

Stage ingénieur de 3^{ème} année, effectué chez Groupama Asset Management, en vue de l'obtention du diplôme d'ingénieur dans la spécialité Mathématiques Appliquées et Calcul Scientifique option Finance (MACS) de l'école d'ingénieurs Sup Galilée

Estimation des Matrices de Variance-Covariance et Théorie des Matrices Aléatoires

Étude théorique des matrices aléatoires - Présentation des principales lois connues - Application de cette théorie au débruitage des matrices de corrélations - Application à la gestion de portefeuilles

Hermann Sosthene TALLA MELI
Maître de stage : Sélim BOUGHALEM
Tuteur Universitaire : Ahmed KEBAIER



<hermann.tallameli@gmail.com, sboughalem@groupama-am.fr,
kebaier@math.univ-paris13.fr>

Groupama Asset Management - Sup Galilée

SOMMAIRE

Introduction	11
1 Présentation de Groupama Asset Management	13
1.1 L'Entreprise : Un gestionnaire activement responsable.	13
1.2 Expertises de Groupama Asset Management	13
1.3 Organigramme de la société	14
2 Loi de Marcenko-Pastur et Application à la Gestion des Portefeuilles	15
2.1 Introduction	15
2.2 Analyse spectrale de matrices aléatoires de grande dimension	16
2.3 Limites des valeurs propres extrêmes	16
2.4 Matrice de covariance d'un échantillon et loi de Marcenko-Pastur (M-P) [26]	17
2.4.1 Intérêt	17
2.4.2 Loi de Marcenko-Pastur dans le cas i.i.d.	17
2.4.2.1 Moments de la Loi de M-P	18
2.4.2.2 Quelques Lemmes sur la théorie des graphes	19
2.4.2.3 Preuve de la loi de M-P dans le cas <i>i.i.d.</i> : preuve du Théorème 2	20
2.4.2.4 Quelques exemples	23
2.4.2.4.1 Cas d'une $\mathcal{N}(0, 1)$	23
2.4.2.4.2 Cas d'une $\mathcal{N}(\mu, \sigma^2 \mathbf{I}_n)$	24
2.4.2.4.3 Cas d'une loi lognormale	25
2.4.2.4.4 Cas d'une loi de Student	25
2.4.3 Généralisation de la loi de M-P dans le cas non iid	26
2.5 Application à la Théorie du Portefeuille	31
2.5.1 Rappels mathématiques	31
2.5.2 portefeuille global de variance minimum [49]	31
2.5.3 Écriture des coefficients d'une matrice symétrique en fonction de ses valeurs propres et de ses vecteurs propres	33
2.5.4 Application du résultat de l'équation 2.57 à l'équation 2.56	34
2.6 Débruitage des matrices de covariance et de corrélation : Reconstruction des matrices de corrélation et de variance-covariance	34
2.6.1 Algorithme de la reconstruction	35
2.6.2 Présentation d'un cas simple : Cas d'une matrice 2×2	35
2.6.3 Calcul des valeurs propres et des vecteurs propres, et Conservation de la trace lors du changement de base	37
2.6.3.1 Calcul des valeurs propres et des vecteurs propres	37
2.6.3.2 Conservation de la trace lors du changement de base	38
2.6.4 Interprétation financière	38
2.6.5 Choix de modification des valeurs propres	40
2.6.5.1 Cas où on remplace les deux valeurs propres par leur moyenne	40
2.6.5.2 Cas où on remplace l'une des deux valeurs propres par 0	41
2.6.5.3 Conclusion	41

2.6.6	Application	41
2.6.7	Tentative de généralisation de la méthode	44
2.6.7.1	Cas d'une matrice de taille 3×3	44
2.6.7.2	Illustration de la méthode dans le cas d'une matrice 3×3	48
2.6.7.3	Généralisation dans le cas d'une matrice $n \times n$	51
2.6.7.4	Application à des données réelles	55
3	Méthode d'estimation du modèle financier, Transformée de Stieltjes et Application à la Gestion des Portefeuilles	60
3.1	Introduction	60
3.2	Enrichissement du modèle : Somme et Produit de deux matrices aléatoires	60
3.2.1	Somme de deux matrices aléatoires : cas où $Y = X + \epsilon$	60
3.2.2	Produit de deux matrices aléatoires : cas où $Y = X\epsilon$	61
3.2.3	Conclusion	61
3.3	Produit de deux matrices aléatoires	61
3.3.1	Introduction	61
3.3.2	Résultats principaux	62
3.3.2.1	Résultats	62
3.3.2.2	LSD de la F -matrice	63
3.3.2.3	Quelques exemples	63
3.3.2.3.1	Produit d'une matrice de Wishart et de l'inverse d'une matrice de Wishart	63
3.3.2.3.2	Produit de deux matrices de Wishart indépendantes	63
3.4	Distribution des rendements	64
3.4.1	Distribution univariée des rendements	64
3.4.2	Distribution multivariée des rendements	65
3.5	Algorithme EM (Espérance-Maximisation) [30]	66
3.5.1	Principe de fonctionnement	66
3.5.2	Applications de l'algorithme EM à une loi de Student multivariée	67
3.5.2.1	Cas où la matrice de corrélation est la matrice identité	67
3.5.2.2	Cas où la matrice de corrélation est différente de la matrice identité	68
3.6	Théorie des Matrices Aléatoires et application de la transformée de Stieltjes	68
3.6.1	Préliminaires	68
3.6.2	Matrices libres	69
3.6.3	Application : Wigner et Marcenko-Pastur	70
3.6.4	Plus d'applications	71
3.6.4.1	Cas d'une matrice de corrélation aléatoire non triviale	71
3.6.4.2	Cas de l'ensemble de Student	72
3.7	Méthode polynomiale pour matrices aléatoires	75
3.7.1	Introduction	75
3.7.2	Exemples de la somme et du produit d'une Wigner et d'une Wishart	75
3.7.3	Applications de la méthode polynomiale	77
3.8	Application financière : Débruitage des matrices de corrélation et distribution empirique des valeurs propres	82
3.8.1	Application sur des données simulées	83
3.8.2	Application sur des rendements de marché	83
3.8.2.1	Estimation des paramètres avec notre jeu de données d'application	83
3.8.2.2	Estimation des paramètres dans le cas général	84
4	Conclusions et Perspectives	85
4.1	Nettoyage des vecteurs propres	85
4.2	Matrices de Levy	85
4.3	Corrélations non linéaires et Copules	86
4.4	Décomposition en valeur singulière (SVD) aléatoire et Analyse par Composante Canonique	86

Conclusion	88
A Quelques résultats en algèbre linéaire	90
A.1 Matrices de variance-covariance	90
A.1.1 Définition	90
A.1.2 Propriétés des matrices de variance-covariance	91
A.1.3 Relation entre $X'X$ et XX'	91
A.2 Formule de l'inverse d'une matrice	91
A.3 Trouver une matrice	92
A.4 Trace d'une matrice inverse	92
A.5 Différence des Traces d'une matrice et de ses sous-matrices majeures	93
A.6 Matrice inverse des matrices complexes	93
A.7 Inégalités de perturbation	93
A.8 Inégalités du rang	94
A.9 Lemme de Borel-Cantelli	94
A.9.1 Résultat	94
A.9.2 Limite supérieure d'ensembles	94
A.9.3 Théorème de Borel-Cantelli (théorie de la mesure)	95
A.9.4 Lemme de Borel-Cantelli (probabilités)	95
B Méthode des moments et méthodologies	96
B.1 Théorème de Convergence des Moments	96
B.2 Méthodologies	97
B.2.1 Méthode des moments	97
B.2.2 Transformée de Stieltjes	97

Remerciements

Tout d'abord, j'aimerais dire un grand merci à **Selim BOUGHALEM** pour l'accueil chaleureux qu'il m'a offert dans l'équipe Gestion SIGMA chez Groupama Asset Management. Son soutien tout au long de ce stage, son expérience en Gestion de portefeuilles, ses connaissances en analyse quantitative et surtout ses conseils m'ont permis de me faire une idée concrète du vaste monde qu'est la gestion. Il a toujours été disponible et nos discussions m'ont toujours été enrichissantes, personnellement et professionnellement.

Je remercie particulièrement tous les membres de l'équipe pour tout leur soutien et leur disponibilités. Je pense à **Pierre Alexis, David, Cyrille** et **Jean** qui n'ont pas hésité à m'expliquer en quoi consiste leur travail. J'ai pu avoir une vue des différents types de gestion et de l'analyse financière.

Je tiens à remercier **Olivier LAFITTE** et toute l'équipe de la MACS pour leur soutien, leur écoute et leurs conseils indéfectibles qu'ils ont pu me donner jusqu'ici. Merci de m'avoir fait partager votre passion pour les mathématiques appliquées.

Un grand merci à **Jonathan Wiltord** qui a été très disponible pendant ma recherche de stage, avec qui j'ai eu de belles discussions amicales, et qui a su me guider dans mes recherches. Je lui souhaite bien de choses pour la suite. Beaucoup de courage et bonne continuation...

Mes remerciements vont à **Bérani CHEKLAME, Pierre MARTIN, Damien TEULON** et à **Isabelle METAYER**, tous stagiaires comme moi, qui ont su garder une ambiance scolaire qu'il nous fallait pour ne pas être déconnecté de ce monde, et qui m'aidaient souvent à prendre des pauses... Je leur souhaite bien de choses pour la suite. Ce fût magnifique...

Merci à mes amis pour tout ce qu'ils ont pu faire pour moi pendant mon cursus. Le soutien, la bonne humeur, les encouragements qu'ils m'ont apporté ont été d'une aide très précieuse. Bon courage pour la suite.

Enfin, je dis un grand merci à toute ma famille, qui m'a toujours soutenu et encouragé jusqu'ici. Il me serait impossible de ne pas la saluer. Malgré la distance qui me sépare d'une partie de ma famille, et tant d'années passées loin d'elle, elle reste ma source inépuisable de motivations et de volonté. Une pensée particulière pour mon oncle qui nous a quittés l'été 2012... PAIX À SON ÂME.

A tous ceux qui m'ont aidé de près ou de loin, je leur adresse mes sincères remerciements.

Notations et abréviations

- X : vecteur ou matrice de données (par exemple les rendements)
 X' : transposée de la matrice X
 p : nombre de lignes de la matrice X
 n : nombre de colonnes de la matrice X
 $c(\text{ou } y)$: rapport $\frac{p}{n}$ qui peut être fini ou infini, mais positif
 \mathbf{S} : matrice de variance-covariance de X
 \mathbf{S}_n : suite de matrices de variance-covariance
 C : matrice de corrélation de X
 λ_i : valeur propre n° i de la matrice de corrélation
 v_i : vecteur propre associé à la valeur propre λ_i
 F^A : distribution spectrale empirique (ESD) de A
 μ : moyenne des données. Scalaire ou vecteur suivant que X est un vecteur ou une matrice
 σ : variance des données (scalaire)
 β_k : moments de la loi de Marcenko-Pastur (MP)
 $\mathbb{E}(X)$: espérance de X
 $\mathbb{V}(X)$: variance de X
 s_y : transformée de Stieltjes
 I_p : matrice identité de taille $p \times p$
 z : nombre complexe de la forme $\lambda + i\epsilon$
 \Im : partie imaginaire (ϵ) du nombre complexe z
 \Re : partie réelle (λ) du nombre complexe z
 σ_x (ou σ_{ii}) : volatilité du titre x (ou du titre i)
 σ_{xy} : covariance des titres x et y
 ρ_{xy} : corrélation des titres x et y
 $\log(x)$: logarithme de x
 $\det(A)$: déterminant de la matrice A
RMT : Théorie des Matrices Aléatoires
ESD : distribution spectrale empirique
LSD : distribution spectrale limite
i.i.d. : indépendant identiquement distribué
portefeuille GMV : portefeuille global de variance minimum

Résumé

Depuis le début des années 2000, beaucoup d'auteurs se sont penchés sur l'utilisation de la Théorie des Matrices Aléatoires (RMT) en finance. Elle sert notamment à la construction des portefeuilles en gestion. Dans ce rapport, notre travail consistera à présenter la Théorie des Matrices Aléatoires à l'estimation de matrices de Variance-Covariance, et son utilisation dans le domaine de la finance. Nous donnerons tout d'abord les principaux résultats théoriques. Par la suite, nous illustrerons ces résultats théoriques via des applications numériques sur des lois bien connues, pour enfin tester la méthode sur des données financières. Un enrichissement du modèle cité dans la théorie sera présenté, et une application numérique sur des données réelles de marché sera donnée pour illustrer la méthode. Nous analyserons les résultats obtenus en les comparant aux résultats obtenus par les méthodes générales utilisées dans le milieu professionnel, pour voir quel a été l'impact de l'enrichissement du modèle.

Introduction

Au point de départ de ce travail, volonté y est de comprendre plus systématiquement les relations entre la Théorie des Matrices Aléatoires et certains aspects de la construction de portefeuilles. Nous avons étudié pour une assez large classe de matrices aléatoires un lien direct entre ces deux théories.

Les matrices aléatoires sont considérées dès 1928 par le statisticien **Wishart**. Il donnera son nom aux matrices de **Wishart** [36], également appelées matrices du LOE. Il s'agit des matrices XX^* , où X est une matrice réelle de taille $n \times p$ dont les entrées sont des gaussiennes réelles standard indépendantes.

Cependant, c'est aux physiciens que l'on doit l'apparition de ce qu'on appellera plus tard la Théorie des Matrices Aléatoires. En effet, dans les années 1950, face à la complexité croissante des spectres nucléaires observés expérimentalement, **Wigner** a suggéré de remplacer l'opérateur hamiltonien du noyau par une matrice aléatoire. Cette hypothèse féconde a conduit au développement rapide d'un nouveau champ de recherche très actif en physique théorique, qui s'est propagé à la théorie des nombres en mathématiques, avec notamment une connexion intéressante avec la fonction zêta de Riemann. En plus de ces exemples, on compte parmi les applications de la Théorie des Matrices Aléatoires les systèmes intégrables, le chaos quantique, la gravité quantique en deux dimensions et aujourd'hui, le sujet qui a fait l'objet de notre étude au sein de Groupama Asset Management à savoir **la gestion de portefeuilles**.

Depuis le début du XXème siècle et l'apparition des statistiques multivariées, beaucoup de travaux ont été consacrés à la mise au point de nouveaux outils statistiques pour traiter de grandes quantités de données. Traditionnellement, ces outils ont été développés dans le cas où l'on dispose d'un grand nombre N d'observations statistiques (ou échantillons) de dimension raisonnable M , et se basent relativement souvent sur l'utilisation de la matrice de covariance empirique des observations. Si les N observations statistiques sont représentées par les vecteurs M -dimensionnels y_1, \dots, y_N , alors la matrice de covariance empirique est donnée par

$$S = \frac{1}{N} \sum_{i=1}^N X_i X_i^*. \quad (2)$$

Aujourd'hui, les progrès des différentes technologies ont entraîné une nette augmentation de la dimension des données à traiter (M grand), sous des contraintes de temps toujours plus fortes (limitations sur N). La plupart des outils classiques (estimateurs, tests) ont été étudiés quand le nombre d'échantillons N est bien plus grand que la dimension des observations M . Or, dans beaucoup de domaines, il n'est pas toujours possible de disposer d'un nombre trop grand d'observations, notamment quand les modèles ne sont stationnaires que sur une courte période de temps. Citons notamment comme exemple le domaine de la finance, où l'on dispose de grands portefeuilles d'actions, mais peu d'échantillons, car là encore les modèles évoluent très vite. Il apparaît dès lors que les outils statistiques classiques affichent des comportements différents, notamment ceux basés sur l'estimation empirique des matrices de covariance.

Dans la RMT, il existe plusieurs ensembles classiques de matrices aléatoires, et chaque ensemble doit son nom à un célèbre mathématicien. Au rang des ensembles les plus connus figurent :

- **Matrice de Wigner** (ou ensemble Hermitien) Les matrices de Wigner [2, 22] sont souvent connues comme des ensembles Hermitiens ou gaussiens [23]. La condition nécessaire est que, les éléments doivent être indépendants et avoir la même variance comme dans le cas gaussien. Ce qui signifie une variance $\frac{1}{N}$ hors de la diagonale, $\frac{\sqrt{2}}{N}$ sur la diagonale, et les éléments indépendants dans la partie triangulaire supérieure de la matrice.

Soit G une matrice aléatoires $N \times N$, avec des éléments indépendants, de moyenne nulle et de variance unitaire. La matrice de Wigner W peut être obtenue de G par la relation

$$W = \frac{G + G'}{\sqrt{2N}}. \quad (3)$$

➡ **Matrice de Wishart** (ou ensemble de Laguerre) Les matrices de Wishart [24], aussi parfois nommées matrices de Grammian [25] ou matrice de covariance, sont très utilisées en statistiques et ont été étudiées par Muirhead [27].

Soit G une $N \times M$ matrice aléatoire, avec des éléments indépendants, de moyenne nulle et de variance unitaire. La matrice de Wishart W peut être obtenue par la formule

$$W = \frac{1}{M} G G'. \quad (4)$$

Dans une matrice de Wishart, les éléments doivent être i.i.d et normalement distribués, comme dans le cas de l'ensemble de Laguerre.

➡ **Matrice de MANOVA** (ou ensemble de Jacobi)

$$J = [W(c_1) + W(c_2)]^{-1} W(c_1) \quad (5)$$

Notons que la définition d'un tel ensemble requiert que les matrices $W(c_1)$ et $W(c_2)$ soient «pures». Dans la littérature, on utilise pour les matrices de MANOVA, la notation $J(c_1, c_2)$, où $c_1 = \frac{N}{M_1}$ et $c_2 = \frac{N}{M_2}$ sont les paramètres associés respectivement aux deux matrices de Wishart. On peut noter que la définition d'une telle matrice demande que $c_1 < 1$ i.e. $M_1 > N$.

Ce stage s'inscrit dans le cadre de l'estimation des matrices de variance-covariance et la théorie des matrices aléatoires. L'objectif du stage est double. D'une part, d'appliquer la théorie des matrices aléatoires à l'estimation de matrices de variance-covariance, et, d'autre part, élaborer et développer des outils informatiques pour établir un protocole de validation des estimations. Ces deux objectifs permettront au final de construire des portefeuilles. Dans notre travail, mention ne sera faite que des matrices de Wishart décrites dans l'équation (4) (qui correspondent aux matrices de covariance). Nous définirons d'abord ce qu'est la Théorie des Matrices Aléatoires, ensuite, nous essaierons d'enrichir la théorie décrite, et nous ferons des applications pour illustrer cette théorie qui propose un modèle probabiliste pour les rendements de marche.

Chapitre 1

Présentation de Groupama Asset Management

1.1 L'Entreprise : Un gestionnaire activement responsable.

Groupama Asset Management, filiale de Groupama, est l'un des principaux acteurs français de la gestion d'actifs pour le compte d'une clientèle professionnelle.

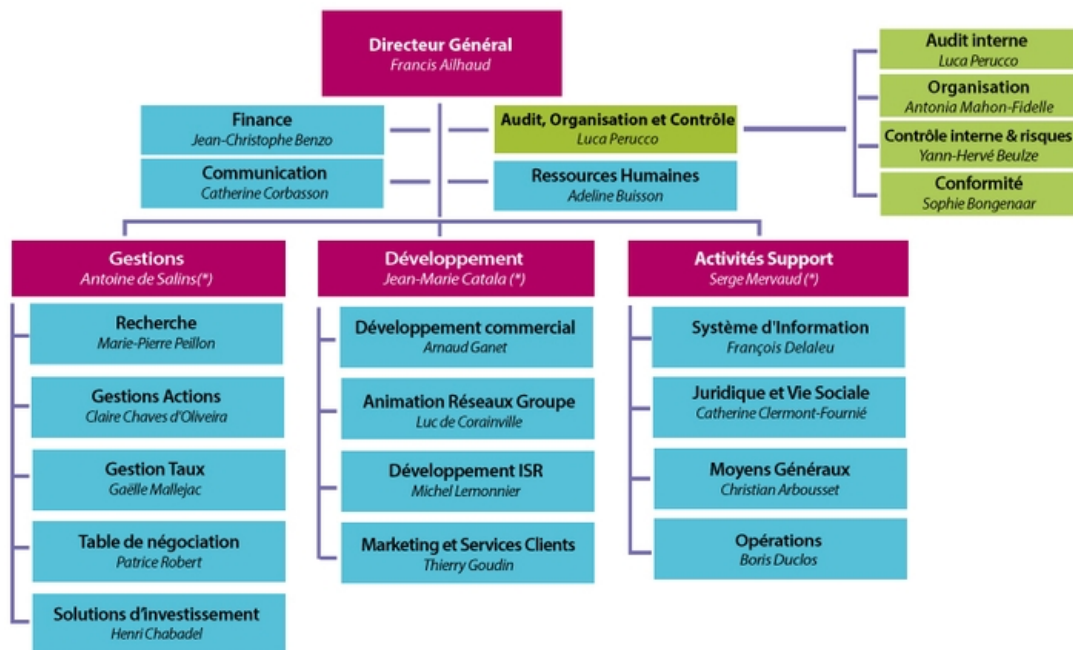
- ➔ **Un investisseur responsable :** Depuis 2001, Groupama Asset Management est engagée dans la promotion de l'investissement responsable (IR) qu'elle développe au cœur de ses métiers. Sa stratégie IR s'inscrit dans le cadre de sa responsabilité sociale auprès de ses clients et de la collectivité, avec l'objectif de se positionner sur l'épargne longue pour un financement plus responsable et durable de l'économie
- ➔ **Une relation personnalisée :** Dotée d'une organisation à taille humaine, Groupama Asset Management offre une structure réactive et de proximité, à l'ensemble de ses clients. Chaque typologie de clientèle de Groupama - investisseur institutionnel- entreprise et distributeur, tant en France qu'en Europe, fait l'objet d'une relation personnalisée avec les équipes commerciales permettant de lui offrir des solutions d'investissement pertinentes et innovantes.
- ➔ **Des solutions globales d'investissement :** A l'écoute de ses clients et des évolutions des marchés, Groupama Asset Management propose des solutions globales d'investissement qui couvrent les principales classes d'actifs. Pour construire une performance régulière dans une vision de long terme, ses équipes de gestion s'appuient sur une solide équipe de recherche.

1.2 Expertises de Groupama Asset Management

Les savoir-faire des équipes de GAM permet d'offrir aux investisseurs des expertises de gestion reconnues :

- ➔ **Gestion Actions :** L'expertise Actions de Groupama AM repose sur une gestion active fondée sur plusieurs moteurs de performance dont, en particulier, l'allocation sectorielle et le « stock picking ».
- ➔ **Gestion Taux :** L'expertise Taux de Groupama AM regroupe les gestions obligataire et monétaire. La gestion obligataire s'appuie sur de nombreuses classes d'actifs : dettes souveraines de la zone euro ou internationale, obligations indexées sur l'inflation de l'euro ou internationale, gestion crédit « Investment Grade » et gestion aggregate. Chacune de ces gestions, qui s'exerce aussi bien dans des portefeuilles ouverts que dédiés, fait l'objet d'une expertise spécifique.
- ➔ **Gestion Sigma :** L'offre de gestions comprend les fonds de gestion directe « Total Return » et « Absolute Return », la multigestion alternative et la multigestion directionnelle.
- ➔ **Gestion dynamique d'allocation :** Ce type de gestion propose une allocation dynamique entre différentes classes d'actifs et différentes zones géographiques, tirant ainsi parti de la pluralité des marchés financiers. Les experts de Groupama AM proposent des profils de gestion définis en fonction du degré de risque maximal auquel l'investisseur est prêt à s'exposer. Pour chaque profil de risque, une répartition souple et dynamique entre les différentes catégories d'actifs est opérée.

1.3 Organigramme de la société



(*) Membres du Comité de Direction Générale

FIGURE 1.1 – Organigramme de Groupama Asset Management

Chapitre 2

Loi de Marcenko-Pastur et Application à la Gestion des Portefeuilles

2.1 Introduction

Comme il a été dit dans l'introduction, la naissance et le développement de la RMT remontent à la fin des années 1940. En mécanique quantique (QM), les niveaux d'énergie d'un système sont décrits par les valeurs propres d'un opérateur Hermitien A sur un espace de Hilbert, appelé Hamiltonien. Pour éviter de travailler avec un opérateur de dimension infini, les chercheurs approchent le système par discrétisation, ce qui revient à une troncature, et qui permet de garder uniquement la partie de l'espace de Hilbert qui est importante pour le problème à l'étude. De là, le comportement de limitation de grandes matrices aléatoires a attiré un intérêt spécial parmi ceux qui travaillaient sur la QM, et plusieurs lois ont été découvertes pendant cette période. Pour plus de détails sur les applications de la RMT dans la QM, voir le livre de Mehta [3] sur les matrices aléatoires.

Depuis la fin des années 1950, la recherche sur l'analyse spectrale limite des matrices aléatoires de grande dimension a attiré un intérêt considérable parmi les mathématiciens, les probabilistes et les statisticiens. Le travail de base a été celui de Wigner [4, 5], qui a découvert la loi semi-circulaire pour une matrice gaussienne (ou matrice de Wigner). Il a démontré que la distribution attendue d'une matrice de Wigner de grande dimension, tend vers une loi appelée loi semi-circulaire. Ce travail a été généralisé par Arnold [6, 7] et Grenander [9] sous plusieurs aspects. Bai et Yin [8] ont prouvé que la distribution spectrale d'une matrice de covariance (sous une bonne normalisation) converge vers la loi semi-circulaire lorsque la dimension est relativement petite devant la taille des données. Allant dans le même sens que Marcenko et Pastur [26] et Pastur [10, 11], plusieurs chercheurs tel Yin, Bai et Krishnaiah [12], Grenander et Silverstein [13], Jonsson [14], Wachter [15, 16], Yin [17], et Yin et Krishnaiah [18] ont développé la théorie asymptotique de l'analyse spectrale des matrices de covariance de grande dimension. Yin, Bai et Krishnaiah [20, 19], Silverstein [21], Wachter [15], Yin [17] et Yin et Krishnaiah [18] ont également fait des recherches sur la distribution spectrale limite de la F -matrice multivariée ou plus généralement, du produit de matrices aléatoires. Au début des années 80, les chercheurs ont beaucoup travaillé sur l'existence de la distribution spectrale limite (LSD) et leurs formes explicites dans certains cas de matrices aléatoires. Ces dernières années, la recherche sur la RMT est plutôt faite sur les théorèmes limites, tel que le théorème central limite pour les statistiques spectrales linéaires, les valeurs propres extrêmes et d'autres sujets.

Dans ce chapitre, nous nous intéresserons tout d'abord à l'analyse spectrale des matrices aléatoires de grande taille, ce qui nous permettra par la suite, de mieux explorer les sujets portant sur les limites des valeurs propres extrêmes, les méthodologies utilisées dans la RMT et surtout et enfin la loi de Marcenko-Pastur (MP), qui est une loi clé pour notre travail.

2.2 Analyse spectrale de matrices aléatoires de grande dimension

Soit A une matrice $m \times m$, ayant des valeurs propres $\lambda_j, j = 1, 2, \dots, m$. Si la matrice A a toutes ses valeurs propres réelles (i.e. est hermitienne), on peut définir la 1-dimension fonction de distribution (distribution spectrale empirique (ESD) de A) :

$$F^A(x) = \frac{1}{m} \# \{j \leq m : \lambda_j \leq x\} \quad (2.1)$$

Ici, $\#E$ signifie la cardinalité de l'ensemble E . Si les valeurs propres de A ne sont pas toutes réelles, on peut définir la 2-dimension fonction de distribution :

$$F^A(x) = \frac{1}{m} \# \{j \leq m : \Re(\lambda_j) \leq x, \Im(\lambda_j) \leq y\}. \quad (2.2)$$

L'un des problèmes majeurs de la Théorie des Matrices Aléatoires est d'établir la convergence d'une suite de distributions spectrales empiriques $\{F^{A_n}\}$ pour une suite donnée de matrices aléatoires $\{A_n\}$. La distribution limite F , qui est d'habitude non aléatoire, est appelée *distribution limite spectrale* (LSD) de la suite $\{A_n\}$.

Dans la théorie, nous travaillerons avec des suites de matrices aléatoires pour lesquelles la dimension (le nombre de colonnes) tend vers l'infini. L'importance de l'ESD est due au fait que, plusieurs statistiques en analyse multivariée peuvent être exprimées comme des fonctions de l'ESD de certaines matrices aléatoires. On a les exemples suivants :

Exemple 1. Soit A une matrice $n \times n$ définie positive. Alors

$$\det(A) = \prod_{j=1}^n \lambda_j = \exp \left(n \int_0^\infty \log x F^A(dx) \right).$$

Exemple 2. Soit une matrice de covariance de la forme $\Sigma = \Sigma_q + \sigma^2 \mathbf{I}$, où la dimension de Σ est p et le rang de Σ_q est q ($q < p$). Supposons que S est la matrice de covariance de n variables i.i.d. Appelons $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p$ les valeurs propres de S . Le test d'hypothèses statistique $H_0 : \text{rank}(\Sigma_q) = q$ contre $H_1 : \text{rank}(\Sigma_q) > q$ est donné par

$$\begin{aligned} T &= \frac{1}{p-q} \sum_{j=q+1}^p \sigma_j^2 - \left(\frac{1}{p-q} \sum_{j=q+1}^p \sigma_j^2 \right)^2 \\ &= \frac{p}{p-q} \int_0^{\sigma_q} x^2 F^S(dx) - \left(\frac{p}{p-q} \int_0^{\sigma_q} x^2 F^S(dx) \right)^2. \end{aligned}$$

On remarque dans les deux exemples quelle peut être l'importance de l'ESD.

2.3 Limites des valeurs propres extrêmes

Dans les applications des théorèmes asymptotiques de l'analyse spectrale des matrices aléatoires de grande dimension, deux problèmes très importants surviennent après avoir trouvé la LSD. Le premier concerne les bornes des valeurs propres extrêmes ; le deuxième concerne le taux de convergence de l'ESD.

Pour le premier problème, la littérature est très vaste. Le premier résultat a été obtenu par Geman [37], qui a prouvé que la plus grande valeur propre d'une matrice de covariance converge presque sûrement vers une certaine limite, sous une condition de croissance sur tous les moments de la distribution. Yin, Bai et Krishnaiah [20] ont prouvé le même résultat sous l'existence du moment d'ordre 4, et Bai, Silverstein et Yin [38] ont démontré que, l'existence du moment d'ordre 4 est également nécessaire pour l'existence de la limite. Pour une matrice de Wigner, Bai et Yin [40] ont trouvé les conditions nécessaire et suffisante pour la convergence presque sûre de la plus grande valeur propre. On peut noter que, pour ce type de matrices, la plus petite valeur propre et la plus grande valeur propre sont symétriques.

Comparé à la convergence presque sûre de la plus grande valeur propre d'une matrice de variance-covariance, un problème relativement plus complexe est de trouver la limite de la plus petite valeur

propre d'une matrice de variance-covariance de dimension élevée. Le premier travail sur ce sujet avait été fait dans Yin, Bai et Krishnaiah [19], où, il était démontré que la limite presque sûre de la plus petite valeur propre d'une matrice de Wishart a une borne inférieure positive lorsque le ratio entre la dimension et le degré de liberté est plus petit que $\frac{1}{2}$. Silverstein [41] a modifié ce travail pour mettre le ratio plus petit que 1. Silverstein [42] a prouvé plus loin, qu'avec une probabilité de 1, la plus petite valeur propre d'une matrice de Wishart tend vers la borne inférieure de la LSD lorsque le ratio dimension sur degré de liberté est plus petit que 1. Cependant, l'approche de Silverstein repose fortement sur la normalité de la distribution et ne peut pas être étendue sur le cas général. La contribution la plus récente sur ce sujet est de Bai et Yin [43], où il est démontré que, sous l'existence du moment d'ordre 4, la plus petite valeur propre (pour $p \leq n$) ou la $p - n + 1^{\text{ème}}$ plus petite valeur propre (lorsque $p > n$) converge vers $a(y) = \sigma^2(1 - \sqrt{y})^2$, où $y = \lim \frac{p}{n} \in (0, \infty)$. Comparé au cas de la plus grande valeur propre d'une matrice de covariance, l'existence du moment d'ordre 4 est nécessaire pour le problème de la plus petite valeur propre. Cependant, ce problème n'a pas encore été résolu.

2.4 Matrice de covariance d'un échantillon et loi de Marcenko-Pastur (M-P) [26]

2.4.1 Intérêt

La matrice de covariance d'un échantillon est une matrice aléatoire très importante dans l'inférence statistique multivariée. Elle est fondamentale dans les tests d'hypothèses, l'analyse en composantes principales, l'analyse factorielle et bien d'autres sujets. La plupart des tests statistiques sont définis par les valeurs propres.

On définit la matrice de covariance d'un échantillon de la manière suivante : Supposons que $x_{jk}, \{j, k = 1, 2, \dots\}$ est un tableau à doubles entrées, de variables aléatoires *i.i.d.* complexes, de moyenne nulle et de variance σ^2 . On écrit $x_j = (x_{1j}, \dots, x_{pj})'$ et $X = (x_1, \dots, x_n)$. La matrice de covariance de l'échantillon est définie par

$$\mathbf{S} = \frac{1}{n-1} \sum_{k=1}^N (x_k - \bar{x})(x_k - \bar{x})^* \quad (2.3)$$

avec $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$. De plus, dans la grande majorité des cas de l'analyse spectrale de matrices aléatoires de grandes dimensions, on définit simplement la matrice de covariance de l'échantillon par :

$$\mathbf{S} = \frac{1}{n} \sum_{k=1}^N x_k x_k^* = \frac{1}{n} X X^* \quad (2.4)$$

parce que $\bar{x} \bar{x}^*$ est une matrice de rang 1 et donc la suppression de \bar{x} n'affecte pas la LSD en raison du Théorème 16 en Annexe.

Dans l'analyse spectrale de matrices aléatoires de grandes dimensions, il est commun de supposer que la dimension p tend vers l'infini proportionnellement aux degrés de liberté n , soit $\frac{p}{n} \rightarrow y \in (0, \infty)$.

2.4.2 Loi de Marcenko-Pastur dans le cas i.i.d.

Dans cette partie, nous considérons la LSD d'une matrice de covariance pour le cas où les variables sous-jacentes sont *i.i.d.*

Théorème 1. *Supposons que les variables $\{x_{ij}\}$ sont des variables aléatoires réelles de moyenne nulle et de variance σ^2 . Considérons aussi que $\frac{p}{n} \rightarrow y \in (0, \infty)$. Alors, avec une probabilité de 1, F^S converge vers la loi de Marcenko-Pastur, qui est définie dans (2.5).*

Yin [17] avait considéré l'existence de la LSD de la suite de matrices $\mathbf{S}_n \mathbf{T}_n$, où \mathbf{T}_n est une matrice aléatoire définie positive, et indépendante de \mathbf{S}_n . Lorsque $\mathbf{T}_n = \mathbf{I}_p$, les résultats de Yin se réduisent au Théorème 1.

On a une généralisation du Théorème 1 dans le cas où les variables $\{x_{ij}\}$ sont des variables aléatoires complexes.

Théorème 2. Supposons que les variables $\{x_{ij}\}$ sont des variables aléatoires complexes de moyenne nulle et de variance σ^2 . Considérons aussi que $\frac{p}{n} \rightarrow y \in (0, \infty)$. Alors, avec une probabilité de 1, F^S converge vers la distribution limite décrite dans le Théorème 1.

2.4.2.1 Moments de la Loi de M-P

On se place dans le cas réel. La loi de M-P a pour fonction de densité :

$$p_y(x) = \begin{cases} \frac{1}{2\pi xy\sigma^2} \sqrt{(b-x)(x-a)} & \text{si } x \in [a, b] \\ 0 & \text{sinon,} \end{cases} \quad (2.5)$$

avec $1 - \frac{1}{y}$ point de masse à l'origine si $y > 1$, $a = \sigma^2(1 - \sqrt{y})^2$ et $b = \sigma^2(1 + \sqrt{y})^2$. Ici, la constante y est la dimension de l'échantillon de l'indice de rapport, et σ^2 le paramètre d'échelle. Si $\sigma^2 = 1$ la loi de M-P est dite loi de M-P standard.

Les moments β_k de la loi de M-P standard sont égaux à $\beta_k = \beta_k(y, \sigma^2) = \int_a^b x^k p_y(x) dx$. Par la suite, on déterminera la forme explicite des β_k . Notons que, $\forall k \geq 1$,

$$\beta_k(y, \sigma^2) = \sigma^{2k} \beta_k(y, 1).$$

Lemme 1. On a

$$\beta_k = \sum_{r=0}^{k-1} \frac{1}{r+1} \binom{k}{r} \binom{k-1}{r} y^r.$$

Preuve. Par définition,

$$\begin{aligned} \beta_k &= \frac{1}{2\pi y} \int_a^b x^{k-1} \sqrt{(b-x)(x-a)} dx \\ &= \frac{1}{2\pi y} \int_{-2\sqrt{y}}^{2\sqrt{y}} (1+y+z)^{k-1} \sqrt{4y-z^2} dz \quad (\text{avec } x = 1+y+z) \\ &= \frac{1}{2\pi y} \sum_{l=0}^{k-1} \binom{k-1}{l} (1+y)^{k-1-l} \int_{-2\sqrt{y}}^{2\sqrt{y}} z^l \sqrt{4y-z^2} dz \\ &= \frac{1}{2\pi y} \sum_{l=0}^{[k-1/2]} \binom{k-1}{2l} (1+y)^{k-1-2l} (4y)^{l+1} \int_{-1}^1 u^{2l} \sqrt{1-u^2} du \quad (\text{avec } z = 2\sqrt{y}u) \\ &= \frac{1}{2\pi y} \sum_{l=0}^{[(k-1)/2]} \binom{k-1}{2l} (1+y)^{k-1-2l} (4y)^{l+1} \int_0^1 w^{l-1/2} \sqrt{1-w} dw \quad (\text{avec } u = \sqrt{w}) \\ &= \sum_{l=0}^{[(k-1)/2]} \frac{(k-1)!}{l!(l+1)!(k-1-2l)!} y^l (1+y)^{k-1-2l} \\ &= \sum_{l=0}^{[(k-1)/2]} \sum_{s=0}^{k-1-2l} \frac{(k-1)!}{l!(l+1)!s!(k-1-2l-s)!} y^{l+s} \\ &= \sum_{l=0}^{[(k-1)/2]} \sum_{s=0}^{k-1-2l} \frac{(k-1)!}{l!(l+1)!(r-l)!(k-1-r-l)!} y^r \\ &= \frac{1}{k} \sum_{r=0}^{(k-1)} \binom{k}{r} y^r \sum_{l=0}^{\min(r, k-1-r)} \binom{s}{l} \binom{k-r}{k-r-l-1} \\ &= \frac{1}{k} \sum_{r=0}^{(k-1)} \binom{k}{r} \binom{k}{r+1} y^r \\ &= \sum_{r=0}^{(k-1)} \frac{1}{r+1} \binom{k}{r} \binom{k-1}{r} y^r. \end{aligned}$$

Par définition, on a $\beta_{2k} \leq b^{2k} = (1 + \sqrt{y})^{4k}$. De ce résultat, on voit facilement que la condition de Carleman (voir Théorème 18) est satisfaite.

N.B. : Notons que pour une compréhension facile, nous avons démontré le lemme dans le cas où $\sigma^2 = 1$. Dans toutes les démonstrations qui suivront dans ce rapport, nous considérerons toujours cette condition. \square

2.4.2.2 Quelques Lemmes sur la théorie des graphes

Pour montrer la convergence de l'ESD d'un échantillon de matrices de covariance de grandes dimensions vers la loi de M-P, nous utiliserons la méthode des moments (voir Annexe B). Pour cela, nous devons définir une classe de Δ -graphes et établir des lemmes concernant certains problèmes de comptage liés aux Δ -graphes.

Supposons i_1, \dots, i_k des nombres entiers positifs (non nécessairement distincts) inférieurs à p et j_1, \dots, j_k des nombres entiers positifs (non nécessairement distincts) inférieurs à n . Un Δ -graphe est défini de la manière suivante : *Tracer 2 lignes parallèles qu'on appellera ligne I et ligne J. Placer les points i_1, \dots, i_k sur la ligne I et les points j_1, \dots, j_k sur la ligne J, et, tracer k segments (vers le bas) de i_u vers j_u , $u = 1, \dots, k$ et k segments (vers le haut) de j_u vers i_{u+1} , $u = 1, \dots, k$ (avec la convention $i_{k+1} = i_1$). Un graphe est noté $G(\mathbf{i}, \mathbf{j})$, où $\mathbf{i} = (i_1, \dots, i_k)$ et $\mathbf{j} = (j_1, \dots, j_k)$. Un exemple de Δ -graphe est donné sur la figure 2.1.*

Deux graphes sont dits isomorphes si l'un devient l'autre par une permutation appropriée sur $(1, 2, \dots, p)$ et une permutation adaptée sur $(1, 2, \dots, n)$. Pour chaque classe d'isomorphisme, il n'y a qu'un graphe, appelé canonique, satisfaisant $i_1 = j_1 = 1$, $i_u \leq \max(i_1, \dots, i_{u-1}) + 1$, et $j_u \leq \max(j_1, \dots, j_{u-1}) + 1$. Un Δ -graphe canonique $G(\mathbf{i}, \mathbf{j})$ est noté par $\Delta(K, r, s)$ si G a $r + 1$ I -sommets qui ne coïncident pas et s J -sommets qui ne coïncident pas.

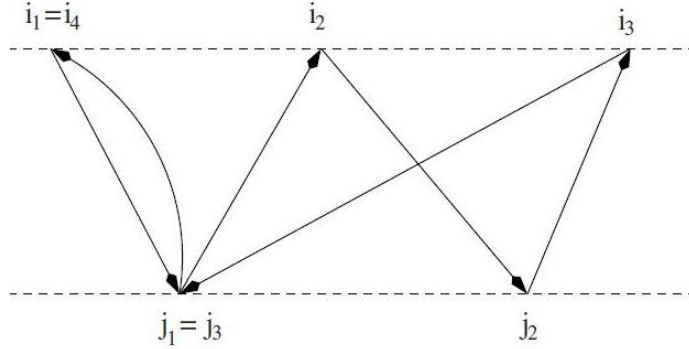


FIGURE 2.1 – Exemple d'un Δ -graphe

Un $\Delta(K, r, s)$ canonique peut être directement défini de la manière suivante :

1. Son ensemble de sommets $V = V_I + V_J$, où $V_I = (1, \dots, r+1)$ est appelé I -sommets, et $V_J = (1, \dots, s)$ est appelé J -sommets.
2. Il existe deux fonctions $f : (1, \dots, k) \mapsto (1, \dots, r+1)$ et $g : (1, \dots, k) \mapsto (1, \dots, s)$, satisfaisant

$$\begin{aligned} f(1) &= 1 = g(1) = f(k+1), \\ f(i) &\leq \max(f(1), \dots, f(i-1)) + 1, \\ g(j) &\leq \max(g(1), \dots, g(j-1)) + 1. \end{aligned}$$

3. Son ensemble de bords $E = (e_{1d}, e_{1u}, \dots, e_{kd}, e_{ku})$, où e_{1d}, \dots, e_{kd} sont appelés les bords bas et les e_{1u}, \dots, e_{ku} sont appelés bords haut.
4. $F(e_{jd}) = (f(j), g(j))$ et $F(e_{ju}) = (g(j), f(j+1))$ pour $j = (1, \dots, k)$.

Dans le cas où $f(j+1) = \max(f(1), \dots, f(j)) + 1$, l'arête $e_{j,u}$ est appelée "innovation haute", et dans le cas où $g(j) = \max(g(1), \dots, g(j-1)) + 1$, le bord $e_{j,d}$ est appelé "innovation basse". Intuitivement, une innovation haute conduit à un nouveau I -sommet et une innovation basse conduit à un nouveau

J -sommet. Nous prenons pour convention que, le premier segment vers le bas est une innovation basse et que le dernier segment vers le haut n'est pas une innovation.

On classe les $\Delta(k, r, s)$ graphes en trois catégories :

1. $\Delta_1(k, r)$: Δ -graphes dans lesquels chaque segment allant vers le bas ne doit coïncider qu'avec un et un seul segment qui va vers le haut. Si on joint les segments qui coïncident, le graphe résultant est un arbre à k branches. Ici, $r + s = k$, et pour la simplicité, on supprime le s .
2. $\Delta_2(k, r, s)$: Δ -graphes qui contiennent au moins un segment "seul"
3. $\Delta_3(k, r, s)$: Δ -graphes qui n'appartiennent ni à $\Delta_1(k, r)$, ni à $\Delta_2(k, r, s)$.

