

Álgebra Lineal

DRAFT

Analysis Series

M. Winklmeier

Work in progress. Use at your own risk.

Índice general

1. Introducción	7
1.1. Ejemplos de sistemas de ecuaciones lineales; matrices de coeficientes	9
1.2. Sistemas lineales 2×2	15
1.3. Resumen	24
1.4. Ejercicios	24
2. $\mathbb{R}^2$ y $\mathbb{R}^3$	27
2.1. Vectores en $\mathbb{R}^2$	27
2.2. Producto interno en $\mathbb{R}^2$	35
2.3. Proyecciones ortogonales en $\mathbb{R}^2$	42
2.4. Vectores en $\mathbb{R}^n$	46
2.5. Vectores en $\mathbb{R}^3$ y el producto cruz	49
2.6. Rectas y planos en $\mathbb{R}^3$	55
2.7. Intersecciones de rectas y planos en $\mathbb{R}^3$	63
2.8. Resumen	70
2.9. Ejercicios	73
3. Sistemas lineales y matrices	79
3.1. Sistemas lineales y eliminación de Gauß y de Gauß-Jordan	79
3.2. Sistemas lineales homogéneos	92
3.3. Matrices y sistemas lineales	94
3.4. Matrices como funciones de $\mathbb{R}^n$ a $\mathbb{R}^m$; composición de matrices	99
3.5. Inversas de matrices	108
3.6. Matrices y sistemas lineales	114
3.7. La transpuesta de una matriz	120
3.8. Matrices elementales	124
3.9. Resumen	130
3.10. Ejercicios	133
4. Determinantes	139
4.1. Determinante de una matriz	139
4.2. Propiedades del determinante	146
4.3. Interpretación geométrica del determinante	154
4.4. Inversa de una matriz	158
4.5. Resumen	162

4.6. Ejercicios	164
5. Espacios vectoriales	167
5.1. Definiciones y propiedades básicas	167
5.2. Subespacios	174
5.3. Combinaciones lineales e independencia lineal	184
5.4. Base y dimensión	198
5.5. Intersecciones y sumas de espacios vectoriales	210
5.6. Resumen	217
5.7. Ejercicios	219
6. Transformaciones lineales y cambio de bases	227
6.1. Transformaciones lineales	227
6.2. Matrices como aplicaciones lineales	238
6.3. Cambio de bases	249
6.4. Transformaciones lineales y sus representaciones matriciales	261
6.5. Resumen	275
6.6. Ejercicios	277
7. Bases ortonormales y proyecciones ortogonales en $\mathbb{R}^n$	283
7.1. Sistemas ortonormales y bases ortogonales	283
7.2. Matrices ortogonales	288
7.3. Complemento ortogonal	293
7.4. Proyecciones ortogonales	299
7.5. El proceso de Gram-Schmidt	304
7.6. Aplicación: Mínimos cuadrados	308
7.7. Resumen	317
7.8. Ejercicios	318
8. Matrices simétricas y diagonalización	321
8.1. Espacios vectoriales complejos	321
8.2. Matrices semejantes	329
8.3. Valores y vectores propios	333
8.4. Propiedades de los valores propios y vectores propios	345
8.5. Matrices simétricas y hermíticas	352
8.6. Aplicación: Secciones cónicas	356
8.6.1. Soluciones de $ax^2 + bxy + cy^2 = d$ como secciones cónicas	369
8.6.2. Soluciones de $ax^2 + bxy + cy^2 + rx + sy = d$	370
8.7. Resumen	374
8.8. Ejercicios	376
A. Números complejos	379
B. Soluciones	385
Índice	399

Este texto surgió de notas de clase del curso MATE-1105, que he dictado varias veces en la Universidad de los Andes, Bogotá. Es una clase dirigida a estudiantes de carreras distintas de las matemáticas. Aunque no es un texto pensado principalmente para matemáticos, intenté incluir pruebas de todos los teoremas, junto con motivaciones y explicaciones adicionales. Espero que las pruebas ayuden a aclarar por qué las cosas funcionan como funcionan. El texto también contiene muchos ejemplos resueltos. Se recomienda al estudiante revisarlos cuidadosamente y producir sus propias soluciones. Para entender bien los temas presentados en el texto, es importante practicar los contenidos de cada sección con los ejercicios al final de cada sección, y luego con los ejercicios al final de cada capítulo. A lo largo del texto hay recuadros con preguntas que tal vez van un poco más allá de un ejercicio. Cada sección termina con una lista de temas que el estudiante debe haber entendido al finalizar la sección. Una buena práctica de estudio puede ser redactar un breve resumen del material presentado en cada sección, a lo mejor con ejemplos propios.

Quiero agradecer a todos los estudiantes y colegas, entre ellos John Goodrick y Otaivin Martínez, que me dieron retroalimentación, señalaron errores e inconsistencias, y ofrecieron sugerencias valiosas sobre cómo mejorarlas. Agradezco especialmente el trabajo de Mateo Bahos, quien leyó todo el manuscrito y aportó muchos ejercicios que seguramente serán útiles para los lectores.

Considero que estas notas de clase siguen siendo un *trabajo en curso*. Si encuentra errores o tiene sugerencias, por favor no dude en contactarme.

Bogotá, agosto de 2026.

DRAFT

Capítulo 1

Introducción

Este capítulo sirve como introducción a los temas principales del álgebra lineal, a saber, el problema de resolver sistemas de ecuaciones lineales con varias incógnitas. No solo nos interesa una forma eficiente de encontrar sus soluciones, sino que también queremos entender cómo podrían verse las soluciones y qué podemos decir acerca de su estructura. Para esto último, será crucial encontrar una interpretación geométrica de los sistemas de ecuaciones lineales. En este capítulo utilizaremos la estrategia de “resolver e insertar” para resolver sistemas lineales. Un formalismo sistemático y eficiente se dará en el Capítulo 3.

Todo lo que discutimos en este capítulo aparecerá de nuevo más adelante, de modo que puede leerlo rápidamente o incluso omitir (partes de) él.

Un *sistema lineal* es un conjunto de ecuaciones para cierto número de incógnitas que deben satisfacerse simultáneamente y en las cuales las incógnitas aparecen únicamente de manera lineal. Si el número de ecuaciones es m y el número de incógnitas es n , entonces lo llamamos un *sistema lineal* $m \times n$. Típicamente las incógnitas se llaman x, y, z o $x_1, x_2, \dots, x_n$. El siguiente es un ejemplo de un sistema lineal de 3 ecuaciones con 5 incógnitas:

$$x_1 + x_2 + x_3 + x_4 + x_5 = 3, \quad 2x_1 + 3x_2 - 5x_3 + x_4 = 1, \quad 3x_1 - 8x_5 = 0.$$

Un ejemplo de un sistema no lineal es

$$x_1x_2 + x_3 + x_4 + x_5 = 3, \quad 2x_1 + 3x_2 - 5x_3 + x_4 = 1, \quad 3x_1 - 8x_5 = 0$$

porque en la primera ecuación tenemos un producto de dos de las incógnitas. También expresiones como x^2 , $\sqrt[3]{x}$, xyz , x/y o $\cos x$ hacen que un sistema sea no lineal.

Ahora discutamos brevemente el caso no trivial más simple: un sistema que consiste en *una ecuación lineal con una incógnita* x . Su forma más general es

$$ax = b \tag{1.1}$$

donde a y b son constantes dadas y queremos encontrar todos los $x \in \mathbb{R}$ que satisfacen (1.1). Claramente, la solución de este problema depende de los coeficientes a y b . Tenemos que distinguir varios casos.

Caso 1. $a \neq 0$. En este caso, hay exactamente *una solución*, a saber $x = b/a$.

Caso 2. $a = 0, b \neq 0$. En este caso, no hay *ninguna solución* porque, cualquiera que sea el valor que elijamos para x , el lado izquierdo ax siempre será cero y por lo tanto no puede ser igual a b .

Caso 3. $a = 0, b = 0$. En este caso, hay *infinitas soluciones*. De hecho, todo $x \in \mathbb{R}$ resuelve la ecuación.

Así vemos que, ya en este caso simple, tenemos tres tipos muy diferentes de solución del sistema (1.1): ninguna solución, exactamente una solución o infinitas soluciones.

Ahora consideremos un sistema de *una ecuación lineal con dos incógnitas* x, y . Su forma más general es

$$ax + by = c. \quad (1.1')$$

Aquí, a, b, c son constantes dadas y queremos encontrar todos los pares x, y de modo que la ecuación se satisfaga. Por ejemplo, si $a = b = 0$ y $c \neq 0$, entonces el sistema no tiene solución, mientras que si, por ejemplo, $a \neq 0$, entonces hay infinitas soluciones porque, sin importar cómo elijamos y , siempre podemos satisfacer el sistema tomando $x = \frac{1}{a}(c - by)$.

Question 1.1

¿Es posible que el sistema tenga exactamente una solución?
(Vuelva a esta pregunta después de haber estudiado el Capítulo 3.)

La forma general de un sistema de dos ecuaciones lineales con una incógnita es

$$a_1x = b_1, \quad a_2x = b_2$$

y la de un sistema de dos ecuaciones lineales con dos incógnitas es

$$a_{11}x + a_{12}y = c_1, \quad a_{21}x + a_{22}y = c_2$$

donde a_1, a_2, b_1, b_2 , respectivamente $a_{11}, a_{12}, a_{21}, a_{22}, c_1, c_2$ son constantes y x , respectivamente x, y , son las incógnitas.

Question 1.2

¿Puede encontrar ejemplos para los coeficientes de tal manera que los sistemas tengan

- | | |
|--------------------------------|-----------------------------------|
| (I) ninguna solución, | (III) exactamente dos soluciones, |
| (II) exactamente una solución, | (IV) infinitas soluciones? |

¿Puede tal vez incluso dar una regla general para cuándo ocurre cada comportamiento?
(Vuelva a esta pregunta después de haber estudiado el Capítulo 3.)

Antes de discutir sistemas lineales generales, discutiremos en este capítulo introductorio el caso especial de un sistema de dos ecuaciones lineales con dos incógnitas. Aunque este es un tipo muy especial de sistema, exhibe muchas propiedades de los sistemas lineales generales y aparece muy frecuentemente en la práctica.

1.1. Ejemplos de sistemas de ecuaciones lineales; matrices de coeficientes

Comencemos con algunos ejemplos de sistemas de ecuaciones lineales.

Ejemplo 1.1. Suponga que un concesionario vende motocicletas y carros. En total tiene 25 vehículos en su tienda, con un total de 80 ruedas. ¿Cuántas motocicletas y carros hay en la tienda?

Solución. Primero damos nombres a las cantidades que queremos calcular. Entonces sea M = número de motocicletas, C = número de carros en el concesionario. Si escribimos la información dada en el ejercicio en fórmulas, obtenemos

$$\begin{aligned} \textcircled{1} \quad M + C &= 25, & (\text{número total de vehículos}) \\ \textcircled{2} \quad 2M + 4C &= 80, & (\text{número total de ruedas}) \end{aligned}$$

ya que suponemos que cada motocicleta tiene 2 ruedas y cada carro tiene 4 ruedas. La ecuación $\textcircled{1}$ nos dice que $M = 25 - C$. Si insertamos esto en la ecuación $\textcircled{2}$, encontramos

$$80 = 2(25 - C) + 4C = 50 - 2C + 4C = 50 + 2C \implies 2C = 30 \implies C = 15.$$

Esto implica que $M = 25 - C = 25 - 15 = 10$. Observe que en nuestros cálculos y argumentos, todas las flechas de implicación van “de izquierda a derecha”, de modo que lo que podemos concluir en este momento es que *el sistema tiene un único candidato posible para una solución y este candidato es $M = 10$, $C = 15$* . No hemos mostrado (todavía) que realmente *sea* una solución. Sin embargo, al insertar estos números en la ecuación original vemos fácilmente que nuestro candidato sí es una solución.

Entonces la respuesta es: hay 10 motocicletas y 15 carros (y no hay otra posibilidad). $\diamond$

Agreguemos una ecuación más al sistema.

Ejemplo 1.2. Suponga que un concesionario vende motocicletas y carros. En total tienen 25 vehículos en su tienda, con un total de 80 ruedas. Además, la tienda los organiza en 7 áreas distintas de la tienda, de modo que en cada área hay ya sea 3 carros o 5 motocicletas. ¿Cuántas motocicletas y carros hay en la tienda?

Solución. De nuevo, sea M = número de motocicletas, C = número de carros. La información del ejercicio conduce al siguiente sistema de ecuaciones:

$$\begin{aligned} \textcircled{1} \quad M + C &= 25, & (\text{número total de vehículos}) \\ \textcircled{2} \quad 2M + 4C &= 80, & (\text{número total de ruedas}) \\ \textcircled{3} \quad M/5 + C/3 &= 7, & (\text{número total de áreas}) \end{aligned}$$

Como en el ejercicio anterior, obtenemos de $\textcircled{1}$ y $\textcircled{2}$ que $M = 10$, $C = 15$. Claramente, esto también satisface la ecuación $\textcircled{3}$. Así que de nuevo la respuesta es: hay 10 motocicletas y 15 carros (y no hay otra posibilidad). $\diamond$

Ejemplo 1.3. Suponga que un concesionario vende motocicletas y carros. En total tienen 25 vehículos en su tienda, con un total de 80 ruedas. Además, la tienda los organiza en 5 áreas distintas de la tienda, de modo que en cada área hay ya sea 3 carros o 5 motocicletas. ¿Cuántas motocicletas y carros hay en la tienda?

Solución. De nuevo, sea M = número de motocicletas, C = número de carros. La información del ejercicio da las siguientes ecuaciones:

$$\begin{aligned} \textcircled{1} \quad M + C &= 25, & (\text{número total de vehículos}) \\ \textcircled{2} \quad 2M + 4C &= 80, & (\text{número total de ruedas}) \\ \textcircled{3} \quad M/5 + C/3 &= 5. & (\text{número total de áreas}) \end{aligned}$$

Como en el ejercicio anterior, obtenemos que $M = 10$, $C = 15$ usando únicamente las ecuaciones $\textcircled{1}$ y $\textcircled{2}$. Sin embargo, esto no satisface la ecuación $\textcircled{3}$; por lo tanto no hay forma de elegir M y C de tal manera que las tres ecuaciones se satisfagan simultáneamente. Por consiguiente, una tienda como la de este ejemplo no existe. $\diamond$

Ejemplo 1.4. Suponga que un parque zoológico tiene aves y gatos. El conteo total de patas de los animales es 60. Alimentar un ave toma 5 minutos; alimentar un gato toma 10 minutos. El tiempo total para alimentar a los animales es 150 minutos. ¿Cuántas aves y gatos hay en el zoológico?

Solución. Sea B = número de aves, C = número de gatos en el zoológico. La información del ejercicio se traduce a las siguientes ecuaciones:

$$\begin{aligned} \textcircled{1} \quad 2B + 4C &= 60, & (\text{número total de patas}) \\ \textcircled{2} \quad 5B + 10C &= 150, & (\text{tiempo total de alimentación}) \end{aligned}$$

La primera ecuación da $B = 30 - 2C$. Al insertar esto en la segunda ecuación, se obtiene

$$150 = 5(30 - 2C) + 10C = 150 - 10C + 10C = 150$$

lo cual siempre es verdadero, independientemente de la elección de B y C . De hecho, por ejemplo $B = 10$, $C = 10$ o $B = 14$, $C = 8$, o $B = 0$, $C = 15$ son soluciones. Concluimos que la información dada en el ejercicio no es suficiente para calcular el número de animales en el parque zoológico. $\diamond$

Observación. La razón de esto es que ambas ecuaciones $\textcircled{1}$ y $\textcircled{2}$ son en esencia la misma ecuación. Si dividimos la primera por 2 y la segunda por 5, entonces terminamos en ambos casos con la ecuación $B + 2C = 30$, de modo que ambas ecuaciones contienen exactamente la misma información.

Algebraicamente, el sistema lineal tiene infinitas soluciones. Pero nuestras variables representan animales y estos solo aparecen en cantidades enteras no negativas, así que tenemos las 16 soluciones diferentes $B = 30 - 2C$ donde $C \in \{0, 1, \dots, 15\}$.

Damos algunos ejemplos más.

Ejemplo 1.5. Encuentre un polinomio P de grado a lo sumo 3 con

$$P(0) = 1, \quad P(1) = 7, \quad P'(0) = 3, \quad P'(2) = 23. \quad (1.2)$$

Solución. Un polinomio de grado a lo sumo 3 queda determinado si conocemos sus 4 coeficientes. En este ejercicio, las incógnitas son los coeficientes del polinomio P . Si escribimos $P(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta$, entonces tenemos que encontrar $\alpha, \beta, \gamma, \delta$ tales que (1.2) se satisfaga. Observe que $P'(x) = 3\alpha x^2 + 2\beta x + \gamma$. Por lo tanto (1.2) es equivalente al siguiente sistema de ecuaciones:

$$\left. \begin{array}{l} P(0) = 1, \\ P(1) = 7, \\ P'(0) = 3, \\ P'(2) = 23. \end{array} \right\} \iff \left\{ \begin{array}{ll} \textcircled{1} & \delta = 1, \\ \textcircled{2} & \alpha + \beta + \gamma + \delta = 7, \\ \textcircled{3} & \gamma = 3, \\ \textcircled{4} & 12\alpha + 4\beta + \gamma = 23. \end{array} \right.$$

Claramente, $\delta = 1$ y $\gamma = 3$. Si insertamos esto en las ecuaciones restantes, obtenemos un sistema de dos ecuaciones para las dos incógnitas α, β :

$$\begin{array}{ll} \textcircled{2'} & \alpha + \beta = 3, \\ \textcircled{4'} & 12\alpha + 4\beta = 20. \end{array}$$

De $\textcircled{2'}$ obtenemos $\beta = 3 - \alpha$. Si insertamos esto en $\textcircled{4'}$, obtenemos que $20 = 12\alpha + 4(3 - \alpha) = 8\alpha + 12$, es decir, $\alpha = (20 - 12)/8 = 1$. Así, la única solución posible es

$$\alpha = 1, \quad \beta = 2, \quad \gamma = 3, \quad \delta = 1.$$

Es fácil verificar que el polinomio $P(x) = x^3 + 2x^2 + 3x + 1$ tiene todas las propiedades deseadas. $\diamond$

Ejemplo 1.6. Un poste mide 5 metros de largo y debe ser recubierto con barniz. Hay dos tipos de barniz disponibles: el azul añade 3 g por cada 50 cm al poste, el rojo añade 6 g por metro al poste. ¿Es posible recubrir el poste con una combinación de los barnices de modo que el peso total añadido sea

- (a) 35 g? (b) 30 g?

Solución. (a) Denotamos por b la longitud del poste que será cubierta de azul y por r la longitud del poste que será cubierta de rojo. Entonces obtenemos el sistema de ecuaciones

$$\begin{array}{ll} \textcircled{1} & b + r = 5 \quad (\text{longitud total}), \\ \textcircled{2} & 6b + 6r = 35 \quad (\text{peso total}), \end{array}$$

La primera ecuación da $r = 5 - b$. Al insertar en la segunda ecuación se obtiene $35 = 6b + 6(5 - b) = 30$, lo cual es una contradicción. Esto muestra que no hay solución.

(b) Como en (a), obtenemos el sistema de ecuaciones

$$\begin{array}{ll} \textcircled{1} & b + r = 5 \quad (\text{longitud total}), \\ \textcircled{2} & 6b + 6r = 30 \quad (\text{peso total}), \end{array}$$

De nuevo, la primera ecuación da $r = 5 - b$. Al insertar en la segunda ecuación se obtiene $30 = 6b + 6(5 - b) = 30$, lo cual siempre es verdadero, independientemente de cómo elijamos b y r , siempre que $\textcircled{1}$ se satisfaga. Esto significa que, para resolver el sistema de ecuaciones, basta resolver únicamente la primera ecuación, pues entonces la segunda se satisface automáticamente. Así tenemos infinitas

soluciones. Cualquier par b, r tal que $b + r = 5$ da una solución. Así, para cualquier b que elijamos, solo tenemos que poner $r = 5 - b$ y tenemos una solución del problema. Por supuesto, también podríamos fijar r y luego elegir $b = 5 - r$ para obtener una solución. Por ejemplo, podríamos elegir $b = 1$, entonces $r = 4$, o $b = 0,00001$, entonces $r = 4,99999$, o $r = -2$ entonces $b = 7$. Claramente, el último ejemplo no tiene sentido para el problema en cuestión, pero aun así satisface nuestro sistema de ecuaciones. $\diamond$

Ejemplo 1.7. Cuando el octano reacciona con oxígeno, el resultado es dióxido de carbono y agua. Encuentre la ecuación de esta reacción.

Solución. Las fórmulas químicas de las sustancias son C_8H_{18} , O_2 , CO_2 y H_2O . Por lo tanto, la ecuación de reacción es



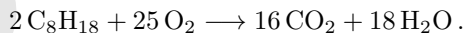
con enteros desconocidos a, b, c, d . Claramente la solución no será única, ya que si tenemos un conjunto de números a, b, c, d que funciona y multiplicamos todos ellos por el mismo número, entonces obtenemos otra solución. Escribamos el sistema de ecuaciones. Para ello observamos que el número de átomos de cada elemento debe ser igual en ambos lados de la ecuación. Obtenemos:

$$\begin{array}{lll} \textcircled{1} & 8a = c, & \text{(carbono)} \\ \textcircled{2} & 18a = 2d, & \text{(hidrógeno)} \\ \textcircled{3} & 2b = 2c + d, & \text{(oxígeno)} \end{array}$$

o, si ponemos todas las variables en el lado izquierdo,

$$\begin{array}{lll} \textcircled{1} & 8a & - c = 0, \\ \textcircled{2} & 18a & - 2d = 0, \\ \textcircled{4} & 2b - 2c & - d = 0. \end{array}$$

Expresemos todas las incógnitas en términos de a : $\textcircled{1}$ y $\textcircled{2}$ muestran que $c = 8a$ y $d = 9a$. Al insertar esto en $\textcircled{3}$ obtenemos $0 = 2b - 2 \cdot 8a - 9a = 2b - 25a$, por lo tanto $b = \frac{25}{2}a$. Si queremos que todos los coeficientes sean enteros, podemos elegir $a = 2$, $b = 25$, $c = 16$, $d = 18$ y la ecuación de reacción se convierte en



$\diamond$

Todos los ejemplos que discutimos en esta sección son los llamados sistemas de ecuaciones lineales. Demos una definición precisa de lo que queremos decir con esto.

Definición 1.8 (Sistema lineal). Un *sistema de ecuaciones lineales* $m \times n$ (o simplemente un *sistema lineal*) es un sistema de m ecuaciones lineales con n incógnitas de la forma

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2, \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m. \end{array} \tag{1.3}$$

Las incógnitas son $x_1, \dots, x_n$, mientras que los números a_{ij} y b_i ($i = 1, \dots, m$, $j = 1, \dots, n$) están dados. Los números a_{ij} se llaman los *coeficientes del sistema lineal* y los números $b_1, \dots, b_n$ se llaman el *lado derecho del sistema lineal*.

Una *solución* del sistema (1.3) es una tupla $(x_1, \dots, x_n)$ tal que las m ecuaciones de (1.3) se satisfacen simultáneamente. El sistema (1.3) se llama *consistente* si tiene al menos una solución. Se llama *inconsistente* si no tiene solución.

En el caso especial en que todos los b_j son iguales a 0, el sistema se llama un *sistema homogéneo*; de lo contrario se llama *inhomogéneo*.

Definición 1.9 (Matriz de coeficientes). La *matriz de coeficientes* A del sistema es la colección de todos los coeficientes a_{ij} en un arreglo como sigue:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \quad (1.4)$$

Los números a_{ij} se llaman las *entradas* o *componentes* de la matriz A . La *matriz aumentada de coeficientes* A del sistema es la colección de todos los coeficientes a_{ij} y el lado derecho; se denota por

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right). \quad (1.5)$$

La matriz de coeficientes no es otra cosa que la colección de los coeficientes a_{ij} ordenados en algún tipo de tabla o rectángulo, de tal manera que el lugar del coeficiente a_{ij} está en la fila i -ésima de la columna j -ésima. La matriz aumentada de coeficientes contiene adicionalmente las constantes del lado derecho.

Notación. En lugar de *fila de una matriz* a veces se dice *renglón de una matriz*.

Observación importante. Hay una correspondencia uno a uno entre sistemas lineales y matrices aumentadas de coeficientes: dado un sistema lineal, es fácil escribir su matriz aumentada de coeficientes y viceversa.

Escribamos las matrices de coeficientes de nuestros ejemplos.

Ejemplo 1.1: Este es un sistema 2×2 con coeficientes $a_{11} = 1$, $a_{12} = 1$, $a_{21} = 2$, $a_{22} = 4$ y lado derecho $b_1 = 25$, $b_2 = 80$. El sistema tiene una solución única. La matriz de coeficientes y la matriz aumentada de coeficientes son

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix}, \quad (A|b) = \left(\begin{array}{cc|c} 1 & 1 & 25 \\ 2 & 4 & 80 \end{array} \right).$$

Ejemplo 1.2: Este es un sistema 3×2 con coeficientes $a_{11} = 1$, $a_{12} = 1$, $a_{21} = 2$, $a_{22} = 4$, $a_{31} = \frac{1}{5}$, $a_{32} = \frac{1}{3}$, y lado derecho $b_1 = 25$, $b_2 = 80$, $b_3 = 7$. El sistema tiene una solución única. La matriz de coeficientes y la matriz aumentada de coeficientes son

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 4 \\ \frac{1}{5} & \frac{1}{3} \end{pmatrix}, \quad (A|b) = \left(\begin{array}{cc|c} 1 & 1 & 25 \\ 2 & 4 & 80 \\ \frac{1}{5} & \frac{1}{3} & 7 \end{array} \right),$$

Ejemplo 1.3: Este es un sistema 3×2 con coeficientes $a_{11} = 1$, $a_{12} = 1$, $a_{21} = 2$, $a_{22} = 4$, $a_{31} = \frac{1}{5}$, $a_{32} = \frac{1}{3}$, y lado derecho $b_1 = 25$, $b_2 = 80$, $b_3 = 7$. El sistema no tiene solución. La matriz de coeficientes es la misma que en el Ejemplo 1.2; la matriz aumentada de coeficientes es

$$(A|b) = \left(\begin{array}{cc|c} 1 & 1 & 25 \\ 2 & 4 & 80 \\ \frac{1}{5} & \frac{1}{3} & 5 \end{array} \right),$$

Ejemplo 1.5: Este es un sistema 4×4 con coeficientes $a_{11} = 0$, $a_{12} = 0$, $a_{13} = 0$, $a_{14} = 1$, $a_{21} = 1$, $a_{22} = 1$, $a_{23} = 1$, $a_{24} = 1$, $a_{31} = 0$, $a_{32} = 0$, $a_{33} = 1$, $a_{34} = 0$, $a_{41} = 12$, $a_{42} = 4$, $a_{43} = 1$, $a_{44} = 0$, y lado derecho $b_1 = 1$, $b_2 = 7$, $b_3 = 3$, $b_4 = 23$. El sistema tiene una solución única. La matriz de coeficientes y la matriz aumentada de coeficientes son

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 12 & 4 & 1 & 0 \end{pmatrix}, \quad (A|b) = \left(\begin{array}{cccc|c} 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 7 \\ 0 & 0 & 1 & 0 & 3 \\ 12 & 4 & 1 & 0 & 23 \end{array} \right).$$

Ejemplo 1.7: Este es un sistema homogéneo 3×4 con coeficientes $a_{11} = 8$, $a_{12} = 0$, $a_{13} = -1$, $a_{14} = 0$, $a_{21} = 18$, $a_{22} = 0$, $a_{23} = 0$, $a_{24} = -2$, $a_{31} = 0$, $a_{32} = 2$, $a_{33} = -2$, $a_{34} = -1$, y lado derecho $b_1 = 0$, $b_2 = 0$, $b_3 = 0$. El sistema tiene una solución única. La matriz de coeficientes y la matriz aumentada de coeficientes son

$$A = \begin{pmatrix} 8 & 0 & -1 & 0 \\ 18 & 0 & 0 & -2 \\ 0 & 2 & -2 & -1 \end{pmatrix}, \quad (A|b) = \left(\begin{array}{cccc|c} 8 & 0 & -1 & 0 & 0 \\ 18 & 0 & 0 & -2 & 0 \\ 0 & 2 & -2 & -1 & 0 \end{array} \right).$$

Vimos que los Ejemplos 1.1, 1.2, 1.5, 1.6 (a) tienen soluciones únicas. En los Ejemplos 1.6 (b) y 1.7 la solución no es única; ¡de hecho, tienen infinitas soluciones! Finalmente, los Ejemplos 1.3 y 1.6(a) no ninguna solución. Así, dado un sistema de ecuaciones lineales $m \times n$, surgen naturalmente dos preguntas importantes:

- **Existencia:** ¿Tiene solución el sistema?
- **Unicidad:** Si el sistema tiene solución, ¿es única?

Más generalmente, quisiéramos poder decir algo acerca de la estructura de las soluciones de los sistemas lineales. Por ejemplo, ¿es posible que haya solo una solución? ¿que haya exactamente dos

soluciones? ¿que haya infinitas soluciones? ¿que no haya ninguna solución? ¿Podemos dar criterios para la existencia y/o unicidad de soluciones? ¿Podemos dar criterios para la existencia de infinitas soluciones? ¿Hay una forma eficiente de calcular todas las soluciones de un sistema lineal dado?

(Spoiler alert: *Un sistema de ecuaciones lineales tiene o bien ninguna, o exactamente una, o infinitas soluciones.* No es posible que tenga, por ejemplo, exactamente 7 soluciones. Esto se discutirá en detalle en el Capítulo 3.)

Antes de responder estas preguntas para sistemas generales $m \times n$ en el Capítulo 3, miraremos más de cerca el caso especial de sistemas 2×2 en la siguiente sección.

Ahora debería haber entendido

- qué es un sistema lineal,
- qué son una matriz de coeficientes y una matriz aumentada de coeficientes,
- su relación con los sistemas lineales,
- que un sistema lineal puede tener distintos tipos de soluciones,
- etc.

Ahora debería ser capaz de

- pasar fácilmente de un sistema lineal $m \times n$ a su matriz de coeficientes (aumentada) y viceversa,
- resolver sistemas lineales mediante el método de “resolver y sustituir”,
- etc.

Ejercicios.

Para los siguientes sistemas de ecuaciones lineales, encuentre al menos una solución (si la hay). ¿Cuáles tienen solución única?

(a) $4x - 6y = 7,$
 $6x - 9y = 12,$

(c) $3x - 5y = 0,$
 $15x - 9y = 0,$

(e) $4x + 14y = 23,$
 $6x + 21y = 30,$

(b) $x + 2y - 3z = -4,$
 $2x + y - 3z = 4,$

(d) $2x + 4y + 6z = 18,$
 $4x + 5y + 6z = 24,$

(f) $6x + 8y = 12,$
 $15x + 20y = 30.$

1.2. Sistemas lineales 2×2

Volvamos a la ecuación del Ejemplo 1.1. Por conveniencia, ahora escribimos x en lugar de B y y en lugar de C . Recordemos que el sistema de ecuaciones que nos interesa resolver es

$$\begin{aligned} \textcircled{1} \quad & x + y = 60, \\ \textcircled{2} \quad & 2x + 4y = 200. \end{aligned} \tag{1.6}$$

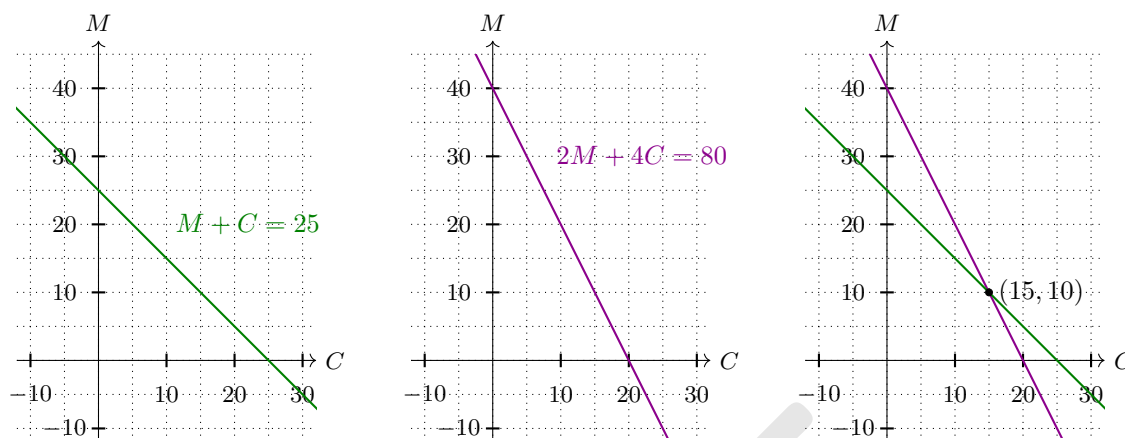


FIGURA 1.1: Gráficas de las rectas L_1 , L_2 que representan las ecuaciones del sistema (1.6) (véase también el Ejemplo 1.1). Su intersección representa la solución única del sistema.

Queremos dar un significado **geométrico** a este sistema de ecuaciones. Para ello pensamos en los pares x, y como puntos (x, y) en el plano. Olvidemos por un momento la ecuación ② y concentrémos solo en ①. Claramente, tiene infinitas soluciones. Si elegimos un x arbitrario, siempre podemos encontrar y de tal manera que ① se satisfaga (basta tomar $y = 60 - x$). De manera similar, si elegimos cualquier y , entonces solo tenemos que tomar $x = 60 - y$ y obtenemos una solución de ①.

¿Dónde en el plano xy se encuentran *todas* las soluciones de ①? Claramente, ① es equivalente a $y = 60 - x$, que identificamos fácilmente como la ecuación de la recta L_1 en el plano xy , la cual pasa por $(0, 60)$ y tiene pendiente -1 . En resumen, un par (x, y) es una solución de ① si y solo si está sobre la recta L_1 , véase la Figura 1.1.

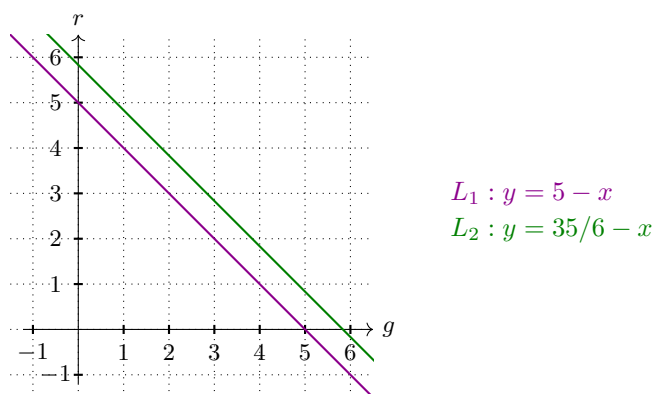
Si aplicamos el mismo razonamiento a ②, encontramos que un par (x, y) satisface ② si y solo si (x, y) está sobre la recta L_2 en el plano xy dada por $y = \frac{1}{4}(200 - 2x)$ (esta es la recta en el plano xy que pasa por $(0, 50)$ con pendiente $-\frac{1}{2}$).

Ahora está claro que un par (x, y) satisface tanto ① como ② si y solo si está sobre ambas rectas L_1 y L_2 . Así, encontrar la solución de nuestro sistema (1.6) es lo mismo que encontrar la intersección de las dos rectas L_1 y L_2 . De la geometría elemental sabemos que hay exactamente tres posibilidades para su intersección:

- (I) L_1 y L_2 no son paralelas. Entonces se intersectan en exactamente un punto.
- (II) L_1 y L_2 son paralelas y no son iguales. Entonces no se intersectan.
- (III) L_1 y L_2 son paralelas e iguales. Entonces $L_1 = L_2$ y se intersectan en infinitos puntos (se intersectan en cada punto de $L_1 = L_2$).

En nuestro ejemplo sabemos que la pendiente de L_1 es -1 y que la pendiente de L_2 es $-\frac{1}{2}$, así que no son paralelas y por lo tanto se intersectan en exactamente un punto. En consecuencia, el sistema (1.6) tiene exactamente una solución.

Si miramos de nuevo el Ejemplo 1.6, vemos que en el Caso (a) tenemos que determinar la intersección

FIGURA 1.2: Ejemplo 1.6. Gráficas de L_1 , L_2 .

de las rectas

$$L_1 : y = 5 - x, \quad L_2 : y = \frac{35}{6} - x.$$

Ambas rectas tienen pendiente -1 , así que son paralelas. Como los términos constantes en ambas rectas no son iguales, no se intersectan en ningún punto, lo cual muestra que el sistema de ecuaciones no tiene solución; véase la Figura 1.2. En el Caso (b), las dos rectas que tenemos que intersectar son

$$G_1 : y = 5 - x, \quad G_2 : y = 5 - x.$$

Vemos que $G_1 = G_2$, así que todo punto sobre G_1 (o G_2) es solución del sistema y por lo tanto tenemos infinitas soluciones.

Observación importante. Que un sistema lineal 2×2 tenga o no una solución única no tiene nada que ver con el lado derecho del sistema, sino solo depende de si las dos rectas son paralelas o no, y esto a su vez depende únicamente de los coeficientes del lado izquierdo.

Consideremos ahora el caso general.

Una ecuación lineal con dos incógnitas

La forma general de una ecuación lineal con dos incógnitas es

$$\alpha x + \beta y = \gamma. \tag{1.7}$$

Para el conjunto de soluciones, hay tres posibilidades:

- (i) *El conjunto de soluciones forma una recta.* Esto ocurre si al menos uno de los coeficientes α o β es distinto de 0. Si $\beta \neq 0$, entonces el conjunto de todas las soluciones es igual a la recta $L : y = -\frac{\alpha}{\beta}x + \frac{\gamma}{\beta}$, que es una recta con pendiente $-\frac{\alpha}{\beta}$. Si $\beta = 0$ y $\alpha \neq 0$, entonces el conjunto de soluciones de (1.7) es una recta paralela al eje y que pasa por $(0, \frac{\gamma}{\alpha})$.

- (II) *El conjunto de soluciones es todo el plano.* Esto ocurre si $\alpha = \beta = \gamma = 0$. En este caso, claramente todo par (x, y) es una solución de (1.7).
- (III) *No hay solución.* Esto ocurre si $\alpha = \beta = 0$ y $\gamma \neq 0$. En este caso, ningún par (x, y) es una solución de (1.7), ya que el lado izquierdo siempre es 0.

En los primeros dos casos, (1.7) tiene infinitas soluciones; en el último caso no tiene solución.

Dos ecuaciones lineales con dos incógnitas

La forma general de dos ecuaciones lineales con dos incógnitas es

$$\begin{aligned} \textcircled{1} \quad Ax + By &= U, \\ \textcircled{2} \quad Cx + Dy &= V. \end{aligned} \tag{1.8}$$

Estamos usando las letras A, B, C, D en lugar de $a_{11}, a_{12}, a_{21}, a_{22}$ con el fin de hacer los cálculos más legibles. Si interpretamos el sistema de ecuaciones como la intersección de dos objetos geométricos, en nuestro caso rectas, ya sabemos que hay exactamente tres tipos posibles de soluciones:

- (I) *Un punto* si $\textcircled{1}$ y $\textcircled{2}$ describen dos rectas no paralelas.
- (II) *Una recta* si $\textcircled{1}$ y $\textcircled{2}$ describen la misma recta; o si una de las ecuaciones es un plano y la otra es una recta.
- (III) *Un plano* si ambas ecuaciones describen un plano.
- (IV) *El conjunto vacío* si las dos ecuaciones describen rectas paralelas pero distintas; o si una de las ecuaciones no tiene solución.

En el caso (I), el sistema tiene exactamente una solución, en los casos (II) y (III) el sistema tiene infinitas soluciones y en el caso (IV) el sistema no tiene solución.

En resumen, tenemos la siguiente observación muy importante.

Observación 1.10. El sistema (1.8) tiene o bien exactamente una solución, o infinitas soluciones, o ninguna solución.

No es posible tener, por ejemplo, exactamente 7 soluciones.

Question 1.3

¿Cuál es la interpretación geométrica de

- (I) un sistema de 3 ecuaciones lineales con 2 incógnitas?
- (II) un sistema de 2 ecuaciones lineales con 3 incógnitas?

¿Qué se puede decir acerca de la estructura de sus soluciones?

Prueba algebraica de la Observación 1.10. Ahora queremos probar algebraicamente la Observación 1.10 y queremos encontrar un criterio sobre A, B, C, D que nos permita decidir fácilmente cuántas soluciones hay. Consideremos los distintos casos.

Caso 1. $B \neq 0$. En este caso podemos resolver ① para y y obtenemos $y = \frac{1}{B}(U - Ax)$. Al insertar en ② encontramos $Cx + \frac{D}{B}(U - Ax) = V$. Si ponemos todos los términos con x en un lado y todos los demás términos en el otro lado, obtenemos

$$\textcircled{2} \quad (AD - BC)x = DU - BV.$$

- (I) Si $AD - BC \neq 0$ entonces hay a lo sumo una solución, a saber $x = \frac{DU - BV}{AD - BC}$ y, en consecuencia, $y = \frac{1}{B}(U - Ax) = \frac{AV - CU}{AD - BC}$. Al insertar estas expresiones para x y y en nuestro sistema de ecuaciones, vemos que efectivamente resuelven el sistema (1.8), de modo que tenemos exactamente una solución.
- (II) Si $AD - BC = 0$ entonces la ecuación ② se reduce a $0 = DU - BV$. Esta ecuación no tiene solución (si $DU - BV \neq 0$) o es verdadera para toda elección posible de x y y (si $DU - BV = 0$). Como ① tiene infinitas soluciones, se sigue que el sistema (1.8) tiene o bien ninguna solución o bien infinitas soluciones.

Caso 2. $D \neq 0$. Este caso es análogo al Caso 1. En este caso podemos resolver ② para y y obtenemos $y = \frac{1}{D}(V - Cx)$. Por lo tanto ① se convierte en $Ax + \frac{B}{D}(V - Cx) = U$. Si ponemos todos los términos con x en un lado y todos los demás términos en el otro lado, obtenemos

$$\textcircled{1} \quad (AD - BC)x = DU - BV.$$

Tenemos los mismos subcasos que antes:

- (I) Si $AD - BC \neq 0$ entonces hay exactamente una solución, a saber $x = \frac{DU - BV}{AD - BC}$ y, en consecuencia, $y = \frac{1}{D}(V - Cx) = \frac{AV - CU}{AD - BC}$.
- (II) Si $AD - BC = 0$ entonces la ecuación ① se reduce a $0 = DU - BV$. Esta ecuación no tiene solución (si $DU - BV \neq 0$) o se cumple para todo x y y (si $DU - BV = 0$). Como ② tiene infinitas soluciones, se sigue que el sistema (1.8) tiene o bien ninguna solución o bien infinitas soluciones.

Caso 3. $B = 0$ y $D = 0$. Observe que en este caso $AD - BC = 0$. En este caso el sistema (1.8) se reduce a

$$Ax = U, \quad Cx = V. \tag{1.9}$$

Vemos que el sistema ya no depende de y . Así, si el sistema (1.9) tiene al menos una solución, entonces automáticamente tenemos infinitas soluciones, ya que podemos elegir y libremente. Si el sistema (1.9) no tiene solución, entonces el sistema original (1.8) tampoco puede tener solución.

Observe que no hay otros casos posibles para los coeficientes. $\square$

En resumen, hemos probado el siguiente teorema.

Teorema 1.11. Consideremos el sistema lineal

$$\begin{aligned} \textcircled{1} \quad Ax + By &= U, \\ \textcircled{2} \quad Cx + Dy &= V. \end{aligned} \tag{1.10}$$

(I) El sistema (1.10) tiene exactamente una solución si y solo si $AD - BC \neq 0$. En este caso, la solución es

$$x = \frac{DU - BV}{AD - BC}, \quad y = \frac{AV - CU}{AD - BC}. \tag{1.11}$$

(II) El sistema (1.10) no tiene solución o tiene infinitas soluciones si y solo si $AD - BC = 0$.

Definición 1.12. El número $d = AD - BC$ se llama el *determinante* del sistema (1.10).

En el Capítulo 4.1 generalizaremos este concepto a sistemas $n \times n$ para $n \geq 3$.

Observación 1.13. Veamos cómo se conecta el determinante con nuestra interpretación geométrica del sistema de ecuaciones. Suponga que $B \neq 0$ y $D \neq 0$. Entonces podemos resolver $\textcircled{1}$ y $\textcircled{2}$ para y con el fin de obtener ecuaciones para un par de rectas

$$L_1: y = -\frac{A}{B}x + \frac{1}{B}U, \quad L_2: y = -\frac{C}{D}x + \frac{1}{D}V.$$

Las dos rectas se intersectan en exactamente un punto si y solo si tienen pendientes diferentes, es decir, si $-\frac{A}{B} \neq -\frac{C}{D}$. Después de multiplicar por $-BD$ vemos que esto es lo mismo que $AD \neq BC$, o, en otras palabras, $AD - BC \neq 0$. Por otro lado, las rectas son paralelas (por lo tanto, o bien son iguales o no tienen intersección) si $-\frac{A}{B} = -\frac{C}{D}$. Este es el caso si y solo si $AD = BC$, o, en otras palabras, si $AD - BC = 0$.

Question 1.4

Considere los casos en que $B = 0$ o $D = 0$ y establezca la conexión entre el Teorema 1.11 y la interpretación geométrica del sistema de ecuaciones.

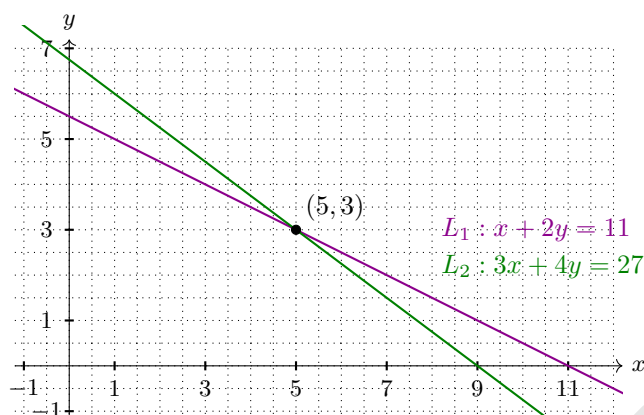
Consideremos algunos ejemplos más.

Ejemplos 1.14. (a)
$$\begin{aligned} \textcircled{1} \quad x + 2y &= 11, \\ \textcircled{2} \quad 3x + 4y &= 27. \end{aligned}$$

Claramente, el determinante es $d = 4 - 6 = -2 \neq 0$. Así que el sistema tiene *exactamente una solución*. Podemos verificar esto fácilmente: la primera ecuación da $x = 11 - 2y$. Al insertar esto en la segunda ecuación se obtiene

$$3(11 - 2y) + 4y = 27 \implies -2y = -6 \implies y = 3 \implies x = 11 - 2 \cdot 3 = 5.$$

Así, la solución es $x = 5, y = 3$. (Si no tuviéramos el Teorema 1.11, tendríamos que verificar que esto no solo es un candidato para una solución, sino que de hecho es una.)

FIGURA 1.3: Ejemplo 1.14(a). Gráficas de L_1 , L_2 y su intersección $(5, 3)$.

Verifique que la fórmula (1.11) se satisface.

- (b) ① $x + 2y = 1$,
 ② $2x + 4y = 5$.

Aquí, el determinante es $d = 4 - 4 = 0$, así que esperamos *o bien ninguna solución o bien infinitas soluciones*. La primera ecuación da $x = 1 - 2y$. Al insertar en la segunda ecuación se obtiene $2(1 - 2y) + 4y = 5$. Vemos que los términos con y se cancelan y obtenemos $2 = 5$, lo cual es una contradicción. Por lo tanto, el sistema de ecuaciones no tiene *ninguna solución*.

- (c) ① $x + 2y = 1$,
 ② $3x + 6y = 3$.

El determinante es $d = 6 - 6 = 0$, así que de nuevo esperamos *o bien ninguna solución o bien infinitas soluciones*. La primera ecuación da $x = 1 - 2y$. Al insertar en la segunda ecuación se obtiene $3(1 - 2y) + 6y = 3$. Vemos que los términos con y se cancelan y obtenemos $3 = 3$, lo cual es verdadero. Por lo tanto, el sistema de ecuaciones tiene *infinitas soluciones* dadas por $x = 1 - 2y$.

Observación. Esto era en cierto modo claro, ya que podemos obtener la segunda ecuación a partir de la primera multiplicando ambos lados por 3, lo cual muestra que ambas ecuaciones contienen la misma información y no perdemos nada si simplemente olvidamos una de ellas.

Ejercicio 1.15. Encuentre todos los $k \in \mathbb{R}$ tales que el sistema

$$\begin{aligned} \textcircled{1} \quad & kx + (15/2 - k)y = 1, \\ \textcircled{2} \quad & 4x + \quad \quad \quad 2ky = 3 \end{aligned}$$

tenga exactamente una solución.

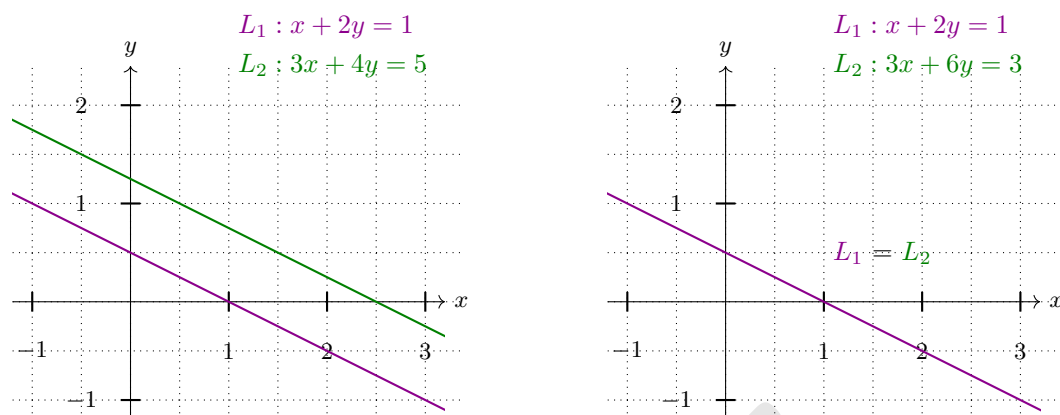


FIGURA 1.4: Imagen de la izquierda: Las rectas L_1, L_2 del Ejemplo 1.14(b) son paralelas y no se intersectan. Por lo tanto, el sistema lineal no tiene solución. Imagen de la derecha: Las rectas L_1, L_2 del Ejemplo 1.14(c) son iguales. Por lo tanto, el sistema lineal tiene infinitas soluciones.

Solución. Solo necesitamos calcular el determinante y encontrar todos los k para los cuales es distinto de cero. Entonces comencemos calculando

$$d = k \cdot 2k - (15/2 - k) \cdot 4 = 2k^2 + 4k - 30 = 2(k^2 + 2k - 15) = 2[(k+1)^2 - 16].$$

Por lo tanto hay exactamente dos valores de k para los cuales $d = 0$, a saber $k = -1 \pm 4$, es decir $k_1 = 3, k_2 = -5$. Para todos los demás k , tenemos que $d \neq 0$. Así que la respuesta es: El sistema tiene exactamente una solución si y solo si $k \in \mathbb{R} \setminus \{-5, 3\}$. $\diamond$

Observación 1.16. (a) Observe que la respuesta no depende del lado derecho del sistema de ecuaciones. Solo los coeficientes del lado izquierdo determinan si hay exactamente una solución o no.

(b) Si quisiéramos, también podríamos calcular la solución x, y en el caso $k \in \mathbb{R} \setminus \{-5, 3\}$. Podríamos hacerlo a mano o usar (1.11). De cualquier forma, encontramos

$$x = \frac{1}{d}[2k - 3(15/2 - k)] = \frac{5k - 45/2}{2k^2 + 4k - 30}, \quad y = \frac{1}{d}[6k - 4] = \frac{6k - 4}{2k^2 + 4k - 30}.$$

Observe que los denominadores son iguales a d y son iguales a 0 exactamente para los valores “prohibidos” de $k = -5$ o $k = 3$.

(c) ¿Qué ocurre si $k = -5$ o $k = 3$? En ambos casos, $d = 0$, así que tendremos o bien ninguna solución o bien infinitas soluciones.

Si $k = -5$, entonces el sistema se convierte en $-5x + 25/2y = 1, \quad 4x - 10y = 3$.

Multiplicando la primera ecuación por $-4/5$ y sin cambiar la segunda ecuación, obtenemos

$$4x - 10y = -\frac{4}{5}, \quad 4x - 10y = 3$$

lo cual claramente no puede satisfacerse simultáneamente.

Si $k = 3$, entonces el sistema se convierte en $3x - 9/2y = 1$, $4x + 6y = 3$.

Multiplicando la primera ecuación por $4/3$ y sin cambiar la segunda ecuación, obtenemos

$$4x - 6y = \frac{4}{3}, \quad 4x - 6y = 3$$

lo cual claramente no puede satisfacerse simultáneamente.

En conclusión, si $k = -5$ o $k = 3$, entonces el sistema lineal no tiene solución.

Debería haber entendido

- la interpretación geométrica de un sistema lineal $m \times 2$ y cómo ayuda a entender la estructura cualitativa de las soluciones,
- cómo el determinante ayuda a decidir si un sistema lineal 2×2 tiene o no una solución única,
- que si un sistema 2×2 tiene solución única depende solo de los coeficientes; *no* depende del lado derecho de la ecuación (los valores concretos de las soluciones, por supuesto, sí dependen del lado derecho de la ecuación),
- etc.

Ahora debería ser capaz de

- pasar fácilmente de un sistema lineal $m \times 2$ a su interpretación geométrica y viceversa,
- calcular el determinante de un sistema lineal 2×2 ,
- determinar si un sistema lineal 2×2 tiene una solución única, ninguna o infinitas soluciones, y calcularlas,
- dar criterios para la existencia/unicidad de soluciones,
- etc.

Ejercicios.

1. Usando el criterio del determinante, diga cuáles sistemas tienen solución única y encuéntrala. En caso de que el determinante sea cero, especifique si el sistema posee infinitas soluciones o ninguna solución.

(a) $6x + y = 3$; $-4x - y = 8$

(b) $5x + 2y = 7$; $2x + 5y = 4$

(c) $4x - 6y = 0$; $2x - 3y = 0$

(d) $2x - 8y = 6$; $-3x + 12y = 4$

(e) $2x - 8y = 6$; $-3x + 12y = -9$

(f) $2y = 4$; $5x - 3y = 1$

2. ¿Para cuáles valores de k se cortan las siguientes rectas exactamente en un punto?

$$y = \frac{k+3}{2}x + \pi - \sqrt{2}, \quad y = \frac{2k-5}{3}x + \sqrt[3]{3}.$$

3. ¿Para cuáles valores de k se cortan las siguientes rectas exactamente en un punto?

$$(k+2)x - 3y = \sqrt{2\pi}, \quad 5kx + (k-1)y = 3 - e^2.$$

1.3. Resumen

Un *sistema lineal* es un sistema de ecuaciones

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

donde $x_1, \dots, x_n$ son las incógnitas y los números a_{ij} y b_i ($i = 1, \dots, m$, $j = 1, \dots, n$) están dados. Los números a_{ij} se llaman los *coeficientes del sistema lineal* y los números $b_1, \dots, b_n$ se llaman el *lado derecho del sistema lineal*.

En el caso especial en que todos los b_i son iguales a 0, el sistema se llama *homogéneo*; de lo contrario se llama *inhomogéneo*.

La *matriz de coeficientes* A y la *matriz aumentada de coeficientes* $(A|b)$ del sistema son

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad (A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{array} \right).$$

La forma general de un sistema lineal 2×2 es

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \tag{1.12}$$

y su *determinante* es

$$d = a_{11}a_{22} - a_{21}a_{12}.$$

El determinante nos dice si el sistema (1.12) tiene una solución única:

- Si $d \neq 0$, entonces (1.12) tiene una solución única.
- Si $d = 0$, entonces (1.12) tiene o bien ninguna o bien infinitas soluciones (depende de b_1 y b_2 cuál caso ocurre).

Observe que d **no** depende del lado derecho del sistema lineal.

1.4. Ejercicios

1. Encuentre el área del triángulo que se encuentra en el primer cuadrante y que está delimitado por las rectas $y = 2x - 4$, $y = -4x + 20$.

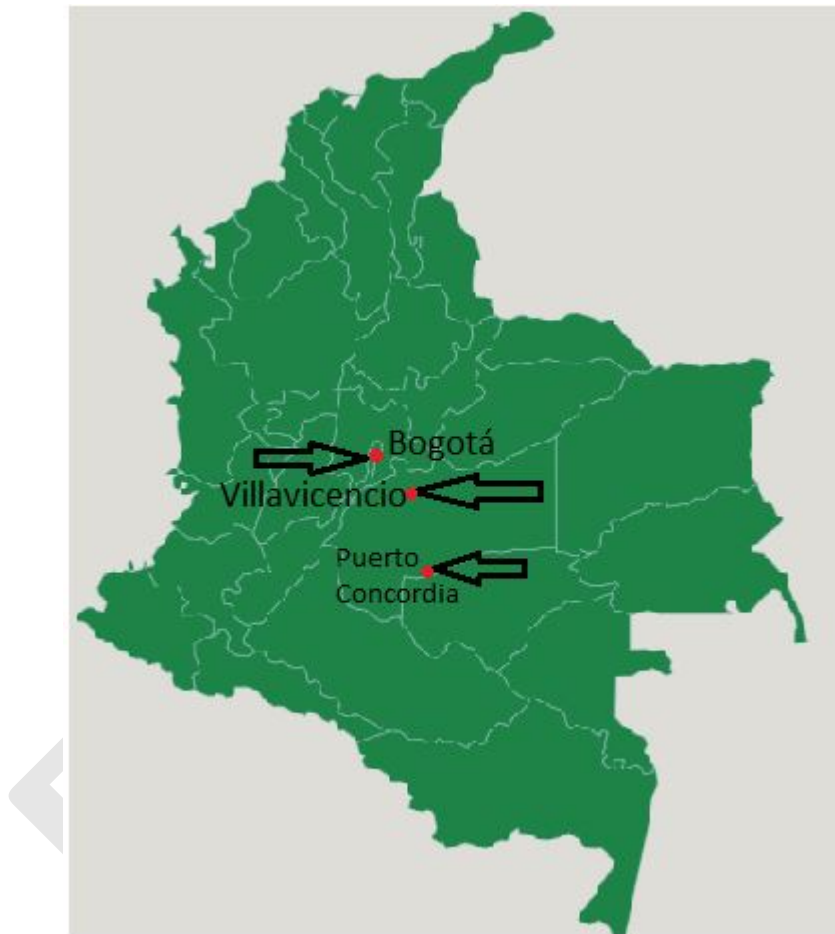
2. Suponga que los puntos $(1, 5)$, $(-1, 3)$ y $(0, 1)$ están sobre la parábola $y = ax^2 + bx + c$. Con esta información, determine los valores de a, b, c .
3. Describa todas las parábolas que pasan por los puntos $(1, 1)$ y $(-1, 4)$.
4. Encuentre todos los valores de $t, k \in \mathbb{R}$ tal que el siguiente sistema sea consistente.

$$2x + 8y = 4,$$

$$5x + 4ky = 20,$$

$$tx + 2y = 1.$$

5. De un número de tres cifras sabemos que sus tres dígitos suman 11, y la suma del primer y tercer dígito es 5. Encuentre todos los números que cumplen la propiedad anterior.
6. El dueño de una tienda vende comida para perros a 40\$ y comida para gatos a 20\$. Haciendo cuentas de la semana observa que por concepto de comida de animales recibió 640\$ y que 22 clientes entraron esa semana a comprar comida de animales. Si se supone que cada cliente tiene una única mascota ¿Cuántos clientes eran dueños de perros y cuántos de gatos?
7. La suma de la cifra de las decenas y la cifra de las unidades de un número de dos dígitos es 12, y si al número se le resta 18, las cifras se invierten. Hallar el número.
8. Sabemos que la distancia entre Bogotá y Puerto Concordia es de 375 km aproximadamente y la distancia entre Villavicencio y Puerto Concordia es de aproximadamente 261 km. Un conductor A parte de Bogotá hacia Villavicencio con una velocidad constante de 57 km/h a las 4:00 am y una hora después, un conductor B parte de Puerto Concordia hacia Bogotá a una velocidad constante de 49 km/h. ¿A qué hora llega el conductor A a Villavicencio? ¿Los conductores A y B se encuentran en carretera? ¿A qué hora lo hacen?. Repita las preguntas si suponemos que el conductor A se mueve a una velocidad de 19 km/h y el conductor B se mueve a una velocidad de 70 km/h.



Capítulo 2

$\mathbb{R}^2$ y $\mathbb{R}^3$

En este capítulo introduciremos los espacios vectoriales $\mathbb{R}^2$, $\mathbb{R}^3$ y $\mathbb{R}^n$. Definiremos operaciones algebraicas en ellos y las interpretaremos geoméricamente. Luego añadiremos a estos espacios una estructura adicional, a saber, un producto interno. Esto nos permite asignar una norma (longitud) a un vector y hablar del ángulo entre dos vectores; en particular, nos da el concepto de ortogonalidad. En la Sección 2.3 definiremos proyecciones ortogonales en $\mathbb{R}^2$ y daremos una fórmula para la proyección ortogonal de un vector sobre otro. Esta fórmula se generaliza fácilmente a proyecciones sobre un vector en $\mathbb{R}^n$ con $n \geq 3$. La Sección 2.5 está dedicada al caso especial y muy importante de $\mathbb{R}^3$, ya que es el espacio que los físicos usan en mecánica clásica para describir nuestro mundo. En las dos últimas secciones estudiamos rectas y planos en $\mathbb{R}^n$ y en $\mathbb{R}^3$. Veremos cómo podemos describirlos mediante fórmulas y aprenderemos a calcular sus intersecciones. Esto conduce naturalmente a la pregunta de cómo resolver sistemas lineales de manera eficiente, cuestión que abordaremos en el próximo capítulo.

2.1. Vectores en $\mathbb{R}^2$

Recordemos que el plano xy es el conjunto de todos los pares (x, y) con $x, y \in \mathbb{R}$. Lo denotaremos por $\mathbb{R}^2$.

Tal vez ya haya encontrado vectores en una clase de física. Por ejemplo, las velocidades y las fuerzas se describen mediante vectores. La velocidad de una partícula indica qué tan rápido se mueve y en qué dirección lo hace. Usualmente, la velocidad se representa mediante una flecha que apunta en la dirección en la que se mueve la partícula y cuya longitud es proporcional a la magnitud de la velocidad.

De manera similar, una fuerza tiene intensidad y dirección, por lo que se representa mediante una flecha que apunta en la dirección en la que actúa y cuya longitud es proporcional a su intensidad. Observe que no importa en qué lugar del espacio $\mathbb{R}^2$ o $\mathbb{R}^3$ pongamos la flecha. Mientras apunte en la misma dirección y tenga la misma longitud, se considera el mismo vector. Llamamos *equivalentes* a dos flechas si tienen la misma dirección y la misma longitud. Un *vector* es el conjunto de todas las flechas que son equivalentes a una flecha dada. Cada flecha específica de este conjunto se llama una *representación* del vector. Una representación especial es la flecha que empieza en el origen $(0, 0)$. Los vectores usualmente se denotan por una letra minúscula con una flecha encima, por ejemplo $\vec{v}$.

Dados dos puntos P, Q en el plano xy , escribimos $\overrightarrow{PQ}$ para el vector representado por la flecha que empieza en P y termina en Q . Por ejemplo, sean $P(2, 1)$ y $Q(4, 4)$ puntos en el plano xy . Entonces la flecha de P a Q es $\overrightarrow{PQ} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

Podemos identificar un punto $P(p_1, p_2)$ en el plano xy con el vector que empieza en el punto $(0, 0)$ y termina en P . Denotamos este vector por $\overrightarrow{OP}$ o $\begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$ o a veces por $(p_1, p_2)^t$ para ahorrar espacio (el superíndice t significa “transpuesto”). p_1 se llama su *coordenada x* o *componente x* y p_2 se llama su *coordenada y* o *componente y*.

Por otro lado, todo vector $\begin{pmatrix} a \\ b \end{pmatrix}$ describe un punto único en el plano xy , a saber, la punta de la flecha que representa el vector dado y empieza en el origen. Claramente, sus coordenadas son (a, b) . Por lo tanto, podemos identificar el conjunto de todos los vectores en $\mathbb{R}^2$ con $\mathbb{R}^2$ mismo.

Observe que la pendiente de la flecha $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ es $\frac{b}{a}$ si $a \neq 0$. Si $a = 0$, entonces el vector es paralelo al eje y .

Por ejemplo, el vector $\vec{v} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$ puede representarse como una flecha cuyo punto inicial está en el origen y cuya punta está en el punto $(2, 5)$. Si ponemos su punto inicial en cualquier otro lugar, encontramos la punta moviéndonos 2 unidades hacia la derecha (paralelamente al eje x) y 5 unidades hacia arriba (paralelamente al eje y).

Un vector muy especial es el vector cero $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Usualmente se denota por $\vec{0}$.

Llamamos *escalares* a los números en $\mathbb{R}$ para distinguirlos de los vectores.

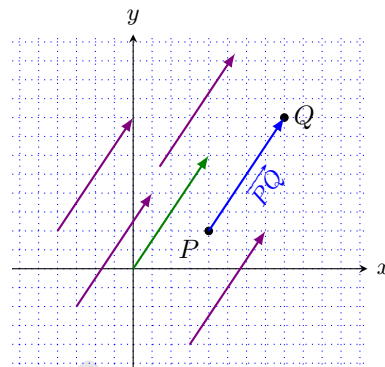


FIGURA 2.1: El vector $\overrightarrow{PQ}$ y varias de sus representaciones. La flecha verde es la representación especial cuyo punto inicial está en el origen.

Álgebra con vectores

Si pensamos en una fuerza y duplicamos su intensidad, entonces el vector correspondiente debería tener el doble de longitud. Si *multiplicamos* la fuerza por 5, entonces la longitud del vector correspondiente debería ser 5 veces mayor; es decir, si por ejemplo se da una fuerza $\vec{F} = (3, 4)^t$, entonces $5\vec{F}$ debería ser $(5 \cdot 3, 5 \cdot 4)^t = (15, 20)^t$.

En general, si se dan un vector $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ y un escalar c , entonces $c\vec{v} = \begin{pmatrix} ca \\ cb \end{pmatrix}$. Observe que el vector resultante siempre es paralelo al original. Si $c > 0$, entonces el vector resultante apunta en la misma dirección que el original; si $c < 0$, entonces apunta en la dirección opuesta, véase la Figura 2.2.

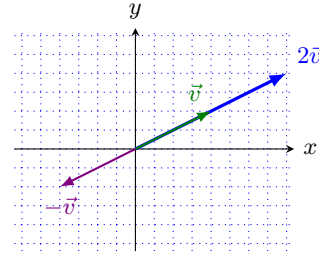


FIGURA 2.2: Multiplicación de un vector por un escalar.

Dados dos puntos $P(p_1, p_2)$, $Q(q_1, q_2)$ en el plano xy . Convénzase de que $\overrightarrow{PQ} = -\overrightarrow{QP}$.

¿Cómo deberíamos *sumar* dos vectores? De nuevo, pensemos en fuerzas. Supongamos que tenemos dos fuerzas $\vec{F}_1$ y $\vec{F}_2$ que actúan sobre la misma partícula. Entonces obtenemos la fuerza resultante si dibujamos la flecha que representa $\vec{F}_1$ y unimos a su punto final el punto inicial de la flecha que representa $\vec{F}_2$. La fuerza total queda entonces representada por la flecha que empieza en el punto inicial de $\vec{F}_1$ y termina en la punta de $\vec{F}_2$.

Convénzase de que obtenemos el mismo resultado si comenzamos con $\vec{F}_2$ y ponemos el punto inicial de $\vec{F}_1$ en la punta de $\vec{F}_2$.

También podríamos pensar en la suma de velocidades. Por ejemplo, si un tren se mueve con velocidad $\vec{v}_t$ y una pasajera en el tren se mueve con velocidad relativa $\vec{v}_p$, entonces su velocidad total con respecto al suelo es la suma vectorial de las dos velocidades.

Supongamos ahora que $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ y $\vec{w} = \begin{pmatrix} p \\ q \end{pmatrix}$. Algebraicamente, obtenemos las componentes de su suma sumando las componentes: $\vec{v} + \vec{w} = \begin{pmatrix} a + p \\ b + q \end{pmatrix}$, véase la Figura 2.3.

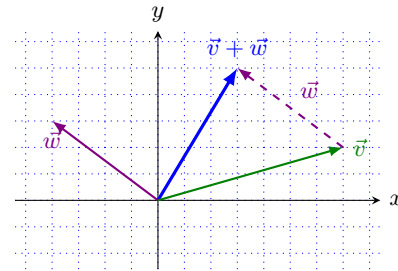


FIGURA 2.3: Suma de dos vectores.

Cuando sume vectores, siempre debería pensar en triángulos (o en polígonos si suma más de dos vectores).

Dados dos puntos $P(p_1, p_2)$, $Q(q_1, q_2)$ en el plano xy . Convénzase de que $\overrightarrow{OP} + \overrightarrow{PQ} = \overrightarrow{OQ}$ y, en consecuencia, $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$.

¿Cómo podría escribir $\overrightarrow{QP}$ en términos de $\overrightarrow{OP}$ y $\overrightarrow{OQ}$?
¿Cuál es su relación con $\overrightarrow{PQ}$?

Nuestra discusión sobre cómo deberían ser el producto de un vector por un escalar y la suma de

dos vectores nos lleva a la siguiente definición formal.

Definición 2.1. Sea $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$, $\vec{w} = \begin{pmatrix} p \\ q \end{pmatrix} \in \mathbb{R}^2$, $c \in \mathbb{R}$. Entonces:

$$\text{Suma vectorial:} \quad \vec{v} + \vec{w} = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} a + p \\ b + q \end{pmatrix},$$

$$\text{Producto por un escalar:} \quad c\vec{v} = c \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ca \\ cb \end{pmatrix}.$$

Es fácil ver que la suma vectorial satisface lo que uno espera de una suma: $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ (asociatividad) y $\vec{v} + \vec{w} = \vec{w} + \vec{v}$ (conmutatividad). Además, tenemos las leyes de distributividad $(a + b)\vec{v} = a\vec{v} + b\vec{v}$ y $a(\vec{v} + \vec{w}) = a\vec{v} + a\vec{w}$. Verifiquemos, por ejemplo, la asociatividad. Para ello, sean $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$. Entonces

$$\begin{aligned} (\vec{u} + \vec{v}) + \vec{w} &= \left[\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right] + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} (u_1 + v_1) + w_1 \\ (u_2 + v_2) + w_2 \end{pmatrix} \\ &= \begin{pmatrix} u_1 + (v_1 + w_1) \\ u_2 + (v_2 + w_2) \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} (v_1 + w_1) \\ (v_2 + w_2) \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \left[\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right] \\ &= \vec{u} + (\vec{v} + \vec{w}). \end{aligned}$$

De la misma manera, verifique la conmutatividad y la distributividad de la suma vectorial.

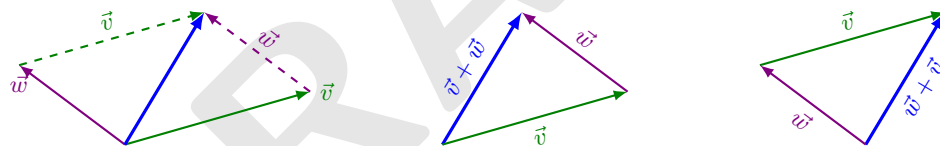


FIGURA 2.4: La imagen ilustra la conmutatividad de la suma vectorial.

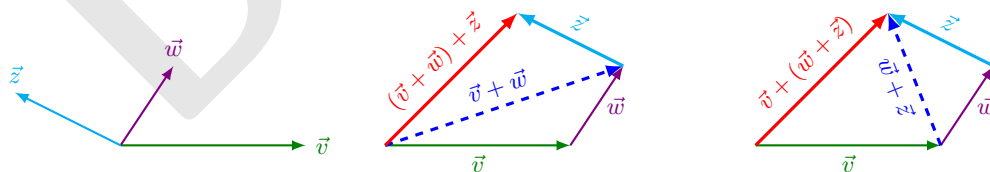


FIGURA 2.5: La imagen ilustra la asociatividad de la suma vectorial.

Dibuje imágenes que ilustren las leyes de distributividad.

Podemos tomar estas propiedades y definir un *espacio vectorial abstracto*. Llamaremos *espacio vectorial* a un conjunto de objetos, llamados *vectores*, con una suma “bien comportada” de sus

elementos y un producto “bien comportado” de sus elementos por escalares. La definición precisa es la siguiente.

Axiomas de espacio vectorial. Sea V un conjunto provisto de dos operaciones

$$\begin{aligned} \text{suma vectorial} \quad + : V \times V &\rightarrow V, & (v, w) &\mapsto v + w, \\ \text{producto de un escalar por un vector} \quad \cdot : \mathbb{K} \times V &\rightarrow V, & (\lambda, v) &\mapsto \lambda \cdot v. \end{aligned}$$

Observe que usualmente escribiremos λv en lugar de $\lambda \cdot v$. Entonces V se llama un *espacio vectorial* sobre $\mathbb{R}$ y sus elementos se llaman *vectores* si se cumple lo siguiente:

- (a) **Asociatividad:** $(u + v) + w = u + (v + w)$ para todo $u, v, w \in V$.
- (b) **Conmutatividad:** $v + w = w + v$ para todo $v, w \in V$.
- (c) **Elemento identidad de la suma:** Existe un elemento $\mathbb{O} \in V$, llamado *identidad aditiva*, tal que para todo $v \in V$ se tiene $\mathbb{O} + v = v + \mathbb{O} = v$.
- (d) **Elemento inverso:** Para todo $v \in V$, tenemos un elemento inverso v' tal que $v + v' = \mathbb{O}$.
- (e) **Elemento identidad de la multiplicación por escalar:** Para todo $v \in V$, se tiene que $1v = v$.
- (f) **Compatibilidad:** Para todo $v \in V$ y $\lambda, \mu \in \mathbb{R}$, se tiene que $(\lambda\mu)v = \lambda(\mu v)$.
- (g) **Leyes de distributividad:** Para todos $v, w \in V$ y $\lambda, \mu \in \mathbb{R}$, se tiene

$$(\lambda + \mu)v = \lambda v + \mu v \quad \text{y} \quad \lambda(v + w) = \lambda v + \lambda w.$$

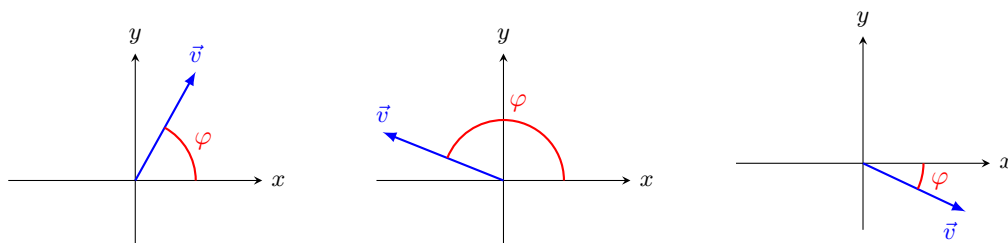
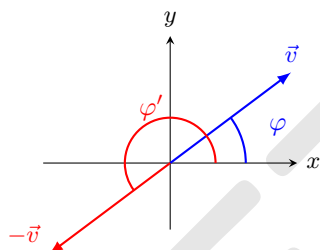
Estos axiomas son fundamentales para el álgebra lineal y volveremos a ellos en el Capítulo 5.1.

Observación 2.2. Si en los axiomas arriba reemplazamos $\mathbb{R}$ por $\mathbb{C}$, entonces obtenemos un *espacio vectorial complejo*. Los discutiremos con más detalles en la Sección ??.

Verifique que $\mathbb{R}^2$ es un espacio vectorial, que su identidad aditiva es $\mathbb{O} = \vec{0}$ y que para todo vector $\vec{v} \in \mathbb{R}^2$, su inverso aditivo es $-\vec{v}$.

Es importante notar que existen espacios vectoriales que no se parecen a $\mathbb{R}^2$ y que no siempre podemos escribir los vectores como columnas como se ve en el siguiente ejemplo.

Ejemplo 2.3. El conjunto de todos los polinomios forma un espacio vectorial (la suma y el múltiplo escalar de polinomios vuelve a ser un polinomio, la suma es asociativa y conmutativa; la identidad aditiva es el polinomio cero y para todo polinomio p , su inverso aditivo es el polinomio $-p$; podemos multiplicar polinomios por escalares y obtener otro polinomio, etc.). En este caso, los vectores son polinomios y no tiene sentido hablar de sus “componentes” o “coordenadas”. (Sin embargo, aprenderemos cómo representar ciertos subespacios del espacio de polinomios como subespacios de algún $\mathbb{R}^n$ en el Capítulo 6.3.)

FIGURA 2.6: Ángulo de un vector con el eje x .FIGURA 2.7: El ángulo de $\vec{v}$ y $-\vec{v}$ con el eje x . Claramente, $\varphi' = \varphi + \pi$.

Después de esta breve excursión sobre espacios vectoriales abstractos, volvamos a $\mathbb{R}^2$. Sabemos que puede identificarse con el plano xy . Esto significa que $\mathbb{R}^2$ tiene más estructura que la de ser solamente un espacio vectorial. Por ejemplo, podemos medir ángulos y longitudes. Observe que estos conceptos *no* aparecen en la definición de espacio vectorial. Son algo adicional a las propiedades de espacio vectorial.

Veamos ahora algunas propiedades geométricas adicionales de los vectores en $\mathbb{R}^2$. Claramente, un vector queda determinado si conocemos su longitud y su ángulo con el eje x . Por el teorema de Pitágoras es claro que la longitud de un vector $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ es $\sqrt{a^2 + b^2}$.

Definición 2.4 (Norma de un vector en $\mathbb{R}^2$). La *longitud* de $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ se denota por $\|\vec{v}\|$. Está dada por

$$\|\vec{v}\| = \sqrt{a^2 + b^2}.$$

Otros nombres para la longitud de $\vec{v}$ son *magnitud de $\vec{v}$* or *norma de $\vec{v}$* .

Como ya mencionamos antes, la pendiente del vector $\vec{v}$ es $\frac{b}{a}$ si $a \neq 0$. Si φ es el ángulo del vector $\vec{v}$ con el eje x , entonces $\tan \varphi = \frac{b}{a}$ si $a \neq 0$. Si $a = 0$, entonces $\varphi = -\frac{\pi}{2}$ o $\varphi = \frac{\pi}{2}$. Recordemos que el rango de $\arctan$ es $(-\pi/2, \pi/2)$, por lo que no podemos simplemente tomar $\arctan$ de la fracción $\frac{a}{b}$ para obtener φ . Observe que $\arctan \frac{b}{a} = \arctan \frac{-b}{-a}$, pero los vectores $\begin{pmatrix} a \\ b \end{pmatrix}$ y $\begin{pmatrix} -a \\ -b \end{pmatrix} = -\begin{pmatrix} a \\ b \end{pmatrix}$ apuntan en direcciones opuestas, por lo que *no* tienen el mismo ángulo con el eje x . De hecho, sus

ángulos difieren en π ; véase la Figura 2.7. De la geometría elemental obtenemos

$$\tan \varphi = \frac{b}{a} \text{ si } a \neq 0 \quad \text{and} \quad \varphi = \begin{cases} \arctan \frac{b}{a} & \text{si } a > 0, \\ \pi + \arctan \frac{b}{a} & \text{si } a < 0, \\ \pi/2 & \text{si } a = 0, b > 0, \\ -\pi/2 & \text{si } a = 0, b < 0. \end{cases}$$

Observe que esta fórmula da ángulos con valores en $[-\pi/2, 3\pi/2)$.

Observación 2.5. Para obtener ángulos con valores en $(-\pi, \pi]$, podemos usar la fórmula

$$\varphi = \begin{cases} \arccos \frac{a}{\sqrt{a^2+b^2}} & \text{si } b > 0, \\ -\arccos \frac{a}{\sqrt{a^2+b^2}} & \text{si } b < 0, \\ \pi & \text{si } a < 0, b = 0. \end{cases}$$

Proposición 2.6 (Propiedades de la norma). Sean $\lambda \in \mathbb{R}$ y $\vec{v}, \vec{w} \in \mathbb{R}^2$. Entonces se cumple lo siguiente:

- (I) $\|\vec{v}\| = 0$ si y solo si $\vec{v} = \vec{0}$.
- (II) $\|\lambda\vec{v}\| = |\lambda|\|\vec{v}\|$,
- (III) $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$ (desigualdad triangular).

Demostración. Sean $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$, $\vec{w} = \begin{pmatrix} c \\ d \end{pmatrix} \in \mathbb{R}^2$ y $\lambda \in \mathbb{R}$.

- (I) Como $\|\vec{v}\| = \sqrt{a^2 + b^2}$, se sigue que $\|\vec{v}\| = 0$ si y solo si $a = 0$ y $b = 0$. Esto ocurre si y solo si $\vec{v} = \vec{0}$.
- (II) $\|\lambda\vec{v}\| = \left\| \lambda \begin{pmatrix} a \\ b \end{pmatrix} \right\| = \left\| \begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix} \right\| = \sqrt{(\lambda a)^2 + (\lambda b)^2} = \sqrt{\lambda^2(a^2 + b^2)} = |\lambda|\sqrt{a^2 + b^2} = |\lambda|\|\vec{v}\|$.
- (III) Posponemos la demostración de la desigualdad triangular hasta el Corolario 2.22, cuando tengamos a nuestra disposición el teorema del coseno. $\square$

Geoméricamente, la desigualdad triangular dice que en el plano el camino más corto para ir de un punto a otro es una línea recta. La Figura 2.8 muestra que es más corto ir directamente desde el origen del vector azul hasta su punta que tomar un desvío a lo largo de $\vec{v}$ y $\vec{w}$. En otras palabras, $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$.

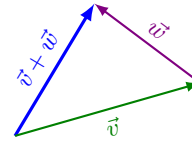


FIGURA 2.8: Desigualdad triangular.

Definición 2.7. Un vector $\vec{v} \in \mathbb{R}^2$ se llama *vector unitario* si $\|\vec{v}\| = 1$.

Observe que todo vector $\vec{v} \neq \vec{0}$ define un vector unitario que apunta en su misma dirección mediante $\|\vec{v}\|^{-1}\vec{v}$.

Observación 2.8. (I) La punta de todo vector unitario está sobre el círculo unitario y, recíprocamente, todo vector cuyo punto inicial es el origen y cuya punta está sobre el círculo unitario es un vector unitario.

(II) Todo vector unitario es de la forma $\begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$ donde φ es su ángulo con el eje x positivo.

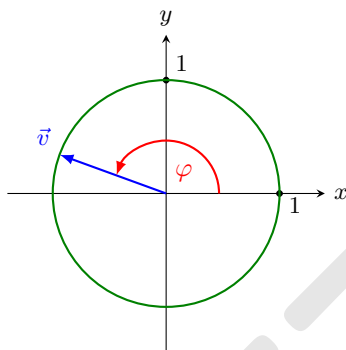


FIGURA 2.9: Vectores unitarios.

Finalmente, definimos dos vectores unitarios muy especiales:

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Claramente, $\vec{e}_1$ es paralelo al eje x , $\vec{e}_2$ es paralelo al eje y and $\|\vec{e}_1\| = \|\vec{e}_2\| = 1$.

Observación 2.9. Todo vector $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ puede escribirse como

$$\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ b \end{pmatrix} = a\vec{e}_1 + b\vec{e}_2.$$

Observación 2.10. Otra notación para $\vec{e}_1$ y $\vec{e}_2$ es $\hat{i}$ y $\hat{j}$.

Debería haber entendido

- el concepto de espacio vectorial abstracto y de vectores,
- el espacio vectorial $\mathbb{R}^2$ y cómo calcular con vectores en $\mathbb{R}^2$,
- la diferencia entre un punto $P(a, b)$ en $\mathbb{R}^2$ y un vector $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ en $\mathbb{R}^2$,
- conceptos geométricos (ángulos, longitud de un vector),
- etc.

Ahora debería ser capaz de

- realizar operaciones algebraicas en el espacio vectorial $\mathbb{R}^2$ y visualizarlas en el plano,
- calcular longitudes y ángulos,
- calcular vectores unitarios y escalar vectores,
- realizar demostraciones abstractas sencillas (por ejemplo, probar que $\mathbb{R}^2$ es un espacio vectorial).
- etc.

Ejercicios.

1. Sean $P(2, 3)$, $Q(-1, 4)$ puntos en $\mathbb{R}^2$ y sea $\vec{v} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$ un vector en $\mathbb{R}^2$.
 - (a) Calcule $\overrightarrow{PQ}$.
 - (b) Calcule $\|\overrightarrow{PQ}\|$.
 - (c) Calcule $\overrightarrow{PQ} + \vec{v}$.
 - (d) Encuentre el ángulo que forma $\vec{v}$ con el eje x .
 - (e) Encuentre el ángulo que forma $\overrightarrow{PQ}$ con el eje x .
2.
 - (a) Determine con la suma vectorial si los puntos $(1, 1)$, $(4, 2)$, $(2, 4)$ y $(-1, 3)$ forman un paralelogramo.
 - (b) Repita el ejercicio anterior con los puntos $(1, -3)$, $(2, 0)$, $(3, -2)$ y $(0, 4)$.
 - (c) Repita el ejercicio anterior con los puntos $(1, 1)$, $(2, 3)$, $(3, 2)$ y $(4, 4)$.

2.2. Producto interno en $\mathbb{R}^2$

En esta sección exploraremos más propiedades geométricas de $\mathbb{R}^2$ e introduciremos el llamado producto interno. Muchas de estas propiedades se trasladan casi literalmente a $\mathbb{R}^3$ y, más generalmente, a $\mathbb{R}^n$. Comencemos con una definición.

Definición 2.11 (Producto interno). Let $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ vectores en $\mathbb{R}^2$. El *producto interno* de $\vec{v}$ y $\vec{w}$ es

$$\langle \vec{v}, \vec{w} \rangle := v_1 w_1 + v_2 w_2.$$

El producto interno también se llama *producto escalar* or *producto punto* y también se puede denotar por $\vec{v} \cdot \vec{w}$.

Usualmente preferimos la notación $\langle \vec{v}, \vec{w} \rangle$, ya que esta notación se usa con frecuencia en física y se extiende de manera natural a espacios vectoriales abstractos con producto interno. Además, la notación con el punto parece sugerir que el producto punto se comporta como un producto usual, mientras que en realidad no es así; véase la Observación 2.14.

Antes de dar propiedades del producto interno y explorar para qué sirve, calculamos primero algunos ejemplos para familiarizarnos con él.

Ejemplos 2.12.

$$(I) \left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \end{pmatrix} \right\rangle = 2 \cdot (-1) + 3 \cdot 5 = -2 + 15 = 13.$$

$$(II) \left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\rangle = 2^2 + 3^2 = 4 + 9 = 13.$$

$$(III) \left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle = 2, \quad \left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = 3.$$

$$(IV) \left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -3 \\ 2 \end{pmatrix} \right\rangle = 0.$$

Proposición 2.13 (Propiedades del producto interno). Sean $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^2$ y $\lambda \in \mathbb{R}$. Entonces se cumple lo siguiente.

$$(I) \langle \vec{v}, \vec{v} \rangle = \|\vec{v}\|^2.$$

En notación de punto: $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$.

$$(II) \langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle.$$

En notación de punto: $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$.

$$(III) \langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle.$$

En notación de punto: $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$.

$$(IV) \langle \lambda \vec{u}, \vec{v} \rangle = \lambda \langle \vec{u}, \vec{v} \rangle.$$

En notación de punto: $(\lambda \vec{u}) \cdot \vec{v} = \lambda(\vec{u} \cdot \vec{v})$.

Demostración. Sean $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ y $\vec{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$.

$$(I) \langle \vec{v}, \vec{v} \rangle = v_1^2 + v_2^2 = \|\vec{v}\|^2.$$

$$(II) \langle \vec{u}, \vec{v} \rangle = u_1 v_1 + u_2 v_2 = v_1 u_1 + v_2 u_2 = \langle \vec{v}, \vec{u} \rangle.$$

$$\begin{aligned} (III) \quad \langle \vec{u}, \vec{v} + \vec{w} \rangle &= \left\langle \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix} \right\rangle \\ &= u_1(v_1 + w_1) + u_2(v_2 + w_2) = u_1 v_1 + u_2 v_2 + u_1 w_1 + u_2 w_2 \\ &= \left\langle \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right\rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle. \end{aligned}$$

$$(IV) \langle \lambda \vec{u}, \vec{v} \rangle = \left\langle \begin{pmatrix} \lambda u_1 \\ \lambda u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\rangle = \lambda u_1 v_1 + \lambda u_2 v_2 = \lambda(u_1 v_1 + u_2 v_2) = \lambda \langle \vec{u}, \vec{v} \rangle. \quad \square$$

Observación 2.14. Observe que la proposición muestra que el producto interno es conmutativo y distributivo, de modo que tiene algunas propiedades del “producto usual” al que estamos acostumbrados por el producto en $\mathbb{R}$ o $\mathbb{C}$, pero hay algunas propiedades que muestran que el producto interno **no** es un producto.

- (a) El producto interno toma dos vectores y devuelve un número, de modo que devuelve un objeto que **no** es del mismo tipo que las dos cosas que introducimos.

- (b) In Example 2.12(iv) vimos que puede ocurrir que $\vec{v} \neq \vec{0}$ y $\vec{w} \neq \vec{0}$, pero aun así $\langle \vec{v}, \vec{w} \rangle = 0$ lo cual es imposible para un producto “decente”.
- (c) Dado un vector $\vec{v} \neq \vec{0}$ y un número $c \in \mathbb{R}$, hay muchas soluciones de la ecuación $\langle \vec{v}, \vec{x} \rangle = c$ para el vector $\vec{x}$, en marcado contraste con el producto usual en $\mathbb{R}$ o $\mathbb{C}$. Mire, por ejemplo, Example 2.12(i) and (ii). Por lo tanto, no tiene sentido escribir algo como $\vec{v}^{-1}$.
- (d) No existe algo como un elemento neutro para la multiplicación escalar.

Veamos ahora por qué el producto interno es útil. De hecho, está relacionado con el ángulo entre dos vectores y nos ayudará a definir proyecciones ortogonales de un vector sobre otro. Comencemos con una definición.

Definición 2.15. Let $\vec{v}, \vec{w}$ vectores en $\mathbb{R}^2$. El *ángulo entre $\vec{v}$ y $\vec{w}$* es el ángulo no negativo más pequeño entre ellos; véase la Figura 2.10. Se denota por $\angle(\vec{v}, \vec{w})$.

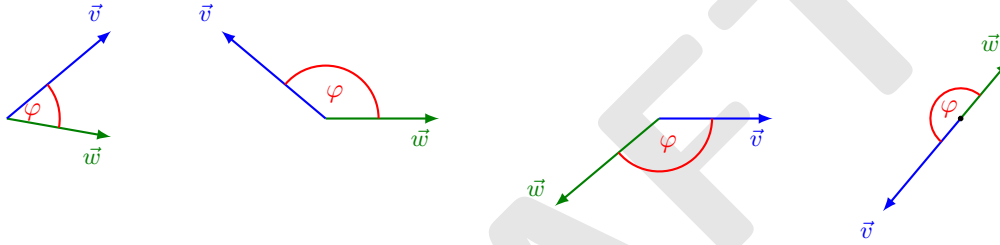


FIGURA 2.10: Ángulo $\varphi := \angle(\vec{v}, \vec{w})$ entre dos vectores $\vec{v}, \vec{w}$.

Las siguientes propiedades del ángulo son fáciles de ver.

- Proposición 2.16.**
- (I) $\angle(\vec{v}, \vec{w}) \in [0, \pi]$ and $\angle(\vec{v}, \vec{w}) = \angle(\vec{w}, \vec{v})$.
 - (II) If $\lambda > 0$, then $\angle(\lambda\vec{v}, \vec{w}) = \angle(\vec{v}, \vec{w})$.
 - (III) If $\lambda < 0$, then $\angle(\lambda\vec{v}, \vec{w}) = \pi - \angle(\vec{v}, \vec{w})$.

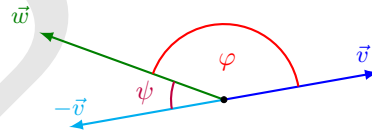


FIGURA 2.11: Ángulo entre el vector $\vec{w}$ y los vectores $\vec{v}$ y $-\vec{v}$. $\varphi = \angle(\vec{w}, \vec{v})$, $\psi = \angle(\vec{w}, -\vec{v}) = \pi - \angle(\vec{w}, \vec{v}) = \pi - \varphi$.

- Definición 2.17.**
- (a) Dos vectores no nulos $\vec{v}$ y $\vec{w}$ se llaman *paralelos* si $\angle(\vec{v}, \vec{w}) = 0$ o π . En este caso usamos la notación $\vec{v} \parallel \vec{w}$.
 - (b) Dos vectores no nulos $\vec{v}$ y $\vec{w}$ se llaman *ortogonales* (or *perpendiculares*) si $\angle(\vec{v}, \vec{w}) = \pi/2$. En este caso usamos la notación $\vec{v} \perp \vec{w}$.

(c) El vector $\vec{0}$ es paralelo y perpendicular a todo vector en $\mathbb{R}^2$.

Las siguientes propiedades deberían ser intuitivamente claras a partir de la geometría. Una demostración formal de (II) y (III) se puede dar fácilmente después del Corolario 2.22. La demostración de (I) se dará después de la Observación 2.26.

Proposición 2.18. *Let $\vec{v}, \vec{w}$ vectores en $\mathbb{R}^2$. Entonces:*

- (I) *Si $\vec{v} \parallel \vec{w}$ y $\vec{w} \neq \vec{0}$, entonces existe $\lambda \in \mathbb{R}$ tal que $\vec{v} = \lambda \vec{w}$.*
- (II) *Si $\vec{v} \parallel \vec{w}$ y $\lambda, \mu \in \mathbb{R}$, entonces también $\lambda \vec{v} \parallel \mu \vec{w}$.*
- (III) *Si $\vec{v} \perp \vec{w}$ y $\lambda, \mu \in \mathbb{R}$, entonces también $\lambda \vec{v} \perp \mu \vec{w}$.*

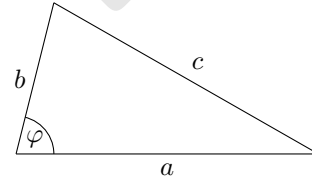
Observación 2.19. (I) Observe que (I) es falsa si no suponemos que $\vec{w} \neq \vec{0}$, porque si $\vec{w} = \vec{0}$, entonces es paralelo a todo vector $\vec{v}$ en $\mathbb{R}^2$, pero no existe $\lambda \in \mathbb{R}$ tal que $\lambda \vec{w}$ pueda llegar a ser distinto de $\vec{0}$.

(II) Observe que la implicación recíproca en (II) y (III) es verdadera solo si $\lambda \neq 0$ y $\mu \neq 0$.

Sin demostración, enunciamos el siguiente teorema, que debería ser conocido.

Teorema 2.20 (Teorema del coseno). *Sean a, b, c los lados de un triángulo y sea φ el ángulo entre los lados a y b . Entonces*

$$c^2 = a^2 + b^2 - 2ab \cos \varphi. \quad (2.1)$$



Teorema 2.21. *Let $\vec{v}, \vec{w} \in \mathbb{R}^2$ and let $\varphi = \angle(\vec{v}, \vec{w})$. Entonces*

$$\langle \vec{v}, \vec{w} \rangle = \|\vec{v}\| \|\vec{w}\| \cos \varphi.$$

Demostración.

Los vectores $\vec{v}$ y $\vec{w}$ definen un triángulo en $\mathbb{R}^2$; véase la Figura 2.12. Ahora aplicamos el teorema del coseno con $a = \|\vec{v}\|$, $b = \|\vec{w}\|$, $c = \|\vec{v} - \vec{w}\|$. Obtenemos

$$\|\vec{v} - \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2 - 2\|\vec{v}\| \|\vec{w}\| \cos \varphi. \quad (2.2)$$

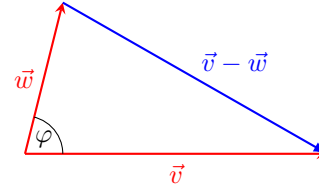


FIGURA 2.12: Triángulo determinado por $\vec{v}$ y $\vec{w}$.

Por otro lado,

$$\begin{aligned} \|\vec{v} - \vec{w}\|^2 &= \langle \vec{v} - \vec{w}, \vec{v} - \vec{w} \rangle = \langle \vec{v}, \vec{v} \rangle - \langle \vec{v}, \vec{w} \rangle - \langle \vec{w}, \vec{v} \rangle + \langle \vec{w}, \vec{w} \rangle = \langle \vec{v}, \vec{v} \rangle - 2\langle \vec{v}, \vec{w} \rangle + \langle \vec{w}, \vec{w} \rangle \\ &= \|\vec{v}\|^2 - 2\langle \vec{v}, \vec{w} \rangle + \|\vec{w}\|^2. \end{aligned} \quad (2.3)$$

La comparación de (2.2) y (2.3) muestra que

$$\|\vec{v}\|^2 + \|\vec{w}\|^2 - 2\|\vec{v}\| \|\vec{w}\| \cos \varphi = \|\vec{v}\|^2 - 2\langle \vec{v}, \vec{w} \rangle + \|\vec{w}\|^2,$$

lo que da la fórmula afirmada. $\square$

Una consecuencia muy importante de este teorema es que ahora podemos determinar si dos vectores son paralelos o perpendiculares entre sí simplemente calculando su producto interno, como se ve en el siguiente corolario.

Corolario 2.22. Sean $\vec{v}, \vec{w} \in \mathbb{R}^2$ y $\varphi = \angle(\vec{v}, \vec{w})$. Entonces:

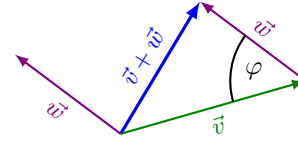
- (I) $\vec{v} \parallel \vec{w} \iff \|\vec{v}\| \|\vec{w}\| = |\langle \vec{v}, \vec{w} \rangle|$.
- (II) $\vec{v} \perp \vec{w} \iff \langle \vec{v}, \vec{w} \rangle = 0$.
- (III) *Desigualdad de Cauchy-Schwarz:* $|\langle \vec{v}, \vec{w} \rangle| \leq \|\vec{v}\| \|\vec{w}\|$.
- (IV) *Desigualdad triangular:*

$$\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|. \quad (2.4)$$

Demostración. Las afirmaciones son claras si uno de los vectores es igual a $\vec{0}$, ya que el vector cero es paralelo y ortogonal a todo vector en $\mathbb{R}^2$. Supongamos ahora que $\vec{v} \neq \vec{0}$ y $\vec{w} \neq \vec{0}$.

- (I) Del Teorema 2.21 tenemos que $|\langle \vec{v}, \vec{w} \rangle| = \|\vec{v}\| \|\vec{w}\|$ si y solo si $|\cos \varphi| = 1$. Esto ocurre si y solo si $\varphi = 0$ o π , es decir, si y solo si $\vec{v}$ y $\vec{w}$ son paralelos.
- (II) Del Teorema 2.21 tenemos que $|\langle \vec{v}, \vec{w} \rangle| = 0$ si y solo si $\cos \varphi = 0$. Esto ocurre si y solo si $\varphi = \pi/2$, es decir, si y solo si $\vec{v}$ y $\vec{w}$ son perpendiculares.
- (III) Por el Teorema 2.21 tenemos que $|\langle \vec{v}, \vec{w} \rangle| = \|\vec{v}\| \|\vec{w}\| |\cos \varphi| \leq \|\vec{v}\| \|\vec{w}\|$ ya que $0 \leq |\cos \varphi| \leq 1$ para $\varphi \in [0, \pi]$.
- (IV) Considere el triángulo cuyos lados son $\vec{v}$, $\vec{w}$ y $\vec{v} + \vec{w}$, y sea φ el ángulo opuesto al lado $\vec{v} + \vec{w}$ (por tanto $\varphi = \pi - \angle(\vec{v}, \vec{w})$). El teorema del coseno da

$$\begin{aligned} \|\vec{v} + \vec{w}\|^2 &= \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\|\vec{v}\| \|\vec{w}\| \cos \varphi \\ &\leq \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\|\vec{v}\| \|\vec{w}\| \\ &= (\|\vec{v}\| + \|\vec{w}\|)^2. \end{aligned}$$



Tomando la raíz cuadrada en ambos lados obtenemos la desigualdad deseada. $\square$

Question 2.1

¿Cuándo se cumple la igualdad en la desigualdad triangular (2.4)? Dibuje una figura y demuestre su afirmación usando los cálculos de la demostración de (IV).

Ejercicio. Demuestre (II) y (III) de la Proposición 2.18 usando el Corolario 2.22.

Ejercicio. (I) Demuestre el Corolario 2.22 (III) sin el teorema del coseno.

Indicación. Empiece con la desigualdad $0 \leq \left\| \|\vec{w}\| \vec{v} - \|\vec{v}\| \vec{w} \right\|^2$ y desarrolle el lado derecho de manera similar a como se hace en la demostración de la Proposición 8.6. Encontrará que

$$0 \leq 2\|\vec{w}\|^2\|\vec{v}\|^2 - 2(\langle \vec{v}, \vec{w} \rangle)^2.$$

(II) Demuestre el Corolario 2.22 (iv) sin el teorema del coseno.

Indicación. Compare con la demostración de la desigualdad triangular en $\mathbb{C}^n$ (Proposición 8.6).

Damos una demostración de (III) y (III) en la Proposición 8.6 sin usar el teorema del coseno, la cual funciona también en el caso complejo.

Ejemplo 2.23. El Teorema 2.21 nos permite calcular fácilmente el ángulo de un vector dado con el eje x y el eje y (véase la Figura 2.13):

$$\cos \varphi_x = \frac{\langle \vec{v}, \vec{e}_1 \rangle}{\|\vec{v}\| \|\vec{e}_1\|}, \quad \cos \varphi_y = \frac{\langle \vec{v}, \vec{e}_2 \rangle}{\|\vec{v}\| \|\vec{e}_2\|}.$$

Si ahora usamos que $\|\vec{e}_1\| = \|\vec{e}_2\| = 1$ and that $\langle \vec{v}, \vec{e}_1 \rangle = v_1$ and $\langle \vec{v}, \vec{e}_2 \rangle = v_2$, entonces podemos simplificar las expresiones a

$$\cos \varphi_x = \frac{v_1}{\|\vec{v}\|}, \quad \cos \varphi_y = \frac{v_2}{\|\vec{v}\|}.$$

Claramente, estas fórmulas también se siguen de trigonometría elemental.

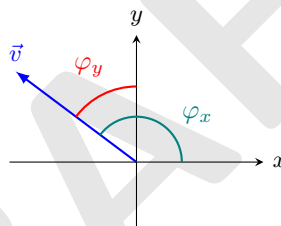


FIGURA 2.13: Ángulo de $\vec{v}$ con los ejes.

Debería haber entendido

- los conceptos de ser paralelo y de ser perpendicular,
- la relación del producto interno con la longitud de un vector y el ángulo entre dos vectores,
- que el producto interno es conmutativo y asociativo, pero que no es un producto,
- etc.

Ahora debería ser capaz de

- calcular el producto interno de dos vectores,
- usar el producto interno para calcular ángulos entre vectores
- usar el producto interno para determinar si dos vectores son paralelos, perpendiculares o ninguna de las dos cosas,
- etc.

Ejercicios.

1. Sea $\vec{v} = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \in \mathbb{R}^2$.
 - (a) Encuentre todos los vectores unitarios cuya dirección es opuesta a la de $\vec{v}$.
 - (b) Encuentre todos los vectores de longitud 3 que tienen la misma dirección que $\vec{v}$.
 - (c) Encuentre todos los vectores que tienen la misma dirección que $\vec{v}$ y que tienen doble longitud de $\vec{v}$.
 - (d) Encuentre todos los vectores con norma 2 que son ortogonales a $\vec{v}$.
2. Para los siguientes vectores $\vec{u}$ y $\vec{v}$ decida si son ortogonales, paralelos o ninguno de los dos. Calcule el coseno del ángulo entre ellos. Si son paralelos, encuentre números reales λ y μ tales que $\vec{v} = \lambda\vec{u}$ y $\vec{u} = \mu\vec{v}$.

$$(a) \quad \vec{v} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \quad \vec{u} = \begin{pmatrix} 5 \\ -2 \end{pmatrix},$$

$$(b) \quad \vec{v} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \quad \vec{u} = \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

$$(c) \quad \vec{v} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \quad \vec{u} = \begin{pmatrix} -8 \\ 6 \end{pmatrix},$$

$$(d) \quad \vec{v} = \begin{pmatrix} -6 \\ 4 \end{pmatrix}, \quad \vec{u} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}.$$

3. (a) Para las siguientes parejas $\vec{v}$ y $\vec{w}$ encuentre todos los $\alpha \in \mathbb{R}$ tal que $\vec{v}$ y $\vec{w}$ son paralelos:

$$(i) \quad \vec{v} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} \alpha \\ -2 \end{pmatrix}, \quad (ii) \quad \vec{v} = \begin{pmatrix} 2 \\ \alpha \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} 1 + \alpha \\ 2 \end{pmatrix},$$

$$(iv) \quad \vec{v} = \begin{pmatrix} \alpha \\ 5 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} 1 + \alpha \\ 2 \end{pmatrix}, \quad (ii) \quad \vec{v} = \begin{pmatrix} 2 \\ \alpha \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} 1 + \alpha \\ 2\alpha \end{pmatrix},$$

- (b) Para las siguientes parejas $\vec{v}$ y $\vec{w}$ encuentre todos los $\alpha \in \mathbb{R}$ tal que $\vec{v}$ y $\vec{w}$ son perpendiculares:

$$(i) \quad \vec{v} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} \alpha \\ -2 \end{pmatrix}, \quad (ii) \quad \vec{v} = \begin{pmatrix} 2 \\ \alpha \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} \alpha \\ 2 \end{pmatrix}, \quad (iii) \quad \vec{v} = \begin{pmatrix} \alpha \\ 5 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} 1 + \alpha \\ 2 \end{pmatrix}.$$

4. Sean $\vec{a} = \begin{pmatrix} 2 \\ \alpha \end{pmatrix}$ y $\vec{b} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

- (a) Encuentre todos los $\alpha \in \mathbb{R}$ tales que:

$$(I) \quad \vec{a} \parallel \vec{b};$$

$$(II) \quad \vec{a} \perp \vec{b};$$

$$(III) \quad \text{el ángulo entre } \vec{a} \text{ y } \vec{b} \text{ es } \frac{\pi}{6};$$

$$(IV) \quad \text{el ángulo entre } \vec{a} \text{ y } \vec{b} \text{ es } \frac{\pi}{4};$$

$$(V) \quad \text{el ángulo entre } \vec{a} \text{ y } \vec{b} \text{ es } \frac{5\pi}{6}.$$

- (b) ¿Hacia donde tiende el ángulo entre $\vec{a}$ y $\vec{b}$ cuando $\alpha \rightarrow \infty$ ó $\alpha \rightarrow -\infty$?

Haga un dibujo de cada caso.

2.3. Proyecciones ortogonales en $\mathbb{R}^2$

Sean $\vec{v}$ y $\vec{w}$ vectores en $\mathbb{R}^2$ y $\vec{w} \neq \vec{0}$. Geométricamente, tenemos una intuición de lo que debería ser la proyección ortogonal de $\vec{v}$ sobre $\vec{w}$ y de que deberíamos poder construirla como se describe en el siguiente procedimiento: Movemos $\vec{v}$ de tal manera que su punto inicial coincida con el de $\vec{w}$. Luego extendemos $\vec{w}$ a una recta y construimos una recta que pasa por la punta de $\vec{v}$ y es perpendicular a $\vec{w}$. El vector desde el punto inicial hasta la intersección de las dos rectas debería ser entonces la *proyección ortogonal de $\vec{v}$ sobre $\vec{w}$* . véase la Figura 2.14

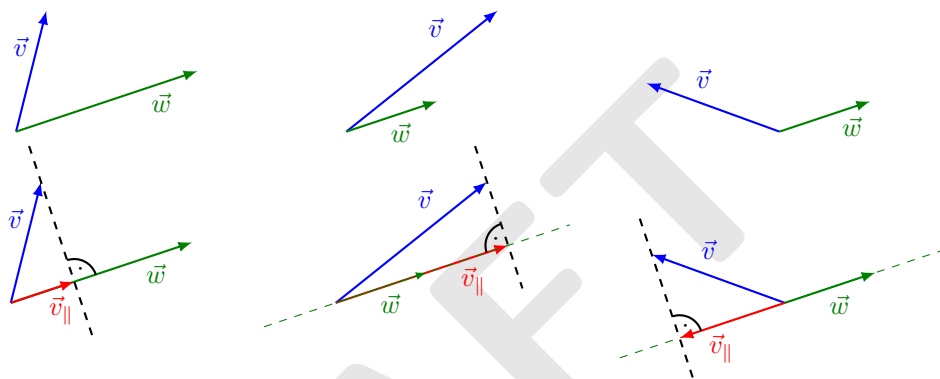


FIGURA 2.14: Algunos ejemplos de la proyección ortogonal de $\vec{v}$ sobre $\vec{w}$ en $\mathbb{R}^2$.

Este procedimiento descompone el vector $\vec{v}$ en una parte paralela a $\vec{w}$ y una parte perpendicular a $\vec{w}$, de manera que su suma nos devuelve $\vec{v}$. La parte paralela es la *proyección ortogonal de $\vec{v}$ sobre $\vec{w}$* .

En el siguiente teorema damos el significado preciso de la proyección ortogonal, mostramos que siempre existe una descomposición como la descrita arriba e incluso derivamos una fórmula para la proyección ortogonal. Una versión más general de este teorema es el Teorema 7.30.

Teorema 2.24 (Proyección ortogonal). Sean $\vec{v}$ y $\vec{w}$ vectores en $\mathbb{R}^2$ y $\vec{w} \neq \vec{0}$. Entonces existen vectores determinados de manera única $\vec{v}_{\parallel}$ y $\vec{v}_{\perp}$ (véase la Figura 2.15) tales que

$$\vec{v}_{\parallel} \parallel \vec{w}, \quad \vec{v}_{\perp} \perp \vec{w} \quad \text{y} \quad \vec{v} = \vec{v}_{\parallel} + \vec{v}_{\perp}. \quad (2.5)$$

El vector $\vec{v}_{\parallel}$ se llama la proyección ortogonal de $\vec{v}$ sobre $\vec{w}$ y está dado por

$$\vec{v}_{\parallel} = \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w}. \quad (2.6)$$

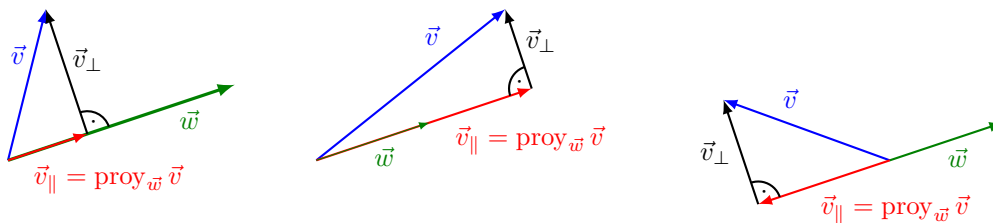


FIGURA 2.15: Ejemplos de descomposiciones de $\vec{v}$ como $\vec{v} = \vec{v}_{\parallel} + \vec{v}_{\perp}$ con $\vec{v}_{\parallel} \parallel \vec{w}$ y $\vec{v}_{\perp} \perp \vec{w}$. Observe que por definición $\vec{v}_{\parallel} = \text{proy}_{\vec{w}} \vec{v}$.

Demostración. Supongamos que tenemos vectores $\vec{v}_{\parallel}$ y $\vec{v}_{\perp}$ que satisfacen (2.5). Como $\vec{v}_{\parallel}$ y $\vec{w}$ son paralelos por definición y como $\vec{w} \neq \vec{0}$, existe $\lambda \in \mathbb{R}$ tal que $\vec{v}_{\parallel} = \lambda \vec{w}$; por lo tanto, para encontrar $\vec{v}_{\parallel}$ basta determinar λ . Para esto, observamos que $\vec{v} = \lambda \vec{w} + \vec{v}_{\perp}$ por (2.5). Tomar el producto interno en ambos lados con $\vec{w}$ conduce a

$$\begin{aligned} \langle \vec{v}, \vec{w} \rangle &= \langle \lambda \vec{w} + \vec{v}_{\perp}, \vec{w} \rangle = \langle \lambda \vec{w}, \vec{w} \rangle + \underbrace{\langle \vec{v}_{\perp}, \vec{w} \rangle}_{= 0 \text{ ya que } \vec{v}_{\perp} \perp \vec{w}} = \langle \lambda \vec{w}, \vec{w} \rangle = \lambda \langle \vec{w}, \vec{w} \rangle = \lambda \|\vec{w}\|^2 \\ \implies \lambda &= \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2}. \end{aligned}$$

Así, si existe una representación de $\vec{v}$ como suma de la forma (2.5), entonces la única posibilidad es

$$\vec{v}_{\parallel} = \lambda \vec{w} = \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w} \quad \text{y} \quad \vec{v}_{\perp} = \vec{v} - \vec{v}_{\parallel} = \vec{v} - \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w}.$$

Esto ya prueba la unicidad de los vectores $\vec{v}_{\parallel}$ y $\vec{v}_{\perp}$. Queda mostrar que, en efecto, tienen las propiedades deseadas. Claramente, por construcción $\vec{v}_{\parallel}$ es paralelo a $\vec{w}$ y $\vec{v} = \vec{v}_{\parallel} + \vec{v}_{\perp}$, ya que definimos $\vec{v}_{\perp} = \vec{v} - \vec{v}_{\parallel}$. Queda verificar que $\vec{v}_{\perp}$ es ortogonal a $\vec{w}$. Esto se sigue de

$$\langle \vec{v}_{\perp}, \vec{w} \rangle = \left\langle \vec{v} - \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w}, \vec{w} \right\rangle = \langle \vec{v}, \vec{w} \rangle - \left\langle \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w}, \vec{w} \right\rangle = \langle \vec{v}, \vec{w} \rangle - \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \langle \vec{w}, \vec{w} \rangle = 0$$

donde en el último paso usamos que $\langle \vec{w}, \vec{w} \rangle = \|\vec{w}\|^2$. $\square$

Notación 2.25. En lugar de $\vec{v}_{\parallel}$ a menudo escribimos $\text{proy}_{\vec{w}} \vec{v}$, en particular cuando queremos enfatizar sobre qué vector estamos proyectando.

Observación 2.26. (I) $\text{proy}_{\vec{w}} \vec{v}$ depende solamente de la dirección de $\vec{w}$. **No** depende de su longitud.

(II) Para todo $c \in \mathbb{R}$, se tiene que $\text{proy}_{\vec{w}}(c\vec{v}) = c \text{proy}_{\vec{w}} \vec{v}$.

(III) Como casos especiales de lo anterior, tenemos $\text{proy}_{\vec{w}}(-\vec{v}) = -\text{proy}_{\vec{w}} \vec{v}$ y $\text{proy}_{-\vec{w}} \vec{v} = \text{proy}_{\vec{w}} \vec{v}$.

(IV) $\vec{v} \parallel \vec{w} \implies \text{proy}_{\vec{w}} \vec{v} = \vec{v}$.

(V) $\vec{v} \perp \vec{w} \implies \text{proy}_{\vec{w}} \vec{v} = \vec{0}$.

(vi) $\text{proy}_{\vec{w}} \vec{v}$ es el único vector en $\mathbb{R}^2$ tal que

$$(\vec{v} - \text{proy}_{\vec{w}} \vec{v}) \perp \vec{v} \quad \text{y} \quad \text{proy}_{\vec{w}} \vec{v} \parallel \vec{w}.$$

Demostración. (I): Por nuestra intuición geométrica, esto debería ser claro. Demos una demostración formal. Supongamos que queremos proyectar $\vec{v}$ sobre $c\vec{w}$ para algún $c \in \mathbb{R} \setminus \{0\}$. Entonces

$$\text{proy}_{c\vec{w}} \vec{v} = \frac{\langle \vec{v}, c\vec{w} \rangle}{\|c\vec{w}\|^2} (c\vec{w}) = \frac{c\langle \vec{v}, \vec{w} \rangle}{c^2 \|\vec{w}\|^2} (c\vec{w}) = \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w} = \text{proy}_{\vec{w}} \vec{v}.$$

Convénzase gráficamente de que no importa si proyectamos $\vec{v}$ sobre $\vec{w}$, sobre $5\vec{w}$ o sobre $-\frac{7}{5}\vec{w}$; solo importa la dirección de $\vec{w}$, no su longitud.

(II): De nuevo, por consideraciones geométricas, esto debería ser claro. El cálculo correspondiente es

$$\text{proy}_{\vec{w}} (c\vec{v}) = \frac{\langle c\vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w} = \frac{c\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w} = c \text{proy}_{\vec{w}} \vec{v}.$$

(III) se sigue directamente de (I) y (II).

(IV), (V) y (VI) follow from the uniqueness of the descomposición of the vector $\vec{v}$ as sum of a vector paralelo y a vector perpendicular to $\vec{w}$. $\square$

Ahora la demostración de la Proposición 2.18 (I) se sigue fácilmente.

Demostración de la Proposición 2.18 (I). Tenemos que mostrar que si $\vec{v} \parallel \vec{w}$ y si $\vec{w} \neq \vec{0}$, entonces existe $\lambda \in \mathbb{R}$ tal que $\vec{w} = \lambda \vec{v}$. De la Observación 2.26 (IV) se sigue que $\vec{v} = \text{proy}_{\vec{w}} \vec{v} = \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w}$, por lo tanto la afirmación se sigue si podemos escoger $\lambda = \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2}$. $\square$

Observación 2.27 (Distancia de un punto a una recta). En Teorema 7.39 probaremos lo siguiente: Sea $L = \{t\vec{v} : t \in \mathbb{R}\}$ una recta que *pasa por el origen* y sea P un punto en $\mathbb{R}^2$. El punto Q en L más cercano a P satisface $\vec{0Q} = \text{proy}_{\vec{v}} \vec{0P}$. La distancia de P a la recta L es $\|\vec{PQ}\|$, es decir,

$$\text{dist}(P, L) = \|\vec{0P} - \text{proy}_{\vec{v}} \vec{0P}\|.$$

En palabras: si proyectamos el vector $\vec{0P}$ sobre la recta L (es decir, sobre su vector director), entonces la punta Q de esta proyección es el punto en la recta L que está más cercano al punto P .

¿Como cambia lo de arriba si L no pasa por el origen?

Terminamos esta sección con algunos ejemplos.

Ejemplo 2.28. Sean $\vec{u} = 2\vec{e}_1 + 3\vec{e}_2$, $\vec{v} = 4\vec{e}_1 - \vec{e}_2$.

(I) $\text{proy}_{\vec{e}_1} \vec{u} = \frac{\langle \vec{u}, \vec{e}_1 \rangle}{\|\vec{e}_1\|^2} \vec{e}_1 = \frac{2}{1^2} \vec{e}_1 = 2\vec{e}_1.$

(II) $\text{proy}_{\vec{e}_2} \vec{u} = \frac{\langle \vec{u}, \vec{e}_2 \rangle}{\|\vec{e}_2\|^2} \vec{e}_2 = \frac{3}{1^2} \vec{e}_2 = 3\vec{e}_2.$

(III) De manera similar, podemos calcular $\text{proy}_{\vec{e}_1} \vec{v} = 4\vec{e}_1$, $\text{proy}_{\vec{e}_2} \vec{v} = -\vec{e}_2$.

$$(IV) \text{proy}_{\vec{u}} \vec{v} = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\|^2} \vec{u} = \frac{\left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ -1 \end{pmatrix} \right\rangle}{\|\vec{u}\|^2} \vec{u} = \frac{8-3}{2^2+3^2} \vec{u} = \frac{5}{13} \vec{u} = \frac{5}{13} \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

$$(V) \text{proy}_{\vec{v}} \vec{u} = \frac{\langle \vec{v}, \vec{u} \rangle}{\|\vec{v}\|^2} \vec{v} = \frac{\left\langle \begin{pmatrix} 4 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\rangle}{4^2+(-1)^2} \vec{v} = \frac{8-3}{17} \vec{v} = \frac{5}{17} \vec{v} = \frac{5}{17} \begin{pmatrix} 4 \\ -1 \end{pmatrix}.$$

Ejemplo 2.29 (Ángulo con los ejes de coordenadas). Sean $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 \setminus \{\vec{0}\}$. Entonces $\cos \angle(\vec{v}, \vec{e}_1) = \frac{a}{\|\vec{v}\|}$, $\cos \angle(\vec{v}, \vec{e}_2) = \frac{b}{\|\vec{v}\|}$, por ende

$$\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix} = \|\vec{v}\| \begin{pmatrix} \cos \angle(\vec{v}, \vec{e}_1) \\ \cos \angle(\vec{v}, \vec{e}_2) \end{pmatrix} = \|\vec{v}\| \begin{pmatrix} \cos \varphi_x \\ \cos \varphi_y \end{pmatrix}$$

y

proyección de $\vec{v}$ sobre el eje $x = \text{proy}_{\vec{e}_1} \vec{v} = \|\vec{v}\| \cos \angle(\vec{v}, \vec{e}_1) \vec{e}_1 = \|\vec{v}\| \cos \varphi_x \vec{e}_1$,

proyección de $\vec{v}$ sobre el eje $y = \text{proy}_{\vec{e}_2} \vec{v} = \|\vec{v}\| \cos \angle(\vec{v}, \vec{e}_2) \vec{e}_2 = \|\vec{v}\| \cos \varphi_y \vec{e}_2$.

Question 2.2

Sean $\vec{w}$ be a vector in $\mathbb{R}^2 \setminus \{\vec{0}\}$.

- (I) ¿Puede describir geoméricamente todos los vectores $\vec{v}$ cuya proyección sobre $\vec{w}$ es igual al vector $\vec{0}$?
- (II) ¿Puede describir geoméricamente todos los vectores $\vec{v}$ cuya proyección sobre $\vec{w}$ tiene longitud 2?
- (III) ¿Puede describir geoméricamente todos los vectores $\vec{v}$ cuya proyección sobre $\vec{w}$ tiene longitud $3\|\vec{w}\|$?

Debería haber entendido

- el concepto de proyecciones ortogonales en $\mathbb{R}^2$,
- por qué la proyección ortogonal de $\vec{w}$ sobre $\vec{w}$ no depende de la longitud de $\vec{w}$,
- etc.

Ahora debería ser capaz de

- calcular la proyección de un vector dado sobre otro vector,
- calcular vectores con una proyección dada sobre otro vector,
- etc.

Ejercicios.

1. Sean $\vec{a} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ y $\vec{b} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$.
 - (a) Calcule $\text{proy}_{\vec{b}} \vec{a}$ y $\text{proy}_{\vec{a}} \vec{b}$.
 - (b) Encuentre todos los vectores $\vec{v} \in \mathbb{R}^2$ tal que $\|\text{proy}_{\vec{a}} \vec{v}\| = 0$. Describa este conjunto geoméricamente.
 - (c) Encuentre todos los vectores $\vec{v} \in \mathbb{R}^2$ tal que $\|\text{proy}_{\vec{a}} \vec{v}\| = 2$. Describa este conjunto geoméricamente.
 - (d) ¿Existe un vector $\vec{x}$ tal que $\text{proy}_{\vec{a}} \vec{x} \parallel \vec{b}$?
 ¿Existe un vector $\vec{x}$ tal que $\text{proy}_{\vec{x}} \vec{a} \parallel \vec{b}$?
 ¿Existe un vector $\vec{x}$ tal que $\text{proy}_{\vec{x}} \vec{a} = \vec{b}$?

2.4. Vectores en $\mathbb{R}^n$

En esta sección extendemos nuestros cálculos de $\mathbb{R}^2$ a $\mathbb{R}^n$. Si $n = 3$, entonces obtenemos $\mathbb{R}^3$, que usualmente sirve como modelo de nuestro mundo físico cotidiano y con el cual probablemente ya está familiarizado por las clases de física. Discutiremos $\mathbb{R}^3$ y algunas de sus particularidades con más detalle en la Sección 2.5.

Primero, definamos $\mathbb{R}^n$.

Definición 2.30. Para $n \in \mathbb{N}$ definimos el conjunto

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \dots, x_n \in \mathbb{R} \right\}.$$

De nuevo, podemos pensar en los vectores como flechas. Como en $\mathbb{R}^2$, podemos identificar cada punto de $\mathbb{R}^n$ con la flecha que empieza en el origen del sistema de coordenadas y termina en el punto dado. El conjunto de todas las flechas con la misma longitud y la misma dirección se llama

un vector en $\mathbb{R}^n$. Así, todo punto $P(p_1, \dots, p_n)$ define un vector $\vec{v} = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$ y viceversa. Como antes,

a veces denotamos los vectores como $(p_1, \dots, p_n)^t$ para ahorrar espacio (vertical). El superíndice t significa “transpuesto”.

$\mathbb{R}^n$ se convierte en un espacio vectorial con las operaciones

$$\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n, \vec{v} + \vec{w} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix}, \quad \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n, c\vec{v} = \begin{pmatrix} cv_1 \\ \vdots \\ cv_n \end{pmatrix}. \quad (2.7)$$

Ejercicio. Muestre que $\mathbb{R}^n$ es un espacio vectorial. Es decir, tiene que mostrar que se cumplen los axiomas de espacio vectorial de la página 31.

Como en $\mathbb{R}^2$, podemos definir la norma de un vector, el ángulo entre dos vectores y un producto interno. Observe que la definición del ángulo entre dos vectores no es diferente de la de $\mathbb{R}^2$, ya que cuando se nos dan dos vectores, estos siempre están en un plano común que nos podemos imaginar como una especie de $\mathbb{R}^2$ rotado. Demos ahora las definiciones formales.

Definición 2.31 (Producto interno; norma de un vector). Para los vectores $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ y

$\vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$, el *producto interno* (o *producto escalar* o *producto punto*) se define como

$$\langle \vec{v}, \vec{w} \rangle = \left\langle \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \right\rangle = v_1 w_1 + \cdots + v_n w_n.$$

The *longitud* de $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$ se denota por $\|\vec{v}\|$ y está dada por

$$\|\vec{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}.$$

Otros nombres para la longitud de $\vec{v}$ son *magnitud de $\vec{v}$* or *norma de $\vec{v}$* .

Como en $\mathbb{R}^2$, tenemos las siguientes propiedades:

- (I) *Simetría del producto interno:* Para todos los vectores $\vec{v}, \vec{w} \in \mathbb{R}^n$, tenemos que $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$.
- (II) *Bilinealidad del producto interno:* Para todos los vectores $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ y todo $c \in \mathbb{R}$, tenemos que $\langle \vec{u}, \vec{v} + c\vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + c\langle \vec{u}, \vec{w} \rangle$.
- (III) *Relación del producto interno con el ángulo entre vectores:* Sean $\vec{v}, \vec{w} \in \mathbb{R}^n$ y sea $\varphi = \angle(\vec{v}, \vec{w})$. Entonces

$$\langle \vec{v}, \vec{w} \rangle = \|\vec{v}\| \|\vec{w}\| \cos \varphi.$$

En particular, tenemos (cf. Proposición 2.18):

- (a) $\vec{v} \parallel \vec{w} \iff \angle(\vec{v}, \vec{w}) \in \{0, \pi\} \iff |\langle \vec{v}, \vec{w} \rangle| = \|\vec{v}\| \|\vec{w}\|,$
- (b) $\vec{v} \perp \vec{w} \iff \angle(\vec{v}, \vec{w}) = \pi/2 \iff \langle \vec{v}, \vec{w} \rangle = 0.$

Observación 2.32. En espacios abstractos con producto interno, el producto interno se usa de hecho para *definir* la ortogonalidad.

- (IV) *Relación del producto interno con la norma:* Para todo vector $\vec{v} \in \mathbb{R}^n$, tenemos $\|\vec{v}\|^2 = \langle \vec{v}, \vec{v} \rangle$.

- (v) *Propiedades de la norma:* Para todos los vectores $\vec{v}, \vec{w} \in \mathbb{R}^n$ y escalares $c \in \mathbb{R}$, tenemos que $\|c\vec{v}\| = |c|\|\vec{v}\|$ y $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$.
- (vi) *Proyecciones ortogonales de un vector sobre otro:* Para todos los vectores $\vec{v}, \vec{w} \in \mathbb{R}^n$ con $\vec{w} \neq \vec{0}$ la proyección ortogonal de $\vec{v}$ sobre $\vec{w}$ es

$$\text{proy}_{\vec{w}} \vec{v} = \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w}. \quad (2.8)$$

Como en $\mathbb{R}^2$, tenemos n “vectores especiales” que son paralelos a los ejes de coordenadas y tienen norma 1:

$$\vec{e}_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \vec{e}_n := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

En el caso especial $n = 3$, los vectores $\vec{e}_1, \vec{e}_2$ y $\vec{e}_3$ se denotan a veces por $\hat{i}, \hat{j}, \hat{k}$.

Para un vector dado $\vec{v} \neq \vec{0}$, ahora podemos determinar fácilmente sus proyecciones sobre los n ejes de coordenadas y su ángulo con los ejes de coordenadas. Por (2.8), la proyección sobre el eje x_j es

$$\text{proy}_{\vec{e}_j} \vec{v} = v_j \vec{e}_j.$$

Sea φ_j el ángulo entre $\vec{v}$ y el eje x_j . Entonces

$$\varphi_j = \angle(\vec{v}, \vec{e}_j) \implies \cos \varphi_j = \frac{\langle \vec{v}, \vec{e}_j \rangle}{\|\vec{v}\| \|\vec{e}_j\|} = \frac{v_j}{\|\vec{v}\|}.$$

Se sigue que $\vec{v} = \|\vec{v}\| \begin{pmatrix} \cos \varphi_1 \\ \vdots \\ \cos \varphi_n \end{pmatrix}$. A veces se usa la notación

$$\hat{v} := \frac{\vec{v}}{\|\vec{v}\|} = \begin{pmatrix} \cos \varphi_1 \\ \vdots \\ \cos \varphi_n \end{pmatrix}$$

para el vector unitario que apunta en la misma dirección que $\vec{v}$. Claramente $\|\hat{v}\| = 1$ porque $\|\hat{v}\| = \|\|\vec{v}\|^{-1}\vec{v}\| = \|\vec{v}\|^{-1}\|\vec{v}\| = 1$. Por lo tanto, $\hat{v}$ es efectivamente un vector unitario que apunta en la misma dirección que el vector original $\vec{v}$.

Debería haber entendido

- el espacio vectorial $\mathbb{R}^n$ y los vectores en $\mathbb{R}^n$,
- los conceptos geométricos (ángulos, longitud de un vector) en $\mathbb{R}^n$,
- que $\mathbb{R}^2$ del capítulo 2.1 es un caso especial de $\mathbb{R}^n$ de esta sección,
- etc.

Ahora debería ser capaz de

- realizar operaciones algebraicas en el espacio vectorial $\mathbb{R}^3$ y, en el caso $n = 3$, visualizarlas en el espacio,
- calcular longitudes y ángulos,
- calcular vectores unitarios y escalar vectores,
- realizar demostraciones abstractas sencillas (por ejemplo, demostrar que $\mathbb{R}^n$ es un espacio vectorial).
- etc.

Ejercicios.

1. Sean $\vec{a} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 3 \end{pmatrix}$ y $\vec{b} = \begin{pmatrix} 0 \\ 4 \\ 5 \\ 1 \end{pmatrix}$. Calcular:

(a) $4\vec{a} + 3\vec{b}$.

(b) $\|3\vec{a} - 2\vec{b}\|$.

(c) $\langle \vec{a} - \vec{b} + 3\vec{e}_1, \vec{b} - 5\vec{e}_4 + \vec{e}_3 \rangle$.

(d) $\text{proy}_{\vec{b}} \vec{a}$.

2.5. Vectores en $\mathbb{R}^3$ y el producto cruz

El espacio $\mathbb{R}^3$ es muy importante, ya que se usa en mecánica para modelar el espacio en el que vivimos. En $\mathbb{R}^3$ podemos definir una operación adicional con vectores, el llamado *producto cruz*. Otro nombre para él es *producto vectorial*. Toma dos vectores y devuelve otro vector. Tiene varias propiedades que lo hacen parecer un producto; sin embargo, veremos que **no** es un producto. Esta es su definición.

Definición 2.33 (Producto cruz). Sea $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \in \mathbb{R}^3$. El *producto cruz* (o *producto vectorial* o *producto exterior*) is

$$\vec{v} \times \vec{w} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} := \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}.$$

Otra notación para el producto cruz es $\vec{v} \wedge \vec{w}$.

Una manera de recordar esta fórmula es la siguiente. Escriba la primera y la segunda componente de los vectores debajo de ellos, de modo que formalmente se obtenga una columna de 5 componentes. Luego trace cruces como en el esquema de abajo, comenzando con la cruz formada por una línea de v_2 a w_3 y luego de w_2 a v_3 . Cada línea representa un producto de las componentes correspondientes; si la línea va de arriba a la izquierda hacia abajo a la derecha, se cuenta con signo positivo; si va

de arriba a la derecha hacia abajo a la izquierda, se cuenta con signo negativo.

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}.$$

The producto cruz is defined only in $\mathbb{R}^3$!

Antes de reunir algunas propiedades sencillas del producto cruz, calculemos algunos ejemplos.

Ejemplos 2.34. Let $\vec{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix}$.

$$\blacksquare \vec{u} \times \vec{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 2 \cdot 7 - 3 \cdot 6 \\ 3 \cdot 5 - 1 \cdot 7 \\ 1 \cdot 6 - 2 \cdot 5 \end{pmatrix} = \begin{pmatrix} 14 - 18 \\ 15 - 7 \\ 6 - 10 \end{pmatrix} = \begin{pmatrix} -4 \\ 8 \\ -4 \end{pmatrix},$$

$$\blacksquare \vec{v} \times \vec{u} = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} \times \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \cdot 3 - 7 \cdot 2 \\ 7 \cdot 1 - 5 \cdot 3 \\ 5 \cdot 2 - 6 \cdot 1 \end{pmatrix} = \begin{pmatrix} 18 - 14 \\ 7 - 15 \\ 10 - 6 \end{pmatrix} = \begin{pmatrix} 4 \\ -8 \\ 4 \end{pmatrix},$$

$$\blacksquare \vec{v} \times \vec{e}_1 = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \cdot 0 - 7 \cdot 0 \\ 7 \cdot 1 - 5 \cdot 0 \\ 5 \cdot 0 - 6 \cdot 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 7 \\ -6 \end{pmatrix},$$

$$\blacksquare \vec{v} \times \vec{v} = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} \times \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 6 \cdot 7 - 7 \cdot 6 \\ 7 \cdot 5 - 5 \cdot 7 \\ 5 \cdot 6 - 6 \cdot 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Proposición 2.35 (Propiedades del producto cruz). Sean $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$ y sea $c \in \mathbb{R}$. Entonces:

- (I) $\vec{u} \times \vec{0} = \vec{0} \times \vec{u} = \vec{0}$.
- (II) $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$.
- (III) $\vec{u} \times (\vec{v} + \vec{w}) = (\vec{u} \times \vec{v}) + (\vec{u} \times \vec{w})$.
- (IV) $(c\vec{u}) \times \vec{v} = c(\vec{u} \times \vec{v})$.
- (V) $\vec{u} \parallel \vec{v} \iff \vec{u} \times \vec{v} = \vec{0}$. En particular, $\vec{v} \times \vec{v} = \vec{0}$.
- (VI) $\langle \vec{u}, \vec{v} \times \vec{w} \rangle = \langle \vec{u} \times \vec{v}, \vec{w} \rangle$.
- (VII) $\langle \vec{u}, \vec{u} \times \vec{v} \rangle = 0$ and $\langle \vec{v}, \vec{u} \times \vec{v} \rangle = 0$, in particular

$$\boxed{\vec{v} \perp \vec{v} \times \vec{u}, \quad \vec{u} \perp \vec{v} \times \vec{u}}$$

En otras palabras: el vector $\vec{v} \times \vec{w}$ es ortogonal tanto a $\vec{v}$ como a $\vec{w}$.

Demostración. Las demostraciones de las fórmulas (i) – (iv) son cálculos sencillos (¡debería hacerlos!).

- (v) La implicación “ $\implies$ ” es fácil de verificar. La otra dirección se sigue fácilmente del Teorema 2.37 de abajo. O se puede ver de la siguiente manera: Supongamos que $\vec{w} \times \vec{v} = \vec{0}$. Si $\vec{v} = \vec{0}$, entonces claramente $\vec{v} \parallel \vec{w}$. Si $\vec{v} \neq \vec{0}$, escribimos $\vec{w} = \vec{a} + \vec{b}$, donde $\vec{a} = \text{proy}_{\vec{v}} \vec{w}$ y $\vec{b} \perp \vec{v}$. Tenemos que mostrar que $\vec{b} = \vec{0}$. Usando que $\vec{v} \times \vec{a} = \vec{0}$ (porque son paralelos), obtenemos que

$$\vec{0} = \vec{v} \times \vec{w} = \vec{v} \times (\vec{a} + \vec{b}) = (\vec{v} \times \vec{a}) + (\vec{v} \times \vec{b}) = \vec{v} \times \vec{b} = \begin{pmatrix} v_2 b_3 - v_3 b_2 \\ v_3 b_1 - v_1 b_3 \\ v_1 b_2 - v_2 b_1 \end{pmatrix}.$$

Además sabemos que $\langle \vec{v}, \vec{b} \rangle = 0$. Así, en total tenemos cuatro ecuaciones lineales para las tres componentes b_1, b_2, b_3 de $\vec{b}$:

$$v_2 b_3 - v_3 b_2 = 0, \quad v_3 b_1 - v_1 b_3 = 0, \quad v_1 b_2 - v_2 b_1 = 0, \quad v_1 b_1 + v_2 b_2 + v_3 b_3 = 0.$$

Como $\vec{v} \neq \vec{0}$, al menos una de sus componentes es distinta de 0. Supongamos que $v_1 \neq 0$. Entonces podemos despejar b_2 y b_3 en la segunda y tercera ecuación y obtenemos que $b_2 = \frac{v_2}{v_1} b_1$ and $b_3 = \frac{v_3}{v_1} b_1$. Si sustituimos esto en la última ecuación, encontramos que

$$0 = v_1 b_1 + \frac{v_2^2}{v_1} b_1 + \frac{v_3^2}{v_1} b_1 = \frac{b_1}{v_1^2} (v_1^2 + v_2^2 + v_3^2) = b_1 \frac{\|\vec{v}\|^2}{v_1}.$$

Como $\vec{v} \neq \vec{0}$, se sigue que $b_1 = 0$, pero entonces también $b_2 = \frac{v_2}{v_1} b_1 = 0$ and $b_3 = \frac{v_3}{v_1} b_1 = 0$. En resumen, $\vec{b} = \vec{0}$, como queríamos mostrar.

- (vi) La demostración es un cálculo largo pero directo:

$$\begin{aligned} \langle \vec{u}, \vec{v} \times \vec{w} \rangle &= \left\langle \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix} \right\rangle \\ &= u_1(v_2 w_3 - v_3 w_2) + u_2(v_3 w_1 - v_1 w_3) + u_3(v_1 w_2 - v_2 w_1) \\ &= u_1 v_2 w_3 - u_1 v_3 w_2 + u_2 v_3 w_1 - u_2 v_1 w_3 + u_3 v_1 w_2 - u_3 v_2 w_1 \\ &= u_2 v_3 w_1 - u_3 v_2 w_1 + u_3 v_1 w_2 - u_1 v_3 w_2 + u_1 v_2 w_3 - u_2 v_1 w_3 \\ &= (u_2 v_3 - u_3 v_2) w_1 + (u_3 v_1 - u_1 v_3) w_2 + (u_1 v_2 - u_2 v_1) w_3 \\ &= \langle \vec{u} \times \vec{v}, \vec{w} \rangle. \end{aligned}$$

- (vii) De (vi) y (v) se sigue que

$$\langle \vec{u}, \vec{u} \times \vec{v} \rangle = \langle \vec{u} \times \vec{u}, \vec{v} \rangle = \langle \vec{0}, \vec{v} \rangle = 0 \quad \text{and} \quad \langle \vec{v}, \vec{u} \times \vec{v} \rangle = -\langle \vec{v}, \vec{v} \times \vec{u} \rangle = 0. \quad \square$$

Observe que el producto cruz es distributivo, pero no es ni conmutativo ni asociativo.

Ejercicio. Demuestre las fórmulas en (i) – (iv) y la implicación “ $\implies$ ” en (v).

Una interpretación geométrica del número $\langle \vec{u}, \vec{v} \times \vec{w} \rangle$ de (vi) se dará en la Proposición 2.39.

Observación 2.36. La propiedad (VII) explica por qué el producto cruz solo tiene sentido en $\mathbb{R}^3$. Dados dos vectores no paralelos $\vec{v}$ y $\vec{w}$, su producto cruz es un vector que es ortogonal a ambos y cuya longitud es $\|\vec{v}\| \|\vec{w}\| \sin \varphi$ (véase el Teorema 2.37; $\varphi = \angle(\vec{v}, \vec{w})$), lo cual debería determinar el resultado de manera única salvo por un factor ± 1 . Este factor tiene que ver con la orientación relativa de $\vec{v}$ y $\vec{w}$ entre sí. Sin embargo, si $n \neq 3$, entonces ocurre una de las siguientes situaciones:

- Si estuviéramos en $\mathbb{R}^2$, el problema es que “no tenemos suficiente espacio”, porque entonces el único vector ortogonal a $\vec{v}$ y $\vec{w}$ al mismo tiempo sería el vector cero $\vec{0}$, y no tendría mucho sentido definir un producto cuyo resultado fuera siempre $\vec{0}$.
- Si estuviéramos en algún $\mathbb{R}^n$ con $n \geq 4$, el problema es que “tenemos demasiadas opciones”. Veremos más adelante en el Capítulo 7.3 que el complemento ortogonal del plano generado por $\vec{v}$ y $\vec{w}$ tiene dimensión $n-2$ y que todo vector en el complemento ortogonal es ortogonal tanto a $\vec{v}$ como a $\vec{w}$. Por ejemplo, si tomamos $\vec{v} = (1, 0, 0, 0)^t$ and $\vec{w} = (0, 1, 0, 0)^t$, entonces todo vector de la forma $\vec{a} = (0, 0, x, y)^t$ es perpendicular tanto a $\vec{v}$ como a $\vec{w}$, y es fácil encontrar infinitos vectores de esta forma que además tengan norma $\|\vec{v}\| \|\vec{w}\| \sin \varphi = 1$ ($\vec{a} = (0, 0, \sin \vartheta, \pm \cos \vartheta)^t$ para arbitrario $\vartheta \in \mathbb{R}$ funciona).

Recordemos que para el producto interno demostramos la fórmula $\langle \vec{v}, \vec{w} \rangle = \|\vec{v}\| \|\vec{w}\| \cos \varphi$ where φ is the angle between the two vectors, see Theorem 2.21. En el siguiente teorema demostraremos una relación similar para el producto cruz.

Teorema 2.37. Sean $\vec{v}, \vec{w}$ vectores en $\mathbb{R}^3$ y sea φ el ángulo entre ellos. Entonces

$$\|\vec{v} \times \vec{w}\| = \|\vec{v}\| \|\vec{w}\| \sin \varphi$$

Demostración. Un cálculo largo pero directo muestra que $\|\vec{v} \times \vec{w}\|^2 = \|\vec{v}\|^2 \|\vec{w}\|^2 - \langle \vec{v}, \vec{w} \rangle^2$. Ahora se sigue del Teorema 2.21 que

$$\begin{aligned} \|\vec{v} \times \vec{w}\|^2 &= \|\vec{v}\|^2 \|\vec{w}\|^2 - \langle \vec{v}, \vec{w} \rangle^2 = \|\vec{v}\|^2 \|\vec{w}\|^2 - \|\vec{v}\|^2 \|\vec{w}\|^2 (\cos \varphi)^2 \\ &= \|\vec{v}\|^2 \|\vec{w}\|^2 (1 - (\cos \varphi)^2) = \|\vec{v}\|^2 \|\vec{w}\|^2 (\sin \varphi)^2. \end{aligned}$$

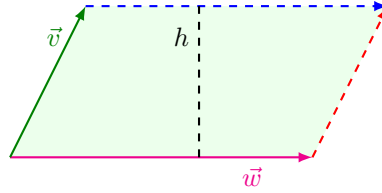
Si tomamos la raíz cuadrada en ambos lados, llegamos a la fórmula afirmada. (No tenemos que preocuparnos por tomar el valor absoluto de $\sin \varphi$ porque $\varphi \in [0, \pi]$, por lo que $\sin \varphi \geq 0$.) $\square$

Ejercicio. Muestre que $\|\vec{v} \times \vec{w}\|^2 = \|\vec{v}\|^2 \|\vec{w}\|^2 - \langle \vec{v}, \vec{w} \rangle^2$.

Aplicación: área de un paralelogramo y volumen de un paralelepípedo

Área de un paralelogramo

Sean $\vec{v}$ y $\vec{w}$ dos vectores en $\mathbb{R}^3$. Entonces definen un paralelogramo (si los vectores son paralelos o uno de ellos es igual a $\vec{0}$, se trata de un *paralelogramo degenerado*).

FIGURA 2.16: Paralelogramo generado por $\vec{v}$ y $\vec{w}$.

