HAL
archives-ouvertes.fr

Fonction d'autocorrélation partielle des processus à temps discret non stationnaires et applications

Sophie Lambert-Lacroix

▶ **To cite this version:**

Sophie Lambert-Lacroix. Fonction d'autocorrélation partielle des processus à temps discret non stationnaires et applications. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1998. Français. tel-00004893

HAL Id: tel-00004893
https://tel.archives-ouvertes.fr/tel-00004893

Submitted on 19 Feb 2004

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

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

THÈSE

présentée par

Sophie LAMBERT-LACROIX

pour obtenir le titre de
Docteur de l'Université Joseph Fourier - Grenoble 1
(*arrêtés ministériels du 5 juillet 1984 et du 30 Mars 1992*)

Spécialité : **Mathématiques Appliquées**

FONCTION D'AUTOCORRÉLATION PARTIELLE DES PROCESSUS À TEMPS DISCRET NON STATIONNAIRES ET APPLICATIONS

Thèse soutenue le 8 juillet 1998 devant la Commission d'Examen

A. LE BRETON	Président du jury
E. MOULINES	Rapporteur
M.C. VIANO	Rapporteur
D.T. PHAM	Examinateur
S. DÉGERINE	Directeur de thèse

Thèse préparée au sein du **L**aboratoire de **M**odélisation et **C**alcul, Institut IMAG

REMERCIEMENTS

Je remercie tout d'abord Monsieur Alain LE BRETON, Professeur à l'Université Joseph Fourier de Grenoble, qui me fait l'honneur de présider le jury.

Mes remerciements vont à Madame Marie-Claude VIANO, Professeur à l'Université de Lille 1, et à Monsieur Éric MOULINES, Professeur à l'ENST Paris, d'avoir accepté de rapporter sur cette thèse. Je les remercie de l'intérêt qu'ils ont témoigné à ce travail.

Je remercie également Monsieur Dinh Tuan PHAM, Directeur de recherche au CNRS, d'avoir accepté de prendre part au jury.

Enfin, je tiens à remercier particulièrement Monsieur Serge DÉGERINE, Professeur à l'Université Joseph Fourier de Grenoble, qui est à l'origine de cette thèse et grâce à qui j'ai pu mener ce travail à son terme. Merci pour son implication tout au long de cette thèse, pour son soutien et ses encouragements.

Table des matières

Introduction **7**

1 Aspect temporel **17**
1.1 Corrélation partielle 18
1.2 Fonction d'autocorrélation partielle 20
1.2.1 Définitions et notations 21
1.2.2 Domaine de variation de $\beta(\cdot,\cdot)$ 22
1.2.3 Décomposition de Wold-Cramér 26
1.3 Algorithme de Levinson-Durbin Généralisé 29
1.4 Modèles autorégressifs 34

2 Aspect spectral **41**
2.1 Spectre évolutif instantané 41
2.2 Propriétés du spectre évolutif instantané 44
2.3 Comparaison avec le spectre évolutif 49
2.4 Comparaison avec le spectre rationnel 51
2.5 Exemples 56
2.5.1 Discrétisation du mouvement Brownien 56
2.5.2 Succession de deux autorégressifs stationnaires 60
2.5.3 Processus uniformément modulé 63
2.5.4 Chirps linéaire 64
2.5.5 Partie réelle d'un chirps linéaire 66

3 Processus périodiquement corrélés **69**
3.1 Fonction d'autocorrélation partielle scalaire 70
3.2 Fonction d'autocorrélation partielle matricielle triangulaire . . 74
3.2.1 Procédé de normalisation et définition 75
3.2.2 Domaine de variation de $\Pi(\cdot)$ 77
3.2.3 Algorithme de type Levinson-Durbin 80
3.3 Correspondance entre $\beta(\cdot,\cdot)$ et $\Pi(\cdot)$ 83
3.4 Modèles autorégressifs périodiques 85

3.5 Extension par la méthode du maximum d'entropie 89
3.5.1 Le problème d'extension 89
3.5.2 La méthode du maximum d'entropie 90
3.5.3 Fonctions de cyclo-corrélation 92

4 Estimation des modèles *ARP* 95
4.1 Les approches scalaires 96
4.1.1 Méthode de Yule-Walker 96
4.1.2 Autocorrélations partielles empiriques 97
4.1.3 Rapports d'énergies résiduelles 99
4.1.4 Méthodes de type Burg 100
4.1.5 Comparaison des méthodes 101
4.2 Les approches vectorielles 104
4.2.1 Méthode de Yule-Walker 105
4.2.2 Matrices d'autocorrélation partielle empiriques triangulaires 105
4.2.3 Rapports d'énergies résiduelles 108
4.2.4 Extensions de la méthode de Burg 109
4.3 Résultats de simulation 114
4.3.1 Présentation des résultats 115
4.3.2 Étude comparative 116

Annexe 131

Bibliographie 147

Introduction

La structure au second ordre d'un processus est usuellement caractérisée par la fonction d'autocovariance. Dans la situation stationnaire, la paramétrisation par la fonction d'autocorrélation partielle est relativement récente. Nous étendons ce résultat au cas non stationnaire. L'avantage de cette fonction est d'être facilement identifiable par rapport à la fonction d'autocovariance qui doit être de type positif. De plus elle conduit naturellement à la définition d'un nouveau spectre dépendant du temps. Ce dernier décrit, à chaque instant, une situation stationnaire dans laquelle le présent est corrélé avec le passé de la même façon que le processus non stationnaire au même instant. L'étude de ses propriétés montre l'intérêt de cette nouvelle approche par rapport à celles de deux autres spectres de même nature. On se restreint ensuite à la classe particulière des processus périodiquement corrélés. La fonction d'autocorrélation partielle fournit une nouvelle paramétrisation qui permet, en particulier, d'étendre de façon naturelle la méthode du maximum d'entropie à cette situation. Enfin nous considérons l'estimation autorégressive dans le cadre de ces processus. L'extension de la méthode des autocorrélations partielles empiriques constitue une nouvelle estimation de ces paramètres. La comparaison avec les procédures existantes, effectuée sur le plan théorique mais aussi par simulation, montre que cette méthode présente de bonnes propriétés. Nous étudions également le lien entre cette approche et celle du cas vectoriel stationnaire.

La notion de corrélation partielle a été introduite au début du siècle par Yule [YUL07]. La fonction d'autocorrélation partielle sera notée $\beta(\cdot,\cdot)$ où $\beta(t,s)$ désigne la corrélation partielle entre $X(t)$ et $X(s)$ dans l'ensemble $\{X(s),\ldots,X(t)\}$. Ce paramètre, souvent appelé coefficient de réflexion, joue un rôle important en traitement du signal et correspond à une notion physique. Pour les processus périodiquement corrélés, et dans le cas non localement déterminable, les quantités introduites par Sakai [SAK83] sous le nom de "coefficients d'autocorrélation partielle normalisés" sont celles données par $\beta(\cdot,\cdot)$. Les paramètres $\beta(t,s)$ sont aussi les éléments clés de la généralisa-

tion des algorithmes de Schur et de Levinson (cf. [KAI82]) aux processus non stationnaires et non localement déterminables. Néanmoins le fait que la fonction $\beta(\cdot,\cdot)$ constitue une autre paramétrisation de la structure au second ordre semble être ignoré.

Dans le cas stationnaire, cette propriété a été établie par Ramsey [RAM 74]. Il s'agit en fait d'un résultat plus ancien d'analyse où les autocorrélations partielles constituent les paramètres d'un système de polynômes orthogonaux sur le cercle unité (cf. [GER60]). C'est dans la situation stationnaire que le paramètre $\beta(k)$, qui correspond alors à $\beta(t, t \Leftrightarrow k)$, est souvent utilisé en statistique. On le rencontre notamment dans le test de Quenouille [QUE49], qui est basé sur le fait que $\beta(k)$ est aussi le coefficient de $X(t \Leftrightarrow k)$ dans la régression de $X(t)$ sur $\{X(t \Leftrightarrow 1), \ldots, X(t \Leftrightarrow k)\}$. C'est aussi une notion très importante pour les modèles autorégressifs. En effet la fonction d'autocorrélation partielle d'un processus autorégressif d'ordre p est caractérisée par $\beta(k) = 0$ pour $k > p$, et Box et Jenkins [BOX94] suggèrent d'utiliser sa représentation graphique à des fins de modélisation. Les autocorrélations partielles sont également à la base de l'algorithme de Levinson [LEV47] et Durbin [DUR60].

