A CLT for Information-Theoretic Statistics of Gram Random Matrices

Malika Kharouf
Joint work with W.Hachem, J.Najim and J.Silverstein

October 12, 2010

The Model: A Non-Centered Random Matrices

Consider a $p \times n$ random matrices:

$$\Sigma_n = \frac{1}{\sqrt{n}} X_n + A_n,$$

where,

- $X_{n_{ij}}$, $1 \leq i \leq p$, $1 \leq j \leq n$ are i.i.d. centered with unit variance and $\mathbb{E}|X_{11}|^{16} < \infty$.
- A_n is a $p \times n$ deterministic matrix with uniformly bounded spectral norm.

The Model: Information-Theoretic Statistics of Gram random matrices

Linear spectral statistics:

$$\mathcal{I}_n(\rho) = \frac{1}{p} \sum_{i=1}^{p} \log \left(\lambda_i^{(n)} + \rho \right) ,$$

where, $\lambda_i^{(n)}$, $i = 1, \dots, p$ are the eigenvalues of the Gram random matrix $\Sigma_n \Sigma_n^*$ and ρ is a nonnegative parameter.

Objective: Understanding the asymptotic distribution of the fluctuations of $\mathcal{I}_n(\rho)$, when the dimensions of the matrix Σ_n converge to infinity at the same pace and obtain a simple form of the variance.

Plan

Motivations:

Mutual Information for Multiple Antenna Radio Channels

Multi-user MIMO scheme

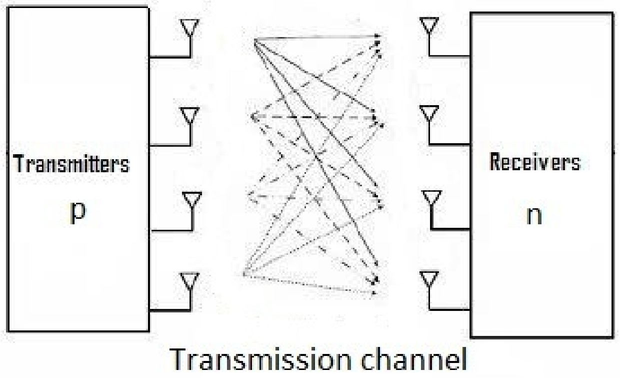

Figure: MIMO Systems

MIMO System: Mathematical Model

The p-dimensional receiver vector $\mathbf{r}_n$ is given by:

$$\mathbf{r}_n = \Sigma_n \mathbf{t}_n + \mathbf{b}_n,$$

where,

- Σ_n represents the channel matrix which assumed to be random.
- $\mathbf{t}_n$ is the n-dimensional transmitter vector.
- $\mathbf{b}_n$ is an additive white Gaussian noise with covariance matrix $\mathbb{E}\mathbf{b}_n\mathbf{b}_n^* = \rho I_p$.

Performance indicator: The Mutual Information:

$$\mathcal{I}_n(\rho) = \frac{1}{p} \log\det\left(\Sigma_n\Sigma_n^* + \rho I_p\right) = \frac{1}{p}\sum_{i=1}^{p} \log\left(\lambda_i + \rho\right)$$

Asymptotic behavior of $\mathcal{I}_n(\rho)$ when $n, p \to \infty$ at the same rate ?

Asymptotic behavior of $\mathcal{I}_n(\rho)$: First-order results

First-order results

Let f_n denotes the ST of $\mu_{\Sigma_n \Sigma_n^*}$, the spectral measure of the eigenvalues of $\Sigma_n \Sigma_n^*$. Then,

$$\mathcal{I}_n(\rho) = -\int_{\rho}^{\infty} f_n(-\omega) d\omega.$$

Then the asymptotic behavior of $\mathcal{I}_n(\rho)$ is closely linked to the asymptotic behavior of f_n as $p, n \to \infty$ with the same pace.

State of the art

- $F^{A_n A_n^*} \to H$, H is a deterministic probability measure. Dozier and Silverstein (04):

$$F^{\Sigma_n \Sigma_n^*} \xrightarrow{\text{weakly}} F,$$

where, F is a deterministic probability measure which the Stieltjes transform is a unique solution of a given coupled equation.

