[redacted]

RIP & NSP & Spark

A has RIP of order k if $\forall \|X\|_0 \le k$ $\quad (1-\delta_k)\|X\|_2^2 \le \|AX\|_2^2 \le (1+\delta_k)\|X\|_2^2$

A has NSP of order k if $\forall h \in N(A)$ $\quad \|h_S\|_2 \le c \dfrac{\|h_{S^c}\|_1}{\sqrt{k}}$ $\quad \forall\, |S| \le k$

A has spark $\ge k+1$ if $\forall h \in N(A)$ $\quad \|h\|_0 \ge k+1$

Relations: $\quad$ RIP $\Rightarrow$ NSP

RIP is like a well conditioned version of spark.

$\text{Spark}(A) \ge k+1 \Rightarrow \forall \|X\|_0 \le k \quad 0 < \|Ax\|_2$ "not in null space"

Activity: $A_{m\times n}$ has RIP of order k if $\exists\ \delta_k$ s.t.

$$(1-\delta_k)\|x\|_2 \le \|Ax\|_2 \le (1+\delta_k)\|x\|_2 \quad \forall\ \|x\|_0 \le k.$$

Do you think the following matrix satisfies RIP of order m?

$$A = \begin{pmatrix} 1 & 1^2 & & 1^n \\ 2 & 2^2 & & 2^n \\ 3 & 3^2 & \cdots & \vdots \\ \vdots & \vdots & & \vdots \\ m & m^2 & & m^n \end{pmatrix}$$

but normalize each column to be length 1.

Is it a good sensing matrix?

RIP & Stable Embeddings

A is $m \times n$ w/ $m \ll n$

A has RIP $\|x\|_2 \approx \|Ax\|_2$ for sparse x

$A : \mathbb{R}^n \to \mathbb{R}^m$.

So A preserves the magnitude of sparse vectors.
As A linear, preserves the distance between sparse points.

Dimension reduction!

Noisy recovery thm

Let A satisfy RIP $(2k)$ with $\delta_{2k} < \sqrt{2}-1$

Let $y = Ax_0 + e$, $\|e\| \leq \varepsilon$.

Then ~~from~~ $\min \|x\|_1$ st $\|Ax - y\| \leq \varepsilon$

satisfies $\|\hat{x} - x\|_2 \leq C_0 \frac{\|x_0 - x_{0,k}\|_1}{\sqrt{k}} + C_2 \varepsilon$.

Proof: Outline (in case $x_{0,k} = 0$)

Technical result: Fix any h. Let Δ be top $2k$ entries

$$\|h_\Delta\|_2 \leq \beta \frac{|\langle Ah_\Delta, Ah \rangle|}{\|h_\Delta\|_2}$$

Optimizer satisfy:

$$\|\hat{x} - x\|_2 \leq C' \frac{|\langle Ah_\Delta, Ah \rangle|}{\|h_\Delta\|_2} \quad w/ \quad h = \hat{x} - x$$

$$\leq C_1 (1+\delta_{2k}) \frac{\cancel{\|h_\Delta\|_2}}{\cancel{\|h_\Delta\|_2}} \|Ah\|$$

As $\|A\hat{x} - y\| \leq \varepsilon$ & $\|Ax_0 - y\| \leq \varepsilon$, $\|A\hat{x} - Ax_0\| \leq 2\varepsilon$

So $\|\hat{x} - x\|_2 \leq C_1 (1+\delta_{2k})\, 2\varepsilon$.

How good is error estimate?

$$\|\hat{x} - x\|_2 \leq C_2 \varepsilon \qquad \text{for } y = Ax + e \text{ w } \|e\|_2 \leq \varepsilon.$$

Suppose we know which k coefs of x were nonzero, S.

$$y = A_S x$$

restriction of A to columns in S.
overdetermined (tall) matrix $m \times k$

Solve for x by least squares

$$\hat{x} = (A_S^t A_S)^{-1} A_S y.$$

How much error

$$\hat{x} - x_0 = (A_S^t A_S)^{-1} A_S (x_0 + e) - x_0$$

$$\hat{x} - x_0 = (A_S^t A_S)^{-1} A_S \, e$$

$$\|\hat{x} - x_0\| \leq \underbrace{\|(A_S^t A_S)^{-1} A_S\|}_{} \, \|e\|_2$$

like inverse of A
its max sing value
is like reciprocal of
A_S's min sing vals

$$\leq \frac{1}{\sqrt{1 - \delta_{2k}}} \|e\|_2.$$

So error estimate optimal up to a constant factor

Activity:

A signal is compressible if

$$|X_0|_{(k)} \leq C_r \cdot k^{-r}$$

a) Compute worst case $\|X_0 - X_{0,S}\|_2$

b) Compute worst case

$$\frac{\|X_0 - X_{0,S}\|_1}{\sqrt{S}}$$

c) Could anyone improve bound $\|\hat{X} - X_0\|_2 \leq C \frac{\|X_0 - X_{0,S}\|_1}{\sqrt{S}}$?

d) Isnt every finite signal compressible ??

