Eigenvalues of Random Matrices under Finite-Rank Perturbations

18.338 Project Presentation

Xiaomin Li [1] Yi Tian [2]

[1]School of Engineering and Applied Sciences, Harvard University

[2]EECS Department, MIT

Random matrices under finite-rank perturbations

- Let X_n be an $n \times n$ Hermite ensemble (Wigner, Wishart, Jacobi...).
- For a fixed $r \geq 1$, let $\theta_1 \geq \cdots \geq \theta_s > 0 > \theta_{s+1} \geq \cdots \geq \theta_r$ be deterministic nonzero real numbers.
- Let P_n be an $n \times n$ Hermite ensemble that has rank r and $\theta_1, \ldots, \theta_r$ as its nonzero eigenvalues.
- X_n and P_n are independent.
- How are the eigenvalues of $X_n + P_n$ and $X_n(I_n + P_n)$ distributed as $n \to \infty$?

Literature: [BGN11]

Benaych-Georges, Florent, and Raj Rao Nadakuditi. "The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices." Advances in Mathematics 227.1 (2011): 494-521.

- Further assume that either X_n or P_n is unitarily invariant.
- Also studied the eigenvectors.
- Phase transition: generalization of [BBAP05].

Theorem (informal, largest eigenvalues)

Let $\widetilde{X}_n$ be either $X_n + P_n$ or $X_n(I_n + P_n)$. Let b be the supremum of the support of the limiting eigenvalue distribution of X_n. For each $1 \le i \le s$, there exists a threshold θ_c and some function f such that

$$\lambda_i(\widetilde{X}_n) \xrightarrow{\text{a.s.}} \begin{cases} f(\theta_i) & \text{if } \theta_i > \theta_c \\ b & \text{otherwise,} \end{cases}$$

while for fixed $i > s$, $\lambda_i(\widetilde{X}_n) \xrightarrow{\text{a.s.}} b$, as $n \to \infty$.

Experiments

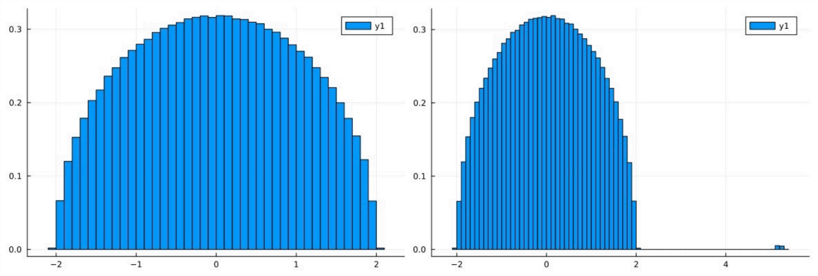

Figure 1: Histograms of the eigenvalues of real Hermite ensembles X_n perturbed by additive $\theta u_n u_n^\top$. Left: $\theta = 0.5$. Right: $\theta = 5$.

Progress

- Confirm the known theoretical results with experiments.

Progress

- Confirm the known theoretical results with experiments.
- Understand the proof techniques in [BGN11].
- Study a new question: perturbing random matrices whose eigenvalue distributions have multiple bulks.

Main idea in the analysis in [BGN11] (1/3)

By unitary invariance, w.l.o.g., suppose $X_n = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Consider rank-2 perturbation $P_n = \theta_1 u_n u_n^\top + \theta_2 v_n v_n^\top$.

Main idea in the analysis in [BGN11] (1/3)

By unitary invariance, w.l.o.g., suppose $X_n = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Consider rank-2 perturbation $P_n = \theta_1 u_n u_n^\top + \theta_2 v_n v_n^\top$.

The eigenvalues of $X_n + P_n$ are solutions to the equation

$$\det(zI - (X_n + P_n)) = 0.$$

Main idea in the analysis in [BGN11] (1/3)

By unitary invariance, w.l.o.g., suppose $X_n = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$.

Consider rank-2 perturbation $P_n = \theta_1 u_n u_n^\top + \theta_2 v_n v_n^\top$.

The eigenvalues of $X_n + P_n$ are solutions to the equation

$$\det(zI - (X_n + P_n)) = 0.$$

If $zI - X_n$ is invertible, we have

$$zI - (X_n + P_n) = (zI - X_n)(I - (zI - X_n)^{-1} P_n).$$

$$\det(zI - (X_n + P_n)) = \det(zI - X_n) \cdot \det(I - (zI - X_n)^{-1} P_n).$$

Main idea in the analysis in [BGN11] (1/3)

By unitary invariance, w.l.o.g., suppose $X_n = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Consider rank-2 perturbation $P_n = \theta_1 u_n u_n^\top + \theta_2 v_n v_n^\top$.

