E9 203 Compressed Sensing & Sparse SP.

17 May 2021.

Last time:
Proof of two Lemmas needed to complete the proof of

Thm. Let $0<\delta<1$ be given. Let $A \in \mathbb{R}^{m\times N}$, $A_{ij} \overset{iid}{\sim} \mathcal{N}(0,\frac{1}{m})$.
If $\boxed{m \geq \kappa_1 k \log\left(\frac{N}{k}\right)}$, then A satisfies RIP of order k with the prescribed δ w.p. $\geq 1-2e^{-\kappa_2 m}\left(\frac{42e}{\delta}\right)^k$
where κ_1 is arbitrary & $\kappa_2 = \frac{\delta^2}{24} - \frac{1}{\kappa_1}$.

Today: Some more remarks, and some more RIP results often found in the literature.

Remark: Other matrices that satisfy RIP.
Only step in the proof where we used the randomness of A was in showing that $\|Ax\|_2^2$ concentrates around $\|x\|_2^2$.

$$\Pr\left\{ \left| \|Ax\|_2^2 - \|x\|_2^2 \right| > \varepsilon \|x\|_2^2 \right\} \leq 2e^{-mC_0(\varepsilon)} \quad — (*)$$

(Need $e^{-m^{\ell}(\cdot)}$, where $\ell \geq 1$ $\because$ we later multiply the RHS by $|Q|$ which grows as e^{+m/κ_1}. $\quad \left(e^{m/\kappa_1} \cdot e^{-mC_0(\varepsilon)}\right)$

A. Subgaussian distributions satisfy an ineq. like (*).

Defn. $X \sim \text{sub}(c^2)$ [Read as: X is subgaussian with parameter c^2] if $\exists\, c>0$ s.t.

$$\underbrace{E\{e^{sX}\}}_{\text{MGF decay at least as fast as a Gaussian.}} \leq e^{c^2 s^2/2} \quad \forall s \in \mathbb{R}.$$

Alt.: X is subgaussian if $\exists\, \beta, \kappa$ s.t.

$$\Pr\{|X| \geq t\} \leq \beta e^{-\kappa t^2} \quad \forall t>0.$$

Examples:
(a) Gaussian, variance $\frac{1}{m}$.
(b) Bernoulli, $\pm\frac{1}{\sqrt{m}}$
(c) Any bounded dist$^{\text{bn}}$ with zero mean & var. $\frac{1}{m}$.
($\exists\, B>0$ s.t. $|X| \leq B$ w.p. 1. Then, $X \sim \text{sub}(B^2)$).

Buldygin-Kozachenko lemma
If $X \sim \text{sub}(c^2)$, then $E\{X\}=0$, $E\{X^2\} \leq c^2$.
When $E\{X^2\} = c^2$, we say that X is strictly subgaussian, and write it as $X \sim \text{ssub}(c^2)$.

Examples:
(i) If $X \sim \mathcal{U}(-1,1)$, then $X \sim \text{ssub}\left(\frac{1}{3}\right)$.
(ii) If X is s.t.

$$P\{X=1\} = \Pr\{X=-1\} = \frac{1-s}{2}$$
$$\Pr\{X=0\} = s, \quad s \in [0,1)$$

Then, for any $s \in \left[0, \frac{2}{3}\right]$, $X \sim \text{ssub}(1-s)$
for any $s \in \left(\frac{2}{3}, 1\right)$, X is NOT ssub.

Lemma: $A \in \mathbb{R}^{m\times N}$, $A_{ij} \overset{iid}{\sim} \text{ssub}\left(\frac{1}{m}\right)$.
Let $y = Ax$ for some $x \in \mathbb{R}^N$.
Then, for any $\varepsilon > 0$,

$$E\{\|y\|_2^2\} = \|x\|_2^2$$

$$\boxed{\Pr\left\{ \left| \|y\|_2^2 - \|x\|_2^2 \right| > \varepsilon \|x\|_2^2 \right\} \leq 2e^{-\frac{m\varepsilon^2}{\kappa^*}}}$$

where $\kappa^* = \frac{2}{1-\log 2} \approx 6.52$.

With the above result, the RIP thm. for Gaussian matrices also holds for $\text{ssub}\left(\frac{1}{m}\right)$ matrices, with

$$\kappa_2 = \frac{\delta^2}{2\kappa^*} - \frac{1}{\kappa_1}.$$

RIP result for subgauss. matrices often found in the literature:

Thm. 9.2 $A \in \mathbb{R}^{m\times N}$ entries subgaussian (c^2) with variance $\frac{1}{m}$. Then $\exists\, C>0$ dep. only on c s.t. the RIC δ_s of A (of order s) satisfies $\delta_s \leq \delta$ w.p. at least $1-\varepsilon$, provided

$$m \geq C\delta^{-2}\left(s \ln\left(\frac{eN}{s}\right) + \ln(2\varepsilon^{-1})\right).$$

Setting $\varepsilon = 2e^{-\frac{\delta^2 m}{2C}}$, yields $+\frac{\delta^2 m}{2C}$

$$\boxed{m \geq 2C\delta^{-2} s \ln\left(\frac{eN}{s}\right)}$$

guarantees $\delta_s \leq \delta$ w.p. at least $1-2e^{-\frac{\delta^2 m}{2C}}$.

