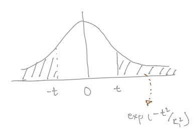
Ref: Vershynin's Book

Next few Weeks

- Sub-gaussian random variables & concentration
- Sub-gaussian rand. vectors & matrices
- Johnson-Lindenstrauss Lemma + Applications
- Matrix Deviation Ineq. + Applications

THM: (Sub-gaussian Properties) Let X be a random variable. Then, the following properties are equivalent - the parameters $K_i > 0$ appearing in these properties differ from each other by at most an absolute constant factor.



① Tails $\Pr\{|X| > t\} \leq 2\exp(-t^2/K_1^2)$ for all $t \geq 0$

② Moments $\|X\|_{L_p} := (\mathbb{E}|X|^p)^{1/p} \leq K_2\sqrt{p}$ for all $p \geq 1$

③ MGF of X^2 is bounded at some point
$\mathbb{E}\exp(X^2/K_3^2) \leq 2$

④ Also, if $\mathbb{E}X = 0$
$\mathbb{E}\exp(\lambda X) \leq \exp(\lambda^2 K_4^2)$ for all λ.

[Recall Moment Generating fn. (MGF) of r.v. Y
$M_Y(\lambda) := \mathbb{E}[\exp(\lambda Y)]$]

Proof: ① ⇒ ②

Assume wlog $K_1 = 1$

$$\mathbb{E}|X|^p = \int_0^\infty \Pr\{|X|^p > t\}\, dt \quad \text{(Tail-Sum Formula.)}$$
$$= \int_0^\infty \Pr\{|X| > u\}\, p\cdot u^{p-1}\, du \quad (t = u^p,\ dt = p\cdot u^{p-1} du)$$
$$\leq p\int_0^\infty 2\exp(-u^2)\, u^{p-1}\, du \quad \text{(via ①)}$$
$$= p\int_0^\infty \exp(-x)\cdot x^{(p/2 - 1)}\, dx \quad (\text{Substitute } x = u^2,\ dx = 2u\,du)$$
$$= p\cdot\Gamma(p/2) \leq p\cdot\left(\frac{p}{2}\right)^{p/2} \quad \text{(Defn. of Gamma fn. + Stirling Approx.)}$$

② ⇒ ③

$$\mathbb{E}\exp(X^2/K_3^2) = \mathbb{E}\left[1 + \sum_{p=1}^\infty \frac{1}{p!}\left(\frac{X^2}{K_3^2}\right)^p\right] \quad \text{Taylor Series Expansion}$$
$$= 1 + \sum_{p=1}^\infty \frac{1}{p!\,K_3^{2p}}\cdot\mathbb{E}X^{2p}$$
$$\leq 1 + \sum_{p=1}^\infty \frac{1}{p!\,K_3^{2p}}\cdot(2p)^p \quad \text{via ②}$$
$$\leq 1 + \sum_{p=1}^\infty \frac{e^p}{\cancel{p^p}}\cdot\frac{2^p\cdot\cancel{p^p}}{K_3^{2p}} \quad p! \geq \left(\frac{p}{e}\right)^p$$
$$= 1 + \sum_{p=1}^\infty \left(\frac{2e}{K_3^2}\right)^p$$

Series Converges for $K_3 > 2e$

③ ⇒ ①

Assume wlog $K_3 = 1$

$$\Pr\{|X| > t\} = \Pr\{e^{X^2} > e^{t^2}\}$$
$$\leq e^{-t^2}\cdot\mathbb{E}e^{X^2} \quad \text{(Markov's Ineq.)}$$
$$\leq 2\cdot e^{-t^2} \quad \text{(via ③)}$$

DEFN: A rand. var. X that satisfies one of the equivalent properties ①-④ is called a <u>sub-gaussian</u> random variable.

Examples.

① Gaussian $X \sim N(0, \sigma^2)$

$\|X\|_{\psi_2} \le C\sigma$

② Bernoulli $\|X\|_{\psi_2} \le \frac{1}{\sqrt{\ln 2}}$

③ Rademacher $\|X\|_{\psi_2} \le \frac{C}{\sqrt{\ln 2}}$

④ Bounded R.V. $\|X\|_{\psi_2} \le \frac{\|X\|_\infty}{\sqrt{\ln 2}}$

DEFN: The sub-gaussian norm of X, denoted by $\|X\|_{\psi_2}$ is defined to be the smallest K_3 in prop ③

$$\|X\|_{\psi_2} := \inf\{t : \mathbb{E}\exp(X^2/t^2) \le 2\}$$

Non Examples: Exponential, Heavy Tail Distribution, Poisson, Cauchy.

