Matrices Aléatoires

Marwa Banna

Doctorante à l'Université Paris-Est Marne-La-Vallée sous la direction de Florence Merlevède et Emmanuel Rio

Colloque Jeunes Probabilistes et Statisticiens

Avril 2014

Introduction et Motivation

- Soit $\mathbf{X}_1, \ldots, \mathbf{X}_n \in \mathbb{R}^N$ des vecteurs i.i.d centrés et de matrice de covariance $\Sigma = \mathbb{E}(\mathbf{X}_1 \mathbf{X}_1^T) = \ldots = \mathbb{E}(\mathbf{X}_n \mathbf{X}_n^T)$.

Introduction et Motivation

- Soit $\mathbf{X}_1, \ldots, \mathbf{X}_n \in \mathbb{R}^N$ des vecteurs i.i.d centrés et de matrice de covariance $\Sigma = \mathbb{E}(\mathbf{X}_1 \mathbf{X}_1^T) = \ldots = \mathbb{E}(\mathbf{X}_n \mathbf{X}_n^T)$.
- La matrice de covariance empirique est définie par

$$\frac{1}{n} \sum_{k=1}^{n} \mathbf{X}_k \mathbf{X}_k^T = \frac{1}{n} \mathcal{X}_n \mathcal{X}_n^T \quad \text{où} \quad \mathcal{X}_n = (\mathbf{X}_1 \ldots \mathbf{X}_n) \in \mathcal{M}_{N \times n}(\mathbb{R}).$$

Introduction et Motivation

- Soit $\mathbf{X}_1, \ldots, \mathbf{X}_n \in \mathbb{R}^N$ des vecteurs i.i.d centrés et de matrice de covariance $\Sigma = \mathbb{E}(\mathbf{X}_1\mathbf{X}_1^T) = \ldots = \mathbb{E}(\mathbf{X}_n\mathbf{X}_n^T)$.
- La matrice de covariance empirique est définie par

$$\frac{1}{n}\sum_{k=1}^{n}\mathbf{X}_k\mathbf{X}_k^T = \frac{1}{n}\mathcal{X}_n\mathcal{X}_n^T \quad \text{où} \quad \mathcal{X}_n = (\mathbf{X}_1 \ldots \mathbf{X}_n) \in \mathcal{M}_{N\times n}(\mathbb{R}).$$

- $\mathbb{E}(\frac{1}{n}\sum_{k=1}^{n}\mathbf{X}_k\mathbf{X}_k^T) = \Sigma$.
- Pour N fixé, la loi des grands nombres implique que

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{n}\mathbf{X}_k\mathbf{X}_k^T = \Sigma \quad \text{p.s.}$$

Introduction et Motivation

- Soit $\mathbf{X}_1, \ldots, \mathbf{X}_n \in \mathbb{R}^N$ des vecteurs i.i.d centrés et de matrice de covariance $\Sigma = \mathbb{E}(\mathbf{X}_1 \mathbf{X}_1^T) = \ldots = \mathbb{E}(\mathbf{X}_n \mathbf{X}_n^T)$.
- La matrice de covariance empirique est définie par

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{X}_k \mathbf{X}_k^T = \frac{1}{n}\mathcal{X}_n \mathcal{X}_n^T \quad \text{où} \quad \mathcal{X}_n = (\mathbf{X}_1 \ldots \mathbf{X}_n) \in \mathcal{M}_{N\times n}(\mathbb{R}).$$

- $\mathbb{E}(\frac{1}{n}\sum_{k=1}^{n} \mathbf{X}_k \mathbf{X}_k^T) = \Sigma$.
- Pour N fixé, la loi des grands nombres implique que

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbf{X}_k \mathbf{X}_k^T = \Sigma \quad \text{p.s.}$$

- Que se passe-t-il lorsque $N := N_n \to \infty$ quand $n \to \infty$?