Thm. 9.12 $A \in \mathbb{R}^{m\times N}$ subgauss. c^2. Then, $\exists\, C_1, C_2 > 0$ depending only on c s.t., for $0<\varepsilon<1$, if

$$m \geq C_1 s \ln\left(\frac{eN}{s}\right) + C_2 \ln(2\varepsilon^{-1})$$

then w.p. at least $1-\varepsilon$, every s-sparse $x \in \mathbb{R}^N$ is exactly recovered from $y = Ax$ via ℓ_1-minimization.

Thm. 9.13 $A \in \mathbb{R}^{m\times N}$ subgauss. c^2 with var. $\frac{1}{m}$ entries.
Then, $\exists\, C_1, C_2 > 0$ depending only on c, & constants $D_1, D_2 > 0$ s.t. for $0<\varepsilon<1$, if

$$m \geq C_1 s \ln\left(\frac{eN}{s}\right) + C_2 \ln(2\varepsilon^{-1}),$$

then the foll. holds w.p. $\geq 1-\varepsilon$ uniformly $\forall x \in \mathbb{C}^N$:
Given $y = Ax + e$ with $\|e\|_2 \leq \eta$ for some $\eta \geq 0$, a soln. $x^\#$ of

$$\min_{z \in \mathbb{R}^N} \|z\|_1 \quad \text{s.t.} \quad \|Az - y\|_2 \leq \eta$$

satisfies

$$\|x - x^\#\|_2 \leq D_1 \frac{\sigma_s(x)_1}{\sqrt{s}} + D_2 \eta$$

$$\|x - x^\#\|_1 \leq D_1 \sigma_s(x)_1 + D_2\sqrt{s}\,\eta.$$

Remark: Setting $\varepsilon = 2e^{-\frac{m}{2C_2}}$ $\Rightarrow$ stable, robust recovery of all vecs. via ℓ_1 min. w.p. $\geq 1-2e^{-m/2C_2}$ provided $m \geq 2C_1 s \ln\left(\frac{eN}{s}\right)$.

The fact that $\|Ax\|_2^2$ concentrates around $\|x\|_2^2$ is the key step in proving:

Johnson Lindenstrauss Lemma (JL Lemma):
Given $0<\varepsilon<1$ and a set Q of $|Q|$ pts $\in \mathbb{R}^N$, and given $m > \frac{24 \log |Q|}{\varepsilon^2}$, there is a Lipschitz fn. $f: \mathbb{R}^N \to \mathbb{R}^m$ s.t.

$$(1-\varepsilon)\|u-v\|_2^2 \leq \|f(u) - f(v)\|_2^2 \leq (1+\varepsilon)\|u-v\|_2^2$$

$\forall u, v \in Q$. Moreover, $f(\cdot)$ can be found in randomized polynomial time (in $|Q|$).

Remarks:
1. "Randomized polynomial time" means that can find f with $O(|Q|)$ complexity, but will need randomization. (That is, not a deterministic construction).
2. Many flavors, with slightly different bounds on m.
3. Trivial result if $m \geq N$: Just use the identity map.
The result is interesting only when $m < N$.
$\Longrightarrow$ We can find a way to REDUCE dimension (from N to m)

without changing the relative distance bet.. any pair of pts by more than an <u>ε fraction of the original distance</u>^{2}.

4. The geometry of the pts. in Q affects (determines) f, but not its <u>existence/non-existence</u>.

<u>Proof</u>: (S. Dasgupta & A. Gupta, "An elementary proof of a theorem by Johnson & Lindenstrauss", Random structures & algos, 22(1): 60-65, 2003.

Take $A \in \mathbb{R}^{m \times N}$, $A_{ij} \overset{iid}{\sim} \underline{\mathcal{N}(0, \frac{1}{m})}$. Let $f(u) = Au$.

Want t.s.t. by randomly choosing A, in $O(|Q|)$ time, will find A s.t.

$$(1-\varepsilon) \|u-v\|_2^2 \leq \|A(u-v)\|_2^2 \leq (1+\varepsilon)\|u-v\|_2^2$$

$\forall\ u, v \in Q$.

Consider $Q' = \{ q \in \mathbb{R}^N,\ q = u-v,\ u, v \in Q \}$.

We know that, if $A_{ij} \overset{iid}{\sim} \mathcal{N}(0, \frac{1}{m})$, from Lemma 1 in the previous class,

$$\Pr\{ \left| \|Aq\|_2^2 - \|q\|_2^2 \right| \geq \varepsilon \|q\|_2^2 \} \leq 2e^{-\frac{m\varepsilon^2}{12}}.$$

Thus, if $m > \frac{24 \log |Q|}{\varepsilon^2}$,

$$2e^{-\frac{m\varepsilon^2}{12}} \leq 2e^{-2\log|Q|} = \underline{\frac{2}{|Q|^2}}$$

By the union bound,

$$\Pr\{ \left| \|Aq\|_2^2 - \|q\|_2^2 \right| > \varepsilon \|q\|_2^2 \text{ for some } q \in Q' \}$$
$$\leq \binom{Q}{2} \frac{2}{|Q|^2} = \frac{|Q|(|Q|-1)}{2} \cdot \frac{2}{|Q|^2} = 1 - \frac{1}{|Q|}.$$

Thus, A "fails", i.e., $\exists$ a diff. $u-v$ that undergoes a large change in its norm when mult. by A, w.p. at most $1 - \frac{1}{|Q|}$.

$\Rightarrow$ A "succeeds" w.p. $\frac{1}{|Q|}$.

$\Rightarrow$ In randomized polynomial time $(O(|Q|))$, will find a matrix that succeeds. $\boxtimes$.