Lemma 2.12 (equivalence of subgaussian properties)

Let X be a random variable. Then the following properties are equivalent with parameters $K_i > 0$ differing from ... by at most an absolute constant:

1. Tails: $P(|X| > t) \leq \exp(1 - \frac{t^2}{K_1^2})$ for all $t \geq 0$
2. Moment: $(E|X|^p)^{\frac{1}{p}} \leq K_2 \sqrt{p}$ for all $p \geq 1$
3. Subexponential moment: $E \exp(\frac{X^2}{K_3^2}) \leq e$

Moreover, if $EX = 0$, $1, 2, 3 \Leftrightarrow$

4. Moment generating function:

$$E \exp(tX) \leq \exp(t^2 K_4^2) \quad \forall t \in \mathbb{R}$$

So far:

Homogeneity (同质), i.e. one may assume w.l.o.g that $K_i = 1$ and then show the other property with constant C

$1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 1$

$2 \Rightarrow 4$: Taylor series:

$$E\exp(tX) = 1 + tEX + \sum_{p=2}^{\infty} \frac{|t|^p EX^p}{p!}$$

$$= 1 + \sum_{p=2}^{\infty} \frac{|t|^p EX^p}{p!}$$

assume $K_2 = 1$

$$\leq 1 + \sum_{p=2}^{\infty} \frac{|t|^p \cdot p^{\frac{p}{2}}}{p!}$$

$p! \geq \left(\frac{p}{e}\right)^p$

$$\leq 1 + \sum_{p=2}^{\infty} \frac{|t|^p \cdot p^{\frac{p}{2}} \cdot e^p}{p^p}$$

$$= 1 + \sum_{p=2}^{\infty} \left(\frac{|t| e}{\sqrt{p}}\right)^p \quad (*)$$

We need to compare this with $\exp(K_4^2 t^2)$

$$\exp(K_4^2 t^2) = 1 + \sum_{k=1}^{\infty} \frac{(K_4 |t|)^{2k}}{k!}$$

$k! \leq k^k \rightarrow$

$$\geq 1 + \sum_{k=1}^{\infty} \left(\frac{K_4 |t|}{\sqrt{k}}\right)^{2k}$$

$$\underset{p=2k}{=} 1 + \sum_{p \in 2\mathbb{N}_+} \left(\frac{K_4 |t|}{\sqrt{\frac{p}{2}}}\right)^p$$

Observation: No odd terms need to control

$\left(\frac{e|t|}{\sqrt{p}}\right)^p$, $p \geq 3$ odd

If $\frac{e|t|}{\sqrt{p}} \le 1$

Then $\left(\frac{e|t|}{\sqrt{p}}\right)^p \le \left(\frac{e|t|}{\sqrt{p}}\right)^{p-1} \le \left(\frac{e|t|}{\sqrt{p-1}}\right)^{p-1}$

If $\frac{e|t|}{\sqrt{p}} > 1$

$$\left(\frac{e|t|}{\sqrt{p}}\right)^p \le \left(\frac{e|t|}{\sqrt{p}}\right)^{p+1} \le \left(\frac{e|t|}{\sqrt{\frac{p+1}{2}}}\right)^{p+1}$$

$$= \left(\frac{\sqrt{2}e|t|}{\sqrt{p+1}}\right)^{p+1}$$

$$\Rightarrow \left(\frac{e|t|}{\sqrt{p}}\right)^p \le \left(\frac{e|t|}{\sqrt{p-1}}\right)^{p-1} + \left(\frac{\sqrt{2}e|t|}{\sqrt{p+1}}\right)^{p+1}$$

$$\le 1 + \sum_{p\in 2\mathbb{N}_+} \left(\frac{e|t|}{\sqrt{p}}\right)^p + \sum_{p\in 2\mathbb{N}_+ +1} \left(\frac{e|t|}{\sqrt{p}}\right)^p$$

$$\le 1 + \sum_{p\in 2\mathbb{N}_+} \left(\frac{e|t|}{\sqrt{p}}\right)^p + \sum_{p\in 2\mathbb{N}_+ +1} \left(\frac{e|t|}{\sqrt{p-1}}\right)^{p-1} + \sum_{p\in 2\mathbb{N}_+} \left(\frac{\sqrt{2}e|t|}{\sqrt{p+1}}\right)^{p+1}$$

$$\overset{\substack{p'=p-1\\ p''=p+1}}{=} 1 + \sum_{p+2\mathbb{N}_+} \left(\frac{e|t|}{\sqrt{p}}\right)^p + \sum_{p'\in 2\mathbb{N}} \left(\frac{e|t|}{\sqrt{p'}}\right)^{p'} + \sum_{\substack{p''\in 2\mathbb{N}\\ p''>2}} \left(\frac{\sqrt{2}e|t|}{\sqrt{p+1}}\right)^{p+1}$$

odd $p''=2$

rename $p', p'' \to p$

$$\leq 1+\sum_{p\in 2\mathbb{N}} \left(2\left(\frac{e|t|}{\sqrt{p}}\right)^p + \left(\frac{\sqrt{2}e|t|}{\sqrt{p}}\right)^p\right)$$

$2+\sqrt{2}^p \leq (2+\sqrt{2})^p$

$$\leq 1+\sum_{p\in 2\mathbb{N}} \left(\frac{(2+\sqrt{2})e|t|}{\sqrt{p}}\right)^p$$

take $K_4 = (\sqrt{2}+1)e$

and hence $\leq \exp(K_4^2 t^2)$

$4 \Rightarrow 1$:

Markov: $\forall \lambda > 0$, $P(X \geq t) = P(e^{\lambda X} \geq e^{\lambda t}) = \dfrac{Ee^{\lambda X}}{e^{\lambda t}}$

prop 4 with $K_4 = 1$

$$\leq e^{-\lambda t + \lambda^2}$$

$\to \lambda = \frac{t}{2}$ (optimal) : $P\{X \geq t\} \leq e^{-\frac{t^2}{4}}$

same for $-X$: $P\{X \leq -t\} \leq e^{-\frac{t^2}{4}}$

$\Rightarrow P(|X|>t) \leq 2e^{-\frac{t^2}{4}} \leq e^{1-\frac{t^2}{4}}$

Remark:

(a) constants 1 and e in prop 1 and 3 are chosen for convenience, any number >0 or >1, respectively, will do.

