

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LA LOI DE WISHART

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN MATHÉMATIQUES

PAR

PILAR MERCEDES AMAYA

SEPTEMBRE 2018

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.10-2015). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je voudrais exprimer ma reconnaissance à mon directeur de mémoire, François Watier, pour m'avoir accompagné et conseillé lors de la création de ce mémoire. Merci de m'avoir donné la chance de me dépasser en tant qu'étudiante et personne.

Il m'est impossible de décrire la gratitude que j'ai envers mes parents pour leur amour inconditionnel et leur patience. Le plus beau cadeau du monde est l'éducation et je les remercie du plus profond de mon cœur de m'avoir soutenue et encouragée tout au long de ma vie étudiante.

Je remercie Ricardo et sa famille pour leur soutien et amour tout au long de ce beau projet.

Finalement, j'adresse mes remerciements aux femmes travaillant aux institutions académiques que j'ai fréquentées lors de mon parcours universitaire qui m'ont énormément aidé : Louise Labelle et Monica Nevins de l'Université d'Ottawa ainsi que Isabella Couture et Gisèle Legault de l'Université du Québec à Montréal.

TABLE DES MATIÈRES

| | |
|---|-----|
| LISTE DES FIGURES | vii |
| LISTE DES SYMBOLES | ix |
| RÉSUMÉ | xi |
| INTRODUCTION | 1 |
| CHAPITRE I | |
| PROPRIÉTÉS DE BASE D'UNE MATRICE ALÉATOIRE DE LOI DE WISHART | 3 |
| 1.1 Matrice aléatoire de loi de Wishart, fonction de densité et fonction de répartition | 3 |
| 1.1.1 Matrice aléatoire de loi de Wishart | 3 |
| 1.1.2 Fonction de densité | 4 |
| 1.1.3 Fonction de répartition | 9 |
| 1.2 Espérance, covariance et variance | 12 |
| 1.2.1 Espérance | 12 |
| 1.2.2 Covariance et variance | 13 |
| 1.3 Fonction caractéristique et fonction génératrice des moments | 14 |
| 1.3.1 Fonction caractéristique | 14 |
| 1.3.2 Fonction génératrice des moments | 15 |
| 1.4 Inférence via la décomposition de Bartlett | 16 |
| CHAPITRE II | |
| LOI DE WISHART INVERSÉE EN LIEN AVEC LA STATISTIQUE BAYÉSIENNE | 21 |
| 2.1 Matrice aléatoire de loi de Wishart Inversée | 21 |
| 2.2 Fonction de densité | 22 |
| 2.3 Approche bayésienne : rappel des éléments de base | 22 |

| | |
|--|----|
| CHAPITRE III | |
| LIENS AVEC D'AUTRES LOIS DE PROBABILITÉS | 29 |
| 3.1 Loi de khi-deux, χ^2 | 29 |
| 3.2 Loi du T^2 de Hotelling | 30 |
| 3.3 Loi du lambda (Λ) de Wilks (généralisation de la loi de de Fisher) . . | 32 |
| 3.4 Loi bêta multivariée | 34 |
| CHAPITRE IV | |
| PROCESSUS DE WISHART | 37 |
| 4.1 Processus stochastique | 37 |
| 4.2 Processus de Wishart | 39 |
| 4.3 Processus d'Orstein-Uhlenbeck au carré | 43 |
| 4.4 Simulation par l'approximation d'Euler | 46 |
| CONCLUSION | 51 |
| ANNEXE A | |
| PRÉLIMINAIRES | 53 |
| A.1 Jacobien de certaines transformations | 53 |
| ANNEXE B | |
| CODE MATLAB | 55 |
| B.1 Code pour l'algorithme d'échantillonnage d'une variable aléatoire de loi de Wishart | 55 |
| B.2 Code pour l'algorithme du processus de Wishart par l'approximation de Euler | 58 |
| RÉFÉRENCES | 63 |

LISTE DES FIGURES

| Figure | Page |
|---|------|
| 4.1 Processus de Wishart à 2 dimensions | 49 |

LISTE DES NOTATIONS

\mathbf{A}^\top : la transposé de la matrice \mathbf{A}

$\mathbf{A} = \mathbf{A}^\top$: la matrice \mathbf{A} est symétrique définie positive

$|\mathbf{A}|$: le déterminant de la matrice \mathbf{A}

$\exp(\mathbf{A})$: exponentielle de la matrice \mathbf{A}

$\text{tr}(\mathbf{A})$: la trace de la matrice \mathbf{A}

$\mathbf{A} \sim N_p(\mathbf{0}, \mathbf{I})$: \mathbf{A} suit une loi normale centrée réduite multidimensionnelle

\mathbf{I}_p : la matrice identité de taille $p \times p$

$\mathbf{A}^{\frac{1}{2}}$: la racine carré définie positive de la matrice \mathbf{A}

\mathbf{A}^{-1} : l'inverse de la matrice \mathbf{A}

Re : partie réelle

RÉSUMÉ

Ce mémoire porte sur la loi de Wishart qui se caractérise particulièrement par son lien avec les matrices symétriques définies positives. Celle-ci est présentée par l'entremise de ses caractéristiques particulières et ses relations avec d'autres lois de probabilités. Par la suite, la loi inverse de Wishart et son implication dans le domaine de la statistique bayésienne sont explorées. Finalement, le processus de Wishart, pour lequel certains modèles stochastiques sont introduits, est présenté. Ces modèles permettront de faire le lien avec le domaine de la mathématique financière. Pour conclure, l'approximation de Euler sera traitée à des fins de modélisation.

Mots clés : Distribution de Wishart, analyse multivariée, décomposition matricielle, processus de Wishart, modèle de Cox-Ingersoll-Ross, processus de Orstein-Uhlenbeck, approximation de Euler.

INTRODUCTION

La loi de Wishart est une loi de probabilité associée aux matrices aléatoires symétriques définies positives. Elle fût introduite par Wishart (1928) et sa découverte poussa de manière importante le développement de l'analyse multivariée. Elle est essentielle à l'estimation de la matrice de variance-covariance d'un échantillon aléatoire provenant d'une loi normale multidimensionnelle. De plus, elle se présente dans l'utilisation de test du rapport de vraisemblance ainsi que la statistique bayésienne dans le contexte de l'analyse multidimensionnelle.

Le processus de Wishart fût introduit par Bru (1991) en commençant sa construction par la somme des carrés de matrice de mouvement brownien. Elle présenta par la suite sa forme générale comme somme des carrés de matrices de processus Ornstein-Uhlenbeck. Ce processus stochastique est un processus de diffusion dans l'espace des matrices symétriques semi-définies positives telles que les matrices de variance-covariance.

La modélisation de la dépendance entre des variables aléatoires est une notion importante dans l'étude de la statistique. Une mesure importante qui capte cette précieuse information est la matrice de variance-covariance. Dans le domaine de la mathématique financière, lors de l'analyse de la variabilité stochastique et covariance dynamique, il est peu plausible que celle-ci soit constante dans le temps. Cette hypothèse peut causer un impact important lors de l'inférence statistique. C'est donc dans ces circonstances que le processus de Wishart joue un rôle clé puisqu'il supporte une matrice de variance-covariance dynamique.

Le mémoire est composé de quatre chapitres. Le premier présente formellement

une variable aléatoire de Wishart et se dédie entièrement à l'analyse de ses traits caractéristiques et de leur construction. Dans ce chapitre, il sera illustré certains de ceux-ci avec un exemple. Dans le second chapitre la loi de Wishart inversée est abordée ainsi que son implication dans la statistique bayésienne. Le troisième chapitre traite des liens de la loi de Wishart avec d'autres loi de probabilité. Enfin, le quatrième chapitre a pour but d'introduire le processus de Wishart tel que Bru (1991) le fit. De plus, cette section abordera certains concepts liés à la mathématique financière et illustrera, par l'entremise de l'approximation de Euler, ce processus.

CHAPITRE I

PROPRIÉTÉS DE BASE D'UNE MATRICE ALÉATOIRE DE LOI DE WISHART

Ce chapitre porte sur les propriétés particulières d'une variable aléatoire de loi de Wishart. Celles-ci seront développées de manière exhaustive. Bien que certains concepts sont hors de l'objectif de compréhension générale de ces caractéristiques, ils seront brièvement introduit afin d'offrir une image représentative du raisonnement derrière chaque propriété étudiée dans ce chapitre.

1.1 Matrice aléatoire de loi de Wishart, fonction de densité et fonction de répartition

1.1.1 Matrice aléatoire de loi de Wishart

Définition 1.1.1. *Soit \mathbf{X} une matrice aléatoire $n \times p$ ($n \geq p$) dont les lignes sont linéairement indépendantes et proviennent d'une loi normale p -dimensionnelle d'espérance $\mathbf{0}$ et variance Σ . C'est-à-dire, \mathbf{X} est de la forme*

$$\begin{pmatrix} \mathbf{x}_1^\top \\ \cdot \\ \cdot \\ \mathbf{x}_n^\top \end{pmatrix}$$

où $\mathbf{x}_i \sim N_p(\mathbf{0}, \Sigma)$, n représente la taille de l'échantillon et p est la dimension des \mathbf{x}_i . Alors, la matrice $\mathbf{S} = \mathbf{X}^\top \mathbf{X}$ est une matrice $p \times p$ définie positive et suit une loi de Wishart écrit $\mathbf{S} \sim W_p(\Sigma, n)$.

Dans l'ensemble de ce document, nous traiterons de la loi de Wishart dont le paramètre n est un nombre entier. Dans ce cas, la loi de Wishart peut être interprétée comme la représentation d'une somme de carrés de n vecteurs aléatoires provenant d'une loi normale multivariée. C'est pour cela qu'elle est souvent décrite comme la généralisation d'une loi multidimensionnelle de khi-deux.

1.1.2 Fonction de densité

La fonction de densité d'une loi de probabilité est à la base de l'étude statistique. Cette sous-section présente deux approches dont la première se réfère aux étapes principales développées par Muirhead (2005). La deuxième, quant à elle, est intimement liée à la fonction Gamma multivariée.

Propriété 1.1.1. Soit \mathbf{S} une matrice aléatoire, $p \times p$, de loi de Wishart avec les paramètres Σ et $n > p - 1$. Sa fonction de densité de probabilité est :

$$f(\mathbf{S}) = \frac{|\mathbf{S}|^{\frac{n-p-1}{2}} \exp(-\frac{1}{2}\text{tr}(\Sigma^{-1}\mathbf{S}))}{\Gamma_p(\frac{n}{2}) 2^{\frac{np}{2}} |\Sigma|^{\frac{n}{2}}}, n \geq p, \mathbf{S} > 0 \quad (1.1)$$

où $\Gamma_p(a)$ est la fonction multivariée tel que :

$$\Gamma_p(a) = \pi^{\frac{p(p-1)}{4}} \prod_{j=1}^p \Gamma[a + (1 - j)]. \quad (1.2)$$

Remarque 1.1.1. Si $n < p$, \mathbf{S} est une matrice singulière et la matrice $\mathbf{S} \sim W_p(\Sigma, n)$ n'a pas de fonction de densité de probabilité.

Démonstration. Considérons la matrice aléatoire de loi normale multivariée, \mathbf{X} , telle qu'à la définition 1.1.1.. Puisque celle-ci est de dimension $n \times p$ avec $n \geq p$, l'on constate que son rang est égal à p avec une probabilité de 1. Donc, si $\mathbf{S} = \mathbf{X}^\top \mathbf{X}$ la densité \mathbf{X} est de la forme :

$$\begin{aligned} & \frac{\exp(-\frac{1}{2}\text{tr}(\mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}))}{(2\pi)^{\frac{np}{2}} |\boldsymbol{\Sigma}|^{\frac{n}{2}}} \\ &= \frac{\exp(-\frac{1}{2}\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{X}^\top \mathbf{X}))}{(2\pi)^{\frac{np}{2}} |\boldsymbol{\Sigma}|^{\frac{n}{2}}} \\ &= \frac{\exp(-\frac{1}{2}\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{S}))}{(2\pi)^{\frac{np}{2}} |\boldsymbol{\Sigma}|^{\frac{n}{2}}} \end{aligned} \quad (1.3)$$

puisque $\text{tr}(\mathbf{X}^\top \boldsymbol{\Sigma}^{-1} \mathbf{X}) = \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{X}^\top \mathbf{X}) = \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{S})$ et l'élément différentiel correspondant à celle-ci est :

$$\frac{\exp(-\frac{1}{2}\text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{S}))}{(2\pi)^{\frac{np}{2}} |\boldsymbol{\Sigma}|^{\frac{n}{2}}} (d\mathbf{X}). \quad (1.4)$$

Observons que \mathbf{X} possède une décomposition orthogonale unique telle que $\mathbf{X} = \mathbf{H}_1 \mathbf{T}$, où \mathbf{T} est une matrice triangulaire supérieure de dimension $p \times p$ dont les éléments sur la diagonale sont positifs et la matrice \mathbf{H}_1 de dimension $n \times p$ est semiorthogonale telle que $\mathbf{H}_1^\top \mathbf{H}_1 = \mathbf{I}_p$. Remarquons que ceci implique $\mathbf{S} = \mathbf{X}^\top \mathbf{X} = (\mathbf{H}_1 \mathbf{T})^\top (\mathbf{H}_1 \mathbf{T}) = \mathbf{T}^\top \mathbf{H}_1^\top \mathbf{H}_1 \mathbf{T} = \mathbf{T}^\top \mathbf{I}_p \mathbf{T} = \mathbf{T}^\top \mathbf{T}$.

Muirhead (2005) introduit une approche qui utilise les propriétés des différentielles afin de calculer le déterminant d'une matrice de dérivées partielles. De plus, il combinera ceci avec les caractéristiques particulières d'une matrice à colonne orthonormale telle que \mathbf{H}_1 afin d'obtenir le lemme suivant.

Lemme 1.1.1. *Soit $\mathbf{X} = \mathbf{H}_1 \mathbf{T}$ tel que définis précédemment et soit $\mathbf{S} = \mathbf{X}^\top \mathbf{X} = \mathbf{T}^\top \mathbf{T} \geq \mathbf{0}$, alors*

$$(d\mathbf{X}) = 2^{-p} |\mathbf{S}|^{(n-p-1)/2} (d\mathbf{S}) (\mathbf{H}_1^\top d\mathbf{H}_1). \quad (1.5)$$

Démonstration. (Muirhead (2005), p.66) □

À présent, nous obtenons l'élément différentiel lié à la densité jointe de \mathbf{S} et \mathbf{H}_1 :

$$\frac{\exp(-\frac{1}{2}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{S}))}{(2\pi)^{\frac{np}{2}}|\boldsymbol{\Sigma}|^{\frac{n}{2}}} \times 2^{-p}|\mathbf{S}|^{(n-p-1)/2}(d\mathbf{S})(\mathbf{H}_1^\top d\mathbf{H}_1). \quad (1.6)$$

Il suffira maintenant d'intégrer relativement à \mathbf{H}_1 sur l'ensemble des matrices telles que celle-ci pour obtenir l'élément différentiel associé à la densité marginale de \mathbf{S} . Nous introduisons cet ensemble de façon formelle par la définition suivante.

Définition 1.1.2. *La variété de Stiefel est l'ensemble des matrices $n \times p$, $1 \leq p \leq n$, telles que :*

$$V_{(p,n)} = \{\mathbf{H}_1(n \times p) : \mathbf{H}_1^\top \mathbf{H}_1 = \mathbf{I}_p\}. \quad (1.7)$$

Muirhead (2005) précise que les entrées de \mathbf{H}_1 doivent être considérés en tant que coordonnées sur une surface de dimension $pn - \frac{1}{2}p(p+1)$ dans un espace euclidien de dimension pn .

Donc on a

$$\int_{V_{(p,n)}} \frac{\exp(-\frac{1}{2}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{S}))}{(2\pi)^{\frac{np}{2}}|\boldsymbol{\Sigma}|^{\frac{n}{2}}} \times 2^{-p}|\mathbf{S}|^{(n-p-1)/2}(d\mathbf{S})(\mathbf{H}_1^\top d\mathbf{H}_1). \quad (1.8)$$

Remarque 1.1.2. Si $n = p$, alors la variété de Stiefel se réfère au groupe orthogonal :

$$V_{(p,p)} = O_p = \{\mathbf{H}(p \times p) : \mathbf{H}^\top \mathbf{H} = \mathbf{I}_p\}. \quad (1.9)$$

Le prochain lemme nous permet de définir l'intégration relativement à \mathbf{H}_1 sur $V_{(p,n)}$.