Proposición 2.38 (Área de un paralelogramo). *El área del paralelogramo generado por los vectores $\vec{v}$ y $\vec{w}$ es*

$$A = \|\vec{v} \times \vec{w}\|. \quad (2.9)$$

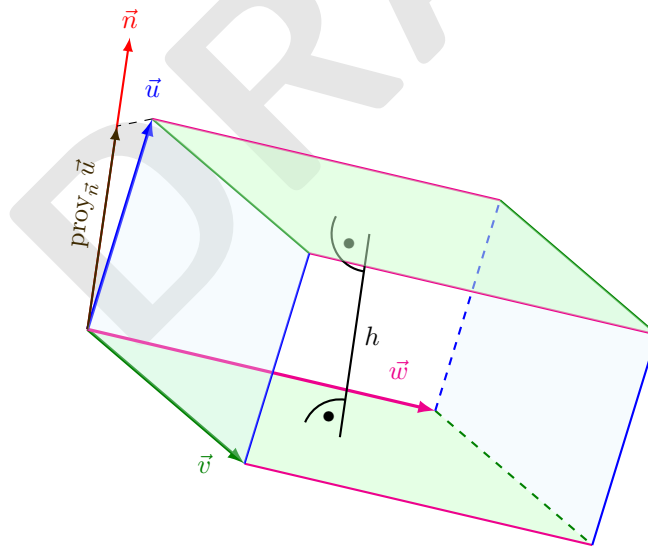
Demostración. El área de un paralelogramo es el producto de la longitud de su base por la altura. Podemos tomar $\vec{w}$ como base. Sea φ el ángulo entre $\vec{w}$ y $\vec{v}$. Entonces obtenemos que $h = \|\vec{v}\| \sin \varphi$ y por lo tanto, con ayuda del Teorema 2.37

$$A = \|\vec{w}\|h = \|\vec{w}\|\|\vec{v}\| \sin \varphi = \|\vec{v} \times \vec{w}\|. \quad \square$$

Observe que en el caso en que $\vec{v}$ y $\vec{w}$ son paralelos, esto da la respuesta correcta $A = 0$.

Volumen de un paralelepípedo

Cualesquiera tres vectores en $\mathbb{R}^3$ definen un paralelepípedo.

FIGURA 2.17: Paralelepípedo generado por $\vec{u}, \vec{v}, \vec{w}$.

Proposición 2.39 (Volumen de un paralelepípedo). *El volumen del paralelepípedo generado por los vectores $\vec{u}$, $\vec{v}$ y $\vec{w}$ es*

$$V = |\langle \vec{u}, \vec{v} \times \vec{w} \rangle|. \quad (2.10)$$

Demostración. El volumen de un paralelepípedo es el producto del área de su base por la altura. Tomemos como base el paralelogramo generado por $\vec{v}, \vec{w}$. Si $\vec{v}$ y $\vec{w}$ son paralelos o uno de ellos es igual a $\vec{0}$, entonces (2.10) es verdadera porque $V = 0$ y $\vec{v} \times \vec{w} = \vec{0}$ en este caso.

Supongamos ahora que no son paralelos. Por la Proposición 2.38 ya sabemos que su base tiene área $A = \|\vec{v} \times \vec{w}\|$. La altura es la longitud de la proyección ortogonal de $\vec{u}$ sobre el vector normal del plano generado por $\vec{v}$ y $\vec{w}$. Ya sabemos que $\vec{v} \times \vec{w}$ es tal vector normal. Por lo tanto obtenemos que

$$h = \|\text{proy}_{\vec{v} \times \vec{w}} \vec{u}\| = \left\| \frac{\langle \vec{u}, \vec{v} \times \vec{w} \rangle}{\|\vec{v} \times \vec{w}\|^2} \vec{v} \times \vec{w} \right\| = \frac{|\langle \vec{u}, \vec{v} \times \vec{w} \rangle|}{\|\vec{v} \times \vec{w}\|^2} \|\vec{v} \times \vec{w}\| = \frac{|\langle \vec{u}, \vec{v} \times \vec{w} \rangle|}{\|\vec{v} \times \vec{w}\|}.$$

Por lo tanto, el volumen del paralelepípedo es

$$V = Ah = \|\vec{v} \times \vec{w}\| \frac{|\langle \vec{u}, \vec{v} \times \vec{w} \rangle|}{\|\vec{v} \times \vec{w}\|} = |\langle \vec{u}, \vec{v} \times \vec{w} \rangle|. \quad \square$$

Observe que en el caso en que dos de los vectores $\vec{u}, \vec{v}$ y $\vec{w}$ son paralelos, la fórmula da la respuesta correcta: el volumen del paralelepípedo es 0.

Corolario 2.40. *Sean $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$. Entonces*

$$|\langle \vec{u}, \vec{v} \times \vec{w} \rangle| = |\langle \vec{v}, \vec{w} \times \vec{u} \rangle| = |\langle \vec{w}, \vec{u} \times \vec{v} \rangle|.$$

Demostración. La fórmula se cumple porque cada una de las expresiones describe el volumen del paralelepípedo generado por los tres vectores dados, ya que podemos tomar cualquiera de los lados del paralelogramo como base. $\square$

Debería haber entendido

- las interpretaciones geométricas del producto cruz,
- por qué existe solo en $\mathbb{R}^3$
- etc.

Ahora debería ser capaz de

- calcular el producto cruz,
- usarlo para decir algo sobre el ángulo entre dos vectores en $\mathbb{R}^3$,
- usarlo para calcular el área de un paralelogramo y el volumen de un paralelepípedo,
- etc.

Ejercicios.

1. (a) Calcule el área del paralelogramo cuyos vértices adyacentes son $A(1, 2, 3)$, $B(2, 3, 4)$, $C(-1, 2, -5)$ y calcule el cuarto vértice.

- (b) Calcule el área del triángulo con los vértices $A(1, 2, 3)$, $B(2, 3, 4)$, $C(-1, 2, -5)$.
- (c) Encuentre un punto P tal que el área del triángulo con vértices B, C, P sea igual a 13. ¿Cuántos tales puntos P hay? Describalos geoméricamente.
2. Calcule el volumen del paralelepípedo determinado por los vectores $\vec{u} = \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} -1 \\ 4 \\ 3 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} 1 \\ -2 \\ 7 \end{pmatrix}$.
3. Use el producto cruz para encontrar el seno del ángulo formado por los vectores $\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$ y $\begin{pmatrix} -3 \\ -2 \\ 4 \end{pmatrix}$.
4. Encuentre todos los vectores $\vec{a} \in \mathbb{R}^3$ tales que $\vec{a} \perp \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ y $\vec{a} \perp \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$. ¿Cuántos de ellos tienen norma 1? ¿Cuáles tales $\vec{a}$ satisfacen que $\vec{a} \times \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 8 \\ -2 \\ -5 \end{pmatrix}$?

2.6. Rectas y planos en $\mathbb{R}^3$

En esta sección discutimos rectas y planos y cómo describirlos mediante fórmulas. En la siguiente sección calcularemos, por ejemplo, intersecciones entre ellos. Trabajaremos principalmente en $\mathbb{R}^3$ y solo daremos algunas indicaciones sobre cómo los conceptos discutidos aquí se generalizan a $\mathbb{R}^n$ con $n \neq 3$. El caso especial $n = 2$ debería ser claro.

La definición formal de rectas y planos se dará en la Definición 5.57, porque esto requiere el concepto de independencia lineal. (Para los curiosos: una *recta* es un subespacio (afín) unidimensional de un espacio vectorial; un *plano* es un subespacio (afín) bidimensional de un espacio vectorial; un *hiperplano* es un subespacio (afín) de dimensión $(n - 1)$ de un espacio vectorial de dimensión n). En esta sección apelamos a nuestro conocimiento e intuición de la geometría elemental.

Rectas

Intuitivamente, es claro qué debería ser una recta en $\mathbb{R}^3$. Para describir completamente una recta en $\mathbb{R}^3$, no es necesario conocer todos sus puntos. Es suficiente conocer o bien

- (a) dos puntos distintos P, Q sobre la recta

or

- (b) un punto P sobre la recta y la dirección de la recta.

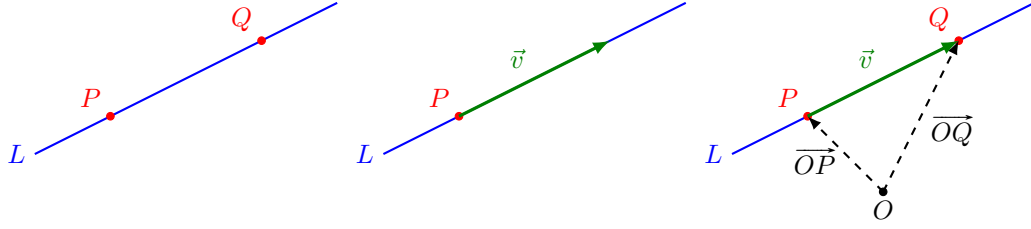


FIGURA 2.18: Recta L dada por: dos puntos P, Q en L ; o por un punto P en L y la dirección de L .

Claramente, ambas descripciones son equivalentes porque: Si tenemos dos puntos distintos P, Q sobre la recta L , entonces su dirección está dada por el vector $\overrightarrow{PQ}$. Si, por otro lado, se nos da un punto P en L y un vector $\vec{v}$ que es paralelo a L , entonces obtenemos fácilmente otro punto Q en L mediante $\overrightarrow{OQ} = \overrightarrow{OP} + \vec{v}$.

Ahora queremos dar fórmulas para la recta.

Ecuación vectorial de una recta

Dados dos puntos $P(p_1, p_2, p_3)$ and $Q(q_1, q_2, q_3)$ con $P \neq Q$, hay exactamente una recta L que pasa por ambos puntos. En fórmulas, esta recta se describe como

$$L = \left\{ \overrightarrow{OP} + t\overrightarrow{PQ} : t \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} p_1 + (q_1 - p_1)t \\ p_2 + (q_2 - p_2)t \\ p_3 + (q_3 - p_3)t \end{pmatrix} : t \in \mathbb{R} \right\}. \quad (2.11)$$

Si se nos da un punto $P(p_1, p_2, p_3)$ en L y un vector $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \neq \vec{0}$ paralelo a L , entonces

$$L = \left\{ \overrightarrow{OP} + t\vec{v} : t \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} p_1 + v_1t \\ p_2 + v_2t \\ p_3 + v_3t \end{pmatrix} : t \in \mathbb{R} \right\}. \quad (2.12)$$

Las fórmulas son fáciles de entender. Dicen lo siguiente: para trazar la recta, primero nos movemos a un punto arbitrario de la recta (este es el término $\overrightarrow{OP}$) y luego nos movemos una cantidad t a lo largo de la recta. Con este procedimiento podemos alcanzar todo punto de la recta y, por otro lado, si hacemos esto, entonces tenemos garantizado terminar sobre la recta.

Las fórmulas (2.11) y (2.12) se llaman *ecuación vectorial* de la recta L . Observe que son iguales si ponemos $v_1 = q_1 - p_1$, $v_2 = q_2 - p_2$, $v_3 = q_3 - p_3$. Usaremos principalmente la notación con las v , ya que es más corta. El vector $\vec{v}$ se llama *vector director* de la recta L .

Question 2.3

¿Es cierto que E pasa por el origen si y solo si $\overrightarrow{OP} = \vec{0}$?

Observación 2.41. Es importante observar que una recta dada tiene muchas parametrizaciones distintas.

- La ecuación vectorial que escribimos depende de los puntos que escojamos en L . Claramente, tenemos infinitas posibilidades para hacerlo.
- Toda recta dada L tiene muchos vectores directores. En efecto, si $\vec{v}$ es un vector director de L , entonces $c\vec{v}$ también lo es para todo $c \in \mathbb{R} \setminus \{0\}$. Sin embargo, todos los posibles vectores directores son paralelos.

Ejercicio. Verifique que las siguientes fórmulas describen todas la misma recta:

$$\begin{aligned} \text{(i)} \quad L_1 &= \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix} : t \in \mathbb{R} \right\}, & \text{(ii)} \quad L_2 &= \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} 12 \\ 10 \\ 8 \end{pmatrix} : t \in \mathbb{R} \right\}, \\ \text{(II)} \quad L_3 &= \left\{ \begin{pmatrix} 13 \\ 12 \\ 11 \end{pmatrix} + t \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix} : t \in \mathbb{R} \right\}. \end{aligned}$$

Question 2.4

- ¿Cómo se puede ver fácilmente si dos rectas dadas son paralelas o perpendiculares entre sí?
- ¿Cómo definiría el ángulo entre dos rectas? ¿Tienen que intersectarse para que se pueda definir un ángulo entre ellas?

Ecuación paramétrica de una recta

De la fórmula (2.12) es claro que un punto (x, y, z) pertenece a L si y solo si existe $t \in \mathbb{R}$ tal que

$$x = p_1 + tv_1, \quad y = p_2 + tv_2, \quad z = p_3 + tv_3. \quad (2.13)$$

Si hubiéramos empezado con (2.11), entonces habríamos obtenido

$$x = p_1 + t(q_1 - p_1), \quad y = p_2 + t(q_2 - p_2), \quad z = p_3 + t(q_3 - p_3). \quad (2.14)$$

El sistema de ecuaciones (2.13) o (2.14) se llama *ecuaciones paramétricas* de L . Aquí, t es el parámetro.

Ecuación simétrica de una recta

Observe que para $(x, y, z) \in L$, las tres ecuaciones en (2.13) deben cumplirse para el mismo t . Si suponemos que $v_1, v_2, v_3 \neq 0$, entonces podemos despejar t y obtenemos que

$$\frac{x - p_1}{v_1} = \frac{y - p_2}{v_2} = \frac{z - p_3}{v_3}. \quad (2.15)$$

Si usamos (2.14), entonces obtenemos

$$\frac{x - p_1}{q_1 - p_1} = \frac{y - p_2}{q_2 - p_2} = \frac{z - p_3}{q_3 - p_3}. \quad (2.16)$$

El sistema de ecuaciones (2.15) o (2.16) se llama *ecuación simétrica* de L .

Si, por ejemplo, $v_1 = 0$ y $v_2, v_3 \neq 0$, entonces la recta es paralela al plano yz y su ecuación simétrica es

$$x = p_1, \quad \frac{y - p_2}{v_2} = \frac{z - p_3}{v_3}.$$

Si $v_1 = v_2 = 0$ y $v_3 \neq 0$, entonces la recta es paralela al eje z y su ecuación simétrica es

$$x = p_1, \quad y = p_2, \quad z \in \mathbb{R}.$$

Representaciones de rectas en $\mathbb{R}^n$.

En $\mathbb{R}^n$, la forma vectorial de una recta es

$$L = \left\{ \overrightarrow{OP} + t\vec{v} : t \in \mathbb{R} \right\}$$

para $P \in L$ fijo y un vector director $\vec{v}$. Su forma paramétrica es

$$x_1 = p_1 + tv_1, \quad x_2 = p_2 + tv_2, \quad \dots, \quad x_n = p_n + tv_n, \quad t \in \mathbb{R},$$

y, suponiendo que todos los v_j son distintos de 0, su forma simétrica es

$$\frac{x_1 - p_1}{v_1} = \frac{x_2 - p_2}{v_2} = \dots = \frac{x_n - p_n}{v_n}.$$

Question 2.5. Forma normal de una recta.

En $\mathbb{R}^2$, también existe la *forma normal* de una recta:

$$L : ax + by = d \tag{2.17}$$

donde a, b y d son números fijos. Esto significa que L consiste en todos los puntos (x, y) cuyas coordenadas satisfacen la ecuación $ax + by = d$.

- (I) Dada una recta en la forma (2.17), encuentre una representación vectorial.
- (II) Dada una recta en representación vectorial, encuentre una forma normal (es decir, escríbala como (2.17)).
- (III) ¿Cuál es la interpretación geométrica de a, b ? (Indicación: dibuje la recta L y el vector $\begin{pmatrix} a \\ b \end{pmatrix}$.)
- (IV) ¿Puede esta forma normal extenderse/generalizarse a rectas en $\mathbb{R}^3$? Si es posible, ¿cómo se puede hacer? Si no es posible, explique por qué no.

Planos

Para conocer completamente un plano E en $\mathbb{R}^3$, es suficiente conocer

- (a) tres puntos P, Q, R en el plano que no yacen sobre una recta común,

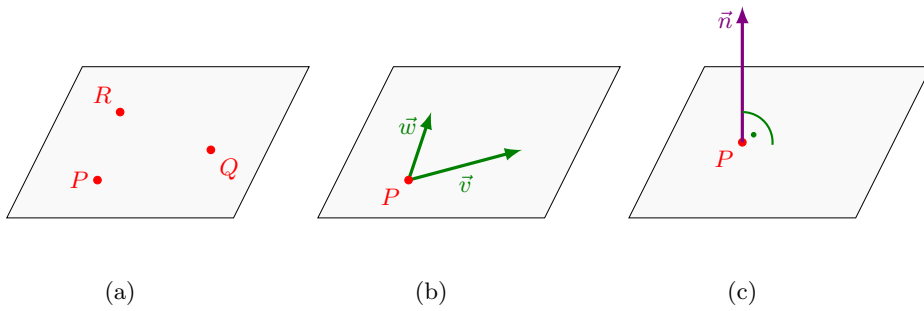


FIGURA 2.19: Plano E dado por: (a) tres puntos P, Q, R en E , (b) un punto P en E y dos vectores $\vec{v}, \vec{w}$ paralelos a E , (c) un punto P en E y un vector $\vec{n}$ perpendicular a E .

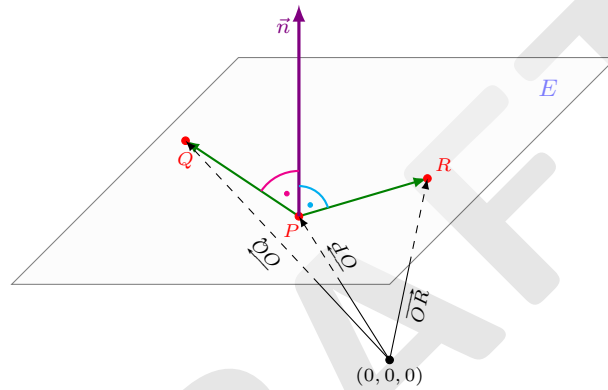


FIGURA 2.20: Plano E dado con tres puntos P, Q, R en E , dos vectores $\overrightarrow{PQ}, \overrightarrow{PR}$ paralelos a E , y un vector $\vec{n}$ perpendicular a E . Observe que $\vec{n} \parallel \overrightarrow{PQ} \times \overrightarrow{PR}$.

o

(b) un punto P en el plano y dos vectores no paralelos $\vec{v}, \vec{w}$ que son ambos paralelos al plano,

o

(c) un punto P en el plano y un vector $\vec{n}$ que es perpendicular al plano,

Primero veamos cómo podemos pasar de una descripción a otra. Claramente, las descripciones (a) and (b) son equivalentes porque, dados tres puntos P, Q, R en E que no yacen sobre una recta, podemos formar los vectores $\overrightarrow{PQ}$ y $\overrightarrow{PR}$. Estos vectores son entonces paralelos al plano E , pero no son paralelos entre sí. (Por supuesto, también podríamos haber tomado $\overrightarrow{QR}$ y $\overrightarrow{QP}$ o $\overrightarrow{RP}$ y $\overrightarrow{RQ}$.) Si, por otro lado, tenemos un punto P en E y dos vectores $\vec{v}$ y $\vec{w}$, paralelos a E y con $\vec{v} \nparallel \vec{w}$, entonces podemos obtener fácilmente otros dos puntos en E , por ejemplo mediante $\overrightarrow{OQ} = \overrightarrow{OP} + \vec{v}$ and $\overrightarrow{OR} = \overrightarrow{OP} + \vec{w}$. Entonces los tres puntos P, Q, R yacen en E y no yacen sobre una recta.

Ecuación vectorial de un plano

En fórmulas, ahora podemos describir nuestro plano E como

$$E = \left\{ (x, y, z) : \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \overrightarrow{OP} + s\vec{v} + t\vec{w} \text{ for some } s, t \in \mathbb{R} \right\}.$$

Como en el caso de la ecuación vectorial de una recta, la fórmula es fácil de entender. Primero nos movemos a un punto arbitrario del plano (este es el término $\overrightarrow{OP}$) y luego nos movemos paralelamente al plano como queramos (este es el término $s\vec{v} + t\vec{w}$). Con este procedimiento podemos alcanzar cualquier punto del plano y, por otro lado, si hacemos esto, entonces tenemos garantizado terminar en el plano.

Question 2.6

¿Es cierto que E pasa por el origen si y solo si $\overrightarrow{OP} = \vec{0}$?

Forma normal de un plano

En $\mathbb{R}^3$, podemos usar el vector normal del plano para describirlo. Supongamos que se nos da un punto P en E y un vector $\vec{n}$ perpendicular al plano. Esto significa que todo vector que es paralelo al plano E debe ser perpendicular a $\vec{n}$. Si tomamos un punto arbitrario $Q(x, y, z) \in \mathbb{R}^3$, entonces $Q \in E$ si y solo si $\overrightarrow{PQ}$ es paralelo a E , lo que significa que $\overrightarrow{PQ}$ es ortogonal a $\vec{n}$. Recordemos que dos vectores son perpendiculares si y solo si su producto interno es 0, así que $Q \in E$ si y solo si

$$\begin{aligned} 0 &= \langle \vec{n}, \overrightarrow{PQ} \rangle = \left\langle \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}, \begin{pmatrix} x - p_1 \\ y - p_2 \\ z - p_3 \end{pmatrix} \right\rangle = n_1(x - p_1) + n_2(y - p_2) + n_3(z - p_3) \\ &= n_1x + n_2y + n_3z - (n_1p_1 + n_2p_2 + n_3p_3) \end{aligned}$$

Si ponemos $d = n_1p_1 + n_2p_2 + n_3p_3$, entonces se sigue que un punto $Q(x, y, z)$ pertenece a E si y solo si sus coordenadas satisfacen

$$n_1x + n_2y + n_3z = d. \quad (2.18)$$

La ecuación (2.18) se llama *forma normal* del plano E y $\vec{n}$ se llama un *vector normal* de E .

Notación 2.42. Para definir E , escribimos $E : n_1x + n_2y + n_3z = d$. Como conjunto, denotamos E por $E = \{(x, y, z) : n_1x + n_2y + n_3z = d\}$.

Ejercicio. Muestre que E pasa por el origen si y solo si $d = 0$.

Observación 2.43. Como antes, observe que la ecuación normal de un plano no es única. Por ejemplo,

$$x + 2y + 3z = 5 \quad \text{and} \quad 2x + 4y + 6z = 10$$

describen el mismo plano. La razón es que “el” vector normal de un plano no es único. Si $\vec{n}$ es un vector normal del plano E , entonces todo $c\vec{n}$ con $c \in \mathbb{R} \setminus \{0\}$ también es un vector normal al plano.

Definición 2.44. El *ángulo entre dos planos* es el ángulo entre sus vectores normales.

Observe que esta definición es consistente con el hecho de que dos planos son paralelos si y solo si sus vectores normales son paralelos.

Observación 2.45. ■ Suponga que se da un plano como en (b) (es decir, conocemos un punto P en E y dos vectores $\vec{v}$ y $\vec{w}$ paralelos a E , pero con $\vec{v} \nparallel \vec{w}$). Para encontrar una descripción como en (c) (es decir, un punto en E y un vector normal), solo tenemos que encontrar un vector $\vec{n}$ que sea perpendicular tanto a $\vec{v}$ como a $\vec{w}$. La Proposición 2.35(vii) nos dice cómo hacerlo: solo necesitamos calcular $\vec{v} \times \vec{w}$. Otra manera de encontrar un $\vec{n}$ adecuado es encontrar una solución del sistema lineal 2×3 dado por $\{\langle \vec{v}, \vec{n} \rangle = 0, \langle \vec{w}, \vec{n} \rangle = 0\}$.

- Suponga que se da un plano como en (c) (es decir, conocemos un punto P en E y un vector normal). Para encontrar vectores $\vec{v}$ y $\vec{w}$ como en (b), podemos proceder de muchas maneras:
 - Encuentre dos soluciones de $\langle \vec{x}, \vec{n} \rangle = 0$ que no sean paralelas.
 - Encuentre dos puntos Q, R en el plano tales que $\overrightarrow{PQ} \nparallel \overrightarrow{PR}$. Eln we can take $\vec{v} = \overrightarrow{PQ}$ and $\vec{w} = \overrightarrow{PR}$.
 - Encuentre una solución $\vec{v} \neq \vec{0}$ de $\langle \vec{n}, \vec{v} \rangle = 0$, que usualmente es fácil de adivinar, y luego calcule $\vec{w} = \vec{v} \times \vec{n}$. El vector $\vec{w}$ es perpendicular a $\vec{n}$ y consecuentemente es paralelo al plano. También es perpendicular a $\vec{v}$ y por lo tanto no es paralelo a $\vec{v}$. En total, este vector $\vec{w}$ hace lo que necesitamos.

Representaciones de planos en $\mathbb{R}^n$.

En $\mathbb{R}^n$, la forma vectorial de un plano es

$$E = \left\{ \overrightarrow{OP} + t\vec{v} + s\vec{w} : t \in \mathbb{R} \right\}$$

para $P \in E$ fijo y dos vectores $\vec{v}, \vec{w}$ paralelos al plano, pero no paralelos entre sí.

Observe que no hay forma normal de un plano en $\mathbb{R}^n$ para $n \geq 4$. La razón es que para $n \geq 4$, hay más que una sola dirección perpendicular al plano dado. Por lo tanto, una forma normal del plano E debe consistir de más que una ecuación (más precisamente debe consistir de $n - 2$ ecuaciones de la forma for $n_1x_1 + \dots + n_nx_n = d$).

Debería haber entendido

- el concepto de rectas y planos en $\mathbb{R}^3$,
- cómo pueden describirse mediante fórmulas,
- etc.

Ahora debería ser capaz de

- pasar fácilmente entre las distintas descripciones de rectas y planos,
- etc.

Ejercicios.

1. Mostrar que las siguientes ecuaciones describen la misma recta:

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} : t \in \mathbb{R} \right\}, \quad \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} 8 \\ 10 \\ 12 \end{pmatrix} : t \in \mathbb{R} \right\}, \quad \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} -4 \\ -5 \\ -6 \end{pmatrix} : t \in \mathbb{R} \right\},$$

$$\left\{ \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix} + t \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} : t \in \mathbb{R} \right\}, \quad \frac{x-1}{4} = \frac{y-2}{5} = \frac{z-3}{6}, \quad \frac{x+3}{4} = \frac{y+3}{5} = \frac{z+3}{6}.$$

2. Dadas líneas L_1 y L_2 y el punto P ,

- (I) determine si L_1 y L_2 son paralelas,
- (II) determine si L_1 y L_2 tienen un punto de intersección,
- (III) determine si P pertenece a L_1 y/o a L_2 ,
- (IV) encuentre una recta paralela a L_2 que pase por P .

(a) $L_1 : \vec{r}(t) = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} + t \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}, \quad L_2 : \frac{x-3}{2} = \frac{y-2}{3} = \frac{z-1}{4}, \quad P(5, 2, 11).$

(b) $L_1 : \vec{r}(t) = \begin{pmatrix} 2 \\ 1 \\ -7 \end{pmatrix} + t \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad L_2 : x = t + 1, y = 3t - 4, z = -t + 2, \quad P(5, 7, 2).$

3. En $\mathbb{R}^3$ considere el plano E dado por $E : 3x - 2y + 4z = 16$.

- (a) Encuentre por lo menos tres puntos que pertenecen a E .
- (b) Encuentre un punto en E y dos vectores $\vec{v}$ y $\vec{w}$ en E que no son paralelos entre si.
- (c) Encuentre un punto en E y un vector $\vec{n}$ que es ortogonal a E .
- (d) Encuentre un punto en E y dos vectores $\vec{a}$ y $\vec{b}$ en E con $\vec{a} \perp \vec{b}$.

4. Para los puntos $P(1, 1, 1)$, $Q(1, 0, -1)$ y los siguientes planos E ,

- (I) encuentre la ecuación del plano.
- (II) determine si P pertenece al plano.
- (III) encuentre una recta que esté ortogonal a E y que contenga al punto Q .

- (a) E es el plano que contiene al punto $A(1, 0, 1)$ y es paralelo a los vectores

$$\vec{v} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \text{ y } \vec{w} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

- (b) E es el plano que contiene los puntos $A(1, 0, 1)$, $B(2, 3, 4)$, $C(3, 2, 4)$.

- (c) E es el plano que contiene el punto $A(1, 0, 1)$ y es ortogonal al vector $\vec{n} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$.

2.7. Intersecciones de rectas y planos en $\mathbb{R}^3$

Intersección de rectas

Dadas dos rectas G y L en $\mathbb{R}^3$, hay tres posibilidades:

- (a) Las rectas se intersecan en exactamente un punto. En este caso, no pueden ser paralelas.
- (b) Las rectas se intersecan en infinitos puntos. En este caso, las rectas tienen que ser iguales. En particular, tienen que ser paralelas.
- (c) Las rectas no se intersecan. Observe que, en contraste con el caso en $\mathbb{R}^2$, las rectas no tienen que ser paralelas para que esto ocurra. Por ejemplo, la recta $L : x = y = 1$ es una recta paralela al eje z que pasa por $(1, 1, 0)$, y $G : x = z = 0$ es una recta paralela al eje y que pasa por $(0, 0, 0)$. Las rectas no se intersecan y no son paralelas.

Ejemplo 2.46. Consideramos cuatro rectas $L_j = \{\vec{p}_j + t\vec{v}_j : t \in \mathbb{R}\}$ con

$$\begin{aligned} \text{(i)} \quad \vec{v}_1 &= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \vec{p}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, & \text{(ii)} \quad \vec{v}_2 &= \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}, \quad \vec{p}_2 = \begin{pmatrix} 2 \\ 4 \\ 7 \end{pmatrix}, \\ \text{(iii)} \quad \vec{v}_3 &= \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad \vec{p}_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, & \text{(iv)} \quad \vec{v}_4 &= \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad \vec{p}_4 = \begin{pmatrix} 3 \\ 0 \\ 5 \end{pmatrix}. \end{aligned}$$

Calcularemos sus intersecciones mutuas.

$$L_1 \cap L_2 = L_1$$

Demostración. Un punto $Q(x, y, z)$ pertenece a $L_1 \cap L_2$ si y solo si pertenece tanto a L_1 como a L_2 . Esto significa que debe existir un $s \in \mathbb{R}$ tal que $\overrightarrow{OQ} = \vec{p}_1 + s\vec{v}_1$ y debe existir un $t \in \mathbb{R}$ tal que $\overrightarrow{OQ} = \vec{p}_2 + t\vec{v}_2$. Observe que s y t son parámetros *diferentes*. Así, buscamos s y t tales que

$$\vec{p}_1 + s\vec{v}_1 = \vec{p}_2 + t\vec{v}_2, \quad \text{that is} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + s \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 7 \end{pmatrix} + t \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}. \quad (2.19)$$

Una vez que hayamos resuelto (2.19) para s y t , los insertamos en las ecuaciones de L_1 y L_2 , respectivamente, para obtener Q . Observe que (2.19) en realidad es un sistema de tres ecuaciones: una ecuación por cada componente de la ecuación vectorial. Escribiéndolo explícitamente y resolviendo cada ecuación para s , obtenemos

$$\begin{aligned} 0 + s &= 2 + 2t, & s &= 2 + 2t, \\ 0 + 2s &= 4 + 4t, & s &= 2 + 2t, \\ 1 + 3s &= 7 + 6t, & s &= 2 + 2t. \end{aligned} \quad \Longleftrightarrow$$

Esto significa que existen infinitas soluciones de (2.19). Dado cualquier punto R en L_1 , existe un $s \in \mathbb{R}$ correspondiente tal que $\overrightarrow{OR} = \vec{p}_1 + s\vec{v}_1$. Ahora, si escogemos $t = (s-2)/2$, entonces se cumple $\overrightarrow{OR} = \vec{p}_2 + t\vec{v}_2$, por lo tanto $R \in L_2$ también. Si, por otro lado, tenemos un punto $R' \in L_2$, entonces existe un $t \in \mathbb{R}$ correspondiente tal que $\overrightarrow{OR'} = \vec{p}_2 + t\vec{v}_2$. Ahora, si escogemos $s = 2 + 2t$, entonces se cumple $\overrightarrow{OR'} = \vec{p}_1 + t\vec{v}_1$, por lo tanto $R' \in L_1$ también. En resumen, mostramos que $L_1 = L_2$. $\square$

Observación 2.47. También podríamos haber visto que los vectores directores de L_1 y L_2 son paralelos. De hecho, $\vec{v}_2 = 2\vec{v}_1$. Entonces basta mostrar que L_1 y L_2 tienen al menos un punto en común para concluir que las rectas son iguales.

$$L_1 \cap L_3 = \{(1, 2, 4)\}$$

Demostración. Como antes, necesitamos encontrar $s, t \in \mathbb{R}$ tales que

$$\vec{p}_1 + s\vec{v}_1 = \vec{p}_3 + t\vec{v}_3, \quad \text{that is} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + s \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}. \quad (2.20)$$

Escribimos esto como un sistema de ecuaciones y obtenemos

$$\begin{array}{lll} \textcircled{1} & 0 + s = -1 + t, & \textcircled{1} \quad s - t = -1, \\ \textcircled{2} & 0 + 2s = 0 + t, & \textcircled{2} \quad 2s - t = 0, \\ \textcircled{3} & 1 + 3s = 0 + 2t, & \textcircled{3} \quad 3s - 2t = -1. \end{array} \quad \Longleftrightarrow$$

De $\textcircled{1}$ se sigue que $s = t - 1$. Al insertar en $\textcircled{2}$ se obtiene $0 = 2(t - 1) - t = t - 2$, de donde $t = 2$. De $\textcircled{1}$ obtenemos entonces que $s = 2 - 1 = 1$. Observe que hasta ahora solo usamos las ecuaciones $\textcircled{1}$ and $\textcircled{2}$. Para ver si realmente encontramos una solución, debemos comprobar si es consistente con $\textcircled{3}$. Al insertar nuestros candidatos para s y t , encontramos que $3 \cdot 1 - 2 \cdot 2 = -1$, lo cual es consistente con $\textcircled{3}$.

Así, L_1 y L_3 se intersecan en exactamente un punto. Para encontrarlo, ponemos $s = 1$ en la ecuación de L_1 :

$$\overrightarrow{OQ} = \vec{p}_1 + 1 \cdot \vec{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix},$$

por lo tanto, el punto de intersección es $Q(1, 2, 4)$.

Para comprobar si este resultado es correcto, podemos poner $t = 2$ en la ecuación de L_3 . El resultado debe ser el mismo. El cálculo correspondiente es:

$$\overrightarrow{OQ} = \vec{p}_3 + 2 \cdot \vec{v}_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix},$$

lo cual confirma que el punto de intersección es $Q(1, 2, 4)$. $\square$

$$L_1 \cap L_4 = \emptyset$$

Demostración. Como antes, necesitamos encontrar $s, t \in \mathbb{R}$ tales que

$$\vec{p}_1 + s\vec{v}_1 = \vec{p}_4 + t\vec{v}_4, \quad \text{that is} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + s \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 5 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}. \quad (2.21)$$

Escribimos esto como un sistema de ecuaciones y obtenemos

$$\begin{array}{lll} \textcircled{1} & s = 3 + t & \textcircled{1} \quad s - t = 3 \\ \textcircled{2} & 2s = t & \textcircled{2} \quad 2s - t = 0 \\ \textcircled{3} & 1 + 3s = 5 + 2t & \textcircled{3} \quad 3s - 2t = 5 \end{array} \quad \Longleftrightarrow$$

De $\textcircled{1}$ se sigue que $s = t + 3$. Al insertar en $\textcircled{2}$ se obtiene $0 = 2(t+3) - t = t + 6$, de donde $t = -6$. De $\textcircled{1}$ obtenemos entonces que $s = -6 + 3 = -3$. Observe que hasta ahora solo usamos las ecuaciones $\textcircled{1}$ and $\textcircled{2}$. Para ver si realmente encontramos una solución, debemos comprobar si es consistente con $\textcircled{3}$. Al insertar nuestros candidatos para s y t , encontramos que $3 \cdot (-3) - 2 \cdot (-6) = 3$, lo cual es inconsistente con $\textcircled{3}$. Por lo tanto concluimos que no hay ningún par de números reales s, t que satisfaga simultáneamente las tres ecuaciones $\textcircled{1}$ – $\textcircled{3}$; por tanto, las dos rectas no se intersecan. $\square$

Ejercicio. Muestre que $L_3 \cap L_4 = \emptyset$.

Intersección de planos

Dados dos planos E_1 y E_2 en $\mathbb{R}^3$, hay dos posibilidades:

- (a) Los planos se intersecan. En este caso, necesariamente se intersecan en infinitos puntos. Su intersección es una recta (si E_1 y E_2 no son paralelos) o un plano (si $E_1 = E_2$).
- (b) Los planos no se intersecan. En este caso, los planos deben ser paralelos y distintos.

Ejemplo 2.48. Consideramos los siguientes cuatro planos:

$$E_1 : x + y + 2z = 3, \quad E_2 : 2x + 2y + 4z = -4, \quad E_3 : 2x + 2y + 4z = 6, \quad E_4 : x + y - 2z = 5.$$

Calcularemos sus intersecciones mutuas.

$$E_1 \cap E_2 = \emptyset$$

Demostración. El conjunto de todos los puntos $Q(x, y, z)$ que pertenecen tanto a E_1 como a E_2 es el conjunto de todos los x, y, z que satisfacen simultáneamente

$$\begin{array}{ll} \textcircled{1} & x + y + 2z = 3, \\ \textcircled{2} & 2x + 2y + 4z = -4. \end{array}$$

Claramente, si x, y, z satisface $\textcircled{1}$, entonces no puede satisfacer $\textcircled{2}$ (el lado derecho sería 6). Podemos ver esto de manera más formal si resolvemos $\textcircled{1}$, por ejemplo para x , y luego sustituimos en $\textcircled{2}$. De $\textcircled{1}$ obtenemos: $x = 3 - y - 2z$. Al sustituir en $\textcircled{2}$ se llega a

$$-4 = 2(3 - y - 2z) + 2y + 4z = 6,$$

lo cual es absurdo. $\square$

Este resultado era de esperarse, ya que los vectores normales de los planos son $\vec{n}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ and $\vec{n}_2 = \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix}$ respectivamente. Como son paralelos, los planos son paralelos y, por lo tanto, o bien son iguales o bien tienen intersección vacía. Ahora vemos que, por ejemplo, $(3, 0, 0) \in E_1$ pero $(3, 0, 0) \notin E_2$, así que los planos no pueden ser iguales. Por lo tanto, tienen intersección vacía.

$$\boxed{E_1 \cap E_3 = E_1}$$

Demostración. El conjunto de todos los puntos $Q(x, y, z)$ que pertenecen tanto a E_1 como a E_3 es el conjunto de todos los x, y, z que satisfacen simultáneamente

$$\begin{aligned} \textcircled{1} \quad & x + y + 2z = 3, \\ \textcircled{2} \quad & 2x + 2y + 4z = 6. \end{aligned}$$

Claramente, ambas ecuaciones son equivalentes: si x, y, z satisface $\textcircled{1}$, entonces también satisface $\textcircled{2}$ y viceversa. Por lo tanto, $E_1 = E_3$. $\square$

$$\boxed{E_1 \cap E_4 = \left\{ \begin{pmatrix} 4 \\ 0 \\ -\frac{1}{2} \end{pmatrix} + t \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} : t \in \mathbb{R} \right\}}.$$

Demostración. Primero observamos que los vectores normales $\vec{n}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ and $\vec{n}_4 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ no son paralelos, así que esperamos que la solución sea una recta en $\mathbb{R}^3$. El conjunto de todos los puntos $Q(x, y, z)$ que pertenecen tanto a E_1 como a E_4 es el conjunto de todos los x, y, z que satisfacen simultáneamente

$$\begin{aligned} \textcircled{1} \quad & x + y + 2z = 3, \\ \textcircled{2} \quad & x + y - 2z = 5. \end{aligned}$$

La ecuación $\textcircled{1}$ muestra que $x = 3 - y - 2z$. Al sustituir en $\textcircled{2}$ se obtiene $5 = 3 - y - 2z + y - 2z = 3 - 4z$, de donde $z = -\frac{1}{2}$. Poniendo esto en $\textcircled{1}$, encontramos que $x + y = 3 - 2z = 4$. Así, en resumen, la intersección consiste en todos los puntos (x, y, z) que satisfacen

$$z = -\frac{1}{2}, \quad x = 4 - y \quad \text{with } y \in \mathbb{R} \text{ arbitrary,}$$

en otras palabras,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 - y \\ y \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ -\frac{1}{2} \end{pmatrix} + \begin{pmatrix} -y \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ -\frac{1}{2} \end{pmatrix} + y \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{with } y \in \mathbb{R} \text{ arbitrary.} \quad \square$$

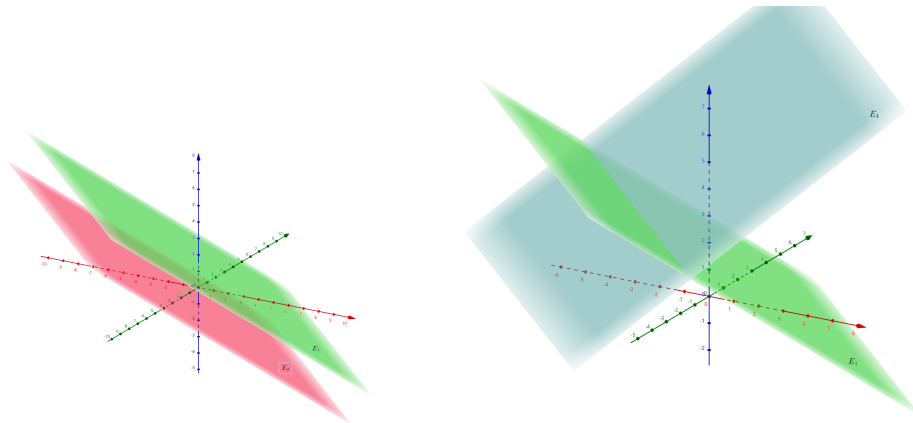


FIGURA 2.21: La figura de la izquierda muestra $E_1 \cap E_2 = \emptyset$; la figura de la derecha muestra $E_1 \cap E_4$, que es una recta.

Intersección de una recta con un plano

Finalmente queremos calcular la intersección de un plano E con una recta L . Hay tres posibilidades:

- (a) El plano y la recta se intersectan en exactamente un punto. Esto ocurre si y solo si L no es paralela a E , lo cual sucede si y solo si L no es perpendicular al vector normal de E .
- (b) El plano y la recta no se intersectan. En este caso, E y L deben ser paralelos, es decir, L debe ser perpendicular al vector normal de E .
- (c) El plano y la recta se intersectan en infinitos puntos. En este caso, L está contenida en E , es decir, E y L deben ser paralelos y deben compartir al menos un punto.

Como ejemplo calculamos $E_1 \cap L_2$. Como L_2 claramente no es paralela a E_1 , esperamos que su intersección consista en exactamente un punto.

$$E_1 \cap L_2 = \{(1/9, 2/9, 4/3)\}$$

Demostración. El conjunto de todos los puntos $Q(x, y, z)$ que pertenecen tanto a E_1 como a L_2 es el conjunto de todos los x, y, z que satisfacen simultáneamente

$$x + y + 2z = 3 \quad \text{and} \quad x = 2 + 2t, \quad y = 4 + 4t, \quad z = 7 + 6t \quad \text{for some } t \in \mathbb{R}.$$

Al sustituir la expresión con t de L_2 en la ecuación del plano E_1 , obtenemos la siguiente ecuación para t :

$$3 = (2 + 2t) + (4 + 4t) + 2(7 + 6t) = 20 + 18t \quad \implies \quad t = -17/18.$$

Al sustituir este t en la ecuación de L_2 se obtiene el punto de intersección $Q(1/9, 2/9, 4/3)$.

Para comprobar nuestro resultado, insertamos las coordenadas en la ecuación de E_1 y obtenemos $x + y + 2z = 1/9 + 2/9 + 2 \cdot 4/3 = 1/3 + 8/3 = 3$ lo cual muestra que $Q \in E_1$. $\square$

Calculemos dos ejemplos más.

Ejemplo 2.49. Consideramos el plano E y las rectas L, G dadas por

$$E : x + 2y - z = 2, \quad L : x + 1 = y = \frac{z-2}{3}. \quad \text{and} \quad G : \frac{x-3}{2} = y-1 = \frac{z-1}{4}.$$

Observe que el vector normal de E y los vectores directores de L y G son

$$\vec{n} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \quad \vec{v}_L = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \vec{v}_G = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}.$$

Observe que $\vec{n} \perp \vec{v}_L$ y $\vec{n} \perp \vec{v}_G$. Por lo tanto $L \parallel E$ y $G \parallel E$ (pero $L \not\parallel G$) y esperamos, para cada una de las rectas, que o bien no se interseque con E o bien esté contenida en E .

$E \cap L = \emptyset$ Las ecuaciones paramétricas de L son

$$x = -1 + t, \quad y = t, \quad z = 2 + 3t.$$

Si sustituimos esto en la fórmula de E , obtenemos

$$2 \stackrel{?}{=} t - 1 + 2t - (2 + 3t) = -7$$

lo cual es absurdo. Por lo tanto, E y L no tienen ningún punto en común. En otras palabras, no se intersecan.

$E \cap G = G$ Las ecuaciones paramétricas de G son

$$x = 3 + 2t, \quad y = 2 + t, \quad z = 5 + 4t.$$

Al sustituir esto en la fórmula de E obtenemos

$$2 \stackrel{?}{=} 3 + 2t + 2(2 + t) - (5 + 4t) = 2$$

lo cual es verdadero para todo $t \in \mathbb{R}$. Por lo tanto, todo punto de G pertenece también a E . En otras palabras, $G \subseteq E$ o $E \cap G = G$.

Observación. Recordemos que en el ejemplo anterior $\vec{n} \perp \vec{v}_L$ y $\vec{n} \perp \vec{v}_G$. Esto significa que L es paralelo a E , por lo tanto o bien cae completamente en E o bien no interseca a E . Lo mismo es cierto para G . Así, si tomamos un punto arbitrario en L y este punto pertenece también a E , entonces la intersección $E \cap L$ no es vacía, pero entonces debemos tener que $E \cap L = L$. Si ese punto no pertenece a E , entonces E y L no pueden intersectarse (porque de lo contrario $L \subseteq E$ y el punto también estaría en E). Por ejemplo, es fácil ver que el punto $P(-1, 0, 2)$ pertenece a L (basta tomar $t = 0$ en su ecuación paramétrica). Pongamos las coordenadas de P en la fórmula de E :

$$-1 + 2 \cdot 0 - 2 = -3 \neq 2.$$

Por lo tanto $P \notin E$ y, por lo tanto, $E \cap L = \emptyset$.

Por otro lado, si escogemos $Q(3, 2, 5)$, entonces vemos fácilmente que $Q \in G$. Al sustituir sus coordenadas en la fórmula de E , encontramos que

$$3 + 2 \cdot 2 - 5 = 2,$$

por lo tanto $Q \in E$ y, en consecuencia, $G \subseteq E$.

Intersección de varias rectas y planos

Si quisiéramos intersecar, por ejemplo, 5 planos en $\mathbb{R}^3$, entonces tendríamos que resolver un sistema de 5 ecuaciones con 3 incógnitas. O si quisiéramos intersecar 7 rectas en $\mathbb{R}^3$, tendríamos que resolver un sistema de 3 ecuaciones con 7 incógnitas. Si los resolvemos como lo hicimos aquí, el proceso podría volverse bastante engorroso. Por eso, el próximo capítulo está dedicado a encontrar una manera sistemática y eficiente de resolver un sistema de m ecuaciones lineales con n incógnitas.

Debería haber entendido

- qué pueden ser geoméricamente las intersecciones de rectas y planos y cómo dependen de su orientación relativa,
- la interpretación de un sistema lineal con tres incógnitas como la intersección de planos en $\mathbb{R}^3$,
- etc.

Ahora debería ser capaz de

- calcular la intersección de rectas y planos,
- etc.

Ejercicios.

1. Considere el plano $E : 2x - y + 3z = 9$ y la recta $L : x = 3t + 1, y = -2t + 3, z = 4t$.
 - (a) Encuentre $E \cap L$.
 - (b) Encuentre una recta G que no interseque ni al plano E ni a la recta L . Pruebe su afirmación. ¿Cuántas rectas con esta propiedad hay?
2. Sea L la recta que pasa por el punto $(1, 1, 1)$ y es paralela al vector $\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}$. Muestre que L no interseca al plano $E : x - 2y - z = 1$.
3. Dado el plano $E : 3x - 4y + 2z = 5$, encuentre
 - (a) un punto $P \in E$ cuya coordenada x es 2;
 - (b) un punto $Q \notin E$ cuya coordenada x es 2;
 - (c) una recta L paralela a E que pase por el punto $(3, 1, 5)$;
 - (d) una recta L perpendicular a E que pase por el punto $(3, 1, 5)$;
4. Dada la recta $L : x = 3t - 2, y = 2t + 5, z = t + 3$, encuentre o diga por qué no existe
 - (a) un punto $P \in L$ cuya coordenada x es 2;
 - (b) un punto $Q \notin L$ cuya coordenada x es 2;
 - (c) un plano E paralelo a L que pase por el punto $(3, 1, 5)$;
 - (d) un plano E perpendicular a L que pase por el punto $(3, 1, 5)$;

- (e) un plano E que contiene a la recta L y que pase por el punto $(2, -5, 1)$;
5. De dos planos E y F en $\mathbb{R}^3$ se sabe que no son paralelos y que los puntos $A(1, 2, 3)$ y $B(4, 0, -3)$ pertenecen a $E \cap F$. Se puede concluir qué es $E \cap F$? Dé dos ejemplos de planos E y F con la propiedad arriba.
6. Un caminante arranca en el tiempo $t = 0$ en el punto $(1, 2)$ con velocidad $\vec{v}_c = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Hay un ciclista en el punto $(12, -3)$ y un punto de refrigerios en el sitio $(16, 7)$. El caminante y el ciclista se mueven en líneas rectas ambos con velocidad constante.
- (a) Muestre que el caminante pasa por el punto de los refrigerios. ¿En cuál tiempo t pasa por este punto?
- (b) ¿En cuál dirección se debe dirigir el ciclista para pasar también por el punto de los refrigerios?
- (c) Supongamos que el ciclista arranca al mismo tiempo que el caminante. ¿Cómo debe el ciclista escoger su velocidad si quiere encontrarse con el caminante en el punto de los refrigerios? ¿Cómo la debe escoger si quiere pasar por el punto antes del caminante?
- (d) Ahora supongamos que el ciclista se mueve con la velocidad $\vec{w} = \begin{pmatrix} 3 \\ 7,5 \end{pmatrix}$. Muestre que el ciclista pasa por el punto de los refrigerios. ¿En qué momento t debe arrancar para encontrarse con el caminante?

2.8. Resumen

El espacio vectorial $\mathbb{R}^n$ está dado por

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \dots, x_n \in \mathbb{R} \right\}.$$

Para puntos $P(p_1, \dots, p_n)$, $Q(q_1, \dots, q_n)$, el vector cuyo punto inicial es P y cuyo punto final es Q es

$$\overrightarrow{PQ} = \begin{pmatrix} q_1 - p_1 \\ \vdots \\ q_n - p_n \end{pmatrix} \quad \text{y} \quad \overrightarrow{OQ} = \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix} \quad \text{donde } O \text{ denota el origen.}$$

En $\mathbb{R}^n$, la *suma* y el *producto por escalares* se definen por

$$\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \vec{v} + \vec{w} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix}, \quad \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad c\vec{v} = \begin{pmatrix} cv_1 \\ \vdots \\ cv_n \end{pmatrix}.$$

La *norma* de un vector es

$$\|\vec{v}\| = \sqrt{v_1^2 + \dots + v_n^2}.$$

Si $\vec{v} = \overrightarrow{PQ}$, entonces $\|\vec{v}\| = \|\overrightarrow{PQ}\| = \text{distancia entre } P \text{ y } Q$.

Para vectores $\vec{v}$ y $\vec{w} \in \mathbb{R}^n$, su *producto interno* es un número real definido por

$$\langle \vec{v}, \vec{w} \rangle = \left\langle \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \right\rangle = v_1 w_1 + \cdots + v_n w_n.$$

Fórmulas importantes que involucran el producto interno

- $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle,$
 $\langle \vec{v}, c\vec{w} \rangle = c\langle \vec{v}, \vec{w} \rangle,$
 $\langle \vec{v}, \vec{w} + \vec{u} \rangle = \langle \vec{v}, \vec{w} \rangle + \langle \vec{v}, \vec{u} \rangle,$
- $\langle \vec{v}, \vec{w} \rangle = \|\vec{v}\| \|\vec{w}\| \cos \varphi,$
- $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$ Desigualdad triangular
- $\vec{v} \perp \vec{w} \iff \langle \vec{v}, \vec{w} \rangle = 0,$
- $\langle \vec{v}, \vec{v} \rangle = \|\vec{v}\|^2.$

El *producto cruz* está definido **solamente** en $\mathbb{R}^3$. Es un vector definido por

$$\vec{v} \times \vec{w} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}.$$

Fórmulas importantes que involucran el producto cruz

- $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u},$
 $\vec{u} \times (\vec{v} + \vec{w}) = (\vec{u} \times \vec{v}) + (\vec{u} \times \vec{w}),$
 $(c\vec{u}) \times \vec{v} = c(\vec{u} \times \vec{v}),$
- $\langle \vec{u}, \vec{v} \times \vec{w} \rangle = \langle \vec{u} \times \vec{v}, \vec{w} \rangle.$
- $\vec{u} \parallel \vec{v} \iff \vec{u} \times \vec{v} = \vec{0}.$
- $\langle \vec{u}, \vec{u} \times \vec{v} \rangle = 0$ y $\langle \vec{v}, \vec{u} \times \vec{v} \rangle = 0$, en particular $\vec{v} \perp \vec{v} \times \vec{u}, \quad \vec{u} \perp \vec{v} \times \vec{u}$.
- $\|\vec{v} \times \vec{w}\| = \|\vec{v}\| \|\vec{w}\| \sin \varphi.$

Aplicaciones

- Área de un paralelogramo generado por $\vec{v}, \vec{w} \in \mathbb{R}^3$: $A = \|\vec{v} \times \vec{w}\|.$
- Volumen de un paralelepípedo generado por $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$: $V = |\langle \vec{u}, \vec{v} \times \vec{w} \rangle|.$

Representaciones de rectas

- **Ecuación vectorial** $L = \{\overrightarrow{OP} + t\vec{v} : t \in \mathbb{R}\}.$
 P es un punto sobre la recta, $\vec{v}$ se llama vector director de L .
- **Ecuación paramétrica** $x_1 = p_1 + tv_1, \dots, x_n = p_n + tv_n, t \in \mathbb{R}.$
 Entonces $P(p_1, \dots, p_n)$ is a point on L y $\vec{v} = (v_1, \dots, v_n)^t$ es un vector director de L .
- **Ecuación simétrica** $\frac{x_1 - p_1}{v_1} = \frac{x_2 - p_2}{v_2} = \dots = \frac{x_n - p_n}{v_n}.$
 Entonces $P(p_1, \dots, p_n)$ is a point on L y $\vec{v} = (v_1, \dots, v_n)^t$ es un vector director de L .
 Si uno o varios de los v_j son iguales a 0, entonces la fórmula anterior debe modificarse.

Representaciones de planos

- **Ecuación vectorial** $E = \{\overrightarrow{OP} + t\vec{v} + s\vec{w} : s, t \in \mathbb{R}\}.$
 P is a point on the line, $\vec{v}$ y $\vec{w}$ are vectors parallel to E with $\vec{v} \nparallel \vec{w}$.
- **Forma normal (¡solo en $\mathbb{R}^3$!)** $E : ax + by + cz = d.$
 El vector $\vec{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ formado por los coeficientes del lado izquierdo es perpendicular a E .
 Además, E pasa por el origen si y solo si $d = 0$.

¡Las parametrizaciones no son únicas! (Una misma recta (o plano) tiene muchas parametrizaciones distintas.)

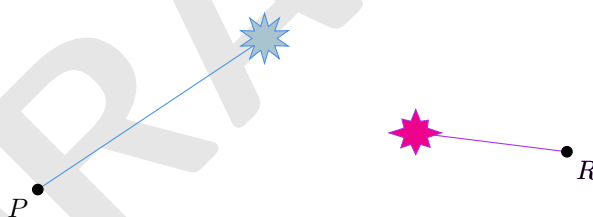
- El ángulo entre dos rectas es el ángulo entre sus vectores directores.
- Dos rectas son paralelas si y solo si sus vectores directores son paralelos.
 Dos rectas son perpendiculares si y solo si sus vectores directores son perpendiculares.
- El ángulo entre dos planos es el ángulo entre sus vectores normales.
- Dos planos son paralelos si y solo si sus vectores normales son paralelos.
 Dos planos son perpendiculares si y solo si sus vectores normales son perpendiculares.
- Una recta es paralela a un plano si y solo si su vector director es perpendicular al vector normal del plano.
 Una recta es perpendicular a un plano si y solo si su vector director es paralelo al vector normal del plano.

2.9. Ejercicios

- Sean $\vec{a} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$ y $\vec{b} = \begin{pmatrix} -1 \\ 4 \end{pmatrix}$. Encuentre vectores $\vec{u}, \vec{w}$ tal que cumplan todas las siguientes condiciones:
 - $\vec{a} = \vec{u} + \vec{w}$.
 - $\vec{u} \parallel \vec{b}$.
 - $\vec{w} \perp \vec{b}$.
- Sea $\vec{a} = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$. Encuentre $\vec{b} \in \mathbb{R}^2$ tal que $\vec{a} \perp \vec{b}$ y $\|\vec{a}\| = \|\vec{b}\|$. Repita el ejercicio cuando $\vec{a} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}$.
- Sea $P(-1, 2)$ un punto en $\mathbb{R}^2$ y, para $R > 0$, sea C_R el círculo centrado en P con radio R .
 - Sea L la recta que pasa por el origen y por $Q(2, 3)$.
Escriba L en forma vectorial y encuentre todos los $R > 0$ tal que
 - C_R no interseca a L ,
 - C_R toca a L en exactamente un punto,
 - C_R interseca a L en exactamente dos puntos.
 - Encuentre todos los puntos $S \in \mathbb{R}^2$ tal que el círculo centrado en S con radio 2 toca la recta L en exactamente un punto.
 - Repita los ejercicios anteriores, pero con la recta G que pasa por los puntos $(0, 1)$ y $(4, 0)$ (en particular, G no pasa por el origen).

Haga bosquejos para los ejercicios arriba. (Cf. Ejercicio 14.)
- Los señales que emite una emisora desde el punto $(0, 0)$ se dejan detectar hasta a una distancia de 100 metros. Si un carro se mueve a lo largo de la recta que pasa por los puntos $(0, 50)$ y $(105, -25)$, ¿podrá detectar los señales?
¿Qué cambia si la emisora está posicionada en el punto $(-20, -60)$? (Cf. Ejercicio 14.)
- Sea $\vec{a} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$ y $\vec{b} = \begin{pmatrix} 1 \\ 2 \\ -4 \end{pmatrix}$. Encuentre escalares x, y tal que $(x\vec{a} + y\vec{b}) \perp \vec{b}$ y $x\vec{a} + y\vec{b} \neq \vec{0}$.
- Sean $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^n$. Responda si son falsas o verdaderas las siguientes afirmaciones. En caso afirmativo demuéstrela y en caso negativo dé un contraejemplo.
 - Si $\langle \vec{a}, \vec{b} \rangle = \langle \vec{a}, \vec{c} \rangle$ y $\vec{a} \neq \vec{0}$, entonces $\vec{b} = \vec{c}$.
 - Si existe un vector $\vec{b}$ con $\langle \vec{a}, \vec{b} \rangle = 0$, entonces $\vec{a} = \vec{0}$.
 - Si $\langle \vec{a}, \vec{b} \rangle = 0$ para todo vector $\vec{b}$, entonces $\vec{a} = \vec{0}$.

- (d) Si $n = 3$. Si $\langle \vec{a}, \vec{b} \rangle = 0$ y $\vec{a} \times \vec{b} = 0$, entonces $\vec{a} = 0$ ó $\vec{b} = 0$ (Haga una interpretación geométrica antes de intentar desarrollar este ítem).
7. Sean $\vec{a} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ y $\vec{b} = \begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix}$. Encuentre por lo menos 3 vectores $\vec{c}$ distintos tales que el volumen del paralelepípedo generado por $\vec{a}, \vec{b}, \vec{c}$ sea 1. ¿Cuál es el lugar geométrico que describen todos los vectores $\vec{c}$ con esta propiedad?
8. En $\mathbb{R}^3$ demuestre que no existe un vector unitario cuyos ángulos directores son $\frac{\pi}{6}, \frac{\pi}{10}, \frac{\pi}{3}$.
9. Un cometa sale disparado del punto $P(1, 0, 3)$ y se mueve con velocidad $(3, 5, -2)^t$, al mismo tiempo un asteroide sale disparado del punto $R(3, 4, -1)$ y se mueve con vector velocidad $(2, 3, 0)^t$. Si suponemos que ambos asteroides se desplazan en línea recta, a velocidad constante y que además ningún objeto celeste perturba su trayectoria, responda las siguientes preguntas:
- 1) ¿Cuáles son las ecuaciones vectoriales que describen las trayectorias del asteroide y del cometa?
 - 2) ¿Los dos asteroides colisionan? ¿En qué tiempo lo hacen?
 - 3) Si el cometa deja en su trayectoria una estela de hielo y el asteroide deja en su recorrido una estela de polvo, ¿el hielo del cometa y el polvo del asteroide se mezclan en algún punto del espacio?.
- (b) Repita las preguntas anteriores considerando que el asteroide parte de $R(10, 5, 2)$ a velocidad $(3, 15, -7)^t$.



10. Considere una pared inclinada que está dada por la ecuación $E : 2x - 3y + 0,5z = 4$.
- (a) Demuestre que el punto $Q(4, 2, 4)$ pertenece a la pared.
 - (b) En el punto $P(2, 0, 1)$ hay un láser. ¿En cuál dirección debe apuntar para marcar el punto Q en la pared?
 - (c) ¿Cuál punto en la pared marcará el láser del literal anterior si apunta en la dirección $\begin{pmatrix} 2/3 \\ 1 \\ 3 \end{pmatrix}$?
11. Una empresa fabrica bultos de peso. Hay tres materiales con los que puede llenar los bultos: Material A tiene una densidad de 1 kg/l, material B tiene una densidad de 2 kg/l y material C tiene una densidad de 3 kg/l. Cada bulto debe pesar 20 kg.

- (a) Interprete la información dada en el ejercicio como un plano en $\mathbb{R}^3$ donde los ejes representan la cantidad de cada material.
- (b) Si en un bulto ya se encuentran 5l del material A y 3l del material B , ¿cuántos litros del material C se debe agregar para completar el bulto?
- (c) ¿Cómo se pueden armar los bultos si se quiere adicionalmente que el volumen de cada bulto es 13l? ¿Cuántas posibilidades de hacerlo hay? Interprete sus cálculos como intersección de dos planos.
- (d) En un bulto ya se encuentran 2l del material A , 5l del material B y 1l del material C . La empresa lo quiere completar con una mezcla de los tres materiales en la que 20 % son del material A , 50 % son del material B , 30 % son del material C . ¿Cuántos litros de esta mezcla hay que agregar para que el bulto pese 20 kg? Interprete sus cálculos como intersección de un plano con una recta.
12. En $\mathbb{R}^3$ sea E un plano y L una recta.
- (a) Muestre que E, L se cortan en exactamente un punto si el vector director de L y el vector normal de E no son perpendiculares.
- (b) Muestre que E, L no se cortan si el vector director de L es perpendicular al vector normal de E y E no contiene ningún punto de L .
- (c) Muestre que L está contenida en E si el vector director de L es perpendicular al vector normal de E y E contiene al menos un punto de L .
13. Encontrar la ecuación vectorial de las rectas que satisfacen:
- (a) Contiene a $(7, 1, 3)$ y $(-1, -2, -3)$.
- (b) Contiene a $(-2, 3, -2)$ y es paralela a $4\vec{e}_2$.
- (c) Contiene a $(-2, 3, 7)$ y es perpendicular a $2\vec{e}_1$.
- (d) Contiene a $(4, 1, -4)$ y es paralela a la recta dada por $\frac{x-2}{3} = \frac{y+1}{6} = \frac{z-5}{2}$.
14. (a) En $\mathbb{R}^2$ considere la recta $L : ax + by = c$ y un punto $P(x_1, y_1)$ exterior a L . Demuestre que la distancia d de P a L viene dada por la fórmula:
- $$d = \frac{|ax_1 + by_1 - c|}{\sqrt{a^2 + b^2}}.$$
- (Hint: Recuerde el significado geométrico del vector $\begin{pmatrix} a \\ b \end{pmatrix}$ en la recta L . Usar proyecciones.)
- (b) En $\mathbb{R}^3$ considere el plano $E : ax + by + cz = d$ y un punto $P(x_1, y_1, z_1)$ exterior a E . Demuestre que la distancia d de P a E viene dada por la fórmula:
- $$d = \frac{|ax_1 + by_1 + cz_1 - d|}{\sqrt{a^2 + b^2 + c^2}}.$$
- (Hint: Es el mismo razonamiento del inciso anterior.)
15. En $\mathbb{R}^3$ considere en forma vectorial la recta $L = \overrightarrow{OP} + t\vec{v}$.

- (a) Encuentre la distancia de la recta L al origen. (*Hint.* Encuentre $t \in \mathbb{R}$ con $\overrightarrow{OP} + t\vec{v} \perp \vec{v}$.)
 (b) Use el ejercicio anterior para encontrar la distancia entre la recta L y el origen donde L es:

$$\begin{pmatrix} 4 \\ 3 \\ 7 \end{pmatrix} + t \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad t \in \mathbb{R}.$$

16. Sea E el plano $3x + y + z = 1$ y $P(-6, 4, 4)$. El siguiente ejercicio pretende encontrar el punto más cercano a P dentro de E .

- (a) Verifique que $P \notin E$.
 (b) Encuentre la ecuación de la recta L que es paralela al vector normal de E y pasa por P .
 (c) Obtenga el punto de intersección entre E y L , llámelo Q .
 (d) Verifique que el valor de la distancia entre el punto obtenido y P da lo mismo que la distancia de P al plano E (Ejercicio 14, parte (b)).
 (e) Justifique el porqué Q es el punto más cercano a P que está dentro de E .

17. Sean $\vec{a}, \vec{b} \in \mathbb{R}^2$ con $\vec{a} \neq \vec{0}$.

- (a) Demuestre que $\|\text{proy}_{\vec{a}} \vec{b}\| \leq \|\vec{b}\|$.
 (b) ¿Qué deben cumplir $\vec{a}$ y $\vec{b}$ para que $\|\text{proy}_{\vec{a}} \vec{b}\| = \|\vec{b}\|$?
 (c) ¿Es cierto que $\|\text{proy}_{\vec{b}} \vec{a}\| \leq \|\vec{b}\|$?

18. Sea $L = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

- (a) Muestre que $(2, 3, 5)$ no pertenece a L .
 (b) Encuentre un plano que contenga a L y pase por $(2, 3, 5)$. ¿Cuántos planos que cumplen la condición anterior existen?

19. Sean

$$L_1 := \frac{x+3}{2} = y-4 = \frac{z+2}{7}$$

y L_2 dada por sus ecuaciones paramétricas:

$$\begin{aligned} x &= -3 + s \\ y &= 2 - 4s \\ z &= 1 + 6s \end{aligned}$$

- (a) ¿ L_1 y L_2 se intersecan?
 (b) Encuentre un plano F que sea perpendicular a L_1 y que pase por $(3, 2, 1)$.
 (c) Encuentre la ecuación normal de un plano E que sea paralelo tanto a L_1 como a L_2 y que pase por el origen. ¿Cuántos planos con esta propiedad hay?

- (d) ¿Existe algún plano que contenga simultáneamente a las rectas L_1, L_2 ? En caso de que la respuesta sea negativa, ¿cuales condiciones deberían cumplir dos rectas para que exista un plano que las contenga a ambas?
20. En $\mathbb{R}^3$ considere el plano E dado por $E : 3x - 2y + 4z = 16$.
- (a) Demuestre que los vectores $\vec{a} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}$ y $\vec{v} = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$ son paralelos al plano E .
- (b) Encuentre números $\lambda, \mu \in \mathbb{R}$ tal que $\lambda\vec{a} + \mu\vec{b} = \vec{v}$.
- (c) Demuestre que el vector $\vec{c} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ no es paralelo al plano E y encuentre vectores $c_{\parallel}$ y $c_{\perp}$ tal que $c_{\parallel}$ es paralelo a E , $c_{\perp}$ es ortogonal a E y $c = c_{\parallel} + c_{\perp}$.
21. Sea E un plano en $\mathbb{R}^2$ y sean $\vec{a}, \vec{b}$ vectores paralelos a E . Demuestre que para todo $\lambda, \mu \in \mathbb{R}$, el vector $\lambda\vec{a} + \mu\vec{b}$ es paralelo al plano.
22. Sea V un espacio vectorial. Demuestre lo siguiente:
- (a) El elemento neutro es único.
- (b) $0v = \mathbf{0}$ para todo $v \in V$.
- (c) $\lambda\mathbf{0} = \mathbf{0}$ para todo $\lambda \in \mathbb{R}$.
- (d) Dado $v \in V$, su inverso $\tilde{v}$ es único.
- (e) Dado $v \in V$, su inverso $\tilde{v}$ cumple $\tilde{v} = (-1)v$.
23. De todos los siguientes conjuntos decida si es un espacio vectorial con su suma y producto usual.
- (a) $V = \left\{ \begin{pmatrix} a \\ a \end{pmatrix} : a \in \mathbb{R} \right\}$,
- (b) $V = \left\{ \begin{pmatrix} a \\ a^2 \end{pmatrix} : a \in \mathbb{R} \right\}$,
- (c) V es el conjunto de todas las funciones continuas $\mathbb{R} \rightarrow \mathbb{R}$.
- (d) V es el conjunto de todas las funciones f continuas $\mathbb{R} \rightarrow \mathbb{R}$ con $f(4) = 0$.
- (e) V es el conjunto de todas las funciones f continuas $\mathbb{R} \rightarrow \mathbb{R}$ con $f(4) = 1$.

DRAFT

Capítulo 3

Sistemas lineales y matrices

Reescribiremos los sistemas lineales como ecuaciones matriciales para resolverlos de manera sistemática y eficiente. Interpretaremos las matrices como transformaciones lineales de $\mathbb{R}^n$ a $\mathbb{R}^m$, lo cual nos permite definir operaciones algebraicas con matrices; específicamente, definiremos la suma y la composición (multiplicación) de matrices, lo cual conduce naturalmente al concepto de la inversa de una matriz. Podemos interpretar una matriz como un sistema que toma una entrada (las variables $x_1, \dots, x_n$) y nos devuelve como salida $b_1, \dots, b_m$ mediante $A\vec{x} = \vec{b}$. A veces se nos da la entrada y queremos encontrar los b_j ; y a veces se nos da la salida $b_1, \dots, b_m$ y queremos encontrar la entrada $x_1, \dots, x_n$ que produce la salida deseada. Esta última pregunta suele ser más difícil que la primera. Veremos que existe una entrada única para cualquier salida dada si y solo si la matriz es invertible. Podemos refinar el concepto de invertibilidad de una matriz. Decimos que A tiene una inversa izquierda si para cualquier $\vec{b}$ la ecuación $A\vec{x} = \vec{b}$ tiene a lo sumo una solución, y decimos que tiene una inversa derecha si $A\vec{x} = \vec{b}$ tiene al menos una solución para cualquier $\vec{b}$.

Discutiremos en detalle la eliminación de Gauß y de Gauß-Jordan, que nos ayuda a encontrar soluciones de un sistema lineal dado y la inversa de una matriz si esta existe. En la Sección 3.7 definimos la transposición de matrices y damos una primera mirada a las matrices simétricas. Estas serán importantes en el Capítulo 8. También veremos la interacción entre transponer una matriz y el producto interno. En la última sección de este capítulo definimos las llamadas matrices elementales, que pueden verse como los bloques de construcción de las matrices invertibles. Las usaremos en el Capítulo 4 para demostrar propiedades importantes del determinante.

3.1. Sistemas lineales y eliminación de Gauß y de Gauß-Jordan

Comenzamos con un sistema lineal como en la Definición 1.8:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \tag{3.1}$$

Recordemos que el sistema se llama *consistente* si tiene al menos una solución; de lo contrario se llama *inconsistente*. De acuerdo con (1.4) y (1.5), su *matriz de coeficientes* asociada y su *matriz aumentada de coeficientes* son

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad (3.2)$$

y

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right). \quad (3.3)$$

Definición 3.1. El conjunto de todas las matrices con m filas y n columnas se denota por $M(m \times n)$. Si queremos enfatizar que la matriz tiene solo entradas reales, entonces escribimos $M(m \times n, \mathbb{R})$ o $M_{\mathbb{R}}(m \times n)$. Otra notación usada con frecuencia es $M_{m \times n}$. Una matriz A se llama una *matriz cuadrada* si su número de filas es igual a su número de columnas.

Para resolver (3.1), podríamos usar la primera ecuación, despejar x_1 e insertar esto en todas las demás ecuaciones. Nos da un nuevo sistema con $m - 1$ ecuaciones para $n - 1$ incógnitas. Luego despejamos x_2 en la siguiente ecuación, lo insertamos en las demás ecuaciones, y continuamos así hasta que solo quede una ecuación. Esto, por supuesto, fallará si por ejemplo $a_{11} = 0$, porque en este caso no podemos despejar x_1 en la primera ecuación. Podríamos salvar nuestro algoritmo diciendo: despejamos en la primera ecuación la primera incógnita cuyo coeficiente es distinto de 0 (o podríamos tomar una ecuación en la que el coeficiente de x_1 sea distinto de 0 y declarar que esta será nuestra primera ecuación. Después de todo, podemos ordenar las ecuaciones como queramos). Incluso con esta modificación, el proceso de resolver y reemplazar es propenso a errores.

Otra idea es manipular las ecuaciones. La pregunta es: ¿qué cambios en las ecuaciones están permitidos sin cambiar la información contenida en el sistema? No queremos destruir información (permitiendo así potencialmente más soluciones) ni introducir más información (eliminando así potencialmente soluciones). O, en términos más matemáticos, ¿qué cambios en el sistema de ecuaciones dado resultan en un sistema equivalente? Aquí llamamos *equivalentes* a dos sistemas si tienen el mismo conjunto de soluciones.

Podemos verificar si el nuevo sistema es equivalente al original si existe una forma de recuperar el original.

Por ejemplo, si intercambiamos la primera y la segunda fila, entonces en realidad no ocurrió nada y terminamos con un sistema equivalente. Podemos volver a la ecuación original simplemente intercambiando otra vez la primera y la segunda fila. Si multiplicamos ambos lados de la primera ecuación por algún factor, digamos por 2, entonces de nuevo nada cambia. Suponga por ejemplo que la primera ecuación es $x + 3y = 7$. Si multiplicamos ambos lados por 2, obtenemos $2x + 6y = 14$. Claramente, si un par (x, y) satisface la primera ecuación, entonces también satisface la segunda y viceversa. Dada la nueva ecuación $3x + 6y = 14$, podemos restaurar fácilmente la anterior simplemente dividiendo ambos lados por 2.

Si tomamos una ecuación y multiplicamos ambos lados por 0, entonces *destruimos* información porque terminamos con $0 = 0$ y no hay forma de recuperar la información que estaba almacenada en la ecuación original. Así que esta no es una operación permitida.

Muestre que elevar al cuadrado ambos lados de una ecuación en general no da una ecuación equivalente. ¿Hay casos en los que sí lo hace?

Elevar al cuadrado una ecuación o tomar el logaritmo en ambos lados u otras cosas de este tipo usualmente no nos interesan porque la ecuación resultante ya no será una ecuación lineal.

Denotemos por R_j la j -ésima fila de nuestro sistema lineal (3.1). La siguiente tabla contiene las llamadas *operaciones elementales por filas*. Son las operaciones “permitidas” porque no alteran la información contenida en un sistema lineal dado, ya que son reversibles. La primera columna describe la operación en palabras, la segunda introduce su notación abreviada y en la última fila damos la operación inversa que nos permite volver al sistema original.

Operación elemental	Notación	Operación inversa
① Intercambiar las filas j y k .	$R_j \leftrightarrow R_k$	$R_j \leftrightarrow R_k$
② Multiplicar la fila j por algún $\lambda \neq 0$.	$R_j \rightarrow \lambda R_j$	$R_j \rightarrow \frac{1}{\lambda} R_j$
③ Reemplazar la fila k por la suma de la fila k y λ veces R_j , y dejar la fila j sin cambiar ($j \neq k$).	$R_k \rightarrow R_k + \lambda R_j$	$R_k \rightarrow R_k - \lambda R_j$

Ejercicio. Muestre que la operación en la tercera columna invierte la operación de la segunda columna.

Ejercicio. Muestre que en lugar de la operación ③ basta tomar ③': $R_k \rightarrow R_k + R_j$, porque ③ se puede escribir como una composición de operaciones de la forma ② y ③'. Muestre cómo se puede hacer esto.

Ejercicio. Muestre que en realidad ① no es necesaria, ya que se puede obtener mediante una composición de operaciones de la forma ② y ③ (o ② y ③'). Muestre cómo se puede hacer esto.

Veamos en un ejemplo cómo funciona esto.

Ejemplo 3.2.

$$\begin{aligned}
 \left\{ \begin{array}{l} x_1 + x_2 - x_3 = 1 \\ 2x_1 + 3x_2 + x_3 = 3 \\ 4x_2 + x_3 = 7 \end{array} \right\} &\xrightarrow{R_2 \rightarrow R_2 - 2R_1} \left\{ \begin{array}{l} x_1 + x_2 - x_3 = 1 \\ x_2 + 3x_3 = 1 \\ 4x_2 + x_3 = 7 \end{array} \right\} \xrightarrow{R_3 \rightarrow R_3 - 4R_2} \left\{ \begin{array}{l} x_1 + x_2 - x_3 = 1 \\ x_2 + 3x_3 = 1 \\ -11x_3 = 3 \end{array} \right\} \\
 &\xrightarrow{R_3 \rightarrow -\frac{1}{11}R_3} \left\{ \begin{array}{l} x_1 + x_2 - x_3 = 1 \\ x_2 + 3x_3 = 1 \\ x_3 = -3/11. \end{array} \right.
 \end{aligned}$$

Aquí podemos detenernos porque ya es bastante fácil leer la solución. Procediendo de abajo hacia arriba, obtenemos

$$x_3 = -3/11, \quad x_2 = 1 - 3x_3 = 20/11, \quad x_1 = 1 + x_3 - x_2 = -12/11.$$

Observe que podríamos continuar nuestras manipulaciones por filas para limpiar aún más el sistema:

$$\begin{aligned} \dots &\rightarrow \left\{ \begin{array}{rcl} x_1 + x_2 - x_3 & = & 1 \\ x_2 + 3x_3 & = & 1 \\ -11x_3 & = & 3 \end{array} \right\} \xrightarrow{R_3 \rightarrow -1/11 R_3} \left\{ \begin{array}{rcl} x_1 + x_2 - x_3 & = & 1 \\ x_2 + 3x_3 & = & 1 \\ x_3 & = & -3/11 \end{array} \right\} \\ &\xrightarrow{R_2 \rightarrow R_2 - 3R_3} \left\{ \begin{array}{rcl} x_1 + x_2 - x_3 & = & 1 \\ x_2 & = & 20/11 \\ x_3 & = & -3/11 \end{array} \right\} \xrightarrow{R_1 \rightarrow R_1 + R_3} \left\{ \begin{array}{rcl} x_1 + x_2 & = & 8/11 \\ x_2 & = & 20/11 \\ x_3 & = & -3/11 \end{array} \right\} \\ &\xrightarrow{R_1 \rightarrow R_1 - R_2} \left\{ \begin{array}{rcl} x_1 + & = & -12/11 \\ x_2 & = & 20/11 \\ x_3 & = & -3/11 \end{array} \right\} \end{aligned}$$

Nuestra estrategia fue aplicar manipulaciones que eliminan sucesivamente las incógnitas en las ecuaciones inferiores, y buscábamos llegar a una forma del sistema de ecuaciones en la que la última ecuación contuviera el menor número posible de incógnitas.

Convénzase de que el primer paso de nuestro proceso de reducción es equivalente a despejar x_1 en la primera ecuación e insertarlo en las otras ecuaciones para eliminarlo allí. El siguiente paso en la reducción es equivalente a despejar x_2 en la nueva segunda ecuación e insertarlo en la tercera ecuación.

Es importante observar que hay infinitas rutas diferentes que conducen al resultado final, pero usualmente algunas son más rápidas que otras.

Analícemos lo que hicimos. Observamos los *coeficientes* del sistema y aplicamos transformaciones de tal manera que se vuelvan 0, porque esto resulta en eliminar las incógnitas correspondientes de las ecuaciones. Así, en el ejemplo anterior podríamos igualmente borrar todos los x_j , quedarnos solo con la matriz aumentada de coeficientes y realizar las operaciones por filas en la matriz. Por supuesto, tenemos que recordar que los números en las primeras columnas son los coeficientes de x_1 , los de la segunda columna son los coeficientes de x_2 , etc. Entonces nuestros cálculos se traducen en lo siguiente:

$$\begin{aligned} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 3 & 1 & 3 \\ 0 & 4 & 1 & 7 \end{array} \right) &\xrightarrow{R_2 \rightarrow R_2 - 2R_1} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 4 & 1 & 7 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 - 4R_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -11 & 3 \end{array} \right) \\ &\xrightarrow{R_3 \rightarrow -1/11 R_3} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & -3/11 \end{array} \right). \end{aligned}$$

Si traducimos esto de vuelta a un sistema lineal, obtenemos

$$\begin{aligned}x_1 + x_2 + x_3 &= 1, \\x_2 + 3x_3 &= 3, \\x_3 &= -3/11,\end{aligned}$$

que se puede resolver fácilmente de abajo hacia arriba.

Hicimos exactamente los mismos cálculos que hicimos con el sistema de ecuaciones, pero en notación matricial se ve mucho más ordenado, ya que no tenemos que escribir las incógnitas todo el tiempo.

Si queremos resolver un sistema lineal, lo escribimos como una matriz aumentada y luego realizamos operaciones por filas hasta llegar a una forma “bonita” en la que podamos leer las soluciones si las hay.

Pero, ¿qué es una forma “bonita”? Recuerde que si un coeficiente es 0, entonces la incógnita correspondiente no aparece en la ecuación.

- Todas las filas que solo contienen ceros deberían estar abajo.
- En la primera ecuación no nula desde abajo, queremos la menor cantidad posible de incógnitas y queremos que sean las últimas incógnitas. Así, como última fila queremos una que solo tenga ceros o una que comience con ceros, hasta que finalmente obtengamos un número no nulo, digamos en la columna k . Este número no nulo siempre se puede hacer igual a 1 dividiendo la fila por él. Ahora sabemos cómo están relacionadas las incógnitas $x_k, \dots, x_n$. Observe que todas las demás incógnitas $x_1, \dots, x_{k-1}$ han desaparecido de la ecuación, ya que sus coeficientes son 0. Si $k = n$, como en nuestro ejemplo anterior, entonces hay una única solución para x_n .
- La segunda fila no nula desde abajo también debería comenzar con ceros hasta llegar a una columna, digamos la columna l , con una entrada no nula que siempre podemos hacer igual a 1. Esta columna debería estar a la izquierda de la columna k (es decir, queremos $l < k$). Porque ahora podemos usar lo que sabemos de la última fila acerca de las incógnitas $x_k, \dots, x_n$ para decir algo acerca de las incógnitas $x_l, \dots, x_{k-1}$.
- Continuamos así hasta que todas las filas estén como queremos.

Observe que la forma de una matriz “bonita” de este tipo se parece un poco a una matriz que tuviera un triángulo formado por solo ceros en su parte inferior izquierda. Puede haber ceros en la parte superior derecha. Si una matriz tiene la forma que acabamos de describir, decimos que está en *forma escalonada por filas*. Demos una definición precisa.

Definición 3.3 (Forma escalonada por filas). Decimos que una matriz $A \in M(m \times n)$ está en *forma escalonada por filas* si:

- Todas sus filas cero son las últimas filas.
- La primera entrada no nula en una fila es 1. Se llama el *pivote* de la fila.
- El pivote de cualquier fila está estrictamente a la derecha del de la fila anterior.

Definición 3.4 (Forma escalonada reducida por filas). Decimos que una matriz $A \in M(m \times n)$ está en *forma escalonada reducida por filas* si:

- A está en forma escalonada por filas.
- Todas las entradas en A que están por encima de un pivote son iguales a 0.

Veamos rápidamente algunos ejemplos.

Ejemplos 3.5.

- (a) Las siguientes matrices están en forma escalonada reducida por filas. Los pivotes están resaltados.

$$\begin{pmatrix} 1 & 6 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 6 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- (b) Las siguientes matrices están en forma escalonada por filas, pero no en forma escalonada reducida por filas. Los pivotes están resaltados.

$$\begin{pmatrix} 1 & 6 & 3 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 6 & 3 & 1 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 6 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- (c) Las siguientes matrices no están en forma escalonada por filas:

$$\begin{pmatrix} 1 & 6 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 6 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 6 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 3 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Ejercicio. ■ Diga por qué las matrices en (b) no están en forma escalonada reducida por filas y use operaciones elementales por filas para transformarlas en una matriz en forma escalonada reducida por filas.

- Diga por qué las matrices en (c) no están en forma escalonada por filas y use operaciones elementales por filas para transformarlas en una matriz en forma escalonada por filas. Transfórmelas más para obtener una matriz en forma escalonada reducida por filas.

Question 3.1

Si intercambiamos dos filas en una matriz, esto corresponde a escribir las ecuaciones dadas en un orden diferente. ¿Cuál es el efecto sobre un sistema lineal si intercambiamos dos columnas?

Recuerde: si traducimos un sistema lineal a una matriz aumentada de coeficientes $(A|b)$, realizamos las operaciones por filas para llegar a una forma escalonada (reducida) por filas $(A'|b')$, y traducimos de vuelta a un sistema lineal, entonces este nuevo sistema contiene exactamente la misma información que el original, pero está “más ordenado” y es fácil determinar su solución.

La pregunta natural ahora es: ¿Podemos siempre transformar una matriz en una matriz en forma escalonada (reducida) por filas? La respuesta es que esto siempre es posible e incluso podemos dar un algoritmo para hacerlo.

Eliminación de Gauß. Sea $A \in M(m \times n)$ y suponga que A no es la matriz cero. La eliminación de Gauß es un algoritmo que transforma A en una forma escalonada por filas. Los pasos son los siguientes:

- Encontrar la primera columna que no consiste enteramente de ceros. Intercambiar filas de manera apropiada de modo que la entrada en esa columna en la primera fila sea distinta de cero.
- Multiplicar la primera fila por un número apropiado de modo que su primera entrada no nula sea 1.
- Usar la primera fila para eliminar todos los coeficientes debajo de su pivote.
- Ahora nuestra matriz se ve como

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & \cdots \\ \vdots & & \vdots & 0 & \ddots \\ 0 & & 0 & 0 & A' \end{pmatrix}.$$

donde $*$ son números arbitrarios y A' es una matriz con menos columnas que A y $m - 1$ filas. Ahora repita el proceso para A' . Observe que al hacer esto las primeras columnas no cambian, ya que solo estamos manipulando ceros.

Eliminación de Gauß-Jordan. Sea $A \in M(m \times n)$. La eliminación de Gauß-Jordan es un algoritmo que transforma A en una forma escalonada reducida por filas. Los pasos son los siguientes:

- Usar la eliminación de Gauß para obtener una forma escalonada por filas de A .
- Usar los pivotes para eliminar los coeficientes no nulos que están en columnas por encima de un pivote.

Por supuesto, si hacemos una reducción a mano, entonces no tenemos que seguir estrictamente los pasos del algoritmo si esto hace que los cálculos sean más fáciles. Sin embargo, estos algoritmos siempre funcionan y por lo tanto pueden programarse para que una computadora los realice.

Definición 3.6. Dos matrices $m \times n$ A y B se llaman *equivalentes por filas* si hay operaciones elementales por filas que transforman A en B . (Claramente entonces B se puede transformar mediante operaciones por filas en A .)

Observación. Sea A una matriz $m \times n$.

- A se puede transformar en infinitas formas escalonadas por filas diferentes.
- Solo hay una forma escalonada reducida por filas a la que A se puede transformar.

Demuestre las afirmaciones anteriores.

Antes de dar ejemplos, observamos que a partir de la forma escalonada por filas podemos saber inmediatamente cuántas soluciones tiene el sistema lineal correspondiente.

Teorema 3.7. Sea $(A|b)$ la matriz aumentada de coeficientes de un sistema lineal $m \times n$ y sea $(A'|b')$ una forma reducida por filas.

- (1) Si hay una fila de la forma $(0 \cdots 0|\beta)$ con $\beta \neq 0$, entonces el sistema no tiene solución.
- (2) Si no hay una fila de la forma $(0 \cdots 0|\beta)$ con $\beta \neq 0$, entonces ocurre una de las siguientes posibilidades:
 - (2.1) Si hay un pivote en cada columna de A' , entonces el sistema tiene exactamente una solución.
 - (2.2) Si hay una columna en A' sin pivote, entonces el sistema tiene infinitas soluciones.

Demostración. (1) Si $(A'|b')$ tiene una fila de la forma $(0 \cdots 0|\beta)$ con $\beta \neq 0$, entonces la ecuación correspondiente es $0x_1 + \cdots + 0x_n = \beta$, que claramente no tiene solución.

- (2) Supongamos ahora que $(A'|b')$ no tiene ninguna fila de la forma $(0 \cdots 0|\beta)$ con $\beta \neq 0$. En el caso (2.1), la matriz transformada es entonces de la forma

$$\left(\begin{array}{cccc|c} 1 & a'_{12} & a'_{13} & & a'_{1n} & b'_1 \\ 0 & 1 & a'_{23} & & a'_{2n} & b'_2 \\ & & 1 & & \vdots & \vdots \\ & & & \ddots & 1 & a'_{(n-1)n} & b'_{n-1} \\ & & & & 0 & 1 & b'_n \\ & & & & & & 0 \\ & & & & & & 0 \\ & & & & & & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & & 0 \end{array} \right). \quad (3.4)$$

Observe que las últimas filas cero aparecen solo si $n < m$. Este sistema claramente tiene la solución única

$$x_n = b'_n, \quad x_{n-1} = b'_{n-1} - a'_{(n-1)n}x_n, \quad \dots, \quad x_1 = b'_1 - a'_{1n}x_n - \cdots - a'_{12}x_2.$$

En el caso (2.2), la matriz transformada es entonces de la forma

$$\left(\begin{array}{cccc|cccc} 0 & \cdots & 0 & \textcolor{violet}{1} & * & \cdots & \cdots & * & b'_1 \\ & & & & & & & & \\ 0 & & & 0 & \textcolor{violet}{1} & * & \cdots & * & b'_2 \\ & & & & & & & & \\ 0 & & & & & 0 & \textcolor{violet}{1} & * & b'_3 \\ & & & & & & & & \vdots \\ & & 0 & & & & & 0 & \textcolor{violet}{1} & * & b'_k \\ & & & & & & & & & & 0 \\ & & & & & & & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{array} \right). \quad (3.5)$$

donde las estrellas representan números. (Si continuamos la reducción hasta llegar a la forma escalonada reducida por filas, entonces los números sobre los 1 deben ser ceros.) Observe que podemos elegir arbitrariamente las incógnitas que corresponden a las columnas sin pivote. Las incógnitas que corresponden a las columnas con pivotes pueden entonces elegirse siempre de manera única para que el sistema se satisfaga. $\square$

Definición 3.8. Las variables que corresponden a columnas sin pivotes se llaman *variables libres*.

Volveremos a este teorema más adelante en la página 115 (el teorema se enuncia nuevamente en el recuadro de color). Del teorema anterior obtenemos como consecuencia inmediata lo siguiente.

Teorema 3.9. *Un sistema lineal tiene o bien ninguna, exactamente una o infinitas soluciones.*

Veamos ahora algunos ejemplos.

Ejemplo 3.10 (Ejemplo con una solución única (sin variables libres)). Consideramos el sistema lineal

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= 12, \\ -x_1 + 2x_2 + 3x_3 &= 15, \\ 3x_1 &\quad - x_3 = 1. \end{aligned} \quad (3.6)$$

Solución. Formamos la matriz aumentada y realizamos reducción por filas.

$$\begin{aligned}
 & \left(\begin{array}{ccc|c} 2 & 3 & 1 & 12 \\ -1 & 2 & 3 & 15 \\ 3 & 0 & -1 & 1 \end{array} \right) \xrightarrow{R_1 \rightarrow R_1 + 2R_2} \left(\begin{array}{ccc|c} 0 & 7 & 7 & 42 \\ -1 & 2 & 3 & 15 \\ 3 & 0 & -1 & 1 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 + 3R_2} \left(\begin{array}{ccc|c} 0 & 7 & 7 & 42 \\ -1 & 2 & 3 & 15 \\ 0 & 6 & 8 & 46 \end{array} \right) \\
 & \xrightarrow{R_1 \leftrightarrow R_2} \left(\begin{array}{ccc|c} -1 & 2 & 3 & 15 \\ 0 & 7 & 7 & 42 \\ 0 & 6 & 8 & 46 \end{array} \right) \xrightarrow{\substack{R_1 \rightarrow -R_1 \\ R_2 \rightarrow \frac{1}{7}R_2}} \left(\begin{array}{ccc|c} 1 & -2 & -3 & -15 \\ 0 & 1 & 1 & 6 \\ 0 & 6 & 8 & 46 \end{array} \right) \\
 & \xrightarrow{R_3 \rightarrow R_3 - 6R_2} \left(\begin{array}{ccc|c} 1 & -2 & -3 & -15 \\ 0 & 1 & 1 & 6 \\ 0 & 0 & 2 & 10 \end{array} \right) \xrightarrow{R_3 \rightarrow \frac{1}{2}R_3} \left(\begin{array}{ccc|c} 1 & -2 & -3 & -15 \\ 0 & 1 & 1 & 6 \\ 0 & 0 & 1 & 5 \end{array} \right).
 \end{aligned}$$

Esto muestra que el sistema (3.6) es equivalente al sistema

$$\begin{aligned}
 x_1 - 2x_2 - 3x_3 &= -15, \\
 x_2 + x_3 &= 6, \\
 x_3 &= 5
 \end{aligned} \tag{3.7}$$

cuya solución es fácil de escribir:

$$x_3 = 5, \quad x_2 = 6 - x_3 = 1, \quad x_1 = -15 + 2x_2 + 3x_3 = 2. \quad \diamond$$

Observación. Si continuamos el proceso de reducción hasta llegar a la forma escalonada reducida por filas, entonces obtenemos

$$\begin{aligned}
 \dots \longrightarrow & \left(\begin{array}{ccc|c} 1 & -2 & -3 & -15 \\ 0 & 1 & 1 & 6 \\ 0 & 0 & 1 & 5 \end{array} \right) \xrightarrow{R_2 \rightarrow R_2 - R_3} \left(\begin{array}{ccc|c} 1 & -2 & -3 & -15 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 5 \end{array} \right) \xrightarrow{R_1 \rightarrow R_1 + 3R_3} \left(\begin{array}{ccc|c} 1 & -2 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 5 \end{array} \right) \\
 & \xrightarrow{R_1 \rightarrow R_1 + 2R_2} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 5 \end{array} \right).
 \end{aligned}$$

Por lo tanto, el sistema (3.6) es equivalente al sistema

$$\begin{aligned}
 x_1 &= 2, \\
 x_2 &= 1, \\
 x_3 &= 5.
 \end{aligned}$$

cuya solución se puede leer inmediatamente como

$$x_3 = 5, \quad x_2 = 1, \quad x_1 = 2.$$

Ejemplo 3.11 (Ejemplo con dos variables libres). Consideramos el sistema lineal

$$\begin{aligned}
 3x_1 - 2x_2 + 3x_3 + 3x_4 &= 3, \\
 2x_1 + 6x_2 + 2x_3 - 9x_4 &= 2, \\
 x_1 + 2x_3 + x_4 - 3x_4 &= 1.
 \end{aligned} \tag{3.8}$$

Solución. Formamos la matriz aumentada y realizamos reducción por filas.

$$\begin{aligned}
 \left(\begin{array}{cccc|c} 3 & -2 & 3 & 3 & 3 \\ 2 & 6 & 2 & -9 & 2 \\ 1 & 2 & 1 & -3 & 1 \end{array} \right) & \xrightarrow{R_2 \rightarrow R_2 - 2R_3} \left(\begin{array}{cccc|c} 3 & -2 & 3 & 3 & 3 \\ 0 & 2 & 0 & -3 & 0 \\ 1 & 2 & 1 & -3 & 1 \end{array} \right) \xrightarrow{R_1 \rightarrow R_1 - 3R_3} \left(\begin{array}{cccc|c} 0 & -8 & 0 & 12 & 0 \\ 0 & 2 & 0 & -3 & 0 \\ 1 & 2 & 1 & -3 & 1 \end{array} \right) \\
 & \xrightarrow{R_1 \leftrightarrow R_3} \left(\begin{array}{cccc|c} 1 & 2 & 1 & -3 & 1 \\ 0 & 2 & 0 & -3 & 0 \\ 0 & -8 & 0 & 12 & 0 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 + 4R_2} \left(\begin{array}{cccc|c} 1 & 2 & 1 & -3 & 1 \\ 0 & 2 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \\
 & \xrightarrow{R_1 \rightarrow R_1 - R_2} \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right).
 \end{aligned}$$

La 3.^a y la 4.^a columna no tienen pivotes y vemos que el sistema (3.8) es equivalente al sistema

$$\begin{aligned}
 x_1 - x_3 &= 1, \\
 x_2 + x_4 &= 0.
 \end{aligned}$$

Claramente podemos elegir x_3 y x_4 (las incógnitas correspondientes a las columnas sin pivote) arbitrariamente. Siempre podremos ajustar x_1 y x_2 de tal manera que el sistema se satisfaga. Para dejar claro que x_3 y x_4 son nuestras variables libres, a veces las llamamos $x_3 = t$ y $x_4 = s$. Entonces toda solución del sistema (3.8) es de la forma

$$x_1 = 1 + t, \quad x_2 = -s, \quad x_3 = t, \quad x_4 = s, \quad \text{para } s, t \in \mathbb{R} \text{ arbitrarios.} \quad \diamond$$

En *forma vectorial* podemos escribir la solución como sigue. Una tupla (x_1, x_2, x_3, x_4) es una solución de (3.8) si y solo si el vector correspondiente es de la forma

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1+t \\ -s \\ t \\ s \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} \quad \text{para algunos } s, t \in \mathbb{R}.$$

Geométricamente, el conjunto de todas las soluciones es un plano afín en $\mathbb{R}^4$.

Observación 3.12. Las soluciones de un sistema inhomogéneo de ecuaciones lineales en notación vectorial son siempre de la forma

$$\vec{x} = \vec{z}_0 + t_1 \vec{y}_1 + \cdots + t_k \vec{y}_k.$$

Esto será importante en el Teorema 3.22. En el ejemplo anterior, $k = 2$ y

$$\vec{z}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{y}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{y}_2 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

Ejemplo 3.13 (Ejemplo sin solución). Consideramos el sistema lineal

$$\begin{aligned} 2x_1 + x_2 - x_3 &= 7, \\ 3x_1 + 2x_2 - 2x_3 &= 7, \\ -x_1 + 3x_2 - 3x_3 &= 2. \end{aligned} \tag{3.9}$$

Solución. Formamos la matriz aumentada y realizamos reducción por filas.

$$\begin{aligned} \left(\begin{array}{ccc|c} 2 & 1 & -1 & 7 \\ 3 & 2 & -2 & 7 \\ -1 & 3 & -3 & 2 \end{array} \right) &\xrightarrow{R_1 \rightarrow R_1 + 2R_3} \left(\begin{array}{ccc|c} 0 & 7 & -7 & 11 \\ 3 & 2 & -2 & 7 \\ -1 & 3 & -3 & 2 \end{array} \right) &\xrightarrow{R_2 \rightarrow R_2 + 3R_3} \left(\begin{array}{ccc|c} 0 & 7 & -7 & 11 \\ 0 & 11 & -11 & 13 \\ -1 & 3 & -3 & 2 \end{array} \right) \\ &\xrightarrow{R_1 \leftrightarrow R_3} \left(\begin{array}{ccc|c} -1 & 3 & -3 & 2 \\ 0 & 11 & -11 & 13 \\ 0 & 7 & -7 & 11 \end{array} \right) &\xrightarrow{R_3 \rightarrow R_3 - \frac{7}{11}R_2} \left(\begin{array}{ccc|c} -1 & 3 & -3 & 2 \\ 0 & 11 & -11 & 13 \\ 0 & 0 & 0 & 30 \end{array} \right). \end{aligned}$$

La última línea nos dice inmediatamente que el sistema (3.9) no tiene solución porque no hay elección de x_1, x_2, x_3 tal que $0x_1 + 0x_2 + 0x_3 = 30$. $\diamond$

Ahora debería haber entendido

- qué significa que dos sistemas lineales sean equivalentes,
- cuáles operaciones por filas transforman un sistema dado en uno equivalente y por qué esto es así,
- cuándo una matriz está en forma escalonada por filas y en forma escalonada reducida por filas,
- por qué el sistema lineal asociado a una matriz en forma escalonada (reducida) es fácil de resolver,
- qué hacen la eliminación de Gauß y la de Gauß-Jordan, y por qué siempre funcionan,
- que la eliminación de Gauß y la de Gauß-Jordan no son nada muy mágico; esencialmente son lo mismo que despejar variables y reemplazar en las ecuaciones restantes. Solo lo hacen de una manera sistemática,
- por qué una matriz dada puede transformarse en muchas formas escalonadas por filas distintas, pero en una única forma escalonada reducida por filas,
- por qué un sistema lineal siempre tiene o bien ninguna, exactamente una o infinitas soluciones,
- etc.

Ahora debería ser capaz de

- identificar si una matriz está en forma escalonada por filas o en forma escalonada reducida por filas,
- usar la eliminación de Gauß o la de Gauß-Jordan para resolver sistemas lineales,

- decir si un sistema tiene ninguna, exactamente una o infinitas soluciones si conoce su forma escalonada,
- etc.

Ejercicios.

- Por medio de operaciones elementales, lleve las siguientes matrices aumentadas a sus formas escalonadas reducidas y encuentre todas las soluciones del sistema de ecuaciones asociado a cada matriz:

(a) $\left(\begin{array}{ccc|c} 1 & 2 & 3 & -2 \\ 2 & -1 & -1 & 1 \end{array}\right)$

(d) $\left(\begin{array}{ccc|c} 5 & 1 & -5 & 1 \\ 4 & 2 & 2 & -4 \\ 1 & 0 & 3 & 2 \\ -3 & 1 & 4 & -1 \end{array}\right)$

(g) $\left(\begin{array}{ccc|c} 0 & 0 & 1 & -3 \\ 0 & 2 & 1 & 1 \\ 3 & 2 & 1 & 7 \end{array}\right)$

(b) $\left(\begin{array}{ccc|c} 2 & 4 & 6 & -1 \\ 4 & 5 & 6 & 2 \\ 2 & 7 & 12 & 1 \end{array}\right)$

(e) $\left(\begin{array}{cccc|c} 1 & 1 & 5 & -1 & 2 \\ 0 & 1 & 2 & 1 & 0 \\ 2 & -1 & 4 & 1 & 13 \end{array}\right)$

(h) $\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 1 \\ 1 & 2 & 3 & 0 & -2 \\ 1 & 2 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 2 \end{array}\right)$

(c) $\left(\begin{array}{cc|c} 6 & 2 & 4 \\ 1 & -2 & -4 \\ 1 & 1 & 2 \end{array}\right)$

(f) $\left(\begin{array}{ccc|c} 0 & 2 & 3 & 1 \\ 2 & -6 & 7 & 3 \\ 1 & -2 & 5 & 2 \end{array}\right)$

(i) $\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 2 & 1 & 0 & 0 & -3 \\ 3 & 2 & 1 & 0 & 2 \\ 4 & 3 & 2 & 1 & 5 \end{array}\right)$

- Encontrar condiciones sobre a, b, c tal que el siguiente sistema tenga solución:

$$\begin{aligned} x + 3y - 2z &= a + 1, \\ 3x - y + z &= b + 6, \\ 5x - 5y + 4z &= c + 11. \end{aligned}$$

- Encuentre un polinomio de grado a lo más 2 que pase por los puntos $(-1, -6)$, $(1, 0)$, $(2, 0)$. ¿Cuántos polinomios con esta propiedad hay?
- ¿Existe un polinomio de grado 1 que pase por los tres puntos del Ejercicio 3.? ¿Cuántos tales polinomios hay?
 - ¿Existe un polinomio de grado 3 que pase por los tres puntos del Ejercicio 3.? ¿Cuántos tales polinomios hay? Dé por lo menos dos polinomios de grado 3.
- Un rolo hace un pequeño tour por Colombia y revisando sus cuentas nota que: En hostales gastó \$30,000 diarios en Medellín, \$20,000 diarios en Villavicencio y \$20,000 diarios en Yopal, en alimentación gastó \$20,000 diarios en Medellín, \$30,000 diarios en Villavicencio y \$20,000 diarios en Yopal y finalmente en desplazamientos gastó en promedio \$10,000 diariamente en cada ciudad. Si se sabe que durante todo el tour gastó \$340,000 en hostales, \$320,000 en alimentación y \$140,000 en desplazamientos, calcule el número de días que estuvo el rolo en cada ciudad.

6. Un criadero del llano tiene chigueros, conejos y ratones arroceros. Cada chiguero consume cada semana en promedio un kilo de frutas, un kilo de hierbas y dos kilos de arroz, cada conejo consume semanalmente en promedio tres kilos de frutas, cuatro kilos de hierbas y cinco kilos de arroz y cada ratón arrocero consume cada semana en promedio dos kilos de frutas, un kilo de hierbas y cinco kilos de arroz. Cada semana se proporcionan 25,000 kilos de frutas, 20,000 de hierbas y 55,000 de arroz. Si las tres especies de roedores se comen todo el alimento ¿Cuántos roedores de cada tipo pueden vivir en el criadero?

3.2. Sistemas lineales homogéneos

En esta sección corta tratamos el caso especial de los sistemas lineales homogéneos. Recordemos que un sistema lineal (3.1) se llama homogéneo si $b_1 = \dots = b_n = 0$. Tal sistema siempre tiene al menos una solución, la llamada *solución trivial* $x_1 = \dots = x_n = 0$. Esto también es claro a partir del Teorema 3.7, ya que sin importar qué operaciones por filas realicemos, el lado derecho siempre permanecerá igual a 0. Observe que si realizamos eliminación de Gauss o de Gauss-Jordan, no es necesario escribir el lado derecho, ya que siempre será 0.

Si adaptamos el Teorema 3.7 al caso especial de un sistema homogéneo, obtenemos lo siguiente.

Teorema 3.14. *Sea A la matriz de coeficientes de un sistema lineal homogéneo $m \times n$ y sea A' una forma reducida por filas.*

- (I) *Si hay un pivote en cada columna, entonces el sistema tiene exactamente una solución, a saber, la solución trivial.*
- (II) *Si hay una columna sin pivote, entonces el sistema tiene infinitas soluciones.*

Corolario 3.15. *Un sistema lineal homogéneo tiene o bien exactamente una solución o infinitas soluciones.*

Veamos un ejemplo.

Ejemplo 3.16 (Ejemplo de un sistema homogéneo con infinitas soluciones). Consideramos el sistema lineal

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 0, \\ 2x_1 + 3x_2 - 2x_3 &= 0, \\ 3x_1 - x_2 - 3x_3 &= 0. \end{aligned} \tag{3.10}$$

Solución. Realizamos reducción por filas en la matriz asociada.

$$\begin{aligned} \begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & -2 \\ 3 & -1 & -3 \end{pmatrix} &\xrightarrow{R_2 \rightarrow R_2 - 2R_1} \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 0 \\ 3 & -1 & -3 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 - 3R_1} \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 0 \\ 0 & -7 & 0 \end{pmatrix} \\ &\xrightarrow{\text{usar } R_2 \text{ para eliminar la 2.ª columna}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{R_2 \rightarrow -R_2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Vemos que la tercera variable es libre, así que ponemos $x_3 = t$. La solución es

$$x_1 = t, \quad x_2 = 0, \quad x_3 = t \quad \text{for } t \in \mathbb{R},$$

o, en forma vectorial

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{for } t \in \mathbb{R}. \quad \diamond$$

Ejemplo 3.17 (Ejemplo de un sistema homogéneo con exactamente una solución). Consideramos el sistema lineal

$$\begin{aligned} x_1 + 2x_2 &= 0, \\ 2x_1 + 3x_2 &= 0, \\ 3x_1 + 5x_2 &= 0. \end{aligned} \quad (3.11)$$

Solución. Realizamos reducción por filas en la matriz asociada.

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 5 \end{pmatrix} \xrightarrow{\text{usar } R_1 \text{ para eliminar la 1.ª columna}} \begin{pmatrix} 1 & 2 \\ 0 & -1 \\ 0 & -1 \end{pmatrix} \xrightarrow{\text{usar } R_2 \text{ para eliminar la 2.ª columna}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 0 & 0 \end{pmatrix} \xrightarrow{R_2 \rightarrow -R_2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Así, la única solución posible es $x_1 = 0$ y $x_2 = 0$. $\diamond$

En la siguiente sección veremos la conexión entre el conjunto de soluciones de un sistema lineal y el sistema lineal homogéneo correspondiente.

Ahora debería haber entendido

- por qué un sistema lineal homogéneo siempre tiene o bien una solución o infinitas soluciones,
- etc.

Ahora debería ser capaz de

- usar la eliminación de Gauß o de Gauß-Jordan para resolver sistemas lineales homogéneos,
- etc.

Ejercicios.

1. Encuentre todas las soluciones de los sistemas homogéneos asociados a las matrices del Ejercicio 1. de la Sección 3.1.
2. Determine el conjunto de soluciones de los siguientes sistemas homogéneos:

$$\begin{array}{lll} \text{(a)} & x + 2y - 3z = 0 & \text{(b)} \quad x + y + z + w = 0 \\ & 2x + 4y - 6z = 0 & \quad -x + 5z - 3w = 0 \\ & -3x - 6y + 9z = 0 & \quad 2x + 3y + 8z = 0 \\ & & \quad x + 2y + 7z - w = 0 \end{array} \quad \text{(c)} \quad \begin{array}{l} 2x - 8y = 0 \\ -x + 4y = 0 \\ 3x - 12y = 0 \end{array}$$

3. Encuentre todos los $r \in \mathbb{R}$ tal que el siguiente sistema tenga solución única:

$$\begin{aligned}(2+r)x - 2y &= 0, \\ 2x + (1-r)y &= 0.\end{aligned}$$

3.3. Matrices y sistemas lineales

Hasta ahora se nos daba un sistema lineal con un lado derecho específico y nos preguntábamos qué x_j debíamos introducir en el sistema para obtener el lado derecho dado. Los problemas de este tipo se llaman *problemas inversos*, ya que se nos da una salida (el lado derecho del sistema; el “estado” que queremos alcanzar) y tenemos que encontrar una entrada adecuada para obtener la salida deseada.

Ahora cambiamos un poco nuestra perspectiva y nos preguntamos: si introducimos ciertos $x_1, \dots, x_n$ en el sistema, ¿qué obtenemos como resultado en el lado derecho? Para investigar esta pregunta, es muy útil escribir el sistema (3.1) en una forma abreviada. Primero observe que podemos verlo como una igualdad de dos vectores con m componentes cada uno:

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}. \quad (3.12)$$

Sea A la matriz de coeficientes y $\vec{x}$ el vector cuyas componentes son $x_1, \dots, x_n$. Entonces escribimos el lado izquierdo de (3.12) como

$$A\vec{x} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}. \quad (3.13)$$

Con esta notación, el sistema lineal (3.1) se puede escribir de manera muy breve como

$$A\vec{x} = \vec{b}$$

$$\text{with } \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Una forma de recordar la fórmula para la multiplicación de una matriz y un vector es que “multiplicamos cada fila de la matriz por el vector columna”, es decir, calculamos “fila por columna”. Por ejemplo, la j -ésima componente de $A\vec{x}$ es “(j -ésima fila de A) por (columna $\vec{x}$)”.

$$A\vec{x} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}. \quad (3.14)$$

Definición 3.18. La fórmula en (3.13) se llama *multiplicación de una matriz y un vector*.

Una matriz $m \times n$ A toma un vector con n componentes y nos devuelve un vector con m componentes.

Observe que algo como $\vec{x}A$ **no** tiene sentido!

Observación 3.19. La fórmula (3.13) se puede interpretar de la siguiente manera. Si A es una matriz $m \times n$ y $\vec{x}$ es un vector en $\mathbb{R}^n$, entonces $A\vec{x}$ es el vector en $\mathbb{R}^m$ que es la suma de las columnas de A ponderadas con los coeficientes dados por $\vec{x}$, ya que

$$\begin{aligned} A\vec{x} &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{pmatrix} + \cdots + \begin{pmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{pmatrix} \\ &= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}. \end{aligned} \quad (3.15)$$

Observación 3.20. Recordemos que $\vec{e}_j$ es el vector que tiene un 1 como su j -ésima componente y ceros en todas las demás. La fórmula (3.13) muestra que para todo $j = 1, \dots, n$

$$A\vec{e}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = j\text{-ésima columna de } A. \quad (3.16)$$

Demostremos algunas propiedades sencillas.

Proposición 3.21. Sea A una matriz $m \times n$, $\vec{x}, \vec{y} \in \mathbb{R}^n$ y $c \in \mathbb{R}$. Entonces

- (I) $A(c\vec{x}) = cA\vec{x}$,
- (II) $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}$,

$$(III) \quad A\vec{0} = \vec{0}.$$

Demostración. Las demostraciones no son difíciles. Se obtienen usando las definiciones y realizando algunos cálculos directos como sigue.

$$(I) \quad A(c\vec{x}) = A \begin{pmatrix} cx_1 \\ \vdots \\ cx_n \end{pmatrix} = \begin{pmatrix} a_{11}cx_1 + \cdots + a_{1n}cx_n \\ a_{21}cx_1 + \cdots + a_{2n}cx_n \\ \vdots \\ a_{m1}cx_1 + \cdots + a_{mn}cx_n \end{pmatrix} = c \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} = cA\vec{x}.$$

$$(II) \quad A(\vec{x} + \vec{y}) = A \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} = \begin{pmatrix} a_{11}(x_1 + y_1) + \cdots + a_{1n}(x_n + y_n) \\ a_{21}(x_1 + y_1) + \cdots + a_{2n}(x_n + y_n) \\ \vdots \\ a_{m1}(x_1 + y_1) + \cdots + a_{mn}(x_n + y_n) \end{pmatrix} \\ = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} + \begin{pmatrix} a_{11}y_1 + \cdots + a_{1n}y_n \\ a_{21}y_1 + \cdots + a_{2n}y_n \\ \vdots \\ a_{m1}y_1 + \cdots + a_{mn}y_n \end{pmatrix} = A\vec{x} + A\vec{y}.$$

(III) Para mostrar que $A\vec{0} = \vec{0}$, podríamos simplemente hacer el cálculo (¡que es muy fácil!) o podemos usar (I):

$$A\vec{0} = A(0\vec{0}) = 0A\vec{0} = \vec{0}. \quad \square$$

Observe que en (III) el $\vec{0}$ del lado izquierdo es el vector cero en $\mathbb{R}^n$, mientras que el $\vec{0}$ del lado derecho es el vector cero en $\mathbb{R}^m$.

La Proposición 3.21 da una idea importante sobre la estructura de las soluciones de los sistemas lineales. Véase también la Observación 3.12.

Teorema 3.22. (I) Sean $\vec{x}$ y $\vec{y}$ soluciones del sistema lineal (3.1). Entonces $\vec{x} - \vec{y}$ es una solución del sistema lineal homogéneo asociado.

(II) Sea $\vec{x}$ una solución del sistema lineal (3.1), sea $\vec{z}$ una solución del sistema lineal homogéneo asociado y sea $\lambda \in \mathbb{R}$. Entonces $\vec{x} + \lambda\vec{z}$ es solución of the system (3.1).

Demostración. Supongamos que $\vec{x}$ y $\vec{y}$ son soluciones de (3.1), es decir,

$$A\vec{x} = \vec{b} \quad \text{and} \quad A\vec{y} = \vec{b}.$$

Por la Proposición 3.21 (I) y (II) tenemos

$$A(\vec{x} - \vec{y}) = A\vec{x} + A(-\vec{y}) = A\vec{x} - A\vec{y} = \vec{b} - \vec{b} = \vec{0}$$

lo cual muestra que $\vec{x} - \vec{y}$ resuelve la ecuación homogénea $A\vec{v} = \vec{0}$. Así queda demostrado (I).

Para demostrar (II), procedemos de manera similar. Si $\vec{x}$ resuelve el sistema inhomogéneo (3.1) y $\vec{z}$ resuelve el sistema homogéneo asociado, entonces

$$A\vec{x} = \vec{b} \quad \text{and} \quad A\vec{z} = \vec{0}.$$

Ahora (II) se sigue de

$$A(\vec{x} + \lambda\vec{z}) = A\vec{x} + \lambda A\vec{z} = A\vec{x} + \lambda A\vec{z} = \vec{b} + \lambda\vec{0} = \vec{b}.$$

□

Corolario 3.23. Sea $\vec{x}$ una solución arbitraria del sistema inhomogéneo (3.1). Entonces el conjunto de todas las soluciones de (3.1) es

$$\{\vec{x} + \vec{z} : \vec{z} \text{ es solución del sistema homogéneo asociado}\}.$$

Esto significa que para encontrar todas las soluciones de un sistema inhomogéneo basta encontrar una solución particular y todas las soluciones del sistema homogéneo correspondiente.

Mostraremos más adelante que el conjunto de todas las soluciones de un sistema homogéneo es un espacio vectorial. Cuando estudie el conjunto de todas las soluciones de ecuaciones diferenciales lineales, encontrará la misma estructura.

Ejemplo 3.24. Consideremos el sistema

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 3, \\ 2x_1 + 3x_2 - 2x_3 &= 3, \\ 3x_1 - x_2 - 3x_3 &= -12. \end{aligned} \tag{3.10'}$$

Solución. Formamos la matriz aumentada y realizamos reducción por filas.

$$\begin{aligned} \left(\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 2 & 3 & -2 & 3 \\ 3 & -1 & -3 & -12 \end{array} \right) & \xrightarrow[R_3 \rightarrow R_3 - 3R_1]{R_2 \rightarrow R_2 - 2R_1} \left(\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & -1 & 0 & -3 \\ 0 & -7 & 0 & -21 \end{array} \right) & \xrightarrow{\text{usar } R_2 \text{ para eliminar la 2.ª columna}} \left(\begin{array}{ccc|c} 1 & 0 & -1 & -3 \\ 0 & -1 & 0 & -3 \\ 0 & 0 & 0 & 0 \end{array} \right) \\ & \xrightarrow{R_2 \rightarrow -R_2} \left(\begin{array}{ccc|c} 1 & 0 & -1 & -3 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right). \end{aligned}$$

Se sigue que $x_2 = 3$ y $x_1 = -3 + x_3$. Si tomamos x_3 como parámetro, la solución general del sistema en forma vectorial es

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{for } t \in \mathbb{R}.$$

Observe que el lado izquierdo del sistema (3.10') es el mismo que el del sistema homogéneo (3.10) en el Ejemplo 3.16, que tiene la solución general

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{for } t \in \mathbb{R}.$$

Esto muestra que, en efecto, obtenemos todas las soluciones de la ecuación inhomogénea como suma de la solución particular $(0, 3, 0)^t$ y todas las soluciones del sistema homogéneo correspondiente. ◇

Ahora debería haber entendido

- que una matriz $m \times n$ puede verse como una función que toma vectores en $\mathbb{R}^n$ y devuelve un vector en $\mathbb{R}^m$,
- la estructura del conjunto de todas las soluciones de un sistema lineal dado,
- etc.

Ahora debería ser capaz de

- calcular expresiones como $A\vec{x}$,
- relacionar las soluciones de un sistema inhomogéneo con las del sistema homogéneo correspondiente,
- etc.

Ejercicios.

1. Sea $A = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 1 & 3 & 0 & 2 \\ -1 & 2 & 1 & 3 \end{pmatrix}$. Para cada uno de los siguientes vectores, verifique si es una solución del sistema homogéneo $A\vec{x} = \vec{0}$:

(a) $\begin{pmatrix} 5 \\ -3 \\ 5 \\ 2 \end{pmatrix}$, (b) $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$, (c) $\begin{pmatrix} 10 \\ -6 \\ 10 \\ 20 \end{pmatrix}$, (d) $\begin{pmatrix} 3 \\ 0 \\ 0 \\ -1 \end{pmatrix}$, (e) $\begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix}$.

Posteriormente, encuentre todas las soluciones del sistema homogéneo. ¿Existe una solución del sistema anterior tal que alguna de sus componentes sea cero pero no sea la solución trivial?

2. En cada ítem, escriba el sistema de ecuaciones lineales correspondiente y obtenga todas sus soluciones:

(a) $\begin{pmatrix} 1 & 1 & 3 & 2 \\ 2 & -1 & 0 & 4 \\ 0 & 3 & 6 & 0 \end{pmatrix} \vec{x} = \begin{pmatrix} 7 \\ 8 \\ 8 \end{pmatrix}$

(b) $\begin{pmatrix} 2 & -1 & 1 \\ 3 & 1 & -2 \\ 1 & -1 & 1 \\ 1 & 5 & 7 \\ 1 & -7 & -5 \end{pmatrix} \vec{x} = \begin{pmatrix} 3 \\ -2 \\ 7 \\ 13 \\ 12 \end{pmatrix}$

(c) $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{pmatrix} \vec{x} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$

(d) $\begin{pmatrix} 3 & 1 & 0 & 2 \\ 2 & 1 & -1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \vec{x} = \begin{pmatrix} 6 \\ 3 \\ 6 \end{pmatrix}$

3. Escriba el siguiente sistema de ecuaciones de la forma $A\vec{x} = \vec{b}$:

$$\begin{aligned} x + 3y + z &= 3 \\ 2x + 7y + 2z &= 5 \\ 2x + 6y + (a^2 - 2)z &= a + 4. \end{aligned}$$

Luego, determine todos los $a \in \mathbb{R}$ tal que el sistema

- (a) tenga única solución,
- (b) tenga infinitas soluciones ó
- (c) no tenga solución.

3.4. Matrices como funciones de $\mathbb{R}^n$ a $\mathbb{R}^m$; composición de matrices

En la sección anterior vimos que una matriz $A \in M(m \times n)$ toma un vector $\vec{x} \in \mathbb{R}^n$ y devuelve un vector $A\vec{x}$ en $\mathbb{R}^m$. Esto nos permite ver A como una función de $\mathbb{R}^n$ a $\mathbb{R}^m$, y por lo tanto podemos definir la suma y la composición de dos matrices. Antes de hacer esto, veamos algunos ejemplos de tales matrices. Como ejemplos trabajamos con matrices 2×2 porque su acción sobre $\mathbb{R}^2$ se puede dibujar en el plano.

Ejemplo 3.25. Consideremos $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Esto define una función T_A de $\mathbb{R}^2$ a $\mathbb{R}^2$ mediante

$$T_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad T_A \vec{x} = A\vec{x}.$$

Observación. Escribimos T_A para denotar la función inducida por A , pero a veces escribiremos simplemente $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ cuando esté claro que consideramos la matriz A como una función.

Calculamos fácilmente

$$T_A \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad T_A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \text{en general} \quad T_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}.$$

Así vemos que T_A representa la reflexión de un vector $\vec{x}$ respecto al eje x .

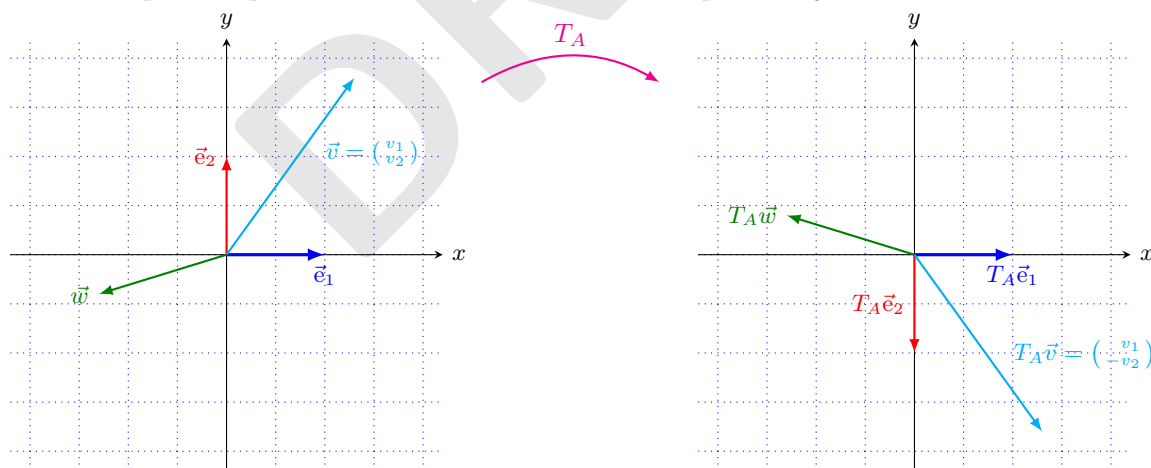


FIGURA 3.1: Reflexión respecto al eje x .

Ejemplo 3.26. Consideremos $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Esto define una función T_B de $\mathbb{R}^2$ a $\mathbb{R}^2$ mediante

$$T_B : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad T_B \vec{x} = B\vec{x}.$$

Calculamos fácilmente

$$T_B \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad T_B \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{en general} \quad T_B \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}.$$

Así vemos que T_B representa la proyección de un vector $\vec{x}$ sobre el eje y .

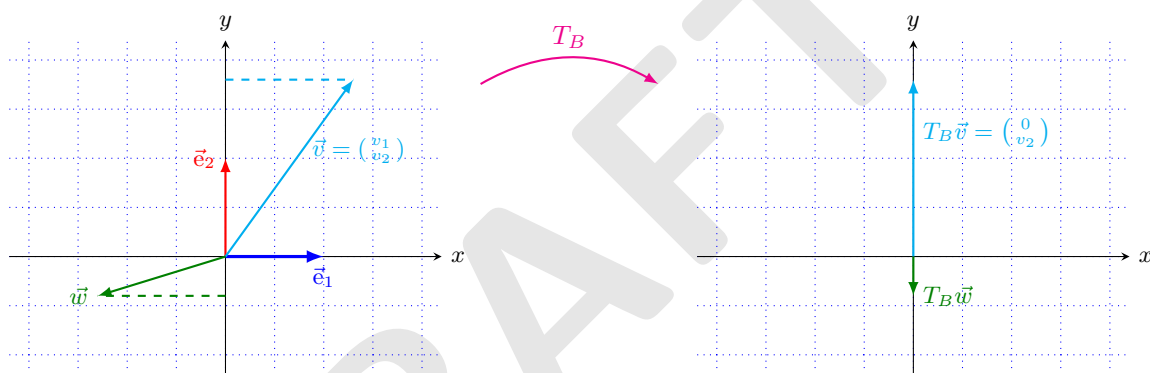


FIGURA 3.2: Proyección ortogonal sobre el eje y .

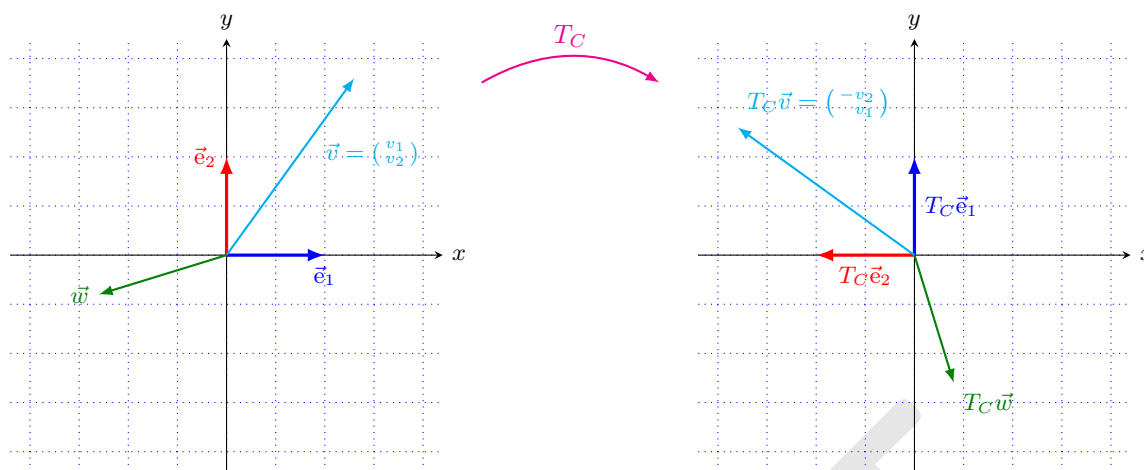
Ejemplo 3.27. Consideremos $C = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Esto define una función T_C de $\mathbb{R}^2$ a $\mathbb{R}^2$ mediante

$$T_C : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad T_C \vec{x} = C\vec{x}.$$

Calculamos fácilmente

$$T_C \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad T_C \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \text{en general} \quad T_C \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}.$$

Así vemos que T_C representa la rotación de un vector $\vec{x}$ en 90° en sentido antihorario.

FIGURA 3.3: Rotación de $\pi/2$ en sentido antihorario.

Al igual que con otras funciones, podemos sumarlas o componerlas. Recuerde de sus clases de cálculo que las funciones se suman “punto a punto”. Esto significa que, si tenemos dos funciones $f, g : \mathbb{R} \rightarrow \mathbb{R}$, entonces la *suma* $f + g$ es una nueva función que se define por

$$f + g : \mathbb{R} \rightarrow \mathbb{R}, \quad (f + g)(x) = f(x) + g(x). \quad (3.17)$$

La multiplicación de una función f por un número c da la nueva función cf definida por

$$cf : \mathbb{R} \rightarrow \mathbb{R}, \quad (cf)(x) = c(f(x)). \quad (3.18)$$

La composición de funciones se define como

$$f \circ g : \mathbb{R} \rightarrow \mathbb{R}, \quad (f \circ g)(x) = f(g(x)). \quad (3.19)$$

Suma de matrices

Veamos cómo se ve esto en el caso de las matrices. Sean A y B matrices. Primero observe que ambas deben partir del mismo espacio $\mathbb{R}^n$ porque queremos aplicarlas al mismo $\vec{x}$, es decir, tanto $A\vec{x}$ como $B\vec{x}$ deben estar definidos. Por lo tanto, A y B deben tener el mismo número de columnas. También deben tener el mismo número de filas porque queremos poder sumar $A\vec{x}$ y $B\vec{x}$. Así, sean $A, B \in M(m \times n)$ y sea $\vec{x} \in \mathbb{R}^n$. Entonces, por la definición de la composición de dos funciones,

obtenemos

$$\begin{aligned}
 (A + B)\vec{x} &:= A\vec{x} + B\vec{x} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\
 &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} + \begin{pmatrix} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n \\ \vdots \\ b_{m1}x_1 + b_{m2}x_2 + \cdots + b_{mn}x_n \end{pmatrix} \\
 &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + b_{m1}x_1 + b_{m2}x_2 + \cdots + b_{mn}x_n \end{pmatrix} \\
 &= \begin{pmatrix} (a_{11} + b_{11})x_1 + (a_{12} + b_{12})x_2 + \cdots + (a_{1n} + b_{1n})x_n \\ (a_{21} + b_{21})x_1 + (a_{22} + b_{22})x_2 + \cdots + (a_{2n} + b_{2n})x_n \\ \vdots \\ (a_{m1} + b_{m1})x_1 + (a_{m2} + b_{m2})x_2 + \cdots + (a_{mn} + b_{mn})x_n \end{pmatrix} \\
 &= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.
 \end{aligned}$$

Vemos que $A + B$ es de nuevo una matriz del mismo tamaño y que las componentes de esta nueva matriz son simplemente la suma de las componentes correspondientes de las matrices A y B .

Multiplicación de una matriz por un escalar

Ahora sea c un número y sea $A \in M(m \times n)$. Entonces we have

$$\begin{aligned}
 (cA)\vec{x} &= c(A\vec{x}) = c \left[\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right] = c \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} \\
 &= \begin{pmatrix} ca_{11}x_1 + \cdots + ca_{1n}x_n \\ \vdots \\ ca_{m1}x_1 + \cdots + ca_{mn}x_n \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ \vdots & \vdots & & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.
 \end{aligned}$$

Vemos que cA es de nuevo una matriz y que las componentes de esta nueva matriz son simplemente el producto de las componentes correspondientes de la matriz A por c .

Proposición 3.28. Sean $A, B, C \in M(m \times n)$, sea $\mathbb{O}$ la matriz cuyas entradas son todas 0 y sean $\lambda, \mu \in \mathbb{R}$. Además, sea $\tilde{A}$ la matriz cuyas entradas son las entradas negativas de A . Entonces tenemos lo siguiente.

- (I) **Asociatividad de la suma de matrices:** $(A + B) + C = A + (B + C)$.
- (II) **Conmutatividad de la suma de matrices:** $A + B = B + A$.
- (III) **Elemento neutro aditivo:** $A + \mathbb{O} = A$.
- (IV) **Inverso aditivo** $A + \tilde{A} = \mathbb{O}$.
- (v) $1A = A$.
- (VI) $(\lambda + \mu)A = \lambda A + \mu A$ y $\lambda(A + B) = \lambda A + \lambda B$.
- (VII) $(\lambda\mu)A = \lambda(\mu A)$.

Demostración. Las afirmaciones de la proposición se pueden demostrar mediante cálculos directos. $\square$

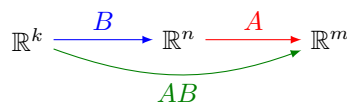
Demuestre la Proposición 3.28.

De la proposición obtenemos inmediatamente el siguiente teorema.

Teorema 3.29. $M(m \times n)$ es un espacio vectorial.

Composición de dos matrices

Ahora calculemos la composición de dos matrices. También se llama el *producto de matrices*. Supongamos que tenemos $A \in M(m \times n)$ y queremos calcular AB para alguna matriz B . Observe que A describe una función de $\mathbb{R}^n \rightarrow \mathbb{R}^m$. Para que AB tenga sentido, necesitamos que B vaya de algún $\mathbb{R}^k$ a $\mathbb{R}^n$, lo cual significa que $B \in M(n \times k)$. La función resultante AB será entonces una aplicación de $\mathbb{R}^k$ a $\mathbb{R}^m$.



Así, sea $B \in M(n \times k)$. Entonces, por la definición de la composición de dos funciones, obtenemos,

para todo $\vec{x} \in \mathbb{R}^k$,

$$\begin{aligned}
 (AB)\vec{x} &= A(B\vec{x}) = A \left[\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nk} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} \right] = A \begin{pmatrix} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1k}x_k \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2k}x_k \\ \vdots \\ b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nk}x_k \end{pmatrix} \\
 &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{pmatrix} \begin{pmatrix} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1k}x_k \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2k}x_k \\ \vdots \\ b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nk}x_k \end{pmatrix} \\
 &= \begin{pmatrix} a_{11}[b_{11}x_1 + b_{12}x_2 + \cdots + b_{1k}x_k] + a_{12}[b_{21}x_1 + b_{22}x_2 + \cdots + b_{2k}x_k] + \cdots + a_{1n}[b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nk}x_k] \\ a_{21}[b_{11}x_1 + b_{12}x_2 + \cdots + b_{1k}x_k] + a_{22}[b_{21}x_1 + b_{22}x_2 + \cdots + b_{2k}x_k] + \cdots + a_{2n}[b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nk}x_k] \\ \vdots \\ a_{m1}[b_{11}x_1 + b_{12}x_2 + \cdots + b_{1k}x_k] + a_{m2}[b_{21}x_1 + b_{22}x_2 + \cdots + b_{2k}x_k] + \cdots + a_{mn}[b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nk}x_k] \end{pmatrix} \\
 &= \begin{pmatrix} [a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1}]x_1 + [a_{11}b_{12} + a_{12}b_{22} + \cdots + a_{1n}b_{n2}]x_2 + \cdots + [a_{11}b_{1k} + a_{12}b_{2k} + \cdots + a_{1n}b_{nk}]x_k \\ [a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1}]x_1 + [a_{21}b_{12} + a_{22}b_{22} + \cdots + a_{2n}b_{n2}]x_2 + \cdots + [a_{21}b_{1k} + a_{22}b_{2k} + \cdots + a_{2n}b_{nk}]x_k \\ \vdots \\ [a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1}]x_1 + [a_{m1}b_{12} + a_{m2}b_{22} + \cdots + a_{mn}b_{n2}]x_2 + \cdots + [a_{m1}b_{1k} + a_{m2}b_{2k} + \cdots + a_{mn}b_{nk}]x_k \end{pmatrix} \\
 &= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} & a_{11}b_{12} + a_{12}b_{22} + \cdots + a_{1n}b_{n2} & \cdots & a_{11}b_{1k} + a_{12}b_{2k} + \cdots + a_{1n}b_{nk} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} & a_{21}b_{12} + a_{22}b_{22} + \cdots + a_{2n}b_{n2} & \cdots & a_{21}b_{1k} + a_{22}b_{2k} + \cdots + a_{2n}b_{nk} \\ \vdots & \vdots & & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1} & a_{m1}b_{12} + a_{m2}b_{22} + \cdots + a_{mn}b_{n2} & \cdots & a_{m1}b_{1k} + a_{m2}b_{2k} + \cdots + a_{mn}b_{nk} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix}.
 \end{aligned}$$

Vemos que AB es una matriz de tamaño $m \times k$, como era de esperarse, ya que la función composición va de $\mathbb{R}^k$ a $\mathbb{R}^m$. La componente $c_{j\ell}$ de la nueva matriz (la entrada en la fila j y la columna ℓ) es

$$c_{j\ell} = \sum_{r=1}^n a_{jr}b_{r\ell}.$$

Así, para calcular esta entrada necesitamos de A solo su fila j -ésima y de B solo su columna ℓ -ésima, y las multiplicamos componente por componente. Puede memorizar esto de nuevo como “fila por columna”, más precisamente:

$$c_{j\ell} = \text{componente en la fila } j \text{ y la columna } \ell \text{ de } AB = (\text{fila } j \text{ de } A) \times (\text{columna } \ell \text{ de } B) \quad (3.20)$$

como en el caso de la multiplicación de un vector por una matriz. De hecho, un vector en $\mathbb{R}^n$ puede verse como una matriz $n \times 1$ (una matriz con n filas y una columna); por lo tanto, (3.13) puede

verse como un caso especial de (3.20).

$$AB = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1\ell} & \dots & b_{1k} \\ b_{21} & \dots & b_{2\ell} & \dots & b_{2k} \\ \vdots & \dots & \vdots & & \vdots \\ \vdots & \dots & \vdots & & \vdots \\ b_{n1} & \dots & b_{n\ell} & \dots & b_{nk} \end{pmatrix} = \begin{pmatrix} c_{11} & \dots & c_{1\ell} & \dots & c_{1k} \\ \vdots & & \vdots & & \vdots \\ c_{j1} & \dots & c_{j\ell} & \dots & c_{jk} \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \dots & c_{m\ell} & \dots & c_{mk} \end{pmatrix}$$

con $c_{j\ell} = a_{j1}b_{1\ell} + a_{j2}b_{2\ell} + \dots + a_{jn}b_{n\ell}$.

Ejemplo 3.30. Sea $A = \begin{pmatrix} 1 & 2 & 3 \\ 8 & 6 & 4 \end{pmatrix}$ y $B = \begin{pmatrix} 7 & 1 & 2 & 3 \\ 2 & 0 & 1 & 4 \\ 2 & 6 & -3 & 0 \end{pmatrix}$. Entonces

$$\begin{aligned} AB &= \begin{pmatrix} 1 & 2 & 3 \\ 8 & 6 & 4 \end{pmatrix} \begin{pmatrix} 7 & 1 & 2 & 3 \\ 2 & 0 & 1 & 4 \\ 2 & 6 & -3 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot 7 + 2 \cdot 2 + 3 \cdot 2 & 1 \cdot 1 + 2 \cdot 0 + 3 \cdot 6 & 1 \cdot 2 + 2 \cdot 1 + 3 \cdot (-3) & 1 \cdot 3 + 2 \cdot 4 + 3 \cdot 0 \\ 8 \cdot 7 + 6 \cdot 2 + 4 \cdot 2 & 8 \cdot 1 + 6 \cdot 0 + 4 \cdot 6 & 8 \cdot 2 + 6 \cdot 1 + 4 \cdot (-3) & 8 \cdot 3 + 6 \cdot 4 + 4 \cdot 0 \end{pmatrix} \\ &= \begin{pmatrix} 17 & 19 & -5 & 11 \\ 76 & 32 & 10 & 48 \end{pmatrix}. \end{aligned}$$

Veamos algunas propiedades de las operaciones algebraicas para matrices que acabamos de introducir.

Proposición 3.31. Sean $A \in M(m \times n)$, $B, C \in M(k \times m)$, $S, T \in M(n \times k)$ y $R \in M(k \times \ell)$. Entonces the following is true.

- (I) **Asociatividad del producto de matrices:** $A(RS) = (AR)S$.
- (II) **Distributividad:** $A(S + T) = AS + AT$ y $(B + C)A = BA + CA$.

Demostración. Las afirmaciones de la proposición se pueden demostrar mediante cálculos directos. $\square$

Demuestre la Proposición 3.31.

Observación muy importante.

La multiplicación de matrices no es conmutativa, es decir, en general

$$AB \neq BA.$$

Que la multiplicación de matrices no sea conmutativa es de esperarse, ya que es la composición de dos funciones (piense en funciones que conoce de sus clases de cálculo. Por ejemplo, sí hace diferencia si primero eleva una variable al cuadrado y luego toma el arctan, o si primero calcula su arctan y luego eleva el resultado al cuadrado).

Veamos un ejemplo. Sea B la matriz del Ejemplo 3.26 y sea C la matriz del Ejemplo 3.27. Recordemos que B representa la proyección ortogonal sobre el eje y y que C representa la rotación antihoraria de 90° . Si tomamos $\vec{e}_1$ (el vector unitario en la dirección x), y primero rotamos y luego proyectamos, obtenemos el vector $\vec{e}_2$. Si, en cambio, primero proyectamos y luego rotamos, obtenemos $\vec{0}$. Esto significa que $BC\vec{e}_1 \neq CB\vec{e}_1$, por lo tanto $BC \neq CB$. Calculemos los productos:

$$BC = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{primero rotación, luego proyección,}$$

$$CB = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \quad \text{primero proyección, luego rotación.}$$

Sea A la matriz del Ejemplo 3.25, sea B la matriz del Ejemplo 3.26 y sea C la matriz del Ejemplo 3.27. Verifique que $AB \neq BA$ y $AC \neq CA$, y comprenda este resultado geoméricamente siguiendo, por ejemplo, a dónde se envían los vectores unitarios.

Observe también que usualmente, cuando AB está definido, la expresión BA no está definida, porque en general el número de columnas de B será diferente del número de filas de A .

Terminamos esta sección con la definición de la llamada matriz identidad.

Definición 3.32. Sea $n \in \mathbb{N}$. Entonces la matriz identidad $n \times n$ es la matriz que tiene unos en su diagonal y ceros en todas las demás entradas:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & \cdots & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}. \quad (3.21)$$

Como notación para la matriz identidad, en la literatura se usan los siguientes símbolos: E_n , id_n , Id_n , I_n , $\mathbf{1}_n$, $\mathbb{1}_n$. El subíndice n puede omitirse si el tamaño de la matriz es claro.

Observación 3.33. Se puede verificar fácilmente que

$$A \text{id}_n = A, \quad \text{id}_n B = B, \quad \text{id}_n \vec{x} = \vec{x}$$

para todo $A \in M(m \times n)$, para todo $B \in M(n \times k)$ y para todo $\vec{x} \in \mathbb{R}^n$.

Ahora debería haber entendido

- qué son la suma y la composición de dos matrices y de dónde vienen las fórmulas,
- por qué la composición de matrices no es conmutativa,
- que $M(m \times n)$ es un espacio vectorial,
- etc.

Ahora debería ser capaz de

- calcular la suma y el producto (composición) de dos matrices,
- etc.

Ejercicios.

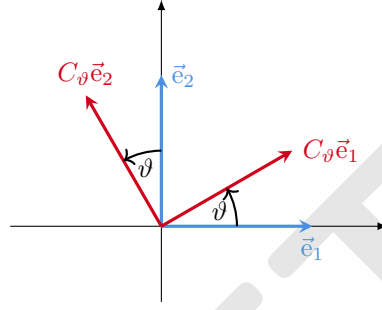
1. Para $A = \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ -1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} -2 & 0 \\ 1 & 4 \\ -7 & 5 \end{pmatrix}$ y $C = \begin{pmatrix} -1 & 1 \\ 4 & 6 \\ -7 & 3 \end{pmatrix}$, calcular:
 - (a) $2A$,
 - (b) $3C - 2A$,
 - (c) $A + B + C$,
 - (d) $2A - 3B + 5C$,
 - (e) una matriz D tal que $A + 2B - 3C + D$ es la matriz de solo ceros.
2. Realice los siguientes cálculos (antes de hacer la multiplicación indicada, especifique cuál será el tamaño de la matriz resultante al hacer el producto):
 - (a) $\begin{pmatrix} 2 & -3 & 5 \\ 1 & 0 & 6 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 & 6 \\ -2 & 3 & 5 \\ 1 & 0 & 4 \end{pmatrix}$,
 - (b) $\begin{pmatrix} 1 & 4 & -2 \\ 3 & 0 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$,
 - (c) $(1 \ 4 \ 0 \ 2) \begin{pmatrix} 3 & -6 \\ 2 & 4 \\ 1 & 0 \\ -2 & 3 \end{pmatrix}$,
 - (d) $\begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix} (0 \ -2 \ 5)$,
 - (e) $\begin{pmatrix} 3 & 2 & 1 & -2 \\ -6 & 4 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 4 & 3 \\ 0 & 5 \\ 2 & 0 \end{pmatrix}$,
 - (f) $\begin{pmatrix} 1 & 6 \\ 0 & 4 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 7 & 1 & 4 \\ 2 & -3 & 5 \end{pmatrix}$.
3. Verifique la ley asociativa de la multiplicación para las matrices

$$A = \begin{pmatrix} 3 & -1 & 4 \\ 1 & 0 & -1 \end{pmatrix}, B = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 0 & -1 \\ -3 & -2 & 0 \end{pmatrix} \text{ y } C = \begin{pmatrix} 1 & 6 \\ -1 & 4 \\ -2 & 3 \end{pmatrix}.$$
4. Encuentre $A \in M(2 \times 2)$ tal que $A \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} = \text{id}_2$.

