Lecture 11: Concentration, moments, Orlicz norms

For a Gaussian $g \sim N(0,\sigma^2)$, the variance σ^2 coincides with its subGaussian parameter. More generally, the subGaussian parameter controls <u>all</u> moments.

<u>Thm</u>: For any RV X TFAE:

① (tails)

$\exists K_1$: $\quad P(|X| \geq t) \leq 2\exp(-t^2/K_1^2) \quad \forall t > 0$

② (moments)

$\exists K_2$: $\quad \|X\|_{L_p} \leq K_2 \sqrt{p} \quad \forall p \geq 1$

③ (MGF of X^2)

$\exists K_3$: $\quad \mathbb{E}\exp(X^2/K_3^2) \leq 2$

Moreover, if $\mathbb{E}X = 0$, then ①-③ are equivalent to:

④ (MGF) $\exists K_4$: $\mathbb{E}\exp(\lambda X) \leq \exp(\lambda^2 K_4^2) \quad \forall \lambda \in \mathbb{R}$

All $K_1, \ldots, K_4$ are proportional by numerical constant.

The proof is a bit long; see Vershynin.
The equivalence ③ defines a norm.

Defn: For any RV X, the subGaussian norm is

$$\|X\|_{\psi_2} \triangleq \inf\left\{K>0 : \mathbb{E}\exp(X^2/K^2) \le 2\right\}$$

Lemma (HW): $\|\cdot\|_{\psi_2}$ is a norm on $V = \{X : \|X\|_{\psi_2} < \infty\}$.

Using this norm, Hoeffding's inequality takes the simple form:

Thm: (Hoeffding Restated)
If $X_1, \dots, X_n$ are independent and subGaussian, then

$$\left\|\sum_{i=1}^n X_i\right\|_{\psi_2}^2 \le C \sum_{i=1}^n \|X_i\|_{\psi_2}^2$$

The nice thing about $\|X\|_{\psi_2}$ is that it makes sense for noncentered X.

<u>Lemma</u>: $\|X-\mathbb{E}X\|_{\psi_2} \le C\|X\|_{\psi_2}$

pf: By the triangle inequality,

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

Observe

$$\|\mathbb{E}X\|_{L^p} = |\mathbb{E}X| \le \mathbb{E}|X| \le \left(\mathbb{E}|X|^p\right)^{1/p} \le C\|X\|_{\psi_2}\sqrt{p}.$$

2

Therefore $\|\mathbb{E}X\|_{\psi_2} \le C\|X\|_{\psi_2}$ ☒

There is a completely parallel story for subexponential RV's.

Thm: For any RV X TFAE:

① (tails)

$\exists K_1$: $P(|X| \geq t) \leq 2\exp(-t/K_1)$ $\forall t > 0$

② (moments)

$\exists K_2$: $\|X\|_{L_p} \leq K_2\, p$ $\forall p \geq 1$

③ (MGF of X^2)

$\exists K_3$: $\mathbb{E}\exp(X/K_3) \leq 2$

Moreover, if $\mathbb{E}X = 0$, then ①-③ are equivalent to

④ (MGF)

$\exists K_4$: $\mathbb{E}\exp(\lambda X) \leq \exp(\lambda^2 K_4^2)$ $\forall |\lambda| \leq \frac{1}{K_4}$

All $K_1, \dots, K_4$ are proportional by numerical constants

Defn: For any RV X, the <u>subexponential norm</u> is

$$\|X\|_{\psi_1} \triangleq \inf\{K > 0 : \mathbb{E}\exp(|X|/K) \leq 2\}$$

Lemma (HW): $\|\cdot\|_{\psi_1}$ is a norm on

$V = \{X : \|X\|_{\psi_1} < \infty\}$

Lemma: $\|X - \mathbb{E}X\|_{\psi_1} \le C\|X\|_{\psi_1}$

The proof is similar to $\|\cdot\|_{\psi_2}$

Observe that ③ directly implies that the square of subGaussian is subexponential.

Lemma: $\|X^2\|_{\psi_1} = \|X\|_{\psi_2}^2$

More generally, the following is true:

Lemma: $\|XY\|_{\psi_1} \le \|X\|_{\psi_2} \cdot \|Y\|_{\psi_2}$

pf: WLOG, can assume $\|X\|_{\psi_2} = \|Y\|_{\psi_2} = 1$

Then

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

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

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

Therefore

$$\mathbb{E}\exp(|XY|) \leq \frac{1}{2}\Big[\underbrace{\mathbb{E}\exp(X^2)}_{\leq 1} + \underbrace{\mathbb{E}\exp(Y^2)}_{\leq 1}\Big]$$

$$\leq 1$$

Thus $\|XY\|_{\psi_1} \leq 1$, as claimed $\blacksquare$

The two norms $\|\cdot\|_{\psi_1}$, $\|\cdot\|_{\psi_2}$ capture two types of tail growth. We can capture more general tails by defining the <u>Orlicz Norm</u>

$$\|X\|_\psi = \inf\{K>0 : \mathbb{E}\,\psi(|X|/K) \leq 1\}$$

One can show that as long as ψ is convex, increasing, and

$$\psi(0)=0, \quad \psi(t) \to \infty \text{ as } t\to\infty$$

this indeed defines a norm on $\{X : \|X\|_\psi < \infty\}$

<u>Ex:</u> (L^p) $\psi(t) = t^p$ <u>Ex:</u> (ψ_2) $\psi_2(t) = e^{t^2}-1$

<u>Ex:</u> (ψ_1) $\psi_1(t) = e^t - 1$