Le nombre de graphes dans chaque isomorphisme pour un $\Delta(kr, s)$ -graphe canonique est donné par le lemme suivant :

Lemme 2. *Pour k, r et s donnés, le nombre de graphes dans une classe isomorphique pour chaque $\Delta(k, r, s)$ -graphe canonique est :*

$$p(p-1)\dots(p-r)n(n-1)\dots(n-s+1) = p^{r+1}n^s[1 + o(n^{-1})] \quad (2.6)$$

Pour un Δ_3 -graphe, nous avons le lemme suivant :

Lemme 3. *Le nombre total de sommets qui ne coïncident pas dans un $\Delta_3(k, r, s)$ -graphe est inférieur ou égal à k .*

Pour un Δ_1 -graphe, on a :

Lemme 4. *Pour k et r , le nombre de $\Delta_1(k, r)$ -graphes est*

$$\frac{1}{r+1} \binom{k}{r} \binom{k-1}{r}. \quad (2.7)$$

2.4.2.3 Preuve de la loi de M-P dans le cas *i.i.d.* : preuve du Théorème 2

Preuve. Cette preuve se décompose principalement en 2 étapes. La première comporte : troncature, centralisation, remise à échelle, la deuxième est une preuve de la loi de M-P par le théorème de convergence des moments.

Objectif final : Notre objectif principal est d'utiliser tous les résultats et les lemmes énoncés précédemment pour montrer que :

$$\beta_k(\mathbf{S}_n) := \int x^k F^{\mathbf{S}_n}(dx) = \beta_k := \int x^k p_y(x) dx \quad (2.8)$$

où $p_y(x)$ est défini dans l'équation 2.5.

Première étape. Troncature, Centralisation, Remise à échelle

Soit C un nombre positif. Définissons :

$$\begin{aligned} \hat{x}_{ij} &= x_{ij} \mathbf{1}_{\{|x_{ij}| \leq C\}}, \\ \tilde{x}_{ij} &= \hat{x}_{ij} - \mathbb{E}(\hat{x}_{11}), \\ \hat{\mathbf{x}}_i &= (\hat{x}_{i1}, \dots, \hat{x}_{ip})', \\ \tilde{\mathbf{x}}_i &= (\tilde{x}_{i1}, \dots, \tilde{x}_{ip})', \\ \hat{\mathbf{S}}_n &= \frac{1}{n} \sum_{i=1}^n \hat{\mathbf{x}}_i \hat{\mathbf{x}}_i^* = \frac{1}{n} \hat{\mathbf{X}} \hat{\mathbf{X}}^*, \\ \tilde{\mathbf{S}}_n &= \frac{1}{n} \sum_{i=1}^n \tilde{\mathbf{x}}_i \tilde{\mathbf{x}}_i^* = \frac{1}{n} \tilde{\mathbf{X}} \tilde{\mathbf{X}}^*. \end{aligned}$$

Notons $F^{\hat{\mathbf{S}}_n} = \text{ESD de } \hat{\mathbf{S}}_n$ et $F^{\tilde{\mathbf{S}}_n} = \text{ESD de } \tilde{\mathbf{S}}_n$. Utilisant le Corollaire 6, on obtient que :

$$L^4(F^{\mathbf{S}}, F^{\hat{\mathbf{S}}_n}) \longrightarrow 4\mathbb{E}(|x_{ij}^2| \mathbf{1}_{\{|x_{ij}| \leq C\}}) \text{ a.s.} \quad (2.9)$$

où L est la distance de Levy entre les 2 ESD (chacune de dimension 2). Pour deux fonctions de distribution F et G , L est défini comme suit :

$$L(F, G) = \inf\{\epsilon : F(x - \epsilon, x - \epsilon) - \epsilon \leq G(x, y) \leq F(x + \epsilon, x + \epsilon) + \epsilon\} \quad (2.10)$$

Notons que la convergence métrique en L implique la convergence en distribution.

En effet :

$$\begin{aligned} L^4(F^S, F^{\hat{S}_n}) &\leq \left(\frac{2}{np} \sum_{ij} (|x_{ij}^2| + |\hat{x}_{ij}^2|) \right) \left(\frac{1}{np} \sum_{ij} (|x_{ij} - \hat{x}_{ij}|^2) \right) \\ &\leq \left(\frac{2}{np} \sum_{ij} (|x_{ij}^2| + |x_{ij}^2|) \right) \left(\frac{1}{np} \sum_{ij} (|x_{ij} - x_{ij} \mathbf{1}_{\{|x_{ij}| \leq C\}}|^2) \right) \\ &\leq \left(\frac{4}{np} \sum_{ij} |x_{ij}^2| \right) \left(\frac{1}{np} \sum_{ij} (|x_{ij}(1 - \mathbf{1}_{\{|x_{ij}| \leq C\}})|^2) \right) \\ &\leq \left(\frac{4}{np} \sum_{ij} |x_{ij}^2| \right) \left(\frac{1}{np} \sum_{ij} (|x_{ij} \mathbf{1}_{\{|x_{ij}| > C\}}|^2) \right) \\ &\leq \left(\frac{4}{np} \sum_{ij} |x_{ij}^2| \right) \left(\frac{1}{np} \sum_{ij} (|x_{ij}^2| \mathbf{1}_{\{|x_{ij}| > C\}}) \right) \\ &\leq \frac{4}{n^2 p^2} \left(\sum_{ij} |x_{ij}^2| \right) \left(\sum_{ij} (|x_{ij}^2| \mathbf{1}_{\{|x_{ij}| > C\}}) \right) \end{aligned}$$

d'où le résultat.

Par le Théorème 16 cité en Annexe, on a :

$$\|F^{\hat{S}_n} - F^{\tilde{S}_n}\| \leq \frac{1}{p} \text{rank}(\mathbb{E}\hat{X}) = \frac{1}{p}. \quad (2.11)$$

En effet, utilisant les définitions de \hat{S}_n et \tilde{S}_n et le Théorème sus-cité, on a :

$$\begin{aligned} \|F^{\hat{S}_n} - F^{\tilde{S}_n}\| &\leq \frac{1}{p} \text{rank}(\hat{X} - \tilde{X}) \\ &\leq \frac{1}{p} \text{rank}(\hat{X} - (\hat{X} - \mathbb{E}\hat{X})) \\ &\leq \frac{1}{p} \text{rank}(\mathbb{E}\hat{X}) = \frac{1}{p}. \end{aligned}$$

Posons $\tilde{\sigma}^2 = \mathbb{E}(|\tilde{x}_{jk}|^2) \rightarrow 1$ lorsque $C \rightarrow \infty$. Utilisant le Corollaire 6, on peut montrer (comme dans le cas précédent) que :

$$L^4(F^{\tilde{S}_n}, F^{\tilde{\sigma}^{-2}\tilde{S}_n}) \longrightarrow 2(1 - \tilde{\sigma}^4), \text{ p.s.} \quad (2.12)$$

Deuxième étape. Preuve de la loi de M-P par le Théorème de Convergence des Moments (MCT)

Toutes les conditions de la première étape étant remplies, on peut maintenant utiliser le théorème de convergence des moments pour montrer la loi de M-P. On a :

$$\begin{aligned} \beta_k(S_n) &= \int x^k F^{S_n}(dx) \\ &= p^{-1} n^{-k} \sum_{\{i_1, \dots, i_k\}} \sum_{\{j_1, \dots, j_k\}} x_{i_1 j_1} \bar{x}_{i_2 j_1} x_{i_2 j_2} \bar{x}_{i_3 j_2} \dots x_{i_k j_k} \bar{x}_{i_1 j_k} \\ &:= p^{-1} n^{-k} \sum_{\mathbf{i}, \mathbf{j}} X_{G(\mathbf{i}, \mathbf{j})}, \end{aligned}$$

où les indices $\mathbf{i} = (i_1, \dots, i_k)$ prennent leurs valeurs entre $1, 2, \dots, p$, et les indices $\mathbf{j} = (j_1, \dots, j_k)$ prennent leurs valeurs entre $1, 2, \dots, n$.

Il reste à montrer que :

$$\mathbb{E}(\beta_k(\mathbf{S}_n)) = p^{-1}n^{-k} \sum_{\mathbf{i}, \mathbf{j}} \mathbb{E}(x_{G(\mathbf{i}, \mathbf{j})}) \quad (2.13)$$

$$= \sum_{r=0}^{k-1} \frac{y_n^r}{r+1} \binom{k}{r} \binom{k-1}{r} + O(n^{-1}) \quad (2.14)$$

(où $y_n = \frac{p}{n}$) et que :

$$\begin{aligned} \mathbb{V}(\beta_k(\mathbf{S}_n)) &= p^{-2}n^{-2k} \sum_{\mathbf{i}_1, \mathbf{j}_1, \mathbf{i}_2, \mathbf{j}_2} [\mathbb{E}(x_{G_1(\mathbf{i}_1, \mathbf{j}_1)} x_{G_2(\mathbf{i}_2, \mathbf{j}_2)} - \mathbb{E}(x_{G_1(\mathbf{i}_1, \mathbf{j}_1)}) \mathbb{E}(x_{G_2(\mathbf{i}_2, \mathbf{j}_2)})] \\ &= O(n^{-2}). \end{aligned} \quad (2.15)$$

(où les graphes G_1 et G_2 sont respectivement définis par $(\mathbf{i}_1, \mathbf{j}_1)$ et $(\mathbf{i}_2, \mathbf{j}_2)$).

La preuve de l'expression (2.14) se fait en utilisant l'indépendance entre X_{G_1} et $X_{G_2} \dots$ La preuve de (2.13) se fait comme suit :

Le Lemme 2 nous dit qu'on peut ré-écrire l'équation (2.12) sous la forme

$$\begin{aligned} \mathbb{E}(\beta_k(\mathbf{S}_n)) &= p^{-1}n^{-k} \sum_{\mathbf{i}, \mathbf{j}} \mathbb{E}(x_{G(\mathbf{i}, \mathbf{j})}) \\ &= p^{-1}n^{-k} \sum_{\Delta(k, r, s)} p(p-1) \dots (p-r)n(n-1) \dots (n-s+1) \mathbb{E}(X_{\Delta(k, r, s)}). \end{aligned} \quad (2.16)$$

L'espérance de l'équation (2.16) peut être écrite sous la forme d'une somme de 3 espérances, chacune correspondant à un Δ_i -graphe canonique (i prenant ses valeurs entre 1, 2 et 3). Ceci implique :

$$\mathbb{E}(\beta_k(\mathbf{S}_n)) = p^{-1}n^{-k} \sum_{\Delta(k, r, s)} K \mathbb{E}(X_{\Delta_1(k, r)}) + K \mathbb{E}(X_{\Delta_2(k, r, s)}) + K \mathbb{E}(X_{\Delta_3(k, r, s)}) \quad (2.17)$$

où $K = p(p-1) \dots (p-r)n(n-1) \dots (n-s+1)$.

➡ Par définition d'un $\Delta_2(k, r, s)$ -graphe (ne contient qu'un et un seul segment), on a :

$$\mathbb{E}(X_{\Delta_2(k, r, s)}) = 0. \quad (2.18)$$

➡ Pour un Δ_3 -graphe, par le Lemme 3, nous avons que $r+s < k$. Étant donné que la variable $x_{\Delta_3(k, r, s)} \leq \left(\frac{2C}{\tilde{\sigma}}\right)^{2k}$, on conclut que

$$p^{-1}n^{-k} \sum_{\Delta_3(k, r, s)} K \mathbb{E}(X_{\Delta_3(k, r, s)}) = O(n^{-1}). \quad (2.19)$$

➡ Pour un Δ_1 -graphe ($s = k - r$), étant donné que $\mathbb{E}(X_{\Delta_1(k, r)}) = 1$, on a :

$$p^{-1}n^{-k} \sum_{\Delta_1(k, r)} K \mathbb{E}(X_{\Delta_1(k, r)}) = p^{-1}n^{-k} \sum_{\Delta_1(k, r)} K. \quad (2.20)$$

Ainsi, par le Lemme 4, on a finalement :

$$\begin{aligned}
\mathbb{E}(\beta_k(\mathbf{S}_n)) &= p^{-1}n^{-k} \sum_{\Delta_1(k,r)} K \\
&= p^{-1}n^{-k} \sum_{\Delta_1(k,r)} p(p-1)\dots(p-r)n(n-1)\dots(n-s+1) \\
&= p^{-1}n^{-k} \sum_{r=0}^{k-1} \frac{1}{r+1} \binom{k}{r} \binom{k-1}{r} p^{r+1} n^s [1 + O(n^{-1})] \\
&= \sum_{r=0}^{k-1} \frac{1}{r+1} \binom{k}{r} \binom{k-1}{r} p^r n^{s-k} + O(n^{-1}) \\
&= \sum_{r=0}^{k-1} \frac{1}{r+1} \binom{k}{r} \binom{k-1}{r} p^r n^{-r} + O(n^{-1}) \\
&= \sum_{r=0}^{k-1} \frac{1}{r+1} \binom{k}{r} \binom{k-1}{r} \left(\frac{p}{n}\right)^r + O(n^{-1}) \\
&= \sum_{r=0}^{k-1} \frac{y_n^r}{r+1} \binom{k}{r} \binom{k-1}{r} + O(n^{-1}) \\
&:= \beta_k.
\end{aligned}$$

où $y_n = \frac{p}{n}$. Ce résultat montre la convergence en loi des $\beta_k(\mathbf{S}_n)$ vers les β_k . D'où la preuve du théorème. □

2.4.2.4 Quelques exemples

Pour illustrer les Théorèmes 1 et/ou 2, nous avons appliqué les résultats pour des variables aléatoires qui suivent diverses lois. Lorsque X suit une certaine loi donnée, en utilisant la normalisation

$$Y := \frac{X - \mu}{\sigma}, \quad (2.21)$$

où μ et σ sont respectivement la moyenne et la variance de X , on trouve que la densité de Y correspond à la densité de Marcenko-Pastur. Cette normalisation connue sous le nom de Théorème Central Limite, est très utilisée en statistique pour établir la convergence en loi de la somme d'une suite de variables aléatoires vers la loi normale. Intuitivement, ce résultat affirme que toute somme de variables aléatoires indépendantes et identiquement distribuées tend vers une variable aléatoire gaussienne. Dans la littérature, il est dit que pour des variables i.i.d., la densité de Marcenko-Pastur approche bien l'histogramme des valeurs propres de la matrice de covariance d'un certain échantillon. La démarche sera la suivante :

- On va générer des variables aléatoires i.i.d. qui suivent une certaine loi.
- On utilise la formule 2.21 pour centrer les variables aléatoires.
- On utilise la formule 4 pour obtenir la matrice de covariance.
- On calcule les valeurs propres de la matrice de covariance précédente.
- On trace l'histogramme de ces valeurs propres et on compare l'allure de cet histogramme à la densité de M-P.

Pour illustrer les résultats, les auteurs considèrent généralement que la moyenne est nulle et que la variance vaut 1. Nous commencerons donc par ce cas, et nous introduirons ensuite des moyennes et des variances non nulles pour vérifier que les résultats collent toujours avec la théorie.

2.4.2.4.1 Cas d'une $\mathcal{N}(0, 1)$

Dans cette partie, on considère que les variables aléatoires suivent une loi normale centrée réduite. On sait que la densité de probabilité d'une telle loi est donnée par

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (2.22)$$

(étant donné que la moyenne μ vaut 0 et la variance σ^2 vaut 1). Sous MATLAB, en utilisant la commande `randn` en plus de la formule 4, nous avons généré de telles variables, et calculé la matrice de variance-covariance pour effectuer le test. La figure 2.2 qui suit nous montre les résultats obtenus :

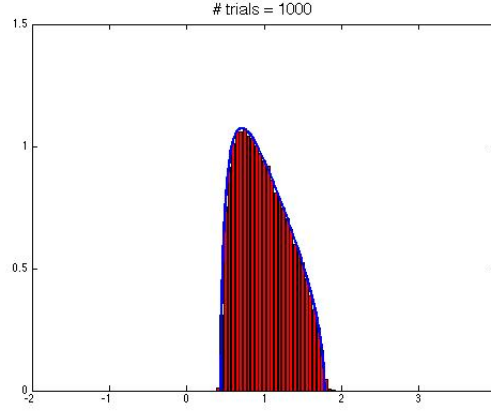


FIGURE 2.2 – Histogramme des valeurs propres d’une matrice dont les données suivent une $\mathcal{N}(0, 1)$, et densité de Marcenko-Pastur

Le résultat obtenu sur la figure 2.2 montre que la densité de Marcenko-Pastur colle bien avec l’histogramme des valeurs propres. Ce qui correspond bien aux résultats obtenus dans la théorie. Nous allons généraliser cette loi en lui attribuant une moyenne différente de zéro et une variance différente de 1.

2.4.2.4.2 Cas d’une $\mathcal{N}(\mu, \sigma^2 \mathbf{I}_n)$

Ici, on va considérer une généralisation de la loi considérée dans la partie 2.4.2.4.1. La densité dans ce cas est donnée par

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (2.23)$$

Sous MATLAB, pour générer de telles variables, il suffit de générer des variables qui suivent une $\mathcal{N}(0, 1)$, les multiplier par la variance, et y rajouter la moyenne. Utilisant la formule 2.21, on obtient la figure qui suit :

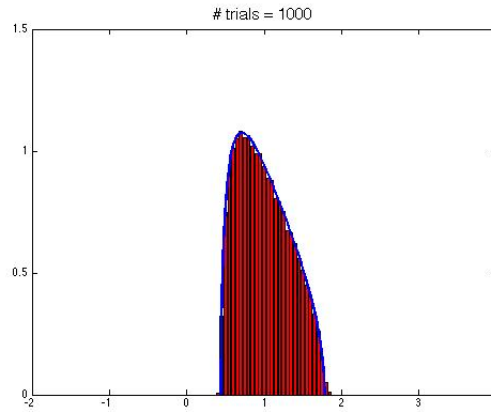


FIGURE 2.3 – Histogramme des valeurs propres d’une matrice dont les données suivent une $\mathcal{N}(\mu, \sigma^2 \mathbf{I}_n)$, et densité de Marcenko-Pastur

Le constat est le même que dans le cas précédent. On remarque que la densité de Marcenko-Pastur approche très bien l’histogramme des valeurs propres. Nous allons actuellement prendre une loi un peu

plus complexe que la loi normale. Cette loi est le logarithme de la loi normale, connue sous le nom de loi lognormale.

2.4.2.4.3 Cas d'une loi lognormale

On sait que la densité d'une loi lognormale s'écrit

$$y = f(x|\mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad (2.24)$$

où μ et σ sont ses paramètres (et non sa moyenne et sa variance). Pour utiliser le **Théorème Central Limite**, nous devons calculer sa moyenne m et sa variance v . Pour cela, on utilise les formules suivantes :

$$m = \exp\left(\mu + \frac{\sigma^2}{2}\right)$$

$$v = \exp(2\mu + \sigma^2) (\exp(\sigma^2) - 1).$$

Il est possible et il serait même préférable de trouver les paramètres μ et σ d'une loi de lognormale, à partir de sa moyenne m et de sa variance v . Étant donné la moyenne m et la variance v d'une loi lognormale, les formules à utiliser pour trouver les paramètres μ et σ sont les suivantes :

$$\mu = \log\left(\frac{m^2}{\sqrt{v + m^2}}\right)$$

$$\sigma = \sqrt{\log\left(\frac{v}{m^2 + 1}\right)}.$$

Étant données une moyenne et une variance, nous avons calculé les paramètres de la loi lognormale, et en utilisant la commande MATLAB *lognrnd*, nous avons généré des variables qui suivent la loi lognormale. Utilisant le résultat 2.21, on obtient la figure 2.4 ci-dessous :

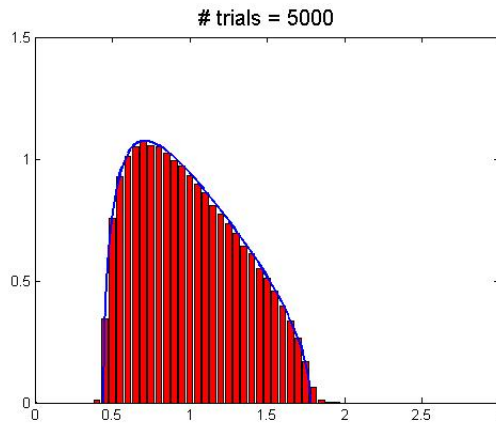


FIGURE 2.4 – Histogramme des valeurs propres d'une matrice dont les données suivent une loi lognormale, et densité de Marcenko-Pastur

On note également qu'ici, la densité de Marcenko-Pastur approche très bien l'histogramme des valeurs propres. On va encore aller plus loin en considérant des variables aléatoires qui suivent une loi de Student.

2.4.2.4.4 Cas d'une loi de Student

On rappelle que, la loi de Student est une loi de probabilité, faisant intervenir le quotient entre une variable suivant une loi normale centrée réduite et la racine carrée d'une variable distribuée suivant la loi du χ^2 . La densité d'une loi de Student à k degrés de libertés s'écrit

$$f(x) = \frac{1}{\sqrt{k\pi}} \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} \left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}}, \text{ pour } k > 0, \quad (2.25)$$

où Γ est la fonction Gamma d'Euler. Notons que pour $k \leq 2$, sa variance est infinie, et vaut $\frac{k}{k-2}$ pour $k > 2$. Pour nos tests, nous nous focaliserons sur les cas où la variance est finie. On a aussi que, pour $k = 1$, l'espérance d'une loi de Student à k degrés de liberté ne peut pas être définie, et pour $k > 1$, elle vaut 0. La commande MATLAB *trnd* et la formule 2.21 nous ont permis d'obtenir la figure 2.5 ci-dessous :

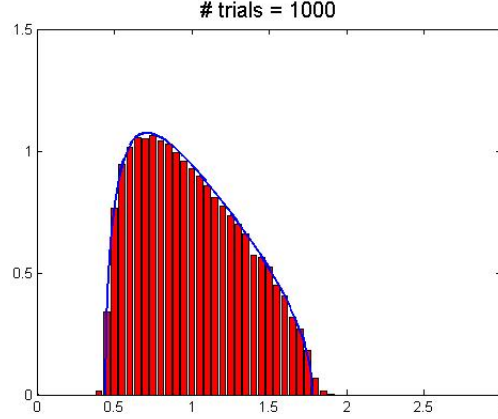


FIGURE 2.5 – Histogramme des valeurs propres d'une matrice dont les données suivent une loi de Student, et densité de Marcenko-Pastur

on remarque que dans ce cas aussi, la densité de Marcenko-Pastur colle encore avec l'histogramme des valeurs propres.

On peut penser qu'en rendant de plus en plus complexe la loi, mais en gardant cette hypothèse de i.i.d. sur les variables aléatoires, on peut toujours s'arranger à faire converger l'histogramme des valeurs propres de la matrice de variance-covariance correspondante vers la densité de la loi de Marcenko-Pastur. Mais que se passe-t-il si on enlève l'hypothèse d'indépendance sur les variables aléatoires ?

2.4.3 Généralisation de la loi de M-P dans le cas non iid

Il est quelques fois intéressant de considérer le cas où les entrées de X_n dépendent de n et pour chaque n , sont indépendantes mais pas identiquement distribuées. Pour ce cas, on a le théorème suivant :

Théorème 3. *Supposons que, pour chaque n , les entrées de X sont des variables complexes indépendantes, avec moyenne μ et variance σ^2 . Supposons $\frac{p}{n} \rightarrow y \in (0, \infty)$, et que, $\forall \eta > 0$,*

$$\frac{1}{\eta^2 np} \sum_{jk} \mathbb{E} \left(|x_{jk}^{(n)}|^2 \mathbf{1}_{\{|x_{jk}^{(n)}| \geq \eta \sqrt{n}\}} \right) \rightarrow 0. \quad (2.26)$$

Alors, avec une probabilité de 1, F^S tend vers la loi de Marcenko-Pastur avec index ration y et indice d'échelle σ^2 .

Preuve. Pour montrer ce théorème, nous allons utiliser la transformée de Stieltjes définie en Annexe. Rappelons que la transformée de Stieltjes de la loi de Marcenko-Pastur s'écrit

$$s_y(z) = -\frac{y + z - 1 - \sqrt{(1 + y - z)^2 - 4y}}{2yz} \quad (2.27)$$

Notre but est de montrer qu'ici, on retrouve bien le même résultat que dans l'équation (B.18), sous certaines hypothèses qu'on mettra bien en avant.

On sait que la transformée de Stieltjes de l'ESD de \mathbf{S}_n , qui est notée s_n , s'écrit :

$$s_n(z) = \frac{1}{p} \text{tr}(\mathbf{S}_n - zI_p)^{-1}.$$

La preuve de ce théorème se fera en trois principales étapes :

1. $\forall z \in \mathbb{C}^+$ fixé, $s_n(z) - \mathbb{E}s_n(z) \rightarrow 0$, presque sûrement.
2. $\forall z \in \mathbb{C}^+$ fixé, $\mathbb{E}s_n(z) \rightarrow s(z)$, la transformée de Stieltjes de la loi de M-P.
3. Sauf pour l'ensemble vide, $s_n(z) \rightarrow s(z) \forall z \in \mathbb{C}^+$.

La troisième étape de la preuve s'obtient des deux premières étapes. Donc on ne fera pas la preuve.

On prendra pour convention :

$$\begin{cases} \mathbb{E}_n s_n(s) &= \mathbb{E}s_n(z) \\ \mathbb{E}_0 s_n(z) &= s_n(z) \end{cases} \quad (2.28)$$

Première étape : Convergence presque sûre de la partie aléatoire

$$s_n(z) - \mathbb{E}s_n(z) \rightarrow 0, \text{ p.s.} \quad (2.29)$$

Notons $\mathbb{E}_k(\cdot)$ l'espérance conditionnelle sachant $\{x_{k+1}, \dots, x_n\}$. Par la formule

$$(A + \alpha\beta^*)^{-1} = A^{-1} - \frac{A^{-1}\alpha\beta^*A^{-1}}{1 + \beta^*A^{-1}\alpha} \quad (2.30)$$

on obtient en utilisant les conventions du système (2.28)

$$\begin{aligned} s_n(z) - \mathbb{E}s_n(z) &= \frac{1}{p} \sum_{k=1}^n [\mathbb{E}_{k-1} \text{tr}(\mathbf{S}_n - zI_p)^{-1} - \mathbb{E}_k \text{tr}(\mathbf{S}_n - zI_p)^{-1}] \\ &= \frac{1}{p} \sum_{k=1}^n \gamma_k, \end{aligned}$$

où, par le Théorème 11,

$$\begin{aligned} \gamma_k &= \mathbb{E}_{k-1} \text{tr}(\mathbf{S}_n - zI_p)^{-1} - \mathbb{E}_k \text{tr}(\mathbf{S}_n - zI_p)^{-1} \\ &= \mathbb{E}_{k-1} [\text{tr}(\mathbf{S}_n - zI_p)^{-1} - \text{tr}(\mathbf{S}_{nk} - zI_p)^{-1}] - \mathbb{E}_k [\text{tr}(\mathbf{S}_n - zI_p)^{-1} - \text{tr}(\mathbf{S}_{nk} - zI_p)^{-1}] \\ &= (\mathbb{E}_{k-1} - \mathbb{E}_k) [\text{tr}(\mathbf{S}_n - zI_p)^{-1} - \text{tr}(\mathbf{S}_{nk} - zI_p)^{-1}] \end{aligned} \quad (2.31)$$

où $\mathbf{S}_{nk} = \mathbf{S}_n - x_k x_k^*$. Appliquant l'égalité (2.30) à $(\mathbf{S}_n - zI_p)^{-1} = (\mathbf{S}_{nk} + x_k x_k^* - zI_p)^{-1}$, on obtient

$$\begin{aligned} (\mathbf{S}_n - zI_p)^{-1} &= (\mathbf{S}_{nk} + x_k x_k^* - zI_p)^{-1} \\ &= (\mathbf{S}_{nk} - zI_p + x_k x_k^*)^{-1} \\ &= (\mathbf{S}_{nk} - zI_p)^{-1} - \frac{(\mathbf{S}_{nk} - zI_p)^{-1} x_k x_k^* (\mathbf{S}_{nk} - zI_p)^{-1}}{1 + x_k x_k^* (\mathbf{S}_{nk} - zI_p)^{-1}} \end{aligned} \quad (2.32)$$

Utilisant l'hypothèse selon laquelle les entrées de X commutent, on a finalement

$$\begin{aligned} (\mathbf{S}_n - zI_p)^{-1} &= (\mathbf{S}_{nk} - zI_p)^{-1} - \frac{(\mathbf{S}_{nk} - zI_p)^{-1} x_k x_k^* (\mathbf{S}_{nk} - zI_p)^{-1}}{1 + x_k x_k^* (\mathbf{S}_{nk} - zI_p)^{-1}} \\ &= (\mathbf{S}_{nk} - zI_p)^{-1} - \frac{x_k^* (\mathbf{S}_{nk} - zI_p)^{-2} x_k}{1 + x_k^* (\mathbf{S}_{nk} - zI_p)^{-1} x_k}. \end{aligned} \quad (2.33)$$

Utilisant le résultat (2.33) dans (2.31), on obtient :

$$\begin{aligned} \gamma_k &= (\mathbb{E}_{k-1} - \mathbb{E}_k) [\text{tr}(\mathbf{S}_n - zI_p)^{-1} - \text{tr}(\mathbf{S}_{nk} - zI_p)^{-1}] \\ &= (\mathbb{E}_{k-1} - \mathbb{E}_k) \left[\text{tr}(\mathbf{S}_{nk} - zI_p)^{-1} - \text{tr} \left(\frac{x_k^* (\mathbf{S}_{nk} - zI_p)^{-2} x_k}{1 + x_k^* (\mathbf{S}_{nk} - zI_p)^{-1} x_k} \right) - \text{tr}(\mathbf{S}_{nk} - zI_p)^{-1} \right] \\ &= -(\mathbb{E}_{k-1} - \mathbb{E}_k) \text{tr} \left(\frac{x_k^* (\mathbf{S}_{nk} - zI_p)^{-2} x_k}{1 + x_k^* (\mathbf{S}_{nk} - zI_p)^{-1} x_k} \right) \\ &= -(\mathbb{E}_{k-1} - \mathbb{E}_k) \frac{x_k^* (\mathbf{S}_{nk} - zI_p)^{-2} x_k}{1 + x_k^* (\mathbf{S}_{nk} - zI_p)^{-1} x_k} \end{aligned}$$

Notons que

$$\left| \frac{x_k^*(\mathbf{S}_{nk} - zI_p)^{-2}x_k}{1 + x_k^*(\mathbf{S}_{nk} - zI_p)^{-1}x_k} \right| \leq \frac{1}{v}. \quad (2.34)$$

Remarquant que $\{\gamma_k\}$ forment une suite de différences de martingales bornées, par le Lemme 5 qui suit, en utilisant l'équation (2.36) du lemme 5 et lemme de Borel-Cantelli (voir Annexe A.9), on obtient le résultat de l'équation (2.29).

Lemme 5. *Soit $\{X_k\}$ une suite complexe de différence de martingale. Alors pour $p > 1$,*

$$\mathbb{E} \left| \sum X_k \right|^p \leq K_p \left(\sum |X_k|^2 \right)^{p/2}, \quad (2.35)$$

On a pour $p = 4$:

$$\begin{aligned} \mathbb{E} |s_n(z) - \mathbb{E}s_n(z)|^4 &\leq \frac{K_4}{p^4} \mathbb{E} \left(\sum_{k=1}^n |\gamma_k|^2 \right)^2 \\ &\leq \frac{4K_4 n^2}{v^4 p^4} = O(n^{-2}). \end{aligned} \quad (2.36)$$

Deuxième étape : Convergence moyenne

Nous allons montrer que

$$\mathbb{E}s_n(z) \rightarrow s(z), \quad (2.37)$$

où $s(z)$ est la transformée de Stieltjes de la loi de M-P, définie à l'équation (B.18), avec $\sigma^2 = 1$.

Par le Théorème 10, on a

$$s_n(z) = \frac{1}{p} \sum_{k=1}^p \frac{1}{\frac{1}{n} \alpha'_k \bar{\alpha}_k - z - \frac{1}{n^2} \alpha'_k X_k^* \left(\frac{1}{n} X_k X_k^* - zI_{p-1} \right)^{-1} X_k \bar{\alpha}_k}, \quad (2.38)$$

où X_k est la matrice obtenue de la matrice X en enlevant la ligne k , et α'_k ($n \times 1$) est la k^{eme} ligne de X . Posons

$$\epsilon_k = \frac{1}{n} \alpha'_k \bar{\alpha}_k - 1 - \frac{1}{n^2} \alpha'_k X_k^* \left(\frac{1}{n} X_k X_k^* - zI_{p-1} \right)^{-1} X_k \bar{\alpha}_k + y_n + y_n z \mathbb{E}s_n(z), \quad (2.39)$$

où $y_n = p/n$. Par (2.38), on a :

$$\mathbb{E}s_n(z) = \frac{1}{1 - z - y_n - y_n z \mathbb{E}s_n(z)} + \delta_n, \quad (2.40)$$

où

$$\delta_n = -\frac{1}{p} \sum_{k=1}^p \mathbb{E} \left(\frac{\epsilon_k}{(1 - z - y_n - y_n z \mathbb{E}s_n(z))(1 - z - y_n - y_n z \mathbb{E}s_n(z) + \epsilon_k)} \right). \quad (2.41)$$

Ces résultats se retrouvent facilement par calculs.

En résolvant (2.40) avec $\mathbb{E}s_n(z)$ comme variable, on obtient deux solutions :

$$\begin{aligned} s_1(z) &= \frac{1 - z - y_n + y_n z \delta_n + \sqrt{(1 - z - y_n - y_n z \delta_n)^2 - 4y_n z}}{2y_n z}, \\ s_2(z) &= \frac{1 - z - y_n + y_n z \delta_n - \sqrt{(1 - z - y_n - y_n z \delta_n)^2 - 4y_n z}}{2y_n z}. \end{aligned}$$

Comparant ces solutions à (2.27), il suffit de montrer que

$$\mathbb{E}s_n(z) = s_1(z) \quad (2.42)$$

et

$$\delta_n \rightarrow 0. \quad (2.43)$$

Montrons d'abord (2.42). Faisant $v \rightarrow \infty$, on sait que $\mathbb{E}s_n(z) \rightarrow 0$ et de là, $\delta_n \rightarrow 0$ par (2.40). Ceci montre que $\mathbb{E}s_n(z) = s_1(z)$ pour tout z avec une grande partie imaginaire. Si (2.42) n'est pas vrai pour tout $z \in \mathbb{C}^+$, alors, par continuité de s_1 et s_2 , il existe un $z_0 \in \mathbb{C}^+$ tel que $s_1(z_0) = s_2(z_0)$, ce qui implique

$$(1 - z_0 - y_n + y_n z_0 \delta_n)^2 - 4y_n z_0 (1 + \delta_n (1 - z_0 - y_n)) = 0.$$

Ainsi,

$$\mathbb{E}s_n(z_0) = s_1(z_0) = \frac{1 - z_0 - y_n + y_n z_0 \delta_n}{2y_n z_0}.$$

En substituant la solution δ_n de l'équation (2.40) dans l'équation ci-dessus, on obtient :

$$\mathbb{E}s_n(z_0) = \frac{1 - z_0 - y_n}{y_n z_0} + \frac{1}{y_n + z_0 - 1 + y_n z_0 \mathbb{E}s_n(z_0)}. \quad (2.44)$$

Notons que pour toute transformée de Stieltjes $s(z)$ de probabilité F définie sur \mathbb{R}^+ et y positif, on a

$$\begin{aligned} \Im(y + z - 1 + yzs(z)) &= \Im\left(z - 1 + \int_0^\infty \frac{yxdF(x)}{x - z}\right) \\ &= v\left(1 + \int_0^\infty \frac{yxdF(x)}{(x - u)^2 + v^2}\right) > 0. \end{aligned} \quad (2.45)$$

L'équation (2.45) nous permet d'affirmer que la partie imaginaire du deuxième terme de l'équation (2.44) est négative. Si $y_n \leq 1$, on montre facilement que $\Im\left(\frac{1 - z_0 - y_n}{y_n z_0}\right) < 0$. Ceci nous dit que $\Im \mathbb{E}s_n(z_0) < 0$, ce qui est impossible étant donné que la partie imaginaire d'une transformée de Stieltjes est toujours positive... **CONTRADICTION** qui implique que (2.42) est vrai dans le cas $y_n \leq 1$.

Dans le cas général, on peut utiliser la méthode suivante : Au vu des équations (2.44) et (2.45), on devrait avoir

$$y_n + z_0 - 1 + y_n z_0 \mathbb{E}s_n(z_0) = \sqrt{y_n z_0}. \quad (2.46)$$

En effet, de l'équation (2.44), on a :

$$\begin{aligned} \mathbb{E}s_n(z_0) - \frac{1 - z_0 - y_n}{y_n z_0} &= \frac{1}{y_n + z_0 - 1 + y_n z_0 \mathbb{E}s_n(z_0)} \\ \frac{\mathbb{E}s_n(z_0) y_n z_0 - 1 + z_0 + y_n}{y_n z_0} &= \frac{1}{y_n + z_0 - 1 + y_n z_0 \mathbb{E}s_n(z_0)} \\ (y_n + z_0 - 1 + y_n z_0 \mathbb{E}s_n(z_0))^2 &= y_n z_0 \\ y_n + z_0 - 1 + y_n z_0 \mathbb{E}s_n(z_0) &= \sqrt{y_n z_0}. \end{aligned}$$

Notons maintenant, \underline{s}_n , la transformée de Stieltjes de la matrice $\frac{1}{n} X^* X$. Sachant que $\frac{1}{n} X^* X$ et $S_n = \frac{1}{n} X X^*$ ont le même ensemble de valeurs propres non nulles, on a la relation suivante entre s_n et \underline{s}_n :

$$s_n(z) = \frac{\underline{s}_n}{y_n} - \frac{1 - \frac{1}{y_n}}{z}.$$

Notons que cette équation est vraie indépendamment du fait que $y_n > 1$ ou $y_n \leq 1$.

De là, on a

$$y_n - 1 + y_n z_0 \mathbb{E}s_n(z_0) = z_0 \mathbb{E}\underline{s}_n(z_0).$$

Substituant ce résultat dans (2.46), on obtient

$$\begin{aligned} z_0 + z_0 \mathbb{E}\underline{s}_n(z_0) &= \sqrt{y_n z_0} \\ z_0(1 + \mathbb{E}\underline{s}_n(z_0)) &= \sqrt{y_n z_0} \\ 1 + \mathbb{E}\underline{s}_n(z_0) &= \frac{\sqrt{y_n z_0}}{z_0} \\ 1 + \mathbb{E}\underline{s}_n(z_0) &= \frac{\sqrt{y_n}}{\sqrt{z_0}}, \end{aligned}$$

ce qui nous mène à une contradiction : La partie imaginaire de l'expression à gauche de l'égalité est positive alors que la partie imaginaire de l'expression à droite est négative. (2.42) est prouvée.

Prouvons maintenant (2.43). On a :

$$\begin{aligned}\delta_n &= -\frac{1}{p} \sum_{k=1}^p \mathbb{E} \left(\frac{\epsilon_k}{(1-z-y_n-y_n z \mathbb{E} s_n(z))(1-z-y_n-y_n z \mathbb{E} s_n(z) + \epsilon_k)} \right) \\ &= -\frac{1}{p} \sum_{k=1}^p \mathbb{E} \left(\frac{\mathbb{E} \epsilon_k}{(1-z-y_n-y_n z \mathbb{E} s_n(z))^2} \right) \\ &\quad + \frac{1}{p} \sum_{k=1}^p \mathbb{E} \left(\frac{\epsilon_k^2}{(1-z-y_n-y_n z \mathbb{E} s_n(z))^2 (1-z-y_n-y_n z \mathbb{E} s_n(z) + \epsilon_k)} \right) \\ &:= J_1 + J_2.\end{aligned}$$

Objectif : Montrer que $J_1 \rightarrow 0$ et $J_2 \rightarrow 0$.

Considérons que (hypothèses fixées) :

$$\begin{cases} |x_{ij}| \leq \eta_n \sqrt{n} \\ \mathbb{E}(x_{ij}) = 0 \\ \mathbb{V}ar(x_{ij}) = 1. \end{cases} \quad (2.47)$$

Utilisant (2.47) et les expressions (2.48) et (2.49) ci-dessous :

$$\mathrm{tr}(A^{-1}) - \mathrm{tr}(A_k^{-1}) = \frac{1 + \alpha'_k A_k^{-2} \alpha_k}{a_{kk} - \alpha'_k A_k^{-1} \alpha_k} \quad (2.48)$$

$$|\mathrm{tr}(A - zI_n)^{-1} - \mathrm{tr}(A_k - zI_{n-1})^{-1}| \leq v^{-1}, \quad (2.49)$$

on a :

$$\begin{aligned}|\mathbb{E} \epsilon_k| &= \left| -\frac{1}{n^2} \mathbb{E} \left\{ \mathrm{tr} \left[X_k^* \left(\frac{1}{n} X_k X_k^* - zI_{p-1} \right)^{-1} X_k \right] \right\} + y_n + y_n z \mathbb{E} s_n(z) \right| \\ &= \left| -\frac{1}{n} \mathbb{E} \left[\mathrm{tr} \left(\frac{1}{n} X_k X_k^* - zI_{p-1} \right)^{-1} \frac{1}{n} X_k X_k^* \right] + y_n + y_n z \mathbb{E} s_n(z) \right| \\ &\leq \frac{1}{n} + \frac{|z|y_n}{n} \left| \mathbb{E} \left[\mathrm{tr} \left(\frac{1}{n} X_k X_k^* - zI_{p-1} \right)^{-1} \right] - s_n(z) \right| \\ &\leq \frac{1}{n} + \frac{|z|y_n}{nv} \rightarrow 0, \end{aligned} \quad (2.50)$$

d'où on conclut que $J_1 \rightarrow 0$.

Pour terminer cette preuve, on va montrer que $J_2 \rightarrow 0$. On a :

$$\begin{aligned}\Im(1-z-y_n-y_n z \mathbb{E} s_n(z) + \epsilon_k) &= \Im \left(\frac{1}{n} \alpha'_k \bar{\alpha}_k - z - \frac{1}{n^2} \alpha'_k X_k^* \left(\frac{1}{n} X_k X_k^* - zI_{p-1} \right)^{-1} X_k \bar{\alpha}_k \right) \\ &= -v \left(1 + \frac{1}{n^2} \alpha'_k X_k^* \left[\left(\frac{1}{n} X_k X_k^* - zI_{p-1} \right)^2 + v^2 I_{p-1} \right]^{-1} X_k \bar{\alpha}_k \right) \\ &< -v\end{aligned}$$

Combinant ce résultat avec (2.45), on obtient

$$\begin{aligned}|J_2| &\leq \frac{1}{pv^3} \sum_{k=1}^p \mathbb{E} |\epsilon_k|^2 \\ &= \frac{1}{pv^3} \sum_{k=1}^p [\mathbb{E} |\epsilon_k - \tilde{\mathbb{E}} \epsilon_k|^2 + \mathbb{E} |\tilde{\mathbb{E}} \epsilon_k - \mathbb{E}(\epsilon_k)|^2 + (\mathbb{E}(\epsilon_k))^2],\end{aligned}$$

où $\tilde{\mathbb{E}}(\cdot)$ représente l'espérance conditionnelle sachant $\{\alpha_j, j = 1, \dots, k-1, k+1, \dots, p\}$. Dans l'estimation de J_1 , on a prouvé que

$$|\mathbb{E}\epsilon_k| \leq \frac{1}{n} + \frac{|z|y}{nv} \rightarrow 0.$$

Notons $A = (a_{ij}) = I_n - \frac{1}{n}X_k^*(\frac{1}{n}X_kX_k^* - zI_{p-1})^{-1}X_k$. Alors, on a :

$$\epsilon_k - \tilde{\mathbb{E}}\epsilon_k = \frac{1}{n} \left(\sum_{k=1}^n a_{ii}(|x_{ki}|^2 - 1) + \sum_{i \neq j} a_{ij}x_{ki}\bar{x}_{kj} \right).$$

Par des calculs élémentaires, on a

$$\begin{aligned} \frac{1}{n^2} \tilde{\mathbb{E}}|\epsilon'_k - \tilde{\mathbb{E}}\epsilon_k|^2 &= \frac{1}{n^2} \left(\sum_{i=1}^n |a_{ii}|^2 (\mathbb{E}|x_{ki}|^4 - 1) + \sum_{i \neq j} [|a_{ij}|^2 \mathbb{E}|x_{ki}|^2 \mathbb{E}|x_{kj}|^2 + a_{ij}^2 \mathbb{E}x_{ki}^2 \mathbb{E}x_{kj}^2] \right) \\ &\leq \frac{1}{n^2} \left(\sum_{i=1}^n |a_{ii}|^2 (\eta_n^2 n) + 2 \sum_{i \neq j} |a_{ij}|^2 \right) \\ &\leq \frac{\eta_n^2}{v^2} + \frac{2}{nv^2}, \end{aligned}$$

où on a utilisé le fait que $|a_{ii}| \leq v^{-1}$.

En utilisant la méthode de décomposition des martingales, on peut montrer que

$$\begin{aligned} \mathbb{E}|\tilde{\mathbb{E}}\epsilon_k - \mathbb{E}\epsilon_k|^2 &= \frac{|z|^2 y^2}{n^2} \mathbb{E} \left| \text{tr} \left(\frac{1}{n} X_k X_k^* - z I_{p-1} \right)^{-1} - \mathbb{E} \left[\text{tr} \left(\frac{1}{n} X_k X_k^* - z I_{p-1} \right)^{-1} \right] \right|^2 \\ &\leq \frac{|z|^2 y^2}{n^2 v^2} \rightarrow 0. \end{aligned}$$

En combinant les trois estimations ci-dessus, on montre la convergence moyenne de la transformée de Stieltjes de l'ESD de S_n .