5. Sean $A = \begin{pmatrix} 3 & -1 & -5 & 2 \\ 4 & 2 & 1 & 0 \end{pmatrix}$ y $C = \begin{pmatrix} 6 & 10 \\ -3 & 2 \end{pmatrix}$. Encuentre por lo menos una matriz B tal que $AB = C$. ¿Cuántas soluciones a la ecuación matricial $AX = C$ existen?

6. En $\mathbb{R}^2$, encuentre la matriz (2×2) que rota el plano cierto ángulo ϑ .

Indicación. Plantee la matriz $C_\vartheta = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ y calcule $C_\vartheta \vec{e}_1$ y $C_\vartheta \vec{e}_2$ para encontrar a, b, c, d .



7. Si $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, encuentre condiciones sobre a, b, c, d tal que $AB = BA$.
8. Sean $A, B \in M(n \times n)$:
- ¿Se cumple la igualdad $A^2 - B^2 = (A - B)(A + B)$? Si su respuesta es negativa, ¿cuáles condiciones sobre A, B puede dar para que se cumpla la igualdad?
 - Lo mismo del inciso anterior para la igualdad $(A + B)^2 = A^2 + 2AB + B^2$.

3.5. Inversas de matrices

Daremos dos motivaciones de por qué nos interesan las inversas de matrices antes de dar la definición formal.

Inversa de una matriz como función

La inversa de una matriz dada es una matriz que “deshace” lo que hizo la matriz original. Revisaremos las matrices de los Ejemplos 3.25, 3.26 y 3.27.

- Supongamos que tenemos la matriz A del Ejemplo 3.25, que representa la reflexión respecto al eje x , y queremos encontrar una matriz que restaure un vector después de haberle aplicado A . Claramente, tenemos que reflejar de nuevo respecto al eje x : reflejar un vector arbitrario $\vec{x}$ dos veces respecto al eje x deja el vector donde estaba. Verifiquemos:

$$AA = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \text{id}_2.$$

Esto significa que para todo $\vec{x} \in \mathbb{R}^2$ tenemos que $A^2 \vec{x} = \vec{x}$; por lo tanto, A es su propia inversa.

- Supongamos que tenemos la matriz C del Ejemplo 3.27, que representa la rotación antihoraria de 90° , y queremos encontrar una matriz que restaure un vector después de haberle aplicado C . Claramente, tenemos que rotar 90° en sentido horario. Supongamos que existe una matriz que representa esta rotación y llamémosla C_{-90° . Por la Observación 3.19, basta saber cómo actúa sobre $\vec{e}_1$ y $\vec{e}_2$ para poder escribirla. Claramente $C_{-90^\circ}\vec{e}_1 = -\vec{e}_2$ y $C_{-90^\circ}\vec{e}_2 = \vec{e}_1$, luego $C_{-90^\circ} = (-\vec{e}_2|\vec{e}_1)$. Verifiquemos:

$$C_{-90^\circ}C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \text{id}_2$$

y

$$CC_{-90^\circ} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \text{id}_2$$

que era de esperarse, porque rotar primero 90° en sentido horario y luego 90° en sentido antihorario deja cualquier vector donde estaba.

- Supongamos que tenemos la matriz B del Ejemplo 3.26, que representa la proyección sobre el eje y . En este caso, no podemos restaurar un vector $\vec{x}$ después de proyectarlo sobre el eje y . Por ejemplo, si sabemos que $B\vec{x} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$, entonces $\vec{x}$ pudo haber sido $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ o $\begin{pmatrix} 7 \\ 2 \end{pmatrix}$ o cualquier otro vector en $\mathbb{R}^2$ cuya segunda componente sea igual a 2. Esto muestra que B no tiene inversa.

Inversa de una matriz para resolver un sistema lineal

Consideremos la siguiente situación. Una tienda de víveres vende dos paquetes diferentes de frutas. El tipo A contiene 1 durazno y 3 mangos, y el tipo B contiene 2 duraznos y 1 mango. Podemos hacer dos tipos diferentes de preguntas:

- Dado cierto número de paquetes de tipo A y de tipo B, ¿cuántos duraznos y cuántos mangos obtenemos?
- ¿Cuántos paquetes de cada tipo necesitamos para obtener un número dado de duraznos y mangos?

La primera pregunta es bastante fácil de responder. Escribamos la información que se nos da. Si

$$\begin{aligned} a &= \text{número de paquetes de tipo A}, & p &= \text{número de duraznos} \\ b &= \text{número de paquetes de tipo B}, & m &= \text{número de mangos.} \end{aligned}$$

entonces

$$\begin{aligned} p &= 1a + 2b \\ m &= 3a + 1b. \end{aligned} \tag{3.22}$$

Usando vectores y matrices, podemos reescribir esto como

$$\begin{pmatrix} p \\ m \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

Sea $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$. Entonces lo anterior se convierte simplemente en

$$\begin{pmatrix} p \\ m \end{pmatrix} = A \begin{pmatrix} a \\ b \end{pmatrix}. \tag{3.23}$$

Si conocemos a y b (es decir, sabemos cuántos paquetes de cada tipo compramos), entonces podemos encontrar los valores de p y m simplemente evaluando $A\begin{pmatrix} a \\ b \end{pmatrix}$, lo cual es relativamente fácil.

Ejemplo 3.34. Supongamos que compramos 1 paquete de tipo A y 3 paquetes de tipo B; entonces calculamos

$$\begin{pmatrix} p \\ m \end{pmatrix} = A \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \end{pmatrix},$$

lo cual muestra que tenemos 9 duraznos y 7 mangos.

Si, por otro lado, conocemos p y m y se nos pide encontrar a y b tales que se cumpla (3.22), tenemos que resolver un sistema lineal, lo cual es mucho más engorroso. Por supuesto, podemos resolver (3.23) usando el proceso de eliminación de Gauß o de Gauß-Jordan, pero si se nos pidiera hacer esto para varias parejas p y m , entonces se volvería largo rápidamente. Sin embargo, si tuviéramos una matriz A' tal que $A'A = \text{id}_2$, entonces esta tarea sería bastante fácil, ya que en este caso podríamos manipular (3.23) como sigue:

$$\begin{pmatrix} p \\ m \end{pmatrix} = A \begin{pmatrix} a \\ b \end{pmatrix} \implies A' \begin{pmatrix} p \\ m \end{pmatrix} = A'A \begin{pmatrix} a \\ b \end{pmatrix} = \text{id}_2 \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Si además supiéramos que $AA' = \text{id}_2$, entonces tenemos que

$$\begin{pmatrix} p \\ m \end{pmatrix} = A \begin{pmatrix} a \\ b \end{pmatrix} \iff A' \begin{pmatrix} p \\ m \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}. \quad (3.24)$$

La tarea de encontrar a y b se reduce de nuevo a realizar una multiplicación matricial. La matriz A' , si existe, se llama la *inversa de A*, y dedicaremos el resto de esta sección a dar criterios para su existencia, investigar sus propiedades y dar una receta para encontrarla.

Ejercicio. Verifique que $A' = \frac{1}{5} \begin{pmatrix} -1 & 2 \\ 3 & -1 \end{pmatrix}$ satisface $A'A = \text{id}_2$.

Ejemplo 3.35. Supongamos que queremos comprar 5 duraznos y 5 mangos. Entonces calculamos

$$\begin{pmatrix} a \\ b \end{pmatrix} = A' \begin{pmatrix} 5 \\ 5 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -1 & 2 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 5 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

lo cual muestra que tenemos que comprar 1 paquete de tipo A y 2 paquetes de tipo B.

Ahora demos la definición precisa de la inversa de una matriz.

Definición 3.36. Una matriz $A \in M(n \times n)$ se llama *invertible* si existe una matriz $A' \in M(n \times n)$ tal que

$$AA' = \text{id}_n \quad \text{and} \quad A'A = \text{id}_n.$$

En este caso A' se llama la *inversa de A* y se denota por A^{-1} . Si A no es invertible, entonces se llama *no invertible* o *singular*.

La razón por la cual en la definición solo admitimos matrices cuadradas (matrices con el mismo número de filas y columnas) se explica en la siguiente observación.

Observación 3.37. (I) Sea $A \in M(m \times n)$ y supongamos que hay una matriz B tal que $BA = \text{id}_n$. Esto significa que si para algún $\vec{b} \in \mathbb{R}^m$ la ecuación $A\vec{x} = \vec{b}$ tiene una solución, entonces esta es única, porque

$$A\vec{x} = \vec{b} \implies BA\vec{x} = B\vec{b} \implies \vec{x} = B\vec{b}.$$

De lo anterior queda claro que $A \in M(m \times n)$ puede tener inversa solo si para todo $\vec{b} \in \mathbb{R}^m$ la ecuación $A\vec{x} = \vec{b}$ tiene *a lo sumo una solución*. Sabemos que si A tiene más columnas que filas, entonces el número de columnas será mayor que el número de pivotes. Por lo tanto, $A\vec{x} = \vec{b}$ tiene o bien ninguna solución o infinitas soluciones (véase el Teorema 3.7). Luego una matriz A con más columnas que filas no puede tener inversa.

(II) De nuevo, sea $A \in M(m \times n)$ y supongamos que hay una matriz B tal que $AB = \text{id}_m$. Esto significa que para todo $\vec{b} \in \mathbb{R}^m$ la ecuación $A\vec{x} = \vec{b}$ es resuelta por $\vec{x} = B\vec{b}$, porque

$$\text{id}_m \vec{b} = \vec{b} \implies AB\vec{b} = \vec{b} \implies A(B\vec{b}) = \vec{b}.$$

De lo anterior queda claro que $A \in M(m \times n)$ puede tener inversa solo si para *todo* $\vec{b} \in \mathbb{R}^m$ la ecuación $A\vec{x} = \vec{b}$ tiene *al menos una solución*. Supongamos que A tiene más filas que columnas. Si aplicamos eliminación de Gauss a la matriz aumentada $(A|\vec{b})$, entonces la última fila de la forma escalonada por filas tiene que ser $(0 \ \dots \ 0 | \beta_m)$. Si escogemos $\vec{b}$ de tal manera que después de la reducción $\beta_m \neq 0$, entonces $A\vec{x} = \vec{b}$ no tiene solución. Tal $\vec{b}$ es fácil de encontrar: solo necesitamos tomar $\vec{e}_m$ (el m -ésimo vector unitario) y hacer los pasos de la eliminación de Gauss hacia atrás. Si tomamos este vector como lado derecho de nuestro sistema, entonces la última fila después de la reducción será $(0 \ \dots \ 0 | 1)$. Por lo tanto, una matriz A con más filas que columnas no puede tener inversa, porque siempre habrá algún $\vec{b}$ tal que la ecuación $A\vec{x} = \vec{b}$ no tenga solución.

En conclusión, mostramos que debemos tener $m = n$ si A debe tener una matriz inversa.

Si $A \in M(m \times n)$ con $n \neq m$, entonces no tiene sentido hablar de una inversa de A , como se explicó arriba. Sin embargo, podemos definir la inversa izquierda y la inversa derecha.

Definición 3.38. Sea $A \in M(m \times n)$.

- (I) Una matriz C se llama una *inversa izquierda* de A si $CA = \text{id}_n$.
- (II) Una matriz D se llama una *inversa derecha* de A si $AD = \text{id}_m$.

Observe que C y D deben ser matrices $n \times m$. Los siguientes ejemplos muestran que las inversas izquierdas y derechas no necesariamente existen, y si existen, no son únicas.

Ejemplos 3.39. (I) $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ no tiene inversa izquierda ni inversa derecha.

(II) $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ no tiene inversa izquierda y tiene inversa derecha $D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$. De hecho,

para todo $x, y \in \mathbb{R}$ la matriz $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ x & y \end{pmatrix}$ es una inversa derecha de A .

(III) $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ no tiene inversa derecha y tiene inversa izquierda $C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. De hecho,

para todo $x, y \in \mathbb{R}$ la matriz $\begin{pmatrix} 1 & 0 & x \\ 0 & 1 & y \end{pmatrix}$ es una inversa izquierda de A .

Observación 3.40. Mostraremos en el Teorema 3.45 que una matriz $A \in M(n \times n)$ es invertible si y solo si tiene una inversa izquierda y una inversa derecha.

Ejemplos 3.41. ■ De los ejemplos al comienzo de esta sección tenemos:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \implies A^{-1} = A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$C = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \implies C^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

$$B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \implies B \text{ no es invertible.}$$

■ Sea $A = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$. Entonces podemos adivinar fácilmente que $A^{-1} = \begin{pmatrix} 1/4 & 0 & 0 & 0 \\ 0 & 1/5 & 0 & 0 \\ 0 & 0 & -1/3 & 0 \\ 0 & 0 & 0 & 1/2 \end{pmatrix}$

es una inversa de A . Es fácil verificar que el producto de estas matrices da id_4 .

- Sea $A \in M(n \times n)$ y supongamos que la k -ésima fila de A consiste solo de ceros. Entonces A no es invertible porque para cualquier matriz $B \in M(n \times n)$, la k -ésima fila de la matriz producto AB será cero, sin importar cómo escojamos B . Así que no hay una matriz B tal que $AB = \text{id}_n$.
- Sea $A \in M(n \times n)$ y supongamos que la k -ésima columna de A consiste solo de ceros. Entonces A no es invertible porque para cualquier matriz $B \in M(n \times n)$, la k -ésima columna de la matriz producto BA será cero, sin importar cómo escojamos B . Así que no hay una matriz B tal que $BA = \text{id}_n$.

Ahora demostremos algunos teoremas sobre matrices inversas. Recordemos que $A \in M(n \times n)$ es invertible si y solo si existe una matriz $A' \in M(n \times n)$ tal que $AA' = A'A = \text{id}_n$. Primero mostraremos que la matriz inversa, si existe, es única.

Teorema 3.42. Sean $A, B \in M(n \times n)$.

- (I) Si A es invertible, entonces su inversa es única.
- (II) Si A es invertible, entonces su inversa A^{-1} es invertible y su inversa es A .
- (III) Si A y B son invertibles, entonces su producto AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$.

Demostración. (I) Supongamos que A es invertible y que A' y A'' son inversas de A . Observe que esto significa que

$$AA' = A'A = \text{id}_n \quad \text{and} \quad AA'' = A''A = \text{id}_n. \quad (3.25)$$

Tenemos que mostrar que $A' = A''$. Esto se sigue de (3.25) y de la asociatividad de la multiplicación de matrices, porque

$$A' = A' \text{id}_n = A'(AA'') = (A'A)A'' = \text{id}_n A'' = A''.$$

- (II) Supongamos que A es invertible y sea A^{-1} su inversa. Para mostrar que A^{-1} es invertible, necesitamos una matriz C tal que $CA^{-1} = A^{-1}C = \text{id}_n$. Esta matriz C es entonces la inversa de A^{-1} . Claramente, $C = A$ sirve. Por lo tanto, A^{-1} es invertible y $(A^{-1})^{-1} = A$.
- (III) Supongamos que A y B son invertibles. Para mostrar que AB es invertible y que $(AB)^{-1} = B^{-1}A^{-1}$, solo tenemos que verificar que $B^{-1}A^{-1}(AB) = (AB)B^{-1}A^{-1} = \text{id}_n$. Vemos que esto es cierto usando la asociatividad del producto de matrices:

$$\begin{aligned} B^{-1}A^{-1}(AB) &= B^{-1}(A^{-1}A)B = B^{-1}\text{id}_n B = B^{-1}B = \text{id}_n, \\ (AB)B^{-1}A^{-1} &= A(BB^{-1})A^{-1} = A\text{id}_n A^{-1} = A^{-1}A = \text{id}_n. \end{aligned} \quad \square$$

Observe que en la demostración adivinamos la fórmula para $(AB)^{-1}$ y luego verificamos que en efecto es la inversa de AB . También podemos calcularla como sigue. Supongamos que C es una inversa izquierda de AB . Entonces

$$C(AB) = \text{id}_n \iff (CA)B = \text{id}_n \iff CA = \text{id}_n B^{-1} = B^{-1} \iff C = B^{-1}A^{-1}.$$

Si D es una inversa derecha de AB , entonces

$$(AB)D = \text{id}_n \iff A(BD) = \text{id}_n \iff BD = A^{-1}\text{id}_n = A^{-1} \iff D = B^{-1}A^{-1}.$$

Como $C = D$, esta es la inversa de AB .

Observación 3.43. En general, la suma de matrices invertibles no es invertible. Por ejemplo, tanto id_n como $-\text{id}_n$ son invertibles, pero su suma es la matriz cero, que no es invertible.

El Teorema 3.44 de la siguiente sección nos mostrará cómo encontrar la inversa de una matriz invertible; véase en particular la sección en la página 116.

Ahora debería haber entendido

- qué significa la invertibilidad de una matriz y por qué no tiene sentido hablar de la inverti-

bilidad de una matriz que no es cuadrada,

- que la invertibilidad de una matriz $n \times n$ es equivalente al hecho de que para todo $\vec{b} \in \mathbb{R}^m$ el sistema lineal asociado $A\vec{x} = \vec{b}$ tiene exactamente una solución.
- etc.

Ahora debería ser capaz de

- adivinar la inversa de matrices invertibles simples, por ejemplo de matrices que tienen una interpretación geométrica clara, o de matrices diagonales,
- verificar si dos matrices dadas son inversas entre sí,
- dar ejemplos de matrices invertibles y no invertibles,
- etc.

Ejercicios.

1. Diga cuáles de las siguientes matrices son mutuamente inversas.

$$A = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1/3 & 1 \\ 1/5 & 1/2 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}, \quad D = \begin{pmatrix} -3 & -1 \\ -5 & -2 \end{pmatrix}, \quad E = \begin{pmatrix} -2 & 1 \\ 5 & -3 \end{pmatrix}.$$

2. Muestre que $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ -2 & -4 & 1 \end{pmatrix}$ es su propia inversa.

3. Verifique que $\begin{pmatrix} 4 & -3 \\ 0 & 0 \\ -1 & 1 \end{pmatrix}$ es una inversa a derecha de la matriz $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \end{pmatrix}$ y úsela para encontrar una solución particular al sistema $A\vec{x} = \vec{b}$.

4. Sea $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$. ¿Es A invertible? *Indicación.* Suponga que $AB = \text{id}_2$ para alguna matriz $B \in M(2 \times 2)$. ¿Cómo se ve la primera columna si uno evalúa el producto AB ?

3.6. Matrices y sistemas lineales

Recordemos el Teorema 3.7:

Para $A \in M(m \times n)$ y $\vec{b} \in \mathbb{R}^m$, considere la ecuación

$$A\vec{x} = \vec{b}. \quad (3.26)$$

Entonces se cumple lo siguiente:

- | | | | |
|-----|---|--------|--|
| (1) | La ecuación (3.26) no tiene solución. | $\iff$ | La forma escalonada reducida por filas del sistema aumentado $(A \vec{b})$ tiene una fila de la forma $(0 \cdots 0 \beta)$ con algún $\beta \neq 0$. |
| (2) | La ecuación (3.26) tiene al menos una solución. | $\iff$ | La forma escalonada reducida por filas del sistema aumentado $(A \vec{b})$ no tiene ninguna fila de la forma $(0 \cdots 0 \beta)$ con algún $\beta \neq 0$. |

En el caso (2), tenemos los siguientes dos subcasos:

- | | | | |
|-------|--|--------|--|
| (2.1) | La ecuación (3.26) tiene exactamente una solución. | $\iff$ | $\# \text{pivotes} = \# \text{columnas}$. |
| (2.2) | La ecuación (3.26) tiene infinitas soluciones. | $\iff$ | $\# \text{pivotes} < \# \text{columnas}$. |

Observe que el caso (1), sin solución, no puede ocurrir para sistemas homogéneos.

El siguiente teorema conecta lo anterior con la invertibilidad de la matriz que representa el sistema.

Teorema 3.44. *Sea $A \in M(n \times n)$. Entonces las siguientes afirmaciones son equivalentes:*

- (I) A es invertible.
- (II) Para todo $\vec{b} \in \mathbb{R}^n$, la ecuación $A\vec{x} = \vec{b}$ tiene exactamente una solución.
- (III) La ecuación $A\vec{x} = \vec{0}$ tiene exactamente una solución.
- (IV) Toda forma escalonada reducida por filas de A tiene n pivotes.
- (V) A es equivalente por filas a id_n .

Completaremos este teorema con un ítem más en el Capítulo 4 (Teorema 4.11).

Demostración. (II) $\Rightarrow$ (III) se sigue si escogemos $\vec{b} = \vec{0}$.

(III) $\Rightarrow$ (IV) Si $A\vec{x} = \vec{0}$ tiene una única solución, entonces, por el Caso (2.1) anterior (o por el Caso (2.1) del Teorema 3.7), el número de pivotes es igual a n (el número de columnas de A) en toda forma escalonada reducida por filas de A .

(IV) $\Rightarrow$ (V) es claro.

(V) $\Rightarrow$ (II) se sigue del caso (2.1) anterior (o por el Teorema 3.7(2.1)) porque ninguna forma reducida por filas de A puede tener una fila formada solo por ceros.

Hasta ahora hemos mostrado que (II)–(V) son equivalentes. Ahora tenemos que conectarlas con (I).

(I) $\Rightarrow$ (II) Supongamos que A es invertible y sea $\vec{b} \in \mathbb{R}^n$. Entonces $A\vec{x} = \vec{b} \iff \vec{x} = A^{-1}\vec{b}$, lo cual muestra la existencia y unicidad de la solución.

(II) $\Rightarrow$ (I) Supongamos que se cumple (II). Construiremos A^{-1} como sigue (esto también nos dice cómo podemos calcular A^{-1} si existe). Recordemos que necesitamos una matriz C tal que $AC = \text{id}_n$.

Esta C será entonces nuestra candidata para A^{-1} (todavía tendríamos que verificar que $CA = \text{id}_n$). Sea us denote the columns of C by $\vec{c}_j$ for $j = 1, \dots, n$, so that $C = (\vec{c}_1 | \dots | \vec{c}_n)$. Recordemos que la k -ésima columna de AC es $A(k$ -ésima columna de $C)$ y que las columnas de id_n son exactamente los vectores unitarios $\vec{e}_k$ (el vector con un 1 como k -ésima componente y ceros en las demás). Entonces $AC = \text{id}_n$ se puede escribir como

$$(A\vec{c}_1 | \dots | A\vec{c}_n) = (\vec{e}_1 | \dots | \vec{e}_n).$$

Por (II) sabemos que las ecuaciones de la forma $A\vec{x} = \vec{e}_j$ tienen solución única. Así que solo necesitamos poner $\vec{c}_j =$ solución única de la ecuación $A\vec{x} = \vec{e}_j$. Con esta elección tenemos entonces, en efecto, que $AC = \text{id}_n$.

Falta mostrar que $CA = \text{id}_n$. Para ello, observe que

$$A = \text{id}_n A \implies A = ACA \implies A - ACA = \mathbf{0} \implies A(\text{id}_n - CA) = \mathbf{0}.$$

Esto significa que $A(\text{id}_n - CA)\vec{x} = \vec{0}$ para todo $\vec{x} \in \mathbb{R}^n$. Como por (II) la ecuación $A\vec{y} = \vec{0}$ tiene la solución única $\vec{y} = \vec{0}$, se sigue que $(\text{id}_n - CA)\vec{x} = \vec{0}$ para todo $\vec{x} \in \mathbb{R}^n$. Pero esto significa que $\vec{x} = CA\vec{x}$ para todo $\vec{x}$, por lo tanto, CA debe ser igual a id_n . $\square$

Teorema 3.45. Sea $A \in M(n \times n)$.

- (I) Si A tiene una inversa izquierda C (es decir, si $CA = \text{id}_n$), entonces A es invertible y $A^{-1} = C$.
- (II) Si A tiene una inversa derecha D (es decir, si $AD = \text{id}_n$), entonces A es invertible y $A^{-1} = D$.

Demostración. (i) Por el Teorema 3.44, basta mostrar que $A\vec{x} = \vec{0}$ tiene como solución única $\vec{0}$. Supongamos entonces que $\vec{x} \in \mathbb{R}^n$ satisface $A\vec{x} = \vec{0}$. Entonces $\vec{x} = \text{id}_n \vec{x} = (CA)\vec{x} = C(A\vec{x}) = C\vec{0} = \vec{0}$. Esto muestra que A es invertible. Además, $C = C(\text{id}_n) = C(AA^{-1}) = (CA)A^{-1} = \text{id}_n A^{-1} = A^{-1}$, por lo tanto $C = A^{-1}$.

- (ii) Por (i) aplicado a D , se sigue que D tiene inversa y que $D^{-1} = A$; por lo tanto, por el Teorema 3.42 (II), A es invertible y $A^{-1} = (D^{-1})^{-1} = D$. $\square$

Cálculo de la inversa de una matriz cuadrada dada

Sea A una matriz cuadrada. La demostración del Teorema 3.44 nos dice cómo encontrar su inversa si existe. Solo necesitamos resolver $A\vec{x} = \vec{e}_k$ para $k = 1, \dots, n$. Esto puede ser engorroso y largo, pero ya sabemos que, si estas ecuaciones tienen soluciones, podemos encontrarlas con la eliminación de Gauß-Jordan. Solo necesitamos formar la matriz aumentada $(A|\vec{e}_k)$, aplicar operaciones por filas hasta llegar a $(\text{id}_n|\vec{e}_k)$. Entonces $\vec{e}_k$ es la solución de $A\vec{x} = \vec{e}_k$ y obtenemos la matriz A^{-1} como la matriz cuyas columnas son los vectores $\vec{c}_1, \dots, \vec{c}_n$. Si no es posible reducir A a la matriz identidad, entonces no es invertible.

Observe que los pasos que debemos realizar para reducir A a la matriz identidad dependen solo de los coeficientes de A y no del lado derecho. Así podemos calcular los n vectores $\vec{c}_1, \dots, \vec{c}_n$ con una sola eliminación de Gauß-Jordan (grande) si aumentamos la matriz dada A con los n vectores $\vec{e}_1, \dots, \vec{e}_n$. Pero la matriz $(\vec{e}_1 | \dots | \vec{e}_n)$ no es nada más que la matriz identidad id_n . Por ende, si tomamos $(A|\text{id}_n)$ y aplicamos la eliminación de Gauß-Jordan, y si podemos reducir A a la matriz identidad, entonces las columnas a la derecha son las columnas de la matriz inversa A^{-1} . Si no podemos llegar a la matriz identidad, entonces A no es invertible.

Ejemplos 3.46. (I) Sea $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Mostraremos que A es invertible al reducir la matriz aumentada $(A|\text{id}_2)$:

$$\begin{aligned} (A|\text{id}_2) &= \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array} \right) \xrightarrow{R_2 \rightarrow R_2 - 3R_1} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array} \right) \xrightarrow{R_1 \rightarrow R_1 + R_2} \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & -2 & -3 & 1 \end{array} \right) \\ &\xrightarrow{R_2 \rightarrow -\frac{1}{2}R_2} \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{array} \right). \end{aligned}$$

Por lo tanto, A es invertible y $A^{-1} = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}$.

Podemos verificar el resultado calculando

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix} = \begin{pmatrix} -2+3 & 1-1 \\ -6+6 & 3-2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} -2+3 & -4+4 \\ 3/2-3/2 & 3-2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(II) Sea $A = \begin{pmatrix} 1 & 2 \\ -2 & -4 \end{pmatrix}$. Mostraremos que A no es invertible al reducir la matriz aumentada $(A|\text{id}_2)$:

$$(A|\text{id}_2) = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ -2 & -4 & 0 & 1 \end{array} \right) \xrightarrow{R_2 \rightarrow R_2 + 2R_1} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{array} \right).$$

Como hay una fila cero en la matriz de la izquierda, concluimos que A no es invertible.

(III) Sea $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{pmatrix}$. Mostraremos que A es invertible al reducir la matriz aumentada $(A|\text{id}_3)$:

$$\begin{aligned} (A|\text{id}_3) &= \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 5 & 5 & 1 & 0 & 0 & 1 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 - 5R_1} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 0 & 1 & 0 \\ 0 & 0 & -4 & -5 & 0 & 1 \end{array} \right) \\ &\xrightarrow{R_2 \rightarrow 4R_2 + 3R_3} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 8 & 0 & -15 & 4 & 3 \\ 0 & 0 & -4 & -5 & 0 & 1 \end{array} \right) \xrightarrow{R_1 \rightarrow 4R_1 + R_3} \left(\begin{array}{ccc|ccc} 4 & 4 & 0 & -1 & 0 & 1 \\ 0 & 8 & 0 & -15 & 4 & 3 \\ 0 & 0 & -4 & -5 & 0 & 1 \end{array} \right) \\ &\xrightarrow{R_1 \rightarrow 2R_1 - R_2} \left(\begin{array}{ccc|ccc} 8 & 0 & 0 & 13 & -4 & -1 \\ 0 & 8 & 0 & -15 & 4 & 3 \\ 0 & 0 & -4 & -5 & 0 & 1 \end{array} \right) \\ &\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 13/8 & -1/2 & -1/8 \\ 0 & 1 & 0 & -15/8 & 1/2 & 3/8 \\ 0 & 0 & 1 & 5/4 & 0 & -1/4 \end{array} \right). \end{aligned}$$

Por lo tanto, A es invertible y $A^{-1} = \begin{pmatrix} 13/8 & -1/2 & -1/8 \\ -15/8 & 1/2 & 3/8 \\ 5/4 & 0 & -1/4 \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 13 & -4 & -1 \\ -15 & 4 & 3 \\ 10 & 0 & 2 \end{pmatrix}$.

Podemos verificar el resultado calculando

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{pmatrix} \begin{pmatrix} 13/8 & -1/2 & -1/8 \\ -15/8 & 1/2 & 3/8 \\ 5/4 & 0 & -1/4 \end{pmatrix} = \cdots = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} 13/8 & -1/2 & -1/8 \\ -15/8 & 1/2 & 3/8 \\ 5/4 & 0 & -1/4 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{pmatrix} = \cdots = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Caso especial: inversa de una matriz 2×2

Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Ya sabemos que A es invertible si y solo si su sistema lineal homogéneo asociado tiene exactamente una solución. Por el Teorema 1.11, este es el caso si y solo si $\det A \neq 0$. Recordemos que $\det A = ad - bc$. Así que supongamos aquí que $\det A \neq 0$.

Caso 1. $a \neq 0$.

$$\begin{aligned} (A | \text{id}_2) &= \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right) \xrightarrow{R_2 \rightarrow aR_2 - cR_1} \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & ad - bc & -c & a \end{array} \right) \\ &\xrightarrow{R_1 \rightarrow R_1 - \frac{b}{ad - bc} R_2} \left(\begin{array}{cc|cc} a & 0 & 1 + \frac{bc}{ad - bc} & -\frac{ab}{ad - bc} \\ 0 & ad - bc & -c & a \end{array} \right) = \left(\begin{array}{cc|cc} a & 0 & \frac{ad}{ad - bc} & -\frac{ab}{ad - bc} \\ 0 & ad - bc & -c & a \end{array} \right) \\ &\xrightarrow{R_1 \rightarrow R_1 - \frac{b}{ad - bc} R_2} \left(\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad - bc} & -\frac{b}{ad - bc} \\ 0 & 1 & -\frac{c}{ad - bc} & \frac{a}{ad - bc} \end{array} \right). \end{aligned}$$

Se sigue que

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \quad (3.27)$$

Caso 2. $a = 0$. Como $0 \neq \det A = ad - bc = bc$ en este caso, se sigue que $c \neq 0$ y cálculos como los anteriores conducen nuevamente a la fórmula (3.27).

Ahora debería haber entendido

- la relación entre la invertibilidad de una matriz cuadrada A y la existencia y unicidad de solución de $A\vec{x} = \vec{b}$,
- que invertir una matriz es lo mismo que resolver un sistema lineal,
- etc.

Ahora debería ser capaz de

- calcular la inversa de una matriz cuadrada si existe,
- usar la inversa de una matriz cuadrada, si existe, para resolver el sistema lineal asociado,
- etc.

Ejercicios.

1. Determine la inversa (si es posible) de las siguientes matrices:

(a) $\begin{pmatrix} 1 & 3 \\ -2 & 6 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{pmatrix}$

(c) $\begin{pmatrix} -1 & 6 \\ 2 & -12 \end{pmatrix}$

(d) $\begin{pmatrix} 3 & 2 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & -1 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 0 & 0 \\ -3 & 0 & 0 \\ 4 & 3 & 1 \end{pmatrix}$

(f) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ 1 & -1 & 2 & 1 \\ 1 & 3 & 3 & 2 \end{pmatrix}$

(g) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 0 & 3 & 3 \end{pmatrix}$

(h) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

2. De los siguientes sistemas de ecuaciones lineales, ¿cuál tiene una solución no trivial?

$$\begin{array}{ll} 2x + y - z = 0, & x - y - z = 0, \\ \text{i) } x - 2y - 3z = 0, & \text{ii) } 2x + y + 2z = 0, \\ -3x - y - z = 0. & -2x + 5y + 6z = 0. \end{array}$$

3. Determine todos los valores de a tal que

$$A = \begin{pmatrix} 1 & 4 & a^2 \\ 1 & 0 & 0 \\ 1 & 2 & 2 \end{pmatrix}$$

es invertible. En dicho caso hallar A^{-1} .

4. Calcular la inversa de la matriz de rotación $C_\vartheta = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix}$. ¿Cómo se interpreta geoméricamente C_ϑ^{-1} ? (ver Sección 3.4, Ejercicio 6.).

5. Calcular la inversa de

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \vartheta & -\sin \vartheta \\ 0 & \sin \vartheta & \cos \vartheta \end{pmatrix}.$$

6. Sea $A \in M(n \times n)$ no invertible. Demuestre que existe $B \in M(n \times n)$, $B \neq \mathbf{0}$ tal que $AB = \mathbf{0}$. (Hint: considere el sistema homogéneo $A\vec{x} = \vec{0}$).

7. Sean $B \in M(6 \times 5)$ y $C \in M(5 \times 6)$. Muestre que BC no puede ser una matriz invertible. (Hint: considere el sistema homogéneo $C\vec{x} = \vec{0}$).

3.7. La transpuesta de una matriz

Definición 3.47. Sea

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in M(m \times n).$$

Entonces su transpuesta A^t es la matriz $n \times m$ cuyas columnas son las filas de A y cuyas filas son las columnas de A , es decir,

$$A^t = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix} \in M(n \times m).$$

Si denotamos $A^t = (\tilde{a}_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$, entonces $\tilde{a}_{ij} = a_{ji}$ para $i = 1, \dots, n$ y $j = 1, \dots, m$.

Ejemplos 3.48. Las transpuestas de

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 7 & 7 \\ 3 & 2 & 4 \end{pmatrix}$$

son

$$A^t = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \quad B^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}, \quad C^t = \begin{pmatrix} 1 & 4 & 7 & 3 \\ 2 & 5 & 7 & 2 \\ 3 & 6 & 7 & 4 \end{pmatrix}.$$

Proposición 3.49. Sean $A, B \in M(m \times n)$. Entonces $(A^t)^t = A$ y $(A + B)^t = A^t + B^t$.

Demostración. Claro. □

Teorema 3.50. Sean $A \in M(m \times n)$ y $B \in M(n \times k)$. Entonces $(AB)^t = B^t A^t$.

Demostración. Observe que tanto $(AB)^t$ como $B^t A^t$ son matrices $m \times k$. Para mostrar que son iguales, solo necesitamos mostrar que son iguales entrada por entrada. Sean $i \in \{1, \dots, m\}$ y

$j \in \{1, \dots, k\}$. Entonces

$$\begin{aligned}
 \text{componente } ij \text{ de } (AB)^t &= \text{componente } ji \text{ de } AB \\
 &= [\text{fila } j \text{ de } A] \times [\text{columna } i \text{ de } B] \\
 &= [\text{columna } j \text{ de } A^t] \times [\text{fila } i \text{ de } B^t] \\
 &= [\text{fila } i \text{ de } B^t] \times [\text{columna } j \text{ de } A^t] \\
 &= \text{componente } ij \text{ de } B^t A^t.
 \end{aligned}$$

□

Teorema 3.51. Sea $A \in M(n \times n)$. Entonces A is invertible if y only if A^t is invertible. En este caso, $(A^t)^{-1} = (A^{-1})^t$.

Demostración. Supongamos que A es invertible. Entonces $AA^{-1} = \text{id}$. Tomando la transpuesta en ambos lados, obtenemos

$$\text{id} = \text{id}^t = (AA^{-1})^t = (A^{-1})^t A^t.$$

Esto muestra que A^t es invertible y que su inversa es $(A^{-1})^t$; véase el Teorema 3.45. Ahora supongamos que A^t es invertible. De lo que acabamos de mostrar se sigue que entonces también su transpuesta $(A^t)^t = A$ es invertible. □

A continuación mostramos una relación importante entre la transposición de una matriz y el producto interno en $\mathbb{R}^n$.

Teorema 3.52. Sea $A \in M(m \times n)$.

- (I) $\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A^t \vec{y} \rangle$ para todo $\vec{x} \in \mathbb{R}^n$ y todo $\vec{y} \in \mathbb{R}^m$.
- (II) Si $\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, B\vec{y} \rangle$ for all $\vec{x} \in \mathbb{R}^n$ y all $\vec{y} \in \mathbb{R}^m$, entonces $B = A^t$.

Demostración. Sea $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ y $B = (b_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$.

- (I) Observe que la k -ésima componente de $A\vec{x}$ es $(A\vec{x})_k = \sum_{j=1}^n a_{kj}x_j$. y que la coordenada ℓ -ésima de $A^t \vec{y}$ es $(A^t \vec{y})_\ell = \sum_{j=1}^m a_{j\ell}y_j$. Entonces

$$\langle A\vec{x}, \vec{y} \rangle = \sum_{k=1}^m (A\vec{x})_k y_k = \sum_{k=1}^m \sum_{j=1}^n a_{kj} x_j y_k = \sum_{j=1}^n \sum_{k=1}^m x_j a_{kj} y_k = \sum_{j=1}^n x_j (A^t \vec{y})_j = \langle \vec{x}, A^t \vec{y} \rangle.$$

- (II) Tenemos que mostrar que para todo $i = 1, \dots, m$ y $j = 1, \dots, n$ se tiene $a_{ij} = b_{ji}$. Tomemos $\vec{x} = \vec{e}_j \in \mathbb{R}^n$ y $\vec{y} = \vec{e}_i \in \mathbb{R}^m$. Si tomamos el producto interno de $A\vec{e}_j$ con $\vec{e}_i$, entonces obtenemos la i -ésima componente de $A\vec{e}_j$. Recordemos que $A\vec{e}_j$ es la j -ésima columna de A , por lo tanto

$$\langle A\vec{e}_j, \vec{e}_i \rangle = a_{ij}.$$

De manera similar, si tomamos el producto interno de $B\vec{e}_i$ con $\vec{e}_j$, entonces obtenemos la j -ésima componente de $B\vec{e}_i$. Como $B\vec{e}_i$ es la i -ésima columna de B , se sigue que

$$\langle \vec{e}_j, B\vec{e}_i \rangle = b_{ji}.$$

Por hipótesis $\langle A\vec{e}_j, \vec{e}_i \rangle = \langle \vec{e}_j, B\vec{e}_i \rangle$, de donde se sigue que $a_{ij} = b_{ji}$, y por lo tanto $B = A^t$. □

Definición 3.53. Sea $A = (a_{ij})_{i,j=1}^n \in M(n \times n)$ una matriz cuadrada.

- (I) A se llama *triangular superior* si $a_{ij} = 0$ si $i > j$.
- (II) A se llama *triangular inferior* si $a_{ij} = 0$ si $i < j$.
- (III) A se llama *diagonal* si $a_{ij} = 0$ cuando $i \neq j$. Diagonal matrices son sometimes denoted by $\text{diag}(c_1, \dots, c_n)$ where the $c_1, \dots, c_n$ son the numbers on the diagonal of the matrix.

Este significa que para una matriz triangular superior, todas sus entradas debajo de la diagonal son cero; para una matriz triangular inferior, todas sus entradas debajo de la diagonal son cero; y para una matriz diagonal, todas las entradas con excepción de las en la diagonal, son cero. Estas matrices se ven como sigue:

$$\begin{pmatrix} a_{11} & & & \\ & a_{22} & * & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}, \quad \begin{pmatrix} a_{11} & & & \\ & a_{22} & & 0 \\ & & \ddots & \\ * & & & a_{nn} \end{pmatrix}, \quad \begin{pmatrix} a_{11} & & & \\ & a_{22} & & 0 \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}$$

matriz triangular superior, matriz triangular inferior, matriz diagonal
 $\text{diag}(a_{11}, \dots, a_{nn})$

Observación 3.54. Una matrix es simultáneamente triangular superior e inferior si y solo si es diagonal.

Ejemplos 3.55.

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 2 & 5 \\ 0 & 0 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 2 & 4 & 2 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 5 & 0 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 8 \end{pmatrix}, \quad E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Las matrices A, B, D, E son triangulares superiores; C, D, E son triangulares inferiores; D, E son diagonales.

Definición 3.56. (I) A matrix $A \in M(n \times n)$ is called *simétrica* si $A^t = A$. El conjunto de todas las matrices simétricas $n \times n$ se denota por $M_{\text{sym}}(n \times n)$.

(II) A matrix $A \in M(n \times n)$ is called *antisimétrica* si $A^t = -A$. El conjunto de todas las matrices antisimétricas $n \times n$ se denota por $M_{\text{asym}}(n \times n)$.

Ejemplos 3.57.

$$A = \begin{pmatrix} 1 & 7 & 4 \\ 7 & 2 & 5 \\ 4 & 5 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 0 & 4 \\ 0 & 4 & 0 \\ 4 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 2 & -5 \\ -2 & 0 & -3 \\ 5 & 3 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 & 8 \\ 0 & 3 & 0 \\ 2 & 0 & 0 \end{pmatrix}.$$

Las matrices A y B son simétricas, C es antisimétrica y D no es ninguna de las dos.

Claramente, toda matriz diagonal es simétrica.

Ejercicio 3.58. ■ Sea $A \in M(n \times n)$. Demuestre que $A + A^t$ es simétrica y que $A - A^t$ es antisimétrica.

- Demuestre que toda matriz $A \in M(n \times n)$ se deja escribir como suma de una matriz simétrica y de una matriz antisimétrica.

Question 3.2

¿Cuántas posibilidades hay de expresar una matriz dada $A \in M(n \times n)$ como suma de una matriz simétrica y de una matriz antisimétrica?

Ejercicio 3.59. Demuestre que las entradas en la diagonal de una matriz antisimétrica son 0.

Ahora debería haber entendido

- por qué $(AB)^t = B^t A^t$,
- qué tiene que ver la transpuesta de una matriz con el producto interno,
- etc.

Ahora debería ser capaz de

- calcular la transpuesta de una matriz dada,
- verificar si una matriz es simétrica, antisimétrica o ninguna de las dos,
- etc.

Ejercicios.

1. (a) Encuentre las transpuestas de las siguientes matrices:

$$(a) \begin{pmatrix} -1 & 4 \\ 10 & 8 \end{pmatrix}, \quad (b) \begin{pmatrix} 1 & 3 \\ -1 & 2 \\ 4 & 5 \end{pmatrix}, \quad (c) \begin{pmatrix} 2 & -1 & 1 & 5 \\ 0 & 0 & 4 & 13 \end{pmatrix}, \quad (d) \begin{pmatrix} 0 \\ -1 \\ 4 \\ 3 \end{pmatrix}.$$

- (b) Para cada una de las matrices del punto anterior, verifique la igualdad $\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A^t \vec{y} \rangle$.

2. Encuentre α, β tales que $\begin{pmatrix} 1 & \beta + 5 & \alpha + 2\beta - 2 \\ \alpha + 2\beta & -1 & 2 \\ 2\alpha + \beta & 2 & 4 \end{pmatrix}$ es una matriz simétrica.

3. ¿Cuáles de las siguientes matrices son antisimétricas?

$$(a) \begin{pmatrix} 2 & -7 \\ 7 & 0 \end{pmatrix}, \quad (b) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad (c) \begin{pmatrix} 3 & -3 & -3 \\ 3 & 3 & -3 \\ 3 & 3 & 3 \end{pmatrix}, \quad (d) \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 2 \\ 1 & -2 & 0 \end{pmatrix}.$$

4. Si $A, B \in M_{sym}(n \times n)$, muestre que $(AB)^t = BA$. ¿Se puede concluir que AB es simétrica?
5. Sean $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathbb{R}^3$ tales que $\langle v_i, v_j \rangle = 0$ si $i \neq j$ y $\langle v_i, v_j \rangle = 1$ si $i = j$. Sea $A = [\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3]$ la matriz cuyas columnas son los vectores dados. Muestre que $AA^t = \text{id}_3$ (ver Sección 3.6, Ejercicio 5.).
6. Muestre que la suma de dos matrices simétricas (antisimétricas) de tamaño $n \times n$ da como resultado una matriz simétrica (antisimétrica).
7. (a) Muestre que una matriz triangular superior (inferior) es invertible si y solo si todas las entradas de la diagonal principal son distintas de 0.
(b) Sea $D \in M(n \times n)$ una matriz diagonal. Determine cuando D es invertible y halle su inversa

3.8. Matrices elementales

En esta sección estudiamos tres tipos especiales de matrices. Se llaman *matrices elementales*. Definámoslas.

Definición 3.60. Para $n \in \mathbb{N}$ definimos las siguientes matrices en $M(n \times n)$:

(I) $S_j(c) = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & c & \\ & & & \ddots \\ & & & & 1 \end{pmatrix}$ para $j = 1, \dots, n$ y $c \neq 0$. Todas las entradas fuera de la diagonal son 0.

(II) $Q_{jk}(c) = \begin{pmatrix} & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{pmatrix}$ para $j, k = 1, \dots, n$ con $j \neq k$ y $c \in \mathbb{R}$. El número c está en la fila j y la columna k . Todas las entradas, excepto c y las de la diagonal, son 0.

(III) $P_{jk} = \begin{pmatrix} & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{pmatrix}$ para $j, k = 1, \dots, n$. Esta matriz se obtiene de la matriz identidad intercambiando las filas j y k (o, equivalentemente, intercambiando las columnas j y k).

Ejemplos 3.61. Veamos algunos ejemplos para $n = 2$.

$$S_1(5) = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}, \quad Q_{21}(3) = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}, \quad P_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Algunos ejemplos para $n = 3$:

$$S_3(-2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad Q_{23}(4) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_{31} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad P_{21} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Veamos cómo actúan estas matrices sobre otras matrices $n \times n$. Sea $A = (a_{ij})_{i,j=1}^n \in M(n \times n)$. Queremos calcular EA , donde E es una matriz elemental.

$$\begin{aligned} \blacksquare S_j(c)A &= \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & c & \\ & & & \ddots \\ & & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ ca_{j1} & ca_{j2} & \dots & ca_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}. \\ \blacksquare Q_{jk}(c)A &= \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \ddots \\ & & & & c \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} + ca_{k1} & a_{j2} + ca_{k2} & \dots & a_{jn} + ca_{kn} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}. \\ \blacksquare P_{jk}A &= \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 0 & \\ & & & \ddots \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 0 \\ & & & & & & & \ddots \\ & & & & & & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kn} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}. \end{aligned}$$

En resumen, vemos que

Proposición 3.62. $\blacksquare S_j(c)$ multiplica la fila j -ésima de A por c .

- $\blacksquare Q_{jk}(c)$ suma c veces la fila k -ésima a la fila j -ésima de A .
- $\blacksquare P_{jk}$ intercambia la fila k -ésima y la fila j -ésima de A .

¡Estas son exactamente las operaciones de la eliminación de Gauß o de Gauß-Jordan! Así vemos que toda operación por filas puede realizarse multiplicando por la izquierda por una matriz elemental apropiada.

Observación 3.63. La forma de las matrices elementales es bastante fácil de recordar si se recuerda que $E \text{id}_n = E$ para toda matriz E , en particular para una matriz elemental. Así, si quiere recordar, por ejemplo, cómo se ve la matriz 5×5 que suma 3 veces la segunda fila a la cuarta, recuerde simplemente que esta matriz es

$$E = E \text{id}_5 = (\text{tomar } \text{id}_5 \text{ y sumar 3 veces su segunda fila a su cuarta fila}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

que es $Q_{42}(3)$.

Question 3.3

¿Cómo actúan las matrices elementales sobre otras matrices si las multiplicamos por la derecha?

Indicación. Hay dos maneras de encontrar la respuesta. Una es realizar la multiplicación de matrices como hicimos en la página 125. O podría usar que $AE = [(AE)^t]^t = [E^t A^t]^t$. Si E es una matriz elemental, entonces E^t también lo es; véase la Proposición 3.65. Como sabe cómo se ve $E^t A^t$, puede entonces deducir cómo se ve su transpuesta.

Como la acción de una matriz elemental se puede “deshacer” (ya que la operación por filas correspondiente se puede deshacer), esperamos que sean invertibles. La siguiente proposición muestra que en efecto lo son y que su inversa es nuevamente una matriz elemental del mismo tipo.

Proposición 3.64. *Toda matriz elemental $n \times n$ es invertible. Más precisamente, para $j, k = 1, \dots, n$ con $j \neq k$ se cumple lo siguiente:*

- (I) $(S_j(c))^{-1} = S_j(c^{-1})$ for $c \neq 0$.
- (II) $(Q_{jk}(c))^{-1} = Q_{jk}(-c)$.
- (III) $(P_{jk})^{-1} = P_{jk}$.

Demostración. Cálculos directos. □

Muestre que la Proposición 3.64 es verdadera. Convéznase de que es verdadera usando su interpretación como operaciones por filas.

Proposición 3.65. *La transpuesta de una matriz elemental $n \times n$ es nuevamente una matriz elemental. Más precisamente, para $j, k = 1, \dots, n$ con $j \neq k$ se cumple lo siguiente:*

- (I) $(S_j(c))^t = S_j(c)$ for $c \neq 0$.
- (II) $(Q_{jk}(c))^t = Q_{kj}(c)$.
- (III) $(P_{jk})^t = P_{jk}$.

Demostración. Cálculos directos. □

Ejercicio 3.66. Muestre que $Q_{jk}(c) = S_k(c^{-1})Q_{jk}(1)S_k(c)$ para $c \neq 0$. Interprete la fórmula en términos de operaciones por filas.

Ejercicio. Muestre que P_{jk} se puede escribir como producto de matrices de la forma $Q_{jk}(c)$ y $S_j(c)$.

Volvamos a la relación entre las matrices elementales y el proceso de eliminación de Gauß-Jordan.

Proposición 3.67. Sea $A \in M(n \times n)$ y sea A' una forma escalonada por filas de A . Entonces existen matrices elementales $E_1, \dots, E_k$ tales que

$$A = E_1 E_2 \cdots E_k A'.$$

Demostración. Sabemos que podemos llegar a A' aplicando operaciones por filas adecuadas a A . Por la Proposición 3.62, estas corresponden a multiplicar A por la izquierda por matrices elementales adecuadas $F_k, F_{k-1}, \dots, F_2, F_1$, es decir,

$$A' = F_k F_{k-1} \cdots F_2 F_1 A.$$

Sabemos que todas las F_j son invertibles; por lo tanto, su producto es invertible y obtenemos

$$A = [F_k F_{k-1} \cdots F_2 F_1]^{-1} A' = F_1^{-1} F_2^{-1} \cdots F_{k-1}^{-1} F_k^{-1} A'.$$

Sabemos que la inversa de toda matriz elemental F_j es nuevamente una matriz elemental; así, si ponemos $E_j = F_j^{-1}$ para $j = 1, \dots, k$, la proposición queda demostrada. $\square$

Corolario 3.68. Sea $A \in M(n \times n)$. Entonces existen matrices elementales $E_1, \dots, E_k$ y una matriz triangular superior U tales que

$$A = E_1 E_2 \cdots E_k U.$$

Demostración. Esto se sigue inmediatamente de la Proposición 3.67 si recordamos que toda forma escalonada reducida por filas de A es una matriz triangular superior. $\square$

El siguiente teorema muestra que toda matriz invertible está “compuesta” de matrices elementales.

Teorema 3.69. Sea $A \in M(n \times n)$. Entonces A es invertible si y solo si se puede escribir como producto de matrices elementales.

Demostración. Supongamos que A es invertible. Entonces la forma escalonada reducida por filas de A es id_n . Por lo tanto, por la Proposición 3.67, existen matrices elementales $E_1, \dots, E_k$ tales que $A = E_1 \cdots E_k \text{id}_n = E_1 \cdots E_k$.

Si, por otro lado, sabemos que A es el producto de matrices elementales, digamos, $A = F_1 \cdots F_\ell$, entonces claramente A es invertible, ya que cada matriz elemental F_j es invertible y el producto de matrices invertibles es invertible. $\square$

Terminamos esta sección con un ejercicio en el que escribimos una matriz invertible 2×2 como producto de matrices elementales. Observe que hay infinitas formas de escribirla como producto de matrices elementales, así como hay infinitas formas de realizar reducción por filas para llegar a la matriz identidad.

Ejemplo 3.70. Escriba la matriz $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ como producto de matrices elementales.

Solución. Usamos la idea de la demostración del Teorema 3.44: aplicamos el proceso de eliminación de Gauß-Jordan y escribimos las transformaciones por filas correspondientes como matrices elementales.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \xrightarrow[\textcolor{blue}{Q_{21}(-3)}]{R_2 \rightarrow R_2 - 3R_1} \underbrace{\begin{pmatrix} 1 & 2 \\ 0 & -2 \end{pmatrix}}_{=Q_{21}(-3)A} \xrightarrow[\textcolor{blue}{Q_{12}(1)}]{R_1 \rightarrow R_1 + R_2} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}}_{=Q_{12}(1)Q_{21}(-3)A} \xrightarrow[\textcolor{blue}{S_2(-\frac{1}{2})}]{R_2 \rightarrow -\frac{1}{2}R_2} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{=S_2(-\frac{1}{2})Q_{12}(1)Q_{21}(-3)A}$$

Así obtenemos que

$$\text{id}_2 = S_2(-\tfrac{1}{2})Q_{12}(1)Q_{21}(-3)A. \quad (3.28)$$

Como las matrices elementales son invertibles, podemos despejar A y obtener

$$\begin{aligned} A &= [S_2(-\tfrac{1}{2})Q_{12}(1)Q_{21}(-3)]^{-1} \text{id}_2 = [S_2(-\tfrac{1}{2})Q_{12}(1)Q_{21}(-3)]^{-1} \\ &= [Q_{21}(-3)]^{-1}[Q_{12}(1)]^{-1}[S_2(-\tfrac{1}{2})]^{-1} \\ &= Q_{21}(3)Q_{12}(-1)S_2(-2). \end{aligned} \quad \diamond$$

Observe que de (3.28) obtenemos gratis la factorización de A^{-1} . Claramente, debemos tener

$$A^{-1} = S_2(-\tfrac{1}{2})Q_{12}(1)Q_{21}(-3). \quad (3.29)$$

Si quisiéramos, ahora podríamos usar (3.29) para calcular A^{-1} . No es en absoluto una sorpresa que en realidad obtengamos primero la factorización de A^{-1} , porque la eliminación de Gauß-Jordan conduce a la inversa de A . Así, A^{-1} es la composición de las matrices que llevan de A a la matriz identidad. (Para pasar de la matriz identidad a A , necesitamos invertir estos pasos.)

Ahora debería haber entendido

- la relación de las matrices elementales con el proceso de Gauß-Jordan,
- por qué una matriz es invertible si y solo si es producto de matrices elementales,
- etc.

Ahora debería ser capaz de

- expresar una matriz invertible como producto de matrices elementales,
- etc.

Ejercicios.

1. Determine cuáles de las siguientes son matrices elementales:

$$\begin{array}{lll}
 \text{(a)} \quad \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} & \text{(b)} \quad \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} & \text{(c)} \quad \begin{pmatrix} -\frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{pmatrix} \\
 \text{(d)} \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} & \text{(e)} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \text{(f)} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}
 \end{array}$$

2. Muestre que cada una de las siguientes matrices es invertible y factorícela como un producto de matrices elementales:

$$\begin{array}{lll}
 \text{(a)} \quad \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}, & \text{(b)} \quad \begin{pmatrix} 2 & 0 & 4 \\ 0 & 1 & 1 \\ 3 & -1 & 1 \end{pmatrix}, & \text{(c)} \quad \begin{pmatrix} 1 & 1 & 0 & 0 \\ 3 & 2 & 4 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, & \text{(d)} \quad \begin{pmatrix} 3 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & -1 \end{pmatrix}.
 \end{array}$$

3. Escriba cada matriz como producto de matrices elementales y una matriz triangular superior:

$$\begin{array}{lll}
 \text{(a)} \quad \begin{pmatrix} 2 & -2 \\ -2 & 6 \end{pmatrix}, & \text{(b)} \quad \begin{pmatrix} 2 & 1 & 3 \\ 0 & -3 & 1 \\ 1 & 0 & 2 \end{pmatrix}, & \text{(c)} \quad \begin{pmatrix} 1 & 0 & 0 \\ 5 & 0 & 0 \\ -2 & 1 & 3 \end{pmatrix}, & \text{(d)} \quad \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}.
 \end{array}$$

4. En los siguientes problemas, encuentre una matriz elemental E tal que $EA = B$:

$$\begin{array}{ll}
 \text{(a)} \quad A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}, B = \begin{pmatrix} 3 & 6 \\ -1 & 3 \end{pmatrix}, & \text{(b)} \quad A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 1 & -2 \end{pmatrix}, \\
 \text{(c)} \quad A = \begin{pmatrix} 1 & 4 \\ 5 & 6 \\ 7 & 8 \end{pmatrix}, B = \begin{pmatrix} 7 & 8 \\ 5 & 6 \\ 1 & 4 \end{pmatrix}, & \text{(d)} \quad A = \begin{pmatrix} 1 & -2 & 5 & 3 \\ 2 & -1 & 0 & 4 \\ 5 & 1 & -3 & 2 \end{pmatrix}, B = \begin{pmatrix} -3 & 0 & 5 & -5 \\ 2 & -1 & 0 & 4 \\ 5 & 1 & -3 & 2 \end{pmatrix}, \\
 \text{(e)} \quad A = \begin{pmatrix} 1 & -3 \\ -1 & 1 \end{pmatrix}, B = \begin{pmatrix} -2 & 0 \\ -1 & 1 \end{pmatrix}, & \text{(f)} \quad A = \begin{pmatrix} 5 & 1 & 2 \\ -1 & 3 & 4 \\ 1 & -2 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 11 & 2 \\ -1 & 3 & 4 \\ 1 & -2 & 0 \end{pmatrix}.
 \end{array}$$

5. (a) Sea $A \in M(3 \times 3)$ una matriz triangular superior (inferior) tal que las entradas de su diagonal son todas distintas de 0. Muestre que A se factoriza como producto de a lo más seis matrices elementales.
- (b) Sean $A, B \in M(3 \times 3)$ matrices triangulares superiores (inferiores). Muestre que AB es una matriz triangular superior (inferior).
- (c) Sea $A \in M(3 \times 3)$ una matriz triangular superior (inferior) tal que las entradas de su diagonal son todas distintas de 0. Muestre que A^{-1} es triangular superior (inferior).

3.9. Resumen

Operaciones elementales por filas (= operaciones que conducen a un sistema equivalente) para resolver un sistema lineal.

Operación elemental	Notación	Operación inversa
① Intercambiar las filas j y k .	$R_j \leftrightarrow R_k$	$R_j \leftrightarrow R_k$
② Multiplicar la fila j por algún $\lambda \in \mathbb{R} \setminus \{0\}$	$R_j \rightarrow \lambda R_j$	$R_j \rightarrow \frac{1}{\lambda} R_j$
③ Reemplazar la fila k por la suma de la fila k y λ veces R_j , y dejar la fila j sin cambios ($j \neq k$)	$R_k \rightarrow R_k + \lambda R_j$	$R_k \rightarrow R_k - \lambda R_j$

Sobre las soluciones de un sistema lineal.

- Un sistema lineal tiene o bien ninguna, exactamente una o infinitas soluciones.
- Si el sistema es homogéneo, entonces tiene o bien exactamente una solución o infinitas soluciones. Siempre tiene al menos una solución, a saber, la solución trivial.
- El conjunto de todas las soluciones de un sistema homogéneo de ecuaciones lineales es un espacio vectorial.
- El conjunto de todas las soluciones de un sistema inhomogéneo de ecuaciones lineales es un espacio vectorial afín.

Para $A \in M(m \times n)$ y $\vec{b} \in \mathbb{R}^m$, considere la ecuación $A\vec{x} = \vec{b}$. Entonces se cumple lo siguiente:

- | | | |
|---------------------------|--------|--|
| (1) Ninguna solución | $\iff$ | La forma escalonada reducida por filas del sistema aumentado $(A \vec{b})$ tiene una fila de la forma $(0 \cdots 0 \beta)$ con algún $\beta \neq 0$. |
| (2) Al menos una solución | $\iff$ | La forma escalonada reducida por filas del sistema aumentado $(A \vec{b})$ no tiene ninguna fila de la forma $(0 \cdots 0 \beta)$ con algún $\beta \neq 0$. |

En el caso (2), tenemos los siguientes dos subcasos:

- | | | |
|--------------------------------|--------|-------------------------|
| (2.1) Exactamente una solución | $\iff$ | # pivotes = # columnas. |
| (2.2) Infinitas soluciones | $\iff$ | # pivots < # columns. |

Álgebra con matrices y vectores

Una matriz $A \in M(m \times n)$ puede verse como una función $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Definición.

$$A\vec{x} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix},$$

$$\begin{aligned}
A + B &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \\
&= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}, \\
AB &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \\
&= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} & \cdots & a_{11}b_{1k} + a_{12}b_{2k} + \cdots + a_{1n}b_{nk} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} & \cdots & a_{21}b_{1k} + a_{22}b_{2k} + \cdots + a_{2n}b_{nk} \\ \vdots & & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1} & \cdots & a_{m1}b_{1k} + a_{m2}b_{2k} + \cdots + a_{mn}b_{nk} \end{pmatrix} \\
&= (c_{j\ell})_{j\ell}
\end{aligned}$$

with

$$c_{j\ell} = \sum_{h=1}^n a_{jh} b_{h\ell}.$$

- Suma de matrices: componente a componente,
- Producto de matriz con vector o de matriz con matriz: “multiplicar fila por columna”.

Propiedades. Sean $A_1, A_2, A_3 \in M(m \times n)$, $B \in M(n \times k)$, $C \in M(k \times r)$ matrices, $\vec{x}, \vec{y} \in \mathbb{R}^n$, $\vec{z} \in \mathbb{R}^k$ y $c \in \mathbb{K}$.

- $A_1 + A_2 = A_2 + A_1$,
- $(A_1 + A_2) + A_3 = A_1 + (A_2 + A_3)$,
- $(AB)C = A(BC)$,
- en general, $AB \neq BA$,
- $A(\vec{x} + c\vec{y}) = A\vec{x} + cA\vec{y}$,
- $(A_1 + cA_2)\vec{x} = A_1\vec{x} + cA_2\vec{x}$,
- $(AB)\vec{z} = A(B\vec{z})$,

Transposición de matrices

Sea $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in M(m \times n)$. Entonces su *transpuesta* es la matriz $A^t = (\tilde{a}_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \in M(n \times m)$ with $\tilde{a}_{ij} = a_{ji}$.

Para $A, B \in M(m \times n)$ y $C \in M(n \times k)$ tenemos

- $(A^t)^t = A$,
- $(A + B)^t = A^t + B^t$,
- $(AC)^t = C^t A^t$,
- $\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A^t \vec{y} \rangle$ para todo $\vec{x} \in \mathbb{R}^n$ y todo $\vec{y} \in \mathbb{R}^m$.

Una matriz A se llama *simétrica* si $A^t = A$ y *antisimétrica* si $A^t = -A$. Observe que solo las matrices cuadradas pueden ser simétricas.

Una matriz $A = (a_{ij})_{i,j=1,\dots,n} \in M(n \times n)$ se llama

- *upper triangular* if $a_{ij} = 0$ whenever $i > j$,
- *lower triangular* if $a_{ij} = 0$ whenever $i < j$,
- *diagonal* si $a_{ij} = 0$ siempre que $i \neq j$.

Claramente, una matriz es diagonal si y solo si es triangular superior e inferior. La transpuesta de una matriz triangular superior es triangular inferior, y viceversa. Toda matriz diagonal es simétrica.

Invertibilidad de matrices

Una matriz $A \in M(n \times n)$ se llama *invertible* si existe una matriz $B \in M(n \times n)$ tal que $AB = BA = \text{id}_n$. En este caso, B se llama la *inversa* de A y se denota por A^{-1} . Si A no es invertible, entonces se llama *singular*.

- La inversa de una matriz invertible A es única.
- Si A es invertible, entonces A^{-1} también lo es y $(A^{-1})^{-1} = A$.
- Si A es invertible, entonces A^t también lo es y $(A^t)^{-1} = (A^{-1})^t$.
- Si A y B son invertibles, entonces AB también lo es y $(AB)^{-1} = B^{-1}A^{-1}$.

Teorema. Sea $A \in M(n \times n)$. Entonces las siguientes afirmaciones son equivalentes:

- (I) A es invertible.
- (II) Para todo $\vec{b} \in \mathbb{R}^n$, la ecuación $A\vec{x} = \vec{b}$ tiene exactamente una solución.
- (III) La ecuación $A\vec{x} = \vec{0}$ tiene exactamente una solución.
- (IV) Toda forma escalonada reducida por filas de A tiene n pivotes.
- (V) A es equivalente por filas a id_n .

Cálculo de A^{-1} usando eliminación de Gauß-Jordan

Sea $A \in M(n \times n)$. Form the augmented matrix $(A | \text{id}_n)$ and use the Gauß-Jordan elimination to reduce A to its reduced row echelon form A' : $(A | \text{id}_n) \rightarrow \cdots \rightarrow (A' | B)$. Si $A' = \text{id}_n$, entonces A es invertible y $A^{-1} = B$. Si $A' \neq \text{id}_n$, entonces A no es invertible.

Inversa de una matriz 2×2

Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Entonces $\det A = ad - bc$. Si $\det A = 0$, entonces A no es invertible. Si $\det A \neq 0$, entonces A es invertible y $A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Matrices elementales

Tenemos los siguientes tres tipos de matrices elementales:

- $S_j(c) = (s_{ik})_{i,k=1,\dots,n}$ para $c \neq 0$, donde $s_{ik} = 0$ si $i \neq k$, $s_{kk} = 1$ para $k \neq j$ and $s_{jj} = c$,
- $Q_{jk}(c) = (q_{i\ell})_{i,\ell=1,\dots,n}$ para $j \neq k$, donde $q_{jk} = c$, $q_{\ell\ell} = 1$ for all $\ell = 1, \dots, n$ y todos los demás coeficientes son iguales a cero,
- $P_{jk} = (p_{i\ell})_{i,\ell=1,\dots,n}$ para $j \neq k$, donde $p_{\ell\ell} = 1$ for all $\ell \in \{1, \dots, n\} \setminus \{j, k\}$, $p_{jk} = p_{kj} = 1$ y todos los demás coeficientes son iguales a cero.

$$S_j(c) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & c & \\ & & & \ddots \\ & & & & 1 \end{pmatrix}, \quad Q_{jk}(c) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & c & \\ & & & \ddots \\ & & & & 1 \end{pmatrix}, \quad P_{jk} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 0 & 1 \\ & & & \ddots \\ & & 1 & 0 \\ & & & & \ddots \\ & & & & & 1 \end{pmatrix}.$$

Relación matriz elemental – operación elemental por filas

Matriz elemental	Operación elemental	Notación
P_{jk}	Intercambiar la fila j con la fila k	$R_j \leftrightarrow R_k$
$S_j(c)$, $c \neq 0$	Multiplicar la fila j por c	$R_j \rightarrow cR_j$
$Q_{jk}(c)$, $j \neq k$	Sumar c veces la fila k a la fila j	$R_k \rightarrow R_k + cR_j$

3.10. Ejercicios

1. Vuelva al Capítulo 1 y haga los ejercicios otra vez utilizando los conocimientos adquiridos en este capítulo.

2. Encuentre las fracciones parciales de $\frac{2x^2 - 4x + 14}{x(x-2)^2}$.

3. Encuentre un sistema lineal 2×3 cuya solución sea

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \quad t \in \mathbb{R}.$$

¿Existen sistemas 3×3 y 4×3 con las mismas soluciones? Dé ejemplos o diga por qué no existen.

¿Existe un sistema 4×3 con las mismas soluciones? Dé ejemplos o diga por qué no existen.

4. Encuentre un sistema lineal 4×4 cuya solución sea

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} + s \begin{pmatrix} 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} + t \begin{pmatrix} 7 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \quad s, t \in \mathbb{R}.$$

5. Considere el sistema lineal

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= b_1, \\ 3x_1 - x_2 + 2x_3 &= b_2, \\ 4x_1 + x_2 + x_3 &= b_3. \end{aligned}$$

Encuentre todo los posibles b_1, b_2, b_3 , o diga por qué no hay, para que el sistema tenga

- (a) exactamente una solución,
- (b) ninguna solución,
- (c) infinitas soluciones.

6. Calcule todas las posibles combinaciones “matriz por vector”:

$$A = \begin{pmatrix} 1 & 0 & 3 & 6 \\ 4 & 8 & 1 & 0 \\ 1 & 4 & 4 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 4 & 8 \\ 1 & 4 \\ 5 & -4 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 3 & 6 \\ 4 & 1 & 0 \\ 1 & 4 & 3 \end{pmatrix}, \quad D = \begin{pmatrix} -1 & 2 & 7 \\ 3 & -2 & 2 \end{pmatrix},$$

$$\vec{r} = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 6 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} 1 \\ 4 \\ 3 \\ 5 \\ -1 \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} 4 \\ 3 \\ 5 \\ -1 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} -3 \\ 5 \end{pmatrix}, \quad \vec{z} = \begin{pmatrix} 1 \\ -2 \\ \pi \end{pmatrix}.$$

7. Sean $A = \begin{pmatrix} 2 & 6 & -1 \\ 1 & -2 & 2 \\ 1 & 2 & -2 \end{pmatrix}$ y $\vec{b} = \begin{pmatrix} 17 \\ 6 \\ 4 \end{pmatrix}$. Encuentre todos los vectores $\vec{x} \in \mathbb{R}^3$ tal que $A\vec{x} = \vec{b}$.

8. Sea $M = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$.

(a) Demuestre que no existe $\vec{y} \neq 0$ tal que $M\vec{y} \perp \vec{y}$.

(b) Encuentre todos los vectores $\vec{x} \neq 0$ tal que $M\vec{x} \parallel \vec{x}$. Para cada tal $\vec{x}$, encuentre $\lambda \in \mathbb{R}$ tal que $M\vec{x} = \lambda\vec{x}$.

9. Calcule todas las posibles combinaciones (matriz)(matriz):

$$A = \begin{pmatrix} 1 & 0 & 3 & 6 \\ 4 & 8 & 1 & 0 \\ 1 & 4 & 4 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 4 & 8 \\ 1 & 4 \\ 5 & -4 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 3 & 6 \\ 4 & 1 & 0 \\ 1 & 4 & 3 \end{pmatrix},$$

$$D = \begin{pmatrix} -1 & 2 & 7 \\ 3 & -2 & 2 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 0 \\ 3 & 6 \end{pmatrix}.$$

10. Determine si las matrices son invertibles. Si lo son, encuentre su matriz inversa.

$$A = \begin{pmatrix} 1 & -2 \\ 2 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} -14 & 21 \\ 12 & -18 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 3 & 6 \\ 4 & 1 & 0 \\ 1 & 4 & 3 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 4 & 6 \\ 2 & 1 & 5 \\ 3 & 5 & 11 \end{pmatrix}.$$

11. Determine si las siguientes matrices son invertibles. Si lo son, encuentre su matriz inversa y escribela como producto de matrices elementales.

$$A = \begin{pmatrix} 1 & 0 \\ 3 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 2 \\ 8 & 6 \end{pmatrix}, \quad C = \begin{pmatrix} 4 & 10 \\ 6 & 15 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 3 & 6 \\ 4 & 1 & 0 \\ 1 & 4 & 3 \end{pmatrix}.$$

12. Una tienda vende dos tipos de cajitas de dulces:

Tipo A contiene 1 chocolate y 3 mentas, Tipo B contiene 2 chocolates y 1 menta.

(a) Dé una ecuación de la forma $A\vec{x} = \vec{b}$ que describe lo de arriba. Diga que significan los vectores $\vec{x}$ y $\vec{b}$.

(b) Calcule, usando el resultado de (a), cuantos chocolates y cuantas mentas contienen:

(i) 1 caja de tipo A y 3 de tipo B,

(iii) 2 caja de tipo A y 6 de tipo B,

(ii) 4 cajas de tipo A y 2 de tipo B,

(iv) 3 cajas de tipo A y 5 de tipo B.

(c) Determine si es posible conseguir

(i) 5 chocolates y 15 mentas,

(iii) 21 chocolates y 23 mentas,

(ii) 2 chocolates y 11 mentas,

(iv) 14 chocolates y 19 mentas.

comprando cajitas de dulces en la tienda. Si es posible, diga cuántos de cada tipo se necesitan.

13. Sea $A_k = \begin{pmatrix} 1 & 3 \\ 2 & k \end{pmatrix}$ y considere la ecuación

$$A_k \vec{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (*)$$

- (a) Encuentre todos los $k \in \mathbb{R}$ tal que $(*)$ tiene exactamente una solución para $\vec{x}$.
 - (b) Encuentre todos los $k \in \mathbb{R}$ tal que $(*)$ tiene infinitas soluciones para $\vec{x}$.
 - (c) Encuentre todos los $k \in \mathbb{R}$ tal que $(*)$ tiene ninguna solución para $\vec{x}$.
 - (d) Haga lo mismo para $A_k \vec{x} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ en vez de $(*)$.
 - (e) Haga lo mismo para $A_k \vec{x} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ en vez de $(*)$ donde $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ es un vector arbitrario distinto de $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.
14. Escriba las matrices invertibles de los Ejercicios 10. y 11. como producto de matrices elementales.
15. Para las siguientes matrices encuentre matrices elementales $E_1, \dots, E_n$ tal que $E_1 \cdot E_2 \cdot \dots \cdot E_n A$ es de la forma triangular superior.
- $$A = \begin{pmatrix} 7 & 4 \\ 3 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 4 & -4 \\ 2 & 1 & 0 \\ 3 & 5 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 0 \\ 2 & 4 & 3 \end{pmatrix}.$$
16. Sea $A \in M(m \times n)$ y sean $\vec{x}, \vec{y} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$. Demuestre que $A(\vec{x} + \lambda \vec{y}) = A\vec{x} + \lambda A\vec{y}$.
17. Demuestre que el espacio $M(m \times n)$ es un espacio vectorial con la suma de matrices y producto con $\lambda \in \mathbb{R}$ usual.
18. Sea $A \in M(n \times n)$.
- (a) Demuestre que $\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A^t \vec{y} \rangle$ para todo $\vec{x} \in \mathbb{R}^n$.
 - (b) Demuestre que $\langle AA^t \vec{x}, \vec{x} \rangle \geq 0$ para todo $\vec{x} \in \mathbb{R}^n$.
19. Sea $A = (a_{ij})_{\substack{i=1, \dots, n \\ j=1, \dots, m}} \in M(m \times n)$ y sea $\vec{e}_k$ el k -ésimo vector unitario en $\mathbb{R}^n$ (es decir, el vector en $\mathbb{R}^n$ cuya k -ésima entrada es 1 y las demás son cero). Calcule $A\vec{e}_k$ para todo $k = 1, \dots, n$ y describa en palabras la relación del resultado con la matriz A .

20. (a) Sea $A \in M(m \times n)$ y suponga que $A\vec{x} = \vec{0}$ para todo $\vec{x} \in \mathbb{R}^n$. Demuestre que $A = 0$ (la matriz cuyas entradas son 0).
 (b) Sea $x \in \mathbb{R}^n$ y suponga que $A\vec{x} = \vec{0}$ para todo $A \in M(n \times n)$. Demuestre que $\vec{x} = \vec{0}$.
 (c) Encuentre una matriz $A \in M(2 \times 2)$ y $\vec{v} \in \mathbb{R}^2$, ambos distintos de cero, tal que $A\vec{v} = \vec{0}$.
 (d) Encuentre matrices $A, B \in M(2 \times 2)$ tal que $AB = 0$ y $BA \neq 0$.

21. Sean $\vec{v} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$ y $\vec{w} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$.

- (a) Encuentre una matriz $A \in M(2 \times 2)$ que mapea el vector $\vec{e}_1$ a $\vec{v}$ y el vector $\vec{e}_2$ a $\vec{w}$.
 (b) Encuentre una matriz $B \in M(2 \times 2)$ que mapea el vector $\vec{v}$ a $\vec{e}_1$ y el vector $\vec{w}$ a $\vec{e}_2$.

22. Sean $A \in M(m, n)$, $B, C \in M(n, k)$, $D \in M(k, l)$.

- (a) Demuestre que $A(B + C) = AB + AC$.
 (b) Demuestre que $A(BD) = (AB)D$.

23. Sean $R, S \in M(n, n)$ matrices invertibles. Demuestre que

$$RS = SR \iff R^{-1}S^{-1} = S^{-1}R^{-1}.$$

24. Sean $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$ y $C = \begin{pmatrix} 9 & 6 \\ -7 & 11 \end{pmatrix}$. Encuentre $X \in M(2 \times 2)$ que satisfice la ecuación

$$AX + 3X - B = C.$$

25. ¿Falso o verdadero? Pruebe sus respuestas.

- (a) Si A es una matriz simétrica invertible, entonces A^{-1} es simétrica.
 (b) Si A, B son matrices simétricas, entonces AB es simétrica.
 (c) Si AB es una matriz simétrica, entonces A, B son matrices simétricas.
 (d) Si A, B son matrices simétricas, entonces $A + B$ es simétrica.
 (e) Si $A + B$ es una matriz simétrica, entonces A, B son matrices simétricas.
 (f) Si A es una matriz simétrica, entonces A^t es simétrica.
 (g) $AA^t = A^tA$.

26. Sea $A \in M(m \times n)$. Demuestre que AA^t y A^tA son matrices simétricas.

27. Sea $A \in M(n \times n)$. Demuestre que $A + A^t$ es simétrica y que $A - A^t$ es antisimétrica.

28. Sea $A \in M(n \times n)$ tal que $A^2 = \mathbb{O}$:
- (a) Muestre que $\text{id}_n - A$ es invertible y encuentre su inversa. (Indicación: Considere $\text{id}_n - A^2$. ¿Por qué en este caso es correcto factorizar por medio de diferencia de cuadrados? Ver Sección 3.4, Ejercicio 8.).
 - (b) Si $\lambda \in \mathbb{R}$, $\lambda \neq 0$, muestre que $\lambda \text{id}_n - A$ es invertible.
 - (c) ¿Cómo se pueden generalizar los incisos anteriores si en lugar de suponer que $A^2 = \mathbb{O}$ suponemos que $A^m = \mathbb{O}$ para algún $m \in \mathbb{N}$?
29. Calcule $(S_j(c))^t, (Q_{ij}(c))^t, (P_{ij})^t$.
30. (a) Sea $P_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in M(2 \times 2)$. Demuestre que P_{12} se deja expresar como producto de matrices elementales de la forma $Q_{ij}(c)$ y $S_k(c)$.
- (b) Pruebe el caso general: Sea $P_{ij} \in M(n \times n)$. Demuestre que P_{ij} se deja expresar como producto de matrices elementales de la forma $Q_{kl}(c)$ y $S_m(c)$.

Observación: El ejercicio demuestra que en verdad solo hay dos tipos de matrices elementales ya que el tercero (las permutaciones) se dejan reducir a un producto apropiado de matrices de tipo $Q_{ij}(c)$ y $S_j(c)$.

Capítulo 4

Determinantes

En esta sección definiremos el determinante de matrices en $M(n \times n)$ para n arbitrario y reconoceremos el determinante para $n = 2$, definido en la Sección 1.2, como un caso especial de nuestra nueva definición. Discutiremos las propiedades principales del determinante y mostraremos que una matriz es invertible si y solo si su determinante es distinto de 0. También daremos una interpretación geométrica del determinante y tendremos una primera idea de su importancia en geometría y en la teoría de la integración. Finalmente usaremos el determinante para calcular la inversa de una matriz invertible y demostraremos la regla de Cramer.

4.1. Determinante de una matriz

Recordemos que en la Sección 1.2, en la página 20, definimos el determinante de una matriz 2×2 por

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Además, sabemos que una matriz A de 2×2 es invertible si y solo si su determinante es distinto de 0, porque ambas afirmaciones son equivalentes a que el sistema homogéneo asociado tenga solo la solución trivial.

En esta sección definiremos el determinante para matrices arbitrarias $n \times n$ y veremos que de nuevo el determinante nos dice si una matriz es invertible o no. Daremos varias fórmulas para el determinante. Como definición usaremos la fórmula de Leibniz porque, en contraste a la fórmula de Laplace en Teorema 4.6, no es recursiva. Primero necesitamos saber qué es una permutación.

Definición 4.1. Una *permutación* de un conjunto M es una biyección $M \rightarrow M$. El conjunto de todas las permutaciones del conjunto $M = \{1, \dots, n\}$ se denota por S_n . Los elementos $\sigma \in S_n$ se pueden escribir como

$$\left\{ \begin{array}{ccccc} 1 & 2 & \cdots & n-1 & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n-1) & \sigma(n) \end{array} \right\}.$$

El *signo* (o *paridad*) de una permutación $\sigma \in S_n$ es

$$\text{sign}(\sigma) = (-1)^{\#\text{inversiones de } \sigma}$$

donde una *inversión* de σ es un par $i < j$ con $\sigma(i) > \sigma(j)$.

Observe que S_n consta de $n!$ permutaciones.

Ejemplos 4.2. (I) S_2 consta de dos permutaciones:

σ	$\begin{Bmatrix} 1 & 2 \\ 1 & 2 \end{Bmatrix}$	$\begin{Bmatrix} 1 & 2 \\ 2 & 1 \end{Bmatrix}$
$\text{sign}(\sigma)$	1	-1

(II) S_3 consta de seis permutaciones:

σ	$\begin{Bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{Bmatrix}$	$\begin{Bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{Bmatrix}$	$\begin{Bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{Bmatrix}$	$\begin{Bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{Bmatrix}$	$\begin{Bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{Bmatrix}$	$\begin{Bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{Bmatrix}$
$\text{sign}(\sigma)$	1	1	1	-1	-1	-1

Por ejemplo, la segunda permutación tiene dos inversiones ($1 < 3$ pero $\sigma(1) > \sigma(3)$ y $2 < 3$ pero $\sigma(2) > \sigma(3)$), la tercera permutación tiene dos inversiones ($1 < 2$ pero $\sigma(1) > \sigma(2)$, $1 < 3$ pero $\sigma(1) > \sigma(3)$), etc.

Definición 4.3. Sea $A = (a_{ij})_{i,j=1,\dots,n} \in M(n \times n)$. Entonces su determinante se define por

$$\det A = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \quad (4.1)$$

La fórmula en la ecuación (4.1) se llama la *fórmula de Leibniz*.

Observación. Otra notación para el determinante es $|A|$.

Observación 4.4. Observe que, según la fórmula,

- (a) el determinante es una suma de $n!$ términos,
- (b) cada término es un producto de n componentes de A ,
- (c) en cada producto hay exactamente un factor de cada fila y de cada columna, y todos esos productos aparecen en la fórmula.

Así que, claramente, la fórmula de Leibniz es una pesadilla computacional ...

¡Igualdad de derechos para filas y columnas!

Muestre que

$$\det A = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}. \quad (4.2)$$

Esto significa: en lugar de poner la permutación en el índice de columna, también podemos ponerla en el índice de fila.

Veamos si esta nueva definición coincide con nuestra definición antigua en el caso $n = 2$. Para ello, calculamos

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \sum_{\sigma \in S_2} \text{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} = a_{11}a_{22} - a_{21}a_{12}$$

que coincide con nuestra definición antigua en Definición 1.12.

Ahora veamos qué nos da la fórmula en el caso $n = 3$. Usando nuestra tabla con las permutaciones en S_3 , encontramos

$$\begin{aligned} \det A &= \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \sum_{\sigma \in S_3} \text{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} a_{3\sigma(3)} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31}. \end{aligned} \quad (4.3)$$

Agrupemos ahora los términos con coeficientes de la primera fila de A .

$$\det A = a_{11}[a_{22}a_{33} - a_{23}a_{32}] - a_{12}[a_{21}a_{33} - a_{23}a_{31}] + a_{13}[a_{21}a_{32} - a_{22}a_{31}]. \quad (4.4)$$

Vemos que los términos entre corchetes son nuevamente determinantes:

- a_{11} se multiplica por el determinante de la matriz 2×2 que se obtiene de A al eliminar la fila 1 y la columna 1.
- a_{12} se multiplica por el determinante de la matriz 2×2 que se obtiene de A al eliminar la fila 1 y la columna 2.
- a_{13} se multiplica por el determinante de la matriz 2×2 que se obtiene de A al eliminar la fila 1 y la columna 3.

Si hubiéramos agrupado los términos por coeficientes de la segunda fila, habríamos obtenido algo similar: cada término a_{2j} estaría multiplicado por el determinante de la matriz 2×2 que se obtiene de A al eliminar la fila 2 y la columna j . Por supuesto, también podríamos agrupar los términos por coeficientes todos de la primera columna. Entonces la fórmula se convertiría en una suma de términos donde los a_{j1} están multiplicados por los determinantes de las matrices que se obtienen de A al eliminar la fila j y la columna 1.

Esto motiva la definición de los llamados *menores* de una matriz.

Definición 4.5. Sea $A = (a_{ij})_{i,j=1,\dots,n} \in M(n \times n)$ y sea M_{ij} la matriz $(n-1) \times (n-1)$ que se obtiene de A al eliminar la fila i y la columna j . Entonces los números $\det M_{ij}$ se llaman *menores de A* y $C_{ij} := (-1)^{i+j} \det(M_{ij})$ se llaman *cofactores de A* .

Observación. Algunos libros usan una notación ligeramente distinta. Llamamos *menores de A* (o *matrices menores de A*) a las matrices $(n-1) \times (n-1)$ M_{ij} que se obtienen de A al eliminar la fila i y la columna j de A . Sin embargo, parece que la mayoría de los textos usa la convención de la Definición 4.5.

Con estas definiciones podemos escribir (4.3) como

$$\det A = \sum_{j=1}^3 (-1)^{1+j} a_{1j} \det M_{1j} = \sum_{j=1}^3 a_{1j} C_{1j}.$$

Esta fórmula se llama la *expansión del determinante de A a lo largo de la primera fila*. También vimos que podemos expandir a lo largo de la segunda o tercera fila, o a lo largo de columnas, así que

$$\det A = \sum_{j=1}^3 (-1)^{k+j} a_{kj} \det M_{kj} = \sum_{j=1}^3 a_{kj} C_{kj}, \quad \text{para } k = 1, 2, 3,$$

$$\det A = \sum_{i=1}^3 (-1)^{i+k} a_{ik} \det M_{ik} = \sum_{i=1}^3 a_{ik} C_{ik}, \quad \text{para } k = 1, 2, 3.$$

La primera fórmula se llama *expansión a lo largo de la k -ésima fila*, y la segunda fórmula se llama *expansión a lo largo de la k -ésima columna*. Con un poco más de esfuerzo se puede mostrar que una fórmula análoga es verdadera para n arbitrario.

Teorema 4.6. *Sea $A = (a_{ij})_{i,j=1,\dots,n} \in M(n \times n)$ y sea M_{ij} la matriz de sus menores. Entonces*

$$\det A = \sum_{j=1}^n (-1)^{k+j} a_{kj} \det M_{kj} = \sum_{j=1}^n a_{kj} C_{kj}, \quad \text{para } k = 1, 2, \dots, n, \quad (4.5)$$

$$\det A = \sum_{i=1}^n (-1)^{i+k} a_{ik} \det M_{ik} = \sum_{i=1}^n a_{ik} C_{ik}, \quad \text{para } k = 1, 2, \dots, n. \quad (4.6)$$

Las fórmulas (4.5) y (4.6) se llaman *expansión de Laplace del determinante*. Más precisamente, (4.5) se llama *expansión a lo largo de la k -ésima fila*, (4.6) se llama *expansión a lo largo de la k -ésima columna*.

Demostración. Primero demostremos (4.5) para el caso $k = 1$. Sea $A \in M(n \times n)$. Observe que por definición $\det A$ es la suma de productos de la forma $\text{sign}(\sigma) a_{1\sigma(1)} \dots a_{n,\sigma(n)}$, véase la Observación 4.4. Así que los términos son exactamente todos los productos posibles de entradas de la matriz A con exactamente un término de cada fila y exactamente un término de cada columna, multiplicados por $+1$ o -1 . Lo mismo ocurre para la fórmula (4.5): a_{11} se multiplica por $\det A_{11}$, pero este último consta de productos con exactamente un factor en cada fila y cada columna de A_{11} , es decir, exactamente un factor de las filas 2 a n y de las columnas 2 a n de A ; a_{12} se multiplica por $\det A_{12}$, pero este último consta de productos con exactamente un factor en cada fila y cada columna de A_{12} , es decir, exactamente un factor de las filas 2 a n y de las columnas 1, 3, hasta n de A ; etc. Así, todos los productos que aparecen en la fórmula de Leibniz (4.1) aparecen también en la fórmula de Laplace (4.5) y viceversa. Por lo tanto, solo queda mostrar que aparecen con el mismo factor 1 o -1 en ambas fórmulas. Sea $\sigma \in S_n$ y pongamos $\tilde{\sigma} : \{2, 3, \dots, n\} \rightarrow \{1, 2, \dots, n\} \setminus \{\sigma(1)\}$, $\tilde{\sigma}(j) = \sigma(j)$. Entonces

$$\begin{aligned} \#(\text{inversiones de } \sigma) &= \#(\text{pares } (i, j) \text{ tales que } i < j \text{ y } \sigma(i) > \sigma(j)) \\ &= \#(\text{pares } (i, j) \text{ tales que } 2 \leq i < j \text{ y } \sigma(i) > \sigma(j)) \\ &\quad + \#(\text{pares } (1, j) \text{ tales que } 1 < j \text{ y } \sigma(1) > \sigma(j)) \\ &= \#(\text{inversiones de } \tilde{\sigma}) + \#(\text{pares } (1, j) \text{ tales que } 1 < j \text{ y } \sigma(1) > \sigma(j)) \\ &= \#(\text{inversiones de } \tilde{\sigma}) + \sigma(1) - 1, \end{aligned}$$

de donde $\text{sign}(\sigma) = (-1)^{\#(\text{inversiones de } \tilde{\sigma}) + \sigma(1) - 1} = \text{sign}(\tilde{\sigma})(-1)^{\sigma(1) - 1} = \text{sign}(\tilde{\sigma})(-1)^{\sigma(1) + 1}$ y por lo tanto

$$\text{sign}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)} \cdots a_{n,\sigma(n)} = (-1)^{\sigma(1)+1}a_{1\sigma(1)} [\text{sign}(\tilde{\sigma})a_{2\sigma(2)} \cdots a_{n,\sigma(n)}].$$

El término entre corchetes es uno de los términos que aparecen en el determinante de $A_{1\sigma(1)}$, luego el producto del lado derecho aparece en la fórmula (4.5) (cuando $j = \sigma(1)$) y es el único término que contiene los factores $a_{11}, a_{12}, \dots, a_{1n}$. En consecuencia, cada término de la fórmula de Leibniz aparece exactamente una vez en la fórmula de Laplace con exactamente el mismo signo, y no hay otros términos en la fórmula de Laplace. Por lo tanto, ambas fórmulas son iguales.

El razonamiento para $k > 1$ es el mismo. Solo necesitamos tomar $\tilde{\sigma}$ como la restricción de σ a $\{1, 2, \dots, n\} \setminus \{k\}$ y observar que $\text{sign}(\sigma) = \text{sign}(\tilde{\sigma})(-1)^{\sigma(k)+k}$. Esto es cierto porque

$$\begin{aligned} \#(\text{inversiones de } \sigma) &= \#(\text{pares } (i, j) \text{ tales que } i < j \text{ y } \sigma(i) > \sigma(j)) \\ &= \#(\text{pares } (i, j) \text{ tales que } i, j \neq k, 1 \leq i < j \text{ y } \sigma(i) > \sigma(j)) \\ &\quad + \#(\text{pares } (i, k) \text{ tales que } i < k \text{ o } j = k \text{ y } \sigma(i) > \sigma(k)) \\ &\quad + \#(\text{pares } (k, j) \text{ tales que } k < j, \text{ o } j = k \text{ y } \sigma(k) > \sigma(j)) \\ &= \#(\text{inversiones de } \tilde{\sigma}) + \#(\text{pares } (1, j) \text{ tales que } 1 < j \text{ y } \sigma(1) > \sigma(j)) \\ &= \#(\text{inversiones de } \tilde{\sigma}) + \sigma(k) - k + 2\ell. \end{aligned}$$

Por lo tanto obtenemos

$$\begin{aligned} \text{sign}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)} \cdots a_{n,\sigma(n)} \\ = (-1)^{\sigma(k)+k}a_{k\sigma(k)} [\text{sign}(\tilde{\sigma})a_{1\sigma(1)} \cdots a_{k-1,\sigma(k-1)}a_{k+1,\sigma(k+1)} \cdots a_{n1,\sigma(n)}]. \end{aligned}$$

El término entre corchetes aparece en el determinante de la submatriz $A_{k\sigma(k)}$, luego aparece en la suma de la expansión a lo largo de la k -ésima fila (en el término con $j = \sigma(k)$).

Para demostrar (4.6), podemos usar los mismos argumentos anteriores aplicados a la “versión por columnas” (4.2) de la fórmula de Leibniz. $\square$

Observe que para calcular, por ejemplo, el determinante de una matriz 5×5 , tenemos que calcular cinco determinantes 4×4 , para cada uno de los cuales tenemos que calcular cuatro determinantes (3×3) , etc. Computacionalmente, es tan largo como la fórmula de Leibniz, pero al menos no tenemos que encontrar primero todas las permutaciones en S_n .

Más adelante veremos cómo calcular el determinante usando eliminación gaussiana. Esto es computacionalmente mucho más eficiente; véase la Observación 4.12.

Ejemplo 4.7. Usamos expansión a lo largo de la segunda columna para calcular

$$\begin{aligned} \det \begin{pmatrix} 3 & 2 & 1 \\ 5 & 6 & 4 \\ 8 & 0 & 7 \end{pmatrix} &= -2 \det \begin{pmatrix} 3 & 1 \\ 5 & 4 \\ 8 & 7 \end{pmatrix} + 6 \det \begin{pmatrix} 3 & 1 \\ 5 & 4 \\ 8 & 7 \end{pmatrix} - 0 \det \begin{pmatrix} 3 & 1 \\ 5 & 4 \\ 8 & 7 \end{pmatrix} \\ &= -2 \det \begin{pmatrix} 5 & 4 \\ 8 & 7 \end{pmatrix} + 6 \det \begin{pmatrix} 3 & 1 \\ 8 & 7 \end{pmatrix} - 0 \det \begin{pmatrix} 3 & 1 \\ 5 & 4 \end{pmatrix} \\ &= -2[5 \cdot 7 - 4 \cdot 8] + 6[3 \cdot 7 - 1 \cdot 8] = -2[35 - 32] + 6[21 - 8] = -6 + 78 = 72. \end{aligned}$$

Obtenemos el mismo resultado si expandimos el determinante, por ejemplo, a lo largo de la primera fila:

$$\begin{aligned}
 \det \begin{pmatrix} 3 & 2 & 1 \\ 5 & 6 & 4 \\ 8 & 0 & 7 \end{pmatrix} &= 3 \det \begin{pmatrix} 5 & 6 & 4 \\ 8 & 0 & 7 \end{pmatrix} - 2 \det \begin{pmatrix} 3 & 6 & 4 \\ 8 & 0 & 7 \end{pmatrix} + 1 \det \begin{pmatrix} 3 & 2 & 4 \\ 5 & 6 & 7 \end{pmatrix} \\
 &= 3 \det \begin{pmatrix} 6 & 4 \\ 0 & 7 \end{pmatrix} - 2 \det \begin{pmatrix} 5 & 4 \\ 8 & 7 \end{pmatrix} + 1 \det \begin{pmatrix} 5 & 6 \\ 8 & 0 \end{pmatrix} \\
 &= 3[6 \cdot 7 - 4 \cdot 0] - 2[5 \cdot 7 - 4 \cdot 8] + [5 \cdot 0 - 6 \cdot 8] \\
 &= 3 \cdot 42 - 2[35 - 32] - 40 = 126 - 6 - 48 = 72.
 \end{aligned}$$

Ejemplo 4.8. Damos un ejemplo de calcular el determinante de una matriz 4×4 . Las flechas en rojo indican en cuál fila o columna expandimos.

$$\begin{aligned}
 \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 6 & 0 & 1 \\ 2 & 0 & 7 & 0 \\ 0 & 3 & 0 & 1 \end{pmatrix} &= \det \begin{pmatrix} 6 & 0 & 1 \\ 0 & 7 & 0 \\ 3 & 0 & 1 \end{pmatrix} - 2 \det \begin{pmatrix} 0 & 0 & 1 \\ 2 & 7 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 3 \det \begin{pmatrix} 0 & 6 & 1 \\ 2 & 0 & 0 \\ 0 & 3 & 1 \end{pmatrix} - 4 \det \begin{pmatrix} 0 & 6 & 0 \\ 2 & 0 & 7 \\ 0 & 3 & 0 \end{pmatrix} \\
 &= 7 \det \begin{pmatrix} 6 & 1 \\ 3 & 1 \end{pmatrix} - 2 \left[7 \det \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \right] + 3 \left[-2 \det \begin{pmatrix} 6 & 1 \\ 3 & 1 \end{pmatrix} \right] - 4 \left[-6 \det \begin{pmatrix} 2 & 7 \\ 0 & 0 \end{pmatrix} \right] \\
 &= 7[6 - 3] - 14[0 - 0] - 6[6 - 3] + 24[0 - 0] = 21 - 18 = 3.
 \end{aligned}$$

Ahora calculamos el determinante de la misma matriz pero escogimos una fila con más ceros en el primer paso. La ventaja es que hay solo dos menores 3×3 cuyos determinantes tenemos que calcular (los otros dos se multiplican por cero, razón por la cuál no tenemos que calcularlos).

$$\begin{aligned}
 \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 6 & 0 & 1 \\ 2 & 0 & 7 & 0 \\ 0 & 3 & 0 & 1 \end{pmatrix} &= -0 \det \begin{pmatrix} 2 & 3 & 4 \\ 0 & 7 & 0 \\ 3 & 0 & 1 \end{pmatrix} + 6 \det \begin{pmatrix} 1 & 3 & 4 \\ 2 & 7 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 0 \det \begin{pmatrix} 1 & 2 & 4 \\ 2 & 0 & 0 \\ 0 & 3 & 1 \end{pmatrix} + \det \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 7 \\ 0 & 3 & 0 \end{pmatrix} \\
 &= 6 \left[-3 \det \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} + 7 \det \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix} \right] + \left[\det \begin{pmatrix} 0 & 7 \\ 3 & 0 \end{pmatrix} - 2 \det \begin{pmatrix} 2 & 3 \\ 3 & 0 \end{pmatrix} \right] \\
 &= 6[-6 + 7] + [-21 + 18] = 6 - 3 = 3.
 \end{aligned}$$

Regla de Sarrus

Terminamos esta sección con la llamada *regla de Sarrus*. De (4.3) sabemos que

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - [a_{12}a_{21}a_{33} + a_{11}a_{22}a_{33} + a_{13}a_{22}a_{31}]$$

lo cual se puede memorizar como sigue: Escriba la matriz A y añádale su primera y segunda columna. Luego sumamos los productos de los tres términos que están en diagonales de arriba a la izquierda hacia abajo a la derecha, y restamos los productos de los términos que están en diagonales de arriba a la derecha hacia abajo a la izquierda, como en la siguiente imagen:

$$\det A = \underline{a_{11}a_{22}a_{33}} + \underline{a_{12}a_{23}a_{31}} + \underline{a_{13}a_{21}a_{32}} - [a_{13}a_{22}a_{31} + a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33}].$$

¡La regla de Sarrus funciona solo para matrices 3×3 !

Convénzase de que también se podrían añadir la primera y la segunda fila debajo de la matriz y hacer cruces.

Ejemplo 4.9 (Regla de Sarrus).

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 8 & 7 \end{pmatrix} &= 1 \cdot 5 \cdot 7 + 2 \cdot 6 \cdot 0 + 3 \cdot 4 \cdot 8 - [3 \cdot 5 \cdot 0 + 6 \cdot 8 \cdot 1 + 7 \cdot 2 \cdot 4] \\ &= 35 + 96 - [48 + 56] = 131 - 106 = 27. \end{aligned}$$

Ahora debería haber entendido

- qué es una permutación,
- cómo derivar la fórmula de expansión de Laplace a partir de la fórmula de Leibniz,
- etc.

Ahora debería ser capaz de

- calcular el determinante de una matriz $n \times n$,
- etc.

Ejercicios.

1. Calcule los determinantes de las siguientes matrices:

$$(a) \begin{pmatrix} -2 & 3 & 1 \\ 0 & 2 & 1 \\ 4 & 6 & 5 \end{pmatrix}$$

$$(b) \begin{pmatrix} 0 & 1 & 4 \\ -2 & 0 & -6 \\ 2 & 1 & 0 \end{pmatrix}$$

$$(c) \begin{pmatrix} 6 & 3 & 5 \\ 3 & -1 & 4 \\ -2 & 1 & -6 \end{pmatrix}$$

$$(d) \begin{pmatrix} -1 & 1 & 0 \\ 2 & 1 & 4 \\ 1 & 5 & 6 \end{pmatrix}$$

$$(e) \begin{pmatrix} 3 & 5 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 2 \end{pmatrix}$$

$$(f) \begin{pmatrix} -3 & 0 & 0 & 0 \\ -4 & 7 & 0 & 0 \\ -5 & 10 & -1 & 0 \\ 2 & 3 & -11 & 6 \end{pmatrix}$$

$$(g) \begin{pmatrix} 2 & 3 & -1 & 20 & \pi \\ 0 & 1 & \sqrt{2} & -11 & 10 \\ 0 & 0 & 4 & -1 & 5 \\ 0 & 0 & 0 & -2 & 50 \\ 0 & 0 & 0 & 0 & 6 \end{pmatrix}$$

$$(h) \begin{pmatrix} -2 & 0 & 0 & 7 \\ 1 & 2 & -1 & 4 \\ 3 & 0 & -1 & 5 \\ 4 & 2 & 3 & 0 \end{pmatrix}$$

2. Sea $A = \begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix}$. Muestre que $\det A = (b-a)(c-a)(c-b)$.

3. Sea $A = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$. Muestre que $\det A = 0$.

4.2. Propiedades del determinante

En esta sección mostraremos propiedades del determinante y demostraremos que una matriz es invertible si y solo si su determinante es distinto de 0.

(D1) El determinante es lineal en sus filas.

Esto significa lo siguiente. Sean $\vec{r}_1, \dots, \vec{r}_n$ los vectores fila de la matriz A y supongamos que $\vec{r}_j = \vec{s}_j + \gamma \vec{t}_j$. Entonces

$$\det A = \det \begin{pmatrix} \vec{r}_1 \\ \vdots \\ \vec{r}_j \\ \vdots \\ \vec{r}_n \end{pmatrix} = \det \begin{pmatrix} \vec{r}_1 \\ \vdots \\ \vec{s}_j + \gamma \vec{t}_j \\ \vdots \\ \vec{r}_n \end{pmatrix} = \det \begin{pmatrix} \vec{r}_1 \\ \vdots \\ \vec{s}_j \\ \vdots \\ \vec{r}_n \end{pmatrix} + \gamma \det \begin{pmatrix} \vec{r}_1 \\ \vdots \\ \vec{t}_j \\ \vdots \\ \vec{r}_n \end{pmatrix}.$$

Esto se demuestra fácilmente expandiendo el determinante a lo largo de la fila j -ésima, o también se puede ver a partir de la fórmula de Leibniz.

(D1') El determinante es lineal en sus columnas.

Esto significa lo siguiente. Sean $\vec{c}_1, \dots, \vec{c}_n$ los vectores columna de la matriz A y supongamos que $\vec{c}_j = \vec{s}_j + \gamma \vec{t}_j$. Entonces

$$\begin{aligned} \det A &= \det(\vec{c}_1 | \dots | \vec{c}_j | \dots | \vec{c}_n) = \det(\vec{c}_1 | \dots | \vec{s}_j + \gamma \vec{t}_j | \dots | \vec{c}_n) \\ &= \det(\vec{c}_1 | \dots | \vec{s}_j | \dots | \vec{c}_n) + \gamma \det(\vec{c}_1 | \dots | \vec{t}_j | \dots | \vec{c}_n). \end{aligned}$$

Esto se demuestra fácilmente expandiendo el determinante a lo largo de la columna j -ésima, o también se puede ver a partir de la fórmula de Leibniz.

(D2) El determinante es alternante en sus filas.

Si se intercambian dos filas de una matriz, entonces el determinante cambia de signo. Esto significa: Sean $\vec{r}_1, \dots, \vec{r}_n$ los vectores fila de la matriz A y $i \neq j \in \{1, \dots, n\}$. Entonces

$$\det A = \det \begin{pmatrix} \vdots \\ \vec{r}_j \\ \vdots \\ \vec{r}_i \\ \vdots \end{pmatrix} = - \det \begin{pmatrix} \vdots \\ \vec{r}_i \\ \vdots \\ \vec{r}_j \\ \vdots \end{pmatrix}.$$

Esto es fácil de ver cuando las dos filas que se van a intercambiar son adyacentes. Por ejemplo, supongamos que $j = i + 1$. Sea A la matriz original y sea B la matriz con las filas i e $i + 1$ intercambiadas. Expandimos el determinante de A a lo largo de la fila i -ésima y el determinante de B a lo largo de la fila $(i + 1)$ -ésima. Note que en ambos casos los menores son iguales, es decir, $M_{ik}^A = M_{(i+1)k}^B$ (usamos los superíndices A y B para distinguir los menores de A y de B). Notamos que

$$\det B = \sum_{k=1}^n (-1)^{(i+1)+k} M_{(i+1)k}^B = \sum_{k=1}^n (-1)(-1)^{i+k} M_{ik}^A = - \sum_{k=1}^n (-1)^{i+k} M_{ik}^A = - \det A.$$

Esto se puede ver también con la fórmula de Leibniz.

Veamos ahora qué ocurre si i y j no son filas adyacentes. Sin pérdida de generalidad podemos suponer que $i < j$. Entonces primero intercambiamos la fila j -ésima ($j - i$) veces con la fila superior hasta que quede en la fila i -ésima. La fila i -ésima original está ahora en la fila $(i + 1)$. Ahora la intercambiamos hacia abajo con sus filas vecinas hasta que se convierta en la fila j . Para hacerlo, necesitamos $j - (i + 1)$ intercambios. En total intercambiamos $[j - i] + [j - (i + 1)] = 2j - 2i + 1$ veces filas adyacentes, así que el determinante de la matriz nueva es

$$\underbrace{(-1) \cdot (-1) \cdot \dots \cdot (-1)}_{2j-2i+1 \text{ veces (un factor por cada intercambio)}} \cdot \det A = (-1)^{2j-2i+1} \det A = - \det A.$$

(D2') El determinante es alternante en sus columnas.

Si se intercambian dos columnas de una matriz, entonces el determinante cambia de signo. Esto significa: Sean $\vec{c}_1, \dots, \vec{c}_n$ los vectores columna de la matriz A y $i \neq j \in \{1, \dots, n\}$. Entonces

$$\det A = \det(\dots | \vec{c}_i | \dots | \vec{c}_j | \dots) = - \det(\dots | \vec{c}_j | \dots | \vec{c}_i | \dots).$$

Esto se sigue de la misma manera que la propiedad alternante para filas.

(D3) $\det \text{id}_n = 1$.

La expansión en la primera fila muestra

$$\det \text{id}_n = 1 \det \text{id}_{n-1} = 1^2 \det \text{id}_{n-2} = \cdots = 1^n = 1.$$

Observación 4.10. Se puede mostrar: Toda función $f : M(n \times n) \rightarrow \mathbb{R}$ que satisface **(D1)**, **(D2)** and **(D3)** (o **(D1')**, **(D2')** and **(D3)**) necesariamente es $\det$.

Veamos ahora algunas propiedades más del determinante.

(D4) $\det A = \det A^t$.

Esto se sigue fácilmente de la fórmula de Leibniz o de la expansión de Laplace (si se expande A a lo largo de la primera fila y A^t a lo largo de la primera columna, se obtienen exactamente los mismos términos). Esto también muestra que **(D1')** se sigue de **(D1)** y que **(D2')** se sigue de **(D2)**, y viceversa.

(D5) Si una fila de A es múltiplo de otra fila, o si una columna es múltiplo de otra columna, entonces $\det A = 0$. En particular, si A tiene dos filas iguales o dos columnas iguales, entonces $\det A = 0$.

Sean $\vec{r}_1, \dots, \vec{r}_n$ las filas de la matriz A y supongamos que $\vec{r}_k = c\vec{r}_j$. Entonces

$$\begin{aligned} \det A &= \det \begin{pmatrix} \vdots \\ \vec{r}_k \\ \vdots \\ \vec{r}_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ c\vec{r}_j \\ \vdots \\ \vec{r}_j \\ \vdots \end{pmatrix} \stackrel{(D2)}{=} -\det \begin{pmatrix} \vdots \\ \vec{r}_j \\ \vdots \\ c\vec{r}_j \\ \vdots \end{pmatrix} \stackrel{(D1)}{=} -c \det \begin{pmatrix} \vdots \\ \vec{r}_j \\ \vdots \\ \vec{r}_j \\ \vdots \end{pmatrix} \\ &\stackrel{(D1)}{=} -\det \begin{pmatrix} \vdots \\ c\vec{r}_j \\ \vdots \\ \vec{r}_j \\ \vdots \end{pmatrix} = -\det \begin{pmatrix} \vdots \\ \vec{r}_k \\ \vdots \\ \vec{r}_j \\ \vdots \end{pmatrix} = -\det A. \end{aligned}$$

Esto demuestra que $\det A = -\det A$, y consecuentemente $\det A = 0$. Si A tiene una columna que es múltiplo de otra, entonces su transpuesta tiene una fila que es múltiplo de otra fila y, con ayuda de **(D4)**, se sigue que $\det A = \det A^t = 0$.

(D6) El determinante de una matriz triangular superior o inferior es el producto de sus entradas diagonales.

Sea A una matriz triangular superior y expandamos su determinante en la primera columna. Entonces solo el primer término de la expansión de Laplace es distinto de 0, porque todos los coeficientes de la primera columna son iguales a 0, excepto posiblemente el de la primera fila. Repetimos esto y obtenemos

$$\begin{aligned} \det A &= \det \begin{pmatrix} c_1 & & & \\ & c_2 & & * \\ & & \ddots & \\ 0 & & & c_n \end{pmatrix} = c_1 \det \begin{pmatrix} c_2 & & & \\ & c_3 & & * \\ & & \ddots & \\ 0 & & & c_n \end{pmatrix} = c_1 c_2 \det \begin{pmatrix} c_3 & & & \\ & c_4 & & * \\ & & \ddots & \\ 0 & & & c_n \end{pmatrix} \\ &= \cdots = c_1 c_2 \cdots c_{n-2} \det \begin{pmatrix} c_{n-1} & 0 \\ 0 & c_n \end{pmatrix} = c_1 c_2 \cdots c_{n-1} c_n. \end{aligned}$$

La afirmación para matrices triangulares inferiores se sigue de (D4) y de lo que acabamos de mostrar, porque la transpuesta de una matriz triangular superior es triangular inferior y las entradas diagonales son las mismas. Or we could repeat the above proof but this time we would expand always in the first row (or last column).

A continuación calculamos el determinante de matrices elementales.

(D7) El determinante de matrices elementales.

- (I) $\det S_j(c) = c$,
- (II) $\det Q_{ij}(c) = 1$,
- (III) $\det P_{ij} = -1$.

La afirmación sobre $S_j(c)$ y $Q_{ij}(c)$ se sigue de (D6), ya que son matrices triangulares. La afirmación para P_{ij} se sigue de (D2) y (D3), porque al intercambiar la fila i y la fila j en P_{ij} obtenemos la matriz identidad, de modo que $\det P_{ij} = -\det \text{id} = -1$.

Ahora calculamos el determinante del producto de una matriz elemental con otra matriz.

(D8) Sea E una matriz elemental y sea $A \in M(n \times n)$. Entonces $\det(EA) = \det E \det A$.

Sea E una matriz elemental y denotemos las filas de A por $\vec{r}_1, \dots, \vec{r}_n$. Tenemos que distinguir entre los tres tipos diferentes de matrices elementales.

Caso 1. $E = S_j(c)$. Sabemos por (D6) que $\det E = \det S_j(c) = c$. Usando la Proposición 3.62 y (D1), encontramos que

$$\det(EA) = \det(S_j(c)A) = \det \begin{pmatrix} \vdots \\ c\vec{r}_j \\ \vdots \end{pmatrix} = c \det \begin{pmatrix} \vdots \\ \vec{r}_j \\ \vdots \end{pmatrix} = c \det A = \det S_j(c) \det A.$$

Caso 2. $E = Q_{ij}(c)$. Sabemos por (D6) que $\det E = \det Q_{ij}(c) = 1$. Usando la Proposición 3.62, (D1) y (D5), encontramos que

$$\begin{aligned} \det(EA) &= \det(Q_{ij}(c)A) = \det \begin{pmatrix} \vdots \\ \vec{r}_i + c\vec{r}_j \\ \vdots \\ \vec{r}_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ \vec{r}_i \\ \vdots \\ \vec{r}_j \\ \vdots \end{pmatrix} + c \det \begin{pmatrix} \vdots \\ \vec{r}_j \\ \vdots \\ \vec{r}_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ \vec{r}_i \\ \vdots \\ \vec{r}_j \\ \vdots \end{pmatrix} \\ &= \det A = \det Q_{ij}(c) \det A. \end{aligned}$$

Caso 3. $E = P_{ij}$. Sabemos por (D6) que $\det E = \det P_{jk} = -1$. Usando la Proposición 3.62 y (D2), encontramos que

$$\begin{aligned} \det(EA) &= \det(P_{jk}A) = \det \begin{pmatrix} P_{jk} & \begin{pmatrix} \vdots \\ \vec{r}_j \\ \vdots \\ \vec{r}_k \\ \vdots \end{pmatrix} \end{pmatrix} = \det \begin{pmatrix} \vdots \\ \vec{r}_k \\ \vdots \\ \vec{r}_j \\ \vdots \end{pmatrix} = -\det \begin{pmatrix} \vdots \\ \vec{r}_j \\ \vdots \\ \vec{r}_k \\ \vdots \end{pmatrix} \\ &= -\det A = \det P_{jk} \det A. \end{aligned}$$

Si repetimos (D8), obtenemos

$$\det(E_1 \cdots E_k A) = \det(E_1) \cdots \det(E_k) \det(A)$$

for elementary matrices $E_1, \dots, E_k$.

(D9) Sea $A \in M(n \times n)$. Entonces A es invertible si y solo si $\det A \neq 0$.

Sea A' la forma escalonada reducida por filas de A . Por la Proposición 3.67 existen matrices elementales $E_1, \dots, E_k$ tal que $A = E_1 \cdots E_k A'$, y por ende

$$\det A = \det(E_1 \cdots E_k A') = \det(E_1) \cdots \det(E_k) \det A'. \quad (4.7)$$

Recordemos que el determinante de una matriz elemental es distinto de cero; por eso (4.7) muestra que $\det A = 0$ si y solo si $\det A' = 0$.

Si A es invertible, entonces $A' = \text{id}$, luego $\det A' = 1 \neq 0$ y por lo tanto también $\det A \neq 0$. Si A no es invertible, entonces la última fila de A' debe ser cero, luego $\det A' = 0$ y por lo tanto también $\det A = 0$.

A continuación mostramos que el determinante es multiplicativo.

(D10) Sean $A, B \in M(n \times n)$. **Entonces** $\det(AB) = \det A \det B$.

Como antes, sea A' la forma escalonada reducida de A . Por la Proposición 3.67 existen matrices elementales $E_1, \dots, E_k$ tal que $A = E_1 \cdots E_k A'$. De (D9) se sigue que

$$\det(AB) = \det(E_1 \cdots E_k A' B) = \det(E_1) \cdots \det(E_k) \det(A' B). \quad (4.8)$$

Si A es invertible, entonces $A' = \text{id}$ y (4.7) muestra que

$$\det(AB) = \det(E_1) \cdots \det(E_k) \det(B) = \det(E_1 \cdots E_k) \det(B) = \det(A) \det(B).$$

Si, por otro lado, A no es invertible, entonces $\det A = 0$. Además, la última fila de A' es cero, entonces también la última fila de $A'B$ es cero, por lo tanto $A'B$ no es invertible y consecuentemente $\det A'B = 0$. Luego $\det(AB) = 0$ por (4.7), y también $\det(A) \det(B) = 0 \det(B) = 0$, entonces también en este caso tenemos que $\det(AB) = \det A \det B$.

(D11) Sea $A \in M(n \times n)$ **una matriz invertible. Entonces** $\det(A^{-1}) = (\det A)^{-1}$.

Si A es invertible, entonces $\det A \neq 0$ y se sigue de (D10) que

$$1 = \det(\text{id}_n) = \det(AA^{-1}) = \det(A) \det(A^{-1}).$$

Despejando $\det(A^{-1})$ obtenemos la fórmula deseada.

Sea $A \in M(n \times n)$. Dé dos demostraciones de $\det(cA) = c^n \det A$ usando una de las siguientes ideas:

- (I) Apply (D1) or (D1') n times.
- (II) Utilice que $cA = \text{diag}(c, c, \dots, c)A$ y aplique (D10) y (D6).

¡El determinante no es aditivo!

Recordemos que $\det(AB) = \det A \det B$. Pero en general

$$\det(A + B) \neq \det A + \det B.$$

Por ejemplo, si $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, entonces $\det A + \det B = 0 + 0 = 0$, but $\det(A + B) = \det \text{id}_2 = 1$.

El siguiente teorema es el Teorema 3.44 junto con (D9).

Teorema 4.11. Sea $A \in M(n \times n)$. Entonces las siguientes afirmaciones son equivalentes:

- (I) A es invertible.
- (II) Para todo $\vec{b} \in \mathbb{R}^n$, la ecuación $A\vec{x} = \vec{b}$ tiene exactamente una solución.
- (III) La ecuación $A\vec{x} = \vec{0}$ tiene exactamente una solución.
- (IV) Toda forma escalonada reducida por filas de A tiene n pivotes.

(v) A es equivalente por filas a id_n .

(vi) $\det A \neq 0$.

On the computational complexity of the determinant.

Observación 4.12. Las propiedades anteriores proporcionan una manera eficiente de calcular el determinante de una matriz $n \times n$. Note que ambas la fórmula de Leibniz y la expansión de Laplace requieren $O(n!)$ pasos ($O(n!)$ significa el “de orden de $n!$ ”. Se puede pensar de ello como “aproximadamente $n!$ ” o “salvo un múltiplo constante es más o menos igual a $n!$ ”. Algo como $O(2(n!))$ todavía es lo mismo que $O(n!)$).

En contraste, reducir una matriz con la eliminación de Gauß-Jordan requiere solamente $O(n^3)$ pasos hasta que llegamos a una forma escalonada. Dado que ella siempre es una matriz triangular superior, su determinante es fácil de calcular.

Si n es grande, entonces n^3 también es grande, pero $n!$ es mucho más grande; por eso la eliminación de Gauß-Jordan es computacionalmente mucho más eficiente que la fórmula de Leibniz o la expansión de Laplace.

Ilustremos esto con un ejemplo.

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 6 \\ 1 & 7 & 8 & 9 \\ 1 & 5 & 3 & 4 \end{pmatrix} &\stackrel{\textcircled{1}}{=} \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 2 \\ 0 & 5 & 5 & 5 \\ 0 & 3 & 0 & 0 \end{pmatrix} \stackrel{\textcircled{2}}{=} 5 \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 3 & 0 & 0 \end{pmatrix} \stackrel{\textcircled{3}}{=} 5 \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 3 & 0 & 0 \end{pmatrix} \\ &\stackrel{\textcircled{4}}{=} 5 \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 3 & 0 & 0 \end{pmatrix} \stackrel{\textcircled{5}}{=} -5 \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \stackrel{\textcircled{7}}{=} -15. \end{aligned}$$

- ① Restamos la primera fila de todas las demás filas. El determinante no cambia.
- ② Factorizamos 5 en la tercera fila.
- ③ Restamos $1/3$ de la última fila a las filas 2 y 3. El determinante no cambia.
- ④ Restamos la fila 3 de la fila 2. El determinante no cambia.
- ⑤ Intercambiamos las filas 2 y 4. Esto da un factor -1 .
- ⑥ Cálculo sencillo.

Ahora debería haber entendido

- las diferentes propiedades del determinante,
- por qué una matriz es invertible si y solo si su determinante es distinto de 0,
- por qué la eliminación de Gauß-Jordan es computacionalmente más eficiente que la fórmula de expansión de Laplace,
- etc.

Ahora debería ser capaz de

- calcular determinantes usando sus propiedades,
- calcular determinantes abstractos,
- usar la factorización de una matriz para calcular su determinante,
- etc.

Ejercicios.

1. Suponga que sabemos que $\det \begin{pmatrix} a & -2b & c \\ 1 & 3 & -1 \\ 0 & 5 & 2 \end{pmatrix} = -1$. Calcule

$$\det \begin{pmatrix} -3a & 6b & -3c \\ -1 & -3 & 1 \\ 0 & 1 & \frac{2}{5} \end{pmatrix}, \quad \det \begin{pmatrix} a & -2b & c \\ 1+2a & 3-4b & -1+2c \\ -a & 5+2b & 2-c \end{pmatrix}, \quad \det \begin{pmatrix} 1 & 3 & -1 \\ 2 & 11 & 0 \\ a+1 & -2b+3 & c-1 \end{pmatrix}.$$

2. Suponga que $\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = 6$. Calcule:

$$\det \begin{pmatrix} a_{31} & a_{32} & a_{33} \\ 3a_{11} - 5a_{31} & 3a_{12} - 5a_{32} & 3a_{13} - 5a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}.$$

3. ¿Para cuáles valores de a la matriz:

$$\begin{pmatrix} 2-a & -2 & 0 \\ 0 & 1 & 1+a \\ a & 2 & 2a \end{pmatrix}$$

no tiene inversa?

4. Sea $A = \begin{pmatrix} 1 & 4 \\ 0 & 4 \end{pmatrix}$. Halle todos los $\lambda \in \mathbb{R}$ tal que $\lambda \text{id}_2 - A$ es no invertible.
5. ¿Falso o verdadero? Sean $A, B \in M(n \times n)$.
- Si n es impar, entonces $\det(-A) = -\det(A)$.
 - Si n es impar y A es antisimétrica, entonces A es no invertible.
 - Si $\det A = 0$ entonces $A = \mathbb{O}$.
 - Si $\det A = 0$ entonces una fila ó columna de A consta de solo ceros.
 - Si $\det A = 0$ entonces por lo menos una entrada de A debe ser 0.
 - Si P es invertible y $A = PBP^{-1}$, entonces $\det A = \det B$.
6. (a) Se dice que $A \in M(n \times n)$ es una *matriz ortogonal* si su inversa es su transpuesta. Sea $A \in M(3 \times 3)$ ortogonal. ¿Cuáles son las posibilidades para $\det A$?

- (b) Se dice que $A \in M(n \times n)$ es una *matriz idempotente* si $A^2 = A$. Sea $A \in M(3 \times 3)$ idempotente. ¿Cuáles son las posibilidades para $\det A$?
7. Sean $A, B \in M(3 \times 3)$ tales que $AB = \mathbb{O}$ y suponga que $\det A \neq 0$. Muestre que $B = \mathbb{O}$.
8. Sean $A, B, C \in M(3 \times 3)$ tales que $AB = AC$ y suponga que $\det A \neq 0$. Muestre que $B = C$. ¿Es cierto también si $\det A = 0$?
9. Sea $A \in M(3 \times 3)$ tal que la suma de sus vectores filas (columnas) da como resultado $\vec{0}$. Muestre que A no es invertible.

4.3. Interpretación geométrica del determinante

En esta breve sección mostramos una interpretación geométrica del determinante. Esto es, por supuesto, solo una pequeña parte de la verdadera importancia del determinante. Escuchará más sobre esto en un curso de cálculo vectorial, cuando se discuta la fórmula de transformación (la regla de sustitución para integrales de dimensión superior), o en un curso de teoría de la medida o geometría diferencial. Aquí nos contentamos con dos hechos básicos.

Área en $\mathbb{R}^2$

Sean $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ y $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ vectores en $\mathbb{R}^2$ y considere la matriz $A = (\vec{a} | \vec{b})$, la matriz cuyas columnas son los vectores dados. Entonces

$$A\vec{e}_1 = \vec{a}, \quad A\vec{e}_2 = \vec{b}.$$

Esto significa que A transforma el cuadrado unitario generado por los vectores unitarios $\vec{e}_1$ y $\vec{e}_2$ en el paralelogramo generado por los vectores $\vec{a}$ y $\vec{b}$. Sea $\text{area}(\vec{a}, \vec{b})$ el área del paralelogramo generado por $\vec{a}$ y $\vec{b}$. Podemos ver $\vec{a}$ y $\vec{b}$ como vectores en $\mathbb{R}^3$ simplemente añadiendo una tercera componente. Entonces la fórmula (2.9) muestra que el área del paralelogramo generado por $\vec{a}$ y $\vec{b}$ es igual a

$$\left\| \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} \right\| = \left\| \begin{pmatrix} a_1 b_2 - a_2 b_1 \\ 0 \\ 0 \end{pmatrix} \right\| = |a_1 b_2 - a_2 b_1| = |\det A|,$$

por lo tanto obtenemos la fórmula

$$\text{area}(\vec{a}, \vec{b}) = |\det A|. \quad (4.9)$$

Es decir, mientras A nos dice como se transforma la forma del cuadrado unitario, $|\det A|$ nos dice como cambia su área, vea Figure 4.1.

También debería notar lo siguiente: El área de la imagen del cuadrado unitario bajo A es cero si y solo si los dos vectores imagen $\vec{a}$ y $\vec{b}$ son paralelos. Esto concuerda con el hecho de que $\det A = 0$ si y solo si las dos rectas descritas por las ecuaciones lineales asociadas son paralelas (o si una ecuación describe todo el plano).

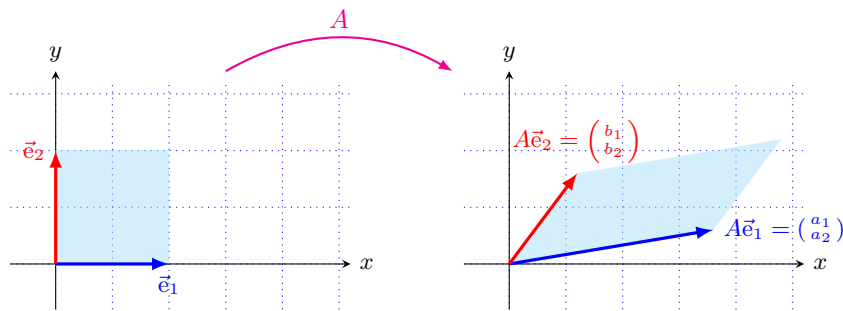


FIGURA 4.1: La figura muestra cómo se transforma el área del cuadrado unitario bajo la transformación lineal A . El área del cuadrado del lado izquierdo es 1; el área del paralelogramo del lado derecho es $|\det A|$.

Volúmenes en $\mathbb{R}^3$

Sean $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ y $\vec{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ vectores in $\mathbb{R}^3$ y considere la matriz $A = (\vec{a} | \vec{b} | \vec{c})$ cuyas columnas son los vectores dados. Entonces

$$A\vec{e}_1 = \vec{a}, \quad A\vec{e}_2 = \vec{b}, \quad A\vec{e}_3 = \vec{c}.$$

Esto significa que A transforma el cubo unitario generado por los vectores unitarios $\vec{e}_1$, $\vec{e}_2$ and $\vec{e}_3$ a un paralelepípedo generado por los vectores $\vec{a}$, $\vec{b}$ y $\vec{c}$. Sea $\text{vol}(\vec{a}, \vec{b}, \vec{c})$ el volumen del paralelepípedo generado por los vectores $\vec{a}$, $\vec{b}$ y $\vec{c}$. According to formula (2.10), $\text{vol}(\vec{a}, \vec{b}, \vec{c}) = |\langle \vec{a}, \vec{b} \times \vec{c} \rangle|$. Calculamos

$$\begin{aligned} |\langle \vec{a}, \vec{b} \times \vec{c} \rangle| &= \left| \left\langle \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \times \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \right\rangle \right| = \left| \left\langle \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_2c_3 - b_3c_2 \\ b_3c_1 - c_3b_1 \\ b_1c_2 - b_2c_1 \end{pmatrix} \right\rangle \right| \\ &= |a_1(b_2c_3 - b_3c_2) - a_2(c_3b_1 - b_3c_1) + a_3(b_1c_2 - b_2c_1)| \\ &= |\det A|, \end{aligned}$$

por lo tanto

$$\text{vol}(\vec{a}, \vec{b}, \vec{c}) = |\det A| \quad (4.10)$$

ya que reconocemos la penúltima línea como la expansión de $\det A$ a lo largo de la primera columna. Es decir, mientras A nos dice como se transforma la forma del cubo unitario, $|\det A|$ nos dice como cambia su volumen,

También debería notar lo siguiente: El volumen de la imagen del cubo unitario bajo A es cero si y solo si los tres vectores imagen están en el mismo plano. Veremos más adelante que esto implica que la imagen de A no es todo $\mathbb{R}^3$, por lo tanto A no puede ser invertible. For details, see Section 6.2.

Lo que vimos para $n = 2$ y $n = 3$ se puede generalizar a $\mathbb{R}^n$ con $n \geq 4$: Una matriz $A \in M(n \times n)$ transforma el cubo unitario en $\mathbb{R}^n$ generado por los vectores unitarios $\vec{e}_1, \dots, \vec{e}_n$ en un paralelepípedo en $\mathbb{R}^n$ y $|\det A|$ nos dice cómo cambia su volumen.

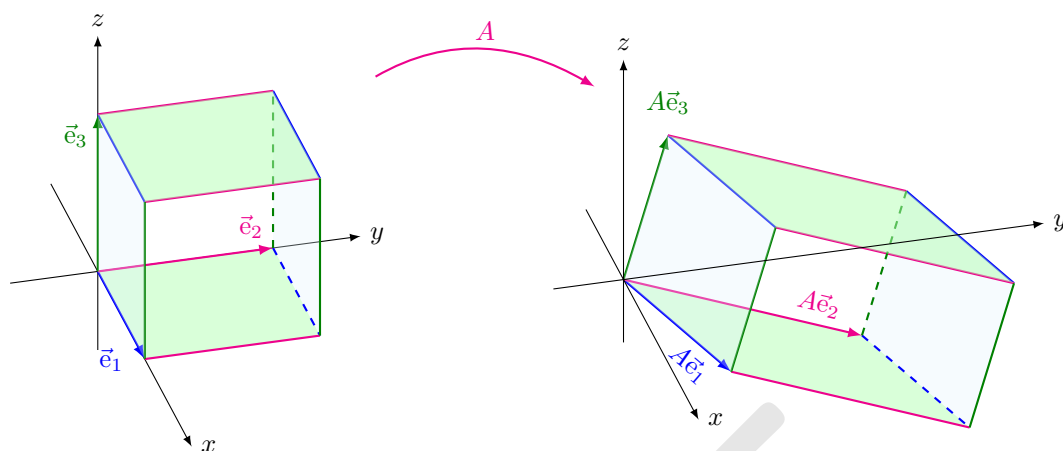
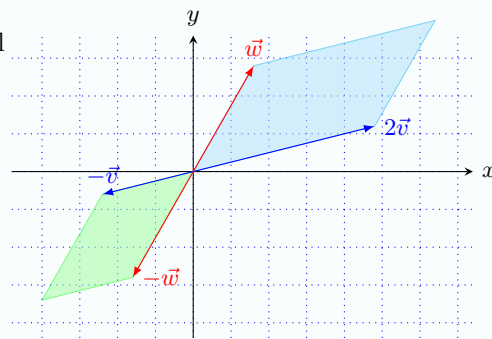


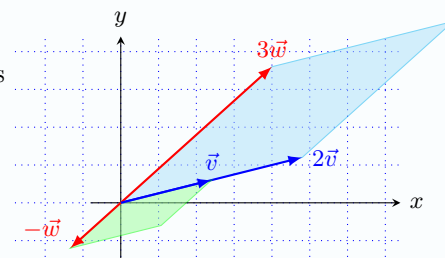
FIGURA 4.2: La figura muestra cómo se transforma el volumen del cubo unitario bajo la transformación lineal A . El volumen del cubo del lado izquierdo es 1; el volumen del paralelepípedo del lado derecho es $|\det A|$.

Ejercicio. Dé dos demostraciones de las siguientes afirmaciones: Una usando la fórmula (4.9) y la linealidad del determinante en sus columnas; y otra usando geometría.

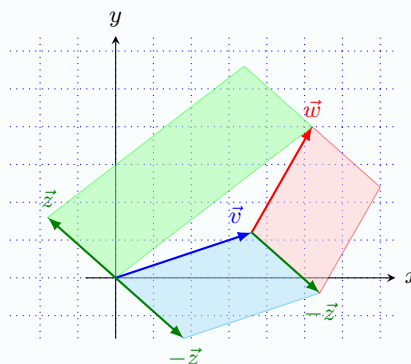
- (I) Muestre que el área del paralelogramo azul es el doble del área del paralelogramo verde.



- (II) Muestre que el área del paralelogramo azul es seis veces el área del paralelogramo verde.



- (III) Muestre que el área del paralelogramo azul y del paralelogramo rojo es igual al área del paralelogramo verde.



Ahora debería haber entendido

- la interpretación geométrica del determinante en $\mathbb{R}^2$ y $\mathbb{R}^3$,
- la estrecha relación entre el determinante y el producto cruz en $\mathbb{R}^3$, y que esta es la razón por la cual el producto cruz aparece en las fórmulas para el área de un paralelogramo y el volumen de un paralelepípedo,
- etc.

Ahora debería ser capaz de

- calcular el área de un paralelogramo y el volumen de un paralelepípedo usando determinantes,
- etc.

Ejercicios.

1. Calcule el volumen del paralelepípedo generado por los vectores $\begin{pmatrix} 1 \\ 3 \\ -5 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix}$.
2. Sea $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & -2 \\ 2 & -2 & 0 \end{pmatrix}$. Calcule el volumen del paralelepípedo generado por $A\vec{e}_1$, $A\vec{e}_2$ y $A\vec{e}_3$.
¿Cómo interpreta geoméricamente el resultado obtenido?
3. Sean $P = (1, 1)$, $Q = (2, 2)$, $R = (0, 3)$ y $W = (-1, 2)$. Muestre que los puntos $PQRW$ forman un paralelogramo y calcule su área.
4. Sean $P = (x_1, y_1)$, $Q = (x_2, y_2)$ y $R = (x_3, y_3)$ puntos de $\mathbb{R}^2$. Muestre que el área de triángulo $\triangle PQR$ está dada por la fórmula:

$$\frac{1}{2} \left| \det \begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{pmatrix} \right|.$$

¿Cuándo este determinante será igual a cero?

4.4. Inversa de una matriz

En esta sección demostramos un método para calcular la inversa de una matriz cuadrada invertible usando determinantes. Aunque la fórmula puede verse bonita, computacionalmente no es eficiente. Aquí está.

Sea $A = (a_{ij})_{i,j=1,\dots,n} \in M(n \times n)$ y sean M_{ij} sus menores; véase la Definición 4.5. Ya sabemos por (4.5) que para todo $k \in \{1, \dots, n\}$ fijo

$$\det A = \sum_{j=1}^n (-1)^{k+j} a_{kj} \det M_{kj}. \quad (4.11)$$

Ahora queremos ver qué ocurre si el k en a_{kj} y en M_{kj} son diferentes.

Proposición 4.13. Sea $A = (a_{ij})_{i,j=1,\dots,n} \in M(n \times n)$ y sean $k, \ell \in \{1, \dots, n\}$ con $k \neq \ell$. Entonces

$$\sum_{j=1}^n (-1)^{\ell+j} a_{kj} \det M_{\ell j} = 0. \quad (4.12)$$

Demostración. Construimos una nueva matriz B a partir de A reemplazando su fila ℓ -ésima por la fila k -ésima. Entonces B tiene dos filas iguales (la fila ℓ y la fila k), por lo tanto $\det B = 0$. Observe que las matrices A y B son iguales en todas partes excepto posiblemente en la fila ℓ -ésima, de modo que sus menores a lo largo de la fila ℓ son iguales: $M_{\ell j}^B = M_{\ell j}^A$ (ponemos superíndices A, B para distinguir los menores de A y los de B). Si expandimos $\det B$ a lo largo de la fila ℓ -ésima, encontramos

$$0 = \det B = \sum_{j=1}^n (-1)^{\ell+j} b_{\ell j} \det M_{\ell j}^B = \sum_{j=1}^n (-1)^{\ell+j} a_{kj} \det M_{\ell j}^A. \quad \square$$

Usando los cofactores C_{ij} de A (véase la Definición 4.5), las fórmulas (4.11) y (4.12) se pueden escribir como

$$\sum_{j=1}^n (-1)^{\ell+j} a_{kj} \det M_{\ell j}^A = \sum_{j=1}^n a_{kj} C_{\ell j} = \begin{cases} \det A & \text{if } k = \ell, \\ 0 & \text{if } k \neq \ell. \end{cases} \quad (4.13)$$

Definición 4.14. Para $A \in M(n \times n)$ definimos su *matriz adjunta* $\text{adj } A$ como la transpuesta de su matriz de cofactores:

$$\text{adj } A := \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{pmatrix}^t = \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{pmatrix}.$$

Atención. Esta matriz no se debe confundir con la “matriz adjunta definida” en Definición 8.8. La matriz adjunta según la Definición 4.14 en inglés se llama “adjugate matrix” o “classical adjoint of A ”, y sirve para calcular la inversa de A (típicamente es difícil calcularla). Por otro lado, la matriz adjunta según la Definición 8.8 en inglés se llama “adjoint matrix of A ”, y tiene que ver con cómo se comporta A en un producto interno (es muy fácil calcularla). Ambas matrices tienen nombres diferentes en inglés, pero tienen el mismo nombre en espa’ nol.

Teorema 4.15. Sea $A \in M(n \times n)$ una matriz invertible. Entonces

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A. \quad (4.14)$$

Demostración. Calculemos $A \operatorname{adj} A$. Por definición de $\operatorname{adj} A$, el coeficiente $c_{k\ell}$ en el producto de matrices $A \operatorname{adj} A$ es exactamente $c_{k\ell} = \sum_{j=1}^n (-1)^{\ell+j} a_{kj} \det M_{\ell j}$, así que por (4.13) se sigue que

$$A \operatorname{adj} A = \begin{pmatrix} \det A & 0 & \dots & 0 \\ 0 & \det A & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \det A \end{pmatrix} = (\det A) \operatorname{id}_n.$$

Reordenando, obtenemos que $A^{-1} = \frac{1}{\det A} \operatorname{adj} A \operatorname{id}_n^{-1} = \frac{1}{\det A} \operatorname{adj} A$. □

Observación 4.16. Observe que la demostración del Teorema 4.15 muestra que $A \operatorname{adj} A = (\det A) \operatorname{id}_n$ es cierto para todo $A \in M(n \times n)$, incluso si no es invertible (en este caso, ambos lados de la fórmula son iguales a la matriz cero).

La fórmula (4.14) puede verse bastante bonita e inocente; sin embargo, tenga presente que para calcular A^{-1} con ella hay que calcular un determinante $n \times n$ y n^2 determinantes de los menores $(n-1) \times (n-1)$ de A . Esto es mucho más que los $O(n^3)$ pasos necesarios en la eliminación de Gauß-Jordan.

Finalmente, demostramos la regla de Cramer para encontrar la solución de un sistema lineal si la matriz correspondiente es invertible.

Teorema 4.17. Sea $A \in M(n \times n)$ una matriz invertible y sea $\vec{b} \in \mathbb{R}^n$. Entonces la solución única $\vec{x}$ de $A\vec{x} = \vec{b}$ está dada por

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} \det A_1^{\vec{b}} \\ \det A_2^{\vec{b}} \\ \vdots \\ \det A_n^{\vec{b}} \end{pmatrix} \quad (4.15)$$

donde $A_j^{\vec{b}}$ es la matriz que se obtiene de A reemplazando su columna j -ésima por el vector $\vec{b}$.

Demostración. Como de costumbre, escribimos C_{ij} para los cofactores de A y M_{ij} para sus menores. Como A es invertible, sabemos que $\vec{x} = A^{-1}\vec{b} = \frac{1}{\det A} \operatorname{adj} A \vec{b}$. Por lo tanto, para $j = 1, \dots, n$

encontramos que

$$x_j = \frac{1}{\det A} \sum_{k=1}^n C_{kj} b_k = \frac{1}{\det A} \sum_{k=1}^n (-1)^{k+j} b_k C_{kj} = \frac{1}{\det A} \det A_j^{\vec{b}}.$$

La última igualdad es cierta porque la penúltima suma es la expansión del determinante de $A_j^{\vec{b}}$ a lo largo de la columna j -ésima. $\square$

Observe que, aunque (4.15) pueda verse bastante bonita, implica el cálculo de $n + 1$ determinantes de matrices $n \times n$, así que requiere $O((n + 1)!)$ pasos.

Ejemplo 4.18. Calculemos la inversa de la matriz $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 8 & 7 \end{pmatrix}$ from Example 4.9, but this time using Theorem 4.17. Ya sabemos que $\det A = 27$. Sus cofactores son

$$\begin{aligned} C_{11} &= \det \begin{pmatrix} 5 & 6 \\ 8 & 7 \end{pmatrix} = -13, & C_{12} &= -\det \begin{pmatrix} 4 & 6 \\ 0 & 7 \end{pmatrix} = -28, & C_{13} &= \det \begin{pmatrix} 4 & 5 \\ 0 & 8 \end{pmatrix} = 32, \\ C_{21} &= -\det \begin{pmatrix} 2 & 3 \\ 8 & 7 \end{pmatrix} = 10, & C_{22} &= \det \begin{pmatrix} 1 & 3 \\ 0 & 7 \end{pmatrix} = 7, & C_{23} &= -\det \begin{pmatrix} 1 & 2 \\ 0 & 8 \end{pmatrix} = -8, \\ C_{31} &= \det \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix} = -3, & C_{32} &= -\det \begin{pmatrix} 1 & 3 \\ 4 & 6 \end{pmatrix} = 6, & C_{33} &= \det \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} = -3. \end{aligned}$$

Por lo tanto

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A = \frac{1}{\det A} \begin{pmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{pmatrix} = \frac{1}{27} \begin{pmatrix} -13 & 10 & -3 \\ -28 & 7 & 6 \\ 32 & -8 & -3 \end{pmatrix}.$$

Ejemplo 4.19. Usemos la regla de Cramer para resolver el siguiente sistema lineal:

$$\begin{aligned} x + y + z &= 8 \\ 4y - z &= -2 \\ 3x - y + 2z &= 0. \end{aligned}$$

Escribimos el sistema anterior en la forma $A\vec{x} = \vec{b}$:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 4 & -1 \\ 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ -2 \\ 0 \end{pmatrix}.$$

Entonces, por la regla de Cramer:

$$x = \frac{\det \begin{pmatrix} 8 & 1 & 1 \\ -2 & 4 & -1 \\ 0 & -1 & 2 \end{pmatrix}}{\det \begin{pmatrix} 1 & 1 & 1 \\ 0 & 4 & -1 \\ 3 & -1 & 2 \end{pmatrix}} = \frac{62}{-8} = -\frac{31}{4}, \quad y = \frac{\det \begin{pmatrix} 1 & 8 & 1 \\ 0 & -2 & -1 \\ 3 & 0 & 2 \end{pmatrix}}{\det \begin{pmatrix} 1 & 1 & 1 \\ 0 & 4 & -1 \\ 3 & -1 & 2 \end{pmatrix}} = \frac{-22}{-8} = \frac{11}{4},$$

$$z = \frac{\det \begin{pmatrix} 1 & 1 & 8 \\ 0 & 4 & -2 \\ 3 & -1 & 0 \end{pmatrix}}{\det \begin{pmatrix} 1 & 1 & 1 \\ 0 & 4 & -1 \\ 3 & -1 & 2 \end{pmatrix}} = \frac{-104}{-8} = 13.$$

Ahora debería haber entendido

- qué es la matriz adjunta y por qué se puede usar para calcular la inversa de una matriz,
- etc.

Ahora debería ser capaz de

- calcular A^{-1} usando $\text{adj } A$.
- resolver sistemas de ecuaciones lineales usando la regla de Cramer.
- etc.

Ejercicios.

1. Usando la regla de Cramer, resuelva los siguientes sistemas de ecuaciones lineales:

$$\begin{array}{llll} \text{(a)} & 2x + 3y = -1, & \text{(b)} & 3x - 2y - 2z = -2, & \text{(c)} & 2x + 3y - z = -5, & \text{(d)} & x - w = 7, \\ & -7x + 4y = 5, & & x + y + 2z = 4, & & 2x + 4y + 6z = 2, & & 2y + z = 2, \\ & & & 2x + y + z = 6, & & x + 2z = 0, & & 4x - y = -3, \\ & & & & & & & 3z - 5w = 2. \end{array}$$

2. Calcular la inversa de las siguientes matrices usando el método de cofactores:

$$\begin{array}{llll} \text{(a)} & \begin{pmatrix} 3 & 5 \\ -1 & 2 \end{pmatrix} & \text{(b)} & \begin{pmatrix} 4 & 2 & 2 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \end{pmatrix} & \text{(c)} & \begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} & \text{(d)} & \begin{pmatrix} 3 & -12 & -2 & -6 \\ 1 & -3 & 0 & -2 \\ -1 & 6 & 1 & 3 \\ -2 & 10 & 2 & 5 \end{pmatrix}. \end{array}$$

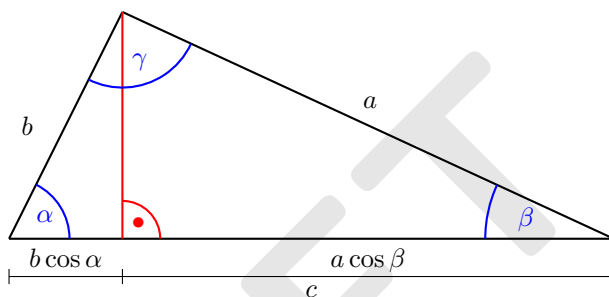
3. Sea

$$A = \begin{pmatrix} 1 & -2 & 3 & 1 \\ -1 & 5 & 0 & 2 \\ 0 & 3 & 1 & -2 \\ -2 & 1 & 2 & 4 \end{pmatrix}.$$

Calcule la entrada a_{32} de A^{-1} .