Un brin d'histoire

- Wishart (1920-1930)
  - La naissance de la statistique mathématique

Un brin d'histoire

- Wishart (1920-1930)
  - La naissance de la statistique mathématique
- Wigner, Mehta, et Dyson (1950-1960)
  - Physique nucléaire

Un brin d'histoire

- Wishart (1920-1930)
  - La naissance de la statistique mathématique
- Wigner, Mehta, et Dyson (1950-1960)
  - Physique nucléaire
- Marčenko, Pastur, Girko, Bai, et Silverstein (1960-1980)

Théorème de Marčenko-Pastur

- Soit $(X_{ij})_{i,j\geqslant 1}$ une famille de variables aléatoires réelles i.i.d tel que $\mathbb{E}(X_{11}) = 0$ et $\mathrm{Var}(X_{11}) = 1$.

Théorème de Marčenko-Pastur

- Soit $(X_{ij})_{i,j\geqslant 1}$ une famille de variables aléatoires réelles i.i.d tel que $\mathbb{E}(X_{11}) = 0$ et $\mathrm{Var}(X_{11}) = 1$.
- Soit $\mathbf{B}_n = \frac{1}{n}\sum_{k=1}^{n} \mathbf{X}_k \mathbf{X}_k^T = \frac{1}{n}\mathcal{X}_n\mathcal{X}_n^T$ où

$$\mathcal{X}_n = (\mathbf{X}_1 \dots \mathbf{X}_n) = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ \vdots & \vdots & & \vdots \\ X_{N1} & X_{N2} & \dots & X_{Nn} \end{pmatrix}.$$

Théorème de Marčenko-Pastur

- Soit $(X_{ij})_{i,j\geqslant 1}$ une famille de variables aléatoires réelles i.i.d tel que $\mathbb{E}(X_{11}) = 0$ et $\mathrm{Var}(X_{11}) = 1$.
- Soit $\mathbf{B}_n = \frac{1}{n}\sum_{k=1}^{n} \mathbf{X}_k \mathbf{X}_k^T = \frac{1}{n}\mathcal{X}_n \mathcal{X}_n^T$ où

$$\mathcal{X}_n = (\mathbf{X}_1 \ldots \mathbf{X}_n) = \begin{pmatrix} X_{11} & X_{12} & \ldots & X_{1n} \\ \vdots & \vdots & & \vdots \\ X_{N1} & X_{N2} & \ldots & X_{Nn} \end{pmatrix}.$$

- La mesure spectrale empirique de $\mathbf{B}_n$ sont respectivement définies par

$$\mu_{\mathbf{B}_n} = \frac{1}{N}\sum_{k=1}^{N} \delta_{\lambda_k}$$

où $\lambda_1, \ldots, \lambda_N$ sont les valeurs propres de $\mathbf{B}_n$.

Théorème de Marčenko-Pastur

- Soit $(X_{ij})_{i,j\geqslant 1}$ une famille de variables aléatoires réelles i.i.d tel que $\mathbb{E}(X_{11}) = 0$ et $\mathrm{Var}(X_{11}) = 1$.
- Soit $\mathbf{B}_n = \frac{1}{n}\sum_{k=1}^{n} \mathbf{X}_k \mathbf{X}_k^T = \frac{1}{n}\mathcal{X}_n\mathcal{X}_n^T$ où

$$\mathcal{X}_n = (\mathbf{X}_1 \dots \mathbf{X}_n) = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ \vdots & \vdots & & \vdots \\ X_{N1} & X_{N2} & \dots & X_{Nn} \end{pmatrix}.$$

- La mesure spectrale empirique de $\mathbf{B}_n$ sont respectivement définies par

$$\mu_{\mathbf{B}_n} = \frac{1}{N}\sum_{k=1}^{N} \delta_{\lambda_k}$$

  où $\lambda_1, \dots, \lambda_N$ sont les valeurs propres de $\mathbf{B}_n$.
- On suppose que $c_n := \frac{N}{n} \xrightarrow[n\to+\infty]{} c \in (0, \infty)$.

