Compressed Sensing of data with a known distribution

Mateo Díaz

Joint work with Mauricio Junca, Felipe Rincón and Mauricio Velasco

Universidad de los Andes
Colombia

May 24, 2016

Introduction: What is Compressed Sensing?

What is Compressed Sensing?

We study the following problem:

Problem

The sensing problem consists on trying to recover a signal $\mathbf{x}_0 \in \mathbb{R}^d$ *from* m *linear measurements encoded in a vector* $\mathbf{y}_0 := \mathbf{A}\mathbf{x}_0$*, where* $\mathbf{A}$ *is a given* $m \times d$ *matrix with* $m < d$*.*

What is Compressed Sensing?

We study the following problem:

Problem

The sensing problem consists on trying to recover a signal $\mathbf{x}_0 \in \mathbb{R}^d$ *from* m *linear measurements encoded in a vector* $\mathbf{y}_0 := \mathbf{A}\mathbf{x}_0$, *where* $\mathbf{A}$ *is a given* $m \times d$ *matrix with* $m < d$.

$$\begin{array}{ccccc} \mathbb{R}^d & \xrightarrow{\mathbf{A}} & \mathbb{R}^m & \xrightarrow{\Delta} & \mathbb{R}^d \\ \mathbf{x}_0 & \longmapsto & \mathbf{y}_0 & \longmapsto & \mathbf{x}_0 \end{array}$$

What is Compressed Sensing?

We study the following problem:

Problem

The sensing problem consists on trying to recover a signal $\mathbf{x}_0 \in \mathbb{R}^d$ *from* m *linear measurements encoded in a vector* $\mathbf{y}_0 := \mathbf{A}\mathbf{x}_0$*, where* $\mathbf{A}$ *is a given* $m \times d$ *matrix with* $m < d$*.*

$$\begin{array}{ccccc} \mathbb{R}^d & \xrightarrow{\mathbf{A}} & \mathbb{R}^m & \xrightarrow{\Delta} & \mathbb{R}^d \\ \mathbf{x}_0 & \longmapsto & \mathbf{y}_0 & \longmapsto & \mathbf{x}_0 \end{array}$$

Since we are collapsing many dimensions, we cannot expect this to work for all signals. However, it should work for a suitable subset.

What operators work for CS?

Such problems can be solved efficiently if the original $\mathbf{x}_0$ is sufficiently sparse; this result was first shown in the seminal work by Candes, Romberg and Tao [CRT06]. They proposed

For compressing

Let $m < d$ the number of measurements and $\mathbf{A} \in \mathbb{R}^{m \times d}$ a random matrix with iid entries $\mathrm{N}(0, 1)$.

What operators work for CS?

Such problems can be solved efficiently if the original $\mathbf{x}_0$ is sufficiently sparse; this result was first shown in the seminal work by Candes, Romberg and Tao [CRT06]. They proposed

For compressing

Let $m < d$ the number of measurements and $\mathbf{A} \in \mathbb{R}^{m\times d}$ a random matrix with iid entries $\mathrm{N}(0,1)$.

For recovery

Given the vector $\mathbf{y}_0 = \mathbf{A}\mathbf{x}_0$ we may attempt to recover the original signal by solving the linear program

$$\Delta(\mathbf{y}_0) = \mathrm{argmin}(\|\mathbf{x}\|_1 : \mathbf{A}\mathbf{x} = \mathbf{y}_0). \tag{P}$$

How good are they?

We say that the method is *successful for* $\mathbf{A}$ *and* $\mathbf{x}_0$ if the program (P) has a unique solution and this solution is $\mathbf{x}_0$.

How good are they?

We say that the method is *successful for* $\mathbf{A}$ *and* $\mathbf{x}_0$ if the program (P) has a unique solution and this solution is $\mathbf{x}_0$.

Definition (Support)

The support of a vector $\mathbf{x} \in \mathbb{R}^d$ is given by

$$\mathrm{supp}(\mathbf{x}) = \{i : x_i \neq 0\}.$$

Definition

A vector $\mathbf{x} \in \mathbb{R}^d$ is s-sparse if $\#\,\mathrm{supp}(\mathbf{x}) \leq s$.

How good are they?

We say that the method is *successful for* $\mathbf{A}$ *and* $\mathbf{x}_0$ if the program (P) has a unique solution and this solution is $\mathbf{x}_0$.

Definition (Support)

The support of a vector $\mathbf{x} \in \mathbb{R}^d$ is given by

$$\mathrm{supp}(\mathbf{x}) = \{i : x_i \neq 0\}.$$

Definition

A vector $\mathbf{x} \in \mathbb{R}^d$ is s-sparse if $\#\,\mathrm{supp}(\mathbf{x}) \leq s$.

The proposed method is very accurate for sparse signals.

How good are they?