Ce document comprend quatre chapitres. Dans le premier, nous présentons la fonction d'autocorrélation partielle dans la situation non stationnaire générale. Le spectre dépendant du temps qui lui est naturellement associé fait l'objet du deuxième chapitre. Le troisième concerne l'étude de la structure des processus périodiquement corrélés et le dernier est consacré à l'estimation autorégressive dans ce cadre. Nous présentons ci-dessous de façon détaillée le contenu de ces chapitres tout en situant les résultats dans leur contexte.

Dans le premier chapitre, après un rappel sur la notion de corrélation partielle, nous introduisons la fonction d'autocorrélation partielle d'un processus non stationnaire. Quelques unes de ses propriétés sont présentées. Un point fondamental est que cette fonction caractérise la structure au second ordre du processus, au même titre que la fonction d'autocovariance $R(\cdot,\cdot)$. L'avantage de la fonction $\beta(\cdot,\cdot)$ est qu'elle est facilement identifiable : son module est inférieur ou égal à 1, l'égalité traduisant les singularités d'ordre fini. Dans ce cas, le processus sera dit localement déterminable. La fonction $R(\cdot,\cdot)$, quand à elle, doit être de type positif. Cette contrainte, assurée par la positivité d'un ensemble de matrices, est clairement plus difficile à contrôler. Le domaine de variation de la fonction $\beta(\cdot,\cdot)$ est noté $\mathcal{D}_\beta$. Partant d'un élément de cet ensemble, nous donnons un procédé constructif qui permet de simuler une séquence d'un processus de structure au second ordre qui lui correspond, avec un coût de stockage réduit. En particulier ce résultat assure la surjectivité dans l'ensemble $\mathcal{D}_\beta$ de l'application qui à $R(\cdot,\cdot)$ associe $\beta(\cdot,\cdot)$. L'extension de l'algorithme de Levinson-Durbin (LDG) au cas

général (éventuellement localement déterminable) donne la correspondance entre la fonction d'autocorrélation partielle et la fonction d'autocovariance. Les relations de cet algorithme permettent d'obtenir l'injectivité de cette application. Par ailleurs, on constate que la bijection entre $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$ est de nature temporelle au sens où leurs restrictions à un domaine de la forme $[s,\ldots,t]^2$ sont également en correspondance biunivoque. La nouvelle approche en termes d'autocorrélations partielles est illustrée en considérant les processus autorégressifs facilement caractérisés, comme dans la situation stationnaire, par cette paramétrisation.

Le deuxième chapitre, intitulé aspect spectral, est consacré à la présentation d'un nouveau spectre et à l'étude de ses propriétés. Il est délicat d'étendre la notion de mesure spectrale au cas non stationnaire. En effet, pour t fixé dans $\mathbb{Z}$, la fonction d'une seule variable $R_t(k) = R(t, t \Leftrightarrow k)$, $k \geq 0$, n'est plus nécessairement une fonction d'autocovariance. Par conséquent on ne peut lui associer une mesure via la transformée de Fourier. Par contre $\beta_t(k) = \beta(t, t \Leftrightarrow k)$, $k \geq 0$, est la fonction d'autocorrélation partielle d'un processus stationnaire. Ce résultat conduit à introduire de façon naturelle un spectre dépendant du temps. Ce dernier, que nous appelons spectre évolutif instantané, est défini par la suite $\{dF_t, t \in \mathbb{Z}\}$ de mesures sur $]\Leftrightarrow\pi, \pi]$ associée à $\{\beta_t(\cdot), t \in \mathbb{Z}\}$. Ainsi à chaque instant t, dF_t est la mesure spectrale d'un processus stationnaire pour lequel les autocorrélations partielles entre le présent et le passé sont identiques à celles du processus non stationnaire à cet instant.

Le spectre évolutif instantané est analysé à travers les propriétés souhaitées par Loynes [Loy68]. Nous constatons qu'il possède des propriétés intéressantes. Ce spectre est une fonction non négative du temps et de la fréquence, qui se réduit au spectre ordinaire lorsque le processus est stationnaire. Il est en correspondance biunivoque avec la fonction d'autocovariance. La distribution marginale de $dF_t(\lambda)$ est celle attendue : $\int_{-\pi}^{\pi} dF_t(\lambda) = Var\{X(t)\}$. D'autre part, lorsqu'on considère un processus constitué de la succession de deux parties stationnaires, on retrouve le spectre de la première et, moyennant des hypothèses de non corrélation entre les deux parties, on retrouve celui de la deuxième au bout d'un temps fini lorsqu'il s'agit d'un modèle autorégressif, sinon infini. Les propriétés qui concernent les transformations sur le processus sont dans l'ensemble vérifiées. Cependant, celles qui sont liées au retournement du temps ne le sont aucunement. Par ailleurs les opérations qui se traduisent facilement en termes d'autocovariance (comme une transformation linéaire) ne sont pas du tout en phase avec le spectre évolutif instantané. En effet, la correspondance entre $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$ est hautement non linéaire et c'est aussi le cas entre $\beta(\cdot,\cdot)$ et $dF_t(\lambda)$. Cependant la correspondance entre $R(\cdot,\cdot)$ et $dF_t(\lambda)$ étant bijective, les spectres

transformés sont toujours "calculables".

L'analyse du spectre évolutif instantané permet de le comparer à deux autres spectres qui lui sont proches. Celui de Mélard [MÉL78] est basé sur la décomposition de Wold-Cramér du processus alors que celui de Grenier [GRE84] repose sur sa représentation $ARMA$. On observe que ces deux derniers sont plus directement reliés à la description temporelle du processus, alors qu'aucune représentation du processus n'a été jusqu'ici associée au spectre évolutif instantané. Néanmoins, le spectre que nous proposons présente des avantages non négligeables. Il est défini en toute généralité alors que celui de Mélard n'existe que pour des processus purement indéterminables et celui de Grenier pour une classe encore plus restrictive. La différence importante est qu'il est en bijection avec la fonction d'autocovariance. En particulier, compte tenu de ses autres propriétés, son indépendance par rapport au temps est équivalente à la stationnarité du processus. Ce n'est plus vrai pour les deux autres alors qu'ils se ramènent au spectre ordinaire lorsque le processus est stationnaire. En effet deux structures au second ordre, l'une étant éventuellement stationnaire, peuvent donner lieu à un même spectre dans leur définition. Sinon les propriétés de Loynes sont dans l'ensemble vérifiées ou infirmées simultanément par les trois spectres. Le comportement du spectre évolutif instantané est illustré sur quelques exemples et lorsque cela est possible, il est comparé avec les deux autres. Notamment on constate qu'il est bien adapté dans le cas de la discrétisation du mouvement Brownien ainsi que pour les "chirps linéaires" pour lesquels il est porté par des droites. La fonction d'autocorrélation partielle et le spectre associé ont fait l'objet d'un rapport interne [DÉG96a] et ont été présentés au congrès IEEE : "International Symposium on Time-Frequency and Time-Scale Analysis" [DÉG96b].

Après cette étude de nature plutôt générale, nous considérons, dans les deux derniers chapitres, les processus périodiquement corrélés ainsi que l'estimation de leur structure. Cette classe de processus a été introduite par Gladysev [GLA61]. Elle est doublement intéressante. Sur le plan pratique elle joue un rôle essentiel en modélisation, en particulier des signaux de télécommunication. Sur le plan théorique elle fournit une nouvelle approche pour l'étude des processus vectoriels stationnaires. En effet, à chaque processus périodiquement corrélé, correspond un processus vectoriel stationnaire et inversement. Ainsi, que ce soit sur le plan de la modélisation ou de l'estimation, nous allons considérer les approches aussi bien scalaires (périodiques) que vectorielles.

