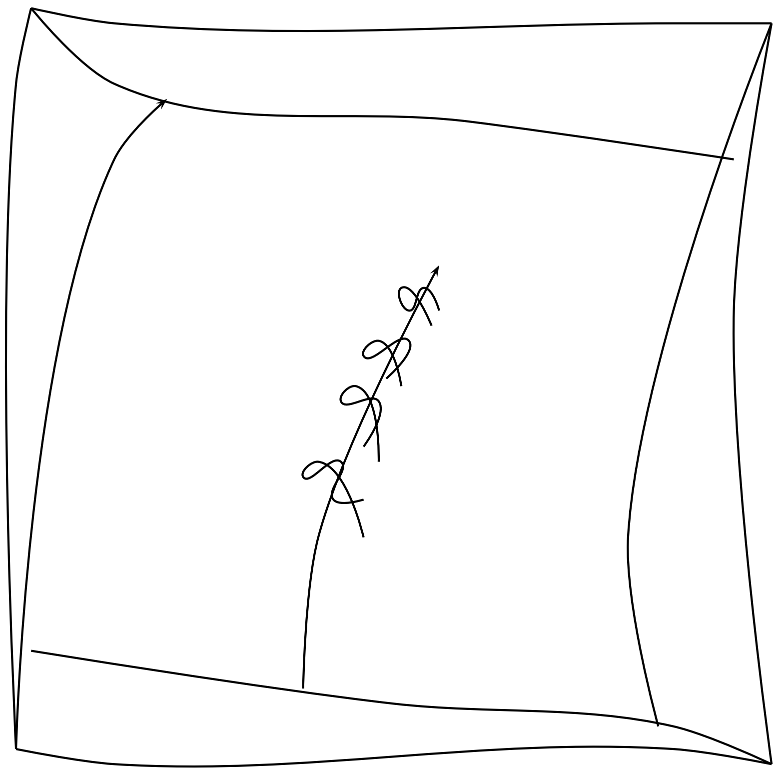
ÁLGEBRA LINEAL

José López/ José Rodríguez

6 de septiembre de 2010



Texto procesado en LaTeX y 90 gráficas en PsTricks.

GARANTÍA

Los autores de este texto garantizan a los estudiantes y a los profesores que todo el material es intuitivo, entendible, de complejidad regulada y que saldrán bien preparados. Nuestra garantía se debe a que hemos revestido todo el formalismo algebraíco de conceptos geométricos, intuitivos y depurados. También sabemos que para el estudiante es muy difícil asimilar muchos conceptos nuevos al mismo tiempo y por ello hemos sido cuidadosos en introducir conceptos nuevos poco a poco. Hemos añadido muchos ejemplos y variados ejercicios de rutina y además talleres de contenidos a veces desafiantes. Por todo esto, el estudiante puede confiar en que sale bien preparado si trabaja el material concienzudamente.

Comenzamos con sistemas de ecuaciones de primer grado. Inmediatamente pasamos al estudio de matrices y de las operaciones que se pueden hacer con ellas. Paralelamente uno se pregunta sobre el significado de lo que se hace y la respuesta puede darse en términos geométricos con sus sofisticaciones: vectores, líneas, planos, paralelepípedos, volúmenes y determinantes, transformaciones lineales, cambio de coordenadas y diagonalización.

José Darío López García.
José del Carmen Rodríguez Santamaría
Universidad de los Andes

ÍNDICE GENERAL

CAPÍTULO 1

ECUACIONES

La validez de las cosas que uno dice depende del contexto dentro del cual se esté hablando. Nosotros fijamos el contexto especificando el conjunto universal a partir del cual se construye lo que se necesite. Nuestro conjunto universal está formado por los números reales, a menos que se especifique lo contrario.

Nuestro objetivo es aprender a resolver un tipo especial de sistemas de ecuaciones, los llamados lineales y que definiremos después. Por ahora, veamos algunos conceptos.

1. ♣ *Definiciones.* *Una* ***ecuación*** *es una igualdad en la cual hay una o varias incógnitas o variables que se denotarán por letras* $x, y, z, ...$ *o letras con subíndices como* $x_1, x_2, ...,$ *y que toman valores en el conjunto universal. Dado un conjunto universal, una* ***identidad*** *es una ecuación que siempre es verdadera, es decir, es válida para todos los elementos del conjunto universal. Por ejemplo:*

$3x + 1 = 2$, *donde* x *toma valores en el conjunto de los reales, es una ecuación.*

$3x = 2x + x$ *es una identidad sobre cualquier conjunto que admita suma.*

$5 = 4 + 1$ *es una identidad pues tenemos tanto a la derecha como a la izquierda del igual dos símbolos que representan el mismo objeto.*

Una ***solución*** *de una ecuación con una variable es un número que al ser reemplazado en la ecuación en el lugar de la variable convierte la ecuación en una identidad. Una solución de* $3x + 1 = 4$ *es* $x = 1$ *porque al reemplazar* 1 *en lugar de* x *en la ecuación, ésta se convierte en* $3(1) + 1 = 4$ *que es la identidad* $4 = 4$. *El conjunto solución de una identidad es todo el conjunto universal. Para verlo, tomemos la identidad* $5 = 4 + 1$. *Si reescribimos dicha identidad como* $5 + x = 4 + 1 + x$, *nos damos cuenta de que es una ecuación que es válida para cualquier número real, por lo que su conjunto solución es el conjunto de los números reales.*

Un ***sistema de ecuaciones*** *es un conjunto de ecuaciones cuyas incógnitas toman valores en el mismo conjunto de números, por ejemplo y haciendo referencia a los números reales, el siguiente es un sistema de ecuaciones:*

$$\left\{ \begin{array}{l} 3x + 4y = 7 \\ 4x + 3y = 7 \end{array} \right. \quad \textit{o simplemente} \quad \begin{array}{l} 3x + 4y = 7 \\ 4x + 3y = 7 \end{array}$$

También podemos escribir todas las ecuaciones de un sistema en un mismo renglón separando éstas por medio de una coma.

Una ***solución*** *de un sistema de ecuaciones es una asignación de números a las variables tal que al reemplazar las variables por sus valores asignados de forma simultánea en todas las ecuaciones, cada una de éstas se convierte en una identidad. Una solución del sistema anterior es* $x = 1, y = 1$ *pues estos valores convierten al sistema de ecuaciones en el sistema de identidades*

$$\begin{cases} 3+4=7 \\ 4+3=7 \end{cases}$$

Es posible que un sistema no tenga soluciones, o que exista una única solución o que existan muchísimas. El ***conjunto solución*** *de una ecuación o de un sistema de ecuaciones es la reunión de todas las soluciones posibles.* ***Resolver un sistema de ecuaciones*** *consiste en hallar el conjunto solución y describirlo de manera simplificada y entendible. Diremos, por ejemplo, que el conjunto solución de un sistema es una línea recta o un plano. Y en general, un sistema de ecuaciones se denomina lineal cuando está asociado a rectas, planos o a sus generalizaciones. Por dicha razón, un* ***sistema lineal*** *puede reconocerse porque contiene únicamente sumas de variables que han sido multiplicadas por números reales, pero no aparecen, digamos, multiplicaciones de variables o variables con exponentes diferentes de uno.*

2. ***Ejemplo*** *El sistema*

$$\begin{cases} 3x^2+4y=7 \\ 3xy=7 \end{cases}$$

no es lineal porque la primera ecuación contiene una variable al cuadrado y además la segunda tiene una multiplicación de dos variables. Todos los ejemplos que siguen son de sistemas lineales.

Para demostrar que un conjunto S es el conjunto solución hay que demostrar que cada elemento de S es solución y que todas las soluciones están en S. En general, eso es difícil y es la razón de ser de este curso. Nuestro punto de partida es notar lo siguiente: la solución de un sistema de ecuaciones se da en términos de igualdades que también definen otro sistema de ecuaciones. Por lo tanto, resolver un sistema de ecuaciones no es más que transformarlo mediante procedimientos adecuados de tal manera que a cada paso sea más clara la solución. Se termina cuando ya no haya forma de añadir claridad, lo cual corresponde a despejar los valores de las incógnitas. Con esto podemos entender la razón de las siguientes definiciones:

3. ♣ ***Definición.*** *Decimos que dos sistemas de ecuaciones son* ***equivalentes*** *si tienen la misma solución.*

4. ***Ejemplo*** *Decir que el sistema*

$$\begin{cases} 3x + 4y = 7 \\ 4x + 3y = 7 \end{cases}$$

y el sistema

$$\begin{cases} x = 1 \\ y = 1 \end{cases}$$

son equivalentes es lo mismo que decir que el primer sistema tiene una solución y que ésta es única. Esto es cierto debido a que, como lo veremos en el capítulo 2, cada ecuación del primer sistema corresponde a una línea recta que no es paralela a la de la otra ecuación. Como dos líneas que no son paralelas se cortan en un único punto, dicho punto es la solución y es única.

En general, el sistema que indique la solución tiene que ser por pura lógica equivalente al sistema original.

5. ♣ ***Definición.*** *Decimos que un procedimiento para transformar un sistema de ecuaciones en otro es **conservativo** si transforma un sistema de ecuaciones cualquiera en otro equivalente.*

Todo procedimiento que aspire a ser conservativo debe cumplir el requisito, ley o regla del **balance**: puesto que una ecuación representa una balanza bien equilibrada; cuando se haga una operación en un lado de una ecuación debe hacerse lo mismo en el otro lado de la misma ecuación. Por ejemplo, si multiplicamos un lado de la ecuación $3x = 1$ por $1/3$, el otro lado también debe multiplicarse por $1/3$ y nos queda $x = 1/3$. Sin embargo, hay que tener cuidado:

6. ***Contraejemplos*** *Hay procedimientos que hacen lo mismo a cada lado de la ecuacion pero no son conservativos porque cambian el conjunto solución:*

a) Un procedimiento no conservativo se ilustra partiendo de la identidad

$$\frac{2+3}{5} = 1 = \frac{1}{1}$$

Esta identidad tiene como conjunto solución al conjunto de los números reales. Sumamos ahora 1 en los numeradores a cada lado de la ecuación y obtenemos

$$\frac{2+3+1}{5} = \frac{1+1}{1}$$

$$\frac{6}{5} = \frac{2}{1} = 2$$

lo cual es una contradicción, cuyo conjunto solución es el vacío. Es por ello que si hacemos la misma operación sobre cada lado de una ecuación, cada lado debe tomarse de forma indivisible, como un todo. Un sistema que contenga una contradicción se denomina *inconsistente.*

b) Elevar al cuadrado cada lado de una ecuación no es un procedimiento conservativo:

Tomamos la ecuación $x = 1$, que tiene como única solución $x = 1$. Elevamos al cuadrado en ambos lados: $x^2 = 1$ que se resuelve con $x = 1$ ó con $x = -1$. El conjunto solución no se conservó. Decimos que nuestro procedimiento creó una ***solución espuria o falsa o no auténtica****. En algunos casos los procedimientos no conservativos que crean soluciones falsas son más fáciles, pero si se usan, hay que descartar las soluciones espurias, como aquí que deberíamos descartar $x = -1$.*

c) Multiplicar por cero a ambos lados de una ecuación no es un procedimiento conservativo. Veamos por qué: tomemos la ecuación $x = 3$. Su única solución es 3. Multipliquemos por cero en todo lado. Nos queda $0 = 0$, cuyo conjunto solución es todo $\mathbb{R}$ pues es una identidad equivalente a $x = x$. Concluimos que multiplicar por cero no conserva soluciones.

7. ***Fabricando procedimientos conservativos.***

¿Por qué al multiplicar a lado y lado de una ecuación por cero se crea una infinidad de soluciones espurias mientras que al multiplicar por $1/3$ no? Hay varias maneras de responder y una de ellas es decir que multiplicar por cero es una operación no invertible, cuyo significado lo podemos entender si pensamos sobre un ejemplo:

Sea $3x = 1$. Multiplicando por $1/3$ en ambos lados llegamos a $x = 1/3$. Si partimos de $x = 1/3$ podemos restaurar la ecuación original multiplicando por 3. Por lo tanto, multiplicar por $1/3$ es reversible. Pero si partimos de $x = 3$ y multiplicamos por cero en ambos lados, llegamos a $0 = 0$. Ahora bien, ninguna operación algebraica puede transformar la ecuación $0 = 0$ en $x = 3$, por lo que la operación multiplicar por cero no es reversible, no tiene inversa. Vemos que los procedimientos conservativos son aquellos que son reversibles, que tienen inverso.

8. ***Ejemplo*** *Los siguientes dos procedimientos son conservativos 1) Sumar un número a ambos lados de una ecuación, pues el procedimiento inverso es restar el mismo número a ambos lados. 2) Multiplicar a lado y lado de una ecuación por un número distinto de cero, pues su inverso es dividir por el mismo número a ambos lados.*

9. ***Ejercicio*** *Demuestre que el encadenamiento de dos procedimientos conservativos da un procedimiento conservativo.*

10. ***Ejemplo*** *Resolvamos la ecuación $ax + b = c$, donde suponemos que $a \neq 0$.*

Para despejar la x, quitamos la b sumando a ambos lados $-b$ y después quitamos la a multiplicando en todos lados por $1/a$ y obtenemos $x = (c - b)/a$. Como usamos únicamente procedimientos conservativos, esta es una solución y es la única.

Es fácil liberarse de la preocupación de la conservación del conjunto solución si podemos demostrar que la solución al sistema original es única. En ese caso, podemos utilizar cualquier método para hallar una solución y si al verificarla nos queda una identidad, entonces podemos estar seguros de haber encontrado todo lo que era necesario encontrar. Esa es una razón por la cual nos interesamos en la unicidad de la solución.

11. *Advertencias.* *Hay sistemas que tienen varias soluciones y hay otros que no tienen solución.*

Por ejemplo, el sistema

$$\begin{cases} x + 4y + 3z = 7 \\ x + 3y + 4z = 7 \end{cases}$$

tiene varias soluciones. Una es $x = 0, \quad y = 1, \quad z = 1$. Otra solución es $x = -7, \quad y = 2, \quad z = 2$. Pero el sistema siguiente no tiene solución pues es una contradicción:

$$\begin{cases} x = 1 \\ x = 2 \end{cases}$$

Estas dos advertencias permiten ver que cuando se hable de unicidad de una solución, se debe demostrar. La estructura típica de una demostración de unicidad puede verse en el teorema siguiente:

12. $\diamond$ ***Teorema.*** *La ecuación $ax + b = c$ tiene una solución y esa solución es única para a, b, c números reales y $a \neq 0$.*

Demostración. Ya sabemos que hay al menos una solución: $s = (c-b)/a$, lo cual puede verificarse por sustitución. Supongamos ahora que existen dos soluciones s y z. Esto significa que $as + b = c$ y que $az + b = c$. Restar lado a lado la segunda ecuación de la primera es un procedimiento conservativo pues es invertible y da $as+b-(az+b) = c-c$ que se puede simplificar en $a(s - z) = 0$. Ahora bien, si $a \neq 0$, podemos multiplicar por su inverso multiplicativo a ambos lados, lo cual es un procedimiento invertible y por lo tanto conservativo, y nos queda $s - z = 0$. Sumando z en ambos lados, lo cual es conservativo por ser invertible, queda: $z = s$. Es decir, si hay dos soluciones y estas son iguales: la solución es única. ■

Enseguida vamos a aprender a utilizar un procedimiento que produce soluciones de sistemas de ecuaciones, lo cual es algo que uno puede verificar, digamos, por sustitución. Pero por ahora no podremos probar que nuestro procedimiento nos permite hallar todas las soluciones.

1.1. Algoritmo solucionador

La palabra *algoritmo* es de origen árabe y honra al matemático árabe de principios del siglo IX Muhammad ibn Musa al-Jwarizmi (Gran Enciclopedia Larousse, 1983, vol I) y se emplea hoy en día con un sentido muy preciso y técnico que en definitiva sirve para designar un conjunto de instrucciones ejecutables por un computador.

Un método que puede usarse para solucionar un sistema de dos ecuaciones lineales con dos incógnitas es despejar una variable de ambas ecuaciones, igualar sacando una ecuación en una sola variable, despejarla y después reemplazar para sacar el valor de la otra variable.

El método anterior puede generalizarse a sistemas con tres incógnitas y con mucho trabajo uno podría tratar de usar el mismo método para sistemas con cuatro variables. Los problemas realistas involucran muchas variables. Eso querría decir que con el método de igualar variables despejadas estos problemas serían irresolubles en la práctica. Para salir del embrollo, debemos desarrollar un método que permita solucionar con igual seguridad tanto sistemas con 2 incógnitas como con un millón. Por supuesto que no estamos pensando en resolver nosotros tales sistemas, sino en saber cómo programar un computador para que él lo haga. Vamos a aprender a combinar ecuaciones de un sistema lineal multiplicando por números no nulos y sumando o restando para resolver sistemas de ecuaciones. Pero antes, vale la pena que nos preguntemos: ¿tiene eso algún peligro?

Por supuesto que podemos sumar o restar ecuaciones lado a lado y el resultado es una ecuación. Igualmente, dos ecuaciones pueden multiplicarse separadamente por números distintos no nulos y después sumarse o restarse y el resultado será una ecuación. Pero, ¿ el procedimiento será conservativo? Veamos que eso no es siempre cierto. El ejemplo es tomar $x = 1$, cuya solución es 1, y restarla de sí misma. El resultado es $0 = 0$, cuya solución es todo $\mathbb{R}$. Eso implica que eso de sumar o restar ecuaciones no garantiza de por sí que uno no cambie el conjunto solución. Así que más nos vale que procedamos con cautela. Todo lo que hay que hacer es asegurarse de no perder información: podemos restar una ecuación de sí misma, pero no podemos olvidar la ecuación que originó todo. Así, por ejemplo, es correcto decir que el sistema de una sola ecuación $x = 1$ es equivalente al sistema conformado por las dos ecuaciones $x = 1$ y $0 = 0$, cuya solución es $x = 1$. La idea entonces es combinar las ecuaciones de tal manera que no se pierda información pero sin aumentar el número de ecuaciones. El siguiente ejemplo ilustra el procedimiento.

13. ***Ejemplo*** *Resolvamos el sistema*

$$\begin{cases} 3x + 4y = 7 \\ 4x + 3y = 7 \end{cases}$$

Estas dos ecuaciones pueden multiplicarse separadamente por números distintos y después sumarse o restarse con el objetivo de aniquilar una variable. Concretamente, multiplicamos ambos lados de la primera ecuación por 4 y separadamente multiplicamos ambos lados de la segunda ecuación por 3, para que los coeficientes de la x en ambas ecuaciones queden iguales. La intención que tenemos es restar las dos ecuaciones para aniquilar la x. Procedamos:

$$\begin{cases} 4(3x + 4y) = 4(7) \\ 3(4x + 3y) = 3(7) \end{cases}$$

Ejecutando:

$$\begin{cases} 12x + 16y = 28 \\ 12x + 9y = 21 \end{cases}$$

Ahora las restamos. Es decir, producimos un nuevo sistema que sea equivalente al original, pero que nos permita despejar la y. Para ello, restamos la segunda ecuación de la primera y el resultado lo ponemos en lugar de la segunda ecuación:

$$\begin{cases} 12x + 16y = 28 \\ 16y - 9y = 28 - 21 \end{cases}$$

que se simplifica en

$$\begin{cases} 12x + 16y = 28 \\ 7y = 7 \end{cases}$$

Multiplicamos la primera ecuación por 1/4 y la segunda por 1/7 en ambos lados y nos produce

$$\begin{cases} 3x + 4y = 7 \\ y = 1 \end{cases}$$

Reemplazamos la solución $y = 1$ en la primera ecuación, $3x + 4y = 7$ y nos queda $3x+4 = 7$ que inmediatamente da $x = 1$. Ahora bien, podemos verificar por sustitución directa que $x = 1$ y $y = 1$ en realidad forman una solución al sistema. Pero, ¿nos faltarán soluciones? No, no nos faltan soluciones pues hemos utilizado procedimientos conservativos, invertibles, tanto para manipular cada ecuación como para tratar con todo el sistema, no perdiendo información. En este momento no tenemos la intención de ser muy rigurosos en justificar el aspecto conservativo de nuestros procedimientos, pues en el capítulo de la inversa podremos estudiar el asunto con comodidad y eficacia.

14. ***Ejercicio*** *Resolver el sistema*

$$\begin{cases} 3x + 6y = 9 \\ 4x + 7y = 11 \end{cases}$$

1.2. Sobre las aplicaciones

Los sistemas de ecuaciones lineales salen naturalmente de la vida real. Traducir dichos problemas en ecuaciones no siempre es una tarea sencilla. Sin embargo, en algunos casos esto puede lograrse aplicando la siguiente regla y si la regla no es suficiente, en todos los casos esta regla evitará errores crasos.

15. ***Regla de oro:***

> Al plantear las ecuaciones de un problema debe tenerse en cuenta que las operaciones indicadas tengan sentido y que todos los términos de cada ecuación se refieran al mismo tipo de elemento y que estén en las mismas unidades.

Por ejemplo, sumar mg de vitamina A con mg de vitamina A tiene mucho sentido. Pero sumar concentración de sal al 20 % con concentración de sal al 30 % no tiene sentido porque mezclar dos vasos de salmuera con las concentraciones indicadas no produce una solución con concentración al 50 %. De forma semejante, sumar onzas con gramos no está permitido, debe hacerse una reducción a las mismas unidades.

16. ***Ejemplo*** *En un resguardo hay dos tipos de pájaros,* P1 *y* P2*. Ellos se alimentan por su cuenta de mosquitos y semillitas pero para asegurar su perfecta nutrición se les administran dos tipos de comida* C1 *y* C2*. Cada pájaro del tipo* P1 *consume por semana* 3 *onzas del alimento* C1 *y* 4 *onzas del alimento* C2*. Cada pájaro del tipo* P2 *consume* 6 *onzas de alimento* C1 *y* 7 *onzas del alimento* C2*. Si se suministran* 9 *onzas del alimento* C1 *y* 11 *onzas del alimento* C2*, ¿cuántos pájaros podrán coexistir?, ¿cómo se modifica la situación si se aumenta el alimento* 100 *veces?*

Solución: notemos que se usan las mismas unidades, onzas. En cuanto a qué está permitido sumar y qué no, tenemos lo siguiente: si uno se prepara para abordar un avión en el cual le cobran a uno por kilos de equipaje, tendría mucho sentido sumar cantidad del producto uno con cantidad del producto dos. Pero no en el contexto del problema, sería lo mismo que sumar diamantes y flores. En cambio sí se puede sumar cantidad del producto consumido por los pájaros de tipo uno con la cantidad consumida del mismo producto por los pájaros de tipo dos. Sea x la cantidad de pájaros P1 y y la del tipo P2 que pueden coexistir. Por tanto, x pájaros del tipo P1 consumirán $3x$ onzas del alimento C1, en tanto que y pájaros P2 consumirán $6y$ onzas del alimento C1. Entre ambos tipos de pájaros deben agotar el alimento C1. Por tanto, $3x + 6y = 9$. Obsérvese la coherencia en las unidades y del buen sentido de la suma: todos los términos de esta ecuación están en onzas y se refieren al alimento C1. Estaría mal formulada la ecuación si todo estuviese en onzas pero en un término hubiese alimento C1 y en otro C2. De igual modo, con el alimento C2, y guardando la coherencia de unidades, tenemos: $4x + 7y = 11$.

Lo dicho queda más claro si lo presentamos en la siguiente tabla:

Consumo de alimento			
	x Pájaros tipo 1	y Pájaros tipo 2	Totales
Alimento C1	$3x$	$6y$	9
Alimento C2	$4x$	$7y$	11

Las dos ecuaciones siguientes definen el sistema:

$$\begin{cases} 3x + 6y = 9 \\ 4x + 7y = 11 \end{cases}$$

con solución $x = 1$ y $y = 1$. El alimento alcanza para exactamente un pájaro de cada tipo. Si se multiplica el alimento de cada tipo por 100, la cantidad de pájaros de cada tipo se multiplicará por 100. Por lo tanto, en el segundo caso, se tendrán 100 pájaros del tipo P1 y otros 100 del tipo P2. En este último caso quedaría el sistema

$$\begin{aligned} 3x + 6y &= 900 \\ 4x + 7y &= 1100. \end{aligned}$$

Por supuesto que uno podría alimentar 50 pájaros de cada tipo y el sobrante echarlo a la basura. Habría un desperdicio. Pero, ¡caramba!, ¿qué pasaría si en lugar de alimentar 100 y 100 deseáramos aprovechar 'mejor' las posibilidades y decidiéramos cultivar 300 pájaros del tipo P1 y nada del tipo P2? de esa forma podríamos sacarle el máximo provecho al alimento C1, pero tendríamos una deficiencia del alimento C2 de 100 unidades, pues cada pájaro del tipo P1 consume 4 unidades del alimento C2. Para no correr el riesgo de que se mueran por causa de desnutrición, podemos decidir alimentar 275 pájaros del tipo P1 y nada del tipo 2. Eso causaría un despilfarro del alimento C1, pues sobrarían $900 - 3 \times 275 = 75$.

Nos aventuramos a asegurar que nuestra solución ($x = 100$ y $y = 100$) utiliza los recursos minimizando la basura, el desperdicio. Esto sucede en el resguardo. Pero, ¿qué sucederá en la naturaleza?, ¿se comportará la naturaleza como si quisiera minimizar el desperdicio? En realidad, no se acostumbra hoy en día a vociferar tales preguntas tan llenas de antropomorfismos. Pero podemos decir lo siguiente: si un ser viviente utiliza mejor los desperdicios que otro, el primero dejará más hijos y llenará la tierra, de tal forma que es difícil imaginar que haya desperdicios no utilizados. Pero supongamos que no hubiese en el mundo más que pájaros del tipo P1 y pájaros del tipo P2: ¿Qué pasaría? Observando que los pájaros del tipo P2 tienen más o menos el doble del tamaño de los pájaros del tipo P1, pues suponemos que el apetito es proporcional al tamaño, y que entre todos tienen casi las mismas necesidades, uno podría decir que francamente no importa si se tienen 2 pájaros del tipo P1 o uno del tipo P2. Esa neutralidad permitiría que el sistema evolucionará caóticamente: las proporciones variarían con el tiempo de un lado al otro como un barco a la deriva.

Este tipo de temas relacionados con la evolución se consideraban antiguamente como dominio exclusivo de la biología pero sucede que, de un lado, la biología se ha vuelto hoy en día materia de seguridad planetaria y por lo tanto de interés universal, y del otro, que se ha inventado una metodología que simula la evolución biológica para resolver todo tipo de problemas de matemáticas, ciencia, tecnología y arte. El magnífico sitio web *http:www.evoljava.com* está a su disposición con material pedagógico para aprender desde cero a programar la evolución usando el lenguaje Java. Esto es complejo e importante no sólo por su aspecto técnico o científico sino además porque a muchas personas les interesa el problema del origen de la vida y de las especies en su relación con la evolución. Y entre menos dogmático se desee ser, es mucho más difícil. Este tema podría llegar a ser una profesión de por vida.

17. $\boxed{\textbf{\textit{Ejercicio}}}$ *Hallar al menos una solución de cada uno de los siguientes sistemas:*

$$\text{a)} \left\{ \begin{array}{l} 2x - 5y = -3 \\ 7x - 4y = 3 \end{array} \right.$$

$$\text{b)} \left\{ \begin{array}{l} 2x - 8y = -4 \\ 4x - 5y = 11 \end{array} \right.$$

$$\text{c)} \left\{ \begin{array}{l} 3x - 2y = 5 \\ 2x + 3y = 12 \end{array} \right.$$

18. $\boxed{\textbf{\textit{Ejercicio}}}$ *Deseamos producir dos tipos de jabones, uno industrial y el otro para manos. Ambos contienen desengrasante y suavizante pero varían en sus proporciones. El tipo industrial necesita 120 gramos de desengrasante y 30 gramos de suavizante mientras que el de manos requiere 40 gramos de desengrasante y 30 de suavizante. Si se tiene en la bodega 160 kilos de desengrasante y 60 de suavizante, ¿cuántos jabones de cada uno se logra hacer?*

1.3. Ecuaciones en forma matricial

El método que hemos aplicado en la sección anterior para resolver un sistema de ecuaciones no se preocupa por distinguir cuál es el sistema original y cuál es el modificado. Pareciera que ignorar el problema de si un procedimiento es conservativo o no, fuese la mejor manera de acabar con todos los problemas. Vamos a remediar eso de una vez por todas y lo haremos a muy bajo costo y de tal forma que el procedimiento pueda automatizarse. Para ello, aprendamos a escribir un sistema en forma matricial. Ahora bien, quedará claro que ya estamos trabajando con sistemas muy específicos. Los sistemas más generales que admiten la representación matricial que vamos a introducir se llaman **lineales**.

Usemos la siguiente convención: al escribir un sistema de ecuaciones se adjudica un lugar fijo a cada variable y todas las variables se escriben explícitamente. Por ejemplo, el sistema:

$4y + 3x = 7$
$4x + 3y = 7$
se reordena en:
$3x + 4y = 7$
$4x + 3y = 7$
Por otra parte el sistema
$3x = 7$
$4x + 3y = 7$
se explícita exhaustivamente en
$3x + 0y = 7$
$4x + 3y = 7$

19. ***Matrices.*** *Si convenimos en que las x van siempre en el primer lugar, y las y siempre van en el segundo lugar, y si cada variable va siempre en su lugar, entonces no hay necesidad de escribirlas. Simplemente se sobreentienden. Con esa convención el sistema*

$$\begin{cases} 3x + 4y = 7 \\ 4x + 3y = 7 \end{cases}$$

se reescribe en forma de arreglo de números

$$\begin{pmatrix} 3 & 4 & \vdots & 7 \\ 4 & 3 & \vdots & 7 \end{pmatrix}$$

A estos arreglos se les llama así: todo el conjunto se llama **matriz aumentada**, la parte izquierda que tiene dos columnas se denomina **matriz** y la parte derecha con una única columna **vector independiente**. La matriz corresponde a los números que acompañan a las incógnitas, y el vector independiente es el que queda al lado derecho del sistema de ecuaciones y que no tienen variables.

Denominamos **entrada** a un lugar específico de la matriz, por ejemplo, la entrada $(2,1)$ representa el coeficiente que en la segunda ecuación acompaña a la x o primera variable y que, en este caso, vale 4.

Solucionemos este sistema pero usando la nueva reescritura matricial. **Las reglas para encontrar soluciones a un sistema en forma matricial** son: Primera, como dos ecuaciones pueden intercambiarse de lugar, entonces dos renglones en su expresión matricial también. Segunda, una fila o renglón de una matriz puede multiplicarse, en todos lados, en todas las entradas, por cualquier número *diferente de cero*. Tercera, dos filas o renglones pueden sumarse o restarse columna por columna (es decir, las x con las x, las y con las y, etc.) Las dos últimas reglas en la práctica se utilizan al tiempo, como en la regla siguiente.

20. ***Regla aniquiladora.*** *Dos filas o renglones de una matriz pueden multiplicarse por separado en toda entrada por números distintos y después restarse o sumarse con el objetivo de aniquilar una variable, es decir, con el fin de convertir el coeficiente respectivo en cero.*

Apliquemos la regla aniquiladora para empezar a solucionar el sistema dado en su nueva reescritura matricial: en la segunda fila o renglón escribimos el resultado de multiplicar la primera ecuación por 4 menos la segunda fila por 3:

$$\begin{matrix} \\ 4R1 - 3R2 \rightarrow \end{matrix} \begin{pmatrix} 3 & 4 & \vdots & 7 \\ 4(3) - 3(4) & 4(4) - 3(3) & \vdots & 4(7) - 3(7) \end{pmatrix}$$

Obsérvese que las operaciones se indican, por fuera de la matriz, exactamente en el mismo renglón donde se ejecutan. Escribir el historial ayuda a corregir errores, los cuales suceden frecuentemente. Ejecutando las operaciones obtenemos:

$$\begin{pmatrix} 3 & 4 & \vdots & 7 \\ 0 & 7 & \vdots & 7 \end{pmatrix}$$

La segunda ecuación dice $7y = 7$. Por supuesto que debemos dividir por 7 para despejar y; esto correspone a:

$$(1/7)R2 \rightarrow \begin{pmatrix} 3 & 4 & \vdots & 7 \\ 0 & 1 & \vdots & 1 \end{pmatrix}$$

Hemos despejado $y = 1$. Para despejar la x procedemos, primero a igualar los coeficientes de las y en ambas ecuaciones. Para eso, multiplicamos la segunda ecuación por 4 y, a continuación, restamos las dos ecuaciones. El resultado de la resta se pone en la primera ecuación y no en la segunda, pues de lo contrario perderíamos la información ya despejada que dice $y = 1$.

$$R1 - 4R2 \rightarrow \begin{pmatrix} 3-0 & 4-4 & \vdots & 7-4 \\ 0 & 1 & \vdots & 1 \end{pmatrix}$$

Simplificamos:

$$\begin{pmatrix} 3 & 0 & \vdots & 3 \\ 0 & 1 & \vdots & 1 \end{pmatrix}$$

Despejamos la x:

$$(1/3)R1 \rightarrow \begin{pmatrix} 1 & 0 & \vdots & 1 \\ 0 & 1 & \vdots & 1 \end{pmatrix}$$

lo cual dice que ya pudimos despejar $x = 1, y = 1$. Esta matriz que queda en el lado izquierdo de este arreglo matricial, y que es cuadrada y tiene unos en la **diagonal** pero ceros en el resto del mundo se llama la matriz identidad y se nota $\mathbb{I}$ aunque a veces se le agrega un subíndice para indicar su tamaño.

Precaución: todos nuestros procedimientos han sido conservativos, invertibles, pero no por azar sino porque hemos sido cuidadosos conservando toda la información. No hubiese sido así si tomáramos, por ejemplo, el triple del primer renglón y lo hubiésemos puesto en el segundo. En general, cada nueva matriz debe tener:

21. ***Conclusión y definición.*** *Resolver un sistema de ecuaciones escrito en forma matricial equivale a combinar apropiadamente los renglones de la matriz, multiplicándolos por números diferentes de cero y sumando o restando para recuperar, si es posible, la matriz identidad en el lado izquierdo del arreglo matricial. Esta forma de combinar renglones, multiplicando por números, sumando y restando, se llama* combinación lineal.

En efecto, ¿qué es solucionar un sistema, por ejemplo, con 3 incógnitas, x, y, z? Es encontrar los valores de dichas incógnitas que satisfagan el sistema, digamos $x = 2$, $y = -1, z = 4$. Si escribimos esta solución en forma matricial tenemos:

$$\begin{pmatrix} 1 & 0 & 0 & \vdots & 2 \\ 0 & 1 & 0 & \vdots & -1 \\ 0 & 0 & 1 & \vdots & 4 \end{pmatrix}$$

Podemos ver que la matriz identidad está en el lado izquierdo de este arreglo. Este método de solucionar ecuaciones se denomina **eliminación Gauss-Jordan**. Pero para no difamar tácitamente a esos dos grandes hombres, más nos vale darnos cuenta de que hemos utilizado la siguiente restricción: en el renglón i de cada nueva matriz, siempre ponemos una combinación lineal que involucre el renglón i de la vieja matriz. Por ejemplo, en la fila uno de la nueva matriz ponemos 3 veces la primera más 4 veces la quinta. Pero no ponemos en la fila uno de la nueva matriz tres veces la tercera menos 4 veces la sexta. Hacemos esto para asegurar que nuestros algoritmos sean conservativos e invertibles.

22. ***Ejemplo*** *Tomemos un sistema representado por*

$$\begin{pmatrix} 2 & 3 & \vdots & 4 \\ 5 & 6 & \vdots & 7 \end{pmatrix}$$

Consideremos el procedimiento de poner en el segundo renglón $5R_1 - 2R_2$. *Este procedimiento es invertible y su inverso consiste en poner en el segundo renglón* $(5/2)R_1 - (1/2)R_2$. *Verifiquémoslo:*

Si aplicamos sobre el arreglo dado el primer procedimiento, obtenemos:

$$\begin{pmatrix} 2 & 3 & \vdots & 4 \\ 5(2)-2(5) & 5(3)-2(6) & \vdots & 5(4)-2(7) \end{pmatrix}$$

que es igual a

$$\begin{pmatrix} 2 & 3 & \vdots & 4 \\ 0 & 3 & \vdots & 6 \end{pmatrix}$$

Si a este arreglo le aplicamos el segundo procedimiento, obtenemos

$$\begin{pmatrix} 2 & 3 & \vdots & 4 \\ (5/2)(2)-0 & (5/2)(3)-(1/2)(3) & \vdots & (5/2)(4)-(1/2)(6) \end{pmatrix}$$

que es exactamente el arreglo original. Es decir, podemos recuperar la información inicial y por ende toda la información se ha conservado.

23. ***Advertencia y ejercicios.*** *Se puede llegar a la matriz identidad cuando existe la solución y ésta es única. Pero no siempre se puede llegar a la matriz identidad, porque no hay solución o tal vez porque hay muchas. Por ejemplo:*

a) Consideremos el sistema de ecuaciones compuesto por una sola ecuación

$$3x + 4y = 5$$

como este sistema tiene infinidad de soluciones de la forma $x = (5 - 4y)/3$, donde y puede tomar cualquier número real, no hay manera de decir x vale tanto y y vale tanto, es decir, no hay forma de llegar a la matriz identidad. Por lo tanto, escriba el sistema correspondiente en forma matricial usando una matriz 2×2 y argumente por qué no se puede llegar a la identidad.

b) Consideremos ahora el sistema de ecuaciones

$$\begin{cases} x = 5 \\ y = 2 \\ x = 4 \end{cases}$$

Vemos que este sistema es inconsistente. ¿Qué pasa cuando lo estudiamos matricialmente? A este sistema se le puede representar por una matriz 3×2, con 3 renglones y dos columnas, la cual puede transformarse en una matriz que contiene tanto la identidad 2×2 como una contradicción (ejercicio).

En resumen, se llega a la identidad cuando la solución existe y es única, de lo contrario o no se llega o se llega pero con una contradicción. Por ahora aceptamos esto como un resultado empírico, pero conservamos el desafío de probarlo.

24. **Ejercicio** *Resolver todos los incisos del ejercicio 17 por medio de matrices.*

25. ***Ejemplo:*** *Resolvamos el sistema siguiente*

$$\begin{pmatrix} 1 & 0 & -1 & \vdots & 0 \\ 0 & 1 & -1 & \vdots & 0 \\ -2 & 0 & 1 & \vdots & -1 \end{pmatrix}$$

Primero ponemos ceros por fuera de la diagonal en la columna uno. Sólo queda recuperar un cero en la tercera fila:

$$\begin{array}{r} \\ \\ R3 + 2R1 \rightarrow \end{array} \begin{pmatrix} 1 & 0 & -1 & \vdots & 0 \\ 0 & 1 & -1 & \vdots & 0 \\ 0 & 0 & -1 & \vdots & -1 \end{pmatrix}$$

Ahora ponemos ceros por fuera de la diagonal de la columna 2: ya está hecho. Multipliquemos la tercera fila por menos uno:

$$\begin{array}{r} \\ \\ -R3 \rightarrow \end{array} \begin{pmatrix} 1 & 0 & -1 & \vdots & 0 \\ 0 & 1 & -1 & \vdots & 0 \\ 0 & 0 & 1 & \vdots & 1 \end{pmatrix}$$

Pongamos ceros por fuera de la diagonal de la columna 3

$$\begin{array}{r} R1 + R3 \rightarrow \\ R2 + R3 \rightarrow \\ \\ \end{array} \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 \\ 0 & 1 & 0 & \vdots & 1 \\ 0 & 0 & 1 & \vdots & 1 \end{pmatrix}$$

Hemos recuperado la matriz identidad, la solución es $x = y = z = 1$. De antemano se planeó esta respuesta para que la aritmética quedara sencilla de entender, pero un sistema puede planearse para que tenga cualquier respuesta.

26. *Ejemplo* *Hemos utilizado en el ejercicio anterior una manera estándar para resolver un sistema en forma matricial, recuperando la matriz identidad de manera muy ordenada, como para programar un computador. Cuando la identidad puede recuperarse, este método siempre funciona. Con todo, un problema específico bien puede resolverse creativamente con un poco de desorden. Resolvamos el sistema:*

$$\begin{pmatrix} 2 & -2 & -1 & \vdots & -5 \\ 0 & 3 & -2 & \vdots & 0 \\ -2 & 1 & 1 & \vdots & 3 \end{pmatrix}$$

Procedemos:

$$\begin{array}{r} R1 + R2 \rightarrow \\ \\ R3 + R1 \rightarrow \end{array} \begin{pmatrix} 2 & 1 & -3 & \vdots & -5 \\ 0 & 3 & -2 & \vdots & 0 \\ 0 & -1 & 0 & \vdots & -2 \end{pmatrix}$$

La tercera ecuación dice que $y = 2$. Ya no volvemos a tocar esa ecuación y despejamos x y z de las otras ecuaciones:

$$\begin{array}{r} \\ R2 + 3R3 \rightarrow \\ \\ \end{array} \begin{pmatrix} 2 & 1 & -3 & \vdots & -5 \\ 0 & 0 & -2 & \vdots & -6 \\ 0 & -1 & 0 & \vdots & -2 \end{pmatrix}$$

$$\begin{array}{r} \\ -R2/2 \rightarrow \\ -R3 \rightarrow \end{array} \begin{pmatrix} 2 & 1 & -3 & \vdots & -5 \\ 0 & 0 & 1 & \vdots & 3 \\ 0 & 1 & 0 & \vdots & 2 \end{pmatrix}$$

$$R1 - R2 + 3R3 \rightarrow \begin{pmatrix} 2 & 0 & 0 & \vdots & -5-2+9=2 \\ 0 & 0 & 1 & \vdots & 3 \\ 0 & 1 & 0 & \vdots & 2 \end{pmatrix}$$

$$(1/2)R1 \rightarrow \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 \\ 0 & 0 & 1 & \vdots & 3 \\ 0 & 1 & 0 & \vdots & 2 \end{pmatrix}$$

que nos dice que $x = 1, z = 3, y = 2$.

27. $\boxed{\textbf{\textit{Ejercicio}}}$ *Primero verifique que todos los siguientes sistemas de ecuaciones escritos en forma matricial tienen la misma solución $x = y = z = 1$, lo cual ha sido diseñado así para que la aritmética resulte sencilla. Después de verificar la solución, resolver todos esos sistemas por Gauss-Jordan.*

$$a) \begin{pmatrix} 1 & 0 & 1 & \vdots & 2 \\ 0 & 1 & 1 & \vdots & 2 \\ 1 & 1 & 0 & \vdots & 2 \end{pmatrix}$$

$$b) \begin{pmatrix} 2 & 0 & 1 & \vdots & 3 \\ 0 & 3 & 1 & \vdots & 4 \\ 1 & 4 & 0 & \vdots & 5 \end{pmatrix}$$

$$c) \begin{pmatrix} 3 & 0 & 2 & \vdots & 5 \\ 0 & 4 & 2 & \vdots & 6 \\ 2 & 5 & 0 & \vdots & 7 \end{pmatrix}$$

$$d) \begin{pmatrix} 3 & 0 & 2 & \vdots & 5 \\ 2 & 4 & 7 & \vdots & 13 \\ 2 & 0 & 5 & \vdots & 7 \end{pmatrix}$$

$$e) \begin{pmatrix} 3 & -2 & 2 & \vdots & 3 \\ 2 & 4 & 7 & \vdots & 13 \\ 2 & -4 & 5 & \vdots & 3 \end{pmatrix}$$

28. $\boxed{\textbf{\textit{Ejercicio}}}$ *Descubra el secreto para lograr que todos los sistemas contengan la misma solución* $(1, 1, 1)$.

29. $\boxed{\textbf{\textit{Ejercicio}}}$ *Diseñe un método para inventar un sistema de ecuaciones que con seguridad tenga una solución prefijada, por ejemplo, $x = 2, y = 5, z = -3$. Póngalo en práctica y resuélvalo para verificar la honestidad del método.*

30. $\boxed{\textbf{\textit{Ejercicio}}}$[1] *En ingeniería química, química, bioquímica y biología molecular es común encontrar el problema de fabricar una solución en cantidad y prescripción de concentraciones dadas teniendo otras soluciones ya preparadas en cantidades y prescripciones determinadas. Para resolver el problema lo único que hay que tener en cuenta es que la cantidad de cada componente se conserva individualmente. Por lo tanto, por cada componente aparece una ecuación que expresa su conservación: en un lado de la ecuación está la contabilidad del componente antes de producir la mezcla y en el otro la contabilidad del componente después de la mezcla. Un caso concreto podría ser: tenemos 60 ml de una mezcla M1 que tiene 40 % del componente A y 60 % del componente B. Además, tenemos 80 ml de la mezcla M2 que contiene 70 % del componente A y 30 % del componente B. Diga qué cantidad de la mezcla M1 y qué de M2 deben combinarse para fabricar, si se puede,*

a) 100 ml de una tercera solución que tenga 10 % del componente A y 90 % del componente B.

b) 100 ml de una tercera solución que tenga 50 % del componente A y 50 % del componente B.

c) 70 ml de una tercera solución que tenga 50 % del componente A y 50 % del componente B.

d) 100 ml de una tercera solución que tenga 55 % del componente A y 45 % del componente B.

e) 70 ml de una tercera solución que tenga 80 % del componente A y 20 % del componente B.

Ayudas: 1) Tenga en cuenta el proceder natural en el laboratorio, primero se hacen las cuentas para saber las fracciones de cada mezcla y después se mira en los frascos sobre el estante para ver si hay suficiente material. 2) Recuerde guardar en todos los términos el mismo tipo, las mismas unidades y que las operaciones tengan sentido. 3) Las concentraciones no se pueden sumar. Sin embargo, a veces se puede decidir de antemano que ciertas concentraciones son imposibles de obtener a partir de mezclas dadas.

31. $\boxed{\textbf{\textit{Ejercicio}}}$ *Inventar problemas semejantes al anterior pero con 3, 4 ó más componentes.*

1.4. Sistemas de m ecuaciones con n incógnitas

Raramente resolveremos en este texto un sistema con más de tres incógnitas, pero es importante tener en cuenta que en las aplicaciones se pueden encontrar sistemas con muchas incógnitas, digamos 10, 20 ó 50. Un ejemplo muy claro que promete sistemas

[1] Inventado propuesto por Jorge Palacio

con millones de ecuaciones es el siguiente: teniendo en cuenta que una ecuación puede interpretarse como una ley de conservación, consideremos una red eléctrica que alimenta una región, o un país o una interconección eléctrica para todo un continente. Para cada nodo, la corriente que entra es igual a la corriente que sale. Es decir, por cada nodo hay una ecuación. El problema consiste en hacer que en cada tomacorriente de cada casa o fábrica haya suficiente corriente para alimentar un motor o un televisor.

Por eso, vale la pena preocuparse por desarrollar métodos que sean al mismo tiempo eficientes e informativos. Y también es necesario pensar si no estaremos causando errores de aproximación que se encadenan unos con otros y que a la larga producen errores no despreciables. Este último problema se estudia en los cursos de análisis numérico. Nosotros ya hemos visto una variante del algoritmo de Gauss-Jordan, con el cual se puede resolver cualquier sistema, pero como ese método es de uso corriente para programar computadores, es importante ver una formulación muy precisa. Por eso, esta sección está dedicada a su formalización.

32. ♣ ***Definiciones.*** *Sea M una matriz aumentada que codifica un sistema de ecuaciones de tal forma que cada ecuación del sistema corresponde a un renglón de la matriz. Las siguientes son* ***operaciones elementales*** *entre renglones de M :*

1) Intercambiar dos ***renglones o filas*** *cualesquiera.*

2) Multiplicar un renglón cualquiera por un escalar no nulo.

3) Reemplazar un renglón cualquiera por la suma del mismo renglón con otro renglón.

- *En la práctica, combinamos las reglas 2 y 3 en una sola: reemplazar un renglón cualquiera por la suma de un múltiplo escalar no nulo del mismo renglón con otro múltiplo escalar de otro renglón.*

33. ***Ejercicio*** *Demuestre que toda operación elemental es invertible y que por lo tanto produce un procedimiento conservativo.*

Decir que las operaciones elementales no alteran las soluciones y que las operaciones elementales son suficientes para resolver sistemas lineales es lo mismo que decir lo siguiente:

34. ⋄ ***Teorema.*** *Dos sistemas son equivalentes si y sólo si la matriz aumentada que representa a un sistema se puede obtener de la matriz aumentada que representa el otro sistema mediante operaciones elementales por renglones.*

35. ***Ejemplo*** *Los siguientes dos sistemas son equivalentes:*

Primer sistema

$$\begin{cases} x - y = 1 \\ 2x + y = 2 \end{cases}$$

Segundo sistema

$$\begin{cases} x = 1 \\ y = 0 \end{cases}$$

En efecto, la matriz aumentada del primer sistema puede transformarse con operaciones elementales en la matriz aumentada del segundo sistema:

$$\begin{pmatrix} 1 & -1 & \vdots & 1 \\ 2 & 1 & \vdots & 2 \end{pmatrix}$$

$$R2 - 2R1 \rightarrow \begin{pmatrix} 1 & -1 & \vdots & 1 \\ 0 & 3 & \vdots & 0 \end{pmatrix}$$

$$R2/3 \rightarrow \begin{pmatrix} 1 & -1 & \vdots & 1 \\ 0 & 1 & \vdots & 0 \end{pmatrix}$$

$$R1 + R2 \rightarrow \begin{pmatrix} 1 & 0 & \vdots & 1 \\ 0 & 1 & \vdots & 0 \end{pmatrix}$$

Lo que deseamos es, por supuesto, encadenar operaciones elementales para llegar a una solución del sistema. Si la solución es única, llegaremos a la identidad. Pero, ¿qué pasa cuando no hay solución o si no existe solución? La respuesta es que no llegamos a la identidad, pero ¿a dónde llegamos? Hay varios estilos de terminar los problemas, pero nosotros fijaremos la atención en un método sencillo y eficaz, el cual está basado en las matrices escalonadas que son aquellas cuyos ceros forman una escalera por debajo de la diagonal y que debe llegar, por lo menos, hasta inmediatamente abajo de la diagonal.

36. ♣ ***Definición.*** *Una matriz, aumentada o no, M, se dice que está en forma* ***escalonada*** *si:*

1) Los renglones compuestos únicamente por ceros están en la parte inferior de la matriz.

2) El número de ceros antes del primer elemento diferente de cero de un renglón es mayor que el del renglón inmediatamente superior.

37. ***Ejemplo*** *La siguiente matriz no es escalonada porque se incumple la condición 2) pues el número de ceros antes del primer elemento diferente de cero del renglón tres es igual al del renglón dos.*

$$\begin{pmatrix} 2 & 0 & 0 & \vdots & 1 \\ 0 & 1 & 1 & \vdots & 3 \\ 0 & 1 & 1 & \vdots & 2 \end{pmatrix}$$

Al decir que la anterior matriz no es escalonada, lo que estamos mencionado es que se puede combinar el segundo y tercer renglón, con operaciones elementales para que nos quede la matriz escalonada siguiente, la cual nos dice inmediatamente que el sistema no tiene solución, pues en el tercer renglón nos queda la contradicción $0 = 1$:

$$R2 - R1 \rightarrow \begin{pmatrix} 2 & 0 & 0 & \vdots & 1 \\ 0 & 1 & 1 & \vdots & 3 \\ 0 & 0 & 0 & \vdots & 1 \end{pmatrix}$$

En general, una matriz escalonada nos dice si hay o no hay solución, y si la solución es única o no. Podemos sacar más información con un poco de trabajo adicional.

38. ♣ ***Definición.*** *Una matriz, aumentada o no, M, se dice que está en forma* ***escalonada reducida*** *si*

1) Está en forma escalonada.

2) El primer elemento diferente de cero en un renglón es el 1.

3) En la columna correspondiente al primer elemento no nulo de un renglón hay solo un elemento diferente de cero, a saber, él mismo.

39. ***Ejemplo*** *De las siguientes tres matrices, la primera no es escalonada reducida pues no es escalonada, la segunda tampoco es escalonada reducida pues el primer elemento diferente de cero en el segundo renglón no es 1, pero la tercera matriz sí es escalonada reducida:*

$$\begin{pmatrix} 1 & 0 & 8 & \vdots & 1 \\ 0 & 1 & 5 & \vdots & 3 \\ 0 & 1 & 2 & \vdots & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 \\ 0 & 2 & 0 & \vdots & 3 \\ 0 & 0 & 1 & \vdots & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 8 & \vdots & 1 \\ 0 & 1 & 5 & \vdots & 3 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

Observemos que la última matriz nos dice que tenemos un sistema con 2 ecuaciones y 3 incógnitas. Esto implica que nos queda una incógnita libre a la cual se le puede dar cualquier valor para que nos quede un sistema de 2 incógnitas con 2 ecuaciones con solución única (pues nos queda la identidad 2×2). Por lo tanto, la tercera matriz nos dice que el sistema asociado tiene infinitas soluciones con **un grado de libertad**, o sea con una y solamente una incógnita libre, que puede ser la z. Concretamente, podemos leer las soluciones despejando x y y en términos de z:

$$x = -8z + 1$$
$$y = -5z + 3.$$

Ahora ya podemos formalizar el algoritmo de Gauss-Jordan tal como se usa en la vida real:

40. ***Algoritmo de Gauss-Jordan para resolver un sistema lineal de ecuaciones con n incógnitas y m ecuaciones:***

Sea el sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\ . \quad . \quad . \quad . \quad . \\ . \quad . \quad . \quad . \quad . \\ . \quad . \quad . \quad . \quad . \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m \end{cases}$$

Obsérvese que no es necesario que el sistema tenga el mismo número de ecuaciones y de incógnitas. Puede ser que haya más ecuaciones que incógnitas o puede ser que haya menos. Eso no importa, pues el algoritmo resuelve todos los casos. Tenemos 5 pasos, válidos tanto para cálculos a mano como para programar un computador:

1) Escribir la correspondiente matriz aumentada asociada al sistema:

$$\begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} & \vdots & b_1 \\ a_{21} & a_{22} & ... & a_{2n} & \vdots & b_2 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & & . & . & . \\ a_{m1} & a_{m2} & ... & a_{mn} & \vdots & b_m \end{pmatrix}$$

2) Escalonar por renglones la matriz aumentada.

3) Decidir la existencia y unicidad de la solución: basándose en la matriz escalonada, decidir si la solución es única (cuando se llega a la matriz cuadrada $\mathbb{I}$), si no hay solución (si se llega a una contradicción, como $1 = 2$) o si hay infinitas soluciones (si hay por lo menos un renglón de sólo ceros y no hay contradicción).

4) Si no hay solución, terminar el algoritmo. Si hay solución, describirla llevando la matriz a la forma reducida escalonada. Es decir, si la solución es única, hay que hallarla. Pero si hay infinitas soluciones, se especifica cuántos grados de libertad hay, es decir, a cuántas incógnitas se les puede dar valores arbitrarios.

5) Escribir la solución al sistema.

41. ***Ejemplo*** *Resolvamos por Gauss-Jordan el sistema*

$$\begin{cases} 2x + y + 2z = 1 \\ x - 2y - z = 2 \\ 3x - 6y - 3z = 3 \end{cases}$$

Solución paso a paso:

a) Matriz asociada

$$\begin{pmatrix} 2 & 1 & 2 & \vdots & 1 \\ 1 & -2 & -1 & \vdots & 2 \\ 3 & -6 & -3 & \vdots & 3 \end{pmatrix}$$

b) Escalonamos. En nuestro proceso usamos el símbolo $R1 \leftrightarrow R2$ que significa que intercambiamos el renglón dos con el uno, lo cual ni quita ni pone:

$$R1 \leftrightarrow R2 \quad \begin{pmatrix} 1 & -2 & -1 & \vdots & 2 \\ 2 & 1 & 2 & \vdots & 1 \\ 3 & -6 & -3 & \vdots & 3 \end{pmatrix}$$

$$\begin{array}{r} \\ R2 - 2R1 \to \\ R3 - 3R1 \to \end{array} \begin{pmatrix} 1 & -2 & -1 & \vdots & 2 \\ 0 & 5 & 4 & \vdots & -3 \\ 0 & 0 & 0 & \vdots & -4 \end{pmatrix}$$

3) Decidimos la existencia y unicidad de la solución: como hay una contradicción en el último renglón que dice que $0 = -4$, concluimos que el sistema es inconsistente, por tanto la solución es el conjunto vacío.

42. *Ejemplo* *Resolvamos por Gauss-Jordan el sistema*

$$\begin{cases} x - 2y + z = 0 \\ -2x + 3y - z = 1 \\ 3x - 2y + z = 2. \end{cases}$$

Solución paso a paso:
1) Escribimos la matriz aumentada asociada

$$\begin{pmatrix} 1 & -2 & 1 & \vdots & 0 \\ -2 & 3 & -1 & \vdots & 1 \\ 3 & -2 & 1 & \vdots & 2 \end{pmatrix}$$

2) Escalonamos

$$\begin{array}{r} \\ R2 + 2R1 \to \\ R3 - 3R1 \to \end{array} \begin{pmatrix} 1 & -2 & 1 & \vdots & 0 \\ 0 & -1 & 1 & \vdots & 1 \\ 0 & 4 & -2 & \vdots & 2 \end{pmatrix}$$

$$\begin{array}{r} \\ \\ R3 + 4R2 \to \end{array} \begin{pmatrix} 1 & -2 & 1 & \vdots & 0 \\ 0 & -1 & 1 & \vdots & 1 \\ 0 & 0 & 2 & \vdots & 6 \end{pmatrix}$$

$$\begin{array}{r} R3/2 \to \\ \\ \\ \end{array} \begin{pmatrix} 1 & -2 & 1 & \vdots & 0 \\ 0 & -1 & 1 & \vdots & 1 \\ 0 & 0 & 1 & \vdots & 3 \end{pmatrix}$$

3) Decidimos la existencia y unicidad: puesto que el último renglón empieza por uno, la solución es única. O también se puede decir que al reducir se puede llegar a la matriz identidad 3×3, por lo que la solución es única.

4) Como existe la solución y es única, la hallamos, para lo cual reducimos la matriz escalonada:

$$\begin{array}{r} R1 - R3 \rightarrow \\ R2 - R3 \rightarrow \\ \\ \end{array} \begin{pmatrix} 1 & -2 & 0 & \vdots & -3 \\ 0 & -1 & 0 & \vdots & -2 \\ 0 & 0 & 1 & \vdots & 3 \end{pmatrix}$$

$$\begin{array}{r} R1 - 2R2 \rightarrow \\ -R2 \rightarrow \\ \\ \end{array} \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 \\ 0 & 1 & 0 & \vdots & 2 \\ 0 & 0 & 1 & \vdots & 3 \end{pmatrix}$$

5) Escribamos la solución: la solución es única y es $x = 1$, $y = 2$, $z = 3$.

43. *Ejemplo* *Resolvamos por Gauss-Jordan el sistema*

$$\begin{cases} 2x - y + z = 1 \\ -4x + 2y - z = 2 \\ -4x + 2y - 2z = -2 \end{cases}$$

Solución:

1) Matriz aumentada

$$\begin{pmatrix} 2 & -1 & 1 & \vdots & 1 \\ -4 & 2 & -1 & \vdots & 2 \\ -4 & 2 & -2 & \vdots & -2 \end{pmatrix}$$

2) Escalonamos la matriz

$$\begin{array}{r} \\ R2 + 2R1 \rightarrow \\ R3 + 2R1 \rightarrow \end{array} \begin{pmatrix} 2 & -1 & 1 & \vdots & 1 \\ 0 & 0 & 1 & \vdots & 4 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

3) Decidimos la existencia y unicidad: el tercer renglón, de sólo ceros, nos indica que el sistema admite infinitas soluciones.

4) Reducimos

$$\begin{array}{r} R1 - R2 \rightarrow \\ \\ \\ \end{array} \begin{pmatrix} 2 & -1 & 0 & \vdots & -3 \\ 0 & 0 & 1 & \vdots & 4 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

$$R1/2 \rightarrow \left(\begin{array}{ccc|c} 1 & -1/2 & 0 & -3/2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

5) Escribimos la solución. El sistema correspondiente es:
$x - (1/2)y = -3/2$
$z = 4$
o bien,
$x = (1/2)y - 3/2$
$z = 4$
de lo cual queda claro que y es una variable libre a la cual se le puede dar cualquier valor. Además, y es la única variable libre, tenemos un grado de libertad.

Notemos que el número de variables es igual al número de columnas de la matriz no aumentada, en este caso, 3. Por otro lado, la matriz reducida quedó con dos renglones, es decir, dos condiciones independientes para el sistema. Por tanto, debe haber una variable libre, tal y como se encontró.

La solución general puede escribirse de varias maneras. La primera es:

$S = \{(x, y, z)/x = -3/2 + (1/2)y,\ z = 4,$ para $x,\ y,\ z \in \mathbb{R}\}$

La segunda es

$S = \{(-3/2 + (1/2)y, y, 4),$ para $y \in \mathbb{R}\}$

La tercera es

$$S = \begin{pmatrix} -3/2 + (1/2)y \\ y \\ 4 \end{pmatrix} = y \begin{pmatrix} 1/2 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -3/2 \\ 0 \\ 4 \end{pmatrix}$$

donde $y \in \mathbb{R}$. De aquí podemos obtener soluciones particulares dando valores a la variable y, por ejemplo, si $y = 0$, se obtiene la solución $(-3/2, 0, 4)$.

44. *Observación*

Si el último renglón de la matriz aumentada escalonada de un sistema cuadrado es de la forma $(0..,0, a : b)$, *entonces el sistema tendrá solución única si* $a \neq 0$*; el sistema no tendrá solución si* $a = 0$ *y* $b \neq 0$*, y el sistema tendrá infinitas soluciones si* $a = b = 0$.

45. ***Ejercicio*** *Adecue la observación anterior al caso de matrices no cuadradas.*

46. ***Ejercicio*** *Resuelva por Gauss-Jordan de cinco pasos los siguientes sistemas:*

$$a) \begin{cases} x + y + z = 2 \\ x - y + z = 1 \\ -x - y + z = 1 \end{cases}$$

$$b)\ \left\{\begin{aligned} x+y+z&=2\\ x-y+z&=1\\ 2x+2z&=3 \end{aligned}\right.$$

$$c)\ \left\{\begin{aligned} x+y+z&=2\\ x-y+z&=1\\ 2x+2z&=1 \end{aligned}\right.$$

47. $\boxed{\textbf{\textit{Ejercicio}}}$ *Para cada una de las siguientes opciones, invente un sistema 3×3 y verifique su predicción usando el algoritmo de Gauss-Jordan de cinco pasos:*
a) Su solución sea única.
b) Tenga infinitas soluciones con un grado de libertad.
c) Tenga infinitas soluciones con dos grados de libertad.
d) No tenga solución.

1.5. Ejercicios de repaso

1. Escriba la solución general del sistema lineal $\left\{\begin{aligned} x-3y+z&=\ \ 1\\ 2x+y-z&=-1 \end{aligned}\right.$

2. Un cuadrado mágico de tamaño n es una matriz $n\times n$ cuyos elementos consisten en todos los enteros entre 1 y n^2, de tal forma que las sumas de los elementos de cada columna, renglón o diagonales son iguales a $\frac{n(n^2+1)}{2}$. Por ejemplo, un cuadrado mágico de tamaño 3 es la matriz $A=\begin{pmatrix} 8 & 1 & 6\\ 3 & 5 & 7\\ 4 & 9 & 2 \end{pmatrix}$. Muestre que no existen cuadrados mágicos de tamaño 2.

3. Encontrar los valores de a, b y c tales que la curva $y = ax^2+bx+c$, pase por los puntos $(3,0), (2,3)$ y $(-1,1)$.

4. El álgebra lineal tiene dos componentes, el algebraico y el geométrico. Comenzando desde el próximo capítulo, combinaremos los dos aspectos para hacer realidad la siguiente idea: un sistema lineal de ecuaciones no es más que una generalización de una ecuación en números reales de la forma $ax=b$. Sólo que en lugar de a irá una matriz A y en vez de x irá una n-tupleta de incógnitas, que si se trata de sistemas de dos por dos, puede ser $\begin{pmatrix} x\\ y \end{pmatrix}$, y que se denota genéricamente como $\vec{X}$ y se lee vector X. Similarmente, en vez de b irá un vector o columna de números, $\vec{B}$. En una palabra, un sistema lineal de ecuaciones que se pueda escribir como una matriz aumentada también se puede reescribir de la forma $A\vec{X}=\vec{B}$.

   a) Use el sentido común para reescribir el primer problema del presente taller de la forma $A\vec{X}=\vec{B}$. Identifique claramente qué es A, $\vec{X}$, $\vec{B}$.

   b) Encuentre el(los) valor(es) del parámetro $k\neq 0$, para el(los) cual(es), el sistema $A\vec{X}=\vec{B}$, tiene soluciones no triviales, es decir, diferentes del vector columna con ceros en todas las entradas, donde $A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$. Este problema tiene la lectura geométrica siguiente: hallar todos los vectores $\vec{X}$ no nulos tales que al multiplicar o aplicar la matriz A sobre $\vec{X}$, lo que resulta es kX que es un alargamiento de $\vec{X}$.

c) Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{pmatrix}$, exhiba la solución general del sistema $A\vec{X} = \vec{B}$, donde $\vec{B}$ es cualquier vector columna de tres componentes.

1.6. Resumen

Los sistemas de ecuaciones lineales tienen una amplia gama de aplicaciones pero la complejidad de los problemas resultantes demanda métodos poderosos y conceptos claros. Hemos elaborado la propuesta de reducir el estudio de los sistemas de ecuaciones lineales al de los problemas matriciales. De esa forma, solucionar un sistema equivale a reducir, según reglas claras y sencillas, consignadas en el algoritmo de Gauss-Jordan, una matriz aumentada cualquiera a una matriz escalonada reducida. Para sistemas lineales tales que su matriz reducida tenga n columnas o variables y r renglones no nulos o condiciones independientes para las variables, se cumple que:

Si la matriz terminal es la identidad, la solución es única. Si hay una contradicción aparente, no hay solución. Si no hay contradicciones y el número de variables o columnas de la matriz no aumentada es mayor que el número de renglones no nulos, $n > r$, hay infinitas soluciones. En este último caso, el número de variables libres es igual al número de columnas o variables menos el número de condiciones o renglones no nulos y se denomina grados de libertad de la solución que es igual a $n - r$. Hemos podido entender por qué el algoritmo de Gauss-Jordan es conservativo, pues no crea ni destruye soluciones. Eso se debe a que todo algoritmo de Gauss-Jordan es una composición de operaciones elementales, cada una de las cuales es reversible, lo cual garantiza que no se crea ni se pierda información.

CAPÍTULO 2

LINEAS, PLANOS, $\mathbb{R}^N$ Y EV=ESPACIOS VECTORIALES

En este capítulo introduciremos un pilar básico del álgebra lineal, los espacios vectoriales. Para mostrar que no hay nada misterioso en esas estructuras, partiremos de ideas geométricas muy sencillas y gradualmente iremos abstrayendo las nociones matemáticas necesarias.

La geometría estudia las relaciones espaciales. Por naturaleza, la geometría es una ciencia aplicada pues nace y se realiza en nuestra relación con el espacio físico ligado a la tierra. Sin embargo, la geometría es redundante, uno puede hacer un curso de geometría sin hacer un solo dibujo, simplemente refiriéndose a relaciones entre números. Con todo, en este curso la geometría es básica, pues hay una correspondencia perfecta entre algunos problemas geométricos y sistemas de ecuaciones lineales. Por otra parte, puesto que todos nosotros tenemos cientos de millares de horas de experiencia en manejo espacial, todos nosotros tenemos un enorme conocimiento que podemos utilizar para dos cosas: una es usar la geometría para dilucidar problemas de álgebra lineal y la otra es usar el álgebra lineal para resolver problemas de geometría.

Un primer objetivo será poder interpretar una ecuación lineal con dos incógnitas como una recta en el plano. Por tanto, resolver un sistema de dos ecuaciones con dos incógnitas equivaldrá a encontrar todos los puntos comunes a dos rectas: puede ser uno sólo, cuando las rectas no son paralelas, o puede no haber solución cuando las rectas son distintas pero paralelas, o puede haber infinitas soluciones cuando las dos ecuaciones representan una misma recta. Observemos que gracias a nuestra interpretación de ecuaciones por rectas, ya nos liberamos, al menos en este caso, de la preocupación de si nuestro algoritmo es conservativo o no. Y lo que es más, iremos reuniendo evidencia a favor de la tesis de que el **algoritmo de Gauss-Jordan** es conservativo.

2.1. Espacios cartesianos

El tipo de geometría que vamos a considerar fue inventado por Descartes y consiste en representar los números sobre el espacio cartesiano que es un espacio con un sistema de direcciones para llegar a cualquier punto.

El espacio unidimensional recto o recta real lo notamos $\mathbb{R}$ y tiene su origen en el punto 0 y todos los números conceptualizan una distancia y una dirección. Los números positivos se ponen al lado derecho del origen, los negativos al lado izquierdo.

El plano cartesiano es el espacio bidimensional notado como $\mathbb{R}^2$.

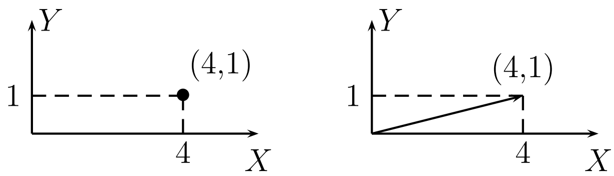


Figura 2.0. El plano cartesiano. Una dupleta representa tanto un punto como un vector, el cual tiene dirección, sentido y magnitud.

El plano cartesiano tiene dos ejes mutuamente perpendiculares, el eje horizontal X que puede considerarse como la línea de los reales $\mathbb{R}$ insertados en el plano, y el eje vertical Y. Esos ejes determinan la representación de cada elemento de $\mathbb{R}^2$ como (x, y), donde x es la componente en X y y es la componente en Y. La componente x se dibuja horizontalmente: positiva hacia la derecha del origen, negativa hacia la izquierda. La componente y se dibuja verticalmente: positiva hacia arriba del origen, negativa hacia abajo. A (x, y) también se le llama dupleta ordenada, pues las dupletas $(2, 1)$ y $(1, 2)$ no son las mismas, es decir, no representan el mismo elemento del plano.

El espacio tridimensional $\mathbb{R}^3$ es el espacio en el que vivimos y se representa por 3 ejes mutuamente perpendiculares y cada punto está dado por una tripleta, la cual tiene 3 coordenadas (x, y, z). Una tripleta también representa un vector.

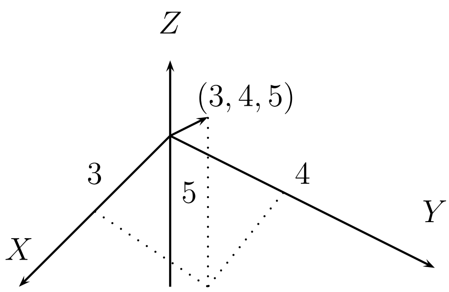


Figura 2.1. El espacio 3D (tres dimensiones): la tripleta $(4, 5, 5)$ o el vector $(4, 5, 5)$.

Es conveniente aprender a representar en la mente al espacio tridimensional, pero orientado de manera arbitraria. Las siguientes recomendaciones pueden ayudar: el espacio tridimensional se dibuja de manera que el piso represente el plano XY y la altura el eje Z. Además, se usa el convenio de la mano izquierda: el pulgar es el eje X positivo. El índice es el eje Z positivo hacia arriba. El dedo del corazón es el eje Y. Se forma un trípode con 3 ángulos rectos, apuntando con el eje X o pulgar hacia el corazón. Ahora, gire la mano de manera arbitraria: tendrá el espacio tridimensional con ejes que apuntan en dirección arbitraria.

Una de las ventajas inmediatas de la idea de Descartes es que soluciona el problema de no poder imaginar, por ejemplo, un espacio de 8 dimensiones. En efecto, los puntos de dicho espacio se representarían por 8-tupletas, las cuales tendrían 8 coordenadas. Una vez hecho esto, es elemental tratar con conjuntos de 8-tupletas y aplicarles las mismas leyes que se aplican a las dupletas en el plano o tripletas en el espacio.

La maquinaria propuesta por Descartes y aceptada universalmente funciona igual de bien en 3 dimensiones como en 93 o en cualquiera otra. Dicho de otra manera, no tenemos hasta ahora una razón matemática para señalar a la dimensión 3, la de nuestro espacio, como privilegiada: no sabemos por qué nuestro espacio es tridimensional. Se han propuesto posibles explicaciones de este misterio, por ejemplo: en dos dimensiones no podrían haber organismos como nosotros, pues el tracto intestinal nos dividiría en dos partes separadas que ya no serían un organismo. Pero si el universo tuviese más de 3 dimensiones espaciales, entonces puede demostrarse que no existirían sistemas planetarios estables: un pequeño meteorito sacaría de órbita a cualquier planeta. Tampoco existirían átomos. La vida como la conocemos no podría existir. La disciplina física que se ha agarrado de lleno con el problema de la dimensionalidad es la **teoría de cuerdas** (Zwiebach, 2004).

En nuestro mundo, las n-tupletas pueden aparecer naturalmente en cualquier momento.

Pero no sólo se desea describir la naturaleza, sino que también se añora entenderla. Eso significa armar relaciones entre las variables de tal manera que mediante el valor de algunas de ellas, llamadas independientes, uno pueda predecir, al menos aproximadamente, el valor de otras, llamadas dependientes. Un modelo es una definición de un conjunto de variables independientes y de sus interrelaciones. Si todo eso se expresa por medio de fórmulas matemáticas, tenemos un modelo matemático. Si adicionalmente las ecuaciones son lineales (de la forma $3x + 4y - 7z = 8$), tenemos un modelo lineal. Pues bien, lo ideal es meter tan pocas variables independientes como sea posible. En la práctica, eso se hace por tanteo. Si con 3 variables uno no puede hacer grandes predicciones, entonces uno agrega otra más a ver si eso sirve de algo.

Nosotros usamos una equivalencia, los elementos de $\mathbb{R}^n$, n-tupletas ordenadas, los representamos indistintamente como un punto con sus coordenadas correspondientes o bien como un vector o flecha que sale del origen y que llega hasta el punto dado. Esta equivalencia permite interpretar un vector como una posición o bien como una abstracción de objetos vectoriales, o sea, como el desplazamiento, la velocidad, la aceleración, la fuerza, el campo magnético, etc. Formalmente tenemos:

48. ♣ *Definición.* *$\mathbb{R}^n$ es el conjunto de las n-tuplas ordenadas o **vectores**. Un elemento $\vec{x} \in \mathbb{R}^n$, lo notamos $(x_1, x_2, ..., x_n)$. $\mathbb{R}^2$ también se llama el plano, y a $\mathbb{R}^3$ el espacio. Al número n también se le llama **dimensión**. Dos n-tuplas ordenadas son iguales si hay igualdad entre las coordenadas. El vector $(4, 2, -5)$ es elemento de $\mathbb{R}^3$, mientras que $(6, -2, -3, 4)$ lo es de $\mathbb{R}^4$. Los vectores $(1, 2, 3)$ y $(2, 1, 3)$ no son iguales. El cero en $\mathbb{R}^3$ es $(0, 0, 0)$ y corresponde con el origen de coordenadas.*

Para conectar nuestros vectores con las aplicaciones en física, debemos notar que nuestros vectores tienen dirección, sentido y magnitud, la cual aprenderemos a medir un poco después. Hay también dos entidades más: segmentos y segmentos dirigidos. Un **segmento** está determinado por un par de puntos sin preferencia de orden, lo notamos $\overline{PQ}$, donde P es un extremo y Q es el otro. Este representa una varilla de hierro que poco importa si se coloca en una dirección o en la otra. Un **segmento dirigido** es un segmento para el que es importante decir cuál es el comienzo y cuál el final. Un segmento dirigido puede representar un desplazamiento, cuya descripción me dice si me muevo de aquí para allá o de allá para acá.

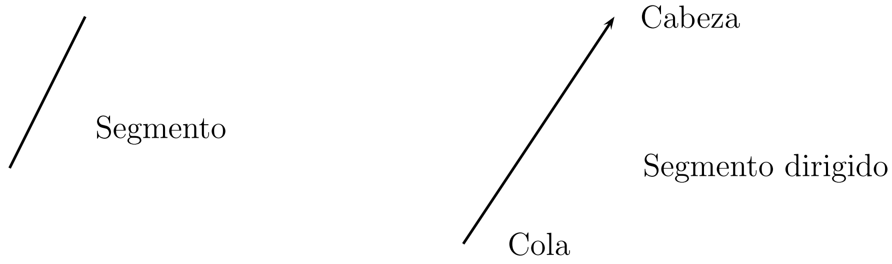


Figura 2.2. Segmento y segmento dirigido.

En un segmento dirigido, al punto de inicio se le llama **cola** y al punto donde termina se le llama **cabeza**. El segmento dirigido se representa por una flecha que va desde la cola hacia y hasta la cabeza. La notación $\overrightarrow{PQ}$ denota un segmento dirigido con cola P y cabeza Q, mientras que la notación $\overleftarrow{PQ}$ denota un segmento dirigido con cola Q y cabeza P. Notemos que $\overleftarrow{PQ} = \overrightarrow{QP}$. Un vector es un segmento dirigido que sale del origen, y como siempre sale del origen, esto nunca se explicita y por lo tanto un vector está determinado por un punto, la cabeza.

2.2. Paralelogramos

Veamos de qué manera hay una relación natural entre operaciones algebraicas con vectores y movimientos geométricos en un paralelogramo. Así vamos fabricando dos mundos paralelos: el algebraico y el geométrico. La idea es aprender a cambiarnos de mundo según nos convenga.

Definamos la suma de vectores y después la multiplicación de un vector por un número. A los números también se les llama **escalares**. Después la resta. Más adelante definiremos el producto punto entre dos vectores, el cual es un producto sorpresivo pues el resultado no es un vector sino un escalar.

49. ♣ *Definición.* *Los vectores se suman coordenada por coordenada. Si estamos en el plano o* $\mathbb{R}^2$*, podríamos tener:* $(1,2)+(3,4)=(1+3,2+4)=(4,6)$*. En general,* $(x_1,x_2,...,x_n)+(y_1,y_2,...,y_n)=(x_1+y_1,x_2+y_2,...,x_n+y_n)$*. Observemos que, nuestra definición sólo se encarga de sumar dos elementos. Por eso decimos que la suma es una*

operación binaria. *Además, está definida de tal forma que la suma de dos elementos del mismo espacio quede dentro del mismo espacio. Por eso decimos que la suma es* ***cerrada***.

Deseamos que la suma esté relacionada con paralelogramos. Eso explica la siguiente definición.

50. ♣ ***Definiciones.*** *Un* ***paralelogramo*** *es una figura plana bordeada de 4 segmentos paralelos dos a dos. (Dos lados son paralelos si las líneas que los contienen nunca se cortan.) Un rombo es un paralelogramo de cuatro lados iguales. Un* ***rectángulo*** *es un rombo con ángulos internos iguales. Un cuadrado es un rectángulo con todos sus lados iguales. Un* ***paralelepípedo*** *es una figura tridimensional de caras planas paralelas dos a dos. Un cubo es una paralelepípedo cuyas caras son cuadrados. Un paralelogramo puede definirse en cualquier dimensión, pero como un paralelogramo está totalmente contenido en un plano, uno se lo imagina que está en* $\mathbb{R}^2$ *y todo le funciona.*

Ahora relacionemos el álgebra con la geometría. Un vector es un desplazamiento, el cual se codifica partiendo del origen y llegando hasta el punto indicado.

Sumar dos vectores es sumar dos desplazamientos, es decir, es desplazar el primer vector y después el segundo vector. Los vectores se suman coordenada por coordenada, pero gráficamente eso equivale a poner la cola del segundo vector sobre la cabeza del primero. La suma es el desplazamiento total, aquel vector que sale del origen y llega hasta la cabeza del segundo vector. Eso de poner la cola contra la cabeza implica tomar el segundo vector, que en principio sale del origen, y trasladarlo paralelamente a sí mismo hasta que llegue a la cabeza del primer vector. La conclusión es que los vectores que se están sumando definen los lados contiguos de un paralelogramo y el resultado de la suma es la diagonal que une el origen con la cabeza del segundo sumando, como en el siguiente gráfico en el cual se suma $(3, 1)$ con $(1, 2)$:

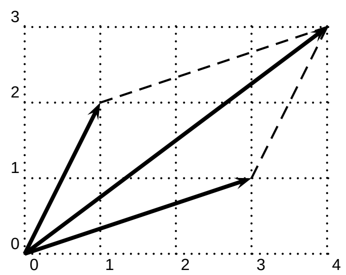


Figura 2.3. Suma de vectores.

51. $\boxed{\textbf{\textit{Ejercicio}}}$ *Justifique gráficamente la apreciación de que la suma es conmutativa, es decir, que para cualquier par de vectores* $\vec{u}$, $\vec{v}$ *se tiene* $\vec{u} + \vec{v} = \vec{v} + \vec{u}$. *Demuestre algebraicamente (a partir de la definición) que la suma es conmutativa.*

52. $\boxed{\textbf{\textit{Ejercicio}}}$ *Justifique gráficamente la apreciación de que la suma es* ***asociativa***, *es decir, que* $(\vec{u}+\vec{v})+\vec{w} = \vec{u}+(\vec{v}+\vec{w})$ *y por lo tanto, la suma binaria puede extenderse a una suma ternaria de manera única, lo cual permite definir* $\vec{u}+\vec{v}+\vec{w}$ *sin ambigüedad. Prueba la asociatividad a partir de la definición de suma. ¿Podrá extenderse la suma a una operación n-aria?*

53. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que en* $\mathbb{R}^n$ *existe un* ***elemento neutro,*** $\vec{o}$*, que al ser sumado a cualquier vector no lo cambia:* $\vec{u} + \vec{o} = \vec{u}$*, para todo* $\vec{u}$*.*

54. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que en* $\mathbb{R}^n$ *para cada vector* $\vec{u}$ *se tiene un vector* $\vec{w}$ *tal que* $\vec{u} + \vec{w} = \vec{o}$*. Demuestre que dicho vector es único y por eso se llama el* ***inverso*** *de* $\vec{u}$ *y se nota* $-\vec{u}$*.*

Cuando un conjunto está provisto de una operación binaria cerrada, asociativa, con elemento neutro y con inverso, decimos que tenemos un **grupo**. Es imposible entender la física moderna en cualquiera de sus facetas sin el concepto de grupo.

55. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que en un grupo el elemento neutro es único.*

En el futuro, en vez de decir que tenemos un conjunto con estructura de grupo diremos que podemos **sumar sin problema**.

Tal vez uno tenga sumas que impliquen un grave problema y así es, en efecto: nosotros hemos definido la suma para vectores, es decir, para segmentos dirigidos que empiezan todos desde el origen. Pero no hemos definido la suma de segmentos dirigidos que empiecen en cualquier parte. ¿Por qué? Porque queremos que nuestras definiciones capten algo importante de lo que pasa en el mundo real. Cuando sumamos dos vectores estamos pensando, por ejemplo, en dos personas que halan una piedra con dos cuerdas y que cada uno hace una fuerza correspondiente. Si la piedra es pequeña, uno puede aproximar la situación diciendo que la fuerza actúa sobre el mismo punto y todo lo representamos como suma de vectores. La suma de las dos fuerzas corresponde a la fuerza que una tercera persona puede hacer para producir el mismo efecto (a muy corto plazo) que las dos personas en conjunto.

Pero si la piedra es grande, lo más seguro es que las cuerdas actúen sobre puntos distantes de la piedra. Uno podría entonces pensar que se puede abstraer la situación diciendo que estamos tratando con vectores generalizados y que éstos también pueden sumarse de una forma muy simple. Pues bien, es verdad que uno puede definir suma de segmentos dirigidos de una manera muy simple, pero no es cierto que el resultado de la suma indique que uno puede reemplazar las dos cuerdas por una tercera y que el resultado sea el mismo en ambos casos. En particular, dos fuerzas que operan en puntos distintos de un cuerpo pueden producir rotaciones, pero una sola fuerza no necesariamente produce rotaciones. Por esta razón, no definimos suma de segmentos dirigidos. ¿Pero qué hacemos entonces con la vida real en la cual hay varias fuerzas que actúan en diversos puntos de un mismo cuerpo? Para estudiar estas situaciones se han desarrollado ciencias complicadas, digamos la estática y la dinámica, la teoría de elasticidad (tanto lineal como no lineal), la teoría de fluidos, la ciencia de resistencia de materiales. Cada una de esas ciencias propone su forma de abstraer, de elaborar los modelos matemáticos y de interpretar los resultados. Por supuesto, en el fondo de todo eso uno encontrará vectores con su suma sin problemas.

Todo esto quiere decir que sumar dos vectores pegando la cabeza del primero con la cola del segundo es un abuso legitimado por la geometría, pero el cual no tiene justificación física en todas las ocasiones.

Pasemos ahora a considerar la multiplicación de un vector por un número.

Cuando uno hace una fuerza, uno puede doblarla, triplicarla, reversarla, etc. Además de suma, los vectores admiten otra operación: la multiplicación escalar. Eso debe implicar que un vector (como entidad matemática) debe poder alargarse, acortarse y reversarse, todo lo cual corresponde simplemente a multiplicarlo por un escalar o número real. Por ejemplo, para alargar un vector dos veces se multiplica cada coordenada del vector por dos.

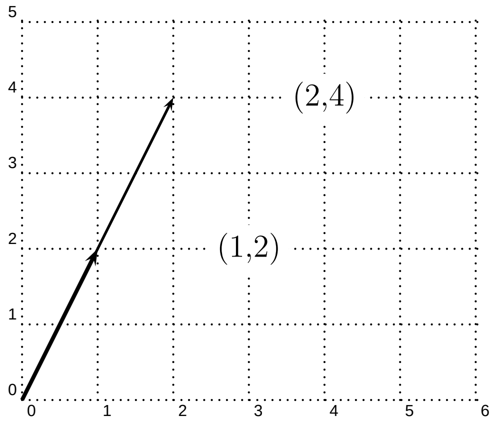


Figura 2.4. El doble de un vector.

Ejemplo: $2(1,2) = (2 \times 1, 2 \times 2) = (2,4)$. Para acortarlo a la mitad se multiplica coordenada por coordenada por $1/2$. Para reversarlo se multiplica por -1, coordenada por coordenada. Observemos que un vector más su revés da cero. Por eso, al revés de un vector se le llama el inverso (aditivo). Ejemplo $-(3,4) = (-3,-4)$ y tenemos que un vector más su inverso da el vector cero:

$$(3,4) + (-3,-4) = (3-3, 4-4) = (0,0)$$

56. ♣ ***Definición.*** *Sea $\vec{u} = (x_1, x_2, ..., x_n)$ un vector de $\mathbb{R}^n$ y λ (lambda) un número o escalar. La multiplicación escalar entre el escalar y el vector la definimos coordenada por coordenada, como $\lambda\vec{u} = (\lambda x_1, \lambda x_2, ..., \lambda x_n)$.*

El uso de letras griegas es usual para representar escalares. Veamos algunas letras junto con otros símbolos usuales en las matemáticas:

α (alfa), β (beta), γ (gamma), ϵ (épsilon), δ (delta), λ (lambda), μ (mu), ν (nu), ρ (ro), σ (sigma), $\exists$ (existe) $\in$ (pertenece) $\forall$ (para todo).

57. $\boxed{\textbf{\textit{Ejercicio}}}$ *Sea $\vec{u} = (2,6), \vec{v} = (7,-3), \vec{w} = (-2,-4)$*

a) Dibujar los vectores $\vec{u}$, $\vec{v}$, $\vec{w}$ en un mismo plano coordenado.

b) Sume $\vec{u} + \vec{v}$ y dibuje la suma. Verifique que la suma de dos vectores es un tercer vector que es la diagonal principal del paralelogramo formado por los dos vectores dados.

c) Repita el ejercicio anterior con $\vec{u} + \vec{w}$ y $\vec{w} + \vec{v}$.

d) Opere y dibuje
$4\vec{u} + 2\vec{v}$
$2\vec{u} + 3\vec{w}$
$3\vec{w} + 5\vec{v}$
e) Calcule y dibuje $-\vec{v}, -\vec{u}, -\vec{w}$.

58. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que en $\mathbb{R}^n$ el **reverso** de un vector da el inverso. Es decir: $(-1)\vec{u} = -\vec{u}$.*

59. ♣ *Definición. Resta de vectores.* *Se define como la suma del inverso. Es lo mismo que con números: $3 - 2 = 3 + (-2)$. Similarmente*
$\vec{u} - \vec{v} = \vec{u} + (-\vec{v})$.
En definitiva da lo mismo que restar coordenada por coordenada. Ejemplo, si
$\vec{u} = (-2, 3)$, $\vec{v} = (1, 2)$,
$\vec{u} - \vec{v} = (-2, 3) - (1, 2) = (-2, 3) + (-1, -2) = (-3, 1)$.

Existe una manera automática de dibujar el vector resta. Se basa en la analogía con los números reales: $5 - 3$ es lo que le falta a 3 para llegar a 5, o sea 2.

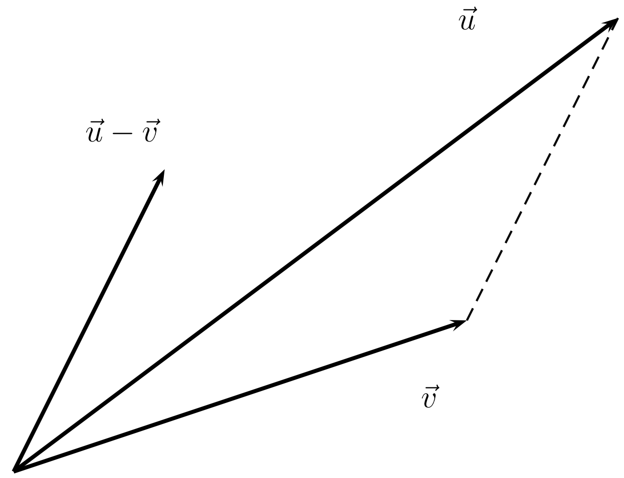


Figura 2.5. Resta gráfica de vectores: $\vec{u} - \vec{v}$ sale del origen y es paralelo al segmento dirigido que va desde $\vec{v}$ hasta $\vec{u}$.

De igual forma, $\vec{u} - \vec{v}$ es el vector que representa lo que le falta a $\vec{v}$ para llegar a $\vec{u}$, es decir, es una flecha cuya cola está en la cabeza de $\vec{v}$ y su cabeza está en la de $\vec{u}$. Pero para nosotros todos los vectores salen del origen. Eso se remedia trazando desde el origen un vector que sea paralelo a la resta gráfica pero con la misma dirección y sentido.

60. $\boxed{\textbf{\textit{Ejercicio}}}$ *Sea $\vec{u} = (2, 6)$, $\vec{v} = (7, -3)$, $\vec{w} = (-2, -4)$. Calcule y grafique las siguientes restas usando la definición de resta como vectores y el equivalente de resta de puntos:*

a) $4\vec{u} - 2\vec{v}$
b) $2\vec{u} - 3\vec{w}$
c) $3\vec{w} - 5\vec{v}$.

2.3. EV=Espacios vectoriales

Nosotros hemos trabajado con $\mathbb{R}^n$ y hemos visto cómo se suman vectores y cómo se multiplica un vector por un escalar. Y lo hemos podido hacer de manera natural y sin problema. Pues bien, hay muchos otros conjuntos fuera de $\mathbb{R}^n$ donde también puede definirse **una suma y una multiplicación escalar sin problema alguno**. A dichas estructuras se les llama **espacios vectoriales**.

61. ♣ *Definición. EV= Espacio vectorial.* *Un espacio vectorial es un conjunto no vacío sobre el cual se ha definido una suma que le da estructura de grupo y una multiplicación por escalares que cumple las siguientes propiedades:*

La multiplicación por un escalar distribuye la suma de vectores. Para todo escalar λ y todo par de vectores $\vec{u}$, $\vec{v}$ se cumple que $\lambda(\vec{u}+\vec{v}) = \lambda\vec{u} + \lambda\vec{v}$.

La multiplicación por un vector distribuye la suma de escalares. Para todo par de escales α, β y cualquier vector $\vec{u}$ se tiene que $(\alpha+\beta)\vec{u} = \alpha\vec{u} + \beta\vec{u}$.

Alargar un vector α veces y el resultado alargarlo β veces, es lo mismo que alargar al vector $\alpha\beta$ veces: $\alpha(\beta\vec{v}) = (\alpha\beta)\vec{v}$.

Alargar un vector una vez es lo mismo que no hacer nada. Para todo $\vec{u}$ se tiene que $1\vec{u} = \vec{u}$.

62. ***Ejercicio*** *Demuestre que $\mathbb{R}^n$ es un espacio vectorial.*

El espacio $\mathbb{R}^n$ puede verse como una elaboración de $\mathbb{R}$ y, en cierto sentido, todos los espacios vectoriales no son más que matrices de $\mathbb{R}^n$, aunque tal vez n pueda ser infinito. Por ahora no profundizaremos en estos temas, pero puede ser muy conveniente dar unos dos o tres ejemplos de espacios vectoriales un poco más abstractos.

63. ***Contraejemplo*** *Nos gusta decir y enfatizar que cuando uno puede **sumar y alargar sin problema**, uno tiene un EV, o sea un conjunto en el cual uno puede operar como si estuviese en $\mathbb{R}^n$. Pero entonces, ¿qué querrá decir sumar o alargar con problemas? Para entender eso, consideremos el siguiente ejemplo, en el cual se definen la suma y la multiplicación escalar de manera especial.*

Para que un conjunto sea EV se requiere, entre otras cosas, que sea no vacío. Como conjunto tomamos el plano, con coordenadas (x, y). Se requiere una suma y una multiplicación escalar. Nuestra suma la definimos como sigue: para sumar, lo único que nos importa es la primera coordenada:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, 0)$$

Similarmente, para multiplicar por un escalar:

$$\lambda(x, y) = (\lambda x, 0)$$

La suma, tal como la hemos definido, es cerrada y lo mismo la multiplicación escalar. Como quien dice, no hay problema ni con la suma, ni con la multiplicación escalar. Pero eso no es suficiente para ser EV. Se requiere que el conjunto con la suma cumpla los axiomas de grupo. Veamos que la suma es conmutativa, es asociativa, no hay elemento neutro, pues si existiese, sería de la forma $(0, w)$, pero $(x, y) + (0, w) = (x, 0) \neq (x, y)$.

Como no hay elemento neutro, no hay necesidad de pensar en inversos, y no tenemos un grupo. Además, para ser EV se requiere que el número uno se multiplique por cualquier vector del mismo número. Veamos si es cierto:

Si $(x, y) \in \mathbb{R}^2$, entonces $1(x, y) = (1x, 0) = (x, 0) \neq (x, y)$. Falso. No tenemos un EV.

64. **_Ejercicio_** *Sobre cualquier $\mathbb{R}^n$ hemos definido una suma, la suma oficial, y una multiplicación escalar, las inducidas por las operaciones naturales de los números reales. ¿Es esa la única manera de definir una suma generalizada y una multiplicación escalar de tal manera que la estructura resultante sea EV?*

65. **_Ejemplo y ejercicio_** *Notamos por $M_{n\times m}$ al conjunto de las matrices con n renglones y m columnas. Definimos la suma de matrices entrada por entrada y la multiplicación escalar de una matriz por un escalar real como la multiplicación de cada entrada por el escalar. Este conjunto es no vacío, permite sumar y multiplicar sin ningún problema y por lo tanto es un EV.*

Demostración: ejercicio.

66. **_Ejercicio_** *Un polinomio de grado 3 es una expresión del tipo $p(x) = a+bx+cx^2+dx^3$. Demuestre que los polinomios de grado 3 no forman un EV, pero que los polinomios de grado menor o igual que 3 sí forman un EV. Dichos polinomios pueden sumarse y multiplicarse por números reales sin ningún problema y siguen siendo polinomios de grado menor o igual que 3.*

67. **_Ejemplo y ejercicio_** *Intuitivamente podemos definir la clase de las funciones continuas como aquellas cuyas gráficas son curvas que no tienen roturas. Defina la suma de dos funciones continuas y averigüe si la suma es cerrada, es decir, si suma de continuas es continua. Lo mismo con la multiplicación escalar. Verifique que la suma y la multiplicación escalar no tienen problema alguno. Concluya declarando si el conjunto de las funciones continuas es o no un EV.*

68. **_Ejemplo y ejercicio_** *Intuitivamente podemos decir que una línea es tangente a una curva cuando es la línea que más se parece a la curva cerca del punto de tangencia. Podemos definir la clase de las funciones derivables como aquellas funciones cuya gráfica es una curva a la cual se le puede, sin ambigüedad alguna, asociar una línea tangente en cada punto. Una función que no es continua no puede ser derivable en los puntos de discontinuidad y lo mismo una función continua que tenga esquinas, pues uno no puede definir cuál es la única línea que más se parece a la función. Usando límites, póngale rigor a nuestra definición intuitiva de función derivable tanto en un punto como en un intervalo abierto. Defina la suma de dos funciones derivables y averigüe si la suma es cerrada, es decir, si la suma de derivables es derivable. Lo mismo con la multiplicación escalar. Concluya declarando si el conjunto de las funciones derivables es o no un EV.*

69. **_Ejercicio_** *Generalice el ejemplo anterior a derivadas de orden superior.*

Estos ejemplos anuncian el enorme cubrimiento del álgebra lineal, la cual contiene infinitas cosas que uno ni se imagina y que se pueden tratar con igual naturalidad que a los elementos de $\mathbb{R}^2$. Siguiendo por este camino de abstracción se puede llegar a la mecánica cuántica que es un pequeño capítulo del álgebra lineal considerada en toda su generalidad (la cual se denomina análisis funcional). No profundizaremos en estos temas, pues sabemos de sobra que es contraproducente que un niño aprenda a caminar sin haber aprendido a gatear.

70. **_Ejercicio_** *Demuestre que en cualquier espacio vectorial el elemento neutro es único. Y que cada inverso es único. Demuestre que el* ***reverso*** *de un vector da el inverso. Es decir,* $(-1)\vec{u} = -\vec{u}$. *Compare con el ejercicio* 58.

71. **_Ejercicio_**

a) Discuta la viabilidad e interés de usar coordenadas cartesianas para representar la posición de un carro y de usar vectores para representar sus desplazamientos.

b) Discuta la viabilidad e interés de usar coordenadas cartesianas para representar puntos en el espacio ocupados por estrellas y de usar vectores para representar sus desplazamientos. (¿Se mueven las estrellas?)

c) Discuta la viabilidad e interés de usar coordenadas cartesianas para representar la dinámica de un conjunto de variables de interés médico, digamos la tensión arterial, la concentración de triglicéridos, y la concentración de azúcar.

d) Compare la forma de orientación de alguien que se guía por coordenadas cartesianas y de alguien que lo hace con ayuda de una brújula.

e) Observe si es verdad que algunas plantas orientan sus hojas para maximizar la iluminación recibida y por lo tanto demarcan el oriente y el occidente. ¿Cómo podremos saber con ayuda de dichas plantas cuál es el oriente y cuál el occidente? ¿Será más beneficioso guiarse por los nidos de termitas que se orientan hacia el polo norte magnético?

f) Explique por qué un perdido en la selva o en el mar termina haciendo círculos.

h) La suma de vectores coordenada por coordenada es equivalente a la suma mediante un paralelogramo. Esto está bien para un plano. ¿Podrá definirse la suma de desplazamientos sobre la superficie de una esfera mediante un paralelogramoide (un paralelogramo curvo) de tal forma que la suma sea conmutativa?

72. **_Ejercicio sobre Tierra plana vs. tierra curva_**

En este curso trabajamos con espacios planos, pero es importante tener en cuenta que no todo es plano. Un buen ejemplo es la Tierra. La siguiente narrativa combina la ficción con la realidad. Los datos reales, fácilmente distinguibles, fueron tomados de la Gran Enciclopedia Larousse (1983).

Reinaba en aquel entonces Carlos V de la casa de Habsburgo sobre Alemania y España, y el rey repartía su tiempo aquí y allá. Eran grandes las riquezas que traían de ultramar y el que nacía, nacía para derrochar, especialmente si tenía nobleza en sus venas. Veía pues el rey a toda la Corte en una sola parranda y preguntó: ¿Hay alguien aquí que esté interesado en algo más que en divertirse? Los cortesanos sintieron pavor pero fueron salvados por uno de ellos, flamenco, que dijo: hay un extranjero que desea hablar con el rey. Fue llevado Fernando de Magallanes ante el rey que, en alemán, dijo: ¡Hablad! Magallanes respondió: puedo demostrar que la Tierra no es plana sino redonda. Hubo consejo y le dieron por misión demostrar su pretensión.

El portugués Magallanes salió de Sanlúcar de Barrameda en 1519 con 5 carabelas. Llegó a América por Brasil, y por allí mismo una tormenta le desbarató una carabela. Pero de todas maneras siguieron y descubrieron el estrecho que hoy se llama de Magallanes. Se aventuró a meterse en las entrañas del océano Pacífico, y después de 3 meses de penurias llegó a las Marianas en 1521, pero poco después fue muerto en una isla de las Filipinas. Agonizante, Magallanes le dijo a Elcano: por Dios y por el rey, cumplid la misión. Elcano respondió: por vos la cumpliré. Viendo Magallanes en los ojos de Elcano su sinceridad, su sabiduría y su firmeza, su corazón se llenó de paz y murió.

Elcano tomó el mando de las carabelas, le dio la vuelta a Asia y también a Africa y llegó a España en el otoño de 1522. Se presentó al rey Carlos V y le entregó su bitácora (las memorias de su recorrido, incluyendo las mediciones de rigor). Le fue anunciado a Elcano la cifra de toda la fortuna que se había invertido en su expedición y lo llevaron a un lugar donde sería ahorcado si su bitácora no demostraba que su misión había sido cumplida.

Le pusieron la soga al cuello y dieron la bitácora a los matemáticos para que la estudiaran.

El rey miraba la escena desde una ventana y detrás de una cortina traída desde Flandes. Elcano estaba ahí con la soga en su cuello esperando el veredicto y el rey se maravilló de su serenidad y de su paciencia. Y pensando en eso, de pronto comprendió el rey que Elcano había cumplido la misión y que él sí sabía de sobra y con certeza que la Tierra era redonda en verdad. Y entonces el rey visualisó a la Tierra como a una manzana en el espacio. Y se preguntó: yo no siento que la Tierra se caiga, ¿qué será lo que la mantiene en su puesto? Y no encontró más respuesta sino a la mano de Dios. Cuando llegó a esa conclusión, el rey se estremeció pues nunca en su vida había percibido a Dios tan cerca de su vida. Inmediatamente se dijo: yo voy a defender a Dios de todos sus enemigos. Los cortesanos querían preguntarle la causa de su estremecimiento, que por demás fue bastante pronunciado, que si tal vez sentía dolor porque ahorcaran a Elcano, pero el rey los dejó callados al ordenarles: hagan otras siete horcas para los matemáticos. Si no me dan la respuesta correcta, que yo ya sé, los ahorcaré.

Los matemáticos estaban tan embolatados en sus quehaceres que ni cuenta se dieron de las otras siete horcas. Cuando los matemáticos terminaron fueron al rey y le dijeron: la Tierra es redonda en verdad. Sabiendo ellos que el rey tenía un refinado gusto por las cosas maravillosas, ellos esperaban que el rey lloraría de alegría. Pero el rey no lloró, ni mostró conmoción alguna. Como los cortesanos vieron que el rey no mandaba a ahorcar a los matemáticos, dedujeron que ellos habían llegado a la respuesta correcta que el rey ya sabía.

