HAL
open science

Contribution des familles exponentielles en traitement des images

Taher Ben Arab

To cite this version:

Taher Ben Arab. Contribution des familles exponentielles en traitement des images. Autre [cond-mat.other]. Université du Littoral Côte d'Opale; Université de Sfax. Faculté des sciences, 2014. Français. NNT : 2014DUNK0357 . tel-01019983

HAL Id: tel-01019983

https://theses.hal.science/tel-01019983v1

Submitted on 7 Jul 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse de DOCTORAT
N° d'ordre : ULCO 2014 .06

Université du Littoral Cote d'Opale
Ecole Doctorale de Lille

Université de Sfax
Faculté des Sciences de Sfax

THÈSE

Présentée

en vue de l'obtention du

DOCTORAT
(En Cotutelle)

Spécialité : Traitement du signal et de l'image et Mathématiques Appliquées

par

Taher Ben Arab

CONTRIBUTION DES FAMILLES EXPONENTIELLES EN TRAITEMENT DES IMAGES

Soutenue le 26 Avril 2014, devant le jury composé de :

Président : M. Abdelhamid Hassairi, Professeur à la Faculté des Sciences de Sfax, Tunisie

Rapporteurs : M. Yassine Ruichek, Professeur à l'Université de Technologie de Belfort-Montbéliard, France
M. Hassine Maatoug, Maître de Conférences HDR à la Faculté des Sciences de Monastir, Tunisie

Examinateur : M. Pierre Gouton, Professeur à la Faculté de Sciences et Techniques Dijon, France

Directeurs : M. Mourad Zribi, Maître de Conférences HDR à l'Université du Littoral Côte d'Opale, France
M. Afif Masmoudi, Professeur à la Faculté des Sciences de Sfax, Tunisie

Invité : M. Atef Masmoudi, Maître Assistant à l'Institut Préparatoire au Etudes d'Ingénieurs de Sfax, Tunisie

DÉDICACES

*A mon cher père **Mohamed** et ma chère mère **Najwa**, en leur disant que je vous dois tout ce que j'ai et j'aurai. Mes parents, nulle expression ne pourrait exprimer ce que je vous dois pour vos efforts déployés et sacrifices pour ma formation. Que dieu leur assure bonne santé et longue vie.*

*A ma très chère épouse **Atizez** et mon adorable fils **Mohamed.** En témoignage de leur soutien, leur amour, leur compréhension, leur encouragement et leurs aides infinies.*
Que Dieu vous garde pour moi et m'accorde santé et longue vie et m'aide à vous rendre heureux autant que vous méritez.

*A mes chers beaux parents **Mustapha** et **Naima** En témoignage de ma grande affection et mes attachements, je vous présente l'expression de ma gratitude et mon grand amour.*
Que dieu vous procure bonne santé et longue vie.

*A mes chers frères **Bilèl, Haikal** et **Morsi,** Que dieu vous procure une vie pleine de bonheur et de prospérité. Que Dieu vous préserve la bonne santé, heureuse vie et un avenir professionnel très brillant.*

*A mon beau frère **Fakreddine** et sa femme **olfa** et leurs adorable enfants **Aimna, Mohamed Aimin** et **Ahmèd.***
*A mon beau frère **Mohamed Fouazi** et sa femme **Maroi** et leurs adorable enfants **Yassin** et **Yesmine.***

*A mon beau frère **Sami** et son femme **Siwar** et leurs adorable fille **Mariem.***
*A ma belle sœur **Inès** et son mari **Houssem** et leurs adorable enfants **Mariem et Mohamed.***
Que Dieu vous préserve la bonne santé, heureuse vie et un avenir professionnel très brillant.
A tous les membres de ma famille sans oublier personne, je dis que je suis fier de vous, espérant que vous êtes fiers de moi.

REMERCIEMENTS

Ce travail de recherche est le résultat d'une convention en cotutelle entre l'Université du Littoral Côte d'Opale (ULCO) France et la Faculté des Sciences de Sfax (FSS) Tunisie. Ce travail a été réalisé au sein du Laboratoire d'Informatique Signal et Image de la Côte d'Opale (LISIC-ULCO) du coté Français et au sein du Laboratoire de Probabilités et Statistique (LPS-FSS) du coté Tunisien.

*J'adresse mes plus vifs remerciements à Monsieur **Mourad Zribi**, Maître de conférences HDR à l'Université du Littoral Côte d'Opale, ainsi qu'à Monsieur **Afif Masmoudi**, Professeur à la faculté des Sciences de Sfax pour avoir assuré la direction scientifique de ce travail de thèse, pour les nombreux et fructueuses discussions, ainsi que ses encouragements qui m'ont été d'une grande aide.*

*J'adresse ensuite mes plus sincères remerciements à Monsieur **Abdelhamid Hassairi**, Professeur à la faculté des Sciences de Sfax d'avoir accepté de présider le jury de cette thèse.*

*Je tiens à remercier Monsieur **Yassine Ruichek**, Professeur à l'Université de Technologie de Belfort Montbéliard, ainsi qu'à Monsieur **Hassine Maatoug**, Maître de conférences HDR à la Faculté des Sciences de Monastir pour avoir accepté la lourde charge de rapporteur, ainsi que pour les remarques constructives qu'ils m'ont fournies.*

*Je remercie Monsieur **Pierre Gouton**, Professeur à l'Université de Bourgogne d'avoir accepté de juger ce travail en tant qu'examinateur.*

*Mes vifs remerciements s'adressent également à Monsieur **Atef Masmoudi** Maître Assistant à l'Institut Préparatoire aux Etudes d'Ingénieurs de Sfax pour m'avoir fait l'honneur de participer à mon jury de thèse.*

Mes sincères remerciements s'adressent également à tous les membres du laboratoire de Probabilités et Statistique de la faculté des Sciences de Sfax, Tunisie, et du laboratoire d'Informatique Signal et Image de la Côte d'Opale de l'Université du Littoral Côte d'Opale, France.

A tous les enseignants qui ont contribué à l'enrichissement de ma formation, qu'ils acceptent mes remerciements profonds et qu'ils trouvent dans cet écrit le fruit de leur engagement et leur travail persévérant.

Trouvez ici l'expression de ma reconnaissance et mes remerciements les plus profonds.

Table des matières

Introduction générale **1**

1 Trace de la fonction variance dans les familles exponentielles **4**

1.1 Introduction 4

1.2 Généralités sur les familles exponentielles 6

1.2.1 Familles exponentielles générales 6

1.2.2 Familles exponentielles naturelles 11

1.2.3 Steepness 18

1.2.4 Lois infiniment divisibles 19

1.3 Introduction de la fonction trace dans les FENs infiniment divisibles 21

1.3.1 Nouvelle caractérisation d'une FEN infiniment divisible 22

1.3.2 Exemples 28

1.3.3 Propriétés asymptotiques de la fonction trace 37

1.4 Application de la fonction trace 42

1.4.1 En traitement du signal 43

1.4.2 En traitement d'images couleur 48

1.5 Conclusion 63

2 Estimation paramétrique des modèles statistiques multivariés 64

2.1 Introduction 64

2.2 Lois multivariées 65

2.2.1 Exemples de lois multivariées 66

2.2.2 Estimation des paramètres 73

2.2.3 Performance des estimateurs 81

2.3 Proposition d'un nouveau modèle multivarié 86

2.3.1 Rappel sur la loi de Riesz 86

2.3.2 Loi de la diagonale modifiée de Riesz 88

2.3.3 Estimation des paramètres 94

2.3.4 Méthode Bayésienne 100

2.3.5 Estimateur EAP pour le modèle proposé. 102

2.3.6 Performance des estimateurs 105

2.4 Conclusion 110

3 Application des familles exponentielles en segmentation d'images couleurs 112

3.1 Introduction 112

3.2 Les représentations classiques de la couleur/ les espaces de couleurs 114

3.2.1 L'espace RVB 115

3.2.2 L'espace TSL 117

3.3 Les différentes approches de la segmentation d'images couleur 120

3.3.1 La segmentation par approche contours 121

3.3.2 La segmentation par approche régions 123

3.4 Modélisation probabiliste en classification 128

3.4.1 Modèle aléatoire de l'image 129

3.4.2 Modélisation par mélange de lois 130

3.4.3 Estimation des paramètres d'un modèle mélange 135

3.4.4 Algorithme EM 137

3.5 Application de l'algorithme EM pour un mélange de lois de la diagonale modifiée de Riesz 148

3.5.1 Estimation des paramètres du mélange MGBK 148

3.5.2 Estimation des paramètres du deuxième mélange MCGBKGU . . 153

3.6 Segmentation d'images couleur 158

3.7 Conclusion 168

Conclusion et Perspectives **169**

Annexe A **172**

Annexe B **178**

Bibliographie **183**

196

Introduction générale

Ce travail s'inscrit dans une thématique de recherche commune au laboratoire de Probabilités et Statistique de l'université de Sfax et au laboratoire d'Informatique, Signal et Image de l'université du Littoral Côte d'Opale. Il présente une synthèse des travaux de recherche que nous avons réalisés dans les domaines de probabilités et statistique et de traitement de signal et d'images.

Les mathématiciens du 19ème et du 20ème siècles ont porté un grand intérêt aux sommes de variables aléatoires indépendantes et identiquement distribuées. De telles études ont abouti à des résultats fondamentaux comme le théorème de Bernoulli, qui a par la suite trouvé sa généralisation dans le théorème de Moivre-Laplace [62]. Laplace et Gauss, qui ont développé la théorie des erreurs d'observations, ont associé la distribution de l'erreur à la somme de variables aléatoires indépendantes, formalisant ainsi le théorème centrale limite. C'est aussi en étudiant les sommes des variables aléatoires indépendantes que Paul Lévy a pu initier la théorie générale des distributions infiniment divisibles en 1937. La théorie moderne porte encore l'empreinte de ce travail pionnier, bien que de nombreux auteurs, comme A. Ya. Khintchine [56], aient contribué à des améliorations et de nouveaux résultats. Au cours des dernières décennies, une masse de nouvelles idées est devenue évidente.

Le concept de lois infiniment divisibles est étendu aux cas des distributions multidimensionnelles [93][98]. Les théorèmes décrivant les distributions infiniment divisibles utilisent par conséquent et très souvent les fonctions caractéristiques, les transformées de Laplace ou les fonctions génératrices plutôt que les distributions elles mêmes. L'ensemble des lois infiniment divisibles représente une classe importante en probabilités et statistique (voir [41][46][52][57][58][59]).

Le traitement d'images est une discipline de l'informatique et des mathématiques appliquées qui étudie les images numériques et leurs transformations, dans le but d'améliorer leur qualité ou d'en extraire de l'information. Le présent travail consiste à étudier des extensions des lois infiniment divisibles en traitement d'images. Des familles exponentielles naturelles ont en effet reçu une attention soutenue d'un point de vue théorique dans les communautés des probabilités et des statistiques, mais ces familles ont cependant été peu étudiées dans le domaine du traitement d'images. Nous cherchons donc ici à évaluer l'apport de ces familles exponentielles dans les problèmes de modélisation des bruits et de segmentation des images couleurs.

Ce mémoire est composé de trois chapitres. Le chapitre 1 est consacré à des familles exponentielles naturelles infiniment divisibles qui seront appliquées à l'imagerie couleur. Nous étudions dans un premier temps des généralités des familles exponentielles et dans un second temps, celles des lois infiniment divisibles, qui ont des propriétés caractéristiques intéressantes. Cette étude nous amène à introduire une nouvelle caractérisation des familles exponentielles naturelles infiniment divisibles pour le problème de la détection de la loi du bruit dans un signal ou dans une image [9][10].

Le chapitre 2 traite le problème de l'estimation paramétrique des modèles multivariés.

Dans un premier temps, nous proposons un nouveau modèle multivarié basé sur la loi de Riesz. La loi de ce nouveau modèle est le produit de convolution d'une loi Gamma univariée et d'une loi Gamma bivariée de Kibble [11]. Dans un deuxième temps, nous étudions le problème de l'estimation des paramètres de ce modèle en utilisant la méthode des moments, la méthode du maximum de vraisemblance et l'approche bayésienne et la performance des estimateurs obtenus en se basant sur l'erreur quadratique moyenne [14].

Le chapitre 3 traite le problème de la généralisation du modèle proposé dans le chapitre 2 dans le cadre du mélange de lois dans le but d'améliorer la qualité de la segmentation des images couleur. En particulier, nous montrons la possibilité de mettre en place un nouvel algorithme paramétrique dans le but d'estimer les paramètres du mélange de ce nouveau modèle [13]. Une présentation de quelques résultats de segmentation statistique multivariée non supervisée d'images couleur caractérisées par la Teinte et la Saturation finalisera ce chapitre [12][13].

La conclusion générale fera l'objet d'un bilan critique des diverses approches proposées pour la caractérisation d'une famille exponentielle naturelle, l'estimation des lois multivariées et la segmentation des images couleur et suggèrera quelques perspectives de recherche.

Chapitre 1

Trace de la fonction variance dans les familles exponentielles

1.1 Introduction

La dispersion représente la variabilité ou l'étendue de différentes valeurs que peut prendre une variable. En statistique, il existe différentes mesures de la dispersion. Les plus courantes sont la variance et l'écart-type ou bien l'écart interquantille. Dans ce chapitre, nous nous interessons à la mesure de la dispersion basée sur la fonction variance dans les familles exponentielles naturelles infiniment divisibles. Les familles exponentielles jouent un rôle important en probabilités et en statistique [32][82][88] notamment en traitement d'images [4][5]. Une famille exponentielle est une classe de lois de probabilités paramétrées. La plupart des lois discrètes et absolument continues standards sont des familles exponentielles. Les familles exponentielles présentent une vaste classe de lois de probabilités très utiles en pratique. Elles représentent un centre d'intérêt pour plusieurs raisons, parmi elles nous

citons deux raisons principales :

1. Plusieurs modèles statistiques sont des familles exponentielles.
2. Les méthodes classiques d'estimation s'appliquent avec succès lorsque le modèle statistique est associé à une famille exponentielle.

Les familles exponentielles apparaissent de façon naturelle dans la recherche de distributions lors des applications statistiques et en particulier dans les méthodes bayésiennes [109]. La famille exponentielle est composée de plusieurs lois paramétriques [66][78] parmis lesquelles nous pouvons citer les lois suivantes : Normale, Exponentielle, Gamma, Bêta, Dirichlet, Bernoulli, Bernoulli Multinomiale, Poisson, Wishart, Inverse Wishart,... etc. Il existe d'autres lois paramétriques dont certains paramètres sont fixes et de valeurs connues, telles que les lois suivantes : Binomiale, Multinomiale et Binomiale Négative [66]. Dans la littérature, il existe d'autres lois paramétriques qui n'appartiennent pas à la famille exponentielle [66] telles que la loi t de Student, la plupart des lois mélanges, ainsi que la famille des distributions uniformes de bornes non fixées.

Dans ce chapitre, nous rappelons, dans un premier temps, les principales définitions et propriétés des familles exponentielles générales, naturelles [7][8][66] et infiniment divisibles. Dans un deuxième temps, nous introduisons une nouvelle caractérisation des familles exponentielles naturelles infiniment divisibles basée sur la fonction trace de la matrice de variance-covariance associée [9][10]. Dans un troisième temps, nous présentons quelques résultats expérimentaux de l'application de la fonction trace pour l'identification de la loi du bruit additif [9] associé à un signal ou à une image.

1.2 Généralités sur les familles exponentielles

Dans cette section, nous commençons par présenter les élements de base des familles exponentielles dans le but d'introduire une nouvelle caractérisation des familles exponentielles naturelles infiniment divisibles. D'abord, nous rappelons quelques propriétés concernant les familles exponentielles générales (FEGs) [103]. Puis, nous présentons la définition des familles exponentielles naturelles (FENs) ainsi que leurs propriétés fondamentales. Enfin, nous rappelons la notion d'infinie divisibilité [46] qui a un intérêt fondamental dans notre travail de recherche pour définir la fonction trace de la matrice de variance-covariance associée à une FEN.

1.2.1 Familles exponentielles générales

Les familles exponentielles générales ont été découvertes en statistique par Koopman [60]. Plusieurs travaux ont montré que ces familles ont un rôle important dans la théorie moderne des statistiques [81][82]. La famille exponentielle générale est caractérisée par le couple (η, T) où η est le vecteur de paramètres et T une statistique.

Dans la suite de cette section, nous citons quelques définitions et théorèmes concernant les familles exponentielles.

Définition 1.2.1. *Un modèle statistique* $\{P_\theta(dx) = P(x,\theta)\nu(dx),\ \theta \in \Theta :$ *ouvert de* $\mathbb{R}^p\}$ *dominé par une mesure* σ*-finie* ν *est dit famille exponentielle générale (FEG) s'il existe des fonctions mesurables* η, T, k *et* h *telles que pour tout* $\theta \in \Theta \subset \mathbb{R}^p$*, on a*

$$P(x,\theta) = h(x)\exp(\langle \eta(\theta), T(x)\rangle - k(\theta)), \tag{1.1}$$

avec

- $T = [T_1, \ldots, T_p]^t : \mathbb{R}^p \longmapsto \mathbb{R}$ *: une fonction mesurable, appelée statistique ;*

- $$\begin{array}{rcl} \eta & : & \Theta \to \mathbb{R}^p \\ \theta & \mapsto & \eta(\theta) = (\eta_1(\theta), \ldots, \eta_p(\theta))^t \end{array}$$

- $k : \Theta \to \mathbb{R}$ *: une fonction mesurable qui ne dépend pas de x ;*

- $h : \mathbb{R} \longmapsto \mathbb{R}^+$ *: une fonction borélienne positive qui ne dépend pas de θ.*

Le terme $\langle ., . \rangle$ représente le produit scalaire dans $\mathbb{R}^p \times \mathbb{R}^p$.

Remarque 1.2.2. *Lorsque $\eta(\theta) = \theta$, on parle de la forme canonique d'une FEG dont la densité de probabilité par rapport à ν est*

$$P(x, \theta) = h(x)\exp(\langle \theta, T(x) \rangle - k(\theta)), \ \forall \theta \in \Theta. \tag{1.2}$$

Exemples

Nous présentons dans cette sous-section quelques exemples de lois univariées qui appartiennent à une FEG.

Exemple 1 : Loi Bêta

La loi Bêta définie sur l'intervalle $[0, 1]$ et de vecteur paramètres $\theta = (a, b) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ a pour densité de probabilité

$$\begin{aligned} P(x, \theta) &= \frac{1}{B(a, b)} x^{a-1}(1-x)^{b-1} \mathbf{1}_{[0,1]}(x) \\ &= \exp((a-1)\ln(x) + (b-1)\ln(1-x) - \ln(B(a, b)))\mathbf{1}_{[0,1]}(x), \end{aligned}$$

avec $B(a, b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ *et* $\Gamma(a) = \displaystyle\int_0^{+\infty} e^{-x} x^{a-1} dx$.

Dans ce cas $\eta(\theta) = (a-1, b-1)^t$, $T(x) = (\ln x, \ln(1-x))^t$, $k(\theta) = \ln B(a, b)$, ν *désigne*

la mesure de Lebesgue sur $[0,1]$ *et* $h(x) = 1$.

Nous présentons dans les figures FIG.1.1, FIG.1.2, F.IG.1.3 et FIG.1.4 les densités de probabilités de la loi Bêta pour différentes valeurs de a et b.

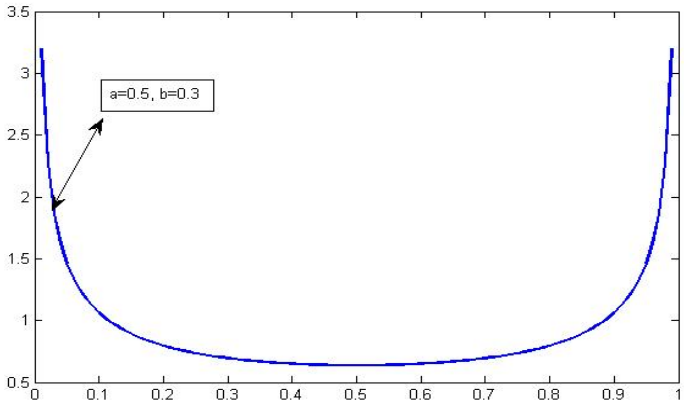


FIG. 1.1 – La distribution est en forme de U pour 0<a<1 et 0<b<1.

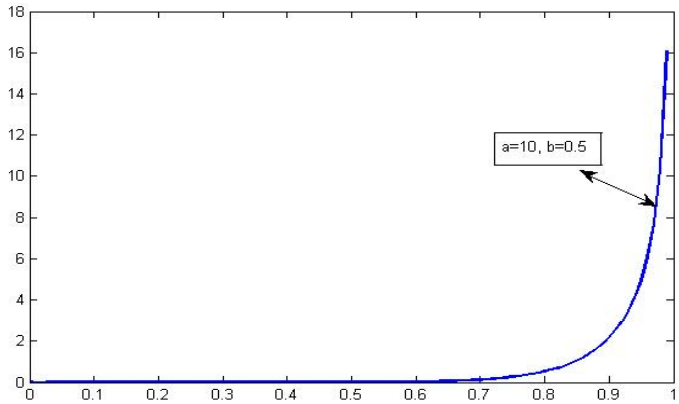


FIG. 1.2 – La distribution est en forme de J pour a>1 et 0<b<1 .

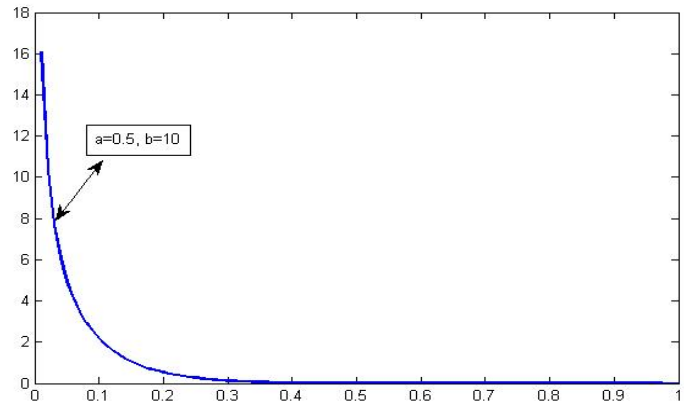


FIG. 1.3 – La distribution est en forme de L pour 0<a<1 et b>1 .

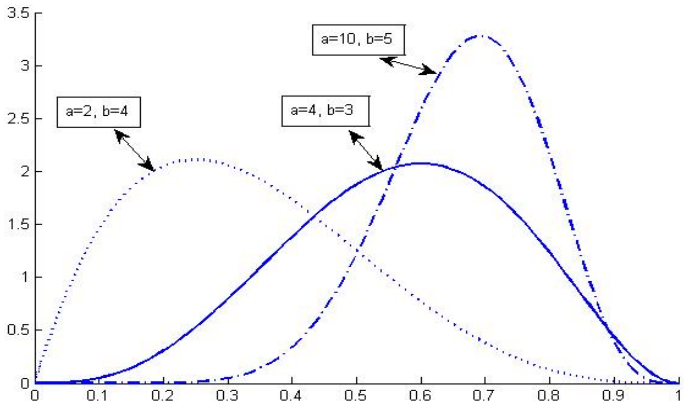


FIG. 1.4 – La distribution est en forme de cloche pour a>1, b>1.

Exemple 2 : Loi Gamma

La loi Gamma de vecteur paramètres $\theta = (a, \lambda) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ *a pour densité de probabilté la fonction*

$$P(x, \theta) = \frac{1}{\lambda^a \Gamma(a)} e^{-\frac{x}{\lambda}} x^{a-1} \mathbf{1}_{\mathbb{R}_+}(x).$$

Dans ce cas $\eta(\theta) = (a-1, \frac{-1}{\lambda})^t$, $T(x) = (\ln x, x)^t$, $k(\theta) = \ln \lambda^a \Gamma(a)$ *et* $h(x) = 1$.

Dans le cas où le paramètre a *est connu, la loi Gamma appartient à la classe des FENs, avec* $\eta(\theta) = \frac{-1}{\lambda}$, $T(x) = x$, $k(\theta) = -a \ln \lambda$, ν *désigne la mesure de Lebesgue sur* $\mathbb{R}_+$ *et* $h(x) = \frac{x^{a-1}}{\Gamma(a)}$.

Nous présentons dans la figure FIG.1.5 la densité de probabilité de la loi Gamma pour différentes valeurs de a et λ.

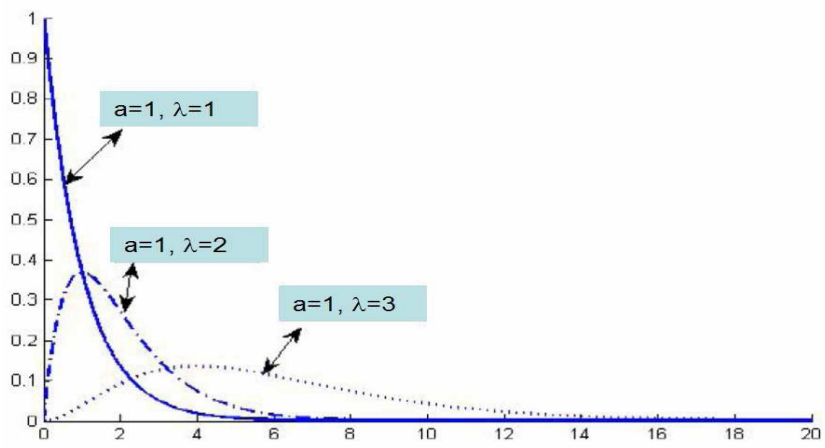


FIG. 1.5 – Variation de la loi Gamma pour différentes valeurs de ses paramètres.

Exemple 3 : Loi Gaussienne

La loi Gaussienne de vecteur paramètres $\theta = (m, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+^*$ *a pour densité de probabilité la fonction*

$$P(x, \theta) = \frac{1}{\sqrt{2\pi}\ \sigma} \exp\Big(- \frac{(x-m)^2}{2\sigma^2} \Big).$$

Dans ce cas $\eta(\theta) = (\frac{m}{\sigma^2}, \frac{-1}{2\sigma^2})^t$, $T(x) = (x, x^2)^t$, $h(x) = \frac{1}{\sqrt{2\pi}}$, ν *désigne la mesure de Lebesgue sur* $\mathbb{R}$ *et* $k(\theta) = \frac{m^2}{2\sigma^2} + \ln \sigma$.

Nous présentons dans la figure FIG.1.6 la densité de probabilité de la loi Gaussienne pour différentes valeurs de m et σ^2.

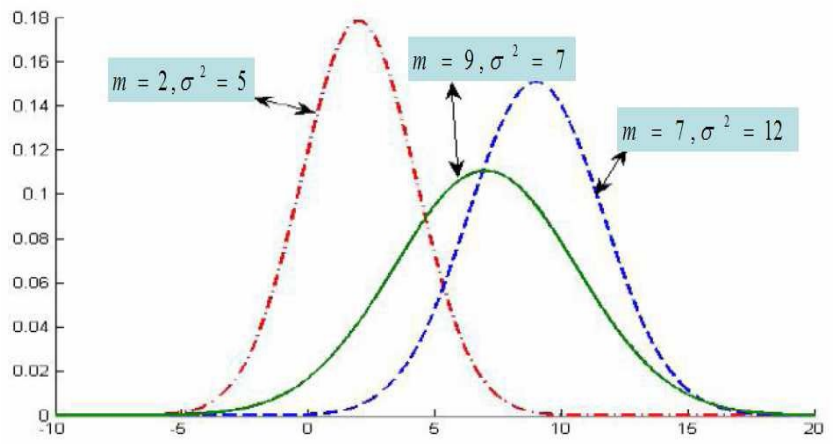


FIG. 1.6 – Variation de la loi Gaussienne pour différentes valeurs de ses paramètres.

Afin d'établir un lien entre les FEGs et la notion d'exhaustivité [66], Darmois [60] a proposé le théorème suivant.

Théorème 1.2.3. ***(Théorème de Darmois)*** *Soit $F = \{P_\theta,\ \theta \in \Theta\}$ une famille exponentille générale engendrée par une mesure σ-finie ν et une statistique T. Soit $(X_1, \ldots, X_n)$ un n-échantillon aléatoire de loi P_θ alors, $\sum_{i=1}^{n} T(X_i)$ est une statistique exhaustive.*

1.2.2 Familles exponentielles naturelles

Les familles exponentielles naturelles ont été introduites par Koopman [60]. Elles représentent une classe des FEGs. Dans la litérature, plusieurs travaux [6][24][25][66][94] ont montré que la FEN présente un intérêt dans la théorie moderne des statistiques.

Dans cette section, nous présentons une caractérisation d'une FEN basée sur la fonction

variance [101]. Pour cette raison nous rappelons la définition de la transformée de Laplace. Dans notre travail de recherche, nous nous plaçons dans un espace vectoriel euclidien $\mathbb{R}^d$.

Définition 1.2.4. *Soit μ une mesure positive définie sur $\mathbb{R}^d$, sa transformée de Laplace L_μ est définie par*

$$\begin{aligned} L_\mu : \mathbb{R}^d &\longrightarrow]0, +\infty] \\ \theta &\longmapsto \int_{\mathbb{R}^d} \exp(\langle \theta, x \rangle) \mu(dx). \end{aligned}$$

Le domaine effectif de L_μ est l'ensemble $D(\mu) = \{\theta \in \mathbb{R}^d \ ; \ L_\mu(\theta) < +\infty\}$.

Par la suite, nous notons par $\Theta(\mu)$ l'intérieur de $D(\mu)$ et nous désignons par $\mathcal{M}(\mathbb{R}^d)$ l'ensemble des mesures μ positives sur $\mathbb{R}^d$ telles que μ n'est pas concentrée sur un hyperplan affine de $\mathbb{R}^d$ et que $\Theta(\mu)$ est non vide.

En utilisant la Définition 1.2.4, nous introduisons la famille exponentielle engendrée par la mesure μ sous forme d'une définition.

Définition 1.2.5. *Soit $\mu \in \mathcal{M}(\mathbb{R}^d)$, l'ensemble*

$$F = F(\mu) = \{P_{\theta,\mu}(dx) = \frac{e^{\langle \theta, x \rangle}}{L_\mu(\theta)} \mu(dx) \ ; \ \theta \in \Theta(\mu)\},$$

est appelé FEN engendrée par μ.

La fonction $k_\mu : \theta \mapsto k_\mu(\theta) = \ln(L_\mu(\theta))$ est appelée fonction génératrice des cumulants de μ. Cette fonction est analytique réelle et strictement convexe sur $\Theta(\mu)$ [66].

Grâce à l'inégalité de Hölder, la fonction k_μ est strictement convexe et infiniment différentiable sur $\Theta(\mu)$, pour $\mu \in \mathcal{M}$ $(\mathbb{R}^d)$. Donc, la différentielle de k_μ notée k'_μ est une fonction injective sur $\Theta(\mu)$ indépendante de μ [66].

Pour définir l'ensemble des mesures qui engendrent une FEN, nous introduisons la proposition suivante.

Proposition 1.2.6. *Si μ et ν sont deux mesures de $\mathcal{M}(\mathbb{R}^d)$, les ensembles $F(\mu)$ et $F(\nu)$ coïncident si et seulement si, il existe un couple $(a, b) \in \mathbb{R}^d \times \mathbb{R}$ tel que*

$$\nu(dx) = e^{\langle a,x\rangle+b}\mu(dx).$$

D'après la Proposition 1.2.6, l'ensemble des mesures qui engendrent F, noté $\mathcal{B}_F$, est donné par

$$\mathcal{B}_F = \{e^{\langle a,x\rangle+b}\mu(dx), (a, b) \in \mathbb{R}^d \times \mathbb{R}\}.$$

D'après cette définition, nous remarquons que $F \subset \mathcal{B}_F$.

Pour définir la fonction variance d'une FEN, nous déterminons ses moments.

Moments d'une FEN

Soit μ une mesure de $\mathcal{M}(\mathbb{R}^d)$, pour tout $\theta \in \Theta(\mu)$, les moments d'ordre n, $n \in \mathbb{N}$, de la loi $P_{\theta,\mu}$ d'une FEN existent et sont définis par

$$\int_{\mathbb{R}^d} x^{\otimes n} P_{\theta,\mu}(dx) = \frac{L_\mu^{(n)}(\theta)}{L_\mu(\theta)},$$

avec $x^{\otimes n}(h_1, \ldots, h_n) = \langle x, h_1\rangle \ldots \langle x, h_n\rangle$, $\forall\ h_1, \ldots, h_n \in \mathbb{R}^d$.

Le vecteur m des moyennes est le gradient de la fonction k_μ au point θ est défini par

$$m = \int_{\mathbb{R}^d} x P_{\theta,\mu}(dx) = k_\mu^{'}(\theta).$$

Nous notons par M_F le domaine des moyennes d'une FEN F engendrée par une mesure $\mu \in \mathcal{M}(\mathbb{R}^d)$. Ce domaine M_F ne dépend pas de la mesure μ. Comme $\mu \in \mathcal{M}(\mathbb{R}^d)$, alors k_μ est strictement convexe sur $\Theta(\mu)$. Ceci implique que $k_\mu^{'}$ est une bijection de $\Theta(\mu)$ sur $M_F = k_\mu^{'}(\Theta(\mu))$ [66]. Nous désignons par ψ_μ la fonction réciproque de $k_\mu^{'}$ définie par

$$\psi_\mu : M_F \to \Theta(\mu) \ \text{ telle que } \quad k_\mu^{'}(\theta) = m \ \Leftrightarrow \theta = \psi_\mu(m). \tag{1.3}$$

A partir de cette fonction ψ_μ, nous obtenons une autre paramétrisation de la famille F appelée paramétrisation par la moyenne et est donnée par

$$F = \{P(m, F) = P(\psi_\mu(m), \mu) \ ; \ m \ \in \ M_F\}. \tag{1.4}$$

Fonction variance d'une FEN

Dans la littérature, la fonction variance [66][78] apparaît comme un outil privilégié efficace pour l'étude d'une FEN. L'importance de cette fonction est due, d'une part, à la caractérisation de la FEN et d'autre part elle a souvent une forme plus simple et plus facile à manipuler que les mesures qui engendrent la FEN ainsi que leurs transformées de Laplace. En effet, grâce à la fonction variance, un certain nombre de classifications des FENs a été réalisé.

D'après la définition des moments d'ordre n de la loi $P_{\theta,\mu}$ d'une FEN F engendrée par une mesure μ de $\mathcal{M}(\mathbb{R}^d)$, la matrice de variance-covariance associée à cette loi est la matrice Hessienne de la fonction k_μ au point θ donnée par l'expression suivante :

$$k''_\mu(\theta) = \int_{\mathbb{R}^d} (x - m) \otimes (x - m) P_{\theta,\mu}(dx)$$

avec $x \otimes y$ représente le produit tensoriel entre x et y.

Définition 1.2.7. *Soit $F = F(\mu)$ une FEN sur $\mathbb{R}^d$ engendrée par une mesure μ de $\mathcal{M}(\mathbb{R}^d)$, la fonction variance V_F de F est définie par*

$$\begin{aligned} V_F \ : \ M_F \ &\to S_+ \\ m \quad &\mapsto \quad V_F(m) = k''_\mu(\psi_\mu(m)) \end{aligned}$$

avec S_+ désigne l'ensemble des matrices carrées $d \times d$ symétriques définies positives.

D'après la définition de la fonction ψ_μ donnée par l'expression 1.3, nous pouvons exprimer

la fonction variance d'une FEN F par

$$V_F(m) = k''_\mu\big(\psi_\mu(m)\big) = \big(\psi_\mu(m)'\big)^{-1}.$$

Le Théorème 1.2.8, dû à Wedderburn [101], prouve que la fonction variance V_F caractérise la famille F.

Théorème 1.2.8. *Si F_1 et F_2 sont deux FENs de domaines des moyennes respectifs, M_{F_1} et M_{F_2}, et si les fonctions variance coïncident sur un ouvert non vide de $M_{F_1} \cap M_{F_2}$, alors les deux FENs sont égales $F_1 = F_2$.*

FENs quadratiques

On dit q'une FEN F est quadratique si sa fonction variance V_F est égale à une fonction polynomiale de degré inférieur ou égale à 2. A une affinité près, Morris [78] a prouvé qu'il existe exactement six familles quadratiques (voir TAB.1.1).

FEN	M_F	$V_F(m)$
Normale	$\mathbb{R}$	p
Poisson	$]0,+\infty[$	m
Gamma	$]0,+\infty[$	$\frac{m^2}{p}$
Binomial	$]0,N[$	$m(1-\frac{m}{N})$
Binomial Négative	$]0,+\infty[$	$m(1+\frac{m}{p})$
Cosinus hyperbolique	$\mathbb{R}$	$m(1+\frac{m^2}{p^2})$

TAB. 1.1 – Les FENs quadratiques.

FENs cubiques

En 1990, Letac et Mora [65] ont montré l'existence de six FENs dont leurs fonctions variance sont des polynômes de degré 3 (cubique, voir TAB.1.2). L'ensemble de ces familles exponentielles naturelles forment la classe de Letac-Mora.

FEN	M_F	$V_F(m)$
Gaussienne inverse	$]0,+\infty[$	$\frac{m^3}{p^2}$
Ressel-Kendall	$]0,+\infty[$	$\frac{m^2}{p}(1+\frac{m}{p})$
Poisson généralisée	$]0,+\infty[$	$m(1+\frac{m}{p})^2$
Binomial négative généralisée	$]0,+\infty[$	$m(1+\frac{m}{p})(1+\frac{a+1}{a}\frac{m}{p})$
Arcsinus stricte	$]0,+\infty[$	$m(1+\frac{m^2}{p^2})$
Arcsinus large	$]0,+\infty[$	$m(1+\dfrac{2m}{p}+\dfrac{a^2+1}{a^2}\dfrac{m^2}{p^2})$

TAB. 1.2 – Les FENs cubiques.

Une généralisation de ces familles quadratiques et cubiques dans le cas multidimensionnel a été réalisée respectivement par Casalis [24], Hassairi et Zarai [49].

Transformation affine sur une FEN

Les transformations affines sont les seules transformations qui conservent la structure d'une FEN [46][49][78].

Désignons par F une FEN et par φ une transformation affine de $\mathbb{R}^d$, c'est à dire

$$\begin{aligned} \varphi : \mathbb{R}^d &\longrightarrow \mathbb{R}^d \\ x &\longmapsto a_\varphi(x)+b, \text{ où } b\in\mathbb{R}^d, a_\varphi\in GL(\mathbb{R}^d). \end{aligned}$$

Si $F = F(\mu)$ alors

$$\varphi(F) = F(\varphi(\mu))$$

est une FEN sur $\mathbb{R}^d$, de domaine des moyennes égal à

$$M_{\varphi(F)} = \varphi(M_F) \text{ et } \varphi(P(m, F)) = P(\varphi(m), \varphi(F)).$$

Si de plus a_φ^* est l'adjoint de a_φ alors, pour tout $m \in M_{\varphi(F)}$, la fonction variance de $\varphi_{(F)}$ est

$$V_{\varphi(F)}(m) = a_\varphi V_F(\varphi^{-1}(m)) a_\varphi^*. \tag{1.5}$$

La proposition suivante étudie une propriété dite condition de symétrie des fonctions variance (Letac [66], 1992).

Proposition 1.2.9. *Soit F une FEN sur $\mathbb{R}^d$ alors, pour tout m dans M_F et tout (α, β) dans $\mathbb{R}^d \times \mathbb{R}^d$, on a*

$$V_F'(m)(V_F(m)(\alpha))(\beta) = V_F'(m)(V_F(m)(\beta))(\alpha). \tag{1.6}$$

On dit alors que V_F vérifie la condition de symétrie.

Preuve. *Comme pour tout m dans M_F*

$$V_F(m) = (\psi_\mu'(m))^{-1},$$

alors, pour tout $h \in \mathbb{R}^d$

$$(\psi_\mu'(m))(h) = (V_F(m))^{-1}(h). \tag{1.7}$$

En calculant la différentielle par rapport à m dans l'équation 1.5 *on a, $\forall\ (h, k) \in \mathbb{R}^d \times \mathbb{R}^d$,*

$$\psi_\mu''(m)(h, k) = -(V_F(m))^{-1}(V_F'(m)(k))V_F(m)^{-1}(h).$$

Le fait que $\psi''_\mu(m)$ est une fonction symétrique, on a

$$\psi''_\mu(m)(h,k) = \psi''_\mu(m)(k,h),$$

alors

$$V_F(m)^{-1}(V'_F(m)(k))(V_F(m)^{-1})(h) = V_F(m)^{-1}(V'_F(m)(h))(V_F(m)^{-1})(k).$$

Si on prend

$$\alpha = (V_F(m))^{-1}(h) \text{ et } \beta = (V_F(m))^{-1}(k),$$

alors on obtient le résultat. □

1.2.3 Steepness

Dans la suite, nous designons par supp(μ) le support de μ (*i.e*; le plus petit ensemble fermé de $\mathbb{R}^d$ verifiant $\mu(\mathbb{R}^d \backslash supp(\mu)) = 0$) et $cs(\mu)$ est l'enveloppe convexe du supp(μ).

Définition 1.2.10. *Soit $\mu \in \mathcal{M}(\mathbb{R}^d)$. On dit que la famille $F = F(\mu)$ est steep si, pour tout $\bar{\theta} \in D(\mu)\backslash\Theta(\mu)$ et tout $\theta \in \Theta(\mu)$, on a*

$$\lim_{\lambda \to 0^+} \langle k'_\mu((1-\lambda)\bar{\theta} + \lambda\theta), \theta - \bar{\theta}\rangle = -\infty.$$

Remarques 1.2.11. *1) Si $D(\mu) = \Theta(\mu)$, alors F est steep. Dans ce cas, certains auteurs [26] disent que $F(\mu)$ est régulière.*

2) La plupart des familles exponentielles standards sont steep.

Le théorème suivant, dû à Barndorff-Nielson [6], caractérise la notion de steepness.

Théorème 1.2.12. *Pour tout μ de $\mathcal{M}(\mathbb{R}^d)$, la FEN $F = F(\mu)$ est steep si et seulement si $int(cs(\mu)) = M_F$, où $int(cs(\mu))$ désigne l'intérieur du support-convexe de μ.*

Notons que pour toute famille exponentielle F steep, le domaine des moyennes M_F est convexe.

Une autre caractérisation de la steepness des familles FENs apparaît dans le thèorème suivant [48].

Théorème 1.2.13. *Soit F une FEN sur $\mathbb{R}^d$. Alors F est steep si et seulement si, pour tout $\overline{m} \in \partial M_F$, on a*

$$\lim_{m \to \overline{m}} \| \psi_\mu(m) \| = +\infty,$$

où ∂M_F désigne la frontière de M_F et $\| \, . \, \|$ désigne une norme sur $\mathbb{R}^d$.

1.2.4 Lois infiniment divisibles

En 1930, P. Lévy [56] a introduit les lois infiniment divisibles. Ces lois ont des propriétés caractéristiques intéressantes et ont fait l'objet de plusieurs travaux [7][46][94]. Le concept des lois infiniment divisibles a été étendu aux cas des distributions multidimensionnelles [93][98]. Les théorèmes décrivant les distributions infiniment divisibles utilisent par conséquent et très souvent les fonctions caractéristiques, les transformées de Laplace ou bien les fonctions génératrices plutôt que les distributions elles mêmes. Une loi infiniment divisible a un rôle important en probabilité et en statistique [41][46][52][57][58][59]. Dans cette section, nous considérons des mesures de probabilités définies sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

Définition 1.2.14. *Une mesure μ est dite infiniment divisible si pour tout $n \in \mathbb{N}^*$, il existe une mesure de probabilite μ_n telle que μ est la puissance $n^{ème}$ de convolution de μ_n (i-e ; $\mu = \mu_n^{*n}$).*

Autrement dit, une variable aléatoire X de loi μ est infiniment divisible si, pour tout entier

$n \in \mathbb{N}^*$, il existe $X_{1,n}, X_{2,n}, ..., X_{n,n}$ des variables aléatoires indépendantes et de même loi μ_n telles que la loi de X égale à la loi de la somme des $X_{i,n}$ ($i.e$; $X \overset{\mathcal{L}}{=} X_{1,n} + \ldots + X_{n,n}$).

Définition 1.2.15. *Soit $\mu \in \mathcal{M}(\mathbb{R}^d)$, on appelle ensemble de Jørgensen de μ l'ensemble*

$$\Lambda(\mu) = \{\lambda > 0 \ ; \ \exists \mu_\lambda \in \mathcal{M}(\mathbb{R}^d) \text{ telle que } \Theta(\mu_\lambda) = \Theta(\mu) \text{ et } L_{\mu_\lambda} = (L_\mu)^\lambda\}.$$

Remarques 1.2.16. 1. *Pour toute mesure μ dans $\mathcal{M}(\mathbb{R}^d)$, si la mesure $\mu_1 = \mu$, alors $1 \in \Lambda(\mu)$.*

2. *$\Lambda(\mu)$ contient l'ensemble $\mathbb{N}^*$ des entiers naturels non nuls.*

On définit dans la proposition 1.2.17 une caractérisation de la notion d'infini divisibilité.

Proposition 1.2.17. *Soit ν une loi de probabilité dans $\mathcal{M}(\mathbb{R}^d)$. Alors ν est infiniment divisible si et seulement si l'ensemble de Jørgensen $\Lambda(\nu) =]0, +\infty[$.*

Exemples :

1. La loi de Poisson $\mathcal{P}(\lambda)$ de paramètre $\lambda \in]0, +\infty[$ de densité $e^{-\lambda}\dfrac{\lambda^k}{k!}$, $k \in \mathbb{N}$ est infiniment divisible. En effet, $\mathcal{P}(\lambda) = \mathcal{P}^{*n}(\frac{\lambda}{n})$.

2. La loi Gaussienne $\mathcal{N}(m, \sigma^2)$ de paramètres $(m, \sigma^2) \in \mathbb{R} \times]0, +\infty[$ est infiniment divisible. En effet, $\mathcal{N}(m, \sigma^2) = \mathcal{N}^{*n}(\frac{m}{n}, \frac{\sigma^2}{n})$.

Afin d'établir une relation entre la mesure de Lévy et la notion d'infini divisibilité, nous introduisons la définition suivante.

Définition 1.2.18. ***(Formule de Lévy-Khintchine).*** *Soit μ une mesure de Borel sur $\mathbb{R}^d \setminus \{0\}$. On dit que μ est une mesure de Lévy si*

$$\int_{\mathbb{R}^d \setminus \{0\}} \min\left(1, \|x\|^2\right) \mu(dx) < \infty.$$

Théorème 1.2.19. *Soit X une variable aléatoire à valeurs dans $\mathbb{R}^d$. Alors X est infiniment divisible s'il existe $b \in \mathbb{R}^d$, une matrice A de taille $d \times d$ et une mesure de Lévy ν_0 tels que la fonction caractéristique de la variable aléatoire X est de la forme*

$$\forall\ u \in \mathbb{R}^d, \quad \varphi_X(u) = \exp\Big(i\langle b, u\rangle - \frac{\langle u, Au\rangle}{2} + \int_{\mathbb{R}^d\setminus\{0\}} \big(e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle \mathbf{1}_{\{|x|\leq 1\}}\big)\nu_0(dx)\Big)$$

Réciproquement, toute application de la forme ci-dessus est la fonction caractéristique d'une variable aléatoire infiniment divisible sur $\mathbb{R}^d$.

Preuve. *Pour la démonstration, le lecteur pourra consulter Sato [93].*

Dans la suite de ce chapitre, nous nous intéressons plus particulièrement aux FENs infiniment divisibles dans lesquelles nous introduisons une nouvelle caractérisation des FENs infiniment divisibles.

1.3 Introduction de la fonction trace dans les FENs infiniment divisibles

Les FENs de lois de probabilité $P_{\theta,\mu}$ offrent une panoplie de modèles très utiles en statistique ainsi qu'en probabilité. Il existe différentes caractérisations des FENs sur $\mathbb{R}^d$. En particulier, Hassairi [46] a étudié la variance généralisée (déterminant de la matrice de variance-covariance) des FENs infiniment divisibles sur $\mathbb{R}^d$. Ghribi et Masmoudi [40] ont montré sous certaines conditions que la variance généralisée caractérise les familles de lois de Bernoullis multivariées. Pommeret [81] a proposé une caractérisation basée sur la diagonalité des matrices de Bhattacharyya associées aux FENs sur $\mathbb{R}^d$. Dans notre travail de recherche, nous proposons une nouvelle caractérisation multidimensionnelle des FENs

infiniment divisibles basée sur la fonction trace de la matrice de variance covariance [10]. Ceci fait l'objet de notre contribution dans ce premier chapitre. La fonction trace fournit plusieurs informations en statistique des FENs, car elle représente la variance totale en terme de l'analyse en composantes principales. De plus, c'est une mesure de variabilité à l'intérieur d'une population multivariée. En pratique, elle intervient dans diverses situations telles que la détection des individus où des observations en dehors d'une limite de surveillance/contrôle où bien la présence des valeurs aberrantes dans un nuage. Dans la suite de ce chapitre, nous montrons l'intérêt de la fonction trace en traitement de signal et des images.

1.3.1 Nouvelle caractérisation d'une FEN infiniment divisible

Soit $F = F(\mu)$ une FEN infiniment divisible sur $\mathbb{R}^d$ engendrée par une mesure de probabilité $\mu \in \mathcal{M}(\mathbb{R}^d)$. Afin de définir la fonction trace, nous rappelons la proposition suivante.

Proposition 1.3.1. *Si μ est une mesure infiniment divisible dans $\mathcal{M}(\mathbb{R}^d)$, alors il existe une matrice positive notée Σ dans $L_s(\mathbb{R}^d)$ et une mesure ν positive telles que*

$$k''_\mu(\theta) = \Sigma + \int_{\mathbb{R}^d} (x \otimes x) e^{\langle \theta, x \rangle} \nu(dx) \tag{1.8}$$

avec, $x \otimes x = x.x^t$ et ν représente la mesure de Lévy associée à la mesure μ.

Preuve. *Pour la démonstration, le lecteur pourra consulter [39], page 342.*

A partir de la Proposition 1.3.1, nous définissons la trace de la fonction variance d'une famille infiniment divisible.

Définition 1.3.2. *Soit $F = F(\mu)$ une FEN infiniment divisible sur $\mathbb{R}^d$, alors la fonction*

trace de la matrice de variance covariance V_F associée à F est définie par

$$tr(k''_\mu(\theta)) = tr(\Sigma) + \int_{\mathbb{R}^d} \| x \|^2 \, e^{\langle \theta, x \rangle} \nu(dx). \tag{1.9}$$

L'application définie par $\theta \longmapsto tr(k''_\mu(\theta))$ est dite fonction trace.

Si de plus la FEN F est steep et que la fermeture de son domaine des moyennes, notée $\overline{M}_F$, possède au moins un point extrême, nous montrons que la matrice Σ est nulle. Ceci, nous permet d'introduire le lemme suivant.

Lemme 1.3.3. *Soit μ une mesure de probabilité infiniment divisible dans $\mathcal{M}(\mathbb{R}^d)$. Supposons que $F = F(\mu)$ est steep et $\overline{M}_F$ admet au moins un point extrême. Alors, il existe une mesure positive ν telle que*

$$k''_\mu(\theta) = \int_{\mathbb{R}^d} (x \otimes x) e^{\langle \theta, x \rangle} \nu(dx). \tag{1.10}$$

Pour démontrer le Lemme 1.3.3., nous avons besoin de rappeler le lemme suivant, dû à Hassairi et Masmoudi [48].

Lemme 1.3.4. *Soit $F = F(\mu)$ une FEN steep et infiniment divisible sur $\mathbb{R}^d$. On suppose que $\overline{M}_F$ admet un point extrême $\overline{m}$. Alors il existe un vecteur u normal extérieur à $\overline{m}$ tel que, pour tout $\theta \in \Theta(\mu)$,*

$$\lim_{\lambda \to +\infty} k''_\mu(\theta + \lambda u) = 0.$$

Preuve du Lemme 1.3.3. *Sans perte de généralités, nous supposons que 0 est un point extrême de $\overline{M}_F = cs(\mu)$.*

Soit $u = (u_1, \ldots, u_d)^t$ un vecteur normal de $\overline{M}_F$ orienté vers l'extérieur au point 0.

Il est clair que

$$M_F \subset \{x \in \mathbb{R}^d;\ \langle u, x \rangle < 0\}$$

et

$$\{x \in \mathbb{R}^d;\ \langle u, x\rangle = 0\} \cap \overline{M}_F = \{0\}.$$

Si $\theta_0 \in \Theta(\mu)$*, alors*

$$\forall \lambda > 0;\ \theta_0 + \lambda u\ \in \Theta(\mu).$$

Comme μ *est infiniment divisible, alors il existe une matrice symétrique* Σ *et une mesure* ν *telles que*

$$k''_\mu(\theta_0 + \lambda u)(h, h) = {}^t h.\Sigma.h + \int_{\mathbb{R}^d} \langle x, h\rangle^2 e^{<\theta_0+\lambda u, x>} \nu(dx).$$

Du fait que

$$\lim_{\lambda \to +\infty} k''_\mu(\theta + \lambda u) = 0,$$

cela implique que

$${}^t h.\Sigma.h = 0 \ et \ \int_{\{x\in\mathbb{R}^d;\ \langle u,x\rangle=0\}} \langle x, h\rangle^2 e^{<\theta_0, x>} \nu(dx) = 0,\ \forall h \in \mathbb{R}^d.$$

Par conséquent,

$$\Sigma = 0.$$

Ainsi,

$$k''_\mu(\theta_0 + \lambda u)(h, h) = \int_{\mathbb{R}^d} \langle x, h\rangle^2 e^{<\theta_0+\lambda u, x>} \nu(dx).$$

□

En s'appuyant sur la Définition 1.3.2 et le Lemme 1.3.3, nous introduisons deux théorèmes qui étudient respectivement le cas où la FEN est infiniment divisible et le cas où la FEN est infiniment divisible et steep [10].