D'où la preuve du théorème. \square

2.5 Application à la Théorie du Portefeuille

2.5.1 Rappels mathématiques

Avant de développer cette partie, voici quelques rappels mathématiques :

$$\mathbb{V}(AX) = A\mathbb{V}A',$$

$$\mathbb{V}(A + X) = \mathbb{V}(X).$$

Si la matrice A est symétrique, on a :

$$\frac{\partial X' A}{\partial X} = A,$$

$$\frac{\partial X' A X}{\partial X} = 2AX.$$

Si la matrice A est inversible, on a :

$$AB = C \iff A^{-1}AB = A^{-1}C \iff B = A^{-1}C.$$

2.5.2 portefeuille global de variance minimum [49]

Considérons un ensemble de N éléments. On définit par μ le vecteur des rendements espérés et par Σ la matrice de covariance correspondante. Le problème de Markowitz peut donc être défini comme suit :

$$\min \frac{1}{2} w' \Sigma w \tag{2.51}$$

sous les contraintes

$$\begin{cases} \mathbf{1}' w = 1 \\ w \in \Omega \cap \mathcal{C} \end{cases} \tag{2.52}$$

où w est le vecteur des poids dans le portefeuille et Ω est l'espace de recherche. Par exemple, si $\Omega = \mathbb{R}^n$, le problème d'optimisation définit le portefeuille global de variance minimum. Si $\Omega = \{w \in \mathbb{R}^n : \mu'w \geq \mu^*\}$, on obtient le portefeuille efficient où μ^* est le rendement espéré désiré de l'investisseur. Le portefeuille tangent est le portefeuille efficient qui maximise le rapport de Sharpe. \mathcal{C} définit l'ensemble des contraintes des poids. On considère deux définitions de \mathcal{C} :

1. $\mathcal{C} = \mathbb{R}^n$. Dans ce cas, la solution n'a pas de contrainte et on la note w^* ou $w^*(\mu, \Sigma)$.
2. Nous pouvons imposer certaines limites $w_i^- \leq w_i \leq w_i^+$ sur le poids de l'actif i . Dans ce cas, on note $\mathcal{C} = \mathcal{C}(w^-, w^+)$ et nous définissons \bar{w} comme la solution correspondante au problème d'optimisation.

Suivant le travail de Jagannathan et Ma [48], on peut montrer que les solutions sous contraintes peuvent être obtenues en résolvant le problème sans contraintes avec une autre spécification de μ et Σ . Nous avons également :

$$\tilde{w} = w^*(\tilde{\mu}, \tilde{\Sigma}) \quad (2.53)$$

Où $\tilde{\mu}$ et $\tilde{\Sigma}$ sont respectivement la perturbation du vecteur d'origine des rendements espérés μ et la perturbation de la matrice de covariance d'origine Σ . Habituellement, l'impact des contraintes des poids est analysé en étudiant la différence entre w^* et \tilde{w} .

Nous aimerions déterminer le portefeuille minimum variance (i.e. le portefeuille ayant la variance la plus faible) dans un cas de contraintes simples. Nous le calculerons à partir du programme d'optimisation suivant :

$$\min_w w' \Sigma w$$

sous la seule contrainte

$$w' e = 1$$

où $e = (1 \ 1 \ \dots \ 1)$. Le lagrangien pour ce problème s'écrit :

$$\mathcal{L} = w' \Sigma w + \mu(1 - w' e).$$

La première condition d'optimalité nous donne la relation suivante :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial w} = 0 &\iff 2\Sigma w - \mu e = 0 \\ &\iff w = \frac{\mu}{2} \Sigma^{-1} e. \end{aligned} \quad (2.54)$$

En la combinant avec la contrainte, on obtient :

$$\begin{aligned} w' e = 1 &\iff \left(\frac{\mu}{2} \Sigma^{-1} e \right)' e = 1 \\ &\iff \frac{\mu}{2} e' \Sigma^{-1} e = 1 \\ &\iff \mu e' \Sigma^{-1} e = 2 \\ &\iff \mu = \frac{2}{e' \Sigma^{-1} e}. \end{aligned} \quad (2.55)$$

En remplaçant 2.55 dans 2.54, on obtient les poids du portefeuille GMV w_g suivants :

$$\begin{aligned} w_g &= \frac{\mu}{2} \Sigma^{-1} e \\ &= \frac{1}{2} \frac{2}{e' \Sigma^{-1} e} \Sigma^{-1} e \\ w_g &= \frac{\Sigma^{-1} e}{e' \Sigma^{-1} e} \end{aligned} \quad (2.56)$$

Notre but est maintenant d'écrire l'équation 2.56 en fonction des valeurs propres et des vecteurs propres de la matrice Σ .

2.5.3 Écriture des coefficients d'une matrice symétrique en fonction de ses valeurs propres et de ses vecteurs propres

On sait que toute matrice réelle symétrique et définie positive est diagonalisable dans une base "orthonormée". Ceci signifie que, si A représente une matrice symétrique définie positive, alors, on peut la décomposer sous la forme $A = PDP^{-1}$, où D est la matrice de valeurs propres (matrice diagonale), et P la matrice de vecteurs propres associés à chaque valeur propre. A étant symétrique, on aura $P^{-1} = P^t$.

Cas 1 : Pour un début, considérons le cas d'une matrice de taille 2×2 :

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$$

Sous de bonnes hypothèses, considérons que λ_1 et λ_2 sont les valeurs propres de la matrice A , et que, $V_1 = (v_{11} \ v_{21})'$ et $V_2 = (v_{12} \ v_{22})'$ sont les vecteurs propres associés respectivement aux valeurs propres λ_1 et λ_2 . On aura donc

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \text{ et } P = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}$$

En effectuant le produit matriciel PDP^t , on obtient :

$$A = PDP^t = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{pmatrix}$$

En effectuant les calculs, on obtient finalement

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} = \begin{pmatrix} \lambda_1 v_{11}^2 + \lambda_2 v_{12}^2 & \lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22} \\ \lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22} & \lambda_1 v_{21}^2 + \lambda_2 v_{22}^2 \end{pmatrix}$$

Cas 2 : Dans le cas d'une matrice de taille 3×3 , on a :

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \text{ et } P = \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix}$$

Pour une bonne vision des choses, nous allons effectuer les détails des calculs. Calculons premièrement le produit $C := PD$. On a :

$$C = PD = \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = \begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} & \lambda_3 v_{13} \\ \lambda_1 v_{21} & \lambda_2 v_{22} & \lambda_3 v_{23} \\ \lambda_1 v_{31} & \lambda_2 v_{32} & \lambda_3 v_{33} \end{pmatrix}$$

Ainsi, étant donné que $A = CP^t$, on aura :

$$A = \begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} & \lambda_3 v_{13} \\ \lambda_1 v_{21} & \lambda_2 v_{22} & \lambda_3 v_{23} \\ \lambda_1 v_{31} & \lambda_2 v_{32} & \lambda_3 v_{33} \end{pmatrix} \begin{pmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{pmatrix}$$

Donc au final, après calculs, pour une matrice 3×3 , on trouve que la matrice A s'écrit

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} = \begin{pmatrix} \lambda_1 v_{11}^2 + \lambda_2 v_{12}^2 + \lambda_3 v_{13}^2 & \lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22} + \lambda_3 v_{13} v_{23} & \lambda_1 v_{11} v_{31} + \lambda_2 v_{12} v_{32} + \lambda_3 v_{13} v_{33} \\ \lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22} + \lambda_3 v_{13} v_{23} & \lambda_1 v_{21}^2 + \lambda_2 v_{22}^2 + \lambda_3 v_{23}^2 & \lambda_1 v_{21} v_{31} + \lambda_2 v_{22} v_{32} + \lambda_3 v_{23} v_{33} \\ \lambda_1 v_{11} v_{31} + \lambda_2 v_{12} v_{32} + \lambda_3 v_{13} v_{33} & \lambda_1 v_{21} v_{31} + \lambda_2 v_{22} v_{32} + \lambda_3 v_{23} v_{33} & \lambda_1 v_{31}^2 + \lambda_2 v_{32}^2 + \lambda_3 v_{33}^2 \end{pmatrix}.$$

Généralisation : Suivant cette logique, on trouve une généralisation du problème dans le cas où la matrice A est de taille $N \times N$. On obtient :

$$a_{ij} = \sum_{k=1}^N \lambda_k v_{ik} v_{jk} \quad \text{pour } i \leq j \quad (2.57)$$

où les $(a_{ij})_{i,j=1,\dots,N}$ sont les coefficients de la matrice A . La matrice A étant symétrique, on obtient facilement tous les autres coefficients.

2.5.4 Application du résultat de l'équation 2.57 à l'équation 2.56

Dans l'équation 2.56, la matrice Σ représente une matrice de covariance, donc symétrique définie positive. Étant donné qu'elle peut se décomposer suivant la forme $\Sigma = PDP^t$, où les matrices D et P ont été définies dans la partie 2.5.3, il est facile de trouver l'inverse de la matrice Σ . On a :

$$\begin{aligned}\Sigma = PDP^t &\iff \Sigma^{-1} = (PDP^t)^{-1} \\ &\iff \Sigma^{-1} = (P^t)^{-1}D^{-1}P^{-1} \\ &\iff \Sigma^{-1} = (P^{-1})^{-1}D^{-1}P^{-1},\end{aligned}\tag{2.58}$$

vu que $P^t = P^{-1}$. Étant donné que $(P^{-1})^{-1} = P$, on a que

$$\Sigma^{-1} = PD^{-1}P^t.\tag{2.59}$$

En appliquant les résultats de l'équation 2.57 à l'équation 2.59, et en gardant les mêmes notations, on obtient :

$$(\Sigma^{-1})_{ij} = \sum_{k=1}^N \frac{1}{\lambda_k} v_{ik} v_{jk}.\tag{2.60}$$

Ainsi, le dénominateur de l'équation 2.56 s'écrira

$$e\Sigma^{-1}e = \sum_{k=1}^N \frac{1}{\lambda_k} \sum_{i=1}^N \sum_{j=1}^N v_{ik} v_{jk}\tag{2.61}$$

car si on a une matrice

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

et un vecteur $e = (1 \ 1)'$, alors le produit $e'Ae$ n'est rien d'autre que la somme de tous les coefficients de la matrice A .

Notre but final étant de déterminer le poids i de notre portefeuille, il nous faudrait exprimer le numérateur de l'équation 2.56 sous forme indicelle. En utilisant le résultat de l'équation 2.60, on a,

$$(\Sigma^{-1}e)_i = \sum_{j=1}^N \frac{1}{\lambda_j} \sum_{k=1}^N v_{ik} v_{jk}.\tag{2.62}$$

En combinant les résultats des équations 2.61 et 2.62, on obtient finalement l'indice l du poids,

$$(w_g)_l = \frac{\sum_{j=1}^N \frac{1}{\lambda_j} \sum_{k=1}^N v_{lk} v_{jk}}{\sum_{k=1}^N \frac{1}{\lambda_k} \sum_{i=1}^N \sum_{j=1}^N v_{ik} v_{jk}}.\tag{2.63}$$

2.6 Débruitage des matrices de covariance et de corrélation : Reconstruction des matrices de corrélation et de variance-covariance

Comme nous l'avons dit dans l'introduction générale, la RMT a été utilisée pour la première fois pour expliquer les degrés d'énergie des systèmes nucléaires complexes. Elle a récemment été utilisée (par des auteurs tel Plerou et al. [62, 63, 64, 65, 66] et Laloux et al. [67, 68]) pour filtrer le bruit dans les séries financières, particulièrement dans les systèmes de grande dimension tel que les marchés actions. Ces auteurs ont principalement proposé des algorithmes de filtrage basés sur la RMT, que nous présenterons dans la suite.

Dans le travail de Laloux et al. [68], le débruitage des matrices de corrélation en utilisant une analyse de la RMT devrait améliorer l'optimisation de portefeuille. Le filtre décrit par Laloux et al. [68]

fonctionne en considérant que les valeurs propres du marché en dessous de la valeur propre aléatoire maximale constituent le bruit. Ces valeurs propres bruitées sont donc filtrées, en étant remplacées par des valeurs constantes, mais en gardant intacte la trace du système (i.e. la somme des valeurs propres).

Allant dans le même sens que Laloux et al. [67, 68], Plerou et al. [62, 63, 64, 65, 66] ont étudié l'application d'un "filtreur" basé sur la méthode de la RMT. Cette méthode de débruitage est un peu différente de celle proposée par Laloux et al. [67, 68]. Au lieu de remplacer les valeurs propres inférieures à la valeur propre aléatoire maximale par des valeurs constantes, Plerou et al. [62, 63, 64, 65, 66] ont proposé de remplacer les valeurs propres inférieures ou égales à la valeur propre maximale bruitée par "0", en conservant la trace lors de la reconstitution de la diagonale principale à ses valeurs originales après filtrage.

Ces auteurs ont démontré dans leurs travaux que, ces techniques de filtrage basées sur une méthode de RMT sont avantageuses dans l'optimisation de portefeuille, tant au niveau de la réduction du risque réalisé des portefeuilles optimisés qu'au niveau de la prévision de ce risque réalisé. Nous donnerons dans la suite un algorithme qui présente cette méthode, et nous illustrerons la dite méthode par plusieurs exemples.

2.6.1 Algorithme de la reconstruction

Dans cette partie, nous essaierons de proposer un algorithme de calcul pour la reconstruction des matrices de corrélation et de covariance. Nous utiliserons les résultats des auteurs précédemment cités dans l'introduction de cette partie pour arriver à nos fins. Pour cela, posons :

- ➔ Σ une matrice de variance-covariance
- ➔ C la matrice de corrélation correspondante au problème
- ➔ D une matrice diagonale composée des valeurs propres de C , la matrice de corrélation
- ➔ P la matrice des vecteurs propres de la matrice de corrélation C
- ➔ P' la matrice transposée de P
- ➔ A la matrice diagonale, composée des inverses des volatilités des différents actifs.

L'algorithme de reconstruction de la matrice de variance-covariance s'écrit comme suit :

Algorithm 1 Algorithme de reconstruction des matrices C et Σ

Require: D, P, A .

- 1: Estimation de la matrice de covariance Σ
 - 2: Calcul de la matrice de corrélation C par la formule $C = A\Sigma A$.
 - 3: Calcul des valeurs et des vecteurs propres de la matrice de corrélation (et si nécessaire, vérification d'égalité par la formule $C = PDP'$).
 - 4: Modification des valeurs propres : $\{\lambda_i \rightarrow \bar{\lambda}_i, i = 1, 2, \dots, n\}$.
 - 5: Obtention d'une matrice de "corrélation" \bar{C} intermédiaire : On remplace dans la matrice D précédemment obtenue, les $\{\lambda_i, i = 1, 2, \dots, n\}$ par les $\{\bar{\lambda}_i, i = 1, 2, \dots, n\}$
 - 6: Obtention de la nouvelle matrice de corrélation C^{new} : Pour obtenir cette matrice, on remplace la diagonale de la matrice \bar{C} par des "1", et on conserve tous les autres éléments de la matrice.
 - 7: Obtention de la nouvelle matrice de Covariance Σ^{new} par la formule $\Sigma^{new} = A^{-1}C^{new}A^{-1}$.
-

2.6.2 Présentation d'un cas simple : Cas d'une matrice 2×2

Pour un début et une bonne compréhension, nous travaillerons d'abord avec des matrices de taille 2×2 , avant de généraliser la méthode à des matrices de taille $n \times n$. La matrice C étant une matrice de corrélation, est, comme nous l'avons dit dans la partie précédente, diagonalisable. On peut la ré-écrire comme $C = PDP'$. Dans cette partie, on considère que la matrice de covariance est connue, étant donné qu'on l'estime dès le départ. Définissons nos différentes matrices pour les calculs :

$$A = \begin{pmatrix} \frac{1}{\sigma_x} & 0 \\ 0 & \frac{1}{\sigma_y} \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

$$P = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}, \quad P' = \begin{pmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{pmatrix}$$

Le calcul des coefficients de C en fonction des coefficients de P , D et P' donne :

$$\begin{aligned} C &= PDP' \\ &= \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} \\ \lambda_1 v_{21} & \lambda_2 v_{22} \end{pmatrix} \begin{pmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 v_{11}^2 + \lambda_2 v_{12}^2 & \lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22} \\ \lambda_1 v_{11} v_{21} + \lambda_1 v_{12} v_{22} & \lambda_1 v_{21}^2 + \lambda_2 v_{22}^2 \end{pmatrix}. \end{aligned} \quad (2.64)$$

La matrice C étant une matrice de corrélation, elle peut s'écrire sous la forme

$$C = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}. \quad (2.65)$$

En identifiant les coefficients de l'équation (2.65) à ceux de l'équation (2.71), on obtient les relations ci-dessous :

$$\begin{cases} 1 &= \lambda_1 v_{11}^2 + \lambda_2 v_{12}^2 \\ 1 &= \lambda_1 v_{21}^2 + \lambda_2 v_{22}^2 \\ \rho &= \lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22} \end{cases} \quad (2.66)$$

En utilisant la méthode de la RMT (i.e. en modifiant les valeurs propres de la matrice C dans la "nouvelle" matrice de corrélation \bar{C}), on obtient comme matrice de corrélation

$$\bar{C} = \begin{pmatrix} \bar{\lambda}_1 v_{11}^2 + \bar{\lambda}_2 v_{12}^2 & \bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22} \\ \bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22} & \bar{\lambda}_1 v_{21}^2 + \bar{\lambda}_2 v_{22}^2 \end{pmatrix}. \quad (2.67)$$

où les $\{\bar{\lambda}_i, i = 1, 2\}$ sont les valeurs propres modifiées de la matrice C . Cette matrice peut également être écrite sous la forme

$$\bar{C} = \begin{pmatrix} 1 & \bar{\rho} \\ \bar{\rho} & 1 \end{pmatrix}, \quad (2.68)$$

ce qui nous permet de trouver les relations suivantes :

$$\begin{cases} 1 &= \bar{\lambda}_1 v_{11}^2 + \bar{\lambda}_2 v_{12}^2 \\ 1 &= \bar{\lambda}_1 v_{21}^2 + \bar{\lambda}_2 v_{22}^2 \\ \bar{\rho} &= \bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22} \end{cases} \quad (2.69)$$

De là, on obtient notre nouvelle matrice de corrélation

$$C^{new} = \begin{pmatrix} 1 & \bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22} \\ \bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22} & 1 \end{pmatrix}, \quad (2.70)$$

Sachant que la relation liant les matrices A , Σ et C s'écrit $\Sigma = A^{-1}CA^{-1}$, on a :

$$\begin{aligned} \Sigma^{new} &= A^{-1}C^{new}A^{-1} \\ &= \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{pmatrix} \begin{pmatrix} 1 & \bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22} \\ \bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22} & 1 \end{pmatrix} \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{pmatrix} \\ &= \begin{pmatrix} \sigma_x & \sigma_x(\bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22}) \\ \sigma_y(\bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22}) & \sigma_y \end{pmatrix} \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{pmatrix} \\ &= \begin{pmatrix} \sigma_x^2 & \sigma_x \sigma_y(\bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22}) \\ \sigma_x \sigma_y(\bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22}) & \sigma_y^2 \end{pmatrix} \end{aligned} \quad (2.71)$$

Dans l'expression de l'équation (2.71), notons $\sigma_{xy}^{new} = \sigma_x \sigma_y (\bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22})$. Nous aimerions trouver la relation entre σ_{xy} , ρ , ρ^{new} et σ_{xy}^{new} . Notons avant que

$$\rho^{new} = \frac{\sigma_{xy}^{new}}{\sigma_x \sigma_y}. \quad (2.72)$$

De là, on a alors

$$\begin{aligned} \sigma_{xy}^{new} &= \sigma_x \sigma_y \rho^{new} \\ &= \frac{\sigma_x \sigma_y}{\sigma_{xy}} \sigma_{xy} \rho^{new} \end{aligned} \quad (2.73)$$

$$\sigma_{xy}^{new} = \frac{\rho^{new}}{\rho} \sigma_{xy} \quad (2.74)$$

On peut exprimer σ_{xy}^{new} en fonction des valeurs propres modifiées $\{\bar{\lambda}_i, i = 1, 2\}$ et des composantes des vecteurs propres. En effet, de la relation (2.73), on a :

$$\sigma_{xy}^{new} = \sigma_x \sigma_y (\bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22}). \quad (2.75)$$

2.6.3 Calcul des valeurs propres et des vecteurs propres, et Conservation de la trace lors du changement de base

Dans cette partie, nous utiliserons des résultats mathématiques basiques pour montrer qu'en utilisant l'algorithme présenté dans la section précédente, la trace de la matrice de corrélation est bien conservée. Pour cela, nous calculerons tout d'abord les valeurs propres et les vecteurs propres de la dite matrice, nous appliquerons la méthode de la RMT, et nous comparerons les traces (avant débruitage et après débruitage).

2.6.3.1 Calcul des valeurs propres et des vecteurs propres

Dans la RMT, lorsqu'on effectue le changement des valeurs propres (pour passer dans la nouvelle base où on a C^{new}), il faudrait que la trace de la matrice de corrélation soit conservée. Vérifions le sur les matrices C et C^{new} citées ci-dessus, dans les équations (2.65) et (2.68). Le calcul des valeurs propres de C donne

$$\begin{aligned} \lambda_1 &= 1 - \rho \\ \lambda_2 &= 1 + \rho \end{aligned}$$

Et celles de C^{new} sont données par

$$\begin{aligned} \lambda_1^{new} &= 1 - \bar{\rho} \\ \lambda_2^{new} &= 1 + \bar{\rho} \end{aligned}$$

En effet, les valeurs propres de C et de C^{new} sont respectivement les λ et les β , solutions des équations $\det(C - \lambda I) = 0$ et $\det(C - \beta I) = 0$. On a par des calculs élémentaires (On ne fera que le calcul pour la matrice C , étant donné que les calculs sont les mêmes pour la matrice C^{new} , à la différence qu'on mettra des $\bar{\rho}$ à la place des ρ) :

$$\begin{aligned} \det(C - \lambda I) = 0 &\iff \begin{vmatrix} 1 - \lambda & \rho \\ \rho & 1 - \lambda \end{vmatrix} = 0 \\ &\iff (1 - \lambda)^2 - \rho^2 = 0 \\ &\iff (1 - \lambda - \rho)(1 - \lambda + \rho) = 0 \\ &\iff 1 - \lambda - \rho = 0 \quad \vee \quad 1 - \lambda + \rho = 0 \\ &\iff \lambda_1 = 1 - \rho \quad \vee \quad \lambda_2 = 1 + \rho \end{aligned} \quad (2.76)$$

Résultat [50] : Soit une matrice M de taille 2×2 inversible et diagonalisable dans une base orthonormée de la forme

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \quad (2.77)$$

Notons \mathcal{U} l'ouvert de $\mathcal{M}_2(\mathbb{R})$ correspondant aux matrices ayant leurs deux valeurs propres distinctes :

$$\mathcal{U} = \{M \in \mathcal{M}_2(\mathbb{R}) \mid \Delta(P_c(M)) \neq 0\}, \quad (2.78)$$

où $P_c(M)$ est le polynôme caractéristique de la matrice M et Δ est son discriminant : $\Delta(ax^2+bx+c) = b^2 - 4ac$. Cet ouvert a deux composantes connexes \mathcal{U}_+ et \mathcal{U}_- correspondant respectivement à $\Delta > 0$ et $\Delta < 0$. Nous ne considérerons que le cas où $\Delta > 0$, étant donné que toutes les valeurs propres avec lesquelles nous travaillons sont réelles. Notons λ_+ et λ_- les valeurs propres de la matrice définie dans (2.77). Nous ne définirons les vecteurs propres que sur $\mathcal{U}_+ \cap \{a \neq \lambda_-, |bc| < (a - \lambda_-)^2\}$, cadre du travail impose. Les vecteurs propres associés à ces valeurs propres s'écrivent alors :

$$v_1 = \begin{pmatrix} d - \lambda_+ \\ a \end{pmatrix} \quad \text{et} \quad v_2 = \begin{pmatrix} b \\ a - \lambda_- \end{pmatrix} \quad (2.79)$$

Appliquant le résultat de l'équation (2.79) à notre exemple, on trouve que les vecteurs propres associés aux valeurs propres λ_1 (qu'on considère comme λ_-) et λ_2 (qu'on considère comme λ_+), s'écrivent

$$v_1 = \begin{pmatrix} 1 - \lambda_2 \\ \rho \end{pmatrix} = \begin{pmatrix} -\rho \\ \rho \end{pmatrix} \quad \text{et} \quad v_2 = \begin{pmatrix} \rho \\ 1 - \lambda_1 \end{pmatrix} = \begin{pmatrix} \rho \\ \rho \end{pmatrix}. \quad (2.80)$$

On remarque donc que la matrice de passage (matrice des vecteurs propres)

$$v = (v_1 \ v_2) = \begin{pmatrix} -\rho & \rho \\ \rho & \rho \end{pmatrix} \quad (2.81)$$

peut s'écrire sous la forme

$$v = \begin{pmatrix} -\alpha & \alpha \\ \alpha & \alpha \end{pmatrix} \quad (2.82)$$

où

$$\alpha = \rho. \quad (2.83)$$

2.6.3.2 Conservation de la trace lors du changement de base

Le résultat précédent démontré, nous pouvons vérifier que la trace de C est égale à la trace de C^{new} . Sachant que la trace d'une matrice est égale à la somme de ses valeurs propres, il nous suffit de vérifier que $\lambda_1 + \lambda_2 = \lambda_1^{new} + \lambda_2^{new}$. On doit aussi trouver que la trace des ces deux matrices vaut N , où N représente la dimension de la matrice. Donc, dans ce cas, on devrait trouver que la trace des deux matrices vaut exactement 2. On a :

$$\begin{aligned} \lambda_1 + \lambda_2 &= 1 - \rho + 1 + \rho \\ &= 2 \end{aligned} \quad (2.84)$$

et

$$\begin{aligned} \lambda_1^{new} + \lambda_2^{new} &= 1 - \bar{\rho} + 1 + \bar{\rho} \\ &= 2 \end{aligned} \quad (2.85)$$

On remarque au vu de (2.84) et de (2.85) que $\lambda_1 + \lambda_2 = \lambda_1^{new} + \lambda_2^{new} = 2$, la dimension des deux matrices. D'où, on peut conclure que, avec le changement dû à la RMT, la trace des deux matrices de corrélation est la même, donc est conservée.

2.6.4 Interprétation financière

Rappelons qu'à l'équation (2.56), nous avons exprimé les poids w_g d'un portefeuille en fonction du vecteur unitaire e et en fonction d'une matrice de covariance Σ . Nous voulons ici exprimer les composantes du vecteur des poids, en fonction :

- ➡ des valeurs propres $\{\lambda_i, i = 1, 2\}$
- ➡ des volatilités
- ➡ des composantes des vecteurs propres.

Notons que les coefficients de corrélation dépendent des valeurs propres et des composantes des vecteurs propres. Avant de déterminer ces poids, donnons nous une matrice de variance covariance Σ de la forme

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix}. \quad (2.86)$$

Le déterminant de cette matrice étant $\det \Sigma = \sigma_x^2 \sigma_y^2 - \sigma_{xy}^2$, on obtient facilement que l'inverse de la matrice Σ s'écrit (en considérant bien évidemment que ce déterminant est non nul) :

$$\Sigma^{-1} = \begin{pmatrix} \frac{\sigma_y^2}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} & -\frac{\sigma_{xy}}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} \\ -\frac{\sigma_{xy}}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} & \frac{\sigma_x^2}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} \end{pmatrix}. \quad (2.87)$$

De là, on a

$$\Sigma^{-1} e = \begin{pmatrix} \frac{\sigma_y^2}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} & -\frac{\sigma_{xy}}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} \\ -\frac{\sigma_{xy}}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} & \frac{\sigma_x^2}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\sigma_y^2 - \sigma_{xy}}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} \\ \frac{\sigma_x^2 - \sigma_{xy}}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} \end{pmatrix} \quad (2.88)$$

et

$$\begin{aligned} e' \Sigma^{-1} e &= \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sigma_y^2 - \sigma_{xy}}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} \\ \frac{\sigma_x^2 - \sigma_{xy}}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} \end{pmatrix} \\ &= \frac{\sigma_x^2 + \sigma_y^2 - 2\sigma_{xy}}{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^2} \end{aligned} \quad (2.89)$$

Combinant (2.88) et (2.89), on obtient l'expression des différents poids en fonction des volatilités :

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} \frac{\sigma_y^2 - \sigma_{xy}}{\sigma_x^2 + \sigma_y^2 - 2\sigma_{xy}} \\ \frac{\sigma_x^2 - \sigma_{xy}}{\sigma_x^2 + \sigma_y^2 - 2\sigma_{xy}} \end{pmatrix}. \quad (2.90)$$

En remplaçant les différentes covariances par leurs expressions (en fonction des valeurs propres et des composantes des vecteurs propres...), on obtient

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} \frac{\sigma_y^2 - \sigma_x \sigma_y (\lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22})}{\sigma_x^2 + \sigma_y^2 - 2\sigma_x \sigma_y (\lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22})} \\ \frac{\sigma_x^2 - \sigma_x \sigma_y (\lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22})}{\sigma_x^2 + \sigma_y^2 - 2\sigma_x \sigma_y (\lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22})} \end{pmatrix}. \quad (2.91)$$

On aimerait enfin exprimer les poids dans la nouvelle base, en fonction des différents coefficients de corrélation (le coefficient de corrélation ρ dans la base de départ et le coefficient de corrélation dans la nouvelle base ρ^{new}). Comme nous l'avons précisé un peu plus haut dans (2.66), les coefficients de corrélation dépendent des valeurs propres, des composantes et des vecteurs propres. En effet,

$$\begin{aligned} \sigma_{xy} &= \sigma_x \sigma_y (\lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22}) \\ &:= \sigma_x \sigma_y \rho \end{aligned} \quad (2.92)$$

ce qui nous permet de déduire facilement le résultat obtenu dans (2.66). Dans la nouvelle base, on a formellement

$$\begin{pmatrix} w_1^{new} \\ w_2^{new} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_y^2 - \sigma_{xy}^{new}}{\sigma_x^2 + \sigma_y^2 - 2\sigma_{xy}^{new}} \\ \frac{\sigma_x^2 - \sigma_{xy}^{new}}{\sigma_x^2 + \sigma_y^2 - 2\sigma_{xy}^{new}} \end{pmatrix}. \quad (2.93)$$

où σ_{xy}^{new} est défini dans (2.74). Utilisant l'expression de σ_{xy}^{new} donnée dans (2.74) dans le vecteur de l'équation (2.93), on obtient finalement,

$$\begin{pmatrix} w_1^{new} \\ w_2^{new} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_y^2 - \frac{\rho^{new}}{\rho} \sigma_{xy}}{\sigma_x^2 + \sigma_y^2 - 2\frac{\rho^{new}}{\rho} \sigma_{xy}} \\ \frac{\sigma_x^2 - \frac{\rho^{new}}{\rho} \sigma_{xy}}{\sigma_x^2 + \sigma_y^2 - 2\frac{\rho^{new}}{\rho} \sigma_{xy}} \end{pmatrix}. \quad (2.94)$$

Ce résultat obtenu, une question vitale se pose. Quel a été l'impact de cette modification sur les facteurs importants (coefficients de corrélation et covariances) de nos matrices ? C'est ce que nous essaierons d'étudier dans la partie qui suit, en considérant principalement deux cas de modification des valeurs propres : le cas où on remplacera les deux valeurs propres par leur moyenne (ce qui n'est pas ce qu'on fera vraiment dans l'application, étant donné qu'on n'a pas de référence ici), et le cas où on remplace l'une des deux valeurs propres par 0.

2.6.5 Choix de modification des valeurs propres

Après changement de base, on s'est rendu compte que ce qu'on modifiait n'était rien d'autre que les corrélations (et donc les covariances). En effet, dans la base de départ, le coefficient de corrélation s'écrit $\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$ et dans la nouvelle base, le coefficient de corrélation s'écrit $\rho^{new} = \frac{\sigma_{xy}^{new}}{\sigma_x \sigma_y}$. Donc, il serait intéressant, pour chaque cas de modification des valeurs propres, d'étudier l'écart entre ces deux corrélations, et essayer de lui donner une interprétation. La différence entre les deux corrélations s'écrit

$$\begin{aligned} \rho - \rho^{new} &= \frac{\sigma_{xy}}{\sigma_x \sigma_y} - \frac{\sigma_{xy}^{new}}{\sigma_x \sigma_y} \\ &= \frac{\sigma_{xy} - \sigma_{xy}^{new}}{\sigma_x \sigma_y} \end{aligned} \quad (2.95)$$

$$\begin{aligned} &= \frac{\sigma_x \sigma_y (\lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22}) - \sigma_x \sigma_y (\bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22})}{\sigma_x \sigma_y} \\ &= (\lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22}) - (\bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22}). \end{aligned} \quad (2.96)$$

Les volatilités σ_x et σ_y étant positives, nous devons juste étudier la différence $\sigma_{xy} - \sigma_{xy}^{new}$ dans l'équation (2.95).

Mais avant, nous devons justifier que ρ^{new} est un coefficient de corrélation. Étant donné que la matrice de passage P de l'ensemble de départ au nouvel ensemble est définie tel que $P^{-1} = P'$ et donc $P'P = PP' = Id$, on a que la matrice P est une matrice orthogonale. Donc le changement conserve les coefficients de corrélation (ils ne sont pas forcément égaux à ceux de la base de départ...). Ainsi, on peut conclure que ρ^{new} est un coefficient de corrélation.

2.6.5.1 Cas où on remplace les deux valeurs propres par leur moyenne

Comme annoncé ci-dessus, nous devons juste étudier la différence entre les corrélations σ_{xy} et σ_{xy}^{new} . Étant données les valeurs propres λ_1 et λ_2 , on les modifiera en $\bar{\lambda}_1$ et $\bar{\lambda}_2$ par la relation $\bar{\lambda}_1 = \bar{\lambda}_2 = \frac{1}{2}(\lambda_1 + \lambda_2)$. On aura donc de (2.96) :

$$\begin{aligned} \sigma_{xy} - \sigma_{xy}^{new} &= \lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22} - (\bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22}) \\ &= \lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22} - \frac{1}{2}(\lambda_1 + \lambda_2) v_{11} v_{21} - \frac{1}{2}(\lambda_1 + \lambda_2) v_{12} v_{22} \\ &= \frac{1}{2} \lambda_1 (v_{11} v_{21} - v_{12} v_{22}) + \frac{1}{2} \lambda_2 (v_{12} v_{22} - v_{11} v_{21}). \end{aligned} \quad (2.97)$$

Des vecteurs propres obtenus dans l'équation (2.80), si on pose

$$v = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} = \begin{pmatrix} -\alpha & \alpha \\ \alpha & \alpha \end{pmatrix}, \quad (2.98)$$

on aura les relations suivantes :

$$\begin{cases} v_{11} &= -v_{22} = -v_{12} = -v_{21} \\ v_{12} &= v_{21} = v_{22} \end{cases} \quad (2.99)$$

En utilisant ces relations dans l'équation (2.97), on obtient

$$\begin{aligned} \sigma_{xy} - \sigma_{xy}^{new} &= \frac{1}{2} \lambda_1 (-v_{11} v_{11} - (v_{11} v_{11})) + \frac{1}{2} \lambda_2 (v_{22} v_{22} - (-v_{22} v_{22})) \\ &= -\lambda_1 v_{11}^2 + \lambda_2 v_{22}^2 \\ &= \lambda_1 v_{22}^2 + \lambda_2 v_{22}^2 \\ &= (\lambda_1 + \lambda_2) v_{22}^2 \end{aligned} \quad (2.100)$$

Étant donné que v_{22}^2 est toujours positif et que $\lambda_1 + \lambda_2 > 0$, on aura que la différence $\sigma_{xy} - \sigma_{xy}^{new}$ est toujours positive, d'où $\sigma_{xy} > \sigma_{xy}^{new}$. On peut conclure qu'en appliquant cette modification sur les valeurs propres, on diminue le coefficient de corrélation entre les deux variables (qui en finance représentent les titres) correspondantes, disons X et Y .

2.6.5.2 Cas où on remplace l'une des deux valeurs propres par 0

Dans cette partie, on aimerait étudier la différence entre les corrélations σ_{xy} et σ_{xy}^{new} , lorsqu'on met l'une des deux valeurs propres à 0. On prendra $\overline{\lambda_1} = 0$ et $\overline{\lambda_2} = \lambda_2$. La différence des corrélations s'écrit alors :

$$\begin{aligned}\sigma_{xy} - \sigma_{xy}^{new} &= \lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22} - (\overline{\lambda_1} v_{11} v_{21} + \overline{\lambda_2} v_{12} v_{22}) \\ &= \lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22} - \lambda_2 v_{12} v_{22} \\ &= \lambda_1 v_{11} v_{21} \\ &= -\lambda_1 v_{11}^2 \quad \text{étant donné que } v_{11} = -v_{21}.\end{aligned}\tag{2.101}$$

On a : $\lambda_1 > 0$, v_{11}^2 est toujours positif, ce qui implique que la différence entre les deux écarts-types est négative, d'où on augmente la corrélation en effectuant ce changement de valeurs propres.

2.6.5.3 Conclusion

On remarque dans cette partie que, lorsqu'on remplace l'une des valeurs propres par leur moyenne, on diminue le coefficient de corrélation. Par contre, lorsqu'on remplace l'une des valeurs propres par 0, on augmente le coefficient de corrélation. Il serait peut être aussi intéressant, d'étudier cette méthode en effectuant des opérations autres que la moyenne des valeurs propres, et la mise à zéro de l'une des valeurs propres. Cette étude ne sera pas effectuée dans ce rapport. On peut vérifier les résultats ci-dessus à l'aide d'une application numérique. On appliquera exactement la procédure décrite ci-dessus avec des données que nous prendrons au hasard.

2.6.6 Application

Nous avons effectué une application numérique sous MATLAB dans le cas d'une matrice 2×2 , pour déterminer les poids d'un portefeuille. Pour cela, avec comme données :

$$\begin{cases} \sigma_x &= 0,15 \\ \sigma_y &= 0,25 \\ \rho &= 0,75 \end{cases}\tag{2.102}$$

nous avons obtenu :

➡ La matrice A

$$A = \begin{pmatrix} \frac{1}{\sigma_x} & 0 \\ 0 & \frac{1}{\sigma_y} \end{pmatrix} = \begin{pmatrix} 6.6667 & 0 \\ 0 & 4 \end{pmatrix}\tag{2.103}$$

➡ La matrice de variance-covariance

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} = \begin{pmatrix} \sigma_x^2 & \rho \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{pmatrix} = \begin{pmatrix} 0.0225 & 0.0281 \\ 0.0281 & 0.0625 \end{pmatrix}\tag{2.104}$$

➡ Matrice de covariance

$$\begin{aligned}C &= \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0.75 \\ 0.75 & 1 \end{pmatrix} \\ &= A \Sigma A = \begin{pmatrix} 1 & 0.75 \\ 0.75 & 1 \end{pmatrix}\end{aligned}\tag{2.105}$$

➡ en appliquant la formule (2.56), on obtient les poids

$$w = \begin{pmatrix} 1.1957 \\ -0.1957 \end{pmatrix}\tag{2.106}$$

où $e = (1 \ 1)'$ dans la formule de l'équation (2.56). On remarque bien que la somme des w_i vaut bien 1.

Nous voulons maintenant appliquer la méthode proposée dans la RMT. Pour cela, nous devons tout d'abord calculer les valeurs propres (lesquelles seront modifiées) de la matrice de corrélation Σ et les vecteurs propres associés.

➡ Calcul des valeurs propres

$$\begin{cases} \lambda_1 = 0.25 \\ \lambda_2 = 1.75 \end{cases} \quad (2.107)$$

➡ Les vecteurs propres associés sont

$$v_1 = \begin{pmatrix} -0.7071 \\ 0.7071 \end{pmatrix} \quad \text{et} \quad v_2 = \begin{pmatrix} 0.7071 \\ 0.7071 \end{pmatrix} \quad (2.108)$$

Notons v la matrice dont les colonnes sont les vecteurs propres v_1 et v_2 :

$$v = \begin{pmatrix} -0.7071 & 0.7071 \\ 0.7071 & 0.7071 \end{pmatrix}. \quad (2.109)$$

➡ Modification des valeurs propres : On remplacera premièrement dans cet exemple, la première valeur propre λ_1 par 0, et on conservera la deuxième valeur propre λ_2 . Ensuite, on remplacera les deux valeurs propres par leur moyenne.

➡ Dans le premier cas, on aura :

☞ les nouvelles valeurs propres sont :

$$\begin{cases} \lambda_1^{new} = 0 \\ \lambda_2^{new} = 1.75 \end{cases} \quad (2.110)$$

notons Λ^{new} la matrice des nouvelles valeurs propres :

$$\Lambda^{new} = \begin{pmatrix} \lambda_1^{new} & 0 \\ 0 & \lambda_2^{new} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1.75 \end{pmatrix} \quad (2.111)$$

☞ Matrice de pseudo-corrélation dans la base intermédiaire :

$$\begin{aligned} \bar{C} &= v * \Lambda^{new} * v' \\ &= \begin{pmatrix} 0.8750 & 0.8750 \\ 0.8750 & 0.8750 \end{pmatrix} \end{aligned} \quad (2.112)$$

☞ Nouvelle matrice de corrélation : Cette matrice représente la matrice \bar{C} de l'équation (2.112) dans laquelle on a remplacé les éléments de la diagonale par des 1

$$C^{new} = \begin{pmatrix} 1 & 0.8750 \\ 0.8750 & 1 \end{pmatrix} \quad (2.113)$$

☞ Nouvelle matrice de variance-covariance

$$\Sigma^{new} = A^{-1} C^{new} A^{-1} \quad (2.114)$$

$$= \begin{pmatrix} 0.0225 & 0.0328 \\ 0.0328 & 0.0625 \end{pmatrix} \quad (2.115)$$

☞ en appliquant la formule (2.56) avec la matrice Σ^{new} ci-dessus, on obtient les poids

$$w^{new} = \begin{pmatrix} 1.5323 \\ -0.5323 \end{pmatrix}. \quad (2.116)$$

☞ On peut également regarder l'écart entre les corrélations ρ et ρ^{new} . Le ρ de départ vaut 0.75 et ρ^{new} vaut 0.8750. On remarque que **cette opération sur les valeurs propres a augmenté le coefficient de corrélation** tout en gardant intacte la trace de la matrice de corrélation.

Nous allons maintenant récapituler tous les résultats obtenus ci-dessus dans le tableau ci-dessous, pour voir quelles sont les différences qui ont été effectuées après la transformation par la méthode de la RMT.

Dans la base de départ					Après changement de base				
σ_{xy}	ρ	val. p.	w	$\text{tr}(C)$	σ_{xy}^{new}	ρ^{new}	val. p.	w^{new}	$\text{tr}(C^{new})$
0.0281	0.75	$\begin{pmatrix} 0.25 \\ 1.75 \end{pmatrix}$	$\begin{pmatrix} 1.1957 \\ -0.1957 \end{pmatrix}$	2	0.0328	0.8750	$\begin{pmatrix} 0 \\ 1.75 \end{pmatrix}$	$\begin{pmatrix} 1.5323 \\ -0.5323 \end{pmatrix}$	2

TABLE 2.1 – Tableau récapitulatif des données avant (σ_{xy}, ρ) et après $(\sigma_{xy}^{new}, \rho^{new})$ le changement de base, et des résultats avant (valeurs propres, poids w et trace de la matrice de corrélation C) et après (valeurs propres, poids w^{new} et trace de la nouvelle matrice de corrélation C^{new}) le changement de base dans le cas où on remplace l’une des valeurs propres par 0.

➡ Nous allons maintenant remplacer les deux valeurs propres par leur moyenne pour voir quel sera le comportement des coefficients de corrélation. Comme nous l’avons dit un peu plus haut, ce n’est pas vraiment ce qu’on fera dans les applications avec les données réelles. En effet, on n’a pas assez de données pour appliquer la vraie procédure (remplacer les valeurs propres inférieures à λ_{max} (borne supérieure de la densité de M-P) par leur moyenne). On aura alors dans le deuxième cas :

☞ les nouvelles valeurs propres sont :

$$\begin{cases} \lambda_1^{new} = 1 \\ \lambda_2^{new} = 1 \end{cases} \quad (2.117)$$

notons Λ^{new} la matrice des nouvelles valeurs propres :

$$\Lambda^{new} = \begin{pmatrix} \lambda_1^{new} & 0 \\ 0 & \lambda_2^{new} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (2.118)$$

☞ Matrice de pseudo-corrélation dans la base intermédiaire :

$$\begin{aligned} \overline{C} &= v * \Lambda^{new} * v' \\ &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \end{aligned} \quad (2.119)$$

☞ Nouvelle matrice de corrélation : Cette matrice représente la matrice \overline{C} de l’équation (2.119) dans laquelle on a remplacé les éléments de la diagonale par des 1

$$C^{new} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (2.120)$$

☞ Nouvelle matrice de variance-covariance

$$\Sigma^{new} = A^{-1} C^{new} A^{-1} \quad (2.121)$$

$$= \begin{pmatrix} 0.0225 & 0 \\ 0 & 0.0625 \end{pmatrix} \quad (2.122)$$

☞ en appliquant la formule (2.56) avec la matrice Σ^{new} ci-dessus, on obtient les poids

$$w^{new} = \begin{pmatrix} 0.7353 \\ 0.2647 \end{pmatrix}. \quad (2.123)$$

☞ On peut également regarder l’écart entre les corrélations ρ et ρ^{new} . Le ρ de départ vaut 0.75 et ρ^{new} vaut 0. On remarque que **cette opération sur les valeurs propres a annulé le coefficient de corrélation** tout en gardant intacte la trace de la matrice de corrélation.