Théorème de Marčenko-Pastur

Theorem

Presque sûrement, pour toute fonction continue et bornée $f : \mathbb{R} \to \mathbb{R}$,

$$\int f \, d\mu_{\mathbf{B}_n} \xrightarrow[n \to +\infty]{} \int f \, d\mu_c$$

où μ_c est la loi de Marčenko-Pastur

$$\left(1 - \frac{1}{c}\right)_+ \delta_0 + \frac{1}{c 2\pi x} \sqrt{(b-x)(x-a)} \mathbf{1}_{[a,b]}(x) dx$$

avec $.+ := \max(0, .)$ *et* $a = (1 - \sqrt{c})^2$ *et* $b = (1 + \sqrt{c})^2$.

La Transformée de Stieltjes

La transformée de Stieltjes $S_\nu : \mathbb{C}_+ \to \mathbb{C}$ d'une loi ν sur $\mathbb{R}$ est définie par

$$S_\nu(z) := \int \frac{1}{x - z} \, d\nu(x) \, .$$

La Transformée de Stieltjes

La transformée de Stieltjes $S_\nu : \mathbb{C}_+ \to \mathbb{C}$ d'une loi ν sur $\mathbb{R}$ est définie par

$$S_\nu(z) := \int \frac{1}{x - z} \, d\nu(x) \, .$$

- $|S_\nu(z)| \leqslant 1/\mathfrak{Im}(z)$ et $\mathfrak{Im}(S_\nu(z)) \geqslant 0$.

La Transformée de Stieltjes

La transformée de Stieltjes $S_\nu : \mathbb{C}_+ \to \mathbb{C}$ d'une loi ν sur $\mathbb{R}$ est définie par

$$S_\nu(z) := \int \frac{1}{x - z} \, d\nu(x) \, .$$

- $|S_\nu(z)| \leqslant 1/\mathfrak{Im}(z)$ et $\mathfrak{Im}(S_\nu(z)) \geqslant 0$.
- La fonction S_ν est analytique sur $\mathbb{C}_+$ et caractérise ν.

$$\nu([a, b]) = \lim_{y \downarrow 0} \frac{1}{\pi} \int_a^b \mathcal{I}m S_\nu(x + iy) \, dx \, .$$

La Transformée de Stieltjes

La transformée de Stieltjes $S_\nu : \mathbb{C}_+ \to \mathbb{C}$ d'une loi ν sur $\mathbb{R}$ est définie par

$$S_\nu(z) := \int \frac{1}{x - z} \, d\nu(x) \, .$$

- $|S_\nu(z)| \leqslant 1/\mathfrak{Im}(z)$ et $\mathfrak{Im}(S_\nu(z)) \geqslant 0$.
- La fonction S_ν est analytique sur $\mathbb{C}_+$ et caractérise ν.

$$\nu([a, b]) = \lim_{y \downarrow 0} \frac{1}{\pi} \int_a^b \mathcal{I}m S_\nu(x + iy) \, dx \, .$$

- Pour une suite de mesures $(\nu_n)_n$, on a

$$\left(\nu_n \xrightarrow[n\to\infty]{\mathcal{L}} \nu \right) \Leftrightarrow \left(\forall z \in \mathbb{C}_+, \; S_{\nu_n}(z) \xrightarrow[n\to\infty]{} S_\nu(z) \right).$$

Preuve par Transformée de Stieltjes

- La transformée de Stieltjes S de la loi de M-P est

$$S(z) = \frac{1}{2cz}\left((1-y) - z + \sqrt{(z-1-c)^2 - 4c}\right).$$

Preuve par Transformée de Stieltjes

- La transformée de Stieltjes S de la loi de M-P est

$$S(z) = \frac{1}{2cz}\left((1-y) - z + \sqrt{(z-1-c)^2 - 4c}\right).$$

- **But:** Montrer pour tout $z \in \mathbb{C}^+$, $\lim_{n\to\infty} S_{\mu_{\mathbf{B}_n}}(z) = S(z)$.