Théorème 1.3.5. *Soient μ_1 et μ_2 deux mesures de probabilités infiniment divisibles dans $\mathcal{M}(\mathbb{R}^d)$. Supposons qu'il existe $\theta_0 \in \Theta(\mu_1) \cap \Theta(\mu_2)$ tel que $k'_{\mu_1}(\theta_0) = k'_{\mu_2}(\theta_0)$, $k''_{\mu_1}(\theta_0) = k''_{\mu_2}(\theta_0)$ et $tr(k''_{\mu_1}(\theta)) = tr(k''_{\mu_2}(\theta))$ $\forall \theta \in \mathrm{B}(\theta_0, R) \subset \Theta(\mu_1) \cap \Theta(\mu_2)$, alors $\mu_1 = \mu_2$.*

Preuve. *Par hypothèse, d'une part, μ_1 et μ_2 sont infiniment divisibles, alors d'après l'équation (1.10) il existe (ν_1, Σ_1) et (ν_2, Σ_2) tels que*

$$k''_{\mu_1}(\theta) = \Sigma_1 + \int_{\mathbb{R}^d} (x \otimes x)\ e^{\langle \theta, x \rangle} \nu_1(dx) \quad et \quad k''_{\mu_2}(\theta) = \Sigma_2 + \int_{\mathbb{R}^d} (x \otimes x)\ e^{\langle \theta, x \rangle} \nu_2(dx).$$

D'autre part, nous avons

$$tr(k''_{\mu_1}(\theta)) = tr(k''_{\mu_2}(\theta)), \forall\ \theta \in \mathrm{B}(\theta_0, R) \subset \Theta(\mu_1) \cap \Theta(\mu_2),$$

où $\mathrm{B}(\theta_0, R)$ désigne la boule ouverte de centre θ_0 et de rayon R.

Cela signifie que

$$\int_{\mathbb{R}^d} \| x \|^2\ \ e^{\langle \theta, x \rangle} \nu_1(dx) = \int_{\mathbb{R}^d} \| x \|^2\ \ e^{\langle \theta, x \rangle} \nu_2(dx) + tr(\Sigma_2 - \Sigma_1),\ \forall\ \theta \in \mathrm{B}(\theta_0, R).$$

En dérivant deux fois par rapport à θ, nous obtenons

$$\int_{\mathbb{R}^d} \| x \|^2\ (x \otimes x)\ e^{\langle \theta, x \rangle} \nu_1(dx) = \int_{\mathbb{R}^d} \| x \|^2\ (x \otimes x)\ e^{\langle \theta, x \rangle} \nu_2(dx),\ \forall\ \theta \in \mathrm{B}(\theta_0, R).$$

En appliquant la fonction trace, nous obtenons

$$\int_{\mathbb{R}^d} \| x \|^4\ \ e^{\langle \theta, x \rangle} \nu_1(dx) = \int_{\mathbb{R}^d} \| x \|^4\ \ e^{\langle \theta, x \rangle} \nu_2(dx),\ \forall\ \theta \in \mathrm{B}(\theta_0, R).$$

D'où

$$\| x \|^4\ \ \nu_1(dx) = \| x \|^4\ \ \nu_2(dx)\ sur\ \mathbb{R}^d \setminus \{0\}.$$

Ceci implique

$$\nu_1 = \nu_2,\ sur\ \mathbb{R}^d \setminus \{0\}.$$

Par conséquent

$$\int_{\mathbb{R}^d} (x \otimes x)\ e^{\langle \theta, x \rangle} \nu_1(dx) = \int_{\mathbb{R}^d} (x \otimes x)\ e^{\langle \theta, x \rangle} \nu_2(dx),$$

Il est clair que

$$k''_{\mu_1}(\theta) = k''_{\mu_2}(\theta) + (\Sigma_2 - \Sigma_1).$$

Puisque

$$k''_{\mu_1}(\theta_0) = k''_{\mu_2}(\theta_0),$$

alors

$$\Sigma_1 = \Sigma_2 \ \textit{et}\ k''_{\mu_1}(\theta) = k''_{\mu_2}(\theta),\ \forall\ \theta\ \in \mathrm{B}(\theta_0, R).$$

Après deux intégrations successives, nous montrons que

$$k_{\mu_1}(\theta) = k_{\mu_2}(\theta) + \langle \beta, \theta \rangle + \gamma;\ \textit{avec}\ \beta\ \in\ \mathbb{R}^d\ \textit{et}\ \gamma\ \in\ \mathbb{R}. \tag{1.11}$$

Comme μ_1 et μ_2 sont deux mesures de probabilité, alors $k_{\mu_1}(0) = k_{\mu_2}(0) = 0$. Ceci implique que $\gamma = 0$.

Par hypothèse $k'_{\mu_1}(\theta_0) = k'_{\mu_2}(\theta_0)$, alors

$$\beta = 0\ \textit{et}\ k_{\mu_1}(\theta) = k_{\mu_2}(\theta),\ \forall\ \theta \in \mathrm{B}(\theta_0, R).$$

La démonstration du Théorème 1.3.5 est achevée. □

La condition initiale du Théorème 1.3.5 est importante pour assurer que la fonction trace caractérise la famille exponentielle naturelle infiniment divisible. En effet, si nous considérons deux lois Gaussiennes centrées $\mu_1 = \mathcal{N}(0, \Sigma_1)$ et $\mu_2 = \mathcal{N}(0, \Sigma_2)$, de matrices de variance-covariance respectivement Σ_1 et Σ_2. De plus, si nous supposons que $\Sigma_1 \neq \Sigma_2$ et $tr(\Sigma_1) = tr(\Sigma_2)$, alors $\forall\ \theta \in \mathbb{R}^d,\ k''_{\mu_1}(\theta) = \Sigma_1 \neq \Sigma_2 = k''_{\mu_2}(\theta),$ ce qui n'est pas le cas car $tr(k''_{\mu_1}(\theta)) = tr(k''_{\mu_2}(\theta))$. Par suite, il n'existe aucun $\theta_0\ \in\ \mathbb{R}^d$ tel que $k''_{\mu_1}(\theta_0) = k''_{\mu_2}(\theta_0)$.

Cela donne un argument raisonnable de l'influence de l'état initial sur la nouvelle caractérisation par la fonction trace.

Dans le Théorème 1.3.6, nous présentons notre deuxième résultat obtenu lorsque nous changeons la condition du Théorème 1.3.5, $k''_{\mu_1}(\theta_0) = k''_{\mu_2}(\theta_0)$, par une nouvelle condition liée au support convexe de la FEN.

Théorème 1.3.6. *Soient μ_1 et μ_2 deux mesures de probabilités steeps et infiniment divisibles dans $\mathcal{M}(\mathbb{R}^d)$. Supposons que les support convexe de μ_1 et μ_2 possèdent au moins un point extrême. Supposons qu'il existe $\theta_0 \in \Theta(\mu_1) \cap \Theta(\mu_2)$ tel que $k'_{\mu_1}(\theta_0) = k'_{\mu_2}(\theta_0)$ et $tr(k''_{\mu_1}(\theta)) = tr(k''_{\mu_2}(\theta)),\ \forall \theta \in \mathrm{B}(\theta_0, R) \subset \Theta(\mu_1) \cap \Theta(\mu_2)$. Alors, $\mu_1 = \mu_2$.*

Preuve. *Par hypothèse, les support convexe de μ_1 et μ_2 possèdent au moins un point extrême. Alors, d'après le Lemme 1.3.3, il existe ν_1 et ν_2 deux mesures positives telles que*

$$k''_{\mu_1}(\theta) = \int_{\mathbb{R}^d} (x \otimes x)\ e^{\langle \theta, x \rangle} \nu_1(dx) \quad et \quad k''_{\mu_2}(\theta) = \int_{\mathbb{R}^d} (x \otimes x)\ e^{\langle \theta, x \rangle} \nu_2(dx).$$

Puisque,

$$tr(k''_{\mu_1}(\theta)) = tr(k''_{\mu_2}(\theta)), \forall\ \theta \in \mathrm{B}(\theta_0, R) \subset \Theta(\mu_1) \cap \Theta(\mu_2).$$

Ceci implique

$$\int_{\mathbb{R}^d} \| x \|^2\ \ e^{\langle \theta, x \rangle} \nu_1(dx) = \int_{\mathbb{R}^d} \| x \|^2\ \ e^{\langle \theta, x \rangle} \nu_2(dx),\ \forall\ \theta \in \mathrm{B}(\theta_0, R).$$

D'où

$$\nu_1 = \nu_2,\ \ on\ \mathbb{R}^d \setminus \{0\}.$$

Alors

$$k''_{\mu_1}(\theta) = k''_{\mu_2}(\theta).$$

Ainsi,

$$k'_{\mu_1}(\theta) = k'_{\mu_2}(\theta) + \beta, \ avec \ \beta \ \in \ \mathbb{R}^d.$$

Comme

$$k'_{\mu_1}(\theta_0) = k'_{\mu_2}(\theta_0),$$

alors,

$$\beta = 0 \ et \ k'_{\mu_1}(\theta) = k'_{\mu_2}(\theta.)$$

Par conséquent,

$$k_{\mu_1}(\theta) = k_{\mu_2}(\theta) + \gamma; \ où \ \gamma \ \in \ \mathbb{R}.$$

Puisque μ_1 et μ_2 sont des probabilités, alors $k_{\mu_1}(0) = k_{\mu_2}(0) = 0$. Ceci implique que $\gamma = 0$.

La démonstration du Théorème 1.3.6 est achevée. □

1.3.2 Exemples

Dans cette section, nous déterminons la fonction trace pour différents exemples de FENs infiniment divisibles (voir [25], [66], [94]).

Famille de lois Multinomiale Négative-Gamma :

Soit $\{e_1, e_2, \ldots, e_d\}$ une base canonique de $\mathbb{R}^d$. On considère la mesure $\mu_0 = \sum_{i=1}^{d} \delta_{e_i}$ et n dans $\mathbb{N}$, avec μ_0^{*n} la puissance $n^{ème}$ de convolution de μ_0 (avec la convention $\mu_0^{*0} = \delta_{e_0}$).

La mesure μ est définie par

$$\mu = \sum_{n=0}^{\infty} \mu_0^{*n} = (\delta_{e_0} - \mu_0)^{*-1}.$$

Il est clair que la transformée de Laplace de μ est de la forme

$$L_\mu(\theta) = (1 - \sum_{i=1}^{d} e^{\theta_i})^{-1}$$

et définie sur $\Theta(\mu) = \{\theta \in \mathbb{R}^d;\ \sum_{i=1}^{d} e^{\theta_i} < 1\}$.

Par conséquent

$$\begin{aligned} tr(k''_\mu(\theta)) &= \frac{\sum_{i=1}^{d} e^{\theta_i}}{1-\sum_{i=1}^{d} e^{\theta_i}} + \frac{\sum_{i=1}^{d} e^{2\theta_i}}{(1-\sum_{i=1}^{d} e^{\theta_i})^2} \\ &= \sum_{i=1}^{d} m_i^2(\theta) + \sum_{i=1}^{d} m_i(\theta), \end{aligned}$$

<u>Famille de lois Gamma-Gaussiennes :</u>

Soit F une FEN de loi Gamma Gaussienne sur $\mathbb{R}^2$ engendrée par une mesure μ définie par

$$\mu(dx, dy) = \frac{1}{\Gamma(a)} x^{a-1} \mathbf{1}_{]0,+\infty[}(x) \frac{1}{\sqrt{2\pi x}} \exp\Big(-\frac{y^2}{2x}\Big) dxdy.$$

La transformée de Laplace de μ est de la forme

$$L_\mu(\theta_1, \theta_2) = \Big(-\theta_1 - \frac{\theta_2^2}{2}\Big)^{-a} \text{ où } \Theta(\mu) = \{\theta = (\theta_1, \theta_2) \in \mathbb{R}^2;\ \theta_1 < 0 \text{ et } \theta_1 + \frac{\theta_2^2}{2} < 0\}.$$

Soit (X, Y) un vecteur aléatoire de $\mathbb{R}^2$ de loi

$$P_{(\theta_1,\theta_2),\mu}(dx, dy) = \frac{\Big(-\theta_1 - \frac{\theta_2^2}{2}\Big)^{-a}}{\Gamma(a)} x^{a-1} \exp(\theta_1 x) \mathbf{1}_{[0,+\infty[}(x) \frac{\exp(\theta_2 y - \frac{y^2}{2x})}{\sqrt{2\pi x}} dxdy,$$

où $a > 0$, $\theta = (\theta_1, \theta_2) \in \Theta(\mu)$. Nous désignons par $\mathcal{L}\ (Y|X)$ la loi conditionnelle de la variable Y sachant la variable X. La variable aléatoire X suit une loi gamma $\gamma_{a,b}$ de paramètre de forme a et de paramètre d'échelle $b = \frac{-1}{\theta_1 + \frac{\theta_2^2}{2}} > 0$. La variable aléatoire $Y|X$ suit une loi Gaussienne $\mathcal{N}(\theta_2 X, X)$.

Le gradient de la fonction k_μ au point $\theta = (\theta_1, \theta_2,)$ est

$$k'_\mu(\theta_1, \theta_2) = \begin{pmatrix} \dfrac{a}{-\theta_1 - \frac{\theta_2^2}{2}} \\ \dfrac{a\theta_2}{-\theta_1 - \frac{\theta_2^2}{2}} \end{pmatrix} = \begin{pmatrix} m_1(\theta) \\ \\ m_2(\theta) \end{pmatrix}.$$

La matrice Hessienne de la fonction k_μ au point θ est de la forme

$$k''_\mu(\theta_1,\theta_2) = \begin{pmatrix} \dfrac{a}{(\theta_1+\frac{\theta_2^2}{2})^2} & \dfrac{a\theta_2}{(\theta_1+\frac{\theta_2^2}{2})^2} \\ \dfrac{a\theta_2}{(\theta_1+\frac{\theta_2^2}{2})^2} & \dfrac{a}{-\theta_1-\frac{\theta_2^2}{2}} + \dfrac{a\theta_2^2}{(\theta_1+\frac{\theta_2^2}{2})^2} \end{pmatrix} = \frac{1}{a}\begin{pmatrix} m_1^2(\theta) & m_1(\theta)m_2(\theta) \\ m_1(\theta)m_2(\theta) & a\ m_1(\theta)+m_2^2(\theta) \end{pmatrix}.$$

Par conséquent, la trace de la fonction variance de la famille F est

$$\begin{aligned} tr(k''_\mu(\theta_1,\theta_2)) &= \frac{a}{-\theta_1-\frac{\theta_2^2}{2}} + \frac{a(1+\theta_2^2)}{(\theta_1+\frac{\theta_2^2}{2})^2} \\ &= \frac{1}{a}(m_1^2\big(\theta)+m_2^2(\theta)\big)+m_1(\theta). \end{aligned}$$

<u>Famille de lois Gaussiennes :</u>

Soit F la FEN engendrée par une loi Gaussienne centrée $\mu = \mathcal{N}(0,\Sigma)$ de matrice de variance-covariance Σ.

La fonction cumulante est définie sur $\mathbb{R}^d$ par

$$k_\mu(\theta) = \frac{1}{2} < \theta, \Sigma\theta > .$$

D'où,

$$tr(k''_\mu(\theta)) = tr(\Sigma) = constante.$$

Nous notons que la fonction trace d'une Famille de lois Gaussienne est une constante et indépendante de sa moyenne.

<u>Famille de lois Inverse Gaussiennes :</u>

La mesure de probabilité μ sur $\mathbb{R}^2$ de la loi Inverse Gaussienne est définie par

$$\mu(dx,dy) = \frac{1}{2\pi x^2}\exp\Big[-\frac{1}{2x}(1+y^2)\Big]\mathbf{1}_{[0,+\infty[}(x)\ dx\ dy.$$

Un calcul assez technique montre que la transformée de Laplace de la mesure μ est définie par

$$L_\mu(\theta_1, \theta_2) = \exp\Big[-\sqrt{-2\theta_1 - \theta_2^2}\,\Big]$$

sur l'ouvert

$$\Theta(\mu) = \{\theta = (\theta_1, \theta_2) \in \mathbb{R}^2;\ 2\theta_1 + \theta_2^2 < 0\}.$$

La fonction cumulante est définie par

$$k_\mu(\theta_1, \theta_2) = -\sqrt{-2\theta_1 - \theta_2^2}.$$

La dérivée de la fonction cumulante est

$$k'_\mu(\theta_1, \theta_2) = \begin{pmatrix} \dfrac{1}{\sqrt{-2\theta_1 - \theta_2^2}} \\ \dfrac{\theta_2}{\sqrt{-2\theta_1 - \theta_2^2}} \end{pmatrix} = \begin{pmatrix} m_1(\theta) \\ \\ m_2(\theta) \end{pmatrix}.$$

Ceci implique que la fonction variance est de la forme

$$\begin{aligned} k''_\mu(\theta_1, \theta_2) &= \begin{pmatrix} \dfrac{1}{(-2\theta_1 - \theta_2^2)^{\frac{3}{2}}} & \dfrac{\theta_2}{(-2\theta_1 - \theta_2^2)^{\frac{3}{2}}} \\ \dfrac{\theta_2}{(-2\theta_1 - \theta_2^2)^{\frac{3}{2}}} & \dfrac{1}{(-2\theta_1 - \theta_2^2)^{\frac{3}{2}}} + \dfrac{\theta_2^2}{(-2\theta_1 - \theta_2^2)^{\frac{3}{2}}} \end{pmatrix} \\ &= \begin{pmatrix} m_1^3(\theta) & m_1^2(\theta)m_2(\theta) \\ m_1^2(\theta)m_2(\theta) & m_1(\theta) + m_1(\theta)m_2^2(\theta) \end{pmatrix}. \end{aligned}$$

Par conséquent,

$$\begin{aligned} tr(k''_\mu(\theta_1, \theta_2)) &= \frac{2}{(-2\theta_1 - \theta_2^2)^{\frac{3}{2}}} + \frac{\theta_2^2}{(-2\theta_1 - \theta_2^2)^{\frac{3}{2}}} \\ &= m_1^3(\theta) + m_1(\theta)(1 + m_2^2(\theta)) \end{aligned}.$$

Famille de lois Gamma :

Soient $a > 0$ et $\rho > 0$ des paramètres réels positifs fixés. On considère la mesure de probabilité μ sur $\mathbb{R}^2$ par

$$\mu(dx, dy) = \frac{(xy)^{a-1}}{\Gamma(a)} f_a(\rho xy) \mathbf{1}_{]0,+\infty[^2}(x, y) dx\ dy,$$

avec $f_a(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(a+k)k!}$, $\forall\ z\ \in\ \mathbb{R}$.

La transformée de Laplace de la mesure μ définie sur l'ouvert $\Theta(\mu) = \{\theta = (\theta_1, \theta_2);\ \theta_1 < 0, \theta_2 < 0, \theta_1\theta_2 - \rho > 0\}$ a pour expression

$$L_\mu(\theta_1, \theta_2) = (\theta_1\theta_2 - \rho)^{-a},$$

La fonction cumulante est de la forme

$$k_\mu(\theta_1, \theta_2) = -a \ln(\theta_1\theta_2 - \rho).$$

La FEN engendrée par μ est une famille de lois Gamma.

Le gradient de $k_\mu(\theta_1, \theta_2)$ est

$$k'_\mu(\theta_1, \theta_2) = \frac{-a}{\theta_1\theta_2 - \rho} \begin{pmatrix} \theta_2 \\ \theta_1 \end{pmatrix} = \begin{pmatrix} m_1(\theta) \\ m_2(\theta) \end{pmatrix}.$$

Un résultat immédiat donne

$$k''_\mu(\theta_1, \theta_2) = \frac{a}{(\theta_1\theta_2 - \rho)^2} \begin{pmatrix} \theta_2^2 & \rho \\ \rho & \theta_1^2 \end{pmatrix} = \frac{1}{a} \begin{pmatrix} m_1^2(\theta) & r m_1^2(\theta) m_2(\theta) \\ r m_1^2(\theta) m_2(\theta) & m_2^2(\theta) \end{pmatrix},$$

avec $r = \frac{\rho}{\theta_1\theta_2}$ représente le coefficient de corrélation de la loi $P_{(\theta,\mu)}$.

Par conséquent,

$$\begin{aligned} tr(k''_\mu(\theta_1, \theta_2)) &= \frac{a}{(\theta_1\theta_2 - \rho)^2}(\theta_1^2 + \theta_2^2) \\ &= \frac{1}{a}(m_1^2(\theta) + m_2^2(\theta)). \end{aligned}$$

Famille de lois de Poisson :

Soit μ une mesure de probabilité sur $\mathbb{R}^d$ définie par

$$\mu(dx) = \sum_{k=0}^{d} \frac{\delta_k(dx)}{k!}, \quad \text{où} \quad k! = \prod_{i=1}^{d} k_i!.$$

La fonction cumulante est définie par

$$k_\mu(\theta) = \sum_{i=1}^{d} e^{\theta_i} \quad sur \quad \Theta(\mu) = \mathbb{R}^d,$$

la famille $F = F(\mu)$ est une famille de lois de Poisson $\mathcal{P}\ (\theta_1) \otimes \ldots \otimes \mathcal{P}\ (\theta_d)$.

Un calcul simple montre que

$$k''_\mu(\theta) = diag(e^{\theta_1}, \ldots, e^{\theta_d}) = diag\Big(m_1(\theta), \ldots, m_d(\theta)\Big). \tag{1.12}$$

$$tr(k''_\mu(\theta)) = \sum_{i=1}^{d} e^{\theta_i} = \sum_{i=1}^{d} m_i(\theta).$$

L'estimateur de la fonction $tr(k''_\mu(\theta))$ est de la forme $\sum_{i=1}^{d} \overline{X}_i$.

Soit Σ une matrice symétrique définie positive donnée par

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{pmatrix}.$$

Soit $X = (X_1, X_2, X_3)$ un vecteur aléatoire qui suit la loi de Poisson sur $\mathbb{R}^3$ définie précédemment. Alors, d'après Letac [66], le vecteur aléatoire $Y = (Y_1, Y_2, Y_3) = \Sigma.X^t$ suit aussi une loi de Poisson. D'après l'équation (1.12), la fonction variance de la FEN engendrée par la loi de Y est définie par

$$V_{F'}(\widetilde{m}(\theta)) = \Sigma.V_F(\Sigma^{-1}m(\theta)).\Sigma,$$

avec $V_F(m(\theta))$ représente la fonction variance de la FEN de loi de Poisson donnée par l'équation (1.12) et $\widetilde{m}(\theta) = (\widetilde{m}_1(\theta), \widetilde{m}_2(\theta), \widetilde{m}_3(\theta))$.

D'où,

$$tr(V_{F'}(\widetilde{m}(\theta))) = (\sigma_{11}^2 + \sigma_{12}^2 + \sigma_{13}^2)\widetilde{m}_1(\theta) + (\sigma_{11}^2 + \sigma_{12}^2 + \sigma_{23}^2)\widetilde{m}_2(\theta) + (\sigma_{13}^2 + \sigma_{23}^2 + \sigma_{33}^2)\widetilde{m}_3(\theta),$$

avec

$$\widetilde{m}_1(\theta) = \frac{(\sigma_{22} - \sigma_{33} - \sigma_{23}^2)m_1(\theta) - (\sigma_{22}\sigma_{33} - \sigma_{13}\sigma_{23})m_2(\theta) + (\sigma_{13}\sigma_{23} - \sigma_{13}\sigma_{22})m_3(\theta)}{2\sigma_{13}\sigma_{12}\sigma_{23} - \sigma_{13}^2\sigma_{22} - \sigma_{33}\sigma_{12}^2 - \sigma_{11}\sigma_{23}^2},$$

$$\widetilde{m}_2(\theta) = \frac{(\sigma_{12}\sigma_{33} - \sigma_{13}^2)m_1(\theta) + (\sigma_{11}\sigma_{33} - \sigma_{13}^2)m_2(\theta) - (\sigma_{11}\sigma_{23} - \sigma_{13}\sigma_{12})m_3(\theta)}{2\sigma_{13}\sigma_{12}\sigma_{23} - \sigma_{13}^2\sigma_{22} - \sigma_{33}\sigma_{12}^2 - \sigma_{11}\sigma_{23}^2}$$

et

$$\widetilde{m}_3(\theta) = \frac{(\sigma_{12}\sigma_{23} - \sigma_{13}\sigma_{22})m_1(\theta) - (\sigma_{11}\sigma_{23} - \sigma_{13}\sigma_{12})m_2(\theta) + (\sigma_{11}\sigma_{22} - \sigma_{12}^2)m_3(\theta)}{2\sigma_{13}\sigma_{12}\sigma_{23} - \sigma_{13}^2\sigma_{22} - \sigma_{33}\sigma_{12}^2 - \sigma_{11}\sigma_{23}^2}.$$

Par conséquent,

$$tr(V_{F'}(\overline{X})) = (\sigma_{11}^2 + \sigma_{12}^2 + \sigma_{13}^2)\overline{X'}_1 + (\sigma_{11}^2 + \sigma_{12}^2 + \sigma_{23}^2)\overline{X'}_2 + (\sigma_{13}^2 + \sigma_{23}^2 + \sigma_{33}^2)\overline{X'}_3,$$

avec

$$\overline{X'}_1 = \frac{(\sigma_{22} - \sigma_{33} - \sigma_{23}^2)\overline{X}_1 - (\sigma_{22}\sigma_{33} - \sigma_{13}\sigma_{23})\overline{X}_2 + (\sigma_{13}\sigma_{23} - \sigma_{13}\sigma_{22})\overline{X}_3}{2\sigma_{13}\sigma_{12}\sigma_{23} - \sigma_{13}^2\sigma_{22} - \sigma_{33}\sigma_{12}^2 - \sigma_{11}\sigma_{23}^2},$$

$$\overline{X'}_2 = \frac{(\sigma_{12}\sigma_{33} - \sigma_{13}^2)\overline{X}_1 + (\sigma_{11}\sigma_{33} - \sigma_{13}^2)\overline{X}_2 - (\sigma_{11}\sigma_{23} - \sigma_{13}\sigma_{12})\overline{X}_3}{2\sigma_{13}\sigma_{12}\sigma_{23} - \sigma_{13}^2\sigma_{22} - \sigma_{33}\sigma_{12}^2 - \sigma_{11}\sigma_{23}^2}$$

et

$$\overline{X'}_3 = \frac{(\sigma_{12}\sigma_{23} - \sigma_{13}\sigma_{22})\overline{X}_1 - (\sigma_{11}\sigma_{23} - \sigma_{13}\sigma_{12})\overline{X}_2 + (\sigma_{11}\sigma_{22} - \sigma_{12}^2)\overline{X}_3}{2\sigma_{13}\sigma_{12}\sigma_{23} - \sigma_{13}^2\sigma_{22} - \sigma_{33}\sigma_{12}^2 - \sigma_{11}\sigma_{23}^2}.$$

<u>Famille de lois Inverse Gaussiennes :</u>

Soit $p > 0$, on considère la mesure de probabilité μ sur $\mathbb{R}^3$ définie par

$$\mu(dx_1, dx_2, dx_3) = \frac{p}{(2\pi)^{\frac{3}{2}} x_1^{\frac{5}{2}}} \exp\Big[-\frac{1}{2x_1}(p^2 + x_2^2 + x_3^2)\Big]\mathbf{1}_{[0,+\infty[}(x_1)\ dx_1\ dx_2\ dx_3.$$

Un calcul assez technique montre que la transformée de Laplace de la mesure μ est définie par

$$L_\mu(\theta_1,\theta_2,\theta_3) = \exp\Big[-p\sqrt{-2\theta_1-\theta_2^2-\theta_3^2}\,\Big]$$

sur l'ouvert

$$\Theta(\mu) = \{\theta = (\theta_1,\theta_2,\theta_3) \in \mathbb{R}^3;\ 2\theta_1+\theta_2^2+\theta_3^2 < 0\}.$$

La fonction cumulante est définie par

$$k_\mu(\theta_1,\theta_2,\theta_3) = -p\sqrt{-2\theta_1-\theta_2^2-\theta_3^2}.$$

La FEN engendrée par μ est une famille de lois Inverse Gaussienne.

La dérivée de la fonction cumulante est

$$k'_\mu(\theta_1,\theta_2,\theta_3) = \begin{pmatrix} \frac{p}{\sqrt{-2\theta_1-\theta_2^2-\theta_3^2}} \\ \frac{p\theta_2}{\sqrt{-2\theta_1-\theta_2^2-\theta_3^2}} \\ \frac{p\theta_3}{\sqrt{-2\theta_1-\theta_2^2-\theta_3^2}} \end{pmatrix} = \begin{pmatrix} m_1(\theta) \\ m_2(\theta) \\ m_3(\theta) \end{pmatrix}.$$

Ceci implique que la fonction variance est de la forme

$$\begin{aligned} k''_\mu(\theta_1,\theta_2,\theta_3) &= \begin{pmatrix} \frac{p}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{3}{2}}} & \frac{p\theta_2}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{3}{2}}} & \frac{p\theta_3}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{3}{2}}} \\ \frac{p\theta_2}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{3}{2}}} & \frac{p}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{1}{2}}} + \frac{p\theta_2^2}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{3}{2}}} & \frac{p\theta_2\theta_3}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{3}{2}}} \\ \frac{p\theta_3}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{3}{2}}} & \frac{p\theta_2\theta_3}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{3}{2}}} & \frac{p}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{1}{2}}} + \frac{p\theta_3^2}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{3}{2}}} \end{pmatrix} \\ &= \frac{m_1(\theta)}{p^2} \begin{pmatrix} m_1^2(\theta) & m_1(\theta)m_2(\theta) & m_1(\theta)m_3(\theta) \\ m_1(\theta)m_2(\theta) & m_2^2(\theta)+p^2 & m_2(\theta)m_3(\theta) \\ m_1(\theta)m_3(\theta) & m_2(\theta)m_3(\theta) & m_3^2(\theta)+p^2 \end{pmatrix}. \end{aligned}$$

Par conséquent,

$$\begin{aligned} tr(k''_\mu(\theta_1,\theta_2,\theta_3)) &= \frac{2p}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{1}{2}}} + \frac{p(1+\theta_2^2+\theta_3^2)}{(-2\theta_1-\theta_2^2-\theta_3^2)^{\frac{3}{2}}} \\ &= \frac{m_1(\theta)}{p^2}\Big[m_1^2(\theta)+m_2^2(\theta)+m_3^2(\theta)+2p^2\Big]. \end{aligned}$$

<u>Famille de lois Gamma-Gaussiennes :</u>

Soit F une FEN de loi Gamma Gaussienne sur $\mathbb{R}^3$ engendrée par une mesure μ de transformée de Laplace (Casalis [25]) définie par

$$L_\mu(\theta_1,\theta_2,\theta_3) = \Big(-\theta_1 - \frac{\theta_2^2}{2} - \frac{\theta_3^2}{2}\Big)^{-1} \text{ où } \Theta(\mu) = \{\theta = (\theta_1,\theta_2,\theta_3) \in \mathbb{R}^3;\ \theta_1 < 0 \text{ et } \theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2} < 0\}.$$

Sa fonction cumulante est définie par

$$k_\mu(\theta_1,\theta_2,\theta_3) = -\ln(-\theta_1 - \frac{\theta_2^2}{2} - \frac{\theta_3^2}{2})$$

Le gradient de la fonction k_μ est

$$k'_\mu(\theta_1,\theta_2,\theta_3) = \begin{pmatrix} \frac{1}{-\theta_1-\frac{\theta_2^2}{2}-\frac{\theta_3^2}{2}} \\ \frac{\theta_2}{-\theta_1-\frac{\theta_2^2}{2}-\frac{\theta_3^2}{2}} \\ \frac{\theta_3}{-\theta_1-\frac{\theta_2^2}{2}-\frac{\theta_3^2}{2}} \end{pmatrix} = \begin{pmatrix} m_1(\theta) \\ m_2(\theta) \\ m_3(\theta) \end{pmatrix}.$$

La matrice Hessienne de la fonction k_μ au point $\theta = (\theta_1, \theta_2, \theta_3)$ est

$$k''_\mu(\theta) = \begin{pmatrix} \frac{1}{(\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2})^2} & \frac{\theta_2}{(\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2})^2} & \frac{\theta_3}{(\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2})^2} \\ \frac{\theta_2}{(\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2})^2} & \frac{-1}{\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2}} + \frac{\theta_2^2}{(\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2})^2} & \frac{\theta_2\theta_3}{(\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2})^2} \\ \frac{\theta_3}{(\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2})^2} & \frac{\theta_2\theta_3}{(\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2})^2} & \frac{-1}{(\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2})^2} + \frac{\theta_3^2}{(\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2})^2} \end{pmatrix}$$

$$= \begin{pmatrix} m_1^2(\theta) & m_1(\theta)m_2(\theta) & m_1(\theta)m_3(\theta) \\ m_1(\theta)m_2(\theta) & -m_1(\theta) + m_2^2(\theta) & m_2(\theta)m_3(\theta) \\ m_1(\theta)m_3(\theta) & m_2(\theta)m_3(\theta) & -m_1(\theta) + m_3^2(\theta). \end{pmatrix}.$$

Par conséquent, la trace de la fonction variance de la famille F est

$$tr(k''_\mu(\theta_1, \theta_2, \theta_3)) = \frac{-2}{\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2}} + \frac{1 + \theta_2^2 + \theta_3^2}{(\theta_1 + \frac{\theta_2^2}{2} + \frac{\theta_3^2}{2})^2}$$

$$= m_1^2(\theta) + m_2^2(\theta) + m_3^2(\theta) - 2m_1(\theta).$$

1.3.3 Propriétés asymptotiques de la fonction trace

Lorsque nous cherchons à étudier une suite de mesures de probabilités provenant de la répétition d'une expérience, une méthode de modélisation consiste à supposer que ces mesures sont des réalisations de variables aléatoires indépendantes équi-distribuées. Comprendre ces mesures et la façon dont elles sont distribuées revient à étudier la loi de probabilité de la variable aléatoire sous-jacente. Nous modélisons le phénomène de la manière suivante : soit $F = F(\mu)$ une FEN infiniment divisible sur $\mathbb{R}^d$ engendrée par une mesure μ et soient $X_1, \ldots, X_n$, n variables aléatoires indépendantes identiquement distribuées (i.i.d) de loi $P_{\theta,\mu} \in F$. Dans ce contexte, reconnaître un estimateur de la fonction trace de la FEN

infiniment divisible F revient à connaître la loi $P_{\theta,\mu} \in F$. Lorsque nous n'avons pas d'idée a priori sur la forme particulière que peut prendre la loi $P_{\theta,\mu} \in F$, nous construisons un estimateur de $P_{\theta,\mu} \in F$. Il s'agit de reconstruire une fonction trace, c'est à dire un estimateur de la fonction trace. Le problème est alors dit non-paramétrique. Dans ce contexte, nous nous sommes intéressés à l'estimation de la fonction trace ainsi que ses propriétés asymptotiques.

La loi forte des grands nombres permet d'affirmer que $tr\big(V_F(\overline{X})\big)$ est un estimateur convergent de la fonction trace $tr\big(k''_\mu(\theta)\big) = tr\big(V_F(m(\theta))\big)$, où $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$. Ce résultat vient du fait que, d'après la loi des grands nombres, lorsque $n \longrightarrow +\infty$, on a $\overline{X}$ converge presque sûrement vers $m(\theta)$. Alors $tr\big(V_F(\overline{X})\big)$ converge presque sûrement vers $tr\big(V_F(m(\theta))\big)$.

Notons que $tr\big(V_F(\overline{X})\big)$ est un estimateur de $tr\big(V_F(m(\theta))\big)$. Il est possible d'obtenir des intervalles de confiance et de tester l'adéquation des données sur différentes lois. Dans tout ce qui suit, nous utilisons $tr\big(V_F((\overline{X}))\big)$ comme estimateur de la fonction trace $tr\big(k''_\mu(\theta)\big) = tr\big(V_F(m(\theta))\big)$, où $V_F(m(\theta)) = \Big(V_{i,j}(m(\theta))\Big)_{1\leq i,j\leq d}$.

Dans les tableaux TAB.1.3 et TAB.1.4 , nous illustrons quelques exemples des estimateurs de la fonction trace associés à des lois de probabilités définies dans la section précédente.

FEN	$tr\big(V_F(\overline{X})\big)$
Négative Multinomiale-Gamma	$tr(BNMG) = \overline{X}_1^2 + \overline{X}_2^2 + \overline{X}_1 + \overline{X}_2.$
Gamma-Gaussienne de paramètre de forme $a > 0$	$tr(BGG) = \frac{1}{a}\big(\overline{X}_1^2 + \overline{X}_2^2\big) + \overline{X}_1,\ a > 0.$
Gaussienne	$tr(BG) = tr(S_n^2) = \frac{1}{n-1}\sum_{i=1}^{n} tr\Big((X_{i,1} - \overline{X_1}).(X_{i,2} - \overline{X_2})^t\Big);$ trace de la matrice de variance-covariance empirique .
Inverse Gaussienne	$tr(BIG) = 2\overline{X}_1^3 + \overline{X}_1\overline{X}_2^2.$
Gamma de paramètre de forme $a > 0$	$tr(BGa) = \frac{1}{a}\big(\overline{X}_1^2 + \overline{X}_2^2\big),\ a > 0.$
Poisson	$tr(BP) = \overline{X}_1 + \overline{X}_2.$

TAB. 1.3 – Différents estimateurs de la fonction trace, pour $d = 2$.

FEN	$tr\big(V_F(\overline{X})\big)$
Gamma-Gaussienne	$tr(GG) = \overline{X}_1^2 + \overline{X}_2^2 + \overline{X}_3^2 - 2\overline{X}_1.$
Gaussienne	$tr(G) = tr(S_n^2) = \frac{1}{n-1}\sum_{i=1}^{n} tr\Big((X_{i,1} - \overline{X_1}).(X_{i,2} - \overline{X_2})^t\Big);$ trace de la matrice de variance-covariance empirique .
Inverse Gaussienne	$tr(IG) = \frac{\overline{X}_1}{p^2}\Big[\overline{X}_1^2 + \overline{X}_2^2 + \overline{X}_3^2 + 2p^2],\ p > 0.$
Poisson	$tr(P) = (\sigma_{11}^2 + \sigma_{12}^2 + \sigma_{13}^2)\overline{X_1'} + (\sigma_{11}^2 + \sigma_{12}^2 + \sigma_{23}^2)\overline{X_2'}$ $+(\sigma_{13}^2 + \sigma_{23}^2 + \sigma_{33}^2)\overline{X_3'}.$

TAB. 1.4 – Différents estimateurs de la fonction trace, pour $d = 3$.

Convergence

Nous prouvons que lorsque la taille de l'échantillon n est suffisamment grande, l'estimateur de la fonction trace $tr\big(V_F(\overline{X})\big)$ est asymptotiquement normal.

Théorème 1.3.7. *Lorsque $n \longrightarrow +\infty$, nous avons*

$$\sqrt{n}\Big(tr\big(V_F(\overline{X})\big) - tr(V_F(m(\theta)))\Big) \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \sigma^2(m(\theta))\Big), \text{ en loi.}$$

Pour démontrer le Théorème 1.3.7, nous avons besoin de rappeler la Méthode Delta sous forme d'un lemme [104].

Lemme 1.3.8. ***Méthode Delta.*** *Soient $\phi : \mathbb{R}^k \longrightarrow \mathbb{R}^p$ une fonction différentiable et r_n une suite de nombres positifs tendant vers $+\infty$. On se donne une suite $(U_n)_{n\in\mathbb{N}}$ de vecteurs aléatoires dans $\mathbb{R}^k$. Si, pour une variable aléatoire U, on a*

$$r_n\Big(U_n - \theta) \xrightarrow{\mathcal{L}} U,$$

alors

$$r_n\Big(\phi(U_n) - \phi(\theta)\Big) \xrightarrow{\mathcal{L}} \nabla\phi(\theta)U.$$

Preuve du Théorème 1.3.7. *D'une part, pour $n \longrightarrow +\infty$, un résultat immédiat du théorème centrale limite est que la variable aléatoire*

$$\sqrt{n}\Big(\begin{pmatrix} \overline{X}_1 \\ \vdots \\ \overline{X}_d \end{pmatrix} - \begin{pmatrix} m_1(\theta) \\ \vdots \\ m_d(\theta) \end{pmatrix}\Big)$$

converge en loi vers une variable aléatoire $\mathcal{U}$, où $\mathcal{U}$ suit une loi normale centrée $\mathcal{N}\Big(0, V_F(m(\theta))\Big)$.

D'autre part, soit la fonction H définie par

$$H : \mathbb{R}^d \to \mathbb{R}^d$$

$$m(\theta) = (m_1(\theta), \cdots, m_d(\theta)) \mapsto (V_{1,1}\big(m(\theta)\big), \cdots, V_{d,d}\big(m(\theta)\big)),$$

avec $(V_{i,i}(m(\theta))$ *représente le* $i^{ème}$ *élément de la diagonale de la matrice de variance-covariance* $V_F(m(\theta))$*. Puisque la fonction* H *est différentiable en* $m(\theta)$*, nous pouvons appliquer la méthode Delta. D'où*

$$\sqrt{n}\Big(H(\overline{X}) - H(m(\theta))\Big) \xrightarrow{\mathcal{L}} H'(m(\theta))\mathcal{U},$$

avec $H'(m(\theta))\mathcal{U} \sim \mathcal{N}\Big(0, {}^t H'(m(\theta))\ V_F(m(\theta))\ H'(m(\theta)))\Big)$ *et* $H'(m(\theta))$ *représente la matice Jacobienne de* $H(m(\theta))$*.*

Soit $e = (1, 1, \ldots, 1)^t$*, lorsque* $n \longrightarrow +\infty$*, d'après ce qui précède nous montrons que*

$$\sqrt{n}\Big(\langle H(\overline{X}), e\rangle - \langle H(m(\theta)), e\rangle\Big) \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \sigma^2(m(\theta))\Big),$$

avec $\sigma^2(m(\theta)) = {}^t e\ {}^t H'(m(\theta))\ V_F(m(\theta))\ H'(m(\theta))\ e$*.*

En utilisant le fait que

$$\langle H(\overline{X}), e\rangle = \sum_{i=1}^{d} V_{ii}(\overline{X}) = tr\big(V_F(\overline{X})\big),$$

et

$$\langle H(m(\theta)), e\rangle = \sum_{i=1}^{d} V_{ii}(m(\theta)) = tr(V_F(m(\theta))).$$

La démonstration du Théorème 1.3.7 est achevée. □

Intervalle de confiance

La construction des intervalles de confiance nécessite la connaissance de la loi des estimateurs à partir desquels ils sont construits. En dehors du cas des échantillons gaussiens, ce n'est pas toujours possible. De plus, il est fréquent que les calculs de lois, effectivement implémentables pour des petits échantillons, deviennent infaisables pour un échantillon de grande taille. On cherche alors à remplacer la loi de l'estimateur par une approximation

plus simple, permettant de construire des intervalles de confiance dont le niveau n'est garanti que asymptotiquement.

Soient $F = F(\mu)$ une FEN infiniment divisible sur $\mathbb{R}^d$ engendrée par une mesure μ et $X_1, \ldots, X_n$, n variables aléatoires indépendentes et identiquement distribuées de loi $P_{\theta,\mu} \in F$. Dans le corollaire suivant nous construisons un intervalle de confiance asymptotique de la fonction trace $tr(V_F(m(\theta)))$.

Corollaire 1.3.9. *Un intervalle de confiance de niveau asymptotique* $(1-\alpha)$ *pour* $tr(V_F(m(\theta)))$ *est de la forme*

$$\Big(tr(V_F(\overline{X})) - u_{\frac{\alpha}{2}}\frac{\widehat{\sigma}(\overline{X})}{\sqrt{n}},\ tr(V_F(\overline{X})) + u_{\frac{\alpha}{2}}\frac{\widehat{\sigma}(\overline{X})}{\sqrt{n}}\Big),$$

où $P(|U| < u_{\frac{\alpha}{2}}) = 1-\alpha$ *et* $U \sim \mathcal{N}(0,1)$.

Preuve. *D'après le résultat du Théorème 1.3.7 et par application du théorème de Slutsky [26], nous montrons que*

$$\sqrt{n}\Big[\sigma(\overline{X})\Big]^{-1}\Big(tr\big(V_F(\overline{X})\big) - tr(V_F(m(\theta)))\Big) \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0,1\Big).$$

Par suite, un intervalle de confiance de niveau asymptotique $(1-\alpha)$ *pour* $tr(V_F(m(\theta)))$ *est de la forme*

$$tr(V_F(\overline{X})) - u_{\frac{\alpha}{2}}\frac{\widehat{\sigma}(\overline{X})}{\sqrt{n}} < tr(V_F(m(\theta))) < tr(V_F(\overline{X})) + u_{\frac{\alpha}{2}}\frac{\widehat{\sigma}(\overline{X})}{\sqrt{n}},$$

où $P(|U| < u_{\frac{\alpha}{2}}) = 1-\alpha$ *et* $U \sim \mathcal{N}(0,1)$. □

1.4 Application de la fonction trace

Les statistiques et les probabilités occupent une place importante dans beaucoup de domaines d'applications tels que le traitement numérique du signal et des images pour ca-

ractériser les processus qui les génèrent. Par exemple, un usage principal du traitement numérique du signal et des images est de réduire les interférences, le bruit, et d'autres composants indésirables dans les données acquises. Réduire l'effet du bruit additif est un problème fondamental et récure dans la plupart des applications en traitement du signal et des images [92]. C'est, d'autre part, un problème théorique central en théorie statistique que ce soit pour des objectifs d'estimation où de détection. Le cas où le signal est un bruit ou un signal source de nature impulsive s'avère particulièrement intéressant à la fois sur le plan théorique [96] et sur le plan pratique comme la réduction de l'effet de bruit atmosphérique en détection Radar ou en communication Haute Fréquence (HF) et l'effet des valeurs aberrantes sur les procédures de traitement statistique d'un signal observé. La détermination de la loi du bruit est un problème simple et compliqué à la fois. Pour le résoudre, nous nous sommes basés sur la méthode de simulation Boostrap non paramétrique [17] modifiée et nous avons utilisé notre nouveau critère de sélection, pour estimer la loi du bruit, basée sur l'estimateur de la fonction trace.

1.4.1 En traitement du signal

Dans le cadre du traitement du signal, la problématique majeure est de réussir à isoler des perturbations (vent, parole, nuage,...) d'un signal additif reçu par une cible. En effet, ces perturbations, également appelées bruit, peuvent provenir soit du système, soit de l'environnement extérieur à celui-ci (brouilleur, échos de sol ou de mer, environnement urbain, interférences, ...). Dans ce cadre, le bruit en général est considéré comme étant un processus aléatoire. Un besoin commun dans le traitement de signal est de générer des signaux additifs qui ressemblent à divers types du bruit aléatoire.

Représentation statistique du signal

Le modèle statistique du signal additif est représenté sous la forme suivante :

$$X = m + \varepsilon, \tag{1.13}$$

où X représente le signal de sortie qui est obtenu comme une fonction linéaire du signal d'entrée m et du bruit ε.

Dans de nombreuses études menées depuis des dizaines d'années, les méthodes d'analyse et les algorithmes associés supposent que le bruit est modélisé comme un bruit Gaussien blanc simple (c'est un bruit qui suit une loi normale) dont les statistiques et les caractéristiques spectrales sont bien connues [110]. Mais, la réalité est très loin de cette hypothèse. En particulier, il existe un bruit impulsif. Contrairement au bruit blanc Gaussien, il est impulsif, non statistique standard et généralement non-stationnaire, non gaussien, avec un comportement de haute fréquence et complexe.

Dans notre travail, nous proposons une méthode pour approximer la loi de Bruit en se basant sur la fonction trace.

Simulation

Nous supposons dans notre étude de simulation, que la distribution du signal d'entré de vecteur moyen m connu et la distribution du vecteur bruit ε appartient à une FEN infiniment divisible engendrée par une mesure $\mu \in \mathcal{M}(\mathbb{R}^2)$. C'est à dire ($\varepsilon \sim P_{\theta,\mu}$).

Nous utilisons les estimateurs proposés de $tr(k''_\mu(\theta)) = tr(V_F(m(\theta)))$, pour les quatre FENs infiniment divisibles représentées dans le TAB.1.3 (Gamma-Gaussienne, Gaussienne, Inverse-Gaussienne, Gamma), afin de déterminer la loi du bruit ε.

Considérons un échantillon aléatoire $\{\varepsilon_1, \ldots, \varepsilon_n\}$ de loi $P_{\theta,\mu}$. Nous adoptons l'approche Boostrap non paramétrique [21] avec une petite modification au niveau de la création de B échantillons Boostrap à partir de l'échantillon initial. C'est à dire, au lieu de créer chaque nouvel échantillon à partir d'un tirage avec remise de l'échantillon initial, nous le remplaçons par une autre méthode définie comme suit : Chaque fois, nous perturbons l'échantillon initial en ajoutant une variable aléatoire u uniforme sur $(0,1)^2$, c'est à dire en créant un nouvel échantillon $\{\varepsilon_1^*, \cdots, \varepsilon_n^*\}$ avec $\varepsilon_i^* = \varepsilon_i + u_i$.

Pour chaque nouvel échantillon obtenu, nous calculons l'estimateur empirique, $tr(V_F(\overline{\varepsilon})) = trep$, correspondant à chaque FEN infiniment divisible (Gamma-Gaussienne, Gaussienne, Inverse-Gaussienne, Gamma).

Pour déterminer l'estimateur de la distribution du bruit ε, nous avons choisi comme critère de comparaison l'erreur quadratique moyenne (EQM) définie par :

$$\mathrm{EQM}(trtp, trep) = \frac{1}{B}\sum_{j=1}^{N}\Big(tr(V_F(m^*(\theta))) - tr(V_F(\overline{\varepsilon_j^*}))\Big)^2, \tag{1.14}$$

où $m^*(\theta) = m(\theta) + (0.5, 0.5)^t$ et $tr(V_F(m(\theta)))) = trtp$ représente la trace de la matrice de variance-covariance de la loi $P_{\theta,\mu}$ avec des valeurs exactes de ces paramètres.

Nous répétons cette expérience pour différentes valeurs de θ.

Nous avons généré les données selon les quatre lois (Gamma-Gaussienne, Gaussienne, Inverse-Gaussienne, Gamma) en choisissant les valeurs des paramètres suivants :

- $a = 2$ pour le modèle Gamma,
- $n = 5000$ la taille de l'échantillon pour chaque loi,
- $B = 100$ le nombre de Boostrap,
- Nous avons choisi cinq valeurs de θ afin de réaliser cette expérience.

Les résultats obtenus sont présentés dans les tableaux TAB 1.5, TAB 1.6, TAB 1.7 et TAB 1.8.

Comme première expérience, nous avons supposé que la loi $P_{\theta,\mu}$ est une Gamma-Gaussienne. Le Tableau TAB 1.5, indique que la valeur de l'EQM est minimale lorsque nous calculons avec l'estimateur de la fonction trace d'une loi Gamma-Gaussienne, pour différentes valeurs de θ. Par suite, la loi du bruit ε est qualifiée par Gamma-Gaussienne.

Dans la seconde expérience, la loi de $P_{\theta,\mu}$ est Gaussienne de moyenne θ et de matrice de variance-covariance $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, avec $\rho = 0.3$. Nous répétons la même procédure que précédemment. Le Tableau TAB 1.6 indique que la valeur de l'EQM calculée avec l'estimateur de la loi Gaussienne est plus faible que celles calculées pour les autre lois. Alors, la loi du bruit ε est Gaussienne.

Dans la troisième expérience, nous avons répété la même procédure avec une loi Inverse Gaussienne. Le Tableau TAB 1.7 indique que la loi du bruit ε est Inverse Gaussienne.

Dans la dernière expérience, nous avons répété la même procédure avec une loi Gamma. le Tableau TAB 1.8 indique que la loi du bruit ε est Gamma.

Le choix de cinq valeurs différentes de θ ne nous permet pas d'affirmer que les résultats obtenus sont efficaces. Pour cette raison, nous avons généralisé notre étude pour différentes valeurs de θ obtenues de façon aléatoire pour chaque loi considérée et nous avons répété la même procédure que précédemment. Dans ce cas, notre étude de simulation montre que la loi du bruit ε dans le cas d'un modèle Gamma-Gaussien converge vers une loi Gamma-Gaussienne avec un pourcentage égal à 97%. Puis, le bruit ε dans le cas d'un modèle Gaussien converge vers une loi Gaussienne, avec un pourcentage égal à 100%. Dans le cas d'un modèle Inverse Gaussien, la loi du bruit ε converge vers une loi Inverse Gaussienne

avec un pourcentage égal à 96%. Enfin, la loi du bruit ε dans le cas d'un modèle Gamma converge vers une loi Gamma avec un pourcentage égal à 97%.

θ_1	θ_2	$trtp$	EQM($trtp$, $tr(BGG)$)	EQM($trtp$, $tr(BG)$)	EQM($trtp$, $tr(BIG)$)	EQM($trtp, tr(BGa)$)
-3	2	12	0.0138	0.3874	6.5072	0.2682
-2	1	3.11	0.0027	0.2174	1.2993	0.1809
-10	2.5	0.59	0.0070	0.1924	1.3064	0.1834
-1	0	5.55	0.0158	0.2906	4.099	0.2916
-10	3	1.02	0.0011	0.1247	0.1327	0.0840

TAB. 1.5 – Différentes valeurs des EQM selon le modèle Gamma-Gaussien bivarié.

