

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la

Recherche Scientifique

Université Frères Mentouri Constantine

FACULTE DES SCIENCES EXACTES
Département de Mathématique

N° d'ordre : .37/DS/2021...

Série: 04 Maths/ 2021.....

Thèse

Pour l'obtention du diplôme de

Doctorat En Sciences

En : **Mathématiques**

Option : Mathématiques appliquées

Intitulée

Les modèles 2D MA à réponse
impulsionnelle finie et à phase non minimale

Présentée par

Beghriche Hanane

Devant le jury composé de :

Mme. S. Belaloui	Pr, Dept. Mathématiques	U.C.1	Présidente
Mme. S. Kharfouchi	Pr, Faculté de Médecine	U.C.3	Rapporteur
Mr. A. Boudeliou	MCA, Dept. Mathématiques	U.C.1	Examineur
Mr. O. Zehrou	Pr, Dept. Mathématiques	U.L.B.M	Examineur

Soutenue le : 29 avril 2021

Remerciements

Comme c'est difficile de trouver les mots exacts qui expriment mes profonds sentiments de gratitude envers les personnes qui étaient à mes côtés pour m'aider à surmonter tous les obstacles et à mener à bien ce travail. Ce n'est que justice de leur écrire cette page.

Ainsi, je tiens à exprimer ma très vive reconnaissance à l'égard de mon encadrante Professeur Soumia Kharfouchi, pour ses conseils, ses encouragements, sa disponibilité et son soutien tout au long de ces années. Les mots me manquent pour la remercier de m'avoir motivée et de m'avoir laissée la liberté nécessaire à l'accomplissement de mes travaux tout en gardant un œil critique et avisé.

Je remercie également Pr. S. Belaloui, professeur à l'université de Constantine 1, qui m'a fait l'honneur de présider le jury de thèse. Mes remerciements les plus respectueux vont aussi à :

- Dr. A. Boudeliou, Maître de Conférences à l'université de Constantine 1.
- Dr. O. Zehrou, Maître de Conférences à l'université de Oum el-Bouaghi.

Qui m'ont fait l'honneur d'accepter de participer à ce jury de thèse. Je leurs suis très reconnaissante pour l'intérêt qu'ils ont porté à mes travaux.

Je dédie ce travail à tous les membres de ma famille, à qui je dois beaucoup, sans leurs patiences, ce travail n'aurait pu voir le jour.

Enfin, Merci à tous ceux qui m'ont aidé en m'accordant leurs temps, leurs encouragements et leurs savoirs.

Hanane Beghriche

RÉSUMÉ

Plusieurs domaines d'application comme : l'astronomie, l'acoustique, le traitement des images et du signal, etc., ont eu recours à l'utilisation des statistiques d'ordre supérieur. Ils ont joué un rôle crucial dans l'identification de système linéaire à phase non minimal. Ainsi, l'objet de cette thèse est l'identification des paramètres d'un modèles 2D MA à phase non minimale à l'aide des cumulants d'ordre supérieur. Nous démontrons d'abord la convergence presque sûre des estimateurs empiriques des statistiques d'ordre supérieurs. Comme cadre pratique, nous abordons le problème de l'identification des modèles à moyennes mobiles 2D (MA) avec des erreurs non gaussiennes basées sur les cumulants seuls sous une hypothèse de phase non minimale en premier lieu et ensuite sur une approche GMM (méthode des moments généralisée). Une étude de simulation vérifie les performances des méthodes proposées.

Mots clé

Statistiques spatiales d'ordre supérieur, phase non minimale, 2D FIR, convergence presque sûre

ABSTRACT

Several fields of application such as: astronomy, acoustics, image and signal processing, etc., have used of higher order statistics. They played a crucial role in the identification of a non-minimal phase linear system. Thus, the object of this thesis is the identification of the parameters of a non-minimal phase 2D MA models using higher order cumulants. First, the almost-sure convergence properties of sample estimates of higher order spatial statistics are derived. As a practical framework, we address the problem of identification of 2D moving average (MA) models with non-Gaussian errors based on cumulants alone under a nonminimum phase assumption first and on a generalized method of moments approach after. A simulation study verifies the performance of the proposed methods.

Key words:

Higher-order spatial statistics, No minimum phase, 2DFR, almost sure convergence.

ملخص

الهدف المرجو من بحثنا في هذه الأطروحة هو تحديد وسائط النموذج باستخدام العزوم المكانية ذات الدرجة العليا. إحصائيات الدرجة العليا (العزوم المكانية) لعبت دورا حاسما في العديد من المجالات التطبيقية، علم الفلك، الصوتيات، معالجة الصور و الإشارات وبشكل اكثر تحديدا لتحديد النظام الخطي في مرحلة الحد غير الأدنى. نبين في هذا العمل التقارب المؤكد تقريبا للمقدرات التجريبية للإحصائيات ذات الدرجة العليا. كإطار عملي، نعالج مشكلة تحديد نماذج المتوسط المتحرك 2DMA

مع أخطاء غير غاوسية التي تعتمد على العزوم مع افتراض مرحلة الحد غير الأدنى في المقام الأول وعلى نهج طريقة اللحظات المعممة في المقام الثاني . وأخيرا أكدت دراسة المحاكاة التي قمنا بها بالتنوع الجيدة لطريقتين المقترحتين.

كلمات البحث الرئيسية:

الإحصائيات المكانية ذات الدرجة العليا، المرحلة غير الدنيا، التقارب المؤكد تقريبا.

Table des matières

1	Généralités	8
1.1	Concepts et définitions	9
1.1.1	Champs aléatoires	9
1.1.2	Stationnarité	9
1.1.3	Champs aléatoires gaussiens	11
1.1.4	Opérateur de retard	11
1.1.5	Ergodicité	12
1.1.6	Notions d'ordre dans \mathbb{Z}^d	12
1.1.7	Indexation des sous-ensembles	13
1.2	Modèles pour les données spatiales	14
1.2.1	Modèle ARMA spatiaux	14
1.2.2	Modèle AR spatial	16
1.2.3	Modèle MA spatial	17
2	Statistiques d'Ordre Supérieur	18
2.1	Introduction aux Statistiques d'Ordre Supérieur	18
2.2	Variables aléatoires réelles scalaires	19
2.2.1	Fonctions caractéristiques d'une variable aléatoire réelle scalaire	19
2.2.2	Moments et cumulants d'une variable aléatoire réelle scalaire	20

2.2.3	Relations entre moments et cumulants	22
2.2.4	Estimation des moments et des cumulants	24
2.2.5	Exemples de variables aléatoires réelles	27
2.3	Variables aléatoires réelles multidimensionnelles	30
2.3.1	Fonctions caractéristiques des variables aléatoires réelles multidimensionnelles	30
2.3.2	Moments et cumulants de vecteurs aléatoires	30
2.3.3	Relations entre moments et cumulants	32
2.3.4	Quelques propriétés des moments et des cumulants	34
2.3.5	Estimation des moments et des cumulants	35
2.4	Multicorrélations	36
2.4.1	Estimation des multicorrélations	37
2.4.2	La motivation pour utiliser des cumulants	38

3 Les modèles moyenne mobile (MA) bidimensionnelle à réponse impulsionnelle

finie		40
3.1	Filtres à réponse impulsionnelle finie	41
3.1.1	Introduction	41
3.1.2	Filtres RIF unidimensionnels (1D)	42
3.1.3	Filtre RIF bidimensionnels (2D)	43
3.2	La phase non minimale	46
3.2.1	Notions importantes	46
3.2.2	Le filtre à phase minimale	48
3.2.3	Le filtre à phase non minimale	49
3.2.4	Autres dénominations	49
3.3	Modèles moyenne mobile (MA) bidimensionnelle	50
3.3.1	Représentation du modèle MA spatial	50
3.3.2	Représentation du modèle 2D RIF MA	50

4 Propriétés asymptotiques des cumulants bidimensionnels	53
4.1 Convergence presque sûre	53
4.2 Normalité asymptotique	59
5 Identification des modèles 2D RIF MA	62
5.1 Identification des modèles 2D RIF MA par les cumulants d'ordre trois seuls	62
5.2 Identification des modèles 2D RIF MA par l'estimateur GMM	67
5.2.1 Aperçu sur la méthode GMM	68
5.2.2 Application de la procédure GMM au modèle 2D RIF MA	70
5.3 Résultats de simulation	73
5.4 Conclusion	77

Introduction générale

Le traitement du signal est un domaine qui a vu l'émergence d'une nouvelle "matière première", qui est l'information. Cette nouvelle discipline se trouve à la jonction des mathématiques, qui lui fournissent un cadre conceptuel, de l'électronique et de l'informatique, qui concourent aux développements des techniques matérielles et logicielles et de la physique, qui décrit les vecteurs de transmission de l'information.

Le traitement du signal a fondé sa problématique sur la théorie des probabilités, sur l'analyse fonctionnelle et sur l'analyse numérique. Ces disciplines sont le cadre naturel de la modélisation en traitement du signal. La théorie des probabilités a été particulièrement mise à contribution pour construire les modèles utilisés en traitement du signal. Cette branche très riche des mathématiques a été très développée depuis son origine dans la "théorie des jeux" pratiquée dès le 17ⁱème siècle par les précurseurs que furent Bernoulli, Fermat, Pascal et bien d'autres. Les concepts de base de la théorie des probabilités sont les notions de variable aléatoire et de fonction (signal) aléatoire. Ces grandeurs sont décrites complètement à travers des lois (ou densités) de probabilité. On se limite bien souvent à une description donnée par les moments.

La plupart des résultats méthodologiques et des techniques de traitement du signal sont fondés sur une description simplifiée des signaux faisant appel aux moments d'ordre 1 et 2. Les fonctions de corrélation et les densités spectrales de puissance (spectres) en sont les outils de base. Ces grandeurs jouent un rôle essentiel dans la description de signaux dans tous les domaines d'application et dans la définition d'opérateurs fondamentaux : filtre adapté, filtre de Wiener, filtre de Kalman. Le statut particulier des signaux gaussiens, qui sont très répandus et pour lesquels la description à l'ordre 2 est complète, a largement contribué à la diffusion des techniques de traitement limitées à l'ordre 2.

Le développement des moyens de traitement a permis dans un passé récent de prendre en compte une description plus subtile des signaux en utilisant des statistiques d'ordre supérieur à 2. Le cadre conceptuel est déjà très riche. Les probabilistes ont depuis longtemps développé une approche générale des variables et des fonctions aléatoires. On peut trouver le fondement de ces modélisations dans ([1], [9], [6], [7], [8], [19], [64]).

Le fort courant de recherche sur ces nouvelles méthodologies qui se développe depuis une quinzaine d'années a suscité une série de colloques internationaux consacrés aux statistiques d'ordre supérieur ([53], [54], [55], [56], [57]), et de nombreuses revues scientifiques leur ont consacré des numéros spéciaux ([51], [65], [66], [67], [68], [79], [70]). Ces études débouchent sur de nouveaux algorithmes de traitement dans de nombreux domaines d'applications. Citons l'astronomie où sont apparues les premières techniques d'ordre supérieur, la prospection sismique et les communications qui ont été un puissant moteur de développement des techniques de déconvolution aveugle, et plus généralement la géophysique, le contrôle non-destructif, les vibrations, la parole, le radar, le sonar. . .

Avant d'aborder l'étude des nouvelles descriptions de la réalité expérimentale utilisant les statistiques d'ordre supérieur, il est bon de s'interroger sur l'opportunité de ces développements. La question est importante et la réponse ne peut être que conjecturale. A l'affirmation indiscutable du caractère plus général de cette approche, on voit objecter à juste titre la complexité accrue, et plus fondamentalement, l'applicabilité douteuse. Le mieux peut être l'ennemi du bien. Nous proposerons de relever ce défi en restant avertis des risques qu'il contient, mais qui sont somme toute associés à toute innovation.

Parmi les raisons invoquées à l'appui de l'utilisation des SOS, outre leur plus grand degré de généralité, nous pensons que l'argument le plus solide est la possibilité qu'elles offrent de résoudre des problèmes non accessibles à l'ordre 2 : nous citerons la séparation de sources, sans modèle a priori, la déconvolution aveugle ou encore le filtrage non linéaire. Laissons aux applications et à l'avenir le soin de trancher. Et consacrons nous à l'étude des SOS qui nous introduiront dans un monde de nouvelles techniques de traitement, et nous apporteront de grandes satisfactions intellectuelles à la mesure de leurs mystères.

Cependant, la plupart des signaux naturels (ou réels) sont des signaux pour lesquels les entrées sont des données spatiales complexes qui présentent un comportement ou des caractéristiques non gaussiens et / ou non linéaires. Ils ont besoin de méthodes de corrélation d'ordre supérieur pour leur analyse. Les statistiques 2D d'ordre supérieur peuvent mesurer la périodicité entre plusieurs points séparés par des distances données. Ils reflètent mieux la géométrie que la fonction de corrélation à deux points (c'est-à-dire les statistiques du second ordre) (voir par exemple ([10],

[18], [59], [30]).

De plus, l'extraction de nouvelles propriétés sur les données spatiales conduit à une amélioration significative des performances de la modélisation prédictive. Ainsi, dans plusieurs cas, il est nécessaire d'utiliser des cumulants spatiaux d'ordre supérieur.

Cependant, pour tirer pleinement parti des nombreuses vertus des cumulants spatiaux expérimentaux, nous devons être capables de confondre les cumulants théoriques et expérimentaux (ou échantillon) lorsqu'il s'agit d'un ensemble de données spatiales fini, et d'établir une forte convergence entre eux. Pour cela, nous étudions dans cet ouvrage :

- la convergence presque sûre des cumulants spatiaux du troisième et du quatrième ordre de l'échantillon. De plus, pour un processus linéaire spatial, nous discutons de la normalité asymptotique des estimations de l'échantillon.

- comme cadre pratique des résultats théoriques proposés, certaines méthodes d'identification des systèmes de moyenne mobile unidimensionnels d'ordre fini (systèmes MA RIF) sont étendues au cas 2D, c'est-à-dire au système 2D MA RIF.

Nous nous concentrerons sur les méthodes basées sur des cumulants d'ordre supérieur sous l'hypothèse de non inversibilité. En effet, des études récentes dans de nombreuses applications scientifiques et techniques ont montré que les valeurs des paramètres sont dans la plage de phase non minimale. Une conclusion clé de ces études est que des cumulants d'ordre supérieur sont nécessaires pour l'identification des systèmes à phase non minimale, ce qui implique que l'hypothèse d'erreurs gaussiennes doit être abandonnée. Dans le cas 1-D, la théorie et les techniques de calcul de la prédiction linéaire ont reçu une grande attention dans la théorie des systèmes et le traitement du signal. Cependant, malgré les efforts de nombreux chercheurs, la majeure partie de cette théorie n'a pas été étendue au cas 2-D. La représentation 2D MA pour les champs aléatoires est pertinente en raison de la grande popularité de leur décomposition de type Wold dans le cas homogène régulier. Les applications de la décomposition de type Wold et de la modélisation 2D MA incluent la modélisation des textures naturelles [24], la segmentation d'image et les problèmes de restauration [39], la modélisation et la récupération d'images [44] et les systèmes radar de traitement adaptatif [24].

Le problème d'identification est souvent recomposé comme celui de trouver une solution unique à un système d'équations non linéaires. Les cumulants d'ordre supérieur sont utiles pour l'identification de systèmes de phase éventuellement non minimaux, car la transformée de Fourier des cumulants conjoints d'ordre k d'un champ aléatoire stationnaire d'ordre k X est le polyspectre d'ordre k de X . En effet, à condition que le k -un cumulant d'ordre des erreurs existe et est non nul pour $k \geq 3$, on peut récupérer la fonction de phase à partir de n'importe quel spectre d'ordre k , voir [43] Lemme 1, [29], [27] pour le cas 1D et [71] et [72] pour le multidimensionnel. Par conséquent, deux approches pour l'identification des paramètres de processus 2D MA non gaussiens sont proposées. La première approche est basée sur des cumulants expérimentaux seuls et sans imposer de phase minimum. Alors qu'une méthode généralisée des moments (GMM) est utilisée dans la seconde approche.

La thèse s'articule autour de cinq chapitres :

Dans le premier chapitre, nous donnons les définitions des différents termes liés à l'ensemble de l'étude dans ce travail, et un aperçu sur les différents modèles spatiaux considérés dans le reste des chapitres.

Nous consacrons le chapitre 2 à la problématique des statistiques d'ordre supérieur à travers l'examen de la théorie des moments et des cumulants. Dans ce but, nous décrivons ces outils statistiques pour une variable aléatoire réelle scalaire ainsi que leur extension au cas multidimensionnel, afin de mieux comprendre les multicorrélation de variables aléatoires.

Dans le chapitre 3, nous traitons en détail les modèles 2D moyenne mobile (MA) à réponse impulsionnelle finie et nous mettons le point en particulier sur le cas où la phase est non minimale.

Aux chapitres 4 et 5, nous abordons les propriétés asymptotiques des estimateurs des cumulants et nous présentons deux approches différentes pour l'identification des paramètres du modèles 2D MA RIF, l'une basée sur les cumulants et l'autre sur les estimateurs GMM. Une étude de simulation vérifie les performances des méthodes proposées.

Chapitre 1

Généralités

La statistique spatiale concerne l'étude de phénomènes observés dans un domaine spatial S . Les domaines d'application sont nombreux : sciences de l'environnement et de la terre, épidémiologie, agronomie, météorologie, économétrie, traitement d'images, etc. Cette diversité fait la richesse de ce domaine. On distingue trois types de données dans la statistique spatiale suivant la nature de S : Les données géostatistiques (S est un sous-espace continu de \mathbb{R}^d , $d \geq 2$), Les données latticielles (S est discret et fixé), Les processus ponctuels S est un processus ponctuel dans \mathbb{R}^d . Ils diffèrent d'abord par la nature de l'unité géographique attachée à chaque unité statistique, soit un lieu précis soit une surface, mais aussi par la qualité aléatoire ou non des positions spatiales.

Dans ce chapitre, nous voulons présenter les principaux concepts méthodologiques qui ont été développés pour permettre l'étude des champs aléatoires, ce dernier étant l'outil de probabilité de modélisation des données spatiales.

Le domaine de la statistique spatiale est trop vaste pour être couvert, même superficiellement, par une simple thèse. Nous avons choisi de restreindre notre étude principalement à l'analyse des processus stochastiques spatiaux discrets. Ce chapitre est organisé en deux sections. Dans la première section, nous donnons des définitions, des notations et des outils qui seront utilisés tout au long de la thèse. Dans la section 2, nous présentons des exemples de modèles spatiaux discrets qui seront des modèles de référence pour les modèles que nous étudierons dans les chapitres suivants.

1.1 Concepts et définitions

Dans cette section, nous introduisons les définitions et notations que nous utiliserons tout au long de cette thèse.

1.1.1 Champs aléatoires

Définition 1.1 *D un domaine sur \mathbb{R}^d , un champ aléatoire sur D est une famille de variable aléatoire sur un espace de probabilité (Ω, A, p) ; tel que pour tout s de D est associée X_s ; le quel est noté $\{X_s; s \in D\}$.*

Un champ aléatoire est totalement caractérisé par la distribution conjointe de n'importe quel sous ensemble fini $(X_{s_1}, X_{s_2}, \dots, X_{s_n})$ de manière cohérente nécessitant

$$P(z(t_i) \in A_i, i = 1, \dots, m, z(s)) = P(z(t_i) \in A_i, i = 1, \dots, m).$$

Une telle spécification est appelée la distribution du processus.

Un champ aléatoire discret X est un processus prenant des valeurs aléatoire sur le réseaux.

1.1.2 Stationnarité

En statistique classique, l'inférence des paramètres est rendue possible par la répétition indépendante des données. En statistiques spatiales, on observe très souvent une réalisation unique des données, par exemple un épisode de pollution à l'ozone, une région agricole particulière, une épidémie végétale, etc... Pour pouvoir réaliser l'inférence statistique pour un événement unique, il faut donc en quelques sortes remplacer l'hypothèse sur les répétitions indépendantes par une hypothèse sur le champ aléatoire qui considère d'une part que certaines de ses caractéristiques sont identiques d'un point à l'autre de l'espace, et d'autre part que l'espérance de certaines grandeurs sont accessibles par des intégrales sur l'espace. On pose donc des hypothèses de stationnarité et d'ergodicité. Une propriété importante de la série chronologique est la stationnarité. Cette propriété est nécessaire pour appliquer certain théorème sur la causalité.

Stationnarité stricte

Définition 1.2 *Un champ aléatoire X_s est dit strictement stationnaire si toutes les lois de finidimensionnelles sont invariantes pour toute translation c -à- d :*

$(X_{s_1}, X_{s_2}, \dots, X_{s_n})$ et $(X_{s_1+h}, X_{s_2+h}, \dots, X_{s_n+h})$ sont les mêmes pour tous n -uplet (S_1, S_2, \dots, S_n) et tout $h \in D$.

Cette hypothèse présente deux défauts majeurs, qui font qu'elle est rarement utilisée : d'une part elle ne dit rien sur l'existence ou non des moments, alors que l'espérance et la variance sont des grandeurs nécessaires pour l'analyse statistique de ces données ; et d'autre part il faut vérifier n supérieur au nombre des données disponibles. Nous préférons l'autre hypothèse de stationnarité du second ordre qui est plus faible.

stationnarité du 2ème ordre

Définition 1.3 *Un champ aléatoire intégrable carré, X_s est un champ stationnaire au second ordre sur \mathbb{R}^d Si X_s est de moyenne constante et de covariance invariante par translation*

$$\begin{aligned} E(X_s) &= m \quad \forall s \in \mathbb{R}^d, \\ \text{cov}(X_u, X_v) &= \gamma(u, v) = \text{cov}(X_{u+\tau}, X_{v+\tau}), \end{aligned}$$

pour chaque u, v et $\tau \in \mathbb{R}^d$.

Comme $\gamma(u, v) = \gamma(u - v, 0)$ pour tout $u, v \in \mathbb{R}^d$ pour un champ aléatoire stationnaire faible, il sera commode de redéfinir la fonction d'autocovariance en fonction d'un argument comme suit,

$$\gamma(h) = \gamma(h, 0).$$

Si X est strictement stationnaire et si $X \in L^2$, alors X est stationnaire du second ordre. L'inverse n'est pas vrai en général mais les deux concepts coïncident si X est gaussien.

