$\boxed{\text{Chapter 5}}$

Quadratic forms and mean estimation

- mean estimation for $N(\mu, \Sigma)$
- Hanson-Wright inequality
- norm concentration w/out isotropy
- projections of random vectors.

As motivation, let's find the sample complexity of estimating the mean of a multivariate normal.

<u>Thm:</u> Suppose $x_i \overset{i.i.d.}{\sim} N(\mu, \Sigma)$. Then

$$\mathbb{P}\left[\left\|\frac{1}{n}\sum_{i=1}^{n} x_i - \mu\right\|_2 \le \sqrt{\frac{tr(\Sigma)}{n}} + \sqrt{\frac{2\log\left(\frac{2}{\delta}\right)\|\Sigma\|_2}{n}}\right] \ge 1-\delta \quad \forall \delta > 0$$

Read Lugosi-Mendelson survey!

pf: Fix a matrix $A \in S^d$ and $z \sim N(0, I)$

Then $\|Az\|_2^2 = z^T A^T A z = \sum_{i,j} (A^T A)_{ij} z_i z_j$

Write the eigenvector-decomposition

$$A = \sum_{i=1}^{d} \gamma_i u_i u_i^T$$

Then

$$z^T A^T A z = z^T\left(\sum_{i=1}^{d} \gamma_i^2 u_i u_i^T\right) z = \sum_{i=1}^{d} \gamma_i^2 \underbrace{\langle u_i, z\rangle}_{N(0,1)}{}^2$$

Notice $\{\langle u_i, z\rangle\}_{i=1}^{m}$ are independent.

Bernstein $\Rightarrow$

$$\mathbb{P}\left[\left|\|Az\|_2^2 - \|A\|_F^2\right| > t\right]$$

$$\leq 2\exp\left[-\frac{1}{2}\left(\frac{t^2}{\sum_i \gamma_i^4} \wedge \frac{4t}{\|A\|_2^2}\right)\right]$$

Finally set

$$A = \frac{1}{\sqrt{n}}\Sigma^{1/2}, \qquad z = \overbrace{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \Sigma^{-1/2}(x_i - \mu)}^{\sim \mathcal{N}(0, I)}$$

Then

$$\mathbb{P}\left[\left|\left\|\frac{1}{n}\sum_{i=1}^{n} x_i - \mu\right\|^2 - \frac{\operatorname{tr}(\Sigma)}{n}\right| > t\right]$$

$$\leq 2\exp\left[-\frac{1}{2}\left(\frac{t^2}{\|\Sigma\|_F^2/n} \wedge \frac{t}{\|\Sigma\|_2}\right)n\right] \qquad \blacksquare$$

We will now generalize results of this type to sub-Gaussians. Namely we want to understand

$$|x^T A x - \mathbb{E}\, x^T A x|$$

where the coordinates of $x = (x_1, \dots, x_n)$ are independent and subGaussian.

<u>Thm</u> (Hanson-Wright)
Let $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ be a random vector with x_i independent, mean zero, and σ-subGaussian. Let $A \in \mathbb{R}^{n \times n}$ be arbitrary. Then

$$\mathbb{P}\left[|x^T A x - \mathbb{E}\, x^T A x| \geq t\right] \leq 2\exp\left(-c \min\left(\frac{t^2}{\sigma^4 \|A\|_F^2}, \frac{t}{\sigma^2 \|A\|_2}\right)\right)$$

See Vershynin 6.1-6.2 for the proof

<u>Cor</u>: Let $B \in \mathbb{R}^{m\times n}$ be a fixed matrix. and let $X = (X_1, \dots, X_n) \in \mathbb{R}^n$ have independent σ-subGaussian coordinates with $\mathbb{E} X_i = 0$ and $\mathbb{E} X_i^2 = 1$

Then

$$\mathbb{P}\left[\left| \|Bx\|_2 - \|B\|_F \right| > t \right] \le 2 e^{-\frac{ct^2}{\sigma^4 \|B\|_2^2}}$$

<u>pf</u>: Apply Hanson-Wright with $A = B^T B$. Compute

- $x^T A x = \|Bx\|_2^2$, $\mathbb{E} x^T A x = \mathbb{E}\, tr(A x x^T) = tr(B^T B)$
$= \|B\|_F^2$

- $\|A\|_2 = \|B\|_2^2$, $\|A\|_F = \|B^T B\|_F \le \|B\|_2 \cdot \|B\|_F$

$$\Rightarrow \mathbb{P}\left[\left| \|Bx\|_2^2 - \|B\|_F^2 \right| \ge u \right] \le 2 \exp\left(-\frac{c}{\sigma^4} \left(\frac{u^2}{\|B\|_2^2 \|B\|_F^2} \wedge \frac{u}{\|B\|_2^2} \right) \right)$$

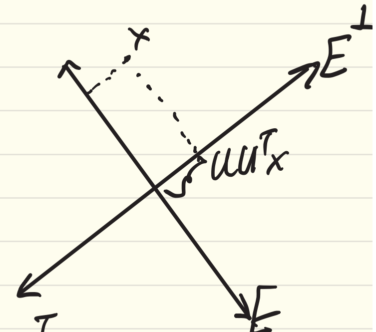
$\Rightarrow$ continue like in the concentration of the norm $\|x\|$. $\blacksquare$

Cor: Let $E \subset \mathbb{R}^n$ be a subspace with $\dim(E)=d$. Let $X=(X_1,\dots,X_n)$ have independent σ-subGaussian coordinates with $\mathbb{E}X_i=0$, $\mathbb{E}X_i^2=1$. Then

$$\mathbb{P}\left[\left|d(X,E)-\sqrt{n-d}\right|>t\right] \le 2\exp\left(-\frac{ct^2}{\sigma^4}\right)$$

pf: Let $U\in\mathbb{R}^{n\times(n-d)}$ have as columns an orthonormal basis for $E^\perp$. Then $\mathrm{Proj}_{E^\perp}=UU^T$ and

$$\mathrm{dist}(x,E)=\|UU^Tx\|_2$$



Then

$$\mathbb{E}\,\mathrm{dist}^2(X,E^\perp) = \mathbb{E}\, X^T UU^T UU^T X$$
$$= \mathbb{E}X^T UU^T X = \mathrm{tr}(UU^T) = n-d$$

$\Rightarrow \mathbb{E}\,\mathrm{dist}(X,E^\perp) \le \sqrt{n-d}$. Apply previous result $B=UU^T$ ∎