(b) $4 \Rightarrow 1$ does not use $EX=0$

the condition only needed for necessities.

Def 2.13 A random variable that satisfies one of the equivalent conditions 1–3 in Lemma 2.12, is called a subgaussian random variable.

The subgaussian norm of X, denoted by $\|X\|_{\psi_2}$, is defined to be smallest K_2 in property 2,

In other words, $\|X\|_{\psi_2} = \sup_{p \geq 1} p^{-\frac{1}{2}} (E|X|^p)^{\frac{1}{p}}$.

① This is indeed a norm:

"$\Leftarrow$": $\|X\|_{\psi_2} = 0 \Leftrightarrow X = 0$ a.s.

"$\Rightarrow$": If not $X=0$ a.s. $\Rightarrow \exists \varepsilon$, s.t. $P(|X| > \varepsilon) = p > 0$

$$\|X\|_{\psi_2} \underset{p\geq 1}{\geq} E|X| \geq \varepsilon P(|X|\geq \varepsilon) > 0$$

② $$\|\lambda X\|_{\psi_2} = \sup_{p\geq 1} p^{-\frac{1}{2}} (E(|\lambda X|^p))^{\frac{1}{p}} = |\lambda| \, \|X\|_{\psi_2}$$

③ Triangle inequality:

$$\|X+Y\|_{\psi_2} = \sup_{p\geq 1} p^{-\frac{1}{2}} (E|X+Y|^p)^{\frac{1}{p}}$$

$$\overset{\text{Minkovski}}{\leq} \sup_{p\geq 1} p^{-\frac{1}{2}} \left((E|X|^p)^{\frac{1}{p}} + (E|Y|^p)^{\frac{1}{p}} \right)$$

$$\leq \sup_{p_1\geq 1} p_1^{-\frac{1}{2}} (E|X|^{p_1})^{\frac{1}{p_1}} + \sup_{p_2\geq 1} p_2^{-\frac{1}{2}} (E|Y|^{p_2})^{\frac{1}{p_2}}$$

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

The class of subgaussian r.v. on a given probability space is thus a normed vector space.

By lemma 2.12, there exist universal constant c, C

s.t. any subgaussian r.v. satisfies:

$$P(|X|>t) \le \exp\left(1-\frac{ct^2}{\|X\|_{\psi_2}^2}\right) \quad \text{for all } t>0$$

$$(E|X|^p)^{\frac{1}{p}} \le \|X\|_{\psi_2}\sqrt{p} \quad \text{for all } p\ge 1$$

$$E\exp\left(\frac{cX^2}{\|X\|_{\psi_2}}\right) \le e$$

(if $EX=0$)

$$E\exp(tX) \le \exp(Ct\|X\|_{\psi_2}^2) \quad \text{for all } t\in\mathbb{R}$$

Moreover, up to abso constant factors, $\|X\|_{\psi_2}$ is the smallest possible number in each of these inequalities.

Example

1. Gaussian

If X is a centered normal r.v. with variance σ^2,

then X is subgaussian with $\|X\|_{\psi_2} \leq C\sigma(1+\omega)$?

2. Bounded RV

Let X be such that $|X| \leq M$ a.s.

Then X is subgaussian with $\|X\|_{\psi_2} \leq M$

indeed $(E|X|^p)^{\frac{1}{p}} \leq M$ for all p

In parameter, a Rademacher r.v. $P(\varepsilon=1)=P(\varepsilon=-1)=\frac{1}{2}$

$\|\varepsilon\|_{\psi_2}=1$

Last time: Normed distribution is well-behaved:

One more reason: rotation variance: ~ makes it easier to work in H-D space

$\boxed{\text{vector of i.i.d. standard normed r.v. is invariant in distribution under rotations.}}$

extreme version:

Given a finite number of inde centered normal X_i, their weight $\sum_i a_i X_i$ is also a centered normed r.v. with variance $\|a\|_2^2$, implying, e.g., $Var(\sum_i X_i) = \sum_i Var(X_i)$

Proof idea: Multi of moment generating function, some words for subgaussian

Lemma 2.14 (Rotation invariance)

consider a finite number of independent, centered subgaussian r.v. X_i. Then $\sum_i X_i$ is also a centered subgaussian random variable.

Moreover, $\|\sum_i X_i\|_{\psi_2}^2 \le C \sum_i \|X_i\|_{\psi_2}^2$

where $C>0$ is an absolute constant.

Pf: One of the equivalent properties if $EX=0$

$E\exp(tX) \le \exp(Ct\|X\|_{\psi_2}^2)$ for all $t \in \mathbb{R}$

$E\exp(t\sum X_i) = E\prod_i \exp(tX_i) = \prod_i E\exp(tX_i) \le \prod \exp(Ct^2\|X_i\|_{\psi_2}^2)$

$= \exp(t^2K^2)$ where $K^2 = C\sum_i \|X_i\|_{\psi_2}^2$

By Lemma 2.12

$(4 \Rightarrow 2)$: $\| \sum_i X_i \|_{\psi_2} \leq C_1 K$

$\uparrow$

absolute constant.