NOTE: $\|X\|_{\psi_2}$ is indeed a norm — verify via prop ② and Minkowski's ineq.

$$\|X+Y\|_{L^p} \le \|X\|_{L^p} + \|Y\|_{L^p}$$

Hence,

$$\|X+Y\|_{\psi_2} \le \|X\|_{\psi_2} + \|Y\|_{\psi_2}$$

Also,

$$\|\alpha X\|_{\psi_2} = \alpha \cdot \|X\|_{\psi_2}$$

Lemma [Centering] $\|X - \mathbb{E}X\|_{\psi_2} \le C \cdot \|X\|_{\psi_2}$

Proof: $\|X - \mathbb{E}X\|_{\psi_2} \le \|X\|_{\psi_2} + \|\mathbb{E}X\|_{\psi_2}$ Δ-ineq.

$\mathbb{E}X$ is a fixed number.

$$\begin{aligned}\|\mathbb{E}X\|_{\psi_2} &\le C.\ |\mathbb{E}X| \\ &\le C.\ \mathbb{E}|X| \quad \text{Jensen's ineq.} \\ &= C.\ \|X\|_{L_1} \\ &\le C.\ \|X\|_{\psi_2} \quad \text{via prop ②}\end{aligned}$$

Therefore,

$$\|X - \mathbb{E}X\|_{\psi_2} \le C'.\ \|X\|_{\psi_2}$$

Hoeffding's Ineq

THM Let $X_1, X_2, \dots, X_N$ be independent, mean-zero, sub-gaussian random variables. Then $\sum_{i=1}^{N} X_i$ is also a sub-gaussian r.v., and

$$\Big\|\sum_{i=1}^{N} X_i\Big\|_{\psi_2}^2 \le C.\ \sum_{i=1}^{N} \|X_i\|_{\psi_2}^2$$

where C is an absolute constant.

Independent sub-gaussians satisfy "Pythagoras' Thm" wrt the sub-gaussian norm

Proof: Consider MGF $S = \sum_{i=1}^{N} X_i$

$$\mathbb{E}\exp(\lambda S) = \mathbb{E}\exp\Big(\lambda \sum_{i=1}^{N} X_i\Big) = \mathbb{E}\prod_{i=1}^{N}\exp(\lambda X_i)$$

$$\underset{\text{ind}}{=} \prod_{i=1}^{N}\mathbb{E}\exp(\lambda X_i) \underset{\text{prop}}{\le} \prod_{i=1}^{N}\exp(\lambda^2 C\|X_i\|_{\psi_2}^2)$$

$$= \exp(\lambda^2 \cdot K^2)$$ where $K^2 = C \cdot \sum_{i=1}^{N} \|X_i\|_{\psi_2}^2$

$\square$

General Hoeffding's Ineq.

THM: let $X_1, \ldots, X_N$ be ind, mean-zero, sub-gauss. Then, for every $t \geq 0$, we have

$$\mathbb{P}\left\{ \left| \sum_{i=1}^{N} X_i \right| > t \right\} \leq 2\exp\left(\frac{-ct^2}{\sum_{i=1}^{N} \|X_i\|_{\psi_2}^2} \right)$$

THM: let $X_1, \ldots, X_N$ be ind, mean zero, sub-gauss, with $K = \max_i \|X_i\|_{\psi_2}$ and let $a = (a_1, \ldots a_N) \in \mathbb{R}^N$. Then, for every $t \geq 0$, we have

$$\mathbb{P}\left\{ \left| \sum_{i=1}^{N} a_i X_i \right| > t \right\} \leq 2\exp\left(\frac{-ct^2}{K^2 \cdot \|a\|^2} \right)$$

Sub-Exponential Distributions

Squares of gaussian r.vs are <ins>not</ins> gaussian, they are χ^2 r.vs.

$$g \sim N(0,1)$$

$$\mathbb{P}_r\{g^2 > t\} = \mathbb{P}_r\{g > \sqrt{t}\}$$

$$\sim \exp\left(\frac{-(\sqrt{t})^2}{2}\right)$$

$$= \exp(-t/2)$$

Heavier Tail

THM: For a rand. var. X, the following properties are equivalent

1. Tails: $\mathbb{P}_r\{|X| > t\} \leq 2\exp(-t/K_1)$ for all $t \geq 0$
2. Moments: $\|X\|_{L_p} := (\mathbb{E}|X|^p)^{1/p} \leq K_2 \cdot p$ for all $p \geq 1$
3. MGF of X: $\mathbb{E}\exp(X/K_3) \leq 2$
4. In addition, if $\mathbb{E}X = 0$, then these props are equivalent to
$$\mathbb{E}\exp(\lambda X) \leq \exp(\lambda^2 K_4^2) \quad \text{for} \quad |\lambda| \leq \frac{1}{K_4}$$

Proof: Vershynin's Book. (Section 2.7)

Defn: A random variable X that satisfies one of the equivalent properties ①-④ in the previous thm. is called a **sub-exponential** r.v.