L'étude de la paramétrisation de la structure au second ordre des processus périodiquement corrélés est considérée au troisième chapitre. Elle permet de mieux appréhender les méthodes d'estimation dans leur conception. Dans un premier temps, cette étude est réalisée en utilisant la fonction

d'autocorrélation partielle introduite au premier chapitre. En effet, un processus périodiquement corrélé de période T est un processus non stationnaire dont la fonction d'autocovariance $R(\cdot,\cdot)$ est périodique de même période. En d'autres termes, la structure de la séquence $\{X(s),\ldots,X(t)\}$ est la même que celle de $\{X(s+T),\ldots,X(t+T)\}$. En particulier, on retrouve les processus stationnaires lorsque $T=1$. Compte tenu de la nature de la bijection entre $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$, il est facile de constater que la même propriété sur les autocorrélations partielles, $\beta(t+T,s+T)=\beta(t,s)$, caractérise également la structure de ces processus. Ainsi on dispose d'une nouvelle paramétrisation qui, comme nous l'avons déjà constaté, est facilement identifiable. Nous donnons le nouvel ensemble de contraintes sur cette paramétrisation obtenu comme la restriction du domaine $\mathcal{D}_\beta$ aux fonctions périodiques. Dans le cadre de ces processus, la correspondance entre les fonctions $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$ est bijective sur des domaines autres que $[s,\ldots,t]^2$. En effet, du fait de la périodicité de ces fonctions, on est amené à les considérer sur les domaines $E(p_1,\ldots,p_T)=\{(t,s), t=1,\ldots,T, 0\leq t\Leftrightarrow s\leq p_t\}$. En particulier, ces ensembles interviennent de façon naturelle pour les modèles autorégressifs où ils généralisent ceux de la forme $[t\Leftrightarrow p,\ldots,t]$ du cas stationnaire. La bijection est alors assurée si et seulement si $p_{i+1}\leq p_i+1$, $i=1,\ldots,T$, $(p_{T+1}=p_1)$. Ces domaines contraints seront notés $E_c(p_1,\ldots,p_T)$. Notons que $[s,\ldots,t]^2$ fait partie de ces ensembles. La restriction de l'Algorithme LDG à la situation périodique (LDP) donne alors cette correspondance sur de tels domaines. Pour certains d'entre eux, cet algorithme coïncide avec celui proposé par Sakai [Sak83] et étendu par Pham [Pha92] au cas localement déterminable.

Nous nous intéressons ensuite à la paramétrisation de la structure des processus périodiquement corrélés par l'intermédiaire de celle liée aux processus vectoriels associés. L'extension de la fonction d'autocorrélation partielle aux processus vectoriels stationnaires est délicate. La difficulté pour définir une matrice d'autocorrélation partielle réside dans le choix de la racine carrée pour normaliser l'autocovariance partielle. Morf et coll. [Mor78a] proposent d'utiliser la racine carrée triangulaire. Dégerine [Dég90] donne les conditions que doivent satisfaire la racine carrée et son inverse pour que la correspondance entre fonction d'autocorrélation partielle et fonction d'autocovariance soit biunivoque et introduit la notion de fonction d'autocorrélation partielle canonique. On trouve ainsi plusieurs extensions de la fonction d'autocorrélation partielle selon le choix du procédé de normalisation. Nous retenons les fonctions dites triangulaires ([Dég90], [Pha92]) qui apparaissent très naturellement dans la correspondance avec les processus périodiquement corrélés. Nous proposons une définition proche de celle de Dégerine dans son approche en termes de variables, mais qui correspond plutôt à

celle de Pham sur le plan algébrique. Nous montrons que la structure du processus vectoriel peut être paramétrisée par ces matrices d'autocorrélation partielle, qui seront notées $\Pi(k)$. La démarche qui conduit à ce résultat est la même que celle utilisée pour la fonction $\beta(\cdot,\cdot)$. En particulier, un procédé constructif permet de simuler une séquence d'un processus vectoriel stationnaire de structure au second ordre donnée. Un algorithme de type Levinson-Durbin, approprié à cette fonction, donne sa correspondance avec la fonction d'autocovariance. Notre approche en termes de variables permet d'obtenir facilement la correspondance biunivoque entre les paramétrisations par les fonctions d'autocorrélation partielle scalaire $\beta(\cdot,\cdot)$ et vectorielle $\Pi(\cdot)$. Notons que cette approche est essentielle pour établir le lien entre les méthodes d'estimation basées sur la structure périodique et celles basées sur la structure vectorielle stationnaire associée.

Nous considérons les processus autorégressifs périodiques (ARP) introduits dans [JON67], dont l'estimation fait l'objet du dernier chapitre. Ces modèles sont une généralisation des processus autorégressifs stationnaires au cas périodique. Les coefficients du filtre ainsi que la variance résiduelle dépendent du temps, mais de façon périodique. Ainsi les paramètres de ces modèles sont constitués de T (si T est la période) filtres et T variances résiduelles, l'ordre étant donné par les T valeurs $(p_1,\ldots,p_T)$ qu'il peut prendre. Comme dans le cas stationnaire, ces modèles peuvent être paramétrisés par les premiers coefficients de $R(\cdot,\cdot)$ (ceux pour un retard inférieur à l'ordre) ou de façon équivalente par ceux de la fonction $\beta(\cdot,\cdot)$. La différence fondamentale entre ces deux caractérisations réside dans le prolongement de ces fonctions au delà de l'ordre du modèle. Pour la fonction $R(\cdot,\cdot)$, on utilise l'analogue des équations de Yule-Walker, alors que la fonction $\beta(\cdot,\cdot)$ est simplement prolongée par zéro. Par contre, contrairement au cas stationnaire, l'utilisation de $R(\cdot,\cdot)$ pose certains problèmes. En effet le prolongement par les équations de Yule-Walker des premiers coefficients d'une fonction d'autocovariance ne conserve pas toujours le caractère positif de cette fonction. Ainsi partant des autocovariances, il n'est pas toujours possible d'ajuster un modèle ARP. On constate là encore l'intérêt de la paramétrisation par $\beta(\cdot,\cdot)$, qui conduit alors à un procédé algorithmique permettant de vérifier s'il est possible d'ajuster un modèle ARP à des coefficients d'autocovariance donnés. Notons par ailleurs que le prolongement par zéro des autocorrélations partielles conserve les contraintes sur $\beta(\cdot,\cdot)$. On considère également la correspondance entre les modèles ARP et les processus vectoriels autorégressifs (VAR). On constate que les modèles ARP sont plus intéressants que les modèles VAR car ils comportent souvent moins de paramètres.

Dans le cas stationnaire, la méthode du maximum d'entropie fait partie des critères qui conduisent à l'estimation autorégressive (cf. [BUR75]).

