$$\boxed{\text{Chapter 7}}$$

Chapter 7

__Matrix deviation inequality and geometric consequences.__

- Matrix Deviation Inequality
- M^*-bound and escape theorem.

We would like to prove the following.

<u>Thm</u>: Let $A \in \mathbb{R}^{m \times n}$ have independent, isotropic, and σ-subGaussian rows. Then for any $T \subseteq \mathbb{R}^n$, we have:

$$(\star) \quad \mathbb{E} \sup_{x \in T} \Big| \|Ax\|_2 - \sqrt{m}\,\|x\|_2 \Big| \leq C \sigma^2 \gamma(T)$$

where $\gamma(T) = \mathbb{E} \sup_{x \in T} |\langle g, x \rangle|$ with $g \sim N(0, I)$

<u>Remark</u>:

- Clearly $\omega(T) \leq \gamma(T)$
- If $T = -T$, then $\omega(T) = \gamma(T)$
- $\omega(T) = \frac{1}{2}\omega(T-T) = \frac{1}{2}\gamma(T-T)$
- (HW) For any $T \subseteq \mathbb{R}^n$ and $y \in T$ it holds:
  $\frac{1}{3}\left[\omega(T) + \|y\|_2\right] \leq \gamma(T) \leq 2\left[\omega(T) + \|y\|_2\right]$

Observe

$$\frac{\mathbb{E}\|Ax\|_2}{\|x\|} = \mathbb{E}\sqrt{\sum_{i=1}^{m} \frac{\langle A_i, x\rangle^2}{\|x\|^2}}$$

independent σ-subGaussian

and $\mathbb{E}\langle A_i, x\rangle^2 = \langle xx^T, \mathbb{E}A_iA_i^T\rangle$

$= \|x\|_2^2$

concentration of norm

$$\Rightarrow (\sqrt{m} - C\sigma^2) \le \frac{\mathbb{E}\|Ax\|_2}{\|x\|_2} \le (\sqrt{m} + C\sigma^2)$$

$$\Rightarrow \quad -C\sigma^2\|x\|_2 \le \mathbb{E}\|Ax\|_2 - \sqrt{m}\|x\|_2 \le C\sigma^2\|x\|_2$$

Notice

$$\gamma(T) = \mathbb{E}\sup_{y\in T} |\langle g, y\rangle| \ge \mathbb{E}|\langle g, x\rangle| = \hat{C}\|x\|_2$$

So $\mathcal{A}$ is amounts to

$$\mathbb{E}\sup_{x\in T}\Big|\|Ax\|_2 - \mathbb{E}\|Ax\|_2\Big| \le C\sigma^2\gamma(T)$$

We will use Talagrand's comparison inequality. Define

$$W_x := \|Ax\|_2 - \sqrt{m}\,\|x\|_2$$

We must show

$$\|W_x - W_y\|_{\psi_2} \le C\sigma^2 \|x-y\|_2 \quad \forall x,y \in T.$$

We prove a few special cases and then prove the general case.

<u>Setting 1</u>: $\|x\|_2 = 1$, $y = 0$

We need to show

$$\big\| \|Ax\|_2 - \sqrt{m} \big\|_{\psi_2} \le C\sigma^2.$$

Observe $\|Ax\|_2 = \|w\|_2$ where $w \in \mathbb{R}^m$ has independent σ-SubGaussian coordinates with $\mathbb{E}\, w_i^2 = 1$ So we already proved this.

Setting 2: $\|x\|_2 = \|y\|_2 = 1$

We must show

$$\big\| \|Ax\|_2 - \|Ay\|_2 \big\|_{\psi_2} \leq C\sigma^2 \|x-y\|_2$$

Let's analyze the squares first

$$Z := \frac{\|Ax\|_2^2 - \|Ay\|_2^2}{\|x-y\|_2} = \frac{\langle A(x-y), A(x+y)\rangle}{\|x-y\|_2} = \langle Au, Av\rangle$$

where $u = \frac{x-y}{\|x-y\|_2}$, $v = x+y$.

