MATH 7251: High-Dimensional Probability **Spring 2022**

Lecture 5

Name: Tyler LaBonte **Professor:** Konstantin Tikhomorov

1 Proof of Subgaussian Concentration

Recall the theorem from last time:

Theorem 1.1

If $X_1, X_2, \dots, X_n$ are mutually independent zero mean subgaussian variables (not necessarily identically distributed), then

$$\left\| \sum_{i=1}^n X_i \right\|_{\Psi_2}^2 \leq C \sum_{i=1}^n \|X_i\|_{\Psi_2}^2 . \tag{1}$$

Additionally, recall that if X is a subgaussian random variable, we have $\mathbb{E}[|X|^p] \leq (c\sqrt{p})^p$ for any $p \geq 1$.

Proof: Of the theorem. We start with $X_1, X_2, \dots, X_n$ independent and zero mean. The first thing to note is that the variables can be assumed to be symmetric without loss of generality. In particular, this implies that all odd moments are zeros. This is because for every subgaussian variable Y of zero mean, $\|Y\|_{\Psi_2}$ is equivalent up to a constant multiple with $\|Y - Y'\|_{\Psi_2}$ where Y' is an independent copy of Y (this is called the subgaussian symmetrization trick and can be proved as an exercise). Also without loss of generality, we can assume that $\sum_{i=1}^n \|X_i\|_{\Psi_2}^2 = 1$.

For every integer $p \geq 1$ and every i, we have

$$\mathbb{E}[|X_i|^p] \leq (\tilde{c} \|X_i\|_{\Psi_2} \sqrt{p})^p. \tag{2}$$

We want to get an upper bound on the even moments of the sum. We have

$$\mathbb{E}\left[\sum_{i=1}^n X_i\right]^p = \sum_{v \in V} \mathbb{E}\left[\prod_{j=1}^p X_{v(j)}\right]. \tag{3}$$

Here, V is the set of ordered p-tuples which contain every index an even number of times (zero is also possible). It corresponds to those expectations which might be nonzero – all the others are zero because the X_i are mutually independent and symmetric. For every representation of p as a sum of positive even integers $p = p_1 + p_2 + \cdots + p_m$ with $p_1 \geq p_2 \geq \cdots \geq p_m$, let $V(p_1, p_2, \dots, p_m)$ be the subset of all p-tuples $v \in V$ such that for some distinct indices $i_1, i_2, \dots, i_m$ we have for every k,

$$|\{\Sigma j \leq p : v(j) = i_k\}| = p_k. \tag{4}$$

This is a partition of V into mutually nonintersecting subsets for all admissible choices of $m \geq 1$ and positive even integers summing up to p. For example, $(5,3,3,5,3,3) \in V$ corresponds to $m = 2, p_1 = 4, p_2 = 2$, that is the subset $V(4,2)$.

For each subset of this form, we can write an upper bound on the expectation. Normalizing the variables, we have

$$\sum_{v \in V(p_1,p_2,\ldots,p_m)} \mathbb{E}\left[\prod_{j=1}^{p} X_{v(j)}\right] = \sum_{v \in V(p_1,p_2,\ldots,p_m)} \prod_{j=1}^{p} \left\|X_{v(j)}\right\|_{\Psi_2} \cdot \mathbb{E}\left[\prod_{j=1}^{p} \frac{X_{v(j)}}{\left\|X_{v(j)}\right\|_{\Psi_2}}\right]. \tag{5}$$

We can bound the last term from above by $\prod_{k=1}^{m} \mathbb{E}\big[Z_k^{p_k}\big]$ where $Z_1, Z_2, \ldots, Z_m$ are mutually independent 1-subgaussian variables. So we have

$$\prod_{k=1}^{m} \mathbb{E}\big[Z_k^{p_k}\big] \leq (\tilde{c}\sqrt{p_k})^{p_k}. \tag{6}$$

Combining these estimates, we get

$$\sum_{v \in V(p_1,p_2,\ldots,p_m)} \mathbb{E}\left[\prod_{j=1}^{p} X_{v(j)}\right] \leq \left(\prod_{k=1}^{m} (\tilde{c}\sqrt{p_k})^{p_k}\right) \sum_{v \in V(p_1,p_2,\ldots,p_m)} \prod_{j=1}^{p} \left\|X_{v(j)}\right\|_{\Psi_2}. \tag{7}$$

