On Random Matrices Related to Quantum Informatics

L. Pastur

Mathematical Division, Institute for Low Temperatures,
Kharkiv, Ukraine

Roscoff, 25–29 June 2012

Outline

- Introduction
- Products of Triangular Matrices
    - Definition
    - Normalized Counting Measure
    - Comments
- Tensor Product Version of Sample Covariance Matrices
    - Definition
    - Normalized Counting Measure
    - Comments

Variations on the theme of "sample" (or "empirical") covariance matrices XX^T, where $X = \{X_{jk}\}_{j,k=1}^n$ are random square matrices. The subject is rather old with a lot of versions and motivations. A recent one is from

$$(Quantum\ Informatics) \cap (Statistical\ Mechnaics).$$

Key words: von Neumann entropy, quantum phase transitions, area law.

Let A be $n \times n$ real symmetric and B be $n \times n$ real anti-symmetric. Set

$$X = A + B,$$

assume a certain distribution for A and B, and study the Normalized Counting Measure (NCM)

$$N_n = n^{-1} \sum_{l=1}^{n} \delta_{\lambda_l^{(n)}}$$

of XX^T as $n \to \infty$, and also extreme eigenvalues, fluctuations of N_n, local statistics, eigenvectors, etc.

Recall that in the convention setting $X = n^{-1/2}Y$, where $\{Y_{jk}\}_{j,k=1}^{n}$ are independent standard Gaussian ($\mathbf{E}\{Y_{jk}\} = 0$, $\mathbf{E}\{Y_{jk}^2\} = 1$) and then N_n tends weakly with probability 1 to the quarter-circle law

$$\rho(\lambda) := N'(\lambda) = \frac{1}{4\pi}\sqrt{\frac{4-\lambda}{\lambda}}\mathbf{1}_{[0,4]}(\lambda)$$

in which $\lambda = 4$ $(\lambda = 0)$ is known as the standard *soft (hard) edge*. This is an old result of *Marchenko-P.*

Write

$$X = (X + X^T)/2 + (X - X^T)/2 := A + B$$

and obtain the simplest example of the above setting.

We need a bit more: the replacement $A \to A + yI$, i.e. $X \to X + yI$.

This is a particular case of results *Silverstein-Dozier 04,*

$y^2 < 1$: similar to quarter-circle law (standard soft and hard edges, the latter at 0);
$y^2 = 1$: upper edge is standard soft, lower edge is non standard hard

$$\rho(\lambda) = \mathrm{Const}\,\lambda^{-1/3}(1 + o(1)),\ \lambda \searrow 0;$$

$y^2 > 1$: both edges are strictly positive and standard soft.

Write in general setting

$$A = \frac{1}{2}A^+ + \frac{1}{2}(A^+)^T + A^0, \; B = \frac{1}{2}B^+ - \frac{1}{2}(B^+)^T$$

where A^+ and B^+ are lower triangular, and A^0 is diagonal.Choose $A^+ = B^+, \; A^0 = yI$ to get

$$A + B = A^+ + yI.$$

Assume that $\{A^+_{jk}\}_{n \geq j > k \geq 1}$ are independent standard Gaussian, $\mathbf{E}\{A^+_{jk}\} = 0, \; \mathbf{E}\{(A^+_{jk})^2\} = 1/n$ to obtain

$$M_n = (A^+ + yI)(A^+ + yI)^T.$$

Theorem

Let M_n be as above. Then its NCM converges weakly with probability 1 to the non-random limit N, whose Stieltjes transform f solves uniquely

$$\log(1+f) = \left(y^2 - z(1+f)\right)^{-1}, \ \Im f \cdot \Im z > 0, \ \Im z \neq 0.$$

$\operatorname{supp} N = [a\_(y), a_+(y)] \subset \mathbb{R}_+$, *$N$ is a. c. and if $\rho = N'$, then*
(i) $y \neq 0$: $a_-(y) = e^{-1} y^4 e^{1/y^2}$, $y \to 0$, $a_+(y) = e(1+y^2)$ $y \to 0$

$$\rho(\lambda) = \mathrm{Const}\, |a_\pm - \lambda|^{1/2}(1+o(1)), \ |a_\pm - \lambda| \to 0,$$

