



UNIVERSITÉ
LAVAL

Exploration des valeurs propres des matrices aléatoires et de leurs propriétés universelles

Travail présenté à M. Aurélien Nicosia dans le cadre du cours MAT-3600 : Projet de fin d'études

Travail réalisé par Thomas Turcotte

25 avril 2025

Table des matières

1	Introduction	1
2	Préliminaires	2
2.1	Matrices aléatoires	2
2.1.1	Matrices de Wishart	2
2.1.2	Matrice de Wigner	3
2.1.3	Cas particulier : EGO	3
2.1.4	Moments des matrices de Wigner	6
2.1.5	Invariance par rotation	7
2.1.6	Complément d'algèbre linéaire	9
2.2	Théorie de la mesure et convergence	9
2.2.1	Théorie de la mesure	10
2.2.2	Lois de probabilité classiques en théorie des matrices aléatoires	11
2.3	Résolvante et transformée de Stieltjes	13
2.3.1	Résolvante d'une matrice	13
2.3.2	Transformée de Stieltjes	14
3	Résultats classiques sur les matrices aléatoires	17
3.1	Convergence de la mesure spectrale pour des matrices de Wigner	17
3.1.1	Démonstration de la loi du demi-cercle	17
3.2	Convergence de la mesure spectrale pour des matrices de Wishart	21
3.3	Universalité pour la convergence de la mesure spectrale	22
4	Simulations	23
4.1	Loi du demi-cercle	23
4.1.1	Cas classique	23
4.1.2	Trois cas non couverts par la loi du demi-cercle	25
4.1.3	Loi du demi-cercle pour une variance infinie	27
4.2	Loi de Marchenko-Pastur	28
5	Matrices aléatoires en finance	31
6	Conclusion	34
A	Code pour l'application en finance	35
	Références	37

1 Introduction

L'étude des matrices aléatoires constitue un domaine relativement récent des mathématiques, qui a émergé au croisement de la statistique, de la physique et de l'analyse. C'est dans les années 1930 que Wishart introduit les premières matrices aléatoires à coefficients réels dans le cadre de la statistique multivariée, en étudiant des matrices de covariance de dimension $N \times N$. Quelques décennies plus tard, vers 1955, Wigner découvre une application inattendue de ces objets en physique nucléaire. Wigner a postulé que l'espacement entre les niveaux d'énergie du noyau d'un atome se comporte approximativement comme l'espacement entre les valeurs propres d'une matrice aléatoire.

Dans ce rapport, on se concentrera exclusivement sur l'étude des valeurs propres de certains types spécifiques de matrices aléatoires. On en déduira deux résultats aisément vérifiables à l'aide de simulations numériques : la loi du demi-cercle et la loi de Marchenko-Pastur. Deux résultats très importants dans la théorie des matrices aléatoires.

Dans la section 2, on introduira diverses notions théoriques qui ne sont pas abordées dans un baccalauréat en mathématiques. Les notions introduites vont être essentielles afin de bien comprendre la démonstration de la loi du demi-cercle. On passera ensuite à la section 3, section dans laquelle on énoncera la loi du demi-cercle et la loi de Marchenko-Pastur. On y démontrera aussi la loi du demi-cercle. Afin d'illustrer ces deux théorèmes, dans la section 4, on présentera diverses simulations des lois du demi-cercle et de Marchenko-Pastur. Dans la section 5, on présentera une application des matrices aléatoires dans le domaine de la finance.

2 Préliminaires

Dans cette section, on définit des concepts mathématiques qui n'ont pas été explorés durant le baccalauréat. L'objectif de cette section est d'introduire les matrices aléatoires, en particulier les matrices de Wishart et les matrices de Wigner. On passera ensuite à plusieurs notions théoriques qui vont permettre d'énoncer et de démontrer la loi du demi-cercle (théorème 3.1). On introduira les notions d'invariance par rotation, de moments des matrices de Wigner, de transformée de Stieltjes et de résolvante d'une matrice.

2.1 Matrices aléatoires

Tel que mentionné dans l'introduction, l'utilisation des matrices aléatoires remonte à environ 100 ans. Commençons par la définition d'une matrice aléatoire et passons ensuite au cas particulier des matrices de Wishart et de Wigner.

Définition 2.1 (Matrice aléatoire). On dit qu'une matrice A est une *matrice aléatoire* si toutes les entrées de A sont des variables aléatoires.

Voici un exemple d'une matrice aléatoire.

Exemple 2.1. Soient X_1, X_2 deux variables aléatoires qui suivent respectivement une loi uniforme sur l'intervalle $[0, 1]$ et une loi uniforme sur l'intervalle $[-1, 1]$, soient aussi X_3, X_4 deux variables aléatoires qui suivent une loi exponentielle de paramètre $\lambda = 2$ et enfin X_5, X_6 deux variables aléatoires qui suivent une loi de Bernoulli de paramètre $p = 1/2$. Alors la matrice

$$A := \begin{pmatrix} X_1 & X_4 & X_3 \\ X_5 & X_2 & X_6 \end{pmatrix},$$

est une matrice aléatoire.

Une matrice aléatoire n'est pas forcément carrée, comme on l'a vu dans l'exemple 2.1. On n'impose pas non plus que les entrées suivent toutes une même loi de probabilité. Cependant, pour la suite du rapport, nous travaillerons avec des matrices aléatoires symétriques, donc carrées, notamment pour pouvoir étudier le spectre de ces matrices.

2.1.1 Matrices de Wishart

Les matrices de Wishart sont un type de matrice aléatoire ayant une structure bien particulière. Elles correspondent aux matrices de covariance pour un jeu de données. Les matrices de Wishart seront particulièrement utiles dans la section 5, section dans laquelle on s'intéressera à des données financières.

Définition 2.2 (Matrice de Wishart). Soit $\{B_{i,j}\}_{1 \leq i \leq N, 1 \leq j \leq p}$ une famille de variables aléatoires indépendantes et identiquement distribuées (i. i. d.) d'espérance nulle et de variance 1. Soit B une matrice aléatoire de taille $N \times p$ (avec $N < p$) telle que l'entrée i, j de B correspond à la variable aléatoire $B_{i,j}$. Une *matrice de Wishart* est une matrice W_N telle que

$$W_N = \frac{1}{p} B B^t. \quad (2.1)$$

La construction d'une matrice de Wishart permet de prendre une matrice aléatoire rectangulaire et de la transformer en une matrice aléatoire carrée dont toutes les valeurs propres sont réelles, puisque W_N est symétrique. Ce résultat découle du théorème spectral.

Remarque 2.1. Les matrices de Wishart sont en fait les matrices de covariance des lignes de B .

Remarque 2.2. Si les variables aléatoires $\{B_{i,j}\}_{1 \leq i \leq N, 1 \leq j \leq p}$ de la définition 2.2 ont une espérance m et une variance σ^2 , on fait la transformation suivante pour retrouver une matrice de Wishart. On pose

$$\hat{B} = \frac{B - m J_{N \times p}}{\sigma},$$

où $J_{N \times p}$ est la matrice $N \times p$ dont chaque entrée vaut 1. On remplace ensuite B par \hat{B} dans l'équation (2.1) et l'on retombe sur une matrice de Wishart.

2.1.2 Matrice de Wigner

Les matrices de Wigner ont été introduites par Wigner dans les années 1950 pour modéliser les écarts entre les niveaux d'énergie du noyau d'un atome. Depuis ce temps, l'étude des matrices de Wigner a pris beaucoup d'importance et de nombreux résultats intéressants découlent de la définition qui suit.

Définition 2.3 (Matrice de Wigner). Soient $\{Z_{i,j}\}_{1 \leq i < j \leq N}$ et $\{Y_{i,i}\}_{1 \leq i \leq N}$ deux familles indépendantes de variables aléatoires indépendantes et identiquement distribuées (i.i.d.) d'espérances nulles et de variances respectives σ_{hd}^2 et σ_{d}^2 . Supposons de plus que tous les moments des variables aléatoires $Z_{i,j}$ et $Y_{i,i}$ sont bornés. Une *matrice de Wigner* est une matrice X_N telle que

$$(X_N)_{i,j} = (X_N)_{j,i} = \begin{cases} Z_{i,j}/\sqrt{N} & \text{si } i < j, \\ Y_{i,i}/\sqrt{N} & \text{si } i = j. \end{cases}$$

Faisons tout de suite la remarque suivante afin de fixer une notation tout au long du rapport.

Remarque 2.3. Afin d'éviter d'alourdir le texte, on dira que l'entrée i, j d'une matrice de Wigner suit une certaine loi de probabilité au lieu de dire que l'entrée i, j suit une certaine loi de probabilité multipliée par un facteur $\frac{1}{\sqrt{N}}$.

On remarque directement que les matrices de Wigner sont des matrices aléatoires possédant une structure particulière. Puisque les matrices de Wigner sont symétriques, toutes leurs valeurs propres sont réelles et c'est ce qui fait que le théorème central de ce rapport, la loi du demi-cercle, est si intéressant. En particulier, puisqu'il nous donne de l'information sur le comportement des valeurs propres lorsque la taille d'une matrice de Wigner tend vers l'infini. On verra dans la section 3 comment interpréter ce théorème à l'aide d'une figure.

Un ensemble de matrices de Wigner très intéressant est l'ensemble des matrices de Wigner dont les entrées $(X_N)_{i,i} \sim \mathcal{N}(0, 2\sigma_{\text{hd}}^2)$ et les entrées $(X_N)_{i,j} \sim \mathcal{N}(0, \sigma_{\text{hd}}^2)$ pour $i \neq j$ (en suivant la notation de la remarque 2.3). Une telle matrice de Wigner est en fait très facile à simuler numériquement. Il suffit de créer une matrice M telle que $M_{i,j} \sim \mathcal{N}(0, \sigma_{\text{hd}}^2/2)$ pour tout i, j . Pour obtenir la matrice de Wigner X_N , il suffit de poser

$$X_N = \frac{1}{\sqrt{N}}(M + M^t). \quad (2.2)$$

On a donc que

$$X_{i,i} \sim N(0, 2\sigma_{\text{hd}}^2),$$

et que

$$X_{i,j} \sim N(0, \sigma_{\text{hd}}^2) \quad \text{pour } i \neq j.$$

La matrice X_N est symétrique, car

$$(X_N)^t = \left(\frac{1}{\sqrt{N}}(M + M^t) \right)^t = \frac{1}{\sqrt{N}}(M^t + (M^t)^t) = \frac{1}{\sqrt{N}}(M^t + M) = \frac{1}{\sqrt{N}}(M + M^t) = X_N.$$

On a donc bel et bien une matrice de Wigner, puisque X_N est symétrique et que les entrées hors diagonale et sur la diagonale sont indépendantes. Des matrices possédant la forme de la matrice de l'équation (2.2) sont bien particulières. La démonstration du théorème principal de ce rapport, la loi du demi-cercle, dont l'énoncé est donné au théorème 3.1, sera faite dans le cas particulier des matrices de Wigner de cette forme avec $\sigma_{\text{d}}^2 = \sigma_{\text{hd}}^2 = \sigma^2$. Ces matrices sont des matrices issues de l'ensemble gaussien orthogonal à un facteur $\frac{1}{\sqrt{N}}$ près. La définition de l'ensemble gaussien orthogonal est donnée dans la prochaine section.

2.1.3 Cas particulier : EGO

Dans la théorie des matrices aléatoires, on s'intéresse parfois à certains groupes de matrices de Wigner. Un de ces groupes est celui des matrices de Wigner issues de l'ensemble gaussien orthogonal de paramètre σ^2 . Une telle matrice de Wigner sera définie sous peu, mais pour l'instant donnons la définition de l'ensemble gaussien orthogonal.

Définition 2.4 (EGO). L'Ensemble Gaussien Orthogonal de paramètre σ^2 (EGO_{σ^2}) est l'ensemble des matrices aléatoires symétriques dont les variables aléatoires sur la diagonale principale sont indépendantes et suivent une loi normale de moyenne 0 et de variance $2\sigma^2$ et les variables aléatoires au-dessus de la diagonale sont indépendantes et suivent une loi normale de moyenne 0 et de variance σ^2 .

Lorsque la valeur du paramètre σ^2 n'a pas d'importance, on parle simplement de l'ensemble gaussien orthogonal.

Puisqu'une matrice de l'ensemble gaussien orthogonal est symétrique, les entrées en dessous de la diagonale principale sont uniquement déterminées par les entrées au-dessus de la diagonale principale.

Soit $\{\xi_{i,j}\}_{j=1}^{\infty}$ une famille de variables aléatoires normales i. i. d. de moyenne 0 et de variance σ^2 . Les premières matrices de l'ensemble gaussien orthogonal de paramètre σ^2 sont

$$\begin{pmatrix} \sqrt{2}\xi_{1,1} & \xi_{1,2} \\ \xi_{1,2} & \sqrt{2}\xi_{2,2} \end{pmatrix}, \quad \begin{pmatrix} \sqrt{2}\xi_{1,1} & \xi_{1,2} & \xi_{1,3} \\ \xi_{1,2} & \sqrt{2}\xi_{2,2} & \xi_{2,3} \\ \xi_{1,3} & \xi_{2,3} & \sqrt{2}\xi_{3,3} \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} \sqrt{2}\xi_{1,1} & \xi_{1,2} & \xi_{1,3} & \xi_{1,4} \\ \xi_{1,2} & \sqrt{2}\xi_{2,2} & \xi_{2,3} & \xi_{2,4} \\ \xi_{1,3} & \xi_{2,3} & \sqrt{2}\xi_{3,3} & \xi_{3,4} \\ \xi_{1,4} & \xi_{2,4} & \xi_{3,4} & \sqrt{2}\xi_{4,4} \end{pmatrix}.$$

Remarque 2.4. Pour fixer la notation par la suite, on dira que X_N est une matrice de Wigner issue de l' EGO_{σ^2} si $X_N = \frac{1}{\sqrt{N}}M_N$ pour M_N une matrice $N \times N$ provenant de l'ensemble gaussien orthogonal de paramètre σ^2 . On a donc une manière simple de définir une matrice de Wigner à partir d'une matrice de l'EGO.

Comme l'étude de la convergence des valeurs propres des matrices aléatoires est l'objectif principal du projet, on pourrait s'intéresser aux valeurs propres d'une matrice 2×2 de l'ensemble gaussien orthogonal pour donner un exemple concret. L'exemple 2.2, provenant de [5], nous montre comment les valeurs propres d'une telle matrice interagissent entre elles.

Exemple 2.2. Soit $M = \begin{pmatrix} X_1 & X_3 \\ X_3 & X_2 \end{pmatrix}$ une matrice issue de l' EGO_1 , c'est-à-dire que $X_1, X_2 \sim \mathcal{N}(0, 2)$ et $X_3 \sim \mathcal{N}(0, 1)$ avec X_1, X_2, X_3 indépendantes. Quelle est la fonction de densité $p(s)$ de la variable aléatoire $S = \lambda_2 - \lambda_1$ où λ_1, λ_2 sont les valeurs propres de M et $\lambda_2 \geq \lambda_1$?