The expression at the front is not too bad, but we need to understand the second term. Instead of estimating it directly, we will compare it with the quantity

$$1 = \left(\sum_{j=1}^{n} \|X_j\|_{\Psi_2}^2\right)^{p/2}. \tag{8}$$

The idea is that we expand this power and compare the numerical coefficients of the summands with the coefficients of the original term. First, we need to contract the set V. Define $\tilde{V}(p_1, p_2, \ldots, p_m)$ as the set of all ordered $p/2$ tuples $\tilde{v} \in \{1, 2, \ldots, n\}^{p/2}$ such that for distinct indices $i_1, i_2, \ldots, i_m$ we have

$$|\{\Sigma j \leq p/2 : \tilde{v}(j) = i_k\}| = p_k/2. \tag{9}$$

For example, $(5,3,3,5,3,3) \in V(4,2)$ would correspond to $(5,3,3) \in \tilde{V}(4,2)$. One can check directly that

$$\left(\sum_{j=1}^{n} \|X_i\|_{\Psi_2}^2\right)^{p/2} \geq \sum_{\tilde{v} \in \tilde{V}(p_1,p_2,\ldots,p_m)} \prod_{j=1}^{p/2} \left\|X_{\tilde{v}(j)}\right\|_{\Psi_2}^2. \tag{10}$$

The crucial part of the argument is to consider a mapping

$$f : V(p_1, p_2, \dots, p_m) \to \tilde{V}(p_1, p_2, \dots, p_m) \tag{11}$$

which for every $v \in V(p_1, p_2, \dots, p_m)$ sets $\tilde{v} \coloneqq f(v)$ obtained from v by removing components repeated an eventh time. For example, suppose we have the 10-tuple (5, 16, 100, 5, 2, 5, 16, 2, 100, 5). If we apply f to this tuple, we obtain (5, 16, 100, 2, 5). Note that

$$\prod_{j=1}^{p/2} \left\| X_{(f(v))_j} \right\|_{\Psi_2}^2 = \prod_{j=1}^{p} \left\| X_{v(j)} \right\|_{\Psi_2}. \tag{12}$$

That is, the twice less representations correspond to the squares in the product. Further, since the function is of course not injective, we can estimate from above the preimage (basically estimating how many elements f maps to the same element). It can be checked that for every $\tilde{v} \in \tilde{V}(p_1, p_2, \dots, p_m)$,

$$|f^{-1}(\tilde{v})| = |\{v \in V(p_1, p_2, \dots, p_m) : f(v) = \tilde{v}\}| \le \binom{p}{p/2} \frac{(p/2)!}{(p_1/2)!(p_2/2)! \dots (p_m/2)!}. \tag{13}$$

Now, we can estimate what we wanted:

$$\sum_{v \in V(p_1, p_2, \dots, p_m)} \prod_{j=1}^{p} \left\| X_{v(j)} \right\|_{\Psi_2} \le \binom{p}{p/2} \frac{(p/2)!}{(p_1/2)!(p_2/2)! \dots (p_m/2)!} \sum_{\tilde{v} \in \tilde{V}(p_1, p_2, \dots, p_m)} \prod_{j=1}^{p/2} \left\| X_{\tilde{v}(j)} \right\|_{\Psi_2}^2. \tag{14}$$

But, by (8) and (10), the product and summation on the right is at most 1, so we can forget about it! Thus,

$$\sum_{v \in V(p_1, p_2, \dots, p_m)} \prod_{j=1}^{p} \left\| X_{v(j)} \right\|_{\Psi_2} \le \binom{p}{p/2} \frac{(p/2)!}{(p_1/2)!(p_2/2)! \dots (p_m/2)!}. \tag{15}$$

Together with (7), this implies

$$\sum_{v \in V(p_1, p_2, \dots, p_m)} \mathbb{E}\left[\prod_{j=1}^{p} X_{v(j)}\right] \le \binom{p}{p/2} \frac{(p/2)!}{(p_1/2)!(p_2/2)! \dots (p_m/2)!} \prod_{k=1}^{m} (\tilde{c}\sqrt{p_k})^{p_k}. \tag{16}$$

Using Stirling's Formula,

$$\sum_{v \in V(p_1, p_2, \dots, p_m)} \mathbb{E}\left[\prod_{j=1}^{p} X_{v(j)}\right] \le \hat{c}^p (\sqrt{p})^p. \tag{17}$$

Thus,

$$\sum_{v \in V} \mathbb{E}\left[\prod_{j=1}^{p} X_{v(j)}\right] \le \hat{c}^p (\sqrt{p})^p \cdot (\text{num of admissible choices of } \boldsymbol{p}). \tag{18}$$

By a stars-and-bars combinatorial argument, the latter number is loosely bounded by $\binom{2p}{p} \le 4^p$.

So finally,

$$\mathbb{E}\left[\left|\sum_{j=1}^{n} X_j\right|^p\right] = c_1^p(\sqrt{p})^p. \tag{19}$$

In a previous lecture we saw that this is a sufficient characterization of a subgaussian random variable. Thus,

$$\left\|\sum_{j=1}^{n} X_j\right\|_{\Psi_2}^2 \leq c_2 \sum_{j=1}^{n} \|X_j\|_{\Psi_2}^2 . \tag{20}$$

$\blacksquare$