1.1.3 Champs aléatoires gaussiens

Définition 1.4 *Un champ aléatoire est dit gaussien si toutes ses lois de dimension fini $L(X_{s_1}, X_{s_2}, \dots, X_{s_n})$ sont gaussiennes. Autrement dit $(X_s)_{s \in S}$ est gaussien si et seulement si toute combinaison linéaire $a_1 X_{s_1} + a_2 X_{s_2} + \dots + a_n X_{s_n}$ suit une loi normale, pour tous $n \in \mathbb{N}^*$, $s_1, \dots, s_n \in S$ et $a_1, \dots, a_n \in \mathbb{R}$. La loi d'un vecteur gaussien $(X_{s_1}, X_{s_2}, \dots, X_{s_n})$ est caractérisée par le vecteur des moyennes $(E[X_{s_1}], E[X_{s_2}], \dots, E[X_{s_n}])$ et la matrice de variance-covariance $(Cov(X_{s_i}, X_{s_j}), 1 \leq i, j \leq N)$.*

Exemple 1.1 *-On appelle bruit blanc faible (bb faible) tout processus $(\epsilon_s)_{s \in S}$ tel que $E(\epsilon_s) = 0$, $Var(\epsilon_s) = \sigma^2 < \infty$ décorrélées : $\forall u \neq v, Cov(\epsilon_u, \epsilon_v) = 0$.*

-On appelle bruit blanc fort (bb fort) tout bb faible tel que les variables $(\epsilon_s)_{s \in S}$ sont indépendantes.

-On appelle bruit blanc gaussien (bb gaussien) tout bruit blanc fort tel que $\forall s, \epsilon_s \rightsquigarrow N(0, \sigma^2)$.

Alors : Un bb fort et un bb gaussien sur S sont strictement stationnaire, un bb faible sur S est stationnaire.

1.1.4 Opérateur de retard

Définition 1.5 *L'opérateur de retard B se définit de la manière suivante*

$$B(X_t) = X_{t-1}.$$

Remarque 1.1 *$B^n(X_t) = X_{t-n}$, pour tout $n \in \mathbb{N}$.*

1.1.5 Ergodicité

L'ergodicité est une propriété qui renforce la notion de stationnarité et qui permet d'obtenir la convergence presque sûre d'une moyenne empirique spatiale (d'une réalisation unique) lorsque le domaine "tend vers l'infini".

Définition 1.6 *Un champ aléatoire stationnaire X_s sur \mathbb{R}^d est dit ergodique pour la moyenne (ou ergodique au premier ordre) si la moyenne spatiale sur un domaine D de X_s converge vers son espérance mathématique lorsque $D \rightarrow \mathbb{R}^d$. c'est-à-dire,*

$$\overline{X_D} = \frac{1}{|D|} \int_D X_s ds \rightarrow E(X_0) = \mu, \text{ lorsque } D \rightarrow \mathbb{R}^d.$$

où $|D|$ désigne le d -volmue de D (sa mesure de Lebesgue) et où la convergence est la convergence moyenne quadratique.

Exemple 1.2 $X = \{X_s, s \in \mathbb{Z}^d\}$, X_s i.i.d est un exemple de processus ergodique.

Après avoir introduit les notions de champs aléatoires, de stationnarité et d'ergodicité dans le cas le plus général (champs aléatoires indexés par \mathbb{R}^d), revenons au contexte de ce travail doctoral qui est l'étude des champs aléatoires indexés par \mathbb{Z}^d . Sur \mathbb{Z}^d , on définit dans le paragraphe suivant, l'ordre partiel habituel et l'ordre lexicographique. Ces notions d'ordre vont nous permettre, même artificiellement, de structurer l'espace selon un certain ordre qui jouera le rôle du passé et du présent ou plus précisément de précédance et de succession des observations.

1.1.6 Notions d'ordre dans \mathbb{Z}^d

Si les processus aléatoires sont indexés par \mathbb{Z} , les notions de futur et de passé sont claires et naturels, cependant lorsque le processus aléatoire est indexé par $\mathbb{Z}^d (d \geq 2)$ plusieurs formulation peuvent être défini en fonction de l'ordre utilisé sur \mathbb{Z}^d . Nous donnons ici deux ordres que nous utilisons.

-L'ordre partiel usuel sur $\mathbb{Z}^d (d \geq 2)$: est défini pour deux points $s = (s_1, s_2, \dots, s_n)$ et $t = (t_1, t_2, \dots, t_n)$ de \mathbb{Z}^d par

$$s \preceq t \quad (\text{respectivement } s \prec t),$$

si pour tout $i = 1, \dots, d$

$$s_i \leq t_i \quad (\text{respectivement } s_i < t_i).$$

Cet ordre est appelé l'ordre des quarts de plan.

-L'ordre lexicographe sur \mathbb{Z}^d : est défini pour deux points $s = (s_1, s')$ et $t = (t_1, t')$ de \mathbb{Z}^d tel que s_1 et t_1 appartiennent à \mathbb{Z} par

$$s \prec\prec t$$

si et seulement si $s_1 \prec t_1$ et $s' \prec\prec t'$ dans \mathbb{Z}^{d-1} .

Cet ordre est appelé l'ordre du demi-plan.

Remarque 1.2 1/ Soient $i = (k; l), j = (m; n)$ deux points dans \mathbb{Z}^2 . On note par $i \leq j$ l'ordre lexicographique défini par $(i \leq j) \iff (k < m) \text{ ou } (k = m; l \leq n)$.

Cet ordre est appelé aussi ordre demi-plan ou ordre total.

2/ Soient $i = (k; l), j = (m; n)$ deux points dans \mathbb{Z}^2 : On appelle ordre quart $i \ll j \iff (k \leq m) \text{ et } (l \leq n)$. Ce qui est considéré comme un cas particulier de l'ordre total.

1.1.7 Indexation des sous-ensembles

Pour deux éléments $a, b \in \mathbb{Z}^d$ tels que $a \leq b$ et $a \neq b$, on note

$$\begin{aligned} s[a, b] &= \{x \in \mathbb{Z}^d : a \leq x \leq b\}, \\ s[a, \infty] &= \{x \in \mathbb{Z}^d : a \leq x\}, \\ s[\infty, a] &= \{x \in \mathbb{Z}^d : x \leq a\}, \\ s \langle a, b \rangle &= s[a, b] \setminus \{a\}, \\ s \langle a, \infty \rangle &= s[a, \infty] \setminus \{a\}. \end{aligned}$$

$s[a, b]$ permet de définir une notion de fenêtre spatiale qui étend la notion classique d'intervalle de temps utilisée dans la période, $s[a, \infty]$ généralise la notion de futur à un cadre spatial, enfin $s[\infty, a]$ permet de donner un équivalent à la notion de passé d'un temps d'observation lorsqu'il s'agit d'un site observé a .

1.2 Modèles pour les données spatiales

Le but de cette section est de donner un aperçu des modèles ARMA, AR, MA, nous l'utilisons comme point de départ.

1.2.1 Modèle ARMA spatiaux

Les modèles ARMA bidimensionnels ont été largement utilisés en spectroscopie d'absorption, traitement des données astronomiques et dans diverses tâches de traitement d'images, mais les techniques d'estimation des paramètres ont été largement limitées au cas de l'excitation gaussienne [Jain 1989; Kashyap 1984], avec les restrictions qui en découlent telles que les modèles symétriques non causaux.

Etant donné une série temporelle X_t , le modèle ARMA est un outil pour comprendre et prédire, éventuellement les valeurs futures de cette série. Le modèle est composé de deux parties : une partie autorégressive (AR) et une partie moyenne mobile (MA). Le modèle est généralement noté par $ARMA(p, q)$, où p est l'ordre de partie AR et q l'ordre de la partie MA. Herman Wold [78] a montré que les processus ARMA pouvaient être utilisés pour modéliser n'importe quelle série stationnaire pour autant que les ordres p et q soient bien choisis. Box et Jenkins [11] ont travaillé pour développer une méthodologie pour l'estimation du modèle d'une série chronologique.

Définition 1.7 *Un champ de carré intégrable (X_s) est appelé modèle ARMA(p, q) spatial de paramètre $p, q \in \mathbb{N}^d$, s'il satisfait l'équation*

$$X_t - \sum_{j \in s \prec (0, p)} \varphi_j X_{t-j} = \varepsilon_s + \sum_{k \in s \prec (0, q)} b_k \varepsilon_{t-k} \quad \forall t \in \mathbb{Z}^d, \quad (1.1)$$

$$(1 - \sum_{j \in s \prec 0, p} \varphi_j B^j) X_t = (1 + \sum_{k \in s \prec 0, q} b_k B^k) \varepsilon_t$$

Où $(\varphi_j)_{j \in s \prec 0, p}$ et $(b_k)_{k \in s \prec 0, q}$ désigne respectivement les paramètres de moyenne autorégressive et mobile avec $\varphi_0 = b_0 = 1$ et $(\varepsilon_s)_{s \in \mathbb{Z}^d}$ est un champ de carré intégrable stationnaire, de variance et vérifiant

$$E(\varepsilon_s) = 0 \quad \text{et} \quad E(\varepsilon_t \varepsilon_s) = 0 \quad \forall s, t \in \mathbb{Z}^d.$$

Si $(q = 0)$ (resp $p = 0$), la somme sur $S \prec 0, q]$ (resp sur $S \prec 0, p]$ est supposée nulle et le champ est appelé champ $AR(p)$ (resp $MA(q)$).

Si $S \prec 0, q]$ et $S \prec 0, p]$, sont définis à partir de l'ordre lexicographique (également appelé ordre de demi plan) le modèle $ARMA$ es appelé demi plan unilatéral.

Définition 1.8 *Un champ $ARMA(p, q)$ $(X_t)_{t \in \mathbb{Z}}$ est dit causal s'il admet une expression unilatérale du type*

$$X_t = \sum_{j \in s[0, \infty]} \psi_j \varepsilon_{t-j} \quad \psi_0 = 1 \quad \forall t \in \mathbb{Z}^d,$$

avec

$$\sum_{j \in s[0, \infty]} |\psi_j| < \infty.$$

Un champ $ARMA$ causal est donc stationnaire.

Notation 1.1 -Polynôme autorégressif :

$$\phi(z) = 1 - \sum_{j \in s[0, p]} \varphi_j z^j \quad z = z_1, z_2, \dots, z_d \in \mathbb{C}^d$$

-Polynôme moyenne mobile :

$$\Theta(z) = 1 + \sum_{k \in s[0, p]} b_k z^k \quad z = z_1, z_2, \dots, z_d \in \mathbb{C}^d.$$

1.2.2 Modèle AR spatial

Nous présentons la définition de modèle $AR(p)$ quadrantal de Tjøstheim ([73], [74]).

Définition 1.9 Un champ de carré intégrable $(X)_{t \in \mathbb{Z}^d}$ est appelé modèle $AR(p)$ spatial de paramètre p dans \mathbb{N}^d s'il satisfait l'équation

$$X_t = \sum_{j \in s \prec (0,p)} \varphi_j X_{t-j} + \varepsilon_t, \quad (1.2)$$

où $(\varepsilon_t)_{t \in \mathbb{Z}^d}$ est un champ de carré intégrable stationnaire au second ordre, de variance $\sigma^2 \succ 0$ et vérifiant $E(\varepsilon_s) = 0$ et $E(\varepsilon_t \varepsilon_s) = 0 \forall s, t \in \mathbb{Z}^d$.

Définition 1.10 Un champ $AR(p)$ $(X_t)_{t \in \mathbb{Z}}$ est dit causal s'il admet une expression unilatérale du type

$$X_t = \sum_{j \in s[0, \infty]} \psi_j \varepsilon_{t-j},$$

avec

$$\sum_{j \in s[0, \infty]} |\psi_j| \prec \infty.$$

Remarque 1.3 Un champ AR causal est donc stationnaire.

1.2.3 Modèle MA spatial

Définition 1.11 Un champ de carré intégrable $(X)_{t \in \mathbb{Z}^d}$ est appelé modèle $MA(q)$ spatial de paramètre q dans \mathbb{N}^d s'il satisfait l'équation

$$X_t = \sum_{k \in s \prec (0,p)} b_k \varepsilon_{t-k} + \varepsilon_t. \quad (1.3)$$

Remarque 1.4 -La définition d'un $MA(q)$ est explicite et ne pose donc pas de problème : le processus X_t est parfaitement défini et est automatiquement stationnaire.

-La représentation est causale par définition.

Définition 1.12 Un champ $MA(q)$ est dit inversible s'il admet une expression du type

$$\varepsilon_t = \sum_{j \in s[0, \infty]} \Phi_j X_{t-j},$$

avec

$$\sum_{j \in s[0, \infty]} |\Phi_j| < \infty.$$

Théorème 1.1 *Un processus à moyenne mobile est inversible si et seulement si son polynôme $\theta(z)$ est tel que $\theta(z) \neq 0$ avec $z \in \mathbb{C}$ tel que $|z| \leq 1$.*

Modèles moyenne mobile sur \mathbb{Z}^2

Définition 1.13 *On dit que le processus $X_{i,j}$, $(i, j) \in \mathbb{Z}^2$ est un processus moyenne mobile sur \mathbb{Z}^2 , s'il satisfait à une équation de récurrence de la forme :*

$$X_{ij} = \sum_{k=-m}^m \sum_{l=-n}^n b_{k,l} \varepsilon_{i-k, j-l} + \varepsilon_{ij}, \quad (1.4)$$

avec $(k, l) \neq (0, 0)$, $m, n \in \mathbb{Z}$ et ε_{ij} est une suite de variables aléatoires indépendantes et de même loi.

Chapitre 2

Statistiques d'Ordre Supérieur

2.1 Introduction aux Statistiques d'Ordre Supérieur

Les statistiques d'ordre supérieur à 2, autrement dit les moments et cumulants d'ordre supérieur à 2, sont utilisées essentiellement en complément aux statistiques d'ordre 2. Elles donnent une description plus complète des données et de leurs propriétés. Les statistiques d'ordre supérieur permettent la résolution de problèmes insolubles à l'ordre 2.

Ont joué un rôle croissant dans les applications de traitement de signal, elles sont utilisées pour la modélisation des signaux et pour l'identification des systèmes à phase non-minimale et des systèmes non-causaux ([47], [26], [28]).

Ce chapitre a pour but de donner les définitions et les propriétés nécessaires à l'introduction des statistiques d'ordre supérieur. Nous proposons également quelques éléments de la théorie de l'estimation des statistiques d'ordre supérieur.

Ce chapitre est organisé en trois sections, la section 1 : Nous définissons les moments et les cumulants d'une variable aléatoire scalaire réelle et les relations reliant ces deux grandeurs. Dans la section 2 est présentée l'extension de ces outils au cas réel multidimensionnel, et dans la section trois on définit les multicorrélations qui sont issues des statistiques d'ordre supérieur de variables aléatoires multidimensionnelles.

Soit $\{x\}$ un processus aléatoire stationnaire, et x une variable aléatoire associée à ce processus. Pour plus de clarté, nous considérons tout d'abord le cas d'une variable aléatoire à valeurs réelles. Puis, on généralisera ce résultat au cas d'une variable aléatoire multidimensionnelle.

Les différents outils statistiques présentés dans ce chapitre ont été étudiés en détail dans divers ouvrages [[9], [36], [62], [13], [20], [52], [76]].

2.2 Variables aléatoires réelles scalaires

2.2.1 Fonctions caractéristiques d'une variable aléatoire réelle scalaire

Lorsque la v.a x admet une densité de probabilité $f_x(x)$, la première fonction caractéristique $P_x(u)$ est définie comme sa transformée de Fourier

$$P_x(u) = E(e^{jux}) = \int_{-\infty}^{+\infty} f_x(x)e^{jux} dx, \quad (2.1)$$

où j désigne la racine carrée de -1 .

La fonction caractéristique est continue en tout point et vaut 1 à l'origine. Elle est donc non nulle dans un voisinage de l'origine, sur lequel on pourra définir son logarithme népérien

$$P'_x(u) = \log P_x(u). \quad (2.2)$$

Cette fonction est appelée seconde fonction caractéristique.

2.2.2 Moments et cumulants d'une variable aléatoire réelle scalaire

Moments

Définition 2.1 Les moments d'ordre k de la variable aléatoire x , s'ils existent [9], sont définis par

$$r_{x(k)} = E(x^k) = \int_{-\infty}^{+\infty} f_x(u)u^k dx. \quad (2.3)$$

Ces moments sont appelés moments "par rapport à l'origine 0". La moyenne arithmétique est le moment d'ordre 1, $r_{x(1)}$.

Définition 2.2 Les moments d'ordre k par rapport à un point quelconque $\alpha \in \mathbb{R}$ sont définis par

$$r_{x(k)} |_{\alpha} = E[(x - \alpha)^k] = \int_{-\infty}^{+\infty} (u - \alpha)^k f_x(u) du . \quad (2.4)$$

Définition 2.3 Les moments définis par l'équation (2.4), calculés par rapport à la valeur moyenne, donc pour $\alpha = r_{x(1)}$, conduisent à ce qu'on appelle les « moments centrés »

$$r'_{x(k)} = E [(x - r_{x(1)})^k] . \quad (2.5)$$

Les relations entre les moments centrés d'ordre k et les moments d'ordre k par rapport à l'origine sont :

$$\begin{aligned} r'_{x(k)} &= \sum_{i=0}^k \frac{k!}{i!(k-i)!} r_{x(k-i)} (-r_{x(1)})^i, \\ r_{x(k)} &= \sum_{i=0}^k \frac{k!}{i!(k-i)!} r'_{x(k-i)} (-r'_{x(1)})^i. \end{aligned}$$

Par exemple, les premiers moments centrés en fonction des moments d'ordre inférieur ou égal à k par rapport à l'origine sont :

$$\begin{aligned}
 r'_{x(1)} &= 0, \\
 r'_{x(2)} &= r_{x(2)} - (r_{x(1)})^2, \\
 r'_{x(3)} &= r_{x(3)} - 3r_{x(2)}r_{x(1)} + 2(r_{x(1)})^3, \\
 r'_{x(4)} &= r_{x(4)} - 4r_{x(3)}r_{x(1)} + 6r_{x(2)}(r_{x(1)})^2 - 3(r_{x(1)})^4.
 \end{aligned}
 \tag{2.6}$$

Il existe une deuxième manière de définir les moments d'une variable aléatoire : à l'aide de la première fonction caractéristique définie en (2.1).

Théorème 2.1 *Le moment d'ordre k par rapport à l'origine de la variable aléatoire x est la dérivée d'ordre k de la première fonction caractéristique calculée en $u = 0$ [41] :*

$$r_{x(k)} = E(x^k) = (-j)^k \left[\frac{dkP_x(u)}{du^k} \right]_{u=0}.
 \tag{2.7}$$

Définition 2.4 *Les moments d'ordre k par rapport à l'origine sont les coefficients de $\frac{(ju)^k}{k!} du$ développement en série de Maclaurin (série de Taylor au voisinage de l'origine) de la première fonction caractéristique :*

$$P_x(u) = \sum_{k=1}^{\infty} \frac{(ju)^k}{k!} r_{x(k)}.$$

Ainsi, la première fonction caractéristique est appelée "fonction génératrice de moments".

cumulants

Définition 2.5 *Formellement, les cumulants, notés $C_{x(1)}, C_{x(2)}, \dots, C_{x(k)}$, de la variable aléatoire x sont définis par l'identité en u de l'équation [36],*

$$\exp \left\{ C_{x(1)}(ju) + C_{x(2)} \frac{(ju)^2}{2!} + \dots + C_{x(k)} \frac{(ju)^k}{k!} + \dots \right\} = 1 + r_{x(1)} \frac{(ju)}{1!} + \dots + r_{x(k)} \frac{(ju)^k}{k!} + \dots = P_x(u).
 \tag{2.8}$$

Comme à la section antérieure, il existe une deuxième manière de définir les cumulants d'une variable aléatoire.

Théorème 2.2 *Le cumulants d'ordre k de la variable aléatoire x est la dérivée d'ordre k de la seconde fonction caractéristique calculée en $u = 0$ [41] :*

$$C_{x(k)} = (-j)^k \left[\frac{d^k P'_x(u)}{du^k} \right]_{u=0}. \quad (2.9)$$

Théorème 2.3 *Les cumulants d'ordre k sont les coefficients de $\frac{(ju)^k}{k!}$ du développement en série de Maclaurin (série de Taylor en voisinage de l'origine) de la seconde fonction caractéristique :*

$$P'_x(u) = \sum_{k=1}^{\infty} \frac{(ju)^k}{k!} C_{x(k)}.$$

La seconde fonction caractéristique est donc une fonction génératrice des cumulants, d'où la dénomination de fonction cumulative.

2.2.3 Relations entre moments et cumulants

Théorème 2.4 *Les cumulants d'ordre k , $C_{x(k)}$, sont liés aux moments par rapport à l'origine d'ordre inférieur ou égal à k par la relation [36] :*

$$C_{x(k)} = k! \sum_{m=0}^k \sum \frac{(-1)^{(p-1)}(p-1)!}{\pi_1! \dots \pi_m!} \left(\frac{r_{x(p_1)}}{p_1!} \right)^{\pi_1} \dots \left(\frac{r_{x(p_m)}}{p_m!} \right)^{\pi_m}, \quad (2.10)$$

où la deuxième somme est effectuée pour toutes les valeurs non-négatives de π_i et p pour lesquelles $p_1\pi_1 + \dots + p_m\pi_m$ et $\pi_1 + \pi_2 + \dots + \pi_m = p$.

-Les cumulants d'ordre k peuvent être calculés à partir des moments d'ordre inférieur ou égal

à k . Pour les ordres 1 à 4, ces relations sont :

$$\begin{aligned}C_{x(1)} &= r_{x(1)}, \\C_{x(2)} &= r_{x(2)} - (r_{x(1)})^2, \\C_{x(3)} &= r_{x(3)} - 3r_{x(2)}r_{x(1)} + 2(r_{x(1)})^3, \\C_{x(4)} &= r_{x(4)} - 4r_{x(3)}r_{x(1)} - 3(r_{x(2)})^2 + 12(r_{x(1)})^2r_{x(2)} - 6(r_{x(1)})^4.\end{aligned}\tag{2.11}$$

-Dans le cas de variables aléatoires centrées ($r_x(1) = 0$), les expressions des cumulants se simplifient en

$$\begin{aligned}C_{x(1)} &= 0, \\C_{x(2)} &= r_{x(2)}, \\C_{x(3)} &= r_{x(3)}, \\C_{x(4)} &= r_{x(4)} - 3(r_{x(2)})^2.\end{aligned}\tag{2.12}$$

Lorsque la variable aléatoire x est gaussienne, sa seconde fonction caractéristique est un polynôme d'ordre 2, ce qui implique la nullité de tous ses cumulants d'ordre supérieur ou égal à 3. Cette propriété caractérise la loi gaussienne. Les variables gaussiennes sont donc entièrement décrites par leurs propriétés au second ordre. Ceci explique pourquoi les chercheurs en traitement du signal se sont longtemps limités au second ordre.

Remarques.

