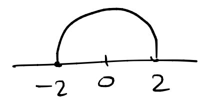
LECTURE 26

Last class: spectrum of a random matrix

W: $n\times n$ symmetric, iid entries

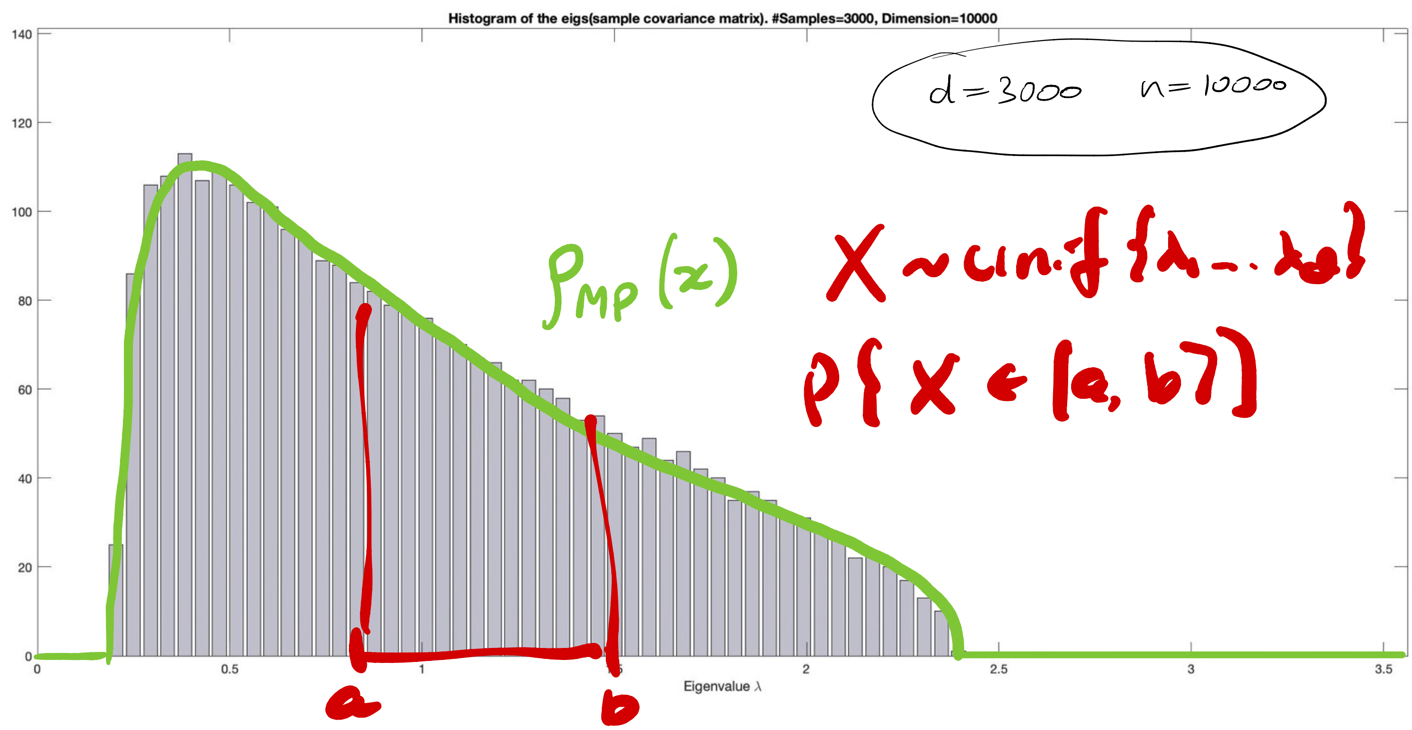
$\Rightarrow$ histogram of eigs$\left(\frac{1}{\sqrt{n}}W\right)$ $\rightarrow$ 

"Semicircle law"

Today: sample covariance matrix

$$\Sigma_n = \frac{1}{n}\sum_{i=1}^{n} X_i X_i^T$$

where $X_i \sim N(0, I_d)$ iid



THM [Marchenko-Pastur law '1967] – informally

If $d/n \to r \in (0,1)$ as $n \to \infty$,

the histogram of the eigenvalues of Σ_n converges to the density

$$\rho_{MP}(x) = \frac{1}{2\pi r x}\sqrt{(x-a)(b-x)} \quad \text{on } [a,b]$$

where $a = (1-\sqrt{r})^2$, $b = (1+\sqrt{r})^2$.