4. El siguiente ejercicio tiene como objetivo demostrar la ley del coseno usando álgebra lineal.

(a) Considere el siguiente triángulo cuyos lados tienen longitudes a, b, c :



Usando trigonometría elemental, deduzca las siguientes relaciones:

$$\begin{aligned} b \cos \alpha + a \cos \beta &= c, \\ c \cos \alpha + a \cos \gamma &= b, \\ c \cos \beta + b \cos \gamma &= a. \end{aligned}$$

(b) Considere las relaciones anteriores como un sistema de ecuaciones con $\cos \alpha$, $\cos \beta$ y $\cos \gamma$ como incógnitas. Calcule el determinante del sistema.

(c) Use la regla de Cramer para despejar $\cos(\gamma)$ y concluya la ley del coseno:

$$c^2 = a^2 + b^2 - 2ab \cos(\gamma).$$

5. Sea $A \in M(n \times n)$.

(a) Muestre que si A no es invertible entonces $A \operatorname{adj} A = \mathbb{O}$.

(b) Muestre que $\det(\operatorname{adj} A) = (\det A)^{n-1}$.

6. Sea $A \in M(n \times n)$ simétrica e invertible. Muestre que $\operatorname{adj} A$ es simétrica. ¿Sigue siendo cierta la afirmación si no suponemos que A es invertible?

4.5. Resumen

El determinante es una función de las matrices cuadradas a los números reales. Más adelante también consideraremos matrices con entradas complejas. En este caso, el determinante es una función de las matrices cuadradas a los números complejos. Sea $A = (a_{ij})_{i,j=1}^n \in M(n \times n)$.

Fórmulas para el determinante

$$\begin{aligned}
 \det A &= \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} && \text{fórmula de Leibniz} \\
 &= \sum_{j=1}^n (-1)^{k+j} a_{kj} \det M_{kj} = \sum_{j=1}^n a_{kj} C_{kj} && \text{expansión de Laplace a lo largo de la fila } k\text{-ésima} \\
 &= \sum_{i=1}^n (-1)^{i+k} a_{ik} \det M_{ik} = \sum_{i=1}^n a_{ik} C_{ik} && \text{expansión de Laplace a lo largo de la columna } k\text{-ésima}
 \end{aligned}$$

con la siguiente notación

- S_n es el conjunto de todas las permutaciones de $\{1, \dots, n\}$,
- M_{ij} son los menores de A (matrices $(n-1) \times (n-1)$ obtenidas de A eliminando la fila i y la columna j),
- $C_{ij} = (-1)^{i+j} \det M_{ij}$ son los cofactores de A .

Inversa de una matriz usando la matriz adjunta

Si $A \in M(n \times n)$ es invertible, entonces

$$A^{-1} = \frac{1}{\det A} \text{adj } A = \frac{1}{\det A} \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{pmatrix}.$$

Interpretación geométrica

El determinante de una matriz A da el volumen orientado de la imagen del cubo unitario bajo A .

- en $\mathbb{R}^2$: área del paralelogramo generado por $\vec{a}$ y $\vec{b} = |\det A|$,
- en $\mathbb{R}^3$: volumen del paralelepípedo generado por $\vec{a}$, $\vec{b}$ y $\vec{c} = |\det A|$.

Propiedades del determinante

- El determinante es lineal en sus filas y columnas.
- El determinante es alternante en sus filas y columnas.
- $\det \text{id}_n = 1$.
- $\det A = \det A^t$.
- Si una fila de A es múltiplo de otra fila, o si una columna es múltiplo de otra columna, entonces $\det A = 0$. En particular, si A tiene dos filas iguales o dos columnas iguales, entonces $\det A = 0$.

- El determinante de una matriz triangular superior o inferior es el producto de sus entradas diagonales.
- Los determinantes de las matrices elementales son

$$\det S_j(c) = c, \quad \det Q_{ij}(c) = 1, \quad \det P_{ij} = -1.$$

- Sea $A \in M(n \times n)$. Entonces A es invertible si y solo si $\det A \neq 0$.
- Sean $A, B \in M(n \times n)$. Entonces $\det(AB) = \det A \det B$.
- Si $A \in M(n \times n)$ es invertible, entonces $\det(A^{-1}) = (\det A)^{-1}$.

Observe, sin embargo, que en general $\det(A + B) \neq \det A + \det B$.

Teorema. Sea $A \in M(n \times n)$. Entonces las siguientes afirmaciones son equivalentes:

- (I) $\det A \neq 0$.
- (II) A es invertible.
- (III) Para todo $\vec{b} \in \mathbb{R}^n$, la ecuación $A\vec{x} = \vec{b}$ tiene exactamente una solución.
- (IV) La ecuación $A\vec{x} = \vec{0}$ tiene exactamente una solución.
- (V) Toda forma escalonada reducida por filas de A tiene n pivotes.
- (VI) A es equivalente por filas a id_n .

4.6. Ejercicios

1. De las siguientes matrices calcule el determinante. Determine si las matrices son invertibles. Si lo son, encuentre su matriz inversa y el determinante de la inversa.

$$A = \begin{pmatrix} 1 & -2 \\ 2 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} -14 & 21 \\ 12 & -18 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 3 & 6 \\ 4 & 1 & 0 \\ 1 & 4 & 3 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 4 & 6 \\ 2 & 1 & 5 \\ 3 & 5 & 11 \end{pmatrix}.$$

2. Sea

$$A = \begin{pmatrix} 3 & 5t^3 & 1 \\ 0 & 2+t & 0 \\ -t & 10-t^4 & t^2 \end{pmatrix}.$$

Determine $t \in \mathbb{R}$ tal que el sistema $A\vec{x} = \begin{pmatrix} \pi \\ \sqrt{2} \\ \sqrt[3]{2} \end{pmatrix}$ tiene exactamente una única solución.

$$3. \text{ Sea } A = \begin{pmatrix} 1+x & y & z & w & r \\ x & 1+y & z & w & r \\ x & y & 1+z & w & r \\ x & y & z & 1+w & r \\ x & y & z & w & 1+r \end{pmatrix}, \text{ muestre que } \det A = 1 + x + y + z + w + r.$$

(Indicación: es más sencillo si usa la propiedad **D1** de la sección 4.2).

4. De las siguientes matrices calcule el determinante. Determine si las matrices son invertibles. Si lo son, encuentre su matriz inversa y el determinante de la inversa.

$$A = \begin{pmatrix} \pi & 3 \\ 5 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 2 & 3 \\ 1 & 3 & 1 \\ 4 & 3 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 2 \\ 1 & 4 & 0 & 3 \\ 1 & 1 & 5 & 4 \end{pmatrix}.$$

5. Encuentre por lo menos cuatro matrices 3×3 cuyo determinante es 18.
6. Use las factorizaciones encontradas en los Ejercicios 14. y 14. en el capítulo 3 para calcular sus determinantes.

7. Escriba la matriz $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 6 \\ -2 & -2 & -6 \end{pmatrix}$ como producto de matrices elementales y calcule el determinante de A usando las matrices elementales encontradas.

8. Determine todos los $x \in \mathbb{R}$ tal que las siguientes matrices son invertibles:

$$A = \begin{pmatrix} x & 2 \\ 1 & x-3 \end{pmatrix}, \quad B = \begin{pmatrix} x & x & 3 \\ 1 & 2 & 6 \\ -2 & 2 & -6 \end{pmatrix}, \quad C = \begin{pmatrix} 11-x & 5 & -50 \\ 3 & -x & -15 \\ 2 & 1 & -x-9 \end{pmatrix}.$$

9. Suponga que una función y satisface $y^{[n]} = b_{n-1}y^{[n-1]} + \dots + b_1y' + b_0y$ donde $b_0, \dots, b_{n-1}$ son coeficientes constantes y $y^{[j]}$ denota la derivada j -ésima de y .

Verifique que $Y' = AY$ donde

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ b_1 & b_2 & \dots & \dots & \dots & b_{n-1} \end{pmatrix}, \quad Y = \begin{pmatrix} y \\ y' \\ y'' \\ \vdots \\ y^{[n-1]} \end{pmatrix}$$

y calcule el determinante de A .

10. Sin usar fórmulas de expansión para determinantes, encuentre para cada una de las matrices dadas parámetros x, y tales que el determinante de las siguientes matrices es igual a cero y explique por qué los parámetros encontrados sirven.

$$N_1 = \begin{pmatrix} x & 2 & 6 \\ 2 & 5 & 1 \\ 3 & 4 & y \end{pmatrix}, \quad N_2 = \begin{pmatrix} 1 & x & y & 2 \\ x & 0 & 1 & y \\ x & 5 & 3 & y \\ 4 & x & y & 8 \end{pmatrix}.$$

11. (a) Calcule $\det B_n$ donde B_n es la matriz en $M(n \times n)$ cuyas entradas en la diagonal son 0 y todas las demás entradas son 1, es decir:

$$B_1 = 0, \quad B_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B_3 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad B_4 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \text{ etc.}$$

¿Cómo cambia la respuesta si en vez de 0 hay x en la diagonal?

- (b) Calcule $\det B_n$ donde B_n es la matriz en $M(n \times n)$ cuyas entradas en la diagonal son 0 y todas las demás entradas satisfacen $b_{ij} = (-1)^{i+j}$, es decir:

$$B_1 = 0, \quad B_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B_3 = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}, \quad B_4 = \begin{pmatrix} 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & -1 \\ -1 & 1 & 0 & 1 \\ 1 & -1 & 1 & 0 \end{pmatrix},$$

$$B_5 = \begin{pmatrix} 0 & 1 & -1 & 1 & -1 \\ 1 & 0 & 1 & -1 & 1 \\ -1 & 1 & 0 & 1 & -1 \\ 1 & -1 & 1 & 0 & 1 \\ -1 & 1 & -1 & 1 & 0 \end{pmatrix}, \text{ etc.}$$

¿Cómo cambia la respuesta si en vez de 0 hay x en la diagonal? Compare con el Ejercicio 12..

Capítulo 5

Espacios vectoriales

En lo que sigue, $\mathbb{K}$ siempre denota un campo. En este capítulo, siempre puede pensar en $\mathbb{K} = \mathbb{R}$, aunque casi todo es cierto también para otros campos, como $\mathbb{C}$, $\mathbb{Q}$ o $\mathbb{F}_p$, donde p es un número primo. Más adelante, en el Capítulo 8, será más útil trabajar con $\mathbb{K} = \mathbb{C}$.

En este capítulo trabajaremos con espacios vectoriales abstractos. Primero discutiremos sus propiedades básicas. Luego, en la Sección 5.2, definiremos subespacios. Estos son subconjuntos de espacios vectoriales que son, a su vez, espacios vectoriales. En la Sección 5.4 introduciremos bases y la dimensión de un espacio vectorial. Estos conceptos son fundamentales en el álgebra lineal, ya que nos permiten clasificar todos los espacios vectoriales de dimensión finita. En cierto sentido, todos los espacios vectoriales de dimensión n sobre el mismo campo $\mathbb{K}$ son iguales. En el Capítulo 6 estudiaremos aplicaciones lineales entre espacios vectoriales.

5.1. Definiciones y propiedades básicas

Primero recordamos la definición de espacio vectorial abstracto de Chapter 2 (p. 31).

Definición 5.1. Sea V un conjunto junto con dos operaciones

$$\begin{aligned} \text{suma de vectores} \quad + : V \times V &\rightarrow V, \quad (v, w) \mapsto v + w, \\ \text{producto de un escalar y un vector} \quad \cdot : \mathbb{K} \times V &\rightarrow V, \quad (\lambda, v) \mapsto \lambda \cdot v. \end{aligned}$$

Observe que usualmente escribiremos λv en lugar de $\lambda \cdot v$. Entonces V (o, más precisamente, $(V, +, \cdot)$) se llama un *espacio vectorial* sobre $\mathbb{K}$ si para todos $u, v, w \in V$ y todos $\lambda, \mu \in \mathbb{K}$ se cumple lo siguiente:

- (a) **Asociatividad:** $(u + v) + w = u + (v + w)$ para todo $u, v, w \in V$.
- (b) **Conmutatividad:** $v + w = w + v$ para todo $u, v \in V$.
- (c) **Elemento neutro de la suma:** Existe un elemento $0 \in V$, llamado la *identidad aditiva*, tal que $0 + v = v + 0 = v$ para todo $v \in V$.

- (d) **Elemento inverso:** Para todo $v \in V$, existe un elemento inverso v' tal que $v + v' = \mathbb{O}$.
- (e) **Elemento neutro de la multiplicación por escalar:** Para todo $v \in V$, tenemos que $1v = v$.
- (f) **Compatibilidad:** Para todo $v \in V$ y $\lambda, \mu \in \mathbb{K}$, tenemos que $(\lambda\mu)v = \lambda(\mu v)$.
- (g) **Leyes distributivas:** Para todos $v, w \in V$ y $\lambda, \mu \in \mathbb{K}$, tenemos

$$(\lambda + \mu)v = \lambda v + \mu v \quad \text{and} \quad \lambda(v + w) = \lambda v + \lambda w.$$

Ya sabemos que $\mathbb{R}^n$ es un espacio vectorial sobre $\mathbb{R}$.

Observación 5.2. (I) Observe que usamos la notación $\vec{v}$ con flecha solamente para el caso especial de vectores en $\mathbb{R}^n$ o $\mathbb{C}^n$. Los vectores en espacios vectoriales abstractos usualmente se denotan sin flecha.

- (II) Si $\mathbb{K} = \mathbb{R}$, entonces V se llama un *espacio vectorial real*. Si $\mathbb{K} = \mathbb{C}$, entonces V se llama un *espacio vectorial complejo*.

Antes de dar ejemplos de espacios vectoriales, mostramos primero algunas propiedades básicas de los espacios vectoriales.

Propiedades 5.3. (I) *El elemento neutro es único.* (Observe que en los axiomas de espacio vectorial solo pedimos la *existencia* de un elemento neutro aditivo; no pedimos unicidad. Así, se podría pensar que puede haber varios elementos que satisfacen (c) en la Definición 5.1. Sin embargo, esto no es posible, como muestra la siguiente demostración.)

Demostración. Supongamos que hay dos elementos neutros $\mathbb{O}$ y $\mathbb{O}'$. Entonces sabemos que para todo v y w en V se cumple lo siguiente:

$$v = v + \mathbb{O}, \quad w = w + \mathbb{O}'.$$

Ahora tomemos $v = \mathbb{O}'$ y $w = \mathbb{O}$. Entonces, usando la conmutatividad, obtenemos

$$\mathbb{O}' = \mathbb{O}' + \mathbb{O} = \mathbb{O} + \mathbb{O}' = \mathbb{O}. \quad \square$$

- (II) Sean $x, y, z \in V$. Si $x + y = x + z$, entonces $y = z$.

Demostración. Sea x' un inverso aditivo de x (esto significa que $x' + x = \mathbb{O}$, el cual debe existir ya que V es un espacio vectorial). Esto se sigue de

$$y = \mathbb{O} + y = (x' + x) + y = x' + (x + y) = x' + (x + z) = (x' + x) + z = \mathbb{O} + z = z. \quad \square$$

- (III) Para todo $v \in V$, su elemento inverso es único. (Observe que en los axiomas de espacio vectorial solo pedimos la *existencia* de un inverso aditivo para cada elemento $x \in V$; no pedimos unicidad. Así, se podría pensar que puede haber varios elementos que satisfacen (d) en la Definición 5.1. Sin embargo, esto no es posible, como muestra la siguiente demostración.)

Demostración. Sea $v \in V$ y supongamos que hay elementos v', v'' en V tales que

$$v + v' = \mathbb{O}, \quad v + v'' = \mathbb{O}.$$

Ahora se sigue de (II) que $v' = v''$ (tome $x = v$, $y = v'$ y $z = v''$ en (II)). $\square$

(IV) Para todo $\lambda \in \mathbb{K}$ tenemos $\lambda\mathbb{O} = \mathbb{O}$.

Demostración. Observe que $\lambda\mathbb{O} = \lambda\mathbb{O} + \mathbb{O}$ and that $\lambda\mathbb{O} = \lambda(\mathbb{O} + \mathbb{O}) = \lambda\mathbb{O} + \lambda\mathbb{O}$, por lo tanto

$$\lambda\mathbb{O} + \mathbb{O} = \lambda\mathbb{O} + \lambda\mathbb{O}.$$

Ahora se sigue de (II) que $\mathbb{O} = \lambda\mathbb{O}$ (tome $x = \lambda\mathbb{O}$, $y = \mathbb{O}$ y $z = \lambda\mathbb{O}$ en (II)). $\square$

(v) Para todo $v \in V$ tenemos que $0v = \mathbb{O}$.

Demostración. La demostración es similar a la anterior. Observe que $0v = 0v + \mathbb{O}0$ and $0v = (0 + 0)v = 0v + 0v$, por lo tanto

$$0v + \mathbb{O} = 0v + 0v.$$

Ahora se sigue de (II) que $\mathbb{O} = 0v$ (tome $x = 0v$, $y = \mathbb{O}$ y $z = 0v$ en (II)). $\square$

(VI) Si $\lambda v = \mathbb{O}$, entonces o bien $\lambda = 0$ or $v = \mathbb{O}$.

Demostración. Si $\lambda = 0$, entonces no hay nada que demostrar. Supongamos ahora que $\lambda \neq 0$. Entonces v es $\mathbb{O}$ porque

$$v = \frac{1}{\lambda}(\lambda v) = \frac{1}{\lambda}\mathbb{O} = \mathbb{O}. \quad \square$$

(VII) Para todo $v \in V$, su inverso es $(-1)v$.

Demostración. Sea $v \in V$. Observe que por (VI), tenemos que $0v = \mathbb{O}$. Por lo tanto

$$\mathbb{O} = 0v = (1 + (-1))v = v + (-1)v.$$

Por lo tanto, $(-1)v$ es un inverso aditivo de v . Por (III), el inverso de v es único; por lo tanto, $(-1)v$ es el inverso de v . $\square$

Observación 5.4. De ahora en adelante, escribimos $-v$ para el inverso aditivo de un vector. Esta notación está justificada por la Propiedad 5.3 (VII).

Ejemplos 5.5. Damos algunos ejemplos importantes de espacios vectoriales.

- $\mathbb{R}$ es un espacio vectorial real. Más generalmente, $\mathbb{R}^n$ es un espacio vectorial real. La demostración es la misma que para $\mathbb{R}^2$ en el Capítulo 2. La asociatividad y la conmutatividad son claras. El elemento neutro es el vector cuyas entradas son todas iguales a cero: $\vec{0} = (0, \dots, 0)^t$. El inverso de un vector dado $\vec{x} = (x_1, \dots, x_n)^t$ es $(-x_1, \dots, -x_n)^t$. Las leyes distributivas son claras, al igual que el hecho de que $1\vec{x} = \vec{x}$ para todo $\vec{x} \in \mathbb{R}^n$.

- $\mathbb{C}$ es un espacio vectorial complejo. Más generalmente, $\mathbb{C}^n$ es un espacio vectorial complejo. La demostración es la misma que para $\mathbb{R}^n$.
- $\mathbb{C}$ también puede verse como un espacio vectorial real.
- $\mathbb{R}$ **no** es un espacio vectorial complejo con la definición usual de las operaciones algebraicas. Si lo fuera, entonces los vectores serían números reales y los escalares serían números complejos. Pero entonces, si tomamos $1 \in \mathbb{R}$ e $i \in \mathbb{C}$, el producto $i1$ tendría que ser un vector, es decir, un número real, lo cual no es el caso.
- $\mathbb{R}$ puede verse como un espacio vectorial sobre $\mathbb{Q}$.
- Para todo $n, m \in \mathbb{N}$, el espacio $M(m \times n)$ de todas las matrices $m \times n$ con coeficientes reales es un espacio vectorial real.

Demostración. Observe que en este caso los vectores son matrices. La asociatividad y la conmutatividad son fáciles de verificar. El elemento neutro es la matriz cuyas entradas son todas iguales a cero. Dada una matriz $A = (a_{ij})_{\substack{i=1,\dots,m, \\ j=1,\dots,n}}$, su inverso (aditivo) es la matriz $-A = (-a_{ij})_{\substack{i=1,\dots,m, \\ j=1,\dots,n}}$. Las leyes distributivas son claras, al igual que el hecho de que $1A = A$ para todo $A \in M(m \times n)$. $\square$

- Para todo $n, m \in \mathbb{N}$, el espacio $M(m \times n, \mathbb{C})$ de todas las matrices $m \times n$ con coeficientes complejos es un espacio vectorial complejo.

Demostración. Como en el caso de las matrices reales. $\square$

- Sea $C(\mathbb{R})$ el conjunto de todas las funciones continuas de $\mathbb{R}$ en $\mathbb{R}$. Definimos la suma de dos funciones f y g de la manera usual como la nueva función

$$f + g : \mathbb{R} \rightarrow \mathbb{R}, \quad (f + g)(x) = f(x) + g(x).$$

El producto de una función f por un número real λ da la nueva función λf definida por

$$\lambda f : \mathbb{R} \rightarrow \mathbb{R}, \quad (\lambda f)(x) = \lambda f(x).$$

Entonces $C(\mathbb{R})$ es un espacio vectorial con estas operaciones.

Demostración. Es claro que estas operaciones satisfacen asociatividad, conmutatividad y distributividad, y que $1f = f$ para toda función $f \in C(\mathbb{R})$. La identidad aditiva es la función cero (la función que es constantemente cero). Para una función dada f , su inverso (aditivo) es la función $-f$. $\square$

- Sea $P(\mathbb{R})$ el conjunto de todos los polinomios. Con la suma usual y el producto usual por escalares, forman un espacio vectorial.

Demuestre que $\mathbb{C}$ es un espacio vectorial sobre $\mathbb{R}$ y que $\mathbb{R}$ es un espacio vectorial sobre $\mathbb{Q}$.

Observe que los conjuntos $M(m \times n)$ y $C(\mathbb{R})$ admiten más operaciones; por ejemplo, podemos multiplicar funciones, o podemos multiplicar matrices, o podemos calcular $\det A$ para una matriz cuadrada. Sin embargo, todas estas operaciones no tienen nada que ver con la pregunta de si son espacios vectoriales o no. Es importante observar que para un espacio vectorial solo necesitamos la suma de dos vectores y el producto de un escalar por un vector, y que estos satisfagan los axiomas de la Definición 5.1.

Damos más ejemplos.

Ejemplos 5.6. ■ Considere $\mathbb{R}^2$, pero cambiamos la suma usual por la nueva suma $\oplus$ definida por

$$\begin{pmatrix} x \\ y \end{pmatrix} \oplus \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x+a \\ 0 \end{pmatrix}.$$

Con esta nueva suma, $\mathbb{R}^2$ **no** es un espacio vectorial. La razón es que no hay identidad aditiva. Para ver esto, supongamos que tuviéramos una identidad aditiva, digamos $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$. Entonces deberíamos tener

$$\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{for all } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

Sin embargo, por ejemplo,

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

■ Considere $\mathbb{R}^2$, pero cambiamos la suma usual por la nueva suma $\oplus$ definida por

$$\begin{pmatrix} x \\ y \end{pmatrix} \oplus \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x+b \\ y+a \end{pmatrix}.$$

Con esta nueva suma, $\mathbb{R}^2$ **no** es un espacio vectorial. Una de las razones es que la suma no es conmutativa. Por ejemplo

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1+1 \\ 0+0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \text{but} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0+0 \\ 1+1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

Muestre que no hay una identidad aditiva $\mathbb{O}$ que satisfaga $\vec{x} \oplus \mathbb{O} = \vec{x}$ for all $\vec{x} \in \mathbb{R}^2$.

■ Sea $V = \mathbb{R}_+ = (0, \infty)$. Convertimos a V en un espacio vectorial real con las siguientes operaciones: Sean $x, y \in V$ y $\lambda \in \mathbb{R}$. Definimos

$$x \oplus y = xy \quad \text{and} \quad \lambda \odot x = x^\lambda.$$

Entonces $(V, \oplus, \odot)$ es un espacio vectorial real.

Demostración. Sean $u, v, w \in V$ y sea $\lambda \in \mathbb{R}$. Entonces:

- (a) **Asociatividad:** $(u \oplus v) \oplus w = (uv) \oplus w = (uv)w = u(vw) = u(v \oplus w) = u \oplus (v \oplus w)$.
- (b) **Conmutatividad:** $v \oplus w = vw = wv = w \oplus v$.
- (c) La **identidad aditiva** de $\oplus$ es 1 porque para todo $x \in V$ tenemos que $1 \oplus x = 1x = x$.
- (d) **Elemento inverso:** Para todo $x \in V$, su elemento inverso es x^{-1} porque $x \oplus x^{-1} = xx^{-1} = 1$, que es el elemento neutro. (Observe que esto está de acuerdo con las Propiedades 5.3 (VI), ya que $(-1) \odot x = x^{-1}$.)
- (e) **Elemento neutro de la multiplicación por escalar:** Para todo $x \in V$, tenemos que $1 \odot x = 1x = x$.
- (f) **Compatibilidad:** Para todo $x \in V$ y $\lambda, \mu \in \mathbb{R}$, tenemos que

$$(\lambda\mu) \odot v = v^{\lambda\mu} = (v^\lambda)^\mu = \mu \odot (v^\lambda) = \lambda \odot (\mu \odot v).$$

- (g) **Leyes distributivas:** For all $x, y \in V$ and $\lambda, \mu \in \mathbb{R}$, we have

$$(\lambda + \mu) \odot x = x^{\lambda+\mu} = x^\lambda x^\mu = (\lambda \odot v)(\mu \odot v) = (\lambda \odot v) \oplus (\mu \odot v)$$

and

$$\lambda \odot (v \oplus w) = (v \oplus w)^\lambda = (vw)^\lambda = v^\lambda w^\lambda = v^\lambda \oplus w^\lambda = (\lambda \odot v) \oplus (\lambda \odot w). \quad \square$$

- El ejemplo anterior se puede generalizar: Sea $f : \mathbb{R} \rightarrow (a, b)$ una función inyectiva. Entonces el intervalo (a, b) se convierte en un espacio vectorial real si definimos la suma de dos vectores $x, y \in (a, b)$ por

$$x \oplus y = f(f^{-1}(x) + f^{-1}(y))$$

y el producto de un escalar $\lambda \in \mathbb{R}$ y un vector $x \in (a, b)$ por

$$\lambda \odot x = f(\lambda f^{-1}(x)).$$

Observe que en el ejemplo anterior teníamos $(a, b) = (0, \infty)$ y $f = \exp$ (es decir: $f(x) = e^x$).

Debería haber entendido

- el concepto de espacio vectorial abstracto,
- que los espacios $\mathbb{R}^n$ son ejemplos de espacios vectoriales, pero hay muchos más,
- que los “vectores” no necesariamente pueden escribirse como columnas (piense en el espacio vectorial de todos los polinomios, etc.)
- etc.

Ahora debería ser capaz de

- dar ejemplos de espacios vectoriales distintos de $\mathbb{R}^n$ o $\mathbb{C}^n$,
- verificar si un conjunto dado con una suma dada y una multiplicación por escalares dada es un espacio vectorial,

- recitar los axiomas de espacio vectorial si lo despiertan en medio de la noche,
- etc.

Ejercicios.

En los siguientes ejercicios, diga si el conjunto dado es un espacio vectorial sobre $\mathbb{R}$. Si no lo es, determiné cuales propiedades de espacio vectorial no se cumplen.

1. $\{(x, y) \in \mathbb{R}^2 : y \leq 0\}$ con la suma y producto escalar usuales.
2. $\{(x, 2x, 3x) : x \in \mathbb{R}\}$ con la suma y producto escalar usuales.
3. $\{(x, y, z) \in \mathbb{R}^3 : xy = 0\}$ con la suma y producto escalar usuales.
4. $\{(x, y) \in \mathbb{R}^2 : x \in \mathbb{R}, y \in \mathbb{Q}\}$ con la suma y producto escalar usuales.
5. El conjunto de puntos de $\mathbb{R}^3$ que están sobre la recta $x = t + 1$, $y = 2t$ y $z = 0$ con la suma y producto escalar usuales en $\mathbb{R}^3$.
6. $\left\{\begin{pmatrix} 0 & -a \\ 0 & b \end{pmatrix} : a, b \in \mathbb{R}\right\}$ con la suma y multiplicación por escalar de matrices.
7. El conjunto de polinomios de grado ≤ 3 con término independiente cero con la suma y multiplicación por escalar de polinomios.
8. El conjunto de polinomios de grado ≤ 2 con coeficiente que acompaña a X no negativo.
9. El conjunto de las funciones derivables en todo $\mathbb{R}$ con la suma y multiplicación por escalar de $C(\mathbb{R})$.
10. $\mathbb{R}^n$ con las siguientes operaciones:

$$\begin{aligned} \text{suma: } \vec{a} \oplus \vec{b} &:= \vec{a} - \vec{b}, \\ \text{producto con escalar: } \lambda \odot \vec{a} &:= \lambda \vec{a}. \end{aligned}$$

11. $\mathbb{R}^2$ con las siguientes operaciones:

$$\begin{aligned} \text{suma: } (a, b) \oplus (c, d) &:= (a + c, -b - d), \\ \text{producto con escalar: } \lambda \odot (a, b) &:= (\lambda a, \lambda b) \end{aligned}$$

12. $\mathbb{R}^2$ con las siguientes operaciones:

$$\begin{aligned} \text{suma: } (a, b) \oplus (c, d) &:= (a + c + 1, b + d), \\ \text{producto con escalar: } \lambda \odot (a, b) &:= (a, \lambda b) \end{aligned}$$

13. Sea $V = \{a\}$ (note que V solo tiene un elemento), sobre V defina las siguientes operaciones:

$$\begin{aligned} \text{suma: } a \oplus a &:= a, \\ \text{producto con escalar: } \lambda \odot a &:= a. \end{aligned}$$

5.2. Subespacios

En esta sección trabajamos principalmente con espacios vectoriales reales, para fijar ideas. Sin embargo, todos los enunciados también son verdaderos para espacios vectoriales complejos. Solo tenemos que reemplazar $\mathbb{R}$ por $\mathbb{C}$ y la palabra *real* por *complejo* en todas partes.

En esta sección investigaremos cuándo un subconjunto de un espacio vectorial dado es, él mismo, un espacio vectorial.

Definición 5.7. Sea V un espacio vectorial y sea $W \subseteq V$ un subconjunto de V . Entonces W se llama un *subespacio* de V si W mismo es un espacio vectorial con la suma y el producto por escalares heredados de V . Un subespacio W se llama un *subespacio propio* si $W \neq \{\mathbb{O}\}$ y $W \neq V$.

Primero observamos los siguientes hechos básicos.

Observación 5.8. Sea V un espacio vectorial.

- V siempre contiene los siguientes subespacios: $\{\mathbb{O}\}$ y el propio V . Sin embargo, no son subespacios propios.
- Si V es un espacio vectorial, W es un subespacio de V y U es un subespacio de W , entonces U es un subespacio de V .

Demuestre estas afirmaciones.

Observación 5.9. Sea W un subespacio de un espacio vectorial V . Sea $\mathbb{O}$ el elemento neutro en V . Entonces $\mathbb{O} \in W$ y es el elemento neutro de W .

Demostración. Como W es un espacio vectorial, debe tener un elemento neutro $\mathbb{O}_W$. De entrada, no es claro que $\mathbb{O}_W = \mathbb{O}$. Sin embargo, como $\mathbb{O}_W \in W \subset V$, sabemos que $0\mathbb{O}_W = \mathbb{O}$. Por otro lado, como W es un espacio vectorial, es cerrado bajo el producto por escalares; por lo tanto $\mathbb{O} = 0\mathbb{O}_W \in W$. Claramente, $\mathbb{O}$ es un elemento neutro en W . Por lo tanto, se sigue que $\mathbb{O} = \mathbb{O}_W$ por la unicidad del elemento neutro de W ; véanse las Propiedades 5.3(i). $\square$

Supongamos ahora que tenemos un espacio vectorial V y, dentro de él, un subconjunto $W \subseteq V$, y queremos verificar si W es un espacio vectorial. En principio tendríamos que verificar los siete axiomas de espacio vectorial de la Definición 5.1. Sin embargo, si W es un subconjunto de V , entonces algunos de los axiomas de espacio vectorial se obtienen automáticamente. Más precisamente, los axiomas (a), (b), (e), (f) and (g) se cumplen automáticamente. Por ejemplo, para demostrar (b), tomamos dos elementos $w_1, w_2 \in W$. También pertenecen a V , ya que $W \subseteq V$, y por lo tanto conmutan: $w_1 + w_2 = w_2 + w_1$.

Incluso podemos demostrar la siguiente proposición:

Proposición 5.10. Sea V un espacio vectorial real y sea $W \subseteq V$ un subconjunto. Entonces W es un subespacio de V si y solo si se cumplen las siguientes tres propiedades:

- (I) $W \neq \emptyset$, es decir, W no es vacío.

- (II) W es cerrado bajo sumas, es decir, si tomamos w_1 y w_2 en W , entonces su suma $w_1 + w_2$ pertenece a W .
- (III) W es cerrado bajo el producto por escalares, es decir, si tomamos $w \in W$ y $\lambda \in \mathbb{R}$, entonces $\lambda w \in W$.

Observe que (II) y (III) se pueden resumir en lo siguiente:

- (iv) W es cerrado bajo sumas y producto por escalares, es decir, si tomamos $w_1, w_2 \in W$ y $\lambda \in \mathbb{R}$, entonces $\lambda w_1 + w_2 \in W$.

Demostración de Proposición 5.10. Supongamos que W es un subespacio; entonces claramente (II) and (III) se cumplen. (I) se cumple porque todo espacio vectorial debe contener al menos la identidad aditiva $\mathbb{O}$.

Supongamos ahora que W es un subconjunto de V tal que las propiedades (I), (II) and (III) se satisfacen. Para mostrar que W es un subespacio de V , necesitamos verificar los axiomas de espacio vectorial (a) - (f) de la Definición 5.1. Por las hipótesis (II) y (III), la suma y el producto por escalares están bien definidos en W . Además, ya nos convencimos de que (a), (b), (e), (f) and (g) se cumplen. Ahora, para la existencia de una identidad aditiva, tomamos un $w \in W$ arbitrario (tal w existe porque W no es vacío por la hipótesis (I)). Por lo tanto $\mathbb{O} = 0w \in W$, donde $\mathbb{O}$ es la identidad aditiva en V . Esta es entonces también la identidad aditiva en W . Finalmente, dado $w \in W \subseteq V$, sabemos por las Propiedades 5.3 (vi) que su inverso aditivo es $(-1)w$, el cual, por nuestra hipótesis (III), pertenece a W . Así hemos verificado que W satisface todos los axiomas de espacio vectorial, por lo tanto es un espacio vectorial. $\square$

La proposición también es verdadera si V es un espacio vectorial complejo. Solo tenemos que reemplazar $\mathbb{R}$ por $\mathbb{C}$ en todas partes.

Para verificar que un $W \subseteq V$ dado es un subespacio, solo hay que verificar (I), (II) and (III) de la proposición anterior. Para verificar que W no es vacío, típicamente se comprueba si contiene a $\mathbb{O}$.

Veamos ejemplos de subespacios.

Ejemplos 5.11. Sea V un espacio vectorial. Suponemos que V es un espacio vectorial real, pero todo funciona también para un espacio vectorial complejo (solo tenemos que reemplazar $\mathbb{R}$ por $\mathbb{C}$ en todas partes).

- (I) $\{0\}$ y V son subespacios de V . Se llaman los *subespacios triviales* de V .
- (II) Fijemos $v \in V$. Entonces el conjunto $W := \{\lambda v : \lambda \in \mathbb{R}\}$ es un subespacio de V .
- (III) Más generalmente, si fijamos $v_1, \dots, v_k \in V$, entonces el conjunto

$$W := \{\alpha_1 v_1 + \dots + \alpha_k v_k : \alpha_1, \dots, \alpha_k \in \mathbb{R}\}$$

es un subespacio de V . Este conjunto se llama el *subespacio generado por* $v_1, \dots, v_k$. Se mostrará en el Teorema 5.24 que, en efecto, es un espacio vectorial.

- (iv) $P(\mathbb{R})$, el conjunto de todos los polinomios reales, es un subespacio de $C(\mathbb{R})$, el conjunto de todas las funciones continuas sobre $\mathbb{R}$.
- (v) Para todo n , los polinomios de grado a lo sumo n , $P_n(\mathbb{R})$, forman un subespacio de $P(\mathbb{R})$ y también de $C(\mathbb{R})$.
- (vi) Si W es un subespacio de V , entonces $V \setminus W$ no es un subespacio. Esto se ve fácilmente si recordamos que W debe contener a $\mathbf{0}$. Pero entonces $V \setminus W$ no puede contener a $\mathbf{0}$; por lo tanto, no puede ser un espacio vectorial.

Algunos ejemplos más:

- Ejemplos 5.12.** (I) El conjunto de todas las soluciones de un sistema homogéneo de ecuaciones lineales es un espacio vectorial.
- (II) El conjunto de todas las soluciones de una ecuación diferencial lineal homogénea es un espacio vectorial.

Demostración. Demostramos solamente (I), ya que la demostración de (II) es similar. Supongamos que el sistema de ecuaciones está dado en forma matricial $A\vec{x} = \vec{0}$ para alguna matriz $A \in (m \times n)$. Denotamos por U el conjunto de todas las soluciones, es decir, $U := \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}$. Claramente, $\vec{0} \in U$. Ahora sean $\vec{y}, \vec{z} \in U$ y $\lambda \in \mathbb{R}$. Tenemos que mostrar que entonces también $\vec{y} + \lambda\vec{z} \in U$. Esto es cierto porque

$$A(\vec{y} + \lambda\vec{z}) = A\vec{y} + \lambda A\vec{z} = \vec{0} + \lambda\vec{0} = \vec{0}.$$

Por lo tanto, el vector $\vec{y} + \lambda\vec{z}$ resuelve la ecuación homogénea, así que pertenece a U , como queríamos mostrar. $\square$

Observe, sin embargo, que el conjunto de todas las soluciones de una ecuación *no* homogénea *no* es un espacio vectorial. Consideremos un sistema de ecuaciones lineales dado en forma matricial por

$$A\vec{x} = \vec{b}$$

donde $A \in M(m \times n)$ y $\vec{b} \in \mathbb{R}^m$ con $\vec{b} \neq \vec{0}$. Denotamos su conjunto de soluciones por

$$U := \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{b}\}.$$

Claramente, $\vec{0} \notin U$ porque $A\vec{0} = \vec{0} \neq \vec{b}$. Esto ya es suficiente para concluir que U no es un espacio vectorial. Pero también podemos ver que U no es cerrado bajo sumas ni bajo multiplicación por escalares. Para ello, sean $\vec{y}, \vec{z} \in U$ y $\lambda \in \mathbb{R} \setminus \{0\}$. Then

$$A(\vec{y} + \lambda\vec{z}) = A\vec{y} + \lambda A\vec{z} = \vec{b} + \lambda\vec{b} = (1 + \lambda)\vec{b} \neq \vec{b}.$$

Sin embargo, U es “casi” un espacio vectorial. Recordemos que si escribimos las soluciones de $A\vec{x} = \vec{b}$ en forma vectorial, siempre son de la forma

$$\vec{x} = \vec{z}_0 + t_1\vec{y}_1 + \cdots + t_k\vec{y}_k$$

donde k es el número de variables libres y $\vec{y}_1, \dots, \vec{y}_k$ son soluciones del sistema *homogéneo* $A\vec{x} = \vec{0}$. Véase la Observación 3.12. Por lo tanto podemos escribir

$$\begin{aligned} U &= \{\vec{x} \in \mathbb{R}^2 : A\vec{x} = \vec{b}\} \\ &= \{\vec{x} = \vec{z}_0 + t_1\vec{y}_1 + \dots + t_k\vec{y}_k : t_1, \dots, t_k \in \mathbb{R}\} \\ &= \vec{z}_0 + \{\vec{x}_0 = t_1\vec{y}_1 + \dots + t_k\vec{y}_k : t_1, \dots, t_k \in \mathbb{R}\} \\ &= \vec{z}_0 + \{\vec{x} \in \mathbb{R}^2 : A\vec{x} = \vec{0}\}. \end{aligned}$$

Esto muestra que U es igual al *espacio vectorial* $W = \{\vec{x} \in \mathbb{R}^2 : A\vec{x} = \vec{0}\}$, pero *desplazado por el vector* $\vec{z}_0$.

Llamaremos a tales conjuntos *subespacios afines*. Son muy importantes en muchas aplicaciones. La definición formal es la siguiente.

Definición 5.13. Sea V un espacio vectorial y sea $W \subseteq V$ un subconjunto. El conjunto W se llama un *subespacio afín* si existe un $z_0 \in V$ y un *subespacio* $W_0 \subseteq V$ tal que

$$W = z_0 + W_0.$$

En otras palabras: W es un subespacio si existe un vector $v_0 \in V$ tal que

$$v_0 + W := \{v_0 + w : w \in W\}$$

es un subespacio de V (podemos tomar $v_0 = -z_0$ y $W_0 = v_0 + W$).

Veamos ejemplos de subespacios afines.

Ejemplos 5.14. Sea V un espacio vectorial. Suponemos que V es un espacio vectorial real, pero todo funciona también para un espacio vectorial complejo (solo tenemos que reemplazar $\mathbb{R}$ por $\mathbb{C}$ en todas partes).

- (I) Todo subespacio de V es también un subespacio afín con $v_0 = \mathbf{0}$.
- (II) Si fijamos z_0 y $v_1, \dots, v_k \in V$, entonces el conjunto $W := \{z_0 + \alpha_1 v_1 + \dots + \alpha_k v_k : \alpha_1, \dots, \alpha_k \in \mathbb{R}\} = z_0 + \{\alpha_1 v_1 + \dots + \alpha_k v_k : \alpha_1, \dots, \alpha_k \in \mathbb{R}\}$ es un subespacio afín de V . En general no será un subespacio.

Ejercicio. Muestre que $W := \{z_0 + \alpha_1 v_1 + \dots + \alpha_k v_k : \alpha_1, \dots, \alpha_k \in \mathbb{R}\}$ es un subespacio afín de V . Muestre que es un subespacio si y solo si $z_0 \in \text{span}\{v_1, \dots, v_k\}$.

- (III) Si W es un subespacio de V , entonces $V \setminus W$ no es un subespacio afín.

Algunos ejemplos más:

Ejemplos 5.15. ■ El conjunto de todas las soluciones de un sistema no homogéneo de ecuaciones lineales es un espacio vectorial afín si no es vacío.

- El conjunto de todas las soluciones de una ecuación diferencial lineal no homogénea es un espacio vectorial afín si no es vacío.

Ahora damos varios ejemplos y no ejemplos de subespacios de $\mathbb{R}^2$ y $\mathbb{R}^3$. Debería intentar generalizarlos a ejemplos en $\mathbb{R}^4, \mathbb{R}^5$, etc., y también intentar proponer sus propios ejemplos.

Ejemplos 5.16 (Ejemplos y no ejemplos de subespacios de $\mathbb{R}^2$).

- $W = \left\{ \begin{pmatrix} \lambda \\ 0 \end{pmatrix} : \lambda \in \mathbb{R} \right\}$ es un subespacio de $\mathbb{R}^2$. Este es, de hecho, un subespacio de la forma (II) del Ejemplo 5.11 with $z = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Observe que geoméricamente W es una recta (es el eje x).
- Para $v_1, v_2 \in \mathbb{R}$ fijos, sea $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and let $W = \{\lambda \vec{v} : \lambda \in \mathbb{R}\}$. Entonces W es un subespacio de $\mathbb{R}^2$. Geométricamente, W es el subespacio trivial $\{\vec{0}\}$ si $\vec{v} = \vec{0}$. De lo contrario, es la recta en $\mathbb{R}^2$ que pasa por el origen y es paralela al vector $\vec{v}$.

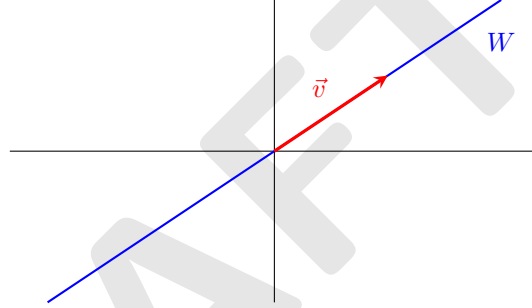


FIGURA 5.1: El subespacio W generado por el vector $\vec{v}$.

- Para fijos $a_1, a_2, v_1, v_2 \in \mathbb{R}$ let $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. Supongamos que $\vec{v} \neq \vec{0}$ y pongamos $W = \{\vec{a} + \lambda \vec{v} : \lambda \in \mathbb{R}\}$. Entonces W es un subespacio afín. Geométricamente, W representa una recta en $\mathbb{R}^2$ paralela a $\vec{v}$ que pasa por el punto (a_1, a_2) . Observe que W es un subespacio si y solo si $\vec{a}$ y $\vec{v}$ son paralelos.

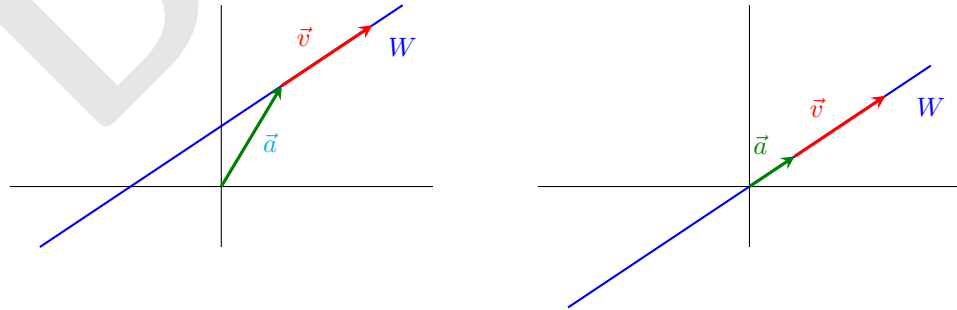


FIGURA 5.2: Bosquejos de $W = \{\vec{a} + \lambda \vec{v} : \lambda \in \mathbb{R}\}$. En la figura del lado izquierdo, $\vec{a} \nparallel \vec{v}$, así que W es un subespacio afín de $\mathbb{R}^2$, pero no un subespacio. En la figura del lado derecho, $\vec{a} \parallel \vec{v}$ y por lo tanto W es un subespacio de $\mathbb{R}^2$.

- $U = \{\vec{x} \in \mathbb{R}^2 : \|\vec{x}\| \geq 3\}$ no es un subespacio de $\mathbb{R}^2$ ya que no contiene a $\vec{0}$.
- $V = \{\vec{x} \in \mathbb{R}^2 : \|\vec{x}\| \leq 2\}$ no es un subespacio de $\mathbb{R}^2$. Por ejemplo, tome $\vec{z} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Entonces $\vec{z} \in V$, sin embargo $3\vec{z} \notin V$.
- $W = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x \geq 0 \right\}$. Entonces W no es un espacio vectorial. Por ejemplo, $\vec{z} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \in W$, pero $(-1)\vec{z} = \begin{pmatrix} -2 \\ 0 \end{pmatrix} \notin W$.

Observe que geométicamente W es un semiplano derecho en $\mathbb{R}^2$.

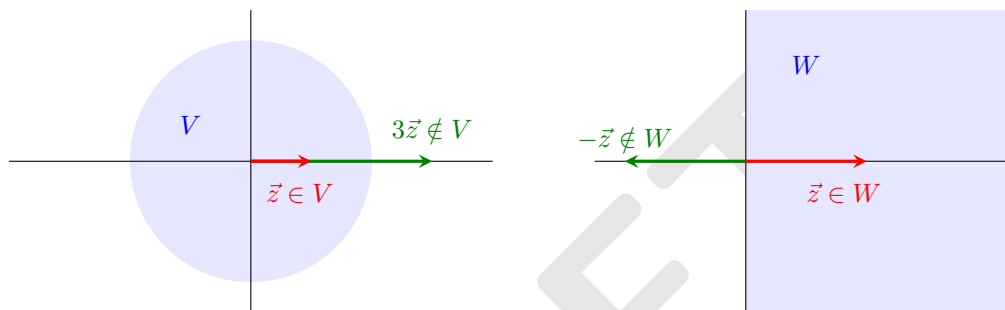


FIGURA 5.3: Los conjuntos V y W en las figuras no son subespacios de $\mathbb{R}^2$.

Ejemplos 5.17 (Ejemplos y no ejemplos de subespacios de $\mathbb{R}^3$).

- Para $x_0, y_0, z_0 \in \mathbb{R}$ fijos, sea $W = \left\{ \lambda \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} : \lambda \in \mathbb{R} \right\}$. Entonces W es un subespacio de $\mathbb{R}^3$.

Geométicamente, W es una recta en $\mathbb{R}^3$ que pasa por el origen y es paralela al vector $\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$.

- Para $a, b, c \in \mathbb{R}$ fijos, el conjunto $W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : ax + by + cz = 0 \right\}$ es un subespacio de $\mathbb{R}^3$.

Demostración. Usamos la Proposición 5.10 para verificar que W es un subespacio de $\mathbb{R}^3$.

Claramente, $\vec{0} \in W$ ya que $0a + 0b + 0c = 0$. Ahora sean $\vec{w}_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$ and $\vec{w}_2 = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$ in W y sea $\lambda \in \mathbb{R}$. Entonces $\vec{w}_1 + \vec{w}_2 \in W$ porque

$$a(x_1 + x_2) + b(y_1 + y_2) + c(z_1 + z_2) = (ax_1 + by_1 + cz_1) + (ax_2 + by_2 + cz_2) = 0 + 0 = 0.$$

También $\lambda\vec{w}_1 \in W$ porque

$$a(\lambda x_1) + b(\lambda y_1) + c(\lambda z_1) = \lambda(ax_1 + by_1 + cz_1) = \lambda 0 = 0.$$

Por lo tanto, W es cerrado bajo suma y producto por escalares, así que es un subespacio de $\mathbb{R}^3$. $\square$

Observación. Observe que W es el conjunto de todas las soluciones de un sistema lineal homogéneo de ecuaciones (una ecuación con tres incógnitas). Por lo tanto, W es un espacio vectorial por el Teorema 3.22, donde se muestra que la suma y el producto por un escalar de dos soluciones de un sistema lineal homogéneo son nuevamente una solución.

Observación. Si $a = b = c = 0$, entonces $W = \mathbb{R}^3$. Si al menos uno de los números $a, b, c \in \mathbb{R}$ es distinto de cero, entonces W es un plano en $\mathbb{R}^3$ que pasa por el origen y tiene vector normal

$$\vec{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

- Para $a, b, c, d \in \mathbb{R}$ fijos, con $d \neq 0$ y al menos uno de los números a, b, c distinto de 0, el conjunto $W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : ax + by + cz = d \right\}$ **no** es un subespacio de $\mathbb{R}^3$; véase la Figura 5.4, pero sí es un subespacio afín.

Demostración. Veamos que W no es un espacio vectorial. Let $\vec{w}_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$ and $\vec{w}_2 = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$ en W . Entonces $\vec{w}_1 + \vec{w}_2 \notin W$ porque

$$a(x_1 + x_2) + b(y_1 + y_2) + c(z_1 + z_2) = (ax_1 + by_1 + cz_1) + (ax_2 + by_2 + cz_2) = d + d = 2d \neq d.$$

(Alternativamente, podríamos haber mostrado que si $\vec{w} \in W$ y $\lambda \in \mathbb{R} \setminus \{1\}$, entonces $\lambda\vec{w} \notin W$; o podríamos haber mostrado que $\vec{0} \notin W$.)

Sabemos que W es un plano en $\mathbb{R}^3$ que tiene vector normal $\vec{n} = (a, b, c)^t$, pero no pasa por el origen. Esto muestra que W es un espacio vectorial afín porque puede escribirse como $W = \vec{v}_0 + W_0$, donde W_0 es el plano paralelo a W que pasa por el origen y $\vec{v}_0$ es un vector arbitrario desde el origen hasta un punto del plano W . (Observe que W_0 es el plano descrito por $ax + by + cz = 0$.)

Observe que ya mostramos en el Corolario 3.23 que W es un espacio vectorial afín. $\square$

Observación. Si $a = b = c = 0$, entonces $W = \emptyset$.

- $W = \{\vec{x} \in \mathbb{R}^3 : \|\vec{x}\| \geq 5\}$ no es un subespacio de $\mathbb{R}^3$ ya que no contiene a $\vec{0}$.
- $W = \{\vec{x} \in \mathbb{R}^3 : \|\vec{x}\| \leq 9\}$ no es un subespacio de $\mathbb{R}^3$. Por ejemplo, tome $\vec{z} = \begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix}$. Entonces $\vec{z} \in W$; sin embargo, por ejemplo, $7\vec{z} \notin W$ (or: $\vec{z} + \vec{z} \notin W$).

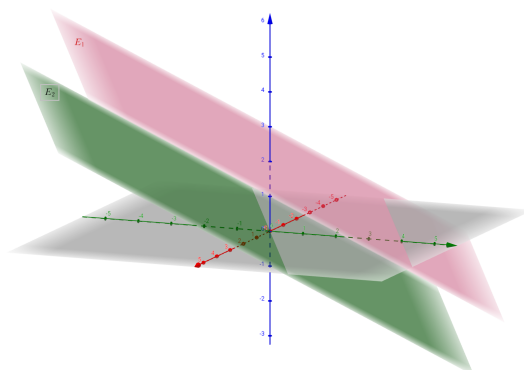


FIGURA 5.4: El plano verde pasa por el origen y es un subespacio de $\mathbb{R}^3$. El plano rojo no pasa por el origen y por lo tanto es un subespacio *afín* de $\mathbb{R}^3$.

- $W = \left\{ \begin{pmatrix} x \\ x^2 \\ x^3 \end{pmatrix} : x \in \mathbb{R} \right\}$. Entonces W no es un espacio vectorial. Por ejemplo, $\vec{a} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in W$, pero $2\vec{a} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \notin W$.

Ejemplos 5.18 (Ejemplos y no ejemplos de subespacios de $M(m \times n)$). Los siguientes conjuntos son ejemplos de subespacios de $M(m \times n)$:

- El conjunto de todas las matrices con $a_{11} = 0$.
- El conjunto de todas las matrices con $a_{11} = 5a_{12}$.
- El conjunto de todas las matrices cuya primera fila es igual a su última fila.

Si $m = n$, entonces también los siguientes conjuntos son subespacios de $M(n \times n)$:

- El conjunto de todas las matrices simétricas.
- El conjunto de todas las matrices antisimétricas.
- El conjunto de todas las matrices diagonales.
- El conjunto de todas las matrices triangulares superiores.
- El conjunto de todas las matrices triangulares inferiores.

Los siguientes conjuntos **no** son subespacios de $M(n \times n)$:

- El conjunto de todas las matrices invertibles.

- El conjunto de todas las matrices no invertibles.
- El conjunto de todas las matrices con determinante igual a 1.

Ejemplos 5.19 (Ejemplos y no ejemplos de subespacios del conjunto de todas las funciones de $\mathbb{R}$ en $\mathbb{R}$). Sea V el conjunto de todas las funciones de $\mathbb{R}$ en $\mathbb{R}$. Entonces V claramente es un espacio vectorial real. Los siguientes conjuntos son ejemplos de subespacios de V :

- El conjunto de todas las funciones continuas.
- El conjunto de todas las funciones diferenciables.
- El conjunto de todas las funciones acotadas.
- El conjunto de todos los polinomios.
- El conjunto de todos los polinomios de grado ≤ 5 .
- El conjunto de todas las funciones f con $f(7) = 0$.
- El conjunto de todas las funciones pares.
- El conjunto de todas las funciones impares.

Los siguientes conjuntos **no** son subespacios de V :

- El conjunto de todos los polinomios de grado 3.
- El conjunto de todos los polinomios de grado ≥ 3 .
- El conjunto de todas las funciones f con $f(7) = 13$.
- El conjunto de todas las funciones f con $f(7) \geq 0$.

Demuestre las afirmaciones anteriores.

Definición 5.20. Para $n \in \mathbb{N}_0$, sea P_n el conjunto de todos los polinomios de grado menor o igual que n .

Observación 5.21. P_n es un espacio vectorial.

Demostración. Claramente, la función cero pertenece a P_n (es un polinomio de grado 0). Para polinomios $p, q \in P_n$ y números $\lambda \in \mathbb{R}$, claramente tenemos que $p + q$ y λp son nuevamente polinomios de grado a lo sumo n , así que pertenecen a P_n . Por la Proposición 5.10, P_n es un subespacio del espacio de todas las funciones reales; por lo tanto es un espacio vectorial. $\square$

Debería haber entendido

- el concepto de subespacio de un espacio vectorial dado,
- por qué solo tenemos que verificar si un subconjunto dado de un espacio vectorial es no vacío, cerrado bajo suma y cerrado bajo multiplicación por escalares si queremos ver si es

un subespacio,

- etc.

Ahora debería ser capaz de

- dar ejemplos y no ejemplos de subespacios de espacios vectoriales,
- verificar si un subconjunto dado de un espacio vectorial es un subespacio,
- etc.

Ejercicios.

En los Ejercicios 1 al 13 diga si el subconjunto W es subespacio del espacio vectorial V .

1. Sea $A \in M(m \times n)$, $V = \mathbb{R}^n$ y $W = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}$.
2. Sea $x \in \mathbb{R}^n$, $V = M(n \times n)$ y $W = \{A \in V : A\vec{x} = \vec{0}\}$.
3. Sea $A \in M(n \times n)$, $V = M(n \times n)$ y $W = \{B \in V : AB = BA\}$.
4. Sea $w \in \mathbb{R}^n$ un vector no nulo, $V = \mathbb{R}^n$ y $W = \{\vec{x} \in V : \vec{x} \perp \vec{w}\}$.
5. Sea $A \in M(n \times n)$, $V = M(n \times n)$ y $W = \{B \in V : AB = \mathbb{O}\}$.
6. $V = \mathbb{R}^3$ y $W = \left\{ \begin{pmatrix} z+2y \\ y+3z \\ z \end{pmatrix} : z, y \in \mathbb{R} \right\}$.
7. $V = M(n \times n)$ y W el conjunto de todas las matrices con determinante cero.
8. $V = M(2 \times 2)$ y $W = \left\{ \begin{pmatrix} a & a+1 \\ 0 & 0 \end{pmatrix} : a, b \in \mathbb{R} \right\}$.
9. $V = \mathbb{R}^3$ y W el plano xy .
10. $V = P_n$ y $W = \{p \in P_n : p''(0) = 0\}$.
11. $V = C(\mathbb{R})$ y $W = \{f \in V : f(x^2) = 2f(x)\}$.
12. $V = C(\mathbb{R})$ y $W = \{f \in V : f(x^2) = (f(x))^2\}$.
13. $V = C(\mathbb{R})$ y $W = \{f \in V : \int_0^1 f = 0\}$.
14. $V = C^1(\mathbb{R})$ (funciones diferenciables en $\mathbb{R}$) y $W = \{f \in V : f'(0) = f'(1)\}$.
15. Sea $V = \mathbb{R}^n$ y $W = \{\vec{x} \in V : \langle x, \vec{e}_1 \rangle + \langle x, \vec{e}_2 \rangle + \cdots + \langle x, \vec{e}_n \rangle = 0\}$. Muestre que W es subespacio vectorial de V .
16. Sea $V = \mathbb{R}^3$:
 - (a) Sea E un plano que pasa por el origen y $W = \{\vec{x} \in V : \vec{x} \perp E\}$. Muestre que W es un subespacio vectorial de V . ¿Puede decir quién es W ?
 - (b) Sea $\vec{w} \in \mathbb{R}^3$ y $W = \{\vec{x} \in V : \text{proy}_w \vec{x} = \vec{0}\}$. Muestre que W es un subespacio vectorial de V . ¿Puede decir quién es W ?

(c) Sea $\vec{w} \in \mathbb{R}^3$ y $W = \{\vec{x} \in V : \vec{x} \times \vec{w} = \vec{0}\}$. Muestre que W es un subespacio vectorial de V . ¿Puede decir quién es W ?

17. En $V = M(2 \times 2)$, sean $W_1 = \{A \in V : a_{22} = 0\}$ y $W_2 = \left\{ \begin{pmatrix} -b & a \\ a & b \end{pmatrix} : a, b \in \mathbb{R} \right\}$.

(a) Muestre que W_1 y W_2 son subespacios vectoriales de V .

(b) Determine $W_1 \cap W_2$ y muestre que también es un subespacio de V .

5.3. Combinaciones lineales e independencia lineal

En esta sección trabajamos principalmente con espacios vectoriales reales, para fijar ideas. Sin embargo, todos los enunciados también son verdaderos para espacios vectoriales complejos. Solo tenemos que reemplazar $\mathbb{R}$ por $\mathbb{C}$ y la palabra *real* por *complejo* en todas partes.

Comenzamos con una definición.

Definición 5.22. Sea V un espacio vectorial real y sean $v_1, \dots, v_k \in V$ y $\alpha_1, \dots, \alpha_k \in \mathbb{R}$. Entonces todo vector de la forma

$$v = \alpha_1 v_1 + \dots + \alpha_k v_k \quad (5.1)$$

se llama una *combinación lineal de los vectores* $v_1, \dots, v_k \in V$.

Ejemplos 5.23. ■ Sea $V = \mathbb{R}^3$ y sean $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$, $\vec{a} = \begin{pmatrix} 9 \\ 12 \\ 15 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}$.

Entonces $\vec{a}$ y $\vec{b}$ son combinaciones lineales de $\vec{v}_1$ y $\vec{v}_2$ porque $\vec{a} = \vec{v}_1 + 2\vec{v}_2$ y $\vec{b} = -\vec{v}_1 + \vec{v}_2$.

■ Sea $V = M(2 \times 2)$ y sean $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $R = \begin{pmatrix} 5 & 7 \\ -7 & 5 \end{pmatrix}$, $S = \begin{pmatrix} 1 & 2 \\ -2 & 3 \end{pmatrix}$.

Entonces R es una combinación lineal de A y B porque $R = 5A + 7B$. S **no** es una combinación lineal de A y B porque claramente toda combinación lineal de A y B es de la forma

$$\alpha A + \beta B = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

de modo que nunca puede ser igual a S , ya que S tiene dos números diferentes en su diagonal.

Definición y Teorema 5.24. *i* Sea V un espacio vectorial real y sean $v_1, \dots, v_k \in V$. Entonces el conjunto de todas sus posibles combinaciones lineales se denota por

$$\text{span}\{v_1, \dots, v_k\} := \{\alpha_1 v_1 + \dots + \alpha_k v_k : \alpha_1, \dots, \alpha_k \in \mathbb{R}\}.$$

Es un subespacio de V y se llama el *span lineal* o el subespacio generado por los vectores $v_1, \dots, v_k$. Los vectores $v_1, \dots, v_k$ se llaman *generadores del subespacio* $\text{span}\{v_1, \dots, v_k\}$.

Observación. Por definición, el espacio vectorial generado por el conjunto vacío es el espacio vectorial que consiste solamente en el vector cero, es decir, $\text{span}\{\} := \{\mathbb{0}\}$.

Observación. Otros nombres para “linear span” que se usan comúnmente son *subespacio generado por* $v_1, \dots, v_k$ o *subespacio generado por* $v_1, \dots, v_k$. En lugar de $\text{span}\{v_1, \dots, v_k\}$ se usa frecuentemente la notación $\text{gen}\{v_1, \dots, v_k\}$. Todos estos nombres y notaciones significan exactamente lo mismo.

Demostración del Teorema 5.24. Tenemos que mostrar que $W := \text{span}\{v_1, \dots, v_k\}$ es un subespacio de V . Para ello usamos nuevamente la Proposición 5.10. Claramente, W no es vacío, ya que al menos $\mathbb{0} \in W$ (solo tenemos que escoger todos los $\alpha_j = 0$). Ahora sean $u, w \in W$ y $\lambda \in \mathbb{R}$. Tenemos que mostrar que $\lambda u + w \in W$. Como $u, w \in W$, existen números reales $\alpha_1, \dots, \alpha_k$ y $\beta_1, \dots, \beta_k$ tales que $u = \alpha_1 v_1 + \dots + \alpha_k v_k$ y $w = \beta_1 v_1 + \dots + \beta_k v_k$. Entonces

$$\begin{aligned} \lambda u + w &= \lambda(\alpha_1 v_1 + \dots + \alpha_k v_k) + \beta_1 v_1 + \dots + \beta_k v_k \\ &= \lambda \alpha_1 v_1 + \dots + \lambda \alpha_k v_k + \beta_1 v_1 + \dots + \beta_k v_k \\ &= (\lambda \alpha_1 + \beta_1) v_1 + \dots + (\lambda \alpha_k + \beta_k) v_k \end{aligned}$$

lo cual pertenece a W , ya que es una combinación lineal de los vectores $v_1, \dots, v_k$. $\square$

Observación. Los generadores de un subespacio dado no son únicos.

Por ejemplo, sean $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Entonces

$$\begin{aligned} \text{span}\{A, B\} &= \{\alpha A + \beta B : \alpha, \beta \in \mathbb{R}\} = \left\{ \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} : \alpha, \beta \in \mathbb{R} \right\}, \\ \text{span}\{A, B, C\} &= \{\alpha A + \beta B + \gamma C : \alpha, \beta, \gamma \in \mathbb{R}\} = \left\{ \begin{pmatrix} \alpha + \gamma & -(\beta + \gamma) \\ \beta + \gamma & \alpha + \gamma \end{pmatrix} : \alpha, \beta, \gamma \in \mathbb{R} \right\}, \\ \text{span}\{A, C\} &= \{\alpha A + \gamma C : \alpha, \gamma \in \mathbb{R}\} = \left\{ \begin{pmatrix} \alpha + \gamma & -\gamma \\ \gamma & \alpha + \gamma \end{pmatrix} : \alpha, \gamma \in \mathbb{R} \right\}. \end{aligned}$$

Vemos que $\text{span}\{A, B\} = \text{span}\{A, B, C\} = \text{span}\{A, C\}$ (en todos los casos consiste exactamente en aquellas matrices cuyas entradas diagonales son iguales y cuyas entradas fuera de la diagonal difieren por un signo menos). Así vemos que ni los generadores ni su número son únicos.

Observación. Si un vector es una combinación lineal de otros vectores, entonces los coeficientes en la combinación lineal no son necesariamente únicos.

Por ejemplo, si A, B, C son las matrices anteriores, entonces $A + B + C = 2A + 2B = 2C$ or $A + 2B + 3C = 4A + 5B = B + 4C$, etc.

Observación 5.25. Sea V un espacio vectorial y sean $v_1, \dots, v_n$ y $w_1, \dots, w_m$ vectores en V . Entonces las siguientes afirmaciones son equivalentes:

(I) $\text{span}\{v_1, \dots, v_n\} = \text{span}\{w_1, \dots, w_m\}$.

(II) $v_j \in \text{span}\{w_1, \dots, w_m\}$ for every $j = 1, \dots, n$ y $w_k \in \text{span}\{v_1, \dots, v_n\}$ for every $k = 1, \dots, m$.

Demostración. (I) $\implies$ (II) es claro.

(II) $\implies$ (I): Observe que $v_j \in \text{span}\{w_1, \dots, w_m\}$ para todo $j = 1, \dots, n$ implica que todo v_j puede escribirse como una combinación lineal de los $w_1, \dots, w_m$. Entonces también toda combinación lineal de $v_1, \dots, v_n$ es una combinación lineal de $w_1, \dots, w_m$. Esto implica que $\text{span}\{v_1, \dots, v_n\} \subseteq \text{span}\{w_1, \dots, w_m\}$. La inclusión recíproca $\text{span}\{w_1, \dots, w_m\} \subseteq \text{span}\{v_1, \dots, v_n\}$ se puede mostrar de manera análoga. Ambas inclusiones juntas muestran que debe haber igualdad. $\square$

Ejemplos 5.26. (I) $P_n = \text{span}\{1, X, X^2, \dots, X^{n-1}, X^n\}$ ya que todo vector en P_n es un polinomio de la forma $p = \alpha_n X^n + \alpha_{n-1} X^{n-1} + \dots + \alpha_1 X + \alpha_0$, por lo tanto es una combinación lineal de los polinomios $X^n, X^{n-1}, \dots, X, 1$.

Ejercicio. Muestre que $\{1, 1 + X, X + X^2, \dots, X^{n-1} + X^n\}$ es también un conjunto de generadores de P_n .

(II) El conjunto de todas las matrices antisimétricas 2×2 está generado por $\left\{ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\}$.

(III) Sea $V = \mathbb{R}^3$ y sean $\vec{v}, \vec{w} \in \mathbb{R}^3 \setminus \{\vec{0}\}$.

- $\text{span}\{\vec{v}\}$ es una recta que pasa por el origen y es paralela a $\vec{v}$.
- Si $\vec{v} \not\parallel \vec{w}$, entonces $\text{span}\{\vec{v}, \vec{w}\}$ es un plano que pasa por el origen y es paralelo a $\vec{v}$ y $\vec{w}$.
Si $\vec{v} \parallel \vec{w}$, entonces es una recta que pasa por el origen y es paralela a $\vec{v}$.

Ejemplo 5.27. Sean $p_1 = X^2 - X + 1, p_2 = X^2 - 2X + 5 \in P_2$, y sea $U = \text{span}\{p_1, p_2\}$. Verifique si $q = 2X^2 - X - 2$ y $r = X^2 + X - 3$ pertenecen a U .

Solución. ▪ Verifiquemos si $q \in U$. Para ello tenemos que verificar si podemos encontrar α, β tales que $q = \alpha p_1 + \beta p_2$. Sustituyendo las expresiones de p_1, p_2, q obtenemos

$$2X^2 - X - 2 = \alpha(X^2 - X + 1) + \beta(X^2 - 2X + 5) = X^2(\alpha + \beta) + X(-\alpha - 2\beta) + \alpha + 5\beta.$$

Comparando coeficientes de las distintas potencias de X , obtenemos el sistema de ecuaciones

$$\begin{aligned} \alpha + \beta &= 2, \\ -\alpha - 2\beta &= -1, \\ \alpha + 5\beta &= -2. \end{aligned}$$

Usamos el proceso de Gauß-Jordan para resolver el sistema:

$$A = \left(\begin{array}{cc|c} 1 & 1 & 2 \\ -1 & -2 & -1 \\ 1 & 5 & -2 \end{array} \right) \longrightarrow \left(\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & -1 & 1 \\ 0 & 4 & -4 \end{array} \right) \longrightarrow \left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{array} \right).$$

Se sigue que $\alpha = 3$ y $\beta = -1$ es una solución, y por lo tanto $q = 2p_1 - p_2$, lo cual muestra que $q \in U$.

- Verifiquemos si $r \in U$. Para ello tenemos que verificar si podemos encontrar α, β tales que $r = \alpha p_1 + \beta p_2$. Sustituyendo las expresiones de p_1, p_2, q obtenemos

$$X^2 + X - 3 = \alpha(X^2 - X + 1) + \beta(X^2 - 2X + 5) = X^2(\alpha + \beta) + X(-\alpha - 2\beta) + \alpha + 5\beta.$$

Comparando coeficientes de las distintas potencias de X , obtenemos el sistema de ecuaciones

$$\begin{aligned}\alpha + \beta &= 1, \\ -\alpha - 2\beta &= 1, \\ \alpha + 5\beta &= -3.\end{aligned}$$

Usamos el proceso de Gauß-Jordan para resolver el sistema:

$$A = \left(\begin{array}{cc|c} 1 & 1 & 1 \\ -1 & -2 & 1 \\ 1 & 5 & -3 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & -1 & 2 \\ 0 & 4 & -4 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 4 \end{array} \right).$$

Vemos que el sistema es inconsistente. Por lo tanto r no es una combinación lineal de p_1 y p_2 , de modo que $r \notin U$. $\diamond$

Definición 5.28. Un espacio vectorial V se llama *finitamente generado* si tiene un conjunto finito de generadores.

Ejemplos 5.29. Los siguientes espacios vectoriales son finitamente generados.

- El espacio vectorial trivial $\{\mathbb{O}\}$ es finitamente generado.
- $\mathbb{R}^n$ porque claramente $\mathbb{R}^n = \text{span}\{\vec{e}_1, \dots, \vec{e}_n\}$, donde $\vec{e}_j$ es el j -ésimo vector unitario.
- $M(m \times n)$ porque está generado por el conjunto de todas las matrices posibles que son 0 en todas partes excepto por un 1 en exactamente una entrada.
- P_n es finitamente generado, como se mostró en el Ejemplo 5.26.
- Sea P el espacio vectorial de todos los polinomios reales. Entonces P no es finitamente generado.

Demostración. Supongamos que P es finitamente generado y sea $q_1, \dots, q_k$ un sistema de generadores de P . Observe que los q_j son polinomios. Denotaremos sus grados por $m_j = \deg q_j$ y ponemos $M = \max\{m_1, \dots, m_k\}$. Entonces cualquier combinación lineal de ellos será un polinomio de grado a lo sumo M , sin importar cómo escojamos los coeficientes. Sin embargo, hay elementos en P que tienen grado mayor, por ejemplo X^{m+1} . Por lo tanto $q_1, \dots, q_k$ no pueden generar todo P . $\square$

Otra demostración usando el concepto de dimensión se dará en el Ejemplo 5.56 (f).

Más adelante, en el Lema 5.53, veremos que todo subespacio de un espacio vectorial finitamente generado es nuevamente finitamente generado.

Ahora nos preguntamos cuál es el menor número de vectores que necesitamos para generar $\mathbb{R}^n$. Sabemos, por ejemplo, que $\mathbb{R}^n = \text{span}\{\vec{e}_1, \dots, \vec{e}_n\}$. Así que en este caso tenemos n vectores que

generan $\mathbb{R}^n$. ¿Podría ser que fueran suficientes menos vectores? Claramente, si quitamos uno de los $\vec{e}_j$, entonces el sistema restante ya no genera $\mathbb{R}^n$, ya que “falta una coordenada”. Sin embargo, ¿podríamos quizá encontrar otros vectores de modo que $n - 1$ o menos vectores sean suficientes para generar todo $\mathbb{R}^n$? La siguiente proposición dice que esto no es posible.

Proposición 5.30. Sean $\vec{v}_1, \dots, \vec{v}_k$ vectores en $\mathbb{R}^n$. Si $\text{span}\{\vec{v}_1, \dots, \vec{v}_k\} = \mathbb{R}^n$, entonces $k \geq n$.

Demostración. Sea $A = (\vec{v}_1 | \dots | \vec{v}_k)$ la matriz cuyas columnas son los vectores dados. Sabemos que existe una matriz invertible E tal que $A' = EA$ está en forma escalonada reducida (la matriz E es el producto de matrices elementales que corresponden a los pasos del proceso de Gauß-Jordan para llegar a la forma escalonada reducida). Ahora, si $k < n$, entonces sabemos que A' debe tener al menos una fila que consiste solamente en ceros. Si podemos encontrar un vector $\vec{w}$ tal que se transforme en $\vec{e}_n$ bajo el proceso de Gauß-Jordan, entonces tendríamos que $A\vec{x} = \vec{w}$ es inconsistente, lo que significa que $\vec{w} \notin \text{span}\{\vec{v}_1, \dots, \vec{v}_k\}$. ¿Cómo encontramos tal vector $\vec{w}$? Bueno, solo tenemos que comenzar con $\vec{e}_n$ y “hacer el proceso de Gauß-Jordan al revés”. En otras palabras, podemos tomar $\vec{w} = E^{-1}\vec{e}_n$. Ahora, si aplicamos el proceso de Gauß-Jordan a la matriz aumentada $(A|\vec{w})$, llegamos a $(EA|E\vec{w}) = (A'|\vec{e}_n)$, que ya sabemos que es inconsistente.

Por lo tanto, $k < n$ no es posible y, por lo tanto, debemos tener que $k \geq n$. $\square$

Observe que la demostración anterior es básicamente la misma que la de la Observación 3.37. Observe que el sistema de vectores $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{R}^n$ es un conjunto de generadores de $\mathbb{R}^n$ si y solo si la ecuación $A\vec{y} = \vec{b}$ tiene solución para todo $\vec{b} \in \mathbb{R}^n$ (como arriba, A es la matriz cuyas columnas son los vectores $\vec{v}_1, \dots, \vec{v}_k$).

Ahora responderemos la pregunta de cuándo los coeficientes de una combinación lineal son únicos. La siguiente observación nos muestra que solo tenemos que responder esta pregunta para el vector cero.

Observación 5.31. Sea V un espacio vectorial, sean $v_1, \dots, v_k \in V$ y sea $w \in \text{span}\{v_1, \dots, v_k\}$. Entonces existen únicos $\beta_1, \dots, \beta_k \in \mathbb{R}$ tales que

$$\beta_1 v_1 + \dots + \beta_k v_k = w \quad (5.2)$$

si y solo si existen únicos $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ tales que

$$\alpha_1 v_1 + \dots + \alpha_k v_k = \mathbb{O}. \quad (5.3)$$

Demostración. Primero observe que (5.3) siempre tiene al menos una solución, a saber, $\alpha_1 = \dots = \alpha_k = 0$. Esta solución se llama la *solución trivial*.

Supongamos que (5.2) tiene dos soluciones distintas, de modo que existen $\gamma_1, \dots, \gamma_k \in \mathbb{R}$ tales que para al menos un $j = 1, \dots, k$ tenemos $\beta_j \neq \gamma_j$ y

$$\gamma_1 v_1 + \dots + \gamma_k v_k = w. \quad (5.2')$$

Restando (5.2) y (5.2') obtenemos

$$(\beta_1 - \gamma_1)v_1 + \dots + (\beta_k - \gamma_k)v_k = w - w = \mathbb{O}$$

donde al menos un coeficiente es distinto de cero. Por lo tanto, (5.3) también tiene más de una solución.

Por otro lado, supongamos que (5.3) tiene una solución no trivial, es decir, al menos uno de los α_j en (5.3) es distinto de cero. Pero entonces, si sumamos (5.2) y (5.3), obtenemos otra solución para (5.2) porque

$$(\alpha_1 + \beta_1)v_1 + \cdots + (\alpha_k + \beta_k)v_k = \mathbb{O} + w = w. \quad \square$$

La demostración muestra que hay tantas soluciones de (5.2) como de (5.3).

También debe observarse que si (5.3) tiene una solución no trivial, entonces automáticamente tiene infinitas soluciones, porque si $\alpha_1, \dots, \alpha_k$ es una solución, entonces también $c\alpha_1, \dots, c\alpha_k$ es una solución para $c \in \mathbb{R}$ arbitrario, ya que

$$c\alpha_1 v_1 + \cdots + c\alpha_k v_k = c(\alpha_1 v_1 + \cdots + \alpha_k v_k) = c\mathbb{O} = \mathbb{O}.$$

De hecho, la discusión anterior debería recordarle la relación entre las soluciones de un sistema no homogéneo y las soluciones de su sistema homogéneo asociado en el Teorema 3.22. Observe que, tal como en el caso de los sistemas lineales, (5.2) podría no tener solución. Esto ocurre si y solo si $w \notin \text{span}\{v_1, \dots, v_k\}$.

Si $V = \mathbb{R}^n$, entonces la observación anterior es exactamente el Teorema 3.22.

Así vemos que solo puede ocurrir uno de los dos casos siguientes: (5.4) tiene exactamente una solución (la trivial) o tiene infinitas soluciones. Observe que esto es análogo a la situación de las soluciones de sistemas lineales homogéneos: o bien tienen solamente la solución trivial, o bien tienen infinitas soluciones. La siguiente definición distingue entre los dos casos.

Definición 5.32. Sea V un espacio vectorial. El conjunto de vectores $v_1, \dots, v_k$ en V se llama *linealmente independiente* si

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = \mathbb{O} \quad (5.4)$$

tiene solamente la solución trivial. Se llaman *linealmente dependientes* si (5.4) tiene más de una solución.

Observación 5.33. El conjunto vacío es linealmente independiente, ya que $\mathbb{O}$ no puede escribirse como una combinación lineal no trivial de vectores del conjunto vacío.

Antes de continuar con la teoría, damos algunos ejemplos.

Ejemplos. (I) Los vectores $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ y $\vec{v}_2 = \begin{pmatrix} -4 \\ -8 \end{pmatrix} \in \mathbb{R}^2$ son linealmente dependientes porque $4\vec{v}_1 + \vec{v}_2 = \vec{0}$.

(II) Los vectores $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ y $\vec{v}_2 = \begin{pmatrix} 3 \\ 0 \end{pmatrix} \in \mathbb{R}^2$ son linealmente independientes.

Demostración. Consideremos la ecuación $\alpha \vec{v}_1 + \beta \vec{v}_2 = \vec{0}$. Esta ecuación es equivalente al siguiente sistema de ecuaciones lineales para α y β :

$$\begin{aligned}\alpha + 3\beta &= 0, \\ 2\alpha + 0\beta &= 0.\end{aligned}$$

Podemos usar el proceso de Gauß-Jordan para obtener todas las soluciones. Sin embargo, en este caso vemos fácilmente que $\alpha = 0$ (de la segunda línea) y luego que $\beta = -\frac{1}{3}\alpha = 0$. Observe que también podríamos haber calculado $\det \begin{pmatrix} 1 & 3 \\ 2 & 0 \end{pmatrix} = -6 \neq 0$ para concluir que el sistema homogéneo anterior tiene solamente la solución trivial. Observe que las columnas de la matriz son exactamente los vectores dados. $\square$

(III) Los vectores $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ y $\vec{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \in \mathbb{R}^3$ son linealmente independientes.

Demostración. Consideremos la ecuación $\alpha \vec{v}_1 + \beta \vec{v}_2 = \vec{0}$. Esta ecuación es equivalente al siguiente sistema de ecuaciones lineales para α y β :

$$\begin{aligned}\alpha + 2\beta &= 0, \\ \alpha + 3\beta &= 0, \\ \alpha + 4\beta &= 0.\end{aligned}$$

Si restamos la primera ecuación de la segunda, obtenemos $\beta = 0$ y luego $\alpha = -2\beta = 0$. Así, nuevamente, este sistema tiene solamente la solución trivial y por lo tanto los vectores $\vec{v}_1$ y $\vec{v}_2$ son linealmente independientes. $\square$

(IV) Sean $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ y $\vec{v}_4 = \begin{pmatrix} 0 \\ 6 \\ 8 \end{pmatrix} \in \mathbb{R}^3$ Entonces

- (a) El sistema $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es linealmente independiente.
- (b) El sistema $\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$ es linealmente dependiente.

Demostración. (a) Consideremos la ecuación $\alpha \vec{v}_1 + \beta \vec{v}_2 + \gamma \vec{v}_3 = \vec{0}$. Esta ecuación es equivalente al siguiente sistema de ecuaciones lineales para α, β y γ :

$$\begin{aligned}\alpha - 1\beta + 0\gamma &= 0, \\ \alpha + 2\beta + 0\gamma &= 0, \\ \alpha + 3\beta + 1\gamma &= 0.\end{aligned}$$

Usamos el proceso de Gauß-Jordan para resolver el sistema. Observe que las columnas de la matriz asociada al sistema anterior son exactamente los vectores dados $\vec{v}_1, \vec{v}_2, \vec{v}_3$.

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 2 & 0 \\ 1 & 3 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & 0 \\ 0 & 3 & 0 \\ 0 & 4 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Por lo tanto, la única solución es $\alpha = \beta = \gamma = 0$ y, en consecuencia, los vectores $\vec{v}_1, \vec{v}_2, \vec{v}_3$ son linealmente independientes.

Observe que también podríamos haber calculado $\det A = 3 \neq 0$ para concluir que el sistema homogéneo tiene solamente la solución trivial.

- (b) Consideremos la ecuación $\alpha\vec{v}_1 + \beta\vec{v}_2 + \delta\vec{v}_4 = \vec{0}$. Esta ecuación es equivalente al siguiente sistema de ecuaciones lineales para α, β y δ :

$$\begin{aligned}\alpha - 1\beta + 0\delta &= 0, \\ \alpha + 2\beta + 6\delta &= 0, \\ \alpha + 3\beta + 8\delta &= 0.\end{aligned}$$

Usamos el proceso de Gauß-Jordan para resolver el sistema. Observe que las columnas de la matriz asociada al sistema anterior son exactamente los vectores dados.

$$\begin{aligned}A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 2 & 6 \\ 1 & 3 & 8 \end{pmatrix} &\rightarrow \begin{pmatrix} 1 & -1 & 0 \\ 0 & 3 & 6 \\ 0 & 4 & 8 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \\ &\rightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.\end{aligned}$$

Así que hay infinitas soluciones. Si tomamos $\delta = t$, entonces $\alpha = \beta = -2t$. En consecuencia, los vectores $\vec{v}_1, \vec{v}_2, \vec{v}_4$ son linealmente dependientes, porque, por ejemplo, $-2\vec{v}_1 - 2\vec{v}_2 + \vec{v}_4 = 0$ (tomando $t = 1$).

Observe que también podríamos haber calculado $\det A = 0$ para concluir que el sistema tiene infinitas soluciones. $\square$

- (v) Las matrices $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ son linealmente independientes en $M(2 \times 2)$.

- (vi) Las matrices $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ son linealmente dependientes en $M(2 \times 2)$ porque $A - B - C = 0$.

Después de estos ejemplos, continuaremos con algunos hechos sobre la independencia lineal. Comenzamos con el caso especial en el que tenemos solamente dos vectores.

Proposición 5.34. *Sean v_1, v_2 vectores en un espacio vectorial V . Entonces v_1, v_2 son linealmente dependientes si y solo si un vector es múltiplo del otro.*

Demostración. Supongamos que v_1, v_2 son linealmente dependientes. Entonces existen $\alpha_1, \alpha_2 \in \mathbb{R}$ tales que $\alpha_1 v_1 + \alpha_2 v_2 = 0$ y al menos uno de α_1 y α_2 es distinto de cero, digamos $\alpha_1 \neq 0$. Entonces tenemos $v_1 + \frac{\alpha_2}{\alpha_1} v_2 = 0$, de donde $v_1 = -\frac{\alpha_2}{\alpha_1} v_2$.

Ahora supongamos, por otro lado, que, por ejemplo, v_1 es un múltiplo de v_2 , es decir, $v_1 = \lambda v_2$ for some $\lambda \in \mathbb{R}$. Entonces $v_1 - \lambda v_2 = 0$, lo cual es una solución no trivial de $\alpha_1 v_1 + \alpha_2 v_2 = 0$, porque podemos tomar $\alpha_1 = 1 \neq 0$ y $\alpha_2 = -\lambda$ (observe que λ puede ser cero). $\square$

La proposición anterior no puede extenderse al caso de tres o más vectores. Por ejemplo, los vectores $\vec{a} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\vec{c} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ son linealmente dependientes porque $\vec{a} + \vec{b} - \vec{c} = \vec{0}$, pero ninguno de ellos es múltiplo de alguno de los otros dos vectores.

Proposición 5.35. *Sea V un espacio vectorial.*

- (I) *Todo sistema de vectores que contiene a $\mathbb{O}$ es linealmente dependiente.*
- (II) *Sean $v_1, \dots, v_k \in V$ y supongamos que existen $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ tales que $\alpha_1 v_1 + \dots + \alpha_k v_k = \mathbb{O}$. Si $\alpha_\ell \neq 0$, entonces v_ℓ es una combinación lineal de los otros v_j .*
- (III) *Si los vectores $v_1, \dots, v_k \in V$ son linealmente dependientes, entonces para todo $w \in V$, los vectores $v_1, \dots, v_k, w$ son linealmente dependientes.*
- (IV) *Si $v_1, \dots, v_k$ son vectores en V y w es una combinación lineal de ellos, entonces $v_1, \dots, v_k, w$ son linealmente dependientes.*
- (V) *Si los vectores $v_1, \dots, v_k \in V$ son linealmente independientes, entonces todo subconjunto de ellos es linealmente independiente.*

Demostración. (I) Sean $v_1, \dots, v_k \in V$. Claramente $1\mathbb{O} + 0v_1 + \dots + 0v_k = \mathbb{O}$ es una combinación lineal no trivial que da $\mathbb{O}$. Por lo tanto, el sistema $\{\mathbb{O}, v_1, \dots, v_k\}$ es linealmente dependiente.

(II) Si $\alpha_\ell \neq 0$, entonces podemos despejar v_ℓ : $v_\ell = -\frac{\alpha_1}{\alpha_\ell} v_1 - \dots - \frac{\alpha_{\ell-1}}{\alpha_\ell} v_{\ell-1} - \frac{\alpha_{\ell+1}}{\alpha_\ell} v_{\ell+1} - \dots - \frac{\alpha_k}{\alpha_\ell} v_k$.

(III) Si los vectores $v_1, \dots, v_k \in V$ son linealmente dependientes, entonces existen $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ tales que al menos uno de ellos es distinto de cero y $\alpha_1 v_1 + \dots + \alpha_k v_k = \mathbb{O}$. Pero entonces también $\alpha_1 v_1 + \dots + \alpha_k v_k + 0w = \mathbb{O}$ lo cual muestra que el sistema $\{v_1, \dots, v_k, w\}$ es linealmente dependiente.

(IV) Supongamos que w es una combinación lineal de $v_1, \dots, v_k$. Entonces existen $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ tal que $w = \alpha_1 v_1 + \dots + \alpha_k v_k$. Por lo tanto obtenemos $w - \alpha_1 v_1 - \dots - \alpha_k v_k = \mathbb{O}$ que es una combinación lineal no trivial, ya que el coeficiente de w es 1.

(V) Supongamos que un subsistema de $v_1, \dots, v_k \in V$ es linealmente dependiente. Entonces, por (III), todo sistema que lo contiene también debe ser linealmente dependiente. En particular, el sistema original de vectores debe ser linealmente dependiente, lo cual contradice nuestra hipótesis. Observe que también el conjunto vacío es linealmente independiente por la Observación 5.33. $\square$

Ahora nos especializamos al caso en que $V = \mathbb{R}^n$. Tomemos vectores $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{R}^n$ y escribamos $(\vec{v}_1 | \dots | \vec{v}_k)$ para la matriz $n \times k$ cuyas columnas son los vectores $\vec{v}_1, \dots, \vec{v}_k$.

Lema 5.36. *Con la notación anterior, las siguientes afirmaciones son equivalentes:*

- (I) $\vec{v}_1, \dots, \vec{v}_k$ son linealmente dependientes.
- (II) Existen $\alpha_1, \dots, \alpha_k$, no todos iguales a cero, tales que $\alpha_1 \vec{v}_1 + \dots + \alpha_k \vec{v}_k = \mathbf{0}$.

$$(III) \text{ Existe un vector } \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix} \neq \vec{0} \text{ tal que } (\vec{v}_1 | \cdots | \vec{v}_k) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix} = \vec{0}.$$

(IV) El sistema homogéneo correspondiente a la matriz $(\vec{v}_1 | \cdots | \vec{v}_k)$ tiene al menos una solución no trivial (y por lo tanto infinitas soluciones).

Demostración. (I) $\Rightarrow$ (II) es simplemente la definición de dependencia lineal. (II) $\Rightarrow$ (III) es solamente reescribir la ecuación vectorial en forma matricial. (IV) solo dice con palabras lo que significa la ecuación en (III). And finally (IV) $\Rightarrow$ (I) holds because every non trivial solution the inhomogeneous system associated to $(\vec{v}_1 | \cdots | \vec{v}_k)$ gives a non-trivial solution of $\alpha_1 \vec{v}_1 + \cdots + \alpha_k \vec{v}_k$. $\square$

Como sabemos que un sistema lineal homogéneo con más incógnitas que ecuaciones tiene infinitas soluciones, obtenemos inmediatamente el siguiente corolario.

Corolario 5.37. Sean $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{R}^n$.

- (I) Si $k > n$, entonces los vectores $\vec{v}_1, \dots, \vec{v}_k$ son linealmente dependientes.
- (II) Si los vectores $\vec{v}_1, \dots, \vec{v}_k$ son linealmente independientes, entonces $k \leq n$.

Observe que (II) **no** dice que si $k \leq n$, entonces los vectores $\vec{v}_1, \dots, \vec{v}_k$ son linealmente independientes. Solo dice que tienen la posibilidad de ser linealmente independientes, mientras que un sistema con más de n vectores siempre es linealmente dependiente.

Ahora nos especializamos aún más al caso en que $k = n$.

Teorema 5.38. Sean $\vec{v}_1, \dots, \vec{v}_n$ vectores en $\mathbb{R}^n$. Entonces las siguientes afirmaciones son equivalentes:

- (I) $\vec{v}_1, \dots, \vec{v}_n$ son linealmente independientes.

$$(II) \text{ La única solución de } (\vec{v}_1 | \cdots | \vec{v}_n) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \vec{0} \text{ es el vector cero } \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \vec{0}.$$

(III) La matriz $(\vec{v}_1 | \cdots | \vec{v}_n)$ es invertible.

(IV) $\det(\vec{v}_1 | \cdots | \vec{v}_n) \neq 0$.

Demostración. La equivalencia de (I) y (II) se sigue del Lema 5.36. La equivalencia de (II), (III) y (IV) se sigue del Teorema 4.11. $\square$

Formule un teorema análogo para vectores linealmente dependientes.

Ahora podemos enunciar cuándo un sistema de n vectores en $\mathbb{R}^n$ genera $\mathbb{R}^n$.

Teorema 5.39. Sean $\vec{v}_1, \dots, \vec{v}_n$ vectores en $\mathbb{R}^n$. y let $A = (\vec{v}_1 | \cdots | \vec{v}_n)$ be the matrix whose columns are the given vectors $\vec{v}_1, \dots, \vec{v}_n$. Entonces las siguientes afirmaciones son equivalentes:

- (I) $\vec{v}_1, \dots, \vec{v}_n$ son linealmente independientes.

$$(II) \mathbb{R}^n = \text{span}\{\vec{v}_1, \dots, \vec{v}_n\}.$$

$$(III) \det A \neq 0.$$

Demostración. (I) $\iff$ (III) se muestra en el Teorema 5.38.

(II) $\iff$ (III): Los vectores $\vec{v}_1, \dots, \vec{v}_n$ generan $\mathbb{R}^n$ si y solo si para todo $\vec{w} \in \mathbb{R}^n$ existen números $\beta_1, \dots, \beta_n$ tales que $\beta_1 \vec{v}_1 + \dots + \beta_n \vec{v}_n = \vec{w}$. En forma matricial, esto significa que $A \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \vec{w}$. Por el Teorema 3.44 sabemos que esto tiene una solución para todo vector $\vec{w}$ si y solo si A es invertible (porque si aplicamos Gauß-Jordan a A , debemos llegar a la matriz identidad). $\square$

La demostración del teorema anterior procede básicamente así: Consideramos la ecuación $A\vec{\beta} = \vec{w}$. ¿Cuándo son linealmente independientes los vectores $\vec{v}_1, \dots, \vec{v}_n$? – Son linealmente independientes si y solo si para $\vec{w} = \vec{0}$ el sistema tiene solamente la solución trivial. Esto ocurre si y solo si la forma escalonada reducida de A es la matriz identidad. Y esto ocurre si y solo si $\det A \neq 0$.

¿Cuándo generan $\mathbb{R}^n$ los vectores $\vec{v}_1, \dots, \vec{v}_n$? – Lo hacen si y solo si para todo vector dado $\vec{w} \in \mathbb{R}^n$ el sistema tiene al menos una solución. Esto ocurre si y solo si la forma escalonada reducida de A es la matriz identidad. Y esto ocurre si y solo si $\det A \neq 0$.

Como una matriz cuadrada A es invertible si y solo si su transpuesta A^t es invertible, el Teorema 5.39 conduce inmediatamente al siguiente corolario.

Corolario 5.40. *Para una matriz $A \in M(n \times n)$ las siguientes afirmaciones son equivalentes:*

- (I) A es invertible.
- (II) Las columnas de A son linealmente independientes.
- (III) Las filas de A son linealmente independientes.

Terminamos esta sección con más ejemplos.

Ejemplos. ■ Recordemos que P_n es el espacio vectorial de todos los polinomios de grado $\leq n$.

En P_3 , consideramos los vectores $p_1 = X^3 - 1$, $p_2 = X^2 - 1$, $p_3 = X - 1$. Estos vectores son linealmente independientes.

Demostración. Sean $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ tales que $\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 = 0$. Esto significa que

$$\begin{aligned} 0 &= \alpha_1(X^3 - 1) + \alpha_2(X^2 - 1) + \alpha_3(X - 1) \\ &= \alpha_1 X^3 + \alpha_2 X^2 + \alpha_3 X - (\alpha_1 + \alpha_2 + \alpha_3). \end{aligned}$$

Comparando coeficientes, se sigue que $\alpha_1 = 0$, $\alpha_2 = 0$, $\alpha_3 = 0$ y $\alpha_1 + \alpha_2 + \alpha_3 = 0$ lo cual muestra que p_1 , p_2 y p_3 son linealmente independientes. $\square$

Si además tomamos $p_4 = X^3 - X^2$, entonces el sistema p_1, p_2, p_3 y p_4 es linealmente dependiente.

Demostración. Como antes, sean $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$ tales que $\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \alpha_4 p_4 = 0$. Esto significa que

$$\begin{aligned} 0 &= \alpha_1(X^3 - 1) + \alpha_2(X^2 - 1) + \alpha_3(X - 1) + \alpha_4(X^3 - X^2) \\ &= (\alpha_1 + \alpha_4)X^3 + (\alpha_2 - \alpha_4)X^2 + \alpha_3X - (\alpha_1 + \alpha_2 + \alpha_3). \end{aligned}$$

Comparando coeficientes, esto es equivalente a $\alpha_1 + \alpha_4 = 0$, $\alpha_2 - \alpha_4 = 0$, $\alpha_3 = 0$ y $\alpha_1 + \alpha_2 + \alpha_3 = 0$. Este sistema de ecuaciones tiene infinitas soluciones. Están dadas por $\alpha_2 = \alpha_4 = -\alpha_1 \in \mathbb{R}$, $\alpha_3 = 0$ (¡verifique esto!). Por lo tanto p_1, p_2, p_3 y p_4 son linealmente dependientes. $\square$

Ejercicio. Muestre que p_1, p_2, p_3 y p_5 son linealmente independientes si $p_5 = X^3 + X^2$.

- En P_2 , consideramos los vectores $p_1 = X^2 + 2X - 1$, $p_2 = 5X + 2$, $p_3 = 2X^2 - 11X - 8$. Estos vectores son linealmente dependientes.

Demostración. Sean $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ tales que $\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 = 0$. Esto significa que

$$\begin{aligned} 0 &= \alpha_1(X^2 + 2X - 1) + \alpha_2(5X + 2) + \alpha_3(2X^2 - 11X - 8) \\ &= (\alpha_1 + 2\alpha_3)X^2 + (2\alpha_1 + 5\alpha_2 - 11\alpha_3)X - \alpha_1 + 2\alpha_2 - 8\alpha_3. \end{aligned}$$

Comparando coeficientes, se sigue que $\alpha_1 + 2\alpha_3 = 0$, $2\alpha_1 + 5\alpha_2 - 11\alpha_3 = 0$, $-\alpha_1 + 2\alpha_2 - 8\alpha_3 = 0$. Escribimos esto en forma matricial y aplicamos Gauß-Jordan:

$$\begin{pmatrix} 1 & 0 & 2 \\ 2 & 5 & -11 \\ -1 & 2 & -8 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 5 & -15 \\ 0 & 2 & -6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -3 \\ 0 & 1 & -3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Esto muestra que el sistema tiene soluciones no triviales (¡encuéntrelas!) y por lo tanto p_1, p_2 y p_3 son linealmente dependientes. $\square$

- En $V = M(2 \times 2)$ considere $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 5 \\ 5 & 0 \end{pmatrix}$. Entonces A, B, C son linealmente dependientes porque $A - B - \frac{2}{5}C = 0$.
- En $V = M(2 \times 3)$ considere $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 1 \end{pmatrix}$. Entonces A, B, C son linealmente independientes.

Ejercicio. ¡Demuéstrelo!

- Encuentre un conjunto de generadores para el espacio vectorial

$$V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x + 2y = 0 \right\}.$$

Solución. Claramente, V es un subespacio de $\mathbb{R}^3$ (es un plano). Sean $\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in V$. Por definición de V , tenemos que $x + 2y = 0$. Podemos despejar x o y en la ecuación anterior; obtenemos $x = -2y$. So

$$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2y \\ y \\ z \end{pmatrix} = y \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Por lo tanto $\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ es un conjunto de generadores para V . ◇

Debería haber entendido

- qué es una combinación lineal,
- el concepto de independencia lineal,
- el concepto de subespacio generado y que este consiste o bien solamente en el vector cero o bien en infinitos vectores,
- geoméricamente el concepto de independencia lineal en $\mathbb{R}^2$ y $\mathbb{R}^3$,
- que los coeficientes en una combinación lineal no son necesariamente únicos,
- qué dice el número de soluciones de $A\vec{x} = \vec{0}$ sobre la independencia lineal de las columnas de A vistas como vectores en $\mathbb{R}^n$,
- qué dice la existencia (o no existencia) de soluciones de $A\vec{x} = \vec{b}$ para todo $\vec{b} \in \mathbb{R}^m$ sobre el subespacio generado por las columnas de A vistas como vectores en $\mathbb{R}^n$,
- por qué una matriz $A \in M(n \times n)$ es invertible si y solo si sus columnas son linealmente independientes,
- etc.

Ahora debería ser capaz de

- verificar si un vector dado es una combinación lineal de un conjunto dado de vectores,
- verificar si un vector dado pertenece al subespacio generado por un conjunto dado de vectores,
- verificar si un conjunto dado de vectores es un generador de un espacio vectorial dado,
- encontrar un conjunto de generadores para un espacio vectorial dado,
- verificar si un conjunto dado de vectores es linealmente independiente,
- etc.

Ejercicios.

En los ejercicios 1 al 10, encontrar un conjunto generador para el espacio dado. Antes de hacerlo, asegúrese que los conjuntos dados efectivamente son espacios vectoriales.

1. $\{\vec{x} \in \mathbb{R}^3 : x + y + z = 0\}$.
2. $\{\vec{x} \in \mathbb{R}^3 : x + y + z = 0, x - y - z = 0\}$.
3. Matrices antisimétricas de tamaño 3×3 .
4. Polinomios de grado ≤ 3 tal que $p''(0) = 0$.
5. Polinomios de grado ≤ 3 tal que $\int_0^1 xp(x)dx = 0$.
6. $\{p \in P_3 : p(0) = p(1), p'(0) = p'(1)\}$.
7. La recta en $\mathbb{R}^2$ dada por $2x + y = 0$.
8. El plano $3x + 2y - z = 0$ en $\mathbb{R}^3$.
9. La recta en $\mathbb{R}^3$ dada por $\frac{x}{2} = \frac{y}{3} = 3z$.
10. $\{p \in P_3 : p(x) = ax^3 + cx^2 + a - 2c\}$.
11. Muestre $\begin{pmatrix} 1 \\ 5 \\ 3 \\ 4 \end{pmatrix} \in \text{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$.
12. Muestre que el conjunto $\{x^3 + 1, x^3 - 5, x^2, x^2 - 1\}$ genera todos los polinomios de grado ≤ 3 tales que su primera derivada evaluada en cero vale cero.
13. Muestre que $\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$ no pertenece al generado de $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$.
14. En los siguientes ejercicios, diga si los vectores dados son linealmente independientes o dependientes.
 - (a) En $\mathbb{R}^3$: $\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \\ -8 \end{pmatrix}$.
 - (b) En $\mathbb{R}^3$: $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.
 - (c) En $M(2 \times 2)$: $\begin{pmatrix} 2 & -1 \\ 4 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -3 \\ 1 & 5 \end{pmatrix}, \begin{pmatrix} 4 & 1 \\ 7 & -5 \end{pmatrix}$.
 - (d) En P_3 : $x^3 + 1, x^3 - 5, x^2$ y $x^2 - 1$.
 - (e) En P_2 : $1 - x, 1 + x, x^2$.
 - (f) En P_3 : $2x, x^3 - 3, 1 + x - 4x^3, x^3 + 18x - 9$.
 - (g) En $C(\mathbb{R})$: $\cos 2t, \sin^2 t, \cos^2 t$.
15. ¿Para qué valores de c son linealmente independientes los vectores $\begin{pmatrix} 1-c \\ -c \end{pmatrix}$ y $\begin{pmatrix} c \\ 1+c \end{pmatrix}$?
16. Determine si el siguiente conjunto $\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \right\}$ es linealmente independiente y describa su generado.

17. ¿Para qué valores de α son linealmente independientes los vectores $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}$, $\begin{pmatrix} 3 \\ \alpha \\ 4 \end{pmatrix}$?
18. Determine condiciones sobre a, b, c para que los vectores $\begin{pmatrix} 1 \\ a \\ a^2 \end{pmatrix}$, $\begin{pmatrix} 1 \\ b \\ b^2 \end{pmatrix}$, $\begin{pmatrix} 1 \\ c \\ c^2 \end{pmatrix}$ sean linealmente independientes (ver Sección 4.1, Ejercicio 2.).
19. ¿Para qué valores de α son linealmente dependientes los vectores $\begin{pmatrix} 2 \\ -5 \\ 3 \end{pmatrix}$, $\begin{pmatrix} -4 \\ 10 \\ -6 \end{pmatrix}$, $\begin{pmatrix} 1 \\ \alpha \\ 0 \end{pmatrix}$?
20. Falso o verdadero:
- (a) Cinco vectores de $\mathbb{R}^4$ pueden ser linealmente independientes.
 - (b) Dos vectores de $\mathbb{R}^3$ pueden generar todo el espacio $\mathbb{R}^3$.
 - (c) Un espacio vectorial puede tener infinitos conjuntos generadores.
 - (d) Sea V un espacio vectorial y $W = \text{span}\{v_1, \dots, v_n\}$. Entonces $W = \text{span}\{v_1, \dots, v_n, w\}$ para todo $w \in W$.
 - (e) En $\mathbb{R}^n$, si $\vec{x}_1, \dots, \vec{x}_k$ son linealmente independientes y A es una matriz invertible entonces los vectores $A\vec{x}_1, \dots, A\vec{x}_k$ son linealmente independientes.

5.4. Base y dimensión

En esta sección trabajamos principalmente con espacios vectoriales reales, para fijar ideas. Sin embargo, todos los enunciados también son verdaderos para espacios vectoriales complejos. Solo tenemos que reemplazar $\mathbb{R}$ por $\mathbb{C}$ y la palabra *real* por *complejo* en todas partes.

Definición 5.41. Sea V un espacio vectorial. Una *base* de V es un conjunto de vectores $\{v_1, \dots, v_n\}$ en V que es linealmente independiente y genera V .

La siguiente observación muestra que una base es un *sistema minimal de generadores de V* y, al mismo tiempo, un *sistema maximal de vectores linealmente independientes*.

Observación. Sea $\{v_1, \dots, v_n\}$ una base de V .

- (I) Sea $w \in V$. Entonces $\{v_1, \dots, v_n, w\}$ no es una base de V , porque este sistema de vectores ya no es linealmente independiente por la Proposición 5.35 (IV).
- (II) Si quitamos uno de los vectores de $\{v_1, \dots, v_n\}$, entonces ya no es una base de V , porque el nuevo sistema de vectores ya no genera V . Por ejemplo, si quitamos v_1 , entonces $v_1 \notin \text{span}\{v_2, \dots, v_n\}$ (de lo contrario, $v_1, \dots, v_n$ serían linealmente dependientes), y por lo tanto $\text{span}\{v_2, \dots, v_n\} \neq V$.

Observación 5.42. Por definición, el conjunto vacío es una base del espacio vectorial trivial $\{\mathbb{O}\}$.

Observación 5.43. Toda base de $\mathbb{R}^n$ tiene exactamente n elementos. Para ver esto, observe que por el Corolario 5.37, una base puede tener a lo sumo n elementos, porque de lo contrario no puede ser linealmente independiente. Por otro lado, si tuviera menos de n elementos, entonces, por la Observación 5.30, no podría generar $\mathbb{R}^n$.

Ejemplos 5.44. ■ Una base de $\mathbb{R}^3$ es, por ejemplo, $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$. Los vectores de esta base son los vectores unitarios estándar. La base se llama la *base estándar* (o *base canónica*) de $\mathbb{R}^3$.

Otros ejemplos de bases de $\mathbb{R}^3$ son

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \right\}$$

Ejercicio. Verifique que los sistemas anteriores son bases de $\mathbb{R}^3$.

Los siguientes sistemas **no** son bases de $\mathbb{R}^3$:

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 9 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Ejercicio. Verifique que los sistemas anteriores no son bases de $\mathbb{R}^3$.

- La *base estándar* en $\mathbb{R}^n$ (o *base canónica* en $\mathbb{R}^n$) es $\{\vec{e}_1, \dots, \vec{e}_n\}$. Recordemos que los $\vec{e}_j$ son los vectores unitarios estándar cuya j -ésima entrada es 1 y todas las demás entradas son 0.

Ejercicio. Verifique que forman una base de $\mathbb{R}^n$.

- La *base estándar* en P_n (o *base canónica* en P_n) es $\{1, X, X^2, \dots, X^n\}$.

Ejercicio. Verifique que forman una base de P_n .

- Sean $p_1 = X$, $p_2 = 2X^2 + 5X - 1$, $p_3 = 3X^2 + X + 2$. Entonces el sistema $\{p_1, p_2, p_3\}$ es una base de P_2 .

Demostración. Tenemos que mostrar que el sistema es linealmente independiente y que genera el espacio P_2 . Sea $q = aX^2 + bX + c \in P_2$. Queremos ver si existen $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ tales que $q = \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3$. Si desarrollamos esta ecuación, encontramos

$$\begin{aligned} aX^2 + bX + c &= \alpha_1 X + \alpha_2 (2X^2 + 5X - 1) + \alpha_3 (3X^2 + X + 2) \\ &= (2\alpha_2 + 3\alpha_3)X^2 + (\alpha_1 + 5\alpha_2 + \alpha_3)X - \alpha_2 + 2\alpha_3. \end{aligned}$$

Comparando coeficientes, obtenemos el siguiente sistema de ecuaciones lineales para los α_j :

$$\left. \begin{array}{l} 2\alpha_2 + 3\alpha_3 = a \\ \alpha_1 + 5\alpha_2 + \alpha_3 = b \\ -\alpha_2 + 2\alpha_3 = c \end{array} \right\} \quad \text{en forma matricial:} \quad \begin{pmatrix} 0 & 2 & 3 \\ 1 & 5 & 1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Ahora aplicamos Gauß-Jordan a la matriz aumentada:

$$\left(\begin{array}{ccc|c} 0 & 2 & 3 & a \\ 1 & 5 & 1 & b \\ 0 & -1 & 2 & c \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 1 & 5 & 1 & b \\ 0 & -1 & 2 & c \\ 0 & 2 & 3 & a \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 11 & b+5c \\ 0 & 1 & -2 & c \\ 0 & 0 & 7 & a+2c \end{array} \right).$$

Así vemos que hay exactamente una solución para cualquier q dado. La existencia de tal solución muestra que $\{p_1, p_2, p_3\}$ genera P_2 . También vemos que para cualquier $q \in P_2$ dado hay exactamente una manera de escribirlo como combinación lineal de p_1, p_2, p_3 . Si tomamos el caso especial $q = 0$, esto muestra que el sistema es linealmente independiente. En resumen, $\{p_1, p_2, p_3\}$ es una base de P_2 . $\square$

- Sean $p_1 = X + 1$, $p_2 = X^2 + X$, $p_3 = X^3 + X^2$, $p_4 = X^3 + X^2 + X + 1$. Entonces el sistema $\{p_1, p_2, p_3, p_4\}$ **no** es una base de P_2 .

Ejercicio. ¡Muestre esto!

- En los espacios $M(m \times n)$, el conjunto de todas las matrices A_{ij} forma una base, donde A_{ij} es la matriz con $a_{ij} = 1$ y todas las demás entradas iguales a 0. Por ejemplo, en $M(2 \times 3)$ tenemos la siguiente base:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

- Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $D = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Entonces $\{A, B, C, D\}$ es una base de $M(2 \times 2)$.

Demostración. Sea $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una matriz arbitraria 2×2 . Consideremos la ecuación $M = \alpha_1 A + \alpha_2 B + \alpha_3 C + \alpha_4 D$. Esto conduce a

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} + \alpha_4 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\ = \begin{pmatrix} \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 & \alpha_4 \\ \alpha_2 + \alpha_3 + \alpha_4 & \alpha_3 + \alpha_4 \end{pmatrix}.$$

Así obtenemos el siguiente conjunto de ecuaciones para los α_j :

$$\left. \begin{array}{l} \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = a \\ \alpha_4 = b \\ \alpha_2 + \alpha_3 + \alpha_4 = c \\ \alpha_3 + \alpha_4 = d \end{array} \right\} \quad \text{en forma matricial:} \quad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

Ahora aplicamos Gauß-Jordan a la matriz aumentada:

$$\begin{aligned} \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & a \\ 0 & 0 & 0 & 1 & b \\ 0 & 1 & 1 & 1 & c \\ 0 & 0 & 1 & 1 & d \end{array} \right) &\longrightarrow \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & a \\ 0 & 1 & 1 & 1 & c \\ 0 & 0 & 1 & 1 & d \\ 0 & 0 & 0 & 1 & b \end{array} \right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & a-b \\ 0 & 1 & 1 & 0 & c-b \\ 0 & 0 & 1 & 0 & d-b \\ 0 & 0 & 0 & 1 & b \end{array} \right) \\ &\longrightarrow \left(\begin{array}{cccc|c} 1 & 1 & 0 & 0 & a-d \\ 0 & 1 & 0 & 0 & c-d \\ 0 & 0 & 1 & 0 & d-b \\ 0 & 0 & 0 & 1 & b \end{array} \right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & a-c \\ 0 & 1 & 0 & 0 & c-d \\ 0 & 0 & 1 & 0 & d-b \\ 0 & 0 & 0 & 1 & b \end{array} \right). \end{aligned}$$

Vemos que hay exactamente una solución para cualquier $M \in M(2 \times 2)$ dado. La existencia de la solución muestra que las matrices A, B, C, D generan $M(2 \times 2)$ y la unicidad muestra que son linealmente independientes si escogemos $M = 0$. $\square$

El siguiente teorema es muy importante. Dice que si V tiene una base que consiste en n vectores, entonces **toda** base consiste exactamente en n vectores.

Teorema 5.45. Sea V un espacio vectorial y sean $\{v_1, \dots, v_n\}$ and $\{w_1, \dots, w_n\}$ bases de V . Entonces $n = m$.

Demostración. Supongamos que $m > n$. Mostraremos que entonces los vectores $w_1, \dots, w_m$ no pueden ser linealmente independientes; por lo tanto, no pueden ser una base de V . Como los vectores $v_1, \dots, v_n$ son una base de V , todo w_j puede escribirse como combinación lineal de ellos. Por lo tanto, existen números a_{ij} tales que

$$\begin{aligned} w_1 &= a_{11}v_1 + a_{12}v_2 + \dots + a_{1n}v_n, \\ w_2 &= a_{21}v_1 + a_{22}v_2 + \dots + a_{2n}v_n, \\ &\vdots \\ w_m &= a_{m1}v_1 + a_{m2}v_2 + \dots + a_{mn}v_n. \end{aligned} \tag{5.5}$$

Ahora consideramos la ecuación

$$c_1w_1 + \dots + c_mw_m = \mathbb{0}. \tag{5.6}$$

Si los $w_1, \dots, w_m$ fueran linealmente independientes, entonces debería seguirse que todos los c_j son 0. Sustituimos (5.5) en (5.6) y obtenemos

$$\begin{aligned} \mathbb{0} &= c_1(a_{11}v_1 + a_{12}v_2 + \dots + a_{1n}v_n) + c_2(a_{21}v_1 + a_{22}v_2 + \dots + a_{2n}v_n) \\ &\quad + \dots + c_m(a_{m1}v_1 + a_{m2}v_2 + \dots + a_{mn}v_n) \\ &= (c_1a_{11} + c_2a_{21} + \dots + c_ma_{m1})v_1 + \dots + (c_1a_{1n} + c_2a_{2n} + \dots + c_ma_{mn})v_n. \end{aligned}$$

Como los vectores $v_1, \dots, v_n$ son linealmente independientes, las expresiones entre paréntesis deben ser iguales a cero. Así encontramos

$$\begin{aligned} c_1a_{11} + c_2a_{21} + \dots + c_ma_{m1} &= 0, \\ c_1a_{12} + c_2a_{22} + \dots + c_ma_{m2} &= 0, \\ &\vdots \\ c_1a_{1n} + c_2a_{2n} + \dots + c_ma_{mn} &= 0. \end{aligned} \tag{5.7}$$

Este es un sistema homogéneo de n ecuaciones para las m incógnitas $c_1, \dots, c_m$. Como $n < m$, debe tener infinitas soluciones. Así que el sistema $\{w_1, \dots, w_m\}$ no es linealmente independiente y, por lo tanto, no puede ser una base de V . Por lo tanto, $m > n$ no puede ser cierto y se sigue que $n \geq m$.

Si suponemos que $n > m$, entonces el mismo argumento anterior, con los papeles de los v_j y los w_j intercambiados, conduce a una contradicción y se sigue que $n \leq m$.

En resumen, mostramos que tanto $n \geq m$ como $n \leq m$ deben ser ciertos. Por lo tanto $m = n$. $\square$

Definición 5.46. ■ Sea V un espacio vectorial finitamente generado. Entonces tiene una base por el Teorema 5.47 de abajo y, por el Teorema 5.45, el número n de vectores necesarios para una base no depende de la base particular escogida. Este número se llama la *dimensión de V* . Se denota por $\dim V$.

- Si un espacio vectorial V no es finitamente generado, entonces ponemos $\dim V = \infty$.
- El conjunto vacío es una base del espacio vectorial trivial $\{\mathbb{O}\}$; por lo tanto, $\dim\{\mathbb{O}\} = 0$.

A continuación mostramos que todo espacio vectorial finitamente generado tiene una base y, por lo tanto, una dimensión bien definida.

Teorema 5.47. *Sea V un espacio vectorial y supongamos que existen vectores $w_1, \dots, w_m \in V$ tales que $V = \text{span}\{w_1, \dots, w_m\}$. Entonces el conjunto $\{w_1, \dots, w_m\}$ contiene una base de V . En particular, V tiene una base finita y $\dim V \leq m$.*

Demostración. Sin pérdida de generalidad podemos suponer que todos los vectores w_j son distintos de $\mathbb{O}$. Comenzamos con el primer vector. Si $V = \text{span}\{w_1\}$, entonces $\{w_1\}$ es una base de V y $\dim V = 1$. De lo contrario, ponemos $V_1 := \text{span}\{w_1\}$ y observamos que $V_1 \neq V$. Ahora verificamos si $w_2 \in \text{span}\{w_1\}$. Si lo está, lo descartamos porque en este caso $\text{span}\{w_1\} = \text{span}\{w_1, w_2\}$, así que no necesitamos w_2 para generar V . Luego verificamos si $w_3 \in \text{span}\{w_1\}$. Si lo está, lo descartamos, etc. Procedemos así hasta encontrar un vector w_{i_2} en nuestra lista que no pertenezca a $\text{span}\{w_1\}$. Tal i_2 debe existir, porque de lo contrario ya tendríamos que $V_1 = V$. Entonces ponemos $V_2 := \text{span}\{w_1, w_{i_2}\}$. Si $V_2 = V$, entonces hemos terminado. De lo contrario, procedemos como antes: Verificamos si $w_{i_2+1} \in V_2$. Si este es el caso, entonces podemos descartarlo porque $\text{span}\{w_1, w_{i_2}\} = \text{span}\{w_1, w_{i_2}, w_{i_2+1}\}$. Luego verificamos w_{i_2+2} , etc., hasta encontrar un w_{i_3} tal que $w_{i_3} \notin \text{span}\{w_1, w_{i_2}\}$ y ponemos $V_3 := \text{span}\{w_1, w_{i_2}, w_{i_3}\}$. Si $V_3 = V$, entonces hemos terminado. Si no, repetimos el proceso. Observe que, después de a lo sumo m repeticiones, esto termina. Esto muestra que podemos extraer del sistema de generadores una base $\{w_1, w_{i_2}, \dots, w_{i_k}\}$ de V . $\square$

El siguiente teorema complementa el anterior.

Teorema 5.48. *Sea V un espacio vectorial finitamente generado. Entonces todo sistema $w_1, \dots, w_m \in V$ de vectores linealmente independientes puede ser completado a una base $\{w_1, \dots, w_m, v_{m+1}, \dots, v_n\}$ de V .*

Demostración. Observe que $\dim V < \infty$ por el Teorema 5.47 y ponga $n = \dim V$. Se sigue que $n \geq m$ porque tenemos m vectores linealmente independientes en V . Si $m = n$, entonces $w_1, \dots, w_m$ ya es una base de V y hemos terminado.

Si $m < n$, entonces $\text{span}\{w_1, \dots, w_m\} \neq V$ y escogemos un vector arbitrario $v_{m+1} \notin \text{span}\{w_1, \dots, w_m\}$ and we define $V_{m+1} := \text{span}\{w_1, \dots, w_m, v_{m+1}\}$. Entonces $\dim V_{m+1} = m + 1$. Si $m + 1 = n$, entonces necesariamente $V_{m+1} = V$ y hemos terminado. Si $m + 1 < n$, entonces escogemos un vector arbitrario $v_{m+2} \in V \setminus V_{m+1}$ y ponemos $V_{m+2} := \text{span}\{w_1, \dots, w_m, v_{m+1}, v_{m+2}\}$. Si $m + 2 = n$, entonces necesariamente $V_{m+2} = V$ y hemos terminado. Si no, repetimos el paso anterior. Observe que después de $n - m$ pasos hemos encontrado una base $\{w_1, \dots, w_m, v_{m+1}, \dots, v_n\}$ of V . $\square$

En resumen, los dos teoremas anteriores dicen lo siguiente:

- Si el conjunto de vectores $v_1, \dots, v_m$ genera el espacio vectorial V , entonces siempre es posible extraer un subconjunto que sea una base de V (necesitamos eliminar $m - n$ vectores).
- Si tenemos un conjunto de vectores linealmente independientes $v_1, \dots, v_m$ en un espacio vectorial finitamente generado V , entonces es posible encontrar vectores $v_{m+1}, \dots, v_n$ tales que $\{v_1, \dots, v_n\}$ sea una base de V (necesitamos agregar $\dim V - m$ vectores).

Corolario 5.49. Sea V un espacio vectorial.

- Si los vectores $v_1, \dots, v_k \in V$ son linealmente independientes, entonces $k \leq \dim V$.
- Si los vectores $v_1, \dots, v_m \in V$ generan V , entonces $m \geq \dim V$.

Ejemplo 5.50. ■ Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in M(2 \times 2)$ y supongamos que queremos completaras a una base de $M(2 \times 2)$ (es claro que A y B son linealmente independientes, así que esto tiene sentido). Como $\dim(M(2 \times 2)) = 4$, sabemos que necesitamos 2 matrices más. Tomamos cualquier matriz $C \notin \text{span}\{A, B\}$, por ejemplo $C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Finally we need a matrix $D \notin \text{span}\{A, B, C\}$. Podemos tomar, por ejemplo, $D = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Entonces A, B, C, D es una base de $M(2 \times 2)$.

Verifique que $D \notin \text{span}\{A, B, C\}$

Encuentre otras matrices C' y D' tales que $\{A, B, C', D'\}$ sea una base de $M(2 \times 2)$.

- Dados los vectores $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 4 \\ 0 \\ 4 \end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\vec{v}_4 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$, $\vec{v}_5 = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$, $\vec{v}_6 = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}$ y queremos encontrar un subconjunto de ellos que forme una base de $\mathbb{R}^3$.

Observe que, de entrada, no es claro que esto sea posible, porque sin cálculos adicionales no sabemos si los vectores dados realmente generan $\mathbb{R}^3$. Si no lo hacen, entonces por supuesto es imposible extraer de ellos una base.

Comencemos. Primero observe que necesitamos 3 vectores para una base, ya que $\dim \mathbb{R}^3 = 3$. Así que comenzamos con el primer vector no nulo, que es $\vec{v}_1$. Vemos que $\vec{v}_2 = 4\vec{v}_1$, así que lo descartamos. Conservamos $\vec{v}_3$ ya que $\vec{v}_3 \notin \text{span}\{\vec{v}_1\}$. Luego, $\vec{v}_4 = \vec{v}_3 - \vec{v}_1$, así que $\vec{v}_4 \in \text{span}\{\vec{v}_1, \vec{v}_3\}$ y lo descartamos. Un pequeño cálculo muestra que $\vec{v}_5 \notin \text{span}\{\vec{v}_1, \vec{v}_3\}$. Por lo tanto, $\{\vec{v}_1, \vec{v}_3, \vec{v}_5\}$ es una base de $\mathbb{R}^3$.

Observación 5.51. Presentaremos una manera más sistemática de resolver ejercicios de este tipo en el Teorema 6.34 y en la Observación 6.35.

Teorema 5.52. Sea V un espacio vectorial con base $\{v_1, \dots, v_n\}$. Entonces todo $x \in V$ puede escribirse de manera única como combinación lineal de los vectores $v_1, \dots, v_n$.

Demostración. Tenemos que mostrar la existencia y unicidad de números $c_1, \dots, c_n$ tales que $w = c_1v_1 + \dots + c_nv_n$.

La *existencia* es clara, ya que el conjunto $\{v_1, \dots, v_n\}$ es un conjunto de generadores de V (¡incluso es una base!).

La *unicidad* se puede mostrar como sigue. Supongamos que hay números $c_1, \dots, c_n$ y $d_1, \dots, d_n$ tales que $w = c_1v_1 + \dots + c_nv_n$ and $w = d_1v_1 + \dots + d_nv_n$. Entonces se sigue que

$$0 = w - w = c_1v_1 + \dots + c_nv_n - (d_1v_1 + \dots + d_nv_n) = (c_1 - d_1)v_1 + \dots + (c_n - d_n)v_n.$$

Entonces todos los coeficientes $c_1 - d_1, \dots, c_n - d_n$ deben ser cero porque los vectores $v_1, \dots, v_n$ son linealmente independientes. Por lo tanto se sigue que $c_1 = d_1, \dots, c_n = d_n$, lo cual muestra la unicidad. Observe que el teorema también es verdadero si $V = \{0\}$, porque por definición la suma vacía es igual a cero. $\square$

Si tenemos un espacio vectorial V y un subespacio $W \subset V$, entonces podemos preguntarnos cuál es la relación entre sus dimensiones, porque W mismo es un espacio vectorial.

Lema 5.53. Sea V un espacio vectorial finitamente generado y sea W un subespacio. Entonces W es finitamente generado y $\dim W \leq \dim V$.

Demostración. Sea V un espacio vectorial finitamente generado con $\dim V = n < \infty$. Sea W un subespacio de V y supongamos que W no es finitamente generado. Entonces podemos construir un sistema arbitrariamente grande de vectores linealmente independientes en W como sigue. Clearly, W cannot be the trivial space, so we can choose $w_1 \in W \setminus \{0\}$ y ponemos $W_1 = \text{span}\{w_1\}$. Entonces W_1 es un subespacio finitamente generado de W ; por lo tanto $W_1 \subsetneq W$ y podemos escoger $w_2 \in W \setminus W_1$. Claramente, el conjunto $\{w_1, w_2\}$ es linealmente independiente. Let us set $W_2 = \text{span}\{w_1, w_2\}$. Como W_2 es un subespacio finitamente generado de W , se sigue que $W_2 \subsetneq W$ y podemos escoger $w_3 \in W \setminus W_2$. Then the vectors w_1, w_2, w_3 are linealmente independientes y ponemos $W_3 = \text{span}\{w_1, w_2, w_3\}$. Continuando con este procedimiento podemos construir subespacios $W_1 \subsetneq W_2 \subsetneq \dots \subsetneq W$ con $\dim W_k = k$ para todo k . En particular, podemos encontrar un sistema de $n+1$ vectores linealmente independientes en $W \subseteq V$, lo cual contradice el hecho de que cualquier sistema de más de $n = \dim V$ vectores en V debe ser linealmente dependiente, véase el Corolario 5.49. Esto también muestra que cualquier sistema de más de n vectores en W debe ser linealmente dependiente. Como una base de W consiste en vectores linealmente independientes, se sigue que $\dim W \leq n = \dim V$. $\square$

Teorema 5.54. Sea V un espacio vectorial finitamente generado y sea $W \subseteq V$ un subespacio. Entonces se cumple lo siguiente:

- (I) $\dim W \leq \dim V$.
- (II) $\dim W = \dim V$ if and only if $W = V$.

Demostración. (I) se sigue inmediatamente del Lema 5.53.

- (II) Si $V = W$, entonces claramente $\dim V = \dim W$. Para mostrar el recíproco, suponemos que $\dim V = \dim W$ y tenemos que mostrar que $V = W$. Como antes, sea $\{w_1, \dots, w_k\}$ una base de W . Entonces estos vectores son linealmente independientes en W y, por lo tanto, también en V . Como $\dim W = \dim V$, sabemos que estos vectores forman una base de V . Por lo tanto $V = \text{span}\{w_1, \dots, w_m\} = W$. $\square$

Observación 5.55. Observe que (I) es verdadero incluso cuando V no es finitamente generado, porque $\dim W \leq \infty = \dim V$ cualquiera que sea $\dim W$. Sin embargo, (II) **no** es verdadero en general para espacios vectoriales de dimensión infinita. In Example 5.56 (f) and (g) mostraremos que $\dim P = \dim C(\mathbb{R})$ a pesar de que $P \neq C(\mathbb{R})$. (Recordemos que P es el conjunto de todos los polinomios y que $C(\mathbb{R})$ es el conjunto de todas las funciones continuas. Así, tenemos $P \subsetneq C(\mathbb{R})$.)

Ahora damos algunos ejemplos de dimensiones de espacios.

Ejemplos 5.56. (a) $\dim \mathbb{R}^n = n$, $\dim \mathbb{C}^n = n$.

- (b) $\dim M(m \times n) = mn$. Esto se sigue porque el conjunto de todas las matrices $m \times n$ A_{ij} que tienen un 1 en la i -ésima fila y la j -ésima columna y todas las demás entradas iguales a cero forma una base de $M(m \times n)$, y hay exactamente mn matrices de este tipo.
- (c) Sea $M_{\text{sym}}(n \times n)$ el conjunto de todas las matrices simétricas $n \times n$. Entonces $\dim M_{\text{sym}}(n \times n) = \frac{n(n+1)}{2}$. Para ver esto, sea A_{ij} la matriz $n \times n$ con $a_{ij} = a_{ji} = 1$ y todas las demás entradas iguales a 0. Observe que $A_{ij} = A_{ji}$. No es difícil ver que el conjunto de todos los A_{ij} con $i \leq j$ forma una base de $M_{\text{sym}}(n \times n)$. La dimensión de $M_{\text{sym}}(n \times n)$ es el número de matrices distintas de este tipo. ¿Cuántas hay? Si fijamos $j = 1$, entonces solo $i = 1$ es posible. Si fijamos $j = 2$, entonces $i = 1, 2$ es posible, etc., hasta que para $j = n$ los valores permitidos para i son $1, 2, \dots, n$. En total tenemos $1 + 2 + \dots + n = \frac{n(n+1)}{2}$ posibilidades. Por ejemplo, en el caso $n = 2$, las matrices son

$$A_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

En el caso $n = 3$, las matrices son

$$A_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$A_{22} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_{33} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Convénzase de que los A_{ij} forman una base de $M_{\text{sym}}(n \times n)$.

- (d) Sea $M_{\text{asym}}(n \times n)$ el conjunto de todas las matrices antisimétricas $n \times n$. Entonces $\dim M_{\text{asym}}(n \times n) = \frac{n(n-1)}{2}$. Para ver esto, para $i \neq j$ sea A_{ij} la matriz $n \times n$ con $a_{ij} = -a_{ji} = 1$ y todas las

demás entradas iguales a 0. No es difícil ver que el conjunto de todos los A_{ij} con $i < j$ forma una base de $M_{\text{asym}}(n \times n)$. ¿Cuántas de estas matrices hay? Si fijamos $j = 2$, entonces solo $i = 1$ es posible. Si fijamos $j = 3$, entonces $i = 1, 2$ es posible, etc., hasta que para $j = n$ los valores permitidos para i son $1, 2, \dots, n-1$. En total tenemos $1 + 2 + \dots + (n-1) = \frac{n(n-1)}{2}$ posibilidades. Por ejemplo, en el caso $n = 2$, la única matriz es

$$A_{12} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

En el caso $n = 3$, las matrices son

$$A_{12} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad A_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Convénzase de que los A_{ij} forman una base de $M_{\text{asym}}(n \times n)$.

Observación. Observe que $\dim M_{\text{sym}}(n \times n) + \dim M_{\text{asym}}(n \times n) = n^2 = \dim M(n \times n)$. Esto no es una coincidencia. Observe que toda matriz $n \times n$ M puede escribirse como

$$M = \frac{1}{2}(M + M^t) + \frac{1}{2}(M - M^t)$$

y que $\frac{1}{2}(M + M^t) \in M_{\text{sym}}(n \times n)$ y $\frac{1}{2}(M - M^t) \in M_{\text{asym}}(n \times n)$. Además, es fácil verificar que $M_{\text{sym}}(n \times n) \cap M_{\text{asym}}(n \times n) = \{0\}$. Por lo tanto $M(n \times n)$ es la *suma directa* de $M_{\text{sym}}(n \times n)$ y $M_{\text{asym}}(n \times n)$. (Para la definición de la *suma directa* de subespacios, véase la Definición 5.59).

- (e) $\dim P_n = n + 1$ ya que $\{1, X, \dots, X^n\}$ es una base de P_n y consiste en $n + 1$ vectores.
- (f) $\dim P = \infty$. Recordemos que P es el espacio de todos los polinomios.

Demostración. Sabemos que para todo $n \in \mathbb{N}$, el espacio P_n es un subespacio de P . Por lo tanto, para todo $n \in \mathbb{N}$, debemos tener que $n + 1 = \dim P_n \leq \dim P$. Esto solo es posible si $\dim P = \infty$. $\square$

- (g) $\dim C(\mathbb{R}) = \infty$. Recordemos que $C(\mathbb{R})$ es el espacio de todas las funciones continuas.

Demostración. Como P es un subespacio de $C(\mathbb{R})$, se sigue que $\dim P \leq \dim(C(\mathbb{R}))$, luego $\dim(C(\mathbb{R})) = \infty$. $\square$

Ahora usamos el concepto de dimensión para clasificar todos los subespacios de $\mathbb{R}^2$ y $\mathbb{R}^3$. Ya sabemos que, por ejemplo, las rectas y los planos que pasan por el origen son subespacios de $\mathbb{R}^3$. Ahora podemos mostrar que no hay otros subespacios propios.

Subespacios de $\mathbb{R}^2$. Sea U un subespacio de $\mathbb{R}^2$. Entonces U debe tener una dimensión. Así tenemos los siguientes casos:

- $\dim U = 0$. En este caso $U = \{\vec{0}\}$ es el subespacio trivial.

- $\dim U = 1$. Entonces U es de la forma $U = \text{span}\{\vec{v}_1\}$ con algún vector $\vec{v}_1 \in \mathbb{R}^2 \setminus \{\vec{0}\}$. Por lo tanto U es una recta paralela a $\vec{v}_1$ que pasa por el origen.
- $\dim U = 2$. En este caso $\dim U = \dim \mathbb{R}^2$. Por lo tanto se sigue que $U = \mathbb{R}^2$ por el Teorema 5.54 (II).
- $\dim U \geq 3$ no es posible porque $0 \leq \dim U \leq \dim \mathbb{R}^2 = 2$.

En conclusión, los únicos subespacios de $\mathbb{R}^2$ son $\{\vec{0}\}$, las rectas que pasan por el origen y el propio $\mathbb{R}^2$.

Subespacios de $\mathbb{R}^3$. Sea U un subespacio de $\mathbb{R}^3$. Entonces U debe tener una dimensión. Así tenemos los siguientes casos:

- $\dim U = 0$. En este caso $U = \{\vec{0}\}$ es el subespacio trivial.
- $\dim U = 1$. Entonces U es de la forma $U = \text{span}\{\vec{v}_1\}$ con algún vector $\vec{v}_1 \in \mathbb{R}^3 \setminus \{\vec{0}\}$. Por lo tanto U es una recta paralela a $\vec{v}_1$ que pasa por el origen.
- $\dim U = 2$. Entonces U es de la forma $U = \text{span}\{\vec{v}_1, \vec{v}_2\}$ con vectores linealmente independientes $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^3$. Por lo tanto U es un plano paralelo a los vectores $\vec{v}_1$ y $\vec{v}_2$ que pasa por el origen.
- $\dim U = 3$. En este caso $\dim U = \dim \mathbb{R}^3$. Por lo tanto se sigue que $U = \mathbb{R}^3$ por el Teorema 5.54 (II).
- $\dim U \geq 4$ no es posible porque $0 \leq \dim U \leq \dim \mathbb{R}^3 = 3$.

En conclusión, los únicos subespacios de $\mathbb{R}^3$ son $\{\vec{0}\}$, las rectas que pasan por el origen, los planos que pasan por el origen y el propio $\mathbb{R}^3$.

Concluimos esta sección con la definición formal de rectas y planos.

Definición 5.57. Sea V un espacio vectorial con $\dim V = n$ y sea $W \subseteq V$ un subespacio. Entonces W se llama una

- *recta* si $\dim W = 1$,
- *plano* si $\dim W = 2$,
- *hiperplano* si $\dim W = n - 1$.

Observe que en $\mathbb{R}^3$ los hiperplanos son exactamente los planos.

Debería haber entendido

- el concepto de base de un espacio vectorial de dimensión finita,
- que un espacio vectorial dado tiene infinitas bases, pero el número de vectores en cualquier base del espacio es el mismo,
- por qué y cómo el concepto de dimensión ayuda a clasificar todos los subespacios de un

espacio vectorial dado,

- por qué una matriz $A \in M(n \times n)$ es invertible si y solo si sus columnas son una base de $\mathbb{R}^n$,
- etc.

Ahora debería ser capaz de

- verificar si un sistema de vectores es una base de un espacio vectorial dado,
- encontrar una base para un espacio vectorial dado,
- extender un sistema de vectores linealmente independientes a una base,
- encontrar la dimensión de un espacio vectorial dado,
- etc.

Ejercicios.

1. Encuentre bases para los espacios dados en los ejercicios del 1 al 10 de la sección 5.3.
2. Determine la dimensión de los siguientes espacios:

(a) En $\mathbb{R}^4$, los vectores $\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$ tales que $w = x + y$.

(b) Todos los vectores de la forma $\begin{pmatrix} a + c \\ a - b \\ b + c \\ -a + b \end{pmatrix}$.

(c) $\left\{ A \in M(2 \times 2) : A \begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\}$.

(d) Las soluciones del sistema homogéneo

$$\begin{aligned} x - 5y &= 0, \\ 2x - 3y &= 0. \end{aligned}$$

(e) Las soluciones del sistema homogéneo

$$\begin{aligned} x - 3y - z &= 0, \\ -2x + 2y - 3z &= 0, \\ 4x - 8y + 5z &= 0. \end{aligned}$$

(f) Las soluciones del sistema homogéneo

$$\begin{aligned} -x + 3y - 2z &= 0, \\ 2x - 6y + 4z &= 0, \\ -3x + 9y - 6z &= 0. \end{aligned}$$

- (g) $V = \text{span}\{\cos 2t, \sin^2 t, \cos^2 t\}$.
3. En los siguientes ejercicios, determine si el conjunto de vectores dado es una base para el espacio vectorial indicado.
- (a) En $M(2 \times 2)$; $\begin{pmatrix} 3 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -5 & 1 \\ 0 & 6 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & -7 \end{pmatrix}$.
- (b) En $W = \{(x, y) \in \mathbb{R}^2 : 3x - y = 0\}$; $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$.
- (c) En $\mathbb{R}_4$; $\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \\ 1 \end{pmatrix}$.
- (d) En P_2 ; $5 - x^2, 3x$.
- (e) En P_3 ; $x^3 - x, x^3 + x^2, x^2 + 1, x - 1$.
4. En $\mathbb{R}^4$, encontrar una base para el subespacio $U = \{\vec{x} \in \mathbb{R}^4 : \langle \vec{x}, (2, -3, 0, 4) \rangle = 0\}$. (Observe que U es un hiperplano en $\mathbb{R}^4$.)
5. En $\mathbb{R}^3$, considere la recta $L : \frac{x}{5} = \frac{y}{3} = 2z$. Encuentre una base para L y complétela a una base de $\mathbb{R}^3$.
6. Encuentre una base para el plano $E : 2x + y - 5z = 0$ y complétela a una base de $\mathbb{R}^3$. Representéla gráficamente (Existe una forma natural de hacerlo. ¿Cuál?).
7. Encuentre una base para $\mathbb{R}^4$ que contenga a los vectores $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$.
8. Muestre que los vectores $\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} -b \\ a \end{pmatrix}$ forman una base de $\mathbb{R}^2$ si $ab \neq 0$. Muestre también que $\begin{pmatrix} a \\ b \end{pmatrix} \perp \begin{pmatrix} -b \\ a \end{pmatrix}$.
9. Demuestre que $\{1 - x^2, 1 + x^2\}$ es una base del subconjunto de P_2 cuya primera derivada evaluada en cero vale cero. Complete esta base a una base de todo P_2 .
10. ¿Para qué valores de α los vectores $\begin{pmatrix} \alpha^5 \\ 1 + \alpha \\ 0 \end{pmatrix}, \begin{pmatrix} \alpha \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 - \alpha \\ 0 \\ 1 \end{pmatrix}$ generan todo $\mathbb{R}^3$?
11. (a) En $M(n \times n)$, muestre que n^2 matrices tales que en todas su entrada a_{nn} vale cero no pueden ser linealmente independientes.
- (b) En P_n , muestre que $n + 1$ polinomios cuya primera derivada evaluada en cero se anula no pueden ser linealmente independientes.
- (c) En P_n , ¿existen $n + 1$ polinomios linealmente independientes tales que el coeficiente de x^0 es 1?

5.5. Intersecciones y sumas de espacios vectoriales

En esta sección construiremos nuevos subespacios a partir de subespacios dados. Veremos que la intersección de dos espacios vectoriales es nuevamente un espacio vectorial, mientras que la unión en general no lo es.

Proposición 5.58. Sean U, W subespacios de un espacio vectorial V . Entonces su intersección $U \cap W$ es un subespacio de V .

Demostración. Claramente, $U \cap W \neq \emptyset$ porque $\mathbf{0} \in U$ y $\mathbf{0} \in W$, de donde $\mathbf{0} \in U \cap W$. Ahora sean $z_1, z_2 \in U \cap W$ y $c \in \mathbb{K}$. Entonces $z_1, z_2 \in U$ y por lo tanto $z_1 + cz_2 \in U$, porque U es un espacio vectorial. Análogamente se sigue que $z_1 + cz_2 \in W$, de donde $z_1 + cz_2 \in U \cap W$. $\square$

Observe que $U \cap W$ es el subespacio más gree que está contenido tanto en U como en W .

Por ejemplo, la intersección de dos planos en $\mathbb{R}^3$ que pasan por el origen es o bien ese mismo plano (si los dos planos originales son el mismo plano), o bien una recta que pasa por el origen. En cualquier caso, es un subespacio de $\mathbb{R}^3$.

Observe, sin embargo, que en general la unión de dos espacios vectoriales no es un espacio vectorial. Por ejemplo, en $\mathbb{R}^2$ las rectas $L : y = 0$ (este es el eje x) y $G : x = 0$ (este es el eje y) son subespacios y su unión $L \cup G$ consiste exactamente en ambos ejes. Esto claramente no es un espacio vectorial, porque no es cerrado bajo sumas. Por ejemplo, $\vec{e}_1 \in L \subseteq L \cup G$ y $\vec{e}_2 \in G \subseteq L \cup G$, pero $\vec{e}_1 + \vec{e}_2 \notin L \cup G$. Para convertirlo en un espacio vectorial, necesitamos incluir todas las combinaciones lineales que faltan. El espacio que obtenemos de esta manera se llama una *suma directa*; véase la Definición 5.59.

Ejercicio. ■ Dé más ejemplos de dos subespacios cuya unión no sea un espacio vectorial.
■ Dé un ejemplo de dos subespacios cuya unión sea un espacio vectorial.

Question 5.1. Unión de subespacios.

¿Puede encontrar un criterio que deban satisfacer los subespacios para que su unión sea un subespacio?

Definamos la *suma* y la *suma directa* de espacios vectoriales.

Definición 5.59. Sean U, W subespacios de un espacio vectorial V . Entonces la *suma* de los espacios vectoriales U y W se define como

$$U + W = \{u + w : u \in U, w \in W\}. \quad (5.8)$$

Si además $U \cap W = \{\mathbf{0}\}$, entonces la suma se llama la *suma directa de U y W* , y se escribe $U \oplus W$ en lugar de $U + W$.

Observación. Sean U, W subespacios de un espacio vectorial V . Entonces $U + W$ es nuevamente un subespacio de V .

Demostración. Claramente, $U + W \neq \emptyset$ porque $\mathbf{0} \in U$ y $\mathbf{0} \in W$, de donde $\mathbf{0} + \mathbf{0} = \mathbf{0} \in U + W$. Ahora sean $z_1, z_2 \in U + W$ y $c \in \mathbb{K}$. Entonces existen $u_1, u_2 \in U$ y $w_1, w_2 \in W$ tales que $z_1 = u_1 + w_1$ y $z_2 = u_2 + w_2$. Por lo tanto

$$z_1 + cz_2 = u_1 + w_1 + c(u_2 + w_2) = (u_1 + cu_2) + (w_1 + cw_2) \in U + W$$

y $U + W$ es un subespacio por la Proposición 5.10. $\square$

Observe que $U + W$ consiste en todas las combinaciones lineales posibles de vectores de U y de W . Obtenemos inmediatamente las siguientes observaciones.

Observación 5.60. (i) Suponga que $U = \text{span}\{u_1, \dots, u_k\}$ y que $W = \text{span}\{w_1, \dots, w_j\}$, entonces $U + W = \text{span}\{u_1, \dots, u_k, w_1, \dots, w_j\}$.

(ii) El espacio $U + W$ es el espacio vectorial más pequeño que contiene tanto a U como a W .

Ejemplos 5.61. (i) Sea V un espacio vectorial y sea $U \subseteq V$ un subespacio. Entonces siempre tenemos:

- (a) $U + \{\mathbf{0}\} = U \oplus \{\mathbf{0}\} = U$,
- (b) $U + U = U$,
- (c) $U + V = V$.

