Lecture 1

Plan: I. Fourier (1)
II. Def of $N(\mu, \Sigma)$
III. Properties

[diff class; HW + quiz]
[bad $N(\mu, \Sigma)$]

Def: Let $f: \mathbb{R}^d \to \mathbb{R}$. Then

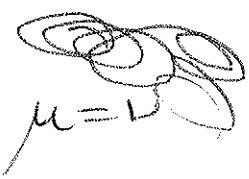
$$\hat{f}(\xi) = \int_{\mathbb{R}^d} e^{-i x \cdot \xi} f(x)\, dx, \quad \xi \in \mathbb{R}^d$$

Properties: (1) $f(x) = \int_{\mathbb{R}^d} e^{i x \cdot \xi} \hat{f}(\xi) \frac{d\xi}{(2\pi)^d}$

(2) $\widehat{\partial_{x_j} f}(\xi) = i \xi_j \hat{f}(\xi)$

(3) $\langle f, g \rangle_{L^2} = \int_{\mathbb{R}^d} f(x) g(x) dx = \int_{\mathbb{R}^d} \hat{f}(\xi) \hat{g}(\xi) d\xi = \langle \hat{f}, \hat{g} \rangle_{L^2}$

(4) μ, ν prob dist $\hat{\mu} = \hat{\nu} \Rightarrow \mu = \nu$



Def: Let $X \sim \mu$ be a random vector in $\mathbb{R}^d$

$\varphi_X(\xi) =$ "characteristic function of X at ξ"

$$= \mathbb{E}[e^{i X \cdot \xi}] = \int_{\mathbb{R}^d} e^{i x \cdot \xi} \mu(x)\, dx$$

$$= \hat{\mu}(\xi)$$

Def: Let $\mu \in \mathbb{R}^d$, $\Sigma \in Sym_{\geq 0}(d)$. Then

$$X \sim \mathcal{N}(\mu, \Sigma)$$

if

$$\varphi_X(\xi) = \exp\left\{ -i\mu \cdot \xi - \frac{1}{2} \xi^T \Sigma \xi \right\}$$

[Σ may be degenerate!]

Properties: $X = (X_1, \ldots, X_d)$

(1) $\mathbb{E}[X_j] = \mu_j$

(2) $\mathbb{E}[(X_j - \mu_j)(X_i - \mu_i)] = \Sigma_{ij}$

(3) If $A \in \mathbb{R}^{m \times d}$, $b \in \mathbb{R}^m$, then

$$AX + b \sim \mathcal{N}(A\mu + b, A\Sigma A^T)$$

(4) If $X \sim \mathcal{N}(0, I_d)$ and $O \in O(d)$ then

$$OX \overset{d}{=} X$$

(5) $X \sim \mathcal{N}(\mu, \Sigma) \iff v \cdot X \sim \mathcal{N}(v^T\mu, v^T\Sigma v)$ $\forall v$.

(6) If $Z = (X, Y) \sim \mathcal{N}(\mu, \Sigma)$ with

$$\mathbb{E}[XY^T] = 0$$

then X, Y are <u>independent</u>.

<u>Proof:</u>

(1) $$\mathbb{E}[X_j] = \mathbb{E}\left[i\,\partial_{\xi_j}\big|_{\xi=0}\, e^{-iX\cdot\xi} \right]$$

$$= i\,\partial_{\xi_j}\big|_{\xi=0}\, e^{-i\mu\cdot\xi - \frac{1}{2}\xi^T\Sigma\xi}$$

$$= \mu_j$$

(2) $$\mathbb{E}[X_i X_j] = -\mathbb{E}\left[\partial_{\xi_i}\partial_{\xi_j}\big|_{\xi=0}\, e^{-iX\cdot\xi} \right]$$

$$= -\partial_{\xi_i}\partial_{\xi_j}\big|_{\xi=0}\, e^{-i\mu\cdot\xi - \frac{1}{2}\xi^T\Sigma\xi}$$

$$= -\partial_{\xi_i}\big|_{\xi=0}\left[\left\{ -i\mu_j - (\Sigma\xi)_j \right\} e^{-i\mu\cdot\xi - \frac{1}{2}\xi\Sigma\xi} \right]$$

$$= \Sigma_{ij} + \mu_i\mu_j$$

(3) $$\mathbb{E}\left[e^{-i(AX+b)\cdot\xi} \right] = \mathbb{E}\left[e^{-i(AX)\cdot\xi} \right] e^{-ib\cdot\xi}$$

$$= \mathbb{E}\left[e^{-iX\cdot(A^T\xi)} \right] e^{-ib\cdot\xi}$$

$$= e^{-i\mu\cdot A^T\xi - \frac{1}{2}(A^T\xi)^T\Sigma A^T\xi - i b\cdot\xi}$$

$$= e^{-i(A\mu+b)\cdot\xi - \frac{1}{2}\xi^T(A\Sigma A^T)\xi}$$

(4) $\mathbb{E}[e^{-i(UX)\cdot\xi}] = \mathbb{E}[e^{-iX\cdot(U^T\xi)}]$

$$= e^{-\frac{1}{2}(U^T\xi)^T I U^T\xi} = e^{-\frac{1}{2}\|\xi\|^2}$$

(5) Trivial!

(6) ~~$\mathbb{E}[XY^T]$~~ Write

$$\Sigma = \mathbb{E}[ZZ^T] = \mathbb{E}[(X\ Y)(X\ Y)^T]$$

$$= \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{XY}^T & \Sigma_{YY} \end{bmatrix}$$

$\mathbb{E}[XY^T] = \Sigma_{XY} = 0$. So

$$\varphi_Z\begin{pmatrix}\xi_X\\ \xi_Y\end{pmatrix} = e^{-i\begin{pmatrix}\mu_X\\ \mu_Y\end{pmatrix}\cdot\begin{pmatrix}\xi_X\\ \xi_Y\end{pmatrix} - \frac{1}{2}\begin{pmatrix}\xi_X\\ \xi_Y\end{pmatrix}^T\begin{pmatrix}\Sigma_{XX} & 0\\ 0 & \Sigma_{YY}\end{pmatrix}\begin{pmatrix}\xi_X\\ \xi_Y\end{pmatrix}}$$

$$= \varphi_X(\xi_X)\,\varphi_Y(\xi_Y)$$

□

Existence:

$d=1$: Define $\nu(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Then if $X \sim \nu$

$$\frac{d}{d\xi}\varphi_X(\xi) = \frac{d}{d\xi}\int_{\mathbb{R}} e^{-ix\cdot\xi}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx$$

$$= \int_{\mathbb{R}}\left(-ix\, e^{-x^2/2}\right) e^{-ix\cdot\xi}\frac{dx}{\sqrt{2\pi}}$$

$$= \int_{\mathbb{R}}\left(i\partial_x e^{-x^2/2}\right) e^{-ix\cdot\xi}\frac{dx}{\sqrt{2\pi}}$$

$$= \int_{\mathbb{R}} e^{-x^2/2}(-i\partial_x)e^{-ix\cdot\xi}\frac{dx}{\sqrt{2\pi}}$$

$$= -\xi\,\varphi_X(\xi)$$

$$\Rightarrow \varphi_X(\xi) = e^{-\frac{1}{2}\xi^2}\underbrace{\varphi_X(0)}_{=1}.$$

$d \geq 1$: Set $\nu_d(x) \equiv \prod_{i=1}^{d}\nu(x_i)$. If $X \sim \nu_d$ the

$$\varphi_X(\xi) = \prod_{i=1}^{d}\varphi_{X_i}(\xi_i) = e^{-\|\xi\|^2/2} \Rightarrow X \sim \mathcal{N}(0, I).$$

Exs: $X \sim \mathcal{N}(0,1)$

(1) $Y = (X, X) \in \mathbb{R}^2$

Q: Is Y Gaussian?

A: Yes $Y = \begin{bmatrix} 1 \\ 1 \end{bmatrix} X \Rightarrow Y \sim \mathcal{N}\left(0, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right)$

(2) Set $Y = \varepsilon X$, $\varepsilon \in \{\pm 1\}$ iid indep of X

Q1: Is Y Gaussian?

Q2: Is $Z = (X, Y)$ Gaussian?

A1: Yes. $\mathbb{E}\left[e^{-i Y \cdot \xi}\right] = \mathbb{E}\left[e^{-i X (\varepsilon \xi)}\right]$.

$$= \mathbb{E}_\varepsilon \left[\mathbb{E}_X \left[e^{-i X (\varepsilon \xi)} \right] \right]$$

$$= \mathbb{E}_\varepsilon \left[e^{-\frac{1}{2}(\varepsilon \xi)^2} \right]$$

$$= e^{-\frac{1}{2}\xi^2} \quad \checkmark$$

Lecture 2: Plan: I. Recall

II. Conditional Expectation = Projection

III. Gaussian Conditioning = Linear

IV. Bayesian Inference

I. $\mu \in \mathbb{R}^d$, $\Sigma \in Sym_{\geq 0}(d)$

$X \sim N(\mu, \Sigma) \iff \mathbb{E}[e^{-i\xi \cdot X}] = e^{-i\mu \cdot \xi - \frac{1}{2}\xi^T \Sigma \xi}$

$\mathbb{E}[X_i] = \mu_i \quad Cov(X_i, X_j) = \Sigma_{ij}$

II. Suppose $(X, Y) \in \mathbb{R}^{l+m} \sim \nu$

Q. Given Y what is our best guess for X?

Warmup. Suppose $Z \overset{\in \mathbb{R}^k}{\sim} \nu$, $\mathbb{E}[f(Z)^2] < \infty$.

Lemma: $\mathbb{E}[f(Z)] = \operatorname{argmin}_c \mathbb{E}[(f(Z) - c)^2]$

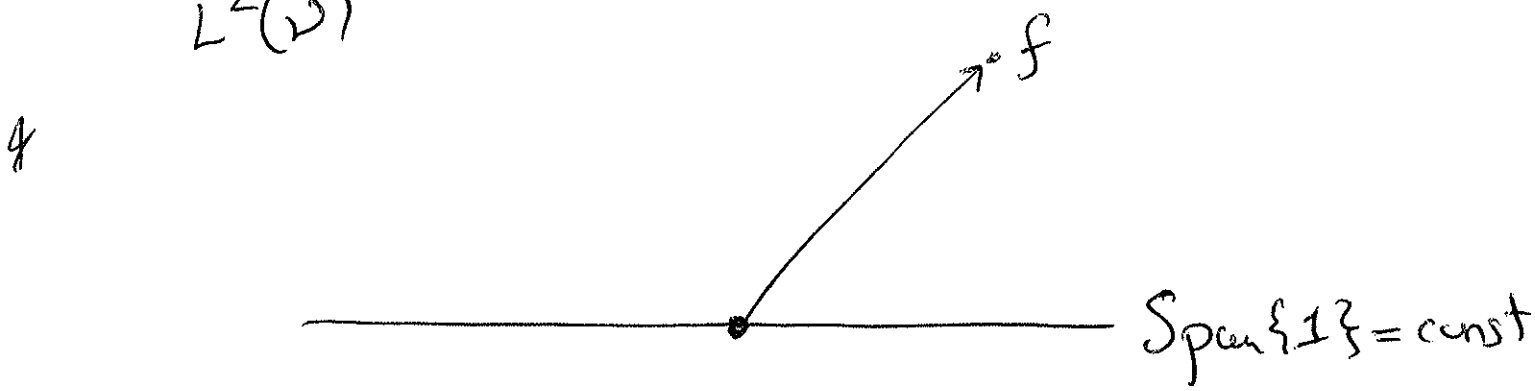
Pf: $\frac{d}{dc} \mathbb{E}[(f(Z) - c)^2] = 2\mathbb{E}[f(Z) - c]$ □

Interp: (1) $\mathbb{E}[f(Z)]$ is the best deterministic MSE guess for $f(Z)$

(2) Define

$$\begin{cases} \mathcal{I}^2(\nu) = \{ f : \mathbb{R}^k \to \mathbb{R} \mid \mathbb{E}[f(Z)^2] < \infty \} \\ \langle f, g \rangle_{L^2(\nu)} = \mathbb{E}[f(Z)g(Z)] = \int_{\mathbb{R}^k} f(z) g(z) \, d\nu(z) \end{cases}$$

$L^2(\nu)$



$$\mathbb{E}[(f(Z) - c)^2] = \| f - c \|^2_{L^2(\nu)}$$

$$\Rightarrow \mathbb{E}[f(Z)] = \inf_{c \in \mathrm{Span}\{1\}} \| f - c \|^2 = \Pi_{\text{const}} f$$

Main Event: $\mathbb{E}[\|Z\|^2] < \infty$

Def: Suppose $Z = (X, Y) \in \mathbb{R}^{k+m} \sim \nu$, $Y \sim \nu_Y$. Then

$$\mathbb{E}[X|Y] = \operatorname*{argmin}_{g \in L^2(\nu_Y)} \mathbb{E}[(X - g(Y))^2]$$

$$= \text{projection of } X \text{ onto } L^2(\nu_Y)$$

Properties: (i) $\mathbb{E}[(X - \mathbb{E}[X|Y])g(Y)] = 0 \quad \forall g$

(ii) $Z = (X, Y, V)$

$$\mathbb{E}[X|Y] = \mathbb{E}[\mathbb{E}[X|Y,V]|Y]$$ "tower property"

(iii) $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$

(iv) $\mathrm{Var}[X] = \mathbb{E}[\mathrm{Var}[X|Y]] + \mathrm{Var}[\mathbb{E}[X|Y]]$

Pf: (i) trivial

(ii) $\mathcal{L}^2(\nu_Y) \subseteq \mathcal{L}^2(\nu_{Y,V}) \Rightarrow \Pi_Y = \Pi_Y \Pi_{Y,V}$.