(ii) $y = 0$: $a_-(0) = 0$, $a_+(0) = e$ and

$$\rho(\lambda) = \begin{cases} \mathrm{Const}\,(e-\lambda)^{1/2}(1+o(1)), & \lambda \nearrow e, \\ (\lambda \log^2 \lambda)^{-1}(1+o(1)), & \lambda \searrow 0. \end{cases}$$

(i) f is not algebraic, cf *Anderson-Zeitouni 08, e.g.* Silverstein-Dozier case

$$f = \left(y^2(1+f)^{-1} - z(1+f)\right)^{-1}.$$

(ii) Most singular hard edge known. Recall the standard hard edge

$$\rho(\lambda) = \mathrm{Const}\,\lambda^{-1/2}(1+o(1)),\ \lambda \searrow 0,$$

of the quarter-circle law and more general Laguerre-type ensembles.
(iii) Implies an interesting quantum phase transition.
(iv) The rate of convergence of minimum eigenvalue of M_n, (à la Tomczak-Jaegerman et al), local statistics, etc?

A short(est) proof of the quarter-circle law for Gaussian vectors is as follows:

(i) Pass to the Stieltjes transform of N_n:

$$g_n(z) := \int \frac{N_n(d\lambda)}{\lambda - z} = n^{-1}\mathrm{Tr}\, G(z),\ G = (M - z)^{-1}$$

(ii) Use the Poincaré inequality to prove

$$\mathbf{Var}\{g_n(z)\} \leq \mathrm{Const}\ / n^2 |\,\mathrm{Im}\, z|^4$$

thereby reducing the problem to the convergence of $\mathbf{E}\{g_n(z)\}$.

(iii) Use the resolvent identity and the integration by parts to prove

$$f_n := \mathbf{E}\{g_n\} = -\frac{1}{z} + \frac{1}{z} f_n - \frac{1}{zn}\mathbf{E}\{g_n \mathrm{Tr}\, M_n G\}.$$

(iv) Use again the resolvent identity and (ii) – (iii) to obtain

$$zf_n^2 + zf_n + 1 = C(z)/n, \ C(z) < \infty, \ \Im z \neq 0.$$

(v) Pass to the limit $n \to \infty$, solve the limiting quadratic equation for $\mathrm{Im}\, f(z)\, \mathrm{Im}\, z > 0$ and recover N from the Stieltjes-Frobenuis inversion formula.

For matrices $Z = \{Z_{jk}^{(n)}\}_{j,k=1}^{n}$ with i.i.d. (but not necessarily Gaussian) entries use the "interpolation trick" (a two-term integration by parts) for

$$n^{-1/2}(\sqrt{1-t}X + \sqrt{t}Z).$$

More general versions

$$H + n^{-1}XTX^T, \text{ and } (A + n^{-1/2}X)T(A + n^{-1/2}X)^T$$

where X has independent entries and H, T and A are given can also be studied by similar approach.

L.Pastur, M. Shcherbina Eigenvalue Distribution of Large Random Matrices AMS. 2011.

Consider the technically simpler case $y = 0$. Use again the Stieltjes transform of N_n and the Poincaré

$$\mathbf{Var}\{g_n(z)\} \leq 1/n^2|\Im z|^4, \tag{1}$$

reducing the problem to the study of

$$f = \lim f_n, \ f_n := \mathbf{E}\{g_n\} = n^{-1}\sum_{j=1}^{n}\mathbf{E}\{G_{jj}\}, \ \Im z \neq 0.$$

The resolvent identity, the integration by parts and (1) imply:

$$\mathbf{E}\{G_{jj}\} \simeq -\frac{1}{z} + \frac{1}{z}\frac{j-1}{n}\mathbf{E}\{G_{jj}\} - \frac{1}{z}\mathbf{E}\{G_{jj}\} \sum_{\sim\sim k=1}^{j-1} \mathbf{E}\{n^{-1}\mathrm{Tr}(A^T GA)_{kk}\}$$

$$\mathbf{E}\{n^{-1}\mathrm{Tr}(A^T GA)_{jj}\} \simeq \frac{1}{n}\sum_{k=j}^{n}\mathbf{E}\{G_{kk}\} - \frac{1}{n}\sum_{k=j}^{n}\mathbf{E}\{G_{kk}\}\mathbf{E}\{n^{-1}\mathrm{Tr}(A^T GA)_{jj}\}.$$

