Mathematics of Compressive Sensing: Random matrices and ℓ_1-recovery

Justin Romberg

Georgia Tech, School of ECE

Dutch-Flemish Numerical Analysis Conference

October 7, 2011
Woudschoten, Zeist, Netherlands

Acquisition as linear algebra

samples M

resolution/ bandwidth N

$$\underbrace{\begin{bmatrix} y \end{bmatrix}}_{\text{data}} = \underbrace{\begin{bmatrix} \quad & \Phi & \quad \end{bmatrix}}_{\text{acquisition system}} \underbrace{\begin{bmatrix} x \end{bmatrix}}_{\text{unknown signal/image}}$$

- Small number of samples $=$ underdetermined system
  Impossible to solve in general
- If x is *sparse* and Φ is *diverse*, then these systems can be "inverted"

Agenda

We will prove (almost from top to bottom) two things:

- That an $M \times N$ iid Gaussian random matrix satisfies

$$(1-\delta)\|x\|_2^2 \;\leq\; \|\Phi x\|_2^2 \;\leq\; (1+\delta)\|x\|_2^2 \quad \forall\ 2S\text{-sparse } x \tag{1}$$

  with (extraordinarily) high probability when

$$M \;\geq\; \mathrm{Const} \cdot S \log(N/S)$$

Agenda

We will prove (almost from top to bottom) two things:

- That an $M \times N$ iid Gaussian random matrix satisfies

$$(1-\delta)\|x\|_2^2 \ \le \ \|\Phi x\|_2^2 \ \le \ (1+\delta)\|x\|_2^2 \quad \forall\ 2S\text{-sparse } x \tag{1}$$

  with (extraordinarily) high probability when

$$M \ \ge \ \mathrm{Const} \cdot S \log(N/S)$$

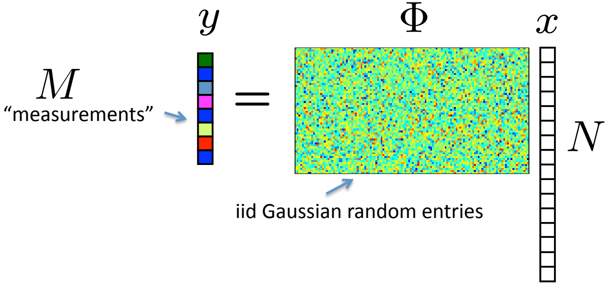
- Suppose an $M \times N$ matrix Φ obeys (1). Let x_0 be an S-sparse vector, and suppose we observe $y = \Phi x_0$. Given y, the solution to

$$\min_x \ \|x\|_{\ell_1} \quad \text{subject to} \quad \Phi x = y$$

  is *exactly* x_0.

Gaussian random matrices

- Each entry of Φ is iid $\mathrm{Normal}(0, M^{-1})$

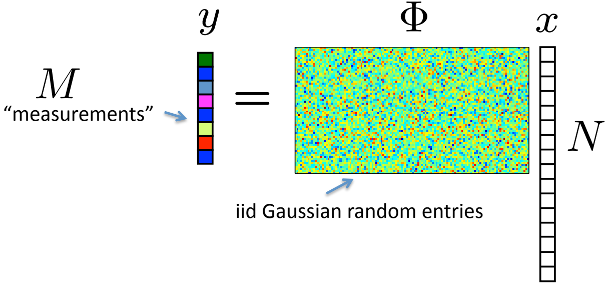


- For *any fixed* $x \in \mathbb{R}^N$, each measurement is

$$y_m \sim \mathrm{Normal}(0, \|x\|_2^2/M)$$

Gaussian random matrices

- Each entry of Φ is iid $\mathrm{Normal}(0, M^{-1})$



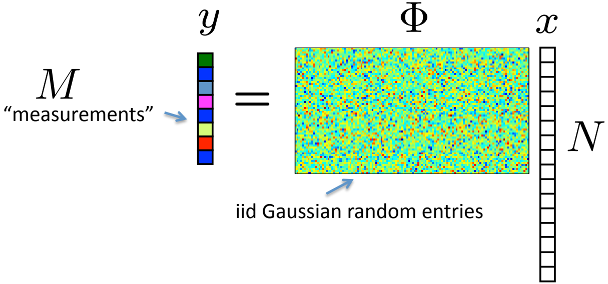
- For *any fixed* $x \in \mathbb{R}^N$, we have

$$\mathrm{E}[\|\Phi x\|_2^2] = \|x\|_2^2$$

the mean of the measurement energy is exactly $\|x\|_2^2$

Gaussian random matrices

- Each entry of Φ is iid $\mathrm{Normal}(0, M^{-1})$

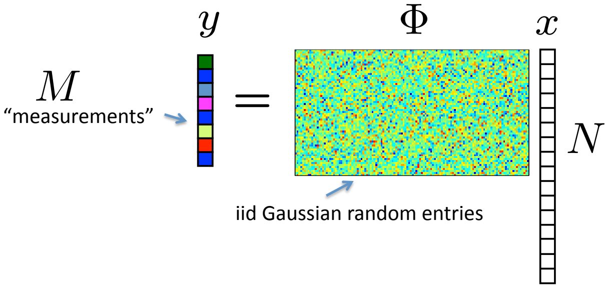


- For *any fixed* $x \in \mathbb{R}^N$, we have

$$\mathrm{P}\left\{ \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| < \delta \|x\|_2^2 \right\} \;\geq\; 1 - 2e^{-M\delta^2/8}$$

