4

Where do we stand?

Rademacher chaos

$$\sum_{i \neq j} a_{ij} \varepsilon_i \varepsilon_j \quad \text{for } P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = \frac{1}{2}, \ \varepsilon_i \text{ i.i.d.}$$

Aim: generalize tail bounds for sums of i.i.d. r.v. to thm's scenario

Lemma 2.10 Let $\xi = (\xi_1, \cdots, \xi_M)$ be a sequence of independent random variables with $E\xi_j = 0$ for all $j = 1, \cdots, M$. Let A_{jk}, $j, k = 1, \cdots, M$, be a doubly indexed seq of ele in a vertex space X. Let $F: X \to \mathbb{R}$ be a convex func. Then

$$EF\Big(\sum_{j \neq k} \xi_j \xi_k A_{jk}\Big) \leq EF\Big(4 \sum_{j,k} \xi_j \underbrace{\xi_k'}_{\text{inde copy of } \xi_k} A_{jk}\Big)$$

Pf: introduce $\delta = (\delta_j)_{j=1}^{M}$ of inde r.v. δ_j with

$P(\delta_j = 0) P(\delta_j = 1) = \frac{1}{2}$

Then for $j \neq k$ $E\, \delta_k (1-\delta_j) = \frac{1}{4}$

This yields:

$$E := EF\Big(\sum_{j\neq k}^{M} \xi_j \xi_k A_{jk}\Big)$$

actually is condi expec

$$= EF\Big(4 \sum_{j\neq k}^{M} E\big(\delta_j (1-\delta_k)\big) \xi_j \xi_k A_{jk}\Big)$$

$$\overset{\text{Jensen}}{\leq} EF\Big(4 \sum_{j\neq k}^{M} \delta_j \xi_j (1-\delta_k) \xi_k A_{jk}\Big)$$

$E(E(F(4\sum \cdots)|\xi))$

Now let $\sigma(\delta) := \{ j \in \{1, \cdots, M\} : \delta_j = 1 \}$

Then, by Fubini

$$E \leq E\Big(\underbrace{E\Big(F\Big(4 \sum_{j\in\sigma(\delta)} \sum_{k\notin\sigma(\delta)} \xi_j \xi_k A_{jk}\Big)\Big|\, \delta\Big)}_{(*)}\Big)$$

Note that in (*), a r.v. ξ_j appears either only as the first factor (if $j \in \sigma(\xi)$) or only as the second factor ($j \notin \sigma(\xi)$)

So if we replace all ξ_k where $k \notin \sigma(\delta)$ by i.i.d ξ_k' Decoupling

- then the first factor never changes
- the value of the expectation is not changed

Hence we have

$$E \leq E(E(E(F(4 \sum_{j \in \sigma(\delta)} \sum_{k \notin \sigma(\delta)} \xi_j \xi_k' A_{j,k}) | \xi, \delta) | \delta))$$

$\exists \delta_0$ and $\sigma := \sigma(\delta_0)$ s.t.

$$E \leq E(E(F(4 \sum_{j \in \sigma(\delta_0)} \sum_{w \notin \sigma(\delta_0)} \xi_j \xi_k' A_{j,k}) | \xi))$$

Goal: introduce missing terms

use: $E\xi_i = 0$, pull out expec using Jensen

$$= E\left(E\left(F\left(4\sum_{j\in\sigma}\left(\sum_{k\notin\sigma}\xi_j\xi_k' A_{jk} + \sum_{k\in\sigma}\xi_j E\xi_k' A_{jk}\right)\Big|\xi\right)\right)\right)$$

$$\overset{\text{Jensen}}{\leq} E\left(E\left(F\left(4\sum_{j\in\sigma}\sum_{k=1}^{M}\xi_j\xi_k' A_{jk}\right)\Big|\xi\right)\right)$$

$$\overset{\text{Fubini}}{=} E\left(E\left(F\left(4\sum_{k=1}^{M}\sum_{j\in\sigma}\xi_j\xi_k' A_{jk} + \sum_{j\notin\sigma}(E\xi_j)\xi_k'\Big|\xi'\right)\right)\right)$$

$$\overset{\text{Jensen}}{\leq} E\left(F\left(4\sum_{j=1}^{M}\sum_{k=1}^{M}\xi_j\xi_k' A_{j,k}\right)\right)$$

Thm 2.11 (tail estimates for Ra chaos)

Let $A \in \mathbb{R}^{M \times M}$ be a symmetric matrix with zero diagonal and ε a Rademacher vector

Consider the Ra chaos:

$$X = \sum_{j,k=1}^{M} \varepsilon_j \varepsilon_k A_{jk}$$

Then

$$P(|X| \geq t) \leq \begin{cases} 2\exp\left(-\dfrac{3t^2}{128\|A\|_F^2}\right) & t \in \left(0, \dfrac{4\|A\|_F^2}{3\|A\|}\right] \\ 2\exp\left(-\dfrac{t}{32\|A\|}\right) & t > \dfrac{4\|A\|_F^2}{3\|A\|} \end{cases}$$

Pf: Define moment generating function:

$$E\exp(\theta X) = E\exp\left(\theta \sum_{j \neq k} \varepsilon_j \varepsilon_k A_{jk}\right)$$

A zero diag

$$\leq E \exp\left(4\theta \sum_{j,k=1}^{n} \varepsilon_j \varepsilon_k' A_{jk}\right)$$

Decoupling

(Lemma 1.10)

for $F = \exp(\theta \cdot)$

$$= E\Big(E\big(\underbrace{\exp\big(4\theta \sum_k \varepsilon_k' \sum_j \varepsilon_j A_{jk}\big)}_{\prod_{k=1}^{n} \exp\left(4\theta \varepsilon_k' \sum_{j=1}^{k} \varepsilon_j A_{jk}\right)} \big| \varepsilon\big)\Big)$$

$E(\exp(\theta \cdot y)) \leq \exp\left(\frac{\theta^2 y_{max}}{2}\right)$ where $y \leq y_{max}$ a.s.