BACKGROUND

① Eigenvalues of a $d \times d$ matrix M:

- $\lambda_i(M - zI) = \lambda_i(M) - z$
- $\lambda_i(M^{-1}) = \dfrac{1}{\lambda_i(M)}$
- $\lambda_i(A+B) \neq \lambda_i(A) + \lambda_i(B)$

② Trace:

- $tr(M) = \sum_{i=1}^{d} M_{ii} = \sum_{k=1}^{d} \lambda_k$ ← eigenvalues of M
- Linearity: $tr(\alpha A + \beta B) = \alpha\, tr(A) + \beta\, tr(B)$
- Cyclic property: $tr(ABC) = tr(CAB) = tr(BCA)$
- $tr(xx^T) = tr(\underbrace{x^T x}_{\text{number}}) = x^T x = \|x\|_2^2$
- If $g \sim N(0, I_d) \Rightarrow \mathbb{E} g^T M g = tr(M)$

③ Rank-one update of the inverse:

- Sherman-Morrison formula: $\forall x, y \in \mathbb{R}^d$:

$$\Big(M + \underbrace{xy^T}_{\text{rank } 1}\Big)^{-1} = M^{-1} - \frac{\overbrace{M^{-1}xy^T M^{-1}}^{\text{rank } 1}}{1 + y^T M^{-1} x}$$

- In particular,

$$y^T (M + xy^T)^{-1} x = \underbrace{y^T M^{-1} x}_{=:q} - \frac{y^T M^{-1} x y^T M^{-1} x}{1 + y^T M^{-1} x} = q - \frac{q^2}{1+q} = \frac{q}{1+q} = 1 - \frac{1}{1+q}$$

$$= \boxed{1 - \frac{1}{1 + y^T M^{-1} x}}$$

③ Stieltjes transform of a r.v. X:

- $s(z) \overset{\text{def}}{=} \mathbb{E}\left[\frac{1}{X - z}\right]$, $z \in \mathbb{C}$.
- Ex: if $X \sim \mathrm{Unif}\{\lambda_1, \dots, \lambda_d\}$, $s(z) = \frac{1}{d} \sum_{i=1}^{d} \frac{1}{\lambda_i - z}$. (*)
- Ex if λ_i = eigs of M, $\frac{1}{\lambda_i - z}$ = eigs of $(M - zI)^{-1}$ $\Rightarrow$ $\boxed{s(z) = \frac{1}{d} \operatorname{tr} (M - zI)^{-1}}$
- Ex: if X has density p_{MP}, $s(z) = \int_a^b \frac{p_{MP}(x)}{x - z} dx = \frac{1 - z - r + \sqrt{(1-z-r)^2 - 4zr}}{2}$ (Ex)

- S.T. determines the distribution of X uniquely:

  (a) if $X \sim \mathrm{Unif}\{\lambda_1, \dots, \lambda_d\}$ then λ_k = poles of $s(z)$ (see (*))

  (b) if X has density $p(x)$, then "Stieltjes inversion":

$$p(x) = \lim_{\varepsilon \downarrow 0} \frac{s(x - i\varepsilon) - s(x + i\varepsilon)}{2i\pi}.$$

Proof of M-P Theorem (p.1) according to [P. Yaskov, A short proof of the M.-P. Theorem '2015]

- $\Sigma_n = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T$ $d\times d$ symmetric.

let $X \sim \mathrm{Unif}\{\lambda_1, \dots, \lambda_d\}$ $\Rightarrow$ Stieltjes transform:

eigenvalues of Σ_n

- $S_n(z) \overset{p.2}{=} \frac{1}{d}\,\mathrm{tr}(\Sigma_n - zI)^{-1} = \underbrace{\frac{n}{d}}_{= r}\,\mathrm{tr}\Big(\underbrace{\boxed{\sum_{i=1}^{n} x_i x_i^T - znI}}_{=: A}\Big)^{-1} = \frac{1}{r}\,\mathrm{tr}\,A^{-1}.$