Les autocorrélations partielles permettent d'étendre de façon naturelle cette méthode à la situation périodique. Dans un premier temps, nous montrons comment cette paramétrisation fournit une nouvelle approche pour résoudre un problème, classique dans le domaine de l'estimation spectrale, qui est celui de l'extension d'une suite de coefficients d'autocovariance. Nous proposons la formulation suivante : étant données les valeurs d'une fonction périodique $R(\cdot,\cdot)$ sur un domaine $E_c(p_1,\ldots,p_T)$, sous quelles conditions sont-elles les premières valeurs d'une fonction d'autocovariance ? Lorsque ces conditions sont satisfaites, comment décrire l'ensemble des extensions d'une telle fonction ? L'Algorithme *LDP*, qui établit la correspondance entre $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$, permet de traiter le problème précédent en terme d'autocorrélations partielles. Les réponses à ces questions sont alors immédiates. Notons que ce problème est une généralisation du problème d'extension d'une suite de matrices d'autocovariance $R_0,\ldots,R_n$, du cas vectoriel stationnaire. Pour les processus stationnaires scalaires, la méthode du maximum d'entropie a été proposée par Burg [Bur75]. Elle consiste à sélectionner la solution du problème d'extension pour laquelle la variance du processus d'innovation est maximum. Lorsque le processus est périodiquement corrélé, la variance du processus d'innovation, qui dépend du temps, est périodique de même période. Notre approche pour étendre la méthode du maximum d'entropie est la suivante : la solution est celle pour laquelle le produit de ces variances, en excluant celles déjà nulles dans le cas localement déterminable, est maximum. Nous constatons que notre méthode coïncide avec celle de Burg [Bur75] pour le cas vectoriel stationnaire et généralise celle résultant du cas dégénéré particulier de Inouye [Ino84]. Par ailleurs elle équivaut à affecter la valeur nulle à la fonction d'autocorrélation partielle là où elle n'est pas connue et par suite la solution est *ARP*. Notre approche est également comparée avec celle de Alpay et coll. [Alp95], qui considèrent le problème d'extension en termes de fonctions de cyclo-corrélation. Ce point de vue sur la méthode du maximum d'entropie dans le cadre des processus périodiquement corrélés a été exposé lors de la "VIII European Signal Processing Conference" [Lam96].

Le dernier chapitre concerne les méthodes qui permettent d'estimer la structure au second ordre des processus *ARP*. Dans la situation scalaire stationnaire, la notion de coefficient d'autocorrélation partielle joue un rôle important dans le domaine de l'estimation autorégressive. On peut citer la méthode de Burg [Bur75] ou celle de Dickinson [Dic78]. Une estimation empirique naturelle de ces paramètres a également été proposée par Dégerine ([Dég86], [Dég93]). Ces coefficients sont aussi à l'origine des méthodes de maximum de vraisemblance approché [Kay83] ou exact [Pha88]. Pour les processus périodiquement corrélés, deux approches sont possibles : scalaire ou vectorielle. En effet la correspondance entre modèles *ARP* et *VAR* étant

bijective, les méthodes d'estimation du cas vectoriel donnent lieu à des méthodes pour la situation périodique et inversement.

Concernant les approches scalaires, on trouve essentiellement trois méthodes. Celle de Pagano [PAG78] constitue une extension de la méthode de Yule-Walker. Elle consiste à ajuster le modèle $ARP(p_1, \ldots, p_T)$ dont la structure est donnée par les autocovariances empiriques biaisées usuelles sur le domaine $E(p_1, \ldots, p_T)$ associé. On retrouve le problème lié à l'utilisation de la fonction d'autocovariance. En effet, bien que ces coefficients soient les premiers coefficients d'une fonction d'autocovariance, l'existence du modèle n'est pas toujours assurée. C'est pourquoi nous préférons retenir les méthodes basées sur les autocorrélations partielles. Celles de Boshnakov [BOS94] et de Sakai [SAK82] sont deux extensions différentes de la méthode de Burg. Nous proposons d'étendre la méthode des autocorrélations partielles empiriques ([DÉG86], [DÉG93]) ($ACPE$) et celle des rapports d'énergies résiduelles de Dickinson [DIC78] (RER) au cas périodique. Les deux extensions de Burg et $ACPE$ sont regroupées dans une même méthodologie permettant de les comparer sur le plan de la conception. En particulier, les contraintes dans la construction récursive des filtres pour les méthodes de type Burg constituent a priori un handicap. Par ailleurs, les singularités d'ordre fini sont estimées presque sûrement par $ACPE$ ou RER, ce qui n'est plus vrai pour les autres méthodes. Cependant dans ce cas, les filtres traduisant les singularités ne sont généralement pas cohérents avec $\hat{\beta}(\cdot, \cdot)$, dont l'estimation devrait alors être remise en cause. Le dernier point est que, contrairement au cas stationnaire, $ACPE$ est la seule méthode qui soit, en toute généralité, récursive sur l'ordre du modèle ARP.

Nous considérons ensuite les approches vectorielles. Ce sont également des extensions des méthodes du cas stationnaire scalaire. Cette généralisation est un point délicat. Comme dans le cadre théorique, la difficulté réside dans le choix de la racine carrée pour normaliser l'autocovariance partielle. Ainsi on trouve plusieurs extensions de la méthode de Burg, selon le choix du procédé de normalisation. Celle de Nuttall [NUT76] (cf. aussi [STR77]) est indépendante de ce choix alors que celle de Morf et coll. [MOR78b] utilise une normalisation de type triangulaire inférieure dans les deux sens progressif et rétrograde. Le même procédé est utilisé par Dickinson [DIC79] pour étendre sa méthode. Dégerine ([DÉG92], [DÉG94]) considère plusieurs extensions des autocorrélations partielles empiriques et donne un cadre général qui regroupe la plupart des méthodes d'estimation utilisant les matrices d'autocorrélation partielle. Nous considérons le rapport entre les approches scalaire et vectorielle. La différence fondamentale est que les méthodes vectorielles n'estiment qu'une sous-classe de modèles ARP. Ce qui est en fait lié à la sur-paramétrisation introduite dans les modèles VAR. On se restreint

donc à cette sous-classe afin d'établir la correspondance entre approches vectorielle et scalaire. Par ailleurs, à chaque méthode du cas périodique considérée précédemment, correspond sur le plan de la conception une méthode vectorielle. On trouve ainsi les homologues de la méthode de Yule-Walker, de *RER* ([DIC79]), et d'*ACPE* ([DÉG92], [PHA92], [DÉG94]). Pour les techniques de Burg, la méthode de Nuttall [NUT76] correspond à celle de Boshnakov [BOS94], et celle de Morf et coll. [MOR78b] à celle de Sakai [SAK82]. Puisque ces méthodes vectorielles sont comparables dans leur conception à celles périodiques, il paraît naturel d'étudier si elles leur sont équivalentes. C'est-à-dire si la structure périodique résultante est la même. Notons que pour les méthodes qui dépendent du procédé de normalisation, on choisira celui qui a permis de définir $\Pi(\cdot)$ au troisième chapitre. Cette équivalence est alors satisfaite pour les méthodes de Yule-Walker, de *RER* et d'*ACPE*, mais pas pour les procédures de Burg.

Les différentes méthodes d'estimation sont ensuite analysées par simulation. Les facteurs déterminants sont essentiellement la proximité du modèle à la singularité, la longueur de la séquence et le type de paramètres étudiés. Toutes les méthodes sont équivalentes lorsque la série observée est longue et le modèle régulier. Comme dans le cas stationnaire, *ACPE* et *RER* conduisent à des résultats similaires. Ces méthodes sont celles qui présentent les meilleurs performances. Pour la méthode de Yule-Walker, on retrouve le défaut classique dû au biais particulièrement sensible sur séries courtes. Ses résultats sont encore plus médiocres lorsque le modèle s'approche de la singularité, y compris pour des séries longues. Pour les extensions de type Burg scalaires, on observe essentiellement les mêmes phénomènes que pour la situation stationnaire. On retrouve leur mauvais comportement pour des séries courtes, voire de longueur moyenne, issues d'un modèle proche de la singularité. Ces méthodes sont également sensibles à l'ordre de grandeur des coefficients des filtres du modèle, même si ce dernier est régulier. Ces mauvais résultats semblent être dûs aux contraintes excessives dans la construction récursive des filtres. En effet, la comparaison de ces méthodes avec celles vectorielles montrent que les procédures vectorielles, introduisant moins de contraintes sur les filtres, améliorent sensiblement les résultats. Cependant leurs performances restent bien inférieures à celles obtenues avec *ACPE* ou *RER*. Notons que toutes les méthodes sont équivalentes, sur les exemples considérés, dans l'estimation des autocorrélations. Les différences portent sur les autocorrélations partielles et surtout sur les paramètres du modèle *ARP* constitués des filtres et des variances résiduelles. Les résultats sur l'estimation autorégressive dans le cadre des processus périodiquement corrélés ont donné lieu à un rapport interne [LAM97a] ainsi qu'à une présentation au seizième Colloque GRETSI sur le Traitement du Signal et des Images [LAM97b].