Bajaron a Elcano de la horca y cuando los matemáticos supieron para qué eran las otras siete horcas, se desmayaron todos, excepto don Beltrán Alonso Rodríguez, duque de Fuensaldaña. Este tipo no podía desmayarse porque el poseía la tradición de los piratas que data desde los romanos: cuando Pompeyo fue encomendado para acabar con ellos, sólo se salvaron aquellos que se aventuraron a abrirse de las costas y meterse mar adentro en el Mediterráneo donde no podían ver las montañas que servían de guía sino sólo contaban con las estrellas. Pues bien, cerca de la costa europea se veían estrellas hacia el norte que no se veían desde la costa de África. Y algo similar pero complementario sucedía desde dicha costa. Eso era un claro efecto de la redondez de la Tierra. Además, adentrándose en el mar, las montañas se van ocultando, lo cual evidencia la redondez del mar.

Elcano fue citado a los 15 días siguientes para que recibiera los honores. El día señalado tocaron trompeta, guardaron silencio y se proclamó:

La Tierra es redonda. Gloria a Dios, al rey, a Magallanes, a Elcano y a su tripulación.

Y salió el bufón y dijo:

La Tierra es redonda como un cilindro.

Hubo risas y hasta algo de nerviosismo pero ni Elcano se defendió, ni tampoco se defendieron los matemáticos. Y con esas risas empezó el festejo.

a) Ofrezca un estimativo del radio de la Tierra, asumiendo que las carabelas navegaron durante 3 años a una velocidad promedio por día de 80 km (personas ordinarias pueden recorrer en cicla, en terreno plano, unos 100 km por día). Tenga en cuenta que el viaje no fue por el ecuador, ni por un círculo máximo sino con muchas vueltas, lo cual alargó el recorrido mínimo, digamos, al doble. Compare su estimativo sobre el radio de la Tierra con el moderno, el cual es de alrededor de 6.000 km.

b) ¿Qué clase de mediciones deberían estar en la bitácora para que los matemáticos pudiesen decidir si la Tierra era redonda o no?

c) ¿Cómo resolvieron los matemáticos y Elcano la objeción del bufón?

d) ¿Para qué se necesitaban 5 carabelas si todas iban juntas? ¿No era acaso suficiente una sola? El Beagle de Darwin era uno solo pero de todas formas le dio vuelta y media al mundo y más o menos en el mismo tiempo que la expedición de Magallanes, después de dar la vuelta al mundo, Darwin regresó a América desde Europa con un único propósito, investigar la evolución natural de las poblaciones de las Galápagos. Sin duda lo que encontró fue el bastión que le permitió sustentar su teoría sobre el surgimiento de las especies por evolución.

e) Si bien Cortés fue guiado en 1513 por un aborigen al océano Pacífico, precisamente por Panamá, nadie le había dado la vuelta al Cono Sur de América y no se sabía que los dos océanos se comunicaban. ¿Cómo estaba tan seguro Magallanes que por el sur se unirían los dos océanos, el Atlántico y el Pacífico? ¿Por qué no se dirigió al norte?

f) ¿Cómo dilucidaron los matemáticos que Magallanes no se había devuelto desde el Cono Sur de América por el Índico hacia el extremo sur de África?

g) ¿Por qué mataron a Magallanes y a Elcano no? ¿Por qué no los mataron a todos?

h) ¿Se debe a Magallanes que en Filipinas una cierta minoría hable español actualmente?

i) El rey Carlos V fue un rey buscapleitos toda su vida. El libreto presentado sugiere que la causa de eso fue una suerte de inspiración divina en la cual el sintió demasiado cerca de sí la presencia de Dios. Por causa de tanta guerra, España se endeudó por 3 eternidades y de la infinita riqueza sacada de las Indias sólo quedó el recuerdo y un país enmiserado, pues por tener mucha riqueza gratis no desarrolló su poder manufacturero y aún más, destruyó el que tenía, que no era poco. Procure elaborar otra explicación a tanta sangre derramada. Tenga en cuenta que Carlos V ganó su primera guerra a los 16 años, la cual fue contra una revolución civil. Los comuneros castellanos se rebelaron contra su rey por ser extranjero, pero Carlos V los demolió en 1521 y quedó con un poder absoluto.

Para saber más. Nosotros estudiamos en este curso una amalgama de álgebra y geometría plana. En el mundo de la complejidad, nuestra materia es fácil pues los temas complicados están asociados a geometría curva y a herramientas analíticas apropiadas. El enfoque adoptado para combatir la complejidad que trae consigo la introducción de la curvatura es el mismo de los cartógrafos: a pequeña escala la tierra se modela plana, *lineal*, y se construyen mapas que se pueden poner sobre una mesa. La curvatura sale de la forma como se curvan los diversos mapas para ser pegados unos con otros. Este tema se estudia en geometría diferencial. Aunque las aplicaciones de esta materia son muy variadas, las más conocidas son en física pues, por ejemplo, la gravitación se reduce según la relatividad general propuesta por Einstein a un efecto de la curvatura del espacio-tiempo, una estructura en la cual el tiempo no se puede disociar del espacio. Esta idea se ha venido desarrollando y ahora cubre toda la física. Un enfoque sobre el tema que es moderno, que ensancha la mente y unifica las ideas, y que además es pedagógico puede encontrarse en Rodríguez (2008), en Internet, y comenzar a trabajarse después de ver cálculo vectorial aunque una hojeada por un estudiante del presente curso le podría servir para ver las ideas generales. Si logra sacar como conclusión que el álgebra lineal está en el fundamento de todo, y que por tanto hay que saberla bien, ya será mucho lo que ha ganado.

2.4. Espacios digitales

Nosotros estamos muy acostumbrados a ver a las tripletas de números reales como puntos del espacio, el cual ha sido direccionado por unos ejes cartesianos. Es natural por tanto, preguntarse qué puede ser una cuadrupleta o una octupleta. Al quedarse corto en la respuesta, se puede llegar a pensar que todo esto no es más que un juego bobo. Por eso, presentamos a continuación una interpretación de las n-tupletas, bajo la cual las cuadrupletas, las octupletas y las trillonitupletas todas tienen un sentido muy claro, muy simple, y de interés tanto tecnológico como matemático. Se trata del **espacio digital**. Este espacio es muy similar a $\mathbb{R}^n$ pero al mismo tiempo es muy diferente. Veamos como se construye.

Tomemos el intervalo $I = [0, n)$ y lo dividimos en n subintervalos iguales. Uno cualquiera de ellos I_j es de la forma $[j, j+1)$. Definimos sobre I la función P_j, que se lee **pulso** en j, como la función que asigna 1 a cualquier elemento sobre I_j y 0 a todo lo demás, tal como se ve en la gráfica siguiente:

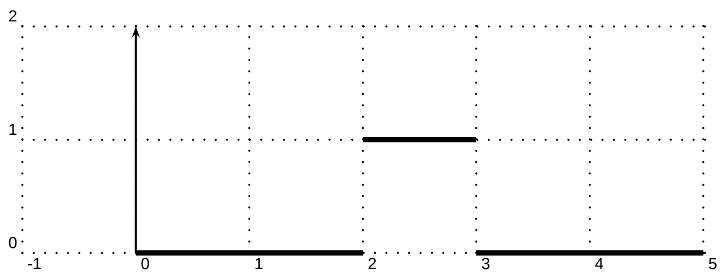


Figura 2.6. La función P_2 es un pulso que vale uno sobre el intervalo [2,3) y cero por fuera.

Si el intervalo base es $[0,5)$, podemos notar a P_2 como $(0,0,1,0,0)$. Pero si el intervalo base es $[0,7)$, P_2 deberá ser notado como $(0,0,1,0,0,0,0)$.

73. $\boxed{\textit{\textbf{Ejercicio}}}$ *Dibuje los siguientes pulsos:*

a) $(0,0,1)$
b) $(0,0,1,0)$
c) $(0,0,1,0,0)$

74. $\boxed{\textit{\textbf{Ejercicio}}}$ *Las siguientes n-tupletas no son pulsos, pero son modificaciones sencillas de pulsos, especifique de cuáles:*

a) $(0,0,3)$
b) $(0,0,5,0)$
c) $(0,0,7,0,0)$

75. $\boxed{\textit{\textbf{Ejercicio}}}$ *Los pulsos son importantes en transmisión de información, un pulso se interpreta como un uno y una ausencia de pulso como un cero, siempre y cuando uno tenga claramente definida la unidad de tiempo básica dada por un reloj. Cuando los pulsos se transmiten en serie se crea un tren de pulsos. ¿A qué tren de pulsos corresponde el número binario 1011010, si cada mensaje tiene 7 pulsos y si los bits (que son biopcionales, todo o nada) se mandan de izquierda a derecha? Bosqueje el tren de pulsos correspondiente.*

76. $\boxed{\textit{\textbf{Ejercicio}}}$ *Formalice la definición de tren de pulsos. Defina la suma de trenes de pulsos y su multiplicación escalar de forma natural, es decir, coordenada por coordenada. Muestre que el conjunto de trenes de pulsos con la suma y la multiplicación escalar así definidos, no forman un espacio vectorial.*

77. $\boxed{\textit{\textbf{Ejercicio}}}$ *Si uno cuenta con un dispositivo de amplitud modulada, se pueden transmitir señales cuya notación es del estilo $(3,4,2,6)$. Esta tetratupleta dice que hemos tomado el intervalo $[0,4)$ y sobre él hemos definido una función o señal que puede notarse como*

$$(3,4,2,6) = 3P_0 + 4P_1 + 2P_2 + 6P_3.$$

*Todo parece indicar que tenemos un espacio vectorial. Podemos sumar y alargar sin problema. Demuéstrelo formalmente, es decir, pruebe que el conjunto de funciones definidas sobre $I = [0,n)$ y constantes sobre cada subintervalo $[i, i+1)$ es un espacio vectorial con la suma y la multiplicación escalar naturales. A dicho espacio lo notamos D_n y lo llamamos el **espacio digital** de dimensión n. Por abuso del lenguaje, los elementos de D_n también se llaman pulsos.*

78. $\boxed{\textit{\textbf{Ejercicio}}}$ *Desde el punto de vista de espacios vectoriales, $\mathbb{R}^n$ y D_n son practicamente lo mismo. Decimos que $\mathbb{R}^n$ y D_n son **isomorfos**. Formalice la definición de isomorfismo de espacios vectoriales y demuestre formalmente que D_n y $\mathbb{R}^n$ son isomorfos. (D_n, el **espacio digital**, es el conjunto de funciones definidas sobre $I = [0,n)$ y constantes sobre cada subintervalo $[i, i+1)$).*

El espacio digital es el referente teórico de la música digital. Veamos algunos detalles del asunto.

El sonido es una onda de presión longitudinal, en la cual las zonas de compresión y de expansión se acomodan a lo largo de la dirección en la cual se transmite. Si uno pone un micrófono para captar sonido, éste lo transforma en una onda eléctrica que viaja a través del cable. Si se visualiza dicha onda por medio de un osciloscopio, la onda se ve como se espera, como una curva que baja y sube innumerables veces. Digitalizar el sonido consiste en generar un tren de pulsos del espacio D_n, que pueden ser tanto positivos como negativos y que tienen diferente amplitud de tal forma que su envolvente sea la onda dada. Si los pulsos son muy gordos, habrá poca fidelidad, pero si son muy flacos consumen mucha información. Hay que lograr un ajuste adecuado. Cuando se transmite sonido en forma digital, sólo se usan pulsos de amplitud uno. Por lo tanto, la amplitud del pulso que codifica la onda de sonido también debe codificarse y se obtendría un sistema de amplitud modulada. O podría usarse un sistema de frecuencia modulada, en el cual los datos son los cambios de frecuencia y no de amplitud.

2.5. La norma de un vector

Definiremos la longitud de un vector de $\mathbb{R}^2$ para que coincida con la noción de longitud del mundo en que vivimos, el cual está caracterizado por el teorema de Pitágoras, descubierto hace 2.500 años. Para ello necesitamos algunos prerrequisitos. La generalización de la definición de norma en $\mathbb{R}^n$ es inmediata. Pero para espacios vectoriales en general, la norma se define de tal manera que cumpla las propiedades naturales que cumple la norma en $\mathbb{R}^n$.

79. ♣ *Definiciones.* *La unidad de longitud es lo que mide una determinada barra cuya longitud sea muy estable a la humedad, el calor, la corrosión. La vara la definió Enrique VIII como la distancia desde la punta de su nariz hasta la punta de su mano bien estirada en posición horizontal. Los franceses definieron el metro como 1/40.000 el círculo máximo de la Tierra que pasaba por un determinado pueblo. La longitud de un segmento en metros es la cantidad de barras de un metro de largo que es necesaria para cubrir el segmento. La unidad de área es un cuadrado de un metro de lado. El área de una figura en metros cuadrados es el número de cuadrados de un metro de lado que son necesarios y suficientes para enladrillar la figura. La unidad de volumen en metros cúbicos es un cubo de un metro de lado. El volumen de una figura en metros cúbicos es el número de cubos unitarios que se necesitan para cubrirla.*

80. ◇ *Teorema.* *El área de un rectángulo de base b y altura h es bh.*

Demostración. Consideremos que b, h sean enteros. Al rectángulo de base b y altura h lo cuadriculamos y se forman bh cuadros de 1 unidad de lado, por lo tanto se requieren bh cuadros para recubrirlo. Es decir, su área es bh. ■

81. $\boxed{\textbf{\textit{Ejercicio}}}$ *Discuta si la anterior demostración es en realidad una demostración o si apenas es su comienzo. En el último caso, haga lo posible por terminarla.*

82. ◇ ***Teorema.*** *El área de una paralelogramo es base por altura.*

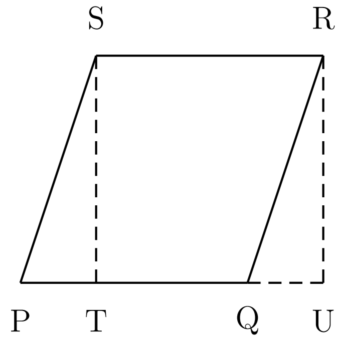


Figura 2.7. Área de un paralelogramo.

Demostración. Consideremos el paralelogramo PQRS. Su altura es lo que mida el segmento TS, y su base es lo que mida PQ. Se fabrica el rectángulo TURS trasladando el triángulo PTS hasta que coincida con el triángulo QUR. El rectángulo nuevo tiene la misma área que el paralelogramo original. Como el área de un rectángulo es base por altura, la del paralelogramo también. ■

83. ♣ ***Definición.*** *Un ángulo se llama recto cuando queda en un vértice de dos líneas que se cruzan formando cuatro ángulos iguales. Medido en grados un ángulo recto mide 90, y en radianes mide $\pi/2$.*

84. ◇ ***Teorema del medio Sol y ejercicio.*** *Los ángulos internos de un triángulo suman 180 grados o π. ¿Por qué se llama este teorema así?*

Demostración. Para probar esto necesitamos el siguiente hecho:

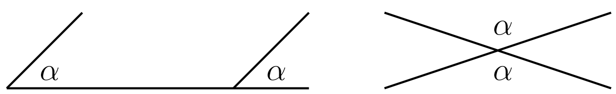


Figura 2.8. Ángulos iguales.

Una escuadra al ser deslizada sobre una línea ni se abre ni se cierra. O mejor dicho: si dos ángulos tienen lados paralelos, entonces son iguales. Por otra parte, cuando dos líneas se cruzan, los ángulos opuestos por el vértice son iguales.

Con los dos resultados anteriores podemos hacer la demostración solicitada.

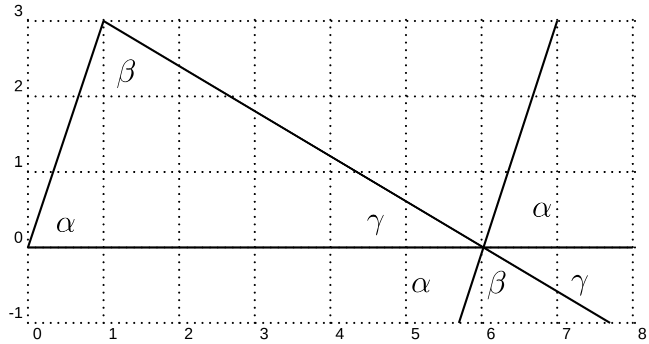


Figura 2.9. Teorema del medio Sol.

En la figura podemos ver tres demostraciones del Teorema. Hay tres maneras de trasladar los ángulos internos para que queden juntos sobre el mismo lado de una línea recta de tal forma que llenen todo ese medio mundo. Por tanto, la suma de los ángulos internos de un triángulo es de 180 grados. ■

85. ◇ ***Teorema de Pitágoras.*** *En un triángulo rectángulo (con uno de sus ángulos recto) de lados* a, b *e hipotenusa* c *se cumple que* $c^2 = a^2 + b^2$.

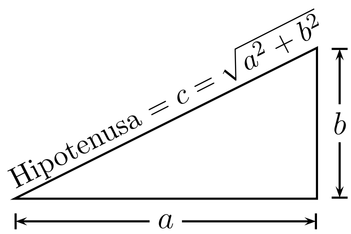


Figura 2.10. Triángulo de Pitágoras.

Demostración. Supongamos que el triángulo rectángulo tiene como vértices A, B, C y como lados a, b, c. Hacemos cuatro copias del triángulo dado y con ellas formamos un cuadrado de lado c de tal forma que quede un cuadrado interior de lado $b - a$. En la figura siguiente, el triángulo inicial está en la esquina inferior derecha, con su hipotenusa apuntando hacia la derecha.

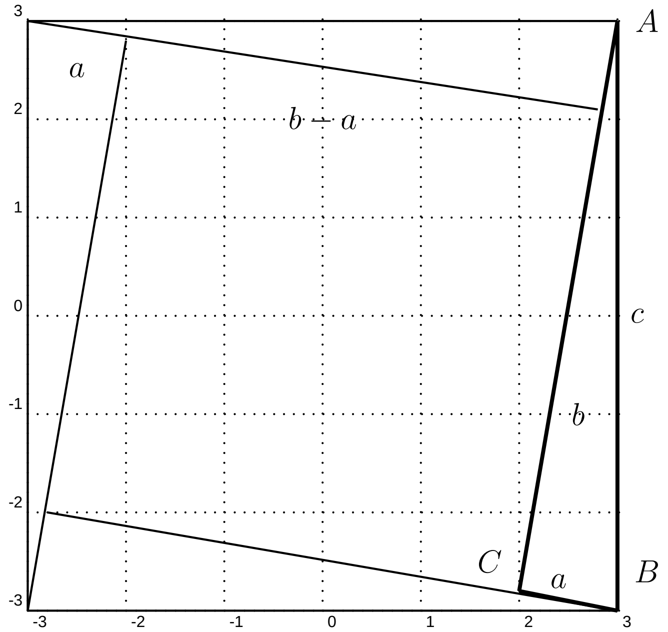


Figura 2.11. Teorema de Pitágoras.

Pero tenemos un problema, ¿son estas figuras realmente cuadrados? Lo que pasa es que un cuadrado no es simplemente una figura que tiene todos los lados iguales. Para que una figura de cuatro lados iguales sea cuadrado se requiere además que los ángulos internos sean todos rectos. Analicemos entonces el vértice C en donde hay un ángulo recto en el interior del triángulo. Pero como el ángulo que queda al lado de una recta es de 180 grados, se deduce que el ángulo interno del cuadrilátero pequeño en el vértice C también es recto. Nuestro razonamiento es válido para todas las esquinas del cuadrilátero pequeño y por tanto es un cuadrado.

Demostremos lo mismo para el cuadrilátero grande, el exterior. Como los ángulos internos de un triángulo suman 180 grados, entonces, en un triángulo rectángulo los dos ángulos no rectos suman $180-90=90$ grados. Por tanto, el ángulo en cada esquina del cuadrilátero grande es recto pues contiene los dos ángulos no rectos del triángulo rectángulo.

Habiendo demostrado que realmente tenemos un cuadrado de lado $b-a$ dentro de otro de lado c, el teorema de Pitágoras es inmediato:

El área del cuadrado grande es c^2, la cual es igual al área del cuadrado pequeño interior, $(b-a)^2$, más el área de 4 triángulos, cada uno de área $(ab)/2$, entonces

$c^2 = (b-a)^2 + 4(ab/2) = b^2 - 2ab + a^2 + 2ab = b^2 + a^2 = a^2 + b^2$ ■

86. ◇ ***Teorema autodemostrado 1.*** *La norma de un vector (a, b) es la distancia desde su cola hasta su cabeza. Se nota como $\|(a, b)\|$. Por lo tanto, armando un triángulo apropiado, se deduce que $\|(a, b)\| = \sqrt{a^2 + b^2}$.*

87. ♣ ***Definición.*** *Un vector unitario es aquel cuyo norma es uno.*

88. $\boxed{\textbf{\textit{Ejercicio}}}$ *Describa todos los vectores de $\mathbb{R}^2$ que sean unitarios.*

89. ***Repaso.*** *Un plp (paralelepípedo) es una figura tridimensional bordeada por 6 caras planas paralelas dos a dos.*

90. $\boxed{\textbf{\textit{Ejercicio}}}$ *Usando Pitágoras defina la norma de un vector en el espacio 3D. Ayuda: forme un plp cuya diagonal principal es el vector dado. Forme un triángulo rectángulo vertical con la diagonal principal, una arista del plp y otra diagonal secundaria. Observe ahora que la diagonal secundaria es la hipotenusa de un triángulo rectángulo en la base del plp. Aplique Pitágoras sobre esos dos triángulos. Generalice a n-dimensiones.*

91. $\boxed{\textbf{\textit{Ejercicio}}}$ *Generalice la definición de norma de un vector para cualquier vector en* $\mathbb{R}^n$.

92. $\boxed{\textbf{\textit{Ejercicio}}}$ *Calcule la norma de los vectores*

$$(2,1),(2,3),(-1,2),(-3,-5),(1,2,3),(-1,2,-3),(-3,-4,5),(2,3,6,5).$$

93. ♣ ***Definición.*** *Sea el punto* P *definido por el vector* $\vec{u}$, *y el punto* Q, *definido por* $\vec{v}$. *La distancia entre dos puntos* P *y* Q *es la norma del vector que va desde* P *a* Q, *calculado por* $\vec{u}-\vec{v}$.

94. ***Ejemplo*** *Sea* $P=(1,3,5), Q=(1,-2,3)$. *Calculemos la distancia de* P *a* Q. *Para ello, hagamos la resta* $(1,3,5)-(1,-2,3)=(1,3,5)+(-1,2,-3)=(0,5,2)$ *y hallemos su norma* $\|(0,5,2)\|=\sqrt{25+4}=\sqrt{29}$.

95. ***Ejemplo*** *La distancia de un punto cualquiera* (x,y,z) *al punto* $(2,3,-1)$ *es la norma del vector* $(x-2,y-3,z+1)$, *es decir,* $\sqrt{(x-2)^2+(y-3)^2+(z+1)^2}$.

96. ***Ejemplo*** *Definamos la circunferencia y hallemos su ecuación.*

Una circunferencia es un conjunto de puntos de $\mathbb{R}^2$ que equidistan de un punto fijo llamado centro. Sea el centro (h,k) y un punto cualquiera de la circunferencia (x,y). Entonces la distancia de (x,y) al centro es constante, igual al radio r. Por tanto, la ecuación es $\sqrt{(x-h)^2+(y-k)^2}=r$ y elevando al cuadrado ambos lados se tiene

$$(x-h)^2+(y-k)^2=r^2,$$

que es la ecuación de la circunferencia con centro en (h,k) y de radio $r>0$.

97. $\boxed{\textbf{\textit{Ejercicio}}}$ *Defina una esfera y halle su ecuación.*

98. $\boxed{\textbf{\textit{Ejercicio}}}$ *Defina la elipse por la regla del jardinero. Tome tres puntillas. Fije dos de ellas en puntos separados, llamados focos, y únalas con un cuerda floja. Con la tercera puntilla tense la cuerda y vaya dibujando la elipse,*

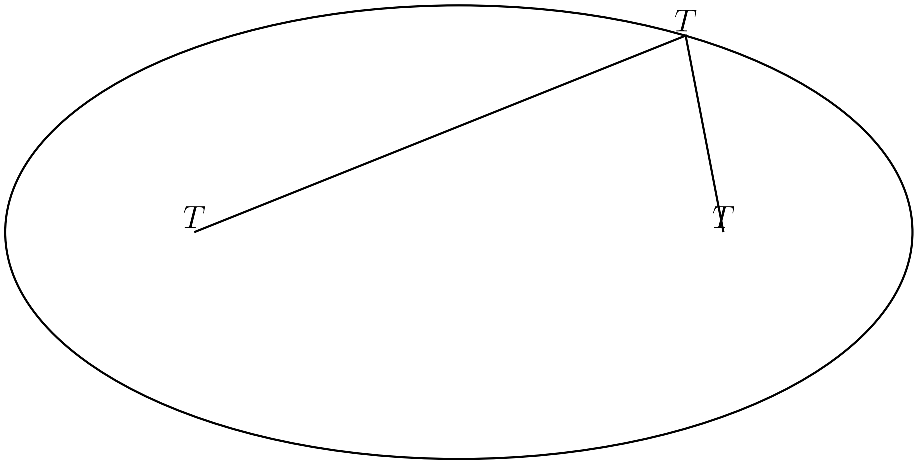


Figura 2.12. *Elipse.*

de tal manera que si los dos focos estuviesen en el mismo punto, quedaría una circunferencia. Demuestre que un punto (x, y) *de la elipse satisface la ecuación*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

donde la elipse se ha centrado en el origen del plano cartesiano, la cuerda del jardinero mide de largo $2a$*, la distancia entre los focos mide* $2c$*, y hemos definido* $b^2 = a^2 - c^2$*.*

99. $\boxed{\textbf{\textit{Ejercicio y definición}}}$ *Demuestre que la definición que tenemos de norma sobre* $\mathbb{R}^n$ *cumple con las siguientes propiedades, las cuales se toman como la definición de norma en un espacio vectorial arbitrario.*

La ***norma*** *es una función* p *definida sobre un espacio vectorial que asocia un número real a cada vector* $\vec{v}$ *tal que*

a) $p(\vec{v}) \geq 0$ *(la norma o largo de un vector no puede ser negativa).*

b) $p(\lambda\vec{v}) = |\lambda| p(\vec{v})$ *para cualquier escalar* λ *y cualquier vector* $\vec{v}$ *(si se alarga un vector, la norma se alarga consecuentemente).*

c) $p(\vec{u} + \vec{v}) \leq p(\vec{u}) + p(\vec{v})$*, para cualquier par de vectores* $\vec{u}$*,* $\vec{v}$*. (Cuando se utiliza la norma para medir distancias, debe cumplirse que el camino directo tiene la distancia más corta entre dos puntos). A esta desigualdad se le llama la desigualdad triangular.*

d) $p(\vec{v}) = 0$ *si y sólo si* $\vec{v}$ *es el vector cero (un vector que no es cero tampoco tiene norma cero).*

En este y en la gran mayoría de textos, la norma se nota como $||\vec{v}||$ *pero en otros libros se usa* $|\vec{v}|$*.*

2.6. El producto interior

Existe una herramienta especialmente diseñada para saber si dos vectores son o no perpendiculares, es decir, si forman un ángulo recto. Se llama **producto interno,**

interior, escalar o punto. Comenzamos con una definición válida para $\mathbb{R}^2$, después la generalizamos a $\mathbb{R}^n$ y luego a un espacio vectorial cualquiera.

100. ♣ ***Definición.*** *Dados dos vectores de $\mathbb{R}^2$, $\vec{u} = (u_1, u_2), \vec{v} = (v_1, v_2)$, definimos el producto interno, escalar o punto $\vec{u} \cdot \vec{v}$ como el número*

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2.$$

101. ***Ejemplo*** *Si $\vec{u} = (3, 5)$ y $\vec{v} = (4, -2)$ entonces*
$\vec{u} \cdot \vec{v} = (3)(4) + (5)(-2) = 12 - 10 = 2.$

102. ***Ejemplo*** *Si $\vec{u} = (1, -1)$ y $\vec{v} = (1, 1)$ entonces*
$\vec{u} \cdot \vec{v} = (1)(1) + (-1)(1) = 1 - 1 = 0.$

103. ***Ejemplo*** *Si $\vec{u} = (1, m)$ y $\vec{v} = (1, -1/m)$ entonces*
$\vec{u} \cdot \vec{v} = (1)(1) + (m)(-1/m) = 1 - 1 = 0.$

104. ***Ejercicio*** *Generalice la definición de producto punto a $\mathbb{R}^n$.*

105. ***Ejercicio*** *Sea el **espacio digital** D_n, el conjunto de funciones definidas sobre $I = [0, n)$ y constantes sobre cada subintervalo $[i, i+1)$. Construya sobre D_n un producto interno que lo haga en todo semejante a $\mathbb{R}^n$. Decimos que $\mathbb{R}^n$ y D_n son **isomorfos** como espacios vectoriales con producto interno. Defina la norma en D_n, de un pulso, de un tren de pulsos, de un elemento cualquiera.*

106. ***Ejercicio*** *Calcule el producto punto entre los pares de vectores. Dibújelos y observe cuáles pares son perpendiculares y cuánto da su producto punto:*

a) $(1, 1), (2, -2)$
b) $(1, 3), (3, -1)$
c) $(1, 7), (2, 4)$
d) $(1, 1, 1), (1, -1, 1)$
e) $(1, 1, 1), (1, -1, 0)$

El producto interno entre dos vectores puede relacionarse con la norma de los vectores y el ángulo entre ellos. Veamos eso con los prerrequisitos básicos:

107. ***Seno y coseno.*** *Recordemos que seno y coseno son relaciones definidas para un círculo de radio 1: dado un ángulo con uno de sus lados sobre el eje $\vec{X}$, el seno es la longitud del segmento vertical abierto por el ángulo, mientras que el coseno es la longitud del correspondiente segmento horizontal.*

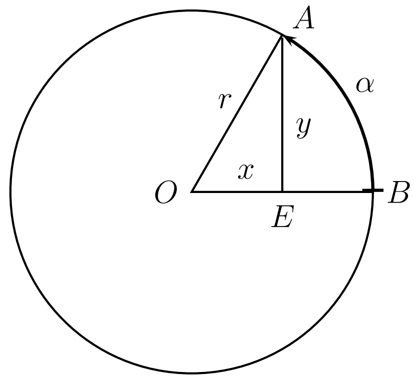


Figura 2.13. $\operatorname{sen}\alpha = y/r$, $\cos\alpha = x/r$.

El teorema de Pitágoras dice que en la gráfica anterior, $x^2+y^2 = 1$*, lo cual significa que* $\operatorname{sen}^2\theta + \cos^2\theta = 1$*. Usando triángulos semejantes, se puede demostrar que en un triángulo recto, el seno es el cociente entre el cateto opuesto y la hipotenusa, mientras que el coseno es el cociente entre el cateto adyacente y la hipotenusa.*

La propiedad más importante del producto punto se basa en el teorema de los cosenos, el cual se basa en la identidad.

108. ***Coseno de una suma.*** $\cos(\gamma_1 + \gamma_2) = \cos\gamma_1 \cos\gamma_2 - \operatorname{sen}\gamma_1 \operatorname{sen}\gamma_2$.

Esta identidad puede demostrarse más luego con ayuda de los números complejos y por ahora la utilizaremos para demostrar el teorema de los cosenos el cual es una generalización del teorema de Pitágoras.

109. ◇ ***Teorema de los cosenos.*** *En un triángulo cualquiera de lados* a, b, c *y de ángulos opuestos* A, B, C *siempre se tiene que* $c^2 = a^2 + b^2 - 2ab\,cos(C)$.

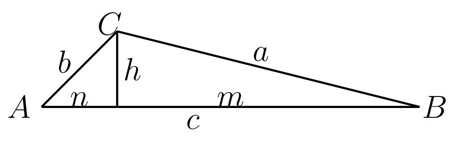


Figura 2.14. Teorema de los cosenos.

Demostración. Pongamos a descansar el triángulo sobre el lado c y del vértice opuesto C bajamos la altura h. La altura separa a c en dos partes, sean m, n tal que $c = m + n$ y por tanto $c^2 = m^2 + n^2 + 2mn$ o sea que $m^2 + n^2 = c^2 - 2mn$. Usando Pitágoras en los dos triángulos rectángulos resultantes, tenemos:

$a^2 = m^2 + h^2$

$b^2 = n^2 + h^2$

Sumando

$a^2 + b^2 = m^2 + h^2 + n^2 + h^2 = m^2 + n^2 + 2h^2 = c^2 - 2mn + 2h^2 = c^2 + 2(h^2 - mn)$

Es decir, $c^2 = a^2 + b^2 - 2(h^2 - mn)$.

Por otro lado, el ángulo C es dividido por h en dos partes γ_1 cuyo lado opuesto es n y γ_2 cuyo lado opuesto es m. Aplicando la identidad del coseno de una suma de ángulos tenemos:

$\cos C = \cos(\gamma_1 + \gamma_2) = \cos\gamma_1 \cos\gamma_2 - \operatorname{sen}\gamma_1 \operatorname{sen}\gamma_2$.

y leyendo dichos valores en el triángulo resulta:

$\cos C = \cos(\gamma_1 + \gamma_2) = (h/b)(h/a) - (n/b)(m/a) = (1/ab)(h^2 - mn)$

Por lo tanto $h^2 - mn = ab\cos C$. Si reemplazamos esta igualdad en

$c^2 = a^2 + b^2 - 2(h^2 - mn)$,

obtenemos

$c^2 = a^2 + b^2 - 2ab\cos C$ tal como reza el teorema. ■

110. ◇ ***Teorema.*** *Si en un triángulo de lados a, b, c se cumple $a^2 + b^2 = c^2$, entonces dicho triángulo es rectángulo. Este teorema define lo que es un ángulo recto para los ingenieros.*

Demostracion: ejercicio.

Ahora podemos establecer lo que podría ser el teorema central del producto punto, el que lo relaciona con normas y ángulos. Aparentemente, la prueba es válida para el plano, pero en realidad es válida para $\mathbb{R}^3$, por qué?

111. ◇ ***Teorema.*** *$\vec{u} \cdot \vec{v} = \|\vec{u}\|\|\vec{v}\| \cos\theta$, donde θ es el ángulo entre los vectores $\vec{u}$ y $\vec{v}$.*

Demostración. Sea $\vec{u} = (u_1, u_2), \vec{v} = (v_1, v_2)$, entonces

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2$$

Utilizando el teorema de los cosenos y midiendo el largo de un segmento por medio de la norma, tenemos:

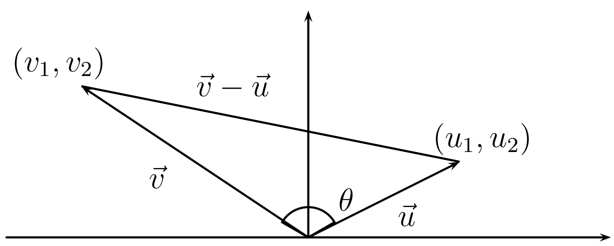


Figura 2.15. Aplicando el teorema de los cosenos.

$\|\vec{u}-\vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\|\vec{u}\|\|\vec{v}\|\cos\theta$. Despejando
$-2\|\vec{u}\|\|\vec{v}\|\cos\theta = \|\vec{u}-\vec{v}\|^2 - \|\vec{u}\|^2 - \|\vec{v}\|^2$
$= \|(u_1-v_1, u_2-v_2)\|^2 - \|(u_1,u_2)\|^2 - \|(v_1,v_2)\|^2$
$= (u_1-v_1)^2 + (u_2-v_2)^2 - u_1^2 - u_2^2 - v_1^2 - v_1^2$
$= u_1^2 - 2u_1v_1 + v_1^2 + u_2^2 - 2u_2v_2 + v_2^2 - u_1^2 - u_2^2 - v_1^2 - v_1^2$
$= -2u_1v_1 - 2u_2v_2$
Resumiendo:
$-2\|\vec{u}\|\|\vec{v}\|\cos\theta = -2u_1v_1 - 2u_2v_2$
al dividir por -2 se tiene el resultado solicitado. ■

112. $\diamond$ ***Teorema.*** *Dos vectores no nulos se cortan en ángulo recto ssi su producto interior es 0. Decimos que los dos vectores son **ortogonales**.*

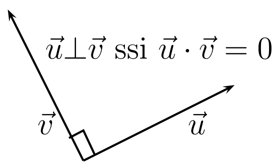


Figura 2.16. Ortogonalidad de vectores.

Demostración. Si los dos vectores se cortan en ángulo recto, 90 ó $\pi/2$, su coseno es cero y, por lo tanto, su producto interno también, pues $\vec{u}\cdot\vec{v} = \|\vec{u}\|\|\vec{v}\|\cos\theta$. Recíprocamente, si el producto punto es cero y ninguno tiene norma cero, el coseno debe ser cero y por lo tanto el ángulo es recto. ■

113. ***Ejercicio + ♣ Definición*** *Compare las dos definiciones siguientes de perpendicularidad y diga cuál es mejor y por qué. Dos vectores $\vec{u},\vec{v}$, son **ortogonales o perpendiculares**, $\vec{u}\perp\vec{v}$, si se cortan en ángulo recto. O bien, dos vectores son perpendiculares si su producto punto es cero.*

114. ***Ejercicio*** *Halle los ángulos entre todos los pares de vectores del ejercicio 106.*

115. ***Ejercicio*** *Demuestre que el vector (a,b) siempre es perpendicular al vector $(b,-a)$ y que lo mismo pasa con los vectores $(1,m),(1,-1/m)$.*

116. ***Ejercicio*** *Halle un vector que quede en el segundo cuadrante, que sea de norma 3 y que sea perpendicular al vector $(1,5)$.*

117. ***Ejercicio*** *Halle un vector en dirección $\theta=\pi/6$ y que mida 8 unidades.*

118. ***Ejercicio*** *Demuestre que si $\vec{u}\perp\vec{v}$ entonces la perpendicularidad se conserva aun si estos vectores se alargan o se acortan: $\lambda\vec{u}\perp\mu\vec{v}$.*

119. ♣ ***Definición.*** *Dos vectores son paralelos ssi uno es múltiplo escalar del otro. Es decir, $\vec{u}$ y $\vec{v}$ son paralelos ssi existe un escalar $\lambda \neq 0$ tal que $\vec{u} = \lambda\vec{v}$, o, $\vec{v} = \lambda\vec{u}$.*

120. ***Ejemplo*** *$(1, 1)$ es paralelo con $(2, 2)$ porque el segundo es 2 veces el primero. Pero $(1, 1)$ y $(1, 3)$ no son paralelos.*

121. $\boxed{\textbf{\textit{Ejercicio}}}$ *En el plano cartesiano, si $\vec{u} \perp \vec{v}$ y si $\vec{v} \perp \vec{w}$, ¿qué relación hay entre $\vec{u}$ y $\vec{w}$? Demuestre su pronóstico.*

Las propiedades del producto interno sobre $\mathbb{R}^n$ se pueden demostrar a partir de la definición operacional que hemos dado. Pero para un EV cualquiera, se hace una definición por propiedades y a partir de ella se deduce todo. Veamos cómo se procede.

El producto interno asocia un número a dos vectores. Pero a pesar de que el resultado no es un vector, se le da el nombre de producto porque cumple la ley distributiva. También cumple otras propiedades listadas en el siguiente teorema, las cuales se toman como definición de producto interior en cualquier espacio vectorial real (con escalares reales).

122. ◇ ***Teorema y ejercicio.*** *El producto interno sobre $\mathbb{R}^n$ cumple las siguientes propiedades, válidas para cualquier par de vectores $\vec{u}$, $\vec{v}$ y para cualquier escalar λ:*

a) Simetría: $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$.

b)Bihomogeneidad escalar: $(\lambda\vec{u})\cdot\vec{v} = \lambda(\vec{v}\cdot\vec{u}) = \vec{v}\cdot(\lambda\vec{u})$. Al multiplicar uno cualquiera de los vectores por un escalar, el producto interno se multiplica por el escalar.

c) Aditividad (distributividad): $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$.

d) No negatividad: $\vec{u} \cdot \vec{u} \geq 0$.

e) Inyectividad: si para todo $\vec{v}$ se tiene que $\vec{u} \cdot \vec{v} = 0$ entonces $\vec{u} = 0$.

Demostración: ejercicio.

123. ♣ ***Definición.*** *Se llama* ***producto interno o interior*** *sobre un espacio vectorial a una función que a cada par de vectores le asocia un número real y que cumple con las propiedades listadas en el teorema anterior, donde además aparecen las propiedades de compatibilidad con la multiplicación por un escalar.*

124. ◇ ***Teorema y ejercicio.*** *El producto interior definido sobre un EV arbitrario cumple con:*

a) $\vec{v} \cdot (\lambda\vec{u}) = \lambda(\vec{v} \cdot \vec{u})$

b) $(\lambda\vec{v}) \cdot (\lambda\vec{u}) = \lambda^2(\vec{v} \cdot \vec{u})$

c) $(\vec{v} + \vec{w}) \cdot \vec{u} = \vec{v} \cdot \vec{u} + \vec{w} \cdot \vec{u}$

d) La función $\sqrt{\vec{u} \cdot \vec{u}}$ cumple las propiedades de una norma por lo que podemos definir $||\vec{u}|| = \sqrt{\vec{u} \cdot \vec{u}}$. Si el producto interno en $\mathbb{R}^n$ se define de la manera usual, la norma definida aquí coincide con la que ya se tenía.

e) No asociatividad: la expresión $\vec{u} \cdot \vec{u} \cdot \vec{w}$ tiene al menos dos interpretaciones no equivalentes.

Demostración: ejercicio.

Habíamos dicho que los pulsos son importantes en tecnología digital, pues un pulso puede interpretarse como un uno y una ausencia de pulso como un cero. Y que con un tren de pulsos puede transmitirse información binaria que es suficiente para transmitir todo tipo de información.

Al definir un pulso, nosotros tomamos el intervalo $[0, n)$, pero ¿qué significa eso tecnológicamente? Que tenemos un reloj que marca el tiempo del sistema y que hemos tomando n unidades de tiempo. Puede ser que un pulso dure un segundo, un milisegundo, un nanosegundo o un femtosegundo. Entre más cortos sean los pulsos, mayor volumen de información pueden transmitir pero más difíciles son de hacer y de controlar. Por ejemplo, se pueden utilizar cristales de cuarzo para estabilizar la frecuencia de los pulsos, pero entre más alta sea la frecuencia, más es el estrés del cristal y podría llegar a romperse. Esto se debe a que cuando se aplica un voltaje ondulatorio a un cristal, el cristal vibra mecánicamente, lo cual crea desplazamientos relativos, tensiones, las cuales pueden dislocarlo. La medida sería entonces disminuir los niveles de potencia, pero si se bajan mucho, quedaría la información tapada por el ruido térmico, un ruido que se genera debido al calor. Por todo esto, nunca ha dejado de ser interesante imaginar circuitos cuyos módulos sean átomos o moléculas con modos de vibración electrónica aislados del movimiento térmico (Boylestad y Nashelsky, 1994).

Tratar de acortar un pulso hasta cero crea severos problemas no sólo técnicos sino también matemáticos. En las matemáticas el tema motivó la definición de la función delta de Dirac como un pulso de duración cero pero de altura infinita para que el área total sea uno. Eso es algo tan escurridizo que su formalización correcta tuvo que esperar hasta casi mediados del siglo XX y se denomina teoría de distribuciones o funciones generalizadas y se estudia como parte del análisis funcional (Yosida, 1978). Demos los primeros pasos en esa dirección.

Recordemos que D_n es el espacio digital sobre el intervalo $[0, n)$, el generado por los pulsos P_i que tienen amplitud uno sobre el intervalo $[i, i+1)$ y cero sobre el resto del intervalo $[0, n)$. Pero ahora tomamos el intervalo $[0, 1)$ y lo dividimos en n subintervalos iguales, cerrados por abajo y abiertos por arriba. A esto se le llama una partición homogénea. Dicha partición induce un **espacio digital** E_n sobre $[0, 1)$, el cual es un espacio vectorial, el del audio digital. Veamos ahora cómo este espacio también tiene su producto interior y cómo se relaciona con D_n.

125. ***Ejercicio*** *Relacione D_2 con E_2 y D_4 con E_4. Decida si la familia D_n es la misma que E_n o si al menos son isomorfos como espacios vectoriales.*

126. ***Ejercicio*** *Un elemento cualquiera de E_n se nota $(c_1, c_2, ..., c_n)$, y el producto interior entre dos elementos de E_n, $\vec{c} = (c_1, c_2, ..., c_n)$ y $\vec{d} = (d_1, d_2, ..., d_n)$ es*

$\vec{c} \cdot \vec{d} = c_1 d_1 + c_2 d_2 + ... + c_n d_n$.

a) Demuestre que nuestra definición satisface todas las propiedades de producto interno.

b) Demuestre que el producto interno que hemos definidio en el espacio digital E_n puede escribirse como una integral. Escriba la norma de un vector usando la forma integral del producto interior de E_n.

c) E_n también es conocido como el conjunto de las funciones escalonadas. Estas funciones se usan para aproximar a las funciones continuas a trozos que también conforman un EV. Demuestre que podemos definir sobre dicho espacio un producto interno, *punto o escalar entre dos funciones, f y g, como el límite cuando n tiende a infinito del producto interno de las funciones escalonadas que aproximan a f y a g y que dicho producto interno toma la sencilla forma siguiente:*

$$f \cdot g = \int_0^1 f(x)g(x)dx$$

127. **Ejercicio** *Dibuje algunas funciones que pertenezcan tanto a E_{16} como a E_8. Dibuje algunas funciones que pertenezcan a E_{16} pero no a E_8. Dibuje una función que no pertenezca a ningún E_n y aproxímela, por instinto y sin cargo de conciencia, por elementos de E_n.*

128. **Ejercicio** *Sea C el conjunto de las funciones continuas definidas sobre el intervalo $[0, 1)$. Atrévase a definir un producto interior en C que respete la relación entre C y E_n.*

a) ¿Qué relación hay entre C y E_n para n grande?

b) ¿Qué pasa cuando n tiende a infinito?

129. **Ejercicio** *Sea T el conjunto de las funciones definidas sobre el intervalo $[0, 1)$, que son continuas a trozos pero con un número finito de discontinuidades. Demuestre que T es un espacio vectorial. ¿Qué relación hay entre C, T y E_n para n grande? Atrévase a definir un producto interior en T que respete su relación con C y con E_n.*

2.7. Líneas en el plano

Primero repasaremos la tecnología más usual para tratar con líneas en el plano. Luego desarrollaremos una tecnología que nos permita tratar con líneas en un espacio de dimensión cualquiera. Nosotros utilizaremos los términos **recta**, **línea** y **línea recta** como sinónimos.

130. ♣ ***Definición.*** *Dos triángulos son semejantes si uno de ellos es la ampliación del otro. O lo que es lo mismo: dos triángulos son semejantes si existen escalas de medida en las cuales los dos triángulos se ven como uno sólo.*

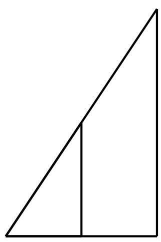


Figura 2.17. Dos triángulos semejantes.

131. ◇ ***Teorema.*** *La ampliación de un triángulo, la cual alarga los lados, conserva los ángulos.*

Demostración. Supongamos que para obtener el triángulo grande multiplicamos los lados del pequeño por el escalar λ. Por el teorema de los cosenos se tiene que para un ángulo cualquiera del triángulo pequeño θ:

$a^2 = b^2 + c^2 - 2bc\cos\theta$. Multiplicando esta ecuación por un escalar λ^2 tenemos:

$\lambda^2 a^2 = \lambda^2 b^2 + \lambda^2 c^2 - 2\lambda^2 bc\cos\theta$.

lo cual puede reescribirse como

$(\lambda a)^2 = (\lambda b)^2 + (\lambda c)^2 - 2(\lambda b)(\lambda c)\cos\theta$.

que representa el teorema de los cosenos para el triángulo grande, pero con el mismo ángulo que el pequeño.

Es decir, los lados del triángulo grande subtienden exactamente los mismos ángulos que el triángulo pequeño. ■

132. ◇ ***Teorema.*** *Al considerar dos triángulos semejantes se tiene que lado uno es a lado uno del otro triángulo como lado dos es a lado dos del otro triángulo.*

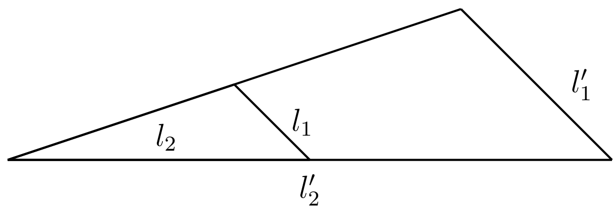


Figura 2.18. Homología y proporcionalidad.

Demostración. Denotemos como l_1, l_2 dos lados del primer triángulo y como l_1', l_2' los dos lados homólogos del segundo triángulo. Se tiene:

(lado uno) es a (lado uno prima)

como

(lado dos) es a (lado dos prima).

O en quebrados,

$l_1/l_2 = l_1'/l_2'$

Demostración:

$l_1' = kl_1$

$l_2' = kl_2$

Por tanto:

$l_1'/l_2' = kl_1/(kl_2) = l_1/l_2$

lo cual implica que

$l_1'/l_1 = l_2'/l_2$

Es bueno aprender a verbalizar esta expresión, por ejemplo: lado uno grande es a lado uno pequeño como lado dos grande es a lado dos pequeño. Puede ser conveniente aclarar que en esta verbalización, cuando decimos *lado grande* o *lado pequeño* realmente estamos diciendo *lo que el lado grande mide* o *lo que el lado pequeño mide.* ■

133. ♣ *La ecuación de la línea.* *Una línea es un conjunto del plano determinado por dos puntos y por triángulos semejantes. Un tercer punto pertenece a la línea si todos los triángulos resultantes son semejantes, como en la figura.*

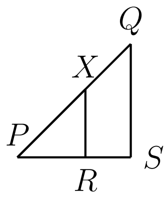


Figura 2.19. Triángulos y rectas.

Un punto X, cualquiera, pertenece a la línea generada por los puntos P y Q si los triángulos PRX, PSQ son semejantes. Eso implica que las medidas de los segmentos siguientes son proporcionales:

$$QS/XR = SP/RP$$

.

Siempre asumiremos triángulos rectángulos, aunque eso no es absolutamente necesario.

134. ***Ejemplo*** *Hallemos la ecuación de la línea que pasa por los puntos $P = (0, 0)$ y $Q = (1, 1)$.*

Solución: Saquemos la ecuación de la recta a partir de triángulos semejantes, pero hagamos una ligera variante en la gráfica, cambiando la posición relativa del punto arbitrario (x, y) con respecto a los puntos dados:

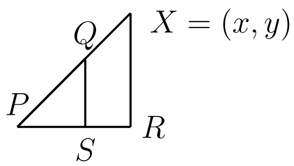


Figura 2.20. La ecuación de una recta.

Un punto X de coordenadas (x, y) está en la línea que pasa por los puntos P y Q siempre y cuando los triángulos PXR y PQS sean semejantes, por tanto $XR/QS = RP/SP$.

Imaginemos el triángulo sobre el plano cartesiano, P en el origen $(0, 0)$,
$Q = (1, 1)$, $S = (1, 0)$, y $X = (x, y)$ entonces $XR = y$, $RP = 1$. Por consiguiente:

$y/1 = x/1$,

o simplemente

$y = x$.

135. *Ejemplo* *Hallemos la ecuación de la línea que pasa por los puntos $P = (1, 3)$ y $Q = (5, 8)$.*

Solución:

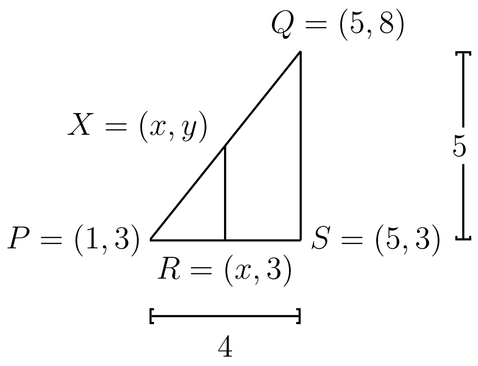


Figura 2.21. La ecuación de una recta que no pasa por el origen.

$QS/XR = SP/RP$ equivale a:
$(8 - 3)/(y - 3) = (5 - 1)/(x - 1)$
$5/(y - 3) = 4/(x - 1)$
reordenando queda:
$4(y - 3) = 5(x - 1)$
$(y - 3) = (5/4)(x - 1)$
$y = (5/4)(x - 1) + 3 = (5/4)x - 5/4 + 12/4 = (5/4)x + 7/4$
que en definitiva nos da:
$y = (5/4)x + 7/4$.

136. ♣ *Definición.* *Cuando la ecuación de una recta se ha escrito de la forma $y = mx + b$, a m se la llama la* ***pendiente****. En el ejemplo anterior la pendiente es $5/4$. Observemos que la pendiente es simplemente cateto opuesto sobre cateto adyacente para el ángulo situado en P. Al ángulo $\theta = Arctg(m)$ se le llama el ángulo de inclinación de la recta. Si la pendiente es $5/4$, el ángulo de inclinación es aproximadamente 51 grados.*

Al coeficiente de x cuando y está despejada se le denomina la pendiente.

137. ⋄ *Teorema.* *Una línea es vertical si y sólo si su ecuación es de la forma $x = k$.*

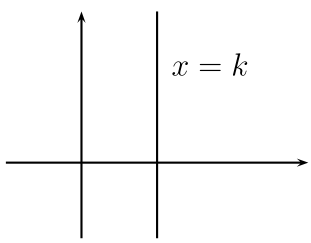


Figura 2.22. La ecuación de una recta vertical.

En efecto, pensemos en la ecuación $x = 5$. En el plano, dicha ecuación representa todos los puntos cuyas coordenadas (x, y) cumplen que $x = 5$. Por ejemplo, los siguientes puntos están en dicha línea: $(5, 3), (5, 8), (5, -1)$ y en general, todos los puntos de la forma $(5, y)$. Por tanto, todos esos puntos están sobre la vertical que pasa exactamente por $(5, 0)$.

138. ⋄ ***Teorema y ejercicio.*** *Toda línea que no es vertical es de la forma* $y = mx+b$ *donde* m *es la pendiente y* b *es el corte con el eje vertical* $\vec{Y}$. *O de otra forma, una línea no vertical que pasa por dos puntos* (x_1, y_1), (x_2, y_2) *tiene pendiente*

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

por lo que su ecuación es
$y - y_1 = m(x - x_1)$
o bien
$y - y_2 = m(x - x_2)$

Demostración. *ejercicio y, por favor, diga exactamente en dónde se usa el hecho de que la recta no sea vertical.*

139. ***Ejemplo*** *Hallemos la ecuación de la línea que pasa por* $(1, 3)$ *y* $(5, 6)$.

Solución: puesto que $y = mx + b$ reemplazamos x por 1, y por 3 y lo mismo con el otro punto:
$3 = m + b \Rightarrow b = 3 - m$
$6 = 5m + b \Rightarrow 6 = 5m + 3 - m = 4m + 3 \Rightarrow m = 3/4 \Rightarrow b = 3 - 3/4 = 9/4$.
Por tanto, la ecuación de la línea es $y = (3/4)x + (9/4)$.

140. ⋄ ***Teorema y ejercicio.*** *La ecuación de una recta cualquiera en el plano es de la forma* $ax + by = c$, *con* a, b, c *números reales, es la* ***ecuación estándar***.

141. ***Ejercicio*** *Halle la ecuación de la línea que pasa por los puntos:*

a) $(2, 3)$ y $(5, -1)$
b) $(2, -4)$ y $(1, 5)$
c) $(-2, 3)$ y $(-5, 3)$
d) $(1, -3)$ y $(1, 6)$
e) $(1, 2)$ y $(-1, 2)$

142. ***Ejercicio*** *Halle la ecuación de la línea que pasa por el punto dado con la pendiente dada:*

a) $(2, 3)$ y -2
b) $(2, -4)$ y 5
c) $(-2, 3)$ y -5
d) $(1, -3)$ y 6
e) $(1, 2)$ y 0

143. **Ejercicio** *Halle la ecuación de la línea que tiene la pendiente dada y el corte con el eje Y dado:*

a) 2 y 1
b) -3 y -1
c) ∞ y -5
d) -1 y -3
e) 0 y -8

144. **Ejercicio** *Halle la ecuación de la línea que pasa por el punto dado y el corte con el eje Y dado:*

a) $(2,3)$ y -8
b) $(2,-4)$ y -3
c) $(-2,3)$ y 4
d) $(1,-3)$ y -3

2.8. Ecuación vectorial de la línea

El enfoque que hemos aprendido en la sección anterior es muy bueno para estudiar líneas en el plano. Pero la generalización de esa metodología al espacio de tres dimensiones es imposible. Aprendamos ahora una tecnología que nos permita tratar con líneas en cualquier dimensión, pero empezaremos con líneas en el plano para después generalizar.

145. ♣ ***Definición.*** *Una línea que pasa por el origen es simplemente el conjunto de todos los vectores que resultan de alargar, acortar o reversar un vector dado, o lo que es lo mismo, que son múltiplos de un vector dado $\vec{D}$, llamado vector director. Informalmente, una línea que pasa por el origen es el conjunto de puntos $\vec{X} = \lambda\vec{D}$. Esto también se interpreta como: para ir desde el origen hasta un punto $\vec{X}$ sobre la línea se camina en la dirección $\vec{D}$ lo que sea necesario, regulando λ, hasta llegar al punto.*

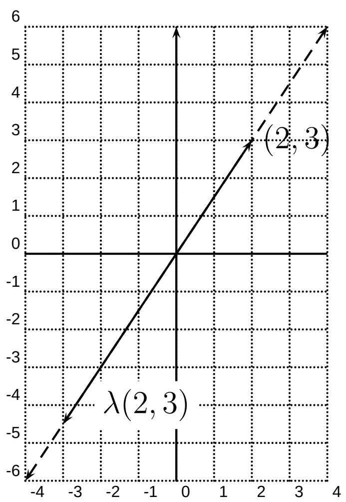


Figura 2.23. Rectas y vectores.

Tenemos en escritura matemática:
$L = \{\vec{X} = (x, y) \in \mathbb{R}^2 : \exists\lambda \in \mathbb{R}, \text{ con } \vec{X} = \lambda\vec{D}\}$

Esta simbología se lee así: la línea L es el conjunto de puntos $\vec{X} = (x, y)$ del plano para los cuales existe un escalar λ de manera que se cumple $(x, y) = \lambda\vec{D}$.

146. ***Ejemplo*** *Definamos en el plano la línea L que pasa por el origen y por el punto $(2, 3)$, es decir, la línea generada por el vector director $(2, 3)$:*

$L = \{\vec{X} = (x, y) \in \mathbb{R}^2 : \exists\lambda \in \mathbb{R}, \text{ tal que se cumple } \vec{X} = \lambda(2, 3)\}$

lo cual lo leemos como: la línea L es el conjunto de puntos $\vec{X}$ del plano tal que existe un escalar λ tal que se cumple $(x, y) = \lambda(2, 3) = (2\lambda, 3\lambda)$. Como la igualdad tiene sentido, coordenada por coordenada, se tiene:

$x = 2\lambda$

$y = 3\lambda$

Si despejamos λ de la primera ecuación $x/2 = \lambda$ mientras que de la segunda $y/3 = \lambda$. Igualando $x/2 = y/3$ la ecuación de la línea es $y = (3/2)x$ que también se escribe $y - (3/2)x = 0$.

147. $\diamond$ ***Teorema y ejercicio.*** *El eje Y es una línea que tiene como ecuación $x = 0$. Todas las demás líneas del plano que pasan por el origen son de la forma $y = mx$, las cuales pasan también por el punto $(1, m)$ formando un segmento cuyo ángulo de inclinación con la horizontal tiene como tangente m y por eso la pendiente es m.*

Demostración: ejercicio.

Ahora, pasemos a considerar líneas del plano que no pasan por el origen. Nuestro punto de partida puede entenderse perfectamente si consideramos la recta $y = 3x + 1$. Tomamos primero dos puntos arbitrarios sobre la línea. Si $x = 0$, $y = 1$ nos da el primer punto $P(0, 1)$, y si $x = 1$, $y = 4$ nos da el segundo punto $Q(1, 4)$.

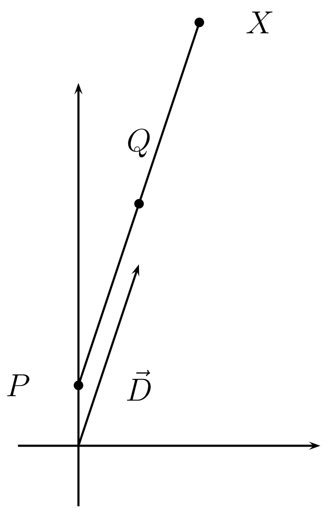


Figura 2.24. El segmento dirigido PQ genera una línea. Este segmento se representa por el vector $\vec{D}$. Cualquier segmento dirigido PX es un múltiplo escalar del vector director D.

Ahora analizamos el vector que empieza en P y termina en Q. Restamos los dos puntos, considerados como vectores $\vec{Q} - \vec{P} = (1,4) - (0,1) = (1,3)$ que corresponde bien con la pendiente de 3 de la recta, si se avanza una unidad en sentido horizontal, se avanza 3 en sentido vertical. A este vector lo llamamos el **vector director** de la línea y lo notamos $\vec{D} = (1,3)$.

Tomemos ahora otros dos puntos arbitrarios sobre la línea, si los restamos, el resultado siempre es un múltiplo del vector director. Algunos ejemplos:

a) $R(2,7)$, $S = (3,10)$, $\vec{S} - \vec{R} = (3,10) - (2,7) = (1,3) = \vec{D}$.

b) $R(1,4)$, $S = (3,10)$, $\vec{S} - \vec{R} = (3,10) - (1,4) = (2,6) = 2(1,3) = 2\vec{D}$.

c)$R(-1,-2)$, $S = (7,22)$, $\vec{S} - \vec{R} = (7,22) - (-1,-2) = (8,24) = 8(1,3) = 8\vec{D}$.

d)$R(a,3a+1)$, $S = (b,3b+1)$, $\vec{S} - \vec{R} = (b,3b+1) - (a,3a+1)$
$= (b-a, 3b+1-(3a+1)) = (b-a, 3b-3a) = (b-a)(1,3) = (b-a)\vec{D}$.

148. ♣ ***Definición.*** *Una línea que pasa por P y tiene* ***vector director*** *$\vec{D}$ es un conjunto de vectores $\vec{X}$ tales que el segmento que parte desde P y llega hasta la cabeza de $\vec{X}$ es un múltiplo del vector director $\vec{D}$. Es decir $\vec{X} - \vec{P} = \lambda\vec{D}$.*

149. ◇ ***Teorema inmediato.*** *Una línea en el plano que pasa por el punto P y con vector director $\vec{D}$ es el conjunto de los puntos $\vec{X}$ de la forma $\vec{X} = \vec{P} + \lambda\vec{D}$. Es decir, si desea llegar a un punto $\vec{X}$ de la línea, llegue desde el origen hasta $\vec{P}$ sobre ella y después quiebre en la dirección del vector director y siga en esa dirección hasta que encuentre el punto buscado.*

150. ***Ejemplo*** *Hallemos la ecuación de la línea que pasa por $(1,4)$ y tiene el vector director $(2,3)$.*

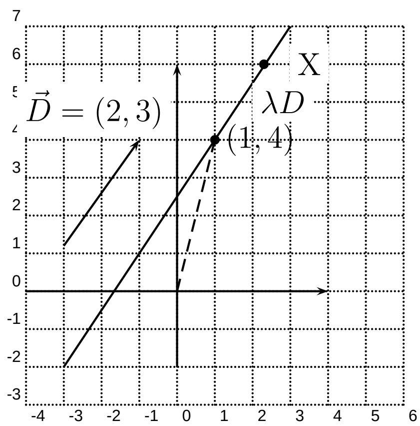


Figura 2.25. Línea que no pasa por el origen.

Para llegar a $\vec{X}$ llegue primero a $(1,4)$ y quiebre después en la dirección $(2,3)$ y camine lo que sea necesario regulando λ hasta llegar a $\vec{X}$:

$\vec{X} = (x,y) = (1,4) + \lambda(2,3) = (1+2\lambda, 4+3\lambda)$.

Igualando coordenada por coordenada tenemos:

$x = 1 + 2\lambda$

$y = 4 + 3\lambda$.

Multiplicando la primera ecuación por 3 y la segunda por 2 y después restando:

$3x - 2y = 3 - 8$ o sea $3x - 2y = -5$ es una línea con pendiente $3/2$, es la misma pendiente de la línea que pasa por $(0,0)$ y tiene vector director $(2,3)$, pues tal línea tiene como ecuación a $3x-2y=0$, y si $x=2$, queda $y=3$, o sea que pasa por cero-cero y luego por $(2,3)$, lo cual da su vector director.

En general, un punto $P(x,y)$ estará en la línea que pasa por $P_o(x_o, y_o)$ y en dirección del vector $\vec{D}$ si el segmento P_oP es paralelo al vector $\vec{D}$, esto es, si existe $\lambda \in \mathbb{R}$ tal que $P_oP = \lambda\vec{D}$, lo cual también se escribe como

$(x - x_o, y - y_o) = \lambda(a,b)$

que da el sistema de ecuaciones

$x - x_o = \lambda a$

$y - y_o = \lambda b$

que en forma vectorial se escribe como

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_o \\ y_o \end{pmatrix} + \lambda \begin{pmatrix} a \\ b \end{pmatrix}$$

o bien

$$\vec{X} = P_o + \lambda\vec{D}$$

.

151. ♣ ***Definición.*** *Dos líneas en el plano son paralelas si nunca se cortan.*

152. ◇ ***Teorema y ejercicio.*** *Dos líneas en el plano son paralelas ssi llevan la misma dirección o, mejor dicho, si sus vectores directores son paralelos.*

Demostración: ejercicio.

El siguiente teorema es una repetición, pero lo dejamos aquí para poder hacer su demostración a partir de la ecuación vectorial de la línea.

153. ◇ ***Teorema.*** *Toda línea en el plano cartesiano es de la forma $x = k$ o $y = mx + b$. El término b es el corte con el eje Y, pues cuando $x = 0$, $y = b$. Como $y = mx + b$ es lo mismo que $mx - y = -b$, y el otro caso es $x = k$, la forma general de una línea es $ax + by = c$.*

Demostración. Las líneas verticales que pasan por el origen todas cumplen la ecuación $x = 0$. Si son verticales y pasan por el punto $(k, 0)$ cumplen la ecuación $x = k$. Las líneas verticales tienen a $(0, 1)$ como vector director. Si el vector director de otra línea no es paralelo a $(0, 1)$, entonces es de la forma (c, d) con $c \neq 0$. Ella pasa por un punto cualquiera $P = (h, k)$. Por lo tanto, un punto cualquiera (x, y) sobre la línea cumple:

$\vec{X} = (x, y) \in L$ ssi $\exists \lambda \in \mathbb{R}$ tal que $(x - h, y - k) = \lambda(c, d) = (c\lambda, d\lambda)$. Igualando coordenadas

$x - h = c\lambda$

$y - k = d\lambda$

Multiplicando la primera ecuación por d y la segunda por c y después restando:

$dx - cy - dh + ck = 0$ o sea $dx - cy = dh - ck$.

Despejando queda $y = (d/c)x - (dh - ck)/c$ que es de la forma $y = mx + b$ con pendiente $m = d/c$, que tiene sentido pues $c \neq 0$. El corte con el eje Y es $b = -(dh - ck)/c$. Esta línea tiene la misma pendiente de la línea que pasa por $(0, 0)$ con vector director (c, d). ■

154. $\diamond$ ***Teorema y ejercicio.*** *Todas las líneas verticales son paralelas. Si dos líneas no son verticales, ellas son paralelas si tienen la misma pendiente. Por ejemplo, la línea* $y = 3x + 1$ *es paralela a la línea* $y = 3x + 7$ *pero ninguna de ellas es paralela a* $y = 2x - 4$.

Demostración: ejercicio.

155. ***Ejercicio*** *Halle la ecuación de la línea que:*

a) Pasa por los puntos $(1, 1), (3, -2)$

b) Es paralela a la línea dada en el inciso a) y pasa por $(7, 5)$

c) Es paralela a la línea dada en el inciso a) pero corta al eje $\vec{Y}$ en 5.

d) Es paralela a la línea dada en el inciso a) pero corta al eje $\vec{X}$ en 8.

e) Es paralela a la línea dada en el inciso a), queda arriba de ésta y guarda una distancia vertical de 5 unidades con dicha línea.

f) Equidista de los puntos $(1, 1)$ y $(3, -2)$.

156. ♣ ***Definición.*** *Dos líneas son perpendiculares si se cortan formando cuatro ángulos iguales. Cualquiera de los ángulos se llama* ***ángulo recto*** *y su medida en grados es 90 y en radianes* $\pi/2$.

157. $\diamond$ ***Teorema.*** *Dos líneas son perpendiculares ssi sus vectores directores son ortogonales. Ejemplo, el eje X es perpendicular al eje Y porque sus vectores directores son $(0,1),(1,0)$ cuyo producto punto da cero.*

Demostración: ejercicio.

158. $\diamond$ ***Teorema y ejercicio.*** *Dos rectas de la forma $y = mx+b$ son perpendiculares ssi el producto de sus pendientes es -1.*

Demostración: ejercicio.

2.9. Proyecciones

Pasamos ahora a calcular la sombra de un vector $\vec{u}$ sobre otro $\vec{v}$, donde imaginamos que la sombra es causada por el Sol del medio día, es decir, que la sombra proyectada por el vector $\vec{u}$ sobre otro $\vec{v}$ genera un ángulo recto con el piso donde cae. El nombre oficial para la sombra es **proyección**:

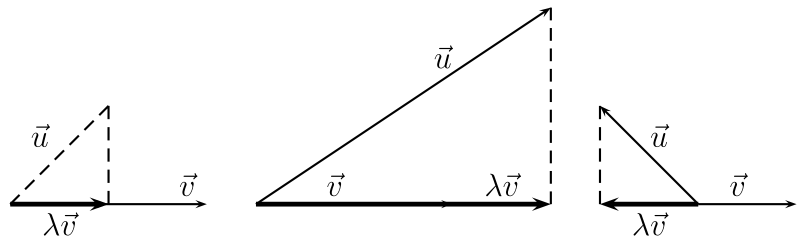


Figura 2.26. La sombra de $\vec{u}$ sobre $\vec{v}$ es el vector reteñido $\lambda\vec{v}$.

Obsérvese que la sombra de $\vec{u}$ sobre $\vec{v}$ es un acortamiento o alargamiento de $\vec{v}$ que puede estar en el mismo sentido o en sentido contrario, pero siempre en la misma dirección de $\vec{v}$. Esto significa que la sombra es un múltiplo escalar de $\vec{v}$.

159. ♣ ***Definición.*** *Sea el vector $\vec{u}$, cuya cabeza es el punto U, y sea $\vec{v}$ otro vector. La **proyección** de $\vec{u}$ sobre $\vec{v}$, ver figura anterior, es el vector $Proy_{\vec{v}}(\vec{u})$, con cabeza en P, y que es múltiplo del vector donde cae la proyección, i.e. $Proy_{\vec{v}}(\vec{u}) = \lambda\vec{v}$, tal que el origen, U y P definen un triángulo rectángulo en P.*

160. $\diamond$ ***Teorema y ejercicio.*** *La proyección o sombra de un vector $\vec{u}$ sobre un vector $\vec{v}$ es $Proy_{\vec{v}}(\vec{u}) = \lambda\vec{v}$, donde $\lambda = (\vec{u}\cdot\vec{v})/(\vec{v}\cdot\vec{v})$.*

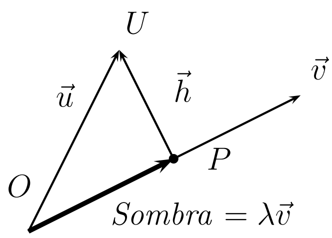


Figura 2.27. La proyección o sombra de $\vec{u}$ sobre $\vec{v}$.

Hagamos dos pruebas de este teorema, una que trabaja con segmentos y otra que se basa en vectores.

Prueba 1. Como estamos en un triángulo rectángulo y necesitamos determinar un cateto adyacente, decimos, cateto adyacente (la proyección) es igual a hipotenusa (el vector $\vec{u}$) por el coseno del ángulo entre la hipotenusa y el cateto adyacente. El coseno lo podemos reemplazar del producto punto para obtener:

$||\text{Proy}_{\vec{v}}(\vec{u})|| = \cos\theta ||\vec{u}||$.

Obsérvese que esta igualdad es entre normas o magnitudes de vectores. Pero además

$\vec{u}\cdot\vec{v} = ||\vec{u}||||\vec{v}||\cos\theta$

por lo que

$\cos\theta = \frac{\vec{u}\cdot\vec{v}}{||\vec{u}||||\vec{v}||}$

y entonces

$||\text{Proy}_{\vec{v}}(\vec{u})|| = \frac{\vec{u}\cdot\vec{v}}{||\vec{u}||||\vec{v}||}||\vec{u}||$

Por otra parte, la proyección es un vector. Esto quiere decir que tiene dirección, la cual está dada por $\vec{v}/||\vec{v}||$, el cual es un vector unitario en la dirección de $\vec{v}$. Entonces,

$\text{Proy}_{\vec{v}}(\vec{u}) = ||\text{Proy}_{\vec{v}}(\vec{u})||\frac{\vec{v}}{||\vec{v}||}$

Reemplazando, simplificando y recordando que $||\vec{v}||^2 = \vec{v}\cdot\vec{v}$ obtenemos la respuesta (ejercicio).

Prueba 2. El punto U es la cabeza de $\vec{u}$ y P la de la $\text{Proy}_{\vec{v}}(\vec{u}) = \lambda\vec{v}$. Sea $\vec{h}$ el segmento PU. Tenemos:

$\vec{u} = \text{Proy}_{\vec{v}}(\vec{u}) + \vec{h} = \lambda\vec{v} + \vec{h}$

$\vec{h} = \vec{u} - \lambda\vec{v}$

Como el origen O, junto con U y P forman un triángulo rectángulo, el vector $OP = \lambda\vec{v}$ debe ser perpendicular a PU que se representa por el vector $\vec{h}$ (que en realidad sale del origen y es paralelo a PU); i.e.:

$\vec{h}\cdot\vec{u} = 0$

$(\vec{u} - \lambda\vec{v})\cdot\vec{v} = 0$

Como vimos en el teorema 122, el producto punto se llama producto porque distribuye a la suma:

$\vec{a}\cdot(\vec{b}+\vec{c}) = \vec{a}\cdot\vec{b} + \vec{a}\cdot\vec{c}$.

Aplicando esta propiedad a la ecuación anterior tenemos

$\vec{u}\cdot\vec{v} - \lambda\vec{v}\cdot\vec{v} = 0$

$\vec{u}\cdot\vec{v} = \lambda\vec{v}\cdot\vec{v}$

por tanto

$\lambda = (\vec{u}\cdot\vec{v})/(\vec{v}\cdot\vec{v})$

y por consiguiente

$\text{Proy}_{\vec{v}}(\vec{u}) = [\frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}]\vec{v} = \frac{\vec{u}\cdot\vec{v}}{||\vec{v}||^2}\vec{v}$

161. $\boxed{\textbf{\textit{Ejercicio y definición}}}$ *Demuestre que* $Proy_{\vec{v}}(\vec{u}) = \dfrac{\vec{u}\cdot\vec{v}}{||\vec{v}||}\dfrac{\vec{v}}{||\vec{v}||}$.

Al escalar $\dfrac{\vec{u}\cdot\vec{v}}{||\vec{v}||}$ *se le llama la* ***componente del vector*** $\vec{u}$ ***sobre*** $\vec{v}$ *y se nota* $Comp_{\vec{v}}\vec{u}$*. Por lo que*

$Proy_{\vec{v}}(\vec{u}) = (Comp_{\vec{v}}\vec{u})\frac{\vec{v}}{||\vec{v}||}$.

Lo cual nos dice que el vector $Proy_{\vec{v}}(\vec{u})$ *es un vector en la dirección de* $\vec{v}$. *La componente es útil para calcular la magnitud del vector proyección, pues basta con calcular el valor absoluto de la componente. Demuestre que*
$||Proy_{\vec{v}}(\vec{u})|| = |Comp_{\vec{v}}\vec{u}| = \frac{|\vec{u}\cdot\vec{v}|}{||\vec{v}||}$.

162. **Ejemplo** *Calculemos la proyección de* $\vec{u} = (1,3)$ *sobre* $\vec{v} = (1,1)$ *y también su norma:*
Utilizando la fórmula
$Proy_{\vec{v}}(\vec{u}) = [\dfrac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}]\vec{v} = \dfrac{\vec{u}\cdot\vec{v}}{||\vec{v}||^2}\vec{v}$
tenemos:
$Proy_{(1,1)}((1,3)) = [((1,3)\cdot(1,1))/((1,1)\cdot(1,1))](1,1)$
$Proy_{(1,1)}((1,3)) = [(1+3)/(1+1)](1,1) = 2(1,1) = (2,2)$.
La norma de la proyección es $\sqrt{8}$, *la cual también puede calcularse por el método de la componente:*
$||Proy_{\vec{v}}(\vec{u})|| = |Comp_{\vec{v}}\vec{u}| = \frac{|\vec{u}\cdot\vec{v}|}{||\vec{v}||}$
lo que nos da:
$||Proy_{\vec{v}}(\vec{u})|| = |(1,3)\cdot(1,1)|/\sqrt{2} = 4/\sqrt{2} = 2\sqrt{2} = \sqrt{8}$.

163. **Ejercicio** *Calcule la proyección del primer vector sobre el segundo y la norma correspondiente en cada uno de los siguientes casos. Utilice el método de la componente para la norma:*

a) $(1,2)$, $(3,3/2)$

b) $(1,-4)$, $(-5,1)$

c) $(-1,6)$, $(3,-3)$

d) $(-1,-5)$, $(-3,-4)$

e) $(1,2)$, $(3,-5)$.

164. **Ejercicio** *Demuestre que*
$Proy_{\vec{v}}(\vec{u}_1+\vec{u}_2) = Proy_{\vec{v}}(\vec{u}_1) + Proy_{\vec{v}}(\vec{u}_2)$
$Proy_{\vec{v}}(\lambda\vec{u}) = \lambda Proy_{\vec{v}}(\vec{u})$.

165. **Ejercicio** *Averigüe (dé un contraejemplo si es falso o pruebe si es verdadero) si se cumple que*
a) $Proy_{\vec{v}_1+\vec{v}_2}(\vec{u}) = Proy_{\vec{v}_1}(\vec{u}) + Proy_{\vec{v}_2}(\vec{u})$
b) $Proy_{\lambda\vec{v}}(\vec{u}) = \lambda Proy_{\vec{v}}(\vec{u})$.

Como en la proyección que hemos definido, cada punto se proyecta en ángulo recto, la proyección es útil para calcular distancias, en algunos casos, su uso se optimiza con la ayuda de la componente.

166. **Ejemplo** *Encontremos la distancia* d *del punto* $U(1,3)$ *a la línea* $y = x$.

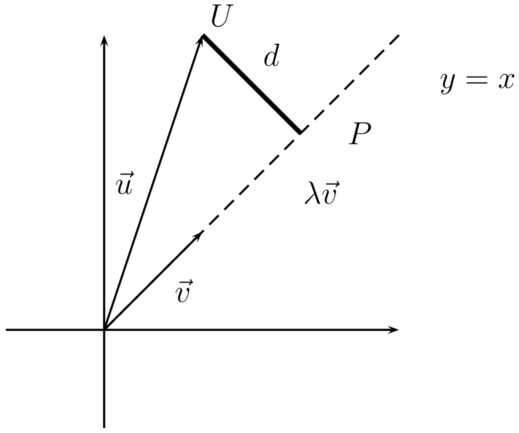


Figura 2.28. La distancia de U a la línea $y = x$ es la magnitud del segmento reteñido, que podemos hallar usando la componente.

Observemos que
$d = \sqrt{||\vec{u}||^2 - |\mathrm{Comp}_v(u)|^2}$
pero $|\mathrm{Comp}_v(u)| = 4/\sqrt{2}$
luego
$d = \sqrt{10-8} = \sqrt{2}$.

167. $\boxed{\textbf{\textit{Ejercicio}}}$ *Encuentre la distancia del punto a la línea, dado:*
a) $y = x$, $(2, 0)$
b) $y = 3x + 1$, $(1, 2)$
c) $y - 3x = 3$, $(2, -3)$
d) $y - 4x = -2$, $(1, 0)$
e) $3x - 7y = 5$, $(4, 5)$
f) $3x + 2y = 1$, $(1, -1)$.

168. $\boxed{\textbf{\textit{Ejercicio}}}$ *Encuentre la ecuación de la recta que está por arriba de la recta dada (las dos rectas son paralelas) y, ademas, a k unidades por encima de ésta, si*
a) Recta: $y = x$, $k = 5$
b) Recta: $y = 2x + 1$, $k = 6$
c) Recta: $y = 3x + 4$, $k = 7$
d) Recta: $y = 7x - 8$, $k = 8$

Todo lo que hemos hecho ha sido desarrollado para el plano. Algunas cosas tienen interpretación directa en el espacio tridimensional, pero para dimensiones superiores, uno debe comenzar con definiciones del siguiente estilo:

169. ♣ ***Definición.*** *En cualquier EV, definimos la* ***recta*** *que pasa por la cabeza de $\vec{P}$ y que tiene vector director $\vec{D}$ como el conjunto de vectores de la forma $\vec{X} = \vec{P} + \lambda\vec{D}$ para $\lambda \in \mathbb{R}$.*

Si en el EV también se ha definido un producto interior, entonces podemos definir la componente de un vector a lo largo de otro y la proyección.

La componente del vector $\vec{u}$ sobre $\vec{v}$ se nota $Comp_{\vec{v}}\vec{u}$ y se define como

$Comp_{\vec{v}}\vec{u} = \frac{\vec{u}\cdot\vec{v}}{||\vec{v}||}$.
Y a la proyección de $\vec{u}$ sobre $\vec{v}$ se nota y se define como
$Proy_{\vec{v}}(\vec{u})$ =$Comp_{\vec{v}}\vec{u}\ \frac{\vec{v}}{||\vec{v}||}$.

Uno podría pensar que los EV de muchas dimensiones, aun infinitas, son arte puro. En realidad, son muy usuales. Demostrémoslo haciendo una pequeña referencia a la síntesis y percepción del sonido.

170. **_Ejercicio para el profesor_** *a) Por pura intuición, y como un buen ejercicio, tome la función $f(x) = sen\, x$, sobre $[0, 2\pi)$ y aproxímela por elementos de los espacios digitales D_4, D_8, D_{16}.*

b) Sea T el espacio vectorial de funciones continuas a trozos sobre $[0, 2\pi)$. Defina el producto interior sobre T como una integral.

c) Halle rigurosamente la proyección de f sobre D_4, D_8, D_{16} y compare los resultados con las aproximaciones obtenidas en a). Recordemos que un espacio digital se construye a partir de los pulsos definidos sobre un intervalo dado.

d) Rigurosamente, ¿qué debe entenderse por digitalización del sonido, es decir, ondas sonoras registradas eléctricamente, que a la larga no son más que funciones de $\mathbb{R}$ en $\mathbb{R}$?

171. **_Ejercicio de investigación_** *Definir las funciones periódicas, probar que forman un espacio vectorial, proveerlo de un producto interior. Demostrar que los senos sen(mx) y los cosenos cos(mx) forman una base infinita de dicho espacio, cuyos elementos son todos mutuamente perpendiculares. Atrévase a decir cuál podrá ser el fundamento matemático de un sintetizador sinusoidal lineal de sonido. Trate de explicar por qué los sintetizadores de sonido digitales, que operan sobre pulsos, y los espacios vectoriales que ellos generan, D_n o E_n, les han ganado a los sinusoidales en el mercado del sonido electrónico (Beauchamp, 2007). (En realidad, hay una variante de los sintetizadores sinusoidales que hace la síntesis de sonido modulando la frecuencia y estos sí han resultado competitivos y muy profesionales).*

2.10. Traslaciones

172. ♣ ***Definición.*** *Una **traslación** no es más que sumar un vector fijo.*

173. **_Ejemplo_** *El punto $(1, 2)$ trasladado $(1, -1)$ se convierte en $(2, 1)$, pues* $(1, 2) + (1, -1) = (2, 1)$.

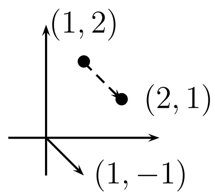


Figura 2.29. El punto $(1, 2)$ trasladado $(1, -1)$ se convierte en $(2, 1)$.

174. *Ejemplo* *Traslación de una circunferencia.*

Una circunferencia centrada en cualquier parte $C = (h, k)$ es una circunferencia centrada en el origen $(0, 0)$ y trasladada el vector C. En efecto, una circunferencia que pasa por el origen tiene como ecuación $x^2 + y^2 = r^2$.

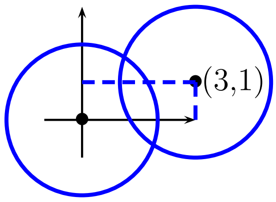


Figura 2.30. Una circunferencia centrada en cero fue trasladada $(3, 1)$.

Si la trasladamos el vector C, queda una circunferencia centrada en C, cuyos puntos tienen como coordenadas (w, z). Podemos hallar la ecuación de dicha circunferencia midiendo distancias. Pero hay otro método: las coordenadas (w, z) son tales que al antitrasladarlas al origen, quedarán sobre la circunferencia centrada en el origen y cumplirán la ecuación de dicha circunferencia. Antitrasladarla equivale a restar (h, k), o sea el punto (w, z) se transforma en $(w-h, z-k)$. Estas nuevas coordenadas cumplen la ecuación de la circunferencia que pasa por el origen:

$(w - h)^2 + (z - k)^2 = r^2$

o cambiando de nombre a las variables queda como es usual

$(x - h)^2 + (y - k)^2 = r^2$.

175. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que una esfera centrada en cualquier parte* $C = (h, k, l)$ *es una esfera centrada en el origen* $(0, 0, 0)$ *y trasladada el vector* C.

La tecnología que hemos aplicado a las traslaciones es una parte especial de otra más general y de la cual presentamos otra aplicación. Imaginemos que encima del plano está superpuesto otro de caucho y que sobre él hemos dibujado una circunferencia con radio 1 y centro en el origen. Estiremos el caucho horizontalmente a veces. Instintivamente sabemos que resulta una elipse. Veamos ahora cómo se prueba que en verdad lo es. Un punto arbitrario (x, y) sobre la supuesta elipse al ser encogido horizontalmente a veces se transforma en un punto de la circunferencia unitaria. Por lo tanto, cumple la ecuación de dicha figura:

$(x/a)^2 + y^2 = 1$, la cual es una elipse.

176. $\boxed{\textbf{\textit{Ejercicio}}}$ *Repetir el razonamiento anterior cuando el caucho se estira*

a) en el sentido vertical únicamente b unidades.

b) horizontalmente a unidades y verticalmente b unidades.

177. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que una línea cualquiera es una traslación de una línea que pasa por el origen.*

178. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que una línea cualquiera es invariante (queda igual) ante una traslación adecuada. Halle una relación entre el vector traslación y el vector director de la línea dada.*

2.11. Sistemas 2×2

Con la geometría que hemos visto podemos ya entender el significado de la cantidad de soluciones de un sistema de ecuaciones lineales (que se pueden representar con matrices) cuando dicho sistema tiene 2 incógnitas.

Si son dos ecuaciones con dos incógnitas, cada ecuación representa una línea.

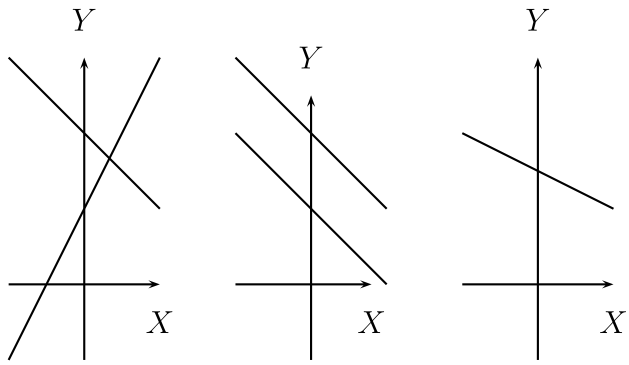


Figura 2.31. Un sistema 2×2 representa un par de líneas, las cuales pueden cortarse en un punto (izquierda), o ser paralelas y nunca cortarse (centro) o coincidir y tener todos los puntos en común (derecha).

Un par de líneas puede constar de: a) dos líneas que tienen pendiente diferente y que por lo tanto se cortan en un punto, dando una solución única; b) dos líneas que tienen igual pendiente pero que no son la misma línea, en ese caso no hay solución; o c) dos líneas que son la misma, es decir, todos los puntos de la una pertenecen a la otra y tenemos un número infinito de soluciones. Con un poco más de precisión, cuando las dos líneas coinciden se dice que hay un número infinito de soluciones con un grado de libertad, lo cual quiere decir que podemos andar por la línea solución para adelante o para atrás.

¿Qué significa en términos prácticos que la solución a un sistema 2×2 sea única? Algo muy práctico es alimentarse bien, imaginemos que sólo necesitáramos **glúcidos** o harinas que vienen en la papa y **proteínas** que vienen en la carne y los cereales. Glúcidos y proteínas deben combinarse en determinadas proporciones y cantidades según la edad y la actividad. Podemos usar dos tipos de alimentos para llenar los requerimientos, el uno rico en proteínas y el otro rico en glúcidos. En general, tendremos un sistema 2×2 con una única solución, es decir, representado por dos líneas que se cortan. La unicidad de la solución significa lo que la mamá no puede alimentar al niño con lo que le venga en gana, sino que debe estudiar algo sobre dietética para combinar correctamente los alimentos. De igual modo, las autoridades no pueden dejar que los restaurantes alimenten a sus clientes con lo más barato de la temporada. Debe hacer una prescripción que refleje un adecuado balance de los diferentes elementos requeridos y cobrar exageradas multas a quien no haga bien las cosas.

¿Y qué significaría el caso de un sistema 2×2 que represente dos líneas paralelas que son la misma línea? Puede tratarse de llenar los requerimientos de glúcidos y de nada más tomando dos alimentos ricos en glúcidos, digamos arroz y papa. Con el deseo de mejorar el buen gusto se puede intercambiar caprichosamente lo uno por lo otro si de lo único que se trata es de llenar los requerimientos de glúcidos, digamos para la comida de la noche, una comida liviana.

Es improbable que se dé el caso de productos naturales que conlleven a un sistema que corresponda a dos líneas que son diferentes pero paralelas y sin puntos comunes, sin solución. Sería más plausible en tecnología de alimentos sintéticos y correspondería a tomar dos productos que tengan el uno, digamos, una proporción de glúcidos del 20 % y que sea igual a la proporción de proteínas, y el otro producto, una proporción de glúcidos del 30 % igual a la de proteínas. Son productos de valor dietético no diferenciado, seguramente no tendrían mercado.

179. **_Ejercicio_** *Invente un significado en la industria que sea de acuerdo con su carrera para un sistema* 2×2 *representado por dos líneas paralelas distintas, otro por dos líneas que se cortan en un único punto y otro por dos líneas que coinciden.*

180. **_Ejemplo_** *Resolvamos el sistema:*

$$\begin{cases} 2x + 3y = 5 \\ 4x + 6y = 10 \end{cases}$$

Dividiendo la segunda ecuación por 2 encontramos la primera. Por lo tanto, se trata de dos líneas superpuestas, todos los puntos de una de ellas son solución de la primera ecuación y por ende de la segunda. Hay infinitas soluciones de la forma $(x, (5 - 2x)/3)$ con un grado de libertad, es decir, con un parámetro libre, la x, en vez de x puede ponerse cualquier número y se tiene una solución. Por ejemplo, con $x = 1$ tenemos la solución $x = 1, y = 1$. Con $x = 25$ tenemos la solución particular $x = 25, y = -15$.

Cuando uno tiene una solución con uno o más parámetros libres, cuando todos se reemplazan por números, da una solución específica que se llama **solución particular**. Por ejemplo; (1,1) es una solución particular al sistema dado.

181. **_Ejercicio_** *Considere las 5 ecuaciones siguientes. Estudie 5 sistemas* 2×2 *generados por varios pares de ellas. Interprete la cantidad de soluciones en términos geométricos. Asegúrese de encontrar los 3 casos posibles. Las líneas son:*

a) $2x - 3y = 5$

b) $2x + 3y = 3$

c) $3x - 2y = 3$

d) $4x - 6y = 10$

e) $6x - 4y = 7$.

182. $\boxed{\textbf{\textit{Ejercicio}}}$ *Trace 5 líneas sobre el plano cartesiano, por lo menos dos paralelas entre ellas. Halle la ecuación de cada línea. Tome algunos pares de esas líneas, prediciendo la cantidad de soluciones de cada sistema correspondiente. Verifique algebraicamente su pronóstico.*

183. $\boxed{\textbf{\textit{Ejercicio}}}$ *Una industria necesita varios tipos de materias primas para poder hacer sus productos. Consideremos tan sólo un producto con dos tipos de materia prima $P1$ y $P2$ que tienen dos tipos de componentes $C1$ y $C2$. La materia prima $P1$ tiene el 10 % del componente $C1$ y el 1 % del componente $C2$ que se necesita para hacer una unidad del producto. La materia prima $P2$ tiene el 40 % del componente $C1$ y material sustituto de componente $C2$ en equivalente al 5 % de los requerimientos para hacer una unidad de producto. ¿Cuántas unidades de cada materia prima se necesitan para hacer un pedido de 300 unidades del producto? Demuestre geométricamente que nuestro problema tiene una solución única.*

2.12. Planos

184. ***Advertencia.*** *El contenido de verdad de las proposiciones depende del contexto que se le dé al discurso. Así por ejemplo, la ecuación $x = 0$ es una expresión que no significa nada. Pero esa ecuación adquiere significado si uno especifica en qué espacio está. Si decimos $x = 0$ en $\mathbb{R}$, estamos diciendo que nos referimos a los puntos de la recta real con coordenada 0. No hay más que un punto en la recta real que satisface dicha ecuación y es el origen. Pero $x = 0$ en el plano $\mathbb{R}^2$ es una ecuación que se satisface por todos los puntos $P(x, y)$ cuya primera coordenada es cero, es decir, todos los puntos de la forma $(0, y)$ como $(0, 0)$, $(0, 1)$, $(0, 10)$, $(0, -8)$. Vemos que todos estos puntos están alineados sobre el eje Y y, por lo tanto, forman una línea. En cambio en $\mathbb{R}^3$ la ecuación $x = 0$ representa todos los puntos (x, y, z) tales que su primera coordenada es cero, o sea $(0, y, z)$ como por ejemplo $(0, 0, 0)$, $(0, 1, 3)$, $(0, 1, 7)$, $(0, 10, -4)$,$(0, 10, 5)$, $(0, -8, 7)$. En general, una línea en el plano 2D (dos dimensiones) es de la forma $ax + by = c$, pero la ecuación $ax + by + cz = d$ no es una línea en 3D (tres dimensiones) sino que es un plano. En particular, $ax + by = c$ representa en 3D el plano $ax + by + 0z = c$. Todo esto vamos a probarlo ahora mismo.*

Hay varias maneras de determinar un plano y una de ellas es la siguiente: un plano está determinado por un punto, donde pasa el plano, y por un vector que es perpendicular al plano y se llama el vector normal del plano. En la siguiente definición continuamos usando la equivalencia entre un vector (una flecha) y su cabeza (un punto), a tal grado que intercambiamos lo uno con lo otro. En el ejemplo que sigue y en adelante usamos dos notaciones equivalentes para un punto que son $P = (x, y, z)$ y $P(x, y, z)$. También adoptamos la misma equivalencia de notaciones para vectores.

185. ♣ ***Definición.*** *Un plano Π es un conjunto de puntos $\vec{X}$ tales que el segmento que une el punto $\vec{X}$ a un punto fijo $\vec{P}$ es perpendicular a un vector fijo llamado vector normal al plano $\vec{N}$. Por tanto, un plano cumple la ecuación $(\vec{X} - \vec{P}) \cdot \vec{N} = 0$. En un ejemplo puede entenderse todo.*

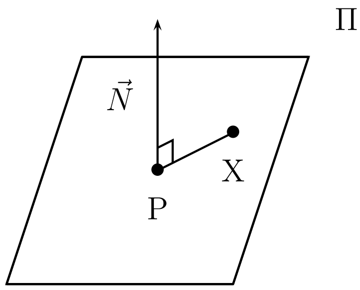


Figura 2.32. Un plano en $\mathbb{R}^3$.

186. ***Ejemplo*** *Determinemos la ecuación del plano que pasa por el punto $\vec{P}(1,2,3)$ y que tiene como vector normal al vector $\vec{N}(4,5,6)$.*

Primero designamos un punto cualquiera del espacio como $\vec{X}(x,y,z)$, el cual además representa al vector que sale del origen y llega a ese punto. Hay puntos del espacio que no pertenecen al plano pero hay otros que sí están en el plano. La condición para que (x,y,z) esté en el plano es que el segmento que va desde $(1,2,3)$ al punto (x,y,z) sea perpendicular al vector $(4,5,6)$. El segmento que va desde $(1,2,3)$ al punto (x,y,z) es simplemente $(x-1, y-2, z-3)$.

Ahora bien, este vector debe ser perpendicular al vector normal $(4,5,6)$, por lo que realizando el producto punto y luego expandiendo:

$(x-1, y-2, z-3) \cdot (4,5,6) = 0$

$4(x-1) + 5(y-2) + 6(z-3) = 0$

$4x + 5y + 6z = 4 + 10 + 18 = 32$, la cual es la ecuación requerida.

En general, y teniendo en mente la misma figura, tomamos $\vec{P}$ de coordenadas $\vec{P}(x_o, y_o, z_o)$, $\vec{N}(a,b,c)$ y un punto arbitrario $\vec{X}(x,y,z)$. De la ecuación

$(\vec{X} - \vec{P}) \cdot \vec{N} = 0$

se tiene:

$(x - x_o, y - y_o, z - z_o) \cdot (a,b,c) = 0$

o bien

$a(x - x_o) + b(y - y_o) + c(z - z_o) = 0$

que es la ecuación del plano que pasa por $\vec{P}(x_o, y_o, z_o)$ y cuyo vector normal es $\vec{N}(a,b,c)$. Esta ecuación es equivalente a

$ax + by + cz = ax_o + by_o + cz_o$,

o bien,

$ax + by + cz = d$ con $d = ax_o + by_o + cz_o$.

A veces es muy conveniente poder visualizar un plano, al menos en la mente. Una manera cómoda de lograrlo es encontrar los cortes del plano con cada uno de los tres ejes.

187. **Ejercicio** *Observando que el eje X cumple con las ecuaciones $y = 0$ y $z = 0$, hallar el corte del plano hallado anteriormente, $4x+5y+6z = 32$, con el eje X. Repita lo mismo con los otros dos ejes. Así se hallan tres puntos que también determinan el plano (¿por qué?). Bosqueje un dibujo que indique el triángulo formado por los tres puntos.*

188. **Ejemplo** *Demostremos que $z = 0$ es un plano. Geométricamente esto es obvio, pues $z = 0$ es una condición cumplida por todos los puntos del 'piso' del espacio 3D, el cual es un plano. La demostración algebraica consiste en demostrar que existe un vector fijo, $\vec{N}$, que es perpendicular a todos los segmentos del plano. Procedamos:*

$z = 0$ puede reescribirse como $0(x-0)+0(y-0)+1(z-0) = 0$, lo cual en lenguaje de perpendicularidad se lee: $(0,0,1)\cdot(x-0, y-0, z-0) = 0$, es decir, el vector fijo $\vec{N} = (0,0,1)$ es perpendicular a todo segmento que empieza en (x,y,z) y termina en $(0,0,0)$.

189. **Ejemplo** *Demostremos que $2x + 4y - 6z = 7$ es un plano. Tenemos que demostrar que existe un vector fijo, $\vec{N}$, que es perpendicular a todos los segmentos del plano. Procedamos:*

$2x + 4y - 6z = 7 =$ puede reescribirse como
$2(x-a) + 4(y-b) - 6(z-c) = 7 - 2a - 4b + 6c$
donde hemos rellenado la expresión, a ambos lados, para que pueda releerse en el lenguaje de la perpendicularidad: $(2,4,-6)\cdot(x-a, y-b, z-c) = 0 = 7-2a-4b+6c$, o sea el vector fijo $\vec{N} = (2,4,-6)$ es perpendicular a todo segmento que empieza en (x,y,z) y termina en (a,b,c). Pero para lograr esto tenemos que resolver
$0 = 7 - 2a - 4b + 6c$.

Esta es una ecuación con 3 incógnitas. Damos a a el valor 0, a b el valor 0 y por tanto c toma el valor $-7/6$. Vemos que hay 2 grados de libertad: a a y a b se les puede dar el valor que queramos que el de c siempre puede hallarse. ¿Qué significa esto? Que estamos buscando un punto en donde anclar al plano. Y ese punto puede anclarse en cualquier parte del plano, el cual tiene dos grados de libertad (adelante-atrás *vs.* izquierda-derecha).

190. **Ejercicio** *Halle el vector normal a cada uno de los planos siguientes $x = 0$, $y = 0$, $z = 0$, $x = 1$, $x = 2$, $x = 3$, $y = -1$, $z = -3$. Describa y dibuje dichos planos con referencia a un sistema de ejes (en el salón de clase).*

191. **Ejercicio** *Hallar las ecuaciones de algunos planos.*

a) Halle el plano que pasa por $(-1,2,-3)$ y que tiene por vector normal a $(-1,0,5)$.

b) Halle el plano que pasa por $(0-,2,-3)$ y que tiene por vector normal a $(3,2,7)$.

c) Halle el plano que pasa por $(-1,-4,-3)$ y que tiene por vector normal a $(-6,3,8)$.

d) Halle el plano que equidista de los puntos $P(1,1,1)$ y $Q(-2,3,5)$.

e) Halle la ecuación del plano que está por encima del plano $x+y+z=1$ y que está exactamente a 5 unidades de distancia de éste.

192. *Ejemplo* *Hallemos el plano determinado por los puntos $P(0,0,1)$, $Q(0,1,0)$ y $R(1,0,0)$.*

Puesto que todo plano es de la forma $ax+by+cz=d$, reemplazamos los puntos en la ecuación, obtenemos 3 ecuaciones, una por cada punto y 4 incógnitas a,b,c,d. Esto quiere decir que hay un grado de libertad, lo cual significa que el vector normal puede acortarse, alargarse, reversarse y, sin embargo, el resultado también es otro vector normal al plano dado.

Reemplazando $(1,0,0)$ en $ax+by+cz=d$ queda $a=d$
Reemplazando $(0,1,0)$ en $ax+by+cz=d$ queda $b=d$
Reemplazando $(0,0,1)$ en $ax+by+cz=d$ queda $c=d$

La ecuación será entonces $dx+dy+dz=d$. Pero d no puede ser cero pues nos quedaría la ecuación $0=0$ que no dice nada. Dividiendo por d obtenemos: $x+y+z=1$.

193. $\boxed{\textbf{\textit{Ejercicio}}}$ *Halle el plano determinado por los 3 puntos:*

a) $(1,0,0),(1,2,3),(1,5,6)$

b) $(0,1,3),(2,1,5),(4,2,-1)$

c) $(1,2,1),(-1,2,-1),(1,0,1)$

d) $(5,0,0),(4,0,0),(1,0,0)$.