Les valeurs propres de M sont données par les racines de son polynôme caractéristique, c'est-à-dire que λ_1 et λ_2 sont les racines du polynôme

$$\lambda^2 - (X_1 + X_2)\lambda + X_1X_2 - X_3^2.$$

On a donc

$$\lambda_{1,2} = \frac{X_1 + X_2 \pm \sqrt{(X_1 + X_2)^2 - 4(X_1X_2 - X_3^2)}}{2} = \frac{X_1 + X_2 \pm \sqrt{(X_1 - X_2)^2 + 4X_3^2}}{2},$$

et

$$S = \sqrt{(X_1 - X_2)^2 + 4X_3^2}.$$

On se rappelle ici que les valeurs propres sont des fonctions de variables aléatoires et sont donc elles aussi des variables aléatoires. On voit tout de suite le contraste entre les valeurs propres d'une matrice déterministe qui sont des nombres réels fixes et les valeurs propres d'une matrice aléatoire qui sont des variables aléatoires et donc des valeurs qui ne sont pas déterministes. Par définition, on a

$$\begin{aligned} p(s) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{e^{-\frac{1}{4}x_1^2}}{2\sqrt{\pi}} \frac{e^{-\frac{1}{4}x_2^2}}{2\sqrt{\pi}} \frac{e^{-\frac{1}{2}x_3^2}}{\sqrt{2\pi}} \delta\left(s - \sqrt{(x_1 - x_2)^2 + 4x_3^2}\right) dx_1 dx_2 dx_3, \\ &= \frac{1}{2^{5/2}\pi^{3/2}} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-\frac{1}{4}x_1^2} e^{-\frac{1}{4}x_2^2} e^{-\frac{1}{2}x_3^2} \delta\left(s - \sqrt{(x_1 - x_2)^2 + 4x_3^2}\right) dx_1 dx_2 dx_3. \end{aligned}$$

Le terme $\delta\left(s - \sqrt{(x_1 - x_2)^2 + 4x_3^2}\right)$ permet de restreindre le domaine \mathbb{R}^3 aux triplets de points (x_1, x_2, x_3) tels que $s = \sqrt{(x_1 - x_2)^2 + 4x_3^2}$. On fait ensuite le changement de variable

$$\begin{cases} x_1 - x_2 &= r \cos(\theta), \\ 2x_3 &= r \sin(\theta), \\ x_1 + x_2 &= \psi, \end{cases}$$

d'où

$$\begin{cases} x_1 &= \frac{r \cos(\theta) + \psi}{2}, \\ x_2 &= \frac{\psi - r \cos(\theta)}{2}, \\ x_3 &= \frac{r \sin(\theta)}{2}. \end{cases}$$

Le jacobien associé à ce changement de variable est

$$\det \begin{pmatrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_1}{\partial \theta} & \frac{\partial x_1}{\partial \psi} \\ \frac{\partial x_2}{\partial r} & \frac{\partial x_2}{\partial \theta} & \frac{\partial x_2}{\partial \psi} \\ \frac{\partial x_3}{\partial r} & \frac{\partial x_3}{\partial \theta} & \frac{\partial x_3}{\partial \psi} \end{pmatrix} = \det \begin{pmatrix} \frac{\cos(\theta)}{2} & -\frac{r \sin(\theta)}{2} & \frac{1}{2} \\ -\frac{\cos(\theta)}{2} & \frac{r \sin(\theta)}{2} & \frac{1}{2} \\ \frac{\sin(\theta)}{2} & \frac{r \cos(\theta)}{2} & 0 \end{pmatrix} = -\frac{r}{4}.$$

On continue maintenant l'intégrale pour trouver $p(s)$.

$$\begin{aligned} p(s) &= \frac{1}{2^{5/2}\pi^{3/2}} \int_0^\infty \int_0^{2\pi} \int_{-\infty}^\infty \exp\left(\frac{-1}{4} \left[\frac{r \cos(\theta) + \psi}{2}\right]^2 - \frac{1}{4} \left[\frac{\psi - r \cos(\theta)}{2}\right]^2 - \frac{1}{2} \left[\frac{r \sin(\theta)}{2}\right]^2\right) \frac{r}{4} \delta(s - r) dr d\theta d\psi, \\ &= \frac{1}{2^{9/2}\pi^{3/2}} \int_0^\infty \int_0^{2\pi} \int_{-\infty}^\infty \exp\left(-\frac{r^2 \cos^2(\theta) + 2r\psi \cos(\theta) + \psi^2}{16} - \frac{r^2 \cos^2(\theta) - 2r\psi \cos(\theta) + \psi^2}{16} - \frac{r^2 \sin^2(\theta)}{8}\right) r \delta(s - r) dr d\theta d\psi, \\ &= \frac{1}{2^{9/2}\pi^{3/2}} \int_0^\infty \int_0^{2\pi} \int_{-\infty}^\infty \exp\left(-\frac{2r^2 \cos^2(\theta) + 2\psi^2}{16} - \frac{r^2 \sin^2(\theta)}{8}\right) r \delta(s - r) dr d\theta d\psi, \\ &= \frac{1}{2^{9/2}\pi^{3/2}} \int_0^\infty \int_0^{2\pi} \int_{-\infty}^\infty \exp\left(-\frac{r^2 + \psi^2}{8}\right) r \delta(s - r) dr d\theta d\psi, \\ &= \frac{1}{2^{9/2}\pi^{3/2}} \int_0^\infty r e^{-\frac{r^2}{8}} \delta(s - r) dr \cdot \int_0^{2\pi} 1 d\theta \cdot \int_{-\infty}^\infty e^{-\frac{\psi^2}{8}} d\psi, \\ &= \frac{1}{2^{9/2}\pi^{3/2}} s e^{-\frac{s^2}{8}} \cdot 2\pi \cdot \underbrace{\left(2\sqrt{2\pi} \int_{-\infty}^\infty \frac{1}{2\sqrt{2\pi}} e^{-\frac{\psi^2}{2(2^2)}} d\psi\right)}_{=1}, \\ &= \frac{2\pi \cdot 2\sqrt{2\pi}}{2^{9/2}\pi^{3/2}} s e^{-\frac{s^2}{8}} = \frac{s e^{-\frac{s^2}{8}}}{4}. \end{aligned}$$

De ce long calcul, on conclut que la fonction de densité de $S = \lambda_2 - \lambda_1$ est $p(s) = \frac{s e^{-\frac{s^2}{8}}}{4}$. On vérifie facilement que $\int_0^\infty p(s) = 1$ et que la fonction de répartition est $P(s) = 1 - e^{-\frac{s^2}{8}}$. On remarque par ailleurs que S suit une loi de Rayleigh de paramètre $\sigma^2 = 4$.

En simulant numériquement 100 000 fois des matrices 2×2 de l'EGO₁, en calculant les valeurs propres et en traçant un histogramme des valeurs de S obtenues, on trouve l'histogramme de la figure 1. On constate bel et bien que l'histogramme coïncide avec la fonction de densité obtenue théoriquement.

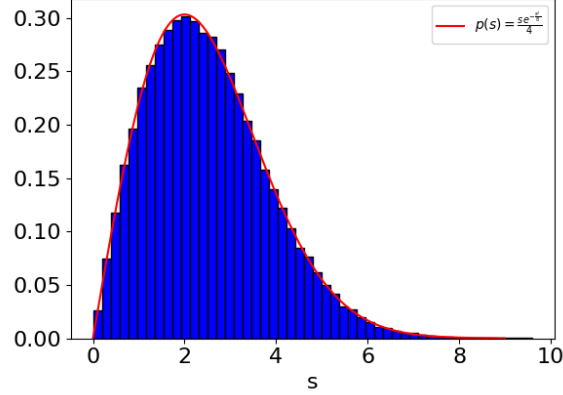


FIGURE 1 – Fonction de densité de l'écart entre les valeurs propres d'une matrice 2×2 issue de l'EGO₁.

2.1.4 Moments des matrices de Wigner

Pour une variable aléatoire Y , on définit le k -ième moment de Y par

$$m_k(Y) = \mathbb{E} [Y^k] .$$

Pour une matrice de Wigner X_N , la notion analogue à celle des moments est donnée dans la définition suivante.

Définition 2.5. Soit X_N une matrice de Wigner, son k -ième moment est donné par

$$\tau_k(X_N) := \frac{1}{N} \mathbb{E} [\text{Tr} (X_N^k)] .$$

La ressemblance entre les moments d'une variable aléatoire et ceux d'une matrice de Wigner n'est à priori pas évidente. Regardons d'un peu plus près la notion de moment pour une matrice de Wigner. On s'intéresse principalement aux valeurs propres des matrices de Wigner. Comme la trace d'une matrice nous donne la somme de ses valeurs propres, la trace nous donne donc de l'information sur les valeurs propres de la matrice de Wigner. Le facteur $\frac{1}{N}$ peut jouer deux rôles principaux. Premièrement, il fait en sorte qu'en calculant le k -ième moment, on calcule la moyenne des valeurs propres de X_N^k . Deuxièmement, ce facteur vient faire en sorte que les moments ne dépendent pas de la taille de la matrice lorsque N devient très grand. Enfin, tout comme le moment d'une variable aléatoire donne de l'information sur la variable, les moments d'une matrice de Wigner donnent de l'information sur la matrice de Wigner.

De cette définition, on peut aisément calculer les premiers moments d'une matrice de Wigner du moment qu'on connaît la loi de probabilité suivie par les entrées de la diagonale principale.

Exemple 2.3. Calculons les premiers moments de X_N pour X_N une matrice provenant de l'EGO _{σ^2} . Pour cet exemple, on oublie la notation introduite dans la remarque 2.3 afin de rendre le calcul plus explicite. On a donc $(X_N)_{i,j} \sim \mathcal{N}(0, \sigma^2/N)$ si $i \neq j$ et $(X_N)_{i,i} \sim \mathcal{N}(0, 2\sigma^2/N)$.

Le premier moment de X_N est

$$\tau_1(X_N) = \frac{1}{N} \mathbb{E} [\text{Tr}(X_N)] = \frac{1}{N} \mathbb{E} \left[\sum_{i=1}^N (X_N)_{i,i} \right] = \frac{1}{N} \sum_{i=1}^N \mathbb{E} [(X_N)_{i,i}] = 0.$$

Le deuxième moment de X_N est

$$\begin{aligned}
\tau_2(X_N) &= \frac{1}{N} \mathbb{E} [\text{Tr}(X_N^2)], \\
&= \frac{1}{N} \mathbb{E} \left[\sum_{k=1}^N \sum_{i=1}^N (X_N)_{k,i} (X_N)_{i,k} \right], \\
&= \frac{1}{N} \mathbb{E} \left[\sum_{k=1}^N \sum_{i=1}^N ((X_N)_{k,i})^2 \right], \\
&= \frac{1}{N} \sum_{k=1}^N \sum_{i=1}^N \mathbb{E} [((X_N)_{k,i})^2], \\
&= \frac{1}{N} \sum_{\substack{k=1 \\ k \neq i}}^N \sum_{i=1}^N \mathbb{E} [((X_N)_{k,i})^2] + \sum_{i=1}^N \mathbb{E} [((X_N)_{i,i})^2].
\end{aligned}$$

On sait que, pour $k \neq i$, les $(X_N)_{k,i} \sim \mathcal{N}(0, \sigma^2/N)$, on a donc $\frac{N}{\sigma^2} ((X_N)_{k,i})^2 \sim \chi_1^2$, d'où $\mathbb{E} [((X_N)_{k,i})^2] = \frac{\sigma^2}{N}$. De manière similaire, on a que $\mathbb{E} [((X_N)_{i,i})^2] = \frac{2\sigma^2}{N}$. On obtient donc

$$\tau_2(X_N) = \frac{1}{N} \left((N^2 - N) \frac{\sigma^2}{N} + N \frac{2\sigma^2}{N} \right) = \frac{\sigma^2(N-1)}{N} + \frac{2\sigma^2}{N} = \sigma^2 + \frac{\sigma^2}{N}.$$

On constate alors que, même lorsque $N \rightarrow \infty$, le deuxième moment de X_N reste borné. On voit pourquoi on a un facteur $\frac{1}{\sqrt{N}}$ sur chacune des entrées de la matrice. Cela permet de s'assurer que les moments restent bornés lorsque N augmente. Ce facteur $\frac{1}{\sqrt{N}}$ nous assure aussi que les valeurs propres ne vont pas dépendre de la taille de la matrice.

La trace normalisée d'une matrice de Wigner à la puissance k converge en probabilité vers le k -ième moment de cette même matrice. Le théorème suivant nous donne le résultat.

Théorème 2.1. Soit X_N une matrice de Wigner. Pour tout $k \in \mathbb{N}$ et pour tout $\varepsilon > 0$,

$$\lim_{N \rightarrow \infty} \mathbb{P} \left(\left| \frac{1}{N} \text{Tr}(X_N^k) - \frac{1}{N} \mathbb{E} [\text{Tr}(X_N^k)] \right| > \varepsilon \right) = 0.$$

On a donc que la trace normalisée de X_N^k converge en probabilité vers le k -ième moment de X_N^k .

2.1.5 Invariance par rotation

On se rappelle tout d'abord que toute matrice de rotation est une matrice orthogonale O . On a donc que O satisfait $O^{-1} = O^t$. On fait la définition suivante.

Définition 2.6. Une matrice aléatoire carrée A est dite *invariante par rotation* si

$$OAO^t \stackrel{\text{en loi}}{=} A,$$

pour toute matrice orthogonale O . En d'autres termes, A est invariante par rotation si $A_{i,j}$ et $(OAO^t)_{i,j}$ suivent la même loi de probabilité pour $1 \leq i, j \leq N$.

Afin de montrer qu'une matrice de Wigner X_N issue de l'EGO $_{\sigma^2}$ est invariante par rotation, montrons d'abord qu'un vecteur v tel que $v_{i,1} \sim \mathcal{N}(0, \sigma^2)$ pour $1 \leq i \leq N$ est tel que $(Ov)_{i,1} \stackrel{\text{en loi}}{=} v_{i,1}$ pour une matrice de rotation O . Pour ce faire, montrons que $(Ov)_{i,1}$ suit la même loi que $v_{i,1}$. De la définition du produit matriciel, on a que

$$(Ov)_{i,1} = \sum_{k=1}^N O_{i,k} v_{k,1}.$$

Comme $v_{k,1} \sim \mathcal{N}(0, \sigma^2)$, on a que $(Ov)_{i,1} \sim \mathcal{N}(0, (O_{i,k})^2 \sigma^2)$. On a donc

$$(Ov)_{i,1} \sim \mathcal{N}\left(0, \sigma^2 \sum_{k=1}^N (O_{i,k})^2\right).$$

Puisque O est orthogonale, $\sum_{k=1}^N (O_{i,k})^2 = 1$, on a finalement que

$$(Ov)_{i,1} \sim \mathcal{N}(0, \sigma^2).$$

On conclut immédiatement que $Ov \stackrel{\text{en loi}}{=} v$. Soit maintenant M une matrice $N \times N$ telle que $M_{i,j} \sim \mathcal{N}(0, \sigma^2/2)$ pour $1 \leq i, j \leq N$. De ce qui vient d'être fait, on sait que chacune des colonnes c_k de M satisfait $Oc_k \stackrel{\text{en loi}}{=} c_k$. Comme la k -ième colonne du produit OM correspond au produit Oc_k et que $Oc_k \stackrel{\text{en loi}}{=} c_k$, il est immédiat que $OM \stackrel{\text{en loi}}{=} M$. Par le travail qui vient d'être fait, il est clair que $OMO^t \stackrel{\text{en loi}}{=} M$. Le calcul

$$OX_N O^t = O \left[\frac{1}{\sqrt{N}} (M + M^t) \right] O^t = \frac{1}{\sqrt{N}} (OMO^t + OM^t O^t) \stackrel{\text{en loi}}{=} \frac{1}{\sqrt{N}} (M + M^t) = X_N,$$

montre finalement que X_N est invariante par rotation. L'invariance par rotation de X_N servira dans la section 3.

Afin d'illustrer l'invariance par rotation, on peut faire une simulation numérique. Prenons une matrice X_{10} où X_{10} est une matrice 10×10 issue de L'EGO₁. On prend une matrice de rotation O générée avec Python. Dans les histogrammes de la figure 2, on considère les entrées 1,1 et 3,5 pour les matrices X_{10} et $OX_{10}O^t$.

