Deux distributions limite de matrices de Gram avec des entrées non i.i.d.

Jamal Najim (ENST et CNRS)

(en collaboration avec Walid Hachem et Philippe Loubaton)

Exposé au GdT “Processus Stochastiques, Matrices Aléatoires” - Laboratoire PMA

Janvier 2006

Matrices de Gram

Matrices du type $\Sigma_n \Sigma_n^T$ où

$$\Sigma_n \quad \text{matrice} \quad N \times n, \quad \Sigma_n = Y_n + \Lambda_n, \quad \mathbb{E}(\Sigma_n) = 0.$$

Régime asymptotique:

$$n \to \infty, \quad \frac{N}{n} \to c \in (0, \infty)$$

Plan de l'exposé

1. Le cas non centré diagonal à profil de variance
   - Y_n à entrées indépendantes (profil de variance),
   - Λ_n pseudo-diagonal
2. Le cas du champ stationnaire gaussien
   - Y_n issu d'un champ stationnaire gaussien.

Le modèle matriciel

Introduisons les matrices Y_n, Λ_n and Σ_n.

$$Y_n = \begin{pmatrix} Y_{11} & \dots & Y_{1n} \\ \vdots & \ddots & \vdots \\ Y_{N1} & \dots & Y_{Nn} \end{pmatrix}, \quad \text{avec} \quad Y_{ij} = \frac{\sigma(i/N, j/n)}{\sqrt{n}} X_{ij} \quad \text{et} \lim_{n\to\infty} \frac{N}{n} = c \in (0,\infty)$$

les X_{ij} étant i.i.d. centrés.

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

Objectif et hypothèses

Objectif

On étudie la limite des distributions spectrales de $\Sigma_n \Sigma_n^T$ et $\Sigma_n^T \Sigma_n$, i.e.:

$$\begin{aligned} D_n &= \frac{1}{N} \sum_{i=1}^{N} \delta_{\mu_i} \quad \text{où} \quad (\mu_1, \cdots, \mu_N) = \mathrm{eigval}(\Sigma_n \Sigma_n^T), \\ \tilde{D}_n &= \frac{1}{n} \sum_{i=1}^{n} \delta_{\tilde{\mu}_i} \quad \text{où} \quad (\tilde{\mu}_1, \cdots, \tilde{\mu}_n) = \mathrm{eigval}(\Sigma_n^T \Sigma_n), \end{aligned}$$

Hypothèses

1. Le profil de variance $\sigma : [0,1]^2 \to (0, \infty)$ est une fonction continue.
2. $\frac{1}{N} \sum_{i=1}^{N} \delta_{(i/n, \Lambda_{ii}^2)} \to H(dt d\lambda)$ où H a un support compact.
3. Hypothèse sur les moments de X_{ij}: il existe $\epsilon > 0$ tel que $\mathbb{E}|X_{ij}|^{4+\epsilon} < \infty$.

Motivations: communication sans fil, systèmes MIMO

La motivation principale réside dans l'étude d'indicateurs de performance de systèmes MIMO (**Multiple-Input Multiple-Output**).

$$n \text{ antennes à l'émission} \left\{ \begin{array}{ccc} < & & > \\ \vdots & & \vdots \\ < & \text{canal} & > \\ \vdots & & \vdots \\ < & & > \end{array} \right\} N \text{ antennes à la réception}$$

Σ_{ij} représente le gain entre l'antenne à l'émission i et l'antenne à la réception j.

La connaissance de la distribution spectrale de $\Sigma_n \Sigma_n^*$ est fondamentale pour calculer des indicateurs de performance tels que l'INFORMATION MUTUELLE du canal:

$$I_n(\sigma^2) = \frac{1}{N} \mathbb{E} \log \det \left(I + \frac{\Sigma \Sigma^*}{\sigma^2} \right)$$

Motivations (suite)

S'il y a des réflecteurs entre les antennes de transmission et les antennes de réception, le gain Σ_{ij} n' est plus centré. La matrice

$$\Sigma_n = Y_n + \Lambda_n$$

est un cas d'école pour lequel on peut étudier le comportement asymptotique de la mesure spectrale.

Outil $\#$ 1: la transformée de Stieltjes

Soit $\mathbb{P} \in \mathcal{P}(\mathbb{R}^+)$ alors sa transformée de Stieltjes est définie par:

$$\mathbf{f}(z) = \int_0^\infty \frac{\mathbb{P}(d\lambda)}{\lambda - z}, \quad z \in \mathbb{C} - \mathbb{R}^+.$$

Connaissant $\mathbf{f}$, on peut reconstruire $\mathbb{P}$:

$$\mathbb{P}[a,b] = \lim_{\eta \to 0^+} \frac{1}{\pi} \int_a^b \mathrm{Im}\mathbf{f}(\xi + \mathbf{i}\eta) d\xi,$$

où a et b sont des points de continuité de $\mathbb{P}$.