Nous allons maintenant récapituler tous les résultats obtenus ci-dessus dans le tableau ci-dessous, pour voir quelles sont les différences qui ont été effectuées après la transformation par la méthode de la RMT.

Dans la base de départ					Après changement de base				
σ_{xy}	ρ	val. p.	w	$\text{tr}(C)$	σ_{xy}^{new}	ρ^{new}	val. p.	w^{new}	$\text{tr}(C^{new})$
0.0281	0.75	$\begin{pmatrix} 0.25 \\ 1.75 \end{pmatrix}$	$\begin{pmatrix} 1.1957 \\ -0.1957 \end{pmatrix}$	2	0	0	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0.7353 \\ 0.2647 \end{pmatrix}$	2

TABLE 2.2 – Tableau récapitulatif des données avant (σ_{xy}, ρ) et après $(\sigma_{xy}^{new}, \rho^{new})$ le changement de base, et des résultats avant (valeurs propres, poids w et trace de la matrice de corrélation C) et après (valeurs propres, poids w^{new} et trace de la nouvelle matrice de corrélation C^{new}) le changement de base dans le cas où on remplace les deux valeurs propres par leur moyenne.

2.6.7 Tentative de généralisation de la méthode

Les exemples proposés dans le cas des matrices de taille 2×2 n'étant pas très parlant, nous allons essayer de généraliser la méthode dans le cas des matrices de taille $n \times n$. Mais avant, nous essaierons de voir si dans le cas des matrices 3×3 , il est possible de tirer des conclusions sur les résultats qu'on obtiendra. Dans la partie 2.6.7.1, nous ferons une étude théorique de la méthode dans le cas des matrices 3×3 , et nous traiterons un exemple numérique et analyserons les résultats obtenus dans la partie 2.6.7.2. Dans la partie 2.6.7.3, nous ferons l'étude théorique dans le cas des matrices de taille $n \times n$ et enfin, dans la partie 2.6.7.4, nous illustrerons les résultats dans le cas des données réelles : les rendements.

2.6.7.1 Cas d'une matrice de taille 3×3

Notons Σ la matrice de variance-covariance. Dans le cas des matrices de taille 3×3 , cette matrice s'écrit sous la forme :

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_y^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_z^2 \end{pmatrix} \quad (2.124)$$

la matrice A dans ce cas s'écrit alors :

$$A = \begin{pmatrix} \frac{1}{\sigma_x} & 0 & 0 \\ 0 & \frac{1}{\sigma_y} & 0 \\ 0 & 0 & \frac{1}{\sigma_z} \end{pmatrix}, \quad (2.125)$$

où σ_x , σ_y et σ_z sont les racines carrées des éléments diagonaux de la matrice de variance covariance Σ , définis à l'équation (2.124). La matrice de corrélation s'obtient par la formule $C = A\Sigma A$. Notons λ_1 , λ_2 et λ_3 les valeurs propres de la matrice C . La matrice diagonale D des valeurs propres s'écrit :

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \quad (2.126)$$

Notons v_1 le vecteur propre associé à la valeur propre λ_1 , v_2 le vecteur propre associé à la valeur propre λ_2 et v_3 le vecteur propre associé à la valeur propre λ_3 . Posons

$$v_1 = \begin{pmatrix} v_{11} \\ v_{21} \\ v_{31} \end{pmatrix}, \quad v_2 = \begin{pmatrix} v_{12} \\ v_{22} \\ v_{32} \end{pmatrix} \quad \text{et} \quad v_3 = \begin{pmatrix} v_{13} \\ v_{23} \\ v_{33} \end{pmatrix} \quad (2.127)$$

La matrice de passage P s'écrit alors (P' représente la matrice transposée de P)

$$P = \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix}, \quad \text{et} \quad P' = \begin{pmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{pmatrix} \quad (2.128)$$

On aimerait trouver les coefficients de C en fonction des coefficients de P , D et P' . La formule $C = PDP'$ nous donne :

$$\begin{aligned}
C &= PDP' \\
&= \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} & \lambda_3 v_{13} \\ \lambda_1 v_{21} & \lambda_2 v_{22} & \lambda_3 v_{23} \\ \lambda_1 v_{31} & \lambda_2 v_{32} & \lambda_3 v_{33} \end{pmatrix} \begin{pmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{pmatrix} \\
&= \begin{pmatrix} u & \rho_{12} & \rho_{13} \\ \rho_{12} & v & \rho_{23} \\ \rho_{13} & \rho_{23} & w \end{pmatrix}, \tag{2.129}
\end{aligned}$$

où

$$\begin{cases} u = \lambda_1 v_{11}^2 + \lambda_2 v_{12}^2 + \lambda_3 v_{13}^2 \\ v = \lambda_1 v_{21}^2 + \lambda_2 v_{22}^2 + \lambda_3 v_{23}^2 \\ w = \lambda_1 v_{31}^2 + \lambda_2 v_{32}^2 + \lambda_3 v_{33}^2 \\ \rho_{12} = \lambda_1 v_{11} v_{21} + \lambda_2 v_{12} v_{22} + \lambda_3 v_{13} v_{23} \\ \rho_{13} = \lambda_1 v_{11} v_{31} + \lambda_2 v_{12} v_{32} + \lambda_3 v_{13} v_{33} \\ \rho_{23} = \lambda_1 v_{21} v_{31} + \lambda_2 v_{22} v_{32} + \lambda_3 v_{23} v_{33} \end{cases} \tag{2.130}$$

Or la matrice C est une matrice de corrélation, donc peut s'écrire sous la forme

$$C = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}, \tag{2.131}$$

ce qui nous permet d'affirmer que

$$\begin{cases} u = \lambda_1 v_{11}^2 + \lambda_2 v_{12}^2 + \lambda_3 v_{13}^2 = 1 \\ v = \lambda_1 v_{21}^2 + \lambda_2 v_{22}^2 + \lambda_3 v_{23}^2 = 1 \\ w = \lambda_1 v_{31}^2 + \lambda_2 v_{32}^2 + \lambda_3 v_{33}^2 = 1 \end{cases} \tag{2.132}$$

En utilisant la méthode de la RMT (i.e. en modifiant les valeurs propres de la matrice C dans la nouvelle matrice de "corrélation" \bar{C} , on obtient :

$$\bar{C} = \begin{pmatrix} x & \bar{\rho}_{12} & \bar{\rho}_{13} \\ \bar{\rho}_{12} & y & \bar{\rho}_{23} \\ \bar{\rho}_{13} & \bar{\rho}_{23} & z \end{pmatrix}. \tag{2.133}$$

où les expressions des $\{\bar{\rho}_{ij}, i, j = 1, 2, 3\}$, de x , y et z sont données par

$$\begin{cases} x = \bar{\lambda}_1 v_{11}^2 + \bar{\lambda}_2 v_{12}^2 + \bar{\lambda}_3 v_{13}^2 \\ y = \bar{\lambda}_1 v_{21}^2 + \bar{\lambda}_2 v_{22}^2 + \bar{\lambda}_3 v_{23}^2 \\ z = \bar{\lambda}_1 v_{31}^2 + \bar{\lambda}_2 v_{32}^2 + \bar{\lambda}_3 v_{33}^2 \\ \bar{\rho}_{12} = \bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22} + \bar{\lambda}_3 v_{13} v_{23} \\ \bar{\rho}_{13} = \bar{\lambda}_1 v_{11} v_{31} + \bar{\lambda}_2 v_{12} v_{32} + \bar{\lambda}_3 v_{13} v_{33} \\ \bar{\rho}_{23} = \bar{\lambda}_1 v_{21} v_{31} + \bar{\lambda}_2 v_{22} v_{32} + \bar{\lambda}_3 v_{23} v_{33} \end{cases} \tag{2.134}$$

et où les $\{\bar{\lambda}_i, i = 1, 2, 3\}$ sont les valeurs propres modifiées de la matrice C .

Pour calculer la matrice de corrélation dans la nouvelle base, nous devons modifier la matrice \bar{C} tel qu'expliqué dans la RMT, i.e. en remplaçant les éléments de sa diagonale par des 1. On a alors :

$$C^{new} = \begin{pmatrix} 1 & \rho_{12}^{new} & \rho_{13}^{new} \\ \rho_{12}^{new} & 1 & \rho_{23}^{new} \\ \rho_{13}^{new} & \rho_{23}^{new} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \bar{\rho}_{12} & \bar{\rho}_{13} \\ \bar{\rho}_{12} & 1 & \bar{\rho}_{23} \\ \bar{\rho}_{13} & \bar{\rho}_{23} & 1 \end{pmatrix}, \tag{2.135}$$

où $\rho_{ij}^{new} = \bar{\rho}_{ij}$, pour $i, j = 1, 2, 3$. Ce qui signifie que par la transformation $\Sigma = A^{-1}CA^{-1}$, on obtiendra

$$\begin{aligned}
\Sigma^{new} &= A^{-1} C^{new} A^{-1} \\
&= \begin{pmatrix} \sigma_x & 0 & 0 \\ 0 & \sigma_y & 0 \\ 0 & 0 & \sigma_z \end{pmatrix} \begin{pmatrix} 1 & \rho_{12}^{new} & \rho_{13}^{new} \\ \rho_{12}^{new} & 1 & \rho_{23}^{new} \\ \rho_{13}^{new} & \rho_{23}^{new} & 1 \end{pmatrix} \begin{pmatrix} \sigma_x & 0 & 0 \\ 0 & \sigma_y & 0 \\ 0 & 0 & \sigma_z \end{pmatrix} \\
&= \begin{pmatrix} \sigma_x^2 & \sigma_x \sigma_y \rho_{12}^{new} & \sigma_x \sigma_z \rho_{13}^{new} \\ \sigma_x \sigma_y \rho_{12}^{new} & \sigma_y^2 & \sigma_y \sigma_z \rho_{23}^{new} \\ \sigma_x \sigma_z \rho_{13}^{new} & \sigma_y \sigma_z \rho_{23}^{new} & \sigma_z^2 \end{pmatrix}.
\end{aligned} \tag{2.136}$$

Notons également la matrice Σ^{new} comme

$$\Sigma^{new} = \begin{pmatrix} \sigma_x^2 & \sigma_{12}^{new} & \sigma_{13}^{new} \\ \sigma_{12}^{new} & \sigma_y^2 & \sigma_{23}^{new} \\ \sigma_{13}^{new} & \sigma_{23}^{new} & \sigma_z^2 \end{pmatrix}. \tag{2.137}$$

où

$$\begin{cases} \sigma_{12}^{new} = \sigma_x \sigma_y \rho_{12}^{new} \\ \sigma_{13}^{new} = \sigma_x \sigma_z \rho_{13}^{new} \\ \sigma_{23}^{new} = \sigma_y \sigma_z \rho_{23}^{new} \end{cases} \tag{2.138}$$

Rappelons que la matrice Σ de départ s'écrit dans l'équation (2.124) comme

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_y^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_z^2 \end{pmatrix}, \tag{2.139}$$

notre but actuel est de trouver les relations qui existent entre σ_x , σ_y , σ_z , σ_{ij} , σ_{ij}^{new} , ρ_{ij} et ρ_{ij}^{new} , pour $i, j = 1, 2, 3$.

Des expressions des équations (2.136) et (2.139), on a,

$$\sigma_{12}^{new} = \sigma_x \sigma_y \rho_{12}^{new} \tag{2.140}$$

$$\begin{aligned}
&= \frac{\sigma_x \sigma_y}{\sigma_{12}} \sigma_{12} \rho_{12}^{new} \\
&= \frac{1}{\rho_{12}} \sigma_{xy} \rho_{12}^{new} \\
\sigma_{12}^{new} &= \frac{\rho_{12}^{new}}{\rho_{12}} \sigma_{xy}
\end{aligned} \tag{2.141}$$

On peut exprimer σ_{xy}^{new} en fonction des volatilités, des valeurs propres modifiées $\{\bar{\lambda}_i, i = 1, 2, 3\}$ et des composantes des vecteurs propres. En effet, de la relation (2.140), on a :

$$\sigma_{12}^{new} = \sigma_x \sigma_y (\bar{\lambda}_1 v_{11} v_{21} + \bar{\lambda}_2 v_{12} v_{22} + \bar{\lambda}_3 v_{13} v_{23}). \tag{2.142}$$

De même, on a :

$$\sigma_{13}^{new} = \sigma_x \sigma_z \rho_{13}^{new} \tag{2.143}$$

$$\begin{aligned}
&= \frac{\sigma_x \sigma_z}{\sigma_{13}} \sigma_{13} \rho_{13}^{new} \\
&= \frac{1}{\rho_{13}} \sigma_{13} \rho_{13}^{new} \\
\sigma_{13}^{new} &= \frac{\rho_{13}^{new}}{\rho_{13}} \sigma_{13}
\end{aligned} \tag{2.144}$$

On peut aussi exprimer σ_{13}^{new} en fonction des volatilités, des valeurs propres modifiées $\{\bar{\lambda}_i, i = 1, 2, 3\}$ et des composantes des vecteurs propres. En effet, de la relation (2.143), on a :

$$\sigma_{13}^{new} = \sigma_x \sigma_z (\bar{\lambda}_1 v_{11} v_{31} + \bar{\lambda}_2 v_{12} v_{32} + \bar{\lambda}_3 v_{13} v_{33}). \tag{2.145}$$

Un raisonnement similaire aux deux autres nous permet de trouver aussi que

$$\sigma_{23}^{new} = \frac{\rho_{23}^{new}}{\rho_{23}} \sigma_{23}. \tag{2.146}$$

Étant donné que l'expression ci-dessus peut aussi s'écrire $\sigma_{23}^{new} = \sigma_y \sigma_z \rho_{23}^{new}$, on peut également exprimer σ_{23}^{new} en fonction des volatilités, des valeurs propres modifiées et des composantes des vecteurs propres. On obtiendra alors

$$\sigma_{23}^{new} = \sigma_y \sigma_z (\bar{\lambda}_1 v_{21} v_{31} + \bar{\lambda}_2 v_{22} v_{32} + \bar{\lambda}_3 v_{23} v_{33}). \quad (2.147)$$

Dans le cas des matrices de taille 3×3 , on remarque qu'il est difficile de conclure sur le choix de modification des valeurs propres, en raison de la complexité de l'étude du signe de l'expression $\sigma_{ij} - \sigma_{ij}^{new}$. Illustrons cette difficulté par l'exemple suivant :

$$\begin{aligned} \sigma_{23} - \sigma_{23}^{new} &= \sigma_y \sigma_z [\lambda_1 v_{21} v_{31} + \lambda_2 v_{22} v_{32} + \lambda_3 v_{23} v_{33} - (\bar{\lambda}_1 v_{21} v_{31} + \bar{\lambda}_2 v_{22} v_{32} + \bar{\lambda}_3 v_{23} v_{33})] \\ &= \sigma_y \sigma_z [(\lambda_1 - \bar{\lambda}_1) v_{21} v_{31} + (\lambda_2 - \bar{\lambda}_2) v_{22} v_{32} + (\lambda_3 - \bar{\lambda}_3) v_{23} v_{33}] \end{aligned} \quad (2.148)$$

On remarque que pour étudier le signe de l'expression 2.148, il faudrait avoir connaître :

- les signes des éléments qui composent les vecteurs propres, et
- les signes des différences $\lambda_i - \bar{\lambda}_i$.

Or on a aucun élément de réponse sur ces deux problèmes, d'où la difficulté.

On aimerait également exprimer les poids w^{new} dans la nouvelle base, pour une matrice 3×3 . Pour cela, nous allons, comme dans le cas d'une matrice 2×2 , effectuer un exercice un peu pénible : calculer l'inverse de la matrice $(\Sigma^{new})^{-1}$. Posons :

$$a = \frac{\sigma_{12}^{new}}{\sigma_x \sigma_y}, \quad b = \frac{\sigma_{13}^{new}}{\sigma_x \sigma_z}, \quad \text{et} \quad c = \frac{\sigma_{23}^{new}}{\sigma_y \sigma_z} \quad (2.149)$$

On a par des calculs simples :

$$D := \det(\Sigma^{new}) = \sigma_x^2 \sigma_y^2 \sigma_z^2 [1 - (a^2 + b^2 + c^2) + 2abc] \quad (2.150)$$

et ainsi, l'inverse $(\Sigma^{new})^{-1}$ de la matrice Σ^{new} s'écrit :

$$\begin{aligned} (\Sigma^{new})^{-1} &= \frac{1}{D} \begin{pmatrix} \sigma_y^2 \sigma_z^2 - \sigma_y^2 \sigma_z^2 c^2 & \sigma_x \sigma_y \sigma_z^2 bc - \sigma_x \sigma_y \sigma_z^2 a & \sigma_x \sigma_y^2 \sigma_z ac - \sigma_x \sigma_y^2 \sigma_z b \\ \sigma_x \sigma_y \sigma_z^2 bc - \sigma_x \sigma_y \sigma_z^2 a & \sigma_x^2 \sigma_z^2 - \sigma_x^2 \sigma_z^2 b^2 & \sigma_x^2 \sigma_y \sigma_z ab - \sigma_x^2 \sigma_y \sigma_z c \\ \sigma_x \sigma_y^2 \sigma_z ac - \sigma_x \sigma_y^2 \sigma_z b & \sigma_x^2 \sigma_y \sigma_z ab - \sigma_x^2 \sigma_y \sigma_z c & \sigma_x^2 \sigma_y^2 - \sigma_x^2 \sigma_y^2 a^2 \end{pmatrix}' \\ &= \frac{1}{D} \begin{pmatrix} \sigma_y^2 \sigma_z^2 (1 - c^2) & \sigma_x \sigma_y \sigma_z^2 (bc - a) & \sigma_x \sigma_y^2 \sigma_z (ac - b) \\ \sigma_x \sigma_y \sigma_z^2 (bc - a) & \sigma_x^2 \sigma_z^2 (1 - b^2) & \sigma_x^2 \sigma_y \sigma_z (ab - c) \\ \sigma_x \sigma_y^2 \sigma_z (ac - b) & \sigma_x^2 \sigma_y \sigma_z (ab - c) & \sigma_x^2 \sigma_y^2 (1 - a^2) \end{pmatrix} \end{aligned} \quad (2.151)$$

Ce résultat obtenu, on peut calculer le produit $(\Sigma^{new})^{-1}e$, où e est le vecteur unitaire de dimension 3×1 . Les calculs donnent :

$$\begin{aligned} (\Sigma^{new})^{-1}e &= \frac{1}{D} \begin{pmatrix} \sigma_y^2 \sigma_z^2 (1 - c^2) & \sigma_x \sigma_y \sigma_z^2 (bc - a) & \sigma_x \sigma_y^2 \sigma_z (ac - b) \\ \sigma_x \sigma_y \sigma_z^2 (bc - a) & \sigma_x^2 \sigma_z^2 (1 - b^2) & \sigma_x^2 \sigma_y \sigma_z (ab - c) \\ \sigma_x \sigma_y^2 \sigma_z (ac - b) & \sigma_x^2 \sigma_y \sigma_z (ab - c) & \sigma_x^2 \sigma_y^2 (1 - a^2) \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{D} \begin{pmatrix} \sigma_y^2 \sigma_z^2 (1 - c^2) + \sigma_x \sigma_y \sigma_z^2 (bc - a) + \sigma_x \sigma_y^2 \sigma_z (ac - b) \\ \sigma_x \sigma_y \sigma_z^2 (bc - a) + \sigma_x^2 \sigma_z^2 (1 - b^2) + \sigma_x^2 \sigma_y \sigma_z (ab - c) \\ \sigma_x \sigma_y^2 \sigma_z (ac - b) + \sigma_x^2 \sigma_y \sigma_z (ab - c) + \sigma_x^2 \sigma_y^2 (1 - a^2) \end{pmatrix}. \end{aligned} \quad (2.152)$$

De là, on a le produit $e'(\Sigma^{new})^{-1}e$:

$$\begin{aligned} e'(\Sigma^{new})^{-1}e &= \frac{1}{D} \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \sigma_y^2 \sigma_z^2 (1 - c^2) + \sigma_x \sigma_y \sigma_z^2 (bc - a) + \sigma_x \sigma_y^2 \sigma_z (ac - b) \\ \sigma_x \sigma_y \sigma_z^2 (bc - a) + \sigma_x^2 \sigma_z^2 (1 - b^2) + \sigma_x^2 \sigma_y \sigma_z (ab - c) \\ \sigma_x \sigma_y^2 \sigma_z (ac - b) + \sigma_x^2 \sigma_y \sigma_z (ab - c) + \sigma_x^2 \sigma_y^2 (1 - a^2) \end{pmatrix} \\ &= \frac{1}{D} [\sigma_y^2 \sigma_z^2 (1 - c^2) + \sigma_x^2 \sigma_z^2 (1 - b^2) + \sigma_x^2 \sigma_y^2 (1 - a^2) + 2\sigma_x^2 \sigma_y \sigma_z (ab - c) \\ &\quad + 2\sigma_x \sigma_y^2 \sigma_z (ac - b) + 2\sigma_x \sigma_y \sigma_z^2 (bc - a)] \\ &:= \frac{\mathcal{H}}{D} \end{aligned} \quad (2.153)$$

Associant les équations (2.152) et (2.153), on trouve que les poids, dans le cas d'une matrice 3×3 , sont donnés par :

$$w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} \frac{\sigma_y^2 \sigma_z^2 (1 - c^2) + \sigma_x \sigma_y \sigma_z^2 (bc - a) + \sigma_x \sigma_y^2 \sigma_z (ac - b)}{\sigma_x \sigma_y \sigma_z^2 (bc - a) + \sigma_x^2 \sigma_z^2 (1 - b^2) + \sigma_x^2 \sigma_y \sigma_z (ab - c)} \mathcal{H} \\ \frac{\sigma_x \sigma_y \sigma_z^2 (ac - b) + \sigma_x^2 \sigma_y \sigma_z (ab - c) + \sigma_x^2 \sigma_y^2 (1 - a^2)}{\mathcal{H}} \end{pmatrix} \quad (2.154)$$

où \mathcal{H} est défini par

$$\mathcal{H} = \sigma_y^2 \sigma_z^2 (1 - c^2) + \sigma_x^2 \sigma_z^2 (1 - b^2) + \sigma_x^2 \sigma_y^2 (1 - a^2) + 2\sigma_x^2 \sigma_y \sigma_z (ab - c) + 2\sigma_x \sigma_y^2 \sigma_z (ac - b) + 2\sigma_x \sigma_y \sigma_z^2 (bc - a) \quad (2.155)$$

2.6.7.2 Illustration de la méthode dans le cas d'une matrice 3×3

Comme dans le cas d'une matrice 2×2 , nous avons effectué une application numérique sous MATLAB dans le cas d'une matrice 3×3 , pour déterminer les poids d'un portefeuille. Pour cela, avec comme données :

$$\begin{cases} \sigma_x = 0.15 \\ \sigma_y = 0.25 \\ \sigma_z = 0.6 \end{cases} \quad \text{et} \quad \begin{cases} \rho_{xy} = 0.75 \\ \rho_{xz} = 0.4 \\ \rho_{yz} = 0.1 \end{cases} \quad (2.156)$$

nous avons obtenu que :

➡ La matrice A s'écrit :

$$A = \begin{pmatrix} \frac{1}{\sigma_x} & 0 & 0 \\ 0 & \frac{1}{\sigma_y} & 0 \\ 0 & 0 & \frac{1}{\sigma_z} \end{pmatrix} = \begin{pmatrix} 6.6667 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1.6667 \end{pmatrix} \quad (2.157)$$

➡ La matrice de variance-covariance s'écrit

$$\begin{aligned} \Sigma &= \begin{pmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z^2 \end{pmatrix} = \begin{pmatrix} \sigma_x^2 & \rho_{xy} \sigma_x \sigma_y & \rho_{xz} \sigma_x \sigma_z \\ \rho_{xy} \sigma_x \sigma_y & \sigma_y^2 & \rho_{yz} \sigma_y \sigma_z \\ \rho_{xz} \sigma_x \sigma_z & \rho_{yz} \sigma_y \sigma_z & \sigma_z^2 \end{pmatrix} \\ &= \begin{pmatrix} 0.0225 & 0.0281 & 0.0360 \\ 0.0281 & 0.0625 & 0.0150 \\ 0.0360 & 0.0150 & 0.36 \end{pmatrix} \end{aligned} \quad (2.158)$$

➡ La matrice de corrélation est donnée par

$$\begin{aligned} C &= \begin{pmatrix} 1 & \rho_{xy} & \rho_{xz} \\ \rho_{xy} & 1 & \rho_{yz} \\ \rho_{xz} & \rho_{yz} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0.75 & 0.4 \\ 0.75 & 1 & 0.1 \\ 0.4 & 0.1 & 1 \end{pmatrix} \\ &= A \Sigma A = \begin{pmatrix} 1 & 0.75 & 0.4 \\ 0.75 & 1 & 0.1 \\ 0.4 & 0.1 & 1 \end{pmatrix} \end{aligned} \quad (2.159)$$

➡ en appliquant la formule (2.56), les poids

$$w = \begin{pmatrix} 1.3217 \\ -0.2563 \\ -0.0655 \end{pmatrix} \quad (2.160)$$

où $e = (1 \ 1 \ 1)'$ dans la formule de l'équation (2.56). On remarque bien que la somme des w_i vaut bien 1.

Nous voulons maintenant appliquer la méthode proposée dans la RMT. Pour cela, nous devons tout d'abord calculer les valeurs propres (lesquelles seront modifiées) de la matrice de corrélation C et les vecteurs propres associés.

➡ Le calcul des valeurs propres de la matrice de corrélation donne :

$$\begin{cases} \lambda_1 = 0.1885 \\ \lambda_2 = 0.9173 \\ \lambda_3 = 1.8942 \end{cases} \quad (2.161)$$

➡ Si on note Λ la matrice des valeurs propres λ_1 , λ_2 et λ_3 ci-dessus, on a :

$$\Lambda = \begin{pmatrix} 0.1885 & 0 & 0 \\ 0 & 0.9173 & 0 \\ 0 & 0 & 1.8942 \end{pmatrix}. \quad (2.162)$$

➡ Les vecteurs propres associés sont

$$v_1 = \begin{pmatrix} 0.7223 \\ -0.6333 \\ -0.2780 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0.0667 \\ 0.4638 \\ -0.8834 \end{pmatrix} \quad \text{et} \quad v_3 = \begin{pmatrix} 0.6884 \\ 0.6196 \\ 0.3772 \end{pmatrix} \quad (2.163)$$

Notons v la matrice dont les colonnes sont les vecteurs propres v_1 , v_2 et v_3 :

$$v = \begin{pmatrix} 0.7223 & 0.0667 & 0.6884 \\ -0.6333 & 0.4638 & 0.6196 \\ -0.2780 & -0.8834 & 0.3772 \end{pmatrix}. \quad (2.164)$$

➡ Modification des valeurs propres : On testera les deux méthodes citées précédemment. Premièrement, on prendra les deux premières valeurs propres égales à 0 et on laissera la dernière valeur propre intacte, et deuxièmement, on remplacera les deux premières valeurs propres par leur moyenne, et on laissera la troisième intacte. Dans le premier cas, on a :

☞ les nouvelles valeurs propres sont

$$\begin{cases} \lambda_1^{new} = 0 \\ \lambda_2^{new} = 0 \\ \lambda_3^{new} = 1.8942 \end{cases} \quad (2.165)$$

notons Λ^{new} la matrice des nouvelles valeurs propres :

$$\Lambda^{new} = \begin{pmatrix} \lambda_1^{new} & 0 & 0 \\ 0 & \lambda_2^{new} & 0 \\ 0 & 0 & \lambda_3^{new} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1.8942 \end{pmatrix} \quad (2.166)$$

☞ La matrice de pseudo-corrélation dans la base intermédiaire :

$$\begin{aligned} \overline{C} &= v * \Lambda^{new} * v' \\ &= \begin{pmatrix} 0.9017 & 0.8079 & 0.4919 \\ 0.8079 & 0.7271 & 0.4427 \\ 0.4919 & 0.4427 & 0.2695 \end{pmatrix} \end{aligned} \quad (2.167)$$

En effet, cette matrice a été obtenue par les calculs qui suivent :

$$\begin{aligned} \overline{C} &= \begin{pmatrix} 0.7223 & 0.0667 & 0.6884 \\ -0.6333 & 0.4638 & 0.6196 \\ -0.2780 & -0.8834 & 0.3772 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1.8942 \end{pmatrix} \begin{pmatrix} 0.7223 & -0.6333 & -0.2780 \\ 0.0667 & 0.4638 & -0.8834 \\ 0.6884 & 0.6196 & 0.3772 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 1.3040 \\ 0 & 0 & 1.1736 \\ 0 & 0 & 0.7145 \end{pmatrix} \begin{pmatrix} 0.7223 & -0.6333 & -0.2780 \\ 0.0667 & 0.4638 & -0.8834 \\ 0.6884 & 0.6196 & 0.3772 \end{pmatrix} \\ &= \begin{pmatrix} 0.9017 & 0.8079 & 0.4919 \\ 0.8079 & 0.7271 & 0.4427 \\ 0.4919 & 0.4427 & 0.2695 \end{pmatrix} \end{aligned}$$

- ☞ Nouvelle matrice de corrélation : Cette matrice représente la matrice \overline{C} de l'équation (2.167) dans laquelle on a remplacé les éléments de la diagonale par des 1

$$C^{new} = \begin{pmatrix} 1 & 0.8079 & 0.4919 \\ 0.8079 & 1 & 0.4427 \\ 0.4919 & 0.4427 & 1 \end{pmatrix} \quad (2.168)$$

- ☞ La nouvelle matrice de variance-covariance

$$\Sigma^{new} = A^{-1} C^{new} A^{-1} \quad (2.169)$$

$$= \begin{pmatrix} 0.0225 & 0.0303 & 0.0443 \\ 0.0303 & 0.0625 & 0.0664 \\ 0.0443 & 0.0664 & 0.3600 \end{pmatrix} \quad (2.170)$$

- ☞ en appliquant la formule (2.56) avec la matrice Σ^{new} ci-dessus, on obtient les poids

$$w^{new} = \begin{pmatrix} 1.3441 \\ -0.2839 \\ -0.0602 \end{pmatrix}. \quad (2.171)$$

- ☞ On peut également regarder l'écart entre les différentes corrélations :

$$\Rightarrow \rho_{xy} = 0.75 \text{ et } \rho_{xy}^{new} = 0.8079$$

$$\Rightarrow \rho_{xz} = 0.9 \text{ et } \rho_{xz}^{new} = 0.4919$$

$$\Rightarrow \rho_{yz} = 0.1 \text{ et } \rho_{yz}^{new} = 0.4427$$

On remarque que cette opération sur les valeurs propres a permis de **diminuer considérablement un des trois coefficients de corrélation** tout en gardant intacte la trace de la matrice de corrélation.

- ➡ Dans le deuxième cas, i.e. le cas où on remplace les deux premières valeurs propres par leur moyenne, on aura (contrairement au cas des matrices de taille 2×2 , nous ne remplacerons pas toutes les valeurs propres par leur moyenne, mais juste deux des trois valeurs propres. Le raisonnement se rapproche un peu plus de celui qu'on appliquera avec des données réelles) :

- ☞ Les nouvelles valeurs propres sont

$$\begin{cases} \lambda_1^{new} = \frac{\lambda_1 + \lambda_2}{2} = 0.5529 \\ \lambda_2^{new} = \frac{\lambda_1 + \lambda_2}{2} = 0.5529 \\ \lambda_3^{new} = 1.8942 \end{cases} \quad (2.172)$$

notons Λ^{new} la matrice des nouvelles valeurs propres :

$$\Lambda^{new} = \begin{pmatrix} \lambda_1^{new} & 0 & 0 \\ 0 & \lambda_1^{new} & 0 \\ 0 & 0 & \lambda_3^{new} \end{pmatrix} = \begin{pmatrix} 0.5529 & 0 & 0 \\ 0 & 0.5529 & 0 \\ 0 & 0 & 1.8942 \end{pmatrix} \quad (2.173)$$

- ☞ La matrice de pseudo-corrélation dans la base intermédiaire (Notons qu'ici, la matrice des vecteurs propres (encore appelée matrice de passage de la première base vers la nouvelle base) v reste la même que celle de l'équation (2.164)) :

$$\begin{aligned} \overline{C} &= v * \Lambda^{new} * v' \\ &= \begin{pmatrix} 1.1885 & 0.5720 & 0.3483 \\ 0.5720 & 1.0678 & 0.3135 \\ 0.3483 & 0.3135 & 0.7438 \end{pmatrix} \end{aligned} \quad (2.174)$$

En effet, pour obtenir ce résultat, on a effectué les calculs ci-dessous :

$$\begin{aligned}
\overline{C} &= \begin{pmatrix} 0.7223 & 0.0667 & 0.6884 \\ -0.6333 & 0.4638 & 0.6196 \\ -0.2780 & -0.8834 & 0.3772 \end{pmatrix} \begin{pmatrix} 0.5529 & 0 & 0 \\ 0 & 0.5529 & 0 \\ 0 & 0 & 1.8942 \end{pmatrix} \begin{pmatrix} 0.7223 & -0.6333 & -0.2780 \\ 0.0667 & 0.4638 & -0.8834 \\ 0.6884 & 0.6196 & 0.3772 \end{pmatrix} \\
&= \begin{pmatrix} 0.3994 & 0.0369 & 1.3040 \\ -0.3502 & 0.2564 & 1.1736 \\ -0.1537 & -0.4884 & 0.7145 \end{pmatrix} \begin{pmatrix} 0.7223 & -0.6333 & -0.2780 \\ 0.0667 & 0.4638 & -0.8834 \\ 0.6884 & 0.6196 & 0.3772 \end{pmatrix} \\
&= \begin{pmatrix} 1.1885 & 0.5720 & 0.3483 \\ 0.5720 & 1.0678 & 0.3135 \\ 0.3483 & 0.3135 & 0.7438 \end{pmatrix}
\end{aligned}$$

☞ Nouvelle matrice de corrélation : Cette matrice représente la matrice \overline{C} de l'équation (2.174) dans laquelle on a remplacé les éléments de la diagonale par des 1

$$C^{new} = \begin{pmatrix} 1 & 0.5720 & 0.3483 \\ 0.5720 & 1 & 0.3135 \\ 0.3483 & 0.3135 & 1 \end{pmatrix} \quad (2.175)$$

☞ La nouvelle matrice de variance-covariance

$$\Sigma^{new} = A^{-1} C^{new} A^{-1} \quad (2.176)$$

$$= \begin{pmatrix} 0.0225 & 0.0215 & 0.0313 \\ 0.0215 & 0.0625 & 0.0470 \\ 0.0313 & 0.0470 & 0.3600 \end{pmatrix} \quad (2.177)$$

☞ en appliquant la formule (2.56) avec la matrice Σ^{new} ci-dessus, on obtient les poids

$$w^{new} = \begin{pmatrix} 0.9929 \\ 0.0366 \\ -0.0296 \end{pmatrix}. \quad (2.178)$$

☞ On peut également regarder l'écart entre les différentes corrélations :

$$\Leftrightarrow \rho_{xy} = 0.75 \text{ et } \rho_{xy}^{new} = 0.5720$$

$$\Leftrightarrow \rho_{xz} = 0.9 \text{ et } \rho_{xz}^{new} = 0.3483$$

$$\Leftrightarrow \rho_{yz} = 0.1 \text{ et } \rho_{yz}^{new} = 0.3135$$

On remarque que cette opération sur les valeurs propres a permis de **diminuer "considérablement" deux des trois coefficients de corrélation** tout en gardant intacte la trace de la matrice de corrélation.

2.6.7.3 Généralisation dans le cas d'une matrice $n \times n$

Dans cette partie, nous cherchons à généraliser au cas des matrices $n \times n$, la méthode démontrée précédemment pour des matrices 2×2 et 3×3 . Étant donné qu'il n'est pas facile de calculer analytiquement l'inverse d'une matrice symétrique définie positive, nous allons conjecturer la méthode pour de telles matrices. La méthode que nous utiliserons se présentera comme suit :

☞ On se donne n volatilités $\{\sigma_{ii}, i = 1, \dots, n\}$, et $\frac{n^2 - n}{2}$ covariances $\{\sigma_{ij}, i, j = 1, \dots, n, \text{ et } i < j\}$.

La matrice de variance-covariance s'écrit ici :

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \ddots & & \vdots \\ \vdots & & \ddots & \\ \sigma_{1n} & \cdots & \sigma_{1n} & \sigma_{nn} \end{pmatrix} \quad (2.179)$$

- ☞ Dans la matrice de variance-covariance Σ , on récupère les n volatilités $\{\sigma_{ii}, i = 1, \dots, n\}$ et on construit la matrice A , qui est une matrice diagonale, et dans laquelle la diagonale est formée par les inverses des dites volatilités :

$$A = \begin{pmatrix} \frac{1}{\sigma_{11}} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \\ 0 & 0 & \dots & \frac{1}{\sigma_{nn}} \end{pmatrix} \quad (2.180)$$

- ☞ On calcule la matrice de corrélation C par la formule $C = A\Sigma A$
 ☞ On calcule les n valeurs propres de la matrice de corrélation, et on construit la matrice diagonale des valeurs propres D :

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \quad (2.181)$$

- ☞ On calcule les n vecteurs propres associés aux valeurs propres précédentes, et on construit la matrice P , à partir de laquelle on obtiendra P' :

$$P = \begin{pmatrix} v_{11} & 0 & \dots & v_{1n} \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \\ v_{n1} & 0 & \dots & v_{nn} \end{pmatrix} \quad \text{d'où} \quad P' = \begin{pmatrix} v_{11} & 0 & \dots & v_{n1} \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \\ v_{1n} & 0 & \dots & v_{nn} \end{pmatrix} \quad (2.182)$$

- ☞ On calcule les coefficients de la matrice C en fonction de ses valeurs propres et des composantes de ses vecteurs propres par la formule $C = PDP'$. En utilisant la formule de l'équation (2.57), on trouve que

$$(C)_{ij} = \sum_{k=1}^n \lambda_k v_{ik} v_{jk} \quad (2.183)$$

- ☞ On calcule la matrice Σ à l'aide de la précédente matrice C obtenue dans (2.183). On sait que cette matrice est symétrique, donc nous n'avons qu'à calculer la partie inférieure de la matrice. Pour cela, on utilise la formule ci-dessous :

$$(\Sigma)_{ij} = \sigma_{ii}\sigma_{jj} \sum_{k=1}^n \lambda_k v_{ik} v_{jk} \quad \text{pour } i \geq j. \quad (2.184)$$

- ☞ Une fois la matrice C obtenue, on modifie les valeurs propres pour obtenir la matrice intermédiaire \overline{C} . Pour cela, on va remplacer les valeurs propres $\{\lambda_i, i = j, \dots, n\}$ (où j est l'indice de la première valeur propre inférieure à λ_{max} , la borne supérieure du spectre de MP) dans l'équation (2.183) par les valeurs propres $\{\overline{\lambda}_i, i = j, \dots, n\}$. On aura alors :

$$(\overline{C})_{ij} = \sum_{k=1}^n \overline{\lambda}_k v_{ik} v_{jk} \quad \text{pour } i \geq j. \quad (2.185)$$

- ☞ Obtention de la matrice C^{new} : on modifie la diagonale de la matrice \overline{C} en remplaçant les éléments correspondants par des 1. On obtient alors :

$$(C^{new})_{ij} = \begin{cases} 1 & \text{Pour } i = j \\ \sum_{k=1}^n \overline{\lambda}_k v_{ik} v_{jk} & \text{pour } i > j \end{cases} \quad (2.186)$$

Pour trouver la partie supérieure de la matrice C^{new} , il suffit juste de se rappeler que cette matrice est symétrique.

- ☞ Toutes ces étapes effectuées, il ne reste plus qu'à trouver la matrice $\Sigma^{new} = A^{-1}C^{new}A^{-1}$, ce qui signifie que pour obtenir les coefficients de la matrice Σ^{new} , on aura juste à calculer $A^{-1} \times (2.186) \times A^{-1}$. On obtient alors :

$$(\Sigma^{new})_{ij} = \begin{cases} \sigma_{ij}^2 & \text{Pour } i = j \text{ (sur la diagonale)} \\ \sigma_{ii}\sigma_{jj} \sum_{k=1}^n \bar{\lambda}_k v_{ik} v_{jk} & \text{pour } i > j \end{cases} . \quad (2.187)$$

- ☞ La matrice Σ^{new} obtenue nous permet de calculer numériquement par la formule de l'équation (2.56) les nouveaux poids w^{new} .
- ☞ Toutes les matrices estimées et les poids obtenus par simulation, on peut aussi, comme dans le cas des matrices 2×2 et 3×3 , généraliser le résultat sur les coefficients de corrélation. On a alors :

$$\sigma_{ij}^{new} = \frac{\rho_{ij}^{new}}{\rho_{ij}} \sigma_{ij} \quad \text{pour } i > j \quad (2.188)$$

$$= \sigma_{ii}\sigma_{jj}\rho_{ij}^{new} \quad \text{pour } i > j. \quad (2.189)$$

Ce résultat peut également être exprimé en fonction des valeurs propres modifiées $\{\bar{\lambda}_k, k = 1, 2, \dots, n\}$. En effet, pour $i > j$, on a :

$$\sigma_{ij}^{new} = \sigma_{ii}\sigma_{jj}\rho_{ij}^{new} \quad (2.190)$$

$$\begin{aligned} &= \sigma_{ii}\sigma_{jj}(C^{new})_{ij} \\ &= \sigma_{ii}\sigma_{jj} \sum_{k=1}^n \bar{\lambda}_k v_{ik} v_{jk} \end{aligned} \quad (2.191)$$

N.B. : Notons que dans le cas des matrices de taille $n \times n$, le choix de modification des valeurs propres se fera toujours selon le choix de l'utilisateur du code, mais dépendra fortement de la borne supérieure de la densité des valeurs propres. En effet, dans le cas des matrices de Wishart, il a été démontré que cette borne vaut $\lambda_{max} = 1 + \frac{N}{T} + 2\sqrt{\frac{N}{T}}$. C'est ce résultat que nous utiliserons en pratique, dans le cas de données réelles. Pour obtenir la nouvelle matrice des valeurs propres Λ^{new} , on comparera les valeurs propres à λ_{max} , et on remplacera toutes celles qui sont inférieures à cette valeur soit par 0, soit par leur moyenne.