θ_1	θ_2	$trtp$	EQM($trtp$, $tr(BGG)$)	EQM($trtp$, $tr(BG)$)	EQM($trtp$, $tr(BIG)$)	EQM($trtp, tr(BGa)$)
0	0	2	0.1013	0.02874	0.1650	0.1526
1	2	2	0.7921	0.0069	1.3393	0.6441
-1	4	2	1.8007	0.0213	1.3064	1.8550
5	4	2	24.0158	0.0290	148.0990	23.2916
30	5	2	350.2154	0.0059	13×10^3	347.1600

TAB. 1.6 – Différentes valeurs des EQM selon le modèle Gaussien bivarié.

θ_1	θ_2	$trtp$	EQM($trtp$, $tr(BGG)$)	EQM($trtp$, $tr(BG)$)	EQM($trtp$, $tr(BIG)$)	EQM($trtp, tr(BGa)$)
-2	1	0.57	0.1766	0.2874	0.01850	0.2526
-1	0	0.97	0.3150	0.4148	0.2331	0.4416
-3	1.5	0.58	0.1453	0.2813	0.0142	0.2457
-5	3	11	2.1586	2.3687	1.9343	2.2696
-15	5	2.41	0.3519	0.6159	0.1608	0.4067

TAB. 1.7 – Différentes valeurs des EQM selon le modèle Inverse Gaussien bivarié.

θ_1	θ_2	$trtp$	EQM($trtp$, $tr(BGG)$)	EQM($trtp$, $tr(BG)$)	EQM($trtp$, $tr(BIG)$)	EQM($trtp, tr(BGa)$)
-1.6	-5	6.13	0.4991	0.4632	33.1650	0.3152
-1	-3	28	1.7300	1.92575	201.3393	0.7411
-4	-4	3.98	0.4815	0.5213	3.3064	0.2758
-1.5	-10	4.13	0.3982	0.2907	32.2767	0.1447
-10	-5	2.11	0.3982	0.3255	0.3896	0.2004

TAB. 1.8 – Différentes valeurs des EQM selon le modèle Gamma bivarié.

1.4.2 En traitement d'images couleur

En quelques années, les images digitales ont envahi notre vie quotidienne. Les caméras numériques permettent en effet de saisir et de manipuler bien plus facilement les images et les films. Leur qualité est maintenant équivalente et souvent supérieure à celles des images obtenues par les procédés photochimiques. Les images digitales offrent de plus l'immense avantage de pouvoir être transmises, améliorées et stockées sur des supports informatiques. Toutefois, les images numériques continuent à présenter certains défauts des photographies. Elles ont en particulier un bruit de fond qui limite leur lisibilité. La technologie digitale

cherche à réduire au maximum ce bruit. Pour comprendre la nature du bruit visuel, il faut rappeler qu'une image numérique se présente comme un rectangle divisé en petits carrés, les pixels. Chaque pixel est le résultat d'une mesure, généralement faite par une matrice de capteurs appelée CCD (Charge Couple Device). Un pixel correspond à un petit carré de la matrice CCD dans lequel le nombre de photons arrivant est compté. Chaque pixel contient trois nombres allant d'habitude de 0 à 255 et indiquant la quantité de rouge, vert et bleu contenues dans le pixel. Une combinaison adéquate de ces trois nombres permet en principe de reproduire n'importe quelle couleur sur écran d'ordinateur.

La figure FIG 1.7 montre un tout petit morceau de neuf pixels extraits d'une image digitale. Dans les pixels rouges, la première composante est élevée, autour de 210, et dans les pixels bleus c'est la troisième composante, bleue, qui domine.

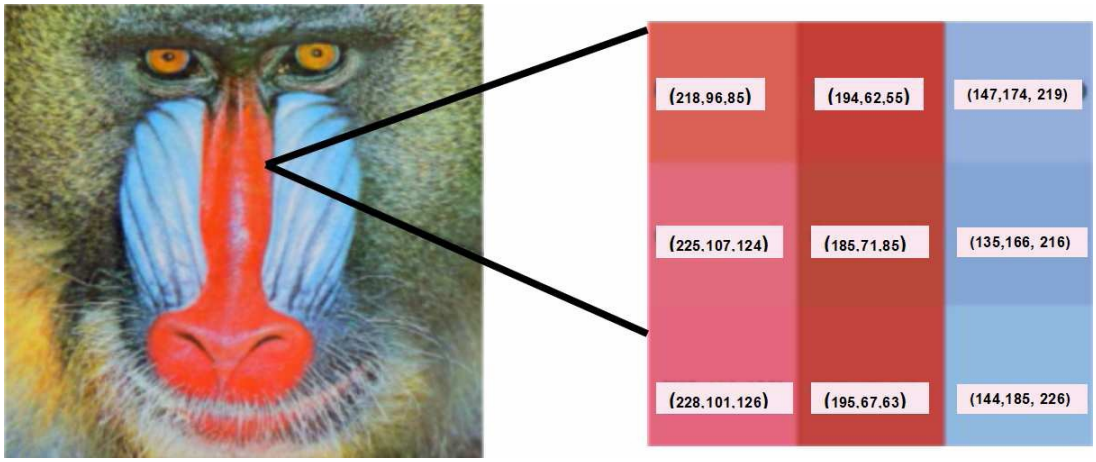


FIG. 1.7 – Image de babouin.

Dans toute image couleur, les trois valeurs des couleurs RVB (Rouge, Vert, Bleu) observées présentent une incertitude due au bruit. Cette incertitude est due aux aléas du comptage des photons arrivant dans chaque capteur. Les valeurs de couleur mesurées sont perturbées car les capteurs reçoivent des photons parasites et subissent des fluctuations électrostatiques lors de leurs charges et décharges. Quand un capteur reçoit beaucoup de photons venant d'une scène bien éclairée, les parasites sont négligeables par rapport au flux de vrais photons. Mais, même dans une photo d'exposition suffisante, les pixels sombres reçoivent très peu de photons et sont donc "bruités".

Dans la figure FIG.1.8, à l'image de gauche, nous avons ajouté l'image du milieu, une image de bruit "blanc ", pour obtenir l'image bruitée de droite. Toutes les images ont du bruit mais il est en général moins accentué que dans cet exemple. Le bruit est dû aux fluctuations de la lumière et à l'activité photoélectronique des capteurs. Notre propre perception génère aussi un bruit, que nous pouvons voir si nous ouvrons les yeux dans l'obscurité, ou simplement en les fermant.

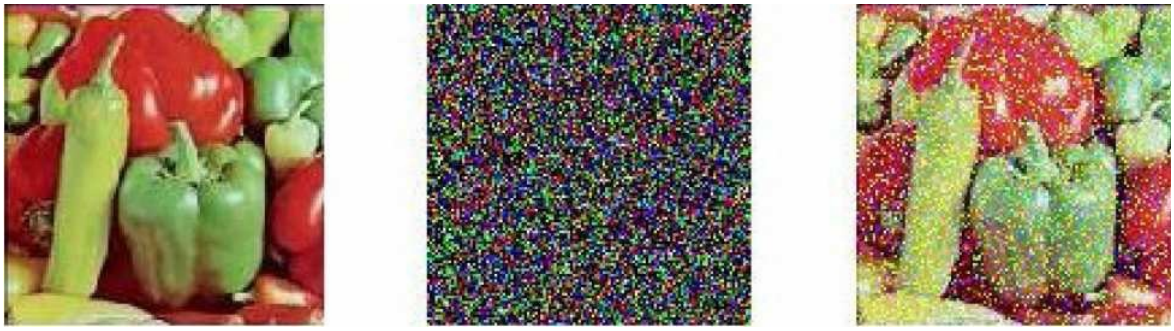


FIG. 1.8 – Influence du bruit sur une image couleur.

La réduction du bruit dans les images couleurs est liée à plusieurs enjeux technologiques. Ces enjeux technologiques sont présents dans pratiquement tout dispositif créant des

images : en microscopie, en imagerie médicale ou satellitaire et bien sûr pour les caméras grand public. La nature du bruit est un vaste sujet délicat de traitement des images qui constitue un outil fondamental dans ce domaine. Dans notre travail, nous nous sommes intéressés plus particulièrement à la modélisation statistique du bruit issue de l'acquisition de d'une image couleur par un capteur. Notre méthode propose l'estimation de la loi du bruit en se basant sur l'estimateur de la fonction trace d'une FEN infiniment divisible.

Représentation statistique d'une image bruitée

Considérons une image $u(x, y, z)$, convoluée par une fonction de transfert $h(x, y, z)$ à laquelle nous avons ajouté un bruit additif ε. L'image de sortie I au point (x, y, z) (pixel) s'écrit sous la forme :

$$I(x, y, z) = (h * u)(x, y, z) + \varepsilon(x, y, z), \ \forall \ (x, y, z) \ \in \ \mathbb{R}^3_+. \tag{1.15}$$

Cette équation peut être écrite de manière vectorielle en notant I l'ensemble des données de l'image de sortie, u l'ensemble des données de l'image d'entrée, H la matrice de transfert et ε le vecteur de bruit. L'équation 1.15 devient alors :

$$I = Hu + \varepsilon. \tag{1.16}$$

Nous supposons que la distribution du vecteur bruit ε appartient à une FEN infiniment divisible engendrée par une mesure $\mu \in \ \mathcal{M}(\mathbb{R}^3)$, $(\varepsilon(t) \sim P_{\theta,\mu})$.

Résultats et commentaires

Pour voir l'intérêt de la fonction trace en traitement d'images, nous allons présenter dans cette section quelques résultats de l'estimation de loi du bruit additif associé à

une image couleur. Pour le cas d'une image couleur, nous représentons chaque pixel par $x_i = (x_i^1, x_i^2, x_i^3)$, pour $i = 1, \ldots, n$ (nombre total de pixels dans l'image) et nous désignons par $m = (m_1, m_2, m_3)$ le vecteur moyenne de l'image. Nous considérons le modèle linéaire suivant : $x_i = m + \varepsilon_i$ avec $\varepsilon_i = (\varepsilon_i^1, \varepsilon_i^2, \varepsilon_i^3)$ le bruit additif centré associé à l'observation (pixel) x_i. Notons que les variables aléatoires $\varepsilon_i^1, \varepsilon_i^2$ et ε_i^3 ne sont pas forcément indépendantes.

Nous supposons que la loi du bruit ϵ_i appartient à une famille exponentielle naturelle infiniment divisible. Donc x_i sera la réalisation d'une loi $P(m, F_j)$ où F_j est une FEN infiniment divisible, $j = 1, \ldots, N$.

Pour déterminer l'estimateur de la distribution du bruit, nous avons généré quatre lois (Gamma-Gaussienne, Gaussienne, Inverse-Gaussienne, Poisson) en choisissant les valeurs des paramètres de la façon suivante :

- n : la taille de l'image.
- $B = 100$ le nombre de Boostrap.
- Choisir cinq valeurs de θ afin de réaliser cette expérience.
- Bruit Gaussien de paramètres $\mu = (0, 0, 0)$ et $\Sigma = (\sigma_{ij})_{1 \leq i,j \leq 3}$, où $\sigma_{ii} = 1$ et $\sigma_{ij} = 0.3$, si $i \neq j$.
- Bruit Gamma-Gaussien du vecteur paramètres $\theta = (-2, 1, 1.2)$.
- Bruit Inverse-Gaussien du vecteur paramètres $\theta = (-1, -0.5, 0.3)$ et $p = 1$.
- Bruit Poisson de paramètres du vecteur paramètres $\theta = (3, 1, 5)$.

Nous présentons dans les Figures FIG.1.9, FIG.1.10, FIG.1.11 et FIG.1.12 l'influence de la nature du bruit sur des images couleurs réelles et sur leurs histogrammes. Nous présentons respectivement dans les figures FIG.1.9 et FIG.1.11 l'image initiale Léna et l'image initiale

Légume, les trois composantes de ces deux images et leurs images bruitées. L'image Léna est de dimension $256 \times 256 \times 3$ et l'image Légume est de dimension $512 \times 512 \times 3$. Dans les figures FIG.1.10 et FIG.1.12, nous présentons les histogrammes 2D pour différentes composantes des deux images avec bruit et sans bruit.

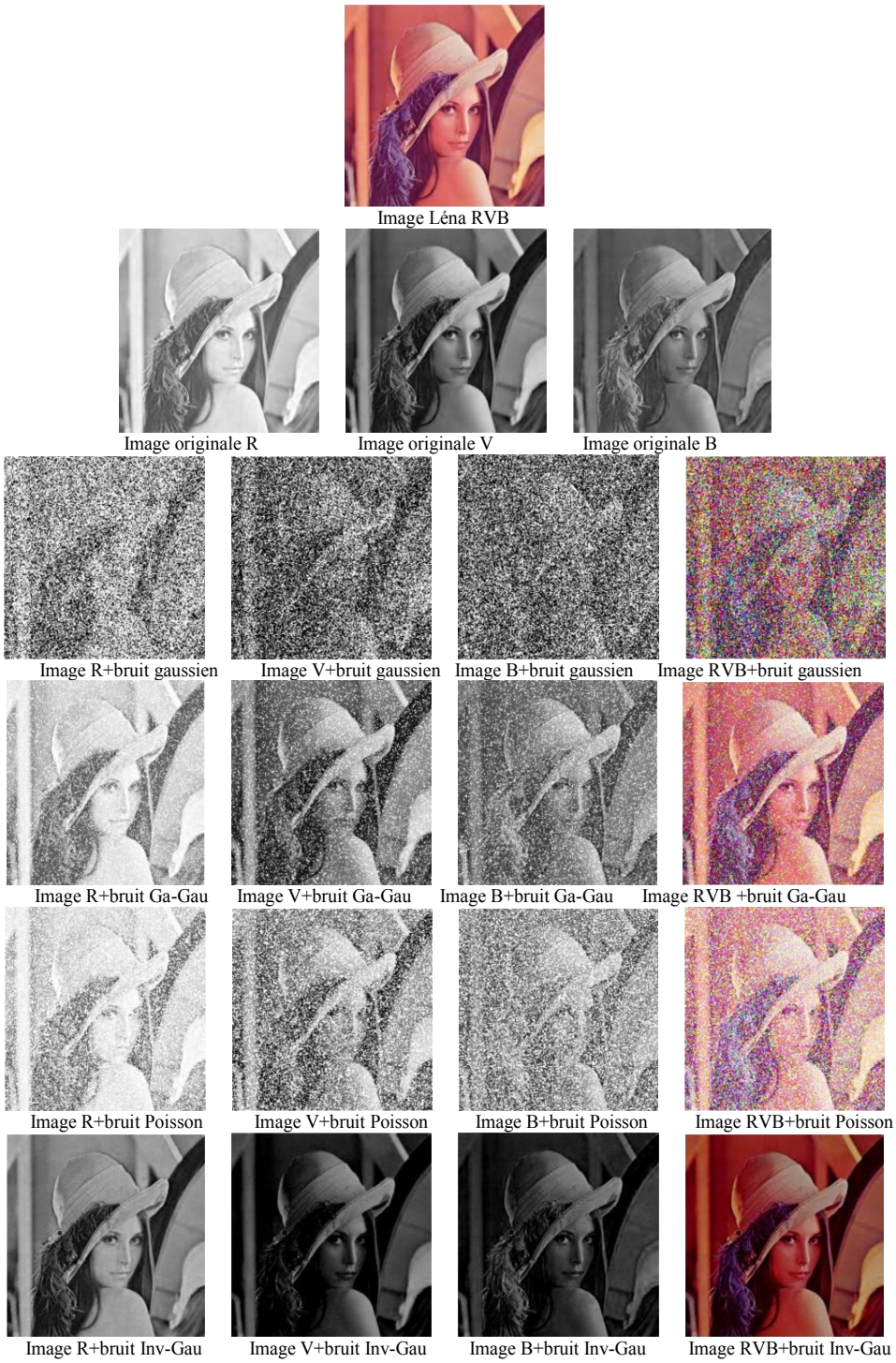


FIG. 1.9 – Influence de la nature du bruit sur l'image Léna.

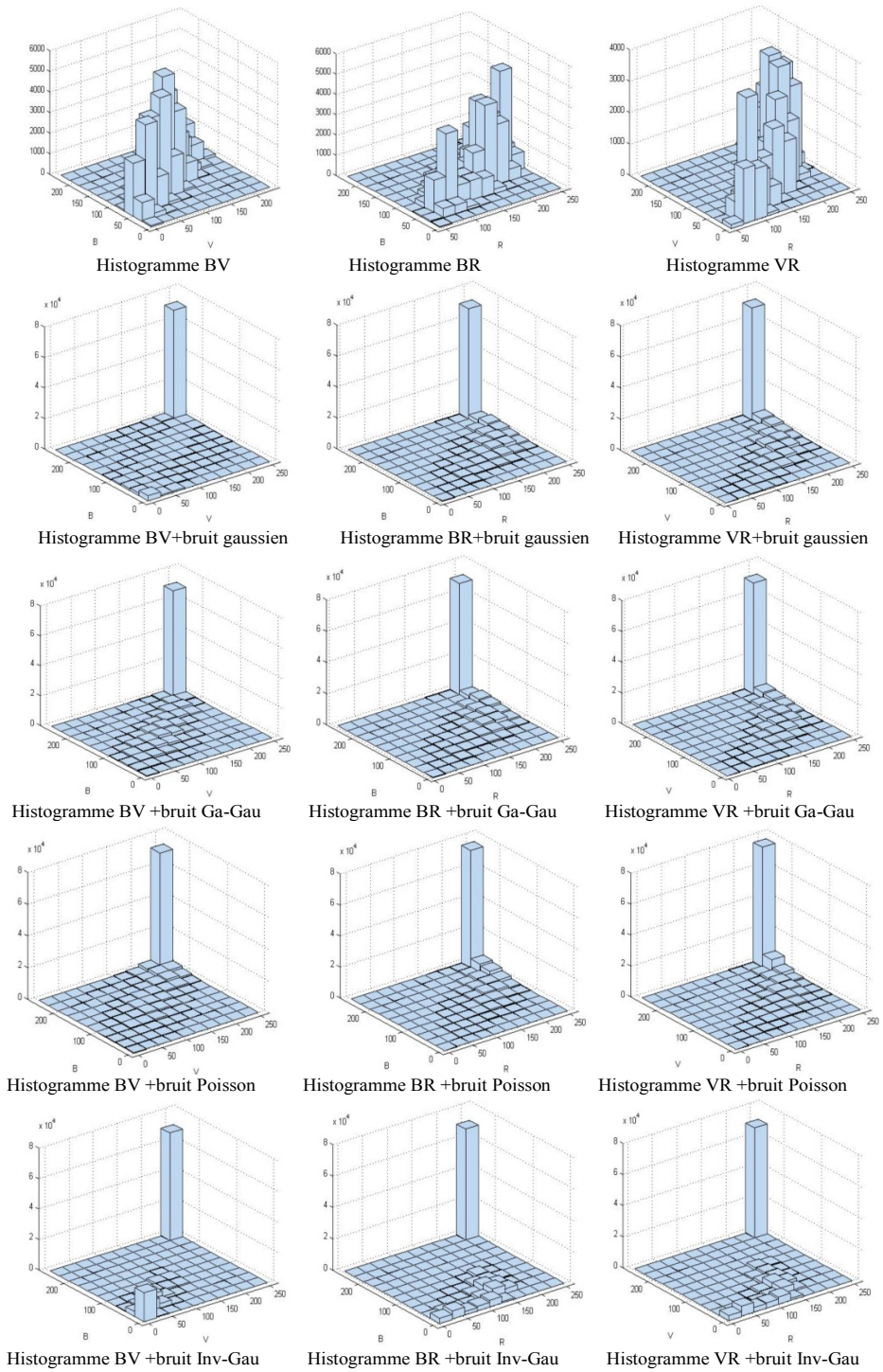


FIG. 1.10 – Histogrammes $2D$ de l'image Léna.

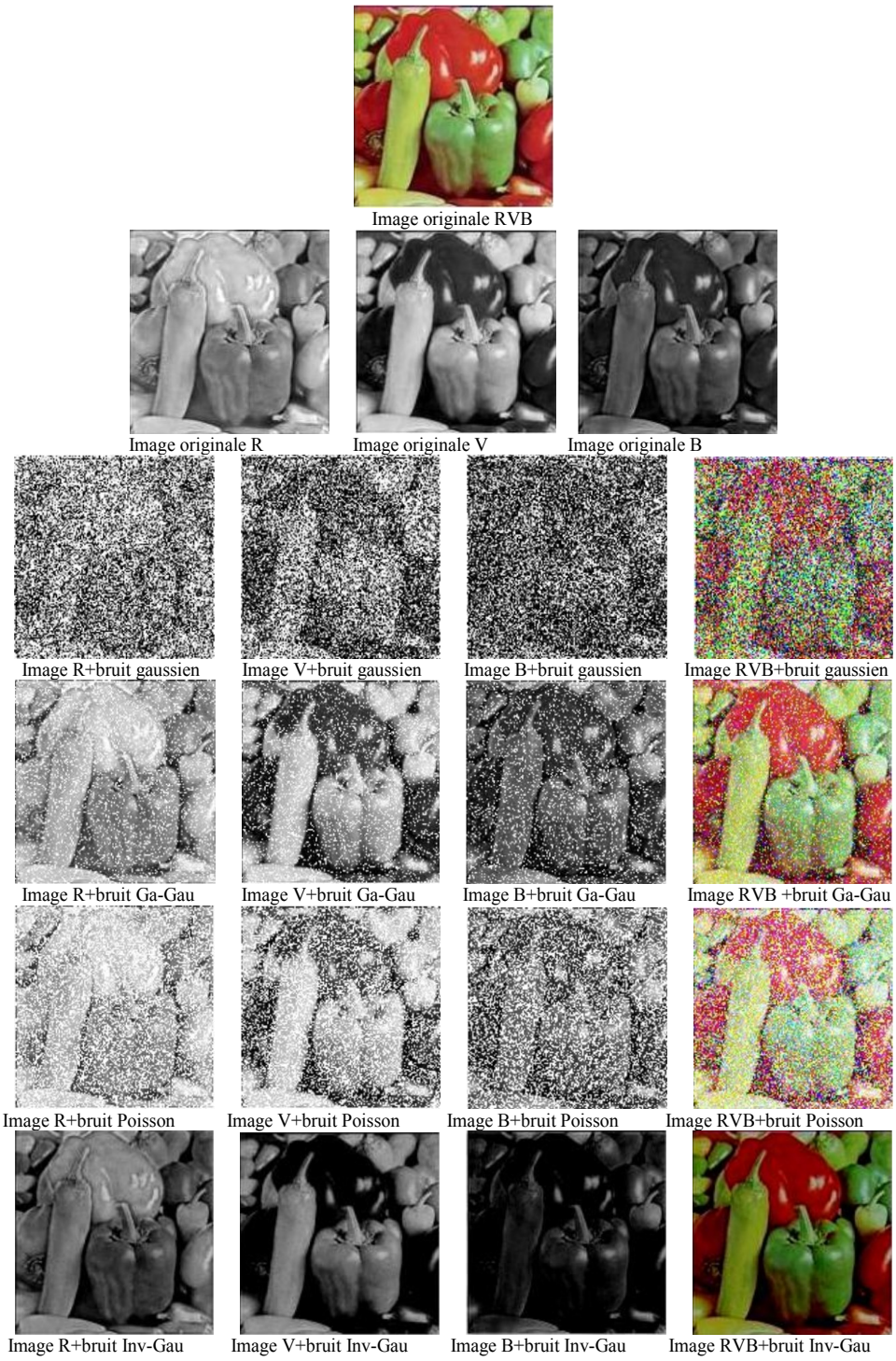


FIG. 1.11 – Influence de la nature du bruit sur l'image Légume.

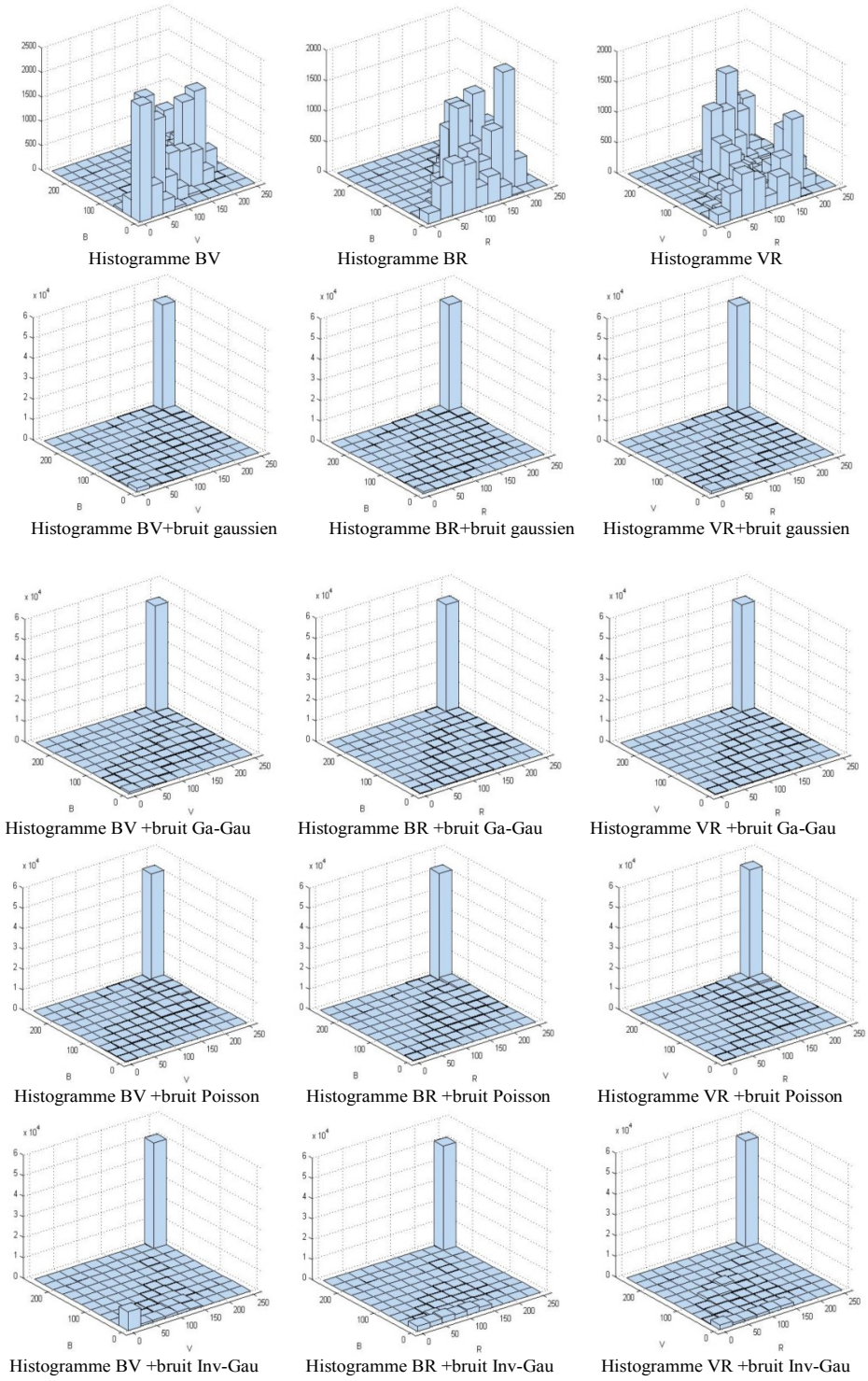


FIG. 1.12 – Histogrammes $2D$ de l'image Légume.

Pour détecter la nature du bruit associé à ces deux images couleurs en se basant sur la fonction trace proposée dans le paragraphe 1.3, nous avons réalisé les quatre étapes suivantes :

• Nous avons supposé que la loi $P_{\theta,\mu}$ est une Gamma-Gaussienne. Puis, nous avons calculé la fonction trace de cette loi notée $trtp$ pour différentes valeurs de θ. Ensuite, nous avons estimé cette fonction trace en utilisant les différents estimateurs présentés dans le tableau TAB.1.4. Enfin, nous avons calculé l'EQM entre $trtp$ et ces différents estimateurs.

• Nous avons supposé que la loi $P_{\theta,\mu}$ est une Gaussienne de moyenne θ et de matrice variance-covariance $\Sigma = (\sigma_{ij})_{1\leq i,j\leq 3}$, avec $\sigma_{ii} = 1$ et $\sigma_{ij} = 0.3$, si $i \neq j$. Puis, nous avons répété la même procédure que celle présentée dans la première étape.

• Nous avons supposé que la loi $P_{\theta,\mu}$ est une Inverse Gaussienne avec $p = 1$. Puis, nous avons répété la même procédure que celle présentée dans la première étape.

• Nous avons supposé que la loi $P_{\theta,\mu}$ est une Poisson. Puis, nous avons répété la même procédure que celle présentée dans la première étape.

D'après les tableaux TAB.1.9 et TAB.1.13 qui représentent respectivement les résultats obtenus par application de ces quatre étapes à ces deux images, nous remarquons que la valeur de l'EQM est minimale pour l'estimateur de la fonction trace d'une loi Gamma-Gaussienne. Donc, nous pouvons qualifier la loi du bruit par une Gamma-Gaussienne. D'après les tableaux TAB.1.10 et TAB.1.14 qui représentent respectivement les résultats obtenus par l'application de ces quatre étapes à ces deux images, nous remarquons que la valeur de l'EQM est minimale pour l'estimateur de la fonction trace d'une loi Gaussienne. Donc, nous pouvons qualifier la loi du bruit par une Gaussienne. D'après les tableaux TAB.1.11 et TAB.1.15 qui représentent respectivement les résultats obtenus par applica-

tion de ces quatre étapes à ces deux images, nous remarquons que la valeur de l'EQM est minimale pour l'estimateur de la fonction trace d'une loi Inverse Gaussienne. Donc, nous pouvons qualifier la loi du bruit par une Inverse Gaussienne. D'après les tableaux TAB.1.12 et TAB.1.16 qui représentent respectivement les résultats obtenus par application de ces quatre étapes à ces deux images, nous remarquons que la valeur de l'EQM est minimale pour l'estimateur de la fonction trace d'une loi Poisson. Donc, nous pouvons qualifier la loi du bruit par une Poisson.

$(\theta_1, \theta_2, \theta_3)$	$trtp$	EQM($trtp$, $tr(GG)$)	EQM($trtp$, $tr(G)$)	EQM($trtp$, $tr(IG)$)	EQM($trtp, tr(P)$)
$(-2, 1, 0)$	5	0.117	0.271	1.201	0.881
$(-1, 0.5, 0.5)$	5.3333	0.0987	0.527	1.35	0.145
$(-6, 2, 1)$	1.0612	0.41910	0.6022	2.54	0.667
$(-0.1, 0.1, 0.1)$	148.1481	0.1521	0.338	1.052	0.271
$(-10, 0, 0)$	0.21	0.1112	0.3458	1.087	0.7519

TAB. 1.9 – Différentes valeurs des EQM selon le modèle Gamma-Gaussien trivarié (Léna).

$(\theta_1, \theta_2, \theta_3)$	$trtp$	EQM($trtp$, $tr(GG)$)	EQM($trtp$, $tr(G)$)	EQM($trtp$, $tr(IG)$)	EQM($trtp, tr(P)$)
$(0, 0, 0)$	3	0.401	0.0241	1.325	0.45782
$(4, 5, 0)$	3	0.598	0.1248	1.179	0.9825
$(2, 1, 0)$	3	0.7542	0.0229	1.4521	0.9257
$(5, 10, 0)$	3	0.888	0.2549	1.0578	0.9876
$(0.5, 0.1, 0.7)$	3	0.2245	0.0457	0.7541	0.7862

TAB. 1.10 – Différentes valeurs des EQM selon le modèle Gaussien trivarié (Léna).

$(\theta_1, \theta_2, \theta_3)$	$trtp$	EQM($trtp$, $tr(GG)$)	EQM($trtp$, $tr(G)$)	EQM($trtp$, $tr(IG)$)	EQM($trtp, tr(P)$)
$(-3, 2, 1)$	8	0.981	0.7267	0.04225	1.0182
$(-0.5, 0.1, 0.1)$	3.0717	0.7587	0.1194	0.05459	1.109
$(-5, 2, 1)$	1.4311	0.18234	1.247	0.04334	0.9801
$(-10, 0.1, 4)$	3.1355	1.581	0.859	0.0118	0.816
$(-25, 4, 5)$	2.2222	1.011	0.5247	0.0372	0.945

TAB. 1.11 – Différentes valeurs des EQM selon le modèle Inverse-Gaussien trivarié (Léna).

$(\theta_1, \theta_2, \theta_3)$	$trtp$	EQM($trtp$, $tr(GG)$)	EQM($trtp$, $tr(G)$)	EQM($trtp$, $tr(IG)$)	EQM($trtp, tr(P)$)
$(0.5, 0.2, 0.4)$	27.0864	2.38	0.9562	4.52	0.22
$(1.2, 1.7, -1.80)$	67.3474	2.248	1.64	3.436	0.929
$(1, 2, 0.03)$	93.8852	2.4512	1.194	3.0664	0.8102
$(1, 0.25, 0.8)$	34.8648	1.1991	0.54606	2.219	0.09618
$(1, 0.44, 0.12)$	29.8663	3.611	1.457	3.7927	0.673

TAB. 1.12 – Différentes valeurs des EQM selon le modèle Poisson trivarié (Léna).

$(\theta_1, \theta_2, \theta_3)$	$trtp$	EQM($trtp$, $tr(GG)$)	EQM($trtp$, $tr(G)$)	EQM($trtp$, $tr(IG)$)	EQM($trtp, tr(P)$)
$(-1, 0, 0)$	3	0.0554	0.724	1.457	2.4587
$(-5, 1, 1)$	0.6875	0.719	1.2146	1.89593	2.189
$(-10, 3, 2)$	1.7143	0.0788	1.194	1.6064	2.1345
$(-0.5, 0.2, 0.3)$	10.5694	0.6177	1.0096	1.36	2.2916
$(-90, 10, 6)$	0.374	0.564	0.9917	1.127	2.146

TAB. 1.13 – Différentes valeurs des EQM selon le modèle Gamma-Gaussien trivarié (Légume).

$(\theta_1, \theta_2, \theta_3)$	$trtp$	EQM($trtp$, $tr(GG)$)	EQM($trtp$, $tr(G)$)	EQM($trtp$, $tr(IG)$)	EQM($trtp, tr(P)$)
$(0, 0, 0)$	3	1.9912	0.0871	3.4172	2.4582
$(2, 1, 0)$	3	1.072	0.0311	1.66723	2.2289
$(5, 11, 10)$	3	1.7547	0.094	1.553	1.994
$(0, 5, 10)$	3	1.086	0.286	2.419	3.916
$(1, 15, 20)$	3	1.21	0.4597	1.837	2.6830

TAB. 1.14 – Différentes valeurs des EQM selon le modèle Gaussien trivarié (Légume).

Le choix de cinq valeurs différentes de θ ne nous permet pas d'affirmer que les résultats obtenus sont efficaces. Pour cette raison, nous avons généralisé notre étude pour une centaine de valeurs de θ de façon aléatoire pour chaque loi considérée et nous avons répété la même procédure que précédemment. Dans ce cas, notre étude nous a monté que la loi du bruit ε_i dans le cas d'un modèle Gamma-Gaussien converge vers une loi Gamma-Gaussienne avec un pourcentage égal à 98%. Puis, le bruit ε_i, dans le cas d'un modèle Gaussien, converge vers une loi Gaussienne, avec un pourcentage égal à 99%. Dans le cas

$(\theta_1, \theta_2, \theta_3)$	$trtp$	EQM($trtp$, $tr(GG)$)	EQM($trtp$, $tr(G)$)	EQM($trtp$, $tr(IG)$)	EQM($trtp, tr(P)$)
$(-15, 3, 2)$	0.6848	1.865	0.7845	0.102	2.82
$(-1, 0, 0)$	3.11	1.74	0.9174	0.0236	2.19
$(-4, 0, 0.7)$	0.8022	1.002	0.1012	1.454	2.1164
$(-0.6, 0.6, 0)$	3.9487	1.618	0.96	0.045	1.9929
$(-100, 7, 6)$	0.2562	1.454	0.27	0.01237	2.2157

TAB. 1.15 – Différentes valeurs des EQM selon le modèle Inverse-Gaussien trivarié (Légume).

$(\theta_1, \theta_2, \theta_3)$	$trtp$	EQM($trtp$, $tr(GG)$)	EQM($trtp$, $tr(G)$)	EQM($trtp$, $tr(IG)$)	EQM($trtp, tr(P)$)
$(3, 0.098, 0.52)$	57.3708	1.945	0.985	2.2347	0.1578
$(0.3, 0.078, 0.73)$	29.2787	0.1809	0.2574	1.2784	0.0027
$(-0.3, -4.078, -2.73)$	1.8759	2.987	1.4789	2.145	0.145
$(-1, -2, -3)$	2.4655	3.257	1.458	5.147	0.9854
$(0.15, -0.29, -0.373)$	15.2009	4.180	1.874	3.1774	0.684

TAB. 1.16 – Différentes valeurs des EQM selon le modèle Poisson trivarié (Légume).

d'un modèle Inverse Gaussien, la loi du bruit ε_i converge vers une loi Inverse Gaussienne avec un pourcentage égal à 97%. Enfin, la loi du bruit ε_i, dans le cas d'un modèle Poisson, converge vers une loi de Poisson avec un pourcentage égal à 99%.

1.5 Conclusion

Nous avons rappelé dans ce chapitre les principales définitions des familles exponentielles, ainsi que leurs propriétés fondamentales. Nous avons introduit une nouvelle caractérisation des familles exponentielles naturelles infiniment divisibles. Cette nouvelle caractérisation se base sur la fonction trace de la matrice variance covariance associée à cette famille. Au niveau du traitement du signal et des images, cette fonction trace nous a permis d'approximer la nature de loi d'un bruit additif associé à un signal ou à une image couleur. Dans un premier temps, nous avons montré à partir de quelques simulations que l'hypothèse portée sur le fait que le signal est perturbé par un bruit gaussien blanc n'est pas toujours vraie et ceci grâce au critère de comparaison basé sur l'erreur quadratique moyenne. Dans un deuxième temps, nous avons traité le problème de l'identification de la nature du bruit affectant une image couleur. L'identification est une étape importante dans tout système d'interprétation de l'information par vision lorsque la nature de la dégradation est inconnue. Certains algorithmes d'analyse supposent que la nature du bruit est Gaussienne. Or, dans la plupart des cas, ces données ne sont pas connues, c'est pourquoi elles doivent être estimées car elles conditionnent la qualité de l'analyse de l'image. L'application de la fonction trace sur des images couleurs, nous a permis de bien déterminer la nature du bruit associé.

Chapitre 2

Estimation paramétrique des modèles statistiques multivariés

2.1 Introduction

La modélisation statistique consiste essentiellement à définir des outils appropriés pour modéliser des données observées, en tenant compte de leur nature aléatoire. Depuis quelques décennies, les méthodes statistiques multivariées sont devenues de plus en plus populaires dans le domaine de traitement des images et ont été utilisées pour de nombreux problèmes tels que le dé-bruitage, la compression et la restauration d'images [20][37][44][61]. Les applications de ces méthodes multivariées ont fait d'énormes progrès grâce au développement de l'informatique.

Dans ce chapitre, nous nous intéressons au problème d'estimation paramétrique pour des modèles multivariés. Plusieurs méthodes d'estimation paramètrique ont été proposées dans la littérature [63]. Parmi elles, nous pouvons citer la méthode des moments [103], la mé-

thode du maximum de vraisemblance [103], la méthode des moindres carrés [103] et la méthode Bayésienne [71].

Ce chapitre se décompose en deux parties. La première partie présente quelques lois multivariées qui appartiennent à la classe des familles exponentielles, ainsi que l'estimation de leurs paramètres en utilisant la méthode des moments et la méthode du maximum de vraisemblance. La deuxième partie présente notre deuxième contribution dans cette thèse qui est la proposition d'un nouveau modèle paramétrique multivarié en se basant sur la loi de Riesz [11], ainsi que l'estimation de ses paramètres en utilisant la méthode des moments, la méthode du maximum de vraisemblance et la méthode Bayésienne [11][13][14].

2.2 Lois multivariées

Dans notre travail, nous nous sommes intéressés à étudier des lois multivariées appartenant à des familles exponentielles [8][25]. L'importance de ces familles est due au fait, d'une part, qu'elles sont facilement manipulables puisqu'elles représentent le coeur de l'ensemble des distributions de probabliltés classiques et, d'autre part, elles possèdent certaines propriétés telles que :

• L'existence d'une statistique complète (et donc également exhaustive minimale) pour le paramètre inconnu de la loi [103].

• L'existence d'un estimateur efficace pour les paramètres de la loi [103].

• Au niveau de l'application elles s'adaptent mieux aux problèmes de traitement d'images [20][38].

2.2.1 Exemples de lois multivariées

2.2.1.a Loi Normale multivariée

On appelle loi Normale multivariée, ou loi multinormale ou loi de Gauss à plusieurs variables, une loi de probabilité qui est la généralisation multidimensionnelle de la loi Normale. La loi Normale multivariée est notamment utilisée dans le traitement d'images médicales. Une seconde application de la loi normale multivariée est la détermination, à partir des intensités dans des IRM du cerveau d'un patient, des différentes classes des tissus (matière grise, matière blanche, liquide céphalo-rachidien) qui le composent. Cette technique est basée sur l'utilisation d'un algorithme espérance-maximisation dans lequel chacune des classes est modélisée par une loi Normale multivariée dont la dimension est égale aux nombres de modalités utilisées pour la classification.

Définition 2.2.1. *Un vecteur aléatoire $X = (X_1, X_2, \ldots, X_d)$ est dit Gaussien si, pour tout $U \in \mathbb{R}^d \setminus \{0\}$, la variable aléatoire réelle $U.X^t$ est une variable aléatoire Gaussienne. C'est à dire si toute combinaison linéaire de ses composantes est une variable aléatoire réelle Gaussienne.*

La loi normale multivariée représente la distribution d'un vecteur aléatoire Gaussien. Soit $X = (X_1, X_2, \ldots, X_d)$ un vecteur aléatoire Gaussien de dimension d. On désigne par $\mu = (\mu_1, \mu_2, \ldots, \mu_d)^t$ son vecteur moyen et par Σ sa matrice de variance-covariance de dimension $d \times d$. Si Σ est inversible alors la densité de X est définie par

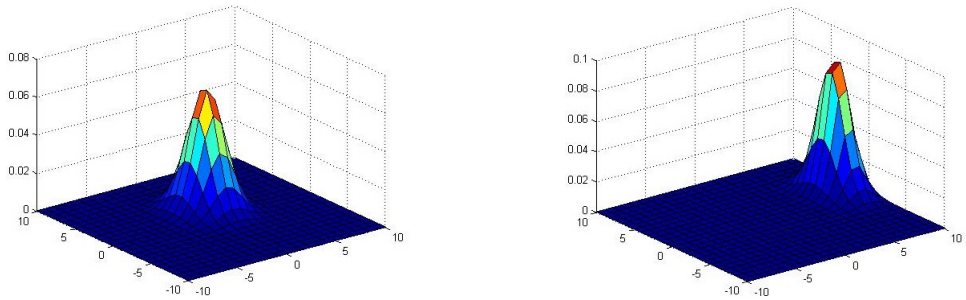
$$f_X(x) = \frac{1}{(2\pi)^{\frac{d}{2}}\sqrt{\det \Sigma}} \exp\Big(\frac{-1}{2}(x-\mu)^t.\Sigma^{-1}.(x-\mu)\Big),\ \forall x \in \mathbb{R}^d. \tag{2.1}$$

Dans la suite de ce travail, nous désignons par $N_d(\mu, \Sigma)$ la loi Normale multivariée de dimension $d \times d$ du vecteur de paramètres (μ, Σ). Les densités marginales de X_i, $i = 1, \ldots, d$,

sont respectivement $N(\mu_i, \sigma_i^2)$, avec $\mu_i = \mathbb{E}(X_i)$ et $\sigma_i^2 = \mathbb{V}ar(X_i)$.

Dans le cas où $\mu = 0 \in \mathbb{R}^d$ et $\Sigma = I_d$ la matrice identité, nous disons que X suit la loi Normale multivariée standard $N_d(0, I_d)$. Dans la littérature, certains auteurs [54] caractérisent la loi Normale multidimensionnelle par le vecteur des paramètres $\theta = (\mu, \Sigma^{-1})$, où Σ^{-1} est l'inverse de la matrice de variance-covariance Σ. La matrice Σ^{-1} est appellée matrice de précision.

Nous présentons dans la figure FIG.2.1 la densité de probabilité d'une loi Normale bivariée pour différentes valeurs des paramètres μ et Σ.



(a) $\mu = (1, 2)$, $\Sigma = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ (b) $\mu = (8, 2)$, $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$

FIG. 2.1 – Représentation de la densité de probabilité d'une loi Normale bivariée.

Dans la suite, nous rappelons la relation entre la loi de Wishart et la loi normale multivariée. D'après [66], les lois de Wishart peuvent être construites à partir des vecteurs Gaussiens.

Définition 2.2.2. *Soient $X_1, \ldots, X_p$, p-vecteurs aléatoires de $\mathbb{R}^d$ indépendants et identiquement distribués selon la loi normale multivariée $N_d(0, \Sigma)$. Si $p \geq d$ alors, la matrice aléatoire définie par $S = X.X^t$ est distribuée selon une loi de Wishart de paramètres p et Σ notée $W_d(p, \Sigma)$ de densité de probabilité qui sera définie ultérieurement par l'équation*

(2.27). L'entier p est appelé degré de liberté de la loi de Wishart.

Pour avoir plus de détails sur les distributions de Wishart, le lecteur pourra consulter [64].

2.2.1.b Loi de Dirichlet multivariée

La loi de Dirichlet [34] a été introduite par Johann Peter Gustav Lejeune Dirichlet. C'est une famille de lois de probabilités absolument continues, par rapport à la mesure de Lebesgue. Elle est paramétrée par le vecteur réel positif α. Cette loi est vue comme une généralisation de la loi Bêta de première espèce.

Définition 2.2.3. *Un vecteur aléatoire $Z = (Z_1, Z_2, \ldots, Z_d)$ à valeurs dans le sous-espace*

$$\Xi = \{(z_1, z_2, \ldots, z_d);\ z_i \geq 0,\ i = 1, \ldots, d;\ \sum_{i=1}^{d} z_i \leq 1\}$$

suit la loi de Dirichlet de paramètres $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d; \alpha_{d+1})$ si sa densité de probabilité est de la forme

$$f_Z(z) = \frac{1}{B(\alpha)} \prod_{i=1}^{d} z_i^{\alpha_i} (1 - \sum_{i=1}^{d} z_i)^{\alpha_{d+1}-1},$$

avec $B(\alpha) = \dfrac{\prod_{i=1}^{d} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{d} (\alpha_i))}$.

Dans le cas où $d = 1$, nous retrouvons la loi Bêta de première espèce $B_1(\alpha_1, \alpha_2)$ qui est définie dans l'Annexe A.

Proposition 2.2.4. *Soient $X_1, X_2 \ldots, X_d$ $d-$variables aléatoires indépendantes qui suivent respectivement les lois Gamma $\gamma(\alpha_i)$, $i = 1, \ldots, d$.*

On considère les variables aléatoires :

$$Z_i = \frac{X_i}{\sum_{i=1}^{d} X_i},\ i = 1, \ldots, d.$$

Alors $(Z_1, Z_2, \ldots, Z_d)$ *suit la loi de Dirichlet de paramètres* $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d; \alpha_{d+1})$.

D'après la Proposition 2.2.4, le résultat obtenu permet d'établir un nouveau résultat sous forme d'une propostion qui est la suivante.

Proposition 2.2.5. *Si* $(Z_1, Z_2, \ldots, Z_d)$ *suit la loi de Dirichlet de paramètres* $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d; \alpha_{d+1})$ *alors la loi marginale de* $(Z_{i_1}, Z_{i_2}, \ldots, Z_{i_k})$, $k \leq d$, *suit la loi de Dirichlet de paramètres* $\alpha = (\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k}; \sum_{\{i \neq i_q;\ q=1,\ldots,k\}} \alpha_i)$.

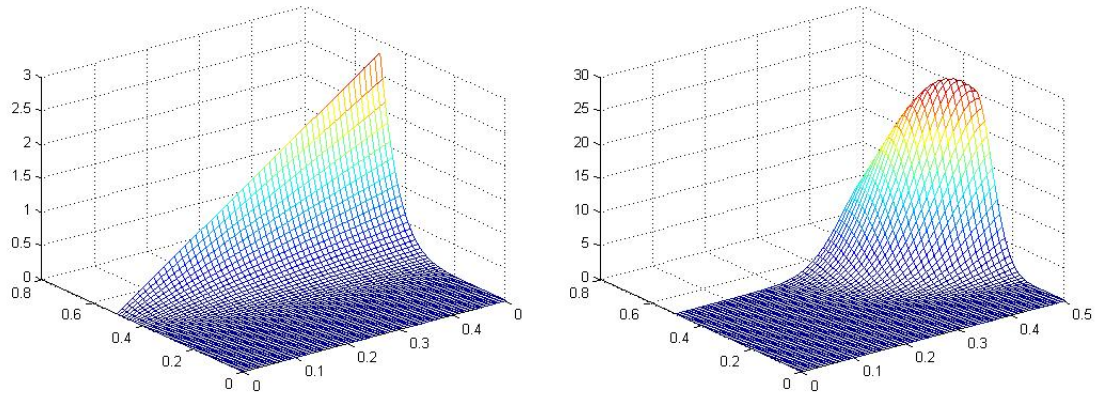
A partir de ces résultats, nous pouvons montrer que la loi marginale de chaque Z_i est une loi Bêta $B_1(\alpha_i, \alpha - \alpha_i)$ avec $\alpha = \sum_{i=1}^{d} \alpha_i$.

Pour simplifier l'étude de la loi de Dirichlet, nous traitons le cas où $d = 2$. Dans ce cas, la loi de Dirichlet de paramètres $\alpha = (\alpha_1, \alpha_2; \alpha_3)$ est la loi Bêta bivariée de première espèce de paramètres $a = \alpha_1$, $b = \alpha_2$ et $c = \alpha_3$, de densité de probabilité

$$f_{(Z_1,Z_2)}(z_1, z_2) = \frac{1}{B(a,b,c)} z_1^{a-1} z_2^{b-1} (1 - z_1 - z_2)^{c-1}, \quad z_1 > 0,\ z_2 > 0,\ z_1 + z_2 < 1, \quad (2.2)$$

avec $B(a,b,c) = \dfrac{\Gamma(a)\Gamma(b)\Gamma(c)}{\Gamma(a+b+c)}$.

Notons que le modèle des lois Bêta bivariées de première espèce est une FEG engendrée par la mesure $\nu(dz_1, dz_2) = \frac{1}{z_1 z_2 (1 - z_1 - z_2)} \mathbf{1}_{\{(z_1,z_2) \in (0,1)^2 / z_1 + z_2 < 1\}}(z_1, z_2) dz_1 dz_2$ et par une statistique $T(z_1, z_2) = (\ln z_1, \ln z_2, \ln(1 - z_1 - z_2))$. La fonction $k(\theta) = -\ln B(a,b,c)$ avec $\theta = (a,b,c) \in]0, +\infty[^3$. La forme de la densité de probabilité du vecteur aléatoire $Z = (Z_1, Z_2)$ varie selon les valeurs de ses paramètres a, b et c (voir FIG.2.2).



(a) $a = 2$, $b = 8$ et $c = 1$. (b) $a = 3$, $b = 11$ et $c = 5$.

FIG. 2.2 – Représentation de la densité d'une loi Bêta bivariée de première espèce.

Moments

Les expressions du moment d'ordre $(1, 1)$, de la covariance et du coefficient de corrélation d'une loi Bêta bivariée de première espèce sont :

$$\mathbb{E}(Z_1 Z_2) = \frac{ab}{(a+b+c+1)(a+b+c)}$$

$$\mathbb{C}ov(Z_1, Z_2) = \frac{-ab}{(a+b+c+1)(a+b+c)^2}$$

$$r = \mathbb{C}orr(Z_1, Z_2) = \frac{\mathbb{C}ov(Z_1, Z_2)}{\sqrt{\mathbb{V}ar(Z_1)\mathbb{V}ar(Z_2)}} = -\sqrt{\frac{ab}{(a+c)(b+c)}}.$$

La loi Bêta bivariée de première espèce est à dépendance négative puisque $-1 < r < 0$.

2.2.1.c Loi de Dirichlet inverse multidimensionnelle

La loi de Dirichlet inverse multidimensionnelle est vue comme une généralisation de la loi Bêta de deuxième espèce définie dans l'Annexe A.

Définition 2.2.6. *Un vecteur aléatoire $Y = (Y_1, Y_2, \ldots, Y_d)$ à valeurs dans le sous-espace*

$$\Xi = \{(y_1, y_2, \ldots, y_d);\ y_i \geq 0,\ i = 1, \ldots, d\}$$

suit la loi de Dirichlet inverse multidimensionnelle de paramètres $\beta = (\beta_1, \beta_2, \ldots, \beta_d; \beta_{d+1})$ *si sa densité de probabilité est de la forme*

$$f_Y(y) = \frac{1}{B(\beta)} \prod_{i=1}^{d} y_i^{\beta_i - 1} (1 + \sum_{i=1}^{d} y_i)^{-\sum_{i=1}^{d+1} \beta_i},$$

avec $B(\beta) = \dfrac{\prod_{i=1}^{d} \Gamma(\beta_i)}{\Gamma(\sum_{i=1}^{d} (\beta_i))}$.