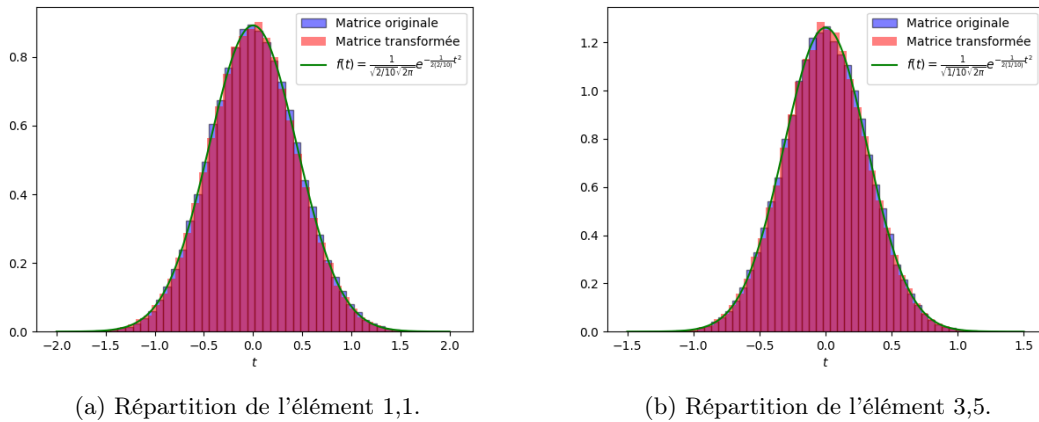


FIGURE 2 – Histogrammes des valeurs prises par X_{10} et $OX_{10}O^t$.

Pour tracer ces histogrammes, on a généré 1 000 000 fois des matrices X_{10} , on a noté la valeur des entrées 1,1 et 3,5 et fait deux histogrammes. On a fait la même chose pour la matrice $OX_{10}O^t$ et l'on constate que, dans le cas d'une entrée sur la diagonale comme à la figure 2a, la distribution de $(X_{10})_{1,1}$ et $(OX_{10}O^t)_{1,1}$ convergent vers la même distribution lorsque le nombre de répétitions augmente. On remarque d'ailleurs que la répartition coïncide avec la fonction de densité d'une loi normale de moyenne nulle et de variance $2/10$. Cette densité est tracée en vert sur la figure 2a. On observe le même phénomène pour une entrée hors diagonale, comme on peut le voir sur la figure 2b. Dans ce cas, les deux distributions coïncident avec la fonction de densité d'une loi normale de moyenne nulle et de variance $1/10$, comme ce qui était attendu. Cette densité est tracée en vert sur la figure 2b. Certaines transformations matricielles préservent l'invariance par rotation, la proposition suivante en donne deux qui seront utiles plus tard.

Proposition 2.1. Soit A une matrice de dimension $N \times N$ invariante par rotation. Alors

1. pour tout $z \in \mathbb{C}$, la matrice $zI_N - A$ est invariante par rotation ;
2. A^{-1} est invariante par rotation.

Démonstration.


1) On a

$$O(zI_N - A)O^t = \underbrace{zOI_NO^t}_{\text{déterministe}} - OAO^t \stackrel{\text{en loi}}{=} zI_N - A.$$

D'où $zI_N - A$ est invariante par rotation.

2) De l'invariance par rotation de A , on a

$$A^{-1} \stackrel{\text{en loi}}{=} (OAO^t)^{-1} = (O^t)^{-1}A^{-1}O^{-1} = OA^{-1}O^t.$$

D'où A^{-1} est invariante par rotation. 


Proposition 2.2. Soit A une matrice invariante par rotation. Alors toutes les entrées de la diagonale principale de A ont la même espérance.

Démonstration. Soit A une matrice de dimension $N \times N$ invariante par rotation. On considère $E_{i,j}$, la matrice élémentaire qui interchange les lignes i et j avec $i \neq j$. Clairement, $E_{i,j}$ est orthogonale. On s'intéresse aux éléments $A_{i,i}$ et $A_{j,j}$ dans le produit

$$E_{i,j}AE_{i,j}.$$

On considère d'abord le produit $E_{i,j}A =: M$. Ce faisant, l'élément $A_{i,i}$ est envoyé en position j, i et l'élément $A_{j,j}$ est envoyé en position i, j . Considérons maintenant le produit $ME_{i,j}$. L'élément $A_{i,i}$ en position j, i est envoyé en position j, j et l'élément $A_{j,j}$ en position i, j est envoyé en position i, i . De l'invariance par rotation de A , on doit avoir

$$A_{j,j} = (E_{i,j}AE_{i,j})_{i,i} \stackrel{\text{en loi}}{=} A_{i,i} \quad \text{et} \quad A_{i,i} = (E_{i,j}AE_{i,j})_{j,j} \stackrel{\text{en loi}}{=} A_{j,j}.$$

Il suit de cette équation que $A_{i,i}$ et $A_{j,j}$ doivent suivre la même loi, ils ont donc la même espérance. En faisant varier i et j , on obtient que tous les éléments de la diagonale principale ont la même espérance. 

2.1.6 Complément d'algèbre linéaire

Cette courte section permet d'énoncer un théorème très utile pour inverser certains blocs d'une matrice par blocs. Ce théorème jouera un rôle clé dans la démonstration du théorème 3.1.

Théorème 2.2 (Complément de Schur). Soit M une matrice inversible de dimension $N \times N$ qu'on divise en quatre blocs comme ceci

$$M = \begin{pmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{pmatrix} \quad \text{et} \quad M^{-1} = Q = \begin{pmatrix} Q_{1,1} & Q_{1,2} \\ Q_{2,1} & Q_{2,2} \end{pmatrix},$$

où $B_{1,1}$ est un bloc $n \times n$ et $B_{2,2}$ est un bloc $(N - n) \times (N - n)$. Alors

$$Q_{1,1}^{-1} = B_{1,1} - B_{1,2}B_{2,2}^{-1}B_{2,1}. \tag{2.3}$$

Bien que le théorème ne semble pas lié au début de la section, c'est un résultat clé pour arriver au théorème 3.1, comme on le verra dans la section 3.

2.2 Théorie de la mesure et convergence

Les théorèmes de la section 3 font intervenir la notion de mesure. On introduit donc diverses notions liées à la théorie de la mesure qui seront utiles dans ce rapport.

2.2.1 Théorie de la mesure

Définition 2.7. Soit X un ensemble et Σ une σ -algèbre sur X . Une fonction ensembliste μ de Σ vers la droite réelle achevée est appelée une *mesure* si les conditions suivantes sont respectées :

- i) pour tout $E \in \Sigma$, $\mu(E) \geq 0$;
- ii) $\mu(\emptyset) = 0$;
- iii) pour toute collection dénombrable $\{E_k\}_{k=1}^{\infty}$ d'ensembles deux à deux disjoints dans Σ , on a

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k).$$

De manière similaire, on définit une mesure de probabilité ainsi.

Définition 2.8. Soit X un ensemble et Σ une σ -algèbre sur X . Une mesure \mathbb{P} est une *mesure de probabilité* si les trois conditions suivantes sont respectées :

- i) $0 \leq \mathbb{P}(E) \leq 1$ pour tout $E \in \Sigma$;
- ii) $\mathbb{P}(\emptyset) = 0$ et $\mathbb{P}(X) = 1$;
- iii) Pour toute collection dénombrable $\{E_k\}_{k=1}^{\infty}$ d'ensembles deux à deux disjoints dans Σ , on a

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k).$$

On remarque qu'une mesure de probabilité est une mesure particulière où la mesure de l'ensemble complet vaut 1.

Définition 2.9. Soit (X, \mathcal{A}) un espace mesurable et $a \in X$. On appelle *mesure de Dirac au point a* , que l'on note δ_a , la mesure de probabilité sur (X, \mathcal{A}) définie par :

$$\text{pour tout } a \in X, \delta_a(A) = \begin{cases} 1 & \text{si } a \in A, \\ 0 & \text{si } a \notin A. \end{cases} \quad (2.4)$$

La mesure de Dirac permet enfin de définir la mesure spectrale, un outil qui servira pour énoncer les théorèmes de la section 3.

Définition 2.10. Notons les valeurs propres d'une matrice aléatoire symétrique A de taille $N \times N$ par λ_i^N , avec $\lambda_1^N \leq \lambda_2^N \leq \dots \leq \lambda_N^N$. Pour $E \subseteq \mathbb{R}$, la *mesure spectrale* de A sur l'ensemble E est donnée par

$$L_N(E) := \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i^N}(E).$$

En particulier, pour $E \subseteq \mathbb{R}$, on a que $L_N(E)$ donne la proportion de valeurs propres de X_N dans E . Si $E = [a, b]$, alors $L_N([a, b])$ donne la proportion de valeurs propres de X_N dans l'intervalle $[a, b]$.

Afin de compléter la section sur la théorie de la mesure, on introduit la convergence faible pour une suite de mesures.

Définition 2.11. Soit $(\mu_n)_{n \in \mathbb{N}}$ une suite de mesures. On dit que la suite $(\mu_n)_{n \in \mathbb{N}}$ *converge faiblement* vers une mesure μ si pour toute fonction f continue et bornée sur \mathbb{R} , on a

$$\lim_{n \rightarrow \infty} \left(\int_{\mathbb{R}} f d\mu_n - \int_{\mathbb{R}} f d\mu \right) = 0.$$

Cette convergence faible sera utile pour caractériser la convergence de la mesure spectrale.

2.2.2 Lois de probabilité classiques en théorie des matrices aléatoires

Dans des théorèmes qui seront énoncés dans la section 3, certaines lois de probabilités qui ne sont pas explorées dans un cours classique de probabilités interviennent. On définit donc deux lois de probabilités intéressantes qui serviront un peu plus tard.

Définition 2.12. La loi du demi-cercle de paramètre $\sigma > 0$ est la mesure de probabilité $\rho(x)dx$ sur \mathbb{R} dont la densité est donnée par

$$\rho(x) = \frac{\sqrt{4\sigma^2 - x^2}}{2\pi\sigma^2} \quad \text{pour } -2\sigma < x < 2\sigma.$$

La fonction de densité est un demi-cercle de rayon 2σ « écrasé » sur l'axe des x . En effet, la constante $\frac{1}{2\pi\sigma^2}$ sert de constante de normalisation pour s'assurer que $\rho(x)$ est bien une fonction de densité. On voit donc bien d'où vient le nom loi du demi-cercle. À la figure 3, le graphe de la fonction de densité est donné pour le cas $\sigma = 1$.

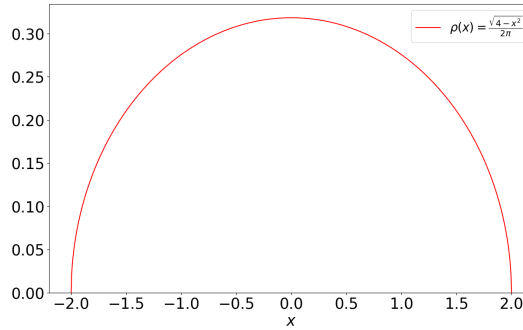


FIGURE 3 – Fonction de densité de la loi du demi-cercle pour $\sigma = 1$.

De la définition de la fonction de densité $\rho(x)$ à la définition 2.12, on déduit directement les propriétés suivantes. Soit X une variable aléatoire qui suit une loi du demi-cercle de paramètre σ . Alors

1. la fonction densité $\rho(x)$ est symétrique par rapport à l'axe des ordonnées ;
2. l'espérance de X est 0 ;
3. la variance de X est σ^2 .


Proposition 2.3. Soit Y une variable aléatoire qui suit une loi bêta de paramètre $\alpha = \beta = \frac{3}{2}$. Alors $X = 4\sigma Y - 2\sigma$ suit une loi du demi-cercle de paramètre σ .

Démonstration. Soient $f(x)$ et $g(y)$ les fonctions de densité de X et Y . Soient aussi $F(x)$ et $G(y)$ les fonctions de répartition de X et Y . Montrons que la fonction de densité de X correspond à la fonction de densité de la loi du demi-cercle de paramètre σ . On a

$$\begin{aligned} F(x) &= \mathbb{P}[X \leq x], \\ &= \mathbb{P}[4\sigma Y - 2\sigma \leq x], \\ &= \mathbb{P}\left[Y \leq \frac{x + 2\sigma}{4\sigma}\right], \\ &= G\left(\frac{x + 2\sigma}{4\sigma}\right). \end{aligned}$$

De la relation $f(x) = F'(x)$, on déduit que

$$\begin{aligned}
f(x) &= F'(x), \\
&= G' \left(\frac{x+2\sigma}{4\sigma} \right), \\
&= g \left(\frac{x+2\sigma}{4\sigma} \right) \cdot \frac{1}{4\sigma}, \\
&= \frac{1}{4\sigma} \frac{8}{\pi} \left(\frac{x+2\sigma}{4\sigma} \right)^{1/2} \left(1 - \frac{x+2\sigma}{4\sigma} \right)^{1/2}, \\
&= \frac{1}{2\pi\sigma^2} (x+2\sigma)^{1/2} (4\sigma - x - 2\sigma)^{1/2}, \\
&= \frac{1}{2\pi\sigma^2} (4\sigma x - x^2 - 2\sigma x + 8\sigma^2 - 2\sigma x - 4\sigma^2)^{1/2}, \\
&= \frac{1}{2\pi\sigma^2} \sqrt{4\sigma^2 - x^2}.
\end{aligned}$$

De l'unicité de la fonction de densité, on conclut immédiatement que X suit une loi du demi-cercle de paramètre σ . 

De cette proposition, on a que la loi du demi-cercle correspond à une loi bêta à un facteur et un décalage près.

On définit maintenant une loi de probabilité associée à une matrice de Wishart. Soit B la matrice $N \times p$ associée à une matrice de Wishart W_N . On pose $\gamma = N/p$.

Définition 2.13. La loi de Marchenko-Pastur de paramètre γ où $0 < \gamma < 1$ est la mesure de probabilité $\varphi(x)dx$ sur \mathbb{R} dont la densité est donnée par

$$\varphi(x) = \frac{\sqrt{(\gamma_+ - x)(x - \gamma_-)}}{2\pi\gamma x} \quad \text{pour } \gamma_- < x < \gamma_+,$$

où $\gamma_{\pm} = (1 \pm \sqrt{\gamma})^2$.

Comparativement à la loi du demi-cercle où la valeur de σ fait seulement en sorte « d'aplatir » le demi-cercle, le paramètre γ dans la loi de Marchenko-Pastur peut complètement changer l'allure de la fonction de densité, comme on peut l'observer à la figure 4.

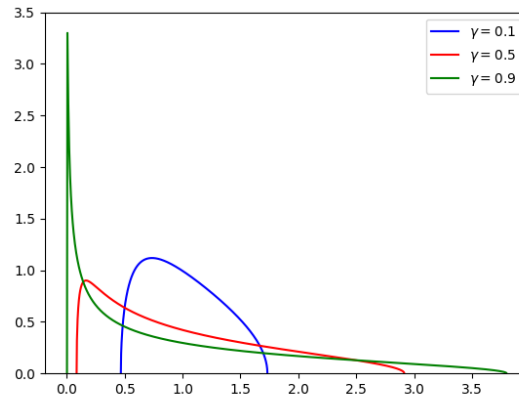



FIGURE 4 – Loi de Marchenko-Pastur pour différentes valeurs de γ .

On remarque que plus γ est près de 1, plus le pic près de 0 est haut et pointu. Tandis que lorsque γ est autour de 0, la courbe semble se rapprocher d'un demi-cercle déformé, il n'y a alors plus de pic prononcé près de 0.

Comparativement à la loi du demi-cercle où l'espérance d'une variable aléatoire qui suit une loi du demi-cercle de paramètre σ est facile à calculer, il est beaucoup moins évident de faire le même calcul pour une variable aléatoire qui suit une loi de Marchenko-Pastur. La proposition suivante donne une expression pour calculer facilement le k -ième moment d'une variable aléatoire qui suit une loi de Marchenko-Pastur.

Proposition 2.4. Soit X une variable aléatoire qui suit une loi de Marchenko-Pastur de paramètre γ . Alors le k -ième moment de X est donné par

$$\mathbb{E}[X^k] = \sum_{r=0}^{k-1} \frac{1}{1+r} \binom{k}{r} \binom{k-1}{r} \gamma^r.$$

Démonstration. La démonstration se trouve dans [2]. 