Suivant toutes les étapes décrites ci-dessous, un algorithme de calcul dans le cas d'une matrice $n \times n$ est proposé ici :

Algorithm 2 Algorithme du calcul des poids GMV dans le cas d'une matrice $n \times n$

Require: r, e

- 1: r : matrice des rendements
 - 2: e : vecteur unitaire de taille $(n, 1)$
 - 3: On estime la matrice de covariance Σ : Sous MATLAB, $\Sigma = cov(r)$
 - 4: Calcul de la matrice A par :
 - 5: **for** i allant de 1 à N **do**
 - 6: $A(i, i) = \frac{1}{\Sigma(i, i)}$
 - 7: **end for**
 - 8: On estime la matrice de corrélation C par $C = A \Sigma A$
 - 9: Calcul des poids GMV par la formule (2.56) et calcul du risque global du portefeuille
 - 10: Application de la RMT :
 - 11: Calcul des valeurs propres et des vecteurs propres de C par $[v, \Lambda] = eig(C)$, où λ est la matrice des valeurs propres (qui sont rangées dans l'ordre croissant) et v la matrice des vecteurs propres associés aux valeurs propres qui sont dans la matrice λ
 - 12: Calcul de la borne max des valeurs propres en utilisant le résultat de Marcenko-Pastur :
 $\lambda_{max} = 1 + \frac{N}{T} + 2\sqrt{\frac{N}{T}}$, laquelle nous permettra de modifier nos valeurs propres
 - 13: Calcul de la nouvelle matrice des valeurs propres $\bar{\lambda}$ en utilisant le résultat de l'étape 12. On distinguera deux cas : On remplacera premièrement les valeurs propres inférieures à λ_{max} par 0 et deuxièmement on remplacera ces valeurs propres par leur moyenne.
 - 14: **if** on veut remplacer ces valeurs propres par 0 **then**
 - 15: **for** i allant de 1 à N **do**
 - 16: **if** s est inférieur à λ_{max} (où $s = \Lambda(i, i)$) **then**
 - 17: $\bar{\Lambda}(i, i) = 0$
 - 18: **else**
 - 19: $\bar{\Lambda}(i, i) = \Lambda(i, i)$
 - 20: **end if**
 - 21: **end for**
 - 22: **end if**
 - 23: **if** on veut remplacer ces valeurs propres par leur moyenne **then**
 - 24: **for** i allant de 1 à N **do**
 - 25: **if** f est inférieur à λ_{max} (où on a posé $f = \Lambda(i, i)$) **then**
 - 26: On rajoute f au vecteur a : $a = [a \ f]$ où a a été initialisé comme étant vide
 - 27: **end if**
 - 28: **end for**
 - 29: on calcule la moyenne m du vecteur a par $m = mean(a)$
 - 30: **for** i allant de 1 à N **do**
 - 31: **if** s est inférieur à λ_{max} (avec $s = \Lambda(i, i)$) **then**
 - 32: $\bar{\Lambda}(i, i) = m$
 - 33: **else**
 - 34: $\bar{\Lambda}(i, i) = \Lambda(i, i)$
 - 35: **end if**
 - 36: **end for**
 - 37: **end if**
 - 38: Construction de la matrice de pseudo-corrélation intermédiaire par $\bar{C} = v * \bar{\Lambda} * v'$
 - 39: Reconstruction de la matrice de corrélation C^{new} obtenue de la matrice \bar{C} dans laquelle on a remplacé sa diagonale par des 1
 - 40: Reconstruction de la nouvelle matrice de covariance par $\Sigma^{new} = A^{-1} * C^{new} * A^{-1}$
 - 41: Calcul des nouveaux poids GMV à l'aide de (2.56) et de Σ^{new} et calcul du nouveau risque global du portefeuille
-

2.6.7.4 Application à des données réelles

La généralisation de cette méthode effectuée, nous l'avons testé sur des données (rendements de marché) réelles pour voir quel est son impact sur les différents poids et les différents coefficients de corrélation calculés. Nous avons tout d'abord considéré $N = 10$ titres, sur une période de temps $T = 783$ jours (3 ans). Utilisant l'algorithme précédent, nous avons calculé les poids de départ, et nous avons utilisé la méthode de la RMT pour obtenir les nouveaux poids. Les résultats ont été récapitulés dans le tableau 2.3 qui suit. Nous avons aussi rajouté dans ce tableau les différentes volatilités, pour regarder si le changement des poids a un lien avec celles-ci. Dans la suite, nous allons citer les différents résultats important pour l'obtention des poids finaux. On a :

- ☞ Après estimation de la matrice de covariance par la méthode décrite à la ligne 3 de l'algorithme, on a calculé la matrice de corrélation à l'aide de la formule citée à la ligne 8. Le calcul des valeurs propres de la matrice de cette matrice de corrélation donne les résultats qui suivent :

$$\Lambda = \begin{pmatrix} 5.8740 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.8371 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2302 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.2685 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5946 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5460 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.3439 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3920 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4423 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4713 \end{pmatrix}$$

- ☞ Calculer la matrice P des vecteurs propres associés aux valeurs propres ci-dessus et obtenir sa transposée.
- ☞ Calcul de la borne maximale des valeurs propres : $\lambda_{max} = 1 + \frac{1}{Q} + 2\sqrt{\frac{1}{Q}} = 1.2388$, où $Q = \frac{T}{N}$, avec N le nombre de titres, et T le nombre de jours considérés. On va ensuite comparer les valeurs propres à cette valeur λ_{max} . Dans un premier temps, on remplacera toutes les valeurs propres inférieures à cette dernière par leur moyenne.
- ☞ La matrice $\bar{\Lambda}$ des valeurs propres modifiées est la suivante :

$$\bar{\Lambda} = \begin{pmatrix} 5.8740 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.4584 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.4584 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.4584 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.4584 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.4584 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.4584 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4584 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4584 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4584 \end{pmatrix}$$

- ☞ Le calcul des nouvelles matrices de corrélation et de variance-covariance se font à l'aide de la matrice $\bar{\Lambda}$ ci-dessus et des lignes 38, 39 et 40 de l'algorithme.
- ☞ À l'aide de la ligne 41, nous avons calculé les nouveaux poids. Comme nous le disions un peu plus haut, nous avons porté dans le tableau qui suit les poids dans la base de départ et la nouvelle base et nous y avons rajouté les volatilités des différents titres. On a :
- ☞ Dans les deux tableaux qui suivent, nous avons comparé les différents coefficients de corrélation calculés dans la base de départ et dans la nouvelle base. On a les deux tableaux qui suivent :
On peut voir que sur des données réelles de marché, le résultat obtenu est un peu semblable à celui obtenu dans nos exemples précédents. En effet, dans les applications proposées dans les cas des matrices de dimension 2×2 et 3×3 , on a remarqué qu'avec l'application de la méthode de la RMT, on réduisait certains coefficients de corrélation. Ici, le constat est le même. On se rend compte que certains de ces coefficients augmentent et d'autres diminuent considérablement. Pour mieux illustrer nos dires, observons la différence entre les matrices C et C^{new} , donnée dans le tableau qui suit :

titres	w	w^{new}	volatilités
1	-0.1228	-0.1013	0.0253
2	0.0287	-0.0412	0.0254
3	0.1958	0.1646	0.0191
4	0.1419	0.2293	0.0173
5	0.1233	0.0371	0.0208
6	-0.0820	-0.0829	0.0244
7	0.1164	0.1304	0.0194
8	0.5145	0.5454	0.0140
9	0.1972	0.2353	0.0178
10	-0.1131	-0.1168	0.0287

TABLE 2.3 – Poids GMV dans l’ancienne (w) et dans la nouvelle (w^{new}) base et volatilités des actifs (notons que les volatilités sont égales dans l’ancienne et la nouvelle base : $\sigma_{ii}^{new} = \sigma_{ii}$) : les nouveaux poids ont été obtenus en remplaçant les valeurs propres inférieures à λ_{max} par leur moyenne.

ρ_{ij}									
	2	3	4	5	6	7	8	9	10
1	0.6161	0.5321	0.6881	0.6049	0.6204	0.5961	0.4286	0.5479	0.6498
2		0.5247	0.6448	0.5648	0.6091	0.5326	0.2903	0.4361	0.5736
3			0.5670	0.4922	0.5487	0.4720	0.3352	0.4234	0.5004
4				0.6779	0.6692	0.6164	0.5065	0.4963	0.6026
5					0.7537	0.5793	0.4060	0.4386	0.6434
6						0.5862	0.3831	0.4566	0.6298
7							0.4224	0.4750	0.5567
8								0.4673	0.4097
9									0.5153

TABLE 2.4 – Coefficients de corrélation dans la base de départ

ρ_{ij}^{new}									
	2	3	4	5	6	7	8	9	10
1	0.5841	0.5371	0.6514	0.6230	0.6328	0.5849	0.4483	0.5164	0.6119
2		0.4957	0.6012	0.5750	0.5841	0.5398	0.4138	0.4766	0.5648
3			0.5528	0.5287	0.5370	0.4963	0.3805	0.4382	0.5193
4				0.6413	0.6514	0.6020	0.4615	0.5316	0.6298
5					0.6230	0.5758	0.4414	0.5084	0.6024
6						0.5848	0.4483	0.5164	0.6119
7							0.4143	0.4772	0.5655
8								0.3658	0.4335
9									0.4993

TABLE 2.5 – Coefficients de corrélation dans la nouvelle base, dans le cas où on a remplacé les valeurs propres inférieures à λ_{max} par leur moyenne.

Commentaires sur le lien entre les poids (anciens et nouveaux) et les volatilités :

Au vu des résultats cités dans le tableau 2.3, on peut noter que, lorsque la volatilité est élevée (comparé aux autres volatilités du tableau 2.3), l’application de la méthode de la RMT (ici, la

$\rho_{ij} - \rho_{ij}^{new}$									
	2	3	4	5	6	7	8	9	10
1	0.0319	-0.0050	0.0366	-0.0181	-0.0124	0.0113	-0.0198	0.0314	0.0379
2		0.0290	0.0436	-0.0102	0.0250	-0.0072	-0.1235	-0.0405	0.0089
3			0.0142	-0.0365	0.0116	-0.0243	-0.0453	-0.0148	-0.0189
4				0.0366	0.0178	0.0144	0.0450	-0.0352	-0.0273
5					0.1307	0.0035	-0.0354	-0.0698	0.0410
6						0.0014	-0.0652	-0.0598	0.0179
7							0.0081	-0.0023	-0.0087
8								0.1015	-0.0238
9									0.0160

TABLE 2.6 – Différence entre les coefficients de corrélation dans l'ancienne base et les coefficients de corrélation dans la nouvelle base.

méthode a été utilisée en remplaçant les valeurs propres inférieures à λ_{max} par leur moyenne) permet de réduire les poids. On peut le voir par exemple sur les titres 2 et 5. Pour le titre 10, on peut noter un léger changement au niveau du poids (l'ancien poids était -0.1131 et le nouveau est -0.1168).

On note également dans le tableau 2.3 que, lorsque la volatilité est petite (également comparé aux autres volatilités du tableau 2.3), en utilisant la méthode de la RMT (ici, la méthode a été utilisée en remplaçant les valeurs propres inférieures à λ_{max} par leur moyenne), les poids ont tendance à augmenter. Notons par exemple les cas des titres 4, 7, 8 et 9. On constate que dans ces cas, les nouveaux poids sont quand même assez élevés comparés aux poids dans l'ancienne base. **On peut se dire qu'étant donné que ces titres ne sont pas trop risqués, on pourrait y investir beaucoup plus, et donc leur donner des poids importants. Dans le cas où les titres sont très risqués, la méthode de la RMT permet de diminuer le poids du titre, donc la somme à y investir.**

- ☞ Maintenant, nous allons remplacer toutes les valeurs propres inférieures à la valeur λ_{max} par 0. La matrice $\bar{\Lambda}$ des valeurs propres modifiées devient alors :

$$\bar{\Lambda} = \begin{pmatrix} 5.8740 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

- ☞ Le calcul des nouvelles matrices de corrélation et de variance-covariance se font à l'aide de la matrice $\bar{\Lambda}$ ci-dessus et des lignes 38, 39 et 40 de l'algorithme.
- ☞ À l'aide de la ligne 41, nous avons calculé les nouveaux poids. Comme nous le disions un peu plus haut, nous avons porté dans le tableau qui suit les poids dans la base de départ et la nouvelle base et nous y avons rajouté les volatilités des différents titres. On a :
- ☞ Nous allons maintenant comparer les différents coefficients de corrélations, ceux de l'ancienne et ceux de la nouvelle base. Les coefficients de corrélation de la base de départ restent inchangés (voir tableau 2.4). Les nouveaux coefficients de corrélation sont donnés dans le tableau qui suit : On peut voir que sur des données réelles de marché, en appliquant la méthode de la RMT dans laquelle on remplace les valeurs propres inférieures à λ_{max} par 0, les résultats obtenus ne sont pas tout à fait semblables à ceux obtenus dans nos exemples précédents. En effet, dans les applications proposées dans les cas des matrices de dimension 2×2 et 3×3 , on a remarqué qu'avec l'application

titres	w	w^{new}	volatilités
1	-0.1228	-0.1247	0.0253
2	0.0287	-0.0501	0.0254
3	0.1958	0.1718	0.0191
4	0.1419	0.2644	0.0173
5	0.1233	0.0353	0.0208
6	-0.0820	-0.1035	0.0244
7	0.1164	0.1381	0.0194
8	0.5145	0.5611	0.0140
9	0.1972	0.2452	0.0178
10	-0.1131	-0.1375	0.0287

TABLE 2.7 – Poids GMV dans l’ancienne (w) et dans la nouvelle (w^{new}) base et volatilités des actifs (notons que les volatilités sont égales dans l’ancienne et la nouvelle base : $\sigma_{ii}^{new} = \sigma_{ii}$) : les valeurs propres inférieures à λ_{max} ont été remplacées par 0.

ρ_{ij}^{new}									
	2	3	4	5	6	7	8	9	10
1	0.6336	0.5825	0.7066	0.6758	0.6864	0.6344	0.4863	0.5601	0.6637
2		0.5377	0.6521	0.6237	0.6335	0.5855	0.4488	0.5170	0.6126
3			0.5996	0.5735	0.5825	0.5383	0.4127	0.4753	0.5632
4				0.6956	0.7065	0.6530	0.5005	0.5766	0.6832
5					0.6757	0.6245	0.4787	0.5514	0.6534
6						0.6343	0.4863	0.5601	0.6637
7							0.4494	0.5176	0.6133
8								0.3968	0.4702
9									0.5416

TABLE 2.8 – Coefficients de corrélation dans la nouvelle base, dans le cas où on a remplacé les valeurs propres inférieures à λ_{max} par 0 dans la nouvelle matrice des valeurs propres.

de la méthode de la RMT, on réduisait certains coefficients de corrélation. Ici, le constat est différent. On remarque que plus des 3/4 des coefficients de corrélation ont augmenté. Pour mieux voir ce résultat, nous observerons la différence entre les matrices C et C^{new} (donc les différences entre les corrélations ρ_{ij} et ρ_{ij}^{new} , pour $i < j$), dans le tableau 2.9 qui suit. Mais avant de regarder cette différence, faisons quelques commentaires sur l’impact de la méthode de la RMT (lorsqu’on remplace les valeurs propres inférieures à λ_{max} par 0) sur les poids des titres. On essaiera de regarder si le fait de mettre à 0 les valeurs propres inférieures à λ_{max} a le même effet que de les remplacer par leur moyenne.

Commentaires sur le lien entre les poids (anciens et nouveaux) et les volatilités :

Dans le tableau 2.7, le constat est à peu près le même que dans le tableau 2.3. On constate que, pour les volatilités élevées (comparés à toutes les autres volatilités citées), les nouveaux poids sont nettement inférieurs aux anciens. Ce résultat s’observe par exemple sur les titres 1, 2, 6, et 10. C’est dire qu’en appliquant la méthode de la RMT dans laquelle on remplace les valeurs propres inférieures à λ_{max} par 0, le résultat est à peu près équivalent à celui obtenu lorsqu’on remplace les valeurs propres inférieures à λ_{max} par leur moyenne. On réduit les poids des titres qui ont une forte volatilité.

On note également dans le tableau 2.7 que, lorsque la volatilité est petite (également comparé aux autres volatilités du tableau 2.7), en utilisant la méthode de la RMT (rappelons juste qu’ici, la méthode a été utilisée en remplaçant les valeurs propres inférieures à λ_{max} par 0), les poids ont

	$\rho_{ij} - \rho_{ij}^{new}$								
	2	3	4	5	6	7	8	9	10
1	-0.0175	-0.0505	-0.0185	-0.0709	-0.0660	-0.0382	-0.0577	-0.0123	-0.0139
2		-0.0129	-0.0073	-0.0589	-0.0244	-0.0529	-0.1585	-0.0808	-0.0389
3			-0.0326	-0.0813	-0.0338	-0.0663	-0.0775	-0.0519	-0.0629
4				-0.0177	-0.0373	-0.0366	0.0059	-0.0802	-0.0806
5					0.0780	-0.0452	-0.0728	-0.1129	-0.0100
6						-0.0481	-0.1031	-0.1035	-0.0339
7							-0.0270	-0.0427	-0.0566
8								0.0705	-0.0605
9									-0.0263

TABLE 2.9 – Différence entre les coefficients de corrélation dans l'ancienne base et les coefficients de corrélation dans la nouvelle base

tendance à augmenter. Notons par exemple les cas des titres 4, 7, 8 et 9. On constate que dans ces cas, les nouveaux poids sont quand même assez élevés comparés aux poids dans l'ancienne base. **On peut se dire qu'étant donné que ces titres ne sont pas trop risqués, on pourrait y investir beaucoup plus, et donc leur donner des poids importants. Dans le cas où les titres sont très risqués, la méthode de la RMT permet de diminuer le poids du titre, donc la somme à y investir.**

Chapitre 3

Méthode d'estimation du modèle financier, Transformée de Stieltjes et Application à la Gestion des Portefeuilles

3.1 Introduction

Dans les chapitres précédents, nous nous sommes atardés à démontrer des résultats importants en RMT, et à proposer quelques exemples simples tels que l'approximation des histogrammes des valeurs propres des matrices de variance-covariance pour des données qui suivent différentes lois par la densité de Marcenko-Pastur, et le calcul des poids des titres dans un portefeuille lorsqu'on effectue un "changement de base". Nous allons à présent essayer d'améliorer les modèles (en perturbant nos données), proposer une méthode d'estimation du modèle financier, et nous appliquerons cette méthode à l'exemple cité dans la partie 3.6.4.2. La méthode d'estimation utilisera un algorithme bien connu en statistique : l'algorithme EM. Dans la littérature, pour faire de l'approximation des matrices de variance-covariance, bon nombre d'auteurs utilisent cet algorithme. Une description concise de cet algorithme sera donnée ici pour nos fins. Depuis le début de notre travail, nous considérons dans nos cas d'étude que la matrice $\frac{G'G}{T}$ est une "bonne" approximation de la matrice de variance-covariance (et donc une variable aléatoire...). Dans ce chapitre, après avoir présenté les différentes façon d'enrichir notre modèle, nous donnerons des résultats importants sur le produit de deux matrices aléatoires, pour ensuite présenter le modèle financier et enfin utiliser l'algorithme EM pour l'estimation des paramètres de ce modèle. Mais avant d'estimer les paramètres du modèle en utilisant cet algorithme, nous ferons des exemples d'applications sur des variables de Student, pour, au vu des résultats obtenus, valider notre choix d'utiliser cet algorithme.

3.2 Enrichissement du modèle : Somme et Produit de deux matrices aléatoires

Comme nous l'avons dit dans l'introduction, nous aimerons enrichir le modèle étudié en additionnant ou en multipliant nos données initiales par une petite perturbation. Pour cela, on étudiera deux cas : un premier cas dans lequel on rajoute une perturbation à nos données initiales, et un deuxième cas dans lequel on multipliera nos données initiales par une autre matrice (également une petite perturbation). Soient X nos v.a. initiales, et soient ϵ des v.a. qu'on supposera être très petites. On suppose que les v.a. X et ϵ sont indépendantes. On considèrera que l'espérance de ϵ est nulle. Notons E_Y , E_X et E_ϵ les estimateurs des différentes matrices de covariance de Y , X et ϵ .

3.2.1 Somme de deux matrices aléatoires : cas où $Y = X + \epsilon$

On aimerait savoir quelle est la relation entre la matrice de variance-covariance de Y , X et ϵ dans le cas où la variable aléatoire Y s'écrit comme la somme des variables aléatoires X et ϵ . Notre but

ici est de savoir si on peut écrire sous une forme facilement manipulable, l'expression de la matrice de variance-covariance de Y en fonction des matrices de variance-covariance de X et ϵ . On a :

$$\begin{aligned}
Y = X + \epsilon &\iff E_Y = \frac{1}{T} Y' Y = \frac{1}{T} (X + \epsilon)' (X + \epsilon) \\
&= \frac{1}{T} (X' + \epsilon') (X + \epsilon) \\
&= \frac{1}{T} (X' X + X' \epsilon + \epsilon' X + \epsilon' \epsilon) \\
&= \frac{1}{T} X' X + \frac{1}{T} \epsilon' \epsilon + \frac{1}{T} X' \epsilon + \frac{1}{T} \epsilon' X \\
E_Y &= E_X + E_\epsilon + \frac{1}{T} X' \epsilon + \frac{1}{T} \epsilon' X
\end{aligned}$$

Le résultat montre que, la matrice de variance-covariance de Y s'écrit comme la somme des matrices de variance-covariance de X et ϵ à laquelle on rajoute d'autres matrices ($\frac{1}{T} X' \epsilon$ et $\frac{1}{T} \epsilon' X$). Étant donné qu'on ne sait pas quelles sont les propriétés de ces deux matrices, ce résultat n'est pas très parlant. La question qui se pose alors est celle de savoir si l'on peut numériquement et "sans problèmes" perturber nos données en utilisant l'opération somme ?

3.2.2 Produit de deux matrices aléatoires : cas où $Y = X\epsilon$

Considérons dans cette partie que les entrées commutent. Comme dans la partie précédente, on aimerait savoir comment s'écrit la matrice de variance-covariance de Y en fonction des matrices de variance-covariance de X et ϵ . On a :

$$\begin{aligned}
Y = X\epsilon &\iff E_Y = \frac{1}{T} Y' Y = \frac{1}{T} (X\epsilon)' (X\epsilon) \\
&= \frac{1}{T} (\epsilon' X') (X\epsilon) \\
&= \frac{1}{T} \epsilon' X' X \epsilon \\
&= \frac{1}{T} X' X \epsilon' \epsilon \\
&= \frac{1}{T} X' X \frac{T}{T} \epsilon' \epsilon \\
E_Y &= T E_X E_\epsilon
\end{aligned}$$

Ce résultat nous montre que, la matrice de variance-covariance de Y vaut le produit des matrices de variance-covariance de X et ϵ , multiplié par T . Ce résultat est très intéressant étant donné que tous les termes sont explicitement connus. Dans les applications, on pourra donc, en utilisant ce calcul, améliorer les modèles en perturbant les données initiales (X) par d'autres données astucieusement choisies (ϵ).

3.2.3 Conclusion

Au vu des résultats obtenus dans les parties 3.2.1 et 3.2.2, nous ne nous intéresserons qu'au cas où, l'on perturbe les données en effectuant le produit $X\epsilon$. En effet, comme nous l'avons vu dans les calculs effectués ci-dessus, dans le cas où on perturbe les données en effectuant la somme, l'expression de la matrice de covariance de Y est difficilement exploitable, alors que, dans le cas de la multiplication, on obtient le produit des matrices de covariance de X et ϵ pondéré par un facteur T .

3.3 Produit de deux matrices aléatoires

3.3.1 Introduction

Dans cette partie, nous allons considérer l'ESD du produit de deux matrices aléatoires, l'une étant une matrice de covariance et l'autre une matrice hermitienne arbitraire. Cette étude se fera en deux parties : la première sera consacrée à l'étude de la LSD de la F -matrice multivariée, qui est le produit d'une matrice de covariance et de l'inverse d'une autre matrice de covariance, indépendantes l'une de

l'autre, et la deuxième sur l'investigation de la LSD d'une matrice de covariance lorsque la population de la matrice de covariance est arbitraire.

Le travail de base dans cette étude a été effectué par Watcher [16], qui considérait la distribution limite des solutions de l'équation

$$\det(X_{1,n_1}X'_{1,n_1} - \lambda X_{2,n_2}X'_{2,n_2}) = 0, \quad (3.1)$$

où X_{j,n_j} est une matrice $p \times n_j$ dont les entrées sont i.i.d. de loi $\mathcal{N}(0, 1)$, et X_{1,n_1} indépendant de X_{2,n_2} . Lorsque $X_{2,n_2}X'_{2,n_2}$ est de rang maximal, les solutions de 3.1 sont $\frac{n_2}{n_1}$ fois les valeurs propres de la F -matrice multivariée

$$\left(\frac{1}{n_1} X_{1,n_1} X'_{1,n_1} \right) \left(\frac{1}{n_2} X_{2,n_2} X'_{2,n_2} \right)^{-1}. \quad (3.2)$$

Yin et Krishnaiah [18] ont établi l'existence de la LSD de la suite matricielle $\{S_n T_n\}$, où S_n est une matrice de Wishart standard de dimension p et n degrés de liberté, T_n une matrice définie positive satisfaisant

$$\begin{cases} p/n & \rightarrow y \in (0, \infty) \\ \beta_k(T_n) & \rightarrow H_k \end{cases} \quad (3.3)$$

et où la suite H_k satisfait la condition de Carleman (Théorème 18). Yin [17] a généralisé ce résultat dans le cas d'une matrice de covariance de variables aléatoires i.i.d. réelles de moyenne 0 et de variance 1. Utilisant les résultats de Yin et Krishnaiah [18], Yin, Bai et Krishnaiah [19] ont montré l'existence de la LSD de la F -matrice multivariée. La forme explicite de la LSD des F -matrices était obtenue dans Bai, Yin et Krishnaiah [39] et Silverstein [46].

3.3.2 Résultats principaux

Présentons les résultats suivants :

3.3.2.1 Résultats

Théorème 4. *Supposons que les entrées de X_n ($p \times n$) sont des variables aléatoires complexes indépendantes satisfaisant (83), que T_n est une suite de matrices hermitiennes indépendantes de X_n , et que l'ESD de T_n converge vers une limite non aléatoire F^T dans un certain sens (probabilité ou presque sûre). Si $p/n \rightarrow y \in (0, \infty)$, alors l'ESD du produit $S_n T_n$ converge en probabilité ou presque sûrement vers une limite non aléatoire, où $S_n = \frac{1}{n} X_n X_n^*$.*

Remarque 1. *Notons que les valeurs propres du produit matriciel $S_n T_n$ sont toutes réelles, bien que cette matrice n'est pas réelle, parce que l'ensemble des valeurs propres de ce produit matriciel est le même que celui de la matrice symétrique $S_n^{1/2} T_n S_n^{1/2}$.*

Ce Théorème contient le résultat de Yin comme cas spécial. Yin [17], a considéré que les entrées de X sont réelles et i.i.d. de moyenne zéro et de variance 1, et la matrice T_n est réelle définie positive, et satisfait, pour chaque k fixé, à

$$\frac{1}{p} \text{tr}(T_n^k) \rightarrow H_k \quad (\text{en probabilité ou presque sûrement}) \quad (3.4)$$

tant que la suite $\{H_k\}$ satisfait à la condition de Carleman.

Dans Silverstein [46], le Théorème 4 a été établi, mais, avec en plus la condition : T_n définie non négative.

La preuve du Théorème qui suit est donnée dans Silverstein and Bai [47]

Théorème 5. *Supposons que les entrées de X_n ($p \times n$) sont des variables aléatoires complexes indépendantes pour chaque n et identiquement distribuées pour tout n et satisfaisant $\mathbb{E}(|x_{11} - \mathbb{E}(x_{11})|^2) = 1$. Supposons aussi que $T_n = \text{diag}(\tau_1, \dots, \tau_p)$, τ_i réel, et la fonction de distribution empirique de $\{\tau_1, \dots, \tau_p\}$ converge presque sûrement vers une fonction de distribution de probabilité H lorsque $n \rightarrow \infty$. Les entrées de X_n et T_n devraient dépendre de n , mais on supprime cette hypothèse pour simplicité. Soit $B_n = A_n + \frac{1}{n} X_n T_n X_n^*$, où A_n est hermitienne, de taille $n \times n$ satisfaisant $F^{A_n} \rightarrow F^A$ presque sûrement.*

Supposons aussi que X_n, T_n et A_n sont indépendants. Lorsque $p = p(n)$ avec $p/n \rightarrow y > 0$ pour $n \rightarrow \infty$, alors, presque sûrement, F^{B_n} l'ESD des valeurs propres de B_n converge lentement, lorsque $n \rightarrow \infty$, vers une d.f. F (non aléatoire), où pour tout $z \in \mathbb{C}^+ \equiv \{z \in \mathbb{C} : \Im z > 0\}$, sa transformée de Stieltjes $s = s(z)$ est l'unique solution dans \mathbb{C}^+ de l'équation

$$s = s_A \left(z - y \int \frac{\tau dH(\tau)}{1 + \tau s} \right), \quad (3.5)$$

où s_A est la transformée de Stieltjes de F^A .

Remarque 2. Notons que le Théorème 5 est plus général que le résultat de Yin dans le sens où on n'a pas besoin de la convergence des moments de l'ESD de \mathbf{T}_n , aussi bien que de la netteté positive \mathbf{T}_n (on n'a pas besoin que \mathbf{T}_n soit définie positive). Notons aussi qu'il permet d'inclure une matrice de perturbation \mathbf{A}_n dans $\mathbf{X}_n^* \mathbf{T}_n \mathbf{X}_n$. Cependant, il est un peu plus restrictif que le résultat de Yin dans le sens où il requiert que la matrice \mathbf{T}_n soit diagonale. La convergence faible de (3.5) a été établie dans Marcenko-Pastur [29] sous des conditions d'existence du moment le plus élevé considéré dans le Théorème 5, mais avec une certaine dépendance entre les entrées de \mathbf{X}_n .

La preuve du Théorème 5 utilise la transformée de Stieltjes qui est définie en Annexe.

3.3.2.2 LSD de la F -matrice

Dans la pratique, la F -matrice est très utilisée. C'est pourquoi il est important de trouver la densité limite des valeurs propres pour de telles matrices. Pour déterminer la LSD de la F -matrice, nous utiliserons le Théorème 6 qui suit.

Théorème 6. Soit $F = S_{n_1} S_{n_2}^{-1}$, où $S_{n_i} (i = 1, 2)$ est une matrice de covariance de dimension p et des données de taille n_i de moyenne 0 et variance 1. Si S_{n_1} et S_{n_2} sont indépendants, $p/n_1 \rightarrow y \in (0, \infty)$ et $p/n_2 \rightarrow y' \in (0, 1)$. Alors la LSD $F_{y,y'}$ de F existe et a une densité de fonction donnée par

$$F'_{y,y'}(x) = \begin{cases} \frac{(1-y')\sqrt{(b-x)(x-a)}}{2\pi x(y+xy')}, & \text{si } a < x < b \\ 0, & \text{ailleurs,} \end{cases} \quad (3.6)$$

$$\text{où } a = \left(\frac{1 - \sqrt{y + y' - yy'}}{1 - y'} \right)^2 \text{ et } b = \left(\frac{1 + \sqrt{y + y' - yy'}}{1 - y'} \right)^2.$$

De plus, si $y > 1$, alors F a un point de masse $1 - \frac{1}{y}$ à l'origine.

3.3.2.3 Quelques exemples

Nous avons effectué des simulations sous MATLAB, pour vérifier que la densité de la F -matrice sus-citée colle avec l'histogramme des valeurs propres du :

- produit d'une matrice de Wishart et de l'inverse d'une matrice de Wishart,
- produit de deux matrices de Wishart indépendantes.

3.3.2.3.1 Produit d'une matrice de Wishart et de l'inverse d'une matrice de Wishart

Considérons deux matrices : une matrice de Wishart et l'inverse d'une matrice de Wishart. Grace à des simulations MATLAB, la figure 3.1 a été obtenue.

Sur la figure 3.1, la densité limite de la F -matrice colle très bien avec l'histogramme des valeurs propres du produit matriciel d'une matrice de Wishart et de l'inverse d'une matrice de Wishart

3.3.2.3.2 Produit de deux matrices de Wishart indépendantes

En considérant deux matrices de Wishart indépendantes, des simulations sous MATLAB ont permis d'obtenir la figure 3.2 qui suit :

On peut voir que, la densité de la F -matrice couvre bien l'histogramme des valeurs propres de ce produit matriciel.

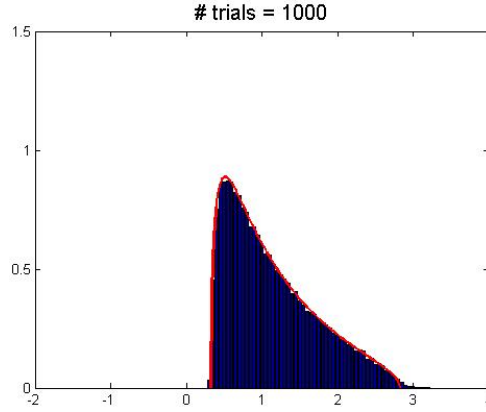


FIGURE 3.1 – Histogramme du produit d’une matrice de Wishart et de l’inverse d’une matrice de Wishart, et densité limite d’une F -matrice

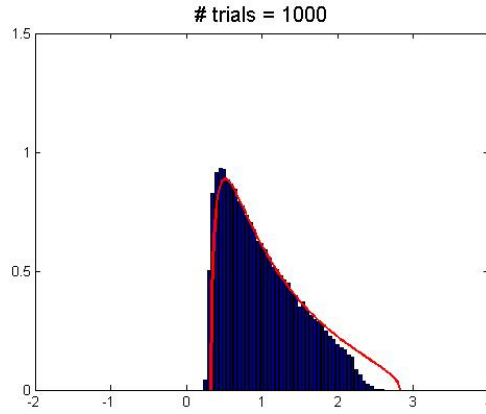


FIGURE 3.2 – Histogramme du produit de deux matrices de Wishart indépendantes, et densité limite d’une F -matrice

3.4 Distribution des rendements

3.4.1 Distribution univariée des rendements

Loin de la supposition des manuels de finance mathématique, les rendements de n’importe quel instrument financier (actions, changes, taux d’intérêt, matières premières, etc) sont loin d’être gaussiens. La distribution inconditionnelle des rendements a des queues lourdes, qui décroissent comme une loi puissance pour des paramètres assez grands. En fait, la fonction de distribution de probabilité empirique des rendements sur une courte période de temps (disons entre quelques minutes et quelques jours) peut raisonnablement être approchée par une distribution de Student

$$P(r) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{1+\mu}{2})}{\Gamma(\frac{\mu}{2})} \frac{a^\mu}{(r^2 + a^2)^{\frac{1+\mu}{2}}} \quad (3.7)$$

où a est un paramètre lié à la variance de la distribution par $\sigma^2 = \frac{a^2}{\mu-2}$, et μ est entre 3 et 5. Nous considérons ici et dans la suite que les rendements sont de moyenne nulle, ce qui est approprié pour des durées assez courtes.

cette distribution inconditionnelle peut cependant induire en erreur, puisque les rendements sont en effet très loin des variables aléatoires *i.i.d.* En d’autres termes, les rendements ne peuvent pas être pris comme des variables aléatoires indépendantes tirées d’une loi de Student. Un tel modèle prévoit que

sur une accumulation de temps, la distribution des rendements est la convolution de distributions de Student, qui converge vers une distribution gaussienne des rendements pour des durées beaucoup plus longues. Mais la volatilité des rendements financiers est elle même une variable dynamique, qui change au fil du temps. En des termes plus formels, le rendement au temps t peut être représenté comme étant le produit d'une composante de la volatilité σ^t et d'une composante directionnelle ξ^t

$$r^t = \sigma^t \xi^t, \quad (3.8)$$

où les ξ^t sont des variables aléatoires *i.i.d* de variance unitaire, et σ^t une variable aléatoire positive avec des composantes positives et négatives. C'est une question de goût de choisir ξ^t gaussien et de garder une haute fréquence, partie imprévisible de σ^t , ou de choisir ξ^t non-gaussien (par exemple distribué suivant une Student) en ne gardant que la fréquence basse, partie prévisible de σ^t .

Insistons sur le fait que dans l'équation (3.8), σ^t et ξ^t ne sont en réalité pas indépendants. La documentation sur les marchés des actions montre que les rendements négatifs du passé ont tendance à faire croître les volatilités futures et vice-versa.

3.4.2 Distribution multivariée des rendements

La distribution univariée des rendements ayant été définie dans la partie précédente, nous voulons étendre cette description à la distribution jointe des rendements de N actifs corrélés. Nous allons tout d'abord nous focaliser sur la distribution jointe des rendements *simultanés* $\{r_1^t, r_2^t, \dots, r_N^t\}$. Tous les marginaux de cette distribution jointe devraient ressembler à la distribution de Student (3.7) ci-dessus ; en outre, elle devrait être compatible avec la (vraie) matrice de corrélation des rendements :

$$C_{ij} = \int \prod_k [dr_k] r_i r_j P(r_1, r_2, \dots, r_N). \quad (3.9)$$

Il y a cependant une extension naturelle de la distribution univariée de Student, qui a une interprétation financière claire. Mais nous ne travaillerons qu'avec des rendements de la forme décrite dans l'équation (3.8) ci-dessus :

$$r_i^t = \sigma^t \xi_i^t,$$

où les ξ_i^t sont des variables aléatoires gaussiennes corrélées de matrice de corrélation \hat{C}_{ij} et la volatilité σ^t est commune à tous les actifs et distribuée suivant

$$P(\sigma) = \frac{2}{\Gamma(\frac{\nu}{2})} \exp \left[-\frac{\sigma_0^2}{\sigma^2} \right] \frac{\sigma_0^\nu}{\sigma^{1+\nu}}, \quad (3.10)$$

où $\sigma_0^2 = \frac{2\nu}{\nu-2}$. La distribution jointe des rendements est ainsi une Student multivariée de distribution P_S :

$$P_S(r_1, r_2, \dots, r_N) = \frac{\Gamma(\frac{N+\nu}{2})}{\Gamma(\frac{\nu}{2}) \sqrt{(\nu\pi)^N \det \hat{C}}} \frac{1}{\left(1 + \frac{\nu}{2} \sum_{ij} r_i (\hat{C}^{-1})_{ij} r_j \right)^{\frac{N+\nu}{2}}} \quad (3.11)$$

Citons quelques propriétés utiles de ce modèle :

- La distribution marginale de chaque r_i est une Student univariée de paramètre ν .
- Lorsque $\nu \rightarrow \infty$, on peut montrer que la distribution multivariée de Student converge vers une distribution gaussienne.
- La matrice de corrélation des r_i est donnée, pour $\nu > 2$, par :

$$C_{ij} = \langle r_i r_j \rangle = \frac{\nu}{\nu-2} \hat{C}_{ij}. \quad (3.12)$$

- Le théorème de Wick pour des variables gaussiennes peut être étendu aux variables de Student. Par exemple, on peut montrer que

$$\langle r_i r_j r_k r_l \rangle = \frac{\nu-2}{\nu-4} [C_{ij} C_{kl} + C_{ik} C_{jl} + C_{il} C_{jk}], \quad (3.13)$$

- La matrice \hat{C}_{ij} peut être estimée à partir des données empirique en utilisant la procédure du maximum de vraisemblance. Étant donnée une série de rendements d'actions r_i , la meilleure matrice \hat{C}_{ij} est donnée par la solution de l'équation ci-dessous :

$$\hat{C}_{ij} = \frac{N + \nu}{T} \sum_{t=1}^T \frac{r_i^t r_j^t}{\nu + \sum_{mn} r_m^t (\hat{C}^{-1})_{mn} r_n^t}. \quad (3.14)$$

Notons que dans la limite gaussienne $\nu \rightarrow \infty$ pour N fixé, le dénominateur de l'expression ci-dessus est simplement donné par ν , et l'expression finale est simplement

$$\hat{C}_{ij} = C_{ij} = \frac{1}{T} \sum_{t=1}^T r_i^t r_j^t, \quad (3.15)$$

Ce modèle de Student multivarié peut être considéré comme trop simple pour décrire des données financières, étant donné qu'il considère qu'il n'y a qu'un seul facteur de volatilité, commun à toutes les actions. On s'attend en réalité à ce que plusieurs facteurs de volatilité soient nécessaires. Cependant, la mise en oeuvre précise de cette idée et la forme résultante de la distribution multivariée n'ont pas encore été explorées en détails et restent un gros sujet de recherche.

Dans la prochaine partie, nous proposerons l'algorithme utilisé dans la pratique pour estimer les paramètres des lois complexes par une méthode de maximum de vraisemblance un peu complexe.

3.5 Algorithme EM (Espérance-Maximisation) [30]

L'algorithme EM, proposé par Dempster et al. (1977) [31], est une classe d'algorithmes, lesquels permettent de trouver le maximum de vraisemblance des paramètres de modèles probabilistes lorsque le modèle dépend des variables latentes non observables. Cet algorithme est souvent utilisé pour la classification de données, l'apprentissage automatique, ou encore l'estimation des modèles de Markov cachés (HMM's)

L'algorithme EM comporte :

- Une étape d'évaluation de l'espérance (E), où l'on calcule l'espérance de la vraisemblance en tenant compte des dernières variables observées,
- Une étape de maximisation (M), où l'on estime le maximum de vraisemblance des paramètres en maximisant la vraisemblance trouvée à l'étape (E).

On utilise ensuite les paramètres trouvés en M comme point de départ d'une nouvelle phase d'évaluation de l'espérance et l'on itère ainsi.

3.5.1 Principe de fonctionnement

Considérons un échantillon $X = (x_1, \dots, x_n)$ d'individus suivant une loi $f(x_i, \theta)$ paramétrée par θ . On cherche à déterminer le paramètre θ maximisant la log-vraisemblance (donc par une méthode de maximum de vraisemblance) donnée par

$$L(X; \theta) = \sum_{i=1}^n \log f(x_i, \theta). \quad (3.16)$$

Cet algorithme est utile lorsque la maximisation de L est complexe mais que, sous réserve de connaître certaines données judicieusement choisies, on peut simplement déterminer θ . Dans ce cas, on s'appuie sur des données complétées par un vecteur $Z = (z_1, \dots, z_n)$ inconnu. En notant $f(z_i|x_i; \theta)$ la probabilité de z_i sachant x_i et le paramètre θ , on peut définir log-vraisemblance complétée comme la quantité

$$L((X, Z); \theta) = \sum_{i=1}^n (\log f(z_i|x_i, \theta) + \log f(x_i; \theta)) \quad (3.17)$$

et donc,

$$L(X; \theta) = L((X, Z); \theta) - \sum_{i=1}^n \log f(z_i|x_i, \theta). \quad (3.18)$$

L'algorithme EM est une procédure itérative basée sur l'espérance des données complétées conditionnellement au paramètre courant. En notant $\theta^{(c)}$ ce paramètre, on peut écrire :

$$\mathbb{E} [L(X; \theta) | \theta^{(c)}] = \mathbb{E} [L((X, Z); \theta) | \theta^{(c)}] - \mathbb{E} \left[\sum_{i=1}^n \log f(z_i | x_i, \theta) | \theta^{(c)} \right], \quad (3.19)$$

ou encore

$$L(X; \theta) = Q(\theta; \theta^{(c)}) - H(\theta, \theta^{(c)}) \quad (3.20)$$

avec $Q(\theta; \theta^{(c)}) = \mathbb{E} [L((X, Z); \theta) | \theta^{(c)}]$ et $H(\theta, \theta^{(c)}) = \mathbb{E} [\sum_{i=1}^n \log f(z_i | x_i, \theta) | \theta^{(c)}]$. On montre que la suite définie par $\theta^{(c+1)} = \arg \max_{\theta} (Q(\theta, \theta^{(c)}))$ fait tendre $L(X; \theta^{(c+1)})$ vers un maximum local. On peut donc définir l'algorithme EM de la manière suivante :

Algorithm 3 Algorithme EM

Require: c, θ, X, Z

- 1: initialisation au hasard de $\theta^{(0)}$
 - 2: $c = 0$
 - 3: **while** L'algorithme n'a pas convergé, **do**
 - 4: Évaluation de l'espérance (étape E) : $Q(\theta; \theta^{(c)}) = \mathbb{E} [L((X, Z); \theta) | \theta^{(c)}]$
 - 5: Maximisation (étape M) : $\theta^{(c+1)} = \arg \max_{\theta} (Q(\theta, \theta^{(c)}))$
 - 6: $c = c+1$
 - 7: **end while**
-

En pratique, pour s'affranchir du caractère local du maximum atteint, on fait tourner l'algorithme EM un grand nombre de fois à partir de valeurs initiales différentes de manière à avoir de plus grandes chances d'atteindre le maximum global de vraisemblance.

Notons que dans le cas des modèles de mélange en particulier, les équations de vraisemblance $\frac{d}{d\theta} \log L(X; \theta) = 0$ n'ont pas de solution explicite. L'algorithme EM permet alors dans certains cas seulement, qui dépendent de la structure de la donnée, d'augmenter (3.16) de façon itérative et de converger vers l'un de ses maxima locaux.

3.5.2 Applications de l'algorithme EM à une loi de Student multivariée

Rappelons que l'algorithme EM est un algorithme très important en statistiques pour l'estimation des paramètres cachés dans une loi. Nous illustrerons les résultats obtenus à l'aide de cet algorithme. Notre cas d'étude est un peu complexe, et correspond au cas où les variables suivent une loi de Student multivariée. Pour résoudre le problème du deuxième cas, une bibliothèque MATLAB (*pmtk3*) a été développée, et nous l'utiliserons pour nos estimations.

Les méthodes classiques d'estimation n'étant pas applicables dans le cas des loi de Student Multivariées, nous utiliserons l'algorithme EM. Dans la littérature, il est dit que, la meilleure distribution qui approche les rendements est la distribution de la loi de Student multivariée, raison pour laquelle nous exposons cet exemple. Notre but ici est d'estimer les paramètres de la loi (μ, Σ, ν) , et de comparer ces paramètres estimés $(\hat{\mu}, \hat{\Sigma}, \hat{\nu})$ à ceux passés en entrées. On considèrera principalement deux cas :

- un cas où on passera en paramètre la matrice identité comme matrice de corrélation (corrélations nulles)
- un cas où on rajoutera des corrélations non nulles

3.5.2.1 Cas où la matrice de corrélation est la matrice identité

Ici, les paramètres passés en entrée sont les suivants :

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \nu = 5, \quad n = 10000 \quad (3.21)$$

Pour générer des variables qui suivent une loi de Student multivariée centrée, nous avons utilisé la commande MATLAB *mvtrnd*. En utilisant la toolbox MATLAB *pmtk3*, nous avons estimé les paramètres des variables générées, et nous avons obtenu les résultats qui suivent :

$$\hat{C} = \hat{\Sigma} = \begin{pmatrix} 1.0444 & -0.0041 \\ -0.0041 & 1.0116 \end{pmatrix}, \quad \hat{\nu} = 4.9878, \quad \hat{\mu} = \begin{pmatrix} 0.0157 \\ 0.0023 \end{pmatrix} \quad (3.22)$$

Comme on peut le remarquer, les estimations observées sont assez voire très proche des paramètres passés en entrée. En effet, on peut voir dans l'équation (3.22), que le paramètre $\hat{\nu}$ obtenu après estimation est très proche de celui passé en entrée (5). Quant à la matrice de variance covariance (notons qu'elle est la même que la matrice de corrélation dans ce cas), on note que l'estimation est aussi assez bonne. On est très proche de la matrice identité.