194. ♣ *Definición.* *Consideremos el plano Π_1 con ecuación $z=3x+4y-5$. Podemos escribir la misma ecuación en forma vectorial:*

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 3x+4y-5 \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -5 \end{pmatrix}.$$

*Esta ecuación tiene la forma general que define la **ecuación vectorial** del plano:*

$$\vec{X} = \alpha\vec{u} + \beta\vec{v} + \vec{P}$$

En el ejemplo considerado $\vec{u}=(1,0,3)$ (en forma de columna), $\vec{v}=(0,1,4)$ y $\vec{P}=(0,0,-5)$.

La ecuación vectorial del plano nos dice que $\vec{X}$ está en el plano que pasa por el punto $\vec{P}$ y que es generado por los vectores $\vec{u}$ y $\vec{v}$. Dando diversos valores a α y a β vamos obtenemos diversos puntos del plano y todo punto del plano se obtiene de esta forma. Por esto decimos que el plano tiene dos grados de libertad, dos parámetros libres, lo cual ya se sabía de antemano pues la ecuación de un plano $ax+by+cz=d$ tiene 3 incógnitas y una sola restricción, por lo que quedan dos incógnitas libres, a las cuales *se les puede dar cualquier valor. Que todo esto sirva para tener muy presente que la ecuación $ax + by + cz = d$ no es y no puede ser la ecuación de una línea en 3D.*

Si consideramos el plano Π_o, que pasa por el origen, que consiste de todos los puntos $\vec{X} = \alpha\vec{u} + \beta\vec{v}$, podemos decir que Π_1 es igual a Π_o trasladado el vector $\vec{P}$.

195. **Ejercicio** *Encuentre las ecuaciones paramétricas de los planos hallados en el ejercicio anterior.*

El problema de hallar el vector normal a otros dos ocurre con frecuencia. Para esto podemos utilizar el producto punto. Pero existe para $\mathbb{R}^3$ una manera estándar de hacerlo y se denomina producto cruz.

196. ♣ ***Definición.*** *En $\mathbb{R}^3$, el* **producto cruz** *o* **producto vectorial** *entre dos vectores $\vec{v} = (v_1, v_2, v_3)$ y $\vec{w} = (\vec{w}_1, \vec{w}_2, \vec{w}_3)$ es un vector notado $\vec{v} \times \vec{w}$ y definido por*
$\vec{v} \times \vec{w} = (v_2w_3 - v_3w_2, -(v_1w_3 - v_3w_1), v_1w_2 - v_2w_1)$.
Notemos que en la coordenada i falta el subíndice i y que hay antisimetría (reversa el signo) en cada coordenada.

197. ◇ ***Teorema y ejercicio.*** *El producto cruz entre dos vectores $\vec{v}$, $\vec{w}$ de $\mathbb{R}^3$ cumple las propiedades:*

a) Es anticonmutativo, $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$

b) Es perpendicular a ambos vectores, $(\vec{v} \times \vec{w}) \perp \vec{v}$, $(\vec{v} \times \vec{w}) \perp \vec{w}$.

c) Su norma es igual al área del paralelogramo determinado por los dos vectores, $||\vec{v} \times \vec{w}|| = ||\vec{v}||\,||\vec{w}|| \operatorname{sen}\theta$, donde θ es el ángulo entre $\vec{v}$ y $\vec{w}$.

d) Si $\vec{i} = (1, 0, 0)$, $\vec{j} = (0, 1, 0)$, $\vec{k} = (0, 0, 1)$ son los generadores de las direcciones positivas de los ejes coordenados X, Y, Z respectivamente, entonces $\vec{i} \times \vec{j} = \vec{k}$. Lo que se enfatiza es que $\vec{i} \times \vec{j}$ es $\vec{k}$ y no es $-\vec{k}$ y se dice que el producto cruz define una orientación llamada de la mano izquierda que es la orientación usual de $\mathbb{R}^3$. En general, se forma un trípode de la forma que $\vec{v}$ va en el pulgar, $\vec{w}$ va con el dedo del corazón y $\vec{v} \times \vec{w}$ va con el índice, todos los dedos son de la mano izquierda.

Demostración: ejercicio.

198. **Ejercicio** *Rehaga el ejercicio 193 usando el producto cruz. De las tres maneras posibles de resolver este ejercicio (por reemplazo, por producto punto y por producto cruz), ¿cuál es más fácil?*

199. **Ejercicio** *Paula y Ricardo compraron su lote y se dispusieron a techarlo. El lote está determinado por líneas que pasan por los puntos siguientes $(0, 1)$, $(1, 0)$, $(2, 0)$, $(0, 2)$. Pusieron una columna de 3 metros en $(0, 1)$ y otra en $(1, 0)$, pero sobre los otros dos puntos pusieron columnas de 5 metros. Mandaron a hacer la vigas metálicas, asumiendo que el techo resultaría un plano. Las armaron y todo salió bien: el techo empalmó perfectamente con las columnas. Por favor, averigüe las longitudes de las vigas y el área del techo. Pero resulta que Alejandro y Manuela compraron el lote de* *al lado, rectilíneo, con esquinas con coordenadas* $(2,0)$, $(0,2)$, $(4,0)$, $(0,5)$. *Sobre los primeros dos puntos pusieron columnas de* 3 *metros, y en los dos últimos pusieron columnas de* 5 *metros. Mandaron a hacer las vigas en la misma parte que sus vecinos Paula y Ricardo, las transportaron al lote, las soldaron en el suelo, que era muy plano, y quedó un gran marco plano, pero al tratar de montarlo sobre las columnas no cuadró ni porque le dieron mucho martillo. Ellos llegaron a la conclusión de que el lote era de mala suerte y lo pusieron en venta y lo venderán así le pierdan el* 30 % *del costo. En realidad, la operación de montaje coincidió con el último día de octubre. Diga si lo que ellos cuentan puede ser cierto y en ese caso explique el misterio (¿por qué a Paula y Ricardo les funcionó todo y a ellos no?) y deles un consejo.*

La definición de proyección y su relación con distancias desarrollada para el plano es válida para cualquier dimensión. Veamos cómo se utiliza en el espacio tridimensional.

200. *Ejemplo* *Calculemos la distancia* d *del punto* $P = (4,5,6)$ *al plano* $x + 2y + 3z = 6$.

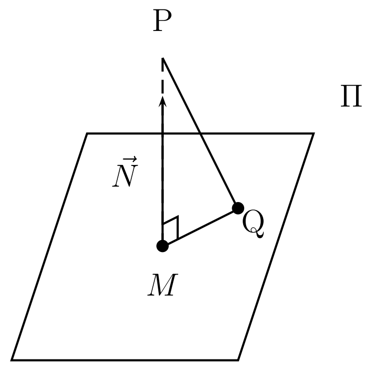


Figura 2.33. Distancia de un punto P a un plano Π.

Utilizamos la siguiente idea: fabricamos un punto Q sobre el plano. Q puede ser cualquiera. Como el vector normal al plano $\vec{N}$ puede ser imaginado en cualquier punto del plano, nos lo imaginamos en el punto M, de tal forma que la línea generada por el vector normal pasa por M y también pasa por P. Se tiene un triángulo rectángulo formado con el segmento normal que va de M a P, el segmento que va de M a Q y el segmento que va de Q a P.

Con la anterior construcción, la distancia del punto P al plano es simplemente la distancia del segmento PM, la cual es la norma de la proyección del vector $\overrightarrow{PQ}$ sobre el normal. Por lo tanto, la distancia es simplemente el valor absoluto de la componente de $\overrightarrow{PQ}$ sobre $\vec{N}$.

Encontremos un punto cualquiera Q sobre el plano. Digamos $x = 1$, $y = 1$ y por tanto $z = 1$. El punto sobre el plano es $Q = (1,1,1)$. El segmento de P a Q es $\overrightarrow{PQ} = (1,1,1) - (4,5,6) = (-3,-4,-5)$. El vector normal $\vec{N}$ al plano es $\vec{N} = (1,2,3)$. En resumen: $d = |\text{Comp}_{\vec{N}}\overrightarrow{PQ}|$

Luego:

$d = |\text{Comp}_{(1,2,3)}((-3,-4,-5))| = |(-3,-4,-5)\cdot(1,2,3)|/\sqrt{1^2+2^2+3^2}$

$= 26/\sqrt{14}$.

Es necesario tener en cuenta que los vectores están todos anclados en el origen. Pero los podemos representar en cualquier punto del espacio por clones adecuados que son segmentos dirigidos. Si los ángulos y magnitudes se conservan, todo quedará bien.

201. **_Ejercicio_** *Rehaga el ejercicio anterior utilizando la noción de componente.*

202. **_Ejercicio_** *Encuentre la distancia d del punto al plano:*

a) $(4, 5, 6)$, $x - 2y + 4z = 6$

b) $(-1, 0, -1)$, $-2x + 3y + 2z = 6$

c) $(0, -2, 3)$, $3x - 5y - 2z = 6$

d) $(2, -3, 4)$, $-2x + 2y - 5z = 6$

e) $(-3, -1, 0)$, $-x - y - 7z = 6$.

Ahora podemos encontrar la distancia entre dos planos paralelos.

203. **_Ejemplo_** *Encontremos la distancia entre dos planos paralelos. El primero es* $x + 2y + 3z = 6$ *y el segundo* $x + 2y + 3z = 32$.

Aclaremos que estos dos planos son paralelos porque tienen el mismo vector normal $\vec{N}(1, 2, 3)$. En general, dos planos son paralelos **ssi (si y sólo si)** sus vectores normales son paralelos, es decir, el uno es múltiplo del otro.

El punto $(4, 5, 6)$ pertenece al segundo plano. Por tanto, la distancia entre los dos planos es igual a la distancia entre el plano $x + 2y + 3z = 6$ y el punto $(4, 5, 6)$. Afortunadamente, nosotros ya resolvimos este problema en el ejemplo 200. Luego la distancia entre los planos es de $26/\sqrt{14}$.

204. **_Ejercicio_** *Rehaga el ejercicio 202 con problemas que consisten en hallar la distancia entre dos planos paralelos.*

2.13. Líneas y planos en $\mathbb{R}^3$

La idea de línea es seguir siempre en la misma dirección. Veamos cómo se implementa esto en tercera dimensión y cómo se resuelven problemas con planos y rectas. Amable lector, si usted siente que va encontrando mucha redundancia, es que usted ya va entendiendo bastante. Nos alegra mucho.

205. ♣ ***Definición.*** *Una línea L en* $\mathbb{R}^3$ *que pasa por el origen es simplemente el conjunto de todos los vectores que resultan de multiplicar uno dado, llamado director, por todos los escalares posibles.*

206. **Ejemplo** *Especifiquemos la línea que pasa por el origen y por el punto* $(2,3,7)$.

En este caso, $L = \{\vec{X} : \vec{X} = \lambda(2,3,7) \text{ con } \lambda \in \mathbb{R}\}$
es decir $\vec{X}(x,y,z) \in L$ ssi $\exists\lambda \in \mathbb{R}$ tal que
$(x,y,z) = \lambda(2,3,7) = (2\lambda, 3\lambda, 7\lambda)$.

Como es natural, la igualdad tiene sentido coordenada por coordenada, por lo tanto, con $\lambda \in \mathbb{R}$:
$x = 2\lambda$
$y = 3\lambda$
$z = 7\lambda$

A este sistema de ecuaciones con un grado de libertad se le denomina **las ecuaciones paramétricas de una recta.** Dando diferentes valores a λ vamos obteniendo puntos sobre la línea. Por ejemplo, para
$\lambda = 0$ obtenemos el punto de la recta $(0,0,0)$,
$\lambda = 1$ obtenemos el punto $(2,3,7)$,
$\lambda = -1$ obtenemos el punto $(-2,-3,-7)$.

También podemos probar que el punto $(6,8,9)$ no está en la recta, pues debe cumplirse que $6 = 2\lambda$, por lo que $\lambda = 3$. Pero también, $8 = 3\lambda$, lo cual no se cumple con $\lambda = 3$.

207. **Ejercicio** *Describa las ecuaciones paramétricas de las líneas que pasan por el origen y además por el punto*

a) $(1,1,4)$
b) $(-1,2,2)$
c) $(2,3,-1)$

Hemos considerado líneas que pasan por el origen. Ahora pasemos a líneas que pasan por cualquier otro punto.

208. ♣ ***Definición.*** *Una línea que pasa por el punto* $\vec{P}$ *y tiene vector director* $\vec{D}$ *es un conjunto de puntos* $\vec{X}$ *tales que el segmento que parte desde* $\vec{P}$ *y llega hasta* $\vec{X}$ *es un múltiplo del vector director* $\vec{D}$. *Es decir,* $\vec{X} - \vec{P} = \lambda\vec{D}$ *o lo que es lo mismo* $\vec{X} = \vec{P} + \lambda D$. *Esta última ecuación se interpreta así: para llegar desde el origen hasta el punto* $\vec{X}$ *de la línea, llegue desde el origen hasta el punto* $\vec{P}$, *quiebre en la dirección de* $\vec{D}$ *y camine la proporción de* $\vec{D}$ *que necesite,* λ, *hasta llegar a* $\vec{X}$.

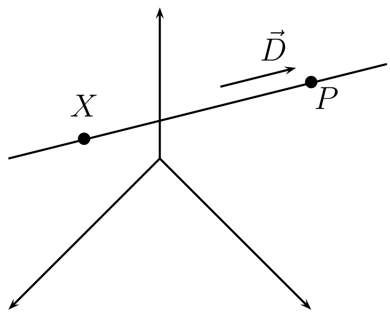


Figura 2.34. Una línea está determinada por un punto P y una dirección que puede ser la de un vector $\vec{D}$ llamado director.

209. ***Ejemplo*** *Hallemos la línea que pasa por* $(5,6,7)$ *y tiene como vector director a* $(2,3,7)$.

Solución: $\vec{X}(x,y,z) \in L$ ssi $\exists\lambda \in \mathbb{R}$ tal que $(x,y,z)-(5,6,7)=\lambda(2,3,7)$, o bien, $(x,y,z)=\lambda(2,3,7)+(5,6,7)=(2\lambda+5,3\lambda+6,7\lambda+7)$. Como la igualdad tiene sentido coordenada por coordenada, tenemos:
$x=2\lambda+5$
$y=3\lambda+6$
$z=7\lambda+7$
A estas ecuaciones se les llama ecuaciones paramétricas de la recta, y a λ el **parámetro**.

210. ***Ejemplo*** *Hallemos las ecuaciones paramétricas de la línea que pasa por los puntos* $(1,2,3)$ *y* $(4,5,6)$.

Solución: necesitamos hallar el vector director. El vector que va de un punto al otro nos puede servir, $\vec{D}=(4,5,6)-(1,2,3)=(3,3,3)$. Como la línea pasa por $(1,2,3)$ la ecuación de la línea es
$(x,y,z)=(1,2,3)+\lambda(3,3,3)$.
Por tanto, con λ como real, las ecuaciones paramétricas de la recta son
$x=3\lambda+1$
$y=3\lambda+2$
$z=3\lambda+3$

Cuando estudiábamos líneas en el plano, pudimos despejar el parámetro λ sin problema alguno. Cuando se trata de líneas en el espacio hay que tener más cuidado. Por ejemplo, despejemos λ de las ecuaciones paramétricas dadas por
$x=2\lambda-1$
$y=3\lambda-8$
$z=7\lambda-1$
Si despejamos λ de las dos primeras ecuaciones nos queda $3x-2y=-3+16$ y si lo hacemos de la segunda y de la tercera queda $7y-3z=-56+3$. En el espacio estas ecuaciones representan planos y no líneas, pues ambas son de la forma $ax+by+cz=d$ con tres incógnitas y sólo una restricción, es decir, con dos incógnitas libres. Tenemos entonces que una línea en 3D se puede entender como la intersección de dos planos. Otra manera de escribir lo mismo es despejando λ de todas las ecuaciones e igualando:

$$(x+1)/2=(y+8)/3=(z+1)/7$$

A este tipo de escritura se le llama las **ecuaciones simétricas** de una recta en 3D. La primera igualdad da un plano y la segunda otro plano, una línea es la intersección de dos planos.

En general, si una recta pasa por el punto $P_o(x_o,y_o,z_o)$ y en la dirección del vector $\vec{D}=(a,b,c)$ entonces la ecuación de dicha recta puede darse en forma paramétrica o en forma simétrica. La forma paramétrica es, con $\lambda \in \mathbb{R}$:
$x=x_o+a\lambda$

$y = y_o + b\lambda$
$z = z_o + c\lambda$
La forma simétrica es:

$$\frac{x - x_o}{a} = \frac{y - y_o}{b} = \frac{y - y_o}{c}$$

si a, b, y c son no nulos.

Todas estas ecuaciones se obtienen de $\vec{X} = P_o + \lambda\vec{D}$, la ecuación de la recta en cualquier dimensión.

211. **Ejercicio** *Halle el vector director, las ecuaciones paramétricas y las simétricas de las líneas que pasan por*

a) $(1, 2, 3)$ y $(2, 3, 4)$
b) $(-1, 0, 2)$ y $(2, -3, -1)$.

212. **Ejercicio** *Halle el vector director, las ecuaciones paramétricas y las simétricas de las líneas que empieza en* $(1, 2, 3)$ *a las* 0 *horas y a las* 8 *va en* $(2, 3, 4)$. *La idea es la de un móvil que va a velocidad constante por una trayectoria que es una línea recta. Se debe formular una relación entre el parámetro* λ *y el tiempo, de tal manera que se cumpla el horario prefijado.*

213. **Ejercicio** *Halle las líneas que son paralelas (tienen el mismo vector director) a las líneas del ejercicio 211 pero que pasan por* $(2, -1, -2)$.

214. **Ejemplo** *Encontremos la distancia d del punto* $P = (2, -1, 5)$ *a la línea con vector director* $\vec{D} = (4, 5, 6)$ *y que pasa por el punto* $(-1, 2, -2)$.

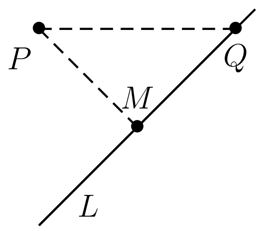


Figura 2.35. Con ayuda de un triángulo rectángulo y proyecciones, o componentes, podemos hallar la distancia de un punto P a una línea L.

Utilizamos la siguiente idea. Encontramos un punto Q sobre la línea. Al proyectar PQ sobre la línea generamos un triángulo rectángulo PQM. La distancia de M a Q puede hallarse por la componente de PQ sobre el vector director de la línea. La distancia del punto a la línea es la magnitud del otro cateto, PM, el cual puede hallarse por Pitágoras. Procedemos:

Encontramos un punto Q sobre la línea, el cual puede ser $Q = (-1, 2, -2)$. Después encontramos el vector $\vec{u}$ definido por el segmento

PQ: $\vec{u} = \vec{P} - \vec{Q} = (2, -1, 5) - (-1, 2, -2) = (3, -3, 7)$.

Hallamos la componente de $\vec{u}$ sobre la línea, i.e., sobre el vector director de la línea $\vec{D}$.

$\mathrm{Comp}_{\vec{D}}(\vec{u}) = |(\vec{u} \cdot \vec{D})|/||\vec{D}||$

$\mathrm{Comp}_{(4,5,6)}((3, -3, 7)) = |(3, -3, 7) \cdot (4, 5, 6)|/\sqrt{4^2 + 5^2 + 6^2}$

$\mathrm{Comp}_{(4,5,6)}((3, -3, 7)) = (12 - 15 + 42)/\sqrt{77} = 39/\sqrt{77}$

Sea M el punto sobre la línea que está en el borde de la proyección del segmento PQ sobre la línea. La distancia del punto P a la línea es exactamente la de P a M, que es la norma del segmento MP, la cual es d. Nuestra geometría es tal que tenemos un triángulo rectángulo con hipotenusa = segmento $PQ = \vec{u} = (3, -3, 7)$, la proyección $P_{\vec{D}}(\vec{u})$, y el segmento MP cuyo largo es d.

Por tanto:

$d = \sqrt{\|(3, -3, 7)\|^2 - (39/\sqrt{77})^2} = \sqrt{67 - \frac{39^2}{77}} \approx 6{,}87.$

215. $\boxed{\textbf{Ejercicio}}$ *Encuentre la distancia del punto $\vec{Q}$ a la línea cuyo vector director $\vec{D}$ y punto $\vec{P}$ son:*

a) $\vec{Q}(1, -1, 5)$; $\vec{D}(1, 2, 0)$, $\vec{P}(-1, 2, 1)$.

b) $\vec{Q}(-1, -2, 4)$; $\vec{D}(-2, -1, 0)$, $\vec{P}(2, 1, 3)$.

c) $\vec{Q}(5, -1, 3)$; $\vec{D}(2, 0, -5)$, $\vec{P}(1, 3, 5)$.

d) $\vec{Q}(-3, -2, 0)$; $\vec{D}(-2, -1, 0)$, $\vec{P}(-2, 1, 0)$.

e) $\vec{Q}(-1, 3, 2)$; $\vec{D}(2, 3, 5)$, $\vec{P}(1, -4, 0)$.

216. **Ejemplo** *Desarrollemos un método para encontrar la distancia entre dos planos paralelos. Encontramos dos puntos, uno en cada plano. Todo lo que tenemos que hacer es proyectar el segmento que los une sobre el vector normal. la norma de la proyección es la distancia requerida, la cual es simplemente la componente en valor absoluto.*

217. **Ejemplo** *Encontremos la distancia entre los planos $\Pi_1 : x + 2y + 3z = 1$ y $\Pi_2 : x + 2y + 3z = 2$.*

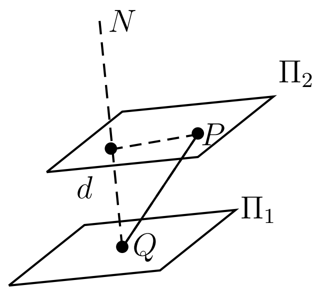


Figura 2.36. Con ayuda de un triángulo rectángulo y componentes, podemos hallar la distancia entre dos planos paralelos.

Si los planos se cortan, la distancia entre ellos es cero. Si hay un vector que sea normal a ambos planos, ellos son paralelos y su distancia se halla como sigue:

Fabricamos un punto P en un plano y otro Q en otro plano. El segmento que los une, PQ, se proyecta sobre el vector normal a los dos planos y el valor absoluto de la componente resultante es la distancia entre ellos.

En nuestro caso, hay un vector normal a ambos planos y es $(1,2,3)$. El punto $P = (4,0,-1)$ pertenece al primer plano, mientras que $Q = (5,0,-1)$ pertenece al segundo plano. El segmento PQ es $(5,0,-1) - (4,0,-1) = (1,0,0)$. Calculamos la componente de PQ sobre el vector normal $\vec{N} = (1,2,3)$:

$\text{Comp}_{\vec{N}}(\vec{u}) = (\vec{u} \cdot \vec{N})/\sqrt{||\vec{N}||}$

$\text{Comp}_{(1,2,3)}(1,0,0) = ((1,0,0) \cdot (1,2,3))/\sqrt{||(1,2,3)||}$

$= 1/\sqrt{14} = \sqrt{14}/14.$

La distancia d entre los dos planos es el valor absoluto de la componente, lo cual es lo mismo.

218. **_Ejercicio_** *Encuentre la distancia entre los siguientes pares de planos:*

a) $3x - 5y + 2z = 1$, $3x - 5y + 2z = 3$
b) $-x + 3y + 2z = 3$, $-x + 3y + 2z = 7$
c) $x - 3y - 2z = 2$, $x - 3y - 2z = 3$
d) $-2x + 4y - 2z = 4$, $-4x + 8y - 4z = 2$
f) $x - 3y + 4z = 1$, $3x - 9y + 12z = 0$.

219. **_Ejemplo_** *Veamos un método para encontrar la distancia entre dos líneas que no se intersecan.*

Este es un ejercicio desafiante. Sin embargo, podemos aplicar un principio *TRIZ* (Altshuller, 2000) que dice: una dimensión más, una salvación más. Esto significa que si el mundo lo abruma a uno con su complejidad, entonces uno se ayuda con otra parte o faceta del mundo. Cuando sea preciso, use o invente herramientas y andamios cuando lo necesite. Miremos cómo se implementa este principio.

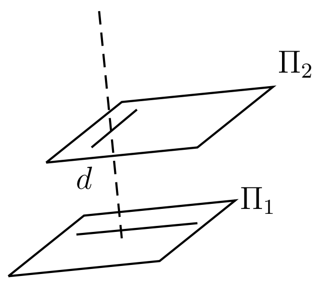


Figura 2.37. Con ayuda de dos planos paralelos podemos hallar la distancia entre dos líneas que nunca se cortan.

Podemos construir dos planos paralelos, tal que uno contenga la primera línea recta y el otro la segunda. Juegue con dos lápices que simulen los vectores directores de las líneas. Deslice un lápiz en forma paralelo con respecto a sí mismo pero siguiendo la dirección del otro lápiz: de ese modo, generará un plano. Deslizando el otro lápiz de igual forma, generará otro plano. Los dos planos son paralelos y ambos tienen el mismo vector normal. Para más señas, la tríada formada por los dos vectores directores y el vector normal, son vectores mutuamente perpendiculares. Como tenemos dos planos paralelos que contienen a las líneas, la distancia entre ellas es la misma que la distancia entre los planos. Y eso es todo.

220. ***Ejemplo y ejercicio*** *Encontremos la distancia entre las líneas en el espacio cuyos vectores directores y puntos son respectivamente* $\vec{D}_1 = (1, 1, -1)$, $P_1 = (4, 0, -1)$ *y* $\vec{D}_2 = (5, -4, 1)$, $P_2 = (5, 0, -1)$.

Necesitamos construir el vector normal de los planos paralelos que contienen a las líneas. El vector normal debe ser perpendicular a ambos vectores directores. Para ello podemos utilizar el producto cruz (inmediato) o el producto punto. Utilicemos el segundo. Sea $\vec{N} = (u, v, w)$ un vector normal. Tenemos: $\vec{N} \cdot \vec{D}_1 = 0$ y $\vec{N} \cdot \vec{D}_2 = 0$, i.e.

$(u, v, w) \cdot (1, 1, -1) = u + v - w = 0$

$(u, v, w) \cdot (5, -4, 1) = 5u - 4v + w = 0$

Sumemos estas dos ecuaciones:

$6u - 3v = 0$.

Puesto que tenemos 3 incógnitas y 2 ecuaciones, debemos tener por lo menos un grado de libertad, i.e. una variable que puede ser fijada como quiera. Sea u tal incógnita y fijemos $u = 1$, entonces de $6u - 3v = 0$ tenemos $v = 2$, y de $u + v - w = 0$ tenemos $w = 3$.

Por tanto, el vector normal a ambos planos es $\vec{N} = (1, 2, 3)$. La ecuación del plano refleja el hecho de que el segmento desde un punto no específico del plano (x, y, z) a un punto fijo del plano es perpendicular al vector normal, por lo tanto su producto punto debe ser cero.

La ecuación del primer plano es $(x-4)+2(y-0)+3(z+1) = 0$ ó $x+2y+3z = 1$. La ecuación del segundo plano es $(x-5)+2(y-0)+3(z+1) = 0$ ó $x+2y+3z = 2$.

La distancia entre las dos líneas es precisamente la distancia entre los dos planos, la cual fue hallada en el ejemplo 217. Observemos que al final todo se reduce a estudiar un triángulo rectángulo determinado por tres segmentos: el primero es el determinado por un punto sobre una línea y otro sobre la otra línea; El segundo es el segmento sobre el vector normal que va de línea a línea, y el tercero es un segmento sobre una de las líneas que completa el triángulo. El problema era saber por qué esto funciona. Y esto ya se entendió.

221. $\boxed{\textbf{\textit{Ejercicio}}}$ *Encuentre la distancia entre los siguientes pares de líneas. Tenga cuidado pues hay un par de líneas que son paralelas y no hemos visto antes nada igual. Damos el vector director y el punto de la primera línea y después lo mismo pero para la segunda:*

a) $(1,1,-1),(0,0,1/3);(2,1,-4/3),(0,0,2/3)$.
b) $(0,-1,2),(-1,-2,4);(-2,-1,0),(2,1,3)$.
c) $(4,0,-19),(7,-1,4);(2,0,-5),(1,3,5)$.
d) $(1,1,2),(-3,-2,0);(-2,-1,0),(-2,1,0)$.
e) $(1,-1,-1),(1,-4,0);(2,3,5),(-1,3,2)$.

222. **_Ejercicio_** *Halle la ecuación paramétrica de la recta intersección entre los planos* $\Pi_1 : 2x - y + z = 0$ *y* $\Pi_2 : x + 3y - z = 2$.

223. **_Ejercicio_** *El vector* $(6,2,1)$ *en unidades apropiadas denota la cantidad de harina, chocolate y huevos gastados en hacer un brownie. Calcule el inventario gastado en* n *brownies.*

224. **_Ejercicio_** *El vector* $(4,2,1)$ *denota la cantidad de dinero gastado en harina, chocolate y huevos para hacer un brownie. Calcule el inventario gastado en* n *brownies sabiendo que se necesita una infraestructura de* $(100,50,40)$ *que indica lo que se gasta en equipo para amasar la harina, para mezclarla con el chocolate y para refrigerar los huevos.*

225. **_Ejercicio_** *El vector* $(6,2,1)$ *en unidades apropiadas denota la cantidad de glúcidos, lípidos y proteínas gastados por hora por una persona en el gimnasio. Calcule el inventario gastado en* n *horas. Reduzca el metabolismo 20 veces, cuando está dormido, y renueve el inventario.*

226. **_Dualidad:_** *Las ecuaciones paramétricas de una línea contienen dos tipos de información, una de posición y otra de velocidad. El primero es el punto de vista de los árboles de la carretera. El segundo es el punto de vista de un carro que recorre la carretera. Todo se puede entender si consideramos las dos ecuaciones siguientes:*

$\vec{X} = (3,8,2) + \lambda(1,2,5)$
$\vec{X} = (3,8,2) + \lambda(2,4,10) = (3,8,2) + 2\lambda(1,2,5)$

La única diferencia entre estas dos expresiones es que la segunda línea tiene un vector director que es el doble de la primera línea. Por lo tanto, las dos líneas son la misma. Pero, !atención¡, si interpretamos λ como el tiempo, vemos inmediatamente que la segunda ecuación describe un carro que se mueve al doble de velocidad que el carro descrito por la primera ecuación.

Obsérvese que estos juegos pueden ser muy complicados. El siguiente sistema de ecuaciones describe una carretera que va en línea recta pero recorrida por un móvil que no va a velocidad constante sino que cada vez se mueve más rápido:

$x = 2\lambda^2$
$y = 3\lambda^2$
$z = 7\lambda^2$

La carretera pasa por el origen en la dirección $(2,3,7)$ pero el móvil está en el origen en el tiempo cero, cuando $\lambda = 1$ el móvil está en $(2,3,7)$. Cuando $\lambda = 2$ el primer móvil va en $(4,6,14)$ y si el tiempo es 10 el móvil está en $(200,300,700)$ recorriendo grandes distancias en tiempos muy pequeños.

227. **Ejemplo** *Hallemos la línea que parte de* $(-3,-6,-12)$ *a las cero horas,* $\lambda=0$*, y que pasa por* $(-1,-3,-5)$ *a las 3 de la tarde (cuando* $\lambda=15$*). Averigüemos luego en qué punto estaremos a la hora 30.*

Solución: La línea está en la dirección de
$\vec{D}=(-1,-3,-5)-(-3,-6,-12)=(-1,-3,-5)+(3,6,12)=(2,3,7)$
y, elaborando el punto de vista de los árboles del camino la ecuación de la recta sería
$\vec{X}=(-3,-6,-12)+\lambda(2,3,7)$.

Cuando $\lambda=15$ esta ecuación produce el punto $(27,39,93)$. Ahora tenemos que regular la velocidad para que una nueva ecuación describa la misma carretera pero con una trayectoria que pase por los puntos especificados a las horas requeridas.

La clave de todo es que si la velocidad es constante, ésta puede codificarse en uno de los posibles vectores directores. Todos los vectores directores son de la forma $\vec{D}=k(2,3,7)=(2k,3k,7k)$ y la ecuación buscada es de la forma
$\vec{X}=(-3,-6,-12)+\lambda(2k,3k,7k)$.

Cuando $\lambda=0$ este carro está en $(-3,-6,-12)$, tal como se necesita. Cuando $\lambda=15$ debemos estar en $(-1,-3,-5)$:
$(-1,-3,-5)=(-3,-6,-12)+15(2k,3k,7k)=(-3,-6,-12)+(30k,45k,105k)$
$(-1,-3,-5)=(30k-3,45k-6,105k-12)$
Despejando k de la primera coordenada
$k=(-1+3)/30=2/30=1/15$
Despejando k de la segunda coordenada
$k=(-3+6)/45=3/45=1/15$
Despejando k de la tercera coordenada
$k=(-5+12)105=7/105=1/15$ (¿Acaso esto pudo haberse hallado más simplemente?)

Como todos los resultados son iguales, vamos bien. La ecuación de la trayectoria solicitada es
$\vec{X}=(-3,-6,-12)+\lambda(2k,3k,7k)=(-3,-6,-12)+(\lambda/15)(2,3,7)$.
Si $\lambda=30$ estamos en la posición
$\vec{X}=(-3,-6,-12)+(\lambda/15)(2,3,7)=(-3,-6,-12)+(30/15)(2,3,7)$
$=(-3,-6,-12)+(2)(2,3,7)=(-3,-6,-12)+(4,6,14)=(1,0,2)$

228. **Ejercicio** *Encuentre la parametrización de la línea que está en* P *a la hora cero,* $\lambda=0$*, y pasa a través de* Q *después de* t *horas; encuentre la posición a la hora* s*. Los datos están dados en el orden* P,Q,t *y* s*.*

a) $(1,2,3),(4,5,6),7,14$.

b) $(-1,0,2),(3,5,7),1,6$.

c) $(2,-1,3),(-1,0,1),1,6$.

d) $(-1,3,-1),(0,1,4),-1,2$.

e) $(0,1,1),(-1,2,0),2,-2$.

2.14. Sistemas 3×3

Estamos interesados en entender lo que significa geométricamente la solución a un sistema de ecuaciones. Muy importante es la capacidad que tiene el análisis geométrico para medir la cantidad de soluciones que un sistema de ecuaciones debe tener. Ya analizamos los sistemas 2×2 y pudimos ver, por ejemplo, que uno puede estar seguro de que un sistema tiene una única solución cuando representa un par de líneas no paralelas. Esa certeza nos permite usar cualquier método o el de Gauss-Jordan, para hallarla. Ahora analizaremos los sistemas 3×3, es decir, un sistema de 3 ecuaciones con 3 incógnitas. El lector debe intuir que la forma de razonar en 2 y 3 dimensiones se generaliza inmediatamente a 8, 20 ó 1.000 dimensiones.

Si tenemos un sistema lineal 3×3, lo que tenemos en cada ecuación es una expresión de la forma $ax + by + cz = d$, es decir, tenemos un plano, ver ejemplo 186. Por lo tanto un sistema 3×3 nos representa un sistema de 3 planos al cual hay que hallarle los puntos comunes o de intersección. Hay varios casos, al igual que con los sistemas 2×2:

1. Los planos se intersecan en un único punto, la solución es única.

2. Los planos se intersecan en una línea, como las hojas de un cuaderno, la solución no es única y se tiene 1 grado de libertad, es decir, se tiene una cantidad infinita de soluciones con un parámetro libre.

3. Los tres planos coinciden, son un mismo plano repetido 3 veces, y se tiene una cantidad infinita de soluciones con 2 grados de libertad para indicar que se puede ir a la derecha o a la izquierda y arriba o abajo en el espacio de soluciones.

4. Los planos son paralelos y son diferentes no teniendo ningún punto en común: el conjunto solución es el conjunto vacío, no hay solución, si se trata de solucionar el sistema, lo que se encuentra es una contradicción.

Veamos ejemplos concretos de cada caso:

1) Los planos se cortan en un único punto. Consideremos el sistema

$$\begin{cases} x = 0 \\ y = 0 \\ z = 0 \end{cases}$$

La solución es $(0, 0, 0)$. Veamos de qué manera tenemos 3 planos que se cortan en un único punto.

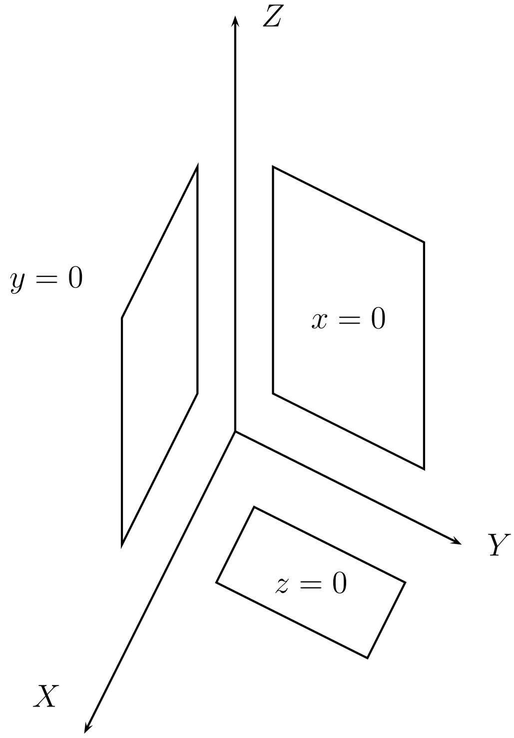


Figura 2.38. Tres planos que se cortan en un único punto $(0, 0, 0)$.

La ecuación $z = 0$ en el espacio nos representa el plano que pasa por (0,0,0) y cuyo vector normal es $(0, 0, 1)$, esto es, el vector que apunta en dirección del eje $\vec{Z}$. Por lo tanto, esta ecuación nos representa el piso o plano XY. La ecuación $x = 0$ en el espacio nos representa el plano que pasa por $(0, 0, 0)$ y cuyo vector normal es $(1, 0, 0)$, o sea el vector que apunta en dirección del eje $\vec{X}$. Por consiguiente, esta ecuación nos representa la pared lateral derecha (uno ve el mundo desde el origen). La ecuación $y = 0$ en el espacio nos representa el plano que pasa por $(0, 0, 0)$ y cuyo vector normal es $(0, 1, 0)$, es decir, el vector que apunta en dirección del eje $\vec{Y}$. Por esta razón, esta ecuación nos representa la pared lateral izquierda. La dos paredes laterales y el piso se intersecan en un único punto $(0, 0, 0)$. Obsérvese que si ponemos el sistema en forma matricial queda como

$$\begin{pmatrix} 1 & 0 & 0 & \vdots & 0 \\ 0 & 1 & 0 & \vdots & 0 \\ 0 & 0 & 1 & \vdots & 0 \end{pmatrix}$$

De igual manera, si al reducir por Gauss-Jordan un sistema de ecuaciones queda la matriz identidad, la solución es única y puede leerse directamente. Recíprocamente, si al reducir una matriz 3×3 queda la matriz identidad, la matriz original representaba un conjunto de planos que se cortan en un único punto.

2) Los planos se intersecan en una línea, como las hojas de un cuaderno, como en el sistema

$$\begin{cases} x = 0 \\ y = 0 \\ x - y = 0 \end{cases}$$

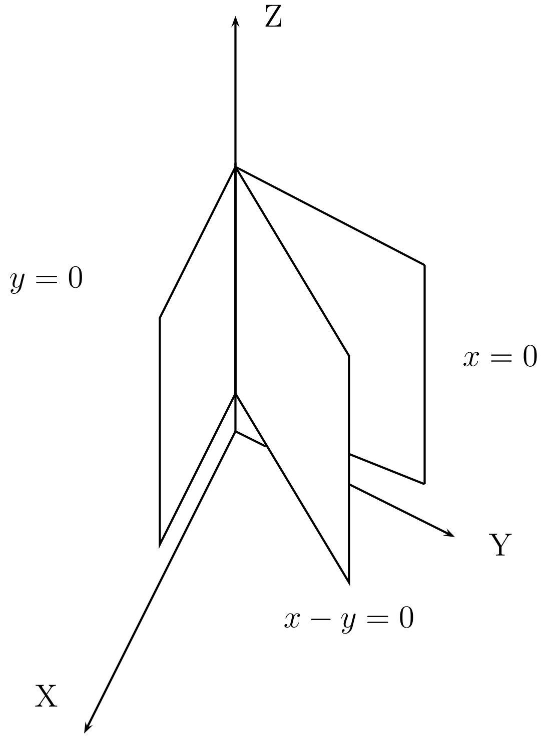


Figura 2.39. Cuaderno, los tres planos tienen en común el lomo del cuaderno.

Notemos que efectivamente estamos en el caso de tres planos que se organizan como las hojas de un cuaderno, $x = 0$ representa una pared lateral que contiene al eje $\vec{Z}$, $y = 0$ es otra pared lateral que también contiene el eje $\vec{Z}$. Por otro lado, la ecuación $x = y$ es lo mismo que $x - y = 0$ que es un plano con vector normal $(1, -1, 0)$. Un punto está en dicho plano si es de la forma

$(x, x, z) = (x, x, 0) + (0, 0, z) = x(1, 1, 0) + z(0, 0, 1)$.

Vemos que cuando $x = 0$ nos queda el eje $\vec{Z}$. Por lo tanto, $x = y$ representa un plano que contiene el eje $\vec{Z}$ y que pasa por la diagonal principal del piso. Todos los planos contienen al eje $\vec{Z}$ y se organizan como un libro. Su solución es una línea, el eje $\vec{Z}$.

Aunque ya conocemos la solución, hallémosla por métodos algebraicos. Comencemos notando que la tercera ecuación dice que $x = y$, y combinando con las dos primeras ecuaciones queda $x = y = 0$. La solución está formada por el conjunto de puntos (x, y, z) de la forma $(0, 0, z)$ pues las dos primeras variables valen cero, pero la tercera no tiene restricción. La solución es el eje $\vec{Z}$ que es una línea.

También podemos proceder por Gauss-Jordan. La matriz del sistema dado puede reducirse hasta

$$\begin{pmatrix} 1 & 0 & 0 & \vdots & 0 \\ 0 & 1 & 0 & \vdots & 0 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

es decir, que la tercera ecuación nos da que $0 = 0$, que puede ignorarse y nos queda un sistema con 3 incógnitas y dos restricciones, el sistema tiene infinitas soluciones con un grado de libertad, en otras palabras, es una línea. La interpretación es la siguiente: si al reducir una matriz 3×3 queda convertida en una matriz con sólo 2 filas o renglones, tenemos el caso de 3 planos que se organizan como las hojas de un cuaderno y viceversa.

3) Los tres planos coinciden, son un mismo plano repetido 3 veces. Ejemplo:

$$\begin{cases} z = 0 \\ 2z = 0 \\ 3z = 0 \end{cases}$$

las dos últimas ecuaciones se simplifican y dan la primera. Esto es, las tres ecuaciones son realmente una misma ecuación. Por lo tanto la solución al sistema es simplemente el conjunto de puntos (x, y, z) cuya tercera coordenada es cero, o sea de la forma

$(x, y, 0) = (x, 0, 0) + (0, y, 0) = x(1, 0, 0) + y(0, 1, 0)$.

Todos estos puntos forman el plano que corresponde al piso, el cual tiene dos grados de libertad, uno en dirección del eje X y otro en dirección del eje Y.

De lo dicho queda claro que al reducir el sistema por Gauss-Jordan nos quedará sólo una ecuación, lo cual produce un problema con 3 incógnitas y una restricción: hay dos incógnitas sin especificar, hay dos grados de libertad, la solución es un plano.

4) Los planos son paralelos y son diferentes nsin tener ningún punto en común.

Ejemplo:

$\Pi_3 : z = 3$
$\Pi_2 : z = 0$
$\Pi_1 : z = -3$

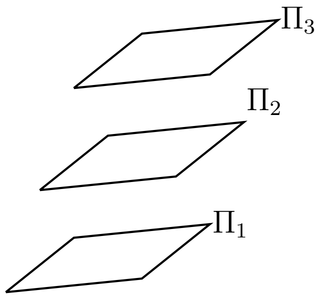


Figura 2.40. Planos paralelos que representan una solución vacía.

Estas ecuaciones representan 3 planos, el piso cero, el piso 3 y el piso -3. Dichos pisos no tienen ningún punto en común y la solución es vacía. Obsérvese que al reducir por Gauss-Jordan un sistema como el dado se llega a una ecuación de la forma $0 = 1$, lo cual es una contradicción.

Analicemos ahora un caso más complicado:

229. *Ejemplo* *Interpretemos la solución al siguiente sistema en términos geométricos:*

$$\begin{cases} 2x - 3y + 4z = 2 \\ -x - y + z = 5 \\ x - 4y + 5z = 7 \end{cases}$$

Lo que queremos hacer es un análisis geométrico que nos dé certeza sobre la cantidad de soluciones que el sistema tiene. Si la solución es única, ésta representa un punto en el espacio, si tiene muchas soluciones estas representarán, o bien, una linea en el espacio, o bien, un plano en el espacio y si no hay solución será porque los planos son paralelos entre sí. Una vez sepamos eso, podemos hallar la solución por nuestro método preferido, Gauss-Jordan.

Veamos cómo podemos utilizar los vectores normales. El plano $2x - 3y + 4z = 2$ tiene como vector normal $\vec{N}_1 = (2, -3, 4)$. El plano $-x - y + z = 5$ tiene como vector normal a $\vec{N}_2 = (-1, -1, 1)$, en tanto que el vector normal de $x - 4y + 5z = 7$ es $\vec{N}_3 = (1, -4, 5)$. Ahora bien, $\vec{N}_1 + \vec{N}_2 = \vec{N}_3$, lo cual significa que $\vec{N}_3$ está sobre el plano Π generado por $\vec{N}_1$ y $\vec{N}_2$. Eso implica que los tres planos son perpendiculares al plano Π. Tal vez los tres planos estén organizados como las hojas de un cuaderno y la solución esté representada por una línea. O tal vez los planos no tengan nada en común entre los tres. En conclusión, o la solución es una línea o es vacía. Pero no puede ser un plano pues los vectores normales apuntan cada uno en una dirección distinta.

Ahora veamos qué da el método de Gauss-Jordan.

Observemos que el sistema fue construido ex profeso para que la tercera ecuación fuese la suma de la primera con la segunda. Por consiguiente al reducir por Gauss-Jordan, la tercera fila desaparecerá. En efecto, el sistema reescrito en forma matricial queda

$$\begin{pmatrix} 2 & -3 & 4 & \vdots & 2 \\ -1 & -1 & 1 & \vdots & 5 \\ 1 & -4 & 5 & \vdots & 7 \end{pmatrix}$$

que al escalonar por renglones se obtiene

$$\begin{array}{r} \\ \\ R1+R2-R3 \rightarrow \end{array} \begin{pmatrix} 2 & -3 & 4 & \vdots & 2 \\ -1 & -1 & 1 & \vdots & 5 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

$$\begin{array}{r} R1-3R2 \rightarrow \\ R1+2R2 \rightarrow \\ \\ \end{array} \begin{pmatrix} 5 & 0 & 1 & \vdots & -13 \\ 0 & -5 & 6 & \vdots & 12 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

Al traducir esta matriz aumentada a un sistema de ecuaciones vemos que tenemos 2 ecuaciones con 3 incógnitas, es decir, 2 planos, los cuales se intersecan en una línea, puesto que sus vectores directores no son paralelos. Ya que la intersección es una línea, hay infinitas soluciones con un grado de libertad. Verifiquémoslo.

Reduciendo la matriz escalonada obtenemos

$$\begin{pmatrix} 1 & 0 & 1/5 & \vdots & -13/5 \\ 0 & 1 & -6/5 & \vdots & -12/5 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

Esto dice que $x + (1/5)z = -13/5$ y que $y - (6/5)z = -12/5$ y que $0z = 0$, o sea, z puede ser cualquiera. Por lo tanto,

$x = -(1/5)z - 13/5$

$y = (6/5)z - 12/5$

$z = z$

esto es

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -13/5 \\ -12/5 \\ 0 \end{pmatrix} + \begin{pmatrix} -1/5 \\ 6/5 \\ 1 \end{pmatrix} z$$

que es una línea con parámetro z. Su vector director es $(-1/5, 6/5, 1)$ y cuando $z = 0$ pasa por el punto $(-13/5, -12/5, 0)$.

Si hubiese alguna duda sobre la fidelidad de la solución, podríamos verificarla por reemplazo. Una vez hecho, concluimos que no hay soluciones espurias. Y nuestro análisis previo sobre los vectores normales nos dice que no faltan soluciones.

Es muy elocuente la forma como el método Gauss-Jordan nos va diciendo cómo son las soluciones, pues nos da una geometría fácil de interpretar y nos produce todas las soluciones, sin faltar y sin añadir ninguna solución espuria. Eso será digno de ser probado en el caso general, y lo haremos en el capítulo de la inversa.

230. $\boxed{\textbf{\textit{Ejercicio}}}$ *Resuelva e interprete geométricamente la solución a los sistemas siguientes:*

$$\text{a)}\ \begin{cases} x - y + 4z = 2 \\ -x + y + z = 5 \\ 5z = 7 \end{cases}$$

$$\text{b)}\ \begin{cases} -3x + 5z = 7 \\ -2x - 2y + 4z = 2 \\ -x + 2y + z = 5 \end{cases}$$

$$\text{c)}\ \begin{cases} 3x - y + 4z = 2 \\ 2x - 4y + 5z = 7 \\ -x - 3y + z = 5 \end{cases}$$

231. $\boxed{\textbf{\textit{Miscelánea}}}$ *Resuelva los siguientes sistemas e interprete geométricamente sus soluciones:*

a) $3x + 4z = 1$, $6x + 8z = 2$ en $\mathbb{R}^2$ y en $\mathbb{R}^3$.

b) $3x + 4z = 1$, $6x + 8z = 2{,}00000001$ en $\mathbb{R}^2$ y en $\mathbb{R}^3$.

c) $3x + 4z = 1$, $6x + 8z = 2$, $9x + 12z = 3$ en $\mathbb{R}^2$ y en $\mathbb{R}^3$.

d) $x + y + z = 0$, $x + y + z = 1$, $x + y + z = 3$ en $\mathbb{R}^3$.

e) $x = 0$, $y = 0$, $x - 3y + 4z = 3$ en $\mathbb{R}^3$.

f) $x = 0$, $y = 0$, $x - 3y + 4z = 3$ en $\mathbb{R}^4$.

g) $x + y + z = 1$, $x - 2y + 3z = 1$, $2x - y + 4z = 2$ en $\mathbb{R}^3$.

h) $x + y + z = 1$, $x - y + z = 1$, $x + y - z = 1$, $3x + y + z = 3$ en $\mathbb{R}^3$.

232. $\boxed{\textbf{\textit{Un buen ejercicio}}}$ *Un sistema de ecuaciones es reducido por Gauss-Jordan para saber el número de ecuaciones independientes. Por tanto, Gauss-Jordan es un procedimiento para limpiar redundancias. Supongamos que tenemos un sistema de ecuaciones con n incógnitas y que después de reducirlo nos quedan i ecuaciones independientes. Establezca un ecuación que relacione el número de incógnitas, m; el número de ecuaciones independientes, i; y el número de parámetros libres o grados de libertad del conjunto solución f. Por favor, ilustre su conclusión con ejemplos adecuados y añada una explicación geométrica.*

233. $\boxed{\textbf{\textit{Ejercicio}}}$ *Halle un polinomio de tercer grado que pase por los puntos* $(1, 1.5)$, $(1.5, 3)$, $(2, 1)$, $(2.5, 4)$.

2.15. Las estaciones (opcional)

Veamos de qué manera las matemáticas que hemos visto nos dan la capacidad de formular ideas muy sencillas con las que podemos tener una visión bastante profunda sobre la naturaleza de uno de los fenómenos planetarios que más incidencia global tiene sobre la vida en la Tierra: las estaciones. Este tema debe dominarse a tal grado que uno pueda contarlo a los sobrinos y a la abuelita. Explicar esto a la luz de una vela, que simule el Sol, y ayudándose de una naranja o una pelota, que simule la Tierra, resulta impresionante.

234. $\boxed{\textbf{\textit{Ejercicio}}}$ *Proponga algunos efectos de las estaciones sobre las migraciones y contramigraciones periódicas, sobre la variabilidad de la vida, sobre la estacionalidad de los fenómenos biológicos. Dé ejemplos con aves, mamíferos, insectos, peces, plantas silvestres y de cultivo. Tenga en cuenta que la estacionalidad de los fenómenos botánicos arrastra la estacionalidad de los demás. Añada consideraciones sobre la conducta de los humanos y del turismo.*

Comencemos mencionando un contraste entre los buenos ciclistas y los aficionados. A los primeros se les ve detenidos en los semáforos haciendo delicadas maniobras para no caerse. A los aficionados eso no les funciona. En cambio ellos pueden viajar largos trechos en bicicleta sin hacer prácticamente nada. ¿Cómo pueden ellos hacer eso si solamente son aficionados? La razón es que hay una ley de la naturaleza que los favorece. En efecto, cuando la bicicleta se mueve, las ruedas giran y lo hacen sobre un plano vertical a la Tierra y en la dirección que uno lleve.

Puede decirse que, en cierto sentido, este sistema es estable: se necesita hacer algo para desequilibrarlo. Esto se logra si uno se desbalancea un poquito hacia algún lado y se crea lo que se llama un torque. Dicha 'estabilidad' del plano de giro de la rueda es lo que oficialmente se denomina como conservación del momento angular.

Ahora bien, en vez de decir que el plano de giro se conserva, lo que se puede decir es que el eje de giro se conserva. El eje de giro es ni más ni menos el vector normal al plano de giro que es el mismo plano que contiene la rueda. Este cambio de punto de vista nos permite alargar o acortar el vector normal para dar a entender que la 'estabilidad' del plano de giro puede ser mayor o menor: es más estable entre más veloz se vaya. El nombre técnico de dicho vector es **vector de momento angular**.

Los trompos y las pirinolas se mantienen de pie mientras giran porque la interacción con el piso no alcanza a interferir notablemente con la rotación y por consiguiente el momento angular se conserva, el cual va en dirección vertical. Los trompos que no son perfectamente simétricos pueden balancearse mientras giran: a este movimiento se llama precesión.

Pensemos ahora en la Tierra: ella gira al rededor del Sol en una órbita elíptica. Pero al mismo tiempo, gira sobre su propio eje. Girando sobre su propio eje se origina el día y la noche. La Tierra gira a muy alta velocidad, nosotros recorremos 40000 km cada día. Y además la Tierra pesa mucho. La consecuencia es que el eje de giro de la Tierra es muy estable. Por supuesto que no es inconmovible, pero es muy estable.

Resumimos todo diciendo que la dirección y la magnitud del vector momento angular se conservan mientras la Tierra recorre su órbita.

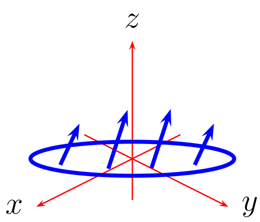


Figura 2.41. Conservación del momento angular.

¿Cómo podemos saber la dirección del eje de giro de la Tierra? Veamos de qué manera se puede medir muy sencillamente y así explicar un hecho simple pero espectacular: si una pared, o una casa, orientada de oriente a occidente es iluminada por el Sol por el sur, con toda seguridad el Sol la iluminará a los seis meses por el norte. Y viceversa. En Bogotá, una pared orientada de oriente a occidente es iluminada en noviembre por el sur y en mayo por el norte.

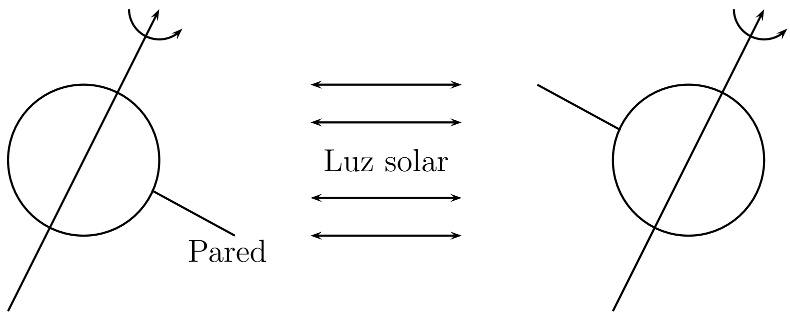


Figura 2.42. Estaciones y patrón de iluminación.

La clara razón es que el eje de la Tierra no es perpendicular al plano orbital: el eje de la Tierra está inclinado. ¿Cómo podemos medir el ángulo de inclinación del eje de giro de la Tierra con respecto al plano orbital? Primero debemos jugar un poco con una naranja, que representará la Tierra, y clavarle un palillo de dientes que haga las veces de un poste de la luz cercano a la casa que habitamos. Segundo: con el índice de la mano derecha imitamos los rayos solares y en la mano izquierda sostenemos la naranja, la vamos moviendo por su órbita y notamos que hay una posición (y otra al otro lado de la órbita) en la cual a medio día, el Sol ilumina perpendicularmente la Tierra. En ese momento el rayo de Sol (el índice) cae exactamente en la dirección del poste de la luz (el palillo de dientes sobre la naranja). Si las cosas no funcionan es porque al mover la naranja no se observó la conservación del momento angular. Lo mejor es clavarle un lápiz a la naranja para representar su eje de giro y para tener en claro que al moverla la dirección del lápiz no debe cambiar.

Imaginemos un poste de luz sobre el Ecuador. El ángulo entre el poste de la luz y el rayo del Sol varía con la época, a veces es cero, el Sol cae perpendicularmente, pero a veces es mucho. El ángulo máximo de inclinación del Sol con respecto al poste corresponde a una situación en la cual el poste de la luz, el eje de giro y el Sol están sobre el mismo plano.

En ese instante puede medirse el ángulo, a medio día, entre el poste de la luz y el rayo de Sol, y dicho ángulo es el que está entre el eje de giro de la Tierra y la normal al plano orbital. Este resultado se basa en un hecho geométrico importante: cuando dos ángulos agudos tienen lados perpendiculares, ellos miden exactamente lo mismo.

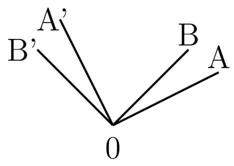


Figura 2.43. Ángulos con lados perpendiculares.

En la gráfica, el ángulo AOB es agudo, lo mismo que $A'OB'$. Además, el lado AO es perpendicular a $A'O$. Y de igual forma, el lado BO es perpendicular al $B'O$. Eso significa que si el ángulo $B'OA'$ mide α, entonces el ángulo $A'OB$ mide $90 - \alpha$. Pero como el ángulo $A'OA$ es recto, entonces el ángulo BOA debe ser el complementario de $A'OB$ y por lo tanto debe medir también α.

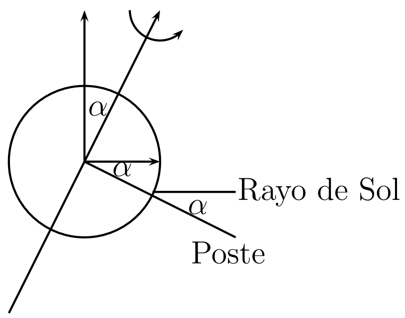


Figura 2.44. La inclinación del eje de giro de la Tierra.

Con esta forma de pensar se ha estimado que el ángulo α entre el eje de giro de la Tierra y la normal al plano orbital es de unos 22 grados. Observemos que si el ángulo α fuese de 0 grados, la mayor parte de Canadá estaría en eterno invierno. En consecuencia, con el ángulo de inclinación del eje de la Tierra se regula la productividad biótica de la Tierra.

235. **_Ejercicio de investigación_** *Investigue si el eje de giro de la Tierra ha cambiado a lo largo de su historia. Si es un movimiento continuo o si ha sido por causa de una hecatombe. Explique cómo se hizo para saberlo, especificando el papel de la geología y de la biología. ¿Qué se espera para el futuro?*

Podemos jugar un poco más con nuestra naranja y si de noche la iluminamos con una linterna podemos observar lo que el polo alejado del Sol está poco o nada iluminado por el Sol, este polo está en invierno. El polo cercano al Sol está siendo fuertemente iluminado, este polo está en verano. Por eso se puede cultivar papa en latitudes exageradas de Siberia, el Sol nunca se oculta y la papa puede crecer y dar fruto en los tres meses que dura el verano.

En realidad, las estaciones no duran cada una tres meses, la órbita es elíptica y el Sol está en uno de sus focos. Cuando la Tierra está más cerca del Sol, ésta recorre su órbita deprisa (para no caerse). Pero cuando la Tierra está alejada, la órbita es recorrida lentamente. Por lo tanto, las estaciones cerca del Sol son cortas. Las estaciones alejadas del Sol, son largas. Midiendo este desbalance podría estimarse la excentricidad de la elipse orbital.

Por otro lado, en el verano los días son largos y en el invierno son cortos. Daría la impresión de que la Tierra girara más despacio en verano y más rápido en invierno. ¿Cómo puede suceder si el momento angular de la Tierra se conserva y ella siempre gira a la misma velocidad angular? Considere la siguiente explicación: debido a que la inclinación del eje de giro de la Tierra no es nula, en casi todos los días hay un hemisferio que no sólo esta más cerca del Sol sino que, además, tiene una área iluminada mayor que la del otro. Es decir, tomando un paralelo a la altura de México en verano, no son 180 los grados iluminados sino que son 200 ó más. Por lo tanto, un punto sobre la superficie al ir girando permanecerá más tiempo a la luz que en la oscuridad. Los días son más largos en verano que en invierno. ¿No es esta una razón para insinuar que los lazos familiares entre los habitantes, de todas las especies, de altas latitudes deberían ser más fuertes que entre sus homólogos del trópico? ¿Será cierto?

Y ahora viene algo preocupante: cuando el Sol se ve al mediodía en enero, la casa “mira hacia la izquierda.”Pero en junio, al mediodía la casa “mira hacia la derecha.”Por lo tanto, hay 12 horas de diferencia entre estos dos mediodías y no 24 como debiera ser. Estas doce horas alargan los días de verano y acortan los de invierno. ¿Qué efecto podrá tener este hecho?

La intrigante sencillez que hemos promovido se ha conseguido a un alto precio pagado a lo largo de muchas generaciones. En particular, la presesión de la tierra fue estudiada cuantitativamente por Hiparco hacia el 125 antes de nuestra era, pero sólo con Newton pudo ofrecerse una explicación del porqué. Aristarco pudo lograr lo que hizo porque se nutrió de una cultura muy desarrollada que ya sabía, por ejemplo, cómo medir la distancia de la Tierra tanto a la Luna como al Sol, algo que fue desarrollado por Aristarco de Samos hacia el año 250 a. J. C. (Gran Enciclopedia Larousse, 1983).

2.16. Ejercicios de repaso

1. Encuentre el ángulo entre los vectores $\vec{u} = (1, 1, 1, 1)$ y $\vec{v} = (1, 0, 0, 0)$.

2. Considere los vectores $\vec{a} = 2\vec{i} + 3\vec{j} - 2\vec{k}$ y $\vec{b} = \vec{i} + \vec{j} + \vec{k}$. Calcule:

   a) $\text{Comp}_{\vec{b}}\vec{a}$.

   b) $\text{Proy}_{\vec{b}}\vec{a}$.

3. Encuentre los valores de t para los cuales el espacio solución de $AX = 0$ es una recta que pasa por el origen o un plano que pasa por el origen, o sólo el origen o

todo $\mathbb{R}^3$, siendo $A = \begin{pmatrix} 1 & 1 & t \\ 1 & t & 1 \\ t & 1 & 1 \end{pmatrix}$.

4. Considere los puntos $A(2,-2,1), \vec{B}(-1,0,3)$ y $\vec{C}(5,-3,4)$, en $\mathbb{R}^3$. Halle:

   a) Las ecuaciones simétricas de la recta L que pasa por los puntos A y C.

   b) La ecuación del plano Π que contiene a los puntos A, B y C.

   c) El área del triángulo cuyos vértices son A, B y C.

   d) Distancia del punto $(-2,3,4)$ al plano Π.

   e) Punto de intersección de la recta L y el plano Π.

5. Considere el paralelepípedo generado por los vectores $\vec{u} = (1,4,-1)$, $\vec{v} = (-1,4,0)$ y $\vec{w} = (6,6,8)$ con uno de sus vértices en el punto $A(3,3,0)$.

   a) Halle las coordenadas de los 7 vértices restantes.

   b) Determine las coordenadas del punto medio del paralelepípedo.

   c) Halle la ecuación de la recta que pasa por dos vértices opuestos del paralelepípedo.

   d) Halle la ecuación del plano que contiene a la cara superior del paralelepípedo.

   e) Halle la distancia entre el punto medio del paralelepípedo y una de sus caras.

   f) Halle el área de una de las caras del paralelepípedo.

   g) Determine el volumen del paralelepípedo.

   h) Halle la ecuación del plano que está 5 unidades por encima de la cara superior del paralelepípedo.

   i) Halle la ecuación de la recta que se encuentra en la cara superior del paralelepípedo y que dista 5 unidades de una de las aristas de la cara superior.

6. Encuentre la ecuación de la recta que pase por el punto $(2,3,4)$ y que sea paralela a los dos planos $\left\{ \begin{array}{r} x - y + z = 1 \\ 2x + y - 3z = 2 \end{array} \right.$

7. Hallar el punto Q del plano Π: $2x+2y+z = -1$ más cercano al punto $P(9,7,3)$.

8. Encuentre las ecuaciones paramétricas del plano en $\mathbb{R}^3$ que pasa por los puntos $(1,0,0), (0,1,0)$ y $(0,0,1)$.

9. Encuentre una ecuación lineal simple en tres variables cuyo conjunto solución sea exactamente el plano dado en el inciso a).

10. Sea $H = \{(x,y,z) \colon x + y = 0\}$ y sea $\vec{v} = (1,1,1)$. Hallar:

    a) $H^{\perp}$.

    b) $\text{Proy}_{H^{\perp}}\vec{v}$.

    c) $\text{Proy}_{H}\vec{v}$.

2.17. Resumen

La capacidad de análisis visual de los mamíferos es algo tan maravilloso, tan eficiente, tan poderoso, que la tecnología moderna a duras penas ha podido imitarla. La geometría es la plataforma que convierte ese poder visual en una herramienta muy eficaz para el análisis de sistemas lineales (y no lineales). Hemos desarrollado el tema para dos y tres variables, lo cual nos indujo a considerar rectas y planos. Pudimos constatar que al interpretar un sistema de ecuaciones en términos de planos o rectas, uno puede predecir si la solución es única o si hay infinitas soluciones sobre una recta o si quedan sobre un plano o si no hay solución. De esa manera, uno sabe de antemano la forma de la solución, la cual puede hallarse por cualquier método. Uno quisiera pensar que lo dicho de alguna manera se extrapola para problemas con más de 3 variables. La verdad es que esta intuición necesita ser elaborada y es un objetivo que enfrentaremos en próximos capítulos.

CAPÍTULO 3

EL DETERMINANTE

En el presente capítulo veremos de qué manera uno puede asociar un número llamado **determinante** a una matriz de tal forma que si el determinante es cero, entonces un sistema de ecuaciones generado por la matriz o no tiene solución o tiene infinitas soluciones, y si el determinante es distinto de cero entonces cualquier sistema generado por la matriz tiene solución única.

3.1. Idea fundamental

Consideremos un sistema lineal 2×2, el cual representa dos líneas. Si las líneas no son paralelas, ellas se cortan en un punto y la solución al sistema es única. Si las líneas son paralelas, pero no coinciden, no hay solución. Pero si las líneas son paralelas y coincidentes, entonces hay un número infinito de soluciones.

Consideremos ahora un sistema lineal de ecuaciones 3×3. Cada ecuación representa un plano. Si los planos se cortan en un único punto, la solución es única. Si los planos tienen una línea en común o un plano, el sistema tiene infinitas soluciones. Y si los planos son paralelos y diferentes, entonces no hay ninguna solución.

Existe una manera de tratar de una sola vez por todas no solo ambos casos sino también las generalizaciones correspondientes a altas dimensiones. La idea es como sigue:

Consideremos el sistema

$$\begin{cases} 2x + 3y = 4 \\ 5x - 6y = 7 \end{cases}$$

Uno puede graficar las dos líneas, verificar que no son paralelas y que, por lo tanto, se cortan en un único punto y que por ende la solución al sistema es única. Veamos ahora cómo se puede predecir que la solución es única con el determinante. Tenemos 7 pasos:

Paso 1. El vector $(2, 3)$ es perpendicular a la línea $2x + 3y = 4$, mientras que el vector $(5, -6)$ lo es a la línea $5x - 6y = 7$.

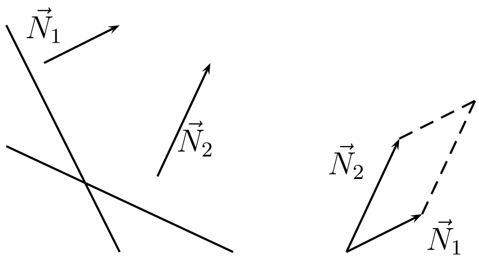


Figura 3.0. A un par de líneas asociamos un paralelogramo determinado por los vectores normales.

Paso 2. Las líneas son paralelas, si lo son los vectores normales y las líneas no son paralelas, si sus vectores normales no lo son. En nuestro ejemplo, los vectores normales no son paralelos, por lo tanto, ya conocemos la respuesta: las líneas se cortan en un único punto y por consiguiente nuestro sistema tiene solución única. Veamos cómo obtenemos la misma respuesta por el determinante.

Paso 3. Para saber si dos vectores son paralelos, formamos con ellos un paralelogramo y calculamos su área, a la cual vamos a llamar determinante: si el determinante es cero, son paralelos, si el determinante no es cero, no son paralelos.

Paso 4. Cada línea tiene una infinidad de vectores normales, pero que el determinante asociado a un sistema 2×2 sea cero o no, es algo que no depende de la escogencia de los vectores normales.

Paso 5. Si el área del paralelogramo determinado por los vectores normales es diferente de cero, entonces el determinante es diferente de cero. En ese caso, el sistema representado por las líneas tiene solución única. Pero si dicha área es cero, las líneas son paralelas y puede ser que nunca se corten, en cuyo caso la solución es vacía; o que las líneas coincidan, y en ese caso hay infinitas soluciones todas sobre la recta dada.

Paso 6. Por lo tanto, si el determinante es diferente de cero, hay solución única. Si el determinante es cero, la solución no es única, puede ser porque no exista solución o puede ser porque haya infinidad de soluciones.

Paso 7. Como todo depende del área del paralelogramo generado por los vectores normales, pues debemos calcular dicha área. Lo haremos en la siguiente sección, pero pondremos la mira en sistemas lineales $n \times n$ cualesquiera. Allá no tendremos área sino volumen generalizado, que también puede llamarse determinante (aunque hay una pequeña diferencia).

3.2. Paralelepípedos = plps

Un **paralelepípedo, plp**, es la generalización a n dimensiones de un paralelogramo. La descripción algebraica de un plp se basa en un sencilla observación:

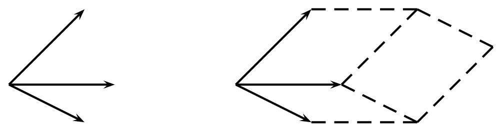


Figura 3.1. Aristas básicas de un plp.

Una cara está determinada por dos vectores pues las aristas opuestas son segmentos paralelos a los vectores dados. Por lo tanto, si se enuncian las aristas o vectores que salen de un vértice cualquiera, quedarán generadas todas las caras que salen de dicho vértice y además, todas las otras caras del plp se construyen por paralelismo. Por lo tanto, un plp está completamente descrito por las aristas que parten de un vértice dado.

236. ♣ *Definición.* *Un **n-plp** P anclado en el origen de $\mathbb{R}^n$ es un conjunto ordenado de n vectores $P = [\vec{u}_1, ..., \vec{u}_n]$, que son las aristas que parten del origen. Ordenamos las aristas diciendo cuál es la primera, la segunda,... Observemos que en esta definición sólo tenemos las aristas. Uno también puede rellenar las caras o rellenar todo el plp. Si queremos rellenar todo el plp, se procede así: $\vec{X} \in P$ relleno ssi $\vec{X} = \alpha_1\vec{u}_1 + ... + \alpha_n\vec{u}_n$ donde los escalares α_i pueden tomar valores entre cero y uno.*

237. ***Ejemplo*** *Un paralelogramo es un 2-plp. El cubo unitario es un 3-plp dado por $[\vec{i}, \vec{j}, \vec{k}]$.*

A los n-plps se les puede calcular el volumen. El 1-volumen de un 1-plp es la norma del único vector que lo forma. El 2-volumen de un 2-plp es su área. El 3-volumen de un 3-plp es su volumen común y corriente. Calculemos el 2-volumen de un 2-plp, es decir, el área de un paralelogramo.

238. ◇ *Teorema.* *Si un 2-plp está dado por $[(a, b), (c, d)]$ su 2-volumen está dado por $ad - bc$ (cuando sea negativo, tómese el valor absoluto).*

Demostración. Inscribimos al 2-plp en el rectángulo que tiene todas sus aristas paralelas a los ejes y que contiene el origen del plp y el vértice opuesto. El área del 2-plp es entonces el área de dicho rectángulo menos el área de todo lo que está por fuera del plp pero por dentro del rectángulo: dos triángulos y dos rectángulos y otros dos triángulos.

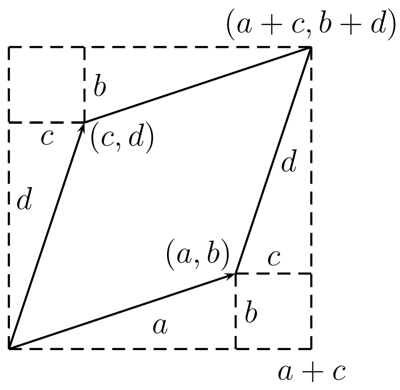


Figura 3.2. Geometría del determinante.

El área es entonces
$A = (a+c)(b+d) - 2[ab/2 + cb + cd/2] = ab + ad + cb + cd - ab - 2bc - dc = ad - bc.$

Observemos que astutamente cuadramos el dibujo para lograr que $A = ad - bc > 0$ y así no tener que preocuparnos por la posibilidad real de obtener una área negativa. Eso se debe a que dibujamos los 2 vectores del 2-plp para que quedaran ordenados como el 2-plp $[\vec{i}, \vec{j}]$: si rotamos 90 grados en dirección contraria a las manecillas del reloj al vector $\vec{i}$, se obtiene el vector $\vec{j}$. Y eso nos garantiza que nuestro procedimiento para calcular el área del 2-plp $[\vec{i}, \vec{j}]$ dé como resultado un número positivo. En general, cada forma de ordenar los vectores de un n-plp se llama una orientación. Todos nuestros $\mathbb{R}^n$ están orientados de la manera natural, la cual es dada por el n-plp unitario, cuyo $j-$ésimo vector tiene un uno en la j-ésima coordenada y cero en las demás. Por ejemplo, la orientación natural de $\mathbb{R}^3$ es la de $[\vec{i}, \vec{j}, \vec{k}] = [(1,0,0), (0,1,0), (0,0,1)]$ que también se llama de la mano izquierda.

239. ***Ejemplo*** *Calculemos el 2-volumen de* $[(1,1),(2,2)]$. *Es* $1 \times 2 - 2 \times 1 = 0$, *lo que significa que dos vectores colineales generan un plp cuyo volumen es cero.*

240. ***Ejercicio*** *Calcule el 2-volumen de los 2-plps dados por*

a) $[(1,1),(2,3)]$
b) $[(1,-1),(2,-3)]$
c) $[(0,1),(1,0)]$
d) $[(1,0),(0,1)]$.

241. ***Ejemplo*** *Usemos la teoría del determinante para saber si el sistema siguiente tiene solución única o no:*

$$\begin{cases} 2x + 3y = 4 \\ 5x - 6y = 7 \end{cases}$$

El vector $(2,3)$ es perpendicular a la línea $2x+3y=4$, en tanto que el vector $(5,-6)$ lo es a la línea $5x-6y=7$. Como $(2,3)$ no es un múltiplo escalar de $(5,-6)$, dichos vectores no son paralelos. Por lo tanto, forman un paralelogramo cuya área no es cero, por consiguiente, la solución es única. Esa respuesta puede sacarse automáticamente como sigue:

El 2-volumen del paralelogramo generado por los vectores normales es, aplicando $ad-cb$ es, $(2)(-6)-(5)(3)=-12-15=-27$. Como el determinante es diferente de cero, los vectores normales no son paralelos, por lo tanto, las líneas correspondientes tampoco, por lo que ellas se cortan en un único punto, de donde se deduce que el sistema tiene solución única.

242. $\boxed{\textbf{\textit{Ejercicio}}}$ *Use la teoría del determinante para decidir si los siguientes sistemas tienen solución única o no.*

a) $x+y=2$, $2x+3y=5$
b) $x-y=4$, $2x-3y=4$
c) $y=4$, $x=2$
d) $x=4$, $y=-1$.

243. ♣ ***Definición.*** *El* ***determinante*** *de un 2-plp definido por* $[(a,b),(c,d)]$ *es* $ad-bc$ *y se nota como*

$$Det(P)=\begin{vmatrix} a & c \\ b & d \end{vmatrix}=ad-bc.$$

El 2-volumen es igual al valor absoluto del determinante.

Nuestra notación se generaliza naturalmente a n-plps: el 3-volumen de un 3-plp $P=[(a,b,c),(d,e,f),(g,h,i)]$ sería el número notado

$$Det(P)=\begin{vmatrix} a & d & g \\ b & e & h \\ c & f & i \end{vmatrix}$$

Calcular este número por geometría para que represente el volumen de una caja es todo un reto que cada quien puede enfrentar. Pero en realidad eso es bastante más complicado y menos general que el siguiente proceso algebraico.

244. $\diamond$ ***Teorema.*** *El 2-det tiene las siguientes propiedades:*

a) $Det(\vec{u},\vec{u})=0$, un plp generado por un solo vector tiene volumen cero.

b) $Det(\lambda\vec{u},\vec{v})=\lambda Det(\vec{u},\vec{v})$, si se alarga una arista del plp, su volumen se amplifica consecuentemente.

c) $Det(\vec{u},\vec{v})=-Det(\vec{v},\vec{u})$, una permutación de los vectores o aristas cambia el signo. Se dice que el determinante es una función alternada.

d) $Det(\vec{u},\vec{v}+\vec{w})=Det(\vec{u},\vec{v})+Det(\vec{u},\vec{w})$, la suma en cada entrada del determinante se reparte. Teniendo en cuenta esta propiedad y la b, se dice que el determinante es multilineal.

e) $Det(\vec{i},\vec{j}) = 1$, el volumen del 2-plp unitario es 1.

f) El determinante tiene signo más o menos, y para hallar el área se toma el valor absoluto.

245. $\boxed{\textbf{\textit{Ejercicio}}}$ *Ilustre gráficamente el teorema anterior para que parezca obvio.*

246. $\boxed{\textbf{\textit{Ejercicio}}}$ *Explique la siguiente contradicción:*

$$det\begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} = 16 - 24 = -8$$

pero por otro lado:

$$Det\begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} = Det\left(2\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\right) = 2Det\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 2(4-6) = -4.$$

247. ♣ *Definición.* *El determinante es un número asociado a un n-plp y que goza de la generalización natural de todas las propiedades dadas en el teorema 244 para 2D.*

248. ***Ejemplo*** *Calculemos el 3-volumen de* $P = [(a,b,c),(d,e,f),(g,h,i)]$.

Desarrollo. Para utilizar las propiedades del determinante, reescribimos el plp P en notación vectorial y no de coordenadas:

$P = [(a\vec{i} + b\vec{j} + c\vec{k}), (d\vec{i} + e\vec{j} + f\vec{k}), (g\vec{i} + h\vec{j} + i\vec{k})]$.

Tenemos

$\text{Det}(P) = \text{Det}(a\vec{i} + b\vec{j} + c\vec{k}, d\vec{i} + e\vec{j} + f\vec{k}, g\vec{i} + h\vec{j} + i\vec{k})$

Ahora repartimos las sumas de la primera entrada:

$\text{Det}(P) = \text{Det}(a\vec{i}, d\vec{i} + e\vec{j} + f\vec{k}, g\vec{i} + h\vec{j} + i\vec{k}) + \text{Det}(b\vec{j}, d\vec{i} + e\vec{j} + f\vec{k}, g\vec{i} + h\vec{j} + i\vec{k}) +$
$+\text{Det}(c\vec{k}, d\vec{i} + e\vec{j} + f\vec{k}, g\vec{i} + h\vec{j} + i\vec{k})$.

Repartamos las sumas de la segunda entrada en el primer renglón de la expresión anterior:

$\text{Det}(a\vec{i}, d\vec{i}+e\vec{j}+f\vec{k}, g\vec{i}+h\vec{j}+i\vec{k}) = \text{Det}(a\vec{i}, d\vec{i}, g\vec{i}+h\vec{j}+i\vec{k}) + \text{Det}(a\vec{i}, e\vec{j}, g\vec{i}+h\vec{j}+i\vec{k}) +$
$+\text{Det}(a\vec{i}, f\vec{k}, g\vec{i} + h\vec{j} + i\vec{k})$.

Repartimos las sumas de la tercera entrada del primer renglón:

$Det(a\vec{i}, d\vec{i}, g\vec{i} + h\vec{j} + i\vec{k}) = \text{Det}(a\vec{i}, d\vec{i}, g\vec{i}) + \text{Det}(a\vec{i}, d\vec{i}, h\vec{j}) + \text{Det}(a\vec{i}, d\vec{i}, i\vec{k})$.

Sacamos las constantes:

$\text{Det}(a\vec{i}, d\vec{i}, g\vec{i} + h\vec{j} + i\vec{k}) = adg\text{Det}(\vec{i},\vec{i},\vec{i}) + adh\text{Det}(\vec{i},\vec{i},\vec{j}) + adi\text{Det}(\vec{i},\vec{i},\vec{k})$.

Observemos que el volumen de un 3-plp apachurrado (que cabe en un plano) es cero. Por esto los tres determinantes anteriores dan cero:

$Det(\vec{i},\vec{i},\vec{i}) = Det(\vec{i},\vec{i},\vec{j}) = Det(\vec{i},\vec{i},\vec{k}) = 0$

Reflexionemos ahora un poco, en un determinante casi todo da cero. ¿Qué es lo que no da cero? Pues sólo aquellos casos en los cuales llegamos a un plp elemental (un cubo permutado) no apachurrado. Por tanto,

$\text{Det}(P) = Det(a\vec{i} + b\vec{j} + c\vec{k}, d\vec{i} + e\vec{j} + f\vec{k}, g\vec{i} + h\vec{j} + i\vec{k})$
$=\text{Det}(a\vec{i}, e\vec{j}, i\vec{k})+ \text{Det}(a\vec{i}, f\vec{k}, h\vec{j})+ \text{Det}(b\vec{j}, d\vec{i}, i\vec{k})+ \text{Det}(b\vec{j}, f\vec{k}, g\vec{i})+$
$+\text{Det}(c\vec{k}, d\vec{i}, h\vec{j})+ \text{Det}(c\vec{k}, e\vec{j}, g\vec{i})$.

Sacamos las constantes:
$= aei\text{Det}(\vec{i}, \vec{j}, \vec{k})+ afh\text{Det}(\vec{i}, \vec{k}, \vec{j})+ bdi\text{Det}(\vec{j}, \vec{i}, \vec{k})+ bfg\text{Det}(\vec{j}, \vec{k}, \vec{i})+$
$+cdh\text{Det}(\vec{k}, \vec{i}, \vec{j})+ ceg\text{Det}(\vec{k}, \vec{j}, \vec{i})$.

Permutemos los vectores para convertirlos en un cubo unitario con determinante positivo:
$aei\text{Det}(\vec{i}, \vec{j}, \vec{k}) = aei$
$afh\text{Det}(\vec{i}, \vec{k}, \vec{j}) = -afhDet(\vec{i}, \vec{j}, \vec{k}) = -afh$
$bdi\text{Det}(\vec{j}, \vec{i}, \vec{k}) = -bdiDet(\vec{i}, \vec{j}, \vec{k}) = -bdi$
$bfg\text{Det}(\vec{j}, \vec{k}, \vec{i}) = -bfgDet(\vec{j}, \vec{i}, \vec{k}) = bfgDet(\vec{i}, \vec{j}, \vec{k}) = bfg$
$cdh\text{Det}(\vec{k}, \vec{i}, \vec{j}) = -cdhDet(\vec{i}, \vec{k}, \vec{j}) = cdhDet(\vec{i}, \vec{j}, \vec{k}) = cdh$
$ceg\text{Det}(\vec{k}, \vec{j}, \vec{i}) = -cegDet(\vec{k}, \vec{i}, \vec{j}) = cegDet(\vec{i}, \vec{k}, \vec{i}) = -cegDet(\vec{i}, \vec{j}, \vec{k})$
$= -ceg$.

Resumiendo:
$\text{Det}(P) = Det(a\vec{i} + b\vec{j} + c\vec{k}, d\vec{i} + e\vec{j} + f\vec{k}, g\vec{i} + h\vec{j} + i\vec{k})$
$= aei - afh - bdi + bfg + cdh - ceg = a(ei - fh) - d(bi - ch) + g(bf - ce)$.

Lo anterior se puede reescribir de una manera muy nemotécnica, si se entiende que el determinante de las matrices con estrellas es el determinante de las matrices 2×2 que quedan al quitar las estrellas:

$$Det(P) = \begin{vmatrix} a & d & g \\ b & e & h \\ c & f & i \end{vmatrix}$$

$$= aDet \begin{pmatrix} * & * & * \\ * & e & h \\ * & f & i \end{pmatrix} - dDet \begin{pmatrix} * & * & * \\ b & * & h \\ c & * & i \end{pmatrix} + gDet \begin{pmatrix} * & * & * \\ b & e & * \\ c & f & * \end{pmatrix}$$

$= a(ei - fh) - d(bi - ch) + g(bf - ce)$

249. *Ejemplo* *Usemos el Det para calcular el 3-volumen del 3-plp* $[(1, -1, 6], (2, 3, 2), (3, 5, 4)]$.

$$Det(P) = \begin{vmatrix} 1 & 2 & 3 \\ -1 & 3 & 5 \\ 6 & 2 & 4 \end{vmatrix}$$

$$= 1Det \begin{pmatrix} * & * & * \\ * & 3 & 5 \\ * & 2 & 4 \end{pmatrix} - 2Det \begin{pmatrix} * & * & * \\ -1 & * & 5 \\ 6 & * & 4 \end{pmatrix} + 3Det \begin{pmatrix} * & * & * \\ -1 & 3 & * \\ 6 & 2 & * \end{pmatrix}$$

$= 1(3 \times 4 - 2 \times 5) - 2(-1 \times 4 - 6 \times 5) + 3(-1 \times 2 - 6 \times 3)$
$= 2 - 2(-34) + 3(-20) = 2 + 68 - 60 = 10.$

250. ♣ ***Definición.*** *El* ***determinante*** *de una matriz $n \times n$ es el determinante del n-plp definido por las columnas de la matriz. De esta forma podemos identificar una matriz con el plp formado por sus columnas.*

251. $\boxed{\textbf{\textit{Ejercicio}}}$ *Calcule los siguientes determinantes.*

a) $\begin{vmatrix} 1 & 0 & 3 \\ -1 & -3 & 5 \\ 6 & 1 & 4 \end{vmatrix}$

b) $\begin{vmatrix} 7 & 5 & 3 \\ -1 & -3 & 3 \\ 2 & 0 & 5 \end{vmatrix}$

c) $\begin{vmatrix} 1 & 2 & -3 \\ -7 & -8 & 2 \\ -1 & -1 & 1 \end{vmatrix}$

d) $\begin{vmatrix} 1 & 0 & 3 \\ 0 & -3 & 0 \\ -6 & -1 & 4 \end{vmatrix}$

252. ◇ ***Teorema.*** *Para una matriz A, $n \times n$, y un escalar k tenemos que*

$$Det(kA) = k^n Det A$$

Demostración. El determinante de A es el determinante del n-plp formado por las columnas de A, los cuales son vectores de $\mathbb{R}^n$. Podemos escribir

$Det(A) = Det(\vec{c}_1, \vec{c}_2, ..., \vec{c}_n)$

donde las $\vec{c}_i$ son las columnas de A. Multiplicar A por k significa multiplicar cada entrada de A por k, y por lo tanto, cada columna también queda multiplicada por k:

$Det(kA) = Det(k\vec{c}_1, k\vec{c}_2, ..., k\vec{c}_n)$

Teniendo en cuenta que el determinante es multilineal, la constante puede salir de la primera columna:

$Det(kA) = kDet(\vec{c}_1, k\vec{c}_2, ..., k\vec{c}_n)$

y también puede salir de la segunda:

$Det(kA) = k^2 Det(\vec{c}_1, \vec{c}_2, ..., k\vec{c}_n)$

y saliendo de todas, obtenemos:

$Det(kA) = k^n Det(\vec{c}_1, \vec{c}_2, ..., \vec{c}_n) = Det(A)$.

Para memorizar este resultado, lo que uno tiene que tener en cuenta es que si uno tiene un rectángulo de área ab, y si cada lado se multiplica por k, el área se convierte en $(ka)(kb) = k^2 ab$. ■

253. $\boxed{\textit{\textbf{Ejercicio}}}$ *Hemos definido el determinante de una matriz $n\times n$ como el Det del n-pln definido por sus columnas. Pero las matrices vienen de sistemas de ecuaciones y nosotros introdujimos los determinantes para que representaran el volumen del n-plp formado por los vectores normales a cada línea, plano o correspondiente generalización. Dichos vectores normales pueden leerse directamente en los renglones de la matriz y no en las columnas. ¿Acaso definimos mal el determinante de una matriz?*

Nos interesa el determinante para decidir si un sistema de ecuaciones tiene solución única o no. Una de las formas como esa decisión puede justificarse es como sigue:

Cada ecuación de un sistema de n ecuaciones con n incógnitas representa una entidad geométrica llamada **hiperplano**. En un sistema $n\times n$ uno considera la intersección de n hiperplanos. Si ellos se cortan en un único punto, el sistema correspondiente tiene solución única. Si hay al menos dos de ellos paralelos, pero que no se toquen, la solución es vacía. Si todos los hiperplanos coinciden, la solución es el hiperplano de coincidencia. Para determinar la forma de intersección entre los hiperplanos, calculamos el determinante del n-plp formado por los vectores normales asociado al conjunto de hiperplanos. Si dicho volumen es cero, hay paralelismo entre los vectores normales y, por tanto, o hay una infinidad de soluciones o no hay solución. Si dicho volumen no es cero, los hiperplanos van cada uno por su lado y, por lo tanto, se cortan en un único punto y la solución al sistema es única.

En cierta forma, ya hemos terminado de ver la teoría referente a sistemas lineales con una única solución. Eso se debe a que una vez sepamos que la solución es única, uno puede buscarla por cualquier método, reemplazar para verificarla y olvidarse de soluciones espurias. Pero cuando la solución no es única, nos queda por averiguar si las soluciones quedan, por ejemplo, sobre un plano o si quedan sobre una línea o si no hay solución. Por eso que la teoría de sistemas lineales es algo que nos va a llevar tiempo y requerirá nuevos puntos de vista que elaboraremos a partir del próximo capítulo.

254. $\boxed{\textit{\textbf{Ejercicio}}}$ *Asocie un sistema de ecuaciones a cada una de las matrices del ejemplo 251 y prediga si dicho sistema tiene solución única o no.*

Según hemos visto, el determinante es un número, un escalar, que determina si la solución de un sistema lineal es única o no. Usando el mismo protocolo que en el cálculo del determinante, uno puede asociar a un par de vectores en el espacio 3D un tercer vector, llamado producto cruz, que es perpendicular a los dos primeros. En el protocolo del producto cruz usamos los símbolos $\vec{i}$, $\vec{j}$, $\vec{k}$ para indicar vectores de largo uno en las direcciones de los ejes coordenados.

255. ***Notación para el producto cruz.*** *Sea $\vec{u} = (u_1, u_2, u_3), \vec{v} = (v_1, v_2, v_3)$, notamos el producto cruz entre esos dos vectores mediante el siguiente simbolismo:*

$$\vec{u}\times\vec{v} = Det\begin{pmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix}$$

Recalquemos que en el primer renglón hay vectores, pero en los otros dos van coordenadas. Por eso, lo que nuestra definición dice es que se calcule simbólicamente este determinante usando las mismas reglas del determinante ordinario, pero que el resultado se interprete como un vector, cuyas componentes son las siguientes:

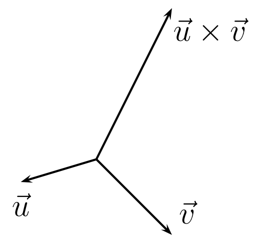


Figura 3.3. El producto cruz.

$$= \vec{i}Det\begin{pmatrix} * & * & * \\ * & u_2 & u_3 \\ * & v_2 & v_3 \end{pmatrix} - \vec{j}Det\begin{pmatrix} * & * & * \\ u_1 & * & u_3 \\ v_1 & * & v_3 \end{pmatrix} + \vec{k}Det\begin{pmatrix} * & * & * \\ u_1 & u_2 & * \\ v_1 & v_2 & * \end{pmatrix}$$

256. **_Ejemplo_** *Calculemos* $(1,2,3)\times(4,5,6)$ *y verifiquemos que es perpendicular a ambos vectores:*

$$\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{vmatrix}$$

$$= \vec{i}Det\begin{pmatrix} * & * & * \\ * & 2 & 3 \\ * & 5 & 6 \end{pmatrix} - \vec{j}Det\begin{pmatrix} * & * & * \\ 1 & * & 3 \\ 4 & * & 6 \end{pmatrix} + \vec{k}Det\begin{pmatrix} * & * & * \\ 1 & 2 & * \\ 4 & 5 & * \end{pmatrix}$$

$$= \vec{i}(-3) - \vec{j}(-6) + \vec{k}(-3) = -3\vec{i} + 6\vec{j} - 3\vec{k} = (-3,6,-3)$$

Verifiquemos la perpendicularidad:
$(1,2,3)\cdot(-3,6,-3) = -3+12-9{=}0$
$(4,5,6)\cdot(-3,6,-3) = -12+30-18 = 0.$

257. **_Ejercicio_** *Demuestre que el producto cruz cumple las propiedades siguientes:*

a) Si los dos vectores que se van a multiplicar están en el plano XY, la tercera coordenada es cero y el producto cruz coincide con el determinante de los vectores considerados como elementos del plano (con dos coordenadas). En ese caso, demuestre que $\|\vec{u}\times\vec{v}\| = \|\vec{u}\|\,||\vec{v}\|\operatorname{sen}\theta$, donde θ es el ángulo entre los dos vectores, lo cual da el área del plp expandido por los dos vectores.

b) Sin un sólo cálculo, demuestre que para cualquier par de vectores del espacio se cumple que $\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \, ||\vec{v}|| \operatorname{sen} \theta$.

c) Demuestre que el 3-volumen de un 3-plp $[\vec{u}, \vec{v}, \vec{w}]$, que es igual al valor absoluto de su determinante, también es igual al valor absoluto de $(\vec{u} \times \vec{v}) \cdot \vec{w}$.

258. ***Ejemplo y ejercicio*** *Cuando el determinante de la matriz asociada a un sistema de n ecuaciones de n incógnitas es diferente de cero, el sistema tiene solución única. Dicha solución puede escribirse en términos de determinantes. Para un sistema 2×2 de la forma*

$$\begin{cases} ax + by = m \\ cx + dy = n \end{cases}$$

la solución puede escribirse como:

$$x = \frac{Det\begin{pmatrix} m & b \\ n & d \end{pmatrix}}{Det\begin{pmatrix} a & b \\ c & d \end{pmatrix}} \qquad\qquad y = \frac{Det\begin{pmatrix} a & m \\ c & n \end{pmatrix}}{Det\begin{pmatrix} a & b \\ c & d \end{pmatrix}}$$

lo cual puede probarse por sustitución directa (ejercicio).

259. ***Ejercicio*** *Observe la estructura de las expresiones de la solución del ejemplo anterior y generalice para sistemas de 3 incógnitas con 3 ecuaciones que tengan solución única. Invente una notación que pueda generalizarse para cualquier sistema de n ecuaciones con n incógnitas.*

260. ***Ejercicio*** *Invente ejercicios que ilustren los descubrimientos hechos en el ejercicio anterior.*

261. ***Ejercicio*** *Demuestre que el determinante de una matriz escalonada superior (que tenga ceros abajo de la diagonal) es el producto de los elementos de la diagonal.*

262. ***Intriga*** *Debido a que en las aplicaciones uno puede encontrarse con sistemas de muchas variables, vale la pena preguntarse cuál podría ser una manera eficiente de calcular determinantes de matrices grandes. A la luz del resultado del ejercicio anterior, valdría la pena escalonar la matriz original (se transforma por Gauss-Jordan en otra que tenga ceros debajo de la diagonal). El determinante de la matriz original seguramente esté relacionado con el determinante de la matriz escalonada. Pero, ¿cómo?*

La orientación tiene importancia en muchas áreas de la ciencia. La idea de determinante surgió de la necesidad de calcular volúmenes de plps en cualquier dimensión. Por eso siempre habíamos tomado el valor absoluto del determinante y el signo lo ignorábamos. Pero en realidad, el signo del determinante tiene una utilidad sin igual y es que distingue la mano derecha de la izquierda.

En efecto, si el dedo pulgar de la mano izquierda representa el eje X ($\vec{i}$), el anular el Y ($\vec{j}$) y el índice el Z ($\vec{k}$), la mano izquierda genera el plp $[\vec{i}, \vec{j}, \vec{k}]$ cuyo determinante es $+1$. Decimos que la **orientación** de la mano izquierda es el $Det[\vec{i}, \vec{j}, \vec{k}]$, en ese orden, y que la mano izquierda tiene orientación positiva. En contraste, la mano derecha tiene orientación negativa: si el dedo pulgar de la mano derecha representa el eje X ($\vec{i}$), el anular el Y ($\vec{j}$) y el índice el Z ($\vec{k}$), la mano derecha genera un plp, para lo cual primero se numera el piso, de izquierda a derecha y después el eje Z y el plp es $(\vec{j}, \vec{i}, \vec{k})$ cuyo determinante es -1. Decimos que la orientación de la mano derecha es el $Det(\vec{i}, \vec{j}, \vec{k})$, en ese orden, y que la mano derecha tiene orientación negativa. Científicamente, la forma como se numeran los elementos de un plp se llama quiralidad y uno habla de quiralidad derecha o izquierda. La quiralidad es notada usualmente como L(*left*) o R(*right*) pero a los químicos les gusta R la primera letra de la palabra latina *rectus* (derecha) o S, la primera letra de la palabra latina *sinister* (izquierda).

El problema de la quiralidad se reduce algunas veces a distinguir si se está a la derecha o a la izquierda de un eje de referencia y, por ende, a definir la forma como un objeto es asimétrico. Un caso que siempre causa curiosidad se relaciona con las mutaciones que cambian la posición de los órganos del cuerpo. En la gran mayoría de humanos, el corazón, el páncreas y el bazo quedan a la izquierda, mientras que el hígado queda a la derecha. Pero de vez en cuando los noticieros reportan un mutante que tiene todo al revés y que se desenvuelve tan bien como los normales.

El problema de la quiralidad en química consiste en predecir por las leyes de la física el origen de la quiralidad molecular, la forma espacial adoptada por algunas moléculas que se orientan, bien como la mano derecha, bien como la izquierda en vez de ser simétricas. La mecánica cuántica es en principio la que tiene la palabra, la cual es, sin exagerar, un capítulo del álgebra lineal y ella predice que no debe haber distinción entre derecha e izquierda. La predicción de la mecánica cuántica es fácil de entender: la diferencia entre derecha e izquierda se percibe en mecánica cuántica por un signo que puede ser positivo o negativo. Pero en mecánica cuántica, dicho signo aparece al cuadrado, por tanto, no importa el signo, lo cual quiere decir que la mecánica cuántica no puede distinguir entre derecha e izquierda y, por lo tanto, no puede explicar estructuras asimétricas. Sin embargo, muchas moléculas tienen quiralidad, asimetría espacial. Este problema ha llegado a ser tan grande que se enmarca dentro de una temática desafiante: ¿cómo nacen las propiedades de los sistemas complejos que desafían las leyes establecidas para sistemas simples? (Primas, 1983).

La orientación o quiralidad es importante en bioquímica: muchos tipos de moléculas de los seres vivos forman estructuras espaciales y pueden quedar como la mano izquierda L o como la derecha R. Todos los aminoácidos, los componentes de las proteínas, son L. La dextrosa es R. Diversas moléculas pueden tener la misma fórmula estructural pero diferente organización espacial y conllevar a diversas propiedades biológicas. La forma como llegamos a saberlo fue porque un compuesto llamado talidomida, que se usaba como sedante en los años 1950, produjo unos 10.000 bebés en todo el mundo con malformaciones genéticas. Investigando, se llegó a la conclusión de que la causa tenía que ver con la organización espacial de los átomos en formas quirales (Talidomida en Wikipedia, 2009).

La quiralidad es importante en antropología: las formas R y L de un sitio orientable de una molécula no son estables sino que, debido al efecto túnel de la mecánica cuántica, una molécula puede cambiar de una estructura a la otra. Eso lleva tiempo y para los aminoácidos la vida media del estado L es alrededor de un millón de años, lo cual permite estimar la edad de fósiles muy antiguos (Snakefly, 2009).

La orientación también es importante en física de partículas elementales, por ejemplo, algunos experimentos no se pueden explicar a menos que se asuman neutrinos de quiralidad izquierda (Halzen y Martin, 1984).

El siguiente ejercicio muestra un detalle de cómo se vive la orientación en geometría diferencial y física matemática (Nash y Sen, 1983).

263. **_Ejercicio_** *La cinta de Möbius es como un cinturón que antes de cerrarlo se le da media vuelta a la correa y al cual se le recorta la parte de cuero que quede fuera de la chapa.*

a) Pruebe experimentalmente que la cinta de Möbius tiene sólo una cara. Es decir, si uno da un recorrido cerrado a lo largo de la cinta, uno termina recorriendo toda la correa por ambas caras del cinturón.

b) Dibuje sobre un pedacito de papel el plp $[\vec{i}, \vec{j}]$*. Ponga el papelito sobre la cinta de tal forma que parezca el plano XY. Transporte el plp sin girarlo a lo largo de la cinta hasta que llegue al mismo lugar. Demuestre que después de un viaje cerrado, el plp original se transforma en* $[\vec{i}, -\vec{j}]$*. Por eso decimos que la cinta de Möbius no es orientable.*

3.3. Ejercicios de repaso

1. Use la geometría del determinante para decidir si los siguientes sistemas tienen solución única o no. Por favor, haga gráficas adecuadas que demuestren la naturalidad de todos los conceptos utilizados.

   (a) $x = 1$, $y = 2$.

   (b) $x - y = 2$, $2x - 9y = 5$.

   (c) $x + y = 1$, $2x - 3y = 4$.

   (d) $x = 0$, $y = 0$, $z = 0$.

   (e) $x = 1$, $y = 1$, $z = 1$.

   (f) $x + y + z = 1$, $x - y + z = 2$, $x - y - z = 3$.

