Bielefeld University

On the smallest singular value of large random matrices with correlated entries

Alexey Naumov

joint work with F. Götze and A. Tikhomirov

Bielefeld 7.08.2013

Random matrix ensemble

- We consider the random matrix $\mathbf{X}_n(\omega) = \{X_{ij}(\omega)\}_{i,j=1}^{n}$ and suppose that the conditions $(\mathbf{C0})$ hold:
  a) random vectors (X_{jk}, X_{kj}) are mutually independent for $1 \le j < k \le n$;
  b) for any $j, k = 1, ..., n$

$$\mathbb{E}\, X_{jk} = 0 \text{ and } \mathbb{E}\, X_{jk}^2 = 1;$$

  c) for any $1 \le j < k \le n$

$$\mathbb{E}(X_{12}X_{21}) = \rho, |\rho| \le 1;$$

- In the case $\rho = 1$ we have a symmetric matrix, if $\rho = 0$ and all random variables are Gaussian we have a matrix from Ginibre ensemble.

Eigenvalues

- Let $\lambda_1, ..., \lambda_n$ be the eigenvalues of the matrix $n^{-1/2}\mathbf{X}_n$. We define
- Empirical spectral measure

$$\mu_n(B) = \frac{1}{n}\#\{i : \lambda_i \in B\}, \quad B \in \mathcal{B}(\mathbb{T}),$$

where $\mathbb{T} = \mathbb{R}$ or $\mathbb{T} = \mathbb{C}$

Elliptic Law

- **Theorem (Elliptic Law)** *Let $\mathbf{X}_n$ satisfy $(\mathbf{C0})$ and $|\rho| < 1$. Then $\mu_n \xrightarrow{weak} \mu$ in probability, and μ has a density g:*

$$g(x,y) = \begin{cases} \frac{1}{\pi(1-\rho^2)}, & x, y \in \left\{u, v \in \mathbb{R} : \frac{u^2}{(1+\rho)^2} + \frac{v^2}{(1-\rho)^2} \leq 1\right\}, \\ 0, & \textit{elsewhere.} \end{cases}$$

  (under additional assumptions by Girko, 1985; N.,2012).
- If (X_{ij}, X_{ji}) - i.i.d.v, $\mathbb{E}\, X_{ij} = \mathbb{E}\, X_{ji} = 0, \mathbb{E}\, X_{ij}^2 = \mathbb{E}\, X_{ji}^2 = 1$, $\mathbb{E}(X_{ij} X_{ji}) = \rho$, $|\rho| \leq 1$ then the elliptic law was proved by H. Nguyen, S. O'Rourke (2012)
- if $\rho = 0$ we have circular law (Girko; Bai; Götze, Tikhomirov;Tao, Vu;...)
- if $\rho \to 1$ we have Wigner's Semi-Circular law

$$g(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4 - x^2}, & -2 \leq x \leq 2, \\ 0, & \text{otherwise.} \end{cases}$$

  (Wigner E.; Pastur;...)

Girko's hermitization

- We know that
$$|\det(\mathbf{A})| = \prod_{i=1}^{n} |\lambda_i(\mathbf{A})| = \prod_{i=1}^{n} s_i(A).$$

Girko's hermitization

- We know that
$$|\det(\mathbf{A})| = \prod_{i=1}^{n} |\lambda_i(\mathbf{A})| = \prod_{i=1}^{n} s_i(A).$$
-
$$U_{\mu_n}(z) = -\int_{\mathbb{C}} \log|z-w|\mu_n(dw) = -\frac{1}{n}\log\left|\det\left(\frac{1}{\sqrt{n}}\mathbf{X}_n - z\mathbf{I}\right)\right|$$
$$= -\frac{1}{2n}\log\det\left(\frac{1}{\sqrt{n}}\mathbf{X}_n - z\mathbf{I}\right)^*\left(\frac{1}{\sqrt{n}}\mathbf{X}_n - z\mathbf{I}\right) = -\int_0^\infty \log x \nu_n(dx)$$
- $\nu_n(\cdot, z)$ is the empirical spectral measure of the singular values of $(n^{-1/2}\mathbf{X}_n - z\mathbf{I})$.