Remarque 2.2.7. *Dans le cas où* $d = 1$*, nous retrouvons la loi Bêta de deuxième espèce* $B_2(\beta_1, \beta_2)$.

Pour simplifier l'étude de la loi de Dirichlet inverse multidimensionnelle, nous traitons le cas où $d = 2$. Dans ce cas, la loi de Dirichlet inverse de paramètres $\beta = (\beta_1, \beta_2; \beta_3)$ est la loi Bêta bivariée de deuxième espèce de paramètres $a = \beta_1$, $b = \beta_2$ et $c = \beta_3$, de densité de probabilité définie par

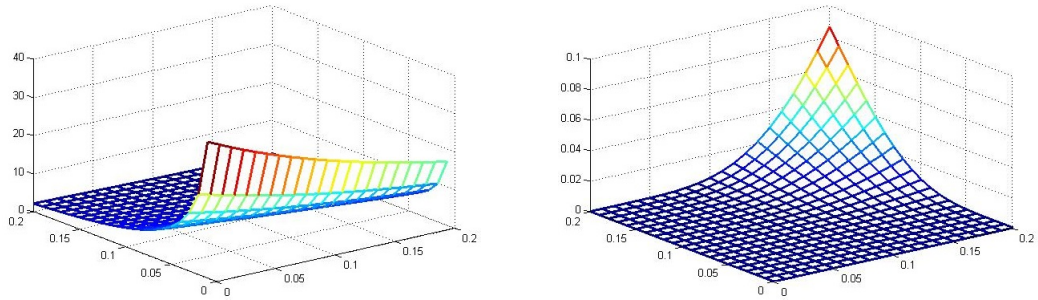
$$f_{(Y_1,Y_2)}(y_1, y_2) = \frac{1}{B(a,b,c)} y_1^{a-1} y_2^{b-1} (1 + y_1 + y_2)^{-(a+b+c)}, \quad y_1 > 0, \ y_2 > 0, \tag{2.3}$$

avec $B(a,b,c) = \dfrac{\Gamma(a)\Gamma(b)\Gamma(c)}{\Gamma(a+b+c)}$.

Un résultat immédiat découle de cette définition est que les lois marginales de Y_1 et Y_2 sont deux lois Bêta univariées de deuxième espèce de paramètres respectivement (a, c) et (b, c). Par suite, on peut construire une variable aléatoire qui suit la loi Bêta bivariée de deuxième espèce de la manière suivante : Soient X_1, X_2 et X_3 trois v.a. indépendantes qui suivent respectivement trois lois Gamma standards (voir Annexe A) de paramètres de forme $a > 0$, $b > 0$, et $c > 0$. Alors les deux variables aléatoires corrélées $Y_1 = \dfrac{X_1}{X_3}$ et $Y_2 = \dfrac{X_2}{X_3}$ suivent respectivement des lois Bêta de deuxième espèce de paramètres de forme (a, c) et (b, c). D'où, le couple de variables aléatoires (Y_1, Y_2) suit la loi Bêta bivariée de

deuxième espèce de densité de probabilité définie par l'équation (2.3).

Notons que le modèle de lois Bêta bivariées de deuxième espèce est une FEG engendrée par la mesure $\nu(dy_1, dy_2) = \frac{1}{y_1 y_2}\mathbf{1}_{]0,+\infty[^2}(y_1, y_2)dy_1dy_2$ et par une statistique $T(y_1, y_2) = (\ln y_1, \ln y_2, \ln(1 + y_1 + y_2))$. La fonction $k(\theta) = -\ln B(a, b, c)$ avec $\theta = (a, b, c) \in]0, +\infty[^3$. La forme de la densité de probabilité du vecteur aléatoire $Y = (Y_1, Y_2)$ varie selon les valeurs de ses paramètres (voir FIG.2.3).



(a) $a = 3, b = 1, c = 3$ (b) $a = 7, b = 8, c = 10$

FIG. 2.3 – Représentation de la densité d'une loi Bêta bivariée de deuxième espèce.

Moments

Les expressions du moment d'ordre $(1, 1)$, de la covariance et du coefficient de corrélation d'une loi Bêta bivariée de deuxième espèce, pour $c > 2$ sont :

$$\mathbb{E}(Y_1Y_2) = \frac{ab}{(c-1)(c-2)}$$

$$\mathbb{C}ov(Y_1, Y_2) = \frac{ab}{(c-1)^2(c-2)}$$

$$r = \mathbb{C}orr(Y_1, Y_2) = \frac{\mathbb{C}ov(Y_1, Y_2)}{\sqrt{\mathbb{V}ar(Y_1)\mathbb{V}ar(Y_2)}} = \sqrt{\frac{ab}{(a+c-1)(b+c-1)}}.$$

La loi Bêta bivariée de deuxième espèce est à dépendance positive puisque $0 < r < 1$.

2.2.2 Estimation des paramètres

Dans le cas des modèles paramétriques, nous supposons que les données sont issues d'une famille de distributions paramétriques connue. L'avantage de l'approche paramétrique est que le modèle à estimer est défini par le moyen d'un ensemble de paramètres θ (par exemple moyenne et variance pour la loi Normale). Dans ces approches, la fonction densité dépend donc de la valeur de ces paramètres. Cette densité en un point x est couramment notée $f(x|\theta)$ ou f_θ. Les méthodes d'estimation des paramètres les plus utilisées sont la méthode du maximum de vraisemblance (MMV) et la méthode des moments (MM). La première est employée généralement pour ses propriétés asymptotiques intéressantes, la seconde pour sa simplicité et l'utilisabilité des estimations résultantes comme valeurs initiales dans le processus itératif de la MMV. Les principes généraux de ces deux méthodes, qui nous apparaissent adéquates pour les lois multivariées, sont décrits dans le paragraphe suivant. Nous présentons également dans cette section les développements et les résultats menant aux différents estimateurs des paramètres des lois multivariées pour les deux méthodes considérées.

Description des méthodes du maximum de vraisemblance et des moments

La méthode du maximum de vraisemblance

La méthode d'estimation du maximum de vraisemblance est une méthode d'estimation ponctuelle puisqu'elle cherche à trouver une valeur estimée pour un paramètre θ à partir d'un ensemble donné [33]. Elle est attractive pour plusieurs raisons. D'abord, elle possède de bonnes propriétés de convergence et d'efficacité quand le nombre d'échantillons est très

grand. Les estimateurs du MMV ont asymptotiquement la variance la plus faible parmi tous les estimateurs sans biais. De plus, l'estimation par la MMV est plus simple que les méthodes alternatives, telles que les méthodes bayésiennes [18]. En effet, de point de vue complexité, la MMV ne nécessite que des techniques de calculs différentiels ou une recherche de gradient, tandis que les méthodes bayésiennes peuvent nécessiter des intégrations multidimensionnelles complexes.

Une caractéristique majeure indésirable du MMV est que ses propriétés avec des échantillons finis peuvent être très différentes de ses propriétés asymptotiques. Bien qu'elles soient convergentes, les estimations des paramètres par la MMV sont bien typiquement biaisées.

Principe

Considérons un modèle statistique paramétré $F = \{P_\theta(dx) = f_\theta(x)\nu(dx);\ \theta \in \Theta \subset \mathbb{R}^k\}$ dominé par une mesure σ-finie ν. Soit $X = (X_1, \ldots, X_n)$ un $n-$échantillon aléatoire de la loi P_θ. La méthode du maximum de vraisemblance consiste à estimer θ telle que la fonction vraisemblance $\theta \longmapsto L(X|\theta) = \prod_{i=1}^{n} f_\theta(X_i)$ soit maximale. Ainsi, on cherche à déterminer le maximum de cette vraisemblance pour que les probabilités des réalisations observées de l'échantillon X soit maximales.

Autrement dit, on appele estimateur de θ par la MMV la statistique T, à valeurs dans $\mathbb{R}^k$,

$$T(X) = \arg\max_{\theta \in \Theta} L(X|\theta).$$

Remarque 2.2.8. *L'unicité du maximum de vraisemblance est liée à la notion d'exhaustivité. L'estimateur du maximum de vraisemblance peut ne pas exister en général [103].*

Méthode du maximum de vraisemblance dans les FENs

Soit $\{P_\theta(dx) = e^{\langle\theta,T(x)\rangle - k(\theta)}\nu(dx)\ ;\ \theta \in \Theta(\mu) \subset \mathbb{R}^k\}$, une famille exponentielle générale engendrée par une mesure σ-finie ν et une statistique T à valeurs dans $\mathbb{R}^d$. Comme P_θ est une loi de probabilité, alors

$$\begin{aligned} k(\theta) &= \ln \int_{\mathbb{R}^d} e^{\langle\theta,T(x)\rangle}\nu(dx) \\ &= \ln \int_{\mathbb{R}^d} e^{\langle\theta,y\rangle}T(\nu)(dx) \\ &= k_{T(\nu)}(\theta), \end{aligned}$$

où $T(\nu)$ désigne la mesure image de ν par T.

La fonction log-vraisemblance d'un n-échantillon aléatoire $X_1, X_2, \ldots, X_n$ de loi P_θ est définie par

$$l(\theta) = \langle\theta, \sum_{i=1}^{n} T(X_i)\rangle - nk(\theta).$$

Supposons que $T(\nu) \in \mathcal{M}(\mathbb{R}^d)$, alors d'après le premier chapitre (sous section Moments d'une FEN) la fonction cumulante $k = k_T(\nu)$ est strictement convexe sur $\Theta = \Theta(T(\nu))$. Ceci implique que la fonction log-vraisemblance $l(\theta)$ est strictement concave. Ainsi que l'estimateur du maximum de vraisemblance $\widehat{\theta}$ de θ est solution de l'équation $\dfrac{\partial l(\theta)}{\partial\theta} = 0$.

La méthode des moments

La méthode des moments [105] est un outil d'estimation intuitif qui date du début des statistiques. Elle consiste à estimer les paramètres recherchés en égalisant certains moments théoriques (qui dépendent de ces paramètres) avec leurs contreparties empiriques. L'égalisation se justifie par la loi des grands nombres qui implique que l'on peut "approcher"

une espérance mathématique par une moyenne empirique. On est donc amené à résoudre un système d'équations.

Principe

Considérons une fonction $h : \mathbb{R}^d \longrightarrow \mathbb{R}^k$ et la statistique s_n de taille k définie par

$$s_n = \frac{1}{n}\sum_{i=1}^{n} h(X^i), \tag{2.4}$$

où h est habituellement choisie de sorte que s_n soit constituée par des moments empiriques d'ordre entier. Soit

$$K(\theta) = \mathbb{E}[s_n] = \mathbb{E}(h(X^1)). \tag{2.5}$$

Un estimateur des moments de θ est construit comme suit

$$\widehat{\theta} = g(s_n), \tag{2.6}$$

où $g(K(\theta)) = \theta$.

2.2.2.a Exemples :

Afin de bien comprendre l'utilisation de ces deux méthodes d'estimation la MMV et la MM, nous présentons dans cette sous section quelques exemples dans le cas des lois bivariées.

<u>Modèle de lois Normales</u>

Estimation par la MMV

Soit $X = (X^1, \ldots, X^n)$ un $n-$échantillon aléatoire de loi $N_2(\mu = (\mu_1, \mu_2), \Sigma)$, où $X^i = (X_1^i, X_2^i)$. Nous cherchons à estimer le vecteur des paramètres inconnus $\theta = (\mu, \Sigma)$. La

fonction log-vraisemblance est définie par

$$l(\theta) = l(X^1, \ldots, X^n; \theta) = -n\ln(2\pi) + \frac{1}{2}\sum_{i=1}^{n} \ln(\det \Sigma^{-1}) - \frac{1}{2}\sum_{i=1}^{n}(X^i - \mu)^t.\Sigma^{-1}.(X^i - \mu). \tag{2.7}$$

Pour le vecteur moyen μ, nous calculons le gradient de la fonction $l(\theta)$ par rapport à μ. En résolvant l'équation suivante

$$\nabla_\mu(l(\theta)) = \sum_{i=1}^{n} -\frac{1}{2}\Sigma^{-1}(\mu - X^i) = 0,$$

nous obtenons l'estimateur $\widehat{\mu}$ de $\mu = (\mu_1, \mu_2)$ donné par

$$\widehat{\mu} = (\widehat{\mu}_1, \widehat{\mu}_2) = \Big(\frac{1}{n}\sum_{i=1}^{n} X_1^i, \frac{1}{n}\sum_{i=1}^{n} X_2^i\Big).$$

Pour obtenir l'estimateur de la matrice Σ^{-1}, nous calculons le gradient de la fonction $l(\theta)$ par rapport à Σ^{-1}. En résolvant l'équation suivante

$$\nabla_{\Sigma^{-1}}(l(\theta)) = n\Sigma - \sum_{i=1}^{n}(X^i - \mu).(X^i - \mu)^t = 0,$$

avec $\nabla_{\Sigma^{-1}}(\ln(\det \Sigma^{-1})) = \Sigma$, nous obtenons l'estimateur $\widehat{\Sigma}^{-1}$ donné par

$$\widehat{\Sigma}^{-1} = n\Big(\sum_{i=1}^{n}(X^i - \widehat{\mu}).(X^i - \widehat{\mu})^t\Big)^{-1}.$$

A partir de l'expression de l'estimateur $\widehat{\Sigma}^{-1}$ de Σ^{-1}, nous déduisons l'estimateur $\widehat{\Sigma}$ du paramètre Σ qui est donné par

$$\widehat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(X^i - \widehat{\mu}).(X^i - \widehat{\mu})^t.$$

Nous remarquons que les estimateurs $\widehat{\mu}$ et $\widehat{\Sigma}$ obtenus par la MMV sont respectivement le vecteur moyen et la matrice de variance-covariance empiriques. Ainsi, d'après la Définition 2.2.2 et dans le cas où les données sont générées par une loi normale bivariée, l'estimateur $\widehat{\Sigma}$ suit une loi de Wichart.

Estimation par la MM

L'estimateur $\widehat{\mu}$ du vecteur moyen μ est donné par

$$\widehat{\mu} = (\widehat{\mu}_1, \widehat{\mu}_2,) = \Big(\frac{1}{n}\sum_{i=1}^{n} X_1^i, \frac{1}{n}\sum_{i=1}^{n} X_2^i\Big). \tag{2.8}$$

L'estimateur de la matrice de variance covariance Σ est donné par

$$\widehat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(X^i - \widehat{\mu}).(X^i - \widehat{\mu})^t. \tag{2.9}$$

Nous remarquons que les estimateurs des paramètres μ et Σ obtenus par la MM sont identiques à ceux obtenus par la MMV.

Modèle de lois Bêta de première espèce

Estimation par la MMV

Soit $Z = (Z^1, \ldots, Z^n)$ un $n-$échantillon aléatoire de loi Bêta bivariée de première espèce du vecteur des paramètres $\theta = (a, b, c)$. La fonction log-vraisemblance est définie par

$$l(\theta) = l(Z; (a, b, c)) = (a-1)\sum_{i=1}^{n}\ln(Z_1^i) + (b-1)\sum_{i=1}^{n}\ln(Z_2^i) + (c-1)\sum_{i=1}^{n}\ln(1 - Z_1^i - Z_2^i) - n\ln(B(a, b, c)). \tag{2.10}$$

Comme le modèle des lois Bêta bivariées de première espèce est une FEG, alors, d'après l'estimation par la MMV, l'estimateur $\widehat{\theta} = (\widehat{a}, \widehat{b}, \widehat{c})$ est solution des équations suivantes :

$$\psi(a) - \psi(a+b+c) = \frac{1}{n}\sum_{i=1}^{n}\ln(Z_1^i), \tag{2.11}$$

$$\psi(b) - \psi(a+b+c) = \frac{1}{n}\sum_{i=1}^{n}\ln(Z_2^i), \tag{2.12}$$

$$\psi(c) - \psi(a+b+c) = \frac{1}{n}\sum_{i=1}^{n} \ln(1 - Z_1^i - Z_2^i), \tag{2.13}$$

où la fontion $\psi(x) = \dfrac{\partial \ln(\Gamma(x))}{\partial x}$ est appelée la fonction digamma.

Estimation par la MM

Pour cet exemple, la fonction h est donnée par

$$h(Z_1, Z_2) = (Z_1, Z_2, Z_1^2, Z_2^2, Z_1 Z_2).$$

Le calcul montre que l'espérence de la fonction h est donnée par

$$K(\theta) = (\frac{a}{a+b+c}, \frac{b}{a+b+c}, \frac{a(a+1)}{(a+b+c)(a+b+c+1)}, \frac{b(b+1)}{(a+b+c)(a+b+c+1)}, \frac{ab}{(a+b+c+1)(a+b+c)}).$$

Ainsi, les estimateurs des paramètres a, b et c sont

$$\widehat{a} = \overline{Z}_1\Big(\frac{\overline{Z}_1(1-\overline{Z}_1)}{S_1^2} - 1\Big), \tag{2.14}$$

$$\widehat{b} = \overline{Z}_2\Big(\frac{\overline{Z}_2(1-\overline{Z}_2)}{S_2^2} - 1\Big), \tag{2.15}$$

$$\widehat{c} = (\overline{Z}_1 - 1)\Big(\frac{\overline{Z}_1^2}{S_1^2} + 1\Big) + \overline{Z}_2\Big(\frac{(1-\overline{Z}_2)^2}{S_2^2} + 1\Big). \tag{2.16}$$

avec $S_j^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(Z_j^i - \overline{Z}_j)^2,\ j = 1, 2.$

Modèle de lois Bêta de deuxième espèce

Estimation par la MMV

Soit $Y = (Y^1, \ldots, Y^n)$ un $n-$échantillon aléatoire qui suit une loi Bêta de deuxième espèce du vecteur des paramètres $\theta = (a, b, c)$. La fonction ln-vraisemblance est définie par

$$l(\theta) = l(Y; (a, b, c)) = (a-1)\sum_{i=1}^{n} \ln(Y_1^i) + (b-1)\sum_{i=1}^{n} \ln(Y_2^i) - (a+b+c)\sum_{i=1}^{n} \ln(1+Y_1^i+Y_2^i) - n\ln(B(a, b, c)). \quad (2.17)$$

Comme le modèle des lois Bêta bivariées de deuxième espèce est une FEG, alors, d'après l'estimation par la MMV, l'estimateur $\widehat{\theta} = (\widehat{a}, \widehat{b}, \widehat{c})$ est solution des équations suivantes :

$$\psi(a) - \psi(a+b+c) = \frac{1}{n}\sum_{i=1}^{n} \ln(Y_1^i) - \frac{1}{n}\sum_{i=1}^{n} \ln(1 + Y_1^i + Y_2^i), \quad (2.18)$$

$$\psi(b) - \psi(a+b+c) = \frac{1}{n}\sum_{i=1}^{n} \ln(Y_2^i) - \frac{1}{n}\sum_{i=1}^{n} \ln(1 + Y_1^i + Y_2^i), \quad (2.19)$$

$$\psi(c) - \psi(a+b+c) = -\frac{1}{n}\sum_{i=1}^{n} \ln(1 + Y_1^i + Y_2^i). \quad (2.20)$$

Remarque 2.2.9. *Du point de vue théorique, une expression explicite des estimateurs des paramètres des lois Bêta de première espèce et des lois Bêta de deuxièmes espèces en utilisant la MMV est difficile à déterminer. Pour cette raison, nous avons adapté dans notre étude la méthode itérative de Newton-Raphson [106] pour résoudre ce problème.*

Estimation par MM

Pour cet exemple, la fonction h est donnée par

$$h(y) = (Y_1, Y_2, Y_1^2, Y_2^2, Y_1Y_2).$$

Un calcul montre que l'espérence de la fonction h est donnée par

$$K(\theta) = \mathbb{E}(h(y)) = (\frac{a}{c-1}, \frac{b}{c-1}, \frac{a(a+1)}{(c-2)(c-1)}, \frac{b(b+1)}{(c-2)(c-1)}, \frac{ab}{(c-1)(c-2)}).$$

Ainsi, les estimateurs des paramètres a, b et c sont

$$\widehat{a} = \overline{Y_1}\Big(1 + \frac{\overline{Y_1}(1+\overline{Y_1})}{S^2_{Y_1}}\Big), \tag{2.21}$$

$$\widehat{b} = \overline{Y_2}\Big(1 + \frac{\overline{Y_2}(1+\overline{Y_2})}{S^2_{Y_2}}\Big), \tag{2.22}$$

$$\widehat{c} = 2 + \frac{\overline{Y_1}(1+\overline{Y_1})}{S^2_{Y_1}}. \tag{2.23}$$

2.2.3 Performance des estimateurs

Dans cette section, nous étudions la performance des estimateurs obtenus par la MMV et la MM sur deux modèles bivariées (Bêta de première espèce et Bêta de deuxième espèce). Les estimateurs peuvent être comparés sur la base de différents critères : le biais, la variance, l'erreur quadratique moyenne et la convergence. Le biais mesure l'écart entre la vraie valeur de la quantité à estimer et la valeur que l'estimateur prend en espérance. La variance quantifie la variabilité autour de l'espérance. Idéalement, un bon estimateur possède un petit biais voire nul, et une petite variance. Il n'est pas possible de réduire simultanement le biais et la variance d'un estimateur, ce qui amène à la définition de l'erreur quadratique moyenne (EQM) qui exprime la combinaison de ces deux quantités. On definit l'EQM de l'estimateur $\widehat{\theta}$ d'un paramètre inconnu θ par :

$$EQM(\widehat{\theta}) = \mathbb{E}[\|\widehat{\theta} - \theta\|^2].$$

Dans de nombreuses études, il est nécessaire de disposer d'échantillons de tailles quelconques de variables de lois connues. Pour cela, nous faisons appel à la simulation qui

consiste à construire à l'aide d'un programme de calcul une suite de nombres $x_1, \ldots, x_n$, chacun suivant la loi voulue indépendamment les uns des autres. Pour créer de tels échantillons, il est nécessaire de disposer a priori d'un générateur de nombres aléatoires u_i uniformément répartis sur l'intervalle $[0, 1]$. Cette méthode nous permet d'estimer le paramètre θ de la loi, et de faire une étude comparative des performances des méthodes d'estimation MMV et MM. La mesure de performance est obtenue en calculant l'EQM de l'estimateur du paramètre θ. Classiquement, cette quantité est estimée à partir d'un certain nombre de réalisations d'un échantillon suivant la loi donnée. Pratiquement, l'estimation du paramètre θ est obtenue de la façon suivante : d'abord nous générons un échantillon de taille n, ensuite nous estimons le paramètre θ par les deux méthodes MMV et MM. Cette étape est réalisée N fois afin de calculer la variation de l'EQM en fonction de la taille de l'échantillon. Afin d'apprécier le comportement de ces estimateurs, nous présentons dans la suite quelques résultats de simulation sur des modèles de lois Bêta de première espèce et de lois Bêta de deuxième espèce. Comme les estimateurs des paramètres du modèle des lois Gaussiennes obtenus par la MMV et la MM sont identiques, nous ne présentons pas des résultats de simulations concernant ce modèle.

Modèle de lois Bêta de première espèce

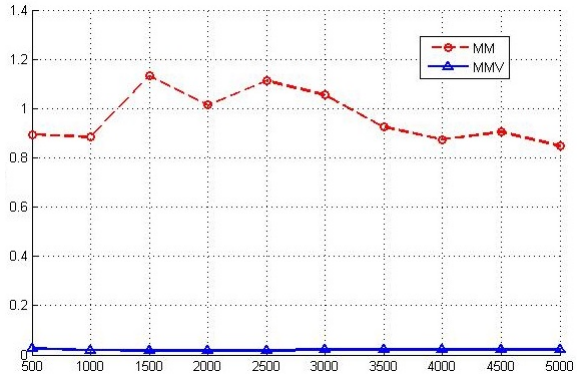
Nous avons généré des réalisations de lois Bêta bivariées de première espèce pour différentes valeurs de a, b et c. Cette expérience a été réalisée pour un échantillon de taille n et pour un nombre de simulations $N = 1000$. Nous présentons dans les figures FIG.2.4.a, FIG.2.4.b et FIG.2.5 la variation de l'EQM en fonction de la taille n de l'échantillon .

FIG.2.4.a est obtenue en fixant $b = 3$ et $c = 0.3$ et en variant a. Les deux courbes dans

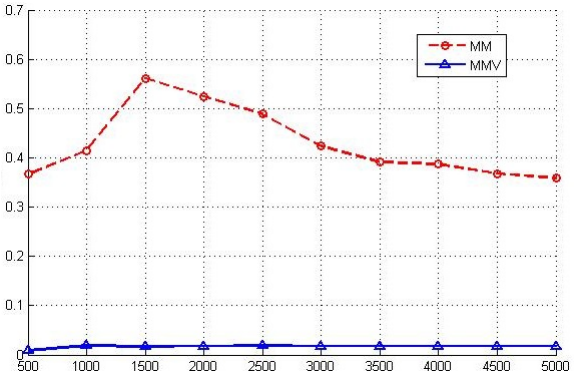
cette figure représentent la variation de l'EQM pour $a = 0.105$.

FIG.2.4.b est obtenue en fixant $a = 3$ et $c = 10$ et en variant b. Les deux courbes dans cette figure représentent la variation de l'EQM pour $b = 2.5$.

FIG.2.5 est obtenue en fixant $a = 3$ et $b = 10$ et en variant c. Les deux courbes dans cette figure représentent la variation de l'EQM pour $c = 45$. Les courbes représentées par "$\circ$" correspondent à la variation de l'EQM en fonction de n obtenues par la MM. Les courbes représentées par "$\triangle$" correspondent à la variation de l'EQM en fonction n obtenues par la MMV. D'après ces trois figures, nous remarquons que les estimateurs des paramètres a, b et c obtenus par la MMV sont plus performants que ceux obtenus par la MM.



(a) EQM du paramètre a.



(b) EQM du paramètre b.

FIG. 2.4 – L'EQM en fonction de n pour les paramètres a et b.

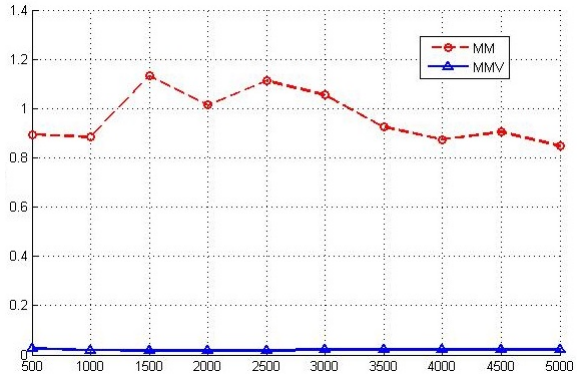


FIG. 2.5 – L'EQM en fonction de n pour le paramètre c.

Modèle de lois Bêta de deuxième espèce

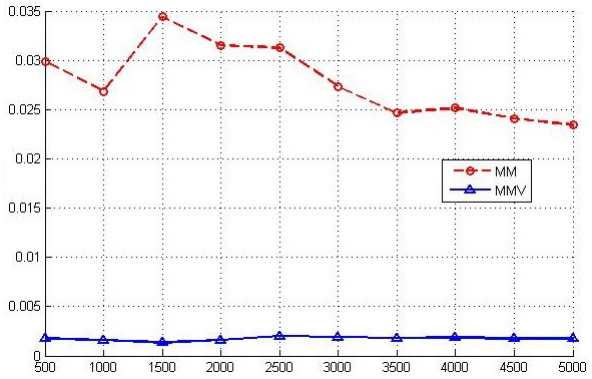
Nous avons généré des réalisations de lois Bêta bivariées de deuxième espèce pour différentes valeurs de a, b et c. Cette expérience a été réalisée pour un échantillon de taille n et pour un nombre de simulations $N = 1000$. Nous présentons dans les figures FIG.2.6.a, FIG.2.6.b et FIG.2.7 la variation de l'EQM en fonction de la taille de l'échantillon n.

FIG.2.6.a est obtenue en fixant $b = 3$ et $c = 2.3$ et en variant a. Les deux courbes dans cette figure représentent la variation de l'EQM pour $a = 1.5$.

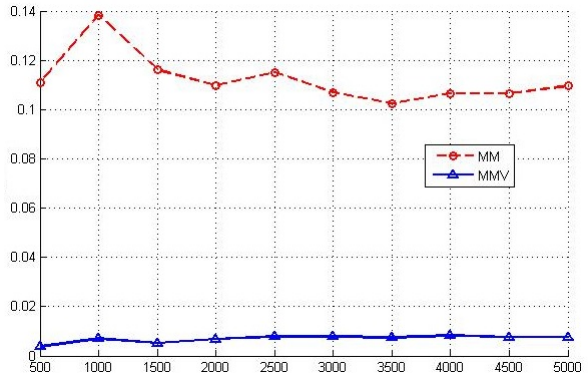
FIG.2.6.b est obtenue en fixant $a = 4$ et $c = 9$ et en variant b. Les deux courbes dans cette figure représentent la variation de l'EQM pour $b = 4$.

FIG.2.7 est obtenue en fixant $a = 10$ et $b = 20$ et en variant c. Les deux courbes dans cette figure représentent la variation de l'EQM pour $c = 45$.

Les courbes représentées par "$\circ$" correspondent à la variation de l'EQM en fonction de n obtenues par la MM. Les courbes représentées par "$\triangle$" correspondent à la variation de l'EQM en fonction de n obtenues par la MMV. D'après ces trois figures, nous remarquons que les estimateurs des paramètres a, b et c obtenus par la MMV sont plus performants que ceux obtenus par la MM.



(a) EQM du paramètre a



(b) EQM du paramètre b

FIG. 2.6 – L'EQM en fonction de n pour les paramètres a et b.

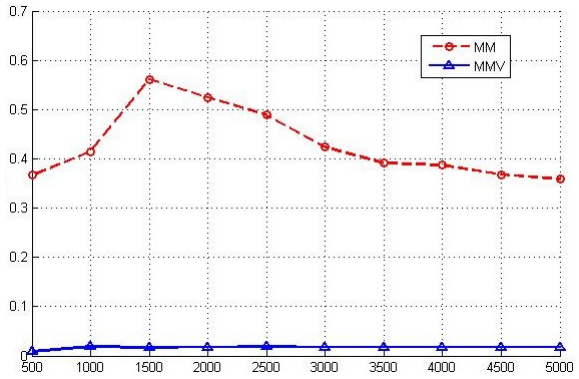


FIG. 2.7 – L'EQM en fonction de n pour le paramètre c.

Afin de généraliser la performance de ces deux méthodes d'estimations sur d'autres modèles de lois multivariées, nous proposons dans la suite de ce chapitre un nouveau modèle

de lois basé sur la loi de Riesz.

2.3 Proposition d'un nouveau modèle multivarié

Les lois de Wishart sur le cône des matrices symétriques définies positives ont fait l'objet de plusieurs études durant ces dernières années [90][95]. Elles sont considérées comme une généralisation des lois Gamma sur $\mathbb{R}$. L'extension de ces lois permet de définir les lois de Riesz sur le cône des matrices symétriques définies positives [47]. Dans ce chapitre, nous nous sommes intéressés à cette loi qui fera l'objet du développement d'un nouveau modèle de lois multivariées. Ce nouveau modèle est appelé modèle de lois de la diagonale modifiée de Riesz. Dans la suite, nous rappelons la loi de Riesz ainsi que ses propriétés. Puis, nous introduisons le nouveau modèle ainsi que les méthodes d'estimation des paramètres qui le caractérisent. Enfin, nous étudions la performance des estimateurs de ces paramètres en se basant sur le calcul de l'erreur quadratique moyenne.

2.3.1 Rappel sur la loi de Riesz

Dans ce paragraphe, nous nous intéressons à la loi de Riesz définie sur le cône des matrices symétriques définies positives [47]. Cette loi est basée sur la notion des puisances généralisées [47]. Dans notre étude, nous nous plaçons dans l'espace vectoriel $L_s(\mathbb{R}^r)$ des matrices carrées d'ordre r symétriques à coéfficients dans $\mathbb{R}$, muni du produit scalaire

$$\langle x, y\rangle = tr(x.y),$$

où $tr(.)$ représente la fonction trace usuelle d'une matrice carrée.

Dans tout ce qui suit, nous désignons par Ω le cône des matrices symétriques définies

positives de $L_s(\mathbb{R}^r)$. Notons que la dimension de $L_s(\mathbb{R}^r)$ est égale à $n = \frac{r(r+1)}{2}$.

Définition 2.3.1. *Soit $s_* = (s_1, \cdots, s_r) \in \mathbb{R}^r$, la puissance généralisée d'un élément $x = (x_{ij})_{1\leq i,j\leq r} \in \Omega$ est définie par $\Delta_{s_*}(x) = \Delta_1(x)^{s_1-s_2}\Delta_2(x)^{s_2-s_3}\ldots\Delta_r(x)^{s_r}$, où $\Delta_k(x)$ est le déterminant de la matrice $P_k(x) = (x_{ij})_{1\leq i,j\leq k}$, $1 \leq k \leq r$.*

Lorsque $s_* = (p, \cdots, p)$, avec $p \in \mathbb{R}$, $\Delta_s(x) = (\det x)^p$.

Définition 2.3.2. *La fonction gamma Ω est définie par l'intégrale suivante*

$$\Gamma_\Omega(s_*) = \int_\Omega e^{-\ tr(x)}\Delta_{s_*-\frac{n}{r}}(x)dx. \tag{2.24}$$

Pour $y \in \Omega$ et $s_* = (s_1, \cdots, s_r) \in \mathbb{R}^r$ tel que $s_i > \frac{(i-1)}{2}$ quelque soit $1 \leq i \leq r$, nous avons

$$\int_\Omega e^{-\langle x,y\rangle}\Delta_{s_*-\frac{n}{r}}(x)dx = \Gamma_\Omega(s_*)\Delta_{s_*}(y^{-1}). \tag{2.25}$$

Notons que la fonction Γ_Ω peut être exprimée comme produit de fonctions gamma ordinaires

$$\Gamma_\Omega(s_*) = \pi^{\frac{r(r-1)}{4}}\prod_{j=1}^{r}\Gamma(s_j - \frac{j-1}{2}), \qquad s_j > \frac{j-1}{2}.$$

Théorème 2.3.3. *Soit $s_* = (s_1, s_2, \ldots, s_r) \in \mathbb{R}^r$ tel que $s_i > \dfrac{i-1}{2}$, $1 \leq i \leq r$, et $\sigma \in \Omega$, alors la mesure absolument continue de Riesz est définie par*

$$L_{R_{s_*}}(dx) = \frac{\Delta_{s_*-\frac{n}{r}}(x)}{\Gamma_\Omega(s_*)}\mathbf{1}_\Omega(x)dx.$$

La FEN F engendrée par la mesure R_{s_*} est définie par l'ensemble

$$F = F(R_{s_*}) = \left\{R(s_*, \sigma) = \frac{e^{-\langle\sigma,x\rangle}}{\Delta_{s_*}(\sigma^{-1})}R_{s_*},\ \sigma \in \Omega\right\}.$$

Gindikin [39] a défini la transformée de Laplace de la mesure de Riesz.

Théorème 2.3.4. *Soit $\theta \in -\Omega$, alors la transformée de Laplace de la mesure absolument continue de Riesz est définie par*

$$L_{R_{s_*}}(\theta) = \int_E e^{\langle \theta, x \rangle} R_{s_*}(dx) = \Delta_{s_*}(-\theta^{-1}). \tag{2.26}$$

Définition 2.3.5. *La matrice aléatoire X suit la loi de Riesz sur Ω, de paramètre $s_* = (s_1, \ldots, s_r)$ tel que $s_i > \dfrac{i-1}{2}$ et $\sigma \in \Omega$, si sa densité de probabilité est définie par*

$$R(s_*, \sigma)(dx) = \frac{e^{-<\sigma,x>} \Delta_{s_* - \frac{n}{r}}(x)}{\Gamma_\Omega(s_*) \Delta_{s_*}(\sigma^{-1})} \mathbf{1}_\Omega(x) dx.$$

Dans le cas où $s_1 = s_2 = \ldots = s_r = p > \dfrac{r-1}{2}$, la loi de Riesz est réduite à une loi de Wishart de densité définie par

$$W(p, \sigma)(dx) = \frac{1}{\Gamma_\Omega(p) \det(\sigma^{-p})} e^{-<\sigma,x>} \det(x)^{p - \frac{n}{r}} \mathbf{1}_\Omega(x) dx. \tag{2.27}$$

Dans le cas où $r = 2$, nous proposons un modèle de la diagonale modifée de la loi de Riesz [11].

2.3.2 Loi de la diagonale modifiée de Riesz

Pour définir notre nouveau modèle, nous avons pris une matrice aléatoire X d'ordre 2 de vecteur paramètres $s_* = (s_1, s_2) \in]0, +\infty[\times]\frac{1}{2}, +\infty[$ et $\theta \in I_2 - \Omega$, avec I_2 la matrice identité d'ordre 2. Nous avons appliqué une transformation affine de matrice $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix}$, où $\Sigma_{11} > 0$, $\Sigma_{22} > 0$ et $\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2 > 0$ à X. Posant $\widetilde{X} = \Sigma X$, le modèle proposé est défini par le vecteur aléatoire $Y = (Y_1, Y_2) = diag\left(\widetilde{X}\right)$, où $diag\left(\widetilde{X}\right)$ représente la diagonale de la matrice aléatoire $\widetilde{X}$.

À partir de la transformée de Laplace de la matrice aléatoire X donnée par

$$L_X(\theta) = \int_E e^{\langle \theta, x \rangle} R_{s_*}(dx) = \Delta_{s_*}((I_2 - \theta)^{-1}), \tag{2.28}$$

nous déduisons la transformée de Laplace du vecteur aléatoire Y qui caractérise notre nouveau modèle. La définition de cette transformée de Laplace est donnée dans la proposition 2.3.6.

Proposition 2.3.6. *Soient* $s_* = (s_1, s_2) \in]0, +\infty[\times]\frac{1}{2}, +\infty[$ *et* $\theta\Sigma \in I_2 - \Omega$, *alors la transformée de Laplace du vecteur aléatoire* Y *est définie par*

$$L_Y(\theta) = [(1 - \theta_{11}\Sigma_{11})(1 - \theta_{22}\Sigma_{22}) - \theta_{11}\theta_{22}\Sigma_{12}^2]^{-s_1}[1 - \theta_{22}\Sigma_{22}]^{-(s_2 - s_1)}. \tag{2.29}$$

Preuve. *D'une part, pour tout* $\theta \in I_2 - \Omega$, *nous avons*

$$L_{\widetilde{X}}(\theta) = E\left(e^{tr(\theta\widetilde{X})}\right) = E\left(e^{tr(\theta\Sigma X)}\right) = L_X(\theta\Sigma) = \Delta_{s_*}((I_2 - \theta\Sigma)^{-1}).$$

et d'autre part, si $\theta_{ij} = 0$, $\forall\ i \neq j$, *alors*

$L_Y(\theta) = \Delta_{s_*}((I_2 - \theta\Sigma)^{-1}) = [1 - \theta_{22}\Sigma_{22}]^{s_1 - s_2}[(1 - \theta_{11}\Sigma_{11})(1 - \theta_{22}\Sigma_{22}) - \theta_{11}\theta_{22}\Sigma_{12}^2]^{-s_2}.\square$

À partir de la transformée de Laplace L_Y, nous déterminons la densité de probabilité du vecteur aléatoire $Y = (Y_1, Y_2)$. Deux cas se présentent pour définir cette densité de probabilité.

2.3.2.a 1er cas : $s_1 = s_2 = s$.

Proposition 2.3.7. *Soient* $s_1 = s_2 = s$, $\theta \in -\Omega$ *et* Σ *une matrice symétrique définie positive. Alors, le vecteur aléatoire* Y *suit la loi Gamma bivariée de Kibble [27] de densité de probabilité définie par*

$$f_Y(y_1, y_2) = \exp(-\frac{\Sigma_{22}y_1 + \Sigma_{11}y_2}{\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2}) \times \frac{(y_1y_2)^{s-1}}{(\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2)^s\Gamma(s)}$$
$$\times f_s(\delta y_1y_2)\mathbf{1}_{[0,+\infty[^2}(y_1, y_2), \tag{2.30}$$

avec $\delta = \dfrac{\Sigma_{12}^2}{(\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2)^2}$ *et la fonction* f_s *est définie par*

$$f_s(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(s+k)k!}, \ \forall \ z \ \in \ \mathbb{R}.$$

Preuve. *Si* $s_1 = s_2 = s$, $\Sigma_{11} > 0$, $\Sigma_{22} > 0$ *et* $\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2 > 0$, *alors la transformée de Laplace du vecteur aléatoire* Y *est définie par*

$$L_Y(\theta) = [(1 - \theta_{11}\Sigma_{11})(1 - \theta_{22}\Sigma_{22}) - \theta_{11}\theta_{22}\Sigma_{12}^2]^{-s}. \tag{2.31}$$

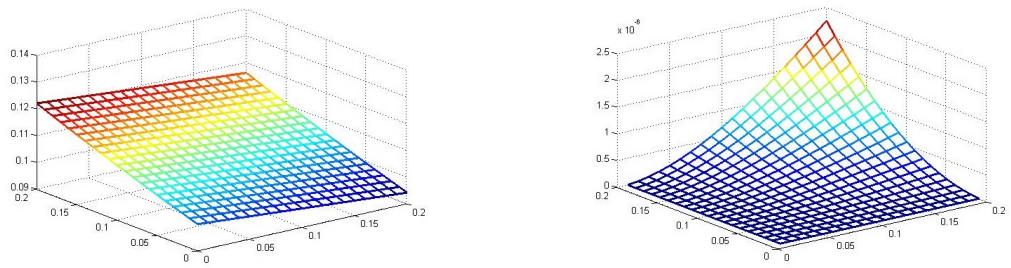
D'après Choua [27], pour tout $\theta \in -\Omega$, *l'équation (2.31) représente la transformée de Laplace de la loi Gamma bivariée de Kibble.* □

La loi Gamma bivariée de Kibble a été introduite en 1933 par Wicksell [102]. En 1941, Kibble a considéré que la distribution bivariée du vecteur aléatoire $Y = (Y_1, Y_2)$ de paramètres $s > 0$, $\Sigma_{11} > 0$, $\Sigma_{22} > 0$ et $\Sigma_{12}^2 \in]0, \Sigma_{11}\Sigma_{22}[$ (i-e,$\Sigma_{1.2} = \Sigma_{11}\Sigma_{22} - \Sigma_{12}^2 > 0$), est de densité f_Y définie par l'équation (2.30).

Ce modèle est qualifié de loi Gamma bivariée car les lois marginales de Y_1 et Y_2 sont des lois $\gamma(s, \Sigma_{11})$ et $\gamma(s, \Sigma_{22})$.

Dans le cas où $\Sigma_{12}^2 = 0$, nous pouvons déduire que la transformée de Laplace de la loi de Kibble peut s'écrire sous la forme d'un produit de deux transformées de Laplace de la loi $\gamma(s, \Sigma_{11})$ et de la loi $\gamma(s, \Sigma_{22})$.

La forme de la densité de probabilité du vecteur aléatoire $Y = (Y_1, Y_2)$ varie selon les valeurs de ses paramètres (voir FIG. 2.8.)



(a) $s = 1,\ \Sigma_{11} = 10, \Sigma_{22} = 3, \Sigma_{12}^2 = 2$ (b) $s = 3,\ \Sigma_{11} = 1, \Sigma_{22} = 10, \Sigma_{12}^2 = 5$

FIG. 2.8 – Représentation de la densité de la loi Gamma bivariée de Kibble.

Moments

Le calcul des moments du vecteur aléatoire Y est obtenu en dérivant la transformée de Laplace définie par l'équation (2.31) au point $(0, 0)$ par rapport à θ_i et $\theta_i\theta_j$ $(i, j = 1, 2)$. D'où, les expressions des moyennes des marges, de la covariance et du coefficient de corrélation de (Y_1, Y_2) qui suit la loi Gamma bivariée de Kibble sont données par

$$E[Y_1] = s\Sigma_{11},\ E[Y_2] = s\Sigma_{22} \tag{2.32}$$

$$\mathbb{V}ar(Y_1) = s\Sigma_{11}^2,\ \mathbb{V}ar(Y_2) = s\Sigma_{22}^2, \tag{2.33}$$

$$\mathbb{C}ov(Y_1, Y_2) = s\Sigma_{12}^2, \quad r = \mathbb{C}orr(Y_1, Y_2) = \frac{\Sigma_{12}^2}{\Sigma_{11}\Sigma_{22}}. \tag{2.34}$$

D'après la définition du coefficient de corrélation donnée par l'équation (2.34), la loi Gamma bivariée de Kibble est à dépendance positive puisque $0 < r < 1$.

2.3.2.b 2eme cas : $s_2 > s_1$.

Dans le cas où $s_2 > s_1$, l'expression de la transformée de Laplace du vecteur aléatoire Y définie par l'équation (2.29) est égale au produit des transformées de Laplace d'une loi Gamma univariée $\gamma(s_2 - s_1, \Sigma_{22})$ et d'une loi Gamma bivariée de Kibble de paramètres s_1,

Σ_{11}, Σ_{22} et Σ_{12}^2. Nous pouvons interpréter d'une autre façon le vecteur aléatoire Y comme la somme d'un vecteur aléatoire $Z = (Z_1, Z_2)$ et d'une variable aléatoire T indépendantes entre eux et qui suivent respectivement la loi Gamma bivariée de Kibble et la loi $\gamma(s_2 - s_1, \Sigma_{22})$.

La densité de probabilité du vecteur aléatoire Y est donnée sous forme d'une proposition.

Proposition 2.3.8. *Soient $0 < s_1 < s_2$ et Σ une matrice symétrique définie positive. Alors, la densité de probabilité de Y est définie par*

$$f_Y(y_1, y_2) = \frac{1}{\Sigma_{22}^{s_2-s_1}} \Big(\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2\Big)^{-s_1} e^{(-\frac{\Sigma_{22}y_1 + \Sigma_{11}y_2}{\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2})} \frac{y_1^{s_1-1} y_2^{s_2-1}}{\Gamma(s_1)\Gamma(s_2)}$$
$$\times \Phi_3\Big(s_2 - s_1; s_2; hy_2, \delta y_1 y_2\Big) \mathbf{1}_{[0,+\infty[^2}(y_1, y_2), \quad (2.35)$$

avec $\delta = \dfrac{\Sigma_{12}^2}{(\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2)^2}$, $h = \dfrac{\Sigma_{12}^2}{\Sigma_{22}(\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2)}$ *et*
$\Phi_3(\alpha; \beta; y_1, y_2) = \displaystyle\sum_{n,m\geq 0} \frac{(\alpha)_n}{(\beta)_{n+m} n! m!} y_1^n y_2^n$, *où* $(\alpha)_n$ *est le symbole de Pochhammer [1] défini par* $(\alpha)_n = \dfrac{\Gamma(\alpha + n)}{\Gamma(\alpha)}$, $\forall n \in \mathbb{N}$ *et* $\alpha > 0$.

Preuve. *Comme Y est la somme d'un vecteur aléatoire Z de densité f_Z et d'une variable aléatoire T de densité f_T, nous pouvons exprimer la densité de probabilité de Y de la façon suivante*

$$\begin{aligned} f_Y(y_1, y_2) &= \int_0^{+\infty} f_Z(y_1, u) f_T(y_2 - u) \boldsymbol{1}_{\{y_2-u>0\}} du \\ &= \frac{\Sigma_{22}^{s_1-s_2} y_1^{s_1-1} \exp(-\frac{\Sigma_{22}y_1}{\Sigma_{1.2}})}{\Sigma_{1.2}^{s_1}\Gamma(s_1)\Gamma(s_2 - s_1)} \int_0^{y_2} \exp(\frac{-(y_2 - u)}{\Sigma_{22}} - \frac{\Sigma_{11}u}{\Sigma_{1.2}})(y_2 - u)^{s_2-s_1-1} \\ &\qquad \times u^{s_1-1} f_{s_1}(\delta y_1 u) du, \quad (2.36) \end{aligned}$$

avec $\delta = \dfrac{\Sigma_{12}^2}{(\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2)^2}$.

En faisant le changement de variable $u = y_2(1-t)$, *l'équation (2.36) devient*

$$f_Y(y_1, y_2) = \frac{\Sigma_{22}^{s_1-s_2} y_1^{s_1-1} y_2^{s_2-1} \exp(-\dfrac{\Sigma_{22}y_1 + \Sigma_{11}y_2}{\Sigma_{1.2}})}{\Sigma_{1.2}^{s_1}\Gamma(s_1)\Gamma(s_2-s_1)} \int_0^1 \exp\Big(y_2 t(\frac{\Sigma_{11}}{\Sigma_{1.2}} - \frac{1}{\Sigma_{22}})\Big)$$
$$\times\, t^{s_2-s_1-1}(1-t)^{s_1-1} f_{s_1}(\delta y_1 y_2(1-t))dt. \quad (2.37)$$

En remplaçant la fonction f_{s_1} *par son expression définie dans la Proposition 2.3.7 et en utilisant le développement en série entière de la fonction exponentielle, nous obtenons*

$$f_Y(y_1, y_2) = \frac{\Sigma_{22}^{s_1-s_2} y_1^{s_1-1} y_2^{s_2-1} \exp(-\dfrac{\Sigma_{22}y_1 + \Sigma_{11}y_2}{\Sigma_{1.2}})}{\Sigma_{1.2}^{s_1}\Gamma(s_1)\Gamma(s_2-s_1)} \sum_{m,n\geq 0} \frac{(y_1y_2\delta)^m (y_2h)^n}{m!n!\Gamma(m+s_1)}$$
$$\times \int_0^1 t^{n+s_2-s_1-1}(1-t)^{m+s_1-1}dt, \quad (2.38)$$

avec $h = \dfrac{\Sigma_{12}^2}{\Sigma_{22}(\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2)}$.

En utilisant le fait que, d'une part $\displaystyle\int_0^1 t^{n+s_2-s_1-1}(1-t)^{m+s_1-1}dt = \frac{\Gamma(n+s_2-s_1)\Gamma(m+s_1)}{\Gamma(m+n+s_2)}$ *et d'autre part* $\Gamma(n+s_2-s_1) = (s_2-s_1)_n\Gamma(s_2-s_1)$ *et* $\Gamma(m+n+s_2) = (s_2)_{m+n}\Gamma(s_2)$, *nous obtenons*

$$f_Y(y_1, y_2) = \frac{\Sigma_{22}^{s_1-s_2} y_1^{s_1-1} y_2^{s_2-1} \exp(-\dfrac{\Sigma_{22}y_1 + \Sigma_{11}y_2}{\Sigma_{1.2}})}{\Sigma_{1.2}^{s_1}\Gamma(s_1)\Gamma(s_2)} \sum_{m,n\geq 0} \frac{(s_2-s_1)_n}{(s_2)_{m+n}m!n!}(y_1y_2\delta)^m(y_2h)^n$$
$$= \frac{\Sigma_{22}^{s_1-s_2} y_1^{s_1-1} y_2^{s_2-1} \exp(-\dfrac{\Sigma_{22}y_1 + \Sigma_{11}y_2}{\Sigma_{1.2}})}{\Sigma_{1.2}^{s_1}\Gamma(s_1)\Gamma(s_2)} \Phi_3\Big(s_2-s_1; s_2; hy_2, \delta y_1y_2\Big). \quad (2.39)$$

□

Remarque 2.3.9. *Nous remarquons que si* $s_1 = s_2 = s$, *la fonction* $\Phi_3\Big(s_2-s_1; s_2; hy_2, \delta y_1y_2\Big) = \Gamma(s)f_s(\delta y_1y_2)$ *et l'expression de la densité de probabilité de* Y *définie par l'équation (2.35) est la même que celle donnée par l'équation (2.30).*

Moments

En dérivant la transformée de Laplace du vecteur aléatoire Y définie par l'équation (2.29)au point $(0,0)$ par rapport à θ_i et $\theta_i\theta_j$ $(i,j=1,2)$, nous obtenons les expressions des moments des marges, de covariance et du coefficient de corrélation de Y qui sont données par

$$E[Y_1] = s_1\Sigma_{11},\ E[Y_2] = s_2\Sigma_{22} \tag{2.40}$$

$$var(Y_1) = s_1\Sigma_{11}^2,\ var(Y_2) = s_2\Sigma_{22}^2 \tag{2.41}$$

$$cov(Y_2, Y_2) = s_1\Sigma_{12}^2,\ r = \mathbb{C}orr(Y_1, Y_2) = \sqrt{\frac{s_1}{s_2}}\frac{\Sigma_{12}^2}{\Sigma_{11}\Sigma_{22}}. \tag{2.42}$$

D'après la définition du coefficient de corrélation r donnée par l'équation (2.42), nous remarquons que $0 < r < 1$. Donc les composantes du vecteur aléatoire $Y = (Y_1, Y_2)$ sont à dépendance positive.

Dans la suite de notre travail, nous supposons que les paramètres s_1 et s_2 sont fixés et que le vecteur aléatoire Y est caractérisé par le vecteur des paramètres inconnus $\theta = (\Sigma_{11}, \Sigma_{22}, \Sigma_{12}^2)$. D'où, un problème d'estimation des paramètres qui se pose.

2.3.3 Estimation des paramètres

Afin de résoudre le problème d'estimation des paramètres du vecteur aléatoire Y, nous avons fait appel à la MMV et à la MM.

2.3.3.a 1er cas : $s_1 = s_2 = s$.