2. Sea A una matriz 4×4 con $\mathrm{Det}(A) = -3$, halle $\mathrm{Det}(-2A)$.

3. Sean A, B matrices 4×4 con $\mathrm{Det}(A) = -3$ y $\mathrm{Det}(B) = 8$. Halle $\mathrm{Det}(-2A + 9B)$.

4. Calcule el volumen del plp generado por los 3 vectores: $(1, 1, 1), (2, -1, 3), (1, 3, -1)$. Use dos métodos, el del determinante y el otro que utiliza tanto el producto cruz como el producto punto.

5. Halle por el método de Kramer las soluciones a los sistemas de ecuaciones del punto uno.

3.4. Resumen

El determinante de una matriz M es un número que predice si un sistema de ecuaciones asociado a M tiene solución única o no. Si el determinante es diferente de cero, el sistema tiene solución única, la cual puede buscarse por cualquier método, por ejemplo, por Kramer. Pero si el determinante es cero, puede pasar que no haya solución o que haya una infinidad de soluciones. Nosotros sabemos esto y también sabemos claramente el porqué.

CAPÍTULO 4

DEPENDENCIA E INDEPENDENCIA LINEAL

Cuando un sistema tiene solución única, lo cual puede averiguarse con el determinante, uno puede hallar la solución por cualquier método, verificarla e interpretarla como un punto en el espacio. Y ya no hay nada más qué hacer. Pero cuando la solución no es única, los problemas empiezan con la incertidumbre de no saber si no hay solución o si hay una infinidad de soluciones que quedan tal vez sobre una recta, o sobre un plano o sobre algo más extendido. Para resolver esa incertidumbre necesitamos profundizar la temática. Curiosamente, eso no implica grandes maravillas sino simplemente aclarar los pequeños detalles que andan por ahí medio perdidos entre tanto material y ponerlos a funcionar en una gran maquinaria.

Un primer detalle es la posibilidad de combinar los elementos de un EV con multiplicación por escalares y sumándolos entre ellos. Comenzamos elaborando esa posibilidad con el estudio del plano $\mathbb{R}^2$. Resulta que su dimensionalidad dos no juega ningún papel especial y que algunos conceptos importantes válidos en el plano tienen sentido en tanto en cualquier dimensión como en cualquier espacio vectorial, como puede ser el espacio de los polinomios de grado menor o igual que n.

4.1. Combinaciones lineales

Los elementos del plano $\mathbb{R}^2$ se notan (x, y) que de acuerdo con Descartes inmediatamente dan la posición donde queda dicho elemento. Hagamos la siguiente observación:

$(x, y) = (x, 0) + (0, y) = x(1, 0) + y(0, 1)$

lo que esto significa es que cualquier punto del plano XY es simplemente un resultado de combinar apropiadamente el conjunto de vectores $\{\vec{i} = (1, 0), \vec{j} = (0, 1)\}$. En el presente contexto, combinar vectores significa simplemente multiplicar por constantes y sumar.

En general, los vectores pueden sumarse, y un vector cualquiera puede multiplicarse por una constante. Por ahora, no tenemos ningún modo general de multiplicar vectores en cualquier EV de tal forma que el resultado del producto sea un vector. Hay que advertir que nosotros hemos considerado el producto interno pero tal producto da números y no vectores; también consideramos el producto cruz, pero lo definimos por una construcción que tenía sentido sólo en 3 dimensiones. Así que no tenemos multiplicación de vectores ni mucho menos división. Con lo permitido hasta ahora, dados dos vectores $\vec{u}, \vec{v}$ todo lo que puede hacerse se reduce a multiplicarlos por un escalar y luego sumarlos, produciendo así lo que llamamos una **combinación lineal** de los dos vectores. Específicamente tenemos:

264. ♣ ***Definición.*** *Dado un conjunto de vectores* $B = \{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ *de un espacio vectorial* V, *una* ***combinación lineal*** *de* B *es una expresión del tipo*
$\vec{w} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + ... + \alpha_n\vec{v}_n$ *con* $\alpha_1, \alpha_2, ..., \alpha_n$ *números reales.*

265. ***Ejemplo y contraejemplo*** *El vector* $(2, 13)$ *es combinación lineal del conjunto* $B = \{(1, 3), (4, 5)\}$ *pues* $6(1, 3) - (4, 5) = (6, 18) - (4, 5) = (2, 13)$ *(se sobreentiende que estamos trabajando en* $\mathbb{R}^2$*). Sin embargo, el vector* $(1, 1)$ *no es combinación lineal del conjunto* $B = \{(1, 0), (3, 0)\}$ *pues si así fuese, existirían constantes* α *y* β *tales que* $(1, 1) = \alpha(1, 0) + \beta(3, 0)$ *lo cual implicaría que* $(1, 1) = (\alpha + 3\beta, 0)$ *de donde se deduce, mirando las segundas coordenadas, que* $1 = 0$ *lo que es una contradicción.*

266. ♣ ***Definición.*** *Dado* B *subconjunto finito de un espacio vectorial* V, *el conjunto generado por* B *es el conjunto que reúne todas las combinaciones lineales que se pueden hacer con los elementos de* B *y se nota gen*(B).

267. ***Ejemplo y contraejemplo*** *Tomemos el plano* $\mathbb{R}^2$*:*

La igualdad
$(x, y) = (x, 0) + (0, y) = x(1, 0) + y(0, 1)$, donde x, y son reales,
dice que el plano $\mathbb{R}^2$ es el conjunto de todas las combinaciones lineales de $\{\vec{i}, \vec{j}\}$ o que el plano es generado por el conjunto de vectores $\{\vec{i}, \vec{j}\}$. Cuando $x > 0, y > 0$ se genera el primer cuadrante. Si $x < 0, y > 0$ se genera el segundo cuadrante. Si $x < 0, y < 0$ se genera el tercer cuadrante. Si $x > 0, y < 0$ se genera el cuarto cuadrante.

Pero, por otra parte:
$(x, y) = -(-x, -y) = -(-x, 0) + [-(0, -y]) = -x(-1, 0) + (-y)(0, -1)$
lo cual dice que el plano también es el conjunto de todas las combinaciones lineales de $\{-\vec{i}, -\vec{j}\}$.

Más aún, el conjunto de vectores $D = \{\vec{u} = (1, 1), \vec{v} = (-1, 1)\}$ también genera todo el plano. Eso se demuestra probando que los vectores que generan el plano, $\{\vec{i}, \vec{j}\}$, también son generados por D. En efecto : $(1, 1) - (-1, 1) = (2, 0)$ por lo que
$\vec{i} = (1/2)(1, 1) - (1/2)(-1, 1)$. Además, $(1, 1) + (-1, 1) = (0, 2)$ por lo que
$\vec{j} = (1/2)(1, 1) + (1/2)(-1, 1)$.

Sin embargo, el plano no es generado por el conjunto de vectores
$D = \{\vec{u} = (1, 1), \vec{v} = (3, 3)\}$,
pues ambos vectores de ese conjunto generan la misma línea. Es decir, las combinaciones lineales de $\{\vec{u}, \vec{v}\}$ producen la misma línea que las combinaciones lineales de $\{\vec{u}\}$ o $\{\vec{v}\}$.

De esta manera se distinguen 2 clases de conjuntos de vectores. Veamos la primera:

268. ♣ ***Definición.*** *Un conjunto de vectores se dice que es* ***LD (linealmente dependiente****) cuando uno cualquiera de ellos es combinación lineal de los otros. Es decir, cuando hay redundancia de vectores en el conjunto.*

Podemos clarificarlo un poco mejor:

269. $\diamond$ ***Teorema.*** *Si un conjunto finito de vectores B es linealmente dependiente, entonces existe la forma de quitarle algún vector, notémoslo $\vec{w}$, de tal forma que el conjunto resultante $B - \{\vec{w}\}$ queda generando el mismo conjunto que B.*

Demostración. Supongamos que un vector cualquiera de ellos, notémoslo $\vec{w}$, es combinación lineal de los otros, $\vec{w} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + ... + \alpha_n\vec{v}_n$, con lo cual estamos diciendo que $B = \{\vec{v}_1, ..., \vec{v}_n, \vec{w}\}$. Entonces una combinación lineal cualquiera de B es una expresión del tipo

$\vec{z} = \gamma_1\vec{v}_1 + ... + \gamma_n\vec{v}_n + \epsilon\vec{w} = \gamma_1\vec{v}_1 + ... + \gamma_n\vec{v}_n + \epsilon(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + ... + \alpha_n\vec{v}_n)$

lo cual es una combinación lineal del conjunto en el cual $\vec{w}$ no aparece. Por lo que $\vec{w}$ no añade nada nuevo y puede ser suprimido sin que B pierda capacidad generatriz. ■

270. $\boxed{\textbf{\textit{Comentarios y ejercicios}}}$ *En la demostración del teorema anterior usamos la poderosa capacidad humana de razonar y no fuimos muy formales. Con todo, es conveniente tener presente lo que la formalidad exige. En primer término, podemos escribir el teorema anterior así: si un conjunto finito B es LD entonces existe $\vec{w} \in B$ tal que gen(B) = gen($B - \{\vec{w}\}$).*

Ahora bien, para demostrar que dos conjuntos A y B son iguales hay que demostrar dos cosas, que $A \subset B$, el primer conjunto está contenido en el segundo (que todos los elementos del primer conjunto están en el segundo) y viceversa, que $B \subset A$ o sea que el segundo conjunto está contenido en el primero (todos los elementos del segundo conjunto están en el primero). Pues bien, nuestra demostración solamente prueba que gen(B) $\subset$ gen($B - \{\vec{w}\}$). Pero faltaría hacer la segunda parte y demostrar que gen($B - \{\vec{w}\}$) $\subset$ gen(B). No lo hicimos porque nos parece obvio que si un conjunto está contenido en otro entonces sus conjuntos generados guardan la misma relación.

Por otro lado, también es cierto que si para un conjunto finito B se tiene que existe $\vec{w} \in B$ tal que gen(B) =gen($B - \{\vec{w}\}$), entonces ese conjunto es LD. Demostrarlo queda de ejercicio.

Reuniendo el teorema con el ejercicio podemos decir sucintamente: un conjunto finito B es LD ssi (si y sólo si) existe $\vec{w} \in B$ tal que gen(B) =gen($B - \{\vec{w}\}$). Vemos que para demostrar una proposición de la forma

$$p \text{ ssi } q$$

hay que demostrar dos cosas, que a partir de p se puede concluir q y que de q se puede concluir p.

No todos los conjuntos de vectores son LD, y como son tan importantes, merecen un nombre.

271. ♣ ***Definición y ejemplo.*** *Un conjunto es LI (**linealmente independiente**) si no es LD. Por ejemplo, el conjunto $\{-\vec{i}, -\vec{j}\}$ es LI. Lo mismo el conjunto $\{(1,2),(1,3)\}$.*

Decimos que un conjunto A es estrictamente mayor que otro B cuando A contiene a B, pero A tiene al menos un elemento más que B.

272. $\diamond$ ***Teorema y ejercicio.*** *Un conjunto de vectores B es LI ssi todo elemento de B es esencial en el sentido que el conjunto generado por B es estrictamente mayor que el generado por el conjunto que resulta de quitarle un vector cualquiera a B.*

Demostración: ejercicio

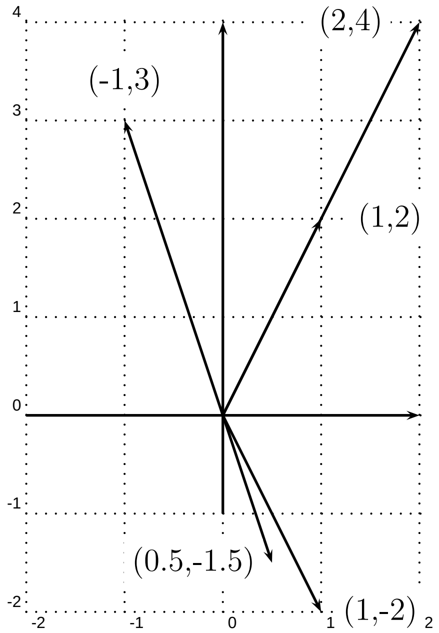


Figura 4.0. Conjuntos LD:
$\{(1,2),(2,4)\},\{(1,2),(2,4),(-1,3)\},\{(-1,3),(0{,}5,-1{,}5)\}$.
Conjuntos LI: {(1,2), (-1,3)}, {(1,2),(0.5,-1.5)},{(-1,3)(0.5,-1.5)},{(2,4),(1,-2)}.

273. ***Interpretación geométrica.*** *Los conceptos LD, LI tienen contraparte geométrica que en algunos casos simplifica las tareas. Citemos algunos ejemplos:*

Primero que todo, consideremos dos vectores no nulos paralelos, el uno múltiplo del otro. El primero determina una línea. El segundo determina la misma línea. Los dos juntos determinan la misma línea: son redundantes. Son dependientes. Pero si los dos vectores dan líneas diferentes, entonces, ellos forman un conjunto LI, pues cada vector da información independiente.

Consideremos el conjunto formado por el vector nulo y otro vector no nulo. El vector nulo genera el origen. El otro vector genera una línea que pasa por el origen. El vector nulo es redundante: no da nada que no esté en el espacio generado por el segundo vector. Este conjunto es dependiente.

El vector $\vec{i} = (1,0)$ genera el eje X, y el vector $\vec{j} = (0,1)$ genera el eje Y. Entre los dos generan todo el plano. Si a dichos ejes se les rota, dentro del plano, por separado para que el ángulo entre ellos quede agudo u obtuso o al revés, de todas formas generarán todo el espacio, a menos que queden superpuestos o paralelos.

Consideremos 3 vectores en el plano en direcciones distintas. Ellos generan 3 líneas diferentes. Y, sin embargo, forman un conjunto LD, la razón es que entre dos líneas generan un plano. La tercera línea no agrega nada nuevo. La tercera línea es redundante. Al igual que la primera es redundante con respecto a las otras dos. En definitiva, en el conjunto de las 3 líneas hay redundancia. El conjunto de los 3 vectores es un conjunto LD.

Consideremos 3 vectores en el espacio no totalmente contenidos en un plano. Forman un conjunto LI. Pero si tomamos 4 vectores en el espacio, el conjunto formado es LD, pues entre los 4 hay uno, al menos, que no añade información independiente de la que ya aportan los otros.

274. **_Ejercicio_** *Por inspección gráfica diga cuáles de los siguientes conjuntos de vectores son linealmente independientes.*

1. $\{(1,1),(2,3)\}$
2. $\{(-1,-1),(1,1)\}$
3. $\{(1,1)\}$
4. $\{(1,1),(0,0)\}$
5. $\{(1,1),(2,2)\}$
6. $\{(1,2),(2,3),(-1,-2)\}$
7. $\{(1,2,3)\}$
8. $\{(1,2,3),(2,4,6)\}$
9. $\{(1,2,3),(2,4,7)\}$
10. $\{(1,2,3),(2,4,7),(1,0,0)\}$
11. $\{(1,2,3),(2,4,7),(1,0,0),(0,0,1)\}$
12. Una fábrica produce sillas y escritorios, cuyas cantidades relativas con respecto a la media las notamos (x,y). Observe cada uno de los conjuntos anteriores con vectores de $\mathbb{R}^2$ y prediga si se trata de fábricas clonadas u originales. Considere ahora los vectores con tres coordenadas y piense en (sillas, escritorios, archivadores) y tome los conjuntos de g) a k) y resuelva la misma pregunta.

Existe una forma totalmente automática de decidir si un conjunto de vectores es LD o no. La idea es simple y su prueba lo es aún más. Observemos el conjunto $S=\{(1,1),(2,2)\}$. Este conjunto es LD puesto que el subconjunto $s=\{(1,1)\}$ genera la misma línea que S. Por eso, $(2,2)$ es redundante con respecto a s: es suficiente ver que $2(1,1)=(2,2)$. O, lo que es lo mismo, $2(1,1)-(2,2)=(0,0)$. Existe otra forma de leer esa igualdad: si comenzamos un viaje desde el origen y viajamos hasta $2(1,1)$ y después viajamos al revés de $(2,2)$, entonces retornamos al origen: podemos hacer un viaje redondo (que llega al mismo lugar de partida) combinando los elementos de S, empezando y terminando en el origen.

275. *Ejemplo* *La suma* $(3,1)+(1,2)=(4,3)$ *dice que el conjunto* $S=\{(1,2),(3,1),(4,3)\}$ *es LD.*

Pero esto es equivalente a decir que $(3,1)+(1,2)-(4,3)=(0,0)$*, lo cual expresa que si comenzamos desde el origen y viajamos hasta* $(3,1)$ *y después caminamos lo equivalente a* $(1,2)$ *y después retrocedemos lo equivalente a* $(4,3)$*, entonces retornamos al origen.*

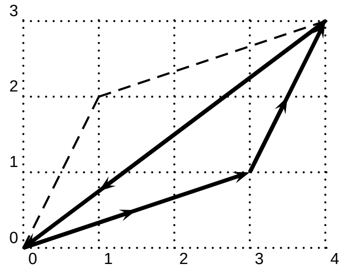


Figura 4.1. Un viaje redondo con un conjunto LD.

Pero cuando el conjunto S es LI, nosotros no podemos hacer un viaje redondo que empiece y termine en el origen a menos que aniquilemos todos los elementos de S multiplicándolos por la constante cero. Veamos cómo funciona eso para el conjunto $\{(1,1),(-1,1)\}$. Supongamos que podemos hacer un viaje redondo a partir del origen, i.e. supongamos que existen escalares α, β tal que

$\alpha(1,1)+\beta(-1,1)=(0,0)$

$(\alpha,\alpha)+(-\beta,\beta)=(0,0)$

$(\alpha-\beta,\alpha+\beta)=(0,0)$

Esto nos da un sistema de dos ecuaciones con dos incógnitas:

$\alpha-\beta=0$, $\alpha+\beta=0$ Sumando obtenemos $\alpha=0$ y restando queda que $\beta=0$. Es

decir, la única forma de hacer un viaje redondo es anulando los recorridos.

276. $\diamond$ ***Teorema del viaje redondo.*** *Un conjunto B es LI ssi el único viaje redondo que permite es el viaje nulo. Un conjunto B es LD ssi existe un viaje no nulo redondo que combina los elementos de B.*

Demostración. Si el conjunto B es LD, uno de sus vectores es redundante, notémoslo $\vec{w}$, puede ser expresado como combinación lineal de los otros y obtenemos una expresión al estilo

$\vec{w}=\alpha_1\vec{v}_1+...+\alpha_n\vec{v}_n$

donde no todos los escalares son cero y con lo cual estamos diciendo que

$B=\{\vec{v}_1,...,\vec{v}_n,\vec{w}\}$. Esta expresión es equivalente a:

$\vec{w}-\alpha_1\vec{v}_1-...-\alpha_n\vec{v}_n-\epsilon\vec{w}=0$

Partiendo de la suposición de que B es LD hemos construido un viaje no nulo y redondo que empieza y termina en el origen.

Recíprocamente: si tenemos partiendo del origen un viaje no nulo y redondo con elementos de B, tenemos una expresión de la forma

$\alpha_1\vec{v}_1 + ... + \alpha_n\vec{v}_n = 0$

con algún coeficiente α_i distinto de cero y con lo cual estamos notando a B de la forma $B = \{\vec{v}_1, ..., \vec{v}_n\}$. Se puede dividir por dicho coeficiente, pues no es cero, y después se puede despejar el correspondiente $\vec{v}_i$:

$\vec{v}_i = -(\alpha_{j_1}/\alpha_i)\vec{v}_{j_1} - ... - (\alpha_{j_{n-1}}/\alpha_i)\vec{v}_{j_{n-1}}$

con lo que estamos probando que si B permite un viaje redondo no nulo entonces B es LD pues tiene un vector redundante. Esa notación complicada corresponde a una renumeración y es la forma de decir que el vector $\vec{v}_i$ ya no aparece en la combinación lineal de la izquierda.

277. **Ejercicio** *Use el teorema del viaje redondo para escribir pruebas formales de los hechos que por intuición usted encontró en el ejercicio 274.*

Cuando hablamos de combinaciones lineales, hablamos de álgebra, pues hacemos operaciones algebraicas. ¿Cuál es el equivalente geométrico de dependencia o independencia lineal?

Supongamos que estamos en el plano, sobre el cual dibujamos un 2-plp. Si uno de sus lados es cero, el área del plp es cero. Si las dos aristas fundamentales del plp son paralelas, formando un conjunto LD, su área también es cero. Pero si las aristas fundamentales no son paralelas, entonces éstas expanden un plp cuya área no es cero y por consiguiente su determinante tampoco es cero. Esto significa que si un 2-plp es generado por un conjunto LI, su determinante no puede ser cero. Pasemos ahora a 3D, en donde tomamos un 3-plp, el cual es expandido por 3 vectores que salen del origen. Si cualquiera de los tres vectores es cero, el plp tiene volumen cero. Si los tres vectores expanden sólo un plano, el plp correspondiente está contenido en un plano con volumen cero y el determinante también es cero. Pero si los tres vectores no están sobre un plano, formando un conjunto LI, el volumen es diferente de cero al igual que su determinante.

278. ◇ ***Teorema y definición.*** *Un conjunto de n vectores en $\mathbb{R}^n$ es LI ssi su determinante es diferente de cero. Un n-plp tiene volumen diferente de cero ssi es generado por un conjunto LI. Un n-plp cuyo determinante es cero y que es generado por un conjunto LD se denomina un* ***plp apachurrado****.*

La prueba de este teorema es simplemente una generalización del siguiente ejemplo: consideremos el 3-plp definido por $[\vec{a}, \vec{b}, \vec{a} + \vec{b}]$, cuyo determinante obedece la relación

$\mathrm{Det}[\vec{a}, \vec{b}, \vec{a} + \vec{b}] = \mathrm{Det}[\vec{a}, \vec{b}, \vec{a}] + \mathrm{Det}[\vec{a}, \vec{b}, \vec{b}]$.

Notemos ahora que el plp $[\vec{a}, \vec{b}, \vec{a}]$ forma un paralelogramo cuyo 3-volumen es cero. En general, el *Det* de un plp con 2 vectores iguales es cero. Esto es una consecuencia directa del axioma de alternancia del Det, si se intercambian dos vectores de un plp o de una matriz, el Det cambia de signo. Por tanto, en $[\vec{a}, \vec{b}, \vec{a}]$ se puede intercambiar el primer vector con el tercero, pero el resultado permanece inalterado y en cambio el Det cambia de signo. Por tanto, tenemos:

$\text{Det}[\vec{a}, \vec{b}, \vec{a}] = -\text{Det}[\vec{a}, \vec{b}, \vec{a}]$

Puesto que estamos trabajando con números reales, esta ecuación es válida sólo para el número cero. Es decir,

$\text{Det}[\vec{a}, \vec{b}, \vec{a}] = 0$.

279. **_Ejercicio_** *Use la tecnología del determinante para demostrar que cada uno de los siguientes conjuntos es LD.*

a) $[(0, 0), (1, 2)]$
b) $[(1, 2), (2, 4)]$
c) $[(1, 2), (-1, -2)]$
d) $[(0, 0, 0), (1, 2, 3), (4, 5, 6)]$
e) $[(1, 2, 3), (2, 4, 6), (7, 8, 9)]$
f) $[(1, 2, 3), (4, 5, 6), (5, 7, 9)]$.

Entramos a estudiar desde el punto de vista geométrico el conjunto de todas las combinaciones lineales generadas por un conjunto finito.

4.2. Subespacios

Si V es un EV, un subespacio vectorial de V es un subconjunto de V que se comporta como si el mismo fuese un EV. Es decir, un subespacio es un EV pequeño metido dentro de otro. Esto también implica que la suma sobre el espacio pequeño es la misma suma del espacio grande pero restringida al espacio pequeño, y lo mismo con la multiplicación escalar.

280. ♣ ***Definición.*** *Si V es un EV, se dice que U es* ***subespacio*** *de V si U es no vacío, si está contenido en V y si la suma y la multiplicación escalar de V restringidas a U son cerradas sobre U.*

281. **_Ejercicio_** *Formalice y demuestre la aseveración de que un sub-EV es un EV metido dentro de otro más grande.*

282. **_Ejemplo_** *El eje X es subespacio vectorial de $\mathbb{R}^2$.*

283. ***Contraejemplo*** *$\mathbb{R}$ no es subespacio de $\mathbb{R}^2$, puesto que los números reales no aparecen en la lista de los elementos del plano. Con todo, uno establece una equivalencia entre $\mathbb{R}$ y el eje X y de esa forma uno dice que $\mathbb{R}$ sí es subespacio del plano.*

284. **_Ejemplo_** *El conjunto de elementos del plano que quedan sobre la línea L definida por $y = 3x$ es un subespacio vectorial de $\mathbb{R}^2$.*

En efecto: dicho conjunto está contenido en el plano, del que hereda una suma y una multiplicación escalar, las cuales debemos demostrar que son cerradas. Para la suma:

Sean $\vec{u}$ y $\vec{v}$ dos elementos de L, debemos demostrar que al sumarlos su resultado está sobre la línea. Procedamos: los elementos de la línea L son de la forma $(l, 3l)$. Entonces, si $\vec{u}$ está en la línea, existen coordenadas (u_1, u_2) tales que $u_2 = 3u_1$, es decir $\vec{u} = (u_1, 3u_1)$. En cuanto a $\vec{v}$, éste se puede escribir como $(v_1, 3v_1)$. Al sumar estos dos elementos obtenemos $(u_1 + v_1, 3u_1 + 3v_1)$, que es lo mismo que $(u_1 + v_1, 3(u_1 + v_1))$ que es de la forma $(z, 3z)$. por lo tanto, la suma es cerrada.

Ahora veamos que la multiplicación escalar es cerrada. Tomamos un elemento de la línea, $(l, 3l)$, lo multiplicamos por un escalar, λ, nos da $(\lambda l, \lambda(3l)) = (\lambda l, 3(\lambda l))$ que es de la forma $(z, 3z)$, por lo que está en la línea. Entonces la multipliación escalar es cerrada.

Como la suma y la multiplicación escalar son operaciones cerradas sobre L, tenemos un subespacio vectorial. Todas las demás propiedades que L necesita para ser *EV*, L las tiene por herencia del plano, el cual es *EV*.

285. ***Contraejemplo*** *La línea $y = 3x + 1$ no es subespacio del plano.*

Es suficiente demostrar que la suma no es cerrada. Tomamos dos elementos determinados $(1, 4)$, $(2, 7)$. Al sumarlos da $(3, 11)$ que no es de la forma $(l, 3l + 1)$, pues debiera haber sido $(3, 10)$. Además, si $x = 0$ entonces $y = 1$, por lo que el elemento (0,0), el cual es el cero del plano considerado como espacio vectorial, no está en la línea. Esa es otra razón que por sí misma es suficiente para decir que la línea no es un subespacio del plano.

286. ***Contraejemplo*** *La circunferencia de radio uno no es EV.*

En efecto, los puntos $(1, 0)$ y $(0, 1)$ pertenecen a dicha circunferencia, pues cumplen con la ecuación $x^2 + y^2 = 1$. Pero al sumarlos da $(1, 1)$, que ya no cumple con dicha ecuación.

287. $\boxed{\textbf{\textit{Ejercicio}}}$ *Decida cuáles de los siguientes conjuntos, definidos por la condición dada, son sub-EV del EV indicado:*

a) $y = 3x + 1$ de $\mathbb{R}^2$.

b) $y = 7x - 2z$ en $\mathbb{R}^3$.

c) $z = 4x - 5y$ en $\mathbb{R}^3$.

d) $y = 3$ en $\mathbb{R}$.

e) $y = x^2$ en $\mathbb{R}^2$.

Hay una manera estándar de generar subespacios, que es simplemente añadir todo lo que sea estrictamente necesario para lograr que uno pueda sumar y alargar sin problema.

288. ♣ ***Recordemos.*** *Dado un conjunto de vectores D de un EV V, se nota por gen(D) el conjunto de todas las combinaciones lineales finitas de D y se llama* ***el conjunto generado por*** *D. Concretamente, si $D = \{\vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$, entonces*

$$gen(D) = \{\lambda_1\vec{u}_1 + \lambda_2\vec{u}_2 + ... + \lambda_n\vec{u}_n : \lambda_i \in \mathbb{R}, i = 1, 2, ..., n\}.$$

Por ejemplo, si $D = \{\vec{i}, \vec{j}\}$, tenemos que gen(D) es todo el plano. En efecto, los elementos de D son LI, y además cualquier elemento (x, y) del plano puede escribirse como $(x, y) = x\vec{i} + y\vec{j}$, es decir, como una combinación lineal de D, lo cual implica que el plano está contenido en gen(D). Por supuesto que gen(D) no es más grande que el plano, pues cualquier combinación de D queda en el plano, pues el plano es cerrado a las combinaciones lineales, pues es un EV.

289. $\diamond$ ***Teorema y ejercicio.*** *El conjunto gen*(D) *generado por* D, *un subconjunto de un EV, por medio de todas sus combinaciones lineales es cerrado bajo la suma y bajo la multiplicación escalar. Es decir, la suma de dos elementos cualesquiera pertenece al conjunto y lo mismo el resultado de multiplicar un elemento del conjunto por un escalar cualquiera. Eso implica que si* D *es un conjunto no vacío de vectores, gen*(D) *es un sub-EV.*

Demostración formal: ejercicio.

290. ***Ejemplo*** *Todos los* $\mathbb{R}^n$ *son sub-EV de* $\mathbb{R}^{n+m}$. *Para poder decir eso hay que sobreentender que los elementos de* $\mathbb{R}^n$ *no pertenecen a* $\mathbb{R}^{n+m}$, *pero que uno los asimila a elementos de* $\mathbb{R}^{n+m}$ *con cero en las coordenadas faltantes. Así, el elemento* $(1, 2, 3)$ *de* $\mathbb{R}^3$ *se asimila a* $(1, 2, 3, 0, 0)$ $\mathbb{R}^5$.

291. ***Contraejemplo.*** *Una línea que no pasa por el origen no es un sub-EV de* $\mathbb{R}^3$.

Demostración. Una línea que no pasa por el origen tiene como ecuaciones paramétricas $x_i = k_i\lambda + b_i$ con $\lambda \in \mathbb{R}$, pero al menos algún $b_i \neq 0$.

Si multiplicamos tal elemento por 2, el resultado se sale de la línea. En efecto, al multiplicar por 2 nos queda:

$x_i = 2k_i\lambda + 2b_i = k_i(2\lambda) + 2b_i = k_i(\mu) + 2b_i$ pero este nuevo elemento no pertenece a la línea original, pues el término independiente debe ser b_i pero es $2b_i$ y no hay forma de arreglar el desperfecto en la coordenada donde el b_i no es cero.

292. ***Ejercicio*** *Pruebe que un plano en* $\mathbb{R}^3$ *es un EV de* $\mathbb{R}^3$ *ssi pasa por el origen.*

4.3. Bases

Cuando un conjunto genera todo el espacio y lo hace sin redundancia, tenemos una base.

293. ♣ ***Definición.*** *Un conjunto de vectores* B *que es LI y que genera un EV* V *se dice que es una* ***base*** *de* V.

Algunos ejemplos aparecen en la siguiente gráfica.

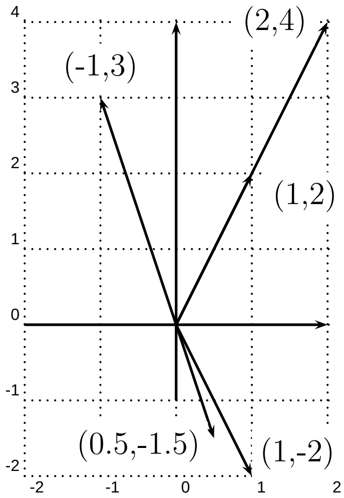


Figura 4.2. No-bases en el plano:
$\{(1,2)\},\{(1,2),(2,4)\},\{(1,2),(2,4),(-1,3)\},\{(-1,3),(0{,}5,-1{,}5)\},\{(1,2),(-1,3),(1-,2)\}$.
Bases : $\{(1,2),(-1,3)\},\{(1,2),(0{,}5,-1{,}5)\},\{(-1,3),(0{,}5,-1{,}5)\},\{(2,4),(1,-2)\}$.

Otros ejemplos:

a) $\{\vec{i},\vec{j}\}$ es una base de $\mathbb{R}^2$. Esta es la **base natural** de $\mathbb{R}^2$. A la base natural también se la llama **canónica**. Esta definición se extiende de forma natural a todos los $\mathbb{R}^n$

b) $\{\vec{i},2\vec{i}\}$ no es una base de $\mathbb{R}^2$.

c) $\{\vec{i},\vec{i}+\vec{j}\}$ es una base de $\mathbb{R}^2$.

d) $\{\vec{i},\vec{j},\vec{i}+\vec{j}\}$ no es una base de $\mathbb{R}^2$.

e) $\{\vec{i},\vec{j}\,\vec{k}\}$ es una base de $\mathbb{R}^3$. Esta es la base natural de $\mathbb{R}^3$.

294. ***Ejemplo*** *Demostremos usando la definición de base que el conjunto* $B=\{(1,0,0),(0,1,0),(0,0,1)\}$ *es una base de* $\mathbb{R}^3$.

Demostración. Tenemos que demostrar que B es LI y que genera todo el espacio, es decir, que $\operatorname{gen}(B)=\mathbb{R}^3$.

Para demostrar que B es LI tomemos una combinación lineal arbitraria de B que dé cero:

$\alpha(1,0,0)+\beta(0,1,0)+\gamma(0,0,1)=(0,0,0)$

Eso es lo mismo que decir que

$(\alpha,\beta,\gamma)=(0,0,0)$.

Ahora bien, dos vectores son iguales cuando sus **componentes** son iguales. Eso da el sistema de ecuaciones

$$\begin{cases}\alpha=0\\ \beta=0\\ \gamma=0\end{cases}$$

que dice que la única forma de hallar una combinación nula con los elementos de B es anulándolos a todos. Es decir, B es LI.

Para demostrar que $\text{gen}(B) = \mathbb{R}^3$ hay que demostrar que todo elemento de $\text{gen}(B)$ está en $\mathbb{R}^3$, y que todo elemento de $\mathbb{R}^3$ está en $\text{gen}(B)$.

Demostremos que todo elemento de $\text{gen}(B)$ está en $\mathbb{R}^3$: un elemento de $\text{gen}(B)$ es una combinación lineal de elementos de B que también son elementos de $\mathbb{R}^3$. Como $\mathbb{R}^3$ es EV, es cerrado bajo la suma y bajo la multiplicación por escalares y por lo tanto es cerrado bajo combinaciones lineades. Por esto todo elemento de $\text{gen}(B)$ está en $\mathbb{R}^3$.

Ahora demostremos que todo elemento de $\mathbb{R}^3$ está en $\text{gen}(B)$. Sea (x, y, z) un elemento arbitrario de $\mathbb{R}^3$, tenemos:

$(x, y, z) = x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1)$

lo cual dice que (x, y, z) es una combinación lineal de los elementos de B, es decir, $(x, y, z) \in \text{gen}(B)$. Hemos demostrado que B no tiene elementos redundantes y que genera todo el espacio $\mathbb{R}^3$, por lo tanto, es una base para dicho espacio.

295. $\boxed{\textbf{\textit{Ejercicio}}}$ *Formule la base natural de $\mathbb{R}^n$ y demuestre formalmente que es una base.*

296. ***Ejemplo*** *Demostremos usando el determinante que $B = \{(1, 0, 1), (2, 1, 0), (0, 3, 1)\}$ es una base de $\mathbb{R}^3$.*

Demostración. El conjunto B define un plp cuyo volumen lo da el valor absoluto del determinante:

$$det(B) = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 1 & 0 & 1 \end{pmatrix} = (1)det\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} - (2)det\begin{pmatrix} 0 & 3 \\ 1 & 1 \end{pmatrix} = 1 - (2)(-3) = 7$$

Como el volumen es 7, el conjunto es LI, pues todos los 3 vectores son esenciales para expandir un volumen.

Probemos ahora que $\text{gen}(B) = \mathbb{R}^3$. Para ello, hay que demostrar que todo elemento de $\text{gen}(B)$ lo es de $\mathbb{R}^3$ y que todo elemento de $\mathbb{R}^3$ lo es de $\text{gen}(B)$. Como todos los elementos de B están en $\mathbb{R}^3$, el cual es EV, cerrado bajo las combinaciones lineales, todo elemento de $\text{gen}(B)$ está en $\mathbb{R}^3$. Probemos ahora que todo elemento de $\mathbb{R}^3$ lo es de $\text{gen}(B)$. Sea $(x, y, z) \in \mathbb{R}^3$. Debemos demostrar que (x, y, z) se puede expresar como una combinación lineal de los elementos de B. Es decir, que existen escalares α, β, γ tales que

$(x, y, z) = \alpha(1, 0, 1) + \beta(2, 1, 0) + \gamma(0, 3, 1)$

lo cual da un sistema de 3 ecuaciones con 3 incógnitas α, β, γ:

$$\begin{cases} x = \alpha + 2\beta \\ y = \beta + 3\gamma \\ z = \alpha + \gamma \end{cases}$$

Este sistema define un conjunto de 3 planos en el espacio α, β, γ. Los vectores normales a dichos planos forman un plp $[(1, 2, 0), (0, 1, 3), (1, 0, 1)]$. El determinante de ese plp es 7. Por lo tanto, los planos se cortan en un único punto. Eso quiere decir que (x, y, z), no importa cuál, siempre se puede descomponer de manera única en B.

Si además, uno quisiera decir exactamente el valor de los escalares, entonces podemos aplicar la regla de Kramer:

Sea

$$M = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 1 & 0 & 1 \end{pmatrix}$$

$$M_x = \begin{pmatrix} x & 2 & 0 \\ y & 1 & 3 \\ z & 0 & 1 \end{pmatrix}$$

$$M_y = \begin{pmatrix} 1 & x & 0 \\ 0 & y & 3 \\ 1 & z & 1 \end{pmatrix}$$

$$M_z = \begin{pmatrix} 1 & 2 & x \\ 0 & 1 & y \\ 1 & 0 & z \end{pmatrix}$$

Tenemos que

$$x = \frac{det(M_x)}{det(M)} \qquad y = \frac{det(M_y)}{det(M)} \qquad z = \frac{det(M_z)}{det(M)}. \blacksquare$$

297. $\boxed{\textbf{\textit{Ejercicio}}}$ *Use el determinante para averiguar si el conjunto* $B = \{(-1, 0, -1), (2, -1, 0), (0, -3, 1)\}$ *es una base de* $\mathbb{R}^3$.

298. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que los pulsos* P_i *forman una base del espacio digital* D_n.

299. $\diamond$ ***Teorema.*** *Si en un EV alguna base tiene un número finito de elementos, entonces todas las bases tienen el mismo número de elementos.*

Demostración. Consideremos sólo $\mathbb{R}^2$:

Sea B una base de $\mathbb{R}^2$. B no puede ser vacía, puesto que el espacio generado por una base vacía es el vacío y $\mathbb{R}^2$ no es vacío.

Por lo tanto, B tiene al menos un elemento, sea (a, b). Ahora bien, ese elemento no es base de $\mathbb{R}^2$. En efecto, el vector (a, b) genera una línea, la cual pasa por el origen pero no es todo $\mathbb{R}^2$. Para más claridad, el elemento $(a, b + 1)$ está en $\mathbb{R}^2$ y sin embargo no es combinación lineal de $\{(a, b)\}$ puesto que las únicas combinaciones lineales de dicho conjunto son los alargamientos: $\{(\lambda a, \lambda b)\}$, pero esos alargamientos forman una línea que no pasa por $(a, b+1)$, pues si pasaran por ese punto, entonces $(a, b+1) = \{(\lambda a, \lambda b)\}$ para algún λ.

Eso implica que, coordenada por coordenada, se tiene:

$a = \lambda a$

$b + 1 = \lambda b$

de la primera ecuación $\lambda = 1$ y reemplazando en la segunda se deduce que $1 = 0$.

Por ende, está mal imaginarse que una línea expanda todo el plano. O sea que B debe tener más de 1 elemento.

Sea $B = \{\vec{u} = (u_1, u_2), \vec{v} = (v_1, v_2)\}$. Entonces, cualquier elemento del plano puede expresarse como combinación lineal de los elementos de B. Sea (z, w) un elemento cualquiera del plano. Debemos probar que siempre se pueden encontrar escalares α, β tales que

$(z, w) = \alpha(u_1, u_2) + \beta(v_1, v_2)$

Esto es equivalente a un sistema de ecuaciones 2×2:

$z = \alpha u_1 + \beta v_1$

$w = \alpha u_2 + \beta v_2$

Las incógnitas son α, β, podemos cambiar dichas incógnitas por x, y:

$z = xu_1 + yv_1$

$w = xu_2 + yv_2$

Reordenando:

$yv_1 = z - xu_1$

$yv_2 = w - xu_2$

despejando, siempre y cuando v_1 y v_2 sean diferentes de cero,

$y = (z - xu_1)/v_1$

$y = (w - xu_2)/v_2$

Revisemos las pendientes de estas líneas para saber sin son paralelas o no. La pendiente de la primera recta es $-u_1/v_1$ y la de la segunda es $-u_2/v_2$. Si estas rectas fuesen paralelas, sus pendientes sería iguales: $-u_1/v_1 = -u_2/v_2$. Que es lo mismo que decir $u_1/u_2 = v_1/v_2$. Pero esto es lo mismo que decir que $u_1 = \lambda v_1$ y que $u_2 = \lambda v_2$ que escrito en forma vectorial se lee

$(u_1, u_2) = \lambda(v_1, v_2)$

que es lo mismo que asegurar que $\vec{u}$ y $\vec{v}$ forman un conjunto LD, lo cual no puede ser, pues estos vectores forman una base, que es un conjunto LI.

Por lo tanto, las dos líneas no son paralelas y se cortan en un único punto: podemos hallar α y β, los números que nos dicen cómo combinar los dos vectores para generar todo el plano. Esos escalares se denominan las **coordenadas** del vector (z, w) en la base dada B. Por consiguiente, ya demostramos que los dos primeros elementos de una base cualquiera en $\mathbb{R}^2$ son LI y expanden todo el plano. Por lo tanto, forman una base.

Como los dos primeros elementos de la base B generan todo el espacio, entonces B no tiene más elementos, pues si tuviese uno más, B no sería LI. En efecto. Si $\vec{w}$ es otro elemento aparte de los dos primeros, entonces, $\vec{w}$ estaría generado por los dos primeros elementos de B. O sea que esos tres elementos generan el mismo espacio que los dos primeros. Es decir, el tercero sobraría: B sería LD y no sería base.

Hemos demostrado que una base cualquiera en $\mathbb{R}^2$ tiene exactamente dos elementos y por consiguiente que todas las bases en $\mathbb{R}^2$ tienen el mismo número de elementos. ■

300. $\boxed{\textbf{\textit{Ejercicio}}}$ *Generalice el ejemplo anterior a* $\mathbb{R}^n$.

301. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que un n-plp tiene determinante diferente de cero ssi las aristas fundamentales del plp forman una base para* $\mathbb{R}^n$. *Demuestre que el determinante de una base nunca es cero.*

302. ♣ ***Definición.*** *Se denomina **dimensión** de un espacio vectorial no nulo al número de elementos de cualquier base de dicho espacio. Si un EV tiene un único elemento que es el elemento neutro de la suma y que se nota $\vec{0}$, entonces no tiene subconjuntos LI y no puede tener base.*

303. ***Ejemplo*** *$\mathbb{R}^n$ tiene dimensión n.*

304. ***Ejercicio*** *Halle la dimensión del espacio digital D_n.*

305. ***Ejercicio*** *Halle la dimensión del EV formado por todos los polinomios de grado menor o igual que 3. Generalice a grado n.*

306. ***Problema.*** *Nosotros ya sabemos cómo hallar un polinomio que pasa por determinados puntos, digamos 6. Tomamos un polinomio de grado 5 pero con coeficientes indeterminados, que son las incógnitas que hay que determinar. Para ello, se reemplaza cada punto en la ecuación del polinomio, de donde salen 6 ecuaciones con 6 incógnitas. Usando esta idea, podemos ver que uno puede aproximar una función cualquiera por polinomios tanto como se desee. Por favor, demuéstrelo o invente contraejemplos.*

307. ***Ejercicio de investigación sobre polinomios de Taylor*** *Investigue en los libros de cálculo la siguiente idea para aproximar funciones derivables por polinomios. La primera derivada indica la pendiente. La segunda, la concavidad. Siguiendo con derivadas de orden superior, uno puede dar cada vez más información sobre la gráfica de la curva. La idea propuesta por Taylor es la de darse cuenta de que, dadas las n primeras derivadas de una función, siempre se puede encontrar un polinomio con las mismas derivadas y resulta que el polinomio dado se parece a la función tanto más cuanto mayor número de derivadas se involucre. El polinomio resultante se le llama polinomio de Taylor. Una aplicación inmediata es esta: en vez de atiborrar los computadores con tablas infinitas sobre los valores de las funciones, digamos, seno, coseno, exponencial, lo que se hace es grabar en memoria el polinomio asociado y calcularlo para cada valor deseado. Por supuesto, hay que decidir el grado del polinomio. La clave es que mientras más alto el grado del polinomio, más precisión. Por eso, se fija la precisión deseada, por ejemplo, para calculadoras es usual que sea de 6 cifras decimales, y en consecuencia se ajusta el grado del polinomio.*

4.4. Planos y sub-EV

Hemos visto la definición geométrica de plano: es un conjunto de puntos tales que cualquier segmento contenido en dicho conjunto es perpendicular a un vector fijo dado, llamado el vector normal del plano. Pudimos demostrar que un plano cumple la ecuación $ax + by + cz = d$. Ahora vamos a dedicar un poco de atención a la siguiente observación: el plano por excelencia $\mathbb{R}^2$ es generado por el conjunto LI $\{\vec{i}, \vec{j}\}$. Podemos generalizar esta situación:

308. ⋄ ***Teorema.*** *Un plano que pasa por el origen en* $\mathbb{R}^3$ *es un subespacio vectorial de dimensión dos, es decir, es generado por un conjunto LI de dos vectores.*

Demostración: ejercicio.

309. ***Ejemplo*** $z = 3x + 5y$ *es la ecuación de un plano* Π *en* $\mathbb{R}^3$ *que pasa por el origen que es un subespacio de dimensión 2. Para verlo, comencemos verificando que ese plano es un subespacio. Veamos que es cerrado a la suma y a la multiplicación escalar:*

Sean dos vectores que estén en Π*:* $(x_1, y_1, 3x_1 + 5y_1)$ *y* $(x_2, y_2, 3x_2 + 5y_2)$*. Sumamos y obtenemos*

$(x_1 + x_2, y_1 + y_2, 3x_1 + 5y_1 + 3x_2 + 5y_2) = (x_1 + x_2, y_1 + y_2, 3(x_1 + x_2) + 5(y_1 + y_2))$

Si denominamos $x_3 = x_1 + x_2$ *y* $y_3 = y_1 + y_2$*, la última expresión adquiere la forma* $(x_3, y_3, 3x_3 + 5y_3)$ *por lo que* z*, siendo la tercera coordenada, es* $z = 3x_3 + 5y_3$*, que denota un elemento del plano* $z = 3x + 5y$*. Vemos que la suma de dos vectores arbitrarios de* Π *también está en* Π*, por lo que este plano es cerrado a la suma.*

Veamos ahora que Π *es cerrado a la multiplicación escalar. Tomemos un elemento de* Π*,* $(x, y, 3x + 5y)$*. Si lo multiplicamos por* λ *obtenemos:* $\lambda(x, y, 3x + 5y) = (\lambda x, \lambda y, \lambda(3x + 5y)) = (\lambda x, \lambda y, 3\lambda x + 5\lambda y))$*. Si denotamos* λx *como* v *y* λy *como* w*, la última expresión se convierte en* $(v, w, 3v + 5w)$ *por lo que la tercera coordenada vale* $z = 3v + 5w$*, lo cual dice que la multiplicación escalar de un elemento de* Π *también es un elemento de* Π*.*

Como Π *es cerrado a la suma y a la multiplicación escalar y es un subconjunto no vacío de un EV, entonces* Π *es un subespacio vectorial.*

Verifiquemos ahora que Π *tiene dimensión dos. Sea* (x, y, z) *que satisface la ecuación de* Π*. Tenemos que* $(x, y, z) = (x, y, 3x + 5y) = x(1, 0, 3) + y(0, 1, 5)$ *es decir, esa figura está generada por el conjunto de vectores* $B = \{(1, 0, 3), (0, 1, 5)\}$ *que es LI pues el uno no es alargamiento del otro. Hemos demostrado que este plano tiene dimensión dos, pues tiene una base de dos elementos.*

310. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que* $x = 0$ *tiene dimensión* 0 *en* $\mathbb{R}$*, dimensión* 1 *en* $\mathbb{R}^2$ *y dimensión* 2 *en* $\mathbb{R}^3$*.*

311. ⋄ ***Teorema.*** *Un plano (que no pasa por el origen) es el resultado de trasladar un subespacio vectorial de dos dimensiones que pasa por el origen.*

312. ***Ejemplo*** $z = 3x + 5y + 8$ *es un plano que es el resultado de trasladar 8 unidades por el eje* Z *al subespacio que cumple la ecuación:* $z = 3x + 5y$*. En efecto: si un punto* (x, y, z) *está en el plano original, entonces*

$(x, y, z) = (x, y, 3x + 5y + 8) = x(1, 0, 3) + y(0, 1, 5) + (0, 0, 8)$

que es el resultado de trasladar el plano que pasa por el origen descrito por

$(x, y, z) = (x, y, 3x + 5y) = x(1, 0, 3) + y(0, 1, 5)$

una distancia por el eje Z *de 8 unidades.*

313. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que las siguientes ecuaciones describen planos que son la traslación de un plano generado por* 2 *vectores que forman un conjunto* LI*:*

a) $z = 2x - 5y$

b) $z = 2x - 5y - 7$

c) $2x - 3y + 4z = 1$

d) $3x - 5y + 2z = 3$

e) $x = 0$

f) $y = 0$

g) $y = 3$.

314. ***Ejemplo*** *Demostremos que el conjunto* $B = \{(2, 1, 1), (-1, 2, -3)\}$ *es una base de* Π *el subespacio de* $\mathbb{R}^3$ *dado por* $x - y - z = 0$ *o bien* $z = x - y$.

Demostración. Primero que todo, ese plano pasa por el origen por lo que es un subespacio de $\mathbb{R}^3$. En segundo lugar, los vectores satisfacen la ecuación del plano, por lo que están en el plano. En tercer lugar, los dos vectores forman un conjunto LI porque el uno no es múltiplo escalar del otro. En cuarto lugar, los dos vectores generan todo el plano. Para verlo, tomemos un vector arbitrario de Π, sea $(x, y, x - y)$. Lo que tenemos que hacer es demostrar que está en el gen(B), es decir, que existen escalares α y β tales que

$(x, y, x - y) = \alpha(2, 1, 1) + \beta(-1, 2, -3)$

Obtenemos un sistema de ecuaciones dado por

$$\begin{cases} x = 2\alpha - \beta \\ y = \alpha + 2\beta \\ x - y = \alpha - 3\beta \end{cases}$$

La primera preocupación es que tenemos un sistema con 2 incógnitas pero con 3 ecuaciones. La única forma de que eso tenga sentido es que una de las ecuaciones sea redundante. Lo bonito es que eso se ve inmediatamente, pues la tercera ecuación es la primera menos la segunda. Como la tercera ecuación es redundante, podemos olvidarla y nos queda un sistema de dos ecuaciones con dos incógnitas:

$$\begin{cases} 2\alpha - \beta = x \\ \alpha + 2\beta = y \end{cases}$$

Si multiplicamos la segunda ecuación por 2 y restamos, obtenemos $-5\beta = x - 2y$. Es decir $\beta = -x/5 + 2y/5$. Sustituyendo este valor en la primera ecuación, obtenemos $2\alpha - (-x/5 + 2y/5) = x$, es decir, $2\alpha = x - x/5 + 2y/5 = 4x/5 + 2y/5$. Lo cual nos da $\alpha = 2x/5 + y/5$. Al final, pudimos demostrar que todo elemento del plano es combinación lineal única de B. Por lo tanto, B genera todo el plano. Concluimos que B es una base del plano Π. **A las constantes α y β las llamamos las coordenadas de (x, y, z) en la base B.**

Hay tres lugares en la demostración que ofrecen retroalimentación para saber si uno va bien o mal. El primero es que los elementos de B deben satisfacer la ecuación del plano. Cierto. El segundo es que debe haber una ecuación redundante. Cierto. Y el tercero es que los valores de α y β hallados deben satisfacer el último sistema de ecuaciones. Veamos si es cierto: $\alpha = 2x/5 + y/5$, $\beta = -x/5 + 2y/5$ dan que $2\alpha - \beta = 2(2x/5 + y/5) - (-x/5 + 2y/5) = 4x/5 + 2y/5 + x/5 - 2x/5 = 5x/5 = x$.

En la segunda ecuación tenemos:

$\alpha + 2\beta y = 2x/5 + y/5 + 2(-x/5 + 2y/5) = 2x/5 + y/5 - 2x/5 + 4y/5 = 5y/5 = y$.

Tal y como debe ser. Uno se demora un poquito más, pero queda seguro de haber hecho las cosas bien. ■

315. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que el conjunto $B = \{(1, 2, -3), (0, 1, -1)\}$ es una base del subespacio de $\mathbb{R}^3$ dado por $x + y + z = 0$.*

Todo plano en el espacio tiene un plano paralelo que pasa por el origen y que es subespacio vectorial generado por dos vectores. Pero el converso también es válido:

316. ⋄ ***Teorema.*** *El espacio generado por un conjunto de dos vectores LI en $\mathbb{R}^3$ es de la forma $ax + by + cz = 0$.*

Demostración. Un plano es el conjunto generado por 2 vectores que forman un conjunto LI. Sean esos vectores $(a_1, a_2, a_3), (b_1, b_2, b_3)$. Sea Π el plano generado por esos dos vectores. Entonces $(x, y, z) \in \Pi$ ssi $(x, y, z) = \alpha(a_1, a_2, a_3) + \beta(b_1, b_2, b_3)$. Coordenada por coordenada tenemos:

$x = \alpha a_1 + \beta b_1$
$y = \alpha a_2 + \beta b_2$
$z = \alpha a_3 + \beta b_3$

Multiplicamos cada ecuación por separado por un escalar tal que en cada ecuación el coeficiente de α sea el mismo:

$a_2 a_3 x = \alpha a_1 a_2 a_3 + \beta b_1 a_2 a_3$
$a_1 a_3 y = \alpha a_2 a_1 a_3 + \beta b_2 a_1 a_3$
$a_1 a_2 z = \alpha a_3 a_1 a_2 + \beta b_3 a_1 a_2$

Restamos la segunda de la primera:

$a_2 a_3 x - a_1 a_3 y = \beta b_1 a_2 a_3 - \beta b_2 a_1 a_3 = \beta(b_1 a_2 a_3 - b_2 a_1 a_3)$

Restamos la tercera de la primera:

$a_2 a_3 x - a_1 a_2 z = \beta b_1 a_2 a_3 - \beta b_3 a_1 a_2 = \beta(b_1 a_2 a_3 - b_3 a_1 a_2)$

Despejamos β de ambas ecuaciones e igualamos:

$\beta = (a_2 a_3 x - a_1 a_3 y)/(b_1 a_2 a_3 - b_2 a_1 a_3) = (a_2 a_3 x - a_1 a_2 z)/(b_1 a_2 a_3 - b_3 a_1 a_2)$

Ahora notamos que esa expresión complicada puede reordenarse para que quede de la forma $ax + by + cz = 0$. ■

317. $\boxed{\textbf{\textit{Ejercicio}}}$ *Revise detalladamente la demostración del teorema anterior y descubra que hay un paso que a veces se puede hacer y que a veces no. Invente un ejemplo para el cual la demostración funcione y un contraejemplo que ilustre que la demostración está incompleta. Añada lo que falte a la demostración para que sea completamente general. Si logra hacerlo, permítanos felicitarlo(a) muy calurosamente.*

318. ⋄ ***Teorema.*** *En $\mathbb{R}^3$ todo conjunto Π que contiene al punto $\vec{P}$ y que es paralelo a otro generado por un conjunto LI de dos vectores cumple con la propiedad: $X \in \Pi$ ssi $(\vec{X} - \vec{P}) \cdot \vec{N} = 0$ para algún vector $\vec{N}$ llamado el vector normal al plano (en la notación hemos usado la dualidad punto-vector).*

319. **Ejercicio** *Haga un dibujo que ilustre el teorema anterior y demuéstrelo. El teorema dice que la definición algebraica del plano coincide con su definición geométrica que conocemos de antaño.*

320. **Ejercicio:** *Resuelva los siguientes sistemas de ecuaciones y caracterice la cantidad de soluciones en términos geométricos:*

a) $x + y + z = 1$, $x + 2y - 3z = 2$, $-x + 2y - z = 0$,

b) $2x + 3y - z = 1$, $x - y - 2z = 5$, $3x + 2y - 3z = 3$,

c) $2x + 3y - z = 1$, $x - y - 2z = 5$, $3x + 2y - 3z = 6$,

d) $x + y + z = 1$, $2x + 2y + 2z = 2$, $3x + 3y + 3z = 3$.

4.5. Espacios de matrices

Tenemos una idea en mente y es armar la infraestructura necesaria para tratar todo sistema lineal de ecuaciones como una generalización de la ecuación para números reales $ax = b$. Nosotros obtendremos $A\vec{X} = \vec{B}$ donde A es una matriz, $\vec{X}$ y $\vec{B}$ denotan vectores. Para poder hacer esto, necesitamos considerar, entre otras cosas, que las matrices también se suman y restan y que se pueden multiplicar por un escalar, es decir, que las matrices forman EV. Al igual que hay muchos $\mathbb{R}^n$, también hay muchos espacios de matrices. En general, definimos $M_{m\times n}$ como el espacio de matrices con m filas y n columnas, tal que las matrices se pueden sumar entrada por entrada y, similarmente, que la multiplicación de una matriz por un escalar es la multiplicación de todas las entradas de la matriz por el escalar.

321. **Ejemplo** *$M_{2\times 2}$ es el espacio de matrices de dos renglones y dos columnas con entradas reales y con la suma y multiplicación escalar definidas entrada por entrada. Formalmente:*

$$M_{2\times 2} = \left\{ \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \ \textit{donde} \ a, b, c, d \in \mathbb{R} \right\}$$

Si uno tiene dos matrices A y B de $M_{2\times 2}$ definidas por

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad\qquad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

la suma es, por definición,

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}$$

y la multiplicación de A por un escalar λ es

$$\lambda A = \lambda \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{pmatrix}$$

Con estas operaciones podemos sumar y alargar sin problemas: $M_{2\times 2}$ es un EV.

$M_{2\times 2}$ tiene dimensión 4. Esa aseveración sale, abusando de la gentileza del lector, de la siguiente igualdad:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix}$$

$M_{2\times 2}$ tiene muchos subespacios vectoriales. El siguiente es el subespacio de las matrices triangulares superiores:

$$S_{2\times 2} = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \text{ donde } a, b, d \in \mathbb{R} \right\}$$

S hereda de $M_{2\times 2}$ la suma y la multiplicación escalar bajo las cuales S es cerrado. Por lo tanto, S es un subespacio vectorial de $M_{2\times 2}$. La dimensión de S es 3.

322. **_Ejercicio_** *En $M_{2\times 2}$ defina el conjunto de las matrices triangulares inferiores. Demuestre que es un subespacio y hállele una base.*

4.6. Complemento ortogonal

323. ♣ ***Definición.*** *Decimos que en un espacio vectorial W, los vectores $\vec{u}$ y $\vec{v}$ son* ***ortogonales*** *si su producto interior es cero. Esto generaliza una definición previa para $\mathbb{R}^n$ ver 112 pág 55.*

324. ♣ ***Definición.*** *Decimos que en un espacio vectorial W, los subespacios H y K generan a todo W si cada elemento $\vec{w} \in W$ puede escribirse como combinación de elementos de los dos subespacios: $\vec{w} = \vec{h} + \vec{k}$ con $\vec{h} \in H$ y $\vec{k} \in K$. Esta definición se generaliza a cualquier número finito de subespacios.*

325. ♣ ***Definición.*** *Decimos que en un espacio vectorial W, $H^\perp$ es el* ***complemento ortogonal*** *de H, si tanto $H^\perp$ como H son subespacios vectoriales de W y si además a) entre $H^\perp$ y H generan a todo V, y b) si todos los elementos del uno son ortogonales a todos los elementos del otro.*

326. **_Ejemplo_** *En el plano, tomamos como H el eje X y como $H^\perp$ el eje Y.*

327. **_Ejemplo_** *La descomposición dada por subespacios a veces es única y a veces no:*

a) Los ejes coordenados X, Y, Z son subespacios, generan todo $\mathbb{R}^3$, los subespacios X, Y general el piso, el piso y Z son complementos ortogonales uno del otro y cada elemento de $\mathbb{R}^3$ se descompone de manera única en piso y Z.

b) X, Z generan una pared, la cual no es complemento ortogonal del piso, pero entre el piso y esa pared generan todo el espacio $\mathbb{R}^3$ pero no de manera única.

328. $\boxed{\textbf{\textit{Ejercicio}}}$ *En $M_{2\times 2}$ definimos el producto interior entre dos matrices de la siguiente forma. Si*

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

entonces definimos ***el producto interno, interior o escalar de*** A ***y*** B *como*

$$A \cdot B = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}$$

a) Demuestre que en realidad nuestra definición cumple con los axiomas de producto interno.

b) Halle el complemento ortogonal del subespacio de las matrices triangulares superiores. Demuestre que cada elemento del espacio grande se descompone de manera única entre el subespacio y su complemento.

329. $\boxed{\textbf{\textit{Ejercicio}}}$ *En $\mathbb{R}^3$, halle el complemento ortogonal de los subespacios señalados y construya una base de todo el espacio uniendo bases del subespacio y de su complemento.*
a) El espacio vectorial que es paralelo a la línea que pasa por $(1, 2, 3)$ y $(4, 5, 6)$.
b) El plano que pasa por el origen y es paralelo a $2x - 4y + 5z = 8$.

4.7. Ejercicios de repaso

1. Determine si $\mathbb{R}^2$, con las operaciones suma y producto por escalar definidas por $(x, y) + (z, w) = (x + z, y + w)$ y $k(x, y) = (2kx, 2ky)$ es un espacio vectorial.

2. Demuestre que si $\{v, w\}$ es un conjunto linealmente independiente en un espacio vectorial V y $u \notin \text{gen}\{v, w\}$ entonces $\{v, w, u\}$ es un conjunto linealmente independiente.

3. Sean $V = \mathbb{R}^3$ y $W = \{(x, y, z); x + 2y + 3z = 0\}$.

   a) Encuentre dos vectores $\vec{v_1}$ y $\vec{v_2}$, que generen a todo W.

   b) ¿Está el vector $\vec{v} = (-1, -1, 1)$ en el gen$\{\vec{v_1}, \vec{v_2}\}$?

4. Sea D el conjunto de matrices de la forma $\begin{pmatrix} 0 & a \\ a & a+b \end{pmatrix}$

   a) Demuestre que D es un subespacio de las matrices de tamaño 2×2.

   b) Halle una *base* para D y determine su *dimensión*.

   c) Halle el complemento ortogonal de D en $M_{2\times 2}$ y muestre cómo se descompone dicho espacio en D y su complemento.

5. Un polinomio de grado n es una función de $\mathbb{R}$ en $\mathbb{R}$ de la forma

$$p(x) = a_0 + a_1x + a_2x^2 + ... + a_nx^n$$

donde todos los a_i son números reales y x es una variable que se instancia con números reales. Sea $P_{\leq n}$ el conjunto de los polinomios de grado menor o igual a n. $P_{\leq n}$ tiene una noción de igualdad, una suma y una mutiplicación por un escalar definidas grado por grado. Formalice matemáticamente estas definiciones. Demuestre que $P_{\leq n}$ es un EV con estas operaciones.

6. ¿El polinomio $1 + x + x^2$ está en el generado por $B = \{1 - x, 1 + x, 1 + x^2\}$?

7. Sean $W = \{p \in P_{\leq 3} \colon p(0) = 0\}$ y $U = \{p \in P_{\leq 3} \colon p(1) = 0\}$. Muestre que W y U son subespacios del espacio $P_{\leq 3}$. Determinar una base para W, una base para U y una base para $W \cap U$, por lo cual hay que probar que este conjunto es un subespacio.

8. Demuestre que el conjunto de todos los polinomios en $P_{\leq n}$ que tienen una tangente horizontal en $x = 0$ es un subespacio de $P_{\leq n}$. Encuentre una base para tal subespacio.

9. Realice lo indicado:

   a) ¿Es el conjunto $\{1 - x, x^2 + 2x - 1, x^2 + 3\}$ una base de $P_{\leq 2}$? Justifique claramente.

   b) En $\mathbb{R}^2$, ¿el conjunto $S = \{(x, y) \in \mathbb{R}^2 \colon \ y = 4 - x^2\}$ es un subespacio vectorial? Justifique claramente.

   c) Sea $C^1[0, 1]$ el conjunto de funciones derivables sobre el intervalo abierto $(0, 1)$ cuya derivada es continua. ¿Es $G = \{f \in C^1[0, 1] \colon f'(x) = xf(x)\}$ un subespacio vectorial de $C^1[0, 1]$? Justifique su respuesta claramente.

10. Sea A una matriz de tamaño $n \times n$ y $\lambda \in \mathbb{R}$. Mostrar que el conjunto $S = \{x : Ax = \lambda x\}$ es un subespacio de $\mathbb{R}^n$. Determine la dimensión de S si $\lambda = 3$ y A es la matriz $A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 0 & -1 \\ 0 & 3 & 4 \end{pmatrix}$.

11. Sea $V = C^1[a, b]$ el conjunto de funciones que van del intervalo $[a, b]$ sobre $\mathbb{R}$ tales que ellas y sus derivadas sean continuas. Sea $H = \{f \in V \mid f(x) + f'(x) = 0\}$. ¿Es H un subespacio vectorial de V?

12. Considere el conjunto $H = \{(x, y, z) \colon (x, y, z) = (t, -3t, t) \colon t \in \mathbb{R}\}$.

    a) Muestre que H es un subespacio vectorial de $\mathbb{R}^3$.

    b) Encuentre una base y la dimensión de H.

    c) Represente geométricamente el espacio H.

4.8. Resumen

Con miras a proveer una plataforma que nos permita pisar sólidamente al atacar problemas en altas dimensiones hemos definido espacio vectorial, dependencia e independencia lineal y bases. El elemento cohesivo ha sido la geometría. Un espacio vectorial es como un plano que pasa por el origen: al combinar vectores, con sumas y multiplicaciones por números, el resultado permanece en el mismo plano. Un conjunto de vectores es linealmente dependiente cuando hay elementos redundantes, es decir, cuando al quitar algún elemento se genera, con combinaciones lineales, el mismo espacio que antes. Un conjunto de vectores es linealmente independiente cuando todos sus elementos son esenciales, o sea que al quitar cualquier vector se genera un espacio menor. Una base es un conjunto no redundante que genera todo el espacio.

CAPÍTULO 5

TL = TRANSFORMACIONES LINEALES

Toda nuestra teoría tiene un objetivo: entender de manera productiva la estructura general de los sistemas lineales de ecuaciones (que son los que siempre hemos visto). Para esto comenzamos introduciendo las matrices, definimos rectas, planos, espacios vectoriales y bases. Ahora definiremos lo que es una TL, transformación lineal. En primera instancia, una TL es una función, correspondencia, asignación o transformación asociada a un sistema lineal de ecuaciones por medio de una matriz. Veremos esto con mucho detalle para después generalizar. Como sucede con las funciones, las TL también pueden componerse, ponerse en serie, una tras de otras, y la consecuencia es que la composición de las TL es otra TL. Por tanto, la composición de las TL genera un producto entre matrices de tal forma que el producto de dos matrices da una matriz que representa la matriz de la TL compuesta.

5.1. Definición de TL

Consideremos el sistema de ecuaciones:

$$\begin{cases} 3x + 4y = 7 \\ 5x + 3y = 9 \end{cases}$$

Reescribamos este sistema usando la siguiente notación:

$$\begin{pmatrix} 3 & 4 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 \\ 9 \end{pmatrix}$$

Asignemos un significado a nuestra nueva notación. Tenemos en primer término la matriz

$$M = \begin{pmatrix} 3 & 4 \\ 5 & 3 \end{pmatrix}$$

Ahora asignaremos a esta matriz una función con propiedades especiales, llamada TL, transformación lineal. Nosotros deseamos que la manera de asignar una TL a una matriz tenga un significado natural, es decir, que esté directamente relacionado con los sistemas de ecuaciones. Por esa razón, la manera natural de asociar por definición una función T, de clase TL, a la matriz M es:

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 5 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x+4y \\ 5x+3y \end{pmatrix}$$

T toma la dupleta $\begin{pmatrix} x \\ y \end{pmatrix}$ y las transforma en la dupleta $\begin{pmatrix} 3x+4y \\ 5x+3y \end{pmatrix}$. Como las dupletas son elementos de $\mathbb{R}^2$, también usamos la siguiente notación, usual para funciones, para indicar cómo es T :

$$T\colon \mathbb{R}^2 \to \mathbb{R}^2$$
$$T((x,y)) = T(x,y) = (3x+4y, 5x+3y).$$

El conjunto sobre el cual opera la TL se denomina **dominio**, y el conjunto sobre el cual caen los resultados de la transformación se denomina **codominio**. Es igual que en funciones, puesto que cada TL es una función. En este ejemplo, el dominio es igual al codominio, y es el conjunto de dupletas, el plano o $\mathbb{R}^2$.

También podemos definir lo que debemos entender por **multiplicar una matriz por un vector**: es lo mismo que considerar la matriz como TL y aplicarla sobre el vector para recuperar el lado izquierdo de un sistema de ecuaciones:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}$$

Lo hecho aquí se extiende de forma natural a cualquier dimensión, es decir, a la multiplicación de cualquier matriz con n columnas por un vector columna con n componentes.

330. $\boxed{\textbf{\textit{Ejercicio}}}$ *Observe la forma como se multiplica una matriz de dos columnas por un vector de dos componentes y descubra dónde y cómo se ha usado el producto punto, interno o interior. Reexprese la regla de multiplicación de una matriz por un vector en términos de dicho producto.*

Ahora podemos decir que resolver el sistema de ecuaciones

$$\begin{cases} 3x+4y = 7 \\ 5x+3y = 9 \end{cases}$$

es lo mismo que exigir encontrar (x,y) tal que $M(x,y) = (7,9) = \vec{B}$, donde M es la matriz o TL asociada al sistema. En general, un sistema de ecuaciones es simplemente $M\vec{X} = \vec{B}$ y se dice que resolverlo es lo mismo que hallar la preimagen de $\vec{B}$ por M. Dicha preimagen es un conjunto que puede tener un único vector, un conjunto de vectores, o puede ser vacío.

Nota de rigor: recalcamos que cuando usamos la notación de funciones para designar una TL, las dupletas son horizontales, (x,y), pero cuando representamos la TL en forma matricial, las dupletas forman un vector columna $\begin{pmatrix} x \\ y \end{pmatrix}$. Lo que es una dupleta horizontal en una notación se convierte en vector columna en la otra. Sin embargo, es conveniente tener presente que en el formalismo que estamos construyendo para matrices, la matriz (x, y) será muy diferente de la matriz $\begin{pmatrix} x \\ y \end{pmatrix}$.

331. $\boxed{\textbf{\textit{Ejercicio}}}$ *Para los sistemas de ecuaciones dados, hallar la matriz asociada y la TL que ésta genera, explicitando el dominio y el codominio:*

a) $\left\{ \begin{array}{l} 8x - 1y = 7 \\ 2x + 3y = 13 \end{array} \right.$

b) $\left\{ \begin{array}{r} -3x + 4y - 4z = 7 \\ 4x + 3y - 2z = 7 \end{array} \right.$

c) $\left\{ \begin{array}{r} -3x + 4y - 4z = 7 \\ 4x + 3y - 2z = 7 \\ -x - 2y + 3z = 4 \end{array} \right.$

332. $\boxed{\textbf{\textit{Ejercicio}}}$ *Hallar la TL asociada a cada una de la siguientes matrices:*

a) $\begin{pmatrix} 3 & -2 \\ -1 & 3 \end{pmatrix}$

b) $\begin{pmatrix} 3 & -2 & 4 \\ -1 & 3 & -7 \end{pmatrix}$

c) $\begin{pmatrix} 3 & -2 & 4 \\ -1 & 3 & -7 \\ 0 & 4 & -5 \end{pmatrix}$

d) $\begin{pmatrix} 3 & -2 \\ -1 & 3 \\ 0 & 4 \end{pmatrix}$

Aunque ya hemos captado la idea de una TL, podemos sacar un gran provecho de la infraestructura que hemos fabricado con EV y demás. Pues bien, un EV es un conjunto en donde se puede sumar y alargar (multiplicación por un escalar). Una transformación lineal es una asignación, o función, que es compatible con la estructura de EV. Respeta la suma y la multiplicación escalar.

Más rigurosamente, si uno tiene dos espacios vectoriales V y W uno puede considerar funciones entre ellos, es decir, maneras de asignar o transformar vectores de V en vectores en W. Para una transformación dada, a V se le llama el dominio, a W se le llama el codominio. La notación usual es $T\colon V \to W$.

333. ♣ ***Definición.*** *Una TL (transformación lineal) $T\colon V \to W$ es una función tal que si $\vec{u}$, $\vec{v}$, $\vec{u}+\vec{v}$ están en V y todas sus imágenes están en W, se tiene que:*

a) $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ y
b) $T(\lambda\vec{u}) = \lambda T(\vec{u})$ para $\lambda \in \mathbb{R}$.

334. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que T es TL ssi $T(\alpha\vec{u}+\beta\vec{v}) = \alpha T(\vec{u}) + \beta T(\vec{v})$ para $\vec{u}$, $\vec{v} \in V$ y α, $\beta \in \mathbb{R}$.*

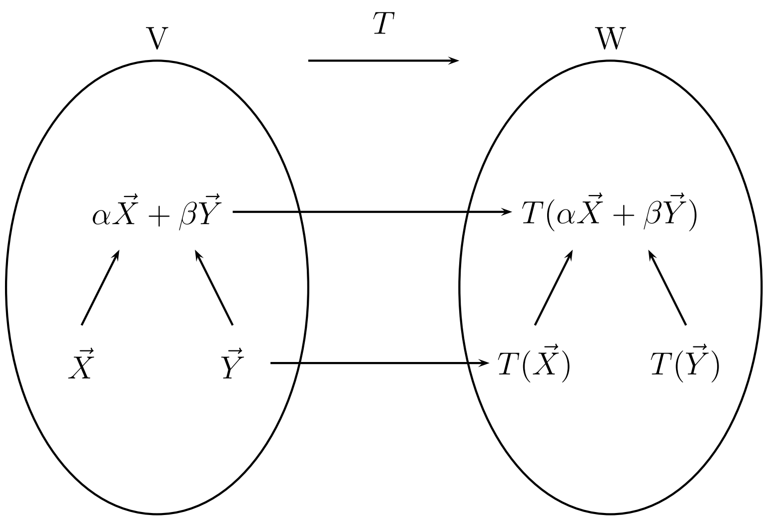


Figura 5.0. Una transformación es lineal cuando da lo mismo primero combinar y después transformar que primero transformar y después combinar.

En otras palabras, T es TL cuando la transformación de una combinación lineal es la combinación lineal de las transformaciones.

335. ***Ejemplo*** $T(\vec{X}) = 8\vec{X}$ *es una TL pues*
$T(\alpha\vec{X}+\beta\vec{Y}) = 8(\alpha\vec{X}+\beta\vec{Y}) = \alpha(8\vec{X}) + \beta(8\vec{Y}) = \alpha T(\vec{X}) + \beta T(\vec{Y})$.

336. ◇ ***Teorema.*** *Cada matriz define una TL. Ilustremos este importante resultado con un ejemplo de una función que ya sabemos que está asociada a una cierta matriz.*

337. ***Ejemplo 1:*** *Verifiquemos que T definida por $T(x,y) = (3x+4y, 4x+3y)$ es TL.*

Necesitamos probar que
$T(\alpha\vec{u} + \beta\vec{v}) = \alpha T(\vec{u}) + \beta T(\vec{v})$
Con las correspondencias obvias, eso es lo mismo que probar que:
$T(\alpha(u_1, u_2) + \beta(v_1, v_2)) = \alpha T(u_1, u_2) + \beta T(v_1, v_2)$
Ahora bien:
$T(u_1, u_2) = (3u_1 + 4u_2, 4u_1 + 3u_2)$
$T(v_1, v_2) = (3v_1 + 4v_2, 4v_1 + 3v_2)$
Por tanto:
$\alpha T(u_1, u_2) = \alpha(3u_1 + 4u_2, 4u_1 + 3u_2) = (3\alpha u_1 + 4\alpha u_2, 4\alpha u_1 + 3\alpha u_2)$
$\beta T(v_1, v_2) = \beta(3v_1 + 4v_2, 4v_1 + 3v_2) = (3\beta v_1 + 4\beta v_2, 4\beta v_1 + 3\beta v_2)$
y al fin:
$\alpha T(u_1, u_2) + \beta T(v_1, v_2)$
$= (3\alpha u_1 + 4\alpha u_2, 4\alpha u_1 + 3\alpha u_2) + (3\beta v_1 + 4\beta v_2, 4\beta v_1 + 3\beta v_2)$
$= (3\alpha u_1 + 3\beta v_1 + 4\alpha u_2 + 4\beta v_2, 4\alpha u_1 + 4\beta v_1 + 3\alpha u_2 + 3\beta v_2)$
$= (3(\alpha u_1 + \beta v_1) + 4(\alpha u_2 + \beta v_2), 4(\alpha u_1 + \beta v_1) + 3(\alpha u_2 + \beta v_2))$
$= T(\alpha u_1 + \beta v_1, \alpha u_2 + \beta v_2)$
$= T((\alpha u_1, \alpha u_2) + (\beta v_1, \beta v_2))$
$= T(\alpha(u_1, u_2) + \beta(v_1, v_2))$.

338. $\boxed{\textbf{\textit{Ejercicio}}}$ *Repita la anterior demostración en notación matricial. Generalice la demostración para que quede claro que en realidad a cada matriz se le asocia de manera natural una TL.*

339. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que las siguientes transformaciones son lineales:*

a) $T(x, y) = (2x + 4y, -3x + 2y)$
b) $T(x, y) = (x - y, 2x - 7y)$
c) $T(x, y) = (-x - y, x + y)$
d) $T(x, y) = (5x - y, -2x + 3y)$
e) $T(x, y, z) = (2x + 3y - z, 2x + 4y - 5z)$
f) La derivada de funciones. Este ejercicio es importante porque la linealidad de la derivada permite definir y tratar los sistemas de ecuaciones diferenciales lineales, que dictaminan la dinámica de muchos sistemas dinámicos, con igual naturalidad y soltura que los sistemas de ecuaciones lineales que estamos viendo. Al hablar de funciones derivables, notamos por $C^n(a, b)$ el conjunto de las funciones cuyo dominio es el intervalo abierto (a, b) y que van sobre $\mathbb{R}$ que tengan n derivadas todas continuas. De igual forma $C^n(\mathbb{R})$ denota el espacio de las funciones que tienen como dominio a todo $\mathbb{R}$ y que tienen n derivadas todas continuas.

340. ***Contraejemplos.*** *Hay funciones que no son TL, como las siguientes:*

$T(x, y) = (x^2, 0)$ no es TL. En efecto:
$T(\lambda x, \lambda y) = ((\lambda x)^2, 0) = \lambda^2(x^2, 0) = \lambda^2 T(x, y) \neq \lambda T(x, y)$ cuando λ diferente de 1 o de 0.

De igual forma $T(x,y)=(x,y+1)$ no es TL pues
$T(\lambda x,\lambda y)=(\lambda x,\lambda y+1)$ lo cual es diferente de
$\lambda T(x,y)=\lambda(x,y+1)=(\lambda x,\lambda y+\lambda)$.

341. **_Ejercicio_** *Formule un test para decidir de una sola mirada si una transformación es TL o no. Ponga a prueba su test sobre varios ejemplos y contraejemplos.*

342. **_Ejercicio_** *En una empresa, una sección toma varios insumos de diferentes tipos y produce varios productos. Para fijar ideas, imaginemos que se tienen 2 insumos A,B y dos productos P,Q. Para hacer un producto tipo P se requieren 2 del tipo A y uno de B. Para hacer un producto tipo Q se requieren 2 de A y 3 de B. Codificar esa información en forma de matriz de tal manera que se responda naturalmente a la siguiente pregunta: si se requieren 3 productos tipo P y 4 tipo Q, ¿cuánto se necesita de cada insumo? Especificar la TL asociada y verbalizar su significado.*

5.2. La matriz de una TL

Nosotros sabemos cómo hallar la matriz asociada a un sistema de ecuaciones lineales, y de ahí sacar una TL. Con todo, uno puede definir una TL sin necesidad de pensar en un sistema de ecuaciones. Por ejemplo, derivar funciones es una TL sobre el espacio de funciones derivables. Con tanta generalidad, pensar en matrices origina serios problemas. Pero en casos muy simples se procede muy sencillamente, por inspección.

343. **_Ejemplo y definición_** *Sea $\vec{X}=(x_1,...,x_n)\in\mathbb{R}^n$. Hallemos la matriz de $I(\vec{X})=\vec{X}$. Puesto que la matriz de I es una matriz tal que a $\vec{X}$ no le hace nada, esa matriz es la matriz identidad, $\mathbb{I}$, la que tiene unos en la diagonal y ceros en las demás entradas. A esta transformación lineal la llamamos la* ***identidad*** *y la notamos I.*

344. **_Ejemplo_** *Calculemos la matriz de $T(x,y)=(3x+4y,4x+3y)$.*
Solución: *se escribe T en notación vertical y se factoriza (x,y) en el sentido de la relación entre los sistemas de ecuaciones y las matrices. Da la matriz*

$$\begin{pmatrix} 3 & 4 \\ 4 & 3 \end{pmatrix}.$$

345. $\diamond$ ***Teorema sobre la caracterización geométrica de una* TLT.** *Para saber cómo funciona una TL sobre $\mathbb{R}^2$ es suficiente saber qué hace ella sobre la base natural.*

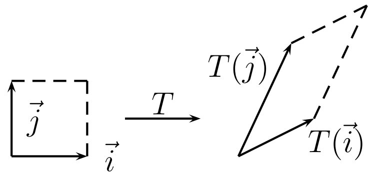


Figura 5.1. Una TL se caracteriza por su efecto sobre la base natural. Podemos pensar en términos de bases o de los plp que ellas generan.

Demostración. Sea T una TL cualquiera de $\mathbb{R}^2$ en $\mathbb{R}^2$.

Tengamos en cuenta que en la base natural del dominio $\mathbb{R}^2$ se tiene que

$$(x, y) = x(1, 0) + y(0, 1)$$

y aplicando T a ambos lados y aplicando la linealidad, obtenemos:

$$T(x, y) = T(x(1, 0) + y(0, 1)) = xT(1, 0) + yT(0, 1).$$

Esta identidad dice que para saber cómo opera una TL es suficiente saber qué hace sobre la base natural o sobre el plp formado por la base natural. ■

346. **_Ejercicio_** *Generalice el teorema anterior a $\mathbb{R}^n$.*

Si reescribimos la identidad $T(x, y) = T(x(1, 0) + y(0, 1)) = xT(1, 0) + yT(0, 1)$ en notación vertical, y la generalizamos a n dimensiones, tenemos el siguiente

347. ***Algoritmo:*** *Para sacar la matriz de una TL T se calcula T de cada uno de los elementos de la base natural $N = \{(1, 0, .., 0), (0, 1, 0, ..., 0), ..., (0, ..., 1)\}$ y dichas imágenes se ponen en posición vertical formando una matriz.*

348. **_Ejemplo_** *Si $T(x, y) = (3x + 4y, 4x + 3y)$ se tiene $T(\vec{i}) = T(1, 0) = (3, 4)$ en tanto que $T(\vec{j}) = T(0, 1) = (4, 3)$ y esos dos vectores resultantes son las columnas de la matriz M de T:*

$$M = \begin{pmatrix} 3 & 4 \\ 4 & 3 \end{pmatrix}$$

349. ***Protesta.*** *Además de la base natural, cada EV tiene un número infinito de bases. ¿Por qué la matriz de una TL tiene que definirse con respecto a la base natural?*

Esta justa protesta exige una respuesta, la cual será estudiada como prerrequisito a la diagonalización.

350. ◇ ***Teorema.*** *Para toda TL T, tenemos $T(\vec{0}) = \vec{0}$.*

Prueba: $T(\vec{0}) = T(\vec{0} + \vec{0}) = T(\vec{0}) + T(\vec{0})$, donde se deduce que $T(\vec{0}) = 2T(\vec{0})$, o $T(\vec{0}) = \vec{0}$ siempre.

Esto significa que cualquier sistema homogéneo de ecuaciones de la forma $MX = 0$ siempre tiene al menos una solución, la solución cero, que también se llama la solución trivial.

351. $\boxed{\textbf{\textit{Ejercicio}}}$ *Describa la cantidad de soluciones de un sistema de ecuaciones en términos de las propiedades de funciones: inyectiva, sobreyectiva. Una función es inyectiva, o 1-1, si a cada vector del codominio le llega máximo un vector del dominio (puede no llegarle nadie). Una función es sobreyectiva si a todos y a cada uno de los elementos del codominio les corresponde alguien en el dominio.*

5.3. Composición de las TL

Una TL es antes que nada una transformación. Las transformaciones pueden tratar de encadenarse o componerse, primero una y después otra. Eso no siempre es posible. Por ejemplo, si una TL toma vectores de un EV de dimensión 6 y produce vectores en un espacio de 5 dimensiones, para poder entrar en una cadena se exige que el próximo eslabón reciba elementos de un espacio de 5 dimensiones.

352. ◇ ***Teorema.*** *Si dos TL pueden componerse, su resultado es una transformación que es TL. Con total claridad, si $S\colon U \to V$ es TL y si $T\colon V \to W$ es TL, entonces la compuesta $T \circ S$ puede hacerse y es TL:*

$$\begin{tikzcd}
U \arrow[r, "S"] \arrow[dr, "T \circ S"'] & V \arrow[d, "T"] \\
 & W
\end{tikzcd}$$

Figura 5.2. La composición o encadenamiento de las TL es TL.

Demostración: ejercicio.

Puesto que la compuesta de las TL es TL, la compuesta tiene una matriz asociada. ¿Cómo hallamos su matriz?

Téngase presente que la notación al componer TL se subordina a la siguiente necesidad: la matriz de una TL, lo mismo que una función opera sobre vectores columna que están a la derecha de ella. Por lo tanto, al componer dos funciones, TL o matrices, la que entra de segunda debe ir a la izquierda para que reciba, por la derecha, lo que la primera produce. Por esta razón, la compuesta del diagrama anterior se denota $T \circ S$ y cuando opera sobre un vector $\vec{X}$ del dominio de S se escribe así:

$$(T \circ S)(\vec{X}) = T(S(\vec{X}))$$

Recordemos que para hallar la matriz de una TL se halla la imagen de cada uno de los vectores de la base natural y los resultados se escriben como columnas formando una matriz. Por tanto, la matriz de una compuesta se halla calculando la imagen de cada uno de los elementos de la base del dominio por la primera TL y el vector resultante se pasa por la segunda TL. Todos los vectores resultantes forman una matriz, la matriz de la compuesta.

353. *Ejemplo* *Dadas dos TL componibles, S y T, hallemos la matriz de la compuesta $T \circ S$:*

Sea $S(x, y) = (2x - 3y, 5x + 4y)$, y sea $T(x, y) = (3x + 2y, -9x + 7y)$
Hallemos la imagen por S de la base natural. Tenemos:
$S(1, 0) = (2, 5), \quad S(0, 1) = (-3, 4)$
y por tanto

$$S = \begin{pmatrix} 2 & -3 \\ 5 & 4 \end{pmatrix}$$

Obsérvese que usamos la misma letra para designar a S tanto como función como matriz. No hay ninguna confusión, pues hay una correspondencia biunívoca entre las dos.

Por otro lado:
$T(1, 0) = (3, -9), \quad T(0, 1) = (2, 7)$
y por tanto:

$$T = \begin{pmatrix} 3 & 2 \\ -9 & 7 \end{pmatrix}$$

Para hallar la matriz de la compuesta simplemente hallamos la imagen de cada uno de los elementos de la base natural por la compuesta:

$$\begin{aligned} T(S(1,0)) &= T(2,5) = T(2(1,0) + 5(0,1)) = 2T(1,0) + 5T(0,1) \\ &= 2(3,-9) + 5(2,7) = (6+10, -18+35) = (16, 17) \\ T(S(0,1)) &= T(-3,4) = T(-3(1,0) + 4T(0,1)) = -3T(1,0) + 4T(0,1) \\ &= -3(3,-9) + 4(2,7) = (-9+8, 27+28) = (-1, 55). \end{aligned}$$

Por lo tanto la matriz de la compuesta $T(S())$ es TS

$$TS = \begin{pmatrix} 16 & -1 \\ 17 & 55 \end{pmatrix}$$

354. *Ejercicio* *Halle $T(S(\vec{X}))$ usando el procedimiento de composición de funciones y por inspección encuentre la matriz asociada. Compare la respuesta con la que acabamos de obtener.*

355. $\boxed{\textbf{\textit{Ejercicio}}}$ *Hemos hallado la matriz asociada a $T \circ S$. Halle la matriz asociada a $S \circ T = S(T())$. Averigüe si las dos composiciones dan la misma matriz. Si lo son, acá y en cualquier otro caso, se dice que la composición de las TL es conmutativa. De lo contrario se dice que la composición de las TL es no conmutativa.*

Ya sabemos encontrar la matriz asociada a cualquier composición de TL. El problema es entonces poder predecir de manera mecánica cuál va a ser la matriz asociada al producto de componer o encadenar dos o más TL sabiendo las matrices componentes.

La solución tiene que ser aquella que convierta la naturalidad en una realidad. En la jerga TRIZ eso se dice así: el mejor diseño es el que no tiene ningún diseño. Revisemos el ejemplo recién resuelto. Hallemos, de una manera ligeramente diferente, la imagen por la compuesta de cada uno de los elementos de la base natural. Primero $T(S(\vec{i}\,))$, sabiendo que:

$$S = \begin{pmatrix} 2 & -3 \\ 5 & 4 \end{pmatrix}$$

$$T = \begin{pmatrix} 3 & 2 \\ -9 & 7 \end{pmatrix}$$

Calculemos la primera columna de la matriz producto TS hallando $S(\vec{i}\,)$ y su resultado pasándolo por T:

$$\begin{aligned} TS(\vec{i}) &= \begin{pmatrix} 3 & 2 \\ -9 & 7 \end{pmatrix} \begin{pmatrix} 2 & -3 \\ 5 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 2 \\ -9 & 7 \end{pmatrix} \begin{pmatrix} 2 \\ 5 \end{pmatrix} \end{aligned}$$

Notemos lo siguiente: el resultado de esta multiplicación es la primera columna de la matriz producto TS. Por tanto, la primera columna de la matriz producto TS tiene en la primera coordenada el resultado de multiplicar, en producto punto, la primera fila o renglón de la matriz T, la cual es $(3, 2)$, por la primera columna de la matriz S que es $(2, 5)$. La primera columna de la matriz producto TS tiene en la segunda coordenada el resultado de multiplicar, en producto punto, la segunda fila o renglón de la matriz T, la cual es $(-9, 7)$, por la primera columna de la matriz S, la cual es $(2, 5)$.

356. $\boxed{\textbf{\textit{Ejercicio}}}$ *Prediga cuáles serán las coordenadas de la segunda columna de la matriz producto TS. Demuestre su predicción.*

357. $\boxed{\textbf{\textit{Ejercicio}}}$ *Generalizando los resultados anteriores, prediga la receta general para el producto de matrices cualesquiera. Demuestre su predicción.*

5.4. Multiplicación de matrices

Esta sección empieza con repaso y complementa lo visto en todo el capítulo. Consideremos el sistema lineal de ecuaciones

$$\begin{cases} 3x + 4y = 7 \\ 4x + 3y = 7 \end{cases}$$

Antes escribíamos este sistema de la forma

$$M = \begin{pmatrix} 3 & 4 & : & 7 \\ 4 & 3 & : & 7 \end{pmatrix}$$

Ahora lo vamos escribir de la siguiente manera:

$$\begin{pmatrix} 3 & 4 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x + 4y \\ 4x + 3y \end{pmatrix} = \begin{pmatrix} 7 \\ 7 \end{pmatrix}$$

Esta reescritura define varios elementos. En primer término a

$$\begin{pmatrix} x \\ y \end{pmatrix}$$

lo llamamos vector columna o vector. Es un elemento de dos coordenadas y por lo tanto vive en $\mathbb{R}^2$. En segundo lugar, estamos definiendo la multiplicación de una matriz por un vector columna (primero la matriz, después el vector), en nuestro caso

$$\begin{pmatrix} 3 & 4 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x + 4y \\ 4x + 3y \end{pmatrix}$$

lo cual dice que la multiplicación de una matriz por un vector columna debe restituir un vector que represente el lado izquierdo de un sistema de ecuaciones. Operacionalmente es como sigue: se toma cada renglón de la matriz y se multiplica en producto punto por el vector columna. El resultado del renglón i por el vector columna es el renglón i de la respuesta de la multiplicación, la cual es un vector columna con tantos renglones cuantos renglones hay en la matriz. Naturalmente que la multiplicación está definida únicamente cuando el producto punto puede ejecutarse, es decir, cuando el número de columnas de la matriz coincide con el número de renglones del vector columna. Además usamos la igualdad: dos vectores columna son iguales cuando son iguales coordenada por coordenada o renglón por renglón.

Haber multiplicado una matriz por un vector nos produjo un vector, en nuestro caso un elemento con dos coordenadas, es decir, un elemento de $\mathbb{R}^2$. Todo lo anterior lo podemos escribir sucintamente diciendo que M es una función, llamada transformación lineal, que toma vectores con dos coordenadas y los transforma en vectores de dos coordenadas.

$$M \colon \mathbb{R}^2 \to \mathbb{R}^2$$
$$M(x, y) = (3x + 4y, 4x + 3y)$$

Cuando usamos notación de funciones, como aquí, los vectores forman tupletas que se escriben horizontalmente, pero cuando usamos notación matricial, los vectores forman una columna y van delante de la matriz.

Una matriz representa una función y las funciones o transformaciones pueden componerse, es decir, pueden ponerse a operar una después de otra. Lo interesante es que al componer 2 matrices como transformaciones, la transformación resultante tiene una matriz asociada llamada la matriz producto. El resultado de la multiplicación se puede calcular automáticamente de acuerdo con la siguiente:

358. ***Regla para multiplicar matrices***

1. Queremos que el producto de matrices T y S, en ese orden, represente la matriz de la compuesta $T \circ S = T(S())$.

2. El orden es importante porque, en general, transformar de un modo y después de otro no es conmutativo (piense en las transformaciones que sufren los alimentos al ser preparados y además estudie rotaciones de 90 grados alrededor de dos ejes perpendiculares).

3. Para hallar la matriz de un producto TS se cuadran las matrices T, S en el siguiente arreglo en escuadra, de tal forma que el producto está en el centro de la escuadra:

$$TS = \begin{pmatrix} & \vdots & S \\ \dots & \dots & \dots \\ T & \vdots & TS \end{pmatrix}$$

Para hallar la entrada p_{ij} del producto, fila i con columna j, se multiplica en producto punto la fila i de T con la columna j de S, como en el siguiente

359. **Ejemplo** *Multipliquemos mecánicamente las dos matrices TS, en ese orden, y verifiquemos que se trata de hallar la matriz de una compuesta. Sea*

$$S = \begin{pmatrix} 2 & -3 \\ 5 & 4 \end{pmatrix}$$

$$T = \begin{pmatrix} 3 & 2 \\ -9 & 7 \end{pmatrix}$$

Para hallar el producto TS, primero preparamos un arreglo en escuadra, en donde T aparece en la esquina izquierda inferior y S en la esquina superior derecha. En el centro de la escuadra pondremos el producto TS.

$$TS = \begin{pmatrix} & & \vdots & 2 & -3 \\ & & \vdots & 5 & 4 \\ \dots & \dots & \dots & .. & .. \\ 3 & 2 & \vdots & \bullet & \bullet \\ -9 & 7 & \vdots & \bullet & \bullet \end{pmatrix}$$

La matriz producto será calculada inmediatamente, pero en el paso anterior ha sido reemplazada por marcas negras para indicar el mecanismo de multiplicación: cada marca define una fila de T y una columna de S. En el producto de matrices, se reemplaza la marca por el producto punto o escalar de dicha fila de T por la susodicha columna de S:

$$TS = \begin{pmatrix} & & \vdots & 2 & -3 \\ & & \vdots & 5 & 4 \\ ... & ... & ... & .. & .. \\ 3 & 2 & \vdots & 3\times 2 + 2\times 5 & 3\times(-3)+2\times 4 \\ -9 & 7 & \vdots & (-9)\times 2 + 7\times 5 & (-9)\times(-3)+7\times 4 \end{pmatrix}$$

Ahora ejecutamos los productos punto y podemos encontrar la matriz producto en la esquina de la escuadra (abajo a la derecha):

$$TS = \begin{pmatrix} & & \vdots & 2 & -3 \\ & & \vdots & 5 & 4 \\ ... & ... & ... & .. & .. \\ 3 & 2 & \vdots & 16 & -1 \\ -9 & 7 & \vdots & 17 & 55 \end{pmatrix}$$

Verifiquemos ahora que se trata de la matriz de una compuesta. Para eso tenemos que calcular $S(x, y)$, producir un vector resultante que ha de ser procesado por T y el resultado debe coincidir con el resultado de multiplicar la matriz TS por (x, y). Veamos. Primero calculemos $S(x, y)$:

$$\begin{pmatrix} 2 & -3 \\ 5 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - 3y \\ 5x + 4y \end{pmatrix}$$

A este vector resultante lo pasamos por T:

$$\begin{pmatrix} 3 & 2 \\ -9 & 7 \end{pmatrix}\begin{pmatrix} 2x - 3y \\ 5x + 4y \end{pmatrix} = \begin{pmatrix} 3(2x - 3y) + 2(5x + 4y) \\ -9(2x - 3y) + 7(5x + 4y) \end{pmatrix}$$

$$= \begin{pmatrix} 6x - 9y + 10x + 8y \\ -18x + 27y + 35x + 28y) \end{pmatrix} = \begin{pmatrix} 16x - y \\ 17x + 55y \end{pmatrix}$$

$$= \begin{pmatrix} 16 & -1 \\ 17 & 55 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

que da exactamente la matriz TS, la cual hallamos mecánicamente. Hemos verificado entonces que el producto de dos matrices corresponde a la composición de las matrices interpretadas como transformaciones.

360. $\boxed{\textbf{\textit{Ejercicio}}}$ *Sean las siguientes matrices:*

$$A = \begin{pmatrix} -2 & -1 \\ 5 & -3 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & -4 \\ 2 & 4 \end{pmatrix}$$

$$C = \begin{pmatrix} 2 & -5 \\ -3 & 1 \end{pmatrix}$$

Calcule mecánicamente los siguientes productos y verifique que se trata de la matriz de una compuesta:

1. AB
2. BA
3. CA
4. CB
5. ABC
6. $A^2 = AA$

Nosotros hemos visto cómo se multiplican dos matrices cuadradas 2×2. El procedimiento usado se extiende naturalmente a matrices de cualquier orden, cuadradas o no, siempre y cuando el producto tenga sentido.

361. ***Ejemplo*** *Sea la TL definida por*

$$S\colon \mathbb{R}^2 \to \mathbb{R}^2$$
$$S(x, y) = (x - y, 2x + 4y)$$

y la TL

$$T\colon \mathbb{R}^2 \to \mathbb{R}^3$$
$$T(x, y) = (2x - 3y, x + 5y, -x - 7y)$$

Usando la tecnología del producto de matrices hallemos la expresión de la compuesta $T \circ S = T(S())$.

Solución: Lo primero que tenemos que verificar es que efectivamente la compuesta está bien definida. Para esto, lo que se necesita es que la salida de la primera función se pueda encadenar con la entrada de la segunda. Para componer TL nosotros exigimos que el codominio o conjunto de llegada de la primera sea exactamente igual al dominio o conjunto de salida de la segunda. Para nuestro ejemplo, el codominio de S es $\mathbb{R}^2$ que es igual al conjunto de salida de T. Por lo tanto, se pueden encadenar, lo cual también podemos resumirlo en el siguiente diagrama:

$$\begin{tikzcd} \mathbb{R}^2 \arrow[r, "S"] \arrow[dr, "T \circ S"'] & \mathbb{R}^2 \arrow[d, "T"] \\ & \mathbb{R}^3 \end{tikzcd}$$

Figura 5.3. La salida de S se puede encadenar con la entrada de T porque ambas son el mismo EV.

La matriz de S es

$$S = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$$

La matriz de T es

$$T = \begin{pmatrix} 2 & -3 \\ 1 & 5 \\ -1 & 7 \end{pmatrix}$$

Hacemos el arreglo en escuadra

$$TS = \begin{pmatrix} & & \vdots & 1 & -1 \\ & & \vdots & 2 & 4 \\ \cdots & \cdots & \cdots & .. & .. \\ 2 & -3 & \vdots & \bullet & \bullet \\ 1 & 5 & \vdots & \bullet & \bullet \\ -1 & 7 & \vdots & \bullet & \bullet \end{pmatrix}$$

En en el lugar de cada marca negra ponemos el producto punto entre la fila de T y la columna de S correspondientes:

$$TS = \begin{pmatrix} & & \vdots & 1 & -1 \\ & & \vdots & 2 & 4 \\ \cdots & \cdots & \cdots & .. & .. \\ 2 & -3 & \vdots & (2)(1)+(-3)(2) & (2)(-1)+(-3)(4) \\ 1 & 5 & \vdots & (1)(1)+(5)(2) & (1)(-1)+(5)(4) \\ -1 & 7 & \vdots & (-1)(1)+(7)(2) & (-1)(-1)+(7)(4) \end{pmatrix}$$

Ejecutamos la aritmética y obtenemos la respuesta: la matriz de la compuesta $T \circ S$ es TS dada por

$$TS = \begin{pmatrix} -4 & -14 \\ 11 & 19 \\ 13 & 29 \end{pmatrix}$$

La compuesta opera sobre vectores que tienen un número de componentes igual al número de columnas de TS, es decir, sobre elementos de $\mathbb{R}^2$ y produce elementos con tres componentes, $\mathbb{R}^3$, lo cual es compatible con el diagrama de la compuesta. La expresión para TS es entonces:

$$T \circ S \colon \mathbb{R}^2 \to \mathbb{R}^3$$
$$TS(x,y) = (-4x - 14y, 11x + 19y, 13x + 29y)$$

Como hay tantos números envueltos, hacemos un chequeo: calculemos TS sobre $(1,3)$ por dos caminos, por definición de la compuesta y usando la matriz producto. Debe dar lo mismo y es cierto: en ambos casos obtenemos $T(S(1,3)) = (-46, 68, 100)$.

362. $\boxed{\textbf{\textit{Ejercicio}}}$ *Sean las siguientes matrices:*

$$A = \begin{pmatrix} -2 & -1 & 3 \\ 5 & -3 & 4 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & -4 \\ 2 & 4 \\ -1 & 3 \end{pmatrix}$$

$$C = \begin{pmatrix} 2 & -5 \\ -3 & 1 \end{pmatrix}$$

Calcule mecánicamente los siguientes productos cuando esto sea posible. Para decidir si la compuesta tiene sentido, verifique dos cosas, que el conjunto de llegada de la primera sea igual al conjunto de salida de la segunda y que el arreglo en escuadra permite definir las marcas negras sin ambigüedad.