Lemme 1.1.2. *Sous les hypothèses du lemme 1.1.1., la forme différentielle :*

$$(\mathbf{H}_1^\top d\mathbf{H}_1) \quad (1.10)$$

définit une mesure invariante par la variété de Stiefel $V_{(p,n)}$. De plus, l'intégrale de celle-ci sur l'ensemble $V_{(p,n)}$ est :

$$\int_{V_{(p,n)}} (\mathbf{H}_1^\top d\mathbf{H}_1) = \text{Vol}[V(p, n)] = \frac{2^p \pi^{\frac{pn}{2}}}{\Gamma_p(\frac{n}{2})}. \quad (1.11)$$

Démonstration. (Muirhead (2005), p.70) □

Finalement,

$$\begin{aligned} & \int_{V_{(p,n)}} \frac{\exp(-\frac{1}{2}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{S}))}{(2\pi)^{\frac{np}{2}} |\boldsymbol{\Sigma}|^{\frac{n}{2}}} \times 2^{-p} |\mathbf{S}|^{(n-p-1)/2} (d\mathbf{S}) (\mathbf{H}_1 d\mathbf{H}_1) \\ &= \frac{\exp(-\frac{1}{2}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{S}))}{(2\pi)^{\frac{np}{2}} |\boldsymbol{\Sigma}|^{\frac{n}{2}}} \times 2^{-p} |\mathbf{S}|^{(n-p-1)/2} \times \frac{2^p \pi^{\frac{pn}{2}}}{\Gamma_p(\frac{n}{2})} (d\mathbf{S}) \\ &= \frac{\exp(-\frac{1}{2}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{S}))}{(2)^{\frac{np}{2}} |\boldsymbol{\Sigma}|^{\frac{n}{2}} \Gamma_p(\frac{n}{2})} \times |\mathbf{S}|^{(n-p-1)/2} (d\mathbf{S}) \end{aligned} \quad (1.12)$$

d'où l'on obtient directement (1.1). □

Il est aussi possible d'obtenir cette fonction de densité par l'intermédiaire de la fonction de Gamma multivariée. Celle-ci est aussi définie sous la forme :

$$\Gamma_p(a) = \int_{\mathbf{A} > \mathbf{0}} \exp(\text{tr}(-\mathbf{A})) |\mathbf{A}|^{a - \frac{(p+1)}{2}} (d\mathbf{A}) \quad \text{Re}(a) > \frac{p-1}{2}, \quad (1.13)$$

où $\mathbf{A} > \mathbf{0}$ indique que \mathbf{A} est une matrice $p \times p$ définie positive non nulle.

Démonstration. Commençons par l'évaluation de l'intégrale suivante :

$$\int_{\mathbf{A} > \mathbf{0}} \exp(\text{tr}(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\mathbf{A})) |\mathbf{A}|^{a - \frac{(p+1)}{2}} (d\mathbf{A}), \quad (1.14)$$

tel que $\boldsymbol{\Sigma} > \mathbf{0}$, une matrice symétrique $p \times p$. Posons le changement de variables $\mathbf{A} = (2^{\frac{1}{2}} \boldsymbol{\Sigma}^{\frac{1}{2}}) \mathbf{Y} (2^{\frac{1}{2}} \boldsymbol{\Sigma}^{\frac{1}{2}})^\top$ où \mathbf{Y} est une matrice symétrique arbitraire. Alors, le

Jacobien $J(\mathbf{A} \rightarrow \mathbf{Y}) = |2^{\frac{1}{2}} \Sigma^{\frac{1}{2}}|^{p+1}$ tel que présenté à la propriété A.1.1. à l'annexe A.

Donc,

$$(d\mathbf{A}) = |2^{\frac{1}{2}} \Sigma^{\frac{1}{2}}|^{p+1} (d\mathbf{Y}) = 2^{\frac{p(p+1)}{2}} |\Sigma|^{\frac{(p+1)}{2}} (d\mathbf{Y}) \quad (1.15)$$

et l'intégrale (1.14) devient :

$$\begin{aligned} & \int_{\mathbf{Y} > \mathbf{0}} \exp(\text{tr}(-\frac{1}{2} \Sigma^{-1} (2^{\frac{1}{2}} \Sigma^{\frac{1}{2}}) \mathbf{Y} (2^{\frac{1}{2}} \Sigma^{\frac{1}{2}})^{\top})) \\ & \times |(2^{\frac{1}{2}} \Sigma^{\frac{1}{2}}) \mathbf{Y} (2^{\frac{1}{2}} \Sigma^{\frac{1}{2}})^{\top}|^{a - \frac{(p+1)}{2}} 2^{\frac{p(p+1)}{2}} |\Sigma|^{\frac{(p+1)}{2}} (d\mathbf{Y}) \end{aligned} \quad (1.16)$$

Simplifions et (1.16) devient :

$$\begin{aligned} & \int_{\mathbf{Y} > \mathbf{0}} \exp(\text{tr}(-\mathbf{Y})) |2\mathbf{Y}\Sigma|^{a - \frac{(p+1)}{2}} 2^{\frac{p(p+1)}{2}} |\Sigma|^{\frac{(p+1)}{2}} (d\mathbf{Y}) \\ & = \int_{\mathbf{Y} > \mathbf{0}} \exp(\text{tr}(-\mathbf{Y})) |\mathbf{Y}|^{a - \frac{(p+1)}{2}} 2^{pa} |\Sigma|^a (d\mathbf{Y}) \\ & = \Gamma_p(a) 2^{pa} |\Sigma|^a. \end{aligned} \quad (1.17)$$

Donc,

$$\int_{\mathbf{A} > \mathbf{0}} \frac{\exp(\text{tr}(-\frac{1}{2} \Sigma^{-1} \mathbf{A})) |\mathbf{A}|^{a - \frac{(p+1)}{2}}}{\Gamma_p(a) 2^{pa} |\Sigma|^a} (d\mathbf{A}) = 1 \quad (1.18)$$

Finalement, prenons $a = \frac{n}{2}$ et obtenons la fonction de densité d'une matrice aléatoire de Wishart, \mathbf{A} , d'où :

$$\int_{\mathbf{A} > \mathbf{0}} \frac{\exp(\text{tr}(-\frac{1}{2} \Sigma^{-1} \mathbf{A})) |\mathbf{A}|^{\frac{n-p-1}{2}}}{\Gamma_p(\frac{n}{2}) 2^{\frac{np}{2}} |\Sigma|^{\frac{n}{2}}} (d\mathbf{A}) = 1. \quad (1.19)$$

On constate donc que la matrice aléatoire \mathbf{A} suit une loi de Wishart puisque par définition d'une fonction de densité de probabilité l'intégrale de celle-ci est égale à 1. \square

1.1.3 Fonction de répartition

Nous présentons maintenant la fonction de répartition d'une matrice aléatoire de Wishart. Cette sous-section se base sur l'approche de Gupta et Nagar (2000) et nous débutons par quelques définitions tirées du chapitre 1 de ce dernier.

Définition 1.1.3. Soit \mathbf{S} une matrice symétrique, $p \times p$, et soit V_k l'espace vectoriel de polynômes homogènes de degré k en $\frac{1}{2}p(p+1)$ éléments distincts de \mathbf{S} . L'espace V_k peut être écrit comme une somme directe de sous-espaces invariants irréductibles V_κ où $\kappa = (k_1, \dots, k_p)$, $k_1 + \dots + k_p = k$, et $k_1 \geq \dots \geq k_p \geq 0$. Alors, le polynôme $(\text{tr}(\mathbf{S}))^k \in V_k$ est associé à une partition unique de polynômes $C_\kappa(\mathbf{S}) \in V_\kappa$ tel que

$$(\text{tr}(\mathbf{S}))^k = \sum_{\kappa} C_\kappa(\mathbf{S}) \quad (1.20)$$

Définition 1.1.4. Le polynôme zonal, $C_\kappa(\mathbf{S})$, est la représentation de $(\text{tr}(\mathbf{S}))^k$ dans le sous-espace V_κ .

Définition 1.1.5. La fonction hypergéométrique généralisée à argument(s) matriciel(s) est définie par

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; \mathbf{S}) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_\kappa \cdots (a_p)_\kappa}{k! (b_1)_\kappa \cdots (b_q)_\kappa} \cdot C_\kappa(\mathbf{S}) \quad (1.21)$$

où \mathbf{S} est une matrice symétrique, \sum_{κ} représente la somme de toutes les partitions de k , $a_i, i = 1, \dots, p; b_j, j = 1, \dots, q$ sont des nombres complexes arbitraires et $C_\kappa(\mathbf{S})$ est le polynôme zonal de \mathbf{S} correspondant à κ .

Remarque 1.1.3. Si p et q sont égales à zéro, alors :

$${}_0F_0(x) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{C_\kappa(\mathbf{X})}{k!} = \sum_{k=0}^{\infty} \frac{(\text{tr}(\mathbf{X}))^k}{k!} = \exp(\text{tr}(\mathbf{X})). \quad (1.22)$$

Propriété 1.1.2. Soit \mathbf{S} une matrice aléatoire, $p \times p$, de loi de Wishart avec les paramètres Σ et $n > p - 1$. Sa fonction de répartition est :

$$\mathbb{P}(\mathbf{S} < \Lambda) = \frac{\Gamma_p[\frac{1}{2}(p+1)]|\Lambda|^{\frac{n}{2}}}{(2)^{\frac{np}{2}}|\Sigma|^{\frac{n}{2}}\Gamma_p[\frac{1}{2}(n+p+1)]} {}_1F_1\left(\frac{n}{2}; \frac{1}{2}(n+p+1); -\frac{1}{2}\Sigma^{-1}\Lambda\right), \quad (1.23)$$

où Λ est une matrice $p \times p$ non-singulière.

Démonstration. Par la définition d'une fonction de répartition d'une matrice aléatoire, nous savons que :

$$\begin{aligned} \mathbb{P}(\mathbf{S} < \Lambda) &= \int_{\mathbf{0} < \mathbf{T} < \Lambda} f_{\mathbf{S}}(\mathbf{T})(d\mathbf{T}) \\ &= \frac{1}{\Gamma_p(\frac{n}{2})2^{\frac{np}{2}}|\Sigma|^{\frac{n}{2}}} \int_{\mathbf{0} < \mathbf{T} < \Lambda} |\mathbf{T}|^{\frac{n-p-1}{2}} \exp(\text{tr}(-\frac{1}{2}\Sigma^{-1}\mathbf{T}))(d\mathbf{T}). \end{aligned} \quad (1.24)$$

De plus, si $\mathbf{S} < \Lambda$ alors $\Lambda = \mathbf{S} + \mathbf{X}$ où \mathbf{X} est une matrice définie positive. Considérons le changement de variable $\mathbf{B} = \Lambda^{-\frac{1}{2}}\mathbf{S}\Lambda^{-\frac{1}{2}}$. Alors, par la propriété A.1.3. de l'annexe A, le Jacobien associé à cette transformation est :

$$J((\mathbf{S}, \mathbf{X}) \rightarrow (\Lambda, \mathbf{B})) = |\Lambda|^{\frac{1}{2}(p+1)}.$$

Donc,

$$\begin{aligned} \mathbb{P}(\mathbf{S} < \Lambda) &= \frac{1}{\Gamma_p(\frac{n}{2})2^{\frac{np}{2}}|\Sigma|^{\frac{n}{2}}} \int_{\mathbf{0} < \mathbf{T} < \Lambda} |\mathbf{T}|^{\frac{n-p-1}{2}} \exp(\text{tr}(-\frac{1}{2}\Sigma^{-1}\mathbf{T}))(d\mathbf{T}) \\ &= \frac{1}{\Gamma_p(\frac{n}{2})2^{\frac{np}{2}}|\Sigma|^{\frac{n}{2}}} \int_{\mathbf{0} < \mathbf{B} < \mathbf{I}_p} |\Lambda^{\frac{1}{2}}\mathbf{B}\Lambda^{\frac{1}{2}}|^{\frac{n-p-1}{2}} \exp(\text{tr}(-\frac{1}{2}\Sigma^{-1}\Lambda^{\frac{1}{2}}\mathbf{B}\Lambda^{\frac{1}{2}})) \\ &\quad \times |\Lambda|^{\frac{p+1}{2}}(d\mathbf{B}) \end{aligned} \quad (1.25)$$

Suite à la remarque 1.1.3., (1.25) devient :

$$\begin{aligned}
\mathbb{P}(\mathbf{S} < \mathbf{\Lambda}) &= \frac{1}{\Gamma_p(\frac{n}{2})2^{\frac{np}{2}}|\Sigma|^{\frac{n}{2}}} \int_{\mathbf{0} < \mathbf{T} < \mathbf{\Lambda}} |\mathbf{T}|^{\frac{n-p-1}{2}} \exp(\text{tr}(-\frac{1}{2}\Sigma^{-1}\mathbf{T})) (d\mathbf{T}) \\
&= \frac{1}{\Gamma_p(\frac{n}{2})2^{\frac{np}{2}}|\Sigma|^{\frac{n}{2}}} \int_{\mathbf{0} < \mathbf{B} < \mathbf{I}_p} |\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{B}\mathbf{\Lambda}^{\frac{1}{2}}|^{\frac{n-p-1}{2}} \exp(\text{tr}(-\frac{1}{2}\Sigma^{-1}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{B}\mathbf{\Lambda}^{\frac{1}{2}})) \\
&\quad \times |\mathbf{\Lambda}|^{\frac{p+1}{2}} (d\mathbf{B}) \\
&= \frac{1}{\Gamma_p(\frac{n}{2})2^{\frac{np}{2}}|\Sigma|^{\frac{n}{2}}} \int_{\mathbf{0} < \mathbf{B} < \mathbf{I}_p} |\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{B}\mathbf{\Lambda}^{\frac{1}{2}}|^{\frac{n-p-1}{2}} |\mathbf{\Lambda}|^{\frac{p+1}{2}} \\
&\quad \times {}_0F_0(-\frac{1}{2}\Sigma^{-1}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{B}\mathbf{\Lambda}^{\frac{1}{2}})(d\mathbf{B}) \\
&= \frac{1}{\Gamma_p(\frac{n}{2})2^{\frac{np}{2}}|\Sigma|^{\frac{n}{2}}} \int_{\mathbf{0} < \mathbf{B} < \mathbf{I}_p} |\mathbf{B}|^{\frac{n-p-1}{2}} |\mathbf{\Lambda}|^{\frac{n}{2}} {}_0F_0((-\frac{1}{2}\Sigma^{-1}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{B}\mathbf{\Lambda}^{\frac{1}{2}}))(d\mathbf{B}) \\
&= \frac{|\mathbf{\Lambda}|^{\frac{n}{2}}}{\Gamma_p(\frac{n}{2})2^{\frac{np}{2}}|\Sigma|^{\frac{n}{2}}} \int_{\mathbf{0} < \mathbf{B} < \mathbf{I}_p} |\mathbf{B}|^{\frac{n-p-1}{2}} {}_0F_0((-\frac{1}{2}\Sigma^{-1}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{B}\mathbf{\Lambda}^{\frac{1}{2}}))(d\mathbf{B}).
\end{aligned} \tag{1.26}$$

Lemme 1.1.3. Soit \mathbf{R} , une matrice symétrique $p \times p$, et \mathbf{S} une matrice définie positive, alors

$$\begin{aligned}
&\int_{\mathbf{0} < \mathbf{S} < \mathbf{I}_p} |\mathbf{S}|^{a-\frac{p+1}{2}} |\mathbf{I}_p - \mathbf{S}|^{b-\frac{p+1}{2}} {}_mF_n(a_1, \dots, a_m; b_1, \dots, b_n; \mathbf{RS})(d\mathbf{S}) \\
&= \frac{\Gamma_p(a)\Gamma_p(b)}{\Gamma_p(a+b)} {}_{m+1}F_{n+1}(a_1, \dots, a_m, a; b_1, \dots, b_n, a+b; \mathbf{R}).
\end{aligned} \tag{1.27}$$

Démonstration. (Gupta et Nagar (2000), p.36)

□

Prenons $a = \frac{n}{2}$ et $b = \frac{p+1}{2}$. Alors (1.26) devient :

$$\begin{aligned}
\mathbb{P}(\mathbf{S} < \mathbf{\Lambda}) &= \frac{|\mathbf{\Lambda}|^{\frac{n}{2}}}{\Gamma_p(\frac{n}{2})2^{\frac{np}{2}}|\mathbf{\Sigma}|^{\frac{n}{2}}} \int_{\mathbf{0} < \mathbf{B} < \mathbf{I}_p} |\mathbf{B}|^{\frac{n}{2}-\frac{p+1}{2}} |\mathbf{I}_p - \mathbf{B}|^{\frac{p+1}{2}-\frac{p+1}{2}} {}_0F_0\left(\left(-\frac{1}{2}\mathbf{\Sigma}^{-1}\mathbf{\Lambda}\mathbf{B}\right)\right)(d\mathbf{B}) \\
&= \frac{|\mathbf{\Lambda}|^{\frac{n}{2}}}{\Gamma_p(\frac{n}{2})2^{\frac{np}{2}}|\mathbf{\Sigma}|^{\frac{n}{2}}} \frac{\Gamma_p(\frac{n}{2})\Gamma_p(\frac{n}{p+1})}{\Gamma_p(\frac{n+p+1}{2})} {}_1F_1\left(\frac{n}{2}; \frac{n}{2} + \frac{p+1}{2}; -\frac{1}{2}\mathbf{\Sigma}^{-1}\mathbf{\Lambda}\right) \\
&= \frac{|\mathbf{\Lambda}|^{\frac{n}{2}}\Gamma_p(\frac{n}{p+1})}{2^{\frac{np}{2}}|\mathbf{\Sigma}|^{\frac{n}{2}}\Gamma_p(\frac{n+p+1}{2})} {}_1F_1\left(\frac{n}{2}; \frac{n+p+1}{2}; -\frac{1}{2}\mathbf{\Sigma}^{-1}\mathbf{\Lambda}\right).
\end{aligned} \tag{1.28}$$