(iii) $Y \equiv \text{const}$

(iv) complete square

III. Suppose $(X, Y) \in \mathbb{R}^{p+m} \sim \mathcal{N}(0, \Sigma)$

$$\Sigma = \mathbb{E}[(XY)(XY)^T] = \begin{bmatrix} \mathbb{E}[XX^T] & \mathbb{E}[XY^T] \\ \mathbb{E}[YX^T] & \mathbb{E}[YY^T] \end{bmatrix}$$

$$= \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}$$

Q: What is $\mathbb{E}[X|Y]$

Lemma: There exists A s.t. $A\Sigma_{YY} = \Sigma_{XY}$ (4)

Pf: $\Sigma_{YY} A^T \overset{?}{=} \Sigma_{YX}$

$\Leftrightarrow \mathrm{Col}(\Sigma_{YY}) \supseteq \mathrm{Col}(\Sigma_{YX})$

$\Leftrightarrow$ ~~$\ker(\Sigma_{YY}) \subseteq \mathrm{Col}(\Sigma_{YX})^\perp$~~

$\Leftrightarrow \ker(\Sigma_{YY})^\perp \supseteq \mathrm{Col}(\Sigma_{YX})$

$\Leftrightarrow \ker(\Sigma_{YY}) \subseteq \mathrm{Col}(\Sigma_{YX})^\perp = \mathrm{Span}\{\mathbb{E}[YX_i]\}^\perp$

So take $v \in \ker(\Sigma_{YY})$. We have

$0 = v^T \Sigma_{YY} v = \mathbb{E}[(Y\cdot v)^2] \Rightarrow Y\cdot v = 0$

$\Rightarrow v \cdot \mathbb{E}[YX_i] = 0 \ \forall i \Rightarrow v \in \mathrm{Col}(\Sigma_{YX})^\perp$. ☐

Thm: $\mathbb{E}[X|Y] = AY \sim \mathcal{N}(0, \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX})$

Pf: $\mathbb{E}[X|Y]$ defined by

$$\mathbb{E}[(X - \mathbb{E}[X|Y])\, g(Y)] = 0 \quad \forall g \in L^2(\sigma_Y).$$

Idea: If $X - \mathbb{E}[X|Y] \perp Y$ then done.

Enough to have

$$\begin{cases} \mathrm{Cov}(X - \mathbb{E}[X|Y], Y) = 0 & (\#) \\ (X - \mathbb{E}[X|Y], Y) \quad \text{Gaussian} \end{cases}$$

Enough to have

$$\mathbb{E}[X|Y] = AY.$$

Constraint (#) gives

$$\mathrm{Cov}(X - AY, Y) = \mathrm{Cov}(X, Y) - \mathrm{Cov}(AY, Y)$$
$$= \Sigma_{XY} - A\Sigma_{YY} \qquad \square$$

II. Suppose we observe

(data x; unknown param)

$$\begin{cases} y = \underset{n\times p}{H}\underset{p\times 1}{w} + \underset{1\times 1}{\sigma}\underset{n\times 1}{\varepsilon} \leftarrow \mathcal{N}(0, I) \\ \text{"prior"} \quad w \sim \mathcal{N}(0, \tau^2 I) \end{cases}$$

If H is fixed then

$$(\varepsilon, y, w) \sim \mathcal{N}(0, \Sigma)$$

with

$$\Sigma = \begin{pmatrix} \Sigma_{yy} & \Sigma_{y\omega} & \Sigma_{y\varepsilon} \\ & \Sigma_{\omega\omega} & \Sigma_{\omega\varepsilon} \\ & & \Sigma_{\varepsilon\varepsilon} \end{pmatrix}$$

$$\Sigma_{yy} = \tau^2 HH^T + \sigma^2 I, \qquad \Sigma_{\omega\varepsilon} = 0$$

$$\Sigma_{y\omega} = \tau^2 H^T \qquad \Sigma_{y\varepsilon} = \sigma^2 I$$

$$\Sigma_{\omega\omega} = \tau^2 I$$

$$\Sigma_{\varepsilon\varepsilon} = I$$

$$E[\omega | y] = (H^T H + \lambda I)^{-1} H^T y$$

$$Cov(\omega | y) = \sigma^2 (H^T H + \lambda I)^{-1}$$

Lecture 3

Plan: I. CLT Statement ①
II. Intuition
III. Lindeberg CLT
IV. GP's & NN's.

I. Def: Suppose X_n, X are RVs. Then $X_n \xrightarrow{d} X$ if

$$\mathbb{E}[f(X_n)] \longrightarrow \mathbb{E}[f(X)] \quad \forall f \text{ bdd. } C^0$$

Thm: (CLT) Let X_i iid with $\mathbb{E}[X_i]=0$, $\mathbb{E}[X_i^2]=1$. Then

$$S_n \equiv \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \Rightarrow Z \sim \mathcal{N}(0,1) \quad \text{as} \quad n \to \infty.$$

Goal: Prove CLT 3 ways

- Lindeberg (today)
- Lévy (HW)
- Stein (Wed)

Q: Why $\mathcal{N}(0,1)$?

Intuition: $h(Z) \geq h(S_n) \quad \forall n$ (why?)

And X, Y iid symm $\Rightarrow h(X) \leq h\left(\frac{X+Y}{\sqrt{2}}\right)$ so

$$S_{2^k} = \frac{1}{\sqrt{2}}\left(S_{2^{k-1}} + S'_{2^{k-1}}\right) \Rightarrow h(S_{2^k}) \geq h(S_{2^{k-1}})$$

II. Thm: (Lindeberg) Suppose $\mathbb{E}[|X_1|^3] < \infty$. Then ② for every $f \in C^3(\mathbb{R})$ with $\|f'''\|_\infty < \infty$ we have

$$\left|\mathbb{E}[f(S_n) - f(Z)]\right| \le \frac{\|f'''\|_\infty}{6n^{3/2}} \left(\mathbb{E}[|X_1|^3] + \mathbb{E}[|Z|^3]\right)$$

Pf: Define for $k = 0, \dots, n$

$$\begin{cases} S_n^{(k)} = \frac{1}{\sqrt{n}}\left(Z_1 + \dots + Z_k + X_{k+1} + \dots + X_n\right) = W_k + \frac{1}{\sqrt{n}} Z_k \\ S_n^{(0)} = S_n, \quad S_n^{(n)} = Z \end{cases}$$

Then

$$f(S_n) - f(Z) = \sum_{k=1}^{n} f(S_n^{(k)}) - f(S_n^{(k-1)})$$

$$\Rightarrow \left|\mathbb{E}[f(S_n) - f(Z)]\right| \le \sum_{k=1}^{n} \left|\mathbb{E}\left[f(S_n^{(k)}) - f(S_n^{(k-1)})\right]\right|$$

We have

$$\left|\mathbb{E}\left[f\left(W_k + \tfrac{1}{\sqrt{n}} Z_k\right) - f\left(W_k + \tfrac{1}{\sqrt{n}} X_k\right)\right]\right|$$

$$= \Big|\mathbb{E}\Big[f(W_k) + \tfrac{1}{\sqrt{n}} Z_k f'(W_k) + \tfrac{1}{2n} Z_k^2 f''(W_k) + \tfrac{1}{6n^{3/2}} Z_k^3 f'''(\tilde{W}_k)$$

$$- \left(f(W_k) + \tfrac{1}{\sqrt{n}} X_k f'(W_k) + \tfrac{1}{2n} X_k^2 f''(W_k) + \tfrac{1}{6n^{3/2}} X_k^3 f'''(\hat{W}_k)\right)\Big]\Big|$$

$$\le \frac{\|f'''\|_\infty}{6n^{3/2}} \left(\mathbb{E}[|Z|^3] + \mathbb{E}[|X_1|^3]\right) \qquad \square$$

Write

$$F_X(z) = \mathbb{P}(X \le z).$$

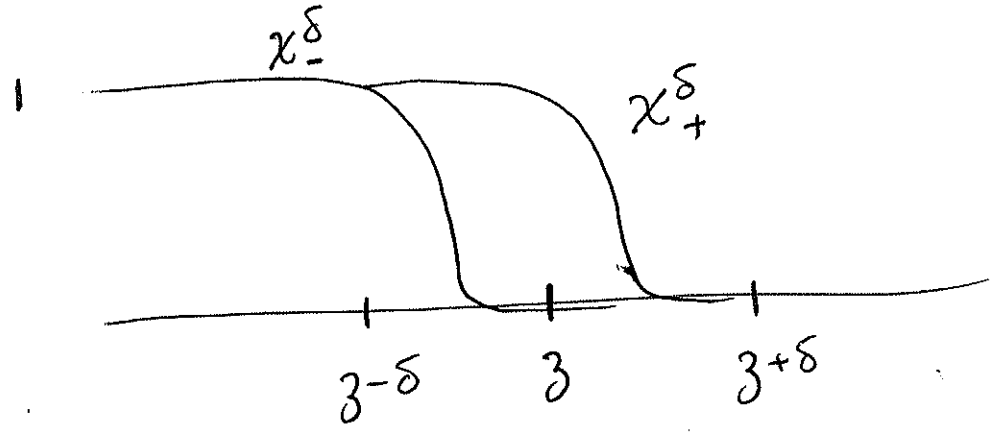
Lemma: Suppose $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ $\forall f$ s.t. f, f', f'', f''' bounded. Then

$$F_{X_n}(z) \longrightarrow F_X(z)$$

for all z where F_X is continuous.

Pf: Fix z and note

$$F_X(z) = \lim_{\delta \downarrow 0} F_X(z-\delta) = \lim_{\delta \downarrow 0} F_X(z+\delta).$$



$\chi_\pm^\delta \in C^\infty$

Note that

$$\mathbb{1}_{(-\infty, z-\delta]} \le \chi_-^\delta \overset{(*)}{\le} \mathbb{1}_{(-\infty, z]} \le \chi_+^\delta \le \mathbb{1}_{(-\infty, z+\delta]}$$

$$(*) \Rightarrow \quad \mathbb{E}[\chi_-^\delta(X_n)] \le F_{X_n}(z) \le \mathbb{E}[\chi_+^\delta(X_n)]$$

$$\Rightarrow \quad \mathbb{E}[X_-^\delta(X)] \le \lim_{n\to\infty} F_{X_n}(z) \le \mathbb{E}[X_+^\delta(X)]$$

$$\Rightarrow F_X(z-\delta) \le \lim_{n\to\infty} F_{X_n}(z) \le F_X(z+\delta)$$

$$\Rightarrow F_X(z) = \lim_{n\to\infty} F_{X_n}(z). \qquad \square$$

Lindeberg/Feller conditions (more general CLT)

Let $\{X_{i,n}\}_{i\le k_n}$ be a triangular array

$$\begin{array}{llll} X_{1,1} & \cdots & X_{k_1 1} & \\ X_{1,2} & \cdots & \cdots & X_{k_2 2} \\ \vdots & & & \vdots \end{array}$$

with $\{X_{i,n}\}_{i=1,\dots,k_n}$ independ (centered) for each fixed n.