Claim: $\mathbb{P}[|Z| \geq s\sqrt{m}] \leq 2\exp\left(-\frac{cs^2}{\sigma^4}\right)$

for any $0 < s \leq \sqrt{m}$

pf: Observe $Z = \sum_{i=1}^{m} \underbrace{\langle A_i, u\rangle \langle A_i, v\rangle}_{\text{independent}}$

and

$$\mathbb{E}\langle A_i, x-y\rangle \langle A_i, x+y\rangle = \mathbb{E}\left(\langle A_i, x\rangle^2 - \langle A_i, y\rangle^2\right) = 0$$

Notice $\langle A_i, u\rangle\langle A_i, v\rangle$ is subexponential with parameter $\|\langle A_i, u\rangle\|_{\psi_2} \cdot \|\langle A_i, v\rangle\|_{\psi_2} \leq 2\sigma^2$ ☒

Getting rid of the squares

Lemma: Let $x, y \in \mathbb{R}^n$ with $\|x\|_2 = \|y\|_2 = 1$.

Then

$$\big\| \|Ax\|_2 - \|Ay\|_2 \big\|_{\psi_2} \leq C\sigma^2 \|x-y\|_2.$$

pf: We want to prove

$$p(s) := \mathbb{P}\left[\frac{\big| \|Ax\|_2 - \|Ay\|_2 \big|}{\|x-y\|_2} \geq s \right] \leq 4\exp\left(-\frac{cs^2}{4\sigma^2}\right)$$

Case 1: $s \leq 2\sqrt{m}$. Observe

$$p(s) = \mathbb{P}\left[\frac{\big| \|Ax\|_2^2 - \|Ay\|_2^2 \big|}{\|x-y\|_2} \geq s(\|Ax\|_2 + \|Ay\|_2) \right]$$

$$= \mathbb{P}\left[|Z| \geq s(\|Ax\|_2 + \|Ay\|_2) \right]$$

$$\leq \mathbb{P}\left[|Z| \geq s\|Ax\|_2 \right]$$

$$\leq \underbrace{\mathbb{P}\left[|Z| \geq \frac{s\sqrt{m}}{2} \right]}_{p_1(s)} + \underbrace{\mathbb{P}\left[\|Ax\|_2 < \frac{\sqrt{m}}{2} \right]}_{p_2(s)}$$

We already prove $p_1(s) \le 2\exp\left(-\frac{cs^3}{\sigma^4}\right)$
[setting 2]

Moreover

$$p_2(s) = \mathbb{P}\left[\|Ax\|_2 < \frac{\sqrt{m}}{2}\right]$$

$$\le \mathbb{P}\left[\left|\|Ax\|_2 - \sqrt{m}\right| > \frac{\sqrt{m}}{2}\right]$$

setting 1

$$\le 2\exp\left(-\frac{cs^2}{\sigma^4}\right).$$

So $p(s) \le 4\exp\left(-\frac{cs^2}{\sigma^4}\right)$

Case 2: $s > 2\sqrt{m}$.

Recall $p(s) = \mathbb{P}\left[\frac{\left|\|Ax\|_2 - \|Ay\|_2\right|}{\|x-y\|_2} \ge s\right]$

Reverse Triangle Inequality

$$\Rightarrow \left|\|Ax\|_2 - \|Ay\|_2\right| \le \|A(x-y)\|_2$$

So

$$p(s) \leq \mathbb{P}\left[\|Au\|_2 \geq s\right]$$

$$\leq \mathbb{P}\left[\|Au\|_2 - \sqrt{m} \geq \frac{s}{2}\right]$$

$$\underset{\text{setting } t}{\leq} 2\exp\left(-\frac{cs^2}{\sigma^4}\right) \qquad \boxtimes$$

Now let's prove the general case

Fix $x, y \in \mathbb{R}^n$. Can rescale and assume

$$\|x\|_2 = 1 \text{ and } \|y\|_2 \geq 1$$

Define $\bar{y} = \frac{y}{\|y\|_2}$.

$$\Rightarrow \|W_x - W_y\|_{\psi_2} \leq \underbrace{\|W_x - W_{\bar{y}}\|_{\psi_2}}_{\leq C\sigma^2 \|x - \bar{y}\|_2} + \|W_{\bar{y}} - W_y\|_{\psi_2}$$