□

1.2 Espérance, covariance et variance

1.2.1 Espérance

Propriété 1.2.1. Soit \mathbf{S} une matrice aléatoire, $p \times p$; de loi de Wishart avec les paramètres $\mathbf{\Sigma}$ et $n > p-1$. Son espérance est de la forme :

$$E[\mathbf{S}] = n\mathbf{\Sigma}. \tag{1.29}$$

Démonstration. Puisque $\mathbf{S} = \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top$ et que les vecteurs \mathbf{x}_i sont indépendants et identiquement distribués selon $N_p(\mathbf{0}, \mathbf{\Sigma})$, nous avons :

$$E[S_{ij}] = E\left[\sum_{r=1}^n x_{ir} x_{jr}\right] = \sum_{r=1}^n E[x_{ir} x_{jr}] = \sum_{r=1}^n \Sigma_{ij} = n\Sigma_{ij}. \tag{1.30}$$

□

De plus,

$$E[\mathbf{S}] = E\left[\sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^\top\right] = \sum_{i=1}^n E[\mathbf{x}_i \mathbf{x}_i^\top] = \sum_{i=1}^n \mathbf{\Sigma} = n\mathbf{\Sigma}. \tag{1.31}$$

1.2.2 Covariance et variance

Propriété 1.2.2. Soit S_{ij} et S_{kl} , deux éléments d'une matrice aléatoire de loi de Wishart. La covariance est :

$$\text{Cov}[S_{ij}, S_{kl}] = n(\Sigma_{ik}\Sigma_{jl} + \Sigma_{il}\Sigma_{jk}) \quad (1.32)$$

Démonstration. Par définition de la covariance :

$$\text{Cov}[S_{ij}, S_{kl}] = E[S_{ij}S_{kl}] - E[S_{ij}]E[S_{kl}] \quad (1.33)$$

Trouvons d'abord $E[S_{ij}S_{kl}]$:

$$\begin{aligned} E[S_{ij}S_{kl}] &= E\left[\sum_{r=1}^n x_{ir}x_{jr} \sum_{s=1}^n x_{ks}x_{ls}\right] = \sum_{r=1}^n \sum_{s=1}^n E[x_{ir}x_{jr}x_{ks}x_{ls}] \\ &= \sum_{r=1}^n E[x_{ir}x_{jr}x_{kr}x_{lr}] + \sum_{\substack{r=1 \\ r \neq s}}^n \sum_{s=1}^n E[x_{ir}x_{jr}x_{ks}x_{ls}] \\ &= \sum_{r=1}^n (E[x_{ir}x_{jr}]E[x_{kr}x_{lr}] + E[x_{ir}x_{kr}]E[x_{jr}x_{lr}] + E[x_{ir}x_{lr}]E[x_{jr}x_{kr}]) \\ &\quad + \sum_{\substack{r=1 \\ r \neq s}}^n \sum_{s=1}^n E[x_{ir}x_{jr}]E[x_{ks}x_{ls}] \\ &= n(\Sigma_{ij}\Sigma_{kl} + \Sigma_{ik}\Sigma_{jl} + \Sigma_{il}\Sigma_{jk}) + n\Sigma_{ij}(n-1)\Sigma_{kl}. \end{aligned} \quad (1.34)$$

Donc, (1.31) devient :

$$\begin{aligned} \text{Cov}[S_{ij}, S_{kl}] &= E[S_{ij}S_{kl}] - E[S_{ij}]E[S_{kl}] \\ &= n(\Sigma_{ij}\Sigma_{kl} + \Sigma_{ik}\Sigma_{jl} + \Sigma_{il}\Sigma_{jk}) + n\Sigma_{ij}(n-1)\Sigma_{kl} - n\Sigma_{ij}n\Sigma_{kl} \\ &= n(\Sigma_{ik}\Sigma_{jl} + \Sigma_{il}\Sigma_{jk}) \end{aligned} \quad (1.35)$$

Alors, par définition de la variance, nous savons que :

$$\text{Var}[S_{ij}] = \text{Cov}[S_{ij}, S_{ij}] = n(\Sigma_{ii}\Sigma_{jj} + \Sigma_{ij}^2) \quad (1.36)$$

□

La prochaine propriété ne sera pas démontrée mais servira à des fins de validation lors de la simulation. Pour plus de détails, se référer à Christensen (2015).

Propriété 1.2.3. *Soit \mathbf{S} une matrice aléatoire, $p \times p$, de loi de Wishart avec les paramètres Σ et $n > p-1$. La matrice de variance-covariance de celle-ci est de la forme :*

$$\text{Cov}[\text{Vec}(\mathbf{S})] = n[\Sigma \otimes \Sigma][\mathbf{T} + \mathbf{I}_{p^2}], \quad (1.37)$$

où $\text{Vec}(\mathbf{S})$ est la transformation de la matrice \mathbf{S} en un vecteur, \otimes est le produit de Kronecker et \mathbf{T} est la matrice $p^2 \times p^2$ de commutation ($\mathbf{I}_{p^2} = \mathbf{T}\mathbf{T}$).

1.3 Fonction caractéristique et fonction génératrice des moments

1.3.1 Fonction caractéristique

Propriété 1.3.1. *Soit \mathbf{S} une matrice aléatoire, $p \times p$, de loi de Wishart avec les paramètres Σ et $n > p-1$. Sa fonction caractéristique, c'est-à-dire la fonction caractéristique conjointe de $S_{11}, S_{12}, \dots, S_{pp}$ est :*

$$\phi_{\mathbf{S}}(\mathbf{Z}) = E[\exp(\text{tr}(i\mathbf{Z}\mathbf{S}))] = |\mathbf{I}_p - 2i\mathbf{Z}\Sigma|^{-\frac{n}{2}} \quad (1.38)$$

où \mathbf{Z} est une matrice symétrique de dimension $p \times p$.

Démonstration.

$$\begin{aligned}
\phi_{\mathbf{S}}(\mathbf{Z}) &= E[\exp(\text{tr}(i\mathbf{Z}\mathbf{S}))] = \int_{\mathbf{S}>\mathbf{0}} \exp(\text{tr}(i\mathbf{Z}\mathbf{S})) f_{\mathbf{S}} d\mathbf{S} \\
&= \int_{\mathbf{S}>\mathbf{0}} \exp(\text{tr}(i\mathbf{Z}\mathbf{S})) \frac{|\mathbf{S}|^{\frac{n-p-1}{2}} \exp(-\frac{1}{2}\text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{S}))}{\Gamma_p(\frac{n}{2}) 2^{\frac{np}{2}} |\boldsymbol{\Sigma}|^{\frac{n}{2}}} d\mathbf{S} \\
&= \frac{1}{\Gamma_p(\frac{n}{2}) 2^{\frac{np}{2}} |\boldsymbol{\Sigma}|^{\frac{n}{2}}} \int_{\mathbf{S}>\mathbf{0}} |\mathbf{S}|^{\frac{n-p-1}{2}} \exp(\text{tr}(i\mathbf{Z}\mathbf{S}) + \text{tr}(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\mathbf{S})) d\mathbf{S} \\
&= \frac{1}{\Gamma_p(\frac{n}{2}) 2^{\frac{np}{2}} |\boldsymbol{\Sigma}|^{\frac{n}{2}}} \int_{\mathbf{S}>\mathbf{0}} |\mathbf{S}|^{\frac{n-p-1}{2}} \exp(\text{tr}(-\frac{1}{2}(\mathbf{I}_p - 2i\mathbf{Z}\boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1}\mathbf{S})) d\mathbf{S}
\end{aligned} \tag{1.39}$$

Remarque 1.3.1. En imposant le changement de variable $\mathbf{A} = \mathbf{T}^{\frac{1}{2}}\mathbf{S}\mathbf{T}^{\frac{1}{2}}$ associé au Jacobien $J(\mathbf{S} \rightarrow \mathbf{A}) = |\mathbf{T}|^{-\frac{1}{2}}$ à la loi de gamma multivariée, on observe que :

$$\int_{\mathbf{S}>\mathbf{0}} |\mathbf{S}|^{a-\frac{p+1}{2}} \exp(\text{tr}(-\mathbf{S}\mathbf{T})) d\mathbf{S} = \Gamma_p(a) |\mathbf{T}|^{-a}. \tag{1.40}$$

Insérons $\mathbf{T} = \frac{1}{2}(\mathbf{I}_p - 2i\mathbf{Z})\boldsymbol{\Sigma}^{-1}$ et $a = \frac{n}{2}$ dans (1.40), et alors (1.39) devient :

$$\phi_{\mathbf{S}}(\mathbf{Z}) = \frac{1}{\Gamma_p(\frac{n}{2}) 2^{\frac{np}{2}} |\boldsymbol{\Sigma}|^{\frac{n}{2}}} \Gamma_p(\frac{n}{2}) \left| \frac{1}{2}(\mathbf{I}_p - 2i\mathbf{Z}\boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1} \right|^{-\frac{n}{2}} = |\mathbf{I}_p - 2i\mathbf{Z}\boldsymbol{\Sigma}|^{-\frac{n}{2}}. \tag{1.41}$$

□

1.3.2 Fonction génératrice des moments

Propriété 1.3.2. Soit \mathbf{S} une matrice aléatoire, $p \times p$, de loi de Wishart avec les paramètres $\boldsymbol{\Sigma}$ et $n > p - 1$. Sa fonction génératrice des moments, c'est-à-dire la fonction génératrice des moments conjointe de $S_{11}, S_{12}, \dots, S_{pp}$ est :

$$M_{\mathbf{S}}(\mathbf{Z}) = E[\exp(\text{tr}(\mathbf{Z}\mathbf{S}))] = |\mathbf{I}_p - 2\mathbf{Z}\boldsymbol{\Sigma}|^{-\frac{n}{2}} \tag{1.42}$$

où \mathbf{Z} est une matrice symétrique de dimension $p \times p$.

Démonstration. Il suffit de poser $\mathbf{T} = \frac{1}{2}(\mathbf{I}_p - 2\mathbf{Z})\boldsymbol{\Sigma}^{-1}$ et $a = \frac{n}{2}$ dans (1.38). □

1.4 Inférence via la décomposition de Bartlett

Nous traiterons dans cette section un exemple d'une matrice aléatoire de loi de Wishart via la décomposition de Bartlett. Par la suite, certaines propriétés de base empiriques seront comparées à leurs valeurs théoriques.

Définition 1.4.1. *Si \mathbf{S} est une matrice aléatoire, $p \times p$, de loi de Wishart avec les paramètres $\Sigma = \mathbf{I}_p$ et $n > p - 1$, alors celle-ci provient d'une loi de Wishart standard.*

Théorème 1.4.1. (Décomposition de Bartlett) *Soit \mathbf{S} une matrice aléatoire, $p \times p$, de loi de Wishart standard et soit $\mathbf{S} = \mathbf{T}^T \mathbf{T}$ la décomposition unique de \mathbf{S} telle que \mathbf{T} est une matrice triangulaire supérieure à éléments sur la diagonale $t_{i,i}$ positifs, pour $1 \leq i \leq p$. Alors, les composantes $t_{i,j}$, $1 \leq i < j \leq p$ de \mathbf{T} sont des variables aléatoire indépendantes telles que $t_{i,i}^2 \sim \chi_{n-i+1}^2$ et $t_{i,j} \sim N(0,1)$ pour $1 \leq i < j \leq p$.*

Démonstration. (Deheuvels (2013), p.128) □

Ce théorème permet donc d'obtenir une matrice aléatoire loi de Wishart standard de manière relativement simple.

Algorithme d'échantillonnage d'une matrice aléatoire de loi de Wishart standard

1. Assigner une dimension fixe p et un entier $n > p - 1$;
2. Obtenir une matrice triangulaire supérieure, \mathbf{T} , à éléments positifs sur la diagonale de la forme $t_{i,i} = \sqrt{\nu_i}$ tels que $\nu_i \sim \chi_{n-i+1}^2$, $i = 1, 2, \dots, p$ et éléments $t_{i,j} \sim N(0,1)$ pour $1 \leq i < j \leq p$;
3. Calculer $\mathbf{S} = \mathbf{T}^T \mathbf{T}$.

Mais qu'en advient-il d'une matrice aléatoire de Wishart non standardisée,
 $\mathbf{S} \sim W_p(n, \Sigma)$?

Théorème 1.4.2. Soit \mathbf{S} une matrice aléatoire, $p \times p$, de loi de Wishart avec les paramètres $\Sigma = \mathbf{I}$ et $n > p - 1$ et \mathbf{P} une matrice $p \times p$ non singulière telle que $\mathbf{A} = \mathbf{P}\mathbf{P}^\top$. Alors, $\mathbf{P}\mathbf{S}\mathbf{P}^\top \sim W_p(\mathbf{P}\mathbf{I}\mathbf{P}^\top = \mathbf{A}, n)$.

Démonstration. Il suffit d'imposer le changement de variable $\mathbf{Y} = \mathbf{P}\mathbf{S}\mathbf{P}^\top$ ainsi que le Jacobien $J(\mathbf{S} \rightarrow \mathbf{Y})$ respectif à la fonction de densité d'une matrice aléatoire de loi de Wishart tel que vue à la propriété 1.1.1. \square

Tel que prescrit par le théorème 1.4.2., il est préférable de considérer une décomposition de la matrice définie positive \mathbf{A} . Donc, dans le contexte présent nous allons utiliser la factorisation de Cholesky. Ceci permettra d'obtenir une matrice triangulaire inférieure, \mathbf{P} , tel que $\mathbf{A} = \mathbf{P}\mathbf{P}^\top$. Pour plus de détails à l'égard de la factorisation de Choleski, se référer à Quarteroni *et al.* (2007).

Algorithme d'échantillonnage d'une variable aléatoire de loi de Wishart non standardisée $W_p(\mathbf{P}\mathbf{I}\mathbf{P}^\top = \mathbf{A}, n)$

1. Assigner une dimension fixe p et un entier $n > p - 1$;
2. Obtenir une matrice triangulaire supérieure, \mathbf{T} , à éléments positifs sur la diagonale de la forme $t_{i,i} = \sqrt{\nu_i}$ tels que $\nu_i \sim \chi_{n-i+1}^2$, $i = 1, 2, \dots, p$ et éléments $t_{i,j} \sim N(0, 1)$ pour $1 \leq i \leq j \leq p$;
3. Calculer $\mathbf{S}_1 = \mathbf{T}^\top \mathbf{T}$;
4. Obtenir la factorisation de Cholesky de \mathbf{A} tel que $\mathbf{A} = \mathbf{P}\mathbf{P}^\top$;
5. Calculer $\mathbf{S}_2 = \mathbf{P}\mathbf{S}_1\mathbf{P}^\top$.

Exemple 1.4.1. Considérons les paramètres $n = 4$ et une matrice définie positive

de dimension 3×3

$$\Sigma = \begin{pmatrix} 1.0000 & 0.8000 & 0.2000 \\ 0.8000 & 1.0000 & 0.5000 \\ 0.2000 & 0.5000 & 1.0000 \end{pmatrix}.$$

Donc, la factorisation de Cholesky de celle-ci est :

$$\mathbf{P} = \begin{pmatrix} 1.0000 & 0.0000 & 0.0000 \\ 0.8000 & 0.6000 & 0.0000 \\ 0.2000 & 0.5667 & 0.7993 \end{pmatrix}.$$

Maintenant, prenons une matrice aléatoire triangulaire supérieure \mathbf{T} construite selon l'étape 2 de l'algorithme d'échantillonnage d'une matrice aléatoire de loi de Wishart telle que

$$\mathbf{T} = \begin{pmatrix} 4.0489 & 0.6910 & -2.0127 \\ 0.0000 & 2.1957 & -0.3321 \\ 0.0000 & 0.0000 & 2.8049 \end{pmatrix}.$$

Alors, nous obtenons une matrice aléatoire $W_3(4, \Sigma)$:

$$\mathbf{S}_2 = \mathbf{P}(\mathbf{T}^\top \mathbf{T})\mathbf{P}^\top = \begin{pmatrix} 16.3937 & 14.7937 & -1.6494 \\ 14.7937 & 15.0854 & -0.1990 \\ -1.6494 & -0.1990 & 6.1505 \end{pmatrix}.$$

Maintenant, nous allons observer la matrice d'espérance et la matrice de variance-covariances d'un échantillon aléatoire de matrices de loi $W_3(\Sigma, 4)$ de taille $n=10\,000$ et nous allons les comparer à leur valeur théorique.