Transformée de Stieltjes et Matrices aléatoires

Le lien entre la TS et les MA se fait par la résolvante:

$$Q_n(z) = (\Sigma_n \Sigma_n^T - zI)^{-1} = (q_{ij}(z))_{1 \leq i,j \leq N}.$$

Alors

$$\mathbf{f}_n(z) = \frac{1}{N} \mathrm{Trace}\, Q_n(z) = \frac{1}{N} \sum_{i=1}^{N} q_{ii}(z) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\lambda_i - z}$$

où les λ_i sont les valeurs propres de $\Sigma_n \Sigma_n^T$.

La fonction $\mathbf{f}_n(z)$ est la Transformée de Stieltjes de la mesure empirique des valeurs propres de $\Sigma_n \Sigma_n^T$.

Résultat important: Si $\mathbf{f}_n(z) \to \mathbf{f}(z)$ qui est elle-même une TS, alors $F^{\Sigma_n \Sigma_n^T} \to F$.

Outil # 2: Comportement asymptotique de formes quadratiques

Soit A une matrice $n \times n$ telle que

$$\sup_n \|A\| \leq K$$

Soit

$$\mathbf{y} = \left(\frac{X_1}{\sqrt{n}}, \cdots, \frac{X_n}{\sqrt{n}} \right)$$

un vecteur tel que les X_i sont i.i.d avec $\mathbb{E}X_1 = 0,\ \mathbb{E}X_1^2 = 1$. Si $\mathbf{y}$ et A sont indépendents alors on a:

$$\mathbf{y} A \mathbf{y}^T - \frac{1}{n} \mathrm{Trace}\, A \xrightarrow[n \to \infty]{} 0$$

presque sûrement.

Deux références proches

Le cas centré (Girko ’90, Boutet de Monvel et al. ’96)

Si $\Lambda_n \equiv 0$, alors l’équation (en π) admet une solution unique dans une certaine classe de noyaux π (noyaux de Stieltjes)

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

et la tranformée de Stieltjes de la mesure spectrale limite est donnée par

$$\mathbf{f}(z) = \int_{[0,1]} \pi_z(dx)$$

Autre formulation: $\pi_z(dx) = k(x,z)dx$

$$k(x,z) = \frac{1}{-z + \int_0^1 \frac{\sigma^2(x,t)}{1+c\int_0^1 \sigma^2(u,t)k(u,z)du} dt}$$

Le cas i.i.d. (Brent Dozier et Silverstein '04)

Si $\sigma(x,y) \equiv \sigma$ est une constante, alors considérons

$$\Sigma_n = Y_n + A_n \quad \text{où} \quad F^{A_n A_n^T} \xrightarrow[n\to\infty]{} H(d\lambda)$$

auquel cas la transformée de Stieltjes de la mesure spectrale limite satisfait l'équation:

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

mais aussi le cas Wigner gaussien ... (Shlyakhtenko '96)

Notion de liberté avec amalgamation

Retour au cas non centré diagonal

Théorème (convergence de la mesure spectrale)

Soit

$$\Sigma_n = Y_n + \Delta_n,$$

alors, les mesures spectrales D_n et $\tilde{D}_n$ de $\Sigma_n \Sigma_n^T$ et $\Sigma_n^T \Sigma_n$ convergent vers des mesures de probabilité déterministes D et $\tilde{D}$:

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

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

où D et $\tilde{D}$ sont caractérisées par leur transformée de Stieltjes $\mathbf{f}(z)$ and $\tilde{\mathbf{f}}(z)$:

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

Caractérisation des transformées de Stieltjes f et $\tilde{\mathbf{f}}$

Considérons le système suivant: pour toute fonction continue bornée g,

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

Ce système a une paire de solutions unique $(\pi, \tilde{\pi})$ dans une certaine classe de solutions (noyaux de Stieltjes). Les transformées de Stieltjes $\mathbf{f}$ et $\tilde{\mathbf{f}}$ sont alors caractérisées par:

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

Quelques éléments de preuve: les mesures empiriques L_n et $\tilde{L}_n$

On introduit les résolvantes

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

Alors

$$\frac{1}{n}\mathrm{Trace} Q_n(z) = \frac{1}{N}\sum_{i=1}^{N} q_{ii}(z)$$

est la transformée de Stieltjes de $D_n = \frac{1}{N}\sum_{i=1}^{N} \delta_{\mu_i}$.

On préfère travailler avec les mesures empiriques suivantes, plus flexibles:

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

Liens entre L_n et $\tilde{L}_n$

Des manipulations matricielles standard sur $q_{ii}(z)$ entraînent

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

Dans le calcul précédent, un petit miracle se produit, basé sur le fait que Λ_n est diagonale. On peut, à partir de là deviner la première equations de (S).

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

On conclut par un argument d'unicité et un de compacité.

Conclusion partielle

Résultat complexe

Malgré la simplicité de la matrice Δ_n de centrage

Une hypothèse difficilement généralisable à un cas non diagonal

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{(i/n,\Lambda_{ii}^2)} \to H(dtd\lambda) \quad \text{où} \quad \Lambda_n = \mathrm{pseudo}-\mathrm{diag}(\Lambda_{ii}, 1 \le i \le N)$$

Une approche possible: les équivalents déterministes

cf. Exposé de P. Loubaton.

