Non-Asymptotic Theory of Random Matrices
Lecture 5: Subgaussian random variables

Lecturer: Roman Vershynin   Scribe: Yuji Nakatsukasa

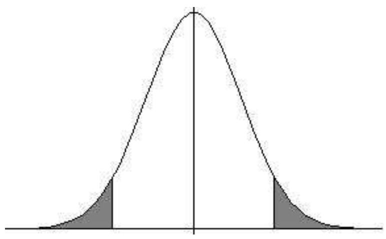
Thursday, January 18, 2007

1 Definition

The topic in this lecture is Subgaussian random variables. We start with the definition, and discuss some properties they hold.

Definition 1 (Subgaussian random variables). *A random variable X is subgaussian if $\exists c, C$ such that*

$$\mathbb{P}(|x| > t) \leq Ce^{-ct^2} \quad \forall t \geq 0. \tag{1}$$



As the name suggests, the notion of subgaussian random variables is a generalization of Gaussian random variables. Both the following well known random variables are subgaussian random variables (r.v's):

Example 2. *The following are examples of subgaussian random variables.*

1. Gaussian r.v's are subgaussian: $g \sim N(0,1) : \mathbb{P}(|g| > t) \leq e^{-t^2/2}$ $\forall t$.
2. Bounded r.v's, Bernoulli variables.

2 Properties

Let us recall Lecture 3. Using Lemma 6, definition (1) can be expressed equivalently in two other ways;

Lemma 3 ((Lecture 3,Lemma 6) Tails/Integrability/Moments)**.**

$$\begin{aligned}(1) &\Leftrightarrow \mathbb{E}\left(e^{c_2X^2}\right) \leq C_2 \\ &\Leftrightarrow (\mathbb{E}|X|^p)^{1/p} \leq C_3\sqrt{p}.\end{aligned}$$

The following Lemma shows that assuming further that the subgaussian r.v is mean zero, there is another equivalent description;

Lemma 4 (Moment Generating Function)**.** *Let X be a mean zero r.v. Then, the following are equivalent;*
(1) X is subgaussian.
(2) $\mathbb{E}e^{tX} \leq e^{ct^2} \quad \forall t \geq 0.$

Note that this is not true when $\mathbb{E}X \neq 0$ (e.g.$X \equiv 1$). Also note that (2) implies that $\mathbb{E}e^{tX} \simeq 1$ when t is small.

Proof. (i) Show (1) $\Rightarrow$ (2). Using Taylor expansion,

$$\mathbb{E}e^{tX} = 1 + t\mathbb{E}(X) + \sum_{k=2}^{\infty} t^k \frac{\mathbb{E}(X^k)}{k!}.$$

Since we are assuming the second term to be zero ($\mathbb{E}(X) = 0$),using Lemma 3 we obtain

$$\begin{aligned}\mathbb{E}e^{tX} &\leq 1 + \sum_{k=2}^{\infty} t^k \frac{(C_3\sqrt{k})^k}{k!} \\ &\leq 1 + \sum_{k=2}^{\infty} \left(\frac{C't}{\sqrt{k}}\right)^k.\end{aligned}$$

1)When $t \leq 1/C'$, since the sum will be smaller than a geometric series,

$$\mathbb{E}e^{tX} \leq 1 + C''t^2 \leq e^{ct^2}.$$

2)When $t \geq 1/C'$, we want to show

$$\mathbb{E}e^{(tX - ct^2)} \leq 1.$$

Here, since X is subgaussian, we know from Lemma 3 that

$$\mathbb{E}e^{c_2X^2} \leq C_2.$$

Here we claim we can set c so that $\mathbb{E}e^{(tX-ct^2)} \le \mathbb{E}e^{c_2X^2}(tX - ct^2 \le c_2X^2)$:

$$tx - ct^2 = -c(t - \frac{x}{2c})^2 + \frac{X^2}{4c} \le \frac{X^2}{4c}.$$

Therefore, by setting $c = 1/4c_2$, we obtain $tX - ct^2 \le c_2X^2$.

$$\therefore \mathbb{E}e^{(tX-ct^2)} \le \mathbb{E}e^{c_2X^2} \le C_2.$$

$$\therefore \mathbb{E}e^{tX} \le C_2e^{ct^2} \le e^{C'''t^2}.$$

The last inequality follows from the fact that t is not too small ($t > 1/C'$).

(ii)Show (2) $\Rightarrow$ (1).

$$\mathbb{P}(X > u) = \mathbb{P}(e^{tX} > e^{tu}) \le \frac{\mathbb{E}(e^{tX})}{e^{tu}},$$

Where we used Markov's inequality. Since we are supposing (2), we have

$$\frac{\mathbb{E}(e^{tX})}{e^{tu}} \le e^{ct^2-tu}.$$

Here optimize in t by setting $t = u/2c$. Then we have

$$\mathbb{P}(X > u) \le e^{-u^2/2c} = e^{-u^2/2c}.$$

□

Using Lemma 4, we can prove the following Theorem, which states that independent and mean-zero subgaussian random variables has another remarkable property (which is trivial in gaussian r.v's (if $\sum a_i^2 = 1$, $\sum a_i g_i = N(0,1)$)).

Theorem 5. *Let $X_1, X_2, \cdots, X_n$ be independent, mean-zero subgaussian random variables. Also let $a_1, a_2, \cdots a_n \in \mathbb{R}$ be such that $\sum_k a_k^2 = 1$. Then, $\sum_k a_k X_k$ is a subgaussian random variable.*

Proof.

$$\mathbb{E}e^{(t\sum_k a_kX_k)} = \mathbb{E}\prod_k e^{ta_kX_k} = \prod_k \mathbb{E}e^{ta_kX_k},$$

where we used the independence of X_k in the last equality. By Lemma 4, $\mathbb{E}e^{tX} \le e^{ct^2}$, for all $t \ge 0$, all k. Therefore,

$$\mathbb{E}e^{(t\sum_k a_kX_k)} \le \prod_k e^{ct^2a_k^2} = e^{\sum_k ct^2a_k^2} = e^{ct^2}.$$

□

We immediately have the following corollary;

Corollary 6. *Let $X_1, X_2, \cdots, X_n$ be independent, mean-zero subgaussian random variables. Then*

$$\mathbb{P}(|\sum_k a_k X_k| > t) \le Ce^{-ct^2/||a||_2^2}, \quad \forall t \ge 0.$$

In this corollary, if we think of a partial case when $X_k = \pm 1$ (Bernoulli r.v's), we obtain (set $a_k = 1/\sqrt{n}$);

$$\mathbb{P}\left(\frac{1}{\sqrt{n}}|\sum \pm 1| > t\right) \le e^{-t^2/2}$$

This is the Hoeffding inequality. This also verifies the Quantitative Central Limit Theorem.