Nous savons, de la section 1.4, que la matrice des espérances théoriques d'une matrice aléatoire de loi $W_p(\Sigma, n)$ est $n \times \Sigma$. Alors, pour notre exemple, nous obtenons :

$$E[\mathbf{S}_2] = 4 \times \begin{pmatrix} 1.0000 & 0.8000 & 0.2000 \\ 0.8000 & 1.0000 & 0.5000 \\ 0.2000 & 0.5000 & 1.0000 \end{pmatrix} = \begin{pmatrix} 4.0000 & 3.2000 & 0.8000 \\ 3.2000 & 4.0000 & 2.0000 \\ 0.8000 & 2.0000 & 4.0000 \end{pmatrix}$$

De plus, en se référant à la propriété 1.2.3., la matrice de variance-covariance d'un échantillon de matrices aléatoires $W_3(\Sigma, 4)$ avec Σ tel que vue au début de l'exemple est :

$$Cov[Vec(\mathbf{S}_2)] = 4 \times [\Sigma \otimes \Sigma][\mathbf{T} + \mathbf{I}_{3^2}] \quad (1.43)$$

$$= \begin{pmatrix} 8.0000 & 6.4000 & 1.6000 & 6.4000 & 5.1200 & 1.2800 & 1.6000 & 1.2800 & 0.3200 \\ 6.4000 & 6.5600 & 2.6400 & 6.5600 & 6.4000 & 2.4000 & 2.6400 & 2.4000 & 0.8000 \\ 1.6000 & 2.6400 & 4.1600 & 2.6400 & 3.2000 & 3.6000 & 4.1600 & 3.6000 & 1.6000 \\ 6.4000 & 6.5600 & 2.6400 & 6.5600 & 6.4000 & 2.4000 & 2.6400 & 2.4000 & 0.8000 \\ 5.1200 & 6.4000 & 3.2000 & 6.4000 & 8.0000 & 4.0000 & 3.2000 & 4.0000 & 2.0000 \\ 1.2800 & 2.4000 & 3.6000 & 2.4000 & 4.0000 & 5.0000 & 3.6000 & 5.0000 & 4.0000 \\ 1.6000 & 2.6400 & 4.1600 & 2.6400 & 3.2000 & 3.6000 & 4.1600 & 3.6000 & 1.6000 \\ 1.2800 & 2.4000 & 3.6000 & 2.4000 & 4.0000 & 5.0000 & 3.6000 & 5.0000 & 4.0000 \\ 0.3200 & 0.8000 & 1.6000 & 0.8000 & 2.0000 & 4.0000 & 1.6000 & 4.0000 & 8.0000 \end{pmatrix} .$$

Donc, avec les paramètres exigés, un échantillon de 10 000 matrices aléatoires de Wishart ont été produites permettant d'observer la matrice des espérances empiriques suivante :

$$\mathbf{E} = \begin{pmatrix} 4.0047 & 3.2105 & 0.8034 \\ 3.2105 & 4.0127 & 1.9994 \\ 0.8034 & 1.9994 & 4.0047 \end{pmatrix}$$

et la matrice de variance-covariance empirique (\mathbf{VC}) :

$$VC = \begin{pmatrix} 8.1695 & 6.5549 & 1.7399 & 6.5549 & 5.3084 & 1.4341 & 1.7399 & 1.4341 & 0.3437 \\ 6.5549 & 6.6841 & 2.7210 & 6.6841 & 6.5478 & 2.2510 & 2.7210 & 2.5010 & 0.8324 \\ 1.7399 & 2.7210 & 4.1497 & 2.7210 & 3.2392 & 3.5897 & 4.1497 & 3.5897 & 1.6126 \\ 6.5549 & 6.6841 & 2.7210 & 6.6841 & 6.5478 & 2.2510 & 2.7210 & 2.5010 & 0.8324 \\ 5.3084 & 6.5478 & 3.2392 & 6.5478 & 8.1295 & 4.0257 & 3.2392 & 4.0257 & 2.0088 \\ 1.4341 & 2.5010 & 3.5897 & 2.5010 & 4.0257 & 4.9429 & 3.5897 & 4.9429 & 3.9952 \\ 1.7399 & 2.7210 & 4.1497 & 2.7210 & 3.2392 & 3.5897 & 4.1497 & 3.5897 & 1.6126 \\ 1.4341 & 2.5010 & 3.5897 & 2.5010 & 4.0257 & 4.9429 & 3.5897 & 4.9429 & 3.9952 \\ 0.3437 & 0.8324 & 1.6126 & 0.8324 & 2.0088 & 3.9952 & 1.6126 & 3.9952 & 7.9385 \end{pmatrix} .$$

La convergence des paramètres estimés à leur valeur théorique est bien évidente mais il est noté qu'un échantillon de taille considérable est nécessaire à l'obtention de résultats raisonnables.

CHAPITRE II

LOI DE WISHART INVERSÉE EN LIEN AVEC LA STATISTIQUE BAYÉSIENNE

L'inférence bayésienne forme une branche de l'analyse statistique particulière en soit. Elle considère un paramètre aléatoire inconnu du modèle avec une loi de probabilité appelée loi a priori. C'est justement à l'étape du choix de cette loi que ce domaine est critiqué et cette aussi à cette étape cruciale, dans un contexte précis, que vient en jeu la loi de Wishart. Nous commençons ce chapitre par l'introduction de la loi Wishart inversé et poursuivons avec certains concepts particuliers à la statistique bayésienne. Ceux-ci permettront de bien comprendre l'implication de la loi de Wishart inversée.

2.1 Matrice aléatoire de loi de Wishart Inversée

Définition 2.1.1. Soit \mathbf{S} une matrice aléatoire, $p \times p$, de loi de Wishart avec les paramètres Σ et $n > p$. Alors son inverse, $\mathbf{U} = \mathbf{S}^{-1}$, suit une loi de Wishart inversée notée $\mathbf{U} \sim WI_p(\Phi, m)$, où Φ est une matrice définie positive, m représente le degré de liberté et p la dimension de la matrice \mathbf{U} .

2.2 Fonction de densité

Définition 2.2.1. Soit \mathbf{S} une matrice aléatoire, $p \times p$, de loi de Wishart avec les paramètres Σ et $n > p - 1$. Alors la fonction de densité de probabilité de son inverse $\mathbf{U} = \mathbf{S}^{-1}$ est :

$$f(\mathbf{U}) = \frac{|\Phi|^{\frac{m}{2}} \exp(-\frac{1}{2}\text{tr}(\Phi\mathbf{U}^{-1}))}{\Gamma_p(\frac{m}{2}) 2^{\frac{mp}{2}} |\mathbf{U}|^{\frac{m+p+1}{2}}}, m(=n) > p - 1, \mathbf{U} > 0 \quad (2.1)$$

Tel que décrit par Nydick (2012), nous pouvons comprendre que la loi de Wishart inversée prend comme paramètre une matrice de somme de carrés et génère des matrices aléatoires de variance-covariance, mais en réalité la loi de Wishart inversée prend comme paramètre l'inverse d'une matrice de variance-covariance et génère des matrices inverses de somme de carrés.

2.3 Approche bayésienne : rappel des éléments de base

La statistique Bayésienne produit une inférence sur un paramètre qui dépendra des observations et de l'information a priori disponible. Dans le contexte de la loi normale multivariée, la loi de Wishart joue un rôle important sous sa forme inverse.

Considérons un jeu de données d'une certaine population et un paramètre (θ) qui représentera une caractéristique de ce dernier. Avant la collecte de données, il y a une incertitude quant aux observations du jeu de données et du paramètre. Cependant, à la suite de l'obtention de la base de données, des informations peuvent en être tirées pour réduire l'incertitude et donc quantifier ce changement. Alors, il sera aussi possible de repérer la valeur qui représentera le mieux la vraie valeur du paramètre d'intérêt.

Quelques notions doivent être introduites à présent :

1. Y est l'ensemble de tous les jeux de données (y) possible.
2. Θ est l'ensemble de toutes les valeurs possibles du paramètre θ .
3. Pour chaque valeur de $\theta \in \Theta$, la loi a priori, notée $f(\theta)$, décrit l'incertitude sur le paramètre θ d'un modèle.
4. Pour chaque valeur de $\theta \in \Theta$ et $y \in Y$, le modèle statistique paramétrique, $f(y|\theta)$, décrit l'incertitude que y aurait été le jeu de données obtenu si la valeur de θ avait été connue.
5. À la suite de l'obtention du jeu de données (y), la valeur de θ est mise à jour. Donc, pour chaque valeur de $\theta \in \Theta$, la loi a posteriori $f(\theta|y)$ décrit l'incertitude que θ soit la vraie valeur du paramètre d'intérêt si le jeu de données est observé.

La loi a posteriori suit précisément le théorème de Bayes et celle-ci est de la forme :

$$f(\theta|y) = \frac{f(y|\theta)f(\theta)}{f(y)}. \quad (2.2)$$

6. Une famille P de lois sur θ est dite conjuguée à un modèle statistique paramétrique, $f(y|\theta)$, si pour toute loi a priori $f \in P$, la loi a posteriori $f(\theta|y) \in P$.

Revenons à la loi normale multivariée. Dans ce contexte, plusieurs variables explicatives sont mesurées pour chaque individu. Alors, il est désirable de pouvoir estimer conjointement certaines propriétés de la population, telles que l'espérance et les corrélations entre certaines variables explicatives.

Nous savons que lorsqu'il est question de la loi normale multivariée, il y a deux paramètres d'intérêt : le vecteur des espérances et la matrice de variance-covariance.

Commençons par la loi a posteriori du vecteur des espérances ($\boldsymbol{\theta}$) si on connaît la matrice de variance-covariance ($\boldsymbol{\Sigma}$).

Le problème est comme suit :

$$f(\boldsymbol{\theta}|\mathbf{y}_1, \dots, \mathbf{y}_n, \boldsymbol{\Sigma}) \propto f(\boldsymbol{\theta}|\boldsymbol{\Sigma})f(\mathbf{y}_1, \dots, \mathbf{y}_n|\boldsymbol{\theta}, \boldsymbol{\Sigma}). \quad (2.3)$$

Premièrement, le modèle statistique paramétrique est de la forme :

$$f(\mathbf{y}_1, \dots, \mathbf{y}_n|\boldsymbol{\theta}, \boldsymbol{\Sigma}) = (2\pi)^{-\frac{np}{2}}|\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\theta})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\theta}) \right\}, \quad (2.4)$$

où $\mathbf{y}_1, \dots, \mathbf{y}_n$ sont des vecteurs indépendants et identiquement distribués selon une loi normale multivariée $N(\boldsymbol{\theta}, \boldsymbol{\Sigma})$. De plus, la formule (2.5) peut être considérée de la façon suivante :

$$\begin{aligned} f(\mathbf{y}_1, \dots, \mathbf{y}_n|\boldsymbol{\theta}, \boldsymbol{\Sigma}) &\propto \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\theta})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\theta}) \right\} \\ &= \exp \left\{ -\frac{1}{2} \left(\sum_{i=1}^n \mathbf{y}_i^\top \boldsymbol{\Sigma}^{-1} \mathbf{y}_i - 2 \sum_{i=1}^n \boldsymbol{\theta}^\top \boldsymbol{\Sigma}^{-1} \mathbf{y}_i + \sum_{i=1}^n \boldsymbol{\theta}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\theta} \right) \right\}, \\ &\propto \exp \left(\sum_{i=1}^n \boldsymbol{\theta}^\top \boldsymbol{\Sigma}^{-1} \mathbf{y}_i - \frac{n}{2} \boldsymbol{\theta}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\theta} \right) \\ &= \exp \left(\boldsymbol{\theta}^\top n \boldsymbol{\Sigma}^{-1} \bar{\mathbf{y}} - \frac{n}{2} \boldsymbol{\theta}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\theta} \right) \end{aligned} \quad (2.5)$$

où \propto représente le symbole de proportionalité.

Maintenant, jetons un coup d'oeil sur la loi a priori (conditionnelle), $f(\boldsymbol{\theta}|\boldsymbol{\Sigma})$. Afin de simplifier les choses, nous voulons trouver une loi a priori conjuguée au modèle statistique. C'est-à-dire, nous désirons que les lois a priori et a posteriori proviennent de la même famille. De plus, nous savons que pour n'importe quelle loi a priori conditionnelle $f(\boldsymbol{\theta}|\boldsymbol{\Sigma})$, la loi a posteriori sera de la forme :

$$f(\boldsymbol{\theta}|\mathbf{y}_1, \dots, \mathbf{y}_n, \boldsymbol{\Sigma}) \propto f(\boldsymbol{\theta}|\boldsymbol{\Sigma}) \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\theta})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \boldsymbol{\theta}) \right\} \quad (2.6)$$

Alors, si $f(\boldsymbol{\theta}|\boldsymbol{\Sigma})$ est conjuguée, elle doit être de la forme $\exp\{-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\sum_{i=1}^n(\mathbf{y}_i - \boldsymbol{\theta})^\top(\mathbf{y}_i - \boldsymbol{\theta})\}$ et ceci peut-être facilement obtenu par la famille de loi normale multivariée. Donc, si $f(\boldsymbol{\theta}|\boldsymbol{\Sigma})$ provient d'une loi normale et que nous savons que $\mathbf{y}_1, \dots, \mathbf{y}_n$ sont i.i.d $N(\boldsymbol{\theta}, \boldsymbol{\Sigma})$, alors $f(\boldsymbol{\theta}|\mathbf{y}_1, \dots, \mathbf{y}_n, \boldsymbol{\Sigma})$ aura aussi une loi normale. Posons $f(\boldsymbol{\theta}) \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Lambda}_0)$.

Clairement $f(\boldsymbol{\theta}|\boldsymbol{\Sigma}) = f(\boldsymbol{\theta})$ et

$$f(\boldsymbol{\theta}) = (2\pi)^{-\frac{p}{2}}|\boldsymbol{\Lambda}_0|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\mu}_0)^\top \boldsymbol{\Lambda}_0^{-1}(\boldsymbol{\theta} - \boldsymbol{\mu}_0)\right\}. \quad (2.7)$$

On développe et on obtient :

$$\begin{aligned} f(\boldsymbol{\theta}) &\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\mu}_0)^\top \boldsymbol{\Lambda}_0^{-1}(\boldsymbol{\theta} - \boldsymbol{\mu}_0)\right\} \\ &= \exp\left\{-\frac{1}{2}(\boldsymbol{\theta}^\top \boldsymbol{\Lambda}_0^{-1} \boldsymbol{\theta} - 2\boldsymbol{\theta}^\top \boldsymbol{\Lambda}_0^{-1} \boldsymbol{\mu}_0 + \boldsymbol{\mu}_0^\top \boldsymbol{\Lambda}_0^{-1} \boldsymbol{\mu}_0)\right\} \\ &\propto \exp(\boldsymbol{\theta}^\top \boldsymbol{\Lambda}_0^{-1} \boldsymbol{\mu}_0 - \frac{1}{2}\boldsymbol{\theta}^\top \boldsymbol{\Lambda}_0^{-1} \boldsymbol{\theta}). \end{aligned} \quad (2.8)$$

Maintenant, mettons les morceaux ensemble en commençant par (2.4) :

$$\begin{aligned} f(\boldsymbol{\theta}|\mathbf{y}_1, \dots, \mathbf{y}_n, \boldsymbol{\Sigma}) &\propto f(\boldsymbol{\theta}|\boldsymbol{\Sigma})f(\mathbf{y}_1, \dots, \mathbf{y}_n|\boldsymbol{\theta}, \boldsymbol{\Sigma}) \\ &= f(\boldsymbol{\theta})f(\mathbf{y}_1, \dots, \mathbf{y}_n|\boldsymbol{\theta}, \boldsymbol{\Sigma}) \\ &\propto \exp(\boldsymbol{\theta}^\top \boldsymbol{\Lambda}_0^{-1} \boldsymbol{\mu}_0 - \frac{1}{2}\boldsymbol{\theta}^\top \boldsymbol{\Lambda}_0^{-1} \boldsymbol{\theta}) \exp(\boldsymbol{\theta}^\top n\boldsymbol{\Sigma}^{-1}\bar{\mathbf{y}} - \frac{n}{2}\boldsymbol{\theta}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\theta}) \\ &\propto \exp(\boldsymbol{\theta}^\top \mathbf{a} - \frac{1}{2}\boldsymbol{\theta}^\top \mathbf{B} \boldsymbol{\theta}), \end{aligned} \quad (2.9)$$

où $\mathbf{a} = \boldsymbol{\Lambda}_0^{-1} \boldsymbol{\mu}_0 + n\boldsymbol{\Sigma}^{-1}\bar{\mathbf{y}}$ et $\mathbf{B} = \boldsymbol{\Lambda}_0^{-1} + n\boldsymbol{\Sigma}^{-1}$.

Ceci implique alors que la loi a posteriori de $\boldsymbol{\theta}$ suit une loi normale multivariée avec une matrice de variance-covariance égale à \mathbf{B}^{-1} et une espérance de $\mathbf{B}^{-1}\mathbf{a}$. Donc, la matrice de précision a posteriori, c'est-à-dire l'inverse de la matrice de

variance-covariance $((\mathbf{B}^{-1})^{-1})$, est en fait la somme de la matrice de précision a priori (Λ_0^{-1}) et la matrice de précision du modèle statistique $(n\Sigma^{-1})$.

La prochaine étape est la loi a posteriori de la matrice de variance-covariance (Σ) si on connaît le vecteur des espérances (θ) .

Le problème est comme suit :

$$f(\Sigma|\mathbf{y}_1, \dots, \mathbf{y}_n, \theta) = f(\Sigma|\theta)f(\mathbf{y}_1, \dots, \mathbf{y}_n|\theta, \Sigma). \quad (2.10)$$

Le modèle statistique paramétrique est tel que vu précédemment, mais sera écrit sous la forme suivante :

$$\begin{aligned} f(\mathbf{y}_1, \dots, \mathbf{y}_n|\theta, \Sigma) &= (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(\Sigma^{-1}\mathbf{S}) \right\} \\ &\propto |\Sigma|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(\Sigma^{-1}\mathbf{S}) \right\}, \end{aligned} \quad (2.11)$$

où $\mathbf{y}_1, \dots, \mathbf{y}_n$ sont des vecteurs indépendants et identiquement distribués selon une loi normale multivariée $N(\theta, \Sigma)$ et $\mathbf{S} = \sum_{i=1}^n (\mathbf{y}_i - \theta)(\mathbf{y}_i - \theta)^\top$.