View this as the finite-difference scheme for

$$f(t,z) = \lim_{n\to\infty, j/n\to t} \mathbf{E}\{G_{jj}\}.$$

Then the limit $j/n \to t \in [0,1]$ yields the equations

$$f(t,z) = -\left(z - \int_0^t h(s,z)ds\right)^{-1}, \; h(t,z) = \left(1 + \int_t^1 f(s,z)ds\right)^{-1},$$

and

$$f(z) = \int_0^1 f(t,z)dt.$$

Denote

$$\varphi(t,z) = \int_t^1 f(s,z)ds$$

to obtain

$$\frac{\partial^2}{\partial t^2}\varphi = \left(\frac{\partial}{\partial t}\varphi\right)^2 (1+\varphi)^{-1}, \; \frac{\partial}{\partial t}\varphi\Big|_{t=0} = z^{-1}, \; \varphi(0,z) = f(z),$$

thus

$$\varphi(t,z) = -1 + e^{-C(t-1)}, \; Ce^{-C} = -z^{-1}.$$

(i) Non-Gaussian case (can be treated similarly to the quarter-law case (interpolation trick etc.)).
(ii) "Scaling asymptotics" of ρ for $\lambda \sim y^2 \to 0$ to determine the type of corresponding quantum phase transition.
(iii) Two term asymptotics of the NCM ν_n of the spectral projection $\mathcal{E}_{M_n}(\Delta),\ \Delta \subset \sigma := \mathrm{supp}\ N$:

$$\begin{aligned}\nu_n(\Delta) &= (\alpha\delta_1 + (1-\alpha)\delta_0) \\ &\quad + n^{-1}\begin{cases} \mu_1(\Delta), & 0 \notin \mathrm{supp}\ N\ (gap,\ no\ pt), \\ \log(n)\ \mu_2(\Delta) & 0 \in \mathrm{supp}\ N\ (no\ gap,\ pt), \end{cases}\end{aligned}$$

where $\alpha = |\Delta|/|\mathrm{supp}\ N|$. This is an analog of the Strong Szegö theorem for Toeplitz operators. Involves the analysis of eigenvectors.

Tensor Product Version of Sample Covariance Matrices

Definition

Consider complex random i.i.d. vectors $\{\varphi_\alpha^j\}_{\alpha,j=1}^{p,k}$, $p = 1, 2...,$ k is fixed, and $\varphi_\alpha^j \in \mathbb{C}^d$ is

- either $d^{-1/2} X_\alpha^j$, and X_α^j is complex Gaussian vectors with i.i.d. standard components
- or uniformly distributed over the unit sphere.

Set

$$\Phi_\alpha = \varphi_\alpha^1 \otimes ... \otimes \varphi_\alpha^k$$

and consider the $d^k \times d^k$ random matrix

$$M_{p,d,k} = \sum_{\alpha=1}^{p} \Phi_\alpha \otimes \Phi_\alpha.$$

We are interested in the (non-random) limit as $p \to \infty$, $d \to \infty$, $p/d^k = p/n \to c \in (0, \infty)$ of the Normalized Counting Measure (NCM)

$$N_{p,d,k} = d^{-k} \sum_{l=1}^{d^k} \delta_{\lambda_l}, \ n = d^k.$$

It is also of interest the limits of the extreme eigenvalues, local statistics,

fluctuations of $N_{p,d,k}$, etc.

Studied by M. Hastings et al (CMP **310** (2012) 25-74) as a model of data hiding and correlation locking schema.

Proved the MP law for the limit N of the expectation of the NCM and the convergence of extreme eigenvalues to the endpoints of the support of N by fairly involved combinatorial analysis of moments $d^{-k}\mathrm{Tr}M_{p,d,k}^m, \ m \in \mathbb{N}$.

Remark. For Gaussian φ's $\Phi_\alpha \in (\mathbb{C}^d)^{\otimes k}$ has just dk independent parameters, while a generic $\Psi \in (\mathbb{C}^d)^{\otimes k}$ has d^k independent parameters. Nevertheless the MP law and the convergence of extreme eigenvalues hold in this case.

We show below that the MP law is valid for the limit with probability 1 of $N_{p,d,k}$ in the above and more general cases (vectors with independent but not necessarily Gaussian components as well as for vectors with log-concave distribution).