Avec cette proposition, on peut tout de suite calculer l'espérance et la variance d'une variable aléatoire qui suit une loi de Marchenko-Pastur.

Exemple 2.4. Soit X une variable aléatoire qui suit une loi de Marchenko-Pastur de paramètre γ . Alors

$$\begin{aligned} \mathbb{E}[X] &= \sum_{r=0}^0 \frac{1}{1+r} \binom{k}{r} \binom{k-1}{r} \gamma^r, \\ &= \frac{1}{1+0} \binom{1}{0} \binom{0}{0} \gamma^0 = 1. \end{aligned}$$

Peu importe le paramètre γ , l'espérance de X est 1. Calculons le deuxième moment de X .

$$\begin{aligned} \mathbb{E}[X^2] &= \sum_{r=0}^1 \frac{1}{1+r} \binom{k}{r} \binom{k-1}{r} \gamma^r, \\ &= \mathbb{E}[X] + \frac{1}{1+1} \binom{2}{1} \binom{1}{1} \gamma, \\ &= 1 + \frac{1}{2} \cdot \frac{2!}{1!1!} \cdot \frac{1!}{0!1!} \gamma = 1 + \gamma. \end{aligned}$$

On conclut alors que le deuxième moment de X est $1 + \gamma$.

De cet exemple, on sait que l'espérance d'une variable aléatoire X qui suit une loi de Marchenko-Pastur de paramètre γ est 1. On peut aussi calculer sa variance en utilisant la formule $\mathbb{V}\text{ar}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. On obtient ainsi que

$$\mathbb{V}\text{ar}[X] = (1 + \gamma) - 1^2 = \gamma.$$

2.3 Résolvante et transformée de Stieltjes

Deux notions centrales de la démonstration de la loi du demi-cercle sont la résolvante d'une matrice et la transformée de Stieltjes d'une mesure. Voyons ces deux notions qui vont revenir dans la section 3.

2.3.1 Résolvante d'une matrice

Définition 2.14. Soit X_N une matrice de Wigner. La *résolvante* de X_N est donnée par

$$G_{X_N}(z) := (zI_N - X_N)^{-1},$$

où I_N est la matrice identité $N \times N$ et z est un nombre complexe dont la partie imaginaire est non nulle.

On peut immédiatement faire un lien entre la résolvante et l'invariance par rotation.

Lemme 2.1. La résolvante d'une matrice de Wigner issue de l'EGO $_{\sigma^2}$ est invariante par rotation.

Démonstration. Une matrice de Wigner issue de l'EGO $_{\sigma^2}$ est invariante par rotation. On applique ensuite la proposition 2.1. 

Encore un fois, ce résultat sera utile dans la démonstration du théorème de Wigner donnée dans la section 3.

2.3.2 Transformée de Stieltjes

Définissons maintenant la transformée de Stieltjes pour une mesure μ .

Définition 2.15. La *transformée de Stieltjes* d'une mesure μ est

$$S_\mu(z) = \int_{\mathbb{R}} \frac{\mu(x)dx}{z - x} \quad z \in \mathbb{C} \setminus \mathbb{R}.$$

On retrouve certaines propriétés intéressantes par rapport à la transformée de Stieltjes. On a d'abord que la fonction S_μ est holomorphe dans le demi-plan $\text{Im}(z) \neq 0$ et encode des informations essentielles sur μ .


Dans le cas où l'on considère la transformée de Stieltjes de la mesure spectrale d'une matrice de Wigner X_N , la résolvante de X_N apparaît dans son expression. En effet, la transformée de Stieltjes de la mesure spectrale de X_N s'exprime à l'aide de la résolvante de X_N . En prenant la mesure spectrale L_N associée à la matrice de Wigner X_N , on a en fait que $S_{L_N}(z)$ correspond à $\frac{1}{N} \text{Tr}(G_{X_N}(z))$ puisque

$$S_{L_N}(z) = \int_{\mathbb{R}} \frac{L_N dx}{z - x} = \int_{\mathbb{R}} \frac{\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i^N}}{z - x} dx = \frac{1}{N} \sum_{i=1}^N \frac{1}{z - \lambda_i^N} = \frac{1}{N} \text{Tr}(G_{X_N}(z)). \quad (2.5)$$

Dans le cas où on considère la transformée de Stieltjes de la mesure spectrale d'une matrice de Wigner, on utilisera l'expression de la transformée de Stieltjes en fonction de la trace de X_N au lieu de la notation intégrale. L'équation (2.5) est valide pour toute valeur de $z \in \mathbb{C} \setminus \mathbb{R}$. En particulier, comme les valeurs propres de X_N sont réelles, le dénominateur de la somme de l'équation (2.5) est toujours non nul.

Le théorème suivant fait le pont entre la section sur la théorie de la mesure et la section sur la transformée de Stieltjes.

Théorème 2.3. Soit μ_n une suite de mesures de probabilités sur \mathbb{R} . Si les mesures de probabilités μ_n sont aléatoires et que, pour tout $z \in \mathbb{C} \setminus \mathbb{R}$, on a que $S_{\mu_n}(z)$ converge vers une limite déterministe $S(z)$ qui est la transformée de Stieltjes d'une mesure de probabilité μ , alors μ_n converge faiblement vers μ .

Démonstration. La démonstration se trouve dans [1]. 

La transformée de Stieltjes constitue un outil puissant pour étudier la convergence de suites de mesures. En effet, la convergence faible d'une suite μ_n vers μ est équivalente à la convergence de leurs transformées de Stieltjes $S_{\mu_n}(z)$ vers $S_\mu(z)$ pour tout z du demi-plan $\text{Im}(z) \neq 0$. Par ailleurs, lorsque la limite $\lim_{n \rightarrow \infty} S_{\mu_n}(z)$ existe, elle est nécessairement la transformée de Stieltjes d'une unique mesure μ , ce qui permet de reconstruire la mesure-limite à partir de la fonction limite.

Puisque la loi du demi-cercle est une loi de probabilité importante pour la suite, calculons sa transformée de Stieltjes. Ce calcul ne donne pas seulement un exemple de calcul de transformée de Stieltjes. Ce calcul sera utile afin de compléter la démonstration du théorème central de ce rapport, c'est-à-dire la loi du demi-cercle énoncée au théorème 3.1.

Exemple 2.5. Calculons la transformée de Stieltjes pour la loi du demi-cercle. Soit $z \in \mathbb{C} \setminus \mathbb{R}$, alors

$$S_\rho(z) = \int_{\mathbb{R}} \frac{d\rho(x)}{z-x} = -\frac{1}{2\pi\sigma^2} \int_{-2\sigma}^{2\sigma} \frac{1}{x-z} \sqrt{4\sigma^2 - x^2} dx.$$

En faisant le changement de variable $x = 2\sigma \cos(y)$, on obtient

$$\begin{aligned} S_\rho(z) &= -\frac{1}{2\pi\sigma^2} \int_{\pi}^0 \frac{1}{2\sigma \cos(y) - z} 2\sigma \sin(y) (-2\sigma \sin(y)) dy, \\ &= -\frac{2}{\pi} \int_0^{\pi} \frac{\sin^2(y)}{2\sigma \cos(y) - z} dy. \end{aligned}$$

De la symétrie de $\sin^2(y)$ et de $\frac{1}{2\sigma \cos(y) - z}$ par rapport à $y = \pi$, on obtient

$$\begin{aligned} S_\rho(z) &= -\frac{1}{\pi} \int_0^{2\pi} \frac{1}{2\sigma \cos(y) - z} \sin^2(y) dy, \\ &= -\frac{1}{\pi} \int_0^{2\pi} \frac{1}{2\sigma \left(\frac{e^{iy} + e^{-iy}}{2} \right) - z} \left(\frac{e^{iy} - e^{-iy}}{2i} \right)^2 dy, \\ &= \frac{1}{4\pi} \int_0^{2\pi} \frac{1}{\sigma(e^{iy} + e^{-iy}) - z} (e^{iy} - e^{-iy})^2 \frac{ie^{iy}}{ie^{iy}} dy. \end{aligned}$$

De la dernière égalité, on reconnaît l'intégrale sur le contour fermé $|\zeta| = 1$. On a donc

$$\begin{aligned} S_\rho(z) &= \frac{1}{4\pi i} \int_{|\zeta|=1} \frac{1}{\sigma(\zeta + \zeta^{-1}) - z} (\zeta - \zeta^{-1})^2 \zeta^{-1} d\zeta, \\ &= \frac{1}{4\pi i} \int_{|\zeta|=1} \frac{1}{\sigma(\zeta + \zeta^{-1}) - z} (\zeta - \zeta^{-1})^2 \zeta^{-1} d\zeta, \\ &= \frac{1}{4\pi i} \int_{|\zeta|=1} \frac{\frac{1}{\zeta^2} (\zeta^2 - 1)^2}{\frac{1}{\zeta} [\sigma(\zeta^2 + 1) - \zeta z]} \frac{1}{\zeta} d\zeta, \\ &= \frac{1}{4\pi i} \int_{|\zeta|=1} \frac{(\zeta^2 - 1)^2}{\zeta^2 (\sigma \zeta^2 - \zeta z + \sigma)} d\zeta, \\ &= \frac{1}{4\pi i} \int_{|\zeta|=1} \frac{(\zeta^2 - 1)^2}{\sigma \zeta^2 \left(\zeta - \frac{z - \sqrt{z^2 - 4\sigma^2}}{2\sigma} \right) \left(\zeta - \frac{z + \sqrt{z^2 - 4\sigma^2}}{2\sigma} \right)} d\zeta. \end{aligned} \tag{2.6}$$

On voit que la dernière intégrale a trois pôles distincts, soient $\zeta_0 = 0$, $\zeta_1 = \frac{z - \sqrt{z^2 - 4\sigma^2}}{2\sigma}$ et $\zeta_2 = \frac{z + \sqrt{z^2 - 4\sigma^2}}{2\sigma}$. On a que ζ_0 est un pôle d'ordre 2 alors que ζ_1 et ζ_2 sont tous deux d'ordre 1. De plus, on a $|\zeta_2| > |\zeta_1|$. Cette affirmation ne sera pas justifiée dans ce rapport, puisque les calculs sont assez longs, mais les détails sont donnés dans [2].

Remarque 2.5. On a que

$$\zeta_1 \zeta_2 = 1.$$

En effet, en effectuant le calcul, on a

$$\zeta_1 \zeta_2 = \left(\frac{z - \sqrt{z^2 - 4\sigma^2}}{2\sigma} \right) \left(\frac{z + \sqrt{z^2 - 4\sigma^2}}{2\sigma} \right) = \frac{z^2 - (z^2 - 4\sigma^2)}{4\sigma^2} = 1.$$

Par la remarque 2.5, on a que $\zeta_1 \zeta_2 = 1$, donc que $|\zeta_1| |\zeta_2| = 1$. Puisqu'on sait que $|\zeta_2| > |\zeta_1|$, on doit avoir

$$|\zeta_1| < 1,$$

donc que ζ_1 est contenu dans l'intérieur de la courbe $|\zeta| = 1$. On peut donc évaluer l'intégrale (2.6) en utilisant le théorème des résidus. On a alors

$$S_\rho(z) = \frac{2\pi i}{4\pi i} \sum_{i=0}^1 \operatorname{Res} \left(\underbrace{\frac{(\zeta^2 - 1)^2}{\sigma \zeta^2 (\zeta - \zeta_1)(\zeta - \zeta_2)}}_{f(\zeta)}, \zeta_i \right). \quad (2.7)$$

Calculons maintenant ces deux résidus.

$i = 0$ Comme ζ_0 est un pôle d'ordre 2, on a que

$$\begin{aligned} \operatorname{Res}(f, \zeta_0) &= \frac{d}{d\zeta} (\zeta - \zeta_0)^2 \frac{(\zeta^2 - 1)^2}{\zeta^2 (\sigma \zeta^2 - \zeta z + \sigma)} \Big|_{\zeta_0}, \\ &= \frac{4\zeta(\zeta^2 - 1)(\sigma \zeta^2 - \zeta z + \sigma) - (2\sigma \zeta - z)(\zeta^2 - 1)^2}{(\sigma \zeta^2 - \zeta z + \sigma)^2} \Big|_{\zeta_0}. \end{aligned}$$

Puisque $\zeta_0 = 0$, on obtient

$$\operatorname{Res}(f, \zeta_0) = \frac{z}{\sigma^2}.$$

$i = 1$ Comme ζ_1 est un pôle d'ordre 1, on a que

$$\operatorname{Res}(f, \zeta_1) = \lim_{\zeta \rightarrow \zeta_1} (\zeta - \zeta_1) \frac{(\zeta^2 - 1)^2}{\sigma \zeta^2 (\zeta - \zeta_1)(\zeta - \zeta_2)} = \frac{(\zeta_1^2 - 1)^2}{\sigma \zeta_1^2 (\zeta_1 - \zeta_2)} = \frac{(\zeta_1^2 - 1)^2}{\sigma \zeta_1 (\zeta_1^2 - \zeta_1 \zeta_2)}.$$

Puisque $\zeta_1 \zeta_2 = 1$, on a

$$\operatorname{Res}(f, \zeta_1) = \frac{\zeta_1^2 - 1}{\sigma \zeta_1} = \frac{\zeta_1 - \zeta_2}{\sigma} = -\frac{\sqrt{z^2 - 4\sigma^2}}{\sigma^2}.$$

En remplaçant les résidus obtenus dans l'équation (2.7), on obtient que

$$S_\rho(z) = \frac{z - \sqrt{z^2 - 4\sigma^2}}{2\sigma^2}. \quad (2.8)$$

Gardons en tête l'équation (2.8). La transformée de Stieltjes de la loi du demi-cercle sera utile pour la démonstration de la loi du demi-cercle.

3 Résultats classiques sur les matrices aléatoires

Après cette section couvrant les préliminaires, on a maintenant tous les outils nécessaires afin de présenter la loi du demi-cercle, la loi de Marchenko-Pastur et de donner une démonstration de la loi du demi-cercle. La loi du demi-cercle a d'abord été démontrée par Wigner en 1955. Comparativement à l'approche suivie dans la démonstration de ce rapport, Wigner a utilisé une approche de type combinatoire afin de démontrer la loi du demi-cercle.

3.1 Convergence de la mesure spectrale pour des matrices de Wigner

Voici enfin le moment tant attendu. Énonçons la loi du demi-cercle.

Théorème 3.1 (Loi du demi-cercle). Soit X_N une matrice de Wigner dont la variance des variables aléatoires en dehors de la diagonale principale est σ^2 . Alors la mesure spectrale de X_N converge faiblement vers la loi du demi-cercle de paramètre $\sigma = \sqrt{\sigma^2}$. En d'autres termes, pour toute fonction continue et bornée f , on a

$$\lim_{N \rightarrow \infty} \left(\int_{\mathbb{R}} f dL_N - \int_{\mathbb{R}} f d\rho \right) = 0.$$

En particulier, dans le cas $f \equiv 1$, on déduit du théorème que pour une matrice de Wigner X_N et $N \rightarrow \infty$, on a que pour $[s, t] \subseteq [-2\sigma, 2\sigma]$,

$$\frac{\text{Nombre de valeurs propres de } X_N \text{ dans } [s, t]}{N} = L_N([s, t]) \rightarrow \int_s^t \frac{\sqrt{4\sigma^2 - x^2}}{2\pi\sigma^2} dx. \quad (3.1)$$

La proportion de valeurs propres de X_N dans l'intervalle $[s, t]$ converge vers l'aire sous la fonction de densité de la loi du demi-cercle de paramètre σ pour l'intervalle $[s, t]$. La figure 5 donne une interprétation visuelle de l'équation (3.1). En effet, le membre de gauche de l'équation (3.1) converge vers l'aire de la région bleue dans la figure 5.