Estimation par la MMV

Soit $(Y^1, \ldots, Y^n)$ un $n-$échantillon aléatoire qui suit la loi Gamma bivariée de Kibble. Le logarithme de la fonction de vraisemblance est défini par

$$l(\theta) = -ns\ln(\Sigma_{1.2}) - \frac{n\Sigma_{22}\overline{Y}_1}{\Sigma_{1.2}} + (s-1)\sum_{i=1}^{n}\ln(Y_1^iY_2^i) - \ln\Gamma(s) - \frac{n\Sigma_{11}\overline{Y}_2}{\Sigma_{1.2}} + \sum_{i=1}^{n}\ln(f_s(\delta Y_1^iY_2^i)), \quad (2.43)$$

avec $\overline{Y}_j = \frac{1}{n}\sum_{i=1}^{n} Y_j^i,\ j = 1, 2,\ \Sigma_{1.2} = \Sigma_{11}\Sigma_{22} - \Sigma_{12}^2$ et $\delta = \dfrac{\Sigma_{12}^2}{\Sigma_{1.2}^2}$.

En dérivant la fonction $l(\theta)$ respectivement par rapport à Σ_{11}, Σ_{22}, et Σ_{12}^2, nous obtenons les équations suivantes :

$$-ns + \frac{n\Sigma_{22}\overline{Y}_1}{\Sigma_{1.2}} - \frac{n\overline{Y}_2}{\Sigma_{22}} + \frac{n\Sigma_{11}\overline{Y}_2}{\Sigma_{1.2}} - 2\frac{\Sigma_{12}^2}{\Sigma_{1.2}^2}H = 0, \quad (2.44)$$

$$-ns + \frac{n\Sigma_{11}\overline{Y}_2}{\Sigma_{1.2}} - \frac{n\overline{Y}_1}{\Sigma_{11}} + \frac{n\Sigma_{22}\overline{Y}_1}{\Sigma_{1.2}} - 2\frac{\Sigma_{12}^2}{\Sigma_{1.2}^2}H = 0, \quad (2.45)$$

$$ns - \frac{n\Sigma_{22}\overline{Y}_1}{\Sigma_{1.2}} - \frac{n\Sigma_{11}\overline{Y}_2}{\Sigma_{1.2}} + \frac{\Sigma_{12}^2 + \Sigma_{11}\Sigma_{22}}{\Sigma_{1.2}^2}H = 0, \quad (2.46)$$

avec $H = \displaystyle\sum_{i=1}^{n} Y_1^iY_2^i\frac{f_{s+1}(\delta Y_1^iY_2^i)}{f_s(\delta Y_1^iY_2^i)}$ et $f_{s_+1} = f_s'$.

En sommant l'équation (2.44) et l'équation (2.46), nous obtenons l'expression suivante

$$\frac{n\overline{Y}_2}{\Sigma_{22}} = \frac{H}{\Sigma_{1.2}}.$$

En remplaçant $\dfrac{H}{\Sigma_{1.2}}$ par $\dfrac{n\overline{Y}_2}{\Sigma_{22}}$ dans les équations (2.44) et (2.45), nous montrons que les estimateurs de Σ_{11} et Σ_{22} par la MMV sont de la forme

$$\widehat{\Sigma}_{11} = \frac{\overline{Y}_1}{s},\ \widehat{\Sigma}_{22} = \frac{\overline{Y}_2}{s}. \quad (2.47)$$

En remplaçant les paramètres Σ_{11} et Σ_{22} par leurs estimateurs dans l'équation (2.46), nous déterminons l'estimateur de Σ_{12}^2 qui est la racine de l'équation suivante :

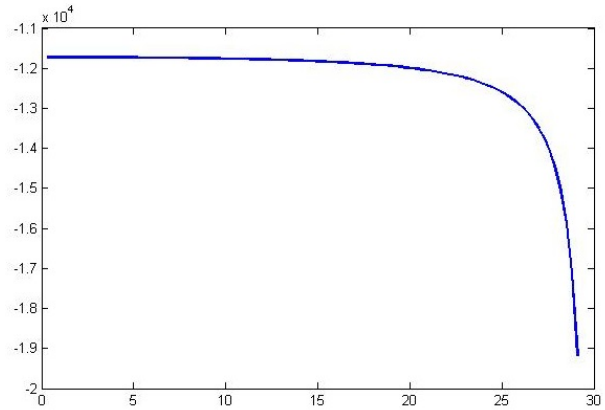
$$\phi(\Sigma_{12}^2) = n - \frac{s}{\overline{Y}_1\overline{Y}_2 - s^2\Sigma_{12}^2}\sum_{i=1}^{n} Y_1^i Y_2^i \frac{f_{s+1}(\widehat{\delta}Y_1^iY_2^i)}{f_s(\widehat{\delta}Y_1^iY_2^i)} = 0, \tag{2.48}$$

avec $\widehat{\delta} = \dfrac{s^4\Sigma_{12}^2}{(\overline{Y}_1\overline{Y}_2 - s^2\Sigma_{12}^2)^2}$.

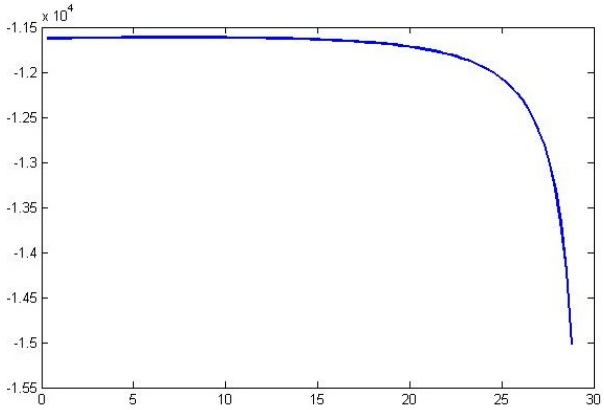
Du point de vu théorique, nous ne pouvons pas déterminer une expression explicite de $\widehat{\Sigma}_{12}^2$. Nous avons adopté une méthode numérique de Newton-Raphson afin de résoudre ce problème. L'estimateur du maximum de vraisemblance existe si la fonction log-vraisemblance est concave sur l'ensemble des paramètres. La propriété de concavité de la fonction $l(\theta)$ par rapport à Σ_{12}^2 n'est plus vérifiée quelque soit le paramètre Σ_{12}^2 appartient à son domaine $]0, \Sigma_{11}\Sigma_{22}[$. En effet, nous pouvons montrer numériquement que la fonction

$$g : \Sigma_{12}^2 \longmapsto l(\widehat{\Sigma}_{11}, \widehat{\Sigma}_{22}, \Sigma_{12}^2)$$

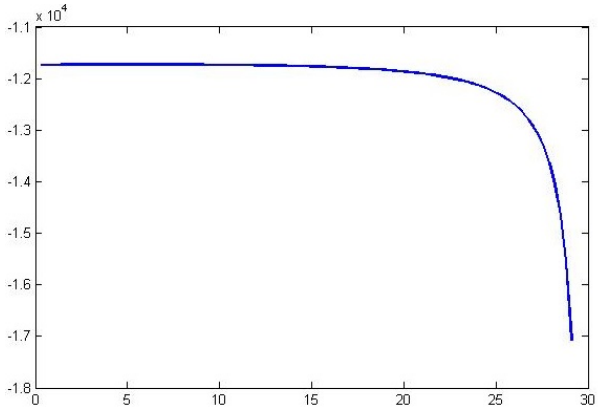
n'est pas forcément concave sur son domaine $]0, \Sigma_{11}\Sigma_{22}[$. Pour ce fait, nous avons simulé un échantillon aléatoire de taille $n = 1000$ distribué selon une loi Gamma bivariée de Kibble, de paramètres $s = 3$, $\Sigma_{11} = 5$, $\Sigma_{22} = 6$ et pour différentes valeurs de $\Sigma_{12}^2 \in]0, \Sigma_{11}\Sigma_{22}[$. Puis nous avons estimé Σ_{11} et Σ_{22}. Nous remarquons que lorsque la valeur de Σ_{12}^2 est proche de $\widehat{\Sigma}_{11}\widehat{\Sigma}_{22}$, la fonction g n'est pas concave (voir FIG.2.9 et FIG.2.10). Dans ce cas, d'après les courbes présentées dans les figures FIG.2.9 et FIG.2.10, nous constatons que le maximum est situé autour de la valeur théorique $\Sigma_{11}\Sigma_{22}$ et il est unique.



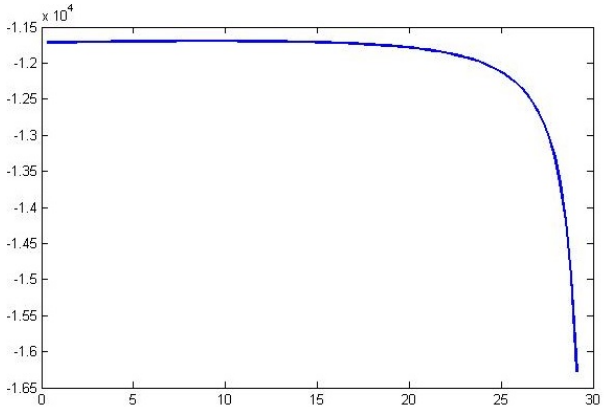
$\Sigma_{12}^2 = 1$



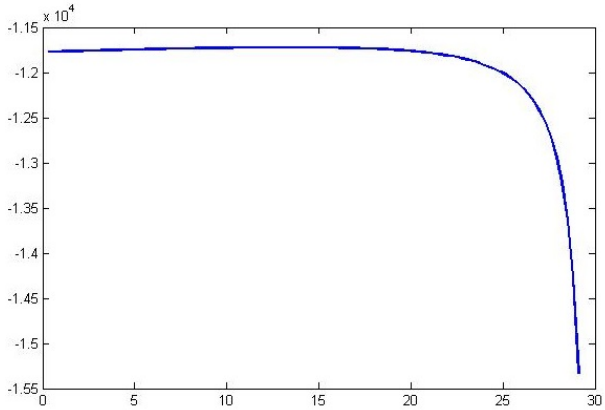
$\Sigma_{12}^2 = 3$



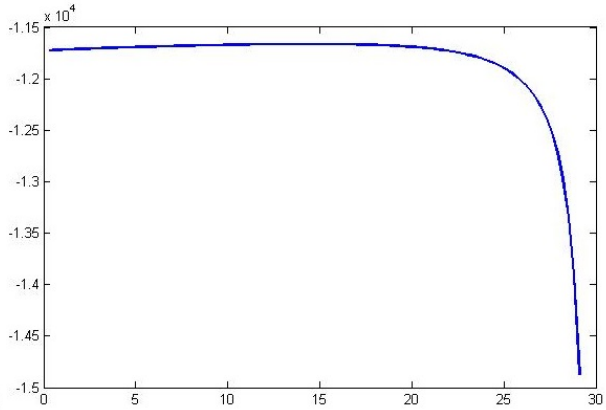
$\Sigma_{12}^2 = 6$



$\Sigma_{12}^2 = 9$

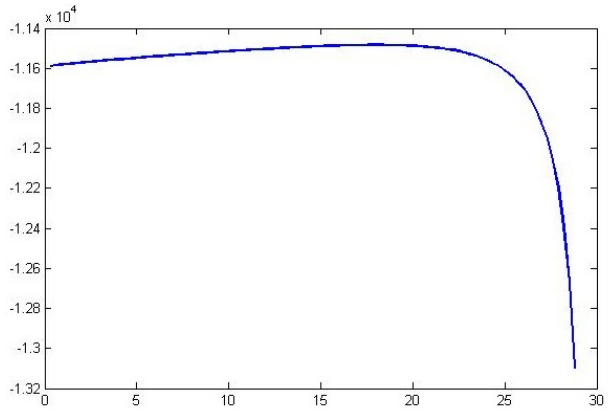


$\Sigma_{12}^2 = 12$

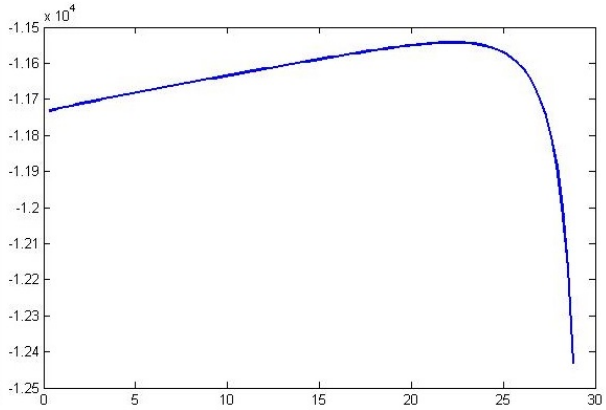


$\Sigma_{12}^2 = 15$

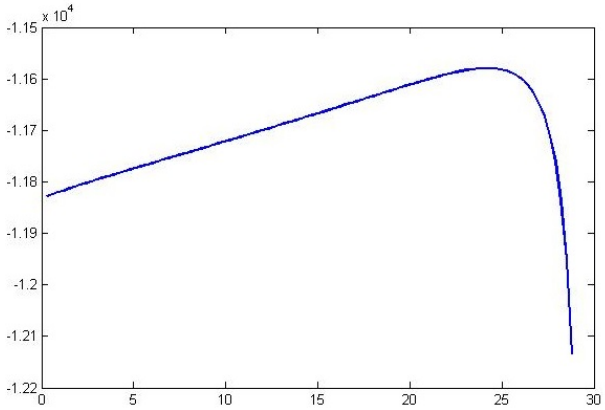
FIG. 2.9 – Représentation de la fonction g pour $s = 3$ et pour différentes valeurs de Σ_{12}^2.



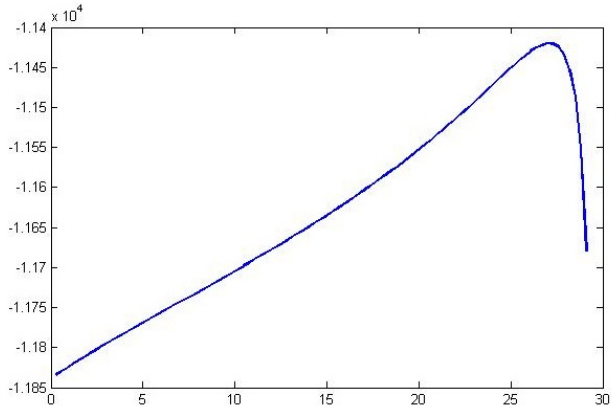
$\Sigma_{12}^2 = 18$



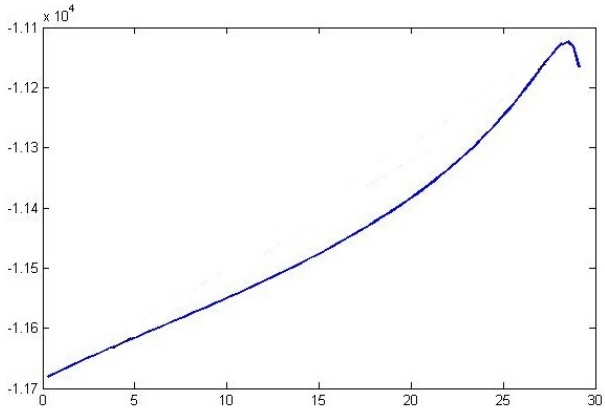
$\Sigma_{12}^2 = 21$



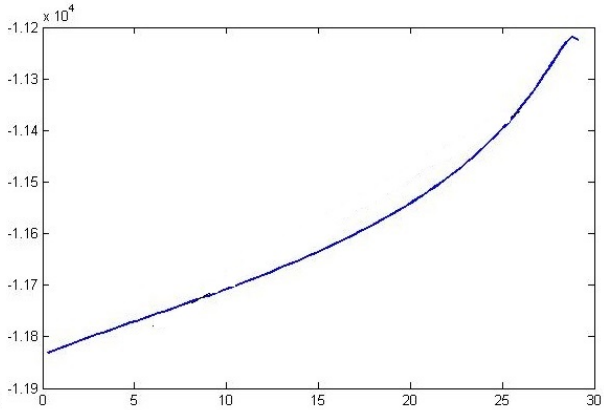
$\Sigma_{12}^2 = 24$



$\Sigma_{12}^2 = 27$



$\Sigma_{12}^2 = 28.5$



$\Sigma_{12}^2 = 29.1$

FIG. 2.10 – Représentation de la fonction g pour $s = 3$ et pour différentes valeurs de Σ_{12}^2.

Estimation par la MM

L'estimateur de $\theta = (\Sigma_{11}, \Sigma_{22}, \Sigma_{12}^2)$ est donné par

$$\widehat{\Sigma}_{11} = \frac{\overline{Y}_1}{s},\ \widehat{\Sigma}_{22} = \frac{\overline{Y}_2}{s} \tag{2.49}$$

$$\widehat{\Sigma}_{12}^2 = \frac{1}{ns}\sum_{i=1}^{n}(Y_1^{(i)} - \overline{Y}_1).(Y_2^{(i)} - \overline{Y}_2)^t. \tag{2.50}$$

2.3.3.b 2eme **cas :** $s_2 > s_1$**.**

Estimation par la MMV

Soit $(Y^1, \ldots, Y^n)$ un $n-$échantillon aléatoire qui suit la même loi que Y. Le logarithme de la fonction de vraisemblance est défini par :

$$l(\theta) = -ns_1 \ln \Sigma_{1.2} + n(s_1 - s_2)\ln(\Sigma_{22}) - \frac{n\Sigma_{22}\overline{Y}_1}{\Sigma_{1.2}} + (s_1 - 1)\sum_{i=1}^{n}\ln(Y_1^i) - \ln\Gamma(s_1) - \frac{n\Sigma_{11}\overline{Y}_2}{\Sigma_{1.2}} + (s_2 - 1)\sum_{i=1}^{n}\ln(Y_2^i) - \ln\Gamma(s_2) + \sum_{i=1}^{n}\ln\Big(\Phi_3(s_2 - s_1; s_2; hY_2^i, \delta Y_1^i Y_2^i)\Big) \tag{2.51}$$

avec $\overline{Y}_j = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} Y_j^i,\ j = 1, 2,\ \delta = \dfrac{\Sigma_{12}^2}{\Sigma_{1.2}^2}$ et $h = \delta\dfrac{\Sigma_{1.2}}{\Sigma_{22}}$.

En dérivant la fonction $l(\theta)$ respectivement par rapport à Σ_{11}, Σ_{22}, et Σ_{12}^2, nous obtenons les équations suivantes :

$$-ns_1 + \frac{n\Sigma_{22}\overline{Y}_1}{\Sigma_{1.2}} + \frac{n\Sigma_{11}\overline{Y}_2}{\Sigma_{1.2}} - \frac{n\overline{Y}_2}{\Sigma_{22}} - \frac{(s_2 - s_1)\Sigma_{12}^2}{s_2\Sigma_{22}\Sigma_{1.2}}\Phi_{1,3} - \frac{2\Sigma_{12}^2}{s_2\Sigma_{1.2}^2}\Phi_{2,3} = 0, \tag{2.52}$$

$$-ns_1 + \frac{n(s_1 - s_2)\Sigma_{12}^2}{\Sigma_{11}\Sigma_{22}} + \frac{n\Sigma_{22}\overline{Y}_1}{\Sigma_{1.2}} + \frac{n\Sigma_{11}\overline{Y}_2}{\Sigma_{1.2}} - \frac{n\overline{Y}_1}{\Sigma_{11}} - \frac{(s_2 - s_1)\Sigma_{12}^2(\Sigma_{11}\Sigma_{22} + \Sigma_{1.2})}{s_2\Sigma_{11}\Sigma_{22}^2\Sigma_{1.2}}\Phi_{1,3} - \frac{2\Sigma_{12}^2}{s_2\Sigma_{1.2}^2}\Phi_{2,3} = 0, \tag{2.53}$$

$$ns_1 - \frac{n\Sigma_{22}\overline{Y}_1}{\Sigma_{1.2}} - \frac{n\Sigma_{11}\overline{Y}_2}{\Sigma_{1.2}} + \frac{(s_2 - s_1)\Sigma_{11}}{s_2\Sigma_{1.2}}\Phi_{1,3} + \frac{\Sigma_{11}\Sigma_{22} + \Sigma_{12}^2}{s_2\Sigma_{1.2}^2}\Phi_{2,3} = 0, \tag{2.54}$$

avec $\Phi_{1,3} = \sum_{i=1}^{n} Y_2^i \frac{\Phi_3\Big(s_2 - s_1 + 1; s_2 + 1; hY_2^i, \delta Y_1^i Y_2^i\Big)}{\Phi_3\Big(s_2 - s_1; s_2; hY_2^i, \delta Y_1^i Y_2^i\Big)}$ et

$$\Phi_{2,3} = \sum_{i=1}^{n} Y_1^i Y_2^i \frac{\Phi_3\Big(s_2 - s_1; s_2 + 1; hY_2^i, \delta Y_1^i Y_2^i\Big)}{\Phi_3\Big(s_2 - s_1; s_2; hY_2^i, \delta Y_1^i Y_2^i\Big)}.$$

Théoriquement, nous ne pouvons pas déterminer une expression explicite des estimateurs $\widehat{\Sigma}_{11}$, $\widehat{\Sigma}_{22}$ et $\widehat{\Sigma}_{12}^2$. Pour cette raison nous adoptons la méthode numérique de Newton-Raphson pour estimer ces différents paramètres.

Estimation par la MM

L'estimateur du vecteur paramètre $\theta = (\Sigma_{11}, \Sigma_{22}, \Sigma_{12}^2)$ est donné par

$$\widehat{\Sigma}_{11} = \frac{\overline{Y}_1}{s_1}, \ \widehat{\Sigma}_{22} = \frac{\overline{Y}_2}{s_2} \tag{2.55}$$

$$\widehat{\Sigma}_{12}^2 = \frac{1}{ns_1} \sum_{i=1}^{n} (Y_1^{(i)} - \overline{Y}_1).(Y_2(i) - \overline{Y}_2)^t \tag{2.56}$$

Puisque nous ne pouvons pas assurer théoriquement l'existance de l'estimateur du maximum de vraisemblace du vecteur des paramètres θ de la loi de la diagonale modifiée de Riesz dans le cas $s_1 = s_2$ et $s_2 > s_1$, nous ferons appel à la méthode d'estimation Bayésienne.

2.3.4 Méthode Bayésienne

L'approche bayésienne peut être présentée comme une généralisation de l'approche classique : les paramètres ne sont plus des paramètres inconnus mais des variables aléatoires dont il faut spécifier la distribution. La distribution donnée au moment de la modélisation

est dite a priori. L'intégration de l'information apportée par les données se fait par le calcul de la distribution dite a posteriori et qui n'est autre que la distribution conditionnelle (calculée par la formule de Bayes) aux valeurs prises par les données [18][71][85].

Principe

Considérons un modèle statistique paramétré $F = \{P_\theta(dx) = f(x|\theta)\nu(dx);\ \theta \in \Theta \subset \mathbb{R}^k\}$ dominé par une mesure σ-finie ν. Soit $X = (X_1, \ldots, X_n)$ un $n-$échantillon aléatoire de la loi P_θ. Soit $x = (x_1, \ldots, x_n)$ une réalisation de l'échatillon aléatoire X. La variable aléatoire condionelle $X_i|\theta$ a pour densité conditionnelle $f(x_i|\theta)$. La fonction vraisemblance associée à l'échantillon aléatoire X est définie par $L(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$. La loi conjointe des observartions et du vecteur des paramètres est donnée par $L(x|\theta)\pi(\theta)$, avec $\pi(\theta)$ est la densité a priori de θ. Par application directe du théorème de Bayes, la loi de θ conditionnelle à X est appelée distribution a posteriori du vecteur des paramètres θ est définie par

$$\pi(\theta|x) = \frac{L(x|\theta)\pi(\theta)}{q(x)},$$

avec $q(x) = \int_\Theta \pi(\theta)L(x|\theta)d\theta$.

Cette distribution a postériori est donc proportionnelle au produit de la vraisemblance par la densité a priori (*i-e*, $\pi(\theta|x) \propto L(x|\theta)\pi(\theta)$).

À partir de cette distribution a posteriori, on peut déduire l'estimateur Bayesien $\widehat{\theta}$ à partir d'un critère d'estimation basé sur l'espérance a posteriori (EAP) donné par $\widehat{\theta} = \mathbb{E}(\theta|X_1, \ldots, X_n)$.

Dans de nombreux cas de lois, le calcul explicite de $\pi(\theta|x)$ est impossible. De même si $\pi(\theta|x)$ est connue, le calcul de $\mathbb{E}(\theta|X_1, \ldots, X_n)$ peut s'avérer très difficile. Dans la littérature,

plusieurs méthodes d'approximation ont été proposées [18].

Dans notre travail, nous nous sommes intéressés plus particulièrement à la méthode de Monte-Carlo [18]. Cette méthode est itérative et consiste à déterminer une approximation du vecteur paramètres θ donné par

$$\begin{aligned}\widehat{\theta} = \mathbb{E}(\theta|X_1,\ldots,X_n) &= \int_\Theta \theta\pi(\theta|x_1,\ldots,x_n)d\theta \\ &= \frac{\displaystyle\int_\Theta \theta L(X_1,\ldots,X_n|\theta)\pi(\theta)d\theta}{\displaystyle\int_\Theta L(X_1,\ldots,X_n|\theta)\pi(\theta)d\theta}.\end{aligned}$$

En appliquant cette méthode, nous obtenons

$$\widehat{\theta} \simeq \frac{\displaystyle\sum_{j=1}^{N}\theta_j\prod_{i=1}^{n}f(x_i|\theta_j)}{\displaystyle\sum_{j=1}^{N}\prod_{i=1}^{n}f(x_i|\theta_j)}$$

où $\theta_1,\ldots,\theta_N \sim \pi(\theta)$.

D'après la loi forte des grands nombres [26], lorsque N tend vers l'infini, $\widehat{\theta}$ converge presque sûrement vers θ.

2.3.5 Estimateur EAP pour le modèle proposé.

Dans cette section, nous allons montrer l'intérêt de la méthode Bayesienne pour l'estimation du vecteur des paramètres $(\Sigma_{11},\Sigma_{22},\Sigma_{12}^2)$ associé à la loi de la diagonale modifiée de Riesz proposée dans la section 2.3.2.

Estimateur EAP dans le cas $s_1 = s_2 = s$.

Soit $(Y^1,\ldots,Y^n)$ un $n-$échantillon aléatoire qui suit la loi Gamma bivariée de Kibble du vecteur des paramètres $\theta = (\Sigma_{11},\Sigma_{22},\Sigma_{12}^2) \in \Theta = \mathbb{R}_+^* \times \mathbb{R}_+^* \times]0,\Sigma_{11}\Sigma_{22}[$, de densité de

probabilité définie par

$$f(y_1^i, y_2^i|\theta) \;=\; \exp(-\frac{\Sigma_{22}y_1^i + \Sigma_{11}y_2^i}{\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2}) \;\times\; \frac{(y_1^i y_2^i)^{s-1}}{(\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2)^s \Gamma(s)} f_s(\delta y_1^i y_2^i)\mathbf{1}_{[0,+\infty[^2}(y_1^i, y_2^i), \tag{2.57}$$

avec $\overline{y}_j = \frac{1}{n}\sum_{i=1}^n y_j^i,\ j = 1, 2,,\ \delta = \dfrac{\Sigma_{12}^2}{(\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2)^2}$ et la fonction f_s est définie par

$$f_s(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(s+k)k!},\ \forall\ z\ \in\ \mathbb{R}.$$

Soit $(y^1, \ldots, y^n)$ une réalisation de l'échantillon aléatoire $(Y^1, \ldots, Y^n)$, alors la fonction de vraisemblance associée à l'échantillon aléatoire est définie par

$$\begin{aligned} L(y^1, \ldots, y^n|\theta) &= \prod_{i=1}^n f((y_1^i, y_2^i)|\theta) \\ &= \exp(-\frac{n\Sigma_{22}\widehat{y}_1 + n\Sigma_{11}\widehat{y}_2}{\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2}) \times \prod_{i=1}^n \frac{(y_1^i y_2^i)^{s-1}}{(\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2)^{ns}\Gamma(s)^n} \prod_{i=1}^n f_s(\delta y_1^i y_2^i). \end{aligned}$$

À partir du domaine de définition du vecteur des paramètres $\theta = (\Sigma_{11}, \Sigma_{22}, \Sigma_{12}^2)$, nous avons supposé que les lois a priori pour les différents paramètres sont :

$$\Sigma_{11} \sim \gamma(a_1, b_1),\ \text{avec } a_1 > 0,\ b_1 > 0,$$

$$\Sigma_{22} \sim \gamma(a_2, b_2),\ \text{avec } a_2 > 0,\ b_2 > 0,$$

$$\Sigma_{12}^2|\Sigma_{11}, \Sigma_{22} \sim\ U(]0, \Sigma_{11}\Sigma_{22}[),$$

avec $U(]0, \Sigma_{11}\Sigma_{22}[)$ représente la loi uniforme sur l'intervalle $]0, \Sigma_{11}\Sigma_{22}[$. Ainsi, la densité a postériori est donnée par

$$f(\theta|(y_1^i, y_2^i)) = \frac{\prod_{i=1}^{n} f((y_1^i, y_2^i)|\theta) f(\Sigma_{11}) f(\Sigma_{22}) f(\Sigma_{12}^2|\Sigma_{11}, \Sigma_{22})}{\int_\Theta \prod_{i=1}^{n} f((y_1^i, y_2^i)|\theta) f(\Sigma_{11}) f(\Sigma_{22}) f(\Sigma_{12}^2|\Sigma_{11}, \Sigma_{22}) d\theta}$$

$$= \exp(-\frac{n\Sigma_{22}\widehat{y}_1 + n\Sigma_{11}\widehat{y}_2}{\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2} - \frac{a_1\Sigma_{11}}{b_1} - \frac{a_2\Sigma_{22}}{b_2}) \times \prod_{i=1}^{n} \frac{(y_1^i y_2^i)^{s-1} f_s(\delta y_1^i y_2^i)}{(\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2)^{ns}\Gamma(s)^n}$$

$$\times \frac{\Sigma_{11}^{a_1-1}\Sigma_{22}^{a_2-1}}{b_1^{a_1} b_2^{a_2}\Gamma(a_1)\Gamma(a_2) \int_\Theta \prod_{i=1}^{n} f((y_1^i, y_2^i)|\theta) f(\Sigma_{11}) f(\Sigma_{22}) f(\Sigma_{12}^2|\Sigma_{11}, \Sigma_{22}) d\theta}. \quad (2.58)$$

En appliquant la méthode de Monte-Carlo, l'estimateur EAP $\widehat{\theta} = (\widehat{\Sigma}_{11}, \widehat{\Sigma}_{22}, \widehat{\Sigma}_{12}^2)$ de θ à l'itération $(p+1)$ est défini par

$$\widehat{\Sigma}_{11}^{(p+1)} = \mathbb{E}(\Sigma_{11}|(y^1, \ldots, y^n), \Sigma_{22}^{(p)}, \Sigma_{12}^{2\ (p)}),$$

$$\widehat{\Sigma}_{22}^{(p+1)} = \mathbb{E}(\Sigma_{22}|(y^1, \ldots, y^n), \Sigma_{11}^{(p+1)}, \Sigma_{12}^{2\ (p)}),$$

$$\widehat{\Sigma}_{12}^{2\ (p+1)} = \mathbb{E}(\Sigma_{12}^2|(y^1, \ldots, y^n), \Sigma_{11}^{(p+1)}, \Sigma_{22}^{(p+1)}).$$

Estimateur EAP dans le cas $s_2 > s_1$.

Nous supposons que les lois a priori pour les paramètres Σ_{11} et Σ_{22} sont deux lois Gamma de paramètres respectifs (c_1, d_1) et (c_2, d_2) et $\Sigma_{12}^2|(\Sigma_{11}, \Sigma_{22})$ suit la loi uniforme $U(]0, \Sigma_{11}\Sigma_{22}[)$.

La densité a posteriori du vecteur des paramètres $\theta = (\Sigma_{11}, \Sigma_{22}, \Sigma_{12}^2)$ est définie par

$$f(\theta|(y_1^i, y_2^i)) = \frac{\prod_{i=1}^{n} f((y_1^i, y_2^i)|\theta) f(\Sigma_{11}) f(\Sigma_{22}) f(\Sigma_{12}^2|\Sigma_{11}, \Sigma_{22})}{\int_\Theta \prod_{i=1}^{n} f((y_1^i, y_2^i)|\theta) f(\Sigma_{11}) f(\Sigma_{22}) f(\Sigma_{12}^2|\Sigma_{11}, \Sigma_{22}) d\theta}$$

avec $f((y_1, y_2)|\theta)$ représente la densité de probabilité de la diagonale modifiée de Riesz définie par l'équation (2.35).

En appliquant la méthode de Monte-Carlo, l'estimateur EAP $\widehat{\theta} = (\widehat{\Sigma}_{11}, \widehat{\Sigma}_{22}, \widehat{\Sigma}^2_{12})$ de θ à l'itération $(p+1)$ est défini par

$$\widehat{\Sigma}_{11}^{(p+1)} = \mathbb{E}(\Sigma_{11}|(y^1, \ldots, y^n), \Sigma_{22}^{(p)}, \Sigma_{12}^{2\ (p)}),$$

$$\widehat{\Sigma}_{22}^{(p+1)} = \mathbb{E}(\Sigma_{22}|(y^1, \ldots, y^n), \Sigma_{11}^{(p+1)}, \Sigma_{12}^{2\ (p)}),$$

$$\widehat{\Sigma}_{12}^{2\ (p+1)} = \mathbb{E}(\Sigma_{12}^2|(y^1, \ldots, y^n), \Sigma_{11}^{(p+1)}, \Sigma_{22}^{(p+1)}).$$

2.3.6 Performance des estimateurs

Dans ce paragraphe, nous étudions la performance des estimateurs des paramètres du modèle proposé obtenu par la MM, MMV et la méthode Bayésienne en utilisant comme critère de comparaison l'erreur quadratique moyenne.

2.3.4.a 1er cas : $s_1 = s_2 = s$.

Nous avons généré des échantillons de tailles $n = 1000$ pour :

i) des variables aléatoires qui suivent la même loi que $Y = (Y_1, Y_2)$ (loi Gamma bivariée de Kibble) pour différentes valeurs de ses paramètres Σ_{11}, Σ_{22} et Σ^2_{12}.

ii) des lois Gamma univariées $\gamma(a_i, b_i)$ (lois a priori) telles que $a_i b_i = \dfrac{\overline{Y}_i}{s}$ et $a_i b_i^2 = \dfrac{S_i^2}{s}$, $i = 1, 2$, avec S_i représente la variance empirique de la variable aléatoire Y_i.

iii) des lois uniformes sur l'intervalle $]0, \Sigma_{11}\Sigma_{22}[$ (lois a priori) telles que $\Sigma_{11}\Sigma_{22} = \dfrac{\overline{Y}_1\overline{Y}_2}{s^2}$.

Cette étude de performance est consacrée à calculer l'erreur quadratique moyenne (EQM) pour chaque estimateur. La valeur de l'EQM pour chaque estimateur est obtenue après $N = 1000$ réalisations. Les tableaux TAB.2.1, TAB.2.2 et TAB.2.3 résument les valeurs de l'EQM obtenues par les différentes méthodes d'estimations ainsi que les vraies valeurs

de ses différents paramètres.

Dans TAB.2.1, nous présentons l'EQM de l'estimateur du paramètre Σ_{11} pour différentes valeurs des autres paramètres. Nous remarquons d'après les résultats présentés dans TAB.2.1 que la valeur de l'EQM obtenue par l'estimateur EAP de Σ_{11} est la plus petite par rapport aux autres estimateurs.

Dans TAB.2.2, nous présentons l'EQM de l'estimateur du paramètre Σ_{22} pour différentes valeurs des autres paramètres. Nous remarquons d'après les résultats présentés dans le TAB.2.2 que la valeur de l'EQM obtenue par l'estimateur EAP de Σ_{22} est la plus petite par rapport aux autres estimateurs.

Dans TAB.2.3, nous présentons l'EQM de l'estimateur du paramètre Σ_{12}^2 pour différentes valeurs des autres paramètres. Nous remarquons d'après les résultats présentés dans TAB.2.3 que la valeur de l'EQM obtenue par l'estimateur EAP de Σ_{12}^2 est la plus faible par rapport aux autres estimateurs.

$(\Sigma_{22}, \Sigma_{12}^2)$	(a_1, b_1)	(a_2, b_2)	EQM-MM	EQM-MMV	EQM-EAP
$(2, 0.3)$	$(1, 2.0297)$	$(1, 2.0134)$	0.064	0.014	0.0094
$(1, 1)$	$(1, 1.9552)$	$(1, 1.0106)$	0.114	0.05	0.00173
$(10, 15)$	$(1, 1.9855)$	$(1, 9.7837)$	0.09	0.0245	0.00987
$(30, 40)$	$(1, 2.0285)$	$(1, 29.8748)$	5×10^{-4}	4.7×10^{-4}	3.23×10^{-4}
$(0.5, 0.9)$	$(1, 1.997)$	$(1, 0.5071)$	0.028	0.0101	0.00848

TAB. 2.1 – L'EQM du paramètre Σ_{11}.

$(\Sigma_{11}, \Sigma_{12}^2)$	(a_1, b_1)	(a_2, b_2)	EQM-MM	EQM-MMV	EQM-EAP
$(5, 3)$	$(1, 5.1327)$	$(1, 24.9516)$	0.0125	0.09824	0.0545
$(0.5, 1)$	$(0.5047, 1.05)$	$(1, 4.9644)$	0.084	0.044	0.01
$(12, 20)$	$(1, 11.3527)$	$(1, 4.7275)$	0.052	0.0172	0.0084
$(1, 2)$	$(1, 1.0116)$	$(1, 4.8316)$	0.01	0.00789	0.0041712
$(20, 50)$	$(1, 20.4757)$	$(1, 4.9529)$	0.08	0.068	0.024

TAB. 2.2 – L'EQM du paramètre Σ_{22}^2.

$(\Sigma_{11}, \Sigma_{22})$	(a_1, b_1)	(a_2, b_2)	EQM-MM	EQM-MMV	EQM-EAP
$(0.5, 4)$	$(1, 0.5217)$	$(1, 4.2354)$	0.02334	0.009867	0.007821
$(2, 1)$	$(1, 2.0989)$	$(1, 1.051)$	0.003359	0.001285	0.000852
$(5, 3)$	$(1, 5.1491)$	$(1, 2.9993)$	0.05641	0.041201	0.0178
$(15, 5)$	$(1, 45.237)$	$(1, 4.9409)$	0.0544	0.011	0.0099823
$(4, 10)$	$(1, 3.9553)$	$(1, 10.3718)$	0.0961	0.072	0.0468

TAB. 2.3 – L'EQM du paramètre Σ_{12}^2.

2.3.4.b 2eme cas : $s_2 > s_1$.

Nous avons généré des échantillons de tailles $n = 1000$ pour les différentes lois suivantes : i) des variables aléatoires qui suivent la même loi que Y (loi de la diagonale modifiée de Riesz) pour différentes valeurs de ses paramètres Σ_{11}, Σ_{22} et Σ_{12}^2. Pour simuler la variable aléatoire Y, nous avons simulé respectivement deux variables aléatoires indépendantes Z et T. Le vecteur aléatoire $Z = (Z_1, Z_2)$ est simulé suivant une loi Gamma bivariée de Kibble de paramètres $s_1, \Sigma_{11}, \Sigma_{22}$ et Σ_{12}^2 et de densité de probabilité définie par l'équation (2.30). La variable aléatoire T est simulée suivant une loi Gamma univariée de paramètres

$s_2 - s_1$ et Σ_{22}.

ii) des lois Gamma univariées $\gamma(a_i, b_i)$ (lois a priori) telles que $a_i b_i = \dfrac{\overline{Y}_i}{s_i}$ et $a_i b_i^2 = \dfrac{S_i^2}{s_1}$, $i = 1, 2$, avec S_i représente la variance empirique du vecteur aléatoire Y_i.

iii) des lois uniformes sur l'intervalle $]0, \Sigma_{11}\Sigma_{22}[$ (loi a priori) telles que $\Sigma_{11}\Sigma_{22} = \dfrac{\overline{Y}_1\overline{Y}_2}{s_1 s_2}$.

Cette étude de performance est consacrée à calculer l'erreur quadratique moyenne (EQM) pour chaque estimateur. La valeur de l'EQM pour chaque estimateur est obtenue après $N = 1000$ réalisations. Les tableaux TAB.2.4, TAB.2.5 et TAB.2.6 résument les valeurs de l'EQM obtenues par les différentes méthodes d'estimation ainsi que les vraies valeurs de ses différents paramètres.

Dans TAB.2.4, nous présentons l'EQM de l'estimateur du paramètre Σ_{11} pour différentes valeurs des autres paramètres. Nous remarquons d'après les résultats présentés dans TAB.2.4 que la valeur de l'EQM obtenue par l'estimateur EAP de Σ_{11} est la plus petite par rapport aux autres estimateurs.

Dans TAB.2.5, nous présentons l'EQM de l'estimateur du paramètre Σ_{22} pour différentes valeurs des autres paramètres. Nous remarquons d'après les résultats présentés dans TAB.2.5 que la valeur de l'EQM obtenue par l'estimateur EAP de Σ_{22} est la plus petite par rapport aux autres estimateurs.

Dans TAB.2.6, nous présentons l'EQM de l'estimateur du paramètre Σ_{12}^2 pour différentes valeurs des autres paramètres. Nous remarquons d'après les résultats présentés dans TAB.2.6 que la valeur de l'EQM obtenue par l'estimateur EAP de Σ_{12}^2 est la plus faible par rapport aux autres estimateurs.

$(\Sigma_{22}, \Sigma_{12}^2)$	(a_1, b_1)	(a_2, b_2)	EQM-MM	EQM-MMV	EQM-EAP
$(2, 0.3)$	$(1, 7.6042)$	$(1, 0.9687)$	0.0451	0.02156	0.01047
$(1, 1)$	$(1, 3.4531)$	$(1, 0.4969)$	0.31857	0.19673	0.1002
$(10, 15)$	$(1, 24.3585)$	$(1, 5.1211)$	0.5469	0.314	0.1386
$(30, 40)$	$(1, 6.0426)$	$(1, 15.1842)$	0.1115	0.008	0.005222
$(0.5, 0.9)$	$(1, 3.7883)$	$(1, 0.2597)$	0.728	0.499	0.09157

TAB. 2.4 – L'EQM du paramètre Σ_{11}.

$(\Sigma_{11}, \Sigma_{12}^2)$	(a_1, b_1)	(a_2, b_2)	EQM-MM	EQM-MMV	EQM-EAP
$(5, 3)$	$(1, 9.2355)$	$(1, 2.0114)$	0.05	0.0354	0.001457
$(0.5, 1)$	$(1, 3.405)$	$(1, 2.0099)$	0.6824	0.49327	0.009897
$(12, 20)$	$(1, 14.4213)$	$(1, 2.013)$	0.451	0.0201	0.00889
$(1, 2)$	$(1, 2.75)$	$(1, 2.0338)$	0.083	0.05946	0.01766
$(20, 50)$	$(1, 22.1492)$	$(1, 2.0285)$	0.16678	0.121	0.00698

TAB. 2.5 – L'EQM du paramètre Σ_{22}^2.

$(\Sigma_{11}, \Sigma_{22})$	(a_1, b_1)	(a_2, b_2)	EQM-MM	EQM-MMV	EQM-EAP
$(0.5, 4)$	$(1, 8.776)$	$(1, 1.9301)$	0.429	0.1998	0.10758
$(2, 1)$	$(1, 2.2397)$	$(1, 0.4979)$	0.007359	0.0047453	0.002056
$(5, 3)$	$(1, 7.8841)$	$(1, 1.5103)$	0.71121	0.34824	0.1056
$(15, 5)$	$(1, 23.237)$	$(1, 2.4686)$	0.8214	0.60217	0.21896
$(4, 10)$	$(1, 27.6910)$	$(1, 5.1292)$	0.20561	0.09542	0.05481

TAB. 2.6 – L'EQM du paramètre Σ_{12}^2.

À partir de ces résultats de simulation, nous pouvons conclure que la méthode Bayésienne est plus performante pour l'estimation des paramètres de notre modèle proposé.

2.4 Conclusion

Ce chapitre a traité le problème de l'estimation des lois multivariées. Dans un premier temps, nous avons rappelé quelques lois multidimensionnelles et quelques méthodes d'estimation telles que la méthode du maximum de vraisemblance, la méthode des moments et la méthode Bayésienne.

Dans un deuxième temps, nous avons proposé un nouveau modèle statistique de lois multidimensionnelles que nous avons appelé modèle de lois de la diagonale modifiée de Riesz qui est basé sur la loi de Riesz multidimensionnelle. La loi de ce nouveau modèle est le produit de convolution d'une loi Gamma univariée et d'une loi Gamma bivariée de Kibble. Dans un troisième temps, nous avons estimé les paramètres de ce nouveau modèle par la méthode Bayésienne, la méthode du maximum de vraisemblance et la méthode des moments. Une étude comparative de la performance des estimateurs a été réalisée en cal-

culant l'erreur quadratique moyenne entre l'estimateur et la valeur réelle de paramètre paramètre.

Dans la suite de notre travail, nous allons traiter le problème de la généralisation de ce nouveau modèle pour le cas du mélange des lois et son intérêt en segmentation d'images couleur. Ceci fera l'objet du chapitre 3.

Chapitre 3

Application des familles exponentielles en segmentation d'images couleurs

3.1 Introduction

La segmentation est une des étapes critiques de l'analyse des images qui conditionne la qualité des traitements des images effectués ultérieurement. Elle permet d'isoler dans l'image les objets sur lesquels doit porter l'analyse. En effet, à partir des résultats de la segmentation, il est possible d'identifier les objets d'une scène. Par rapport aux images à niveaux de gris, la segmentation d'images couleur est un domaine de recherche assez récent. En effet, pendant plusieurs années, beaucoup d'attentions ont été focalisées sur la segmentation d'images à niveaux de gris [107]. Cependant, des études montrent que du fait de l'adaptation à la lumière, l'oeil humain peut seulement discerner quelques douzaines de niveaux d'intensités dans une image complexe, mais reconnaît des milliers de variations chromatiques. Ainsi, l'utilisation des attributs couleurs pour la segmentation

d'images devrait fournir de meilleurs résultats que l'attribut d'intensité, dès lors que ces attributs fournissent une caractéristique naturelle sur les scènes observées.

Dans ce chapitre, nous nous sommes intéressé plus particulièrement à la segmentation statistique multivariée des images couleurs. Les approches fondées sur des modélisations statistiques donnent de bons résultats dans ce domaine [13][29]. Les modélisations statistiques offrent un cadre méthodologique souple et riche, permettant d'aborder à l'aide d'outils variés de nombreux problèmes d'imagerie couleur. La segmentation consiste à subdiviser une image couleur en différentes régions significatives à l'aide d'une règle de décision. Nous nous plaçons dans le cas de la segmentation non supervisée où nous n'avons aucune connaissance a priori sur les régions qui représentent les différentes classes de l'image [20][37][111]. Les algorithmes les plus utilisés dans le cadre de la segmentation non supervisée sont issus de la famille d'algorithmes basés sur l'algorithme EM (Expectation-Maximisation) [33][70]. En effet, la loi de probabilité de l'image couleur à segmenter est souvent supposée un mélange de lois connues (Gaussienne [5], lois d'un système de Pearson [73]). Cependant, lorsque les distributions du mélange ne peuvent pas être approchées par des modèles connus, des approches non paramétriques peuvent résoudre ce problème [42][111]. Notre étude réside dans l'intégration de l'algorithme K-means et des modèles qui appartiennent aux familles exponentielles générales multivariées dans l'approche de segmentation à fin de régulariser et d'assurer une meilleure identification du mélange. Nous montrons dans ce chapitre la possibilité de mettre en place un nouvel algorithme paramétrique dans le but d'estimer les paramètres du mélange afin d'effectuer une segmentation non supervisée des images couleur. D'abord, nous rappelons la notion d'espace de couleurs nécessaire à la visualisation colorée d'images. Puis, nous présentons quelques différentes

approches de segmentation. Ensuite, nous définissons la modélisation probabiliste en classification. En outre, nous introduisons la généralisation des lois de la diagonale modifiée de Riesz dans le cas du modèle de mélange. Enfin, nous présentons quelques résultats de segmentation statistique multivariée non supervisée d'images couleur obtenues par notre introduction.

3.2 Les représentations classiques de la couleur/ les espaces de couleurs

La couleur des pixels d'une image acquise par une caméra couleur est codée dans l'espace de représentation RVB (Rouge, Vert, Bleu) et ainsi associée à trois composantes couleur. Un pixel dans une image donne ainsi naissance à un point dans un espace tridimensionnel dont les coordonnées sont les niveaux de ces trois composantes. Par extension, l'ensemble des pixels de l'image donne naissance à des nuages de points permettant de définir des classes de pixels. La couleur est une notion très complexe car elle fait intervenir des phénomènes physiques (propriétés de la lumière et des matériaux), physiologiques (structure et fonctionnement de l'oeil humain) et psychologiques (interprétation par le cerveau). Différentes théories ont donc été proposées et plusieurs études ont été et sont encore menées pour tenter de modéliser cette information aussi riche que complexe. Elles ont donné naissance à de nombreux modèles [100] (RVB, TSL, YIQ, YUV, AC1C2, XYZ, CIELAB, CIELUV, I1I2I3,...) basés sur des propriétés spécifiques. Il existe donc de nombreux espaces de représentation de la couleur qui ont été construits en se basant sur ces propriétés physiques, physiologiques ou psychologiques de la perception de la couleur. L'analyse des images

couleur n'est donc pas restreinte à l'espace d'acquisition RVB. Il est également reconnu que les résultats d'un traitement dépendent du choix de l'espace couleur. De nombreux auteurs tentent ainsi de déterminer l'espace couleur le mieux adapté à leur problématique et leurs travaux ont montré qu'il n'existe pas d'espaces couleur adaptés à toutes les problématiques d'analyse d'images. Dans cette section, nous présentons les deux espaces RVB et TSL (Teinte, Saturation, Luminosité).

3.2.1 L'espace RVB

L'espace de couleurs RVB demeure le plus répandu. En effet, il est implémenté dans la plupart des outils matériels de visualisation (écran, vidéoprojection...). Dans cet espace, un pixel est codé par trois composantes Rouge, Vert et Bleu, à valeurs à l'intérieur d'un cube unité (FIG.3.1.a). Cet espace a été développé en fonction des connaissances liées à la vision humaine, les cônes étant plus sensibles à ces trois couleurs. Ce modèle est additif, ce qui signifie que toutes les couleurs sont déduites à partir du noir (R=V=B= 0) en ajoutant plus ou moins certaines composantes (FIG.3.1.b). Dans cet espace, chaque composante est donc définie par une valeur entre 0 et 1 (ou entre 0 et 255 selon le standard). De ce fait elles doivent être normalisées de la même façon, ce qui est quelquefois contraignant. Le principal inconvénient de ce modèle réside dans la manipulation même des couleurs. En effet, si on veut augmenter la luminosité d'une couleur, il faut incrémenter proportionnellement chaque composante étant donnée la corrélation entre les plans R, V et B. Ces contraintes font du modèle additif des couleurs (RVB), un espace de couleurs peu approprié à la représentation des images multicomposantes sous la forme d'une image de couleurs. Il s'avère donc préférable pour nos opérations sur les observations astronomiques

multibandes, d'utiliser le modèle psychovisuel TSL basé sur la perception de la couleur par l'oeil.

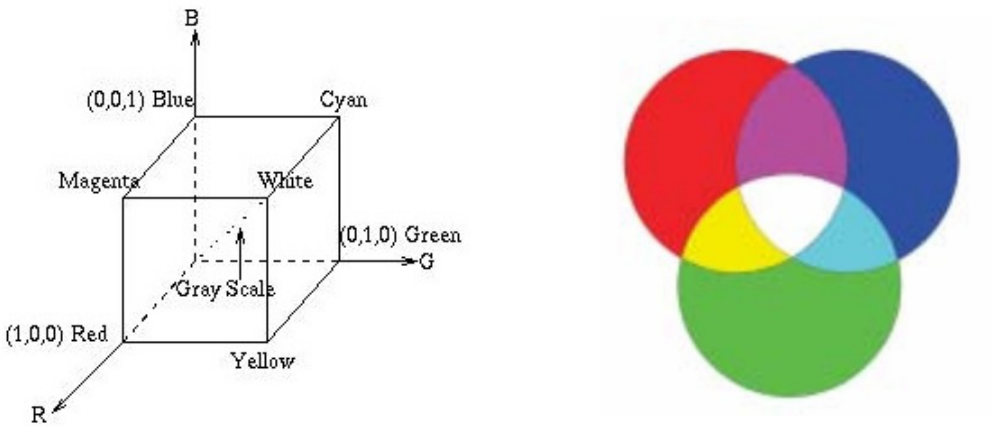


FIG. 3.1 – (a).Cube RVB, (b).Composition additive des couleurs.

Chaque point de l'image est repérable dans un espace tridimensionnel par ses trois composantes chromatiques R, V, B, dont les expressions sont les suivantes :

$$R = \int_{-\infty}^{+\infty} S(\lambda)T_r(\lambda)s(\lambda)d\lambda$$

$$V = \int_{-\infty}^{+\infty} S(\lambda)T_v(\lambda)s(\lambda)d\lambda$$

$$B = \int_{-\infty}^{+\infty} S(\lambda)T_b(\lambda)s(\lambda)d\lambda$$

où $S(\lambda)$ est la distribution spectrale de la source lumineuse secondaire, $T_r(\lambda)$, $T_v(\lambda)$ et $T_b(\lambda)$ sont les transmittances des filtres rouge, vert et bleu, et $s(\lambda)$ est la sensibilité du capteur.