Write

$$S_n^2 = \sum_{i=1}^{k_n} \mathrm{Var}(X_{n,i})$$

This array satisfies Lindeberg Condition if

$$\forall \varepsilon > 0 \qquad L_n(\varepsilon) \equiv \frac{1}{S_n^2} \sum_{i=1}^{k_n} \mathbb{E}\left[X_{n,i}^2 \, 1\!\!1_{\{|X_{n,i}| > \varepsilon S_n\}}\right] \to 0$$

(LC)

Thm: (LC) $\Rightarrow$ $\frac{1}{s_n}\sum_{i=1}^{k_n} X_{i,n} \xrightarrow{d} \mathcal{N}(0,1)$ as $n\to\infty$.

Moreover suppose

$$\lim_{n\to\infty} \max_{1\le k\le k_n} \frac{\mathrm{Var}[X_{i,n}]}{s_n^2} = 0 \qquad (FC)$$

Thm: (EC) & (LEC) $\Leftrightarrow$ CLT.

Thm: (MCLT) Suppose X_i are iid in $\mathbb{R}^d$ with

$$\mathbb{E}[X_i] = 0 \qquad \mathbb{E}[X_i X_i^T] = \Sigma \qquad \mathbb{E}[\|X_j\|^3] < \infty \ \forall j.$$

Then

$$S_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i \Rightarrow \mathcal{N}(0,\Sigma)$$

Pf: (ideas?)

Application: Let $X_1, X_2, \ldots$ iid with U/w density.

Set

$$B_i = 1\!\!1_{\{X_i > X_1,\ldots,X_{i-1}\}}$$

We have

$$\mathbb{P}(B_i = 1) = 1/i$$

Also B_i are independent!

Set

$$R_n = \sum_{i \le n} B_i$$

Thm: $H_n = \sum_{i=1}^{n} \frac{1}{i}$; $s_n^2 = H_n - \sum_{i \le n} \frac{1}{i^2}$. Then

$$\frac{R_n - H_n}{s_n} \Rightarrow \mathcal{N}(0,1)$$

Pf: 1) Lindeberg automatically true if $|X_{n,i}| \le C$.

2) Set $X_{n,i} = B_i - 1/i$ Then

$\mathbb{E}[X_{n,i}] = 0$ $\quad Var(X_{n,i}) = \frac{1}{i}(1 - \frac{1}{i}) \Rightarrow s_n^2 = H_n - \sum_{i=1}^{n} \frac{1}{i^2}$.

Lemma: B_i indep

Pf: $\rho_i = \#\{j \le i \mid X_j \le i\}$ = rank of X_i in first i

$B_i = 1\!\!1_{\{\rho_i = i\}}$. Note

$$(\rho_1, \dots, \rho_n) \xleftrightarrow{1:1} (1 \dots n) \dots$$

$r_i = \#\{j \le n \mid X_j \le X_i\}$

$\rho_i = \#\{j \le i \mid X_j \le X_i\}$

[filler] NNGP, Poisson

$\pi_n \in S_n$ uniform ; $\pi_n \xrightarrow{1:1} \rho = (\rho_1, \dots, \rho_n)$ $\square$

Lecture 4:

Plan: I. Stein's Lemma ①

↑ II. CLT mod OU

↳ III. OU intro

(IV. Poisson Limits)

I. Thm: Let $X_1, X_2, \ldots$ iid mean 0 variance 1, with $\mathbb{E}|X_1|^3 < \infty$. Then for every h with h, h', h'', h''' bounded we have

$$|\mathbb{E}[h(S_n)] - \mathbb{E}[h(Z)]| \leq \frac{\|h'''\|_\infty}{2\sqrt{n}} \mathbb{E}[|X_1|^3]$$

Goal: New Proof (+ Sketch $h''' \longrightarrow h'$)

Lemma: (Stein) A random variable $W \sim \mathcal{N}(0,1) \Leftrightarrow$

$$\mathbb{E}[Wf(W)] = \mathbb{E}[f'(W)]$$

for any bounded, absolutely continuous f.

Pf: ~~$\mathbb{E}[Wf(W)]$~~ Suppose $W \sim \mathcal{N}(0,1)$. Then

$$\mathbb{E}[Wf(W)] = \int W f(W) e^{-W^2/2} \frac{dW}{\sqrt{2\pi}}$$

$$= \int f(W) \left(-\partial_W e^{-W^2/2}\right) \frac{dW}{\sqrt{2\pi}}$$

$$= \int f'(W) e^{-W^2/2} \frac{dW}{\sqrt{2\pi}}$$

$$= \mathbb{E}[f'(W)]$$

Conversely, take

$$f(w) = e^{-i w \cdot \xi}$$

Then

$$-i\xi \varphi_W(\xi) = -i \varphi_W'(\xi) \quad \Rightarrow \quad \varphi_W(\xi) = e^{-\xi^2/2}$$ □

Exercise: $f: \mathbb{R}^d \to \mathbb{R}$ $\quad W \sim N(0, \Sigma)$ $\quad \Leftrightarrow$

$$\mathbb{E}[W f(W)] = \Sigma \, \mathbb{E}[\nabla f(W)]$$

Interp: Write $(A f)(w) = f'(w) - w f(w)$ (Stein operator)

Then $W \sim N(0,1) \Leftrightarrow \mathbb{E}[A f(W)] = 0 \quad \forall f$.

II. Now suppose we want to bound

$$\mathbb{E}[h(W)] - \mathbb{E}[h(Z)], \qquad Z \sim N(0,1).$$

Idea: For each $s \geq 0$ define

$$(T_s h)(w) \equiv \mathbb{E}_Z\left[h\left(e^{-s} w + \sqrt{1 - e^{-2s}}\, Z\right)\right]$$ "OU Semigroup"

Q: What is $\lim_{s \to 0} (T_s h)(w)$?

A: $T_0 h = h$!

Q: What is $\lim_{s \to \infty} (T_s h)(w)$

A: $= \mathbb{E}_Z[h(Z)]$.

Interp: T_s interpolates for Ito Gaussian average by "blurring" ω.

Claim: $(T_s(T_t h))(\omega) = (T_{s+t} h)(\omega)$

Pf: We have

$$T_s(T_t h)(\omega) = \mathbb{E}_Z\left[(T_t h)(e^{-s}\omega + \sqrt{1-e^{-2s}}\, Z)\right]$$

$$= \mathbb{E}_Z\left[\mathbb{E}_{Z'}\left[h\left(e^{-t}\{e^{-s}\omega + \sqrt{1-e^{-2s}}\, Z\} + \sqrt{1-e^{-2t}}\, Z'\right)\right]\right]$$

$$= \mathbb{E}_{Z,Z'}\left[h\left(e^{-(t+s)}\omega + \sqrt{e^{-2t} - e^{-2(s+t)}}\, Z + \sqrt{1-e^{-2t}}\, Z'\right)\right]$$

$$= \mathbb{E}_Z\left[h\left(e^{-(t+s)}\omega + \sqrt{1-e^{-2(s+t)}}\, Z\right)\right]$$

$$= (T_{t+s} h)(\omega) \qquad \square$$

Claim: Def: $\mathcal{L}u \equiv \lim_{s \downarrow 0} \frac{T_s u - u}{s}$ "generator"

Claim: $\mathcal{L}u(\omega) = u''(\omega) - \omega u'(\omega) = (\mathcal{A}u')(\omega)$

Pf: We have

$$T_s u(\omega) - u(\omega) = \mathbb{E}_Z\left[u(e^{-s}\omega + \sqrt{1-e^{-2s}}\, Z) - u(\omega)\right]$$

$$= \mathbb{E}_Z\Big[u\big((1-s)\omega + \sqrt{1-2s+o(s^2)}\,Z\big) - u(\omega)\Big] \quad (4)$$

$$= \mathbb{E}_Z\Big[u'(\omega)\big(-s\omega + \sqrt{s}\,(1+o(s))Z\big) + \tfrac{1}{2}u''(\omega)\,s\,Z^2 + o(s^{3/2})\Big]$$

$$= s\left(-\omega u'(\omega) + \tfrac{1}{2}u''(\omega)\right) + o(s^{3/2})$$

□

$$\partial_s T_s = \mathcal{L} T_s$$

III. Now we seek to estimate

$$\left|\mathbb{E}[h(S_n)] - \mathbb{E}[h(Z)]\right| = \left|T_0 h(\omega) - T_\infty h\right|$$

Idea: Find f_h so that

$$T_0 h(\omega) - T_\infty h(\omega) = f_h'(\omega) - \omega f_h(\omega) = \mathcal{A} f$$

$$-\int_0^\infty \partial_s T_s h$$

We have

$$\mathcal{A} f_h(\omega) = -\int_0^\infty \partial_s T_s h = -\mathcal{L} T_s h$$

We have

$$\mathcal{A} f_h = T_0 h - T_\infty h = -\int_0^\infty \partial_s T_s h \, ds$$

$$= -\mathcal{L} \int_0^\infty T_s h \, ds$$

$$= -\mathcal{A}\left(\left(\int_0^\infty T_s h \, ds \right)' \right)$$

$$\Rightarrow f_h'(w) = -\int_0^\infty \partial_w (T_s h)(w) \, ds$$

$$= -\int_0^\infty e^{-s} \, \mathbb{E}\left[h'(e^{-s} w + \sqrt{1-e^{-2s}} \, Z) \right] ds$$

We have

$$\|f\|_\infty \le \|h'\|_\infty, \quad \|f'\|_\infty \le \|h''\|_\infty, \quad \|f''\|_\infty \le \|h'''\|_\infty$$

Now we can complete proof of CLT.

$W = S_n$

$$\mathbb{E}[W f(W)] = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}[X_i f(S_n)]$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}\left[X_i f\left(W_i + \frac{1}{\sqrt{n}} X_i \right) \right]$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{E}\left[\frac{1}{\sqrt{n}} f'(W^{(i)})\right] + \mathbb{E}[X_i R_i] \overset{\leq}{} \frac{\|f''\|_\infty}{2} \frac{Y_i^2}{n}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[f'(W)\right] + O\left(\frac{\|f''\|_\infty \mathbb{E}|X_1|^3}{2\sqrt{n}}\right)$$

Lecture 5

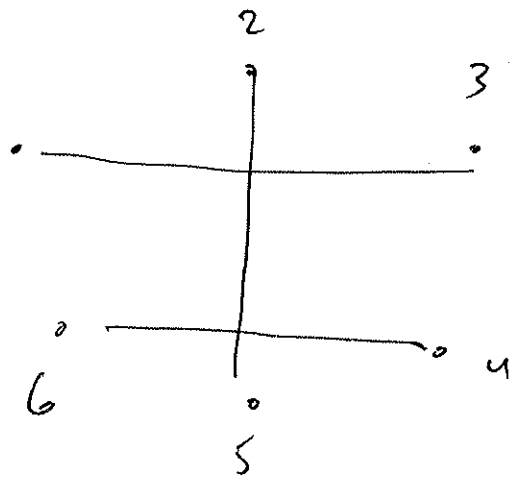
Plan: I. Wick
II Cumulant Intro
III. GOE Def
IV. Wick for GOE

Def: Let $k \geq 1$ even. A pair partition of $[k] = \{1, \dots, k\}$ is a partition $\pi \in S_k$ with $k/2$ blocks of size 2.

eg: $k = 6$

"blocks"

$\pi = \{(13), (25), (46)\}$



Thm: (Wick)
Suppose $X = (X_1, \dots, X_d) \sim \mathcal{N}(0, \Sigma)$. Then

$$\mathbb{E}[X_1 \cdots X_d] = \sum_{\pi \in PP(d)} \prod_{(ij) \in \pi} \Sigma_{ij}$$

