

# SDS 384 11: Theoretical Statistics

## Lecture 6: Lipschitz continuous functions

---

Purnamrita Sarkar  
Department of Statistics and Data Science  
The University of Texas at Austin

# Recall-Lipschitz functions of Gaussian random variables

## Definition

A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is  $L$ -Lipschitz w.r.t the Euclidean norm if

$$|f(x) - f(y)| \leq L\|x - y\|_2 \quad \forall x, y \in \mathbb{R}^n$$

## Theorem (LG:Lipschitz functions of Gaussians)

Let  $(X_1, \dots, X_n)$  be a vector of iid  $N(0, 1)$  random variables. Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be  $L$ -Lipschitz w.r.t the Euclidean norm. Then  $f(X) - E[f(X)]$  is sub-gaussian with parameter at most  $L$ , i.e.  $\forall t \geq 0$ ,

$$P(|f(X) - E[f(X)]| \geq t) \leq e^{-\frac{t^2}{2L^2}}$$

- So a  $L$ -Lipschitz function of  $n$  gaussian random variables behave like a subgaussian with variance proxy  $L^2$ .

## Proof – (Courtesy Tao, Maurey and Pisier)

### Proof.

- WLOG assume  $E[F(X)] = 0$  and  $L = 1$ . Assume for simplicity that  $F$  is smooth
- We will just prove the upper tail  $P(F(X) \geq \lambda) \leq C \exp(-c\lambda^2)$ .
- All we need is

$$E[e^{tF(X)}] \leq e^{C't^2} \quad \text{for } t > 0 \quad (1)$$

- Lipschitz property implies the gradient  $|\nabla F(x)| \leq 1 \forall x \in \mathbb{R}^n$

## Proof contd.

### Proof contd.

- Consider an iid copy  $Y$ .
- Jensen's inequality implies  $E[e^{-tF(Y)}] \geq e^{-tE[F(Y)]} = 1$
- $E[e^{tF(X)}] \leq E[e^{t(F(X)-F(Y))}]$

$$F(X) - F(Y) = \int_0^{\pi/2} \frac{d}{d\theta} F(\underbrace{X \sin \theta + Y \cos \theta}_{X_\theta}) d\theta$$

- $$= \frac{\pi}{2} E_\theta [F'(X_\theta) \cdot X'_\theta]$$

$$e^{t(F(X)-F(Y))} \leq E_\theta \left[ e^{\frac{\pi}{2} t F'(X_\theta) \cdot X'_\theta} \right]$$

- $X'_\theta = X \cos \theta - Y \sin \theta$ . Also note that  $X_\theta, X'_\theta \stackrel{iid}{\sim} N(0, 1)$

## Proof contd.

### Proof contd.

- $$e^{t(F(X)-F(Y))} \leq \frac{2}{\pi} \int_0^{\pi/2} e^{\frac{\pi}{2} t F'(X_\theta) X'_\theta} d\theta$$
- $X'_\theta = X \cos \theta - Y \sin \theta$ . Also note that  $X_\theta, X'_\theta \stackrel{iid}{\sim} N(0, I_n)$   
$$E[e^{t(F(X)-F(Y))}] \leq \frac{2}{\pi} \int_0^{\pi/2} E[e^{\frac{\pi}{2} t F'(X_\theta) \cdot X'_\theta}] d\theta$$
- $$\begin{aligned} &= \frac{2}{\pi} \int_0^{\pi/2} E_{X_\theta} E_{X'_\theta} [e^{\frac{\pi}{2} t F'(X_\theta) \cdot X'_\theta} | X_\theta] d\theta \\ &\leq e^{\frac{\pi^2 t^2}{4}} \end{aligned}$$
- The last step is true because conditioned on  $X_\theta$ ,  $F'(X_\theta) \cdot X'_\theta \sim N(0, \sigma^2)$  where  $\sigma \leq 1$ .
- This proves Eq 1.

## Example 1

- Remember our friend chi square r.v.s? Consider  $\{X_i\}_{i=1}^n \stackrel{\text{iid}}{\sim} N(0, 1)$ .
- We proved that  $Y = \sum_i X_i^2$  is subexponential and we got the bound

$$P(|Y/n - 1| \geq \epsilon) \leq 2e^{-n\epsilon^2/8}.$$

- Lets try to prove a similar bound with the LG theorem.
- Let  $\underline{x} = (x_1, \dots, x_n)$  and  $f(\underline{x}) = \|\underline{x}\|_2$ .
- Note that Euclidian norm is 1-Lipschitz.
- So we have  $P(f(X) - E[f(X)] \geq t) \leq e^{-t^2/2}$  for  $t \geq 0$ .
- Since  $E[\sqrt{Y}] \leq \sqrt{E[Y]}$ , we have  $E[\sqrt{Y}] \leq \sqrt{E[Y]} = \sqrt{n}$ .
- $P(f(X) \geq E[f(X)] + t) \geq P(\sqrt{Y} \geq \sqrt{n} + t) = P(Y/n \geq (1 + \epsilon)^2)$
- Since  $(1 + \epsilon^2) \leq 1 + 3\epsilon$ ,  
$$e^{-n\epsilon_0^2/18} \geq P(Y/n \geq (1 + \epsilon_0/3)^2) \geq P(Y/n \geq 1 + \epsilon_0)$$

## Example 2: order statistics

### Example

Consider a sequence of independent r.v.s  $X = \{X_1, \dots, X_n\}$ . Let  $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}$ .  $P(|X_{(k)} - E[X_{(k)}]| \geq \epsilon) \leq 2e^{-\epsilon^2/2}$

### Proof.

- First note that  $|X_{(k)} - Y_{(k)}| \leq \|X - Y\|_2$ . (How?)
- So the order statistics are 1-Lipschitz.