1/ Les moments centrés et les cumulants sont identiques jusqu'à l'ordre 3 inclus (voir également les équations (2.6) et (2.11)). A l'exception de ce cas, le calcul des cumulants d'ordre k nécessite le calcul des moments d'ordre inférieur ou égal à k .

2/. Le cumulants d'ordre un est la moyenne ; le cumulants d'ordre deux est la variance (moment d'ordre deux après soustraction du carré de la moyenne).

3/ Le cumulants d'ordre trois (skewness en anglais, asymétrie en français) est nul lorsque la variable aléatoire est centrée et sa densité de probabilité est symétrique.

4/ Le cumulante d'ordre quatre appelé (kurtosis), la valeur du kurtosis caractérise le comportement à l'infini des densité de probabilité.

2.2.4 Estimation des moments et des cumulants

Soit x une variable aléatoire réelle scalaire et N réalisations indépendantes de x notées x_1, \dots, x_N

Estimateurs des moments

Un estimateur des moments d'ordre k par rapport à l'origine de la variable x est :

$$\hat{r}_{x(k)} = \frac{1}{N} \sum_{n=1}^N x_n^k. \quad (2.13)$$

Théorème 2.5 *L'estimateur $\hat{r}_{x(k)}$ a les propriétés suivantes [62] :*

- $\hat{r}_{x(k)}$ est non biaisé : $E(\hat{r}_{x(k)}) = r_{x(k)}$.

- La variance de l'estimateur est : $\text{var} [r_{x(k)}] = \frac{1}{N} \left(r_{x(2k)} - (r_{x(k)})^2 \right)$.

- L'estimateur des moments est un estimateur consistant car $\lim_{N \rightarrow \infty} \text{var} [\hat{r}_{x(k)}] = 0$, donc l'estimateur $\hat{r}_{x(k)}$ converge vers $r_{x(k)}$ avec une probabilité de 1 quand $N \rightarrow \infty$.

- $\hat{r}_{x(k)}$ est asymptotiquement normal.

Un estimateur des moments centrés d'ordre k de la variable aléatoire x est :

$$\hat{r}'_{x(k)} = \frac{1}{N} \sum_{n=1}^N (x_n - \hat{r}_{x(1)})^k.$$

Il a été démontré que $\hat{r}'_{x(k)}$ est un estimateur non biaisé de $r'_{x(k)}$ en négligeant les termes d'ordre $N^{-\frac{1}{2}}$ [36].

Estimateurs des cumulants

Un estimateur des cumulants d'ordre k est obtenu en remplaçant dans l'équation (2.10) les moments par leurs estimateurs :

$$C_{x^{(k)}} = k! \sum_{m=0}^k \sum \frac{(-1)^{(p-1)}(p-1)!}{\pi_1! \dots \pi_m!} \left(\frac{\hat{r}_{x^{(p_1)}}}{p_1!}\right)^{\pi_1} \dots \left(\frac{\hat{r}_{x^{(p_m)}}}{p_m!}\right)^{\pi_m}. \quad (2.14)$$

Cependant, ces estimateurs sont généralement biaisés. comme le montre l'exemple ci-dessous.

Exemple 2.1 Soit x une variable aléatoire centrée (donc $r_{x^{(1)}} = 0$). A partir de la quatrième équation donnée dans les relations (2.12), le cumulante d'ordre 4 de cette variable s'écrit :

$$C_{x^{(4)}} = r_{x^{(4)}} - 3(r_{x^{(2)}})^2.$$

A partir de l'équation (2.14), un estimateur de ce cumulante est donc

$$\hat{C}_{x^{(4)}} = \hat{r}_{x^{(4)}} - 3(\hat{r}_{x^{(2)}})^2.$$

En utilisant les estimateurs donnés par l'équation (2.13) nous obtenons :

$$\begin{aligned} \hat{C}_{x^{(4)}} &= \frac{1}{N} \sum_{n=1}^N x_n^4 - 3 \left(\frac{1}{N} \sum_{n=1}^N x_n^2 \right)^2 \\ \hat{C}_{x^{(4)}} &= \frac{1}{N} \sum_{n=1}^N x_n^4 - 3 \left(\frac{1}{N^2} \sum_{n,m=1}^N x_n^2 x_m^2 \right) \\ \hat{C}_{x^{(4)}} &= \frac{1}{N} \sum_{n=1}^N x_n^4 - \frac{3}{N^2} \sum_{n=m=1}^N x_n^2 x_m^2 - \frac{3}{N^2} \left(\sum_{n=1}^N x_n^2 \sum_{m=1, n \neq m}^N x_m^2 \right). \end{aligned}$$

Grâce aux propriétés développées dans le théorème (2.5), la moyenne de $\hat{C}_{x^{(4)}}$ est donnée par :

$$E[\hat{C}_{x^{(4)}}] = \left(N \frac{1}{N} - \frac{3N}{N^2} \right) r_{x^{(4)}} - N(N-1) \frac{3}{N^2} (r_{x^{(2)}})^2.$$

En utilisant les équations (2.12) :

$$E[\hat{C}_{x(4)}] = C_{x(4)} - \frac{3}{N}(C_{x(4)} + 2(r_{x(2)})^2),$$

ce qui conduit à un estimateur biaisé du $C_{x(4)}$.

Cependant, ces estimateurs sont asymptotiquement non biaisés, puisque le biais tend vers 0 lorsque $N \rightarrow \infty$.

Dans les situations pratiques où le nombre de données est grand, $\hat{C}_{x(4)}$ est un estimateur approximativement non biaisé. Toutefois, si l'échantillon est de taille petite taille, le biais peut devenir gênant. Pour ces situations, il existe des estimateurs non biaisés des cumulants : les k -statistiques.

-Estimateurs non biaisés des cumulants : les k -statistiques

Des estimateurs non biaisés des cumulants sont les k -statistiques [36], [20]. Les k -statistiques sont des fonctions symétriques qui ont été introduites par Sir Ronald Fisher pour caractériser les distributions.

Par exemple, en notant $s_r = \sum_{n=1}^N x_n^r$, les premières k -statistiques sont :

$$\hat{C}_{x(1)} = \frac{1}{N}s_1$$

$$\hat{C}_{x(2)} = \frac{1}{N(N-1)} [Ns_2 - (s_1)^2]$$

$$\hat{C}_{x(3)} = \frac{1}{N(N-1)(N-2)} [N^2s_3 - 3Ns_1s_2 - +2(s_1)^2]$$

$$\hat{C}_{x(4)} = \frac{1}{N(N-1)(N-2)(N-3)} [(N^3 + N^2)s_4 - 4(N^2 + N)s_1s_3 - 3(N^2 + N)(s_2)^2 + 12Ns_2(s_1)^2 - 6(s_1)^4].$$

Théorème 2.6 Les k -statistiques $C_{x(k)}$ ont les propriétés suivantes [21] :

- $C_{x(k)}$ sont des estimateurs non biaisés des cumulants : $E(C_{x(k)}) = r_{x(k)}$.

-Les variances de $C_{x(k)}$ jusqu'à l'ordre 4 sont :

$$\text{var} [C_{x(2)}] = \frac{1}{N}C_{x(4)} - \frac{2}{N-1} (C_{x(2)})^2.$$

$$\text{var} [C_{x(3)}] = \frac{1}{N}C_{x(6)} + \frac{9}{N-1} \left(C_{x(4)}k_{x(2)} + (C_{x(3)})^2 \right) + \frac{6N}{(N-1)(N-2)} (C_{x(2)})^3.$$

$$\text{var} [C_{x(4)}] = \frac{1}{N}C_{x(8)} + \frac{1}{N-1} \left(16C_{x(6)}C_{x(2)} + 48C_{x(5)}C_{x(3)} + 34(C_{x(4)})^2 \right) + \frac{8N}{(N-1)(N-2)} 9C_{x(4)}(C_{x(2)})^2 + 18(C_{x(3)})^2C_{x(2)} + \frac{24N(N+1)}{(N-1)(N-2)(N-3)} (C_{x(2)})^4.$$

-Les k -statistiques sont des estimateurs consistants puisque $\lim_{N \rightarrow \infty} \text{var} [C_{x(k)}] = 0$, donc l'estimateur $C_{x(k)}$ converge vers 0 avec une probabilité de 1 quand $N \rightarrow \infty$.

-L'estimateur des moments est un estimateur consistant car $\lim_{n \rightarrow \infty} \text{var} [\hat{r}_{x(k)}] = 0$, donc l'estimateur $\hat{r}_{x(k)}$ converge vers $r_{x(k)}$ avec une probabilité de 1 quand $N \rightarrow \infty$.

- $C_{x(k)}$ est asymptotiquement normal.

2.2.5 Exemples de variables aléatoires réelles

- Variable gaussienne

Soit x une variable gaussienne de moyenne α et variance σ^2 , notée $X \rightarrow N(\alpha, \sigma^2)$.

La densité de probabilité $f_x(u) = \frac{1}{\sigma\sqrt{2\pi}} \exp -\left(\frac{u-\alpha}{2\sigma^2}\right)^2$.

La fonction de répartition $F_x(u) = \frac{1}{2} \left[1 + h\left(\frac{u-\alpha}{\sigma\sqrt{2}}\right) \right]$ où $h(z) = \frac{2}{\pi} \int_0^{\pi} e^{-t^2} dt$

La première fonction caractéristique $P_x(t) = e^{j\alpha t - \frac{\sigma^2 t^2}{2}}$.

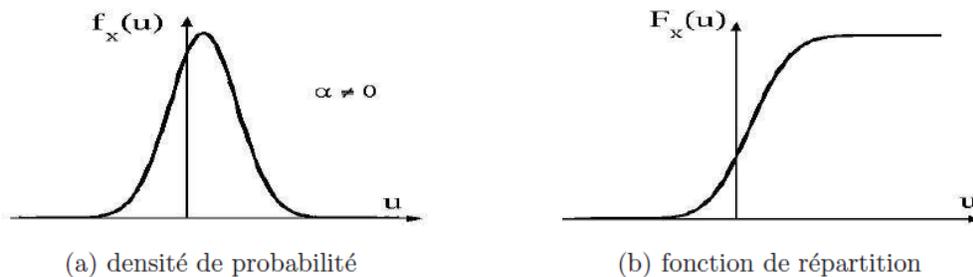


Figure 1 : Variable gaussienne.

Les moments, les moments centrés et les cumulants jusqu'à l'ordre 4 d'une variable aléatoire gaussienne sont donnés dans le tableau.

- Les moments centrés d'ordre impair sont nuls, $r_{x(2k+1)} = 0$, et les moments centrés d'ordre pair sont, $r_{x(2k)} = \sigma^{2k} \frac{(2k)!}{k!2^k}$.

- Les cumulants d'ordre supérieur à 2 d'une variable gaussienne sont nuls.

L'ordre k	moment	moment centré	cumulant
1	α	0	α
2	$\alpha^2 + \sigma^2$	σ^2	σ^2
3	$\alpha(\alpha^2 + 3\sigma^2)$	0	0
4	$\alpha^4 + 6\alpha^2\sigma^2 + 3\sigma^4$	$3\sigma^4$	0

- Variable exponentielle

Soit x une variable aléatoire positive distribuée selon la loi exponentielle.

- La densité de probabilité $f_x(u) = \begin{cases} \lambda e^{-u\lambda} & , u \geq 0 \\ 0 & , u < 0 \end{cases}$
- La fonction de répartition $F_x(u) = \begin{cases} 1 - e^{-u\lambda} & , u \geq 0 \\ 0 & , u < 0 \end{cases}$

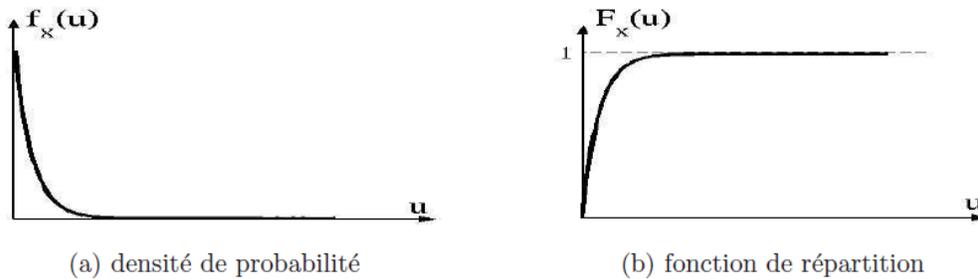


Figure 2 : Variable exponentielle.

-La première fonction caractéristique $\phi_x(t) = \frac{j}{j+\sigma t}$ avec $\sigma = \frac{1}{\lambda}$.

Les moments et les cumulants jusqu'à l'ordre 4 d'une variable aléatoire exponentielle sont données dans le tableau suivant :

L'ordre k	moment	moment centré	cumulant
1	σ	0	σ
2	$2\sigma^2$	σ^2	σ^2
3	$6\sigma^3$	$2\sigma^3$	$2\sigma^3$
4	$24\sigma^4$	$9\sigma^4$	$6\sigma^4$

Les moments d'ordre k par rapport à l'origine vérifient la relation $r_x(k) = k!\sigma^k$, et les cumulants d'ordre k vérifient la relation $C_x(k) = (k - 1)!\sigma^k$.

Après avoir défini les moments et les cumulants d'une variable aléatoire réelle scalaire, nous étudierons à la section suivante l'extension de ces outils au cas réel multidimensionnel.

2.3 Variables aléatoires réelles multidimensionnelles

2.3.1 Fonctions caractéristiques des variables aléatoires réelles multidimensionnelles

On peut représenter les variables aléatoires à plusieurs dimensions par le vecteur colonne

$$x^T = (x_1, \dots, x_N).$$

De la même façon que dans le cas scalaire, on définit la fonction caractéristique conjointe de N variables aléatoires x_n par la relation

$$P_x(u) = E\left(e^{j \sum_n u_n x_n}\right) = E(e^{ju^T x}), \quad (2.15)$$

où $u^T = (u_1, \dots, u_N)$. Si les composantes x_n du vecteur aléatoire x admettent une densité conjointe $f_x(v)$, alors la fonction caractéristique de x est donnée par la transformée de Fourier de cette densité.

$$P_x(u) = \int_{-\infty}^{+\infty} f_x(v) e^{ju^T v} dv. \quad (2.16)$$

Comme précédemment, la seconde fonction caractéristique est $P'_x(u) = \log P_x(u)$. Les fonctions caractéristiques peuvent servir à générer les moments et les cumulants.

2.3.2 Moments et cumulants de vecteurs aléatoires

Moments

Définition 2.6 *Les moments croisés d'ordre k du vecteur aléatoire x sont définis de la même manière que dans le cas scalaire :*

$$r_{x(i_1, i_2, \dots, i_k)} = E(x_{i_1} x_{i_2} \dots x_{i_k}) = \int_{\mathbb{R}^n} f_x(\mathbf{u}) u_{i_1} u_{i_2} \dots u_{i_k} d\mathbf{u}. \quad (2.17)$$

Théorème 2.7 *Le moment croisé d'ordre k du vecteur aléatoire x est la dérivée d'ordre k de la première fonction caractéristique calculée en $u = 0$:*

$$r_{x(i_1, i_2, \dots, i_k)} = (-j)^k \left[\frac{d^k P_x(u)}{du_{i_1} du_{i_2} \dots du_{i_k}} \right]_{u=0}. \quad (2.18)$$

Définition 2.7 *Les moments croisés d'ordre k sont les coefficients de $(j)^k \frac{u_{i_1} u_{i_2} \dots u_{i_k}}{k!}$ du développement en série de Maclaurin de la première fonction caractéristique :*

$$P_x(u) = \sum_{k=1}^{\infty} \frac{(j)^k u_{i_1} u_{i_2} \dots u_{i_k}}{k!} r_{x(i_1, i_2, \dots, i_k)}.$$

Après avoir donné les définitions des moments de vecteur aléatoire réelle, nous nous intéressons maintenant à celles de ses cumulants. Ces grandeurs statistiques, équivalentes aux moments, sont préférées dans de nombreux traitements utilisant les statistiques d'ordre supérieur.

cumulants

Comme dans le cas scalaire, les cumulants sont définis formellement par une identité. Pour les vecteurs aléatoires, nous utiliserons une définition équivalente à celle donnée pour les moments.

Définition 2.8 *Les cumulants croisés d'ordre k sont les coefficients de $(j)^k \frac{u_{i_1} u_{i_2} \dots u_{i_k}}{k!}$ du développement en série de Maclaurin de la seconde fonction caractéristique :*

$$P'_x(u) = \sum_{k=1}^{\infty} \frac{(j)^k u_1^{k_1} u_2^{k_2} \dots u_k^{k_n}}{k!} C_{x(i_1, i_2, \dots, i_K)},$$

avec $k_1, k_2, \dots, k_n \in \mathbb{N}$, tel que $k = k_1 + k_2 + \dots + k_n$.

Théorème 2.8 *Le cumulant croisé d'ordre k du vecteur aléatoire x est la dérivée d'ordre k de la deuxième fonction caractéristique calculée en $u = 0$:*

$$C_{x(i_1, i_2, \dots, i_k)} = (-j)^k \left[\frac{d^k P'_x(u)}{du_{i_1} du_{i_2} \dots du_{i_k}} \right]_{u=0} \quad (2.19)$$

Notation 2.1 Les cumulant d'ordre k sont aussi notés

$$C_{x(i_1, i_2, \dots, i_k)} = cum[x_{i_1}, x_{i_2}, \dots, x_{i_k}]$$

Dans cette thèse nous utilisons, en fonction du contexte, les deux notations.

Définition 2.9 La matrice de covariance $C'xx$ d'un vecteur aléatoire n dimensionnel $x^T = (x_1, x_2, \dots, x_n)$ est définie par :

$$C'xx = E[(x - E[x])(x - E[x])^T] = cum([x_i; x_j] / 1 \leq i \leq n; 1 \leq j \leq n).$$

2.3.3 Relations entre moments et cumulants

Théorème 2.9 Les cumulants croisés d'un vecteur aléatoire n dimensionnel, $Cum[x_1, x_2, \dots, x_n]$, sont liés aux moments croisés d'ordre inférieur ou égal à n par la relation

$$cum[x_1, x_2, \dots, x_n] = \sum (-1)^{q-1} (q-1)! E \left[\prod_{i \in v_1} x_i \right] E \left[\prod_{j \in v_2} x_j \right] \dots E \left[\prod_{k \in v_p} x_k \right]. \quad (2.20)$$

où la sommation s'étend sur tous les ensembles $\{v_1, v_2, \dots, v_p : 1 \leq p \leq n\}$ formant une partition de $\{1, 2, \dots, n\}$. et q est le nombre d'éléments composant la partition.

Par exemple, les cumulants croisés s'expriment en fonction des moments croisés par :

$$\begin{aligned} cum[x_i] &= E[x_i], \\ cum[x_i, x_j] &= E[x_i x_j] - E[x_i] E[x_j], \\ cum[x_i, x_j, x_k] &= E[x_i x_j x_k] - E[x_i] E[x_j x_k] - E[x_j] E[x_i x_k] - E[x_k] E[x_i x_j] + 2E[x_i] E[x_j] E[x_k], \\ cum[x_i, x_j, x_k, x_l] &= E[x_i x_j x_k x_l] - E[x_i] E[x_j x_k x_l] - E[x_j] E[x_i x_k x_l] - E[x_k] E[x_i x_j x_l] \\ &\quad - E[x_l] E[x_i x_j x_k] - E[x_i x_j] E[x_k x_l] + 2E[x_i x_j] E[x_k] E[x_l] \\ &\quad + 2E[x_i x_k] E[x_j] E[x_l] + 2E[x_i x_l] E[x_j] E[x_k] + 2E[x_j x_k] E[x_i] E[x_l] \\ &\quad + 2E[x_j x_l] E[x_i] E[x_k] + 2E[x_k x_l] E[x_i] E[x_j] - 6E[x_i] E[x_j] E[x_k] E[x_l]. \end{aligned} \quad (2.21)$$

Pour écrire l'ensemble de ces relations de façon plus compacte, il est utile d'introduire des conventions d'écriture. On conviendra d'écrire une somme de k termes se déduisant les uns des autres par permutation d'indices par une notation crochet [20]

$$[3] a_{ij}b_{kl} = a_{ij}b_{kl} + a_{ik}b_{jl} + a_{il}b_{jk}.$$

La présence du crochet entraîne une sommation implicite. Les termes à k indices sont ici des tenseurs d'ordre k complètement symétriques.

Dans le cas non centré, les cumulants d'ordre 1, 2, 3 et 4 sont alors donnés en fonction des moments par les expressions compactes

$$\begin{aligned} cum [x_i] &= r_{x(i)}, \\ cum [x_i, x_j] &= r_{x(i,j)} - r_{x(i)}r_{x(j)}, \\ cum [x_i, x_j, x_l] &= r_{x(i,j,l)} - [3] r_{x(i)}r_{x(j,l)}, \\ cum [x_i, x_j, x_l, x_m] &= r_{x(i,j,l,m)} - [4] r_{x(i)}r_{x(j,l,m)} - [3] r_{x(i,j)}r_{x(l,m)} + 2 [6] r_{x(i)}r_{x(j)}r_{x(l,m)} - 6r_{x(i)}r_{x(j)}r_{x(l)}r_{x(m)}. \end{aligned} \tag{2.22}$$

Dans le cas centré, ces expressions se simplifient en

$$\begin{aligned} cum [x_i] &= 0, \\ cum [x_i, x_j] &= r_{x(i,j)}, \\ cum [x_i, x_j, x_l] &= r_{x(i,j,l)}, \\ cum [x_i, x_j, x_l, x_m] &= r_{x(i,j,l,m)} - [3] r_{x(i,j)}r_{x(l,m)}. \end{aligned} \tag{2.23}$$

Il est intéressant de comparer ces expressions avec celles du cas scalaire : pour les retrouver, il suffit de remplacer $n[m]$ par nm . Dans le cas centré, aux ordres 5 et 6

$$\begin{aligned} cum [x_i, x_j, x_l, x_m, x_n] &= r_{x(i,j,l,m,n)} - [10] r_{x(i,j)}r_{x(l,m,n)}, \\ cum [x_i, x_j, x_l, x_m, x_n, x_o] &= r_{x(i,j,l,m,n,o)} - [15] r_{x(i,j)}r_{x(l,m,n,o)} - [10] r_{x(i,j,l)}r_{x(m,n,o)} + 2 [15] r_{x(i,j)}r_{x(l,m)}r_{x(n,o)}. \end{aligned} \tag{2.24}$$

Théorème 2.10 *Les moments croisés d'un vecteur aléatoire n dimensionnel, $E[x_1x_2\dots x_n]$, sont*

liés aux cumulants croisés d'ordre inférieur ou égal à n par la relation suivante

$$E[x_1 x_2 \dots x_n] = \sum cum [\{x_i\}_{i \in v_1}] cum [\{x_j\}_{j \in v_2}] \dots cum [\{x_l\}_{l \in v_p}], \quad (2.25)$$

où la sommation s'étend sur tous les ensembles $\{v_1, v_2, \dots, v_p : 1 \leq p \leq n\}$ formant une partition de $\{1, 2, \dots, n\}$

A l'ordre 4, on a alors immédiatement

$$\begin{aligned} r_{x(i,j,l,m)} &= cum [x_i, x_j, x_l, x_m] + [4] cum [x_i] cum [x_j, x_l, x_m] + \\ &\quad [3] cum [x_i, x_j] cum [x_l, x_m] + [6] cum [x_i] cum [x_j] cum [x_l, x_m] + \\ &\quad cum [x_i] cum [x_j] cum [x_l] cum [x_m]. \end{aligned}$$

2.3.4 Quelques propriétés des moments et des cumulants

★ Multiplication des variables par des constantes : $E(\alpha_1 x_1, \alpha_2 x_2, \dots, \alpha_N x_N) = \prod_{i=1}^N \alpha_i E(x_1, x_2, \dots, x_N)$, (relation similaire pour les cumulants).