Gaussian random matrices

- Each entry of Φ is iid $\mathrm{Normal}(0, M^{-1})$



- For *all $2S$-sparse* $x \in \mathbb{R}^N$, we have

$$\mathrm{P}\left\{\max_x \left|\|\Phi x\|_2^2 - \|x\|_2^2\right| < \delta\|x\|_2^2\right\} \;\geq\; 1 - 2e^{c\cdot S\log(N/S)}e^{-M\delta^2/8}$$

So we can make this probability close to 1 by taking

$$M \;\geq\; \mathrm{Const}\cdot S\,\log(N/S)$$

Random projection of a fixed vector

For Gaussian random Φ operating on a *fixed* $x \in \mathbb{R}^N$

$$\|\Phi x\|_2^2 \approx \|x\|_2^2$$

Theorem: Let Φ be an $M \times N$ matrix whose entries are iid Gaussian

$$\Phi_{i,j} \sim \mathrm{Normal}(0, 1/M).$$

Set $v = \Phi x$. Then

$$\mathrm{E}\,\|v\|_2^2 = \|x\|_2^2,$$

as

$$\mathrm{E}\left[\sum_{m=1}^{M} v_m^2\right] = \sum_{m=1}^{M} \mathrm{E}[v_m^2] = \sum_{m=1}^{M} \frac{1}{M}\|x\|_2^2 = \|x\|_2^2,$$

since $v_m = \langle x, \phi_m \rangle \sim \mathrm{Normal}(0, M^{-1}\|x\|_2^2)$

Random projection of a fixed vector

For Gaussian random Φ operating on a *fixed* $x \in \mathbb{R}^N$

$$\|\Phi x\|_2^2 \approx \|x\|_2^2$$

Theorem: Let Φ be an $M \times N$ matrix whose entries are iid Gaussian

$$\Phi_{i,j} \sim \mathrm{Normal}(0, 1/M).$$

Set $v = \Phi x$. Then

$$\mathrm{E}\,\|v\|_2^2 = \|x\|_2^2,$$

and for any $0 < \delta \leq 1$

$$\begin{aligned}\mathrm{P}\left\{\left|\|v\|_2^2 - \|x\|_2^2\right| > \delta\right\} &\leq 2\exp\left(-\frac{(\delta^2 - \delta^3)M}{4}\right) \\ &\leq 2\exp\left(-\delta^2 M/8\right)\end{aligned}$$

for $\delta \leq 1/2$.

The Markov inequality

Let Y be a positive random variable. Then for all $t > 0$

$$\mathrm{P}\left\{Y \geq t\right\} \;\leq\; \frac{\mathrm{E}[Y]}{t}$$

The Markov inequality

Let Y be a positive random variable. Then for all $t > 0$

$$\mathrm{P}\{Y \geq t\} \;\leq\; \frac{\mathrm{E}[Y]}{t}$$

Proof:

$$\begin{aligned}
\mathrm{E}[Y] &= \int_0^\infty y\, f_Y(y)\; dy \\
&\geq \int_t^\infty y\, f_Y(y)\; dy \\
&\geq t \int_t^\infty f_Y(y)\; dy \\
&= t\,\mathrm{P}\{Y \geq t\}\,.
\end{aligned}$$

The Markov inequality

Let Y be a positive random variable. Then for all $t > 0$

$$\mathrm{P}\left\{Y \geq t\right\} \;\leq\; \frac{\mathrm{E}[Y]}{t}$$

Also:

$$\begin{aligned}
\mathrm{P}\left\{Y^2 \geq t^2\right\} &\leq \frac{\mathrm{E}[Y^2]}{t^2} \\
\mathrm{P}\left\{Y^3 \geq t^3\right\} &\leq \frac{\mathrm{E}[Y^3]}{t^3} \\
\mathrm{P}\left\{e^{\lambda Y} \geq e^{\lambda t}\right\} &\leq \frac{\mathrm{E}[e^{\lambda Y}]}{e^{\lambda t}} \qquad \lambda > 0 \\
&\vdots \\
\mathrm{P}\left\{\phi(Y) \geq \phi(t)\right\} &\leq \frac{\mathrm{E}[\phi(y)]}{\phi(t)}
\end{aligned}$$

for any strictly monotonic $\phi(\cdot)$.