Induction $d = 1, 2$ trivial

Pf: $\mathbb{E}[X_1 \cdots X_d] = \left(\prod_{k=1}^{d} i \partial_{\xi_k} \mathbb{E}\left[e^{-iX \cdot \xi}\right] \right) \Big|_{\xi = 0}$

$$= \prod_{k=1}^{d} i\partial_{\xi_k} \; e^{-\frac{1}{2}\xi^T \Sigma \xi} \Big|_{\xi=0}$$

$$= \left[\left(\prod_{k=1}^{d-1} i\partial_{\xi_k} \right) -i (\Sigma \xi)_d \right] e^{-\frac{1}{2}\xi^T \Sigma \xi} \Big|_{\xi=0}$$

$$= \left[\sum_{m=1}^{d-1} \Sigma_{md} \prod_{k \neq i,d} i\partial_{\xi_k} \; e^{-\frac{1}{2}\xi^T \Sigma \xi} \right] \Big|_{\xi=0}$$

$$= \sum_{m=1}^{d-1} \Sigma_{md} \sum_{\pi \in PP([d] \setminus \{m,d\})} \prod_{(i,j)\in\pi} \Sigma_{ij} \qquad \square$$

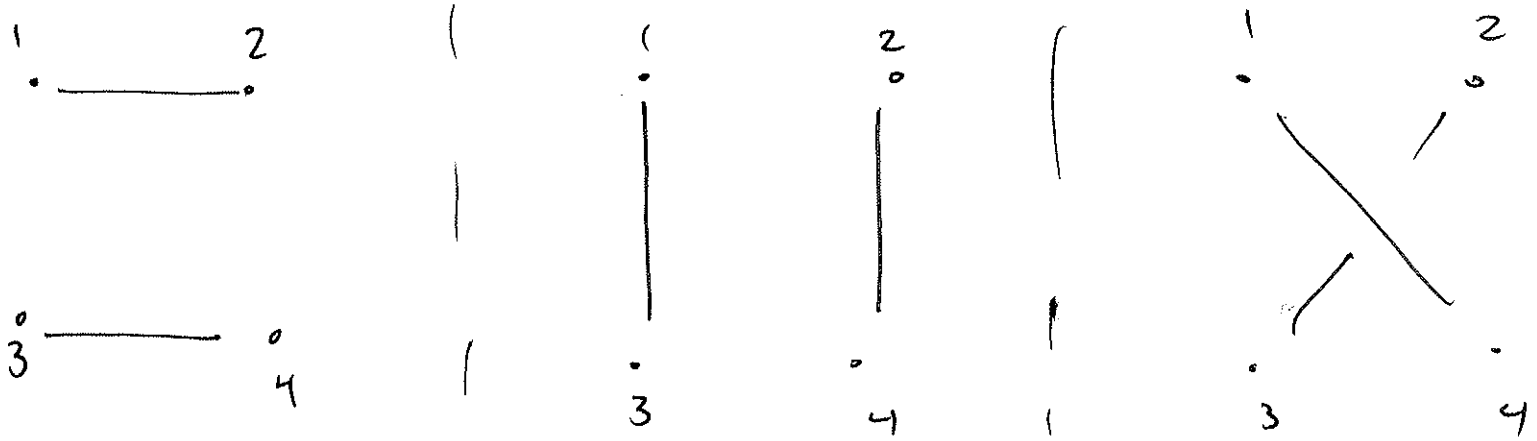
Ex 1: Suppose $X \sim \mathcal{N}(0,1)$. Compute $\mathbb{E}[X^q]$.

$$\mathbb{E}[X^q] = \mathbb{E}[X_1 \cdots X_q] \qquad (X_1, \ldots, X_q) = \mathcal{N}(0, \mathbb{1}\mathbb{1}^T)$$

$$= \sum_{\pi \in PP(q)} \prod_{(i,j)\in\pi} 1$$

$$= |PP(q)| = \begin{cases} 0 & q \text{ odd} \\ (q-1)!! & q \text{ even} \end{cases}.$$

Ex 2: Write out "diagrams" for $\mathbb{E}[X_1 X_2 X_3 X_4]$.



$$\Sigma_{12}\Sigma_{34} + \Sigma_{13}\Sigma_{24} + \Sigma_{14}\Sigma_{23}$$

Ex 3: $\mathrm{Var}[X^T A X] = \sum_{i,j,\ell,e} A_{ij} A_{\ell e}\, \mathbb{E}[X_i X_j X_\ell X_e] = 2\,\mathrm{tr}((A\Sigma)^2) + (\underbrace{\mathbb{E}()}_{\mathrm{tr}(A\Sigma)})^2$

Context: Def: Suppose $X = (X_1, \dots, X_d)$ any RV with all finite moments. Then

$$\kappa(X) \equiv \kappa(X_1, \dots, X_d) = \text{"joint/mixed cumulant"}$$

$$= i^m \frac{\partial^m}{\partial \xi_1 \cdots \partial \xi_m} \log \mathbb{E}\left[e^{i \xi \cdot X}\right]\Big|_{\xi = 0}$$

Thm: We have

(0) $\kappa(X_1) = \mathbb{E}[X_1]$, $\kappa(X_1, X_2) = \mathrm{Cov}(X_1, X_2)$

(i) $\kappa(\cdot, \dots, \cdot)$ is multi-linear

(ii) $\kappa(X_1, \dots, X_d + c) = \kappa(X_1, \dots, X_d) + c\,\delta_{d,1}$

(iii) If $\exists\, \ell \subset [d] = \ell \cup \ell^c$, $\ell \neq \emptyset$, $\ell \neq [d]$ and $X_\ell \perp\!\!\!\perp X_{\ell^c}$ then $\kappa(X) = 0$

(iv) $\mathbb{E}[X_1 \cdots X_d] = \sum_{\pi \in P(d),\ \pi = (\pi_1, \dots, \pi_B)} \prod_{b=1}^{B} \kappa(X_{\pi_b})$

(v) $\kappa(X_{\cdot|Y}) = \sum_{\substack{\pi \in P(d) \\ \pi=(\pi_1 \dots \pi_b)}} \kappa(\kappa(X_{\pi_1}|Y), \dots, \kappa(X_{\pi_b}|Y))$ (4)

III. Def: (GOE) Let $N \geq 1$ and take

$$M = \frac{1}{\sqrt{2}}(A + A^T), \quad A = (a_{ij}) \quad a_{ij} \overset{iid}{\sim} \mathcal{N}(0, \tfrac{1}{N})$$

i.e. $M = (m_{ij}) \Rightarrow$ $m_{ij} = m_{ji}$

$m_{ij} \sim \mathcal{N}(0,1)$ $i \neq j$

$m_{ij} \sim \mathcal{N}(0,2)$ $i = j$.

Goal: Study eigenvalues

$$\lambda_1 = \lambda_1(M) \geq \dots \geq \lambda_N = \lambda_N(M)$$

as $N \to \infty$.

NB: $\mathbb{E}[M_{ij} M_{ke}] = \frac{1}{N}(\delta_{ik}\delta_{je} + \delta_{ie}\delta_{jk})$

Claim: Fix $u \in \mathbb{R}^N$. Then $\mathbb{E}[\|Mu\|^2] \overset{?}{=}$

Pf: $\mathbb{E}[\|Mu\|^2] = \mathbb{E}\left[\sum_{i=1}^N (Mu)_i^2\right]$

$$= \mathbb{E}\left[\sum_{i=1}^N \left(\sum_{j=1}^N M_{ij} u_j\right)^2\right]$$

$$\frac{1}{N}\mathrm{tr}(M^2) = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[(M^2)_{ii}\right]$$

$$= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[\sum_{j=1}^{N} M_{ij}M_{ji}\right]$$

Lecture 6

Plan: I. Semi-Circle Thm ①

II. Warmup

III. Paths = moments

IV. Proof

Recall: $M = (m_{i\bar{j}})_{i,\bar{j}=1,\dots,N}$

$$m_{i\bar{j}} = \begin{cases} \mathcal{N}(0, 1/N) & , \ i > j \\ \mathcal{N}(0, 2/N) & , \ i = j \\ m_{ji} & , \ i < j \end{cases}$$

Thm: (Wigner) Write λ_i = eigenvals of M.

For all $a < b$ we have

$$\frac{1}{N} \#\{i \le N \mid a \le \lambda_i \le b\} \to \int_a^b \rho_{sc}(x)\, dx$$

where

$$\rho_{sc}(x) = \begin{cases} \sqrt{4 - x^2} & |x| \le 2 \\ 0 & \text{o/w} \end{cases}.$$

Proof Idea: (1) If limiting histogram ρ exists

$$k^{\text{th}} \text{ moment} = \frac{1}{N} \sum_{i=1}^{N} \lambda_i^k \xrightarrow{N \to \infty} \int_{\mathbb{R}} x^k \rho(x)\, dx$$

(2) Compute as $N \to \infty$

$$\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\lambda_i^k\right] = \mathbb{E}\left[\frac{1}{N}\mathrm{tr}(M^k)\right] = \mathbb{E}[\text{poly Gaussians}] \to m_k$$

(3) Check

$$\frac{\mathrm{Var}\left[\frac{1}{N}\sum_{i=1}^{N}\lambda_i^k\right]}{\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\lambda_i^k\right]^2} \to 0$$

(4) Find that $\rho = \rho_{sc}$ has $\int x^k d\rho = m_k$.

Warmup: $\mathbb{E}[M_{ab}M_{cd}] = \frac{1}{N}(\delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc})$.

$$\mathbb{E}\left[\frac{1}{N}\mathrm{tr}(M^2)\right] = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}(M^2)_{ii}\right]$$

$$= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[\sum_{j=1}^{N}M_{ij}M_{ji}\right]$$

$$= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{N}(\delta_{ij}+1)$$

$$= 1 + O(1/N)$$

$$\mathbb{E}\left[\frac{1}{N}\mathrm{tr}(M^4)\right] = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}(M^4)_{ii}\right]$$

$$= \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{abc}m_{ia}m_{ab}m_{bc}m_{ci}\right]$$

$$= \frac{1}{N} \sum_{i=1}^{N} \sum_{abc} \frac{1}{N^2} \Big((\delta_{i\bar{i}} \delta_{ab} + \delta_{ib} \delta_{a\bar{a}})(\delta_{bc} \delta_{c\bar{i}} + \delta_{bi} \delta_{c\bar{c}}) + \quad \{12\}\{34\} \quad (3)$$

$$(\delta_{i\bar{b}} \delta_{ac} + \delta_{ic} \delta_{a\bar{b}})(\delta_{a\bar{c}} \delta_{bi} + \delta_{ai} \delta_{b\bar{c}}) + \quad \{13\}\{24\}$$

$$(\delta_{ic} \delta_{ai} + \delta_{ii} \delta_{ac})(\delta_{ab} \delta_{bc} + \delta_{ac} \delta_{bb}) \Big) \quad \{14\}\{23\}$$

$\{12\}\{34\} = \delta_{ib} +$ 2 or more indices match

$\{13\}\{24\} =$ 2 or more match

$\{14\}\{23\} = \delta_{ac} +$ 2 or more match

$$= \frac{1}{N^3} \sum_{i=1}^{N} \sum_{abc} (\delta_{ib} + \delta_{ac}) + O(1/N) = 2 + O(1/N).$$

General Moments:

Recall: $(AB)_{ij} = \sum_{\ell} a_{i\ell} b_{\ell j}$

$$\Rightarrow (M^{2k})_{\substack{i \\ j}} = \sum_{i_2 \cdots i_{2k}} m_{ii_2} m_{i_2 i_3} \cdots m_{i_{2k} j}$$

Interp: Construct graph $G = (V = [N], \text{all edges})$

$$(M^{2k})_{ij} = \sum_{\substack{\text{paths of } m \text{ steps} \\ \pi = (\pi(0), \ldots, \pi(m)) \\ \pi(0) = i \; \pi(m) = j}} \prod_{a=2}^{m} \overset{\omega(\pi)}{m_{\pi(a-1)\pi(a)}}$$

So

$$\mathbb{E}\left[\frac{1}{N}\operatorname{tr}(M^{2k})\right] = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}(M^{2k})_{ii}\right]$$

$$= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[\sum_{\substack{\text{paths } \pi \text{ of} \\ 2k \text{ steps} \\ \text{from } i \to i}} \omega(\pi)\right]$$

$$= \sum_{\substack{\text{paths } \pi \text{ of } 2k \\ \text{steps } i \to i}} \mathbb{E}[\omega(\pi)] \qquad \omega(\pi) = \prod_{a=1}^{2k} m_{\pi(a-1)\pi(a)}$$

$\pi = (i, \pi(1), \ldots, \pi(2k-1), i)$

Lemma: $\mathbb{E}[\omega(\pi)] = 0$ unless π visits every edge even # of times. Further

$$|\mathbb{E}[\omega(\pi)]| \le \alpha_k N^{-k}$$

Back to $\mathbb{E}\left[\frac{1}{N}\operatorname{tr}(M^2)\right] = \frac{1}{N}\sum_{\pi=(i,\pi(1),i)} \mathbb{E}[m_{i\pi(1)} m_{\pi(1)i}]$

$= 1$



Lecture 7:

Plan: I. Stiltjes Intro
II. 2 Big Ideas ①
III. Density via Stiltjes.