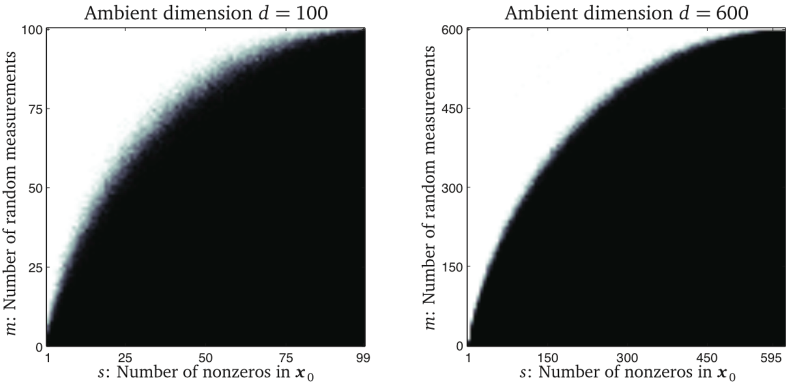


Figure: Recovery frequency, taken from [ALMT14]

The main questions of our research

We consider an extension of the original problem:

The main questions of our research

We consider an extension of the original problem:

Problem

Consider the sensing problem with the additional hypothesis that $\mathsf{x}_0 \in \mathbb{R}^d$ *is drawn from a known distribution* $\mathcal{F}$. ***Is it possible to take advantage of this new information? How to do so?***

The main questions of our research

We consider an extension of the original problem:

Problem

Consider the sensing problem with the additional hypothesis that $\mathbf{x}_0 \in \mathbb{R}^d$ is drawn from a known distribution $\mathcal{F}$. ***Is it possible to take advantage of this new information? How to do so?***

This appears to be a natural extension of the original problems. In real applications we can have an estimation of the distribution $\mathcal{F}$. A good example are Head MRI.

Agenda

Prelims

Natural questions

Recall

$$\Delta(y_0) = \operatorname{argmin}(\|x\|_1 : Ax = y_0). \tag{P}$$

Natural questions

Recall

$$\Delta(y_0) = \text{argmin}(\|x\|_1 : Ax = y_0). \tag{P}$$

Q1 For a given s, what is the probability of obtaining an $\mathbf{A}$ for which (P) successfully recovers all s-sparse vectors?

Q2 Can we formalize the phase transition of the probability of success?

A characterization for a successful problem

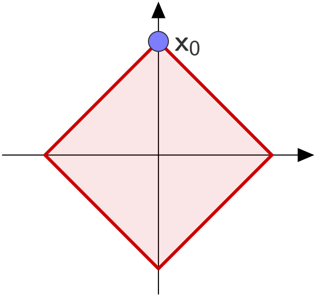
Definition (Descent Cone)

For a point $x_0 \in \mathbb{R}^d$ and $f : \mathbb{R}^d \to \mathbb{R}$ a convex function let $D(f, x)$ be the descent cone of the function f at x_0, given by $D(f, x_0) := \text{Cone}\{x - x_0 : f(x) \leq f(x_0)\}$.

A characterization for a successful problem

Definition (Descent Cone)

For a point $x_0 \in \mathbb{R}^d$ and $f : \mathbb{R}^d \to \mathbb{R}$ a convex function let $D(f, x)$ be the descent cone of the function f at x_0, given by $D(f, x_0) := \text{Cone}\{x - x_0 : f(x) \leq f(x_0)\}$.



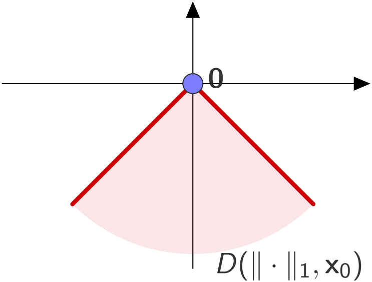


Figure: The figure on the left is the ℓ_1-ball and the one on the right is the descent cone generated by the ℓ_1-norm and $\mathbf{e}_1$.

A characterization for a successful problem

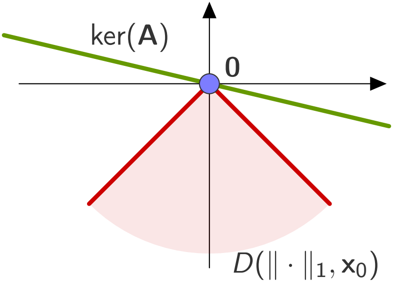
Theorem ([CRPW12])

The compressed sensing method (P) *is successful for* $\mathbf{A}$ *and* $\mathbf{x}_0$ *if and only if* $D(\|\cdot\|_1, x_0) \cap ker(\mathbf{A}) = \{0\}$.

A characterization for a successful problem

Theorem ([CRPW12])

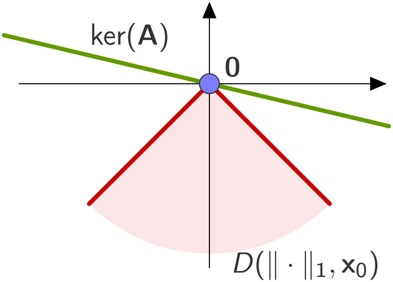
The compressed sensing method (P) *is successful for* $\mathbf{A}$ *and* $\mathbf{x}_0$ *if and only if* $D(\|\cdot\|_1, x_0) \cap \ker(\mathbf{A}) = \{0\}$.