Comme nous l'avons dit un peu plus haut, nous présenterons un cas où la matrice de corrélation passée en paramètre est différente de la matrice identité.

3.5.2.2 Cas où la matrice de corrélation est différente de la matrice identité

On va considérer dans cette partie une matrice de corrélation dans laquelle les coefficients de corrélation sont différents de 0. Dans ce cas, les paramètres d'entrées considérés sont les suivants :

$$C = \begin{pmatrix} 1 & 0.4 \\ 0.4 & 1 \end{pmatrix}, \quad \nu = 5, \quad n = 10000 \quad (3.23)$$

une estimation des paramètres des variables générées, nous a donné les résultats qui suivent :

$$\hat{\Sigma} = \hat{C} = \begin{pmatrix} 0.9919 & 0.4111 \\ 0.4111 & 1.0350 \end{pmatrix}, \quad \hat{\nu} = 5.0190, \quad \hat{\mu} = \begin{pmatrix} -0.0145 \\ -0.0061 \end{pmatrix} \quad (3.24)$$

On remarque que le constat est le même que dans le cas précédent. On peut voir dans l'équation (3.24) que le paramètre $\hat{\nu}$ est très proche de celui qu'on a passé en entrée. La matrice de variance covariance obtenue est également bien estimée par l'algorithme EM.

Au vu des résultats obtenus dans les deux exemples sus-cités, nous allons nous appuyer sur l'algorithme EM pour estimer les paramètres des lois, étant donné que les paramètres estimés des lois utilisées, sont assez proches des paramètres passés en entrée. Dans la suite, nous estimerons les paramètres de la distribution des rendements décrite dans la partie 3.4.2 à l'aide de l'algorithme EM, et nous analyserons le comportement des résultats.

3.6 Théorie des Matrices Aléatoires et application de la transformée de Stieltjes

3.6.1 Préliminaires

La RMT se propose de faire de la statistique des valeurs propres λ_α des matrices aléatoires de grande dimension. En particulier, la densité des valeurs propres $\rho(\lambda)$ est définie comme suit :

$$\rho_N(\lambda) = \frac{1}{N} \sum_{\alpha=1}^N \delta(\lambda - \lambda_\alpha), \quad (3.25)$$

où λ_α sont les valeurs propres de la matrice symétrique \mathbf{H} . Il est important d'introduire la résolvante $G_H(z)$ (qui a été défini en Annexe B.2.2) de \mathbf{H} (encore appelé transformée de Stieltjes), où z est un nombre complexe. De cette transformée de Stieltjes, on peut extraire le spectre par la formule

$$\rho_N(\lambda) = \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \Im(G_H(\lambda - i\epsilon)). \quad (3.26)$$

Dans la limite lorsque $N \rightarrow \infty$, il arrive souvent (mais pas toujours) que la densité des valeurs propres ρ_N converge presque sûrement vers une unique densité $\rho_\infty(\lambda)$ bien définie. Ceci signifie que la densité ρ_N devient indépendante de la réalisation spécifique de la matrice \mathbf{H} , pourvu que \mathbf{H} soit un échantillon typique dans son ensemble. Cette propriété, appelée "ergodicité" ou "auto-pondération", est extrêmement

importante pour les applications pratiques dans le sens où le résultat asymptotique $\rho_\infty(\lambda)$ peut être utilisé pour décrire la densité des valeurs propres d'un cas *simple*.

Plusieurs autres transformations, au delà de la résolvante $G(z)$, se révèlent être utiles pour notre travail. Une de ses transformations est la "Blue fonction" $B(z)$, qui est l'inverse de $G(z)$, i.e. : $B[G(z)] = G[B(z)] = z$. La R -transformation est simplement liée à la Bleue fonction par :

$$R(z) = B(z) - \frac{1}{z}. \quad (3.27)$$

Il est facile de vérifier que $R(z)$ obéit à la propriété ci-dessous :

$$R_{aH}(z) = aR_H(az) \quad (3.28)$$

où a est un nombre réel arbitraire. En outre, $R(z)$ peut être étendu pour des z assez grands comme

$$R(z) = \sum_{k=1}^{\infty} c_k z^{k-1} \quad (3.29)$$

où les coefficients c_k peuvent être considérés comme cumulants. Par exemple, $c_1 = \int d\lambda \lambda \rho(\lambda)$. Lorsque $c_1 = 0$, $c_2 = \int d\lambda \lambda^2 \rho(\lambda)$

Le dernier objet dont nous aurons besoin est plus encombrant. C'est la S -transformation qui se définit comme suit [58] :

$$S(z) = -\frac{1+z}{z} \eta^{-1}(1+z) \quad \eta(y) \equiv -\frac{1}{y} G\left(\frac{1}{y}\right). \quad (3.30)$$

Dans la suite, nous passerons en revue plusieurs résultats de RMT sur la densité $\rho_\infty(\lambda)$ qui peut être obtenue en utilisant un concept étonnamment efficace : la liberté des matrices [59].

3.6.2 Matrices libres

La "liberté" est la généralisation sur les matrices, du concept de variables aléatoires. Nous pouvons dire que, deux matrices \mathbf{A} et \mathbf{B} sont mutuellement libres si leurs bases propres sont liées l'une à l'autre par une rotation aléatoire, ou plus simplement, si les vecteurs propres de \mathbf{A} et \mathbf{B} sont presque sûrement orthogonaux. Une définition plus précise et complète peut être trouvée par exemple dans [69], mais la définition simplifiée donnée ici, et les exemples suivants, seront suffisants pour notre travail.

Exemple 3. *Considérons deux matrices \mathbf{A} et \mathbf{B} , et choisissons une certaine matrice de rotation \mathbf{O} dans le groupe orthogonal $O(N)$, uniformément à la mesure de Haar. Alors \mathbf{A} et $\mathbf{O}'\mathbf{B}\mathbf{O}$ sont mutuellement libres.*

Exemple 4. *Cet exemple est plus intéressant et pas très compréhensible. Considérons deux matrices \mathbf{H}_1 et \mathbf{H}_2 choisies indépendamment dans l'ensemble GOE , i.e. l'ensemble des matrices symétriques tel que toutes les entrées soient des variables gaussiennes i.i.d. Étant donné que la mesure de cet ensemble de matrices aléatoires est invariable par transformation orthogonale, la matrice de rotation \mathbf{O}_1 diagonalisant \mathbf{H}_1 est une matrice de rotation aléatoire sur $O(N)$ (ce raisonnement est en réalité une méthode numérique commode pour générer des matrices de rotations aléatoires). La rotation $\mathbf{O}_1'\mathbf{O}_2$ de la base propre de \mathbf{H}_1 à celle de \mathbf{H}_2 est donc aussi aléatoire et \mathbf{H}_1 et \mathbf{H}_2 sont mutuellement libres.*

Maintenant, la liberté des matrices nous permettra de donner des résultats permettant de déterminer le spectre de la somme de matrices, connaissant le spectre de chacune des matrices, pourvu qu'elles soient mutuellement libres. Plus précisément, si $R_A(z)$ et $R_B(z)$ sont les R -transformations de deux matrices libres \mathbf{A} et \mathbf{B} , alors :

$$R_{A+B}(z) = R_A(z) + R_B(z). \quad (3.31)$$

Ce résultat généralise clairement la règle de convolution pour la somme de deux variables aléatoires indépendantes, pour lesquelles le logarithme de la fonction caractéristique est additif. Une fois que $R_{A+B}(z)$ est connu, on peut en principe inverser la R -transformation pour obtenir la densité des valeurs propres de $\mathbf{A} + \mathbf{B}$.

Il existe un résultat analogue pour le produit des matrices aléatoires non négatives. Dans ce cas, la S -transformation est multiplicative :

$$S_{A+B}(z) = S_A(z)S_B(z). \quad (3.32)$$

Dans la suite, nous montrerons comment ces règles permettent d'obtenir très facilement des densités de valeurs propres bien connues pour de grandes matrices, et nous présenterons aussi quelques résultats plus récents.

3.6.3 Application : Wigner et Marcenko-Pastur

Commençons par la célèbre loi du demi-cercle de Wigner pour les matrices Gaussiennes Orthogonales. Comme indiqué ci-dessus, deux telles matrices \mathbf{H}_1 et \mathbf{H}_2 sont mutuellement libres. En outre, grâce à la stabilité des variables Gaussiennes par addition, la matrice $(\mathbf{H}_1 + \mathbf{H}_2)/\sqrt{2}$ est dans le même ensemble. On a alors :

$$R_{\sqrt{2}H}(z) = R_{H_1+H_2}(z) = R_{H_1}(z) + R_{H_2}(z) = 2R_H(z). \quad (3.33)$$

Utilisant le résultat de l'équation (3.28) ci-dessus avec $a = \sqrt{2}$, on trouve que $R(z)$ satisfait à

$$2R_H(z) = \sqrt{2}R_H(\sqrt{2}z) \rightarrow R_H(z) = z. \quad (3.34)$$

où nous avons considéré la normalisation standard $\text{tr}(\mathbf{H}^2) = 1$. On peut facilement vérifier que $R(z) = z$ est en effet la R -transformation de demi-cercle de Wigner. Il existe un autre Théorème Central Limite permettant d'obtenir ce résultat.

Supposons que $\mathbf{H}_i, i = 1, \dots, \mathcal{N}$ sont des matrices aléatoires de "petites" traces, tel que chaque élément ait une variance égale à ϵ^2 avec $\epsilon \rightarrow 0$.

$$G(z) = \frac{1}{z} + 0 + \epsilon^2 \frac{1}{z^3} + O\left(\frac{\epsilon^3}{z^4}\right) \rightarrow \frac{1}{z} \approx G - \epsilon^2 G^3. \quad (3.35)$$

De là,

$$B(z) \approx \frac{1}{z - \epsilon^2 z^3} \rightarrow R(z) = B(z) - \frac{1}{z} \approx \epsilon^2 z + O(\epsilon^3 z^2) \quad (3.36)$$

Maintenant, si ces \mathcal{N} matrices sont mutuellement libres, avec $\epsilon = \mathcal{N}^{-1/2}$ et $\mathcal{N} \rightarrow \infty$, alors la R -transformation de la somme de telles matrices est :

$$R(z) = \mathcal{N}\epsilon^2 z + O(\mathcal{N}\epsilon^3 z^2) \rightarrow_{\mathcal{N} \rightarrow \infty} z. \quad (3.37)$$

Donc la somme de \mathcal{N} "petites" matrices centrées a un spectre de Wigner pour \mathcal{N} assez large, avec quelques corrections lors des calculs.

Dans l'exemple suivant, on va considérer la corrélation empirique dans le cas où la vraie matrice de corrélation est l'identité : $\mathbf{C} = \mathbf{I}$. Alors \mathbf{E} est par définition la somme de matrices de rang 1 $\delta E_{ij}^t = (r_i^t r_j^t)/T$, où les r_i^t sont des variables aléatoires indépendantes de variance unitaire. De là, $\delta \mathbf{E}^t$ a une valeur propre égale à q (pour N assez grand) associée à la direction \mathbf{r}^t , et $N - 1$ valeurs propres nulles correspondant à l'hyperplan perpendiculaire à \mathbf{r}^t . Les différents $\delta \mathbf{E}^t$ sont donc mutuellement libres et on peut utiliser l'astuce de la R -transformation. Sachant que

$$\delta G^t(z) = \frac{1}{N} \left(\frac{1}{z - q} + \frac{N - 1}{z} \right) \quad (3.38)$$

Inversant $\delta G(z)$, la "Blue" transformation élémentaire donne :

$$\delta B(z) = \frac{1}{z} + \frac{q}{N(1 - qz)} \rightarrow \delta R(z) = \frac{q}{N(1 - qz)}. \quad (3.39)$$

Utilisant l'addition des R -transformations, on déduit :

$$B_E(z) = \frac{1}{z} + \frac{q}{1 - qz} \rightarrow G_E(z) = \frac{(z + q - 1) - \sqrt{(z + q - 1)^2 - 4zq}}{2zq}, \quad (3.40)$$

qui reproduit le très connu résultat de Marcenko-Pastur pour la densité des valeurs propres (pour $q < 1$) [26] :

$$\rho_E(\lambda) = \frac{\sqrt{4\lambda q - (\lambda + q - 1)^2}}{2\pi\lambda q}, \quad \lambda \in [(1 - \sqrt{q})^2, (1 + \sqrt{q})^2]. \quad (3.41)$$

La caractéristique remarquable de ce résultat est qu'il ne devrait avoir aucune valeur propre hors de l'intervalle $[(1 - \sqrt{2})^2, (1 + \sqrt{2})^2]$ lorsque $N \rightarrow \infty$. On pourrait vérifier que $\rho_E(\lambda)$ converge vers $\delta(\lambda - 1)$ lorsque $q = 1/Q \rightarrow 0$, ou $T \gg N$. Rappelons que la fonction δ est définie telle que :

$$\delta(\lambda) = \begin{cases} 1 & \text{en } \lambda \\ 0 & \text{ailleurs} \end{cases} \quad (3.42)$$

Lorsque $q > 1$, certaines valeurs propres nulles apparaissent dans le spectre, qui devient alors :

$$\rho_E(\lambda) = (1 - Q)\delta(\lambda) + \frac{\sqrt{4\lambda Q - (\lambda + Q - 1)^2}}{2\pi\lambda}. \quad (3.43)$$

Utilisant $G_E(z)$, on peut directement montrer que $(1/N)\text{tr}(\mathbf{E}^{-1}) = G_E(0)$ est donné par $(1 - q)^{-1}$, pour $q < 1$. Le résultat de Marcenko-Pastur est important à cause de son degré d'universalité : comme le résultat du demi-cercle de Wigner, ce résultat tient lorsque les variables r_i^t sont i.i.d., avec un moment d'ordre deux fini.

3.6.4 Plus d'applications

3.6.4.1 Cas d'une matrice de corrélation aléatoire non triviale

En général, les variables aléatoires sont décrites par une matrice de corrélation C non triviale, différente de la matrice identité I . Le résultat de Marcenko-Pastur pour le spectre d'une matrice empirique \mathbf{E} peut être étendu à une matrice générale \mathbf{C} , et ouvrir la voie pour caractériser le spectre ρ_C même avec de l'information partielle $Q = T/N < \infty$. Considérant que les rendements r_i^t sont gaussiens, la matrice empirique E peut se décomposer sous la forme

$$\mathbf{E} = \sqrt{C}\hat{X}[\sqrt{C}\hat{X}]^T \quad (3.44)$$

où \hat{X} est une matrice rectangulaire $N \times T$, de v.a. gaussiennes décorréliées, de variance unitaire. Mais sachant que les valeurs propres de la matrice $\sqrt{C}\hat{X}[\sqrt{C}\hat{X}]^T$ sont les mêmes que celles de la matrice $C\hat{X}\hat{X}^T$, nous pouvons utiliser la S -transformation, avec $A = C$ et $B = \hat{X}\hat{X}^T$ qui sont mutuellement libres, et où le spectre de B est par construction, donné par la loi de Marcenko-Pastur. Ceci nous permet de donner une expression pour la résolvante de \mathbf{E} :

$$G_E(z) = \int d\lambda \rho_C(\lambda) \frac{1}{z - \lambda(1 - q + qzG_E(z))}. \quad (3.45)$$

On peut vérifier que si $\rho_C(\lambda) = \delta(\lambda - 1)$, on retrouve le résultat connu de M-P, qui est

$$G_E(z) = \frac{(z + q - 1) - \sqrt{(z + q - 1)^2 - 4zq}}{2zq}. \quad (3.46)$$

L'équation (3.45) ci-dessus peut se ré-écrire sous la forme

$$zG_E(z) = ZG_C(Z), \quad \text{avec} \quad Z = \frac{z}{1 + q(zG_E(z) - 1)}, \quad (3.47)$$

expression facile à utiliser pour une évaluation numérique [56]. De ces équations, on peut évaluer $G_E(0)$ qui vaut

$$G_E(0) = \text{tr}\mathbf{E}^{-1} = \frac{\text{tr}\mathbf{C}^{-1}}{1 - q}. \quad (3.48)$$

Pour reconstruire le spectre de C de celui de E , il faudrait utiliser une approche paramétrique, et non pas chercher à inverser la relation ci-dessus. Il est facile d'inverser lorsque la dimension est inférieure ou égale à 3, mais au delà, c'est impossible (voir encore [56]).

Notons que, dans le cas où la matrice \mathbf{C} a des valeurs propres isolées (par exemple trop grandes...), la méthode ci-dessus ne peut pas être appliquée. La méthode proposée ne décrit que les parties continues du spectre. Il a par exemple été démontré (dans [57]) que, pour une matrice \mathbf{C} ayant une grande valeur propre qui n'est pas dans le spectre de Wishart, les statistiques de cette valeur propre suivent une gaussienne de largeur $\sim T^{-1/2}$.

3.6.4.2 Cas de l'ensemble de Student

Supposons maintenant que les rendements r_i^t suivent une loi de Student multivariée définie dans la partie 3.4.2. Étant donné que dans ce cas $r_i^t = \sigma_t \xi_i^t$, la matrice de corrélation empirique peut s'écrire comme

$$E_{ij} = \frac{1}{T} \sum_t \sigma_t^2 \xi_i^t \xi_j^t, \quad \langle \xi_i \xi_j \rangle \equiv \widehat{C}_{ij}. \quad (3.49)$$

Dans le cas où $\widehat{\mathbf{C}} = \mathbf{1}$, ceci peut être vu comme une somme de projecteurs mutuellement libres, et on peut utiliser l'astuce de la R-transformation. Ceci permet d'obtenir les deux équations ci-dessous pour la résolvante de \mathbf{E} , obtenus pour la première fois par Marcenko-Pastur, pour N et T tendant vers l'infini :

$$\lambda = \frac{G_R}{G_R^2 + \pi^2 \rho_E^2} + \int ds P(s) \frac{\mu(s - q\mu G_R)}{(s - q\mu G_R)^2 + \pi^2 \rho_E^2} \quad (3.50)$$

$$0 = \rho \left(-\frac{1}{G_R^2 + \pi^2 \rho_E^2} + \int ds P(s) \frac{q\mu^2}{(s - q\mu G_R)^2 + \pi^2 \rho_E^2} \right) \quad (3.51)$$

où G_R est la partie réelle de la résolvante, et $P(s) = s^{\mu/2-1} e^{-s} \Gamma(\mu/2)$ est la distribution de $s = \mu/\sigma^2$ dans le cas d'une distribution de Student ; Cependant, le résultat ci-dessus tient également pour d'autres distributions de σ , correspondant à la classe des distributions multivariées "elliptiques". En effet, utilisant le résultat selon lequel toute matrice $\sigma_t^2 \eta_i^t \eta_j^t$ est un projecteur, sa résolvante s'écrit simplement

$$\begin{aligned} G_t(z) &= \frac{1}{N} \left(\frac{1}{z - \frac{\bar{\mu}}{Q_{s_t}}} + \frac{N-1}{z} \right) \\ &= \frac{1}{N} \left(\frac{\frac{\bar{\mu}}{Q_{s_t}}}{z(z - \frac{\bar{\mu}}{Q_{s_t}})} \right) + \frac{1}{z} \end{aligned} \quad (3.52)$$

où $Q = \frac{T}{N}$ et $\sigma_t^2 = \frac{\bar{\mu}}{s_t}$ (avec $\bar{\mu} = \frac{\mu}{2} - 1$), tel que $P(s) = s^{\frac{\mu}{2}-1} \frac{e^{-s}}{\Gamma(\frac{\mu}{2})}$ dans le cas d'une Student. Avant de démontrer le résultat ci-dessus, il est nécessaire d'avoir quelques connaissances sur les projecteurs :

Proposition 1. *Si P est une matrice $\mathbb{C}^{m \times m}$ / $P^2 = P$, alors P est un projecteur. Et dans ce cas, $\text{tr}(P) = \text{rank}(P)$.*

Proposition 2. *Si P est un projecteur, alors $I - P$ est aussi un projecteur.*

Preuve. En effet :

$$\begin{aligned} (I - P)^2 &= (I - P)(I - P) \\ &= I^2 - IP - PI + P^2 \\ &= I - P - P + P \\ &= I - P. \end{aligned}$$

□

Proposition 3. *Pour P projecteur, on a le résultat général suivant :*

$$(zI - P)^{-1} = \frac{1}{z} I + \frac{1}{z(z-1)} P \quad (3.53)$$

Proposition 4. *Pour deux matrices A et B et un scalaire α , on a la formule suivante :*

$$(A + \alpha B^*)^{-1} = A^{-1} - \frac{A^{-1} \alpha B^* A^{-1}}{1 + B^* A^{-1} \alpha}. \quad (3.54)$$

Preuve. Preuve de la Proposition 3

Ce résultat se retrouve facilement en utilisant la formule donnée dans la Proposition 4. En effet, en

posant $A = zI$ et $\alpha B^* = -P$, on obtient :

$$\begin{aligned}
(A + \alpha B^*)^{-1} &= (zI - P)^{-1} = \frac{1}{z}I - \frac{(zI)^{-1}(-P)(zI)^{-1}}{1 + (-P)(zI)^{-1}} \\
&= \frac{1}{z}I + \frac{\frac{1}{z}I^{-1}P\frac{1}{z}I^{-1}}{1 - P\frac{1}{z}I^{-1}} \\
&= \frac{1}{z}I + \frac{\frac{1}{z}IP\frac{1}{z}I}{1 - P\frac{1}{z}I} \\
&= \frac{1}{z}I + \frac{\frac{1}{z^2}P}{1 - \frac{1}{z}P} \\
&= \frac{1}{z}I + \frac{1}{z^2}P \times \frac{z}{(zI - P)} \\
&= \frac{1}{z}I + \frac{1}{z} \frac{P}{P^{-2}(zP^2 - P)} \\
&= \frac{1}{z}I + \frac{1}{z} \frac{P}{P^{-1}(zP - P)} \text{ car } P^2 = P \\
&= \frac{1}{z}I + \frac{1}{z} \frac{P}{P^{-1}P(z - 1)} \\
&= \frac{1}{z}I + \frac{1}{z(z - 1)}P
\end{aligned}$$

□

Utilisant ce résultat et le fait que $\text{tr}(\alpha A + \beta B) = \alpha \text{tr}(A) + \beta \text{tr}(B)$, on a :

$$\text{tr}(zI - P)^{-1} = \frac{1}{z}\text{tr}(I) + \frac{1}{z(z - 1)}\text{tr}(P). \quad (3.55)$$

$G_t(z)$ étant posé dans l'équation 3.52, il est maintenant question de déterminer la fonction bleue $B_t(z)$, pour ensuite poser le système qui nous permettra de déterminer la densité (ou plus précisément, de retrouver les équations 3.50 et 3.51).

Posons dans l'équation 3.52, $h_t = \frac{\bar{\mu}}{Qs_t}$. On a :

$$\begin{aligned}
G_t(B_t(z)) = z &\iff \frac{1}{N} \left(\frac{1}{B_t(z) - h_t} + \frac{N - 1}{B_t(z)} \right) = z \\
&\iff \frac{1}{B_t(z) - h_t} + \frac{N - 1}{B_t(z)} = Nz \\
&\iff \frac{B_t(z) + NB_t(z) - Nh_t - B_t(z) + h_t}{B_t(z)(B_t(z) - h_t)} = Nz \\
&\iff NB_t(z) - Nh_t + h_t = NzB_t(z)(B_t(z) - h_t) \\
&\iff N(B_t(z) - h_t) + h_t = NzB_t(z)(B_t(z) - h_t) \\
&\iff B_t(z) = \frac{1}{z} + \frac{h_t}{N(zB_t(z) - zh_t)}.
\end{aligned}$$

$B_t(x)$ trouvé, nous pouvons maintenant, utilisant les propriétés de la R -transformation, trouver la fonction bleue. On a :

$$B(x) = \frac{1}{x} + \frac{1}{T} \sum_t \frac{\frac{\bar{\mu}}{s_t}}{\left(1 - \frac{x\bar{\mu}}{Qs_t}\right)} = \frac{1}{x} + \int ds P(s) \frac{\frac{\bar{\mu}}{s}}{\left(1 - \frac{x\bar{\mu}}{Qs}\right)} \quad (3.56)$$

Cette égalité est due au fait qu'on considère (T et N assez grand) la limite lorsque T et N tendent vers

l'infini, pour Q fixé. Étant donné que $B(G(z)) = z$, que $z = \lambda + i\epsilon$ et que $G(z) = G_R(\lambda) + i\pi\rho(\lambda)$, on a :

$$\begin{aligned}
B(G(z)) &= \frac{1}{G(z)} + \int ds P(s) \frac{\frac{\bar{\mu}}{s}}{\left(1 - \frac{G(z)\bar{\mu}}{Qs}\right)} \\
z &= \frac{1}{G(z)} + \int ds P(s) \frac{\frac{\bar{\mu}}{s}}{\left(1 - \frac{G(z)\bar{\mu}}{Qs}\right)} \\
&= \frac{1}{G(z)} + \int ds P(s) \frac{\bar{\mu}}{\left(s - \frac{G(z)\bar{\mu}}{Q}\right)} \\
\lambda + i\epsilon &= \frac{1}{G_R + i\pi\rho} + \int ds P(s) \frac{\bar{\mu}}{\left(s - \frac{(G_R + i\pi\rho)\bar{\mu}}{Q}\right)} \\
&= \frac{G_R - i\pi\rho}{G_R^2 + \pi^2\rho^2} + \int ds P(s) \frac{\bar{\mu}}{\left(s - \frac{\bar{\mu}G_R}{Q} - i\frac{\pi\rho\bar{\mu}}{Q}\right)} \\
&= \frac{G_R}{G_R^2 + \pi^2\rho^2} - i\frac{\pi\rho}{G_R^2 + \pi^2\rho^2} + \int ds P(s) \frac{\bar{\mu} \left(s - \frac{\bar{\mu}G_R}{Q} + i\frac{\pi\rho\bar{\mu}}{Q}\right)}{\left(s - \frac{\bar{\mu}G_R}{Q}\right)^2 + \frac{\bar{\mu}^2\pi^2\rho^2}{Q^2}} \\
\lambda + i\epsilon &= \frac{G_R}{G_R^2 + \pi^2\rho^2} + \int ds P(s) \frac{\bar{\mu} \left(s - \frac{\bar{\mu}G_R}{Q}\right)}{\left(s - \frac{\bar{\mu}G_R}{Q}\right)^2 + \frac{\bar{\mu}^2\pi^2\rho^2}{Q^2}} \\
&\quad + i\pi\rho \left[-\frac{1}{G_R^2 + \pi^2\rho^2} + \int ds P(s) \frac{\frac{\bar{\mu}^2}{Q}}{\left(s - \frac{\bar{\mu}G_R}{Q}\right)^2 + \frac{\bar{\mu}^2\pi^2\rho^2}{Q^2}} \right]
\end{aligned}$$

En identifiant les parties réelle et imaginaire, on obtient le système d'équations qui suit :

$$\begin{cases}
\lambda &= \frac{G_R}{G_R^2 + \pi^2\rho^2} + \int ds P(s) \frac{\bar{\mu} \left(s - \frac{\bar{\mu}G_R}{Q}\right)}{\left(s - \frac{\bar{\mu}G_R}{Q}\right)^2 + \frac{\bar{\mu}^2\pi^2\rho^2}{Q^2}} \\
\epsilon &= \pi\rho \left[-\frac{1}{G_R^2 + \pi^2\rho^2} + \int ds P(s) \frac{\frac{\bar{\mu}^2}{Q}}{\left(s - \frac{\bar{\mu}G_R}{Q}\right)^2 + \frac{\bar{\mu}^2\pi^2\rho^2}{Q^2}} \right]
\end{cases} \quad (3.57)$$

Faisant tendre ϵ vers 0, on obtient bien les équations voulues. Les résultats importants sont les suivants :

1. Il n'y a plus de borne supérieure du spectre : $\rho_E(\lambda) \sim \lambda^{-1-\mu/2}$ lorsque $\lambda \rightarrow \infty$.
2. Mais il existe une borne inférieure du spectre pour tous les μ .

Le cas $\hat{C} \neq 1$ peut également être traité en utilisant la S-transformation.

Au lieu d'utiliser l'estimateur habituel (Pearson) de la matrice de corrélation, on pourrait utiliser la procédure du maximum de vraisemblance, décrite à l'équation (3.14). Étonnamment, le spectre $\rho_{ML}(\lambda)$ correspondant est complètement différent [61], et est donné par le résultat standard de Marcenko-Pastur ! La raison intuitive est que l'estimateur du maximum de vraisemblance donné à l'équation (3.14) renormalise effectivement les rendements par la volatilité journalière σ_t lorsque σ_t est assez grand. Donc, toutes les anomalies provoquées par des grosses journées (i.e. $\sigma_t \gg \sigma_0$) disparaissent.

Enfin, mentionnons que dans la littérature récente, on a considéré un autre ensemble de matrices aléatoires dont les éléments sont distribués suivant une Student, où au lieu d'avoir une volatilité σ_t dépendant du temps, c'est la volatilité globale qui est aléatoire et distribuée selon (3.10).

Toute la théorie présentée, nous allons maintenant proposer une application financière du modèle proposé dans la partie 3.4.2. Dans cette application, nous utiliserons l'algorithme EM pour estimer les paramètres de la distribution des rendements ($\hat{\Sigma}$, $\hat{\nu}$ et $\hat{\mu}$) et certains résultats de J.-P. Bouchaud et Marc Potters pour trouver la meilleure densité qui approche l'histogramme des valeurs propres de la matrice de corrélation. Pour terminer, nous comparerons la densité obtenue à celle de MP pour voir quelles sont les différences.

3.7 Méthode polynomiale pour matrices aléatoires

3.7.1 Introduction

N. Raj Rao et Alan Edelman [35] ont défini une classe de matrices aléatoires algébriques. Ce sont des matrices aléatoires pour lesquelles la transformée de Stieltjes de la distribution limite des valeurs propres est algébrique, i.e. qu'elle satisfait à une équation polynomiale (à deux variables). Les matrices de Wigner et de Wishart dont les distributions limites des valeurs propres sont respectivement données par la loi du demi cercle et la loi de Marčenko-Pastur sont des cas spéciaux.

Dans l'article [35], les auteurs ont développé une méthode polynomiale qui permet de décrire une classe de matrices algébriques. Cette méthode est tout d'abord basée sur une approche analytique, laquelle donnera un polynôme qui sera ensuite utilisé numériquement (en utilisant une toolbox développée par les deux auteurs, qui s'appelle RMTTool random matrix calculator) pour obtenir la densité limite des valeurs propres.

3.7.2 Exemples de la somme et du produit d'une Wigner et d'une Wishart

N. Raj Rao et Alan Edelman [35] ont proposé une méthode forte et assez pratique permettant de calculer la distribution limite des valeurs propres d'une large classe de matrices aléatoires. Cette méthode permet tout d'abord de se pencher sur des cas de matrices aléatoires pour lesquelles on a des densités célèbres (tel la densité du demi-cercle, la densité de M-P, de McKay, et bien d'autres...), et ensuite d'en déduire d'autres densités... En particulier, nous encoderons la distribution limite des valeurs propres comme solutions de polynômes à deux variables. Les opérations canoniques sur les matrices aléatoires deviennent donc les opérations sur les polynômes de 2 variables.

Exemple 5. Nous allons illustrer les propos ci-dessus avec l'exemple qui suit :
Considérons une matrice de Wigner, qui sous MATLAB s'écrit

$$G = \text{sign}(\text{randn}(N))/\text{sqrt}(N);$$

$$A = (G + G')/\text{sqrt}(2);$$

dont les valeurs propres convergent vers la loi du demi-cercle lorsque $N \rightarrow \infty$, et la matrice de Wishart, qui s'écrit sous MATLAB

$$G = \text{randn}(N, 2 * N)/\text{sqrt}(2 * N);$$

$$A = G * G';$$

dont les valeurs propres convergent vers la loi de M-P. Les distributions de valeurs propres associées ont pour transformées de Stieltjes $m_A(z)$ et $m_B(z)$, qui sont solutions des équations $L_{mz}^A(m, z) = 0$ et $L_{mz}^B(m, z) = 0$, respectivement, où

$$L_{mz}^A(m, z) = m^2 + zm + 1$$

$$L_{mz}^B(m, z) = m^2z - (-2z + 1)m + 2.$$

La somme et le produit des données indépendantes de ces matrices aléatoires ont une distribution limite de valeurs propres dont la transformée de Stieltjes est solution des équations polynomiales à deux variables $L_{mz}^{A+B}(m, z) = 0$ et $L_{mz}^{AB}(m, z) = 0$ respectivement, qui peuvent être calculées de $L_{mz}^A(m, z)$ et $L_{mz}^B(m, z)$ uniquement. Pour obtenir $L_{mz}^{A+B}(m, z)$, on applique la transformation appelée "Add Atomic Wishart", et pour obtenir $L_{mz}^{AB}(m, z)$, on applique la transformation appelée "Multiply Wishart" (voir figure 3.4). Mais comme nous l'avons dit, nous ne nous intéresserons dans les applications qu'à $L_{mz}^{AB}(m, z)$.

Exemple 6. On aimerait calculer ici $L_{mz}^{AB}(m, z)$ dans le cas où $c = 0.5$. On obtient en appliquant la transformation "Multiply Wishart"

$$L_{mz}^{AB}(m, z) = L_{mz}^A \left((0.5 - 0.5zm)m, \frac{z}{0.5 - 0.5zm} \right) \quad (3.58)$$

Substituant $L_{mz}^A(m, z) = m^2 + zm + 1$ dans (3.58) et en ne conservant que le numérateur de l'expression obtenue, on trouve le polynôme de deux variables

$$L_{mz}^{AB}(m, z) = m^4z^2 - 2m^3z + m^2 + 4mz + 4. \quad (3.59)$$

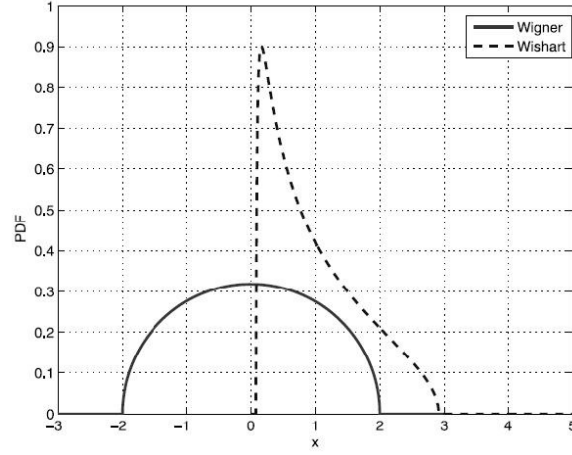


FIGURE 3.3 – Densité limite des valeurs propres pour les matrices GOE et les matrices de Wishart

B	Operation	$L_{mz}^B(m, z)$	MATLAB code
<i>Stochastic transformations</i>			
$A + G'TG$	“Add Atomic Wishart”	$L_{mz}^A(m, z - \alpha_m)$, where $\alpha_m = c \sum_{i=1}^d \frac{p_i \lambda_i}{1 + \lambda_i m}$, with $\sum_i p_i = 1$.	function LmzB = AplusWish(LmzA, c, p, lambda) syms m z alpha = z - c * sum(p .* (lambda ./ (1 + lambda * m))); temp_pol = subs(LmzA, z, z - alpha); LmzB = irreducLuv(temp_pol, m, z);
$A \times W(c)$	“Multiply Wishart”	$L_{mz}^A(\alpha_m, z m, \frac{z}{\alpha_m z})$, where $\alpha_{m,z} = (1 - c - c z m)$.	function LmzB = AtimesWish(LmzA, c) syms m z z1 alpha = (1 - c - c * z1 * m); temp_pol = subs(LmzA, m, m * alpha); temp_pol = subs(temp_pol, z, z1 / alpha); temp_pol = subs(temp_pol, z1, z); % Replace dummy variable LmzB = irreducLuv(temp_pol, m, z);
$(A^{1/2} + \sqrt{s}G) \times (A^{1/2} + \sqrt{s}G)'$	“Grammian”	$L_{mz}^A(\frac{m}{\alpha_m}, \alpha_m^2 z + \alpha_m s(c - 1))$, where $\alpha_m = 1 + s c m$.	function LmzB = AgramG(LmzA, c, s) syms m z alpha = (1 + s * c * m); beta = alpha * (z * alpha + s * (c - 1)); temp_pol = subs(subs(LmzA, m, m / alpha), z, beta); LmzB = irreducLuv(temp_pol, m, z);

FIGURE 3.4 – Lois opérationnelles sur l’obtention des polynômes de deux variables (et leur programmation sur MATLAB). La matrice gaussienne G est de taille $N \times L$, $W(c) = GG'$ est la matrice de Wishart, $c = N/L \rightarrow 0$, lorsque $N, L \rightarrow \infty$, et la matrice T est une matrice aléatoire atomique diagonale.

La densité du produit des deux matrices sus-citées est donnée sur la figure 3.5 qui suit. Ces résultats peuvent aussi s’obtenir en effectuant directement les opérations matricielles ordinaires (produit de deux matrices aléatoires).

En utilisant les interconnexions de la figure B.1, N. Raj Rao et Alan Edelman [35] ont donné les résultats des représentations des polynômes de deux variables de certaines distributions connues. On citera par exemple ici les cas les différents polynômes obtenus pour la loi de Marcenko-Pastur (voir figure 3.6) :

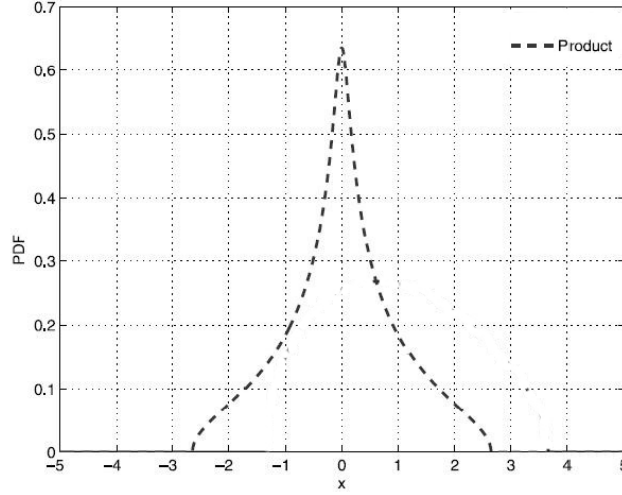


FIGURE 3.5 – Densité limite des valeurs propres pour la somme et le produit des matrices GEO et de Wishart

The Marčenko–Pastur distribution	
L	Bivariate polynomials
L_{mz}	$czm^2 - (1 - c - z)m + 1$
L_{gz}	$czg^2 + (1 - c - z)g + 1$
L_{rg}	$(cg - 1)r + 1$
L_{sy}	$(cy + 1)s - 1$
$L_{\mu z}$	$\mu^2zc - (zc + 1 - z)\mu + 1$
$L_{\eta z}$	$\eta^2zc + (-zc + 1 - z)\eta - 1$

FIGURE 3.6 – Représentation des polynômes de deux variables de deux distributions algébriques : Celle de Marcenko-Pastur et celle du demi-cercle.

3.7.3 Applications de la méthode polynomiale

Nous illustrerons dans cette partie, l'utilisation de la méthode polynomiale décrite un peu plus haut avec certains exemples. Les exemples correspondront aux cas cités dans le tableau de la figure 3.7 ci dessous :

La documentation sur l'implémentation sous MATLAB de la méthode polynomiale est disponible via le package RMTTool [45] . Nous nous intéresserons principalement au cas des matrices aléatoires de Jacobi. Cet ensemble a été défini à l'équation 5. Il est défini en fonction de deux matrices indépendantes de Wishart $\mathbf{W}_1(c_1)$ et $\mathbf{W}_2(c_2)$ comme $\mathbf{J} = (\mathbf{I} + \mathbf{W}_2(c_2)\mathbf{W}_1^{-1}(c_1))^{-1}$. Nous travaillerons dans la suite de cette partie principalement avec les polynômes L_{mz} . Les indices des polynômes ne doivent pas être confondus avec la taille des matrices. Les différents cas des matrices listées dans le tableau de la figure 3.7 sont donnés ci-dessous :

➡ Commençons avec $\mathbf{A}_1 = \mathbf{I}$. Le polynôme de deux variables obtenu par la transformée de Stieltjes de sa fonction de distribution est donné par

$$L_{mz}^1(m, z) = (1 - z)m - 1. \quad (3.60)$$

Nous ne donnons pas de représentation graphique de la densité pour ce polynôme parce que c'est

Transformation	Numerical MATLAB code	Symbolic MATLAB code
Initialization	<code>% Pick n, c1, c2 N1=n/c1; N2=n/c2;</code>	<code>% Define symbolic variables syms m c z;</code>
$\mathbf{A}_1 = \mathbf{I}$	<code>A1 = eye(n,n);</code>	<code>Lmz1 = m*(1-z)-1;</code>
$\mathbf{A}_2 = \mathbf{W}_1(c_1) \times \mathbf{A}_1$	<code>G1 = randn(n,N1)/sqrt(N1); W1 = G1*G1'; A2 = W1*A1;</code>	<code>Lmz2 = AtimesWish(Lmz1,c1);</code>
$\mathbf{A}_3 = \mathbf{A}_2^{-1}$	<code>A3 = inv(A2);</code>	<code>Lmz3 = invA(Lmz2);</code>
$\mathbf{A}_4 = \mathbf{W}_2(c_2) \times \mathbf{A}_3$	<code>G2 = randn(n,N2)/sqrt(N2); W2 = G2*G2'; A4 = W2*A3;</code>	<code>Lmz4 = AtimesWish(Lmz3,c2);</code>
$\mathbf{A}_5 = \mathbf{A}_4 + \mathbf{I}$	<code>A5 = A4+I;</code>	<code>Lmz5 = shiftA(Lmz4,1);</code>
$\mathbf{A}_6 = \mathbf{A}_5^{-1}$	<code>A6 = inv(A5);</code>	<code>Lmz6 = invA(Lmz5);</code>

FIGURE 3.7 – Étapes de construction d’une matrice de Jacobi à l’aide de la matrice identité et de deux matrices de Wishart, et commandes MATLAB utilisées pour générer la densité des valeurs propres associées à cette matrice. Les fonctions utilisées pour générer les polynômes de deux variables (pour chacune des matrices listées) sont données dans le tableau de la figure 3.4.

une droite (plus précisément un segment de l’axe des abscisses) qui ne nous donne aucune information sur notre étude.

- ➔ Pour $\mathbf{A}_2 = \mathbf{W}_1(c_1) \times \mathbf{A}_1$, en utilisant le code MATLAB du tableau de la figure 3.7 correspondant à cette opération, on obtient

$$L_{mz}^2(m, z) = zc_1m^2 - (1 - z - c_1)m + 1. \quad (3.61)$$

Ce produit matriciel représente le produit entre une matrice de Wishart et la matrice identité. Donc le résultat ne sera rien d’autre que la matrice de wishart de départ. En utilisant les commandes développées dans le package RMTTool, nous avons obtenu la figure 3.8 qui suit :

On observe que la densité obtenue par la méthode utilisée ici colle parfaitement avec l’histogramme des valeurs propres. Pour un nombre de simulations assez élevé, on verra très bien cette adhérence de la densité à l’histogramme des valeurs propres.

- ➔ Pour $\mathbf{A}_3 = \mathbf{A}_2^{-1}$, on obtient le polynôme suivant (toujours en utilisant le code MATLAB du tableau de la figure 3.7 correspondant à cette opération)

$$L_{mz}^3(m, z) = z^2c_1m^2 + (c_1z + z - 1)m + 1. \quad (3.62)$$

La matrice \mathbf{A}_2 étant le produit d’une matrice de Wishart et de la matrice identité, reste toujours une matrice de Wishart. Ainsi, la matrice \mathbf{A}_3 n’est rien d’autre que l’inverse d’une matrice de Wishart. À l’aide du package sus-cité, on a obtenu la figure 3.9 qui suit :

On remarque que sur cette figure aussi, la densité limite colle parfaitement avec l’histogramme des valeurs propres. Comme pour le cas précédent, pour un nombre de simulations assez élevé, on verra très bien cette adhérence de la densité à l’histogramme des valeurs propres.

