Lecture 2/03/16

Wednesday, March 2, 2016 1:16 PM

Todays Agenda

→ Recap the CS algo ✓
→ Reprove the noiseless setting ✓
→ Show that gaussian matrices satisfy good RIP ✓

noiseless version.

1. We pick a sensing matrix $\phi \in \mathbb{R}^{m \times n}$
2. Adversary picks x, $\|x\|_0 = k$, $x \in \mathbb{R}^n$
3. We observe $\phi x = y \in \mathbb{R}^m$
4. Can we recover x from y?

Try to make ϕ as small as possible

Algo [Candes – Tao]

I. Choose ϕ to be (δ, k) – RIP matrix

$\Rightarrow \left[\forall \; \|x\|_2 = 1, \; x \text{ is } k\text{-sparse}, \quad 1 - \delta \le \|\phi x\|_2^2 \le 1 + \delta \right]$

II. $y = \phi x$ is observed

III. Solve LP :- $x^* \equiv \min_{\tilde{x}} \|\tilde{x}\|_1$
$\phi \tilde{x} = y$ } Linear Program

IV. Output x^*.

Theorem
if $(2k, \delta) < \frac{1}{3}$, then $x^* = x$

Proof :-
Consider error vector

$$h = x^* - x$$

We'll show $\|h\|_2 = 0$

break up co-ordinates into disjoint sets
$T_0, T_1, T_2, \ldots$

T_0 = k - non-zero coordinates in unknown

T_1 = k - largest coordinates of $h_{[n]\setminus T_0}$

T_2 = " " " " $h_{[n]\setminus T_0 \setminus T_1}$

So, $h = \sum_{j \geq 0} h_{T_j}$

Part ① bound $\|h_{T_2 \cup T_3 \cup \dots}\|_2$

Part ② bound $\|h_{T_0 \cup T_1}\|_2$

$$\|h_{T_2}\|_2 \leq \sqrt{k}\, \|h_{T_2}\|_\infty$$

$$\leq \frac{1}{\sqrt{k}} \|h_{T_1}\|_1$$

(max coordinate in T_2 < min coord T_1)

$$\|h_{T_3}\|_2 \leq \frac{1}{\sqrt{k}} \|h_{T_2}\|_1$$

$\vdots$

Ⓐ $\|h_{(T_0 \cup T_1)^c}\|_2 \leq \sum_{j \geq 2} \|h_{T_j}\|_2 \leq \frac{1}{\sqrt{k}} \|h_{T_1 \cup T_2 \cup \dots}\|_1 = \underbrace{\frac{1}{\sqrt{k}} \|h_{T_0^c}\|_1}$

Now, $\|$

$\|h_{T_2} + h_{T_3} + \dots\|_2$

We optimized the $\|x^*\|_1$! ← Opt soln to LP

$$\|x^*\|_1 \leq \|x\|_1$$

$$\Rightarrow \|x - h\|_1 \leq \|x\|_1$$

$$\underbrace{\|(x-h)_{T_0}\|_1}_{\geq \downarrow} + \underbrace{\|(x-h)_{T_0^c}\|_1}_{= \downarrow} \leq \|x_{T_0}\|_1$$

$$\|x_{T_0}\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1 \leq \|x_{T_0}\|_1$$

Ⓑ

$\Rightarrow \boxed{\|h_{T_0^c}\|_1 \leq \|h_{T_0}\|_1}$

$\|x_{T_0}\|_1 \leq \|(x-h)_{T_0}\|_1 + \|h_{T_0}\|_1$

Ⓐ $\Rightarrow \|h_{(T_0 \cup T_1)^c}\|_2 \leq \frac{1}{\sqrt{k}} \|h_{T_0}\|_1$

(2k, δ) RIP

(A) $\Rightarrow \quad \| h_{(T_0 \cup T_1)^c} \|_2 \quad \quad \sqrt{K}$

To bound part (2) Use property that ϕ is $(2k, \delta)$ RIP

$$(1-\delta)\|h_{T_0 \cup T_1}\|_2^2 \le \|\phi h_{T_0 \cup T_1}\|_2^2 = \|\phi h - \sum_{j \ge 2} \phi h_{T_j}\|_2^2$$

$$= \langle \phi h_{T_0 \cup T_1}, \phi h - \sum_{j \ge 2} \phi h_{T_j} \rangle$$

$\phi h = 0 \Rightarrow$ //

$$\underbrace{\langle \phi h_{T_0 \cup T_1}, -\sum_{j \ge 2} \phi h_{T_j} \rangle}$$

To bound this, notice

$$4\langle \phi h_{T_0}, \phi h_{T_j} \rangle = \underbrace{\|\phi(h_{T_0} + h_{T_j})\|_2^2} - \underbrace{\|\phi(h_{T_0} - h_{T_j})\|_2^2}$$