Smallest singular value

- Let $s_k := s_k(\mathbf{W})$ be the singular values of $\mathbf{W} = \mathbf{X} + \mathbf{M}_n$, $s_n(\mathbf{W}) \leq \ldots \leq s_1(\mathbf{W})$, and

$$s_1(\mathbf{W}) = ||\mathbf{W}|| = \sup_{\mathbf{x}:||\mathbf{x}||_2=1} ||\mathbf{W}\mathbf{x}||_2, \quad s_n(\mathbf{W}) = \inf_{\mathbf{x}:||\mathbf{x}||_2=1} ||\mathbf{W}\mathbf{x}||_2.$$

Smallest singular value

- Let $s_k := s_k(\mathbf{W})$ be the singular values of $\mathbf{W} = \mathbf{X} + \mathbf{M}_n$, $s_n(\mathbf{W}) \leq ... \leq s_1(\mathbf{W})$, and

$$s_1(\mathbf{W}) = ||\mathbf{W}|| = \sup_{\mathbf{x}:||\mathbf{x}||_2=1} ||\mathbf{W}\mathbf{x}||_2, \quad s_n(\mathbf{W}) = \inf_{\mathbf{x}:||\mathbf{x}||_2=1} ||\mathbf{W}\mathbf{x}||_2.$$

- **T h e o r e m.** Let X_{jk}, $1 \leq j, k \leq n$, satisfy the conditions **(C0)**. Assume that the squares of X_{jk}'s are uniformly integrable , i.e.

$$\max_{j,k} \mathbb{E} |X_{jk}|^2 \mathbb{I}\{|X_{jk}| > M\} \to 0 \quad \text{as} \quad M \to \infty.$$

  Let $\mathbf{X} = \{X_{jk}\}$ be $n \times n$ random matrix with entries X_{jk} and let $\mathbf{M}_n$ denote a non-random matrix with $\|\mathbf{M}_n\| \leq Kn^Q =: K_n$ for some $K > 0$ and $Q \geq 0$. Then there exist constants $C, A, B > 0$ depending on K, Q such that

$$\mathbb{P}(s_n \leq n^{-B}) \leq Cn^{-A},$$

Compressible and incompressible vectors

- Following Rudelson and Vershynin we shall partition the unit sphere $\mathcal{S}^{(n-1)} := \{\mathbf{x} : ||\mathbf{x}||_2 = 1\}$ into two sets of so-called compressible and incompressible vectors.
- Let $\delta, r \in (0, 1)$. A vector $x \in \mathbb{R}^n$ is called δ-sparse if $|\operatorname{supp}(\mathbf{x})| \leq \delta n$.
- A vector $\mathbf{x} \in \mathcal{S}^{(n-1)}$ is called (δ, r)-compressible if $\mathbf{x}$ within Euclidean distance r from the set of all δ-sparse vectors.
- A vector $\mathbf{x} \in \mathcal{S}^{(n-1)}$ is called (δ, r)-incompressible if it is not (δ, r)-compressible.

Compressible and incompressible vectors

- Following Rudelson and Vershynin we shall partition the unit sphere $\mathcal{S}^{(n-1)} := \{\mathbf{x} : ||\mathbf{x}||_2 = 1\}$ into two sets of so-called compressible and incompressible vectors.
- Let $\delta, r \in (0,1)$. A vector $x \in \mathbb{R}^n$ is called δ-sparse if $|\operatorname{supp}(\mathbf{x})| \leq \delta n$.
- A vector $\mathbf{x} \in \mathcal{S}^{(n-1)}$ is called (δ, r)-compressible if $\mathbf{x}$ within Euclidean distance r from the set of all δ-sparse vectors.
- A vector $\mathbf{x} \in \mathcal{S}^{(n-1)}$ is called (δ, r)-incompressible if it is not (δ, r)-compressible.
- There exists a set $\sigma(\mathbf{x}) \in [n]$ of cardinality $|\sigma(x)| \geq \frac{1}{2} n\delta$ such that

$$\frac{r}{\sqrt{2n}} \leq |x_k| \leq \frac{1}{\sqrt{\delta n}} \text{ for any } k \in \sigma(x),$$

Subgaussian distribution

- for compressible vectors vectors

$$\inf_{\mathbf{x}\in\mathcal{C}(\delta,r)} ||\mathbf{W}\mathbf{x}||_2 \geq n^{1/2} \text{ with high probability}$$

