Compressed Sensing

Setup: $x_0 \in \mathbb{R}^n$ unknown signal

$\|x_0\|_0 = s << n$, so x_0 is sparse

measurement $a_i \cdot x_0 = b_i$ or $a_i \cdot x_0 = b_i + \epsilon_i$ noise

Given a_i & b_i, find x.

Want:

$$\min \|x\|_0 \text{ such that } Ax = b \quad \text{w/} \quad A = \begin{pmatrix} -a_1- \\ -a_2- \\ \vdots \\ -a_m- \end{pmatrix}$$

Can't minimize $\|x\|_0$ b/c NP-hard (combinatorially expensive)

Instead

$$\min \|x\|_1 \text{ such that } Ax = b$$

~~Plan:~~ ~~If~~

When do these problems succeed at finding x_0? Stable to noise?
Stable to not-quite sparse signals?

Key ideas for recovery conditions

spark
incoherence
null space property
restricted isometry property

Spark

The <u>spark</u> of $A \in \mathbb{R}^{m \times n}$ is smallest # lin'ly dependent columns

Theorem:

[For any $y \in \mathbb{R}^m$, there exists at most one k-sparse signal x st $y = Ax$] if and only if $spark(A) > 2k$.

Proof:

$\Rightarrow$: Assume at most one k-sparse x st $y = Ax$.
Suppose $spark(A) \leq 2k$. $\exists\, h \in N(A)$ st $\|h\|_0 \leq 2k$
write $h = x - x'$ w/ $\|x\|_0 \leq k$ & $\|x'\|_0 \leq k$. $Ax = Ax'$.
Contradiction.

$\Leftarrow$: Suppose $spark(A) > 2k$. Suppose for some y $\exists$ x, x' st $\|x\|_0 \leq k$, $\|x'\|_0 \leq k$
st $y = Ax = Ax'$. Then $x - x' \in N(A)$. $\|x - x'\|_0 \leq 2k$. ~~contradiction w spark(A) ≤ 2k~~
So $Ah = 0$ for $h = x - x'$. $\|h\|_0 \leq 2k \Rightarrow h = 0 \Rightarrow x = x'$.

1) Setup: Let $x_0 \in \mathbb{R}^n$, $\|x_0\|_0 = s$.

Let $A \in \mathbb{R}^{m \times n}$, $m < n$.

Let $b = Ax_0$.

Find x_0 by solving

$$\min \|x\|_0 \text{ s.t. } Ax = b. \qquad (*)$$

Claim: Fix A. If spark$(A) > 2s$, then $\forall x_0$, x_0 is unique soln to $(*)$.

Fix A If spark$(A) \leq 2s$, $\exists\, x_0$ such that x_0 not unique soln to $(*)$

Gist: If x_0 & x_1 are s-sparse solns of $Ax = b$

$x_0 - x_1 \in N(A)$ & $\|x_0 - x_1\|_0 \leq 2s$.

Q: Why isnt the following a counter example to Thm 1.1?

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad (**)$$

A has spark 2 ~~3 2 4~~ (take $k = s = 1$)

but there is exactly one soln to $(**)$ with sparsity ≤ 1:

$$x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

A: Spark condition is equivalent to $(*)$ working for all x_0

If spark is too low, $(*)$ may work for some x_0 but will necessarily fail for some other x_0.

Coherence

Let $A = \begin{pmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{pmatrix}$

$$\mu(A) = \max_{i \neq j} \frac{|\langle a_i, a_j \rangle|}{\|a_i\|_2 \, \|a_j\|_2} \quad \text{is coherence}$$

Low values of coherence are good.

Fact: $\mu(A) \in \left[\sqrt{\dfrac{n-m}{m(n-1)}}, 1 \right]$

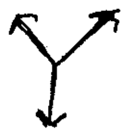
Welch bound

4) A is $m \times n$ $\quad \boxed{\begin{matrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{matrix}}$

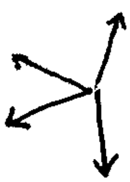
$a_i \in \mathbb{R}^m$ is column of A.

Low coherence;  in $\mathbb{R}^2$

lowest coherence with 3 vectors in $\mathbb{R}^2$

 Coherence is $\frac{1}{2}$

~~lowest coherence with 4 vect~~
lowest coherence with 4 vectors in $\mathbb{R}^2$?



 ? No coherence $= 1$ in this case

Large coherence: parallel vectors

5) A is $m \times n$ $\boxed{a_1 - - - - - a_n}$

$\mu(A) \geq \sqrt{\dfrac{n-m}{m(n-1)}}$ Welch bound

If $n \gg m$, this bound becomes $\mu(A) \geq \dfrac{1}{\sqrt{m}}$

This bound appears to depend on wrong variable ??!!

~~The ambient dimension~~
If n is larger, vectors must start "clumping"
So minimal incoherence should be ~~decaying~~ growing
in n

Resolution: Equality in welch bound can only occur when $n \leq \binom{m+1}{2}$

~~[illegible]~~

In that case $\mu(A) \gtrsim \frac{1}{\sqrt{m}}$ — ambient dimension controls best incoherence up to a certain n.
Then n controls best incoherence

Lemma: For any $A \in \mathbb{R}^{m \times n}$, $\text{Spark}(A) \geq 1 + \frac{1}{\mu(A)}$

Proof (by Gershgorin's Circle Theorem):

WLOG, assume A has unit magnitude columns.

Let $\Lambda \subseteq \{1 \cdots n\}$ be st $|\Lambda| = p$.

Let $G = A_\Lambda^t A_\Lambda$. Then $g_{ii} = 1 \ \forall i$ & $|g_{ij}| \leq \mu(A) \ \forall i \neq j$.

By Gershgorin's Circle Thm: if $\sum_{j \neq i} |g_{ij}| < |g_{ii}|$, then

G is PSD. This cols of A_Λ are lin'ly independent

~~We have $p < 1 + \frac{1}{\mu(A)} \ \forall p < \text{spark}(A)$~~

Hence if $(p-1)\mu < 1$ then $\text{spark}(A) > p$

if $p < 1 + \frac{1}{\mu}$ then $\text{spark}(A) > p$

So $\text{Spark}(A) \geq 1 + \frac{1}{\mu}$

Thm: If $k < \frac{1}{2}\left(1 + \frac{1}{\mu(A)}\right)$ then

$\forall y \in \mathbb{R}^m$ there is at most one x st $y = Ax$ & $\|x\|_0 \leq k$.

With bound $\mu(A) \geq \frac{1}{\sqrt{m}}$ that means

can only guarantee recovery for $k \lesssim \sqrt{m}$.