1. AB
2. BA
3. CA
4. CB
5. ABC
6. $A^2 = AA$

5.5. Ejercicios de repaso

1. Compruebe que la matriz $A = \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}$ satisface la ecuación cuadrática $X^2 + X - 2\mathbb{I} = 0$ donde X^2 denota el producto de matrices XX para matrices cuadradas. ¿Podría hallar otra matriz B que sea solución de dicha ecuación?

2. Sea A una matriz $m \times n$ y sea e_j el vector columna $n \times 1$ cuya $j-$ésima componente es 1 y cuyas otras componentes son ceros. Mostrar que:

a) $A \cdot e_j$ es la j-ésima columna de A.

b) Si $A\vec{x} = 0$, $\forall \vec{x} \in \mathbb{R}^n$, entonces $A = 0$.

c) Si A y B son matrices del mismo tamaño y $A\vec{x} = B\vec{x}$, $\forall \vec{x} \in \mathbb{R}^n$, entonces $A = B$.

3. Una matriz A cuadrada es nilpotente si $A^r = 0$ para algún entero positivo r.

   a) Dar un ejemplo de una matriz 2×2 no nula que sea nilpotente.

   b) Mostrar que, si existe B tal que $AB = BA = \mathbb{I}$, entonces A no es nilpotente.

4. Considere la transformación lineal

   $T\colon \mathbb{R}^2 \to \mathbb{R}^2$, definida por $T(X) = AX$, donde $A = \begin{pmatrix} 5 & 4 \\ 1 & 2 \end{pmatrix}$

   a) Halle los valores de λ para los cuales $T(X) = \lambda X, \;\; X \neq 0$.

   b) Halle un vector $\vec{v} \in \mathbb{R}^2$, $\vec{v} \neq \vec{0}$, tal que $T(\vec{v}) = 6\vec{v}$.

5. Sea $P_{\leq n}$ el EV de los polinomios de grado menor o igual a n. Consideremos las transformaciones dadas por

   $T_1\colon P_1 \to P_2$ con $T_1(p(x)) = xp(x)$

   $T_2\colon P_2 \to P_2$ con $T_2(p(x)) = p(2x+1)$

   a) Calcule $T_1(7x+3)$.

   b) Calcule $T_2(4x^2 - 3x + 1)$.

   c) Calcule $T_2(T_1(9x-1))$.

   d) Calcule $T_2(T_1(p(x)))$.

   e) ¿Hay algún polinomio para el cual $T_1(T_2(p(x)))$ esté bien definido?

   f) Demuestre que T_1 y T_2 son lineales.

   g) Halle $T_1(a+bx)$, $T_2(a+bx+cx^2)$, $T_2(T_1(a+bx)))$ e invente un procedimiento matricial para obtener por matrices los mismos resultados.

5.6. Resumen

Todo sistema lineal de ecuaciones está naturalmente asociado a una matriz, digamos F, la cual representa una función con propiedades especiales. Es compatible con las operaciones de un EV, la suma y la multiplicación por escalares:

$$F(\alpha\vec{x} + \beta\vec{y}) = \alpha F(\vec{x}) + \beta F(\vec{y})$$

La función resultante se denomina transformación lineal (TL). La compuesta, cuando existe, de dos TL es TL y por tanto también tiene matriz: la matriz de una compuesta es el producto de matrices. Todo esto es tan natural que se denota simplemente como: la matriz de $F \circ G$ es el producto de las matrices FG. La multiplicación de matrices es no conmutativa, en general, $FG \neq GF$.

5.7. Gran taller de repaso

Con este taller podrá revisar y aplicar la teoria desarrollada en el capitulo. Asuma cada ejercicio como un reto. [1]

1. Encontrar los coeficientes a, b, c para que la parábola $y = ax^2 + bx + c$ pase por los puntos $(1, 4), (2, 8)$ y $(3, 14)$.

   Rta. $a = 1, b = 1, c = 2$.

2. Una fábrica de muebles tiene dos divisiones: un taller donde se fabrican las partes de los muebles y una división de ensamble, donde se unen las partes para obtener el producto terminado. Suponga que se tienen 12 empleados en el taller y 20 en la división, y que cada empleado trabaja 8 horas diarias. Suponga que se producen sólo sillas y mesas. Una silla requiere $\frac{384}{17}$ horas de maquinado y $\frac{480}{17}$ horas de ensamblaje. Una mesa requiere $\frac{240}{17}$ horas de maquinado y $\frac{640}{17}$ horas de ensamblaje. Suponga que se tiene una demanda ilimitada de estos productos y que el fabricante quiere mantener ocupados todos los empleados. ¿Cuántas sillas y cuántas mesas al día puede producir esa fábrica?

   Rta. 3 sillas y 2 mesas.

3. En 3 islas vive una población *estable* de 35.000 aves. Cada año, el 10 % de población de la isla A emigra a la isla B, el 20 % de la población de la isla B emigra a la isla C y el 5 % de la población de la isla C emigra a la isla A. Calcule el número de aves en cada isla, si no *varía* la población de cada isla de un año al siguiente.

4. Para cada una de las siguientes afirmaciones diga si es falsa (F) o verdadera (V) según sea el caso. Si es falsa, justifique su respuesta mediante un ejemplo o mediante el uso de la teoría. Si es verdadera, justifique únicamente mediante la teoría (teoremas, definiciones, etc).

   a) El generado de cualesquiera dos vectores no nulos en $\mathbb{R}^2$ es todo $\mathbb{R}^2$.

   b) El generado de cualesquiera tres vectores no nulos y no paralelos en $\mathbb{R}^3$ es todo $\mathbb{R}^3$.

   c) Si $\{\vec{v_1}, \vec{v_2}, ..., \vec{v_k}\}$ son vectores en $\mathbb{R}^2$ tales que el gen$\{\vec{v_1}, \vec{v_2}, ..., \vec{v_k}\} = \mathbb{R}^2$, entonces $k = 2$.

   d) Existen exactamente dos vectores perpendiculares a un vector no nulo dado en $\mathbb{R}^n$.

   e) El ángulo entre dos vectores no nulos en $\mathbb{R}^n$ es menor que 90°, si y sólo si, el producto punto de los vectores es positivo.

   f) Si $\vec{u} \cdot \vec{v} = 0$, entonces $\vec{u} = 0$, o bien, $\vec{v} = 0$.

[1] Con un especial agradecimiento al profesor Oscar Casas.

g) Si $x \cdot y = x \cdot z$, entonces ¿siempre $y = z$? Justifique su respuesta.

Sugerencia: Construya unos ejemplos.

h) $A^2 - B^2 = (A - B)(A + B)$, para cualesquiera matrices cuadradas A y B.

i) Si $\vec{v}$, $\vec{w}$ son vectores de $\mathbb{R}^n$ con la misma magnitud, entonces la magnitud de $\vec{v} - \vec{w}$ es cero.

j) Todo sistema lineal de ecuaciones con el mismo número de ecuaciones que de incógnitas tiene siempre solución única.

k) Todo sistema lineal de ecuaciones con el mismo número de ecuaciones que de variables tiene al menos una solución.

l) Un sistema lineal con una matriz cuadrada asociada A tiene solución única, si y sólo si, la matriz A es equivalente por renglones a la matriz identidad.

m) Un sistema lineal con más ecuaciones que variables siempre admite un número infinito de soluciones.

5. Encontrar todos los escalares c, si existen, para los cuales se tiene que:

   a) El vector $(2, 6)$ es paralelo al vector $(c, -3)$.

   b) El vector $(c, -c, 4)$ es paralelo al vector $(-2, 2, 20)$.

   c) El vector (c^2, c^3, c^4) es paralelo al vector $(1, -2, 4)$.

   d) El vector $(13, -15)$ es una combinación lineal de los vectores $(1, 5)$ y $(3, c)$.

   e) El vector $\vec{i} + c\vec{j}$ es una combinación lineal de los vectores $\vec{i} + 2\vec{j} + \vec{k}$ y $3\vec{i} + 6\vec{j} + 3\vec{k}$.

   f) El vector $(c, -3, 5)$ es ortogonal al vector $(-1, 3, 4)$.

   g) El vector $(c, -c, c)$ es ortogonal tanto al vector $(-1, 3, 4)$ como al vector $(2, 1, -1)$.

   h) Los puntos $(2, 0, 4)$, $(4, c, -c)$ y $(6, c, c)$ son los vértices de un triángulo equilátero.

6. Mostrar que el punto medio de la hipotenusa de un triángulo rectángulo es equidistante de los tres vértices.

7. Encuentre la distancia del punto $P(1, -1, 2)$ al plano $2x + y - z = 1$.

8. Determine el punto Q del plano $2x + y - z = 1$ que se encuentra más cerca del origen.

9. Pruebe que si $\vec{b} + \vec{c} = 0$ y $||\vec{a}|| = ||\vec{b}||$, entonces $(\vec{a} - \vec{b}) \cdot (\vec{a} - \vec{c}) = 0$.

   Sugerencia: Distribuya el producto punto y use las hipótesis.

10. Sean $\vec{u} = (-2, 5)$ y $\vec{v} = (\alpha, -2)$. Determine α tal que:

    a) $\vec{u}$ y $\vec{v}$ sean ortogonales.

b) El ángulo entre $\vec{u}$ y $\vec{v}$ es de $2\pi/3$.

c) $\vec{u}$ y $\vec{v}$ sean paralelos.

d) El ángulo entre $\vec{u}$ y $\vec{v}$ es de $\pi/3$.

Rta. a) $\alpha = -5$, b) $\alpha = \frac{80\pm 58\sqrt{3}}{13}$, c) $\alpha = \frac{4}{5}$, d) para ningún α.

11. Para vectores $\vec{u}$, $\vec{v}$ y $\vec{w}$ en $\mathbb{R}^n$ y para escalares r y s, probar que, si $\vec{w}$ es ortogonal tanto a $\vec{u}$ como a $\vec{v}$, entonces $\vec{w}$ es ortogonal a $r\vec{u} + s\vec{v}$.

12. Encontrar los valores de a, b y c para los cuales la parábola $y = ax^2 + bx + c$ pasa por los puntos $(1, -4), (0, -5)$ y $(2, 3)$.

13. Hallar los valores de A y B tales que $\frac{x}{2x^2 + x - 1} = \frac{x}{(2x-1)(x+1)} = \frac{A}{2x-1} + \frac{B}{x+1}$.

14. Consideremos el sistema lineal

$$\begin{cases} x_1 + 2x_2 - 3x_3 + x_4 = 2 \\ 3x_1 + 6x_2 - 8x_3 - 2x_4 = 1. \end{cases}$$

a) Codifique el sistema de la forma matricial $A\vec{X} = \vec{B}$.

b) Halle la solución más general, $\vec{s}_h$, en términos de parámetros libres, del sistema $A\vec{X} = \vec{0}$. Este sistema, igualado al vector cero, se llama **homogéneo**.

c) Halle por tanteo una solución específica $\vec{x}_p$ del sistema $A\vec{X} = \vec{B}$. Una solución específica del sistema $A\vec{X} = \vec{B}$ recibe el nombre de **solución particular**. ¿Es única la solución particular?

d) Halle la solución más general, $\vec{s}_g$, del sistema completo $A\vec{X} = \vec{B}$ llamado **no homogéneo**.

e) Demuestre que $\vec{s}_g = \vec{s}_h + \vec{x}_p$. De esa forma se demuestra que toda solución general es la suma de una solución al sistema homogéneo más una solución particular. En términos geométricos: la solución de un sistema $A\vec{X} = \vec{B}$ se puede descomponer como la suma de un subespacio vectorial (que pasa por cero y es solución del sistema homogéneo) más un vector de traslación. Haga un dibujo que ayude a entender si tal descomposición es o no única, es decir, si el vector traslación o la solución particular es o no única.

15. Considere el *sistema lineal* con un parámetro k: $kx + y + z = 0$, $x + ky + z = 1$.

a) Comprobar que para $k = -1$, el vector $(1, 1/2, 1/2)$, satisface el sistema.

b) ¿Para qué valor(es) del parámetro k el sistema es *inconsistente* (que contiene una contradicción)?

c) ¿Para qué valor(es) del parámetro k el sistema admite *infinitas soluciones*?

d) Para $k = -1$ escriba la solución del sistema en la forma $Y_g = Y_h + Y_p$, donde Y_h es la solución del *sistema homogéneo* y Y_p es una solución particular del *no homogéneo.*

16. Si un sistema lineal tiene *matriz aumentada* equivalente por filas a la matriz $\begin{pmatrix} 1 & 2 & 1 & | & -2 \\ 0 & 1 & 2 & | & -2 \\ 0 & 0 & a^2-4 & | & a+2 \end{pmatrix}$ determine el(los) valor(es) de a para el(los) cual(es) el sistema tiene:

    a) Única solución.
    b) Infinitas soluciones.
    c) Inconsistencia.

17. Sea $H = \{(x, y) : xy = 0\}$. ¿Es H con suma y producto por escalar usuales un subespacio vectorial de $\mathbb{R}^2$? Justifique su respuesta.

18. Determine si el conjunto de matrices 3×3 de la forma $\begin{pmatrix} x & 0 & x \\ 0 & x & 0 \\ x & 0 & x \end{pmatrix}$, donde cada x puede ser un escalar cualquiera, es un espacio vectorial con suma y producto por escalar usual. En caso afirmativo halle una base para tal espacio y encuentre su dimensión.

19. Decidir cuáles de los siguientes conjuntos son subespacios de $\mathbb{R}^n$ ::

    a) $H_1 = \{X \in \mathbb{R}^n \colon A \cdot X = 0, A \in M_{n\times n}\}$.
    b) $H_2 = \{(a_1, a_2, a_3, ..., a_n) \colon a_1^2 + a_2^2 + a_3^2 + ... + a_n^2 = 1\}$.
    c) $H_3 = \{(a_1, a_2, a_3, ..., a_n) \colon a_1 + a_2 + a_3 + ... + a_n = 0\}$.

    Rta. a) Si, b) No, c) Si.

20. Extender el conjunto $\left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} \right\}$ para formar una base del espacio $\mathbb{R}^4$.

    Rta. Añada el vector $(1, 0, 0, 0)$ y verifique que estos cuatro vectores forman una base.

21. Determine cuáles de los siguientes subconjuntos del espacio vectorial de matrices 2×2 son subespacios.

    a) $\left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$

    b) $\left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \mid a + b = 0 \right\}$

    c) $\left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \mid a + b = 5 \right\}$

    d) $\left\{ \begin{pmatrix} a & c \\ 0 & b \end{pmatrix} \mid a + b = 0, \quad c \in \mathbb{R} \right\}$

22. Determine si el vector v pertenece al espacio generado por los vectores del conjunto H.

    a) $v = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \quad H = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

    b) $v = x - x^3, \quad H = \left\{ x^2,\ 2x + x^2,\ x + x^3 \right\}$

    c) $v = \begin{pmatrix} 0 & 1 \\ 4 & 2 \end{pmatrix}, \quad H = \left\{ \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 2 & 3 \end{pmatrix} \right\}$

23. Determine si los siguientes conjuntos de vectores son linealmente dependientes o independientes.

    a) $\left\{ \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ -4 \\ 14 \end{pmatrix} \right\} \quad V = \mathbb{R}^3$

    b) $\left\{ \begin{pmatrix} 1 \\ 7 \\ 7 \end{pmatrix} \begin{pmatrix} 2 \\ 7 \\ 7 \end{pmatrix}, \begin{pmatrix} 3 \\ 7 \\ 7 \end{pmatrix} \right\} \quad V = \mathbb{R}^3$

    c) $\left\{ 3 - x + 9x^2, 5 - 6x + 3x^2, 1 + x - 5x^2 \right\} \quad V = P_2$

    d) $\left\{ 8 + 3x + 3x^2, x + 2x^2, 2 + 2x + x^2, 8 - 2x + 5x^2 \right\} \quad V = P_2$

24. Encuentre una base para las soluciones del siguiente sistema de ecuaciones.

$$\begin{pmatrix} x_1 & - & 4x_2 & + & 3x_3 & - & x_4 & = & 0 \\ 2x_1 & - & 8x_2 & + & 6x_3 & - & 2x_4 & = & 0 \end{pmatrix}$$

25. Encuentre una base para cada uno de los siguientes subespacios y halle su dimensión:

    a) El subespacio $H = \left\{ a_2x^2 + a_1x + a_0 \colon a_2 - 2a_1 = a_0 \right\}$ de P_2

    b) El subespacio $H = \{p(x) \in P_3 \colon p(7) = 0\}$

    c) El subespacio $H = \{p(x) \in P_3 \colon p(7) = 0, \quad p(5) = 0\}$

    d) El subespacio $H = \{p(x) \in P_3 \colon p(7) = 0, \quad p(5) = 0, \quad p(3) = 0\}$

    e) El subespacio $H = \{p(x) \in P_3 \colon p(7) = 0 \quad p(5) = 0, \quad p(3) = 0, \quad p(1) = 0\}$

26. Determine si la transformación dada es o no lineal

    a) $T \colon \mathbb{R}^2 \to \mathbb{R}^2, \qquad T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$

    b) $T \colon \mathbb{R}^3 \to \mathbb{R}^2, \qquad T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ z \end{pmatrix}$

    c) $T \colon \mathbb{R}^2 \to \mathbb{R}^2, \qquad T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} xy \\ x \end{pmatrix}$

d) $T\colon M_{2\times 2} \to M_{2\times 2},\ T(A) = AB$; B una matriz fija

e) $T\colon M_{2\times 2} \to M_{2\times 2},\ T(A) = AA^t$ donde A^t indica la **matriz transpuesta** de A que se obtiene de A intercambiando las filas por columnas, como en el siguiente ejemplo:

$$\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix}^t = \begin{pmatrix} a & c & e \\ b & d & f \end{pmatrix}$$

f) $T\colon P_2 \to P_1, \qquad T(a + bx + cx^2) = b + cx;$

Rta. a) Sí, b) No, c) No, d) Sí, e) Sí, f) Sí.

27. Sea $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ una transformación lineal tal que
$$T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}; \quad T\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -4 \\ 0 \\ 5 \end{pmatrix},$$
calcule

a) $T\begin{pmatrix} 2 \\ 5 \end{pmatrix}$ y $T\begin{pmatrix} -3 \\ 9 \end{pmatrix}$

b) $T\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ y $T\begin{pmatrix} -3 \\ 7 \end{pmatrix}$

Respuesta: a) $T\begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 8 \\ 8 \\ 7 \end{pmatrix}, T\begin{pmatrix} -3 \\ 9 \end{pmatrix} = \begin{pmatrix} 21 \\ 10 \\ -5 \end{pmatrix}$

28. Sea $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ una transformación lineal tal que
$$T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}; \quad T\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -4 \\ 0 \\ 5 \end{pmatrix},$$

29. Sea V un espacio vectorial. Muestre que si $\vec{v} \in V$ y r es un escalar que cumple $r\vec{v} = 0$, entonces $r = 0$ o $\vec{v} = 0$.

30. Sea $\{v_1, v_2, \ldots, v_n\}$ un conjunto de vectores linealmente independientes. Muestre que los vectores $v_1, v_1 + v_2, v_1 + v_2 + v_3, \ldots, v_1 + v_2 + v_3 + \cdots + v_n$ son linealmente independientes.

31. Encuentre un vector $\vec{v}$ en $\mathbb{R}^3$ que sea linealmente independiente a los vectores $(2, 1, 2)$ y $(-1, 3, 4)$

32. Decida cuáles de los siguientes conjuntos son subespacio de $F(\mathbb{R}) = \{f\colon \mathbb{R} \to \mathbb{R}\}$:

a) $K_1 = \{f\colon \mathbb{R} \to \mathbb{R}\colon \forall x\ (f(x) = f(-x)\}$

b) $K_2 = \{f\colon \mathbb{R} \to \mathbb{R}\colon f(1) = 0\}$

c) $K_3 = \{f\colon \mathbb{R} \to \mathbb{R}\colon f(x) = ax + b, \quad a, b \in \mathbb{R}, b \neq 0\}$

Rta. a) Sí, b) Sí, c) No.

33. Considere el espacio $M_{n\times n}$ de todas las matrices $n \times n$. ¿Es $T = \{a_{ij} = 0, \forall i \geq j\}$ un subespacio de $M_{n\times n}$?

Rta. Sí.

34. Determine cuáles de los siguientes subconjuntos dotados con las operaciones usuales son subespacios del espacio vectorial dado.

 a) $\left\{\begin{pmatrix} a & b \\ a-b & a+b \end{pmatrix} \mid a, b \in \mathbb{R}\right\}, \quad V = M_{2\times 2}(\mathbb{R}).$

 b) $U = \left\{p(x) \in P_{\leq 3}\colon \quad \int_0^1 p(x)dx = 0\right\}, \quad V = P_{\leq 3}.$

Rta. a) Sí. b) Sí.

35. En el espacio vectorial real de los polinomios de grado menor o igual a 3 se consideran los subconjuntos

 $F_0 = \{p(x) \mid p(0) = 0\}$

 $F_1 = \{p(x) \mid p(1) = 0\}$

 $F_{-1} = \{p(x) \mid p(-1) = 0\}$

 a) Probar que F_1, F_0 y F_{-1} son subespacios de $P_{\leq 3}$.

 b) Hallar una base para cada subespacio.

 Respuesta: b) F_0 =gen$\{x, x^2, x^3\}$, F_1 =gen$(\{-1 + x, -1 + x^2, -1 + x^3\})$,
 F_{-1} = gen$(\{1 + x, -1 + x^2, 1 + x^3\})$

36. Calcule explícitamente los siguientes espacios:

 a) El espacio generado por el vector $(1, 2, 3)$ en $\mathbb{R}^3$.

 b) El espacio generado por los polinomios $\{p_1(x), p_2(x)\}$, donde $p_1(x) = 3 + x^2 + x^3$ y $p_2(x) = x + 1$.

 c) El espacio generado por los vectores $\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 7 \end{pmatrix}\right\}$ en $\mathbb{R}^3$.

 Rta. a) Recta que pasa por el origen en dirección del vector $(1, 2, 3)$, b) Todos los polinomios de la forma $3a + b + bx + ax^2 + ax^3$, $a, b \in \mathbb{R}$, c) $\mathbb{R}^3$.

37. a) Sean x_1 y x_2 soluciones al sistema no homogéneo $AX = b$. Muestre que $x_1 - x_2$ es una solución del homogéneo.

b) Sea x una solución particular del sistema no homogéneo $AX = b$ y sea y otra solución cualquiera del sistema $AX = b$. Muestre que existe una solución h del sistema homogéneo $AX = 0$ tal que $y = x + h$.

Rta. a) Halle $A(x_1 - x_2)$, b) Utilice la parte a).

38. Determine si el conjunto de vectores es linealmente dependiente o independiente.

    a) $\{(3, -1, 2), (1, 5, 3), (1, -11, -4)\}$ en $\mathbb{R}^3$.

    b) $\{x - 1, x + 1, x^2\}$ en $P_{\leq 2}$.

39. Considere H y K dos subespacios de V y defina $H+K = \{h+k \mid h \in H,\ \ k \in K\}$.

    a) Muestre que $H + K$ es un subespacio de V.

    b) Si $H \cap K = \{0\}$, muestre que $\dim(H + K) = \dim(H) + \dim(K)$.

40. Sea A una matriz de tamaño $n \times n$ y $\lambda \in \mathbb{R}$. Mostrar que el conjunto $S = \{x \in \mathbb{R}^n \colon Ax = \lambda x\}$ es un subespacio de $\mathbb{R}^n$. Determine la dimensión de S si $\lambda = 3 \quad \text{y} \quad A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 0 & -1 \\ 0 & 3 & 4 \end{pmatrix}$.

41. Sea $P_{\leq 3}$ el espacio vectorial de los polinomios de grado menor o igual a 3; y sean $W = \{p \in P_{\leq 3} \colon p(0) = 0\}$ y $U = \{p \in P_{\leq 3} \colon p(1) = 0\}$. Muestre que W y U son subespacios del espacio $P_{\leq 3}$. Determinar una base para W, una base para U y una base para $W \cap U$.

42. Sean $V = \mathbb{R}^3$ y $W = \{(x, y, z) : x + 2y + 3z = 0\}$.

    a) Encuentre dos vectores $\vec{v_1}$ y $\vec{v_2}$, que generen a todo W.

    b) ¿Está el vector $\vec{v} = (-1, -1, 1)$ en el gen$\{\vec{v_1}, \vec{v_2}\}$?

43. Determine si la transformación dada es o no lineal

    a) $T \colon \mathbb{R}^2 \to \mathbb{R}^2, \qquad T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$

    b) $T \colon \mathbb{R}^3 \to \mathbb{R}^2, \qquad T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ z \end{pmatrix}$

    c) $T \colon \mathbb{R}^2 \to \mathbb{R}^2, \qquad T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} xy \\ x \end{pmatrix}$

    d) $T \colon M_{2\times 2} \to M_{2\times 2},\ \ T(A) = AB$; B una matriz fija

    e) $T \colon M_{2\times 2} \to M_{2\times 2},\ T(A) = AA^t$

    f) $T \colon P_2 \to P_1,\ T(a + bx + cx^2) = b + cx$

44. Se dice que la matriz A conmuta con la matriz B si $AB = BA$. Muestre que si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ conmuta con toda matriz 2×2, entonces existe un escalar k tal que $A = k\mathbb{I}$.

    Sugerencia: Considere por ejemplo los productos AB, BA, con $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

CAPÍTULO 6

DETERMINANTE DE UNA TL

Regresamos ahora a la teorización sobre álgebra lineal para poder avanzar en aplicaciones. Los resultados de esta sección los necesitamos para hallar soluciones a sistemas de ecuaciones, poder medir la cantidad de soluciones a un sistema de ecuaciones, para hallar la inversa de una matriz y para diagonalización de matrices. También representan una poderosa herramienta para retroalimentación en contra de errores que son tan comunes en álgebra lineal.

6.1. TL y determinantes

363. ♣ *Definición.* *El determinante $Det(T)$ de una TL $T\colon \mathbb{R}^n \to \mathbb{R}^n$ es el determinante de la matriz de T. La definición se aplica sólo cuando el dominio y el codominio coinciden.*

Recordemos que para hallar la matriz de una TL T se toma cada uno de los elementos de la base natural y se pasa por T. El resultado se escribe verticalmente formando una matriz, la matriz de T. Dicho de otra manera, para hallar la matriz de T se toma el n-plp unitario (la base natural), se transforma por T y el n-plp resultante, en notación vertical, da la matriz de T. Una vez que uno tiene la matriz, se la interpreta como un plp y se le saca el determinante y ese es $Det(T)$.

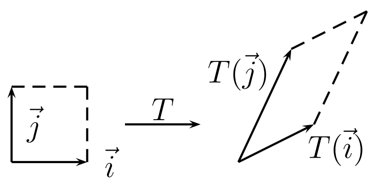


Figura 6.0. El determinante de una TL T es el determinante de la imagen del plp definido por la base natural.

En resumen, el determinante de una transformación lineal T es el volumen (signado) de la imagen por T del plp unitario.

364. **Ejemplo** *Calculemos el determinante de $T(x,y) = (2x-3y, -x+4y)$. Primero calculamos la imagen del cuadrado unitario $T(1,0) = (2,-1)$ y $T(0,1) = (-3,4)$. Por lo tanto, la matriz de T es:*

$$\begin{pmatrix} 2 & -3 \\ -1 & 4 \end{pmatrix}$$

cuyo determinante es $Det(T) = (2)(4) - (-3)(-1) = 8 - 3 = 5$.

Hemos encontrado que $Det(T) = 5$. Esto significa que T toma el 2-plp unitario y lo transforma en un 2-plp cuyo 2-volumen es 5. Por tanto, el volumen del 2-plp unitario se multiplicó por 5. Esto también implica que el volumen de cualquier 2-plp también se amplifica 5 veces:

365. **Ejercicio** *Muestre que para la T del ejercicio previo siempre se da que: $Det(T(P)) = Det(T)Det(P)$ donde P es cualquier 2-plp. Puede mostrarse que la prueba requerida depende tan sólo de la linealidad de T y de la multilinealidad de Det. Por tanto, este ejercicio puede generalizarse a cualquier T y a cualquier n-plp.*

366. $\diamond$ **Teorema.** *El determinante de cualquier TL puede entenderse como un factor de ampliación. Más precisamente, sea $T\colon \mathbb{R}^n \to \mathbb{R}^n$ una TL. Sea P cualquier n-plp. Entonces:*

$Det(T(P)) = Det(T)Det(P)$.

En particular $vol(T(P)) = |Det(T)Det(P)|$

367. $\diamond$ **Teorema.** *Demostremos que si A y B son matrices cuadradas $n \times n$ entonces $Det(AB) = Det(A)Det(B)$.*

Demostración. Cada matriz es la matriz de una TL asociada. Por lo tanto, podemos ver a A y a B como TL. Puesto que ambas tienen el mismo dominio y codominio, se pueden componer y la matriz asociada a la compuesta es el producto de las matrices AB. Para hallar el $Det(AB)$ lo que hacemos es tomar cualquier n-plp P y pasarlo por B y lo que dé, se pasa por A. Cuando pasa por B obtenemos que P se transforma en $B(P)$ y que

$Det(B(P)) = Det(B)Det(P)$

pues el $Det(B)$ funciona como factor de ampliación.

Al pasar a $B(P)$ por A se obtiene:

$Det(A(B(P))) = Det(A)Det(B(P)) = Det(A)Det(B)Det(P)$

es decir que

$Det((AB)(P) = Det(A)Det(B)Det(P)$

por lo que el determinante o factor de ampliación asociado al producto de matrices es el producto de los determinantes. ■

368. **Ejercicio** *Demuestre que el $Det(T) = 0$ si T se define por $T(x,y) = (x,0)$.*

369. ***Ejemplo*** *Ilustremos lo que significa* $Det(T) = 0$. *Para ello, consideremos la* TL *dada por* $T(x, y) = (x, 0)$ *cuyo determinante es cero.*

Vemos que T toma un vector (x, y) y le calcula la sombra vertical sobre el eje $\vec{X}$. Geométricamente, T toma todo el plano y lo apachurra verticalmente contra el eje $\vec{X}$. Muy en particular, T toma el vector $(0, 1)$ y lo procesa en $(0, 0)$. Es decir, lo apachurra. Eso es lo mismo que decir que son muchos los vectores del plano que son apachurrados sobre el mismo vector del eje $\vec{X}$. Por ejemplo, $T(3, 1) = T(3, 2) = T(3, 5) = (3, 0)$. Eso quiere decir que la ecuación $T(\vec{X}) = \vec{Y}$ tiene o bien muchas soluciones, o bien, no tiene ninguna. Veamos:

La ecuación $T(x, y) = (3, 0)$ tiene una infinidad de soluciones, todas del tipo $(3, y)$. Por lo tanto, esta ecuación tiene una infinidad de soluciones con un grado de libertad, pues podemos dar a y el valor que queramos.

En particular, la ecuación $T(x, y) = (0, 0)$ tiene una infinidad de soluciones con un grado de libertad, de la forma $(0, y)$.

Por otra parte, la ecuación $T(x, y) = (3, 1)$ no tiene ni siquiera una solución, pues para que pueda haber solución, la segunda coordenada del vector del lado derecho de la ecuación debe ser cero.

Releamos los anteriores resultados en términos de sistemas de ecuaciones:

$T(x, y) = (x, 0) = (3, 0)$ es equivalente al sistema dado por:

$$\begin{cases} x = 3 \\ 0 = 0 \end{cases}$$

lo cual no da información ni restricciones de ningún tipo sobre y: éste puede ser cualquiera, mientras que x debe ser 3.

De igual forma, $T(x, y) = (x, 0) = (0, 0)$ dice que

$$\begin{cases} x = 0 \\ 0 = 0 \end{cases}$$

o sea que y queda libre: las soluciones son del tipo $(0, y)$ que son una infinidad, o sea, todo el eje $\vec{Y}$.

Por otra parte, $T(x, y) = (x, 0) = (3, 1)$ es equivalente al sistema

$$\begin{cases} x = 3 \\ 0 = 1 \end{cases}$$

que da una contradicción. Es decir, el sistema no tiene solución.

370. $\diamond$ ***Teorema y definición.*** *Si el* $Det(T) = 0$ *quiere decir que* T *toma el n-plp unitario y al transformarlo lo apachurra, pues se le pierde el volumen. Pero si* $Det(T) \neq 0$, *quiere decir que* T *toma al n-plp unitario y al transformarlo no lo apachurra. En otras palabras, si* $Det(T) \neq 0$, T *toma al n-plp unitario, el cual es una base, y lo transforma en otra base. Si el* $Det(T) = 0$, *decimos que* T ***es apachurrante o que apachurra****.*

Ahora bien, si T apachurra al n-plp unitario, eso significa que hay muchos vectores que al ser transformados caen sobre el mismo vector. Por lo tanto, T no es inyectiva, no es 1-1, y por tanto, la ecuación $T(\vec{X}) = \vec{Y}$ no tiene garantía de tener solución única. Puede incluso que no tenga solución.

Pero si $Det(T) \neq 0$, eso implica que T no apachurra al n-plp unitario y por ende no apachurra a nadie, y por tanto T es 1-1, es inyectiva. Eso quiere decir que si la ecuación $T(\vec{X}) = \vec{Y}$ tiene solución, ésta es única. Pero, ¿cuándo hay solución? Cuando la transformación es sobre, es decir, cuando para cualquier Y en el co-dominio existe un X en el dominio tal que $T(X) = Y$. ¿Cuándo una TL es sobre?

371. $\diamond$ ***Teorema y definición.*** *Una transformación lineal $T\colon \mathbb{R}^n \to \mathbb{R}^n$ es 1-1 ssi es sobre. Por tanto, si el sistema $T(X) = Y$ tiene solución, es única.*

Demostración. Si T es 1-1, ésta no puede apachurrar al plp unitario, i.e. la imagen del plp unitario es una base que expande todo el codominio. Por lo tanto, T es sobre. Supongamos ahora que T es sobre. Esto quiere decir que la imagen de T llena todo el codominio. Esto implica que la imagen del plp unitario expande todo $\mathbb{R}^n$. Puesto que el plp unitario tiene n vectores, su imagen tiene máximo n vectores y debe expandir un espacio de dimensión n. Por lo tanto, la imagen del plp unitario es una base de $\mathbb{R}^n$ y no puede ocurrir ningún apachurramiento, dos vectores del dominio no pueden caer sobre el mismo vector del codominio: T es 1-1. Por lo tanto, si la ecuación $T(X) = Y$ tiene solución, es única. ■

372. $\boxed{\textbf{\textit{Ejercicio}}}$ *Una TL puede entenderse dinámicamente como un proceso y geométricamente como una deformación o transformación. Una TL tiene inversa cuando se puede reversar. Oficialmente, una TL $T\colon V \to W$ tiene inversa cuando tiene inversa como función, es decir, cuando existe $G\colon W \to V$ tal que $G \circ F = \mathbb{I}$ y $F \circ G = \mathbb{I}$, donde $\mathbb{I}$ es la identidad en el dominio correspondiente. Demuestre que una transformación lineal T tiene inversa ssi su determinante está definido y no es cero. Dé ejemplos. ¿Cómo saber si dicha inversa es única, o si es TL o no?*

373. ♣ ***Definición.*** *El determinante de un sistema de n ecuaciones con n variables es el determinante de la matriz asociada a dicho sistema.*

374. ***Ejemplo*** *El sistema:*

$$\begin{cases} 2x + 4y = 3 \\ 3x - 7y = 5 \end{cases}$$

tiene como matriz asociada:

$$M = \begin{pmatrix} 2 & 4 \\ 3 & -7 \end{pmatrix}$$

El determinante de M es $2 \times (-7) - 3 \times 4 = -14 - 12 = -26$.

375. $\diamond$ ***Teorema.*** *Si el determinante de un sistema de n ecuaciones en n variables es cero, hay un apachurramiento de por medio, y el sistema o bien no tiene solución, o bien tiene una infinidad de soluciones. En particular, el sistema $T\vec{X} = \vec{0}$ tiene la solución $\vec{0}$ y además otra infinidad de soluciones. Si el determinante de un sistema de ecuaciones es diferente de cero, no hay ningún apachurramiento de por medio y el sistema tiene una solución que es única, es decir, la matriz asociada al sistema tiene inversa.*

376. ***Ejercicio imponente*** *Proponga una demostración rigurosa algebraica al teorema anterior que copie la sencillez y la naturalidad de un ataque geométrico. Primero para el plano $\mathbb{R}^2$ y después en general.*

377. ***Ejemplo*** *Desde el punto de vista de las TL y sus determinantes estudiemos el sistema:*

$$\begin{cases} 2x + 4y = 3 \\ 3x - 7y = 5 \end{cases}$$

A este sistema le corresponde la matriz:

$$M = \begin{pmatrix} 2 & 4 \\ 3 & -7 \end{pmatrix}$$

El determinante de M es $2 \times (-7) - 3 \times 4 = -14 - 12 = -26$. M es una TL que va del plano al plano.

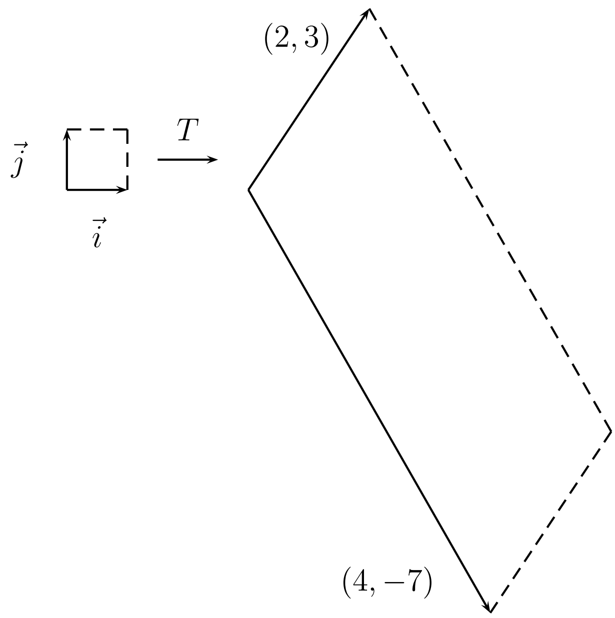


Figura 6.1. Plástica lineal. T transforma o deforma el plp unitario en un plp.

M toma el cuadrado unitario formado por $\vec{i}, \vec{j}$ y lo transforma en el paralelogramo definido por los vectores $(2, 3), (4, -7)$. El cuadrado unitario tiene área 1, pero el paralelogramo en que M transforma dicho cuadrado tiene área 26. El signo menos quiere decir que mientras que en el cuadrado unitario uno va de $\vec{i}$ a $\vec{j}$ contrario a las manecillas del reloj, en el paralelogramo imagen uno va de $M(\vec{i})$ hacia $M(\vec{j})$ con las manecillas del reloj.

Puesto que el determinante de M no es cero, M no apachurra al plano al transformarlo. Por lo tanto, T es 1-1 y sobre. Eso quiere decir que a ningún vector del codominio le llega más de un vector y que a todo punto del plano imagen le llega un vector del plano preimagen. Por tanto, todo sistema de la forma $M\vec{W} = \vec{B}$ tiene solución. Además, la solución es única. Pues si no fuese así, la TL M apachurraría algún vector sobre otro y el determinante sería cero.

378. ***Ejemplo*** *Desde el punto de vista de las TL y sus determinantes estudiemos el sistema:*

$$\begin{cases} 2x + 4y = a \\ 4x + 8y = b \end{cases}$$

A este sistema le corresponde la matriz:

$$M = \begin{pmatrix} 2 & 4 \\ 4 & 8 \end{pmatrix}$$

El determinante de M es $2 \times (8) - 4 \times 4 = 16 - 16 = 0$. M es una TL que va del plano al plano que toma el cuadrado unitario formado por $\vec{i}, \vec{j}$ y lo transforma en el paralelogramo definido por los vectores $(2, 4), (4, 8)$.

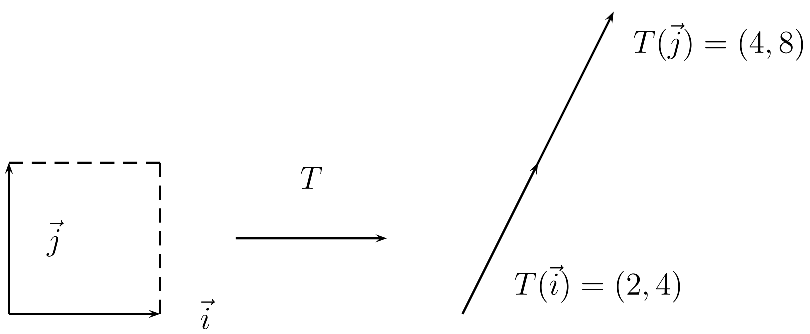


Figura 6.2. Colapso, apachurramiento.

Estos dos vectores son colineales pues la segunda coordenada de cada vector es el doble del primero. Esos dos vectores generan la misma línea que pasa por el origen cuya ecuación es $y = 2x$. El cuadrado unitario tiene área 1, pero el paralelogramo en que M transforma dicho cuadrado tiene área 0. Por eso decimos que M **apachurra** el 3-plp o cubo unitario.

Demostremos ahora que si M transforma el cuadrado unitario en un segmento de recta con área cero, entonces, M apachurrará todo el plano sobre una línea. En efecto: el cuadrado unitario es generado por la base natural $\vec{i}, \vec{j}$. Las imágenes de estos dos vectores son colineales, o paralelos, es decir, $M(\vec{i}) = kM(\vec{j})$. Tomemos un punto cualquiera del dominio, con coordenadas arbitrarias (x, y). Su imagen es

$M(x, y) = M(x\vec{i} + y\vec{j}) = M(x\vec{i}) + M(y\vec{j}) = xM(\vec{i}) + yM(\vec{j})$

$= xkM(\vec{j}) + yM(\vec{j}) = (xk + y)M(\vec{j})$

En resumen:
$M(x, y) = (xk + y)M(\vec{j})$
lo cual dice que todas las imágenes caen sobre un alargamiento o un acortamiento de $M(\vec{j})$, o sea, caen sobre una línea: M toma todo el plano y lo apachurra sobre una línea.

Puesto que el determinante de M es cero, M apachurra al plano al transformarlo. En este caso, lo apachurra sobre una línea. Vemos que no a todo punto del plano imagen le llega un vector del plano preimagen. Por tanto, todo sistema de la forma $M\vec{X} = \vec{B}$ a veces tiene solución y en ese caso tiene muchas porque al vector de la imagen le caen encima todos aquellos que son apachurrados sobre él. Pero puede ser que el sistema no tenga solución, pues M apachurra al plano sobre la recta $y = 2x$ definida por los puntos $(2, 4)$ y $(4, 8)$. Y si un vector en el plano imagen no está en dicha línea, pues no le llega nadie del plano preimagen y el sistema no tiene solución. Verifiquémoslo.

Resolvamos

$$\begin{cases} 2x + 4y = 5 \\ 4x + 8y = 10 \end{cases}$$

Como podemos simplificar la segunda ecuación por dos, el anterior sistema es equivalente a

$$\begin{cases} 2x + 4y = 5 \\ 2x + 4y = 5 \end{cases}$$

y este nuevo sistema tiene una ecuación repetida que no aporta información. Podemos tacharla. Sólo queda la ecuación:
$2x + 4y = 5$
despejando $y = (5 - 2x)/4$.

Es decir, cualquier par de valores de la forma $(x, (5 - 2x)/4)$ satisface el sistema. Por ejemplo, $(0, 5/4)$ o bien, $(1, 3/4)$. Infinitas soluciones que forman una línea, la cual tiene un grado de libertad.

Todo eso fue así porque tomamos el sistema $M\vec{X} = \vec{B}$ con $\vec{B} = (5, 10)$, o sea, sobre la línea $y = 2x$ que es sobre la cual M apachurra todo el plano preimagen.

Resolvamos ahora el sistema:

$$\begin{cases} 2x + 4y = 5 \\ 4x + 8y = 11 \end{cases}$$

Como podemos simplificar la segunda ecuación por dos, el anterior sistema es equivalente al sistema

$$\begin{cases} 2x + 4y = 5 \\ 2x + 4y = 5{,}5 \end{cases}$$

Por transitividad $5 = 5{,}5$.

Eso es una contradicción, no hay solución al sistema. ¿Por qué? Porque el vector $(5, 11)$ no queda sobre la línea $y = 2x$ donde queda apachurrado todo el plano preimagen.

379. $\boxed{\textbf{\textit{Ejercicio}}}$ *Estudie los siguientes sistemas tal como se hizo en los ejemplos anteriores.*

a) $2x - 3y = 5, 3x - 7y = 6$
b) $2x - 3y = 0, 3x - 7y = 0$
c) $2x - 3y = 5, 4x - 6y = 6$
d) $2x - 3y = 0, 4x - 6y = 0$
e) $2x - 3y + 4z = 1, 4x - 5y + 2z = 3, 6x - 8y + 6z = 4$
f) $2x - 3y + 4z = 0, 4x - 5y + 2z = 0, 6x - 8y + 6z = 0$

Nosotros hemos tomado el n-plp unitario como instrumento para predecir si una TL tiene inversa o no, y si los sistemas de ecuaciones asociados tienen única solución o si tal vez no tengan o si tal vez tengan una infinidad. En realidad, el cubo unitario no es necesariamente el mejor instrumento de predicción, del cual podemos saber gracias al siguiente resultado.

380. $\diamond$ ***Teorema del apachurramiento.*** *Dada una TL, $T\colon \mathbb{R}^n \to \mathbb{R}^m$, entonces $Det(T) = 0$ si y sólo si existe una base (un n-plp en el dominio de T, cuyas aristas forman un conjunto LI), tal que al ser transformado por T, un número positivo de aristas son apachurradas contra cero mientras que el resto son transformadas en un conjunto linealmente independiente.*

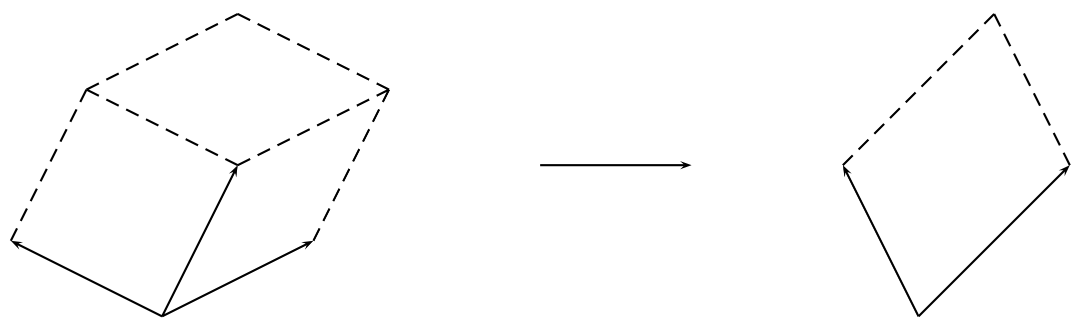


Figura 6.3. Discriminación lineal.

Comenzamos aclarando que una base en un espacio de 3 dimensiones puede ser apachurrada contra un plano, pero eso no implica que sus aristas sean apachurradas contra cero. En contraste, en nuestro teorema, las aristas apachurradas son apachurradas contra cero.

La mejor manera de entender este teorema tan espectacular es verlo funcionando.

381. *Ejemplo* *Consideremos el sistema de ecuaciones*

$$\begin{cases} 2x + 4y + 5z = 25 \\ 3x + 4y - 8z = -13 \\ 5x + 8y - 3z = 12 \end{cases}$$

Este sistema contiene una TL asociada cuya matriz es:

$$M = \begin{pmatrix} 2 & 4 & 5 \\ 3 & 4 & -8 \\ 5 & 8 & -3 \end{pmatrix}$$

Esta matriz toma vectores con 3 coordenadas (x, y, z) en $\mathbb{R}^3$ y produce vectores con 3 coordenadas dadas por $(2x+4y+5z, 3x+4y-8z, 5x+8y-3z)$. Por lo tanto, define una $TL\ T\colon \mathbb{R}^3 \to \mathbb{R}^3$. Concretamente $T(x, y, z) = (2x + 4y + 5z, 3x + 4y - 8z, 5x + 8y - 3z)$.

Para saber si hay apachurramiento o no del cubo unitario calculemos el

$Det(T) = 2(-12 + 64) - 4(-9 + 40) + 5(24 - 20) = 104 - 124 + 20 = 0$.

Hay apachurramiento del cubo unitario. Pero es posible que todo el cubo sea transformado en un plp, pero que ninguna arista del cubo unitario sea apachurrada sobre cero, porque una de sus aristas quede apachurrada sobre el plano formado por las imágenes de los otros dos vectores. Hallemos un 3-plp tal que las aristas apachurradas caigan sobre cero. Se procede así:

Resolvemos el problema: hallar $\vec{v} : T(\vec{v}) = \vec{0}$. Es decir, hallar (x, y, z) tales que

$(2x + 4y + 5z, 3x + 4y - 8z, 5x + 8y - 3z) = (0, 0, 0)$

Procedemos con Gauss-Jordan para resolver dicho sistema. El vector independiente es siempre cero, y aunque combinemos los renglones, sigue siendo cero. Por lo tanto, es siempre cero y se escribe:

$$\begin{pmatrix} 2 & 4 & 5 \\ 3 & 4 & -8 \\ 5 & 8 & -3 \end{pmatrix}$$

$$\begin{array}{r} R2 - R1 \to \\ 3R1 - 2R2 \to \\ R1 + R2 - R3 \to \end{array} \begin{pmatrix} 1 & 0 & -13 \\ 0 & 4 & 31 \\ 0 & 0 & 0 \end{pmatrix}$$

$$R2/4 \to \begin{pmatrix} 1 & 0 & -13 \\ 0 & 1 & 31/4 \\ 0 & 0 & 0 \end{pmatrix}$$

Eso significa que el sistema original es equivalente a:

$x - 13z = 0$

$y + (31/4)z = 0$

o lo que es lo mismo

$x = 13z$

$y = -(31/4)z$

Por lo tanto el conjunto de vectores que se aniquilan sobre cero quedando totalmente apachurrados es de la forma $(13z, (-31/4)z, z) = z(13, -31/4, 1)$. Por lo tanto, podemos armar un plp cuya primera arista sea $(13, -31/4, 1)$, la cual quedará totalmente apachurrada. Completemos ahora el 3-plp con otros dos vectores que no serán apachurrados, es decir, que sus imágenes formen un conjunto LI.

La TL dada está definida por

$$T(x, y, z) = (2x + 4y + 5z, 3x + 4y - 8z, 5x + 8y - 3z)$$

en notación vertical eso se lee así:

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + 4y + 5z \\ 3x + 4y - 8z \\ 5x + 8y - 3z \end{pmatrix} = x\begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} + y\begin{pmatrix} 4 \\ 4 \\ 8 \end{pmatrix} + z\begin{pmatrix} 5 \\ -8 \\ -3 \end{pmatrix}$$

lo cual significa que T transforma todo el dominio en una combinación lineal arbitraria de los vectores columna ahí expresados. Nosotros ya sabemos que esos 3 vectores no pueden expandir un espacio 3D, pues de lo contrario el cubo unitario no se apachurraría. Por lo tanto, dicho conjunto es LD. Ellos expanden un espacio 2D, generado por un conjunto LI con 2 vectores. Vemos claramente que los dos primeros vectores forman un conjunto $\{(2, 3, 5), (4, 4, 8)\}$ que es LI puesto que no son múltiplos. Por lo tanto, el tercer vector es una combinación lineal de los otros dos:

$(5, -8, -3) = -13(2, 3, 5) + (31/4)(4, 4, 8)$

Por tanto, T transforma el espacio 3D en un nuevo espacio generado por estos dos vectores, $\{(2, 3, 5), (4, 4, 8)\}$. Dicho espacio tiene dimensión 2, un número que se llama el **rango de** T.

Estamos hallando una base, es decir, un 3-plp cuyos vectores forman un conjunto LI tal que una de sus aristas se apachurre contra cero y las imágenes de las otras dos formen un conjunto LI. El vector que se apachurra es $(13, -31/4, 1)$ y los otros dos vectores pueden tomarse como cualquiera de las preimágenes de los vectores $\{(2, 3, 5), (4, 4, 8)\}$.

382. $\boxed{\textbf{\textit{Ejercicio}}}$ *Para terminar la demostración anterior encuentre dos vectores* $\vec{X}, \vec{Y}$ *tales que* $T(\vec{X}) = (2, 3, 5)$ *y* $T(Y) = (4, 4, 8)$. *Pruebe que el conjunto* $\{\vec{X}, \vec{Y}, \vec{Z}\}$ *con* $\vec{Z} = (13, -31/4, 1)$ *es una base de* $\mathbb{R}^3$. *Por lo tanto, el plp definido por dicho conjunto es tal que una de sus aristas se apachurra sobre cero mientras que las imágenes de los otros dos generan la imagen de* T. *El primero se apachurra sobre cero, pero los otros dos conservan su independencia lineal al ser transformados.*

6.2. Núcleo e imagen de una TL

Introducimos dos conceptos asociados a una TL, ambos de tremenda importancia: el núcleo y la imagen.

383. ♣ ***Definición.*** *Sea una transformación lineal $T\colon \mathbb{R}^n \to \mathbb{R}^m$. El conjunto de vectores que se apachurran totalmente por T se denomina* ***espacio nulo, núcleo o Kernel*** *de T y se nota Ker(T) y su dimensión se denomina nulidad y se nota ν (nu). El conjunto formado por todas las imágenes de T se denomina el* ***espacio imagen****, se nota Im(T) y su dimensión se denomina* ***rango*** *y se nota ρ (ro).*

En el anterior ejemplo, la nulidad es 1 y el rango 2. El espacio de salida es 3D. No es una coincidencia que $1 + 2 = 3$.

384. ◇ ***Teorema de las dimensiones.*** *Para toda TL, $T\colon \mathbb{R}^n \to \mathbb{R}^m$ se tiene que la nulidad + el rango = dimensión del espacio dominio.*
Es usual escribir esta ecuación como $n = \rho + \nu$

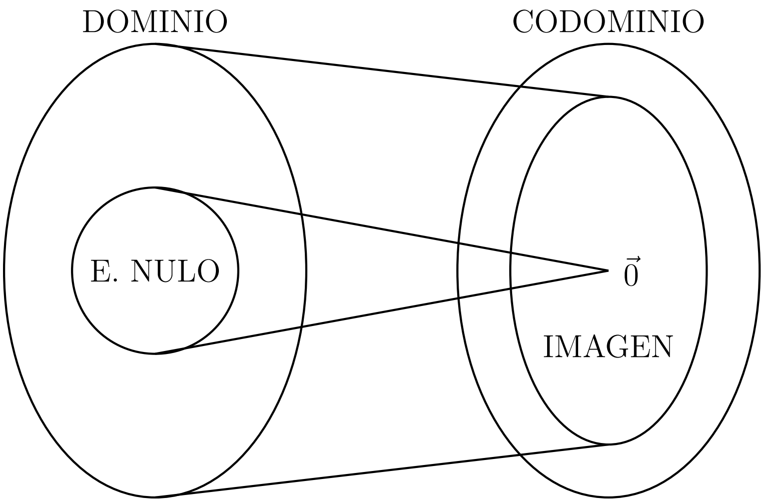


Figura 6.4. Espacio nulo e imagen.

Demostración. Existe un n-plp, cuyos vectores forman una base del dominio, tal que el número de aristas apachurradas es igual a la nulidad ν, mientras que las otras aristas, que son $\rho = n - \nu$, tienen imágenes que forman un conjunto LI que expanden todo el espacio imagen. Claramente $n = \nu + \rho$. ■

Para traducir los anteriores resultados en términos de sistemas de ecuaciones necesitamos formalizar una definición y añadir un teorema adicional:

385. ♣ ***Definición.*** *Un sistema de la forma* $M\vec{X} = \vec{0}$ *se llama* ***homogéneo****. Un sistema de la forma* $M\vec{X} = \vec{B}$*, con* $B \neq 0$*, se llama* ***no homogéneo****. Un vector específico* $\vec{P}$ *que resuelve el sistema no homogéneo se llama* ***solución particular****. Una familia de vectores* $\vec{X}_g$ *se dice que es la* ***solución general*** *al sistema, si cada miembro de la familia es solución y si cada solución es un miembro de la familia.*

386. ⋄ ***Teorema y ejercicio.*** *La solución a un sistema de ecuaciones* $M\vec{X} = \vec{B}$ *se puede descomponer como* $\vec{X} = \vec{N} + \vec{P}$ *donde* $\vec{N}$ *representa el espacio nulo de* M*, solución de* $M\vec{X} = \vec{0}$ *y* P *es cualquier solución a* $M\vec{X} = \vec{B}$*. Se acostumbra decir que la solución a un sistema lineal de ecuaciones puede expresarse como la suma de una solución particular del no homogéneo,* $\vec{P}$*, más la solución general a la ecuación homogénea,* $M\vec{X} = 0$*.*

Demostración: Ejercicio.

387. ***Ejemplo de todo a todo*** *Estudiemos el sistema*

$$\begin{cases} x + 3y = 7 \\ 3x + 9y = 21 \end{cases}$$

Multiplicando la primera ecuación por 3, obtenemos la segunda. Por tanto, la segunda ecuación es redundante y podemos olvidarla. La primera ecuación da:
$y = \frac{7}{3} - \frac{x}{3}$
lo cual significa que el conjunto solución al sistema consiste de los pares $(x, y) = (x, \frac{7}{3} - \frac{x}{3})$. Esta solución puede ser escrita también como
$(0, \frac{7}{3}) + (x, -\frac{x}{3}) = (0, \frac{7}{3}) + x(1, -\frac{1}{3})$.
Por tanto, la solución es una línea que pasa por el punto $(0, \frac{7}{3})$ y cuyo vector director es $(1, -\frac{1}{3})$ y que tiene pendiente -1/3. Verifiquemos que (0,7/3) es en verdad una solución: $x + 3y = 7$ se convierte en $0 + 3(\frac{7}{3}) = 7$ que es una identidad.

El sistema homogéneo es
$x + 3y = 0$
$3x + 9y = 0$
Este sistema es redundante: multiplicamos la primera ecuación por 3 y obtenemos la segunda, la cual podemos olvidar. La solución a la primera ecuación es $y = -x/3$, la cual es una línea que pasa por el origen y cuya pendiente es $-1/3$. Está generada por el vector $(1, -1/3)$. Observemos que la solución al sistema no homogéneo es la solución al sistema homogéneo, $x(1, 1/3)$, más una solución específica particular, $(0, 7/3)$.

Ambos sistemas tienen asociada la misma TL, cuya matriz es

$$M = \begin{pmatrix} 1 & 3 \\ 3 & 9 \end{pmatrix}$$

El núcleo es el conjunto de vectores que caen sobre cero, i.e. son la solución a la homogénea, que es la línea $\{(x, y)\colon x, y \in \mathbb{R}, y = -x/3\}$.

La **imagen de una matriz** es, por definición, la imagen de la TL asociada. En consecuencia,

$$Im(M) = \begin{pmatrix} 1 & 3 \\ 3 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+3y \\ 3x+9y \end{pmatrix}$$

$$= x\begin{pmatrix} 1 \\ 3 \end{pmatrix} + y\begin{pmatrix} 3 \\ 9 \end{pmatrix} = x\begin{pmatrix} 1 \\ 3 \end{pmatrix} + 3y\begin{pmatrix} 1 \\ 3 \end{pmatrix} = (x+3y)\begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

o, Im(M) es el sub-EV generado por $(1,3)$. Observemos que es una línea con un grado de libertad a pesar de que tiene como coeficiente $(x+3y)$: no importa cómo combine uno x con y, siempre tenemos un múltiplo de $(1,3)$, o elementos de la forma $t(1,3)$. Esto significa que el punto $(x,y) \in Im(M)$ ssi $x = t$, $y = 3t$. Por tanto, la imagen es la línea $y = 3x$ con pendiente 3. Puesto que la línea $y = -x/3$ es solución del sistema homogéneo, con pendiente $-1/3$, somos testigos de que para esta matriz el kernel y la imagen son perpendiculares.

El $Det(M) = 9 - 9 = 0$, mostrando que M apachurra a $\mathbb{R}^2$ contra la imagen, la línea $y = 3x$. En particular, la línea $y = -x/3$ es apachurrada contra cero siguiendo la dirección del vector director. Eso nos motivaría a pensar que el conjunto solución al sistema no homogéneo, la línea $y = 7/3 - x/3$, es apachurrada contra la intersección de esta línea con la línea imagen $y = 3x$. Vamos a ver si eso es cierto o no.

El punto intersección de la línea solución con la línea imagen es la solución al sistema
$y = 7/3 - x/3$
$y = 3x$
Esto implica que $7/3 - x/3 = 3x$, ó, $7 - x = 9x$, ó $x = 7/10$ mientras que $y = 21/10$. Nosotros estamos pretendiendo que la imagen de la línea solución $y = 7/3 - x/3$ es $(7/10, 21/10)$. ¿Es verdad?

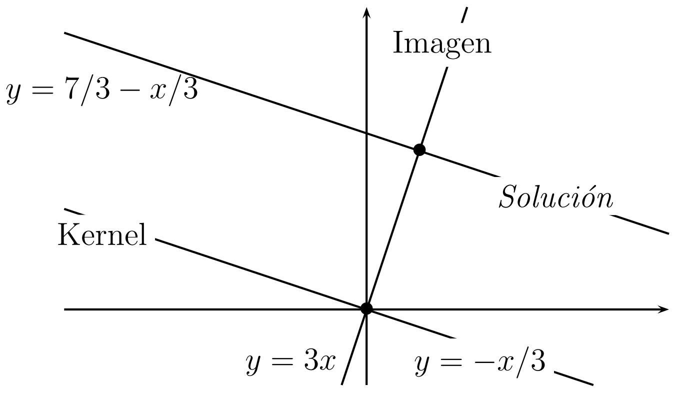


Figura 6.5. Interpretación geométrica de un sistema lineal.

Veamos. La imagen de la línea es

$$\begin{pmatrix} 1 & 3 \\ 3 & 9 \end{pmatrix} \begin{pmatrix} x \\ 7/3 - x/3 \end{pmatrix} = \begin{pmatrix} x + 3(7/3 - x/3) \\ 3x + 9(7/3 - x/3) \end{pmatrix} = \begin{pmatrix} x + 7 - x \\ 3x + 21 - 3x \end{pmatrix}$$

$$= \begin{pmatrix} 7 \\ 21 \end{pmatrix}$$

Por lo tanto, hemos malinterpretado algo, hemos imaginado que M apachurra todo contra la línea imagen de M en una dirección paralela al Kernel. Pero, M no es así de simple. Para verlo, consideremos la imagen por M de los vectores $(1, -1/3)$ y $(1, 3)$. El primer vector está en el núcleo, por lo tanto, es apachurrado contra cero. Pero la imagen del segundo vector es:

$$M = \begin{pmatrix} 1 & 3 \\ 3 & 9 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 10 \\ 30 \end{pmatrix} = 10 \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

por lo que M no apachurra $(1, 3)$ sino que lo elonga 10 veces.

Tenemos entonces una interpretación geométrica de M. Ella apachurra los vectores en la dirección $y = -x/3$ pero elonga 10 veces los vectores en la dirección $y = 3x$.

Adicionalmente, tenemos la siguiente confirmación: la matriz M genera la TL

$M : \mathbb{R}^2 \to \mathbb{R}^2$ tal que Dim(Ker(M))+Dim(Im(M)) $= 1 + 1 = 2 =$ dimensión del dominio, lo cual se escribe usualmente como $\nu + \rho = n$ y que se puede leer como: *nulidad más rango igual a número de incógnitas.*

Número de grados de libertad de la solución = 1 =Dim(kernel).

388. $\boxed{\textbf{\textit{Ejercicio}}}$ *En el ejemplo previo, el Kernel y la imagen de M fueron perpendiculares mutuamente: ¿Es una coincidencia o un teorema por descubrir? Para enriquecer la pregunta, estudie los sistemas*

a) $x + 3y = 7$; $3x + 9y = 21$

b) $ax + by = 7$; $bx + cy = 21$

389. ***Ejemplo de todo a todo*** *Resolvamos el sistema*

$$\begin{cases} 2x + 4y + 5z = 25 \\ 3x + 4y - 8z = -13 \\ 5x + 8y - 3z = 12 \end{cases}$$

Podemos proceder por Gauss-Jordan:

$$\begin{pmatrix} 2 & 4 & 5 & \vdots & 25 \\ 3 & 4 & -8 & \vdots & -13 \\ 5 & 8 & -3 & \vdots & 12 \end{pmatrix}$$

$$\begin{array}{r} R2 - R1 \to \\ 3R1 - 2R2 \to \\ R1 + R2 - R3 \to \end{array} \begin{pmatrix} 1 & 0 & -13 & \vdots & -38 \\ 0 & 4 & 31 & \vdots & 101 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

$$R2/4 \to \begin{pmatrix} 1 & 0 & -13 & \vdots & -38 \\ 0 & 1 & 31/4 & \vdots & 101/4 \\ 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

Eso significa que el sistema original es equivalente a:

$x - 13z = -38$
$y + (31/4)z = 101/4$
o lo que es lo mismo
$x = -38 + 13z$
$y = 101/4 - (31/4)z$

En notación vertical la solución se puede escribir como:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -38 + 13z \\ 101/4 - (31/4)z \\ z \end{pmatrix} = \begin{pmatrix} -38 \\ 101/4 \\ 0 \end{pmatrix} + z \begin{pmatrix} 13 \\ -31/4 \\ 1 \end{pmatrix}$$

Esa forma de escribir da directamente una solución particular
$x = -38, y = 101/4, z = 0$
y una solución del espacio nulo con un grado de libertad: $z(13, -31/4, 1)$.

Debemos aclarar que cuando el espacio nulo no es el vector cero, la solución puede escribirse de muchas maneras. Por ejemplo, para obtener, en este caso, una solución particular podemos dar a z el valor de $z = 3$ y por tanto $y = 2, x = 1$ y la solución podría escribirse:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + z \begin{pmatrix} 13 \\ -31/4 \\ 1 \end{pmatrix}$$

Vemos que el Kernel de la TL asociada, llamémosla M, es generado por el vector $(13, -31/4, 1)$. La dimensión del Kernel es uno. Por lo tanto, la imagen debe estar generada por dos vectores para que su dimensión sea dos.

El determinante de M es cero. Por lo tanto, la TL M apachurra el espacio $\mathbb{R}^3$ contra la imagen:

$$\mathrm{Im}(M) = x \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} + y \begin{pmatrix} 4 \\ 4 \\ 8 \end{pmatrix} + z \begin{pmatrix} 5 \\ -8 \\ -3 \end{pmatrix}$$

Estos tres vectores generan el subespacio $\mathrm{Im}(M)$, pero deben formar un conjunto LD. Los dos primeros vectores $\{(2, 4, 5), (3, 4, -8)\}$ forman un conjunto LI porque uno no es múltiplo del otro. Por lo tanto, esos dos vectores generan toda la imagen de M. Un conjunto generado por dos vectores es un plano, cuyo vector normal es el producto cruz de dos vectores no paralelos sobre el plano: el vector normal al plano es $(-52, 31, -4)$. La imagen de una TL es un EV, por lo tanto, el plano pasa por el origen. Por lo que la ecuación del plano es $-52x + 31y - 4z = 0$. En conclusión, M apachurra el espacio 3D sobre el plano Π cuya ecuación es $-52x + 31y - 4z = 0$, pero aún no sabemos qué es lo que M hace sobre el plano. Aprenderemos más acerca de esto en un próximo capítulo sobre diagonalización.

De otro lado, podemos decir que un sistema de la forma $MX = B$, con la M de este problema, a veces tiene solución y a veces no. El sistema tiene solución cuando $B \in \Pi$ y en ese caso la solución no es única: a cualquier solución podemos añadir el Kernel, el cual es la línea generada por $(13, -31/4, 1)$. Pero si B no está en Π, no hay solución y el sistema sería inconsistente.

Un sistema de ecuaciones es reducido por Gauss-Jordan para saber el número de ecuaciones independientes del sistema. Por lo que Gauss-Jordan es un procedimiento para descartar redundancias. De esa forma uno puede mecánicamente formar una base para el espacio dado. Además, sabemos que la dimensión del Kernel es igual al número de parámetros libres del conjunto solución al sistema de ecuaciones. Por otro lado, el rango es la dimensión de la imagen, la cual es también igual al número de vectores de cualquier base de la imagen. La suma de esos dos números es igual al número de incógnitas del sistema. Hay varias relaciones de ese tipo que envuelven una variedad de descriptores, tal como lo muestra el siguiente teorema y ejercicio.

390. $\diamond$ ***Teorema.*** *Si una matriz es sometida al proceso de depuración por Gauss-Jordan, el número de filas independientes es igual al número de columnas independientes.*

Demostración. Una matriz representa una LT $T\colon \mathbb{R}^n \to \mathbb{R}^m$. Sabemos que la dimensión del Kernel, ν, más la de la imagen ρ, es igual a la dimensión del dominio n. Tenemos $\nu + \rho = n$. Equivale a $n - \nu = \rho$. La primera cantidad, $n - \nu$, es el número de incógnitas menos el número de parámetros libres, que da ρ. ¿Pero qué es ρ? es la dimensión de la imagen o el número de columnas independientes. ■

391. $\boxed{\textbf{\textit{Gran ejercicio}}}$ *Supongamos que tenemos un sistema de ecuaciones, cuya matriz asociada sea M, con n incógnitas y que al hacer la reducción quedamos con i ecuaciones independientes. Desenmarañe varias relaciones que pueden incluir: el número de incógnitas, el número de parámetros libres, la dimensión del Kernel de M, la dimensión de la imagen de M, el número de filas independientes de M, el número de filas redundantes de M, el número de columnas independientes de M, el número de columnas redundantes de M.*

392. $\boxed{\textbf{\textit{Ejercicios para lucirse}}}$ *Usando el poder de la maquinaria desarrollada, estudie los siguientes sistemas de ecuaciones. Para cada TL asociada especifique el dominio, el rango, el espacio nulo y construya bases para dichos subespacios, verifique que $n = \nu + \rho$ y ofrezca una interpretación geométrica de la TL asociada:*

a) $$\begin{cases} 2x - y = 3 \\ 4x - 2y = 6 \end{cases}$$

b) $$\begin{cases} x - y = 1 \\ 4x - 4y = 4 \end{cases}$$

c) $$\begin{cases} -x + 4y + 2z = 25 \\ 2x - y - 5z = -10 \\ x + 3y - 3z = 15 \end{cases}$$

d) $$\begin{cases} 2x + 4y + 5z = 25 \\ 4x + 8y + 10z = 50 \\ 6x + 12y + 15z = 75 \end{cases}$$

e) $\begin{cases} 2x + 4y + 5z = 25 \\ 3x + 4y - 8z = -13 \\ 5x + 12y - 3z = 12 \end{cases}$

393. ***Tema de investigación para exponer.*** *Investigue los sistemas de Lindenmayer y los fractales. Puede empezar con el libro de Prusinkiewicz (1989). Internet también es muy rico en literatura sobre el tema, el cual puede explorarse en www.Google.com/. Para poder descifrar la carátula del libro correctamente necesita estudiar y entender este tema propuesto. Es un tema bellísimo.*

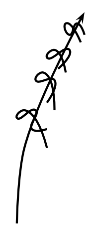


Figura 6.6. Arte y matemáticas.

6.3. Ejercicios de repaso

1. Definimos el **espacio columna** de una matriz M de n columnas y m filas como el espacio generado por las columnas de la matriz tomadas como vectores de $\mathbb{R}^m$. Demuestre que el espacio columna de M es igual a $\mathrm{Im}(T)$.

2. Sea A una matriz de tamaño $n \times n$. Demuestre que los vectores renglón de la matriz A son *linealmente dependientes* si y sólo si, los vectores columna de A son *linealmente dependientes.*

3. Si $T\colon \mathbb{R}^4 \to \mathbb{R}^3$ es una transformación lineal, ¿puede ser T inyectiva? (Justifique completamente su respuesta).

4. Encuentre el núcleo, la imagen, la nulidad y el rango de las siguientes transformaciones lineales:

   a) $T\colon \mathbb{R}^2 \to \mathbb{R}, \qquad T\begin{pmatrix} x \\ y \end{pmatrix} = x + y$

   b) $T\colon \mathbb{R}^3 \to \mathbb{R}, \qquad T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x + y + z$

   c) $T\colon \mathbb{R}^3 \to \mathbb{R}^2, \qquad T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y + z \\ x - y + 2z \end{pmatrix}$

Respuesta: a) Ker(T) =gen$\left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\}$, nulidad=1, Imagen($T$) $= \mathbb{R}$, Rango= 1,

5. Considere la matriz $A = \begin{pmatrix} a & b & c \\ 1 & 1 & 1 \end{pmatrix}$. Encontrar condiciones sobre $a, b,$ y c para que
   a) Rango de A sea igual a 1. b) Rango de A sea igual a 2.

6. Sea T: $P_3 \to P_2$ la transformación definida por $T(p(x)) = p'(x)$ donde $p'(x)$ denota la derivada de $p(x)$. Muestre que T es lineal, halle el núcleo e imagen de tal transformación y una base para el núcleo y una para la imagen. Invente una forma matricial de calcular $T(p(x))$, lo cual implica fabricar la base natural del dominio y transformarla por T.

6.4. Resumen

A cada paralelepípedo (plp) se le puede asociar su volumen. El determinante calcula el volumen y además da un signo (que puede interpretarse como orientación, sea por la mano derecha o la izquierda). Una TL es una función que transforma plps que salen del origen en plps que salen del origen. Representando una TL por su matriz asociada naturalmente, e identificando dicha matriz con un plp, tenemos que a cada TL le corresponde un número, su determinante. Su significado es, aparte del signo, el siguiente: el determinante de una TL dada T es el factor de ampliación de los volúmenes de los plps al ser transformados:

$$\mathrm{Det}(T(plp)) = \mathrm{Det}(T)\mathrm{Det}(plp)$$

Si el determinante de una TL es cero, la TL apachurra a todo n-plp. Por tanto, si un sistema de ecuaciones tiene a T como TL asociada, el sistema puede carecer de soluciones o bien tiene un número infinito de soluciones. La solución general a un sistema de ecuaciones puede descomponerse como la suma de una solución particular y de la solución al sistema homogéneo. Un sistema de ecuaciones siempre tiene solución y es única ssi el Det de la TL asociada es diferente de cero.

CAPÍTULO 7

LA MATRIZ INVERSA

La inversa de una transformación lineal es una función que deshace lo que la primera hizo. No toda transformación lineal tiene inversa. Pero cuando la inversa de una TL existe, podemos probar que es otra TL. Como una TL en dimensión finita es una función que puede representarse por una matriz, el problema de la TL inversa es el de la matriz inversa, es decir, el de encontrar otra matriz que al multiplicarla por la primera nos dé la identidad. Un algoritmo para el cálculo de la inversa puede verse directamente en el ejemplo 405.

394. ♣ ***Definición.*** *Si existe (como función) la inversa de una TL, $T: V \to W$, se denota $T^{-1}: W \to V$ la cual deshace lo que T hizo: $T^{-1}(T(\vec{X})) = \vec{X}$ y viceversa: $T(T^{-1}(\vec{Y})) = \vec{Y}$ donde $\vec{X}$ está en el dominio de T, que es V, y $\vec{Y}$ en su codominio, W.*

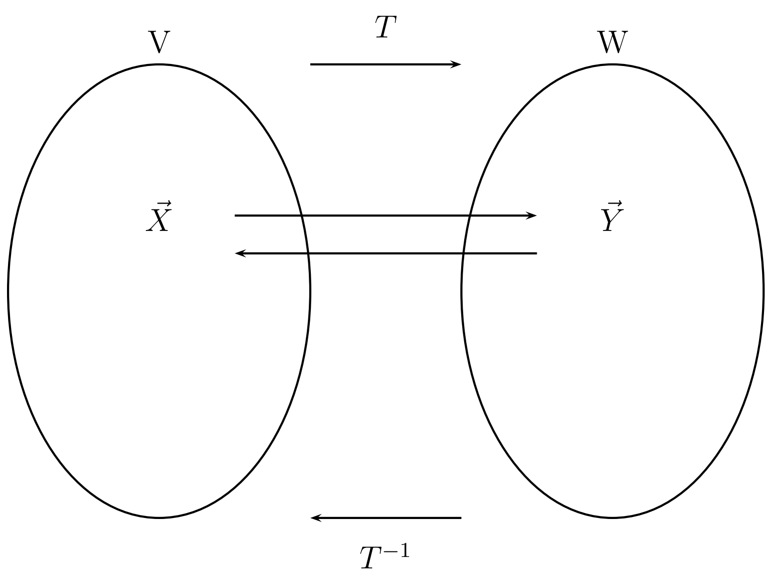


Figura 7.0. La transformación inversa.

395. ♣ ***Definición.*** *T es una TL invertible si existe su inversa como función.*

396. ***Ejemplo*** *$T(x) = 3x$ es una TL de $\mathbb{R}$ en $\mathbb{R}$. Ella multiplica por 3 un número dado. Por consiguiente, su inversa debe ser aquella función que divide por 3 a cualquier número: $T^{-1}(y) = (1/3)y$. En efecto, si las componemos, tenemos:*
$T^{-1}(T(x)) = T^{-1}(3x) = (1/3)(3x) = x$
y además,
$T(T^{-1}(y)) = T((1/3)y) = 3(1/3)y = y$.

Obsérvese que la inversa de una TL no necesariamente existe, pensemos en esto: si T denota una proyección, digamos, $T(x, y) = (x, 0)$, entonces, esta función asocia a cada vector su componente x dejando de lado su componente y. Por ejemplo, $T(3, 4) = (3, 0)$ y $T(3, 5) = (3, 0)$. Ahora preguntamos: ¿cuál es $T^{-1}(3, 0)$? Es $(3, 4)$ o es $(3, 5)$. Puesto que no se puede escoger un único valor, entonces T no tiene inversa. Pero si la inversa llegase a existir, de antemano se sabe que debe ser TL. En efecto:

397. ◇ ***Teorema.*** *Si T es una TL y es invertible, su inversa también es TL.*

Demostración. Sea T^{-1} la función inversa de T. Sea $T(\vec{x}) = \vec{z}, T(\vec{y}) = \vec{w}$, que es lo mismo que decir que: $T^{-1}(\vec{z}) = \vec{x}, T^{-1}(\vec{w}) = \vec{y}$.
Por tanto:
$$\begin{aligned} T^{-1}(\alpha\vec{z} + \beta\vec{w}) &= T^{-1}(\alpha T(\vec{x}) + \beta T(\vec{y})) = T^{-1}(T(\alpha\vec{x} + \beta\vec{y})) \\ &= \alpha\vec{x} + \beta\vec{y} = \alpha T^{-1}(\vec{z}) + \beta T^{-1}(\vec{w}), \end{aligned}$$
lo cual demuestre que T^{-1} reparte la suma y es compatible con la multiplicación escalar y por lo tanto es lineal. ■

398. ◇ ***Teorema y ejercicio.*** *La matriz de la TL inversa de M es una matriz M^{-1} tal que*
$MM^{-1} = MM^{-1} = \mathbb{I}$

Demostración: ejercicio.

7.1. El cálculo de la inversa

Para hallar la inversa hay varios caminos. Uno de ellos es plantear el producto de matrices dado en el teorema anterior y sacar n^2 ecuaciones y resolverlas. Eso es mucho trabajo. Existe un método mucho más fácil conocido como el algoritmo de Gauss-Jordan para la inversa.

Recordamos que el algoritmo de Gauss-Jordan para resolver un sistema de ecuaciones consiste en recobrar la matriz identidad. Por otra parte, un sistema de ecuaciones se escribe como $M\vec{X} = \vec{B}$, o sea que solucionarlo corresponde a hallar la imagen inversa de $\vec{B}$ por M. En efecto, si la inversa de M existe, multiplicando la ecuación a resolver a ambos lados por la inversa tenemos:

$M^{-1}M\vec{X} = M^{-1}\vec{B}$
$\mathbb{I}\vec{X} = M^{-1}\vec{B}$
$\vec{X} = M^{-1}\vec{B}$

Todo esto nos hace pensar que al resolver un sistema de ecuaciones con solución única, implícitamente se ha resuelto la inversa. Los siguientes teoremas nos dicen que nuestra intuición es correcta. Nuestro primer paso será demostrar que el algoritmo de Gauss-Jordan para resolver sistemas de ecuaciones puede automatizarse.

399. Ejemplo *Sea una matriz cualquiera*

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Verifiquemos que la operación de sumar las dos filas de M y poner el resultado en su segunda fila es equivalente a multiplicar por la izquierda a M por la matriz

$$S = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

Demostración.

$$SM = \begin{pmatrix} & & \vdots & a & b \\ & & \vdots & c & d \\ \dots & \dots & \dots & .. & .. \\ 1 & 0 & \vdots & a & b \\ 1 & 1 & \vdots & a+c & b+d \end{pmatrix}$$

Vemos que en la segunda fila aparece la suma de las dos primeras. ■

400. Ejercicio *Halle la matriz S que al operar sobre M resta las dos primeras filas y pone el resultado en la primera fila.*

401. ♣ Definición. *Se dice que G es un algoritmo lineal si G es una cadena de matrices o TL que cambia matrices en matrices.*

402. Ejemplo *Combinar ecuaciones para resolver un sistema de ecuaciones es un algoritmo lineal.*

403. ◇ Teorema. *Todo algoritmo lineal G que se aplica sobre una matriz M es equivalente a multiplicar a M por una matriz.*

Demostración. Una matriz es un arreglo rectangular que lo escribimos de esa forma para nuestra conveniencia. Pero también podríamos haberla escrito en forma lineal, formando un gran vector. Visto de esa forma, una matriz como un gran vector, usamos ahora el hecho que toda TL tiene una matriz asociada. Por tanto, toda TL sobre matrices tiene otra matriz y por consiguiente todo algoritmo lineal también tiene una matriz. Visto con más calma, toda operación entre filas para reducir una matriz tiene una matriz asociada, y al ir encadenando dichas operaciones lo que realmente hacemos es ir multiplicando las matrices asociadas correspondientes. Al terminar toda la reducción, tenemos un gran producto de matrices asociadas, la cual da una gran matriz, que es aquella a la que se refiere este teorema. ■

404. ⋄ ***Teorema de Gauss-Jordan para la inversa.*** *Un algoritmo lineal que transforma una matriz M en la identidad $\mathbb{I}$, también transforma a la identidad en la inversa. Para implementar este teorema se pone un arreglo que contenga la matriz M y a su lado la identidad, se transforma a M en la identidad y lo que se le haga a M se va haciendo directamente sobre $\mathbb{I}$. El resultado es que en vez de $\mathbb{I}$ queda M^{-1}. Eso se nota $(M|\mathbb{I}) \sim (\mathbb{I}|M^{-1})$.*

Demostración. Un algoritmo que transforma una matriz M en la identidad $\mathbb{I}$ equivale a multiplicar por una matriz, digamos H:

$HM = \mathbb{I}$

Por lo tanto $H = M^{-1}$.

El algoritmo se implementa con el método de la matriz aumentada, como podemos verlo en el ejemplo siguiente.

405. ***Ejemplo*** *Calculemos la matriz inversa de la siguiente matriz, en donde hemos dejado el término independiente del sistema del cual proviene:*

$$\begin{pmatrix} 3 & 4 & \vdots & 7 \\ 4 & 3 & \vdots & 7 \end{pmatrix}$$

Para ello cuadramos una supermatriz que contenga la identidad:

$$\begin{pmatrix} 3 & 4 & \vdots & 1 & 0 & \vdots & 7 \\ 4 & 3 & \vdots & 0 & 1 & \vdots & 7 \end{pmatrix}$$

Ahora transformamos la matriz M en la identidad y lo mismo que le hacemos a M le hacemos a la identidad que la acompaña en el centro del arreglo matricial. Cuando terminemos, en el lado izquierdo tendremos la identidad, en el centro la inversa de M y en el lado derecho la solución:

$$\begin{array}{r} 4R1 \rightarrow \\ 3R2 \rightarrow \end{array} \begin{pmatrix} 12 & 16 & \vdots & 4 & 0 & \vdots & 28 \\ 12 & 9 & \vdots & 0 & 3 & \vdots & 21 \end{pmatrix}$$

$$R1 - R2 \rightarrow \left(\begin{array}{cc:cc:c} 12 & 16 & 4 & 0 & 28 \\ 0 & 7 & 4 & -3 & 7 \end{array}\right)$$

$$\begin{array}{r} R1/4 \rightarrow \\ R2/7 \rightarrow \end{array} \left(\begin{array}{cc:cc:c} 3 & 4 & 1 & 0 & 7 \\ 0 & 1 & 4/7 & -3/7 & 1 \end{array}\right)$$

$$R1 - 4R2 \rightarrow \left(\begin{array}{cc:cc:c} 3 & 0 & -9/7 & 12/7 & 3 \\ 0 & 1 & 4/7 & -3/7 & 1 \end{array}\right)$$

$$R1/3 \rightarrow \left(\begin{array}{cc:cc:c} 1 & 0 & -3/7 & 4/7 & 1 \\ 0 & 1 & 4/7 & -3/7 & 1 \end{array}\right)$$

De acuerdo con la ética matricial, verificamos la respuesta. La solución $x = 1, y = 1$ se verifica directamente. Lo que queda por verificar es la inversa. Debemos multiplicar las dos matrices cuyo producto debe ser la identidad:

$$TS = \left(\begin{array}{ccccc} & & \vdots & -3/7 & 4/7 \\ & & \vdots & 4/7 & -3/7 \\ \ldots & \ldots & \ldots & .. & .. \\ 3 & 4 & \vdots & 7/7 & 0 \\ 4 & 3 & \vdots & 0 & 7/7 \end{array}\right)$$

que indica que el producto de las dos matrices es la identidad: la aritmética es correcta y la inversa es la que se dijo que era.

406. ***Ejemplo*** *Demos una mirada al ejercicio anterior desde otra perspectiva que nos permita entender mucho mejor el significado de un algoritmo lineal.*

Comencemos con un sistema de ecuaciones cuya matriz A es:

$$A = \begin{pmatrix} 3 & 4 \\ 4 & 3 \end{pmatrix}$$

Nuestro propósito es convertir la matriz del sistema en la identidad. Para lograrlo, la primera cosa que debemos hacer es multiplicar la primera fila por 4 y la segunda por 3. Para hacer esa tarea podemos multiplicar por la izquierda por la matriz

$$Z_1 = \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}$$

¿Cómo podemos estar seguros de que esta es la matriz correcta? Por verificación directa o también haciendo la tarea sobre la identidad. Recordemos que la tarea era multiplicar la primera fila por 4 y la segunda por 3: si hacemos esa operación sobre la identidad, obtenemos la matriz Z_1.

Después ponemos en la fila 2 a $R_1 - R_2$, fila uno menos fila dos. Eso se hace automáticamente multiplicando por

$$Z_2 = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$$

Después, dividimos la fila uno por 4 y la dos por 7. Esto produce

$$Z_3 = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/7 \end{pmatrix}$$

Ahora, ponemos sobre la fila uno, $R_1 - 4R_2$, una tarea que se hace con

$$Z_4 = \begin{pmatrix} 1 & -4 \\ 0 & 1 \end{pmatrix}$$

Finalmente dividimos la primera fila por

$$Z_5 = \begin{pmatrix} 1/3 & 0 \\ 0 & 1 \end{pmatrix}$$

A estas matrices, las Z, las llamamos matrices de Gauss.

Por eso, nosotros hacemos, en el orden correcto, una cadena de multiplicaciones que convierten la matriz A en la identidad:

$Z_5Z_4Z_3Z_2Z_1A = \mathbb{I}$

La cadena de multiplicaciones $Z_5Z_4Z_3Z_2Z_1$ es una matriz, llamémosla Z. Tenemos $ZA = \mathbb{I}$ y Z por tanto debe ser la inversa de A: $Z = A^{-1}$. Si aplicamos el mismo algoritmo sobre la identidad, tenemos todo: $ZI = Z = A^{-1}$.

Resumiendo: una cadena de multiplicaciones de matrices de Gauss que convierte una matriz A en la identidad, es en último término una multiplicación por una sola matriz que es la inversa de A. Si aplicamos el mismo algoritmo sobre la identidad, obtendremos, por supuesto, la matriz inversa A^{-1}.

407. ***Ejemplo*** *Hemos encontrado en el último ejemplo que fue necesario hacer una cadena de multiplicaciones, $Z_5Z_4Z_3Z_2Z_1$. Aprendamos un método eficiente para hacerlo en el orden correcto:*

$$\begin{pmatrix}
 & & \vdots & 4 & 0 \\
 & & \vdots & 0 & 3 \\
\cdots & \cdots & \cdots & .. & .. \\
1 & 0 & \vdots & 4 & 0 \\
1 & -1 & \vdots & 4 & -3 \\
\cdots & \cdots & \cdots & .. & .. \\
1/4 & 0 & \vdots & 1 & 0 \\
0 & 1/7 & \vdots & 4/7 & -3/7 \\
\cdots & \cdots & \cdots & .. & .. \\
1 & -4 & \vdots & -9/7 & 12/7 \\
0 & 1 & \vdots & 4/7 & -3/7 \\
\cdots & \cdots & \cdots & .. & .. \\
1/3 & 0 & \vdots & -3/7 & 4/7 \\
0 & 1 & \vdots & 4/7 & -3/7
\end{pmatrix}$$

La matriz que queda abajo a la derecha es la inversa de la que queda arriba a la derecha, lo cual ratifica lo visto en ejemplos anteriores.

408. $\boxed{\textbf{\textit{Ejercicio}}}$ *Utilice el algoritmo de Gauss-Jordan para hallar la inversa, si existe, de cada una de las siguientes matrices:*

a) $\begin{pmatrix} 1 & 4 \\ 2 & 7 \end{pmatrix}$

b) $\begin{pmatrix} 2 & 8 \\ 8 & 2 \end{pmatrix}$

c) $\begin{pmatrix} 0.2 & 0.8 \\ 0.8 & 0.2 \end{pmatrix}$

d) $\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$

Ayuda para la última matriz, ejecute las siguientes operaciones:

1. Ponga en la fila 1: $\cos\theta R_1 + \operatorname{sen}\theta R_2$ y use la identidad $\operatorname{sen}^2\theta + \cos^2\theta = 1$.
2. Ponga en la fila 2: $\operatorname{sen}\theta R_1 - R_2$ y cambie $\operatorname{sen}^2\theta - 1$ por $-\cos^2\theta$
3. Multiplique la fila 2 por $-1/\ \cos\theta$.

409. $\boxed{\textbf{\textit{Ejercicio}}}$ *Descomponga cada inversa hallada en el problema anterior en un producto de matrices de Gauss. Verifique su respuesta.*

410. *Intriga.* *El algoritmo de Gauss-Jordan para hallar la inversa podría ejecutarse de varias maneras: ¿Darán todas esas maneras la misma respuesta final? O en otras palabras, ¿es la inversa única? La respuesta es afirmativa. De igual forma que el inverso del número* 5 *es* 1/5 *y ningún otro, la inversa de una matriz es una sola (si existe) y todos los métodos dan la misma respuesta.*

411. ⋄ ***Teorema. La inversa es única.*** *Prueba: Sean B y C inversas de A, entonces*

$AB = BA = \mathbb{I}$ *y* $AC = CA = \mathbb{I}$.

Debemos probar que $B = C$. Puesto que $BA = \mathbb{I}$, entonces $BAC = IC = C$. Pero $AC = \mathbb{I}$, por tanto, $BAC = BI = B$. En conclusión, $B = C$.

412. **Ejercicio** *Nosotros utilizamos en el ejemplo previo el hecho de que el producto de matrices es asociativo. Esto significa que el producto de matrices está definido para dos matrices, y que por lo tanto, para multiplicar 3 matrices, debemos multiplicar las dos primeras, en el orden correcto, y el resultado multiplicarlo por la tercera. O podemos multiplicar la segunda por la tercera, en el orden correcto, y el resultado multiplicarlo por la primera. Resulta que ambos métodos dan lo mismo: a esto se llama asociatividad de la multiplicación de matrices. Escrito formalmente $ABC = A(BC) = (AB)C$ donde A, B, C son matrices que se pueden multiplicar en el orden señalado. Recalque dónde usamos la asociatividad en la demostración del teorema anterior. Demuestre tal propiedad: puede hacerse directamente sobre matrices o sobre composición de las TL.*

Cuando la matriz inversa existe, es única, lo cual significa que todos los métodos inventados o por inventar dan la misma única respuesta. Además del método de Gauss-Jordan existe otro método muy popular llamado de los cofactores. En la explicación que daremos, usamos la notación M_{ij} para denotar la entrada de la matriz M que queda en la fila i columna j. Además, usamos la **transpuesta** de una matriz definida así: si M tiene en su fila i columna j la entrada M_{ij} entonces la transpuesta de M, notada M^t, se define por $(M^t)_{ij} = M_{ji}$, es decir, la transpuesta de una matriz es la matriz reflejada con respecto a la diagonal. También, usamos **submatrices** que resultan de tachar en una matriz una determinada columna y una fila.

413. ***Ejemplo*** *Sea*

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

Tenemos que $M_{12} = b$, $M_{21} = d$.
Y la transpuesta de M es

$$M^t = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix}$$

Por otro lado, si tachamos en M la fila 2 y la columna 2 nos queda la submatriz

$$\begin{pmatrix} a & g \\ & \\ c & i \end{pmatrix}$$

414. ***La inversa por cofactores.*** *En el caso 2×2 apareció el determinante en el cálculo de la matriz inversa como indicador o negador de su existencia. Eso era de esperarse, pues si una matriz apachurra al cubo unitario, entonces no puede ser uno a uno, inyectiva, y a varios elementos del dominio le hace corresponder el mismo elemento del codominio. Por lo tanto, no puede tener inversa. Resulta que si la matriz inversa existe, todas sus entradas siempre pueden darse en términos de determinantes y de determinantes de submatrices, tal como lo explicamos enseguida:*

El algoritmo general para hallar la inversa por medio de subdeterminantes o cofactores es como sigue:

0. Una matriz que no es cuadrada, no tiene inversa. En efecto, una matriz no cuadrada apachurra o bien deja a alguien sin imagen.

1. Si la matriz M es cuadrada, se calcula su determinante. Si es cero, no hay inversa y el algoritmo se termina. Si el determinante es diferente de cero, la inversa existe y hay que buscarla.

2. Se fabrica una matriz C de cofactores $[C_{ij}]$. El cofactor C_{ij} es $(-1)^{i+j}$ multiplicado por el determinante de la matriz que queda de tachar la fila i y la columna j.

3. Se fabrica la matriz adjunta de M, notada $\text{Adj}(M)$ y que es la transpuesta de la matriz de cofactores. Si la matriz de cofactores es $[C_{ij}]$, su transpuesta es $[C_{ji}]$ que es la matriz adjunta.

4. La matriz inversa es: uno sobre el determinante de M por la matriz adjunta. O también: la matriz inversa es uno sobre el determinante por la matriz de cofactores transpuesta.

Para probar que este algoritmo funciona lo mejor es probar que $M(\text{Adj}(M)) = (\text{det}M)\mathbb{I}$. De lo cual se concluye que $M\frac{1}{\text{Det}M}\text{Adj}(M) = \mathbb{I}$, y como la inversa es única, se deduce que

$$M^{-1} = \frac{1}{\text{Det}M}\text{Adj}(M)$$

.

415. ***Ejemplo*** *Calculemos por el método de los cofactores la matriz inversa de*

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

1. Calculamos su determinante: $\text{Det}M = ad - bc$ y la matriz inversa existe si su determinante es diferente de cero, y en ese caso debemos calcularla.

2. Fabricamos la matriz de cofactores. El cofactor c_{11} es $(-1)^{i+j} = (-1)^{1+1} = 1$ multiplicado por el determinante que resulta de tachar la fila 1 y la columna 1. Al tachar lo que toque, queda la submatriz de un solo elemento d. Su determinante es d. Por lo tanto, $c_{11} = d$. El cofactor c_{12} es $(-1)^{i+j} = (-1)^{1+2} = -1$ multiplicado por el determinante que resulta de tachar la fila 1 y la columna 2. Se tacha y queda la submatriz de un solo elemento c. Su determinante es c. Por lo tanto, $c_{12} = -c$. El cofactor c_{21} es $(-1)^{i+j} = (-1)^{2+1} = -1$ multiplicado por el determinante que resulta de tachar la fila 2 y la columna 1. Se tacha y queda la submatriz de un solo elemento b. Su determinante es b. Por lo tanto, $c_{21} = -b$. El cofactor c_{22} es $(-1)^{i+j} = (-1)^{2+2} = 1$ multiplicado por el determinante que resulta de tachar la fila 2 y la columna 2. Se tacha y queda la submatriz de un solo elemento a. Su determinante es a. Por lo tanto, $c_{22} = a$.

3. La matriz de cofactores es:

$$C = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

4. La matriz inversa es: uno sobre el determinante de M por la matriz de cofactores transpuesta. La inversa es la siguiente:

$$M^{-1} = (1/DetM) \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

O bien, la matriz inversa es uno sobre el determinante por la adjunta, donde la adjunta es la matriz de cofactores transpuesta:

$$\text{Adj}(M) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Obsérvese que

$$M(\text{Adj}(M)) = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} = (\det M)\mathbb{I}.$$

En definitiva

416. $\boxed{\textbf{\textit{Ejercicio}}}$ *Tome las matrices siguientes e inviértalas por el método de cofactores. Verifique su respuesta multiplicando las dos matrices y recobrando la identidad.*

a) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$

b) $\begin{pmatrix} -1 & 0 & 1 \\ 2 & 0 & -1 \\ -2 & 3 & 0 \end{pmatrix}$

417. $\boxed{\textbf{\textit{Ejercicio}}}$ *¿Para qué valores de λ la matriz siguiente es invertible?*

$$\begin{pmatrix} 1-\lambda & 2 & 3 \\ 0 & 1-\lambda & 1 \\ 0 & 1 & -\lambda \end{pmatrix}$$

418. **Ejercicio de investigación** *Demuestre en general que el método de los cofactores funciona. Es decir, que en realidad sí produce la inversa. Lo que hay que hacer es demostrar que la inversa producida por el método de los cofactores multiplicada por la matriz dada produce la identidad. O más fácil, la matriz dada multiplicada por su adjunta da el determinante por la identidad.*

7.2. La descomposición *LU*

Ya sabemos cómo usar el método de Gauss-Jordan para invertir una matriz de tamaño 2×2. Uno intuye que para matrices más grandes todo lo que uno tiene que hacer es trabajar un poco más siguiendo la misma metodología. Pue sí, eso es cierto aunque no completamente. Lo que sucede es que, en general, la solución a un problema grande no es la solución agrandada de un problema pequeño. ¿Por qué? Porque siempre nacen, como mínimo, responsabilidades administrativas para manejar volúmenes grandes de información, por ejemplo, en relación con una memoria limitada. Eso es tan serio que hay toda una ciencia llamada análisis numérico (Burden y Faires, 1985) que trata tanto problemas de gran volumen de información como de problemas que no sabemos o no podemos resolver mediante fórmulas.

En esta sección desarrollaremos un ejemplo de encontrar la inversa de una matriz 3×3 por el método de Gauss-Jordan con el propósito de entender el teorema de descomposición LU que tiene utilidad en análisis numérico.

419. **Ejemplo** *Calculemos la matriz inversa de*

$$M = \begin{pmatrix} 1 & -1 & 1 \\ 0 & -1 & -2 \\ 1 & 1 & -1 \end{pmatrix}$$

Solución: planteamos la matriz aumentada:

$$\begin{pmatrix} 1 & -1 & 1 & \vdots & 1 & 0 & 0 \\ 0 & -1 & -2 & \vdots & 0 & 1 & 0 \\ 1 & 1 & -1 & \vdots & 0 & 0 & 1 \end{pmatrix}$$

Ahora transformamos la matriz M en la identidad y lo que le hacemos a M, se lo hacemos a la identidad:

$$\begin{array}{r} \\ \\ R_1 - R_3 \rightarrow \end{array} \begin{pmatrix} 1 & -1 & 1 & \vdots & 1 & 0 & 0 \\ 0 & -1 & -2 & \vdots & 0 & 1 & 0 \\ 0 & -2 & 2 & \vdots & 0 & 0 & -1 \end{pmatrix}$$

$$-R_2 \rightarrow \begin{pmatrix} 1 & -1 & 1 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 2 & \vdots & 0 & -1 & 0 \\ 0 & -2 & 2 & \vdots & 1 & 0 & -1 \end{pmatrix}$$

$$\begin{array}{r} \\ \\ 2R_2+R_3 \rightarrow \end{array} \begin{pmatrix} 1 & -1 & 1 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 2 & \vdots & 0 & -1 & 0 \\ 0 & 0 & 6 & \vdots & 1 & -2 & -1 \end{pmatrix}$$

$$\begin{array}{r} R_1+R_2 \rightarrow \\ \\ \\ \end{array} \begin{pmatrix} 1 & 0 & 3 & \vdots & 1 & -1 & 0 \\ 0 & 1 & 2 & \vdots & 0 & -1 & 0 \\ 0 & 0 & 6 & \vdots & 1 & -2 & -1 \end{pmatrix}$$

$$\begin{array}{r} 2R_1-R_3 \rightarrow \\ \\ \\ \end{array} \begin{pmatrix} 2 & 0 & 0 & \vdots & 1 & 0 & 1 \\ 0 & 1 & 2 & \vdots & 0 & -1 & 0 \\ 0 & 0 & 6 & \vdots & 1 & -2 & -1 \end{pmatrix}$$

$$\begin{array}{r} \\ 3R_2-R_3 \rightarrow \\ \\ \end{array} \begin{pmatrix} 2 & 0 & 0 & \vdots & 1 & 0 & 1 \\ 0 & 3 & 0 & \vdots & -1 & -1 & 1 \\ 0 & 0 & 6 & \vdots & 1 & -2 & 1 \end{pmatrix}$$

$$\begin{array}{r} (1/2)R_1 \rightarrow \\ (1/3)R_2 \rightarrow \\ (1/6)R_3 \rightarrow \end{array} \begin{pmatrix} 1 & 0 & 0 & \vdots & 1/2 & 0 & 1/2 \\ 0 & 1 & 0 & \vdots & -1/3 & -1/3 & 1/3 \\ 0 & 0 & 1 & \vdots & 1/6 & -2/6 & 1/6 \end{pmatrix}$$

Ahora codificamos cada transformación entre los renglones de M como una matriz que se halla aplicando sobre la identidad la operación dada:

La operación de poner en la tercera fila $R_1 - R_3$ corresponde a

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

La operación de poner en la segunda fila $-R_2$ corresponde a

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

La operación de poner en la tercera fila $2R_2 + R_3$ corresponde a

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

La operación de poner en la primera fila $R_1 + R_2$ corresponde a

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

La operación de poner en la primera fila $2R_1 - R_3$ corresponde a

$$\begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

La operación de poner en la segunda fila $3R_2 - R_3$ corresponde a

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

La operación de multiplicar la primera fila por $1/2$, la segunda por $1/3$, la tercera por $1/6$ corresponde a

$$\begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & -1 \\ 0 & 0 & 1/4 \end{pmatrix}$$

Ahora multiplicamos todas las matrices en el orden correcto:

$$\left(\begin{array}{rrrcrrr}
 & & & \vdots & 1 & 0 & 0 \\
 & & & \vdots & 0 & 1 & 0 \\
 & & & \vdots & 1 & 0 & -1 \\
\cdots & \cdots & \cdots & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\
1 & 0 & 0 & \vdots & 1 & 0 & 0 \\
0 & -1 & 0 & \vdots & 0 & -1 & 0 \\
0 & 0 & 1 & \vdots & 1 & 0 & -1 \\
\cdots & \cdots & \cdots & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\
1 & 0 & 0 & \vdots & 1 & 0 & 0 \\
0 & 1 & 0 & \vdots & 0 & -1 & 0 \\
0 & 2 & 1 & \vdots & 1 & -2 & -1 \\
\cdots & \cdots & \cdots & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\
1 & 1 & 0 & \vdots & 1 & -1 & 0 \\
0 & 1 & 0 & \vdots & 0 & -1 & 0 \\
0 & 0 & 1 & \vdots & 1 & -2 & -1 \\
\cdots & \cdots & \cdots & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\
2 & 0 & -1 & \vdots & 1 & 0 & 1 \\
0 & 1 & 0 & \vdots & 0 & -1 & 0 \\
0 & 0 & 1 & \vdots & 1 & -2 & -1 \\
\cdots & \cdots & \cdots & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\
1 & 0 & 0 & \vdots & 1 & 0 & 1 \\
0 & 3 & -1 & \vdots & -1 & -1 & 1 \\
0 & 0 & 1 & \vdots & 1 & -2 & -1 \\
\cdots & \cdots & \cdots & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\
1/2 & 0 & 0 & \vdots & 1/2 & 0 & 1/2 \\
0 & 1/3 & 0 & \vdots & -1/3 & -1/3 & 1/3 \\
0 & 0 & 1/6 & \vdots & 1/6 & -1/3 & -1/6
\end{array}\right)$$

Como de costumbre, la matriz inversa la leemos abajo a la derecha. Debido a que hay muchos cálculos, verificamos la respuesta, la matriz por su inversa debe dar la identidad:

$$\begin{pmatrix} & & & \vdots & 1/2 & 0 & 1/2 \\ & & & \vdots & -1/3 & -1/3 & 1/3 \\ & & & \vdots & 1/6 & -1/3 & -1/6 \\ ... & ... & ... & .. & .. & .. & .. \\ 1 & -1 & 1 & \vdots & 1 & 0 & 0 \\ 0 & -1 & -2 & \vdots & 0 & 1 & 0 \\ 1 & 1 & -1 & \vdots & 0 & 0 & 1 \end{pmatrix}$$

Ahora notemos que algunas matrices de Gauss que hemos encontrado son del tipo U, i.e., (*upper triangular*) triangular superior: todas sus entradas que no son cero están por encima de la diagonal o en ella. Y hay otras que son L (*lower triangular*) triangular inferior, cuyas entradas no nulas están por debajo de la diagonal o en ella. Así podemos formular el siguiente Teorema.