★ Addition des variables : $E(x_1 + y_1, x_2, \dots, x_N) = E(x_1, x_2, \dots, x_N) + E(y_1, x_2, \dots, x_N)$, (relation similaire pour les cumulants).

★ Invariance par translation d'une variable pour les cumulants d'ordre supérieur à un : $cum_k(x_1 + c, x_2, \dots, x_N) = cum_k(x_1, x_2, \dots, x_N)$ (cette relation n'est pas vérifiée pour les moments).

<Q★ Sommes de variables indépendantes : si x_i et y_i sont deux variables aléatoires indépendantes,

$cum_k(x_1 + y_1, x_2 + y_2, \dots, x_N + y_N) = cum_k(x_1, x_2, \dots, x_N) + cum_k(y_1, y_2, \dots, y_N)$, (cette relation n'est pas vérifiée pour les moments).

★ Indépendance des composantes : Si les premières composantes x_1, \dots, x_p sont indépendantes des dernières : $cum_k(x_1, x_2, \dots, x_N) = cum_k(x_1, x_2, \dots, x_p) + cum_k(x_{p+1}, \dots, x_N)$.

(cette relation n'est pas vérifiée pour les moments).

★ Permutation des variables si on modifie l'ordre des variables (x_1, x_2, \dots, x_N) les moments et cumulants sont inchangés.

★ Tous les cumulants d'ordre supérieur ou égal à trois d'une variable gaussienne sont tous nuls.

2.3.5 Estimation des moments et des cumulants

Estimateurs des moments

Un estimateur non biaisé des moments croisés d'ordre k par rapport à l'origine du vecteur aléatoire x est :

$$\widehat{r}_{x(i_1, i_2, \dots, i_k)} = \frac{1}{N} \sum_{k=1}^N x_{i_1} x_{i_2} \dots x_{i_k}. \quad (2.26)$$

Estimateurs des cumulants

Des estimateurs non biaisés des cumulants croisés sont donnés par les k -statistiques. Une description plus détaillée sur les k -statistiques pour des variables multidimensionnelles est donnée en [20] et [36].

Les premières k -statistiques pour un vecteur aléatoire n dimensionnel sont :

$$cum(x_p, x_q) = \frac{1}{N(N-1)} [N s_{pq} - s_p s_q],$$

$$cum(x_p, x_q, x_r) = \frac{1}{N(N-1)(N-2)} [N^2 s_{pqr} - N s_p s_{qr} - N s_q s_{pr} - N s_r s_{pq} + 2 s_p s_q s_r],$$

$$cum(x_p, x_q, x_r, x_t) = \frac{1}{N(N-1)(N-2)(N-3)} [N^2(N+1) s_{pqrt} - N(N+1) s_p s_{qrt} - N(N+1) s_q s_{prt} - N(N+1) s_r s_{pqt} - N(N+1) s_t s_{pqr} - N(N+1) s_{pq} s_{rt} - N(N+1) s_{pr} s_{qt} - N(N+1) s_{pt} s_{qr} + 2 s_p s_q s_r s_t + 2 s_p s_r s_{qt} + 2 s_p s_t s_{rq} + 2 s_q s_r s_{pt} + 2 s_q s_t s_{pr} + 2 s_r s_t s_{pq} - 6 s_p s_q s_r s_t],$$

où $s_p = \sum_{j=1}^N x_{pj}$, $s_{pq} = \sum_{j=1}^N x_{pj} x_{qj}$, $s_{pqr} = \sum_{j=1}^N x_{pj} x_{qj} x_{rj}$, $s_{pqrt} = \sum_{j=1}^N x_{pj} x_{qj} x_{rj} x_{tj}$. et les index $p; q; r; t \in \{1, \dots, n\}$ ne sont pas nécessairement distincts.

Exemple 2.2 Vecteur aléatoire gaussien n dimensionnel.

Soit $x = (x_1, x_2, \dots, x_n)^T$ un vecteur aléatoire gaussien de moyenne $E[x] = \alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$ et de matrice de covariance $C'_{xx} = \{\sigma_{ij}\}_{i,j=1}^n$, $x \mapsto N_{\mathbb{R}^2}(\alpha, C_{xx})$

La densité de probabilité est :

$$f_x(u) = \frac{1}{(2\pi)^{\frac{n}{2}} [\det C_{xx}]} \exp \left\{ -\frac{1}{2} (u - \alpha)^T C_{xx}^{-1} (u - \alpha) \right\}.$$

La première fonction caractéristique [40] s'exprime :

$$P_x(t) = \exp \left\{ jt^T \alpha - \frac{1}{2} t^T C_{xx} t \right\}.$$

Les moments croisés centrés d'ordre impair sont nuls, $r_{x(k_1, k_2, \dots, k_n)} = 0$, et les moments croisés centrés d'ordre pair sont

$$r_{x(k_1, k_2, \dots, k_n)} = \sum (\sigma_{ij} \sigma_{lm} \dots \sigma_{xz}),$$

où $k = k_1 + k_2 + \dots + k_n = 2\lambda$, $i, j, l, m, \dots, xz \in \{1, 2, \dots, n\}$ et la somme \sum est effectuée sur toutes les permutations de $\{1, 2, \dots, 2\lambda\}$ qui donnent $(2\lambda - 1)! / (2^{\lambda-1} (\lambda - 1)!)$ termes dans la somme [32].

Les cumulants croisés $C_{x(k_1, k_2, \dots, k_n)}$ d'ordre $k = k_1 + k_2 + \dots + k_n$ supérieur ou égal à 3 d'un vecteur aléatoire n dimensionnel gaussien sont nuls.

2.4 Multicorrélations

Les propriétés d'ordre supérieur des processus aléatoires peuvent être décrites, comme le sont les propriétés d'ordre 2, dans le domaine du temps ou dans le domaine des fréquences. L'étude dans le domaine du temps conduit aux multicorrélations, l'étude dans le domaine des fréquences aux multispectres.

Les multicorrélations sont définies à partir des cumulants d'ordre supérieur à 2 comme la corrélation est définie à partir du cumulants d'ordre 2 [13], [9], [42], [45], [50], [49]].

Définition 2.10 Soit un processus aléatoire $x(t)$ à valeurs réelles, la multicorrélation d'ordre p de $x(t)$ est le cumulants d'ordre p des valeurs (variables aléatoires) de $x(t)$ aux instants

$$t_1, t_2, \dots, t_{p-1}$$

$$C_{x(p)} = \text{cum} [x(t_1), x(t_2), \dots, x(t_{p-1})], \quad (2.27)$$

où $t = (t_1, t_2, \dots, t_{p-1})$.

Dans le cas général la multicorrélation d'ordre p est fonction d'un temps absolu t_0 et des $p - 1$ temps t_i . Lorsque le processus $x(t)$ est stationnaire au sens strict, ses statistiques sont invariantes par changement de l'origine des temps. Un des temps peut être pris comme origine (nous choisissons $t = t_0$ ici). La multicorrélation n'est alors fonction que des $p - 1$ écarts de temps $\tau_i = t - t_i$. La multicorrélation d'ordre p est alors

$$C_{x(p)}(\tau) = cum [x(t), x(t + \tau_1), \dots, x(t + \tau_{p-1})], \quad (2.28)$$

où $\tau = (\tau_1, \tau_2, \dots, \tau_{p-1})$. On retrouve la corrélation classique (cumulant d'ordre 2)

$$C_{x(2)}(\tau_1) = cum [x(t), x(t + \tau_1)],$$

la bicorrélation (cumulant d'ordre 3)

$$C_{x(3)}(\tau_1, \tau_2) = cum [x(t), x(t + \tau_1), x(t + \tau_2)]. \quad (2.29)$$

Les multicorrélations, définies à partir des cumulants, peuvent s'exprimer en fonction des moments en utilisant les relations entre les cumulants et les moments données ci-dessus. Lorsque dans le cas centrés, la corrélation et la bicorrélation s'expriment directement en fonction des moments

$$\begin{aligned} C_{x(2)}(\tau_1) &= E [x(t)x(t + \tau_1)], \\ C_{x(3)}(\tau_1, \tau_2) &= E [x(t)x(t + \tau_1)x(t + \tau_2)]. \end{aligned}$$

2.4.1 Estimation des multicorrélations

Définition de l'estimateur

Nous allons présenter les expressions pour le cas de la bicorrélation. Nous nous limitons aux signaux à valeurs réelles. Pour un processus $x(n)$, à valeurs réelles, centré connu à travers une

réalisation de N , on estime la bicorrélation par:

$$\widehat{C_{x(3)}}(\tau_1, \tau_2) = \frac{1}{N} \sum_{t=1}^N x(t)x(t + \tau_1)x(t + \tau_2). \quad (2.30)$$

Où N est le nombre de points du processus observé.

Propriétés de l'estimateur

L'estimateur est asymptotiquement non biaisé si

$$NK(N) \rightarrow 1 \quad \text{quand } N \rightarrow \infty.$$

Pour obtenir un estimateur non biaisé il faut prendre

$$K(N) = \frac{1}{N - \max(\tau_1, \tau_2)}.$$

Asymptotiquement, la variance de l'estimateur $\widehat{C_{x(3)}}(\tau_1, \tau_2)$ peut s'écrire

$$\text{var}(\widehat{C_{x(3)}}(\tau_1, \tau_2)) = \frac{1}{N} \sum_{\tau} [C_{x(2)}(\tau_1)]^3 + \frac{1}{N} \sum_k C_{x(6)}(\tau, \tau_1, \tau_2, \tau + \tau_1, \tau + \tau_2), \quad (2.31)$$

où $C_{x(2)}(\tau_1)$, $C_{x(6)}(\tau, \tau_1, \tau_2, \tau + \tau_1, \tau + \tau_2)$ sont les cumulants d'ordre 2 et 6 respectivement.

2.4.2 La motivation pour utiliser des cumulants

La motivation pour utiliser des cumulants plutôt que des moments découle des faits suivants :

A/ Les cumulants, mais pas les moments, des processus linéaires, peuvent être exprimés comme corrélations.

B/ Tous les cumulants d'ordre supérieur ou égal à trois d'une variable gaussienne sont tous nuls.

C/ Certaines considérations techniques concernant l'ergodicité et l'existence de polyspectres sont plus facilement satisfaits par des cumulants plutôt que par des moments ([12];[14]).

Cependant, les expressions pour les cumulants d'ordre supérieur sont plus complexes. Heureusement, les cumulants de troisième et quatrième ordre suffisent pour la plupart des problèmes intéressants.

Chapitre 3

Les modèles moyenne mobile (MA) bidimensionnelle à réponse impulsionnelle finie

Alternativement aux processus autorégressifs spatiaux, les modèles de moyenne mobile sont d'une grande importance théorique et pratique, et ont été étudiées dans différents contextes. La pertinence des modèles MA est principalement due à la grande popularité de la décomposition de type Wold d'un champ aléatoire homogène régulier. Il est bien connu que les modèles de représentation MA pour les champs aléatoires sont moins informatifs que les modèles AR ou ARMA.

Cependant, du point de vue de l'estimation, il est très intéressant de connaître les performances de nouvelles propositions pour l'estimation des paramètres d'un modèle.

Dans ce chapitre, nous nous intéressons à la présentation du modèle 2D MA à l'aide de filtre de réponse impulsionnelle finie.

3.1 Filtres à réponse impulsionnelle finie

3.1.1 Introduction

Les filtres numériques ont une large gamme d'applications dans les domaines de communication, de traitement d'images, de reconnaissance de formes, Les filtres sont un élément clé dans la plupart des systèmes électroniques et plus généralement de tout dispositif de traitement du signal. Un filtre numérique fonctionne sur des signaux échantillonnés à temps discret, à la différence des filtres analogiques qui opèrent sur des signaux analogiques en temps continu. La plupart des systèmes électroniques contiennent un grand nombre de filtres qui répondent aux exigences variées. Le filtrage est une forme de traitement du signal, obtenu par l'émission du signal utile par l'intermédiaire d'un ensemble de circuits électroniques, ce qui modifie de façon adaptative son spectre de fréquence, de phase, et donc sa forme. Le filtre est conçu pour extraire certaines informations relatives à ce signal, généralement pour retourner, à partir d'un signal incident, un signal qui est plus intelligible en termes de l'information qu'il contient et que l'on veut sélectionner.

Les avantages des filtres sont considérables : les modèles peuvent être personnalisés, précis et adaptables ; au-delà de leur fonction utile, ils peuvent être facilement intégrés dans des ensembles existants et peuvent générer des réductions de coûts dans la conception des composants.

Au cours des dernières années et avec l'amélioration rapide de la technologie informatique, le traitement du signal numérique bidimensionnel (2D) est devenu plus important. En particulier, la conception de filtres numériques 2D a reçu beaucoup d'attention dans divers domaines.

Les filtres numériques peuvent être classés en plusieurs groupes, selon les critères de classification utilisés. Les deux principaux types de filtres numériques, classés selon la longueur de la réponse impulsionnelle, sont les filtres numériques à réponse impulsionnelle finie (filtres RIF) en anglais Finite Impulse Reponse filter (FIR) est un filtre dont la réponse impulsionnelle est de durée finie. et les filtres numériques à réponse impulsionnelle infinie (filtres RII). Chaque type de filtres a des avantages et des inconvénients qui doivent être examinés avec soin lors de leur conception.

impulsionnelle finie

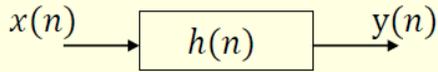
Les filtres numériques à réponse impulsionnelle finie sont également connus sous le nom de filtres numériques non récursifs car ils n'ont pas de réaction, bien que des algorithmes récursifs peuvent être utilisés pour concevoir les filtres RIF. Ces filtres ont de nombreux avantages tels que la stabilité, la possibilité d'obtenir une phase linéaire.

Nous présentons dans cette section les notions générales de filtre à réponse impulsionnelle finie des signaux monodimensionnelles et bidimensionnelles.

3.1.2 Filtres RIF unidimensionnels (1D)

Présentation du filtre RIF

De façon générale le filtre à réponse impulsionnelle finie est décrit par :



$$y[n] = \sum_{k=0}^N h_k x[n - k], \quad (3.1)$$

avec $x[n]$ représente les valeurs du processus d'entrée et $y[n]$ les valeurs du processus de sortie.

Où N : Nombre de coefficients (ordre du filtre), h_k : Coefficients de la fonction de transfert du filtre.

Fonction de transfert de filtre RIF

Définition 3.1 *Considérons alors la transformée en Z du signal de sortie :*

$$Y(Z) = H(Z)X(Z),$$

où $H(Z)$ est la fonction de transfert en Z définie par :

$$H(Z) = \sum_{k=0}^N h_k Z^{-k}.$$

Réalisation de filtre RIF

Les filtres RIF peuvent être réalisés à l'aide de trois éléments ou opérations de base. Soit l'élément gain, l'élément de sommation et le retard unitaire. Ces éléments sont suffisants pour réaliser tous les filtres. La réalisation présentée dans la figure ci-contre est une réalisation du filtre RIF.

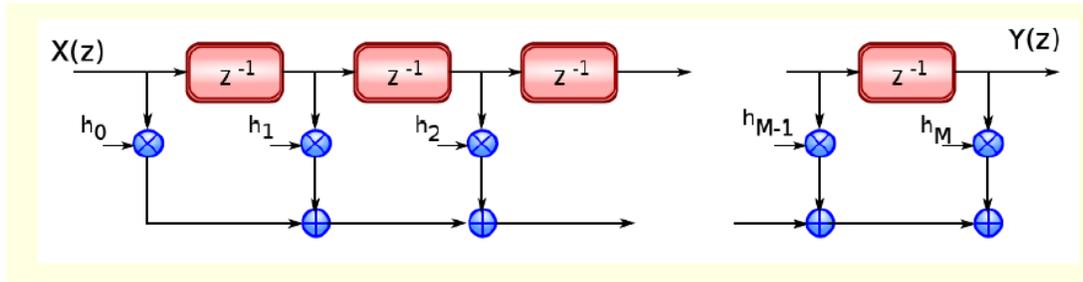


Figure 3 : Réalisation d'un filtre RIF.

3.1.3 Filtre RIF bidimensionnels (2D)

Présentation du filtre RIF

Définition 3.2 Un filtre RIF à deux dimensions est défini par une relation entre la sortie $y(m, n)$ et l'entrée $x(m, n)$ qui s'écrit :

$$y(m, n) = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} h_{ij} x(m-i, n-j), \quad (3.2)$$

où l'ensemble des coefficients h_{ij} constitue une matrice de dimension $N * M$.

Fonction de transfert de filtre RIF

La fonction de transfert bidimensionnel en $Z = (Z_1, Z_2)$ donne la relation entre l'entrée et la sortie

$$\begin{aligned} H(Z_1, Z_2) &= \frac{Y(Z_1, Z_2)}{X(Z_1, Z_2)} \\ &= \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} h_{ij} Z^{-i} Z^{-j}}{1 + \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} a_{ij} Z^{-i} Z^{-j}}. \end{aligned}$$

Dans le cas d'un filtre RIF $\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} a_{ij} = 0$. La fonction de transfert des filtres à réponse impulsionnelle finie (RIF) est de type polynomial, de ce fait, cette fonction de transfert n'a pas de pôles à l'extérieur du cercle unité. Cette absence de pôles permet d'obtenir des filtres inconditionnellement stables.

Donc la fonction de transfert à deux variables correspondante, $h(Z_1, Z_2)$ s'exprime en fonction de cette matrice par [Francis Cottet] :

$$H(Z_1, Z_2) = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} h_{ij} Z^{-i} Z^{-j}. \quad (3.3)$$

La réponse fréquentielle est donnée par :

$$\begin{aligned} H(\omega_1, \omega_2) &= \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{M-1} h(n_1, n_2) e^{-j(n_1\omega_1 + n_2\omega_2)} \\ &= S(\omega_1, \omega_2) e^{j\theta(\omega_1, \omega_2)}, \end{aligned}$$

avec $S(\omega_1, \omega_2) = |H(\omega_1, \omega_2)|$ et $\theta(\omega_1, \omega_2) = \arg(H(\omega_1, \omega_2))$, et $S(\omega_1, \omega_2)$, la réponse d'amplitude de $(H(\omega_1, \omega_2))$, est une fonction à valeurs réelles.

Exemple 3.1 *Tracé de la réponse en fréquence* $H(e^{ju}, e^{jv}) = \frac{1}{4}(1 + \cos(u))(1 + \cos(v))$ du filtre 2D RIF, avec $H(e^{j0}, e^{j0}) = 1$

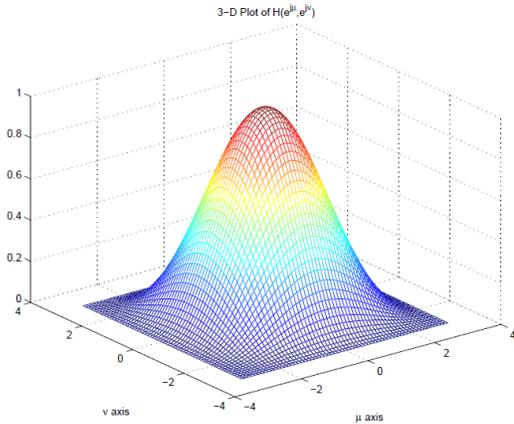


Figure 4 : Tracé 3D de $H(e^{ju}, e^{jv})$.

Remarque 3.1 *La plupart du temps, les expressions des transformées en Z utilisées en traitement du signal sont des fractions rationnelles de la variable Z . On appelle « pôles » les racines du dénominateur de transformée et « zéros » numérateur.*

Propriétés

- **Stabilité** : le filtre RIF ne pose jamais de problème de stabilité (est toujours stable) en raison de l'absence de pôles.
- **Approximation** : Toute fonction de filtrage numérique stable et causale peut être approchée par la fonction de transfert d'un filtre RIF.
- **Phase linéaire** : Les filtres RIF peuvent générer des filtres à phase linéaire
- Un filtre RIF est moins sélectif qu'un filtre RII du même ordre. C'est-à-dire que la transition entre la bande passante et la bande rejetée est moins rapide que dans le cas du filtre RII.

3.2 La phase non minimale

3.2.1 Notions importantes

Stabilité

- La stabilité BIBO (bounded input-bounded output) exige qu'à toute entrée bornée le filtre associe une sortie bornée. Autrement dit, un filtre est stable, si pour toute entrée $x(n)$ vérifiant

$$|x(n)| \prec +\infty$$

la sortie $y(n)$ vérifie

$$|y(n)| \prec +\infty$$

Il existe d'autres méthodes pour définir la stabilité, données dans les définitions suivantes :

Définition 3.3 *Un filtre est stable, si tous les pôles de sa fonction de transfert en Z dans le plan complexe se trouvent à l'intérieur du cercle de rayon 1. Autrement dit, il faut que le cercle unité appartienne à la région de convergence.*

Définition 3.4 *Dans le cas continu, un filtre est stable, si tous les pôles de sa fonction de transfert se trouvent dans la moitié gauche de la représentation du plan de Z .*

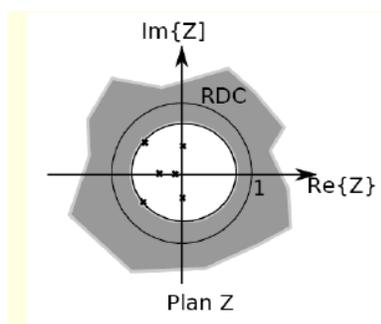


Figure 5 : Stabilité.