Chapitre 1

Aspect temporel

Dans ce chapitre, $X(\cdot) = \{X(t), t \in \mathbb{Z}\}$ désigne un processus (non stationnaire) scalaire centré à valeurs complexes, défini sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, et qui admet des moments d'ordre deux. Seule l'étude de la structure au second ordre de $X(\cdot)$ étant considérée, il est naturel d'utiliser une approche géométrique. C'est-à-dire qu'on se place dans l'espace de Hilbert complexe engendré par les composantes du processus, $\mathcal{M} = \overline{\mathcal{L}}\{X(t), t \in \mathbb{Z}\}$. Les éléments de $\mathcal{M}$ sont des variables aléatoires complexes centrées et le produit hermitien est défini par la covariance $\langle U, V \rangle = E\left\{U\overline{V}\right\}$. Le carré de la norme représente alors la variance, $\|U\|^2 = Var(U)$, et l'orthogonalité correspond à la non corrélation. La structure au second ordre de $X(\cdot)$ est donc caractérisée par sa fonction d'autocovariance $R(\cdot,\cdot) = \{R(t,s) = \langle X(t), X(s)\rangle, (t,s) \in \mathbb{Z}^2\}$. L'objet de ce chapitre est de montrer qu'elle l'est également par la fonction d'autocorrélation partielle $\beta(\cdot,\cdot) = \{\beta(t,s), (t,s) \in \mathbb{Z}^2\}$. Rappelons que $\beta(t,s)$ est la corrélation partielle entre $X(t)$ et $X(s)$ dans l'ensemble $\{X(s), \ldots, X(t)\}$.

Même dans le cas stationnaire, la paramétrisation par les autocorrélations partielles est relativement récente. Ramsey [Ram74], après avoir défini les coefficients $\beta(t,s)$ pour un processus quelconque, se restreint au cas stationnaire pour préciser le domaine de variation de la fonction d'autocorrélation partielle et établir sa correspondance avec la fonction d'autocovariance. Ce résultat est également observé par Burg en 1975 [Bur75], dans le cadre du traitement du signal, où $\Leftrightarrow\!\!\!\!\!\beta(k)$ est appelé coefficient de réflexion. En fait cette équivalence est un résultat classique en théorie des polynômes orthogonaux (cf. [Ger60]). On peut aussi reconnaître l'usage des coefficients $\beta(t,s)$ dans des articles du traitement du signal : [Des90], [Lev81] ou [Kai82]. Nous constatons que les quantités $\beta(t,s)$ sont les éléments clés de la généralisation des algorithmes de Schur et Levinson-Durbin [Lev81] aux processus non stationnaires dans le cas non localement déterminable. Cependant, il

semble que la correspondance bijective entre les fonctions $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$ soit méconnue.

Dans la première section, nous donnons quelques rappels sur la notion de corrélation partielle. Cette approche purement géométrique permet d'établir de façon simple les propriétés de la fonction $\beta(\cdot,\cdot)$ introduite à la deuxième section. Dans cette dernière, après avoir décrit le domaine de variation $\mathcal{D}_\beta$ de cette fonction, nous présentons un procédé qui permet de construire une séquence d'un processus de fonction d'autocorrélation partielle connue. Cette procédure prouve en particulier l'injectivité dans l'ensemble $\mathcal{D}_\beta$ de l'application qui à $R(\cdot,\cdot)$ associe $\beta(\cdot,\cdot)$. Dans la troisième section, l'extension de l'algorithme de Levinson-Durbin au cas non stationnaire général permet d'obtenir la surjectivité de cette application. Enfin dans la dernière section, nous illustrons cette approche en considérant le cas des processus autorégressifs.

1.1 Corrélation partielle

On peut reconnaître le coefficient de corrélation partielle dans l'article de Yule [YUL07]. Il s'agit maintenant d'une notion classique en statistique, tout du moins pour des variables aléatoires scalaires réelles, et pour une utilisation standard. Il existe en effet des développements sur ce type de coefficients liés à la trigonométrie sphérique (cf. [DES93]), comme l'identité de Yule présentée ci-après, souvent méconnus dans la littérature statistique, mais très répandus en traitement du signal. La corrélation partielle joue en effet un rôle central dans les algorithmes récursifs rapides pour la résolution de problèmes de moindres carrés. C'est ainsi que le concept a été étendu à des vecteurs aléatoires (cf. [MOR78a]).

Nous nous plaçons ici dans le cadre de variables aléatoires scalaires complexes en ne retenant que les propriétés qui nous seront utiles pour la suite. Soient W un élément de $\mathcal{M}$ et $\mathcal{N}$ un sous-espace vectoriel fermé de $\mathcal{M}$, nous notons $W/\mathcal{N}$ la projection orthogonale (au sens de $\langle\cdot,\cdot\rangle$) de W sur $\mathcal{N}$, $\varepsilon(W/\mathcal{N}) = W \Leftrightarrow W/\mathcal{N}$, l'erreur de prédiction linéaire associée et $\eta(W/\mathcal{N}) = \varepsilon(W/\mathcal{N})/\left\|\varepsilon(W/\mathcal{N})\right\|$, le vecteur unitaire correspondant, lorsque $\|\varepsilon(W/\mathcal{N})\| \neq 0$, sinon on convient de poser $\eta(W/\mathcal{N}) = \varepsilon(W/\mathcal{N}) = \vec{0}$.

Définition 1.1.1. *La corrélation partielle entre deux éléments U et V de $\mathcal{M}$ sachant un sous-espace fermé $\mathcal{N}$, notée $\rho(U,V/\mathcal{N})$, est définie par :*

$$\rho(U,V/\mathcal{N}) = \langle \eta(U/\mathcal{N}), \eta(V/\mathcal{N}) \rangle .$$

On dit aussi que $\rho(U,V/\mathcal{N})$ est la corrélation partielle entre U et V dans l'ensemble $\{\mathcal{N},U,V\}$. Notons que $\rho(U,V/\mathcal{N}) = \overline{\rho(V,U/\mathcal{N})}$. Cette corrélation

partielle mesure la qualité de la liaison linéaire entre U et V qui ne résulte pas de liaisons (linéaires) éventuelles entre U ou V d'une part et les éléments de $\mathcal{N}$ d'autre part. D'un point de vue géométrique $\rho(U, V/\mathcal{N})$ est le "cosinus d'angle" entre les deux erreurs de prédiction linéaire de U et V par $\mathcal{N}$.

On note $\{W\}$ et $\{\mathcal{N}, W\}$ les sous-espaces de $\mathcal{M}$ engendrés respectivement par W et $\{\mathcal{N}, W\}$. Le signe $\oplus$ représente une somme orthogonale, $\rho(U, V) = \rho(U, V/\{\vec{0}\})$ la corrélation habituelle, et $|\cdot|$ le module.

Proposition 1.1.1. *La corrélation partielle satisfait les propriétés suivantes :*

a) $|\rho(U, V/\mathcal{N})| \leq 1$,

b) $|\rho(U, V/\mathcal{N})| = 1 \Leftrightarrow \eta(U/\mathcal{N}) = \rho(U, V/\mathcal{N})\eta(V/\mathcal{N})$ *et* $\rho(U, V/\mathcal{N}) \neq 0$,

c) $\varepsilon(U/\mathcal{N}) = \|\varepsilon(U/\mathcal{N})\| \, \rho(U, V/\mathcal{N})\eta(V/\mathcal{N}) \oplus \varepsilon(U/\{\mathcal{N}, V\})$,

d) $\|\varepsilon(U/\{\mathcal{N}, V\})\|^2 = \left[1 \Leftrightarrow |\rho(U, V/\mathcal{N})|^2\right] \|\varepsilon(U/\mathcal{N})\|^2$.