Maintenant, considérons la loi a priori (conditionnelle), $f(\Sigma|\theta)$. Nous savons que la matrice de variance-covariance doit être symétrique et, de plus, définie positive. Il est donc nécessaire que la loi a priori considère ces deux choses. C'est ici que la loi de Wishart vient en jeu. Nous savons que cette dernière modélise la loi de matrices de somme de carrés avec les propriétés nécessaires. C'est-à-dire, pour $\mathbf{x}_1, \dots, \mathbf{x}_{n_0}$ indépendants et identiquement distribués selon une loi normale multivariée $N(\mathbf{0}, \Sigma_0)$ et $n_0 > p_0$, $\mathbf{X}^\top \mathbf{X}$ est une matrice symétrique définie positive. Heureusement, la loi de Wishart inversée est une loi a priori conjuguée pour la matrice des variance-covariance et la loi de Wishart est une loi a priori conjuguée pour la matrice de précision $(\Sigma^{-1})!$

Donc, on note $f(\Sigma|\boldsymbol{\theta}) = f(\Sigma)$ puisque la loi a priori de Σ est indépendante de $\boldsymbol{\theta}$ et Σ suit une loi $WI_{p_0}(\mathbf{S}_0, n_0)$.

De plus,

$$f(\Sigma) \propto |\Sigma|^{-(n_0+p+1)/2} \exp \left\{ -\frac{1}{2} \text{tr}(\Sigma^{-1} \mathbf{S}_0) \right\}. \quad (2.12)$$

Combinons (2.11), (2.12) et (2.13) :

$$\begin{aligned} f(\Sigma|\mathbf{y}_1, \dots, \mathbf{y}_n, \boldsymbol{\theta}) &\propto f(\Sigma|\boldsymbol{\theta}) f(\mathbf{y}_1, \dots, \mathbf{y}_n|\boldsymbol{\theta}, \Sigma) \\ &\propto |\Sigma|^{-(n_0+p+1)/2} \exp \left\{ -\frac{1}{2} \text{tr}(\Sigma^{-1} \mathbf{S}_0) \right\} |\Sigma|^{-\frac{n}{2}} \\ &\times \exp \left\{ -\frac{1}{2} \text{tr}(\Sigma^{-1} \mathbf{S}) \right\} \\ &\propto |\Sigma|^{-(n_0+n+p+1)/2} \exp \left\{ -\frac{1}{2} \text{tr}((\mathbf{S}_0 + \mathbf{S}) \Sigma^{-1}) \right\}. \end{aligned} \quad (2.13)$$

Ceci implique que la loi a posteriori suit une loi de Wishart inversée avec une matrice de sommes des carrés égale à la somme de la matrice de sommes des carrés a priori, \mathbf{S}_0 , et la matrice de sommes des carrés du modèle statistique, \mathbf{S} . On observe aussi que le nombre de degrés de liberté (taille de l'échantillon a posteriori) est la somme de la taille de l'échantillon a priori et la taille de l'échantillon de données.

CHAPITRE III

LIENS AVEC D'AUTRES LOIS DE PROBABILITÉS

Dans ce chapitre, nous traiterons de la relation entre la loi de Wishart et quatre lois de probabilités. Nous débuterons par la loi associée à celle-ci de manière naturelle, c'est-à-dire dans le cas univarié. Par la suite, certaines lois multivariées intéressantes seront étudiées en terminant par la loi bêta multivariée. Celle-ci, en relation à la loi de Fisher, joue un rôle important lors de l'analyse de régression multidimensionnelle.

3.1 Loi de khi-deux, χ^2

La loi de khi-deux représente la somme des carrés de n vecteurs aléatoires indépendants de loi normale univariée. Parallèlement, la loi de Wishart représente la somme des carrés de n matrices aléatoires indépendantes de loi normale multivariée. Donc l'on constate que la loi de Wishart est équivalente à la loi de khi-deux dans le contexte multivarié.

Naturellement, si $\mathbf{S} \sim W_1(\Sigma, n)$ où $\Sigma = \sigma^2 = 1$ nous obtenons une loi de khi-deux avec n degrés de liberté (matrice \mathbf{X} est donc un vecteur) donc

$$\mathbf{S} = \mathbf{X}^\top \mathbf{X} = \sum_{i=1}^n x_i^2 \sim \chi_n^2 \quad (3.1)$$

De plus, il a été noté que la variance des éléments d'une matrice de Wishart a la forme suivante $Var[S_{ij}] = n(\sigma_{ii}\sigma_{jj} + \sigma_{ij}^2)$. Par conséquent, si $\mathbf{S} \sim W_1(\Sigma, n)$ où $\Sigma = \sigma^2 = 1$ alors $Var[\mathbf{S}] = n(1 \times 1 + (1)^2) = 2n$ et ceci représente évidemment la variance d'une variable aléatoire de loi khi-deux.

Propriété 3.1.1. *Soit $\mathbf{S} \sim W_p(\Sigma, n)$ et \mathbf{x} un vecteur aléatoire, $p \times 1$, non nul indépendant de \mathbf{S} . Alors,*

$$\frac{\mathbf{x}^\top \mathbf{S} \mathbf{x}}{\mathbf{x}^\top \Sigma \mathbf{x}} \sim \chi_n^2. \quad (3.2)$$

Démonstration. (Gupta et Nagar (2000), p.96) □

On constate donc par la propriété 3.1.1. que :

$$\mathbf{x}^\top \mathbf{S} \mathbf{x} \sim \mathbf{x}^\top \Sigma \mathbf{x} \times \chi_n^2. \quad (3.3)$$

où $\mathbf{x}^\top \Sigma \mathbf{x}$ peut être considéré en tant que paramètre d'échelle de la loi χ_n^2 .

3.2 Loi du T^2 de Hotelling

La loi du T^2 de Hotelling est en fait une généralisation de la loi de Student au carré dans le contexte de l'analyse multidimensionnelle. Elle se rencontre lors du test sur l'égalité des moyennes de deux populations.

Définition 3.2.1. *Soit \mathbf{x} un vecteur aléatoire, telle que $\mathbf{x} \sim N_p(\mathbf{0}, \Sigma)$ et \mathbf{S} une matrice de Wishart indépendante de \mathbf{x} telle que $\mathbf{S} \sim W_p(\Sigma, n)$. Alors, la loi T^2 de Hotelling peut s'obtenir comme suit :*

$$\mathbf{Y} = n\mathbf{x}^\top \mathbf{S}^{-1} \mathbf{x} \sim T_{p,n}^2. \quad (3.4)$$

Remarque 3.2.1 Si $p=1$, $T_{1,n}^2 = F(1, n)$ est une variable de Student au carré avec n degrés de liberté.

Propriété 3.2.1. Si \mathbf{Y} est une matrice aléatoire telle que $\mathbf{Y} \sim T_{p,n}^2$ alors

$$\frac{n-p+1}{pn} \mathbf{Y} \sim F(p, n-p+1). \quad (3.5)$$

où F représente la loi de Fisher.

Démonstration. Puisque

$$\mathbf{Y} = n\mathbf{x}^\top \mathbf{S}^{-1} \mathbf{x} \sim T_{p,n}^2, \quad (3.6)$$

ceci implique que

$$\frac{\mathbf{Y}}{n} = \mathbf{x}^\top \mathbf{S}^{-1} \mathbf{x}, \quad (3.7)$$

et donc

$$\mathbf{x}^\top \mathbf{S}^{-1} \mathbf{x} = \mathbf{x}^\top \mathbf{S}^{-1} \mathbf{x} \left(\frac{\mathbf{x}^\top \Sigma^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{S}^{-1} \mathbf{x}} \right) = \mathbf{x}^\top \Sigma^{-1} \mathbf{x} \left(\frac{\mathbf{x}^\top \Sigma^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{S}^{-1} \mathbf{x}} \right)^{-1}. \quad (3.8)$$

Lemme 3.2.1. Soit \mathbf{S} une matrice aléatoire de loi de Wishart de paramètres Σ et $n > p-1$ et \mathbf{x} un vecteur aléatoire $p \times 1$ indépendant de \mathbf{S} , alors

$$\left(\frac{\mathbf{x}^\top \Sigma^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{S}^{-1} \mathbf{x}} \right) \sim \chi_{n-p+1}^2 \quad (3.9)$$

Démonstration. (Eaton (2007b), p.313) □

Ensuite, si nous considérons $\Sigma = \mathbf{I}_p$ et \mathbf{x} un vecteur $p \times 1$, (3.7) devient :

$$\mathbf{x}^\top \mathbf{S}^{-1} \mathbf{x} = \mathbf{x}^\top \mathbf{I}_p^{-1} \mathbf{x} \left(\frac{\mathbf{x}^\top \mathbf{I}_p^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{S}^{-1} \mathbf{x}} \right)^{-1} = \frac{\chi_p^2}{\chi_{n-p+1}^2} \quad (3.10)$$

Finalement, multiplions (3.6) par $\frac{1}{n-p+1}$

$$\frac{\mathbf{Y}}{n} \times \frac{\frac{1}{p}}{\frac{1}{n-p+1}} = \mathbf{x}^\top \mathbf{S}^{-1} \mathbf{x} \times \frac{\frac{1}{p}}{\frac{1}{n-p+1}} = \frac{\chi_p^2}{\chi_{n-p+1}^2} \times \frac{\frac{1}{p}}{\frac{1}{n-p+1}} = \frac{\frac{\chi_p^2}{p}}{\frac{\chi_{n-p+1}^2}{n-p+1}} \quad (3.11)$$

□

3.3 Loi du lambda (Λ) de Wilks (généralisation de la loi de de Fisher)

La loi du lambda de Wilks s'associe à la validation d'hypothèses de l'Analyse de Variance multivariée (MANOVA) en traitant directement sur les matrices résiduelles.

Définition 3.3.1. Si $\mathbf{A} \sim W_p(\Sigma, m)$ et $\mathbf{B} \sim W_p(\Sigma, N)$ indépendantes et $m, n \geq p$ ($|\mathbf{A}| \neq 0$), alors la loi du lambda de Wilks peut s'obtenir comme suit :

$$\mathbf{U} = \frac{|\mathbf{A}|}{|\mathbf{A} + \mathbf{B}|} = \frac{1}{|\mathbf{A}^{-1}\mathbf{B} + \mathbf{I}|} = \prod_{i=1}^p (1 + \theta_i)^{-1} \sim \Lambda(p, m, n) \quad (3.12)$$

où θ_i sont les valeurs propres de $\mathbf{A}^{-1}\mathbf{B}$.

Ceci est en lien avec la loi de Fisher dans quatre circonstances :

1. Quand le paramètre $p = 1$. C'est-à-dire, que les matrices \mathbf{A} et \mathbf{B} suivent une loi khi-deux avec m et n degrés de liberté respectivement. Alors,

$$\frac{1 - \Lambda(1, m, n)}{\Lambda(1, m, n)} = \frac{1 - \frac{|\mathbf{A}|}{|\mathbf{A} + \mathbf{B}|}}{\frac{|\mathbf{A}|}{|\mathbf{A} + \mathbf{B}|}} = \frac{|\mathbf{A} + \mathbf{B}| - |\mathbf{A}|}{|\mathbf{A} + \mathbf{B}|} \times \frac{|\mathbf{A} + \mathbf{B}|}{|\mathbf{A}|} = \frac{|\mathbf{A} + \mathbf{B}| - |\mathbf{A}|}{|\mathbf{A}|} \quad (3.13)$$

Rappelons-nous que si $p = 1$, nous avons $\mathbf{A} = \mathbf{X}^\top \mathbf{X} = \sum_{i=1}^m x_i^2$ et $\mathbf{B} = \mathbf{Y}^\top \mathbf{Y} = \sum_{i=1}^n y_i^2$. Alors, (3.12) devient :

$$\frac{1 - \Lambda(1, m, n)}{\Lambda(1, m, n)} = \frac{|\mathbf{A} + \mathbf{B}| - |\mathbf{A}|}{|\mathbf{A}|} = \frac{\sum_{i=1}^n y_i^2}{\sum_{i=1}^m x_i^2} \sim \frac{\chi_n^2}{\chi_m^2} = \frac{n}{m} F(n, m) \quad (3.14)$$

c'est-à-dire,

$$\mathbf{U} = \frac{\chi_m^2}{\chi_m^2 + \chi_n^2} = \frac{1}{1 + \frac{\chi_n^2}{\chi_m^2}} = \frac{1}{1 + \frac{n}{m} F(n, m)} \quad (3.15)$$

donc,

$$\frac{n}{m} F(n, m) = \frac{1 - U}{U}. \quad (3.16)$$

De plus, il est à noter que la relation entre la distribution de la loi Fisher et la distribution de la loi bêta comme suit : si $\mathbf{X} \sim F(d_1, d_2)$ alors

$\frac{\frac{d_1}{d_2} \mathbf{X}}{1 + \frac{d_1}{d_2} \mathbf{X}} \sim \text{Beta}(\frac{d_1}{2}, \frac{d_2}{2})$. Celle-ci nous permet d'observer que

$$\frac{\frac{d_1}{d_2} \mathbf{X}}{1 + \frac{d_1}{d_2} \mathbf{X}} = \frac{\frac{d_1}{d_2} \frac{d_2}{d_1} \frac{\chi_{d_1}^2}{\chi_{d_2}^2}}{1 + \frac{d_1}{d_2} \frac{d_2}{d_1} \frac{\chi_{d_1}^2}{\chi_{d_2}^2}} = \frac{\frac{\chi_{d_1}^2}{\chi_{d_2}^2}}{1 + \frac{\chi_{d_1}^2}{\chi_{d_2}^2}} = \frac{\chi_{d_1}^2}{\chi_{d_1}^2 + \chi_{d_2}^2} \quad (3.17)$$

qui est similaire à la formule (3.14) et nous permet de conclure que lorsque $p = 1$ alors $U \sim \text{Beta}(\frac{m}{2}, \frac{n}{2})$.

2. Quand le paramètre $n = 1$ (\mathbf{Y} est un vecteur $1 \times p$). Nous savons que,

$$U = \frac{|\mathbf{A}|}{|\mathbf{A} + \mathbf{B}|} = |\mathbf{A}^{-1}\mathbf{B} + \mathbf{I}|^{-1} = |\mathbf{A}^{-1}\mathbf{Y}^\top \mathbf{Y} + \mathbf{I}|^{-1} \quad (3.18)$$

Considérons $\mathbf{C} = \mathbf{A}^{-1}\mathbf{Y}^\top$ et $\mathbf{D} = \mathbf{Y}$. Il est clair que \mathbf{C} est de dimension $p \times 1$ puisque \mathbf{A} est $p \times p$ et \mathbf{Y}^\top est $p \times 1$. Nous introduisons maintenant la lemme 3.3.1.

Lemme 3.3.1. *Soit $\mathbf{C}_{p \times n}$ et $\mathbf{D}_{n \times p}$ des matrices complexes. Alors,*

$$|\mathbf{CD} + \mathbf{I}_p| = |\mathbf{DC} + \mathbf{I}_n| \quad (3.19)$$

Démonstration. (Eaton (2007a), p.43) □

Nous obtenons

$$U = \frac{|\mathbf{A}|}{|\mathbf{A} + \mathbf{B}|} = |\mathbf{A}^{-1}\mathbf{B} + \mathbf{I}|^{-1} = |\mathbf{A}^{-1}\mathbf{Y}^\top \mathbf{Y} + \mathbf{I}|^{-1} = (\mathbf{Y}\mathbf{A}^{-1}\mathbf{Y}^\top + 1)^{-1} \quad (3.20)$$

Maintenant, trouvons la distribution de $\mathbf{YA}^{-1}\mathbf{Y}^\top$. On observe que

$$\mathbf{YA}^{-1}\mathbf{Y}^\top = \frac{\mathbf{YA}^{-1}\mathbf{Y}^\top}{\mathbf{YY}^\top} \mathbf{YY}^\top. \quad (3.21)$$

Donc, sous les hypothèses du théorème 3.2.1, nous savons que $\frac{\mathbf{YY}^\top}{\mathbf{YA}^{-1}\mathbf{Y}^\top}$ suit une loi χ_{m-p+1}^2 et $\mathbf{YY}^\top \sim \chi_p^2$. Alors, il s'en suit que :

$$\mathbf{U} = (1 + \mathbf{YA}^{-1}\mathbf{Y}^\top)^{-1} = \frac{1}{1 + \frac{\chi_p^2}{\chi_{m-p+1}^2}}. \quad (3.22)$$

Par conséquent, $\mathbf{U} \sim \text{Beta}(\frac{m-p+1}{2}, \frac{p}{2})$.

Les démonstrations des 2 prochaines circonstances sont présentées au chapitre 2 de Seber (1984) :

3. Quand le paramètre $p = 2$. Nous avons

$$\frac{1 - \sqrt{\Lambda(2, m, n)}}{\sqrt{\Lambda(2, m, n)}} = \frac{n}{m-1} F(2n, 2(m-1)) \quad (3.23)$$

4. Quand le paramètre $n = 2$. Nous avons

$$\frac{1 - \sqrt{\Lambda(p, m, 2)}}{\sqrt{\Lambda(p, m, 2)}} = \frac{p}{m-p+1} F(2p, 2(m-p+1)) \quad (3.24)$$

3.4 Loi bêta multivariée

Dans un cours de régression classique, il est toujours enseigné le test de l'importance globale d'un modèle. Celui-ci considère la possibilité qu'aucune variable exogène contribue à expliquer la variable dépendante et se caractérise par le test d'hypothèse suivant :

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

vs

$$(3.25)$$

H_1 : au moins un des coefficients n'est pas égal à 0

Pour répondre à cette question, on se réfère à la somme des carrés du modèle de régression et la somme de carré des résidus ainsi qu'à la statistique du test sous l'hypothèse nulle. Nous avons alors :

$$G = \frac{\frac{SS_{reg}}{p}}{\frac{SS_{res}}{n-p-1}} \sim F(p, n-p-1), \quad (3.26)$$

où F indique la loi de Fisher.