$\Big\Vert$ if $\exists$ limit

- $S_{n-1}(z) = \underbrace{\frac{n-1}{d}}_{\approx r}\,\mathrm{tr}\Big(\underbrace{\boxed{\sum_{i=1}^{n-1} x_i x_i^T - znI}}_{=: B}\Big)^{-1} \simeq \frac{1}{r}\,\mathrm{tr}\,B^{-1}$ $\Rightarrow \mathrm{tr}\,A^{-1} \simeq \mathrm{tr}\,B^{-1}$. **(*)**

- $x_n^T A^{-1} x_n = x_n^T (B + x_n x_n^T)^{-1} x_n = 1 - \dfrac{1}{1 + x_n^T B^{-1} x_n}$ (Sherman-Morrison p. 3)

- x_n is independent of B $\Rightarrow$ conditionally on B,

$$\mathbb{E}\, x_n^T B^{-1} x_n \overset{p.2}{=} \mathrm{tr}\,B^{-1} \overset{(*)}{\simeq} \mathrm{tr}\,A^{-1}.$$

$\wr\wr$ concentration

$x_n^T B^{-1} x_n$

- $\Rightarrow$ $x_n^T A^{-1} x \simeq 1 - \dfrac{1}{1 + \mathrm{tr}\,A^{-1}}$

Similarly, $x_k^T A^{-1} x \simeq \overset{4}{\text{same}}$. Sum them up $\Rightarrow$

$$\underbrace{\sum_{k=1}^{n} x_k^T A^{-1} x_k}_{\parallel} \simeq n\left[1 - \frac{1}{1 + \mathrm{tr}(A^{-1})}\right]$$ $\leftarrow$ number

$$\mathrm{tr}\left(\sum_{k=1}^{n} x_k^T A^{-1} x_k\right) = \mathrm{tr}\Big(\underbrace{\sum_{k=1}^{n} x_k x_k^T}_{= A + znI} A^{-1}\Big)$$ (linearity & cyclic property)

$$= \operatorname{tr}\left[(A + znI)A^{-1}\right] = \operatorname{tr}\left(I_d + znA^{-1}\right) = d + zn \operatorname{tr} A^{-1}$$

- We showed:

$$d + zn \operatorname{tr} A^{-1} \approx n\left[1 - \frac{1}{1 + \operatorname{tr} A^{-1}}\right]$$

$\Rightarrow$ Quadratic equation in $\operatorname{tr} A^{-1}$. Solve, use $s_n(z) = \frac{1}{r} \operatorname{tr} A^{-1} \Rightarrow$

$$s_n(z) = \frac{1 - z - r + \sqrt{(1-z-r)^2 - 4zr}}{2}$$

Same as for ρ_{MP} (p. 3)

Uniqueness $\Rightarrow$ QED.

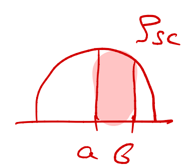
Rigorous Statements: Move them to the front.

- Semicircle law: With probability 1, $\forall$ interval $[a,b) \subset \mathbb{R}$:

$$\lim_{n\to\infty} \frac{\#\{i : \lambda_i(\frac{1}{\sqrt{n}}W) \in [a,b)\}}{n} = \int_a^b \rho_{sc}(x)\,dx$$

fraction of eigs $\in [a,b]$

fraction of area



$\parallel$

$\mathbb{P}\{X \in [a,b]\}$ where $X \sim \mathrm{Unif}\{\lambda_1, \dots, \lambda_n\}$

- Marchenko-Pastur law: with probability 1, $\forall$ interval $[a,b] \subset \mathbb{R}$:

$$\lim_{n\to\infty} \frac{\#\{i : \lambda_i(\Sigma_n) \in [a,b)\}}{d} = \int_a^b \rho_{MP}(x)\,dx$$

fraction of eigs $\in [a,b]$

fraction of area