State of the art

- $F^{A_n A_n^*} \to H$, H is a deterministic probability measure. Dozier and Silverstein (04):

$$F^{\Sigma_n \Sigma_n^*} \xrightarrow{\text{weakly}} F,$$

  where, F is a deterministic probability measure which the Stieltjes transform is a unique solution of a given coupled equation.
- V. L. Girko (91), Hachem-Loubaton-Najim (07) : Look for a deterministic approximation of the Stieltjes transform f_n of $F^{\Sigma_n \Sigma_n^*}$. $\exists$ a $p \times p$ deterministic valued function $T_n(\rho)$ such that:

$$f_n(-\rho) - \frac{1}{p} \mathrm{Tr}\, T_n(-\rho) \xrightarrow[n \to \infty]{a.s} 0$$

Fundamental equations

Theorem (Girko '91, Hachem-Loubaton-Najim '07)

The following system of two equations

$$\begin{cases} \delta_n(\rho) = \frac{1}{n}\mathrm{Tr}\left(\rho\left(1+\tilde{\delta}_n(\rho)\right) I_p + \frac{A_n A_n^*}{1+\delta_n(\rho)}\right)^{-1} \triangleq \frac{1}{n}\mathrm{Tr}\, T_n(\rho) \\ \tilde{\delta}_n(\rho) = \frac{1}{n}\mathrm{Tr}\left(\rho\left(1+\delta_n(\rho)\right) I_n + \frac{A_n^* A_n}{1+\tilde{\delta}_n(\rho)}\right)^{-1} \triangleq \frac{1}{n}\mathrm{Tr}\, \tilde{T}_n(\rho), \end{cases}$$

admits a unique solution $(\delta_n, \tilde{\delta}_n)$ *in* $\mathcal{S}(\mathbb{R}^+)^2$. *Moreover,*

$$\int_{\mathbb{R}^+} f(\lambda) dF^{\Sigma_n \Sigma_n^*}(\lambda) - \int_{\mathbb{R}^+} f(\lambda)\pi_n(d\lambda) \xrightarrow[n\to\infty]{a.s} 0, \quad \forall f \in \mathcal{C}_B(\mathbb{R}^+),$$

where π_n *is the positive measure where* δ_n *is the Stieltjes transform.*

First order result: Deterministic equivalents

Theorem (Hachem-Loubaton-Najim '07)

Let $V_n(\rho) = \int_{\mathbb{R}^+} \log(\lambda + \rho)\pi_n(d\lambda)$. *Then we have:*

$$\mathbb{E}\mathcal{I}_n(\rho) - V_n(\rho) \xrightarrow[n,p\to\infty,\frac{p}{n}\to c>0]{} 0.$$

Moreover, $V_n(\rho)$ *admits a closed-form expression*

$$V_n(\rho) = \frac{1}{p}\sum_{i=1}^{p} \log\left(\rho\left(1+\tilde{\delta}_n\right) + \frac{\mu_{n,i}^2}{1+\delta_n}\right) + \frac{n}{p}\log\left(1+\delta_n\right) - \frac{\rho n}{p}\delta_n\tilde{\delta}_n,$$

where $\mu_{n,i}$ *are the singular values of the mean matrix* A_n.

In the non-centered case, the first-order asymptotic study of the mutual information depends mainly on the limiting behavior of the singular values of the mean matrix A_n.

Study of the fluctuations

CLT for $p\left(\mathcal{I}_n(\rho) - V_n(\rho)\right)$

In order to study the CLT for $p\left(\mathcal{I}_n(\rho) - V_n(\rho)\right)$ we study separately two quantities:

- The random quantity $p\left(\mathcal{I}_n(\rho) - \mathbb{E}\mathcal{I}_n(\rho)\right)$ from which the fluctuations arise and,
- The deterministic quantity $p\left(\mathbb{E}\mathcal{I}_n(\rho) - V_n(\rho)\right)$ which yields a bias.

Asymptotic distribution of the fluctuations: Definition of the variance

Theorem (Hachem-Kharouf-Najim-Silverstein '10)

Let $\vartheta = \mathbb{E}X_{11}^2$, $\kappa = \mathbb{E}|X_{11}|^4 - 2 - \vartheta^2$ *and let*
$\gamma = \frac{1}{n}\mathrm{Tr}\, T^2$, $\tilde{\gamma} = \frac{1}{n}\mathrm{Tr}\, \tilde{T}^2$, $\underline{\gamma} = \frac{1}{n}\mathrm{Tr}\, T\,\bar{T}$, $\underline{\tilde{\gamma}} = \frac{1}{n}\mathrm{Tr}\, \tilde{T}\,\bar{\tilde{T}}$. *Denote by*

$$\begin{aligned}\Theta_n^2 \quad = \quad & -\log\left(\left(1 - \frac{1}{n\left(1+\tilde{\delta}\right)}\mathrm{Tr}\, TAA^*T\right)^2 - \rho^2\gamma\tilde{\gamma}\right) \\ & -\log\left(\left|1 - \vartheta\frac{1}{n\left(1+\tilde{\delta}\right)}\mathrm{Tr}\, \bar{T}\bar{A}A^*T\right|^2 - |\vartheta|^2\rho^2\underline{\gamma\tilde{\gamma}}\right) \\ & +\kappa\frac{\rho^2}{n^2}\sum_i t_{ii}^2\sum_j \tilde{t}_{jj}^2\end{aligned}$$

Then Θ_n^2 *is well defined.*

Some remarks

- The variance is the sum of tree terms: the first term would be the same in the Gaussian case.
- The variance depends on the singular values of the main matrix as well as on its singular vectors.
- In the circular case ($X_{ij} \stackrel{\mathcal{D}}{=} X_{ij}e^{i\alpha}$ for all α), the second term disappears.