Proved last week

$$\le 4\delta \|h_{T_0}\|_2 \|h_{T_j}\|_2$$

Overall, we can conclude,

$$\left|\langle \phi h_{T_0}, -\sum_{j \ge 2} \phi h_{T_j} \rangle\right| \le \delta \|h_{T_0}\|_2 \sum_{j \ge 2} \|h_{T_j}\|_2$$

$$\le \delta \|h_{T_0 \cup T_1}\|_2 \quad "$$

Similarly for $\langle \phi h_{T_1}, \sum_{j \ge 2} \phi h_{T_j} \rangle$

Crude Δ^{le} Inequality,

$$\left|\langle \phi h_{T_0 \cup T_1}, \sum_{j \ge 2} \phi h_{T_j} \rangle\right| \le 2\delta \|h_{T_0 \cup T_1}\|_2 \sum_{j \ge 2} \|h_{T_j}\|_2$$

$$\Rightarrow \quad (1-\delta)\|h_{T_0 \cup T_1}\|_2^2 \le 2\delta \|h_{T_0 \cup T_1}\|_2 \sum_{j \ge 2} \|h_{T_j}\|_2$$

$$\le 2\delta \|h_{T_0 \cup T_1}\|_2 \cdot \frac{1}{\sqrt{K}} \|h_{T_0}\|_1 \quad (A \,\&\, B)$$

$$\Rightarrow \quad \|h_{T_0 \cup T_1}\|_2 \le \frac{2\delta}{1-\delta} \frac{1}{\sqrt{K}} \|h_{T_0}\|_1$$

$$\overset{CS}{\le} \frac{2\delta}{1-\delta} \frac{1}{\sqrt{K}} \cdot \sqrt{K} \|h_{T_0}\|_2$$

$$\Rightarrow (1-\delta)\,\|h_{T_0\cup T_1}\|_2 \le 2\delta\,\|h_{T_0\cup T_1}\|_2$$

$$\Rightarrow (1-3\delta)\,\|h_{T_0\cup T_1}\|_2 \le 0$$

$$\Rightarrow \|h_{T_0\cup T_1}\|_2 = 0 \quad (\delta < \tfrac{1}{3})$$

$$\Rightarrow h = 0 \;!!$$

How to design good RIP matrices

$\phi \in \mathbb{R}^{m\times n}$ is (δ, k) RIP if

$$(1-\delta) \le \|\phi x\|_2^2 \le (1+\delta) \qquad \forall x \in \mathbb{R}^n, \ \|x\|_2 = 1, \ \|x\|_0 \le k$$

Want small 'm'.

Reminiscent of Johnson – Lindenstrauss Lemma

like JL but for all K-sparse vectors.

$$\phi = \begin{array}{c|c} & \overset{\longleftarrow\ n\ \longrightarrow}{} \\ m & \begin{bmatrix} \text{---}\ g_1\ \text{---} \\ \text{---}\ g_2\ \text{---} \\ \vdots \\ \text{---}\ g_m\ \text{---} \end{bmatrix} \end{array}$$

$\phi = [\, g_{ij} \,]$ where $g_{ij} = N(0,1)$ random variable

a) Fix $x \in \mathbb{R}^n$, x is k-sparse, $\|x\|_2 = 1$

b) Show that whp $\|\phi x\|_2^2 \approx 1 \pm \delta$

c) Union bound over all x, $\|x\|_0 \le k$, $\|x\|_2 = 1$.

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad y = \phi x = \begin{bmatrix} \text{---}\ g_1\ \text{---} \\ \text{---}\ g_2\ \text{---} \\ \vdots \\ \text{---}\ g_m\ \text{---} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

k

$$y_i = \sum_{j=1}^{k} g_{ij} x_j$$

each $g_{ij} \sim N(0,1)$

if $Z \sim N(0,1)$
then $\alpha Z \sim N(0, \alpha^2)$

$\Rightarrow$ $g_{ij} x_j \sim N(0, x_j^2)$

$\Rightarrow y_i \sim N(0, \sum x_j^2)$
$\sim N(0,1)$.

$$y \sim \left.\begin{bmatrix} N(0,1) \\ N(0,1) \\ \vdots \\ N(0,1) \end{bmatrix}\right\} m$$

$$E[\|y\|^2] = E[\sum y_i^2] = \sum E[y_i^2] = m$$

$R = \|y\|^2$ is random variable

$E[R] = m$

$\mathrm{Var}[R] = E[R^2] - E[R]^2 \leq O(m)$

<u>Bernstein's Inequality</u>

$$\Pr[|R - m| > t\sqrt{m}] \leq e^{-t^2}$$

Set $t = \epsilon\sqrt{m}$

$$\Pr[|R - m| > \epsilon m] \leq e^{-\epsilon^2 m}$$

$\Rightarrow$ For fixed u,
wp $\geq 1 - e^{-\epsilon^2 m}$, $\|\phi u\|_2^2 \in [1 \pm \epsilon] m$

Want $\forall$ k-sparse vectors, length to be preserved.