Causalité

Définition 3.5 *Un filtre est causal s'il n'est pas nécessaire de connaître les échantillons futurs ($x(t + m), m \succ 0$) de l'entrée $x(t)$ pour en calculer la sortie $y(t)$ à l'instant t . L'effet ne peut pas précéder la cause ceci se traduit par le fait que la réponse impulsionnelle doit être nulle pour $t \prec 0$.*

- Un filtre causal est un filtre dont la sortie ne dépend que des valeurs intérieures (passés + présent).

Par exemple $y(n) = x(n - 4) + 2x(n - 1) + x(n)$.

- Un filtre anti causal est un filtre dont la sortie ne dépend que des valeurs futures ou de présent.

Par exemple $y(n) = x(n + 4) + 2x(n) + 3x(n + 2)$.

- Un filtre non causal est un filtre dont la sortie ne dépend que des valeurs passées et de présent ainsi que des valeurs futures ou de présent.

Par exemple $y(n) = 2x(n + 4) + 4x(n) + 3x(n - 2)$.

- La stabilité et la causalité peuvent s'exprimer dans les différentes représentations des filtres linéaires et stationnaires.

Une condition nécessaire et suffisante de stabilité est que la réponse impulsionnelle soit sommable en module, c'est - à - dire :

$$\sum_{n=-\infty}^{+\infty} |h(n)| \prec +\infty,$$

et la condition de causalité est simplement

$$h(n) = 0 \quad \text{pour} \quad n \prec 0.$$

L'inversibilité

Un filtre est inversible si nous pouvons déterminer de manière unique son entrée à partir de sa sortie, nous obtenons le système d'identité

$$H_{inv}H = 1.$$

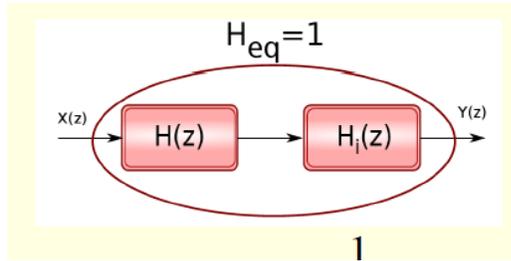


Figure 6 : Système inverse.

Pour un filtre linéaire stationnaire discret ayant une fonction de transfert $H(z)$ rationnelle, la fonction de transfert du filtre inverse est

$$H_{inv}(x) = \frac{1}{H(x)}.$$

Les pôles de $H(z)$ sont les zéros de $H_{inv}(z)$ et les zéros de $H(z)$ sont les pôles de $H_{inv}(z)$. Il s'ensuit qu'un filtre linéaire et stationnaire sera stable, causal

et causalement inversible si et seulement si :

1. le nombre de pôles est égal au nombre de zéros,
2. les pôles et les zéros sont dans le disque de rayon 1.

3.2.2 Le filtre à phase minimale

Un système linéaire et stationnaire sera dit stable, causal et causalement inversible ou à phase minimale si

- 1- Le filtre linéaire et stationnaire est stable et causal ;
- 2- son inverse (permettant de retrouver l'entrée à partir de la sortie) est stable et causal.

Définition 3.6 *Un filtre est de phase minimale si tous les pôles et les zéros de sa fonction de transfert sont dans la moitié gauche de représentation du plan Z (en temps discret, respectivement, à l'intérieur du cercle unitaire du plan (Z)).*

3.2.3 Le filtre à phase non minimale

Définition 3.7 *Un filtre à phase non minimum est un filtre causal et stable d'ont l'inverse est instable, non inversible.*

Définition 3.8 *Dans le cas continu, un filtre de phase non minimale a tous les zéros de sa fonction de transfert dans la moitié droite de représentation du plan Z .*

Un système à phase non minimale donné aura une contribution de phase plus élevée que le système à phase minimale avec une réponse en amplitude équivalente.

3.2.4 Autres dénominations

Phase maximale

Définition 3.9 *Un filtre à phase maximale est l'opposé d'un filtre à phase minimale. Un filtre causal et stable est un système à phase maximale si son inverse est causal et instable. C'est, les zéros du système à temps discret sont en dehors du cercle unitaire .*

Phase linéaire

Définition 3.10 *Un filtre à phase linéaire est un filtre dont la réponse en phase est linéaire - le temps de propagation de groupe est constant.*

3.3 Modèles moyenne mobile (MA) bidimensionnelle

3.3.1 Représentation du modèle MA spatial

Soit $x = (x(i, j))_{(i, j) \in \mathbb{Z}^2}$ un champ aléatoire à valeur réelle. Alors le modèle moyenne mobile bidimensionnel peut être représenté par :

$$x(n, m) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} b(i, j) \varepsilon(n - i, m - j), \quad (3.4)$$

où la suite $\{\varepsilon(n, m) : (n, m) \in \mathbb{Z}^2\}$ une famille de variables aléatoires centrées, indépendantes et identiquement distribuées et de même variance σ^2 et avec $b(i, j)$ désignant les paramètres du modèle. Nous supposons que $b(0, 0) = 1$. En pratique, le processus aléatoire observé est de dimension finie. Pour le support

$$D = \{(i, j) : 0 \leq i \leq S - 1, 0 \leq j \leq T - 1\},$$

le processus observé est l'ensemble $\{x(n, m), (n, m) \in D\}$. En particulier, pour $(i, j) \in S_{(q_1, q_2)}$ où $S_{(q_1, q_2)} = \{(n, m) : 0 \leq n \leq q_1, 0 \leq m \leq q_2\}$ et q_1, q_2 sont des nombre naturels connus, le modèle (3.4) peut être écrit sous la forme :

$$x(n, m) = \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b(i, j) \varepsilon(n - i, m - j). \quad (3.5)$$

Le modèle obtenu par la relation dans (3.5) avec l'hypothèse de bruit blanc sur les erreurs est appelé modèle *MA* spatial.

3.3.2 Représentation du modèle 2D RIF MA

On considère le processus représenté à l'aide de la figure

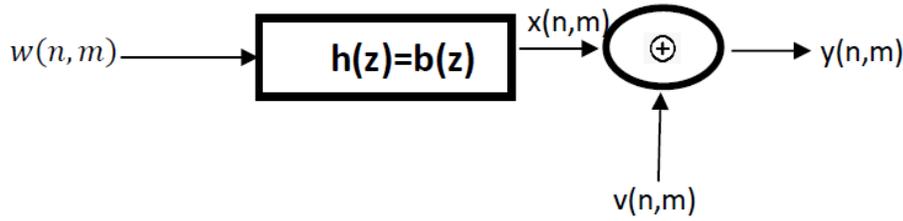


Figure 7 : Processus 2D RIF MA.

L'entrée $w(n, m)$ est un processus indépendant identiquement distribuée (i.i.d) non gaussien, $h(z_1, z_2)$ la réponse impulsionnelle du filtre où $z = (z_1, z_2)$, et $x(n, m)$ la sortie. On trouve l'équation suivante

$$x(n, m) = \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} h(i, j)w(n - i, m - j). \quad (3.6)$$

-Le modèle 2D MA RIF est donné par

$$x(n, m) = \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b(i, j)w(n - i, m - j), \quad (3.7)$$

et $h(i, j) = b(i, j)$, où $b(i, j)$ sont les coefficients MA. Le signal $x(n, m)$ observé en bruit additif noté $y(n, m)$

$$y(n, m) = x(n, m) + v(n, m)$$

où $v(n, m)$ est un processus gaussien indépendant de $w(n, m)$.

- L'inverse du modèle 2D RIF MA présenté par l'équation (3.7), est donné par

$$w(n, m) = \sum_{i=r_1}^{r_2} \sum_{j=r_1}^{r_2} c(i, j)x(n - i, m - j)$$

avec $r_1 < 0$ et $r_2 > 0$. On normalise $(c(i, j))$ en posant $(c(i, j) = 1)$, la partie à phase minimale est représentée par la partie causale du filtre ($i > 0, j > 0$) et la partie à phase maximale par la partie anti-causale ($i < 0, j < 0$).

Remarque 3.2 - Si $y(n, m) = x(n, m) + v(n, m)$, où $x(n, m)$ est un processus non gaussien,

Chapitre 3. Les modèles moyenne mobile (MA) bidimensionnelle à réponse impulsionnelle finie

et $v(m, n)$ est un processus gaussien indépendant de $x(m, n)$, alors le cumulants de $y(m, n)$ est identique au cumulants du signal $x(m, n)$.

-Les cumulants du processus linéaire portent les informations d'amplitude et de phase du système linéaire. Ainsi, dans le calcul des cumulants du processus bruyant observé $y(m, n)$, nous transformons les données observées asymptotiquement dans un domaine (rapport signal sur bruit) élevé qui préserve tous les informations pertinentes sur le signal.

Chapitre 4

Propriétés asymptotiques des cumulants bidimensionnels

Dans ce chapitre, la convergence presque sûre des estimateurs des cumulants spatiaux empiriques d'ordre trois et quatre est discutée ainsi que la normalité asymptotique pour un processus linéaire spatial.

4.1 Convergence presque sûre

Nous considérons le moment (cumulant) empirique d'ordre 3 $\hat{r}_3(\nu^1; \nu^2)$ de $X_{\mathbf{t}}$, à savoir,

$$\hat{r}_{3X}(\nu^1; \nu^2) = \frac{1}{N^d} \sum_{\substack{t_i=1 \\ i=1, \dots, d}}^N X_{\mathbf{t}} X_{\mathbf{t}+\nu^1} X_{\mathbf{t}+\nu^2}. \quad (4.1)$$

Notre principale hypothèse sur $X_{\mathbf{t}}$ est

$$(A) \quad C_{4X}(\nu^1; \nu^2; \nu^3) \in L_1(\mathbb{R}^3) \text{ et } C_{6X}(\nu^1; \nu^2; \nu^3; \nu^4; \nu^5) \in L_1(\mathbb{R}^5).$$

Avant de présenter la convergence presque sûre de $\hat{r}_3(\nu^1; \nu^2)$, nous prouvons le lemme suivant

Lemme 4.1 Si $(X_{\mathbf{t}})_{\mathbf{t} \in \mathbb{Z}^d}$, $EX_{\mathbf{t}} = 0$, est un processus stationnaire du 6-ième ordre, et $r_{2X}(\nu)$, $r_{3X}(\nu^1; \nu^2)$, $C_{4X}(\nu^1; \nu^2; \nu^3)$ et $C_{6X}(\nu^1; \nu^2; \nu^3; \nu^4; \nu^5) \in L_1$, alors pour $\mathbf{t} = (t_1, \dots, t_d)$, $\nu^1 = (\nu_1^1, \nu_2^1, \dots, \nu_d^1)$ et $\nu^2 = (\nu_1^2, \nu_2^2, \dots, \nu_d^2)$

$$Y_{\mathbf{t}}^{(\nu^1, \nu^2)} = X_{\mathbf{t}} X_{\mathbf{t} + \nu^1} X_{\mathbf{t} + \nu^2} - r_{3X}(\nu^1; \nu^2)$$

est un processus stationnaire, et la fonction de densité spectrale $f_{Y^{(\nu^1, \nu^2)}}(\lambda)$ satisfait

$$\sup_{\lambda \in \mathbb{Z}^d} f_{Y^{(\nu^1, \nu^2)}}(\lambda) < \infty.$$

Preuve Il est bien clair que $Y_{\mathbf{t}}^{(\nu^1, \nu^2)}$ est un processus stationnaire. Posons pour tout $\tau \in \mathbb{Z}^d$, $R_{Y^{(\nu^1, \nu^2)}}(\tau) = EY_{\mathbf{t}}^{(\nu^1, \nu^2)} Y_{\mathbf{t} + \tau}^{(\nu^1, \nu^2)} = r_{6X}(\nu^1, \nu^2, \tau, \tau + \nu^1, \tau + \nu^2) - [r_{3X}(\nu^1, \nu^2)]^2$.

Par l'usage de la relation de conversion entre les moments d'ordre supérieurs et les cumulants nous obtenons (voir la preuve du Lemme 1 de [33])

$$\begin{aligned} R_{Y^{(\nu^1, \nu^2)}}(\tau) &= C_{6X}(\nu^1, \nu^2, \tau, \tau + \nu^1, \tau + \nu^2) \\ &\quad + \{r_{2X}(\nu^1) C_{4X}(\tau - \nu^2, \tau + \nu^1 - \nu^2, \nu^2)\}_{15} \\ &\quad + \{r_{2X}(\nu^1) r_{2X}(\tau - \nu^2) r_{2X}(\nu^2 - \nu^1)\}_{15}, \end{aligned}$$

où la notation $\{\cdot\}_j$ désigne la somme de tous les j différents termes obtenus en interchangeant les arguments des termes entre parenthèses, et par la propriété de sommabilité absolue de $r_{2X}(\nu)$, $r_{3X}(\nu^1, \nu^2)$, $C_{4X}(\nu^1, \nu^2, \nu^3)$ et $C_{6X}(\nu^1, \nu^2, \nu^3, \nu^4, \nu^5)$, nous avons

$$\sum_{\tau \in \mathbb{Z}^d} |R_{Y^{(\nu^1, \nu^2)}}(\tau)| < +\infty.$$

Si $f_{Y^{(\nu^1, \nu^2)}}(\cdot)$ est la fonction de densité spectrale de $Y_{\mathbf{t}}^{(\nu^1, \nu^2)}$, par la théorie spectrale (voir [31]), $f_{Y^{(\nu^1, \nu^2)}}(\lambda)$ a la représentation

$$f_{Y^{(\nu^1, \nu^2)}}(\lambda) = \frac{1}{(2\pi)^d} \sum_{\mathbf{n} \in \mathbb{Z}^d} R_{Y^{(\nu^1, \nu^2)}}(\mathbf{n}) e^{-i \langle \lambda, \mathbf{n} \rangle},$$

où $\langle \lambda, \mathbf{n} \rangle = \sum_{i=1}^d \lambda_i n_i$, d'où

$$\sup_{\lambda \in \mathbb{Z}^d} f_{Y^{(\nu^1, \nu^2)}}(\lambda) \leq \frac{1}{(2\pi)^d} \sum_{\mathbf{n} \in \mathbb{Z}^d} |R_{Y^{(\nu^1, \nu^2)}}(\mathbf{n})| < +\infty.$$

■

Le premier résultat important de cette section est la convergence presque sûre du moment (cumulant) empirique d'ordre 3

Théorème 4.1 *Si $(X_{\mathbf{t}})_{\mathbf{t} \in \mathbb{Z}^d}$, $EX_{\mathbf{t}} = 0$, est un processus stationnaire du 6-ième ordre, et $r_{2X}(\nu)$, $r_{3X}(\nu^1; \nu^2)$, $C_{4X}(\nu^1; \nu^2; \nu^3)$ et $C_{6X}(\nu^1; \nu^2; \nu^3; \nu^4; \nu^5) \in L_1$, alors pour $\mathbf{t} = (t_1, \dots, t_d)$, $\nu^1 = (\nu_1^1, \nu_2^1, \dots, \nu_d^1)$ et $\nu^2 = (\nu_1^2, \nu_2^2, \dots, \nu_d^2)$*

(1)

$$E \left\{ \frac{1}{N^{2d}} \left| \sum_{\substack{t_i=1 \\ i=1, \dots, d}}^N X_{\mathbf{t}} X_{\mathbf{t}+\nu^1} X_{\mathbf{t}+\nu^2} - r_3(\nu^1; \nu^2) \right|^2 \right\} = O\left(\frac{1}{N^d}\right) \quad (4.2)$$

(2) pour $k > \frac{3}{4}$,

$$\frac{1}{N^d} \left| \sum_{\substack{t_i=1 \\ i=1, \dots, d}}^N X_{\mathbf{t}} X_{\mathbf{t}+\nu^1} X_{\mathbf{t}+\nu^2} - r_3(\nu^1; \nu^2) \right| \stackrel{p.s}{=} o\left(\frac{1}{N^{d(1-k)}}\right) \quad (4.3)$$

où *p.s* représente la convergence presque sûrement.

Preuve Par la théorie spectrale (voir [31]), nous avons

$$R_{Y^{(\nu^1, \nu^2)}}(\tau) = \int_{[0, 2\pi]^d} e^{i\langle \lambda, \tau \rangle} f_{Y^{(\nu^1, \nu^2)}}(\lambda) d\lambda,$$

De plus

$$E \left\{ \left| \sum_{\substack{t_i=1 \\ i=1, \dots, d}}^N X_{\mathbf{t}} X_{\mathbf{t}+\nu^1} X_{\mathbf{t}+\nu^2} - r_3(\nu^1, \nu^2) \right|^2 \right\}$$

$$\begin{aligned}
 &= E \left\{ \left| \sum_{i=1, \dots, d}^N Y_{\mathbf{t}}^{(\nu^1, \nu^2)} \right|^2 \right\} \\
 &= E \left\{ \sum_{i=1, \dots, d}^N Y_{\mathbf{t}}^{(\nu^1, \nu^2)} Y_{\mathbf{s}}^{(\nu^1, \nu^2)} \right\} \\
 &= \sum_{i=1, \dots, d}^N R_{Y^{(\nu^1, \nu^2)}}(\mathbf{t} - \mathbf{s}),
 \end{aligned}$$

alors

$$\begin{aligned}
 &E \left\{ \left| \sum_{i=1, \dots, d}^N X_{\mathbf{t}} X_{\mathbf{t} + \nu^1} X_{\mathbf{t} + \nu^2} - r_3(\nu^1, \nu^2) \right|^2 \right\} \\
 &= \sum_{i=1, \dots, d}^N \int_{[0, 2\pi]^d} e^{i\langle \lambda, \mathbf{t} - \mathbf{s} \rangle} f_{Y^{(\nu^1, \nu^2)}}(\lambda) d\lambda \\
 &\leq \sup_{\lambda \in \mathbb{Z}^d} f_{Y^{(\nu^1, \nu^2)}}(\lambda) \sum_{i=1, \dots, d}^N \int_{[0, 2\pi]^d} e^{i\langle \lambda, \mathbf{t} - \mathbf{s} \rangle} d\lambda \\
 &\leq C \times N^d,
 \end{aligned}$$

où C représente constante indépendante de N , alors

$$\begin{aligned}
 &E \left\{ \frac{1}{N^{2d}} \left| \sum_{i=1, \dots, d}^N X_{\mathbf{t}} X_{\mathbf{t} + \nu^1} X_{\mathbf{t} + \nu^2} - r_3(\nu^1, \nu^2) \right|^2 \right\} \\
 &\leq C \times \frac{N^d}{N^{2d}} = \frac{C}{N^d},
 \end{aligned}$$

qui est (4.2).

Suivant [33], considérons (4.3) sur la base de (4.2) et soit

$$S(N) = \frac{1}{N^{kd}} \left| \sum_{\substack{t_i=1 \\ i=1, \dots, d}}^N Y_{\mathbf{t}}^{(\nu^1, \nu^2)} \right|.$$

Pour $k > \frac{3}{4}$, il existe β qui satisfait $\beta > 1$ et,

$$k - \frac{1}{2} > \frac{1}{2\beta} > \frac{1}{4}.$$

Maintenant, considérons les deux cas où $N = M^\beta$ et pour N quelconque.

(I) Quand $N = M^\beta$, nous avons

$$\begin{aligned}
 E |S(N)|^2 &= E \frac{1}{N^{2kd}} \left| \sum_{\substack{t_i=1 \\ i=1,\dots,d}}^N Y_{\mathbf{t}}^{(\nu^1, \nu^2)} \right|^2 \\
 &= \frac{1}{N^{2kd-2d}} E \frac{1}{N^{2d}} \left| \sum_{\substack{t_i=1 \\ i=1,\dots,d}}^N Y_{\mathbf{t}}^{(\nu^1, \nu^2)} \right|^2 \\
 &\leq \frac{1}{N^{2kd-2d}} \times \frac{C}{N^d} = CM^{\beta d(1-2k)},
 \end{aligned}$$

comme $\beta d(1-2k) < -d \leq -1$, alors $\sum_{M=1}^{+\infty} M^{\beta d(1-2k)}$ est une série de Riemman convergente, par conséquent

$$\sum_{M=1}^{+\infty} E |S(M^\beta)|^2 < +\infty.$$

De l'inégalité de Chebychev et le lemme de Borel-Cantelli lemma, il s'ensuit

$$\lim_{M \rightarrow \infty} S(M^\beta) \stackrel{p.s.}{=} 0. \tag{4.4}$$

(II) Pour le cas général (N quelconque), et $N_M \leq N \leq N_{M+1}$, où $N_M = M^\beta$, nous avons

$$|S(N) - N^{-kd} M^{kd\beta} S(M^\beta)| \leq M^{-\beta kd} \left| \sum_{\substack{t_i=M^\beta+1 \\ i=1,\dots,d}}^N Y_{\mathbf{t}}^{(\nu^1, \nu^2)} \right|,$$

et

$$\begin{aligned}
 &E \left\{ \sup_{N_M \leq N \leq N_{M+1}} |S(N) - N^{-kd} M^{kd\beta} S(M^\beta)|^2 \right\} \\
 &\leq M^{-2\beta kd} \sup_{N_M \leq N \leq N_{M+1}} E \left| \sum_{\substack{t_i=M^\beta+1 \\ i=1,\dots,d}}^N Y_{\mathbf{t}}^{(\nu^1, \nu^2)} \right|^2 \\
 &\leq CM^{-2\beta kd} \sup_{N_M \leq N \leq N_{M+1}} [N - M^\beta]^d \\
 &\leq CM^{-2\beta kd} [(M+1)^\beta - M^\beta]^d,
 \end{aligned}$$

remarquant que $(M + 1)^\beta - M^\beta \leq (2M)^\beta$, nous déduisons que

$$E \sup_{N_M \leq N \leq N_{M+1}} |S(N) - N^{-kd} M^{kd\beta} S(M^\beta)|^2 \leq C_1 M^{\beta d(1-2k)},$$

où C_1 représente une constante indépendante de M . Comme $\beta d(1 - 2k) < -d \leq -1$, nous obtenons

$$\sum_{M=1}^{+\infty} E \sup_{N_M \leq N \leq N_{M+1}} |S(N) - N^{-kd} M^{kd\beta} S(M^\beta)|^2 < \infty,$$

De l'inégalité de Chebychev et le lemme de Borel-Cantelli lemma, il s'ensuit

$$\lim_{M \rightarrow \infty} \sup_{N_M \leq N \leq N_{M+1}} |S(N) - N^{-kd} M^{kd\beta} S(M^\beta)|^2 \stackrel{p.s}{=} 0. \quad (4.5)$$

De (4.4), (4.5) et

$$S(N) \leq \sup_{N_M \leq N \leq N_{M+1}} |S(N) - N^{-kd} M^{kd\beta} S(M^\beta)| + N^{-kd} M^{kd\beta} S(M^\beta),$$

nous avons, pour $k > \frac{3}{4}$,

$$S(N) = \frac{1}{N^{kd}} \left| \sum_{\substack{t_i=1 \\ i=1, \dots, d}}^N Y_{\mathbf{t}}^{(\nu^1, \nu^2)} \right| \xrightarrow{p.s} 0,$$

ce qui achève la preuve. ■

Remarque 4.1 *Il n'est pas difficile de généraliser le Théorème (4.1) aux estimateurs des moments et des cumulants d'ordre k , pourvu qu'on prenne $Y_{\mathbf{t}}^{(\nu^1; \dots; \nu^{k-1})} = X_{\mathbf{t}} X_{\mathbf{t}+\nu^1} \dots X_{\mathbf{t}+\nu^{k-1}} - r_{kX}(\nu^1; \dots; \nu^{k-1})$, et que les cumulants d'ordre $2k$ -soient absolument sommables.*

A présent, nous établissons la normalité asymptotique de $\hat{r}_{3X}(\nu^1, \nu^2)$ pour un processus linéaire spatial.