A characterization for a successful problem

Theorem ([CRPW12])

The compressed sensing method (P) *is successful for* $\mathbf{A}$ *and* $\mathbf{x}_0$ *if and only if* $D(\|\cdot\|_1, x_0) \cap ker(\mathbf{A}) = \{0\}$.



Therefore, the probability of success is equivalent to the probability of a random subspace not intersecting a cone.

Probability of success

Wait, random vector subspaces? Is there a way to measure them?

Probability of success

Wait, random vector subspaces? Is there a way to measure them?

Yes, it turns out that if $\mathbf{A} \in \mathbb{R}^{m \times d}$ is a random matrix with iid entries $\mathrm{N}(0, 1)$ then $\ker(\mathbf{A})$ is uniformly distributed over the Grassmannian $\mathrm{Gr}(d - m, \mathbb{R}^d)$.

Probability of success

Wait, random vector subspaces? Is there a way to measure them?

Yes, it turns out that if $\mathbf{A} \in \mathbb{R}^{m \times d}$ is a random matrix with iid entries $\mathrm{N}(0,1)$ then $\ker(\mathbf{A})$ is uniformly distributed over the Grassmannian $\mathrm{Gr}(d-m, \mathbb{R}^d)$. It follows that if K is a fixed subspace of dimension $d-m$ and $\mathbf{Q} \in \mathrm{O}(d)$, chosen with the Haar measure on $\mathrm{O}(d)$, then,

$$\mathbb{P}_{\mathbf{A}}\{(\mathrm{P}) \text{ is successful for } \mathbf{x}_0 \text{ and } \mathbf{A}\} = \mathbb{P}_{\mathbf{Q}}\left\{D(\|\cdot\|_1, \mathbf{x}_0) \cap \mathbf{Q}K = \{\mathbf{0}\}\right\}.$$

Probability of success

Wait, random vector subspaces? Is there a way to measure them?

Yes, it turns out that if $\mathbf{A} \in \mathbb{R}^{m \times d}$ is a random matrix with iid entries $\mathrm{N}(0, 1)$ then $\ker(\mathbf{A})$ is uniformly distributed over the Grassmannian $\mathrm{Gr}(d - m, \mathbb{R}^d)$. It follows that if K is a fixed subspace of dimension $d - m$ and $\mathbf{Q} \in \mathrm{O}(d)$, chosen with the Haar measure on $\mathrm{O}(d)$, then,

$$\mathbb{P}_{\mathbf{A}}\{(\mathrm{P}) \text{ is successful for } \mathbf{x}_0 \text{ and } \mathbf{A}\} = \mathbb{P}_{\mathbf{Q}}\left\{D(\|\cdot\|_1, \mathbf{x}_0) \cap \mathbf{Q}K = \{\mathbf{0}\}\right\}.$$

But, is there a formula to this probability?

Intrinsic volumes

A recent paper, [ALMT14], gives answers to this questions relating them with the intrinsic volumes of the descent cone $D(\|\cdot\|_1, \mathbf{x}_0)$.

Definition (Projection $\pi_C(x)$)

Let $C \subseteq \mathbb{R}^d$ be any closed convex set, define $\pi_C(x) = \text{argmin}(\|x - y\|_2 : y \in C)$.

Intrinsic volumes

A recent paper, [ALMT14], gives answers to this questions relating them with the intrinsic volumes of the descent cone $D(\|\cdot\|_1, \mathbf{x}_0)$.

Definition (Projection $\pi_C(x)$)

Let $C \subseteq \mathbb{R}^d$ be any closed convex set, define $\pi_C(x) = \text{argmin}(\|x - y\|_2 : y \in C)$.

Definition (Intrinsic Volumes)

Let C a polyhedral cone in $\mathbb{R}^d$. For each $0 \leq k \leq d$, define the kth intrinsic volume $\nu_k(C)$ is given by

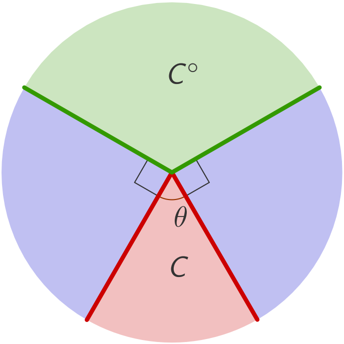
$$\nu_k(C) := \mathbb{P}(\pi_C(g) \text{ lies in the interior of a } k\text{-dimensional face of } C),$$

where g is a standard normal vector in $\mathbb{R}^d$.

