Limiting spectral distribution of the eigenvalues of a Gram matrix with a given variance profile

Jamal Najim

(joint work with Walid Hachem and Philippe Loubaton)

Ecole Nationale Supérieure des Télécommunications and CNRS
najim@tsi.enst.fr

Outline

1. The model
2. Motivations
3. Main result
4. Elements of proof
5. Hand waving

The matrix model

We introduce the matrices Y_n, Λ_n and Σ_n.

$$Y_n = \begin{pmatrix} Y_{11} & \dots & Y_{1n} \\ \vdots & \ddots & \vdots \\ Y_{N1} & \dots & Y_{Nn} \end{pmatrix}, \quad \lim_{n\to\infty} \frac{N}{n} = c\,(< 1)$$

with $Y_{ij} = \frac{\sigma(i/N, j/n)}{\sqrt{n}} X_{ij}$ and the X_{ij}'s being i.i.d. centered.

$$\Lambda_n = \begin{pmatrix} \Lambda_{11} & 0 & & & \dots & 0 \\ 0 & \Lambda_{22} & 0 & & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ 0 & \dots & 0 & \Lambda_{NN} & 0 \dots & 0 \end{pmatrix}$$

$$\Sigma_n = Y_n + \Lambda_n$$

Aim and assumptions

Aim

We study the limit of the spectral distribution of $\Sigma_n \Sigma_n^T$ and $\Sigma_n^T \Sigma_n$, that is the limit of:

$$D_n = \frac{1}{N} \sum_{i=1}^{N} \delta_{\mu_i} \quad \text{where} \quad (\mu_1, \cdots, \mu_N) = \text{eigval}(\Sigma_n \Sigma_n^T),$$

$$\tilde{D}_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{\tilde{\mu}_i} \quad \text{where} \quad (\tilde{\mu}_1, \cdots, \tilde{\mu}_n) = \text{eigval}(\Sigma_n^T \Sigma_n),$$

Assumptions

1. The variance profile $\sigma : [0,1]^2 \to (0, \infty)$ is a continuous function.
2. $\frac{1}{N} \sum_{i=1}^{N} \delta_{(i/n, \Lambda_{ii}^2)} \to H(dt d\lambda)$ where H has a compact support.
3. X_{ij}'s moments: $\exists \epsilon > 0, \quad \mathbb{E}|X_{ij}|^{4+\epsilon} < \infty$.

Wireless communication

The main motivation lies in the performance analysis of **Multiple-Input Multiple-Output** (MIMO) digital communication systems:

$$n \text{ transmit antennas} \left\{ \begin{array}{ccc} < & & > \\ \vdots & & \vdots \\ < & \text{channel} & > \\ \vdots & & \vdots \\ < & & > \end{array} \right\} N \text{ receive antennas}$$

Σ_{ij} represents the gain between transmit antenna i and receive antenna j.

The knowledge of the spectral distribution of $\Sigma_n \Sigma_n^T$ is fundamental to compute performance indexes such as

- the CAPACITY of the channel,
- the SINR (Signal-to-Interference-plus-Noise Ratio).

Wireless communication (2)

If there are reflectors between the transmit and receive antenna, The gain Σ_{ij} is no longer centered and the matrix $\Sigma_n = Y_n + \Lambda_n$ is a simplified model for channels with reflectors.

More ressources on Wireless applications of Random matrices on the MALCOM project web page:

`http://www.tsi.enst.fr/~najim/aci-malcom/`

References

There exists a huge amount of references. Here is a subective sample.

On Gram matrices

- Marčenko and Pastur ('67)
- Boutet de Monvel, Khorunzhy, Vasilchuk ('96): study of $\frac{1}{n}G_nG_n^T$.
- Girko ('01)
- Brent Dozier and Silverstein ('04): study of $\left(\frac{X_n}{\sqrt{n}} + R_n\right)\left(\frac{X_n}{\sqrt{n}} + R_n\right)^T$.