$y_{max} = 4 \left| \sum_{j \neq k} \varepsilon_j A_{jk} \right|$

$$\leq \exp\left(8\theta^2 \sum_{k=1}^{n} \left(\sum_{j=1}^{n} \varepsilon_j A_{jk}\right)^2\right)$$

By symmetric of A: $\sum_k \left(\sum_j \varepsilon_j A_{jk}\right)^2 = \sum_k \sum_j \varepsilon_j A_{jk} \sum_e \varepsilon_e A_{ek}$

$$= \varepsilon^T A^2 \varepsilon$$

The matrix A^2 is $\begin{cases} \text{sym} \\ \text{posi semide} \end{cases}$

Again estimat moment generating func:

(now deco doesn't work since A^2 is not zero-diag)

$B = A^2$

$$E \exp(\kappa \varepsilon^* B \varepsilon)$$

κ and θ? $\kappa = 8\theta^2$ (see recall)

↑kappa

$$= E \exp\left(\kappa \sum_j B_{jj} + \kappa \sum_{j \neq k} B_{jk} \varepsilon_j \varepsilon_k\right)$$

decoupling

$$\leq \exp(\kappa \operatorname{tr}(B)) \, E \exp\left(4\kappa \sum \varepsilon_j \varepsilon_k' B_{jk}\right)$$

homework

$$\leq \exp(\kappa \operatorname{tr} B) \, E \exp\left(8\kappa^2 \sum_k \left(\sum_j \varepsilon_j B_{jk}\right)^2\right)$$

Now use positive semidef

$$\sum_k \left(\sum_j \varepsilon_j B_{jk}\right)^2 = \varepsilon^* B^2 \varepsilon \overset{B = PDP^*}{=\!=} \varepsilon^* P D^2 P^* \varepsilon$$

$$D_{ii} = \lambda_i(B) \qquad \sum \lambda_i^2(B) \cdot (P\varepsilon_i)^2 \leq \lambda_{max}(B) \varepsilon^* P D P^* \varepsilon$$

$$= \|B\| \, \varepsilon^* B \varepsilon$$

Hence

$$E \exp(k \varepsilon^* B \varepsilon) \le \exp(k \operatorname{tr} B)\, E \exp(8 k^2 \|B\| \varepsilon^* B \varepsilon)$$

$$= \exp(k \operatorname{tr}(B))\, E\left((\exp k \varepsilon^* B \varepsilon)^{8k\|B\|}\right)$$

$$\le \exp(k \operatorname{tr}(B)) \left[E \exp(k \varepsilon^* B \varepsilon)\right]^{8k\|B\|}$$

Jensen for $\phi(x) = x^{\frac{1}{8k\|B\|}}$ provided $8k\|B\| < 1$ concave

$$\Rightarrow E \exp(k \varepsilon^* B \varepsilon) \le \exp\left(\frac{k \operatorname{tr}(B)}{1 - 8k\|B\|}\right)$$

$$\forall\ 0 < k < \frac{1}{8\|B\|}$$

recall:

$$E(\exp(\theta X)) \le E(\exp(8\theta^2 \varepsilon B \varepsilon))$$

$\rightarrow$ specify $k = 8\theta^2$

$$\Rightarrow E \exp \theta \lambda \le \exp\left(\frac{8\theta^2 \operatorname{tr}(A^2)}{1-64\theta^2 \|A^2\|} \right) \quad \leftarrow B = A^2$$

$$0 < \theta < \frac{1}{8\|A\|}$$

use fact: $\|A^2\| = \sup_{x \in S^{n-1}} \langle x, A^2 x \rangle$

$$\ge \sup_{x \in S^{n-1}} \langle Ax, Ax \rangle$$

$$= \sup_{x \in S^{n-1}} \|Ax\|^2 = \|A\|^2$$

$$\operatorname{tr} A^2 = \operatorname{tr}\left(\sum_i A_{ji} A_{ik} \right)_{j,k}^2$$

$$= \sum_j \sum_i A_{ij} A_{ji} = \sum_{ij} A_{ij}^2 = \|A\|_F^2$$

$$\Rightarrow E\exp\theta X \le \exp\left(\frac{8\theta^2\|A\|_F^2}{1-64\theta^2\|A\|^2}\right) \quad 0<\theta<\frac{1}{8\|A\|}$$

Now assume $0<\theta<\frac{1}{16\|A\|}$

$\Rightarrow$ dominator $\ge 1-\frac{1}{4}=\frac{3}{4}$

$$P(X\ge t) = P(\exp\theta X \ge \exp\theta t)$$

$$\underset{\text{Markov}}{\le} e^{-\theta t} E\exp(\theta X)$$

$$\le \exp\left(-\theta t + \frac{8\theta^2\|A\|_F^2}{1-64\theta^2\|A\|^2}\right)$$

$$\le \exp\left(-\theta t + \frac{32}{3}\theta^2\|A\|_F^2\right)$$

opt in θ