3.2.2 L'espace TSL

Le sigle « TSL » parfois écrit « TLS » fait référence aux termes Teinte, Saturation, Luminosité. L'abréviation anglaises pour Hue Saturation and Intensity HSI est parfois utilisée. Le modèle TSL est un modèle de représentation dit "naturel", c'est-à-dire proche de la perception physiologique de la couleur par l'oeil humain. En effet, le modèle RVB aussi adapté soit-il pour la représentation informatique de la couleur ou bien l'affichage sur les périphériques de sortie, ne permet pas de sélectionner facilement une couleur. Assurément, le réglage de la couleur en RVB dans les outils informatiques se fait en générale à l'aide de trois glisseurs ou bien de trois cases avec les valeurs relatives de chacune des composantes primaires, or l'éclaircissement d'une couleur demande d'augmenter proportionnellement les valeurs respectives de chacune des composantes. Ainsi, le modèle TSL a-t-il été mis au point afin de pallier cette lacune du modèle RVB. Le modèle TSL consiste à décomposer la couleur selon des critères physiologiques :

• **la teinte** (en anglais Hue), correspondant à la perception de la couleur (T-shirt mauve ou orange),

• **la saturation**, décrivant la pureté de la couleur, c'est-à-dire son caractère vif ou terne (T-shirt neuf ou délavé),

• **la luminance**, indiquant la quantité de lumière de la couleur, c'est-à-dire son aspect clair ou sombre (T-shirt au soleil ou à l'ombre).

Voici une représentation graphique du modèle TSL, dans lequelle la teinte est représentée par un cercle chromatique et la luminance et la saturation par deux axes :

Le modèle représentation graphique du modèle TSL a été mis au point dans le but de

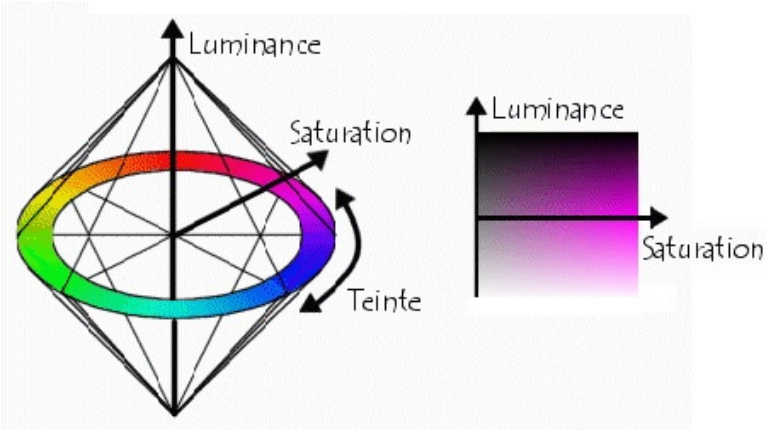


FIG. 3.2 – Représentation graphique du modèle TSL.

permettre un choix interactif rapide d'une couleur, pour autant il n'est pas adapté à une description quantitative d'une couleur. Il existe d'autres modèles naturels de représentation proches du modèle TSL :

• **HSB** : Hue, Saturation, Brightness soit Teinte, Saturation, Brillance en français. La **brillance** décrit la perception de la lumière émise par une surface.

• **HSV** : Hue, Saturation, Value soit Teinte, Saturation, Valeur en français.

• **HCI** : Hue, Chrominance, Intensity soit Teinte, Chrominance, Intensité en français.

L'avantage essentiel du modèle TSL est de séparer distinctement l'information de luminosité, placée sur un axe vertical, de la teinte et de la saturation, situées sur un même plan. L'intérêt de ce modèle pour le traitement d'images et les compositions colorées est qu'il permet d'améliorer la qualité visuelle d'une image par la modification de sa luminosité et de sa saturation, sans altérer ses teintes. La Figure FIG.3.3 représente, de haut en bas et de gauche à droite, la composition colorée RVB et la transformation du RVB en Teinte/Saturation/Luminosité. L'image en bas a été obtenue par une transformation

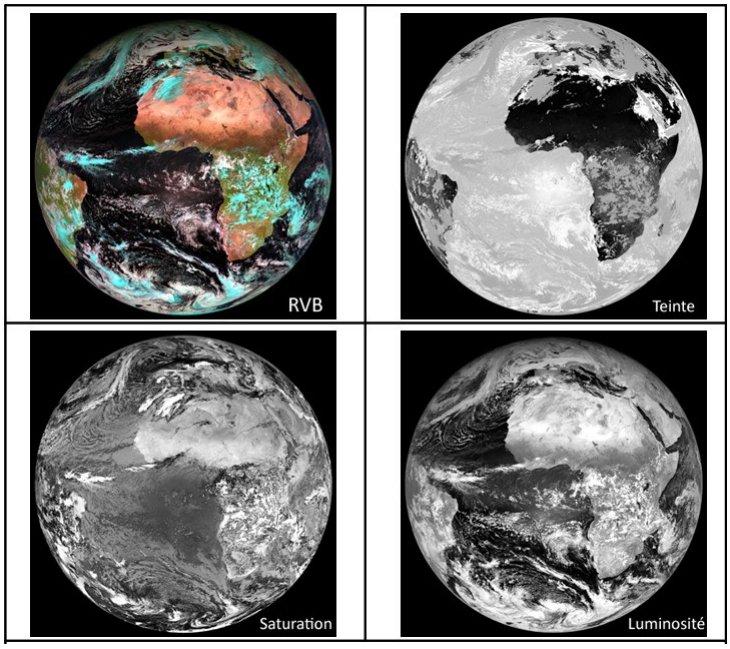
inverse (TSL vers RVB), après la modification du contraste de la composante 'saturation'.

D'après [72], les composantes du modèle TSL (ou HSI) sont définies par :

$$\text{Teinte} = T = \cos^{-1}\left(\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right); \quad \text{si } B < G, \text{ sinon } T = 2\pi - T$$

$$\text{Luminance} = L = \frac{R+G+B}{3},$$

$$\text{Saturation } = S = 1 - \frac{\min(R,G,B)}{L}.$$



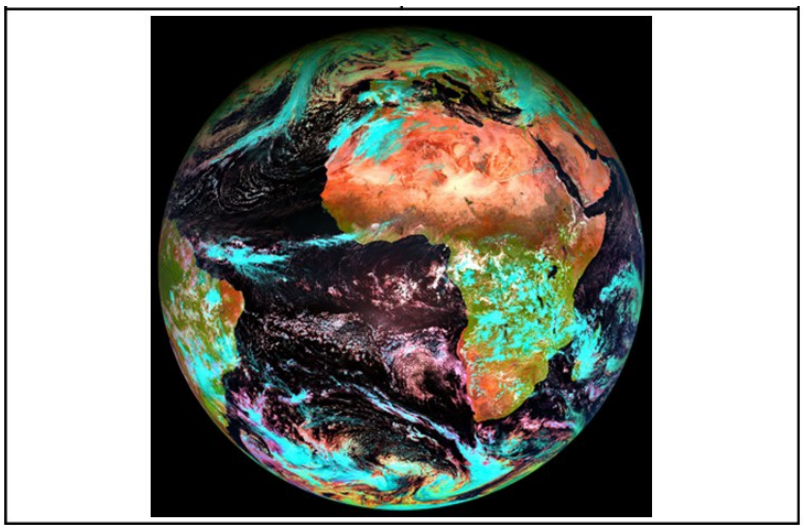


FIG. 3.3 – Qualité visuelle d'une image.

3.3 Les différentes approches de la segmentation d'images couleur

La segmentation d'une image couleur consiste à partitionner une image en régions disjointes de pixels, c'est-à-dire en ensembles de pixels à la fois connexes et proches au sens colorimétrique. Il s'agit d'une étape indispensable à la plupart des algorithmes d'analyse d'images. La littérature recense un nombre incommensurable de techniques de segmentation, qui se différencient par les hypothèses sur lesquelles elles reposent. Il n'existe pas d'algorithmes de segmentation universels, le choix d'une approche est guidé notamment par le type d'application, les contraintes de mise en oeuvre et le type d'images à traiter. Il existe de très nombreuses méthodes de segmentation qui varient fortement selon leurs façons d'aborder le problème de la segmentation d'images [3][15][67][80]. Dans la suite de cette section, nous rappelons les approches contours et les approches régions.

3.3.1 La segmentation par approche contours

La segmentation par approche contours s'intéresse aux contours de l'objet dans l'image [50][89]. La plupart des algorithmes qui lui sont associés sont locaux, c'est-à-dire qu'ils fonctionnent au niveau du pixel. La détection de contour n'est en général qu'une étape préliminaire dans la reconnaissance d'objets car elle fournit plus d'informations sur les régions que les régions elles-même et doit donc être complétée par un algorithme de segmentation région. Cependant nous pouvons dire que les contours constituent des indices riches, au même titre que les points d'intérêts, pour toute interprétation ultérieure de l'image et méritent donc d'être traités à part. Les approches contours se caractérisent par le fait qu'elles ne prennent en compte que des informations sur le contour des objets. Les premiers modèles de segmentation s'appuient sur des détections de changement rapide d'intensité lumineuse ou de couleur, c'est-à-dire les contours de l'image. L'application des détecteurs de contours sous la forme de filtres dérivateurs permet d'obtenir les contours des objets présents dans la scène. Nous pouvons citer les approches se basant sur les différences finies comme l'opérateur de gradient, l'opérateur Laplacien [74], les filtres de Sobel [97], Prewitt [83], Roberts [86] ou bien des approches reposant sur des critères d'optimalité comme les filtres de Canny-Deriche [22]. Mais ce genre de techniques est peu exploitable car elles donnent des contours non fermés, bruités ou des contours non détectés (voir FIG.3.4).

FIG. 3.4 – Détection de contours sur l'image Léna.

La segmentation par approche contour présente certaines limites. Les principales limites des méthodes de détection de contour sont les suivantes [2] :

• Les contours extraits selon les méthodes classiques souvent ne correspondent pas nécessairement à la limite des objets. Dans de nombreuses images de basse qualité, certaines méthodes produisent des faux contours.

• Les techniques de détection de contour dépendent de l'information contenue dans le voisinage local de l'image. Il n'y a pas d'information globale.

• Dans la plupart des cas, les stratégies de détection des contours ignorent l'organisation d'ordre supérieur qui peut être utilement présent dans l'image.

• Après l'extraction des points du contour, ces derniers sont reliés afin de déterminer les frontières. Le processus de fermeture des contours peut parfois conduire à des discontinuités et des lacunes dans l'image.

• Il est souvent difficile d'identifier et de classer les contours parasites. Par conséquent, de nombreux auteurs préfèrent l'approche régions pour la segmentation des images couleur.

3.3.2 La segmentation par approche régions

Les approches régions sont des techniques de segmentation qui consistent à identifier et à localiser les ensembles connexes de pixels. Autrement dit, contrairement aux approches frontières (contours) qui recherchent des dissimilarités, les approches régions recherchent plutôt la similarité. Ces approches produisent des régions fermées R_i homogènes et cohérentes d'une image I, ces régions sont dans la mesure du possible liées à un objet ou une partie d'un objet représenté dans l'image I [112]. Les pixels de chaque région doivent respecter des critères d'homogénéité et de connexité. L'homogénéité d'une région R_i est

définie par un prédicat d'uniformité, noté $Pred(R_i)$. Ce prédicat est vrai si R_i est homogène, et il est faux dans le cas contraire. Les régions doivent respecter les quatre conditions suivantes :

- $\bigcup_i R_i = I$,
- R_i est constituée de pixels connexes pour tout i,
- $Pred(R_i) = vrai$ pour tout i,
- $Pred(R_i \cup R_j) = faux$ pour tout $i \neq j$, R_i et R_j sont adjacents dans I.

La première condition implique que chaque pixel de l'image doit appartenir à une région R_i et que l'union de toutes les régions correspond à l'image entière. La deuxième condition est relative à la structure des régions. Elle définit une région comme un sous-ensemble de pixels connexes. La troisième condition exprime que chaque région doit respecter un prédicat d'uniformité. La dernière condition implique la non réalisation de ce même prédicat pour toute réunion de deux régions adjacentes. Le résultat de la segmentation est une image dans laquelle est attribuée à chaque pixel une étiquette correspondant à la région à laquelle il appartient. Les méthodes de reconstruction de régions peuvent se diviser en deux grandes familles, selon que l'analyse s'effectue dans le plan image ou dans un espace couleur [28][61]. Ces deux familles de méthodes de segmentation d'images couleur représentent une liste de méthodes non exhaustive, car nous trouvons dans la littérature une grande diversité de méthodes et de manières à les classifier [3][67].

Nous pouvons tout d'abord noter des méthodes qui effectuent un traitement global sur l'image comme par exemple un seuillage [43] [91] sur l'intensité de l'image pour déterminer les objets ayant une intensité d'image proche (voir FIG.3.5).

FIG. 3.5 – Segmentation par seuillage d'histogramme.

Le seuillage peut aussi être déterminé en étudiant l'histogramme de l'intensité de l'image pour repérer les différents modes de la distribution de l'intensité et calculer ainsi des seuils optimaux afin de détecter les différents objets de la scène. Mais cette méthode ne fonctionne que si les objets ont une intensité homogène et de plus, qui soit différente de celle du fond. Certaines méthodes consistent à définir directement les caractéristiques des objets que l'on cherche à segmenter, comme par exemple leur couleur. Les régions obtenues auront donc une couleur homogène. Ce type de méthodes est efficace si les objets sont très homogènes, ce qui est rarement le cas. D'autres approches consistent à faire grandir un ensemble de germes choisis dans l'image. Ce sont les approches par croissance de régions [43]. Un critère d'homogénéité est choisi afin d'augmenter les régions. Quand plusieurs régions se rejoignent, leurs frontières délimitent les contours des régions finales. Un des principaux inconvénients de cette méthode est que les frontières obtenues ne correspondent pas forcément à des limites précises dans l'image. De plus, elle est assez sensible au nombre et au positionnement des germes initiaux. L'idée de ce type de segmentation est de partir d'un point d'amorce (appelé seed) que l'on étend en ajoutant les points de la frontière qui satisfont le critère d'homogénéité afin de constituer des régions.

FIG. 3.6 – Segmentation par croissance de régions.

Il existe des méthodes dites de division/fusion (Split and Merge en anglais) [51] qui consistent soit à diviser de manière très fine une image et ensuite à fusionner les régions adjacentes selon des critères d'homogénéité, soit à diviser l'image de façon itérative tant que les régions ne sont pas assez homogènes. On utilise souvent des arbres ou des graphes pour effectuer cette étape.

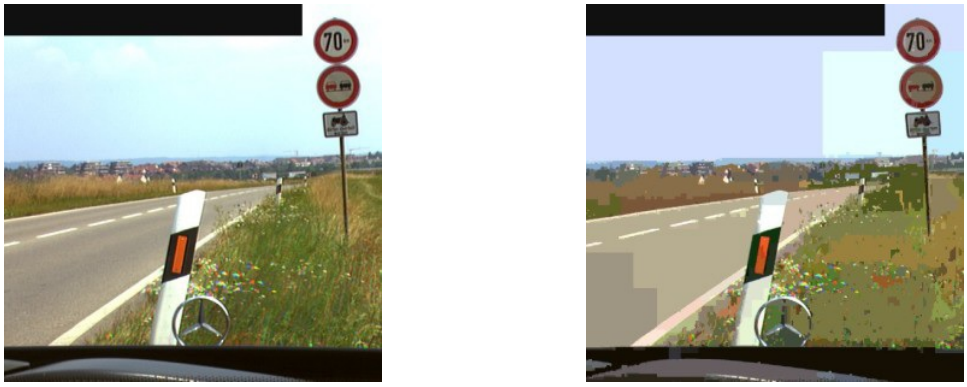


FIG. 3.7 – Segmentation par division/fusion.

Des méthodes morphologiques comme l'algorithme de la ligne de partage des eaux (LPE, ou watersheds en anglais) considèrent le gradient de l'image comme une surface topologique

et partent de germes initiaux pour immerger la surface [19]. Ainsi, les zones où se rencontre l'eau des différents bassins versants forment les frontières entre les régions. Ce genre de méthode est très rapide mais fournit un nombre très grand de régions qu'il faudra par la suite fusionner pour obtenir une segmentation correcte des objets de la scène (voir FIG.3.8).

FIG. 3.8 – Segmentation par la ligne de partage des eaux.

Nous pouvons également indiquer les méthodes de classification (plus souvent appelées clustering en anglais) qui consistent à regrouper en sous-ensembles les pixels qui possèdent des caractéristiques proches. Citons par exemple les méthodes de K-means [45] [69] qui consistent à séparer les pixels en K groupes en minimisant une distance entre un pixel donné et le représentant d'un groupe. Ces méthodes fournissent donc une partition de l'image mais ne permettent pas de distinguer quelles classes font partie de l'objet qui nous intéresse, et quelles classes font partie du fond (voir FIG.3.9).

FIG. 3.9 – Segmentation en 4 classes par le K-means.

Les méthodes de classification peuvent être séparées selon deux groupes, les méthodes supervisées et non supervisées, c'est-à-dire, avec ou sans apprentissage. Dans la classification non-supervisée [79][87][108], c'est le logiciel qui gère les sites des références sur lesquels est basée la classification, l'utilisateur fixe seulement le nombre de classes. Dans la classification supervisée, c'est l'utilisateur qui définit le nombre de classes, mais aussi les sites des références sur lesquels se base le logiciel pour les regroupements (c'est d'ailleurs la différence majeure entre la classification non-supervisée et celle supervisée). Dans notre travail, nous considérons qu'aucune connaissance a priori n'est connue sur les images couleurs utilisées, donc nous nous intéressons aux méthodes de segmentation d'images couleur par classification de pixels qui sont fondées sur la classification non-supervisées dans un cadre probabiliste.

3.4 Modélisation probabiliste en classification

Dans cette section, nous abordons le problème de la classification par une approche probabiliste. Cette approche, comme son nom l'indique, fait appel à la modélisation probabiliste.

L'objectif est d'établir une classification automatique des pixels en classes homogènes. Ici, le sens donné à l'homogénéité des classes se base sur l'analyse de la distribution de probabilité de l'image. Cette approche probabiliste présente deux avantages majeurs. D'une part, elle permet d'avoir accès à des probabilités d'appartenance des individus aux différentes classes. C'est d'ailleurs à partir de ces probabilités que s'établit la classification. D'autre part, le cadre formel de cette approche permet de proposer des solutions théoriques au problème de choix du nombre de classes, qui est en pratique inconnu. Dans le paragraphe 3.4.1, nous présentons le modèle sous sa forme générale et nous donnons la règle de classification des pixels dans les classes. Une fois le modèle identifié, il s'agit d'estimer ses paramètres. La méthode classique du maximum de vraisemblance ne pouvant pas être directement utilisée, pour cela nous faisons appel à l'algorithme EM, qui sera présenté dans le paragraphe 3.4.4. Dans le paragraphe 3.4.2, nous rappelons la modélisation par mélange de lois tout en précisant quelques exemples de mélanges. Le paragraphe 3.4.3 fera l'objet de l'estimation des paramètres du mélange.

3.4.1 Modèle aléatoire de l'image

Le modèle probabiliste le plus couramment utilisé en segmentation d'images couleur est le suivant : une image est modélisée par un couple $(X = (X_s)_{s\in S}, Z = (Z_s)_{s\in S})$ de champs aléatoires, où S représente l'ensemble des pixels d'une image : $S = \{s(i,j); 1 \leq i \leq N_c$ et $1 \leq i \leq N_l\}$. L'ensemble $\{X = (X_s)_{s\in S}\}$ représente le champ d'observations où chaque X_s est un vecteur aléatoire réel. L'ensemble $\{Z = (Z_s)_{s\in S}\}$ définit le champ de classes où chaque Z_s est à valeurs dans un ensemble fini de classes $\Omega = \{1, \ldots, K\}$ avec K désigne le nombre de classes. Ce modèle est un modèle hiérarchique [55] composé de

deux niveaux de champs aléatoires, c'est à dire, le processus de la région cachée Z et le processus observable X. Ce cadre probabiliste permet de construire la règle de decision optimale $\delta^*(x)$, dite règle de Bayes, qui minimise le risque conditionnel $R(\delta(x) \mid x)$ pour chaque observation x. En associant un coût nul à une bonne affectation et un coût de 1 à une mauvaise affectation, le risque conditionnel s'écrit :

$$R(\delta(x) \mid x) = 1 - P(Z = \delta(x) \mid X = x)$$

La regle $\delta^*(x)$ consiste donc à affecter l'observation x à la classe la plus probable a posteriori :

$$\delta^*(x) = \arg \max_{k=1,\ldots,K} P(Z = k \mid X = x)$$

Cette règle porte également le nom de MAP (maximum a posteriori).

En classification probabiliste, la règle de décision repose donc sur les probabilités a posteriori.

3.4.2 Modélisation par mélange de lois

Les modèles fondés sur les mélanges de lois sont riches, permettant ainsi la modélisation des situations variées et relativement aisées à manipuler car ils s'appuient sur des modèles élémentaires plus simples et bien connus. Ils sont utilisés avec succès en classification d'images [68][111].

Présentation et interprétation

Le modèle de mélange suppose que chaque classe est caractérisée par une distribution de probabilité. Dans un modèle de mélange, on considère que les données $x_1, \ldots, x_n$ constituent un échantillon de n réalisations indépendantes du vecteur aléatoire X à valeurs dans

$\mathbb{R}^p$ dont la fonction de densité peut s'écrire de la façon suivante :

$$f(x) = \sum_{k=1}^{K} \pi_k f_k(x), \tag{3.1}$$

où K est le nombre de classes, f_k est la densité de la distribution de X conditionnellement à la classe k et les π_k sont les proportions du mélange ($\pi_k \in (0,1)$ et $\sum_{k=1}^{K} \pi_k = 1$).

La définition générale de la loi mélange recouvre une famille très large de lois de probabilités. Nous supposons que les composantes f_k $(k = 1, \ldots, K)$ du mélange sont paramétrées par θ_k $(k = 1, \ldots, K)$. La densité de probabilité de la loi mélange donnée par l'équation (3.1) s'écrit alors :

$$f(x, \Phi) = \sum_{k=1}^{K} \pi_k f(x, \theta_k)$$

où $\Phi = (\pi_1, \ldots, \pi_k, \theta_1, \ldots, \theta_k)$ désigne le vecteur des paramètres du mélange.

Les composantes $f(x, \theta_k)$ appartiennent souvent à une même famille de lois dont le choix dépend des données à représenter. Par exemple, Mendoza Rosas et Cruz-Reyna [76] modélisent le temps de repos d'un volcan entre deux éruptions (données continues) par un mélange de lois exponentielles. Il existe des cas où les composantes d'un mélange n'appartiennent pas à la même famille de lois : Dasgupta et Raftery [31] introduisent une composante uniforme dans un mélange Gaussien pour modéliser des données bruitées.

Si l'on se place dans le cadre du modèle de mélange paramétrique (Gaussien ou non), la formule de Bayes permet d'écrire :

$$P(Z = k \mid X = x, \Phi) = \frac{\pi_k f(x, \theta_k)}{f(x)}$$

où $f(x) = \sum_{k=1}^{K} \pi_k f_k(x, \theta_k)$ est une quantité commune à chacune des classes. La règle de décision de Bayes peut alors être écrite de la façon suivante :

$$\delta^*(x) = \arg \max_{k=1,\ldots,K} \{\pi_k f(x, \theta_k)\}.$$

À la vue de cette reformulation, il apparaît clairement que, dans le cadre du modèle de mélange, le problème de la classification se résume à l'estimation des paramètres du modèle. Nous citons dans la suite quelques exemples de modèles de mélanges de lois.

3.4.2.a Modèle de mélange Gaussien

Dans la littérature, le modèle de mélange Gaussien est le plus utilisé en classification d'images. La densité de probabilité $f(x, \theta_k)$ $(k = 1, \ldots, K)$ d'une loi Gausienne $N_p(\mu_k, \Sigma_k)$, où μ_k est le vecteur moyen et Σ_k est la matrice de variance-covariance est donnée par :

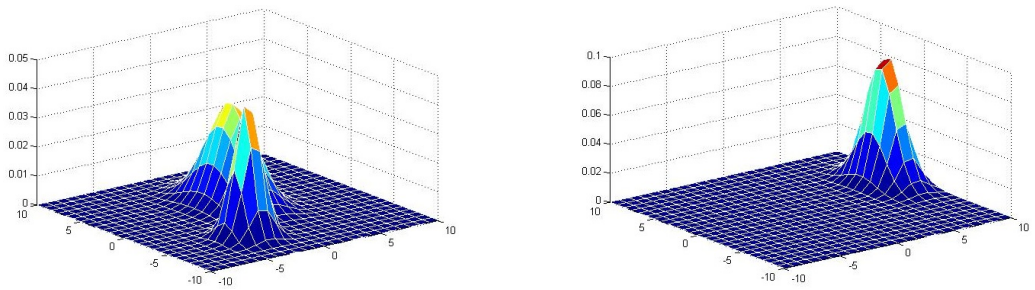
$$f(x, \theta_k) = \frac{1}{2\pi\sqrt{\det \Sigma_k}} \exp\left(\frac{-1}{2}(x - \mu_k)^t.\Sigma_k^{-1}.(x - \mu_k)\right)$$

avec $\theta_k = (\mu_k, \Sigma_k)$.

Nous présentons dans les figures FIG.3.10.a et FIG.3.10.b deux mélanges Gaussiens pour $p = 2$. Les valeurs des paramètres de chaque mélange sont données dans le tableau TAB.3.1.

Paramètres	π_1	π_2	μ_1	μ_2	Σ_1	Σ_2
FIG.3.10.a	0.7	0.3	$\begin{pmatrix} 3 \\ 4 \end{pmatrix}$	$\begin{pmatrix} 5 \\ 6 \end{pmatrix}$	$\begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}$	$\begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix}$
FIG.3.10.b	0.4	0.6	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 8 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$	$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$

TAB. 3.1 – Valeurs des paramètres des mélanges Gaussiens.



(a) (b)

FIG. 3.10 – Mélanges Gaussiens de deux composantes.

3.4.2.b Modèle de mélange de lois Bêta de première espèce

Dans le cas $p = 2$, la densité de probabilité $f((x_{1,k}, x_{2,k}), \theta_k)$ $(k = 1, \ldots, K)$ de la loi Bêta bivariée de première espèce est définie par :

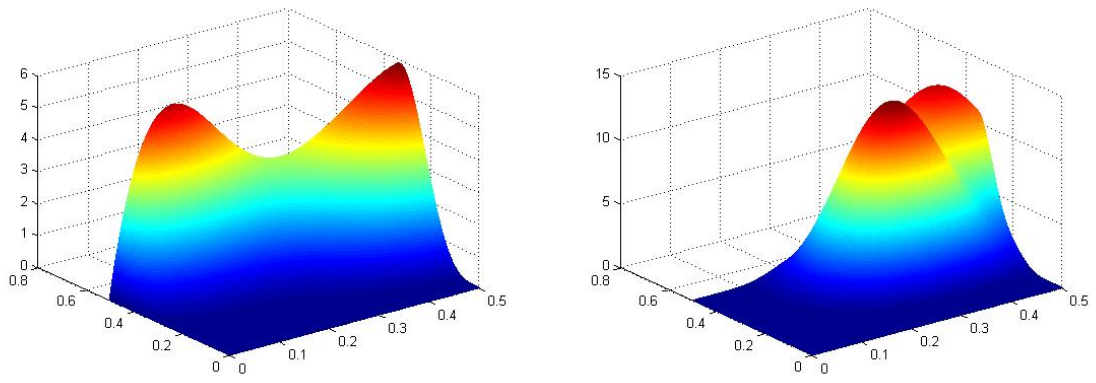
$$f((x_{1,k}, x_{2,k}), \theta_k) = \frac{1}{B(a_k, b_k, c_k)} x_{1,k}^{a_k-1} x_{2,k}^{b_k-1} (1 - x_{1,k} - x_{2,k})^{c_k-1} \mathbf{1}_D(x_{1,k}, x_{2,k}) \tag{3.2}$$

avec $D = \{x_{1,k} > 0,\ x_{2,k} > 0;\ x_{1,k} + x_{2,k} < 1\}$, $B(a_k, b_k, c_k) = \dfrac{\Gamma(a_k)\Gamma(b_k)\Gamma(c_k)}{\Gamma(a_k + b_k + c_k)}$ *et* $\theta_k = (a_k, b_k, c_k)$, $k = 1, \ldots, K$.

Nous présentons dans les figures FIG.3.11.a et FIG.3.11.b deux mélanges de lois Bêta de première espèce pour $p = 2$. Les valeurs des paramètres de chaque mélange sont données dans le tableau TAB.3.2.

Paramètres	π_1	π_2	a_1	a_2	b_1	b_2	c_1	c_2
FIG.3.11.a	0.4	0.6	8	2	5	9	3	4
FIG.3.11.b	0.5	0.5	10	12	10	11	13	5

TAB. 3.2 – Valeurs des paramètres des mélanges Bêta de première espèce.



(a) (b)

FIG. 3.11 – Mélanges de lois Bêta de première espèce de deux composantes.

3.4.2.c Modèle de mélange de lois Bêta de deuxième espèce

Dans le cas $p = 2$, la densité de probabilité $f((x_{1,k}, x_{2,k}), \theta_k)$ $(k = 1, \ldots, K)$ de la loi Bêta bivariée de deuxième espèce est définie par :

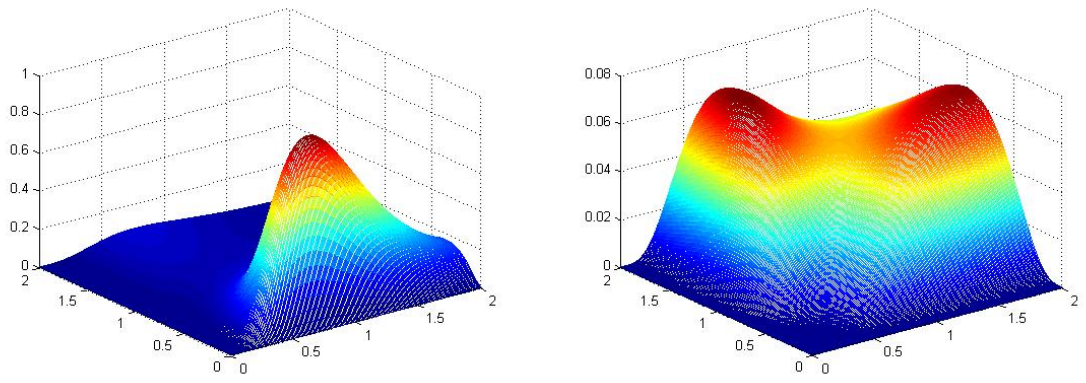
$$f((x_{1,k}, x_{2,k}), \theta_k) = \frac{1}{B(a_k, b_k, c_k)} x_{1,k}^{a-1} x_{2,k}^{b_k-1} (1 + x_{1,k} + x_{2,k})^{-(a_k+b_k+c_k)}, \quad x_{1,k} > 0, \; x_{2,k} > 0, \tag{3.3}$$

$$avec\ B(a_k, b_k, c_k) = \frac{\Gamma(a_k)\Gamma(b_k)\Gamma(c_k)}{\Gamma(a_k + b_k + c_k)}\ et\ \theta_k = (a_k, b_k, c_k),\ k = 1, \ldots, K.$$

Nous présentons dans les figures FIG.3.12.a et FIG.3.12.b deux mélanges de lois Bêta de deuxième espèce pour $p = 2$. Les valeurs des paramètres de chaque mélange sont données dans le tableau TAB.3.3.

Paramètres	π_1	π_2	a_1	a_2	b_1	b_2	c_1	c_2
FIG.3.12.a	0.4	0.6	15	3	6	8	4	7
FIG.3.12.b	0.8	0.2	5	13	9	6	2	5

TAB. 3.3 – Valeurs des paramètres des mélanges Bêta de deuxième espèce.



(a) (b)

FIG. 3.12 – Mélanges de lois Bêta de deuxième espèce de deux composantes.

3.4.3 Estimation des paramètres d'un modèle mélange

Dans ce paragraphe, nous décrivons la méthode classique du maximum de vraisemblance pour estimer les paramètres du modèle mélange. Cette méthode consiste à rechercher les valeurs des paramètres qui maximisent la vraisemblance (ou plutôt le logarithme de la vraisemblance) des données observées.

Données complètes et vraisemblance complète

Nous considérerons que les données observées $x_1, \ldots, x_n$ ne correspondent qu'à une connaissance partielle des données complètes $h_1, \ldots, h_n$ qui sont supposées être de la forme $h_j = (x_j, z_j), j = 1, \ldots, n$, ou z_j est le numéro de la classe à laquelle appartient l'observation x_j. Nous verrons que dans le cas de la classification non supervisée, cette dernière information sera cachée et z_j sera appelé une donnée manquante. On appellera vraisemblance complète, la vraisemblance calculée à partir des données complètes $(x_1, z_1), \ldots, (x_n, z_n)$ et elle sera notée $L((x, z); \Phi)$, ou plus simplement $L(\Phi)$.

Estimation par maximum de vraisemblance

La méthode du maximum de vraisemblance (MMV) permet d'estimer le paramètre Φ du modèle par $\widehat{\Phi}_{MMV}$:

$$\widehat{\Phi}_{MMV} = \arg\max_{\Phi} L(\Phi).$$

Les n données $(x_j, z_j), j = 1, \ldots, n$, sont supposées indépendantes, on peut exprimer la vraisemblance du paramètre Φ par le produit de toutes les densités marginales. Pour simplifier le calcul on utilise le logarithme néperien de la fonction vraisemblance. Dans le cadre du modèle de mélange, la fonction log-vraisemblance s'écrit alors de la façon suivante :

$$l(\Phi) = \ln(L(\Phi) = \sum_{j=1}^{n} \ln(\sum_{k=1}^{K} \pi_k f(x_j, \theta_k)).$$

Si de plus, les labels des observations sont connus, ($i.e.$ dans le cas supervisé), cette fonction $l(\Phi)$ peut s'écrire sous la forme suivante :

$$l(\Phi) = \sum_{j=1}^{n} \sum_{k=1}^{K} \delta_{kj} \ln(\pi_k f(x_j, \theta_k)),$$

où $\delta_{kj} = 1$ si l'observation x_j appartient à la classe k et $\delta_{kj} = 0$ sinon.

La maximisation de la fonction log-vraisemblance d'un modèle de mélange dans le cas non-supervisé conduit en général à des équations de vraisemblance qui ne possèdent pas de solutions analytiques. Il existe différents algorithmes permettant de maximiser la fonction log-vraisemblance quand les labels sont inconnus. L'algorithme le plus utilisé d'entre eux est l'algorithme itératif EM (Expectation-Maximization) [26].

3.4.4 Algorithme EM

L'algorithme EM repose sur l'idée qu'il est plus facile de maximiser la vraisemblance complète $L((x,z);\Phi)$ que la vraisemblance $L(x;\Phi)$ et se base sur la relation suivante entre les deux vraisemblances :

$$L(x;\theta) = L((x,z);\Phi) - \ln f((x,z) \mid x,\Phi).$$

Cependant, la vraisemblance complète n'est pas calculable du fait que z n'est pas totalement connue. Dempster, Laird et Rubin [33] ont proposé pour maximiser cette vraisemblance, une procédure itérative qui se base sur la maximisation de l'espérance conditionnelle de la vraisemblance pour une valeur du paramètre courant Φ. L'algorithme consiste donc simplement à construire une suite $(\Phi^{(q)})$ qui verifie :

$$\Phi^{(q+1)} = \arg\max_{\Phi} Q(\Phi,\Phi^{(q)}),$$

où $Q(\Phi,\Phi^{(q)}) = E[L((x,z);\Phi) \mid x,\Phi^{(q)}]$. Si l'on note $\tau_{kj} = E[Z_k \mid X = x_j,\Phi^{(q)}]$ l'espérance du k^{eme} élément du vecteur aléatoire Z conditionnellement à $X = x_j$ et à $\Phi^{(q)}$, on peut alors écrire :

$$Q(\Phi,\Phi^{(q)}) = \sum_{j=1}^{n}\sum_{k=1}^{K} t_{kj}^{(q)} \ln(\pi_k f(x_j,\theta_k))$$

De plus, il a été montré que la suite ainsi générée $\Phi^{(q)}$ fait croitre la vraisemblance $L(x;\theta)$ et converge vers un maximum local de celle-ci sous certaines conditions de régularité. On pourra consulter [33] pour un plus simple développement du principe de cet algorithme d'estimation.

Les étapes de l'algorithme EM

Partant d'une solution initiale $\Phi^{(0)}$, l'algorithme EM altèrne entre deux étapes appelées

respectivement **E**, pour expectation, et **M**, pour maximisation :

Étape E : cette étape consiste à estimer, à l'itération q, l'espérance de Z_k sachant la valeur du paramètre estimé à l'étape précédente et $X = x_j$:

$$\begin{aligned} \tau_{kj}^{(q)} &= E[Z_k \mid X = x_j, \widehat{\Phi}^{(q)}] \\ &= P[Z = k \mid X = x_j, \widehat{\Phi}^{(q)}] \end{aligned} .$$

La formule de Bayes permet finalement d'exprimer les probabilités τ_{kj} d'appartenance des x_j aux classes k, de la façon suivante :

$$\tau_{kj}^{(q)} = \frac{\pi_k^{(q)} f(x_j, \widehat{\theta}_k^{(q)})}{\sum_{k=1}^{K} \pi_k^{(q)} f(x_j, \widehat{\theta}_k^{(q)})}$$

Étape M : Cette étape consiste à estimer à chaque itération $(q+1)$ les proportions du mélange par :

$$\widehat{\pi}_k^{(q+1)} = \frac{\sum_{j=1}^{n} \tau_{kj}^{(q)}}{n}$$

et les paramètres $\theta_1, \ldots, \theta_K$ seront obtenus en résolvant les équations de vraisemblance correspondantes au modèle de mélange retenu.

Remarque 3.4.1. *Le calcul des valeurs initiales des paramètres du modèle mélange $\Phi^{(0)}$ est probablement l'étape la plus délicate. L'algorithme EM a la propriété d'augmenter la vraisemblance des observations. Il converge donc vers un optimum. Cependant rien ne garantit que cet optimum soit global. En fait, l'algorithme EM converge vers le plus proche optimum local. Le résultat de l'algorithme dépend donc de l'initialisation. Ce problème a été discuté par de nombreux auteurs [16][77] qui ont proposé plusieurs solutions pour faciliter la convergence de l'algorithme et éviter de tomber sur des maxima locaux. Sans information a priori sur les paramètres, plusieurs stratégies d'initialisation existent :*

- Effectuer plusieurs initialisations des paramètres au hasard et comparer les vraisemblances obtenues après convergence pour éliminer les maxima locaux.

- Initialiser les probabilités a posteriori τ_{kj} d'appartenance aux classes à 0 ou à 1 selon la partition initiale et démarrer l'EM par l'étape M à la première itération (ce qui permettra d'initialiser les paramètres).

- Utiliser les estimateurs des paramètres par la méthode des moments.

Dans la suite, nous présontons quelques exemples d'application de l'algorithme EM pour le cas où $p = 2$.

3.4.4.a Application de l'algorithme EM à un mélange de lois Gaussiennes

Les paramètres à estimer sont les probabilités a priori, les vecteurs moyens et les matrices de variance-covariances de chaque classe. À l'étape M de l'itération $(q+1)$, nous obtenons :

$$\widehat{\pi}_k^{(q+1)} = \frac{1}{n}\sum_{j=1}^{n}\tau_{kj}^{(q)}, \quad \widehat{\mu}_k^{(q+1)} = \frac{1}{n}\sum_{j=1}^{n}\tau_{kj}^{(q)}x_j$$

$$\widehat{\Sigma}_k^{(q+1)} = \frac{1}{n}\sum_{j=1}^{n}\tau_{kj}^{(q)}(x_j - \widehat{\mu}_k^{(q+1)})(x_j - \widehat{\mu}_k^{(q+1)})^t.$$

Nous présentons dans la figure FIG.3.13 l'estimation d'un mélange de trois lois Gaussiennes en utilisant l'algorithme EM. Comme nous avons montré dans la sous section 2.2.2.a du chapitre 2, que l' estimateur du vecteur des paramètres θ d'une loi Gausienne obetnu par la méthode des moments (MM) et par la méthode du maximum de vraisemblence (MMV) est le même, nous avons initialisé θ par l'estimateur obtenu par la MM. Les valeurs des paramètres de ce mélange ainsi que les EQMs entre ses valeurs et leurs valeurs estimées sont définis dans le tableau TAB.3.4.

Dans le tableau TAB.3.5, nous présontons pour chaque mélange l'EQM entre la vraie

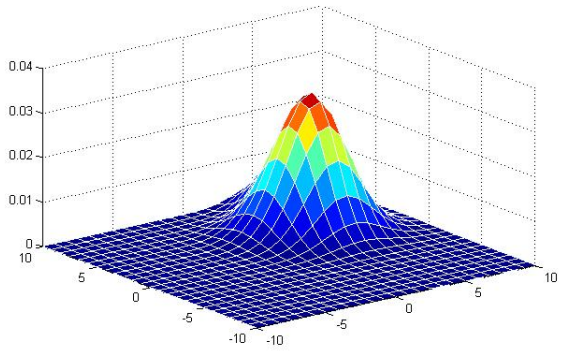
densité mélange et celle estimée.

Paramètres	π_1	π_2	π_3	μ_1	μ_2	μ_3	Σ_1	Σ_2	Σ_2
Mélange 1	0.3	0.3	0.4	$\begin{pmatrix} 3 \\ 4 \end{pmatrix}$	$\begin{pmatrix} 5 \\ 6 \end{pmatrix}$	$\begin{pmatrix} 4 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$	$\begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix}$	$\begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}$
EQM	0.0011	0.009	0.0039	0.036	0.0028	2.98	1.91	1.5904	2.709
Mélange 2	0.2	0.5	0.3	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 8 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$	$\begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}$
EQM	0.0012	3.32×10^{-4}	2.64×10^{-4}	0.0057	0.004	2.93	1.0004	0.6164	4.2185
Mélange 3	0.2	0.4	0.4	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$
EQM	0.111	0.0309	0.00248	0.0051	0.0026	2.199	0.1031	0.3403	1.2838

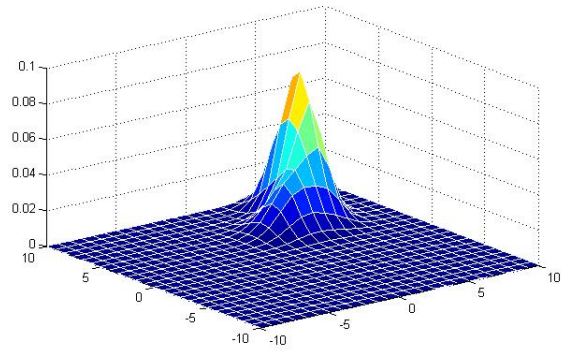
TAB. 3.4 – Valeurs estimées des paramètres du mélange Gaussien.

Mélange	Mélange 1	Mélange 2	Mélange 3
EQM	3.75×10^{-5}	5.14×10^{-5}	6.278×10^{-5}

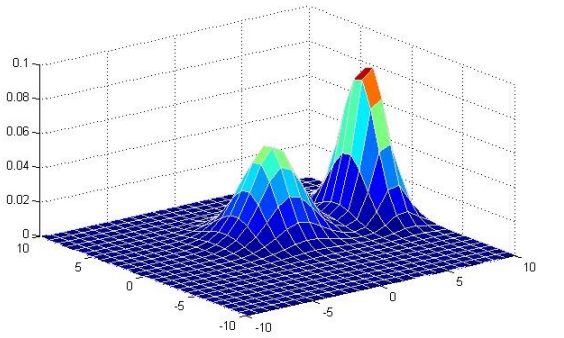
TAB. 3.5 – Valeurs de l'EQM pour le mélange de trois lois Gaussiennes.



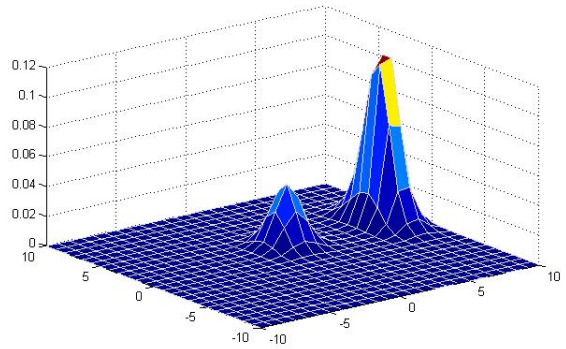
Mélange 1



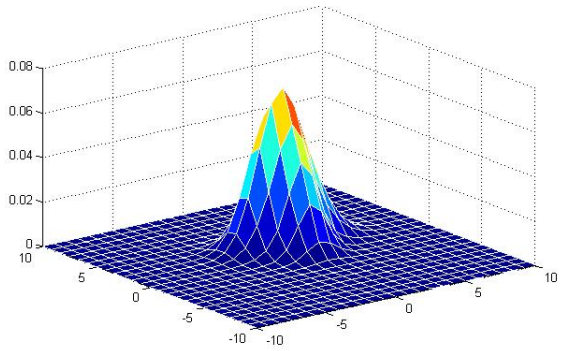
Mélange 1 estimé



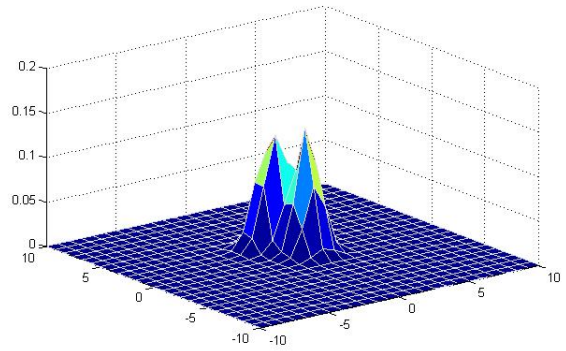
Mélange 2



Mélange 2 estimé



Mélange 3



Mélange 3 estimé

FIG. 3.13 – Estimation de mélange de trois lois Gaussiennes.

3.4.4.b Application de l'algorithme EM à un mélange de lois Bêta de première espèce

Les paramètres à estimer sont les probabilités a priori π_k, le vecteur $\theta_k = (a_k, b_k, c_k)$ pour chaque classe $k = 1, \ldots, K$.

À l'étape **E** de l'itération (q), la probabilité a postériori $\tau_{kj}^{(q)}$ est définie par

$$\tau_{kj}^{(q)} = \frac{\pi_k^{(q)} f((x_{1,k}, x_{2,k}); \theta_k)}{\sum_{k=1}^{K} \pi_k^{(q)} f((x_{1,k}, x_{2,k}); \theta_k),}$$

avec $f((x_{1,k}, x_{2,k}); \theta_k)$ définie par l'équation (3.2).

À l'étape **M** de l'itération $(q+1)$, l'estimateur de la probabilité a priori est donné par

$$\widehat{\pi}_k^{(q+1)} = \frac{1}{n} \sum_{j=1}^{n} \tau_{kj}^{(q)}$$

et les estimateurs de a_k, b_k et c_k sont les solutions des trois équations suivantes :

$$\sum_{j=1}^{n} \tau_{k,j}^{(q)} [\psi(a_k + b_k + c_k) - \psi(a_k) + \ln(X_{1,k}^j)] = 0, \tag{3.4}$$

$$\sum_{j=1}^{n} \tau_{k,j}^{(q)} [\psi(a_k + b_k + c_k) - \psi(b_k) + \ln(X_{2,k}^j)] = 0, \tag{3.5}$$

$$\sum_{j=1}^{n} \tau_{k,j}^{(q)} [\psi(a_k + b_k + c_k) - \psi(c_k) - \ln(1 - X_{1,k}^j - X_{2,k}^j)] = 0, \tag{3.6}$$

avec ψ représente la fonction digamma.

Théoriquement, nous ne pouvons pas déterminer une expression explicite des estimateurs des paramètres $(\widehat{a}_k, \widehat{b}_k, \widehat{c}_k)$, pour $k = 1, 2, \ldots, K$. Pour résoudre ce problème, nous avons appliqué la méthode numérique de Newton Raphson.

Nous présentons dans la figure FIG.3.14 l'estimation d'un mélange de trois lois Bêta de

première espèce en utilisant l'algorithme EM avec deux méthodes d'initialisation du vecteur des paramètres $\theta = (a, b, c)$ de la loi Bêta bivariée de première espèce. La première méthode est basée sur l'initialisation par l'estimateur $\widehat{\theta}$ obetnu par la MM du vecteur des paramètres θ. La deuxième méthode est basée sur l'initialisation par l'estimateur $\widehat{\theta}$ obetnu par la MMV du vecteur des paramètres θ. Ces estimateurs sont présentés dans la sous section 2.2.2.a du chapitre 2. Les paramètres de ce mélange et la valeur de l'EQM pour chaque paramètre sont présentés dans le tableau TAB.3.6.

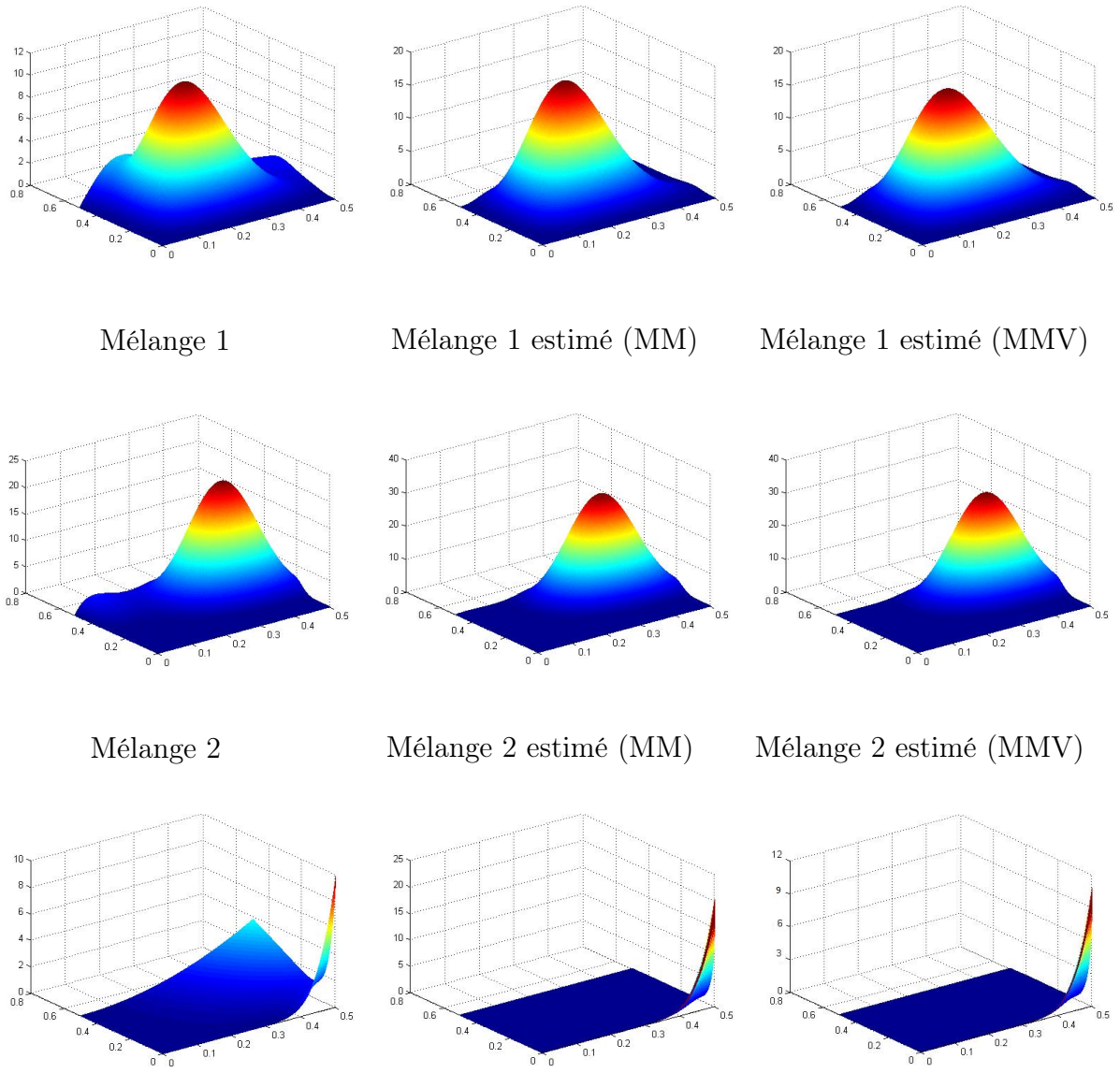
Dans le tableau TAB.3.7, nous présontons pour chaque mélange l'EQM entre la vraie densité mélange et celle estimée. D'après ce tableau, nous constatons que l'initialisation des paramètres dans l'algorithme EM par la MMV est plus performante que celle obtenue lorsque ses paramètres sont initialisés par la MM.

Paramètres	π_1	π_2	π_3	a_1	a_2	a_3	b_1	b_2	b_3	c_1	c_2	c_3
Mélange 1	0.1	0.3	0.6	8	2	4	5	9	5	3	4	11
EQM(MM)	0.15	0.457	0.748	0.7052	0.011	1.011	0.044	1.6789	0.011	0.0048	0.0013	0.017
EQM(MMV)	0.01	0.09	0.16	1.35	0.0064	0.0047	0.39	1.81	4.36×10^{-5}	0.017	0.0136	0.057
Mélange 2	0.7	0.1	0.2	13	2	4	12	11	12	14	9	3
EQM(MM)	0.87	0.009	0.04	0.0172	0.163	0.0062	0.0085	0.695	1.012	0.143	0.659	0.122
EQM(MMV)	0.9	0.01	0.04	0.003	0.045	0.547	0.0049	0.608	0.547	0.013	0.391	6.31×10^{-5}
Mélange 3	0.5	0.2	0.3	3	22	44	2	1	12	1	13	5
EQM(MM)	0.25	0.01	0.219	0.457	0.995	0.3624	0.087	0.054	0.0088	0.0215	0.4538	0.0048
EQM(MMV)	0.25	5.5×10^{-4}	0.226	0.553	0.099	0.00515	0.00987	0.035	0.0198	0.027	0.44	0.0043

TAB. 3.6 – Valeurs des paramètres de mélange de trois Bêta bivariées de première espèce.