Asymptotic distribution of the fluctuations: The CLT

Theorem (Hachem-Kharouf-Najim-Silverstein '10)

The following convergence holds true:

$$\frac{p}{\Theta_n}\left(\mathcal{I}_n(\rho) - \mathbb{E}\mathcal{I}_n(\rho)\right) \xrightarrow[p,n\to\infty]{\mathcal{D}} \mathcal{N}(0,1),$$

where $\mathcal{D}$ stands for convergence in distribution.

Proof of the CLT: The approach

REFORM (**RE**solvent **FOR**mula and **M**artingale).

- $\mathcal{I}_n(\rho) - \mathbb{E}\mathcal{I}_n(\rho)$ as a sum of increments of martingale.
- Identification of the variance.

CLT for martingales

Theorem

Let $\Gamma_1^{(n)}, \ldots, \Gamma_n^{(n)}$ be a sequence of increments of martingale with respect to a given filtration $\mathcal{F}_1^{(n)}, \ldots, \mathcal{F}_n^{(n)}$. Assume that there exists a sequence of nonnegative real numbers $(\Theta_n^2)_n$ uniformly bounded away from zero and from infinity. Assume that:

- $$\sum_{j=1}^{n} \mathbb{E}\left(\Gamma_j^{(n)^2} | \mathcal{F}_{j-1}^{(n)}\right) - \Theta_n^2 \xrightarrow[n\to\infty]{\mathcal{P}} 0.$$
- *The Lyapunov's condition*

$$\exists \alpha > 0, \quad \frac{1}{\Theta_n^{2(1+\alpha)}} \sum_{j=1}^{n} \mathbb{E}|\Gamma_j^{(n)}|^{2+\alpha} \xrightarrow[n\to\infty]{} 0, \quad \textit{holds.}$$

Then $\Theta_n^{-1} \sum_{j=1}^{n} \Gamma_j^{(n)}$ converges in distribution to $\mathcal{N}(0,1)$.

Sum of martingale differences

We have,

$$\mathcal{I}_n - \mathbb{E}\mathcal{I}_n = \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j-1}) \left(-\log(1+\xi_j)\right) \triangleq \sum_{j=1}^{n} \Gamma_j,$$

where,

$$\xi_j = \frac{\eta_j^* Q_j \eta_j - \left(\frac{1}{n}\mathrm{Tr}\, Q_j + a_j^* Q_j a_j\right)}{1 + \frac{1}{n}\mathrm{Tr}\, Q_j + a_j^* Q_j a_j}.$$

with η_j, a_j are resp. the jth columns of matrices Σ_n and A_n, Q_j is the resolvent of the matrix $\Sigma_j \Sigma_j^*$ and $\mathbb{E}_j$ stands for the conditional expectation with respect to the σ-algebra $\mathcal{F}_j^{(n)} = \sigma(x_1, \ldots, x_j)$.

Sum of the conditional variances

Some properties of the function log,

$$\sum_{j=1}^{n} \mathbb{E}_{j-1}\left((\mathbb{E}_j - \mathbb{E}_{j-1})\log(1+\xi_j)\right)^2 - \sum_{j=1}^{n} \mathbb{E}_{j-1}\left(\mathbb{E}_j \xi_j\right)^2 \xrightarrow[p,n\to\infty]{\mathcal{P}} 0$$

where (recall)

$$\xi_j = \frac{\eta_j^* Q_j \eta_j - \left(\frac{1}{n}\mathrm{Tr}\, Q_j + a_j^* Q_j a_j\right)}{1 + \frac{1}{n}\mathrm{Tr}\, Q_j + a_j^* Q_j a_j}.$$

Study of the sum of conditional variances

Standard calculations remain the problem to the study of the asymptotic behavior of the quantities:

$$\frac{1}{n}\mathrm{Tr}\,(\mathbb{E}_j Q_n)^2 \quad \text{and} \quad a_j^*(\mathbb{E}_j Q_n)^2 a_j,$$

where Q_n is the resolvent of $\Sigma_n \Sigma_n^*$ matrix.

Outline of the proof

A good comprehension of the asymptotic behavior of these terms requires a specific study of bilinear forms of type $u_n^* Q(\rho) v_n$ where at least u_n or v_n is a given column of the deterministic mean matrix A_n.