Cette forme est équivalente à :

$$G = \frac{SS_{reg}}{SS_{res}} \sim \frac{p}{n-p-1} F(p, n-p-1). \quad (3.27)$$

Alors, la loi G possède une fonction de densité de probabilité :

$$g(x) = \frac{1}{Beta(p/2, n-p-1/2)} \frac{x^{\frac{p}{2}-1}}{(1+x)^{(n-p-1+p)/2}}, 0 \leq x < \infty \quad (3.28)$$

où

$$Beta(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}. \quad (3.29)$$

Considérons maintenant la fonction de densité de probabilité de la variable $G^* = \frac{G}{1+G}$ ($= \frac{SS_{reg}}{SS_{reg}+SS_{res}}$). Celle-ci est de la forme :

$$g(y) = \frac{1}{Beta(p/2, n-p-1/2)} y^{\frac{p}{2}-1} (1-y)^{\frac{(n-p-1)}{2}-1}, 0 \leq y < 1 \quad (3.30)$$

pour laquelle nous reconnaissons bien la fonction de densité d'une variable aléatoire de loi bêta, $Beta(p/2, n-p-1/2)$.

Dans le cas multivarié, on peut penser aux matrices indépendantes non-singulières $S(SS_{reg})$ et $U(SS_{res})$ telles que $S \sim W_p(\Sigma, n_S)$ et $U \sim W_p(\Sigma, n_U)$. Donc, de manière similaire au cas univarié, nous pouvons considérer $G = SU^{-1}$ et $G^* = S(S+U)^{-1}$. Notons que G et G^* ne sont plus des matrices symétriques. et que S

et \mathbf{U} sont des matrices définies positives. Alors, $\mathbf{S} + \mathbf{U}$ est aussi une matrice définie positive avec probabilité 1. Nous pouvons transformer celles-ci pour obtenir des matrices symétriques :

$$\mathbf{G} = \mathbf{U}^{-\frac{1}{2}}\mathbf{S}\mathbf{U}^{-\frac{1}{2}}, \quad (3.31)$$

et

$$\mathbf{G}^* = (\mathbf{S} + \mathbf{U})^{-\frac{1}{2}}\mathbf{S}(\mathbf{S} + \mathbf{U})^{-\frac{1}{2}}. \quad (3.32)$$

Nous obtenons alors la densité de probabilité conjointe des $\frac{1}{2}p(p+1)$ éléments de la matrice \mathbf{G}^* qui maintenant est associé à loi bêta multivariée, $\mathbf{B}(n_{\mathbf{S}}/2, n_{\mathbf{U}}/2)$ de densité :

$$f(\mathbf{Y}) = \frac{1}{\mathbf{B}_p(n_{\mathbf{S}}/2, n_{\mathbf{U}}/2)} |\mathbf{Y}|^{\frac{(n_{\mathbf{S}}-p-1)}{2}} |\mathbf{I}_p - \mathbf{Y}|^{\frac{(n_{\mathbf{U}}-p-1)}{2}}, \mathbf{0} < \mathbf{Y} < \mathbf{I}_p. \quad (3.33)$$

CHAPITRE IV

PROCESSUS DE WISHART

4.1 Processus stochastique

Dans cette section, les notions de base d'un processus stochastique présentées seront tirées de E. Shreve (2003). Ces processus sont au coeur de la mathématique financière et permettront d'introduire le processus de Wishart. Il est à noter que l'objectif n'est pas de faire une analyse en profondeur du calcul stochastique, mais d'avoir les outils nécessaires pour étudier la construction d'un processus de Wishart. Bien que ceux-ci soient énoncés dans le cas univariée, ils se traduisent naturellement dans le cas multidimensionnel. Aucune preuve ne sera présentée mais les hypothèses telles que précisées dans cette section seront supposées vérifiées dans les sections subséquentes.

Définition 4.1.1. *Soit \mathbb{R} , l'ensemble des nombres réels. On appelle un processus stochastique à temps continu une famille de variables aléatoires $\{X_t, t \in \mathbb{R}\}$.*

Définition 4.1.2. *Soit un espace mesurable (Ω, \mathcal{F}) . On dit que $(\mathcal{F}_t)_{t \geq 0}$ est une filtration s'il s'agit d'une famille croissante de sous-tribus de \mathcal{F} . C'est-à-dire, $\mathcal{F}_t \subseteq \mathcal{F}$ et $\mathcal{F}_s \subseteq \mathcal{F}_t$, si $s \leq t$.*

Définition 4.1.3. *Un processus $(X_t)_{t \geq 0}$ est dit adapté à la filtration $(\mathcal{F}_t)_{t \geq 0}$ si pour chaque t , X_t est \mathcal{F}_t -mesurable.*

Définition 4.1.4. Une fonction $t \mapsto X_t(\omega)$ est appelée une trajectoire de X pour chaque $\omega \in \Omega$ fixé.

Définition 4.1.5. Un mouvement brownien standard est un processus stochastique $(B_t)_{t \geq 0}$, démarrant à zéro ($B_0 = 0$ p.s.) dont les trajectoires sont continues, les accroissements disjoints sont indépendants et pour tout $s > 0$ $(B_{t+s} - B_t) \sim N(0, s)$.

Définition 4.1.6. Soit $(B_t)_{t \geq 0}$, un mouvement brownien standard. Pour $t > 0$, la variation quadratique de celui-ci est :

$$\langle B \rangle_t = \sum_{i=1}^{2^n} \left\{ B\left(\frac{it}{2^n}\right) - B\left(\frac{(i-1)t}{2^n}\right) \right\}^2 \quad (4.1)$$

De plus, on peut montrer que le mouvement brownien est un processus à variation quadratique fini tel que pour $t \geq 0$,

$$\langle B \rangle_t = t \quad (4.2)$$

Définition 4.1.7. Soit $(B_t)_{t \geq 0}$, un mouvement brownien standard. Le processus stochastique conditionnel $(Z_t)_{0 \leq t \leq 1}$ où

$$Z_t = B_t - tB_1 \quad (4.3)$$

est appelé pont brownien.

Définition 4.1.8. Un processus $H = (H_t)_{0 \leq t \leq T}$ est un processus élémentaire si celui-ci est de la forme :

$$H_t(\omega) = \sum_{j=0}^{n-1} e_j(\omega) I_{[t_j, t_{j+1})}(t) \quad (4.4)$$

où t_0, \dots, t_n est une partition de $[0, T]$ et pour chaque j , e_j est une variable aléatoire bornée et \mathcal{F}_{t_j} -mesurable.

Définition 4.1.9. Si H est un processus élémentaire, alors l'intégrale stochastique sur $[0, T]$ de celui-ci est un processus stochastique défini par :

$$\int_0^T H_s dB_s = \sum_{j=0}^{n-1} e_j (B_{t_{j+1}} - B_{t_j}) \quad (4.5)$$

Définition 4.1.10. Un processus X_t est un processus d'Itô s'il existe X_0 , μ_s un processus intégrable et σ_s un processus adapté à la filtration (\mathcal{F}_t) tels que :

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s \quad (4.6)$$

Théorème 4.1.1. (Formule d'Itô) Soit u , une fonction continûment différentiable par rapport à t et deux fois continûment différentiable par rapport à x et X un processus d'Itô de la forme donnée à la définition (5.1.5). Posons $Y_t = u(t, X_t)$. Alors pour tout t , on a :

$$dY_t = \frac{\partial u}{\partial t}(t, X_t)dt + \frac{\partial u}{\partial x}(t, X_t)[\mu_t dt + \sigma_t dB_t] + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(t, X_t) \sigma_t^2 dt \quad (4.7)$$

Proposition 4.1.1. (Intégration par parties) Soit X et Y deux processus d'Itô de la forme

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s \quad (4.8)$$

et

$$Y_t = Y_0 + \int_0^t \mu'_s ds + \int_0^t \sigma'_s dB_s \quad (4.9)$$

alors

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + d\langle X, Y \rangle_t \quad (4.10)$$

où $d\langle X, Y \rangle_t$ représente la covariation entre le processus X et Y définie par :

$$d\langle X, Y \rangle_t = (\sigma_t \sigma'_t) dt \quad (4.11)$$

4.2 Processus de Wishart

Nous présentons maintenant le processus de Wishart tel que Bru (1991) l'a introduit. Plusieurs notions énoncées dans la section précédente seront maintenant transposées sous leurs formes multidimensionnelles. Pour la suite, nous supposons que les matrices sont à valeurs réelles et les paramètres n et p sont entiers.

Définition 4.2.1. Soit \mathbf{N}_t une matrice $n \times p$ ($n \geq p$) dont les lignes sont des mouvements browniens indépendants et proviennent d'une distribution multidimensionnelle plus grand ou égale à 1 et dont les valeurs au temps initial sont $\mathbf{N}_0 = \mathbf{C}$. Alors, un processus de Wishart est un processus matriciel écrit $WIS(n, p, \mathbf{s}_0)$ de la forme :

$$\mathbf{S}_t = \mathbf{N}_t^\top \mathbf{N}_t, \quad \mathbf{s}_0 = \mathbf{C}^\top \mathbf{C} \quad (4.12)$$

Donc, on constate que si le temps est fixe, cela revient à l'étude d'une matrice aléatoire, $\mathbf{S}_{p \times p}$, distribuée selon une loi de Wishart. De plus, nous observons que l'équation différentielle associé à (4.9) est :

$$d\mathbf{S}_t = d\mathbf{N}_t^\top \mathbf{N}_t + \mathbf{N}_t^\top d\mathbf{N}_t + n\mathbf{I}_p dt \quad (4.13)$$

en utilisant la propriété d'intégration par parties présentée à la proposition 4.1.1. dans le cas multidimensionnel. Celle-ci se traduit comme suit : soit \mathbf{X} et \mathbf{Z} , deux processus d'Itô, tel que

$$\begin{aligned} d\mathbf{X}_t &= \mathbf{X}_t \mathbf{A}_X dt + d\mathbf{B}_t \mathbf{Q}_X, & \mathbf{X}_0 &= \mathbf{x}_0 \\ d\mathbf{Z}_t &= \mathbf{Z}_t \mathbf{A}_Z dt + d\mathbf{B}_t \mathbf{Q}_Z, & \mathbf{Z}_0 &= \mathbf{z}_0 \end{aligned} \quad (4.14)$$

où \mathbf{B}_t est un mouvement brownien multidimensionnel $n \times p$ ($n \geq p$) et \mathbf{A} et \mathbf{Q} sont des matrices $p \times p$. Alors,

$$d\langle \mathbf{X}, \mathbf{Z} \rangle_t = d\mathbf{X}_t \mathbf{Z}_t + \mathbf{X}_t d\mathbf{Z}_t + d\langle \mathbf{X}, \mathbf{Z} \rangle_t, \quad t \geq 0 \quad (4.15)$$

où $\langle \mathbf{X}, \mathbf{Z} \rangle_t$ représente la covariation des processus \mathbf{X} et \mathbf{Z} . Dans le cas multidimensionnel, nous avons :

$$d\langle \mathbf{X}, \mathbf{Z} \rangle_t = \mathbf{Q}_X^\top d\langle \mathbf{B}_t^\top, \mathbf{B}_t \rangle_t \mathbf{Q}_Z \quad (4.16)$$

De plus,

$$d\langle \mathbf{B}_t^\top, \mathbf{B}_t \rangle_{(ij)} = \sum_{j=1}^n d\langle B_{(ji),t}, B_{(jk),t} \rangle_t = n \mathbf{1}_{\{i=k\}} dt \quad (4.17)$$

puisque nous savons, par la remarque de la définition 4.1.5, que $\langle B_{(ab),t}, B_{(cd),t} \rangle_t = t$ si et seulement si $a = c$ et $b = d$. Donc,

$$d\langle \mathbf{B}_t^\top, \mathbf{B}_t \rangle = n\mathbf{I}_p dt \quad (4.18)$$

et nous voyons bien que (4.10) se réfère au cas où $d\mathbf{X}_t = d\mathbf{Z}_t = d\mathbf{N}_t$.

Maintenant considérons $d\mathbf{B}_t = (\mathbf{S}_t^{\frac{1}{2}})^{-1} \mathbf{N}_t^\top d\mathbf{N}_t$. Celle-ci est bien une équation d'un mouvement brownien multidimensionnel de taille $p \times p$ où $(\mathbf{S}_t^{\frac{1}{2}})^{-1}$ représente l'inverse de la racine carré de la matrice symétrique positive \mathbf{S}_t . Nous voyons, par conséquent, que :

$$\begin{aligned} d\mathbf{B}_t^\top &= (((\mathbf{S}_t^{\frac{1}{2}})^{-1}) (\mathbf{N}_t^\top d\mathbf{N}_t))^\top \\ &= (\mathbf{N}_t^\top d\mathbf{N}_t)^\top ((\mathbf{S}_t^{\frac{1}{2}})^{-1})^\top \\ &= d\mathbf{N}_t^\top \mathbf{N}_t (\mathbf{S}_t^{\frac{1}{2}})^{-1} \end{aligned} \quad (4.19)$$

donc \mathbf{S}_t satisfait l'équation différentielle :

$$d\mathbf{S}_t = \mathbf{S}_t^{\frac{1}{2}} d\mathbf{B}_t + d\mathbf{B}_t^\top \mathbf{S}_t^{\frac{1}{2}} + n\mathbf{I}_p dt \quad (4.20)$$

Remarque 4.2.1. Si $p = 1$, ceci est équivalent à un processus du carré de Bessel de la forme :

$$dS_t = 2S_t^{\frac{1}{2}} dB_t + nIdt \quad (4.21)$$

Considérons à ce moment $\mathbf{S}_t = \mathbf{X}_t^\top \mathbf{X}_t$ ($\mathbf{s}_0 = \mathbf{x}_0^\top \mathbf{x}_0$) où $\mathbf{X}_t = \mathbf{X}_t \mathbf{A} dt + d\mathbf{N}_t \mathbf{Q}$ et $d\mathbf{B}_t = (\mathbf{S}_t^{\frac{1}{2}})^{-1} \mathbf{X}_t^\top d\mathbf{N}_t \mathbf{Q} ((\mathbf{Q}^\top \mathbf{Q})^{\frac{1}{2}})^{-1}$. Alors, \mathbf{B}_t multidimensionnelle $p \times p$. Donc, trouvons l'équation différentielle associée à \mathbf{S}_t . Commençons par $d\mathbf{B}_t$ puisque nous

désirons isoler $d\mathbf{N}_t$.

$$\begin{aligned}
d\mathbf{B}_t &= (\mathbf{S}_t^{\frac{1}{2}})^{-1} \mathbf{X}_t^\top d\mathbf{N}_t \mathbf{Q} (\mathbf{Q}^\top \mathbf{Q})^{-\frac{1}{2}} \\
d\mathbf{B}_t (\mathbf{Q} (\mathbf{Q}^\top \mathbf{Q})^{-\frac{1}{2}})^{-1} &= (\mathbf{S}_t^{\frac{1}{2}})^{-1} \mathbf{X}_t^\top d\mathbf{N}_t \\
d\mathbf{B}_t (\mathbf{Q}^\top \mathbf{Q})^{\frac{1}{2}} \mathbf{Q}^{-1} &= (\mathbf{S}_t^{\frac{1}{2}})^{-1} \mathbf{X}_t^\top d\mathbf{N}_t \quad (4.22) \\
((\mathbf{S}_t^{\frac{1}{2}})^{-1} \mathbf{X}_t^\top)^{-1} d\mathbf{B}_t (\mathbf{Q}^\top \mathbf{Q})^{\frac{1}{2}} \mathbf{Q}^{-1} &= d\mathbf{N}_t \\
(\mathbf{X}_t^\top)^{-1} \mathbf{S}_t^{\frac{1}{2}} d\mathbf{B}_t (\mathbf{Q}^\top \mathbf{Q})^{\frac{1}{2}} \mathbf{Q}^{-1} &= d\mathbf{N}_t.
\end{aligned}$$

Alors, nous trouvons facilement $d\mathbf{N}_t^\top$:

$$\begin{aligned}
d\mathbf{N}_t^\top &= ((\mathbf{X}_t^\top)^{-1} \mathbf{S}_t^{\frac{1}{2}} d\mathbf{B}_t (\mathbf{Q}^\top \mathbf{Q})^{\frac{1}{2}} \mathbf{Q}^{-1})^\top \\
&= (\mathbf{Q}^{-1})^\top ((\mathbf{Q}^\top \mathbf{Q})^{\frac{1}{2}})^\top d\mathbf{B}_t^\top (\mathbf{S}_t^{\frac{1}{2}})^\top ((\mathbf{X}_t^\top)^{-1})^\top \quad (4.23) \\
&= (\mathbf{Q}^\top)^{-1} (\mathbf{Q}^\top \mathbf{Q})^{\frac{1}{2}} d\mathbf{B}_t^\top \mathbf{S}_t^{\frac{1}{2}} \mathbf{X}_t^{-1}.
\end{aligned}$$