420. $\diamond$ ***Propuesta.*** *Cuando una matriz M es invertible, puede descomponerse como un producto de una matriz triangular inferior por una triangular superior.*

Demostración. Nuestro teorema se basa en una interpretación del ejemplo anterior en el cual demostramos que toda inversa puede escribire como un producto de matrices de Gauss. Ahora bien, nosotros elaboramos los cálculos de tal forma que primero trabajamos la parte abajo de la diagonal, asegurándonos que estuviese rellena de ceros. Eso se logró con matrices L. Y después, elaboramos la parte de arriba de la diagonal con matrices U. Cabe aclarar que también tocamos la diagonal, lo cual se hace con matrices que clasifican tanto de L como de U.

Por lo tanto, nuestro ordenado trabajo puede escribirse simbólicamente como

$M^{-1} = U_n...U_2U_1L_m...L_2L_1$

donde las L_i son matrices tipo L y las U_j son de tipo U.

Puede probarse que el producto de matrices tipo U es U y que el producto de matrices tipo L es L. Por tanto, todo puede comprimirse en

$M^{-1} = UL$

Ahora tomemos la inversa a cada lado. Tengamos presente que la inversa de un producto de matrices cambia el orden del producto, pues una matriz corresponde a una TL y cuando se componen las TL el camino inverso reversa el orden de aparición.

Si tomamos la inversa a ambos lados obtenemos

$M = (UL)^{-1} = L^{-1}U^{-1}$

pero como la inversa de una matriz tipo U existe y es U y la inversa de un matriz tipo L existe y es L, podemos escribir

$M = LU$

lo cual dice que la matriz se pudo descomponer como un producto de una matriz L por una U. ■

421. $\boxed{\textbf{\textit{Ejercicio}}}$ *Ilustre con ejemplos apropiados las aseveraciones gratuitas formuladas en la demostración de la propuesta anterior, cómo que una matriz triangular tiene inversa (Alerta: para ello se requiere que todo elemento de la diagonal sea diferente de cero, lo cual lo cumplen todas nuestras matrices. ¿Por qué?).*

422. ◇ ***Objeción a la propuesta de descomposición* LU.** *En el proceso de reducción de un sistema de ecuaciones pueden aparecer filas de ceros que habría que poner en útimo lugar o soluciones desordenadas que habría que ordenar. Eso se hace intercambiando filas. Esta operación es generada por una matriz que no es triangular. Para verlo, consideremos el siguiente ejemplo:*

Sea el sistema que indica una solución desordenada:

$$\begin{cases} 3y = 4 \\ 5x = 6 \end{cases}$$

que se escribe en forma matricial

$$\begin{pmatrix} 0 & 3 & \vdots & 4 \\ 5 & 0 & \vdots & 6 \end{pmatrix}$$

El sistema se ordena intercambiando el lugar de las filas. Esto se puede hacer por medio de la matriz

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

la cual no es una matriz triangular.

423. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre que la inversa de una matriz que intercambia renglones es su propia inversa.*

Aunque una matriz que intercambia renglones no es triangular, dicha matriz puede descomponerse como producto de matrices triangulares

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

El problema es que esta descomposición intercala matrices tipo U con matrices tipo L y por eso crea problemas en nuestra descomposición LU. Se ha optado por dejar estos casos de lado y lo ponemos como advertencia.

424. ◇ ***Teorema.*** *Cuando una matriz M es invertible y el proceso de inversión no involucra intercambio de filas, M puede descomponerse como un producto $M = LU$ donde L es una una matriz triangular inferior, U es una triangular superior y tanto L como U están libres de ceros en la diagonal.*

425. $\boxed{\textbf{\textit{Intriga}}}$ *El procedimiento de Gauss-Jordan de reducción de matrices puede ser aplicado a sistemas con matriz asociada que no es invertible. Eso parece implicar que el teorema de descomposición LU podría extenderse a matrices muy arbitrarias, pero ¿qué tanto?*

426. $\diamond$ ***Teorema.*** *Para una matriz A invertible se tiene que*

$$Det(A^{-1}) = (Det(A))^{-1}$$

lo cual dice que si A estira entonces A^{-1} encoge por el mismo factor y al revés.

Demostración. Por definición de inversa tenemos:

$A^{-1}A = \mathbb{I}$

Tomamos determinantes a ambos lados:

$Det(A^{-1}A) = Det(\mathbb{I}) = 1$

Como sabemos que el determinante de un producto de matrices es el producto de matrices (cuando se pueden multiplicar), entonces

$Det(A^{-1})Det(A) = 1$

es decir, $Det(A^{-1}) = (Det(A))^{-1}$. ■

Sellemos este capítulo con una demostración formal del hecho de que el algoritmo de Gauss-Jordan es conservativo.

427. $\diamond$ ***Teorema.*** *El algoritmo de Gauss-Jordan que lleva un sistema arbitrario de ecuaciones lineales a forma reducida escalonada es conservativo, es decir, no crea ni destruye soluciones, transformando sistemas en sistemas equivalentes que tienen las mismas soluciones.*

Demostración. Nosotros hemos utilizado el algoritmo de Gauss-Jordan para encontrar una única solución cuando existe o, en su defecto, para llevar al sistema hasta un lugar donde sea cómodo leer la solución. Dicha forma la hemos llamado escalonada reducida.

Toda operación de Gauss Jordan puede ser codificada como una matriz bien sea triangular superior o triangular inferior, ambas libres de ceros en la diagonal, o como una matriz de intercambio de renglones. Cualquiera de éstas matrices tiene inversa. Eso significa que no combinamos renglones a ojo cerrado sino que lo hacemos con la firme intención de no perder información.

Con toda esta maquinaria tan poderosa, demostrar que nuestro algoritmo conservativo es inmediato. Formalmente, lo que tenemos que demostrar es que $\vec{S}$ es solución del sistema $A\vec{X} = \vec{B}$ ssi $\vec{S}$ es solución del sistema $GA\vec{X} = G\vec{B}$ donde G es una matriz usada en el algoritmo.

Supongamos que $\vec{S}$ es solución del sistema $A\vec{X} = \vec{B}$. Eso implica que $A\vec{S} = \vec{B}$. Podemos mutiplicar a ambos lados por G y obtener $GA\vec{S} = G\vec{B}$, lo cual dice que $\vec{S}$ es solución del sistema transformado. Nuestro algoritmo no destruye soluciones.

Supongamos ahora que tenemos una solución al sistema transformado, es decir, que existe $\vec{S}$ tal que $GA\vec{S} = G\vec{B}$. Como cada G de las que nosotros usamos en nuestro algoritmo es invertible, podemos multiplicar por la inversa de G a ambos lados y obtener $G^{-1}(GA\vec{S}) = G^{-1}(G\vec{B})$. Aplicamos la asociatividad del producto matricial para obtener $(G^{-1}G)A\vec{S} = (G^{-1}G)\vec{B})$ que es equivalente a $\mathbb{I}A\vec{S} = \mathbb{I}\vec{S}$, es decir $A\vec{S} = \vec{S}$, lo cual expresa que si tenemos una solución al sistema transformado, dicha solución es también solución del sistema original. Así hemos demostrado que nuestro algoritmo no crea soluciones nuevas.

Puesto que nuestro algoritmo no crea ni destruye soluciones, es conservativo y sirve para hallar la soluciones de un sistema lineal de ecuaciones. ■

7.3. Ejercicios de repaso

1. *a*) Mostrar que la matriz $\begin{pmatrix} 2 & -3 \\ 5 & -7 \end{pmatrix}$ es invertible y encuentre su inversa.

   b) Escriba la solución al sistema
   $2x_1 - 3x_2 = 4$
   $5x_1 - 7x_2 = -3$

2. Sea A una matriz invertible y c un escalar, ¿es cierto que $\big(cA\big)^{-1} = \frac{1}{c}A^{-1}$?

3. Si $AB = 0$ y B es una matriz invertible, donde la matriz 0 denota la matriz que tiene cero en toda entrada, ¿es cierto que $A = 0$?

4. Si A y B son matrices cuadradas invertibles, ¿es cierto que $A+B$ es invertible y $(A+B)^{-1} = A^{-1} + B^{-1}$?

5. ¿Qué figura en el plano $\mathbb{R}^2$ forman los puntos (x, y) para los cuales la matriz $\begin{pmatrix} x & y & 1 \\ 1 & 2 & 1 \\ 3 & -3 & 1 \end{pmatrix}$ no es invertible?

6. Sea A una matriz invertible. Demuestre que si $A = \text{adj}(A)$, entonces $\text{Det}(A) = 1$.

7. Hallar A^{-1} si $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 5 \\ 0 & 0 & 4 \end{pmatrix}$. Expresar A^{-1} como un producto de matrices de Gauss y A como un producto LU.

8. Suponga que A es una matriz 3×3 cuyo espacio nulo es una recta que pasa por el origen en $\mathbb{R}^3$. ¿Es posible que $\text{Im}(A)$ sea también una recta que pasa por el origen? Explique claramente.

9. ¿Existen valores de r y s para los cuales el rango de A sea uno o dos, siendo $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r-2 & 2 \\ 0 & s-1 & r+2 \\ 0 & 0 & 3 \end{pmatrix}$?

10. Demuestre que si A es una matriz 3×4 entonces los vectores columna son linealmente dependientes. Extienda el enunciado anterior y su justificación a una matriz real de tamaño $m \times n$ con $n > m$.

11. Una matriz A cuadrada se dice idempotente si $A^2 = A$.

    a) Dar un ejemplo de una matriz idempotente no nula y distinta de la identidad.

    b) Mostrar que si A es a la vez idempotente e invertible, entonces A es la matriz identidad.

7.4. Resumen

Cuando una TL es invertible, su inversa es una TL y tiene una matriz, la matriz inversa. Dicha matriz puede calcularse por un algoritmo de Gauss-Jordan o bien por el método de los cofactores, que se basa en subdeterminantes.

CAPÍTULO 8

CAMBIO DE COORDENADAS

Una base permite la existencia de un sistema de coordenadas en un EV, cada vector tiene su conjunto de coordenadas que lo identifica y distingue de todos los demás. Pero un EV tiene un número infinito de bases: ¿Cuál de ellas es la mejor? En principio, no hay ninguna preferencial. Pero demostraremos que para problemas muy específicos hay bases que permiten soluciones muy sencillas. De hecho, hasta ahora hemos usado la **base natural** del espacio $\mathbb{R}^n$:

$N = \{(1,0,...,0),(0,1,0,...,0),(0,...,1)\} = \{\vec{e}_1, \vec{e}_2, ..., \vec{e}_k, ..., \vec{e}_n\}$, y con ella hemos trabajado siempre. Ahora entramos a considerar bases cualesquiera.

8.1. Bases arbitrarias

Una base $B = \{\vec{e}_i\}_{i=1,2,...,n}$ es un conjunto de vectores LI que genera todo el espacio, es decir, si $\vec{X} \in V$, existen escalares α_1, α_2, α_n tal que

$\vec{X} = \alpha_1\vec{e}_1 + ... + \alpha_n\vec{e}_n$

A los coeficientes α_i les llamamos las coordenadas del vector $\vec{X}$ en la base dada. Notamos al vector de coordenadas $(\alpha_1, ..., \alpha_n)$ como $\vec{X}_B$. Al notar las coordenadas de esa manera, se está asumiendo que uno ha fijado un orden sobre la base dada. A una base con un orden prefijado la llamamos **marco**, marco de coordenadas o marco de referencia. Pero usaremos marco y base como sinónimos cuyo significado será claro del contexto.

428. ***Ejemplo*** *El conjunto $B = \{(2,1),(2,-1)\}$ es una base para $\mathbb{R}^2$. Ese conjunto genera todo el plano y no tiene información redundante. La combinación lineal*

$1(2,1) - 2(2,-1) = (2,1) + (-4,2) = (2-4, 1+2) = (-2,3)$

nos dice que las coordenadas de $(-2,3)$ en la base B son el par ordenado $(1,-2)$. Podemos escribir

$$\begin{pmatrix} -2 \\ 3 \end{pmatrix}_B = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

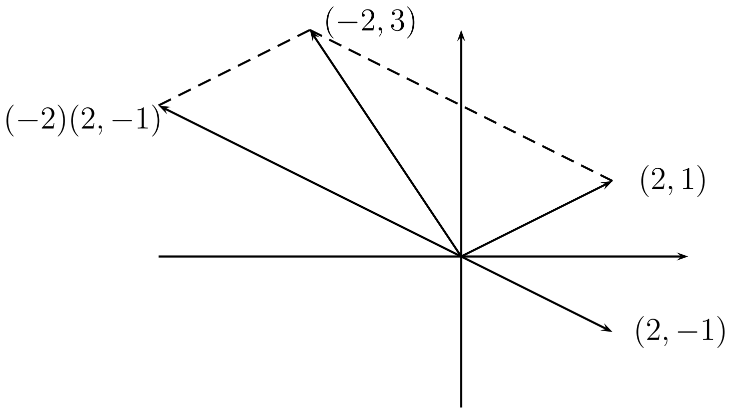


Figura 8.0. En la base $B = \{(2,1),(2,-1)\}$, $(-2,3)_B = (1,-2)$.

La noción de coordenadas tiene sentido solamente si uno ha numerado los vectores de la base en cuestión. Uno podría usar una notación diferente para las coordenadas de un vector y que no se confundan con un vector. De hecho, en algunos textos se usa la siguiente convención: los vectores se escriben entre paréntesis redondos y las coordenadas entre paréntesis cuadrados, como $[\alpha_1, ..., \alpha_n]$. Esta sabia medida será ignorada, la razón es que nos guiaremos por el contexto, sin el cual nada tiene sentido.

En general, las coordenadas de un vector y el vector serán diferentes. Pero en la base natural son lo mismo. Por ejemplo, el vector (x, y) puede descomponerse como

$(x, y) = x(1, 0) + y(0, 1)$

lo cual dice que en la base natural de $\mathbb{R}^2$ las coordenadas de (x, y) son (x, y).

Con frecuencia usaremos las coordenadas de un vector en operaciones matriciales. En tal caso, debemos asegurar que lo que se escriba tenga sentido: si las coordenadas de un vector van delante de una matriz (a su derecha), las coordenadas deberán aparecer en forma de columna. Pero si las coordenadas aparecen detrás de la matriz (a su izquierda) las coordenadas deben aparecer en forma de fila. De esa forma los productos tendrán sentido. Y todo eso será hecho sin muchas ceremonias. Además, y como parte del contexto, una base siempre está numerada. Y hay formas de numerar mejores que otras: ¡qué tal que numeráramos la base natural al revés!

Veamos ahora cómo se hace un cambio de coordenadas.

429. ***Ejemplo para memorizar el procedimiento*** *Sea* $B = \{(1, 1), (-1, 1)\}$*una base del plano* $\mathbb{R}^2$*. Para* $\vec{X}$ *tenemos:*

$\vec{X} = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2$

Si $\vec{X} = (x_1, x_2)$, las coordenadas de $\vec{X}$ en la base natural son (x_1, x_2). Esa ecuación puede escribirse en notación vertical como sigue:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1\alpha_1 - 1\alpha_2 \\ 1\alpha_1 + 1\alpha_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}.$$

Interpretemos dicha ecuación: tenemos una matriz que es una máquina de procesamiento que acepta coordenadas (α_1, α_2) en la base B y produce como producto

(x_1, x_2) que son las coordenadas de $\vec{X}$ en la base natural. Por lo tanto, la matriz que hemos encontrado es la matriz de cambio de coordenadas de la base B a la base natural N. Siguiendo la inteligente notación de Fisher (1970), notamos a esa matriz como

$I_N^B =$ **Matriz de cambio de coordenadas** de la base B a la base N.

Tenemos

$$I_N^B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Esta notación se justifica diciendo que un vector no cambia en absoluto por causa de un cambio de coordenadas, y como conserva su identidad, usamos la I. Una justificación más oficial será ofrecida más adelante comenzando con en el teorema 443.

430. ◇ ***Teorema y notación.*** *Para hallar las* ***coordenadas*** $X_N = (X)_N$ *de un vector en la base natural sabiendo sus coordenadas* $\vec{X}_B$ *en una base* B*, se usa una* ***matriz de paso o de cambio de base o de cambio de coordenadas*** I_N^B *que se construye poniendo la base* B *en notación vertical. Dicha matriz conserva la identidad del vector (por eso usamos* I*), acepta coordenadas en la base* B *y produce coordenadas en la base natural:*

$(X)_N = B(X)_B$

Observamos y recalcamos: la matriz de cambio de base o de cambio de coordenadas desde la base B a la base natural N es simplemente la base B en notación vertical. Es decir, se numera la base y por orden se pone cada vector de dicha base en posición vertical formando una matriz.

431. $\boxed{\textbf{\textit{Ejercicio}}}$ *Demuestre el teorema en su forma general.*

432. ◇ ***Teorema.*** *Las coordenadas de cualquier vector en una base* B *dada son únicas.*

Demostración. En la ecuación $(X)_N = B(X)_B$ podemos despejar $(X)_B$. En efecto, B es una base, y por lo tanto es un n-plp, cuyo volumen no es cero. Por tanto, B como matriz tiene inversa, B^{-1}. Multiplicando la citada ecuación por esta inversa, se tiene:

$B^{-1}(X)_N = B^{-1}B(X)_B = I(x)_B = (X)_B$.

Por tanto, sabiendo $X = X_N$, nosotros podemos encontrar inmediatamente $(X)_B$. Podemos hacer eso de una forma única porque la inversa de una matriz, si existe, es única. ■

433. ***Ejemplo*** *Sea* $B = \{(1,1), (-1,1)\}$*una base del plano* $\mathbb{R}^2$.

Si las coordenadas de un vector $\vec{v}$ en la base B son $(3,5)$ las coordenadas en la base natural son:

$$\vec{v}_N = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \times 3 + 1 \times 5 \\ 1 \times 3 - 1 \times 5 \end{pmatrix} = \begin{pmatrix} 8 \\ -2 \end{pmatrix}.$$

Verifiquemos que las coordenadas de $\vec{v}$ en la base B son $[3,5]$, por lo tanto,
$\vec{X} = 3\vec{e}_1 + 5\vec{e}_2$
lo cual también lo podemos escribir verticalmente:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 5\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1\times 3 + 1\times 5 \\ 1\times 3 - 1\times 5 \end{pmatrix} = \begin{pmatrix} 8 \\ -2 \end{pmatrix}.$$

434. $\boxed{\textbf{\textit{Ejercicio}}}$ *Si las coordenadas de un vector $\vec{v}$ en una base (ordenada) B son $(2,7)$, encuentre la matriz de cambio de base desde B a la base natural y las coordenadas del vector dado en la base N, si B es*

a) $B = \{(1,1),(2,-1)\}$

b) $B = \{(-1,1),(-3,-4)\}$

c) $B = \{(2,-2),(-1,-5)\}$.

435. ***Ejemplo*** *Dados $B_1 = \{(5,2),(-1,3)\}$ y $X_{B_1} = (2,3)$, encontremos X_{B_2} si $B_2 = \{(1,1),(2,5)\}$.*

Solución: $X = X_N = 2(5,2) + 3(-1,3) = (7,13)$.
Necesitamos encontrar X_{B_2}, es decir, expresar ese vector en la base B_2:
$(7,13) = \alpha(1,1) + \beta(2,5)$
Igualando coordenada por coordenada:
$7 = \alpha + 2\beta$
$13 = \alpha + 5\beta$
Restemos: $-6 = -3\beta$, lo que implica que $\beta = 2$ y $\alpha = 3$.
Respuesta: $X_{B_2} = (3,2)$.

436. $\boxed{\textbf{\textit{Ejercicio}}}$ *Dado $B_1 = \{(-1,-2),(-1,3)\}$ y $X_{B_1} = (5,7)$, encontrar X_{B_2} si*

a) $B_2 = \{(1,1),(2,5)\}$
b) $B_2 = \{(1,1),(-1,4)\}$
c) $B_2 = \{(1,1),(2,-3)\}$.

437. $\diamond$ ***Teorema.*** *Si una matriz de cambio de coordenadas, I_N^B, cambia de coordenadas desde la base B a coordenadas en la base N, la matriz inversa $(I_N^B)^{-1}$ cambia las coordenadas en la base N a coordenadas en la base B:*

$I_B^N = (I_N^B)^{-1}$
$(X)_B = I_B^N (X)_N = (I_N^B)^{-1}(X)_N$.

438. $\boxed{\textbf{\textit{Ejercicio}}}$ *Pruebe el teorema.*

439. $\boxed{\textbf{\textit{Ejercicio}}}$ *Si las coordenadas del vector $\vec{v}$ en la base natural son $(2, 7)$, encuentre la matriz de cambio de coordenadas de N a B y las coordenadas del vector dado en la base B, si B es*

a) $B = \{(1, 1), (2, -1)\}$
b) $B = \{(-1, 1), (-3, -4)\}$
c) $B = \{(2, -2), (-1, -5)\}$.

440. ***Ejemplo*** *Si $B_1 = \{(5, 2), (-1, 3)\}$ y $B_2 = \{(1, 1), (2, 5)\}$, encontremos $I_{B_2}^{B_1}$.*

Solución: El trabajo de $I_{B_2}^{B_1}$ es traducir las coordenadas en la base B_1 a coordenadas en la base B_2: $X_{B_2} = I_{B_2}^{B_1} X_{B_1}$. Sea

$$I_{B_2}^{B_1} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}$$

Sea X el primer vector de B_1, $X = (5, 2)$, entonces las coordenadas de X en la base B_1 son $X_{B_1} = (1, 0)$, porque $(5, 2) = 1(5, 2) + 0(-1, 3)$.

Tenemos: $X_{B_2} = (5, 2)_{B_2}$ y
$X_{B_2} = I_{B_2}^{B_1} X_{B_1}$
ó

$$X_{B_2} = (5, 2)_{B_2} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

de tal forma que $(5, 2)_{B_2} = (\alpha, \beta)$, y por tanto $(5, 2) = \alpha(1, 1) + \beta(2, 5)$. Por consiguiente
$5 = \alpha + 2\beta$
$2 = \alpha + 5\beta$.
Restando, $3 = -3\beta$,
$\beta = -1$,
$\alpha = 2 - 5\beta = 2 + 5 = 7$.

Para encontrar γ y δ, analizamos el segundo vector de B_1, $(-1, 3)$, cuyas coordenadas $(-1, 3)_{B_1}$ en B_1 son $(0, 1)$.

Sea $X = (-1, 3)$, entonces $X_{B_1} = (0, 1)$. Tenemos: $X_{B_2} = (-1, 3)_{B_2}$ y

$$X_{B_2} = (-1, 3)_{B_2} = I_{B_2}^{B_1} X_{B_1} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \gamma \\ \delta \end{pmatrix}$$

así que $(-1, 3) = \gamma(1, 1) + \delta(2, 5)$. Por consiguiente
$-1 = \gamma + 2\delta$
$3 = \gamma + 5\delta$.
Restando, $-4 = -3\delta$,
$\delta = 4/3$,
$\gamma = 3 - 5\gamma = 3 - 5(4/3) = 3 - 20/3 = -11/3$.
En conclusión:

$$I_{B_2}^{B_1} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} = \begin{pmatrix} 7 & -11/3 \\ -1 & 4/3 \end{pmatrix}$$

441. **Ejercicio** *Encontrar* $I_{B_2}^{B_1}$ *si* $B_1 = \{(4,2),(-1,2)\}$ *y*

a) $B_2 = \{(4,2),(1,5)\}$.
a) $B_2 = \{(1,-2),(1,2)\}$.
a) $B_2 = \{(-3,2),(3,5)\}$.

442. **Ejercicio** *Si*

$$I_{B_2}^{B_1} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} = \begin{pmatrix} 7 & -11/3 \\ -1 & 4/3 \end{pmatrix}$$

a) Encontrar B_1 si $B_2 = \{(1,1),(2,5)\}$,
b) Encontrar B_2 si $B_1 = \{(5,2),(-1,3)\}$.

Usamos la notación $I_{B_2}^{B_1}$ para denotar la matriz de cambio de coordenadas de la base B_1 a la base B_2. La I indica dos cosas. Primera, que una matriz de cambio de base conserva la identidad del vector, es decir, que el vector no es transformado sino que lo único que cambia es su descomposición debido a que se cambia de base. Y segunda, porque $I_{B_2}^{B_1}$ es la matriz de la TL, I: $V \to V$ cuando en el dominio se usa la base B_1 y en el codominio la base B_2. En la base natural N a ambos lados, la matriz de I es la matriz identidad $\mathbb{I} = I_N^N$. ¿Qué pasa si cambiamos de base en el dominio o en el codominio o en ambos lados? Veamos:

443. $\diamond$ **Teorema.** *Una matriz de cambio de coordenadas $I_{B_2}^{B_1}$ de la base B_1 a la base B_2 es la matriz de la TL identidad I tomando B_1 como base en el dominio y a B_2 como la base en el codominio.*

Predemostración. $I_{B_2}^{B_1}$ no puede cambiar la identidad de ningún vector. La matriz simplemente toma las coordenadas de cualquier vector en la base del dominio y las reescribe en la base del codominio, es una matriz de cambio de base. Muy pronto veremos una maquinaria poderosa para probar rigurosamente este y otros teoremas.

444. $\diamond$ **Teorema.** *Para encontrar la matriz $I_{B_2}^{B_1}$ de cambio de coordenadas de la base B_1 a la base B_2, es suficiente cambiar de coordenadas de la base B_1 a la base N y después de la base Na la base B_2. Esto es, uno debe multiplicar ambas matrices,* $I_{B_2}^{N} I_N^{B_1}$, en el orden correcto. *La matriz $I_N^{B_1}$ es la base B_1 en notación vertical.*

445. **Ejercicio** *Pruebe el teorema anterior.*

446. **Ejercicio** *Diseñe un procedimiento de Gauss-Jordan para encontrar la matriz de cambio de base $I_{B_2}^{B_1}$.*

447. **Ejercicio** *Si* $[-1,2],[-3,5]$ *son las coordenadas de dos vectores en la base* B_1*, encuentre la matriz de cambio de coordenadas de la base* B_1 *a la base* B_2 *y las coordenadas en* B_2 *de cada uno de los vectores, si las bases son:*

a) $B_1 = \{(1,1),(2,-1)\}, B_2 = \{(-1,1),(-3,-4)\}$
b) $B_1 = \{(-1,1),(-3,-4)\}, B_2 = \{(2,-2),(-1,-5)\}$
c) $B_1 = \{(2,-2),(-1,-5)\}, B_2 = \{(1,1),(2,-1)\}$.

448. **Ejercicio** *Si* $[-1,2,3],[-3,5,6],[4,1,-1]$ *son las coordenadas de tres vectores en la base* B_1*, encuentre la matriz de cambio de base de* B_1 *a* B_2 *y las coordenadas en* B_2 *de cada uno de dichos vectores si las bases son:*

$B_1 = \{(2,2,1),(-1,5,6)(-2,-2,3\}$
$B_2 = \{(1,1,1),(2,-1,1),(-1,-1,2)\}$.

449. **Ejemplo** *Dado* $M = I_{B_2}^{B_1}$ *y* B_1*, encontremos* B_2*.*

Primera solución: $M = I_{B_2}^{B_1} = I_{B_2}^{N} I_{N}^{B_1} = (B_2)^{-1}B_1$.
Por tanto, $(B_2)^{-1}B_1 = M$ y multiplicando en ambos lados a la izquierda por B_2 obtenemos
$B_2(B_2)^{-1}B_1 = B_2M$ que es lo mismo que $B_1 = B_2M$. Multiplicando por M^{-1} a la derecha obtenemos $B_1M^{-1} = B_2MM^{-1} = B_2$
Segunda solución:
$B_2 = I_N^{B_2} = I_N^{B_1} I_{B_1}^{B_2} = B_1(I_{B_2}^{B_1})^{-1} = B_1M^{-1}$.

450. **Ejercicio** *Encontrar* B_1 *si* $B_2 = \{(1,1),(2,5)\}$ *y*

$$I_{B_2}^{B_1} = \begin{pmatrix} 7 & -11/3 \\ -1 & 4/3 \end{pmatrix}.$$

451. **Ejercicio** *Encontrar* B_2 *si* $B_1 = \{(5,2),(-1,3)\}$ *y*

$$I_{B_2}^{B_1} = \begin{pmatrix} 7 & -11/3 \\ -1 & 4/3 \end{pmatrix}.$$

452. **Ejercicio** *a) Dado* $M = I_{B_2}^{B_1}$ *y* B_2*, encontrar* B_1*; b) Dado* $M = I_{B_1}^{B_2}$ *y* B_1*, encontrar* B_2*.*

8.2. Rotaciones en $\mathbb{R}^2$

En esta sección hallamos la matriz de una rotación en $\mathbb{R}^2$, lo cual será muy importante en lo que sigue.

453. ♣ ***Definición.*** *Una* ***rotación*** *es una* TL *que conserva ángulos y normas.*

454. ***Ejemplo*** *Encontremos la matriz que en $\mathbb{R}^2$ causa una rotación contraria a las manecillas del reloj por un ángulo θ.*

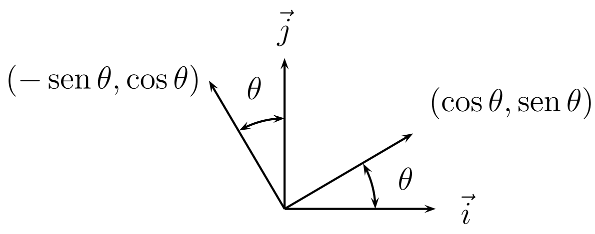


Figura 8.1. La base natural ha sido rotada un ángulo θ.

Para saber qué hace una TL es suficiente saber qué hace ella sobre una base cualquiera. Tomemos la base natural. El vector $\vec{i} = (1,0)$ es rotado pero no alargado, por lo tanto, se cambia en $(\cos\theta, \operatorname{sen}\theta)$, mientras que el vector $\vec{j} = (0,1)$ se cambia en $(-\operatorname{sen}\theta, \cos\theta)$. Por lo tanto, la matriz R de esa rotación es:

$$R = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

En el caso particular en que $\theta = \pi/2$ ó $90°$, la matriz de rotación es

$$R(\pi/2) = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos(\pi/2) & -\operatorname{sen}(\pi/2) \\ \operatorname{sen}(\pi/2) & \cos(\pi/2) \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Verifiquemos que esta rotación convierte $\vec{i} = (1,0)$ en $\vec{j} = (0,1)$:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

455. $\boxed{\textit{Ejercicio}}$ *Calcule el determinante de una rotación y explique el porqué de la respuesta. ¿Es toda TL con $Det(T) = 1$ una rotación? Ofrezca ejemplos y contraejemplos. ¿Qué más se necesita para que una TL con determinante 1 sea una rotación?*

456. $\boxed{\textit{Ejercicio}}$ *Halle la matriz inversa de una rotación de ángulo θ y demuestre que es igual a la matriz de rotación de ángulo $-\theta$.*

8.3. Números que rotan

Ofrecemos una construcción de los números complejos que nos enseña que las matemáticas son una combinación de arte y de ingeniería.

Los números reales pueden dibujarse sobre una recta. Dicho de otra forma, a una recta puede dársele una estructura aritmética, gracias a la cual los elementos de la línea se puedan sumar, restar, multiplicar, dividir, sacar raíz cuarta, logaritmo y hallar el seno y el coseno y todo lo demás. Pues bien, ¿qué tiene la línea que el plano no tenga? ¿qué nos impide darle una estructura aritmética al plano?

Nada.

Veamos. Cada punto del plano se direcciona con dos coordenadas (x, y) donde x, y son reales, tal como lo hemos hecho siempre. Podemos definir la suma coordenada por coordenada, como de costumbre: $(x_1, y_2) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$.

457. **_Ejercicio_** *Demuestre que el plano con la suma es un grupo conmutativo, es decir, que podemos sumar sin problema y que en la suma el orden no importa.*

Teniendo solucionado lo de la suma y la resta, nos queda la multiplicación. Para definirla, nos basamos en que los puntos del plano son vectores, de tal forma que la multiplicación de un vector (x, y) por un escalar real λ es $(\lambda x, \lambda y)$. Dicho de otra forma, ya sabemos cómo multiplicar un número real por un punto del plano. Para definir la multiplicación de un punto del plano por otro punto del plano nos basamos en algo que todos sabemos:

458. **_Ejercicio_** *Demuestre que el plano con la suma y producto escalar son un espacio vectorial sobre los reales, lo cual quiere decir que los escalares son los números reales.*

¿Qué significa que el plano tenga estructura de espacio vectorial? Significa que existen TL del plano en el plano. Y en general, las TL se componen y al ser representadas por matrices, la composición es igual al producto de matrices. Entonces lo que tenemos que hacer es definir una correspondencia entre cada punto del plano y una TL o matriz de tal manera que la multiplicación entre puntos del plano quede definida por el producto de las matrices que lo representan.

La elección de representar un punto del plano por una matriz se ha resuelto por medio de ligeras modificaciones de las rotaciones. Eso se hace gracias a las **coordenadas polares**. Dichas coordenadas definen un punto del plano por dos valores: la longitud, módulo o norma del radio vector que va del origen, el complejo $(0, 0)$, al punto dado, y el ángulo polar que dicho radiovector forma con la parte positiva del eje X en dirección contraria a las manecillas del reloj. Debemos advertir que el ángulo polar no se puede definir de manera única: dado un ángulo polar, si se le suma cualquier número de vueltas, o sea $2n\pi$, con $n \in \mathbb{Z}$, tenemos el mismo punto del plano.

En una palabra, podemos representar un punto del plano (x, y) como $(r, \theta)_p$ donde $r = \sqrt{x^2 + y^2}$ y θ es el ángulo cuya tangente es y/x y el subíndice p nos permite reconocer que se trata de coordenadas polares.

459. **_Ejemplos_** *El punto* $(1, 0)$ *se representa como* $(1, 0)_p$, *el punto* $(-1, 0)$ *como* $(1, \pi)_p$, *el punto* $(0, 1)$ *como* $(1, \pi/2)_p$ *y el punto* $(0, -1)$ *como* $(1, 3\pi/2)_p$.

Ahora, a cada punto del plano $(r,\theta)_p$ le hacemos corresponder la matriz dada por

$$(r,\theta)_p \leftrightarrow \begin{pmatrix} r\,cos\theta & -rsen\theta \\ rsen\theta & r\,cos\theta \end{pmatrix}$$

460. **Ejemplos** *El punto $(1,0)$ tiene módulo 1 y ángulo polar 0, por tanto*

$$(1,0) \leftrightarrow \begin{pmatrix} cos(0) & -sen(0) \\ sen(0) & cos(0) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

El número complejo i tiene módulo 1 y ángulo polar $\pi/2$, por lo tanto, es la matriz

$$(0,1) \leftrightarrow \begin{pmatrix} cos(\pi/2) & -sen(\pi/2) \\ sen(\pi/2) & cos(\pi/2) \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Si se componen dos rotaciones, obtenemos una rotación:

461. **Ejercicio** *Demuestre que si se multiplican dos puntos del plano $(r_1,\theta_1)_p$ y $(r_2,\theta_2)_p$, usando la representación matricial y multiplicando en cualquier orden las matrices respectivas, se obtiene el punto del plano $(r_1+r_2,\theta_1+\theta_2)_p$.*

462. **Ejercicio** *Demuestre que la multiplicación es cerrada, conmutativa, asociativa, distribuye la suma, que $(1,0)$ es el elemento neutro de la multiplicación, es decir, que $(1,0)$ multiplicado por (x,y) da (x,y).*

Por todas las propiedades aritméticas demostradas, de ahora en adelante llamaremos a los puntos del plano como **complejos o números complejos** y al plano mismo como **plano complejo** y se nota $\mathbb{C}$.

Ha resultado muy útil la notación polinómica en la cual (x,y) se representa por el polinomio $x+iy$ donde el número complejo i es el indicador de la dirección vertical, es decir $(0,1)$ o bien $(1,\pi/2)_p$. De igual forma, debido a que $(1,0)$ es el elemento neutro de la multiplicación, se le llama 1 y de esa manera los números reales se convierten en un subconjunto de los complejos y por ello a los reales se les pinta sobre el horizontal, cuyos puntos son de la forma $(x,0)$.

463. **Ejercicio** *Verifique que con $i=(0,1)=(1,\pi/2)_p$ se tiene que $i^2=ii=-1$.*

Ahora podemos multiplicar números complejos como si fuesen polinomios. Si un número complejo es $w=(a,b)=a+bi$ y el otro es $z=c+di$ entonces su producto es $wz=(a+bi)(c+di)=ac+adi+dci+bdi^2=ac+(ad+dc)i-bd=ac-bd+i(ad+dc)=$
$=(ac-bd,ad+dc)$.

Habíamos dicho que el plano complejo con la multiplicación escalar con números reales es un EV, pero en vez de escalares reales se pueden tomar escalares complejos.

464. **Ejercicio** *Demuestre que los complejos son un espacio vectorial sobre los complejos.*

465. **Ejercicio** *Demuestre que los complejos sobre los reales son un EV de dimensión* dos, *pero los complejos sobre los complejos tienen dimensión* uno.

De la multiplicación se deriva la división: dividir significa multiplicar por el inverso multiplicativo.

466. $\boxed{\textbf{Ejercicio}}$ *Halle el inverso multiplicativo de $(r,\theta)_p$. Ayuda: el inverso de una rotación es la rotación al revés. El inverso de multiplicar todo el mundo por un escalar no nulo r es multiplicar todo el mundo por $1/r$.*

Existe una notación de los números complejos que da inmediatamente su estructura multiplicativa: si $z \in \mathbb{C}$, es decir, si z es un punto del plano complejo, entonces $z = re^{i\theta}$ donde r es el módulo de z y θ es el ángulo polar.

Si tenemos dos complejos
$z = re^{i\theta}$
$w = se^{i\phi}$
entonces la multiplicación simplemente multiplica los módulos y suma los ángulos:
$zw = re^{i\theta}se^{i\phi} = rse^{i(\theta+\phi)}$.
Se dice que un número es unitario si tiene módulo uno. En tal caso:
$z = e^{i\theta}$
Multipliquemos un número cualquiera $w = se^{i\phi}$ por uno unitario $z = e^{i\theta}$:
$zw = e^{i\theta}se^{i\phi} = se^{i(\theta+\phi)}$
lo que pasa es que el número w sufre una rotación. Por supuesto que el radio no varía.

467. $\boxed{\textbf{Investigación}}$ *Hemos dicho que un complejo $(r,\theta)_p$ puede notarse como $re^{i\theta}$. En realidad, nuestra notación se basa en que existe la forma de definir la función exponencial compleja, y al evaluar $re^{i\theta}$ da $(r,\theta)_p$ o bien $rcos\theta + isen\theta$. Investigue, pues, cómo se define la exponencial compleja y cómo se prueba que $re^{i\theta} = rcos\theta + isen\theta$. Se puede encontrar una introducción al tema en (Steward, 2003).*

El módulo de un complejo se lee directamente en su expresión polar pero también puede recuperarse mediante la siguiente estrategia:

Definimos el conjugado de $z = re^{i\theta}$ como $\bar{z} = re^{-i\theta}$, es decir $\bar{z}$ tiene el mismo módulo que z pero el ángulo está medido al revés, en dirección negativa, en la misma dirección que las manecillas del reloj. A veces resulta más cómodo denotar el conjugado de z por z^*, por ejemplo, cuando el complejo viene con coordenadas que vienen de cálculos complicados. Ahora bien:
$z\bar{z} = re^{i\theta}re^{-i\theta} = r^2e^{i(\theta-\theta)} = r^2e^{i(0)}$,
el resultado que nos da es un número complejo que tiene norma r^2 y ángulo polar cero. Como tiene ángulo polar cero está sobre la parte positiva del eje X por lo que es un número real y como tiene módulo r^2, pues es en definitiva el número real r^2. Por consiguiente, el módulo cuadrado, largo cuadrado o norma cuadrado de un complejo es $z\bar{z}$.

Los números complejos tienen aplicaciones importantes. Por ejemplo, si uno tiene un cuerpo que rota alrededor de otro, la proyección sobre el eje X da una onda. Por esa razón, los complejos, las rotaciones y las ondas están íntimamente relaciondas. Eso se utiliza para facilitar cálculos, digamos, del comportamiento de sistemas ondulatorios sometidos a fuerzas también ondulatorias usando la transformada de Laplace (Zill, 2008). La relación de los complejos con las ondas también ha servido de base en mecánica cuántica para postular que la física fundamental también está ligada a los complejos. La razón es que tanto la radiación como la materia tienen propiedades de onda, lo cual se evidencia cuando se crea interferencia (Feynman, 1998).

8.4. Bases ortonormales

Al rotar la base natural uno obtiene una nueva base que se parece muchísimo a la natural: sus elementos son mutuamente perpendiculares y tienen norma uno. Este tipo de bases es muy útil y recibe un nombre:

468. ♣ ***Definición.*** *Una* ***base es ortogonal*** *cuando sus elementos son mutuamente perpendiculares. Ejemplo: la base natural es ortogonal. Toda rotación de la base natural es una base ortogonal. Si se toma la base natural y se rota y si cada vector es alargado positivamente, se obtiene una base ortogonal.*

469. ♣ ***Definición.*** *Una* ***base es ortonormal*** *cuando es ortogonal y cuando cada uno de sus vectores es unitario, es decir, tiene norma uno. Ejemplo: la rotación de la base natural produce una base ortonormal. Para convertir una base ortogonal en una ortonormal hay que dividir cada vector por su norma. Aunque nadie lo diga, es mejor numerar las bases ortonormales de tal forma que su Det sea* más uno. *Si uno tiene una base ortonormal con determinante menos uno, simplemente se intercambia un par de vectores y la nueva numeración da una base con Det = +1. O uno puede reversar un elemento de la base.*

470. ***Ejemplo*** *La base* $B_1 = \{(1,-1),(1,1)\}$*es ortogonal porque el producto punto* $(1,-1)\cdot(1,1) = 0$*, pero no es una base ortonormal pues la norma de cada uno de sus elementos es* $\sqrt{2}$*. Pero obtenemos una base ortonormal si dividimos cada vector por su norma* $\sqrt{2}$*. La base resultante es* $B_2 = \{(\sqrt{2}/2, -\sqrt{2}/2), (\sqrt{2}/2, \sqrt{2}/2)\}$.

471. ***Ejemplo*** *Consideremos el plano* Π *en el espacio 3D tal que sus puntos* (x,y,z) *satisfagan* $-x-2y+3z=0$*. Este conjunto es un EV pues es un plano que pasa por el origen. Este plano tiene dimensión dos, pues es generado por un conjunto LI con dos elementos:*

$-x-2y+3z=0$ implica que $x=-2y+3z$ y por tanto un punto $(x,y,z)\in\Pi$ si y sólo si

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2y+3z \\ y \\ z \end{pmatrix}$$

$$= \begin{pmatrix} -2y \\ y \\ 0 \end{pmatrix} + \begin{pmatrix} 3z \\ 0 \\ z \end{pmatrix} = y\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$

Es decir, el conjunto $B_1 = \{\vec{v}_1 = (-2,1,0), \vec{v}_2 = (3,0,1)\}$ es una base del plano, pero no es una base ortonormal porque $(-2,1,0)\cdot(3,0,1)$ no es cero.

Encontremos una base ortonormal para Π. Observemos que la norma del vector $(-2,1,0)$ es $\sqrt{5}$. Por tanto, podemos tomar $\vec{u}_1 = (-2\sqrt{5}/5, \sqrt{5}/5, 0)$ como un primer vector de nuestra base ortonormal. El segundo vector $\vec{u}_2$ puede ser construido de varias maneras. La primera es proyectar el segundo vector de la base original, $\vec{v}_2$ sobre $\vec{u}_1$. Esa proyección se sustrae de $\vec{v}_2$ y luego se normaliza para obtener $\vec{u}_2$.

O también, podemos notar que $(-1,-2,3)$ es normal al plano y que tiene norma $\sqrt{14}$ y que por tanto, la normal unitaria $\vec{n}$ es

$$\vec{n} = (-\sqrt{14}/14, -2\sqrt{14}/14, 3\sqrt{14}/14)$$

El segundo vector $\vec{u}_2$ de la base ortonormal del plano puede ser tomado como:
$\vec{u}_2 = \vec{u}_1 \times \vec{n} = (-2\sqrt{5}/5, \sqrt{5}/5, 0) \times (-\sqrt{14}/14, -2\sqrt{14}/14, 3\sqrt{14}/14)$
$= \sqrt{70}/70(3,6,5)$.

Vale la pena verificar que este vector en verdad tiene norma uno y que satisface la ecuación del plano.

Para terminar, una base ortonormal del plano es
$B_2 = \{(-2\sqrt{5}/5, \sqrt{5}/5, 0), \sqrt{70}/70(3,6,5)\}$.

Además, podemos construir una base ortonormal B_3 para el espacio $\mathbb{R}^3$:
$B_3 = \{\sqrt{70}/70(3,6,5), \sqrt{5}/5(-2,1,0), \sqrt{14}/14(-1,-2,3)\}$.

Nuestra forma de listar los elementos de la base define una numeración con la cual la base tiene determinante $+1$, de tal manera que podemos pensar como sigue: el primer vector de la base juega el papel del eje X, el vector normal juega el papel del eje Y y el tercero el papel del eje Z. Y además, el piso XY coincide con el plano.

472. **Ejercicio** *Construya una base ortonormal con determinante $+1$ para $\mathbb{R}^3$ tal que dos de sus vectores sean una base del plano siguiente:*

a) $z = -4x + 3xy$
b) $x - 7y + 2z = 0$
c) $2x - 3y - 9z = 0$.

473. **Ejemplo** *Sea B una base ortogonal de V. Si separamos a B en dos subconjuntos disjuntos, los subespacios generados por cada uno de ellos forman subespacios tales que cada uno es el complemento ortogonal del otro.*

474. ◇ ***Procedimiento para hallar las coordenadas de un vector en una base ortonormal.***

Supongamos que tenemos una base ortonormal $B = \{\vec{e}_1, ..., \vec{e}_n\}$ y un vector $\vec{v}$. Para hallar las coordenadas de $\vec{v}$ en B se procede como sigue. En primer término escribimos la ecuación que nos dice que $\vec{v}$ está en gen(B):
$\vec{v} = a_1\vec{e}_1 + a_2\vec{e}_2 + ... + a_j\vec{e}_j + ... + a_n\vec{v}_n$

Para hallar a_j multiplicamos la anterior ecuación en producto punto por $\vec{e}_j$:
$\vec{e}_j \cdot \vec{v} = \vec{e}_j \cdot (a_1\vec{e}_1 + a_2\vec{e}_2 + ... + a_j\vec{e}_j + ... + a_n\vec{v}_n)$
$= a_1(\vec{e}_j \cdot \vec{e}_1) + a_2(\vec{e}_j \cdot \vec{e}_2) + ... + a_j(\vec{e}_j \cdot \vec{e}_j) + ... + a_n(\vec{e}_j \cdot \vec{v}_n)$

Como se trata de una base ortonormal, todos los productos punto dan cero excepto aquel en $a_j(\vec{e}_j \cdot \vec{e}_j)$ lo cual da a_j pues el producto punto de $\vec{e}_j$ por el mismo da la norma cuadrado que es uno, pues la base está normalizada. Nos quedamos con

$\vec{e}_j \cdot \vec{v} = a_j$
o bien
$a_j = \vec{e}_j \cdot \vec{v}$
Este resultado tambien puede escribirse como sigue:

475. $\diamond$ ***Teorema.*** *Para una base ortonormal $B = \{\vec{e}_1, ..., \vec{e}_n\}$ y un vector $\vec{v}$ se tiene:*

$$\vec{v} = (\vec{e}_1 \cdot \vec{v})\vec{e}_1 + (\vec{e}_2 \cdot \vec{v})\vec{e}_2 + ... + (\vec{e}_j \cdot \vec{v})\vec{e}_j + ... + (\vec{e}_n \cdot \vec{v})\vec{v}_n$$

Ya tenemos alguna práctica sobre cómo construir bases ortonormales en dos y tres dimensiones. Pero en muchas aplicaciones se requiere trabajar con espacios de polinomios o de otras funciones. Por ejemplo, para estudiar fenómenos periódicos se usa la base formada por el conjunto de funciones

$$\mathbb{F} = \{1, \cos x, \cos 2x, ..., \cos nx, ..., \operatorname{sen} x, \operatorname{sen} 2x, ..., \operatorname{sen} nx, ...\}$$

Como vemos, este conjunto es infinito, lo cual exige generalizar nuestra teoría para EV de dimensión finita a dimensión infinita. En particular, es necesario entender qué quiere decir una suma infinita, una suma que nunca termina de hacerse. Porque la idea es expresar las funciones en la base $\mathbb{F}$, lo cual es un problema más sencillo una vez que uno la ortonormaliza. La expresión de una función en tal base ortonormalizada se llama la expansión en serie de Fourier de la función (Zill, 2007).

Es usual que uno tenga una base, pero que no sea ortonormal. Uno puede ortonormalizarla usando el siguiente

476. $\diamond$ ***Procedimiento para construir una base ortonormal a partir de una base cualquiera.***

Sea una base cualquiera que bien puede ser infinita:

$$B = \{\vec{b}_1, ..., \vec{b}_n, ...\}$$

Para construir una base ortonormal

$$B^{\Vdash} = \{\vec{e}_1, ..., \vec{e}_n\}$$

a partir de B se procede así:

1. El primer elemento de $B^{\Vdash}$ es simplemente el primer elemento de B normalizado:

$$\vec{e}_1 = \frac{\vec{b}_1}{||\vec{b}_1||}$$

2. Para hallar $\vec{e}_2$ consideramos el subespacio generado por $\{\vec{b}_1, \vec{b}_2\}$ y decidimos que $\vec{e}_2$ debe ser elemento de dicho subespacio, como lo es $\vec{e}_1$. Eso es equivalente a decir que
   $\vec{b}_2 = a_1\vec{e}_1 + a_2\vec{e}_2$

Conocemos el valor de a_1. Por el teorema anterior $a_1 = \vec{e}_1 \cdot \vec{b}_2$, pero $a_2\vec{e}_2$ no se conoce, lo reeplazamos por $\vec{h}_2$:

$\vec{b}_2 = a_1\vec{e}_1 + \vec{h}_2$

Despejemos $\vec{h}_2$:

$$\vec{h}_2 = \vec{b}_2 - a_1\vec{e}_1 = \vec{b}_2 - (\vec{e}_1 \cdot \vec{b}_2)\vec{e}_1$$

El vector $\vec{h}_2$ es perpendicular a $\vec{e}_1$, pues es un alargamiento de $\vec{e}_2$, por tanto el segundo elemento de $B^{\Vdash}$ es $\vec{h}_2$ normalizado:

$$\vec{e}_2 = \frac{\vec{h}_2}{||\vec{h}_2||}$$

3. Supongamos que hemos construido los primeros $j-1$ elementos de $B^{\Vdash}$.

   Para construir $\vec{e}_j$ consideramos el subespacio generado por $\{\vec{b}_1, \vec{b}_2, ..., \vec{b}_{j-1}, \vec{b}_j\}$ y queremos que coincida con el generado por $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_{j-1}, \vec{e}_j\}$. Eso es equivalente a decir que

   $\vec{b}_j = a_1\vec{e}_1 + a_2\vec{e}_2 + ... + a_{j-1}\vec{e}_{j-1} + a_j\vec{e}_j$

   pero $a_j\vec{e}_j$ no se conoce, lo reeplazamos por $\vec{h}_j$:

   $\vec{b}_j = a_1\vec{e}_1 + a_2\vec{e}_2 + ... + a_{j-1}\vec{e}_{j-1} + \vec{h}_j$

   Despejemos $\vec{h}_j$:

   $\vec{h}_j = \vec{b}_j - a_1\vec{e}_1 - a_2\vec{e}_2 - ... - a_{j-1}\vec{e}_{j-1}$

   que también se escribe como

$$\vec{h}_j = \vec{b}_j - (\vec{e}_1 \cdot \vec{b}_j)\vec{e}_1 - (\vec{e}_2 \cdot \vec{b}_j)\vec{e}_2 - ... - (\vec{e}_{j-1} \cdot \vec{b}_j)\vec{e}_{j-1}$$

   El vector $\vec{h}_j$ es perpendicular al subespacio generado por $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_{j-1}\}$, pues es un alargamiento de $\vec{e}_j$, por tanto el j-ésimo elemento de $B^{\Vdash}$ es $\vec{h}_j$ normalizado:

$$\vec{e}_j = \frac{\vec{h}_j}{||\vec{h}_j||}$$

4. Si se itera este procedimiento, se obtienen todos los elementos de la base ortonormalizada. El problema es que para una base infinita uno nunca termina. Por esto que en aplicaciones uno fija la exactitud deseada y determina en consecuencia qué tantos elementos de la base debe considerar.

477. *Ejemplo* *Construyamos una base ortonormal* $B^{\Vdash} = \{\vec{e}_1, \vec{e}_2\}$ *para el subespacio generado por* $B = \{(1,1,1), (-2,3,5)\}$

El primer vector de $B^{\Vdash}$ es el primer vector de B pero normalizado:

$$\vec{e}_1 = (1/\sqrt{3})(1,1,1)$$

El segundo vector $\vec{e}_2$ debe ser tal que
$(-2,3,5) = a_1\vec{e}_1 + a_2\vec{e}_2$
dando a $a_2\vec{e}_2$ el nombre de $\vec{h}_2$ obtenemos
$(-2,3,5) = a_1\vec{e}_1 + \vec{h}_2$
y como $a_1 = \vec{e}_1 \cdot (-2,3,5)$
$\vec{h}_2 = (-2,3,5) - (\vec{e}_1 \cdot (-2,3,5))\vec{e}_1$
$= (-2,3,5) - ((1/\sqrt{3})(1,1,1) \cdot (-2,3,5))(1/\sqrt{3})(1,1,1)$
$= (-2,3,5) - ((1/3)(1,1,1) \cdot (-2,3,5))(1,1,1)$
$= (-2,3,5) - (1/3)(-2+3+5)(1,1,1)$
$= (-2,3,5) - (1/3)(6)(1,1,1)$
$= (-2,3,5) - 2(1,1,1)$
$= (-4,1,3)$
El segundo vector de la base buscada es entonces $\vec{h}_2$ normalizado:

$$\vec{e}_2 = (1/\sqrt{26})(-4,1,3).$$

Podemos ver que $\vec{e}_2$ está en el generado por B pues h_2 lo está dado que
$(-4,1,3) = -2(1,1,1) + (-2,3,5)$.

478. $\boxed{\textbf{\textit{Ejercicio}}}$ *En $\mathbb{R}^4$ construya una base ortonormal para el subespacio generado por* $\{(1,1,1,1),(-2,3,5,7),(-1,4,2,1)\}$.

479. ♣ ***El principio de inducción matemática.*** *Nosotros utilizamos un esquema nuevo para demostrar que toda base es ortonormalizable, así tenga un número infinito de elementos siempre y cuando dichos elementos se puedan listar. Si hubiésemos procedido como antes, hubiésemos mostrado cómo se ortonormaliza una base con uno, dos y tres elementos y hubiésemos dicho que para casos más grandes, se sigue el mismo patrón. Pero aquí procedimos de otra forma, de la forma que se considera rigurosa y que se llama* ***demostración por inducción****. En efecto, lo que hicimos puede expresarse como sigue:*

para todo $n \in \mathbb{N} = \{0,1,2,3,...\}$ si B_n es una base, entonces puede ortonormalizarse

Este enunciado se generaliza en el siguiente:

para todo $n \in \mathbb{N} = \{0,1,2,3,...\}$ se tiene que la proposición $p(n)$ es cierta.

Para demostrar por inducción la validez de una proposición de esta clase, demuestre dos cosas. Primera, que $p(1)$ es cierta y, segunda, que siempre que se asuma que $p(j-1)$ es cierta entonces puede demostrarse que $p(j)$ también lo es.

Porque si las dos cosas se cumplen, entonces como $p(1)$ es cierta, entonces $p(2)$ también. Como $p(2)$ es cierta, entonces $p(3)$ también y así sucesivamente con la certeza de poder llegar hasta cualquier n no importa cuán grande sea.

8.5. La matriz transpuesta

En algunos casos, para encontrar la inversa de una matriz es suficiente transponerla, es decir, convertir sus columnas en filas.

480. ♣ ***Definición.*** *Al poner la columna i-ésima de la matriz M como la fila i-ésima, se forma una nueva matriz, M^T, llamada la matriz transpuesta de M. El número de filas de M es igual al número de columnas de M^T y el número de columnas de M es igual al número de filas de M^T. Si las entradas de M se denotan como M_{ij}, las entradas de M^T son M_{ji}.*

481. ***Ejemplo***

$$M = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}.$$

$$M^T = \begin{pmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{pmatrix}.$$

482. $\boxed{\textbf{\textit{Ejercicio}}}$ *Calcule MM^T y M^TM y encuentre la inversa de M si*

$$M = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

483. ***Ejemplo*** *Sea*

$$M = \begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 0 \end{pmatrix}$$

entonces

$$M^T = \begin{pmatrix} 2 & 5 & 8 \\ 3 & 6 & 9 \\ 4 & 7 & 0 \end{pmatrix}$$

484. ♣ ***Definiciones.*** *Una **matriz es simétrica** si es igual a su transpuesta. Es decir, M es simétrica si $M^T = M$. Una **matriz es antisimétrica** si es igual a menos su transpuesta. Es decir, M es antisimétrica si $M = -M^T$, o bien que $M^T = -M$. Ambos tipos de matrices deben ser matrices cuadradas.*

485. ***Ejemplos*** *La matriz*

$$M = \begin{pmatrix} 2 & 5 & 8 \\ 5 & 6 & 7 \\ 8 & 7 & 0 \end{pmatrix} = M^T$$

es simétrica. La matriz

$$M = \begin{pmatrix} 0 & 5 & 8 \\ -5 & 0 & 7 \\ -8 & -7 & 0 \end{pmatrix} = -M^T$$

es antisimétrica. Toda matriz antisimétrica tiene ceros sobre la diagonal.

486. ◇ ***Teorema.*** *Una matriz cuadrada y su transpuesta tienen igual determinante.*

Demostración. Consideremos sólo matrices dos por dos:

$$\text{Det}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - cb = \text{Det}\begin{pmatrix} a & c \\ b & d \end{pmatrix} \blacksquare$$

487. $\boxed{\textbf{\textit{Ejercicio}}}$ *Pruebe el teorema para el caso de una matriz* 3×3.

El siguiente resultado permite pruebas elegantes de diversos teoremas.

488. ◇ ***Teorema.*** *La matriz transpuesta cumple*

$$M(\vec{u}) \cdot \vec{v} = \vec{u} \cdot M^T(\vec{v})$$

donde la TL asociada a M es $M\colon \mathbb{R}^m \to \mathbb{R}^n$, la TL asociada a M^T va en sentido contrario, $\vec{u}$ está en el dominio de M y $\vec{v}$ en su codominio.

Por tanto, toda matriz simétrica S cumple con la identidad

$$S(\vec{u}) \cdot \vec{v} = \vec{u} \cdot S(\vec{v})$$

489. ***Ejemplo*** *Si $\vec{u} = (u_1, u_2, u_3), \vec{v} = (v_1, v_2, v_3)$ y*

$$M^T = \begin{pmatrix} 2 & 5 & 8 \\ 3 & 6 & 9 \\ 4 & 7 & 0 \end{pmatrix}$$

entonces

$$M(\vec{u}) \cdot \vec{v} = \begin{pmatrix} 2 & 5 & 8 \\ 3 & 6 & 9 \\ 4 & 7 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 2u_1 + 5u_2 + 8u_3 \\ 3u_1 + 6u_2 + 9u_3 \\ 4u_1 + 7u_2 + 0u_3 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

$$= (2u_1 + 5u_2 + 8u_3)v_1 + (3u_1 + 6u_2 + 9u_3)v_2 + (4u_1 + 7u_2 + 0u_3)v_3$$

$= u_1(2v_1 + 3v_2 + 4v_3) + u_2(5v_1 + 6v_2 + 7v_3) + u_3(8v_1 + 9v_2 + 0v_3)$

$$= \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \cdot \begin{pmatrix} 2v_1 + 3v_2 + 4v_3 \\ 5v_1 + 6v_2 + 7v_3 \\ 8v_1 + 9v_2 + 0v_3 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

$= \vec{u} \cdot M^T(\vec{v})$.

490. $\boxed{\textbf{\textit{Ejercicio}}}$ *Pruebe el teorema en su forma general.*

491. ⋄ ***Teorema y ejercicios.*** $(AB)^T = B^T A^T$. *Es decir, la transpuesta de un producto es el producto de las transpuestas, pero en orden inverso.*

Demostración. Las matrices dadas definen una composición de TL como sigue:

B: $\mathbb{R}^r \to \mathbb{R}^s$, A: $\mathbb{R}^s \to \mathbb{R}^t$ y como la TL asociada a una transpuesta va en sentido contrario, igual que la inversa, entoces la transpuesta de un producto debe ir en sentido contrario al del producto.

Ahora bien, sabemos que para una matriz M: $M\vec{x} \cdot \vec{y} = \vec{x} \cdot M^T\vec{y}$. Por tanto, $(AB)\vec{x} \cdot \vec{y} = \vec{x} \cdot (AB)^T\vec{y}$.

Por otro lado, $(AB)\vec{x} \cdot \vec{y} = A(B(\vec{x})) \cdot \vec{y} = (B\vec{x}) \cdot (A^T\vec{y}) = \vec{x} \cdot B^T(A^T\vec{y}) = \vec{x} \cdot (B^T A^T)\vec{y}$.

Por consiguiente, para todo par de vectores $\vec{x}$, $\vec{y}$ tenemos:

$$\vec{x} \cdot (AB)^T\vec{y} = \vec{x} \cdot (B^T A^T)\vec{y}$$

De aquí podemos concluir que $(AB)^T = B^T A^T$ pero debemos hacerlo correctamente, pues uno puede proceder erróneamente diciendo, por ejemplo, que todo lo que hay que hacer es simplificar por $\vec{x}$ y por $\vec{y}$. Eso no puede hacerse porque simplificar quiere decir multiplicar por el inverso y resulta que un vector no tiene inverso en el producto interior. Dicho de otra forma: de $\vec{x} \cdot \vec{y} = \vec{x} \cdot z$ no puede concluirse que $\vec{y} = \vec{z}$ (ejercicio).

La forma correcta de proceder es la siguiente: como $\vec{x} \cdot (AB)^T\vec{y} = \vec{x} \cdot (B^T A^T)\vec{y}$ es cierto para todo par de vectores $\vec{x}$ en el dominio de la TL asociada a B y $\vec{y}$ en el codomino de la TL asociada a A, entonces tomamos en el lugar de $\vec{x}$ a $\vec{e}_i$, el elemento i de la base natural del dominio de la TL asociada a B, y en vez de $\vec{y}$ a ϵ_j, el elemento j de la base natural del codominio de la TL asociada a A. Obtenemos:

$$\vec{e}_i \cdot (AB)^T\vec{\epsilon}_j = \vec{e}_i \cdot (B^T A^T)\vec{\epsilon}_j$$

Ahora bien, para cualquier matriz M, $\vec{e}_i \cdot M\vec{\epsilon}_j$ es M_{ij}, la entrada de la matriz M en la fila i y columna j (ejercicio). Eso significa que

$$((AB)^T)_{ij} = (B^T A^T)_{ij}$$

es decir, que las dos matrices $(AB)^T$ y $B^T A^T$ son iguales entrada por entrada, lo cual significa que ellas son iguales en todo sentido. ■

492. ⋄ ***Teorema.*** *Para una matriz simétrica M, Ker$(M) \perp$ Im(M).*

Demostración. Si M es simétrica, entonces es cuadrada y la TL asociada puede tomarse con dominio igual al codominio. Si $\vec{y}\in$Ker(M), entonces $M\vec{y}=\vec{0}$. Por tanto
$M\vec{x}\cdot\vec{y}=\vec{x}\cdot M^T\vec{y}=\vec{x}\cdot M\vec{y}=\vec{x}\cdot\vec{0}=0$, y $M\vec{x}$ pertenece a Im(M) mientras que $\vec{y}$ está en el Kernel, su producto punto es cero, y por lo tanto, la imagen es perpendicular al Kernel. ■

493. *Aplicación* *Este teorema nos ayuda a evitarnos trabajo. Si debemos calcular la imagen y el Kernel de una matriz simétrica, calculamos uno de ellos y el otro lo hallamos por ortogonalidad.*

494. *Ejercicio* *Resuelva la siguiente objeción: el teorema dice que el Kernel y la Imagen de una matriz simétrica son mutuamente perpendiculares, pero no que entre los dos generen a todo el espacio, es decir, que los dos se complementen ortogonalmente. Por lo tanto, la aplicación es abusiva y no sirve de nada.*

495. *Ejemplo* *Calculemos el Kernel y la imagen de la TL cuya matriz es*

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$$

El Kernel es el subespacio vectorial que cae sobre cero. Puesto que el determinante de la matriz no es cero, la transformación lineal que ésta representa es 1-1 y sobre, y por tanto, el único vector que cae sobre cero es cero que es el único elemento del Kernel. Como la Imagen y el Kernel son mutuamente perpendiculares, la Imagen es todo $\mathbb{R}^2$. Observemos que esto es cierto para cualquier matriz invertible.

496. *Ejercicio* *Encuentre el Kernel y la imagen de la TL cuya matriz es*

a) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$

b) $\begin{pmatrix} 1 & 2 \\ 2 & -3 \end{pmatrix}$

c) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$

d) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 5 \end{pmatrix}$

8.6. Aplicaciones

497. $\diamond$ ***Teorema y definición.*** *En $\mathbb{R}^n$, toda base ortonormal escrita en notación vertical forma una matriz Q tal que $Q^{-1} = Q^T$, donde Q^T representa la matriz transpuesta de Q. Decimos que una matriz O es* ***ortogonal*** *cuando sus columnas representan una base ortonormal. En ese caso: $OO^T = O^TO = \mathbb{I}$ puesto que $O^T = O^{-1}$.*

Demostración. Sea $Q = \{\vec{b}_1, \vec{b}_2, ..., \vec{b}_n\}$ una base ortonormal de $\mathbb{R}^n$. Por lo tanto, cada vector tiene norma uno, es decir $\vec{b}_i \cdot \vec{b}_i = 1$, mientras que vectores diferentes son perpendiculares: $\vec{b}_i \cdot \vec{b}_j = 0$, cuando $i \neq j$. Esto se acostumbra resumir diciendo: $\vec{b}_i \cdot \vec{b}_j = \delta^i_j$, donde δ^i_j es uno si $i = j$ y es cero si $i \neq j$. El símbolo δ^i_j se lee *delta i j* y se le conoce como el **Delta de Kronecker**.

La i-ésima columna de la matriz Q es el i-ésimo vector de la base Q. La j-ésima fila de Q^T es la j-ésima columna de Q. Por tanto Q^TQ tiene en su entrada (i, j) el valor $\vec{b}_i \cdot \vec{b}_j$ que es uno sobre la diagonal y cero en cualquier otro lado. Por tanto, $Q^TQ = I$ y la inversa de Q es Q^T. ■

498. ***Ejemplo y ejercicio*** *En $\mathbb{R}^2$ tomamos la base $\{(1, 2), (-2, 1)\}$. Sus vectores son perpendiculares, pero no son unitarios. Los normalizamos para obtener una base ortonormal:*
$\{(1/\sqrt{5}, 2/\sqrt{5}), (-2/\sqrt{5}, 1/\sqrt{5})\}$.

La base puede ser representada como una matriz:

$$Q = \begin{pmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$$

cuya transpuesta es

$$Q^T = \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$$

El producto de esas dos matrices es, en cualquier orden, la identidad (ejercicio). Por tanto, Q y Q^T son la inversa la una de la otra.

499. $\diamond$ ***Teorema y ejercicio.*** *En el plano tenemos:*

a) La matriz $R(\theta)$ de una rotación es ortogonal. Es decir, la matriz de $R^{-1}(\theta)$es simplemente la transpuesta de $R(\theta)$.

b) La inversa $R^{-1}(\theta)$ de una rotación $R(\theta)$ con ángulo θ es la rotación al revés $R(-\theta)$ con el mismo ángulo.

c) Sea B el resultado de rotar por $R(\theta)$ a la base natural N. La matriz de $R(\theta)$ también representa la matriz de cambio de coordenadas del marco rotado, B, a la base natural N: $I^B_N = R(\theta)$.

Demostración: ejercicio.

500. **Ejercicio** *¿Conserva una rotación el producto punto? Si una TL conserva el producto punto, ¿es una rotación? Es decir, si A es una matriz de una rotación, ¿$A\vec{x} \cdot A\vec{y} = \vec{x} \cdot \vec{y}$? Y si $T(x) \cdot T(y) = \vec{x} \cdot \vec{y}$, ¿es T una rotación?*

501. ***Más entendimiento.*** *Hay una diferencia notable entre saber y entender. Como es difícil captar qué quiere decir, resolveremos el siguiente ejercicio de 4 maneras diferentes. Con eso queremos decir que entender implica manejar una perspectiva global que le permita a uno resolver un problema de muy diversas maneras, unas buenas para un propósito y otras para otro. Entender es necesario: los humanos no estamos diseñados para producir resultados sin equivocaciones. Pero hay que producirlos: ¿cómo? haciendo el mismo problema por métodos muy diferentes y viendo que los resultados concuerdan.*

502. ***Ejemplo*** *Encontremos la ecuación de una elipse rotada E. La elipse original O obedece a la ecuación $x^2/9 + y^2/4 = 1$ y es rotada $\pi/4$ en sentido contrario a las manecillas del reloj.*

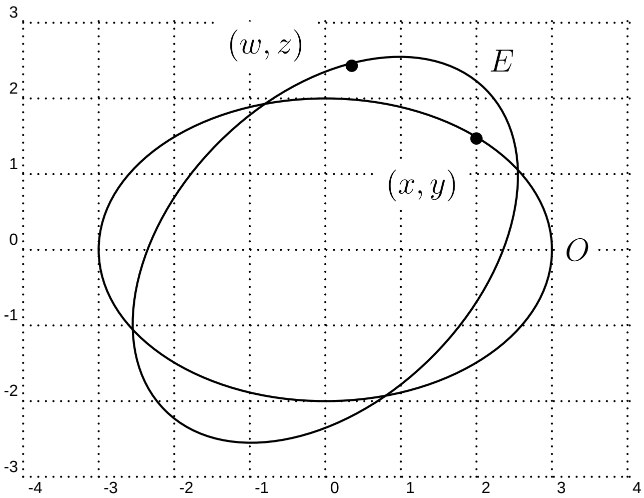


Figura 8.2. La rotación de una elipse.

Primer método. La matriz R de una rotación al contrario de las manecillas del reloj y con ángulo $\pi/4$ es

$$M = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

Puesto que las columnas de M son vectores que forman una base ortonormal, su inversa es la transpuesta:

$$M^{-1} = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

Por lo tanto, el punto $(x,y) \in E$ ssi $(w,z) = M^{-1}(x,y) \in O$.

$$M^{-1}(x,y) = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sqrt{2}x/2 + \sqrt{2}y/2 \\ -\sqrt{2}x/2 + \sqrt{2}y/2 \end{pmatrix}$$

Eso quiere decir que el punto $(x, y) \in E$ ssi la dupleta
$(\sqrt{2}x/2 + \sqrt{2}y/2, -\sqrt{2}x/2 + \sqrt{2}y/2)$
satisface la ecuación O:
$(\sqrt{2}x/2 + \sqrt{2}y/2)^2/9 + (-\sqrt{2}x/2 + \sqrt{2}y/2)^2/4 = 1$
Equivalentemente
$13x^2 + 13y^2 - 10xy = 72$
Notemos: un término de la forma xy es un distintivo de las rotaciones.

Es conveniente hacer un test: por inspección vemos que el punto con coordenadas $(\sqrt{3}/2, \sqrt{3}/2)$ está en la elipse E, por lo que debe satisfacer su ecuación. Veamos:

$$\begin{aligned} 13(\sqrt{3}/2)^2 + 13(\sqrt{3}/2)^2 - 10(\sqrt{3}/2)(\sqrt{3}/2) &= (13)(9)/2 + (13)(9)/2 - 10(9/2) \\ &= (13)(9) - (5)(9) \\ & 72 \end{aligned}$$

Pasamos el test.

Segundo método. La matriz M es también la notación vertical de la base $B = (\sqrt{2}/2)\{(1,1), (-1,1)\}$. Sea $P_B = (w, z)$ las coordenadas del punto P en la base B, mientras que $P_N = (x, y)$ representa las coordenadas del mismo punto P en la base natural. Ese par de coordenadas está relacionado por
$(w, z) = P_B = I_B^N P_N = B^{-1}(x, y)$

$$B^{-1}(x, y) = \begin{pmatrix} \sqrt{2}x/2 + \sqrt{2}y/2 \\ -\sqrt{2}x/2 + \sqrt{2}y/2 \end{pmatrix}$$

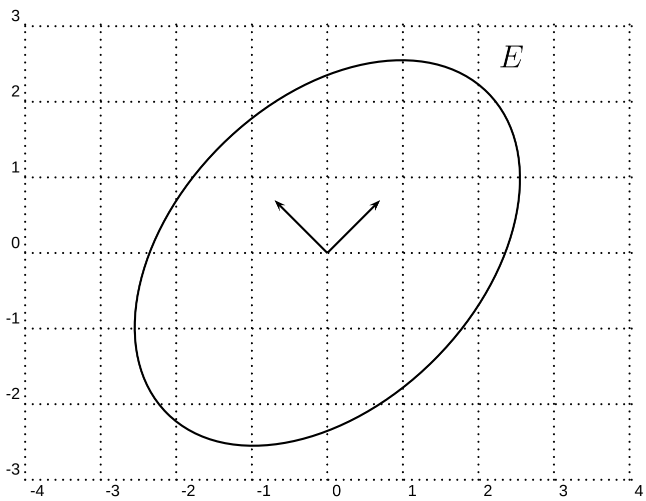


Figura 8.3. Desde la base $B = (\sqrt{2}/2)\{(1,1), (-1,1)\}$, la elipse E se ve derechita.

Por tanto
$w = \sqrt{2}x/2 + \sqrt{2}y/2$
$z = -\sqrt{2}x/2 + \sqrt{2}y/2$

La ecuación de E en la base B es simplemente $w^2/9 + z^2/4 = 1$ porque en la base B uno ve una elipse normal con eje mayor 3 y eje menor 2, por lo tanto, para encontrar la ecuación de E en la base natural, es suficiente sustituir w y z en esta ecuación por

su equivalente en la base natural. La ecuación resultante es exactamente la que ya habíamos encontrado:
$(\sqrt{2}x/2+\sqrt{2}y/2)^2/9+(-\sqrt{2}x/2+\sqrt{2}y/2)^2/4=1$

Tercer método. Comenzamos con una base un tanto diferente de la anterior: nuestra nueva base W es

$$W=\begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

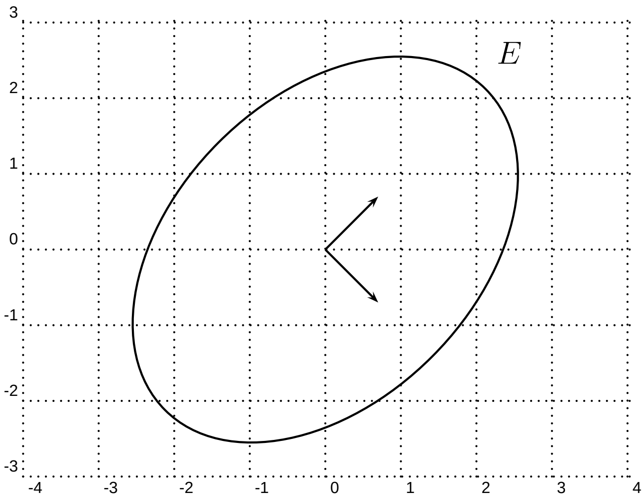


Figura 8.4. Desde la base $W=(\sqrt{2}/2)\{(1,-1),(1,1)\}$, la elipse E se ve derechita.

Observemos que la base, al igual que la anterior, también tiene $Det(W)=1$ y además es ortonormal.

Sea $P_W=(u,v)$ las coordenadas de un punto P en la base W, en tanto que $P_N=(x,y)$ representa las coordenadas del mismo punto P en la base natural. Ese par de coordenadas se relaciona por
$(u,v)=P_W=I_W^N P_N=W^{-1}(x,y)$
donde

$$W^{-1}=\begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

por lo que

$$W^{-1}(x,y)=\begin{pmatrix} \sqrt{2}x/2-\sqrt{2}y/2 \\ \sqrt{2}x/2+\sqrt{2}y/2 \end{pmatrix}$$

Por tanto,
$u=\sqrt{2}x/2-\sqrt{2}y/2$
$v=\sqrt{2}x/2+\sqrt{2}y/2$

La ecuación de E en la base W es $u^2/4+v^2/9=1$. Por lo que la ecuación de E en la base natural es:
$(\sqrt{2}x/2-\sqrt{2}y/2)^2/4+(\sqrt{2}x/2+\sqrt{2}y/2)^2/9=1$

lo cual coincide con la ecuación hallada anteriormente:
$(\sqrt{2}x/2+\sqrt{2}y/2)^2/9+(-\sqrt{2}x/2+\sqrt{2}y/2)^2/4=1$

503. ***El cuarto método.*** *Puesto que hemos venido guiándonos por el contexto, hemos sido muy perezosos con la notación. Pero ahora veamos de qué manera una mejor notación es necesaria. Enfaticemos que cuando ponemos una columna en frente de una matriz, esa columna realmente representa las coordenadas de un vector en una base dada, que por defecto es la natural. Y si escribimos una fila al lado izquierdo de la matriz, podemos decir que pusimos las coordenadas de un vector, pero en posición horizontal y debemos referirlo como transpuesto* $(\vec{x})^T$. *Ejemplo:*

$$\vec{x}=(\vec{x})_N=\begin{pmatrix}1\\2\\3\end{pmatrix}$$

$$(\vec{x})^T=\begin{pmatrix}1&2&3\end{pmatrix}$$

Notemos ahora que si

$$M=\begin{pmatrix}a&b\\c&d\end{pmatrix}$$

entonces

$$\begin{pmatrix}x&y\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

$$=\begin{pmatrix}x&y\end{pmatrix}\begin{pmatrix}ax+by\\cx+dy\end{pmatrix}=x(ax+by)+y(cx+dy)=ax^2+(b+c)xy+dy^2$$

En una ecuación: $(\vec{x})^T M\vec{x}=ax^2+(b+c)xy+dy^2$
Podemos concluir que la información contenida en la ecuación de una elipse pudo codificarse en la matriz M. Hemos convertido la geometría en álgebra.

Ahora, recordemos que, como se vio en el segundo método, la matriz

$$B=\begin{pmatrix}\sqrt{2}/2&-\sqrt{2}/2\\\sqrt{2}/2&\sqrt{2}/2\end{pmatrix}$$

codifica la información para una base ortonormal B, y por tanto su inversa es igual a su transpuesta. Las coordenadas en la base B se denotan como $(P)_B=(w,z)$ y la elipse luce como $w^2/9+z^2/4=1$. En forma matricial, eso es equivalente a

$$\begin{pmatrix}w&z\end{pmatrix}\begin{pmatrix}1/9&0\\0&1/4\end{pmatrix}\begin{pmatrix}w\\z\end{pmatrix}=1$$

o
$((\vec{x})_B)^T M(\vec{x})_B=1$

donde

$$M = \begin{pmatrix} 1/9 & 0 \\ 0 & 1/4 \end{pmatrix}$$

Podemos asociar biunívocamente la elipse a la matriz M si hacemos la aclaración de que M se refiere a la base B. Por tanto, tenemos que la matriz asociada a la elipse es

$$M_B^B = \begin{pmatrix} 1/9 & 0 \\ 0 & 1/4 \end{pmatrix}$$

y por tanto

$((\vec{x})_B)^T M_B^B (\vec{x})_B = 1.$

Para encontrar la ecuación de la elipse en coordenadas cartesianas, necesitamos solamente cambiar de coordenadas en la base B a la base natural, una tarea que es resuelta por la identidad $(\vec{x})_B = I_B^N (\vec{x})_N$. Además, B es ortonormal y por tanto $(I_B^N)^T = (I_B^N)^{-1} = I_N^B = B$. Adicionalmente, recordemos que la transposición invierte el orden de un producto. Tenemos:

$$\begin{aligned}
((\vec{x})_B)^T M_B^B (\vec{x})_B &= (I_B^N (\vec{x})_N)^T \; M_B^B \; (I_B^N (\vec{x})_N) \\
&= ((\vec{x})_N)^T \; (I_B^N)^T \; M_B^B \; (I_B^N (\vec{x})_N) \\
&= ((\vec{x})_N)^T \; (I_B^N)^{-1} M_B^B \; (I_B^N (\vec{x})_N) \\
&= ((\vec{x})_N)^T \; (I_N^B) \; M_B^B \; (I_B^N (\vec{x})_N) \\
&= ((\vec{x})_N)^T \; (I_N^B \; M_B^B \; I_B^N)(\vec{x})_N \\
&= ((\vec{x})_N)^T \; M_N^N \; (\vec{x})_N
\end{aligned}$$

en donde hemos aplicado que la transpuesta de un producto es el producto de las transpuestas pero en orden inverso, que la transpuesta de una matriz ortogonal coincide con su inversa, que la inversa de un cambio de base es el cambio de base al revés, que el producto de matrices es asociativo y que $M_N^N = I_N^B M_B^B I_B^N = BMB^T$.

En conclusión, la elipse en la base B se lee $((\vec{x})_B)^T M_B^B (\vec{x})_B = 1$, o

$$\begin{pmatrix} w & z \end{pmatrix} \begin{pmatrix} 1/9 & 0 \\ 0 & 1/4 \end{pmatrix} \begin{pmatrix} w \\ z \end{pmatrix} = 1$$

Pero por otro lado, la elipse en la base natural se expresa como $((\vec{x})_N)^T M_N^N (\vec{x})_N = 1$ o equivalentemente

$$1 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} 1/9 & 0 \\ 0 & 1/4 \end{pmatrix} \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Test: todo eso debe ser igual a $13x^2+13y^2-10xy=72$. Veamos:

$$1=(\sqrt{2}/2)^2\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{9} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$=(1/2)\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{9} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}\begin{pmatrix} x+y \\ -x+y \end{pmatrix}$$

$$=(1/2)\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x/9+y/9 \\ -x/4+y/4 \end{pmatrix}$$

$$=(1/2)\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} x/9+y/9+x/4-y/4 \\ x/9+y/9-x/4+y/4 \end{pmatrix}$$

$$=(1/2)\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} (4x+4y+9x-9y)/36 \\ (4x+4y-9x+9y)/36 \end{pmatrix}$$

$$=(1/72)\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 4x+4y+9x-9y \\ 4x+4y-9x+9y \end{pmatrix}$$

$$=(1/72)\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 13x-5y \\ -5x+13y \end{pmatrix}$$

Por tanto:

$1=(1/72)(13x^2-5xy-5xy+13y^2)$

Y al final obtenemos: $72=13x^2-10xy+13y^2$.

504. $\boxed{\textbf{\textit{Ejercicio}}}$ *Siempre hemos escuchado que $y=1/x$ representa una hipérbola. Demuéstrelo. Para ello, tome la hipérbola $x^2-y^2=1$, rótela en sentido antihorario un ángulo $\pi/4$ y encuentre la ecuación de la figura rotada. Use cuatro métodos distintos para resolver el problema.*

8.7. Ejercicios de repaso

1. Sean $B=\{(2,2),(4,-1)\}$ y $B^*=\{(1,3),(-1,1)\}$. Muestre que B y B^* son bases de $\mathbb{R}^2$ y encuentre la matriz de transición de B a B^* y la de B^* a B.

2. Sean $B=\{6+3x,10+2x\}$ y $B^*=\{2,3+2x\}$ bases de $P_{\leq 2}$.

   a) Halle la matriz de transición de B a B^* y la de B^* a B.

   b) Encuentre $[-4+x]_B$ y use una matriz encontrada anteriormente para hallar $[-4+x]_{B^*}$.

3. Sea $V=$gen$(\{\text{sen}(x),\cos(x)\})$. Muestre que $B^{'}=\{2\,\text{sen}(x)+\cos(x),3\cos(x)\}$ es base de V y encuentre la matriz de transición de la base $B=\{\text{sen}(x),\cos(x)\}$ a la base B'.

4. Sea T una transformación lineal de V en U, espacios vectoriales. Demuestre que si $\{v_1,v_2,...,v_m\}$ es una base de V y $T(v_1)=T(v_2)=...=T(v_n)=\vec{0}$ entonces $T(w)=\vec{0}$ para todo $w\in V$. Encuentre la matriz de la transformación.

5. Sea T una transformación lineal de V en V, espacio vectorial. Demuestre que si $\{v_1, v_2, ..., v_m\}$ es una base de V y $T(v_1) = v_1, T(v_2) = v_2, ..., T(v_n) = v_n$ entonces $T(w) = w$, para todo $w \in V$. Encuentre la matriz de la transformación.

6. Obtenga una base ortonormal de $\mathbb{R}^4$ que contenga los vectores $\left\{\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}, 0\right\}$, $\left\{0, \frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right\}$.

7. Sea A una matriz 2×2 tal que $A^T A = \mathbb{I}$, muestre que la transformación lineal asociada preserva producto punto, longitudes y ángulos.

8. Sea $H = \{(x, y, z, w) \in \mathbb{R}^4\colon x = z,\ y = w\}$ un subespacio vectorial de $\mathbb{R}^4$.

   a) Halle una base ortonormal B_1 para H.

   b) Halle una base ortonormal B_2 para $H^\perp$.

   c) Exprese el vector $\vec{v} = (1, 0, 0, 1)$ como $\vec{v} = \vec{h} + \vec{p}$, donde $\vec{h} \in H$ y $\vec{p} \in H^\perp$.

   d) Escriba las coordenadas del vector $\vec{a} = (3, -2, 3, -2)$ en la base B_1.

9. Sea $H = \{(x, y, z, w) : x = y, w = 3y\}$ y sea $\vec{v} = (-1, 2, 3, 1)$. Encuentre una base ortonormal para H y para su complemento ortogonal $H^\perp$. Escriba $\vec{v} = \vec{h} + \vec{p}$, con $\vec{h}$ y $\vec{p}$ elementos de H y de $H^\perp$, respectivamente.

10. Sean A y B matrices de tamaño $n \times n$, con $A \neq O$ y $B \neq O$, donde O es la matriz de ceros. Mostrar que si A es *simétrica* y B *antisimétrica*, entonces el conjunto $\{A, B\}$, es *linealmente independiente*.

11. Considere la matriz $R = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$.

    a) Calcule R^n, n un entero positivo.

    b) Para qué valores de θ la matriz R es invertible?

    c) Halle R^{-1}.

Rta. a) $\begin{pmatrix} \cos(n\theta) & -\operatorname{sen}(n\theta) \\ \operatorname{sen}(n\theta) & \cos(n\theta) \end{pmatrix}$, b) para todo θ, c) $R^{-1} = R^T$.

12. Pruebe que si A es simétrica e invertible, A^{-1} es simétrica.

Sugerencia: Haga ver que $(A^{-1})^T = A^{-1}$.

13. Sea S_n el espacio vectorial de matrices simétricas $n \times n$, halle la dimensión de este espacio.

8.8. Resumen

Hemos aprendido cómo cambiar de coordenadas de una base a otra con la ayuda de matrices apropiadas. Para cambiar de coordenadas de una base B a la base natural, usamos una matriz que es la misma base B escrita en forma vertical, un vector al lado del otro. Para cambiar de la base natural a la base B, usamos como matriz la inversa de B. Para cambiar de B_1 a base B_2, usamos la matriz $B_2^{-1}B_1$. Si B es ortonormal, con elementos mutuamente perpendiculares y de norma uno, $B^{-1} = B^T$. Similarmente, si R es una rotación, $R^{-1} = R^T$. Nuestras bases siempre se numeran de tal forma que las matrices asociadas tengan determinante positivo.

CAPÍTULO 9

TL EN CUALQUIER PAR DE BASES

Hemos adquirido la libertad de cambiar de coordenadas de una base a otra base cualquiera. Podemos usar esa libertad para expresar TL en forma matricial, pero con respecto a un par de bases bien escogidas. De esa forma, la información que conlleva la matriz puede ser más transparente, más útil y más fácil de entender y manejar. Comencemos con un ejemplo para mostrar el significado de la transparencia. La utilidad será demostrada ampliamente en todo lo que sigue, especialmente en diagonalización.

9.1. Fundamento

Sabemos cómo asociar una matriz T a una transformación lineal T. La asociación es tan clara y natural que nosotros usamos la misma letra tanto para la matriz como para la TL. Eso ha sido posible porque hemos usado la base natural en todo lado. Ahora, nos liberamos de esa restricción y consideraremos el problema de asociar una matriz con referencia a cualquier par de bases, una en el dominio y otra en el codominio.

Con el símbolo

$$T_{B_2}^{B_1}$$

representaremos la matriz asociada a una TL, T: $V \to W$ tal que tenemos en el dominio V a la base B_1 y en el codominio W a la base B_2. Observemos que las bases se ponen indicando una dirección vertical. Cuando T: $V \to W$ y $B_1 = B_2 = B$, T_B^B se refiere a la matriz de T con respecto a B.

La matriz $T_{B_2}^{B_1}$ se define para que llene el requisito siguiente:

$$(T(X))_{B_2} = T_{B_2}^{B_1}(X_{B_1})$$

lo cual indica que el trabajo de $T_{B_2}^{B_1}$ es transformar las coordenadas $(X)_{B_1}$ de cualquier vector X en la base B_1 en coordenadas $(T(X))_{B_2}$ de $T(X)$ en la base B_2. Los siguientes ejemplos muestran que conceptualmente no hay nada qué agregar a lo que ya hemos visto para poder obtener $T_{B_2}^{B_1}$. Sin embargo, elaboraremos algunos teoremas que nos ayudarán a ganar comprensión y poder mecánico.

505. Ejemplo *Calculemos T_N^N, donde N es la base natural o canónica de $\mathbb{R}^2$, dado que $B_1 = \{(1,1),(-2,2)\}$, $B_2 = \{(-1,-1),(-2,2)\}$, y que*

$$T_{B_2}^{B_1} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Solución: Para encontrar T_N^N nosotros necesitamos encontrar $T(\vec{i})$ y $T(\vec{j})$ y poner los vectores resultantes en notación vertical en forma de matriz. Necesitamos además expresar cada elemento de la base natural como una combinación lineal de la base B_1. Por eso que necesitamos calcular T de cada uno de los elementos de dicha base. Nuestro punto de partida es:

$(T(X))_{B_2} = T_{B_2}^{B_1}(X)_{B_1}$

Si tomamos $X = (1,1)$, entonces $(X)_{B_1} = (1,0)$ porque $(1,1) = 1(1,1) + 0(-2,2)$, y

$$(T(X))_{B_2} = T_{B_2}^{B_1}(X)_{B_1} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Por lo que,

$(T(X))_{B_2} = (T(1,1))_{B_2} = 1(-1,-1) + 3(-2,2) = (-7,5)$.

Similarmente,

$(T(-2,2))_{B_2} = 2(-1,-1) + 4(-2,2) = (-10,6)$.

pero por otro lado, descompongamos los elementos de la base natural en la base B_1

$\vec{i} = (1,0) = \alpha(1,1) + \beta(-2,2)$

Encontramos, $\alpha = 1/2$ y $\beta = -1/4$. Por tanto,

$\vec{i} = (1,0) = 1/2(1,1) - 1/4(-2,2)$

y

$T(\vec{i}) = T(1,0) = 1/2T(1,1) - 1/4T(-2,2) = 1/2(-7,5) - 1/4(-10,6)$

$= (-7/2 + 5/2, 5/2 - 3/2) = (-1,1)$.

Ahora $\vec{j}$:

$\vec{j} = (0,1) = \alpha(1,1) + \beta(-2,2)$

Uno encuentra que, $\alpha = 1/2$ y $\beta = 1/4$. Por tanto,

$\vec{j} = (0,1) = 1/2(1,1) + 1/4(-2,2)$

y

$T(\vec{j}) = T(0,1) = 1/2T(1,1) + 1/4T(-2,2) = 1/2(-7,5) + 1/4(-10,6)$

$= (-7/2 - 5/2, 5/2 + 3/2) = (-6,4)$.

Resumiendo: $T(\vec{i}) = (-1,1)$ y $T(\vec{j}) = (-6,4)$. Entonces:

$$T = \begin{pmatrix} -1 & -6 \\ 1 & 4 \end{pmatrix}$$

Es hora de hacer un test: hagamos el problema al revés.

506. Ejemplo *Calculemos $T_{B_2}^{B_1}$ dado que $B_1 = \{(1,1),(-2,2)\}$, $B_2 = \{(-1,-1),(-2,2)\}$ y que*

$$T = \begin{pmatrix} -1 & -6 \\ 1 & 4 \end{pmatrix}$$

Solución: debemos encontrar

$$T_{B_2}^{B_1} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}$$

dado que
$T_{B_2}^{B_1} X_{B_1} = (T(X))_{B_2}$.
Si $X = (1, 1)$, entonces $X_{B_1} = (1, 0)$
Por tanto

$$T_{B_2}^{B_1}(1, 0) = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = (T(1, 1))_{B_2}$$

Encontremos $(T(1, 1))_{B_2}$. Comencemos con $T(1, 1)$:

$$T(1, 1) = \begin{pmatrix} -1 & -6 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -7, 5 \end{pmatrix}$$

Así, $(T(1, 1))_{B_2} = (\alpha, \beta) = (-7, 5)_{B_2}$, esto es

$(-7, 5) = \alpha(-1, -1) + \beta(-2, 2)$

o lo que es lo mismo

$$\begin{aligned} -7 &= -\alpha - 2\beta \\ 5 &= -\alpha + 2\beta \end{aligned}$$

por tanto, $\alpha = 1$ y $\beta = 3$.

Debemos tomar ahora el segundo vector de la base B_1: Si $X = (-2, 2)$, entonces $X_{B_1} = (0, 1)$.
Por lo que

$$T_{B_2}^{B_1}(0, 1) = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \gamma \\ \delta \end{pmatrix} = (T(-2, 2))_{B_2}$$

Encontremos $(T(-2, 2))_{B_2}$:

$$T(-2, 2) = \begin{pmatrix} -1 & -6 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} -2 \\ 2 \end{pmatrix} = \begin{pmatrix} -10, 6 \end{pmatrix}$$

De tal forma que, $(T(-2, 2))_{B_2} = (\gamma, \delta) = (-10, 6)_{B_2}$, lo cual quiere decir que

$(-10, 6) = \gamma(-1, -1) + \delta(-2, 2)$

o equivalentemente que

$$\begin{aligned} -10 &= -\gamma - 2\delta \\ 6 &= -\gamma + 2\delta \end{aligned}$$

Por tanto, $\gamma = 2$ y $\delta = 4$.
En conclusión:

$$T_{B_2}^{B_1} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Lo cual demuestra que hicimos bien las cosas y que hemos entendido.

507. $\boxed{\textbf{\textit{Ejercicio}}}$ *Hacer al derecho y al revés, como en los dos pasados ejemplos, las siguientes tareas:*

a) Calcular T_N^N dado que $B_1 = \{(1,2),(-2,3)\}$, $B_2 = \{(-2,-1),(-1,2)\}$, y que

$$T_{B_2}^{B_1} = \begin{pmatrix} -1 & -1 \\ 2 & -1 \end{pmatrix}$$

b) Calcular T_N^N dado que $B_1 = \{(-2,5),(2,0)\}$, $B_2 = \{(-1,-1),(-1,1)\}$, y que $T_{B_2}^{B_1}$ es la misma del inciso a).

c) Calcular T_N^N dado que $B_1 = \{(5,-1),(3,7)\}$, $B_2 = \{(0,-1),(1,-1)\}$, y que

$$T_{B_2}^{B_1} = \begin{pmatrix} -7 & -2 \\ -2 & 7 \end{pmatrix}.$$

9.2. Reflexiones en el plano

La matriz

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

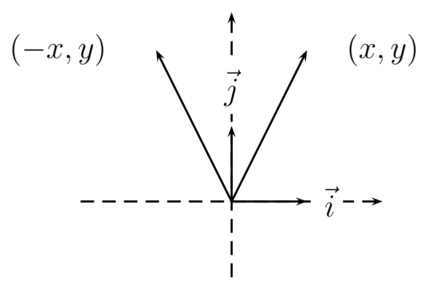


Figura 9.1. Una reflexión con respecto al eje Y.

transforma al vector (x, y) de $\mathbb{R}^2$ en $(-x, y)$, también de $\mathbb{R}^2$. En palabras, esta matriz es una TL que refleja cada vector de $\mathbb{R}^2$ con respecto al eje $\vec{Y}$. Por ejemplo, la imagen de $\vec{i} = (1, 0)$ es $-\vec{i} = (-1, 0)$, pero la imagen de $\vec{j} = (0, 1)$ es $\vec{j}$, mientras que la de (1,1) es (-1,1) y que la de (-2,3) es (2,3). Por lo que podemos mirar la matriz y exclamar: ¡es una reflexión! Lo mismo debemos poder decir de cualquier reflexión con respecto a cualquier eje. Por ejemplo, probaremos eso para la matriz

$$\begin{pmatrix} -4/5 & -3/5 \\ -3/5 & 4/5 \end{pmatrix}$$

la cual es una reflexión con respecto a la línea $y = -3x$. Sin embargo, uno no ve dicha información en esa matriz. Pero si escribimos la misma TL con respecto a una base apropiada, $B = \{(1, 1/3), (-1, 3)\}$ la matriz correspondiente será (ejercicio)

$$R_B^B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

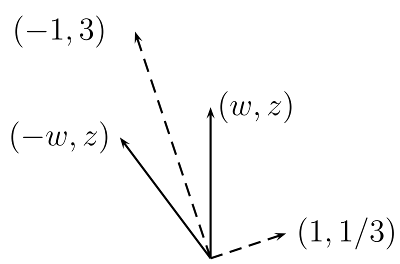


Figura 9.2. Una reflexión con respecto al eje $y = -3x$ vista desde la base $B = \{(1, 1/3), (-1, 3)\}$.

Esta expresión dice que la matriz reescrita en la base B, tanto en el dominio como en el codominio, es claramente una reflexión. La única corrección pertinente es que la reflexión es con respecto a la línea, una información que está contenida en la base B: el segundo vector de la base B, $(-1, 3)$, genera el eje de reflexión, mientras que el primero, $(1, 1/3)$, genera la línea perpendicular. El orden que hemos dado a la numeración de los elementos de la base se ajusta a nuestro estilo de que las bases tengan determinante positivo.

9.3. El papel de las bases

Ya sabemos de sobra que una TL está completamente determinada por su efecto sobre los elementos de la base natural. Pero ya nos dimos cuenta de que la base natural no solamente no tiene nada de especial sino que a veces no es la mejor para expresar una información dada. Es por tanto la hora de resaltar el papel de las otras bases.

508. ⋄ ***Teorema.*** *Para determinar completamente una TL, $T\colon V \to W$ es suficiente saber qué hace ella sobre una base cualquiera del dominio V.*

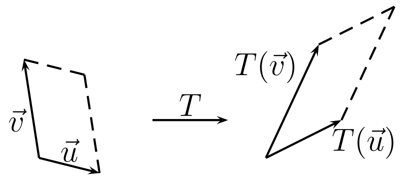


Figura 9.3. Las TL se conocen por sus efectos sobre una base cualquiera.

Demostración. Sea $B = \{\vec{e}_1, ..., \vec{e}_n\}$ una base del EV V. La base B expande todo el espacio. Por tanto, si $\vec{X} \in V$ entonces $\vec{X} = \alpha_1\vec{e}_1 + ... + \alpha_n\vec{e}_n$ y por tanto

$T(\vec{X}) = T(\alpha_1\vec{e}_1 + ... + \alpha_n\vec{e}_n)$ pero como T es una TL, entonces:

$T(\vec{X}) = \alpha_1 T(\vec{e}_1) + ... + \alpha_n T(\vec{e}_n)$

Esto significa: si conocemos el efecto de una TL T sobre una base cualquiera B del dominio, sabremos qué hace T sobre cualquier elemento X del espacio V, para lo cual es suficiente combinar las imágenes de los elementos de la base poniendo como coeficientes las coordenadas de X en la misma base B. ■

En términos geométricos, el último teorema dice que para saber qué hace una TL es suficiente saber la imagen de un n-plp de volumen no nulo.

Ahora, aplicaremos este teorema para encontrar la matriz de una TL con respecto a una base arbitraria. Denotamos cualquier matriz N como $N = (n_{ij})$, lo cual significa que la entrada de la matriz n_{ij} está en la intersección de la fila i con la columna j.