- For incompressible vectors

$$\mathbb{P}\left(\inf_{\mathbf{x}\in\mathcal{IC}(\delta,r)} ||\mathbf{W}\mathbf{x}||_2 < \varepsilon n^{-1/2}\right) \leq \frac{1}{\delta n}\sum_{k=1}^{n}\mathbb{P}(\mathrm{dist}(\mathbf{W}_k, H_k) < r^{-1}\varepsilon),$$

where $\mathbf{W}_1, \ldots, \mathbf{W}_n$ denote the columns of $\mathbf{W}$ and let $\mathcal{H}_k$ denotes the span of all column vectors except the kth.

Procedure for two moments

- We start with $\delta_n^{(1)} \backsim \frac{1}{\ln n}, r_n^{(1)} \backsim n^{-Q_1}, Q_1 > 0$ and consider the set $\mathcal{C}_0 := \mathcal{C}(\delta_n^{(1)}, r_n^{(1)})$. We may show that

$$\mathbb{P}(\inf_{\mathbf{x}\in\mathcal{C}_0} \|\mathbf{W}\mathbf{x} - \mathbf{y}\|_2/\|\mathbf{x}\|_2 \leq \tau\sqrt{n}, \mathcal{E}_K) \leq \exp\{-cn\}.$$

Procedure for two moments

- We start with $\delta_n^{(1)} \backsim \frac{1}{\ln n}, r_n^{(1)} \backsim n^{-Q_1}, Q_1 > 0$ and consider the set $\mathcal{C}_0 := \mathcal{C}(\delta_n^{(1)}, r_n^{(1)})$. We may show that

$$\mathbb{P}(\inf_{\mathbf{x}\in\mathcal{C}_0} \|\mathbf{W}\mathbf{x} - \mathbf{y}\|_2/\|\mathbf{x}\|_2 \leq \tau\sqrt{n}, \mathcal{E}_K) \leq \exp\{-cn\}.$$

- Now we take

$$\delta_n^{(2)} \backsim \frac{\ln(n\delta_n^{(1)})}{\ln n} \text{ and } r_n^{(2)} \backsim n^{-Q_2}, Q_2 > 0.$$

and show that for $\mathcal{C}_1 = \mathcal{C}(\delta_n^{(2)}, r_n^{(2)}) \cap \mathcal{IC}(\delta_n^{(1)}, r_n^{(1)})$

$$\mathbb{P}(\inf_{\mathbf{x}\in\mathcal{C}_1} \|\mathbf{W}\mathbf{x} - \mathbf{y}\|_2/\|\mathbf{x}\|_2 \leq \gamma\sqrt{n}, \mathcal{E}_K) \leq \exp\{-cn\ln(n\delta_n^{(1)})\}.$$

Procedure for two moments

- We start with $\delta_n^{(1)} \backsim \frac{1}{\ln n}, r_n^{(1)} \backsim n^{-Q_1}, Q_1 > 0$ and consider the set $\mathcal{C}_0 := \mathcal{C}(\delta_n^{(1)}, r_n^{(1)})$. We may show that

$$\mathbb{P}(\inf_{\mathbf{x}\in\mathcal{C}_0} \|\mathbf{W}\mathbf{x} - \mathbf{y}\|_2/\|\mathbf{x}\|_2 \leq \tau\sqrt{n}, \mathcal{E}_K) \leq \exp\{-cn\}.$$

- Now we take

$$\delta_n^{(2)} \backsim \frac{\ln(n\delta_n^{(1)})}{\ln n} \text{ and } r_n^{(2)} \backsim n^{-Q_2}, Q_2 > 0.$$

  and show that for $\mathcal{C}_1 = \mathcal{C}(\delta_n^{(2)}, r_n^{(2)}) \cap \mathcal{IC}(\delta_n^{(1)}, r_n^{(1)})$

$$\mathbb{P}(\inf_{\mathbf{x}\in\mathcal{C}_1} \|\mathbf{W}\mathbf{x} - \mathbf{y}\|_2/\|\mathbf{x}\|_2 \leq \gamma\sqrt{n}, \mathcal{E}_K) \leq \exp\{-cn\ln(n\delta_n^{(1)})\}.$$