4.2 Normalité asymptotique

Théorème 4.2 *Soit*

$$X_{\mathbf{t}} = \sum_{\mathbf{j} \in \mathbb{Z}^d} \psi_{\mathbf{j}} \varepsilon_{\mathbf{t}-\mathbf{j}},$$

où les $\varepsilon_{\mathbf{j}}$ sont indépendamment et identiquement distribuées avec $E(\varepsilon_{\mathbf{j}}) = 0$, $E(\varepsilon_{\mathbf{j}}^2) = 1$, $E(\varepsilon_{\mathbf{j}}^3) = \gamma_3$, $E(\varepsilon_{\mathbf{j}}^6) < \infty$ et $\sum_{\mathbf{j} \in \mathbb{Z}^d} |\psi_{\mathbf{j}}| < \infty$, alors

$$N^{\frac{d}{2}} (\hat{r}_{3X}(\nu^1; \nu^2) - r_{3X}(\nu^1; \nu^2)) \rightarrow \mathcal{N}(0, \sigma_X^2(\nu^1; \nu^2)), \quad (4.6)$$

où

$$\begin{aligned} \sigma_X^2(\nu^1; \nu^2) &= \sum_{\substack{\tau_i = -\infty \\ i=1, \dots, d}}^{+\infty} \{C_{6X}(\nu^1; \nu^2; \tau; \tau + \nu^1; \tau + \nu^2) \\ &+ \{r_{2X}(\nu^1) C_{4X}(\tau - \nu^2; \tau + \nu^1 - \nu^2; \nu^2)\}_{15} \\ &+ \{r_{2X}(\nu^1) r_{2X}(\tau - \nu^2) r_{2X}(\nu^2 - \nu^1)\}_{15}\}. \end{aligned}$$

Preuve Pour $\mathbf{k} = (k, \dots, k) \in \mathbb{Z}^d$ tel que $k > 0$, soit

$$\begin{aligned} X_{\mathbf{t}, \mathbf{k}} &= \sum_{\substack{s_i = -k \\ i=1, \dots, d}}^{+k} \psi_{\mathbf{s}} \varepsilon_{\mathbf{t}-\mathbf{s}}, \\ \hat{r}_{3, \mathbf{kX}}(\nu^1, \nu^2) &= \frac{1}{N^d} \sum_{\substack{t_i=1 \\ i=1, \dots, d}}^N X_{\mathbf{t}, \mathbf{k}} X_{\mathbf{t}+\nu^1, \mathbf{k}} X_{\mathbf{t}+\nu^2, \mathbf{k}}, \end{aligned}$$

$$E(X_{\mathbf{t}, \mathbf{k}} X_{\mathbf{t}+\nu^1, \mathbf{k}} X_{\mathbf{t}+\nu^2, \mathbf{k}}) = r_{3, \mathbf{kX}}(\nu^1, \nu^2).$$

De la symétrie de $r_{3X}(\nu^1, \nu^2)$, nous pouvons supposer $0 \leq \nu_i^1 \leq \nu_i^2$ pour tout $i = 1, \dots, d$, alors

$$E(X_{\mathbf{t}, \mathbf{k}} X_{\mathbf{t}+\nu^1, \mathbf{k}} X_{\mathbf{t}+\nu^2, \mathbf{k}}) = \begin{cases} \gamma_3 \sum_{\substack{s_i = -k \\ i=1, \dots, d}}^{k-\nu_i^2} \psi_{\mathbf{s}} \psi_{\nu^1+\mathbf{s}} \psi_{\nu^2+\mathbf{s}}, \\ \nu_i^2 \leq 2k \text{ pour au moins un } i \\ 0, \quad \text{si non} \end{cases}.$$

Pour $\nu_i^2 \leq 2k$ et pour au moins un $i = 1, \dots, d$, nous avons

$$r_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2) = E\hat{r}_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2).$$

Nous déduisons que $X_{\mathbf{t},\mathbf{k}}X_{\mathbf{t}+\nu^1,\mathbf{k}}X_{\mathbf{t}+\nu^2,\mathbf{k}}$ est indépendante de $X_{\mathbf{s},\mathbf{k}}X_{\mathbf{s}+\nu^1,\mathbf{k}}X_{\mathbf{s}+\nu^2,\mathbf{k}}$ pourvu que $|s_i - t_i| > 2k + \nu_i^2$ pour au moins un $i = 1, \dots, d$ par l'indépendance des $\varepsilon_{\mathbf{t}}$. Donc, si nous posons $\mathbf{m} = 2\mathbf{k} + \nu^2$, $X_{\mathbf{t},\mathbf{k}}X_{\mathbf{t}+\nu^1,\mathbf{k}}X_{\mathbf{t}+\nu^2,\mathbf{k}} - r_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2)$ est un champ stationnaire \mathbf{m} -dependant de moyenne, il s'en suit que

$$N^{-\frac{d}{2}} \left(\sum_{i=1, \dots, d}^N \{X_{\mathbf{t},\mathbf{k}}X_{\mathbf{t}+\nu^1,\mathbf{k}}X_{\mathbf{t}+\nu^2,\mathbf{k}} - r_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2)\} \right) \rightarrow \mathcal{N}\left(0, \sigma_{X,\mathbf{k}}^2(\nu^1, \nu^2)\right), \quad (4.4)$$

$$\text{où } \sigma_{X,\mathbf{k}}^2(\nu^1, \nu^2) = \lim_{N \rightarrow \infty} N^d$$

$$\times \text{Var} \left(N^{-d} \sum_{i=1, \dots, d}^N \{X_{\mathbf{t},\mathbf{k}}X_{\mathbf{t}+\nu^1,\mathbf{k}}X_{\mathbf{t}+\nu^2,\mathbf{k}} - r_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2)\} \right),$$

il n'est pas difficile de montrer que

$$\begin{aligned} \sigma_{X,\mathbf{k}}^2(\nu^1, \nu^2) &= \sum_{i=1, \dots, d}^{+\infty} \{C_{6,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2, \tau, \tau + \nu^1, \tau + \nu^2) \\ &+ \{r_{2,\mathbf{k}\mathbf{X}}(\nu^1) C_{4,\mathbf{k}\mathbf{X}}(\tau - \nu^2, \tau + \nu^1 - \nu^2, \nu^2)\}_{15} \\ &+ \{r_{2,\mathbf{k}\mathbf{X}}(\nu^1) r_{2,\mathbf{k}\mathbf{X}}(\tau - \nu^2) r_{2,\mathbf{k}\mathbf{X}}(\nu^2 - \nu^1)\}_{15}\}. \end{aligned}$$

Maintenant pour montrer (4.6) nous utilisons la Proposition 6.3.9 de [15]. Pour cela, remarquons que ;

$$\begin{aligned} &N^{\frac{d}{2}} [\hat{r}_{3X}(\nu^1, \nu^2) - r_{3X}(\nu^1, \nu^2)] \\ &= N^{\frac{d}{2}} \{\hat{r}_{3X}(\nu^1, \nu^2) - \hat{r}_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2) \\ &- E[\hat{r}_{3X}(\nu^1, \nu^2) - \hat{r}_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2)]\} \\ &+ N^{\frac{d}{2}} [\hat{r}_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2) - r_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2)] \\ &= T_{k,N} + Z_{k,N}, \end{aligned}$$

où $T_{k,N} = N^{-\frac{d}{2}} \{\hat{r}_{3X}(\nu^1, \nu^2) - \hat{r}_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2) - E[\hat{r}_{3X}(\nu^1, \nu^2) - \hat{r}_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2)]\}$ et $Z_{k,N} = N^{-\frac{d}{2}} [\hat{r}_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2) - r_{3,\mathbf{k}\mathbf{X}}(\nu^1, \nu^2)]$. Mais selon la Proposition 6.3.9 de [15], il nous reste à prouver que :

(i) $Z_{k,N} \Rightarrow Z_k$ quand $N \rightarrow \infty$ pour chaque $k = 1, 2, \dots$

(ii) $Z_k \Rightarrow Z$ quand $k \rightarrow \infty$, et

(iii) $\lim_{k \rightarrow \infty} \lim_{N \rightarrow \infty} \sup P(|T_{k,N}| > \epsilon) = 0$ pour tout $\epsilon > 0$,

après quoi nous pouvons immédiatement déduire (4.6).

D'abord, (4.4) établit (i). Maintenant prouvons (iii). Selon l'inégalité de Chebychev, pour prouver

(iii) il suffit de montrer que

$$\lim_{k \rightarrow \infty} \lim_{N \rightarrow \infty} \sup \text{Var}(T_{k,N}) = 0. \quad (4.7)$$

En effectuant quelques calculs, nous obtenons

$$\begin{aligned} & N^{\frac{d}{2}} \{ \hat{r}_{3X}(\nu^1, \nu^2) - \hat{r}_{3,\mathbf{k}X}(\nu^1, \nu^2) \} \\ &= N^{-\frac{d}{2}} \sum_{i=1, \dots, d}^N \left\{ \underbrace{X_{\mathbf{t},\mathbf{k}} X_{\mathbf{t}+\nu^1, \mathbf{k}} u_{\mathbf{t}+\nu^2, \mathbf{k}}}_{S_1} + \underbrace{X_{\mathbf{t},\mathbf{k}} u_{\mathbf{t}+\nu^1, \mathbf{k}} X_{\mathbf{t}+\nu^2, \mathbf{k}}}_{S_2} \right. \\ & \left. + \underbrace{u_{\mathbf{t},\mathbf{k}} X_{\mathbf{t}+\nu^1, \mathbf{k}} X_{\mathbf{t}+\nu^2, \mathbf{k}}}_{S_3} \right\} = S_1 + S_2 + S_3, \end{aligned}$$

où $u_{\mathbf{t},\mathbf{k}} = X_{\mathbf{t}} - X_{\mathbf{t},\mathbf{k}} = \sum_{i=1, \dots, d}^{|s_i| > k} \psi_{\mathbf{s}} \varepsilon_{\mathbf{t}-\mathbf{s}}$. C'est assez facile de montrer que quand $k \rightarrow \infty$,

$$\text{Var}(S_1 + S_2 + S_3) \leq 3 \{ \text{Var} S_1 + \text{Var} S_2 + \text{Var} S_3 \} \rightarrow 0,$$

ce qui donne (4.7) et donc (iii). Il reste à montrer. De (4.4), $Z_k \sim \mathcal{N}(0, \sigma_{X,\mathbf{k}}^2(\nu^1, \nu^2))$. Remarquons

que si on pose $\sigma_X^2(\nu^1, \nu^2) = \lim_{k \rightarrow \infty} \sigma_{X,\mathbf{k}}^2(\nu^1, \nu^2)$, ce n'est alors pas difficile de montrer que

$$\begin{aligned} \sigma_X^2(\nu^1, \nu^2) &= \sum_{i=1, \dots, d}^{+\infty} \{ C_{6X}(\nu^1, \nu^2, \tau, \tau + \nu^1, \tau + \nu^2) + \\ & \{ r_{2X}(\nu^1) r_{2X}(\tau - \nu^2) r_{2X}(\nu^2 - \nu^1) \}_{15} \} < \infty, \end{aligned}$$

Alors, en utilisant le résultat du Problem 6.16 de [15] $Z_k \Rightarrow Z$, où $Z \sim \mathcal{N}(0, \sigma_X^2(\nu^1, \nu^2))$. De

(i), (ii) et (iii)

$$N^{\frac{d}{2}} (\hat{r}_{3X}(\nu^1, \nu^2) - r_{3X}(\nu^1, \nu^2)) \rightarrow \mathcal{N}(0, \sigma_X^2(\nu^1, \nu^2)).$$

■

Chapitre 5

Identification des modèles 2D RIF MA

Dans la suite, nous proposons deux approches pour l'identification des paramètres de processus 2D MA non gaussiens. La première approche est basée sur des cumulants expérimentaux seuls et sans imposer de phase minimum. Alors qu'une méthode généralisée des moments (GMM) est utilisée dans la seconde approche.

5.1 Identification des modèles 2D RIF MA par les cumulants d'ordre trois seuls

Dans cette section, nous abordons le problème de l'identification des modèles moyenne mobile 2D (MA) avec erreurs non-Gaussiennes basée sur les cumulants expérimentaux seuls et sans phase minimale. L'algorithme proposé est une extension, au cas 2D, de celui proposé dans [82] pour le cas unidimensionnel.

Considérons le processus 2D RIF MA donné par l'équation suivante :

$$x(n, m) = \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b(i, j) w(n - i, m - j), \quad (5.1)$$

où $b(i, j)$, $i = 1, \dots, q_1$, $j = 1, \dots, q_2$ sont les coefficients MA inconnus, et le signal est observé en bruit additif :

$$y(n, m) = x(n, m) + v(n, m). \quad (5.2)$$

On admet les hypothèses suivantes :

i] La série de bruit génératrice $w(n, m)$, inobservée, est un processus i.i.d. non-Gaussien de moyenne nulle avec au moins un cumulants $0 < |\gamma_{m,w}| < \infty$, $m > 2$.

ii] Le système est de phase non minimale (c.à.d. $x(n, m)$ est non inversible).

iii] Le bruit additif $v(n, m)$ est un modèle ARMA Gaussien avec spectre de puissance inconnu et est indépendant de l'entrée $w(n, m)$ et donc des sorties $x(n, m)$.

Pour obtenir les conditions de phase minimale pour le modèle (5.1), nous avons besoin de l'écrire sous une forme polybomiale de la manière suivante : soit $\mathbf{B} = (B_1, B_2)$ l'opérateur retard, pour n'importe quel $\mathbf{i} = (i_1, i_2) \in \mathbb{Z}^2$, $\mathbf{z} = (z_1, z_2) \in \mathbb{C}^2$, considérons $B^{\mathbf{i}}X(t_1, t_2) = B_1^{i_1} B_2^{i_2} X(t_1, t_2) = X(t_1 - i_1, t_2 - i_2)$ et $\mathbf{z}^{\mathbf{i}} = z_1^{i_1} z_2^{i_2}$.

Définissons $\Theta(\mathbf{z}) = \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b(i, j) z_1^i z_2^j$, alors le modèle (5.1) peut être écrit sous forme condensée comme suit :

$$x(n, m) = \Theta(\mathbf{B}) w(n, m).$$

(5.1) est de phase minimale (ou inversible) si la condition suivante est satisfaite :

\mathbf{C}_0 . aucunes des racines de $\Theta(z_1, z_2)$ ne se trouvent dans le polydisque unité fermé ($|z_1| \leq 1, |z_2| \leq 1$).

Si nous considérons un support de prédiction semi-infini de (5.1) d'ordre connu (q_1, q_2) , nous pouvons déduire une relation entre les paramètres MA $b(i, j)$, $i = 0, \dots, q_1$, $j = 0, \dots, q_2$, en arrangeant ses termes par l'ordre lexicographique, et certaines fonctions des statistiques. En effet, il n'est pas difficile de voir que sous la condition (iii),

$$C_{ky}(\nu^1; \dots; \nu^{k-1}) = C_{kx}(\nu^1; \dots; \nu^{k-1}), \quad (5.3)$$

et que les cumulants d'ordre k sont liés à la réponse impulsionnelle par

$$C_{ky}(\nu^1; \dots; \nu^{k-1}) = \gamma_{k,w} \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} h(i, j) \times h[(i, j) + \nu^1] \times \dots \times h[(i, j) + \nu^{k-1}], \quad (5.4)$$

avec $\nu^1 = (t_1, t_2)$, $\nu^2 = \dots = \nu^{k-1} = \mathbf{0}$ (où $\mathbf{0} = (0, 0)$) et $k = 3$ dans (5.4)), nous avons alors,

$$C_{3y}((t_1, t_2); \mathbf{0}) = \gamma_{3,w} \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} h^2(i, j) h(i + t_1, j + t_2) \quad (5.5)$$

Nous substituons $h(i + t_1, j + t_2) = b(i + t_1, j + t_2)$ dans (5.5) pour obtenir une nouvelle équation normale qui est,

$$C_{3y}((t_1, t_2), \mathbf{0}) = \gamma_{3,w} \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b(i, j) \times h^2(i - t_1, j - t_2), \text{ pour tout } (t_1, t_2) \quad (5.6)$$

où nous avons utilisé $h(i, j) = b(i, j) = 0$ pour $i < 0$ ou $j < 0$ ou $i > q_1$ ou $j > q_2$. Rappelons [71] que

$$h(i, j) = b(i, j) = \frac{C_{3y}((q_1, q_2); (i, j))}{C_{3y}((q_1, q_2); \mathbf{0})}. \quad (5.7)$$

Par substitution de (5.7) dans (5.6) il vient

$$\begin{aligned} & \gamma_{3,w} \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b(i, j) C_{3y}^2((q_1, q_2); (i - t_1, j - t_2)) \\ &= C_{3y}((t_1, t_2); \mathbf{0}) \times C_{3y}^2((q_1, q_2); \mathbf{0}), \text{ pour tout } (t_1, t_2) \end{aligned} \quad (5.8)$$

Nous nous référons à (5.8) comme l'équation normale basée sur les cumulants spatiaux pour un modèle 2D MA pur. Pour retirer $\gamma_{3,w}$ de l'équation (5.8) nous utilisons encore (5.5) et (5.7) pour obtenir

$$\gamma_{3,w} = \frac{C_{3y}^2((q_1, q_2); \mathbf{0})}{C_{3y}((q_1, q_2); (q_1, q_2))}, \quad (5.9)$$

où nous avons utilisé $b(0, 0) = 1$ et $C_{3y}((q_1, q_2); (t_1, t_2)) = 0$ pour $t_1 < 0$ ou $t_2 < 0$ ou $t_1 > q_1$ ou $t_2 > q_2$, maintenant par la substitution de (3.9) dans (3.8), nous obtenons

$$\begin{aligned} & \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b(i, j) C_{3y}^2((q_1, q_2); (i - t_1, j - t_2)) \\ &= C_{3y}((t_1, t_2); \mathbf{0}) \times C_{3y}((q_1, q_2); (q_1, q_2)). \end{aligned} \quad (5.10)$$

Si nous organisons les vecteurs et les matrices dans l'ordre lexicographique, concaténant l'équation normale (5.10) pour $-q_1 \leq t_1 \leq q_1$ et $-q_2 \leq t_2 \leq q_2$ nous obtenons l'équation matricielle simple suivante

$$\begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{bmatrix} \mathbf{b} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix}, \quad (5.11)$$

voir l'annexe D pour les expressions de \mathbf{C}_1 , \mathbf{C}_2 , \mathbf{c}_1 et \mathbf{c}_2 .

Le théorème suivant garantit l'identification des paramètres 2D RIF MA (sans imposer de phase minimum) en résolvant (5.11).

Théorème 5.1 *Sous les conditions i) -iii), étant donné les vrais cumulants des sorties $C_{3y}(\mathbf{q}; \mathbf{t})$ et $C_{3y}(\mathbf{r}; \mathbf{0})$ du modèle (5.1) et (5.2) pour les retards ($0 \leq t_1 \leq q_1$ et $0 \leq t_2 \leq q_2$) et ($-q_1 \leq \tau_1 \leq q_1$ et $-q_2 \leq \tau_2 \leq q_2$) respectivement, les $(q_1 + 1)(q_2 + 1)$ inconnues $b(i, j)$, $i = 0, \dots, q_1, j = 0, \dots, q_2$ peuvent être récupérées par la solution de (5.11) d'une manière unique*

Preuve La preuve est très similaire à la preuve du Théorème 1 de [82] dans le cas du modèle 1D RIF MA et donc omise. ■

En pratique, seuls les estimateurs empiriques de $C_{3y}((q_1, q_2); (t_1, t_2))$ et $C_{3y}((\tau_1, \tau_2); \mathbf{0})$ sont disponibles à partir l'échantillon fini $\{y(n, m), 1 \leq n \leq N_1, 1 \leq m \leq N_2\}$. Cependant, il sera possible de remplacer les cumulants théoriques par les cumulants expérimentaux lors de la résolution de (5.11) pour \mathbf{b} ce qui aiderait à l'identifier. Ainsi, l'algorithme d'identification des paramètres 2D MA est résumé dans l'algorithme I comme suit.

Algorithme I.

Etape 1 : A partir des sorties $y(n, m)$, $1 \leq n \leq N_1$, $1 \leq m \leq N_2$, former les estimateurs

$$\hat{C}_{3y}((q_1, q_2); (t_1, t_2)) = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1 - q_1} \sum_{j=1}^{N_2 - q_2} y(i, j) \times \\ y(i + q_1, j + q_2) y(i + t_1, j + t_2);$$

pour $0 \leq t_1 \leq q_1$ et $0 \leq t_2 \leq q_2$,

et

$$\hat{C}_{3y}((\tau_1, \tau_2); \mathbf{0}) = \\ \frac{1}{N_1 N_2} \sum_{i=\max(1, -\tau_1)}^{\min(N_1, N_1 - \tau_1)} \sum_{j=\max(1, -\tau_2)}^{\min(N_2, N_2 - \tau_2)} y^2(i, j) \times \\ y(i + \tau_1, j + \tau_2),$$

pour $-q_1 \leq \tau_1 \leq q_1$ et $-q_2 \leq \tau_2 \leq q_2$.

Etape 2 : En utilisant les estimateurs $\hat{C}_{3y}((q_1, q_2); (t_1, t_2))$ et $\hat{C}_{3y}((\tau_1, \tau_2); \mathbf{0})$, résoudre de manière itérative le système (5.11) pour obtenir les estimateurs $\hat{b}(i, j)$ de $b(i, j)$ $i = 0, \dots, q_1$, $j = 0, \dots, q_2$.