$M = \frac{1}{\sqrt{2}}(A + A^T)$ $\quad A = (a_{ij})_{i,j=1,\dots,N}$ $\quad a_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0, 1/N)$

Thm: (Wigner, Marchenko-Pastur) For $a < b$

$$\frac{1}{N} \#\{\lambda_i : \lambda_i \in (a,b)\} \xrightarrow{N\to\infty} \int_a^b \rho_{sc}(x)\,dx$$

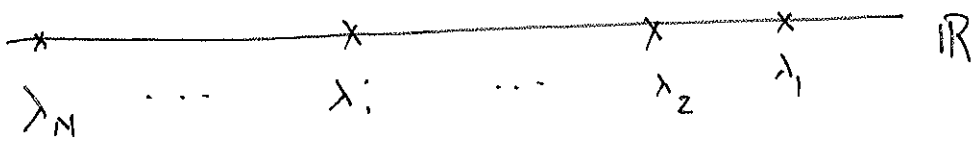
where $\rho_{sc}(x) = \frac{1}{2\pi}\sqrt{4-x^2}\,\mathbb{1}_{|x|\le 2}$.

Goal: Novel proof via "Stiltjes Transform"

$$m_N(z) \equiv \frac{1}{N} tr((M - zI)^{-1}) = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\lambda_j - z} \quad \left(= \mathbb{E}_\lambda\left[\frac{1}{\lambda - z}\right]\right)$$

$\mathbb{C}$

$\cdot z$



Interp: ① When $|z| >> 1$ $\quad (M - zI)^{-1} = -\frac{1}{z}\left(I - \frac{M}{z}\right)^{-1}$

$$= -\frac{1}{z}\sum_{k\ge 0}\frac{M^k}{z^k}$$

$\Rightarrow m_N(z) = -\frac{1}{z}\sum_{k\geq 0}\frac{1}{z^k}\frac{1}{N}\mathrm{tr}(M^k)$ "encodes moments"

$= -\frac{1}{z} + O\left(\frac{1}{|z|^2}\right)$ as $|z| \to \infty$.

② $\frac{1}{\lambda - z}$ = "electric field @ z from e^- @ λ"

$\Rightarrow m_N(z)$ = total e-field @ z from e^- @ λ_i's.

Proof Idea: For every N

$$\frac{1}{N}\#\{i : \lambda_i \in (a,b)\} = \frac{1}{\pi}\int_a^b \mathrm{Im}\left(m_N(x+i\varepsilon)\right)dx + O(\sqrt{\varepsilon}) \qquad (*)$$

So

$$\frac{1}{N}\#\{i : \lambda_i \in (a,b)\} \to \int_a^b \lim_{\varepsilon \downarrow 0}\lim_{N\to\infty} \mathrm{Im}\left(m_N(x+i\varepsilon)\right)dx$$

Moreover

$$m_N(z) \to m(z) \qquad (**) \text{ "via cavity method"}$$

and

$$m(z) = \left(-m(z) - z\right)^{-1} \Rightarrow m(z)^2 + zm(z) + 1 = 0$$

Today: We prove $(*)$.

Lemma: $\|(M - zI)^{-1}\| = \max_i \frac{1}{|\lambda_i - z|} \leq \frac{1}{|\mathrm{Im}(z)|} \Rightarrow |m_N(z)| \leq \frac{1}{|\mathrm{Im}(z)|}$

Lemma: $\frac{1}{N}\#\{i : |\lambda_i - a| \leq \eta\} \leq 2\eta\ \mathrm{Im}\left(m_N(a+i\eta)\right)$ "density UB"

Step 1: (Let $\Gamma = \partial\{[a,b]\times[-\varepsilon,\varepsilon]\}$. Then)

$$\frac{1}{N}\#\{i:\lambda_i\in(a,b)\} = \frac{1}{\pi}\int_a^b \mathrm{Im}\left(m_N(x+i\varepsilon)\right) + \frac{1}{N\pi}\sum_{i=1}^{N}\arctan\left(\frac{\varepsilon}{\lambda_i-a}\right) - \arctan\left(\frac{\varepsilon}{\lambda_i-b}\right).$$

Pf: Then

$$\frac{1}{N}\#\{i:\lambda_i\in(a,b)\} = -\frac{1}{2\pi i}\oint_\Gamma m_N(z)\,dz$$

$$\text{RHS} = -\frac{1}{2\pi i}\left[\underbrace{\int_a^b \left(m_N(x-i\varepsilon) - m_N(x+i\varepsilon)\right)dx}_{A} + \underbrace{\int_{-\varepsilon}^{\varepsilon}\left(m_N(b+iy) - m_N(a-iy)\right)i\,dy}_{B}\right]$$

Note that

$$m_N(\bar z) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i-\bar z} = \frac{1}{N}\sum_{i=1}^{N}\overline{\frac{1}{\lambda-z_i}} = \overline{m_N(z)}.$$

So

$$A = \int_a^b -2i\,\mathrm{Im}\left(m_N(x+i\varepsilon)\right)dx \qquad \checkmark$$

Next,

$$m_N(b+iy) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{(\lambda_i-b)-iy} = \frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_i-b+iy}{(\lambda_i-b)^2+y^2}$$

$$\Rightarrow \int_{-\varepsilon}^{\varepsilon} m_N(b+iy)\,i\,dy = \frac{1}{N}\sum_{i=1}^{N} 2i\tan^{-1}\left(\frac{\varepsilon}{\lambda_i-b}\right) \qquad \square$$

Since $|\tan^{-1}(t)| \le \min\left(\frac{\pi}{2}, |t|\right)$ we find

$$\left| \underbrace{\frac{1}{N}\#\{i : \lambda_i \in (a,b)\} - \frac{1}{\pi}\int_a^b \mathrm{Im}\,(m_N(x+i\varepsilon))\,dx}_{(\#)} \right| \le \frac{1}{N}\sum_{i=1}^{N} \min\left(1, \frac{2\varepsilon}{\pi d_i}\right) \qquad (4)$$

where

$$d_i = \min\{|\lambda_i - a|, |\lambda_i - b|\}$$

So

$$(\#) \le \frac{1}{N}\sum_{i=1}^{N} \min\left(1, \frac{2\varepsilon}{\pi d_i}\right)\left(1\!\!1_{\{d_i \le \eta\}} + 1\!\!1_{\{d_i \ge \eta\}}\right)$$

$$\le \frac{\#\{i : d_i \le \eta\}}{N} + \frac{2\varepsilon}{\pi\eta}$$

$$\le 2\eta\left(\mathrm{Im}\,(m_N(a+i\eta)) + \mathrm{Im}\,(m_N(b+i\eta))\right) + \frac{2\varepsilon}{\pi\eta}$$

Now suppose $\exists m : \mathbb{H} \to \mathbb{R}$ with $\mathrm{Im}(m) \le K$ and $m(x) = \lim_{\varepsilon \downarrow 0} m(x+i\varepsilon)$

$$\sup_{|x| \le L} |m_N(x+i\varepsilon) - m(x+i\varepsilon)| \xrightarrow{N\to\infty} 0 \qquad \forall L, \varepsilon$$

Then with $\eta = \sqrt{\varepsilon}$

$$\lim_{N\to\infty} \mathbb{E}\left[\left|\frac{1}{N}\#\{i : \lambda_i \in (a,b)\} - \frac{1}{\pi}\int_a^b \mathrm{Im}\,(m_N(x+i\varepsilon))\,dx\right|\right] \le C\cdot\sqrt{\varepsilon}$$

□

So we find limit density

$$\rho(x) = \lim_{\varepsilon \downarrow 0} \lim_{N\to\infty} \frac{1}{\pi}\,\mathrm{Im}\,(m_N(x+i\varepsilon))$$

Now the goal is to compute limit of $m_N(x+iq)$ as $N \to \infty$. 2 steps: ① $\mathbb{E}$, ② Var

Lemma: Suppose $A = \begin{pmatrix} \alpha & \beta^T \\ \beta & \Delta \end{pmatrix}$. Then

$$\gamma = (A^{-1})_{11} = (\alpha - \beta^T \Delta^{-1} \beta)^{-1} \quad \text{and}$$

$$\operatorname{tr}(A^{-1}) = \operatorname{tr}(\Delta^{-1}) + \gamma(1 + \beta^T \Delta^{-2} \beta)$$

Now write

$$M = \begin{pmatrix} M_{11} & \overbrace{M_{12} \cdots M_{1N}}^{\beta} \\ \beta^T & M^{(1)} \end{pmatrix}$$

Goal: relate $(M - zI)^{-1}$ to $R = (M^{(1)} - zI)^{-1}$

Then

$$((M - zI)^{-1})_{11} = \frac{1}{M_{11} - z - \beta^T (M^{(1)} - zI)^{-1} \beta}$$

$$\mathbb{E}[m_N(z)] = \mathbb{E}\left[((M - zI)^{-1})_{11} \right]$$

$$= \mathbb{E}\left[\frac{1}{M_{11} - z - \beta^T (M^{(1)} - zI)^{-1} \beta} \right] \quad \left(\lesssim \frac{1}{|\operatorname{Im}(z)|} \right)$$

Observe: ① $M_{11} = \mathcal{O}(1/\sqrt{N})$

② $\beta \sim \mathcal{N}(0, \frac{1}{N} I_{N-1}) \Rightarrow \beta^T (M^{(1)} - zI)^{-1} \beta$

$$= \frac{1}{N} \operatorname{tr}((M^{(1)} - zI)) + \mathcal{O}(1/\sqrt{N})$$

So ~~if we knew~~ let's write

$$\begin{cases} M_{11} - z - \beta^T R \beta = -z - \mathbb{E}[m_N] + \text{Error} \\ \text{Error} = M_{11} - (\beta^T R \beta - \frac{1}{N} \operatorname{tr}(R)) + \frac{1}{N} \operatorname{tr}(R) \end{cases}$$

~~$*(m_N - \mathbb{E}[m_N])$~~

$$\mathbb{E}[|\text{Error}|] \leq C\left(\frac{1}{\sqrt{N}} + \frac{1}{\varepsilon\sqrt{N}} + \frac{1}{N\varepsilon}\right)$$

and get

$$\mathbb{E}[m_N] = \mathbb{E}\left[\frac{1}{-z - \mathbb{E}[m_N] + \text{Error}}\right]$$

From here get

$$m = \frac{1}{-z - m} \quad \Rightarrow \quad m^2 + zm + 1 = 0$$

$$\Rightarrow \quad m(z) = \frac{-z \pm \sqrt{z^2 - 4}}{2}$$

Only + gives correct boundary

Lecture 8

Plan: I. Recall + Goal
II. Gaussian Poincaré
III. Schur Complement

I. $M = \frac{1}{\sqrt{2}}(A + A^T)$ $\quad A = (a_{ij})_{i,j=1,\dots,N}$ $\quad a_{ij} \overset{i.i.d}{\sim} \mathcal{N}(0, 1/N)$

$$m_N(z) \equiv \frac{1}{N} tr\left((M - zI)^{-1}\right) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\lambda_i - z}$$

Showed: If $\exists\, m : \mathbb{H} \to \mathbb{R}$ s.t.

$$\sup_{|x| \le L} |m_N(x + i\varepsilon) - m(x + i\varepsilon)| \to 0 \qquad \forall \varepsilon, L > 0$$

with $|Im(m)| \le K$ then

$$\mathbb{E}\left[\left|\frac{1}{N}\#\{i : \lambda_i \in (a,b)\} - \int_a^b \lim_{\varepsilon \downarrow 0} \frac{1}{\pi} Im(m(x + i\varepsilon))\, dx\right|\right] \to 0.$$

Def: Let $m(z)$ be the unique solution to

$$m(z)(-z - m(z)) = 1 \iff m(z)^2 + z m(z) + 1 = 0$$

with $|m(z)| < 1$.