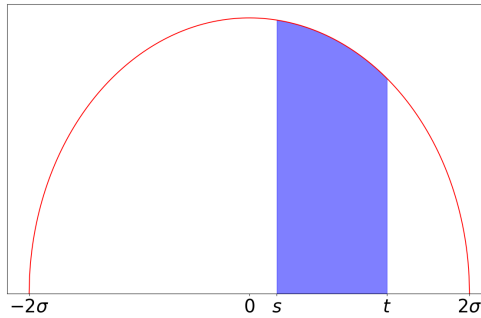


FIGURE 5 – Aire sous la fonction de densité de la loi du demi-cercle de paramètre σ dans l'intervalle $[s, t]$.

Donnons maintenant la démonstration du théorème. Puisque la démonstration du cas général est très complexe et demande des notions très avancées, on donne la démonstration dans le cas d'une matrice de Wigner issue de l'EGO $_{\sigma^2}$. Ce restreindre à ce cas fait intervenir diverses notions introduites dans la section 2, comme l'invariance par rotation et la transformée de Stieltjes.

3.1.1 Démonstration de la loi du demi-cercle

L'idée derrière la démonstration est la suivante : on va démontrer la convergence de la mesure spectrale L_N en utilisant la convergence de sa transformée de Stieltjes lorsque $N \rightarrow \infty$. En effet, si l'on montre que la transformée de Stieltjes de la mesure spectrale converge vers une fonction déterministe et que celle-ci est exactement la transformée de Stieltjes de la loi du demi-cercle, alors le résultat sera démontré grâce au théorème 2.3. Voici le plan de la démonstration :

1. Trouver une limite qui ne dépend pas de N à la transformée de Stieltjes de L_N . On montre aussi que la transformée de Stieltjes converge vers cette limite déterministe qu'on note $\mathfrak{s}_{L_N}(z)$. Ici, l'indice N sert seulement de notation. On verra dans l'étape 2 de la démonstration pourquoi on conserve un tel indice.
2. On montre ensuite que $\mathfrak{s}_{L_N}(z)$ satisfait une équation fonctionnelle, ce qui donne une forme close pour $\mathfrak{s}_{L_N}(z)$.
3. Comme on a déjà calculé la transformée de Stieltjes pour la loi du demi-cercle à l'exemple 2.5, il reste à conclure que $\mathfrak{s}_{L_N}(z)$ correspond à la transformée de Stieltjes de L_N . Ce qui permet de faire appel au 2.3 et de conclure.

Allons-y pour la démonstration du théorème 3.1 !

Démonstration de la loi du demi-cercle.

Étape 1.

On se rappelle que, par l'équation (2.5), la transformée de Stieltjes de la mesure spectrale de X_N est donnée par

$$S_{L_N}(z) = \frac{1}{N} \text{Tr}(G_{X_N}(z)),$$

pour tout $z \in \mathbb{C} \setminus \mathbb{R}$. À l'aide d'un développement de Taylor, on trouve une limite déterministe pour $S_{L_N}(z)$. Voyons comment y arriver. On a

$$\begin{aligned} S_{L_N}(z) &= \frac{1}{N} \text{Tr}(G_{X_N}(z)), \\ &= \frac{1}{N} \sum_{i=1}^N \frac{1}{z - \lambda_i^N}, \\ &= \frac{1}{zN} \sum_{i=1}^N \frac{1}{1 - \frac{\lambda_i^N}{z}}, \\ &= \frac{1}{zN} \sum_{i=1}^N \sum_{k=0}^{\infty} \left(\frac{\lambda_i^N}{z} \right)^k, \\ &= \frac{1}{zN} \sum_{k=0}^{\infty} \sum_{i=1}^N \left(\frac{\lambda_i^N}{z} \right)^k, \\ &= \sum_{k=0}^{\infty} \frac{1}{z^{k+1}} \frac{1}{N} \text{Tr}(X_N^k). \end{aligned}$$

La quatrième égalité se justifie par le fait qu'en prenant z tel que $|z|$ est assez grand, on a $\left| \frac{\lambda_i^N}{z} \right| < 1$ et la série de Taylor converge. La dernière égalité se justifie par le fait que la trace d'une matrice A correspond à la somme de ses valeurs propres et que, pour $j \in \mathbb{N}$, les valeurs propres de A^j sont simplement les valeurs propres de A à la puissance j . Par la loi des grands nombres, on a que

$$\frac{1}{N} \text{Tr}(X_N^k) = \frac{1}{N} \mathbb{E} [\text{Tr}(X_N^k)] = \tau_k(X_N), \quad (3.2)$$

lorsque N est grand. On a donc

$$\begin{aligned}\lim_{N \rightarrow \infty} S_{L_N}(z) &= \lim_{N \rightarrow \infty} \sum_{k=0}^{\infty} \frac{1}{z^{k+1}} \frac{1}{N} \text{Tr}(X_N^k), \\ &= \sum_{k=0}^{\infty} \frac{1}{z^{k+1}} \lim_{N \rightarrow \infty} \frac{1}{N} \text{Tr}(X_N^k), \\ &= \sum_{k=0}^{\infty} \frac{1}{z^{k+1}} \tau_k(X_N) =: \mathfrak{s}_{L_N}(z).\end{aligned}$$

La troisième égalité découle de l'équation (3.2). De ce calcul, on obtient un résultat très fort. On obtient que lorsque $N \rightarrow \infty$, la transformée de Stieltjes de la mesure spectrale converge vers $\mathfrak{s}_{L_N}(z)$. Ceci complète alors l'étape 1.

Étape 2.

On cherche maintenant à trouver une forme close pour $\mathfrak{s}_{L_N}(z) = \sum_{k=0}^{\infty} \frac{1}{z^{k+1}} \tau_k(X_N)$. Pour ce faire, l'idée est d'utiliser la résolvante de X_N et de faire appel au théorème 2.2. L'utilisation de la résolvante va permettre de trouver une équation fonctionnelle pour $\mathfrak{s}_{L_N}(z)$ et donc de passer à l'étape suivante. Posons

$$M = (G_{X_N}(z))^{-1} = zI_N - X_N,$$

on a donc

$$M = \begin{pmatrix} M_{1,1} & \cdots & M_{1,N} \\ \vdots & \ddots & \vdots \\ M_{1,N} & \cdots & M_{N,N} \end{pmatrix} = \begin{pmatrix} z & & 0 \\ & \ddots & \\ 0 & & z \end{pmatrix} - \begin{pmatrix} X_{1,1} & \cdots & X_{1,N} \\ \vdots & \ddots & \vdots \\ X_{1,N} & \cdots & X_{N,N} \end{pmatrix} = zI_N - X_N.$$

Découpons M comme

$$M = \begin{pmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{pmatrix},$$

où le bloc $B_{1,1}$ est de dimension 1×1 et le bloc $B_{2,2}$ est de dimension $(N-1) \times (N-1)$. Par le théorème 2.2, on a

$$[(G_{X_N}(z))_{1,1}]^{-1} = B_{1,1} - B_{1,2} B_{2,2}^{-1} B_{2,1}. \quad (3.3)$$

En explicitant l'équation (3.3), on a

$$\frac{1}{(G_{X_N}(z))_{1,1}} = M_{1,1} - \sum_{k=2}^N \sum_{\ell=2}^N M_{1,k} (B_{2,2}^{-1})_{k,\ell} M_{\ell,1}. \quad (3.4)$$

On s'intéresse maintenant à l'espérance de l'équation (3.4). Par construction de M , on a $\mathbb{E}[M_{1,1}] = z$. De plus, les entrées de $B_{2,2}$ sont indépendantes des entrées $M_{1,i}$ et $M_{i,1}$ pour $1 \leq i \leq N$. En prenant l'espérance

sur l'équation (3.4), on obtient

$$\begin{aligned}
\mathbb{E} \left[\frac{1}{(G_{X_N}(z))_{1,1}} \right] &= \mathbb{E} \left[M_{1,1} - \sum_{k=2}^N \sum_{\ell=2}^N M_{1,k} (B_{2,2}^{-1})_{k,\ell} M_{\ell,1} \right], \\
&= \mathbb{E} [M_{1,1}] - \mathbb{E} \left[\sum_{k=2}^N \sum_{\ell=2}^N M_{1,k} (B_{2,2}^{-1})_{k,\ell} M_{\ell,1} \right], \\
&= z - \sum_{k=2}^N \sum_{\ell=2}^N \mathbb{E} [M_{1,k} (B_{2,2}^{-1})_{k,\ell} M_{\ell,1}], \\
&= z - \sum_{k=2}^N \sum_{\ell=2}^N \mathbb{E} [(B_{2,2}^{-1})_{k,\ell}] \mathbb{E} [M_{1,k} M_{\ell,1}], \\
&= z - \sum_{k=2}^N \sum_{\ell=2}^N \mathbb{E} [(B_{2,2}^{-1})_{k,\ell}] \text{Cov} [M_{1,k}, M_{\ell,1}], \\
&= z - \sum_{k=2}^N \sum_{\ell=2}^N \mathbb{E} [(B_{2,2}^{-1})_{k,\ell}] \frac{\sigma^2}{N} \delta_{k\ell}, \\
&= z - \sigma^2 \mathbb{E} \left[\frac{1}{N} \text{Tr}(B_{2,2}^{-1}) \right].
\end{aligned} \tag{3.5}$$

On observe que $\frac{1}{N-1} \text{Tr}(B_{2,2}^{-1})$ est la transformée de Stieltjes de la sous-matrice de X_N de dimension $(N-1) \times (N-1)$. On s'attend à ce que la variation entre la transformée de Stieltjes d'une matrice de Wigner de dimension $N \times N$ et celle d'une matrice de Wigner de dimension $(N-1) \times (N-1)$ soit petite lorsque $N \rightarrow \infty$. En d'autres termes, quand $N \rightarrow \infty$, on a l'égalité $S_{X_N}(z) = S_{X_{N-1}}(z)$ puisque le terme dont on prend l'espérance dans l'équation (3.6) est très près de $\frac{1}{N-1} \text{Tr}(B_{2,2}^{-1})$ pour N grand, on a donc

$$\frac{1}{N} \text{Tr}(B_{2,2}^{-1}) = \frac{1}{N-1} \text{Tr}(B_{2,2}^{-1}).$$

Comme $B_{2,2}^{-1}$ est la résolvante d'une matrice de Wigner $(N-1) \times (N-1)$, le membre de droite de l'équation (3.6) devient

$$z - \sigma^2 \mathfrak{s}_{X_N}(z). \tag{3.7}$$

Il reste maintenant à trouver de quelle manière on peut écrire $\mathbb{E} \left[\frac{1}{(G_{X_N}(z))_{1,1}} \right]$. En fait, on a l'approximation

$$\mathbb{E} \left[\frac{1}{(G_{X_N}(z))_{1,1}} \right] \approx \frac{1}{\mathbb{E} [(G_{X_N}(z))_{1,1}]}, \tag{3.8}$$

qui devient très près d'une égalité pour N grand. Ouvrons une parenthèse pour justifier l'approximation de l'équation (3.8). Si une variable aléatoire X se concentre fortement autour de son espérance, alors pour une fonction convexe f , on a l'approximation

$$\mathbb{E} [f(X)] \approx f(\mathbb{E} [X]).$$

On applique ce résultat à la fonction convexe $f(x) = 1/x$, qui est convexe pour des valeurs de x positives. Sous l'hypothèse que $(G_{X_N}(z))_{1,1}$ est concentré autour de son espérance et prend principalement des valeurs positives, on obtient

$$\mathbb{E} \left[\frac{1}{(G_{X_N}(z))_{1,1}} \right] \approx \frac{1}{\mathbb{E} [(G_{X_N}(z))_{1,1}]}. \tag{3.8}$$

En supposant des petites variations, on obtient l'égalité de l'équation (3.8). Fermons la parenthèse et retournons à la démonstration. Comme la résolvante est invariante par rotation et que, par la proposition 2.2, les entrées sur la diagonale de la résolvante ont la même espérance. Donc

$$\mathbb{E} [(G_{X_N}(z))_{1,1}] = \frac{1}{N} \mathbb{E} [\text{Tr}(G_{X_N}(z))] = \mathfrak{s}_{L_N}(z).$$

En se rappelant ce qui a été fait, l'équation (3.5) devient

$$\frac{1}{\mathfrak{s}_{L_N}(z)} = z - \sigma^2 \mathfrak{s}_{L_N}(z).$$

On obtient par la suite une expression quadratique en $\mathfrak{s}_{L_N}(z)$ pour tout $z \in \mathbb{C} \setminus \mathbb{R}$:

$$\sigma^2 (\mathfrak{s}_{L_N}(z))^2 - z \mathfrak{s}_{L_N}(z) + 1 = 0,$$

d'où l'on tire

$$\mathfrak{s}_{L_N}(z) = \frac{z \pm \sqrt{z^2 - 4\sigma^2}}{2\sigma^2},$$

à l'aide de la formule pour trouver les racines d'une expression quadratique. On garde le signe $-$ du \pm . Une telle justification est faite en détail dans [6]. La forme close de la transformée de Stieltjes de X_N est donc

$$\mathfrak{s}_{L_N}(z) = \frac{z - \sqrt{z^2 - 4\sigma^2}}{2\sigma^2}. \quad (3.9)$$

Étape 3.

De l'étape 2, on sait que $S_{L_N}(z)$ converge vers $\mathfrak{s}_{L_N}(z) = \frac{z - \sqrt{z^2 - 4\sigma^2}}{2\sigma^2}$. De l'exemple 2.5, on sait que la transformée de Stieltjes de la loi du demi-cercle est $\frac{z - \sqrt{z^2 - 4\sigma^2}}{2\sigma^2}$. La transformée de Stieltjes de la mesure spectrale $S_{L_N}(z)$ converge ainsi vers la transformée de Stieltjes de la loi du demi-cercle. Ainsi, par le théorème 2.3, on conclut que L_N converge faiblement vers la loi du demi-cercle. Ce dernier résultat met fin à la démonstration.



3.2 Convergence de la mesure spectrale pour des matrices de Wishart

Passons maintenant à un théorème qui caractérise la mesure spectrale d'une matrice de Wishart lorsque $N \rightarrow \infty$. Dans le cas d'une matrice de Wishart, la mesure spectrale correspond à la même mesure spectrale que la mesure spectrale d'une matrice de Wigner, puisque la définition 2.10 est valide pour toute matrice aléatoire symétrique.

Une première observation quant aux valeurs propres d'une matrice de Wishart découle de la proposition suivante.

Proposition 3.1. Soit W_N une matrice de Wishart. Alors W_N est semi-définie positive.

Démonstration. On se rappelle qu'une matrice symétrique M de dimension $N \times N$ est semi-définie positive si et seulement si $v^t M v \geq 0$ pour tout $v \in \mathbb{R}^n$. C'est cette caractérisation que l'on va exploiter. Soit $v \in \mathbb{R}^N$. On a

$$v^t W_N v = v^t \frac{1}{p} B B^t v = \frac{1}{p} v^t B B^t v = \frac{1}{p} (B^t v)^t (B^t v) = \frac{1}{p} \|B^t v\|^2 \geq 0.$$



Le fait que W_N soit semi-définie positive donne des informations sur les valeurs propres de W_N puisqu'une matrice symétrique dont toutes les entrées sont réelles est semi-définie positive si et seulement si toutes ses valeurs propres sont non négatives. Comme les valeurs propres d'une matrice de Wishart sont non négatives, la mesure spectrale ne peut pas converger vers la loi du demi-cercle, parce que le support de la loi du demi-cercle admet des valeurs négatives. Il faut donc trouver une autre loi vers laquelle la mesure spectrale converge. C'est ce qui est fait dans le théorème suivant.