Les propriétés a) et b) sont des conséquences immédiates de la définition compte tenu de la convention qui conduit à $\rho(U, \vec{0}/\mathcal{N}) = 0$. La propriété c) est fondamentale ; elle est obtenue en décomposant la projection $U/\{\mathcal{N}, V\}$ sur la somme orthogonale $\{\mathcal{N}, V\} = \mathcal{N} \oplus \{\eta(V/\mathcal{N})\}$. La propriété d), qui découle de c), montre que $|\rho(U, V/\mathcal{N})|^2$ mesure le gain relatif sur la variance de l'erreur de prédiction de U par l'adjonction de V à $\mathcal{N}$.

Il n'existe pas de relations d'ordre entre les autocorrélations partielles obtenues en faisant varier le sous-espace $\mathcal{N}$. Par exemple il est clair que $|\rho(U, V)| = 1$ implique $|\rho(U, V/\mathcal{N})| = 1$ si $U \notin \mathcal{N}$ (cf. Propriété b) ci-dessus), mais on peut avoir $|\rho(U, V/\mathcal{N})| = 1$ et $|\rho(U, V)| = 0$. De même, dans le cas réel, le signe de $\rho(U, V/\mathcal{N})$ peut varier avec $\mathcal{N}$ et en particulier être différent de celui de $\rho(U, V)$. Il est facile de construire des exemples illustrant ces points en remarquant qu'un vecteur W peut se définir par la somme des deux composantes, $W = W/\mathcal{N} \oplus \varepsilon(W/\mathcal{N})$, avec un choix arbitraire de ces composantes.

Citons cependant, sans la démontrer car nous ne l'utiliserons pas ici, l'identité de Yule [Yul07],

$$\rho(U, V/\{\mathcal{N}, W\}) = \frac{\rho(U, V/\mathcal{N}) \Leftrightarrow \rho(U, W/\mathcal{N})\overline{\rho(V, W/\mathcal{N})}}{\left[1 \Leftrightarrow |\rho(U, W/\mathcal{N})|^2\right]^{\frac{1}{2}} \left[1 \Leftrightarrow |\rho(V, W/\mathcal{N})|^2\right]^{\frac{1}{2}}}.$$

La Propriété fondamentale c) permet d'énoncer la proposition suivante.

Proposition 1.1.2. *Deux éléments U et V de $\mathcal{M}$ et un sous-espace fermé $\mathcal{N}$ satisfont, avec la convention $0^{-1/2} = 0$,*

$$\eta(U/\{\mathcal{N}, V\}) = \left[1 \Leftrightarrow |\rho(U, V/\mathcal{N})|^2\right]^{-\frac{1}{2}} \{\eta(U/\mathcal{N}) \Leftrightarrow \rho(U, V/\mathcal{N})\eta(V/\mathcal{N})\},$$

$$\eta(V/\{\mathcal{N},U\}) = \left[1 \Leftrightarrow |\rho(V,U/\mathcal{N})|^2\right]^{-\frac{1}{2}} \{\eta(V/\mathcal{N}) \Leftrightarrow \rho(V,U/\mathcal{N})\eta(U/\mathcal{N})\}\,.$$

Les relations de la Proposition 1.1.2 définissent une rotation hyperbolique dans le plan (cf. Figure 1.1 pour le cas réel) qui associe aux deux vecteurs unitaires, $\eta(U/\mathcal{N})$ et $\eta(V/\mathcal{N})$, deux autres vecteurs $\eta(U/\{\mathcal{N},V\})$ et $\eta(V/\{\mathcal{N},U\})$, également unitaires, selon la matrice de passage :

$$\frac{1}{\sqrt{1 \Leftrightarrow |\rho|^2}} \begin{bmatrix} 1 & \Leftrightarrow\rho \\ \Leftrightarrow\overline{\rho} & 1 \end{bmatrix}, \qquad \rho = \rho(U,V/\mathcal{N}).$$

D'après la proposition ci-dessus,

$$\langle \eta(U/\{\mathcal{N},V\}), \eta(V/\{\mathcal{N},U\})\rangle = \Leftrightarrow\rho(U,V/\mathcal{N}) = \Leftrightarrow\langle \eta(U/\mathcal{N}), \eta(V/\mathcal{N})\rangle\,.$$

Cette relation apparaît clairement sur la Figure 1.1.

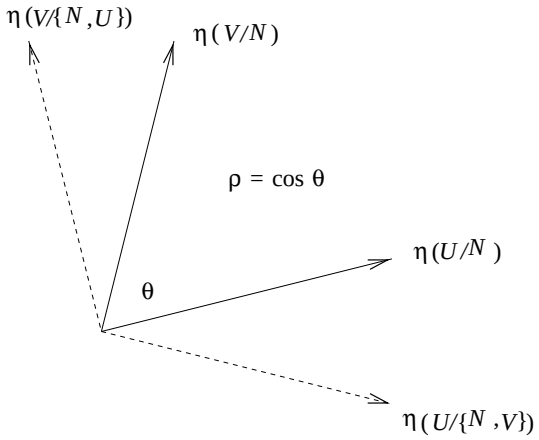


Figure 1.1 : *Rotation hyperbolique*

1.2 Fonction d'autocorrélation partielle

Dans la première sous-section, nous introduisons les innovations partielles progressives et rétrogrades permettant de définir la fonction d'autocorrélation partielle. Le domaine de variation de cette fonction est décrit à la deuxième sous-section. Dans la dernière sous-section, nous rappelons la décomposition de Wold-Cramér. A cette occasion, nous précisons une classification des différents types de processus utilisée dans la suite.

1.2.1 Définitions et notations

La structure au second ordre du processus $X(\cdot)$ est usuellement paramétrisée par la fonction d'autocovariance,

$$R(t,s) = E\left\{X(t)\overline{X(s)}\right\} = \langle X(t), X(s)\rangle\,, \quad (t,s) \in \mathbb{Z}^2.$$

Cette fonction satisfait la symétrie hermitienne, $R(s,t) = \overline{R(t,s)}$. Elle est donc caractérisée par sa restriction aux couples (t,s) satisfaisant $s \leq t$. Par la suite nous veillerons à conserver l'ordre $s \leq t$ dans l'utilisation de ces variables afin de faciliter la reconnaissance des aspects progressifs et rétrogrades. La fonction $R(\cdot,\cdot)$ est de type positif, c'est-à-dire que pour $s \leq t$ la matrice $R_{s,t} = \{R(s+i, s+j)\}_{i,j=0,\ldots,t-s}$ est définie non négative (comme matrice de covariance du vecteur aléatoire $[X(s), \ldots, X(t)]^T$). La fonction d'autocorrélation partielle que nous présentons ci-dessous caractérise également la structure au second ordre de $X(\cdot)$, mais son domaine de variation se décrit de façon beaucoup plus simple. Pour $t \in \mathbb{Z}$, introduisons les sous-espaces de $\mathcal{M}$ suivants :

$$\begin{aligned}
\mathcal{M}(s;t) &= \mathcal{L}\left\{X(u), s \leq u \leq t\right\}, \quad s \leq t,\\
\mathcal{M}(\Leftrightarrow\infty;t) &= \overline{\mathcal{L}}\left\{X(u), u \leq t\right\},\\
\mathcal{M}(\Leftrightarrow\infty) &= \bigcap_t \mathcal{M}(\Leftrightarrow\infty;t).
\end{aligned}$$

Soient $X^f(t;s)$ et $X^b(s;t)$ les projections orthogonales de $X(t)$ et $X(s)$ respectivement sur $\mathcal{M}(s;t \Leftrightarrow 1)$ et $\mathcal{M}(s+1;t)$ avec la convention $X^f(t;t) = X^b(t;t) = \vec{0}$. Alors $\varepsilon^f(t;s) = X(t) \Leftrightarrow X^f(t;s)$ et $\varepsilon^b(s;t) = X(s) \Leftrightarrow X^b(s;t)$ sont les innovations partielles d'ordre $t \Leftrightarrow s \geq 0$ respectivement progressive et rétrograde. Notons que dans les notations du type $\varepsilon^f(t;s)$ ou $\varepsilon^b(s;t)$, la première variable indique l'instant t ou s en lequel est considérée l'erreur alors que la seconde représente l'instant limite s ou t sur lequel porte la régression. Introduisons les innovations normées,

$$\eta^f(t;s) = \frac{\varepsilon^f(t;s)}{\sigma^f(t;s)}, \quad \sigma^{f2}(t;s) = \left\|\varepsilon^f(t;s)\right\|^2 = Var\left\{\varepsilon^f(t;s)\right\},$$

$$\eta^b(s;t) = \frac{\varepsilon^b(s;t)}{\sigma^b(s;t)}, \quad \sigma^{b2}(s;t) = \left\|\varepsilon^b(s;t)\right\|^2 = Var\left\{\varepsilon^b(s;t)\right\},$$

avec la convention $\eta^f(t;s) = \vec{0}$ lorsque $\varepsilon^f(t;s) = \vec{0}$ et son homologue dans le sens rétrograde.