Thm: $\forall \varepsilon > 0$, $z \in \mathbb{C}$ $Im(z) = \varepsilon$

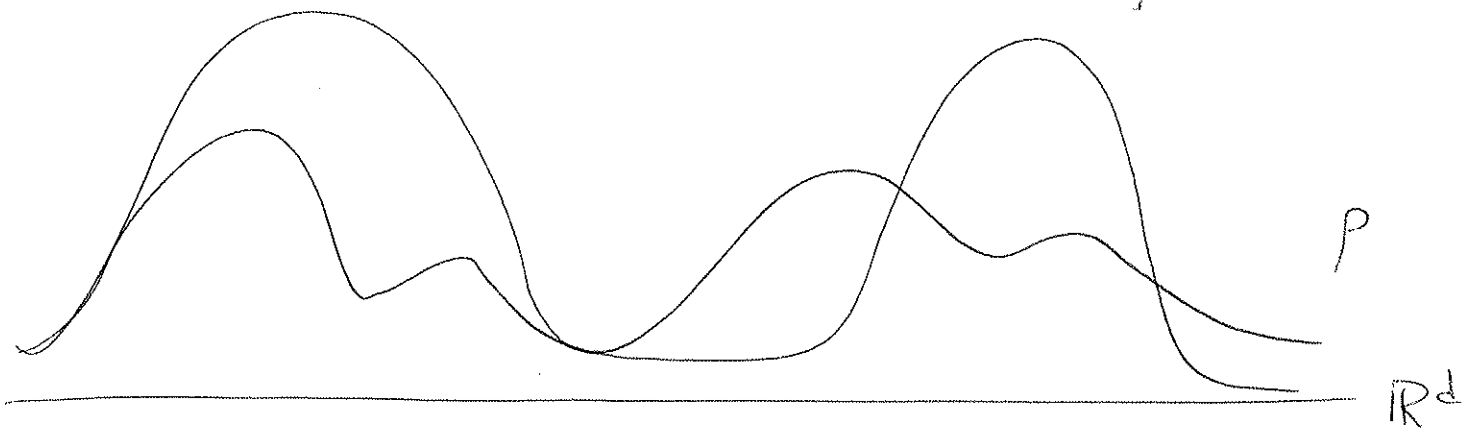
(a) $\mathbb{E}\left[|m_N(z) - \mathbb{E}[m_N(z)]|\right] \le \frac{3}{N\varepsilon^2}$

(b) $|\mathbb{E}[m_N(z)] - m(z)| \le \frac{7(|z| + \varepsilon^{-1})}{\varepsilon^5 \sqrt{N}}$

II. Q: How to bound $\mathrm{Var}\left[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i - \tau}\right]$?

↑ non-iid!

Idea: (Poincaré) Consider density $\rho(x) > 0$



Def: C_ρ = "Poincare Constant of ρ"

$= \inf_{C>0}$ s.t. $\mathrm{Var}_\rho[f] \le C\, \mathbb{E}_\rho[\|\nabla f\|^2] \quad \forall f \in C^1$

Interp: C_ρ = "biggest ~~bottleneck~~ valley of ρ"

Thm: Let $X = (X_i \sim N(0, \sigma_i^2))_{i=1}^n$ indep. Then

$$\mathrm{Var}[f(X)] \le \frac{\pi^2}{8}\sum_{i=1}^{n}\sigma_i^2\, \mathbb{E}[(\partial_i f(X))^2]$$

Pf: WLOG $\&\ \sigma_i^2 = 1 \ \forall i$. Note

$$\mathrm{Var}[f(X)] = \frac{1}{2}\mathbb{E}[(f(X) - f(Y))^2], \quad X, Y \overset{iid}{\sim} X$$

Now interpolate

$$X_\theta = X\sin\theta + Y\cos\theta, \quad \frac{d}{d\theta}X_\theta = X\cos\theta - Y\sin\theta$$

For each θ $(X_\theta, \frac{d}{d\theta} X_\theta)$ Gaussian with mean 0

$$\text{Cov} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

We have

$$f(X) - f(Y) = \int_0^{\pi/2} \langle \nabla f(X_\theta), \frac{d}{d\theta} X_\theta \rangle \, d\theta$$

$$\Rightarrow (f(X) - f(Y))^2 \le \frac{\pi}{2} \int_0^{\pi/2} \langle \nabla f(X_\theta), \frac{d}{d\theta} X_\theta \rangle^2 d\theta$$

$$\Rightarrow \text{Var}[f(X)] \le \frac{\pi}{4} \int_0^{\pi/2} \mathbb{E}\left[\langle \nabla f(X_\theta), \frac{d}{d\theta} X_\theta \rangle^2 \right] d\theta$$

$$= \frac{\pi}{4} \int_0^{\pi/2} \mathbb{E}\left[\|\nabla f(X_\theta)\|^2 \right] d\theta$$

$$= \frac{\pi^2}{8} \mathbb{E}\left[\|\nabla f(X)\|^2 \right] \qquad \square$$

Props: $\text{Var}\left[\frac{1}{N} \text{Re}\left(\text{tr}(M - zI)^{-1} \right)\right]$, $\text{Var}\left[\text{Im}\left(\frac{1}{N}\text{tr}(M - zI)^{-1}\right)\right]$

$$\le \frac{\pi^2}{4\varepsilon^4 N^2}$$

$$\Rightarrow \mathbb{E}\left[|m_N(t) - \mathbb{E}[m_N(t)]| \right] \le \frac{\pi}{\sqrt{2}\varepsilon^2 N} < \frac{3}{N\varepsilon^2}.$$

Pf: Let $f(M) = \text{Re}\left(\text{tr}(\overbrace{(M - zI)^{-1}}^{R}) \right)$. The indep variables are $H = \begin{cases} m_{ij}, & i > j \\ m_{ii} \end{cases}$.

$$\mathrm{Var}[f(M)] \le \sum_{i\le j} \frac{1}{N} \mathbb{E}\left[(\partial_{M_{ij}} f(M))^2\right] + \sum_i \frac{2}{N} \mathbb{E}\left[(\partial_{M_{ii}} f(M))^2\right] \quad (4)$$

To evaluate compute

$$\frac{d}{dt}\Big|_{t=0} \mathrm{tr}(tH + M - zI)^{-1} = -\mathrm{tr}(HR^2)$$

$$\Rightarrow \frac{\partial f}{\partial M_{ij}} = -2R^2_{ij} \qquad \frac{\partial f}{\partial M_{ii}} = -(R^2)_{ii}$$

$$\Rightarrow \mathrm{Var}[f(M)] \le \frac{2}{N} \sum_{i,j} |(R^2)_{ij}|^2 = \frac{2}{N} \mathrm{tr}\left((R^2)(R^2)^*\right)$$

$$= \frac{2}{N} \cdot \sum_{i=1}^N \frac{1}{|\lambda - z|^4} \le \frac{2}{(\mathrm{Im}(z))^4}$$

□

II. Our goal is now to study

~~$\mathbb{E}[m_N(z)] = \mathbb{E}\left[\frac{1}{N}\mathrm{tr}((M-zI)^{-1})\right]$~~

~~$= \mathbb{E}[((M-zI)^{-1})_{11}]$~~

Goal

Claim: $$\left|\mathbb{E}[m_N(z)] - \frac{1}{-z - \mathbb{E}[m_N(z)]}\right| \le \frac{7}{\varepsilon^4 \sqrt{N}}.$$

Note

Write

$$M - zI = \begin{pmatrix} M_{11} - z & \beta \\ \beta^T & M^{(1)} - z \end{pmatrix} \qquad \beta \sim \mathcal{N}(0, \tfrac{1}{N} I_{N-1})$$

Idea: $\mathbb{E}[m_N(z)] = \mathbb{E}[\frac{1}{N} \mathrm{tr}((M - zI)^{-1})_{11}]$

Lemma: (Schur Complement) Suppose

$$A = \begin{matrix} 1 \\ n-1 \end{matrix} \begin{pmatrix} \overset{1}{\alpha} & \overset{n-1}{\beta} \\ \beta^T & \Delta \end{pmatrix}$$

Then

$$\begin{cases} (A^{-1})_{11} = (\alpha - \beta^T \Delta^{-1} \beta)^{-1} \\ \mathrm{tr}(A^{-1}) = \mathrm{tr}(\Delta^{-1}) + (A^{-1})_{11} (1 + \beta^T \Delta^2 \beta) \end{cases}$$

$$= \mathbb{E}[(M_{11} - z - \beta^T (M^{(1)} - z) \beta)^{-1}]$$

Key Points:

① $\mathbb{E}[M_{11}^2] = \frac{2}{N}$

② $\mathbb{E}[(\beta^T (M^{(1)} - z)^{-1} \beta - \frac{1}{N} \mathrm{tr}((M^{(1)} - z)^{-1}))^2] = O(1/N)$

③ $\mathbb{E}[(\frac{1}{N} \mathrm{tr}((M^{(1)} - z)^{-1}) - \frac{1}{N} \mathrm{tr}((M - z)^{-1}))^2] = O(1/N)$

④ $\mathbb{E}[(\frac{1}{N} \mathrm{tr}((M - zI)^{-1}) - \mathbb{E})^2] = O(1/N)$

Hence

$$\mathbb{E}[m_N(z)] = \left(-z - \mathbb{E}[m_N(z)] + \Delta_{N,\varepsilon}\right)^{-1}$$

where

$$\mathbb{E}[\Delta_{N,\varepsilon}^2] = O(1/N).$$

Lecture 9

Plan: I. GP Def (1)

II. Examples

III. Feature Maps & Kernels

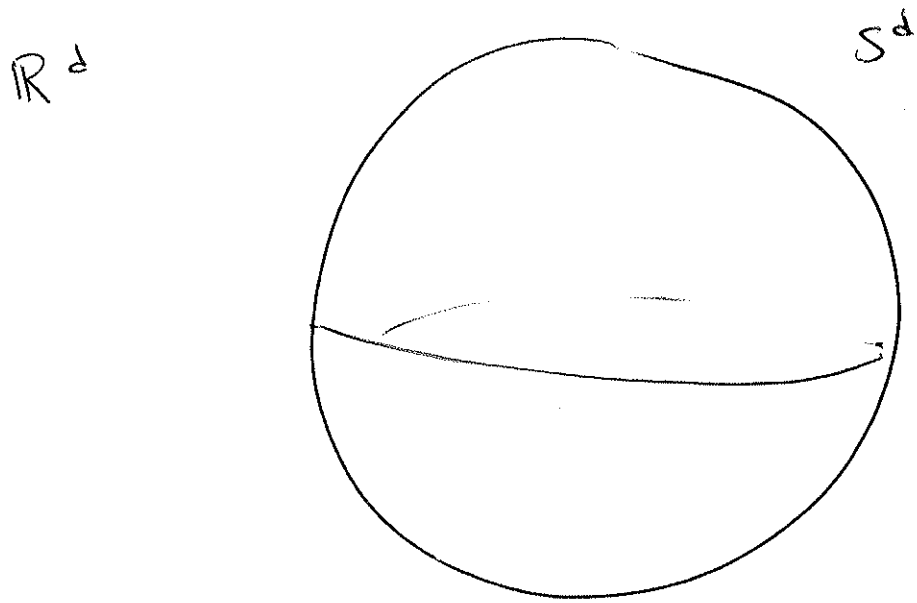
IV. Maxima + Canonical Metric

Def: Let T be a set. A GP on T is a collection of RVs $(X_t)_{t \in T}$ s.t.

$$\begin{cases} \forall t_1, \dots, t_k \in T \quad (X_{t_1}, \dots, X_{t_k}) \sim \mathcal{N}(\mu(t), \Sigma(t)) \\ \mu_t = \text{"mean of } X_t\text{"} \qquad \Sigma_{st} = \text{"covariance of } X_s, X_t\text{"} \end{cases}$$