Etape 3 : Enfin pour satisfaire la condition normalisée $b(0, 0) = 1$, nous prenons $\hat{b}_1(i, j) = \frac{\hat{b}(i, j)}{\hat{b}(0, 0)}$ comme estimateur final de $b(i, j)$ $0 \leq i, j \leq 1$, $(i, j) \neq (0, 0)$.

Remarque 5.1 *Le Théorème (4.1) et le Théorème (4.2) ont établi la convergence presque sûre et la normalité asymptotique des estimations $\hat{C}_{3y}((q_1, q_2); (t_1, t_2))$ et $\hat{C}_{3y}((\tau_1, \tau_2); \mathbf{0})$. Cette forte consistance garantit que l'estimation des paramètres $b(i, j)$, $(i = 0, \dots, q_1, j = 0, \dots, q_2)$ aboutit à des estimateurs consistants fortement à condition que le processus de sortie $y(i, j)$ satisfasse les conditions de stationnarité du Théorème (4.1). En effet, les deux paramètres 2D MA sont identifiés comme des fonctions mesurables de $C_{3y}((q_1, q_2), (t_1, t_2))$ et $C_{3y}((\tau_1, \tau_2); \mathbf{0})$, et $\hat{C}_{3y}((q_1, q_2); (t_1, t_2))$ et $\hat{C}_{3y}((\tau_1, \tau_2); \mathbf{0})$, sont des estimateurs consistants de $C_{3y}((q_1, q_2); (t_1, t_2))$ et $C_{3y}((\tau_1, \tau_2); \mathbf{0})$ respectivement. Par conséquent, les paramètres du modèle 2D MA seront eux-mêmes consistants. De plus, la distribution asymptotique de nos estimateurs de paramètres sera gaussienne.*

Remarque 5.2 Dans les applications pratiques, le vrai ordre (q_1, q_2) n'est pas connu a priori.

Dans la preuve du théorème (5.1), on montre que le rang de la matrice cumulée $\begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{bmatrix}$ est $R = (q_1 + 1)(q_2 + 1)$. Alors, (q_1, q_2) peut être déterminé en utilisant la décomposition en valeurs singulières (SVD) en testant

$$\hat{C}_{3y}^R((q_1, q_2); \mathbf{0}) \neq 0, \hat{C}_{3y}^{\varkappa+1}((k_1, k_2); \mathbf{0}) \approx 0, \quad (5.12)$$

où $\varkappa = (k_1 + 1)(k_2 + 1)$. Voir [81] pour plus de détails sur le SVD basé sur (5.12).

5.2 Identification des modèles 2D RIF MA par l'estimateur GMM

Dans cette section, nous examinerons les méthodes généralisées d'estimation des moments (GMM) qui fournissent des estimateurs pratiques et efficaces. Comme motivation alternative, les estimateurs GMM sont souvent disponibles là où une analyse de vraisemblance est extrêmement difficile, voire impossible (par exemple, s'il y a beaucoup d'incertitude sur la forme distributionnelle). L'estimateur GMM est dérivé directement d'un ensemble d'hypothèses minimales, les conditions dites de moment que le modèle devrait satisfaire. D'autre part, une difficulté pratique avec la méthode habituelle de ML est son coût de calcul sur une taille d'échantillon moyenne à grande. De plus, l'estimation GMM inclut certains estimateurs bien connus en tant que cas spéciaux, de sorte qu'ils peuvent être analysés commodément à un niveau théorique. Dans un cadre spatial, Kelejian et Prucha (1999) ont introduit un estimateur de moments généralisés pour le paramètre autorégressif du processus de perturbation spatiale qui est simple à calculer et reste réalisable sur le plan calcul même pour de grands échantillons. Récemment (2010), ils ont introduit une nouvelle classe d'estimateurs GMM pour le paramètre autorégressif d'un processus de perturbation spatialement autorégressive qui permet des innovations hétéroscédastiques. Ils ont prouvé que leurs estimateurs GMM sont encore une fois simples en termes de calcul, même dans de grands échantillons. Par conséquent, notre objectif est de proposer des estimateurs GMM pour les paramètres d'un modèle 2D RIF MA qui sont simples à calculer et qui ont des propriétés

d'échantillonnage souhaitables.

5.2.1 Aperçu sur la méthode GMM

Dans cette section, nous allons considérer une méthodologie GMM. Pour introduire la classe des estimateurs GMM, introduisons le vecteur $N_1 N_2 \times 1$ de variables aléatoires $\underline{w}(i, j)$ observé à la position (i, j) , soit $\underline{\theta} \in \Omega$ un vecteur inconnu de coefficients de dimension $a \times 1$, et soit $h(\underline{\theta}, \underline{w}(i, j))$ une fonction vectorielle de dimension $r \times 1$. Définissons pour

$$\mathbf{N} = (N_1, N_2), g_{\mathbf{N}}(\underline{\theta}) = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} h(\underline{\theta}, \underline{w}(i, j)).$$

De plus supposons que $\underline{\theta}_0$ est la vraie valeur de $\underline{\theta}$ et que les conditions sur les moments peuvent être formulé par

$$E_{\underline{\theta}_0} [h(\underline{\theta}, \underline{w}(i, j))] = 0.$$

L'estimateur GMM est défini par la résolution du problème de minimisation

$$\hat{\underline{\theta}}_{\mathbf{N}} = \arg \min_{\underline{\theta} \in \Omega} [g_{\mathbf{N}}(\underline{\theta})]' W_{\mathbf{N}} [g_{\mathbf{N}}(\underline{\theta})],$$

où la matrice de poids $W_{\mathbf{N}}$ est définie positive.

Comme cadre convenable, les hypothèses suivantes ont été admises.

I.1 Le processus $(\underline{w}(i, j), (i, j) \in \mathbb{Z}^2)$ st strictement stationnaire, ergodic et admet des moments d'ordre quatre.

I.2 L'espace des paramètres Ω est compact et $\underline{\theta}_0$ est dans son intérieur.

I.3 Pour tout $\underline{\theta} \in \Omega$, $h(\underline{\theta}, \cdot)$ est une fonction mesurable.

I.4 $h(\underline{\theta}, \underline{w}(i, j))$ est continue et continuellement différentiable en $\underline{\theta}$ sur un voisinage de $\underline{\theta}_0$.

I.5 $W_{\mathbf{N}}$ est une suite de matrices $r \times r$ positives qui convergent en probabilité vers une matrice positive définie déterministe W .

I.6 $|h_k(\underline{\theta}, \underline{w}(i, j))| \leq f(\underline{w}(i, j))$ pour tout $\underline{\theta} \in \Omega$ et $k = 1, \dots, r$ où $f(\cdot)$ est une fonction non négative tel que $E_{\underline{\theta}_0} \{f(\underline{w}(i, j))\} < \infty$.

I.7 $\underline{\theta}_0$ est l'unique solution de $E_{\underline{\theta}_0} [h(\underline{\theta}, \underline{w}(i, j))] = 0$ c-à-d, $E_{\underline{\theta}_0} [h(\underline{\theta}, \underline{w}(i, j))] = 0$ si et seulement si $\underline{\theta} = \underline{\theta}_0$.

I.8 $\sqrt{N_1 N_2} g_{\mathbf{N}}(\underline{\theta}_0) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}(0, \mathcal{S})$.

I.9 Pour toute suite $(\underline{\theta}_{\mathbf{N}}^*, \mathbf{N} \in \mathbb{N}^2)$ tel que $\underline{\theta}_{\mathbf{N}}^* \xrightarrow{P} \underline{\theta}_0$, $\left. \frac{\partial g_{\mathbf{N}}(\underline{\theta})}{\partial \underline{\theta}'} \right|_{\underline{\theta}_{\mathbf{N}}^*}$ converge en probabilité vers \mathcal{D}' , où $\mathcal{D}' = E \left. \frac{\partial h(\underline{\theta}, \underline{w}(i, j))}{\partial \underline{\theta}'} \right|_{\underline{\theta}_0}$ avec les colonne de \mathcal{D}' lineairement independants.

Soit $Dg_{\mathbf{N}}(\underline{\theta})$ une matrice de dimension $r \times a$ de dérivées avec les éléments typiques $Dg_{\mathbf{N}}(\underline{\theta})_{ij} = \frac{\partial g_{\mathbf{N}}(\underline{\theta})_i}{\partial \theta_j}$. Alors $\hat{\underline{\theta}}_{\mathbf{N}}$ peut aussi être défini en termes de la solution du premier ordre du problème de minimisation

$$Dg_{\mathbf{N}}(\hat{\underline{\theta}}_{\mathbf{N}})' W_{\mathbf{N}} [g_{\mathbf{N}}(\hat{\underline{\theta}}_{\mathbf{N}})] = 0.$$

Quand $r = a$, le modèle est exactement identifiable, la matrice de poids devient sans importance et $\hat{\underline{\theta}}_{\mathbf{N}}$ est défini par l'équation $g_{\mathbf{N}}(\hat{\underline{\theta}}_{\mathbf{N}}) = 0$.

Etant données ces hypothèses, l'estimateur GMM est consistant et asymptotiquement normalement distribué avec

$$\sqrt{N_1 N_2} (\hat{\underline{\theta}}_{\mathbf{N}} - \underline{\theta}_0) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}(0, \mathcal{V}), \quad (5.13)$$

où $\mathcal{V} = (\mathcal{D}\mathcal{S}^{-1}\mathcal{D}')^{-1}$.

Avant d'appliquer ce résultat au modèle 2D RIF MA, il est utile de, d'abord, commenter le sens et le rôle de ces conditions de régularité. L'hypothèse **(I.1)** est nécessaire pour vérifier le théorème central limite donné dans **(I.8)** et appliquer le théorème ergodic plusieurs fois. L'hypothèse **(I.2)** restreint les paramètres à un sous-ensemble fermé et borné de l'espace euclidien, ainsi, lorsqu'il est vérifié, l'estimateur peut être caractérisé en fonction de sa condition de premier ordre Les hypothèses **(I.3)** et **(I.4)** sont des conditions de régularité mathématique qui garantissent que les fonctions de moment sont continues et ont des variances finies, et que dans un voisinage de $\underline{\theta}_0$ ils peuvent être différenciés. L'hypothèse **(I.7)** est une condition d'identification clé qui exclut à la fois l'échec d'identification locale. Les hypothèses **(I.9)** et **(I.8)** donnent (5.13).

5.2.2 Application de la procédure GMM au modèle 2D RIF MA

Les résultats de la sous section 2.2 suggèrent une seconde estimation dans le cadre de la méthode GMM. Dans le cas du modèle 2D RIF MA (1) donné par (5.1) et (5.2), nous pouvons dériver des conditions d'orthogonalité pour estimer $\gamma = \left(\frac{b^2(0,0)}{b(1,1)}, \frac{b^2(0,1)}{b(1,1)}, \frac{b^2(1,0)}{b(1,1)}, b(1,1) \right)'$ à partir de (5.4), (5.7) et (5.9). En effet, pour $k = 3$, $\nu^1 = (-t_1, -t_2)$ et $\nu^2 = \mathbf{0}$ l'équation (5.4) donne,

$$C_{3y}((-t_1, -t_2); \mathbf{0}) = \gamma_{3,w} \sum_{i=0}^1 \sum_{j=0}^1 h^2(i, j) h(i - t_1, j - t_2),$$

comme $C_{3y}((-t_1, -t_2); \mathbf{0}) = C_{3y}((t_1, t_2); (t_1, t_2))$, et par l'usage de (5.7) et (5.9) nous obtenons,

$$\begin{aligned} C_{3y}((t_1, t_2); (t_1, t_2)) &= \frac{b^2(0,0)}{b(1,1)} C_{3y}((q_1, q_2); (-t_1, -t_2)) \\ &+ \frac{b^2(0,1)}{b(1,1)} C_{3y}((q_1, q_2); (-t_1, 1 - t_2)) \\ &+ \frac{b^2(1,0)}{b(1,1)} C_{3y}((q_1, q_2); (1 - t_1, -t_2)) \\ &+ b(1,1) C_{3y}((q_1, q_2); (1 - t_1, 1 - t_2)), \end{aligned}$$

pour $0 \leq t_1 \leq 1$ et $0 \leq t_2 \leq 1$. Ceci représente un système de quatre conditions d'orthogonalité donné par,

$$\begin{cases} h(\gamma, (i, j)) = A(i, j) \times \gamma - H(i, j) \\ E(h(\gamma, (i, j))) = 0 \end{cases}$$

où

$$\begin{aligned} A(i, j) &= y(i, j)y(i + 1, j + 1) \\ &\times \begin{pmatrix} y(i, j) & y(i, j + 1) & y(i + 1, j) & y(i + 1, j + 1) \\ y(i, j - 1) & y(i, j) & y(i + 1, j - 1) & y(i + 1, j) \\ y(i - 1, j) & y(i - 1, j + 1) & y(i, j) & y(i, j + 1) \\ y(i - 1, j - 1) & y(i - 1, j) & y(i, j - 1) & y(i, j) \end{pmatrix}, \end{aligned}$$

et

$$H(i, j) = y(i, j) \times \begin{pmatrix} y^2(i, j) \\ y^2(i, j + 1) \\ y^2(i + 1, j) \\ y^2(i + 1, j + 1) \end{pmatrix}.$$

Etant donné les observations $\{y(n, m), 1 \leq n \leq N_1, 1 \leq m \leq N_2\}$, définissons $g_{\mathbf{N}}(\boldsymbol{\gamma}, (i, j)) = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} h(\boldsymbol{\gamma}, (i, j))$, où $\mathbf{N} = (N_1, N_2)$, et supposons que $\boldsymbol{\gamma}_0$, la vraie valeur de $\boldsymbol{\gamma}$, est à l'intérieur de l'espace des paramètres compact Γ . L'estimateur GMM de $\boldsymbol{\gamma}$ est alors défini en résolvant le problème de minimisation,

$$\hat{\boldsymbol{\gamma}}_{\mathbf{N}}^{GMM} = \arg \min_{\boldsymbol{\gamma} \in \Omega} [g_{\mathbf{N}}(\boldsymbol{\gamma}, (i, j))] W_{\mathbf{N}} [g_{\mathbf{N}}(\boldsymbol{\gamma}, (i, j))], \quad (5.14)$$

où la matrice de poids $W_{\mathbf{N}}$ est définie positive.

Comme il y a quatre paramètres pour quatre conditions d'orthogonalité, le système est exactement identifié et l'estimateur GMM $\boldsymbol{\gamma}_{\mathbf{N}}^{GMM}$ de $\boldsymbol{\gamma}$ basé sur les $N_1 N_2$ observations est défini par l'équation $g_{\mathbf{N}}(\boldsymbol{\gamma}_{\mathbf{N}}^{GMM}, (i, j)) = 0$. Ainsi, il suffit de résoudre le système $\frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} h(\boldsymbol{\gamma}_{\mathbf{N}}^{GMM}, (i, j)) = 0$, qui peut être réécrite sous la forme

$$\left[\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} A(i, j) \right] \times \boldsymbol{\gamma}_{\mathbf{N}}^{GMM} = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} H(i, j),$$

par conséquent

$$\boldsymbol{\gamma}_{\mathbf{N}}^{GMM} = \left[\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} A(i, j) \right]^{-1} \times \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} H(i, j). \quad (5.15)$$

La consistance et la normalité asymptotique de $\boldsymbol{\gamma}_{\mathbf{N}}^{GMM}$ peuvent être obtenus après vérification de quelques conditions de régularité. D'abord, une bonne manière de choisir la matrice de poids $W_{\mathbf{N}}$ qui résout le problème de minimisation (5.14) est en posant

$$W_{\mathbf{N}} \equiv \tilde{W}_{\mathbf{N}} = \left[\frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} h(\boldsymbol{\gamma}_{\mathbf{N}}^{(1)}, (i, j)) h'(\boldsymbol{\gamma}_{\mathbf{N}}^{(1)}, (i, j)) \right]^{-1}$$

où $\gamma_{\mathbf{N}}^{(1)} = \left(\frac{\hat{b}_1^2(0,0)}{\hat{b}_1(1,1)}, \frac{\hat{b}_1^2(0,1)}{\hat{b}_1(1,1)}, \frac{\hat{b}_1^2(1,0)}{\hat{b}_1(1,1)}, \hat{b}_1(1,1) \right)'$ avec $\hat{b}_1(0,0)$, $\hat{b}_1(0,1)$, $\hat{b}_1(1,0)$, $\hat{b}_1(1,1)$ étant les estimateurs consistants de $b(0,0)$, $b(0,1)$, $b(1,0)$, $b(1,1)$, respectivement, calculés dans Algorithm I. Comme $h(\gamma, (i, j))$ est une fonction mesurable de $X(i, j)$, $0 \leq i, j \leq 1$ et γ est supposé être tel que les conditions du Theorem (5.1) sont vérifiées, nous obtenons

$$\tilde{W}_{\mathbf{N}} \xrightarrow{P} \tilde{W} = [Eh(\gamma_0, (i, j)) h'(\gamma_0, (i, j))]^{-1},$$

il est clair que $\tilde{W}_{\mathbf{N}}$ et \tilde{W} sont définis positives. Donc, la consistance et la normalité asymptotique de $\gamma_{\mathbf{N}}^{GMM}$ peuvent être établies dans le théorème suivant :

Théorème 5.2 *Considérons un modèle 2D RIF MA(1) et supposons que les conditions du Theorem (5.1) sont vérifiées. Alors,*

$$\sqrt{N_1 N_2} (\gamma_{\mathbf{N}}^{GMM} - \gamma_0) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \Sigma),$$

où $\Sigma = V^{-1} \Omega V^{-1}$ avec $V = E(A(i, j))$ et $\Omega = \tilde{W}^{-1}$.

Preuve A la lumière de (5.15) nous pouvons écrire,

$$(\gamma_{\mathbf{N}}^{GMM} - \gamma_0) = \left[\frac{1}{\sqrt{N_1 N_2}} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} A(i, j) \right]^{-1} \times \frac{1}{\sqrt{N_1 N_2}} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} h(\gamma_0, (i, j)), \quad (5.16)$$

donc ce n'est pas difficile de montrer que, sous les conditions du Théorème (5.1), $h(\gamma_0, (i, j))$ est une différence de martingale bidimensionnelle avec une variance finie $\Omega = \tilde{W}^{-1}$. Ainsi, d'après le théorème central limite des différences de martingale de Huang $\frac{1}{\sqrt{N_1 N_2}} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} h(\gamma_0, (i, j))$ converge vers une distribution normale multivariée de moyenne zero et de matrice de covariance Ω . D'après (5.16) et la procédure de Cramer-Wold device, le résultat s'en suit. ■

Maintenant pour dériver l'estimation du paramètre d'intérêt \mathbf{b} à partir de $\gamma_{\mathbf{N}}^{GMM}$, notons d'abord que $\gamma_{\mathbf{N}}^{GMM}(i)$ et $\gamma_{\mathbf{N}}^{GMM}(4)$ $i = 1, 2, 3$ ont le même signe donc nous calculons simplement directement $\hat{b}_{GMM}(1, 1) = \gamma_{\mathbf{N}}^{GMM}(4)$; $|\hat{b}_{GMM}(0, 0)| = [\gamma_{\mathbf{N}}^{GMM}(1) \gamma_{\mathbf{N}}^{GMM}(4)]^{\frac{1}{2}}$; $|\hat{b}_{GMM}(0, 1)| = [\gamma_{\mathbf{N}}^{GMM}(2) \gamma_{\mathbf{N}}^{GMM}(4)]^{\frac{1}{2}}$ et $|\hat{b}_{GMM}(1, 0)| = [\gamma_{\mathbf{N}}^{GMM}(3) \gamma_{\mathbf{N}}^{GMM}(4)]^{\frac{1}{2}}$ en gardant les signes des estimateurs de $b(0, 0)$, $b(0, 1)$ et $b(1, 0)$ donnés dans l'Algorithm I.

Finalement pour vérifier la condition de normalité $b(0,0) = 1$ nous prenons $\frac{\hat{b}_{GMM}(i,j)}{\hat{b}_{GMM}(0,0)}$ comme estimateur final de $b(i,j)$ $0 \leq i, j \leq 1, (i,j) \neq (0,0)$.

5.3 Résultats de simulation

Dans cette section, nous évaluons la proximité des estimateurs des paramètres de \mathbf{b} à leurs valeurs attendues par une étude de Monte Carlo et comparons les performances des deux estimateurs proposés. Cette étude permet de vérifier et de corroborer, d'une part, l'efficacité de l'algorithme d'identification précédent et d'autre part la validité des résultats théoriques proposés présentés dans les sections précédentes.

Considérons le modèle 2D MA de premier ordre,

$$\begin{aligned} x(n, m) = & b_1 w(n, m-1) + b_2 w(n-1, m) + \\ & b_3 w(n, m-2) + w(n, m), \end{aligned} \quad (5.17)$$

Le bruit additif $v(n, m)$ est un processus ARMA Gaussien donné par

$$\begin{aligned} v(n, m) = & 0.3v(n, m-1) - 0.08v(n-2, m-1) + \\ & 0.54 + e(n, m) - 1.25e(n, m-1), \end{aligned}$$

le signal $x(n, m)$ est observé en bruit additif :

$$y(n, m) = x(n, m) + v(n, m),$$

où $e(n, m) \sim \mathcal{N}(0, 0.5)$ et $w(n, m)$ sont i.i.d $(0, 1)$ et suit une distribution lambda généralisée (GLD) avec une asymétrie $\kappa_3 = 0$ et un kurtosis $\kappa_4 = 0,85$. Les vraies valeurs des paramètres et $(b_1 = 1, b_2 = 0,25, b_3 = -1,65)$ (le cas de phase non minimale) et $(b_1 = 0,25, b_2 = 0,35, b_3 = 0,15)$ (le cas de phase minimale), Proposition 1 de [3] donne un ensemble convenable de conditions sur les paramètres b_1, b_2 et b_3 pour satisfaire C_0). Notez que dans l'algorithme d'identification proposé, la phase minimale n'est pas imposée.

De plus, les erreurs avec des caractéristiques non gaussiennes sont la clé de l'identification. C'est pourquoi nous choisissons une distribution GLD erreurs considérées dans [58] qui peut être simulée de manière flexible et peut accepter une large gamme de valeurs pour les paramètres d'asymétrie et d'aplatissement excessif (Kurtosis).

Les Tableaux 1 et 2 présentent les colonnes moyennes, des colonnes des erreurs quadratiques moyennes, des colonnes d'aplatissement et des colonnes d'asymétrie des estimateurs b_1, b_2 et b_3 sur 500 répétitions de Monte Carlo pour les tailles d'échantillon (100,100), (500,500) et (1000,1000) dans la phase minimale et la phase non minimale respectivement. Les valeurs moyennes de l'estimateur donné par l'algorithme I b_i et l'estimateur GMM b_i^{GMM} , ($i = 1,2,3$), dans les deux cas, sur les 500 répliques étaient toujours proche des valeurs vraies pour toutes les tailles de grille.