La fonction d'autocorrélation partielle décrit, pour (t,s) parcourant $\mathbb{Z}^2$, la corrélation partielle entre $X(t)$ et $X(s)$ dans l'ensemble $\{X(s), \ldots, X(t)\}$. De façon plus précise, avec les notations ci-dessus, on pose la définition suivante.

Définition 1.2.1. *La fonction d'autocorrélation partielle du processus* $X(\cdot)$, *notée* $\beta(\cdot,\cdot)$, *est définie pour* $(t,s) \in \mathbb{Z}^2$ *par* :

$$\beta(t,s) = \begin{cases} \left\langle \eta^f(t; s+1), \eta^b(s; t \Leftrightarrow 1) \right\rangle & si\ s < t \\ \|X(t)\|^2 & si\ s = t. \\ \left\langle \eta^b(t; s \Leftrightarrow 1), \eta^f(s; t+1) \right\rangle & si\ s > t \end{cases}$$

1.2.2 Domaine de variation de $\beta(\cdot,\cdot)$

On a convenu de poser $\beta(t,t) = Var\{X(t)\}$ au lieu de 1 afin que la fonction $\beta(\cdot,\cdot)$ caractérise à elle seule la structure au second ordre de $X(\cdot)$. Le module de cette fonction, pour $s \neq t$, est inférieur ou égal à 1, l'égalité traduisant les singularités d'ordre fini. En effet, pour $s < t$, $|\beta(t,s)| = 1$ si et seulement si s est le plus grand entier tel que $X(t)$ appartienne à l'ensemble $\mathcal{L}\{X(s), \ldots, X(t \Leftrightarrow 1)\}$. La corrélation partielle est alors par convention égale à zéro lorsqu'elle n'est pas définie, c'est-à-dire aux points $(t, s \Leftrightarrow k)$ et $(t+k, s)$, $k \geq 1$. De même, si une variable $X(t)$ est nulle presque sûrement, on a $\beta(t, t \Leftrightarrow k) = \beta(t+k, t) = 0$ pour $k \geq 0$. Plus précisément, la fonction $\beta(\cdot,\cdot)$ vérifiant la symétrie hermitienne, $\beta(s,t) = \overline{\beta(t,s)}$, est donc caractérisée par sa restriction à l'ensemble $\left\{(t,s) \in \mathbb{Z}^2, s \leq t\right\}$ où elle satisfait les contraintes :

(i) $\beta(t,t) \geq 0$ et

$$\beta(t,t) = 0 \Rightarrow \beta(t,s) = 0, \quad s < t,$$

(ii) $|\beta(t,s)| \leq 1$, $s < t$ et

$$|\beta(t,s)| = 1 \Rightarrow \beta(t+k,s) = \beta(t, s \Leftrightarrow k) = 0, \quad k \geq 1.$$

Nous notons $\mathcal{D}_\beta$ l'ensemble des fonctions définies sur $\mathbb{Z}^2$ qui satisfont les contraintes ci-dessus. Nous verrons en effet qu'il s'agit du domaine de variation de $\beta(\cdot,\cdot)$. La Figure 1.2 illustre la simplicité de ce domaine. Utilisant la symétrie hermitienne, on considère le demi-plan de $\mathbb{Z}^2$ défini par l'ensemble $\left\{(t,s) \in \mathbb{Z}^2, s \leq t\right\}$. Chaque symbole représente les différents types de valeurs prises par $\beta(\cdot,\cdot)$ au point de $\mathbb{Z}^2$ correspondant. Ainsi dans le cas où le processus ne présente pas de singularités d'ordre fini, la fonction $\beta(\cdot,\cdot)$ est réelle strictement positive sur la diagonale $t = s$, et son module est strictement inférieur à 1 hors de cette diagonale. Dans le cas contraire, les bandes de zéros horizontales et verticales proviennent de la convention $0^{-1} = 0$.

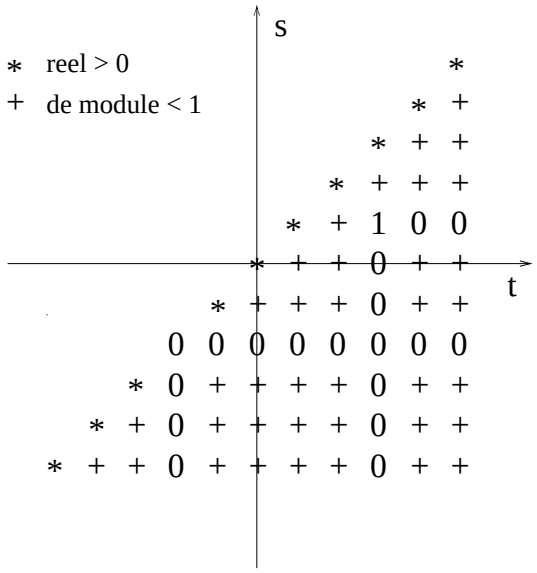


Figure 1.2 : *Illustration de* $\mathcal{D}_\beta$

La double récurrence satisfaite par les innovations normées ci-après est fondamentale. En particulier elle est à la base du procédé de construction donné par la Proposition 1.2.2.

Proposition 1.2.1. *Les innovations partielles normées satisfont, pour* $s < t$*, les relations* ($0^{-1/2} = 0$ *par convention*)

$$\eta^f(t;s) = \left[1 \Leftrightarrow |\beta(t,s)|^2\right]^{-\frac{1}{2}} \left\{\eta^f(t;s+1) \Leftrightarrow \beta(t,s)\eta^b(s;t \Leftrightarrow 1)\right\},$$

$$\eta^b(s;t) = \left[1 \Leftrightarrow |\beta(s,t)|^2\right]^{-\frac{1}{2}} \left\{\eta^b(s;t \Leftrightarrow 1) \Leftrightarrow \beta(s,t)\eta^f(t;s+1)\right\},$$

et les variances résiduelles sont données par

$$\sigma^{f2}(t;s) = \beta(t,t) \prod_{j=1}^{t-s} \left[1 \Leftrightarrow |\beta(t,t \Leftrightarrow j)|^2\right], \tag{1.1}$$

$$\sigma^{b2}(s;t) = \beta(s,s) \prod_{j=1}^{t-s} \left[1 \Leftrightarrow |\beta(s,s+j)|^2\right]. \tag{1.2}$$

Notons que ces relations, ainsi que les expressions des variances $\sigma^{f2}(t;s)$ et $\sigma^{b2}(s;t)$, apparaissent dans [Lev81] et [Kai82]. La double récurrence résulte de la Proposition 1.1.2 appliquée à $U = X(t)$, $V = X(s)$ et $\mathcal{N} =$

$\{X(s+1),\ldots,X(t\Leftrightarrow 1)\}$. L'expression (1.1) découle de l'itération de la relation $\sigma^{f2}(t;s) = \sigma^{f2}(t;s+1)\left[1\Leftrightarrow|\beta(t,s)|^2\right]$ donnée par le point d) de la Proposition 1.1.1. Il en est de même pour $\sigma^{b2}(s;t)$.