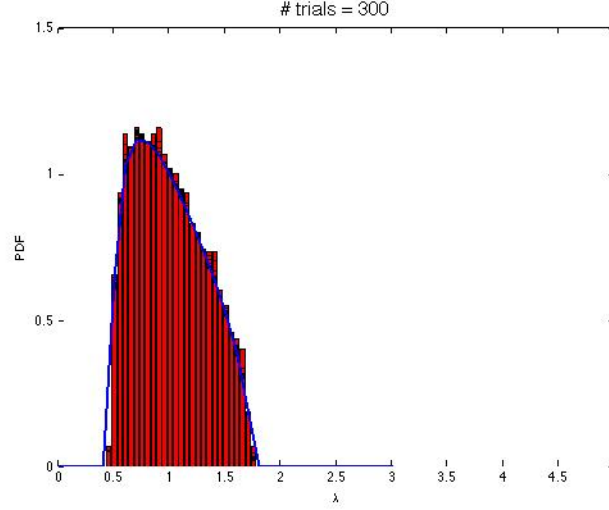


FIGURE 3.8 – Histogramme des valeurs propres de \mathbf{A}_2 (produit d’une matrice de Wishart avec la matrice identité), et, densité limite obtenue via la méthode polynomiale.

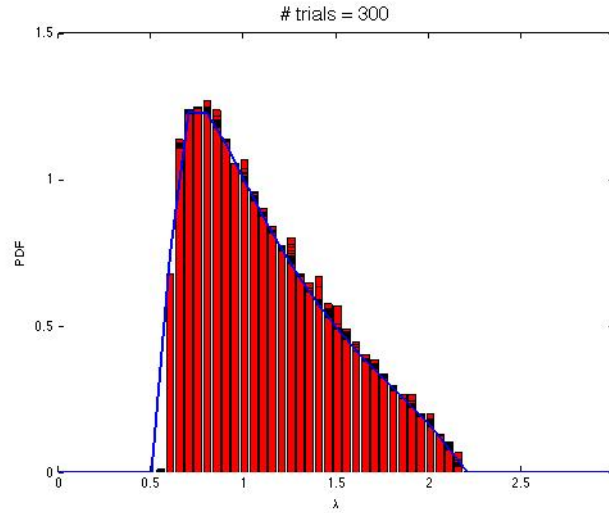


FIGURE 3.9 – Histogramme des valeurs propres de \mathbf{A}_3 (inverse du produit d’une matrice de Wishart avec la matrice identité), et, densité limite obtenue via la méthode polynomiale.

- ➔ Pour $\mathbf{A}_4 = \mathbf{W}_2(c_2) \times \mathbf{A}_3$, on a, en utilisant la même procédure que celle utilisée dans les cas précédents, le polynôme de deux variables ci-dessous :

$$L_{mz}^4(m, z) = (c_1 z^2 + c_2 z) m^2 + (c_1 z + z - 1 + c_2) m + 1. \quad (3.63)$$

La figure 3.10 qui suit présente l’histogramme des valeurs propres de la matrice \mathbf{A}_4 , et la densité limite obtenue grâce au polynôme de l’équation 3.63.

La matrice \mathbf{A}_4 est le produit d’une matrice de covariance (Wishart) et de l’inverse d’une matrice de covariance (Wishart). Dans la littérature, des auteurs ont donné une forme analytique de la densité des valeurs propres de ce type de matrices. Le constat est le même que dans les cas précédents. On observe toujours que la densité limite obtenue grâce au polynôme de l’équation 3.63 colle bien avec l’histogramme des valeurs propres.

- ➔ Pour $\mathbf{A}_5 = \mathbf{A}_4 + \mathbf{I}$, les formules du tableau de la figure 3.7 ont permis d’obtenir le polynôme

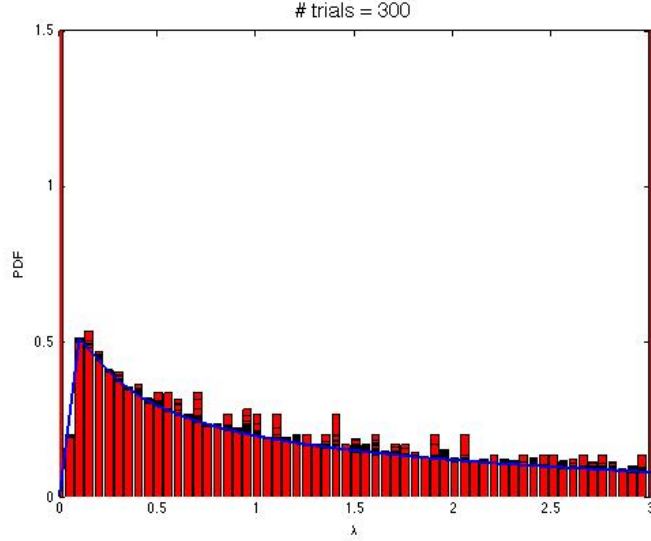


FIGURE 3.10 – Histogramme des valeurs propres de \mathbf{A}_4 (produit d’une matrice de Wishart avec l’inverse du produit d’une matrice de Wishart avec la matrice identité), et, densité limite obtenue via la méthode polynomiale.

suivant :

$$L_{mz}^5(m, z) = ((z - 1)^2 c_1 + c_2(z - 1))m^2 + (c_1(z - 1) + z - 2 + c_2)m + 1; \quad (3.64)$$

Dans ce cas, on perturbe la matrice \mathbf{A}_4 en lui rajoutant la matrice identité. Les transformations du package RMTTool nous ont permis d’obtenir la figure 3.11 qui suit.

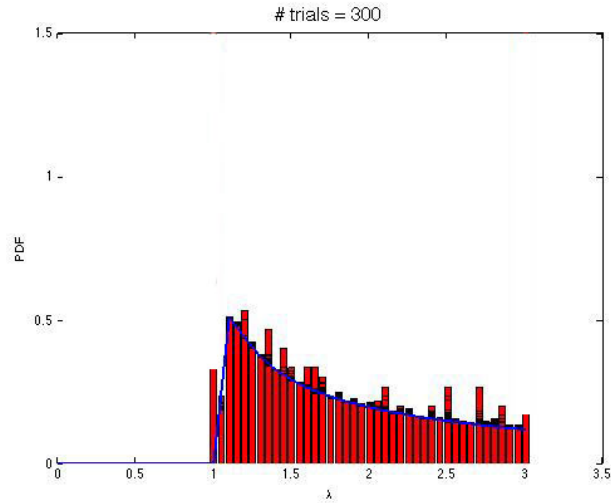


FIGURE 3.11 – Histogramme des valeurs propres de \mathbf{A}_5 (somme de la matrice \mathbf{A}_4 et de la matrice identité), et, densité limite obtenue via la méthode polynomiale.

Dans ce cas, le résultat tient encore comme dans les cas précédents. On observe toujours que la densité limite obtenue grâce au polynôme de l’équation 3.64 colle bien avec l’histogramme des valeurs propres.

Notons sur la figure 3.11 qu’en 1, il y a un pic au niveau de la valeur propre correspondante. Dans la littérature, l’explication donnée porte sur la valeur du paramètre c_2 de la deuxième matrice de

Wishart. En effet, il est dit que, lorsque la valeur de c_2 est supérieure à 1, on a un pic en 1 et si elle est inférieure à 1, il n'y a pas de pic.

- ➔ Comme dernier cas de figure, en utilisant tous les résultats sus-cités, on va étudier le cas de la matrice de Jacobi qui est $\mathbf{J} = \mathbf{A}_6 = \mathbf{A}_5^{-1}$. La procédure utilisée précédemment a permis d'obtenir le polynôme suivant :

$$\begin{aligned} L_{mz}^J(m, z) &\equiv L_{mz}^6(m, z) \\ &= (c_1 z + z^3 c_1 - 2c_1 z^2 - c_2 z^3 + c_2 z^2) m^2 \\ &\quad + (-1 + 2z + c_1 - 3c_1 z + 2c_1 z^2 + c_2 z - 2c_2 z^2) m \\ &\quad - c_2 z - c_1 + 2 + c_1 z. \end{aligned} \quad (3.65)$$

On peut noter au vu de cette équation que, plus l'opération sur les matrices devient complexe, plus le polynôme devient long et pénible à obtenir. Les fonctionnalités du package RMTTool nous ont permis d'obtenir la figure 3.12 qui suit :

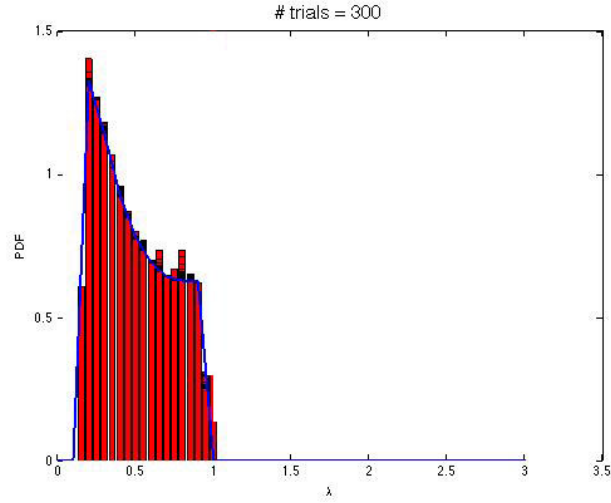


FIGURE 3.12 – Histogramme des valeurs propres de \mathbf{A}_6 (inverse de la matrice \mathbf{A}_5), et, densité limite obtenue via la méthode polynomiale.

Là encore, le constat est le même. La densité obtenue à l'aide de la méthode polynomiale colle très bien avec l'histogramme des valeurs propres. On observe aussi un pic en 1 comme dans le cas précédent, ce qui est aussi dû au fait qu'on ait pris $c_2 > 1$.

Utilisant les arguments théorétiques, il est clair que les ensembles de matrices aléatoires $\mathbf{A}_1, \dots, \mathbf{A}_6$ ne sont définis que pour $c_1 < 1$. Il y aura une masse atomique de poids $1 - \frac{1}{c_2}$ en 1 lorsque $c_2 > 1$ (voir par exemple les figures ci-dessus). La densité limite de chaque ensemble peut être exprimée comme

$$f_{A_i}(x) = \frac{\sqrt{(x - a_-)(a_+ - x)}}{2\pi l_2(x)} \quad \text{pour } a_- < x < a_+, \quad (3.66)$$

pour $i = 2, \dots, 6$, où a_-, a_+ et les polynômes $l_2(x)$ sont exprimés dans le tableau de la figure 3.13 qui suit :

	$l_2(x)$	a_{\pm}
A_2	xc_1	$(1 \pm \sqrt{c_1})^2$
A_3	x^2c_1	$\frac{1}{(1 \mp \sqrt{c_1})^2}$
A_4	$c_1x^2 + c_2x$	$\frac{1 + c_1 + c_2 - c_1c_2 \pm 2\sqrt{c_1 + c_2 - c_1c_2}}{(1 - c_1)^2}$
A_5	$c_1(x - 1)^2 + c_2(x - 1)$	$\frac{c_1^2 - c_1 + 2 + c_2 - c_1c_2 \pm 2\sqrt{c_1 + c_2 - c_1c_2}}{(1 - c_1)^2}$
A_6	$(c_1x + x^3c_1 - 2c_1x^2 - c_2x^3 + c_2x^2)$	$\frac{(1 - c_1)^2}{c_1^2 - c_1 + 2 + c_2 - c_1c_2 \mp 2\sqrt{c_1 + c_2 - c_1c_2}}$

FIGURE 3.13 – Paramètres pour déterminer la fonction de densité limite des valeurs propres en utilisant l'équation 3.66.

3.8 Application financière : Débruitage des matrices de corrélation et distribution empirique des valeurs propres

Ayant présenté tous les outils théoriques, nous allons dans cette partie présenter une analyse des matrices de corrélation empirique pour une loi de Student. Dans la littérature, plusieurs études portent sur la comparaison entre le spectre empirique des données et les résultats prédits par la RMT. Nous allons dans cette partie, appliquer un résultat proposé par Jean-Philippe Bouchaud et Marc Potters dans [32], qui stipule que, le spectre de la matrice de corrélation \mathbf{C} a une distribution de loi puissance (correspondant à une coexistence de larges et de petits secteurs d'activités en finance). Ces auteurs ont proposé la densité ci-dessous pour C :

$$\rho_C(\lambda) = \frac{\mu A}{(\lambda - \lambda_0)^{(1+\mu)}} \Theta(\lambda - \lambda_{\min}), \quad (3.67)$$

avec A et λ_0 reliés à λ_{\min} par la normalisation de ρ_C et par $\text{tr}(\mathbf{C}) = N$. En utilisant l'équation 3.47, on pourrait tracer le spectre $\rho_E(\lambda)$. Dans la figure ?? qui suit, on visualisera le spectre de MP et la distribution de loi puissance donnée dans l'équation 3.67. Pour obtenir la loi puissance, Jean-Philippe Bouchaud et Marc Potters ont paramétré la distribution en utilisant un facteur α , et ont fixé comme hypothèses : $\alpha = \lambda_{\min} \in [0, 1]$, $A = (1 - \alpha^2)$ et $\lambda_0 = 2\alpha - 1$.

Proposons un algorithme pour l'obtention du spectre empirique ρ_E à partir du spectre théorique ρ_C présenté dans l'équation 3.67 :

Algorithm 4 Algorithme pour l'obtention du spectre empirique ρ_E

Require: C , ρ_C , X , z

- 1: C : matrice de corrélation théorique (généralement inconnue).
 - 2: ρ_C : spectre théorique (correspond au spectre de la matrice C). Il est défini dans notre cas à l'équation (3.67).
 - 3: X : matrice de données (rendements dans notre cas), qui permettra d'obtenir la matrice de corrélation empirique (matrice \mathbf{E}).
 - 4: z : nombre complexe qui s'écrit $z = \lambda + i\epsilon$.
 - 5: Utiliser le spectre ρ_C pour obtenir la résolvante $G_E(z)$ à l'aide de l'équation (3.47). Étant donné qu'il est difficile de résoudre l'équation (3.47) analytiquement, il est conseillé et il serait préférable de faire une résolution numérique.
 - 6: De la résolvante $G_E(z)$ (transformée de Stieltjes) de \mathbf{E} , on peut extraire le spectre ρ_E .
 - 7: Tracer ρ_E et comparer cette densité à l'allure de l'histogramme des valeurs propres de la matrice de covariance estimée des données X .
-

3.8.1 Application sur des données simulées

Nous avons estimé par l'algorithme les paramètres de la loi de Student multivariée décrite dans la partie 3.4.2. Pour obtenir nos variables, nous avons généré des variables gaussiennes corrélées, nous avons estimé la matrice de covariance de ces variables corrélées, laquelle nous a permis de générer les variables de Student multivariée. Mais un problème se pose : dans la littérature, il n'existe pas encore de façon de générer des variables de Student multivariée ayant des volatilités différentes. La commande MATLAB actuelle utilisée pour générer des variables de Student multivariée transforme les données de telle sorte que toutes les variances soient égales à 1. Utilisant cette normalisation, nous avons tout d'abord généré des variables qui suivent cette loi multivariée, pour ensuite estimer les paramètres des données générées et tracer l'histogramme des valeurs propres de la matrice de covariance estimée ($\hat{\Sigma}$), et enfin tracer les densités (théorique et empirique) proposées par Bouchaud et Potters. La figure 3.14 qui suit présente les résultats :

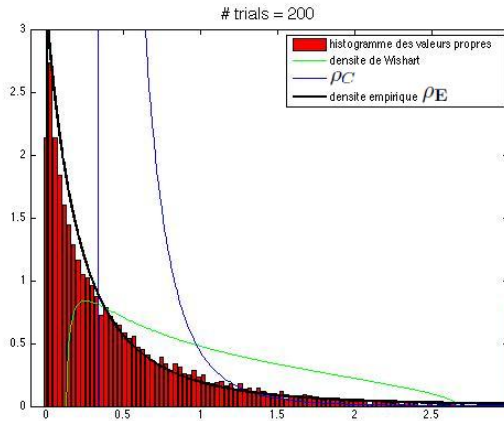


FIGURE 3.14 – Histogramme des valeurs propres de la matrice de covariance empirique, et densités de la loi de MP en vert (pour $c = N/M$), de la loi puissance (ρ_E) en bleu avec comme paramètres $\mu = 2$ et $\lambda = 0.35$ et densité empirique (ρ_C) en noir.

On remarque sur cette figure que, pour des données simulées, la densité empirique obtenue en utilisant l'algorithme 4 approche bien l'histogramme des valeurs propres de la matrice de covariance empirique. Nous allons à présent vérifier si le résultat est le même avec les rendements de marché que nous utilisons.

3.8.2 Application sur des rendements de marché

Dans cette partie, nous nous attarderons à présenter tout d'abord les résultats obtenus avec le jeu de données qu'on utilise depuis le début de ce travail, et ensuite nous analyserons les résultats obtenus avec plusieurs jeux de rendements pour enfin tirer une conclusion sur l'utilisation de la méthode dans le cadre des rendements utilisés.

3.8.2.1 Estimation des paramètres avec notre jeu de données d'application

Comme dans les autres applications, nous allons considérer ici le même jeu de données avec lesquelles nous travaillons depuis le départ. Rappelons qu'il s'agit d'une matrice X de taille 750×10 , où 750 représente le nombre de jours et 10 représente le nombre de titres. Rappelons également que les paramètres d'une loi de Student multivariée a comme paramètres un vecteur μ , un scalaire ν et une matrice C . Comme dans le cas de simulation, nous allons estimer les paramètres du modèle, et nous regarderons si l'histogramme des valeurs propres a la même allure que la densité proposée par Bouchaud et Potters. Pour ce jeu de données, une estimation des paramètres donne les résultats suivants :

- le paramètre $\hat{\nu} = 6.4306$,

– le paramètre $\hat{\mu}$ est s'écrit

$$\hat{\mu} = \begin{pmatrix} -0.0012 \\ 0.0007 \\ 0.0001 \\ 0.0007 \\ 0.0011 \\ 0.0013 \\ 0.0004 \\ 0.0004 \\ -0.0010 \\ 0.0005 \end{pmatrix},$$

– et la matrice de covariance $\hat{\Sigma}$ s'écrit

$$\hat{\Sigma} = 10^{-3} \times \begin{pmatrix} 0.6739 & 0.4071 & 0.2599 & 0.3069 & 0.3102 & 0.3708 & 0.3066 & 0.1465 & 0.2531 & 0.4865 \\ 0.4071 & 0.6264 & 0.2678 & 0.2959 & 0.3039 & 0.3753 & 0.2625 & 0.1050 & 0.1979 & 0.4440 \\ 0.2599 & 0.2678 & 0.3708 & 0.1965 & 0.1902 & 0.2565 & 0.1847 & 0.0829 & 0.1460 & 0.2841 \\ 0.3069 & 0.2959 & 0.1965 & 0.3170 & 0.2385 & 0.2785 & 0.2151 & 0.1186 & 0.1549 & 0.3076 \\ 0.3102 & 0.3039 & 0.1902 & 0.2385 & 0.4351 & 0.3708 & 0.2278 & 0.1107 & 0.1584 & 0.3829 \\ 0.3708 & 0.3753 & 0.2565 & 0.2785 & 0.3708 & 0.6017 & 0.2743 & 0.1226 & 0.1877 & 0.4395 \\ 0.3066 & 0.2625 & 0.1847 & 0.2151 & 0.2278 & 0.2743 & 0.3977 & 0.1122 & 0.1632 & 0.3218 \\ 0.1465 & 0.1050 & 0.0829 & 0.1186 & 0.1107 & 0.1226 & 0.1122 & 0.1676 & 0.1116 & 0.1583 \\ 0.2531 & 0.1979 & 0.1460 & 0.1549 & 0.1584 & 0.1877 & 0.1632 & 0.1116 & 0.3264 & 0.2636 \\ 0.4865 & 0.4440 & 0.2841 & 0.3076 & 0.3829 & 0.4395 & 0.3218 & 0.1583 & 0.2636 & 0.8895 \end{pmatrix}.$$

L'estimation des paramètres effectuée, nous avons calculé les valeurs propres de la matrice $\hat{\Sigma}$ (elles sont normalement égales aux valeurs propres de la matrice de corrélation, d'après les résultats d'algèbre linéaire). Nous obtenons les résultats qui suivent :

$$\Lambda = \begin{pmatrix} 0.0030 \\ 0.0003 \\ 0.0003 \\ 0.0001 \\ 0.0003 \\ 0.0001 \\ 0.0001 \\ 0.0002 \\ 0.0002 \\ 0.0002 \end{pmatrix}.$$

On remarque que, les valeurs propres de la matrice de covariance obtenues en supposant que les rendements suivent une loi de Student multivariée sont très proches de 0. Sur un tracé de l'histogramme des valeurs propres, se résultat se voit mieux. En effet, on observe juste un pic au point 0.

3.8.2.2 Estimation des paramètres dans le cas général

Les résultats obtenus dans la partie 3.8.2.1 n'étant pas satisfaisant, nous avons décidé d'estimer les paramètres des données dans le cas où on considère tous les titres du portefeuille (environ 60). Une estimation des paramètres a montré que dans 95% des cas, la loi de Student n'est pas la loi appropriée pour décrire les rendements ici utilisés. En effet, lorsqu'on estime le paramètre ν de la loi, on se rend compte qu'il varie entre 0.3 et 1.8. Or il devrait être strictement supérieur à deux pour que sa variance puisse exister. Cette condition n'étant pas remplie, nous pouvons directement conclure en disant que le modèle décrit n'est pas approprié pour nos données.

Chapitre 4

Conclusions et Perspectives

La théorie des matrices aléatoires étant encore très nouvelle dans le domaine de la finance, plusieurs questions sont jusqu'aujourd'hui sujets de recherches (ou ont déjà été résolus mais sont difficiles à mettre en oeuvre numériquement), et s'avèrent être très prometteurs. On peut par exemple citer les matrices de "Lévy", le "nettoyage" des vecteurs propres, les matrices de corrélation dépendant du temps ou encore les corrélations non linéaires.

4.1 Nettoyage des vecteurs propres

La théorie des matrices aléatoires a beaucoup contribué et contribue à l'amélioration de la reconstruction des matrices de corrélation à partir des données empiriques. Cependant, l'amélioration est limitée par le fait que la plupart des résultats de la RMT concerne les valeurs propres mais disent peu de choses à propos des vecteurs propres. Il n'est pas évident de formuler des résultats naturels pour la structure de ces vecteurs propres. De là symétrie, on peut juste dire que la plus grande valeur propre de la matrice de corrélation d'un ensemble d'actions pourrait être uniforme. Comme perspective d'étude, on pourrait par exemple utiliser les modèles de facteur, ou les modèles d'arbre ultramétrique. Bien que la connaissance de l'influence du bruit sur le spectre des valeurs propres est tout à fait satisfaisant, la "vraie" façon de mesurer le bruit sur les vecteurs propres est largement inexploré (le cas de la valeur propre supérieure a été traité dans la partie IV de [32]). Des techniques statistiques pour "nettoyer" les vecteurs propres avec une structure non triviale sont nécessaires (pour une tentative récente, voir [51]). Les résultats concernant la structure des vecteurs propres sont difficiles dès qu'on a plus d'hypothèses sur l'invariance statistique sous les transformations orthogonales. Par exemple, la structure des vecteurs propres des matrices de Lévy est extrêmement riche et montre numériquement des transitions de localisation intéressantes. Cependant, on ne dispose pas de beaucoup de résultats analytiques.

4.2 Matrices de Levy

La plupart des résultats cités dans ce rapport (ou dans les articles cités dans la bibliographie) sur le spectre des matrices aléatoires est en grande partie universelle en ce qui concerne la distribution des éléments matriciels. Bien que beaucoup de ces résultats soient faciles à obtenir sous la supposition que les variables aléatoires considérées pour leur construction sont gaussiennes, cette supposition n'est pas cruciale. Par exemple, la distribution du demi-cercle de Wigner tient pour n'importe quelle grandes matrices symétriques composées d'éléments *i.i.d.*, pourvu que celles-ci aient un moment d'ordre deux fini.

On s'attend cependant à ce que les résultats changent quand la queue de distribution de ces éléments est si lourde que le deuxième moment diverge, correspondant à un index de queue μ inférieur à 2. La généralisation de la distribution Wigner dans ce cas (appelé matrices de Levy, parce que l'ensemble correspondant est stable par addition) a été mise au point dans [52] en utilisant des méthodes heuristiques [53]. Leur résultat sur $\rho(\lambda)$ a récemment été rigoureusement prouvé dans [54]. La caractéristique remarquable est que le support du spectre n'est plus borné ; en réalité, $\rho(\lambda)$ décroît pour λ grand avec exactement la même queue que celle de la distribution de chaque élément.

Remarquons que, bien que les matrices de Levy soient stables par addition par construction, deux telles matrices ne sont pas mutuellement libres. Le problème vient en particulier des grandes valeurs

propres ; les vecteurs propres correspondants sont semblables à ceux de la base canonique. Donc on ne peut pas supposer que la base propre change par une rotation aléatoire. Un ensemble différent peut cependant être construit, où chaque matrice de Levy sera rotée aléatoirement avant d'être sommée (voir [53]). Dans ce cas, "la liberté (freeness)" est imposée à la main et les R-transformations sont additives. Le point fixe $R(z) = z$ dans le cas de Wigner est alors $R(z) = z^{\mu-1}$. Le spectre des valeurs propres diffère cependant de celui obtenu dans [52, 54], bien que les queues asymptotiques soient les mêmes : $\rho(\lambda) \propto \lambda^{-1-\mu}$

La généralisation du résultat de Marcenko-Pastur dans le cas des matrices ayant des queues lourdes est également un travail récent [55]. Ici encore, le spectre perd à la fois ses bords supérieur et inférieur pour toutes les valeurs finies de $Q = T/N$ dès que la variance des variables aléatoires r_i^t diverge, i.e. lorsque $\mu < 2$.

4.3 Corrélations non linéaires et Copules

Une caractérisation complète de la structure de corrélation des données revient à déterminer une copule. En particulier, les corrélations non-linéaires, telles que $\langle r_i^2 r_j^2 \rangle_c$, ou les corrélations de queue, peuvent être très importantes pour certaines applications (par exemple, le risque d'un portefeuille d'options). Plusieurs auteurs ont montré que, la copule de Student, qui est intuitive et populaire, ne décrit pas les données de marché. La construction d'un modèle adéquat pour des corrélations linéaires et non-linéaires d'actionss est toujours un problème très ouvert.

4.4 Décomposition en valeur singulière (SVD) aléatoire et Analyse par Composante Canonique

Il est souvent intéressant de considérer des matrices de corrélation non symétriques, ou même rectangulaires, entre N variables "d'entrée" X et M variables de "sortie" Y . La matrice de corrélation empirique sur une période de temps T est définie par

$$\mathcal{E}_{ia} = \frac{1}{T} \sum_{t=1}^T X_i^t Y_a^t. \quad (4.1)$$

On aimerait savoir ce qu'on pourrait dire du spectre de la valeur singulière de \mathcal{E} dans le cas spécial où $N, M, T \rightarrow \infty$, avec $n = N/T$ et $m = M/T$ fixés. Tandis que l'hypothèse nulle naturelle pour des matrices de corrélation est $\mathbf{C} = \mathbf{1}$, ce qui mène à la densité de Marcenko-Pastur, l'hypothèse nulle pour des matrices de corrélations croisées entre les ensembles des variables d'entrée et de sortie est $\mathcal{C} = \mathbf{0}$. Cependant, dans le cas général, les variables d'entrée et de sortie peuvent très bien être corrélées entre elles, par exemple si on choisit des variables d'entrée redondantes. Pour établir un résultat universel, il faudrait donc considérer les composants principaux exacts normalisés pour les variables types X et Y :

$$\hat{X}_\alpha^t = \frac{1}{\sqrt{\lambda_\alpha}} \sum_i V_{\alpha,i} X_i^t \quad \text{et} \quad \hat{Y}_\beta^t = \frac{1}{\sqrt{\lambda_\beta}} \sum_i V_{\beta,i} Y_i^t \quad (4.2)$$

où λ_α , $V_{\alpha,i}$ et λ_β , $V_{\beta,i}$ sont respectivement les valeurs propres et les vecteurs propres de \mathbf{E}_X et \mathbf{E}_Y . On définit maintenant la $M \times N$ matrice de corrélation normalisée comme étant $\hat{\mathcal{E}} = \hat{Y} \hat{X}'$. On peut alors utiliser les astuces suivantes :

- Les valeurs propres non zéro de $\hat{\mathcal{E}}' \hat{\mathcal{E}}$ sont les mêmes que celles de $\hat{X}' \hat{X} \hat{Y}' \hat{Y}$
- $\mathbf{A} = \hat{X}' \hat{X}$ et $\mathbf{B} = \hat{Y}' \hat{Y}$ sont deux $T \times T$ matrices mutuellement libres, avec N (M) valeurs propres exactement égales à 1 (dus à la construction de \hat{X} et de \hat{Y}), et $(T - N)^+$ ($(T - M)^+$) égales à 0.
- Les S -transformations sont multiplicatives, permettant ainsi d'obtenir le spectre de \mathbf{AB}

En raison de la simplicité des spectres de \mathbf{A} et \mathbf{B} , le calcul de la S -transformation est particulièrement facile [60]. Le résultat final de la densité des valeurs singulières (i.e. la racine carrée des valeurs propres de \mathbf{AB}) est :

$$\rho(c) = \max(1 - n, 1 - m) \delta(c) + \max(m + n - 1, 0) \delta(c - 1) + \Re \frac{\sqrt{(c^2 - \gamma_-)(\gamma_+ - c^2)}}{\pi c(1 - c^2)}, \quad (4.3)$$

où $n = N/T$, $m = M/T$ et γ_{\pm} sont donnés par

$$\gamma_{\pm} = n + m - 2mn \pm 2\sqrt{mn(1-n)(1-m)}, \quad 0 \leq \gamma_{\pm} \leq 1 \quad (4.4)$$

Les valeurs de c permises sont toutes entre 0 et 1, comme elles devraient l'être puisque ces valeurs singulières peuvent être interprétées comme des coefficients de corrélation. Lorsque $T \rightarrow \infty$ pour N et M fixés, toutes les valeurs singulières tendent vers 0, comme cela devrait être lorsqu'il n'y a aucune vraie corrélation entre X et Y ; l'intervalle permis lorsque $n, m \rightarrow 0$ devient :

$$c \in \left[\frac{|m-n|}{\sqrt{m} + \sqrt{n}}, \frac{m-n}{\sqrt{m} - \sqrt{n}} \right], \quad (4.5)$$

montrant que pour N, M fixés, l'ordre de grandeur des valeurs singulières décroît comme $T^{-1/2}$.

Notons que l'on peut considérer un ensemble de référence différente, où on considère deux vecteurs indépendants X et Y avec des vraies matrices de corrélation C_X et C_Y égales à $\mathbf{1}$. Le spectre direct SVD dans ce cas peut aussi être calculé comme la S -convolution de deux distributions de Marcenko-Pastur de paramètres m et n . Cette alternative (la référence définie ci-dessus) n'est cependant pas très pratique, dans le sens où elle mélange la possible structure de corrélation non triviale des variables d'entrée et des variables de sortie avec des corrélations croisées entre ces variables.

Mentionnons deux problèmes qui valent le détour : problèmes concernant la décomposition en valeurs singulières aléatoires. Dans le cas où il n'y a aucune corrélation, que ce soit entre N variables d'entrée et M variables de sortie, le spectre des valeurs singulières empiriques avec T observations est donné par l'équation (4.3), qui est l'analogue du résultat de Marcenko-Pastur. Dans les cas pratiques, cependant, il pourrait y avoir de véritables corrélations entre les variables d'entrée et de sortie, décrites par un "vrai" spectre non-trivial des valeurs singulières. La modification de ce spectre, due à la mesure du bruit (i.e l'analogue de l'équation (3.45)) est, à notre connaissance, pas connue. Une analyse détaillée de l'influence de queues lourdes dans la distribution des variables d'entrée et de sortie pourrait également être assez satisfaisante.

Conclusion

J'ai étudié théoriquement et numériquement l'estimation des matrices de variance-covariance via la Théorie des Matrices Aléatoires, pour plusieurs lois données : La loi normale (centrée réduite et cas général), la loi lognormale et la loi de Student. La majeure partie du travail a ensuite été consacrée à l'estimation d'une matrice de covariance dont les variables aléatoires suivent une loi de Student multivariée. Le cas d'une loi normale centrée réduite était déjà traité analytiquement et numériquement dans la littérature. Ces résultats ont été le point de départ pour ce travail, et nous ont permis de généraliser la méthode à d'autres lois, grâce à des travaux effectués par plusieurs auteurs. Les résultats obtenus pour ces différentes lois ont été cités, avec la procédure utilisée.

Ce stage de fin d'études a été formateur sur plusieurs points. Tout d'abord, il m'a appris ce qu'est vraiment le monde de l'entreprise. Ensuite, il m'a fait découvrir ce qu'est une salle de marché, et plus précisément le front office, et m'a enfin donné la possibilité d'utiliser et de développer mes compétences et connaissances sur un projet concret, de plus en plus présent de nos jours pour certains métiers en salle de marché.

Pendant mon stage, j'ai eu la chance de prendre part à un projet neuf dans le domaine de la recherche en finance, vu qu'il ne date que des années 2000. J'ai pu découvrir et approfondir les résultats déjà obtenus et les différentes pistes actuellement explorées par les chercheurs, et appliquer les résultats obtenus sur des données concrètes. C'est donc un bon point de départ pour ma carrière professionnelle puisqu'il m'a permis de voir et de comprendre les outils qui pourront me servir dans l'aboutissement et l'accomplissement de mes tâches quotidiennes.

Cette ouverture sur le monde du travail n'en est plus que bénéfique puisqu'elle m'a donné les clés pour me rendre compte plus clairement de ce que pourra être la suite de ma vie professionnelle.

Humainement, mon intégration au sein de l'entreprise m'a donné l'opportunité de développer mes capacités au travail en équipe et m'a sensibilisé à l'importance de la communication entre les différents acteurs et de travailler dans un environnement professionnel exigeant. Cette intégration m'a aussi permis de développer encore plus ma capacité d'autonomie dans le sens où dans un travail de recherche, c'est après plusieurs essais non fructueux, qu'on peut se poser (et la poser aussi aux autres membres de l'équipe) la question de savoir où est le problème.

En finance, j'ai découvert ce qu'est la gestion d'actifs. C'est une activité qui consiste à gérer les capitaux (détenus en propre ou confiés par un investisseur tiers) dans le respect des contraintes réglementaires et contractuelles, en appliquant les consignes et/ou politiques d'investissements définies par le détenteur des actifs gérés, pour en tirer le meilleur rendement possible en fonction d'un niveau de risque choisi. Mon sujet étant plus axé mathématiques, j'ai beaucoup plus découvert la gestion d'actifs **quantitative**. Ce sujet étant toujours en constante évolution, m'a permis de mettre en oeuvre les connaissances acquises pendant mon cursus (licence : algèbre et analyse, et cursus d'ingénieurs : analyse), et d'apprendre d'autres connaissances requises pour l'exercice de ce métier. C'est dire que, effectuer ce stage, fût une manière de faire converger mes envies et mes attentes.

Annexe A

Quelques résultats en algèbre linéaire

A.1 Matrices de variance-covariance

A.1.1 Définition

En théorie des probabilités et en statistiques, la matrice de variance-covariance est une matrice dont l'élément en position (i, j) correspond à la covariance entre le $i^{\text{ème}}$ et le $j^{\text{ème}}$ éléments d'un vecteur aléatoire X . Chaque élément du vecteur est une variable aléatoire scalaire, avec un nombre fini de valeurs empiriques observées ou avec un nombre fini ou infini de valeurs potentielles indiquées par une distribution de probabilité théorique commune à toutes les variables aléatoires.

Intuitivement, la matrice de variance-covariance généralise la notion de variance en plusieurs dimensions. Définissons le vecteur \mathbf{X} comme

$$\mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}. \quad (\text{A.1})$$

Si les entrées de ce vecteur sont des variables aléatoires, chacune d'elle ayant une variance finie, alors, la matrice de covariance Σ est la matrice dont l'élément (i, j) est la covariance

$$\Sigma_{ij} = \text{cov}(X_i, X_j) = \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)] \quad (\text{A.2})$$

où $\mu_i = \mathbb{E}(X_i)$ est l'espérance du $i^{\text{ème}}$ élément de \mathbf{X} . En d'autres termes, on a :

$$\Sigma = \begin{bmatrix} \mathbb{E}[(X_1 - \mu_1)(X_1 - \mu_1)] & \mathbb{E}[(X_1 - \mu_1)(X_2 - \mu_2)] & \cdots & \mathbb{E}[(X_1 - \mu_1)(X_n - \mu_n)] \\ \mathbb{E}[(X_2 - \mu_2)(X_1 - \mu_1)] & \mathbb{E}[(X_2 - \mu_2)(X_2 - \mu_2)] & \cdots & \mathbb{E}[(X_2 - \mu_2)(X_n - \mu_n)] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}[(X_n - \mu_n)(X_1 - \mu_1)] & \mathbb{E}[(X_n - \mu_n)(X_2 - \mu_2)] & \cdots & \mathbb{E}[(X_n - \mu_n)(X_n - \mu_n)] \end{bmatrix}. \quad (\text{A.3})$$

La définition donnée ci-dessus est équivalente à l'écriture matricielle suivante :

$$\Sigma = \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])'] \quad (\text{A.4})$$

Cette forme peut être vue comme une généralisation du calcul de la variance en grande dimension. Rappelons que pour une variable aléatoire scalaire X ,

$$\sigma_i^2 = \text{var}(X_i) = \mathbb{E}[(X_i - \mathbb{E}(X_i))^2] = \mathbb{E}[(X_i - \mathbb{E}(X_i))(X_i - \mathbb{E}(X_i))]$$

En effet, les entrées de la diagonale de la matrice de covariance Σ sont les variances de chaque élément du vecteur \mathbf{X} .

A.1.2 Propriétés des matrices de variance-covariance

Pour $\Sigma = \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])']$ et $\mu = \mathbb{E}(\mathbf{X})$, où \mathbf{X} est un vecteur aléatoire de dimension p , considérons une autre variable aléatoire \mathbf{Y} de dimension q . On a les propriétés qui suivent :

1. $\Sigma = \mathbb{E}(\mathbf{X}\mathbf{X}') - \mu\mu'$
2. Σ est symétrique et définie semi-positive.
3. $\text{cov}(\mathbf{A}\mathbf{X} + \mathbf{a}) = \mathbf{A} \text{cov}(\mathbf{X}) \mathbf{A}'$
4. $\text{cov}(\mathbf{X}, \mathbf{Y}) = \text{cov}(\mathbf{Y}, \mathbf{X})'$
5. $\text{cov}(\mathbf{X}_1 + \mathbf{X}_2, \mathbf{Y}) = \text{cov}(\mathbf{X}_1, \mathbf{Y}) + \text{cov}(\mathbf{X}_2, \mathbf{Y})$
6. Si $p = q$, alors $\text{var}(\mathbf{X} + \mathbf{Y}) = \text{var}(\mathbf{X}) + \text{cov}(\mathbf{X}, \mathbf{Y}) + \text{cov}(\mathbf{Y}, \mathbf{X}) + \text{var}(\mathbf{Y})$
7. $\text{cov}(\mathbf{A}\mathbf{X} + \mathbf{a}, \mathbf{B}'\mathbf{Y} + \mathbf{b}) = \mathbf{A} \text{cov}(\mathbf{X}, \mathbf{Y}) \mathbf{B}$
8. Si \mathbf{X} et \mathbf{Y} indépendants et décorrélés, alors $\text{cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{0}$

où \mathbf{X} , \mathbf{X}_1 et \mathbf{X}_2 sont des vecteurs aléatoires de taille $p \times 1$, \mathbf{Y} est un vecteur aléatoire de taille $q \times 1$, \mathbf{a} est un vecteur $q \times 1$, \mathbf{b} est un vecteur $p \times 1$, et \mathbf{A} et \mathbf{B} sont des matrices de taille $q \times p$. La matrice de covariance est un instrument important dans plusieurs domaines qu'on ne citera pas ici.

A.1.3 Relation entre $X'X$ et XX'

Pour un vecteur aléatoire (ou même matrice aléatoire de taille $p \times n$ (cas que nous avons utilisé pendant tout notre travail)) \mathbf{X} , on peut bien évidemment définir deux matrices de variance-covariance empirique : $\Sigma_p = \frac{1}{n}X'X$ et $\Sigma_n = \frac{1}{p}XX'$. Les matrices $\frac{1}{n}X'X$ et $\frac{1}{p}XX'$ ont les mêmes valeurs propres non nulles (si elles sont différentes, elles le seront toutes à un même facteur près (T/N)) ; C'est juste que, si $n > p$, alors $\frac{1}{p}XX'$ aura $n - p$ valeurs propres nulles (et vice versa si $n < p$). Les distributions des valeurs propres des matrices de covariance étant bien définies dans les deux cas, nous pouvons aussi très bien écrire une formule sur la relation entre les transformées de Stieltjes des deux distributions :

$$m_{\Sigma_n}(z) = \frac{1}{z} \left(\frac{p}{n} - 1 \right) + \frac{p}{n} m_{\Sigma_p}(z) \quad (\text{A.5})$$

A.2 Formule de l'inverse d'une matrice

Soit $A = (a_{ij})$ une matrice $n \times n$. Notons A_{ij} le cofacteur de a_{ij} . L'expansion de Laplace du déterminant stipule que, pour tout j ,

$$\det(A) = \sum_{i=1}^n a_{ij} A_{ij}. \quad (\text{A.6})$$

Notons $A^a = (A_{ij})'$ la matrice adjacente de A . En appliquant la formule ci-dessus, on obtient

$$AA^a = \det(A)I_n. \quad (\text{A.7})$$

Ces deux résultats prouvent les théorèmes qui suivent.

Théorème 7. *Soit A une matrice $n \times n$ avec déterminant non nul. Alors A est inversible et on a*

$$A^{-1} = \frac{1}{\det(A)} A^a. \quad (\text{A.8})$$

Théorème 8. *On a*

$$\text{tr}(A^{-1}) = \sum_{k=1}^n \frac{A_{kk}}{\det(A)}. \quad (\text{A.9})$$

A.3 Trouver une matrice

Le résultat suivant est connu comme étant la méthode de Hua :

$$\begin{pmatrix} I & \mathbf{O} \\ -CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ \mathbf{O} & D - CA^{-1}B \end{pmatrix} \quad (\text{A.10})$$

Dans la pratique, cette formule peut être considérée comme effectuer une élimination Gaussienne sur les lignes de la matrice $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, pour éliminer le bloc (2-1). Une transformation similaire sur les colonnes tient aussi. Une application importante de cette transformation est le théorème suivant :

Théorème 9. *Si A est une matrice carrée non singulière, alors*

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A) \det(D - CA^{-1}B). \quad (\text{A.11})$$

Ce théorème est obtenu de l'équation (A.10) en calculant le déterminant de chaque coté de l'égalité. Notons que la tranformation (A.10) ne change pas le rang de la matrice.

A.4 Trace d'une matrice inverse

Pour une matrice A de taille $n \times n$, définissons A_k , une sous-matrice majeure d'ordre $n - 1$, qui résulte de la matrice A dans laquelle on supprime la k^{eme} ligne et la k^{eme} colonne. Utilisant les équations (A.8) et (A.11), on a les résultats ci-dessous :

Théorème 10. *Si les matrices A et A_k , $k = 1, 2, \dots, n$ sont non singulières, et si on écrit $A^{-1} = [a^{kl}]$, alors*

$$a^{kk} = \frac{1}{a_{kk} - \alpha'_k A_k^{-1} \beta_k} \quad (\text{A.12})$$

et donc

$$tr(A^{-1}) = \sum_{k=1}^n \frac{1}{a_{kk} - \alpha'_k A_k^{-1} \beta_k}, \quad (\text{A.13})$$

où a_{kk} est le k^{eme} élément de la diagonale de A , A_k est défini ci-dessus, α'_k est le vecteur obtenu de la k^{eme} ligne de la matrice A , en lui enlevant le k^{eme} élément, et β_k est le vecteur obtenu de la k^{eme} colonne de A , en lui enlevant le k^{eme} élément.

Si A est une matrice $n \times n$ symétrique et non singulière, et toutes ses sous-matrices majeures (d'ordre $(n - 1)$) A_k , $k = 1, 2, \dots, n - 1$ sont non singulières, alors de l'équation (A.13), on a immédiatement que,

$$tr(A^{-1}) = \sum_{k=1}^n \frac{1}{a_{kk} - \alpha'_k A_k^{-1} \alpha_k}. \quad (\text{A.14})$$

Si A est une matrice hermitienne singulière $n \times n$, et toutes ses sous-matrices majeures A_k , $k = 1, 2, \dots, n - 1$ d'ordre $(n - 1)$ sont singulières, nous avons

$$tr(A^{-1}) = \sum_{k=1}^n \frac{1}{a_{kk} - \alpha_k^* A_k^{-1} \alpha_k}, \quad (\text{A.15})$$

où $*$ représente la transposée des conjugués complexes des matrices ou des vecteurs.