Théorème 3.2 (Loi de Marchenko-Pastur). Soient $\lambda_1^N, \lambda_2^N, \dots, \lambda_N^N$ les valeurs propres de $W_N = \frac{1}{p} B B^t$, une matrice de Wishart où B est une matrice $N \times p$. Supposons que $N \rightarrow \infty$ et $p \rightarrow \infty$ avec $N/p = \gamma$. Alors la mesure spectrale de W_N converge faiblement vers la loi de Marchenko-Pastur de paramètre γ .

Démonstration. La démonstration est donnée dans [2]. ✎

Ce théorème complète la discussion sur les résultats en lien avec les matrices de Wishart. Dans la section 4, on verra des simulations numériques pour illustrer le théorème 3.2.

En particulier, dans le cas $f \equiv 1$, on déduit du théorème que pour une matrice de Wishart W_N et $N \rightarrow \infty$ avec $\gamma = N/p$, on a que pour $[s, t] \subseteq [\gamma_-, \gamma_+]$,

$$\frac{\text{Nombre de valeurs propres de } W_N \text{ dans } [s, t]}{N} = L_N([s, t]) \rightarrow \int_s^t \frac{\sqrt{(\gamma_+ - x)(x - \gamma_-)}}{2\pi\gamma x} dx. \quad (3.10)$$

La proportion de valeurs propres de W_N dans l'intervalle $[s, t]$ converge vers l'aire sous la fonction de densité de la loi de Marchenko-Pastur de paramètre γ pour l'intervalle $[s, t]$. La figure 6 donne une interprétation visuelle de l'équation (3.10) lorsque $\gamma = 1/3$. En effet, le membre de gauche de l'équation (3.10) converge vers l'aire de la région bleue dans la figure 6.

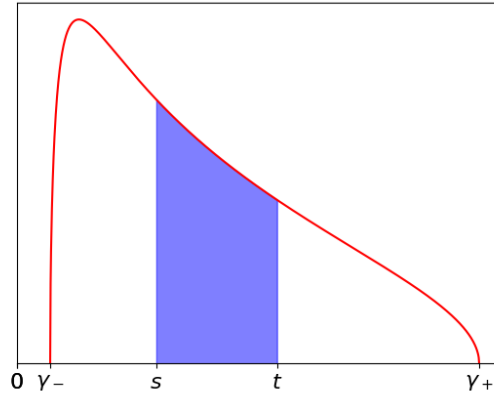


FIGURE 6 – Aire sous la fonction de densité de la loi de Marchenko-Pastur de paramètre $\gamma = 1/3$ dans l'intervalle $[s, t]$.

Dans cette section, on a énoncé les deux théorèmes importants. On a aussi été en mesure de donner une jolie interprétation de ces théorèmes à l'aide de figures. Ceci met fin à cette section. Dans la prochaine section, on verra la loi du demi-cercle et la loi de Marchenko-Pastur émerger lorsque la taille de la matrice augmente.

3.3 Universalité pour la convergence de la mesure spectrale

Introduisons maintenant le terme *universalité*. On parle d'universalité lorsque les résultats obtenus ne dépendent pas de la loi de probabilité utilisée. Plus précisément, ce qu'on veut dire, c'est que la loi du demi-cercle n'est pas vraie que pour une matrice de Wigner issue de l'EGO $_{\sigma^2}$, mais pour toute matrice de Wigner dont l'espérance des variables aléatoires en entrées est nulle et dont tous les moments sont finis. La loi du demi-cercle n'impose pas non plus que les variables aléatoires soient continues ou discrètes, ce qui donne beaucoup de liberté dans le choix de variables aléatoires.

Pour une matrice de Wigner, on pourrait aussi bien prendre des variables aléatoires suivant une loi uniforme sur l'intervalle $[-1, 1]$, une loi uniforme discrète sur l'ensemble $\{-2, -1, 0, 1, 2\}$ ou encore une loi normale centrée.

On vient de faire un commentaire sur l'universalité pour la loi du demi-cercle, mais on observe aussi le même phénomène pour la loi de Marchenko-Pastur. Il suffit que les variables aléatoires en entrées soient centrées et de variance 1, ce qu'on peut toujours faire en multipliant les variables aléatoires par un facteur approprié.

Maintenant que l'universalité est introduite, on pourra voir dans la section suivante une démonstration de ce phénomène.

4 Simulations

4.1 Loi du demi-cercle

Dans cette section, on s'intéressera à des simulations de la loi du demi-cercle et de la loi de Marchenko-Pastur. Comme mentionné après avoir énoncé les deux théorèmes dans la section 3, on a fourni une interprétation lorsque $f \equiv 1$. C'est ce qu'on va exploiter dans cette section. Lorsque $f \equiv 1$, on doit s'attendre à ce que la région couverte par les bandes de l'histogramme corresponde à l'aire sous la fonction de densité théorique qui est soit la loi du demi-cercle, soit la loi de Marchenko-Pastur.

4.1.1 Cas classique

Une des plus simples matrices de Wigner est une matrice dont les entrées suivent une loi de Rademacher. La fonction de masse de cette loi de probabilité est définie par

$$p(k) = \begin{cases} \frac{1}{2} & \text{si } k = -1, \\ \frac{1}{2} & \text{si } k = 1, \\ 0 & \text{sinon.} \end{cases}$$

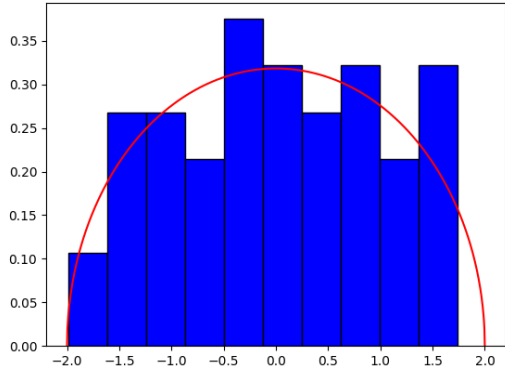
En résumé, une variable aléatoire qui suit une loi de Rademacher retourne 1 ou -1 avec une probabilité $1/2$ pour les deux valeurs. Voici un exemple de matrice de Wigner 5×5 dont les entrées suivent une loi de Rademacher

$$\frac{1}{\sqrt{5}} \begin{pmatrix} -1 & -1 & 1 & 1 & 1 \\ -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 \end{pmatrix}.$$

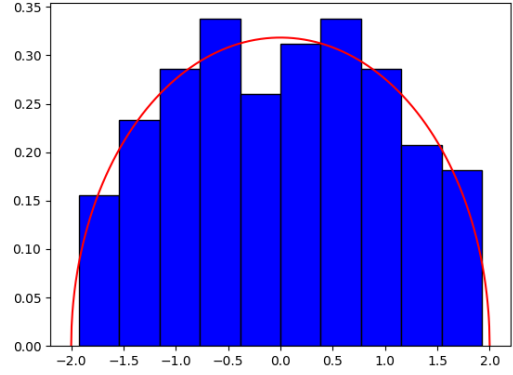
Si X est une variable aléatoire qui suit une loi de Rademacher, alors on a $\mathbb{E}[X] = 0$ et $\text{Var}[X] = 1$. Donc l'histogramme des valeurs propres lorsque $N \rightarrow \infty$ devrait ressembler à la fonction de densité $f(t) = \frac{\sqrt{4-t^2}}{2\pi}$. Voyons comment la loi du demi-cercle émerge lorsque la taille d'une matrice de Wigner de Rademacher augmente. À la figure 7, lorsque N augmente, on voit que l'histogramme des valeurs propres se rapproche de plus en plus de la densité attendue qui est tracée en rouge. Pour toutes les simulations de la section 4.1, la courbe en rouge représente toujours la loi du demi-cercle de paramètre σ où $\sigma = \sqrt{\sigma^2}$ et σ^2 est la variance des entrées de la matrice de Wigner en dehors de la diagonale principale.

On constate tout de suite que l'histogramme « colle » assez bien à ce qui est attendu. Bien que l'histogramme ne compte que 10 boîtes, ces boîtes donnent à l'histogramme une forme assez près de celle attendue.

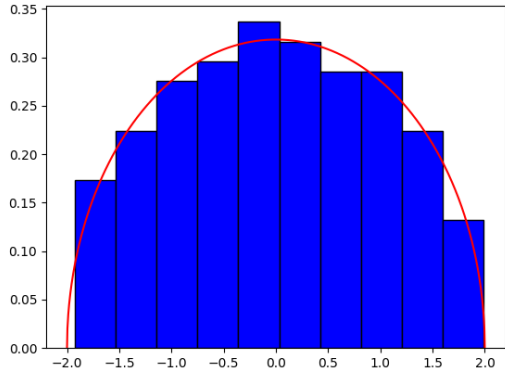
Voyons aussi quelle est la forme de l'histogramme des valeurs propres lorsque $N = 10\,000$. Dans la figure 8, le cas $N = 10\,000$ est tracé pour deux matrices de Wigner dont les entrées suivent une loi de Rademacher dans un cas et une loi normale dans l'autre.



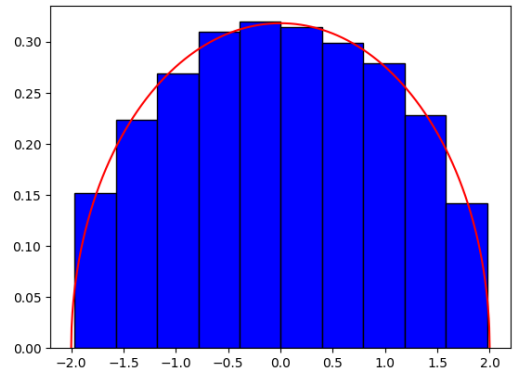
(a) Matrice 50×50 .



(b) Matrice 100×100 .

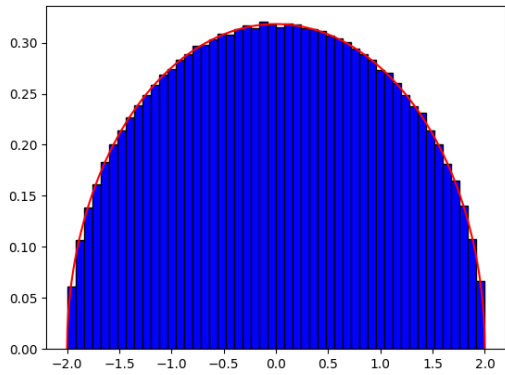


(c) Matrice 250×250 .

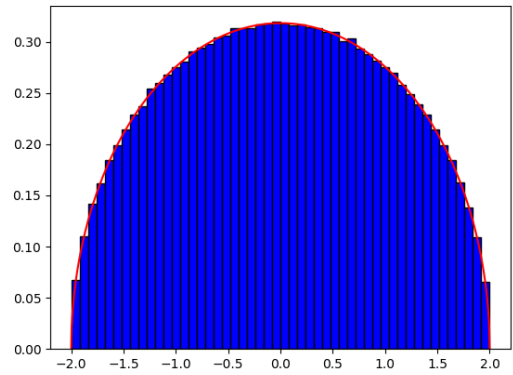


(d) Matrice 500×500 .

FIGURE 7 – Histogrammes des valeurs propres d'une matrice de Wigner de Rademacher pour des matrices de différentes tailles.



(a) Matrice de Wigner de Rademacher.



(b) Matrice de Wigner issue de l'EGO₁.

FIGURE 8 – Histogramme des valeurs propres pour deux matrices de Wigner $10\,000 \times 10\,000$.

Comme on s'y attendait, la convergence ne dépend pas de la loi de probabilité que suivent les variables aléatoires. C'est ce qui découle du théorème 3.1, puisqu'on n'impose pas que les entrées de la matrice de Wigner suivent une loi de probabilité particulière. On se réfère à la discussion de la section 3.3 sur l'universalité pour plus de détails.

4.1.2 Trois cas non couverts par la loi du demi-cercle

Cas 1 (familles de variables aléatoires non i. i. d.) :

Soyons fous, créons une matrice aléatoire symétrique dont les entrées suivent soit une loi normale centrée et réduite, soit une loi de Rademacher. La probabilité que l'entrée i, j de la matrice de Wigner suive une loi normale est $\frac{1}{2}$. Par complémentarité, la probabilité que l'entrée i, j suive une loi de Rademacher est elle aussi $\frac{1}{2}$. Pour $N = 10\,000$, l'histogramme des valeurs propres d'une telle matrice est donné à la figure 9.

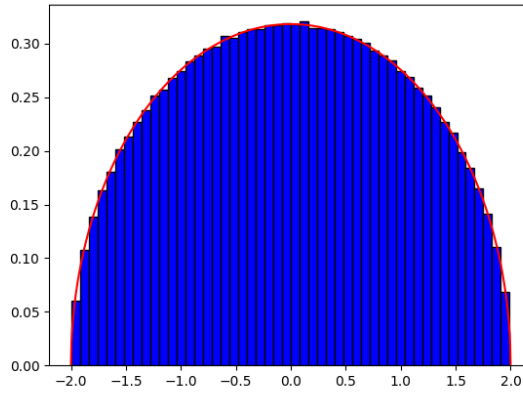


FIGURE 9 – Histogramme des valeurs propres pour une matrice de Wigner dont la loi des entrées est choisie aléatoirement entre une loi normale centrée et réduite et une loi de Rademacher.

Bien qu'une telle matrice ne satisfasse pas les hypothèses du théorème 3.1 puisque les entrées sur la diagonale ne sont plus identiquement distribuées, on observe tout de même que la mesure spectrale d'une telle matrice aléatoire semble converger vers la loi du demi-cercle.

Cas 2 (variables aléatoires non centrées sur la diagonale) :

Depuis le début du rapport, on travaille avec des lois de probabilité d'espérance nulle. Cependant, même si l'espérance des variables aléatoires sur la diagonale principale est m au lieu de 0, le théorème 3.1 s'applique, mais à un décalage près. Voyons ce qu'on veut dire par décalage. Soit X_N une matrice de Wigner dont les variables aléatoires sur la diagonale principale ont une espérance $m \neq 0$. Alors

$$X_N = X_N - \frac{m}{\sqrt{N}}I_N + \frac{m}{\sqrt{N}}I_N = \underbrace{X_N - \frac{m}{\sqrt{N}}I_N}_{=: Y_N} + \frac{m}{\sqrt{N}}I_N.$$

Dans ce cas, Y_N est une matrice de Wigner puisque $\mathbb{E}[(Y_N)_{i,i}] = \mathbb{E}\left[(X_N)_{i,i} - \frac{m}{\sqrt{N}}\right] = \mathbb{E}[(X_N)_{i,i}] - \frac{m}{\sqrt{N}} = \frac{m}{\sqrt{N}} - \frac{m}{\sqrt{N}} = 0$. Comme on s'intéresse à la mesure spectrale de X_N , on doit trouver les valeurs propres de X_N . Puisque Y_N est une matrice de Wigner et que $X_N = Y_N + \frac{m}{\sqrt{N}}I_N$, est-il possible de déterminer la mesure spectrale de X_N en sachant que la mesure spectrale de Y_N converge vers la loi du demi-cercle lorsque $N \rightarrow \infty$? Rappelons-nous que les valeurs propres de X_N sont les valeurs $\lambda \in \mathbb{R}$ telles que $\det(X_N - \lambda I_N) = 0$.