Mélange	Mélange 1	Mélange 2	Mélange 3
EQM(MMV)	2.014×10^{-6}	2.34×10^{-7}	1.01×10^{-7}
EQM(MM)	4.86×10^{-5}	3.92×10^{-5}	2.9584×10^{-6}

TAB. 3.7 – Valeurs de l'EQM pour le mélange de trois Bêta bivariées de première espèce.



Mélange 1 | Mélange 1 estimé (MM) | Mélange 1 estimé (MMV)

Mélange 2 | Mélange 2 estimé (MM) | Mélange 2 estimé (MMV)

Mélange 3 | Mélange 3 estimé (MM) | Mélange 3 estimé (MMV)

FIG. 3.14 – Estimation de mélange de trois Bêta bivariées de première espèce.

3.4.4.c Application de l'algorithme EM à un mélange de lois Bêta de deuxième espèce

Les paramètres à estimer sont les probabilités a priori π_k et le vecteur $\theta_k = (a_k, b_k, c_k)$ pour chaque classe $k = 1, \ldots, K$.

À l'étape **E** de l'itération (q), la probabilité a postériori $\tau_{kj}^{(q)}$ est définie par

$$\tau_{kj}^{(q)} = \frac{\pi_k^{(q)} f((x_{1,k}, x_{2,k}); \theta_k)}{\sum_{k=1}^{K} \pi_k^{(q)} f((x_{1,k}, x_{2,k}); \theta_k),}$$

avec $f((x_{1,k}, x_{2,k})$ définie par l'équation (3.3).

À l'étape **M** de l'itération $(q+1)$, l'estimateur de la probabilité a priori est donné par

$$\widehat{\pi}_k^{(q+1)} = \frac{\sum_{j=1}^{n} \tau_{kj}^{(q)}}{n}$$

et les estimateurs de a_k, b_k et c_k sont les solutions des trois équations suivantes :

$$\sum_{j=1}^{n} \tau_{k,j}^{(q)} [\psi(a_k + b_k + c_k) - \psi(a_k) + \ln(X_{1,k}^j) - \ln(1 + X_{1,k}^j + X_{2,k}^j)] = 0, \tag{3.7}$$

$$\sum_{j=1}^{n} \tau_{k,j}^{(q)} [\psi(a_k + b_k + c_k) - \psi(b_k) + \ln(X_{2,k}^j) - \ln(1 + X_{1,k}^j + X_{2,k}^j)] = 0, \tag{3.8}$$

$$\sum_{j=1}^{n} \tau_{k,j}^{(q)} [\psi(a_k + b_k + c_k) - \psi(c_k) - \ln(1 + X_{1,k}^j + X_{2,k}^j)] = 0, \tag{3.9}$$

avec ψ représente la fonction digamma.

Comme nous ne pouvons pas déterminer une expression explicite des estimateurs des paramètres $(\widehat{a}_k, \widehat{b}_k, \widehat{c}_k)$, pour $k = 1, 2, \ldots, K$, nous avons appliqué la méthode numérique de Newton Raphson.

Nous présentons dans la figure FIG.3.15 l'estimation d'un mélange de trois lois Bêta de deuxième espèce avec deux méthodes d'initialisation du vecteur des paramètres $\theta = (a, b, c)$ de la loi Bêta bivariée de deuxième espèce. La première méthode est basée sur l'initialisation par l'estimateur $\widehat{\theta}$ obetnu par la MM du vecteur des paramètres θ. La deuxième méthode est basée sur l'initialisation par l'estimateur $\widehat{\theta}$ obetnu par la MMV du vecteur des paramètres θ. Ces estimateurs sont présentés dans la sous section 2.2.2.a du chapitre

2. Les paramètres de ce mélange et l'EQM pour chaque paramètre sont présentés dans le tableau TAB.3.8. Dans le tableau TAB.3.9, nous présontons pour chaque mélange l'EQM entre la vraie densité mélange et celle estimée. D'après ce tableau, nous constatons que l'initialisation des paramètres dans l'algorithme EM par la MMV est plus performante que celle obtenue lorsque ses paramètres sont initialisés par la MM.

Paramètres	π_1	π_2	π_3	a_1	a_2	a_3	b_1	b_2	b_3	c_1	c_2	c_3
Mélange 1	0.2	0.4	0.4	5	3	4	6	2	8	4	7	7
EQM(MM)	0.59	0.16	0.138	0.4529	0.987	1.335	0.561	0.708	0.60	0.378	1.558	0.619
EQM(MMV)	0.04	0.148	0.36	0.0029	0.0627	0.0075	0.0024	0.005	0.008	0.18	0.97	0.0094
Mélange 2	0.2	0.7	0.1	15	13	20	8	1	4	3	9	10
EQM(MM)	0.04	0.44	0.746	0.262	0.309	0.0039	0.414	0.0457	0.0096	1.873	0.656	0.0184
EQM(MMV)	0.0294	0.298	0.455	0.018	0.013	0.055	0.079	0.042	0.00206	1.373	0.463	0.044
Mélange 3	0.4	0.2	0.4	5	3	5	2	11	7	13	1	1
EQM(MM)	0.36	0.058	0.486	0.0084	0.0409	0.0539	0.0066	0.3456	1.557	0.075	1.2013	1.68
EQM(MMV)	0.089	0.0125	0.156	0.0055	4.19×10^{-4}	0.68	7.30×10^{-4}	0.0153	1.007	0.001	0.5831	0.619

TAB. 3.8 – Valeurs des paramètres estimés de mélange de trois lois Bêta bivariées de deuxième espèce.

Mélange	Mélange 1	Mélange 2	Mélange 3
EQM (MMV)	2.26×10^{-6}	4.91×10^{-4}	1.056×10^{-5}
EQM(MM)	2.62×10^{-6}	9.34×10^{-3}	1.9109×10^{-6}

TAB. 3.9 – Valeurs de l'EQM pour le mélange de trois Bêta bivariées de deuxième espèce.

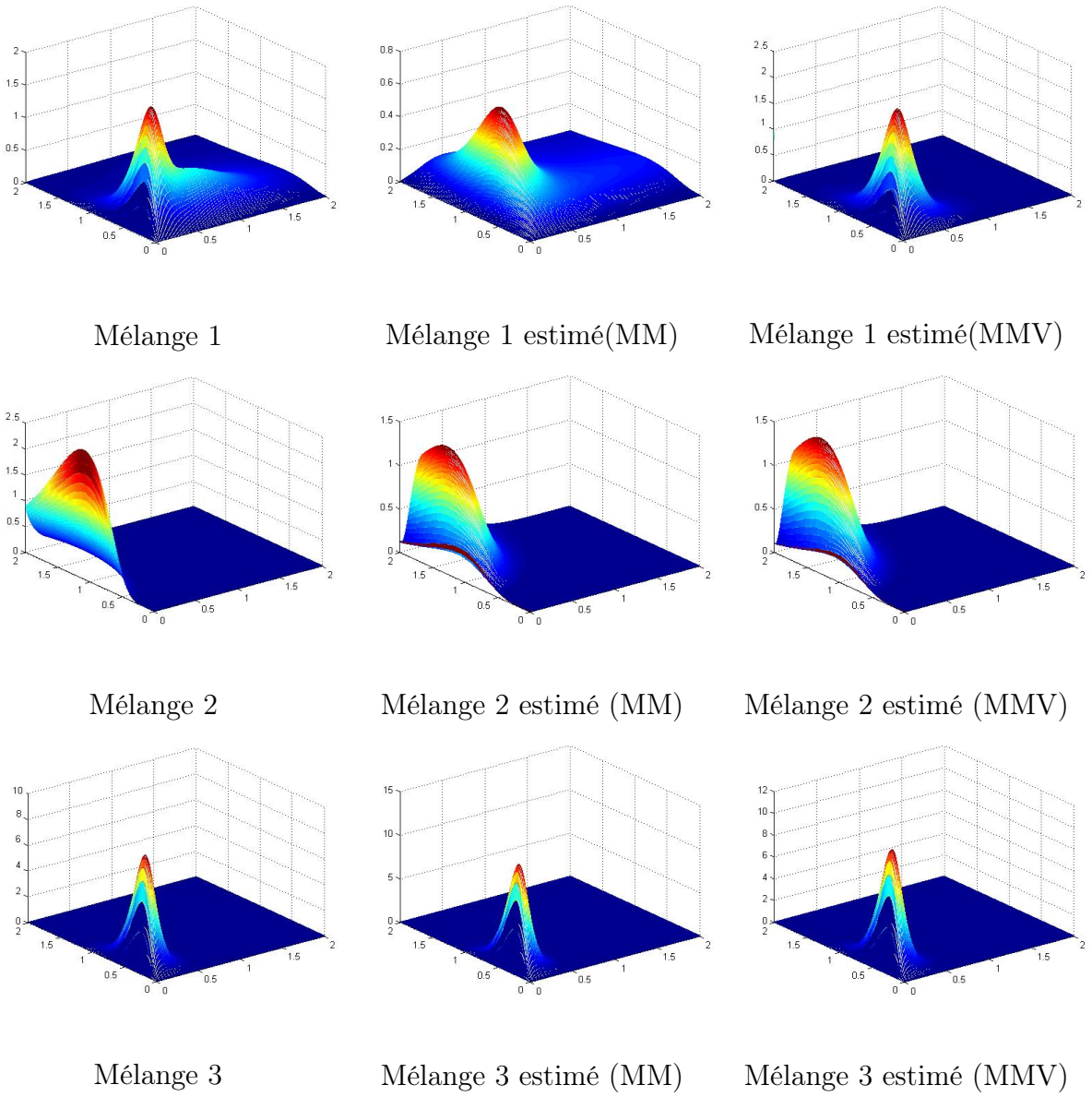


FIG. 3.15 – Estimation de mélange de trois Bêta bivariées de deuxième espèce.

3.5 Application de l'algorithme EM pour un mélange de lois de la diagonale modifiée de Riesz

Ce paragraphe fait l'objet de notre contribution concernant la généralisation du modèle des lois de la diagonale modifiée de Riesz proposé dans le chapitre 2 dans le cadre du mélange des lois. Ce modèle est caractérisé par le vecteur des paramètres fixés $(s_1, s_2) \in \mathbb{R}_+^* \times]\frac{1}{2}, +\infty[$ et le vecteur des paramètres $\theta = (\Sigma_{11}, \Sigma_{22}, \Sigma_{12}^2) \in D$, avec $D = \{\Sigma_{11} > 0,\ \Sigma_{22} > 0$ et$\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2 > 0\}$. La généralisation de ce modèle sera traité sous deux conditions : $s_1 = s_2 = s$ et $s_2 > s_1$. Si $s_1 = s_2 = s$, le modèle du mélange est appelé mélange Gamma bivarié de Kibble, noté MGBK. Si $s_2 > s_1$, le modèle du mélange est appelé mélange de produit de convolution d'une loi Gamma bivariée de Kibble de paramètres $s_1, \Sigma_{11}, \Sigma_{22}$, et Σ_{12}^2 et d'une loi Gamma univariée de paramètres $s_2 - s_1$ et Σ_{22}, noté MCGBKGU.

3.5.1 Estimation des paramètres du mélange MGBK

Soit $(X^1, \ldots, X^n)$ un $n-$échantillon aléatoire de loi Gamma bivariée de Kibble de vecteur paramètres $\theta = (\Sigma_{11}, \Sigma_{22}, \Sigma_{12}^2)$. La densité du mélange Gamma bivariée de Kibble est donnée par :

$$\begin{aligned} f((x_1, x_2), \Phi) &= \sum_{k=1}^{K} \pi_k f((x_{1,k}, x_{2,k}); \theta_k) \\ &= \sum_{k=1}^{K} \pi_k \exp(-\frac{\Sigma_{22,k} x_{1,k} + \Sigma_{11,k} x_{2,k}}{\Sigma_{11,k}\Sigma_{22,k} - \Sigma_{12,k}^2}) \frac{(x_{1,k} x_{2,k})^{s-1}}{(\Sigma_{11,k}\Sigma_{22,k} - \Sigma_{12,k}^2)^s \Gamma(s)} \\ &\quad \times f_s(\delta_k x_{1,k} x_{2,k}), \end{aligned}$$

où $\delta_k = \dfrac{\Sigma_{12,k}}{\Sigma^2_{1.2,k}}$, π_k la probabilité a priori et $\theta_k = (\Sigma_{11,k}, \Sigma_{22,k}, \Sigma^2_{12,k})$.

Pour estimer le vecteur des paramètres du mélange

$\Phi = (\pi_1, \ldots, \pi_K, \Sigma_{11,1}, \ldots, \Sigma_{11,K}, \Sigma_{22,1}, \ldots, \Sigma_{22,K}, \Sigma^2_{12,1}, \ldots, \Sigma^2_{12,K})$, nous appliquons l'algorithme EM.

À l'étape **E** de l'itération (q), la probabilité a postériori $\tau^{(q)}_{kj}$ est définie par

$$\tau^{(q)}_{kj} = \frac{\pi^{(q)}_k f((x_{1,k}, x_{2,k}); \theta^{(q)}_k)}{\displaystyle\sum_{k=1}^{K} \pi^{(q)}_k f((x_{1,k}, x_{2,k}); \theta^{(q)}_k)}.$$

À l'étape **M** de l'itération $(q+1)$, l'estimateur $\widehat{\Phi}^{(q+1)}$ du vecteur des paramètres Φ est défini par $\widehat{\Phi}^{(q+1)} = \arg\max_{\Phi} Q(\Phi, \Phi^{(q)})$ avec

$$\begin{aligned} Q(\Phi, \Phi^{(q)}) &= \sum_{j=1}^{n}\sum_{k=1}^{K} t^{(q)}_{kj} \ln(\pi_k f((x^j_{1,k}, x^j_{2,k}); \theta_k) \\ &= \sum_{j=1}^{n}\sum_{k=1}^{K} t_{kj}[\ln(\pi_k) - s\ln(\Sigma_{1.2,k}) - \frac{\Sigma_{22,k}x^j_{1,k} + \Sigma_{11,k}x^j_{2,k}}{\Sigma_{1.2,k}} + (s-1)\ln(x^i_{1,k}x^i_{2,k}) \\ &\quad - \ln\Gamma(s) + \ln(f_s(\delta_k x^i_{1,k} x^i_{2,k}))]. \end{aligned}$$

Ceci, nous permet de déterminer l'estimateur de la probabilité a priori par

$$\widehat{\pi}^{(q+1)}_k = \frac{1}{n}\sum_{j=1}^{n} \tau^{(q)}_{kj}$$

et les estimateurs de $\Sigma_{11,k}, \Sigma_{22,k}$ et $\Sigma^2_{12,k}$, k=1,...,K, qui sont les solutions des trois équations suivantes :

$$\sum_{j=1}^{n^{(k)}} \tau^{(q)}_{k,j}\Big[-s + \frac{\Sigma_{22,k}X^j_{1,k}}{\Sigma_{1.2,k}} - \frac{X^j_{2,k}}{\Sigma_{22,k}} + \frac{\Sigma_{11,k}X^j_{2,k}}{\Sigma_{1.2,k}} - 2X^j_{1,k}X^j_{2,k}\frac{\Sigma_{22,k}\Sigma^2_{12,k}}{\Sigma^2_{1.2,k}}\frac{f_{s+1}(\delta_k X^j_{1,k}X^j_{2,k})}{f_s(\delta_k X^j_{1,k}X^j_{2,k})}\Big] = 0, \tag{3.10}$$

$$\sum_{j=1}^{n^{(k)}} \tau^{(q)}_{k,j}\Big[-s + \frac{\Sigma_{11,k}X^j_{2,k}}{\Sigma_{1.2,k}} - \frac{X^j_{1,k}}{\Sigma_{11,k}} + \frac{\Sigma_{22,k}X^j_{1,k}}{\Sigma_{1.2,k}} - 2X^j_{1,k}X^j_{2,k}\frac{\Sigma_{11,k}\Sigma^2_{12,k}}{\Sigma^2_{1.2,k}}\frac{f_{s+1}(\delta_k X^j_{1,k}X^j_{2,k})}{f_s(\delta_k X^j_{1,k}X^j_{2,k})}\Big] = 0, \tag{3.11}$$

$$\sum_{j=1}^{n^{(k)}} \tau_{k,j}^{(q)} \Big[s - \frac{\Sigma_{22,k} X_{1,k}^j + \Sigma_{11,k} X_{2,k}^j}{\Sigma_{1.2,k}} + X_{1,k}^j X_{2,k}^j \frac{\Sigma_{12,k}^2 + \Sigma_{11,k}\Sigma_{22,k}}{\Sigma_{1.2,k}^2} \frac{f_{s+1}(\delta_k X_{1,k}^j X_{2,k}^j)}{f_s(\delta_k X_{1,k}^j X_{2,k}^j)} \Big] = 0, \tag{3.12}$$

avec $f_{s_1}(x) = \sum_{l=0}^{\infty} \frac{x^l}{\Gamma(s_1+l)l!}$, $\forall\ x\ \in\ \mathbb{R}$, $f_{s_+1} = f'_s$, et $\Sigma_{1.2,k} = \Sigma_{11,k}\Sigma_{22,k} - \Sigma_{12,k}^2$.

Comme les équations (3.10), (3.11) et (3.12) n'admettent pas des solutions explicites des estimateurs des paramètres $\widehat{\Sigma}_{11,k}$, $\widehat{\Sigma}_{22,k}$ et $\widehat{\Sigma}_{12,k}^2$ pour $k = 1, 2, \ldots, K$, nous appliquons la méthode numérique de Newton Raphson pour résoudre ce problème.

Exemple

Nous présentons dans la figure FIG.3.17 l'estimation d'un mélange de trois lois Gamma bivaiées de Kibble en utilisant l'algorithme EM avec trois choix d'initialisation du vecteur des paramètres $\theta_k = (\Sigma_{11,k}, \Sigma_{22,k}, \Sigma_{12,k}^2)$, pour $k = 1, 2, 3$. Le premier choix d'initialisation de θ_k est donné par la MM. Le deuxième choix d'initialisation de θ_k est donné par la MMV. Le troisième choix d'initialisation de θ_k est donné par l' EAP. Les paramètres de ce mélange et la valeur de l'EQM pour chaque paramètre sont présentés dans le tableau TAB.3.10 avec $s = 2$ pour $k = 1, 2, 3$.

Dans le tableau TAB.3.11, nous présontons l'EQM entre la densité du mélange de trois lois Gamma bivariées de Kibble avec les vrais paramètres et sa densité estimée. D'après ce tableau, nous constatons que l'EQM est minimale lorsque les paramètres du mélange MGBK sont initialisés par l'EAP dans l'algorithme EM.

Paramètres	π_1	π_2	π_3	$\Sigma_{11,1}$	$\Sigma_{11,2}$	$\Sigma_{11,3}$	$\Sigma_{22,1}$	$\Sigma_{22,2}$	$\Sigma_{22,3}$	$\Sigma^2_{12,1}$	$\Sigma^2_{12,2}$	$\Sigma^2_{12,3}$
Mélange 1	0.7	0.2	0.1	2.5	4	9	3.5	6	12	6	5	2
EQM (MM)	0.911	0.067	1.76	1.726	0.56	0.0122	1.985	0.6228	4.1468	0.955	1.8102	1.0884
EQM (MMV)	0.0689	0.0497	0.0106	1.4968	0.196	0.00863	0.957	0.0233	1.0339	0.6575	0.7902	0.05
EQM (EAP)	0.0524	0.0242	0.0054	0.014	0.0195	0.001744	0.019	0.00134	0.336	0.075	0.042	0.032
Mélange 2	0.5	0.3	0.2	4	14	11	12	22	2	2	9	5
EQM (MM)	0.66	0.482	1.438	2.08	1.66	1.0244	2.58	2.395	2.0106	2.12	2.45	0.84
EQM (MMV)	0.0458	0.0718	0.045	1.845	0.69	0.718	1.08	1.245	0.95	1.02	1.81	0.4884
EQM (EAP)	0.0033	0.0203	0.004	1.28	0.107	0.658	0.25	0.8974	0.65895	0.002	0.982	0.112
Mélange 3	0.2	0.2	0.6	8	4	5	1.3	2	2	0.2	2	5
EQM (MM)	0.17	0.27	0.0189	1.6228	0.6429	1.254	1.228	0.1468	0.585	1.0912	1.82	0.7884
EQM (MMV)	0.091	0.2482	0.01887	1.248	0.1549	1.094	1.0118	0.0982	0.055	0.912	1.392	0.577
EQM (EAP)	0.0241	0.0982	0.0098	0.846	0.10854	0.8754	0.9008	0.018	0.01829	0.092	0.812	0.0487

TAB. 3.10 – Valeurs de l'EQM pour les différents paramètres du mélange MGBK.

Mélange	Mélange 1	Mélange 2	Mélange 3
EQM (MM)	2.26×10^{-6}	4.91×10^{-4}	9.056×10^{-5}
EQM(MMV)	2.62×10^{-6}	9.34×10^{-5}	7.9109×10^{-5}
EQM(EAP)	1.02×10^{-7}	3.34×10^{-5}	8.9109×10^{-6}

TAB. 3.11 – Valeurs de l'EQM pour le mélange MGKB.

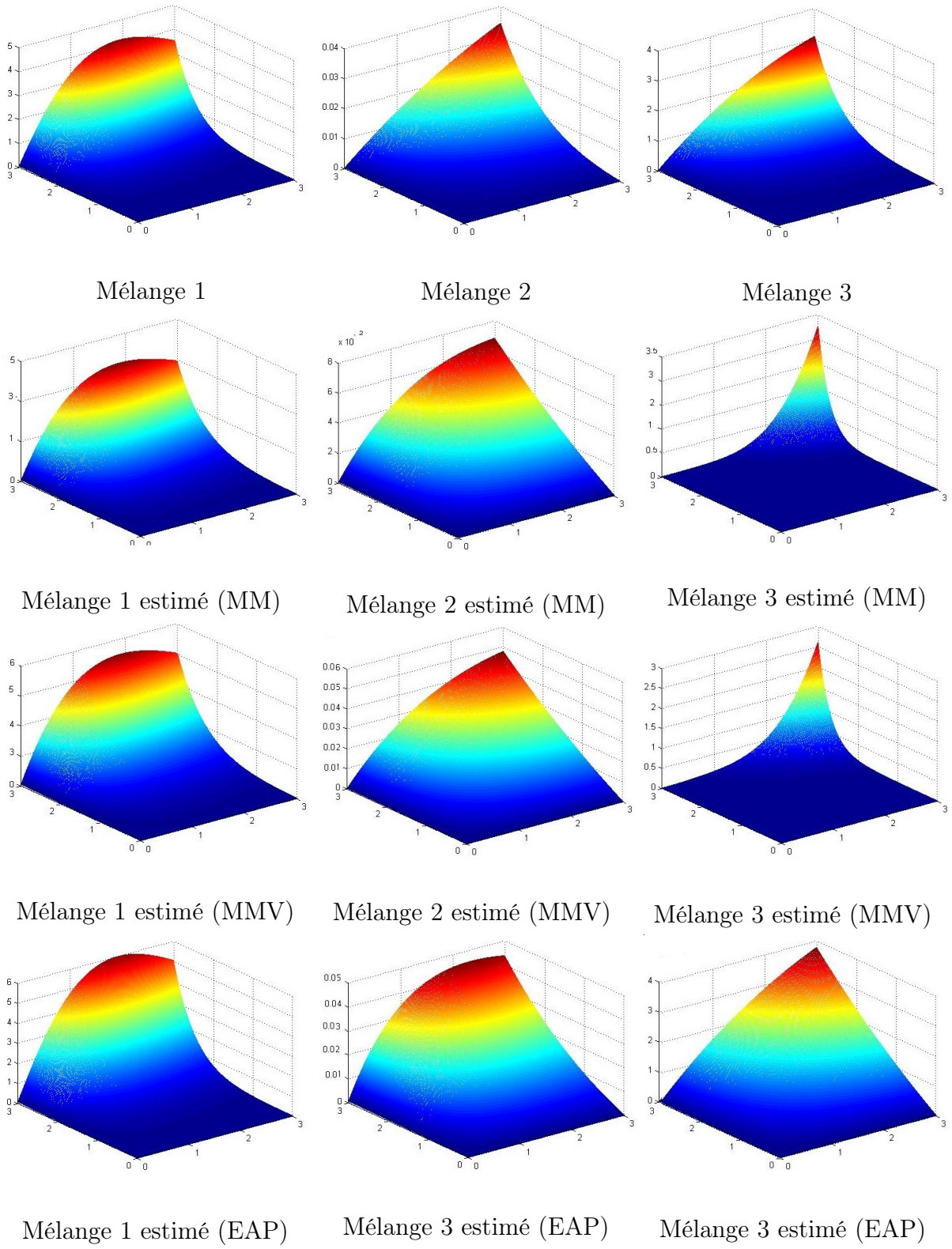


FIG. 3.16 – Estimation du mélange MGBK.

3.5.2 Estimation des paramètres du deuxième mélange MCGBKGU

Soit $(X^1, \ldots, X^n)$ un $n-$échantillon aléatoire de loi MCGBKGU de vecteur paramètres $\theta = (\Sigma_{11}, \Sigma_{22}, \Sigma_{12}^2)$. La densité du mélange MCGBKGU est donnée par :

$$\begin{aligned} f((x_1, x_2), \Phi) &= \sum_{k=1}^{K} \pi_k f((x_{1,k}, x_{2,k}); \theta_k) \\ &= \sum_{k=1}^{K} \frac{\pi_k}{\Sigma_{22,k}^{s_2-s_1}} \Big(\Sigma_{11,k}\Sigma_{22,k} - \Sigma_{12,k}^2\Big)^{-s_1} \frac{x_{1,k}^{s_1-1} x_{2,k}^{s_2-1}}{\Gamma(s_1)\Gamma(s_2)} \\ &\quad \times e^{(-\frac{\Sigma_{22,k}x_{1,k}] + \Sigma_{11,k}x_{2,k}}{\Sigma_{11,k}\Sigma_{22,k} - \Sigma_{12,k}^2})} \Phi_3\Big(s_2 - s_1; s_2; h_k x_{2,k}, \delta_k x_{1,k} x_{2,k}\Big), \end{aligned}$$

où $\delta_k = \dfrac{\Sigma_{12,k}}{\Sigma_{1.2,k}^2}$, $\Sigma_{1.2,k} = \Sigma_{11,k}\Sigma_{22,k} - \Sigma_{12,k}^2$, $h_k = \dfrac{\Sigma_{12,k}^2}{\Sigma_{22,k}(\Sigma_{11,k}\Sigma_{22,k} - \Sigma_{12,k}^2)^2}$, π_k la probabilité a priori, $k = 1, \ldots, K$, et $\theta_k = (\Sigma_{11,k}, \Sigma_{22,k}, \Sigma_{12,k}^2)$.

Pour estimer le vecteur des paramètres du mélange

$\Phi = (\pi_1, \ldots, \pi_K, \Sigma_{11,1,}, \ldots, \Sigma_{11,K}, \Sigma_{22,1}, \ldots, \Sigma_{22,K}, \Sigma_{12,1}^2, \ldots, \Sigma_{12,K}^2)$, nous appliquons l'algorithme EM.

À l'étape **E** de l'itération (q), la probabilité a postériori $\tau_{kj}^{(q)}$ est définie par

$$\tau_{kj}^{(q)} = \frac{\pi_k^{(q)} f((x_{1,k}, x_{2,k}); \theta_k^{(q)})}{\sum_{k=1}^{K} \pi_k^{(q)} f((x_{1,k}, x_{2,k}); \theta_k^{(q)})}.$$

À l'étape **M** de l'itération $(q+1)$, l'estimateur $\widehat{\Phi}^{(q+1)}$ du vecteur des paramètres Φ est défini par $\widehat{\Phi}^{(q+1)} = \arg\max_{\Phi} Q(\Phi, \Phi^{(q)})$ avec

$$\begin{aligned} Q(\Phi, \Phi^{(q)}) &= \sum_{j=1}^{n}\sum_{k=1}^{K} t_{kj}^{(q)} \ln(\pi_k f((x_{1,k}^j, x_{2,k}^j); \theta_k) \\ &= \sum_{j=1}^{n}\sum_{k=1}^{K} t_{kj}[\ln(\pi_k) - s_1 \ln \Sigma_{1.2,k} + (s_1 - s_2)\ln(\Sigma_{22}) - \frac{\Sigma_{22,k}x_{1,k}^j}{\Sigma_{1.2,k}} \\ &\quad +(s_1 - 1)\ln(x_{1,k}^j) - \ln\Gamma(s_1) - \frac{\Sigma_{11}x_{2,k}^j}{\Sigma_{1.2,k}} + (s_2 - 1)\ln(x_{2,k}^i) - \ln\Gamma(s_2) \\ &\quad + \ln(\Phi_3(s_2 - s_1; s_2; h_k x_{2,k}^i, \delta x_{1,k}^i x_{2,k}^i). \end{aligned}$$

Ceci, nous permet de déterminer l'estimateur de la probabilité a priori

$$\widehat{\pi}_k^{(q+1)} = \frac{\sum_{j=1}^{n} \tau_{kj}^{(q)}}{n}$$

et les estimateurs de $\Sigma_{11,k}, \Sigma_{22,k}$ et $\Sigma_{12,k}^2$, k=1,...,K, qui sont les solutions des trois équations suivantes :

$$\sum_{j=1}^{n^{(k)}} \tau_{k,j}^{(q)} \Big[-s_1 + \frac{\Sigma_{22,k} X_{1,k}^j}{\Sigma_{1.2,k}} + \frac{\Sigma_{11,k} X_{2,k}^j}{\Sigma_{1.2,k}} - \frac{X_{2,k}}{\Sigma_{22,k}} - \frac{(s_2 - s_1)\Sigma_{12,k}^2 X_{1,k}^j}{s_2 \Sigma_{22,k} \Sigma_{1.2,k}} \times \frac{\Phi_3\Big(s_2 - s_1 + 1; s_2 + 1; h_k X_{2,k}^j, \delta_k X_{1,k}^j X_{2,k}^j\Big)}{\Phi_3\Big(s_2 - s_1; s_2; h_k X_{2,k}^j, \delta_k X_{1,k}^j X_{2,k}^j\Big)} - \frac{2\Sigma_{12,k}^2 X_{1,k}^j X_{2,k}^j}{s_2 \Sigma_{1.2,k}^2} \times \frac{\Phi_3\Big(s_2 - s_1; s_2 + 1; h_k X_{2,k}^i, \delta X_{1,k}^j X_{2,k}^j\Big)}{\Phi_3\Big(s_2 - s_1; s_2; h_k X_{2,k}^j, \delta_k X_{1,k}^k X_{2,k}^j\Big)} \Big] = 0, \quad (3.13)$$

$$\sum_{j=1}^{n^{(k)}} \tau_{k,j}^{(q)} \Big[-s_1 + \frac{(s_1 - s_2)\Sigma_{12,k}^2}{\Sigma_{11,k}\Sigma_{22,k}} + \frac{\Sigma_{22,k} X_{2,k}^j}{\Sigma_{1.2,k}} + \frac{\Sigma_{11,k} X_{2,k}^j}{\Sigma_{1.2,k}} - \frac{X_{1,k}^j}{\Sigma_{11,k}} - \frac{(s_2 - s_1) X_{1,k}^j \Sigma_{12,k}^2 (\Sigma_{11,k}\Sigma_{22,k} + \Sigma_{1.2,k}) \Phi_3\Big(s_2 - s_1 + 1; s_2 + 1; h_k X_{2,k}^j, \delta X_{1,k}^j X_{2,k}^j\Big)}{s_2 \Sigma_{11,k} \Sigma_{22,k}^2 \Sigma_{1.2,k} \Phi_3\Big(s_2 - s_1; s_2; h_k X_{2,k}^j, \delta_k X_{1,k}^j X_{2,k}^j\Big)} - \frac{2\Sigma_{12,k}^2}{s_2 \Sigma_{1.2,k}^2} X_{1,k}^j X_{2,k}^j \frac{\Phi_3\Big(s_2 - s_1; s_2 + 1; h_k X_{2,k}^i, \delta X_{1,k}^j X_{2,k}^j\Big)}{\Phi_3\Big(s_2 - s_1; s_2; h_k X_{2,k}^j, \delta_k X_{1,k}^k X_{2,k}^j\Big)} \Big] = 0, \quad (3.14)$$

$$\sum_{j=1}^{n^{(k)}} \tau_{k,j}^{(q)} \Big[s_1 - \frac{\Sigma_{22,k} X_{1,k}}{\Sigma_{1.2,k}} - \frac{\Sigma_{11,k} X_{2,k}}{\Sigma_{1.2,k}} + \frac{(s_2 - s_1)\Sigma_{11,k} X_{1,k}^j \Phi_3\Big(s_2 - s_1 + 1; s_2 + 1; h_k X_{2,k}^j, \delta X_{1,k}^j X_{2,k}^j\Big)}{s_2 \Sigma_{1.2} \Phi_3\Big(s_2 - s_1; s_2; h_k X_{2,k}^j, \delta_k X_{1,k}^j X_{2,k}^j\Big)} + \frac{\Sigma_{11,k}\Sigma_{22,k} + \Sigma_{12,k}^2}{s_2 \Sigma_{1.2,k}^2} X_{1,k}^j X_{2,k}^j \frac{\Phi_3\Big(s_2 - s_1; s_2 + 1; h_k X_{2,k}^i, \delta X_{1,k}^j X_{2,k}^j\Big)}{\Phi_3\Big(s_2 - s_1; s_2; h_k X_{2,k}^j, \delta_k X_{1,k}^k X_{2,k}^j\Big)} \Big] = 0. \quad (3.15)$$

Comme les équations (3.13), (3.14) et (3.15) n'admettent pas des solutions explicites des estimateurs des paramètres $\widehat{\Sigma}_{11,k}$, $\widehat{\Sigma}_{22,k}$ et $\widehat{\Sigma}_{12,k}^2$ pour $k = 1, 2, \ldots, K$, nous appliquons la méthode numérique de Newton Raphson pour résoudre ce problème.

Exemple

Nous présentons dans la figure FIG.3.17 l'estimation d'un mélange de trois lois MCGBKGU en utilisant l'algorithme EM avec trois choix d'initialisation du vecteur des paramètres $\theta_k = (\Sigma_{11,k}, \Sigma_{22,k}, \Sigma^2_{12,k})$, pour $k = 1, 2, 3$. Le premier choix d'initialisation de θ_k est donné par la MM. Le deuxième choix d'initialisation de θ_k est donné par la MMV. Le troisième choix d'initialisation de θ_k est donné par l' EAP. Les paramètres de ce mélange et la valeur de l'EQM pour chaque paramètre sont présentés dans le tableau TAB.3.12 avec $s_1 = 3$ et $s_2 = 4$ pour $k = 1, 2, 3$.

Dans le tableau TAB.3.13, nous présontons l'EQM entre la densité du mélange de trois lois bivariées du MCGBKGU avec les vrais paramètres et sa densité estimée. D'après ce tableau, nous constatons que l'EQM est minimale lorsque les paramètres du mélange MGBK sont initialisés par l'EAP dans l'algorithme EM.

Paramètres	π_1	π_2	π_3	$\Sigma_{11,1}$	$\Sigma_{11,2}$	$\Sigma_{11,3}$	$\Sigma_{22,1}$	$\Sigma_{22,2}$	$\Sigma_{22,3}$	$\Sigma^2_{12,1}$	$\Sigma^2_{12,2}$	$\Sigma^2_{12,3}$
Mélange 1	0.7	0.2	0.1	2.5	4	9	3.5	6	12	6	5	2
EQM (MM)	0.891	0.064	0.981	1.025	0.576	0.0722	0.9445	0.6228	1.18	0.55	0.812	0.884
EQM (MMV)	0.881	0.0066	0.694	1.0054	0.1296	0.026	0.911	0.0233	1.03	0.0578	0.592	0.05
EQM (EAP)	0.451	0.0041	0.00101	0.0021	0.025	0.00794	1.9×10^{-3}	0.0034	0.036	2.1×10^{-3}	2.2×10^{-5}	0.0215
Mélange 2	0.5	0.3	0.2	4	14	11	12	22	2	2	9	5
EQM (MM)	0.98	0.7	1.784	1.99	1.52	1	1.78	2.22	1.952	2.65	2.548	0.9199
EQM (MMV)	0.054	0.44	0.15	1.245	0.989	0.0998	1.2	0.92	0.094	1.001	0.608	0.336
EQM (EAP)	0.001	0.0034	0.0056	0.98	0.1902	0.065	0.52	0.444	0.033	0.0016	0.427	0.0257
Mélange 3	0.2	0.2	0.6	8	4	5	1.3	2	2	0.2	2	5
EQM (MM)	0.9081	0.22	0.9837	1.8331	1.698	1.1177	2.042	1.68	1.85	1.4512	1.192	1.71
EQM (MMV)	0.701	0.1919	0.5437	1.6228	1.6429	1.004	1.6228	1.108	1.751	1.0912	1.08102	1.08
EQM (EAP)	0.401	0.1004	0.1437	1.284	0.991	0.894	1.07	1.048	1.0055	0.1912	0.11	0.67

TAB. 3.12 – Valeurs de l'EQM pour les différents paramètres du mélange MCGBKGU.

Mélange	Mélange 1	Mélange 2	Mélange 3
EQM (MM)	0.9894	1.91	0.8825
EQM(MMV)	0.742	0.934	0.793
EQM(EAP)	0.102	3.34×10^{-2}	0.008

TAB. 3.13 – Valeurs de l'EQM pour le mélange MCGBKBU.

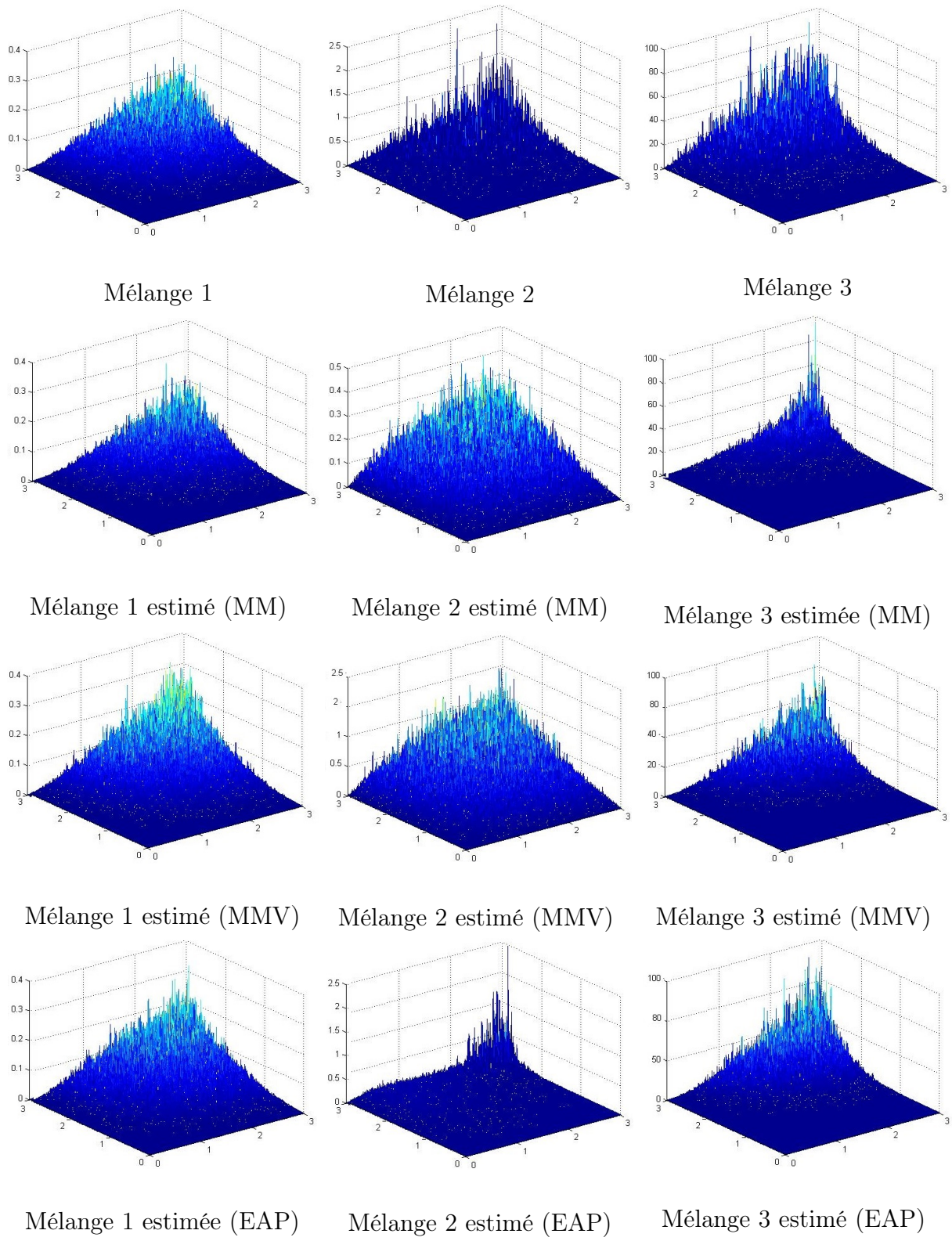


FIG. 3.17 – Estimation de mélange MCGBKBU.

3.6 Segmentation d'images couleur

L'emploi de la couleur en segmentation d'images est un sujet de recherche relativement récent. Bien que l'on trouve plusieurs sortes d'algorithmes de segmentation d'images couleur, la littérature n'est pas aussi riche que pour la segmentation des images en niveaux de gris. La couleur étant une information que l'on attache directement et naturellement aux régions, une segmentation couleur devrait permettre d'obtenir des régions qui sont davantage significatives. La segmentation couleur pose des problèmes relatifs au traitement de l'information couleur contenue dans les images. Si le système de représentation RVB est généralement la base de l'acquisition des images couleur, il n'est pas forcément le mieux adapté à l'analyse. En 1955, Thierry Carron [23] a adapté certain nombre de transformations de l'espace RVB initial en un nouvel espace de représentation tel que l'espace TLS. Il a montré que la segmentation des images couleur présentées dans cet espace est la plus appropriée dont laquel l'information couleur est représentée par les valeurs de la Teinte (T) et de la Saturation (S). Donc, la perception humaine de l'image peut être caractérisée par un vecteur aléatoire bivarié qui concerne la Teinte et la Saturation. Il est à noter également que cette transformation est inversible. Ceci va faciliter les phases d'élaboration des algorithmes de segmentation en permettant la synthèse des images dans la base TLS. Dans la sgmentation d'images, nous considérons que toute l'image est caractérisée par un modèle de mélange fini. Dans la littérature, plusieurs auteurs ont supposé que la composante (T, S) suit une loi Gaussienne bivariée [84], ou bien une loi Bêta bivariée de première espèce [53] ou bien une loi Bêta bivariée de deuxième espèce [13].

Puisque la composante (T, S) ne prend pas des valeurs négatives, et comme la loi de la

diagonale modifiée de Riesz ne prend que des valeurs dans l'intervalle $]0,+\infty[\times]0,+\infty[$, alors dans ce chapitre, nous proposons que la composante (T,S) suivent cette loi.

L'algorithme de la segmentation proposé dans ce chapitre est composé des différentes étapes suivantes :

• Conversion de l'image couleur RVB en une image de type TLS.

• Initialisation et régularisation des différents attributs de l'image (classes, moyenne, variance,...), formée par les deux composantes (T,S), assurée par l'utilisation de l'histogramme $2D$ de l'image (T,S) et l'algorithme K-means. La probabilité a priori π_k pour chaque classe k, $k=1,\ldots,K$, est choisie $\frac{1}{K}$.

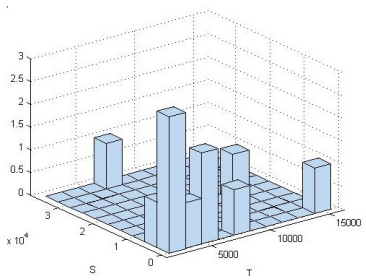
• Estimation des différents paramètres de chaque classe en utilisant l'algorithme EM dans le cas des lois bivariées.

• Segmentation statistique, en utilisant la règle de Bayes.

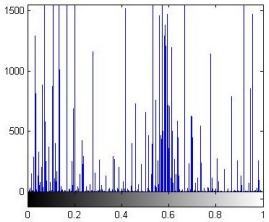
• Restauration de l'image segmentée dans l'espace TLS.

• Conversion de l'image segmentée TLS en une image segmentée RVB.

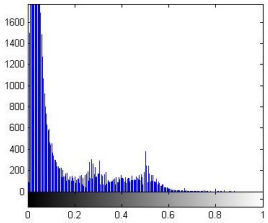
Nous présentons dans la suite quelques résultats de segmentation obtenus par la méthode proposée sur deux images couleur satellitaires notées Sat 1 et Sat 2 (voir FIG.3.22.a et FIG.3.23.a). Les images utilisées sont des images Spot de taille $256\times256\times3$ pixels codées en RVB. La conversion de ces deux images en TLS est représentée dans FIG.3.22.b et FIG.3.23.d. Dans la figure FIG.3.18, nous présontons l'histogrammes $2D$ en composantes (T,S) et les histogrammes $1D$ de la composante T et de la composante S de l'image Sat 1. Dans la figure FIG.3.19, nous présontons l'histogrammes $2D$ en composantes (T,S) et les histogrammes $1D$ de la composante T et de la composante S de l'image Sat 2.



Composantes (T, S).

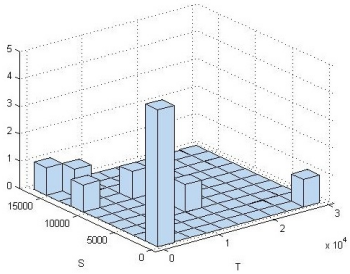


Composante T.

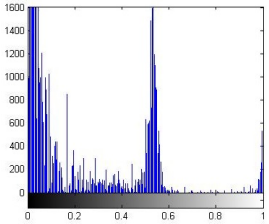


Composante S.

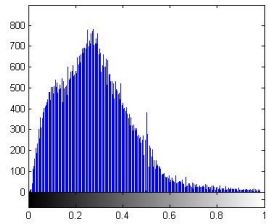
FIG. 3.18 – Histogrammes de l'image Sat 1.



Composantes (T, S).



Composante T.



Composante S.

FIG. 3.19 – Histogramme de l'image Sat 2.

Initialisation

Selon les histogrammes présentés dans les figures FIG 3.18 et FIG 3.19, nous choisissons le nombre de classes $K = 3$ et nous appliquons l'algorithme $K-$means pour initialiser les paramètres du mélange pour chaque classe. Ensuite, dans un premier temps, nous supposons que les composantes (T, S) de chaque classe sont représentées par des lois Gaussiennes bivariées. Dans un deuxième temps, nous supposons que les composantes (T, S) de chaque classe sont représentées par des lois de la diagonale modifiée de Riesz.

Estimation des paramètres du mélange

Nous utilisons l'algorithme EM pour estimer les paramètres du mélange pour ces deux cas. Dans le cas d'un mélange Gaussien, nous avons initialisé les paramètres par l'utilisation de la MM. Dans le cas du mélange des lois de la diagonale modifiée de Riesz, nous avons initialisé les paramètres par l'utilisation de l'EAP.

Nous présentons respectivement dans les tableaux TAB.3.14 et TAB.3.15 les valeurs des paramètres initiales et estimées par l'algorithme EM après 10 iterations pour chaque classe des deux images Sat 1 et Sat 2.

Pour l'image Sat 1, les paramètres s_1 et s_2 des lois MGBK et MCGBKGU sont fixés arbitrairement par :

- Classe 1/MGBK : $s_1 = s_2 = 7$, Classe 2/MGBK : $s_1 = s_2 = 8$, Classe 3/MGBK : $s_1 = s_2 = 4$,
- Classe 1/MCGBKGU : $s_1 = 7$ et $s_2 = 18$, Classe 2/MCGBKGU : $s_1 = 8$ et $s_2 = 18$ Classe 3/MCGBKGU : $s_1 = 4$ et $s_2 = 18$.

Pour l'image Sat 2, les paramètres s_1 et s_2 de la loi de la diagonale modifiée des lois MGBK et MCGBKGU sont fixés arbitrairement par :

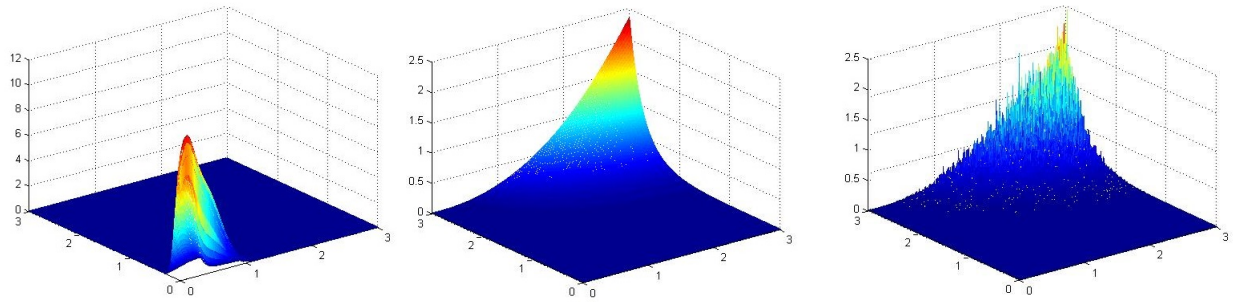
– Classe 1/MGBK : $s_1 = s_2 = 5$, Classe 2/MGBK : $s_1 = s_2 = 2$, Classe 3/MGBK : $s_1 = s_2 = 9$,

– Classe 1/MCGBKGU : $s_1 = 5$ et $s_2 = 10$, Classe 2/MCGBKGU : $s_1 = 2$ et $s_2 = 10$ Classe 3/MCGBKGU : $s_1 = 9$ et $s_2 = 18$.

Nous présontons dans les figures FIG.3.20 et FIG.3.21 l'estimation de la densité de la loi mélange pour chaque image.

Paramètre	Valeur initiale à partir de K-means			Valeur estimée par l'EM		
	$classe1$	$classe2$	$classe3$	$classe1$	$classe2$	$classe3$
MGAUB						
π	1/3	1/3	1/3	0.9363	0.0038	0.0017
μ_1	0.0152	0.2642	0.1878	0.0091	0.3657	0.1106
μ_2	0.0175	0.0857	0.0477	0.0162	0.0607	0.0001
σ_1	0.0013	0.0596	0.1178	0.6613	0.0482	0.0675
σ_2	0.0005	0.0083	0.0217	0.4483	0.0018	10^{-4}
ρ	0.0003	0.0016	0.0278	0.1807	0.0001	10^{-5}
MGBK						
π	1/3	1/3	1/3	0.0819	0.5768	0.3413
Σ_{11}	0.17	0.0528	0.0939	1.1879	2.3531	1.4263
Σ_{22}	0.19	0.0171	0.0239	0.0031	0.0354	0.0332
Σ_{12}^2	0.003	0.0003	0.0014	0.0027359	0.0010	0.0290
MCGBKGU						
π	1/3	1/3	1/3	0.325	0.5326	0.1424
Σ_{11}	0.012	0.0441	0.0247	0.0016	0.0082	1.4701
Σ_{22}	0.0037	0.0016	0.0037	0.0525	0.0172	0.0033
Σ_{12}^2	20.11×10^{-5}	1.598×10^{-4}	0.0096	2.1682	2.2646	4.5947

TAB. 3.14 – Valeurs des paramètres du mélange de l'image Sat 1.

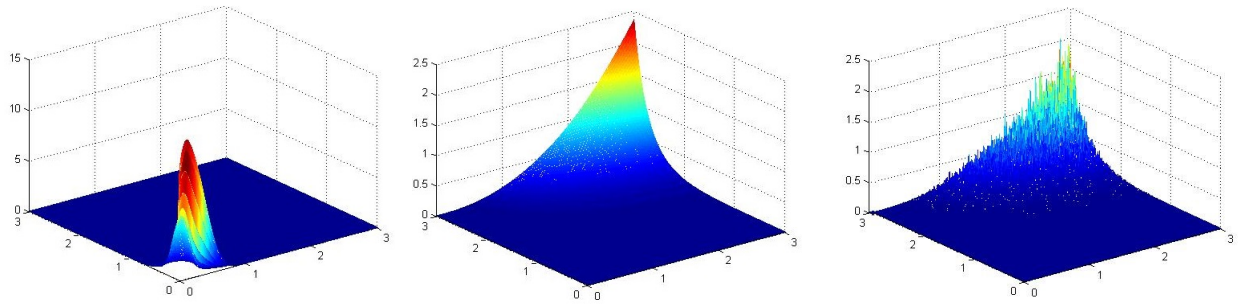


MGAUB MGBK MCGBKGU

FIG. 3.20 – Estimation de la densité de la loi mélange pour l'image Sat 1.