The Markov inequality

Let Y be a positive random variable. Then for all $t > 0$

$$\mathrm{P}\{Y \geq t\} \;\leq\; \frac{\mathrm{E}[Y]}{t}$$

Chernoff-type bound:

$$\mathrm{P}\{Y \geq t\} \;\leq\; \frac{\mathrm{E}[e^{\lambda Y}]}{e^{\lambda t}} \quad \text{for any} \quad \lambda > 0.$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\mathrm{P}\left\{\|v\|_2^2 > 1 + \delta\right\} \leq \frac{\mathrm{E}[e^{\lambda\|v\|_2^2}]}{e^{\lambda(1+\delta)}}$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\mathrm{P}\left\{\|v\|_2^2 > 1 + \delta\right\} \leq \frac{\mathrm{E}[e^{\lambda\|v\|_2^2}]}{e^{\lambda(1+\delta)}}$$

$$= \frac{\mathrm{E}[e^{\lambda\sum_m v_m^2}]}{e^{\lambda(1+\delta)}}$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\mathrm{P}\left\{\|v\|_2^2 > 1 + \delta\right\} \leq \frac{\mathrm{E}[e^{\lambda\|v\|_2^2}]}{e^{\lambda(1+\delta)}}$$

$$= \frac{\mathrm{E}[e^{\lambda \sum_m v_m^2}]}{e^{\lambda(1+\delta)}}$$

$$= \frac{\mathrm{E}[e^{\lambda v_1^2} e^{\lambda v_2^2} \cdots e^{\lambda v_M^2}]}{e^{\lambda(1+\delta)}}$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\mathrm{P}\left\{\|v\|_2^2 > 1 + \delta\right\} \leq \frac{\mathrm{E}[e^{\lambda\|v\|_2^2}]}{e^{\lambda(1+\delta)}}$$

$$= \frac{\mathrm{E}[e^{\lambda\sum_m v_m^2}]}{e^{\lambda(1+\delta)}}$$

$$= \frac{\mathrm{E}[e^{\lambda v_1^2} e^{\lambda v_2^2} \cdots e^{\lambda v_M^2}]}{e^{\lambda(1+\delta)}}$$

$$= \frac{\mathrm{E}[e^{\lambda v_1^2}]\,\mathrm{E}[e^{\lambda v_2^2}] \cdots \mathrm{E}[e^{\lambda v_M^2}]}{e^{\lambda(1+\delta)}}$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$
\begin{aligned}
\mathrm{P}\left\{\|v\|_2^2 > 1+\delta\right\} &\leq \frac{\mathrm{E}[e^{\lambda\|v\|_2^2}]}{e^{\lambda(1+\delta)}} \\
&= \frac{\mathrm{E}[e^{\lambda\sum_m v_m^2}]}{e^{\lambda(1+\delta)}} \\
&= \frac{\mathrm{E}[e^{\lambda v_1^2} e^{\lambda v_2^2} \cdots e^{\lambda v_M^2}]}{e^{\lambda(1+\delta)}} \\
&= \frac{\mathrm{E}[e^{\lambda v_1^2}]\,\mathrm{E}[e^{\lambda v_2^2}] \cdots \mathrm{E}[e^{\lambda v_M^2}]}{e^{\lambda(1+\delta)}} \\
&= \frac{(\mathrm{E}[e^{\lambda v_1^2}])^M}{e^{\lambda(1+\delta)}} \quad \text{(since } v_m \text{ i.i.d.)}
\end{aligned}
$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\mathrm{P}\left\{\|v\|_2^2 > 1 + \delta\right\} \;\leq\; \frac{(\mathrm{E}[e^{\lambda v_1^2}])^M}{e^{\lambda(1+\delta)}}, \qquad v_1 \sim \mathrm{Normal}(0, M^{-1})$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\mathrm{P}\left\{\|v\|_2^2 > 1 + \delta\right\} \;\leq\; \frac{(\mathrm{E}[e^{\lambda v_1^2}])^M}{e^{\lambda(1+\delta)}}, \qquad v_1 \sim \mathrm{Normal}(0, M^{-1})$$

It is known that

$$\mathrm{E}[e^{\lambda v_1^2}] \;=\; \frac{1}{\sqrt{1 - 2\lambda/M}} \qquad \text{for } \lambda < M/2.$$

A first upper concentration bound ...

For $v = \Phi x$, $\|x\|_2 = 1$, we have that

$$\mathrm{P}\left\{\|v\|_2^2 > 1+\delta\right\} \;\leq\; \frac{(\mathrm{E}[e^{\lambda v_1^2}])^M}{e^{\lambda(1+\delta)}}, \qquad v_1 \sim \mathrm{Normal}(0, M^{-1})$$

And so

$$\mathrm{P}\left\{\|v\|_2^2 > 1+\delta\right\} \;\leq\; \left(\frac{e^{-2\lambda(1+\delta)/M}}{1-2\lambda/M}\right)^{M/2} \qquad \forall\ \lambda < M/2$$

A first upper concentration bound ...

We have

$$\mathrm{P}\left\{\|v\|_2^2 > 1+\delta\right\} \;\; \leq \;\; \left(\frac{e^{-2\lambda(1+\delta)/M}}{1-2\lambda/M}\right)^{M/2} \qquad \forall\ \lambda < M/2$$

Choose

$$\lambda = \frac{M\delta}{2(1+\delta)}$$

(easy to see that in this case $\lambda < M/2$).

A first upper concentration bound ...