Example

Consider a convex cone $C \subseteq \mathbb{R}^2$. The standard normal distribution is symmetric with respect to $\mathbf{0}$. Therefore,

$$\nu_2(C) = \theta/2\pi, \quad \nu_1(C) = 1/2, \quad \nu_0(C) = (\pi - \theta)/2\pi.$$



Tail functionals

Definition

Let $C \subset \mathbb{R}^d$ a closed convex cone. For each $s \in \{0, 1, \ldots, d\}$, the sth tail functional is defined as

$$t_s(C) := \sum_{j=s}^{d} \nu_j(C).$$

Similarly, the sth half-tail functional is defined as

$$h_s(C) := \sum_{\substack{j=s \\ j-s \text{ even}}}^{d} \nu_j(C).$$

Probability of success in terms of the intrinsic volumes

It appears that the intrinsic volumes encode a lot of statistical information about the cone.

Probability of success in terms of the intrinsic volumes

It appears that the intrinsic volumes encode a lot of statistical information about the cone.

Theorem (Kinematic formula [SW08])

Let $C \subset \mathbb{R}^d$ a closed convex cone and $L \subset \mathbb{R}^d$ a linear subspace with dimension $d - m$, then

$$\mathbb{P}_{\mathbf{Q}}\left\{C \cap \mathbf{Q}L = \{\mathbf{0}\}\right\} = 1 - 2h_{m+1}(C).$$

Probability of success in terms of the intrinsic volumes

It appears that the intrinsic volumes encode a lot of statistical information about the cone.

Theorem (Kinematic formula [SW08])

Let $C \subset \mathbb{R}^d$ a closed convex cone and $L \subset \mathbb{R}^d$ a linear subspace with dimension $d-m$, then

$$\mathbb{P}_{\mathbf{Q}}\left\{C \cap \mathbf{Q}L = \{\mathbf{0}\}\right\} = 1 - 2h_{m+1}(C).$$

Remark

For each closed convex cone $C \in \mathbb{R}^d$ which is not a linear subspace

$$2h_s(C) \geq t_s(C) \geq 2h_{s+1}(C) \qquad \text{for } s = 0, 1, 2, \cdots, d-1.$$

Thus, $1 - t_m(D(\|\cdot\|_1, \mathbf{x}_0))$ is very close to the probability of perfect recovery.

Statistical dimension

Definition (Statistical dimension)

Let $C \subseteq \mathbb{R}^d$ a polyhedral cone, define the statistical dimension $\delta(C)$ as

$$\delta(C) := \sum_{k=0}^{d} k\nu_k(C).$$

Phase transition and the Statistical Dimension

It appears that the statistical dimension perfectly describes the inflection point of the phase transition.

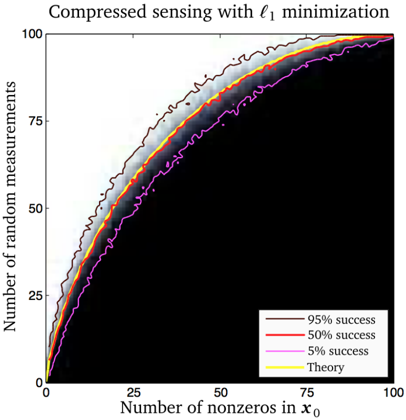


Figure: Recovery frequency with the statistical dimension, taken from [ALMT14]

The problem

Remark

We do not know how to compute the intrinsic volumes for descent cones :(

Weighted Compressed Sensing

Main idea

We may use a weighted ℓ^1 norm to deform the descent cone in a suitable way.

Main idea

We may use a weighted ℓ^1 norm to deform the descent cone in a suitable way.

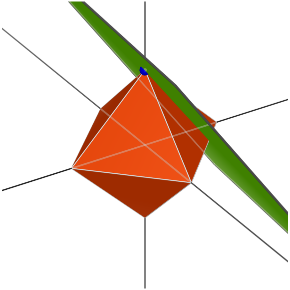


Figure: The blue points are $\mathbf{x}'$, the polytopes are the ℓ_1 and $\ell_1^{\mathbf{w}}$ balls and the green planes are the $\ker(\mathbf{A}) + \mathbf{x}'$.

Weighted CS

Definition

(ℓ^1_w norm) For a given $w = (w_1, \ldots, w_d) \in \mathbb{R}^d_+$ vector of weights, the ℓ^1_w norm of a vector x is given by

$$\|x\|_1^w = \sum_{i=1}^{d} w_i |x_i|.$$

Weighted CS

Definition

(ℓ_w^1 norm) For a given $w = (w_1, \ldots, w_d) \in \mathbb{R}_+^d$ vector of weights, the ℓ_w^1 norm of a vector x is given by

$$\|x\|_1^w = \sum_{i=1}^{d} w_i |x_i|.$$

New Algorithm for Recovery

With this in mind, the new algorithm $\Delta_{\mathbf{w}}$ that we propose for recovery X_0 from $AX_0 = y_0$ is

$$\Delta_{\mathbf{w}}(y_0) := \operatorname{argmin}(\|x\|_1^{\mathbf{w}} : Ax = y_0). \qquad (P_{\mathbf{w}})$$

Analogous definitions

Definition (Expected intrinsic volumes)

For a fixed vector $\mathbf{w} \in \mathbb{R}^d_{>0}$ and $\mathbf{X}_0 \sim \mathcal{F}$ a random vector, we define the kth expected intrinsic volume as

$$\bar{\nu}_k(\mathbf{w}) = \mathbb{E}_{\mathbf{X}_0}\left[\nu_k(D\left(\text{supp}(\mathbf{X}_0), \mathbf{w}\right))\right],$$

for $k = 0 \ldots d$.