Nuestro trabajo es generalizar lo que hemos hecho para hallar la matriz de una TL T con respecto a la base natural. Todo vector de la base natural del dominio debe ser transformado por T y el resultado se escribe verticalmente. Todo fue muy fácil porque las coordenadas de un vector en la base natural eran a la larga el mismo vector. Hagamos una pequeña generalización que podrá ser entendida en el siguiente ejemplo.

509. ***Ejemplo*** *Sea $T(x, y) = (2x-5y, 4x+3y)$ y sea $B = \{\vec{b}_1 = (1, 1), \vec{b}_2 = (-1, 1)\}$ una base del plano $\mathbb{R}^2$. Por lo tanto, si $\vec{x} = (x_1, x_2)$, tenemos*

$$X_N = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$X_B = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$$

donde

$\vec{X} = \alpha_1\vec{b}_1 + \alpha_2\vec{b}_2$

por lo que

$T(\vec{X}) = T(\alpha_1\vec{b}_1 + \alpha_2\vec{b}_2) = \alpha_1 T(\vec{b}_1) + \alpha_2 T(\vec{b}_2)$

o, en notación vertical

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \alpha_1 T\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 T\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \alpha_1 \begin{pmatrix} -3 \\ 7 \end{pmatrix} + \alpha_2 \begin{pmatrix} -7 \\ -1 \end{pmatrix}$$

$$= \begin{pmatrix} -3\alpha_1 - 7\alpha_2 \\ 7\alpha_1 - 1\alpha_2 \end{pmatrix} = \begin{pmatrix} -3 & -7 \\ 7 & -1 \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$$

Vemos que :
$T(X_N) = (T(X))_N = M(X_B)$
donde la matriz M es simplemente la imagen de la base B escrita verticalmente:

$$M = \begin{pmatrix} -3 & -7 \\ 7 & -1 \end{pmatrix}$$

Notemos que $M = T_N^B$, porque de un lado $T(X_N) = (T(X))_N = M(X_B)$, y del otro $T(X_N) = (T(X))_N = T_N^B(X_B)$. Por tanto, para encontrar la matriz de una LT T con respecto a una base B del dominio y con respecto a la base natural N del codominio, podemos usar la siguiente receta (que no vale la pena memorizar):

Para encontrar la matriz T_N^B de una LT T: $V \to W$ con respecto a una base numerada B del dominio V y a la base natural N del codominio W, aplicamos T sobre cada elemento de la base B, y el resultado es reescrito en notación vertical creando una matriz que se denota T_N^B.

Cuando multiplicamos esa matriz por las coordenadas $\vec{X}_B$ del vector $\vec{X}$ en la base B, obtenemos las coordenadas de $T(\vec{X})$ en la base natural: $(T(\vec{X}))_N = T_N^B \vec{X}_B$. Por esa razón, la matriz que representa a T debe denotarse de tal manera que quede claro de una sola mirada que ella recibe coordenadas en la base B, $(\vec{X})_B$, y produce las coordenadas de $T(\vec{X})$ en la base N, $(T(\vec{X}))_N$.

Podemos automatizar la receta anterior si tan sólo consideramos que T_N^B es la matriz formada por $T(B)$ que es también TB donde T es realmente $[T]_N^N$, la matriz de la LT T con respecto a las bases naturales. Concretamente, la matriz T del ejemplo previo es

$$T_N^N = \begin{pmatrix} 2 & -5 \\ 4 & 3 \end{pmatrix}$$

De otra parte, sea B la matriz que codifica en notación vertical la base B:

$$B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Si multiplicamos $T_N^N B$, obtenemos, como se espera, que resulte la matriz T_N^B:

$$[T]_N^N(B) = \begin{pmatrix} & & \vdots & 1 & -1 \\ & & \vdots & 1 & 1 \\ ... & ... & ... & .. & .. \\ 2 & -5 & \vdots & -3 & -7 \\ 4 & 3 & \vdots & 7 & -1 \end{pmatrix}$$

Hemos ilustrado el siguiente principio:

510. $\diamond$ ***Teorema.*** *Para encontrar la matriz T_N^B de $T\colon V \to W$ con respecto a la base B del dominio y a la base natural N en el codominio, multiplicamos la matriz ordinaria de T (con respecto a la base natural) por la matriz formada por la base B en posición vertical.*

Hay aún otro comentario importante. La matriz B formada por los vectores de la base B en posición vertical es la matriz de cambio de coordenadas de la base B a la base N. Por lo tanto, hemos demostrado que $T_N^B = T_N^N I_N^B$, una expresión que dice que para encontrar la matriz T_N^B, nosotros encontramos la matriz ordinaria de T (con respecto a las bases naturales) y después hacemos una interface con un traductor en el dominio que hace un cambio de coordenadas de la base B a la base natural.

511. $\boxed{\textbf{\textit{Ejercicio}}}$ *Especifique el teorema anterior para el caso en el cual $B = N$ y demuestre que se obtiene algo que ya conocíamos.*

512. $\boxed{\textbf{\textit{Ejercicio}}}$ *Encuentre la matriz T_N^B donde:*

a) $T(x,y) = (2x - 8y, 2x + 3y)$, y $B = \{(1,1), (-1,1)\}$

b) $T(x,y) = (x - 7y, 4x - 4y)$, y $B = \{(1,1), (-1,2)\}$

c) $T(x,y) = (5x - 4y, -4x + 6y)$, y $B = \{(1,1), (2,1)\}$

d) $T(x,y) = (2x - 9y, 4x + 7y)$, y $B = \{(1,1), (-2,4)\}$

e) $T(x,y) = (-x - 8y, 4x + 2y)$, y $B = \{(1,1), (-1,-7)\}$

El próximo nivel de generalización es cuando trabajamos con bases arbitrarias en cualquier lado de T. Ese nivel puede ser superado si procedemos con claridad.

513. ***Nuestra tarea.*** *Sea $T\colon V \to W$ una TL y $B_1 = \{\vec{c}_k\}$ una base cualquiera de V y $B_2 = \{\vec{d}_j\}$ una base cualquiera de W. Nuestra tarea es entonces encontrar la matriz de T pero con respecto a las bases dadas. Concretamente:*

Encontrar la matriz $T_{B_2}^{B_1}$ que reciba las coordenadas $\vec{X}_{B_1}$ de un vector $\vec{X}$ en la base del dominio y produzca las coordenadas $(T(\vec{X}))_{B_2}$ de $T(\vec{X})$ en la base del codominio, lo cual lo notamos como

$$(T(\vec{X}))_{B_2} = T_{B_2}^{B_1} \vec{X}_{B_1}$$

La solución a este proyecto se da enseguida.

514. ***Notación.*** *Introducimos dos nuevas notaciones:*

- $B_1 = \{\vec{c}_1, \vec{c}_2, ..., \vec{c}_n\}$ donde n es la dimensión del espacio V podrá ser notado como $B_1 = \{\vec{c}_k\}$, y de modo similar para cualquier otra base.

- Una combinación lineal de la forma $\vec{v} = \alpha_1\vec{d}_1 + \alpha_2\vec{d}_2 + ... + \alpha_n\vec{d}_n$ podrá ser notada como

$$\vec{v} = \sum_{j=1}^{n} \alpha_j \vec{d}_j$$

o como $\vec{v} = \sum\limits_{j=1}^{n} \alpha_j \vec{d}_j$ o como $\vec{v} = \sum\limits_{j} \alpha_j \vec{d}_j$ o como $\vec{v} = \sum \alpha_j \vec{d}_j$. Las dos primeras sumatorias se leen: sumatoria desde $j = 1$ hasta n. La tercera se lee: sumatoria sobre las j, y la cuarta se dice simplemente sumatoria.

515. ***Ejemplos*** *La nueva notación permite acortar la escritura de las demostraciones, para lo cual se usan con frecuencia resultados del siguiente tipo:*

- $\sum \lambda a_i = \lambda \sum a_i$ que dice que para aumentar el sueldo de un año hay que aumentar el sueldo de cada mes.
- $\sum a_i + \sum b_i = \sum(a_i + b_i)$, lo cual dice que lo que gana un grupo de parejas es igual a lo que ganan los maridos más lo que ganan las esposas y también es igual a la suma de lo que ganan todas las parejas.
- $\sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij}$ que dice que si en el mes i, semana j, uno gana a_{ij}, entonces lo que gana en el año es igual a lo que gana en todos los meses, reuniendo lo que se gana en las semanas del mes, y que también es igual a lo que se gana durante la primera semana sumando durante todos los meses, más lo que se gana en la semana número 2 durante todos los meses y así con la semana 3 y 4.
- Si tenemos dos vectores del mismo espacio, descompuestos en la misma base B, $\vec{v} = \sum_i \alpha_i \vec{b}_i$ y $\vec{w} = \beta_i \vec{b}_i$, podemos definir el producto punto con respecto a la base dada B como $\sum_i \alpha_i \beta_i$ que consiste en multiplicar coordenada por coordenada y después sumar los resultados.
- Si tenemos la matriz $M = (m_{ij})$ y un vector $\vec{X} = \alpha_1\vec{d}_1 + \alpha_2\vec{d}_2 + ... + \alpha_n\vec{d}_n$ que representa la descomposición de $\vec{X}$ en la base $\{\vec{d}_k\}$, entonces la expresión $\sum m_{ij}\alpha_j$ da la coordenada i del vector que resulta de multiplicar la matriz por el vector. Más exactamente, $\sum m_{ij}\alpha_j$ es el producto entre la fila i de la matriz por el vector de coordenadas del vector $\vec{X}$ (para lo cual se exige que el número de columnas de la matriz sea igual al número de coordenadas de $\vec{X}$).
- Si uno tiene dos matrices (t_{ij}) y (s_{kl}) que se pueden multiplicar, la multiplicación se puede escribir como $\sum_j t_{ij}s_{jl}$ lo cual da la entrada il de la matriz producto y que es igual al producto punto entre la fila i de (t) por la columna l de (s).

516. $\diamond$ ***Teorema para entender y memorizar.*** *Sea una TL $T\colon V \to W$ y dos bases cualesquiera $B_1 = \{\vec{c}_k\}$ en el dominio y $B_2 = \{\vec{d}_j\}$ en el codominio. Para encontrar la matriz $T_{B_2}^{B_1}$ de T con respecto a las bases dadas se hace lo siguiente: se toma $\vec{c}_1$, el primer vector de la base del dominio, se aplica T sobre $\vec{c}_1$ para obtener*

$T(\vec{c}_1)$. Se descompone dicho vector en la base de llegada B_2 para obtener $(T(\vec{c}_1))_{B_2}$ que son las coordenadas de $T(\vec{c}_1))$ en la base de llegada. Esas coordenadas se escriben como la primera columna de la matriz buscada, $T_{B_2}^{B_1}$. Se itera lo mismo con el segundo vector y con todos los demás hasta terminar.

Antes de ver la demostración, veamos dos ejemplos.

517. *Ejemplo* *$B = \{(1,3),(2,-5)\}$ es una base del plano. Supongamos que T es tal que $T(1,3) = 5(1,3)$ y que $T(2,-5) = 7(2,-5)$. Encontremos la matriz de T con respecto a B, es decir, T_B^B.*

Encontramos la imagen por T de cada uno de los vectores de la base del dominio y la descomponemos en la base del codominio y las coordenadas resultantes las ponemos en notación vertical formando una matriz.

Comencemos con $(1,3)$. Su imagen es $5(1,3)$ que descompuesta en la base B produce $T(1,3) = 5(1,3) = 5(1,3) + 0(2,-5)$. Por lo que las coordenadas correspondientes son $(5,0)$, que forman la primera columna de la matriz buscada.

Similarmente, $T(2,-5) = 7(2,-5) = 0(1,3)+7(2,-5)$. Las coordenadas correspondientes son $(0,7)$. Por tanto,

$$T_B^B = \begin{pmatrix} 5 & 0 \\ 0 & 7 \end{pmatrix}.$$

518. *Ejemplo* *Sea $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ tal que con respecto a cierta base $B_1 = \{\vec{c}_1, \vec{c}_2\}$ en el dominio y a la base $B_2 = \{\vec{d}_1, \vec{d}_2\}$ en el codominio, $T(\vec{c}_1) = 5\vec{d}_2 + 7\vec{d}_2$ y*
$T(\vec{c}_2) = -3\vec{d}_2 + 4\vec{d}_2$, entonces

$(T(\vec{c}_1))_{B_2} = (5,7)$
$(T(\vec{c}_2))_{B_2} = (-3,4)$

y por tanto

$$T_{B_2}^{B_1} = \begin{pmatrix} 5 & -3 \\ 7 & 4 \end{pmatrix}.$$

Ahora veamos la demostración del teorema.

Demostración. Tomemos $\vec{c}_1$, el primer elemento de la base B_1 en el dominio y pasémoslo por T para obtener $T(\vec{c}_1)$ que debe ser descompuesto en la base B_2 del codominio: $T(\vec{c}_1) = \sum_j t_{j1}\vec{d}_j$. Las coordenadas de $T(\vec{c}_1)$ en la base B_2 son $(T(\vec{c}_1))_{B_2} = (t_{j1})$. El subíndice 1 es para denotar que hemos encontrado la primera columna de la matriz buscada $T_{B_2}^{B_1}$.

Hacemos lo mismo con $\vec{c}_k$. Puesto que la imagen de $\vec{c}_k$ es un vector en W, puede ser descompuesto en la base $\{\vec{d}_j\}$:

$T(\vec{c}_k) = \sum_j t_{jk}\vec{d}_j$.

Notemos que hemos puesto el subíndice de $\vec{c}_k$ en segundo lugar en t_{jk} porque queremos que la imagen de un vector número k de la base sea la columna número k de la matriz, y el número de la columna se escribe en segundo lugar en los subíndices de las entradas de la matriz, puesto que el primer lugar es para la fila. En resumen:

$$T_{B_2}^{B_1} = (t_{jk})$$

Demostremos ahora que esta matriz toma las coordenadas de $\vec{X}$ en la abse B_1 y produce las coordenadas de $T(\vec{X})$ en la abse B_2.

Sea $\vec{X}$ en el dominio que puede ser descompuesto en la base dada: $\vec{X} = \sum_k \alpha_k c_k$. Las coordenadas de $\vec{X}$ en B_1 son los (α_k). Pasemos $\vec{X}$ por T para obtener

$T(\vec{X}) = T(\sum_k \alpha_k \vec{c}_k) = \sum_k \alpha_k T(\vec{c}_k)$

donde hemos aplicado la linealidad de T. Ahora recordamos que $T(\vec{c}_k) = \sum_j t_{jk}\vec{d}_j$:

$$T(\vec{X}) = \sum_k \alpha_k \sum_j t_{jk}\vec{d}_j = \sum_k \sum_j \alpha_k t_{jk}\vec{d}_j = \sum_j (\sum_k \alpha_k t_{jk})\vec{d}_j = \sum_j (\sum_k t_{jk}\alpha_k)\vec{d}_j.$$

Lo cual dice que la j-ésima coordenada de $(T(\vec{X}))_{B_2}$ es $\sum_k t_{jk}\alpha_k$ que es el producto punto de la fila j de (t_{jk}) con el vector formado por las coordenadas de $\vec{X}$ en B_1. Eso demuestra que la matriz (t_{jk}) toma las coordenadas $\vec{X}_{B_1}$ y produce $T(\vec{X}))_{B_2}$, las coordenadas de $T(\vec{X})$ en la base B_2:

$$(T(\vec{X}))_{B_2} = T_{B_2}^{B_1}\vec{X}_{B_1}. \blacksquare$$

519. $\boxed{\textbf{\textit{Ejercicio}}}$ *Supongamos que $T: \mathbb{R}^2 \to \mathbb{R}^2$ y consideremos $B_1 = \{(1,3),(2,-5)\}$ como la base del dominio de T mientras que $B_2 = \{(2,6),(6,-15)\}$ es la base en el codominio. T es tal que $T(1,3) = 6(1,3)$ y que $T(2,-5) = 12(2,-5)$. Encuentre la matriz de T con respecto a B_1 y B_2, i.e., encuentre $T_{B_2}^{B_1}$.*

Como hemos podido ver, este teorema permite la resolución inmediata de un tipo especial de problemas como los considerados en los ejemplos. Pero hay muchos más problemas para los cuales el teorema no da solución simple. Para enfrentar problemas más difíciles se requiere estar bien armados, tal como lo aprenderemos en el próximo capítulo.

9.4. Ejercicios de repaso

1. Sea V =gen$(\{1, \mathrm{sen}(x), \cos(x)\})$ y sea $T: V \to V$ definida por $T(f(x)) = f'(x)$ donde f' denota la derivada. Encuentre la matriz de la transformación tomando $\{1, \sin(x), \cos(x)\}$ como base de V.

2. Si T es una transformación lineal tal que

   $T(1,0,0) = (2,3,4,5)$,

   $T(0,1,0) = (1,2,1,0)$ y

   $T(0,0,1) = (0,0,0,1)$, encuentre:

a) $T(1,2,3)$.

b) Una expresión algebraica para T.

c) Espacio nulo, espacio imagen, rango y nulidad de T.

3. Determine si las siguientes afirmaciones son verdaderas o falsas y justifique clara y adecuadamente su respuesta:

   a) Si $A^T A = A$ entonces A es simétrica y $A = A^2$.

   b) Si X_1 es solución de $AX = b$ y X_2 es solución de $AX = b$ entonces $X_1 + X_2$ es solución de $AX = 0$.

   c) Si $A_{3\times 3}$ es una matriz cuyo espacio nulo es una recta que pasa por el origen en $\mathbb{R}^3$, entonces la imagen de A es una recta que pasa por el origen en $\mathbb{R}^3$.

4. Encuentre una transformación lineal $T\colon \mathbb{R}^3 \to \mathbb{R}^3$, cuyo núcleo sea el plano $2x + 3y + z = 0$.

5. Considere la transformación $T\colon P_{\leq 2} \to P_{\leq 2}$ definida por $T(p(x)) = [xp(x)]'$. Sean $B = \{1, 1+x, 1+x^2\}$ y $B' = \{1, 1+2x, x^2\}$ bases del espacio de salida y del espacio de llegada respectivamente.

   a) Encuentre la matriz de representación de T con respecto a las bases B, B'.

   b) Utilice la matriz anterior para encontrar la imagen de $q(x) = x^2 - x$.

6. Sea $T\colon \mathbb{R}^2 \to \mathbb{R}^2$, la reflexión con respecto a la recta $y = -x$ de un vector $\vec{u} \in \mathbb{R}^2$.

   a) Obtenga explícitamente el **operador lineal** (TL de un espacio en sí mismo) que la define.

   b) Halle la matriz de representación de esta transformación en la base canónica.

   c) Trace la imagen del triángulo con vértices en $(-1,4)$, $(3,1)$ y (2,6).

   d) ¿Es T una transformación inyectiva? ¿Es T una transformación sobreyectiva? ¿Es T un isomorfismo?

   e) Sea $T\colon P_{\leq 2} \to P_{\leq 2}$ la transformación lineal definida por $T(p(x)) = p(x-1)$. Considerar las bases $B = \{x^2, x, 1\}$, $B' = \{x, x+1, x^2-1\}$. Encontrar las matrices de representación R_B, $R_{B'}$ de T y una matriz C invertible tal que $R_{B'} = C^{-1} R_B C$.

7. Encuentre la representación matricial de la transformación lineal dada, halle el núcleo, la imagen, el rango y la nulidad de la transformación.

   a) $T\colon \mathbb{R} \to P_3$; $\quad T(a) = a + ax + ax^3$

   b) $T\colon P_4 \to P_4$; $\quad T(p(x)) = xp'(x)$

   c) $T\colon P_3 \to \mathbb{R}^3$; $\quad T(ax^3 + bx^2 + cx + d) = [a-b, a+c, b-c]$

   d) $T\colon \mathbb{R}^2 \to \mathbb{R}^3$; $\quad T[x,y] = [x-y, 2x+y, y]$
   con las bases $B = \{[2,1],[1,2]\}$ $\quad B' = \{[1,-1,0],[0,2,0],[0,2,5]\}$

9.5. Resumen

Lo que hemos llamado en capítulos anteriores la matriz de una TL ha resultado ser la matriz asociada a esa TL, pero con respecto a la base natural. En este capítulo hemos definido la matriz de una TL con referencia a un par cualquiera de bases. Para ello, se encuentra la imagen por T de cada uno de los vectores de la base del dominio y se descompone en la base del codominio y las coordenadas resultantes las ponemos en notación vertical formando una matriz. Eso permite expresar una TL de manera transparente. Por ejemplo, una reflexión puede ser identificada inmediatamente, si se encuentra la base apropiada para representarla.

CAPÍTULO 10

LA MATRIZ DE LA COMPUESTA

En el capítulo 5 aprendimos a asociar la matriz de una compuesta con el **producto de matrices**, pero con respecto a las bases naturales. En este capítulo aprenderemos la generalización de ese resultado a bases cualesquiera.

10.1. Diagramas

520. $\diamond$ ***Teorema.*** *La matriz de la compuesta de dos TL es el producto de las matrices de los componentes en el orden adecuado y teniendo en cuenta las bases en cada lado. Más formalmente:*

Consideremos dos TL:
$S\colon U \to V$
$T\colon V \to W$

$$\begin{tikzcd} U \arrow[r, "S"] \arrow[dr, "T\circ S"'] & V \arrow[d, "T"] \\ & W \end{tikzcd}$$

Figura 10.1. La composición de dos TL.

y que B_1 es base de U, B_2 es de V y B_3 es de W. Entonces

$$(T\circ S)_{B_3}^{B_1} = (T)_{B_3}^{B_2}(S)_{B_2}^{B_1}.$$

Demostración. Démosles nombres a los elementos de las bases:
la base de U es $B_1 = \{\vec{b}_k\}$, la de V es $B_2 = \{\vec{c}_j\}$ y la de W es $E_3 = \{\vec{d}_i\}$.

Decir que la matriz de S, con respecto a las bases B_1 en el dominio y B_2 en el codominio, es (s_{jk}), es lo mismo que decir que $S(\vec{b}_k) = \sum_j s_{jk}\vec{c}_j$.

De otra parte, si la matriz de T, con respecto a las bases B_2 en el dominio y B_3 en el codominio, es (t_{ij}), eso significa que $T(\vec{c}_j) = \sum_i t_{ij}\vec{d}_i$.

Encontremos ahora la matriz de la compuesta $T \circ S : U \to W$. Lo que tenemos que hacer es hallar las coordenadas en la base B_3 de $(T \circ S)(\vec{b}_k)$.

La imagen por $T \circ S$ de cada elemento de la base de U, $\{\vec{b}_k\}$, es:

$$\begin{aligned}(T \circ S)(\vec{b}_k) &= T(S(\vec{b}_k)) = T(\textstyle\sum_j s_{jk}\vec{c}_j) = \sum_j s_{jk}T(\vec{c}_j) = \sum_j s_{jk}\sum_i t_{ij}\vec{d}_i\\ &= \textstyle\sum_j\sum_i s_{jk}t_{ij}\vec{d}_i\\ &= \textstyle\sum_i(\sum_j t_{ij}s_{jk})\vec{d}_i\end{aligned}$$

lo cual significa que la columna k de la matriz de la compuesta $T \circ S$ tiene en la fila i un término igual a $\sum_j t_{ij}s_{jk}$ que es igual al producto punto entre la fila i de la matriz de T por la columna k de la matriz de S. Enfaticemos ese resultado:

$(T \circ S)_{ik} =$ (fila i de T)(columna k de S).

Lo que esta ecuación significa es que la matriz de la compuesta $(T \circ S)$ es precisamente la multiplicación de la matriz de T por la matriz de S.

El siguiente diagrama resume todo:

$$\begin{tikzcd} U \arrow[r, "S^{B_1}_{B_2}"] \arrow[dr, "(T\circ S)^{B_1}_{B_3}"'] & V \arrow[d, "T^{B_2}_{B_3}"] \\ & W \end{tikzcd}$$

Figura 10.2. La matriz de la compuesta.

$$(T \circ S)^{B_1}_{B_3} = T^{B_2}_{B_3}S^{B_1}_{B_2}. \blacksquare$$

521. ***Ejemplo*** *Leamos, interpretemos y probemos el siguiente resultado:*

$$((T \circ S)(\vec{X})_{B_1})_{B_3} = T^{B_2}_{B_3}S^{B_1}_{B_2}\vec{X}_{B_1}$$

Este teorema dice que la matriz de la compuesta hallada en el teorema anterior toma las coordenadas en la base B_1 de un vector $\vec{X}$ y produce las coordenadas de $T(S(\vec{X}))$ en la base B_3 y da el procedimiento para hacerlo: se toman las coordenadas de entrada, se pasan por la matriz de S y el resultado se pasa por la matriz de T, tamando en cada paso las bases adecuadas.

Demostración. Usando la misma notación del teorema anterior tenemos: si $\vec{X} = \sum_k \alpha_k \vec{b}_k$ entonces $(\vec{X})_{B_1}$, las coordenadas de $\vec{X}$ en la base B_1, son los (α_k).

Teniendo en cuenta que $S(\vec{b}_k) = \sum_j s_{jk}\vec{c}_j$ obtenemos:

$T(S(\vec{X})) = T(S(\sum_k \alpha_k \vec{b}_k)) = T(\sum_k \alpha_k S(\vec{b}_k)) = T(\sum_k \alpha_k \sum_j s_{jk}\vec{c}_j)$
$= \sum_k \alpha_k \sum_j s_{jk} T(\vec{c}_j)$

Reemplazando $T(\vec{c}_j)$ por $\sum_i t_{ij}\vec{d}_i$ obtenemos

$T(S(\vec{X})) = \sum_k \alpha_k \sum_j s_{jk} \sum_i t_{ij}\vec{d}_i = \sum_k \sum_j \sum_i \alpha_k s_{jk} t_{ij}\vec{d}_i = \sum_k \sum_j \sum_i t_{ij} s_{jk} \alpha_k \vec{d}_i$

Podemos organizar este resultado de dos maneras. La primera:

$T(S(\vec{X})) = \sum_i \sum_k (\sum_j t_{ij} s_{jk})\alpha_k \vec{d}_i$

que dice que la matriz de $T \circ S$ es el producto de T por S. Y la segunda manera es

$T(S(\vec{X})) = \sum_i \sum_j (t_{ij}(\sum_k (s_{jk}\alpha_k))\vec{d}_i$

que dice que las coordenadas de $T(S(\vec{X}))$ en la base E_3 se hallan alimentando S con las coordenadas de $\vec{X}$ en la base B_1, y procesando el resultado por T. ■

10.2. Cambio de coordenadas y TL

Podemos utilizar el teorema anterior para obtener una solución elegante al problema de expresar una TL con respecto a bases arbitrarias. Procediendo por sentido común tenemos:

Podemos imaginar que T es un mensaje que contiene la información para ejecutar cierta tarea (ejecutar cierta TL sobre cierto *input*). Esa información puede venir, digamos, en griego. ¿Cómo podremos utilizar dicha información para procesar un *input* que viene en latin y producir un *output* que debe ir in alemán? Pues muy fácil: tomamos el input en latín, lo traducimos al griego, lo procesamos en griego, obtenemos un output en griego, lo traducimos al alemán y listo. Con formalidad, eso se expresa en el teorema siguiente:

522. ◇ ***Teorema.*** *Sea $T\colon \mathbb{R}^n \to \mathbb{R}^m$ una LT. Sea B_1 una base del dominio $\mathbb{R}^n$, y B_2 una base del codominio $\mathbb{R}^m$. Si la matriz de T con respecto a la base natural es T_N^N, entonces la matriz de T con respecto a la base B_1 en el dominio y a B_2 en el codominio es*

$$T_{B_2}^{B_1} = I_{B_2}^N T_N^N I_N^{B_1}$$

donde las matrices de traducción o de cambio de base son $I_N^{B_1}$ que traduce de la base B_1 a N, y la matriz $I_{B_2}^N$ que traduce de la base N a la base B_2, en tanto que T_N^N es la matriz de T con respecto a la base natural del domino y a la base natural del codominio (pueden ser diferentes).

Demostración. Cada espacio tiene su base natural correspondiente, denotada N. Entonces, el diagrama siguiente es conmutativo, es decir, el diagrama presenta dos caminos que producen los mismos resultados. Un camino representa el lado derecho de la ecuación a probar y el otro el lado izquierdo. ■

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "T_N^N"] & \mathbb{R}^m \arrow[d, "I_{B_2}^N"] \\
\mathbb{R}^n \arrow[u, "I_N^{B_1}"] \arrow[r, "T_{B_2}^{B_1}"] & \mathbb{R}^m
\end{tikzcd}$$

Figura 10.3. Diagrama conmutativo para $T_{B_2}^{B_1} = I_{B_2}^N T_N^N I_N^{B_1}$.

523. *Explicación.* *Si sabemos cómo es la matriz de una TL T con respecto a las bases naturales, tanto en el dominio como en el codominio, podemos encontrar la matriz de T con respecto a la base arbitraria B_1 en el dominio y a B_2 en el codominio adaptando traductores apropiados: debemos poner un traductor en el dominio que pase de la base B_1 a la base natural y otro traductor en el codominio que pase de la base natural a la base B_2. Esos traductores son las matrices de cambio de base. $I_N^{B_1}$ es simplemente la base B_1 escrita como una matriz, mientras que el traductor $I_{B_2}^N$ es la matriz $(B_2)^{-1}$.*

524. *Explicación dual.* *La igualdad*

$$T_{B_2}^{B_1} = I_{B_2}^N T_N^N I_N^{B_1}$$

puede ser leída en términos puramente geométricos: la matriz de $T\colon V \to W$ con respecto a las bases B_1 en el dominio y B_2 en el codominio pueden encontrarse como sigue: B_1 representa un plp en el dominio. T transforma ese plp en otro en el codominio. Ese nuevo plp se representa fielmente por $T_N^N I_N^{B_1}$. Cada uno de los vectores de ese plp puede ser descompuesto en la base B_2 gracias a $I_{B_2}^N$, y las coordenadas correspondientes se escriben verticalmente formando una matriz, la cual es $T_{B_2}^{B_1}$. El trabajo de dicha matriz se expresa por la igualdad $(T(\vec{X}))_{B_2} = T_{B_2}^{B_1} \vec{X}_{B_1}$.

El diagrama conmutativo del teorema anterior se usa juntamente con los siguientes resultados formando una maquinaria exquisitamente poderosa:

525. $\diamond$ ***Teorema y ejercicio.*** *Para calcular $I_{B_2}^N T_N^N I_N^{B_1}$ se tiene en cuenta que:*

- *La matriz $I_N^{B_1}$ es simplemente la base B_1 puesta como matriz, el primer vector es la primera columna, el segundo vector es la segunda columna y así sucesivamente.*

- *La matriz T_N^N tiene como columna k la imagen por T del vector e_k de la base natural.*
- *La matriz $I_{B_2}^N$ es la inversa de la matriz B_2, para lo cual hay un procedimiento automático de Guass-Jordan.*

Demostración. La primera aserción generaliza el ejemplo 429, página 208 hecho para dos dimensiones y se prueba de manera general así:

Sea $\vec{X}$ en cualquier $\mathbb{R}^n$ y sea $B = \{\vec{b}_i\}$ una base. Descomponemos $\vec{X}$ en la base: $\vec{X} = \sum \alpha_i \vec{b}_i$. Tomamos la coordenada j-ésima de dicha igualdad:

$(\vec{X})_j = \sum \alpha_i b_{ji}$

donde b_{ji} es la coordenada j-ésima del vector i-ésimo de la base. Esta ecuación puede escribirse como

$(\vec{X})_j = \sum b_{ji} \alpha_i$

lo cual se interpreta directamente como la multiplicación de la base B escrita como matriz, escribiendo en orden los vectores como columnas, por las coordenadas de $\vec{X}$ en la base B. Por lo tanto, la matriz B recibe las coordenadas de X en la base B, las α, y produce las coordenadas de X en la base natural, o sea, el mismo $\vec{X}$. Hemos demostrado que la matriz I_N^B es simplemente la base B puesta como matriz, el primer vector es la primera columna, el segundo vector es la segunda columna y así sucesivamente.

Demostración general de la segunda propiedad: Ejercicio.

Para demostrar la tercera aserción tenemos en cuenta que $I_{B_2}^N$ es la inversa de $I_N^{B_2}$, y como la matriz de $I_N^{B_2}$ es B_2 puesta como matriz, entonces, la matriz de $I_{B_2}^N$ es la inversa de la matriz B_2. ∎

526. **Ejercicio** *Usando diagramas conmutativos apropiados, justifique las igualdades siguientes, encuentre el trabajo de cada matriz y señale la forma expedita de calcular cada término:*

a) $R_N^N = I_N^B R_B^B I_B^N$

b) $R_B^N = R_B^B I_B^N$

c) $R_B^N = I_B^N R_N^N$

d) $R_N^B = I_N^B R_B^B$

e) $R_N^B = R_N^N I_N^B$.

Veamos cómo se utilzan estos teoremas para resolver problemas de geometría.

527. **Ejemplo** *Encontremos la matriz de $T(x, y)$ con respecto a la base natural de $\mathbb{R}^2$, si $B = \{(1, 1), (-2, 2)\}$ y*

$$T_N^B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Solución: **Método 1.** Podemos ver de la matriz que

$T(1,1) = 1\vec{i} + 3\vec{j}$
$T(-2,2) = 2\vec{i} + 4\vec{j}$
Eso se debe a que la matriz T_N^B tiene como columnas las coordenadas en la base natural de las imágenes por T de cada uno de los vectores de la base B.

Debemos encontrar $T(\vec{i})$ y $T(\vec{j})$. Tenemos:
$\vec{i} = \alpha(1,1) + \beta(-2,2)$
por lo que
$1 = \alpha - 2\beta$
$0 = \alpha + 2\beta$
sumando y despejando α obtenemos $\alpha = 1/2$ y $\beta = -1/4$.
por tanto $\vec{i} = 1/2(1,1) - 1/4(-2,2)$ y
$T(\vec{i}) = 1/2T(1,1) - 1/4T(-2,2) = 1/2(1,3) - 1/4(2,4) = (0,1/2)$
Similarmente,
$\vec{j} = \alpha(1,1) + \beta(-2,2)$
entonces,
$0 = \alpha - 2\beta$
$1 = \alpha + 2\beta$
sumando y despejando α obtenemos $\alpha = 1/2$ y $\beta = 1/4$.
Por tanto, $\vec{j} = 1/2(1,1) + 1/4(-2,2)$ y
$T(\vec{j}) = 1/2T(1,1) + 1/4T(-2,2) = 1/2(1,3) + 1/4(2,4) = (1,5/2)$.
Por consiguiente, la matriz de T es:

$$T = T_N^N = \begin{pmatrix} 0 & 1 \\ 1/2 & 5/2 \end{pmatrix}.$$

Método 2. Hay que hallar $T = T_N^N$ y como tenemos $T_N^{B_1}$ usamos la siguiente identidad dada por un diagrama conmutativo $T_N^N = T_N^{B_1} I_{B_1}^N$.

La matriz de $I_{B_1}^N$ es la inversa de B_1 que la hallamos por cofactores: la inversa es uno sobre el determinante por la matriz de cofactores transpuesta. Como

$$B = \begin{pmatrix} 1 & -2 \\ 1 & 2 \end{pmatrix}$$

el determinante es 4, la matriz de cofactores es:

$$C = \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}$$

La matriz de cofactores transpuesta es

$$C^T = \begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}$$

La inversa de B_1 es

$$(B_1)^{-1} = (1/4)\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}$$

Como $T_N^N = T_N^{B_1} I_{B_1}^N$ tenemos que

$$T_N^N = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}(1/4)\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix} = (1/4)\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1/2 & 5/2 \end{pmatrix}.$$

528. $\boxed{\textbf{\textit{Ejercicio}}}$ *Encontrar por dos métodos diferentes $T(x,y)$ si*

a) $B_1 = \{(1,2),(-2,3)\}$ y

$$T_N^{B_1} = \begin{pmatrix} 1 & -1 \\ 3 & 4 \end{pmatrix}$$

b) $B_1 = \{(-1,1),(2,5)\}$ y

$$T_N^{B_1} = \begin{pmatrix} 1 & 5 \\ -3 & 1 \end{pmatrix}$$

c) $B_1 = \{(-1,-1),(-2,2)\}$ y

$$T_N^{B_1} = \begin{pmatrix} -2 & -1 \\ -3 & 5 \end{pmatrix}$$

Continuemos con la depuración de nuestra maquinaria.

529. $\diamond$ ***Teorema.*** *El determinante de una TL T es independiente de la base.*

Demostración. Hemos desarrollado en capítulos previos un método para calcular el determinante de una TL T con respecto a la base natural: $DetT = DetT_N^N$. Ahora probaremos que también podemos utilizar cualquier base y da lo mismo:

$$DetT_N^N = DetT_B^B$$

Para verlo, comencemos notando que
$T_B^B = I_B^N T_N^N I_N^B$

Recordando que $DetT$ puede entenderse como el factor de amplificación de un volumen signado de un plp dado, entonces concluimos que $Det(S \circ T) = DetSDetT$. Por tanto,
$DetT_B^B = DetI_B^N DetT_N^N DetI_N^B$.

Recordemos ahora que si una TL T es invertible, tiene determinante diferente de cero, y entonces $Det(T^{-1}) = 1/DetT$, porque si un proceso amplifica por el factor k, el proceso inverso amplifica por el factor $1/k$. Reconociendo que $I_B^N = (I_N^B)^{-1}$, concluimos que $Det(I_N^B)^{-1} = 1/DetI_B^N$. Por tanto:

$$\begin{aligned} DetT_B^B &= DetI_B^N DetT_N^N Det(I_B^N)^{-1} \\ &= DetI_B^N DetT_N^N (1/DetI_B^N) = DetT_N^N \ \blacksquare \end{aligned}$$

530. $\boxed{\textit{Ejercicio}}$ *Ilustre el teorema anterior con un cálculo explícito de DetT en la base natural y en la base dada B:*

a) $T(x, y) = (3x - 5y, 4x + 4y)$, $B = \{(1, 1), (-1, 1)\}$

b) $T(x, y) = (2x - 1y, 4x + 3y)$, $B = \{(1, 1), (-2, 2)\}$

c) $T(x, y) = (x - 4y, 3x - 4y)$, $B = \{(2, 2), (-1, 1)\}$

d) $T(x, y) = (-1x - 3y, -3x - 3y)$, $B = \{(2, 2), (-2, 2)\}$

e) $T(x, y) = (4x - 5y, 4x + 7y)$, $B = \{(1, 1), (-1, 1)\}$.

531. $\boxed{\textit{Ejercicio}}$ *Sea* $\vec{X} = (x_1, x_2, ..., x_n)$, $\vec{Y} = y_1, y_2, ..., y_n)$ *vectores de* $\mathbb{R}^n$. *Nosotros definimos el producto punto, interno o interior como* $\vec{X}\cdot\vec{Y} = \sum x_i y_i$. *Claramente, esa es una definición ligada a la base natural. Si damos otra base B y tenemos las coordenadas de* $\vec{X}$, $\vec{Y}$ *en esa base, ¿cómo ha de calcularse el producto punto? Considere primero el caso de una base ortonormal.*

532. ***Ejemplo*** *Estudiemos la reflexión R con respecto a la línea* $y = x$.

a) Para encontrar la matriz de esta TL con respecto a las bases naturales, debemos descubrir qué hace R sobre la base natural $\{(1, 0), (0, 1)\}$. Esa reflexión intercambia el eje $\vec{X}$ con el eje $\vec{Y}$. Por tanto, $R(1, 0) = (0, 1)$ y $R(0, 1) = (1, 0)$.

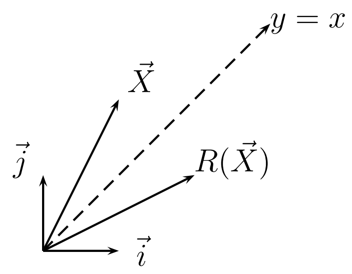


Figura 10.4. Una reflexión R con respecto al eje $y = x$ intercambia $\vec{i}$ y $\vec{j}$ y a $\vec{X}$ lo transforma en $R(\vec{X})$.

Por tanto, la matriz de T con respecto a la base natural es

$$R = R_N^N = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

escrita de esa forma, la matriz de R enfatiza el hecho de que R intercambia los ejes. Observemos que $R^2 = \mathbb{I}$ y que la inversa de R es ella misma.

b) Encontremos una base B, en la cual la matriz correspondiente enfatice que R es una reflexión. Por supuesto, un elemento de la base debe ser un vector que genere el eje de reflexión, en este caso, el eje es la línea $y = x$, mientras que el otro vector de esa base podría ser perpendicular al primero, un generador de la línea $y = -x$. Por tanto, $B = \{(1, -1), (1, 1)\}$. A pesar de que notamos a B como conjunto, es, por contexto, un conjunto numerado, ordenado. Escogemos el orden sugerido por la lista de B porque queremos que el vector $(1, -1)$ juegue el papel del eje $\vec{X}$ en relación con una reflexión con respecto al eje $\vec{Y}$.

Por tanto:
$R_B^B = I_B^N R_N^N I_N^B$
Tomando en cuenta que

$$I_N^B = B = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

y que

$$I_B^N = B^{-1} = \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

Obtenemos

$$R_B^B = B^{-1} R_N^N B \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

Al final

$$R_B^B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

R es claramente una reflexión. Observemos que $(R_B^B)^2 = \mathbb{I}$ y que la inversa de R_B^B es ella misma.

533. ***Intriga.*** *Dos veces una reflexión da la identidad, por lo tanto una reflexión es su propia inversa. Además, hemos encontrado que eso es cierto en dos bases diferentes. ¿Es una coincidencia o es cierto siempre?*

534. ***Ejemplo*** *Estudiemos la reflexión R con respecto a la línea $y = -3x$.*

Esta vez ya no podemos encontrar a simple vista la matriz de R con respecto a alguna base. Encontremos, en primer lugar, una base B en la cual R sea fácilmente calculable.

Por supuesto, $B = \{\vec{b}_1 = (1, 1/3), \vec{b}_2 = (-1, 3)\}$, porque $(-1, 3)$ genera la línea que sirve de eje de reflexión, mientras que el otro vector ayuda a formar una base ortonormal y con determinante positivo. De esa forma, podemos pensar que el primer vector es un generador del eje horizontal y que el segundo lo es del eje vertical, con la dirección positiva apuntando como uno la espera.

En la base B tenemos:
$R(\vec{b}_1) = -\vec{b}_1$
$R(\vec{b}_2) = \vec{b}_2$
Esto significa que
$(R(\vec{b}_1))_B = (-1, 0)$

$(R(\vec{b}_2))_B = (0, 1)$
Por lo tanto:

$$R_B^B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Construyamos ahora la matriz de R con respecto a la base natural. Usamos la igualdad
$R_N^N = I_N^B R_B^B I_B^N$.
Tomando en cuenta que

$$I_N^B = B = \begin{pmatrix} 1 & -1 \\ 1/3 & 3 \end{pmatrix}$$

y que

$$I_B^N = B^{-1} = \begin{pmatrix} 9/10 & 3/10 \\ -1/10 & 3/10 \end{pmatrix}$$

Entonces

$$R_N^N = BR_N^N B^{-1} = \begin{pmatrix} 1 & -1 \\ 1/3 & 3 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 9/10 & 3/10 \\ -1/10 & 3/10 \end{pmatrix}$$

Como fue prometido, el resultado neto es:

$$R_N^N = \begin{pmatrix} -8/10 & -6/10 \\ -6/10 & 8/10 \end{pmatrix} = \begin{pmatrix} -4/5 & -3/5 \\ -3/5 & 4/5 \end{pmatrix}$$

Nos damos cuenta de que cada vector columna tiene norma uno, como debe ser, y que la matriz es simétrica. Esa es la matriz de una TL definida por
$R(x, y) = (-4x/5 - 3y/5, -3x/5 + 4y/5)$.
¿Adivinaría usted que esta expresión representa una reflexión? ¿Hay alguna manera de saberlo?

535. ***Ejemplo*** *Calculemos la ecuación del círculo $R(C)$ que es la imagen del círculo C cuya ecuación es $(x-8)^2 + (y+10)^2 = 1$ por la reflexión R con respecto a la línea $y = x$.*

Primero saquemos la respuesta por sentido común: la reflexión de un círculo debe ser un círculo con el mismo radio y centrado en el punto que resulta de reflejar el centro del círculo inicial. Como se refleja con respecto a la línea $y = x$, la reflexión de $(8, -10)$ es $(-10, 8)$. Por lo que el círculo reflejado es $(x + 10)^2 + (y - 8)^2 = 1$.

Procedamos ahora con nuestra maquinaria:
El punto (w, z) pertenece a $R(C)$ ssi $R^{-1}((w, z))$, la imagen inversa de (w, z), satisface la ecuación del círculo C. Así debemos encontrar la reflexión inversa. Afortunadamente la inversa de una reflexión es la misma reflexión: $R^{-1} = R$. Por lo tanto,
$(w, z) \in R(C)$ ssi $R^{-1}(w, z) \in C$ ssi $R(w, z) \in C$
Recordando que

$$R = R_N^N = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

entonces $R(w,z)$ está dado por una multiplicación:

$$R(w,z) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} w \\ z \end{pmatrix} = \begin{pmatrix} z, w \end{pmatrix}$$

Estas dupletas deben satisfacer la ecuación $(x-8)^2+(y+10)^2 = 1$, lo cual significa que $(z-8)^2+(w+10)^2 = 1$, o si uno prefiere $(y-8)^2+(x+10)^2 = 1$, tal como debe ser. Ahora que sabemos que nuestra maquinaria sirve para los casos fáciles, usémosla para un caso más difícil.

536. **Ejercicio** *Enriquezca el ejemplo anterior con un dibujo apropiado.*

537. ***Ejemplo*** *Estudiemos la ecuación de la imagen de una elipse que es reflejada con respecto a la línea $y = -3x$.*

Podemos fabricar una elipse si magnificamos un círculo de manera desigual por ambos ejes. Podemos hacer una magnificación utilizando $M(x,y) = (5x, 4y)$. Nuestra elipse E será el resultado de magnificar por M un círculo C con ecuación $x^2+y^2 = 1$.
Vemos que $(w,z) \in E$ *ssi* $M^{-1}(w,z) \in C$.
Claramente, $M^{-1}(w,z) = (w/5, z/4)$ es un punto en el círculo C ssi
$(w/5)^2+(z/4)^2 = 1$.
O, si uno prefiere, la ecuación de la elipse es
$(x/5)^2+(y/4)^2 = 1$.
pero por otro lado, la matriz de reflexión R con respecto a la línea $y = -3x$ y con respecto a la base natural es

$$R = R_N^N = \begin{pmatrix} -8/10 & -6/10 \\ -6/10 & 8/10 \end{pmatrix}$$

Esta matriz es su propia inversa. Por tanto, $(w, z \in R(E)$ ssi
$R^{-1}(w,z) = R(w,z) \in E$.
Esto sucede cuando $(-8w/10 - 6z/10, 6w/10+8z/10)$ satisface la ecuación de E, $(x/5)^2+(y/4)^2 = 1$, i.e., ssi
$((-8w/10-6z/10)/5)^2 + ((6w/10+8z/10)/4)^2 = 1$.

538. **Ejercicio** *Dibuje una figura para ilustrar el ejercicio anterior. Invente un test para verificar que todo ha sido bien hecho.*

539. **Ejercicio** *Sea i el número complejo $\sqrt{-1}$. Estudie el efecto de la magnificación $M(x,y) = (x, iy)$ sobre el círculo.*

540. **Ejercicio** *Una cónica es el resultado de la intersección de un cono (de dos faldas) con un plano. Puede dar un círculo, una elipse, una parábola, una hipérbola, o un par de líneas que se intersecan en el origen. Haga un dibujo ilustrando las cónicas.*

541. **Ejercicio** *Estudie las reflexiones con respecto a las siguientes líneas y encuentre la ecuación de la imagen por la reflexión de la cónica dada:*

a) $y = -x$, cónica: $y = x^2$
b) $y = -2x$, cónica: $y^2 - x^2 = 1$
c) $y = -x/2$, cónica: $y = 7x$
d) $y = 3x + 1$, $x^2 = y^2$. Ayuda: sea cauteloso.

Recordemos que una base es ortogonal cuando sus vectores son mutuamente perpendiculares y que es ortonormal cuando además tienen norma uno. Es recomendable que cuando uno tenga la libertad de escoger los vectores a su gusto, que el primero sea escogido en el primer cuadrante, $\mathbb{R}^2$, o en el primer octante, en $\mathbb{R}^3$ y que tenga determinante positivo. De esa forma uno puede imaginarse la base como si fuese la base natural.

542. **Ejercicio** *Considere el plano* Π*, cuya ecuación es* $x + 2y + 3z = 0$.

a) Verifique que Π es un sub-EV. Pruebe que su dimensión es dos.

b) Encuentre una base ortogonal B de $\mathbb{R}^3$ tal que dos de sus vectores sean base del plano y que el otro vector sea el vector normal del plano. Ordene B de tal manera que tenga determinante positivo.

c) Defina R como la reflexión de espejo con respecto al plano π. Encuentre R_B^B.

d) Encuentre R.

e) Encuentre R^{-1}. Compare con $(R_B^B)^{-1}$.

f) La ecuación de un círculo en el plano $\mathbb{R}^2$ es $x^2 + y^2 = 1$. De forma semejante, la ecuación de una esfera en $\mathbb{R}^3$ es $x^2 + y^2 + z^2 = 1$. Defina un elipsoide E como el resultado de la magnificación de una esfera por medio de M donde $M(x, y, z) = (ax, by, cz)$. Encuentre la ecuación del elipsoide.

g) Encuentre la ecuación de la imagen de E por R.

h) Repita la misma tarea con un hiperboloide, que es un círculo magnificado por N donde $N(x, y, z) = (ax, by, cz)$ donde uno o dos coeficientes de a o b o c pueden valer $i = \sqrt{-1}$.

10.3. Rotaciones en 3D

Ya sabemos que una rotación en $\mathbb{R}^2$ tiene una matriz, con respecto a la base natural, que es igual a

$$R = R_N^N = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

Podemos extender esa rotación en el plano a todo el espacio tridimensional si consideramos que el plano es el sub-EV de $\mathbb{R}^3$ generado por los ejes $\vec{X}$ y $\vec{Y}$. La nueva rotación rota los ejes horizontales común y corriente, mientras que al eje $\vec{Z}$ no lo mueve. Por tanto, si $\vec{k} = (0, 0, 1)$ (en la base natural), entonces $R(\vec{k}) = \vec{k}$. Esta nueva rotación se denota también por R y tiene como matriz

$$R = R_N^N = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

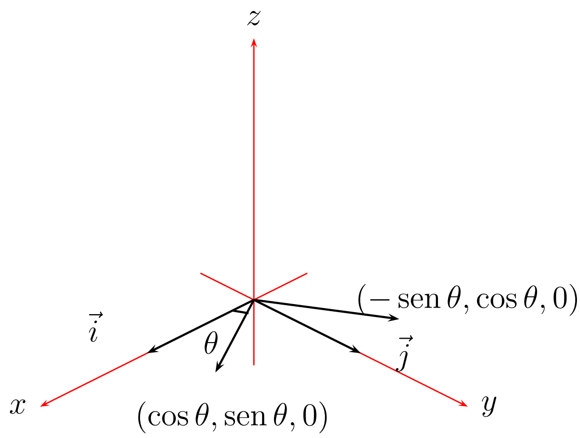


Figura 10.5. Una rotación de ángulo θ alrededor del eje Z.

De ahora en adelante, toda rotación es una rotación alrededor de algún eje.

543. $\boxed{\textbf{\textit{Ejercicio}}}$ *Generalice a 3D el teorema 499, página 227 sobre la ortogonalidad de las matrices de rotación.*

544. ***Ejemplo*** *Encontremos la matriz de rotación en sentido contrario a las manecillas del reloj de ángulo* $\pi/3$ *alrededor del eje generado por* $(1, 1, 1)$.

Imaginemos que la línea generada por $(1, 1, 1)$ es un nuevo eje Z. Este vector es normal al plano $x + y + z = 0$. Debemos encontrar de alguna manera una base para ese plano, el cual es un sub-EV de $\mathbb{R}^3$ que pasa por el origen. Podemos tomar cualquier vector de ese plano, digamos, $(1, 0, -1)$ como un elemento de la base y otro puede ser $(1, 0, -1) \times (1, 1, 1) = (1, -2, 1)$.

La tríada $(1, -2, 1), (1, 0, -1), (1, 1, 1)$ forma una base de $\mathbb{R}^3$ con la misma orientación que la base natural, tal que los dos primeros vectores forman una base del plano. Como todos esos vectores son mutuamente perpendiculares, la correspondiente base ortonormal se obtiene normalizando

$$B = I_N^B = \begin{pmatrix} \sqrt{6}/6 & \sqrt{2}/2 & \sqrt{3}/3 \\ -2\sqrt{6}/6 & 0 & \sqrt{3}/3 \\ \sqrt{6}/6 & -\sqrt{2}/2 & \sqrt{3}/3 \end{pmatrix}$$

Como es una base ortonormal, su inversa es su transpuesta:

$$B^{-1} = B^T = I_B^N = \begin{pmatrix} \sqrt{6}/6 & -2\sqrt{6}/6 & \sqrt{6}/6 \\ \sqrt{2}/2 & 0 & -\sqrt{2}/2 \\ \sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3 \end{pmatrix}$$

El determinante de B es igual a $+1$, por lo que podemos pensar en B como si fuese la base natural N. Por tanto, la rotación de $\pi/3$ alrededor de la línea generada por $(1,1,1)$ vista desde la base B es:

$$R_B^B = \begin{pmatrix} cos(\pi/3) & -\operatorname{sen}(\pi/3) & 0 \\ \operatorname{sen}(\pi/3) & cos(\pi/3) & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

esto significa que si movemos apropiadamente la cabeza para ver la rotación en acción, se vería exactamente como si se hiciera sobre el plano ordinario.

Conociendo la matriz de la rotación con respecto a la base B, la matriz de la rotación con respecto a la base natural es

$R = R_N^N = I_N^B R_B^B I_B^N.$

La matriz I_N^B es simplemente B, la cual es ortonormal, y por tanto

$I_B^N = (I_B^N)^{-1} = B^{-1} = B^T.$

Por lo que:
$R = BR_B^B B^T =$

$$= \begin{pmatrix} \sqrt{6}/6 & \sqrt{2}/2 & \sqrt{3}/3 \\ -2\sqrt{6}/6 & 0 & \sqrt{3}/3 \\ \sqrt{6}/6 & -\sqrt{2}/2 & \sqrt{3}/3 \end{pmatrix} \begin{pmatrix} 1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{6}/6 & -2\sqrt{6}/6 & \sqrt{6}/6 \\ \sqrt{2}/2 & 0 & -\sqrt{2}/2 \\ \sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3 \end{pmatrix}$$

545. **_Ejercicio_** *El ejemplo anterior involucra una buena dosis de conceptos y de cálculos. ¿Cómo podremos saber que en realidad lo hicimos todo bien? Podemos explotar la siguiente idea: si uno itera una rotación de $\pi/3$ tres veces, le da una rotación de π. Pero iterar una rotación 3 veces es lo mismo que elevar la matriz correspondiente al cubo. Pruebe que la matriz R del ejercicio anterior, elevada al cubo, da la matriz de una rotación de ángulo π.*

10.4. Trayectorias sobre el plano

Una trayectoria es una función que a cada instante de tiempo t le asocia una posición en un espacio dado.

546. ***Ejemplo*** *Comencemos con una partícula que orbita con velocidad angular uniforme formando un círculo sobre el plano generado por los ejes horizontales. Su trayectoria se describe por*

$\vec{r}(t) = (\cos t, \operatorname{sen} t, 0)$, $0 \leq t \leq 2\pi$.

Si $t = 0$, la partícula está en $(1, 0, 0)$. Si $t = \pi/2$ la partícula está en el punto $(0, 1, 0)$. Si $t = 2\pi$ la partícula alcanza su punto de partida. Observemos que el círculo se describe en sentido antihorario.

547. ***Ejercicio*** *Encuentre la trayectoria de una partícula puntual que orbita con velocidad angular uniforme alrededor del origen formando un círculo en el plano YZ.*

548. ***Ejercicio*** *Someta la trayectoria $\vec{r}(t) = (\cos t, \operatorname{sen} t, 0)$ al efecto de la magnificación $M(x, y) = (3x, 5y, 0)$ y encuentre la trayectoria del movimiento resultante.*

549. ***Ejemplo*** *Consideremos ahora el caso en el cual el movimiento circular es sobre el plano $x + y + z = 0$.*

Ya sabemos que existe una base ortonormal B de $\mathbb{R}^3$ con la misma orientación que la base natural y tal que sus dos primeros vectores son una base del plano (ver ejemplo 544):

$$B = (\vec{e}_1, \vec{e}_2, \vec{e}_3) = \begin{pmatrix} \sqrt{6}/6 & \sqrt{2}/2 & \sqrt{3}/3 \\ -2\sqrt{6}/6 & 0 & \sqrt{3}/3 \\ \sqrt{6}/6 & -\sqrt{2}/2 & \sqrt{3}/3 \end{pmatrix}$$

Si la partícula es vista desde la base B, el movimiento luce como un movimiento circular uniforme sobre un plano:

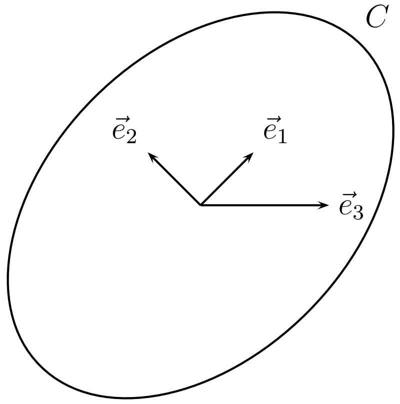


Figura 10.6. Bases apropiadas simplifican las descripciones.

$(\vec{r}(t))_B = (\cos t, \operatorname{sen} t, 0)$.
Por tanto, $(\vec{r}(t))_N = I_N^B(\vec{r}(t))_B = B(r(t))_B$, esto es

$$(\vec{r}(t))_N = \begin{pmatrix} \sqrt{6}/6 & \sqrt{2}/2 & \sqrt{3}/3 \\ -2\sqrt{6}/6 & 0 & \sqrt{3}/3 \\ \sqrt{6}/6 & -\sqrt{2}/2 & \sqrt{3}/3 \end{pmatrix} \begin{pmatrix} \cos t \\ \operatorname{sen} t \\ 0 \end{pmatrix}$$

550. $\boxed{\textbf{\textit{Ejercicio}}}$ *Traslade el movimiento de la partícula del ejercicio previo al plano* $x + y + z = 1$ *y encuentre la trayectoria resultante.*

551. $\boxed{\textbf{\textit{Ejercicio}}}$ *Encuentre la ecuación del elipsoide* $x^2/4 + y^2/9 + z^2/16 = 1$ *que es rotado alrededor del eje generado por* $(1, 1, 1)$ *un ángulo de* $\pi/3$.

10.5. Ejercicios de repaso

1. En $\mathbb{R}^2$, halle la matriz con respecto a la base natural de la reflexión con respecto a la línea $y = 9x$. Halle la imagen del punto $(2, -8)$. Halle la imagen de la parábola $y = x^2$. Halle la imagen del círculo $x^2 + y^2 = 4$. Halle la imagen del círculo $(x - 1)^2 + (y - 9)^2 = 1$. Ilustre las discusiones con dibujos apropiados.

2. En $\mathbb{R}^3$, halle la matriz con respecto a la base natural de la reflexión con respecto al plano $y - x + 3z = 2$. Halle la imagen del punto $(2, -8, 1)$. Halle la imagen de la línea $y = x + 4$. Halle la imagen del círculo del plano $x^2 + y^2 = 4$. Halle la imagen de la esfera $(x - 1)^2 + (y - 9)^2 + (z - 3)^2 = 1$. Ilustre las discusiones con dibujos apropiados.

3. En $\mathbb{R}^3$, halle la matriz con respecto a la base natural de la rotación de ángulo $\pi/6$ con respecto al eje determinado por el vector $(1, -2, 3)$. Halle la imagen de la línea $x + y + z = 1, x - y + z = 2$. Halle la imagen de la esfera

   $(x - 1)^2 + (y - 9)^2 + (z - 3)^2 = 1$.

   Ilustre las discusiones con dibujos apropiados.

10.6. Resumen

La libertad para representar una TL con respecto a cualquier par de bases ha hecho posible imponentes tareas geométricas gracias a que con bases apropiadas se simplifican las descripciones. En particular, hemos podido extender la regla de multiplicar matrices para hallar la matriz de la compuesta con respecto a bases arbitrarias. De eso hemos podido deducir naturalmente la forma de cambiar de coordenadas y la representación de las TL con respecto a cualquier base.

CAPÍTULO 11

DIAGONALIZACIÓN

La **diagonal** de una matriz A es el conjunto de las entradas A_{ii}. Una **matriz diagonal** es una matriz cuadrada cuyas entradas fuera de la diagonal son todas cero, es decir, toda la información relevante está sobre la diagonal. La matriz identidad es diagonal. El objetivo del presente capítulo es demostrar que en muchos casos, una matriz es equivalente a una matriz diagonal, en el sentido de que ambas representan la misma información, la misma TL, pero con respecto a bases diferentes.

Matrices equivalentes son igual de importantes. Pero puede pasar que para problemas y preguntas específicas, una forma en particular sea más importante que las otras. Por ejemplo, una matriz diagonal es muy transparente. Además, las matrices diagonales son muy fáciles de multiplicar, una propiedad que nos será muy útil para el estudio de los sistemas dinámicos. Por eso que en el presente capítulo estudiaremos el problema de encontrar el equivalente diagonal de una matriz.

A lo largo del capítulo estaremos tratando con matrices cuadradas, es decir, con TL de un espacio es sí mismo, o con **operadores lineales**.

Una concisa introducción a la diagonalización puede encontrarse en el ejemplo 581.

552. ***Ejemplo*** *La importancia de la diagonalización se aprecia fácilmente al resolver tareas geométricas. Hemos aprendido en el capítulo pasado que una elipse E que obedece la ecuación $x^2/9 + y^2/4 = 1$ y que es rotada $\pi/4$ en sentido antihorario se convierte en otra elipse que tiene como ecuación $13x^2 + 13y^2 - 10xy = 72$. A dicha elipse podemos asociar la matriz*

$$E = \begin{pmatrix} 13/72 & -5/72 \\ -5/72 & 13/72 \end{pmatrix}$$

tal que

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 13/72 & -5/72 \\ -5/72 & 13/72 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{13}{72}x^2 + \frac{13}{72}y^2 - \frac{10}{72}xy = 1$$

Todo eso se refería a la base natural. Por eso, preferimos notar E como E_N^N. Sin embargo, una elipse rotada puede ser inmediatamente reconocida como una elipse, algo que no pasa con nuestra ecuación. El remedio es reescribir la ecuación de la elipse rotada en una base apropiada, B, desde la cual se vea derecha. La base B, escrita en forma matricial, es

$$B = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

Puesto que la matriz B representa una base ortonormal, su inversa es la transpuesta:

$$B^{-1} = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

Si las coordenadas de un punto en la base B se denotan como (w, z), la elipse luce como

$w^2/9 + z^2/4 = 1$

y podemos decir inmediatamente que se trata de una elipse. Esta información puede ser codificada directamente en la matriz

$$E_B^B = \begin{pmatrix} 1/9 & 0 \\ 0 & 1/4 \end{pmatrix}$$

donde hemos enfatizado que esa matriz representa, en la base B, la misma información que la matriz E_N^N.

De todas maneras, las dos matrices representan la misma información, pero con referencia a bases diferentes y, por lo tanto, están relacionadas por una matriz de cambio de base. En nuestro caso:

$E_B^B = I_B^N E_N^N I_N^B = B^{-1} E_N^N B$
$E_N^N = I_N^B E_B^B I_B^N = B E_B^B B^{-1}$

En el caso de la elipse, pudimos fabricar una base ortonormal y por tanto la matriz inversa B^{-1} era simplemente la transpuesta. Pero no siempre las cosas son tan fáciles porque no siempre se goza de tanta simetría como con la elipse.

11.1. Matrices diagonales

553. ♣ *Definiciones.* *Una matriz cuadrada M es* ***diagonal*** *cuando todas sus entradas no nulas están sobre la diagonal. Las entradas en la diagonal pueden ser iguales o distintas de cero. Una matriz diagonal puede ser notada como un vector, de tal forma que la entrada M_{ii} de la matriz está asociada con la entrada i-ésima del vector.*

554. ***Ejemplo*** $D(1, 2, 3)$ *es la matriz diagonal*

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

555. **Ejemplo** *La matriz diagonal*

$$D(1,4,1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

representa una LT en la base natural. Veamos cómo se ve desde la base

$$B = \begin{pmatrix} 1 & -1 & 1 \\ 0 & -1 & -2 \\ 1 & 1 & -1 \end{pmatrix}$$

Esta base representa la matriz de cambio de coordenadas I_N^B. La matriz inversa representa el cambio contrario

$$I_B^N = B^{-1} = \begin{pmatrix} 1/2 & 0 & 1/2 \\ -1/3 & -1/3 & 1/3 \\ 1/6 & -1/3 & -1/6 \end{pmatrix}$$

Tenemos: $D_B^B = I_B^N D_N^N I_N^B$

Después de los reemplazos adecuados, obtenemos:

$$D_B^B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 1 & 3 \end{pmatrix}$$

Observemos que esta matriz no es simétrica. Sin embargo, cuando se trata de una base ortonormal tenemos el resultado siguiente:

556. $\boxed{\textbf{Ejercicio}}$ *Pruebe que para una matriz diagonal D el producto $O^T DO$ es simétrico, donde O es una base ortonormal.*

557. $\boxed{\textbf{Ejercicio}}$ *Sea la base*

$$B = \begin{pmatrix} 1 & -1 & 1 \\ 0 & -1 & -2 \\ 1 & 1 & -1 \end{pmatrix}$$

Reexpresar en la base B las siguientes matrices diagonales:

a) $D(1,2,1)$

b) $D(-1,2,1)$

c) $D(1,-1,1)$

d) $D(1,-3,-4)$

e) $D(-2,1,-4)$

558. $\boxed{\textbf{\textit{Ejercicio}}}$ *Pruebe que las matrices diagonales conmutan entre ellas, esto es,* $D_1D_2 = D_2D_1$. *Proponga una explicación geométrica de ese hecho. Pruebe que esa propiedad se conserva aun si su naturaleza diagonal se esconde detrás de un cambio de coordenadas. Demuéstrelo.*

11.2. Magnificaciones

Una matriz diagonal, interpretada como la matriz de cierta TL puede verse como una magnificación. Para ilustrar eso, notemos que para la matriz diagonal $D(1,2,3)$ tenemos:

$D(1,0,0) = 1(1,0,0)$
$D(0,1,0) = 2(0,1,0)$
$D(0,0,1) = 3(0,0,1)$

Por tanto, D amplifica una vez en la dirección X, dos veces en la dirección Y. y tres veces en la dirección Z. D es una magnificación anisotrópica, es decir, que magnifica de forma diferente en direcciones diferentes. Puesto que nunca se especificó base alguna, se está tratando con la base natural. Por supuesto, una magnificación no debe estar esclavizada a dicha base. La gráfica siguiente junto a las definiciones y ejercicios formalizarán dicha idea. Todo es algo muy simple. Sin embargo, una magnificación es tratada en la literatura como algo muy esotérico, con nombres complicados. Por la fuerza de la costumbre, aprenderemos a usarlos.

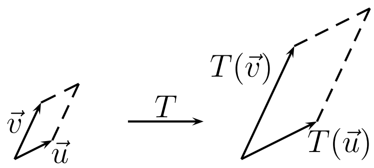


Figura 11.1. Una magnificación es simple y exige una codificación simple.

559. ♣ ***Definición.*** *Si hay un escalar λ y un vector $\vec{v} \neq \vec{0}$ tal que para una TL $T\colon V \to V$ se tenga $T(\vec{v}) = \lambda\vec{v}$, entonces al escalar se le llama* **eigen-*valor*** *y al vector se le llama* **eigen-*vector*.** *La palabra alemana* eigen *(léase aiguen) significa propio. El vector cero se excluye, pues siempre se tiene que $T(\vec{0}) = \vec{0}$ para toda TL T. Sin embargo, recordemos que el número cero sí puede ser un valor propio que corresponde a un apachurramiento.*

560. ***Ejemplo*** *Sea $T(x,y) = (x+3y, 2x+6y)$, entonces*
$T(1,2) = (1+6, 2+12) = (7,14) = 7(1,2)$.

Por lo tanto, 7 es un eigen-*valor de T y el correspondiente* eigen-*vector es* $(1,2)$. *También decimos que* $(7,(1,2))$ *es un* ***par propio***.

561. **Ejercicio** *Demuestre que si $\vec{v}$ es un vector propio de una TL T, entonces $\gamma\vec{v}$, con $\gamma \in \mathbb{R}$ y $\gamma \neq 0$, también lo es. Encuentre el correspondiente* eigen-*valor. Demuestre que si $(\lambda, \vec{v})$ y $(\mu, \vec{w})$ son pares propios, entonces $\vec{v} + \vec{w}$ no es un vector propio ni de λ ni de μ a menos que $\lambda = \mu$. Este resultado nos permite hablar de* **eigen-*espacio***, *el cual es un espacio generado por un conjunto de* eigen-*vectores asociados al mismo* eigen-*valor.*

562. ♣ ***Definiciones.*** *Una TL T es* ***diagonal*** *o es una* ***magnificación*** *cuando existe una base B donde T_B^B sea una matriz diagonal. En otras palabras, una TL T es diagonal ssi existe una base $B = \{\vec{v}_1, ..., \vec{v}_n\}$tal que $T(\vec{v}_i) = \lambda_i \vec{v}_i$ para ciertos escalares $\{\lambda_i\}$.*

563. ♣ ***Definiciones.*** *Dos matrices, M y N, son* ***similares*** *si ambas representan la misma TL pero con respecto a, quizá, bases diferentes A y B: $M = T_B^B$ y $N = T_A^A$. Una de las bases es usualmente la natural porque en esa base generalmente se definen las TL: $T(x, y, ..., z) = ...$ Por ejemplo, las dos TL del ejemplo siguiente son similares.*

564. **Ejemplo** *Sea B la base natural rotada $\pi/4$ al contrario de las manecillas del reloj. La matriz diagonal*

$$T_B^B = \begin{pmatrix} 1/9 & 0 \\ 0 & 1/4 \end{pmatrix}$$

representa una TL con respecto a B. Por supuesto, ver ejemplo 552, dicha transformación lineal es una magnificación. Si reescribimos dicha TL con respecto a la base natural obtenemos:

$$T_N^N = T = \begin{pmatrix} 13/72 & -5/72 \\ -5/72 & 13/72 \end{pmatrix}$$

por lo tanto, la definición de T es
$T(x, y) = (13x/72 - 5y/72, -5x/72 + 13y/72)$.

T ha sido representada en dos bases diferentes que codifican la misma información en forma de matrices similares.

565. ◇ ***Teorema.*** *Si dos matrices M y N son similares, entonces existe una matriz invertible K tal que $M = K^{-1}NK$.*

566. **Ejercicio** *Pruebe el anterior teorema y especifique la matriz K.*

567. ◇ ***Teorema.*** *Dos matrices similares tienen el mismo determinante.*

Demostración. Si dos matrices, M y N, son similares, entonces ellas representan la misma TL con respecto a bases apropiadas: $M = T_B^B$ y $N = T_C^C$, por lo tanto, existe una matriz de cambio de base K que relaciona a M y a N, de tal forma que $M = K^{-1}NK$. Tenemos:

$$\begin{aligned} Det(M) &= Det(K^{-1}NK) = Det(K^{-1})Det(N)Det(K) \\ &= (1/Det(K))Det(N)Det(K) = Det(N). \ \blacksquare \end{aligned}$$

568. $\boxed{\textbf{\textit{Ejercicio}}}$ *Sabemos que las bases no juegan papel alguno en la definición de un determinante de una* TL*, es decir, el determinante de una* TL *es algo asociado a la* TL *independiente de cualquier base. Formule la definición de determinante de una* TL *de tal forma que lo dicho sea obvio. Ayuda: el determinante debe ser función de un espacio de* TL *en* $\mathbb{R}$ *y reúne ciertas propiedades.*

569. ***Ejemplo*** *Sea la matriz* M *similar a la matriz diagonal* $D = (1, 2, 0)$*, entonces* M *es no invertible, pues su determinante es cero, por lo que* M *apachurra contra cero alguna parte de su dominio. Es claro que el Kernel tiene dimensión uno y que la imagen tiene dimensión dos.*

570. $\boxed{\textbf{\textit{Ejercicio}}}$ *Cierta matriz* M *es similar a una matriz diagonal. Por favor, diga cuándo la matriz original es invertible, la dimensión del Kernel y de la Imagen, si:*

a) $D(1, 0, 1)$

b) $D(-1, -1, -1)$

c) $D(1, 0, 0)$

d) $D(1, 0, 3, 0)$

e) $D(2, 3, 0, -1, 0)$.

571. $\boxed{\textbf{\textit{Ejercicio}}}$ *Para calcular el Kernel y la Imagen de una* TL *dada, hemos apelado en capítulos anteriores a la base natural. Sin embargo, hemos usado en el ejercicio anterior la creencia de que dichos espacios pueden calcularse en cualquier representación, en particular en una base propia, cuando existe, porque las respuestas son inmediatas. ¿Es eso correcto? En ese caso, pruébelo, o bien dé un contraejemplo.*

11.3. El *eigen*-problema

El ***eigen*-problema** es encontrar, si es posible, una representación diagonal de una TL T. O mejor dicho, es dilucidar si una TL es una magnificación.

572. ***Ejemplo*** *Sea* $T(x, y) = (x + 3y, 2x + 6y)$*. Debemos resolver* $T(x, y) = \lambda(x, y) = (\lambda x, \lambda y)$*. Esto es equivalente al sistema de ecuaciones:*

$x + 3y = \lambda x$
$2x + 6y = \lambda y$

Multiplicando la primera ecuación por dos y usando transitividad, obtenemos $2\lambda x = \lambda y$, la que es una ecuación con tres incógnitas. Una solución es $\lambda = 1$, con $x = 1$, $y = 2$. Otra solución es $\lambda = 0$. En ese caso, $x + 3y = 0$, por tanto, podemos tomar $x = -3$, $y = 1$. En concreto, el problema es resuelto por los pares propios $(1, (1, 2))$ y $(0, (-3, 1))$. En un momento aprenderemos a probar que no existe otro λ que resuelva el problema, es decir, que cualquier otro valor propio es un múltiplo de alguno de los dos ya hallados. Observemos que el conjunto $\{(1,2),(-3,1)\}$ es una base de $\mathbb{R}^2$y por tanto T puede escribirse con respecto a B como $T_B^B = D(1,0)$. De esa forma, T es una magnificación o una TL diagonal.

573. $\boxed{\textbf{\textit{Ejercicio}}}$ *Resuelva el problema propio para*
$T(x,y) = (x+3y, 3x+9y)$.

574. ***Contraejemplo*** *El problema de diagonalización no siempre puede ser resuelto. Como hemos visto, la esencia de una matriz diagonal es su carácter de magnificación. Consideremos ahora una rotación. Ella es en verdad una TL, pero ninguna dirección es amplificada pues todo es rotado. Trabajemos con una rotación concreta, digamos de ángulo π. La matriz correspondiente es*

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Para diagonalizar esa matriz, debemos resolver el problema propio

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}$$

eso es equivalente a

$-y = \lambda x$
$x = \lambda y.$

Multiplicando la primera ecuación por λ, la segunda por -1 e igualando, obtenemos:
$-\lambda y = \lambda^2 x = -x$
$\lambda^2 x + x = 0$
Factorizando
$(\lambda^2+1)x = 0$
Por lo que λ debe cumplir $(\lambda^2+1) = 0$, cuya solución es $\lambda = \sqrt{-1} = i$ que no está en los reales. Por lo tanto, la matriz dada no es diagonalizable sobre los reales.

575. $\boxed{\textbf{\textit{Ejercicio}}}$ *Dé un ejemplo de una rotación con al menos un valor propio real. Caracterice todas las rotaciones con valores propios reales. Pruebe que en $\mathbb{R}^2$ casi todas las rotaciones carecen de valores propios reales. Pruebe que en $\mathbb{R}^3$ todas las rotaciones tienen al menos un valor propio real.*

576. $\boxed{\textbf{\textit{Ejercicio}}}$ *Construya una TL $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ tal que sea una magnificación en la dirección Z pero una rotación en el plano XY. Prediga cuál será el resultado del* eigen-*problema para dicha TL T.*

11.4. Diagramas conmutativos

Esta sección es autocontenida, puede servir de repaso y es especial para aquellos que necesitan un camino bien cortico al tema de la diagonalización.

577. $\diamond$ ***Teorema.*** *Una TL es diagonal cuando existe una base $B = \{\vec{b}_1, ..., \vec{b}_n\}$ tal que todos sus elementos son vectores propios. Tal base, cuando existe, se llama **base propia**. La matriz F_B^B de una TL diagonal F con respecto a una base propia B es la matriz diagonal compuesta por los valores propios, listados en correspondencia con los vectores de la base.*

Demostración. Sea $F\colon \mathbb{R}^n \to \mathbb{R}^n$ una TL tal que existe una base B con n vectores $\vec{b}_1, ..., \vec{b}_n$ y con $F(\vec{b}_i) = \lambda_i \vec{b}_i$ para escalares apropiados λ_i. Recordemos que nuestras bases siempre son ordenadas. Todos los $\vec{b}_i$ son diferentes de $\vec{0}$, pero algunos λ_i pueden ser cero.

Veamos cómo se escribe la matriz de F con respecto a una base propia B: cada elemento de la base $\vec{b}_i$ es pasado a través de F y el resultado se descompone en la base B, las coordenadas correspondientes se escriben en forma vertical en una matriz, la cual es la matriz de F con respecto a B. Tenemos que $F(b_i) = \lambda_i \vec{b}_i$, cuya descomposición en la base propia B produce $\lambda_i \vec{b}_i = 0\vec{b}_1 + ... + \lambda_i \vec{b}_i + ... + 0\vec{b}_n$ y, por consiguiente, las coordenadas de $F(b_i)$ en la base B son $(F(b_i))_B = (0, ..., \lambda_i, ..., 0)$. Esas coordenadas se escriben en forma vertical como la columna i. Vemos que terminamos con una matriz diagonal:

$$F_B^B = \begin{pmatrix} \lambda_1 & 0 & ... & ... & 0 \\ 0 & \lambda_2 & 0 & ... & 0 \\ \vdots & ... & \vdots & ... & \vdots \\ 0 & ... & ... & ... & \lambda_n \end{pmatrix}$$

En una palabra, F_B^B es una matriz diagonal compuesta de los valores propios, en el mismo orden en el cual la base fue numerada. ∎

Si escribimos la matriz F con respecto a la base natural, tenemos la matriz ordinaria de F, también notada F_N^N. ¿Cómo se relacionan esas dos matrices F_B^B y F_N^N? Hay una manera muy natural de entender la relación:

578. $\diamond$ ***Teorema.*** *Sea $F\colon \mathbb{R}^n \to \mathbb{R}^n$ una TL que admite una base propia B, y sea N la base natural. Obtenemos la siguiente relación:*

$$F_N^N = I_N^B F_B^B I_B^N$$

Demostración. Veamos la naturalidad de esa identidad.

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "F_B^B"] & \mathbb{R}^n \arrow[d, "I_N^B"] \\
\mathbb{R}^n \arrow[u, "I_B^N"] \arrow[r, "F_N^N"] & \mathbb{R}^n
\end{tikzcd}$$

Figura 11.2. Diagrama conmutativo.

En el diagrama adjunto, tenemos dos caminos: el primero, en la parte inferior, llega en un solo paso, mientras que el segundo da un viaje de tres pasos. Las dos rutas son equivalentes. Decimos que el **diagrama es conmutativo**. Cada parte de cada ruta tiene su significado: I_B^N es la matriz de cambio de coordenadas desde la base natural a la base propia. F_B^B acepta coordenadas de $\vec{X}$ en la base propia y produce las coordenadas de $F(\vec{X})$ en la base propia. I_N^B cambia de coordenadas de la base propia a la base natural. Encadenando todo obtenemos: las coordenadas son recibidas en la base natural, se cambian a la base propia, son procesadas por F en la **base propia,** el resultado sale en la base propia y se cambia de la base propia a la base natural. Por lo tanto, la maquinaria acepta, en su totalidad, las coordenadas de $\vec{X}$ en la base natural y produce las coordenadas de $F(\vec{X})$ en la misma base. ■

579. **_Ejercicio_** *Con la misma notación del último teorema, dibuje el diagrama de composiciones de las TL adecuadas para que el resultado siguiente resulte evidente:*

$$F_B^B = I_B^N F_N^N I_N^B.$$

11.5. Uso del determinante

Con los ejercicios, por demás simples, que hemos visto, queda claro que el problema propio puede dar lugar a tremendos problemas algebraicos. Afortunadamente, contamos con una muy buena manera de usar el determinante para facilitarnos la vida. Veámosla.

El problema propio $F(\vec{v}) = \lambda\vec{v}$ es equivalente a $F(\vec{v}) - \lambda\vec{v} = \vec{0}$. Factoricemos $\vec{v}$. Para hacerlo, no podemos proceder como sigue: $(F - \lambda)\vec{v} = \vec{0}$ porque un número no puede ser restado de una TL. Para obtener matriz menos matriz procedemos como se debe:

$(F - \lambda\mathbb{I})\vec{v} = \vec{0}$

donde $\mathbb{I}$ es la identidad. Ahora, esta ecuación siempre tiene la solución trivial $\vec{v} = \vec{0}$. Por lo que a nosotros lo que nos interesa realmente es encontrar soluciones no triviales. Si llamamos $M = F - \lambda\mathbb{I}$, estamos buscando soluciones no triviales a $M\vec{v} = \vec{0}$. Pero eso implicaría que buscamos un vector que sea apachurrado contra cero. Por lo tanto, $Det(M) = 0$, y esa ecuación da un polinomio en términos de λ, cuyas raíces son los valores propios de F.

Podemos resumir:

580. ◇ ***Teorema.*** *Una condición necesaria para que una TL pueda ser diagonalizada es que $Det(F - \lambda\mathbb{I}) = 0$ para algún número λ.*

581. *Ejemplo* *Diagonalicemos la matriz*

$$F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Solución: Esta matriz define una TL que notamos F. Queremos encontrar vectores no nulos que satisfagan la ecuación $F(\vec{v}) = \lambda\vec{v}$ que es lo mismo que encontrar vectores no nulos que satisfagan $F(\vec{v} - \lambda\mathbb{I})\vec{v} = \vec{0}$, por lo que $Det(F(\vec{v}) - \lambda\mathbb{I}) = 0$.

$$Det\begin{pmatrix} 0-\lambda & 1 \\ 1 & 0-\lambda \end{pmatrix} = (-\lambda)^2 - 1 = \lambda^2 - 1 = 0,$$

lo cual implica que hay dos soluciones para λ, es decir, dos valores propios 1 y -1. Para encontrar los correspondientes vectores propios, procedemos primero con $\lambda = 1$:

Necesitamos encontrar vectores $\vec{v}$ tales que $(F - \lambda\mathbb{I})\vec{v} = (F - \mathbb{I})\vec{v} = \vec{0}$. Eso produce

$$\begin{pmatrix} 0-1 & 1 \\ 1 & 0-1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Notemos que la primera ecuación se transforma en la segunda si se multiplica por -1. Por consiguiente, cuando reduzcamos la matriz por Gauss-Jordan, la segunda ecuación se desaparece, pues no da información adicional, en perfecto acuerdo con lo que se espera de una matriz con determinante igual a cero.

Ahora, tenemos una sola ecuación $-v_1 + v_2 = 0$. Como hay dos incógnitas, hay un grado de libertad, lo que significa que uno puede elegir a su antojo el valor de una de las coordenadas del vector propio. O lo que es lo mismo, un vector propio que es multiplicado por un escalar también da un vector propio asociado al mismo valor propio.

Tomando la ecuación $-v_1 + v_2 = 0$, que es lo mismo que $v_1 = v_2$, dando al v_1 el valor de 1 obtenemos que v_2 toma el mismo valor. Por lo tanto, un vector asociado al valor propio 1 es $(1, 1)$.

Ahora analizamos el valor propio -1:

$$\begin{pmatrix} 0-(-1) & 1 \\ 1 & 0-(-1) \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Olvidamos la segunda ecuación y en la primera, $v_1 + v_2 = 0$, ponemos $v_1 = 1$ por lo que $v_2 = -1$, y obtenemos el vector propio $(1, -1)$. Notemos que el segundo vector es perpendicular al primero: esto es lo que pasa cuando empezamos con una matriz simétrica de coeficientes reales (probaremos esto más adelante). Eso implica que encontrar el primer vector propio de una matriz simétrica 2×2 determina inmediatamente el segundo por ortogonalidad.

Ahora manufacturamos una base propia con determinante igual a $+1$. Eso se hace dividiendo cada vector propio por su norma. La norma de $(1, 1)$, usando Pitágoras, es $\sqrt{2}$, la cual es la misma que de $(1, -1)$.

Por tanto, la base propia es $B = \{(1/\sqrt{2}, 1/\sqrt{2}), (1/\sqrt{2}, -1/\sqrt{2})\}$. Cuando se escribe eso, ya se está teniendo en cuenta que la base propia ha sido numerada:

$(1/\sqrt{2}, 1/\sqrt{2})$ es el primer vector y que $(1/\sqrt{2}, -1/\sqrt{2})$ es el segundo, pero el determinante de la base es

$$Det \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} = -1$$

Para lograr que el determinante quede positivo, numeramos la base en el orden inverso. La base es entonces

$B = \{(1/\sqrt{2}, -1/\sqrt{2}), (1/\sqrt{2}, 1/\sqrt{2})\}$.

Por tanto, la matriz de cambio de coordenadas I_N^B es simplemente la matriz formada por la base en notación vertical

$I_N^B = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$

Recordemos que $F_N^N = I_N^B F_B^B I_B^N$, donde la matriz I_B^N es inversa de I_N^B. Puede verificarse en el caso presente que la inversa es la transpuesta. Eso es válido porque la base B tiene vectores mutuamente perpendiculares, cuya norma es uno. Por tanto:

$I_B^N = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$

Y la representación diagonal de F es

$F_B^B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

582. **Ejercicio** *Con las matrices encontradas en el ejercicio previo, verificar que* $F_N^N = I_N^B F_B^B I_B^N$.

583. ♣ ***Definiciones.*** *Dada una matriz M, el polinomio que resulta de evaluar $Det(M - \lambda\mathbb{I}) = 0$ se denomina el* ***polinomio característico de** M. La* ***multiplicidad algebraica*** *de un valor propio es igual al número de veces que el valor propio se repite como raíz del polinomio característico de la matriz. Por ejemplo, si el polinomio característico es $p(\lambda) = (\lambda - 3)^2(\lambda - 4) = 0$ entonces 3 tiene multiplicada algebraica 2, mientras que 4 tiene multiplicada algebraica 1.*

584. **Ejemplo** *La siguiente matriz no es diagonalizable en números reales:*

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

En efecto, el polinomio asociado (el **polinomio característico**) es

$$Det \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 + 1 = 0$$

cuyas raíces son i, $-i$. Esto era de esperar puesto que la matriz representa una rotación de 90 grados, la cual transforma $(0,1)$ en $(1,0)$ y $(0,1)$ en $(-1,0)$. Es claro que una rotación no multiplica ningún vector por un escalar. Por lo que al hacer el proceso de diagonalización, lo que uno encuentra es un polinomio que no tiene raíces reales. Sin embargo, recordemos que los números complejos están asociados a ondas y por lo tanto no ha de ser sorpresa que matrices como la presentada en este ejemplo aparezcan en sistemas oscilantes (Zill, 2008).

585. $\boxed{\textbf{\textit{Ejercicio}}}$ *Las siguientes matrices diagonales representan magnificaciones en una base dada. Encontrar su representación en la base natural y resolver el correspondiente problema propio para dichas matrices, encontrando los valores propios, los vectores propios y revertiendo la matriz diagonal original:*

1. $D(1,-2)$, $B = \{(1,1), (-1,1)\}$
2. $D(0,0)$, $B = \{\sqrt{2}/2(1,1), \sqrt{2}/2(-1,1)\}$
3. $D(1,0)$, $B = \{(1,1), (-1,2)\}$
4. $D(2,-1)$, $B = \{(1,0), (1,1)\}$
5. $D(-2,1)$, $B = \{\sqrt{10}/10(1,3), \sqrt{10}/10(3,-1)\}$.

586. $\boxed{\textbf{\textit{Ejercicio}}}$ *Encuentre todos los valores propios de las siguientes matrices y escriba la forma diagonal asociada, si eso es posible. Matrices de este tipo son necesarias para modelar el fenómeno de* ***resonancia*** *cuando un modo natural de oscilación de un sistema es estimulado desde el exterior, como cuando una copa vibra inducida por el sonido de una guitarra (Kittel y otros, 1965).*

a) $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$

b) $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$

c) $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

d) $\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

e) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

587. $\boxed{\textbf{\textit{Ejercicio}}}$ *Analice el siguiente razonamiento: Hemos mostrado, como se esperaba, que una rotación de ángulo $\pi/2$ no tiene valores propios reales. Sin embargo, una magnificación por el número complejo $i = \sqrt{-1}$ puede causar una rotación. Por lo tanto, una rotación no es diagonalizable en reales pero sí en complejos. En general, la solución al problema propio siempre comienza con la búsqueda de las raíces de un polinomio y un polinomio de grado n siempre tiene n raíces, que pueden ser complejas. Hemos probado que toda matriz es diagonalizable en complejos.*

588. ***Raíces de polinomios.*** *Aunque hemos venido tratando con matrices* 2×2*, nuestra teoría es aplicable a cualquier dimensión. El problema es la complejidad de la aritmética que resulta. Por tal razón, uno puede optar por soluciones asistidas por computador. Hay, con todo, un caso en el cual la aritmética es manejable a mano y es el caso de dimensión* 3*. Después de encontrar el polinomio* $Det(M - \lambda \mathbb{I}) = 0$*, debemos encontrar las raíces. En tres dimensiones, el polinomio característico es cúbico y puede verse que siempre tiene al menos una raíz real. ¿Cómo podemos encontrarla? Podemos explorar la recta real para ver si de casualidad encontramos una raíz. Podemos mejorar si recordamos que un polinomio es una función continua. Por lo que si encontramos un valor en el cual el polinomio es positivo y otro valor en el cual el polinomio es negativo, podemos estar seguros de que existe una raíz entre los dos valores: esa es una aplicación del teorema del valor intermedio para funciones continuas (método de bisección de Newton).*

Existe la posibilidad de que el polinomio tenga raíces racionales. Entonces, podemos aplicar el siguiente Teorema. si un polinomio $a_n x^n + ... + a_1 x + a_o = 0$ con coeficientes enteros tiene una raíz racional, r, entonces, ella es de la forma

r=(Divisores de a_o)/(Divisores de a_n).

589. ***Ejemplo*** *La ecuación* $4x + 3 = 0$ *tiene como posibles raíces racionales a los quebrados de la forma (divisores de* 3*)/(divisores de* 4*). De hecho, la única raíz es* $-3/4$*. Use este ejemplo como ayuda nemotécnica para recordar qué es divisor de qué.*

590. ***Ejemplo*** *Una matriz tiene como polinomio característico a:*
$x^3 + 2x^2 - x - 2 = 0$*. Puesto que es cúbico, tiene por lo menos una raíz real. Los posibles candidatos son (divisores de* 2*)/ (divisores de* 1*). Una lista completa es:* $1, -1, 2, -2$*. De hecho, una sustitución directa da que las raíces son* $1, -1, -2$*.*

591. ***Ejemplo*** *El polinomio característico de una matriz es:* $x^3 - x^2 - 2x + 2 = 0$*. Las posibles raíces racionales son de la forma (divisores de* 2*)/(divisores de* 1*), que son* $1, -1, 2, -2$*. Después de chequear, la única raíz que sirvió es* $x = 1$*. Por tanto,* $x - 1$ *es un factor. Podemos dividir el polinomio* $x^3 - x^2 - 2x + 2 = 0$ *por* $x - 1$ *para encontrar el cociente* $x^2 - 2$*, cuyas raíces son:* $\sqrt{2}, -\sqrt{2}$*.*

592. ***Ejemplo y definición*** *Para la ecuación* $(x-1)^3 = x^3 - 3x^2 + 3x - 1 = 0$*, podemos comenzar con las posibles raíces racionales* $+1, -1$*. Un chequeo dice que* 1 *es raíz, pero que* -1 *no lo es. En dónde están las otras dos raíces? Dividimos* $x^3 + 3x^2 + 3x + 1$ *por* $x - 1$ *para obtener* $(x-1)^2 = x^2 + 2x + 1$*. Este polinomio tiene como posibles raíces a* 1*,* -1*. Un chequeo directo dice que* 1 *es raíz pero que* -1 *no lo es. Dividimos por* $(x - 1)$ *para obtener* $x - 1$ *como cociente, cuya única raíz es* 1*. Por lo que el polinomio tiene* 3 *raíces pero todas iguales. En nuestro ejemplo,* 1 *es una raíz de multiplicidad algebraica* 3*.*

593. ***Ejemplo*** *Para el polinomio* $x^2 - 3x/2 + 1/2 = 0$*, debemos multiplicar todo por* 2 *para obtener* $2x^2 - 3x + 1 = 0$ *cuyas posibles raíces racionales son* $1, -1, 1/2, -1/2$*. Un chequeo dice que* -1 *y* $-1/2$ *no son raíces, pero que* 1 *y* $1/2$ *sí lo son.*

594. $\boxed{\textbf{\textit{Ejercicio}}}$ *Encuentre las raíces de los siguientes polinomios:*

a) $x^3 - 5x^2 + 7x - 3 = 0$

b) $x^3 - 6x^2 + 11x - 6 = 0$

c) $x^3 + x^2 - 2x - 2 = 0$

d) $12x^3 - 31x^2 + 15x - 2 = 0$

e) $x^3 - 31x^2/12 + 5x/4 - 1/6 = 0.$

595. $\boxed{\textbf{\textit{Ejercicio}}}$ *Resuelva el problema propio para las siguientes matrices y verifique que conmutan entre ellas.*

a) $\begin{pmatrix} 6 & -2 & 0 \\ -3 & 6 & -3 \\ 0 & -1 & 6 \end{pmatrix}$

b) $\begin{pmatrix} 4 & -4 & -2 \\ -6 & 3 & -6 \\ -1 & -2 & 5 \end{pmatrix}$

c) $\begin{pmatrix} 0 & -4 & 6 \\ -6 & 3 & -6 \\ 3 & -2 & -3 \end{pmatrix}$

d) $\begin{pmatrix} 2/3 & -2/3 & 8/3 \\ -1 & 2 & -1 \\ 4/3 & -1/3 & -2/3 \end{pmatrix}$

e) $\begin{pmatrix} -2/3 & 2/3 & 4/3 \\ 1 & 0 & 1 \\ 2/3 & 1/3 & -4/3 \end{pmatrix}$

596. $\boxed{\textbf{\textit{Ejercicio}}}$ *Pruebe que si dos matrices tienen la misma base propia, entonces necesariamente conmutan entre ellas.*

597. $\boxed{\textbf{\textit{Ejercicio}}}$ *Sea*

$$Q = \begin{pmatrix} -1/\sqrt{2} & -1/(3\sqrt{2}) & 2/3 \\ 1/\sqrt{2} & -1/(3\sqrt{2}) & 2/3 \\ 0 & 4/(3\sqrt{2}) & 1/3 \end{pmatrix}$$

a) Pruebe que Q es la matriz de alguna rotación.

b) Encuentre el eje de rotación y el ángulo correspondiente.

c) Bosqueje la gráfica de la cuádrica $x^2 + 4y^2 + 9z^2 = 1$.

d) Encuentre la ecuación de la figura que es el resultado de rotar por Q a $x^2 + 4y^2 + 9z^2 = 1$ y bosquéjela.

11.6. Matrices simétricas

Recordemos que: dada una matriz cuadrada M, su transpuesta, M^T, es el resultado de reflejar la matriz M con respecto a la diagonal, es decir, si $M = (a_{ij})$, $M^T = (a_{ji})$. Una matriz que es igual a su transpuesta se llama **simétrica.** El determinante de una matriz es igual al de su transpuesta. Para toda matriz M, tenemos que $M(\vec{u}) \cdot \vec{v} = \vec{u} \cdot M^T(\vec{v})$. Por lo tanto, toda matriz simétrica S satisface la identidad:

$$S(\vec{u}) \cdot \vec{v} = \vec{u} \cdot S(\vec{v})$$

El problema propio para matrices simétricas es muy agradable y ofrece la posibilidad de retroalimentarse para saber si uno va resolviendo un problema bien o mal. Eso es muy importante, pues los cálculos asociados al proceso de diagonalización son muy tediosos y por lo tanto muy propensos a generar errores. Por ello es mejor tener métodos de retroalimentación, como los presentados por los siguientes teoremas:

598. ◇ ***Teorema.*** *Los vectores propios de una matriz simétrica real son ortogonales entre ellos, si los valores propios correspondientes son diferentes.*

Demostración. Sea S una matriz simétrica tal que $S(\vec{u}) = \lambda\vec{u}$ y $S(\vec{v}) = \gamma\vec{v}$, y tomamos $\lambda \neq \gamma$, por lo tanto

$S(\vec{u}) \cdot \vec{v} = \vec{u} \cdot S(\vec{v})$

pero el lado izquierdo es igual a $\lambda\vec{u} \cdot \vec{v}$ y el lado derecho es igual a $\vec{u} \cdot \gamma\vec{v}$. Por consiguiente

$\lambda\vec{u} \cdot \vec{v} = \vec{u} \cdot \gamma\vec{v} = \gamma\vec{u} \cdot \vec{v}$

Puesto que $\lambda \neq \gamma$, debemos concluir que $\vec{u} \cdot \vec{v} = 0$ y que a valores propios diferentes corresponden vectores propios ortogonales. ■

599. ◇ ***Teorema.*** *Los valores propios de una matriz simétrica real son todos reales.*

Prueba en dos pasos. **Paso 1**. Es posible definir un producto punto, interno o interior para vectores con coordenadas complejas.

Demostración. Sea S una matriz simétrica y sea $\vec{v}$ un vector propio con λ el correspondiente valor propio. Debemos probar que se trata de un número real. ¿Cómo lo haremos? Es un placer explicitar que toda nuestra teoría ha sido construida para números reales y que, por tanto, debemos reconstruir todo de nuevo para tener derecho de hablar de valores propios complejos. Necesitamos extender nuestra teoría a números complejos y trabajar en ese mundo más grande. Ese trabajo monumental ya ha sido hecho:

Hemos definido el producto punto entre vectores con componentes reales como sigue: si $\vec{u} = (u_1, ..., u_n)$ y $\vec{v} = (v_1, ..., v_n)$ entonces

$\vec{u} \cdot \vec{v} = u_1 v_1 + ... + u_n v_n$

Queremos definir el producto punto entre vectores con coordenadas complejas. La experiencia ha demostrado que tal definición debe llenar un requisito muy simple para que sea una extensión útil: se necesita que $\|\vec{u}\|^2 = \vec{u} \cdot \vec{u} > 0$, siempre que $\vec{u} \neq 0$. Eso se logra si se define el producto punto entre vectores con coordenadas complejas como sigue. Si $\vec{Z} = (z_1, ..., z_n)$y $\vec{W} = (w_1, ..., w_n)$ entonces

$\vec{Z} \cdot \vec{W} = z_1 w_1^* + ... + z_n w_n^*$

donde el símbolo w^* denota el complejo conjugado de $a + bi$ que es $a - bi$.

Por ejemplo, si $z = 1 + i$, entonces $\|z\|^2 = (1 + i)(1 - i) = 1 - i^2 = 1 + 1 = 2 > 0$.

Similarmente, si $\vec{Z} = (1 + i, 2 + 4i)$ y $\vec{W} = (1 - 3i, 3 - 4i)$, tenemos que

$\vec{Z} \cdot \vec{W} = (1 + i)(1 - 3i)^* + (2 + 4i)(3 - 4i)^*$

$= (1 + i)(1 + 3i) + (2 + 4i)(3 + 4i) = -12 + 24i$.

de otra parte,

$\|\vec{Z}\|^2 = \vec{Z} \cdot \vec{Z} = (1 + i, 2 + 4i) \cdot (1 + i, 2 + 4i)^*$

$(1 + i, 2 + 4i) \cdot (1 - i, 2 - 4i) = (1 - i^2) + (4 - 16i^2) = 1 + 1 + 4 + 16 = 22 > 0$. ■

600. $\boxed{\textbf{\textit{Ejercicio}}}$ *Pruebe que el nuevo producto punto cumple con las dos siguientes propiedades:*

a) $(\lambda\vec{Z}) \cdot \vec{W} = \lambda(\vec{Z} \cdot \vec{W})$
b) $\vec{Z} \cdot \lambda\vec{W} = \lambda^*(\vec{Z} \cdot \vec{W})$

Continuamos con la prueba del teorema 599. **Paso dos.** Si λ es valor propio de una matriz simétrica real, entonces es igual a su conjugado: $\lambda = \lambda^*$.

Demostración. Supongamos ahora que la matriz simétrica S tiene un valor propio λ, i.e., $S(\vec{Z}) = \lambda\vec{Z}$ para cierto vector $\vec{Z}$. Aún no sabemos si λ es real o no. Un número real es un número complejo que es igual a su conjugado, pues no tiene parte imaginaria.

Puesto que S es simétrica, ella cumple:

$S\vec{Z} \cdot \vec{Z} = \vec{Z} \cdot S\vec{Z}$

Puesto que $\vec{Z}$ es un vector propio de S, tenemos que

$(\lambda\vec{Z}) \cdot \vec{Z} = \vec{Z} \cdot (\lambda\vec{Z})$

Aplicando esas dos propiedades al ejercicio precedente obtenemos:

$\lambda(\vec{Z} \cdot \vec{Z}) = \lambda^*(\vec{Z} \cdot \vec{Z})$

Ahora tenemos dos posibilidades: que $\lambda = 0$, en cuyo caso el correspondiente valor propio es real. O bien, $\lambda \neq 0$ y en ese caso $\vec{Z} \cdot \vec{Z} = \|\vec{Z}\|^2 > 0$, podemos dividir entonces

$\lambda(\vec{Z} \cdot \vec{Z}) = \lambda^*(\vec{Z} \cdot \vec{Z})$

por $\|\vec{Z}\|$ a ambos lados para obtener: $\lambda = \lambda^*$ lo cual significa que λ es un número real. ■

601. $\diamond$ ***Teorema de los ejes principales.*** *Toda matriz simétrica real (con entradas reales, sin imaginarios) S, es diagonalizable, lo cual significa que tiene una base propia con respecto a la cual se ve como diagonal, D. Esto es equivalente a decir: Sea S una matriz simétrica. Entonces existe una matriz ortogonal B que diagonaliza a S, es decir $S = B^T DB$. Cada elemento de la base genera un eje que se llama un eje principal y de ahí el nombre del teorema. En las aplicaciones, próximo capítulo, veremos la importancia de dichos ejes.*

Demostración: Investigación. Pruebe que uno puede formar una base de vectores propios B que expanden todo el espacio dominio de la TL codificada por S. En la base B, S se ve como diagonal, D. Al relacionar D y S se tiene que $S = BDB^T$.