We have

$$\mathrm{P}\left\{\|v\|_2^2 > 1+\delta\right\} \;\leq\; \left(\frac{e^{-2\lambda(1+\delta)/M}}{1-2\lambda/M}\right)^{M/2} \qquad \forall\ \lambda < M/2$$

Choose

$$\lambda = \frac{M\delta}{2(1+\delta)}$$

(easy to see that in this case $\lambda < M/2$).

And so

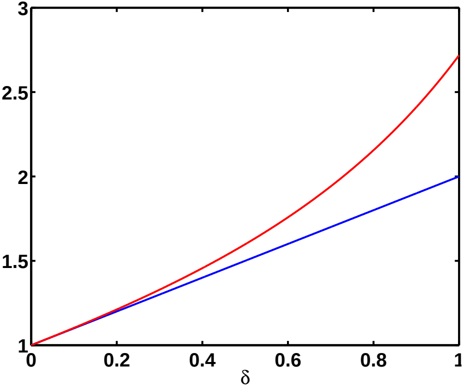
$$\mathrm{P}\left\{\|v\|_2^2 > 1+\delta\right\} \;\leq\; \left((1+\delta)e^{-\delta}\right)^{M/2}.$$

The upper concentration bound

We have

$$\mathrm{P}\left\{\|v\|_2^2 > 1+\delta\right\} \;\leq\; \left((1+\delta)e^{-\delta}\right)^{M/2}.$$

blue: $1+\delta$, red: $e^{\delta-(\delta^2-\delta^3)/2}$



The upper concentration bound

We have

$$\mathrm{P}\left\{\|v\|_2^2 > 1+\delta\right\} \;\leq\; \left((1+\delta)e^{-\delta}\right)^{M/2}.$$

and so

$$\mathrm{P}\left\{\|v\|_2^2 > 1+\delta\right\} \;\leq\; e^{-(\delta^2-\delta^3)M/4}$$

The lower concentration bound

The lower bound follows the exact same sequence of steps (work them out at home!):

$$
\begin{aligned}
\mathrm{P}\left\{\|v\|_2^2 < 1-\delta\right\} &\leq \left(\frac{e^{2(1-\delta)\lambda/M}}{1+2\lambda/M}\right)^{M/2} \\
&\leq \left((1-\delta)e^{\delta}\right)^{M/2} \quad \text{by taking} \quad \lambda = \frac{M\delta}{2(1-\delta)} \\
&\leq e^{-(\delta^2-\delta^3)M/4}
\end{aligned}
$$

The Johnson-Lindenstrauss Lemma

We have shown that for any *fixed* $x \in \mathbb{R}^N$

$$(1-\delta)\|x\|_2^2 \;\le\; \|\Phi x\|_2^2 \;\le\; (1+\delta)\|x\|_2^2$$

with probability exceeding $1 - 2e^{-c\delta^2 M}$.

(Can take $c = 1/8$.)

A simple application of the union bound means that for any set of K vectors $x_1, x_2, \ldots, x_K$, the above holds with probability exceeding $1 - Ke^{-\delta^2 M/8}$...

The Johnson-Lindenstrauss Lemma

Theorem: (J&L, 1984): Let $\mathcal{Q}$ be a arbitrary set of Q vectors in $\mathbb{R}^N$, and let Φ be an $M \times N$ random linear mapping. Then

$$(1-\delta)\|x_1 - x_2\|_2^2 \;\leq\; \|\Phi(x_1 - x_2)\|_2^2 \;\leq\; (1+\delta)\|x_1 - x_2\|_2^2$$

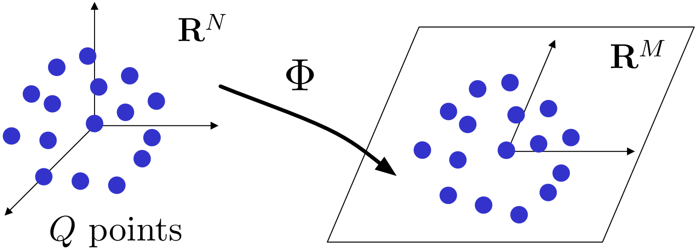
for *all* $x_1, x_2 \in \mathcal{Q}$ with

$$\mathrm{P}\left\{\text{Failure}\right\} \;\leq\; 2Q^2 e^{-\delta^2 M/8} \;\leq\; \epsilon$$

when

$$M \;\geq\; \frac{8}{\delta^2}\left[2\log(Q) + \log\left(\frac{1}{\epsilon}\right) + 0.7\right]$$

The Johnson-Lindenstrauss Lemma



Φ embeds to precision δ with probability ϵ when

$$M \;\geq\; \frac{8}{\delta^2}\left[2\log(Q) + \log\left(\frac{1}{\epsilon}\right) + 0.7\right]$$

Concentration bound

We have: For any *fixed* $x \in \mathbb{R}^N$

$$(1-\delta)\|x\|_2^2 \;\leq\; \|\Phi x\|_2^2 \;\leq\; (1+\delta)\|x\|_2^2$$

with probability exceeding $1 - 2e^{-c\delta^2 M}$.

(Can take $c = 1/8$.)

We want: this for *all* $2S$-sparse x simultaneously...