The eigenvalues of $X_n + P_n$ are solutions to the equation

$$\det(zI - (X_n + P_n)) = 0.$$

If $zI - X_n$ is invertible, we have

$$zI - (X_n + P_n) = (zI - X_n)(I - (zI - X_n)^{-1} P_n).$$
$$\det(zI - (X_n + P_n)) = \det(zI - X_n) \cdot \det(I - (zI - X_n)^{-1} P_n).$$

This means z is an eigenvalue of $X_n + P_n$ and not an eigenvalue of X_n if and only if $\det(I - (zI - X_n)^{-1} P_n) = 0$.

Main idea in the analysis in [BGN11] (2/3)

$$\begin{aligned}
&\det(I - (zI - X_n)^{-1} P_n) \\
= \ &\det(I - (zI - X_n)^{-1}(\theta_1 u_n u_n^\top + \theta_2 v_n v_n^\top)) \\
\overset{(i)}{=} \ &\det\left(I - [u_n, v_n]^\top \operatorname{diag}\big((z-\lambda_1)^{-1}, \ldots, (z-\lambda_n)^{-1}\big)[\theta_1 u_n, \theta_2 v_n]\right) \\
= \ &\det\left(\begin{bmatrix} 1 - \theta_1 \sum_{i=1}^n u_i^2 (z-\lambda_i)^{-1} & \theta_2 \sum_{i=1}^n u_i v_i (z-\lambda_i)^{-1} \\ \theta_1 \sum_{i=1}^n u_i v_i (z-\lambda_i)^{-1} & 1 - \theta_2 \sum_{i=1}^n v_i^2 (z-\lambda_i)^{-1} \end{bmatrix}\right),
\end{aligned}$$

where (i) is due to Sylvester's determinant identity.

Main idea in the analysis in [BGN11] (2/3)

$$
\begin{aligned}
& \det(I - (zI - X_n)^{-1} P_n) \\
= \; & \det(I - (zI - X_n)^{-1}(\theta_1 u_n u_n^\top + \theta_2 v_n v_n^\top)) \\
\overset{(i)}{=} \; & \det\left(I - [u_n, v_n]^\top \mathrm{diag}\big((z-\lambda_1)^{-1}, \ldots, (z-\lambda_n)^{-1}\big)[\theta_1 u_n, \theta_2 v_n]\right) \\
= \; & \det\left(\begin{bmatrix} 1 - \theta_1 \sum_{i=1}^n u_i^2 (z-\lambda_i)^{-1} & \theta_2 \sum_{i=1}^n u_i v_i (z-\lambda_i)^{-1} \\ \theta_1 \sum_{i=1}^n u_i v_i (z-\lambda_i)^{-1} & 1 - \theta_2 \sum_{i=1}^n v_i^2 (z-\lambda_i)^{-1} \end{bmatrix} \right),
\end{aligned}
$$

where (i) is due to Sylvester's determinant identity. In the limit,

$$
\sum_{i=1}^n u_i v_i (z-\lambda_i)^{-1} \overset{a.s.}{\to} 0, \quad \sum_{i=1}^n u_i^2 (z-\lambda_i)^{-1} \overset{a.s.}{\to} n^{-1} \sum_{i=1}^n (z-\lambda_i)^{-1}.
$$

Main idea in the analysis in [BGN11] (2/3)

$$
\begin{aligned}
& \det(I - (zI - X_n)^{-1} P_n) \\
= \ & \det(I - (zI - X_n)^{-1} (\theta_1 u_n u_n^\top + \theta_2 v_n v_n^\top)) \\
\overset{(i)}{=} \ & \det\left(I - [u_n, v_n]^\top \operatorname{diag}\big((z - \lambda_1)^{-1}, \ldots, (z - \lambda_n)^{-1}\big)[\theta_1 u_n, \theta_2 v_n]\right) \\
= \ & \det\left(\begin{bmatrix} 1 - \theta_1 \sum_{i=1}^n u_i^2 (z - \lambda_i)^{-1} & \theta_2 \sum_{i=1}^n u_i v_i (z - \lambda_i)^{-1} \\ \theta_1 \sum_{i=1}^n u_i v_i (z - \lambda_i)^{-1} & 1 - \theta_2 \sum_{i=1}^n v_i^2 (z - \lambda_i)^{-1} \end{bmatrix} \right),
\end{aligned}
$$

where (i) is due to Sylvester's determinant identity. In the limit,

$$
\sum_{i=1}^n u_i v_i (z - \lambda_i)^{-1} \overset{a.s.}{\to} 0, \quad \sum_{i=1}^n u_i^2 (z - \lambda_i)^{-1} \overset{a.s.}{\to} n^{-1} \sum_{i=1}^n (z - \lambda_i)^{-1}.
$$