Preuve par Transformée de Stieltjes

- La transformée de Stieltjes S de la loi de M-P est

$$S(z) = \frac{1}{2cz}\left((1-y) - z + \sqrt{(z-1-c)^2 - 4c}\right).$$

- **But:** Montrer pour tout $z \in \mathbb{C}^+$, $\lim_{n\to\infty} S_{\mu_{\mathbf{B}_n}}(z) = S(z)$.
    1. $S_{\mu_{\mathbf{B}_n}}(z) - \mathbb{E}\big(S_{\mu_{\mathbf{B}_n}}(z)\big) \to 0$ presque sûrement

Preuve par Transformée de Stieltjes

- La transformée de Stieltjes S de la loi de M-P est

$$S(z) = \frac{1}{2cz}\left((1-y) - z + \sqrt{(z-1-c)^2 - 4c}\right).$$

- **But:** Montrer pour tout $z \in \mathbb{C}^+$, $\lim_{n\to\infty} S_{\mu_{\mathbf{B}_n}}(z) = S(z)$.
  1. $S_{\mu_{\mathbf{B}_n}}(z) - \mathbb{E}\big(S_{\mu_{\mathbf{B}_n}}(z)\big) \to 0$ presque sûrement
  2. $\mathbb{E}\big(S_{\mu_{\mathbf{B}_n}}(z)\big) - \mathbb{E}\big(S_{\mu_{\mathbf{G}_n}}(z)\big) \to 0$ où

$$\mathbf{G}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{Y}_i \mathbf{Y}_i^T \quad \text{et} \quad \mathbf{Y}_i = (Y_{1i}, \ldots, Y_{Ni})^T,$$

$(Y_{i,j})_{i,j\geqslant 1}$ étant une suite i.i.d. Gaussienne centrée réduite indépendante de $(X_{i,j})_{i,k\geqslant 1}$.

Preuve par Transformée de Stieltjes

- La transformée de Stieltjes S de la loi de M-P est

$$S(z) = \frac{1}{2cz}\left((1-y) - z + \sqrt{(z-1-c)^2 - 4c}\right).$$

- **But:** Montrer pour tout $z \in \mathbb{C}^+$, $\lim_{n\to\infty} S_{\mu_{\mathbf{B}_n}}(z) = S(z)$.
  1. $S_{\mu_{\mathbf{B}_n}}(z) - \mathbb{E}\big(S_{\mu_{\mathbf{B}_n}}(z)\big) \to 0$ presque sûrement
  2. $\mathbb{E}\big(S_{\mu_{\mathbf{B}_n}}(z)\big) - \mathbb{E}\big(S_{\mu_{\mathbf{G}_n}}(z)\big) \to 0$ où

$$\mathbf{G}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{Y}_i \mathbf{Y}_i^T \quad \text{et} \quad \mathbf{Y}_i = (Y_{1i}, \ldots, Y_{Ni})^T,$$

     $(Y_{i,j})_{i,j\geqslant 1}$ étant une suite i.i.d. Gaussienne centrée réduite indépendante de $(X_{i,j})_{i,k\geqslant 1}$.
  3. $\mathbb{E}\big(S_{\mu_{\mathbf{G}_n}}(z)\big) - S(z) \to 0$ presque sûrement.