Paramètre	Valeur initiale à partir de K-means			Valeur estimé par l'EM		
	*classe*1	*classe*2	*classe*3	*classe*1	*classe*2	*classe*3
MGAUB						
π	1/3	1/3	1/3	0.9295	0.0314	0.0392
μ_1	0.0131	0.1808	0.085	0.011	0.0998	0.0001
μ_2	0.026	0.1941	0.0996	0.0225	0.2174	0.78
σ_1	0.004	0.0443	0.075	0.0003	0.0147	0.001
σ_2	0.0025	0.0185	0.0437	0.0021	0.0092	0.3367
ρ	0.0005	0.007	0.0349	0.0005	0.0003	10^{-5}
MGBK						
π	1/3	1/3	1/3	10^{-5}	0.6895	0.3105
Σ_{11}	0.015	0.0452	0.0425	6.4627	1.1649	1.6137
Σ_{22}	0.29	0.0442	0.0498	0.0034	0.0105	0.0521
Σ_{12}^2	0.0005	0.0018	0.0017	0.0021	0.0013	0.0487
MCGBKGU						
π	1/3	1/3	1/3	0.3921	0.6061	0.0018
Σ_{11}	0.0134	0.0916	0.0092	0.063	0.1215	0.0045
Σ_{22}	0.0092	7.756×10^{-4}	0.0201	0.007	0.0011	0.078
Σ_{12}^2	7.3505×10^{-5}	4.61×10^{-5}	0.0018	5.3608	0.8577	1.9893

TAB. 3.15 – Valeurs des paramètres du mélange de l'image Sat 2.



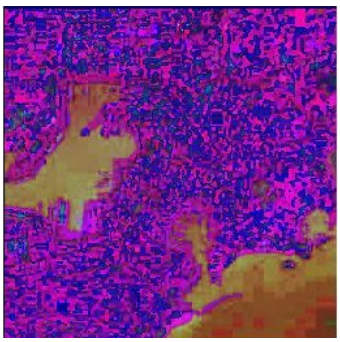
MGAUB MGBK MCGBKGU

FIG. 3.21 – Estimation de la densité de la loi mélange pour l'image Sat 2.

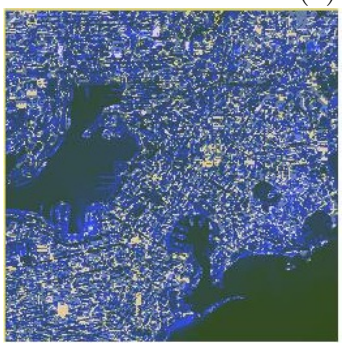
Segmentation statistique des images couleur

Une fois les paramètres du mélange dans les différents cas du modèle (MGAUB, MGBK et MCGBK) de chaque image sont estimés, nous appliquons la règle de classification de Bayes sur ces deux images dans la base (T, S). Les résultats obtenus de la segmentation de ces deux images sont présentés dans l'espace TLS (voir FIG.3.22.c, FIG.3.22.d, FIG.3.22.e, FIG.3.23.c, FIG.3.23.d et FIG.3.23.e). Ensuite, nous représentons les images segmentées dans l'espace RVB (voir FIG.3.22.f, FIG.3.22.g, FIG.3.22.h, FIG.3.23.f, FIG.3.23.g et FIG.3.23.h).

(a) RVB

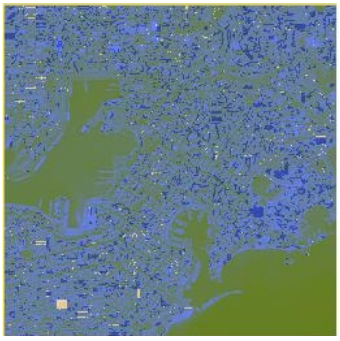


(b) TLS



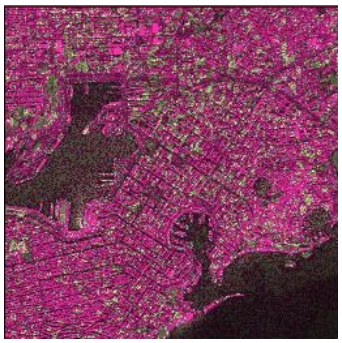
(c) MGBK

(d) MCGBKGU



(e) MGAUB

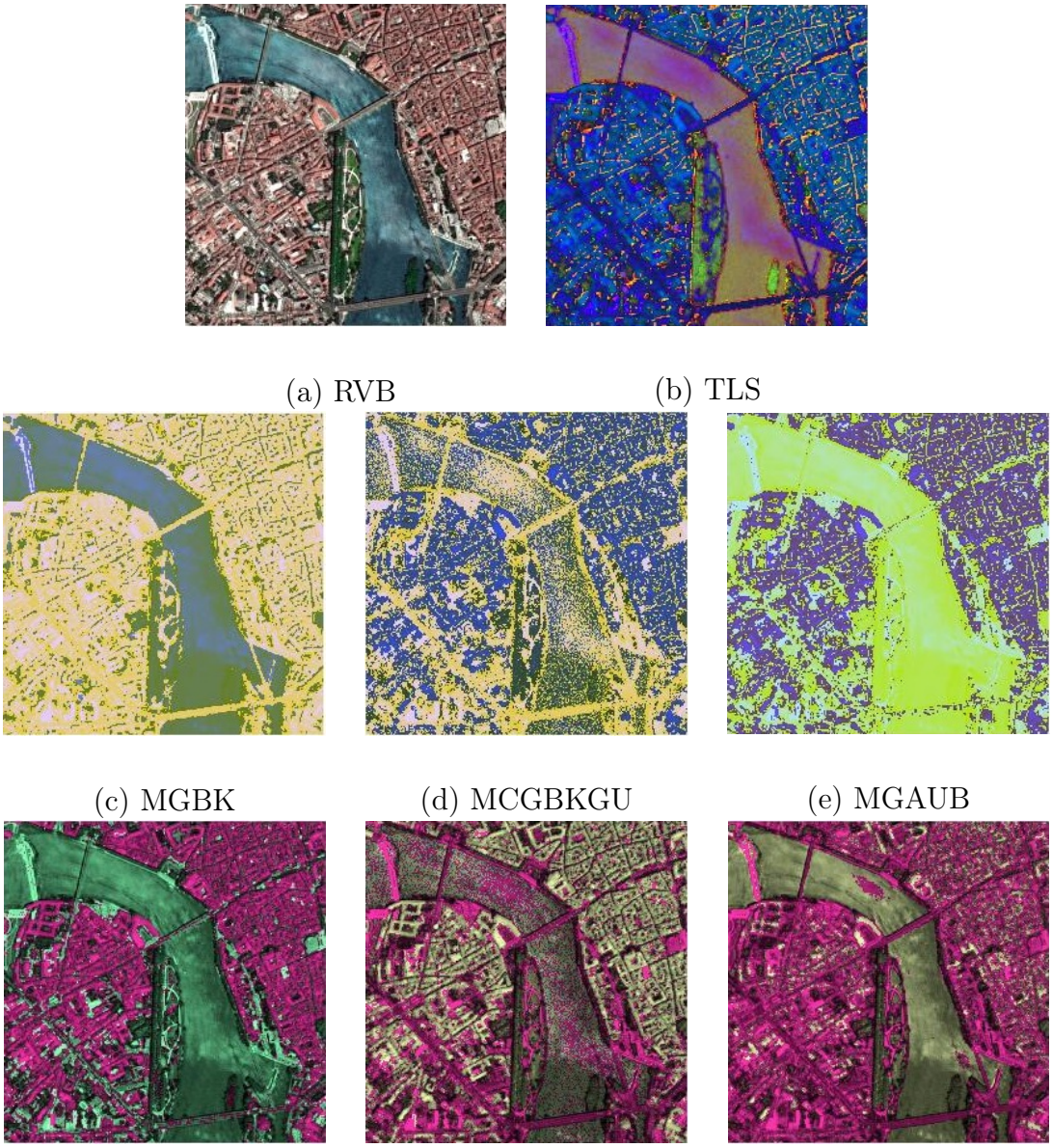
(f) MGBK



(g) MCGBKGU

(h) MGAUB

FIG. 3.22 – Résultats de la segmentation de l'image Sat 1.



(a) RVB (b) TLS

(c) MGBK (d) MCGBKGU (e) MGAUB

(f) MGBK (g) MCGBKGU (h) MGAUB

FIG. 3.23 – Résultats de la segmentation de l'image Sat 2.

Evaluation de la performance

L'évaluation de la performance de l'image restaurée en RVB est faite par des tests subjectifs ou bien des tests objectifs [35] de la qualité de l'image. Les méthodes de test objectif de la qualité de l'image sont souvent utilisés, car les résultats numériques d'une mesure objective sont facilement calculés et permettent une comparaison de la consistance des différents algorithmes. Il existe plusieurs mesures de la qualité d'image disponibles pour l'évaluation de la performance de la méthode de la segmentation d'images. Une étude approfondie des mesures de la qualité est donnée par [35]. Pour l'évaluation de la performance de l'algorithme de segmentation développé, nous considérons les mesures de qualité suivantesde :

- La Distance Maximale (DM), qui est définie par $DM = \max_{i,j}[I(i,j) - I_s(i,j)]$, avec $I(i,j)$ représente la valeur du pixel de l'image originale I et $I_s(i,j)$ représente la valeur du pixel de l'image I_s segmentée. Si la valeur de la DM est grande, alors on a un nombre de pixels mal classés est assez élevé. Ceci introduit une mauvaise segmentation.

- L'EQM, c'est l'eurrer quadratique moyenne qui est définie pour deux image I et I_s de tailles $m \times n$ par $\frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\|I(i,j) - I_s(i,j)\|^2$.

- Peak Signal to Noise Ratio (PSNR), qui est une mesure de distoration utilisée en imagerie. Il s'agit de qualifier la performance des codeurs en mesurant la qualité de reconstruction de l'image par rapport à l'image originale. Cette mesure est définie par $PSNR = 10\ln_{10}(\frac{255^2}{EQM})$. Si la valeur de la PSNR est petite, alors on a un nombre de pixels mal classés est assez élevé. Ceci introduit une mauvaise segmentation.

Afin d'évaluer la performance de l'algorithme de segmentation proposé pour l'image Sat 1 et l'image Sat 2 codées en RVB, nous calculons la DM, l'EQM et PSNR. Dans le tableau

TAB.3.16 ; nous donnons les valeurs de ces mesures.

Image Sat 1	MGBK	MCGBKGU	MGAUB
DM	0.5499	0.74512	0.8915
EQM	0.4101	0.551	0.8402
PSNR	8.9124	5.8919	1.7417
Image Sat 2	MGBK	MCGBKGU	MGAUB
DM	0.6194	0.7019	0.9313
EQM	0.5928	0.6892	0.9516
PSNR	5.6194	3.4476	0.9313

TAB. 3.16 – Valeurs de la DM, l'EQM et PSNR pour les images Sat 1 et Sat 2.

D'après les résultats présentés dans ce tableau, nous remarquons que pour les différents critères d'évaluation adoptés pour chaque image couleur satellitaire codée en RVB (Sat 1 et Sat 2) l'algorithme de segmentation proposé dans le cas d'un mélange de la diagonale modifiée des lois de Riesz est plus performant que celui qui utilise un mélange Gaussien bivarié.

3.7 Conclusion

Dans ce chapitre, nous avons proposé une méthode de segmentation statistique adaptée à des images couleur en utilisant le mélange de lois de la diagonale modifiée de Riesz. La segmentation a été réalisée après avoir estimer les paramètres du mélange de l'image couleur caractérisée par la composante (T, S). Ces paramètres ont été estimés par l'utilisation

de l'algorithme EM dans un cadre bivarié. L'initialisation de cet algorithme est basée sur l'utilisation de l'algorithme K-means et l'EAP. L'utilisation du modèle mélange proposé en segmentation a permis d'amiliorer la segmentation des images satellitaires que lorsque nous utilisons le modèle de mélange Gaussien.

Conclusion et perspectives

Ce mémoire évalue l'apport des familles exponentielles multivariées dans le domaine de traitement du signal et des images couleur. Dans un premier temps, nous avons traité le problème de la caractérisation des familles exponentielles naturelles infiniment divisibles, dans le cadre multidimensionnel. Après avoir introduit les principales définitions et propriétés des familles exponentielles infiniment divisibles, nous avons proposé une nouvelle caractérisation des familles exponentielles infiniment divisibles basée sur la fonction trace de la matrice de variance-covariance associée. Cette proposition a permis d'identifier la loi du bruit additif associé à un signal ou à une image.

Dans un deuxième temps, nous nous sommes intéressés au problème de l'estimation paramétrique pour des modèles multivariés qui appartiennent à la classe des familles exponentielles. D'abord, nous avons rappelé quelques lois multivariées ainsi que l'estimation de leurs paramètres en utilisant la méthode des moments et la méthode du maximum de vraisemblance. Ensuite, nous avons proposé un nouveau modèle bivarié basé sur la loi de Riesz. Enfin, nous avons fait une étude comparative des différentes méthodes d'estimations des paramètres de ce nouveau modèle.

Dans un troisième temps, nous avons consacré notre travail à la segmentation statistique

des images couleur. Ces images couleur ont été traités dans l'espace TLS où nous n'avons travaillé qu'avec les deux composantes (T, S). Dans ce cas, nous avons introduit une nouvelle modélisation de l'image dans cette base (T, S) par un mélange de lois de la diagonale modifiée de Riesz. Pour estimer les paramètres de ce mélange, nous avons appliqué l'algorithme EM. L'initialisation de cet algorithme est basée sur l'utilisation de l'algorithme $K-$means et l'EAP. Les résultats de segmentation obtenus par cette nouvelle modélisation qui est basée sur la loi de Riesz ont permis d'obtenir une qualité de segmentation supérieure à celle obtenue par un modèle de mélange Gaussien.

Différentes perspectives se dégagent de ces traveaux. Premièrement, appliquer la fonction trace au problème de la réduction de la dimension de l'espace. Deuxièmement, il est possible d'étendre le modèle de la diagonale modifiée de Riesz pour le cas $r > 2$. Troisièmement, le problème de la segmentation de l'image couleur dans la base des trois composantes (T, S, L) pourra être traité. Enfin, nous pouvons introduire la méthode bootstrap dans le cadre de la segmentation multivariée.

Annexe A

Dans cet Annexe, nous présentons quelques lois undimensionnelles qui appartiennent à des FEG, ainsi que leurs propriétés.

Loi Gamma

Dans la théorie des probabilités et des statistiques, une distribution Gamma, ou loi Gamma, est un type de lois de probabilité de variables aléatoires réelles positives. La famille des distributions Gamma inclut entre autres les lois exponentielles, la loi de la somme des variables aléatoires indépendantes suivant une même loi exponentielle, ainsi que la loi du χ^2. La loi Gamma permet donc de modéliser une grande variété de phénomènes pour des grandeurs positives.

Définition.1 Une variable aléatoire X suit une loi Gamma de paramètres $(a, \lambda) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$, notée $\gamma(a, \lambda)$, si sa densité de probabilité peut se mettre sous la forme

$$f_X(x) = \frac{\lambda^a}{\Gamma(a)} e^{-\lambda x} x^{a-1} \mathbf{1}_{\mathbb{R}_+}(x)$$

avec $\Gamma(a) = \int_0^{+\infty} \exp(-x) x^{a-1} dx$, λ est un paramètre d'échelle et a est un paramètre de forme.

Le paramètre a influe sur la forme de la distribution (voir FIG.1.5).

Remarques.1.

1) Si le paramètre d'échelle prend la valeur $\lambda = 1$, alors la loi Gamma est dite standard (notée $\gamma(a)$).

2) Si le paramètre de forme prend la valeur $a = 1$, alors la loi Gamma est dite la loi exponentielle de paramètre λ.

Moments

Si X suit la loi $\gamma(a, \lambda)$, alors

$$\mathbb{E}(X^k) = \frac{\Gamma(a+k)}{\lambda^a \Gamma(a)}.$$

Nous en déduisons que :

$$\mathbb{E}(X) = \frac{a}{\lambda} \quad \text{et} \quad \mathbb{V}ar(X) = \frac{a}{\lambda^2}.$$

Dans le cas d'une loi exponentielle de paramètre λ, nous avons

$$\mathbb{E}(X) = \frac{1}{\lambda} \quad \text{et} \quad \mathbb{V}ar(X) = \frac{1}{\lambda^2}.$$

Le paramètre λ est donc interprétable comme l'inverse de la valeur moyenne de X ou l'inverse de son écart-type. Ceci donne naissance à une forme légèrement différente de la densité paramétrisée par la moyenne ou l'écart-type λ, définie par

$$f_X(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \mathbf{1}_{\mathbb{R}_+}(x),$$

ou plus généralement

$$\widetilde{f}_X(x) = \frac{1}{\lambda^a \Gamma(a)} e^{-\frac{x}{\lambda}} x^{a-1} \mathbf{1}_{\mathbb{R}_+}(x).$$

La forme de la densité de la variable aléatoire X varie selon les paramètres a et λ comme le montre la figure FIG.1.5.

Dans tout ce qui suit, la loi Gamma $\gamma(a, \lambda)$ a pour densité de probabilité $\widetilde{f}$.

Loi Normale

La loi de probabilité la plus utilisée en statistique est la loi Normale, encore appelée loi de Gauss ou de Laplace-Gauss.

Définition.2 Une variable aléatoire X suit la loi Normale $N(m, \sigma^2)$ dans $\mathbb{R}$ si sa densité de probabilité a pour forme :

$$f_X(x) = \frac{1}{\sqrt{2\pi}\ \sigma} \exp\Big(- \frac{(x-m)^2}{2\sigma^2}\Big),\ \forall\ x \in \mathbb{R},$$

où $m \in \mathbb{R}$ est la moyenne et $\sigma^2 > 0$ la variance de X.

Dans le cas où les paramètres $m = 0$ et $\sigma^2 = 1$, la variable aléatoire X est qualifiée de Gausienne centrée réduite.

La forme de la densité de probabilité de la variable aléatoire X varie selon les paramètres m et σ^2 (voir FIG.1.6).

Proposition.1 Soit X une variable aléatoire qui suit la loi Normale centrée réduite, alors, la variable aléatoire $\dfrac{X^2}{2}$ suit la loi $\gamma(\frac{1}{2}, 1)$.

Moments

Si X suit la loi $N(m, \sigma^2)$, alors

$$\mathbb{E}((X-m)^{2k+1}) = 0 \quad \text{et} \quad \mathbb{E}((X-m)^{2k}) = \frac{2^k \sigma^{2k}}{\sqrt{\pi}} \Gamma\Big(\frac{2k+1}{2}\Big),\ \forall\ k\ \in \mathbb{N}.$$

Par conséquent, les moments d'odre impair d'une variable Normale centrée sont nuls, et les moments d'ordre pair s'expriment seulement en fonction de la variance. Autrement dit, la loi d'une variable aléatoire Normale est caractérisée par son moment d'ordre 2.

Loi Bêta de première espèce

Dans la théorie des probabilités et statistiques, la loi Bêta est une loi de probabilité absolument continue, définie sur $[0, 1]$, caractérisée par deux paramètres de forme a et b.

La loi Bêta est un cas spécial de la loi de Dirichlet. La famille des lois Bêta admet une grande variété de formes et elle permet de modéliser de nombreuses distributions à support fini.

Définition.3 Une variable aléatoire X suit la loi Bêta univariée sur $[0,1]$, de paramètres de forme $a > 0$ et $b > 0$, dite première espèce, si sa densité de probabilité est définie par

$$f_X(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}\mathbf{1}_{[0,1]}.$$

avec $B(a,b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

Nous désignons par $B_1(a,b)$ la loi Bêta univariée première espèce de paramètres de forme a et b. Cette loi peut être définie sur tout intervalle $[c,d]$ en faisant la transformation affine $H = c + dX$.

Remarque.2

Dans le cas $a = 1$ et $b = 1$, la loi Bêta sur $[0,1]$ coïncide avec la loi uniforme sur $[0,1]$.

La forme de la densité de la variable aléatoire X varie selon les paramètres de forme a et b (voir les figures FIG.1.1, FIG.1.2, FIG.1.3 et FIG.1.4).

Moments

Le moment non centré d'ordre $k \in \mathbb{N}^*$ de la variable aléatoire X est

$$\mathbb{E}(X^k) = \frac{B(a+k,b)}{B(a,b)}.$$

Nous en déduisons,

$$\mathbb{E}(X) = \frac{a}{a+b} \quad \mathbb{V}ar(X) = \frac{ab}{(a+b+1)(a+b)^2}.$$

Loi Bêta de deuxième espèce

Définition.4. La variable aléatoire X suit la loi Bêta univariée sur $\mathbb{R}_+$, de paramètres de

forme $a > 0$ et $b > 0$ est dite Bêta deuxième espèce si sa densité de probabilité est définie par

$$f_X(x) = \frac{1}{B(a,b)} x^{a-1}(1+x)^{-a-b}\mathbf{1}_{\mathbb{R}_+}.$$

Nous désignons par $B_2(a,b)$ la loi Bêta univariée deuxième espèce de paramètres de forme a et b.

Les lois Bêta univariées deuxième espèce peuvent être construites à partir des lois Gamma univariées, comme nous le montre la proposition suivante.

Proposition.3. Soient X_1 et X_2 deux v.a. indépendantes qui suivent respectivement la loi $\gamma(a,\lambda)$ et la loi $\gamma(b,\lambda)$, $a > 0,\ b > 0,\ \lambda > 0$. Alors la variable alétoire $Y = \dfrac{X_1}{X_2}$ suit la loi $B_2(a,b)$.

Preuve En se basant sur l'hypothèse de l'indépendance entre les deux variables aléatoires X_1 et X_2, la loi du couple (X_1, X_2) a pour densité

$$f_{(X_1,X_2)}(x_1,x_2) = \frac{1}{\Gamma(a)\Gamma(b)} e^{-(x_1+x_2)} x_1^{a-1} x_2^{b-1} \mathbf{1}_{\mathbb{R}_+^* \times \mathbb{R}_+^*}(x_1,x_2).$$

Considérons la transformation $\delta : \mathbb{R}_+ \times \mathbb{R}_+ \longrightarrow \mathbb{R}_+ \times \mathbb{R}_+$ définie par

$$\delta(x_1,x_2) = (U = X_1, Y = \frac{X_1}{X_2}).$$

Comme le Jacobien $J(\delta^{-1})$ de la transformation δ^{-1} est $\frac{u}{y^2}$; alors la densité du couple (U,Y) est

$$f_{(U,Y)}(u,y) = \frac{1}{\Gamma(a)\Gamma(b)} \frac{u^{a+b-1}}{y^{b+1}} e^{-u(1+\frac{1}{y})}.$$

En effectuant une intégration par rapport à u, la densité de Y est alors définie par

$$f_Y(y) = \frac{1}{B(a,b)} y^{a-1}(1+y)^{-(a+b)}\mathbf{1}_{\mathbb{R}_+^*}.$$

Proposition.4. Soit X une variable aléatoire qui suit la loi $B_1(a,b)$, alors la variable aléatoire $J = \dfrac{X}{1-X}$ suit la loi $B_2(a,b)$.

En se basant sur la proposition prècédente, nous pouvons construire des lois Bêta première espèce à partir des lois Gamma.

Proposition.5. Soient X_1 et X_2 deux v.a. indépendantes qui suivent respectivement la loi $\gamma(a,\lambda)$ et la loi $\gamma(b,\lambda)$, $a > 0,\ b > 0,\ \lambda > 0$. Alors la variable alétoire $U = \dfrac{X_1}{X_1 + X_2}$ suit la loi $B_1(a,b)$.

Moments

Le moment non centré d'ordre $k \in \mathbb{N}^*$ de la variable aléatoire Y est

$$\mathbb{E}(Y^k) = \frac{B(a+k, b-k)}{B(a,b)}.$$

Nous en déduisons que,

$$\mathbb{E}(Y) = \frac{a}{b-1} \quad \text{pour}(b > 1) \quad \mathbb{V}ar(Y) = \frac{a(a+b-1)}{(b-1)^2(b-2)}, \quad \text{pour}(b > 2).$$

La variable aléatoire Y admet des moments d'ordre 1 et 2 que pour $b > 2$.

Annexe B

Teinte

La teinte est la forme pure d'une couleur, c'est-à-dire sans adjonction, ni de blanc, ni de noir, ni de gris. Les teintes sont représentées sur le pourtour du cercle chromatique (dit aussi "roue chromatique", "roue des couleurs" ou encore "cercle des couleurs"). On peut parler également de tons. Du fait que l'on représente le plus souvent cette composante sur le cercle chromatique, elle est généralement considérée comme étant de nature angulaire (allant de 0° à 360° exclu en partant généralement du rouge pur). Cependant, en informatique ou en mathématiques, on peut la considérer, soit comme un pourcentage (c'est-à-dire, un nombre réel allant de 0 à 1, soit comme un nombre entier allant généralement de 0 à 255 (lorsqu'il est codé sur un byte de 8 bits ou sur 1 octet). Certaines transformations atypiques utilisent d'autres échelles comme, par exemple, des nombres entiers allant de 0 à 1530. On notera cependant que si le cercle chromatique répartit de manière équitable les teintes pures (rouge, jaune, vert, cyan, bleu, magenta, puis rouge), ça ne constitue pas une obligation concernant le TSL. Ainsi, dans certains domaines tels que l'imagerie satellite ou la colorimétrie, on utilise parfois des cercles chromatiques répartissant différemment les teintes sur le cercle (de manière non linéaire par exemple). Ceci permet des ajustements

pour une meilleure corrélation avec la réalité physique ou perceptuelle, ou encore, dans le but de regrouper certaines tranches afin d'en faire ressortir une information recherchée. Le TSL n'utilise donc pas son cercle chromatique comme une définition objective des teintes physiques mais comme une représentation théorique abstraite.

Saturation

En théorie des couleurs, la saturation, ou, "pureté", est l'intensité d'une teinte spécifique : une teinte hautement saturée a une couleur vive et intense tandis qu'une teinte moins saturée paraît plus fade et grise. On retrouvera également cette notion sous les termes de "chroma", de "chromie" ou encore, de "chrominance" (bien que cette dernière notion soit sujette à controverse en ce qui concerne le TSL). Une teinte totalement "désaturée" (i.e., les nuances du gris) sera dite "achromatique". Cependant, cette composante est dite "conservative" car si toutes les teintes désaturées font bien référence au même gris (qu'importe la composante de clarté "L"), ce gris possède obligatoirement une teinte. Donc, si l'on sature le gris, la teinte initiale réapparaîtra. Par conséquent, les saturations maximales et minimales ne constituent pas des cas particuliers. En effet, dans le cas de l'utilisation d'une saturation maximale, on pourrait considérer que la composante de saturation est superflue puisqu'elle n'influence pas la description de la couleur. Et inversement, dans le cas de l'échelle des gris, on pourrait considérer que la teinte est une information inutile puisqu'elle est non perceptible. Mais, en raison de cette propriété conservative, dans les deux cas, les deux composantes sont nécessaires au TSL pour définir une couleur. Cependant, certains algorithmes ou codages en informatique peuvent adapter leur gestion de ces cas extrêmes à des fins de compression sans perte en les considérant comme des cas particuliers. Cette composante indépendante est de nature linéaire bien que son échelle

puisse, comme pour les autres composantes, ne pas être linéaire à des fins d'ajustements pratiques. On la considère généralement comme étant le rayon du cercle chromatique des teintes pures ou de la sphère (selon les représentations que l'on donne au TSL). Dans cette approche, elle pourra être relative et signée afin de devenir négative. Si ceci dénature quelque peu le TSL, cette technique permettra d'obtenir à peu de frais mathématiques la teinte opposée sur le cercle chromatique, c'est-à-dire, la couleur complémentaire. On la considère donc généralement comme étant un pourcentage, c'est-à-dire, une valeur en nombres réels allant de 0 à 1. En informatique, on l'utilisera généralement sous la forme d'un nombre entier allant de 0 à 255 (ou de -127 à $+128$), voire, de 0 à 127 ou 128 selon les cas. Cependant, il est à noter qu'il y a une singularité remarquable lorsqu'on l'utilise avec des nombres entiers sur une échelle qui n'est pas impaire. En effet, dans le domaine des nombres entiers, il n'existe pas de valeur centrale à un nombre pair. Par exemple, la moitié de 255 est 127,5, donc, le "gris pur" ne pourra pas être défini précisément puisqu'il faudra choisir entre la valeur 127 et la valeur 128. Si l'on considère les teintes comme étant la périphérie d'un cercle et que l'on connait également sa taille exacte (i.e., l'échelle), on peut en déduire l'échelle de la saturation qui est son rayon avec ce calcul : $S = \frac{T}{2\pi}$. La représentation sous la forme d'un carré de la relation entre les teintes et la saturation est à considérer correctement car, si elle montre bien la propriété conservative de ces deux composantes du TSL, elle n'en demeure pas moins erronée du fait que le gris pur est représenté sous la forme d'une ligne (la dernière en bas) et non pas d'un point unique comme le définit le TSL. C'est donc une représentation pratique (en informatique en particulier) mais trompeuse. Ceci dit, la représentation sous la forme d'un cercle (avec ses rayons de teintes saturées) ne peut pas être considérée correcte dans le domaine des nombres entiers

car la géométrie dans les espaces matriciels comporte quelques singularités. Par exemple, le théorème de Pythagore ne s'applique pas dans les espaces non euclidiens. En effet, dans un espace matriciel, les points de la périphérie d'un cercle ne sont pas tous à la même distance de son centre.

Lumière/Luminosité/Luminance/Clarté

Cette composante fait référence à l'ajout d'une intensité du blanc ou du noir aux teintes (saturées ou non). On parle également "d'intensité lumineuse de la couleur" bien que cela soit sujet à controverse car le TSL ne décrit pas le comportement physique de la lumière mais décrit une couleur. Les termes usités font en effet polémique par leur référence à la réalité physique ou perceptuelle de la lumière. Mais d'un point de vue purement perceptuel de la couleur, le TSL se révèle également inexacte car certaines couleurs sont par nature plus sombres ou claires que d'autres. Par exemple, nous savons que le rouge est plus lumineux que le bleu, mais qu'il est plus sombre que le vert, qui est la couleur pour laquelle l'oeil est le plus sensible. Cette polémique est issue en partie de la traduction du terme anglophone "lightness". Par conséquent, nous préfèrerons utiliser dans cet article le terme de "clarté" de la couleur qui se révèle plus significatif pour faire référence à cette composante du TSL. Tout comme ses homologues, cette composante peut ne pas être régulière à des fins d'adaptations spécifiques. Par exemple, nous pourrions avec plus de "gradients" entre la teinte et le noir pur que de la teinte au blanc pur ou inversement !. On notera également que ces deux extrêmes (i.e., le noir et le blanc) constituent, à l'instar de la saturation, des cas singuliers non particuliers car cette composante est également conservative. Et, tout comme pour la saturation, certaines techniques de compression sans perte du codage en informatique peuvent exister. On peut donc, à partir du noir ou du

blanc pur, retrouver, et la teinte, et la saturation initiale de la couleur en modifiant la clarté. Cependant, tout comme pour le gris, le TSL ne définit qu'un seul noir et qu'un seul blanc (qui sont donc des couleurs uniques à part entières mais qui peuvent être définies de bien des manières selon T et S). Cette composante est également de nature linéaire et peut donc être considérée tant comme un pourcentage que comme une valeur entière relative ou non. Mais, contrairement à la saturation, on peut également la considérer comme étant une donnée angulaire (représentant les latitudes sur la représentation sphérique) à l'instar de la Teinte. Nous retrouvons donc les deux représentations possibles du rapport entre T et L (i.e., sous la forme d'un cercle ou d'un carré) avec la même problématique que pour les représentations de la saturation, à savoir que le carré fait ressortir la propriété conservative tout en multipliant le noir et le blanc ce qui est donc incorrect. Le fait que la clarté puisse être considérée, soit comme la hauteur du système, soit comme une latitude fait qu'il existe de nombreuses représentations tridimensionnelles possibles du TSL. Cependant, il est préférable de le représenter dans un espace non euclidien en considérant une sphère dont la couleur est un rayon partant du centre gris. On remarque également que lorsqu'elle est considérée comme représentant la hauteur du système, si sa valeur est sur une échelle relative, le point 0 est le gris. Sinon, on considère la valeur minimale comme représentant normalement le noir. Si on la considère comme étant une latitude, on peut procéder de la même manière. Mais en la considérant comme une valeur angulaire, selon où l'on place l'angle 0, les valeurs négatives peuvent, à l'instar de la saturation, inverser la teinte.

Bibliographie

[1] Abramowitz M., Stegun I.A., Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. *Dover, New York, ninth dover printing, tenth gpo printing edition*, 1964.

[2] Acharya T., Image Processing : Principles and Applications, *John Wiley and Sons*, 2005.

[3] Anwander A., Segmentation d'images couleur par un opérateur gradient vectoriel multiéchelle et contour actif : Application à la quantification des phases minéralogiques du clinker de ciment, *thèse de doctorat de L'INSTITUT NATIONAL DES SCIENCES APPLIQUÉES DE LYON*, 2001.

[4] Banerjee A., Dhillon I., Ghosh J., Merugu S., An Information Theoretic Analysis of Maximum Likelihood Mixture Estimation for Exponential Families. *In International Conference on Machine Learning*, pages 57-64, 2004.

[5] Banfield J.D., Raftery A.E., Model-Based Gaussian and Non-Gaussian. Clustering, *Biometrics*, vol. 49, No. 3, 1993, pp. 803-821, 1993.

[6] Barndorff-Nielsen O., Information and Exponential Families in Statistical Theory. *Wiley and Sons*, Inc, New York, 1978.

[7] Barndorff-Nielsen O.E, Mikosh T, Resnick S.I. Lévy Processes. *Theory and Applications*, Birkhäuser, Boston, Mass., U.S.A, 2001.

[8] Balakrishnan N, Johnson N.L, Kotz S., Continuous multivariate distributions. *Models and Applications, Wiley* 1985, New York, 2nd edition(2000).

[9] Ben Arab T., Zribi M., Masmoudi A., Application of the Variance-covariance Matrix in Natural Exponential to Identify Noise in Signal, *2012 IEEE International Conference on Systems, Man, and Cybernetics*, October 14-17, 2012 in COEX, Seoul, Korea, pp 1054-1057.

[10] Ben Arab T., Masmoudi A., Zribi M., Trace of the variance-covariance matrix in natural exponential families. *Communications in Statistics-Theory and Methods*, Accepted, 2012.

[11] Ben Arab T., Zribi M., Masmoudi A., Parameter Estimation Of The Diagonal Of The Modified Riesz Distribution, *IEEE 17th International Conference on Intelligent Engineering Systems*, June 19-21, in Costa Rica 2013.

[12] Ben Arab T., Zribi M., Masmoudi A., Unsupervised Satellite Image Segmentation using a Bivariate Beta type-II Mixture Model. *IEEE 9th International Conference on Computational Cybernetics*, July 8-10, Tihany, Hungary, Page(s) : 95-100, 2013.

[13] Ben Arab T., Zribi M., Masmoudi A., Finite Bivariate Gamma Mixtures for Color Image Segmentation. *International Conference on advances in Mathmatics ICAAM*, December 16-19, Hammamet-Tunisia, 2013.

[14] Ben Arab .T, Masmoudi A., Zribi M. Bayesian Approach for Parameter Estimation of the Diagonal of the Modified Riesz Distribution. *Neurocomputing*, Soumis.

[15] Benois J., Barba D., Image segmentation by region-contour cooperation for image coding, *ICPR92*, Vol C, pp 331-334, La Haye 1992.

[16] Biernacki C., Celeux G., Govaert G., Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Computational Statistics and Data Analysis*, pages 561-575. 2003.

[17] Bickel P., Freedman D., Some asymptotic theory for the bootstrap, *Annals of statistics*, 9, (6), 1196-1217, 1981.

[18] Bernardo J., Smith A.F.M., Bayesian Theory. *John Wiley*, New York, 1994.

[19] Beucher S., Meyer, F., The morphological approach to segmentation : the watershed transformation. *Mathematical Morphology in Image Processing (Ed. E.R. Dougherty*, pages 433-481, 1993.

[20] Bouveyron C., Modélisation et classification des données de grande dimension : Application à l'analyse d'image. *Thèse soutenue à l'université Joseph Fourier, INRIA Rhône-Alpes*, 28 septembre, 2006.

[21] Bradley E., The Jackknife, the Bootstrap, and Other Resampling Plans, *Society for Industrial and Applied Mathematics*, U.S., 31 janvier 1982.

[22] Canny J., A computational approach to edge detection. *IEEE Pattern Analysis and Machine Intelligence*, vol.8, 6, pp. 679-698, 1986.

[23] Carron T., Segmentations d'images couleur dans la base Teinte-Luminance-Saturation : approche numérique et symbolique *Thèse de l'université de Savoie*, 1955.

[24] Casalis M., Les familles exponentielles sur $\mathbb{R}^2$ de fonction-variance $V(m) = am \otimes m + B(m) + c$. *C. R. Acad.Sci*, Paris Sér. I Math. 314 635-638, 1992.

[25] Casalis M., The $2d + 4$ simple quadratic natural exponential families on $\mathbb{R}^d$. *Ann. Statist* 1996 ; 24, Pages 1828-1854.

[26] Feller W., An introduction to probability theory and its applications. *Wiley and Sons*, Vol. 2, 2nd edition, Inc, New York, 1971.

[27] Chao-Wei C., Wen-Jang H., A note on characterizations of the bivariate gamma distribution, *Journal of Statistical Planning and Inference*, 2003.

[28] Cheng H.D., Jiang X.H., Sun Y., Wang J., Color image segmentation : advances and projects. *Pattern Recognition*, 34(12) :2259-2281, 2001.

[29] Chesnaud C., Technique statistique de segmentation par contour actif et mise en oeuvre rapide. *Thèse soutenue à l'université des droits, d'économie et des sciences d'Aix-Marseille*, Février, 2000.

[30] Courrègue Ph., Philoche J.-L, Priouret P., Regression linéaire et estimation par la méthode des moindres carrés, *Ann. Inst. Henri Poincaré* , vol 7, n° 4, pp 253-270, 1971.

[31] Dasgupta A., Raftery A.E, Detecting features in spatial point processes with clutter via model-based clustering, *Journal of the American Statistical Association*, 93, 441, 1998.

[32] Delmas J.P., Relation entre les algorithmes d'estimations itératives EM et ICE avec exemples d'aplications, *Quinzième collogue Gresti-Juan-les-pins* , 18-21 Septembre 1995.

[33] Dempster A.P.,n Laird N M., Rubin D.B., Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society*, B, 39, pp.1-38, 1977.

[34] Epimaco A.C.Jr., Roberto N.P., Khursheed A., Generalization of the Dirichlet Distribution. *9th National Convention on Statistics (NCS)*.October 4-5, 2004.

[35] Eskicioglu A.M, Fisher P.S., Image Quality Measures and their Performance, *IEEE Transactions On comm...*, Vol.43, No.12, pp.2959-2965, 1995.

[36] Faraut J., Korányi A., Analysis on symmetric cones, *Oxford university press*, 1994.

[37] Fontaine M., Segmentation non supervisée d'images couleur par analyse de la connexité des pixels. *Thèse soutenue à l'université des sciences et technologies de Lille 1*, France, décembre 2001.

[38] Gaver P., Multivariate gamma distributions generated by mixture. *Sankhya*, Series A, 32, 123-126, 1970.

[39] Gikhman I.I, Skorohoda A.V., The Theory of Stochastic Processes 2. *Springer*, New York, 1973.

[40] Ghribi A., Masmoudi A., Characterization of multinomial exponential families by generalized variance. *Statistics and Probability Letters*, Vol. 80.2010, 11-12, p. 939-944, 2010.

[41] Gupta A.K, Ofori-Nyarko S., Improved estimation of generalized variance and precision. *Statistics*, 26, 99-109., 1995.

[42] Hadrich A., Zribi M., Masmoudi A., A proposed normalized B-spline density estimator and it application in unsupervised statistical image segmentation. *SMC 2012*, 480-483, 2012.

[43] Haralick R.M, Shapiro L.G., Computer and robot vision, *Addison-Wesley Pub. Co.*, volume 1, 1992.

[44] Harold P., Méthodes et algorithmes de dématriçage et de filtrage du bruit pour la photographie numérique.*PhD thesis, Université de Paris Est*, 2009.

[45] Hartigan J. A., Clustering algorithms, *Wiley*, 1975-351, 1975.

[46] Hassairi A., Generalized Variance And Exponential Families, *Annals of Statistics*, 27, Pages 374-385, 1999.

[47] Hassairi A., Lajmi S., Riesz exponential families on symmetric cones, *J. Theoret. Probab.*, **14**, 927-948, 2001.

[48] Hassairi A., Masmoudi A., Extension of the variance function of a steep exponential family. *J. Multivariate analysis* , 239-256, 2005.

[49] Hassairi, A., Zarai, M., Characterization of the simple cubic multivariate exponential families. *Journal of Functional analysis*, 69-89, 2006.

[50] Herbulot A., Mesures statistiques non-paramétriques pour la segmentation d'images et de vidéos et minimisation par contours actifs, *thèse de doctorat de l'Université de Nice Sophia-Antipolis*, 2007.

[51] Horowitz S.L., Pavlidis T., Picture Segmentation by a directed split-and-merge procedure, *Proceedings of the 2nd International Joint Conference on Pattern Recognition*, pp. 424-433, Copenhagen, Denmark 1974.

[52] Iliopoulos G., Kourouklis S., On improved interval estimation for the generalized variance. *Journal of Statistical Planning and Inference*, 66, 305-320, 1998.

[53] Jagadesh B.N. et al, A Robust Skin Colour Segmentation Using Bivariate Pearson Type II (Bivariate Beta) Mixture Model, *I.J. Image, Graphics and Signal Processing*, 2012.

[54] Joong-Ho, Won Johan L.-J.K., Rajaratnam B., Maximum likelhhood covariance estimation with a condidion number constraint. *Technical Report N°. 2009-10, Department of Statistics Stanford university, Stanford, California 94305-4065.* August 2009.

[55] Kelly P.A. et al. Adaptative segmentation of speckeled images using a hierarchical random fied model, *IEEE transactions on Acoustic, Speech and Signal processing*, pp. 1628-1641,1988.

[56] Khinchin A.Ya., Paul Lévy Processus stochastiques et mouvement brownien [Stochastic processes and Brownian motion]. *Uspekhi Mat. Nauk*, 5 :2(36), 219-219, 1950.

[57] Kokonendji C.C., Pommeret D., Estimateurs de la variance généralisée pour des familles exponentielles non Gaussiennes. *C.R. Acad. Sci*, Paris Sér. I 332, 351-356, 2001.

[58] Kokonendji C.C, Pommeret D., Comparing UMVU and ML estimators of the generalized variance for natural exponential families. *Statistics*, 41, 547-558, 2007.

[59] Kokonendji C.C, Seshadri V., On the determinant of the second derivative of a Laplace transform. *Ann. Statist*, 24, Pages 1813-1827, 1996.

[60] Koopman P.O., On distributions admitting a sufficient statistic. *Transactions of the American Mathematical Society*, vol. 39, pages 399-409, 1936.

[61] Lambert, P., Macaire, L., Filtering and segmentation : the specificity of color images (keynote). *In Proc. of the Int. Conf. on Color in Graphics and Image Processing*, volume 1, pages 57-64, Saint-Etienne, 2000.

[62] Lantier D., Trotoux D., La loi des grands nombres : le théorème de De Moivre-Laplace, Contribution à une approche historique de l'enseignement des mathématiques, *collection Les publications de l'IREM de Besançon*, ISBN : 2-909963-136, pp. 259-294.

[63] Lejeune M., Statistique : La théeorie et ses applications. *Springer*, pages 448, 27 janv 2011.

[64] Letac G., A characterization of the Wishart exponential families by an invariance property. *Th. Prob.*, vol. 2, No 1, p. 71-86. 1989.

[65] Letac G, Mora M., Natural Real Exponential Families With Cubic variance Functions. *Ann. Statist*, 18, Pages 1-37,1990.

[66] Letac G., Lectures on natural exponential families and their variance functions. *Instituto de Matemática pura e aplicada : monografias de matemática*, Rìo. 50, 1992.

[67] Lezoray O., Segmentation d'images couleur par morphologie mathématique et classification de données par réseaux de neurones : application à la classification de cellules en cytologie des séreuses, *thèse de doctorat de l'Université de Caen*, 2000.

[68] Lourme A., Contribution à la Classification par Modèles de Mélange et Classification Simultanée d'Echantillons d'Origines Multiples, *thèse de doctorat de l'Université*

Lille 1, 2011.

[69] MacQueen J., Some Methods for Classification and Analysis of MultiVariate Observations, *Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1, page 281-297, 1967.

[70] Marchic N., Pieczynski W., Estimation of mixture and unsupervised segmentation of images. *Proceeding of IGARSS'91*, pp. 1083-1086, Helsinki, Finland, Jun 1991.

[71] Maritz J.S., Lwin T., Empirical Bayes Methods.*Chapman and Hall*, 2nd ed., 1989.

[72] Martial GRIMM M., Proposition d'une nouvelle méthode de transformation de signaux vidéo RVB $\dashrightarrow$ HSI. Application aux systèmes de vision couleur temps réel, *Thèse de doctorat de l'Université HENRY POINcARÉ NANCY I*, 1994.

[73] Marzouki A. Segmentation statistique des images radar. *Thèse soutenu à l'université des sciences et technologie de Lille*, 1996.

[74] Marr D., Hildreth E., Theory of edge detection. *Proc. Royal Soc.* London, vol. 20, pp. 187-217, 1980.

[75] Masmoudi A., Conditional Natural Exponential Families. *Statistics and Probability Letters*, 76, Pages 1882-1888, 2006.

[76] Mendoza-Rosas A.T., De la Cruz-Reyna S., A mixture of exponentials distribution for a simple and precise assessment of the volcanic hazard, *Nat. Hazards Earth Syst. Sci.*, 9, 425-431, 2009.

[77] McLachlan G.J., Peel D., Finite Mixture Models. *John Wiley and Sons*, New York, 2000.

[78] Morris C.N., Natural Exponential Families With Quadratic Variance-Function. *Ann. Statist*, 10, Pages 65-80, 1982.

[79] Pal S., Image Segmentation Using the Rough-Set-Initialized EM Algorithm, *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, vol. 40, no. 11, november 2002.

[80] Parka S.H, Yuna I.D, Leea S.U, Color image segmentation based on $3-D$ clustering : morphological approach, *Pattern Recognition,* Volume 31, Issue 8, 1, Pages 1061-1076, August 1998.

[81] Pommeret D., Bhattacharyya matrices and exponential families. *Journal of multivariate analysis* , 63, 1, page 105-118, 1997.

[82] Pommeret D., A Construction Of The UMVU Estimator for Simple Quadratic Natural Exponential Families. *J. Multivar. Anal.* 85, Pages 217-233, 2003.

[83] Prewitt J.M.S., Object enhancement and extraction. *Picture processing and Psychopictorics*, pp. 75-149, Academic Press, New York, 1970.

[84] Rajkumar G.V.S., Srinivasa Rao K., Srinivasa Rao P., Studies on Colour Image Segmentation T echniqu e Based on Finite Left Truncated Bivariate Gaussian Mixture Model with K-Means, *Global Journal of Computer Science and Technology* , Volume 11 Issue 18 Version 1.0 October 2011.

[85] Robbins H., An empirical bayes approach to statistics, *Proc. Third Berkley Symposium on Mathematical Statistics*, 1, 157-163, 1956.

[86] Roberts L.G., Machine perception of three dimensional solids. *Optical and Electro-optical Information Processing*, pp. 159-197, MIT Press, Cambridge, 1965.

[87] Rosenberger C., Chehdi K., Unsupervised segmentation of multi-spectral images, *International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS)*, September 2-5, Ghent, Belgium, 2003.

[88] Rousseeuw, P.J. and Van Driessen, K., A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 41, 212-223, 1999.

[89] Rousselle J.J., Les contours actifs, une méthode de segmentation, application à l'imagerie médicale, *thèse de doctorat de l'Université François Rabelais de Tours*, 2003.

[90] Rowe D.B., Multivariate Bayesian statistics : models for source separation and signal unmixing. *Chapman and Hall/CRC*, 2003.

[91] Sahoo P.K., Soltani S., Wong A. K.C., Chen Y.C., A survey of thresholding techniques, *Computer Vision, Graphics, and Image Processing*, Volume 41, Issue 2, Pages 233-260, 1988.

[92] Sangwine S.J., Horne R.E.N, The Color Image Processing Hand Book, *Chapmann and Hall*, UK, 1998.

[93] Sato K., Lèvy Processes and Infinitely Divisible Distributions. *Cambridge University Press*, Cambridge, 1999.

[94] Seshadri V., The Inverse Gaussian distribution, *McGill University* , Montreal, Oxford, 1993.

[95] Skriver H., Nielsen A.A, Conradsen K., Evaluation of the Wishart test statistics for polarimetric SAR data. *International Geoscience and Remote Sensing Symposium (IGARSS)*, Vol.2, 699-701, 2003.

[96] Smith Ph.D, Steven W., The Scientist and Engineer's Guide to Digital Signal Processing. *California Technical Publishing*, USA, 1998.

[97] Sobel I., Neighbourhood coding of binary images for fast contour following and general array binary processing. *Computer Graphics and Image Processing*, vol.8, pp. 127-135, 1978.

[98] Tomasy, J., Kozubowski, J., Nolan, P. Infinite divisibility of skew Gaussian and Laplace Laws. *Statistics and Probability Letters*, 78, Issue 6, 654-660, 2008.

[99] Travis D., Effective color displays, theory, practice. *Academic press*, 1991.

[100] Vandenbroucke N., Segmentation d'images couleur par classification de pixels dans des espaces d'attributs colorimètriques adaptés. Application à l'analyse d'images de football. *Thèse de doctorat, Université Lille 1*, décembre 2000.

[101] Wedderburn R.W.M., Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biornetrika*,1974.

[102] Wicksell, S.D., On Correlation Functions of Type III. *Biometrika*, vol 25, 121-133, 1933.

[103] *Wintenberger.fr/cours.L3MAS/2012/StatMathPoly2012.pdf.*

[104] Yao J., Université de Rennes 1, *http ://perso.univ-rennes1.fr/jian-feng.yao/pedago.*

[105] *www.iecn.u-nancy.fr/ tindel/enseignement/esial/estim-ponctuelle.pdf.*

[106] Ypma T.J, Historical developement of the Newton-Raphson method. *SIAM Review*, 37, 531-551, 1995.

[107] Zhi-kai H., De-Hui L., Unsupervised Image Segmentation Using EM algorithm by Histogram, *ICIC* , Vol. 4681, pp. 1275-1282, 2007.

[108] Zhong Y., Zhang L., Huang B., Li P., An Unsupervised Artificial Immune Classifier for Multi/Hyperspectral Remote Sensing Imagery, *IEEE Trans. Geosc. Remote Sens.*, Vol. 44, No. 2, Febr. 2006.

[109] Zhu Q.Y., Modèle bayésiens et application à l'estimation des caractéristiques de produits finis et au contrôle de la qualité. *Thèse soutenue à l'Ecole Nationale des Ponts et Chaussées*, Paris, Juillet 1991.

[110] Zoubeida M., Performances des Detecteurs Distribues CFAR Dans Un Clutter α-Stable Bases Sur des Methodes de Pretraitement : La Transformee en Ondelettes et La Compression Non Lineaire.*Thèse soutenue à l'université de Constantine département d'electronique*, Algérie, 2007.

[111] Zribi M., Les Fonctions Spéciales et les Représentations des Groupes pour la Reconnaissance des Formes : Application à l'Imagerie Médicale, *Thèse de doctorat de l'Université de Rennes I*, 1997.

[112] Zucker, S., Region growing, childhood and adolescence. *Computer Vision Graphics and Image Processing*, 5, 382-399, 1976.

Résumé : Cette thèse est consacrée à l'évaluation des familles exponentielles pour les problèmes de la modélisation des bruits et de la segmentation des images couleurs.
Dans un premier temps, nous avons développé une nouvelle caractérisation des familles exponentielles naturelles infiniment divisibles basée sur la fonction trace de la matrice de variance covariance associée. Au niveau application, cette nouvelle caractérisation a permis de détecter la nature de la loi d'un bruit additif associé à un signal où à une image couleur. Dans un deuxième temps, nous avons proposé un nouveau modèle statistique paramétrique multivarié basé sur la loi de Riesz. La loi de ce nouveau modèle est appelée loi de la diagonale modifiée de Riesz. Ensuite, nous avons généralisé ce modèle au cas de mélange fini de lois. Enfin, nous avons introduit un algorithme de segmentation statistique d'image couleur, à travers l'intégration de la méthode des centres mobiles (K-means) au niveau de l'initialisation pour une meilleure définition des classes de l'image et l'algorithme EM pour l'estimation des différents paramètres de chaque classe qui suit la loi de la diagonale modifiée de la loi de Riesz.

Mots clés : Algorithme EM, Estimateur, Familles exponentielles naturelles, Fonction trace, Loi de Riesz, Méthode des centres mobiles, Segmentation statistique.

Abstract : This thesis is dedicated to the evaluation of the exponential families for the problems of the noise modeling and the color images segmentation. First, we developed a new characterization of the infinitely divisible natural exponential families based on the trace function of the associated variance-covariance matrix. At the application level, this new characterization allowed to detect the nature of the law of an additive noise associated with a signal or with a color image. Second, we proposed a new parametric multivariate statistical model based on Riesz's distribution. The law of this new model is called the modified diagonal Reisz distribution. Then we generalized this model in the case of a finished mixture of distribution. Finally we introduced an algorithm of statistical segmentation of color images through the integration of the k-means method at the level of the initialization for a better definition of the image classes and the algorithm EM for the estimation of the different parameters of every class which follows the modified diagonal Reisz distribution.

Keywords : EM algorithm, Estimator, Natural exponential families, Riesz distribution, Statistical segmentation, Trace function.