Si U y W son subespacios de V , entonces

- (a) $U \subseteq U + W$ y $W \subseteq U + W$.
- (b) $U + W = U$ if y only if $W \subseteq U$.

(ii) Sean U y W rectas en $\mathbb{R}^2$ que pasan por el origen. Entonces son subespacios de $\mathbb{R}^2$ y tenemos que $U + W = U$ si las rectas son paralelas y $U + W = \mathbb{R}^2$ si no son paralelas.

(iii) Sean U y W rectas en $\mathbb{R}^3$ que pasan por el origen. Entonces son subespacios de $\mathbb{R}^3$ y tenemos que $U + W = U$ si las rectas son paralelas; de lo contrario, $U + W$ es el plano que contiene ambas rectas.

(iv) Sea U una recta y sea W un plano en $\mathbb{R}^3$, ambos pasando por el origen. Entonces son subespacios de $\mathbb{R}^3$ y tenemos que $U + W = W$ si la recta U está contenida en W . Si no, entonces $U + W = \mathbb{R}^3$.

Demuestre las afirmaciones de los ejemplos anteriores.

Recordemos que la intersección de dos subespacios es nuevamente un subespacio; véase la Proposición 5.58. La fórmula para la dimensión de la suma de dos espacios vectoriales en la siguiente proposición se puede entender como sigue: Si sumamos las dimensiones de los dos espacios vectoriales, entonces contamos dos veces la parte que es común a ambos espacios; por lo tanto, tenemos que restar su dimensión para obtener la dimensión correcta de la suma de los espacios vectoriales.

Proposición 5.62. Sean U, W subespacios de un espacio vectorial V . Entonces

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

En particular, $\dim(U + W) = \dim U + \dim W$ si $U \cap W = \{\mathbb{O}\}$.

Demostración. Sean $\dim U = k$ y $\dim W = m$. Recordemos que $U \cap W$ es un subespacio de V y que $U \cap W \subseteq U$ y $U \cap W \subseteq W$. Sea $v_1, \dots, v_\ell$ una base de $U \cap W$. Por el Teorema 5.48, podemos completarla a una base $v_1, \dots, v_\ell, u_{\ell+1}, \dots, u_k$ of U . De manera similar, podemos completarla a una base $v_1, \dots, v_\ell, w_{\ell+1}, \dots, w_m$ of W . Ahora afirmamos que $v_1, \dots, v_\ell, u_{\ell+1}, \dots, u_k, w_{\ell+1}, \dots, w_m$ es una base de $U + W$.

- Primero mostramos que los vectores $v_1, \dots, v_\ell, u_{\ell+1}, \dots, u_k, w_{\ell+1}, \dots, w_m$ generan $U + W$. Esto se sigue de la Observación 5.60 y de

$$\begin{aligned} U + W &= \text{span}\{v_1, \dots, v_\ell, u_{\ell+1}, \dots, u_k\} + \text{span}\{v_1, \dots, v_\ell, w_{\ell+1}, \dots, w_m\} \\ &= \text{span}\{v_1, \dots, v_\ell, u_{\ell+1}, \dots, u_k, v_1, \dots, v_\ell, w_{\ell+1}, \dots, w_m\} \\ &= \text{span}\{v_1, \dots, v_\ell, u_{\ell+1}, \dots, u_k, w_{\ell+1}, \dots, w_m\}. \end{aligned}$$

- Ahora mostramos que los vectores $v_1, \dots, v_\ell, u_{\ell+1}, \dots, u_k, w_{\ell+1}, \dots, w_m$ son linealmente independientes. Sean $\alpha_1, \dots, \alpha_n, \beta_{\ell+1}, \dots, \beta_m \in \mathbb{R}$ tales que

$$\alpha_1 v_1 + \dots + \alpha_\ell v_\ell + \alpha_{\ell+1} u_{\ell+1} + \dots + \alpha_k u_k + \beta_{\ell+1} w_{\ell+1} + \dots + \beta_m w_m = \mathbb{O}.$$

Se sigue que

$$\underbrace{\alpha_1 v_1 + \dots + \alpha_\ell v_\ell + \alpha_{\ell+1} u_{\ell+1} + \dots + \alpha_k u_k}_{\in U} = - \underbrace{(\beta_{\ell+1} w_{\ell+1} + \dots + \beta_m w_m)}_{\in W} \quad (5.9)$$

y por lo tanto $-(\beta_{\ell+1} w_{\ell+1} + \dots + \beta_m w_m) \in U \cap W$; de aquí que debe ser una combinación lineal de los vectores $v_1, \dots, v_\ell$, porque estos son una base de $U \cap W$. Así podemos encontrar $\gamma_1, \dots, \gamma_\ell \in \mathbb{R}$ tales que $\gamma_1 v_1 + \dots + \gamma_\ell v_\ell = -(\beta_{\ell+1} w_{\ell+1} + \dots + \beta_m w_m)$. Esto implica que

$$\gamma_1 v_1 + \dots + \gamma_\ell v_\ell + \beta_{\ell+1} w_{\ell+1} + \dots + \beta_m w_m = \mathbb{O}.$$

Como los vectores $v_1, \dots, v_\ell, w_{\ell+1}, \dots, w_m$ forman una base de W , son linealmente independientes, y concluimos que $\gamma_1 = \dots = \gamma_\ell = \beta_{\ell+1} = \dots = \beta_m = 0$. Sustituyendo en (5.9), obtenemos

$$\alpha_1 v_1 + \dots + \alpha_\ell v_\ell + \alpha_{\ell+1} u_{\ell+1} + \dots + \alpha_k u_k = \mathbb{O},$$

de donde $\alpha_1 = \dots = \alpha_k = 0$.

Se sigue que

$$\begin{aligned} \dim(U + W) &= \#\{v_1, \dots, v_\ell, u_{\ell+1}, \dots, u_k, w_{\ell+1}, \dots, w_m\} \\ &= \ell + (k - \ell) + (m - \ell) \\ &= k + m - \ell \\ &= \dim U + \dim W - \dim(U \cap W) \end{aligned}$$

porque los vectores $v_1, \dots, v_\ell, u_{\ell+1}, \dots, u_k$ son una base de U . □

Ejemplos 5.63. En $\mathbb{R}^3$ considere los subespacios E, F, G dados por $E : 2x - y + 3z = 0$, $F = \text{span}\{\vec{v}, \vec{w}\}$ y $G = \text{span}\{\vec{a}\}$ donde

$$\vec{v} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{a} = \begin{pmatrix} -4 \\ 3 \\ 2 \end{pmatrix}.$$

Encuentre $E \cap F$, $E + F$, $E \cap G$, $E + G$ y $F \cap G$, $F + G$, y sus dimensiones.

Solución. Claramente, E y F son planos en $\mathbb{R}^3$ y G es una recta.

$E \cap F$ Observe que los vectores normales $\vec{n}_E$ de E y $\vec{n}_F$ de F son

$$\vec{n}_E = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}, \quad \vec{n}_F = \vec{v} \times \vec{w} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}.$$

Solución 1. La forma normal de F es $F : x + 2y - z = 0$. Un punto (x, y, z) pertenece a $E \cap F$ si y solo si sus coordenadas satisfacen simultáneamente las ecuaciones de E y F . Por lo tanto obtenemos el siguiente sistema de ecuaciones:

$$\begin{aligned} 2x - y + 3z &= 0, \\ x + 2y - z &= 0. \end{aligned}$$

Un cálculo corto (Gauß-Jordan) muestra que el conjunto de soluciones es la recta $H : x = -y = -z$, o, en forma vectorial,

$$H = \text{span}\{\vec{b}\} \quad \text{donde} \quad \vec{b} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}. \quad (*)$$

Solución 2. También podemos usar las formas vectoriales de E y F . Para escribir E en forma vectorial, solo necesitamos escoger dos vectores $\vec{r}, \vec{s}$ que sean paralelos a E , por ejemplo

$$E = \text{span}\{\vec{r}, \vec{s}\} \quad \text{donde} \quad \vec{r} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \vec{s} = \begin{pmatrix} 1 \\ 5 \\ 1 \end{pmatrix}.$$

Un vector $\vec{x} = (x, y, z)^t$ pertenece a $E \cap F$ si y solo si $\vec{x}$ es una combinación lineal de $\vec{v}, \vec{w}$ y una combinación lineal de $\vec{r}, \vec{s}$, es decir, si y solo si existen $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ tales que $\alpha\vec{v} + \beta\vec{w} = \gamma\vec{r} + \delta\vec{s}$, o

$$\alpha\vec{v} + \beta\vec{w} - \gamma\vec{r} - \delta\vec{s} = \vec{0}.$$

Escribiendo esto como un sistema para las incógnitas $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, obtenemos

$$\begin{aligned} \beta - \gamma - \delta &= 0, \\ \alpha - 2\gamma - 5\delta &= 0, \\ 2\alpha + \beta - \delta &= 0. \end{aligned}$$

Un cálculo directo muestra que la solución general es $\alpha = t, \beta = -t, \gamma = -2t, \delta = t$. Por lo tanto

$$E \cap F = \{t\vec{v} - t\vec{w} : t \in \mathbb{R}\} = \{t(\vec{v} - \vec{w}) : t \in \mathbb{R}\} = \text{span}\{\vec{v} - \vec{w}\} = \text{span}\{\vec{b}\} = H$$

o equivalentemente

$$E \cap F = \{-2t\vec{r} + t\vec{s} : t \in \mathbb{R}\} = \{t(-2\vec{r} - \vec{s}) : t \in \mathbb{R}\} = \text{span}\{-2\vec{r} - \vec{s}\} = \text{span}\{\vec{b}\} = H$$

con $\vec{b}$ y H como en (*).

[E + F] *Solución 1.* Sabemos que $E + F = \text{span}\{\vec{v}, \vec{w}, \vec{r}, \vec{s}\}$. Ahora, de manera similar al Ejemplo 5.50, vemos que los vectores $\vec{v}, \vec{w}, \vec{r}$ son linealmente independientes; por lo tanto, la dimensión de $E + F$ es mayor o igual que 3. Como es un subespacio de $\mathbb{R}^3$, debe ser igual a $\mathbb{R}^3$. (También podríamos usar el Teorema 6.34 y la Observación 6.35 para encontrar un sistema de generadores de $E + F$.)

Solución 2. Sabemos que

$$\dim(E + F) = \dim E + \dim F - \dim(E \cap F) = 2 + 2 - 1 = 3.$$

Como $E + F \subseteq \mathbb{R}^3$, se sigue que $E + F = \mathbb{R}^3$.

[E ∩ G] Sustituimos la forma paramétrica de G en la forma normal de E y obtenemos como condición para la intersección que

$$0 = 2(-4t) - 3t + 3 \cdot (-t) = -14t.$$

La única solución es $t = 0$, lo cual implica que $E \cap G = \{\vec{0}\}$.

[E + G] Es fácil ver que los tres vectores $\vec{r}, \vec{s}, \vec{a}$ son linealmente independientes; por lo tanto generan $\mathbb{R}^3$ y de ahí $E + G = \text{span}\{\vec{r}, \vec{s}, \vec{a}\} = \mathbb{R}^3$.

Alternativamente, podríamos usar que $\dim(E + G) = \dim E + \dim G - \dim(E \cap G) = 2 + 1 - 0 = 3$ para concluir que $E + G = \mathbb{R}^3$.

[F ∩ G] Es fácil ver que los tres vectores $\vec{v}, \vec{w}, \vec{a}$ son linealmente dependientes. De hecho, $\vec{a} = 2\vec{v} - 4\vec{w}$. Por lo tanto $G \subseteq F$ y, en consecuencia, $G \cap F = G$.

[F + G] De lo anterior se sigue que $F + G = \text{span}\{\vec{v}, \vec{w}, \vec{a}\} = \text{span}\{\vec{v}, \vec{w}\} = F$.

De lo anterior, es claro que $\dim(E \cap F) = 1$, $\dim(E + F) = 3$, $\dim(E \cap G) = 0$, $\dim(E + G) = 3$, $\dim(F \cap G) = 2$, $\dim(F + G) = 3$.

◇

Ejemplo 5.64. En $\mathbb{R}^5$ considere los subespacios $U = \{(x, y, z, r, s) : 2x + y + 5z + 4s = 0\}$ y $W = \{(x, y, z, r, s) : x + y + 4z + 3s = 0, 3x + y + 6z + 4r + 6s = 0\}$. Encuentre $U \cap W$, $U + W$ y sus dimensiones.

Solución. **[U ∩ W]** Es fácil ver que $\dim U = 4$ y $\dim W = 3$.

Solución 1. Un punto (x, y, z, r, s) pertenece a $U \cap W$ si y solo si satisface el siguiente sistema de ecuaciones lineales

$$\begin{aligned} 2x + y + 5z + 4s &= 0, \\ x + y + 4z + 3s &= 0, \\ 3x + y + 6z + 4r + 6s &= 0. \end{aligned}$$

Esto se puede resolver con eliminación de Gauß-Jordan

$$\begin{pmatrix} 2 & 1 & 5 & 0 & 4 \\ 1 & 1 & 4 & 0 & 3 \\ 3 & 1 & 6 & 4 & 6 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & -1 & -3 & 0 & -2 \\ 1 & 1 & 4 & 0 & 3 \\ 0 & -2 & -6 & 4 & -3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & -1 & -3 & 0 & -2 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 4 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 4 & 1 \end{pmatrix}$$

Por lo tanto, $\dim(U \cap W) = 2$ y

$$U \cap W = \left\{ (x, y, z, r, s) : \begin{array}{l} x + z + s = 0 \\ y + 3z + 2s = 0 \\ 4r + s = 0 \end{array} \right\} = \text{span} \left\{ \begin{pmatrix} -4 \\ -8 \\ 0 \\ -1 \\ 4 \end{pmatrix}, \begin{pmatrix} -1 \\ -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}. \quad (5.10)$$

Solución 2. Podemos usar formas vectoriales de U y W . Escogemos cualquier conjunto de vectores linealmente independientes $\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4$ en U y $\vec{w}_1, \vec{w}_2, \vec{w}_3$ en W . Entonces $U = \text{span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4\}$ y $W = \text{span}\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$. Por ejemplo, podemos tomar

$$\vec{u}_1 = \begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \vec{u}_2 = \begin{pmatrix} 0 \\ -5 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \vec{u}_3 = \begin{pmatrix} 0 \\ -4 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \vec{u}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \vec{w}_1 = \begin{pmatrix} -3 \\ -3 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \vec{w}_2 = \begin{pmatrix} -2 \\ 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \vec{w}_3 = \begin{pmatrix} -1 \\ -3 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Entonces $\vec{x} \in U \cap W$ si es una combinación lineal tanto de los $\vec{u}_j$ como de los $\vec{w}_j$, es decir,

$$\begin{aligned} \vec{x} &= \alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \alpha_3 \vec{u}_3 + \alpha_4 \vec{u}_4 \\ &= \beta_1 \vec{w}_1 + \beta_2 \vec{w}_2 + \beta_3 \vec{w}_3 \end{aligned}$$

para algunos $\alpha_j, \beta_j \in \mathbb{R}$. Si tomamos la diferencia de los lados derechos, obtenemos

$$\vec{0} = \alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \alpha_3 \vec{u}_3 + \alpha_4 \vec{u}_4 - \beta_1 \vec{w}_1 - \beta_2 \vec{w}_2 - \beta_3 \vec{w}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & -3 & -2 & -1 \\ -2 & -5 & -4 & 0 & -3 & 2 & -3 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}.$$

Resolvemos este sistema usando eliminación de Gauß-Jordan:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -3 & -2 & -1 \\ -2 & -5 & -4 & 0 & -3 & 2 & -3 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 & 0 \end{pmatrix} \rightarrow \dots \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & -3 & -2 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -2 & 0 \end{pmatrix}.$$

La solución general es

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = t \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + s \begin{pmatrix} -4 \\ 0 \\ 4 \\ -1 \\ -2 \\ 1 \\ 0 \end{pmatrix}$$

y por lo tanto

$$\begin{aligned} U \cap W &= \{(-t - 4s)\vec{u}_1 + t\vec{u}_2 + 4s\vec{u}_3 - s\vec{u}_4 : t, s \in \mathbb{R}\} \\ &= \{t(-\vec{u}_1 + \vec{u}_2) + s(-4\vec{u}_1 + 4\vec{u}_3 - \vec{u}_4) : t, s \in \mathbb{R}\} \\ &= \text{span}\{-\vec{u}_1 + \vec{u}_2, -4\vec{u}_1 + 4\vec{u}_3 - \vec{u}_4\} \end{aligned}$$

o equivalentemente

$$\begin{aligned} U \cap W &= \{-2s\vec{w}_1 + s\vec{w}_2 + t\vec{w}_3 : t, s \in \mathbb{R}\} = \{s(-2\vec{w}_1 + \vec{w}_2) + t\vec{w}_3 : t, s \in \mathbb{R}\} \\ &= \text{span}\{-2\vec{w}_1 + \vec{w}_2, \vec{w}_3\}. \end{aligned}$$

Este es, por supuesto, el mismo resultado que en (5.10).

$\boxed{U + W}$ Sabemos que $\dim(U + W) = \dim U + \dim W - \dim(U \cap W) = 4 + 3 - 2 = 5 = \dim \mathbb{R}^5$, por lo tanto $U + W = \mathbb{R}^5$. $\diamond$

Ahora debería haber entendido

- el concepto de suma y suma directa de dos subespacios,
- por qué la fórmula $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$ tiene sentido,
- etc.

Ahora debería ser capaz de

- encontrar la intersección de dos espacios vectoriales y su dimensión,
- encontrar la suma de dos espacios vectoriales y su dimensión,
- decidir si la suma de dos espacios vectoriales es una suma directa,
- etc.

Ejercicios.

1. En $\mathbb{R}^4$, sean $U = \text{span} \left\{ \begin{pmatrix} 1 \\ 5 \\ 2 \\ -5 \end{pmatrix}, \begin{pmatrix} -3 \\ 4 \\ 1 \\ -3 \end{pmatrix} \right\}$ y $V = \{\vec{x} \in \mathbb{R}^4 : \langle \vec{x}, (1, 2, -2, -1)^t \rangle = 0, \langle \vec{x}, (2, 5, -5, 1)^t \rangle = 0\}$. Determine $U \cap V$, $U + V$ y $\dim(U + V)$.
2. En $\mathbb{R}^3$ muestre que si E, F son planos no paralelos que pasan por el origen, entonces $E + F = \mathbb{R}^3$.
3. Sea V el subespacio de las matrices triangulares superiores y W el subespacio de las matrices triangulares inferiores. Muestre que $M(n \times n) = V + W$. ¿Es directa esta suma?
4. Muestre que $M(3 \times 3) = M_{\text{sym}}(3 \times 3) \oplus M_{\text{asym}}(3 \times 3)$. (*Pista:* Basta demostrar que $M_{\text{sym}}(3 \times 3) \cap M_{\text{asym}}(3 \times 3) = \{\mathbf{0}\}$ ¿Por qué es suficiente mostrar esto?).

5. Sean U, V subespacios de $\mathbb{R}^n$, responda las siguientes preguntas.
 - (a) Si $\dim U + \dim V = n$ ¿se sigue que $U + V = \mathbb{R}^n$?
 - (b) Si $\mathbb{R}^n = U + V$ y $n = \dim U + \dim V$ ¿se sigue que $U \cap V = \{\vec{0}\}$?
6. Sea $V \subseteq \mathbb{R}^n$ un subespacio de dimensión k . Demuestre que existe V' subespacio de $\mathbb{R}^n$ tal que $\mathbb{R}^n = V \oplus V'$ (Hint: Escoja una base de V y complétela a una base de $\mathbb{R}^n$, ¿cómo se debe tomar V' ?).
7. Sean $U, V \subseteq \mathbb{R}^n$ subespacios. Muestre que $\vec{0}$ se escribe de forma única como la suma de un elemento de U con un elemento de V si y solo si $U \cap V = \{\vec{0}\}$. (En tal caso $U + V = U \oplus V$.)
8. Suponga que $U, V, W \subseteq \mathbb{R}^5$ con $\dim U = 2$, $\dim V = 3$ y $\dim W = 4$.
 - (a) ¿Cuáles son las posibilidades para $\dim U \cap V$ y $\dim U + V$? Dé ejemplos para cada caso.
 - (b) ¿Cuáles son las posibilidades para $\dim U \cap W$ y $\dim U + W$? Dé ejemplos para cada caso.
 - (c) ¿Cuáles son las posibilidades para $\dim V \cap W$ y $\dim V + W$? Dé ejemplos para cada caso.

5.6. Resumen

Sea V un espacio vectorial sobre $\mathbb{K}$ y sean $v_1, \dots, v_k \in V$.

Combinaciones lineales e independencia lineal

- Un vector w se llama una *combinación lineal* de los vectores $v_1, \dots, v_k$ si existen escalares $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ tales que

$$w = \alpha_1 v_1 + \dots + \alpha_k v_k.$$

- El conjunto de todas las combinaciones lineales de los vectores $v_1, \dots, v_k$ es un subespacio de V , llamado el *espacio generado por los vectores* $v_1, \dots, v_k$ o el *subespacio generado por los vectores* $v_1, \dots, v_k$. Notación:

$$\begin{aligned} \text{span}\{v_1, \dots, v_k\} &:= \text{span}\{v_1, \dots, v_k\} := \{w \in V : w \text{ es combinación lineal de } v_1, \dots, v_k\} \\ &= \{\alpha_1 v_1 + \dots + \alpha_k v_k : \alpha_1, \dots, \alpha_k \in \mathbb{K}\}. \end{aligned}$$

- Los vectores $v_1, \dots, v_k$ se llaman *linealmente independientes* si la ecuación

$$\alpha_1 v_1 + \dots + \alpha_k v_k = \vec{0}$$

tiene solamente la solución trivial $\alpha_1 = \dots = \alpha_k = 0$.

Base y dimensión

- Un sistema $v_1, \dots, v_m$ de vectores en V se llama una *base de V* si es linealmente independiente y $\text{span}\{v_1, \dots, v_m\} = V$.
- Un espacio vectorial V se llama *finitamente generado* si tiene una base finita. En este caso, toda base de V tiene el mismo número de vectores. El número de vectores necesarios para una base de un espacio vectorial V se llama la *dimensión de V* .

- Si V no es finitamente generado, ponemos $\dim V = \infty$.
- Para $v_1, \dots, v_k \in V$, se sigue que $\dim(\text{span}\{v_1, \dots, v_k\}) \leq k$ con igualdad si y solo si los vectores $v_1, \dots, v_k$ son linealmente independientes.
- Si V es finitamente generado, entonces todo sistema linealmente independiente de vectores $v_1, \dots, v_k \in V$ se puede extender a una base de V .
- Si $V = \text{span}\{v_1, \dots, v_k\}$, entonces V tiene una base que consiste en un subsistema de los vectores dados $v_1, \dots, v_k$.
- Si U es un subespacio de V , entonces $\dim U \leq \dim V$.
- Si V es finitamente generado y U es un subespacio de V , entonces $\dim U = \dim V$ si y solo si $U = V$. Esta afirmación es falsa si $\dim V = \infty$.
- $\dim\{\mathbb{O}\} = 0$ y $\{\mathbb{O}\}$ tiene la base única $\emptyset$.

Ejemplos de dimensiones de algunos espacios vectoriales:

- $\dim\{\mathbb{O}\} = 0$,
- $\dim \mathbb{R}^n = n$, $\dim \mathbb{C}^n = n$,
- $\dim M(m \times n) = mn$,
- $\dim M_{\text{sym}}(n \times n) = \frac{n(n+1)}{2}$,
- $\dim M_{\text{asym}}(n \times n) = \frac{n(n-1)}{2}$,
- $\dim P_n = n + 1$,
- $\dim P = \infty$,
- $\dim C(\mathbb{R}) = \infty$.

Independencia lineal, propiedad generadora y bases en $\mathbb{R}^n$ y $\mathbb{C}^n$

Sean $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{R}^n$ o $\mathbb{C}^n$ y sea $A = (\vec{v}_1 | \dots | \vec{v}_k) \in M(n \times k)$ la matriz cuyas columnas son los vectores dados.

- $\text{span}\{\vec{v}_1, \dots, \vec{v}_k\} = \mathbb{R}^n$ si y solo si el sistema $A\vec{x} = \vec{b}$ tiene al menos una solución para todo $\vec{b} \in \mathbb{R}^n$.
- Los vectores $\vec{v}_1, \dots, \vec{v}_k$ son linealmente independientes si y solo si el sistema $A\vec{x} = \vec{0}$ tiene solamente la solución trivial $\vec{x} = \vec{0}$.
- Los vectores $\vec{v}_1, \dots, \vec{v}_k$ son una base de $\mathbb{R}^n$ si y solo si $k = n$ y A es invertible.

Intersecciones y sumas de subespacios

Sea V un espacio vectorial y sean U, W subespacios de V . Entonces

$$U \cap W := \{v \in V : v \in U \text{ y } v \in W\},$$

$$U \cup W := \{v \in V : v \in U \text{ o } v \in W\},$$

$$U + W := \{u + w : u \in U, w \in W\}.$$

Observe que $U \cap W \subseteq U \subseteq U \cup W \subseteq U + W$ and $W \cap U \subseteq W \subseteq U \cup W \subseteq U + W$.

- $U \cap W$ y $U + W$ son subespacios de V
- $U \cup W$ en general **no** es un subespacio.
- La suma de U y W se llama una *suma directa* y se denota por $U \oplus W$ si $U \cap W = \{0\}$.
- $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$.

5.7. Ejercicios

1. Sea X el conjunto de todas las funciones de $\mathbb{R}$ a $\mathbb{R}$. Demuestre que X con la suma y producto con números en $\mathbb{R}$ es un espacio vectorial.

De los siguientes subconjuntos de X , diga si son subespacios de X .

- (a) Todas las funciones acotadas de $\mathbb{R}$ a $\mathbb{R}$.
- (b) Todas las funciones constantes.
- (c) Todas las funciones continuas.
- (d) Todas las funciones continuas con $f(3) = 0$.
- (e) Todas las funciones continuas con $f(3) = 4$.
- (f) Todas las funciones con $f(3) > 0$.
- (g) Todas las funciones pares.
- (h) Todas las funciones impares.
- (i) Todos los polinomios.
- (j) Todas las funciones no negativas.
- (k) Todos los polinomios de grado ≥ 4 .

2. Sean $A \in M(m \times n)$ y sea $\vec{a} \in \mathbb{R}^k$.

- (a) Demuestre que $U = \{A\vec{x} : \vec{x} \in \mathbb{R}^n\}$ es un subespacio de $\mathbb{R}^m$.
- (b) ¿Los conjuntos $R = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = (1, 1, \dots, 1)^t\}$ y $S = \{\vec{x} \in \mathbb{R}^n : A\vec{x} \neq 0\}$ son subespacios de $\mathbb{R}^n$?

3. Sean $A \in M(m \times n)$ y sea $\vec{a} \in \mathbb{R}^k$.

- (a) ¿El conjunto $T = \{\vec{x} \in \mathbb{R}^k : \langle \vec{x}, \vec{a} \rangle = 0\}$ es un subespacio de $\mathbb{R}^k$?
 (b) ¿Los conjuntos

$$S_1 = \{\vec{x} \in \mathbb{R}^k : \|\vec{x}\| = 1\}, \quad B_1 = \{\vec{x} \in \mathbb{R}^k : \|\vec{x}\| \leq 1\}, \quad F = \{\vec{x} \in \mathbb{R}^k : \|\vec{x}\| \geq 1\}$$

son subespacios de $\mathbb{R}^k$?

4. Considere el conjunto $\mathbb{R}^2$ con las siguientes operaciones:

$$\oplus : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_2 \\ x_2 + y_1 \end{pmatrix},$$

$$\odot : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \lambda \odot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}.$$

¿Es $\mathbb{R}^2$ con esta suma y producto con escalares un espacio vectorial?

5. Considere el conjunto $\mathbb{R}^2$ con las siguientes operaciones:

$$\boxplus : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \boxplus \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ 0 \end{pmatrix},$$

$$\boxdot : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \lambda \boxdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}.$$

¿Es $\mathbb{R}^2$ con esta suma y producto con escalares un espacio vectorial?

6. (a) Sea $V = (-\frac{\pi}{2}, \frac{\pi}{2})$ y defina suma $\oplus : V \times V \rightarrow V$ y producto con escalar $\odot : \mathbb{R} \times V \rightarrow V$ por

$$x \oplus y = \arctan(\tan(x) + \tan(y)), \quad \lambda \odot x = \arctan(\lambda \tan(x))$$

para todo $x, y \in V$, $\lambda \in \mathbb{R}$. Demuestre que $(V, \oplus, \odot)$ es un espacio vectorial sobre $\mathbb{R}$.

- (b) Una generalización de la construcción en (a) es lo siguiente:

Sea V un conjunto y $f : \mathbb{R}^n \rightarrow V$ una función biyectiva. Entonces V es un espacio vectorial con suma y producto con escalar definido así:

$$x \oplus y = f(f^{-1}(x) + f^{-1}(y)), \quad \lambda \odot x = f(\lambda f^{-1}(x))$$

para todo $x, y \in V$, $\lambda \in \mathbb{R}$.

7. Sea U un subespacio de $\mathbb{R}^n$. Demuestre que $\mathbb{R}^n \setminus U$ no es un subespacio de $\mathbb{R}^n$.
8. Sean $m, n \in \mathbb{N}$. Demuestre que $M(m \times n, \mathbb{R})$ con la suma y producto con números en $\mathbb{R}$ es un espacio vectorial.
- De los siguientes subconjuntos de $M(n \times n)$, diga si son subespacios.

- (a) Todas las matrices con $a_{11} = 0$.
- (b) Todas las matrices con $a_{11} = 3$.
- (c) Todas las matrices con $a_{12} = \mu a_{11}$ para un $\mu \in \mathbb{R}$ fijo.
- (d) Todas las matrices cuya primera columna coincide con la última columna.

Para los siguientes numerales supongamos que $n = m$.

- (e) Todas las matrices simétricas (es decir, todas las matrices A con $A^t = A$).
- (f) Todas las matrices que no son simétricas.
- (g) Todas las matrices antisimétricas (es decir, todas las matrices A con $A^t = -A$).
- (h) Todas las matrices diagonales.
- (i) Todas las matrices triangulares superiores.
- (j) Todas las matrices triangulares inferiores.
- (k) Todas las matrices invertibles.
- (l) Todas las matrices no invertibles.
- (m) Todas las matrices con $\det A = 1$.

9. Demuestre que

$$V = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} : \begin{array}{l} x_1 + x_2 - 2x_3 - x_4 = 0, \\ x_1 - x_2 + x_3 + 7x_4 = 0 \end{array} \right\}$$

es un subespacio de $\mathbb{R}^4$.

10. Demuestre que

$$W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} : \begin{array}{l} 3x_1 - x_2 - 2x_3 - x_4 = 3, \\ 4x_1 + x_2 + x_3 + 7x_4 = 5 \end{array} \right\}$$

es un subespacio afín de $\mathbb{R}^4$.

11. Considere los sistemas de ecuaciones lineales

$$(1) \quad \begin{cases} x + 2y + 3z = 0, \\ 4x + 5y + 6z = 0, \\ 7x + 8y + 9z = 0 \end{cases}, \quad (2) \quad \begin{cases} x + 2y + 3z = 3, \\ 4x + 5y + 6z = 9, \\ 7x + 8y + 9z = 15 \end{cases}.$$

Sea U el conjunto de todas las soluciones de (1) y W el conjunto de todas las soluciones de (2). Note que se pueden ver como subconjuntos de $\mathbb{R}^3$.

- (a) Demuestre que U es un subespacio de $\mathbb{R}^3$ y descríballo geométricamente.

- (b) Demuestre que W no es un subespacio de $\mathbb{R}^3$.
 (c) Demuestre que W es un subespacio afín de $\mathbb{R}^3$ y descríballo geoméricamente.

12. (a) Sean $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} -2 \\ 5 \end{pmatrix} \in \mathbb{R}^2$. Escriba $v = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$ como combinación lineal de v_1 y v_2 .

- (b) ¿Es $v = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}$ combinación lineal de $v_1 = \begin{pmatrix} 1 \\ 7 \\ 5 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix}$?

- (c) ¿Es $A = \begin{pmatrix} 13 & -5 \\ 50 & 8 \end{pmatrix}$ combinación lineal de

$$A_1 = \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 1 \\ -2 & 2 \end{pmatrix}, A_3 = \begin{pmatrix} 2 & 1 \\ 5 & 0 \end{pmatrix}, A_4 = \begin{pmatrix} 1 & -1 \\ 5 & 2 \end{pmatrix}?$$

13. (a) ¿Los vectores $v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $v_2 = \begin{pmatrix} 2 \\ 2 \\ 5 \end{pmatrix}$, $v_3 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$ son linealmente independientes en $\mathbb{R}^3$?

- (b) ¿Los vectores $v_1 = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 7 \\ 2 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix}$ son linealmente independientes en $\mathbb{R}^3$?

- (c) ¿Los vectores $p_1 = X^2 - X + 2$, $p_2 = X + 3$, $p_3 = X^2 - 1$ son linealmente independientes en P_2 ? ¿Son linealmente independientes en P_n para $n \geq 3$?

- (d) ¿Los vectores $A_1 = \begin{pmatrix} 1 & 3 & 1 \\ -2 & 2 & 3 \end{pmatrix}$, $A_2 = \begin{pmatrix} 1 & 7 & 3 \\ 2 & -1 & 2 \end{pmatrix}$, $A_3 = \begin{pmatrix} 1 & -1 & 0 \\ 5 & 2 & 8 \end{pmatrix}$ son linealmente independientes en $M(2 \times 3)$?

14. Sean $\vec{v}_1, \dots, \vec{v}_k, \vec{w} \in \mathbb{R}^n$. Suponga que $\vec{w} \neq \vec{0}$ y que $\vec{w} \in \mathbb{R}^n$ es ortogonal a todos los vectores $\vec{v}_j$. Demuestre que $\vec{w} \notin \text{span}\{\vec{v}_1, \dots, \vec{v}_m\}$. ¿Se sigue que el sistema $\vec{w}, \vec{v}_1, \dots, \vec{v}_m$ es linealmente independiente?

15. Determine si $\text{span}\{a_1, a_2, a_3, a_4\} = \text{span}\{v_1, v_2, v_3\}$ para

$$a_1 = \begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix}, a_2 = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}, a_3 = \begin{pmatrix} 1 \\ 2 \\ 13 \end{pmatrix}, a_4 = \begin{pmatrix} 2 \\ 1 \\ 11 \end{pmatrix}, v_1 = \begin{pmatrix} 5 \\ -3 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \\ 8 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}.$$

16. (a) ¿Las siguientes matrices generan el espacio de todas las matrices simétricas 2×2 ?

$$A_1 = \begin{pmatrix} 2 & 0 \\ 0 & 7 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 13 & 0 \\ 0 & 5 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix},$$

Si no lo hacen, encuentre un $M \in M_{\text{sym}}(2 \times 2) \setminus \text{span}\{A_1, A_2, A_3\}$.

- (b) ¿Las siguientes matrices generan el espacio de todas las matrices simétricas 2×2 ?

$$B_1 = \begin{pmatrix} 2 & 0 \\ 0 & 7 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 13 & 0 \\ 0 & 5 \end{pmatrix}, \quad B_3 = \begin{pmatrix} 0 & 3 \\ -3 & 0 \end{pmatrix},$$

- (c) ¿Las siguientes matrices generan el espacio de las matrices triangulares superiores 2×2 ?

$$C_1 = \begin{pmatrix} 6 & 0 \\ 0 & 7 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 0 & 3 \\ 0 & 5 \end{pmatrix}, \quad C_3 = \begin{pmatrix} 10 & -7 \\ 0 & 0 \end{pmatrix}.$$

Si no, encuentre una matriz M triangular superior que no pertenece a $\text{span}\{C_1, C_2, C_3\}$.

17. Sea $n \in \mathbb{N}$ y sea V el conjunto de las matrices simétricas $n \times n$ con la suma y producto con $\lambda \in \mathbb{R}$ usual.

- (a) Demuestre que V es un espacio vectorial sobre $\mathbb{R}$.
 (b) Encuentre matrices que generan V . ¿Cuál es el número mínimo de matrices que se necesitan para generar V ?

18. Determine si los siguientes conjuntos de vectores son bases del espacio vectorial indicado.

(a) $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} -2 \\ 5 \end{pmatrix}$; $\mathbb{R}^2$.

(b) $A = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 5 & 3 \\ 1 & 2 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 1 \\ -2 & 2 \end{pmatrix}$, $D = \begin{pmatrix} 2 & 1 \\ 5 & 0 \end{pmatrix}$; $M(2 \times 2)$.

(c) $p_1 = 1 + x$, $p_2 = x + x^2$, $p_3 = x^2 + x^3$, $p_4 = 1 + x + x^2 + x^3$; P_3 .

19. (a) Es F el plano dado por $F : 2x - 5y + 3z = 0$. Demuestre que F es subespacio de $\mathbb{R}^3$ y encuentre vectores $\vec{u}$ y $\vec{w} \in \mathbb{R}^3$ tal que $F = \text{span}\{\vec{u}, \vec{w}\}$.

(b) Sean $v_1 = \begin{pmatrix} 1 \\ 7 \\ 3 \end{pmatrix}$, $v_2 = \begin{pmatrix} -5 \\ 1 \\ 2 \end{pmatrix} \in \mathbb{R}^3$. Sea E el plano $E = \text{span}\{v_1, v_2\}$. Escriba E en la forma $E : ax + by + cz = d$.

(c) Encuentre un vector $w \in \mathbb{R}^3$, distinto de v_1 y v_2 , tal que $\text{span}\{v_1, v_2, w\} = E$.

(d) Encuentre un vector $v_3 \in \mathbb{R}^3$ tal que $\text{span}\{v_1, v_2, v_3\} = \mathbb{R}^3$.

20. (a) Encuentre una base para el plano $E : x - 2y + 3z = 0$ en $\mathbb{R}^3$.

(b) Complete la base encontrada en (i) a una base de $\mathbb{R}^3$.

21. Sea $F := \{(x_1, x_2, x_3, x_4)^t : 2x_1 - x_2 + 4x_3 + x_4 = 0\}$.

(a) Demuestre que F es un subespacio de $\mathbb{R}^4$

(b) Encuentre una base para F y calcule $\dim F$.

- (c) Complete la base encontrada en (ii) a una base de $\mathbb{R}^4$.
22. Sea $G := \{(x_1, x_2, x_3, x_4)^t : 2x_1 - x_2 + 4x_3 + x_4 = 0, x_1 - x_2 + x_3 + 2x_4 = 0\}$.
- (a) Demuestre que G es un subespacio de $\mathbb{R}^4$
- (b) Encuentre una base para G y calcule $\dim G$.
- (c) Complete la base encontrada en (ii) a una base de $\mathbb{R}^4$.
23. Sean $v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 4 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 4 \\ 2 \\ 5 \end{pmatrix}$, $v_4 = \begin{pmatrix} 2 \\ 8 \\ 3 \end{pmatrix}$, $v_5 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.
- Determine si estos vectores generan el espacio $\mathbb{R}^3$. Si lo hacen, escoja una base de $\mathbb{R}^3$ de los vectores dados.
24. Sean $C_1 = \begin{pmatrix} 6 & 0 \\ 0 & 7 \end{pmatrix}$, $C_2 = \begin{pmatrix} 6 & 3 \\ 0 & 12 \end{pmatrix}$, $C_3 = \begin{pmatrix} 6 & -3 \\ 0 & 2 \end{pmatrix}$, $C_4 = \begin{pmatrix} 12 & -9 \\ 0 & -1 \end{pmatrix}$.
- Determine si estas matrices generan el espacio de las matrices triangulares superiores 2×2 . Si lo hacen, escoja una base de las matrices dadas.
25. Sean $p_1 = x^2 + 7$, $p_2 = x + 1$, $p_3 = 3x^3 + 7x$. Determine si los polinomios p_1, p_2, p_3 son linealmente independientes. Si lo son, complételos a una base en P_3 .
26. Para los siguientes conjuntos, determine si son espacios vectoriales. Si lo son, calcule su dimensión.
- (a) $M_1 = \{A \in M(n \times n) : A \text{ es triangular superior}\}$.
- (b) $M_2 = \{A \in M(n \times n) : A \text{ tiene ceros en la diagonal}\}$.
- (c) $M_3 = \{A \in M(n \times n) : A^t = -A\}$.
- (d) $M_4 = \{p \in P_5 : p(0) = 0\}$.
27. Para los siguientes sistemas de vectores en el espacio vectorial V , determine la dimensión del espacio vectorial generado por ellos y escoja un subsistema de ellos que es base del espacio vectorial generado por los vectores dados. Complete este subsistema a una base de V .
- (a) $V = \mathbb{R}^3$, $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 3 \\ 2 \\ 7 \end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$.
- (b) $V = P_4$, $p_1 = x^3 + x$, $p_2 = x^3 - x^2 + 3x$, $p_3 = x^2 + 2x - 5$, $p_4 = x^3 + 3x + 2$.
- (c) $V = M(2 \times 2)$, $A = \begin{pmatrix} 1 & 4 \\ -2 & 5 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 0 \\ 1 & 4 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 12 \\ -7 & 11 \end{pmatrix}$, $D = \begin{pmatrix} 9 & -12 \\ 10 & 1 \end{pmatrix}$.

28. Sea V un espacio vectorial. ¿Falso o verdadero?
- (a) Suponga $v_1, \dots, v_k, u, z \in V$ tal que z es combinación lineal de los $v_1, \dots, v_k$. Entonces z es combinación lineal de $v_1, \dots, v_k, u$.
 - (b) Si u es combinación lineal de $v_1, \dots, v_k \in V$, entonces $v_1, \dots, v_k, u$ es un sistema de vectores linealmente dependientes.
 - (c) Si $v_1, \dots, v_k \in V$ es un sistema de vectores linealmente dependientes, entonces v_1 es combinación lineal de los $v_2, \dots, v_k$.
29. (a) ¿Es $\mathbb{C}^n$ un espacio vectorial sobre $\mathbb{R}$?
(b) ¿Es $\mathbb{C}^n$ un espacio vectorial sobre $\mathbb{Q}$?
(c) ¿Es $\mathbb{R}^n$ un espacio vectorial sobre $\mathbb{C}$?
(d) ¿Es $\mathbb{R}^n$ un espacio vectorial sobre $\mathbb{Q}$?
(e) ¿Es $\mathbb{Q}^n$ un espacio vectorial sobre $\mathbb{R}$?
(f) ¿Es $\mathbb{Q}^n$ un espacio vectorial sobre $\mathbb{C}$?
30. Sea V un espacio vectorial y sean $U, W \subseteq V$ subespacios.
- (a) Demuestre que $U \cap W$ es un subespacio.
 - (b) Demuestre que $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$.
 - (c) Suponga que $U \cap W = \{0\}$. Demuestre que $\dim(U \oplus W) = \dim U + \dim W$.

DRAFT

Capítulo 6

Transformaciones lineales y cambio de bases

En la primera sección de este capítulo definiremos aplicaciones lineales entre espacios vectoriales y discutiremos sus propiedades. Estas son funciones que “se comportan bien” con respecto a la estructura de espacio vectorial. Por ejemplo, las matrices $m \times n$ pueden verse como aplicaciones lineales de $\mathbb{R}^m$ a $\mathbb{R}^n$. Demostraremos la llamada *fórmula de la dimensión* para aplicaciones lineales. En la Sección 6.2 estudiaremos el caso especial de las matrices. Uno de los resultados principales será la fórmula de la dimensión (6.4). En la Sección 6.4 veremos que, después de escoger una base, toda aplicación lineal entre espacios vectoriales de dimensión finita puede representarse como una matriz. Esto nos permitirá trasladar resultados sobre matrices al caso de las transformaciones lineales.

Como en capítulos anteriores, trabajamos con espacios vectoriales sobre $\mathbb{R}$ o $\mathbb{C}$. Recordemos que $\mathbb{K}$ siempre denota $\mathbb{R}$ o $\mathbb{C}$.

6.1. Transformaciones lineales

Definición 6.1. Sean U, V espacios vectoriales sobre el mismo campo $\mathbb{K}$. Una función $T : U \rightarrow V$ se llama una *transformación lineal* si para todos $x, y \in U$ y $\lambda \in \mathbb{K}$ se cumple lo siguiente:

$$T(x + y) = Tx + Ty, \quad T(\lambda x) = \lambda Tx. \quad (6.1)$$

Otros términos para *transformación lineal* son *función lineal*, *aplicación lineal* u *operador lineal*.

Observación. Observe que muy a menudo se escribe Tx en lugar de $T(x)$ cuando T es una función lineal.

Observación 6.2. (i) Claramente, (6.1) es equivalente a

$$T(x + \lambda y) = Tx + \lambda Ty \quad \text{para todos } x, y \in U \text{ y } \lambda \in \mathbb{K}. \quad (6.1)'$$

(ii) Se sigue inmediatamente de la definición que

$$T(\lambda_1 v_1 + \cdots + \lambda_k v_k) = \lambda_1 T v_1 + \cdots + \lambda_k T v_k$$

para todos $v_1, \dots, v_k \in V$ y $\lambda_1, \dots, \lambda_k \in \mathbb{K}$.

- (III) La condición (6.1) dice que **una aplicación lineal respeta las estructuras de espacio vectorial de su dominio y de su espacio de llegada.**

Ejercicio 6.3. Sean U, V espacios vectoriales sobre $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$). Denotemos por $\mathcal{L}(U, V)$ el conjunto de todas las aplicaciones lineales de U en V . Muestre que $\mathcal{L}(U, V)$ es un espacio vectorial sobre $\mathbb{K}$. Esto significa que debe mostrar que la suma de dos aplicaciones lineales es una aplicación lineal, que un múltiplo escalar de una aplicación lineal es una aplicación lineal y que se cumplen los axiomas de espacio vectorial.

Ejemplos 6.4 (Aplicaciones lineales). (a) Toda matriz $A \in M(m \times n)$ puede identificarse con una aplicación lineal $\mathbb{R}^n \rightarrow \mathbb{R}^m$.

- (b) La derivación es una aplicación lineal, por ejemplo:

- (i) Sea $C(\mathbb{R})$ el espacio de todas las funciones continuas y sea $C^1(\mathbb{R})$ el espacio de todas las funciones continuamente diferenciables. Entonces

$$T : C^1(\mathbb{R}) \rightarrow C(\mathbb{R}), \quad Tf = f'$$

es una aplicación lineal.

Demostración. Primero observe que $f' \in C(\mathbb{R})$ si $f \in C^1(\mathbb{R})$, así que la aplicación T está bien definida. Ahora queremos ver que es lineal. Tomamos entonces $f, g \in C^1(\mathbb{R})$ y $\lambda \in \mathbb{R}$. Obtenemos

$$T(\lambda f + g) = (\lambda f + g)' = (\lambda f)' + g' = \lambda f' + g' = \lambda Tf + Tg. \quad \square$$

- (ii) Las siguientes aplicaciones también son lineales. Observe que su acción es la misma que la de T anterior, pero cambiamos los espacios vectoriales sobre los que actúa.

$$R : P_n \rightarrow P_{n-1}, \quad Rf = f', \quad S : P_n \rightarrow P_n, \quad Sf = f'.$$

- (c) La integración es una aplicación lineal. Por ejemplo:

$$I : C([0, 1]) \rightarrow C([0, 1]), \quad f \mapsto If \quad \text{donde} \quad (If)(x) = \int_0^x f(t) dt.$$

Demostración. Claramente, I está bien definida, ya que la integral de una función continua es nuevamente continua. Para mostrar que I es lineal, fijamos $f, g \in C(\mathbb{R})$ y $\lambda \in \mathbb{R}$. Obtenemos para todo $x \in \mathbb{R}$:

$$\begin{aligned} (I(\lambda f + g))(x) &= \int_0^x (\lambda f + g)(t) dt = \int_0^x \lambda f(t) + g(t) dt = \lambda \int_0^x f(t) dt + \int_0^x g(t) dt \\ &= \lambda (If)(x) + (Ig)(x). \end{aligned}$$

Como esto es verdadero para todo x , se sigue que $I(\lambda f + g) = \lambda(If) + (Ig)$. $\square$

(d) Como ejemplo de una aplicación lineal de $M(n \times n)$ en sí mismo, consideramos

$$T : M(n \times n) \rightarrow M(n \times n), \quad T(A) = A + A^t.$$

Demostración de que T es una aplicación lineal. Sean $A, B \in M(n \times n)$ y sea $c \in \mathbb{R}$. Entonces

$$\begin{aligned} T(A + cB) &= (A + cB) + (A + cB)^t = A + cB + A^t + (cB)^t = A + cB + A^t + cB^t \\ &= A + A^t + c(B + B^t) = T(A) + cT(B). \end{aligned} \quad \square$$

(e) Un primer ejemplo de una **transformación no lineal**. Sea

$$T : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad T \begin{pmatrix} x \\ y \end{pmatrix} = xy.$$

$$\text{Entonces } T \begin{pmatrix} 1 \\ 0 \end{pmatrix} = T \begin{pmatrix} 0 \\ 1 \end{pmatrix} = xy = 0, \text{ pero } T \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \neq T \begin{pmatrix} 1 \\ 0 \end{pmatrix} + T \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

(f) Un segundo ejemplo de una **transformación no lineal**. Sea

$$T : \mathbb{R}^3 \rightarrow \mathbb{R}^2 \quad T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 3z \\ 2|y| \end{pmatrix}.$$

$$\text{Entonces } T \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \text{ pero } T \left(-3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) = T \begin{pmatrix} 0 \\ -3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \end{pmatrix} \neq \begin{pmatrix} 0 \\ -6 \end{pmatrix} = -3T \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

El siguiente lema muestra que una aplicación lineal siempre envía el vector cero al vector cero.

Lema 6.5. Si T es una aplicación lineal, entonces $T\mathbf{0} = \mathbf{0}$.

Demostración. $T\mathbf{0} = T(\mathbf{0} - \mathbf{0}) = T\mathbf{0} - T\mathbf{0} = \mathbf{0}$. □

Definición 6.6. Sea $T : U \rightarrow V$ una aplicación lineal.

(I) T se llama *inyectiva* (o *uno a uno*) si

$$x, y \in U, \quad x \neq y \quad \implies \quad Tx \neq Ty.$$

(II) T se llama *sobreyectiva* si para todo $v \in V$ existe al menos un $x \in U$ tal que $Tx = v$.

(III) T se llama *biyectiva* si es inyectiva y sobreyectiva.

(IV) El *kernel* de T (o *núcleo* de T o *espacio nulo* de T) es

$$\ker(T) := \{x \in U : Tx = \mathbf{0}\}.$$

A veces se usan las notaciones $N(T)$ o N_T para $\ker(T)$.

(v) La *imagen* de T (o *rango* de T) es

$$\operatorname{Im}(T) := \{y \in V : y = Tx \text{ para algún } x \in U\}.$$

A veces se usan las notaciones $\operatorname{Rg}(T)$, $\operatorname{R}(T)$ o $T(U)$ para $\operatorname{Im}(T)$.

Atención. Es preferible usar “imagen” en lugar de “rango” porque “rango” también tiene otro significado, vea Definición 6.12.

Observación 6.7. (I) Observe que $\ker(T)$ es un subconjunto de U y $\operatorname{Im}(T)$ es un subconjunto de V . En la Proposición 6.11 mostraremos que incluso son subespacios.

(II) Claramente, T es inyectiva si y solo si para todos $x, y \in U$ se cumple lo siguiente:

$$Tx = Ty \implies x = y.$$

(III) Si T es una aplicación lineal inyectiva, entonces su inversa $T^{-1} : \operatorname{Im}(T) \rightarrow U$ existe y también es lineal.

Ejercicio 6.8. Sean U, V, W espacios vectoriales sobre $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$).

- Suponga que $T : U \rightarrow V$ y $S : V \rightarrow W$ son funciones lineales. Muestre que su composición $ST : U \rightarrow W$ también es una función lineal.
Al comparar $\operatorname{Im}(ST)$ y $\operatorname{Im}(S)$, ¿qué puede concluir?
Al comparar $\ker(ST)$ y $\ker(S)$, ¿qué puede concluir?
- Suponga que $T : U \rightarrow V$ es una función lineal invertible, de modo que podemos definir su función inversa $T^{-1} : \operatorname{Im}(T) \rightarrow U$. Muestre que esta también es una función lineal.

El siguiente lema es muy útil.

Lema 6.9. Sea $T : U \rightarrow V$ una aplicación lineal.

- (I) T es inyectiva si y solo si $\ker(T) = \{\mathbb{O}\}$.
 (II) T es sobreyectiva si y solo si $\operatorname{Im}(T) = V$.

Demostración. (I) Por el Lema 6.5, sabemos que $\mathbb{O} \in \ker(T)$. Supongamos que T es inyectiva. Entonces el único vector $v \in U$ que satisface $Tv = \mathbb{O}$ es $v = \mathbb{O}$. Por lo tanto $\ker(T)$ contiene solamente a $\mathbb{O}$, de donde $\ker(T) = \{\mathbb{O}\}$.

Ahora supongamos que $\ker(T) = \{\mathbb{O}\}$ y sean $x, y \in U$ con $Tx = Ty$. Por la Observación 6.7, es suficiente mostrar que $x = y$. Por hipótesis, $\mathbb{O} = Tx - Ty = T(x - y)$, de donde $x - y \in \ker(T) = \{\mathbb{O}\}$. Por lo tanto $x - y = \mathbb{O}$, lo cual significa que $x = y$.

(II) se sigue directamente de las definiciones de sobreyectividad y de imagen de una aplicación lineal. □

Ejemplos 6.10 (Núcleos e imágenes de las aplicaciones lineales de los Ejemplos 6.4).

- (a) Discutiremos el caso de las matrices al comienzo de la Sección 6.2.

- (b) Si $T : C^1(\mathbb{R}) \rightarrow C(\mathbb{R})$, $Tf = f'$, entonces es fácil ver que el núcleo de T consiste exactamente en las funciones constantes. Además, T es sobreyectiva porque toda función continua es la derivada de otra función, pues para cada $f \in C(\mathbb{R})$ podemos poner $g(x) = \int_0^x f(t) dt$. Entonces $g \in C^1(\mathbb{R})$ y $Tg = g' = f$, lo cual muestra que $\text{Im}(T) = C(\mathbb{R})$.
- (c) Para el operador de integración del Ejemplo 6.4(c), tenemos que $\ker(I) = \{0\}$ y $\text{Im}(I) = C^1(\mathbb{R})$. En otras palabras, I es inyectiva pero no sobreyectiva.

Demostración. Primero demostramos la afirmación sobre la imagen de I . Supongamos que $g \in \text{Im}(I)$. Entonces g es de la forma $g(x) = \int_0^x f(t) dt$ para alguna $f \in C(\mathbb{R})$. Por el teorema fundamental del cálculo, se sigue que $g \in C^1(\mathbb{R})$, así que hemos demostrado $\text{Im}(I) \subseteq C^1(\mathbb{R})$. Para mostrar la otra inclusión, sea $g \in C^1(\mathbb{R})$. Entonces g es diferenciable y $g' \in C(\mathbb{R})$ y, de nuevo por el teorema fundamental del cálculo, tenemos que $g(x) = \int_0^x g'(t) dt$, así que $g \in \text{Im}(I)$ y se sigue que $C^1(\mathbb{R}) \subseteq \text{Im}(I)$.

Ahora supongamos que $Ig = 0$. Si derivamos, encontramos que $0 = (Ig)'(x) = \frac{d}{dx} \int_0^x g(t) dt = g(x)$ para todo $x \in \mathbb{R}$; por lo tanto, $g \equiv 0$ (la función cero), de donde $\ker(I) = \{0\}$. $\square$

- (d) Sea $T : M(n \times n) \rightarrow M(n \times n)$, $T(A) = A + A^t$. Entonces $\ker T = M_{\text{asym}}(n \times n)$ (= el espacio de todas las matrices antisimétricas $n \times n$) y $\text{Im} T = M_{\text{sym}}(n \times n)$ (= el espacio de todas las matrices simétricas $n \times n$).

Demostración. Primero demostramos la afirmación sobre la imagen de T . Claramente, $\text{Im}(T) \subseteq M_{\text{sym}}(n \times n)$ porque para toda $A \in M(n \times n)$ tenemos que $T(A)$ es simétrica, ya que $(T(A))^t = (A + A^t)^t = A^t + (A^t)^t = A^t + A = T(A)$. Para demostrar $M_{\text{sym}}(n \times n) \subseteq \text{Im}(T)$ tomamos alguna $B \in M_{\text{sym}}(n \times n)$. Entonces $T(\frac{1}{2}B) = \frac{1}{2}B + (\frac{1}{2}B)^t = \frac{1}{2}B + \frac{1}{2}B = B$, donde usamos que B es simétrica. En resumen, mostramos que $\text{Im}(T) = M_{\text{sym}}(n \times n)$.

La afirmación sobre el núcleo de T se sigue de

$$A \in \ker T \iff T(A) = 0 \iff A + A^t = 0 \iff A = -A^t \iff A \in M_{\text{asym}}(n \times n). \quad \square$$

Proposición 6.11. Sea $T : U \rightarrow V$ una aplicación lineal. Entonces

- (I) $\ker(T)$ es un subespacio de U .
 (II) $\text{Im}(T)$ es un subespacio de V .

Demostración. (I) Por el Lema 6.5 tenemos que $0 \in \ker(T)$. Sean $x, y \in \ker(T)$ y $\lambda \in \mathbb{K}$. Entonces $x + \lambda y \in \ker(T)$ porque

$$T(x + \lambda y) = Tx + \lambda Ty = 0 + \lambda 0 = 0.$$

De ahí concluimos que $\ker(T)$ es un subespacio de U por la Proposición 5.10.

- (II) Claramente, $0 \in \text{Im}(T)$. Sean $v, w \in \text{Im}(T)$ y $\lambda \in \mathbb{K}$. Entonces existen $x, y \in U$ tales que $Tx = v$ y $Ty = w$. Entonces $v + \lambda w = Tx + \lambda Ty = T(x + \lambda y) \in \text{Im}(T)$, de donde $v + \lambda w \in \text{Im}(T)$. Por lo tanto, $\text{Im}(T)$ es un subespacio de V por la Proposición 5.10. $\square$

Como ahora sabemos que $\ker(T)$ y $\text{Im}(T)$ son subespacios, la siguiente definición tiene sentido.

Definición 6.12. Sea $T : U \rightarrow V$ una aplicación lineal. Definimos

$$\dim(\ker(T)) = \text{nulidad de } T, \quad \dim(\operatorname{Im}(T)) = \text{rango de } T.$$

A veces se usan las notaciones $\nu(T) = \dim(\ker(T))$ y $\rho(T) = \dim(\operatorname{Im}(T))$.

Atención. El “rango” de la Definición 6.6 es un subespacio (en inglés es “range”), mientras el “rango” de la Definición 6.12 es la su dimensión, o sea, es un número (en inglés es “rnk”).

Ejemplo. Sea $T : P_3 \rightarrow P_3$ definida por $Tp = p'$. Entonces $\operatorname{Im}(T) = \{q \in P_3 : \deg q \leq 2\}$ y $\ker(T) = \{q \in P_3 : \deg q = 0\}$. En particular, $\dim(\operatorname{Im}(T)) = 3$ y $\dim(\ker(T)) = 1$.

Demostración. ■ Primero mostramos la afirmación sobre la imagen de T . Sabemos que la derivación disminuye en 1 el grado de un polinomio. De ahí que $\operatorname{Im}(T) \subseteq \{q \in P_3 : \deg q \leq 2\}$. Por otro lado, sabemos que todo polinomio de grado ≤ 2 es la derivada de un polinomio de grado ≤ 3 . Así se sigue la afirmación.

- Ahora mostramos la afirmación sobre el núcleo de T . Recordemos que $\ker(T) = \{p \in P_3 : Tp = 0\}$. Así, el núcleo de T consiste exactamente en aquellos polinomios cuya primera derivada es 0. Estos son exactamente los polinomios constantes, es decir, los polinomios de grado 0. □

Lema 6.13. Sea $T : U \rightarrow V$ una aplicación lineal entre dos espacios vectoriales U, V y sea $\{u_1, \dots, u_k\}$ una base de U . Entonces $\operatorname{Im} T = \operatorname{span}\{Tu_1, \dots, Tu_k\}$.

Demostración. Claramente, $Tu_1, \dots, Tu_k \in \operatorname{Im} T$. Como la imagen de T es un espacio vectorial, todas las combinaciones lineales de estos vectores también deben pertenecer a $\operatorname{Im} T$, lo cual muestra que $\operatorname{span}\{Tu_1, \dots, Tu_k\} \subseteq \operatorname{Im} T$. Para mostrar la otra inclusión, sea $y \in \operatorname{Im} T$. Entonces existe $x \in U$ tal que $y = Tx$. Expresemos x como combinación lineal de los vectores de la base: $x = \alpha_1 u_1 + \dots + \alpha_k u_k$. Entonces we obtain

$$y = Tx = T(\alpha_1 u_1 + \dots + \alpha_k u_k) = \alpha_1 Tu_1 + \dots + \alpha_k Tu_k \in \operatorname{span}\{Tu_1, \dots, Tu_k\}.$$

Como y era arbitrario en $\operatorname{Im} T$, concluimos que $\operatorname{Im} T \subseteq \operatorname{span}\{Tu_1, \dots, Tu_k\}$. Así, en resumen, hemos demostrado la afirmación. □

Proposición 6.14. Sean U, V espacios vectoriales sobre $\mathbb{K}$ y sea $T : U \rightarrow V$ una aplicación lineal. Sean $x_1, \dots, x_k \in U$ y pongamos $y_1 := Tx_1, \dots, y_k := Tx_k$. Entonces se cumple lo siguiente.

- (I) Si $x_1, \dots, x_k$ son linealmente dependientes, entonces $y_1, \dots, y_k$ también son linealmente dependientes.
- (II) Si $y_1, \dots, y_k$ son linealmente independientes, entonces $x_1, \dots, x_k$ también son linealmente independientes.
- (III) Suponga además que T es invertible. Entonces $x_1, \dots, x_k$ son linealmente independientes si y solo si $y_1, \dots, y_k$ son linealmente independientes.

En general, la implicación “Si $x_1, \dots, x_k$ son linealmente independientes, entonces $y_1, \dots, y_k$ son linealmente independientes.” es *falsa*. ¿Puede dar un ejemplo?

Demostración de la Proposición 6.14. (I) Supongamos que los vectores $x_1, \dots, x_k$ son linealmente dependientes. Entonces existen $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ tales que $\lambda_1 x_1 + \dots + \lambda_k x_k = \mathbb{O}$ y at least one $\lambda_j \neq 0$. Pero entonces

$$\begin{aligned}\mathbb{O} &= T\mathbb{O} = T(\lambda_1 x_1 + \dots + \lambda_k x_k) = \lambda_1 T x_1 + \dots + \lambda_k T x_k \\ &= \lambda_1 y_1 + \dots + \lambda_k y_k,\end{aligned}$$

de donde los vectores $y_1, \dots, y_k$ son linealmente dependientes.

(II) se sigue directamente de (I).

(III) Supongamos que los vectores $y_1, \dots, y_k$ son linealmente independientes. Entonces también lo son $x_1, \dots, x_k$ por (I). Ahora supongamos que $x_1, \dots, x_k$ son linealmente independientes. Observe que T es invertible, así que T^{-1} existe. Por lo tanto, podemos aplicar (I) a T^{-1} para concluir que el sistema $y_1, \dots, y_k$ es linealmente independiente. (Observe que $x_j = T^{-1}y_j$.) $\square$

Ejercicio 6.15. Suponga que $T : U \rightarrow V$ es una aplicación lineal inyectiva y que $\{u_1, \dots, u_\ell\}$ es un conjunto de vectores linealmente independientes en U . Muestre que $\{Tu_1, \dots, Tu_\ell\}$ es un conjunto de vectores linealmente independientes en V .

La siguiente proposición es **muy** útil y se usa en la demostración del Teorema 6.18.

Proposición 6.16. Sean U, V espacios vectoriales sobre $\mathbb{K}$ con $\dim U = k < \infty$.

- (I) Si $T : U \rightarrow V$ es una transformación lineal, entonces $\dim \operatorname{Im}(T) \leq \dim U$.
- (II) Si $T : U \rightarrow V$ es una transformación lineal inyectiva, entonces $\dim \operatorname{Im}(T) = \dim U$.
- (III) Si $T : U \rightarrow V$ es una transformación lineal biyectiva, entonces $\dim U = \dim V$.

Demostración. Sea $u_1, \dots, u_k$ una base de U .

(I) Por el Lema 6.13, sabemos que $\operatorname{Im} T = \operatorname{span}\{Tu_1, \dots, Tu_k\}$. Por lo tanto $\dim \operatorname{Im} T \leq k = \dim U$ por el Teorema 5.47.

(II) Supongamos que T es inyectiva. Mostraremos que $Tu_1, \dots, Tu_k$ son linealmente independientes. Sean $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ tales que $\alpha_1 Tu_1 + \dots + \alpha_k Tu_k = \mathbb{O}$. Entonces

$$\mathbb{O} = \alpha_1 Tu_1 + \dots + \alpha_k Tu_k = T(\alpha_1 u_1 + \dots + \alpha_k u_k).$$

Como T es inyectiva, se sigue que $\alpha_1 u_1 + \dots + \alpha_k u_k = \mathbb{O}$, de donde $\alpha_1 = \dots = \alpha_k = 0$, lo cual muestra que los vectores $Tu_1, \dots, Tu_k$ son efectivamente linealmente independientes. Por lo tanto son una base de $\operatorname{span}\{Tu_1, \dots, Tu_k\} = \operatorname{Im} T$ y concluimos que $\dim \operatorname{Im} T = k = \dim U$.

(III) Como T es biyectiva, es sobreyectiva e inyectiva. La sobreyectividad significa que $\operatorname{Im} T = V$, y la inyectividad de T implica que $\dim \operatorname{Im} T = \dim U$ por (II). En conclusión,

$$\dim U = \dim \operatorname{Im} T = \dim V.$$

$\square$

La proposición anterior nos dice, por ejemplo, que no existe una aplicación lineal inyectiva de $\mathbb{R}^5$ a $\mathbb{R}^3$; o que no existe una aplicación lineal sobreyectiva de $\mathbb{R}^3$ a $M(2 \times 2)$.

Observación 6.17. La Proposición 6.16 también es verdadera para $\dim U = \infty$. En este caso, (i) se cumple claramente cualquiera que sea $\dim \operatorname{Im}(T)$. Para demostrar (ii), debemos mostrar que $\dim \operatorname{Im}(T) = \infty$ si T es inyectiva. Observe que para todo $n \in \mathbb{N}$ podemos encontrar un subespacio U_n de U con $\dim U_n = n$ y definimos T_n como la restricción de T a U_n , es decir, $T_n : U_n \rightarrow V$. Como la restricción de una aplicación inyectiva es inyectiva, se sigue de (ii) que $\dim \operatorname{Im}(T_n) = n$. Por otro lado, $\operatorname{Im}(T_n)$ es un subespacio de V ; por lo tanto $\dim V \geq \dim \operatorname{Im}(T_n) = n$ por Theorem 5.54 y Remark 5.55. Como esto es verdadero para todo $n \in \mathbb{N}$, se sigue que $\dim V = \infty$. La demostración de (iii) es la misma que en el caso de dimensión finita.

Teorema 6.18. Sean U, V espacios vectoriales sobre $\mathbb{K}$ y sea $T : U \rightarrow V$ una aplicación lineal. Además, sean $E : U \rightarrow U$, $F : V \rightarrow V$ aplicaciones lineales biyectivas. Entonces se cumple lo siguiente:

- (i) $\operatorname{Im}(T) = \operatorname{Im}(TE)$, in particular $\dim(\operatorname{Im}(T)) = \dim(\operatorname{Im}(TE))$.
- (ii) $\ker(TE) = E^{-1}(\ker(T))$ y $\dim(\ker(T)) = \dim(\ker(TE))$.
- (iii) $\ker(T) = \ker(FT)$, in particular $\dim(\ker(T)) = \dim(\ker(FT))$.
- (iv) $\operatorname{Im}(FT) = F(\operatorname{Im}(T))$ y $\dim(\operatorname{Im}(T)) = \dim(\operatorname{Im}(FT))$.

En resumen tenemos

$$\begin{array}{ll} \ker(FT) = \ker(T), & \ker(TE) = E^{-1}(\ker(T)), \\ \operatorname{Im}(FT) = F(\operatorname{Im}(T)), & \operatorname{Im}(TE) = \operatorname{Im}(T). \end{array} \quad (6.2)$$

and

$$\begin{array}{l} \dim \ker(T) = \dim \ker(FT) = \dim \ker(TE) = \dim \ker(FTE), \\ \dim \operatorname{Im}(T) = \dim \operatorname{Im}(FT) = \dim \operatorname{Im}(TE) = \dim \operatorname{Im}(FTE). \end{array} \quad (6.3)$$

Demostración. (i) Sea $v \in V$. Si $v \in \operatorname{Im}(T)$, entonces existe $x \in U$ tal que $Tx = v$. Pongamos $y = E^{-1}x$. Entonces $v = Tx = TEE^{-1}x = TEy \in \operatorname{Im}(TE)$. Por otro lado, si $v \in \operatorname{Im}(TE)$, entonces existe $y \in U$ tal que $TEy = v$. Pongamos $x = Ey$. Entonces $v = TEy = Tx \in \operatorname{Im}(T)$.

(ii) Para mostrar que $\ker(TE) = E^{-1}\ker(T)$, observe que

$$\ker(TE) = \{x \in U : Ex \in \ker(T)\} = \{E^{-1}u : u \in \ker(T)\} = E^{-1}(\ker(T)).$$

Se sigue que

$$E^{-1} : \ker T \rightarrow \ker(TE)$$

es una biyección lineal y por lo tanto $\dim \ker T = \dim \ker(TE)$ por la Proposición 6.16(iii) (o la Observación 6.17 en el caso de dimensión infinita), tomando E^{-1} como T , $\ker(T)$ como U y $\ker(TE)$ como V .

(III) Sea $x \in U$. Entonces $x \in \ker(FT)$ si y solo si $FTx = \mathbb{O}$. Como F es inyectiva, sabemos que $\ker(F) = \{\mathbb{O}\}$, de donde se sigue que $Tx = \mathbb{O}$. Pero esto es equivalente a $x \in \ker(T)$.

(IV) Para mostrar que $\text{Im}(FT) = F \text{Im}(T)$, observe que

$$\text{Im}(FT) = \{y \in V : y = FTx \text{ para algún } x \in U\} = \{Fv : v \in \text{Im}(T)\} = F(\text{Im}(T)).$$

Se sigue que

$$F : \text{Im } T \rightarrow \text{Im}(FT)$$

es una biyección lineal y por lo tanto $\dim \text{Im } T = \dim \text{Im}(FT)$ por la Proposición 6.16(III) (o la Observación 6.17 en el caso de dimensión infinita), tomando F como T , $\text{Im}(T)$ como U y $\text{Im}(FT)$ como V . $\square$

Observación 6.19. En general, $\ker(T) = \ker(TE)$ y $\ker(T) = \ker(FT)$ son afirmaciones falsas. Tomemos por ejemplo $U = V = \mathbb{R}^2$, $T = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ y $E = F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Entonces claramente se satisfacen las hipótesis del teorema y

$$\ker(T) = \text{span} \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \quad \text{Im}(T) = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\},$$

pero

$$\ker(TE) = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}, \quad \text{Im}(FT) = \text{span} \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

Dibuje una figura para visualizar el ejemplo anterior, teniendo en cuenta que T representa la proyección sobre el eje x y que E y F son una rotación de 45° y un “estiramiento” por el factor $\sqrt{2}$.

Terminamos esta sección con uno de los teoremas principales del álgebra lineal. En la siguiente sección volveremos a demostrarlo para el caso especial en que T está dada por una matriz, en el Teorema 6.33. El teorema siguiente puede considerarse una versión libre de coordenadas del Teorema 6.33.

Teorema 6.20. Sean U, V espacios vectoriales con $\dim U = n < \infty$ y sea $T : U \rightarrow V$ una aplicación lineal. Entonces

$$\dim(\ker(T)) + \dim(\text{Im}(T)) = n. \quad (6.4)$$

Demostración. Sea $k = \dim(\ker(T))$ y sea $\{u_1, \dots, u_k\}$ una base de $\ker(T)$. La completamos a una base $\{u_1, \dots, u_k, w_{k+1}, \dots, w_n\}$ of U y ponemos $W := \text{span}\{w_{k+1}, \dots, w_n\}$. Observe que por construcción $\ker(T) \cap W = \{\mathbb{O}\}$. (¡Demuestre esto!) Consideremos $\tilde{T} = T|_W$, la restricción de T a W .

Se sigue que $\tilde{T}$ es inyectiva porque si $\tilde{T}x = \mathbb{O}$ para algún $x \in W$ entonces también $Tx = \tilde{T}x = \mathbb{O}$, de donde $x \in \ker(T) \cap W = \{\mathbb{O}\}$. Se sigue de la Proposición 6.16(II) que

$$\dim \text{Im } \tilde{T} = \dim W = n - k. \quad (6.5)$$

Para completar la demostración, basta mostrar que $\text{Im } \tilde{T} = \text{Im } T$. Recordemos que, por el Lema 6.13, la imagen de una aplicación lineal está generada por las imágenes de una base del espacio vectorial inicial. Por lo tanto encontramos que

$$\begin{aligned}\text{Im } T &= \text{span}\{Tu_1, \dots, Tu_k, Tw_{k+1}, \dots, Tw_n\} = \text{span}\{Tw_{k+1}, \dots, Tw_n\} \\ &= \text{span}\{\tilde{T}w_{k+1}, \dots, \tilde{T}w_n\} \\ &= \text{Im } \tilde{T}\end{aligned}$$

donde en el segundo paso usamos que $Tu_1 = \dots = Tu_k = \mathbb{O}$ y por lo tanto no contribuyen al subespacio generado, y en el tercer paso usamos que $Tw_j = \tilde{T}w_j$ para $j = k+1, \dots, n$. Así mostramos que $\text{Im } \tilde{T} = \text{Im } T$; en particular, sus dimensiones son iguales y la afirmación se sigue de (6.5) porque, recordando que $k = \dim \ker(T)$,

$$n = \dim \text{Im } \tilde{T} + k = \dim \text{Im } T + \dim \ker T. \quad \square$$

Observe que una forma alternativa de demostrar el teorema anterior es demostrar primero el Teorema 6.33 para matrices y luego usar los resultados sobre representaciones de aplicaciones lineales de la Sección 6.4 para concluir la fórmula (6.4).

Ahora debería haber entendido

- qué es una aplicación lineal y por qué son las aplicaciones naturales que se consideran en espacios vectoriales,
- qué significan inyectividad, sobreyectividad y biyectividad,
- qué son el núcleo y la imagen de una aplicación lineal,
- por qué la fórmula de la dimensión (6.4) es verdadera,
- etc.

Ahora debería ser capaz de

- dar ejemplos de aplicaciones lineales,
- verificar si una función dada es una aplicación lineal,
- encontrar bases y la dimensión de núcleos e imágenes de una aplicación lineal dada,
- etc.

Ejercicios.

De los ejercicios 1 al 14 determinar si la función dada es una transformación lineal. Si lo es demuéstrelo, en caso contrario dé un ejemplo donde no se cumpla la linealidad.

$$1. \quad T: \mathbb{R}^4 \rightarrow \mathbb{R}^3, \quad T \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} x + 2y \\ 3y - 5z \\ w \end{pmatrix}.$$

$$2. T : \mathbb{R}^3 \rightarrow \mathbb{R}^2, T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 1 \end{pmatrix}.$$

$$3. T : \mathbb{R}^2 \rightarrow \mathbb{R}^2, T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x^2 \end{pmatrix}.$$

$$4. T : \mathbb{R}^2 \rightarrow \mathbb{R}^2, T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix}.$$

$$5. T : \mathbb{R}^4 \rightarrow \mathbb{R}, T \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = |w|.$$

$$6. T : \mathbb{R}^4 \rightarrow \mathbb{R}^3, T \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} x - y + 2z + 3w \\ y + 4z + 3w \\ x + 6z + 6w \end{pmatrix}.$$

$$7. T : P_3 \rightarrow P_4, T(p) = \int_0^x p(t) dt.$$

$$8. T : \mathbb{R}^3 \rightarrow P_3, T \begin{pmatrix} a \\ b \\ c \end{pmatrix} = (a - 3b)x^3 + (b + 2c).$$

$$9. T : M(3 \times 3) \rightarrow M(3 \times 3), T(A) = A^t - A.$$

$$10. T : M(2 \times 2) \rightarrow M(2 \times 2), T(A) = A \begin{pmatrix} 2 & 4 \\ 0 & 1 \end{pmatrix}.$$

$$11. \text{ Sea } g \in C^1(\mathbb{R}) \text{ y } T : C^1(\mathbb{R}) \rightarrow C(\mathbb{R}) \text{ dada por: } T(f) = (gf)'.$$

$$12. T : M(3 \times 3) \rightarrow \mathbb{R}, T \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} + a_{22} + a_{33}.$$

$$13. \text{ Sea } \vec{x}_0 \in \mathbb{R}^n \text{ y } T : M(m \times n) \rightarrow \mathbb{R}^m \text{ dada por } T(A) = A\vec{x}_0.$$

$$14. T : M(n \times n) \rightarrow \mathbb{R}, T(A) = \det A.$$

15. De los ejercicios anteriores, (excepto el Ejercicio 11.) determine $\text{Im } T$, $\ker T$ y sus dimensiones.

16. Sea $\vec{w} \in \mathbb{R}^n$ un vector no nulo y $T : \mathbb{R}^n \rightarrow \mathbb{R}$ dada por $T(\vec{x}) = \langle \vec{x}, \vec{w} \rangle$. Demuestre que T es una transformación lineal y determine las dimensiones de $\ker T$ e $\text{Im } T$.

17. Sea $\vec{w} \in \mathbb{R}^n$ un vector no nulo y $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ dada por $T(\vec{x}) = \text{proy}_{\vec{w}} \vec{x}$. Demuestre que T es lineal, encuentre $\text{Im } T$, $\ker T$ y sus dimensiones.

18. Sea $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por $T(\vec{x}) = \vec{x} \times \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, ¿ T es una transformación lineal? En caso afirmativo encuentre $\ker T$ e $\operatorname{Im} T$ y sus dimensiones. (Pista: Sale muy sencillo si lo piensa geoméricamente).

19. Sea $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ una transformación lineal tal que:

$$T \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}, \quad T \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ -9 \end{pmatrix},$$

encontrar $T \begin{pmatrix} 7 \\ 11 \end{pmatrix}$ y aún más general, determinar $T \begin{pmatrix} x \\ y \end{pmatrix}$. ¿ T es inyectiva?. ¿Cómo cambia la respuesta si $T \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$?

20. ¿Existe una transformación lineal $T : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ tal que

$$T \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad T \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ -1 \\ -3 \end{pmatrix}, \quad T \begin{pmatrix} 1 \\ -14 \\ 8 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 0 \\ -5 \end{pmatrix}?$$

6.2. Matrices como aplicaciones lineales

En esta sección trabajaremos principalmente con espacios vectoriales reales, para fijar ideas. Sin embargo, todas las afirmaciones también son verdaderas para espacios vectoriales complejos. Solo tenemos que reemplazar en todas partes $\mathbb{R}$ por $\mathbb{C}$ y la palabra *real* por *complejo*.

Sea $A \in M(m \times n)$. Ya sabemos que podemos ver A como una aplicación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$. Por lo tanto, $\ker(A)$ y $\operatorname{Im}(A)$, así como los términos *inyectividad* y *sobreyectividad*, están definidos.

Estrictamente hablando, deberíamos distinguir entre una matriz y la aplicación lineal inducida por ella. Así, deberíamos escribir $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ para la aplicación $x \mapsto Ax$. La razón es que, si vemos A directamente como una aplicación lineal, esto implica que tácitamente ya hemos escogido una base en $\mathbb{R}^n$ y en $\mathbb{R}^m$; véase la Sección 6.4 para más detalles. Sin embargo, usualmente abusaremos de la notación y escribiremos A en lugar de T_A .

Si vemos una matriz A como una aplicación lineal y al mismo tiempo como un sistema de ecuaciones lineales, entonces obtenemos lo siguiente.

Observación 6.21. Sea $A \in M(m \times n)$ y denotemos las columnas de A por $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{R}^m$. Entonces se cumple lo siguiente.

- (I) $\ker(A) =$ todas las soluciones $\vec{x}$ del sistema homogéneo $A\vec{x} = \vec{0}$.

- (II) $\text{Im}(A) = \text{todos los vectores } \vec{b} \text{ tales que el sistema } A\vec{x} = \vec{b} \text{ tiene solución}$
 $= \text{span}\{\vec{a}_1, \dots, \vec{a}_n\}.$

En consecuencia,

- (III) A es inyectiva $\iff \ker(A) = \{\vec{0}\}$
 $\iff$ el sistema homogéneo $A\vec{x} = \vec{0}$ tiene solamente la solución trivial $\vec{x} = \vec{0}$.
- (IV) A es sobreyectiva $\iff \text{Im}(A) = \mathbb{R}^m$
 $\iff$ para todo $\vec{b} \in \mathbb{R}^m$, el sistema $A\vec{x} = \vec{b}$ tiene al menos una solución.

Demostración. Todas las afirmaciones deberían ser claras, excepto quizás la segunda igualdad en (II). Esto se sigue de

$$\begin{aligned} \text{Im } A &= \{A\vec{x} : \vec{x} \in \mathbb{R}^n\} = \left\{ (\vec{a}_1 | \dots | \vec{a}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \right\} \\ &= \{x_1\vec{a}_1 + \dots + x_n\vec{a}_n : x_1, \dots, x_n \in \mathbb{R}\} \\ &= \text{span}\{\vec{a}_1, \dots, \vec{a}_n\}, \end{aligned}$$

véase también la Observación 3.19. □

Para practicar un poco, demostramos las dos observaciones siguientes de dos maneras.

Observación 6.22. Sea $A \in M(m \times n)$. Si $m > n$, entonces A no puede ser sobreyectiva.

Demostración con Gauß-Jordan. Sea A' la forma escalonada reducida por filas de A . Entonces debe existir una matriz invertible E tal que $A = EA'$ y la última fila de A' debe ser cero porque puede tener a lo más n pivotes. Pero entonces $(A'|\vec{e}_m)$ es inconsistente, o lo que significa que $(A|E^{-1}\vec{e}_m)$ es inconsistente. De ahí que $E^{-1}\vec{e}_m \notin \text{Im } A$, por lo que A no puede ser sobreyectiva. (Basicamente estamos diciendo que claramente A' no es sobreyectiva porque podemos encontrar fácilmente un lado derecho tal que $A'\vec{x}' = \vec{b}'$ es inconsistente. Es suficiente escoger cualquier $\vec{b}'$ cuya última coordenada es diferente de 0. El vector más fácil es $\vec{e}_m$. Ahora haga el proceso de Gauß-Jordan hacia atrás sobre este vector para obtener un lado derecho $\vec{b}$ tal que $A\vec{x} = \vec{b}$ sea inconsistente.) □

Demostración usando el concepto de dimensión. Ya vimos que $\text{Im } A$ es el subespacio generado por sus columnas. Por lo tanto $\dim \text{Im } A \leq \# \text{columnas de } A = n < m = \dim \mathbb{R}^m$, por lo tanto $\text{Im } A \subsetneq \mathbb{R}^m$. □

Observación 6.23. Sea $A \in M(m \times n)$. Si $m < n$, entonces A no puede ser inyectiva.

Demostración con Gauß-Jordan. Sea A' la forma escalonada reducida por filas de A . Entonces A' can have at most m pivots. Como A' tiene más columnas que pivotes, el sistema homogéneo $A\vec{x} = \vec{0}$ tiene infinitas soluciones; entonces también $\ker A$ contiene infinitos vectores y, en particular, A no puede ser inyectiva. □

Demostración usando el concepto de dimensión. Ya vimos que $\text{Im } A$ es el subespacio generado por sus n columnas en $\mathbb{R}^m$. Como $n > m$, se sigue que los vectores columna son linealmente dependientes

en $\mathbb{R}^m$, de donde $A\vec{x} = \vec{0}$ tiene una solución no trivial. Por lo tanto $\ker A$ no es trivial y se sigue que A no es inyectiva. $\square$

Observe que las observaciones **no** implican que A sea sobreyectiva si $m \leq n$, ni que A sea inyectiva si $n \leq m$. ¡Encuentre ejemplos!

Del Teorema 3.44 obtenemos el siguiente teorema muy importante para el caso especial $m = n$.

Teorema 6.24. Sea $A \in M(n \times n)$ una matriz cuadrada. Entonces las siguientes afirmaciones son equivalentes.

- (I) A es invertible.
- (II) A es inyectiva, es decir, $\ker A = \{\vec{0}\}$.
- (III) A es sobreyectiva, es decir, $\text{Im } A = \mathbb{R}^n$.

En particular, A es inyectiva si y solo si A es sobreyectiva si y solo si A es biyectiva.

Definición 6.25. Sea $A \in M(m \times n)$ y sean $\vec{c}_1, \dots, \vec{c}_n$ las columnas de A y $\vec{r}_1, \dots, \vec{r}_m$ las filas de A . Definamos

- (I) $C_A := \text{span}\{\vec{c}_1, \dots, \vec{c}_n\} =: \text{espacio columna de } A \subseteq \mathbb{R}^m$,
- (II) $R_A := \text{span}\{\vec{r}_1, \dots, \vec{r}_m\} =: \text{espacio fila de } A \subseteq \mathbb{R}^n$,

La siguiente proposición se sigue inmediatamente de la definición anterior y de la Observación 6.21(II).

Proposición 6.26. For $A \in M(m \times n)$ it follows that

- (I) $R_A = C_{A^t}$ y $C_A = R_{A^t}$,
- (II) $C_A = \text{Im}(A)$ y $R_A = \text{Im}(A^t)$.

La siguiente proposición se sigue directamente de la teoría general de la Sección 6.1. Daremos otra demostración al final de esta sección.

Proposición 6.27. Sea $A \in M(m \times n)$, $E \in M(n \times n)$, $F \in M(m \times m)$ y supongamos que E y F son invertibles. Entonces

- (I) $C_A = C_{AE}$.
- (II) $R_A = R_{FA}$.

Demostración. (I) Observe que $C_A = \text{Im}(A) = \text{Im}(AE) = C_{AE}$, donde en la primera y en la tercera ecuación usamos la Proposición 6.26, y en la segunda ecuación usamos el Teorema 6.18.

- (II) Recuerde que, si F es invertible, entonces F^t también es invertible. Con la Proposición 6.26(I) y lo que ya probamos en (I), obtenemos $R_{FA} = C_{(FA)^t} = C_{A^t F^t} = C_{A^t} = R_A$. $\square$

Obtenemos inmediatamente la siguiente proposición.

Proposición 6.28. Sea $A, B \in M(m \times n)$.

(I) Si A y B son equivalentes por filas, entonces

$$\dim(\ker(A)) = \dim(\ker(B)), \quad \dim(\operatorname{Im}(A)) = \dim(\operatorname{Im}(B)), \quad \operatorname{Im}(A^t) = \operatorname{Im}(B^t), \quad R_A = R_B.$$

(II) Si A y B son equivalentes por columnas, entonces

$$\dim(\ker(A)) = \dim(\ker(B)), \quad \dim(\operatorname{Im}(A)) = \dim(\operatorname{Im}(B)), \quad \operatorname{Im}(A) = \operatorname{Im}(B), \quad C_A = C_B.$$

Demostración. Solo demostraremos (I). La afirmación (II) se puede demostrar de manera similar (o se puede deducir fácilmente de (I) aplicando (I) a las matrices transpuestas). Que A y B sean equivalentes por filas significa que podemos transformar B en A mediante transformaciones por filas. Como las transformaciones por filas se pueden representar mediante multiplicación por matrices elementales por la izquierda, existen matrices elementales $F_1, \dots, F_k \in M(m \times m)$ tales que $A = F_1 \dots F_k B$. Observe que todas las F_j son invertibles, consecuentemente $F := F_1 \dots F_k$ es invertible y $A = FB$. Por lo tanto, todas las afirmaciones en (I) se siguen del Teorema 6.18 y Proposición 6.27. $\square$

La proposición anterior es muy útil para calcular el núcleo de una matriz A : sea A' la forma escalonada reducida por filas de A . Entonces podemos aplicar la proposición a A y a A' , y vemos que $\ker(A) = \ker(A')$.

De hecho, sabemos esto desde el primer capítulo de este curso, pero entonces no teníamos a nuestra disposición palabras elegantes como “núcleo”. No dice otra cosa que: las soluciones de un sistema homogéneo no cambian si aplicamos transformaciones por filas, que es exactamente la razón por la cual funciona la eliminación de Gauß-Jordan.

En los Ejemplos 6.36 y 6.37 calcularemos el núcleo y la imagen de una matriz. Ahora demostraremos dos lemas técnicos.

Lema 6.29. Sea $A \in M(m \times n)$. Entonces existen matrices elementales $E_1, \dots, E_k \in M(n \times n)$ and $F_1, \dots, F_\ell \in M(m \times m)$ tales que

$$F_1 \dots F_\ell A E_1 \dots E_k = A''$$

donde A'' es de la forma

$$A'' = \left(\begin{array}{c|c} \overbrace{\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \end{pmatrix}}^r & \overbrace{\begin{pmatrix} 0 & & \\ & & \\ & & \end{pmatrix}}^{n-r} \\ \hline \begin{pmatrix} 0 & & & \\ & & & \\ & & & \end{pmatrix} & \begin{pmatrix} 0 & & \\ & & \\ & & \end{pmatrix} \end{array} \right) \begin{array}{l} \left. \vphantom{\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}} \right\} r \\ \left. \vphantom{\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}} \right\} m-r \end{array} \quad (6.6)$$

Demostración. Como $C_A = \text{Im}(A)$ por la Proposición 6.26, la primera igualdad es clara.

Now let $F_1, \dots, F_\ell, E_1, \dots, E_k$ y A', A'' be as in Lemma 6.29 y set $F := F_1 \cdots F_\ell$ and $E := E_1 \cdots E_k$. Entonces

$$\begin{aligned} \dim(R_A) &= \dim(R_{FAE}) = \dim(R_{A''}) = r = \dim(C_{A''}) = \dim(C_{FAE}) \\ &= \dim(C_A). \end{aligned}$$

□

Como consecuencia inmediata obtenemos el siguiente teorema, que es un caso especial del Teorema 6.20; véase también el Teorema 6.47.

Teorema 6.33. Sea $A \in M(m \times n)$. Entonces

$$\dim(\ker(A)) + \dim(\text{Im}(A)) = n. \quad (6.8)$$

Demostración. Con la notación anterior, obtenemos

$$\begin{aligned} \dim(\ker(A)) &= \dim(\ker(A'')) = n - r, \\ \dim(\text{Im}(A)) &= \dim(\text{Im}(A'')) = r \end{aligned}$$

y se sigue la fórmula deseada. □

Para el cálculo de una base de $\text{Im}(A)$, el siguiente teorema es útil.

Teorema 6.34. Sea $A \in M(m \times n)$ y let A' be its reduced row-echelon form con columns $\vec{c}_1, \dots, \vec{c}_n$ and $\vec{c}_1', \dots, \vec{c}_n'$ respectively. Suponga que las columnas con pivotes de A' son las columnas $j_1 < \dots < j_k$. Entonces $\dim(\text{Im}(A)) = k$ y una base de $\text{Im}(A)$ está dada por las columnas $\vec{c}_{j_1}, \dots, \vec{c}_{j_k}$ de A .

Demostración. Sea E una matriz invertible tal que $A = EA'$. Por hipótesis sobre las columnas con pivotes de la matriz A' , sabemos que $\dim(\text{Im}(A')) = k$ y que una base de $\text{Im}(A')$ está dada por las columnas $\vec{c}_{j_1}', \dots, \vec{c}_{j_k}'$. Por el Teorema 6.18 se sigue que $\dim(\text{Im}(A)) = \dim(\text{Im}(A')) = k$. Ahora observe que por la definición de E tenemos que $E\vec{c}_\ell' = \vec{c}_\ell$ para todo $\ell = 1, \dots, n$; en particular esto se tiene para las columnas con pivotes de la matriz A' . Adicionalmente, como E es invertible y los vectores $\vec{c}_{j_1}', \dots, \vec{c}_{j_k}'$ son linealmente independientes, se sigue del Teorema 6.14 que los vectores $\vec{c}_{j_1}, \dots, \vec{c}_{j_k}$ son linealmente independientes. Claramente pertenecen a $\text{Im}(A)$, por lo que $\text{span}\{\vec{c}_{j_1}, \dots, \vec{c}_{j_k}\} \subseteq \text{Im}(A)$. Como ambos espacios tienen la misma dimensión, necesariamente son iguales. □

Observación 6.35. El teorema anterior se puede usar para determinar una base de un subespacio dado en la forma $U = \text{span}\{\vec{v}_1, \dots, \vec{v}_k\} \subseteq \mathbb{R}^m$ as follows: Defina la matriz $A = (\vec{v}_1 | \dots | \vec{v}_k)$. Entonces claramente $U = \text{Im } A$ y podemos aplicar el Teorema 6.34 para encontrar una base de U .

Ejemplo 6.36. Encuentre $\ker(A)$, $\text{Im}(A)$, $\dim(\ker(A))$, $\dim(\text{Im}(A))$ y R_A para

$$A = \begin{pmatrix} 1 & 1 & 5 & 1 \\ 3 & 2 & 13 & 1 \\ 0 & 2 & 4 & -1 \\ 4 & 5 & 22 & 1 \end{pmatrix}.$$

Solución. Primero, reduzcamos por filas la matriz A :

$$\begin{aligned}
 A = \begin{pmatrix} 1 & 1 & 5 & 1 \\ 3 & 2 & 13 & 1 \\ 0 & 2 & 4 & -1 \\ 4 & 5 & 22 & 1 \end{pmatrix} &\xrightarrow[\substack{Q_{41}(-4)}]{Q_{21}(-1)} \begin{pmatrix} 1 & 1 & 5 & 1 \\ 0 & -1 & -2 & -2 \\ 0 & 2 & 4 & -1 \\ 0 & 1 & 2 & -3 \end{pmatrix} \xrightarrow[\substack{Q_{42}(1)}]{Q_{32}(2)} \begin{pmatrix} 1 & 1 & 5 & 1 \\ 0 & -1 & -2 & -2 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & -5 \end{pmatrix} \\
 &\xrightarrow[\substack{Q_{43}(-1)}]{S_2(-1)} \begin{pmatrix} 1 & 1 & 5 & 1 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\substack{Q_{12}(-1)}]{S_4(1/5)} \begin{pmatrix} 1 & 0 & 3 & -1 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\substack{Q_{24}(-2)}]{Q_{14}(1)} \begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} =: A'.
 \end{aligned}$$

Ahora se sigue inmediatamente que $\dim R_A = \dim C_A = 3$ y

$$\begin{aligned}
 \dim(\operatorname{Im}(A)) &= \# \text{ columnas con pivotes de } A' = 3, \\
 \dim(\ker(A)) &= 4 - \dim(\operatorname{Im}(A)) = 1
 \end{aligned}$$

(o: $\dim(\operatorname{Im}(A)) = \# \text{ filas no nulas de } A' = 3$, o: $\dim(\operatorname{Im}(A)) = \dim(R_A) = 3$ o: $\dim(\ker(A)) = \# \text{ columnas sin pivote de } A' = 1$).

Núcleo de A : Sabemos que $\ker(A) = \ker(A')$ por el Teorema 6.18 o Proposition 6.28. De la forma explícita de A' es claro que $A'\vec{x} = 0$ si y solo si $x_4 = 0$, x_3 arbitrario, $x_2 = -2x_3$ y $x_1 = -3x_3$. Por lo tanto

$$\ker(A) = \ker(A') = \left\{ \begin{pmatrix} -3x_3 \\ -2x_3 \\ x_3 \\ 0 \end{pmatrix} : x_3 \in \mathbb{R} \right\} = \operatorname{span} \left\{ \begin{pmatrix} -3 \\ -2 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Imagen de A : Las columnas con pivotes de A' son las columnas 1, 2 y 4. Por lo tanto, por el Teorema 6.34, una base de $\operatorname{Im}(A)$ está dada por las columnas 1, 2 y 4 de A :

$$\operatorname{Im}(A) = \operatorname{span} \left\{ \begin{pmatrix} 1 \\ 3 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \right\}. \quad (6.9)$$

Método alternativo para calcular la imagen de A : Podemos usar *operaciones por columnas* de A para obtener $\operatorname{Im} A$. (Se se siente más cómodo con operaciones de filas, podría aplicar operaciones por filas a la matriz A^t y entonces transponer la matriz resultante.) Obtener $(C_j$ denota las “columna

jésima de A):

$$\begin{aligned}
 A = \begin{pmatrix} 1 & 1 & 5 & 1 \\ 3 & 2 & 13 & 1 \\ 0 & 2 & 4 & -1 \\ 4 & 5 & 22 & 1 \end{pmatrix} &\xrightarrow{\substack{C_2 \rightarrow C_2 - C_1 \\ C_3 \rightarrow C_3 - 5C_1 \\ C_4 \rightarrow C_4 - C_1}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & -1 & -2 & -2 \\ 0 & 2 & 4 & -1 \\ 4 & 1 & 2 & -3 \end{pmatrix} \xrightarrow{\substack{C_3 \rightarrow C_3 - 2C_2 \\ C_4 \rightarrow C_4 - 2C_2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & -1 & 0 & 0 \\ 0 & 2 & 0 & -5 \\ 4 & 1 & 0 & -5 \end{pmatrix} \\
 &\xrightarrow{\substack{C_4 \rightarrow -1/5C_4 \\ C_1 \rightarrow C_1 + 3C_2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 3 & 2 & 0 & 1 \\ 7 & 1 & 0 & 1 \end{pmatrix} \xrightarrow{\substack{C_1 \rightarrow C_1 - 3C_4 \\ C_2 \rightarrow C_2 - 2C_4}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 4 & -1 & 0 & 1 \end{pmatrix} \\
 &\xrightarrow{\substack{C_3 \leftrightarrow C_4 \\ C_2 \rightarrow -C_2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 4 & 1 & 1 & 0 \end{pmatrix} =: \tilde{A}.
 \end{aligned}$$

Se sigue que

$$\operatorname{Im}(A) = \operatorname{Im}(\tilde{A}) = \operatorname{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}. \quad (6.9')$$

- Explique por qué funciona el método con operaciones por columnas.
- Muestre mediante un cálculo explícito que los espacios en (6.9) y (6.9') son iguales.

Ejemplo 6.37. Encuentre una base de $\operatorname{span}\{p_1, p_2, p_3, p_4\} \subseteq P_3$ y su dimensión para

$$\begin{aligned}
 p_1 &= x^3 - x^2 + 2x + 2, & p_2 &= x^3 + 2x^2 + 8x + 13, \\
 p_3 &= 3x^3 - 6x^2 - 5, & p_4 &= 5x^3 + 4x^2 + 26x - 9.
 \end{aligned}$$

Solución. Primero identificamos P_3 con $\mathbb{R}^4$ mediante $ax^3 + bx^2 + cx + d \hat{=} (a, b, c, d)^t$. Los polinomios p_1, p_2, p_3, p_4 corresponden a los vectores

$$\vec{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 2 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 8 \\ 13 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 3 \\ -6 \\ 0 \\ -5 \end{pmatrix}, \quad \vec{v}_4 = \begin{pmatrix} 5 \\ 4 \\ 26 \\ -9 \end{pmatrix}.$$

Ahora usamos la Observación 6.35 para encontrar una base de $\operatorname{span}\{v_1, v_2, v_3, v_4\}$. Para ello consideramos la matriz A cuyas columnas son los vectores $\vec{v}_1, \dots, \vec{v}_4$:

$$A = \begin{pmatrix} 1 & 1 & 3 & 5 \\ -1 & 2 & -6 & 4 \\ 2 & 8 & 0 & 26 \\ 2 & 13 & -5 & -9 \end{pmatrix}.$$

Claramente, $\text{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\} = \text{Im}(A)$, por lo tanto es suficiente encontrar una base de $\text{Im}(A)$. Aplicando transformaciones por filas a A , obtenemos

$$A = \begin{pmatrix} 1 & 1 & 3 & 5 \\ -1 & 2 & -6 & 4 \\ 2 & 8 & 0 & 26 \\ 2 & 13 & -5 & -9 \end{pmatrix} \longrightarrow \cdots \longrightarrow \begin{pmatrix} 1 & 0 & 4 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = A'.$$

Las columnas con pivotes de A' son la primera y la segunda columna; por lo tanto, por el Teorema 6.34, una base de $\text{Im}(A)$ está dada por sus primeras dos columnas, es decir, los vectores $\vec{v}_1$ y $\vec{v}_2$.

Se sigue que $\{p_1, p_2\}$ es una base de $\text{span}\{p_1, p_2, p_3, p_4\} \subseteq P_3$, por ende $\dim(\text{span}\{p_1, p_2, p_3, p_4\}) = 2$.

Observación 6.38. Usemos la abreviatura $\pi = \text{span}\{p_1, p_2, p_3, p_4\}$. El cálculo anterior muestra de hecho que cualesquiera dos vectores de p_1, p_2, p_3, p_4 forman una base de π . Para ver esto, observe que claramente cualesquiera dos de ellos son linealmente independientes, de modo que la dimensión del espacio que generan es 2. Por otro lado, este espacio generado es un subespacio de π que tiene la misma dimensión 2. Por lo tanto deben ser iguales.

Observación 6.39. Si quisiéramos completar p_1, p_2 a una base de P_3 , tenemos (al menos) las dos opciones siguientes:

- (I) Para encontrar $q_3, q_4 \in P_3$ tales que p_1, p_2, q_3, q_4 forman una base de P_3 , podemos usar el proceso de la reducción que aplicamos para encontrar A' . Suponga que E es una matriz invertible tal que $A = EA'$. Podemos encontrar tal E al recordar las operaciones de filas que transforman A en A' . Sean $\vec{e}_j$ los vectores unitarios estándar de $\mathbb{R}^4$. Entonces ya sabemos que $\vec{v}_1 = E\vec{e}_1$ y $\vec{v}_2 = E\vec{e}_2$. Si ponemos $\vec{w}_3 = E\vec{e}_3$ y $\vec{w}_4 = E\vec{e}_4$, entonces $\vec{v}_1, \vec{v}_2, \vec{w}_3, \vec{w}_4$ forman una base de $\mathbb{R}^4$. Esto se debe a que $\vec{e}_1, \dots, \vec{e}_4$ son linealmente independientes y E es inyectiva. De ahí que $E\vec{e}_1, \dots, E\vec{e}_4$ are linealmente independientes too (by Proposition 6.14).
- (II) Si ya tenemos algunos conocimientos sobre complementos ortogonales, como se discute en el Capítulo 7, entonces sabemos que cualquier base del complemento ortogonal de $\text{span}\{\vec{v}_1, \vec{v}_2\}$ los completa a una base de $\mathbb{R}^4$, que luego solo tenemos que traducir de nuevo a vectores en P_3 . Para encontrar two linealmente independientes vectors which are orthogonal to $\vec{v}_1$ y $\vec{v}_2$ tenemos que encontrar soluciones linealmente independientes del sistema homogéneo de dos ecuaciones con cuatro incógnitas

$$\begin{aligned} x_1 - x_2 + 2x_3 + 2x_4 &= 0, \\ x_1 + 2x_2 - 6x_3 + 4x_4 &= 0 \end{aligned}$$

o, en notación matricial, $P\vec{x} = 0$ donde P es la matriz 2×4 cuyas filas son $\vec{v}_1$ y $\vec{v}_2$. Como claramente $\text{Im}(P) \subseteq \mathbb{R}^2$, se sigue que $\dim(\text{Im}(P)) \leq 2$ y consecuentemente $\dim(\ker(P)) \geq 4 - 2 = 2$.

Observación 6.40. En la Sección 7 definiremos el complemento ortogonal $U^\perp$ of a subspace $U \subseteq \mathbb{R}^n$ (Definition 7.18). Consiste en todos los vectores $\vec{w}$ que son ortogonales a todo $\vec{u} \in U$. Mostraremos en el Teorema 7.22 que para toda matriz $A \in M(m \times n)$

$$\ker(A) = (R_A)^\perp.$$

Como $R(A) = \text{Im } A^t$, esto muestra la relación importante

$$\ker(A) = (\text{Im } A^t)^\perp.$$

Ejemplo 6.41. Sea $T : \mathbb{R}^5 \rightarrow \mathbb{R}^4$ dada por

$$T \begin{pmatrix} x \\ y \\ z \\ w \\ r \end{pmatrix} = \begin{pmatrix} 3x + 2y - 5r \\ 2y + z - w \\ 5x + 2y - w \\ w + 3r \end{pmatrix}.$$

Queremos escribir T en la forma $A\vec{x}$. Observe que T se puede expresar en la forma

$$T \begin{pmatrix} x \\ y \\ z \\ w \\ r \end{pmatrix} = \begin{pmatrix} 3 & 2 & 0 & 0 & -5 \\ 0 & 2 & 1 & 0 & -1 \\ 5 & 2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \\ r \end{pmatrix},$$

Esta forma de expresar T no es arbitraria, ya que si $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \in \text{Im } T$ entonces:

$$\begin{aligned} x + 2y - 5r &= y_1, \\ 2y + z - w &= y_2, \\ 5x + 2y - w &= y_3, \\ w + 3r &= y_4 \end{aligned}$$

y sabemos por la Sección 3.3 que este sistema se puede escribir en la forma:

$$\begin{pmatrix} 3 & 2 & 0 & 0 & -5 \\ 0 & 2 & 1 & 0 & -1 \\ 5 & 2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \\ r \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}.$$

Ahora debería haber entendido

- cuál es la relación entre las soluciones de un sistema homogéneo y el núcleo de la matriz de coeficientes asociada,
- cuál es la relación entre los lados derechos admisibles de un sistema de ecuaciones lineales y la imagen de la matriz de coeficientes asociada,
- por qué la fórmula de la dimensión (6.8) se cumple y por qué es solo un caso especial de (6.4),
- por qué funciona el proceso de Gauß-Jordan,

- etc.

Ahora debería ser capaz de

- calcular una base del núcleo de una matriz y su dimensión,
- calcular una base de la imagen de una matriz y su dimensión,
- etc.

Ejercicios.

- Encuentre una base para el espacio generado de los siguientes conjuntos:

(a) $\left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}.$

(b) $\left\{ \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -4 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} 10 \\ -5 \\ 5 \end{pmatrix} \right\}.$

(c) $\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ 5 \end{pmatrix} \right\}.$

(d) $\left\{ \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 2 & 0 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix} \right\}.$

(e) $\{X^3 - X^2 + X + 2, 2X^3 + 4X + 2, 2X^3 + X^2 + 5X + 1, X^3 + 2X^2 + 4X - 1\}.$

- En los siguientes ejercicios, exprese T del Ejemplo 6.41 de la forma $A\vec{x}$. Determine $\text{Im } T$, $\ker T$ y sus dimensiones.

(a) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3, T(\vec{x}) = \vec{x} \times \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}.$

(b) $T : \mathbb{R}^4 \rightarrow \mathbb{R}^3, T \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} x + y \\ x - z \\ 2y + w \end{pmatrix}.$

(c) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^5, T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y \\ 5x + 4y + 9z \\ 2x - 3y - z \\ x + z \\ 2y + 2z \end{pmatrix}.$

(d) Sea $w = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}$ y $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por $T(\vec{x}) = \text{proy}_{\vec{w}} \vec{x}.$

- Encuentre una matriz $A(3 \times 3)$ tal que su núcleo es el plano $E : x + 2y - z = 0$.

4. Encuentre una matriz $A(3 \times 3)$ tal que su imagen es el plano $E : 2x - y + 3z = 0$.
5. Sea $A \in M(m \times n)$ tal que para todo $\vec{b} \in \mathbb{R}^m$, el sistema $A\vec{x} = \vec{b}$ tiene solución. ¿Cuánto vale $\dim(\text{Im } A)$?, ¿qué se puede decir de $\ker A$?
6. Sean $A \in M(m \times n)$ y $B \in M(n \times k)$.
 - (a) Muestre que $\dim(\text{Im } AB) \leq \dim \text{Im } A$.
 - (b) Muestre que $\dim(\text{Im } AB) \leq \dim \text{Im } B$.
 - (c) Muestre que $\dim(\ker AB) \geq \dim \ker B$.
 - (d) Muestre que $\dim(\ker AB) \geq \dim \ker A$.

¿Cuándo se tiene igualdad en los ejercicios anteriores? Encuentre ejemplos para igualdad y ejemplos donde hay desigualdad estricta.
7. Sea $A \in M(n \times n)$ tal que $A^2 = A$, muestre que $\ker A \oplus \text{Im } A = \mathbb{R}^n$. (Pista: Basta mostrar que $\ker A \cap \text{Im } A = \{\vec{0}\}$ ¿Por qué?)

6.3. Cambio de bases

En esta sección, trabajamos principalmente con espacios vectoriales reales para fijar ideas. Sin embargo, todos los enunciados también son válidos para espacios vectoriales complejos. Solo tenemos que reemplazar en todas partes $\mathbb{R}$ por $\mathbb{C}$ y la palabra *real* por *complejo*.

Usualmente representamos los vectores en $\mathbb{R}^n$ como columnas de números, por ejemplo

$$\vec{v} = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix}, \quad \text{o, más generalmente,} \quad \vec{w} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \quad (6.10)$$

Tales columnas de números se interpretan usualmente como las coordenadas cartesianas del extremo del vector si su punto inicial está en el origen. Así, por ejemplo, podemos visualizar $\vec{v}$ como el vector que obtenemos cuando nos movemos 3 unidades a lo largo del eje x , 2 unidades a lo largo del eje y y -1 unidad a lo largo del eje z .

Si tomamos $\vec{e}_1, \vec{e}_2, \vec{e}_3$ como los vectores unitarios que son paralelos a los ejes x, y y z , respectivamente, entonces podemos escribir $\vec{v}$ como una suma ponderada de ellos:

$$\vec{v} = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} = 3\vec{e}_1 + 2\vec{e}_2 - \vec{e}_3. \quad (6.11)$$

Así, la columna de números que usamos para describir $\vec{v}$ en (6.10) puede verse como una manera conveniente de abreviar la suma en (6.11).

Sin embargo, a veces puede tener más sentido describir cierto vector no mediante sus coordenadas cartesianas. Por ejemplo, pensemos en un tablero de ajedrez infinitamente grande (esto es $\mathbb{R}^2$). Entonces la torre se mueve a lo largo de los ejes cartesianos, mientras que el alfil se mueve a lo largo de las diagonales, es decir, a lo largo de $\vec{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\vec{b}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ y el caballo se mueve en

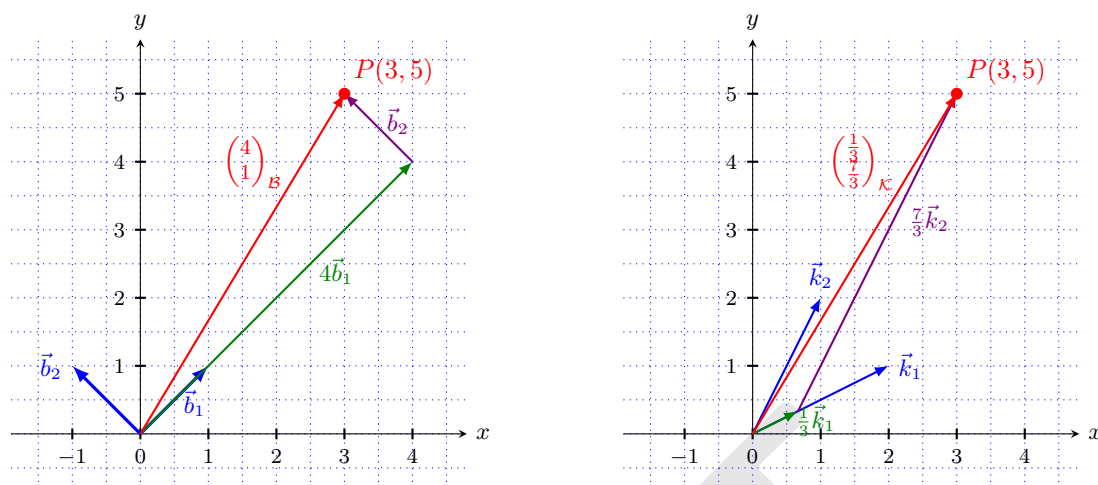


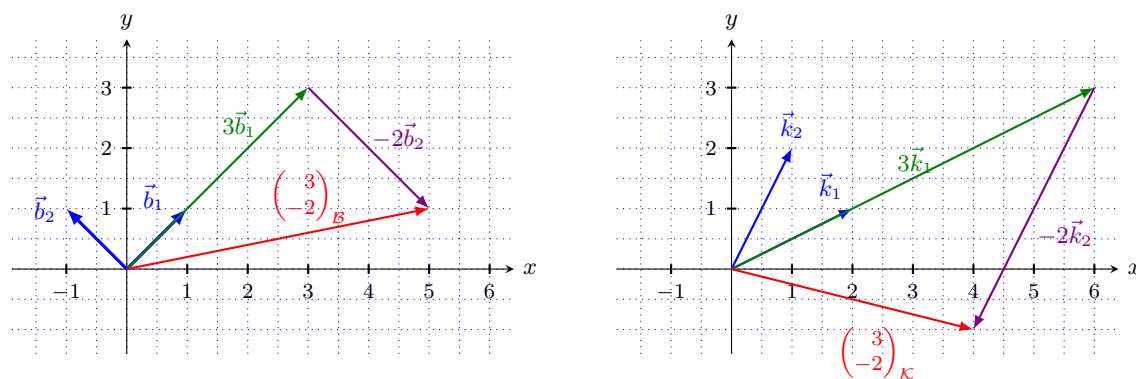
FIGURA 6.1: Las figuras muestran el punto $(3, 5)$ en coordenadas de “alfil” y de “caballo”. Los vectores para el alfil son $\vec{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\vec{b}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ y $\vec{x}_B = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$. Los vectores para el caballo son $\vec{k}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\vec{k}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}_B$ y $\vec{x}_K = \begin{pmatrix} 1/3 \\ 7/3 \end{pmatrix}_K$.

direcciones paralelas a $\vec{k}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\vec{k}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Suponemos que en nuestro juego de ajedrez imaginario la torre, el alfil y el caballo pueden moverse en múltiplos arbitrarios de sus direcciones. Supongamos que las tres piezas están situadas en el origen del tablero y queremos moverlas a la casilla $(3, 5)$. Para la torre, esto es muy fácil. Solo tiene que moverse 3 pasos a la derecha y luego 5 pasos hacia arriba. Denotaría su movimiento como $\vec{v}_R = \begin{pmatrix} 3 \\ 5 \end{pmatrix}_R$ donde ponemos el índice R para indicar que los números en este vector columna corresponden al sistema de coordenadas natural de la torre. El alfil no puede hacer esto. Solo puede moverse a lo largo de las diagonales. Entonces, ¿qué tiene que hacer? Tiene que moverse 4 pasos en la dirección de $\vec{b}_1$ y 1 paso en la dirección de $\vec{b}_2$. Así, denotaría su movimiento con respecto a su sistema de coordenadas de alfil como $\vec{v}_B = \begin{pmatrix} 4 \\ 1 \end{pmatrix}_B$. Finalmente, el caballo tiene que moverse $\frac{1}{3}$ pasos en la dirección de $\vec{k}_1$ y $\frac{7}{3}$ pasos en la dirección de $\vec{k}_2$ para alcanzar el punto $(3, 5)$. Así, denotaría su movimiento con respecto a su sistema de coordenadas de caballo como $\vec{v}_K = \begin{pmatrix} 1/3 \\ 7/3 \end{pmatrix}_K$. Véase la Figura 6.1.

Ejercicio. Verifique que $\vec{v}_B = \begin{pmatrix} 4 \\ 1 \end{pmatrix}_B = 4\vec{b}_1 + 1\vec{b}_2 = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ y que $\vec{v}_K = \begin{pmatrix} 1/3 \\ 7/3 \end{pmatrix}_K = 1/3\vec{k}_1 + 7/3\vec{k}_2 = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$.

Aunque los tres vectores $\vec{v}$, $\vec{v}_B$ y $\vec{v}_K$ se ven muy diferentes, describen el mismo vector, solo desde tres perspectivas distintas (la perspectiva de la torre, del alfil y del caballo). Debemos recordar que tienen que interpretarse como combinaciones lineales de los vectores que describen sus movimientos. Lo que acabamos de hacer fue realizar un cambio de bases en $\mathbb{R}^2$: en lugar de describir un punto del plano en coordenadas cartesianas, usamos coordenadas de “alfil” y de “caballo”.

También podemos ir en la otra dirección y transformar de coordenadas de “alfil” o de “caballo” a coordenadas cartesianas. Supongamos que sabemos que el alfil se mueve 3 pasos en su dirección $\vec{b}_1$ y -2 pasos en su dirección $\vec{b}_2$; ¿dónde termina? En su sistema de coordenadas, se desplaza mediante

FIGURA 6.2: Las figuras muestran los vectores $\begin{pmatrix} 3 \\ -2 \end{pmatrix}_{\mathcal{B}}$ y $\begin{pmatrix} 3 \\ -2 \end{pmatrix}_{\mathcal{K}}$.

el vector $\vec{u} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}_{\mathcal{B}}$. En coordenadas cartesianas este vector es

$$\vec{u} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}_{\mathcal{B}} = 3\vec{b}_1 - 2\vec{b}_2 = \begin{pmatrix} 3 \\ 3 \end{pmatrix} + \begin{pmatrix} 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}.$$

Si movemos el caballo 3 pasos en su dirección $\vec{k}_1$ y -2 pasos en su dirección $\vec{k}_2$, es decir, si lo movemos a lo largo de $\vec{w} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}_{\mathcal{K}}$ según su sistema de coordenadas, entonces, en coordenadas cartesianas, este vector es

$$\vec{w} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}_{\mathcal{K}} = 3\vec{k}_1 - 2\vec{k}_2 = \begin{pmatrix} 6 \\ 3 \end{pmatrix} + \begin{pmatrix} -2 \\ -4 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}.$$

¿Pueden el alfil y el caballo alcanzar todos los puntos del plano? Si es así, ¿de cuántas maneras? La respuesta es sí, y pueden hacerlo exactamente de una manera. La razón es que, tanto para el alfil como para el caballo, cada conjunto de vectores de dirección forma una base de $\mathbb{R}^2$ (¡verifíquelo!).

Precisemos el concepto de cambio de base. Supongamos que se nos da una base ordenada $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ de $\mathbb{R}^n$. Si escribimos

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}} \quad (6.12)$$

entonces lo interpretamos como un vector expresado con respecto a la base $\mathcal{B}$ y

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}} := x_1\vec{b}_1 + \dots + x_n\vec{b}_n. \quad (6.13)$$

Si el vector columna no lleva ningún índice, entonces lo interpretamos como un vector con respecto a la base canónica $\vec{e}_1, \dots, \vec{e}_n$ de $\mathbb{R}^n$. Ahora queremos encontrar una manera de calcular las coordenadas cartesianas (es decir, aquellas con respecto a la base canónica) si se nos da un vector en coordenadas respecto a $\mathcal{B}$, y viceversa.

Resultará que la siguiente matriz será muy útil:

$$A_{\mathcal{B} \rightarrow \text{can}} = (\vec{v}_1 | \dots | \vec{v}_n) = \text{matriz cuyas columnas son los vectores de la base } \mathcal{B}.$$

Explicaremos el índice “ $\mathcal{B} \rightarrow \text{can}$ ” en un momento.

Transición de la representación con respecto a una base dada a coordenadas cartesianas.

Supongamos que se nos da un vector como en (6.13). ¿Cómo obtenemos sus coordenadas cartesianas?

Esto es bastante directo. Solo necesitamos recordar qué significa la notación $(\cdot)_{\mathcal{B}}$. Denotaremos por $\vec{x}_{\mathcal{B}}$ la representación del vector con respecto a la base $\mathcal{B}$ y por $\vec{x}$ su representación con respecto a la base estándar de $\mathbb{R}^n$.

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}} = x_1 \vec{b}_1 + x_2 \vec{b}_2 + \dots + x_n \vec{b}_n = (\vec{b}_1 | \vec{b}_2 | \dots | \vec{b}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A_{\mathcal{B} \rightarrow \text{can}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A_{\mathcal{B} \rightarrow \text{can}} \vec{x}_{\mathcal{B}},$$

es decir

$$\vec{x} = A_{\mathcal{B} \rightarrow \text{can}} \vec{x}_{\mathcal{B}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{\text{can}}. \quad (6.14)$$

El último vector (el que contiene $y_1, \dots, y_n$) describe el mismo vector que $\vec{x}_{\mathcal{B}}$, pero lo hace con respecto a la base estándar de $\mathbb{R}^n$. La matriz $A_{\mathcal{B} \rightarrow \text{can}}$ se llama la *matriz de transición de la base $\mathcal{B}$ a la base canónica* (lo cual explica el subíndice “ $\mathcal{B} \rightarrow \text{can}$ ”). La matriz también se llama la *matriz de cambio de coordenadas*.

Transición de coordenadas cartesianas a la representación con respecto a una base dada.

Supongamos que se nos da un vector $\vec{x}$ en coordenadas cartesianas. ¿Cómo calculamos sus coordenadas $\vec{x}_{\mathcal{B}}$ con respecto a la base $\mathcal{B}$?

Solo necesitamos recordar que la relación entre $\vec{x}$ y $\vec{x}_{\mathcal{B}}$, de acuerdo con (6.14), es

$$\vec{x} = A_{\mathcal{B} \rightarrow \text{can}} \vec{x}_{\mathcal{B}}.$$

En este caso, conocemos las entradas del vector $\vec{x}$. Así que solo necesitamos invertir la matriz $A_{\mathcal{B} \rightarrow \text{can}}$ para obtener las entradas de $\vec{x}_{\mathcal{B}}$:

$$\vec{x}_{\mathcal{B}} = A_{\mathcal{B} \rightarrow \text{can}}^{-1} \vec{x}.$$

Esto requiere, por supuesto, que $A_{\mathcal{B} \rightarrow \text{can}}$ sea invertible. Pero esto está garantizado por el Teorema 5.39, ya que sabemos que sus columnas son linealmente independientes. Así se sigue que la matriz de transición de la base canónica a la base $\mathcal{B}$ está dada por

$$A_{\text{can} \rightarrow \mathcal{B}} = A_{\mathcal{B} \rightarrow \text{can}}^{-1}.$$

Observe que también podríamos hacer esto “a mano”: Se nos da $\vec{x} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{can}$ y queremos encontrar las entradas $x_1, \dots, x_n$ del vector $\vec{x}_{\mathcal{B}}$ que describe el mismo vector. Es decir, necesitamos números $x_1, \dots, x_n$ tales que

$$\vec{x} = x_1 \vec{b}_1 + \dots + \vec{b}_n x_n.$$

Si conocemos los vectores $\vec{b}_1, \dots, \vec{b}_n$, entonces podemos escribir esto como un sistema de ecuaciones lineales de tamaño $n \times n$ y luego resolverlo para $x_1, \dots, x_n$, lo cual, por supuesto, en realidad es lo

mismo que aplicar la inversa de la matriz $A_{\mathcal{B} \rightarrow can}$ al vector $\vec{x} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{can}$.

Ahora supongamos que tenemos dos bases ordenadas $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ y $\mathcal{C} = \{\vec{c}_1, \dots, \vec{c}_n\}$ de $\mathbb{R}^n$ y que se nos da un vector $\vec{x}_{\mathcal{B}}$ con respecto a la base $\mathcal{B}$. ¿Cómo podemos calcular su representación $\vec{x}_{\mathcal{C}}$ con respecto a la base $\mathcal{C}$? La manera más sencilla es usar la base canónica de $\mathbb{R}^n$ como base auxiliar. Entonces primero calculamos el vector dado $\vec{x}_{\mathcal{B}}$ con respecto a la base canónica; llamamos a este vector $\vec{x}$. Luego pasamos de $\vec{x}$ a $\vec{x}_{\mathcal{C}}$. De acuerdo con las fórmulas anteriores, esto es

$$\vec{x}_{\mathcal{C}} = \vec{A}_{can \rightarrow \mathcal{C}} \vec{x} = A_{can \rightarrow \mathcal{C}} A_{\mathcal{B} \rightarrow can} \vec{x}_{\mathcal{B}}.$$

Por lo tanto, la matriz de transición de la base $\mathcal{B}$ a la base $\mathcal{C}$ es

$$A_{\mathcal{B} \rightarrow \mathcal{C}} = A_{can \rightarrow \mathcal{C}} A_{\mathcal{B} \rightarrow can}.$$

Ejemplo 6.42. Volvamos a nuestro ejemplo del tablero de ajedrez imaginario. Tenemos la “base de alfil” $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ donde $\vec{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\vec{b}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ y la “base de caballo” $\mathcal{K} = \{\vec{k}_1, \vec{k}_2\}$ $\vec{k}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\vec{k}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Entonces las matrices de transición a la base canónica son

$$A_{\mathcal{B} \rightarrow can} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad A_{\mathcal{K} \rightarrow can} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

sus inversas son

$$A_{can \rightarrow \mathcal{B}} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad A_{can \rightarrow \mathcal{K}} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

y las matrices de transición de $\mathcal{B}$ a $\mathcal{K}$ y de $\mathcal{K}$ a $\mathcal{B}$ son

$$A_{\mathcal{B} \rightarrow \mathcal{K}} = \frac{1}{3} \begin{pmatrix} 1 & -3 \\ 1 & 3 \end{pmatrix}, \quad A_{\mathcal{K} \rightarrow \mathcal{B}} = \frac{1}{2} \begin{pmatrix} 1 & 3 \\ -1 & 3 \end{pmatrix}.$$

- Dado un vector $\vec{x} = \begin{pmatrix} 2 \\ 7 \end{pmatrix}_{\mathcal{B}}$ en coordenadas de alfil, ¿cuáles son sus coordenadas de caballo?

Solución. $(\vec{x})_{\mathcal{K}} = A_{\mathcal{B} \rightarrow \mathcal{K}} \vec{x}_{\mathcal{B}} = \frac{1}{3} \begin{pmatrix} 1 & -3 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 7 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -19 \\ 23 \end{pmatrix}_{\mathcal{K}}.$

◇

- Dado un vector $\vec{y} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}_{\mathcal{K}}$ en coordenadas de caballo, ¿cuáles son sus coordenadas de alfil?

Solución. $(\vec{y})_{\mathcal{B}} = A_{\mathcal{K} \rightarrow \mathcal{B}} \vec{y}_{\mathcal{K}} = \frac{1}{2} \begin{pmatrix} 1 & 3 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}_{\mathcal{B}}.$

◇

- Dado un vector $\vec{z} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ en coordenadas estándar, ¿cuáles son sus coordenadas de alfí?

Solución. $(\vec{z})_{\mathcal{B}} = A_{can \rightarrow \mathcal{B}} \vec{z} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}_{\mathcal{B}}$.

◇

Ejemplo 6.43. Recordemos el ejemplo de la página 109, donde teníamos una tienda que vendía distintos tipos de paquetes de comida. El paquete de tipo A contiene 1 durazno y 3 mangos, y el paquete de tipo B contiene 2 duraznos y 1 mango. Planteamos dos tipos de preguntas:

Pregunta 1. Si compramos a paquetes de tipo A y b paquetes de tipo B, ¿cuántos duraznos y mangos obtendremos? Podríamos reformular esta pregunta para que sea más similar a la Pregunta 2: ¿cuántos duraznos y mangos necesitamos para llenar a paquetes de tipo A y b paquetes de tipo B?

Pregunta 2. ¿Cuántos paquetes de tipo A y de tipo B tenemos que comprar para obtener p duraznos y m mangos?

Recordemos que teníamos la relación

$$M \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} m \\ p \end{pmatrix}, \quad M^{-1} \begin{pmatrix} m \\ p \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{donde} \quad M = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \quad \text{y} \quad M^{-1} = \frac{1}{5} \begin{pmatrix} -1 & 2 \\ 3 & -1 \end{pmatrix}. \quad (6.15)$$

Podemos ver estos problemas en dos sistemas de coordenadas distintos. Tenemos la “base de frutas” $\mathcal{F} = \{\vec{p}, \vec{m}\}$ y la “base de paquetes” $\mathcal{P} = \{\vec{A}, \vec{B}\}$ donde

$$\vec{m} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{p} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \vec{A} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad \vec{B} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Observe que $\vec{A} = \vec{m} + 3\vec{p}$, $\vec{B} = 2\vec{m} + \vec{p}$, y que $\vec{m} = \frac{1}{5}(-\vec{A} + 3\vec{B})$ y $\vec{p} = \frac{1}{5}(2\vec{A} - \vec{B})$ (esto significa, por ejemplo, que un mango es tres quintos de un paquete B menos un quinto de un paquete A).

Un ejemplo para la primera pregunta es: ¿Cuántos duraznos y mangos necesitamos para obtener 1 paquete de tipo A y 3 paquetes de tipo B? Claramente, necesitamos 7 duraznos y 6 mangos. Así, el punto que queremos alcanzar está en “coordenadas de paquetes” $\begin{pmatrix} 1 \\ 3 \end{pmatrix}_{\mathcal{P}}$ y en “coordenadas de frutas” $\begin{pmatrix} 7 \\ 6 \end{pmatrix}_{\mathcal{F}}$. Esto se esquematiza en la Figura 6.3.

Un ejemplo para la segunda pregunta es: ¿cuántos paquetes de tipo A y de tipo B tenemos que comprar para obtener 5 duraznos y 5 mangos? Usando (6.15), encontramos que necesitamos 1 paquete de tipo A y 3 paquetes de tipo B. Así, el punto que queremos alcanzar está en “coordenadas de paquetes” $\begin{pmatrix} 1 \\ 3 \end{pmatrix}_{\mathcal{P}}$ y en “coordenadas de frutas” $\begin{pmatrix} 5 \\ 5 \end{pmatrix}_{\mathcal{F}}$. Esto se esquematiza en la Figura 6.4.

En el resto de esta sección aplicaremos estas ideas para introducir coordenadas en espacios vectoriales abstractos (finitamente generados) V con respecto a una base dada. Esto nos permite identificar, en cierto sentido, V con $\mathbb{R}^n$ o $\mathbb{C}^n$ para un n apropiado.

Supongamos que se nos da un espacio vectorial real V con una base ordenada $\mathcal{B} = \{v_1, \dots, v_n\}$. Dado un vector $w \in V$, sabemos que existen números reales determinados de manera única $\alpha_1, \dots, \alpha_n$ tales que

$$w = \alpha_1 v_1 + \dots + \alpha_n v_n.$$

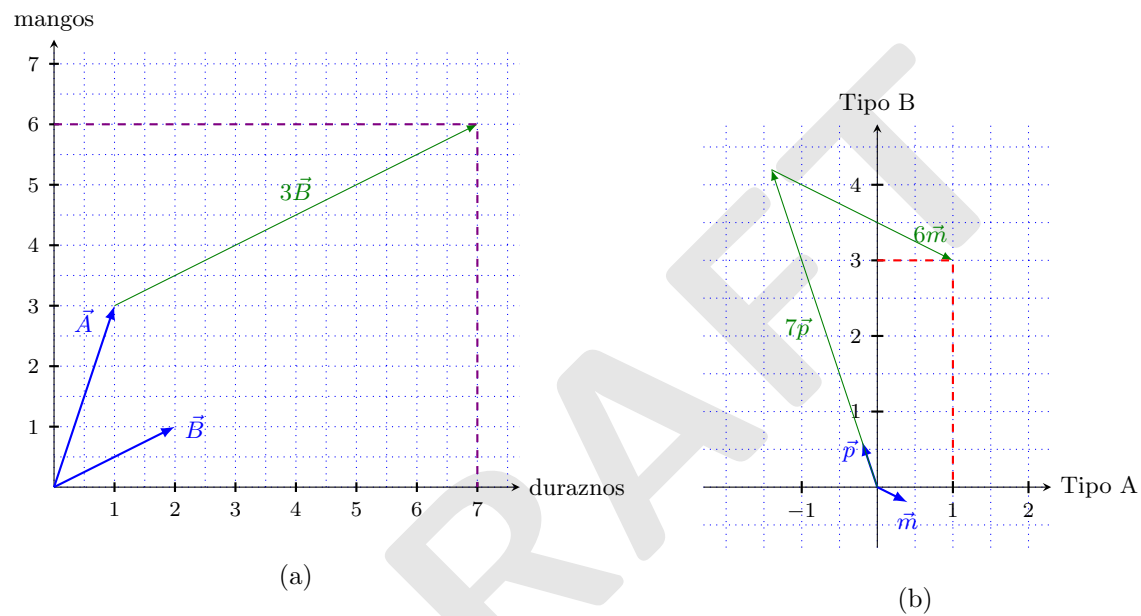


FIGURA 6.3: ¿Cuántos duraznos y mangos necesitamos para obtener 1 paquete de tipo A y 3 paquetes de tipo B? Respuesta: 7 duraznos y 6 mangos. La Figura (a) describe la situación en el “plano de frutas”, mientras que la Figura (b) describe la misma situación en el “plano de paquetes”. En ambas figuras vemos que $\vec{A} + 3\vec{B} = 7\vec{p} + 6\vec{m}$.

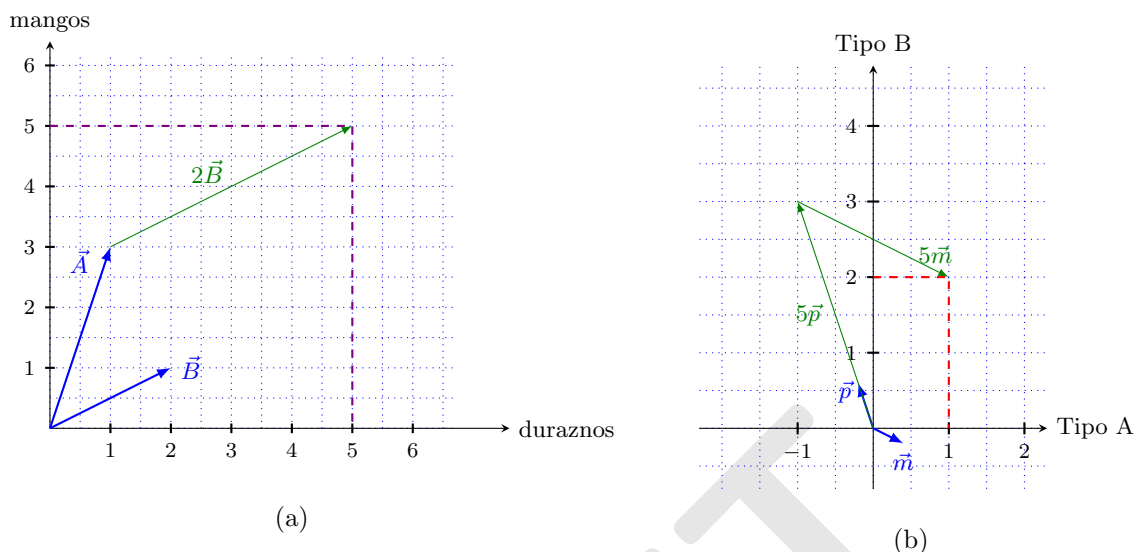


FIGURA 6.4: ¿Cuántos paquetes de tipo A y de tipo B necesitamos para obtener 5 duraznos y 5 mangos? Respuesta: 1 paquete de tipo A y 2 paquetes de tipo B. La Figura (a) describe la situación en el “plano de frutas”, mientras que la Figura (b) describe la misma situación en el “plano de paquetes”. En ambas figuras vemos que $\vec{A} + 2\vec{B} = 5\vec{p} + 5\vec{m}$.

Así, si se nos da w , podemos encontrar los números $\alpha_1, \dots, \alpha_n$. Por otro lado, si se nos dan los números $\alpha_1, \dots, \alpha_n$, podemos reconstruir fácilmente el vector w (basta sustituir en el lado derecho de la ecuación anterior). Por lo tanto, tiene sentido escribir

$$w = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}_{\mathcal{B}}$$

donde nuevamente el índice $\mathcal{B}$ nos recuerda que la columna de números debe entenderse como los coeficientes con respecto a la base $\mathcal{B}$. De esta manera, identificamos V con $\mathbb{R}^n$, ya que cada vector columna da un vector w en V y cada vector w da un vector columna en $\mathbb{R}^n$. Observe que si comenzamos con algún w en V , calculamos sus coordenadas con respecto a una base dada y luego regresamos a V , recuperamos nuestro vector original w .

Ejemplo 6.44. En P_2 , considere las bases $\mathcal{B} = \{p_1, p_2, p_3\}$, $\mathcal{C} = \{q_1, q_2, q_3\}$, $\mathcal{D} = \{r_1, r_2, r_3\}$ donde

$$p_1 = 1, p_2 = X, p_3 = X^2, \quad q_1 = X^2, q_2 = X, q_3 = 1, \quad r_1 = X^2 + 2X, r_2 = 5X + 2, r_3 = 1.$$

Queremos escribir el polinomio $\pi(X) = aX^2 + bX + c$ con respecto a la base dada.

- Base $\mathcal{B}$: Claramente, $\pi = cp_1 + bp_2 + ap_3$, por lo tanto $\pi = \begin{pmatrix} c \\ b \\ a \end{pmatrix}_{\mathcal{B}}$.

- Base $\mathcal{C}$: Claramente, $\pi = aq_1 + bq_2 + cq_3$, por lo tanto $\pi = \begin{pmatrix} a \\ b \\ c \end{pmatrix}_{\mathcal{C}}$.
- Base $\mathcal{D}$: Esto requiere algunos cálculos. Recordemos que necesitamos números $\alpha, \beta, \gamma \in \mathbb{R}$ tales que

$$\pi = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}_{\mathcal{D}} = \alpha r_1 + \beta r_2 + \gamma r_3.$$

Esto conduce a la siguiente ecuación

$$aX^2 + bX + c = \alpha(X^2 + 2X) + \beta(5X + 2) + \gamma = \alpha X^2 + (2\alpha + 5\beta)X + 2\beta + \gamma.$$

Comparando coeficientes obtenemos

$$\left. \begin{array}{l} \alpha = a, \\ 2\alpha + 5\beta = b, \\ 2\beta + \gamma = c \end{array} \right\} \quad \text{en forma matricial:} \quad \begin{pmatrix} 1 & 0 & 0 \\ 2 & 5 & 0 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}. \quad (6.16)$$

Note que las columnas de la matriz en el lado derecho son exactamente las representaciones vectoriales de r_1, r_2, r_3 con respecto a la base $\mathcal{C}$ y el vector columna $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ es exactamente la representación vectorial de π con respecto a la base $\mathcal{C}$. La solución del sistema es

$$\alpha = a, \quad \beta = -\frac{2}{5}a + \frac{1}{5}b, \quad \gamma = \frac{2}{5}a - \frac{1}{5}b + c,$$

por lo tanto

$$\pi = \begin{pmatrix} a \\ -\frac{2}{5}a + \frac{1}{5}b \\ \frac{2}{5}a - \frac{1}{5}b + c \end{pmatrix}_{\mathcal{D}}.$$

También podríamos haber encontrado la solución haciendo una vuelta por $\mathbb{R}^3$ de la siguiente manera: Identificamos los vectores q_1, q_2, q_3 con los vectores de la base canónica $\vec{e}_1, \vec{e}_2, \vec{e}_3$ de $\mathbb{R}^3$. Entonces los vectores r_1, r_2, r_3 y π corresponden a

$$\vec{r}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \vec{r}_2 = \begin{pmatrix} 0 \\ 5 \\ 2 \end{pmatrix}, \quad \vec{r}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{\pi}' = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Sea $R = \{\vec{r}_1, \vec{r}_2, \vec{r}_3\}$. Para encontrar las coordenadas de $\vec{\pi}'$ con respecto a la base $\vec{r}_1, \vec{r}_2, \vec{r}_3$, observamos que

$$\vec{\pi}' = A_{R \rightarrow \text{can}} \vec{\pi}'_R$$

donde $A_{R \rightarrow \text{can}}$ es la matriz de transición de la base R a la base canónica de $\mathbb{R}^3$, cuyas columnas consisten en los vectores $\vec{r}_1, \vec{r}_2, \vec{r}_3$. Así vemos que esta es exactamente la misma ecuación que la de (6.16).

Damos un ejemplo en un espacio de matrices.

Ejemplo 6.45. Considere las matrices

$$R = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \quad T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 2 & 3 \\ 3 & 0 \end{pmatrix}.$$

(I) Muestre que $\mathcal{B} = \{R, S, T\}$ es una base de $M_{sym}(2 \times 2)$ (el espacio de todas las matrices simétricas de tamaño 2×2).

(II) Escriba Z en términos de la base $\mathcal{B}$.

Solución. (I) Claramente, $R, S, T \in M_{sym}(2 \times 2)$. Como ya sabemos que $\dim M_{sym}(2 \times 2) = 3$, basta mostrar que R, S, T son linealmente independientes. Entonces consideremos la ecuación

$$0 = \alpha R + \beta S + \gamma T = \begin{pmatrix} \alpha + \beta & \alpha + \gamma \\ \alpha + \gamma & \alpha + 3\beta \end{pmatrix}.$$

Obtenemos el sistema de ecuaciones

$$\left. \begin{array}{l} \alpha + \beta = 0, \\ \alpha + \gamma = 0, \\ \alpha + 3\beta = 0 \end{array} \right\} \quad \text{en forma matricial:} \quad \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 3 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (6.17)$$

Haciendo algunos cálculos, se sigue que $\alpha = \beta = \gamma = 0$. De ahí mostramos que R, S, T son linealmente independientes y por lo tanto son una base de $M_{sym}(2 \times 2)$.

(II) Para escribir Z en términos de la base $\mathcal{B}$, necesitamos encontrar $\alpha, \beta, \gamma \in \mathbb{R}$ tales que

$$Z = \alpha R + \beta S + \gamma T = \begin{pmatrix} \alpha + \beta & \alpha + \gamma \\ \alpha + \gamma & \alpha + 3\beta \end{pmatrix}.$$

Obtenemos el sistema de ecuaciones

$$\left. \begin{array}{l} \alpha + \beta = 2, \\ \alpha + \gamma = 3, \\ \alpha + 3\beta = 0 \end{array} \right\} \quad \text{en forma matricial:} \quad \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 3 & 0 \end{pmatrix}}_{=A} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}. \quad (6.18)$$

Por lo tanto

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = A^{-1} \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 3 & 0 & -1 \\ -1 & 0 & 1 \\ -3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix},$$

$$\text{de donde } Z = 3R - S = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}_{\mathcal{B}}. \quad \diamond$$

Ahora damos una solución alternativa (que es esencialmente la misma que la anterior) haciendo un rodeo por $\mathbb{R}^3$. Sea $\mathcal{C} = \{A_1, A_2, A_3\}$ donde $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, $A_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Esta es claramente una base de $M_{sym}(2 \times 2)$. La identificamos con la base estándar $\vec{e}_1, \vec{e}_2, \vec{e}_3$ de $\mathbb{R}^3$. Entonces los vectores R, S, T en esta base se ven como

$$R' = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad S' = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}, \quad T' = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{y} \quad Z' = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}.$$

- (I) Para mostrar que R, S, T son linealmente independientes, solo tenemos que mostrar que los vectores R', S' y T' son linealmente independientes en $\mathbb{R}^3$. Para ello, consideramos la matriz A cuyas columnas son estos vectores. Observe que esta es la misma matriz que apareció en (6.18). Es fácil mostrar que esta matriz es invertible (¡ya calculamos su inversa!). Por lo tanto los vectores R', S', T' son linealmente independientes en $\mathbb{R}^3$, por lo tanto R, S, T son linealmente independientes en $M_{sym}(2 \times 2)$.
- (II) Ahora, para encontrar la representación de Z en términos de la base $\mathcal{B}$, solo necesitamos encontrar la representación de Z' en términos de la base $\mathcal{B}' = \{R', S', T'\}$. Esto se hace de la siguiente manera:

$$Z'_{\mathcal{B}'} = A_{can \rightarrow \mathcal{B}'} Z' = A^{-1} Z' = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}.$$

Ahora debería haber entendido

- el significado geométrico de un cambio de bases en $\mathbb{R}^n$,
- cómo un espacio vectorial abstracto de dimensión finita puede representarse como $\mathbb{R}^n$ o $\mathbb{C}^n$ y que la representación depende de la base elegida de V ,
- cómo cambia la representación vectorial si se reordena la base elegida,
- etc.

Ahora debería ser capaz de

- realizar un cambio de base en $\mathbb{R}^n$ y $\mathbb{C}^n$ dada una base,
- representar vectores en un espacio vectorial de dimensión finita V como vectores columna después de elegir una base,
- etc.

Ejercicios.

1. Sea $\mathcal{B} = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \right\}$. Muestre que $\mathcal{B}$ es una base de $M(2 \times 2)$ y encuentre $[A]_{\mathcal{B}}$ para $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.
2. Sea $\mathcal{B} = \{X^2 - 1, X^2 + X + 1, X^2\}$. Muestre que $\mathcal{B}$ es base de P_2 . Encuentre $[p(X)]_{\mathcal{B}}$ para $p(X) = a + bX + cX^2$.
3. Sea $\mathcal{B} = \{1, e^x, e^{-x}\}$ y $V = \text{span}\{1, e^x, e^{-x}\}$.
 - (a) Muestre que $\sinh x, \cosh x \in V$.

(b) Encuentre $A \in M(3 \times 3)$ tal que

$$\begin{aligned} 1 &= a_{11} + a_{12}e^x + a_{13}e^{-x}, \\ \sinh x &= a_{21} + a_{22}e^x + a_{23}e^{-x}, \\ \cosh x &= a_{31} + a_{32}e^x + a_{33}e^{-x}. \end{aligned}$$

(c) Muestre que $\mathcal{B}' = \{1, \sinh x, \cosh x\}$ es base de V .

(d) Encuentre $A_{\mathcal{B} \rightarrow \mathcal{B}'}$ y $A_{\mathcal{B}' \rightarrow \mathcal{B}}$.

4. Muestre que $\mathcal{B} = \{1, X - 1, (X - 1)^2\}$ es base de P_2 y escriba $a + bX + cX^2$ en términos de $\mathcal{B}$. Aún más general, muestre que $\mathcal{B} = \{1, X - 1, (X - 1)^2, \dots, (X - 1)^n\}$ es base de P_n y obtenga $A_{\mathcal{B} \rightarrow \text{can}}$, $A_{\text{can} \rightarrow \mathcal{B}}$ donde $\text{can} = \{1, X^2, \dots, X^n\}$.