La Transformée de Stieltjes

$$S_{\mu_{\mathbf{B}_n}}(z) = \int \frac{1}{x-z}\, d\mu_{\mathbf{B}_n}(x) = \frac{1}{N}\sum_{k=1}^{N} \frac{1}{\lambda_k - z} \\ = \frac{1}{N}\mathrm{Tr}\,(\mathbf{B}_n - z\mathbf{I})^{-1} = f(\mathbf{X}_1, \ldots, \mathbf{X}_n). \tag{1}$$

$$S_{\mu_{\mathbf{G}_n}}(z) = \frac{1}{N}\mathrm{Tr}\,(\mathbf{G}_n - z\mathbf{I})^{-1} = f(\mathbf{Y}_1, \ldots, \mathbf{Y}_n). \tag{2}$$

La Transformée de Stieltjes

$$S_{\mu_{\mathbf{B}_n}}(z) = \int \frac{1}{x-z}\, d\mu_{\mathbf{B}_n}(x) = \frac{1}{N}\sum_{k=1}^{N} \frac{1}{\lambda_k - z} = \frac{1}{N}\mathrm{Tr}\,(\mathbf{B}_n - z\mathbf{I})^{-1} = f(\mathbf{X}_1, \ldots, \mathbf{X}_n). \tag{1}$$

$$S_{\mu_{\mathbf{G}_n}}(z) = \frac{1}{N}\mathrm{Tr}\,(\mathbf{G}_n - z\mathbf{I})^{-1} = f(\mathbf{Y}_1, \ldots, \mathbf{Y}_n). \tag{2}$$

Montrer que

$$\mathbb{E}f(\mathbf{X}_1, \ldots, \mathbf{X}_n) - \mathbb{E}f(\mathbf{Y}_1, \ldots, \mathbf{Y}_n) \to 0$$

La Méthode de Lindeberg (Chatterjee 2006)

- Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction trois fois dérivable.

La Méthode de Lindeberg (Chatterjee 2006)

- Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction trois fois dérivable.
- Soient $(X_k)_{k\in\mathbb{Z}}$ et $(Y_k)_{k\in\mathbb{Z}}$ deux suites indépendantes de variables aléatoires indépendantes bornées centrées réduites. On note les vecteurs aléatoires

$$\mathbf{X} = (X_1, \dots, X_n) \quad \text{et} \quad \mathbf{Y} = (Y_1, \dots, Y_n).$$

La Méthode de Lindeberg (Chatterjee 2006)

- Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction trois fois dérivable.
- Soient $(X_k)_{k\in\mathbb{Z}}$ et $(Y_k)_{k\in\mathbb{Z}}$ deux suites indépendantes de variables aléatoires indépendantes bornées centrées réduites. On note les vecteurs aléatoires

$$\mathbf{X} = (X_1, \dots, X_n) \quad \text{et} \quad \mathbf{Y} = (Y_1, \dots, Y_n).$$

Comment peut-on approximer $\mathbb{E}(f(\mathbf{X}))$ par $\mathbb{E}(f(\mathbf{Y}))$?

La Méthode de Lindeberg (Chatterjee 2006)

- Avec la notation $\mathbf{Z}_i = (X_1, \ldots, X_i, Y_{i+1}, \ldots, Y_n)$, on a

$$\mathbb{E}f(\mathbf{X}) - \mathbb{E}f(\mathbf{Y}) = \sum_{i=1}^{n} \left(\mathbb{E}f(\mathbf{Z}_i) - \mathbb{E}f(\mathbf{Z}_{i-1})\right)$$

La Méthode de Lindeberg (Chatterjee 2006)

- Avec la notation $\mathbf{Z}_i = (X_1, \ldots, X_i, Y_{i+1}, \ldots, Y_n)$, on a

$$\mathbb{E}f(\mathbf{X}) - \mathbb{E}f(\mathbf{Y}) = \sum_{i=1}^{n} \big(\mathbb{E}f(\mathbf{Z}_i) - \mathbb{E}f(\mathbf{Z}_{i-1})\big)$$