Thm: Let A be $N \times n$ w/ iid $N(0,1)$ entries

$\forall\, t \geq 0$,

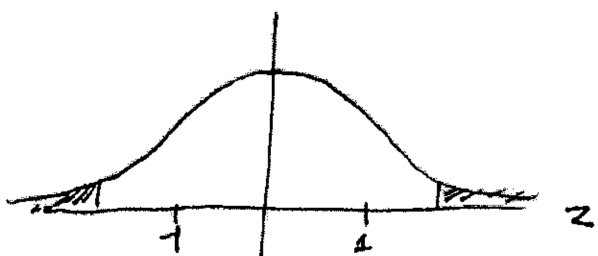
$$\sqrt{N} - \sqrt{n} - t \leq \sigma_{min}(A) \leq \sigma_{max}(A) \leq \sqrt{N} + \sqrt{n} + t$$

w/ probability at least $1 - 2e^{-t^2/2}$.

"Tall random matrices are approximate isometries"

Note: taking t big provides worse control on singular values but better guarantee on probability bound is true.

Compare to: Let $z \sim N(0,1)$ $P(|z| \in (-t,t)) \geq 1 - \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ for $t > 1$



If we want more certainty on window of Z we need wider window

Theorem: Let $A_{m \times n}$ have iid $N(0,1)$ entries.

Let $\bar{A} = \frac{1}{\sqrt{m}} A$. $\forall k \in [1,n]$ $\forall \delta \in (0,1)$

if $m \geq \frac{C_1}{\delta^2} k \log\left(\frac{en}{k}\right)$ then $\delta_k(\bar{A}) < \delta$

with probability at least $1 - 2e^{-c\delta^2 m}$. Here, C_1, c are universal constants.

Proof: ~~$\forall t \geq 0$~~ ~~$\sqrt{m} - \sqrt{k} - t \leq \sigma_{min}$~~

Consider a fixed subset $T \subset [1,n]$ w $|T| = k$

By concentration of singular values,

$$\sqrt{m} - \sqrt{k} - t \leq \sigma_{min}(A_T) \leq \sigma_{max}(A_T) \leq \sqrt{m} + \sqrt{k} + t \quad \text{w/ prob } 1-2e^{-ct^2}$$

So

$$1 - \frac{\sqrt{k}}{\sqrt{m}} - \frac{t}{\sqrt{m}} \leq \sigma_{min}(\bar{A}_T) \leq \sigma_{max}(\bar{A}_T) \leq 1 + \sqrt{\tfrac{k}{m}} + \tfrac{t}{\sqrt{m}} \quad \text{w/ prob } 1-2e^{-ct^2}$$

So

$$\|A_T^t A_T - I_k\| \leq 3\delta_0 \quad \text{where } \delta_0 = \sqrt{\frac{k}{m}} + \frac{t}{\sqrt{m}}$$

where t, m will be st $\delta_0 < \delta/3$

So <u>all</u> subsets of size k are st $\|A_T^t A_T - I_k\| < \delta$

with probability at least $1 - \binom{n}{k} 2e^{-ct^2}$ provided $\sqrt{\frac{k}{m}} + \frac{t}{\sqrt{m}} < \delta/3$.

$$1 - 2e^{k \log(en/k) - ct^2}$$

Note: $\binom{n}{k} \leq \left(\frac{en}{k}\right)^k$

Choose t st $t = \frac{1}{\sqrt{2}} \sqrt{k \log(en/k)} + \frac{\delta}{6}\sqrt{m}$

So prob. at least

$$1 - 2e^{-c\frac{\delta^2 m}{6}}$$

All we need is $\sqrt{\frac{k}{m}} + C_1' \sqrt{\frac{k \log(en/k)}{m}} < \delta/3$

Choose $m \geq \frac{C_1' k \log en/k}{\delta^2}$ ■

Application to compressed sensing

To get robust recovery, setting $\delta_{2k}(A) < \sqrt{2} - 1$

If $m \gtrsim \tilde{C}\, k \log n/k$ then A iid $N(0,1)$ satisfies RIP w/ prob $1 - 2e^{-\tilde{c} m}$

If you want a RIP constant twice as good, need 4 times the # measurements. Dont need arb small δ in order to get recovery