Ex 1: $T = \{1, \dots, d\}$ GP = Gaussian on $\mathbb{R}^d$

Ex 2: $T = S^{d-1} \times S^{N-1}$



Let M - $N \times d$ w/ iid $\mathcal{N}(0,1)$ entries. Then

$X_{u,v} = v^T M u, \quad u \in \mathbb{R}^d$

is GP on S^{d-1}. We have

$\mu_{(u,v)} = \mathbb{E}[X_{u,v}] = 0 \quad \mathbb{E}[X_{u,v} X_{u',v'}] = \mathbb{E}[v^T M u (v')^T M u']$

$(\Sigma_{(u,v),(u',v')})_{ij} = \mathbb{E}[(X_u)_i (X_v)_j] = \mathbb{E}[(u^T M)_i (M^T v)_j]$

$= u^T \mathbb{E}[MM^T]_{ij} v$

$= \delta_{ij} u^T v = \delta_{ij} \langle u, v \rangle \Rightarrow \Sigma_{u,v} = \langle u, v \rangle I.$

$\mathbb{E} \sum_{i,j} v_i u_j M_{ij} \sum_{i',j'} v'_{i'} u'_{j'} M_{i'j'}$

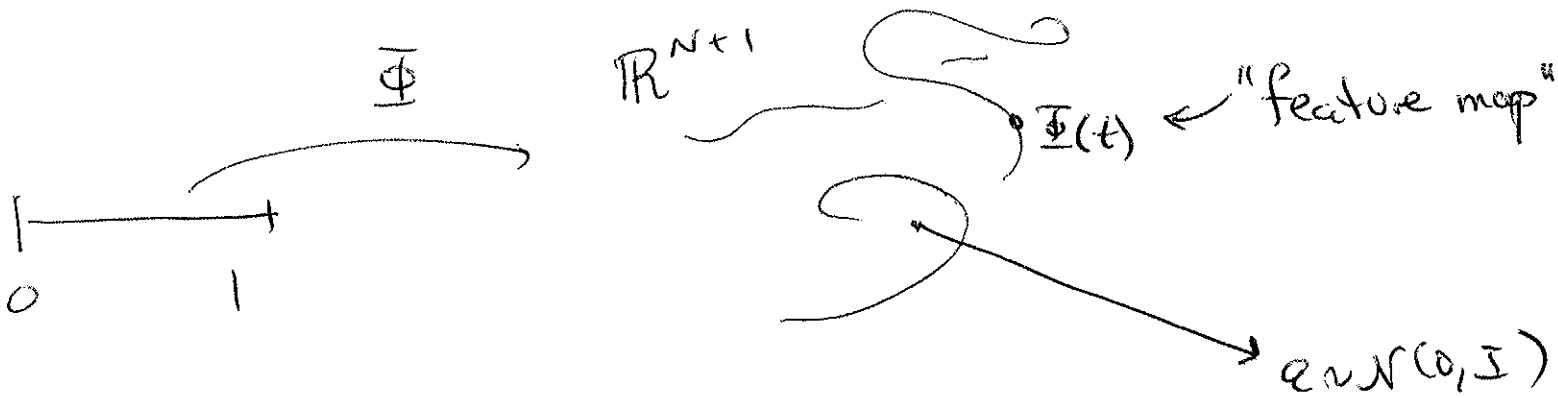
$= \langle u, u' \rangle \langle v, v' \rangle$

Q: What is $\sup_{u \in S^{d-1}, v \in S^{d-1}} X_{u,v}$? = top singular value of M!

Ex 3: $T = [0,1] \quad X_t = \sum_{j=0}^{N} a_j t^j \quad a_j \overset{iid}{\sim} \mathcal{N}(0,1)$

Q: Why GP?

A: $X_t = \langle \vec{a}, \overline{\Phi}_t \rangle, \quad \overline{\Phi}_t = (1, t, \ldots, t^N)$



Thm:

Ex 4: Let $\Phi: T \to \mathcal{H} = (V, \langle , \rangle)$ with

Def $\|\Phi_t\| = \sqrt{\langle \Phi_t, \Phi_t \rangle} < \infty \quad \forall t$. Then

$X_t = \langle \vec{a}, \Phi_t \rangle = \sum_{j \geq 0} a_j \varphi_j(t)$ is GP w/ $\vec{a} \sim N(0, I)$

↑ expand in ONB

$\mu_t = E[X_t] = 0$

$\Sigma_{s,t} = E[\Phi_t^T \vec{a} \vec{a}^T \Phi_s] = \langle \Phi_t, \Phi_s \rangle$

Thm: Every (centered) GP arises through some Φ!

Q: Φ for Ex 1, 2 ?

Ex 5: $T = [0,1]$ $\mathcal{H} = L^2([0,1], dt) = \{f: [0,1] \to \mathbb{R} \mid \int f^2 < \infty\}$

$\langle f, g \rangle = \int_0^1 f(t) g(t) dt$

$\Phi_t = \mathbb{1}_{[0,t]}$. Then

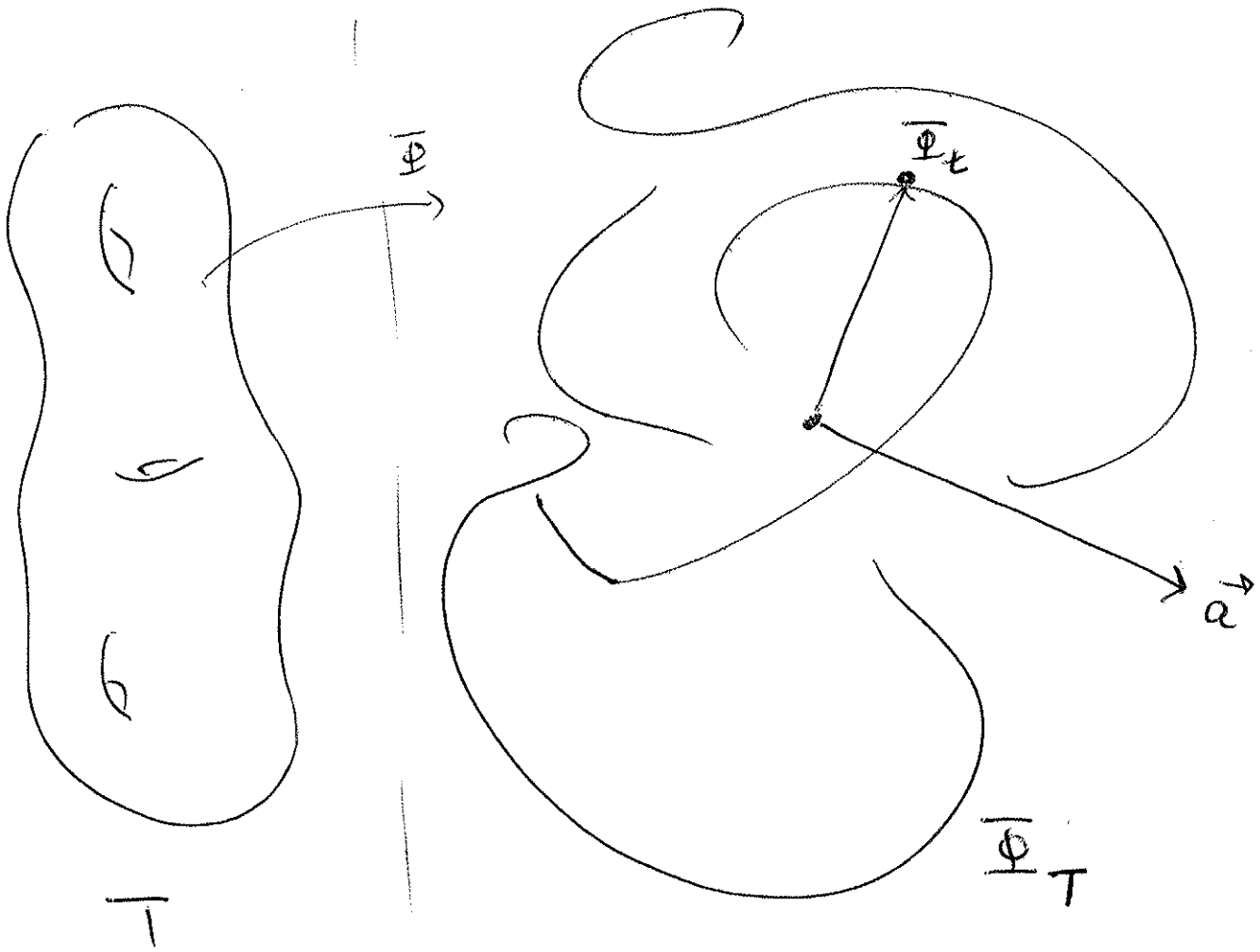
$\Sigma_{s,t} = \langle \mathbb{1}_{[0,t]}, \mathbb{1}_{[0,s]} \rangle_{L^2} = \int_0^1 \mathbb{1}_{[0,t]}(\tau) \mathbb{1}_{[0,s]}(\tau)$

$= \min\{t, s\}$

IV. Def: Let T be a set and $(X_t)_{t\in T}$ be $GP(\mu, \Sigma)$. The canonical metric associated to X is

$$d_X^2(s,t) \equiv \mathbb{E}\left[|X_s - X_t|^2\right] = \Sigma_{s,s} + \Sigma_{t,t} - 2\Sigma_{s,t}.$$

Idea: The canonical metric endows T with intrinsic geometry. All GP questions are naturally about this geometry.



Q: How to write d_X via $\overline{\Phi}$?

A: $d_X^2(s,t) = \|\bar{\sigma}_s - \bar{\sigma}_t\|^2$ = dist in ambient space!

[also get Riemmanian metric]

Def: Let X be a centered GP on T. Define

$$\mathbb{E}\left[\sup_{t\in T} X_t\right] := \sup\left\{\mathbb{E}\left[\max_{t\in F} X_t\right] : F \subset T \text{ finite}\right\}$$

Prop: Let $X_1, \dots, X_n$ be centered Gaussians w/ $\mathrm{Var}[X_i] \le \sigma_i^2$. Then

$$\mathbb{E} \max_{i\le n} X_i \le \sigma\sqrt{2\log n}$$

Pf: Use Jensen:

$$\begin{aligned} e^{\lambda\, \mathbb{E}[\max_{i\le n} X_i]} &\le \mathbb{E}[e^{\lambda \max_i X_i}] \\ &\le \mathbb{E}[\max_i e^{\lambda X_i}] \\ &\le \sum_i \mathbb{E}[e^{\lambda X_i}] \\ &= \sum_i e^{\lambda^2 \mathrm{Var}\, X_i/2} \\ &\le n\, e^{\lambda^2\sigma^2/2} \end{aligned}$$

$\Rightarrow$

$$\mathbb{E}[\max_i X_i] \le \frac{\log n}{\lambda} + \frac{\lambda \sigma^2}{2}$$

Take $\lambda = \sqrt{2\log n}/\sigma$ □

[wide NN, $(f, \nabla f)$

Prop: $X_1, \dots, X_n$ iid $N(0,1)$ then as $n \to \infty$

$$\mathbb{E}[\max_{i \le n} X_i] = \sqrt{2\log n}\,(1 + o(1))$$

Pf: $\mathbb{P}(X > t) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-x^2/2}\,dx \ge \frac{e^{-3/2}}{\sqrt{2\pi}\, t}\, e^{-t^2/2}$. Set

$t = \sqrt{(2-\delta)\log n}$. Then

$$n\,\mathbb{P}(X > t) \ge c n^{\delta/2}/t \to \infty \quad \text{S}$$

$$\mathbb{P}(\max_i X_i \le t) = (1 - \mathbb{P}(X > t))^n \le e^{-n\mathbb{P}(X>t)} \to 0$$

$$\mathbb{E}[\max_i X_i] \ge \mathbb{E}[\mathbb{1}_{\max > t} \max_i X_i]$$

$$\ge t\,\mathbb{P}(\max_i X_i > t)$$

$$= \sqrt{(2-\delta)\log n}\,(1 - o(1))$$

□

Q: How many paths $i \to i$ of $2k$ steps are there that visit m distinct vertices?

