Notes for High-Dimensional Probability Second Edition by Roman Vershynin

Gallant Tsao

July 24, 2025

Contents

9 Deviations of Random Matrices on Sets

The main question in this chapter is: How does an $m \times n$ matrix act on a general set $t \subset \mathbb{R}^n$?

9.1 Matrix Deviation Inequality

Take an $m \times n$ random matrix X with independent, isotropic, and subgaussian rows. The concentration of the norm (Theorem 3.1.1) tells us that for any fixed vector $x \in \mathbb{R}^n$, the approximation

$$\|Ax\|_2 \approx \sqrt{m}\|x\|_2$$

holds with high probability.
Let's ask something more general: Is it true that with high probability, the equation above holds *simultaneously* for many vectors $x \in \mathbb{R}^n$? To quantify how many, pick some bounded set $T \subset \mathbb{R}^n$ and ask if the approximation holds simultaneously for all $x \in T$. It turns out that the maximal error is about $\gamma(T)$, the Gaussian complexity of T.

Theorem 9.1.1 (Matrix deviation inequality)**.** Let A be an $m \times n$ random matrix with independent, isotropic and subgaussian rows A_i. Then for any subset $T \subset \mathbb{R}^n$,

$$\mathbb{E}\left[\sup_{x \in T} \left| \|Ax\|_2 - \sqrt{m}\|x\|_2 \right|\right] \leq CK^2\gamma(T),$$

where $\gamma(T)$ is the Gaussian complexity from Section 7.5.3, defined as

$$\gamma(T) = \mathbb{E}\left[\sup_{x \in T} |\langle g, x\rangle|\right], \; g \sim N(0, I_n),$$

and $K = \max_i \|A_i\|_{\psi_2}$.

The plan is to deduce this from Talagrand's comparison inequality (Corollary 8.5.8). To do that, we just have to check the random process

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

indexed by vectors $x \in \mathbb{R}^n$ has subgaussian increments. Here is the claim:

Theorem 9.1.2 (Subgaussian increments)**.** Let A be an $m \times n$ random matrix with independent, isotropic and subgaussian rows A_i. Then the random process Z_x defined above has subgaussian increments:

$$\|Z_x - Z_y\|_{\psi_2} \leq CK^2\|x - y\|_2 \text{ for all } x, y \in \mathbb{R}^n,$$

here $K = \max_i \|A_i\|_{\psi_2}$.

Once we have proved this theorem, we plug it into Talagrand's comparison inequality (Exercise 8.37 (a)) and get

$$\mathbb{E}\left[\sup_{x \in T} |Z_x|\right] \leq CK^2\gamma(T)$$

which directly gives Theorem 9.1.1. So, all we have to do is prove Theorem 9.1.2 - and it is in fact easier since it's for fixed x and y.

Proof of Theorem 9.1.2. This argument will be a bit longer than usual, so we'll (hopefully) make it easier by starting with simpler cases and building up from there.
Step 1: Unit vector x and zero vector y. If $\|x\|_2 = 1$ and $y = 0$, the inequality in the theorem statement becomes

$$\left\| \|Ax\|_2 - \sqrt{m} \right\|_{\psi_2} \leq CK^2.$$

The random vector $Ax \in \mathbb{R}^m$ has independent, subgaussian coordinates $\langle A_i, x\rangle$, which satisfy

$$\mathbb{E}\left[\langle A_i, x\rangle^2\right] = 1$$

by isotropy. So, the equation above follows from the concentration of the norm (Theorem 3.1.1).
Step 2: Unit vectors x, y and the squared process. □