Then, $\det(I - (zI - X_n)^{-1} P_n) = 0$ if and only if

$$
1 - \theta_1 n^{-1} \sum_{i=1}^n (z - \lambda_i)^{-1} = 0, \text{ or } 1 - \theta_2 n^{-1} \sum_{i=1}^n (z - \lambda_i)^{-1} = 0.
$$

Main idea in the analysis in [BGN11] (3/3)

z is an eigenvalue of $X_n + P_n$ and not an eigenvalue of X_n

$$\iff \det(I - (zI - X_n)^{-1} P_n) = 0$$

$$\iff 1 - \theta_1 n^{-1} \sum_{i=1}^{n} (z - \lambda_i)^{-1} = 0, \text{ or } 1 - \theta_2 n^{-1} \sum_{i=1}^{n} (z - \lambda_i)^{-1} = 0.$$

Main idea in the analysis in [BGN11] (3/3)

z is an eigenvalue of $X_n + P_n$ and not an eigenvalue of X_n

$$\iff \det(I - (zI - X_n)^{-1}P_n) = 0$$

$$\iff 1 - \theta_1 n^{-1} \sum_{i=1}^{n} (z - \lambda_i)^{-1} = 0, \text{ or } 1 - \theta_2 n^{-1} \sum_{i=1}^{n} (z - \lambda_i)^{-1} = 0.$$

Let $\mu_{X_n} := n^{-1} \sum_{i=1}^{n} \delta_{\lambda_i(X_n)}$. Then, its Cauchy transform is precisely

$$G_{\mu_{X_n}}(z) = \int (z - t)^{-1} \mathrm{d}\mu_{X_n}(t) = n^{-1} \sum_{i=1}^{n} (z - \lambda_i)^{-1}.$$

Main idea in the analysis in [BGN11] (3/3)

z is an eigenvalue of $X_n + P_n$ and not an eigenvalue of X_n

$$\iff \det(I - (zI - X_n)^{-1} P_n) = 0$$
$$\iff 1 - \theta_1 n^{-1} \sum_{i=1}^{n} (z - \lambda_i)^{-1} = 0, \text{ or } 1 - \theta_2 n^{-1} \sum_{i=1}^{n} (z - \lambda_i)^{-1} = 0.$$

Let $\mu_{X_n} := n^{-1} \sum_{i=1}^{n} \delta_{\lambda_i(X_n)}$. Then, its Cauchy transform is precisely

$$G_{\mu_{X_n}}(z) = \int (z - t)^{-1} \mathrm{d}\mu_{X_n}(t) = n^{-1} \sum_{i=1}^{n} (z - \lambda_i)^{-1}.$$

In the limit of $n \to \infty$, $G_{\mu_{X_n}} \to G_{\mu_X}$, where μ_X is the limiting eigenvalue distribution of X_n. Hence, new limiting eigenvalue z of $X_n + P_n$ satisfy

$$G_{\mu_X}(z) = \theta_1^{-1}, \text{ or } G_{\mu_X}(z) = \theta_2^{-1}.$$

Results in the additive case in [BGN11]

Theorem (informal, largest eigenvalues)

Let $\widetilde{X}_n$ be either $X_n + P_n$. Let b be the supremum of the support of the limiting eigenvalue distribution of X_n. For each $1 \leq i \leq s$,

$$\lambda_i(\widetilde{X}_n) \xrightarrow{\text{a.s.}} \begin{cases} G_{\mu_X}^{-1}(\theta_i^{-1}) & \text{if } \theta_i > (G_{\mu_X}(b^+))^{-1}, \\ b & \text{otherwise}, \end{cases}$$

while for fixed $i > s$, $\lambda_i(\widetilde{X}_n) \xrightarrow{\text{a.s.}} b$, as $n \to \infty$.

Results in the additive case in [BGN11]

Theorem (informal, largest eigenvalues)

Let $\widetilde{X}_n$ be either $X_n + P_n$. Let b be the supremum of the support of the limiting eigenvalue distribution of X_n. For each $1 \leq i \leq s$,

$$\lambda_i(\widetilde{X}_n) \xrightarrow{\text{a.s.}} \begin{cases} G_{\mu_X}^{-1}(\theta_i^{-1}) & \text{if } \theta_i > (G_{\mu_X}(b^+))^{-1}, \\ b & \text{otherwise}, \end{cases}$$

while for fixed $i > s$, $\lambda_i(\widetilde{X}_n) \xrightarrow{\text{a.s.}} b$, as $n \to \infty$.

