MATH888: High-dimensional probability and statistics Fall 2021

Lecture 9 — September 27, 2021

Sebastien Roch, UW-Madison *Scribe: Yuxin Sun*

1 Overview

In the last lecture we applied Bernstein's inequality to prove the concentration property of the vector with independent sub-gaussian coordinates. In this lecture, we will focus on the concentration of quadratic forms.

2 Conditional Expectation

The slides for this section (`https://people.math.wisc.edu/~roch/hdps/roch-hdps-slides9.pdf`) provides a quick review of basic facts of conditional expectation. These materials are based on Williams [Wil91], but the results can be found in any graduate-level probability textbook. The following fact will be widely used throughout this course.

Lemma 1. *Let Y and Z be independent random variables with $\mathbf{E}[Z] = 0$. Then, for every convex function F, one has*

$$\mathbf{E}[F(Y)] \leq \mathbf{E}[F(Y + Z)].$$

Proof. By the tower law of expectation and Jensen's inequality, we have that

$$\begin{aligned}
\mathbf{E}[F(Y + Z)] &= \mathbf{E}[\mathbf{E}[F(Y + Z) \mid Y]] \geq \mathbf{E}[F(\mathbf{E}[Y + Z \mid Y])] \\
&= \mathbf{E}[F(\mathbf{E}[Y \mid Y] + \mathbf{E}[Z \mid Y])] \\
&= \mathbf{E}[F(Y + \mathbf{E}[Z])] \\
&= \mathbf{E}[F(Y)].
\end{aligned}$$

□

3 Symmetrization

In this section, we briefly discuss the symmetrization technique. More details can be found in section 6.4 of [Ver18]. A random variable Z is *symmetric* if Z and $-Z$ have the same distribution. A simple example of a symmetric random variable is the well known *symmetric Bernoulli*, which takes values -1 and $+1$ with equal probability $1/2$ each, i.e,

$$\mathbf{Pr}[Z = 1] = \mathbf{Pr}[Z = -1] = 1/2.$$

Fact 2 ([Ver18])**.** *Let X be a random variable and Z be an independent symmetric Bernoulli random variable. Then,*

1. *Let X' be an independent copy of X. We have that $X - X'$ is symmetric.*

2. *If X is symmetric, then ZX has the same distribution of X.*

As an application, we will use the symmetrization technique to prove the following lemma.

Lemma 3. *Let X be a real random variable such that $\mathbf{E}[X] = 0$ and $|X| \le c$. Then, for every $\lambda \in \mathbb{R}$, we have that*

$$\mathbf{E}[\exp(\lambda X)] \le \exp(2\lambda^2 c^2).$$

Proof. Let X' be an independent copy of X and Z be an independent symmetric Bernoulli random variable. Then, we have that

$$\begin{aligned}\mathbf{E}[\exp(\lambda X)] &\le \mathbf{E}[\exp(\lambda(X - X'))] = \mathbf{E}[\exp(\lambda Z(X - X'))] = \mathbf{E}[\mathbf{E}[\exp(\lambda Z(X - X')) \mid X, X']] \\ &\le \mathbf{E}[\exp(\lambda^2 (X - X')^2/2)] \le \exp(2\lambda^2 c^2),\end{aligned}$$

where the first inequality comes from Lemma 1, the first equality comes from Fact 2, the second equality comes from the tower law of expectation, and the second inequality comes from the fact that $\mathbf{E}[\exp(\lambda Z)] \le \exp(\lambda^2/2), \forall \lambda \in \mathbb{R}$. □

4 Hanson-Wright Inequality

Theorem 4 (Hanson-Wright inequality)**.** *Let $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$ be a random vector with independent mean zero sub-gaussian coordinates X_i. Let $A \in \mathbb{R}^{n \times n}$. Then, for every $t \ge 0$, we have that*

$$\mathbf{Pr}\left[\left|X^T A X - \mathbf{E}[X^T A X]\right| \ge t\right] \le 2 \exp\left(-c \min\left(\frac{t^2}{K^4 \|A\|_F^2}, \frac{t}{K^2 \|A\|_2}\right)\right),$$

where $K = \max_i \|X_i\|_{\psi_2}$, $\|A\|_F$ and $\|A\|_2$ denote the Frobenius norm and spectral norm of matrix A respectively.

We briefly describe the proof idea in this section and will prove Theorem 4 in detail in the next two lectures. As many times before, the proof will be based on bounding the moment generating function of $X^T A X$. We will first apply the decoupling technique to replace the term $X^T A X$ to $X^T A X'$, where X' is an independent copy of X. Then, we will compare the moment generating functions of the decoupled $X^T A X'$ to the case where X is the Gaussian random vector.

References

[Ver18] R. Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.

[Wil91] David Williams. *Probability with martingales.* Cambridge university press, 1991.