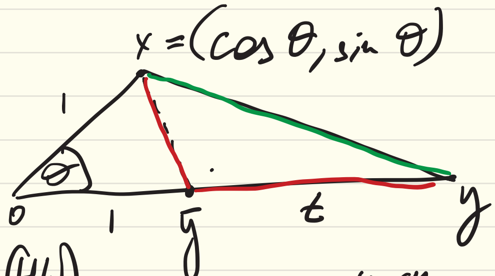
Notice

$$\|W_{\bar{y}} - W_y\|_{\psi_2} = \overbrace{\|\bar{y} - y\|_2}^{= \|y\| - 1} \|W_{\bar{y}}\|_{\psi_2}$$

We know: (Setting 1)

$$\|W_{\bar{y}}\|_{\psi_2} \leq C\sigma^2$$

So $$\|W_x - W_y\|_{\psi_2} \leq C\sigma^2(\|x-\bar{y}\|_2 + \|\bar{y}-y\|_2)$$
$$\leq C\sqrt{2}\sigma^2\|x-y\|_2$$



(HW)
Cor: Let $A \in \mathbb{R}^{m \times n}$ with A_i independent, isotropic, σ^2-subGaussian. Then for any $T \subseteq \mathbb{R}^n$ and $u > 0$, we have

$$\sup_{x \in T} \left| \|Ax\|_2 - \sqrt{m}\|x\|_2 \right| \leq C\sigma^2(w(T) + u\,\mathrm{rad}(T))$$

w.p. $1 - 2\exp(-u^2)$.

Some Consequences:

① Bounds on singular values

Let $T = S^{n-1}$ ← sphere. Then we learn

$$\sqrt{m} - C\sigma^2(\sqrt{n}+u) \leq \|Ax\|_2 \leq \sqrt{m} + C\sigma^2(\sqrt{n}+u)$$

$\forall x \in S^{n-1}$ w.p. $1-2\exp(-u^2)$

② Diameter of random projections

Define $P = \frac{1}{\sqrt{n}} A$.

Then

$$\mathbb{E}\,\mathrm{diam}(PT) \leq \sqrt{\frac{m}{n}}\,\mathrm{diam}(T) + C\sigma^2 S(T)$$

pf:

$$\mathbb{E} \sup_{x\in T} \|Ax\|_2 \leq \sqrt{m} \sup_{x\in T} \|x\|_2 + C\sigma^2 \gamma(T)$$

replace T by $T-T$ and divide by $\sqrt{n}$.

$\boxtimes$

③ Johson-Lindenstrauss

Let $S = \{x_i\}_{i=1}^{N} \subseteq \mathbb{R}^n$.

Set $T := \left\{ \frac{x-y}{\|x-y\|} : x, y \in S \right\}$

Recall $\gamma(T) \le C\sqrt{\log n}$. So

$$\sup_{x,y\in S} \left| \frac{\|Ax-Ay\|_2}{\|x-y\|_2} - \sqrt{m} \right| \le C\sqrt{\log N}$$

w.h.p.

$\Rightarrow$

$\forall x, y \in S$

$$\left(1 - C\sqrt{\frac{\log N}{m}}\right)\|x-y\|_2 \le \frac{1}{\sqrt{m}}\|Ax-Ay\| \le \left(1 + C\sqrt{\frac{\log N}{m}}\right)\|x-y\|_2$$

Accuracy ε requires $m \gtrsim \frac{\log N}{\varepsilon^2}$

Prop: Fix a set $T \subseteq \mathbb{R}^n$. Let $A \in \mathbb{R}^{m \times n}$ have independent, isotropic, σ-subGaussian rows A_i. Then with probability 0.99 we have

$$\|x-y\|_2 - \delta \le \frac{1}{\sqrt{m}} \|Ax - Ay\|_2 \le \|x-y\|_2 + \delta \quad \forall_{x,y \in T}$$

where
$$\delta = \frac{C\sigma^2 w(T)}{\sqrt{m}}$$

pf: Apply matrix deviation to $T-T$ ∎

Remark: Choose $m \geq \frac{C\sigma^4}{\varepsilon^2} d(T)$

$$d(T) = \frac{w(T)^2}{\mathrm{diam}(T)^2}$$

Then
$$\delta = C\varepsilon \, \mathrm{diam}(T).$$

Random Sections:

The following two consequences of the matrix deviation inequality will be the main tools for sparse recovery.