A: $\binom{N}{m} \cdot \alpha_m \simeq \alpha_m \cdot N^m$ (other than i)

Hence:

$$\mathbb{E}\left[\frac{1}{N}\mathrm{tr}(M^{2k})\right] = \sum_{\substack{\pi=(\pi(0)=i,\,\pi(1),\ldots,\pi(2k-1),\,\pi(2k)=i) \\ \pi \text{ visits } k \text{ distinct vertices}}} \mathbb{E}[w(\pi)] + o(1)$$

~~A path $\pi = i \to i$ w/ k distinct vertices~~

Def: Given a path $\pi = (i, \pi(1), \ldots, \pi(2k-1), i)$ which has k edges and it traverses every edge it sees exactly 2 times.

Def: $e_1, e_2, \ldots, e_k$. $\Big($ $\#\pi$ visit each k edges $= \#$

Define a pair-partition of $\{1, \ldots, 2k\}$ by pairing ij iff $e_i = e_j$.

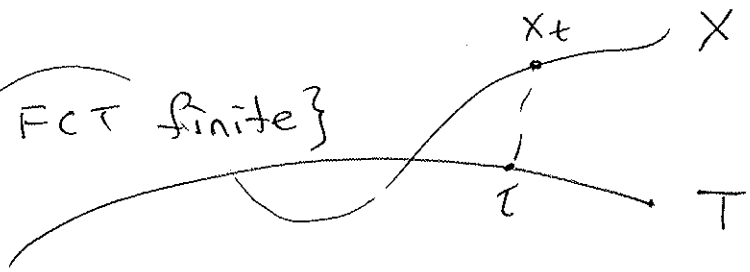
Claim: $\#\pi = \#$ non-cross part of $[k]$ = Catalan's.

Lecture 10

Recall: $X_t = \langle a, \Phi_t \rangle = \sum_{j \geq 0} a_j \varphi_j(t)$ is GP, $t \in T$ $\quad a_j \sim N(0,1)$

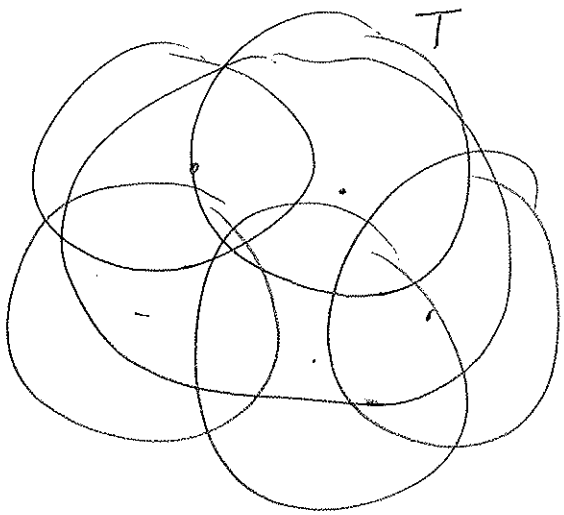
$d_X^2(s,t) = \mathbb{E}[|X_s - X_t|^2]$

$\mathbb{E}[\sup_{t \in T} X_t] = \sup\{\mathbb{E}[\sup_{t \in F} X_t],\ F \subset T \text{ finite}\}$



Def: An ε-net in (T,d) is a collection of points $t_1, \dots, t_K \in T$ s.t.

$$\bigcup_{k=1}^{K} B_\varepsilon(t_k) \supseteq T.$$



The ~~ε-entropy covering #~~ at scale ε

$N(T,d,\varepsilon)$ = size of minimal ε-net.

Thm: (Dudley + Sudakov) ~~Assume~~

We have

$$\mathbb{E}[\sup_{t \in T} X_t] \lesssim 16 \int_0^\infty \sqrt{\log N(T, d_X, \varepsilon)}\, d\varepsilon \qquad (*)$$

and $\forall \varepsilon > 0$

$$\mathbb{E}[\sup_{t \in T} X_t] \gtrsim \varepsilon \sqrt{\log N(T, d_X, \varepsilon)}$$

Pf of (*): Fix $F \subset T$ finite. Write

$$D = \max_{s,t \in F} d_X(s,t)$$

and define $\varepsilon_k = D 2^{-k}$. and set

π_k = ε_k-net on T

We have

$$\pi_0 = \{t_0\}$$

and for all $t \in T$ set

$$\pi_k(t) = \text{nearest point in } \pi_k \text{ to } t$$

We have

$$X_t = X_t - X_{t_0} + X_{t_0}$$

$$= X_{t_0} + \sum_{k \geq 1} X_{\pi_k(t)} - X_{\pi_{k-1}(t)}$$

$$\Rightarrow \sup_{t \in F} X_t \leq X_{t_0} + \sum_{k \geq 1} \sup_{t \in F} X_{\pi_k(t)} - X_{\pi_{k-1}(t)}$$

$$\Rightarrow \mathbb{E}\left[\sup_{t \in F} X_t\right] \leq \sum_{k \geq 1} \mathbb{E}\left[\sup_{t \in F} X_{\pi_k(t)} - X_{\pi_{k-1}(t)}\right]$$

(how much can ↙)

Key point: $d_X(\pi_k(t), \pi_{k-1}(t)) \leq 3\varepsilon_k$

$\mathbb{E}\left[\left|X_{\pi_k(t)} - X_{\pi_{k-1}(t)}\right|^2\right]$ "

And so

$$\mathbb{E}\left[\sup_{t \in F} X_{\pi_k(t)} - X_{\pi_{k-1}(t)}\right] \leq \mathbb{E}\left[\sup_{s,t \in \pi_k(F) \times \pi_{k-1}} X_s - X_t\right]$$

$$\leq 3\varepsilon_k \sqrt{2 \log \mathcal{N}(T, d_X, \pi_k)^2}$$

$$= 6\varepsilon_k \sqrt{\log \mathcal{N}(T, d_X, \pi_k(\varepsilon_k))} \leq 12 \int_{\varepsilon_{k+1}}^{\varepsilon_{k-1}} \sqrt{\log \mathcal{N}(T, d_X, \varepsilon)}\, d\varepsilon$$

□

Ex: Consider $X_u = u^T M u$, $u \in S^{N-1}$ with

$$M = \frac{1}{\sqrt{2}}(A + A^T), \quad A = (a_{ij}) \quad a_{ij} \sim \mathcal{N}(0,1)$$

Q: Is X a GP? A: Y

Q: ~~Φ_t~~ ~~Φ_t~~ ~~Φ_t~~ ?

Q: $d_X^2(u,v) \stackrel{?}{=} \mathbb{E}[|X_u - X_v|^2]$

$$= \mathbb{E}\Big[\Big|\sum_{ij}(u_i u_j - v_i v_j) M_{ij} M_{kl}\Big|^2\Big]$$

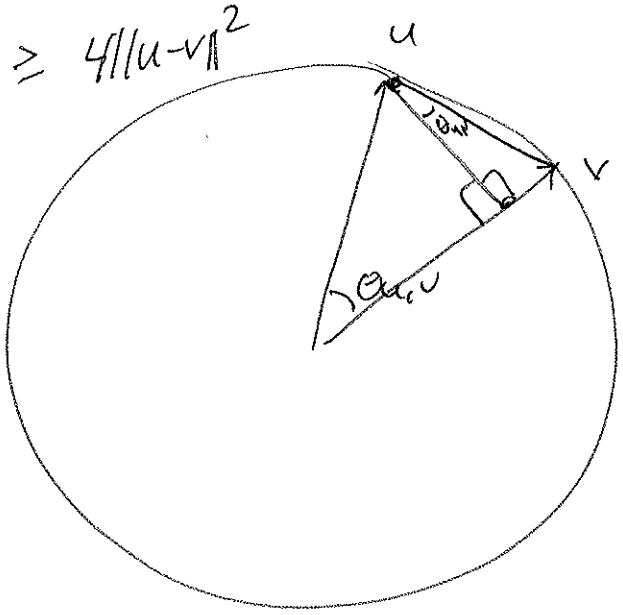
$$= \sum_{\substack{ij \\ i'j'}} (u_i u_j - v_i v_j)(u_{i'} u_{j'} - v_{i'} v_{j'})(\delta_{ii'}\delta_{jj'} + \delta_{ij}\delta_{ji'})$$

$$= 2 \sum_{ij} u_i^2 u_j^2 + v_i^2 v_j^2 - 2 u_i v_i u_j v_j$$

$$= 2\left(1 - \langle u, v\rangle^2\right)$$

$$= 4 \sin^2 \theta_{u,v} \geq 4\|u - v\|^2$$

$$\leq 4\sqrt{2}\,\|u - v\|^2$$

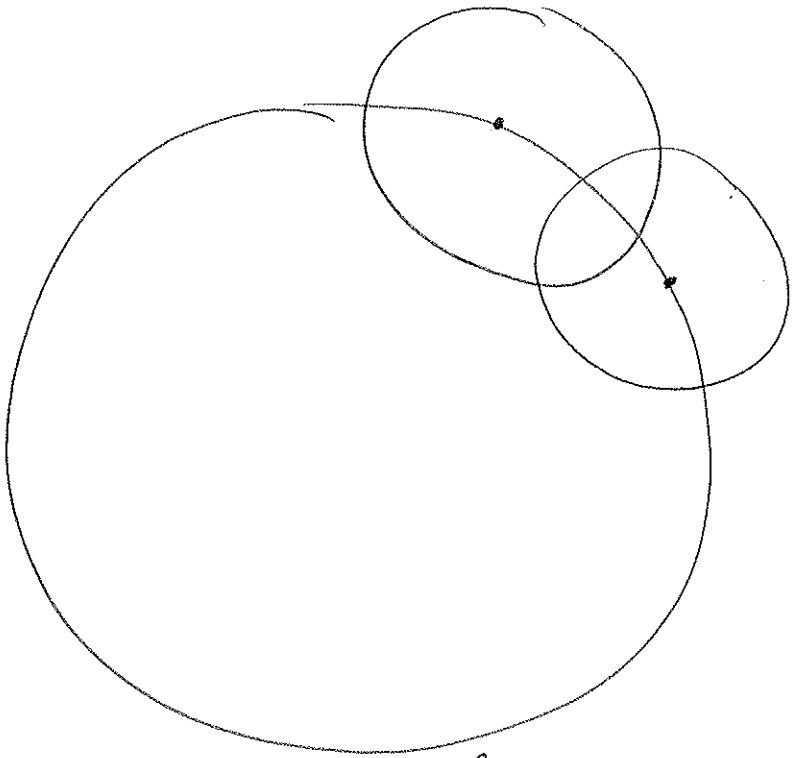


So need

$$N(S^{d-1}, \|\cdot\|, \varepsilon) \leq ?$$

Lemma: $N(S^{d-1}, \|\cdot\|, \varepsilon) \leq \left(1 + \frac{2}{\varepsilon}\right)^d$

Pf: ~~Suppose we have n balls of radius ε that cover S^{d-1} with centers $t_1, t_2, \ldots, t_n$~~ a minimal collection of



Let $u_1, \ldots, u_n$ be a maximal ε-separated i.e.

$$\|u_i - u_j\| > \varepsilon \quad \forall i \neq j.$$

Then (1) $u_1, \ldots, u_n$ ε-net and

(2) $B_{\varepsilon/2}(u_i) \cap B_{\varepsilon/2}(u_j) = \emptyset$ if $i \neq j$.

(3) $\bigcup_{i=1}^{n} B_{\varepsilon/2}(u_i) \subset B_0(1 + \varepsilon/2)$

So

$$W(S^{d-1}, \|\cdot\|, \varepsilon) \le n$$

and

$$n \cdot \mathrm{vol}\left(B^d_{\varepsilon/2}\right) \le \mathrm{vol}\left(B^d_{1+\varepsilon/2}\right)$$

$$\Leftrightarrow \quad n \cdot \left(\frac{\varepsilon}{2}\right)^d \mathrm{vol}\left(B_1^d\right) \le \mathrm{vol}\left(1+\varepsilon/2\right)^d \mathrm{vol}\left(B_1^d\right)$$

$$\Leftrightarrow \quad n \le \left(1+\frac{2}{\varepsilon}\right)^d .$$

So

$$\mathbb{E}\left[\sup_{u \in S^{d-1}} u^T M u\right] \le 16 \int_0^1 \sqrt{d \log\left(1+\frac{2}{\varepsilon}\right)}\, d\varepsilon$$

$$\le C \cdot \sqrt{d} \quad !$$

And

$$\mathbb{E}\left[\sup_{u \in S^{d-1}} u^T M u\right] \ge C' \sqrt{d} .$$