We define $\bar{t}_k$ and $\bar{h}_k$ as the tail and the half-tail of the expected intrinsic volumes. It is easy to prove that the success probability is given by

$$1 - 2\bar{h}_{m+1}(\mathbf{w}).$$

Expected statistical dimension

With this in mind, we define

Definition

(**Expected statistical dimension**) For a fixed vector $\mathbf{w} \in \mathbb{R}^d_{>0}$ and a $\mathbf{X}_0 \sim \mathcal{F}$ a random vector, the expected statistical dimension is given by

$$\overline{\delta}(\mathbf{w}) := \mathbb{E}_{\mathbf{X}_0}\left[\delta(D(\mathrm{supp}(\mathbf{X}_0), \mathbf{w}))\right]. \quad (1)$$

We propose to choose the weights $\mathbf{w}$ so as to minimize the expected statistical dimension.

How to choose the weights?

Definition

For $I \subset [d]$ define $q_I := \mathbb{P}(supp(X_0) = I)$ and for $j \in [d]$ let $\beta_j := \sum_{I \ni j} q_I$.

How to choose the weights?

Definition

For $I \subset [d]$ define $q_I := \mathbb{P}(supp(X_0) = I)$ and for $j \in [d]$ let $\beta_j := \sum_{I \ni j} q_I$.

Theorem

For any $w \in \mathbb{R}_+^d$ and any $\tau > 0$ the following inequality holds

$$\begin{aligned} \overline{\delta}(\mathbf{w}) \leq \quad & \mathbb{E}_{X_0}|\operatorname{supp}(X_0)| + \textstyle\sum_{j=1}^d \beta_j(\tau w_j)^2 \\ & + \textstyle\sum_{j=1}^d \left((1-\beta_j)\left[\sqrt{\frac{2}{\pi}} \int_{\tau w_j}^{\infty} (u-\tau w_j)^2 e^{-\frac{u^2}{2}} du\right]\right). \end{aligned} \tag{2}$$

Moreover, the right hand side is minimized if $\lambda_i := \tau w_i$ satisfies the equation

$$\lambda_i \frac{\beta_i}{(1-\beta_i)} = \sqrt{\frac{2}{\pi}} \int_{\lambda_i}^{\infty} (u-\lambda_i) e^{-\frac{u^2}{2}} du.$$

Estimating intrinsic volumes

The question

We don't know how to compute the intrinsic volumes... so let's estimate them with Monte Carlo simulations!

The question

We don't know how to compute the intrinsic volumes... so let's estimate them with Monte Carlo simulations!

Question

Given a descent cone C and a point $\mathbf{x}$, what is the dimension of the face containing $\pi_C(\mathbf{x})$?

The main problem here is that the number of possible cells grows exponentially. Fortunately, there is a lot of symmetry in the cells.

The result

Theorem (Intrinsic dimension)

Let $\mathbf{z} = (z_0, z_1, \ldots, z_{d-k}) \in \mathbb{R}^{d-k+1}$*. Consider the permutation* $i_1, i_2, \ldots, i_{d-k}$ *of the numbers* $1, 2, \ldots, d-k$ *that satisfies* $|z_{i_1}|/w_{i_1} \geq |z_{i_2}|/w_{i_2} \geq \cdots \geq |z_{i_{d-k}}|/w_{i_{d-k}}$*. For* $j = 0, 1, \ldots, d-k-1$ *define*

$$b_j := w_{i_1}|z_{i_1}| + \cdots + w_{i_j}|z_{i_j}| - \frac{a^2 + w_{i_1}^2 + \cdots + w_{i_j}^2}{w_{i_{j+1}}}|z_{i_{j+1}}|,$$

and

$$b_{d-k} := w_{i_1}|z_{i_1}| + \cdots + w_{i_{d-k}}|z_{i_{d-k}}|.$$

Then the numbers b_j *satisfy* $b_0 \leq b_1 \leq \cdots \leq b_{d-k}$*, and the projection of the point* $\mathbf{z}$ *onto the cone* $\overline{D}(I, \mathbf{w})$ *lands in the interior of a face of dimension* l*, where* l *is such that* $b_{l-1} < az_0 \leq b_l$ *(where by convention* $b_{-1} = -\infty$ *and* $b_{d-k+1} = \infty$*).*

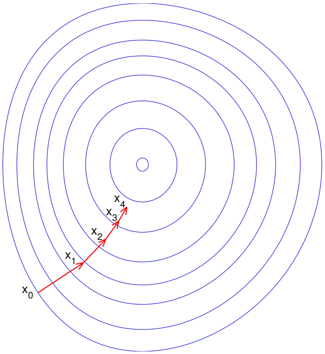
A numerical method to find the weights

What to use?