- On note $\mathbf{Z}_i^0 = (X_1, \ldots, X_{i-1}, 0, Y_{i+1}, \ldots, Y_n)$. Grâce à la formule de Taylor, on a

$$\begin{aligned} f(\mathbf{Z}_i) - f(\mathbf{Z}_{i-1}) &= (X_i - Y_i)\partial_i f(\mathbf{Z}_i^0) + (X_i^2 - Y_i^2)\partial_i^2 f(\mathbf{Z}_i^0) \\ &\quad + \frac{1}{6} X_i^3 \partial_i^3 f(\mathbf{Z}_i^*) + \frac{1}{6} Y_i^3 \partial_i^3 f(\mathbf{Z}_i^{**}) \end{aligned}$$

La Méthode de Lindeberg (Chatterjee 2006)

- Avec la notation $\mathbf{Z}_i = (X_1, \ldots, X_i, Y_{i+1}, \ldots, Y_n)$, on a

$$\mathbb{E}f(\mathbf{X}) - \mathbb{E}f(\mathbf{Y}) = \sum_{i=1}^{n} \left(\mathbb{E}f(\mathbf{Z}_i) - \mathbb{E}f(\mathbf{Z}_{i-1})\right)$$

- On note $\mathbf{Z}_i^0 = (X_1, \ldots, X_{i-1}, 0, Y_{i+1}, \ldots, Y_n)$. Grâce à la formule de Taylor, on a

$$\left|\mathbb{E}f(\mathbf{Z}_i) - \mathbb{E}f(\mathbf{Z}_{i-1})\right| \leqslant \frac{1}{6}\|X_i\|_\infty^3 \mathbb{E}|\partial_i^3 f(\mathbf{Z}_i^*)| + \frac{1}{6}\|Y_i\|_\infty^3 \mathbb{E}|\partial_i^3 f(\mathbf{Z}_i^{**})|$$

La Méthode de Lindeberg (Chatterjee 2006)

- Avec la notation $\mathbf{Z}_i = (X_1, \ldots, X_i, Y_{i+1}, \ldots, Y_n)$, on a

$$\mathbb{E} f(\mathbf{X}) - \mathbb{E} f(\mathbf{Y}) = \sum_{i=1}^{n} \left(\mathbb{E} f(\mathbf{Z}_i) - \mathbb{E} f(\mathbf{Z}_{i-1}) \right)$$

- On note $\mathbf{Z}_i^0 = (X_1, \ldots, X_{i-1}, 0, Y_{i+1}, \ldots, Y_n)$. Grâce à la formule de Taylor, on a

$$\left| \mathbb{E} f(\mathbf{Z}_i) - \mathbb{E} f(\mathbf{Z}_{i-1}) \right| \leqslant \frac{1}{6} \|X_i\|_\infty^3 \mathbb{E} |\partial_i^3 f(\mathbf{Z}_i^*)| + \frac{1}{6} \|Y_i\|_\infty^3 \mathbb{E} |\partial_i^3 f(\mathbf{Z}_i^{**})|$$

La Méthode de Lindeberg (Chatterjee 2006)

- Avec la notation $\mathbf{Z}_i = (X_1, \ldots, X_i, Y_{i+1}, \ldots, Y_n)$, on a

$$\mathbb{E}f(\mathbf{X}) - \mathbb{E}f(\mathbf{Y}) = \sum_{i=1}^{n} \big(\mathbb{E}f(\mathbf{Z}_i) - \mathbb{E}f(\mathbf{Z}_{i-1})\big)$$

- On note $\mathbf{Z}_i^0 = (X_1, \ldots, X_{i-1}, 0, Y_{i+1}, \ldots, Y_n)$. Grâce à la formule de Taylor, on a

$$\left|\mathbb{E}f(\mathbf{Z}_i) - \mathbb{E}f(\mathbf{Z}_{i-1})\right| \leqslant \frac{1}{6}\|X_i\|_\infty^3 \mathbb{E}|\partial_i^3 f(\mathbf{Z}_i^*)| + \frac{1}{6}\|Y_i\|_\infty^3 \mathbb{E}|\partial_i^3 f(\mathbf{Z}_i^{**})|$$

- On doit contrôler la borne pour la dérivée partielle d'ordre 3 de f.

Merci de votre attention!