602. $\boxed{\textit{\textbf{Ejercicio}}}$ *Diagonalice, si tan sólo es posible, las siguientes matrices. Ellas son 2×2, para practicar con matrices 3×3, pase al capítulo siguiente donde tendremos de sobra. Las matrices son:*

a) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

b) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$

c) $\begin{pmatrix} 1 & 2 \\ 2 & -3 \end{pmatrix}$

d) $\begin{pmatrix} -1 & 5 \\ 5 & -3 \end{pmatrix}$

e) $\begin{pmatrix} -1 & 6 \\ 6 & -1 \end{pmatrix}$

603. $\boxed{\textit{\textbf{Investigación}}}$ *Las rotaciones no producen magnificaciones (en números reales), por lo tanto no pueden ser similares a una matriz diagonal. Por ello, es posible que toda TL tenga en parte una rotación y en parte una magnificación. ¿Será eso verdad? Por favor, dé ejemplos, contraejemplos, pruebas y todo lo que haga falta para dilucidar esa intriga.*

11.7. Reconociendo las cónicas

Sabemos que la ecuación de una cónica que ha sido rotada incluye términos de la forma xy. Por ejemplo, la ecuación $13x^2 - 10xy + 13y^2 = 72$ describe una elipse rotada en sentido contrario a las manecillas del reloj un ángulo de $\pi/4$. Ahora enfrentaremos el problema inverso: determinar la naturaleza de una cónica junto con el ángulo de rotación de la figura cuya ecuación en la base natural es $ax^2 + by + cy^2 + dx + ey = f$. Para poder hacer eso haremos uso de todo el poder de la diagonalización.

604. ***Ejemplo*** *Identifiquemos la cónica y el ángulo de rotación si su ecuación en la base natural es* $13x^2 - 10xy + 13y^2 = 72$.

La ecuación puede reescribirse matricialmente como:

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 13 & -5 \\ -5 & 13 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 72$$

lo cual es de la forma $X^T M X = k$, donde X es un vector, M una matriz, k un escalar. Para identificar la cónica, debemos escribir la ecuación en otra base desde la cual ya no aparezca el término xy. Esto se logra si reescribimos la matriz asociada a la cónica en una matriz en la cual adquiera una forma diagonal. Una matriz diagonal representa una magnificación, i.e., una TL que deja fijas ciertas direcciones, en las cuales ella produce la multiplicación por un escalar. Por eso necesitamos resolver el problema propio $M\vec{v} = \lambda\vec{v}$ or $(M - \lambda\mathbb{I})\vec{v} = \vec{0}$ donde $\vec{0}$ representa el vector nulo. Puesto que $\vec{v} = \vec{0}$ es una solución y necesitamos una no trivial, exigimos que $(M - \lambda\mathbb{I})$ sea no inyectiva, o sea que apachurre ciertos vectores contra cero, por tanto, $Det(M - \lambda\mathbb{I})$ debe ser cero. Este es nuestro comienzo:

$$Det \begin{pmatrix} 13-\lambda & -5 \\ -5 & 13-\lambda \end{pmatrix} = 0$$

$(13-\lambda)^2 - 25 = (13 - \lambda - 5)(13 - \lambda + 5) = 0$
cuyas raíces son $\lambda_1 = 8$ y $\lambda_2 = 18$.
Si reemplazamos el valor $\lambda_1 = 8$ en la ecuación $(M - \lambda\mathbb{I})\vec{v} = 0$, obtenemos:

$$\begin{pmatrix} 13-8 & -5 \\ -5 & 13-8 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0$$

o, simplificando:

$$\begin{pmatrix} 5 & -5 \\ -5 & 5 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0$$

Observemos que el determinante de esta matriz es cero, como debe ser. Esto se debe a que la segunda ecuación es redundante y que nos queda una ecuación con dos incógnitas, por lo que una incógnita puede ser fijada a voluntad: sea $v_1 = 1$ por lo que $v_2 = 1$. Por lo tanto, el primer vector propio es $(1, 1)$. Reemplazando el segundo valor propio debe producir el segundo vector propio $(-1, 1)$, puesto que a valores propios diferentes corresponden vectores propios perpendiculares. ¿Es esto verdad? Veamos:

$$\begin{pmatrix} 13-18 & -5 \\ -5 & 13-18 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0$$

$$\begin{pmatrix} -5 & -5 \\ -5 & -5 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0$$

eso se reduce a la ecuación $-5v_1 - 5v_2 = 0$, la cual se satisface con $(-1, 1)$. Con estos dos vectores propios formamos una matriz de tal forma que su primer vector unitario quede en el primer cuadrante y el segundo en el segundo cuadrante. De esa manera el determinante queda positivo. Podemos formar pues la base ortonormal

$$B = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

Esa base es el resultado de rotar la base natural un ángulo de $\pi/4$. Esta matriz también representa la matriz de cambio de coordenadas I_N^B, cuya inversa es su transpuesta: $(I_N^B)^{-1} = (I_N^B)^T = I_B^N$.

De otro lado, cada elemento $\vec{b}_i$ de la base B obedece a la ecuación $M\vec{b}_i = \lambda\vec{b}_i$. Por lo tanto, la matriz adquiere la forma diagonal D en la base B:

$$D = M_B^B = \begin{pmatrix} 8 & 0 \\ 0 & 18 \end{pmatrix}$$

Las dos representaciones están ligadas por las identidades

$M_B^B = I_B^N M_N^N I_N^B$
$M_N^N = I_N^B M_B^B I_B^N$

Usando la segunda identidad, podemos reescribir la ecuación de la cónica $X^T MX = 72$ en la base B. De hecho, la ecuación original es exactamente:

$(X_N)^T M_N^N X_N = 72$.

Podemos hacer las sucesivas transformaciones

$(X_N)^T I_N^B M_B^B I_B^N X_N = 72$
$(X_N)^T (I_B^N)^T D(I_B^N X_N) = 72$
$(I_B^N X_N)^T D X_B = 72$
$X_B^T D(X_B) = 72$

Si a las coordenadas X_B de un punto X en la base B las llamamos (u_1, u_2), entonces obtenemos

$$\begin{pmatrix} u_1 & u_2 \end{pmatrix} \begin{pmatrix} 8 & 0 \\ 0 & 18 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 72, \text{ ó}$$

$8(u_1)^2 + 18(u_2)^2 = 72$
$(u_1)^2/9 + (u_2)^2/4 = 1$.

Conclusión: la ecuación en la base natural representa una elipse con ancho 3 en la dirección $(1, 1)$, y ancho 2 en la dirección $(-1, 1)$. El ángulo de rotación es $\pi/4$. Notemos la correspondencia natural: la primera coordenada u_1 fue dada por el primer vector de la base, el cual iba en la dirección $(1, 1)$ con valor propio 8, mientras que la segunda coordenada u_2 iba en la dirección $(-1, 1)$, cuyo valor propio es 18.

El mérito de una base ortonormal es que cuando es considerada como matriz, su inversa es su transpuesta. Sin embargo, uno también puede trabajar en una base ortogonal cuyos elementos tienen igual largo, y aun en el caso en el cual el primer vector no esté en el primer cuadrante.

605. ***Ejemplo*** *La cónica* $11x^2/10 + 19y^2/10 + 6xy/10 = 1$ *está representada por*

$$M = M_N^N = \begin{pmatrix} 11/10 & 3/10 \\ 3/10 & 19/10 \end{pmatrix}$$

El problema del valor propio produce $\lambda_1 = 1$, $\lambda_2 = 2$ con base propia B e inversa B^{-1} dadas, respectivamente, por

$$B = \begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix}$$

$$B^{-1} = \begin{pmatrix} 3/10 & 1/10 \\ -1/10 & 3/10 \end{pmatrix}$$

La matriz diagonal correspondiente es

$$D = M_B^B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

En la base B la cónica tiene la ecuación:
$u_1^2 + 2u_2^2 = 1$
$u_1^2 + u_2^2/(1/2) = 1$
$u_1^2 + u_2^2/(\sqrt{2}/2)^2 = 1$

Por lo tanto: La cónica es una elipse orientada en la dirección $(3,1)$ con ancho 1. En la dirección normal, $(-1,3)$, tiene ancho $\sqrt{2}/2$. Su ángulo de rotación es $Arctan(1/3)$.

606. $\boxed{\textbf{\textit{Ejercicio}}}$ *Pruebe que la base B dada por*

$$B = \begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix}$$

es una base propia de todas las cónicas dadas, encuentre los valores propios, la forma diagonal de las cónicas, la ecuación asociada, el ángulo de rotación y haga un dibujo (el papel de $\vec{i}$ es jugado ahora por $(3,1)$ y el de $\vec{j}$ por $(-1,3)$). Las cónicas están representadas por una ecuación cuyo lado derecho es uno, pero en caso de inconsistencia la ecuación debe igualarse a -1. En la base natural, las cónicas están representadas por las matrices:

a) $\begin{pmatrix} 21/10 & 3/10 \\ 3/10 & 29/10 \end{pmatrix}$

b) $\begin{pmatrix} -6/10 & 12/10 \\ 12/10 & 26/10 \end{pmatrix}$

c) $\begin{pmatrix} 8/10 & -6/10 \\ -6/10 & -8/10 \end{pmatrix}$

d) $\begin{pmatrix} -14/10 & -12/10 \\ -12/10 & -46/10 \end{pmatrix}$

e) $\begin{pmatrix} -12/10 & -6/10 \\ -6/10 & -28/10 \end{pmatrix}$

f) Tome varios pares (A, B) de esas matrices y realice los productos AB y BA. ¿Qué observa usted? ¿ Podría explicarlo?

607. $\boxed{\textbf{\textit{Ejercicio}}}$ *¿Es nuestro formalismo aplicable a parábolas? Dé ejemplos y contraejemplos. ¿Puede nuestro formalismo ser ligeramente modificado para ser aplicable a parábolas? Para entender la pregunta, tome la parábola $y = x^2$ en coordenadas cartesianas y rótela un ángulo de $\pi/4$ en sentido contrario a las manecillas del reloj. Después enfrente el problema propio con la ecuación resultante.*

11.8. Ejercicios de repaso

1. Si la matriz de la transformación T: $\mathbb{R}^2 \to \mathbb{R}^2$, en la base canónica de $\mathbb{R}^2$, es diagonalizable y $(1, 2)$, $(3, 1)$ son vectores propios de T y $T(5, -5) = (2, -1)$, halle los **autovalores** (sinónimo de valores propios y de eigen-valores) de T y la matriz de la transformación, en la base canónica de $\mathbb{R}^2$.

2. ¿Para qué valores de a la matriz $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & a & 0 \end{pmatrix}$ es diagonalizable?

3. Considere la matriz $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. ¿Es A diagonalizable? En caso afirmativo, hallar su forma diagonal D y obtener una matriz P invertible tal que $P^{-1}AP = D$.

4. Muestre que si $B^2 = \mathbb{I}$, entonces los valores propios de B son 1 y -1.

5. Encuentre una matriz $A_{2\times 2}$ simétrica con valores propios $\lambda_1 = -1$, $\lambda_2 = 2$ y vectores propios correspondientes $\vec{v_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ y $\vec{v_2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

6. Si $A_{n\times n}$ es diagonalizable y $\lambda_1, ..., \lambda_n$ son los valores propios de A entonces $Det(A) = \lambda_1\lambda_2...\lambda_n$.

11.9. Resumen

Una matriz diagonal es casi la más simple y además tiene una interpretación geométrica simple: es una magnificación anisotrópica (que no funciona igual en todas las direcciones). La libertad de expresar una TL en cualquier base se explota al máximo en la diagonalización, un programa que consiste en fabricar una base especial, una base propia, que permita una representación diagonal de dicha TL. Vimos que esto se usa para reconocer una cónica cuya naturaleza está escondida detrás de una rotación. Pero la diagonalización tiene muchas aplicaciones, algunas de las cuales las veremos en el próximo capítulo y otras se ven en un curso de ecuaciones diferenciales.

CAPÍTULO 12

APLICACIONES

Las aplicaciones de la diagonalización son ilimitadas y muchas veces sorpresivas. Comenzaremos con una aplicación a figuras tridimensionales, que es un tema con aplicación directa a la ingeniería, a la industria y que también tiene que ver con algunos problemas filosóficos acerca de la naturaleza del hombre. Después veremos cómo se usa la diagonalización para estudiar máximos y mínimos y para atacar un problema en mecánica relacionado con la rotación de cuerpos rígidos. Terminamos con una aplicación al estudio de los sistemas dinámicos.

12.1. Ajuste de patrones

Un **patrón** es un objeto que sirve de **modelo** o **referencia** para comparar objetos de un conjunto dado. Por ejemplo, un círculo en el plano es un conjunto de puntos que equidistan perfectamente de otro punto dado. Cuando nosotros vemos una figura, inmediatamente la comparamos con la batería de patrones que hay en la mente, círculos, cuadrados, triángulos y demás. No existe en la naturaleza elementos perfectamente circulares, pero nosotros decimos que algo es circular porque el patrón circular es el que mejor se ajusta a la figura. Eso quiere decir que hay una diferencia entre el patrón y el objeto a clasificar y que el patrón círculo es el que minimiza la diferencia.

Lo anterior implica que las diferencias se pueden cuantificar, medir de alguna manera. En algunos casos esto puede ser algo muy complicado. Por ejemplo, la ciencia no es más que un estudio del grado y calidad del ajuste entre las teorías, patrones preconcebidos, y los datos de la naturaleza. Pero en otros caso, el problema de ajuste puede ser tratable e incluso sencillo, pues se reduce a la solución de un sistema de ecuaciones lineales.

608. ***Ejemplo*** ***Línea de regresión de mínimos cuadrados.***

Investiguemos la sospecha de que el ingreso mensual tiene una dependencia lineal con respecto al tiempo de experiencia. Los datos los reportamos por pares de la forma (tiempo de experiencia en años, sueldo en unidades arbitrarias):
$(3,5),(2,5),(2,4),(4,9)(3,7),(5,11)$.

Solución: Nosotros dibujamos los puntos sobre un plano cartesiano, lo cual se denomina un **diagrama de dispersión**.

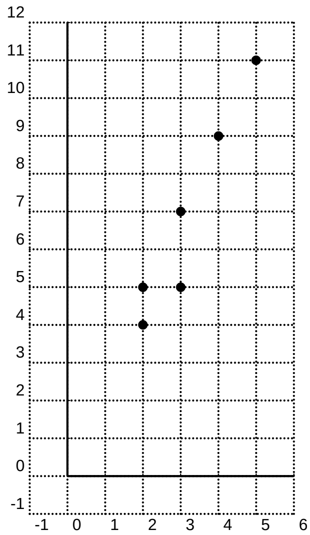


Figura 12.1. Los años de experiencia en el eje horizontal, el sueldo en el eje vertical.

Mirando el diagrama de dispersión, uno se forja una idea sobre aquella línea recta que mejor se ajusta a los datos, es decir, aquella que mejor parece representar los datos.

Cuando la línea viene de unos datos particulares, de una muestra, nosotros usamos para la línea la ecuación con letras latinas:

$y = a + bx$.

Pero cuando nosotros formulamos un modelo, un patrón, usamos letras griegas

$y = \alpha + \beta x$

En nuestro caso, la línea que mejor se ajusta a nuestros datos parece pasar por el punto más cercano al origen y por el más lejano. Esos puntos son: $(2, 4)$ y $(5, 11)$, los cuales nos generan una línea (en letras latinas):

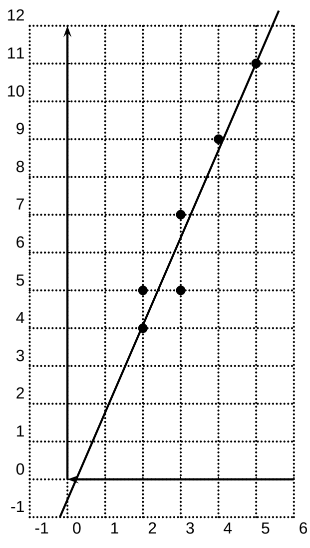


Figura 12.2. Estimación visual de la línea que mejor representa los datos, aquella que causa un mínimo de descontento global.

La pendiente de dicha línea es
$b = \dfrac{y_2 - y_1}{x_2 - x_1} = \dfrac{11-4}{5-2} = \dfrac{7}{3} = 2{,}33$
Por lo tanto, la ecuación de la línea es
$y = a + bx = a + 2{,}33x$
Para hallar a estudiamos un punto cualquiera, por ejemplo, $(2,4)$:
$4 = a + 2{,}33(2) = a + 4{,}66$
por lo que $a = 4 - 4{,}66 = -0{,}66$
Por consiguiente, nuestra **estimación visual de la línea de regresión** es
$y = -0{,}66 + 2{,}33x$.
Mirando la nueva gráfica, vemos que es razonable probar la creencia de que el ingreso mensual es proporcional a la experiencia, es decir, que se justifica un modelo lineal. Esta idea es importante, pues demuestra que al problema de ajuste de patrones comienza con ideas preconcebidas.

Para hallar exactamente la línea que mejor se ajusta a los datos es usual sacar una fórmula general y después aplicarla. Consideremos el problema de ajustar una línea de la forma $y = \alpha + \beta x$ a unos datos de la forma (x_i, y_i). Para un x_i dado el correspondiente y observado es y_i pero el modelo predice $\alpha + \beta y_i$. Hay una discrepancia, un error, entre lo predicho por el modelo y lo observado. Dicho error se mide por la expresión:

$$\epsilon^2(\alpha, \beta) = \sum_i (y_i - (\alpha + \beta y_i))^2 = \sum_i (y_i - \alpha - \beta y_i)^2$$

y se llama error cuadrático. El patrón buscado es aquella línea recta que minimiza el error cuadrático y se llama **línea de regresión de mínimos cuadrados**. Para minimizar una función como ésta, derivamos e igualamos a cero, pero como la función depende de dos variables, usamos derivadas parciales en las que cuando se deriva parcialmente con respecto a una variable, las demás se consideran constantes. Las derivadas parciales, digamos con respecto a α se notan como $\dfrac{\partial}{\partial \alpha}$. Precedemos:

$\dfrac{\partial \epsilon}{\partial \alpha} = 2 \sum_i (y_i - \alpha - \beta x_i)(-1) = 0$
Dividimos a ambos lados por -2 y individualizamos la sumatoria:
$\dfrac{\partial \epsilon}{\partial \alpha} = \sum_i y_i - \sum_i \alpha - \beta \sum_i x_i = 0$
El término $\sum_i \alpha$ es igual a la suma de n términos cada uno de los cuales vale α, por lo tanto este término vale $n\alpha$.
El término $\sum_i y_i$ es igual a $n\bar{Y}$ donde $\bar{Y}$ es la media de los y_i pues $\bar{Y} = (1/n)\sum_i y_i$. Similarmente, $\sum_i x_i = n\bar{X}$. Reemplazando obtenemos:
$n\bar{Y} - n\alpha - n\beta\bar{Y} = 0$
Dividiendo por n y despejando α obtenemos:

$$\alpha = \bar{Y} - \beta\bar{X}$$

Ahora derivamos parcialmente con respecto a β:

$\frac{\partial \epsilon}{\partial \beta} = 2\sum_i (y_i - \alpha - \beta x_i)(-x_i) = 0$
Dividiendo por -2 y expandiendo obtenemos:
$\frac{\partial \epsilon}{\partial \beta} = \sum_i x_i y_i - \alpha \sum_i x_i - \beta \sum_i (x_i)^2 = 0$
Utilizamos la definición de la media de los x_i:
$\sum_i x_i y_i - n\alpha \bar{X} - \beta \sum_i (x_i)^2 = 0$
Reemplazamos α por $\bar{Y} - \beta \bar{X}$:
$\sum_i x_i y_i - n(\bar{Y} - \beta \bar{X})\bar{X} - \beta \sum_i (x_i)^2 = 0$
es decir
$\sum_i x_i y_i - n\bar{X}\bar{Y} + n\beta(\bar{X})^2 - \beta \sum_i (x_i)^2 = 0$
o sea
$\sum_i x_i y_i - n\bar{X}\bar{Y} + \beta[n(\bar{X})^2 - \sum_i (x_i)^2] = 0$
Ahora despejamos β: $\beta = -\frac{\sum_i x_i y_i - n\bar{X}\bar{Y}}{n(\bar{X})^2 - \sum_i (x_i)^2}$
El signo menos lo usamos en el denominador y al final

$$\beta = \frac{\sum_i x_i y_i - n\bar{X}\bar{Y}}{\sum_i (x_i)^2 - n(\bar{X})^2}$$

Ahora construimos la siguiente tabla para aplicar nuestras fórmulas al problema específico:

Regresión de ingresos mensuales (y) *vs* tiempo de experiencia x					
	x_i	y_i	$x_i y_i$	$(x_i)^2$	$(y_i)^2$
	3	5	15	9	25
	2	5	10	4	25
	2	4	8	4	16
	4	9	36	16	81
	3	7	21	9	49
	5	11	55	25	121
Sumas	$\sum x_i$ =19	$\sum y_i$ =41	$\sum x_i y_i$ =145	$\sum (x_i)^2 = 67$	$\sum (y_i)^2$ =317

Calculamos los promedios. Como en este caso tenemos 6 pares de datos:
$\bar{x} = \sum x/6 = 19/6 = 3.17$
$\bar{y} = \sum y/6 = 41/6 = 6.83$

La línea de regresión (mín. cuadrados): $y = a + bx$ se determina por
$\beta = \frac{SS_{xy}}{SS_{xx}} = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2}$;
$\alpha = \bar{y} - \beta \bar{x}$.

Para nuestro ejemplo, esto se convierte en

$$b = \frac{145 - 6(3.17)(6.83)}{67 - 6(3.17)^2} = \frac{15.09}{6.71} = 2.24$$

$a = \bar{y} - b\bar{x}$

$a = 6.83 - (2.24)(3.17) = 6.83 - 7.10 = -0.27$

Así, la línea de regresión de mínimos cuadrados es

$y = -0.27 + 2.24x$

Comparamos esta respuesta con la que habíamos sacado visualmente:

$y = -0.66 + 2.33x$

Comparando las dos ecuaciones, la sacada por cálculo gráfico y la sacada por fórmulas, vemos que hay concordancia y por tanto podemos confiar en que hemos hecho bien las cosas.

Observemos que nosotros podemos hallar la línea de regresión para datos que se ajustan a cualquier modelo, digamos a una parábola. Por lo tanto, la línea de regresión de por sí no significa mucho. Lo que falta por hacer es demostrar que dicha línea es un buen modelo y que compite con otros. Por otra parte, no hay límite ni a la variabilidad ni a la complejidad de los modelos estudiados. Tenemos muchísima experiencia con este tipo de problemas y es parte de la estadística (Jobson, 1991).

609. **_Ejercicio_** *Halle la línea de regresión de mínimos cuadrados que mejor se ajusta al conjunto de puntos dados por* $\{(0, 10), (2, 9), (3, 7), (4, 4), (3, 4)\}$

12.2. Cuádricas

En esta sección estudiamos un tipo especial de figuras tridimensionales que representan generalizaciones a la parábola, la elipse, la hipérbola, etc.

610. **_Ejemplo_** *Una de las figuras 3D más simples es el cilindro, un ejemplo de los cuales es*

$y^2 + z^2 = 1.$

En dos dimensiones, esta ecuación representa un círculo, pero en tres dimensiones no hay restricción alguna a la variable x, la cual puede asumir cualquier valor. Por tanto, $y^2 + z^2 = 1$ representa un cilindro cuyo eje es el eje X.

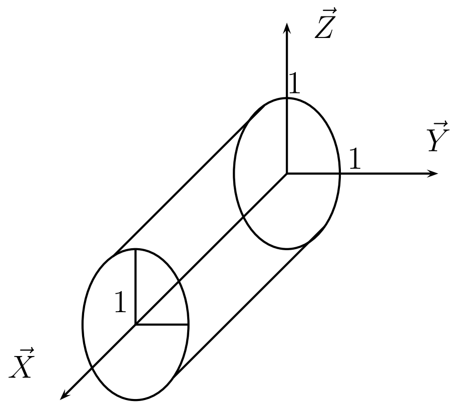


Figura 12.3. El cilindro $y^2 + z^2 = 1$.

611. ***Ejemplo*** *El hiperboloide de un manto.*

Consideremos ahora una expresión más compleja: $x^2 + y^2 - z^2 = 1$. Podemos reconocer que figura es con una simple manipulación: $x^2 + y^2 = z^2 + 1$. De esa forma, nos damos cuenta de que tenemos un círculo en el plano XY cuyo radio se relaciona con el valor de z: mientras más grande sea z, el radio del círculo es más grande, es decir, a medida que subimos el plano XY, la intersección con éste es un círculo cada vez más grande. Oficialmente se dice: la traza de $x^2 + y^2 = z^2 + 1$ con $z = constante$, es un círculo que es función creciente y par de z. Además, para poder apreciar mejor la silueta de esta figura, fijamos $y = 0$, y nos queda una hipérbole: $x^2 = z^2 + 1$. Por tanto, podemos hablar de una figura que se llama hiperboloide de un manto:

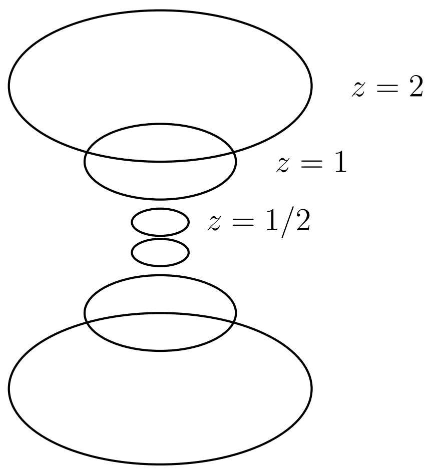


Figura 12.4. Trazas horizontales del hiperboloide $x^2 + y^2 = z^2 + 1$.

Si pegamos todos esos círculos obtenemos una figura tridimensional, la cual puede visualizarse por medio de una espiral:

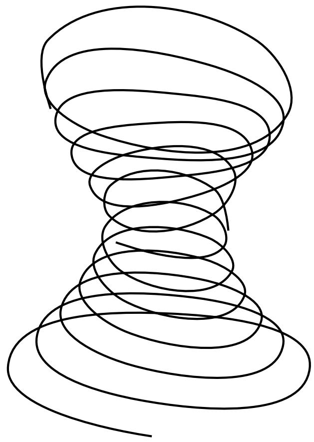


Figura 12.5. Una espiral trazadora sobre el hiperboloide $x^2 + y^2 = z^2 + 1$.

612. ***Ejemplo*** *La esfera.*

Para visualizar la ecuación $x^2 + y^2 + z^2 = 1$, podemos reconocer que $x^2 + y^2 = 1 - z^2$ representa un círculo cuyo radio decrece con z, la cual debe cumplir $z \in [-1, 1]$.

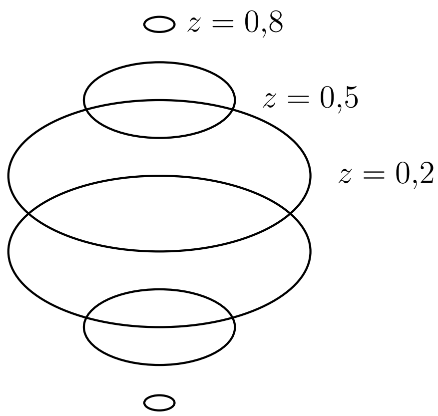


Figura 12.6. Bosquejo de una esfera.

Las trazas con $z = k$ dan círculos y al pegar todas esas curvas se obtiene una esfera, en este caso de radio 1.

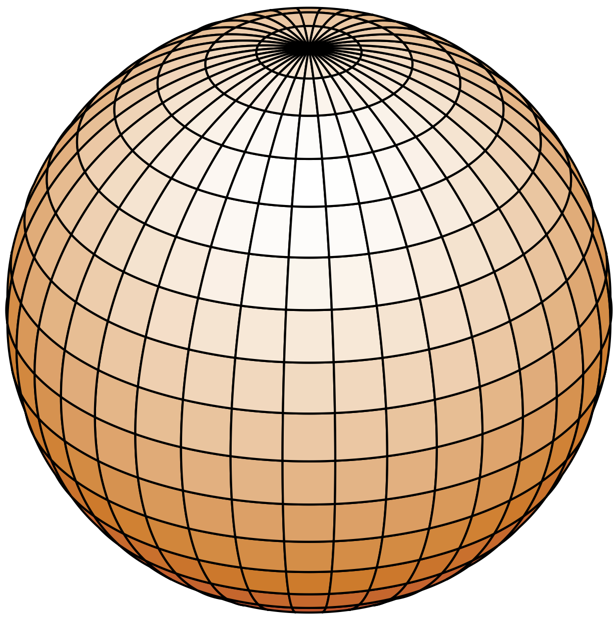


Figura 12.7. La esfera.

613. ***Ejemplo*** *Hiperboloide con dos mantos.*

Ahora presentamos un estudio de $x^2 + y^2 - z^2 = 1$, que nos da una figura que se conoce como hiperboloide de dos mantos. Se llama hiperboloide porque si damos a x el valor de cero o si le damos a y el mismo valor, nos queda una hipérbola. Por lo que la silueta de esta figura es una hipérbola vista desde dos ángulos diferentes. Podemos reorganizar para ver otra figura sencilla: $x^2 + y^2 = z^2 - 1$ es la ecuación de un círculo cuando $z^2 - 1 > 0$. Por lo tanto, en $-1 < z < 1$ no hay figura. Para valores de $z > 1$, tenemos un círculo cuyo radio crece con z. No hay que hacer el estudio de $z < -1$ pues la figura permanece lo mismo si intercambiamos z por $-z$, es decir, la figura es simétrica con respecto al plano XY. Por esa razón, la figura tiene dos partes, se dice dos mantos, una arriba del plano XY y otra simétrica, abajo de él. Además, la figura es invariante si intercambiamos x por y. Eso quiere decir que no se puede distinguir la dirección x de la de y.

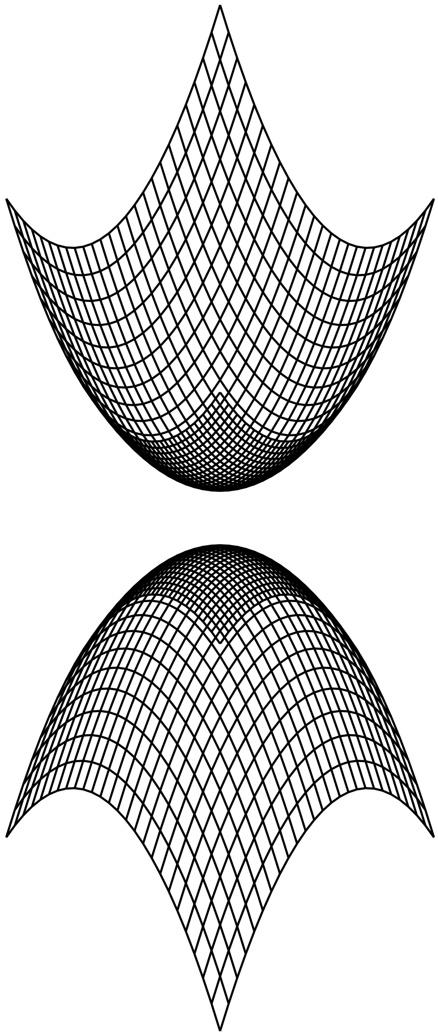


Figura 12.8. El hiperboloide de dos hojas $x^2 + y^2 = z^2 - 1$.

614. $\boxed{\textbf{\textit{Ejercicio}}}$ *Bosqueje las figuras:*

a) $x^2 + y^2 + z^2 = 9$

b) $x^2 + y^2 - 4z^2 = 0$

c) $4x^2 - y^2 - 4z^2 = 1$

d) $x^2 + 9y^2 - 4z^2 = 1$

e) $-x^2 + y^2/9 - 4z^2 = 1$

f) Observe que la mayoría de figuras pertenece a la misma clase. ¿Es eso una coincidencia o un hecho?

Pasemos ahora a explorar el poder del álgebra lineal. Nuestra tarea primaria es demostrar que no nos falta ninguna figura, lo cual quiere decir que hemos estado trabajando con ecuaciones polinómicas de segundo grado y si añadimos una ecuación del tipo $13x^2 - 10xy + 13y^2 = 72$, que no hemos visto, nada ganaremos. En general, todas las complicaciones posibles son ligeras modificaciones de lo ya visto.

615. ♣ *Definición.* *Una **k-cuádrica** es una ecuación polinómica de segundo grado en k variables. Por ejemplo,* $13x^2 - 10xy + 13y^2 = 72$ *es una 2-cuádrica, mientras que* $13x^2 - 10xy + 13y^2 - 72z^2 = 1$ *es una 3-cuádrica.*

616. ***Ejemplo*** *Veamos de qué manera un simple cambio de base puede ser usado para reescribir una expresión complicada en otra forma cuya interpretación es inmediata. Consideremos la cuádrica* $x^2 + 2y^2 + 3z^2 + 2\sqrt{2}yz = 1$. *La matriz asociada a ella es:*

$$D_B^B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & \sqrt{2} \\ 0 & \sqrt{2} & 3 \end{pmatrix}$$

Esta matriz es simétrica con valores propios 1, 4, 1. El valor propio 4 tiene como espacio propio el generado por $(0, \sqrt{2}, 2)$. El valor propio 1 tiene un espacio propio de 2 dimensiones, generado por una base con 2 elementos. Escogemos una base para este sub-EV tal que su unión con $(0, 2, \sqrt{2})$ sea una base ortogonal:

$$C = \begin{pmatrix} 0 & 1 & 6 \\ \sqrt{2} & -2 & 2 \\ 2 & \sqrt{2} & -\sqrt{2} \end{pmatrix}$$

Esta base da origen a una ortonormal

$$B = \begin{pmatrix} 0 & 1/\sqrt{7} & 6/\sqrt{42} \\ 1/\sqrt{3} & -2/\sqrt{7} & 2/\sqrt{42} \\ 2/\sqrt{6} & \sqrt{2}/\sqrt{7} & -\sqrt{2}/\sqrt{42} \end{pmatrix}$$

cuya transpuesta coincide con su inversa. La matriz diagonal asociada a esa cuádrica es, por tanto, $D = D(4, 1, 1)$. Eso implica que desde la base B la cuádrica se vea como

$4u_1^2 + u_2^2 + u_3^2 = 1$, la cual es un elipsoide, cuyos anchos principales son 1/2 en la dirección $(0, \sqrt{2}, 2)$, 1 en la dirección $(1, -2, \sqrt{2})$, y 1 en la dirección $(6, 2, -\sqrt{2})$.

Obsérvese que pudimos escoger nuestra base en el sub-EV de dos dimensiones a capricho, y eso se debe a que en este caso una sección transversal es un círculo. Si permanecemos en el plano que contiene el círculo, podemos cambiar de una base ortonormal a otra, pero la ecuación del círculo permanecería invariable. Eso explica nuestra libertad.

617. $\boxed{\textit{Ejercicio}}$ *Estudie las siguientes cuádricas. Encuentre los valores propios, los vectores propios correspondientes, la base propia ortonormalizada, la forma diagonal, y finalmente haga un dibujo. Las cuádricas son:*

a) $4x^2/3 + 4y^2/3 + 4z^2/3 + 2xy/3 - 2xz/3 - 2yz/3 = 1$

b) $x^2/3 + 4y^2/3 + z^2/3 + 2xy/3 - 8xz/3 - 2yz/3 = 1$

c) $x^2/3 + y^2/3 + z^2/3 - 4xy/3 + 4xz/3 + 4yz/3 = 1$

d) $-4x^2/3 - 7y^2/3 - 4z^2/3 + 10xy/3 - 4xz/3 - 10yz/3 = 1$

e) $-2x^2/3 + y^2/3 - 2z^2/3 - 4xy/3 - 2xz/3 + 4yz/3 = 1$.

618. **Ejercicio: al derecho y al revés.** *Aplique la magnificación propuesta a la cuádrica dada y recobre la original por una transformación inversa:*

a) $D(1, 2, 3)$, $x^2 + y^2 + z^2 = 1$

b) $D(1, -2, 3)$, $x^2 + y^2 + z^2 = 1$

c) $D(1, 1/2, 1/3)$, $x^2 - y^2 + z^2 = 1$

d) $D(1, 0, 3)$, $x^2 + y^2/4 + z^2/9 = 1$

e) $D(1, -1/2, 3)$, $x^2 - y^2 - z^2 = 1$.

Hemos estado trabajando en la construcción de figuras que están centradas en el origen, y que pueden ser rotadas y agrandadas o achicadas. Consideremos ahora una nueva variante:

619. **Ejercicio** *Revise la sección de 'traslaciones' al principio del curso y use la teoría correspondiente para trasladar las cuádricas dadas al punto dado:*

a) $x^2 + 3xz - z^2 = 1$; $(1, 2, 3)$

b) $-x^2 - y^2 + 3xy + 4z^2 = 1$; $(-1, 2, -3)$

c) $x^2 - y^2 + 5xy - 5xz3z^2 = 1$; $(4, 2, 3)$

d) $x^2 - y^2 + 3xy - 9z^2 = 1$; $(1, 5, 3)$

e) $x^2 - y^2 - 3xy - 5xz + z^2 = 1$; $(1, 2, 7)$.

Ahora enfrentemos el problema inverso: dada una cuádrica con términos de primer grado, encontremos una traslación que la reduzca a una cuádrica ordinaria centrada fuera del origen.

620. **Ejemplo** *Encontremos la traslación asociada a la cuádrica*

$s^2 - 11s + 3su - u^2 + 3u = 0$

En este caso, la traslación asume la siguiente forma: $s = x - a$, $u = z - b$. Por lo que la cuádrica se convierte en

$(x-a)^2 - 11(x-a) + 3(x-a)(z-b) - (z-b)^2 + 3(z-b) = 0$

Nuestro objetivo es cancelar los términos de la forma cx o kz, los cuales son distintivos de las traslaciones. Por lo tanto, desarrollamos nuestra expresión:

$x^2 + 3xz - z^2 + (-2a - 11 - 3b)x + (-3a + 2b + 3)z + a^2 + 11a + 3ab - b^2 - 3b = 0$

por lo tanto:

$-2a - 11 - 3b = 0$

$-3a + 2b + 3 = 0$

de lo cual encontramos que $a = -1$ y $b = -3$ y por tanto

$a^2 + 11a + 3ab - b^2 - 3b = -1$

Obtenemos

$x^2 + 3xz - z^2 + 0x + 0z - 1 = 0$

o

$x^2 + 3xz - z^2 = 1$

la cual es una cuádrica que ha sido rotada y cuyo estudio ya sabemos hacer. Tenemos un cilindro en el plano XZ, que es generado por una hipérbola.

621. $\boxed{\textbf{\textit{Ejercicio}}}$ *Las siguientes expresiones representan cuádricas que han sido trasladadas desde el origen al punto dado P. Encuentre la traslación y reescriba la cuádrica como vista desde los ejes trasladados, desde la cual se ve como una cuádrica ordinaria:*

a) $x^2 + 5xz - z^2 - 12x + 9z = 18$

b) $-x^2 - y^2 + +4z^2 + 9y = 1$

c) $x^2 - y^2 + 3z^2 - 16z = 1$

d) $x^2 - y^2 + 3xy - 9z^2 - 9z = 1$

e) $x^2 - y^2 - 3xy - 5xz + z^2 + 4x = 1$

12.3. Ejes principales

Los ejes principales tienen una vívida intepretación en la dinámica de cuerpos rígidos. No hay que pensar que eso es complicado. Veamos cómo es eso de entendible y de bonito.

Un cuerpo rígido es aquel que no se deforma al irse moviendo sea libremente o bajo el influjo de fuerzas externas. No existen cuerpos verdaderamente rígidos, pero sí existen cuerpos aproximadamente rígidos para los cuales se aplica la teoría.

El movimiento de un cuerpo rígido, como un avión de combate de geometría fija, se describe en cada instante por el movimiento de su centro de masa y por una rotación alrededor de un eje. Al siguiente instante, el centro de masa puede trasladarse y el eje de rotación puede cambiar.

Si una partícula puntual se mueve horizontalmente, hay dos cantidades que se conservan en ausencia de fuerzas externas, el momento lineal, que es un vector, y la energía cinética que es un escalar. El momento es $\vec{p} = m\vec{v}$ y la energía cinética es $K = (1/2)mv^2$.

Similarmente, si un cuerpo está rotando, aunque su centro masa no se traslade, sus partículas se están moviendo y, por lo tanto, debe haber algún tipo de indicadores, que muestren que rotar despacio es muy diferente de rotar rápido, como lo saben muy bien los karatecas que golpean con la pierna en rotación o los tenistas de mesa que muñequean al golpear la bola. Y por otro lado, hacer girar una puerta despacio es más fácil que hacerlo rápido y, también, el esfuerzo necesario parece ser más grande mientras más cerca se esté del eje de rotación de la puerta.

Los estudios han indicado que la dinámica de la rotación de cuerpos rígidos se describe con dos indicadores fundamentales: el momento angular y el torque. El primero es un vector que va paralelo al eje de rotación y de magnitud proporcional a la velocidad angular, y el segundo relaciona la fuerza y el brazo o distancia perpendicular al eje de giro.

Formalmente, el vector de momento angular, $\vec{J}$ se define como $\vec{J} = \sum M_n \vec{r}_n \times \vec{v}_n$, donde M_n es la masa de la partícula n, $\vec{r}_n$ es su distancia al eje de rotación y v_n es su velocidad. En particular, para un anillo, como una rueda de bicicleta de masa total M, de radio R y que gira con respecto a su eje con velocidad angular $\vec{\omega}$, se tiene que el momento angular es $\vec{J} = MR^2\vec{\omega}$. La velocidad angular es un vector perpendicular al anillo, lo mismo que $\vec{J}$. La cantidad $I = MR^2$ es una constante llamada el momento de inercia, por lo que

$$\vec{J} = I\vec{\omega}$$

Eso implica que en un anillo, el momento angular y la velocidad angular son paralelas.

Por otra parte, el torque $\vec{N}$ es $\vec{N} = \sum \vec{r}_n \times \vec{F}_n$. Entonces tenemos la ley fundamental que dice que si hay una fuerza que produzca un torque, entonces o cambia la dirección del eje de giro o cambia la velocidad angular o cambian ambos:

$$d\vec{J}/dt = \vec{N}.$$

Si no hay fuerza externa, el torque es cero y el momento angular se conserva. Eso es lo que permite que el que va en la bicicleta adquiera una cierta velocidad y no se caiga sin necesidad de pedalear. Esto también permite que la niña que gira sobre sus patines lo haga cada vez más rápido a medida que recoge los brazos.

En general, para cuerpos que son más complicados que los anillos, la relación entre las componentes del vector del momento angular $\vec{J} = (J_x, J_y, J_z)$ y las de la velocidad angular $\vec{\omega} = (\omega_x, \omega_y, \omega_z)$ se da por

$$\vec{J} = [I]\vec{\omega}$$

pero ahora $[I]$ ya no es un escalar sino una matriz, llamada **tensor de inercia** y está dado por

$$[I] = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix}$$

donde I es simétrico.

En componentes, la ecuación $\vec{J} = [I]\vec{\omega}$ se puede escribir como
$J_x = I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z$
$J_y = I_{yx}\omega_x + I_{yy}\omega_y + I_{yz}\omega_z$
$J_z = I_{zx}\omega_x + I_{zy}\omega_y + I_{zz}\omega_z$

Lo que esta ecuación quier decir es que, en general, la velocidad angular y el momento angular no son paralelos, como en un anillo.
Por otro lado, la energía cinética K asociada a la rotación se escribe como

$$K = \frac{1}{2}(\omega_x^2 I_{xx} + \omega_y^2 I_{yy} + \omega_z^2 I_{zz} + 2\omega_x\omega_y I_{xy} + 2\omega_y\omega_z I_{yz} + 2\omega_z\omega_x I_{zx})$$

donde cada uno de los términos de la forma $I_{\mu\nu}$ es una integral que involucra la distribución de masa con su distancia al eje de rotación. La expresión anterior también puede escribirse como

$$K = [\vec{\omega}]^T [I][\vec{\omega}]$$

Vemos que para $K = constante$, se tiene una forma cuadrática simétrica con coeficientes reales, por lo que puede diagonalizarse en reales. Además, todos los coeficientes son positivos, y por lo tanto, la forma cuadrática representa un elipsoide. Los vectores propios de $[I]$ son paralelos a los ejes del elipsoide, los cuales se llaman **ejes principales**.

Ahora bien, la energía cinética no puede depender de los ejes de referencia. Por eso tomamos la base que más simplifique las cosas: aquella dada por los ejes principales. Si llamamos a los valores propios como I_1, I_2, I_3, se tiene entonces que

$$K = \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2)$$

Sobre la misma base, el vector de momento angular se escribe como

$J_1 = I_1\omega_1$
$J_2 = I_2\omega_2$
$J_3 = I_3\omega_3$

lo cual permite escribir la energía cinética como

$$K = \frac{1}{2I_1}J_1^2 + \frac{1}{2I_2}J_2^2 + \frac{1}{2I_3}J_3^2$$

En el siguiente teorema tenemos un resultado que amerita el nombre de ejes principales que se ha dado a los ejes propios del momento de inercia.

622. $\diamond$ ***Teorema.*** *Cuando el cuerpo gira alrededor de un eje principal, $\vec{J}$ es paralelo a $\vec{\omega}$.*

Demostración. Un eje principal está definido por la ecuación

$$[I]\vec{u} = \lambda\vec{u}.$$

Se tiene que, si el cuerpo gira con respecto a un eje principal, entonces $\vec{\omega}$ y dicho eje son paralelos. Eso implica que

$$[I]\vec{\omega} = \lambda\vec{\omega}.$$

Eso quiere decir que si el cuerpo gira con respecto a un eje principal, el tensor de inercia se percibe como un escalar y por tanto,

$$\vec{J} = [I]\omega$$

se convierte en

$$\vec{J} = \lambda\omega.$$

es decir, el momento angular y la velocidad angular son paralelas. ■

Pero, ¿cuál es el misterio de que $\vec{\omega}$ y $\vec{J}$ sean o no paralelos? La razón es que desde el punto de vista del sistema de referencia determinado por los ejes principales, se tiene que

$$\frac{d\vec{J}}{dt} + \omega \times \vec{J} = \vec{N}.$$

Por lo que aunque no haya fuerza exterior que cree un torque, la discrepancia entre la velocidad angular y el momento angular hace las veces de torque.

Es el caso del trompo: como no es totalmente simétrico como una esfera, el tensor de inercia no es reemplazable por un escalar y por tal razón, habrá un ángulo entre la velocidad angular y el momento angular. ¿Cuál es el efecto? Que mientras que el trompo gira anclado en un mismo punto del piso, su eje de rotación va dando la vuelta. A eso se llama **precesión**.

Es posible que a uno se le ocurra pensar que todos estos temas ya han sido sobreestudiados y que no hay nada nuevo que decir. Pues no, ese no es el caso. Citemos varios ejemplos:

Debido a que la Tierra no es una esfera perfecta, toda la teoría se le aplica y resulta que las observaciones cuadran bien, pero no perfectamente con la teoría. Y nadie sabe por qué. De forma similar, creemos que el calentamiento global está modulado por la forma como se mezclan las aguas profundas de los océanos, cuyos movimientos están influenciados por la rotación de la Tierra. Pero de eso es poco lo que se sabe, es un tema abierto. Para acabar de completar, el campo magnético de la Tierra y su rotación deben admitir alguna conexión, cuya forma exacta es tema de estudios avanzados (Kittel y otros, 1965).

12.4. Máximos y mínimos

Hemos venido aprendiendo cómo se asocian matrices simétricas a algunas cuádricas. Esa asociación puede extenderse a las cuádricas derivadas de los máximos y mínimos de funciones reales. La idea fundamental es fácilmente comprensible para funciones $f : \mathbb{R} \to \mathbb{R}$, cuya gráfica es una curva suave. Un punto de la curva se llama **crítico** si la línea tangente en ese punto es horizontal. Un punto crítico puede ser mínimo o máximo o punto de inflexión.

Hay varias tecnologías para dilucidar la naturaleza de un punto crítico. Una de ellas es el polinomio de Taylor, el cual sirve para asociar a un punto crítico de una función la parábola tangente, algo de la forma $y = a(x-b)^2 + c$. Si la parábola abre hacia arriba, cuando $a > 0$, tenemos un mínimo. Si la parábola abre hacia abajo, cuando $a < 0$, tenemos un máximo local. Si la parábola es degenerada y es realmente una línea, uno no sabe qué hacer, pues todo puede darse.

Desarrollemos la teoría correspondiente para funciones con dos variables, del tipo $f\colon \mathbb{R}^2 \to \mathbb{R}$. En este caso un **punto crítico** es aquel cuyo plano tangente es horizontal. En vez de la parábola tangente, ahora tenemos el paraboloide tangente, el cual es una figura tridimensional cuyas secciones con ciertos planos son parábolas. Los tres tipos fundamentales de paraboloides son los siguientes:

1. Un minimoide: $z = ax^2 + by^2$ con $a > 0$ y $b > 0$. Por ejemplo, $z = x^2 + y^2$. Observemos que para valores constantes y positivos de z, la figura describe un círculo, cuyo tamaño crece a medida que crece z. Por lo que un minimoide es media cáscara de huevo que abre hacia arriba: en el punto crítico tenemos un mínimo.

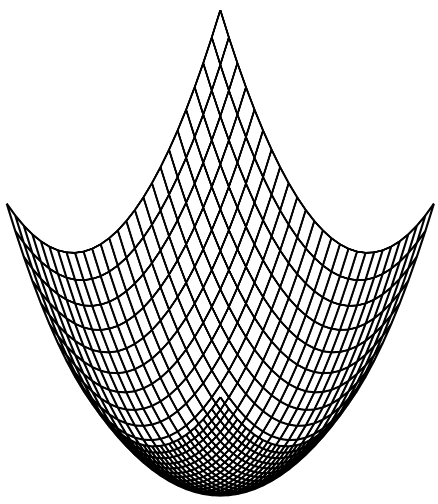


Figura 12.9. Minimoide $z = x^2 + y^2$.

Lo mismo para con $z = 4x^2 + 9y^2$. Las secciones horizontales son elipses, que son deformaciones de un círculo. Una elipse tiene un círculo promedio, que en este caso es una función monótona de z: mientras más grande z, más grande es la elipse. Cuando

$z = 0$, la elipse se reduce a un punto. Como podemos ver, el paraboloide abre hacia arriba. Si se trata de un paraboloide tangente a un punto crítico, estaríamos en el caso de un mínimo.

2. Un maximoide: $z = ax^2 + by^2$ con $a < 0$ y $b < 0$. Por ejemplo,
$z = -(4x^2 + 9y^2) = -4x^2 - 9y^2$.
Este es un minimoide patas arriba: tenemos un máximo, en el punto crítico.

3. Una silla: $z = ax^2 + by^2$ con $a > 0$ y $b < 0$ o $a < 0$ y $b > 0$. Por ejemplo, $z = -x^2 + y^2$ o $z = 4x^2 - 9y^2$. Analicemos el caso más simple: $z = -x^2 + y^2$. Para $y = 0$ tenemos una parábola en x que abre hacia abajo. Si leemos la ecuación como $z = y^2 - x^2$, para x fijo, tenemos una parábola en y que está bajo el plano $z = 0$ la cantidad x^2. Mientras más distante esté x del origen, más abajo estará el mínimo de la parábola. Por lo que una silla se determina por dos parábolas. Una silla se ve maximoide desde en lado y minimoide desde el otro. Una silla tiene la estructura adecuada para que el vaquero se siente extendiendo sus piernas a lo largo del máximo mientras su cuerpo adquiere estabilidad por el mínimo traversal.

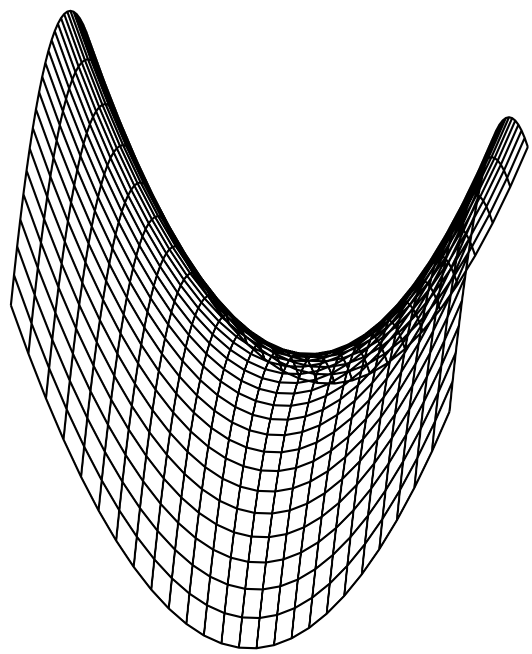


Figura 12.10. La silla $z = -x^2 + y^2$. La dirección X sale desde la página.

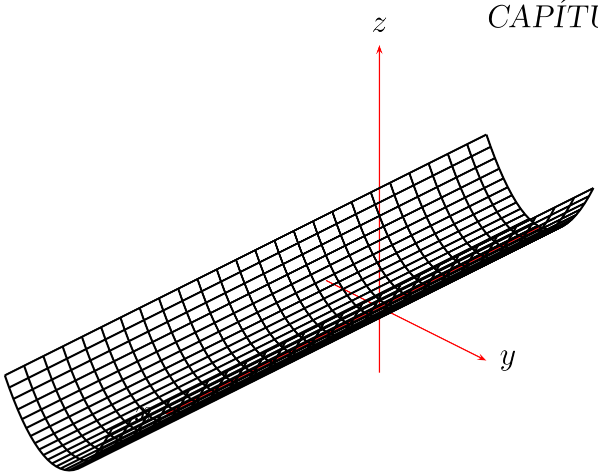


Figura 12.11. El paraboloide degenerado $z = y^2$.

Hay además unos casos patógenos que es bueno saber:

a) Un mínimo o máximo degenerado: sucede cuando el paraboloide tangente no tiene las 3 variables, como $z = y^2$. En este caso, tenemos que el mínimo queda no en un punto sino en una línea.

b) Un extremoide defraudado. Eso pasa cuando el paraboloide tangente a la figura en un punto crítico es un plano, el cual no sirve para dilucidar la naturaleza del punto. En efecto: puede verse que el paraboloide tangente a la figura $z = 4x^4 + 9y^4$ en $(0, 0)$ es un plano, $z = 0$. Sin embargo, la figura tiene un mínimo. Similarmente, el paraboloide tangente a $z = -4x^4 - 9y^4$ también es $z = 0$ y sin embargo la figura tiene un máximo. También pasa lo mismo con la 'silla' $z = -4x^4 + 9y^4$. Vemos, pues, que el paraboloide tangente que termina siendo un plano no tiene ningún mérito para descifrar la naturaleza de un punto crítico.

623. ***Ejemplos*** *Los siguientes paraboloides son tangentes en* $(0, 0)$ *a ciertas funciones, en donde ellas tienen un punto crítico. Decidamos la naturaleza del punto crítico, si se trata de máximo, mínimo o silla:*

a) $z = x^2/4 - y^2/9$, una silla.

b) $z = x^2/16 - y^2/25$, una silla.

c) $z = 3x^2 - y^2/9$, una silla.

d) $z = -4x^2 - y^2/16$, un máximo.

e) $z = x^2/4 + 7y^2$, un mínimo.

f) $z = x^2/9 - y^2/4$, una silla.

g) $z = -x^2/9 - 3y^2$, un máximo.

h) $z = 5$, no se sabe.

624. **Ejercicio** *Los siguientes paraboloides son tangentes en $(0,0)$ a ciertas funciones, en donde ellas tienen un punto crítico. Decida, si es posible, la naturaleza del punto crítico:*

a) $z = -x^2/4 - y^2/9$

b) $z = -x^2/16 - y^2/25$

c) $z = 3x^2 + y^2/9$

d) $z = -4x^2 + y^2/16$

e) $z = 2x^2/4 - 7y^2$

f) $z = -x^2/9 + y^2/4$

g) $z = -x^2/9 - 3y^2$

h) $z = -7$

i) $z = +7$

Observemos que en esta muestra de puntos críticos, la mayoría son sillas. ¿Es eso una coincidencia o qué?

625. **Ejercicio** *Decimos que un punto crítico es **estable** cuando es un mínimo o máximo. Si se trata de una silla, decimos que es **inestable**. Pruebe que cuando una cuádrica depende de n variables, la probabilidad de que el sistema sea estable en un punto crítico es $2/2^n = 1/2^{n-1}$. Esto parece implicar que cuando muchas variables afectan un sistema, la probabilidad de inestabilidad es muy alta. ¿Cómo se ataca este problema en la industria? ¿En las empresas? ¿En las sociedades? ¿Será la tendencia a la inestabilidad una razón por la cual en la mayoría de culturas se apela a dioses y agüeros para remediar la incapacidad de conservar la estabilidad o el desarrollo sostenible?*

626. **Ejemplo** *Supongamos que el paraboloide tangente a cierta figura en un punto crítico es $z = x^2 + y^2 - 4xy$. ¿Tenemos aquí un máximo o un mínimo o qué? Este es un problema de diagonalización. La figura es un paraboloide rotado sobre el plano XY. El paraboloide está descrito por*

$$z = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Los valores propios de la matriz asociada son 3 y -1. Esto significa que hay una base ortonormal que genera las coordenadas (u, v), desde la cual la figura se ve como:

$$z = \begin{pmatrix} u & v \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = 3u^2 - v^2,$$

la cual es una silla. Por lo tanto, el punto crítico se clasifica como silla.

627. **Ejercicio** *Reescriba los siguientes paraboloides en una base diagonalizante y decida su naturaleza:*

a) $z = -x^2/4 - y^2/9 + 4xy$

b) $z = -x^2/16 - y^2/25 + 7xy$

c) $z = 3x^2 + y^2/9 - 2xy$

d) $z = -4x^2 + y^2/16 + xy/5$

e) $z = 2x^2/4 - 7y^2 - 7xy$

f) $z = -x^2/9 + y^2/4 + 8xy$

g) $z = -x^2/9 - 3y^2 - xy$

628. **Ejercicio** *Pruebe que $z = p(x, y)$, donde p es un polinomio de segundo grado, define un punto crítico estable si el determinante de la matriz asociada es positivo. ¿Qué pasa en otras dimensiones?*

12.5. Sistemas dinámicos lineales

Un sistema dinámico es un sistema que cambia con el tiempo. Algunos sistemas se estudian más fácilmente si se modela el tiempo como una variable continua. De toda maneras, los computadores han impuesto la necesidad de modelar el tiempo como variable discreta. Una descripción cruda de un sistema dinámico produciría una lista de los estados ocupados por el sistema en tiempos sucesivos (discretos). ¿Será posible inferir el estado de un sistema en un futuro, si se conoce una condición inicial? Eso es perfectamente posible si sabemos cómo inferir E_{t+1}, el estado del sistema en el tiempo $t + 1$, sabiendo el estado E_t del sistema en el tiempo t. La inferencia se hace por un operador, el operador de evolución, O_t, que dictamina la evolución del sistema:

$E_{t+1} = O_t(E_t)$

Un **sistema es lineal** cuando O_t es una matriz. En ese caso, uno puede saber el estado en un tiempo lejano por medio de iteraciones del operador de evolución:

$E_{t+k} = M(...(M(E_t))) = M^k(E_t)$

El problema es que multiplicar matrices es algo tedioso, por lo que uno puede ayudarse de los computadores. Pero veamos qué da de sí el proceso de diagonalización. Veamos cómo se calcula F^n para una matriz F que puede ser diagonalizada.

Si la matriz F puede ser diagonalizada, entonces puede ser escrita como el producto de 3 matrices:

$F = PDP^{-1}$

donde D es una matriz diagonal. Elevando a la potencia n a ambos lados:

$F^n = (PDP^{-1})^n = (PDP^{-1})(PDP^{-1})...(PDP^{-1})(PDP^{-1})$

$= PD(P^{-1}P)D(P^{-1}...P)D(P^{-1}P)DP^{-1})$

$= PD(\mathbb{I})D(\mathbb{I}...\mathbb{I})D(\mathbb{I})DP^{-1})$
$= PDD...DDP^{-1} = PD^nP^{-1}$

Por tanto, el problema de iterar una matriz diagonalizable se convierte en el problema de iterar una matriz diagonal.

629. ⋄ ***Teorema.*** *D^n de una matriz diagonal D es la matriz diagonal que en su diagonal pone la potencia n de cada uno de sus elementos.*

630. ***Ejemplo*** *Si la matriz diagonal:*

$$D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{pmatrix}$$

Entonces

$$D^8 = \begin{pmatrix} 3^8 & 0 & 0 \\ 0 & 5^8 & 0 \\ 0 & 0 & 7^8 \end{pmatrix}$$

631. ***Investigación*** *Demuestre el teorema anterior en su forma general usando el principio llamado Inducción matemática.*

Vamos a discutir ahora un requisito para ser diagonalizable.

632. ⋄ ***Teorema.*** *Sea M una matriz diagonalizable que tenga al menos un valor propio no nulo. Entonces M^n es diferente de la matriz cero para cualquier n finito.*

Demostración. Puesto que M es diagonalizable se puede escribir como $M = PDP^{-1}$ donde D es una matriz diagonal y P es una matriz de cambio de base. Por lo que $M^n = PD^nP^{-1}$. Como D^n tiene en su diagonal los valores propios a la potencia n, y al menos uno de ellos no es cero, entonces D^n nunca será la matriz cero por lo que M^n tampoco. ■

Este teorema significa que si una matriz A es tal que existe un natural n tal que A^n es la matriz cero, entonces A no puede ser diagonalizable ni en reales ni en complejos.

633. ***Ejemplo y definición*** *Sea la TL $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T(\vec{i}) = \vec{j}$, $T(\vec{j}) = \vec{0}$, entonces la matriz de T^2 es la matriz cero. En efecto, $T(T(\vec{i})) = T(\vec{j}) = \vec{0}$ en tanto que $T(T(\vec{j})) = T(\vec{0}) = \vec{0}$. Como T^2 manda una base sobre cero, manda a todo el mundo sobre cero y de esa forma es idénticamente cero. La matriz de T en la base natural es*

$$T = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

la cual cumple $T^2 = 0$. Una matriz M tal que $M^n = 0$ para algún n se llama ***nilpotente****. Las matrices nilpotentes no son diagonalizables ni en reales ni en complejos.*

634. ***Ejemplo y definiciones*** *Tratemos de diagonalizar la matriz T del ejemplo anterior para ver qué aprendemos:*

$$T = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Tenemos que

$$T - \lambda \mathbb{I} = \begin{pmatrix} -\lambda & 0 \\ 1 & -\lambda \end{pmatrix}$$

tiene como determinante $\lambda^2 = 0$, por lo que tiene un único valor propio, el cual es 0. Para hallar un vector propio asociado a dicho valor procedemos:

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

La primera ecuación no da nada y de la segunda obtenemos $x = 0$ por lo que y puede tomar cualquier valor, digamos 1. Por tanto, al valor propio 0 le corresponde el vector propio $(0, 1)$.

El problema es que nos falta otro vector propio para completar una base, puesto que T tiene un único valor propio de **multiplicidad algebraica** 2 pues se repite dos veces. También se usa el término **multiplicidad geométrica** de un valor propio para denotar la dimensión del espacio propio asociado a un valor propio. En nuestro ejemplo, el valor propio 0 tiene multiplicidad geométrica 1, pero multiplicidad algebraica 2, por lo que origina un déficit de vectores propios. Y por eso no es diagonalizable.

Ahora bien, el hecho de que una matriz tenga valores propios repetidos no es ningún obstáculo para que ella sea diagonalizable. Por ejemplo, la matriz identidad es diagonal y es diagonalizable y tiene un único valor propio 1 con multiplicidad algebraica n, la dimensión del espacio, y ese único valor propio tiene multiplicidad geométrica n: no hay déficit de valores propios para fabricar una base.

Nuestra conclusión es entonces:

635. ⋄ ***Teorema.*** *Una matriz que denota una TL $T\colon \mathbb{R}^n \to \mathbb{R}^n$ es diagonalizable ssi hay suficientes vectores propios para construir una base de $\mathbb{R}^n$. Eso es equivalente a decir que la multiplicidad algebraica de cada valor propio es igual a su multiplicidad geométrica. O que la suma de las dimensiones de todos los espacios propios es igual a la dimensión del espacio.*

636. ***Ejercicio*** *Decimos que un sistema es un **sistema dinámico discreto** cuando el número de estados que puede ocupar es discreto y cuya transición puede modelarse por una matriz. Pruebe que si un sistema es lineal con operador de evolución una matriz constante M, y si tiene un sólo estado de equilibrio, entonces tenderá rápidamente a él desde cualquier condición inicial. Un sistema dinámico puede tener valores propios complejos, los cuales generan oscilaciones. Para verlo, recuerde que las potencias de i son: $i = 1$, $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, $i^5 = i$, $i^6 = -1$, etc., en ritmo periódico.*

637. $\boxed{\textbf{\textit{Ejercicio}}}$ *Estudie la conducta del sistema dinámico cuya condición inicial y operador de evolución son:*

a) $\begin{pmatrix} 20 \\ 10 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$

b) $\begin{pmatrix} 10 \\ 20 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}$

c) $\begin{pmatrix} 20 \\ 20 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$

La teoría de sistemas dinámicos discretos se ha beneficiado mucho de la pregunta siguiente: las matrices nilpotentes parecen ser un obstáculo a la diagonalización, ¿qué pasa si a una matriz no diagonalizable le quitamos su parte nilpotente, si la tiene? Esa pregunta fue respondida por el el siguiente teorema.

638. $\diamond$ ***Teorema de descomposición de Jordan-Chevalley.*** *Excepto por un cambio de base, toda matriz invertible puede descomponerse como la suma de una matriz diagonal más una matriz nilpotente. Es decir, si A es una matriz invertible, entonces puede descomponerse como $A = M(D+N)M^{-1}$ donde M es una matriz de cambio de base, D es una matriz diagonal y N es una matriz nilpotente cuyas únicas entradas no nulas están encima de la diagonal y son unos.*

Demostración: investigación (Lang, 2002).

639. ***Ejemplo y ejercicio*** *Sea $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Si A es la matriz de transición de un sistema dinámico que toma los estados del presente y los lleva a la siguiente unidad de tiempo, lo que nos interesa es poder calcular los estados en cualquier momento futuro. Para ello hay que hallar las potencias de A. Por ejemplo, hallemos A^{2000}.*

Solución: La matriz A es invertible pero no es diagonalizable (ejercicio), pero puede descomponerse como la suma de una matriz diagonal más una nilpotente que tiene unos encima de la diagonal:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Vemos que $A = \mathbb{I} + N$ donde $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, cuyo cuadrado da la matriz de ceros: $N^2 = 0$. Ahora recordamos el desarrollo del binomio:

$$(a+b)^n = a^n + na^{n-1}b + \frac{n(n-1)}{2}a^{n-2}b^2 + ... + b^n$$

y lo aplicamnos sobre la descomposición de A:

$$A^{2000} = (\mathbb{I}+N)^{2000} = \mathbb{I}^{2000} + 2000\mathbb{I}^{1999}N^1 + \tfrac{n(n-1)}{2}\mathbb{I}^{1998}N^2 + ... + N^{2000}$$

Como la matriz N es tal que su cuadrado y todas las potencias superiores son cero, nos quedamos con:

$A^{2000} = (\mathbb{I} + N)^{2000} = \mathbb{I} + 2000N$

Reemplazando obtenemos:

$$A^{2000} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 2000 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2000 \\ 0 & 1 \end{pmatrix}.$$

640. ♣ ***Fuerte advertencia.*** *Nosotros utilizamos el desarrollo del binomio que todos conocemos que es válido para números. ¿Podemos aplicarlo a matrices? Podemos, pero con una condición: que las matrices conmuten, es decir, que para matrices A y B, se cumpla que $AB = BA$. Para matrices no conmutativas hay que tener más cuidado. Por ejemplo: $(A+B)^2 = A^2 + AB + BA + B^2$ que es igual a $A^2 + 2AB + B^2$ solamente cuando $AB = BA$.*

Una de las implicaciones de la no conmutatividad está en teoría de partículas elementales: si la interacción entre partículas se representa por matrices conmutativas, los mediadores de la interacción no tienen masa, como sucede con la interacción electromagnética entre electrones, la cual es mediada por el fotón que no tiene masa y por consiguiente viaja a la velocidad de la luz. Pero no es así con la iteracción fuerte entre neutrones, los cuales tienen como mediadores a los mesones que tienen masa y que viajan a velocidades menores que las de la luz y cuyo rango de interacción es sólo a muy cortas distancias (Nash y Sen, 1983).

641. $\boxed{\textbf{\textit{Ejercicio}}}$ *Halle A^{2000} si $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Ayuda: descomponga a A como una diagonal D más una nilpotente N y calcule N^3. Luego aplique a $(D+N)^{2000}$ el desarrollo del binomio.*

12.6. Ejercicios de repaso

1. Grafique la cónica $8x^2 - 4xy + 5y^2 - 2x + y = 1$.

2. Efectúe una rotación de ejes para identificar la cónica: $5x^2 + 4xy + 2y^2 + 24 = 0$. Determine la matriz Q ortogonal que diagonaliza la matriz A asociada a la forma cuadrática y escriba la ecuación de la cónica en el nuevo sistema de ejes. Trace su gráfica.

12.7. Resumen

Las aplicaciones del álgebra lineal no tienen límite ni en cuanto a cantidad ni en cuanto a variabilidad. Hemos pasado revista a algunas aplicaciones más bien fáciles de captar y que al mismo tiempo arrojan cierta luz sobre problemas muy complicados.

12.8. Gran taller de repaso

1. Conteste falso (F) o verdadero (V) según sea el caso. Si la proposición es verdadera dé una argumentación matemática. Si es falsa, proporcione un ejemplo:

   a) Cualesquiera dos vectores en $\mathbb{R}^2$ generan todo $\mathbb{R}^2$.

   b) El producto punto de un vector consigo mismo da la magnitud del vector.

   c) Si $\vec{v}$ y $\vec{w}$ son vectores en $\mathbb{R}^n$ de la misma magnitud, entonces la magnitud del vector $\vec{v}-\vec{w}$ es cero.

   d) Si las matrices A y B son invertibles, entonces $A+B$ es invertible.

   e) Existen exactamente dos vectores perpendiculares a cualquier vector no nulo en $\mathbb{R}^n$.

   f) Si la imagen bajo una transformación lineal T de un n-plp B en $\mathbb{R}^n$ tiene volumen 12, la n-plp B tiene volumen $\frac{12}{|A|}$.

   g) Toda matriz invertible es diagonalizable.

   h) Si $|A|=2$ y $|B|=3$, entonces $|A+B|=5$.

   i) Los vectores propios de una matriz A son los mismos que los de la matriz A^T.

   j) El conjunto de puntos (x,y) en el plano para los cuales la matriz $A=\begin{pmatrix} x & y & 1\\ 1 & 2 & 1\\ 3 & 3 & 1\end{pmatrix}$ es **singular** (que está sola, que no tiene inversa) describen una recta con pendiente $m=\frac{1}{2}$.

   k) Si la proyección del vector $\vec{b}$ sobre el subespacio W es $\vec{b}$ mismo, entonces $\vec{b}$ es ortogonal a todo vector en W.

   l) Todo subespacio no trivial de $\mathbb{R}^n$ posee una base ortonormal.

   m) Toda matriz ortogonal tiene espacio nulo $\{0\}$.

   n) Si B y B' son bases ortonormales, entonces la matriz de cambio de base $I_{B'}^{B}$ es una matriz ortogonal.

   ñ) La proyección de un vector $\vec{b}$ sobre el gen$(\{\vec{a}\})$ es un múltiplo escalar de $\vec{b}$.

   o) Si los vectores $\vec{b}$ y $\vec{c}$ tienen la misma proyección sobre un subespacio W, entonces $\vec{b}=\vec{c}$.

   p) Toda matriz ortogonal tiene espacio nulo $\{0\}$.

   q) Si A es simétrica y ortogonal, entonces $A^2=\mathbb{I}$.

   r) Para toda escogencia de las bases B y B' se tiene que $(DetI_{B'}^{B})=1$.

2. Sea $H=\left\{\begin{pmatrix} x\\ y\\ z\end{pmatrix} :\ x=-t, y=t, z=2t;\ t\in\mathbb{R}\right\}$:

   a) Halle $H^{\perp}$ el complemento ortogonal de H en $\mathbb{R}^3$.

b) Encuentre una base ortonormal B de $\mathbb{R}^3$.

c) Sea $H = \text{gen}(\vec{b})$, donde $\vec{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3$, y sea $\vec{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \mathbb{R}^3$, halle $\text{Proy}_{H^\perp}\vec{a}$.

d) Con referencia al inciso anterior, escriba el vector $\vec{a}$ como $a_H + a_{H^\perp}$.

3. Encuentre la transformación lineal $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ que refleja un vector $\vec{v} \in \mathbb{R}^3$ sobre el plano $x - y + z = 0$.

4. Escriba la tercera columna de la matriz $\begin{pmatrix} 2/7 & 3/\sqrt{3} & . \\ 3/7 & -2/\sqrt{13} & . \\ 6/7 & 0 & . \end{pmatrix}$ para que sea una matriz ortogonal.

5. Encontrar una matriz C invertible tal que $D = C^{-1}AC$ es una diagonalización ortogonal de la matriz simétrica $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

6. Mostrar que toda matriz simétrica cuyos valores propios son 0 y 1 únicamente, es la matriz de una **proyección**. Es decir, hay que mostrar que existe una base $B = \{b_1, b_2\}$ tal que la imagen de uno de ellos es $\vec{0}$ y la del otro es él mismo.

7. Use el método de los mínimos cuadrados para encontrar la línea recta que mejor aproxima los datos: $(1, 1)$, $(10, 8)$, $(14, 12)$, $(16, 20)$.

8. Encontrar la matriz de cambio de coordenadas de la base B a la base B', donde $V = P_{\leq 3}$, $B = \{x^3 + x^2 + 1, 2x^3 - x^2 + x + 1, x^3 - x + 1, x^2 + x + 1\}$ y $B' = \{x^3, x^2, x, \}$.

9. Sea $T\colon W \to W$ donde $W = \text{gen}(\{e^x, xe^x\})$ y T es la transformación derivada, $B = \{e^x, xe^x\}$ y $B' = \{2xe^x, 3e^x\}$. Encontrar las matrices de representación R_B, $R_{B'}$ y la matriz C tal que $R_{B'} = C^{-1}R_BC$.

CAPÍTULO 13

RESPUESTAS

Capítulo 1

14. $x = 1$, $y = 1$
17. *a*) $x = 1$, $y = 1$
b) $x = 54/11$, $y = 19/11$
c) $x = 3$, $y = 2$
18. 1.000 y 1.000.
23. a)

$$\begin{pmatrix} 3 & 4 & \vdots & 5 \\ 0 & 0 & \vdots & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & \vdots & 5 \\ 0 & 1 & \vdots & 0 \\ 0 & 0 & \vdots & 1 \end{pmatrix}$$

Capítulo 2

57. b) $\vec{u} + \vec{v} = (9, 3)$
d) $4\vec{u} + 2\vec{v} = (4, 16)$
$2\vec{u} + 3\vec{w} = (-2, 0)$
$3\vec{w} + 5\vec{v} = (29, -27)$
$-\vec{v} = (-7, 3)$,
$-\vec{w} = (-2, -6)$, $-\vec{w} = (2, 4)$
60. a) (-6,30)
b) $(10, 24)$
c) $(-41, 3)$

72 d. Cinco barcos pueden usarse como sigue: póngalos en las esquinas de un pentágono regular, que debe ser lo suficientemente grande como para que los barcos puedan ocultarse unos de otros debido a la curvatura de la Tierra. De esa forma, la curvatura de la Tierra puede medirse en cada punto y uno también podría saber si la Tierra es como una esfera o como un cilindro.

81. Nuestra prueba es válida para números naturales solamente. Es necesario extenderla a racionales y después a reales.

90. La norma cuadrado es
$\|(a, b, c)\|^2 = a^2 + b^2 + c^2$
92. $\sqrt{5}$, $\sqrt{13}$, $\sqrt{5}$, $\sqrt{34}$, $\sqrt{14}$, $\sqrt{14}$, $\sqrt{50}$

97. Los puntos sobre una esfera tienen una distancia igual al centro, y es igual al radio de la esfera. Por lo tanto, la ecuación de la esfera con centro en (a, b, c) y radio r es
$(x - a)^2 + (x - b)^2 + (x - c)^2 = r^2$.
104. $(u_1, u_2, u, ..., u_n) \cdot (v_1, v_2, u, ..., v_n)$
$= (u_1 v_1 + u_2 v_2 + u_n v_n)$
106. a) 0, b) 0, c) 30 d) 1, e) 0
114. a) $180 = \pi$, b) $90 = \pi/2$,
c) Arccos$(30/\sqrt{1000})$,
d) Arccos$(1/3)$ e) $90 = \pi/2$
116. $(3\sqrt{26}(-5, 1)$
121. Son paralelas.
141. a) $y = (-4/3)(x - 2) + 3$
b) $y = -9(x - 2) - 4$
c) $y = (2/3)(x + 2) + 3$
d) $x = 1$
e) $y = 2$
142. a) $y = -2(x - 2) + 3$
b) $y = 5(x - 2) - 4$
c) $y = -5(x + 2) + 3$

d) $y = 6(x-1) - 3$
e) $y = 2$
143. a) $y = 2x + 1$
b) $y = -3x - 1$
c) $x = 0$
d) $y = -x - 3$
e) $y = -8$
144. a) $y = (11/2)x - 8$
b) $y = -x - 3$
c) $y = 2x + 4$
d) $y = -3$
e) $y = (1/4)x - 2$
155. a) $y = -3x/2 + 5/2$
b) $y = -3x/2 + 31/2$
c) $y = -3x/2 + 5/2$
d) $y = -3x/2 + 12$
e) $y = -3x/2 + 15/2$
163. a) $(1.6, 0.8)$
167. a) $\sqrt{2}$
b) $\sqrt{10}/5$

c) Dos puntos sobre una línea: $(0,3)$ y $(-1,0)$. Por lo tanto, el vector director es $(1,3)$. El segmento de $(2,3)$ a $(0,3)$ es $(2,-6)$ cuya norma es $\sqrt{40}$. Proyectamos el vector $(2,-6)$ sobre $(1,3)$ para obtener $(-8/5, -24/5)$, cuya norma es $\sqrt{640/25}$. Formamos un triángulo rectángulo y aplicamos Pitágoras para encontrar la distancia 3.79.

181. $a+b$ se cruzan en $(2, -1/3)$.

$c+e$: son paralelas.

$a+d$: las dos líneas son la misma: un grado de libertad.

$b+d$ se cruzan en $(8, 5/3)$.

183. Ración diaria: 100/11.2 de papas y 100/43.8 de pollo.

187. $(1,0,0)$, $(0,1,0)$, $(1,0,0)$, $(1,0,0)$, $(0,1,0)$, $(0,0,1)$.

191. a) $-x+1) + 5(z+3) = 0$
b) $3x + 2(y+2) + 7(z+3) = 0$
c) $-6(x+1) + 3(y+4) + 8(z+3) = 0$
193. a) $-3(x+1) = 0$
b) $-2x + 16(y-1) - 16(z-3) = 0$
c) $-4(x-1) - 4(z-1) = 0$

d) Los puntos pertenecen al eje X: hay demasiados planos y un grado de libertad.

199. El cuento es verídico: las vigas de la primera casa miden la raíz cuadrada de 2, 5, 8, 5. El techo es un plano: $-4x - 4y + 6z = 10$. Las columnas de la casa segunda no generan un plano, el cual se determina con 3 puntos no colineales y no cuatro.

202. a) $36/\sqrt{21}$.

204 a) Encuentre la distancia entre $x - 2y + 4z = 6$ y $x - 2y + 4z = 18$

207. a) $(1-t, 1-t, 4-4t)$ o $(t, t, 4t)$
b) $(-1+t, 2-2t, 2-2t)$ o $(-t, 2t, 2t)$
c) $(2-2t, 3-3t, -1+t)$ o $(2t, 3t, -t)$

212. a) Director $(1,1,1)$.

Paramétricas: $x = 1 + t$, $y = 2 + t$, $z = 3 + t$

Simétricas: $x - 1 = y - 2 = z - 3$

b) Director $(3, -3, -3)$.

Paramétricas: $x = -1 + 3t$, $y = -3t$, $z = 2 - 3t$

Simétricas $(x+1)/3 = -y/3 = (2 - z)/3$

215. e) El segmento de $(1,-4,0)$ a $(2,-7,-2)$ es $(-1,3,2)$, cuya norma es $\sqrt{57}$. La proyección de $(2,-7,-2)$ sobre $(2,3,5)$ es $(-27/57)(2,3,5)$, cuya norma es $27/\sqrt{57}$. El teorema de Pitágoras da la distancia del punto $(-1,3,2)$ a la línea y es 6.65.

218. d) $P = (0,0,-2)$ está en el primer plano, $Q = (0,0,1/2)$ en el segundo. El segmento $PQ = (0,0,-5/2)$ se proyecta sobre el vector normal $(-2,4,-2)$. La norma del vector resultante es la distancia, la cual es $5/\sqrt{24}$.

221. a) $\sqrt{14}/14$.

223. $(6n, 2n, n)$.

224. $(4n + 100, 2n + 50, n + 40)$

225. $(6,2,1), (1/20)(6,2,1)$

228. a) $D = k(3,3,3)$, $k = 1/7$; $\vec{X} = (1,2,3) + \lambda(1/7)(3,3,3)$; $\vec{X}(14) = (7,8,9)$.

230. a) Inconsistente: los tres planos tienen intersección vacía.

b) La intersección de los 3 planos es un único punto: $(0, 9, 7/5)$

c) La intersección de los 3 planos es una línea, como en las hojas de un libro. El vector director de la línea es (11,-7/10,1) y ella pasa por (1,-17/10,0).

231. b) Tenemos 2 líneas paralelas en $\mathbb{R}^2$, y dos planos paralelos en $\mathbb{R}^3$. Las dos líneas (y los dos planos) casi coinciden. Pero su intersección es vacía. Pero en un mundo donde predomine la incertidumbre, ellos parecerían iguales.

232. Tenemos u incógnitas restringidas por i ecuaciones: nos quedamos con $u - i$ grados de libertad: $u = i + f$.

233. Tenemos 4 puntos, podemos tomar un polinomio de tercer grado, con 4 coeficientes a determinar. De esa forma obtenemos un sistema 4×4. El polinomio es: $-2075 - 155{,}5x + 4{,}25x^2 + 0{,}11x^3$.

Capítulo 3

240. a) 1, b)-1, c)-1, d) 1

279. f) Los vectores son coplanares.

274. LI: a, c, g, i, j. Los vectores representan especies diferentes.

287. Sub-EV: b, c.

305. 4; $n + 1$

320. a) Tres planos que se intersecan en un único punto: $(5/6, 1/3, -1/6)$.

b) Un punto, $(46/10, -12/5, 1)$

c) Tres planos que son como las hojas de un libro. Un grado de libertad.

d) Todas las expresiones representan el mismo plano: hay dos grados de libertad.

Capítulo 5

331. a)

$$[T] = \begin{pmatrix} 8 & -1 \\ 2 & 3 \end{pmatrix}$$

$T(x, y) = (8x - y, 2x + 3y)$

T: $\mathbb{R}^2 \to \mathbb{R}^2$

332. a) $T(x, y) = (3x - 2y, -x + 3y, 4y)$

342.

$$T = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}$$

351. Un sistema generado por una TL inyectiva puede carecer de soluciones, pero cuando la tiene, es única. Si una TL es sobre, el sistema tiene solución, pero la unicidad no se garantiza. Una TL siempre envía el cero sobre cero, por lo que cero tiene una preimagen no nula, la TL es automáticamente no 1-1.

355.

$$ST = \begin{pmatrix} 16 & -1 \\ 17 & 55 \end{pmatrix}$$

Ni la multiplicación de matrices, ni la composición de TL son conmutativas, esto implica que el orden es importante.

357. Una entrada (i, j) del producto matricial AB es simplemente el producto punto entre la fila i de A y la columna j de B.

360. a)

$$AB = \begin{pmatrix} -4 & 4 \\ -1 & 32 \end{pmatrix}$$

b)

$$BA = \begin{pmatrix} -22 & 11 \\ 16 & -14 \end{pmatrix}$$

c)

$$CA = \begin{pmatrix} -29 & 13 \\ 11 & 0 \end{pmatrix}$$

d)

$$CB = \begin{pmatrix} -8 & -28 \\ -1 & 8 \end{pmatrix}$$

e)

$$ABC = \begin{pmatrix} -18 & 25 \\ 100 & -27 \end{pmatrix}$$

Capítulo 6

379. a) Solución única, $x = 17/2$

b) $y = x = 0$.

c) Sistema inconsistente.

d) la solución es la línea $2x + 3y = 0$

388. El Kernel está generado por $(1,-1/3)$, mientras que la imagen está generada por $(1,2)$. No son perpendiculares mutuamente, pero para matrices simétricas eso siempre funciona. La imagen es generada por $(1,2)$. La imagen es igual al Kernel. Entonces $M^2=0$.

b) $DetM=0$, La imagen es generada por $(1,4)$, el Kernel por $(1,1)$.

c) $DetM=0$, la imagen es generada por $\{(-1,2,1),(4,-1,3)\}$, el Kernel por $(18/7,1/7,1)$.

d) $detM=0$, la imagen es generada por $(1,2,3)$, el Kernel por $\{(-5/2,0,1),(-2,1,0)\}$.

Capítulo 7

400. Ejecute la operación sobre la identidad. El resultado, en el caso 3×3, es:

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

408. a)

$$\begin{pmatrix} -7 & 4 \\ 2 & 1 \end{pmatrix}$$

b)

$$(1/30)\begin{pmatrix} -1 & 4 \\ 4 & -1 \end{pmatrix}$$

c)

$$(1/3)\begin{pmatrix} -1 & 4 \\ 4 & -1 \end{pmatrix}$$

d)

$$\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

409. Las descomposiciones no son únicas: para a) presentamos dos. La matriz inversa aparece abajo a la derecha:

$$\left(\begin{array}{rrcrr} & & \vdots & 1 & 0 \\ & & \vdots & -2 & 1 \\ \cdots & \cdots & \cdots & .. & .. \\ 1 & 4 & \vdots & -7 & 4 \\ 0 & 1 & \vdots & -2 & 1 \\ \cdots & \cdots & \cdots & .. & .. \\ 1 & 0 & \vdots & -7 & 4 \\ 0 & -1 & \vdots & 2 & -1 \end{array}\right)$$

$$\left(\begin{array}{rrcrr} & & \vdots & 1 & 0 \\ & & \vdots & 2 & -1 \\ \cdots & \cdots & \cdots & .. & .. \\ 1 & -4 & \vdots & -7 & 4 \\ 0 & 1 & \vdots & 2 & -1 \end{array}\right)$$

d)

$$\left(\begin{array}{rrcrr} & & \vdots & \cos\theta & \operatorname{sen}\theta \\ & & \vdots & 0 & 1 \\ \cdots & \cdots & \cdots & .. & .. \\ 1 & 0 & \vdots & \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -1 & \vdots & \operatorname{sen}\theta\cos\theta & -\cos^2\theta \\ \cdots & \cdots & \cdots & .. & .. \\ 1 & 0 & \vdots & \cos\theta & \operatorname{sen}\theta \\ 0 & -1/\cos\theta & \vdots & \operatorname{sen}\theta & \cos\theta \end{array}\right)$$

416. a) $detM=0$

b)

$$1/3\begin{pmatrix} -3 & 6 & 0 \\ 2 & -2 & 1 \\ -6 & -3 & 0 \end{pmatrix}$$

c)

$$-1/5\begin{pmatrix} -5 & 0 & 0 \\ 9 & -1 & 2 \\ 27 & -1 & 3 \end{pmatrix}$$

d)

$$-1\begin{pmatrix} 2 & -1 & 0 \\ 3 & -1 & 0 \\ 1 & -3 & 1 \end{pmatrix}$$

Capítulo 8

434. a)

$$\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}; \begin{pmatrix} 16 \\ -5 \end{pmatrix}$$

b)

$$\begin{pmatrix} -1 & -3 \\ 1 & -4 \end{pmatrix}; \begin{pmatrix} -23 \\ -26 \end{pmatrix}$$

c)

$$\begin{pmatrix} 2 & -1 \\ -2 & -5 \end{pmatrix}; \begin{pmatrix} -3 \\ -39 \end{pmatrix}$$

436. a) $(-82/3, 23/3)$
b) $(-37/5, 23/5)$
c) $(-14/5, -23/5)$
439. a)

$$\begin{pmatrix} 1/3 & 2/3 \\ 1/3 & -1/3 \end{pmatrix}; \begin{pmatrix} 11/3 \\ -5/3 \end{pmatrix}$$

b)

$$\begin{pmatrix} -4/7 & 3/7 \\ -1/7 & -1/7 \end{pmatrix}; \begin{pmatrix} 13/7 \\ -9/7 \end{pmatrix}$$

c)

$$\begin{pmatrix} 5/12 & -1/12 \\ -1/6 & -1/6 \end{pmatrix}; \begin{pmatrix} 1/4 \\ -3/2 \end{pmatrix}$$

441. a)

$$\begin{pmatrix} 1 & -7/18 \\ 0 & 5/9 \end{pmatrix}$$

b)

$$\begin{pmatrix} 3/2 & -1 \\ 5/2 & 0 \end{pmatrix}$$

442. a) $\{(5,2), (-1,3)\}$
b) $\{(1,1), (2,5)\}$

446. $I_{B_2}^{B_1} = I_{B_2}^{N} I_N^{B_1} = (B_2)^{-1} B_1$. Por lo tanto, para encontrar $I_{B_2}^{B_1}$ uno empieza aumentando B_2 del lado derecho con B_1. La reducción produce la matriz solicitada.

447. a)

$$\begin{pmatrix} -1/7 & -11/7 \\ -2/7 & -1/7 \end{pmatrix}; \begin{pmatrix} -3 \\ 0 \end{pmatrix}; \begin{pmatrix} -52/7 \\ 1/7 \end{pmatrix}$$

b)

$$\begin{pmatrix} -1/2 & -25/12 \\ 0 & 7/6 \end{pmatrix}; \begin{pmatrix} -11/3 \\ 7/3 \end{pmatrix}; \begin{pmatrix} 589/60 \\ 35/6 \end{pmatrix}$$

448.

$$\begin{pmatrix} 5/3 & 8 & 3 \\ 0 & -2 & 0 \\ -1/3 & 5/3 & 5/3 \end{pmatrix}$$

452. a) $B_1 = B_2 I_{B_2}^{B_1}$
b) $B_2 = B_1 I_{B_1}^{B_2}$

455. Una rotación no cambia volúmenes: su determinante es 1.

496. a) El Kernel es la línea $y = -x/2$ y la imagen es la línea ortogonal $y = 2x$.

500. a) verdadero: toda rotación conserva el producto punto, pues conserva normas y ángulos.

b) La siguiente matriz no es una rotación, es una reflexión pero también conserva el producto punto:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

504. La matriz asociada a la hipérbola $x^2 - y^2 = 1$ es

$$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = M_N^N$$

$M_B^B = I_B^N M_N^N I_B^N = BMB^T$ donde B es una base ortonormal

$$B = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

.

Po tanto,

$$M_B^B = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}$$

cuya respectiva ecuación es $xy = 1$.

Capítulo 9

507. a)

$$T_N^N = \begin{pmatrix} -1 & 3/7 \\ 17/7 & 9/7 \end{pmatrix}$$

b)

$$T_N^N = \begin{pmatrix} 1 & 1/5 \\ 0 & 3/5 \end{pmatrix}$$

c)

$$T_N^N = \begin{pmatrix} -7/38 & 41/38 \\ 29/19 & -26/19 \end{pmatrix}$$

512. a)

$$T_N^B = \begin{pmatrix} -6 & -10 \\ 5 & 1 \end{pmatrix}$$

519.

$$\begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$$

Capítulo 10

526. c) $R_B^N = (I)_B^N (R)_N^N$ significa que si uno sabe la matriz de una TL dada con respecto a la base natural en ambos lados, entonces para obtener la matriz de la TL con respecto a la base natural en el dominio y a B en el codominio, uno necesita un traductor a la izquierda que cambie de coordenadas de N a B.

528. a) $T(x, y) = (5x/7 + y/7, x/7 + 10y/7)$

b) $T(x, y) = (y, 16x/7 - 5y/7)$

c) $T(x, y) = (5x/4+3y/4, x/4+11y/4)$

530. a)

$$T_N^N = \begin{pmatrix} 3 & -5 \\ 4 & 4 \end{pmatrix}$$

Su determinante es 32 y permanece invariante si cambiamos a la base B:

$$T_B^B = \begin{pmatrix} 3 & -4 \\ 5 & 4 \end{pmatrix}$$

541. a) R es una reflexión y una base natural para R podría ser $B = \{(1, 1), (-1, 1)\}$ y la matriz de R es:

$$R_B^B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$R = R_N^N = I_N^B R_B^B I_B^N = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

tenemos $R(x, y) = (-y, -x)$. R es su propia inversa: $R^{-1}(x, y) = (-y, x)$. Sea P la parábola original $y = x^2$ y sea Q la parábola reflejada. Un punto (x, y) estará en Q ssi $R^{-1}(x, y) \in P$ ssi $(-y, -x)$ satisface $y = x^2$, i.e. $-x = (-y)^2$ o $y = \sqrt{-x}$ para $x \leq 0$.

547. $\vec{r}(t) = (0, \cos t, \operatorname{sen} t)$

548. $\vec{r}(t) = (3 \cos t, 5 \operatorname{sen} t, 0)$

550. El vector normal al plano es $(1, 1, 1)$, el vector de traslación. La trayectoria resultante es $(r(t))_N + (1, 1, 1)$

551. La matriz de rotación de R fue encontrada en el ejemplo 544 y ejercicio 545. La matriz asociada al elipsoide E es

$$E = \begin{pmatrix} 1/4 & 0 & 0 \\ 0 & 1/9 & 0 \\ 0 & 0 & 1/16 \end{pmatrix}$$

sea F el elipsoide rotado. Un punto (x, y, z) estará en F ssi $R^{-1}(x, y, z)$ satisface la ecuación E.

Capítulo 11

556. Usemos la notación $\langle , \rangle$ para el producto punto: $\langle O^T DOx, y\rangle = \langle DOx, Oy\rangle = \langle Ox, DOy\rangle = \langle x, O^T DOy\rangle$.

557. a)

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 4/3 & 2/3 \\ 0 & 1/3 & 5/3 \end{pmatrix}$$

b)

$$\begin{pmatrix} 0 & 1 & -1 \\ 2/3 & 2/3 & 4/3 \\ -1/3 & 2/3 & 4/3 \end{pmatrix}$$

c)

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/3 & -4/3 \\ 0 & -2/3 & -1/3 \end{pmatrix}$$

d)

$$\begin{pmatrix} -3/2 & -5/2 & 5/2 \\ -5/3 & -2 & -1 \\ 5/6 & -1/2 & 5/2 \end{pmatrix}$$

e)

$$\begin{pmatrix} -3 & -1 & 1 \\ -2/3 & -5/3 & 8/3 \\ 1/3 & 4/3 & -1/3 \end{pmatrix}$$

558. $D(a, b, ..., c)D(e, f, ..., g) = D(ae, bf, ..., cg) = D(e, f, ..., g)D(a, b, ..., c)$

570. a) No invertible, dim Ker$= 1$, dim Imagen$= 1$

573. Eigen pares: $(0, (1, 1/3)), (10, (1/3, 1))$

575. En $\mathbb{R}^3$ todas las rotaciones tienen un eje de rotación, por lo que en esa dirección no hay transformación y el valor propio correspondiente es 1.

585. a)

$$D_N^N = \begin{pmatrix} -1/2 & 3/2 \\ 3/2 & -1/2 \end{pmatrix}$$

586. El único valor propio de todas las matrices es 1, pero no hay suficientes vectores propios para formar una base.

587. Una matriz siempre tiene suficientes valores propios, pero le pueden faltar vectores propios, y por eso muchas matrices no son diagonalizables.

594. a) $1, 1, 3$.

b) $1, 2, 3$.

c) -1, $\sqrt{2}$, $-\sqrt{2}$.

d)2, $1/4$, $1/3$

e) 2, $1/4$, $1/3$

595.

Todas las matrices tienen la misma base propia:

$$B = \begin{pmatrix} -1 & 1 & 2 \\ 0 & 1 & -3 \\ 1 & 0 & 1 \end{pmatrix}$$

Las formas diagonales son:

a) D$(9, 6, 3)$

b) D$(9, 6, -3)$

c) D$(9, -6, -3)$

d) D$(3, -2, 1)$

e) D$(-1, -2, 1)$

596. Sea $M(e_i) = \alpha_i e_i$ y $N(e_i) = \beta_i e_i$ y que $x = \sum \lambda_i e_i$. Entonces

$MN(x) = \sum \lambda_i \alpha_i \beta_i e_i = \sum \lambda_i \beta_i \alpha_i e_i = NM(x)$.

602. a) Forma diagonal $D(2, -3\sqrt{2})$. base propia: $\{(2, 1-\sqrt{2}), (1-\sqrt{2}, -2)\}$

b) $D(-\sqrt{26}, -3\sqrt{26}$

c) $D(5, -7)$. Base propia: $\{(1, 1), (-1, 1)\}$

d) $D(5, 0)$. Base propia: $\{(1/2, 1), (-2, 1)\}$

e) $D(2, 0)$, Base propia $\{(1, 1), (-1, 1)\}$

606. Las matrices diagonales y las cónicas están dadas por:

a)D$(2, 3)$, una elipse

b)D$(-1, 3)$, una hipérbole

c)$(1, -1)$, una hipérbole

d) $(-1, -5)$, (elipse).

e) $(-1, -3)$, (elipse).

f) Los productos son conmutativos.

614. a) Si $x^2 + y^2 - z^2 = 0$ entonces $x^2 + y^2 = z^2$. Esto significa que tenemos un círculo cuyo radio aumenta con z. La figura es simétrica con respecto al eje Z. Para ver la silueta, tomamos la traza de la figura con el plano XZ, i.e. cuando $y = 0$. Obtenemos $x^2 = z^2$ o $x = \pm y$. Lo mismo pasa con $x = 0$. Por tanto, tenemos un cono.

b) Cono.

c) Hiperboloide de dos mantos.

d) Hiperboloide de un manto.

e) Hiperboloide de dos mantos.

Capítulo 12

617. La base ortonormal propia es la misma para todos los casos:

$$B = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{3} & 1/\sqrt{6} \\ 0 & -1/\sqrt{3} & -2/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{3} & -1/\sqrt{6} \end{pmatrix}$$

Las formas diagonales y las figuras son:
a) D$(1, 2, 1)$, elipsoide.
b) D$(-1, 2, 1)$, hiperboloide de una hoja.
c) D$(1, -1, 1)$, hiperboloide de una hoja.
d) D$(-2, 1, -4)$, hiperboloide de dos hojas.
e) D$(-1, -1, 1)$, hiperboloide de dos hojas.

618. a) $x^2 + y^2/4 + z^2/9 = 1$
b) $x^2 + y^2/4 + z^2/9 = 1$
c) $x^2 - 4y^2 + 9z^2 = 1$
d) Mal definido.
e) $x^2 - 4y^2 - z^2/9 = 1$

619. a)
$^{(}x - 1)^2 + 3(x - 1)(z - 3) - (z - 3)^2 = 1$

621. a) $x' = x - 1$ y $z' = z + 2$

624. a) Máximo
b) Máximo
c) Mínimo
d) Silla
e) Silla
f) Silla
g) Máximo
h) No se sabe
i) No se sabe

627. a) $z = -2{,}18u^2 + 1{,}82v^2$, silla
b) $-0{,}23u^2 + 0{,}131v^2$, silla
c) $z = -0{,}201u^2 + 3{,}312v^2$, silla
d) $z = 4{,}01u^2 + 0{,}074v^2$, mínimo
e) $z = -1{,}174u^2 + 7{,}67v^2$, silla
f) $z = 3{,}93u^2 + 4{,}07v^2$, mínimo
g) $z = -3{,}08u^2 - 0{,}027v^2$, máximo
h) Desconocida
i) Desconocida

628. Tenemos un punto crítico estable ssi todos los valores propios tienen igual signo. En dos dimensiones, esto es equivalente a decir que el determinante de la matriz diagonal sea positivo. Puesto que el determinante es invariante al cambio de base, el determinante de una matriz diagonal es igual al de otra que le sea similar. Por lo tanto, la estabilidad se decide por el determinante de la matriz original.

637. a) Eigen pares $(4, (1, 1{,}5))$, y $(-1, (-1, 1))$. El primer valor propio domina toda la conducta a largo plazo y define una proporción de 2 a 3.
b) Eigen pares:
$(5{,}82, (1, 1{,}41))$, y $(0{,}171, (-0{,}7, 1))$.
c) Proporciones asintóticas 1 a 1.

BIBLIOGRAFÍA

[1] Altshuller G (2000) *The innovation algorithm: TRIZ, systematic innovation and technical creativity*. Translated, edited and annotated by Lev Shulyak and Steven Rodman. Technical Innovation Center, Inc., Worcester, MA.

[2] Boylestad R, L Nashelsky (1994) *Electrónica de circuitos.* Prentice Hall, México.

[3] Beauchamp J W (Editor) (2007) *Analysis, synthesis, and perception of musical sounds*. Springer, NY.

[4] Burden R, Faires D (1985) *Análisis numérico.* Grupo Editorial Iberoamérica. México.

[5] Feynman R (1998) *Seis piezas fáciles: la física explicada por un genio.* Introducción de Paul Davies. Traducción castellana de Javier García Sanz. Barcelona.

[6] Fisher R (1970) *An introduction to linear algebra*. Dickenson Publishing.

[7] Jobson J D (1991) *Applied Multivariate Data Analysis.* Springer, NY.

[8] Lang S (2002) *Algebra*. Springer-Verlag, NY.

[9] *Gran Enciclopedia Larousse* (1983), Editorial Planeta, Barcelona.

[10] Hans F, Martin A (1984) *Quarks and leptons: an introductory course in modern particle physics.* John Wiley and Sons. NY.

[11] *Mechanics. Berkeley Physics Course-Vol 1.* McGraw-Hill Book Company, NY.

[12] Nash C, S Sen (1983) *Topology and the geometry of physics.* AP.

[13] J Hanan (1989) *Lindenmayer systems, fractals, and plants*. Springer en NY en 1989.

[14] Primas Hans (1983) *Chemistry, quantum mechanics and reductionism.* Springer. Berlin.

[15] Rodríguez José (2008) *Electromagnetismo y geometría.* http://arxiv.org/abs/0806.1492 Citado el 20 de noviembre del 2009.

[16] Snakefly (2009) *How to date a fossil.* http://snakefly.tripod.com/Date.html Citado el 20 de noviembre del 2009.

[17] Steward J (2003) *Calculus and early transcendentals*, 5e. Thomson.

[18] Wikipedia (2009) *Talidomida*, http://es.wikipedia.org/wiki/Talidomida Citado el 20 de noviembre del 2009.

[19] Yosida K (1978) *Functional analysis.* Springer-Verlag. Berlin.

[20] Zill Dennis G (2007) *Ecuaciones diferenciales con aplicaciones de modelado.* Traducción Francisco Sánchez Fragoso, 8^a ed. International Thomson Editores, México.

[21] Zill Dennis G, M Cullen (2008) *Matemáticas avanzadas para ingeniería.* Traducción Emilio Sordo Zabay, 3^a ed. McGrawHill. México.

[22] Zwiebach B (2004) *A first course in string theory.* Cambridge University Press.

ÍNDICE ALFABÉTICO