For Wigner matrices, μ_X is the semicircle distribution, and

$$G_{\mu_X}(z) = \frac{z - \operatorname{sign}(z)\sqrt{z^2 - 4\sigma^2}}{2\sigma^2}.$$

For $0 < z \leq 2\sigma$, $G_{\mu_X}^{-1}(z) = \sigma^2 z + z^{-1}$.

Experiments

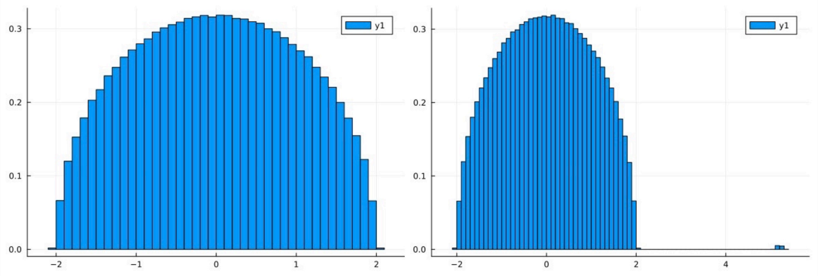

Figure 2: Histograms of the eigenvalues of real Hermite ensembles X_n perturbed by additive $\theta u_n u_n^\top$. Left: $\theta = 0.5$ (theory: no spike). Right: $\theta = 5$ (theory: spike at 5.2).

Perturbing random matrices with two bulks

Consider $X_{2n} = \mathrm{diag}(Y_n, Z_n)$ where Y_n, Z_n are diagonal and two bulks.

Rank-2 perturbation $P_{2n} = \theta_1 u_{2n} u_{2n}^\top + \theta_2 v_{2n} v_{2n}^\top$.

Perturbing random matrices with two bulks

Consider $X_{2n} = \mathrm{diag}(Y_n, Z_n)$ where Y_n, Z_n are diagonal and two bulks.

Rank-2 perturbation $P_{2n} = \theta_1 u_{2n} u_{2n}^\top + \theta_2 v_{2n} v_{2n}^\top$.

z is a limit eigenvalue of $X_n + P_n$ and not a limit eigenvalue of X_n if and only if $G_{\mu_X}(z) = \theta_1^{-1}$, or $G_{\mu_X}(z) = \theta_2^{-1}$. So, reduces to finding G_{μ_X}.

Perturbing random matrices with two bulks

Consider $X_{2n} = \text{diag}(Y_n, Z_n)$ where Y_n, Z_n are diagonal and two bulks.

Rank-2 perturbation $P_{2n} = \theta_1 u_{2n} u_{2n}^\top + \theta_2 v_{2n} v_{2n}^\top$.

z is a limit eigenvalue of $X_n + P_n$ and not a limit eigenvalue of X_n if and only if $G_{\mu_X}(z) = \theta_1^{-1}$, or $G_{\mu_X}(z) = \theta_2^{-1}$. So, reduces to finding G_{μ_X}.

Let $\mu_{Y_n} = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i(X_{2n})}$ and $\mu_{Z_n} = \frac{1}{n}\sum_{i=n+1}^{2n} \delta_{\lambda_i(X_{2n})}$ Then,

$$\mu_{X_n} = \frac{1}{2n}\sum_{i=1}^{2n} \delta_{\lambda_i(X_{2n})} = \frac{1}{2}(\mu_{Y_n} + \mu_{Z_n}),$$

Perturbing random matrices with two bulks

Consider $X_{2n} = \text{diag}(Y_n, Z_n)$ where Y_n, Z_n are diagonal and two bulks.

Rank-2 perturbation $P_{2n} = \theta_1 u_{2n} u_{2n}^\top + \theta_2 v_{2n} v_{2n}^\top$.

z is a limit eigenvalue of $X_n + P_n$ and not a limit eigenvalue of X_n if and only if $G_{\mu_X}(z) = \theta_1^{-1}$, or $G_{\mu_X}(z) = \theta_2^{-1}$. So, reduces to finding G_{μ_X}.