5. Sean $\mathcal{B}_1 = \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$ y $\mathcal{B}_2 = \left\{ \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 4 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} \right\}$.

(a) Muestre que $\mathcal{B}_1$ y $\mathcal{B}_2$ son bases de $\mathbb{R}^3$.

(b) Sea $\vec{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. Encuentre $[\vec{v}]_{\mathcal{B}_1}$ y $[\vec{v}]_{\mathcal{B}_2}$.

(c) Obtenga $A_{\mathcal{B}_1 \rightarrow \mathcal{B}_2}$ y $A_{\mathcal{B}_2 \rightarrow \mathcal{B}_1}$.

6. Sea $\vartheta \in (-\pi, \pi]$. Muestre que $\mathcal{B}_\vartheta = \left\{ \begin{pmatrix} \cos \vartheta \\ \sin \vartheta \end{pmatrix}, \begin{pmatrix} -\sin \vartheta \\ \cos \vartheta \end{pmatrix} \right\}$ es una base de $\mathbb{R}^2$. Encuentre $A_{\mathcal{B}_\vartheta \rightarrow \text{can}}$ y $A_{\text{can} \rightarrow \mathcal{B}_\vartheta}$. ¿Cómo se interpreta geoméricamente $\mathcal{B}_\vartheta$?

7. Sean a, b tal que $ab \neq 0$.

(a) Muestre que $\mathcal{B} = \left\{ \frac{1}{\sqrt{a^2+b^2}} \begin{pmatrix} a \\ b \end{pmatrix}, \frac{1}{\sqrt{a^2+b^2}} \begin{pmatrix} -b \\ a \end{pmatrix} \right\}$ es base de $\mathbb{R}^2$.

(b) Muestre que existe $\vartheta \in (-\pi, \pi]$ tal que $\mathcal{B} = \mathcal{B}_\vartheta$. (Hint: Interpretación geométrica de $\mathcal{B}$)

8. Sea $\mathcal{B}_\vartheta$ como en Ejercicio 6..

(a) Si $\vartheta = \frac{\pi}{6}$, escriba $\begin{pmatrix} -3\sqrt{3} \\ -3 \end{pmatrix}$ en términos de $\mathcal{B}_\vartheta$.

(b) Si $\vartheta = \frac{\pi}{4}$, escriba $\begin{pmatrix} 1 \\ -1 \end{pmatrix}_{\mathcal{B}_\vartheta}$ en términos de la base canónica.

9. Sean $\vartheta_1, \vartheta_2 \in (-\pi, \pi]$, ¿cómo se interpreta geoméricamente $A_{\mathcal{B}_{\vartheta_1} \rightarrow \mathcal{B}_{\vartheta_2}}$?

6.4. Transformaciones lineales y sus representaciones matriciales

Sean U, V espacios vectoriales sobre $\mathbb{K}$ y sea $T : U \rightarrow V$ una transformación lineal. Recordemos que T satisface

$$T(\lambda_1 x_1 + \cdots + \lambda_k x_k) = \lambda_1 T(x_1) + \cdots + \lambda_k T(x_k)$$

para todo $x_1, \dots, x_k \in U$ y $\lambda_1, \dots, \lambda_k \in \mathbb{K}$. Esto muestra que, para conocer T , en realidad basta conocer cómo actúa T sobre una base de U . Supongamos que se nos da una base $\mathcal{B} = \{u_1, \dots, u_n\}$ de U y tomemos un vector arbitrario $w \in U$. Entonces existen $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ determinados de manera única tales que $w = \lambda_1 u_1 + \cdots + \lambda_n u_n$. Por lo tanto

$$Tw = T(\lambda_1 u_1 + \cdots + \lambda_n u_n) = \lambda_1 Tu_1 + \cdots + \lambda_n Tu_n. \quad (6.19)$$

Así, Tw es una combinación lineal de los vectores $Tu_1, \dots, Tu_n \in V$ y los coeficientes son exactamente $\lambda_1, \dots, \lambda_n$.

Supongamos que se nos da una base $\mathcal{C} = \{v_1, \dots, v_m\}$ de V . Entonces sabemos que, para cada $j = 1, \dots, n$, el vector Tu_j es una combinación lineal de los vectores de la base $v_1, \dots, v_m$ de V . Por lo tanto, existen números $a_{ij} \in K$ determinados de manera única ($i = 1, \dots, m, j = 1, \dots, n$) tales que $Tu_j = a_{j1}v_1 + \cdots + a_{jm}v_m$, es decir

$$\begin{aligned} Tu_1 &= a_{11}v_1 + a_{21}v_2 + \cdots + a_{m1}v_m, \\ Tu_2 &= a_{12}v_1 + a_{22}v_2 + \cdots + a_{m2}v_m, \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ Tu_n &= a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{mn}v_m. \end{aligned} \quad (6.20)$$

Definamos la matriz A_T y el vector $\vec{\lambda}$ por

$$A_T = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in M(m \times n), \quad \vec{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{R}^n.$$

Observe que la primera columna de A_T es la representación vectorial de Tu_1 con respecto a la base $v_1, \dots, v_m$, la segunda columna es la representación vectorial de Tu_2 , y así sucesivamente.

Ahora volvamos al cálculo de Tw y a su conexión con la matriz A_T . De (6.19) y (6.20) obtenemos

$$\begin{aligned}
 Tw &= \lambda_1 Tu_1 + \lambda_2 Tu_2 + \cdots + \lambda_n Tu_n \\
 &= \lambda_1(a_{11}v_1 + a_{21}v_2 + \cdots + a_{m1}v_m) \\
 &\quad + \lambda_2(a_{12}v_1 + a_{22}v_2 + \cdots + a_{m2}v_m) \\
 &\quad + \cdots \cdots \cdots \\
 &\quad + \lambda_n(a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{mn}v_m) \\
 &= (a_{11}\lambda_1 + a_{12}\lambda_2 + \cdots + a_{1n}\lambda_n)v_1 \\
 &\quad + (a_{21}\lambda_1 + a_{22}\lambda_2 + \cdots + a_{2n}\lambda_n)v_2 \\
 &\quad + \cdots \cdots \cdots \\
 &\quad + (a_{m1}\lambda_1 + a_{m2}\lambda_2 + \cdots + a_{mn}\lambda_n)v_m.
 \end{aligned}$$

El cálculo muestra que, para cada k , el coeficiente de v_k es la k -ésima componente del vector $A_T \vec{\lambda}$! Ahora podemos ir un paso más allá. Recordemos que la elección de la base $\mathcal{B}$ de U y de la base $\mathcal{C}$ de V nos permite escribir w y Tw como vectores columna:

$$w = \vec{w}_{\mathcal{B}} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}_{\mathcal{B}}, \quad Tw = \begin{pmatrix} a_{11}\lambda_1 + a_{12}\lambda_2 + \cdots + a_{1n}\lambda_n \\ a_{21}\lambda_1 + a_{22}\lambda_2 + \cdots + a_{2n}\lambda_n \\ \vdots \\ a_{m1}\lambda_1 + a_{m2}\lambda_2 + \cdots + a_{mn}\lambda_n \end{pmatrix}_{\mathcal{C}}.$$

Esto muestra que

$$(Tw)_{\mathcal{C}} = A_T \vec{w}_{\mathcal{B}}.$$

Por razones que ahora esperamos sean claras, la matriz A_T se llama la *representación matricial de T con respecto a las bases $\mathcal{B}$ y $\mathcal{C}$* .

Así, toda transformación lineal $T : U \rightarrow V$ puede representarse como una matriz $A_T \in M(m \times n)$. Por otro lado, toda matriz $A(m \times n)$ induce una transformación lineal $T_A : U \rightarrow V$.

Observación muy importante. Esta identificación de matrices de tamaño $m \times n$ con transformaciones lineales $U \rightarrow V$ depende de la elección de la base! Véase el Ejemplo 6.48.

Resumamos lo que hemos encontrado hasta ahora.

Teorema 6.46. Sean U, V espacios vectoriales de dimensión finita y sea $\mathcal{B} = \{u_1, \dots, u_n\}$ una base ordenada de U , y sea $\mathcal{C} = \{v_1, \dots, v_m\}$ una base ordenada de V . Entonces se cumple lo siguiente:

- (I) Toda transformación lineal $T : U \rightarrow V$ puede representarse como una matriz $A_T \in M(m \times n)$ tal que

$$(Tw)_{\mathcal{C}} = A_T \vec{w}_{\mathcal{B}}$$

donde $(Tw)_{\mathcal{C}}$ es la representación de $Tw \in V$ con respecto a la base $\mathcal{C}$ y $\vec{w}_{\mathcal{B}}$ es la representación de $w \in U$ con respecto a la base $\mathcal{B}$. Las entradas a_{ij} de A_T pueden calcularse como en (6.20).

(II) Toda matriz $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in M(m \times n)$ induce una transformación lineal $T : U \rightarrow V$ definida por

$$T(u_j) = a_{1j}v_1 + \dots + a_{mj}v_m, \quad j = 1, \dots, n.$$

(III) $T = T_{A_T}$ and $A = A_{T_A}$. Esto significa: Si comenzamos con una transformación lineal $T : U \rightarrow V$, calculamos su representación matricial A_T y luego la transformación lineal $T_{A_T} : U \rightarrow V$ inducida por A_T , entonces recuperamos nuestra transformación original T . Si, por otro lado, comenzamos con una matriz $A \in M(m \times n)$, calculamos la transformación lineal $T_A : U \rightarrow V$ inducida por A y luego calculamos su representación matricial A_{T_A} , entonces recuperamos nuestra matriz original A .

Demostración. Ya mostramos (I) y (II) en el texto anterior al teorema. Para ver (III), comencemos con una transformación lineal $T : U \rightarrow V$ y sea $A_T = (a_{ij})$ la representación matricial de T con respecto a las bases $\mathcal{B}$ y $\mathcal{C}$. Para T_{A_T} , la transformación lineal inducida por A_T , se sigue que

$$T_{A_T}u_j = a_{1j}v_1 + \dots + a_{mj}v_m = Tu_j, \quad j = 1, \dots, n$$

Como esto es cierto para todos los vectores de la base y tanto T como T_{A_T} son lineales, deben ser iguales.

Si, por otro lado, se nos da una matriz $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in M(m \times n)$, entonces la transformación lineal T_A inducida por A actúa sobre los vectores de la base $u_1, \dots, u_n$ como sigue:

$$T_A u_j = T_{A_T} u_j = a_{1j}v_1 + \dots + a_{mj}v_m.$$

Pero entonces, por definición de la representación matricial A_{T_A} de T_A , se sigue que $A_{T_A} = A$. $\square$

Veamos esta “identificación” de matrices con transformaciones lineales de manera un poco más formal. Al elegir una base $\mathcal{B} = \{u_1, \dots, u_n\}$ en U e identificar así U con $\mathbb{R}^n$, en realidad estamos definiendo una biyección lineal

$$\Psi : U \rightarrow \mathbb{R}^n, \quad \Psi(\lambda u_1 + \dots + \lambda_n u_n) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Recordemos que denotamos el vector del lado derecho por $\vec{u}_{\mathcal{B}}$.

Lo mismo ocurre si elegimos una base $\mathcal{C} = \{v_1, \dots, v_m\}$ de V . Obtenemos una biyección lineal

$$\Phi : V \rightarrow \mathbb{R}^m, \quad \Phi(\mu v_1 + \dots + \mu_m v_m) = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix}.$$

Con estas transformaciones lineales, encontramos que

$$A_T = \Phi \circ T \circ \Psi^{-1} \quad \text{and} \quad T_A = \Phi^{-1} \circ A \circ \Psi.$$

Las transformaciones Ψ y Φ “traducen” los espacios U y V a $\mathbb{R}^n$ y $\mathbb{R}^m$, donde las bases elegidas sirven como “diccionario”. De esta manera, “traducen” transformaciones lineales $U : U \rightarrow V$ a matrices $A \in M(m \times n)$ y viceversa. En un diagrama esto se ve así:

$$\begin{array}{ccc} U & \xrightarrow{T} & V \\ \Psi \downarrow & & \downarrow \Phi \\ \mathbb{R}^n & \xrightarrow{A_T} & \mathbb{R}^m \end{array}$$

Así, para ir de U a V , podemos hacer el rodeo por $\mathbb{R}^n$ y $\mathbb{R}^m$. El diagrama anterior se llama *diagrama conmutativo*. Esto significa que no importa qué camino tomemos para ir de una esquina del diagrama a otra, siempre que nos movamos en las direcciones de las flechas. Observe que en este caso incluso podemos ir en las direcciones opuestas de las flechas que representan Ψ y Φ , porque son biyecciones.

¿Cuál es la utilidad de una representación matricial de una transformación lineal? A veces los cálculos son más fáciles en el mundo de las matrices. Por ejemplo, sabemos cómo calcular la imagen y el kernel de una matriz. Por lo tanto, usando el Teorema 6.18:

- Si queremos calcular $\text{Im } T$, solo necesitamos calcular $\text{Im } A_T$ y luego usar Φ para “traducir de regreso” a la imagen de T . En fórmula: $\text{Im } T = \text{Im}(\Phi^{-1}A_T\Psi) = \text{Im}(\Phi^{-1}A_T) = \Phi^{-1}(\text{Im } A_T)$.
- Si queremos calcular $\ker T$, solo necesitamos calcular $\ker A_T$ y luego usar Ψ para “traducir de regreso” al kernel de T . En fórmula: $\ker T = \ker(\Phi^{-1}A_T\Psi) = \ker(A_T\Psi) = \Psi^{-1}(\ker A_T)$.
- Si $\dim U = \dim V$, es decir, si $n = m$, entonces T es invertible si y solo si A_T es invertible. Esto ocurre si y solo si $\det A_T \neq 0$.

Resumamos. Del Teorema 6.24 obtenemos de nuevo el siguiente teorema muy importante, vea Teorema 6.20 y Proposición 6.16.

Teorema 6.47. Sean U, V espacios vectoriales y sea $T : U \rightarrow V$ una transformación lineal. Entonces

$$\dim U = \dim(\ker T) + \dim(\text{Im } T). \quad (6.21)$$

Si $\dim U = \dim V$, entonces las siguientes condiciones son equivalentes:

- (I) T es invertible.
- (II) T es inyectiva, es decir, $\ker T = \{\mathbb{O}\}$.
- (III) T es sobreyectiva, es decir, $\text{Im } T = V$.

Observe que si T es biyectiva, entonces necesariamente se tiene $\dim U = \dim V$.

Veamos algunos ejemplos.

Ejemplo 6.48. Consideramos el operador de derivación

$$T : P_3 \rightarrow P_3, \quad Tp = p'.$$

Observe que, en este caso, los espacios vectoriales U y V son ambos iguales a P_3 .

- (I) Represente T con respecto a la base $\mathcal{B} = \{p_1, p_2, p_3, p_4\}$ y encuentre su kernel donde $p_1 = 1$, $p_2 = X$, $p_3 = X^2$, $p_4 = X^3$.

Solución. Solo necesitamos evaluar T en los elementos de la base y luego escribir nuevamente el resultado como combinación lineal de la base. Como en este caso las bases son “sencillas”, los cálculos son bastante simples:

$$Tp_1 = 0, \quad Tp_2 = 1 = p_1, \quad Tp_3 = 2X = 2p_2, \quad Tp_4 = 3X^2 = 3p_3.$$

Por lo tanto, la representación matricial de T es

$$A_T^{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

El kernel de A_T es claramente $\text{span}\{\vec{e}_1\}$, de donde $\ker T = \text{span}\{p_1\} = \text{span}\{1\}$. $\diamond$

- (II) Represente T con respecto a la base $\mathcal{C} = \{q_1, q_2, q_3, q_4\}$ y encuentre su kernel donde $q_1 = X^3$, $q_2 = X^2$, $q_3 = X$, $q_4 = 1$.

Solución. Nuevamente solo necesitamos evaluar T en los elementos de la base y luego escribir el resultado como combinación lineal de la base.

$$Tq_1 = 3X^2 = 3q_2, \quad Tq_2 = 2X = 2q_3, \quad Tq_3 = X = q_4, \quad Tq_4 = 0.$$

Por lo tanto, la representación matricial de T es

$$A_T^{\mathcal{C}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

El kernel de A_T es claramente $\text{span}\{\vec{e}_4\}$, de donde $\ker T = \text{span}\{q_4\} = \text{span}\{1\}$. $\diamond$

- (III) Represente T con respecto a la base $\mathcal{B}$ en el dominio de T (en el P_3 en el “lado izquierdo”) y en la base $\mathcal{C}$ en el espacio de la salida (en el P_3 en el “lado derecho”).

Solución. Calculamos

$$Tp_1 = 0, \quad Tp_2 = 1 = q_4, \quad Tp_3 = 2X = 2q_3, \quad Tp_4 = 3X^2 = 3q_2.$$

Por lo tanto, la representación matricial de T es

$$A_T^{\mathcal{B}, \mathcal{C}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

El kernel de A_T es claramente $\text{span}\{\vec{e}_1\}$, de donde $\ker T = \text{span}\{p_1\} = \text{span}\{1\}$. $\diamond$

(iv) Represente T con respecto a la base $\mathcal{D} = \{r_1, r_2, r_3, r_4\}$ y encuentre su kernel donde

$$r_1 = X^3 + X, \quad r_2 = 2X^3 + X^2 + 2X, \quad r_3 = 3X^3 + X^2 + 4X + 1, \quad r_4 = 4X^3 + X^2 + 4X + 1.$$

Solución 1. Nuevamente solo necesitamos evaluar T en los elementos de la base y luego escribir el resultado como combinación lineal de la base. Esta vez los cálculos son un poco más tediosos.

$$\begin{aligned} Tr_1 &= 3X^2 + 1 &= -8r_1 + 2r_2 + r_4, \\ Tr_2 &= 6X^2 + 2X + 2 &= -14r_1 + 4r_2 + 2r_3, \\ Tr_3 &= 9X^2 + 2X + 4 &= -24r_1 + 5r_2 + 2r_3 + 2r_4, \\ Tr_4 &= 12X^2 + 2X + 4 &= -30r_1 + 8r_2 + 2r_3 + 2r_4. \end{aligned}$$

Por lo tanto, la representación matricial de T es

$$A_T^{\mathcal{D}} = \begin{pmatrix} -8 & -14 & -24 & -30 \\ 2 & 4 & 5 & 8 \\ 0 & 2 & 2 & 2 \\ 1 & 0 & 2 & 2 \end{pmatrix}.$$

Para calcular el kernel de A_T , aplicamos el proceso de Gauß-Jordan y obtenemos

$$A_T^{\mathcal{D}} = \begin{pmatrix} -8 & -14 & -24 & -30 \\ 2 & 4 & 5 & 8 \\ 0 & 2 & 2 & 2 \\ 1 & 0 & 2 & 2 \end{pmatrix} \rightarrow \cdots \rightarrow \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

El kernel de A_T es claramente $\text{span}\{-2\vec{e}_1 - \vec{e}_2 + \vec{e}_4\}$, de donde $\ker T = \text{span}\{-2r_1 - r_2 + r_4\} = \text{span}\{1\}$. $\diamond$

Solución 2. Ya tenemos la representación matricial $A_T^{\mathcal{C}}$ y podemos usarla para calcular $A_T^{\mathcal{D}}$. Para ello definamos los vectores

$$\vec{\rho}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{\rho}_2 = \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \quad \vec{\rho}_3 = \begin{pmatrix} 3 \\ 1 \\ 4 \\ 1 \end{pmatrix}, \quad \vec{\rho}_4 = \begin{pmatrix} 4 \\ 1 \\ 4 \\ 1 \end{pmatrix}.$$

Observe que estos vectores son las representaciones de nuestros vectores de base $r_1, \dots, r_4$ en la base $\mathcal{C}$. La matriz de cambio de bases de $\mathcal{C}$ a $\mathcal{D}$ y su inversa son, en coordenadas,

$$S_{\mathcal{D} \rightarrow \mathcal{C}} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 4 & 4 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad S_{\mathcal{C} \rightarrow \mathcal{D}} = S_{\mathcal{D} \rightarrow \mathcal{C}}^{-1} = \begin{pmatrix} 0 & -2 & 1 & -2 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 1 \end{pmatrix}.$$

Se sigue que

$$\begin{aligned} A_T^{\mathcal{D}} &= S_{\mathcal{C} \rightarrow \mathcal{D}} A_T^{\mathcal{C}} S_{\mathcal{D} \rightarrow \mathcal{C}} \\ &= \begin{pmatrix} 0 & -2 & 1 & -2 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 4 & 4 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -8 & -14 & -24 & -30 \\ 2 & 4 & 5 & 8 \\ 0 & 2 & 2 & 2 \\ 1 & 0 & 2 & 2 \end{pmatrix}. \end{aligned}$$

Veamos cómo se ve esto en diagramas. Definimos las dos biyecciones de P_3 con $\mathbb{R}^4$ dadas por la elección de las bases $\mathcal{C}$ y $\mathcal{D}$ mediante $\Psi_{\mathcal{C}}$ y $\Psi_{\mathcal{D}}$:

$$\begin{aligned} \Psi_{\mathcal{C}} : P_3 &\rightarrow \mathbb{R}^4, & \Psi_{\mathcal{C}}(q_1) &= \vec{e}_1, \Psi_{\mathcal{C}}(q_2) = \vec{e}_2, \Psi_{\mathcal{C}}(q_3) = \vec{e}_3, \Psi_{\mathcal{C}}(q_4) = \vec{e}_4, \\ \Psi_{\mathcal{D}} : P_3 &\rightarrow \mathbb{R}^4, & \Psi_{\mathcal{D}}(r_1) &= \vec{e}_1, \Psi_{\mathcal{D}}(r_2) = \vec{e}_2, \Psi_{\mathcal{D}}(r_3) = \vec{e}_3, \Psi_{\mathcal{D}}(r_4) = \vec{e}_4. \end{aligned}$$

Entonces tenemos los siguientes diagramas:

$$\begin{array}{ccc} P_3 & \xrightarrow{T} & P_3 \\ \Psi_{\mathcal{C}} \downarrow & & \downarrow \Psi_{\mathcal{C}} \\ \mathbb{R}^4 & \xrightarrow{A_T^{\mathcal{C}}} & \mathbb{R}^4 \end{array} \quad \begin{array}{ccc} P_3 & \xrightarrow{T} & P_3 \\ \Psi_{\mathcal{D}} \downarrow & & \downarrow \Psi_{\mathcal{D}} \\ \mathbb{R}^4 & \xrightarrow{A_T^{\mathcal{D}}} & \mathbb{R}^4 \end{array}$$

Ya conocemos todo en el diagrama de la izquierda y queremos calcular $A_T^{\mathcal{D}}$ en el diagrama de la derecha. Podemos juntar los diagramas de la siguiente manera:

$$\begin{array}{ccccc} P_3 & & \xrightarrow{T} & & P_3 \\ & \searrow \Psi_{\mathcal{C}} & & \swarrow \Psi_{\mathcal{C}} & \\ \Psi_{\mathcal{D}} \downarrow & & & & \downarrow \Psi_{\mathcal{D}} \\ \mathbb{R}^4 & \xrightarrow{S_{\mathcal{D} \rightarrow \mathcal{C}}} & \mathbb{R}^4 & \xrightarrow{A_T^{\mathcal{C}}} & \mathbb{R}^4 & \xrightarrow{S_{\mathcal{C} \rightarrow \mathcal{D}}} & \mathbb{R}^4 \\ & & & \searrow A_T^{\mathcal{D}} & & \swarrow & \end{array}$$

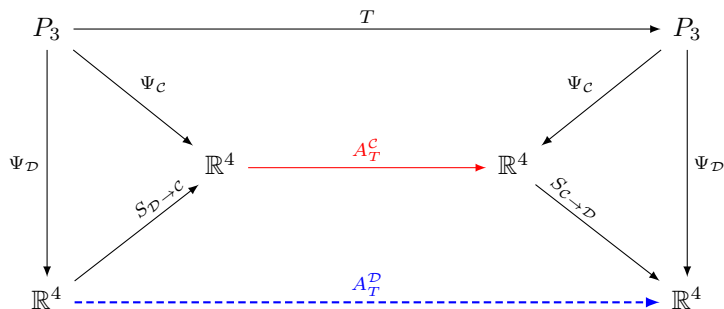
También podemos ver que las transformaciones de cambio de base $S_{\mathcal{D} \rightarrow \mathcal{C}}$ y $S_{\mathcal{C} \rightarrow \mathcal{D}}$ son

$$S_{\mathcal{D} \rightarrow \mathcal{C}} = \Psi_{\mathcal{C}} \circ \Psi_{\mathcal{D}}^{-1}, \quad S_{\mathcal{C} \rightarrow \mathcal{D}} = \Psi_{\mathcal{D}} \circ \Psi_{\mathcal{C}}^{-1}.$$

Para $A_T^{\mathcal{D}}$ obtenemos

$$A_T^{\mathcal{D}} = \Psi_{\mathcal{D}} \circ T \circ \Psi_{\mathcal{D}}^{-1} = S_{\mathcal{D} \rightarrow \mathcal{C}} \circ A_T^{\mathcal{C}} \circ S_{\mathcal{C} \rightarrow \mathcal{D}}.$$

Otra manera de dibujar el diagrama anterior es



◇

Observe que las matrices A_T^B , A_T^C , A_T^D y $A_T^{B,C}$ se ven todas diferentes, pero describen la misma transformación lineal. La razón por la que se ven diferentes es que en cada caso usamos bases distintas para describirlas.

Ejemplo 6.49. El siguiente ejemplo no es muy aplicado, pero sirve para practicar un poco más. Consideramos el operador dado por

$$T: M(2 \times 2) \rightarrow P_2, \quad T\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = (a+c)X^2 + (a-b)X + a-b+d.$$

Muestre que T es una transformación lineal y represente T con respecto a las bases $\mathcal{B} = \{B_1, B_2, B_3, B_4\}$ de $M(2 \times 2)$ y $\mathcal{C} = \{p_1, p_2, p_3\}$ de P_2 donde

$$B_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad B_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

y

$$p_1 = 1, \quad p_2 = X, \quad p_3 = X^2.$$

Encuentre bases para $\ker T$ y $\operatorname{Im} T$, y sus dimensiones.

Solución. Primero verificamos que T es efectivamente una transformación lineal. Para ello, tomamos matrices $A_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$ y $A_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$ y $\lambda \in \mathbb{R}$. Entonces

$$\begin{aligned} T(\lambda A_1 + A_2) &= T\left(\lambda \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} + \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}\right) = T\left(\lambda \begin{pmatrix} \lambda a_1 + a_2 & \lambda b_1 + b_2 \\ \lambda c_1 + c_2 & \lambda d_1 + d_2 \end{pmatrix}\right) \\ &= (\lambda a_1 + a_2 + \lambda c_1 + c_2)X^2 + (\lambda a_1 + a_2 - \lambda b_1 - b_2)X + \lambda a_1 + a_2 - (\lambda b_1 + b_2) + \lambda d_1 + d_2 \\ &= \lambda[(a_1 + c_1)X^2 + (a_1 - b_1)X + a_1 - b_1 + d_1] + [(a_2 + c_2)X^2 + (a_2 - b_2)X + a_2 - b_2 + d_2] \\ &= \lambda T(A_1) + T(A_2). \end{aligned}$$

Esto muestra que T es una transformación lineal.

Ahora calculamos su representación matricial con respecto a las bases dadas.

$$\begin{aligned} TB_1 &= X^2 + X + 1 = p_1 + p_2 + p_3, \\ TB_2 &= -X = -p_2, \\ TB_3 &= X^2 = p_3, \\ TB_4 &= 1 = p_1. \end{aligned}$$

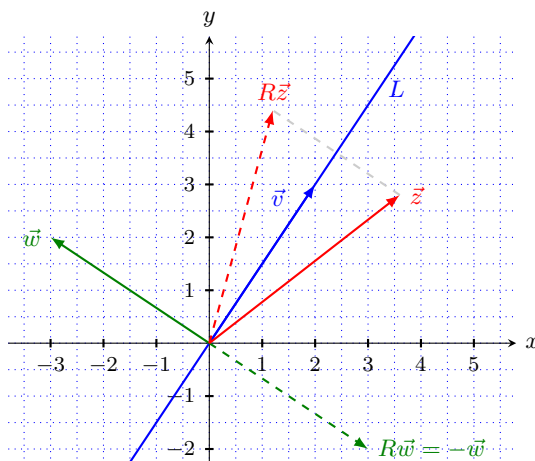


FIGURA 6.5: La figura muestra la reflexión R respecto de la recta L . El vector $\vec{v}$ es paralelo a L , de modo que $R\vec{v} = \vec{v}$. El vector $\vec{w}$ es perpendicular a L , de modo que $R\vec{w} = -\vec{w}$.

Por lo tanto, la representación matricial de T es

$$A_T = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Para determinar el kernel y la imagen de A_T , aplicamos el proceso de Gauß-Jordan:

$$A_T = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Así, la imagen de A_T es $\mathbb{R}^3$ y su kernel es $\ker A_T = \text{span}\{\vec{e}_1 + \vec{e}_2 - \vec{e}_3 - \vec{e}_4\}$. Por lo tanto, $\text{Im } T = P_2$ y $\ker T = \text{span}\{B_1 + B_2 - B_3 - B_4\} = \text{span}\left\{\begin{pmatrix} -1 \\ -1 \end{pmatrix}\right\}$. Para sus dimensiones encontramos $\dim(\text{Im } T) = 3$ y $\dim(\ker T) = 1$. $\diamond$

Ejemplo 6.50 (Reflexión en $\mathbb{R}^2$). En $\mathbb{R}^2$, considere la recta $L : 3x - 2y = 0$. Sea $R : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la transformación que toma un vector en $\mathbb{R}^2$ y lo refleja respecto de la recta L , véase la Figura 6.5. Encuentre la representación matricial de R con respecto a la base estándar de $\mathbb{R}^2$.

Observación. Observe que L es la recta que pasa por el origen y es paralela al vector $\vec{v} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

Solución 1 (usar coordenadas adaptadas al problema). Claramente, hay dos direcciones que son especiales en este problema: la dirección paralela y la dirección ortogonal a la recta. Por lo tanto, una base adaptada al ejercicio es $\mathcal{B} = \{\vec{v}, \vec{w}\}$ donde $\vec{v} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ y $\vec{w} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$. Claramente, $R\vec{v} = \vec{v}$ y $R\vec{w} = -\vec{w}$. Por lo tanto, la representación matricial de R con respecto a la base $\mathcal{B}$ es

$$A_R^{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Para obtener la representación A_R con respecto a la base estándar, solo necesitamos realizar un cambio de base. Recordemos que las matrices de cambio de bases están dadas por

$$S_{\mathcal{B} \rightarrow \text{can}} = (\vec{v}|\vec{w}) = \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}, \quad S_{\text{can} \rightarrow \mathcal{B}} = S_{\mathcal{B} \rightarrow \text{can}}^{-1} = \frac{1}{13} \begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}.$$

Therefore

$$A_R = S_{\mathcal{B} \rightarrow \text{can}} A_R^{\mathcal{B}} S_{\text{can} \rightarrow \mathcal{B}} = \frac{1}{13} \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix} = \frac{1}{13} \begin{pmatrix} -5 & 12 \\ 12 & 5 \end{pmatrix}. \quad \diamond$$

Solución 2 (reducir el problema a una reflexión conocida). El problema sería fácil si se nos pidiera calcular la representación matricial de la reflexión respecto del eje x . Esta sería simplemente $A_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Ahora podemos proceder como sigue: primero rotamos $\mathbb{R}^2$ alrededor del origen de manera que la recta L sea paralela al eje x ; luego reflejamos respecto del eje x y finalmente rotamos de regreso. El resultado es el mismo que reflejar respecto de L . Supongamos que Rot es la matriz de rotación. Entonces

$$A_T = \text{Rot}^{-1} \circ A_0 \circ \text{Rot}. \quad (6.22)$$

¿Cómo podemos calcular Rot ? Sabemos que $\text{Rot}\vec{v} = \vec{e}_1$ y que $\text{Rot}\vec{w} = \vec{e}_2$. Se sigue que $\text{Rot}^{-1} = (\vec{v}|\vec{w}) = \begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}$. Observe que, salvo por un factor numérico, esto es $S_{\mathcal{B} \rightarrow \text{can}}$. Podemos calcular fácilmente que $\text{Rot} = (\text{Rot}^{-1})^{-1} = \frac{1}{13} \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}$. Si insertamos esto en (6.22), encontramos nuevamente $A_R = \begin{pmatrix} -5 & 12 \\ 12 & 5 \end{pmatrix}$. $\diamond$

Solución 3 (cálculo directo). Podemos formar un sistema de ecuaciones lineales para encontrar A_T . Escribimos $A_R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ con números desconocidos a, b, c, d . Nuevamente usamos que sabemos que $A_T\vec{v} = \vec{v}$ y $A_T\vec{w} = -\vec{w}$. Esto da las siguientes ecuaciones:

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \vec{v} = A_T\vec{v} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2a + 3b \\ 2c + 3d \end{pmatrix},$$

$$\begin{pmatrix} -3 \\ 2 \end{pmatrix} = \vec{w} = -A_T\vec{w} = -\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 3a - 2b \\ 3c - 2d \end{pmatrix}$$

lo cual da el sistema

$$2a + 3b = 2, \quad 2c + 3d = 3, \quad 3a - 2b = -3, \quad 3c - 2d = 2.$$

Su solución única es $a = -\frac{5}{13}$, $b = c = \frac{12}{13}$, $d = \frac{5}{13}$, entonces $A_R = \begin{pmatrix} -5 & 12 \\ 12 & 5 \end{pmatrix}$. $\diamond$

Ejemplo 6.51 (Reflexión y proyección ortogonal en $\mathbb{R}^3$). En $\mathbb{R}^3$, considere el plano $E : x - 2y + 3z = 0$. Sea $R : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ la transformación que toma un vector en $\mathbb{R}^3$ y lo refleja respecto del plano E y sea $P : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ la proyección ortogonal sobre E . Encuentre la representación matricial de R con respecto a la base estándar de $\mathbb{R}^3$.

Observación. Observe que E es el plano que pasa por el origen y es ortogonal al vector $\vec{n} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$.

Además, si tomamos $\vec{v} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ y $\vec{w} = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}$, entonces es fácil ver que $\{\vec{v}, \vec{w}\}$ es una base de E .

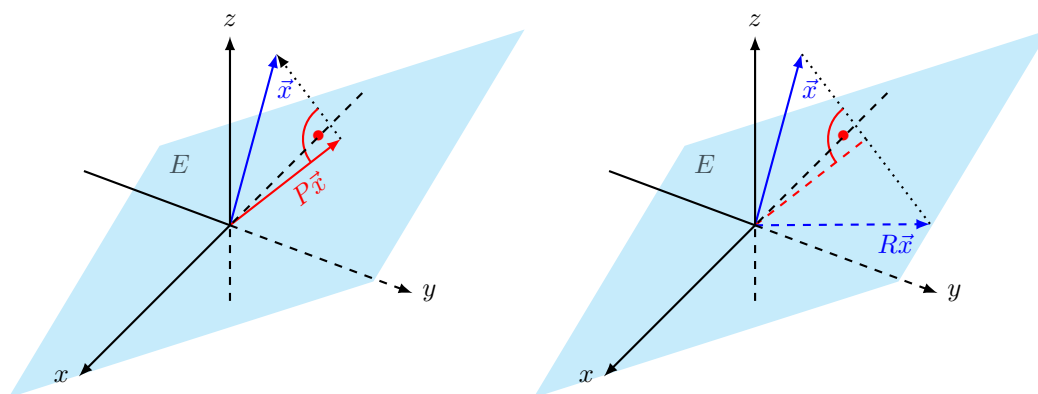


FIGURA 6.6: La figura muestra el plano $E : x - 2y + 3z = 0$ y, para el vector $\vec{x}$, muestra su proyección ortogonal $P\vec{x}$ sobre E y su reflexión $R\vec{x}$ respecto de E ; véase el Ejemplo 6.51.

Solución 1 (usar coordenadas adaptadas al problema). Claramente, una base adaptada al ejercicio es $\mathcal{B} = \{\vec{v}, \vec{w}, \vec{n}\}$ porque para estos vectores tenemos $R\vec{v} = \vec{v}$, $R\vec{w} = \vec{w}$, $R\vec{n} = -\vec{n}$, y $P\vec{v} = \vec{v}$, $P\vec{w} = \vec{w}$, $P\vec{n} = \vec{0}$. Por lo tanto, la representación matricial de R con respecto a la base $\mathcal{B}$ es

$$A_R^{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

y la de P es

$$A_P^{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Para obtener las representaciones A_R y A_P con respecto a la base estándar, solo necesitamos realizar un cambio de base. Recordemos que las matrices de cambio de bases están dadas por

$$S_{\mathcal{B} \rightarrow \text{can}} = (\vec{v} | \vec{w} | \vec{n}) = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 3 & -2 \\ 0 & 2 & 3 \end{pmatrix}, \quad S_{\text{can} \rightarrow \mathcal{B}} = S_{\mathcal{B} \rightarrow \text{can}}^{-1} = \frac{1}{28} \begin{pmatrix} 13 & 2 & -3 \\ -3 & 6 & 5 \\ 2 & -4 & 6 \end{pmatrix}.$$

Por lo tanto

$$\begin{aligned} A_R &= S_{\mathcal{B} \rightarrow \text{can}} A_R^{\mathcal{B}} S_{\text{can} \rightarrow \mathcal{B}} = \frac{1}{28} \begin{pmatrix} 2 & 0 & 1 \\ 1 & 3 & -2 \\ 0 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 13 & 2 & -3 \\ -3 & 6 & 5 \\ 2 & -4 & 6 \end{pmatrix} \\ &= \frac{1}{7} \begin{pmatrix} 6 & 2 & -3 \\ 2 & 3 & 6 \\ -3 & 6 & -2 \end{pmatrix} \end{aligned}$$

y

$$\begin{aligned}
A_R &= S_{\mathcal{B} \rightarrow \text{can}} A_P^{\mathcal{B}} S_{\text{can} \rightarrow \mathcal{B}} = \frac{1}{28} \begin{pmatrix} 2 & 0 & 2 \\ 1 & 3 & -1 \\ 0 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 13 & 2 & -3 \\ -3 & 6 & 5 \\ 2 & -4 & 6 \end{pmatrix} \\
&= \frac{1}{14} \begin{pmatrix} 13 & 2 & -3 \\ 2 & 10 & 6 \\ -3 & 6 & 5 \end{pmatrix}. \quad \diamond
\end{aligned}$$

Solución 2 (reducir el problema a una reflexión conocida). El problema sería fácil si se nos pidiera calcular la representación matricial de la reflexión respecto del plano xy . Esta sería simplemente $A_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Ahora podemos proceder como sigue: primero rotamos $\mathbb{R}^3$ alrededor del origen de manera que el plano E sea paralelo al plano xy ; luego reflejamos respecto del plano xy y finalmente rotamos de regreso. El resultado es el mismo que reflejar respecto del plano E . Dejamos los detalles al lector. Un procedimiento análogo funciona para la proyección ortogonal. $\diamond$

Solución 3 (cálculo directo). Finalmente, podemos formar un sistema de ecuaciones lineales para encontrar A_R . Escribimos $A_R = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ con incógnitas a_{ij} . Nuevamente usamos que sabemos que $A_R \vec{v} = \vec{v}$, $A_R \vec{w} = \vec{w}$ y $A_R \vec{n} = -\vec{n}$. Esto da un sistema de 9 ecuaciones lineales para las nueve incógnitas a_{ij} , que puede resolverse. $\diamond$

Observación 6.52. Otra solución más es la siguiente. Sea Q la proyección ortogonal sobre $\vec{n}$. Ya sabemos cómo calcular su matriz de representación:

$$Q\vec{x} = \frac{\langle \vec{x}, \vec{n} \rangle}{\|\vec{n}\|^2} \vec{n} = \frac{x - 2y + 3z}{14} \vec{n} = \frac{1}{14} \begin{pmatrix} 1 & -2 & 3 \\ -2 & 4 & -6 \\ 3 & -6 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Por lo tanto $A_Q = \frac{1}{14} \begin{pmatrix} 1 & -2 & 3 \\ -2 & 4 & -6 \\ 3 & -6 & 9 \end{pmatrix}$. Geométricamente, es claro que $P = \text{id} - Q$ y $R = \text{id} - 2Q$. De esto se sigue que

$$A_P = \text{id} - A_Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{14} \begin{pmatrix} 1 & -2 & 3 \\ -2 & 4 & -6 \\ 3 & -6 & 9 \end{pmatrix} = \frac{1}{14} \begin{pmatrix} 13 & 2 & -3 \\ 2 & 10 & 6 \\ -3 & 6 & 5 \end{pmatrix}$$

y

$$A_R = \text{id} - 2A_Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{7} \begin{pmatrix} 1 & -2 & 3 \\ -2 & 4 & -6 \\ 3 & -6 & 9 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 6 & 2 & -3 \\ 2 & 3 & 6 \\ -3 & 6 & -2 \end{pmatrix}.$$

Cambio de bases como representación matricial de la identidad

Finalmente observemos que una matriz de cambio de bases no es otra cosa que la matriz identidad escrita con respecto a bases distintas. Para ver esto, sean $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ y $\mathcal{C} = \{\vec{w}_1, \dots, \vec{w}_w\}$

bases de $\mathbb{R}^n$. Definimos las biyecciones lineales Ψ_B y Ψ_C como sigue:

$$\begin{aligned}\Psi_B : \mathbb{R}^n &\rightarrow \mathbb{R}^n, & \Psi_B(\vec{e}_1) &= \vec{v}_1, \dots, \Psi_B(\vec{e}_n) = \vec{v}_n, \\ \Psi_C : \mathbb{R}^n &\rightarrow \mathbb{R}^n, & \Psi_C(\vec{e}_1) &= \vec{w}_1, \dots, \Psi_C(\vec{e}_n) = \vec{w}_n.\end{aligned}$$

Además definimos las matrices de cambio de bases

$$S_{B \rightarrow can} = (\vec{v}_1 | \dots | \vec{v}_n), \quad S_{C \rightarrow can} = (\vec{w}_1 | \dots | \vec{w}_n).$$

Observe que estas matrices son exactamente las representaciones matriciales de Ψ_B y Ψ_C . Ahora consideremos el diagrama

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{\text{id}} & \mathbb{R}^n \\ \Psi_B^{-1} \downarrow & & \downarrow \Psi_C^{-1} \\ \mathbb{R}^n & \xrightarrow{A_{\text{id}}} & \mathbb{R}^n \end{array}$$

Por lo tanto

$$A_{\text{id}} = \Psi_C^{-1} \circ \text{id} \circ \Psi_B = \Psi_C \circ \Psi_B^{-1} = S_{C \rightarrow can}^{-1} \circ S_{B \rightarrow can} = S_{can \rightarrow C} \circ S_{B \rightarrow can} = S_{B \rightarrow C}.$$

Ahora debería haber entendido

- por qué toda transformación lineal entre espacios vectoriales de dimensión finita puede escribirse como una matriz y por qué la matriz depende de las bases elegidas,
- cómo cambia la representación matricial si cambian las bases elegidas,
- en particular, cómo cambia la representación matricial si se reordenan las bases elegidas,
- etc.

Ahora debería ser capaz de

- representar una transformación lineal entre espacios vectoriales de dimensión finita como una matriz,
- usar la representación matricial de una transformación lineal para calcular su kernel y su imagen,
- interpretar una matriz como una transformación lineal entre espacios vectoriales de dimensión finita,
- etc.

Ejercicios.

1. De los ejercicios 1 al 14 (exceptuando el 11.) de la sección 6.1 obtenga la representación matricial de T en las respectivas bases canónicas.
2. Encuentre la representación matricial en la respectiva base canónica de las siguientes transformaciones. En P_n tome la base $\{1, X, X^2, \dots, X^{n+1}\}$.

- (a) En P_4 , $D(p) = Xp' - p$.
- (b) En P_4 , $D(p) = p''$.
- (c) En P_n , $D(p) = p^{(m)}$, la m -ésima derivada de p .
- (d) En $M(3 \times 3)$, $T(A) = A - A^t$.
- (e) En P_n , $J : P_n \rightarrow \mathbb{R}$ dada por $J(p) = \int_0^1 p(t)dt$.
3. Para cada transformación lineal dada, encuentre su representación matricial en las bases indicadas:
- (a) $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $T(x, y) = (y, 0)$ de base canónica a $\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$.
- (b) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y+z \\ x-y \end{pmatrix}$ de $\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$ a $\mathcal{B}_2 = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$.
- (c) $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $T(\vec{x}) = \vec{x} \times \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ en la base $\mathcal{B} = \left\{ \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 2 \end{pmatrix} \right\}$.
- (d) $T : P_2 \rightarrow \mathbb{R}^2$, $T(aX^2 + bX + c) = \begin{pmatrix} 2a+b+c \\ b-3c \end{pmatrix}$ de $\mathcal{B}_1 = \{X^2 + 1, X^2 + X, X^2 + X + 1\}$ a $\mathcal{B}_2 = \left\{ \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$.
4. Sea $V = \text{span}\{\cos x, \sin x, x \cos x, x \sin x\}$ y $\mathcal{B} = \{\cos x, \sin x, x \cos x, x \sin x\}$.
- (a) Demuestre que $\mathcal{B}$ es base de V .
- (b) Para $D : V \rightarrow V$ dada por $D(f) = f'$, obtenga $[D]_{\mathcal{B}}$. ¿ D es invertible?
5. Sea $V = \text{span}\{1, e^x, e^{-x}\}$, $\mathcal{B}_1 = \{1, e^x, e^{-x}\}$ y $\mathcal{B}_2 = \{1, \cosh x, \sinh x\}$. Considere $D : V \rightarrow V$ dada por $D(f) = f'$, obtenga $[D]_{\mathcal{B}_2}^{\mathcal{B}_1}$. ¿ D es invertible?
6. Sea $\vec{w}$ un vector no nulo y $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por $T(\vec{x}) = \text{proy}_{\vec{w}} \vec{x}$. Encuentre una base $\mathcal{B}$ de $\mathbb{R}^3$ tal que $[T]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.
7. En $\mathbb{R}^3$, sean $E : x + y + z = 0$ y $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por $T(\vec{x}) = \text{reflexión de } \vec{x} \text{ con respecto a } E$.
- (a) Encuentre una base $\mathcal{B}$ tal que $[T]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.
- (b) Obtenga $[T]_{\text{can}}$.
- (c) Describa T en las coordenadas usuales.

8. (a) Sea $S : \mathbb{R}^4 \rightarrow \mathbb{R}$ una transformación lineal tal que $Se_1 = 4$, $Se_2 = -3$, $Se_3 = 0$ y $Se_4 = \pi$. Muestre que existe $\vec{w} \in \mathbb{R}^4$ tal que $S\vec{x} = \langle \vec{x}, \vec{w} \rangle$ para todo $\vec{x} \in \mathbb{R}^4$.

(b) Sea $S : \mathbb{R}^4 \rightarrow \mathbb{R}$ una transformación lineal tal que $S \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = 1$, $S \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = -2$, $S \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = 3$

y $S \begin{pmatrix} 1 \\ 0 \\ 0 \\ 4 \end{pmatrix} = -1$. Encuentre $\vec{w} \in \mathbb{R}^4$ tal que $S\vec{x} = \langle \vec{x}, \vec{w} \rangle$ para todo $\vec{x} \in \mathbb{R}^4$.

9. Sea $T : V \rightarrow W$ una transformación lineal y suponga que $\mathcal{B} = \{v_1, \dots, v_n\}$ es una base de V . Si para cada $i \in \{1, 2, \dots, n\}$ se tiene que $T(v_i) = \vec{0}$, muestre que $T = \mathcal{O}$ (la transformación que a todo elemento de V lo envía al vector cero de W).
10. Sea $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ transformación lineal tal que $T(\vec{e}_i) = \vec{e}_{i+1}$ si $1 \leq i < n$ y $T(\vec{e}_n) = 0$. Muestre que $T^n = \mathcal{O}$ ¿Cómo es la representación matricial de $T, T^2, \dots, T^{n-1}$ en la base canónica?

6.5. Resumen

Transformaciones lineales

Una función $T : U \rightarrow V$ entre dos espacios vectoriales sobre $\mathbb{K}$, U y V , se llama *transformación lineal* (o *función lineal* o *aplicación lineal*) si satisface

$$T(u_1 + \lambda u_2) = T(u_1) + \lambda T(u_2) \quad \text{para todo } u_1, u_2 \in U \text{ y } \lambda \in \mathbb{K}.$$

El conjunto de todas las transformaciones lineales de U a V se denota por $\mathcal{L}(U, V)$.

- La composición de transformaciones lineales es una transformación lineal.
- Si una transformación lineal es invertible, entonces su inversa es una transformación lineal.
- Si U, V son espacios vectoriales sobre $\mathbb{K}$, entonces $\mathcal{L}(U, V)$ es un espacio vectorial sobre $\mathbb{K}$. Esto significa: si $S, T \in \mathcal{L}(U, V)$ y $\lambda \in \mathbb{K}$, entonces $S + \lambda T \in \mathcal{L}(U, V)$.

Para una transformación lineal $T : U \rightarrow V$ definimos los siguientes conjuntos

$$\begin{aligned} \ker T &= \{u \in U : Tu = \mathcal{O}\} \subseteq U, \\ \text{Im } T &= \{Tu : u \in U\} \subseteq V. \end{aligned}$$

$\ker T$ se llama *kernel de T* o *kernel de T* . Es un subespacio de U . $\text{Im } T$ se llama *imagen de T* o *imagen de T* . Es un subespacio de V .

La transformación lineal T se llama *inyectiva* si $Tu_1 = Tu_2$ implica $u_1 = u_2$ para todo $u_1, u_2 \in U$. La transformación lineal T se llama *sobreyectiva* si para cada $v \in V$ existe algún $u \in U$ tal que $Tu = v$. La transformación lineal T se llama *biyectiva* si es inyectiva y sobreyectiva.

Sea $T : U \rightarrow V$ una transformación lineal.

- Las siguientes condiciones son equivalentes:

- (I) T es inyectiva.
- (II) $Tu = \mathbb{O}$ implica que $u = \mathbb{O}$.
- (III) $\ker T = \{\mathbb{O}\}$.

■ Las siguientes condiciones son equivalentes:

- (I) T es sobreyectiva.
- (II) $\operatorname{Im} T = V$.

■ Si T es biyectiva, entonces necesariamente $\dim U = \dim V$. En otras palabras: si $\dim U \neq \dim V$, entonces no existe ninguna biyección entre ellos.

Sean U, V espacios vectoriales sobre $\mathbb{K}$ y sea $T : U \rightarrow V$ una transformación lineal. Además, sean $E : U \rightarrow U$, $F : V \rightarrow V$ transformaciones lineales biyectivas. Entonces

$$\begin{array}{ll} \ker(FT) = \ker(T), & \ker(TE) = E^{-1}(\ker(T)), \\ \operatorname{Im}(FT) = F(\operatorname{Im}(T)), & \operatorname{Im}(TE) = \operatorname{Im}(T), \end{array}$$

and

$$\begin{array}{l} \dim \ker(T) = \dim \ker(FT) = \dim \ker(TE) = \dim \ker(FTE), \\ \dim \operatorname{Im}(T) = \dim \operatorname{Im}(FT) = \dim \operatorname{Im}(TE) = \dim \operatorname{Im}(FTE). \end{array}$$

Si $\dim U = n < \infty$, entonces

$$\dim(\ker(T)) + \dim(\operatorname{Im}(T)) = n.$$

Transformaciones lineales y matrices

Toda matriz $A \in M_{\mathbb{K}}(m \times n)$ representa una transformación lineal de $\mathbb{K}^n$ a $\mathbb{K}^m$ mediante

$$T_A : \mathbb{K}^n \rightarrow \mathbb{K}^m, \quad \vec{x} \mapsto A\vec{x}.$$

Muy a menudo escribimos A en lugar de T_A .

Por otro lado, toda transformación lineal $T : U \rightarrow V$ entre espacios vectoriales de dimensión finita U y V tiene una representación matricial. Sea $\mathcal{B} = \{u_1, \dots, u_n\}$ una base de U y $\mathcal{C} = \{v_1, \dots, v_m\}$ una base de V . Supongamos que $Tu_j = a_{1j}v_1 + \dots + a_{mj}v_m$. Entonces la representación matricial de T con respecto a las bases $\mathcal{B}$ y $\mathcal{C}$ es $A_T = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in M(m \times n)$. Observe que la representación matricial de T depende de las bases elegidas en U y V .

Si definimos las funciones Ψ y Φ como

$$\Psi : U \rightarrow \mathbb{K}^n, \quad \Psi(\alpha_1 u_1 + \dots + \alpha_n u_n) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad \Phi : V \rightarrow \mathbb{K}^m, \quad \Phi(\beta_1 v_1 + \dots + \beta_m v_m) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix},$$

entonces estas funciones son lineales y $\Phi \circ A_T \circ \Psi = T$ and $\Psi^{-1} \circ T \circ \Phi^{-1} = A_T$. En un diagrama esto es

$$\begin{array}{ccc} U & \xrightarrow{T} & V \\ \Psi \downarrow & & \downarrow \Phi \\ \mathbb{R}^n & \xrightarrow{A_T} & \mathbb{R}^m \end{array}$$

Matrices

Sea $A \in M(m \times n)$.

- El *espacio columna* C_A de A es el subespacio generado por sus vectores columna. Es igual a $\text{Im } A$.
- El *espacio fila* R_A de A es el subespacio generado por sus vectores fila. Es igual al complemento ortogonal de $\ker A$.
- $\dim R_A = \dim C_A = \dim(\text{Im } A) =$ número de columnas con pivotes en cualquier forma escalonada de A .

Kernel e imagen de A :

- $\dim(\ker A) =$ número de variables libres = número de columnas sin pivotes en cualquier forma escalonada por filas de A .

$\ker A$ es igual al conjunto solución de $A\vec{x} = \vec{0}$, que puede determinarse, por ejemplo, con la eliminación de Gauß o de Gauß-Jordan.

- $\dim(\text{Im } A) = \dim C_A =$ número de columnas con pivotes en cualquier forma escalonada por filas de A .

$\text{Im}(A)$ puede encontrarse por cualquiera de los siguientes dos métodos:

- (I) **Reducción por filas** de A . Las columnas de la matriz original A que corresponden a las columnas de la forma escalonada reducida por filas de A son una base de $\text{Im } A$.
- (II) **Reducción por columnas** de A . Las columnas restantes son una base de $\text{Im } A$.

6.6. Ejercicios

1. Determine si las siguientes funciones son lineales. Si lo son, calcule el kernel y la dimensión del kernel.

(a) $A : \mathbb{R}^3 \rightarrow M(2 \times 2)$, $A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + y & x - z \\ x + y - 3z & z \end{pmatrix}$,

(b) $B : \mathbb{R}^3 \rightarrow M(2 \times 2)$, $A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2xy & x - z \\ x + y - 3z & z \end{pmatrix}$,

(c) $D : P_3 \rightarrow P_4$, $Dp = p' + xp$,

$$(d) \quad T : P_3 \rightarrow M(2 \times 3), \quad T(ax^3 + bx^2 + cx + d) = \begin{pmatrix} a+b & b+c & c+d \\ 0 & a+d & 0 \end{pmatrix},$$

$$(e) \quad T : P_3 \rightarrow M(2 \times 3), \quad T(ax^3 + bx^2 + cx + d) = \begin{pmatrix} a+b & b+c & c+d \\ 0 & a+d & 3 \end{pmatrix}.$$

2. Sean U, V espacios vectoriales sobre $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$) y sea $T : U \rightarrow V$ una función lineal invertible. Entonces podemos considerar su función inversa $T^{-1} : \text{Im}(T) \rightarrow U$. Demuestre que es una función lineal.
3. Sean U, V, W espacios vectoriales sobre $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$) y sean $T : U \rightarrow V, S : V \rightarrow W$ funciones lineales. Demuestre que la composición $ST : U \rightarrow W$ también es una función lineal.
4. Sean U, V espacios vectoriales sobre $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$) y sea $T : U \rightarrow V$ una transformación lineal y c un escalar diferente de 0. Demuestre que
 - (a) $\text{Im}(cT) = \text{Im}(T)$,
 - (b) $\ker(cT) = \ker(T)$.
5. Sean U, V espacios vectoriales sobre $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$) y sean $S, T : U \rightarrow V$ transformaciones lineales.
 - (a) Demuestre que $\ker S \cap \ker T \subseteq \ker(S + T)$.
 - (b) Encuentre ejemplos de S, T para que $\ker S \cap \ker T = \ker(S + T)$.
 - (c) Encuentre ejemplos de S, T para que $\ker S \cap \ker T \subsetneq \ker(S + T)$.
 - (d) Demuestre que $\text{Im}(S + T) \subseteq (\text{Im } S) + (\text{Im } T)$.
 - (e) Encuentre ejemplos de S, T para que $\text{Im}(S + T) = (\text{Im } S) + (\text{Im } T)$.
 - (f) Encuentre ejemplos de S, T para que $\text{Im}(S + T) \subsetneq (\text{Im } S) + (\text{Im } T)$.

Hay alguna relación entre $\ker(S + T)$ y $(\ker S) + (\ker T)$?

Hay alguna relación entre $\text{Im}(S + T)$ y $\text{Im } S \cup \text{Im } T$?

6. Sean U, V espacios vectoriales sobre $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$). Con $\mathcal{L}(U, V)$ denotamos el conjunto de todas las transformaciones lineales de U a V . Demuestre que $\mathcal{L}(U, V)$ es un espacio vectorial sobre $\mathbb{K}$. ¿Qué se puede decir sobre $\dim \mathcal{L}(U, V)$?
7. Sean U, V espacios vectoriales sobre $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$). Sabemos de Ejercicio 6. que $\mathcal{L}(U, V)$ es un espacio vectorial. Fije un vector $v_0 \in V$. Demuestre que la siguiente función es una función lineal:

$$\Phi_{v_0} : \mathcal{L}(U, V) \rightarrow U, \quad \Phi_{v_0}(T) := T(v_0).$$

8. Sean

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad F = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

- Demuestre que E y F son invertibles. Describa como actúan geoméricamente en $\mathbb{R}^2$.
- Calcule $\text{Im}(A)$, $\ker(A)$ y sus dimensiones. Dibuja $\text{Im}(A)$ y $\ker(A)$, diga qué objetos geométricos son.
- Calcule $\text{Im}(A)$, $\text{Im}(FA)$, $\text{Im}(AE)$ y sus dimensiones. Dibújalos y diga cual es la relación entre ellos.
- Calcule $\ker(A)$, $\ker(FA)$, $\ker(AE)$ y sus dimensiones. Dibújalos y diga cual es la relación entre ellos.

9. De los siguientes matrices, calcule kernel, imagen y las dimensiones correspondientes.

$$A = \begin{pmatrix} 1 & 4 & 7 & 2 \\ 2 & 5 & 8 & 4 \\ 3 & 6 & 9 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 5 & 1 \\ 3 & 2 & 13 & 1 \\ 0 & 2 & 7 & -1 \\ 4 & 5 & 25 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 9 \end{pmatrix}.$$

10. Sea $A \in M(m \times n)$. Demuestre:

- A inyectiva $\implies m \geq n$.
- A sobreyectiva $\implies n \geq m$.

Demuestre que la implicación “ $\Leftarrow$ ” en (i) y (ii) en general es falsa.

11. Sea $A \in M(m \times n)$ y suponga que A es invertible. Demuestre que $m = n$.

12. Sea $A \in M(n \times 1)$ y $B \in M(1 \times n)$ ambas no nulas. Describa $\text{Im}(AB)$.

13. Sean $m, n \in \mathbb{N}$ y $A \in M(m \times n)$.

- ¿Cuáles son las dimensiones posibles de $\ker A$ y $\text{Im } A$?
- Para cada $j = 0, 1, 2, 3$ encuentre una matriz $A_j \in M(2 \times 3)$ con $\dim(\ker A_j) = j$, es decir: encuentre matrices A_0, A_1, A_2, A_3 con $\dim(\ker A_0) = 0$, $\dim(\ker A_1) = 1$, ... Si tal matriz no existe, explique por qué no existe.

14. (a) Encuentre una transformación lineal de $M(5 \times 5)$ a $M(3 \times 3)$ diferente de la transformación nula.

(b) Encuentre por lo menos dos diferentes funciones lineales biyectivas de $M(2 \times 2)$ a P_3 .

(c) Existe una función lineal biyectiva $S : M(2 \times 2) \rightarrow P_k$ para $k \in \mathbb{N}$, $k \neq 3$?

15. Sean V y W espacios vectoriales.

- Sea $U \subset V$ un subespacio y sean $u_1, \dots, u_k \in U$. Demuestre que $\text{gen}\{u_1, \dots, u_k\} \subset U$.

(b) Sean $u_1, \dots, u_k, w_1, \dots, w_m \in V$. Demuestre que lo siguiente es equivalente:

- 1) $\text{gen}\{u_1, \dots, u_k\} = \text{gen}\{w_1, \dots, w_m\}$.
- 2) Para todo $j = 1, \dots, k$ tenemos $u_j \in \text{gen}\{w_1, \dots, w_m\}$ y para todo $\ell = 1, \dots, m$ tenemos $w_\ell \in \text{gen}\{u_1, \dots, u_k\}$.
- 3) Sean $v_1, v_2, v_3, \dots, v_m \in V$ y sea $c \in \mathbb{R}$. Demuestre que $\text{gen}\{v_1, v_2, v_3, \dots, v_m\} = \text{gen}\{v_1 + cv_2, v_2, v_3, \dots, v_m\}$.

(c) Sean $v_1, \dots, v_k \in V$ y sea $A : V \rightarrow W$ una función lineal invertible. Demuestre que $\dim \text{gen}\{v_1, \dots, v_k\} = \dim \text{gen}\{Av_1, \dots, Av_k\}$. ¿Es verdad si A no es invertible?

16. (a) Sean $\vec{v}_1 = \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ y sea $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Demuestre que $\mathcal{B}$ es una base de $\mathbb{R}^3$ y escriba los vectores $\vec{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\vec{y} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ en términos de la base $\mathcal{B}$.

17. Sean $R = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$, $S = \begin{pmatrix} 3 & 2 \\ 0 & 7 \end{pmatrix}$, $T = \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix}$. Demuestre que $\mathcal{B} = \{R, S, T\}$ es una base del espacio de las matrices triangulares superiores y exprese las matrices

$$K = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, L = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

en términos de la base $\mathcal{B}$.

18. Sean $\vec{a}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\vec{a}_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $\vec{b}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $\vec{b}_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \in \mathbb{R}^2$ y sean $\mathcal{A} = \{\vec{a}_1, \vec{a}_2\}$, $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$.

(a) Demuestre que $\mathcal{A}$ y $\mathcal{B}$ son bases de $\mathbb{R}^2$.

(b) Sea $(\vec{x})_{\mathcal{A}} = \begin{pmatrix} 7 \\ 8 \end{pmatrix}$. Encuentre $(\vec{x})_{\mathcal{B}}$ y $\vec{x}$ (en la representación estándar).

(c) Sea $(\vec{y})_{\mathcal{B}} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$. Encuentre $(\vec{y})_{\mathcal{A}}$ y $\vec{y}$ (en la representación estándar).

19. Sea $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ una base de $\mathbb{R}^2$ y sean $\vec{x}_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, $\vec{x}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $\vec{x}_3 = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$ (dados en coordenadas cartesianas).

(a) Si se sabe que $\vec{x}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}_{\mathcal{B}}$, $\vec{x}_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}_{\mathcal{B}}$, es posible calcular $\vec{b}_1$ y $\vec{b}_2$? Si sí, calcúlelos. Si no, explique por qué no es posible.

(b) Si se sabe que $\vec{x}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}_{\mathcal{B}}$, $\vec{x}_3 = \begin{pmatrix} 6 \\ 2 \end{pmatrix}_{\mathcal{B}}$, es posible calcular $\vec{b}_1$ y $\vec{b}_2$? Si sí, calcúlelos. Si no, explique por qué no es posible.

- (c) ¿Existen $\vec{b}_1$ y $\vec{b}_2$ tal que $\vec{x}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}_{\mathcal{B}}$, $\vec{x}_2 = \begin{pmatrix} 6 \\ 2 \end{pmatrix}_{\mathcal{B}}$? Si sí, calcúlelos. Si no, explique por qué no es posible.
- (d) ¿Existen $\vec{b}_1$ y $\vec{b}_2$ tal que $\vec{x}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}_{\mathcal{B}}$, $\vec{x}_3 = \begin{pmatrix} 2 \\ 5 \end{pmatrix}_{\mathcal{B}}$? Si sí, calcúlelos. Si no, explique por qué no es posible.
20. (a) Demuestre que la siguiente función es lineal:
- $$\Phi : M(2 \times 2) \rightarrow M(2 \times 2), \quad \Phi(A) = A^t$$
- (b) Sea $\mathcal{B} = \{E_1, E_2, E_3, E_4\}$ la base estándar¹ de $M(2 \times 2)$. Encuentre la matriz que representa a Φ con respecto a esta base.
- (c) Sean $R = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $U = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ y sea $\mathcal{C} = \{R, S, T, U\}$. Demuestre que $\mathcal{C}$ es una base de $M(2 \times 2)$ y escriba Φ como matriz con respecto a esta base.
21. (a) Demuestre que $T : P_3 \rightarrow P_3$, $Tp = p'$ es una función lineal.
- (b) Determine $\ker(T)$, $\text{Im}(T)$, $\dim(\ker(T))$, $\dim(\text{Im}(T))$.
- (c) Sea $\mathcal{B} = \{1, X, X^2, X^3\}$ la base estándar de P_3 . Encuentre la matriz que representa a T con respecto a esta base.
- (d) Sean $q_1 = X+1$, $q_2 = X-1$, $q_3 = X^2+X$, $q_4 = X^3+1$. Demuestre que $\mathcal{C} = \{q_1, q_2, q_3, q_4\}$ es una base de P_3 .
- (e) Encuentre la matriz con respecto a la base $\mathcal{C}$ que representa a T .
22. Sean $T : P_3 \rightarrow P_4$ dada por $T(p) = \int_0^x p(t)dt$ y $D : P_4 \rightarrow P_3$ dada por $D(p) = p'$
- (a) Muestre que T , D son transformaciones lineales y para cada una encuentre su kernel, su imagen y las dimensiones del kernel y la imagen.
- (b) ¿Se cumple que $T(D(p)) = p$ para todo $p \in P_4$? En caso de que la respuesta sea negativa, ¿en cuáles casos se cumple?
- (c) Repetir lo del inciso anterior para $D(T(p))$ donde $p \in P_3$.
23. Sea $\vec{w} \in \mathbb{R}^n$ un vector no nulo. Muestre que existe $T : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que $T\vec{w} \neq 0$. Calcule $\dim(\ker T)$ y $\dim(\text{Im } T)$.
24. En $\mathbb{R}^n$, sea $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ una transformación lineal diferente de la trivial.
- (a) Muestre que existe $\vec{w} \in \mathbb{R}^n$ tal que $\varphi(\vec{x}) = \langle \vec{x}, \vec{w} \rangle$ para todo $\vec{x} \in \mathbb{R}^n$. ¿Cuál es la dimensión de $\ker \varphi$? ¿Si $n = 2$ ó $n = 3$ como luce $\ker \varphi$?
- (b) Sean $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$ y $v_3 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$. Encuentre algún $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ tal que $\varphi(\vec{v}_1)$, $\varphi(\vec{v}_2)$ y $\varphi(\vec{v}_3)$ son todos diferentes de 0.
- (c) Sean $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ vectores de $\mathbb{R}^2$ todos distintos de $\vec{0}$. Muestre que existe $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ tal que φ no se anula en ninguno de ellos.

¹ $E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

DRAFT

Capítulo 7

Bases ortonormales y proyecciones ortogonales en $\mathbb{R}^n$

En este capítulo trabajaremos en $\mathbb{R}^n$ y **no** en espacios vectoriales arbitrarios, ya que queremos explorar con más detalle sus propiedades geométricas. En particular discutiremos la ortogonalidad. Observe que en un espacio vectorial arbitrario no tenemos el concepto de ángulos ni de ortogonalidad. Todo lo que discutiremos aquí puede extenderse a espacios con producto interno, donde el producto interno se usa para *definir* ángulos. Recordemos que mostramos en el Teorema 2.21 que, para vectores no nulos $\vec{x}, \vec{y} \in \mathbb{R}^n$, el ángulo φ entre ellos satisface la ecuación

$$\cos \varphi = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \|\vec{y}\|}.$$

En un espacio con producto interno general $(V, \langle \cdot, \cdot \rangle)$, esta ecuación se usa para definir el ángulo entre dos vectores. En particular, se dice que dos vectores son ortogonales si su producto interno es 0. Los espacios con producto interno son útiles, por ejemplo, en física, y tal vez en un futuro no tan lejano haya un capítulo en estas notas de clase sobre ellos.

Primero definiremos qué es el complemento ortogonal de un subespacio de $\mathbb{R}^n$ y veremos que la suma directa de un subespacio y su complemento ortogonal nos da todo $\mathbb{R}^n$.

Ya sabemos qué es la proyección ortogonal de un vector $\vec{x}$ sobre otro vector $\vec{y} \neq \vec{0}$ (véase la Sección 2.3). Como es independiente de la norma de $\vec{y}$, también podemos considerarla como la proyección ortogonal de $\vec{x}$ sobre la recta generada por $\vec{y}$. En este capítulo generalizaremos el concepto de proyección ortogonal sobre una recta a la proyección ortogonal sobre un subespacio arbitrario.

Como aplicación, discutiremos el método de mínimos cuadrados para la aproximación de datos.

7.1. Sistemas ortonormales y bases ortogonales

Recordemos que dos vectores $\vec{x}$ y $\vec{y}$ son *ortogonales* (o *perpendiculares*) entre sí si y solo si $\langle \vec{x}, \vec{y} \rangle = 0$. En este caso escribimos $\vec{x} \perp \vec{y}$.

Definición 7.1. (I) Un conjunto de vectores $\vec{x}_1, \dots, \vec{x}_k \in \mathbb{R}^n$ se llama un *conjunto ortogonal* si son ortogonales dos a dos; en fórmulas podemos escribir esto como

$$\langle \vec{x}_j, \vec{x}_\ell \rangle = 0 \quad \text{para } j \neq \ell.$$

(II) Un conjunto de vectores $\vec{x}_1, \dots, \vec{x}_k \in \mathbb{R}^n$ se llama un *conjunto ortonormal* si son ortonormales dos a dos; en fórmulas podemos escribir esto como

$$\langle \vec{x}_j, \vec{x}_\ell \rangle = \begin{cases} 1 & \text{para } j = \ell, \\ 0 & \text{para } j \neq \ell. \end{cases}$$

La diferencia entre un conjunto ortogonal y uno ortonormal es que en este último requerimos además que cada vector del conjunto satisfaga $\langle \vec{x}_j, \vec{x}_j \rangle = 1$, es decir, que $\|\vec{x}_j\| = 1$. Por lo tanto, un conjunto ortogonal puede contener vectores de longitudes arbitrarias, incluyendo el vector $\vec{0}$, mientras que en un conjunto ortonormal todos los vectores deben tener longitud 1. Observe que todo sistema ortonormal también es un sistema ortogonal. Por otro lado, todo sistema ortogonal que no contenga $\vec{0}$ puede convertirse en uno ortonormal normalizando cada vector (es decir, dividiendo cada vector por su norma).

Ejemplos 7.2. (I) Los siguientes sistemas son sistemas ortogonales, pero no sistemas ortonormales, ya que la norma de al menos uno de sus vectores es diferente de 1:

$$\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right\}.$$

(II) Los siguientes sistemas son sistemas ortonormales:

$$\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right\}.$$

Lema 7.3. *Todo sistema ortonormal es linealmente independiente.*

Demostración. Sea $\vec{x}_1, \dots, \vec{x}_k$ un sistema ortonormal y consideremos

$$\vec{0} = \alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \dots + \alpha_{n-1} \vec{x}_{n-1} + \alpha_n \vec{x}_n.$$

Tenemos que mostrar que todos los α_j deben ser cero. Para hacer esto, tomamos el producto interno en ambos lados con los vectores $\vec{x}_j$. Comencemos con $\vec{x}_1$. Obtenemos

$$\begin{aligned} \langle \vec{0}, \vec{x}_1 \rangle &= \langle \alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \dots + \alpha_{n-1} \vec{x}_{n-1} + \alpha_n \vec{x}_n, \vec{x}_1 \rangle \\ &= \alpha_1 \langle \vec{x}_1, \vec{x}_1 \rangle + \alpha_2 \langle \vec{x}_2, \vec{x}_1 \rangle + \dots + \alpha_{n-1} \langle \vec{x}_{n-1}, \vec{x}_1 \rangle + \alpha_n \langle \vec{x}_n, \vec{x}_1 \rangle. \end{aligned}$$

Como $\langle \vec{0}, \vec{x}_1 \rangle = 0$, $\langle \vec{x}_1, \vec{x}_1 \rangle = \|\vec{x}_1\|^2 = 1$ y $\langle \vec{x}_2, \vec{x}_1 \rangle = \dots = \langle \vec{x}_{n-1}, \vec{x}_1 \rangle = \langle \vec{x}_n, \vec{x}_1 \rangle = 0$, se sigue que

$$0 = \alpha_1 + 0 + \dots + 0 = \alpha_1.$$

Ahora podemos repetir este proceso con $\vec{x}_2, \vec{x}_3, \dots, \vec{x}_n$ para mostrar que $\alpha_2 = \dots = \alpha_n = 0$. $\square$

Observación. El lema muestra que todo sistema ortogonal de n vectores en $\mathbb{R}^n \setminus \{\vec{0}\}$ es una base de $\mathbb{R}^n$.

Definición 7.4. Una base ortonormal de $\mathbb{R}^n$ es una base cuyos vectores forman además un conjunto ortogonal. Ocasionalmente escribiremos ONB por “orthonormal basis”.

Ejemplos 7.5 (Bases ortonormales de $\mathbb{R}^n$).

- (I) La base canónica $\vec{e}_1, \dots, \vec{e}_n$ es una base ortonormal de $\mathbb{R}^n$.
- (II) Los siguientes sistemas son ejemplos de bases ortonormales de $\mathbb{R}^2$:

$$\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}, \quad \left\{ \frac{1}{\sqrt{13}} \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \frac{1}{\sqrt{13}} \begin{pmatrix} -3 \\ 2 \end{pmatrix} \right\}, \quad \left\{ \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \frac{1}{5} \begin{pmatrix} -4 \\ 3 \end{pmatrix} \right\}.$$

- (III) Los siguientes sistemas son ejemplos de bases ortonormales de $\mathbb{R}^3$:

$$\left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \right\}, \quad \left\{ \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \frac{1}{\sqrt{10}} \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{35}} \begin{pmatrix} 1 \\ -5 \\ 3 \end{pmatrix} \right\}.$$

Ejercicio 7.6. Muestre que toda base ortonormal de $\mathbb{R}^2$ es de la forma $\left\{ \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} \right\}$ o $\left\{ \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \begin{pmatrix} \sin \varphi \\ -\cos \varphi \end{pmatrix} \right\}$ para algún $\varphi \in \mathbb{R}$. Véase también el Ejercicio 7.13.

Veremos en el Corolario 7.27 que todo sistema ortonormal en $\mathbb{R}^n$ puede completarse a una base ortonormal. En la Sección 7.5 mostraremos cómo construir una base ortonormal de un subespacio de $\mathbb{R}^n$ a partir de una base dada. En particular se sigue que todo subespacio de $\mathbb{R}^n$ tiene una base ortonormal.

Las bases ortonormales son muy útiles. Entre otras cosas, es muy fácil escribir un vector dado $\vec{w} \in \mathbb{R}^n$ como una combinación lineal de una base de este tipo. Recordemos que, si se nos da una base arbitraria $\vec{z}_1, \dots, \vec{z}_n$ de $\mathbb{R}^n$ y queremos escribir un vector $\vec{x}$ como combinación lineal de esta base, entonces tenemos que encontrar coeficientes $\alpha_1, \dots, \alpha_n$ tales que $\vec{x} = \alpha_1 \vec{z}_1 + \dots + \alpha_n \vec{z}_n$, lo cual significa que debemos resolver un sistema de tamaño $n \times n$ para determinar los coeficientes. Sin embargo, si la base dada es una base ortonormal, entonces calcular los coeficientes se reduce a evaluar n productos internos, como muestra el siguiente teorema.

Teorema 7.7 (Representación de un vector con respecto a una ONB). Sea $\vec{x}_1, \dots, \vec{x}_n$ una base ortonormal de $\mathbb{R}^n$ y sea $\vec{w} \in \mathbb{R}^n$. Entonces

$$\vec{w} = \langle \vec{w}, \vec{x}_1 \rangle \vec{x}_1 + \langle \vec{w}, \vec{x}_2 \rangle \vec{x}_2 + \dots + \langle \vec{w}, \vec{x}_n \rangle \vec{x}_n.$$

Demostración. Como $\vec{x}_1, \dots, \vec{x}_n$ es una base de $\mathbb{R}^n$, existen $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tales que

$$\vec{w} = \alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \dots + \alpha_n \vec{x}_n.$$

Ahora tomemos el producto interno en ambos lados con $\vec{x}_j$ para $j = 1, \dots, n$. Observe que $\langle \vec{x}_k, \vec{x}_j \rangle = 0$ si $k \neq j$ y que $\langle \vec{x}_j, \vec{x}_j \rangle = \|\vec{x}_j\|^2 = 1$.

$$\begin{aligned}\langle \vec{w}, \vec{x}_j \rangle &= \langle \alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \dots + \alpha_n \vec{x}_n, \vec{x}_j \rangle \\ &= \alpha_1 \langle \vec{x}_1, \vec{x}_j \rangle + \alpha_2 \langle \vec{x}_2, \vec{x}_j \rangle + \dots + \alpha_n \langle \vec{x}_n, \vec{x}_j \rangle \\ &= \alpha_j \langle \vec{x}_j, \vec{x}_j \rangle = \alpha_j.\end{aligned}$$

□

Note que la demostración de este teorema es esencialmente la misma que la del Lema 7.3. De hecho, el Lema 7.3 se sigue del teorema anterior si elegimos $\vec{w} = \vec{0}$.

Ejercicio 7.8. Si $\vec{x}_1, \dots, \vec{x}_n$ es una base ortogonal, pero no necesariamente ortonormal, de $\mathbb{R}^n$, entonces para todo $\vec{w} \in \mathbb{R}^n$ tenemos que

$$\vec{w} = \frac{\langle \vec{w}, \vec{x}_1 \rangle}{\|\vec{x}_1\|^2} \vec{x}_1 + \frac{\langle \vec{w}, \vec{x}_2 \rangle}{\|\vec{x}_2\|^2} \vec{x}_2 + \dots + \frac{\langle \vec{w}, \vec{x}_n \rangle}{\|\vec{x}_n\|^2} \vec{x}_n.$$

(Puede usar una versión modificada de la demostración del Teorema 7.7, o bien definir $y_j = \|\vec{x}_j\|^{-1} \vec{x}_j$, mostrar que $\vec{y}_1, \dots, \vec{y}_n$ es una base ortogonal y aplicar la fórmula del Teorema 7.7.)

Ahora debería haber entendido

- qué es un sistema ortogonal,
- qué es un sistema ortonormal,
- qué es una base ortonormal,
- por qué las bases ortogonales son útiles,
- etc.

Ahora debería ser capaz de

- verificar si un conjunto dado de vectores es un sistema ortogonal/ortonormal,
- verificar si un conjunto dado de vectores es una base ortogonal/ortonormal del espacio dado,
- verificar si una base dada es una base ortogonal u ortonormal,
- dar ejemplos de bases ortonormales,
- encontrar los coeficientes de un vector dado con respecto a una base ortonormal u ortogonal dada.
- etc.

Ejercicios.

1. De los ejercicios anteriores, verifique si el conjunto dado es una base ortonormal del espacio vectorial V al que se refiere.

$$(a) \quad V = \mathbb{R}^2, \quad \left\{ \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\}.$$

(b) $V = \mathbb{R}^2$, $\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$.

(c) En $\mathbb{R}^2$ considerar $V =$ la recta $3x - 2y = 0$, $\left\{ \frac{1}{\sqrt{13}} \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\}$.

(d) En $\mathbb{R}^3$ considerar $V = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} \right\}$, $\left\{ \frac{1}{\sqrt{17}} \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{357}} \begin{pmatrix} 8 \\ -2 \\ -17 \end{pmatrix} \right\}$.

(e) En $\mathbb{R}^3$ considerar $V = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 6 \\ 17 \end{pmatrix} \right\}$, $\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \right\}$.

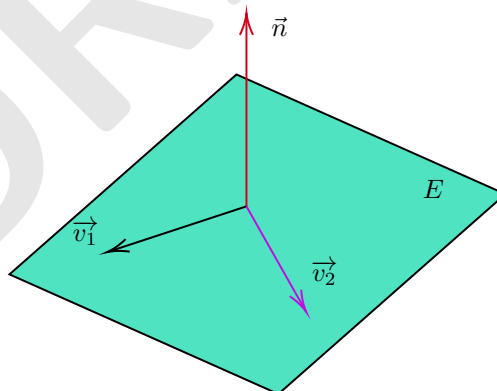
2. ¿Para qué valores de a, b es el conjunto una base ortogonal de $\mathbb{R}^3$?

$$\left\{ \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ a \\ 2 \end{pmatrix}, \begin{pmatrix} b-5a \\ 4 \\ 1 \end{pmatrix} \right\}$$

3. En $\mathbb{R}^2$, sea $\vec{v}_1 = \begin{pmatrix} a \\ b \end{pmatrix}$ un vector no nulo. ¿Cuántas bases ortogonales de $\mathbb{R}^2$ que contienen a $\vec{v}_1$ existen? ¿Cuántas bases ortogonales $\{\vec{v}_1, \vec{v}_2\}$ de $\mathbb{R}^2$ existen tales que $\|\vec{v}_1\| = \|\vec{v}_2\|$?

4. El siguiente ejercicio pretende obtener una base ortonormal del plano $E : ax + by + cz = 0$ con herramientas vistas hasta ahora.

(a) Considere un vector $\vec{v}_1$ paralelo a E con $\vec{v}_1 \neq \vec{0}$. Sea $\vec{n}$ algún vector normal de E y tome $\vec{v}_2 = \vec{v}_1 \times \vec{n}$. Demuestre $\left\{ \frac{\vec{v}_1}{\|\vec{v}_1\|}, \frac{\vec{v}_2}{\|\vec{v}_2\|} \right\}$ es una base ortonormal de E , (observe que $\left\{ \frac{\vec{v}_1}{\|\vec{v}_1\|}, \frac{\vec{v}_2}{\|\vec{v}_2\|}, \frac{\vec{n}}{\|\vec{n}\|} \right\}$ es una base ortonormal de $\mathbb{R}^3$).



(b) Para el plano $E : x + 2y + 3z = 0$, obtenga una base ortonormal y complétela a una base ortonormal de $\mathbb{R}^3$.

(c) Escriba $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ en términos de la base que obtuvo del inciso anterior.

(d) Sea $L : \frac{x}{3} = \frac{y}{2} = \frac{z}{5}$. ¿Puede obtener una base ortonormal de $\mathbb{R}^3$ que contenga algún vector director de L ?

5. Sea $\mathcal{B}$ cualquier base de $\mathbb{R}^n$ y sean $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$. Si $\langle [\vec{v}_1]_{\mathcal{B}}, [\vec{v}_2]_{\mathcal{B}} \rangle = 0$ ¿se sigue que $\vec{v}_1 \perp \vec{v}_2$?

7.2. Matrices ortogonales

Ya vimos que es muy fácil expresar un vector dado como combinación lineal de los elementos de una base ortonormal. En esta sección queremos explorar las propiedades de las matrices de transición entre dos bases ortonormales de $\mathbb{R}^n$.

Sean $\mathcal{B} = \{\vec{u}_1, \dots, \vec{u}_n\}$ y $\mathcal{C} = \{\vec{w}_1, \dots, \vec{w}_n\}$ bases ortonormales de $\mathbb{R}^n$. Sea $Q = A_{\mathcal{B} \rightarrow \mathcal{C}}$ la matriz de transición de la base $\mathcal{B}$ a la base $\mathcal{C}$. Sabemos que sus entradas q_{ij} son los números determinados de manera única tales que

$$\vec{u}_1 = \begin{pmatrix} q_{11} \\ \vdots \\ q_{n1} \end{pmatrix}_{\mathcal{C}} = q_{11}\vec{w}_1 + \dots + q_{n1}\vec{w}_n, \quad \dots, \quad \vec{u}_n = \begin{pmatrix} q_{1n} \\ \vdots \\ q_{nn} \end{pmatrix}_{\mathcal{C}} = q_{1n}\vec{w}_1 + \dots + q_{nn}\vec{w}_n.$$

Como $\mathcal{C}$ es una base ortonormal, se sigue que $q_{ij} = \langle \vec{u}_j, \vec{w}_i \rangle$; véase el Teorema 7.7. Por lo tanto

$$A_{\mathcal{B} \rightarrow \mathcal{C}} = \begin{pmatrix} \langle \vec{u}_1, \vec{w}_1 \rangle & \langle \vec{u}_2, \vec{w}_1 \rangle & \dots & \langle \vec{u}_n, \vec{w}_1 \rangle \\ \langle \vec{u}_1, \vec{w}_2 \rangle & \langle \vec{u}_2, \vec{w}_2 \rangle & \dots & \langle \vec{u}_n, \vec{w}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{u}_1, \vec{w}_n \rangle & \langle \vec{u}_2, \vec{w}_n \rangle & \dots & \langle \vec{u}_n, \vec{w}_n \rangle \end{pmatrix}.$$

Si intercambiamos el papel de $\mathcal{B}$ y $\mathcal{C}$ y usamos que $\langle \vec{w}_i, \vec{u}_j \rangle = \langle \vec{u}_j, \vec{w}_i \rangle$, entonces obtenemos

$$A_{\mathcal{C} \rightarrow \mathcal{B}} = \begin{pmatrix} \langle \vec{w}_1, \vec{u}_1 \rangle & \langle \vec{w}_2, \vec{u}_1 \rangle & \dots & \langle \vec{w}_n, \vec{u}_1 \rangle \\ \langle \vec{w}_1, \vec{u}_2 \rangle & \langle \vec{w}_2, \vec{u}_2 \rangle & \dots & \langle \vec{w}_n, \vec{u}_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{w}_1, \vec{u}_n \rangle & \langle \vec{w}_2, \vec{u}_n \rangle & \dots & \langle \vec{w}_n, \vec{u}_n \rangle \end{pmatrix} = \begin{pmatrix} \langle \vec{u}_1, \vec{w}_1 \rangle & \langle \vec{u}_1, \vec{w}_2 \rangle & \dots & \langle \vec{u}_1, \vec{w}_n \rangle \\ \langle \vec{u}_2, \vec{w}_1 \rangle & \langle \vec{u}_2, \vec{w}_2 \rangle & \dots & \langle \vec{u}_2, \vec{w}_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{u}_n, \vec{w}_1 \rangle & \langle \vec{u}_n, \vec{w}_2 \rangle & \dots & \langle \vec{u}_n, \vec{w}_n \rangle \end{pmatrix}.$$

Esto muestra que $A_{\mathcal{C} \rightarrow \mathcal{B}} = (A_{\mathcal{B} \rightarrow \mathcal{C}})^t$. Si usamos que $A_{\mathcal{C} \rightarrow \mathcal{B}} = (A_{\mathcal{B} \rightarrow \mathcal{C}})^{-1}$, entonces encontramos que

$$(A_{\mathcal{B} \rightarrow \mathcal{C}})^{-1} = (A_{\mathcal{B} \rightarrow \mathcal{C}})^t.$$

De estos cálculos obtenemos el siguiente lema.

Lema 7.9. Sean $\mathcal{B} = \{\vec{u}_1, \dots, \vec{u}_n\}$ y $\mathcal{C} = \{\vec{w}_1, \dots, \vec{w}_n\}$ bases ortonormales de $\mathbb{R}^n$ y sea $Q = A_{\mathcal{B} \rightarrow \mathcal{C}}$ la matriz de transición de la base $\mathcal{B}$ a la base $\mathcal{C}$. Entonces

$$Q^t = Q^{-1}.$$

Definición 7.10. Una matriz $A \in M(n \times n)$ se llama *matriz ortogonal* si es invertible y $A^t = A^{-1}$.

Proposición 7.11. Sea $Q \in M(n \times n)$. Entonces las siguientes condiciones son equivalentes:

- (I) Q es una matriz ortogonal.
- (II) Q^t es una matriz ortogonal.
- (III) Q^{-1} existe y es una matriz ortogonal.

Demostración. (I) $\implies$ (II): Supongamos que Q es ortogonal. Entonces es invertible; por lo tanto, también Q^t es invertible por el Teorema 3.51 y $(Q^t)^{-1} = (Q^{-1})^t = (Q^t)^t = Q$ se cumple. Por lo tanto, Q^t es una matriz ortogonal.

(II) $\implies$ (I): Supongamos que Q^t es una matriz ortogonal. Entonces $(Q^t)^t = Q$ must be an ortogonal matrix too by what we just proved.

(I) $\implies$ (III): Supongamos que Q es ortogonal. Entonces es invertible y $(Q^{-1})^{-1} = (Q^t)^{-1} = (Q^{-1})^t$ donde en el segundo paso usamos el Teorema 3.51. Por lo tanto, Q^{-1} es una matriz ortogonal.

(III) $\implies$ (I): Supongamos que Q^{-1} es una matriz ortogonal. Entonces its inverse $(Q^{-1})^{-1} = Q$ must be an ortogonal matrix too by what we just proved. $\square$

Por el Lema 7.9, toda matriz de transición de una ONB a otra ONB es una matriz ortogonal. El recíproco también es cierto, como muestra el siguiente teorema.

Teorema 7.12. Sea $Q \in M(n \times n)$. Entonces:

- (I) Q es una matriz ortogonal si y solo si sus columnas son una base ortonormal de $\mathbb{R}^n$.
- (II) Q es una matriz ortogonal si y solo si sus filas son una base ortonormal de $\mathbb{R}^n$.
- (III) If Q es una matriz ortogonal, entonces $|\det Q| = 1$.

Demostración. (I) Supongamos que Q es una matriz ortogonal y sean $\vec{c}_j$ sus columnas. Ya sabemos que foman una base de $\mathbb{R}^n$, ya que Q es invertible. Para mostrar que también son un sistema ortonormal, calculamos

$$\text{id} = Q^t Q = \begin{pmatrix} \vec{c}_1 \\ \vdots \\ \vec{c}_n \end{pmatrix} (\vec{c}_1 \mid \cdots \mid \vec{c}_n) = \begin{pmatrix} \langle \vec{c}_1, \vec{c}_1 \rangle & \langle \vec{c}_1, \vec{c}_2 \rangle & \cdots & \langle \vec{c}_1, \vec{c}_n \rangle \\ \langle \vec{c}_2, \vec{c}_1 \rangle & \langle \vec{c}_2, \vec{c}_2 \rangle & \cdots & \langle \vec{c}_2, \vec{c}_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{c}_n, \vec{c}_1 \rangle & \langle \vec{c}_n, \vec{c}_2 \rangle & \cdots & \langle \vec{c}_n, \vec{c}_n \rangle \end{pmatrix}. \quad (7.1)$$

Como el producto es igual a la matriz identidad, se sigue que todos los elementos de la diagonal deben ser iguales a 1 y todos los demás elementos deben ser iguales a 0. Esto significa que $\langle \vec{c}_j, \vec{c}_j \rangle = 1$ para $j = 1, \dots, n$ y $\langle \vec{c}_j, \vec{c}_k \rangle = 0$ para $j \neq k$; por lo tanto, las columnas de Q son una base ortonormal de $\mathbb{R}^n$.

Ahora supongamos que las columnas $\vec{c}_1, \dots, \vec{c}_n$ de Q son una base ortonormal de $\mathbb{R}^n$. Entonces clearly (7.1) holds which shows that Q es una matriz ortogonal.

(II) Las filas de Q son las columnas de Q^t ; por lo tanto, son una base ortonormal de $\mathbb{R}^n$ por (I) y la Proposición 7.11 (II).

(III) Recordemos que $\det Q^t = \det Q$. Por lo tanto obtenemos

$$1 = \det \text{id} = \det(QQ^t) = (\det Q)(\det Q^t) = (\det Q)^2,$$

lo cual prueba la afirmación. $\square$

Claramente, no toda matriz R con $|\det R| = 1$ es una matriz ortogonal. Por ejemplo, si $R = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, entonces $\det R = 1$, pero $R^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ es diferente de $R^t = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

Question 7.1

Suponga que $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{R}^n$ son ortogonales dos a dos y sea $R \in M(n \times n)$ la matriz cuyas columnas son estos vectores. ¿Puede calcular $R^t R$ y $R R^t$? ¿Cuáles son las condiciones sobre los vectores para que R sea invertible? Si es invertible, ¿cuál es su inversa? (Debería poder responder las preguntas anteriores más o menos fácilmente si $\|\vec{a}_j\| = 1$ para todo $j = 1, \dots, n$, porque en este caso R es una matriz ortogonal.)

Ejercicio 7.13. Muestre que toda matriz ortogonal de tamaño 2×2 es de la forma $Q = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$ o $Q = \begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix}$. Compare esto con el Ejercicio 7.6.

Ejercicio 7.14. Use los resultados de la Sección 4.3 para probar que $|\det Q| = 1$ si Q es una matriz ortogonal de tamaño 2×2 o 3×3 .

Se puede mostrar que toda matriz ortogonal representa una rotación (si su determinante es 1) o la composición de una rotación y una reflexión (si su determinante es -1).

Matrices ortogonales en $\mathbb{R}^2$. Sea $Q \in M(2 \times 2)$ una matriz ortogonal con columnas $\vec{c}_1$ y $\vec{c}_2$. Recordemos que $Q\vec{e}_1 = \vec{c}_1$ y $Q\vec{e}_2 = \vec{c}_2$. Como $\vec{c}_1$ es un vector unitario, es de la forma $\vec{c}_1 = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$ para algún $\varphi \in \mathbb{R}$. Como $\vec{c}_2$ también es un vector unitario y además debe ser ortogonal a $\vec{c}_1$, solo existen las dos opciones posibles $\vec{c}_2^+ = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}$ o $\vec{c}_2^- = \begin{pmatrix} \sin \varphi \\ -\cos \varphi \end{pmatrix}$, véase la Figura 7.1.

- En el primer caso, $\det Q = \det(\vec{c}_1 | \vec{c}_2^+) = \det \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} = \cos^2 \varphi + \sin^2 \varphi = 1$ y Q representa la rotación por φ en sentido antihorario.
- En el segundo caso, $\det Q = \det(\vec{c}_1 | \vec{c}_2^-) = \det \begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix} = -\cos^2 \varphi - \sin^2 \varphi = -1$ y Q representa la rotación por φ en el sentido antihorario, seguida por una reflexión en la dirección dada por $\vec{c}_1$ (o: reflexión en el eje x seguida por la rotación por φ en el sentido antihorario).

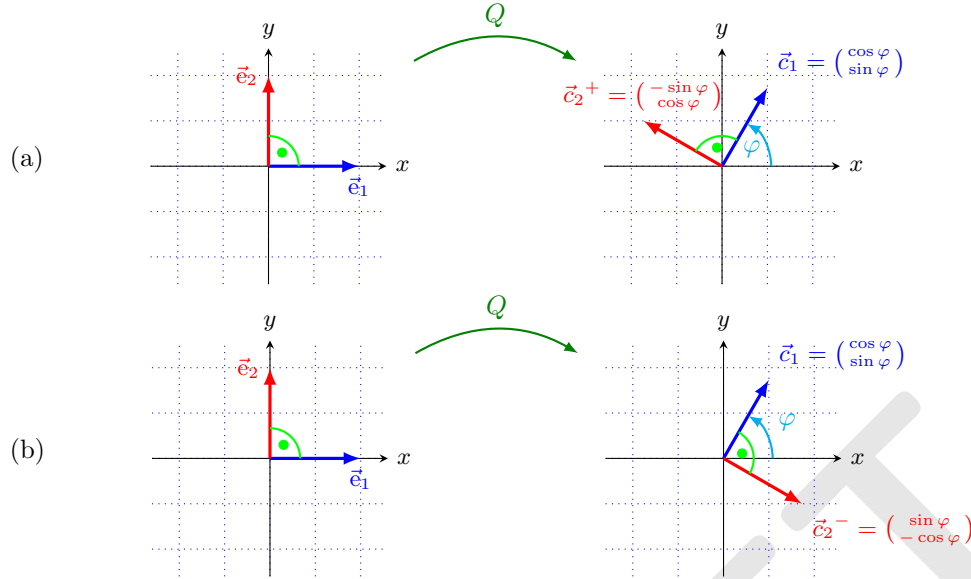


FIGURA 7.1: En el caso (a), Q representa una rotación y $\det A = 1$. En el caso (b), representa una rotación seguida de una reflexión y $\det Q = -1$.

Ejercicio 7.15. Sea Q una matriz ortogonal de tamaño $n \times n$. Muestre lo siguiente.

- (I) Q preserva productos internos, es decir, $\langle \vec{x}, \vec{y} \rangle = \langle Q\vec{x}, Q\vec{y} \rangle$ para todo $\vec{x}, \vec{y} \in \mathbb{R}^n$.
- (II) Q preserva lengths, es decir, $\|\vec{x}\| = \|Q\vec{x}\|$ para todo $\vec{x} \in \mathbb{R}^n$.
- (III) Q preserva angles, es decir, $\angle(\vec{x}, \vec{y}) = \angle(Q\vec{x}, Q\vec{y})$ para todo $\vec{x}, \vec{y} \in \mathbb{R}^n \setminus \{\vec{0}\}$.

Ejercicio 7.16. Sea $Q \in M(n \times n)$

- (I) Suponga que Q preserva productos internos, es decir, $\langle \vec{x}, \vec{y} \rangle = \langle Q\vec{x}, Q\vec{y} \rangle$ para todo $\vec{x}, \vec{y} \in \mathbb{R}^n$. Muestre que Q es una matriz ortogonal.
- (II) Suponga que Q preserva longitudes, es decir, $\|\vec{x}\| = \|Q\vec{x}\|$ para todo vector. Muestre que Q es una matriz ortogonal.

El Ejercicio 7.15, junto con el Ejercicio 7.16, muestra lo siguiente.

Una matriz Q es una matriz ortogonal si y solo si preserva longitudes, si y solo si preserva ángulos. Es decir,

$$\begin{aligned}
 Q \text{ es ortogonal} &\iff Q^t = Q^{-1} \\
 &\iff \langle Q\vec{x}, Q\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle \quad \text{fo all } \vec{x}, \vec{y} \in \mathbb{R}^n \\
 &\iff \|Q\vec{x}\| = \|\vec{x}\| \quad \text{fo all } \vec{x} \in \mathbb{R}^n.
 \end{aligned}$$

Definición 7.17. Una transformación lineal $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ se llama una *isometría* si $\|T\vec{x}\| = \|\vec{x}\|$ para todo $\vec{x} \in \mathbb{R}^n$.

Observe que toda isometría es inyectiva, ya que $T\vec{x} = \vec{0}$ si y solo si $\vec{x} = \vec{0}$; por lo tanto, necesariamente $n \leq m$.

Ahora debería haber entendido

- que una matriz es ortogonal si y solo si representa un cambio de bases entre dos bases ortonormales,
- que una matriz ortogonal representa una composición de rotaciones y reflexiones,
- etc.

Ahora debería ser capaz de

- verificar si una matriz dada es una matriz ortogonal,
- construir matrices ortogonales,
- etc.

Ejercicios.

1. Verifique que las siguientes matrices son ortogonales.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \vartheta & -\sin \vartheta \\ 0 & \sin \vartheta & \cos \vartheta \end{pmatrix}, \quad \begin{pmatrix} \cos \vartheta & 0 & -\sin \vartheta \\ 0 & 1 & 0 \\ \sin \vartheta & 0 & \cos \vartheta \end{pmatrix}, \quad \begin{pmatrix} \cos \vartheta & -\sin \vartheta & 0 \\ \sin \vartheta & \cos \vartheta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

¿Cuál es la interpretación geométrica de cada una? Ver 3.4 ejercicio 6..

2. Para el plano $E : 2x + y - z = 0$, obtenga una base $\mathcal{B}$ ortonormal de $\mathbb{R}^3$ tal que sus dos primeros vectores sean una base de E . Obtenga $A_{\mathcal{B} \rightarrow \text{can}}$ y $A_{\text{can} \rightarrow \mathcal{B}}$.
3. Encuentre por lo menos seis isometrías distintas de $\mathbb{R}^2$ a $\mathbb{R}^3$.
4. Sean $A, B \in M(n \times n)$:
 - (a) Si AB es ortogonal. ¿Se puede concluir que A y B deben ser matrices ortogonales?
 - (b) Si A, B son matrices ortogonales. ¿Se puede concluir que AB es una matriz ortogonal?
5. Sea $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $Q \in M(n \times n)$
 - (a) Demuestre que T es una isometría si y solo si $\langle Tx, Ty \rangle = \langle x, y \rangle$ para todo $x, y \in \mathbb{R}^n$ (por ende T preserva ángulos). (Hint: Basta hacer lo mismo que en el ejercicio 7.16 parte (ii)).
 - (b) Muestre que Q es una matriz ortogonal si y solo si Q es una isometría
 - (c) Sea $\mathcal{B}_1 = \{\vec{e}_1, \dots, \vec{e}_n\}$ la base canónica de $\mathbb{R}^n$ y suponga que T es una isometría. Muestre que $\{T\vec{e}_1, \dots, T\vec{e}_n\}$ es un sistema ortonormal de vectores.
6. Sea $\vec{x} \in \mathbb{R}^n$. Muestre que $T(\vec{x}) = \langle x, \vec{e}_1 \rangle T\vec{e}_1 + \dots + \langle x, \vec{e}_n \rangle T\vec{e}_n$.

7.3. Complemento ortogonal

En esta sección aprenderemos cómo encontrar todos los vectores que son ortogonales a un subespacio dado U de $\mathbb{R}^n$. Este conjunto se llama el *complemento ortogonal* de U . Comenzamos con su definición formal.

Definición 7.18. Sea U un subespacio de $\mathbb{R}^n$.

- (I) Sea U un subespacio de $\mathbb{R}^n$. Decimos que un vector $\vec{x} \in \mathbb{R}^n$ es *perpendicular a U* si es perpendicular a todo vector en U . En este caso escribimos $\vec{x} \perp U$.
- (II) El *complemento ortogonal* de U se denota por $U^\perp$ y es el conjunto de todos los vectores que son perpendiculares a todo vector en U , es decir

$$U^\perp = \{\vec{x} \in \mathbb{R}^n : \vec{x} \perp U\} = \{\vec{x} \in \mathbb{R}^n : \vec{x} \perp \vec{u} \text{ para todo } \vec{u} \in U\}.$$

Comenzamos con algunas observaciones sencillas.

Observación 7.19. Sea U un subespacio de $\mathbb{R}^n$.

- (I) $U^\perp$ es un subespacio de $\mathbb{R}^n$.
- (II) $U \cap U^\perp = \{\vec{0}\}$.
- (III) $(\mathbb{R}^n)^\perp = \{\vec{0}\}$ y $\{\vec{0}\}^\perp = \mathbb{R}^n$.

En Teorema 7.25 se probará que $(U^\perp)^\perp = U$.

Demostración. (I) Claramente, $\vec{0} \in U^\perp$. Sean $\vec{x}, \vec{y} \in U^\perp$ y sea $c \in \mathbb{R}$. Entonces, para todo $\vec{u} \in U$, tenemos que $\langle \vec{x} + c\vec{y}, \vec{u} \rangle = \langle \vec{x}, \vec{u} \rangle + c\langle \vec{y}, \vec{u} \rangle = 0$, por lo tanto $\vec{x} + c\vec{y} \in U^\perp$ y $U^\perp$ es un subespacio por el Teorema 5.10.

(II) Sea $\vec{x} \in U \cap U^\perp$. Entonces se sigue que $\vec{x} \perp \vec{x}$, por lo tanto $\|\vec{x}\|^2 = \langle \vec{x}, \vec{x} \rangle = 0$ lo cual muestra que $\vec{x} = \vec{0}$ y por lo tanto $U \cap U^\perp$ contiene solamente el vector $\vec{0}$.

(III) La inclusión $\{\vec{0}\} \subseteq (\mathbb{R}^n)^\perp$ es clara. Para mostrar igualdad, tome $\vec{x} \in (\mathbb{R}^n)^\perp$. Entonces $\vec{x} \perp \vec{y}$ para todo $\vec{y} \in \mathbb{R}^n$, en particular también $\vec{x} \perp \vec{x}$. Por lo tanto $\|\vec{x}\|^2 = \langle \vec{x}, \vec{x} \rangle = 0$, lo cual muestra que $\vec{x} = \vec{0}$. Se sigue que $(\mathbb{R}^n)^\perp = \{\vec{0}\}$.

Es claro que $\langle \vec{x}, \vec{0} \rangle = 0$ para todo $\vec{x} \in \mathbb{R}^n$; por lo tanto $\mathbb{R}^n \subseteq \{\vec{0}\}^\perp \subseteq \mathbb{R}^n$, lo cual prueba que $\{\vec{0}\}^\perp = \mathbb{R}^n$. $\square$

Ejemplos 7.20. (I) El complemento ortogonal de una recta en $\mathbb{R}^2$ es de nuevo una recta; véase la Figura 7.2.

(II) El complemento ortogonal de una recta en $\mathbb{R}^3$ es el plano perpendicular a la recta dada. El complemento ortogonal de un plano en $\mathbb{R}^3$ es la recta perpendicular al plano dado; véase la Figura 7.2.

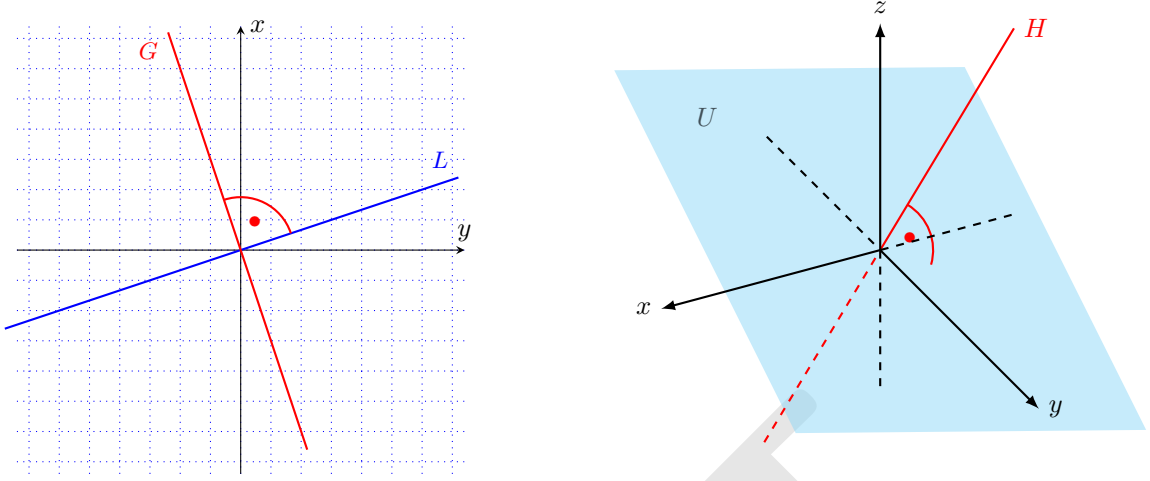


FIGURA 7.2: La figura de la izquierda muestra el complemento ortogonal de la recta L en $\mathbb{R}^2$, que es la recta G . La figura de la derecha muestra el complemento ortogonal del plano U en $\mathbb{R}^3$, que es la recta H . Observe que el complemento ortogonal de H es U .

El siguiente objetivo es mostrar que $\dim U + \dim U^\perp = n$ y establecer un método para calcular $U^\perp$. Para ello, el siguiente lema es útil. Nos dice que, para verificar que algún $\vec{x}$ es perpendicular a U , no tenemos que comprobar que $\vec{x} \perp \vec{u}$ para todo $\vec{u} \in U$, sino que basta comprobarlo para un conjunto de vectores $\vec{u}$ que genera U .

Lema 7.21. Sea $U = \text{span}\{\vec{u}_1, \dots, \vec{u}_k\} \subseteq \mathbb{R}^n$. Entonces $\vec{x} \in U^\perp$ si y solo si $\vec{x} \perp \vec{u}_j$ para todo $j = 1, \dots, k$.

Demostración. Supongamos que $\vec{x} \perp U$; entonces $\vec{x} \perp \vec{u}$ para todo $\vec{u} \in U$, en particular para los vectores generadores $\vec{u}_1, \dots, \vec{u}_k$. Ahora supongamos que $\vec{x} \perp \vec{u}_j$ para todo $j = 1, \dots, k$. Sea $\vec{u} \in U$ un vector arbitrario en U . Entonces existen $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ tales que $\vec{u} = \alpha_1 \vec{u}_1 + \dots + \alpha_k \vec{u}_k$. Así obtenemos

$$\langle \vec{x}, \vec{u} \rangle = \langle \vec{x}, \alpha_1 \vec{u}_1 + \dots + \alpha_k \vec{u}_k \rangle = \langle \vec{x}, \alpha_1 \vec{u}_1 \rangle + \dots + \alpha_k \langle \vec{x}, \vec{u}_k \rangle = 0.$$

Como $\vec{u}$ puede elegirse arbitrariamente en U , se sigue que $\vec{x} \perp U$. □

Teorema 7.22. Sea $A \in M(m \times n)$. Entonces

$$\ker(A) = (R_A)^\perp = (\text{Im } A^t)^\perp.$$

Demostración. Sean $\vec{r}_1, \dots, \vec{r}_n$ las filas de A . Como $R_A = \text{span}\{\vec{r}_1, \dots, \vec{r}_n\}$, basta mostrar que $\vec{x} \in \ker(A)$ si y solo si $\vec{x} \perp \vec{r}_j$ para todo $j = 1, \dots, n$.

Por definición, $\vec{x} \in \ker(A)$ si y solo si

$$\vec{0} = A\vec{x} = \begin{pmatrix} \vec{r}_1 \\ \vdots \\ \vec{r}_m \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} \langle \vec{r}_1, \vec{x} \rangle \\ \vdots \\ \langle \vec{r}_m, \vec{x} \rangle \end{pmatrix}.$$

Esto ocurre si y solo si $\langle \vec{r}_j, \vec{x} \rangle = 0$ for all $j = 1, \dots, m$, es decir, si y solo si $\vec{x} \perp \vec{r}_j$ para todo $j = 1, \dots, m$. $\square$

Demostración alternativa del Teorema 7.22. Observe que $R_A = C_{A^t} = \text{Im}(A^t)$. Así que tenemos que mostrar que $\ker(A) = (\text{Im}(A^t))^\perp$. Recordemos que $\langle Ax, y \rangle = \langle x, A^t y \rangle$. Por lo tanto

$$\begin{aligned} x \in \ker(A) &\iff Ax = 0 \\ &\iff Ax \perp \mathbb{R}^m \\ &\iff \langle Ax, y \rangle = 0 \text{ para todo } y \in \mathbb{R}^m \\ &\iff \langle x, A^t y \rangle = 0 \text{ para todo } y \in \mathbb{R}^m \\ &\iff x \in (\text{Im}(A^t))^\perp. \end{aligned} \quad \square$$

El teorema anterior conduce a un método para calcular el complemento ortogonal de un subespacio dado U de $\mathbb{R}^n$ como sigue.

Lema 7.23. Sea $U = \text{span}\{\vec{u}_1, \dots, \vec{u}_k\} \subseteq \mathbb{R}^n$ y sea A la matriz cuyas filas son los vectores $\vec{u}_1, \dots, \vec{u}_k$. Entonces

$$U^\perp = \ker A. \quad (7.2)$$

Demostración. Sea $\vec{x} \in \mathbb{R}^n$. Por el Lema 7.21 sabemos que $\vec{x} \in U^\perp$ si y solo si $\vec{x} \perp \vec{u}_j$ para todo $j = 1, \dots, k$. Esto ocurre si y solo si

$$\begin{aligned} \langle \vec{u}_1, \vec{x} \rangle &= 0 \\ \langle \vec{u}_2, \vec{x} \rangle &= 0 \\ \vdots &= \vdots \\ \langle \vec{u}_k, \vec{x} \rangle &= 0 \end{aligned} \quad \text{lo cual se deja escribir en forma matricial como} \quad \begin{pmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_k \end{pmatrix} \vec{x} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

que es lo mismo que $A\vec{x} = \vec{0}$ por definición de la matriz A . En conclusión, $\vec{x} \perp U$ si y solo si $A\vec{x} = \vec{0}$, es decir, si y solo si $\vec{x} \in \ker A$. $\square$

En el Ejemplo 7.28 calcularemos el complemento ortogonal de un subespacio de $\mathbb{R}^4$.

Los siguientes dos teoremas son los resultados principales de esta sección.

Teorema 7.24. Para todo subespacio $U \subseteq \mathbb{R}^n$ se tiene que

$$\dim U + \dim U^\perp = n. \quad (7.3)$$

Demostración. Sea $\vec{u}_1, \dots, \vec{u}_k$ una base de U . Observe que $k = \dim U$. Entonces tenemos en particular $U = \text{span}\{\vec{u}_1, \dots, \vec{u}_k\}$. Como en el Lema 7.21 consideramos la matriz $A \in M(k \times n)$ cuyas filas son los vectores $\vec{u}_1, \dots, \vec{u}_k$. Entonces $U^\perp = \ker A$, así que

$$\dim U^\perp = \dim(\ker A) = n - \dim(\text{Im } A).$$

Observe que $\dim(\text{Im } A)$ es la dimensión del espacio columna de A , que es igual a la dimensión del espacio fila de A por la Proposición 6.32. Como los vectores $\vec{u}_1, \dots, \vec{u}_k$ son linealmente independientes, esta dimensión es igual a k . Por lo tanto $\dim U^\perp = n - k = n - \dim U$. Reordenando obtenemos la fórmula deseada $\dim U^\perp + \dim U = n$.

(También podríamos haber dicho que la forma reducida de A no puede tener ninguna fila cero porque sus filas son linealmente independientes. Por lo tanto la forma escalonada reducida debe tener k pivotes y por ende

$$\dim U^\perp = \dim(\ker A) = n - \#(\text{pivotes de la forma reducida de } A) = n - k = n - \dim U.$$

Básicamente volvimos a probar la Proposición 6.32.) □

Teorema 7.25. *Sea $U \subseteq \mathbb{R}^n$ un subespacio de $\mathbb{R}^n$. Entonces se cumple lo siguiente.*

(I) $U \oplus U^\perp = \mathbb{R}^n$.

(II) $(U^\perp)^\perp = U$.

Demostración. (I) Recordemos que $U \cap U^\perp = \{\vec{0}\}$ por la Observación 7.19, por lo tanto, la suma es una suma directa. Ahora mostremos que $U + U^\perp = \mathbb{R}^n$. Como $U + U^\perp \subseteq \mathbb{R}^n$, solo tenemos que mostrar que $\dim(U + U^\perp) = n$ porque el único subespacio de dimensión n de $\mathbb{R}^n$ es $\mathbb{R}^n$ mismo; véase el Teorema 5.54. De la Proposición 5.62 y del Teorema 7.24 obtenemos

$$\dim(U + U^\perp) = \dim(U) + \dim(U^\perp) - \dim(U \cap U^\perp) = \dim(U) + \dim(U^\perp) = n$$

donde usamos que $\dim(U \cap U^\perp) = \dim\{\vec{0}\} = 0$.

(II) Primero mostremos que $U \subseteq (U^\perp)^\perp$. Para ello, fijemos $\vec{u} \in U$. Entonces, para todo $\vec{y} \in U^\perp$, tenemos que $\langle \vec{x}, \vec{y} \rangle = 0$, por lo tanto $\vec{x} \perp U^\perp$, that is, $\vec{x} \in (U^\perp)^\perp$. Observe que $\dim(U^\perp)^\perp = n - \dim U^\perp = n - (n - \dim U) = \dim U$. Como ya sabemos que $U \subseteq (U^\perp)^\perp$, se sigue que deben ser iguales por el Teorema 5.54. □

La siguiente proposición muestra que todo subespacio de $\mathbb{R}^n$ tiene una base ortonormal. Más adelante daremos otra demostración de este hecho cuando introduzcamos el proceso de Gram-Schmidt en la Sección 7.5.

Proposición 7.26. *Todo subespacio $U \subseteq \mathbb{R}^n$ con $\dim U > 0$ tiene una base ortonormal.*

Demostración. Sea U un subespacio de $\mathbb{R}^n$ con $\dim U = k > 0$. Entonces $\dim U^\perp = n - k$ y podemos elegir una base $\vec{w}_{k+1}, \dots, \vec{w}_n$ de $U^\perp$. Sea $A_0 \in M((n - k) \times n)$ la matriz cuyas filas son los vectores $\vec{w}_{k+1}, \dots, \vec{w}_n$. Como $U = (U^\perp)^\perp$, sabemos que $U = \ker A_0$. Escojamos cualquier $\vec{u}_1 \in \ker A_0$ con $\vec{u}_1 \neq \vec{0}$. Entonces $\vec{u}_1 \in U$. Ahora formamos la nueva matriz $A_1 \in M((n - k + 1) \times n)$ agregando $\vec{u}_1$ como nueva fila a la matriz A_0 . Observe que las filas de A_1 son linealmente independientes,

de modo que $\dim \ker(A_1) = n - (n - k + 1) = k - 1$. Si $k - 1 > 0$, escogemos cualquier vector $\vec{u}_2 \in \ker A_1$ con $\vec{u}_2 \neq \vec{0}$. Este vector es ortogonal a todas las filas de A_1 ; en particular, pertenece a U (ya que es ortogonal a $\vec{w}_{k+1}, \dots, \vec{w}_n$) y es perpendicular a $\vec{u}_1 \in U$. Ahora formamos la matriz $A_2 \in M((n - k + 2) \times n)$ agregando el vector $\vec{u}_2$ como fila a A_1 . Nuevamente, las filas de A_2 son linealmente independientes y por lo tanto $\dim(\ker A_2) = n - (n - k + 2) = k - 2$. Si $k - 2 > 0$, escogemos cualquier vector $\vec{u}_3 \in \ker A_2$ con $\vec{u}_3 \neq \vec{0}$. Este vector es ortogonal a todas las filas de A_2 ; en particular, pertenece a U (ya que es ortogonal a $\vec{w}_{k+1}, \dots, \vec{w}_n$) y es perpendicular a $\vec{u}_1, \vec{u}_2 \in U$. Continuamos este proceso hasta obtener vectores $\vec{u}_1, \dots, \vec{u}_k \in U$ que son ortogonales dos a dos y la matriz $A_k \in M(n \times n)$ consiste en filas linealmente independientes, de modo que su kernel es trivial. Por construcción, $\vec{u}_1, \dots, \vec{u}_k$ es un sistema ortogonal de k vectores en U , ninguno de los cuales es igual a $\vec{0}$. Por lo tanto, son linealmente independientes y por ello forman una base ortogonal de U , ya que $\dim U = k$. Para obtener una base ortonormal solo tenemos que normalizar cada uno de los vectores. $\square$

Corolario 7.27. *Todo sistema ortonormal en $\mathbb{R}^n$ se deja completar a una base ortonormal.*

Demostración. Sea $\vec{w}_1, \dots, \vec{w}_k$ un sistema ortonormal en $\mathbb{R}^n$ y sea $W = \text{span}\{\vec{w}_1, \dots, \vec{w}_k\}$. Por la Proposición 7.26 podemos encontrar una base ortonormal $\vec{u}_1, \dots, \vec{u}_{n-k}$ de $W^\perp$ (tome $U = W^\perp$ en la proposición). Entonces $\vec{w}_1, \dots, \vec{w}_k, \vec{u}_1, \dots, \vec{u}_{n-k}$ es una base ortonormal de $U \oplus U^\perp = \mathbb{R}^n$. $\square$

Concluimos esta sección con algunos ejemplos.

Ejemplo 7.28. Encuentre una base para el complemento ortogonal de

$$U = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Solución. Recordemos que $\vec{x} \in U^\perp$ si y solo si es perpendicular a los vectores que generan U . Por lo tanto $\vec{x} \in U^\perp$ if and only if it belongs to the kernel of the matrix whose rows are the generators of U . Así calculamos

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -2 & -2 & -4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \end{pmatrix}.$$

Por lo tanto, una base de $U^\perp$ está dada por

$$\vec{w}_1 = \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{w}_2 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}. \quad \diamond$$

Ejemplo 7.29. Encuentre una base ortonormal para el complemento ortogonal de

$$U = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Solución. Usaremos el método de la Proposición 7.26. Otra solución de este ejercicio se dará en el Ejemplo 7.44. A partir de la solución del Ejemplo 7.28 podemos tomar el primer vector de la base $\vec{w}_1$. Lo añadimos a la matriz de la solución del Ejemplo 7.28 y reducimos la nueva matriz (observe que los primeros pasos son idénticos a la reducción de la matriz original). Obtenemos

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & -2 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 2 & 5 \end{pmatrix}$$

cuyo kernel está generado por

$$\begin{pmatrix} 5 \\ 1 \\ -5 \\ 2 \end{pmatrix}.$$

Por lo tanto, una base ortonormal de $U^\perp$ está dada por

$$\vec{y}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{y}_2 = \frac{1}{\sqrt{55}} \begin{pmatrix} 5 \\ 1 \\ -5 \\ 2 \end{pmatrix}. \quad \diamond$$

Ahora debería haber entendido

- el concepto de complemento ortogonal,
- en particular la interpretación geométrica del complemento ortogonal de un subespacio (al menos en $\mathbb{R}^2$ y $\mathbb{R}^3$),
- etc.

Ahora debería ser capaz de

- encontrar el complemento ortogonal de un subespacio dado de $\mathbb{R}^n$,
- encontrar una base ortogonal de un subespacio dado de $\mathbb{R}^n$,
- etc.

Ejercicios.

1. Encuentre el complemento ortogonal de los siguientes conjuntos:

(a) $\text{span} \left\{ \begin{pmatrix} -1 \\ 5 \end{pmatrix} \right\}$

(b) La intersección de los planos $x + 2y + 5z = 0$, $2x - 3y - 4z = 0$.

(c) $\text{span} \left\{ \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \right\}$

(d) La imagen de la transformación lineal $T : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ dada por:

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + y - 3z \\ 3x + 2y - 5z \\ x - y \\ -x - 3y + 4z \end{pmatrix}$$

2. En $\mathbb{R}^4$, encuentre una base ortonormal del hiperplano $x - y + w = 0$.
3. Se dan un vector $\vec{v}$ y un subespacio W . En cada caso escriba $\vec{v}$ como la suma de un elemento en W con un elemento de $W^\perp$. Adicional encuentre una base ortonormal para $W^\perp$.

(a) $v = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ y $W = \text{span} \left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$.

(b) $\vec{v} = \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}$ y W es el plano $2x + y + 2z = 0$.

(c) $\vec{v} = \begin{pmatrix} 10 \\ -1 \\ 6 \end{pmatrix}$ y W es la recta $x = 0, \frac{y}{2} = 3z$.

(d) $\vec{v} = \begin{pmatrix} 1 \\ -1 \\ 3 \\ 2 \end{pmatrix}$ y $W = \left\{ \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \in \mathbb{R}^4 : y = 2x + w, z = x - 2w \right\}$.

4. Sea U un subespacio de $\mathbb{R}^n$. Muestre que $(U^\perp)^\perp = U$.
5. Sea $A \in M(n \times m)$ y W el espacio generado por las columnas de A . Muestre que todo $\vec{v} \in \mathbb{R}^n$ puede ser escrito como la suma de dos elementos $\vec{a}, \vec{b}$ tales que $\vec{a} \in \ker A^t$ y $\vec{b} \in W$.

7.4. Proyecciones ortogonales

Recordemos que en la Sección 2.3 discutimos la proyección ortogonal de un vector sobre otro en $\mathbb{R}^2$. Esto claramente puede extenderse a dimensiones superiores. Sean $\vec{v}, \vec{w} \in \mathbb{R}^n$ con $\vec{w} \neq \vec{0}$. Entonces

$$\text{proy}_{\vec{w}} \vec{v} := \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w} \quad (7.4)$$

es el único vector en $\mathbb{R}^n$ que es paralelo a $\vec{w}$ y satisface que $\vec{v} - \text{proy}_{\vec{w}} \vec{v}$ es ortogonal a $\vec{w}$. Ya sabemos que la proyección es independiente de la longitud de $\vec{w}$. Así, $\text{proy}_{\vec{w}} \vec{v}$ debe considerarse como la proyección de $\vec{v}$ sobre el subespacio unidimensional generado por $\vec{w}$.

En esta sección queremos generalizar esto a proyecciones ortogonales sobre subespacios de dimensión mayor; por ejemplo, puede pensarse en la proyección en $\mathbb{R}^3$ sobre un plano dado. Entonces, dado un subespacio U de $\mathbb{R}^n$, queremos definir la *proyección ortogonal* como la función de $\mathbb{R}^n$ en $\mathbb{R}^n$ que asigna a cada vector $\vec{v}$ su proyección ortogonal sobre U . Comenzamos con el análogo del Teorema 2.24.

Teorema 7.30 (Proyección ortogonal). Sea $U \subseteq \mathbb{R}^n$ un subespacio y sea $\vec{v} \in \mathbb{R}^n$. Entonces existen vectores determinados de manera única $\vec{v}_{\parallel}$ y $\vec{v}_{\perp}$ tales que

$$\vec{v}_{\parallel} \in U, \quad \vec{v}_{\perp} \perp U \quad y \quad \vec{v} = \vec{v}_{\parallel} + \vec{v}_{\perp}. \quad (7.5)$$

El vector $\vec{v}_{\parallel}$ se llama la proyección ortogonal de $\vec{v}$ sobre U ; se denota por $\text{proy}_U \vec{v}$.

Daremos una fórmula para la proyección ortogonal en la Observación 7.32 usando bases ortonormales y otra fórmula en Teorema 7.41.

Demostración. Primero mostramos la existencia de los vectores $\vec{v}_{\parallel}$ y $\vec{v}_{\perp}$. Si $U = \mathbb{R}^n$, tomamos $\vec{v}_{\parallel} = \vec{v}$ y $\vec{v}_{\perp} = \vec{0}$. Si $U = \{\vec{0}\}$, tomamos $\vec{v}_{\parallel} = \vec{0}$ y $\vec{v}_{\perp} = \vec{v}$. En caso contrario, sea $0 < \dim U = k < n$. Elijamos bases ortonormales $\vec{u}_1, \dots, \vec{u}_k$ de U y $\vec{w}_{k+1}, \dots, \vec{w}_n$ de $U^{\perp}$. Esto es posible por el Teorema 7.24 y la Proposición 7.26. Entonces $\vec{u}_1, \dots, \vec{u}_k, \vec{w}_{k+1}, \dots, \vec{w}_n$ es una base ortonormal de $\mathbb{R}^n$ y, para todo $\vec{v} \in \mathbb{R}^n$, encontramos con ayuda del Teorema 7.7 that

$$\vec{v} = \underbrace{\langle \vec{u}_1, \vec{v} \rangle \vec{u}_1 + \dots + \langle \vec{u}_k, \vec{v} \rangle \vec{u}_k}_{\in U} + \underbrace{\langle \vec{w}_{k+1}, \vec{v} \rangle \vec{w}_{k+1} + \dots + \langle \vec{w}_n, \vec{v} \rangle \vec{w}_n}_{\in U^{\perp}}.$$

Si tomamos $\vec{v}_{\parallel} = \langle \vec{u}_1, \vec{v} \rangle \vec{u}_1 + \dots + \langle \vec{u}_k, \vec{v} \rangle \vec{u}_k$ y $\vec{v}_{\perp} = \langle \vec{w}_{k+1}, \vec{v} \rangle \vec{w}_{k+1} + \dots + \langle \vec{w}_n, \vec{v} \rangle \vec{w}_n$, entonces tienen las propiedades deseadas.

A continuación mostramos la unicidad de la descomposición de $\vec{v}$. Supongamos que existen vectores $\vec{v}_{\parallel}$ y $\vec{z}_{\parallel} \in U$ y $\vec{v}_{\perp}$ y $\vec{z}_{\perp} \in U^{\perp}$ tales que $\vec{v} = \vec{v}_{\parallel} + \vec{v}_{\perp}$ y $\vec{v} = \vec{z}_{\parallel} + \vec{z}_{\perp}$. Entonces $\vec{v}_{\parallel} + \vec{v}_{\perp} = \vec{z}_{\parallel} + \vec{z}_{\perp}$ y, reordenando, encontramos que

$$\underbrace{\vec{v}_{\parallel} - \vec{z}_{\parallel}}_{\in U} = \underbrace{\vec{z}_{\perp} - \vec{v}_{\perp}}_{\in U^{\perp}}.$$

Como $U \cap U^{\perp} = \{\vec{0}\}$, se sigue que $\vec{v}_{\parallel} - \vec{z}_{\parallel} = \vec{0}$ y $\vec{z}_{\perp} - \vec{v}_{\perp} = \vec{0}$, y por lo tanto $\vec{z}_{\parallel} = \vec{v}_{\parallel}$ y $\vec{z}_{\perp} = \vec{v}_{\perp}$. $\square$

Definición 7.31. Sea U un subespacio de $\mathbb{R}^n$. Entonces definimos la *proyección ortogonal sobre U* como la función que envía $\vec{v} \in \mathbb{R}^n$ a su proyección ortogonal sobre U . Usualmente se denota por P_U , así que

$$P_U : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad P_U \vec{v} = \text{proy}_U \vec{v}.$$

Observación 7.32 (Fórmula para la proyección ortogonal). La demostración del Teorema 7.30 indica cómo podemos calcular la proyección ortogonal sobre un subespacio dado $U \subseteq \mathbb{R}^n$. Si $\vec{u}_1, \dots, \vec{u}_k$ es una base ortonormal de U , entonces

$$P_U \vec{v} = \langle \vec{u}_1, \vec{v} \rangle \vec{u}_1 + \dots + \langle \vec{u}_k, \vec{v} \rangle \vec{u}_k. \quad (7.6)$$

Esto muestra que P_U es una transformación lineal, ya que $P_U(\vec{x} + c\vec{y}) = P_U \vec{x} + cP_U \vec{y}$ se sigue fácilmente de (7.6).

Ejercicio. Si $\vec{u}_1, \dots, \vec{u}_k$ es una base ortogonal de U (pero no necesariamente ortonormal), muestre que

$$P_U \vec{v} = \frac{\langle \vec{u}_1, \vec{v} \rangle}{\|\vec{u}_1\|^2} \vec{u}_1 + \dots + \frac{\langle \vec{u}_k, \vec{v} \rangle}{\|\vec{u}_k\|^2} \vec{u}_k. \quad (7.7)$$

Observación 7.33 (Fórmula para la proyección ortogonal cuando $\dim U = 1$). Si $\dim U = 1$, obtenemos nuevamente la fórmula (7.4) que ya conocemos de la Sección 2.3. Para ver esto, elija $\vec{w} \in U$ con $\vec{w} \neq \vec{0}$. Entonces $\vec{w}' = \|\vec{w}\|^{-1}\vec{w}$ es una base ortonormal de U y, de acuerdo con (7.6) tenemos que

$$\text{proy}_{\vec{w}} \vec{v} = \text{proy}_U \vec{v} = \langle \vec{w}', \vec{v} \rangle \vec{w}' = \langle \|\vec{w}\|^{-1}\vec{w}, \vec{v} \rangle (\|\vec{w}\|^{-1}\vec{w}) = \|\vec{w}\|^{-2} \langle \vec{w}, \vec{v} \rangle \vec{w} = \frac{\langle \vec{w}, \vec{v} \rangle}{\|\vec{w}\|^2} \vec{w}.$$

Observación 7.34 (Teorema de Pitágoras). Sea U un subespacio de $\mathbb{R}^n$, sea $\vec{v} \in \mathbb{R}^n$ y sean $\vec{v}_{\parallel}$ y $\vec{v}_{\perp}$ como en el Teorema 7.30. Entonces

$$\|\vec{v}\|^2 = \|\vec{v}_{\parallel}\|^2 + \|\vec{v}_{\perp}\|^2.$$

Demostración. Usando que $\vec{v}_{\parallel} \perp \vec{v}_{\perp}$, encontramos

$$\begin{aligned} \|\vec{v}\|^2 &= \langle \vec{v}, \vec{v} \rangle = \langle \vec{v}_{\parallel} + \vec{v}_{\perp}, \vec{v}_{\parallel} + \vec{v}_{\perp} \rangle = \langle \vec{v}_{\parallel}, \vec{v}_{\parallel} \rangle + \langle \vec{v}_{\parallel}, \vec{v}_{\perp} \rangle + \langle \vec{v}_{\perp}, \vec{v}_{\parallel} \rangle + \langle \vec{v}_{\perp}, \vec{v}_{\perp} \rangle \\ &= \langle \vec{v}_{\parallel}, \vec{v}_{\parallel} \rangle + \langle \vec{v}_{\perp}, \vec{v}_{\perp} \rangle = \|\vec{v}_{\parallel}\|^2 + \|\vec{v}_{\perp}\|^2. \end{aligned}$$

□

Ejercicio 7.35. Sea U un subespacio de $\mathbb{R}^n$ con base $\vec{u}_1, \dots, \vec{u}_k$ y sea $\vec{w}_{k+1}, \dots, \vec{w}_n$ una base de $U^{\perp}$. Encuentre la representación matricial de P_U con respecto a la base $\vec{u}_1, \dots, \vec{u}_k, \vec{w}_{k+1}, \dots, \vec{w}_n$.

Ejercicio 7.36. Sea U un subespacio de $\mathbb{R}^n$. Muestre que $P_{U^{\perp}} = \text{id} - P_U$. (Puede mostrar esto directamente o usando la representación matricial de P_U del Ejercicio 7.35.)

Ejercicio 7.37. Sea U un subespacio de $\mathbb{R}^n$. Muestre que $(P_U)^2 = P_U$. (Puede mostrar esto directamente o usando la representación matricial de P_U del Ejercicio 7.35.)

Ejercicio 7.38. Sea U un subespacio de $\mathbb{R}^n$.

- (I) Encuentre $\ker P_U$ y $\text{Im } P_U$.
- (II) Encuentre $P_{U^{\perp}} P_U$ y $P_U P_{U^{\perp}}$.

En el Teorema 7.30 usamos el concepto de ortogonalidad para definir la proyección ortogonal de $\vec{v}$ sobre un subespacio dado. Obtuvimos una descomposición de $\vec{v}$ en una parte paralela al subespacio dado y una parte ortogonal a él. El siguiente teorema muestra que la proyección ortogonal de $\vec{v}$ sobre U nos da el punto en U que está más cerca de $\vec{v}$.

Teorema 7.39. Sea U un subespacio de $\mathbb{R}^n$ y sea $\vec{v} \in \mathbb{R}^n$. Entonces $P_U \vec{v}$ es el punto en U que está más cerca de $\vec{v}$, es decir,

$$\|\vec{v} - P_U \vec{v}\| \leq \|\vec{v} - \vec{u}\| \quad \text{para todo } \vec{u} \in U.$$

Demostración. Sean $\vec{v} \in \mathbb{R}^n$ y $\vec{u} \in U \subseteq \mathbb{R}^n$. Observe que $\vec{v} - P_U \vec{v} \in U^{\perp}$ y que $P_U \vec{v} - \vec{u} \in U$, ya que ambos vectores pertenecen a U . Por lo tanto, el teorema de Pitágoras muestra que

$$\|\vec{v} - \vec{u}\|^2 = \|\vec{v} - P_U \vec{v} + P_U \vec{v} - \vec{u}\|^2 = \|\vec{v} - P_U \vec{v}\|^2 + \|P_U \vec{v} - \vec{u}\|^2 \geq \|\vec{v} - P_U \vec{v}\|^2.$$

Tomando raíz cuadrada en ambos lados se obtiene la desigualdad deseada.

□

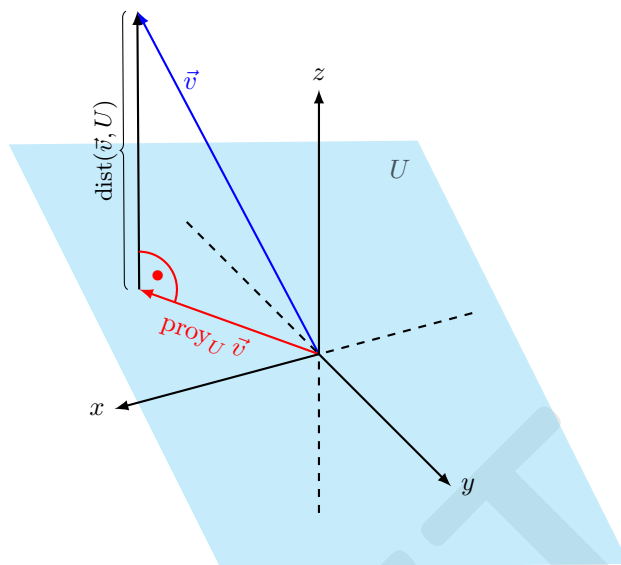


FIGURA 7.3: La figura muestra la proyección ortogonal del vector $\vec{v}$ sobre el subespacio U (que es un vector) y la distancia de $\vec{v}$ a U (que es un número. Es la longitud del vector $(\vec{v} - \text{proy}_U \vec{v})$).

Definición 7.40. Sea U un subespacio de $\mathbb{R}^n$ y sea $\vec{v} \in \mathbb{R}^n$. Entonces definimos la *distancia de $\vec{v}$ a U* como

$$\text{dist}(\vec{v}, U) := \|\vec{v} - P_U \vec{v}\|.$$

Esta es la distancia más corta de $\vec{v}$ a cualquier punto de U .

En la Observación 7.32 ya encontramos una fórmula para la proyección ortogonal P_U de un vector $\vec{v}$ sobre un subespacio dado U . Sin embargo, esta fórmula requiere tener una base ortonormal de U . Queremos dar otra fórmula para P_U que no requiera conocer una base ortonormal.

Teorema 7.41. Sea U un subespacio de $\mathbb{R}^n$ con base $\vec{u}_1, \dots, \vec{u}_k$ y sea $B \in M(n \times k)$ la matriz cuyas columnas son los vectores de esta base. Entonces se cumple lo siguiente.

- (I) B es inyectiva.
- (II) $B^t B : \mathbb{R}^k \rightarrow \mathbb{R}^k$ es una biyección.
- (III) La proyección ortogonal sobre U está dada por la fórmula

$$P_U = B(B^t B)^{-1} B^t.$$

Demostración. (i) Por construcción, las columnas de B son linealmente independientes. Por lo tanto, la solución única de $B\vec{x} = \vec{0}$ es $\vec{x} = \vec{0}$ lo que demuestre que B es inyectiva.

- (ii) Observe que $B^t B \in M(k \times k)$ y suponga que $B^t B\vec{x} = \vec{0}$ para algún $\vec{x} \in \mathbb{R}^k$. Entonces se sigue para todo $\vec{y} \in \mathbb{R}^k$ que $B\vec{y} = \vec{0}$, porque

$$0 = \langle \vec{y}, B^t B\vec{y} \rangle = \langle (B^t)^t \vec{y}, B\vec{y} \rangle = \langle B\vec{y}, B\vec{y} \rangle = \|B\vec{y}\|^2.$$

Como B es inyectiva, esto implica $\vec{y} = \vec{0}$, de modo que $B^t B$ es inyectiva. Como es una matriz cuadrada, se sigue que incluso es biyectiva.

- (III) Observe que, por construcción, $\text{Im } B = U$. Ahora sea $\vec{x} \in \mathbb{R}^n$. Observe que $P_U \vec{x} \in \text{Im } B$. Por lo tanto existe exactamente un $\vec{z} \in \mathbb{R}^k$ tal que $P_U \vec{x} = B\vec{z}$. Además, $\vec{x} - P_U \vec{x} \perp U = \text{Im } B$, por lo tanto, para todo $\vec{y} \in \mathbb{R}^k$ tenemos que

$$0 = \langle \vec{x} - P_U \vec{x}, B\vec{y} \rangle = \langle \vec{x} - B\vec{z}, B\vec{y} \rangle = \langle B^t \vec{x} - B^t B\vec{z}, \vec{y} \rangle.$$

Como esto es cierto para todo $\vec{y} \in \mathbb{R}^k$, se sigue que $B^t \vec{x} - B^t B\vec{z} = \vec{0}$. Ahora recordamos que $B^t B$ es invertible, así que podemos despejar $\vec{z}$ y obtener $\vec{z} = (B^t B)^{-1} B^t \vec{x}$. Esto finalmente da

$$P_U \vec{x} = B\vec{z} = B(B^t B)^{-1} B^t \vec{x}.$$

Como esto se cumple para todo $\vec{x} \in \mathbb{R}^n$, la fórmula (III) queda probada. $\square$

Ahora debería haber entendido

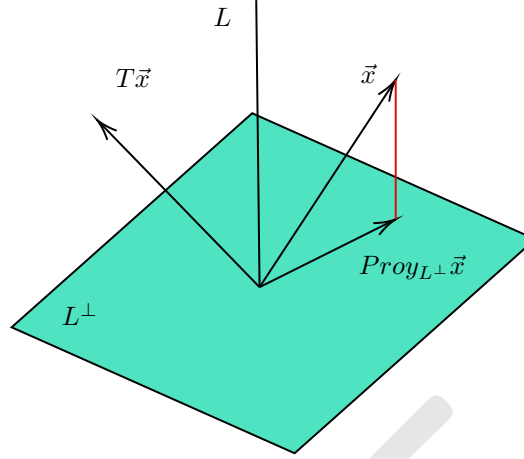
- el concepto de proyección ortogonal sobre un subespacio de $\mathbb{R}^n$,
- la interpretación geométrica de las proyecciones ortogonales y cómo se relaciona con la distancia de un punto a un subespacio,
- etc.

Ahora debería ser capaz de

- calcular la proyección ortogonal de un punto sobre un subespacio,
- calcular la distancia de un punto a un subespacio,
- etc.

Ejercicios.

1. Sea $E : x + 2y - z = 1$. Encuentre la distancia del punto $P(6, 1, 2)$ al plano E .
2. Sea $L : \frac{x-1}{2} = y + 2 = \frac{z+1}{3}$. Encuentre la distancia del punto $P(6, -1, 0)$ a la recta L .
3. Sea $L : \frac{x}{2} = \frac{y}{3} = z$.
 - (a) Encuentre $L^\perp$.
 - (b) Sea $T(\vec{x}) =$ media rotación de $\vec{x}$ con respecto a la recta L . Encuentre una fórmula explícita para T . (Hint: ¿Cuál es la proyección de $T\vec{x}$ en L y en $L^\perp$? Vea la gráfica).



4. Sea W un subespacio de $\mathbb{R}^n$ y P_W la proyección ortogonal sobre W .
 - (a) Sean $\{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_k\}$ una base ortonormal de W . Recuerde que esta base se puede completar a una base ortonormal $\mathcal{B}$ de $\mathbb{R}^n$. ¿Cómo es la representación matricial de P_W con respecto a la base $\mathcal{B}$?
 - (b) Pruebe que $[P_W]_{can}$ es una matriz simétrica.
 - (c) Muestre que $\langle P_W \vec{x}, \vec{y} \rangle = \langle \vec{x}, P_W \vec{y} \rangle$ para todos $\vec{x}, \vec{y} \in \mathbb{R}^n$.
 - (d) Muestre que $P_W P_W \vec{x} = P_W \vec{x}$ para todo $\vec{x} \in \mathbb{R}^n$.
5. Sean V, W subespacios de $\mathbb{R}^n$ tales que $W \subseteq V$. Muestre que $V^\perp \subseteq W^\perp$.
6. Sean V, W subespacios de $\mathbb{R}^n$ tales que $W \subseteq V$. Muestre que $P_V P_W = P_W P_V = P_W$. (Hint: $P_V P_W = P_W$ es directo. Para probar que $P_W P_V \vec{x} = P_W \vec{x}$ para todo $\vec{x} \in \mathbb{R}^n$, escriba $\vec{x} = \vec{v} + \vec{v}^\perp$ donde $\vec{v} \in V$ y $\vec{v}^\perp \in V^\perp$.)
7. Sea $W \subseteq \mathbb{R}^m$ un subespacio de dimensión n . Muestre que existe una isometría $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tal que $\text{Im } T = W$. (Hint: Empiece escogiendo una base ortonormal para W , ver ejercicio 5. de la sección 7.2).

7.5. El proceso de Gram-Schmidt

En esta sección describiremos el llamado proceso de ortonormalización de Gram-Schmidt. En términos generales, convierte una base dada de un subespacio de $\mathbb{R}^n$ en una base ortonormal, proporcionando así otra demostración de que todo subespacio de $\mathbb{R}^n$ tiene una base ortonormal (Corolario 7.27).

Teorema 7.42. Sea U un subespacio de $\mathbb{R}^n$ con base $\vec{u}_1, \dots, \vec{u}_k$. Entonces existe una base ortonormal $\vec{x}_1, \dots, \vec{x}_k$ de U tal que

$$\text{span}\{\vec{u}_1, \dots, \vec{u}_j\} = \text{span}\{\vec{x}_1, \dots, \vec{x}_j\} \quad \text{para todo } j = 1, \dots, k.$$

Demostración. La demostración es constructiva, es decir, no solo probamos la existencia de tal base, sino que incluso mostramos cómo obtenerla. La idea es construir la nueva base $\vec{x}_1, \dots, \vec{x}_k$ paso a paso. Para simplificar un poco la notación, ponemos $U_j = \text{span}\{\vec{u}_1, \dots, \vec{u}_j\}$ for $j = 1, \dots, k$. Observe que $\dim U_j = j$ y que $U_k = U$.