Thm (M^*-bound) Consider $T \subseteq \mathbb{R}^n$ and let $A \in \mathbb{R}^{m \times n}$ have independent, isotropic, σ-subGaussian rows. Then $E = \ker(A)$ satisfies

$$\mathbb{E}\, \mathrm{diam}(T \cap E) \leq \frac{C\sigma^2 w(T)}{\sqrt{m}}$$

pf: We'll actually prove

$$\mathbb{E} \sup_z \mathrm{diam}(T \cap (z+E)) \leq \frac{C\sigma^2 w(T)}{\sqrt{m}}$$

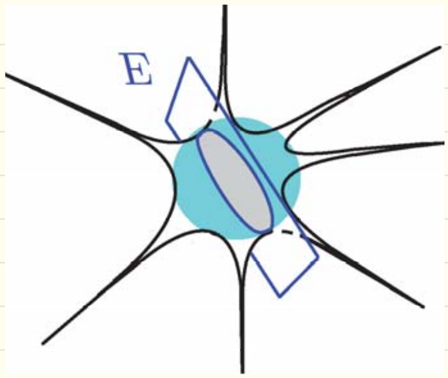
Matrix deviation with $T-T$ gives

$$\mathbb{E} \sup_{x,y \in T} \big| \|Ax - Ay\| - \sqrt{m}\|x-y\| \big| \leq C\sigma^2 \overbrace{\gamma(T-T)}^{=2w(T)}$$

$$\rotatebox{90}{$\leq$}$$

$$\mathbb{E} \sup_z \sup_{x,y \in (z+E)\cap T} \sqrt{m}\,\|x-y\| = \sqrt{m}\,\mathbb{E} \sup_z \mathrm{diam}((z+E)\cap T)$$

$\boxtimes$



Ex: Let $T = B_1^n$. Then

$$\mathbb{E}\,\mathrm{diam}(T \cap E) \le C\sqrt{\frac{\log n}{m}}$$

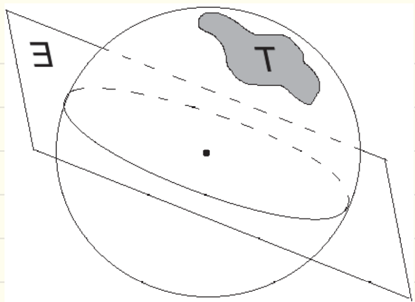
so if $m = \delta n$, then $\dim E = (1-\delta)n$

but $\mathbb{E}\,\mathrm{diam}(T \cap E) \le C\sqrt{\frac{\log n}{\delta n}}$

Remark: How big must m be to guarantee

$$\mathbb{E}\,\mathrm{diam}(T \cap E) \le \varepsilon\,\mathrm{diam}(T)$$

Answer: $m \ge C\frac{\sigma^4}{\varepsilon^2} \cdot \underbrace{d(T)}_{= \frac{w(T)^2}{\mathrm{diam}\,T}}$



<u>Thm</u> (Escape Thm) Consider $T \subseteq S^{n-1}$ and let $A \in \mathbb{R}^{m \times n}$ have independent, isotropic, σ-subGaussian rows. If

$$m \geq C \sigma^4 w(T)^2$$

then $E = \ker(A)$ satisfies $T \cap E = \emptyset$ w.p. $1 - 2\exp\left(\frac{-cm}{\sigma^4}\right)$.

<u>pf</u>: High Probability Matrix Deviation gives $\sup_{x \in T} \left| \|Ax\|_2 - \sqrt{m}\|x\|_2 \right| \leq C\sigma^2 (w(T) + u)$ w.p. $1 - \exp(-u^2)$. In this event, for and $x \in T \cap E$, it holds

$$\sqrt{m} \leq C\sigma^2 (w(T) + u)$$

Set $u = \frac{\sqrt{m}}{2C\sigma^2}$ and get a contradiction $\square$