Lecture 5: Applications to Random Matrix Theory

Isabella Zhu

20 February 2025

§1 TLDR

Recall from last lecture we talked about $\max_{X \in B_2} v^T X$. We will use similar technique for a result in random matrix theory and also in *packing* (the dual to ϵ-nets).

§2 Subgaussian Random Vectors

When do we have $v^T X \sim \mathrm{subG}(\sigma^2)$ for $v \in B_2^d$?

Lemma 2.1

If $X = \begin{pmatrix} X_1 & X_2 & \dots & X_d \end{pmatrix}^T \in \mathbb{R}^d$ and $X_i \sim \mathrm{subG}(\sigma^2)$ independent, then for any $v \in B_2^d$, we have

$$v^T X \sim \mathrm{subG}(\sigma^2).$$

Proof. We have

$$\mathbb{E}[e^{v^T X}] = \mathbb{E}[e^{s \sum v_j X_j}] = \prod_{i=1}^{d} \mathbb{E}[e^{s v_j X_j}] \le \prod_{i=1}^{d} e^{s^2 v_j^2 \sigma^2} 2 \le e^{s^2 \sigma^2 / 2}.$$

§2.1 Application: Operator Norm of Random Matrix

Let matrix A be a $m \times n$ random matrix. Note A is a linear operator from $\mathbb{R}^n \to \mathbb{R}^m$. We equip distance metrics on both: l_p on $\mathbb{R}^n$ and l_q on $\mathbb{R}^m$.

What we care about is how much does our linear operator *distort distances*?

Definition 2.2. We have

$$||A||_{p \to q} = \sup_{x \in B_p^n} |Ax|_q$$

The **operator norm** is defined as $||A||_{2 \to 2}$.

Remark 2.3. Notice that for every $z \in \mathbb{R}^m$, we have $|z|_2 = \sup_{y \in B_2^m} y^T z$, so

$$||A||_{2\to 2} = \sup_{x \in B_2^n} |Ax|_2 = \sup_{y \in B_2^m, x \in B_2^n} y^T Ax.$$

Note the similarity of this problem to the ball stuff from lecture 4.

Consider a random matrix A such that $A_{i,j} \sim \text{subG}(\sigma^2)$. We take two quarter-nets. Let N_n be the $\frac{1}{4}$-net for B_2^n and let N_m be the $\frac{1}{4}$-net for B_2^m. Both have sizes at most 9^n and 9^m.

We have

$$||A||_{op} = \max_{x \in B_2^n} |Ax|_2 \le \max_{z \in N_n} |Az|_2 + \frac{1}{4} \max_{x \in B_2} |Ax|_2$$

Therefore,

$$\frac{3}{4}||A||_{op} \le \max_{z \in N_n} |Az|_2 = \max_{z \in N_n, y \in B_2^m} y^T Az$$

We do the same thing again, replacing y with w instead of x with z.

$$\frac{3}{4}||A||_{op} \le \max_{z \in N_n, w \in N_m} w^T Az + \frac{1}{4}||A||_{op}$$

Thus, we have

$$||A||_{op} \le 2 \max_{z \in N_n, w \in N_m} w^T Az \le \sigma\sqrt{n+m}.$$

We can also calculate the MGF

$$\mathbb{E}[e^{sw^T Az}] = \prod_{i,j} \mathbb{E}[e^{sw_i A_{ij} z_j}] \le \prod_{i,j} e^{\sigma^2 s^2 w_i^2 z_j^2/2} = e^{\sigma^2 s^2 |w|_2^2 |z|_2^2/2} \le e^{\sigma^2 s^2/2}$$

where we exploit subgaussianity of A_{ij} and use the fact that w and z are in the unit ball. To reiterate, the key takeaway is

$$\boxed{||A||_{op} \le \sigma\sqrt{m+n}}$$

with high probability.

§3 Packing

Definition 3.1. *(Packing)* Fix $K \subset \mathbb{R}^d$, $\epsilon > 0$, and distance metric $d(.,.)$. A set P is called an ϵ-**packing** of K w.r.t. $d(.,.)$ if

1. $P \subset K$
2. $d(z, z') \ge \epsilon$ for any two distinct points in P.

Packing is the dual of ϵ-nets. We will discuss this later.

Definition 3.2. *(Covering and packing numbers)* Fix $K \subset \mathbb{R}^d$, $\epsilon > 0$, and distance metric $d(.,.)$.

- The **covering number** N_ϵ of K is the size of its smallest ϵ-net.
- The **packing number** P_ϵ of K is the size of its largest ϵ-packing.

Proposition 3.3

We have $P_{2\epsilon} \leq N_\epsilon \leq P_\epsilon$.

We will first show that $P_{2\epsilon} \leq N_\epsilon$. For every point in $P_{2\epsilon}$, we will inject it into N_ϵ. For every $z \in P_{2\epsilon}$, choose $\pi(z) \in N_\epsilon$ such that $d(z, \pi(z)) \leq \epsilon$. For distance bounding reasons, $\pi(z)$ is injective.

Next we will show that $N_\epsilon \leq P_\epsilon$ by showing that P_ϵ is also an ϵ-net. Suppose it isn't, then there exists a point $x \in K$ which is distance larger than ϵ for any $z \in P_\epsilon$ so we can just add z into the packing, contradiction.