The sub-exponential norm of X, denoted by $\|X\|_{\psi_1}$, is defined to be the smallest K_3 in prop. ③, i.e.,

$$\|X\|_{\psi_1} := \inf\{t : \mathbb{E}\exp(X/t) \leq 2\}.$$

Properties:

1. Any sub-gaussian r.v. is sub-exponential.
2. Square of a sub-gaussian is sub-exp.
$$\|X^2\|_{\psi_1} \leq \|X\|_{\psi_2}^2$$
3. Centering. $\|X - \mathbb{E}X\|_{\psi_1} \leq c \cdot \|X\|_{\psi_1}$ Proof: Exercise
4. Product of sub-gaussians is sub-exponential.
$$\|XY\|_{\psi_1} \leq \|X\|_{\psi_2} \cdot \|Y\|_{\psi_2}$$

Proof of ②: $K = \|X\|_{\psi_2}$

$$\mathbb{E}\exp(X^2/K^2) \leq 2$$

with $L := K^2$

$$Y = X^2$$

$$\mathbb{E}\exp(Y/L) \leq 2$$

$\boxed{\text{Proof of ④}}$ Assume, wlog, $\|X\|_{\psi_2} = \|Y\|_{\psi_2} = 1$

We will show that

$\mathbb{E}\exp(X^2) \leq 2$ & $\mathbb{E}\exp(Y^2) \leq 2$

implies $\mathbb{E}\exp|XY| \leq 2$

Using Young's ineq. $|ab| \leq \frac{a^2}{2} + \frac{b^2}{2}$ for all $a, b \in \mathbb{R}$

$$\mathbb{E}\exp|XY| \leq \mathbb{E}\exp\left(\frac{X^2}{2} + \frac{Y^2}{2}\right)$$

$$= \mathbb{E}\exp\left(\frac{X^2}{2}\right) \cdot \exp\left(\frac{Y^2}{2}\right)$$

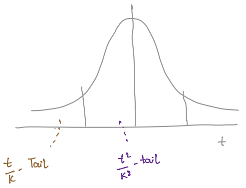
$$\le \frac{1}{2}\,\mathbb{E}\left[\exp\left(\frac{X^2}{2}\right) + \exp\left(\frac{Y^2}{2}\right)\right]$$

$$\le \frac{1}{2}\left[2+2\right] = 2 \quad \square$$

Bernstein's Ineq.

THM: Let $X_1, X_2 \ldots, X_N$ be ind, mean-zero, sub-exp, r.vs. Then, for any $t \geq 0$

$$\Pr\left\{\left|\sum_{i=1}^{N} X_i\right| > t\right\} \leq 2\exp\left(-c \cdot \min\left\{\frac{t^2}{\sum_i \|X_i\|_{\psi_1}^2}, \frac{t}{\max_i \|X_i\|_{\psi_1}}\right\}\right)$$



COR: (Avg of Sub-exp RVs)

Let $X_1, \ldots X_N$ be ind, mean-zero, sub-exp r.vs. Then, for every $t \geq 0$,

$$\Pr\left\{\left|\frac{1}{N}\sum_{i=1}^{N} X_i\right| > t\right\} \leq 2\exp\left(-c\,N \min\left\{\frac{t^2}{K^2}, \frac{t}{K}\right\}\right)$$

where

$$K = \max_i \|X\|_{\psi_1}$$

COR: (Weighted Sum of Sub-exp.)

Let $X_1, \ldots X_N$ be ind, mean-zero, sub-exp. r.vs and $a = (a_1, \ldots a_N) \in \mathbb{R}^N$. Then, for any $t \geq 0$,

$$\Pr\left\{\left|\sum_{i=1}^{N} a_i X_i\right| > t\right\} \leq 2\exp\left(-c \min\left\{\frac{t^2}{K^2\|a\|^2}, \frac{t}{K\|a\|_\infty}\right\}\right)$$

$$a = \left(\frac{1}{\sqrt{N}}, \frac{1}{\sqrt{N}} \cdots \frac{1}{\sqrt{N}}\right)$$

"CLT"

$$\Pr\left\{\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} X_i\right| \geq t\right\} \leq \begin{cases} 2\exp(-ct^2) & t \leq c\sqrt{N} \\ 2\exp(-t\sqrt{N}) & t > c\sqrt{N} \end{cases}$$