- Tomemos $\vec{x}_1 = \|\vec{u}_1\|^{-1}\vec{u}_1$. Entonces claramente $\|\vec{x}_1\| = 1$ y $\text{span}\{\vec{u}_1\} = \text{span}\{\vec{x}_1\} = U_1$.
- El vector $\vec{x}_2$ debe ser un vector normalizado en U_2 que sea ortogonal a $\vec{x}_1$, es decir, debe ser ortogonal a U_1 . Así que simplemente tomamos $\vec{u}_2$ y le restamos su proyección sobre U_1 :

$$\vec{w}_2 = \vec{u}_2 - \text{proy}_{U_1} \vec{u}_2 = \vec{u}_2 - \text{proy}_{\vec{x}_1} \vec{u}_2 = \vec{u}_2 - \langle \vec{x}_1, \vec{u}_2 \rangle \vec{x}_1.$$

Claramente $\vec{w}_2 \in U_2$ porque es una combinación lineal de vectores en U_2 . Además, $\vec{w}_2 \perp U_1$ porque

$$\langle \vec{w}_2, \vec{x}_1 \rangle = \langle \vec{u}_2 - \langle \vec{x}_1, \vec{u}_2 \rangle \vec{x}_1, \vec{x}_1 \rangle = \langle \vec{u}_2, \vec{x}_1 \rangle - \langle \vec{x}_1, \vec{u}_2 \rangle \langle \vec{x}_1, \vec{x}_1 \rangle = \langle \vec{u}_2, \vec{x}_1 \rangle - \langle \vec{x}_1, \vec{u}_2 \rangle = 0.$$

Por lo tanto, el vector $\vec{x}_2$ que buscamos es

$$\vec{x}_2 = \|\vec{w}_2\|^{-1}\vec{w}_2.$$

Como $\vec{x}_2 \in U_2$, se sigue que $\text{span}\{\vec{x}_1, \vec{x}_2\} \subseteq U_2$. Ambos espacios tienen dimensión 2, así que deben ser iguales.

- El vector $\vec{x}_3$ debe ser un vector normalizado en U_3 que sea ortogonal a $U_2 = \text{span}\{\vec{x}_1, \vec{x}_2\}$. Así que simplemente tomamos $\vec{u}_3$ y le restamos su proyección sobre U_2 :

$$\vec{w}_3 = \vec{u}_3 - \text{proy}_{U_2} \vec{u}_3 = \vec{u}_3 - (\text{proy}_{\vec{x}_1} \vec{u}_3 + \text{proy}_{\vec{x}_2} \vec{u}_3) = \vec{u}_3 - (\langle \vec{x}_1, \vec{u}_3 \rangle \vec{x}_1 + \langle \vec{x}_2, \vec{u}_3 \rangle \vec{x}_2).$$

Claramente $\vec{w}_3 \in U_3$ porque es una combinación lineal de vectores en U_3 . Además, $\vec{w}_3 \perp U_2$ porque para $j = 1, 2$ obtenemos

$$\begin{aligned} \langle \vec{w}_3, \vec{x}_j \rangle &= \langle \vec{u}_3 - (\langle \vec{x}_1, \vec{u}_3 \rangle \vec{x}_1 + \langle \vec{x}_2, \vec{u}_3 \rangle \vec{x}_2), \vec{x}_j \rangle \\ &= \langle \vec{u}_3, \vec{x}_j \rangle - \langle \vec{x}_1, \vec{u}_3 \rangle \langle \vec{x}_1, \vec{x}_j \rangle - \langle \vec{x}_2, \vec{u}_3 \rangle \langle \vec{x}_2, \vec{x}_j \rangle \\ &= \langle \vec{u}_3, \vec{x}_j \rangle - \langle \vec{x}_j, \vec{u}_3 \rangle \langle \vec{x}_j, \vec{x}_j \rangle = \langle \vec{u}_3, \vec{x}_j \rangle - \langle \vec{x}_j, \vec{u}_3 \rangle = 0. \end{aligned}$$

Por lo tanto, el vector $\vec{x}_3$ que buscamos es

$$\vec{x}_3 = \|\vec{w}_3\|^{-1}\vec{w}_3.$$

Como $\vec{x}_3 \in U_3$, se sigue que $\text{span}\{\vec{x}_1, \vec{x}_2, \vec{x}_3\} \subseteq U_3$. Como ambos espacios tienen dimensión 3, deben ser iguales.

Repetimos esto k veces hasta haber construido la base $\vec{x}_1, \dots, \vec{x}_k$.

Observe que el procedimiento general es el siguiente:

- Supongamos que ya hemos construido $\vec{x}_1, \dots, \vec{x}_\ell$. Entonces primero construimos

$$\vec{w}_{\ell+1} = \vec{u}_{\ell+1} - P_{U_\ell} \vec{u}_{\ell+1}.$$

Este vector satisface $\vec{w}_{\ell+1} \in U_{\ell+1}$ y $\vec{w}_{\ell+1} \perp U_\ell$. Observe que $\vec{w}_{\ell+1} \neq \vec{0}$ porque, de lo contrario, tendríamos que $\vec{u}_{\ell+1} = P_{U_\ell} \vec{u}_{\ell+1} \in U_\ell$, lo cual es imposible porque $\vec{u}_{\ell+1}, \vec{u}_\ell, \dots, \vec{u}_1$ son linealmente independientes. Entonces $\vec{x}_{\ell+1} = \|\vec{w}_{\ell+1}\|^{-1} \vec{w}_{\ell+1}$ tiene todas las propiedades deseadas. $\square$

Ejemplo 7.43. Sea $U = \text{span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ donde

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} -1 \\ 4 \\ \sqrt{2} \\ 3 \\ 2 \end{pmatrix}, \quad \vec{u}_3 = \begin{pmatrix} -2 \\ 5 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Queremos encontrar una base ortonormal $\vec{x}_1, \vec{x}_2, \vec{x}_3$ de U usando el proceso de Gram-Schmidt.

Solución. (I) $\vec{x}_1 = \|\vec{u}_1\|^{-1} \vec{u}_1 = \frac{1}{2} \vec{u}_1$.

$$(II) \quad \vec{w}_2 = \vec{u}_2 - \text{proy}_{\vec{x}_1} \vec{u}_2 = \vec{u}_2 - \langle \vec{x}_1, \vec{u}_2 \rangle \vec{x}_1 = \vec{u}_2 - 4\vec{x}_1 = \vec{u}_2 - 2\vec{u}_1 = \begin{pmatrix} -3 \\ 2 \\ \sqrt{2} \\ 1 \\ 0 \end{pmatrix}$$

$$\Rightarrow \quad \vec{x}_2 = \|\vec{w}_2\|^{-1} \vec{w}_2 = \frac{1}{4} \begin{pmatrix} -3 \\ 2 \\ \sqrt{2} \\ 1 \\ 0 \end{pmatrix}.$$

$$(III) \quad \vec{w}_3 = \vec{u}_3 - \text{proy}_{\text{span}\{\vec{x}_1, \vec{x}_2\}} \vec{u}_3 = \vec{u}_3 - [\langle \vec{x}_1, \vec{u}_3 \rangle \vec{x}_1 + \langle \vec{x}_2, \vec{u}_3 \rangle \vec{x}_2] = \vec{u}_3 - [2\vec{x}_1 + 4\vec{x}_2]$$

$$= \begin{pmatrix} -2 \\ 5 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} -3 \\ 2 \\ \sqrt{2} \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ -\sqrt{2} \\ -2 \\ 0 \end{pmatrix}$$

$$\Rightarrow \quad \vec{x}_3 = \|\vec{w}_3\|^{-1} \vec{w}_3 = \frac{1}{\sqrt{10}} \begin{pmatrix} 0 \\ -2 \\ \sqrt{2} \\ 2 \\ 0 \end{pmatrix}.$$

Por lo tanto, la base ortonormal deseada de U es

$$\vec{x}_1 = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{x}_2 = \frac{1}{4} \begin{pmatrix} -3 \\ 2 \\ \sqrt{2} \\ 1 \\ 0 \end{pmatrix}, \quad \vec{x}_3 = \frac{1}{\sqrt{10}} \begin{pmatrix} 0 \\ -2 \\ \sqrt{2} \\ 2 \\ 0 \end{pmatrix}.$$

$\diamond$

Observe que obtendremos una base diferente si cambiamos el orden de la base dada $\vec{u}_1, \vec{u}_2, \vec{u}_3$.

Ejemplo 7.44. Daremos otra solución del Ejemplo 7.29. Se nos pidió encontrar una base ortonormal del complemento ortogonal de

$$U = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Del Ejemplo 7.28 ya sabemos que

$$U^\perp = \text{span}\{\vec{w}_1, \vec{w}_2\} \quad \text{donde} \quad \vec{w}_1 = \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{w}_2 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}.$$

Usamos el proceso de Gram-Schmidt para obtener una base ortonormal $\vec{x}_1, \vec{x}_2$ de U .

$$(I) \quad \vec{x}_1 = \|\vec{v}_1\|^{-1} \vec{v}_1 = \frac{1}{\sqrt{5}} \vec{v}_1.$$

$$(II) \quad \vec{y}_2 = \vec{w}_2 - \text{proy}_{\vec{x}_1} \vec{w}_2 = \vec{w}_2 - \langle \vec{x}_1, \vec{w}_2 \rangle \vec{x}_1 = \vec{w}_2 - \frac{2}{\sqrt{5}} \vec{x}_1 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} - \frac{2}{5} \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -5 \\ -1 \\ 5 \\ -2 \end{pmatrix}$$

$$\Rightarrow \quad \vec{x}_2 = \|\vec{y}_2\|^{-1} \vec{y}_2 = \frac{1}{\sqrt{55}} \begin{pmatrix} -5 \\ -1 \\ 5 \\ -2 \end{pmatrix}$$

Por lo tanto

$$\vec{x}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{x}_2 = \frac{1}{\sqrt{55}} \begin{pmatrix} -5 \\ -1 \\ 5 \\ -2 \end{pmatrix}.$$

Ahora debería haber entendido

- por qué funciona el proceso de Gram-Schmidt,
- etc.

Ahora debería ser capaz de

- aplicar el proceso de Gram-Schmidt para generar una base ortonormal de un subespacio dado,
- etc.

Ejercicios.

1. Mediante proceso de Gram-Schmidt obtenga una base ortonormal del plano $x - y + z = 0$.

2. Encuentre una base ortonormal de $\mathbb{R}^4$ que contenga una base del subespacio generado por los vectores $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix}$.
3. Sea $W = \text{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ y $\vec{v} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$. Encuentre el elemento en W más cercano a $\vec{v}$ y determine la distancia de $\vec{v}$ a W .
4. Sea $A = \begin{pmatrix} 1 & 2 & 2 & -5 \\ 3 & 2 & 1 & -2 \\ 2 & 0 & -1 & 3 \\ 7 & -2 & 1 & 4 \end{pmatrix}$. Determine una base ortonormal para el espacio columna de A .
5. Sea $E : 2x + 3y + z = 0$ y P_E la proyección ortogonal sobre E . Encuentre bases $\mathcal{B}_1$ y $\mathcal{B}_2$ tales que $[P_E]_{\mathcal{B}_2}^{\mathcal{B}_1} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$.

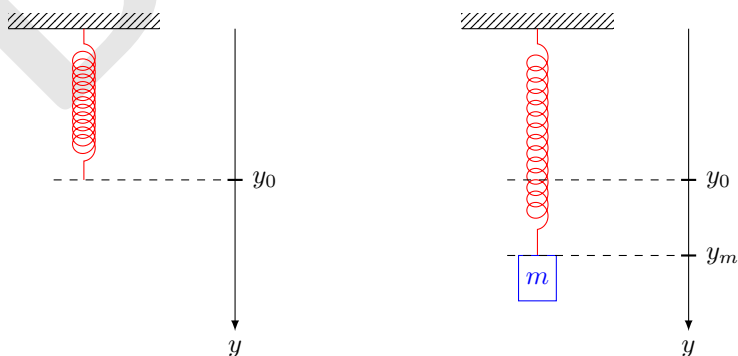
7.6. Aplicación: Mínimos cuadrados

En esta sección queremos presentar el método de mínimos cuadrados para ajustar una función lineal a ciertas mediciones. Veamos un ejemplo.

Ejemplo 7.45. Supongamos que queremos medir la constante de Hook k de un resorte. Por la ley de Hook sabemos que

$$y = y_0 + km \quad (7.8)$$

donde y_0 es la elongación del resorte sin ninguna masa colgada y y es la elongación del resorte cuando le colgamos la masa m .



Supongamos que medimos la elongación para diferentes masas. Si la ley de Hook es válida y si nuestras mediciones fueran perfectas, entonces los puntos medidos deberían estar sobre una recta con pendiente k . Sin embargo, las mediciones nunca son perfectas y los puntos más bien estarán dispersos alrededor de una recta. Supongamos que obtuvimos las siguientes mediciones.

m	2	3	4	5
y	4.5	5.1	6.1	7.9

La Figura 7.4 contiene una gráfica de estas mediciones en el plano m - y .

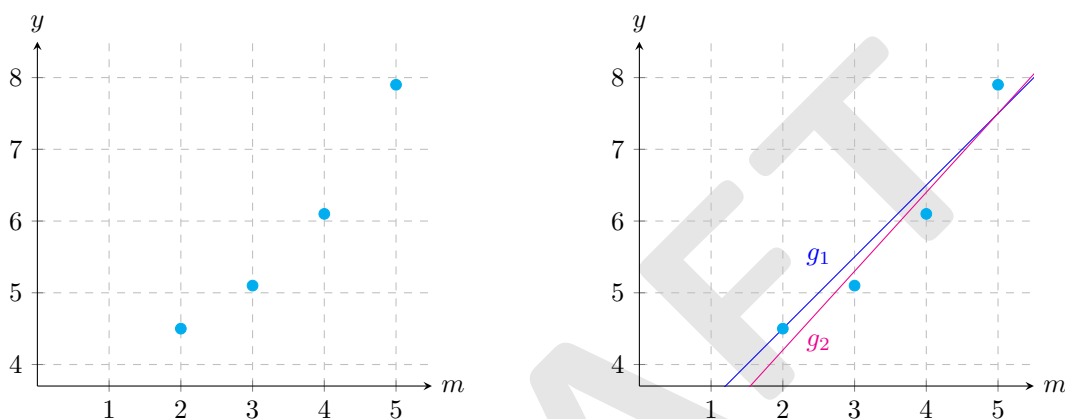


FIGURA 7.4: La gráfica de la izquierda muestra los datos medidos. En la gráfica de la derecha añadimos las dos funciones $g_1(x) = x + 2,5$, $g_2(x) = 1,1x + 2$ que parecen ser candidatas razonables para aproximaciones lineales de los datos medidos.

La gráfica nos da cierta confianza de que la ley de Hook se cumple, ya que los puntos parecen estar más o menos sobre una recta. ¿Cómo ajustamos de la mejor manera una recta a los puntos? La pendiente parece estar alrededor de 1. Podríamos hacer las siguientes conjeturas:

$$g_1(x) = x + 2,5 \quad \text{o} \quad g_2(x) = 1,1x + 2$$

¿Cuál de las dos funciones es la mejor aproximación? ¿Hay otras aproximaciones que sean aún mejores?

La respuesta a estas preguntas depende mucho de cómo medimos qué tan “buena” es una aproximación. Una forma muy común es la siguiente: Para cada punto medido, tomamos la diferencia $\Delta_j := m_j - g(m_j)$ entre el valor medido y el valor de nuestra función de prueba. Luego elevamos al cuadrado todas estas diferencias, las sumamos y después tomamos la raíz cuadrada $\left(\sum_{j=1}^n (m_j - g(m_j))^2\right)^{\frac{1}{2}}$; véase también la Figura 7.5. El número resultante será nuestra medida de qué tan buena es nuestra conjetura.

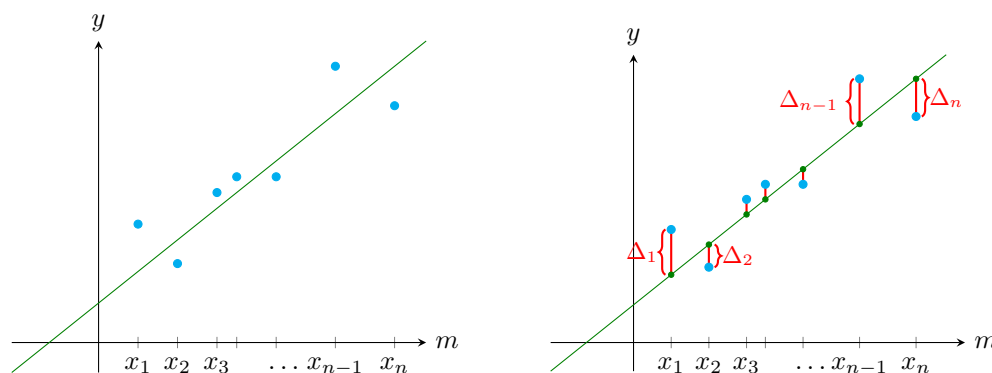


FIGURA 7.5: La gráfica de la izquierda muestra puntos para los cuales queremos encontrar una función lineal aproximante. La gráfica de la derecha muestra una función lineal de este tipo y cómo medir el error o la discrepancia entre los puntos medidos y la recta propuesta. Una medida del error es $(\sum_{j=1}^n \Delta_j^2)^{\frac{1}{2}}$.

Antes de hacer esto para nuestros datos, hacemos algunas observaciones sencillas.

- (I) Si todos los puntos medidos están sobre una recta y tomamos esta recta como nuestra candidata, entonces este método da el error total 0, como debe ser.
- (II) Tomamos los cuadrados de los errores en cada punto medido para que el error siempre se cuente como positivo. De lo contrario, podría ocurrir que los errores se cancelen entre sí. Si simplemente sumáramos los errores, entonces el error total podría ser 0 aunque la recta aproximante esté bastante lejos de todos los puntos medidos.
- (III) Hay otras formas de medir el error; por ejemplo, se podría usar $\sum_{j=1}^n |m_j - g(m_j)|$, pero resulta que el método con los cuadrados tiene muchas ventajas. (Véase algún curso de optimización para más detalles.)

Ahora calculemos los errores para nuestros puntos medidos y nuestras dos funciones propuestas.

m	2	3	4	5	m	2	3	4	5
y (medido)	4.5	5.1	6.1	7.9	y (medido)	4.5	5.1	6.1	7.9
$g_1(m)$	4.5	4.5	6.5	7.5	$g_2(m)$	4.2	5.3	6.4	7.5
$y - g_1$	0	0.6	-0.4	0.4	$y - g_2$	0.3	-0.2	-0.3	0.4

Por lo tanto, para los errores obtenemos

$$\text{Error para la función } g_1 : \Delta^{(1)} = [0^2 + 0,6^2 + (-0,4)^2 + 0,4^2]^{\frac{1}{2}} = [0,68]^{\frac{1}{2}} \approx 0,825,$$

$$\text{Error para la función } g_2 : \Delta^{(2)} = [0,3^2 + (-0,2)^2 + (-0,3)^2 + 0,4^2]^{\frac{1}{2}} = [0,38]^{\frac{1}{2}} \approx 0,616,$$

así que nuestra segunda conjetura parece estar más cerca de la mejor aproximación lineal a nuestros puntos medidos que la primera conjetura. Este ejercicio continuará en la p. 312.

Ahora surge la pregunta de cómo podemos encontrar la aproximación lineal óptima.

Mejor aproximación lineal. Supongamos que se nos dan datos medidos $(x_1, y_1), \dots, (x_n, y_n)$ y queremos encontrar una función lineal $g(x) = ax + b$ tal que el error total

$$\Delta := \left[\sum_{j=1}^n (y_j - g(x_j))^2 \right]^{\frac{1}{2}} \quad (7.9)$$

sea mínimo. En otras palabras, tenemos que encontrar los parámetros a y b tales que Δ sea lo más pequeño posible. La clave aquí es reconocer el lado derecho de (7.9) como la norma de un vector (aquí es crucial la forma particular en que elegimos medir el error). Reescribamos (7.9) como sigue:

$$\begin{aligned} \Delta &= \left[\sum_{j=1}^n (y_j - g(x_j))^2 \right]^{\frac{1}{2}} = \left[\sum_{j=1}^n (y_j - (ax_j - b))^2 \right]^{\frac{1}{2}} = \left\| \begin{pmatrix} y_1 - (ax_1 - b) \\ y_2 - (ax_2 - b) \\ \vdots \\ y_n - (ax_n - b) \end{pmatrix} \right\| \\ &= \left\| \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - \begin{bmatrix} a \\ a \\ \vdots \\ a \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\|. \end{aligned}$$

Pongamos

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{y} \quad \vec{u} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \quad (7.10)$$

Observe que estos son vectores en $\mathbb{R}^n$. Entonces

$$\Delta = \|\vec{y} - [a\vec{x} + b\vec{u}]\|$$

y la pregunta es cómo tenemos que elegir a y b para que esto sea lo más pequeño posible. En otras palabras, buscamos el punto en el espacio vectorial generado por $\vec{x}$ y $\vec{u}$ que está más cerca de $\vec{y}$. Por el Teorema 7.39, este punto está dado por la proyección ortogonal de $\vec{y}$ sobre ese plano.

Para calcular esta proyección, tome $U = \text{span}\{\vec{x}, \vec{u}\}$ y sea P la proyección ortogonal sobre U . Entonces, por nuestro razonamiento,

$$P\vec{y} = a\vec{x} + b\vec{u}. \quad (7.11)$$

Ahora veamos cómo podemos calcular fácilmente a y b a partir de (7.11).¹ En lo que sigue supondremos que $\vec{x}$ y $\vec{u}$ son linealmente independientes, de modo que U es un plano. Esta suposición parece razonable porque que sean linealmente dependientes significaría que $x_1 = \dots = x_n$ (en nuestro ejemplo con el resorte esto significaría que siempre usamos la misma masa en el experimento). Observe que si $\vec{x}, \vec{u}$ no fueran linealmente independientes, entonces la matriz A de abajo tendría solo una columna; todo lo demás funciona igual.

Recordemos que por el Teorema 7.41 la proyección ortogonal sobre U está dada por

$$P = A(A^t A)^{-1} A^t$$

¹Por supuesto, simplemente podría calcular $P\vec{y}$ y luego plantear el sistema lineal $n \times 2$ para encontrar los coeficientes a y b .

donde A es la matriz $n \times 2$ cuyas columnas consisten en los vectores $\vec{x}$ y $\vec{u}$. Por lo tanto, (7.11) se convierte en

$$A(A^t A)^{-1} A^t \vec{y} = a\vec{x} + b\vec{u} = A \begin{pmatrix} a \\ b \end{pmatrix}. \quad (7.12)$$

Como por nuestra suposición las columnas de A son linealmente independientes, es inyectiva. Por lo tanto podemos concluir de (7.12) que

$$(A^t A)^{-1} A^t \vec{y} = \begin{pmatrix} a \\ b \end{pmatrix}$$

que es la fórmula para los números a y b que estábamos buscando.

Resumamos nuestro razonamiento anterior en un teorema.

Teorema 7.46. Sean dados $(x_1, y_1), \dots, (x_n, y_n)$. La función lineal $g(x) = ax + b$ que minimiza el error total

$$\Delta := \left[\sum_{j=1}^n (y_j - g(x_j))^2 \right]^{\frac{1}{2}} \quad (7.13)$$

está dada por

$$\begin{pmatrix} a \\ b \end{pmatrix} = (A^t A)^{-1} A^t \vec{y} \quad (7.14)$$

donde $\vec{y}, \vec{x}$ y $\vec{u}$ son como en (7.10) y A es la matriz $n \times 2$ cuyas columnas son los vectores $\vec{x}$ y $\vec{u}$.

En la Observación 7.47 mostraremos cómo esta fórmula puede derivarse con métodos del cálculo.

Exercise 7.45 continued. Usemos el Teorema 7.46 para calcular la mejor aproximación lineal a los datos del Exercise 7.45. Observe que en este caso los m_j corresponden a los x_j del teorema y escribiremos $\vec{m}$ en lugar de $\vec{x}$. En este caso tenemos

$$\vec{m} = \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \quad \vec{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad A = (\vec{x} | \vec{u}) = \begin{pmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} 4,5 \\ 5,1 \\ 6,1 \\ 7,9 \end{pmatrix},$$

Por lo tanto,

$$A^t A = \begin{pmatrix} 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} 54 & 14 \\ 14 & 4 \end{pmatrix}, \quad (A A^t)^{-1} = \frac{1}{5} \begin{pmatrix} 2 & -7 \\ -7 & 27 \end{pmatrix}$$

y por lo tanto

$$\begin{aligned} \begin{pmatrix} a \\ b \end{pmatrix} &= (A^t A)^{-1} A^t \vec{y} = \frac{1}{10} \begin{pmatrix} 2 & -7 \\ -7 & 27 \end{pmatrix} \begin{pmatrix} 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 4,5 \\ 5,1 \\ 6,1 \\ 7,9 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} -3 & -1 & 1 & 3 \\ 13 & 6 & -1 & -8 \end{pmatrix} \begin{pmatrix} 4,5 \\ 5,1 \\ 6,1 \\ 7,9 \end{pmatrix} \\ &= \begin{pmatrix} 1,12 \\ 1,98 \end{pmatrix}. \end{aligned}$$

Concluimos que la mejor aproximación lineal es

$$g(m) = 1,12m + 1,98.$$

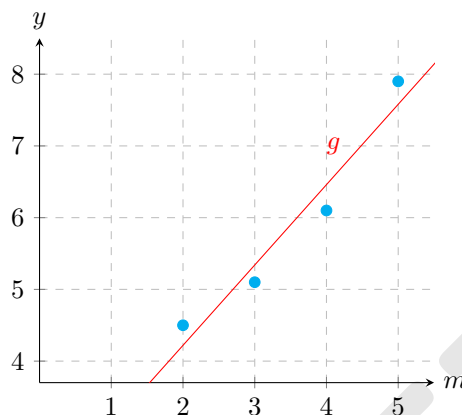


FIGURA 7.6: La gráfica muestra los datos medidos y la aproximación lineal $g(m) = 1,12m + 1,98$ calculada con el Teorema 7.46.

El método anterior puede generalizarse a otros tipos de funciones. Mostraremos cómo puede adaptarse al caso de funciones polinomiales y exponenciales.

Funciones polinomiales.. Supongamos que se nos dan datos medidos $(x_1, y_1), \dots, (x_n, y_n)$ y queremos encontrar un polinomio de grado k que ajuste mejor los puntos de datos. Sea $p(x) = a_k x^k + a_{k-1} x^{k-1} + \dots + a_1 x + a_0$ el polinomio deseado. Definimos los vectores

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \vec{\xi}_k = \begin{pmatrix} x_1^k \\ x_2^k \\ \vdots \\ x_n^k \end{pmatrix}, \quad \vec{\xi}_{k-1} = \begin{pmatrix} x_1^{k-1} \\ x_2^{k-1} \\ \vdots \\ x_n^{k-1} \end{pmatrix}, \quad \dots, \quad \vec{\xi}_1 = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \vec{\xi}_0 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Si los vectores $\vec{\xi}_k, \dots, \vec{\xi}_0$ son linealmente independientes, entonces

$$\begin{pmatrix} a_k \\ \vdots \\ a_1 \\ a_0 \end{pmatrix} = (A^t A)^{-1} A^t \vec{y}$$

donde $A = (\vec{\xi}_k \mid \dots \mid \vec{\xi}_0)$ es la matriz $n \times (k+1)$ cuyas columnas son los vectores $\vec{\xi}_k, \dots, \vec{\xi}_0$. Observe que por nuestra suposición $k < n$ (de lo contrario, los vectores $\vec{\xi}_k, \dots, \vec{\xi}_0$ no pueden ser linealmente independientes).

Observación. En general, se deberían tener muchos más puntos de datos que el grado del polinomio que se quiere ajustar; de lo contrario puede ocurrir el problema de sobreajuste. Por ejemplo, supongamos que la curva que buscamos es $f(x) = 0,1 + 0,2x$ y que se nos dan solo tres mediciones: $(0, 0,25)$, $(1, 0)$, $(3, 1)$. Entonces un ajuste lineal nos daría $g(x) = \frac{2}{7}x + \frac{1}{28} \approx 0,23x + 0,036$. El ajuste con una función cuadrática da $p(x) = \frac{1}{4}x^2 - \frac{1}{2}x + \frac{1}{4}$, que coincide perfectamente con los puntos de datos, pero está lejos de la curva que buscamos. La razón es que tenemos demasiados parámetros libres en el polinomio, de modo que ajusta demasiado bien los datos. (Observe que, para cualesquiera $n + 1$ puntos dados $(x_1, y_1), \dots, (x_{n+1}, y_{n+1})$ con $x_1 \neq \dots, x_{n+1}$, existe exactamente un polinomio p de grado $\leq n$ tal que $p(x_j) = y_j$ para todo $j = 1, \dots, n + 1$.) Si tuviéramos muchos más puntos de datos e intentáramos ajustar un polinomio a una función lineal, entonces el coeficiente principal debería volverse muy pequeño, pero este efecto no aparece si tenemos muy pocos puntos de datos.

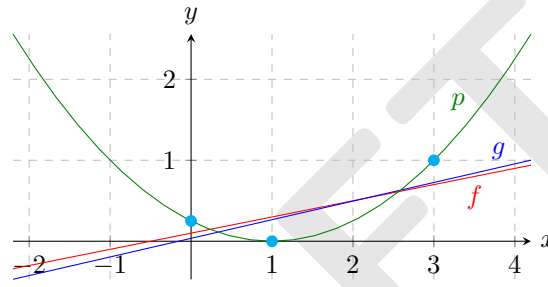


FIGURA 7.7: Ejemplo de sobreajuste cuando tenemos demasiadas variables libres para un conjunto dado de puntos de datos. Los puntos marcan las mediciones que se supone aproximan la curva roja f . Ajustar el polinomio p de grado 2 conduce a la curva verde. La curva azul g es el resultado de un ajuste lineal.

Funciones exponenciales.. Supongamos que se nos dan datos medidos $(x_1, y_1), \dots, (x_n, y_n)$ y queremos encontrar una función de la forma $g(x) = ce^{kx}$ para ajustar nuestros puntos de datos. Sin pérdida de generalidad podemos suponer que $c > 0$ (de lo contrario ajustamos $-g$). Entonces solo necesitamos definir $h(x) = \ln(g(x)) = \ln c + kx$, de modo que podamos usar el método para ajustar una función lineal a los puntos de datos $(x_1, \ln(y_1)), \dots, (x_n, \ln(y_n))$ con el fin de obtener c y k .

Observación 7.47. Mostremos cómo la fórmula del Teorema 7.46 puede derivarse con métodos analíticos. Recordemos que el problema es el siguiente: Sean dados $(x_1, y_1), \dots, (x_n, y_n)$. Encontrar una función lineal $g(x) = ax + b$ que minimice el error total

$$\Delta := \left[\sum_{j=1}^n (y_j - g(x_j))^2 \right]^{\frac{1}{2}} = \left[\sum_{j=1}^n (y_j - [ax_j + b])^2 \right]^{\frac{1}{2}}$$

Consideremos Δ como función de a y b . Entonces tenemos que encontrar el mínimo de

$$\Delta(a, b) = \left[\sum_{j=1}^n (y_j - [ax_j + b])^2 \right]^{\frac{1}{2}}$$

como función de las dos variables a, b . Para simplificar los cálculos un poco, observamos que es suficiente minimizar el cuadrado de Δ porque $\Delta(a, b) \geq 0$ para todo a, b , y por lo tanto es minimal si y solo si su cuadrado sea mínimo. Así que queremos encontrar a, b que minimicen

$$F(a, b) := (\Delta(a, b))^2 = \sum_{j=1}^n (y_j - ax_j - b)^2. \quad (7.15)$$

Para ello, tenemos que derivar F . Como $F : \mathbb{R}^2 \rightarrow \mathbb{R}$, la derivada será una función con valores vectoriales. Obtenemos

$$\begin{aligned} DF(a, b) &= \left(\frac{\partial F}{\partial a}(a, b), \frac{\partial F}{\partial b}(a, b) \right) = \left(\sum_{j=1}^n -2x_j(y_j - ax_j - b), \sum_{j=1}^n -2(y_j - ax_j - b) \right) \\ &= 2 \left(a \sum_{j=1}^n x_j^2 + b \sum_{j=1}^n x_j - \sum_{j=1}^n x_j y_j, a \sum_{j=1}^n x_j + nb - \sum_{j=1}^n y_j \right). \end{aligned}$$

Ahora necesitamos encontrar los puntos críticos, es decir, a, b tales que $DF(a, b) = 0$. Esto ocurre para

$$\left\{ \begin{array}{l} a \sum_{j=1}^n x_j^2 + b \sum_{j=1}^n x_j = \sum_{j=1}^n x_j y_j \\ a \sum_{j=1}^n x_j + bn = \sum_{j=1}^n y_j \end{array} \right\} \quad \text{es decir} \quad \begin{pmatrix} \sum_{j=1}^n x_j^2 & \sum_{j=1}^n x_j \\ \sum_{j=1}^n x_j & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n x_j y_j \\ \sum_{j=1}^n y_j \end{pmatrix}. \quad (7.16)$$

Ahora podemos multiplicar ambos lados por la izquierda por la inversa de la matriz y obtener la solución para a, b . Esto muestra que F tiene un único punto crítico. Como F tiende a infinito cuando $\|(a, b)\| \rightarrow \infty$, la función F debe tener efectivamente un mínimo en este punto crítico. Para más detalles, véase un curso de cálculo vectorial u optimización.

Observamos lo siguiente: Si, como antes, ponemos

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \vec{u} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad A = (\vec{x} | \vec{u}) = \begin{pmatrix} x_1 & 1 \\ \vdots & \\ x_n & 1 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

entonces

$$\sum_{j=1}^n x_j^2 = \langle \vec{x}, \vec{x} \rangle, \quad \sum_{j=1}^n x_j = \langle \vec{x}, \vec{u} \rangle, \quad n = \langle \vec{u}, \vec{u} \rangle, \quad \sum_{j=1}^n x_j y_j = \langle \vec{x}, \vec{y} \rangle, \quad \sum_{j=1}^n y_j = \langle \vec{u}, \vec{y} \rangle.$$

Por lo tanto, las expresiones en la ecuación (7.16) pueden reescribirse como

$$\begin{aligned} \begin{pmatrix} \sum_{j=1}^n x_j^2 & \sum_{j=1}^n x_j \\ \sum_{j=1}^n x_j & n \end{pmatrix} &= \begin{pmatrix} \langle \vec{x}, \vec{x} \rangle & \langle \vec{x}, \vec{u} \rangle \\ \langle \vec{u}, \vec{x} \rangle & \langle \vec{u}, \vec{u} \rangle \end{pmatrix} = \begin{pmatrix} \vec{x} \\ \vec{u} \end{pmatrix} (\vec{x} | \vec{u}) = A^t A \\ \begin{pmatrix} \sum_{j=1}^n x_j y_j \\ \sum_{j=1}^n y_j \end{pmatrix} &= \begin{pmatrix} \langle \vec{x}, \vec{y} \rangle \\ \langle \vec{u}, \vec{y} \rangle \end{pmatrix} = \begin{pmatrix} \vec{x} \\ \vec{u} \end{pmatrix} \vec{y} = A^t \vec{y} \end{aligned}$$

y reconocemos que la ecuación (7.16) es la misma que

$$A^t A \begin{pmatrix} a \\ b \end{pmatrix} = A^t \vec{y}$$

que se vuelve nuestra ecuación (7.14) si multiplicamos ambos lados de la ecuación del lado izquierdo por $(A^t A)^{-1}$.

Ahora debería haber entendido

- qué es el método de mínimos cuadrados,
- cómo se relaciona con las proyecciones ortogonales,
- qué es el sobreajuste,
- etc.

Ahora debería ser capaz de

- ajustar una función lineal a puntos de datos dados,
- ajustar un polinomio a puntos de datos dados,
- ajustar una función exponencial a puntos de datos dados,
- etc.

Ejercicios.

1. Una bola rueda a lo largo del eje x con velocidad constante. A lo largo de la trayectoria de la bola se miden las coordenadas x de la bola en ciertos tiempos t . Las siguientes mediciones son (t en segundos, x en metros):

x	1.5	2.0	3.0	4.0	4.5	6
t	1.4	2.3	4.7	6.6	7.4	10.8

- (a) Dibuje los puntos en el plano tx .
- (b) Use el método de mínimos cuadrados para encontrar la posición inicial x_0 y la velocidad v de la bola.
- (c) Dibuje la recta en el bosquejo anterior. ¿Dónde/Cómo se ven x_0 y v ?

Hint. Recuerde que $x(t) = x_0 + vt$ para un movimiento con velocidad constante.

2. Se supone que una sustancia química inestable decae según la ley $P(t) = P_0 e^{kt}$. Suponga que se hicieron las siguientes mediciones:

t	1	2	3	4	5
P	7.4	6.5	5.7	5.2	4.9

Con el método de mínimos cuadrados aplicado a $\ln(P(t))$, encuentre P_0 y k que mejor corresponden con las mediciones. Dé una estimada para $P(8)$.

3. Con el método de mínimos cuadrados encuentre el polinomio $y = p(x)$ de grado 2 que mejor aproxima los siguientes datos:

x	-2	-1	0	1	2	3	4
y	15	8	2.8	-1.2	-4.9	-7.9	-8.7

7.7. Resumen

Sea U un subespacio de $\mathbb{R}^n$. Entonces su complemento ortogonal se define por

$$U^\perp = \{\vec{x} \in \mathbb{R}^n : \vec{x} \perp \vec{u} \text{ para todo } \vec{u} \in U\}.$$

Para cualquier subespacio $U \subseteq \mathbb{R}^n$ se cumple lo siguiente:

- $U^\perp$ es un espacio vectorial.
- $U^\perp = \ker A$, donde A es cualquier matriz cuyas filas están formadas por una base de U .
- $(U^\perp)^\perp = U$.
- $\dim U + \dim U^\perp = n$.
- $U \oplus U^\perp = \mathbb{R}^n$.
- U tiene una base ortonormal. Una forma de construir tal base es construir primero una base arbitraria de U y luego aplicar el proceso de ortogonalización de Gram-Schmidt para obtener una base ortonormal.

Proyección ortogonal sobre un subespacio $U \subseteq \mathbb{R}^n$

Sea $P_U : \mathbb{R}^n \rightarrow \mathbb{R}^n$ la proyección ortogonal sobre U . Entonces

- P_U es una transformación lineal.
- $P_U \vec{x} \parallel U$ para todo $\vec{x} \in \mathbb{R}^n$.
- $\vec{x} - P_U \vec{x} \perp U$ para todo $\vec{x} \in \mathbb{R}^n$.
- Para todo $\vec{x} \in \mathbb{R}^n$, el punto en U más cercano a $\vec{x}$ está dado por $\vec{x} - P_U \vec{x}$ y $\text{dist}(\vec{x}, U) = \|\vec{x} - P_U \vec{x}\|$.
- Fórmulas para P_U :
 - Si $\vec{u}_1, \dots, \vec{u}_k$ es una base de U , entonces

$$P_U = \langle \vec{u}_1, \cdot \rangle + \dots + \langle \vec{u}_k, \cdot \rangle,$$

es decir $P_U \vec{x} = \langle \vec{u}_1, \vec{x} \rangle + \dots + \langle \vec{u}_k, \vec{x} \rangle$ para todo $\vec{x} \in \mathbb{R}^n$.

- si B es cualquier matriz cuyas columnas forman una base de U , entonces $P_U = B(B^t B)^t$.

Matrices ortogonales

Una matriz $Q \in M(n \times n)$ se llama *matriz ortogonal* si es invertible y si $Q^{-1} = Q^t$. Observe que las siguientes afirmaciones para una matriz $Q \in M(n \times n)$ son equivalentes:

- (I) Q es una matriz ortogonal.
- (II) Q^t es una matriz ortogonal.
- (III) Q^{-1} es una matriz ortogonal.
- (IV) Las columnas de Q son una base ortonormal de $\mathbb{R}^n$.
- (V) Las filas de Q son una base ortonormal de $\mathbb{R}^n$.
- (VI) Q preserva productos internos, es decir $\langle \vec{x}, \vec{y} \rangle = \langle Q\vec{x}, Q\vec{y} \rangle$ for all $\vec{x}, \vec{y} \in \mathbb{R}^n$.
- (VII) Q preserva longitudes, es decir $\|\vec{x}\| = \|Q\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$.

Toda matriz ortogonal representa una rotación (en este caso su determinante es 1) o una composición de una rotación con una reflexión (en este caso su determinante es -1).

7.8. Ejercicios

1. (a) Complete $\begin{pmatrix} 1/4 \\ \sqrt{15/16} \end{pmatrix}$ a una base ortonormal para $\mathbb{R}^2$. ¿Cuántas posibilidades hay para hacerlo?
- (b) Complete $\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$ a una base ortonormal para $\mathbb{R}^3$. ¿Cuántas posibilidades hay para hacerlo?
- (c) Complete $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}$ a una base ortonormal para $\mathbb{R}^3$. ¿Cuántas posibilidades hay para hacerlo?
2. Encuentre una base para el complemento ortogonal de los siguientes espacios vectoriales. Encuentre la dimensión del espacio y la dimensión de su complemento ortogonal.
 - (a) $U = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \right\} \subseteq \mathbb{R}^4$,
 - (b) $U = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \right\} \subseteq \mathbb{R}^4$.
3. (a) Sea $U = \{(x, y, z)^t \in \mathbb{R}^3 : x + 2y + 3z = 0\} \subseteq \mathbb{R}^3$.
 - 1) Sea $\vec{v} = (0, 2, 5)^t$. Encuentre el punto $\vec{x} \in U$ que esté más cercano a $\vec{v}$ y calcule la distancia entre $\vec{v}$ y $\vec{x}$.
 - 2) ¿Hay un punto $\vec{y} \in U$ que esté a una distancia máxima de $\vec{v}$?
 - 3) Encuentre la matriz que representa la proyección ortogonal sobre U (en la base estándar).

(b) Sea $W = \text{gen}\{(1, 1, 1, 1)^t, (2, 1, 1, 0)^t\} \subseteq \mathbb{R}^4$.

- 1) Encuentre una base ortogonal de W .
- 2) Sean $\vec{a}_1 = (1, 2, 0, 1)^t$, $\vec{a}_2 = (11, 4, 4, -3)^t$, $\vec{a}_3 = (0, -1, -1, 0)^t$. Para cada $j = 1, 2, 3$ encuentre el punto $\vec{w}_j \in W$ que esté más cercano a $\vec{a}_j$ y calcule la distancia entre $\vec{a}_j$ y $\vec{w}_j$.
- 3) Encuentre la matriz que representa la proyección ortogonal sobre W (en la base estándar).

4. Sean $\vec{v} = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 1 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 5 \end{pmatrix}$, $\vec{a} = \begin{pmatrix} 0 \\ 3 \\ 4 \\ 0 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 3 \end{pmatrix}$, $\vec{c} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$, $\vec{d} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$.

- (a) Demuestre que $\vec{v}$ y $\vec{w}$ son linealmente independientes y encuentre una base ortonormal de $U = \text{span}\{\vec{v}, \vec{w}\} \subseteq \mathbb{R}^4$.
- (b) Demuestre que $\vec{a}$, $\vec{b}$, $\vec{c}$ y $\vec{d}$ son linealmente independientes. Use el proceso de Gram-Schmidt para encontrar una base ortonormal de $U = \text{span}\{\vec{a}, \vec{b}, \vec{c}, \vec{d}\} \subseteq \mathbb{R}^5$. Encuentre una base de $U^\perp$.

5. Encuentre una base ortonormal de $U^\perp$ donde $U = \text{gen}\{(1, 0, 2, 4)^t\} \subseteq \mathbb{R}^4$.

6. (a) Sea $\varphi \in \mathbb{R}$ y sean $\vec{v}_1 = \begin{pmatrix} \cos \varphi \\ -\sin \varphi \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} \sin \varphi \\ \cos \varphi \end{pmatrix}$. Demuestre que $\vec{v}_1, \vec{v}_2$ es una base ortonormal de $\mathbb{R}^2$.
- (b) Sea $\alpha \in \mathbb{R}$. Encuentre la matriz $Q(\alpha) \in M(2 \times 2)$ que describe rotación por α contra las manecillas del reloj.
- (c) Sean $\alpha, \beta \in \mathbb{R}$. Explique por qué es claro que $Q(\alpha)Q(\beta) = Q(\alpha + \beta)$. Use esta relación para concluir las identidades trigonométricas

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta, \quad \sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta.$$

7. Sean $O(n) = \{Q \in M(n \times n) : Q \text{ es matriz ortogonal}\}$ y $SO(n) = \{Q \in O(n) : \det Q = 1\}$.

- (a) Demuestre que $O(n)$ con la composición es un grupo. Es decir, hay que probar que:
 - 1) Para todo $Q, R \in O(n)$, la composición QR es un elemento en $O(n)$.
 - 2) Existe un $E \in O(n)$ tal que $QE = Q$ y $EQ = Q$ para todo $Q \in O(n)$.
 - 3) Para todo $Q \in O(n)$ existe un elemento inverso $\tilde{Q}$ tal que $\tilde{Q}Q = Q\tilde{Q} = E$.
- (b) ¿Es $O(n)$ conmutativo (es decir, se tiene $QR = RQ$ para todo $Q, R \in O(n)$)?
- (c) Demuestre que $SO(n)$ con la composición es un grupo.

8. Sea $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una isometría. Demuestre que T es inyectivo y que $m \geq n$.

DRAFT

Capítulo 8

Matrices simétricas y diagonalización

En este capítulo trabajamos principalmente en $\mathbb{R}^n$ y en $\mathbb{C}^n$. Escribimos $M_{\mathbb{R}}(n \times n)$ o $M_{\mathbb{C}}(n \times n)$ solo si importa si la matriz bajo consideración es una matriz real o compleja.

La primera sección se dedica a $\mathbb{C}^n$. Ya sabemos que es un espacio vectorial sobre $\mathbb{C}$, pero ahora introducimos adicionalmente el producto interno en él. Además definimos matrices hermíticas y matrices unitarias en $\mathbb{C}^n$ que son los análogos a matrices simétricas y matrices ortogonales en $\mathbb{R}^n$. Definimos valores propios y vectores propios en la Sección 8.3. Resulta que es más conveniente trabajar sobre $\mathbb{C}$ porque los valores propios son ceros del llamado polinomio característico y en $\mathbb{C}$ todo polinomio tiene un cero. El teorema principal es el Teorema 8.49 que dice que una matriz de $n \times n$ es diagonalizable si y solo si tiene suficientes vectores propios para generar $\mathbb{C}^n$ (o $\mathbb{R}^n$). Resulta que toda matriz simétrica y toda matriz hermítica es diagonalizable.

Terminamos el capítulo con una aplicación de la diagonalización ortogonal a la solución de ecuaciones cuadráticas en dos variables.

8.1. Espacios vectoriales complejos

En esta sección introducimos $\mathbb{C}^n$ como un espacio con producto interno porque algunos cálculos para valores propios later en este capítulo son más naturales en $\mathbb{C}^n$ que en $\mathbb{R}^n$. Puede saltarse de la mayoría de esta sección. La parte más importante es la definición del producto interno en $\mathbb{C}^n$, la noción de ortogonalidad derivado de él, y el concepto de matrices hermíticas y de matrices unitarias.

En analogía a $\mathbb{R}^n$, definimos el *espacio vectorial* $\mathbb{C}^n$ como el conjunto

$$\mathbb{C}^n = \left\{ \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} : z_1, \dots, z_n \in \mathbb{C} \right\}$$

junto con la suma y multiplicación por un escalar $c \in \mathbb{C}$:

$$\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} + \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} := \begin{pmatrix} w_1 + z_1 \\ \vdots \\ w_n + z_n \end{pmatrix}, \quad c \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} := \begin{pmatrix} cz_1 \\ \vdots \\ cz_n \end{pmatrix}.$$

Es fácil verificar que $\mathbb{C}^n$ junto con estas operaciones satisface los axiomas de espacio vectorial de Definición 5.1 con $\mathbb{K} = \mathbb{C}$, por lo tanto es un espacio vectorial complejo. En particular, tenemos conceptos como independencia lineal de vectores, base y dimensión de $\mathbb{C}^n$, etc.

A continuación introducimos un producto interno en $\mathbb{C}^n$. Como en el caso de los vectores reales, nos gustaría interpretar $\langle \vec{z}, \vec{z} \rangle$ como el cuadrado de la norma de $\vec{z}$. En particular, debería ser un número real no negativo. Para $\mathbb{C}^1 = \mathbb{C}$, los vectores son simplemente números complejos $\vec{z} = z_1$ y nos gustaría tener $\langle \vec{z}, \vec{z} \rangle = |z_1|^2 = z_1 \bar{z}_1$ donde $\bar{z}$ es el conjugado complejo del número complejo z . Esto nos motiva a definir el producto interno en $\mathbb{C}^n$ como sigue.

Definición 8.1 (Producto interno y norma de un vector en $\mathbb{C}^n$). Para vectores $\vec{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$

y $\vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n$ el *producto interno* (o *producto escalar* o *producto punto*) se define como

$$\langle \vec{z}, \vec{w} \rangle = \left\langle \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \right\rangle = \sum_{j=1}^n z_j \bar{w}_j = z_1 \bar{w}_1 + \cdots + z_n \bar{w}_n.$$

La *longitud* de $\vec{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{C}^n$ se denota por $\|\vec{z}\|$ y está dada por

$$\|\vec{z}\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

Otros nombres para la longitud de $\vec{z}$ son *magnitud de $\vec{z}$* o *norma de $\vec{z}$* .

Ejercicio 8.2. Muestre que el producto escalar de Definición 8.1 puede verse como una extensión del producto escalar en $\mathbb{R}^n$ en el siguiente sentido: Si las componentes de $\vec{z}$ y $\vec{v}$ resultan ser reales, entonces también pueden verse como vectores en $\mathbb{R}^n$. La afirmación es que su producto escalar como vectores en $\mathbb{R}^n$ es igual a su producto escalar en $\mathbb{C}^n$. Lo mismo es cierto para sus normas.

Propiedades 8.3. (I) Norma de un vector: Para todos los vectores $\vec{z} \in \mathbb{C}^n$, tenemos que

$$\langle \vec{z}, \vec{z} \rangle = \|\vec{z}\|^2.$$

(II) Simetría del producto interno: Para todos los vectores $\vec{v}, \vec{w} \in \mathbb{C}^n$, tenemos (¡nótese la conjugación compleja en el lado derecho!)

$$\langle \vec{v}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{v} \rangle}.$$

(III) Sesquilinealidad del producto interno: Para todos los vectores $\vec{u}, \vec{v}, \vec{z} \in \mathbb{C}^n$ y todo $c \in \mathbb{C}$, tenemos que

$$\langle \vec{v} + c\vec{w}, \vec{z} \rangle = \langle \vec{v}, \vec{z} \rangle + c\langle \vec{w}, \vec{z} \rangle \quad y \quad \langle \vec{v}, \vec{w} + c\vec{z} \rangle = \langle \vec{v}, \vec{w} \rangle + \bar{c}\langle \vec{v}, \vec{z} \rangle.$$

(IV) Para todos los vectores $\vec{v}, \vec{w} \in \mathbb{C}^n$ y $c \in \mathbb{C}$, tenemos que $\|c\vec{v}\| = |c|\|\vec{v}\|$.

Demostración. Sea $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$, $\vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$, $\vec{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{C}^n$ y sea $c \in \mathbb{C}$.

$$(I) \quad \langle \vec{z}, \vec{z} \rangle = z_1 \bar{z}_1 + \cdots + z_n \bar{z}_n = |z_1|^2 + \cdots + |z_n|^2 = \|\vec{z}\|^2.$$

$$(II) \quad \langle \vec{v}, \vec{w} \rangle = v_1 \bar{w}_1 + \cdots + v_n \bar{w}_n = \overline{\bar{v}_1 w_1 + \cdots + \bar{v}_n w_n} = \overline{w_1 \bar{v}_1 + \cdots + w_n \bar{v}_n} = \overline{\langle \vec{w}, \vec{v} \rangle}.$$

(III) Un cálculo directo muestra que

$$\begin{aligned} \langle \vec{v} + c\vec{w}, \vec{z} \rangle &= (v_1 + cw_1)\bar{z}_1 + \cdots + (v_n + cw_n)\bar{z}_n \\ &= v_1 \bar{z}_1 + \cdots + v_n \bar{z}_n + c w_1 \bar{z}_1 + \cdots + c w_n \bar{z}_n \\ &= \langle \vec{v}, \vec{z} \rangle + c \langle \vec{w}, \vec{z} \rangle. \end{aligned}$$

La segunda ecuación puede mostrarse mediante un cálculo análogo. En lugar de repetirlos, también podemos usar la propiedad de simetría del producto interno:

$$\langle \vec{v}, \vec{w} + c\vec{z} \rangle = \overline{\langle \vec{w} + c\vec{z}, \vec{v} \rangle} = \overline{\langle \vec{w}, \vec{v} \rangle + c\langle \vec{z}, \vec{v} \rangle} = \overline{\langle \vec{w}, \vec{v} \rangle} + \bar{c}\overline{\langle \vec{z}, \vec{v} \rangle} = \langle \vec{v}, \vec{z} \rangle + \bar{c}\langle \vec{v}, \vec{z} \rangle.$$

(IV) $\|c\vec{z}\|^2 = \langle c\vec{z}, c\vec{z} \rangle = c\bar{c}\langle \vec{z}, \vec{z} \rangle = |c|^2\|\vec{z}\|^2$. Tomando la raíz cuadrada en ambos lados, obtenemos la igualdad deseada $\|c\vec{z}\| = |c|\|\vec{z}\|$. $\square$

Para $\mathbb{C}^n$ no hay teorema del coseno y en general no tiene demasiado sentido hablar del ángulo entre dos vectores complejos (la ortogonalidad todavía tiene sentido!).

Definición 8.4. Sea $\vec{z}, \vec{v} \in \mathbb{C}^n$.

(I) Los vectores $\vec{z}, \vec{v}$ se llaman *ortogonales* o *perpendiculares* si $\langle \vec{z}, \vec{v} \rangle = 0$. En este caso escribimos $\vec{z} \perp \vec{v}$.

(II) Si $\vec{v} \neq \vec{0}$, entonces la *proyección ortogonal de $\vec{z}$ sobre $\vec{v}$* es $\text{proy}_{\vec{v}} \vec{z} = \frac{\langle \vec{z}, \vec{v} \rangle}{\|\vec{v}\|^2} \vec{v}$.

La siguiente proposición muestra que la ortogonalidad funciona en $\mathbb{C}^n$ como se espera.

Proposición 8.5. Sea $\vec{z}, \vec{v} \in \mathbb{C}^n$.

(I) Teorema de Pitágoras: Si $\vec{z} \perp \vec{v}$, entonces $\|\vec{z} + \vec{v}\|^2 = \|\vec{z}\|^2 + \|\vec{v}\|^2$.

(II) Si $\vec{v} \neq \vec{0}$, entonces $\vec{z} = \vec{z}_{\parallel} + \vec{z}_{\perp}$ con $\vec{z}_{\parallel} := \text{proy}_{\vec{v}} \vec{z}$ y $\vec{z}_{\perp} := \vec{z} - \text{proy}_{\vec{v}} \vec{z}$ y

$$\text{proy}_{\vec{v}} \vec{z} \parallel \vec{v}, \quad y \quad \vec{z} - \text{proy}_{\vec{v}} \vec{z} \perp \vec{v}.$$

Además, $\|\text{proy}_{\vec{v}} \vec{z}\| \leq \|\vec{z}\|$.

(III) Si $\vec{v} \neq \vec{0}$, entonces $\mathbb{C}^n \rightarrow \mathbb{C}^n$, $\vec{z} \mapsto \text{proy}_{\vec{v}} \vec{z}$ es una transformación lineal.

Demostración. (I) Si $\vec{z} \perp \vec{v}$, entonces $\|\vec{z} + \vec{v}\|^2 = \langle \vec{z}, \vec{z} \rangle + \langle \vec{z}, \vec{v} \rangle + \langle \vec{v}, \vec{z} \rangle + \langle \vec{v}, \vec{v} \rangle = \langle \vec{z}, \vec{z} \rangle + \langle \vec{v}, \vec{v} \rangle = \|\vec{z}\|^2 + \|\vec{v}\|^2$.

(II) Es claro que $\vec{z} = \vec{z}_{\parallel} + \vec{z}_{\perp}$ y que $\vec{z}_{\parallel} \parallel \vec{v}$ por definición de $\vec{z}_{\parallel}$ y $\vec{z}_{\perp}$. Que $\vec{z}_{\perp} \perp \vec{v}$ se sigue de

$$\langle \vec{z}_{\perp}, \vec{v} \rangle = \langle \vec{z} - \text{proy}_{\vec{v}} \vec{z}, \vec{v} \rangle = \langle \vec{z}, \vec{v} \rangle - \langle \text{proy}_{\vec{v}} \vec{z}, \vec{v} \rangle = \langle \vec{z}, \vec{v} \rangle - \frac{\langle \vec{z}, \vec{v} \rangle}{\|\vec{v}\|^2} \langle \vec{v}, \vec{v} \rangle = \langle \vec{z}, \vec{v} \rangle - \langle \vec{z}, \vec{v} \rangle = 0.$$

Finalmente, por el teorema de Pitágoras,

$$\|\vec{z}\|^2 = \|(\vec{z} - \text{proy}_{\vec{v}} \vec{z}) + \text{proy}_{\vec{v}} \vec{z}\|^2 = \|\vec{z} - \text{proy}_{\vec{v}} \vec{z}\|^2 + \|\text{proy}_{\vec{v}} \vec{z}\|^2 \geq \|\text{proy}_{\vec{v}} \vec{z}\|^2.$$

(III) Suponga que $\vec{v} \neq \vec{0}$ y sean $\vec{z}_1, \vec{z}_2 \in \mathbb{C}^n$ y $c \in \mathbb{C}$. Entonces

$$\begin{aligned} \text{proy}_{\vec{v}}(\vec{z}_1 + c\vec{z}_2) &= \frac{\langle \vec{z}_1 + c\vec{z}_2, \vec{v} \rangle}{\|\vec{v}\|^2} \vec{v} = \frac{\langle \vec{z}_1, \vec{v} \rangle + c\langle \vec{z}_2, \vec{v} \rangle}{\|\vec{v}\|^2} \vec{v} = \frac{\langle \vec{z}_1, \vec{v} \rangle \vec{v}}{\|\vec{v}\|^2} + \frac{c\langle \vec{z}_2, \vec{v} \rangle \vec{v}}{\|\vec{v}\|^2} \\ &= \text{proy}_{\vec{v}} \vec{z}_1 + c \text{proy}_{\vec{v}} \vec{z}_2. \end{aligned}$$

□

Question 8.1

¿Qué cambia si en la definición de la proyección ortogonal ponemos $\langle \vec{v}, \vec{z} \rangle$ en lugar de $\langle \vec{z}, \vec{v} \rangle$?

Ahora mostremos la desigualdad triangular. Nótese que las siguientes desigualdades (8.1) y (8.2) fueron demostradas para espacios vectoriales reales en Corollary 2.22 usando el teorema del coseno.

Proposición 8.6. Para todos los vectores $\vec{v}, \vec{w} \in \mathbb{C}^n$ y $c \in \mathbb{C}$, tenemos la desigualdad de Cauchy-Schwarz (que es un caso especial de la llamada desigualdad de Hölder)

$$|\langle \vec{v}, \vec{w} \rangle| \leq \|\vec{v}\| \|\vec{w}\| \quad (8.1)$$

y la desigualdad triangular

$$\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|. \quad (8.2)$$

Demostración. Primero mostramos (8.1). Es obviamente cierto si $\vec{w} = \vec{0}$ porque en este caso ambos lados de la desigualdad son iguales a 0. Así que supongamos ahora que $\vec{w} \neq \vec{0}$. Nótese que para cualquier $\lambda \in \mathbb{C}$ tenemos que

$$0 \leq \|\vec{v} - \lambda \vec{w}\|^2 = \langle \vec{v} - \lambda \vec{w}, \vec{v} - \lambda \vec{w} \rangle = \|\vec{v}\|^2 - \lambda \langle \vec{w}, \vec{v} \rangle - \bar{\lambda} \langle \vec{v}, \vec{w} \rangle + |\lambda|^2 \|\vec{w}\|^2.$$

Si elegimos $\lambda = -\frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2}$, obtenemos

$$\begin{aligned} 0 &\leq \|\vec{v}\|^2 - \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \langle \vec{w}, \vec{v} \rangle - \frac{\overline{\langle \vec{v}, \vec{w} \rangle}}{\|\vec{w}\|^2} \langle \vec{v}, \vec{w} \rangle + \frac{|\langle \vec{v}, \vec{w} \rangle|^2}{\|\vec{w}\|^4} \|\vec{w}\|^2 \\ &= \|\vec{v}\|^2 - 2 \frac{|\langle \vec{v}, \vec{w} \rangle|^2}{\|\vec{w}\|^2} + \frac{|\langle \vec{v}, \vec{w} \rangle|^2}{\|\vec{w}\|^4} \|\vec{w}\|^2 \\ &= \|\vec{v}\|^2 - \frac{|\langle \vec{v}, \vec{w} \rangle|^2}{\|\vec{w}\|^2} = \frac{1}{\|\vec{w}\|^2} [\|\vec{v}\|^2 \|\vec{w}\|^2 - |\langle \vec{v}, \vec{w} \rangle|^2] \end{aligned}$$

Se sigue que $\|v\|^2\|w\|^2 - |\langle \vec{v}, \vec{w} \rangle|^2 \geq 0$, por lo tanto $\|v\|^2\|w\|^2 \geq |\langle \vec{v}, \vec{w} \rangle|^2$. Obtenemos la desigualdad deseada tomando raíz cuadrada.

Ahora mostremos la desigualdad triangular. Es esencialmente lo mismo que para vectores en $\mathbb{R}^n$, cf. Corollary 2.22.

$$\begin{aligned} \|\vec{v} + \vec{w}\|^2 &= \langle \vec{v} + \vec{w}, \vec{v} + \vec{w} \rangle = \langle \vec{v}, \vec{v} \rangle + \langle \vec{v}, \vec{w} \rangle + \langle \vec{w}, \vec{v} \rangle + \langle \vec{w}, \vec{w} \rangle \\ &= \langle \vec{v}, \vec{v} \rangle + \langle \vec{v}, \vec{w} \rangle + \overline{\langle \vec{v}, \vec{w} \rangle} + \langle \vec{w}, \vec{w} \rangle \\ &= \|\vec{v}\|^2 + 2 \operatorname{Re} \langle \vec{v}, \vec{w} \rangle + \|\vec{w}\|^2 \\ &\leq \|\vec{v}\|^2 + 2|\langle \vec{v}, \vec{w} \rangle| + \|\vec{w}\|^2 \leq \|\vec{v}\|^2 + 2\|\vec{v}\| \|\vec{w}\| + \|\vec{w}\|^2 = (\|\vec{v}\| + \|\vec{w}\|)^2. \end{aligned}$$

En la primera desigualdad usamos que $\operatorname{Re} a \leq |a|$ para cualquier número complejo a y en la segunda desigualdad usamos (8.1). Si tomamos raíz cuadrada en ambos lados obtenemos la desigualdad triangular. $\square$

Observación 8.7. Observe que la elección de λ en la demostración de (8.1) no es tan arbitraria como puede parecer. Observe que para este particular λ

$$\vec{v} - \lambda \vec{w} = \vec{v} - \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w} = \vec{v} - \operatorname{proy}_{\vec{w}} \vec{v}.$$

Por lo tanto, esta elección de λ minimiza la norma de $\vec{v} - \lambda \vec{w}$ y $\vec{v} - \operatorname{proy}_{\vec{w}} \vec{v} \perp \vec{w}$. Por lo tanto, por Pitágoras,

$$\begin{aligned} \|\vec{v}\|^2 &= \|(\vec{v} - \operatorname{proy}_{\vec{w}} \vec{v}) + \operatorname{proy}_{\vec{w}} \vec{v}\|^2 = \|\vec{v} - \operatorname{proy}_{\vec{w}} \vec{v}\|^2 + \|\operatorname{proy}_{\vec{w}} \vec{v}\|^2 \\ &\geq \|\operatorname{proy}_{\vec{w}} \vec{v}\|^2 = \left\| \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \vec{w} \right\|^2 = \frac{|\langle \vec{v}, \vec{w} \rangle|^2}{\|\vec{w}\|^2} \end{aligned}$$

lo cual muestra que $\|\vec{v}\|^2\|\vec{w}\|^2 \geq |\langle \vec{v}, \vec{w} \rangle|^2$.

Otra forma de ver esta desigualdad es

$$\begin{aligned} 0 &\leq \|\vec{v} - \operatorname{proy}_{\vec{w}} \vec{v}\|^2 = \langle \vec{v} - \operatorname{proy}_{\vec{w}} \vec{v}, \vec{v} - \operatorname{proy}_{\vec{w}} \vec{v} \rangle = \langle \vec{v} - \operatorname{proy}_{\vec{w}} \vec{v}, \vec{v} \rangle = \|\vec{v}\|^2 - \langle \operatorname{proy}_{\vec{w}} \vec{v}, \vec{v} \rangle \\ &= \|\vec{v}\|^2 - \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{w}\|^2} \langle \vec{w}, \vec{v} \rangle = \|\vec{v}\|^2 - \frac{|\langle \vec{v}, \vec{w} \rangle|^2}{\|\vec{w}\|^2} \end{aligned}$$

lo cual nuevamente da $\|\vec{v}\|^2\|\vec{w}\|^2 \geq |\langle \vec{v}, \vec{w} \rangle|^2$.

Clases importantes de matrices

Recordemos que para una matriz $A \in M_{\mathbb{R}}(m \times n)$ definimos su transpuesta A^t . La propiedad importante de A^t es que es la única matriz tal que

$$\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A^t \vec{y} \rangle \quad \text{para todo } \vec{x} \in \mathbb{R}^n, \vec{y} \in \mathbb{R}^m.$$

En el caso complejo, para una matriz dada $A \in M_{\mathbb{C}}(m \times n)$ queremos una matriz A^* tal que

$$\langle A\vec{x}, \vec{y} \rangle = \langle \vec{x}, A^* \vec{y} \rangle \quad \text{para todo } \vec{x} \in \mathbb{C}^n, \vec{y} \in \mathbb{C}^m.$$

Es fácil verificar que debemos tomar $A^* = \overline{A}^t$, donde $\overline{A}$ es la matriz que obtenemos de A tomando el conjugado complejo de cada entrada. Claramente, si todas las entradas de A son números reales, entonces $A^t = A^*$.

Definición 8.8. La matriz A^* se llama *matriz adjunta* de A .

Atención. Esta matriz no se debe confundir con la “matriz adjunta definida” en Definición 4.14. Ver comentario después de Definición 4.14.

Lema 8.9. Sea $A \in M(n \times n)$. Entonces $\det(A^*) = \overline{\det A} = \text{conjugado complejo de } \det A$.

Demostración. $\det A^* = \det(\overline{A})^t = \det \overline{A} = \overline{\det A}$. La última igualdad se sigue directamente de la definición del determinante. $\square$

Una matriz con entradas reales es simétrica si y solo si $A = A^t$. El análogo para matrices complejas son las matrices hermíticas.

Definición 8.10. Una matriz $A \in M(n \times n)$ se llama *hermítica* si $A = A^*$.

Ejemplos 8.11. ■ $A = \begin{pmatrix} 1 & 2+3i \\ 5 & 1-7i \end{pmatrix} \implies A^* = \begin{pmatrix} 1 & 5 \\ 2-3i & 1+7i \end{pmatrix}$. La matriz A no es hermítica.

■ $A = \begin{pmatrix} 1 & 2+3i \\ 2-3i & 5 \end{pmatrix} \implies A^* = \begin{pmatrix} 1 & 2+3i \\ 2-3i & 1+7i \end{pmatrix}$. La matriz A es hermítica.

Ejercicio 8.12. ■ Muestre que las entradas en la diagonal de una matriz hermítica deben ser reales.

■ Muestre que el determinante de una matriz hermítica es un número real.

Otra clase importante de matrices reales son las matrices ortogonales. Recordemos que una matriz $Q \in M_{\mathbb{R}}(n \times n)$ es una matriz ortogonal si y solo si $Q^t = Q^{-1}$. Vimos que si Q es ortogonal, entonces sus columnas (o filas) forman una base ortonormal para $\mathbb{R}^n$ y que $|\det Q| = 1$, por lo tanto $\det Q = \pm 1$.

El análogo en espacios vectoriales complejos son las llamadas matrices unitarias.

Definición 8.13. Una matriz $Q \in M(n \times n)$ se llama *unitaria* si $Q^* = Q^{-1}$.

Está claro a partir de la definición que una matriz es unitaria si y solo si sus columnas (o filas) forman una base ortonormal para $\mathbb{C}^n$, cf. Teorema 7.12.

Proposición 8.14. Sea $Q \in M(n \times n)$.

(I) Lo siguiente es equivalente:

- (a) Q es unitaria.
- (b) $\langle Q\vec{x}, Q\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle$ para todo $\vec{x}, \vec{y} \in \mathbb{R}^n$.
- (c) $\|Q\vec{x}\| = \|\vec{x}\|$ para todo $\vec{x} \in \mathbb{R}^n$.

(II) Si Q es unitaria, entonces $|\det Q| = 1$.

Demostración. (I) (a) $\implies$ (b): Suponga que Q es una matriz unitaria y sean $\vec{x}, \vec{y} \in \mathbb{C}^n$. Entonces

$$\langle Q\vec{x}, Q\vec{y} \rangle = \langle Q^*Q\vec{x}, \vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle.$$

(b) $\implies$ (a): Fijemos $\vec{x} \in \mathbb{C}^n$. Entonces tenemos $\langle Q\vec{x}, Q\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle$ para todo $\vec{y} \in \mathbb{C}^n$, por lo tanto

$$0 = \langle Q\vec{x}, Q\vec{y} \rangle - \langle \vec{x}, \vec{y} \rangle = \langle Q^*Q\vec{x}, \vec{y} \rangle - \langle \vec{x}, \vec{y} \rangle = \langle Q^*Q\vec{x} - \vec{x}, \vec{y} \rangle = \langle (Q^*Q - \text{id})\vec{x}, \vec{y} \rangle.$$

Como esto es cierto para cualquier $\vec{y} \in \mathbb{C}^n$, se sigue que $(Q^*Q - \text{id})\vec{x} = 0$. Como $\vec{x} \in \mathbb{C}^n$ era arbitrario, concluimos que $Q^*Q - \text{id} = 0$, en otras palabras, que $Q^*Q = \text{id}$.

(b) $\implies$ (c): Se sigue de (b) que $\|Q\vec{x}\|^2 = \langle Q\vec{x}, Q\vec{x} \rangle = \langle \vec{x}, \vec{x} \rangle = \|\vec{x}\|^2$, por lo tanto $\|Q\vec{x}\| = \|\vec{x}\|$.

(c) $\implies$ (b): Observe que el producto interno de dos vectores en $\mathbb{C}^n$ puede expresarse compesadamente en términos de normas como sigue

$$\langle \vec{a}, \vec{b} \rangle = \frac{1}{4} \left[\|\vec{a} + \vec{b}\|^2 - \|\vec{a} - \vec{b}\|^2 + i\|\vec{a} + i\vec{b}\|^2 - i\|\vec{a} - i\vec{b}\|^2 \right]$$

como puede verificarse fácilmente. Por lo tanto encontramos

$$\begin{aligned} \langle Q\vec{x}, Q\vec{y} \rangle &= \frac{1}{4} \left[\|Q\vec{x} + Q\vec{y}\|^2 - \|Q\vec{x} - Q\vec{y}\|^2 + i\|Q\vec{x} + iQ\vec{y}\|^2 - i\|Q\vec{x} - iQ\vec{y}\|^2 \right] \\ &= \frac{1}{4} \left[\|Q(\vec{x} + \vec{y})\|^2 - \|Q(\vec{x} - \vec{y})\|^2 + i\|Q(\vec{x} + i\vec{y})\|^2 - i\|Q(\vec{x} - i\vec{y})\|^2 \right] \\ &= \frac{1}{4} \left[\|\vec{x} + \vec{y}\|^2 - \|\vec{x} - \vec{y}\|^2 + i\|\vec{x} + i\vec{y}\|^2 - i\|\vec{x} - i\vec{y}\|^2 \right] \\ &= \langle \vec{x}, \vec{y} \rangle. \end{aligned}$$

(II) Suponga que Q es unitaria. Entonces

$$1 = \det \text{id} = \det QQ^* = (\det Q)(\det Q^*) = (\det Q)(\overline{\det Q}) = |\det Q|^2. \quad \square$$

Ejemplos 8.15. ■ La matriz $Q = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$ es unitaria porque $QQ^* = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, por lo tanto $Q^* = Q^{-1}$. Nótese que $\det Q = -i^2 = 1$.

■ La matriz $Q = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{i\beta} \end{pmatrix}$ es unitaria porque $QQ^* = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{i\beta} \end{pmatrix} \begin{pmatrix} e^{-i\alpha} & 0 \\ 0 & e^{-i\beta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, por lo tanto $Q^* = Q^{-1}$. Nótese que $\det Q = e^{i(\alpha+\beta)}$, por lo tanto $|\det Q| = 1$.

Ejemplos 8.16. ■ La matrix $Q = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$ es unitaria porque $QQ^* = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, por lo tanto $Q^* = Q^{-1}$. Note que $\det Q = -i^2 = 1$.

■ La matrix $Q = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{i\beta} \end{pmatrix}$ es unitaria porque $QQ^* = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{i\beta} \end{pmatrix} \begin{pmatrix} e^{-i\alpha} & 0 \\ 0 & e^{-i\beta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, mostrando que $Q^* = Q^{-1}$. Observe que $\det Q = e^{i(\alpha+\beta)}$, por lo tanto $|\det Q| = 1$.

Ahora debería haber entendido

- la estructura de $\mathbb{C}^n$ como espacio vectorial,
- el producto interno en $\mathbb{C}^n$,
- que el concepto de ortogonalidad tiene sentido en $\mathbb{C}^n$ y funciona como en $\mathbb{R}^n$,
- por qué matrices hermiticas en $\mathbb{C}^n$ juegan el papel de simétricas en $\mathbb{R}^n$,
- por qué matrices unitarias en $\mathbb{C}^n$ juegan el papel de ortogonales en $\mathbb{R}^n$,
- etc.

Ahora debería ser capaz de

- calcular con vectors en $\mathbb{C}^n$,
- verificar si vectors en $\mathbb{C}^n$ son ortogonales,
- calcular la proyección ortogonal de un vector sobre otro,
- verificar si una matriz dada es hermitica,
- verificar si una matriz dada es unitaria,
- etc.

Ejercicios.

1. Sea

$$A = \begin{pmatrix} i & -3+i & 2i \\ 2 & 4i & 2+2i \\ 4+i & 2-i & 6-i \end{pmatrix}.$$

Encuentre bases para la imagen del espacio columna y el kernel de A .

2. Sea

$$A = \begin{pmatrix} 1-i & 1+i & 8 \\ 3i & -3 & -12+12i \\ 4+i & -1+4i & 12+20i \end{pmatrix}$$

Encuentre bases ortonormales para la imagen del espacio columna y el kernel de A .

3. Sean $v_1 = \begin{pmatrix} 1 \\ 1-i \\ -3 \\ i \end{pmatrix}$ y $v_2 = \begin{pmatrix} i \\ 1+i \\ 1 \\ 3+3i \end{pmatrix}$ y $V = \text{span}\{v_1, v_2\}$. Verifique que $v_1 \perp v_2$ y obtenga una base ortonormal de $V^\perp$.

4. Encuentre $a, b, c, d, e, f \in \mathbb{R}$ tales que la matriz

$$A = \begin{pmatrix} 1 & 3a+ib & 1+3i \\ 7a-5ib-4 & 3+ic & 5e+3if+2i \\ 1-3i & 4e-6if+2-8i & 4-ic \end{pmatrix}$$

sea hermitiana.

5. Verifique que la matriz

$$\frac{1}{2} \begin{pmatrix} 1 & -i & -1+i \\ i & 1 & 1+i \\ 1-i & 1+i & 0 \end{pmatrix}$$

es unitaria.

6. Considere $V = \mathbb{C}^2$ y $T : V \rightarrow V$ dada por $T \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} -z_2 \\ \bar{z}_1 \end{pmatrix}$.

(a) ¿ T es transformación lineal si V se considera sobre $\mathbb{K} = \mathbb{R}$?

(b) ¿ T es transformación lineal si V se considera sobre $\mathbb{K} = \mathbb{C}$?

7. Sean $\vec{x}, \vec{y} \in V$ donde $V = \mathbb{R}^n$ ó $\mathbb{C}^n$.

(a) Si $V = \mathbb{C}^n$, muestre que $\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + 2\operatorname{Re}\langle \vec{x}, \vec{y} \rangle + \|\vec{y}\|^2$.

(b) Si $V = \mathbb{R}^n$, muestre que $\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$ si y solo si $\vec{x} \perp \vec{y}$.

(c) Si $V = \mathbb{C}^n$ ¿sigue siendo válida la afirmación del inciso anterior?

8. Sean $A, B \in M(n \times n)$ hermítianas.

Muestre que AB es hermítiana si y solo si $AB = BA$.

9. Muestre que $\dim \mathbb{C}^n = 2n$ si se considera $\mathbb{C}^n$ como un espacio vectorial sobre $\mathbb{R}$.

10. Sea $A \in M(n \times n)$ hermítiana. Muestre que $\langle A\vec{x}, \vec{x} \rangle \in \mathbb{R}$ para todo $\vec{x} \in \mathbb{C}^n$.

8.2. Matrices semejantes

Definición 8.17. Sean $A, B \in M(n \times n)$ matrices (reales o complejas). Se llaman *semejantes* o *similares* si existe una matriz invertible C tal que

$$A = C^{-1}BC. \quad (8.3)$$

En este caso escribimos $A \sim B$.

Ejercicio 8.18. Muestre que $A \sim B$ si y solo si existe una matriz invertible $\tilde{C}$ tal que

$$A = \tilde{C}B\tilde{C}^{-1}. \quad (8.4)$$

Question 8.2

Suponga que A y B son semejantes. ¿La matriz C en (8.3) es única o es posible que existan matrices invertibles diferentes $C_1 \neq C_2$ tales que $A = C_1^{-1}BC_1 = C_2^{-1}BC_2$?

Observación 8.19. La semejanza es una *relación de equivalencia* en el conjunto de todas las matrices cuadradas. Esto significa que satisface las siguientes tres propiedades. Sean $A_1, A_2, A_3 \in M(n \times n)$. Entonces:

- (I) *Reflexividad*: $A \sim A$ para todo $A \in M(n \times n)$.
- (II) *Simetría*: Si $A_1 \sim A_2$, entonces también $A_2 \sim A_1$.
- (III) *Transitividad*: Si $A_1 \sim A_2$ y $A_2 \sim A_3$, entonces también $A_1 \sim A_3$.

Demostración. (I) La reflexividad es clara. Solo necesitamos elegir $C = \text{id}$.

(II) Supongamos que $A_1 \sim A_2$. Entonces existe una matriz invertible C tal que $A_1 = C^{-1}A_2C$. Multiplicar por la izquierda por C y por la derecha por C^{-1} da $CA_1C^{-1} = A_2$. Sea $\tilde{C} = C^{-1}$. Entonces $\tilde{C}$ es invertible y $\tilde{C}^{-1} = C$. Por lo tanto obtenemos $\tilde{C}^{-1}A_1\tilde{C} = A_2$ lo cual muestra que $A_2 \sim A_1$.

(III) *Transitividad*: Si $A_1 \sim A_2$ y $A_2 \sim A_3$, entonces existen matrices invertibles C_1 y C_2 tales que $A_1 = C_1^{-1}A_2C_1$ and $A_2 = C_2^{-1}A_3C_2$. Se sigue que

$$A_1 = C_1^{-1}A_2C_1 = C_1^{-1}C_2^{-1}A_3C_2C_1 = (C_1C_2)^{-1}A_3C_1C_2.$$

Tomando $C = C_1C_2$ se muestra que $A_1 = C^{-1}A_3C$, por lo tanto $A_1 \sim A_3$. $\square$

Podemos interpretar $A \sim B$ como sigue: Sea C una matriz invertible con $A = C^{-1}BC$. Como C es una matriz invertible, sus columnas $\vec{c}_1, \dots, \vec{c}_n$ forman una base de $\mathbb{R}^n$ (o $\mathbb{C}^n$) y podemos ver C como la matriz de transición de la base canónica a la base $\vec{c}_1, \dots, \vec{c}_n$. Como B es la representación matricial de la transformación $\vec{x} \mapsto B\vec{x}$ con respecto a la base canónica de $\mathbb{R}^n$, la ecuación $A = C^{-1}BC$ dice que A representa la misma transformación lineal, pero con respecto a la base $\vec{c}_1, \dots, \vec{c}_n$.

Por otro lado, si A y B son representaciones matriciales de la misma transformación lineal, pero con respecto a bases posiblemente diferentes, entonces $A = C^{-1}BC$, donde C es la matriz de transición entre las dos bases. Por lo tanto A y B son semejantes.

Así hemos mostrado:

Dos matrices A y $B \in M(n \times n)$ son semejantes si y solo si representan la misma transformación lineal. La matriz C en $A = C^{-1}BC$ es la matriz de transición entre las dos bases usadas en las representaciones A y B .

Por lo tanto, el siguiente hecho no es muy sorprendente.

Proposición 8.20. Si $A, B \in M(n \times n)$ son semejantes, entonces $\det A = \det B$.

Demostración. Sea $C \in M(n \times n)$ invertible tal que $A = C^{-1}BC$. Entonces,

$$\det A = \det C^{-1}BC = \det(C^{-1}) \det B \det C = (\det C)^{-1} \det B \det C = \det B. \quad \square$$

Ejercicio 8.21. Muestre que $\det A = \det B$ no implica que A y B sean semejantes.

Ejercicio 8.22. Suponga que A y B son semejantes. Muestre que $\dim(\ker A) = \dim(\ker B)$ y que $\dim(\text{Im } A) = \dim(\text{Im } B)$. ¿Por qué esto no es sorprendente?

Question 8.3

Suponga que A y B son semejantes. ¿Cuál es la relación entre $\ker A$ y $\ker B$? ¿Cuál es la relación entre $\operatorname{Im} A$ y $\operatorname{Im} B$?

Hint. Theorem 6.18.

Una clase muy conveniente de matrices son las matrices diagonales, porque es bastante fácil calcular con ellas. Estrechamente relacionadas están las llamadas matrices diagonalizables.

Definición 8.23. Una matriz $A \in M(n \times n)$ se llama *diagonalizable* si es semejante a una matriz diagonal.

En otras palabras, una matriz A es diagonalizable si existe una matriz diagonal D y una matriz invertible C con

$$C^{-1}AC = D. \quad (8.5)$$

¿Cómo podemos decidir si una matriz A es diagonalizable? Sabemos que es diagonalizable si y solo si es semejante a una matriz diagonal, es decir, si y solo si existe una base $\vec{c}_1, \dots, \vec{c}_n$ tal que la representación de A con respecto a estos vectores es una matriz diagonal. En este caso, (8.5) se satisface si las columnas de C son los vectores de la base $\vec{c}_1, \dots, \vec{c}_n$.

Denotemos las entradas diagonales de D por $d_1, \dots, d_n$. Entonces es fácil ver que $D\vec{e}_j = d_j\vec{e}_j$. Esto significa que si aplicamos D a algún $\vec{e}_j$, entonces la imagen $D\vec{e}_j$ es paralela a $\vec{e}_j$. Como D no es otra cosa que la representación de A con respecto a la base $\vec{c}_1, \dots, \vec{c}_n$, tenemos $A\vec{c}_j = d_j\vec{c}_j$.

Podemos hacer esto más formal: Tome la ecuación (8.5) y multiplique ambos lados desde el lado izquierdo por C para obtener $AC = CD$. Recordemos que para cualquier matriz B , tenemos que $B\vec{e}_j = j$ -ésima columna de B . Obtenemos

$$\begin{aligned} AC\vec{e}_j &= A\vec{c}_j, \\ CD\vec{e}_j &= C(d_j\vec{e}_j) = d_jC(\vec{e}_j) = d_j\vec{c}_j. \end{aligned} \quad \xrightarrow{AC=CD} \quad A\vec{c}_j = d_j\vec{c}_j.$$

En resumen, encontramos:

Una matriz $A \in M(n \times n)$ es diagonalizable si y solo si podemos encontrar una base $\vec{c}_1, \dots, \vec{c}_n$ de $\mathbb{R}^n$ (o $\mathbb{C}^n$) y números $d_1, \dots, d_n$ tales que

$$A\vec{c}_j = d_j\vec{c}_j, \quad j = 1, \dots, n.$$

En este caso $C^{-1}AC = D$ (o equivalentemente $A = CDC^{-1}$), donde $D = \operatorname{diag}(d_1, \dots, d_n)$ y $C = (\vec{c}_1 | \dots | \vec{c}_n)$.

Los vectores $\vec{c}_j$ se llaman vectores propios de A y los números d_j se llaman valores propios de A . Se discutirán con más detalle en la siguiente sección, donde también veremos cómo calcularlos.

La diagonalización de una matriz es muy útil cuando queremos calcular potencias de la matriz.

Proposición 8.24. Sea $A \in M(n \times n)$ una matriz diagonalizable y sea C una matriz invertible y $D = \operatorname{diag}(d_1, \dots, d_n)$ tal que $A = CDC^{-1}$. Entonces $A^k = C \operatorname{diag}(d_1^k, \dots, d_n^k) C^{-1}$ para todo $k \in \mathbb{N}_0$. Si A es invertible, entonces todos los d_j son diferentes de 0 y la fórmula es cierta para todo $k \in \mathbb{Z}$.

Demostración. Sea $k \in \mathbb{N}_0$. Entonces

$$\begin{aligned}
 A^k &= (C \operatorname{diag}(d_1, \dots, d_n) C^{-1})^k \\
 &= C \operatorname{diag}(d_1, \dots, d_n) C^{-1} C \operatorname{diag}(d_1, \dots, d_n) C^{-1} \cdots C \operatorname{diag}(d_1, \dots, d_n) C^{-1} \\
 &= C \operatorname{diag}(d_1, \dots, d_n) \operatorname{diag}(d_1, \dots, d_n) \cdots \operatorname{diag}(d_1, \dots, d_n) C^{-1} \\
 &= C (\operatorname{diag}(d_1, \dots, d_n))^k C^{-1} \\
 &= C \operatorname{diag}(d_1^k, \dots, d_n^k) C^{-1}.
 \end{aligned}$$

Si todos los $d_j \neq 0$, entonces D es invertible con inversa

$$D^{-1} = (\operatorname{diag}(d_1, \dots, d_n))^{-1} = \operatorname{diag}(d_1^{-1}, \dots, d_n^{-1}).$$

Por lo tanto A es invertible y $A^{-1} = (C D C^{-1})^{-1} = C D^{-1} C^{-1}$ y obtenemos para $k \in \mathbb{Z}$ con $k < 0$

$$\begin{aligned}
 A^k &= A^{-|k|} = (A^{-1})^{|k|} = (C D^{-1} C^{-1})^{|k|} = C (D^{-1})^{|k|} C^{-1} = C D^{-|k|} C^{-1} \\
 &= C \operatorname{diag}(d_1^k, \dots, d_n^k) C^{-1}.
 \end{aligned}$$

□

La Proposición 8.24 es útil, por ejemplo, cuando describimos sistemas dinámicos mediante matrices o cuando resolvemos ecuaciones diferenciales lineales con coeficientes constantes en dimensiones superiores.

Ahora debería haber entendido

- que matrices semejantes representan la misma transformación lineal,
- por qué matrices semejantes tienen el mismo determinante,
- por qué una matriz es diagonalizable si y solo si $\mathbb{R}^n$ (o $\mathbb{C}^n$) admite una base formada por vectores propios de A ,
- etc.

Ahora debería ser capaz de

- etc.

Ejercicios.

1. Sean $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 2 & -1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 12 & 110 & 7 \\ 5 & 16 & -6 \\ 6 & 57 & 4 \end{pmatrix}$ y $C = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 3 & 1 \\ 1 & 4 & -1 \end{pmatrix}$. Verifique que $AC = CB$ y concluya que A, B son matrices semejantes.
2. Encuentre tres matrices que son semejantes a la matriz A del Ejercicio 1.. Para cada una de ellas, encuentre el determinante y la traza.
3. De las siguientes afirmaciones, diga en cada una si es verdadera ó falsa. Si es verdadera justifique por qué, si es falsa encuentre un contraejemplo.

- (a) Sean $A, B \in M(n \times n)$ tales que $\det A = \det B$, entonces A, B son matrices semejantes.
- (b) Sean $D_1, D_2 \in M(n \times n)$ matrices diagonales tales que $D_1 \neq D_2$, entonces D_1, D_2 no son matrices semejantes.
- (c) Si $A, B \in M(n \times n)$ son matrices equivalentes por filas entonces A, B son matrices semejantes.
4. (a) Sean $\lambda_1, \lambda_2 \in \mathbb{C}$ y considere

$$D_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad D_2 = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix}.$$

Muestre que D_1, D_2 son semejantes. ¿Como puede generalizar este resultado a matrices D_1, D_2 diagonales de cualquier tamaño?

- (b) Muestre que dos matrices diagonales $D_1, D_2 \in M(n \times n)$ son semejantes si, salvo el orden, tienen exactamente los mismos valores en la diagonal.
5. Defina la siguiente función

$$\text{tr} : M(n \times n) \rightarrow \mathbb{R}, \quad \text{tr } A := a_{11} + a_{22} + \cdots + a_{nn}. \quad (*)$$

Note que para una matriz $A \in M(n \times n)$ el número $\text{tr } A$ es la suma de los elementos de la diagonal de A : este número se llama *traza de A*, ver definición 8.35.

- (a) Muestre que tr definida en $(*)$ es una transformación lineal.
- (b) Sean $A, B \in M(n \times n)$. Muestre que $\text{tr } AB = \text{tr } BA$.
- (c) Sean $A, B \in M(n \times n)$ matrices semejantes. Muestre que $\text{tr } A = \text{tr } B$.
- (d) Sea $A \in M(n \times n)$ tal que $\text{tr}(A^t A) = 0$. Muestre que $A = \mathbb{O}$.
6. Encuentre matrices A y B con $\text{tr } A = \text{tr } B$ que no son semejantes.
7. Encuentre matrices A y B con $\det A = \det B$ que no son semejantes.

8.3. Valores y vectores propios

Definición 8.25. Sea V un espacio vectorial y sea $T : V \rightarrow V$ una transformación lineal. Un número λ se llama *valor propio* de T si existe un vector $\vec{v} \neq \mathbb{O}$ tal que

$$T\vec{v} = \lambda\vec{v}. \quad (8.6)$$

El vector $\vec{v}$ se llama entonces un *vector propio*.

La razón por la cual excluimos $\vec{v} = \mathbb{O}$ en la definición anterior es que para todo λ se cumple que $T\mathbb{O} = \mathbb{O} = \lambda\mathbb{O}$, de modo que (8.6) se satisfaría para cualquier λ si se nos permitiera elegir $\vec{v} = \mathbb{O}$, en cuyo caso la definición no tendría mucho sentido.

Ejercicio 8.26. Muestre que 0 es un valor propio de T si y solo si $\dim(\ker T) \geq 1$, es decir, si y solo si T no es invertible. Muestre que v es un vector propio con valor propio 0 si y solo si $v \in \ker T \setminus \{\mathbb{O}\}$.

Ejercicio 8.27. Muestre que todos los valores propios de una matriz unitaria tienen norma 1.

Question 8.4

Sean V, W espacios vectoriales y sea $T : V \rightarrow W$ una transformación lineal. ¿Por qué **no** tiene sentido hablar de valores propios de T si $V \neq W$?

Enumeremos algunas propiedades de los vectores propios que son fáciles de ver.

- (I) Un vector v es un vector propio de T si y solo si $Tv \parallel v$.
- (II) Si v es un vector propio de T con valor propio $\lambda \neq 0$, entonces $v \in \operatorname{Im} T$ porque $v = \frac{1}{\lambda}Tv$.
- (III) Si v es un vector propio de T con valor propio λ , entonces todo múltiplo no nulo de v es un vector propio con el mismo valor propio porque

$$T(cv) = cTv = c\lambda v = \lambda(cv).$$

- (IV) Podemos generalizar (III) como sigue: Si $v_1, \dots, v_k$ son vectores propios de T con el mismo valor propio λ , entonces toda combinación lineal no nula es un vector propio con el mismo valor propio porque

$$T(\alpha_1 v_1 + \dots + \alpha_k v_k) = \alpha_1 T v_1 + \dots + \alpha_k T v_k = \alpha_1 \lambda v_1 + \dots + \alpha_k \lambda v_k = \lambda(\alpha_1 v_1 + \dots + \alpha_k v_k).$$

La propiedad (IV) dice que el conjunto de todos los vectores propios con el mismo valor propio es casi un subespacio. Lo único que falta es el vector cero $\mathbb{O}$. Esto motiva la siguiente definición.

Definición 8.28. Sea V un espacio vectorial y sea $T : V \rightarrow V$ una transformación lineal con valor propio λ . Entonces el *espacio propio* de T correspondiente a λ es

$$\begin{aligned} \operatorname{Eig}_\lambda(T) &:= \operatorname{Eig}(T, \lambda) := \{v \in V : v \text{ es vector propio de } T \text{ con valor propio } \lambda\} \cup \{\mathbb{O}\} \\ &= \{v \in V : Tv = \lambda v\}. \end{aligned}$$

La dimensión de $\operatorname{Eig}_\lambda(T)$ se llama la *multiplicidad geométrica* of λ .

Proposición 8.29. Sea $T : V \rightarrow V$ una transformación lineal y sea λ un valor propio de T . Entonces

$$\operatorname{Eig}_\lambda(T) = \ker(T - \lambda \operatorname{id}).$$

Demostración. Sea $v \in V$. Entonces

$$\begin{aligned} v \in \operatorname{Eig}_\lambda(T) &\iff Tv = \lambda v \iff Tv - \lambda v = \mathbb{O} \iff Tv - \lambda \operatorname{id} v = \mathbb{O} \\ &\iff (T - \lambda \operatorname{id})v = \mathbb{O} \iff v \in \ker(T - \lambda \operatorname{id}). \end{aligned} \quad \square$$

Observe que la Proposición 8.29 muestra nuevamente que $\text{Eig}_\lambda(T)$ es un subespacio de V . Además muestra que λ es un valor propio de T si y solo si $T - \lambda \text{id}$ no es invertible. Para el caso especial $\lambda = 0$ tenemos que $\text{Eig}_0(T) = \ker T$.

Ejemplos 8.30. (a) Sea V un espacio vectorial y sea $T = \text{id}$. Entonces para todo $v \in V$ tenemos que $Tv = v = 1v$. Por lo tanto T tiene un único valor propio, a saber $\lambda = 1$, y $\text{Eig}_1(T) = \ker(T - \text{id}) = \ker 0 = V$. Su multiplicidad geométrica es $\dim(\text{Eig}_1(T)) = \dim V$.

(b) Sea $V = \mathbb{R}^2$ y sea R la reflexión respecto al eje x . Si $\vec{v}$ es un vector propio de R , entonces $R\vec{v}$ debe ser paralelo a $\vec{v}$. Esto ocurre si y solo si $\vec{v}$ es paralelo al eje x , en cuyo caso $R\vec{v} = \vec{v}$, o si $\vec{v}$ es perpendicular al eje x , en cuyo caso $R\vec{v} = -\vec{v}$. Todos los demás vectores cambian de dirección bajo una reflexión. Por lo tanto tenemos los valores propios $\lambda_1 = 1$ and $\lambda_2 = -1$ and $\text{Eig}_1(R) = \text{span}\{\vec{e}_1\}$, $\text{Eig}_{-1}(R) = \text{span}\{\vec{e}_2\}$. Cada valor propio tiene multiplicidad geométrica 1.

Observe que la representación matricial de R con respecto a la base canónica de $\mathbb{R}^2$ es $A_R = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $A_R \vec{x} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$. Por lo tanto $A_R \vec{x}$ es paralelo a $\vec{x}$ si y solo si $x_1 = 0$ (en cuyo caso $\vec{x} \in \text{span}\{\vec{e}_2\}$) or $x_2 = 0$ (en cuyo caso $\vec{x} \in \text{span}\{\vec{e}_1\}$).

(c) Sea $V = \mathbb{R}^2$ y sea R la rotación de 90° . Entonces claramente $R\vec{v} \not\parallel \vec{v}$ para todo $\vec{v} \in \mathbb{R}^2 \setminus \{\vec{0}\}$. Por lo tanto R no tiene valores propios.

Observe que la representación matricial de R con respecto a la base canónica de $\mathbb{R}^2$ es $A_R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Si consideramos A_R como una matriz real, entonces no tiene valores propios. Sin embargo, si consideramos A_R como una matriz compleja, entonces tiene los valores propios $\pm i$, como veremos más adelante.

(d) Let $A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. Como siempre, identificamos A con la transformación lineal $\mathbb{R}^6 \rightarrow \mathbb{R}^6$, $\vec{x} \mapsto A\vec{x}$. No es difícil ver que los valores propios y los espacios propios de A son

$$\begin{aligned} \lambda_1 &= 1, & \text{Eig}_1(A) &= \text{span}\{\vec{e}_1\}, & \text{multiplicidad geom.: } 1, \\ \lambda_2 &= 5, & \text{Eig}_5(A) &= \text{span}\{\vec{e}_2, \vec{e}_3, \vec{e}_4\}, & \text{multiplicidad geom.: } 3, \\ \lambda_3 &= 8, & \text{Eig}_8(A) &= \text{span}\{\vec{e}_6, \vec{e}_7\}. & \text{multiplicidad geom.: } 2. \end{aligned}$$

Muestre las afirmaciones anteriores.

(e) Sea $V = C^\infty(\mathbb{R})$ el espacio de todas las funciones infinitamente diferenciables de $\mathbb{R}$ en $\mathbb{R}$ y sea $T : V \rightarrow V$, $Tf = f'$. Análogamente al Ejemplo 6.4 podemos mostrar que T es una transformación lineal. Los valores propios de T son aquellos $\lambda \in \mathbb{R}$ tales que existe una función $f \in C^\infty(\mathbb{R})$ con $f' = \lambda f$. Sabemos que para todo $\lambda \in \mathbb{R}$ esta ecuación diferencial tiene una solución y que toda solución es de la forma $f_\lambda(x) = ce^{\lambda x}$ para algún número real c . Por lo tanto, todo $\lambda \in \mathbb{R}$ es un valor propio de T con espacio propio $\text{Eig}_\lambda(T) = \text{span}\{g_\lambda\}$ donde g_λ es la función dada por $g_\lambda(x) = e^{\lambda x}$. En particular, la multiplicidad geométrica de cualquier $\lambda \in \mathbb{R}$ es 1.

- (f) Sea $V = C^\infty(\mathbb{R})$ el espacio de todas las funciones infinitamente diferenciables de $\mathbb{R}$ en $\mathbb{R}$ y sea $T : V \rightarrow V$, $Tf = f''$. Es fácil ver que T es una transformación lineal. Los valores propios de T son aquellos $\lambda \in \mathbb{R}$ tales que existe una función $f \in C^\infty(\mathbb{R})$ con $f'' = \lambda f$. Si $\lambda > 0$, entonces la solución general de esta ecuación diferencial es $f_\lambda(x) = a e^{\sqrt{\lambda}x} + b e^{-\sqrt{\lambda}x}$. Si $\lambda < 0$, la solución general es $f_\lambda(x) = a \cos \sqrt{\lambda}x + b \sin \sqrt{\lambda}x$. Si $\lambda = 0$, la solución general es $f_0(x) = ax + b$. Por lo tanto, todo $\lambda \in \mathbb{R}$ es un valor propio de T con multiplicidad geométrica 2.

Escriba los espacios propios para un λ dado.

Encuentre los valores propios y los espacios propios si consideramos el espacio vectorial de funciones infinitamente diferenciables de $\mathbb{R}$ en $\mathbb{C}$.

En los ejemplos anteriores era relativamente fácil adivinar los valores propios. Pero, ¿cómo calculamos los valores propios de, por ejemplo, $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ o de la transformación lineal $T : M(n \times n) \rightarrow M(n \times n)$, $T(A) = A + A^t$?

Como cualquier transformación lineal en un espacio vectorial de dimensión finita V puede “traducirse” a una matriz eligiendo una base de V , es suficiente encontrar valores propios de matrices, como muestra el siguiente teorema.

Teorema 8.31. *Sea V un espacio vectorial de dimensión finita con base $\mathcal{B} = \{v_1, \dots, v_n\}$ y sea $T : V \rightarrow V$ una transformación lineal. Si A_T es la representación matricial de T con respecto a la base $\mathcal{B}$, entonces los valores propios de T y A_T coinciden y un vector $v = c_1 v_1 + \dots + c_n v_n$ es un vector propio de T con valor propio λ si y solo si $\vec{x} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ es un vector propio de A_T con el mismo valor propio λ . En particular, las dimensiones de los espacios propios de T y de A_T coinciden.*

Demostración. Sea $\mathbb{K} = \mathbb{R}$ si V es un espacio vectorial real y $\mathbb{K} = \mathbb{C}$ si V es un espacio vectorial complejo y sea $\Phi : V \rightarrow \mathbb{K}^n$ la transformación lineal definida por $\Phi(v_j) = \vec{e}_j$, ($j = 1, \dots, n$).

Esto significa que Φ “traduce” un vector $v = c_1 v_1 + \dots + c_n v_n$ al vector columna $\vec{x} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$, cf. Section 6.4.

$$\begin{array}{ccc} V & \xrightarrow{T} & V \\ \Phi \downarrow & & \downarrow \Phi \\ \mathbb{K}^n & \xrightarrow{A_T} & \mathbb{K}^n \end{array}$$

Recordemos que $T = \Phi^{-1} A_T \Phi$. Sea λ un valor propio de T con vector propio v , es decir, $Tv = \lambda v$. Expresamos v como combinación lineal de los vectores de la base $\mathcal{B}$ como $v = c_1 v_1 + \dots + c_n v_n$. Por lo tanto,

$$Tv = \lambda v \iff \Phi^{-1} A_T \Phi v = \lambda v \iff A_T \Phi v = \Phi \lambda v \iff A_T(\Phi v) = \lambda(\Phi v)$$

lo cual ocurre si y solo si λ es un valor propio de A_T y $\Phi v \in \text{Eig}_\lambda(A_T)$. □

La demostración muestra que $\text{Eig}_\lambda(A_T) = \Phi(\text{Eig}_\lambda(T))$ como era de esperarse.

Corolario 8.32. *Supongamos que A y B son matrices semejantes y sea C una matriz invertible con $A = C^{-1}BC$. Entonces A y B tienen los mismos valores propios y para todo valor propio λ tenemos que $\text{Eig}_\lambda(B) = C \text{Eig}_\lambda(A)$.*

Ahora volvamos a la pregunta de cómo calcular los valores propios y vectores propios de una matriz dada A . Recordemos que λ es un valor propio de A si y solo si $\ker(A - \lambda \text{id}) \neq \{\vec{0}\}$, vea Proposition 8.29. Como $A - \lambda \text{id}$ es una matriz cuadrada, esto ocurre si y solo si $\det(A - \lambda \text{id}) = 0$.

Definición 8.33. La función $\lambda \mapsto \det(A - \lambda \text{id})$ se llama el *polinomio característico* de A . Usualmente se denota por p_A .

Antes de discutir el polinomio característico y mostrar que efectivamente es un polinomio, describiremos cómo encontrar los valores propios y vectores propios de una matriz cuadrada dada A .

Procedimiento para encontrar los valores propios y vectores propios de una matriz cuadrada dada A .

- Calcule el polinomio característico $p_A(\lambda) := \det(A - \lambda \text{id})$.
- Encuentre los ceros $\lambda_1, \dots, \lambda_k$ del polinomio característico. Estos son los **valores propios** de A .
- Para cada valor propio λ_j calcule $\ker(A - \lambda_j)$, por ejemplo usando eliminación de Gauß-Jordan. Esto da los **espacios propios**.

Ejemplo 8.34. Encuentre los valores propios y los espacios propios de $A = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$.

Solución. ■ El polinomio característico de A es

$$p_A(\lambda) = \det(A - \lambda \text{id}) = \det \begin{pmatrix} 2-\lambda & 1 \\ 3 & 4-\lambda \end{pmatrix} = (2-\lambda)(4-\lambda) - 3 = \lambda^2 - 6\lambda + 5.$$

- Ahora podemos completar el cuadrado o usar la fórmula de solución para ecuaciones cuadráticas para encontrar los ceros de p_A . Aquí elegimos completar el cuadrado.

$$p_A(\lambda) = \lambda^2 - 6\lambda + 5 = (\lambda - 3)^2 - 4 = (\lambda - 5)(\lambda - 1).$$

Por lo tanto, los **valores propios** de A son $\lambda_1 = 5$ y $\lambda_2 = 1$.

- Ahora calculamos los espacios propios usando eliminación de Gauß.

$$* \quad A - 5 \text{id} = \begin{pmatrix} 2-5 & 1 \\ 3 & 4-5 \end{pmatrix} = \begin{pmatrix} -3 & 1 \\ 3 & -1 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 + R_1} \begin{pmatrix} -3 & 1 \\ 0 & 0 \end{pmatrix} \xrightarrow{R_1 \rightarrow -R_1} \begin{pmatrix} 3 & -1 \\ 0 & 0 \end{pmatrix}.$$

$$\text{Por lo tanto, } \ker(A - 5 \text{id}) = \text{span} \left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\}.$$

$$* A - 1 \text{ id} = \begin{pmatrix} 2-1 & 1 \\ 3 & 4-1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 - 3R_1} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

$$\text{Por lo tanto, } \ker(A - 1 \text{ id}) = \text{span} \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

En resumen, tenemos dos valores propios,

$$\lambda_1 = 5, \quad \text{Eig}_5(A) = \text{span} \left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\}, \quad \text{multiplicidad geom.: } 1,$$

$$\lambda_2 = 1, \quad \text{Eig}_1(A) = \text{span} \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}, \quad \text{multiplicidad geom.: } 1.$$

◇

Si ponemos $\vec{v}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\vec{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ podemos verificar nuestro resultado calculando

$$A\vec{v}_1 = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 15 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 3 \end{pmatrix} = 5\vec{v}_1,$$

$$A\vec{v}_2 = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \vec{v}_2.$$

Antes de dar más ejemplos, mostramos que el polinomio característico es efectivamente un polinomio. Primero necesitamos una definición.

Definición 8.35. Sea $A = (a_{ij})_{i,j=1}^n \in M(n \times n)$. La *traza de A* es la suma de sus entradas en la diagonal:

$$\text{tr } A := a_{11} + a_{22} + \dots + a_{nn}.$$

Observación. Observe que el ejercicio 5. de la sección 8.2 muestra que la traza de matrices semejantes coincide, entonces si V es un espacio de dimensión finita y si T es una transformación lineal, tiene sentido definir $\text{tr } T$ como la traza de la representación matricial de T en cualquier base de V .

Teorema 8.36. Sea $A = (a_{ij})_{i,j=1}^n \in M(n \times n)$ y sea $p_A(\lambda) = \det(A - \lambda \text{ id})$ el polinomio característico de A . Entonces se cumple lo siguiente.

(i) p_A es un polinomio de grado n .

(i) Sea $p_A(\lambda) = c_n \lambda^n + c_{n-1} \lambda^{n-1} + \dots + c_1 \lambda + c_0$. Entonces tenemos fórmulas para los coeficientes c_n, c_{n-1} and c_0 :

$$c_n = (-1)^n, \quad c_{n-1} = (-1)^{n-1} \text{tr } A, \quad c_0 = \det A.$$

Demostración. Por definición,

$$p_A(\lambda) = \det(A - \lambda \text{ id}) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix}.$$

De acuerdo con la Observación 4.4, el determinante es la suma de productos donde cada producto consiste en un signo y n factores elegidos de modo que contiene una entrada de cada fila y de cada columna de $A - \lambda \text{id}$. Por lo tanto es claro que p_A es un polinomio en λ . El término con más factores λ es el de la forma

$$(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda). \quad (8.7)$$

Todos los demás términos contienen a lo sumo $n - 2$ factores con λ . Para ver esto, supongamos por ejemplo que en uno de los términos el factor de la primera fila no es $(a_{11} - \lambda)$ but some a_{1j} . Entonces no puede haber otro factor de la j -ésima columna, en particular el factor $(a_{jj} - \lambda)$ no puede aparecer. Así que este término ya tiene dos factores sin λ ; por lo tanto, el grado del término como polinomio en λ puede ser a lo sumo $n - 2$. Esto muestra que

$$p_A(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) + \text{términos de grado a lo más } n - 2. \quad (8.8)$$

Si expandimos el primer término y ordenamos por potencias de λ , obtenemos

$$(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) = (-1)^n \lambda^n + (-1)^{n-1} \lambda^{n-1} (a_{11} + \cdots + a_{nn}) \\ + \text{términos de grado a lo más } n - 2.$$

Insertando esto en (8.7), encontramos que

$$p_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \lambda^{n-1} (a_{11} + \cdots + a_{nn}) + \text{términos de grado a lo más } n - 2, \quad (8.9)$$

por ende $\deg(p_A) = n$.

La fórmula (8.9) también muestra la afirmación sobre c_n and c_{n-1} . La fórmula para c_0 se sigue de

$$c_0 = p_A(0) = \det(A - 0 \text{id}) = \det A. \quad \square$$

Obtenemos inmediatamente el siguiente corolario muy importante.

Corolario 8.37. *Una matriz $n \times n$ puede tener a lo sumo n valores propios diferentes.*

Demostración. Sea $A \in M(n \times n)$. Entonces los valores propios de A son exactamente los ceros de su polinomio característico. Como tiene grado n , puede tener a lo sumo n ceros. $\square$

Ahora entendemos por qué trabajar con espacios vectoriales complejos es más adecuado cuando nos interesan los valores propios. Son precisamente los ceros del polinomio característico. Aunque un polinomio puede no tener ceros reales, siempre tiene ceros si permitimos que sean números complejos. De hecho, cualquier polinomio siempre puede factorizarse sobre $\mathbb{C}$.

Sea $A \in M(n \times n)$ y sea p_A su polinomio característico. Entonces existen números complejos $\lambda_1, \dots, \lambda_k$ y enteros $m_1, \dots, m_k \geq 1$ tales que

$$p_A(\lambda) = (\lambda_1 - \lambda)^{m_1} \cdot (\lambda_2 - \lambda)^{m_2} \cdots (\lambda_k - \lambda)^{m_k}.$$

Los números $\lambda_1, \dots, \lambda_k$ son precisamente los valores propios complejos de A and $m_1 + \cdots + m_k = \deg p_A = n$.

Definición 8.38. El número entero m_j se llama la *multiplicidad algebraica* del valor propio λ_j .

El siguiente teorema es muy importante, pero omitimos su demostración.

Teorema 8.39. Sea $A \in M(n \times n)$ y sea λ un valor propio de A . Entonces

multiplicidad geométrica de $\lambda \leq$ multiplicidad algebraica de λ .

Ejemplo 8.40. Let $A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 0 & 0 \\ 0 & 0 & 5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 8 \end{pmatrix}$. Como $A - \lambda \text{id}$ es una matriz triangular superior,

su determinante es el producto de las entradas en la diagonal. Obtenemos

$$p_A(\lambda) = \det(A - \lambda \text{id}) = (1 - \lambda)(5 - \lambda)^3(8 - \lambda)^2.$$

Por lo tanto, los valores propios de A son $\lambda_1 = 1$, $\lambda_2 = 5$, $\lambda_3 = 8$. Calculemos los espacios propios.

■ $A - 1 \text{id} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 7 \end{pmatrix} \xrightarrow{\text{permutar filas}} \begin{pmatrix} 0 & 4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. Esta matriz está en forma escalonada por filas y podemos ver fácilmente que $\text{Eig}_1(A) = \ker(A - 1 \text{id}) = \text{span}\{\vec{e}_1\}$ que tiene dimensión 1.

■ $A - 5 \text{id} = \begin{pmatrix} -4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \xrightarrow{\text{permutar filas}} \begin{pmatrix} -4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. Esta matriz está en forma escalonada por filas y podemos ver fácilmente que $\text{Eig}_5(A) = \ker(A - 5 \text{id}) = \text{span}\{\vec{e}_2\}$ que tiene dimensión 1.

■ $A - 8 \text{id} = \begin{pmatrix} -7 & 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & 1 & 0 & 0 & 0 \\ 0 & 0 & -3 & 1 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. Esta matriz está en forma escalonada por filas y podemos ver fácilmente que $\text{Eig}_8(A) = \ker(A - 8 \text{id}) = \text{span}\{\vec{e}_5, \vec{e}_6\}$ que tiene dimensión 2.

En resumen, tenemos

$$\begin{array}{llll} \lambda_1 = 1, & \text{Eig}_1(A) = \text{span}\{\vec{e}_1\}, & \text{multiplicidad geom.: } 1, & \text{multiplicidad alg.: } 1, \\ \lambda_2 = 5, & \text{Eig}_5(A) = \text{span}\{\vec{e}_2\}, & \text{multiplicidad geom.: } 1, & \text{multiplicidad alg.: } 3, \\ \lambda_3 = 8, & \text{Eig}_8(A) = \text{span}\{\vec{e}_5, \vec{e}_6\}, & \text{multiplicidad geom.: } 2, & \text{multiplicidad alg.: } 2. \end{array}$$

Ejemplo 8.41. Encuentre los valores propios complejos y los espacios propios de $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Solución. Del Ejemplo 8.30 ya sabemos que R no tiene valores propios reales. El polinomio característico de R es

$$p_R(\lambda) = \det(R - \lambda) = \det \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 + 1 = (\lambda - i)(\lambda + i).$$

Por lo tanto, los valores propios son $\lambda_1 = -i$ y $\lambda_2 = i$. Calculemos los espacios propios.

- $R - (-i) \text{id} = \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 + iR_1} \begin{pmatrix} i & -1 \\ 0 & 0 \end{pmatrix}$. Por lo tanto, $\text{Eig}_{-i}(R) = \ker(R + i \text{id}) = \text{span} \left\{ \begin{pmatrix} 1 \\ i \end{pmatrix} \right\}$.
- $R - i \text{id} = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 + iR_1} \begin{pmatrix} -i & -1 \\ 0 & 0 \end{pmatrix}$. Por lo tanto, $\text{Eig}_i(R) = \ker(R - i \text{id}) = \text{span} \left\{ \begin{pmatrix} 1 \\ -i \end{pmatrix} \right\}$.

◇

Ejemplo 8.42. Encuentre la diagonalización de $A = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$.

Solución. Necesitamos encontrar una matriz invertible C y una matriz diagonal D tales que $D = C^{-1}AC$. Por el Ejemplo 8.34, A tiene los valores propios $\lambda_1 = 5$ y $\lambda_2 = 1$, por lo tanto A es efectivamente diagonalizable. Sabemos que las entradas diagonales de D son los valores propios de A , por lo tanto $D = \text{diag}(5, 1)$ y las columnas de C son los vectores propios correspondientes $\vec{v}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\vec{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, por lo tanto

$$D = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix} \quad \text{and} \quad D = C^{-1}AC.$$

Alternativamente, podríamos haber elegido $\tilde{D} = \text{diag}(1, 5)$. Entonces la $\tilde{C}$ correspondiente es $\tilde{C} = (\vec{v}_2 | \vec{v}_1)$ porque la j -ésima columna de la matriz invertible debe ser un vector propio correspondiente a la j -ésima entrada de la matriz diagonal, mostrando que

$$\tilde{D} = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}, \quad \tilde{C} = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix} \quad \text{and} \quad \tilde{D} = \tilde{C}^{-1}A\tilde{C}. \quad \diamond$$

Observe que, salvo el orden de los elementos diagonales, la matriz D está determinada de manera única por A . Sin embargo, para la matriz C tenemos más opciones. Por ejemplo, si multiplicamos cada columna de C por una constante arbitraria diferente de 0, todavía funciona.

Ejemplo 8.43. Sea $V = M(2 \times 2)$ y sea $T : V \rightarrow V$, $T(M) = M + M^t$. Encuentre los valores propios y los espacios propios de T .

Solución. Let $M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $M_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $M_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $M_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Entonces $\mathcal{B} = \{M_1, M_2, M_3, M_4\}$ es una base de $M(2 \times 2)$. La representación matricial de T con respecto a ella es

$$A_T = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

El polinomio característico es

$$\begin{aligned}\det(A_T - \lambda \text{id}) &= \det \begin{pmatrix} 2-\lambda & 0 & 0 & 0 \\ 0 & 1-\lambda & 1 & 0 \\ 0 & 1 & 1-\lambda & 0 \\ 0 & 0 & 0 & 2-\lambda \end{pmatrix} = (2-\lambda) \det \begin{pmatrix} 1-\lambda & 1 & 0 \\ 1 & 1-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix} \\ &= (2-\lambda)^2 [(1-\lambda)^2 - 1] = \lambda(\lambda-2)^3.\end{aligned}$$

Por lo tanto hay dos valores propios: $\lambda_1 = 0$ y $\lambda_2 = 2$.

Encontremos los espacios propios.

$$\begin{aligned}\blacksquare A_T - 0 \text{id} &= A_T \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \xrightarrow[\substack{R_1 \rightarrow \frac{1}{2}R_1 \\ R_4 \rightarrow \frac{1}{2}R_4}]{R_3 \rightarrow R_3 - R_2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_3 \leftrightarrow R_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ \blacksquare A_T - 2 \text{id} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 + R_3} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.\end{aligned}$$

$$\text{Por lo tanto, } \text{Eig}_0(A_T) = \text{span} \left\{ \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \right\} \text{ and } \text{Eig}_2(A_T) = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Esto significa que los valores propios de T son 0 y 2 y que los espacios propios son $\text{Eig}_0(T) = \text{span} \{M_2 - M_3\} = \text{span} \left\{ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\}$ and

$$\text{Eig}_0(T) = \text{span} \{M_2 - M_3\} = \text{span} \left\{ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\} = M_{\text{asym}}(2 \times 2),$$

$$\text{Eig}_2(T) = \text{span} \{M_1, M_2 + M_3, M_4\} = \text{span} \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} = M_{\text{sym}}(2 \times 2),$$

Observación. Podríamos haber calculado los espacios propios de T directamente, sin calcular primero los de A_T , como sigue.

- Una matriz M pertenece a $\text{Eig}_0(T)$ si y solo si $T(M) = 0$. Este es el caso si y solo si $M + M^t = 0$, lo cual significa que $M = -M^t$. Así, $\text{Eig}_0(T)$ es el espacio de todas las matrices antisimétricas 2×2 .
- Una matriz M pertenece a $\text{Eig}_2(T)$ si y solo si $T(M) = 2M$. Esto significa que $M + M^t = 2M$. Este es el caso si y solo si $M = M^t$. Así, $\text{Eig}_0(T)$ es el espacio de todas las matrices simétricas 2×2 .

◇

Ahora debería haber entendido

- el concepto de valores propios y vectores propios,
- por qué una matriz $n \times n$ puede tener a lo sumo n valores propios,
- por qué la restricción de A a cualquiera de sus espacios propios actúa como un múltiplo de la identidad,
- qué dice el polinomio característico de una matriz sobre sus valores propios,
- por qué una matriz $n \times n$ es diagonalizable si y solo si $\mathbb{K}^n$ tiene una base formada por vectores propios de A ,
- etc.

Ahora debería ser capaz de

- calcular el polinomio característico de una matriz cuadrada A ,
- calcular los valores propios y vectores propios de una matriz cuadrada A ,
- diagonalizar una matriz diagonalizable,
- etc.

Ejercicios.

1. Para las siguientes matrices, encuentre los vectores propios, los espacios propios, una matriz invertible C y una matriz diagonal D tal que $C^{-1}AC = D$.

$$A_1 = \begin{pmatrix} -3 & 5 & -20 \\ 2 & 0 & 8 \\ 2 & 1 & 7 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -2 & 0 & 1 \\ 0 & 2 & 0 \\ 9 & 0 & 6 \end{pmatrix}, \quad A_3 = \begin{pmatrix} -2 & 0 & -1 \\ 0 & 2 & 0 \\ 9 & 0 & 6 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \\ 1 & 3 & 2 \end{pmatrix}.$$

2. Sea $D : P_3 \rightarrow P_3$ dada por $Dp = p'$. Encuentre el polinomio característico y los valores propios de D .
3. Sea $D : P_3 \rightarrow P_3$ dada por $Dp = p + xp' + p''$. Encuentre el polinomio característico y los valores propios de D .
4. Sea

$$A = \begin{pmatrix} 3 & -1 \\ -2 & 4 \end{pmatrix}.$$

- (a) Calcule $(A^{-1})^n$ para cualquier $n \in \mathbb{N}$.
- (b) ¿Que se puede decir de $(A^{-1})^n$ cuando $n \rightarrow \infty$?
5. Sea $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por:

$$T(\vec{x}) = \text{reflexión de } \vec{x} \text{ con respecto a la recta } y = x.$$

Calcule los valores propios de T y muestre que T es diagonalizable. (Hint: basta escoger una base adecuada de $\mathbb{R}^2$).

6. Sea $V = C[0, 1]$ y $T : V \rightarrow V$ dada por

$$(Tf)(x) = \int_0^x f(t) \, dt.$$

Muestre que T no tiene valores propios.

7. Sea $A \in M(n \times n)$. Muestre que:

(a) A es diagonalizable si y solo si A^{-1} es diagonalizable.

(b) A es diagonalizable si y solo si A^t es diagonalizable.

8. Sea $A \in M(n \times n)$ invertible y λ un valor propio de A . Note que $\lambda \neq 0$. Muestre que $\frac{1}{\lambda}$ es un valor propio de A^{-1} . (Esto dice que si A tiene valores propios $\mu_1, \dots, \mu_k$ entonces los valores propios de A^{-1} son $\frac{1}{\mu_1}, \dots, \frac{1}{\mu_k}$)

9. Sea $A \in M(n \times n)$. Muestre que A y A^t tienen los mismos valores propios. (Hint: analice el polinomio característico)

10. Sea $\vec{v} \in \mathbb{R}^3$ un vector no nulo y $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por $T(\vec{x}) = \vec{v} \times \vec{x}$. Muestre que 0 es el único valor propio real.

11. Sea W un subespacio de $\mathbb{R}^n$ con $\dim W = m$ y P_W la proyección ortogonal de $\mathbb{R}^n$ sobre W . ¿Cuáles son los autovalores de P_W ? ¿Cuál es el polinomio característico de P_W ? ¿Es P_W diagonalizable? (Hint: Empiece escogiendo una base ortonormal de W).

12. Sea $A \in M(n \times n)$ tal que $A^m = \text{id}_n$ para algún $m \in \mathbb{N}$.

(a) Muestre que si $\lambda \in \mathbb{C}$ es un valor propio de A entonces $\lambda^m = 1$.

(b) Encuentre cuatro matrices distintas A tales que $A^3 = \text{id}_3$.

13. Sea $A \in M(n \times n)$ tal que $A^m = \mathbb{O}$ para algún $m \in \mathbb{N}$.

(a) Muestre que $\lambda \text{id}_n - A$ es invertible para todo $\lambda \in \mathbb{C} - \{0\}$. (Hint: La prueba es la misma del ejercicio 28. sección 3.5)

(b) Encuentre el polinomio característico de A . Observe que en ejercicio 2., $D^4 = \mathbb{O}$.

14. Sea $A \in M(n \times n)$ una matriz tal que $A^2 = A$, muestre que A es diagonalizable. (Hint: Por el ejercicio 7. de la sección 6.2; elija una base de $\ker A$ y complétela a una base de $\mathbb{R}^n$ por medio de una base de $\text{Im } A$)

15. Sea $A \in M(n \times n)$ distinta de la matriz nula y $T : M(n \times n) \rightarrow M(n \times n)$ dada por $T(X) = XA$. Muestre que A y T tienen los mismos valores propios.

16. Sea $A \in M(n \times n)$ tal que todos sus valores propios son 0. ¿Se puede concluir que $A = \mathbb{O}$? ¿Cómo cambia la respuesta si suponemos que A es diagonalizable?

8.4. Propiedades de los valores propios y vectores propios

En esta sección reunimos propiedades importantes de los vectores propios.

Proposición 8.44. Sea $A \in M(n \times n)$ y sean $\lambda_1, \lambda_2, \dots, \lambda_k$ valores propios de A distintos dos a dos, con vectores propios $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$. Entonces los vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ son linealmente independientes.

Demostración. Demostramos la afirmación por inducción.

Base de la inducción: $k = 2$. Supongamos que $\lambda_1 \neq \lambda_2$ son valores propios de A con vectores propios $\vec{v}_1$ y $\vec{v}_2$. Por lo tanto $A\vec{v}_1 = \lambda_1\vec{v}_1$ y $A\vec{v}_2 = \lambda_2\vec{v}_2$, y $\vec{v}_1 \neq \vec{0} \neq \vec{v}_2$. Sean α_1, α_2 números tales que

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 = \vec{0}. \quad (8.10)$$

Supongamos que $\alpha_1 \neq 0$. Entonces $\vec{v}_1 = \frac{\alpha_2}{\alpha_1}\vec{v}_2$ y

$$\lambda_1\vec{v}_1 = A\vec{v}_1 = A\left(\frac{\alpha_2}{\alpha_1}\vec{v}_2\right) = \frac{\alpha_2}{\alpha_1}A\vec{v}_2 = \frac{\alpha_2}{\alpha_1}\lambda_2\vec{v}_2 = \lambda_2\frac{\alpha_2}{\alpha_1}\vec{v}_2 = \lambda_2\vec{v}_1 \implies \vec{0} = (\lambda_1 - \lambda_2)\vec{v}_1.$$

Como $\lambda_1 \neq \lambda_2$ y $\vec{v}_1 \neq \vec{0}$, la última igualdad es falsa y por lo tanto debemos tener $\alpha_1 = 0$. Entonces, por (8.10), $\vec{0} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 = \alpha_2\vec{v}_2$, por lo tanto también $\alpha_2 = 0$ lo que muesgra que $\vec{v}_1$ y $\vec{v}_2$ son linealmente independientes.

Paso inductivo: Supongamos que ya sabemos para algún $j < k$ que los vectores $\vec{v}_1, \dots, \vec{v}_j$ son linealmente independientes. Tenemos que mostrar que entonces también los vectores $\vec{v}_1, \dots, \vec{v}_{j+1}$ son linealmente independientes. Para ello, sean $\alpha_1, \alpha_2, \dots, \alpha_{j+1}$ tales que

$$\vec{0} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \dots + \alpha_j\vec{v}_j + \alpha_{j+1}\vec{v}_{j+1}. \quad (8.11)$$

Por un lado aplicamos A a ambos lados de la ecuación y usamos el hecho de que los vectores son vectores propios. Por otro lado multiplicamos ambos lados por λ_{j+1} y luego comparamos los dos resultados.

$$\begin{aligned} \text{aplicar } A : \quad \vec{0} &= A(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \dots + \alpha_j\vec{v}_j + \alpha_{j+1}\vec{v}_{j+1}) \\ &= \alpha_1A\vec{v}_1 + \alpha_2A\vec{v}_2 + \dots + \alpha_jA\vec{v}_j + \alpha_{j+1}A\vec{v}_{j+1} \\ &= \alpha_1\lambda_1\vec{v}_1 + \alpha_2\lambda_2\vec{v}_2 + \dots + \alpha_j\lambda_j\vec{v}_j + \alpha_{j+1}\lambda_{j+1}\vec{v}_{j+1} \end{aligned} \quad (1)$$

$$\text{multiplicar por } \lambda_{j+1} : \quad \vec{0} = \alpha_1\lambda_{j+1}\vec{v}_1 + \alpha_2\lambda_{j+1}\vec{v}_2 + \dots + \alpha_j\lambda_{j+1}\vec{v}_j + \alpha_{j+1}\lambda_{j+1}\vec{v}_{j+1} \quad (2)$$

La diferencia (1)-(2) da

$$\vec{0} = \alpha_1(\lambda_1 - \lambda_{j+1})\vec{v}_1 + \alpha_2(\lambda_2 - \lambda_{j+1})\vec{v}_2 + \dots + \alpha_j(\lambda_j - \lambda_{j+1})\vec{v}_j.$$

Observe que el término con $\vec{v}_{j+1}$ se canceló. Por la hipótesis de inducción, los vectores $\vec{v}_1, \dots, \vec{v}_j$ son linealmente independientes, por lo tanto

$$\alpha_1(\lambda_1 - \lambda_{j+1}) = 0, \quad \alpha_2(\lambda_2 - \lambda_{j+1}) = 0, \quad \dots, \quad \alpha_j(\lambda_j - \lambda_{j+1}) = 0.$$

También sabemos que λ_{j+1} no es igual a ninguno de los otros λ_ℓ , por lo tanto se sigue que

$$\alpha_1 = 0, \quad \alpha_2 = 0, \quad \dots, \quad \alpha_j = 0.$$

Insertando esto en (8.11) se obtiene que también $\alpha_{j+1} = 0$ y la demostración queda completa. $\square$

Observe que la proposición muestra nuevamente que una matriz $n \times n$ puede tener a lo sumo n valores propios diferentes.

Corolario 8.45. Sea $A \in M(n \times n)$ y sean $\mu_1, \dots, \mu_k$ los valores propios distintos de A . Si en cada $\text{Eig}_{\mu_j}(A)$ elegimos vectores linealmente independientes $\vec{v}_1^j, \dots, \vec{v}_{\ell_j}^j$, entonces el sistema de todos esos vectores es linealmente independiente. En particular, si elegimos bases en $\text{Eig}_{\mu_j}(A)$, vemos que la suma de los espacios propios es una suma directa

$$\text{Eig}_{\mu_1}(A) \oplus \dots \oplus \text{Eig}_{\mu_k}(A)$$

$$\text{y } \dim(\text{Eig}_{\mu_1}(A) \oplus \dots \oplus \text{Eig}_{\mu_k}(A)) = \dim(\text{Eig}_{\mu_1}(A)) + \dots + \dim(\text{Eig}_{\mu_k}(A)).$$

Demostración. Sean $\alpha_j^{(m)}$ números tales que

$$\begin{aligned} \vec{0} &= \alpha_1^{(1)} \vec{v}_1^1 + \dots + \alpha_{\ell_1}^{(1)} \vec{v}_{\ell_1}^1 + \alpha_1^{(2)} \vec{v}_1^2 + \dots + \alpha_{\ell_2}^{(2)} \vec{v}_{\ell_2}^2 + \dots + \alpha_1^{(k)} \vec{v}_1^k + \dots + \alpha_{\ell_k}^{(k)} \vec{v}_{\ell_k}^k \\ &= \vec{w}_1 + \vec{w}_2 + \dots + \vec{w}_k \end{aligned}$$

con $\vec{w}_j = \alpha_1^{(j)} \vec{v}_1^j + \dots + \alpha_{\ell_j}^{(j)} \vec{v}_{\ell_j}^j \in \text{Eig}_{\mu_j}$. La Proposición 8.44 implica que $\vec{w}_1 = \dots = \vec{w}_k = \vec{0}$. Pero entonces también todos los coeficientes $\alpha_j^{(m)} = 0$ porque, para m fijo, los vectores $\vec{v}_1^{(m)}, \dots, \vec{v}_{\ell_m}^{(m)}$ son linealmente independientes. Ahora todas las afirmaciones son claras. $\square$

Una clase muy especial de matrices son las matrices diagonales.

Teorema 8.46. (I) Sea $D = \text{diag}(d_1, \dots, d_n) = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$ sea una matriz diagonal.

Entonces los valores propios de D son precisamente los números $d_1, \dots, d_n$ y la multiplicidad geométrica de cada valor propio es igual a su multiplicidad algebraica.

(II) Let $B = \begin{pmatrix} d_1 & & * \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$ and $C = \begin{pmatrix} d_1 & & 0 \\ * & & \\ & \ddots & \\ & & d_n \end{pmatrix}$ sean matrices triangulares superior y

triangulares inferior, respectivamente. Entonces los valores propios de D son precisamente los números $d_1, \dots, d_n$ y la multiplicidad algebraica de un valor propio es igual al número de veces que aparece en la diagonal. En general, no se puede decir nada sobre las multiplicidades geométricas.

Demostración. (I) Como el determinante de una matriz diagonal es el producto de sus elementos diagonales, obtenemos para el polinomio característico de D

$$p_D(\lambda) = \det(D - \lambda) = \det \begin{pmatrix} d_1 - \lambda & & 0 \\ & \ddots & \\ 0 & & d_n - \lambda \end{pmatrix} = (d_1 - \lambda) \cdots (d_n - \lambda).$$

Como los ceros del polinomio característico son los valores propios de D , hemos mostrado que los números en la diagonal de D son precisamente sus valores propios. La multiplicidad algebraica de un valor propio μ es igual al número de veces que se repite en la diagonal de D . La multiplicidad geométrica de μ es igual a $\dim(\ker(D - \mu \text{id}))$. Observe que $D - \mu \text{id}$ es una matriz diagonal y la j -ésima entrada en su diagonal es 0 si y solo si $\mu = d_j$. No es difícil ver que la dimensión del kernel de una matriz diagonal es igual al número de ceros en su diagonal. Así, en resumen, para un valor propio μ de A tenemos:

$$\begin{aligned} \text{multiplicidad algebraica de } \mu &= \text{número de veces que } \mu \text{ aparece en la diagonal de } D \\ &= \text{multiplicidad geométrica de } \mu. \end{aligned}$$

- (II) Como el determinante de una matriz triangular es el producto de sus elementos diagonales, obtenemos para el polinomio característico de B

$$p_B(\lambda) = \det(B - \lambda \text{id}) = \det \begin{pmatrix} d_1 - \lambda & & * \\ & \ddots & \\ 0 & & d_n - \lambda \end{pmatrix} = (d_1 - \lambda) \cdots (d_n - \lambda).$$

y análogamente para C . El razonamiento para las multiplicidades algebraicas de los valores propios es como en el caso de una matriz diagonal. Sin embargo, en general la multiplicidad algebraica y la geométrica de un valor propio de una matriz triangular pueden ser diferentes, como muestra el Ejemplo 8.40. $\square$

Ejemplo 8.47. Let $D = \begin{pmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}$. Entonces $p_D(\lambda) = (1 - \lambda)(5 - \lambda)^3(8 - \lambda)^2$.

Los valores propios son 1 (con mult. geom. = mult. alg. = 1), 5 (con mult. geom. = mult. alg. = 3) y 8 (con mult. geom. = mult. alg. = 2),

Teorema 8.48. Si A y B son matrices semejantes, entonces tienen el mismo polinomio característico. En particular, tienen los mismos valores propios con las mismas multiplicidades algebraicas. Además, también las multiplicidades geométricas son iguales.

Demostración. Sea C una matriz invertible tal que $A = C^{-1}BC$. Por lo tanto,

$$A - \lambda \text{id} = C^{-1}BC - \lambda \text{id} = C^{-1}BC - \lambda C^{-1}C = C^{-1}(B - \lambda \text{id})C$$

y obtenemos para el polinomio característico de A

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda \text{id}) = \det(C^{-1}(B - \lambda \text{id})C) = \det(C^{-1}) \det(B - \lambda \text{id}) \det C = \det(B - \lambda \text{id}) \\ &= p_B(\lambda). \end{aligned}$$

Esto muestra que A y B tienen los mismos valores propios y que sus multiplicidades algebraicas coinciden.

Ahora sea μ un valor propio. Entonces,

$$\begin{aligned}\text{Eig}_\mu(A) &= \ker(A - \mu \text{id}) = \ker(C^{-1}(B - \mu \text{id})C) = \ker((B - \mu \text{id})C) = C^{-1} \ker(B - \mu \text{id}) \\ &= C^{-1} \text{Eig}_\mu(B)\end{aligned}$$

donde en el penúltimo paso usamos que C^{-1} es invertible. La invertibilidad de C^{-1} también muestra que $\dim(C^{-1} \text{Eig}_\mu(B)) = \dim(\text{Eig}_\mu(B))$, por lo tanto $\dim \text{Eig}_\mu(A) = \dim(\text{Eig}_\mu(B))$, lo cual prueba que la multiplicidad geométrica de μ como valor propio de A es igual a la de B . $\square$

A continuación probamos un teorema muy importante sobre la diagonalización de matrices.

Teorema 8.49. *Sea $A \in M_{\mathbb{K}}(n \times n)$ con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$. Entonces las siguientes condiciones son equivalentes.*

- (I) *A es diagonalizable, es decir, existe una matriz diagonal D y una matriz invertible C tales que $C^{-1}AC = D$.*
- (II) *Para todo valor propio de A , sus multiplicidades geométrica y algebraica son iguales.*
- (III) *A tiene un conjunto de n vectores propios linealmente independientes.*
- (IV) *$\mathbb{K}^n$ tiene una base formada por vectores propios de A .*

Demostración. Sean $\mu_1, \dots, \mu_k$ los valores propios distintos de A y denotemos las multiplicidades algebraicas de μ_j por $m_j(A)$ y $m_j(D)$, y las multiplicidades geométricas por $n_j(A)$ y $n_j(D)$.

(I) $\implies$ (II): Por hipótesis, A y D son semejantes, así que tienen los mismos valores propios por el Teorema 8.48 y

$$m_j(A) = m_j(D) \quad \text{y} \quad n_j(A) = n_j(D) \quad \text{para todo } j = 1, \dots, k,$$

y el Teorema 8.46 muestra que

$$m_j(D) = n_j(D) \quad \text{para todo } j = 1, \dots, k,$$

porque D es una matriz diagonal. Por lo tanto concluimos que también

$$m_j(A) = n_j(A) \quad \text{para todo } j = 1, \dots, k.$$

(II) $\implies$ (III): Recordemos que las multiplicidades geométricas $n_j(A)$ son las dimensiones del kernel de $A - \mu_j \text{id}$. Así, en cada $\ker(A - \mu_j \text{id})$ podemos elegir una base formada por $n_j(A)$ vectores. En total tenemos $n_1(A) + \dots + n_k(A) = m_1(A) + \dots + m_k(A) = n$ such vectors and they are linealmente independientes by Corollary 8.45.

(III) $\implies$ (IV): Esto es claro porque $\dim \mathbb{K}^n = n$.

(IV) $\implies$ (I): Sea $\mathcal{B} = \{\vec{c}_1, \dots, \vec{c}_n\}$ una base de $\mathbb{K}^n$ formada por vectores propios de A y sean $d_1, \dots, d_n$ los valores propios correspondientes, es decir, $A\vec{c}_j = d_j\vec{c}_j$. Observe que los d_j no necesariamente son distintos dos a dos. Entonces la matriz $C = (\vec{c}_1 | \dots | \vec{c}_n)$ es invertible y $C^{-1}AC$ es

la representación de A en la base $\mathcal{B}$, por lo tanto $C^{-1}AC = \text{diag}(d_1, \dots, d_n)$. Más detalladamente, usando que $\vec{c}_j = C\vec{e}_j$ y $C^{-1}\vec{c}_j = \vec{e}_j$,

$$j\text{-ésima columna de } C^{-1}AC = C^{-1}AC\vec{e}_j = C^{-1}A\vec{c}_j = C^{-1}(d_j\vec{c}_j) = d_jC^{-1}\vec{c}_j = d_j\vec{e}_j,$$

por lo tanto $D = (d_1\vec{e}_1 | \dots | d_n\vec{e}_n) = \text{diag}(d_1, \dots, d_n)$. $\square$

Una consecuencia inmediata del Teorema 8.49 es la siguiente.

Corolario 8.50. *Si una matriz $A \in M(n \times n)$ tiene n valores propios diferentes, entonces es diagonalizable.*

Demostración. Si A tiene n valores propios diferentes $\lambda_1, \dots, \lambda_n$, entonces para cada uno de ellos la multiplicidad algebraica es igual a 1. Además,

$$1 \leq \text{multiplicidad geométrica} \leq \text{multiplicidad algebraica} = 1$$

para cada valor propio. Por lo tanto la multiplicidad algebraica y la geométrica para cada valor propio son iguales (ambas son iguales a 1) y la afirmación se sigue del Teorema 8.49. $\square$

Corolario 8.51. *Si la matriz $A \in M(n \times n)$ es diagonalizable, entonces su determinante es igual al producto de sus valores propios.*

Demostración. Sean $\lambda_1, \dots, \lambda_n$ los valores propios (no necesariamente distintos) de A y sea C una matriz invertible tal que $C^{-1}AC = D := \text{diag}(\lambda_1, \dots, \lambda_n)$. Entonces,

$$\det A = \det(CDC^{-1}) = (\det C)(\det D)(\det C^{-1}) = \det D = \prod_{j=1}^n \lambda_j. \quad \square$$

Teorema 8.52. *Sea $A \in M(n \times n)$ y sean $\mu_1, \dots, \mu_k$ sus valores propios distintos. Entonces A es diagonalizable si y solo si*

$$\mathbb{K}^n = \text{Eig}_{\mu_1}(A) \oplus \dots \oplus \text{Eig}_{\mu_k}(A). \quad (8.12)$$

donde $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$ según si A actúa sobre $\mathbb{R}$ o sobre $\mathbb{C}$.

Demostración. Denotemos la multiplicidad algebraica de cada μ_j por $m_j(A)$ y su multiplicidad geométrica por $n_j(A)$.

Si A es diagonalizable, entonces las multiplicidades geométrica y algebraica son iguales para cada valor propio. Por lo tanto,

$$\begin{aligned} \dim(\text{Eig}_{\mu_1}(A) \oplus \dots \oplus \text{Eig}_{\mu_k}(A)) &= \dim(\text{Eig}_{\mu_1}(A)) + \dots + \dim(\text{Eig}_{\mu_k}(A)) \\ &= n_1(A) + \dots + n_k(A) = m_1(A) + \dots + m_k(A) = n. \end{aligned}$$

Como todo subespacio n -dimensional de $\mathbb{K}^n$ es igual a $\mathbb{K}^n$, queda probado (8.12).

Supongamos ahora que (8.12) es cierta. Tenemos que mostrar que A es diagonalizable. En cada Eig_{μ_j} elegimos una base $\mathcal{B}_j$. Por (8.12), la colección de todos esos vectores de base forma una base de $\mathbb{K}^n$. Por lo tanto encontramos una base de $\mathbb{K}^n$ formada por vectores propios de A . Consecuentemente A es diagonalizable por el Teorema 8.49. $\square$

El teorema anterior dice que A es diagonalizable si y solo si hay suficientes vectores propios de A para generar $\mathbb{K}^n$. Este es el caso si y solo si $\mathbb{K}^n$ se descompone en la suma directa de subespacios en cada uno de los cuales A actúa simplemente multiplicando cada vector por un número (a saber, por el valor propio correspondiente).

Para practicar un poco las nociones de multiplicidad algebraica y geométrica, terminamos esta sección con una demostración alternativa del Teorema 8.49.

Demostración alternativa del Teorema 8.49. Probemos $(\text{I}) \Rightarrow (\text{IV}) \Rightarrow (\text{III}) \Rightarrow (\text{II}) \Rightarrow (\text{I})$.

$(\text{I}) \Rightarrow (\text{IV})$: Esto ya se discutió después de la Definición 8.23. Sea $D = \text{diag}(d_1, \dots, d_n)$ y sean $\vec{c}_1, \dots, \vec{c}_n$ las columnas de C . Claramente forman una base de $\mathbb{K}^n$ porque C es invertible. Por hipótesis sabemos que $AC = CD$. Por lo tanto tenemos que

$$A\vec{c}_j = j\text{-ésima columna de } AC = j\text{-ésima columna de } CD = d_j \cdot (j\text{-ésima columna de } C) = d_j \vec{c}_j.$$

Por lo tanto los vectores $\vec{c}_1, \dots, \vec{c}_n$ son linealmente independientes y son todos vectores propios de A , y por lo tanto incluso forman una base de $\mathbb{K}^n$.

$(\text{IV}) \Rightarrow (\text{III})$: Claro.

$(\text{III}) \Rightarrow (\text{II})$: Supongamos que $\vec{v}_1, \dots, \vec{v}_n$ es una base de $\mathbb{K}^n$ formada por vectores propios de A . Claramente, cada uno de ellos debe pertenecer a algún espacio propio de A . Sea ℓ_j el número de esos vectores que pertenecen a $\text{Eig}_{\mu_j}(A)$. De aquí se sigue que $\ell_j \leq n_j(A)$ porque los vectores son linealmente independientes y $n_j(A) = \dim \text{Eig}_{\mu_j}(A)$. So by Theorem 8.39 we have $\ell_j \leq n_j(A) \leq m_j(A)$ donde $m_j(A)$ es la multiplicidad algebraica de μ_j . Sumando sobre todos los vectores propios, obtenemos

$$n = \ell_1 + \dots + \ell_k \leq n_1(A) + \dots + n_k(A) \leq m_1(A) + \dots + m_k(A) = n$$

La primera igualdad se cumple porque los vectores son una base de $\mathbb{K}^n$ y la última igualdad se cumple por definición de la multiplicidad algebraica. Por lo tanto todos los signos $\leq$ son en realidad igualdades y $n_1(A) + \dots + n_k(A) = m_1(A) + \dots + m_k(A)$. Se sigue que

$$\begin{aligned} 0 &= n_1(A) + \dots + n_k(A) - [m_1(A) + \dots + m_k(A)] \\ &= [n_1(A) - m_1(A)] + \dots + [n_k(A) - m_k(A)]. \end{aligned}$$

Como $n_j(A) - m_j(A) \leq 0$ para todo $j = 1, \dots, k$, cada uno de los términos debe ser cero, lo que muestra que $n_j(A) = m_j(A)$ como se deseaba.

$(\text{II}) \Rightarrow (\text{I})$: Para cada $j = 1, \dots, k$ elijamos una base $\mathcal{B}_j$ de $\text{Eig}_{\mu_j}(A)$. Observe que cada base tiene $n_j(A)$ vectores. Por el Corolario 8.45, el sistema formado por todos estos vectores de base es linealmente independiente. Además, el número total de estos vectores es $n_1(A) + \dots + n_k(A) = m_1(A) + \dots + m_k(A) = n$ donde usamos la hipótesis de que las multiplicidades algebraica y geométrica son iguales para cada valor propio. Por lo tanto la colección de todos esos vectores forma una base de $\mathbb{K}^n$. Que A es diagonalizable se sigue ahora como en la demostración de $(\text{IV}) \Rightarrow (\text{I})$: $\square$

Ahora debería haber entendido

- por qué los vectores propios de valores propios diferentes de una matriz A son linealmente independientes,
- más generalmente, por qué la suma de los espacios propios es incluso una suma directa,
- por qué una matriz es diagonalizable si y solo si el espacio vectorial tiene una base formada por vectores propios de A ,
- multiplicidades algebraica y geométrica,

- etc.

Ahora debería ser capaz de

- verificar si una matriz dada es diagonalizable,
- si es diagonalizable, encontrar su diagonalización,
- etc.

Ejercicios.

1. Para cada una de las siguientes matrices, determine si son diagonalizables. Si lo es, encuentre una D que es semejante. $D = CAC^{-1}$.

$$A_1 = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} -1 & 4 & 2 & -7 \\ 0 & 5 & -3 & 6 \\ 0 & 0 & -5 & 1 \\ 0 & 0 & 0 & 11 \end{pmatrix}, \quad A_5 = \begin{pmatrix} 3 & 2 & 5 & 1 \\ 2 & 0 & 2 & 6 \\ 5 & 2 & 7 & -1 \\ 1 & 6 & -1 & 3 \end{pmatrix}.$$

2. Sea $T : M(2 \times 2) \rightarrow M(2 \times 2)$ dada por

$$T(A) = \frac{1}{2}(A - A^t).$$

Muestre que T es diagonalizable.