Our objective is to use a gradient descent algorithm to minimize the expected statistical dimension in terms of the weights.

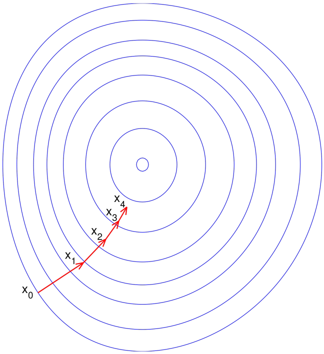
What to use?

Our objective is to use a gradient descent algorithm to minimize the expected statistical dimension in terms of the weights. How does this works?



What to use?

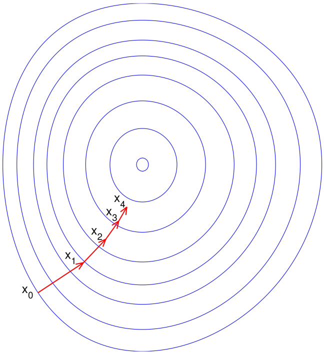
Our objective is to use a gradient descent algorithm to minimize the expected statistical dimension in terms of the weights. How does this works?



$$\mathbf{w}_{k+1} = \mathbf{w}_k - \tau \nabla \bar{\delta}(\mathbf{w_k})$$

What to use?

Our objective is to use a gradient descent algorithm to minimize the expected statistical dimension in terms of the weights. How does this works?



$$\mathbf{w}_{k+1} = \mathbf{w}_k - \tau \nabla \bar{\delta}(\mathbf{w_k})$$

We have a big problem:

- how to compute $\nabla \bar{\delta}$?

We were able to estimate this quantity using Monte Carlo simulations based on similar ideas.

Experimental results

Recovering Juan Valdez

We "trained" our algorithm using the rows of the following binary image to obtain the empirical distribution and then we performed a row-by-row reconstruction. Which means that we solved 244 optimization problems, one for each row.

Figure: Original image with dimensions 750×244

Recovering Juan Valdez

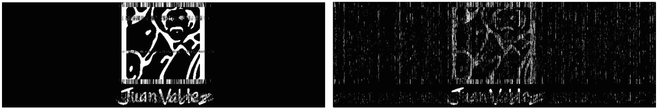


Figure: Recovery with 150 measures (20%)

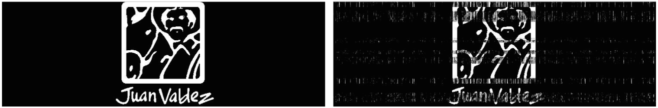


Figure: Recovery with 225 measures (30%)

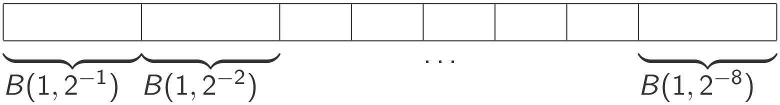
Figure: Recovery with 525 measures (70%)

Vector of Bernoulli Variables

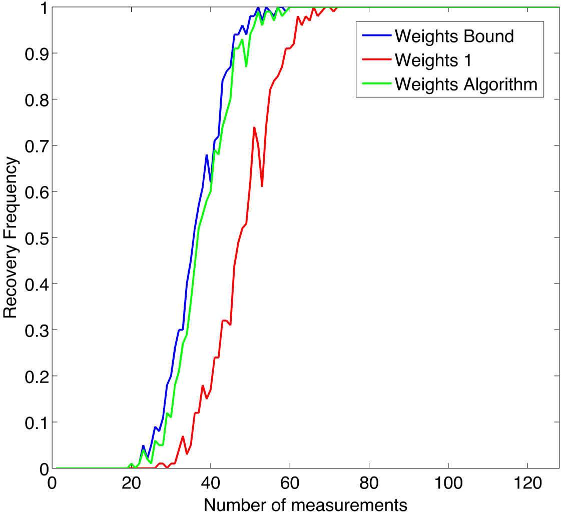
For this experiment we generated random vectors $X_0 \in \mathbb{R}^{128}$. We partioned the entries of X in 8 blocks, in every block the entries were iid random variables with Bernoulli distribution, where the distribution parameter was defined by the block.