- There exists an absolute constant $\delta_2 > 0$ such that $\delta_n^{(2)} \geq \delta_2$. This implies

$$\mathcal{IC}(\delta_n^{(2)}, r_n^{(2)}) \subset \mathcal{IC}(\delta_2, r_n^{(2)}) =: \mathcal{C}_2.$$

Bound via distance

- Let $\mathbf{X}_1, \dots, \mathbf{X}_n$ denote the columns of $\mathbf{W}$ and let $\mathcal{H}_k$ denotes the span of all column vectors except the kth. Then for every $\eta > 0$

$$\mathbb{P}(\inf_{\mathbf{x} \in \mathcal{C}_2} \|\mathbf{W}\mathbf{x}\|_2 \leq \eta \left(\frac{r_n^{(2)}}{\sqrt{n}} \right)^2, \mathcal{E}_K) \leq \frac{1}{n\delta_2} \sum_{k=1}^{n} \mathbb{P}(\mathrm{dist}(\mathbf{W}_k, \mathcal{H}_k) < \eta \frac{r_n^{(2)}}{\sqrt{n}}, \mathcal{E}_K).$$

- $\mathrm{dist}(\mathbf{W}_k, \mathcal{H}_k) \geq |(\mathbf{W}_k, \mathbf{h})|$, where $\mathbf{h}$ is unit vector orthogonal to $\mathcal{H}_k$

Bound for the distance

- For arbitrary matrix $\mathbf{A}$

$$\operatorname{dist}(\mathbf{A}_1, \mathcal{H}_1) \geq \frac{|(\mathbf{B}^{-T}\mathbf{v}, \mathbf{u}) - a_{11}|}{\sqrt{1 + \|\mathbf{B}^{-T}\mathbf{v}\|_2^2}},$$

where

$$\begin{pmatrix} a_{11} & \mathbf{v}^T \\ \mathbf{u} & \mathbf{B} \end{pmatrix} \tag{1}$$

- Set $\mathbf{A} := \mathbf{W}$. Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{B}$ are determined by (1). Then

$$\sup_{a \in \mathbb{R}} \Pr\left\{ \frac{|(\mathbf{B}^{-T}\mathbf{v}, \mathbf{u}) - a|}{\sqrt{1 + \|\mathbf{B}^{-T}\mathbf{v}\|^2}} \leq \varepsilon, \text{ and } \|\mathbf{B}\|_2 \leq K_n \right\} \leq Cn^{-A}$$

with $0 < \varepsilon \leq n^{-B}$ for some constants $A > 0$ and $B > 0$.

Quadratic form

- $$\mathbf{Q} = \begin{pmatrix} \mathbf{O}_{n-1} & \mathbf{B}^{-T} \\ \mathbf{B}^{-1} & \mathbf{O}_{n-1} \end{pmatrix} \quad \mathbf{w} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix},$$

  where $\mathbf{O}_{n-1}$ is $(n-1) \times (n-1)$ matrix with zero entries. Using the definition of $\mathbf{Q}$ we may write

  $$(\mathbf{B}^{-T}\mathbf{v}, \mathbf{u}) = \frac{1}{2}(\mathbf{Q}\mathbf{w}, \mathbf{w}).$$

Decoupling

- Introduce vector

$$\mathbf{w}' = \begin{pmatrix} \mathbf{u}' \\ \mathbf{v}' \end{pmatrix},$$

  where $\mathbf{u}', \mathbf{v}'$ are independent copies of $\mathbf{u}, \mathbf{v}$ respectively. We need the following statement.

- $$\sup_{v \in \mathbb{R}} \mathbb{P}_{\mathbf{w}} \left(|(\mathbf{Qw}, \mathbf{w}) - v| \leq 2\varepsilon \right) \leq \mathbb{P}_{\mathbf{w},\mathbf{w}'} \left(|(\mathbf{QP}_{J^c}(\mathbf{w} - \mathbf{w}'), \mathbf{P}_J \mathbf{w}) - u| \leq 2\varepsilon \right),$$

  where u doesn't depend on $\mathbf{P}_J \mathbf{w} = (\mathbf{P}_J \mathbf{u}, \mathbf{P}_J \mathbf{v})^T$.