A single $2S$-dimensional subspace

Theorem: Let V be a $2S$-dimensional subspace of $\mathbb{R}^N$. Then

$$\mathrm{P}\left\{\sup_{x \in V} \left|\|\Phi x\|_2^2 - \|x\|_2^2\right| > \delta\right\} \;\leq\; 2 \cdot 9^{2S} \cdot e^{-c'\delta^2 M}$$

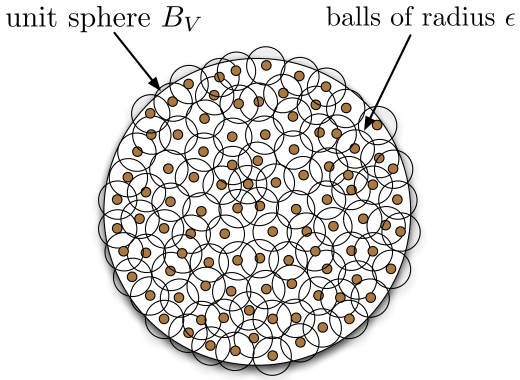
where the constant $c' = c/4$ with c from the previous theorem.

As before, it is enough to prove this for

$$x \in B_V = \{x \in V : \|x\|_2 = 1\}$$

Covering the sphere

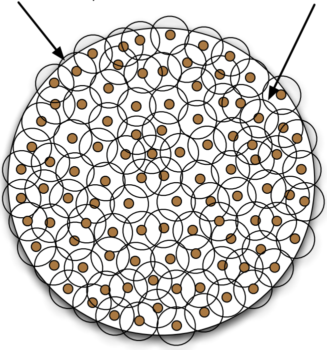
An ϵ-net for B_V:



for every $x \in B_V$, there is a $y \in \mathrm{Net}$ such that $\|x - y\|_2 \leq \epsilon$

$N(B_V, \epsilon)$ is the *size of the smallest ϵ-net*

Covering the sphere



It is a fact that

$$N(B_V, \epsilon) \;\leq\; \left(1 + \frac{2}{\epsilon}\right)^{2S}$$

From discrete to continuous

Lemma: Fix $0 \le \epsilon < 1/2$, and let $\mathcal{N}_\epsilon$ be the minimal ϵ-net for B_V. Then

$$\sup_{x \in B_V} \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| \;\le\; \frac{1}{1-2\epsilon} \max_{y \in \mathcal{N}_\epsilon} \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right|$$

A single $2S$-dimensional subspace

Theorem: Let V be a $2S$-dimensional subspace of $\mathbb{R}^N$. Then

$$\mathrm{P}\left\{\sup_{x\in V}\left|\|\Phi x\|_2^2 - \|x\|_2^2\right| > \delta\right\} \;\leq\; 2\cdot 9^{2S}\cdot e^{-c'\delta^2 M}$$

where the constant $c' = c/4$ with c from the previous theorem.

A single $2S$-dimensional subspace

Theorem: Let V be a $2S$-dimensional subspace of $\mathbb{R}^N$. Then

$$\mathrm{P}\left\{\sup_{x\in V}\left|\|\Phi x\|_2^2 - \|x\|_2^2\right| > \delta\right\} \;\leq\; 2\cdot 9^{2S}\cdot e^{-c'\delta^2 M}$$

where the constant $c' = c/4$ with c from the previous theorem.

So Φ is "well-conditioned" on V when

$$M \;\geq\; \mathrm{Const}\cdot S$$

A single $2S$-dimensional subspace

Theorem: Let V be a $2S$-dimensional subspace of $\mathbb{R}^N$. Then

$$\mathrm{P}\left\{\sup_{x\in V}\left|\|\Phi x\|_2^2 - \|x\|_2^2\right| > \delta\right\} \;\leq\; 2\cdot 9^{2S}\cdot e^{-c'\delta^2 M}$$

where the constant $c' = c/4$ with c from the previous theorem.

We want this for *all subspaces* in which $2S$-sparse signals live...

A single $2S$-dimensional subspace

Theorem: Let V be a $2S$-dimensional subspace of $\mathbb{R}^N$. Then

$$\mathrm{P}\left\{\sup_{x\in V}\left|\|\Phi x\|_2^2 - \|x\|_2^2\right| > \delta\right\} \;\le\; 2\cdot 9^{2S}\cdot e^{-c'\delta^2 M}$$

where the constant $c' = c/4$ with c from the previous theorem.

We want this for *all subspaces* in which $2S$-sparse signals live...

There are $\binom{N}{2S} \le \left(\frac{Ne}{2S}\right)^{2S}$ such subspaces...

All $2S$-dimensional subspaces

For $\Gamma \subset \{1, \ldots, N\}$, let

$$B_\Gamma = \left\{ x \in \mathbb{R}^N \;:\; x_\gamma = 0,\ \gamma \notin \Gamma,\ \|x\|_2 = 1 \right\}.$$

Theorem:

$$\mathrm{P}\left\{ \max_{|\Gamma| \le 2S} \sup_{x \in B_\Gamma} \left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| > \delta \right\} \;\le\; 2 \left(\frac{Ne}{2S} \right)^{2S} 9^{2S}\, e^{-c' \delta^2 M}$$

SUCCESS!!!