Dans notre rapport, nous utiliserons très souvent la résolvante d'une matrice hermitienne $X = (x_{jk})$ (i.e. $A = (X - zI)^{-1}$), où z est un nombre complexe, avec une partie imaginaire positive. Dans ce cas, nous avons

$$tr((X - zI)^{-1}) = \sum_{k=1}^n \frac{1}{x_{kk} - z - x_k^* H_k^{-1} x_k}, \quad (\text{A.16})$$

où H_k est la matrice obtenue de la matrice $X - zI$ en lui enlevant la k^{eme} ligne et la k^{eme} colonne, et x_k est la k^{eme} colonne de X dans laquelle on a enlevé le k^{eme} élément.

A.5 Différence des Traces d'une matrice et de ses sous-matrices majeures

Théorème 11. *Si les matrices A et A_k sont non singulières et symétriques, alors*

$$\text{tr}(A^{-1}) - \text{tr}(A_k^{-1}) = \frac{1 + \alpha'_k A_k^{-2} \alpha_k}{a_{kk} - \alpha'_k A_k^{-1} \alpha_k}. \quad (\text{A.17})$$

Ai la matrice A est hermitienne, alors α'_k est remplacé par α_k^ dans l'égalité ci-dessus.*

A.6 Matrice inverse des matrices complexes

Théorème 12. *Si les matrices Hermitiennes A et B commutent et sont tel que $A^2 + B^2$ est non singulière, alors, la matrice complexe $A + iB$ est non singulière et*

$$(A + iB)^{-1} = (A - iB)(A^2 + B^2)^{-1}. \quad (\text{A.18})$$

Ce résultat peut facilement être vérifié.

Soit $z = u + iv$, $v > 0$, et soit A une matrice hermitienne $n \times n$. Alors

$$|\text{tr}(A - zI_n)^{-1} - \text{tr}(A_k - zI_{n-1})^{-1}| \leq v^{-1}. \quad (\text{A.19})$$

A.7 Inégalités de perturbation

Théorème 13. *1. Soient \mathbf{A} et \mathbf{B} deux matrices normales $n \times n$, avec les valeurs propres λ_k et δ_k , $k = 1, 2, \dots, n$ respectivement. Alors*

$$\min \sum_{k=1}^n |\lambda_k - \delta_{\pi(k)}|^2 \leq \text{tr}[(\mathbf{A} - \mathbf{B})(\mathbf{A} - \mathbf{B})^*] \leq \max_{\pi} \sum_{k=1}^n |\lambda_k - \delta_{\pi(k)}|^2, \quad (\text{A.20})$$

où $\pi = (\pi(1), \dots, \pi(n))$ est une permutation de $1, 2, \dots, n$.

2. Dans 1., si \mathbf{A} et \mathbf{B} sont deux matrices $n \times p$ et λ_k et δ_k , $k = 1, 2, \dots, n$ représentent leurs valeurs singulières, alors la conclusion de l'équation A.20 reste vraie. Si les valeurs singulières sont rangées dans l'ordre décroissant, alors on a

$$\sum_{k=1}^n |\lambda_k - \delta_k|^2 \leq \text{tr}[(\mathbf{A} - \mathbf{B})(\mathbf{A} - \mathbf{B})^*], \quad (\text{A.21})$$

où $\nu = \min\{p, n\}$.

Théorème 14. *Soit $\{\lambda_k\}$ et $\{\delta_k\}$ deux ensembles de nombres complexes et leurs distributions empiriques notées par F et \bar{F} . Alors, quelque soit $\alpha > 0$, on a*

$$L(F, \bar{F})^{\alpha+1} \leq \min_{\pi} \frac{1}{n} \sum_{k=1}^n |\lambda_k - \delta_{\pi(k)}|^{\alpha}, \quad (\text{A.22})$$

où L est la distance de Levy entre deux fonctions de distribution F et G de dimension 2, définie à l'équation 2.10

En combinant les Théorèmes 13 et 13 avec $\alpha = 2$, on obtient les deux corollaires qui suivent :

Corollaire 5. *Soit \mathbf{A} et \mathbf{B} deux matrices normales $n \times n$ d'ESD respectifs $F^{\mathbf{A}}$ et $F^{\mathbf{B}}$. Alors*

$$L^3(F^{\mathbf{A}}, F^{\mathbf{B}}) \leq \frac{1}{n} \text{tr}[(\mathbf{A} - \mathbf{B})(\mathbf{A} - \mathbf{B})^*]. \quad (\text{A.23})$$

Corollaire 6. *Soient \mathbf{A} et \mathbf{B} deux matrices $p \times n$ et $F^{\mathbf{S}}$ et $F^{\bar{\mathbf{S}}}$ les ESD de $\mathbf{S} = \mathbf{A}\mathbf{A}^*$ et $\bar{\mathbf{S}} = \mathbf{B}\mathbf{B}^*$. Alors*

$$L^4(F^{\mathbf{S}}, F^{\bar{\mathbf{S}}}) \leq \frac{2}{p^2} (\text{tr}(\mathbf{A}\mathbf{A}^* + \mathbf{B}\mathbf{B}^*)) (\text{tr}[(\mathbf{A} - \mathbf{B})(\mathbf{A} - \mathbf{B})^*]). \quad (\text{A.24})$$

A.8 Inégalités du rang

Dans le cas où les variables ne sont pas iid, ces théorèmes sont très utiles :

Théorème 15. Soient A et B deux matrices hermitiennes $n \times n$. Alors

$$\|F^A - F^B\| \leq \frac{1}{n} \text{rang}(A - B). \quad (\text{A.25})$$

Théorème 16. Soient A et B deux matrices complexes de taille $n \times p$. Alors

$$\|F^{AA^*} - F^{BB^*}\| \leq \frac{1}{p} \text{rang}(A - B). \quad (\text{A.26})$$

En général, si F et D sont deux matrices hermitiennes d'ordres $p \times p$ et $n \times n$ respectivement, alors nous avons

$$\|F^{F+ADA^*} - F^{F+BDB^*}\| \leq \frac{1}{p} \text{rang}(A - B). \quad (\text{A.27})$$

A.9 Lemme de Borel-Cantelli

Le théorème de Borel-Cantelli ou lemme de Borel-Cantelli, nommé d'après les mathématiciens Émile Borel et Francesco Paolo Cantelli, est un résultat de théorie de la mesure très utilisé en théorie des probabilités.

A.9.1 Résultat

En théorie des probabilités, ce théorème concerne une suite d'événements et stipule que :

Lemme 6. (Borel-Cantelli) Si la somme des probabilités d'une suite $(A_n)_{n \geq 0}$ d'événements d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ est finie, alors la probabilité qu'une infinité d'entre eux se réalisent simultanément est nulle.

L'indépendance des événements n'est pas nécessaire. Par exemple, considérons une suite $(X_n)_{n \geq 1}$ de variables aléatoires, telle que, pour tout $n \geq 1$,

$$\mathbb{P}(X_n = 0) = \frac{1}{n^2}.$$

La somme des $\mathbb{P}(X_n = 0)$ est finie¹, donc d'après le lemme de Borel-Cantelli la probabilité que $X_n = 0$ se produise pour une infinité d'indices n est 0. En d'autres termes, avec une probabilité de 1, X_n est non nul à partir d'un certain rang (aléatoire) n_0 . On a donc appliqué le lemme de Borel-Cantelli à la suite d'événements $(A_n)_{n \geq 0}$ définie par

$$A_n = \{X_{n+1} = 0\} = \{\omega \in \Omega \mid X_{n+1}(\omega) = 0\}.$$

A.9.2 Limite supérieure d'ensembles

Définition : La limite supérieure d'une suite $(A_n)_{n \geq 0}$ de parties d'un ensemble Ω est l'ensemble $\limsup_n A_n$ des éléments $\omega \in \Omega$ tels que l'assertion $\{\omega \in A_k\}$ soit vérifiée pour une infinité d'indices $k \geq 0$.

En d'autres termes, on peut dire que $\omega \in \limsup_n A_n$ si et seulement si l'ensemble $\{k \geq 0 \mid \omega \in A_k\}$ est infini, ou bien non borné. Une formulation équivalente est la suivante : pour tout $n \geq 0$, on peut trouver $k \geq n$ tel que $\omega \in A_k$. Cette dernière formulation fournit une écriture commode de la limite supérieure d'ensembles à l'aide d'opérations élémentaires sur les ensembles :

$$\limsup_n A_n = \bigcap_{n \geq 0} \left(\bigcup_{k \geq n} A_k \right).$$

1. En fait, elle vaut $\zeta(2) = \frac{\pi^2}{6}$

Sous l'influence de la terminologie anglo-saxonne, on dira aussi parfois que $\omega \in \limsup_n A_n$ si et seulement si $\{\omega \in A_k\}$ "infiniment souvent" ou bien "infinitely often", d'où la notation rencontrée dans certains ouvrages :

$$\mathbb{P}\left(\limsup_n A_n\right) = \mathbb{P}(A_n \text{ i.o.}).$$

Finalement, remarquons que la définition " $\omega \in \limsup_n A_n$ si et seulement si ω appartient à une infinité de A_k " peut induire en erreur : si par exemple toutes les parties A_k sont égales, il se peut que ω appartienne à A_k pour une infinité d'indices k , et il se peut donc que ω appartienne à $\limsup_n A_n$, sans pour autant qu' ω appartienne à une infinité de A_k (puisque'il n'existe, au fond, qu'un seul A_k).

A.9.3 Théorème de Borel-Cantelli (théorie de la mesure)

Pour un espace mesuré général (X, \mathcal{A}, μ) , le lemme de Borel-Cantelli prend la forme suivante :

Théorème 17. (Borel-Cantelli) Soit $(A_n)_{n \geq 0}$ une suite dans \mathcal{A} . Si

$$\sum_{n \geq 0} \mu(A_n) < +\infty,$$

alors

$$\mu(\limsup_n A_n) = 0.$$

A.9.4 Lemme de Borel-Cantelli (probabilités)

Un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ est un cas particulier d'espace mesuré, en ce qu'on suppose, de plus, que $\mathbb{P}(\Omega)=1$, alors que la seule restriction du même ordre sur (X, \mathcal{A}, μ) est $\mu(X) \geq 0$. En particulier, le lemme de Borel-Cantelli donné dans la partie A.9.1 est une forme affaiblie du Théorème de Borel-Cantelli donné à la section précédente. Peut-être le lemme de Borel-Cantelli est-il plus populaire en probabilités, où il est crucial dans la démonstration, par Kolmogorov, de la loi forte des grands nombres (s'il ne faut donner qu'un seul exemple). Dans le cadre probabiliste, une formulation plus formelle du lemme donné en langage intuitif dans la partie A.9.1 pourrait donc s'écrire :

Lemme 7. (Borel-Cantelli) Dans un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, considérons une suite $(A_n)_{n \geq 0}$ d'éléments de \mathcal{A} . Si

$$\sum_{n \geq 0} \mathbb{P}(A_n) < +\infty,$$

alors

$$\mathbb{P}(\limsup_n A_n) = 0.$$

Annexe B

Méthode des moments et méthodologies

B.1 Théorème de Convergence des Moments

Une des méthodes les plus populaires dans la RMT est la méthode des moments définie dans l'Annexe B.2.1, qui utilise le Théorème de Convergence des Moments (MCT). Soit $\{F_n\}$ une suite de fonctions de distribution avec tous ses moments finis. Le MCT examine sous quelles conditions la convergence des moments de tous ordres fixés implique la convergence faible de la suite des fonctions de distribution $\{F_n\}$.

Notons le $k^{\text{ème}}$ moment de la fonction de distribution F_n par

$$\beta_{n,k} = \beta_k(F_n) := \int x^k dF_n(x). \quad (\text{B.1})$$

Lemme 8. (*Unicité de la limite.*) Une suite de fonctions de distribution $\{F_n\}$ converge lentement vers une limite si les conditions suivantes sont satisfaites :

1. Chaque F_n a ses moments de tous ordres finis.
2. Pour chaque $k \geq 0$ fixé, $\beta_{n,k}$ converge vers une limite finie β_k lorsque $n \rightarrow \infty$.
3. Si deux fonctions F et G , continues à droite et non décroissantes ont la même suite de moments β_k , alors $F = G + \text{const.}$

Lemme 9. (*M. Riesz*). Soit β_k une suite de moments de la fonction de distribution F . Si

$$\liminf_{k \rightarrow \infty} \frac{1}{k} \beta_{2k}^{\frac{1}{2k}} < \infty, \quad (\text{B.2})$$

alors F est uniquement déterminé par la suite $\{\beta_k, k = 0, 1, \dots\}$.

Ce Lemme est un corollaire du prochain Lemme dû à Carleman. Mais avant d'énoncer ce Lemme, énonçons d'abord le Théorème suivant :

Théorème 18. Soit $\{\beta_k = \beta_k(F)\}$ une suite de moments d'une fonction de distribution F . La condition de Carleman stipule que :

$$\sum \beta_{2k}^{-1/2k} = \infty. \quad (\text{B.3})$$

Le Lemme dû à Carleman est le suivant :

Lemme 10. Soit $\{\beta_k = \beta_k(F)\}$ une suite de moments de la fonction de distribution F . Si la condition de Carleman définie dans le Théorème 18 est satisfaite, alors F est uniquement déterminé par la suite $\{\beta_k, k = 0, 1, \dots\}$.

Généralement, la condition B.4 ne peut pas être modifiée (affaiblie). Cependant, pour des distributions unilatérales, elle peut être affaiblie. C'est ce que dit le prochain corollaire. Pour chacune des hypothèses dans le Corollaire, nous considérons que les distributions sont les fonctions de distribution de variables aléatoires non négatives. Il est facile de voir que le corollaire est vrai pour des distributions unilatérales si on change les moments β_k en leur valeur absolue.

Corollaire 7. Soient F et G deux fonctions de distribution avec $F(0_-) = G(0_-)$, $\beta_k(F) = \beta_k(G) = \beta_k$, pour tout entier $k \geq 1$, et

$$\sum \beta_{2k}^{-1/2k} = \infty. \quad (\text{B.4})$$

Alors $F = G$.

B.2 Méthodologies

Les valeurs propres d'une matrice peuvent être vues comme des fonctions continues des entrées de la matrice. Mais ces fonctions n'ont pas de forme «fermée» lorsque la dimension de la matrice est plus grande que 4. Donc des méthodes spéciales sont requises pour les comprendre. Il existe trois principales méthodes utilisées :

- ➔ Méthode des moments,
- ➔ Transformation de Stieltjes,
- ➔ Décomposition polynômes orthogonaux de la densité exacte des valeurs propres.

Mais, dans notre étude, nous nous focaliserons principalement sur les deux premières méthodes que sont la méthode des moments et la transformée de Stieltjes qui sont implicitement liées (Nous le verrons plus tard).

B.2.1 Méthode des moments

Notons par $\{F_n\}$ une suite de fonctions de répartition, et notons le moment d'ordre k de F_n par

$$\beta_{n,k} = \beta_k(F_n) := \int x^k dF_n(x). \quad (\text{B.5})$$

La méthode des moments est basée sur le théorème de convergence des moments.

Soit A une matrice Hermitienne $n \times n$, $\lambda_1 \leq \dots \leq \lambda_n$ ses valeurs propres. L'ESD, F^A de A est défini comme dans l'équation (2.1), en remplaçant m par n . Ainsi, le moment d'ordre k de F^A peut s'écrire comme

$$\beta_{n,k}(A) = \int_{-\infty}^{\infty} x^k F^A(dx) = \frac{1}{n} \text{tr}(A^k). \quad (\text{B.6})$$

Cette expression joue un rôle fondamental dans la RMT. Par le théorème de convergence des moments (voir lemmes B.1, B.2 et B.3), montrer que l'ESD d'une suite de matrices aléatoires $\{A_n\}$ converge vers une certaine limite, revient à montrer que, pour chaque k fixé, la suite $\{\frac{1}{n} \text{tr}(A^k)\}$ converge vers une limite β_k , et à montrer que la condition de Carleman (Théorème 18) est vérifiée,

$$\sum_{k=1}^{\infty} \beta_{2k}^{-1/2k} = \infty. \quad (\text{B.7})$$

Notons que, dans la plupart des cas, la LSD a un support fini, et de là, la fonction caractéristique de la LSD est analytique et la condition nécessaire pour le théorème de convergence des moments tient automatiquement.

B.2.2 Transformée de Stieltjes

Soit A une matrice Hermitienne $n \times n$ et F_n son ESD. La transformation de Stieltjes de F_n est donnée par :

$$s_n(z) = \int \frac{1}{x-z} dF_n(x) = \frac{1}{n} \text{tr}(A - zI)^{-1}. \quad (\text{B.8})$$

où $X := \text{tr}(A - zI)^{-1}$ est la résolvante de la matrice A . Utilisant le Théorème ?? (formule de l'inverse d'une matrice), on obtient

$$s_n(z) = \frac{1}{n} \sum_{k=1}^n \frac{1}{a_{kk} - z - \alpha_k^*(A_k - zI)^{-1} \alpha_k} \quad (\text{B.9})$$

où A_k est la matrice $(n-1) \times (n-1)$ obtenue de la matrice A en enlevant la k^e ligne et la k^e colonne. α_k correspond à la k^e colonne de A dans laquelle on a enlevé le k^e élément.

Si le dénominateur de l'équation (B.9) peut s'écrire sous la forme $g(z, s_n(z)) + o(1)$ pour une certaine fonction g , alors la LSD F existe, et sa transformation de Stieltjes est solution de l'équation

$$s = \frac{1}{g(z, s)}. \quad (\text{B.10})$$

La transformée de Stieltjes peut être vue comme l'espérance

$$s_n(z) = \mathbb{E}_x \left[\frac{1}{x - z} \right],$$

où la variable aléatoire x a pour fonction de distribution $F_n(x)$. Par conséquent, pour toute fonction continue inversible $h(x)$ sur le support de $dF_n(x)$, la transformée de Stieltjes $s_n(z)$ peut aussi être écrite en utilisant la distribution de la variable aléatoire $y = h(x)$ comme

$$s_n(z) = \mathbb{E}_x \left[\frac{1}{x - z} \right] = \mathbb{E}_y \left[\frac{1}{h^{-1}(y) - z} \right], \quad (\text{B.11})$$

où $h^{-1}(\cdot)$ est l'inverse de $h(\cdot)$ (i.e. $h(h^{-1}(x)) = x$). Réciproquement, pour $y = h(x)$, on obtient la relation

$$\mathbb{E}_y \left[\frac{1}{y - z} \right] = \mathbb{E}_x \left[\frac{1}{h(x) - z} \right]. \quad (\text{B.12})$$

Pour $z \in D \equiv \{z \in \mathbb{C} : \Im(z) > 0\}$, on peut trouver la densité de probabilité $f_n(x)$ à partir de la transformée de Stieltjes qui est définie comme dans l'équation (B.8). On a le théorème suivant (formule d'inversion) :

Théorème 19. *La formule d'inversion de Stieltjes-Perron est donnée par*

$$f_n(x) = dF_n(x) = \frac{1}{\pi} \lim_{\epsilon \rightarrow 0} \Im s_n(x + i\epsilon). \quad (\text{B.13})$$

Définissons $S_G(z) = \int \frac{1}{\lambda - z} dG(z)$. On a le théorème ci-dessous

Théorème 20. *Soient a et b deux points de G . Alors :*

$$G[a, b] = \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \int_a^b \Im S_G(x + i\epsilon) dx \quad (\text{B.14})$$

Preuve. Le Théorème 20 est le même que le Théorème 19, à la seule différence que dans le Théorème 20, on se restreint sur un intervalle. Pour une meilleure compréhension de la preuve, nous ne démontrerons que l'équation (B.14). On a :

$$\begin{aligned} \frac{1}{\pi} \int \Im S_G(x + i\epsilon) &= \frac{1}{\pi} \int_a^b \Im \int \left(\frac{dG(\lambda)}{\lambda - x - i\epsilon} \right) dx \\ &= \frac{1}{\pi} \int_a^b \Im \int \left(\frac{\lambda - x}{(\lambda - x)^2 + \epsilon^2} + i \frac{\epsilon}{(\lambda - x)^2 + \epsilon^2} \right) dG(\lambda) dx \\ &= \frac{1}{\pi} \int_a^b \int \frac{\epsilon dG(\lambda)}{(\lambda - x)^2 + \epsilon^2} dx \\ &= \frac{1}{\pi} \int dG(\lambda) \int_a^b \frac{\epsilon}{(\lambda - x)^2 + \epsilon^2} dx \\ &= \frac{1}{\pi} \int dG(\lambda) \left[\arctan \left(\frac{\lambda - a}{\epsilon} \right) - \arctan \left(\frac{\lambda - b}{\epsilon} \right) \right]. \end{aligned} \quad (\text{B.15})$$

En prenant la limite de (B.15) lorsque $\epsilon \rightarrow 0$, et en utilisant le théorème de convergence dominée, on obtient

$$\begin{aligned} \frac{1}{\pi} \int \Im S_G(x + i\epsilon) &= \frac{1}{\pi} \int dG(\lambda) \left[\frac{\pi}{2} - \left(-\frac{\pi}{2} \right) \right] \\ &= \frac{1}{\pi} \int dG(\lambda) \pi \\ &= \int dG(\lambda) \\ &= G\{[a, b]\} \end{aligned} \quad (\text{B.16})$$

□

Comme nous venons de le voir à travers la démonstration ci-dessus, la transformée de Stieltjes est très importante. Elle permet par exemple de :

- ➔ Établir la convergence de l'ESD d'une suite de matrices. Pour cela, on a juste besoin de montrer que la convergence de leur transformée de Stieltjes et la LSD peuvent être trouvés par la limite de la transformée de Stieltjes. On a le théorème suivant :

Théorème 21. Soit $\{G_n\}$ une suite de fonctions à variations bornées, et $G(-\infty) = 0$ pour tout n . Alors :

$$\lim_{n \rightarrow \infty} s_{G_n}(z) = s(z) \quad \forall z \in D \quad (\text{B.17})$$

si et seulement si il existe une fonction G à variations bornées avec $G(-\infty) = 0$, et une transformation de Stieltjes tel que G_n converge lentement vers G .

- ➔ Trouver la fonction de densité. Pour cela, on a le théorème suivant :

Théorème 22. Soit G une fonction à variations bornées et $x_0 \in \mathbb{R}$. Supposons que $\lim_{z \in D \rightarrow x_0} \Im s_G(z)$ existe et est notée $\Im s_G(x_0)$. Alors G est différentiable en x_0 et sa dérivée est $\frac{1}{\pi} \Im s_G(x_0)$.

De la transformée de Stieltjes découlent plusieurs autres transformées très importantes. On note par exemple ¹ (On ne citera que celles qui nous serviront. Pour avoir toutes les transformées et les différentes interconnections, voir la figure B.1) :

- fonction verte : $G(z) = -S(z)$; $\rho(\lambda) = \lim_{\epsilon \rightarrow 0} \frac{1}{\pi} \Im(G(x + i\epsilon))$
- fonction bleue : $B[G(z)] = z$
- R-transformation : $R(z) = B(z) - \frac{1}{z}$, ou encore $\frac{1}{G(z)} + R(G(z)) = z$
- S-transformation : $\mathbf{S}(z) = -\frac{1+z}{z} \eta^{-1}(1+z)$ où $\eta(y) = -\frac{1}{y} G\left(-\frac{1}{y}\right)$

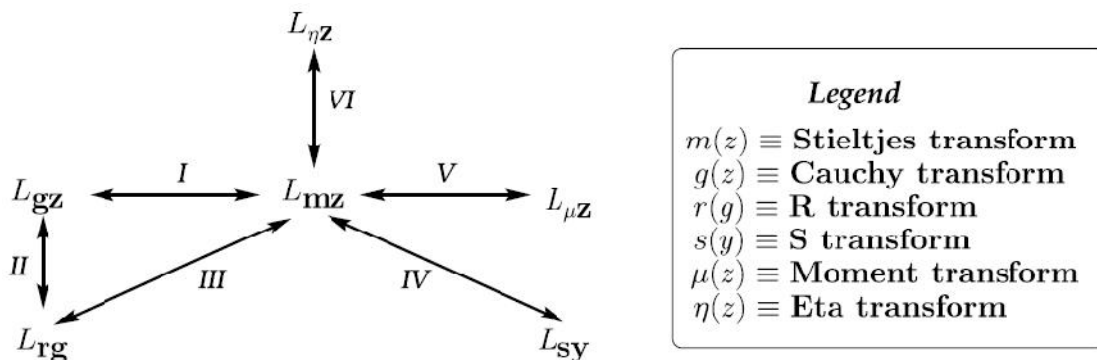


FIGURE B.1 – Différentes interconnections entre les polynômes à deux variables.

Exemple 7. Soit $z = u + iv$, avec $v > 0$, et $s(z)$ la transformée de Stieltjes de la loi de M-P. On a :

$$s(z) = \frac{\sigma^2(1-y) - z + \sqrt{(z - \sigma^2 - y\sigma^2)^2 - 4y\sigma^4}}{2yz\sigma^2} \quad (\text{B.18})$$

Preuve. Pour $y < 1$, on a :

$$s(z) = \int_a^b \frac{1}{x-z} \frac{1}{2\pi xy\sigma^2} \sqrt{(b-x)(x-a)} dx \quad (\text{B.19})$$

1. <http://www.cmap.polytechnique.fr/~euroschoolmathfi/J-P-Bouchaud-Cours3.pdf>

où $a = \sigma^2(1 - \sqrt{y})^2$ et $b = \sigma^2(1 + \sqrt{y})^2$.

En effectuant le CDV $x = \sigma^2(1 + y + 2\sqrt{y}\cos w)$, et en posant $\zeta = e^{iw}$, on a :

$$\begin{aligned}
s(z) &= \int_a^b \frac{1}{x-z} \frac{1}{2\pi xy\sigma^2} \sqrt{(b-x)(x-a)} dx \\
&= \int_0^\pi \frac{2}{\pi} \frac{1}{(1+y+2\sqrt{y}\cos w)(\sigma^2(1+y+2\sqrt{y}\cos w)-z)} \sin^2 w dw \\
&= \frac{1}{\pi} \int_0^{2\pi} \frac{\left(\frac{e^{iw}-e^{-iw}}{2i}\right)^2}{(1+y+2\sqrt{y}(e^{iw}+e^{-iw}))(\sigma^2(1+y+2\sqrt{y}(e^{iw}+e^{-iw}))-z)} dw \\
&= -\frac{1}{4i\pi} \oint_{|\zeta|=1} \frac{(\zeta-\zeta^{-1})^2}{(1+y+\sqrt{y}(\zeta+\zeta^{-1}))(\sigma^2(1+y+\sqrt{y}(\zeta+\zeta^{-1}))-z)} d\zeta \\
&= -\frac{1}{4i\pi} \oint_{|\zeta|=1} \frac{(\zeta^2-1)^2}{\zeta((1+y)\zeta+\sqrt{y}(\zeta^2+1))(\sigma^2(1+y)\zeta+\sqrt{y}\sigma^2(\zeta^2+1)-z\zeta)} d\zeta. \tag{B.20}
\end{aligned}$$

La fonction intégrante a cinq pôles (les ζ qui annulent le dénominateur de la fonction intégrante de l'équation (B.20)...) simples en :

$$\begin{aligned}
\zeta_0 &= 0 \\
\zeta_1 &= \frac{-(1+y)+(1-y)}{2\sqrt{y}} \\
\zeta_2 &= \frac{-(1+y)-(1-y)}{2\sqrt{y}} \\
\zeta_3 &= \frac{-\sigma^2(1+y)+z+\sqrt{\sigma^4(1-y)^2-2\sigma^2(1+y)z+z^2}}{2\sigma^2\sqrt{y}} \\
\zeta_4 &= \frac{-\sigma^2(1+y)+z-\sqrt{\sigma^4(1-y)^2-2\sigma^2(1+y)z+z^2}}{2\sigma^2\sqrt{y}}.
\end{aligned}$$

Les résidus correspondant à ces pôles sont :

$$\frac{1}{y\sigma^2}, \mp \frac{1-y}{yz} \text{ et } \pm \frac{1}{\sigma^2 yz} \sqrt{\sigma^4(1-y)^2-2\sigma^2(1+y)z+z^2}.$$

Remarquant que $\zeta_3\zeta_4 = 1$ et reprenant la définition de la racine carrée des nombres complexes, nous savons que la partie réelle et la partie imaginaire de $\sqrt{\sigma^4(1-y)^2-2\sigma^2(1+y)z+z^2}$ et $-\sigma^2(1+y)+z$ ont le même signe et de là, $|\zeta_3| > 1$, $|\zeta_4| < 1$. On a également $|\zeta_1| = |-\sqrt{y}| < 1$ et $|\zeta_2| = \left|-\frac{1}{\sqrt{y}}\right| > 1$.

Par l'intégration de Cauchy, nous obtenons :

$$\begin{aligned}
s(z) &= -\frac{1}{2} \left(\frac{1}{y\sigma^2} - \frac{1}{\sigma^2 yz} \sqrt{\sigma^4(1-y)^2-2\sigma^2(1+y)z+z^2} - \frac{1-y}{yz} \right) \\
&= \frac{\sigma^2(1-y)-z+\sqrt{(z-\sigma^2-y\sigma^2)^2-4y\sigma^4}}{2yz\sigma^2}. \quad \text{CQFD pour } y < 1...
\end{aligned}$$

Lorsque $y > 1$, étant donné que la loi de M-P a un "point mass" $1 - \frac{1}{y}$ en 0, $s(z)$ devient $s(z)$ de l'équation (65) auquel on rajoute $-\frac{y-1}{yz}$. En appliquant la même procédure que celle du cas $y < 1$, on trouve qu'ici, $|\zeta_3| = |-\sqrt{y}| > 1$ et $|\zeta_4| = \left|\frac{1}{\sqrt{y}}\right| < 1$ et ainsi, le résidu à ζ_4 devrait être compté dans l'intégrale. On trouve au final le résultat par le même cheminement. Pour le cas $y = 1$, on utilise la continuité. \square

Bibliographie

- [1] <http://www.groupama-am.fr>
- [2] Eugene P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Math.*, 62 :548-564, 1955.
- [3] Mehta, M. L. (2004). *Random Matrices*, third edition. Elsevier/Academic Press, Amsterdam.
- [4] Wigner, E. P. (1955). Characteristic vectors bordered matrices with infinite dimensions. *Ann. Math.* 62, 548-564.
- [5] Wigner, E. P. (1958). On the distributions of the roots of certain symmetric matrices. *Ann. Math.* 67, 325-327.
- [6] Arnold, L. (1967). On the asymptotic distribution of the eigenvalues of random matrices. *J. Math. Anal. Appl.* 20, 262-268
- [7] Arnold, L. (1971). On Wigner's semicircle law for the eigenvalues of random matrices. *Z. Wahrsch. Verw. Gebiete* 19, 191-198.
- [8] Bai, Z. D. and Yin, Y. Q. (1988a). A convergence to the semicircle law. *Ann. Probab.* **16**(2), 863-875.
- [9] Grenander, U. (1963). *Probabilities on Algebraic Structures*. John Wiley & Sons, Inc., New York-London ; Almqvist & Wiksells, Stockholm-Göteborg-Uppsala.
- [10] Pastur, L. A. (1972). On the spectrum of random matrices. *Teoret. Mat. Fiz.* 10, 102-112, (*Teoret. Mat. Phys.* 10, 67-74).
- [11] Pastur, L. A. (1973). Spectra of random self-adjoint operators. *Uspekhi Mat. Nauk*, 28(1), 4-63, (*Russian Math. Surv.* 28(1), 1-67).
- [12] Bai, Z. D., Yin, Y. Q., and Krishnaiah, P. R. (1986). On limiting spectral distribution of product of two random matrices when the underlying distribution is isotropic. *J. Multivariate Anal.* 19, 189-200.
- [13] Grenander, U. and Silverstein, J. W. (1977). Spectral analysis of networks with random topologies. *SIAM J. Appl. Math.* 32(2), 499-519.
- [14] Jonsson, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* 12, 1-38.
- [15] Wachter, K. W. (1978). The strong limits of random matrix spectra for sample matrices of independent elements. *Ann Probab.* 6(1), 1-18.
- [16] Wachter, K. W. (1980). The limiting empirical measure of multiple discriminant ratios. *Ann. Statist.* 8, 937-957.
- [17] Yin, Y. Q. (1986). Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.* 20, 50-68.
- [18] Yin, Y. Q. and Krishnaiah, P. R. (1983). A limit theorem for the eigenvalues of product of two random matrices. *J. Multivariate Anal.* 13, 489-507.
- [19] Yin, Y. Q., Bai, Z. D., and Krishnaiah, P. R. (1983). Limiting behavior of the eigenvalues of a multivariate F matrix. *J. Multivariate Anal.* 13, 508-516.
- [20] Yin, Y. Q., Bai, Z. D., and Krishnaiah, P. R. (1988). On the limit of the largest eigenvalue of the large dimensional sample covariance matrix. *Probab. Theory Relat. Fields* 78, 509-521.
- [21] Silverstein, J. W. (1985a). The limiting eigenvalue distribution of a multivariate F matrix. *SIAM J. Math. Anal.* 16(3), 641-646.

- [22] Eugene P. Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67 :325-327, 1958.
- [23] Madan Lal Mehta. *Random Matrices*. Academic Press, Boston, second edition, 1991.
- [24] J. Wishart. The generalized product moment distribution in samples from a normal multivariate population. *Biometrika*, 20 A :32-52, 1928.
- [25] V. L. Girko. *Spectral Theory of Random Matrices (Russian)*. Nauka, Moscow, 1988.
- [26] V. A. Marcenko and L. A. Pastur, Math. USSR, Sb. 1, 457-483 (1967).
- [27] Robb J. Muirhead. *Aspects of Multivariate Statistical Theory*. John Wiley & Sons, New York, 1982.
- [28] V.A. Marčenko and L.A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Math USSR Sbornik*, 1 :457-483, 1967.
- [29] Marcenko, V. A. and Pastur, L. A. (1967). Distribution for some sets of random matrices. *Math. USSR-Sb.* 1, 457–483
- [30] fr.wikipedia.org/wiki/Algorithme_espérance-maximisation
- [31] A.P. Dempster, M.M. Laird et Donald Rubin, «Maximum Likelihood from incomplete Data via the EM Algorithm», *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, n° 1, 1977, p. 1-38.
- [32] Jean-Philippe Bouchaud, Marc Potters. Financial Applications of Random Matrix Theory : a Short review.
- [33] Bai, Z. D. and Silverstein, J. W., (1998). No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.* **26**(1), 316-345.
- [34] Jonsson, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12**, 1-38.
- [35] N. Raj Rao and Alan Edelman. The Polynomial Method for Random Matrices.
- [36] J. Wishart, The Generalised Product Moment Distribution in Samples from a Normal Multivariate Population, *Biometrika*, 20A (1928) 32-52.
- [37] Geman, S. (1980). A limit theorem for the norm of random matrices. *Ann. Probab.* 8, 252–261.
- [38] Bai, Z. D., Silverstein, J. W., and Yin, Y. Q. (1988). A note on the largest eigenvalue of a large dimensional sample covariance matrix. *J. Multivariate Anal.* 26, 166–168.
- [39] Bai, Z. D., Yin, Y. Q., and Krishnaiah, P. R. (1987). On the limiting empirical distribution function of the eigenvalues of a multivariate F matrix. *Theory Probab. Appl.* 32, 537–548.
- [40] Bai, Z. D. and Yin, Y. Q. (1988b). Necessary and sufficient conditions for the almost sure convergence of the largest eigenvalue of Wigner matrices. *Ann. Probab.* 16(4), 1729–1741.
- [41] Silverstein, J. W. (1984a). Comments on a result of Yin, Bai and Krishnaiah for large dimensional multivariate F matrices. *J. Multivariate Anal.* 15, 408–409.
- [42] Silverstein, J. W. (1984b). Some limit theorems on the eigenvectors of large dimensional sample covariance matrices. *J. Multivariate Anal.* 15, 295–324.
- [43] Bai, Z. D. and Yin Y. Q. (1993). Limit of the smallest eigenvalue of large dimensional covariance matrix. *Ann. Probab.* 21(3), 1275–1294.
- [44] M. Potters, J.-P. Bouchaud, L. Laloux, *Acta Phys. Pol. B* 36, 2767 (2005).
- [45] N.R. Rao, RMTTool : a random matrix and free probability calculator in MATLAB. <http://www.mit.edu/~raj/rmtool/>.
- [46] Silverstein, J. W. (1995). Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *J. Multivariate Anal.* 5, 331–339.
- [47] Silverstein, W. J. and Bai, Z. D. (1995). On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *J. Multivariate Anal.* 54, 175–192.
- [48] JAGANNATHAN J. and MA T. (2003), Reduction in Large Portfolios : Why Imposing the Wrong Constraints Helps, *Journal of Finance*, 58(4), pp. 1651-1683.
- [49] Thierry Roncalli, Understanding the Impact of Weights Constraints in Portfolios Theory.
- [50] lmba.math.univ-brest.fr/perso/yves.coudene/agreg-diag.pdf

- [51] I. Dimov, P. Kolm, L. Maclin and D. Shider, *Hidden Noise Structure and Random Matrix Models of stock Correlations*, arXiv :09091383
- [52] P. Cizeau and J. P. Bouchaud, Phys. Rev. E, 50 1810 (1994)
- [53] Z. Burda and J. Jurkiewicz, Chapter 13 of this book, arXiv0909.5228
- [54] G. Ben Arous, A. Guionnet, Comm. Math. Phys. 278 715 (2008)
- [55] S. Belinschi, A. Dembo, A. Guionnet, Comm Math. Phys. 289 1023 (2009)
- [56] Z. Burda, A. Görlich, A. Jarosz and J. Jurkiewicz, Physica A, 343, 295-310 (2004)
- [57] J. Baik, G Ben Arous, and S. Péché. Ann. Probab., 33 (2005) 1643.
- [58] for a recent review, see : A. Tulino, S. Verdù, *Random Matrix Theory and Wireless Communications*, Foundations and Trends in Communication and information Theory, **1**, 1-182 (2004).
- [59] D. V. Vioculescu, K. J. Dykema, A. Nica, *Free Random Variables*, AMS, Providence, RI 1992.
- [60] J.-P. Bouchaud, L. Laloux, M. A. Miceli and M. Potters, Eur. Phys. J. B **55** (2007) 201
- [61] G. Biroli, J.-P. Bouchaud, M. Potters, Acta Phys. Pol. B **38**, 4009 (2007).
- [62] Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L., Guhr, T., and Stanley, H. (2002). Random matrix approach to cross correlations in financial data. Phys. Rev. E, 65(6) :066126.
- [63] Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L., and Stanley, H. (1999). Universal and nonuniversal properties of cross correlations in financial time series. Phys. Rev. Lett., 83(7) :1471–1474.
- [64] Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L., and Stanley, H. (2000a). Econo- physics : financial time series from a statistical physics point of view. Physica A, 279 :443–456.
- [65] Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L., and Stanley, H. (2000b). A random matrix theory approach to financial cross-correlations. Physica A, 287 :374–382.
- [66] Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L., and Stanley, H. (2001). Collective behaviour of stock price movements - a random matrix theory approach. Physica A, 299 :175–180.
- [67] Laloux, L., Cizeau, P., Bouchaud, J., and Potters, M. (1999). Noise dressing of financial correlation matrices. Phys. Rev. Lett., 83(7) :1467–1470.
- [68] Laloux, L., Cizeau, P., Potters, M., and Bouchaud, J. (2000). Random matrix theory and financial correlations. Int. J. Theor. Appl. Finance, 3(3) :391–397.
- [69] for a recent review, see : A. Tulino, S. Verdù, *Random Matrix Theory and Wireless Communications*, Foundations and Trends in communication and Information Theory, **1**, 1-182 (2004).

Table des figures

1.1	Organigramme de Groupama Asset Management	14
2.1	Exemple d'un Δ -graphe	19
2.2	Histogramme des valeurs propres d'une matrice dont les données suivent une $\mathcal{N}(0, 1)$, et densité de Marcenko-Pastur	24
2.3	Histogramme des valeurs propres d'une matrice dont les données suivent une $\mathcal{N}(\mu, \sigma^2 \mathbf{I}_n)$, et densité de Marcenko-Pastur	24
2.4	Histogramme des valeurs propres d'une matrice dont les données suivent une loi lognormale, et densité de Marcenko-Pastur	25
2.5	Histogramme des valeurs propres d'une matrice dont les données suivent une loi de Student, et densité de Marcenko-Pastur	26
3.1	Histogramme du produit d'une matrice de Wishart et de l'inverse d'une matrice de Wishart, et densité limite d'une F -matrice	64
3.2	Histogramme du produit de deux matrices de Wishart indépendantes, et densité limite d'une F -matrice	64
3.3	Densité limite des valeurs propres pour les matrices GOE et les matrices de Wishart	76
3.4	Lois opérationnelles sur l'obtention des polynômes de deux variables (et leur programmation sur MATLAB). La matrice gaussienne \mathbf{G} est de taille $N \times L$, $W(c) = GG'$ est la matrice de Wishart, $c = N/L \rightarrow > 0$, lorsque $N, L \rightarrow \infty$, et la matrice \mathbf{T} est une matrice aléatoire atomique diagonale.	76
3.5	Densité limite des valeurs propres pour la somme et le produit des matrices GEO et de Wishart	77
3.6	Représentation des polynômes de deux variables de deux distributions algébriques : Celle de Marcenko-Pastur et celle du demi-cercle.	77
3.7	Étapes de construction d'une matrice de Jacobi à l'aide de la matrice identité et de deux matrices de Wishart, et commandes MATLAB utilisées pour générer la densité des valeurs propres associées à cette matrice. Les fonctions utilisées pour générer les polynômes de deux variables (pour chacune des matrices listées) sont données dans le tableau de la figure 3.4.	78
3.8	Histogramme des valeurs propres de \mathbf{A}_2 (produit d'une matrice de Wishart avec la matrice identité), et, densité limite obtenue via la méthode polynomiale.	79
3.9	Histogramme des valeurs propres de \mathbf{A}_3 (inverse du produit d'une matrice de Wishart avec la matrice identité), et, densité limite obtenue via la méthode polynomiale.	79
3.10	Histogramme des valeurs propres de \mathbf{A}_4 (produit d'une matrice de Wishart avec l'inverse du produit d'une matrice de Wishart avec la matrice identité), et, densité limite obtenue via la méthode polynomiale.	80
3.11	Histogramme des valeurs propres de \mathbf{A}_5 (somme de la matrice \mathbf{A}_4 et de la matrice identité), et, densité limite obtenue via la méthode polynomiale.	80
3.12	Histogramme des valeurs propres de \mathbf{A}_6 (inverse de la matrice \mathbf{A}_5), et, densité limite obtenue via la méthode polynomiale.	81
3.13	Paramètres pour déterminer la fonction de densité limite des valeurs propres en utilisant l'équation 3.66.	82
3.14	Histogramme des valeurs propres de la matrice de covariance empirique, et densités de la loi de MP en vert (pour $c = N/M$), de la loi puissance ($\rho_{\mathbf{E}}$) en bleu avec comme paramètres $\mu = 2$ et $\lambda = 0.35$ et densité empirique ($\rho_{\mathbf{C}}$) en noir.	83

B.1	Différentes interconnexions entre les polynômes à deux variables.	99
-----	---	----

Liste des tableaux

2.1	Tableau récapitulatif des données avant (σ_{xy}, ρ) et après $(\sigma_{xy}^{new}, \rho^{new})$ le changement de base, et des résultats avant (valeurs propres, poids w et trace de la matrice de corrélation C) et après (valeurs propres, poids w^{new} et trace de la nouvelle matrice de corrélation C^{new}) le changement de base dans le cas où on remplace l'une des valeurs propres par 0.	43
2.2	Tableau récapitulatif des données avant (σ_{xy}, ρ) et après $(\sigma_{xy}^{new}, \rho^{new})$ le changement de base, et des résultats avant (valeurs propres, poids w et trace de la matrice de corrélation C) et après (valeurs propres, poids w^{new} et trace de la nouvelle matrice de corrélation C^{new}) le changement de base dans le cas où on remplace les deux valeurs propres par leur moyenne.	44
2.3	Poids GMV dans l'ancienne (w) et dans la nouvelle (w^{new}) base et volatilités des actifs (notons que les volatilités sont égales dans l'ancienne et la nouvelle base : $\sigma_{ii}^{new} = \sigma_{ii}$) : les nouveaux poids ont été obtenus en remplaçant les valeurs propres inférieures à λ_{max} par leur moyenne.	56
2.4	Coefficients de corrélation dans la base de départ	56
2.5	Coefficients de corrélation dans la nouvelle base, dans le cas où on a remplacé les valeurs propres inférieures à λ_{max} par leur moyenne.	56
2.6	Différence entre les coefficients de corrélation dans l'ancienne base et les coefficients de corrélation dans la nouvelle base.	57
2.7	Poids GMV dans l'ancienne (w) et dans la nouvelle (w^{new}) base et volatilités des actifs (notons que les volatilités sont égales dans l'ancienne et la nouvelle base : $\sigma_{ii}^{new} = \sigma_{ii}$) : les valeurs propres inférieures à λ_{max} ont été remplacées par 0.	58
2.8	Coefficients de corrélation dans la nouvelle base, dans le cas où on a remplacé les valeurs propres inférieures à λ_{max} par 0 dans la nouvelle matrice des valeurs propres.	58
2.9	Différence entre les coefficients de corrélation dans l'ancienne base et les coefficients de corrélation dans la nouvelle base	59