Vector of Bernoulli Variables

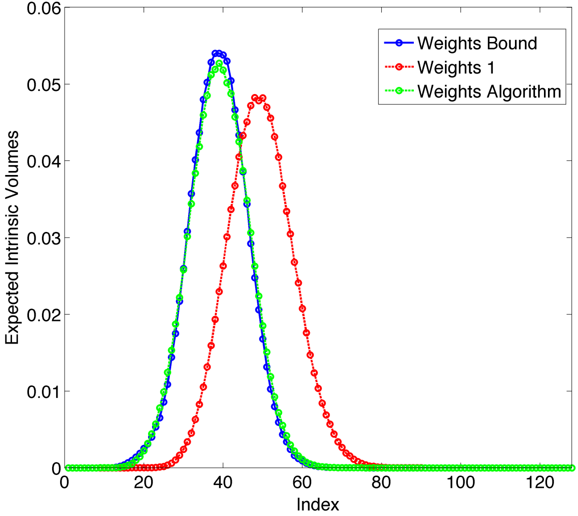
For this experiment we generated random vectors $X_0 \in \mathbb{R}^{128}$. We partioned the entries of X in 8 blocks, in every block the entries were iid random variables with Bernoulli distribution, where the distribution parameter was defined by the block.



Vector of Bernoulli Variables



Vector of Bernoulli Variables



A non-sharp case

For the last experiment we didn't obtain a positive result. In this experiment we sampled a random vector $X_0 \in \mathbb{R}^{128}$ with uniform distribution over the following 4 supports.

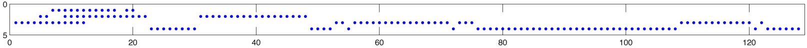
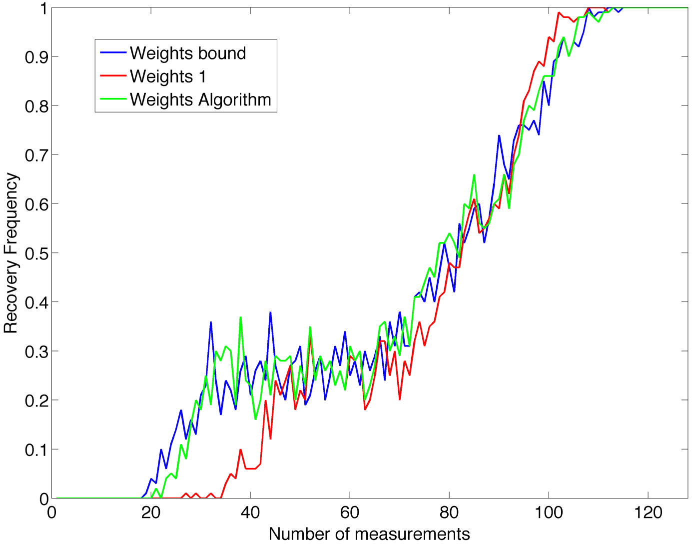


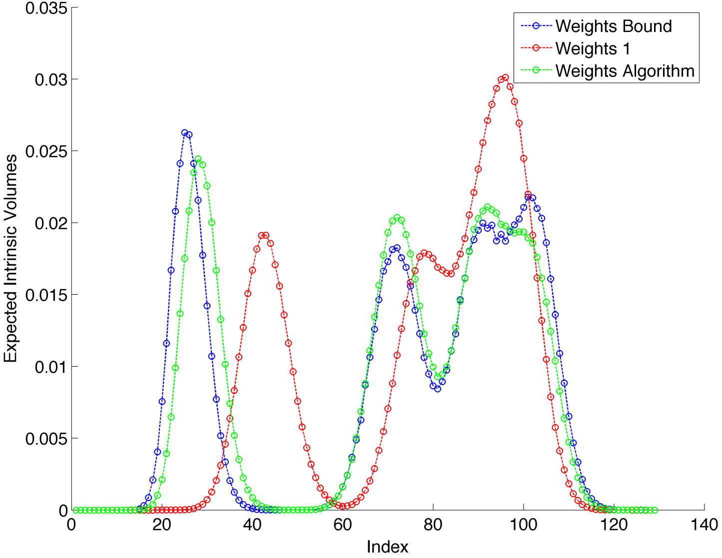
Figure: The four possible signals of the distribution, each row represents one of them. A blue point means 1 and white means 0.

Note that for this case, the number of possible outcomes for the signal was significantly smaller compared with the other experiments.

A non-sharp case

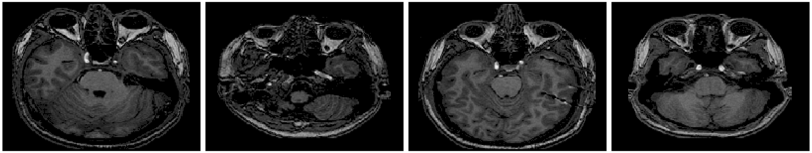


A non-sharp case



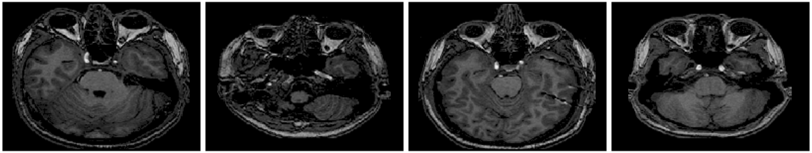
Brain MRI

We took 47 real Brain MRI from 5 patients. We performed a leave-one-out cross-validation to measure the frequency of perfect recovery for several m (number of measures). All the images had size 215×184.