The approach used above for the quarter-circle law and its "triangular" analog does not apply to the tensor product version.

We will use an elaboration of the Marchenko-P. and Girko approach. Its version for $k = 1$ is given by Pajor-P. It is applicable to log-concave distributed φ_α's (not necessarily Gaussian) and any k.

(i) Observe that

$$M = \sum_{\alpha=1}^{p} L_\alpha, \; L_\alpha = (x, \varphi_\alpha)\varphi_\alpha.$$

(ii) Use either martingale differences (or Poincaré for Gaussian) to prove

$$\mathbf{Var}\{g_n(z)\} = o(1), \Im z \neq 0, \; n \to \infty, p \to \infty, p/n \in [0, \infty)$$

(iii) Use the resolvent identity to write

$$f_n = -z^{-1} + (zn)^{-1} \sum_{\alpha=1}^{p} (G\varphi_\alpha, \varphi_\alpha)$$

(iv) Use the rank one perturbation formulas:

$$G = G_\alpha - \frac{G_\alpha L_\alpha G_\alpha}{1 + (G_\alpha \varphi_\alpha, \varphi_\alpha)}, \quad G_\alpha = G|_{\varphi_\alpha = 0}$$

implying

$$(G\varphi_\alpha, \varphi_\alpha) = \frac{(G_\alpha \varphi_\alpha, \varphi_\alpha)}{1 + (G_\alpha \varphi_\alpha, \varphi_\alpha)}.$$

to rewrite (iii) as

$$f_n = -z^{-1} + (zn)^{-1} \sum_{\alpha=1}^{p} \frac{(G_\alpha \varphi_\alpha, \varphi_\alpha)}{1 + (G_\alpha \varphi_\alpha, \varphi_\alpha)}.$$

(v) Use the independence of G_α and φ_α and to obtain:

$$\mathbf{E}_\alpha\{(G_\alpha\varphi_\alpha,\varphi_\alpha)\} = n^{-1}\mathrm{Tr}G_\alpha,\ \mathbf{Var}\{(G_\alpha\varphi_\alpha,\varphi_\alpha)\} \le \mathrm{Const}/n|\Im z|^2.$$

(iv) Use (ii) and (v) to replace $(G_\alpha\varphi_\alpha,\varphi_\alpha)$ in (iv) by its expectation $f_{\alpha n} := \mathbf{E}\{n^{-1}\mathrm{Tr}G_\alpha\}$.
(v) Use the rank one perturbation formula of (iv) to find that $f_{\alpha n} = f_n + O(1/n)$ and get the "pre"– limiting quadratic equation

$$f_n = -\frac{1}{z} + \frac{c}{z}\frac{f_n}{1+f_n} + o(1),\ \Im z \neq 0,\ c = p/n$$

equivalent to the above.

For any $n \times n$ matrix A we need random vectors $\varphi \in \mathbb{C}^n$ possessing
(i) isotropy

$$\mathbf{E}\{\,(A\varphi,\varphi)\} = n^{-1}\mathrm{Tr}\,A;$$

(ii) vanishing of fluctuations of $(A\varphi,\varphi)$ ("good" vectors)

$$\mathbf{Var}\{\,(A\varphi,\varphi)\} = ||A||\delta_n,\ \delta_n = O(1),\ n \to \infty.$$

Lemma

Let $\varphi \in \mathbb{C}^d$ be a random vector as above and A is $d^k \times d^k$ matrix. If φ^1,...φ^k ate k independent copies of φ then the random vector $\Phi = \varphi^1 \otimes ... \otimes \varphi^k$ also possesses the above properties in which $n = d^k$ and δ_n is replaced by $C_k\delta_d$, where C_k depends only on k.

Proof is based on the martingale-differences.

Study the extreme eigenvalues, both for $c > 1$ (both edges are standard soft) and $c = 1$ (lower edge is standard soft). Have likely differerent rates of convergence (depending on k).

Example: for Gaussian vectors

$$\mathbf{Var}\{g_n\} \leq \frac{C(z)k}{n^{1+1/k}}, \ 0 < C(z) < \infty, \ \mathrm{Im}\, z \neq 0,$$

thus, different scaling of fluctuations of linear eigenvalue statistics (CLT), etc.