3. (a) Sea $D : P_n \rightarrow P_n$ dada por $Dp = p'$. ¿Es D diagonalizable?
(b) Sea $D : P_n \rightarrow P_n$ dada por $Dp = p + xp' + p''$. ¿Es D diagonalizable?
4. Sean $A, B \in M(n \times n)$.
(a) Si A, B son diagonalizables, ¿se sigue que $A + B$ es diagonalizable?
(b) Si AB es diagonalizable, ¿se sigue que A o B son diagonalizables?
(c) Si A, B son diagonalizables, ¿se sigue que AB es diagonalizable?
5. Calcule A^{50} para

$$A = \begin{pmatrix} -29 & 20 & -4 \\ 0 & 1 & 0 \\ 210 & -140 & 29 \end{pmatrix}.$$

6. ¿Para cuáles valores de $k, t \in \mathbb{R}$, la matriz

$$\begin{pmatrix} i & -1 & 0 \\ 0 & 3k + 2t & k - 4t - 5 \\ 0 & 0 & 5k - 8t \end{pmatrix}$$

es diagonalizable?

7. Sea $A \in M(n \times n)$ diagonalizable y sean $d_1, d_2, \dots, d_k$ todos sus valores propios distintos. Muestre que $(A - d_1 \text{id}_n)(A - d_2 \text{id}_n) \dots (A - d_k \text{id}_n) = \mathbb{O}_{n \times n}$. ¿Sigue siendo cierta la afirmación si no suponemos que A es diagonalizable?
8. Sea $A \in M(n \times n)$ triangular superior ó inferior. ¿Cuál es el polinomio característico de A ? ¿Puede dar condiciones de cuando A es diagonalizable?
9. Sean $A, B, C \in M(2 \times 2)$ y sea

$$V = \begin{pmatrix} A & C \\ \mathbb{O}_{2 \times 2} & B \end{pmatrix}$$

- (a) Muestre que el polinomio característico de V es la multiplicación de los polinomios característicos de A y B .
- (b) Si $C = \mathbb{O}_{2 \times 2}$, muestre que V es diagonalizable si y solo si A, B son diagonalizables.
- (c) ¿Es cierta la conclusión del inciso anterior si no suponemos que $C = \mathbb{O}_{2 \times 2}$?

8.5. Matrices simétricas y hermíticas

En esta sección trataremos matrices simétricas y hermíticas. Los resultados principales son que todos los valores propios de una matriz hermítica son reales, que los vectores propios correspondientes a valores propios diferentes son ortogonales y que toda matriz hermítica es diagonalizable. Observe que las matrices simétricas son un caso especial de las hermíticas, así que siempre que mostremos algo sobre matrices hermíticas, lo mismo será cierto para matrices simétricas.

Teorema 8.53. *Sea A una matriz hermítica. Entonces todo valor propio λ de A es real.*

Demostración. Sea A hermítica, es decir, $A^* = A$, y sea λ un valor propio de A con vector propio $\vec{v}$. Entonces $\vec{v} \neq \vec{0}$ y $A\vec{v} = \lambda\vec{v}$. Tenemos que mostrar que $\lambda = \bar{\lambda}$. Por lo tanto

$$\lambda \|\vec{v}\|^2 = \lambda \langle \vec{v}, \vec{v} \rangle = \langle \lambda \vec{v}, \vec{v} \rangle = \langle A\vec{v}, \vec{v} \rangle = \langle \vec{v}, A^* \vec{v} \rangle = \langle \vec{v}, A\vec{v} \rangle = \langle \vec{v}, \lambda \vec{v} \rangle = \bar{\lambda} \langle \vec{v}, \vec{v} \rangle = \bar{\lambda} \|\vec{v}\|^2.$$

Como $\vec{v} \neq \vec{0}$, se sigue que $\lambda = \bar{\lambda}$, lo cual significa que la parte imaginaria de λ es 0; por lo tanto $\lambda \in \mathbb{R}$. □

Teorema 8.54. *Sea A una matriz hermítica y sean λ_1, λ_2 dos valores propios diferentes de A con vectores propios $\vec{v}_1$ y $\vec{v}_2$, es decir $A\vec{v}_1 = \lambda_1\vec{v}_1$ and $A\vec{v}_2 = \lambda_2\vec{v}_2$. Entonces $\vec{v}_1 \perp \vec{v}_2$.*

Demostración. La demostración es similar a la demostración del Teorema 8.53. Tenemos que mostrar que $\langle \vec{v}_1, \vec{v}_2 \rangle = 0$. Observe que, por el Teorema 8.53, los valores propios λ_1, λ_2 son reales.

$$\lambda_1 \langle \vec{v}_1, \vec{v}_2 \rangle = \langle \lambda_1 \vec{v}_1, \vec{v}_2 \rangle = \langle A\vec{v}_1, \vec{v}_2 \rangle = \langle \vec{v}_1, A^* \vec{v}_2 \rangle = \langle \vec{v}_1, A\vec{v}_2 \rangle = \langle \vec{v}_1, \lambda_2 \vec{v}_2 \rangle = \bar{\lambda}_2 \langle \vec{v}_1, \vec{v}_2 \rangle = \lambda_2 \langle \vec{v}_1, \vec{v}_2 \rangle.$$

Como $\lambda_1 \neq \lambda_2$ por hipótesis, se sigue que $\langle \vec{v}_1, \vec{v}_2 \rangle = 0$. □

Corolario 8.55. *Sea A una matriz hermítica y sean λ_1, λ_2 dos valores propios diferentes de A . Entonces $\text{Eig}_{\lambda_1}(A) \perp \text{Eig}_{\lambda_2}(A)$.*

El siguiente teorema es uno de los teoremas más importantes del Álgebra Lineal.

Teorema 8.56. *Toda matriz hermítica es diagonalizable.*

Teorema 8.56*. *Toda matriz simétrica es diagonalizable.*

Posponemos la demostración de estos teoremas hasta el final de esta sección. Como consecuencia inmediata obtenemos el siguiente teorema muy importante.

Teorema 8.58. *Una matriz es hermítica si y solo si es diagonalizable unitariamente, es decir, existe una matriz unitaria Q y una matriz diagonal D tal que $D = Q^{-1}AQ = Q^*AQ$.*

La formulación del teorema anterior para matrices reales es:

Teorema 8.58*. *Una matriz es simétrica si y solo si es diagonalizable ortogonalmente, es decir, existe una matriz ortogonal Q y una matriz diagonal D tal que $D = Q^{-1}AQ = Q^tAQ$.*

En ambos casos, $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ donde los $\lambda_1, \dots, \lambda_n$ son los valores propios de A y las columnas de Q son los vectores propios correspondientes.

Demostración. Sea A una matriz hermítica. Del Teorema 8.56 sabemos que A es diagonalizable. Por lo tanto,

$$\mathbb{C}^n = \text{Eig}_{\mu_1}(A) \oplus \dots \oplus \text{Eig}_{\mu_k}(A)$$

donde $\mu_1, \dots, \mu_k$ son los valores propios distintos de A . En cada espacio propio $\text{Eig}_{\mu_j}(A)$ podemos elegir una base ortonormal $\mathcal{B}_j$ formada por n_j vectores $\vec{v}_1^j, \dots, \vec{v}_{n_j}^j$, donde n_j es la multiplicidad geométrica de μ_j . Sabemos que los espacios propios son ortogonales dos a dos por el Corolario 8.55. Por lo tanto el sistema de todos estos vectores forma una base ortonormal $\mathcal{B}$ de $\mathbb{C}^n$. Por lo tanto, la matriz Q cuyas columnas son los vectores de esta base es una matriz unitaria y $Q^{-1}AQ = D$.

Supongamos ahora que A es diagonalizable unitariamente. Tenemos que mostrar que A es hermítica. Sea Q una matriz unitaria y sea D una matriz diagonal tal que $D = Q^*AQ$. Entonces $A = QDQ^*$ y

$$A^* = (QDQ^*)^* = (Q^*)^*D^*Q^* = QDQ^* = A$$

donde usamos que $D^* = D$ porque D es una matriz diagonal cuyas entradas en la diagonal son números reales, ya que son los valores propios de A . $\square$

La demostración del Teorema 8.58* es la misma.

Corolario 8.60. *Si una matriz A es hermítica (o simétrica), entonces su determinante es el producto de sus valores propios.*

Demostración. Esto se sigue del Teorema 8.56 (o el Teorema 8.56*) y el Corolario 8.51. $\square$

Proof of Theorem 8.56. Sea $A \in M_{\mathbb{C}}(n \times n)$ una matriz hermítica y sean $\mu_1, \dots, \mu_k$ los valores propios distintos de A con multiplicidades geométricas $n_1, \dots, n_k$. Por el Teorema 8.52 es suficiente mostrar que

$$\mathbb{C}^n = \text{Eig}_{\mu_1}(A) \oplus \dots \oplus \text{Eig}_{\mu_k}(A).$$

Denotemos el lado derecho por U , es decir, $U := \text{Eig}_{\mu_1}(A) \oplus \cdots \oplus \text{Eig}_{\mu_k}(A)$. Entonces tenemos que mostrar que $U^\perp = \{\vec{0}\}$. Para obtener una contradicción, supongamos que esto no es cierto y sea $\ell = \dim(U^\perp)$. En cada $\text{Eig}_{\mu_j}(A)$ elegimos una base ortogonal $\vec{v}_1^{(j)}, \dots, \vec{v}_{n_j}^{(j)}$ y elegimos una base ortogonal $\vec{w}_1, \dots, \vec{w}_\ell$ en $U^\perp$. El conjunto de todos estos vectores es una base ortonormal $\mathcal{B}$ de $\mathbb{C}^n$ porque todos los espacios propios son ortogonales entre sí y a $U^\perp$. Sea Q la matriz cuyas columnas son estos vectores: $Q = (\vec{v}_1^{(1)} | \cdots | \vec{v}_{n_k}^{(k)} | \vec{w}_1 | \cdots | \vec{w}_\ell)$. Entonces Q es una matriz unitaria porque sus columnas son una base ortogonal de $\mathbb{C}^n$. A continuación definamos $B = Q^{-1}AQ$. Entonces B es hermítica porque $B^* = (Q^{-1}AQ)^* = Q^*A^*(Q^{-1})^* = Q^{-1}AQ = B$ donde usamos que $A = A^*$ por hipótesis y que $Q^{-1} = Q^*$ porque es una matriz unitaria. Por otro lado, B , al ser la representación matricial de A con respecto a la base $\mathcal{B}$, es de la forma

$$B = \begin{pmatrix} \mu_1 & & & & & \\ & \ddots & & & & \\ & & \mu_1 & & & \\ & & & \ddots & & \\ & & & & \mu_2 & \\ & & & & & \ddots \\ & & & & & & \mu_k & \\ & & & & & & & \ddots \\ & & & & & & & & C \end{pmatrix}.$$

Todos los espacios vacíos son 0 y C es una matriz $\ell \times \ell$ (es la representación matricial de la restricción de A a $U^\perp$ con respecto a la base $\vec{w}_1, \dots, \vec{w}_\ell$). El polinomio característico de C tiene al menos un cero, por lo tanto C tiene al menos un valor propio λ . Claramente, λ es entonces también un valor propio de B y, si $\vec{y} \in \mathbb{C}^\ell$ es un vector propio de C , obtenemos un vector propio de B con el mismo valor propio poniendo 0s como sus primeras $n - \ell$ componentes y $\vec{y}$ como sus últimas ℓ componentes. Como A y B tienen los mismos valores propios, λ debe ser igual a uno de los valores propios $\mu_1, \dots, \mu_k$, digamos $\lambda = \mu_{j_0}$. Pero entonces la dimensión del espacio propio $\text{Eig}_{\mu_{j_0}}(B)$ es estrictamente mayor que la dimensión de $\text{Eig}_{\mu_{j_0}}(A)$ which contradicts Theorem 8.48. Por lo tanto $U^\perp = \{\vec{0}\}$ y el teorema queda demostrado. $\square$

Demostración de Teorema 8.56.* La demostración es esencialmente la misma que la del Teorema 8.56. Solo tenemos que observar que, usando la notación de la demostración anterior, la matriz C es simétrica (porque B es simétrica). Si vemos C como una matriz compleja, tiene al menos un valor propio λ porque en $\mathbb{C}$ su polinomio característico tiene al menos un cero complejo. Sin embargo, como C es hermítica, todos sus valores propios son reales, por lo tanto λ es real; así, es un valor propio de C si la vemos como una matriz real. $\square$

Ahora debería haber entendido

- por qué las matrices hermíticas y simétricas tienen valores propios reales,
- por qué los vectores propios de valores propios diferentes de una matriz hermítica son perpendiculares entre sí,
- por qué una matriz hermítica/simétrica es diagonalizable ortogonalmente,

- que, salvo una rotación y quizá una reflexión, los espacios propios de una matriz hermítica están generados por los ejes coordenados,
- etc.

Ahora debería ser capaz de

- encontrar valores propios y vectores propios de matrices hermíticas/simétricas,
- diagonalizar matrices simétricas,
- escribir $\mathbb{C}^n$ (o $\mathbb{R}^n$) como suma directa de los espacios propios de una matriz hermítica (o simétrica) dada,
- etc.

Ejercicios.

1. Diagonalice ortogonalmente las siguientes matrices:

$$(a) \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}, \quad (b) \begin{pmatrix} 6 & 2 & 4 \\ 2 & 3 & 2 \\ 4 & 2 & 9 \end{pmatrix}, \quad (c) \begin{pmatrix} 2 & -1 & 0 \\ -1 & 3 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \quad (d) \begin{pmatrix} 1 & -3 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

2. De una matriz simétrica $A \in M(3 \times 3)$ se sabe que el polinomio característico es $p(\lambda) = \lambda^3 - 5\lambda^2 + 8\lambda - 4$.

- (a) Determine los valores propios de A y las multiplicidades geométricas y algebraicas.
- (b) Se sabe que $\ker(A - \text{id}) = \text{gen}\{3\vec{e}_2 - 4\vec{e}_3\}$. Encuentre los espacios propios de A .
- (c) Encuentre una matriz A que cumple lo arriba.

3. Diagonalice

$$A = \begin{pmatrix} 5 & 3(1+i) \\ 3(1-i) & 2 \end{pmatrix}.$$

4. (a) Dé una matriz simétrica tal que su kernel es el plano $x - 3y + 2z = 0$. ¿Cuál debe ser la imagen de la matriz que escogió?
- (b) Dé una matriz simétrica que tenga por imagen el plano $5x - y + z = 0$. ¿Cuál es su kernel?
- (c) Caracterice todas las matrices $M(3 \times 3)$ que tienen un único valor propio.
5. Obtenga una base ortogonal de $\mathbb{R}^n$ de vectores propios de T donde T es la transformación lineal dada en el Ejercicio 3. en Sección 7.4.
6. Sean $A, B \in M_{sym}(2 \times 2)$ y $C \in M(2 \times 2)$ todas matrices con entradas reales. Considere

$$V = \begin{pmatrix} A & C^t \\ C & B \end{pmatrix}.$$

Muestre que $V \in M_{sym}(2 \times 2)$ y que además V es diagonalizable.

7. Sea $A, B \in M(n \times n)$ con entradas complejas. Muestre que:
- (a) AA^* y A^*A son diagonalizables.
 - (b) Si A, B son hermíticas y $AB = BA$ entonces AB es diagonalizable.
8. Sean $A, B \in M_{sym}(2 \times 2)$ y $V = \begin{pmatrix} A & \mathbb{O} \\ \mathbb{O} & B \end{pmatrix}$ donde $\mathbb{O}$ es la matriz cero. Muestre que $V \in M_{sym}(4 \times 4)$ y que además, los valores propios de V son los valores propios de A junto con los valores propios de B .

8.6. Aplicación: Secciones cónicas

En esta sección estudiaremos ecuaciones cuadráticas en x y y . Recordemos que sabemos cómo tratar ecuaciones lineales en dos variables. La forma más general es

$$ax + by = d \quad (8.13)$$

con constantes a, b, d . A solución es a tuple (x, y) which satisfies (8.13). Podemos ver el conjunto de todas las soluciones como un subconjunto del plano $\mathbb{R}^2$. Como (8.13) es una ecuación lineal (un sistema 1×2 de ecuaciones lineales), sabemos que tenemos las siguientes posibilidades para el conjunto solución:

- (a) una *recta* si $a \neq 0$ o $b \neq 0$,
- (b) el *plano* $\mathbb{R}^2$ si $a = 0, b = 0$ y $d = 0$,
- (c) el *conjunto vacío* (sin solución) si $a = 0, b = 0$ y $d \neq 0$,

Ahora consideraremos la ecuación cuadrática

$$ax^2 + bxy + cy^2 = d \quad (8.14)$$

con constantes a, b, c, d .

En lo que sigue siempre supondremos que $d \geq 0$. Esto no implica pérdida de generalidad, porque si $d < 0$, podemos multiplicar ambos lados de (8.14) por -1 y reemplazar a, b, c por $-a, -b, -c$. El conjunto de soluciones no cambia.

De nuevo, queremos identificar las soluciones con subconjuntos de $\mathbb{R}^2$ y averiguar qué tipo de figuras son. La ecuación (8.14) no es lineal, así que debemos ver qué relación tiene (8.14) con lo que hemos estudiado hasta ahora. Resulta que el lado izquierdo de (8.14) puede escribirse como un producto interno

$$\left\langle G \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle \quad \text{con} \quad G = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}. \quad (8.15)$$

Question 8.5

La matriz G de (8.15) no es la única elección posible. Encuentre todas las matrices posibles G tales que $\langle G \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle = ax^2 + bxy + cy^2$.

La matriz G es muy conveniente porque es simétrica. Esto significa que, salvo una transformación *ortogonal*, es una matriz diagonal. Así que, una vez que sabemos cómo resolver el problema cuyo

G es diagonal, entonces lo sabemos para el caso general, pues las soluciones difieren solo por una rotación y quizá una reflexión. Esto nos motiva a estudiar primero el caso en que G es diagonal, es decir, cuyo $b = 0$.

Ecuación cuadrática sin término mixto ($b = 0$).

Si $b = 0$, entonces (8.14) se convierte en

$$ax^2 + cy^2 = d \quad (8.16)$$

con constantes $d \geq 0$ y $a, c \in \mathbb{R}$.

Observación 8.61. El conjunto solución es simétrico con respecto al eje x y al eje y , porque si algún (x, y) es una solución de (8.16), entonces también lo son $(-x, y)$ y $(x, -y)$.

Definamos

$$\alpha := \sqrt{|a|}, \quad \gamma := \sqrt{|c|}, \quad \text{por lo tanto} \quad \alpha^2 = \begin{cases} a & \text{si } a \geq 0, \\ -a & \text{si } a < 0 \end{cases} \quad \text{y} \quad \gamma^2 = \begin{cases} c & \text{si } c \geq 0, \\ -c & \text{si } c < 0. \end{cases}$$

Tenemos que distinguir varios casos según si los coeficientes a, c son positivos, negativos o 0.

Caso 1.1: $a > 0$ y $c > 0$. En este caso, la ecuación (8.16) se convierte en

$$\alpha^2 x^2 + \gamma^2 y^2 = d. \quad (8.16.1.1)$$

(I) Si $d > 0$, entonces (8.16.1.1) es la ecuación de una **elipse** cuyos ejes son paralelos a los ejes x y y . La intersección con el eje x está en $\pm \frac{\sqrt{d}}{\alpha} = \pm \sqrt{d/a}$ y la intersección con el eje y está en $\pm \frac{\sqrt{d}}{\gamma} = \pm \sqrt{d/c}$.

(II) Si $d = 0$, entonces la única solución de (8.16.1.1) es **el punto $(0, 0)$** .

Observación 8.62. Observe que la longitud de los semiejes de la elipse es proporcional a $\sqrt{d}$. Por lo tanto, cuando d disminuye, la elipse de (I) se hace más pequeña y para $d = 0$ degenera al punto $(0, 0)$ de (II).

Caso 1.2: $a < 0$ y $c < 0$. En este caso, la ecuación (8.16) se convierte en

$$-\alpha^2 x^2 - \gamma^2 y^2 = d. \quad (8.16.1.2)$$

(I) Si $d > 0$, entonces (8.16.1.2) **no tiene solución** porque el lado izquierdo siempre es menor o igual que 0, mientras que el lado derecho es estrictamente positivo.

(II) Si $d = 0$, entonces la única solución de (8.16.1.2) es **el punto $(0, 0)$** .

Caso 2.1: $a > 0$ y $c < 0$. En este caso, la ecuación (8.16) se convierte en

$$\alpha^2 x^2 - \gamma^2 y^2 = d. \quad (8.16.2.1)$$

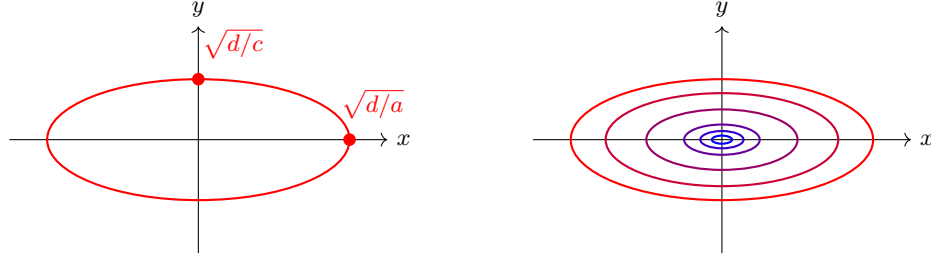


FIGURA 8.1: Solución de (8.16) para $\det G > 0$. Si $a > 0, b > 0$, entonces la solución es un elipse (si $d > 0$) o el punto $(0, 0)$ (si $d = 0$). La gráfica derecha muestra elipses con a y c fijos pero d decreciente (de rojo a azul). Si $a < 0, b < 0, d > 0$, entonces no hay solución.

- (I) Si $d > 0$, entonces (8.16.2.1) es la ecuación de una **hipérbola**. Si $x = 0$, la ecuación no tiene solución. En efecto, necesitamos $|x| \geq \frac{\sqrt{r}}{\alpha}$ para que la ecuación tenga solución. Por lo tanto la hipérbola no interseca el eje y (de hecho, la hipérbola no puede pasar por la franja $-\frac{\sqrt{d}}{\alpha} < y < \frac{\sqrt{d}}{\alpha}$).

- *Intersección con los ejes coordenados:* No hay intersección con el eje y . Intersección con el eje x en $x = \pm \frac{\sqrt{d}}{\alpha} = \pm \sqrt{d/a}$.
- *Asíntotas:* Para $|x| \rightarrow \infty$ y $|y| \rightarrow \infty$, la hipérbola tiene las asíntotas

$$y = \pm \frac{\alpha}{\gamma} x.$$

Observe que la asíntota no depende de d .

Demostración. Se sigue de (8.16.2.1) que $|x| \rightarrow \infty$ si y solo si $|y| \rightarrow \infty$, porque de lo contrario la diferencia $\alpha^2 x^2 - \gamma^2 y^2$ no puede ser constante. Dividiendo (8.16.2.1) por x^2 y por γ^2 y reordenando se llega a

$$\frac{y^2}{x^2} = \frac{\alpha^2}{\gamma^2} - \frac{d}{\gamma^2 x^2} \approx \frac{\alpha^2}{\gamma^2}, \quad \text{por lo tanto} \quad y \approx \pm \frac{\alpha}{\gamma} x. \quad \square$$

- (II) Si $d = 0$, entonces (8.16.2.1) se convierte en $\alpha^2 x^2 + \gamma^2 y^2 = 0$, y su solución es el **par de rectas $y = \pm \frac{\alpha}{\gamma} x$** .

Observación 8.63. Observe que el punto de intersección de la hipérbola con el eje x es proporcional a $\sqrt{d}$. Por lo tanto, cuyo d disminuye, los puntos de intersección se acercan al 0 y el giro se vuelve más agudo. Si $d = 0$, el punto de intersección llega a 0 y la hipérbola se convierte en dos ángulos que se ven como dos rectas que se cruzan.

Caso 2.2: $a < 0$ y $c > 0$. En este caso, la ecuación (8.16) se convierte en

$$-\alpha^2 x^2 + \gamma^2 y^2 = d. \quad (8.16.2.2)$$

Este caso es el mismo que el Caso 2.1, solo con los papeles de x y y intercambiados. Así obtenemos:

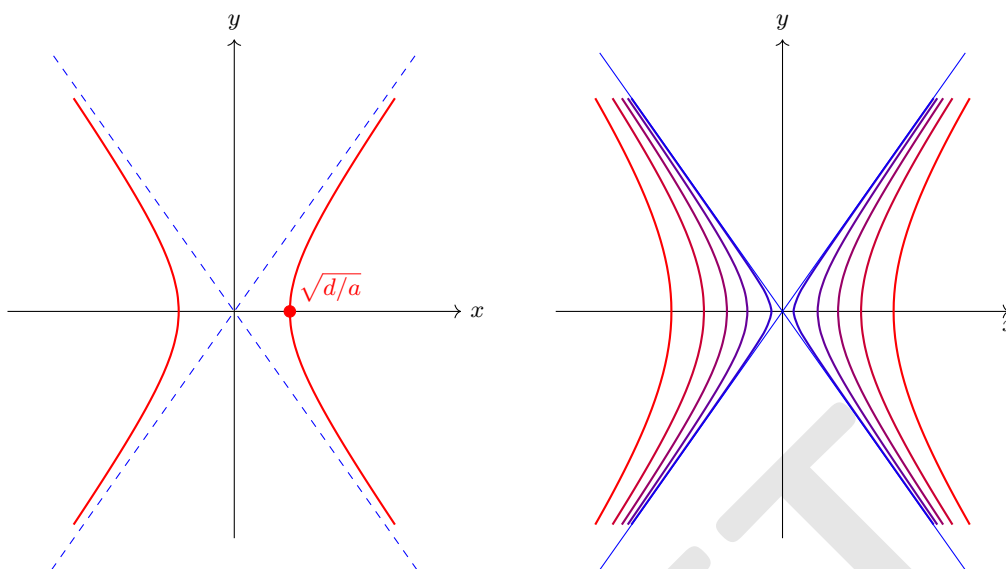


FIGURA 8.2: Solución de (8.16) para $\det G < 0$. Las soluciones son hipérbolas (si $d > 0$) o un conjunto de dos rectas que se intersecan. La figura de la izquierda muestra una solución para $a > 0$, $c < 0$ y $d > 0$. La figura de la derecha muestra hipérbolas con a y c fijos pero con d decreciente. El par azul de rectas que pasan por el origen corresponde al caso $d = 0$.

(I) Si $d > 0$, entonces (8.16.2.1) es la ecuación de una **hipérbola**.

- *Intersección con los ejes coordenados:* No hay intersección con el eje x . Intersección con el eje y en $y = \pm \frac{\sqrt{d}}{\gamma} = \pm \sqrt{d/c}$.
- *Asíntotas:* Para $|x| \rightarrow \infty$ y $|y| \rightarrow \infty$, la hipérbola tiene las asíntotas $y = \pm \frac{\alpha}{\gamma}x$.

(II) Si $d = 0$, entonces (8.16.2.1) se convierte en $\alpha^2 x^2 + \gamma^2 y^2 = 0$, y su solución es el **par de rectas $y = \pm \frac{\alpha}{\gamma}x$** .

Caso 3.1: $a > 0$ y $c = 0$. Entonces (8.16) se convierte en $\alpha^2 x^2 = d$.

- Si $d > 0$, las soluciones son las **dos rectas paralelas $x = \pm \frac{\sqrt{d}}{\alpha}$** .
- Si $d = 0$, la solución es la **recta $x = 0$** .

Caso 3.3: $a < 0$ y $c = 0$. Entonces (8.16) se convierte en $-\alpha^2 x^2 = d$.

- Si $d > 0$, **no hay solución**.
- Si $d = 0$, la solución es **la recta $x = 0$** .

Caso 3.2: $a = 0$ y $c > 0$. Entonces (8.16) se convierte en $\gamma^2 y^2 = d$.

- Si $d > 0$, las soluciones son las **dos rectas paralelas $y = \pm \frac{\sqrt{d}}{\gamma}$** .
- Si $d = 0$, la solución es la **recta $y = 0$** .

Caso 3.4: $a = 0$ y $c < 0$. Entonces (8.16) se convierte en $-\gamma^2 y^2 = d$.

- Si $d > 0$, **no hay solución**.
- Si $d = 0$, la solución es **la recta $y = 0$** .

Caso 3.5: $a = 0$ y $c = 0$. Entonces (8.16) se convierte en $0 = d$.

- Si $d > 0$, no hay solución.
- Si $d = 0$, la solución es $\mathbb{R}^2$.

Observe que en los Casos 1.1 y 1.2, $\det G = ac > 0$, en los Casos 2.1 y 2.2, $\det G = ac < 0$ y en todos los casos restantes $\det G = 0$.

Ecuación cuadrática con término mixto.

Ahora queremos resolver (8.14) sin la suposición de que $b = 0$. Sea $G = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$ y $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$. Entonces (8.14) es equivalente a

$$\langle G\vec{x}, \vec{x} \rangle = d. \quad (8.17)$$

Si G fuera diagonal, entonces podríamos dar inmediatamente la solución. Sabemos que G es simétrica, por lo tanto sabemos que G puede diagonalizarse ortogonalmente. En otras palabras, existe una base ortogonal de $\mathbb{R}^2$ con respecto a la cual G tiene una representación como matriz diagonal. Incluso podemos elegir esta base de modo que sea una rotación de la base canónica $\vec{e}_1$ y $\vec{e}_2$ (sin una reflexión adicional).

Sean λ_1, λ_2 valores propios de G y sea $D = \text{diag}(\lambda_1, \lambda_2)$. Elegimos una matriz ortogonal Q tal que

$$D = Q^{-1}GQ. \quad (8.18)$$

Denotemos las columnas de Q por $\vec{v}_1$ y $\vec{v}_2$. Son vectores propios normalizados de G con valores propios λ_1 y λ_2 , respectivamente. Recordemos que para una matriz ortogonal Q siempre tenemos que $\det Q = \pm 1$. Podemos suponer que $\det Q = 1$, porque si no, simplemente podemos multiplicar una de sus columnas por -1 . Esta columna sigue siendo entonces un vector propio normalizado de G con el mismo valor propio, por lo tanto (8.18) sigue siendo válida. Con esta elección garantizamos que Q es una rotación.

De (8.18) se sigue que $G = QDQ^{-1} = QDQ^*$. Así obtenemos de (8.17) que

$$d = \langle G\vec{x}, \vec{x} \rangle = \langle QDQ^*\vec{x}, \vec{x} \rangle = \langle DQ^*\vec{x}, Q^*\vec{x} \rangle = \langle D\vec{x}', \vec{x}' \rangle = \langle D\vec{x}', \vec{x}' \rangle = \lambda_1 x'^2 + \lambda_2 y'^2$$

donde $\vec{x}' = \begin{pmatrix} x' \\ y' \end{pmatrix} = Q^*\vec{x} = Q^{-1}\vec{x}$.

Observe que el vector columna $\begin{pmatrix} x' \\ y' \end{pmatrix}$ es la representación de $\vec{x}$ con respecto a la base $\vec{v}_1, \vec{v}_2$ (recordemos que son vectores propios de G). Por lo tanto, la solución de (8.14) es una de las soluciones que encontramos para el caso $b = 0$, solo que ahora los ejes de simetría de las figuras ya no son los ejes x y y , sino las direcciones de los vectores propios de G . En otras palabras: como Q es una rotación, obtenemos las soluciones de $ax^2 + bxy + cy^2 = d$ rotando con la matriz Q las soluciones de $ax^2 + cy^2 = d$.

Procedimiento para encontrar las soluciones de $ax^2 + bxy + cy^2 = d$.

- Escriba la matriz simétrica $G = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$.
- Encuentre los valores propios λ_1 y λ_2 y vectores propios de G y defina la matriz diagonal $D = \text{diag}(\lambda_1, \lambda_2)$. y la matriz ortogonal Q tal que $\det Q = 1$ y $D = Q^{-1}GQ$.
- Forma cuadrática sin términos mixtos: $d = \lambda_1 x'^2 + \lambda_2 y'^2$ donde x', y' son las componentes de $\vec{x}' = Q^{-1}\vec{x}$.
- Gráfica de la solución: En el sistema de coordenadas xy , indique el eje x' (paralelo a $\vec{v}_1$) y el eje y' (paralelo a $\vec{v}_2$). Observe que estos ejes son una rotación de los ejes x y y . Las soluciones son entonces, dependiendo de los valores propios, una elipse, hipérbola, etc., cuyos ejes de simetría son los ejes x' y y' .

Si queremos conocer solo la forma de la solución, basta calcular los valores propios λ_1, λ_2 de G , o incluso solo $\det G$. Recordemos que siempre suponemos $d \geq 0$.

- Si $\det G > 0$, entonces obtenemos una **elipse** (que puede ser degenerada).
 - Si $\lambda_1 > 0$ y $\lambda_2 > 0$, entonces la solución es una elipse con longitud de sus ejes $\sqrt{d/\lambda_1}$ y $\sqrt{d/\lambda_2}$. Si $d = 0$ la elipse es solo el punto $(0, 0)$.
 - Si $\lambda_1 < 0$ y $\lambda_2 < 0$, entonces no hay solución (si $d > 0$) o la solución es solo el punto $(0, 0)$ (si $d = 0$).
- Si $\det G < 0$, entonces obtenemos una **hipérbola** (que puede ser degenerada).
 - Si $\lambda_1 > 0$ y $\lambda_2 < 0$, entonces la solución es una hipérbola que interseca el eje x' en $\sqrt{d/\lambda_1}$ y no tiene intersección con el eje y' .
 - Si $\lambda_1 < 0$ y $\lambda_2 > 0$, entonces la solución es una hipérbola que interseca el eje x' en $\sqrt{d/\lambda_2}$ y no tiene intersección con el eje x' .

En ambos casos, las asíntotas de la hipérbola tienen pendiente $\pm\sqrt{\lambda_1/\lambda_2}$. Si $d = 0$, la hipérbola degenera al par de rectas $y = \pm\sqrt{\lambda_1/\lambda_2} x$.

- Si $\det G = 0$, entonces obtenemos ya sea el conjunto vacío, uno de los ejes, dos rectas paralelas a uno de los ejes, o $\mathbb{R}^2$.

Definición 8.64. Los ejes de simetría se llaman *ejes principales*.

Ejemplo 8.65. Considere la ecuación

$$10x^2 + 6xy + 2y^2 = 4. \quad (8.19)$$

- (I) Escriba la ecuación en forma matricial.
- (II) Haga un cambio de coordenadas de modo que la ecuación cuadrática (8.19) no tenga término mixto.
- (III) Describa la solución de (8.19) en términos geométricos y haga un bosquejo. Indique los ejes principales y las intersecciones importantes.

Solución. (i) Primero escribimos (8.19) en la forma $\langle G\vec{x}, \vec{x} \rangle$ con una matriz simétrica G . Definamos $G = \begin{pmatrix} 10 & 3 \\ 3 & 2 \end{pmatrix}$. Entonces (8.19) es equivalente a

$$\left\langle G \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = 4. \quad (8.20)$$

(ii) Ahora calculamos los **valores propios** de G . Son las raíces del polinomio característico $\det(G - \lambda)$.

$$0 = \det(G - \lambda) = (10 - \lambda)(2 - \lambda) - 9 = \lambda^2 - 12\lambda + 11 = (\lambda - 6)^2 - 25 = (\lambda - 1)(\lambda - 11).$$

Por lo tanto, los valores propios de G son

$$\lambda_1 = 1, \quad \lambda_2 = 11.$$

A continuación necesitamos los **vectores propios** normalizados. Para ello, calculamos $\ker(G - \lambda_j)$ usando eliminación de Gauß:

$$\begin{aligned} \bullet \quad G - \lambda_1 &= \begin{pmatrix} 9 & 3 \\ 3 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 3 & 1 \\ 0 & 0 \end{pmatrix} \Rightarrow \vec{v}_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ -3 \end{pmatrix}, \\ \bullet \quad G - \lambda_2 &= \begin{pmatrix} -1 & 3 \\ 3 & -9 \end{pmatrix} \rightarrow \begin{pmatrix} -1 & 3 \\ 0 & 0 \end{pmatrix} \Rightarrow \vec{v}_2 = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ 1 \end{pmatrix}. \end{aligned}$$

(Recordemos que para matrices simétricas los vectores propios correspondientes a valores propios diferentes son ortogonales. Si resuelve un ejercicio de este tipo, puede ser buena idea verificar si los vectores son en efecto ortogonales entre sí.)

Observación. Con la información obtenida hasta ahora, ya podemos bosquejar la solución.

- La solución es una elipse porque ambos valores propios son positivos.
- Los ejes principales (ejes de simetría) son paralelos a los vectores $\vec{v}_1$ u $\vec{v}_2$. La elipse los interseca en $\pm\sqrt{4/1} = \pm 2$ a lo largo del eje paralelo a $\vec{v}_1$ y en $\pm\sqrt{4/11} = \pm 2/\sqrt{11}$ a lo largo del eje paralelo a $\vec{v}_2$.

Pongamos

$$Q = (\vec{v}_1 | \vec{v}_2) = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 3 \\ -3 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 11 \end{pmatrix},$$

entonces

$$Q^{-1} = Q^t \quad \text{y} \quad D = Q^{-1}GQ = Q^tGQ.$$

Observe que $\det Q = 1$, por lo tanto es una rotación en $\mathbb{R}^2$. Es una rotación por el ángulo $\arctan(-3)$.

Si definimos

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = Q^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} x - 3y \\ 3x + y \end{pmatrix},$$

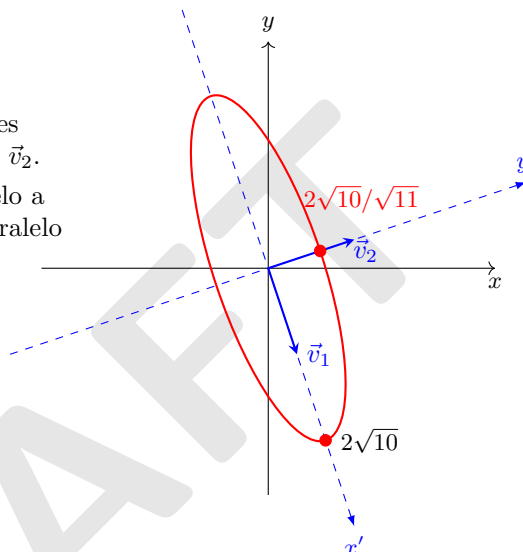
entonces (8.20) gives

$$4 = \left\langle G \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = \left\langle DQ^t \begin{pmatrix} x \\ y \end{pmatrix}, Q^t \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = \left\langle D \begin{pmatrix} x' \\ y' \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix} \right\rangle,$$

y por lo tanto

$$4 = x'^2 + 11y'^2 = \frac{1}{10}(x - 3y)^2 + \frac{11}{10}(3x + y)^2.$$

- (III) La solución de (8.19) es una elipse cuyos ejes principales son paralelos a los vectores $\vec{v}_1$ y $\vec{v}_2$. x' es la coordenada a lo largo del eje paralelo a $\vec{v}_1$, y' es la coordenada a lo largo del eje paralelo a $\vec{v}_2$.



◇

Ejemplo 8.66. Considere la ecuación

$$-\frac{47}{17}x^2 - \frac{32}{17}xy + \frac{13}{17}y^2 = 2. \quad (8.21)$$

- (I) Escriba la ecuación en forma matricial.
- (II) Haga un cambio de coordenadas de modo que la ecuación cuadrática (8.21) no tenga término mixto.
- (III) Describa la solución de (8.21) en términos geométricos y haga un bosquejo. Indique los ejes principales y las intersecciones importantes.

Solución. (I) Primero escribimos (8.21) en la forma $\langle G\vec{x}, \vec{x} \rangle$ con una matriz simétrica G . Definamos $G = \frac{1}{17} \begin{pmatrix} -47 & -16 \\ -16 & 13 \end{pmatrix}$. Entonces (8.21) es equivalente a

$$\left\langle G \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = 2. \quad (8.22)$$

(II) Ahora calculamos los **valores propios** de G . Son las raíces del polinomio característico

$$0 = \det(G - \lambda) = \left(-\frac{47}{17} - \lambda\right)\left(\frac{13}{17} - \lambda\right) - \frac{128}{17^2} = \lambda^2 + \frac{34}{17}\lambda - \frac{611}{17^2} - \frac{256}{17^2} = \lambda^2 + 2\lambda - 3 = (\lambda - 1)(\lambda + 3).$$

Por lo tanto, los valores propios de G son

$$\lambda_1 = -3, \quad \lambda_2 = 1.$$

A continuación necesitamos los **vectores propios** normalizados. Para ello, calculamos $\ker(G - \lambda_j)$ usando eliminación de Gauß:

$$\begin{aligned} \bullet \quad G - \lambda_1 &= \frac{1}{17} \begin{pmatrix} 4 & -16 \\ -16 & 64 \end{pmatrix} \rightarrow \frac{1}{17} \begin{pmatrix} 1 & -4 \\ 0 & 0 \end{pmatrix} \quad \Rightarrow \quad \vec{v}_1 = \frac{1}{\sqrt{17}} \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \\ \bullet \quad G - \lambda_2 &= \frac{1}{17} \begin{pmatrix} -64 & -16 \\ -16 & -4 \end{pmatrix} \rightarrow \frac{1}{17} \begin{pmatrix} 4 & 1 \\ 0 & 0 \end{pmatrix} \quad \Rightarrow \quad \vec{v}_2 = \frac{1}{\sqrt{17}} \begin{pmatrix} -1 \\ 4 \end{pmatrix}. \end{aligned}$$

Observación. Con la información obtenida hasta ahora, ya podemos bosquejar la solución.

- La solución es una hipérbola porque los valores propios tienen signos opuestos.
- Los ejes principales (ejes de simetría) son paralelos a los vectores $\vec{v}_1$ y $\vec{v}_2$. Las intersecciones de la hipérbola con el eje paralelo a $\vec{v}_2$ son $\pm\sqrt{2}$.

Pongamos

$$Q = (\vec{v}_1 | \vec{v}_2) = \frac{1}{\sqrt{17}} \begin{pmatrix} 4 & -1 \\ 1 & 4 \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} -3 & 0 \\ 0 & 1 \end{pmatrix},$$

entonces

$$Q^{-1} = Q^t \quad \text{y} \quad D = Q^{-1}GQ = Q^t G Q.$$

Observe que $\det Q = 1$, por lo tanto Q es una rotación de $\mathbb{R}^2$. Es una rotación por el ángulo $\arctan(1/4)$.

Si definimos

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = Q^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{17}} \begin{pmatrix} 4x + y \\ -x + 4y \end{pmatrix},$$

entonces (8.22) gives

$$2 = \left\langle G \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = \left\langle D Q^t \begin{pmatrix} x \\ y \end{pmatrix}, Q^t \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = \left\langle D \begin{pmatrix} x' \\ y' \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix} \right\rangle,$$

por lo tanto

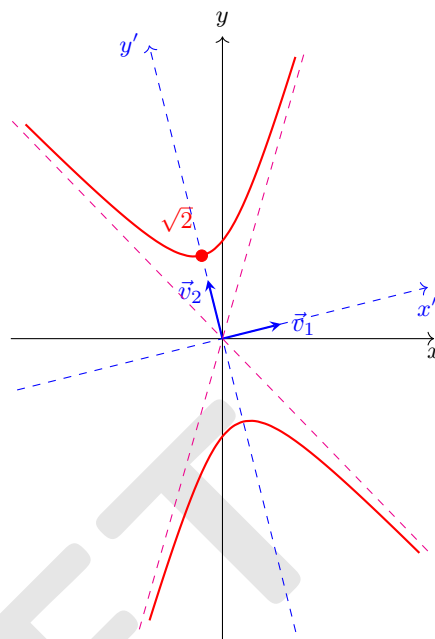
$$2 = -3x'^2 + y'^2 = -\frac{3}{17}(4x + y)^2 + \frac{1}{17}(-x + 4y)^2.$$

- (III) La solución de la ecuación (8.19) es una hipérbola cuyos ejes principales son paralelos a los vectores $\vec{v}_1$ y $\vec{v}_2$.

x' es la coordenada a lo largo del eje paralelo a $\vec{v}_1$,

y' es la coordenada a lo largo del eje paralelo a $\vec{v}_2$.

El ángulo entre los ejes x y x' es $\arctan(1/4)$.



Asíntotas de la hipérbola. Para calcular las pendientes de las asíntotas de la hipérbola, primero calculamos en el sistema de coordenadas $x'-y'$. Nuestro punto de partida es la ecuación $2 = -3x'^2 + y'^2$.

$$2 = -3x'^2 + y'^2 \iff \frac{y'^2}{x'^2} = 3 + \frac{1}{2x'^2} \iff \frac{y'}{x'} = \pm \sqrt{3 + \frac{1}{2x'^2}}.$$

Vemos que $|y'| \rightarrow \infty$ si y solo si $|x'| \rightarrow \infty$ y que $\frac{y'}{x'} \approx \pm\sqrt{3}$. Así, las pendientes de las asíntotas en coordenadas $x'-y'$ son $\pm\sqrt{3}$.

¿Cómo encontramos la pendiente en coordenadas x, y ?

- **Método 1: Usar Q .** Sabemos que si rotamos nuestra hipérbola mediante la transformación lineal Q^{-1} (es decir, si rotamos por $\arctan(1/4)$), entonces obtenemos una hipérbola cuyos ejes de simetría son los ejes x y y y cuyas asíntotas tienen pendientes ± 3 . Por lo tanto, para obtener las asíntotas de nuestra hipérbola, solo necesitamos aplicar Q a los vectores $\vec{w}_1$ y $\vec{w}_2$, que son paralelos a las nuevas asíntotas. Los vectores resultantes son entonces paralelos a nuestra hipérbola original. En nuestro caso $\vec{w}_1 = \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}$, $\vec{w}_2 = \begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix}$. Por lo tanto

$$\vec{w}'_1 = Q\vec{w}_1 = \frac{1}{\sqrt{17}} \begin{pmatrix} 4 & 1 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix} = \frac{1}{\sqrt{17}} \begin{pmatrix} 4 + \sqrt{3} \\ -1 + 4\sqrt{3} \end{pmatrix},$$

$$\vec{w}'_2 = Q\vec{w}_2 = \frac{1}{\sqrt{17}} \begin{pmatrix} 4 & 1 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix} = \frac{1}{\sqrt{17}} \begin{pmatrix} 4 - \sqrt{3} \\ -1 - 4\sqrt{3} \end{pmatrix}.$$

Por lo tanto, las pendientes de las asíntotas de nuestra hipérbola son

$$\frac{-1 + 4\sqrt{3}}{4 + \sqrt{3}} \quad \text{y} \quad \frac{-1 - 4\sqrt{3}}{4 - \sqrt{3}}.$$

- **Método 2: Sustituir en las fórmulas.** Las asíntotas son rectas que satisfacen $\frac{y'}{x'} = \pm\sqrt{3}$. Usando $x' = \frac{1}{\sqrt{17}}(4x - y)$ y $y' = \frac{1}{\sqrt{17}}(x + 4y)$, obtenemos

$$\begin{aligned} \pm\sqrt{3} &= \frac{y'}{x'} = \frac{\frac{1}{\sqrt{17}}(x + 4y)}{\frac{1}{\sqrt{17}}(4x - y)} = \frac{x + 4y}{4x - y} \\ \iff \pm\sqrt{3}(4x - y) &= x + 4y \\ \iff (\pm 4\sqrt{3} - 1)x &= (4 \pm \sqrt{3})y \\ \iff \frac{y}{x} &= \frac{-1 \pm 4\sqrt{3}}{4 \pm \sqrt{3}}. \end{aligned}$$

- **Método 3: Sumar ángulos.** Sabemos que el ángulo entre el eje x' y una asíntota es $\arctan \sqrt{3}$ y el ángulo entre el eje x' y el eje x es $\arctan(1/4)$. Por lo tanto, el ángulo entre la asíntota y el eje x es $\arctan \sqrt{3} + \arctan(1/4)$ (vea Figure 8.3.)

$$-3x^2 + y^2 = 2$$

$$-\frac{47}{17}x^2 - \frac{32}{17}16xy + \frac{13}{17}y^2 = 2$$

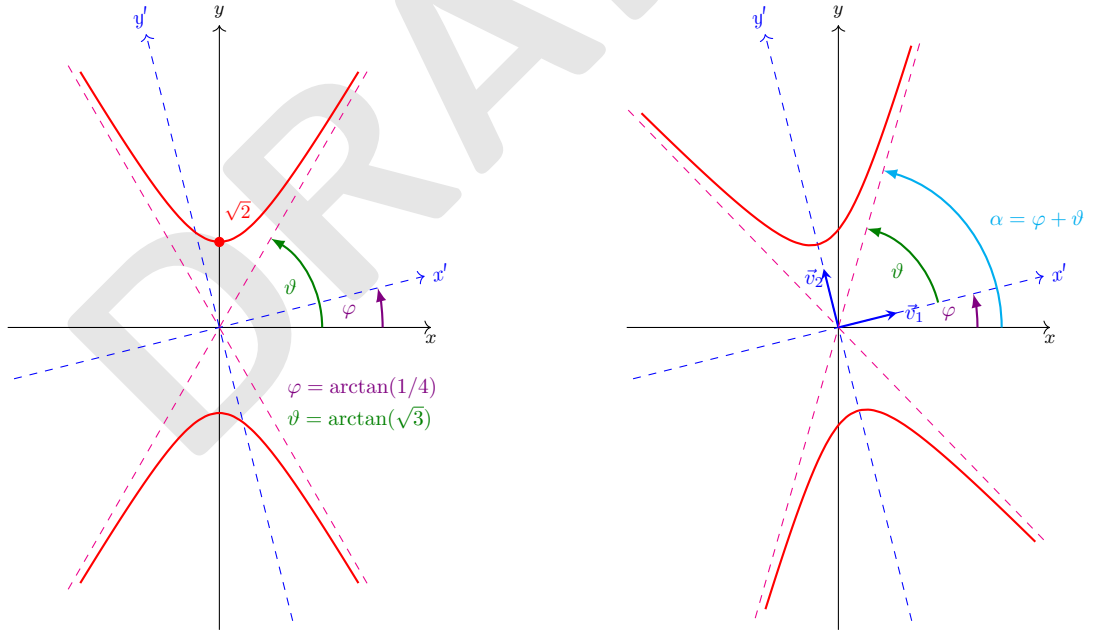


FIGURA 8.3: La figura de la derecha (nuestra hipérbola) se obtiene de la figura de la izquierda aplicándole la transformación Q (es decir, rotándola por $\arctan(1/4)$).

◇

Ejemplo 8.67. Considere la ecuación

$$9x^2 - 6xy + y^2 = 25. \quad (8.23)$$

- (I) Escriba la ecuación en forma matricial.
- (II) Haga un cambio de coordenadas de modo que la ecuación cuadrática (8.23) no tenga término mixto.
- (III) Describa la solución de (8.23) en términos geométricos y haga un bosquejo. Indique los ejes principales y las intersecciones importantes.

Solution 1. ■ Primero escribimos (8.21) en la forma $\langle G\vec{x}, \vec{x} \rangle$ con una matriz simétrica G . Definamos $G = \begin{pmatrix} 9 & -3 \\ -3 & 1 \end{pmatrix}$. Entonces (8.23) es equivalente a

$$\left\langle G \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = 25. \quad (8.24)$$

- Ahora calculamos los **valores propios** de G . Son las raíces del polinomio característico

$$0 = \det(G - \lambda) = (9 - \lambda)(1 - \lambda) - 9 = \lambda^2 - 10\lambda = \lambda(\lambda - 10).$$

Por lo tanto, los valores propios de G son

$$\lambda_1 = 0, \quad \lambda_2 = 10.$$

A continuación necesitamos los **vectores propios** normalizados. Para ello, calculamos $\ker(G - \lambda_j)$ usando eliminación de Gauß:

$$\begin{aligned} \bullet \quad G - \lambda_1 &= \begin{pmatrix} 9 & -3 \\ -3 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 3 & -1 \\ 0 & 0 \end{pmatrix} \implies \vec{v}_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \\ \bullet \quad G - \lambda_2 &= \begin{pmatrix} -1 & -3 \\ -3 & -9 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix} \implies \vec{v}_2 = \frac{1}{\sqrt{10}} \begin{pmatrix} -3 \\ 1 \end{pmatrix}. \end{aligned}$$

Observación. Con la información obtenida hasta ahora, ya podemos bosquejar la solución.

- La solución consiste en dos rectas paralelas porque uno de los valores propios es cero y el otro es positivo.
- Las rectas son paralelas a $\vec{v}_1$ y sus intersecciones con el eje paralelo a $\vec{v}_1$ son $\pm\sqrt{25/10} = \pm\sqrt{5/2}$.

Pongamos

$$Q = (\vec{v}_1 | \vec{v}_2) = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 10 \end{pmatrix},$$

entonces

$$Q^{-1} = Q^t \quad \text{y} \quad D = Q^{-1}GQ = Q^tGQ.$$

Observe que $\det Q = 1$, por lo tanto Q es una rotación en $\mathbb{R}^2$. Es una rotación por el ángulo $\arctan(3)$.

Si definimos

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = Q^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} x + 3y \\ -3x + y \end{pmatrix},$$

entonces (8.24) gives

$$25 = \left\langle G \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = \left\langle DQ^t \begin{pmatrix} x \\ y \end{pmatrix}, Q^t \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = \left\langle D \begin{pmatrix} x' \\ y' \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix} \right\rangle,$$

por lo tanto

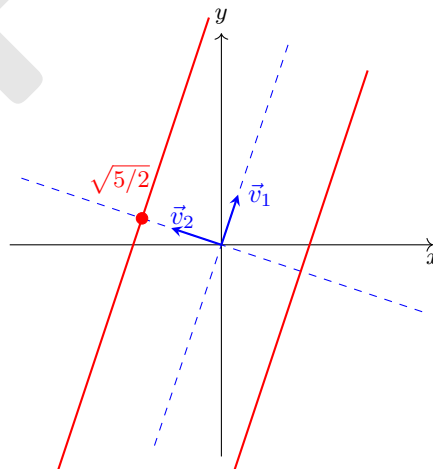
$$25 = 10y'^2 = (-3x + y)^2.$$

- La solución de (8.19) son dos rectas paralelas al vector $\vec{v}_1$ que intersecan el eje y' en $\pm\sqrt{25/10} = \pm\sqrt{5/2}$.

x' es la coordenada a lo largo del eje paralelo a $\vec{v}_1$,

y' es la coordenada a lo largo del eje paralelo a $\vec{v}_2$.

El ángulo entre los ejes x y x' es $\arctan(3)$.



Solution 2. Observe que

$$25 = 9x^2 - 6xy + y^2 = (3x - y)^2 \quad \Longleftrightarrow \quad 5 = |3x - y|.$$

Por lo tanto, la solución son dos rectas paralelas dadas por

$$y = 3x \pm 5$$

lo cual coincide con el resultado anterior.

8.6.1. Soluciones de $ax^2 + bxy + cy^2 = d$ como secciones cónicas

La razón por la cual el título de esta sección es “sección cónica” es que la mayoría de los conjuntos solución de las ecuaciones cuadráticas pueden obtenerse como la intersección de un cono doble con planos.

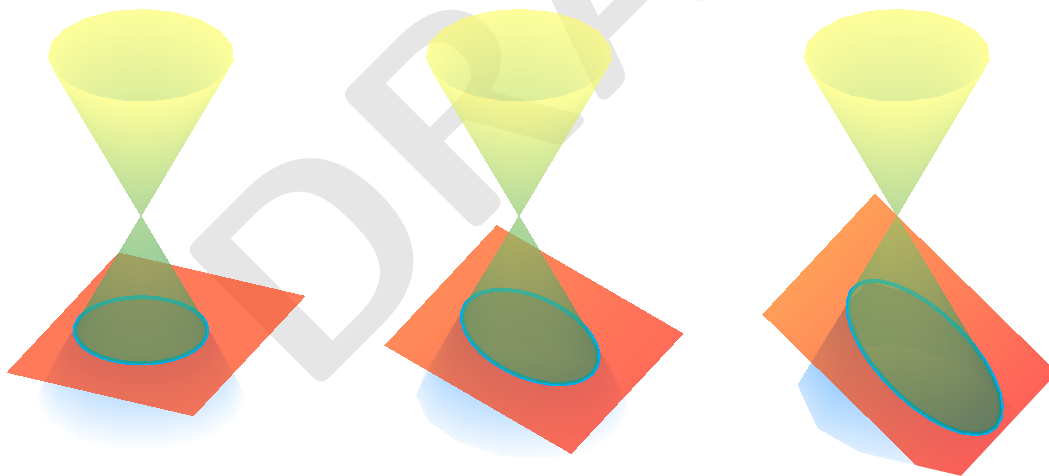


FIGURA 8.4:

Elipses. El plano en la imagen de la izquierda es paralelo al plano xy . Por lo tanto, la intersección con el cono es un círculo. Si el plano comienza a inclinarse, la intersección se convierte en una elipse. Cuanto más inclinado está el plano, más alargada es la elipse. Mientras el plano aún no sea paralelo a la superficie del cono, interseca solo la parte superior o la parte inferior del cono, y la intersección es una elipse.

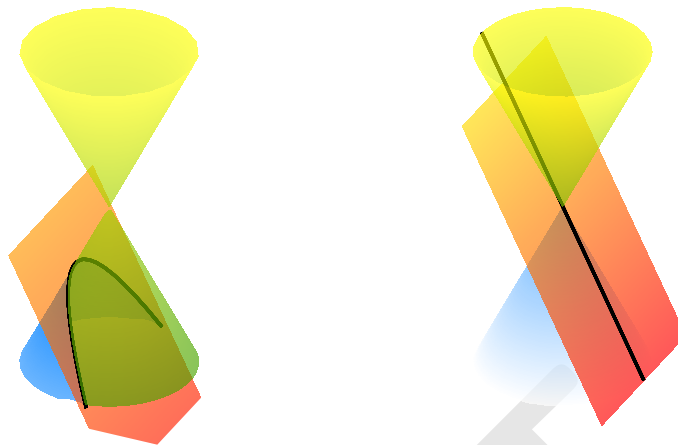


FIGURA 8.5: **Parábola.** Si el plano es paralelo a la superficie del cono y no pasa por el origen, entonces la intersección con el cono es una parábola (esta **no** es una posible solución de (8.14)). Si el plano es paralelo a la superficie del cono y pasa por el origen, entonces el plano es tangente al cono y la intersección es una recta.

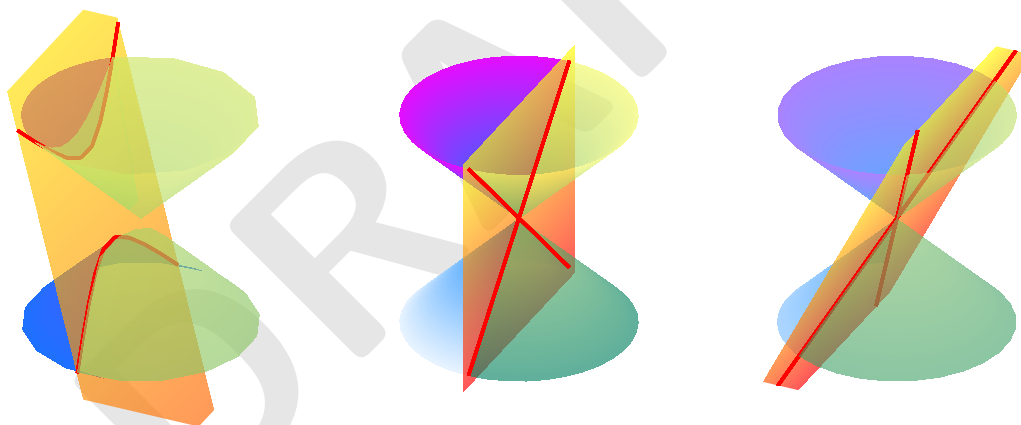


FIGURA 8.6: **Hipérbola.** Si el plano es más inclinado que el cono, entonces interseca tanto la parte superior como la inferior del cono. La intersección es una hipérbola. Si el plano pasa por el origen, entonces la hipérbola degenera en dos rectas que se intersecan. El plano en la imagen del medio es paralelo al plano yz . Por lo tanto, la intersección con el cono es un círculo.

8.6.2. Soluciones de $ax^2 + bxy + cy^2 + rx + sy = d$

Discutamos brevemente el caso en que la ecuación cuadrática (8.14) contiene términos lineales:

$$ax^2 + bxy + cy^2 + rx + sy = d \quad (8.25)$$

Queremos encontrar una transformación de modo que (8.25) pueda escribirse sin los términos lineales rx y sy .

Sea $G = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$ y sean λ_1, λ_2 sus valores propios. Además, sea $D = \text{diag}(\lambda_1, \lambda_2)$ y sea Q una matriz ortogonal con $\det Q = 1$ y $D = Q^{-1}GQ$.

En lo que sigue suponemos que G es invertible.

Método 1. Primero eliminamos el término mixto bxy .

Si ponemos $\vec{x}' = Q^{-1}\vec{x}$, entonces $ax^2 + bxy + cy^2 = \lambda_1 x'^2 + \lambda_2 y'^2$. Como x' y y' son lineales en x y y , la ecuación (8.25) se convierte en

$$\lambda_1 x'^2 + \lambda_2 y'^2 + r'x' + s'y' = d'.$$

Ahora solo necesitamos completar cuadrados en el lado izquierdo para obtener

$$\lambda_1(x' + r'/2)^2 + \lambda_2(y' + s'/2)^2 - (r'/2)^2 - (s'/2)^2 = d'.$$

Observe que esto siempre puede hacerse si λ_1 y λ_2 no son 0 (aquí usamos que G es invertible).

Si ponemos $d'' = d' + (r'/2)^2 + (s'/2)^2$, $x'' = x' + r'/2$, $y'' = y' + s'/2$, entonces

$$\lambda_1 x''^2 + \lambda_2 y''^2 = d''. \quad (8.26)$$

Como $\vec{x}'' = \begin{pmatrix} r'/2 \\ s'/2 \end{pmatrix} + \vec{x}' = \begin{pmatrix} r'/2 \\ s'/2 \end{pmatrix} + Q^{-1}\vec{x}$ vemos que la solución es la solución de $\lambda_1 x^2 + \lambda_2 y^2 = d''$, pero rotada por Q y trasladada por el vector $\begin{pmatrix} r'/2 \\ s'/2 \end{pmatrix}$.

Método 2. Primero eliminamos los términos lineales rx y sy .

Hagamos el ansatz $x = x_0 + \tilde{x}$ y $y = y_0 + \tilde{y}$. Sustituyendo en (8.25) se obtiene

$$\begin{aligned} d &= a(x_0 + \tilde{x})^2 + b(x_0 + \tilde{x})(y_0 + \tilde{y}) + c(y_0 + \tilde{y})^2 + r(x_0 + \tilde{x}) + s(y_0 + \tilde{y})^2 \\ &= a\tilde{x}^2 + b\tilde{x}\tilde{y} + c\tilde{y}^2 + [2ax_0 + by_0 + r]\tilde{x} + [2cy_0 + bx_0 + s]\tilde{y} + ax_0^2 + bx_0y_0 + cy_0^2. \end{aligned} \quad (8.27)$$

Queremos que los términos lineales en $\tilde{x}$ y $\tilde{y}$ desaparezcan, así que necesitamos $2ax_0 + by_0 + r = 0$ y $2cy_0 + bx_0 + s = 0$. En forma matricial esto es

$$-\begin{pmatrix} r \\ s \end{pmatrix} = \begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = 2G \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Supongamos que G es invertible. Entonces podemos resolver para x_0 y y_0 y obtenemos $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = -\frac{1}{2}G^{-1}\begin{pmatrix} r \\ s \end{pmatrix}$. Ahora, si ponemos $\tilde{d} = d - ax_0^2 - bx_0y_0 - cy_0^2$, entonces (8.27) se convierte en

$$\tilde{d} = a\tilde{x}^2 + b\tilde{x}\tilde{y} + c\tilde{y}^2 \quad (8.28)$$

que ahora está en la forma de (8.14) (si $\tilde{d}$ es negativo, entonces debemos multiplicar ambos lados de (8.28) por -1 . En este caso, los valores propios de G cambian de signo, por lo tanto D también cambia de signo, pero Q no). Por lo tanto, si ponemos $\vec{x}' = Q^{-1}\vec{\tilde{x}}$, entonces

$$\tilde{d} = \lambda_1 x'^2 + \lambda_2 y'^2$$

y $\vec{x}' = Q^{-1}\vec{\tilde{x}} = Q^{-1}(\vec{x} - \vec{x}_0) = Q^{-1}\vec{x} - Q^{-1}\vec{x}_0 = Q^{-1}\vec{x} + \frac{1}{2}Q^{-1}G^{-1}\begin{pmatrix} r \\ s \end{pmatrix}$. Así, nuevamente vemos que la solución de (8.25) es la solución de $\lambda_1 x^2 + \lambda_2 y^2 = \tilde{d}$, pero rotada por Q y trasladada por el vector $\frac{1}{2}Q^{-1}G^{-1}\begin{pmatrix} r \\ s \end{pmatrix}$.

Ejemplo 8.68. Encuentre las soluciones de

$$10x^2 + 6xy + 2y^2 + 8x - 2y = 4. \quad (8.19')$$

Solución. Sabemos por el Ejemplo 8.65 que

$$G = \begin{pmatrix} 10 & 3 \\ 3 & 2 \end{pmatrix}, \quad Q = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 3 \\ -3 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 11 \end{pmatrix}$$

y que

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = Q^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} x - 3y \\ 3x + y \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} x \\ y \end{pmatrix} = Q \begin{pmatrix} x' \\ y' \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} x' + 3y' \\ -3x' + y' \end{pmatrix}.$$

Método 1. Con la notación anterior, sabemos por el Ejemplo 8.65 que (8.19') es

$$\begin{aligned} 4 &= 10x^2 + 6xy + 2y^2 + 8x - 2y = x'^2 + 11y'^2 + 8x - 2y \\ &= x'^2 + 11y'^2 + \frac{8}{\sqrt{10}}(x' + 3y') - \frac{2}{\sqrt{10}}(-3x' + y') \\ &= x'^2 + \frac{14}{\sqrt{10}}x' + 11y'^2 + \frac{22}{\sqrt{10}}y' \\ &= \left(x' + \frac{7}{\sqrt{10}}\right)^2 + 11\left(y' + \frac{1}{\sqrt{10}}\right)^2 - 6, \end{aligned}$$

por lo tanto

$$\left(x' + \frac{7}{\sqrt{10}}\right)^2 + 11\left(y' + \frac{1}{\sqrt{10}}\right)^2 = 4 + 6 = 10.$$

Esta es una elipse orientada como la del Ejemplo 8.65, pero trasladada $7/\sqrt{10}$ en la dirección x' y $-1/\sqrt{10}$ en la dirección y' . Las longitudes de los semiejes son $\sqrt{10}$ y $\sqrt{\frac{10}{11}}$.

Método 2. Observe que

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = -\frac{1}{2}G^{-1} \begin{pmatrix} r \\ s \end{pmatrix} = -\frac{1}{2} \cdot \frac{1}{11} \begin{pmatrix} 2 & -3 \\ -3 & 10 \end{pmatrix} \begin{pmatrix} 8 \\ -2 \end{pmatrix} = -\frac{1}{22} \begin{pmatrix} 22 \\ -44 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

Pongamos $\tilde{x} = x - x_0 = x + 1$ y $\tilde{y} = y - y_0 = y - 2$. Entonces

$$\begin{aligned} 4 &= 10x^2 + 6xy + 2y^2 + 8x - 2y = 10(\tilde{x} - 1)^2 + 6(\tilde{x} - 1)(\tilde{y} + 2) + 2(\tilde{y} + 2)^2 + 8(\tilde{x} - 1) - 2(\tilde{y} + 2) \\ &= 10\tilde{x}^2 - 20\tilde{x} + 1 + 6\tilde{x}\tilde{y} + 12\tilde{x} - 6\tilde{y} - 12 + 2\tilde{y}^2 + 8\tilde{y} + 8 + 8\tilde{x} - 8 - 2\tilde{y} - 4 \\ &= 10\tilde{x}^2 + 6\tilde{x}\tilde{y} + 2\tilde{y}^2 - 15 \end{aligned}$$

por lo tanto

$$19 = 10\tilde{x}^2 + 6\tilde{x}\tilde{y} + 2\tilde{y}^2 = \tilde{x}'^2 + 11\tilde{y}'^2$$

con

$$\begin{pmatrix} \tilde{x}' \\ \tilde{y}' \end{pmatrix} = Q^{-1} \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} \tilde{x} + 3\tilde{y} \\ 3\tilde{x} - \tilde{y} \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} (x+1) + 3(y-2) \\ 3(x+1) - (y-2) \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} x + 3y - 5 \\ 3x - y + 5 \end{pmatrix}.$$

◇

Ahora debería haber entendido

- que una matriz simétrica 2×2 que no es un múltiplo de la identidad determina dos direcciones distinguidas en $\mathbb{R}^2$, a saber, las paralelas a sus vectores propios,
- por qué un cambio de variables es útil para encontrar soluciones de una ecuación cuadrática en dos variables,
- etc.

Ahora debería ser capaz de

- encontrar las soluciones de ecuaciones cuadráticas en dos variables,
- hacer un cambio de coordenadas tal que la ecuación transformada no tenga término mixto,
- bosquejar la solución en el plano xy ,
- etc.

Ejercicios.

1. Encuentre una substitución ortogonal que diagonalice las formas cuadráticas dadas y encuentre la forma diagonal. Haga un bosquejo de las soluciones. Si es un elipse, calcule las longitudes de los ejes principales y el ángulo que tienen con el eje x . Si es una hipérbola, calcule en ángulo que tiene las asíntotas con el eje x .
 - (a) $10x^2 - 6xy + 2y^2 = 4$,
 - (b) $x^2 - 9y^2 = 2$,
 - (c) $x^2 - 9y^2 = 20$ (compare la solución con la del literal anterior!)
 - (d) $11x^2 - 16xy - y^2 = 30$.
 - (e) $x^2 + 4xy + 4y^2 = 4$.
 - (f) $xy = 1$.
 - (g) $5x^2 - 2xy + 5y^2 = -4$.
 - (h) $x^2 - 2xy + 4y^2 = 0$.
2. Encuentre la fórmula de una elipse cuyos semiejes tienen magnitudes $\frac{1}{2}$ y $\frac{1}{3}$ y cuyos ejes principales son paralelos a $\begin{pmatrix} 1 & 2 \end{pmatrix}$ y $\begin{pmatrix} -2 & 1 \end{pmatrix}$.
3. Encuentre la fórmula de una elipse cuyos semiejes tienen magnitudes 3 y 1 y cuyo primer eje principal tiene un ángulo de 30° con el eje x .

8.7. Resumen

$\mathbb{C}^n$ como espacio con producto interno

$\mathbb{C}^n$ es un espacio con producto interno si definimos

$$\langle \vec{z}, \vec{w} \rangle = \sum_{j=1}^n z_j \overline{w_j}.$$

Para todo $\vec{v}, \vec{w}, \vec{z} \in \mathbb{C}^n$ y $c \in \mathbb{C}$ tenemos:

- $\langle \vec{v}, \vec{z} \rangle = \overline{\langle \vec{z}, \vec{v} \rangle},$
- $\langle \vec{v} + c\vec{w}, \vec{z} \rangle = \langle \vec{v}, \vec{z} \rangle + c\langle \vec{w}, \vec{z} \rangle, \quad \langle \vec{z}, \vec{v} + c\vec{w} \rangle = \langle \vec{z}, \vec{v} \rangle + \overline{c}\langle \vec{z}, \vec{w} \rangle,$
- $\langle \vec{z}, \vec{z} \rangle = \|\vec{z}\|^2,$
- $\langle \vec{v}, \vec{z} \rangle \leq \|\vec{v}\| \|\vec{z}\|,$
- $\|\vec{v} + \vec{z}\|^2 \leq \|\vec{v}\|^2 + \|\vec{z}\|^2.$

La adjunta de una matriz $A \in M_{\mathbb{C}}(n \times n)$ es $A^* = \overline{(A^t)} = (\overline{A})^t$ (= transpuesta y conjugada compleja). La matriz A se llama *hermítica* si $A^* = A$. La matriz Q se llama *unitaria* si es invertible y $Q^* = Q^{-1}$.

Observe que $\det A^* = \overline{\det A}$.

Valores propios y vectores propios

Definición. Sea $A \in M(n \times n)$. Entonces λ se llama un *valor propio* de A con *vector propio* $\vec{v}$ si $\vec{v} \neq \vec{0}$ y $A\vec{v} = \lambda\vec{v}$. El conjunto de todas las soluciones de $A\vec{v} = \lambda\vec{v}$ para un valor propio λ se llama el espacio propio de A asociado a λ . Se denota por $\text{Eig}_{\lambda}(A)$.

Los valores propios de A son exactamente los ceros del *polinomio característico*

$$p_A(\lambda) = \det(A - \lambda).$$

Es un polinomio de grado n . Como todo polinomio de grado ≥ 1 tiene al menos una raíz compleja, toda matriz compleja tiene al menos un valor propio (pero hay matrices reales sin valores propios). Además, una matriz $n \times n$ tiene a lo sumo n valores propios. Si factorizamos p_A , obtenemos

$$p_A(\lambda) = (\lambda - \mu_1)^{m_1} \cdots (\lambda - \mu_k)^{m_k}$$

donde $\mu_1, \dots, \mu_k$ son los valores propios *distintos* de A . El exponente m_j se llama *multiplicidad algebraica* de μ_j . La *multiplicidad geométrica* de μ_j es $\dim(\text{Eig}_{\mu_j}(A))$. Observe que

- multiplicidad geométrica $\leq$ multiplicidad algebraica,
- la suma de todas las multiplicidades algebraicas es $m_1 + \cdots + m_k = n$.

Matrices semejantes.

- Dos matrices $A, B \in M(n \times n)$ se llaman *semejantes* si existe una matriz invertible C tal que $A = C^{-1}BC$.

- Una matriz A se llama *diagonalizable* si es semejante a una matriz diagonal.

Caracterización de la diagonalizabilidad. Sea $A \in M_{\mathbb{C}}(n \times n)$ y sean $\mu_1, \dots, \mu_k$ los valores propios distintos de A . Definimos $n_j = \dim(\text{Eig}_{\mu_j}(A)) =$ multiplicidad geométrica de μ_j y $m_j =$ multiplicidad algebraica de μ_j . Entonces las siguientes condiciones son equivalentes:

- (I) A es diagonalizable.
- (II) $\mathbb{C}^n$ tiene una base formada por vectores propios de A .
- (III) $\mathbb{C}^n = \text{Eig}_{\mu_1}(A) \oplus \dots \oplus \text{Eig}_{\mu_k}(A)$.
- (IV) $n_j = m_j$ para todo $j = 1, \dots, k$.
- (V) $n_1 + \dots + n_k = n$.

Lo mismo es cierto para matrices simétricas reemplazando $\mathbb{C}^n$ por $\mathbb{R}^n$.

Propiedades de matrices unitarias. Sea Q una matriz unitaria $n \times n$. Entonces:

- $|\det Q| = 1$,
- Si $\lambda \in \mathbb{C}$ es un valor propio de Q , entonces $|\lambda| = 1$.
- Q es diagonalizable unitariamente (no demostramos este hecho), por lo tanto $\mathbb{C}^n$ tiene una base formada por vectores propios de Q . Pueden elegirse mutuamente ortogonales.

Además, $Q \in M(n \times n)$ es unitaria si y solo si $\|Q\vec{z}\| = \|\vec{z}\|$ para todo $\vec{z} \in \mathbb{C}^n$.

Propiedades de matrices hermíticas. Sea $A \in M_{\mathbb{C}}(n \times n)$ una matriz hermítica $n \times n$. Entonces:

- $\det A \in \mathbb{R}$,
- Si λ es un valor propio de A , entonces $\lambda \in \mathbb{R}$.
- A es diagonalizable unitariamente por lo tanto $\mathbb{C}^n$ tiene una base formada por vectores propios de A . Pueden elegirse mutuamente ortogonales.

Además, $A \in M(n \times n)$ es hermítica si y solo si $\langle A\vec{v}, \vec{z} \rangle = \langle \vec{v}, A\vec{z} \rangle$ para todo $\vec{v}, \vec{z} \in \mathbb{C}^n$.

Propiedades de matrices simétricas. Sea $A \in M_{\mathbb{R}}(n \times n)$ una matriz simétrica $n \times n$. Entonces:

- A es diagonalizable ortogonalmente. por lo tanto $\mathbb{R}^n$ tiene una base formada por vectores propios de A . Pueden elegirse mutuamente ortogonales.

Además, A es simétrica si y solo si $\langle A\vec{v}, \vec{z} \rangle = \langle \vec{v}, A\vec{z} \rangle$ para todo $\vec{v}, \vec{z} \in \mathbb{R}^n$.

Solución de $ax^2 + bxy + cy^2 = d$. La ecuación puede reescribirse como $\langle G\vec{x}, \vec{x} \rangle = d$ con la matriz simétrica

$$G = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}.$$

Sean λ_1, λ_2 los valores propios de G y supongamos que $d \geq 0$. Entonces las soluciones son:

- una **elipse** si $\det G > 0$, más precisamente,
 - una elipse con longitudes de sus ejes $\sqrt{d/\lambda_1}$ y $\sqrt{d/\lambda_2}$ si $\lambda_1, \lambda_2 > 0$ y $d > 0$,
 - el punto $(0, 0)$ si $d = 0$,
 - el conjunto vacío si $\lambda_1, \lambda_2 < 0$ y $d > 0$,
- **hipérbola** si $\det G < 0$, más precisamente,
 - una hipérbola si $d > 0$,
 - dos rectas que se cruzan en el origen si $d = 0$,
- dos rectas paralelas, una recta o $\mathbb{R}^2$ si $\det G = 0$.

8.8. Ejercicios

1. Sea Q una matriz unitaria. Demuestre que todos sus autovalores tienen norma 1.
2. Muestre que A y B son semejantes si y solo si A^t, B^t son semejantes.
3. Encuentre todas las matrices que son semejantes a la identidad.
4. Son las matrices $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ y $B = \begin{pmatrix} 5 & 1 \\ 0 & 1 \end{pmatrix}$ semejantes? *Sugerencia.* Ejercicio 5.(c).
5. Sea A una matriz con autovalores $\mu_1, \dots, \mu_k$ y sea c una constante.
 - (a) ¿Qué se puede decir sobre los autovalores de cA ? ¿Qué se puede decir sobre los autovalores de $A + cid$?
6. Dados la matriz A y los vectores u y w :

$$A = \begin{pmatrix} 25 & 15 & -18 \\ -30 & -20 & 36 \\ -6 & -6 & 16 \end{pmatrix}, \quad u = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad w = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$
 - (a) Diga si los vectores u y w son autovectores de A . Si lo son, cuáles son los vectores propios correspondientes?
 - (b) Puede usar que $\det(A - \lambda) = -\lambda^3 + 21\lambda^2 - 138\lambda + 280$. Calcule todos los autovalores de A .
7. Encuentre el polinomio característico y los valores y espacios propios de la matriz

$$A = \begin{pmatrix} 1 & -1 & 2 & 1 \\ 2 & 4 & 1 & -1 \\ 3 & -3 & 6 & 3 \\ 4 & 2 & 5 & 1 \end{pmatrix},$$

8. Considere la matriz

$$A = \begin{pmatrix} 1 & -3 & 2 & 1 \\ 2 & 2 & 1 & -1 \\ 3 & -3 & 6 & 3 \\ 4 & 2 & 5 & 1 \end{pmatrix},$$

- (a) Sin calcular el polinomio característico, demuestre que $\lambda = 1$ es un valor propio y encuentre su espacio propio.
 (b) Encuentre el polinomio característico y todos los valores y espacios propios de la matriz A .

9. Sea $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. Calcule $e^A := \sum_{n=0}^{\infty} \frac{1}{n!} A^n$.

Sugerencia. Encuentre una matriz invertible C y una matriz diagonal D tal que $A = C^{-1}DC$ y use esto para calcular A^n .

10. (a) Sea $\Phi : M(2 \times 2, \mathbb{R}) \rightarrow M(2 \times 2, \mathbb{R})$, $\Phi(A) = A^t$. Encuentre los valores propios y los espacios propios de Φ .
 (b) Sea P_2 el espacio vectorial de polinomios de grado menor o igual a 2 con coeficientes reales. Encuentre los valores propios y los espacios propios de $T : P_2 \rightarrow P_2$, $Tp = p' + 3p$.
 (c) Sea R la reflexión en el plano $P : x + 2y + 3z = 0$ en $\mathbb{R}^3$. Calcule los valores propios y los espacios propios de R .
11. Consideramos una cuerda de longitud L que está fija en ambos extremos. Al ser excitada, su elongación vertical satisface la ecuación diferencial parcial $\frac{\partial^2}{\partial t^2} u(t, x) = \frac{\partial^2}{\partial x^2} u(t, x)$. Si hacemos el ansatz $u(t, x) = e^{i\omega t} v(x)$ para algún número ω y una función v que depende solo de x , obtenemos $-\omega^2 v = v''$. Si ponemos $\lambda = -\omega^2$, vemos que debemos resolver el siguiente problema de valores propios:

$$T : V \rightarrow V, \quad Tv = v''$$

con

$$V = \{f : [0, L] \rightarrow \mathbb{R}, f \text{ es dos veces diferenciable y } f(0) = f(L) = 0\}.$$

- (a) Muestre que V es un espacio vectorial.
 (b) Muestre que T es un operador lineal bien definido.
 (c) Encuentre los valores propios y los espacios propios de T .
12. Encuentre los valores propios y los espacios propios de las siguientes matrices $n \times n$:

$$A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & \cdots & 1 & 2 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & 1 & \cdots & 1 & n \end{pmatrix}.$$

Compare con el Ejercicio 11..

13. Sea $A \in M(n \times n, \mathbb{C})$ una matriz hermíticaa tal que todos sus autovalores son estrictamente mayores a 0. Sea $\langle \cdot, \cdot \rangle$ el producto interno estandar en $\mathbb{C}^n$. Demuestre que A induce un producto interno en $\mathbb{C}^n$ a través de

$$\mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}, \quad (x, y) := \langle Ax, y \rangle.$$

DRAFT

Apéndice A

Números complejos

Un *número complejo* es una expresión de la forma

$$a + ib$$

donde $a, b \in \mathbb{R}$ y i se llama la *unidad imaginaria*. El número a se llama la *parte real* de z , denotada por $\operatorname{Re}(z)$ y b se llama la *parte imaginaria* de z , denotada por $\operatorname{Im}(z)$.

El conjunto de todos los números complejos a veces se llama el *plano complejo* y se denota por $\mathbb{C}$:

$$\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}.$$

Un número complejo puede visualizarse como un punto en el plano $\mathbb{R}^2$, donde a es la coordenada sobre el *eje real* y b es la coordenada sobre el *eje imaginario*.

Sean $a, b, x, y \in \mathbb{R}$. Definimos las operaciones algebraicas de suma y producto para números complejos $z = a + ib$, $w = x + iy$:

$$\begin{aligned} z + w &= (a + ib) + (x + iy) := a + x + i(b + y), \\ zw &= (a + ib)(x + iy) := ax - by + i(ay + bx). \end{aligned}$$

Ejercicio A.1. Muestre que si identificamos el número complejo $z = a + ib$ con el vector $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$, entonces la suma de números complejos es la misma que la suma de vectores en $\mathbb{R}^2$.

Daremos una interpretación geométrica de la multiplicación de números complejos más adelante, después de la fórmula (A.5).

De la definición anterior se sigue que $i^2 = -1$. Además, podemos ver los números reales $\mathbb{R}$ como un subconjunto de $\mathbb{C}$ si identificamos un número real x con el número complejo $x + 0i$.

Sean $a, b \in \mathbb{R}$ y $z = a + ib$. Entonces el *conjugado complejo* de z es

$$\bar{z} = a - ib$$

y su *módulo* o *norma* es

$$|z| = \sqrt{a^2 + b^2}.$$

Geoméricamente, el conjugado complejo se obtiene de z mediante una reflexión respecto al eje x , y su norma es la distancia del punto representado por z al origen del plano complejo.

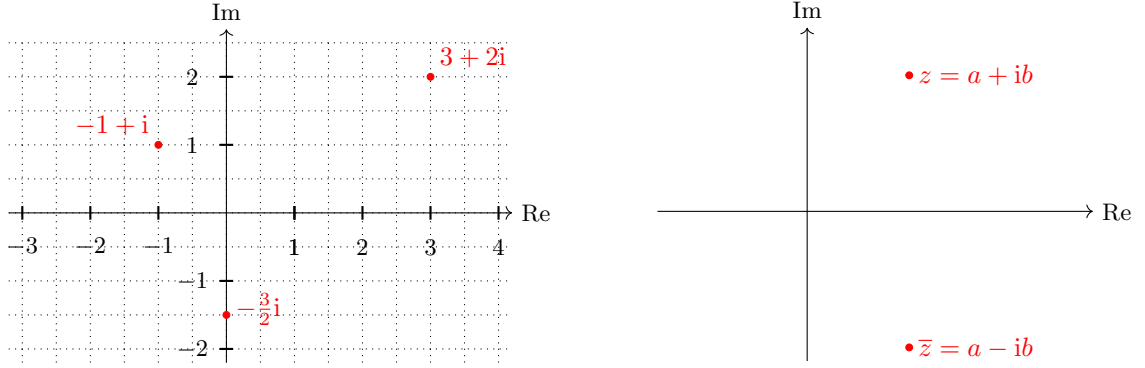


FIGURA A.1: Plano complejo.

Propiedades A.2. Sean $a, b, x, y \in \mathbb{R}$ y sean $z = a + ib$, $w = x + iy$. Entonces:

- (I) $z = \operatorname{Re} z + i \operatorname{Im} z$.
- (II) $\operatorname{Re}(z + w) = \operatorname{Re}(z) + \operatorname{Re}(w)$, $\operatorname{Im}(z + w) = \operatorname{Im}(z) + \operatorname{Im}(w)$.
- (III) $\overline{\overline{z}} = z$, $\overline{z + w} = \overline{z} + \overline{w}$, $\overline{zw} = \overline{z} \overline{w}$.
- (IV) $z\overline{z} = |z|^2$.
- (V) $\operatorname{Re} z = \frac{1}{2}(z + \overline{z})$, $\operatorname{Im} z = \frac{1}{2i}(z - \overline{z})$.

Demostración. (I) y (II) deberían ser claros. Para (III), observe que $\overline{\overline{z}} = \overline{a - ib} = a + ib$,

$$\begin{aligned}\overline{z + w} &= \overline{a + x + i(b + y)} = a + x - i(b + y) = a - ib + x - iy = \overline{a + ib} + \overline{x + iy} = \overline{z} + \overline{w}, \\ \overline{zw} &= \overline{ax - by + i(ay + bx)} = ax - by + i(ay + bx) = (a - ib)(x - iy) = (\overline{a + ib})(\overline{x + iy}) = \overline{z} \overline{w}.\end{aligned}$$

(IV) se sigue de

$$z\overline{z} = (a + ib)(\overline{a + ib}) = (a + ib)(a - ib) = a^2 + b^2 + i(ab - ba) = a^2 + b^2 = |z|^2$$

and (V) se sigue de

$$\begin{aligned}z + \overline{z} &= a + ib + (\overline{a + ib}) = a + ib + a - ib = 2a = 2\operatorname{Re}(z), \\ z - \overline{z} &= a + ib - (\overline{a + ib}) = a + ib - (a - ib) = 2ib = 2i\operatorname{Im}(z).\end{aligned}$$

□

Llamamos *real* a un número complejo si es de la forma $z = a + i0$ para algún $a \in \mathbb{R}$, y lo llamamos *puramente imaginario* si es de la forma $z = 0 + ib$ para algún $b \in \mathbb{R}$. Por lo tanto

$$\begin{aligned}z \text{ es real} &\iff z = \overline{z} &\iff z = \operatorname{Re}(z) \\ z \text{ es puramente imaginario} &\iff z = -\overline{z} &\iff z = i\operatorname{Im}(z).\end{aligned}$$

Resulta que $\mathbb{C}$ es un *campo*, es decir, satisface

- (a) **Asociatividad de la suma:** $(u + v) + w = u + (v + w)$ para todo $u, v, w \in \mathbb{C}$.
- (b) **Conmutatividad de la suma:** $v + w = w + v$ para todo $u, v \in \mathbb{C}$.
- (c) **Elemento identidad de la suma:** Existe un elemento 0, llamado *identidad aditiva*, tal que para todo $v \in \mathbb{C}$ tenemos $0 + v = v + 0 = v$.
- (d) **Inverso aditivo:** Para todo $z \in \mathbb{C}$, tenemos un elemento inverso $-z$ tal que $z + (-z) = 0$.
- (e) **Asociatividad de la multiplicación** $(uv)w = u(vw)$ para todo $u, v, w \in \mathbb{C}$.
- (f) **Conmutatividad de la multiplicación** $vw = wv$ para todo $u, v \in \mathbb{C}$.
- (g) **Elemento identidad de la suma:** Existe un elemento 1, llamado *identidad multiplicativa*, tal que para todo $v \in \mathbb{C}$ tenemos $1 \cdot v = v + 1 = v$.
- (h) **Inverso multiplicativo:** Para todo $z \in \mathbb{C} \setminus \{0\}$, tenemos un elemento inverso z^{-1} tal que $z \cdot z^{-1} = 1$.
- (i) **Leyes distributivas:** Para todo $u, v, w \in \mathbb{C}$ tenemos

$$u(w + v) = uw + uv.$$

Es fácil verificar que se cumplen la conmutatividad, la asociatividad y la distributividad. Claramente, la identidad aditiva es $0 + i0$ y la identidad multiplicativa es $1 + 0i$. Si $z = a + ib$, entonces su inverso aditivo es $-a - ib$. Si $z \in \mathbb{C} \setminus \{0\}$, entonces $z^{-1} = \frac{\bar{z}}{|z|^2} = \frac{a-ib}{a^2+b^2}$. Esto puede verse fácilmente si recordamos que $|z|^2 = z\bar{z}$.

La demostración del siguiente teorema está fuera del alcance de estas notas.

Teorema A.3 (Teorema fundamental del álgebra). *Todo polinomio complejo no constante tiene al menos una raíz compleja.*

Obtenemos inmediatamente el siguiente corolario.

Corolario A.4. *Todo polinomio complejo p puede escribirse en la forma*

$$p(z) = c(z - \lambda_1)^{n_1} \cdot (z - \lambda_2)^{n_2} \cdots (z - \lambda_k)^{n_k} \quad (\text{A.1})$$

donde $\lambda_1, \dots, \lambda_k$ son las raíces distintas de p . Observe que $n_1 + \cdots + n_k = \deg(p)$.

Los enteros $n_1, \dots, n_k$ se llaman la *multiplicidad* de la raíz correspondiente.

Demostración. Sea $n = \deg(p)$. Si $n = 0$, entonces p es constante y claramente es de la forma (A.1). Si $n > 0$, entonces, por el Teorema A.3, existe $\mu_1 \in \mathbb{C}$ tal que $p(\mu_1) = 0$. Por lo tanto existe algún polinomio q_1 tal que $p(z) = (z - \mu_1)q_1(z)$. Claramente, $\deg(q_1) = n - 1$. Si q_1 es constante, terminamos. Si q_1 no es constante, entonces debe tener un cero μ_2 . Por lo tanto $q_1(z) = (z - \mu_2)q_2(z)$ con algún polinomio q_2 con $\deg(q_2) = n - 2$. Si repetimos este proceso n veces, finalmente obtenemos que

$$p(z) = c(z - \mu_1)(z - \mu_2) \cdots (z - \mu_n).$$

Ahora solo tenemos que agrupar todos los términos con el mismo μ_j y obtenemos la forma (A.1). $\square$

Funciones de números complejos

Es más o menos claro cómo formar un polinomio complejo. También podemos extender a los números complejos las funciones que admiten una representación mediante series de potencias. Para ello, recordemos (de algún curso de cálculo) que una serie de potencias es una expresión de la forma

$$\sum_{n=0}^{\infty} c_n (z - a)^n \quad (\text{A.2})$$

donde los c_n son los coeficientes y a es el punto donde la serie de potencias está centrada. En nuestro caso, son números complejos y z es un número complejo. Recordemos que una serie $\sum_{n=0}^{\infty} a_n$ se llama *absolutamente convergente* si y solo si $\sum_{n=0}^{\infty} |a_n|$ es convergente. Puede demostrarse que toda serie absolutamente convergente de números complejos es convergente. Además, para toda serie de potencias de la forma (A.2) existe un número $R > 0$ o $R = \infty$, llamado el *radio de convergencia* tal que la serie converge absolutamente para todo $z \in \mathbb{C}$ con $|z - a| < R$ y diverge para z con $|z - a| > R$. Esto significa que la serie converge absolutamente para todo z en el disco abierto de radio R centrado en a , y diverge fuera del disco cerrado de radio R centrado en a . Para z en la frontera, la serie puede converger o divergir. Observe que $R = 0$ and $R = \infty$ are allowed. Si $R = 0$, entonces la serie converge solo para $z = a$, y si $R = \infty$, entonces la serie converge para todo $z \in \mathbb{C}$.

Funciones importantes que conocemos de los números reales y que tienen una serie de potencias son el seno, el coseno y la función exponencial. Podemos usar su representación mediante series de potencias para definir las también para números complejos.

Definición A.5. Sea $z \in \mathbb{C}$. Entonces definimos

$$\sin z = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1}, \quad \cos z = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n}, \quad e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n. \quad (\text{A.3})$$

Observe que para todo z las series en (A.3) son absolutamente convergentes porque, por ejemplo, para la serie de la función seno tenemos $\sum_{n=0}^{\infty} \left| \frac{(-1)^n}{(2n+1)!} z^{2n+1} \right| = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} |z|^{2n+1}$ es convergente porque $|z|$ es un número real y sabemos que la serie del coseno es absolutamente convergente para todo argumento real. Por lo tanto, la serie del seno es absolutamente convergente para todo $z \in \mathbb{C}$, y por lo tanto converge. El mismo argumento muestra que las series del coseno y de la exponencial son convergentes para todo $z \in \mathbb{C}$.

Observación A.6. Como la serie de la función seno contiene solo potencias impares de z , esta es una función impar y el coseno es una función par porque contiene solo potencias pares de z . En fórmulas:

$$\sin(-z) = -\sin z, \quad \cos(-z) = \cos z.$$

A continuación mostramos la relación entre las funciones trigonométricas y la función exponencial.

Teorema A.7 (Fórmulas de Euler). Para todo $z \in \mathbb{C}$ tenemos que

$$\begin{aligned} e^{iz} &= \cos z + i \sin z, \\ \cos(z) &= \frac{1}{2}(e^{iz} + e^{-iz}), \\ \sin(z) &= \frac{1}{2i}(e^{iz} - e^{-iz}). \end{aligned}$$

Demostración. Mostremos la fórmula para e^{iz} . En el cálculo usaremos que $i^{2n} = (i^2)^n = (-1)^n$ and $i^{2n+1} = (i^2)^n i = (-1)^n i$ and

$$\begin{aligned} e^{iz} &= \sum_{n=0}^{\infty} \frac{1}{n!} (iz)^n = \sum_{n=0}^{\infty} \frac{1}{n!} i^n z^n = \sum_{n=0}^{\infty} \frac{1}{(2n)!} i^{(2n)} z^{2n} + \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} i^{(2n+1)} z^{2n+1} \\ &= \sum_{n=0}^{\infty} \frac{1}{(2n)!} (-1)^n z^{2n} + \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} i(-1)^n z^{2n+1} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} + i \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} \\ &= \cos z + i \sin z. \end{aligned}$$

Observe que el tercer paso requiere una justificación adecuada (véase algún curso de cálculo integral). Para la demostración de la fórmula para $\cos z$, observamos que de lo que acabamos de probar se sigue que

$$\begin{aligned} \frac{1}{2}(e^{iz} + e^{-iz}) &= \frac{1}{2}(\cos(z) + i \sin(z) + \cos(-z) + i \sin(-z)) = \frac{1}{2}(\cos(z) + i \sin(z) + \cos(z) - i \sin(z)) \\ &= \cos(z). \end{aligned}$$

La fórmula para la función seno se sigue análogamente. □

Ejercicio. Sean $z, w \in \mathbb{C}$. Muestre lo siguiente.

- (I) $e^z e^w = e^{z+w}$. *Sugerencia.* Utilice el producto de Cauchy.
- (II) Utilice las fórmulas de Euler para probar $\cos \alpha \cos \beta = \frac{1}{2}(\cos(\alpha - \beta) + \cos(\alpha + \beta))$, $\sin \alpha \sin \beta = \frac{1}{2}(\cos(\alpha - \beta) - \cos(\alpha + \beta))$, $\sin \alpha \cos \beta = \frac{1}{2}(\sin(\alpha + \beta) + \sin(\alpha - \beta))$.
- (III) $(\cos z)^2 + (\sin z)^2 = 1$.
- (IV) $\cosh(z) = \cos(iz)$, $\sinh(z) = -i \sin(iz)$. En particular, $\sin$ and $\cos$ **no** son funciones acotadas en $\mathbb{C}$.
- (V) Muestre que la función exponencial es $2\pi i$ periódica.

Representación polar de números complejos

Sea $z \in \mathbb{C}$ con $|z| = 1$ y sea φ el ángulo entre el eje real positivo y la recta que conecta el origen y z . Se llama el *argumento* de z , y se denota por $\arg(z)$. Observe que el argumento solo está determinado módulo 2π . Esto significa que, si sumamos o restamos cualquier múltiplo entero de 2π al argumento, obtenemos otro argumento válido.

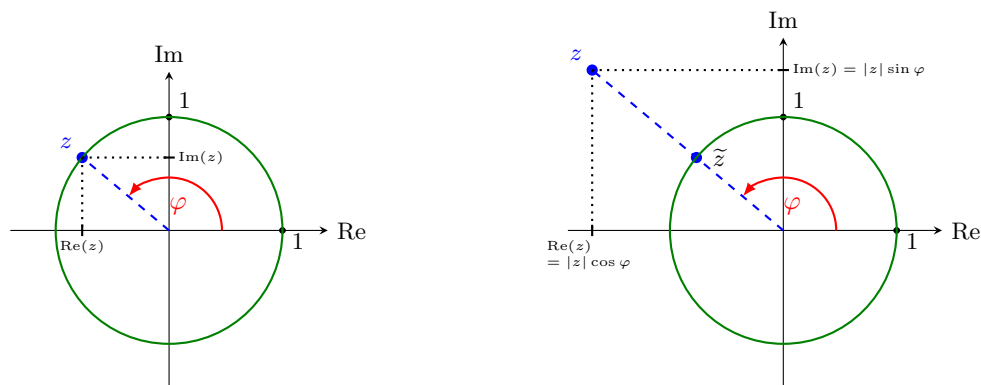


FIGURA A.2: Imagen izquierda: Si $|z| = 1$, entonces $z = \cos \varphi + i \sin \varphi = e^{i\varphi}$.
 Imagen derecha: Si $z \neq 0$, entonces $z = |z| \cos \varphi + i |z| \sin \varphi = |z| e^{i\varphi}$.

Entonces el real and imaginary part of z are $\operatorname{Re}(z) = \cos \varphi$ and $\operatorname{Im}(z) = \sin \varphi$, y por lo tanto $z = \cos \varphi + i \sin \varphi = e^{i\varphi}$. Vimos en la Observación 2.5 cómo podemos calcular el argumento de un número complejo.

Ahora sea $z \in \mathbb{C} \setminus \{0\}$ y nuevamente sea φ el ángulo entre el eje real positivo y la recta que conecta el origen con z . Sea $\tilde{z} = \frac{z}{|z|}$. Entonces $|\tilde{z}| = 1$ y por lo tanto $\tilde{z} = e^{i\varphi}$. Se sigue que

$$z = |z| e^{i\varphi}. \quad (\text{A.4})$$

(A.4) se llama la *representación polar* de z .

Ahora podemos dar una interpretación geométrica del producto de dos números complejos. Sean $z, w \in \mathbb{C} \setminus \{0\}$ y sean $\alpha = \arg z$ y $\beta = \arg w$. Entonces

$$zw = |z| e^{i\alpha} |w| e^{i\beta} = |z| |w| e^{i(\alpha+\beta)}. \quad (\text{A.5})$$

Esto muestra que el producto zw es el número complejo cuya norma es el producto de las normas de z y w , y cuyo argumento es la suma de los argumentos de z y w .

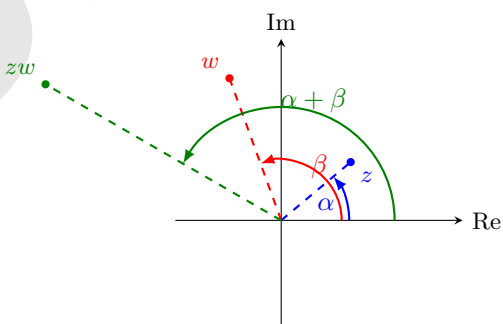


FIGURA A.3: Interpretación geométrica de la multiplicación de dos números complejos.

Apéndice B

Soluciones

Soluciones de algunos ejercicios del Capítulo 1

Soluciones de Sección 1.1

- (a) No tiene solución.
- (b) Infinitas soluciones.
- (c) Solución única $x = 0, y = 0$.
- (d) Infinitas soluciones.
- (e) No tiene solución.
- (f) Infinitas soluciones.

Soluciones de Sección 1.2

- 1.(a) Solución única. $x = \frac{1}{2}, y = -30$.
- 1.(b) Solución única. $x = \frac{9}{7}, y = \frac{2}{7}$.
- 1.(c) Infinitas soluciones.
- 1.(d) Ninguna solución.
- 1.(e) Infinitas soluciones.
- 1.(f) Solución única. $x = \frac{7}{5}, y = 2$.
- 2. $k \in \mathbb{R} \setminus \{19\}$.
- 3. $k \in \mathbb{R} \setminus \{8 - \sqrt{66}, 8 + \sqrt{66}\}$.

Soluciones de Sección 1.4

- 1. 6.
- 2. $a = 3, b = 1, c = 1$.
- 3. Si $y = ax^2 + bx + c$ es tal parábola, entonces $a + c = \frac{5}{2}$ y $b = -\frac{3}{2}$.
- 4. Las opciones posibles son:

- $t \neq \frac{1}{2}, k = 10$.

- $t = \frac{1}{2}, k \neq 5$.

- $t = \frac{1}{4}, k = 10$.

5. 065, 164, 263, 362, 461, 560.

6. 10 clientes eran dueños de perros y 12 de gatos.

7. 75.

8. Si la velocidad del conductor A es de $57 \frac{km}{h}$ y la del conductor B es de $49 \frac{km}{h}$.

- El conductor A llega a Villavicencio a las 6am.
- Los dos conductores no se encuentran en carretera.

Si la velocidad del conductor A es de $19 \frac{km}{h}$ y la del conductor B es de $70 \frac{km}{h}$.

- El conductor A llega a Villavicencio a las 10am.
- Los dos conductores se encuentran en carretera a las 9am.

Soluciones de algunos ejercicios del Capítulo 2

Soluciones de Sección 2.1

- 1.(a) $\overrightarrow{PQ} = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$.
- 1.(b) $\sqrt{10}$.
- 1.(c) $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$.
- 1.(d) $-\arctan \frac{2}{3}$.
- 1.(e) $\pi - \arctan \frac{1}{3}$.
- 2.(a) Sí forman un paralelogramo.
- 2.(b) No forman un paralelogramo.
- 2.(c) Sí forman un paralelogramo.

Soluciones de Sección 2.2

- 3.(a) Note que hay tres maneras cómo acercarse a este problema: dos vectores $\vec{a}$, $\vec{b}$ son paralelos
 - 1) si y solo si existen escalares μ, λ tal que $\mu\vec{a} = \lambda\vec{b}$;
 - 2) si y solo si $\langle \vec{a}, \vec{b} \rangle = \|\vec{a}\| \|\vec{b}\|$;
 - 3) si y solo si las rectas que generan son paralelas, es decir, no se intersecan en exactamente un punto, lo que se puede verificar con el determinante.
- 4.(a)(I) $\alpha = -2$.
- 4.(a)(II) $\alpha = 2$.
- 4.(a)(III) $\alpha = -2(2 \pm \sqrt{3})$.
- 4.(a)(IV) $\alpha = 0$.
- 4.(a)(V) No existe tal α .
- 4.(b) Cuando $\alpha \rightarrow \infty$ el ángulo entre $\vec{a}$ y $\vec{b}$ tiende a $\frac{3\pi}{4}$. Cuando $\alpha \rightarrow -\infty$ el ángulo entre $\vec{a}$ y $\vec{b}$ tiende a $\frac{\pi}{4}$.

Soluciones de Sección 2.3

- 1.(a) $\text{proy}_{\vec{a}} \vec{b} = \frac{11}{29} \vec{b}$, $\text{proy}_{\vec{b}} \vec{a} = \frac{11}{10} \vec{a}$.
- 1.(b) Todos los $\vec{v}$ que son perpendiculares a $\vec{a}$, es decir todo los vectores de la forma $\vec{v} = t \begin{pmatrix} -3 \\ 1 \end{pmatrix}$.
- 1.(c) Todos los $\vec{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ tales que $|x + 3y| = 2\sqrt{10}$. Note que son todos los vectores de la forma $\vec{v} = \frac{2}{\|\vec{a}\|} \vec{a} + t \begin{pmatrix} -3 \\ 1 \end{pmatrix}$. Observe la relación con los vectores de (ii). ¿Por qué es así?
- 1.(d) No. Sí. No.

Soluciones de Sección 2.4

- 1.(a) $\begin{pmatrix} 8 \\ 16 \\ 15 \\ 15 \end{pmatrix}$.
- 1.(b) $\sqrt{210}$.
- 1.(c) -50 .
- 1.(d) $\frac{1}{6} \vec{b}$.

Soluciones de Sección 2.5

- 1.(a) $2\sqrt{26}$.
- 1.(b) $\sqrt{26}$.
- 1.(c) Todos los $P(x, y, z)$ tales que $\|\vec{BC} \times \vec{BP}\| = 13$, es decir, $\left\| \begin{pmatrix} -7+9y-5z \\ 6-9x+3z \\ -1+5x-3y \end{pmatrix} \right\| = 13$. Por ejemplo, $P((41 \pm \sqrt{7117})/106, 0, 0)$ sirve.
2. 188.
3. $\sqrt{\frac{5}{29}}$.
4. Todos los $\vec{a}$ que son paralelos a $\begin{pmatrix} 3 \\ 7 \\ 2 \end{pmatrix}$. Los que tienen norma 1 son $\frac{1}{\sqrt{62}} \begin{pmatrix} 3 \\ 7 \\ 2 \end{pmatrix}$ y $-\frac{1}{\sqrt{62}} \begin{pmatrix} 3 \\ 7 \\ 2 \end{pmatrix}$.

Soluciones de Sección 2.6

- 2.(a)(I) No son paralelas.
- 2.(a)(II) No tienen ningún punto en común.
- 2.(a)(III) P pertenece a L_1 , P no pertenece a L_2 .
- 2.(a)(IV) $\frac{x-5}{2} = \frac{y-2}{3} = \frac{z-11}{4}$.
- 2.(b)(I) No son paralelas.
- 2.(b)(II) Se cruzan en exactamente un punto, a saber en $(4, 5, -1)$.
- 2.(b)(III) P pertenece a L_1 , P no pertenece a L_2 .
- 4.4.a(I) $x - y - z = 0$.
- 4.4.a(II) P no pertenece a E .
- 4.4.a(III) $\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + t \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} : t \in \mathbb{R} \right\}$.
- 4.4.b(I) $3x + 3y - 4z = -1$.
- 4.4.b(II) P no pertenece a E .
- 4.4.b(III) $\frac{x-1}{3} = \frac{y}{3} = -\frac{z+1}{4}$.
- 4.4.c(I) $3x + 2y + z = 4$.

4.4.c(II) P no pertenece a E .

4.4.c(III) $x = 3t + 1$, $y = 2t$, $z = t - 1$, $t \in \mathbb{R}$.

Soluciones de Sección 2.7

1.(a) Se intersecan en $(\frac{5}{2}, 2, 2)$.

1.(b) Existen infinitas rectas con dicha propiedad.

5. $E \cap F$ es un plano ó es una recta que pasa por A y B .

Soluciones de Sección 2.9

1. *Hint.* Es rápido si usa proyecciones.

2. $\vec{b} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ y $\vec{b} = \begin{pmatrix} -3 \\ -1 \end{pmatrix}$.

6.(a) Falso.

6.(b) Falso.

6.(c) Verdadero.

6.(d) Verdadero.

9.(a)ii. Colisionan en el punto $(7, 10, -1)$ y en tiempo $t = 2$.

9.(b)ii. No colisionan.

9.(b)iii. Las dos estelas se mezclan en el punto $(13, 20, -5)$.

15.(a) $3\sqrt{6}$.

16.(c) $Q(-3, 5, 5)$.

16.(d) La distancia obtenida es $\sqrt{11}$.

18.(b) Solo existe un único plano que cumple tal condición.

19.(a) No.

19.(c) $34x - 5y - 9z = 0$ es el único plano con las propiedades deseadas.

20.(b) $\lambda = \mu = \frac{1}{2}$.

Soluciones de algunos ejercicios del Capítulo 3

Soluciones de Sección 3.1

$$1.(a) \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & -1 \end{array} \right).$$

$$1.(c) \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

$$1.(e) \left(\begin{array}{cccc|c} 1 & 0 & 3 & 0 & 5 \\ 0 & 1 & 2 & 0 & -3/2 \\ 0 & 0 & 0 & 1 & 3/2 \end{array} \right).$$

$$1.(g) \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -3 \end{array} \right).$$

$$1.(g) \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 5 \end{array} \right).$$

2. Los $a, b, c \in \mathbb{R}$ tales que $a - 2b + c = 0$.

3. $-x^2 + 3x - 2$.

5. El rolo pasó 6 días en Medellín, 4 días en Villavicencio y 4 días en Yopal.

Soluciones de Sección 3.2

$$2.(a) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = s \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$

$$2.(b) \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = s \begin{pmatrix} 5 \\ 6 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 3 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

$$2.(c) \begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$

$$3. r \in \mathbb{R} - \{-3, 2\}$$

Soluciones de Sección 3.3

1.(a) Sí.

1.(b) No.

1.(c) No.

Las soluciones del sistema homogéneo son

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = t \begin{pmatrix} 5 \\ -3 \\ 5 \\ 2 \end{pmatrix}$$

2. (a) No tiene solución.

(b) No tiene solución.

3.(a) $a \in \mathbb{R} \setminus \{2, -2\}$.

3.(b) $a = 2$.

3.(c) $a = -2$.

Soluciones de 3.4

$$1.(e) D = \begin{pmatrix} 0 & 0 \\ 8 & 5 \\ -6 & -3 \end{pmatrix}.$$

2.(b) No es posible efectuar la multiplicación.

$$2.(d) \begin{pmatrix} 0 & -2 & 5 \\ 0 & -6 & 15 \\ 0 & 2 & -5 \end{pmatrix}.$$

$$2.(f) \begin{pmatrix} 19 & -17 & 34 \\ 8 & -12 & 20 \\ -8 & -11 & 7 \end{pmatrix}.$$

$$4. A = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}.$$

$$6. \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix}.$$

7. $a = d, c = 0$.

Soluciones de Sección 3.5

3. Una solución particular al sistema $A\vec{x} = \vec{b}$ usando la inversa a derecha es

$$\vec{x} = \begin{pmatrix} 4 & -3 \\ 0 & 0 \\ -1 & 1 \end{pmatrix} \vec{b}.$$

Soluciones de Sección 3.6

$$1.(a) \frac{1}{12} \begin{pmatrix} 6 & -3 \\ 2 & 1 \end{pmatrix}.$$

1.(e) No tiene inversa.

$$1.(f) \frac{1}{9} \begin{pmatrix} 21 & -3 & -3 & -6 \\ 4 & -1 & -4 & 1 \\ -1 & -2 & 1 & 2 \\ -15 & 6 & 6 & 3 \end{pmatrix}.$$

$$1.(h) \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/5 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

2.

(I) Solo tiene solución trivial.

(II) Tiene solución no trivial.

3. A tiene inversa si $a \neq 2, -2$ y en tal caso

$$A^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{4-a^2} & \frac{2-a^2}{2(a^2-4)} & \frac{a^2}{2(a^2-4)} \\ \frac{1}{a^2} - 4 & \frac{1}{a^2-4} & \frac{-2}{a^2-4} \end{pmatrix}.$$

5.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \vartheta & \sin \vartheta \\ 0 & -\sin \vartheta & \cos \vartheta \end{pmatrix}.$$

Soluciones de Sección 3.7

2. $\alpha = \frac{3}{2}, \beta = \frac{7}{2}$.

3.

(a) No.

(b) Sí.

(c) No.

(d) Sí.

4. No se puede concluir que AB es simétrica.

Soluciones de Sección 3.8

1.a) Sí. 1.c) No. 1.f) No.

3.a) $Q_{21}(-2) \begin{pmatrix} 2 & -3 \\ 0 & 0 \end{pmatrix}$.

3.(c) $Q_{21}(5)P_{23}Q_{21}(2) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$.

4.b) $E = Q_{31}(-4)$.

4.c) $E = P_{13}$.

4.e) $E = Q_{12}(3)$.

Soluciones de Sección 3.10

2. $-\frac{7}{2x} + \frac{11}{2(x-2)} - \frac{7}{(x-2)^2}$.

7. $\begin{pmatrix} 5 \\ 3/2 \\ 3 \end{pmatrix}$.

10. La matriz B no es invertible. La matriz D es invertible y $D^{-1} = \frac{1}{57} \begin{pmatrix} 3 & 15 & -6 \\ -12 & -3 & 24 \\ 15 & -1 & -11 \end{pmatrix}$.

24. $X = \frac{1}{19} \begin{pmatrix} 50 & 29 \\ -29 & 36 \end{pmatrix}$.

Soluciones de algunos ejercicios del Capítulo 4

Soluciones de Sección 4.1

1. (a) -4,
1. (c) 47,
1. (e) 6,
1. (g) 96.

Soluciones de Sección 4.2

1. $-3/5, -1, -1$.
2. 18.
3. Para todo $a \in \mathbb{R}$ la matriz tiene inversa.
- 6.(a) 1 y -1 .
- 6.(b) 0 y 1.
7. *Hint.* Use que A es invertible.

Soluciones de Sección 4.3

2. 0.
3. 3.
4. Elija por ejemplo P como punto inicial y desarrolle $\det[\overrightarrow{PQ}, \overrightarrow{PR}]$. Compare con el determinante de la fórmula dada.

Soluciones de Sección 4.4

- 1.b) $x = \frac{10}{7}, y = \frac{26}{7}, z = -\frac{4}{7}$.
- 1.d) $x = \frac{21}{29}, y = \frac{171}{29}, z = -\frac{284}{29}, w = -\frac{182}{29}$.
- 2.b) $\frac{1}{14} \begin{pmatrix} 3 & -6 & 2 \\ 2 & 10 & -8 \\ -1 & 2 & 4 \end{pmatrix}$.
- 2.d) $\begin{pmatrix} 1 & 0 & 2 & 0 \\ -1 & 1 & 2 & -2 \\ 1 & 0 & -3 & 3 \\ 2 & -2 & -2 & 3 \end{pmatrix}$.
3. $-\frac{4}{37}$.
5. Use la relación entre A y $\text{adj } A$.

Soluciones de Sección 4.6

2. $t \in \mathbb{R} \setminus \{0, -\frac{1}{3}, -2\}$.

4.c) Es invertible y su inversa es

$$\frac{1}{38} \begin{pmatrix} 1 & -69 & 10 & 27 \\ 8 & 18 & 4 & -12 \\ 7 & 11 & -6 & -1 \\ -11 & -1 & 4 & 7 \end{pmatrix}.$$

9. $(-1)^{n+1}b_1$.

Soluciones de algunos ejercicios del Capítulo 5

Soluciones de Sección 5.1

2. Sí es espacio vectorial.
4. No es espacio vectorial.
6. Sí es espacio vectorial.
7. Sí es espacio vectorial.
9. Sí es espacio vectorial.
10. No es espacio vectorial.
13. Sí es espacio vectorial.

Soluciones de Sección 5.2

1. Sí es subespacio de V .
3. Sí es subespacio de V .
5. Sí es subespacio de V .
7. No es subespacio de V .
8. No es subespacio de V .
11. Sí es subespacio de V .
14. Sí es subespacio de V .

$$17.(b) W_1 \cap W_2 = \left\{ \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix} : a, b \in \mathbb{R} \right\}.$$

Soluciones de Sección 5.3

2. $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.
4. $\{1, x, x^3\}$.
6. $\{2x^3 - 3x^2 + x, 1\}$.
8. $\left\{ \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 5 \end{pmatrix} \right\}$
9. $\left\{ \begin{pmatrix} 2 \\ 3 \\ \frac{1}{3} \end{pmatrix}, \begin{pmatrix} 6 \\ 9 \\ 1 \end{pmatrix} \right\}$.

- 14.(a) Son linealmente dependientes.
- 14.(c) Son linealmente dependientes.
- 14.(e) Son linealmente independientes.
- 14.(g) Son linealmente dependientes.
16. El conjunto dado es linealmente dependiente y su generado es el plano $x - y + z = 0$
- 20.(a) Falso.
- 20.(c) Verdadero.

20.(d) Verdadero.

20.(e) Verdadero.

Soluciones de Sección 5.4

- 2.(b) Tiene dimensión 3.
- 2.(e) Su dimensión es 0.
- 2.(f) Tiene dimensión 2.
- 2.(g) Tiene dimensión 2.
- 3.(a) El conjunto dado es linealmente dependiente.
- 3.(b) El conjunto dado sí es base de W .
- 3.(c) El conjunto dado es base de $\mathbb{R}^4$.
- 3.(d) El conjunto dado solo tiene dos vectores.
- 3.(e) El conjunto dado es linealmente dependiente.
4. Una base del subespacio dado es $\left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}$.
6. *Hint.* Si tiene una base de E , completela a una base de $\mathbb{R}^3$ con el vector normal del plano.
10. $\alpha \in \mathbb{R} - \{-1, 2\}$.
- 11.(c) Sí existen. *Hint.* Modifique un poco la base canónica de P_n .

Soluciones de Sección 5.5

$$1. U \cap V = \text{span} \left\{ \begin{pmatrix} 7 \\ -3 \\ 0 \\ 1 \end{pmatrix} \right\}$$

$$U + V = \{ \vec{x} \in \mathbb{R}^4 : \langle \vec{x}, (-2, -1, 1, 1) \rangle = 0 \} \text{ y } \dim(U + V) = 3.$$

3. La suma de V y W no es directa.

4. *Hint.* Recuerde cuánto valen $\dim M_{\text{sym}}(3 \times 3)$ y $\dim M_{\text{asym}}(3 \times 3)$.

5. (a) No.

5. (b) Sí.

Soluciones de Sección 5.7

- 1.(a) Sí es subespacio.
- 1.(d) Sí es subespacio.
- 1.(f) No es subespacio.
- 1.(h) Sí es subespacio.
- 1.(j) No es subespacio.
- 4. No.
- 5. No.
- 26.(b) $n^2 - n$.
- 26.(d) 5.
- 29.(a) Sí.
- 29.(c) No.
- 29.(a) Sí.

Soluciones de algunos ejercicios del Capítulo 6

Soluciones de Sección 6.1

2. No es transformación lineal.
 4. Sí es transformación lineal.
 6. Sí es transformación lineal.
 8. Sí es transformación lineal.
 10. Sí es transformación lineal.
 12. Sí es transformación lineal.
 14. No es transformación lineal.
 15. 4. $\text{Im } T, \ker T$ son ambos el eje x .
 15. 6. $\text{Im } T$ es el plano $x + y - z = 0$ y

$$\ker T = \text{span} \left\{ \begin{pmatrix} 6 \\ 4 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -6 \\ -3 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

15. 8. $\text{Im } T = \text{span}\{x^3, 1\}$ y

$$\ker T = \text{span} \left\{ \begin{pmatrix} 6 \\ 2 \\ -1 \end{pmatrix} \right\}.$$

- 15.)10. $\text{Im } T = M(2 \times 2)$ y $\ker T = \{\mathbb{O}_{2 \times 2}\}$.

- 15.)12. $\text{Im } T = \mathbb{R}$ y $\ker T$ son las matrices de

$$\text{la forma} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & -(a_{11} + a_{22}) \end{pmatrix}$$

17. $\text{Im } T = \text{span}\{\vec{w}\}$ y
 $\ker T = \{\vec{x} \in \mathbb{R}^n : \vec{x} \perp \vec{w}\}$.

20. No existe.

Soluciones de Sección 6.2

- 2.(a) $\text{Im } T$ es el plano $x - 2z = 0$ y

$$\ker T = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \right\}.$$

$$2.(b) \text{Im } T = \mathbb{R}^3 \text{ y } \ker T = \text{span} \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \\ 2 \end{pmatrix} \right\}.$$

$$2.(c) \text{Im } T = \text{span} \left\{ \begin{pmatrix} 1 \\ 5 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 4 \\ -3 \\ 0 \\ 2 \end{pmatrix} \right\} \text{ y}$$

$$\ker T = \text{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}.$$

- 2.(d) $\text{Im } T = \text{span}\{\vec{w}\}$ y $\ker T$ es el plano $x + 3y - z = 0$.

5. Observe que $\text{Im } A = \mathbb{R}^m$.

7. *Hint.* Recuerde que
 $\dim \text{Im } A + \dim \ker A = n$.

Soluciones de Sección 6.3

$$2. [p(X)]_{\mathcal{B}} = \begin{pmatrix} b-a & \\ & b \\ c-2b+a & \end{pmatrix}.$$

$$3.(a) \text{Recuerde que } \sinh x = \frac{e^x - e^{-x}}{2},$$

$$\cosh x = \frac{e^x + e^{-x}}{2}.$$

$$3.(b) A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}$$

$$4. [a + bX + cX^2]_{\mathcal{B}} = \begin{pmatrix} c + b + a \\ 2c + b \\ c \end{pmatrix}.$$

$$6. A_{\mathcal{B}_{\vartheta} \rightarrow \text{can}} = \begin{pmatrix} \cos \vartheta & \sin \vartheta \\ -\sin \vartheta & \cos \vartheta \end{pmatrix} \text{ y}$$

$$A_{\text{can} \rightarrow \mathcal{B}_{\vartheta}} = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix}.$$

$$8.(a) \left[\begin{pmatrix} -3\sqrt{3} \\ -3 \end{pmatrix} \right]_{\mathcal{B}_{\vartheta}} = \begin{pmatrix} -3 \\ 0 \end{pmatrix}.$$

$$8.(b) \begin{pmatrix} 1 \\ -1 \end{pmatrix}_{\mathcal{B}_{\vartheta}} = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}.$$

9. *Hint.* Use la relación

$$A_{\mathcal{B}_{\vartheta_1} \rightarrow \mathcal{B}_{\vartheta_2}} = A_{\mathcal{B}_{\vartheta_1} \rightarrow \text{can}} A_{\text{can} \rightarrow \mathcal{B}_{\vartheta_2}}.$$

Soluciones de Sección 6.4

$$1. 4. \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$1. 6. \begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 1 & 4 & 3 \\ 1 & 0 & 6 & 6 \end{pmatrix}.$$

1. 8. $\begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -3 & 0 \end{pmatrix}.$

1. 10. $\begin{pmatrix} 2 & 0 & 0 & 0 \\ 4 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 4 & 1 \end{pmatrix}.$

1. 12. $(1 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 1).$

2.(a) $\begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$

2.(e) $(1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n+1}).$

3.(b) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 1 \end{pmatrix}.$

3.(d) $\begin{pmatrix} 0 & 4/5 & 2/5 \\ 3 & 7/5 & 16/5 \end{pmatrix}.$

4.(a) *Hint.* Suponga una combinación lineal de vectores de $\mathcal{B}$ igualada a 0, evalúe en $x = 0$, $x = \pi$.

4.(b) $[D]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$

8.(a) $\vec{w} = (4, -3, 0, \pi)^t.$

8.(b) $\vec{w} = (19, -18, 3, -5)^t.$

Soluciones de Sección 6.6

22.(b) Se cumple para los polinomios con coeficiente independiente cero.

22.(c) Se cumple para los polinomios en P_3 .

23. Hay varias opciones, la más natural es considerar $T\vec{x} = \langle \vec{x}, \vec{w} \rangle$.

24.(a) Observe que si $n = 2$ entonces $\dim \ker \varphi = 1$ y si $n = 3$ entonces $\dim \ker \varphi = 2$.

24.(b) Elija φ tal que su kernel (que es una recta) no pase por ninguno de los vectores dados.

Soluciones de algunos ejercicios del Capítulo 7

Soluciones de Sección 7.1

- 1.(a) No es base ortonormal de V
- 1.(c) Sí es base ortonormal de V .
- 1.(d) Sí es base ortonormal de V .
2. $a = -\frac{5}{8}$, $b = -\frac{21}{8}$.
- 4.(d) *Hint.* Escoja una base ortonormal del plano $3x + 2y + 5z = 0$.
5. Falso.

Soluciones de Sección 7.2

1. La matriz $\begin{pmatrix} \cos \vartheta & -\sin \vartheta & 0 \\ \sin \vartheta & \cos \vartheta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ rota en un ángulo ϑ el plano xy .
2. *Hint.* Recuerde que $A_{B \rightarrow \text{can}}$ es ortogonal.
- 4.(a) Falso.
- 4.(a) Verdadero.

Soluciones de Sección 7.3

- 1.(a) $\text{span} \left\{ \begin{pmatrix} 5 \\ 1 \end{pmatrix} \right\}$.
- 1.(c) La recta $t \begin{pmatrix} 7 \\ 5 \\ 1 \end{pmatrix}$.
- 1.(d) Una base del complemento ortogonal es $\left\{ \begin{pmatrix} -5 \\ 3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -7 \\ 5 \\ 0 \\ 1 \end{pmatrix} \right\}$.
- 3.(a) $\begin{pmatrix} 3 \\ 5 \end{pmatrix} = -\begin{pmatrix} 1 \\ -1 \end{pmatrix} + 4\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Una base ortonormal de $W^\perp$ es $\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$.
- 3.(c) $\begin{pmatrix} 10 \\ -1 \\ 6 \end{pmatrix} = \begin{pmatrix} 10 \\ -1 \\ 6 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. Una base ortonormal de $W^\perp$ es $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{37}} \begin{pmatrix} 0 \\ 1 \\ -6 \end{pmatrix} \right\}$.

Soluciones de Sección 7.4

1. $\frac{5}{\sqrt{6}}$.
2. $\sqrt{59}$.
- 3.(b) $T(\vec{x}) = \vec{x} - 2 \text{proy}_{L^\perp} \vec{x}$.
- 4.(a) $[P_W]_{\mathcal{B}} = \begin{pmatrix} \text{id}_k & \mathbb{O}_{k \times (n-k)} \\ \mathbb{O}_{(n-k) \times k} & \mathbb{O}_{(n-k) \times (n-k)} \end{pmatrix}$
- 4.(c) Recuerde como se expresa $[P_W]_{\text{can}}$ en términos de $[P_W]_{\mathcal{B}}$.
- 4.(c) No hace falta hacer cuentas, aplique el inciso anterior.

Soluciones de Sección 7.5

1. Partiendo de la base $\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$ se obtiene: $\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$.
2. Sea W el generado de $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix}$. Encuentre una base para $W^\perp$ y aplique Gram-Schmidt en la base dada de W y en la base obtenida de $W^\perp$.
3. $\text{proy}_W \vec{v} = \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \end{pmatrix}$ y la distancia de $\vec{v}$ a W es $\frac{1}{\sqrt{2}}$.
4. Observe que $\dim \text{Im } A = 3$ y una base de $\text{Im } A$ es $\left\{ \begin{pmatrix} 1 \\ 3 \\ 2 \\ 7 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -1 \\ 1 \end{pmatrix} \right\}$. Aplicando Gram-Schmidt se obtiene $\left\{ \frac{1}{3\sqrt{7}} \begin{pmatrix} 1 \\ 3 \\ 2 \\ 7 \end{pmatrix}, \frac{1}{\sqrt{35}} \begin{pmatrix} 11/3 \\ 4 \\ 1/3 \\ 7/3 \end{pmatrix}, \frac{1}{\sqrt{15}} \begin{pmatrix} 2 \\ -1 \\ 3 \\ 1 \end{pmatrix} \right\}$.

Soluciones de Sección 7.8

2. $\dim U = \dim U^\perp = 2$. Una base para $U^\perp$ es

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

(a) $3 \frac{1}{14} \begin{pmatrix} 13 & -2 & -3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{pmatrix}.$

(b) $3 \frac{1}{4} \begin{pmatrix} 3 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 3 \end{pmatrix}.$

5. Una base ortonormal para $U^\perp$ es

$$\frac{1}{\sqrt{17}} \left\{ \begin{pmatrix} 4 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{357}} \begin{pmatrix} 2 \\ 0 \\ -17 \\ 8 \end{pmatrix} \right\}.$$

Soluciones de algunos ejercicios del Capítulo 8

Soluciones de Sección 8.1

1. Al reducir la matriz se obtiene la identidad.
2. Una base ortonormal para $\text{Im } A$ es $\left\{ \frac{1}{\sqrt{28}} \begin{pmatrix} 1-i \\ 3i \\ 4+i \end{pmatrix} \right\}$ y una base ortonormal para $\ker A$ es $\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{17}} \begin{pmatrix} -2-2i \\ -2+2i \\ 1 \end{pmatrix} \right\}$.
4. $a = 1, b = 0, c = 0, e = 2, f = -2$.
- 6.(a) Sí.
- 6.(b) No.
- 7.(c) La afirmación del inciso (b) en $\mathbb{C}^n$ no es válida.

Soluciones de Sección 8.2

- 3.(a) Falsa.
- 3.(b) Falsa.
- 3.(c) Falsa.
- 5.(c) Use que $B = CAC^{-1}$ y la propiedad del inciso (b).
- 5.(d) Observe que las entradas de la diagonal de $A^t A$ son las normas al cuadrado de los vectores columna de A .

Soluciones de Sección 8.3

2. Su polinomio característico es λ^4 .
3. El polinomio característico de D es $(\lambda - 1)(\lambda - 2)(\lambda - 3)(\lambda - 4)$.
4. Si empieza diagonalizando A se obtiene $A = \frac{1}{3} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$.
5. Los valores propios de T son 1 y -1 .
6. Resuelva la ecuación diferencial $y = \lambda y'$ sujeta a la condición inicial $y(0) = 0$.
8. Si λ es valor propio de A , existe $\vec{x} \neq 0$ tal que $A\vec{x} = \lambda\vec{x}$. Multiplique por A^{-1} .
10. Observe que $\vec{x} \neq 0$ es un vector propio de T si $\vec{v} \times \vec{x} = \lambda\vec{x}$. ¿Para cuáles λ la igualdad anterior es cierta?
11. Vea el ejercicio 4.(a) de la sección 7.4
- 13.(b) El polinomio característico de A es λ^n .

14. Los valores propios de A son 0 y 1.
16. Si A no es diagonalizable la afirmación es falsa.

Soluciones de Sección 8.4

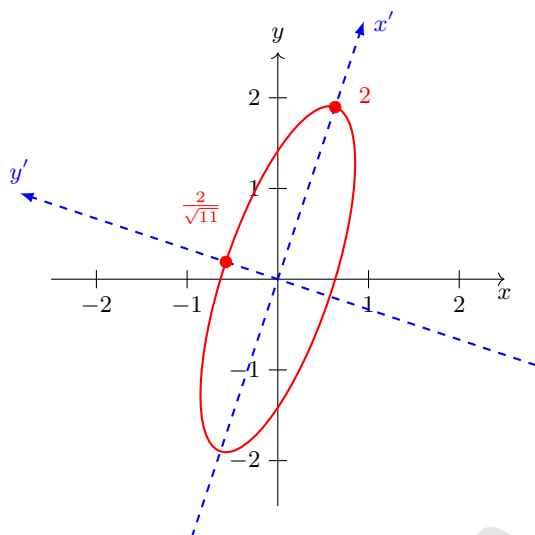
- 3.(a) No.
- 3.(a) Sí.
4. Las tres afirmaciones son falsas.
5. $A^{20} = \text{id}_{3 \times 3}$
6. $k \neq 5t$ ó $t = 5, k = 25$.
7. Expresar $A = QDQ^{-1}$ donde D es la matriz de valores propios de A , por ende $(A - d_1 \text{id}_n)(A - d_2 \text{id}_n) \dots (A - d_k \text{id}_n) = Q(D - d_1 \text{id}_n)(D - d_2 \text{id}_n) \dots (D - d_k \text{id}_n)Q^{-1}$.
- 9.(c) Si no suponemos que $C \neq \mathbb{O}_{2 \times 2}$ la conclusión de 9.(b) es falsa.

Soluciones de Sección 8.5

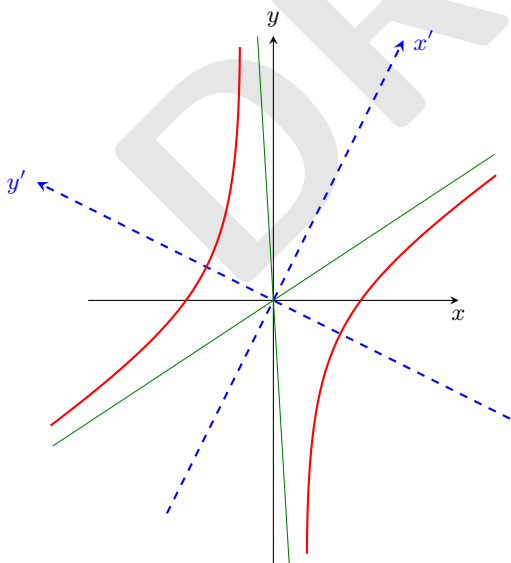
- 2.(a) Los valores propios de A son 1 y 2.
- 2.(b) $E_1 = \text{span} \left\{ \begin{pmatrix} 0 \\ 3 \\ -4 \end{pmatrix} \right\}$ y E_5 es el plano $3y - 4z = 0$.
3. $A = \frac{1}{3} \begin{pmatrix} 1+i & 1 \\ 1 & -1+i \end{pmatrix} \begin{pmatrix} 8 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1-i & 1 \\ 1 & -(1+i) \end{pmatrix}$.
5. Escoja una base ortonormal de $L^\perp$ y complétela a una base de $\mathbb{R}^3$ con un vector unitario de L . La representación matricial de T en esta base será $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.
6. Use el teorema 8.56*
- 7.(a) Use el teorema 8.56
- 7.(b) Use el teorema 8.56 y el ejercicio 8. de la sección 8.1.

Soluciones de Sección 8.6

1.(a) $(x')^2 + 11(y')^2 = 4$ con ejes de rotación $\frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \frac{1}{\sqrt{10}} \begin{pmatrix} -3 \\ 1 \end{pmatrix}$.

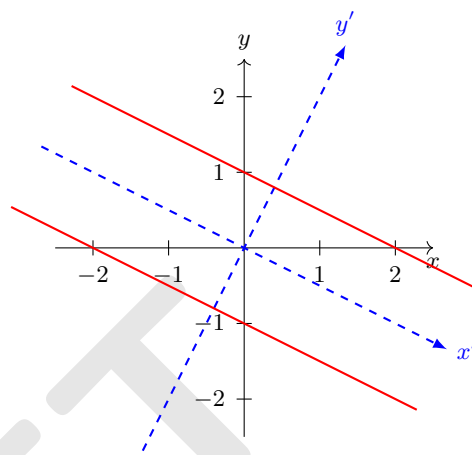


1.(d) $-(x')^2 + 3(y')^2 = 6$ con ejes de rotación $\frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ y asíntotas $y = \frac{-(1+2\sqrt{3})}{2-\sqrt{3}}x, y = \frac{2\sqrt{3}-1}{2+\sqrt{3}}x$.



1.(e) $5(y')^2 = 4$ con ejes de rotación

$$\frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$



1.(g) $4(x')^2 + 6(y')^2 = -4$.

3. $3x^2 - 4\sqrt{3}xy + 7y^2 = 9$

Soluciones de Sección 8.8

6.(b) Los autovalores de A son 4, 7, 10.

9. $D = \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix}$ y $e^A = C^{-1} \begin{pmatrix} 1 & 0 \\ 0 & e^5 \end{pmatrix} C$.

10.(a) Los valores propios de Φ son 1 y -1 y los espacios propios son $M_{sym}(n \times n)$, $M_{asym}(n \times n)$.

10.(b) El único valor propio de T es 3 con espacio propio asociado $\text{span}\{1\}$.

10.(c) Los valores propios son 1, -1 y los espacios propios son el plano $x + 2y + 3z = 0$

y la recta $t \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

11.c Los autovalores de T son los valores de la forma $-\frac{k^2\pi^2}{L^2}$ donde $k \in \mathbb{N}$ y para cada $-\frac{k^2\pi^2}{L^2}$, su espacio propio es $\text{span}\{\sin \frac{k^2\pi^2}{L^2}t\}$.

Índice alfabético

$|A|$ (determinante de A), 140

$\|\cdot\|$, 47

$\|\cdot\|$, 32, 322

$\oplus$, 210

$\langle \cdot, \cdot \rangle$, 322

$\langle \cdot, \cdot \rangle$, 35, 47

$\angle(\vec{v}, \vec{w})$, 37

$\|$, 37

$\perp$, 37, 293, 323

$\sim$, 329

$\times$, 49

$\wedge$, 49

$\mathbb{C}$, 379

$\mathbb{C}^n$, 321

$M(m \times n)$, 80

$\mathbb{R}^2$, 27, 30

$\mathbb{R}^3$, 49

$\mathbb{R}^n$, 46

$\text{Eig}_\lambda(T)$, 334

Im , 379

$\mathcal{L}(U, V)$, 228, 278

$M_{\text{asym}}(n \times n)$, 122

$M_{\text{sym}}(n \times n)$, 122

P_n , 182

Re , 379

S_n , 139

$U^\perp$, 293

$\text{dist}(\vec{v}, U)$, 302

$\text{adj } A$, 158

A^* , 326

$\arg$, 383

gen , 184

p_A , 337

$\text{proy}_U \vec{v}$, 300

$\text{proy}_{\vec{w}} \vec{v}$, 43, 48, 323

span , 184

A^t , 120

$\hat{v}$, 48

$\vec{v}_{\parallel}$, 42

$\vec{v}_{\perp}$, 42

adjunta

matriz (adjoint), 326

matriz (adjugate), 158

afín

subespacio, 177

ángulo entre $\vec{v}$ y $\vec{w}$, 37

ángulo entre dos planos, 61

antisimétrica, 122

aplicación lineal, *véase* transformación lineal

aproximación por mínimos cuadrados, 308

argumento de un número complejo, 383

aumentada

matriz de coeficientes, 13, 80

base, 198

cambio de, 249

estándar de $\mathbb{R}^n$, 199

estándar de P_n , 199

base ortogonal, 285

biyectiva, 229

cambio de bases, 249

cambio de coordenadas

matriz, 252

campo, 380

cofactor, 141

combinación lineal, 184

complemento ortogonal, 293, 293

componente de una matriz, 13

composición de funciones, 101

conjugado complejo, 379

conjunto vacío, 185, 187, 189, 199, 202

desigualdad de Cauchy-Schwarz, 39, 324

- desigualdad de Hölder, 324
- desigualdad triangular, 33, 39, 324
- determinante, 20, 140
 - expansión a lo largo de la k ésima fila/columna, 142
 - expansión de Laplace, 142
 - fórmula de Leibniz, 140
 - regla de Sarrus, 144
- diagonal, 122
- diagonalizable, 331
- diagonalización ortogonal, 353
- diagrama, 264, 336
 - conmutativo, 264
- dimensión, 202
- distancia de $\vec{v}$ a un subespacio, 302
- ecuación simétrica, 58
- ecuación vectorial para rectas, 56
- ecuaciones paramétricas, 57
- ejes principales, 361
- elemental
 - matriz, 124
- eliminación gaussiana, 85
- eliminación de Gauß-Jordan, 85
- entrada, 13
- espacio vectorial, 167
- equivalente por filas, 85
- escalar, 28
- espacio columna, 240
- espacio fila, 240
- espacio nulo, 229
- espacio propio, 334
- espacio vectorial, 31
 - generado, 184
 - intersección, 210
 - polinomios, 182
 - subespacio, 174
 - suma, 210
 - suma directa, 210
- expansión a lo largo de la k ésima fila/columna, 142
- expansión de Laplace, 142
- fórmula de Leibniz, 140
- fórmulas de Euler, 382
- fila, 13
- finitamente generado, 187
- forma escalonada por filas, 83
- forma escalonada reducida por filas, 83
- forma normal
 - plano, 60
 - recta, 58
- forma vectorial de la solución, 89
- generador, 184
- Gram-Schmidt, 304
- hiperplano, 55, 207
- idempotente, 154
- identidad aditiva, 381
- identidad aditiva, 31, 167
- identidad multiplicativa, 381
- imagen de una aplicación lineal, 230
- inversa, 116
- inversa por la derecha, 111
- inversa por la izquierda, 111
- inverso aditivo, 169
- invertible, 110
- inyectiva, 229
- isometría, 291
- lado derecho, 12, 79
- linealmente dependiente, 189
- linealmente independiente, 189
- longitud de un vector, *véase* norma de un vector
- módulo, 379
- magnitud de un vector, *véase* norma de un vector
- matrices semejantes, 329
- matrix
 - orthogonal, 153
- matrix inversa
 - $2 \times$, 118
- matriz, 80
 - adjunta (adjoint), 326
 - adjunta (adjugate), 158
 - antisimétrica, 122
 - cambio de coordenadas, 252
 - coeficientes, 80
 - cofactor, 141
 - cuadrada, 80

- diagonal, 122
- diagonalizable, 331
- elemental, 124
- equivalente por filas, 85
- espacio columna/fila, 240
- forma escalonada por filas, 83
- forma escalonada reducida por fila, 83
- hermítica, 326, 353
- idempotente, 154
- inversa por la derecha, 111
- inversa por la izquierda, 111
- inversa, 116
- invertible, 110
- menor, 141
- multiplicación por un vector, 95
- producto, 103
- producto por escalar, 102
- simétrica, 122
- simétrica, 353
- singular, 110
- suma, 101
- transición, 252
- transpuesta, 120
- triangular inferior, 122
- triangular superior, 122
- unitaria, 326
- matriz de cambio de coordenadas, 252
- matriz de coeficientes, 80
 - aumentada, 80
- matriz de coeficientes aumentada, 13, 80
- matriz de coeficientes matrix, 13
- matriz de transición, 252
- matriz elemental, 124
- matriz hermítica, 326
- matriz ortogonal, 289
- menor, 141
- multiplicidad
 - algebraica, 339
 - geométrica, 334
- multiplicidad algebraica, 339
- multiplicidad geométrica, 334
- mínimos cuadrados
 - aproximación, 308
- número complejo, 379
- norma, 379
- norma de un vector, 32, 47, 322
- nulidad, 232
- núcleo, 229
- ONB, 285
- operaciones elementales de filas, 81
- operaciones de filas, 81
- operador lineal, *véase* transformación lineal
- ortogonal
 - base, 285
 - complemento, 293
 - matriz, 153, 289
 - proyección, 299
- ortogonalización, 304
- paralelepípedo, 53
- paralelogramo, 52
- parte imaginaria de z , 379
- parte real de z , 379
- permutación, 139
- pivote, 83
- plano, 55, 207
 - ángulo entre dos planos, 61
 - forma normal, 60
 - vector normal, 60
- plano complejo, 379
- polinomio característico, 337
- proceso de Gram-Schmidt, 304
- producto
 - interno, 47, 322
- producto cruz, 49
- producto de vector en $\mathbb{R}^2$ por escalar, 30
- producto escalar, 35, 47, 322
- producto interno, 35, 47, 322
- producto punto, 35, 47, 322
- producto vectorial, 49
- proyección ortogonal, 299, 300
- proyección ortogonal en $\mathbb{R}^2$, 42
- proyección ortogonal en $\mathbb{R}^n$, 48, 300, 323
- proyección ortogonal sobre un plano en $\mathbb{R}^3$, 270
- radio de convergencia, 382
- rango, 230, 232
- recta, 55, 207
 - ecuación simétrica, 58
 - ecuación vectorial, 56

- ecuaciones paramétricas, 57
- forma normal, 58
- vector director, 56
- reflexión en $\mathbb{R}^3$, 270
- reflexión in $\mathbb{R}^2$, 269
- relación de equivalencia, 329
- renglón, 13
- representación polar de un número complejo, 384
- representación matricial de una transformación lineal, 262
- Sarrus
 - regla de, 144
- sesquilineal, 323
- signo de una permutación, 139
- simétrica
 - matriz, 122
- sistema
 - ortogonal, 284
 - ortonormal, 284
- sistema lineal, 79
 - consistente, 80
 - homogéneo, 92
 - inhomogéneo, 92
- sistema ortogonal, 284
- sistema ortonormal, 284
- sobreajuste, 314
- sobreyectiva, 229
- solución
 - forma vectorial, 89
- solución trivial, 92, 188
- span, 184
- span lineal, 184
- subespacio, 174
 - afín, 177
- subespacio propio, 174
- subespacio trivial, 175
- sum directa, 210
- suma de funciones, 101
- suma directa, 206
- suma vectorial en $\mathbb{R}^2$, 30
- systema lineal, 12
 - consistente, 12
 - homogéneo, 12
 - inhomogéneo, 12
- solución, 12
- Teorema de Pitágoras, 301, 323
- transformaciones lineales
 - representación matricial, 261
- transformación lineal, 227
 - representación matricial, 262
- transpuesta
 - matriz, 120
- traza, 333, 338
- triangular inferior, 122
- triangular superior, 122
- unidad imaginaria, 379
- uno a uno, 229
- valor propio, 331, 333
- variables libres, 87
- vector, 31
 - en $\mathbb{R}^2$, 27
 - norma, 32, 47, 322
 - ortogonal, 37
 - perpendicular, 37
 - unitario, 33
- vector director, 56
- vector normal de un plano, 60
- vector propio, 331, 333
- vector unitario, 33
- vectores
 - ortogonales, 323
 - paralelos, 37
 - perpendiculares, 323
- vectores ortogonales, 37, 323
- vectores perpendiculares, 323