20170058 **Keonwoo Kim** · May 31, 2020
MAS583(A) Random Matrix Theory and Its Application

Homework 4

1. (Cauchy interlacing law) We may write A_N as follows since A_N is Hermitian:

$$A_N = \begin{bmatrix} a_{11} & \mathbf{v}^\dagger \\ \mathbf{v} & A_{N-1} \end{bmatrix},$$

where $\dagger$ denotes the conjugate transpose. Since A_{N-1} is also Hermitian, it can be unitarily diagonalized:

$$A_{N-1} = U^\dagger D U, \qquad U \text{ unitary}, \quad D = \operatorname{diag}(\lambda_1(A_{N-1}), \ldots, \lambda_{N-1}(A_{N-1})).$$

Letting

$$V = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & U \end{bmatrix},$$

we have

$$V A_N V^\dagger = \begin{bmatrix} a_{11} & (U\mathbf{v})^\dagger \\ U\mathbf{v} & D \end{bmatrix}.$$

Letting $\mathbf{w} = (w_1, \ldots, w_{N-1})^T = U\mathbf{v}$, the characteristic polynomial of A_N is

$$\begin{aligned} p(t) &:= \det(tI - A_N) = \det(tI - V A_N V^\dagger) \\ &= (t - a_{11}) \det(tI - D) + \sum_{i=1}^{N-1} \overline{w_i}\, w_i \det(tI - D^{(i)}) \\ &= (t - a_{11})(t - \lambda_1(A_{N-1})) \cdots (t - \lambda_{N-1}(A_{N-1})) \\ &\quad - \sum_{i=1}^{N-1} |w_i|^2 \prod_{1 \le j \le N-1,\, j \ne i} (t - \lambda_j(A_{N-1})). \end{aligned}$$

Assume that w_i's are nonzero and

$$\lambda_1(A_{N-1}) > \cdots > \lambda_{N-1}(A_{N-1}).$$

Then,

$$p(\lambda_i(A_{N-1})) = -|w_i|^2 \prod_{1\le j\le N-1,\, j\ne i} (\lambda_i(A_{N-1}) - \lambda_j(A_{N-1})) \begin{cases} < 0, & \text{if } i \text{ is odd} \\ > 0, & \text{if } i \text{ is even} \end{cases}.$$

Since $p(t)$ is monic polynomial of degree n, by the intermediate value theorem, $p(t)$ has N distinct roots $\lambda_i(A_N)$, $i = 1, \dots, N$, where $\lambda_{i+1}(A_N) < \lambda_i(A_{N-1}) < \lambda_i(A_N)$ for any $i = 1, \dots, N-1$.

For the general case, consider the perturbation

$$A_N^{(\epsilon)} = A_N + V^\dagger \begin{bmatrix} 0 & \epsilon & \cdots & \epsilon \\ \epsilon & -\epsilon & & \mathbf{0} \\ \vdots & & \ddots & \\ \epsilon & \mathbf{0} & & -(N-1)\epsilon \end{bmatrix} V = V^\dagger \begin{bmatrix} a_{11} & \mathbf{w}^T + \boldsymbol{\epsilon}^T \\ \mathbf{w} + \boldsymbol{\epsilon} & D^{(\epsilon)} \end{bmatrix} V$$

where $\epsilon > 0$, $\boldsymbol{\epsilon} = \epsilon(1, \dots, 1)^T$ and $D^{(\epsilon)} = D - \epsilon\,\mathrm{diag}(1, \dots, N-1)$. Then,

$$\begin{aligned} \det(tI - A_N^{(\epsilon)}) &= \det(tI - V A_N^{(\epsilon)} V^\dagger) \\ &= (t - a_{11}) \prod_{i=1}^{N-1} (t - (\lambda_i(A_{N-1}) - i\epsilon)) \\ &\quad - \sum_{i=1}^{N-1} |w_i + \epsilon|^2 \prod_{1\le j\le N-1,\, j\ne i} (t - (\lambda_j(A_{N-1}) - j\epsilon)). \end{aligned}$$

Assume $w_i + \epsilon \ne 0$ for any $i = 1, \dots, N-1$. As above, putting $t = \lambda_i(A_{N-1}) - i\epsilon$, $i = 1, \dots, N-1$, the characteristic polynomial of $A_N^{(\epsilon)}$ has N distinct roots satisfying the following:

$$\begin{aligned} \lambda_N(A_N^{(\epsilon)}) &< \lambda_{N-1}(A_{N-1}) - (N-1)\epsilon < \lambda_{N-1}(A_N^{(\epsilon)}) \\ &< \cdots < \lambda_2(A_N^{(\epsilon)}) < \lambda_1(A_1) - \epsilon < \lambda_1(A_N^{(\epsilon)}). \end{aligned} \qquad (*)$$

(Note that $\lambda_{i+1}(A_{N-1}) \le \lambda_i(A_{N-1})$ implies $\lambda_{i+1}(A_{N-1}) - (i+1)\epsilon < \lambda_i(A_{N-1}) - i\epsilon$.) By the Courant-Fischer min-max principle from the linear algebra, we have

$$\lambda_k(A_N) = \min_{\dim V = k-1} \max_{x \in V^\perp, \|x\|=1} x^\dagger A x.$$

(Note that the Hermitian matrix is a (compact) self-adjoint operator.) However, the operator norm of $A_N - A_N^{(\epsilon)}$ is $O(\epsilon)$, which implies

$$|x^\dagger (A_N - A_N^{(\epsilon)}) x| \le \|x\| \|(A_N - A_N^{(\epsilon)}) x\| \le O(\epsilon) \|x\|^2$$

so that from the min-max principle, for any $k = 1, \dots, N$,

$$|\lambda_k(A_N) - \lambda_k(A_N^{(\epsilon)})| \le O(\epsilon).$$

Therefore, as $\epsilon \downarrow 0$, $(*)$ implies

$$\lambda_N(A_N) \le \lambda_{N-1}(A_{N-1}) \le \lambda_{N-1}(A_N) \le \cdots \le \lambda_2(A_N) \le \lambda_1(A_{N-1}) \le \lambda_1(A_N),$$

that is, the eigenvalues of A_N and A_{N-1} interlace.