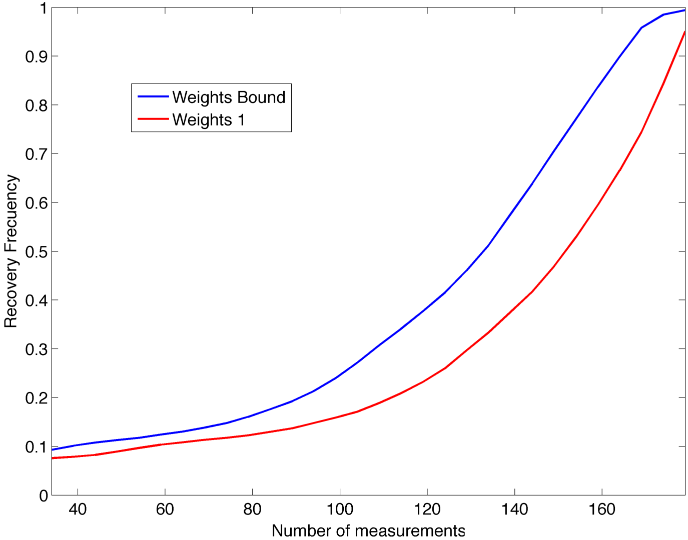
Brain MRI

We took 47 real Brain MRI from 5 patients. We performed a leave-one-out cross-validation to measure the frequency of perfect recovery for several m (number of measures). All the images had size 215×184.



In the experiment we calculated the weights separately for each row.

Brain MRI



Weights Bound
Weights 1
Recovery Frecuency
Number of measurements
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
40
60
80
100
120
140
160

Conclusions and future work

Conclusions and future work

- We showed that it is possible to use statistical information, about the signal we wish to sample, to reduce the number of measures.

Conclusions and future work

- We showed that it is possible to use statistical information, about the signal we wish to sample, to reduce the number of measures.
- Given a distribution of the signal, we proposed a method to do this with weights. We describe two ways to find the weights

Conclusions and future work

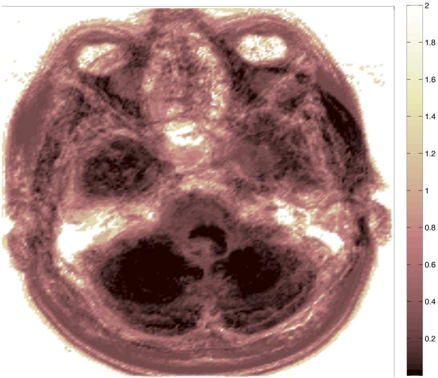
- We showed that it is possible to use statistical information, about the signal we wish to sample, to reduce the number of measures.
- Given a distribution of the signal, we proposed a method to do this with weights. We describe two ways to find the weights
- Our results with the numerical weights suggest that our analytic weights are a local optimum or at least close to one.

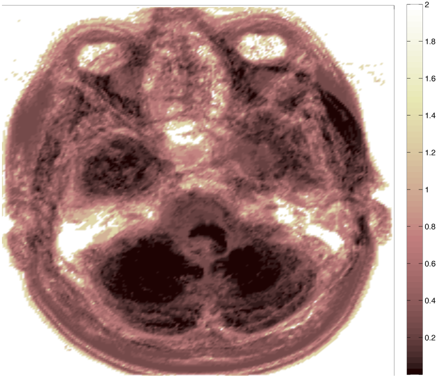
Conclusions and future work

- We showed that it is possible to use statistical information, about the signal we wish to sample, to reduce the number of measures.
- Given a distribution of the signal, we proposed a method to do this with weights. We describe two ways to find the weights
- Our results with the numerical weights suggest that our analytic weights are a local optimum or at least close to one.
- In the future we would like to have a method to find weights when minimize the statistical doesn't work

Conclusions and future work

- We showed that it is possible to use statistical information, about the signal we wish to sample, to reduce the number of measures.
- Given a distribution of the signal, we proposed a method to do this with weights. We describe two ways to find the weights
- Our results with the numerical weights suggest that our analytic weights are a local optimum or at least close to one.
- In the future we would like to have a method to find weights when minimize the statistical doesn't work
- We would like to generalize these ideas for different settings, e.g., matrix completion or matrix demixing.


2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2



Thank you very much!

References

Dennis Amelunxen, Martin Lotz, Michael B. McCoy, and Joel A. Tropp.
Living on the edge: phase transitions in convex programs with random data.
Information and Inference, 2014.

Venkat Chandrasekaran, Benjamin Recht, PabloA. Parrilo, and AlanS. Willsky.
The convex geometry of linear inverse problems.
Foundations of Computational Mathematics, 12(6):805–849, 2012.

Emmanuel J Candès, Justin Romberg, and Terence Tao.
Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information.
Information Theory, IEEE Transactions on, 52(2):489–509, 2006.

Rolf Schneider and Wolfgang Weil.
Stochastic and integral geometry.
Springer Science & Business Media, 2008.