Morally, # vectors to union bound over

$$\leq \underbrace{\binom{n}{k}}_{\text{choosing coordinates}} \underbrace{O(1)^k}_{\text{discretization}}$$

$$= n^{O(k)}$$

$m = \Theta\left(k \log \frac{n}{k}\right)$

So, we're good if $\underbrace{e^{-\epsilon^2 m} \cdot n^{O(k)}}_{\text{overall failure probability}} < 1/2$

We've shown (except for union bound) that $m = \Theta(k \log n)$ suffices!

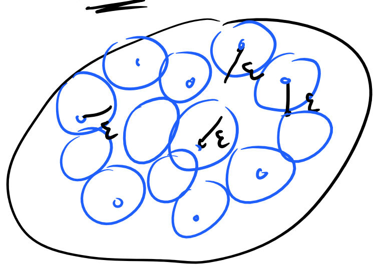
How to union bound over infinitely many vectors ??

Fix k coordinates as $\{1, 2, \ldots k\}$ ← WLOG.

there are ∞ many unit vectors $\|x\| = 1$

$X = \{x : \sum_{i=1}^{k} x_i^2 = 1\}$ are unit vectors!

Find a small set N ("net") st $\forall x \in X$, $\exists\, n \in N$ st $\|x - n\|_2 \le \varepsilon$



X

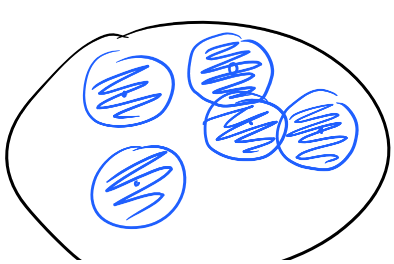
balls of radius ϵ around net points should cover X

How small can N be ?

→ greedy ϵ-net.

Start with arbitrary x_0, add it to N

As long as $\exists\, x \in X$ st $d(x, N) > \varepsilon$, add x to N.



$\forall\, n_1, n_2 \in N$,

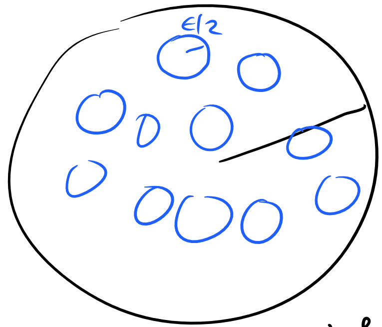
we have $\|n_1 - n_2\|_2 > \epsilon$

$\Rightarrow B(n_i, \frac{\epsilon}{2}) \cap B(n_j, \frac{\epsilon}{2}) = \phi$

$i \neq j \in N$

$\Rightarrow B(n_i, \frac{\epsilon}{2}) \cap B(n_j, \frac{\epsilon}{2}) = \phi$
$\forall i, j \in N$

Moreover, $N \subseteq X$, all points are in unit ball $B(0,1)$



\# net points (ie)

$$|N| \leq \frac{\text{Volume}(B(0,1))}{\text{Volume}(B(\epsilon/2))}$$

$$\propto \left(\frac{2}{\epsilon}\right)^k$$

Vol $\propto r^k$ in k dimensional space.

We'll now show that all points in N are good for ϕ

(ie) $\forall\ n \in N$,

$$1 - \frac{\delta}{2} \leq \|\phi n\|_2^2 \leq 1 + \frac{\delta}{2}$$

Union bound as long as

$$\underbrace{e^{-\epsilon^2 m}}_{} \cdot \underbrace{\binom{n}{k}\left(\frac{2}{\epsilon}\right)^k}_{} < \frac{1}{2}$$

If ϕ is $1 \pm \frac{\delta}{2}$ good for N, will show ϕ is good for X also.

$\downarrow$

Extremality argument.

let $1+A$ be the largest distortion for ϕ.

Meaning $\exists\ \tilde{x}$ st $\|\phi \tilde{x}\|_2 = (1+A)$, $\tilde{x} \in X$

$\boxed{\text{Want to bound } A}$

$\exists\ \tilde{n} \in N$ st $\|\tilde{x} - \tilde{n}\|_2 \leq \frac{\delta}{4}$ (Find ϵ-net with $\epsilon = \frac{\delta}{4}$)

$$\|\phi \tilde{x}\|_2 = 1 + A = \|\phi \tilde{n} + \phi(\tilde{x} - \tilde{n})\|_2$$

$$\leq \|\phi \tilde{n}\|_2 + \|\phi(\tilde{x} - \tilde{n})\|_2$$

$$1 + A \leq 1 + \frac{\delta}{2} + \frac{\delta}{4}(1+A)$$

normalizing + extremality

$$A \le \frac{\delta}{2} + \frac{\delta}{4} + \frac{A\delta}{4}$$

$$\Rightarrow A\left(1 - \frac{\delta}{4}\right) \le \frac{3\delta}{4}$$

$$\Rightarrow A \le \delta$$

recap: Union bound over $\binom{n}{k}$ coordinates

↳ for each k coordinates, compute a $\frac{\delta}{4}$-net for X

Ensure Φ is $\frac{\delta}{2}$ good for all net points

⇓

(extremality argument)

Φ is δ good for all points.