If u_n and v_n are deterministics, Hachem-Loubaton-Najim-Vallet (preprint'10)

$$u_n^* Q(\rho) v_n \approx u_n^* T(\rho) v_n$$

Asymptotic behavior of the bias:

Theorem (Hachem-Kharouf-Najim-Silverstein '10)

We have,

$$p\left(\mathbb{E}\mathcal{I}_n(\rho) - V_n(\rho)\right) - \mathcal{B}_n(\rho) \xrightarrow[p,n\to\infty]{} 0$$

where,

$$\mathcal{B}_n(\rho) = \kappa\, Cte(\rho, \delta, \tilde{\delta})$$

$\kappa = \mathbb{E}|X_{11}|^4 - 2 - \vartheta^2.$

Outline of the proof of the bias term

The bias term is given by

$$\begin{aligned}
\chi_n(\rho) &= p\left(\mathbb{E}\mathcal{I}_n(\rho) - V_n(\rho)\right) \\
&= p\int_\rho^\infty \frac{d}{d\omega}\mathbb{E}\log\det\left(\Sigma_n\Sigma_n^* + \omega I_p\right)d\omega \\
&\qquad -p\int_\rho^\infty \frac{d}{d\omega}\left(\int_{\mathbb{R}^+}\log\left(\lambda+\omega\right)\pi_n(d\lambda)\right)d\omega \\
&= \int_\rho^\infty \mathrm{Tr}\,\left(\mathbb{E}Q_n(\omega) - T_n(\omega)\right)d\omega.
\end{aligned}$$

Then it remains to study the asymptotic behavior of $\mathrm{Tr}\,\left(\mathbb{E}Q_n(\omega) - T_n(\omega)\right)$. We prove,

$$\mathrm{Tr}\,\left(\mathbb{E}Q_n(\omega) - T_n(\omega)\right) - \kappa Cte(\rho,\delta,\tilde{\delta}) \xrightarrow[n\to\infty]{} 0$$

Case of a non-centered separable random matrix model

The non-centered separable case

$$\Sigma_n = \frac{1}{\sqrt{n}} D_n^{1/2} X_n \tilde{D}_n^{1/2} + A_n,$$

where, $D_n^{1/2}$ and $\tilde{D}_n^{1/2}$ are resp. $p \times p$ and $n \times n$ deterministic diagonal matrices with nonnegative entries.

First-order asymptotic behavior

$$V_n(\rho) = \frac{1}{p} \log\det T_n^{-1}(\rho) + \frac{1}{p} \log\left(I_n + \delta_n \tilde{D}_n\right) - \frac{\rho n}{p} \delta_n \tilde{\delta}_n,$$

where, $\delta_n = \frac{1}{n}\mathrm{Tr}\; T(\rho)$ and $\tilde{\delta}_n(\rho) = \frac{1}{n}\mathrm{Tr}\; \tilde{T}(\rho)$, with

$$\begin{aligned} T_n(\rho) &= \left(\rho\left(I_p + \tilde{\delta}_n D_n\right) + A_n \left(I_n + \delta_n \tilde{D}_n\right)^{-1} A_n^*\right)^{-1} \\ \tilde{T}_n(\rho) &= \left(\rho\left(I_n + \delta_n \tilde{D}_n\right) + A_n^* \left(I_p + \tilde{\delta}_n D_n\right)^{-1} A_n\right)^{-1} \end{aligned}$$

The non-centered separable case

The variance:

$$\begin{aligned}
\Theta_n^2 &= -\log\left(\Omega_n(\rho) - \rho^2\gamma\tilde{\gamma}\right) - \log\left(\bar{\Omega}_n(\rho) - |\vartheta|^2\rho^2\underline{\gamma\tilde{\gamma}}\right) \\
&\quad + \kappa\frac{\rho^2}{n^2}\sum_i d_i^2 t_{ii}^2 \sum_j \tilde{d}_j^2 \tilde{t}_{jj}^2
\end{aligned}$$

where:

$$\Omega_n(\rho) = \left(1 - \frac{1}{n}\mathrm{Tr} D_n^{1/2} T_n A_n \left(I_n + \delta\tilde{D}_n\right)^{-1} \tilde{D}_n \left(I_n + \delta\tilde{D}_n\right)^{-1} A_n^* T_n D_n^{1/2}\right)$$

and

$$\bar{\Omega}_n(\rho) = \left|1 - \vartheta\frac{1}{n}\mathrm{Tr} D_n^{1/2} \bar{T}_n \bar{A}_n \left(I_n + \delta\tilde{D}_n\right)^{-1} \tilde{D}_n \left(I_n + \delta\tilde{D}_n\right)^{-1} A_n^* T_n D_n^{1/2}\right|$$

Thank you !