- $$\varepsilon_0^{1/2} \sqrt{1 + ||\mathbf{B}^{-T} \mathbf{v}||_2^2} \leq ||\mathbf{B}^{-1}||_2 \leq \varepsilon_0^{-1} ||\mathbf{QP}_{\mathcal{J}^c}(\mathbf{w} - \mathbf{w}')||_2$$

Small ball probability

- $$\sup_{\substack{\mathbf{w}_0=(\mathbf{a},\mathbf{b})^T\in\mathcal{IC}(\delta,r_n^{(2)})\\ w\in\mathbb{R}}} \mathbb{P}_{\mathbf{P}_{\mathcal{J}}\mathbf{w}}\left(|(\mathbf{w}_0,\mathbf{P}_{\mathcal{J}}\mathbf{w})-w|\le\varepsilon_0^{-3/2}\varepsilon\right),$$

  where

  $$\mathbf{w}_0=\frac{1}{||\mathbf{Q}\mathbf{P}_{\mathcal{J}^c}(\mathbf{w}-\mathbf{w}')||_2}\begin{pmatrix}\mathbf{B}^{-T}\mathbf{P}_{\mathcal{J}^c}(\mathbf{v}-\mathbf{v}')\\ \mathbf{B}^{-1}\mathbf{P}_{\mathcal{J}^c}(\mathbf{u}-\mathbf{u}')\end{pmatrix}=\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\end{pmatrix}.$$

Small ball probability

- Let us fix a vector $\mathbf{w}_0$ and a number w. We can rewrite

$$(\mathbf{w}_0, P_{\mathcal{J}}\mathbf{w}) = \sum_{i \in \mathcal{J}} (a_i X_i + b_i Y_i),$$

where $||\mathbf{a}||_2^2 + ||\mathbf{b}||_2^2 = 1$.

Small ball probability

- Let us fix a vector $\mathbf{w}_0$ and a number w. We can rewrite

$$(\mathbf{w}_0, P_{\mathcal{J}}\mathbf{w}) = \sum_{i\in\mathcal{J}}(a_iX_i + b_iY_i),$$

  where $||\mathbf{a}||_2^2 + ||\mathbf{b}||_2^2 = 1$.
- There exists $\mathcal{J}_0 \in \mathcal{J}$ of cardinality $|\mathcal{J}_0| \backsim n$ such that

$$\frac{r_n^{(2)}}{\sqrt{2n}} \leq |a_i| \leq \frac{1}{\sqrt{\delta n}} \text{ for any } i \in \mathcal{J}_0,$$

Small ball probability

- Let us fix a vector $\mathbf{w}_0$ and a number w. We can rewrite

$$(\mathbf{w}_0, P_{\mathcal{J}}\mathbf{w}) = \sum_{i \in \mathcal{J}} (a_i X_i + b_i Y_i),$$

  where $||\mathbf{a}||_2^2 + ||\mathbf{b}||_2^2 = 1$.
- There exists $\mathcal{J}_0 \in \mathcal{J}$ of cardinality $|\mathcal{J}_0| \backsim n$ such that

$$\frac{r_n^{(2)}}{\sqrt{2n}} \leq |a_i| \leq \frac{1}{\sqrt{\delta n}} \text{ for any } i \in \mathcal{J}_0,$$

- Use the inequalities for Levy concentration function (see Petrov).

- Let $X_1, X_2, \ldots$ be independent random variables with

$$\mathbb{E}\, X_k = 0 \text{ and } \mathbb{E}\, X_k^2 = \sigma_k^2 > 0. \tag{2}$$

We denote $\sigma^2 = \sum_{k=1}^n \sigma_k^2$ and

$$Q(X, \lambda) = \sup_{a \in \mathbb{R}} \mathbb{P}(|X - a| \le \lambda).$$

- Assume that the condition (2) holds and let $S_n = \sum_{k=1}^n X_k$. Then

$$Q(S_n, \lambda) \le \frac{\sqrt{\lambda}}{(2\sigma^2 - 8 \sum_{i=1}^n \mathbb{E}\, X_j^2 \mathbb{I}(|X_j| \ge \lambda/2))^{1/2}}.$$

THANK YOU!