De plus, nous remarquons que :

$$\begin{aligned}
d\mathbf{X}_t^\top &= (\mathbf{X}_t \mathbf{A} dt + d\mathbf{N}_t \mathbf{Q})^\top \\
&= (d\mathbf{N}_t \mathbf{Q})^\top + (\mathbf{X}_t \mathbf{A} dt)^\top \quad (4.24) \\
&= \mathbf{Q}^\top d\mathbf{N}_t^\top + \mathbf{A}^\top \mathbf{X}_t^\top dt
\end{aligned}$$

Maintenant, nous savons par la méthode d'intégration par parties que :

$$\begin{aligned}
d\mathbf{S}_t &= d(\mathbf{X}_t^\top \mathbf{X}_t) \\
&= d\mathbf{X}_t^\top \mathbf{X}_t + \mathbf{X}_t^\top d\mathbf{X}_t + \langle \mathbf{X}^\top, \mathbf{X} \rangle_t \\
&= (\mathbf{Q}^\top d\mathbf{N}_t^\top + \mathbf{A}^\top \mathbf{X}_t^\top dt) \mathbf{X}_t + \mathbf{X}_t^\top (\mathbf{X}_t \mathbf{A} dt + d\mathbf{N}_t \mathbf{Q}) + \mathbf{Q}^\top n \mathbf{I}_p dt \mathbf{Q} \\
&= (\mathbf{Q}^\top (\mathbf{Q}^\top)^{-1} (\mathbf{Q}^\top \mathbf{Q})^{\frac{1}{2}} d\mathbf{B}_t^\top \mathbf{S}_t^{\frac{1}{2}} \mathbf{X}_t^{-1} \mathbf{X}_t + \mathbf{A}^\top \mathbf{X}_t^\top \mathbf{X}_t dt) + \quad (4.25) \\
&\quad (\mathbf{X}_t^\top \mathbf{X}_t \mathbf{A} dt + \mathbf{X}_t^\top (\mathbf{X}_t^\top)^{-1} \mathbf{S}_t^{\frac{1}{2}} d\mathbf{B}_t (\mathbf{Q}^\top \mathbf{Q})^{\frac{1}{2}} \mathbf{Q}^{-1} \mathbf{Q}) + n \mathbf{Q}^\top \mathbf{Q} dt \\
&= (\mathbf{Q}^\top \mathbf{Q})^{\frac{1}{2}} d\mathbf{B}_t^\top \mathbf{S}_t^{\frac{1}{2}} + \mathbf{A}^\top \mathbf{S}_t dt + \mathbf{S}_t \mathbf{A} dt + \mathbf{S}_t^{\frac{1}{2}} d\mathbf{B}_t (\mathbf{Q}^\top \mathbf{Q})^{\frac{1}{2}} + n \mathbf{Q}^\top \mathbf{Q} dt
\end{aligned}$$

Ceci nous permet d'introduire formellement un processus de Wishart généralisé.

Définition 4.2.2. Soit un mouvement brownien multidimensionnel B_t $p \times p$ et Q et A des matrices $p \times p$. De plus, considérons la valeur initiale s_0 , une matrice définie positive et n plus grand ou égale à 0. Alors, un processus de Wishart généralisé, S_t , satisfait l'équation différentielle :

$$dS_t = Q^T dB_t^T S_t^{\frac{1}{2}} + S_t^{\frac{1}{2}} dB_t Q + (A^T S_t + S_t A)dt + nQ^T Qdt, \quad S_0 = s_0 \quad (4.26)$$

4.3 Processus d'Orstein-Uhlenbeck au carré

Nous poursuivons notre étude avec le processus de Orstein-Uhlenbeck au carré dans le cas univarié. Par la suite, son analogue dans le cas multivarié est présenté. Ceci permettra de voir le lien direct avec le processus de Wishart. Nous présentons, formellement une équation différentielle stochastique telle que vue lors de la création de modèle en finance. Les équations telles que celle-ci sont dirigées par un mouvement brownien et sont essentielles lors de l'étude de la modélisation de trajectoires aléatoires dynamiques, tel que l'évolution d'un actif risqué.

Définition 4.3.1. Une équation différentielle stochastique est définie par

$$dX_t = \beta(t, X_t)dt + \gamma(t, X_t)dB_t, \quad X_0 = x_0, t \in [0, T], x \in \mathbb{R} \quad (4.27)$$

où $(B_t, t \geq 0)$ est un mouvement brownien standard et $\beta(t, x)$ et $\gamma(t, x)$ sont des fonctions qui représentent le terme de dérive et le terme de diffusion, respectivement.

Nous introduisons le modèle Cox-Ingersoll-Ross. Ce dernier est utilisé pour modéliser l'évolution dans le temps d'un taux d'intérêt court terme (R_t).

Définition 4.3.2. Un processus, R_t , de Cox-Ingersoll-Ross est défini par

$$dR_t = a(b - R_t)dt + (\sigma^2 R_t)^{\frac{1}{2}} dB_t, \quad R_0 = r_0 \quad \text{et} \quad \sigma^2 < 2ab \quad (4.28)$$

où B_t est un mouvement brownien, b , a et σ sont des constantes positives qui représentent la moyenne à long terme, la vitesse à laquelle le processus converge vers la moyenne, et la volatilité du taux d'intérêt court terme respectivement.

Bien qu'il soit possible de trouver une solution unique selon la valeur initiale, il n'est pas possible de la trouver explicitement, mais nous pouvons faire appel à la simulation de Monte Carlo. Une caractéristique intéressante de ce modèle est qu'il ne peut pas prendre de valeurs négatives. L'on remarque que la partie représentant la diffusion se dirige vers zéro quand le taux d'intérêt lui-même approche zéro. Donc, quand la diffusion s'atténue, la courbe du taux d'intérêt dépendra du terme de dérive qui se rapporte à la valeur de la variable a (puisque $R(t) \rightarrow 0$) et qui l'empêche donc de prendre des valeurs sous zéro étant donnée la condition $\sigma^2 < 2ab$. Un processus caractérisé par ceci est appelé un processus de retour à la moyenne.

Le modèle de Cox-Ingersoll-Ross est en fait l'extension du modèle de Ornstein-Uhlenbeck. Nous présentons, alors, celui-ci.

Définition 4.3.3. *Un processus $(R_t, t \geq 0)$ est un processus de Ornstein-Uhlenbeck si R_t satisfait l'équation différentielle stochastique de Ornstein-Uhlenbeck définie par :*

$$dR_t = \beta R_t dt + \eta dB_t, \quad R_0 = \tau, \quad (4.29)$$

où B_t est un mouvement brownien, β et η sont des constantes positives qui représentent la moyenne à long terme et la volatilité, respectivement.

La relation entre ces modèles est constatée lors de l'étude de la somme de carrés du processus de Ornstein-Uhlenbeck, c'est-à-dire dR_t^2 .

C'est ici que la formule de Itô telle que décrite à la définition 4.1.7. vient en jeu. Considérons le cas dans lequel $Y_t = u(t, R_t) = R_t^2$.

Alors, nous avons :

$$\begin{aligned}
dR_t &= \beta R_t dt + \eta dB_t \\
dY_t = dR_t^2 &= \frac{\partial}{\partial t} R_t^2 dt + \frac{\partial}{\partial r} R_t^2 (\beta R_t dt + \eta dB_t) + \frac{1}{2} \frac{\partial^2}{\partial r^2} R_t^2 \eta^2 dt \\
&= 0 + 2R_t(\beta R_t dt + \eta dB_t) + \frac{1}{2} 2\eta^2 dt \\
&= (2R_t(\beta R_t) + \eta^2) dt + 2R_t \eta dB_t \\
&= (2\beta Y_t + \eta^2) dt + 2Y_t^{\frac{1}{2}} \eta dB_t.
\end{aligned} \tag{4.30}$$

Maintenant, considérons la somme des carrés d'un nombre J de processus d'Ornstein-Uhlenbeck, c'est-à-dire, $Y_t = R_{1t}^2 + \dots + R_{Jt}^2$.

Donc,

$$\begin{aligned}
dY_t &= \sum_{j=1}^J dR_{jt}^2 \\
&= \sum_{j=1}^J [2\beta R_{jt}^2 + \eta^2] dt + \sum_{j=1}^J 2R_{jt} \eta dB_t \\
&= [2\beta Y_t + J\eta^2] dt + 2Y_t^{\frac{1}{2}} \eta dB_t
\end{aligned} \tag{4.31}$$

Si nous comparons (4.28) et (4.31) :

$$\begin{aligned}
dR_t &= a(b - R_t) dt + (\sigma^2 R_t)^{\frac{1}{2}} dB_t \\
dY_t &= [2\beta Y_t + J\eta^2] dt + 2Y_t^{\frac{1}{2}} \eta dB_t
\end{aligned} \tag{4.32}$$

On voit clairement que $a = -2\beta$, $ab = J\eta^2$ et $\sigma = 2\eta$.

Maintenant, considérons le modèle d'un taux d'intérêt influencé par plusieurs risques de marché (n plus grand que 1). Le modèle de Ornstein-Uhlenbeck se caractérise, dans le cas multidimensionnel par l'équation différentielle suivante :

$$d\mathbf{R}_t = \mathbf{R}_t \mathbf{A} dt + d\mathbf{B}_t \mathbf{Q}, \mathbf{R}_0 = r_0 \tag{4.33}$$

où \mathbf{B}_t est une matrice de mouvements browniens $p \times p$, \mathbf{A} et \mathbf{Q} sont des matrices $n \times n$.

Observons le cas où $dY_t = dR_t^\top R_t$, $Y_0 = R_0^\top R_0$. Une fois de plus, nous considérons la propriété d'intégration par parties multidimensionnelle.

Alors,

$$\begin{aligned}
d(Y_t) &= d(R_t^\top R_t) \\
&= dR_t^\top R_t + R_t^\top dR_t + d\langle R_t^\top, R_t \rangle_t \\
&= (R_t A dt + dB_t Q)^\top R_t + R_t^\top (R_t A dt + dB_t Q) + nQ^\top Q dt \\
&= (R_t A dt)^\top R_t + (dB_t Q)^\top R_t + R_t^\top R_t A dt + R_t^\top dB_t Q + Q^\top Q dt \\
&= A^\top R_t^\top R_t dt + Q^\top dB_t^\top R_t + R_t^\top R_t A dt + R_t^\top dB_t Q + nQ^\top Q dt \\
&= A^\top Y_t dt + Q^\top dB_t^\top Y_t^{\frac{1}{2}} + Y_t A dt + Y_t^{\frac{1}{2}} dB_t Q + nQ^\top Q dt \\
&= Q^\top dB_t^\top Y_t^{\frac{1}{2}} + Y_t^{\frac{1}{2}} dB_t Q + (A^\top Y_t + Y_t A) dt + nQ^\top Q dt \\
&= (Q^\top Q)^{\frac{1}{2}} dB_t^\top Y_t^{\frac{1}{2}} + Y_t^{\frac{1}{2}} dB_t (Q^\top Q)^{\frac{1}{2}} + (A^\top Y_t + Y_t A) dt + nQ^\top Q dt
\end{aligned} \tag{4.34}$$

La solution (Y_t) de cette équation différentielle est un processus de Ornstein-Uhlenbeck au carré et nous reconnaissons bien que celle-ci correspond exactement à un processus de Wishart généralisé.

4.4 Simulation par l'approximation d'Euler

Rappelons l'équation différentielle stochastique (4.27) et considérons sa forme multidimensionnelle :

$$dX_t = \beta(t, X_t)dt + \gamma(t, X_t)dB_t \tag{4.35}$$

dont la solution est un processus d'Itô, X_t , sur l'intervalle $0 \leq t \leq T$. De plus, considérons la discrétisation de l'intervalle $[0, T]$: $0 = t_0 \leq t_1 \leq t_2 \leq \dots \leq t_N = T$, avec $T = 1$ tel que $X_{t_0} = X_0$ et $t_{i+1} - t_i = \frac{1}{N}$.

L'approximation de Euler est une méthode récursive permettant d'approximer la

solution d'une équation différentielle selon le schéma suivant :

$$\tilde{\mathbf{X}}_{t_{i+1}} = \tilde{\mathbf{X}}_{t_i} + \beta(t_i, \tilde{\mathbf{X}}_i)(t_{i+1} - t_i) + \gamma(t_i, \tilde{\mathbf{X}}_i)(\mathbf{B}_{t_{i+1}} - \mathbf{B}_{t_i}), \quad (4.36)$$

pour lequel $i = 0, 1, \dots, N - 1$ et $\tilde{\mathbf{X}}_{t_0} = \mathbf{X}_{t_0}$. On observe que pour simuler ce processus étant donnée la valeur initiale, il suffit de simuler l'accroissement d'un mouvement brownien. Pour plus de détails, ce référer à Iacus (2008).

Rappelons qu'un mouvement brownien multidimensionnel est de la forme :

$$\mathbf{B} = \{\mathbf{B}_t = (\mathbf{B}_t^1, \dots, \mathbf{B}_t^n)^T, t \in [0, \infty)\} \quad (4.37)$$

dont les composantes sont indépendentes et $\mathbf{B}_0^i = 0$ pour $i = 1, \dots, n$. De plus, les accroissements sont des variables gaussiennes aléatoires indépendentes tels que si $0 \leq t < s$ alors $\mathbf{B}_s - \mathbf{B}_t \sim N_p(\mathbf{0}, (s - t)\mathbf{I})$.

Tel que définie précédemment, l'équation différentielle stochastique associée à un processus de Wishart généralisé est de la forme :

$$d\mathbf{S}_t = \mathbf{Q}^T d\mathbf{B}_t^T \mathbf{S}_t^{\frac{1}{2}} + \mathbf{S}_t^{\frac{1}{2}} d\mathbf{B}_t \mathbf{Q} + (\mathbf{A}^T \mathbf{S}_t + \mathbf{S}_t \mathbf{A})dt + n\mathbf{Q}^T \mathbf{Q}dt, \quad \mathbf{S}_0 = \mathbf{s}_0 \quad (4.38)$$

Par conséquent, la méthode de Euler offre une discrétisation du processus de Wishart de la forme :

$$\begin{aligned} \tilde{\mathbf{S}}_{t_{i+1}} = & \tilde{\mathbf{S}}_{t_i} + \mathbf{Q}^T (\mathbf{B}_{t_{i+1}}^T - \mathbf{B}_{t_i}^T) \tilde{\mathbf{S}}_{t_i}^{\frac{1}{2}} + \tilde{\mathbf{S}}_{t_i}^{\frac{1}{2}} (\mathbf{B}_{t_{i+1}} - \mathbf{B}_{t_i}) \mathbf{Q} + \\ & (\mathbf{A}^T \tilde{\mathbf{S}}_{t_i} + \tilde{\mathbf{S}}_{t_i} \mathbf{A})(t_{i+1} - t_i) + n\mathbf{Q}^T \mathbf{Q}(t_{i+1} - t_i) \end{aligned} \quad (4.39)$$

où $\tilde{\mathbf{S}}_{t_0} = \mathbf{S}_{t_0}$.

Par construction, le schéma d'Euler implique que plus N est grand plus $\tilde{\mathbf{S}}$ se rapproche de \mathbf{S} puisque l'approximation sera faite à une fréquence plus élevée. Alors, si les partitions de l'intervalle de temps sont suffisamment petites, $\tilde{\mathbf{S}}$ sera semi-définie positive dans toute cet intervalle. Bien que le schéma d'Euler soit

explicite, il n'impose pas que le processus soit strictement positif pour tout temps dans l'intervalle $[t_0, T]$ tel qu'un processus de Wishart. Par la nature d'un mouvement brownien, la racine carrée de la matrice $\tilde{\mathbf{S}}$ ne sera pas réalisable s'il implique des valeurs négatives. Ceci provoquera l'arrêt de la simulation du processus avant même d'atteindre le temps T et donc certaines restrictions sont nécessaires.

Supposons un processus discret de Wishart de la forme $(\tilde{\mathbf{S}}_{t_0}, \tilde{\mathbf{S}}_{t_1}, \dots, \tilde{\mathbf{S}}_{t_j})$ associé au mouvement brownien discretisé $(\mathbf{B}_{t_0}, \mathbf{B}_{t_1}, \dots, \mathbf{B}_{t_j})$ tel que $j < N - 1$ et que lors de la production de $\tilde{\mathbf{S}}_{t_{j+1}}$, celle-ci ne soit plus semi-définie positive. Dans ce cas, un pont brownien sera considéré conditionnellement sur le processus de Wishart aux temps t_j et t_{j+1} afin d'obtenir $\mathbf{B}_{t_{j+1/2}}$. C'est-à-dire, à partir du pont brownien rejoignant ces deux temps il est possible de simuler le processus de Wishart associé au mouvement brownien au temps $t_{j+1/2}$ (mi-temps) tel que :

$$\mathbf{B}_{t_{j+1/2}} = \frac{(t_{j+1} - t_{j+1/2})\mathbf{B}_{t_j} + (t_{j+1/2} - t_j)\mathbf{B}_{t_{j+1}}}{t_{j+1} - t_j} + \sqrt{\frac{(t_{j+1} - t_{j+1/2})(t_{j+1/2} - t_j)}{t_{j+1} - t_j}} \mathbf{Z} \quad (4.40)$$

où $\mathbf{Z} \sim N_p(\mathbf{0}, \mathbf{I})$. Pour plus de détails à l'égard de la simulation d'un pont brownien, se referer à Belezza Sousa *et al.* (2012).