Let $\mu_{Y_n} = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i(X_{2n})}$ and $\mu_{Z_n} = \frac{1}{n}\sum_{i=n+1}^{2n} \delta_{\lambda_i(X_{2n})}$ Then,

$$\mu_{X_n} = \frac{1}{2n}\sum_{i=1}^{2n} \delta_{\lambda_i(X_{2n})} = \frac{1}{2}(\mu_{Y_n} + \mu_{Z_n}),$$

Let μ_X, μ_Y, μ_Z be the limiting distributions of $\mu_{X_{2n}}, \mu_{Y_n}, \mu_{Z_n}$. We have

$$G_{\mu_X} = \frac{1}{2}(G_{\mu_Y} + G_{\mu_Z}).$$

Numerical example: Wigner matrices

Suppose μ_Z, μ_Y are two semicircle distributions in $[-2, 2], [6, 10]$. Then

$$G_{\mu_Z}(z) = \frac{z - \operatorname{sign}(z)\sqrt{z^2 - 4}}{2}, \quad z \leq -2, \text{ or } z \geq 2,$$

$$G_{\mu_Y}(z) = \frac{z - 8 - \operatorname{sign}(z-8)\sqrt{(z-8)^2 - 4}}{2}, \quad z \leq 6, \text{ or } z \geq 10.$$

Numerical example: Wigner matrices

Suppose μ_Z, μ_Y are two semicircle distributions in $[-2, 2], [6, 10]$. Then

$$G_{\mu_Z}(z) = \frac{z - \operatorname{sign}(z)\sqrt{z^2 - 4}}{2}, \quad z \leq -2, \text{ or } z \geq 2,$$
$$G_{\mu_Y}(z) = \frac{z - 8 - \operatorname{sign}(z-8)\sqrt{(z-8)^2 - 4}}{2}, \quad z \leq 6, \text{ or } z \geq 10.$$

Hence,

$$G_{\mu_X}(z) = \frac{2z - 8 - \operatorname{sign}(z)\sqrt{z^2 - 4} - \operatorname{sign}(z-8)\sqrt{(z-8)^2 - 4}}{4},$$

where $z \leq -2$, $2 \leq z \leq 6$, or $z \geq 10$.

Numerical example: Wigner matrices

Suppose μ_Z, μ_Y are two semicircle distributions in $[-2,2], [6,10]$. Then

$$G_{\mu_Z}(z) = \frac{z - \operatorname{sign}(z)\sqrt{z^2-4}}{2}, \quad z \le -2, \text{ or } z \ge 2,$$
$$G_{\mu_Y}(z) = \frac{z - 8 - \operatorname{sign}(z-8)\sqrt{(z-8)^2-4}}{2}, \quad z \le 6, \text{ or } z \ge 10.$$

Hence,

$$G_{\mu_X}(z) = \frac{2z - 8 - \operatorname{sign}(z)\sqrt{z^2-4} - \operatorname{sign}(z-8)\sqrt{(z-8)^2-4}}{4},$$

where $z \le -2$, $2 \le z \le 6$, or $z \ge 10$. In particular,

$$G(-2) = -G(10) = \sqrt{6} - 3 \approx -0.55, \; G(2) = -G(6) = \sqrt{2} - 1 \approx 0.41.$$

Numerical example: Wigner matrices

Suppose μ_Z, μ_Y are two semicircle distributions in $[-2,2], [6,10]$. Then

$$G_{\mu_Z}(z) = \frac{z - \operatorname{sign}(z)\sqrt{z^2-4}}{2}, \quad z \le -2, \text{ or } z \ge 2,$$

$$G_{\mu_Y}(z) = \frac{z - 8 - \operatorname{sign}(z-8)\sqrt{(z-8)^2-4}}{2}, \quad z \le 6, \text{ or } z \ge 10.$$

Hence,

$$G_{\mu_X}(z) = \frac{2z - 8 - \operatorname{sign}(z)\sqrt{z^2-4} - \operatorname{sign}(z-8)\sqrt{(z-8)^2-4}}{4},$$

where $z \le -2$, $2 \le z \le 6$, or $z \ge 10$. In particular,

$$G(-2) = -G(10) = \sqrt{6} - 3 \approx -0.55, \; G(2) = -G(6) = \sqrt{2} - 1 \approx 0.41.$$

- $|\theta_i|^{-1} \ge 0.55$: no new eigenvalue;
- $0.41 \le |\theta_i|^{-1} < 0.55$: a spike outside $[-2,10]$ but none in $[2,6]$;
- $|\theta_i|^{-1} < 0.41$, both a spike outside $[-2,10]$ and one in $[2,6]$.

Jupyter notebook

Jinho Baik, Gérard Ben Arous, and Sandrine Péché.
Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices.
The Annals of Probability, 33(5):1643–1697, 2005.

Florent Benaych-Georges and Raj Rao Nadakuditi.
The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices.
Advances in Mathematics, 227(1):494–521, 2011.