On wireless communication and random matrices

- Tse et al. (Hanly, Evans, Zeitouni '99-'00)
- Telatar ('99)
- Verdu et al. (Shamai, Tulino, etc. '99-'00-'04)

Convergence of the spectral distribution

Theorem

The spectral distributions D_n and $\tilde{D}_n$ of $\Sigma_n \Sigma_n^T$ and $\Sigma_n^T \Sigma_n$ converge toward deterministic probability distributions D and $\tilde{D}$:

$$D_n \quad \xrightarrow[n\to\infty]{\text{distribution}} \quad D \quad (a.s.)$$

$$\tilde{D}_n \quad \xrightarrow[n\to\infty]{\text{distribution}} \quad \tilde{D} \quad (a.s.)$$

where D and $\tilde{D}$ are characterized by their Stieltjes transform $\mathbf{f}(z)$ and $\tilde{\mathbf{f}}(z)$:

$$\mathbf{f}(z) = \int_0^\infty \frac{D(dx)}{x - z} \quad \text{and} \quad \tilde{\mathbf{f}}(z) = \int_0^\infty \frac{\tilde{D}(dx)}{x - z} \quad \text{for} \quad Im(z) > 0.$$

Characterization of D and $\tilde{D}$ via their Stieltjes Transform

Consider the following system: for every continuous and bounded function g,

$$(S)\quad \begin{cases} \int g\, d\pi_z = \int \frac{g(u,\lambda)}{-z(1+\int \sigma^2(u,\cdot)d\tilde{\pi}_z)+\frac{\lambda}{1+c\int \sigma^2(\cdot,cu)d\pi_z}} H(du,d\lambda) \\ \int g\, d\tilde{\pi}_z = c\int \frac{g(cu,\lambda)H(du,d\lambda)}{-z(1+c\int \sigma^2(\cdot,cu)d\pi_z)+\frac{\lambda}{1+\int \sigma^2(u,\cdot)d\tilde{\pi}_z}} + (1-c)\int_c^1 \frac{g(u,0)\, du}{-z(1+c\int \sigma^2(\cdot,u)d\pi_z)} \end{cases}$$

This system has a unique pair of solutions $(\pi, \tilde{\pi})$ among a certain class of solutions (Stieltjes kernels). The Stieltjes transforms $\mathbf{f}$ and $\tilde{\mathbf{f}}$ are then characterized by:

$$\mathbf{f}(z) = \int \pi(z, dt, d\lambda) \quad \text{and} \quad \tilde{\mathbf{f}}(z) = \int \tilde{\pi}(z, dt, d\lambda).$$

Two important special cases

The centered case

If $\Lambda_n \equiv 0$, then the system (S) becomes

$$\int g d\pi_z = \int_{[0,1]} \frac{g(u)}{-z + \int_0^1 \frac{\sigma^2(u,t)}{1 + c\int_0^1 \sigma^2(x,t)\pi_z(dx)} dt} du.$$

Same equation as in the Gaussian field case (Boutet de Monvel et al.)

The i.i.d. case

If $\sigma(x,y) \equiv \sigma$ is a constant, then Assumption 2 becomes: $\frac{1}{N}\sum_{i=1}^N \delta_{\Lambda_{ii}^2} \to H_\Lambda(d\lambda)$ and one can directly obtain an equation involving the Stieltjes transform:

$$f(z) = \int \frac{H_\Lambda(d\lambda)}{-z(1 + c\sigma^2 f(z)) + (1-c)\sigma^2 + \frac{\lambda}{1 + c\sigma^2 f(z)}}.$$

This equation is the same as in Brent Dozier and Silverstein.