Ensuite, les valeurs propres de $\tilde{\mathbf{S}}_{t_{j+1/2}}$ sont évaluées afin de confirmer qu'elle soit bien semi-définie positive. Si ceci est bien le cas, les futures simulations seront produites aux temps $(t_{j+1}, t_{j+3/2}, t_{j+2}, \dots, T)$; sinon, un nouveau pont brownien sera construit afin d'évaluer $\mathbf{B}_{t_{j+1/4}}$. Cette procédure est refaite jusqu'à l'obtention d'un processus discret de Wishart allant jusqu'au temps T pour lequel la matrice Wishart est toujours semi-définie positive.

Exemple 4.4.1. Considérons un processus de Wishart à deux dimensions établi selon l'approximation de Euler sur un intervalle de temps $[0, 1]$. Nous rappelons

l'équation différentielle stochastique de ce dernier :

$$dS_t = Q^T dB_t^T S_t^{\frac{1}{2}} + S_t^{\frac{1}{2}} dB_t Q + (A^T S_t + S_t A)dt + nQ^T Q dt, S_0 = s_0 \quad (4.41)$$

Prenons

$$Q = \begin{pmatrix} 1.00 & 4.00 \\ 0.00 & -2.00 \end{pmatrix} \quad A = \begin{pmatrix} -3.00 & 0.00 \\ 0.00 & -3.00 \end{pmatrix} \quad s_0 = \begin{pmatrix} 1.00 & 0.00 \\ 0.00 & 1.00 \end{pmatrix}.$$

Nous observons dans la figure 4.1 la trajectoire des éléments de la matrice S en commençant par une partition telle que $t_{i+1} - t_i = \frac{1}{1000}$.

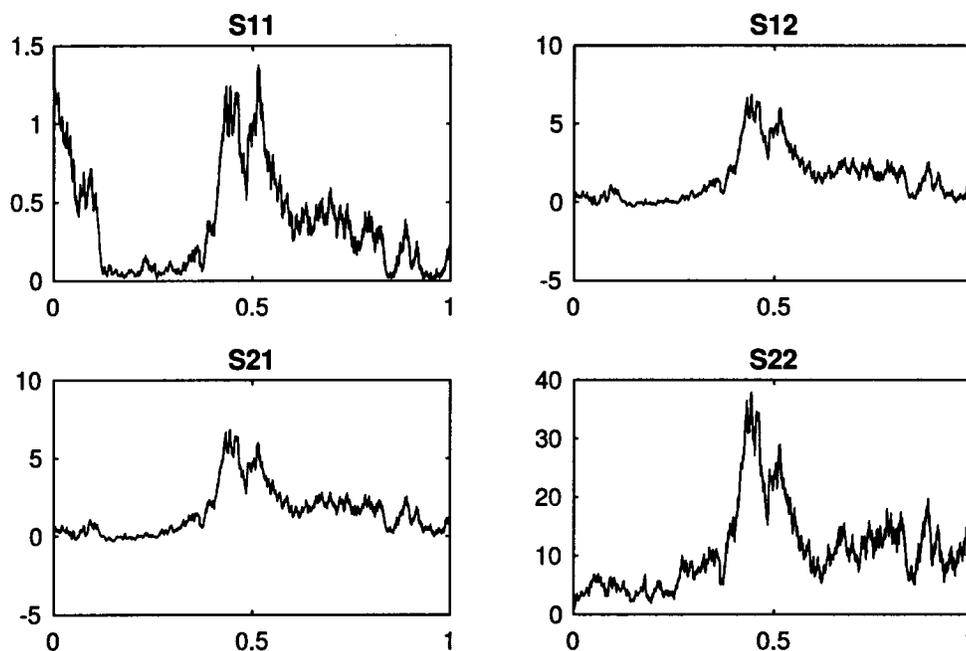


Figure 4.1 Processus de Wishart à 2 dimensions

Cette trajectoire consiste d'un peu plus de 69 000 évaluations. C'est-à-dire, que l'intervalle entre deux temps à été réduit au moins une fois.

La simulation d'Euler est un processus discrétisation qui converge vers le processus de Wishart et est une façon relativement simple d'estimer son parcours dynamique. Malheureusement, elle comporte certaines lacunes, en particulier, lors de l'évaluation de la racine carrée de la matrice de Wishart, qui n'est pas toujours bien définie. Une discrétisation exacte d'un processus de Wishart fut présentée en 2013. Au cœur de cette impressionnante découverte, se trouve la décomposition de la forme canonique d'un processus de Wishart. Cette décomposition est définie par ses auteurs comme : *«the infinitesimal generator of $WIS_d(x, \alpha, 0, \mathbf{I}_p^n)$ is the sum of commuting operators that are associated to elementary SDEs that can be sampled exactly»* (Ahdida et Alfonsi, 2013). Ceci leur permettra, par la suite, de généraliser leur découverte pour tout processus de Wishart sans restrictions sur les paramètres.

CONCLUSION

Au début de ce mémoire, nous avons élaboré, expliqué et montré plusieurs notions fondamentales d'une matrice aléatoire de loi de Wishart. Dans un contexte multidimensionnel, ceci s'avère problématique puisqu'il n'est pas si simple de les illustrer. Un exemple a été présenté via la décomposition de Bartlett et ceci a permis d'observer l'espérance et la matrice de variance-covariance empiriques d'un échantillon de taille considérable.

Ensuite, il a été pour but d'offrir un avant-goût de l'implication de la forme d'une matrice aléatoire de loi de Wishart inversée dans le domaine de la statistique bayésienne. Nous avons examiné les lois a posteriori des paramètres associés à la loi normale multivariée. Il a été observé que la loi de Wishart est en fait une distribution a priori conjuguée pour la matrice de précision et par conséquent la distribution a posteriori de la matrice de variance-covariance suit une loi de Wishart inversé.

Après, son apport à d'autres lois de probabilité en commençant par son équivalent dans le context univarié à été étudié. Nous avons observé, entre autres, un lien direct entre ces lois en considérant $\mathbf{x}^T \Sigma \mathbf{x}$ en tant que paramètre d'échelle de la loi χ_n^2 .

Finalement, la mathématique financière, un sujet très riche, fait appel par moment au processus de Wishart. Nous l'avons révélé tel que Bru (1991) l'introduit. Ce processus dynamique a été illustré par le biais de l'approximation de Euler. Cette méthode offre des résultats satisfaisants mais cette approximation n'est pas toujours bien définie à cause de la racine carrée de la matrice $\tilde{\mathbf{S}}$. Afin de résoudre à

cela, un pont brownien a été considéré et la trajectoire d'un processus de Wishart à deux dimensions a été observée sur un intervalle de temps $[0, 1]$.

Son association aux matrices définies positives, telle que la matrice de variance-covariance, confirme son importance dans plusieurs aspects de la statistique. Il est évident que, par son intime lien à des processus stochastiques tel que le mouvement brownien cette loi de probabilité joue toujours un rôle important dans le domaine de la statistique.

ANNEXE A

PRÉLIMINAIRES

A.1 Jacobien de certaines transformations

Au chapitre 1 et 2, plusieurs transformations matricielles seront nécessaires pour changer d'une fonction de densité à une autre. Ces transformations sont associées au calcul de différents Jacobiens qui permettront d'étudier les caractéristiques d'une matrice aléatoire de Wishart. C'est pour cela que ce sujet est brièvement présenté dans cette section. Pour plus de détails, ce référer à Mathai (1997) et Gupta et Nagar (2000).

Définition A.1.1. Soit \mathbf{Y} et \mathbf{X} des matrices aléatoires tel que $\mathbf{Y} = (y_1, \dots, y_p)^\top$ et $\mathbf{X} = (x_1, \dots, x_p)^\top$. Supposons la transformation $\mathbf{X} = F(\mathbf{Y})$. Alors, le jacobien de la transformation de \mathbf{X} à \mathbf{Y} ou \mathbf{Y} en tant que fonction de \mathbf{X} est définie par :

$$J(\mathbf{Y} \rightarrow \mathbf{X}) = \det \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_1}{\partial x_p} \\ \vdots & & \vdots \\ \frac{\partial y_p}{\partial x_1} & \dots & \frac{\partial y_p}{\partial x_p} \end{pmatrix} \text{ ou } (d\mathbf{Y}) = J(d\mathbf{X}), J \neq 0$$

où $\frac{\partial y_i}{\partial x_j}$ ($i, j = 1, \dots, p$) représentent les dérivées partielles des éléments de \mathbf{Y} par rapport aux éléments de \mathbf{X} .

Propriété A.1.1. Si $Y = AXA^\top$ tel que Y, X sont des matrices $p \times p$ symétriques et $A_{p \times p}$, $|A| \neq 0$, alors

$$dY = |A|^{p+1} dX \quad (\text{A.1})$$

Propriété A.1.2. Si $Y = X^{-1}$ tel que X soit une matrice $p \times p$ symétrique alors

$$dY = |X|^{-(p+1)} dX \quad (\text{A.2})$$

Propriété A.1.3. Soit Y, X, U, V sont des matrices $p \times p$ symétriques, définies positives. Si $U = X + Y$ et $V = U^{-\frac{1}{2}} Y (U^{-\frac{1}{2}})^\top$ où $U = U^{\frac{1}{2}} (U^{\frac{1}{2}})^\top$ alors

$$J((X, Y) \rightarrow (U, V)) = |U|^{\frac{(p+1)}{2}} dX \quad (\text{A.3})$$

ANNEXE B

CODE MATLAB

B.1 Code pour l'algorithme d'échantillonnage d'une variable aléatoire de loi de Wishart

```
1 function [ matWish,A3, choles] = Wishart(n,V)
2 % fonction donnant une matrice aléatoire de Wishart
3 % n est le degré de liberté
4 % V est la matrice Sigma
5
6 % Factorisation de Cholesky de la Sigma
7 p= size(V,1);
8 choles= (chol(V))';
9
10 % Création de la matrice triangulaire supérieur T,
11 % représenté ici par A3
12 A= zeros(p);
13 rchi=zeros(p,1);
14
15 for i=1:p
16 rchi(i,1)=chi2rnd(n-i+1);
17 end
```

```
18
19 % Racine carrée de variable aléatoire de Khi deux sur
20 % la diagonale
21 rchi2=sqrt(rchi);
22 mat1=diag(rchi2);
23
24 % Variable aléatoire de loi normale de la partie
25 % supérieur de la matrice
26 nor=randn(p);
27 mat2= tril(nor,-1);
28
29 A2=A+mat1+mat2;
30 A3=A2';
31
32 %Création de la matrice de Wishart.
33 matWish=choles*(A3'*A3)*(choles');
34 end
35
36 % Part 2
37 fonction[B,espetheo,C,vartheo]=Wishart_esp_var(reps,V,n)
38 % fonction donnant l'espérance théorique et observée
39 % ainsi que la matrice de variance covariance théorique
40 % et observée
41 % reps est le nombre de répétition
42 % n est le degré de liberté
43 % V est la matrice Sigma
44
45 p= size(V,1);
```

```
46 elements=p*p;
47 matrice=cell(reps,1);
48 matrice2=zeros(reps,elements);
49
50 for ii=1:reps
51 matrice{ii} = Wishart(n,V);
52 end
53
54 for k=1:reps
55 matrice2(k,:)=reshape(matrice{k},1,elements);
56 end
57
58 % La matrice d'espérance observée
59 moyenne=mean(matrice2);
60 B=reshape(moyenne,p,p);
61 % La matrice d'espérance théorique
62 espetheo=n*V;
63
64 % La matrice de variance covariance observée
65 C=cov(matrice2);
66 % La matrice de variance théorique
67 K=commutation(p,p);
68 vartheo=n*(eye(elements)+K)*kron(V,V);
69 end
70
71
72 function K = commutation(m,n)
73 %function donnant la matrice de commutation
```

```

74 K = zeros(m*n, m*n);
75 m0 = 1:(m*n);
76
77 N = reshape(m0, m,n)';
78 n0 = N(:);
79
80 for i = 1:(m*n)
81     K(m0(i), n0(i)) = 1;
82 end
83 end

```

B.2 Code pour l'algorithme du processus de Wishart par l'approximation de Euler

```

1 function [S,t] = motion8( Q,A,n,Sone,dt0 )
2 % Fonction générant un processus de Wishart dynamique
3 % Q et A sont des matrice n par n
4 % n est un nombre entier
5 % Sone est la valeur de la matrice de Wishart au temps 0
6 % dt0 est l'écart initial entre 2 points dans
7 % l'intervalle de temps
8
9 for j=1:1000
10 S{j} = zeros(2);
11 end
12 S{1}=Sone;
13
14 for k=1:1000

```

```
15 t{k} = zeros(1);
16 end
17 t{1}=0;
18
19 interval=dt0;
20 B1=[0 0;0 0];
21
22 i=2;
23 time=0;
24
25 while time<=1
26 if (interval<eps)
27     disp('Intervalle trop petit et converge vers 0');
28     return
29     % Arrêt de simulation si intervalle entre 2 temps
30     % est trop court
31 end
32 B2=B1+sqrt(interval)*normrnd(0,1,2,2);
33 %Formule de Euler.
34 S{i}=S{i-1} + (Q.)*(B2.'-B1.)*sqrtm(S{i-1})
35 + sqrtm(S{i-1})*(B2-B1)*Q + ((A.)*S{i-1} +
36 S{i-1}*A)*interval + n*(Q.)*Q*interval;
37 %Obtient les valeurs propres de la matrice de Wishart.
38 vp=eig(S{i});
39 mineig=min(vp);
40
41 if mineig>0 %Vérifie que la plus petite valeur propre
42             %soit positive.
```

```
13 time=time+interval;
14 t{i}=time;
15 B1=B2;
16 i=i+1;
17
18 else
19
20 while mineig<0
21 interval2=(1/2)*interval;
22 if (interval2<eps)
23     disp('Intervalle trop petit et converge vers 0');
24     return
25     % Arrêt de simulation si intervalle entre 2 temps
26     % est trop court
27 end
28 % Création du pont brownien si la plus petite valeur
29 % propre est négative.
30 Bdiff=0.5*(B2-B1)+(sqrt(interval2*interval2*
31     (1/interval)))*normrnd(0,1,2,2);
32
33 interval=interval2;
34 S{i}=S{i-1} + (Q.')(Bdiff.)*sqrtm(S{i-1}) +
35     sqrtm(S{i-1})*(Bdiff)*Q + ((A.)*S{i-1} +
36     S{i-1}*A)*interval + n*(Q.)*Q*interval;
37
38 vp=eig(S{i});
39 mineig=min(vp);
40 end
```

```
71 time=time+interval;  
72 t{i}=time;  
73 B1=B2;  
74 i=i+1;  
75 end  
76 end  
77 end
```


RÉFÉRENCES

- Ahdida, A. et Alfonsi, A. (2013). Exact and high-order discretization schemes for wishart processes and their affine extensions. *The Annals of Applied Probability*, 23(3), 1025–1073. Récupéré de <http://www.jstor.org/stable/23473267>
- Beleza Sousa, J., Esquivel, M. L. et Gaspar, R. M. (2012). Brownian bridge vectorial simulation.
- Bru, M.-F. (1991). Wishart processes. *Journal of Theoretical Probability*, 4(4), 725–751. <http://dx.doi.org/10.1007/BF01259552>. Récupéré de <https://doi.org/10.1007/BF01259552>
- Christensen, R. (2015). Covariance of the wishart distribution with applications to regression. Récupéré de <http://www.math.unm.edu/~fletcher/Wishart.pdf>
- Deheuvels, P. (2012- 2013). Cours d'analyse statistique multivariée. Récupéré de <http://wintenberger.fr/cours/ModelesLineaires/Anova-2012.pdf>
- E. Shreve, S. (2003). *Stochastic Calculus for Finance II*, volume I, II.
- Eaton, M. L. (2007a). *Chapter 1 : Vector Space Theory*, Dans *Multivariate Statistics*, volume Volume 53 de *Lecture Notes–Monograph Series*, (p. 1–69). Institute of Mathematical Statistics : Beachwood, Ohio, USA
- Eaton, M. L. (2007b). *Chapter 8 : The Wishart Distribution*, Dans *Multivariate Statistics*, volume Volume 53 de *Lecture Notes–Monograph Series*, (p. 302–333). Institute of Mathematical Statistics : Beachwood, Ohio, USA
- Gupta, A. K. et Nagar, D. K. (2000). *Matrix Variate Distributions*. Chapman and Hall/CRC.
- Iacus, S. M. (2008). *Simulation and Inference for Stochastic Differential Equations*. Springer-Verlag New York.
- Mathai, A. (1997). *Jacobians of Matrix Transformations and Functions of Matrix Argument*. World Scientific Pub. Récupéré de https://books.google.ca/books?id=Ci0hr2_nYjYC
- Muirhead, R. J. (2005). *Aspects of Multivariate Statistical Theory*. A John Wiley and Sons, INC., Publication.

Nydick, S. W. (2012). *The Wishart and Inverse Wishart Distributions*.

Quarteroni, A., Sacco, R. et F., S. (2007). *Méthode Numériques*. Springer-Verlag Mailand.

Seber, G. A. F. (1984). *Multivariate Observations*. New Jersey, USA : A John Wiley and Sons, INC., Publication.

Wishart, J. (1928). The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, 20A(1/2), 32–52. Récupéré de <http://www.jstor.org/stable/2331939>