De plus et comme prévu, l'augmentation de la taille de l'échantillon améliore considérablement les résultats en termes de réduction des biais. De plus, les moyennes des coefficients d'aplatissement et de asymétrie sur les 500 répétitions ont tendance à être plus proches des valeurs de référence 0 et 0,85 respectivement lorsque la taille de l'échantillon augmente pour les deux estimateurs.

Les observations des tableaux 1 et 2 confirment la cohérence et la normalité asymptotique des estimateurs proposés comme indiqué dans la remarque (5.1) et le théorème (5;2). Elles suggèrent également que l'estimateur GMM produit des estimations de \mathbf{b} qui sont un peu plus précises que celles données par l'algorithme I. Cependant, malgré la charge de calcul raisonnable de l'estimateur GMM de \mathbf{b} , un avantage majeur de l'estimateur donné par l'algorithme I est qu'il ne nécessite pas d'estimateur cohérent préliminaire.

Chapitre 5. Identification des modèles 2D RIF MA

Tableau1 : Echantillon de distribution des estimateurs \hat{b}_1 , \hat{b}_2 et \hat{b}_3 et \hat{b}_1^{gmm} , \hat{b}_2^{gmm} et \hat{b}_3^{gmm}

dans le cas de phase minimale avec $\kappa_3 = 0.85$ et $\kappa_4 = 3$.

$N =$			
(100, 100)			
<i>true b_i</i>	0.25	0.35	0.15
	\hat{b}_1 (\hat{b}_1^{gmm})	\hat{b}_2 (\hat{b}_2^{gmm})	\hat{b}_3 (\hat{b}_3^{gmm})
<i>mean</i>	0.2501 (0.2498)	0.3502 (0.3501)	0.150 (0.153)
<i>rmse</i>	0.18×10^{-4} (0.12×10^{-4})	0.0029 (0.002)	0.0038 (0.0025)
<i>kurt</i>	2.5910 (3.2985)	3.1432 (3.176)	2.879 (2.764)
<i>skew</i>	0.0151 (0.0099)	0.0092 (0.0111)	0.0098 (0.0089)

$N =$			
(500, 500)			
<i>true b_i</i>	0.25	0.35	0.15
	\hat{b}_1 (\hat{b}_1^{gmm})	\hat{b}_2 (\hat{b}_2^{gmm})	\hat{b}_3 (\hat{b}_3^{gmm})
<i>mean</i>	0.25009 (0.2500)	0.3501 (0.3497)	0.15006 (0.1502)
<i>rmse</i>	0.119×10^{-4} (0.099×10^{-4})	0.0015 (0.0014)	0.0029 (0.0025)
<i>kurt</i>	2.7910 (3.321)	2.9143 (3.156)	3.1787 (2.8691)
<i>skew</i>	0.0105 (0.0098)	0.0100 (0.0113)	0.0091 (0.0079)

$N =$			
(1000, 1000)			
<i>true b_i</i>	0.25	0.35	0.15
	\hat{b}_1 (\hat{b}_1^{gmm})	\hat{b}_2 (\hat{b}_2^{gmm})	\hat{b}_3 (\hat{b}_3^{gmm})
<i>mean</i>	0.25001 (0.2489)	0.35005 (0.35005)	0.150 (0.1499)
<i>rmse</i>	0.119×10^{-4} (0.074×10^{-4})	0.0015 (0.0013)	0.0029 (0.0025)
<i>kurt</i>	2.7910 (3.146)	2.9143 (3.09)	3.1787 (3.234)
<i>skew</i>	0.0105 (0.0098)	0.0100 (0.0054)	0.0091 (0.0123)

Tableau 2 : Echantillon de distribution des estimateurs \hat{b}_1 , \hat{b}_2 et \hat{b}_3 et \hat{b}_1^{gmm} , \hat{b}_2^{gmm} et \hat{b}_3^{gmm}

dans le cas non minimum avec $\kappa_3 = 0.85$ et $\kappa_4 = 3$.

$N =$ (100, 100)

<i>true b_i</i>	1	0.25	-1.65
	\hat{b}_1 (\hat{b}_1^{gmm})	\hat{b}_2 (\hat{b}_2^{gmm})	\hat{b}_3 (\hat{b}_3^{gmm})
<i>mean</i>	1.0003 (1.0005)	0.2499 (0.2503)	-1.6501 (-1.6498)
<i>rmse</i>	0.9229×10^{-4} (0.9×10^{-4})	0.00218 (0.002)	0.00079 (0.00062)
<i>kurt</i>	3.2110 (2.763)	2.8942 (3.245)	3.1787 (3.0985)
<i>skew</i>	0.0099 (0.0134)	0.0210 (0.0125)	0.0081 (0.0099)

$N =$ (500, 500)

<i>true b_i</i>	1	0.25	-1.65
	\hat{b}_1 (\hat{b}_1^{gmm})	\hat{b}_2 (\hat{b}_2^{gmm})	\hat{b}_3 (\hat{b}_3^{gmm})
<i>mean</i>	1 (0.999)	0.2499 (0.2508)	-1.6502 (-1.65)
<i>rmse</i>	0.8589×10^{-4} (0.812×10^{-4})	0.00191 (0.0019)	0.00059 (0.0005)
<i>kurt</i>	2.9902 (3.146)	2.6942 (2.885)	3.0840 (3.1328)
<i>skew</i>	0.0089 (0.0121)	0.0110 (0.0085)	0.0071 (0.0099)

$N =$ (1000, 1000)

<i>true b_i</i>	1	0.25	-1.65
	\hat{b}_1 (\hat{b}_1^{gmm})	\hat{b}_2 (\hat{b}_2^{gmm})	\hat{b}_3 (\hat{b}_3^{gmm})
<i>mean</i>	1 (1)	0.25005 (0.2481)	-1.65004 (-1.65)
<i>rmse</i>	0.8143×10^{-4} (0.835×10^{-4})	0.00132 (0.0011)	0.00055 (0.00043)
<i>kurt</i>	3.002 (3.0904)	2.9941 (3.281)	3.1840 (2.871)
<i>skew</i>	0.0009 (0.0012)	0.0120 (0.0088)	0.0069 (0.005)

5.4 Conclusion

Ce chapitre présente de nouvelles approches pour prendre en compte des données complexes de traitement de signaux non linéaires et / ou non gaussiens. L'originalité de ce travail a été l'investigation des propriétés asymptotiques de statistiques d'ordre élevé dans un contexte spatial. Les propriétés de convergence des moments de troisième et quatrième ordre et les estimations des cumulants ont été décrites et prouvées. L'étude a également présenté de nouvelles façons d'identifier un système à phase non minimale 2-D linéaire en utilisant des cumulants d'ordre supérieur seuls et avec des estimateurs GMM. Une étude de Monte Carlo a conduit à un estimateur de \mathbf{b} , un avantage majeur de l'estimateur donné par l'algorithme I est qu'il ne nécessite pas d'estimateur cohérent préliminaire.

APPENDICE

Posons $\mathbf{q} = (q_1, q_2)$ et $R = (q_1 + 1)(q_2 + 1)$, l'expression de \mathbf{C}_1 , \mathbf{C}_2 , \mathbf{c}_1 et \mathbf{c}_2 sont :

$$\mathbf{c}_1 = \begin{pmatrix} C_{3y}((-q_1, -q_2), \mathbf{0}) \times C_{3y}(\mathbf{q}, \mathbf{q}) \\ C_{3y}((-q_1, -q_2 + 1), \mathbf{0}) \times C_{3y}(\mathbf{q}, \mathbf{q}) \\ \vdots \\ C_{3y}((0, -1), \mathbf{0}) \times C_{3y}(\mathbf{q}, \mathbf{q}) \end{pmatrix}_{(R-1) \times 1}$$

$$\mathbf{c}_2 = \begin{pmatrix} C_{3y}((0, 0), \mathbf{0}) \times C_{3y}(\mathbf{q}, \mathbf{q}) \\ C_{3y}((0, 1), \mathbf{0}) \times C_{3y}(\mathbf{q}, \mathbf{q}) \\ \vdots \\ C_{3y}(\mathbf{q}, \mathbf{0}) \times C_{3y}(\mathbf{q}, \mathbf{q}) \end{pmatrix}_{R \times 1}$$

$$\mathbf{C}_1 = \begin{pmatrix} C_{3y}^2(\mathbf{q}, \mathbf{q}) & 0 & \dots & 0 & 0 \\ \vdots & C_{3y}^2(\mathbf{q}, \mathbf{q}) & & & \\ \vdots & & \ddots & & \\ C_{3y}^2(\mathbf{q}, (0, 1)) & \dots & \dots & C_{3y}^2(\mathbf{q}, \mathbf{q}) & 0 \end{pmatrix}_{(R-1) \times R}$$

$$\mathbf{C}_2 = \begin{pmatrix} C_{3y}^2(\mathbf{q}, \mathbf{0}) & C_{3y}^2(\mathbf{q}, (0, 1)) & \dots & C_{3y}^2(\mathbf{q}, \mathbf{q}) \\ 0 & C_{3y}^2(\mathbf{q}, \mathbf{0}) & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & C_{3y}^2(\mathbf{q}, \mathbf{0}) \end{pmatrix}_{R \times R} .$$

Bibliographie

- [1] AKAIKE. H. (1966). “Note on higher spectra”, *Ann. of the Inst. of Stat. Math.*, vol.18, pp. 123–126.
- [2] Bashkansky, M., and Reintjes, J. (2000). Statistics and reduction of speckle in optical coherence tomography. *Optics Letters*, **25** (8), 545–547.
- [3] Basu, R., and Reinsel, G. C. (1993). Properties of the spatial unilateral first-order ARMA model. *Adv. Appl. Prob.* **25**, 631-648.
- [4] Besag, J. E., (1972). On the correlation structure of some two dimensional stationary processes. *Biometrika*. **59**, 43-48.
- [5] Besag, J. E., (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B* **36**, 192–23.
- [6] Billinger. D. R, (1981) *Time Series, Data Analysis and Theory*, Holden-Day.
- [7] Billinger. D. R, Rosenblatt M., (1966). *Asymptotic theory of kth-order spectra. In : Harris B (ed) Spectral analysis of time series.* Wiley, New York : 189–232.
- [8] Billinger. D. R, Rosenblatt M., (1967). *Asymptotic theory of kth-order spectra. Harris B (ed) Spectral analysis of time series* pp. 153–188.
- [9] Blanc-Lapierre, A et Fortet, R. (1953). *Théorie des fonctions aléatoires.* Ed. Masson, Paris. pp 48.
- [10] Boucher A., (2009). Considering complex training images with search tree partitioning. *Comput Geosci*, **35**,1151–1158.

Bibliographie

- [11] Box, G et Jenkins, G. (1970). Time series analysis : Forecasting and Daviscontrol. Holden Day.
- [12] Brillinger, D.R. (1965). Introduction to Polyspectra. *Ann. Math. Statist.* **36** :1351-74.
- [13] Brillinger, D. R. (1981). Time Series, Data Analysis and Theory, Holden-Day.
- [14] Briuinger, D.R., and Rosenblatt, M. (1967). Computation and interpretation of k-th order spectra. In *SpectralAnalysis of Time Series*. Edited by B. Harris. New York : Wiley. 189-232.
- [15] Brockwell.
- [16] Brodatz, P. (1966). *Textures : A Photographic Album for Artists and Designers*. New York : Dover Publications.
- [17] Burckhardt C.B. (1978). Speckle in Ultrasound B Mode Scans, *IEEE Trans. Sonics Ultrasonics*, **25**, 1-6.
- [18] Chugunova TL, and Hu Ly., (2008). Multiple-point simulations constrained by continuous auxiliary data. *Math Geosci*, **40**,133–146.
- [19] Cramer. H. (1946). Mathematical methods of statistics, Princeton University.
- [20] Cullagh, P. Mc. (1987). Tensor Methods in Statistics, Monographs on Statistics and Applied Probability. Chapman and Hall.
- [21] De Lathauwer. L. (1997). Signal processing based on multilinear algebra. Thèse de doctorat, Université Catholique de Leuven.
- [22] Dimitrakopoulos R., Mustapha H., and Gloaguen, E., (2010). High-order Statistics of Spatial Random Fields : Exploring Spatial Cumulants for Modeling Complex Non-Gaussian and Non-linear Phenomena. *Math Geosci*, **42**, 65–99.
- [23] Francos, J. M., Narasimhan, A., and Woods, J. W. (1995) Maximum likelihood parameter estimation of textures using a Wold-decomposition based model. *IEEE Transactions on Image Processing*, **4** (12), 1655 1666. ;
- [24] Francos, J.M. and Nehorai, A. (2003). Interference mitigation in STAP using the two-dimensional wold decomposition model. *IEEE Tran. Signal Proces.*, **51**, 2461-2470.
- [25] Giannakis, G. B. (1987). “Cumulants : A powerful tool in signal processing”, Proceedings of The IEEE, vol. 75, no. 9, pp. 1333–1334.

Bibliographie

- [26] Giannakis, G. B, et Mendel, J. M (1989). Identification of non-minimum phase systems using higherorder statistics. *IEEE Trans. Acoust. Speech Signal Process.*, 37 no. 3, 360–377.
- [27] Giannakis, G. B. & Mendel, J. M. (1990) Cumulant-based order determination of non-Gaussian ARMA models. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **38** (8), 1411-1423.
- [28] Giannakis, G. B, et Swami, A. (1990). On estimating non-causal non-minimum phase ARMA models of non-gaussian processes. *IEEE Trans. Acoust. Speech Signal Process.*, 38 no. 3, 478–495.
- [29] Giannakis, G. B. & Swami, A, (1992). Identifiability of general ARMA processes using linear cumulant-based estimators. *Automatica*, **28** (4), 771-779.
- [30] Gloaguen, E., and Dimitrakopoulos, R, (2009). Two-dimensional conditional simulations based on the wavelet decomposition of training images. *Math Geosci*, **41**(6), 679–701.
- [31] Guyon, X. (1995) *Random fields on a net work*. Probability and its applications. Springer, New York.
- [32] Holmquist, B. (1988). Moments and cumulants of the multivariate normal distribution. *Stochastic Analysis and Applications*, 6, 273–278.
- [33] Hongwei, L., and Qiansheng, C. (1998). Convergence properties of higher-order moment and cumulant estimates. *Journal of electronics*, **15** (3), 242-247.
- [34] Huang, D. E.A. Swanson, C.P. Lin, J.S. Schuman, W.G. Stinson, W. Chang, M.R. Hee, T. Flotte, K. Gregory, C.A. Pulifito, and J.G. Fujimoto, (1991). *Optical coherence tomography. Science*, 254 :1178–1181.
- [35] Kendall, M. G, et Stuart, A. (1963). *The Advanced Theory of Statistics, Volume 1 - Distribution Theory, Second Edition*. Charles Griffin and Company Ltd, London. pp 70.
- [36] Kendall, M. G, et Stuart, A. (1963). *The Advanced Theory of Statistics, Volume 1 - Distribution Theory, Second Edition*. Charles Griffin and Company Ltd, London, chap 13.
- [37] Kourgly A., Belhadj-Aissa A., (2003). Segmentation texturale des images urbaines par le biais de l’analyse variographique, *Téledétection*, **3**, (2-3-4), 337–348.

Bibliographie

- [38] Kourgly A., Analyse variographique et synthèse de textures appliquées aux images numériques, Thèse de Doctorat d'Etat, (2007). Université des Sciences et de la Technologie « Houari Boumediene », Algérie, 181 pages.
- [39] Krishnamurthy, R., Woods, J. W., and Francos, J. M. (1996). Adaptive restoration of textured images with mixed spectra. *IEEE Trans. Image Process*, **5** (4), 648-52.
- [40] Kshirsahar, A. N. (1972). Multivariate Analysis. Marcel Dekker Inc., New York, chap 2.
- [41] Lacoume, J. L, Amblard, P. O et Comon. P. (1997). Statistiques d'ordre supérieur pour le traitement du signal. Ed. Masson, Paris, chap 1.
- [42] LACOUME, J. L, GAETA, M, AMBLARD, P. O. (1992). "From order 2 to HOS : new tools and applications", in Proc. European Signal Processing Conf. EUSIPCO, Brussels, pp. 91-98.
- [43] Lii, K.S., and Rosenblatt, M. (1982). Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes. *Ann. Star.* **10**,1195-1208.
- [44] Liu, F., and Picard, R.W. (1996). Detecting and segmenting periodic motion. Media lab vision and modelling tr-400, MIT.
- [45] Mendel, J. M (1991). Tutorial on higher-order statistics (spectra) in signal processing and system theory : Theoretical results and some applications, *Proc. IEEE*, 79 (3) : 278-305.
- [46] Mirowski, P.W., Trtzlaff, D.M., Davies, R.C., McCormick, D.S., Williams, N., and Signer, C, (2008). Stationary scores on training images for multipoint geostatistics. *Math Geosci*, **41**, 447-474.
- [47] Nikias, C. L.(1988). ARMA bispectrum approach to nonminimum phase system identification. *IEEETrans. Acoust. Speech Signal Process.*, 36 no. 4, 513-524.
- [48] Nikias, C. L, and Mendel, J. M, (1993). Signal processing with higher-order spectra, *IEEE Signal Processing Magazine*, **10** (3), 10-37.
- [49] NIKIAS, C. L, PETROPULU, A.B. (1993). Higher-Order Spectra Analysis, Signal Processing Series. Prentice-Hall, Englewood Cliffs.
- [50] NIKIAS, C.L, RAGHUVVEER, M. (1985). "Bispectrum estimation : a digital signal processing framework", Proc. of IEEE, vol. 75, pp. 869-891.

Bibliographie

- [51] Numéro spécial non gaussien, non linéaire”, *Traitement du Signal*, Vol.7, 5, (1990).
- [52] Picinbono, B. (1994). *Signaux aléatoires : Tome 2 : Fonctions aléatoires et modèles*. Ed. Dunod,Paris.
- [53] Proc. 1st International workshop on higher-order statistics, Vail, Colorado, U.S.A., june (1989).
- [54] Proc. 2nd International workshop on higher-order statistics, Chamrousse, France, july (1991).
- [55] Proc. 3rd International workshop on higher-order statistics, Lake Tahoe, California, U.S.A., july (1993).
- [56] Proc. IEEE Signal Processing ATHOS Workshop on Higher-Order statistics, Begur, Girona, Spain, june (1995).
- [57] Proc. IEEE Signal Processing Workshop on Higher-Order statistics, Banff, Canada, july (1997).
- [58] Ramberg, J. and Schmeiser, B. (1975). An Approximate Method for Generating Asymmetric Random Variables. *Communications of the ACM*, **17**(2), 78-82.
- [59] Remy, N., Boucher, A., and Wu, J, (2009). *Applied geostatistics with SGeMS : A user's guide*. Cambridge University Press, Cambridge.
- [60] Ripley, B. D., (1981). *Spatial Statistics*. John Wiley & Sons, New York.
- [61] Rosenblatt, M, (1985). *Stationary Sequences and Random Fields*. Birkhäuser, Boston. MR0885090.
- [62] Serfling, R.G. (1980). *Approximations Theorems of Mathematical Statistics*. Wiley.
- [63] Schmitt, J.M., Xiang, S.H., and Yung, K.M. (1999). Speckle in optical coherence tomography. *J. Biomed. Opt.*, 4 :95–105.
- [64] Shiriyayen. A. (1960). “Some problems in the spectral theory of higher-order moments”, *Th. Prob. and Appl.*, vol. 5, no. 3, 1960.
- [65] ”Special issue higher-order statistics”, *IEEE trans. ASSP*, vol.38, 7, (1990).
- [66] “Special issue higher-order statistics”, *IEEE trans. AC*, vol.35, 1, (1990).

Bibliographie

- [67] “Special issue on applications of high-order statistics”, IEE Proceedings-Part F, vol.140, 6, (1993).
- [68] “Special issue higher-order statistics”, Signal Processing, vol.36, 3, (1994).
- [69] “Special issue higher-order statistics”, Journ. Franklin Institute, vol. 333(B), 3, (1996).
- [70] “Special issue higher-order statistics”, Signal Processing, vol.53, 2-3, (1996).
- [71] Swami, A., (1990). Modeling of multi-dimensional non-Gaussian processes using cumulants. *J. Multidimensional Signals and Systems*, **1**, 1-37.
- [72] Swami, A., Giannakis, G.B., and Mendel, J.M, (1989). Modeling and Parameter Estimation of Multidimensional Non-Gaussian Processes Using Cumulants. *Multidimensional Systems and Signal Processing.*, **1** (1), 11-37.
- [73] TJOOSTHEIM, D. (1978). Statistical spatial séries modelling. *Advances in Appl. Probability*, 10(1) :130-154.
- [74] TJOOSTHEIM D. (1983). Statistical spatial séries modelling. II Some further results on unilatéral lattice processes. *Adv. in Appl. Probab.*, 15(3) :562-584.
- [75] Treitz, P., and Howarth, P, (2000). Hight spatial remote sensing data for forest ecosystem classification : an examination of spatial scale. *Remote sensing of environment*, **72**, 268-289.
- [76] Vakhania, N.N et Kandelaki, N.B. (1996). Random vectors with values in complex Hilbert spaces. *Theory Probab. Appl.*, 41 no. 1, 116–131.
- [77] Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. New York, NY, USA : Springer-Verlag.
- [78] Wold, H. (1938). A study in the Analysis of Stationary Time Series. Almgrist and Wiksell.
- [79] Wollstein, G., Paunescu, L. A., Ko, T. H., Fujimoto, J. G., Kowaleviev, A., Hartl, I., Beaton, S., Ishikawa, H., Mtttox, C., Singh, O., Duker, J., Drexler, W., and Schuman, J. S. (2005). Ultrahigh-resolution optical coherence tomography in glaucoma. *Ophthalmology*, **112** (2), 229-237.
- [80] Yuan, J., (1999) Testing Gaussianity and linearity for random fields in the frequency domain. *J. Times Series Analysis*. **21** (6), 723-737

Bibliographie

- [81] Zhang, X. D., and Zhang, Y. S. (1993) Singular value decompositionbased MA order determination of non-Gaussian ARMA models. *IEEE Trans. Signal Processing*, **41**, 2657-2664.
- [82] Zhang, X. D., and Zhang, Y. S. (1994) FIR system identification using higher-order statistics alone. *IEEE Trans. Signal Processing*, **42**, 2854-2858.
- [83] Zeimer, R. C., Wilensky, J. T., Gieser, D. K., and Viana, M. A. (1991) Association between intraocular pressure peaks and progression of visual field loss. *Ophthalmology*, **98**, 64-69.