All $2S$-dimensional subspaces

Theorem:

$$\mathrm{P}\left\{\sup_{\text{all } 2S \text{ sparse } x} \left|\|\Phi x\|_2^2 - \|x\|_2^2\right| > \delta\right\} \leq 2\left(\frac{Ne}{2S}\right)^{2S} 9^{2S}\, e^{-c'\delta^2 M}$$
$$= e^{\log 2 + 2S\log(Ne/2S) + 2S\log 9 - c'\delta^2 M}$$

Which is to say

$$(1-\delta)\|x\|_2^2 \;\leq\; \|\Phi x\|_2^2 \;\leq\; (1+\delta)\|x\|_2^2 \quad \forall\ 2S - \text{sparse } x$$

with high probability when

$$M \;\geq\; \frac{\mathrm{Const}}{\delta^2}\cdot S\log(N/S)$$

SUCCESS!!!

Next up ...

Theorem: Let Φ be an $M \times N$ matrix that is an approximate isometry for $3S$-sparse vectors. Let x_0 be an S-sparse vector, and suppose we observe $y = \Phi x_0$. Given y, the solution to

$$\min_x \ \|x\|_1 \quad \text{subject to} \quad \Phi x = y$$

is *exactly* x_0.

Moving to the solution

$$\min_x \ \|x\|_1 \quad \text{such that} \quad \Phi x = y$$

Call the solution to this $x^\sharp$. Set

$$h = x^\sharp - x_0.$$

Moving to the solution

$$\min_x \ \|x\|_1 \quad \text{such that} \quad \Phi x = y$$

Call the solution to this $x^\sharp$. Set

$$h = x^\sharp - x_0.$$

Two things must be true:

- $\Phi h = 0$
  Simply because both $x^\sharp$ and x_0 are feasible: $\Phi x^\sharp = y = \Phi x_0$
- $\|x_0 + h\|_1 \leq \|x_0\|_1$
  Simply because $x_0 + h = x^\sharp$, and $\|x^\sharp\|_1 \leq \|x_0\|_1$

Moving to the solution

$$\min_{x} \ \|x\|_1 \quad \text{such that} \quad \Phi x = y$$

Call the solution to this $x^\sharp$. Set

$$h = x^\sharp - x_0.$$

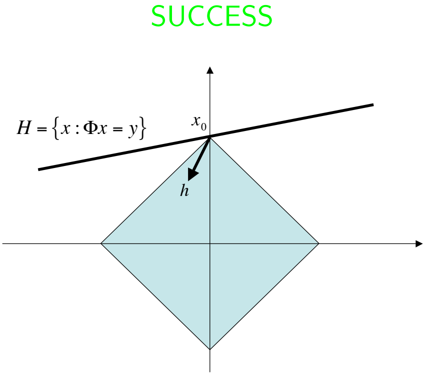
Two things must be true:

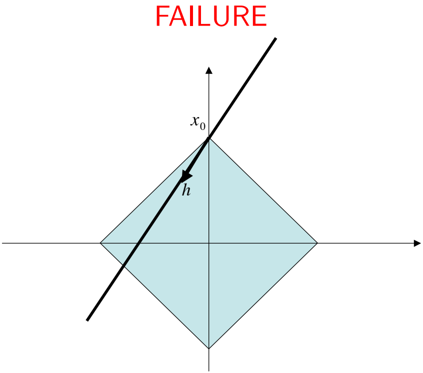
- $\Phi h = 0$
  Simply because both $x^\sharp$ and x_0 are feasible: $\Phi x^\sharp = y = \Phi x_0$

- $\|x_0 + h\|_1 \leq \|x_0\|_1$
  Simply because $x_0 + h = x^\sharp$, and $\|x^\sharp\|_1 \leq \|x_0\|_1$

We'll show that if Φ is $3S$-RIP, then these conditions are *incompatible* unless $h = 0$

Geometry





Two things must be true:

- $\Phi h = 0$
- $\|x_0 + h\|_1 \leq \|x_0\|_1$

Cone condition

For $\Gamma \subset \{1, \ldots, N\}$, define $h_\Gamma \in \mathbb{R}^N$ as

$$h_\Gamma(\gamma) = \begin{cases} h(\gamma) & \gamma \in \Gamma \\ 0 & \gamma \notin \Gamma \end{cases}$$

Let Γ_0 be the support of x_0. For any "descent vector" h, we have

$$\|h_{\Gamma_0^c}\|_1 \;\leq\; \|h_{\Gamma_0}\|_1$$

Cone condition

For $\Gamma \subset \{1, \ldots, N\}$, define $h_\Gamma \in \mathbb{R}^N$ as

$$h_\Gamma(\gamma) = \begin{cases} h(\gamma) & \gamma \in \Gamma \\ 0 & \gamma \notin \Gamma \end{cases}$$

Let Γ_0 be the support of x_0. For any "descent vector" h, we have

$$\|h_{\Gamma_0^c}\|_1 \;\leq\; \|h_{\Gamma_0}\|_1$$

Why? The triangle inequality..