The empirical measures L_n and $\tilde{L}_n$

Let us introduce the resolvents

$$Q(z) = (\Sigma_n \Sigma_n^T - zI)^{-1} = (q_{ij}(z))_{i,j \le N} \quad \text{and} \quad \tilde{Q}(z) = (\Sigma_n^T \Sigma_n - zI)^{-1} = (\tilde{q}_{ij}(z))_{i,j \le n}$$

Then $\frac{1}{n}\mathrm{Trace} Q(z)$ is the Stieltjes transform of D_n. We deal instead with the empirical measures

$$\begin{aligned} L_n(z, dx\, d\lambda) &= \frac{1}{N} \sum_{i=1}^{N} q_{ii}(z) \delta_{\left(\frac{i}{N}, \Lambda_{ii}^2\right)}(dx\, d\lambda) \\ \tilde{L}_n(z, dx\, d\lambda) &= \frac{1}{n} \sum_{i=1}^{N} \tilde{q}_{ii}(z) \delta_{\left(\frac{i}{n}, \Lambda_{ii}^2\right)}(dx\, d\lambda) + \frac{1}{n} \sum_{i=N+1}^{n} \tilde{q}_{ii}(z) \delta_{\frac{i}{n}}(dx) \otimes \delta_0(d\lambda) \end{aligned}$$

Links between L_n and $\tilde{L}_n$

Somehow standard matrix manipulations on $q_{ii}(z)$ yield

$$\begin{aligned}\int g dL_n &= \frac{1}{N}\sum_{i=1}^{N} g(i/n, \Lambda_{ii}^2) q_{ii}(z) \\ &\approx \frac{1}{N}\sum_{i=1}^{N} \frac{g(i/N, \Lambda_{ii}^2)}{-z - z\int \sigma^2(i/N, \cdot) d\tilde{L}_n + \frac{\Lambda_{ii}^2}{1+\frac{N}{n}\int \sigma^2(\cdot, i/n) dL_n}}\end{aligned}$$

In the previous computation, a tiny miracle occurs based on the fact that Λ_n is diagonal. Moreover one can guess the first equation of the system (S):

$$\int g\, d\pi_z = \int \frac{g(u,\lambda)}{-z(1+\int \sigma^2(u,\cdot) d\tilde{\pi}_z) + \frac{\lambda}{1+c\int \sigma^2(\cdot, cu) d\pi_z}} H(du, d\lambda)$$

End of proof: a compacity argument

- From every subsequence of $(L_n, \tilde{L}_n)$, there exists a converging subsequence:

$$L_{\phi(n)} \xrightarrow[n\to\infty]{\text{weak}} \pi_\phi \quad \text{and} \quad \tilde{L}_{\phi(n)} \xrightarrow[n\to\infty]{\text{weak}} \tilde{\pi}_\phi.$$

  where π_ϕ and $\tilde{\pi}_\phi$ are Stieltjes kernels.

- Due to the links between L_n and $\tilde{L}_n$, $(\pi_\phi, \tilde{\pi}_\phi)$ is a solution of (S) and the unicity of the solutions yield:

$$(\pi_\phi, \tilde{\pi}_\phi) = (\pi, \tilde{\pi}).$$

- Necessarily, the whole sequence converges:

$$L_n \xrightarrow[n\to\infty]{\text{weak}} \pi \quad \text{and} \quad \tilde{L}_n \xrightarrow[n\to\infty]{\text{weak}} \tilde{\pi}.$$

Forthcoming work and open issues

The non-centered gaussian field case

Consider the matrix $\Gamma_n = G_n + T_n$ where T_n is a Toeplitz matrix and G_n is a matrix with gaussian correlated entries:

$$\mathrm{cov}(G_{ij}, G_{i',j'}) = \gamma(i-i', j-j'), \quad \text{with} \sum_{(k,l)\in\mathbb{Z}^2} |\gamma(k,l)| < \infty.$$

- Prove that the limiting spectral distribution of $\Gamma_n \Gamma_n^T$ is characterized by the same equations as in the current work (as already known in the centered case).

The general non-centered case with variance profile

Consider the matrix $\Sigma_n = Y_n + A_n$ where Y_n is as previously and A_n is deterministic, not necessarily (pseudo)-diagonal. This model is important in Wireless communication

- Study the spectral distribution of $\Sigma_n \Sigma_n^T$ (which might not converge) in terms of deterministic equivalents, as in Girko's work for instance.