On affirme que les valeurs propres de X_N sont entièrement déterminées par celles de Y_N . En effet,

$$\begin{aligned}
0 &= \det(X_N - \lambda I_N), \\
\iff 0 &= \det\left(Y_N + \frac{m}{\sqrt{N}} I_N - \lambda I_N\right), \\
\iff 0 &= \det\left(Y_N - \underbrace{\left(\lambda - \frac{m}{\sqrt{N}}\right)}_{=: \hat{\lambda}} I_N\right), \\
\iff 0 &= \det(Y_N - \hat{\lambda} I_N).
\end{aligned}$$

De la dernière équivalence, il découle que les $\hat{\lambda}$ sont les valeurs propres de Y_N . Les valeurs propres λ de X_N satisfont donc l'égalité $\lambda = \hat{\lambda} + \frac{m}{\sqrt{N}}$ et sont entièrement déterminées par les valeurs propres $\hat{\lambda}$ de Y_N . Puisque les valeurs propres de X_N sont les valeurs propres de Y_N auxquelles on ajoute $\frac{m}{\sqrt{N}}$, la mesure spectrale de X_N correspond à celle de Y_N avec un décalage de $\frac{m}{\sqrt{N}}$. La mesure spectrale converge donc vers la loi du demi-cercle décalée de $\frac{m}{\sqrt{N}}$ dont la densité est donnée par

$$\hat{\rho}(x) = \frac{\sqrt{4\sigma^2 - (x - \frac{m}{\sqrt{N}})^2}}{2\pi\sigma^2} \quad \text{pour } \frac{m}{\sqrt{N}} - 2\sigma < x < \frac{m}{\sqrt{N}} + 2\sigma. \quad (4.1)$$

Puisque m est un nombre fini et que, dans l'énoncé du théorème 3.1 $N \rightarrow \infty$, alors le quotient $\frac{m}{\sqrt{N}} \rightarrow 0$ lorsque $N \rightarrow \infty$. Ainsi, on observe le même phénomène que dans le cas où l'espérance des variables aléatoires sur la diagonale principale est 0 et la densité donnée à l'équation (4.1) correspond à celle de la loi du demi-cercle lorsque $N \rightarrow \infty$. Illustrons ce phénomène avec une simulation dont la matrice pour laquelle nous cherchons à déterminer la mesure spectrale est une matrice de Wigner de l'EGO₁, mais dont l'espérance des variables aléatoires sur la diagonale principale est 1 au lieu de 0. L'histogramme des valeurs propres d'une telle matrice est donné à la figure 10.

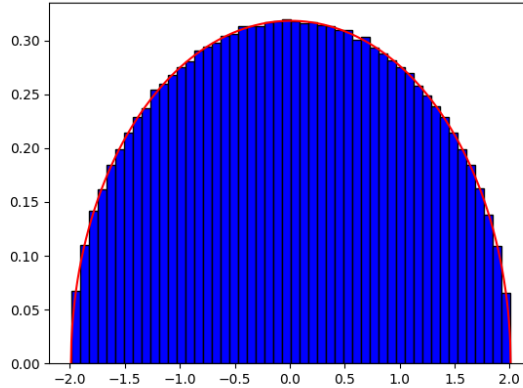


FIGURE 10 – Histogramme des valeurs propres pour une matrice aléatoire symétrique dont les entrées sur la diagonale principale suivent une loi normale de moyenne 1 et de variance 2.

On remarque bien évidemment qu'une telle matrice n'est pas une matrice de Wigner, puisqu'elle ne satisfait pas la définition d'une matrice de Wigner, soit la définition 2.3. Cependant, la démarche qu'on vient de faire nous assure bien que la mesure spectrale d'une telle matrice converge vers la loi du demi-cercle lorsque $N \rightarrow \infty$.

Cas 3 (variables aléatoires non centrées partout) :

Nous avons traité le cas où les entrées sur la diagonale principale d'une matrice aléatoire symétrique avaient

une espérance non nulle. Regardons maintenant ce qui se passe lorsque toutes les entrées d'une matrice aléatoire symétrique ont toutes une espérance non nulle. En fait, on n'a pas trouvé de résultat comme la loi du demi-cercle dans ce cas-là malheureusement. Cependant, la modification suivante permet de retomber sur le cas des matrices de Wigner.

Soit A_N une matrice $N \times N$ comme dans la définition 2.3, mais dont les familles de variables aléatoires $\{Z_{i,j}\}_{1 \leq i < j \leq N}$ et $\{Y_{i,i}\}_{1 \leq i \leq N}$ ont respectivement des espérances m_{hd} et m_{d} . Soit J_N la matrice $N \times N$ dont chaque entrée vaut 1. Alors la matrice

$$A_N - \frac{m_{\text{hd}}}{\sqrt{N}} J_N - \frac{m_{\text{d}} - m_{\text{hd}}}{\sqrt{N}} I_N, \quad (4.2)$$

est une matrice de Wigner. La loi du demi-cercle s'applique donc à cette matrice. Prenons par exemple une matrice aléatoire symétrique dont les entrées suivent une loi de Poisson de paramètre $\lambda = 2$. Notons par A_N cette matrice. Alors

$$X_N := A_N - 2J_N, \quad (4.3)$$

est une matrice de Wigner comme prévu par l'équation (4.2). Le théorème 3.1 nous garantit alors que la mesure spectrale de X_N converge vers la loi du demi-cercle lorsque $N \rightarrow \infty$. La figure illustre ce qui vient d'être dit pour une matrice X_N comme dans l'équation (4.3) avec $N = 10\,000$.

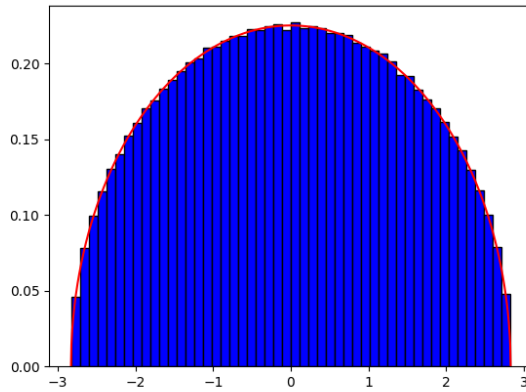


FIGURE 11 – Histogramme des valeurs propres pour une matrice X_N comme dans l'équation (4.3) avec $N = 10\,000$.

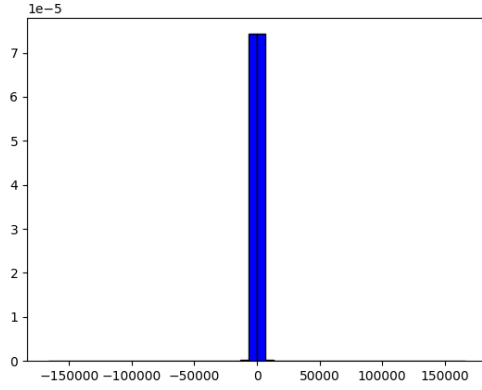
On constate bel et bien que la mesure spectrale converge vers la loi du demi-cercle lorsque N devient grand.

4.1.3 Loi du demi-cercle pour une variance infinie

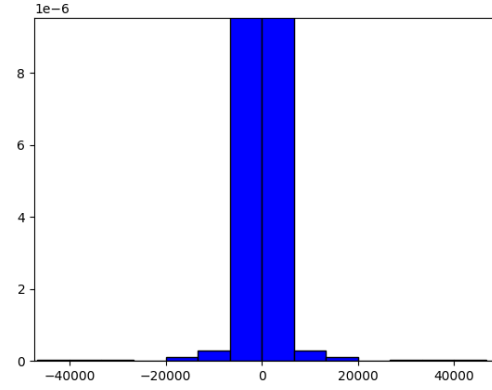
Du théorème 3.1, on sait que la loi du demi-cercle s'applique seulement lorsque la variance des familles de variables aléatoires de la définition 2.3 est bornée. Que se passe-t-il si la variance n'est plus bornée ? Voyons ce qui se produit avec une matrice comme dans la définition 2.3, mais dont les familles de variables aléatoires $\{Z_{i,j}\}_{1 \leq i < j \leq N}$ et $\{Y_{i,i}\}_{1 \leq i \leq N}$ suivent des lois de Cauchy standard, c'est-à-dire que la fonction de densité des variables aléatoires des deux familles est

$$f(x) = \frac{1}{\pi(1+x^2)} \quad \text{pour } x \in \mathbb{R}.$$

Dans ce cas, la variance des entrées de la matrice est infinie. Notons par A_N une telle matrice. À la figure 12, on voit clairement que l'histogramme des valeurs propres pour la matrice A_N avec $N = 10\,000$ ne converge pas vers un demi-cercle.



(a) Histogramme des valeurs propres de $A_{10\,000}$.



(b) Zoom sur l'histogramme des valeurs propres de $A_{10\,000}$.

FIGURE 12 – Histogramme des valeurs propres pour $A_{10\,000}$ sans zoom et avec zoom.

À la table 1, on constate que pour certains N , le module de la valeur propre ayant le plus grand module augmente lorsque N augmente pour une matrice de Wigner dont les entrées suivent une loi de Cauchy. On pourrait ainsi s'attendre à ce que pour N assez grand, le module de la valeur propre ayant le plus grand module devienne arbitrairement grand. L'hypothèse que la variance des entrées d'une matrice de Wigner est finie est donc essentielle afin que le théorème 3.1 soit valide pour toute matrice de Wigner.

N	$\max_i \{ \lambda_i \}$
50	139
100	119
250	3 884
500	2 748
1 000	47 627
10 000	167 002

TABLE 1 – Module de la valeur propre ayant le plus grand module pour certaines valeurs de N .

Dans cette présente section, on a exploré la loi du demi-cercle pour différentes lois de probabilités. On a aussi mis à l'épreuve certaines hypothèses nécessaires pour que la loi du demi-cercle soit valide. On conclut aussi que la variance finie pour les entrées de la matrice de Wigner est essentielle comme on vient de le voir. On pourrait pousser la réflexion un peu plus loin, jusqu'à supposer que seulement le deuxième moment des entrées de la matrice soit borné et, dans ce cas, le théorème 3.1 tient encore tel que mentionné dans [1]. Cependant, on ne s'attarde pas sur ce cas dans ce rapport. C'est ainsi que l'on conclut cette section sur les simulations de la loi du demi-cercle.

4.2 Loi de Marchenko-Pastur

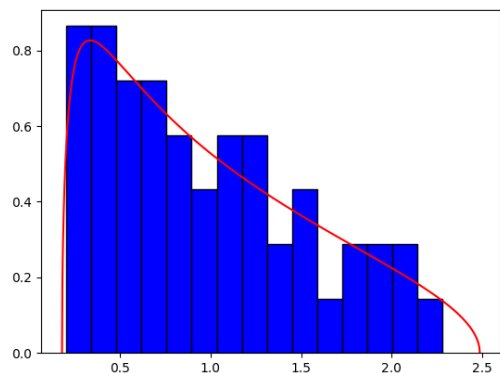
Passons maintenant aux simulations de la loi de Marchenko-Pastur. Pour ces simulations, on considérera que $\gamma = 1/3$. Plus précisément, la matrice B qui servira à créer la matrice de Wishart sera une matrice $1 \cdot m \times 3 \cdot m$ pour $m \in \mathbb{N}$. On a que $m \in \mathbb{N}$ et $m \rightarrow \infty$, on a donc bel et bien $\gamma = 1/3$ lorsque la dimension tend vers l'infini. On construit la matrice B pour que $B_{i,j} \sim \mathcal{N}(0, 1)$. Comme chacune des entrées a une espérance nulle et un écart type de 1, les hypothèses de la définition sont satisfaites et on crée une matrice de Wishart à l'aide de l'équation (2.1).

À la figure 13, on observe la convergence de l'histogramme des valeurs propres vers la loi de Marchenko-Pastur lorsque la taille de la matrice de Wishart augmente. Dans cette figure, la courbe en rouge représente

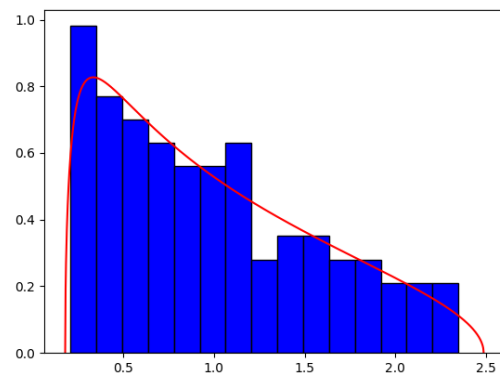
la loi de Marchenko-Pastur pour $\gamma = 1/3$. Dans ce cas, l'équation de la courbe rouge est donnée par

$$\frac{\sqrt{(\gamma_+ - x)(x - \gamma_-)}}{2\pi\gamma x} \quad \text{pour } \gamma_- < x < \gamma_+,$$

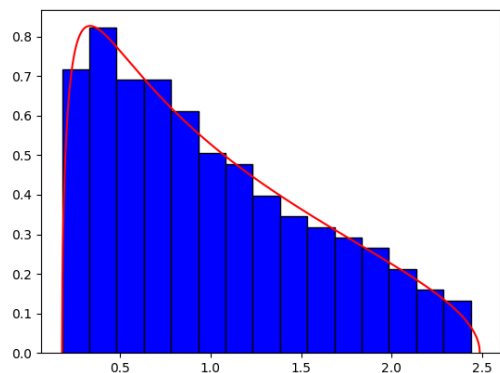
où $\gamma_{\pm} = (1 \pm \sqrt{\gamma})^2$.



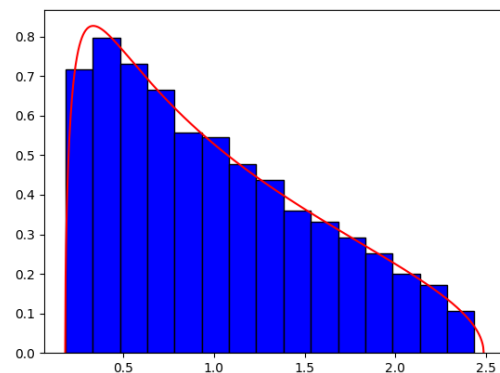
(a) Matrice 50×50 .



(b) Matrice 100×100 .



(c) Matrice 250×250 .



(d) Matrice 500×500 .

FIGURE 13 – Histogrammes des valeurs propres d'une matrice de Wishart dont chacune des entrées de la matrice B suit une loi normale de moyenne 0 et de variance 1.

Voyons maintenant comment se comporte l'histogramme des valeurs propres d'une matrice de Wishart de dimension $10\,000 \times 10\,000$. Comme précédemment, on prend $\gamma = 1/3$ et chaque entrée de B suit une loi normale de moyenne 0 et de variance 1. On retrouve cet histogramme à la figure 14.

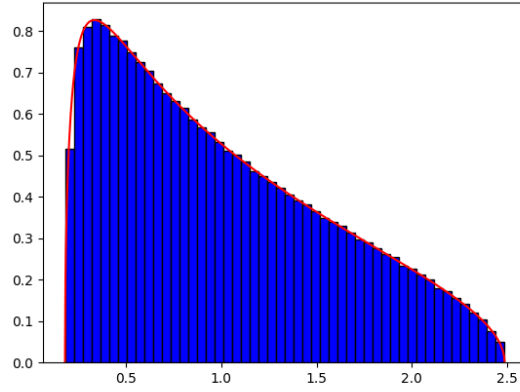


FIGURE 14 – Histogramme des valeurs propres d’une matrice de Wishart $10\,000 \times 10\,000$.

Dans ce cas, on constate que la dimension de la matrice de Wishart est assez grande pour que l’histogramme colle très bien à la densité attendue.

On aurait aussi pu travailler avec plusieurs autres valeurs de paramètre γ , mais dans tous les cas, on aurait retrouvé des résultats similaires. Comme dans la section 4.1, il aurait été possible de tester pourquoi les différentes hypothèses sont nécessaires afin que la mesure spectrale converge vers la loi de Marchenko-Pastur, mais cela aurait été redondant. On fait seulement un commentaire afin de mentionner qu’il est possible de faire des tests semblables en enlevant certaines hypothèses ou en les poussant à la limite.

5 Matrices aléatoires en finance

On veut maintenant appliquer la théorie des matrices aléatoires en finance. Pour ce faire, on considère certaines sociétés du S&P 500. Cet exemple d'application s'inspire de ce qui est fait dans [4].

On considère 488 sociétés du S&P 500. Pour ces sociétés, on s'intéresse à la variation quotidienne du prix de fermeture de leurs actions. On s'intéresse à tous les jours où la bourse a été ouverte du premier janvier 2020 au premier janvier 2025. C'est en tout 1258 jours, mais afin de considérer les variations, il n'y aura que 1257 variations pour cette période.