$$\begin{aligned} \|x_0\|_1 \geq \|x_0 + h\|_1 &= \|x_0 + h_{\Gamma_0} + h_{\Gamma_0^c}\|_1 \\ &\geq \|x_0 + h_{\Gamma_0^c}\|_1 - \|h_{\Gamma_0}\|_1 \\ &= \|x_0\|_1 + \|h_{\Gamma_0^c}\|_1 - \|h_{\Gamma_0}\|_1 \end{aligned}$$

Cone condition

For $\Gamma \subset \{1, \ldots, N\}$, define $h_\Gamma \in \mathbb{R}^N$ as

$$h_\Gamma(\gamma) = \begin{cases} h(\gamma) & \gamma \in \Gamma \\ 0 & \gamma \notin \Gamma \end{cases}$$

Let Γ_0 be the support of x_0. For any "descent vector" h, we have

$$\|h_{\Gamma_0^c}\|_1 \;\leq\; \|h_{\Gamma_0}\|_1$$

We will show that if Φ is $3S$-RIP, then

$$\Phi h = 0 \quad \Rightarrow \quad \|h_{\Gamma_0}\|_1 \leq \rho \|h_{\Gamma_0^c}\|_1$$

for some $\rho < 1$, and so $h = 0$.

Some basic facts about ℓ_p norms

- $\|h_\Gamma\|_\infty \leq \|h_\Gamma\|_2 \leq \|h_\Gamma\|_1$
- $\|h_\Gamma\|_1 \leq \sqrt{|\Gamma|} \cdot \|h_\Gamma\|_2$
- $\|h_\Gamma\|_2 \leq \sqrt{|\Gamma|} \cdot \|h_\Gamma\|_\infty$

Dividing up $h_{\Gamma_0^c}$

Recall that Γ_0 is the support of x_0

Fix $h \in \mathrm{Null}(\Phi)$. Let

$$\begin{aligned}
\Gamma_1 &= \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\
\Gamma_2 &= \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\
&\vdots
\end{aligned}$$

Dividing up $h_{\Gamma_0^c}$

Recall that Γ_0 is the support of x_0

Fix $h \in \mathrm{Null}(\Phi)$. Let

$$\begin{aligned} \Gamma_1 &= \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ \Gamma_2 &= \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ &\vdots \end{aligned}$$

Then

$$0 = \|\Phi h\|_2 = \|\Phi(\sum_{j \geq 1} h_{\Gamma_j})\|_2 \geq \|\Phi(h_{\Gamma_0} + h_{\Gamma_1})\|_2 - \|\sum_{j \geq 2} \Phi h_{\Gamma_j}\|_2$$

Dividing up $h_{\Gamma_0^c}$

Recall that Γ_0 is the support of x_0

Fix $h \in \mathrm{Null}(\Phi)$. Let

$$\begin{aligned}\Gamma_1 &= \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ \Gamma_2 &= \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ &\vdots\end{aligned}$$

Then

$$\begin{aligned}0 = \|\Phi h\|_2 = \|\Phi(\sum_{j\geq 1} h_{\Gamma_j})\|_2 &\geq \|\Phi(h_{\Gamma_0} + h_{\Gamma_1})\|_2 - \|\sum_{j\geq 2} \Phi h_{\Gamma_j}\|_2 \\ &\geq \|\Phi(h_{\Gamma_0} + h_{\Gamma_1})\|_2 - \sum_{j\geq 2} \|\Phi h_{\Gamma_j}\|_2\end{aligned}$$

Dividing up $h_{\Gamma_0^c}$

Recall that Γ_0 is the support of x_0

Fix $h \in \mathrm{Null}(\Phi)$. Let

$$\begin{aligned}\Gamma_1 &= \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c},\\ \Gamma_2 &= \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c},\\ &\vdots\end{aligned}$$

Then

$$\|\Phi(h_{\Gamma_0} + h_{\Gamma_1})\|_2 \;\leq\; \sum_{j \geq 2} \|\Phi h_{\Gamma_j}\|_2$$

Dividing up $h_{\Gamma_0^c}$

Recall that Γ_0 is the support of x_0

Fix $h \in \mathrm{Null}(\Phi)$. Let

$$\begin{aligned} \Gamma_1 &= \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ \Gamma_2 &= \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ &\vdots \end{aligned}$$

Applying the $3S$-RIP gives

$$\begin{aligned} \sqrt{1-\delta_{3S}}\|h_{\Gamma_0}+h_{\Gamma_1}\|_2 &\le \|\Phi(h_{\Gamma_0}+h_{\Gamma_1})\|_2 \\ &\le \sum_{j\ge 2}\|\Phi h_{\Gamma_j}\|_2 \;\le\; \sum_{j\ge 2}\sqrt{1+\delta_{2S}}\|h_{\Gamma_j}\|_2 \end{aligned}$$

Dividing up $h_{\Gamma_0^c}$

Recall that Γ_0 is the support of x_0

Fix $h \in \mathrm{Null}(\Phi)$. Let

$$\begin{aligned}
\Gamma_1 &= \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\
\Gamma_2 &= \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\
&\vdots
\end{aligned}$$