1. On considère le modèle $\Sigma_n = Y_n + A_n$
2. A_n générale, lignes et colonnes bornées en norme L^2
3. On cherche un équivalent déterministe à la transformée de Stieltjes.

Modèle matriciel

On considère une matrice Z_n de dimensions $N \times n$, dont les entrées sont données par:

$$Z^n_{j_1 j_2} = \frac{1}{\sqrt{n}} \sum_{(k_1,k_2)\in\mathbb{Z}^2} h(k_1,k_2) U(j_1 - k_1, j_2 - k_2),$$

où

1. $h \in \ell^1(\mathbb{Z}^2)$ i.e.

$$\sum_{(k_1,k_2)\in\mathbb{Z}^2} |h(k_1,k_2)| < \infty,$$

2. $U(i,j)$ famille gaussienne complexe i.i.d. (0,1):

$$U(i,j) = X + \mathbf{i}X', \quad X, X' \sim \mathcal{N}\left(0, \frac{1}{\sqrt{2}}\right).$$

avec X et X' indépendants.

Corrélation entre les entrées

La covariance entre $Z^n_{j_1 j_2}$ et $Z^n_{j'_1 j'_2}$ est donnée par:

$$\begin{aligned}\operatorname{cov}(Z^n_{j_1 j_2}, Z^n_{j'_1 j'_2}) &= \frac{1}{n} \sum_{(k_1,k_2)\in\mathbb{Z}^2} h(k_1,k_2) h^*(k_1-(j_1-j'_1), k_2-(j_2-j'_2)), \\ &= \frac{\kappa(j_1-j'_1, j_2-j'_2)}{n}.\end{aligned}$$

Objectif

1. Etudier la mesure spectrale de $Z_n Z_n^*$ (limite?),
2. Etudier un équivalent déterministe pour $(Z_n+B_n)(Z_n+B_n)^*$,

Une remarque: Mesure spectrale de $Z_n Z_n^*$

(Girko '90, Boutet de Monvel, Khorunzhiy '96) La mesure spectrale limite est déterminée par sa transformée de Stieltjes $\mathbf{f}$, obtenue

$$\mathbf{f}(z) = \int_{[0,1]} \pi_z(dx)$$

où π_z satisfait la même équation que dans le cas à profil de variance.

$$\forall g \in C([0,1]), \quad \int g \, d\pi_z = \int_0^1 \frac{g(u)}{-z + \int_0^1 \frac{|\Phi|^2(u,t)}{1+c\int_0^1 |\Phi|^2(x,t)\,\pi_z(dx)} dt} du.$$

avec (ce qui joue le rôle du) profil de variance:

$$\Phi(t_1, t_2) = \sum_{(l_1,l_2)\in\mathbb{Z}^2} h(l_1, l_2) e^{2\pi\mathbf{i}(l_1 t_1 - l_2 t_2)}$$

On affine l'objectif:

Comment déduire TOUS les résultats pour des matrices de Gram issues de champs stationnaires gaussiens À PARTIR des résultats sur les matrices à profil de variance?

Deux étapes

1. Couplage par périodisation de la matrice initiale

Considérons $\tilde{Z}_n$ où

$$\tilde{Z}^n_{j_1 j_2} = \frac{1}{\sqrt{n}} \sum_{(k_1,k_2)\in\mathbb{Z}^2} h(k_1,k_2) U(j_1 - k_1 \bmod N, j_2 - k_2 \bmod n),$$

alors $\tilde{Z}_n$ est une bonne approximation de Z_n au sens de l'inégalité de Bai:

$$\begin{aligned} & \mathcal{L}^4(F^{(Z_n+B_n)(Z_n+B_n)^*}, F^{(\tilde{Z}_n+B_n)(\tilde{Z}_n+B_n)^*}) \\ \leq \quad & \frac{2}{N^2} \mathrm{Tr}(Z_n - \tilde{Z}_n)(Z_n - \tilde{Z}_n)^* \times \ldots \\ \xrightarrow{P} \quad & 0. \end{aligned}$$

où $\mathcal{L}$ distance de Lévy entre probas.

Éléments de preuve -suite-

Similarité avec une matrice à entrée indépendantes

Considérons la matrice de Fourier $p \times p$:

$$F^p_{j_1,j_2} = \frac{1}{\sqrt{p}} \exp 2\mathbf{i}\pi \left(\frac{j_1 j_2}{p} \right).$$

Alors les entrées de $Y_n = F_N \tilde{Z}_n F_n^*$ ont la forme:

$$Y^n_{l_1 l_2} = \frac{1}{\sqrt{n}} \Phi \left(\frac{l_1}{N}, \frac{l_2}{n} \right) X^n_{l_1 l_2}$$

où Φ est définie précédemment et les $X^n_{l_1 l_2}$ sont des gaussiennes i.i.d complexes.

Conclusion

Tous les résultats obtenus dans le cas de matrices à profil de variance se “transportent” au cas de matrices stationnaires gaussiennes.