Ce qu'on va faire, c'est créer une matrice 488×1257 qu'on note B . Chaque ligne est associée à une société et chaque colonne est associée à une variation entre deux journées consécutives où la bourse est ouverte. Voyons un exemple pour 3 sociétés sur une période de 5 jours.

Exemple 5.1. On considère les sociétés Best-Buy, Costco et Walmart dont les indices sont respectivement BBY, COST et WMT. Les prix de fermeture pour les 3 sociétés du 31 mars au 4 avril 2025 se trouvent dans le tableau suivant :

	31 mars	1 ^{er} avril	2 avril	3 avril	4 avril
BBY	73.61	74.37	75.73	62.22	60.44
COST	945.78	954.40	965.08	967.08	916.48
WMT	87.79	88.83	89.76	87.26	83.19

On calcule les variations et l'on obtient

	1a – 31m	2a – 1a	3a – 2a	4a – 3a
BBY	0.76	1.36	−13.51	−1.78
COST	8.62	10.68	2.00	−50.60
WMT	1.04	0.93	−2.5	−4.07

On vient ensuite centrer et réduire chacune des lignes de la matrice 3×4 du dernier tableau. On calcule la moyenne et l'écart-type de chacune des lignes. Pour chaque ligne, on vient soustraire la moyenne associée à cette ligne et l'on divise par l'écart-type associé à la ligne. La nouvelle matrice 3×4 obtenue est alors la matrice \hat{B} (on verra ce qu'est la matrice \hat{B} sous peu) associée à la matrice de Wishart.

On fait cette même construction pour la matrice 488×1257 . L'idée exploitée est la suivante : on considère la matrice de Wishart associée à \hat{B} , c'est-à-dire la matrice $W_{488} = \frac{1}{1257} \hat{B} \hat{B}^t$. On se demande si la mesure spectrale d'une telle matrice converge, ou du moins s'approche de la loi de Marchenko-Pastur. Plus précisément, est-ce que le théorème 3.2 s'applique à W_{488} . Si ce n'est pas le cas, on pourra conclure qu'au moins une des hypothèses du théorème 3.2 n'est pas vérifiée.

Afin d'obtenir la matrice W_{488} , diverses étapes ont été réalisées. Le travail a été fait sur Python. Le code utilisé pour cette section se trouve dans l'annexe A. La librairie YFinance de l'étape 1 est une librairie qui permet d'aller récupérer de nombreuses données par rapport à la bourse. Voici maintenant les grandes étapes à réaliser afin d'obtenir la matrice désirée :

1. Récupérer les données des sociétés à l'aide de la librairie YFinance. On traite ensuite les données. Pour la période choisie, il y a certaines sociétés dont les données ne sont pas disponibles pour tous les jours de la période. Ces sociétés ont été retirées. Sur les 503 sociétés du S&P 500, seulement 488 ont été conservées pour faire l'analyse.
2. Créer une matrice dont chaque ligne correspond à une société et chaque colonne correspond au prix de fermeture pour chaque société à une journée fixée. Pour ce faire, on a utilisé la librairie Pandas.
3. Calculer les variations quotidiennes des prix de fermeture pour les différentes journées de la période. La matrice des variations est la matrice B introduite plus tôt.

4. Centrer et réduire chaque ligne de la matrice B pour obtenir une matrice \hat{B} et créer la matrice de Wishart associée à la matrice \hat{B} . C'est-à-dire la matrice $W_{488} = \frac{1}{1257} \hat{B} \hat{B}^t$. On fait une telle modification afin d'essayer de récupérer l'hypothèse que les variables de la matrice B ont une espérance nulle et une variance de 1.

5. Calculer les valeurs propres de W_{488} .

Une fois ce travail fait, on peut ensuite calculer les valeurs propres de W_{488} . Si les entrées de \hat{B} sont bel et bien indépendantes et identiquement distribuées, alors on devrait s'attendre à ce que la mesure spectrale converge vers la loi de Marchenko-Pastur de paramètre $\gamma = 488/1257 \approx 0.388$.

On constate que ce n'est pas le cas, en effet certaines valeurs propres sont beaucoup plus grandes que γ_+ . On retrouve ces grandes valeurs propres dans la table 2.

2.68	3.00	3.21	3.30
3.49	3.81	4.40	5.11
5.73	7.35	7.81	8.35
9.74	23.33	28.42	160.63

TABLE 2 – Valeurs propres retirées.

On constate alors que la valeur propre 160.63 est beaucoup plus grande que ce qui est prédit par le théorème 3.2. Cependant, les grandes valeurs propres ne sont pas les seules qui diffèrent du résultat attendu. En effet, il y a beaucoup de valeurs propres qui sont inférieures à γ_- . Sur 488 valeurs propres, 142 se situent en dessous de γ_- , ce qui représente 29% de toutes les valeurs propres. Dans la même lancée, 3.3% des valeurs propres sont supérieures à γ_+ , ce qui ne laisse que 67.7% des valeurs propres qui sont dans l'intervalle attendu.

Pour illustrer l'histogramme des valeurs propres, on a enlevé les 16 valeurs propres de la table 2 afin de rendre l'histogramme lisible. L'histogramme des 472 valeurs propres restantes ainsi que la loi de Marchenko-Pastur pour $\gamma = 488/1257$ sont illustrés à la figure 15.

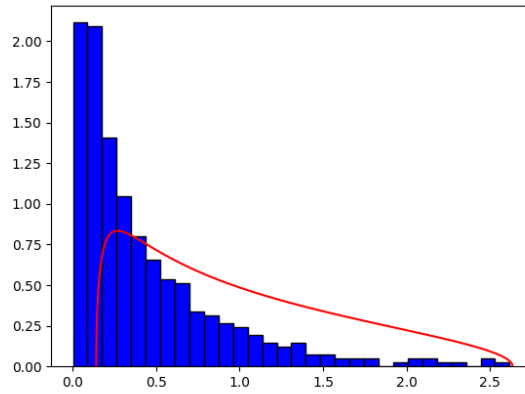


FIGURE 15 – Histogramme des 472 valeurs propres de W_{488} conservées.

On constate que l'histogramme des valeurs propres ne converge pas vers la loi de Marchenko-Pastur. Sur 488 valeurs propres, seulement 330 se trouvent dans l'intervalle $[\gamma_-, \gamma_+]$. On peut donc s'attendre à ce qu'une ou plusieurs hypothèses du théorème 3.2 ne soient pas satisfaites. Comme une des hypothèses imposées sur la matrice \hat{B} associée à une matrice de Wishart est que les entrées de \hat{B} soient indépendantes, il est possible que les valeurs propres en dehors de l'intervalle $[\gamma_-, \gamma_+]$ témoignent d'une dépendance entre certaines variations. On pourrait donc avoir qu'à certains moments les variations sont toutes négatives, comme lorsque la valeur des actions de nombreuses sociétés diminue. On constate d'ailleurs que c'est le cas lors d'un krach boursier.

Les entrées pourraient aussi ne pas être toutes identiquement distribuées, cela pourrait aussi expliquer en partie la différence entre le résultat attendu et ce qui est observé.

Soyez rassurés, le fait que l'histogramme des valeurs ne colle pas à la loi de Marchenko-Pastur n'est pas négatif. C'est même l'opposé. Ce sont les valeurs propres qui se trouvent à l'extérieur de l'intervalle $[\gamma_-, \gamma_+]$ qui donnent de l'information sur le marché. Les valeurs propres près de 0 suggèrent qu'il y a des portefeuilles qui ont soit un rendement non nul ou que le risque qui leur est associé est presque nul. Cette analyse est faite de manière plus approfondie dans [\[3\]](#).

6 Conclusion

Tout au long de ce rapport, on a pu voir intervenir les diverses notions qui ont été introduites dans la section contenant les préliminaires. Ces outils ont permis de démontrer la loi du demi-cercle, mais aussi de faire les simulations ainsi que d'appliquer des résultats sur les matrices aléatoires au domaine de la finance.

La démonstration de la loi du demi-cercle s'appuie sur un ensemble de concepts essentiels, tels que la transformée de Stieltjes, la résolvante, l'invariance par rotation, ainsi que la convergence des suites transformée de Stieltjes de mesures. Ce résultat central permet de relier ces différentes notions dans un cadre unifié, en établissant la convergence de la mesure spectrale associée à une matrice de Wigner.

De plus, les diverses hypothèses de la loi du demi-cercle ont été mises à l'épreuve dans des simulations. On a ainsi pu explorer certains cas limite de la loi du demi-cercle. Ces simulations ont permis de mettre en évidence la différence entre la loi du demi-cercle et la loi de Marchenko-Pastur. Cette différence ne vient pas que du fait que les mesures spectrales ne convergent pas vers la même loi, mais aussi du fait que les matrices de Wigner et de Wishart ont une structure fondamentalement différente.

De la section sur l'application en finance, il a été possible de conclure certains résultats malgré que les résultats attendus n'aient pas été observés. En effet, malgré que la mesure spectrale de la matrice considérée ne convergeait pas vers la mesure spectrale, on a tout de même déduit de l'information à partir des valeurs propres très près de 0. Ces valeurs propres indiquent que le risque associé à certains portefeuilles est presque nul ou que le rendement de certains portefeuilles est non nul.

Cependant, la finance n'est pas le seul domaine où les matrices aléatoires peuvent servir d'outil important. On peut notamment trouver des applications des matrices aléatoires en physique quantique, en intelligence artificielle et en analyse complexe où elle est liée à la fonction zêta de Riemann.

Une ouverture intéressante pour la suite serait de s'intéresser à des matrices de Wigner dont les entrées sont des variables aléatoires complexes. Ce faisant, il est possible de généraliser la loi du demi-cercle à des matrices de Wigner qui sont hermitiennes au lieu d'être simplement symétriques comme celles considérées dans ce rapport. Aussi, une loi particulièrement intéressante émerge dans le contexte des matrices aléatoires non hermitiennes. En effet, sous certaines hypothèses, la mesure spectrale converge vers le disque unité du plan complexe. Ce résultat, bien que très intéressant, nécessite des outils mathématiques bien plus complexes que ceux présentés dans ce rapport. Pour ce qui est de l'application en finance, il serait intéressant d'explorer le même problème que celui présenté dans ce rapport en utilisant d'autres sociétés ou bien encore de changer la période considérée.

En conclusion, les matrices aléatoires offrent un cadre théorique riche, avec de nombreuses applications potentielles. Bien que ce rapport se soit principalement concentré sur les matrices de Wigner et de Wishart à entrées réelles, la théorie dans son ensemble est vaste, en pleine expansion, et constitue un domaine de recherche particulièrement actif.

Annexe A Code pour l'application en finance

Les codes utilisés pour tracer les histogrammes se trouvent sur GitHub à l'adresse <https://github.com/TomasTurcotte/Projet-de-fin-d-etudes>. Comme le code utilisé dans la section 5 est plus complexe que les autres, on le donne dans le rapport.

```
1 import yfinance as yf
2 import pandas as pd
3 import numpy as np
4 import matplotlib.pyplot as plt
5 np.set_printoptions(suppress=True)
6
7
8 # Recherche des indices des sociétés du S&P500
9 tickers = pd.read_html('https://en.wikipedia.org/wiki/List_of_S%26P_500_companies')[0]
10 liste_tickers = tickers.Symbol.to_list()
11
12
13 # On enlève les indices dont les données ne sont pas récupérables
14 liste_tickers.remove('BRK.B')
15 liste_tickers.remove('BF.B')
16
17
18 # Téléchargement des données et création d'une matrice
19 donnees = yf.download(liste_tickers, '2020-01-01', '2025-01-01',
20                        auto_adjust=True)['Close']
21 Matrice_donnees = np.transpose(donnees.to_numpy())
22
23
24 # On enlève les sociétés dont certaines des valeurs de fermeture ne sont pas
25 # disponible sur la période de temps choisie
26 Liste_lignes = []
27 for j in range(Matrice_donnees.shape[1]):
28     for i in range(Matrice_donnees.shape[0]):
29         if np.isnan(Matrice_donnees[i, j]) == True:
30             if i not in Liste_lignes:
31                 Liste_lignes.append(i)
32
33 compteur = 0
34 for indice in Liste_lignes:
35     Matrice_donnees = np.delete(Matrice_donnees, indice - compteur, 0)
36     compteur = compteur + 1
37
38
39 # Calcul des variations entre chacune des journées ou la bourse est ouverte
40 for i in range(Matrice_donnees.shape[1] - 1):
41     Matrice_donnees[:, i] = Matrice_donnees[:, i + 1] - Matrice_donnees[:, i]
42
43 Matrice_donnees_finale = np.delete(Matrice_donnees, Matrice_donnees.shape[1] - 1, 1)
44 print('Dimensions de la matrice de données finale = ', Matrice_donnees_finale.shape)
45
46
47 # Standardisation de la matrice de données et calcul de la matrice de covariance
48 Matrice_donnees_finale = (Matrice_donnees_finale -
```

```

49             Matrice_donnees_finale.mean(1, keepdims=True))/\
50             Matrice_donnees_finale.std(1, keepdims=True)
51 Matrice_covariance = (1/Matrice_donnees.shape[1])*(
52     Matrice_donnees_finale @ np.transpose(Matrice_donnees_finale))
53
54
55 # Paramètre gamma pour Marchenko-Pastur
56 gamma = Matrice_donnees_finale.shape[0]/Matrice_donnees_finale.shape[1]
57 gamma_plus = (1 + np.sqrt(gamma))**2
58 gamma_moins = (1 - np.sqrt(gamma))**2
59 t = np.linspace(gamma_moins, gamma_plus, 1000)
60
61
62 # Calcul et traitement des valeurs propres
63 Valeurs_propres = np.linalg.eig(Matrice_covariance)[0]
64 Liste_1 = []
65 Liste_2 = []
66 Liste_3 = []
67 Liste_4 = []
68 for vp in Valeurs_propres:
69     if (vp < gamma_plus):
70         Liste_1.append(vp)
71     if (vp < gamma_moins):
72         Liste_3.append(vp)
73     elif (vp > gamma_moins):
74         Liste_4.append(vp)
75     if (vp > gamma_plus):
76         Liste_2.append(vp)
77 Liste_2.sort()
78
79 print('Nombre de valeurs propres total : ', Valeurs_propres.shape[0])
80 print('Nombre de valeurs propres admissibles : ', len(Liste_1))
81 print('Valeurs propres retirées : ', Liste_2)
82 print('Nombre de valeurs propres inférieures à gamma_- : ', len(Liste_3))
83 print('Nombre de valeurs propres entre gamma_- et gamma_+ : ', len(Liste_4))
84 plt.hist(Liste_1, density=True, bins=30, color='blue', edgecolor='black')
85 plt.plot(t, (np.sqrt((gamma_plus-t)*(t-gamma_moins)))/(2*np.pi*gamma*t),
86     color='red')
87 plt.show()

```

Références

- [1] G. W. ANDERSON, A. GUIONNET et O. ZEITOUNI, *An Introduction to Random Matrices* (Cambridge studies in advanced mathematics). Cambridge University Press, 2012, t. 118.
- [2] Z. BAI et J. W. SILVERSTEIN, *Spectral analysis of large dimensional random matrices*, 2nd edition. Springer series in statistics, 2010.
- [3] M. CUCURINGU, *Lecture 4: PCA in high dimensions, random matrix theory and financial applications*, sept. 2019.
- [4] L. LALOUX, P. CIZEAU, J.-P. BOUCHAUD et M. POTTERS, « Noise Dressing of Financial Correlation Matrices », *Physical Review Letters*, 1999.
- [5] G. LIVAN, M. NOVAES et P. VIVO, *Introduction to random matrices : theory and practice* (SpringerBriefs in mathematical physics). Springer, 2018, t. 26.
- [6] M. POTTERS et J.-P. BOUCHAUD, *A First Course in Random Matrix Theory: for Physicists, Engineers and Data Scientists*. Cambridge University Press, 2020.