Applying the $3S$-RIP gives

$$\|h_{\Gamma_0} + h_{\Gamma_1}\|_2 \;\leq\; \sqrt{\frac{1+\delta_{2S}}{1-\delta_{3S}}} \sum_{j\geq 2} \|h_{\Gamma_j}\|_2$$

Dividing up $h_{\Gamma_0^c}$

Recall that Γ_0 is the support of x_0

Fix $h \in \mathrm{Null}(\Phi)$. Let

$$\begin{aligned} \Gamma_1 &= \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ \Gamma_2 &= \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ &\vdots \end{aligned}$$

Then

$$\|h_{\Gamma_0} + h_{\Gamma_1}\|_2 \ \le \ \sqrt{\frac{1+\delta_{2S}}{1-\delta_{3S}}} \sum_{j\ge 2} \sqrt{2S}\|h_{\Gamma_j}\|_\infty$$

since $\|h_{\Gamma_j}\|_2 \le \sqrt{2S}\|h_{\Gamma_j}\|_\infty$

Dividing up $h_{\Gamma_0^c}$

Recall that Γ_0 is the support of x_0

Fix $h \in \mathrm{Null}(\Phi)$. Let

$$\begin{aligned} \Gamma_1 &= \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ \Gamma_2 &= \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ &\vdots \end{aligned}$$

Then

$$\|h_{\Gamma_0} + h_{\Gamma_1}\|_2 \;\leq\; \sqrt{\frac{1+\delta_{2S}}{1-\delta_{3S}}} \sum_{j \geq 1} \frac{1}{\sqrt{2S}} \|h_{\Gamma_j}\|_1$$

since $\|h_{\Gamma_j}\|_\infty \leq \frac{1}{2S} \|h_{\Gamma_{j-1}}\|_1$

Dividing up $h_{\Gamma_0^c}$

Recall that Γ_0 is the support of x_0

Fix $h \in \mathrm{Null}(\Phi)$. Let

$$\begin{aligned} \Gamma_1 &= \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ \Gamma_2 &= \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ &\vdots \end{aligned}$$

Which means

$$\|h_{\Gamma_0} + h_{\Gamma_1}\|_2 \ \le \ \sqrt{\frac{1+\delta_{2S}}{1-\delta_{3S}}}\,\frac{\|h_{\Gamma_0^c}\|_1}{\sqrt{2S}}$$

Dividing up $h_{\Gamma_0^c}$

Recall that Γ_0 is the support of x_0

Fix $h \in \mathrm{Null}(\Phi)$. Let

$$\begin{aligned}
\Gamma_1 &= \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\
\Gamma_2 &= \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\
&\vdots
\end{aligned}$$

Working to the left

$$\|h_{\Gamma_0}\|_2 \;\leq\; \|h_{\Gamma_0} + h_{\Gamma_1}\|_2 \;\leq\; \sqrt{\frac{1+\delta_{2S}}{1-\delta_{3S}}}\,\frac{\|h_{\Gamma_0^c}\|_1}{\sqrt{2S}}$$

Dividing up $h_{\Gamma_0^c}$

Recall that Γ_0 is the support of x_0

Fix $h \in \mathrm{Null}(\Phi)$. Let

$$\begin{aligned} \Gamma_1 &= \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ \Gamma_2 &= \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c}, \\ &\vdots \end{aligned}$$

Working to the left

$$\frac{\|h_{\Gamma_0}\|_1}{\sqrt{S}} \;\le\; \|h_{\Gamma_0}\|_2 \;\le\; \|h_{\Gamma_0} + h_{\Gamma_1}\|_2 \;\le\; \sqrt{\frac{1+\delta_{2S}}{1-\delta_{3S}}}\,\frac{\|h_{\Gamma_0^c}\|_1}{\sqrt{2S}}$$

Wrapping it up

We have shown

$$\|h_{\Gamma_0}\|_1 \le \sqrt{\frac{1+\delta_{2S}}{1-\delta_{3S}}}\sqrt{\frac{S}{2S}}\|h_{\Gamma_0^c}\|_1$$

$$= \rho\|h_{\Gamma_0^c}\|_1$$

for

$$\rho = \sqrt{\frac{1+\delta_{2S}}{2(1-\delta_{3S})}}$$

Wrapping it up

We have shown

$$\|h_{\Gamma_0}\|_1 \le \sqrt{\frac{1+\delta_{2S}}{1-\delta_{3S}}}\sqrt{\frac{S}{2S}}\|h_{\Gamma_0^c}\|_1$$

$$= \rho\|h_{\Gamma_0^c}\|_1$$

for

$$\rho = \sqrt{\frac{1+\delta_{2S}}{2(1-\delta_{3S})}}$$

Taking $\delta_{2S} \le \delta_{3S} < 1/3 \quad \Rightarrow \quad \rho < 1.$

SUCCESS!!

Theorem: Let Φ be an $M \times N$ matrix that is an approximate isometry for $3S$-sparse vectors. Let x_0 be an S-sparse vector, and suppose we observe $y = \Phi x_0$. Given y, the solution to

$$\min_x \ \|x\|_1 \quad \text{subject to} \quad \Phi x = y$$

is *exactly* x_0.