Remarquons que $\sigma^{f2}(t;s)$ ne dépend que des autocorrélations partielles entre $X(t)$ d'une part et les variables $X(u)$, pour $s \leq u < t$, d'autre part. La Proposition 1.2.3 de la sous-section qui suit montre qu'il en est de même des coefficients de l'expression de $\varepsilon^f(t;s)$ en fonction d'une base orthonormée naturelle du passé de $X(t)$ et plus généralement de la description même de $X(t)$ sur son passé.

Nous montrons ci-dessous que le domaine de variation de $\beta(\cdot,\cdot)$ est l'ensemble $\mathcal{D}_\beta$ décrit plus haut.

Proposition 1.2.2. *Soient $\{Z(t), t \in \mathbb{Z}\}$ une suite orthonormale et $\beta(\cdot,\cdot)$ un élément de $\mathcal{D}_\beta$. Alors, pour tout s de $\mathbb{Z}$, la séquence $\{X(t), t \geq s\}$ définie pour $t = s, s+1, \ldots$, par :*

$$\eta^f(t;s) = Z(t) \text{ si } \sigma^f(t;s) > 0,\ \vec{0} \text{ sinon,}$$

pour $k = s,\ldots,t\Leftrightarrow 1$:

$$\begin{aligned}
\eta^f(t;k+1) &= \left[1\Leftrightarrow|\beta(t,k)|^2\right]^{\frac{1}{2}}\eta^f(t;k) + \beta(t,k)\eta^b(k;t\Leftrightarrow 1),\\
\eta^b(k;t) &= \left[1\Leftrightarrow|\beta(k,t)|^2\right]^{-\frac{1}{2}}\left\{\eta^b(k;t\Leftrightarrow 1)\Leftrightarrow\beta(k,t)\eta^f(t;k+1)\right\},\\
X(t) &= [\beta(t,t)]^{\frac{1}{2}}\eta^f(t;t), \quad \eta^b(t;t) = \eta^f(t;t),
\end{aligned}$$

admet une fonction d'autocorrélation partielle qui coïncide avec $\beta(\cdot,\cdot)$ sur le domaine $\left\{(u,t)\in\mathbb{Z}^2, u,t = s, s+1,\ldots\right\}$.

Démonstration.- On suppose avoir construit la séquence $\{X(u), u = s,\ldots, t\Leftrightarrow 1\}$ en fonction des variables $Z(u)$, $u = s,\ldots,t\Leftrightarrow 1$, de telle sorte que ses autocorrélations partielles coïncident avec la restriction de $\beta(\cdot,\cdot)$ à $\{(u,v)\in\mathbb{Z}^2, u,v = s,\ldots,t\Leftrightarrow 1\}$. De plus on dispose de la base $\{\eta^b(k;t\Leftrightarrow 1), k = s,\ldots,t\Leftrightarrow 1\}$ de l'espace $\mathcal{M}(s;t\Leftrightarrow 1)$ correspondant. Notons que ces hypothèses n'engagent à rien pour $t = s$, mais qu'elles sont aussi vérifiées pour $t = s+1$ après avoir effectué le premier pas de récurrence :

$$\eta^f(s;s) = \eta^b(s;s) = Z(s); \qquad X(s) = \beta(s,s)^{\frac{1}{2}}\eta^f(s;s).$$

L'algorithme, à travers la première relation de récurrence, définit la variable $X(t)$ sous la forme,

$$X(t) = \sum_{k=s}^{t-1}\sigma^f(t;k+1)\beta(t,k)\eta^b(k;t\Leftrightarrow 1) + \sigma^f(t;s)\eta^f(t;s),$$

où $\eta^f(t;s) = Z(t)$ est orthogonal à $\mathcal{M}(s;t\Leftrightarrow 1)$. Il s'en suit que cette définition équivaut à $X(t) = X^f(t;s) + \sigma^f(t;s)\eta^f(t;s)$ où les notations sont en accord

avec leur signification habituelle. En particulier $\sigma^f(t;s)$ est la norme de $\varepsilon^f(t;s)$. Compte tenu du choix de la base de $\mathcal{M}(s;t \Leftrightarrow 1)$, on a

$$\begin{aligned}
\varepsilon^f(t;s+1) &= X(t) \Leftrightarrow X^f(t;s+1) \\
&= \sigma^f(t;s+1)\beta(t,s)\eta^b(s;t \Leftrightarrow 1) + \sigma^f(t;s)\eta^f(t;s), \\
\left\|\varepsilon^f(t;s+1)\right\|^2 &= \sigma^{f2}(t;s+1)\left|\beta(t,s)\right|^2 + \sigma^{f2}(t;s) = \sigma^{f2}(t;s+1),
\end{aligned}$$

$$\left\langle \varepsilon^f(t;s+1), \eta^b(s;t \Leftrightarrow 1)\right\rangle = \sigma^f(t;s+1)\beta(t,s).$$

Ainsi $\sigma^{f2}(t;s+1)$ est la norme de $\varepsilon(t;s+1)$ et $\beta(t,s)$ est la corrélation partielle entre $X(t)$ et $X(s)$. Cette récurrence se poursuit pour $j = s+2, \ldots, t$, avec

$$\begin{aligned}
\varepsilon^f(t;j) &= \sum_{k=s}^{j-1} \sigma^f(t;k+1)\beta(t,k)\eta^b(k;t \Leftrightarrow 1) + \sigma^f(t;s)\eta^f(t;s), \\
\left\|\varepsilon^f(t;j)\right\|^2 &= \sum_{k=s}^{j-1} \sigma^{f2}(t;k+1)\left|\beta(t,k)\right|^2 + \sigma^{f2}(t;s) \\
&= \sum_{k=s+1}^{j-1} \sigma^{f2}(t;k+1)\left|\beta(t,k)\right|^2 + \sigma^{f2}(t;s+1) = \ldots = \sigma^{f2}(t;j),
\end{aligned}$$

$$\left\langle \varepsilon^f(t;j), \eta^b(j \Leftrightarrow 1; t \Leftrightarrow 1)\right\rangle = \sigma^f(t;j)\beta(t, j \Leftrightarrow 1).$$

Ainsi $\sigma^{f2}(t;j)$ est la norme de $\varepsilon(t;j)$ et $\beta(t, j \Leftrightarrow 1)$ est la corrélation partielle entre $X(t)$ et $X(j \Leftrightarrow 1)$. Enfin

$$\|X(t)\|^2 = \sum_{k=s}^{t-1} \sigma^{f2}(t;k+1)\left|\beta(t,k)\right|^2 + \sigma^{f2}(t;s) = \beta(t,t).$$

La première partie de l'hypothèse de récurrence à l'étape t est acquise. Par conséquent la deuxième relation de récurrence de l'algorithme construit effectivement la nouvelle base $\left\{\eta^b(k;t), k = s, \ \ldots, t\right\}$ de l'espace $\mathcal{M}(s;t)$.$\diamondsuit$

Cet algorithme permet de simuler facilement une séquence de structure donnée avec un coût de stockage réduit. En effet la variable $X(t)$ est obtenue en fonction du passé résumé par $\eta^b(k;t \Leftrightarrow 1)$, $k = s, \ldots, t \Leftrightarrow 1$, de l'innovation apportée par $Z(t)$ et des valeurs $\beta(t,k)$, $k = s, \ldots, t$. Il est clair que l'on peut procéder dans le sens inverse du temps. Par ailleurs, partant de $t = 0$, une séquence bilatérale $\{X(t), t \in \mathbb{Z}\!\!\!\mathbb{Z}\}$, dans laquelle $X(\Leftrightarrow t)$ et $X(t)$ sont générés alternativement, peut être associée de façon analogue à un élément $\beta(\cdot,\cdot)$ de $\mathcal{D}_\beta$. A l'étape t, on dispose de $X(\Leftrightarrow t+1), \ldots, X(t \Leftrightarrow 1)$, c'est-à-dire de $\eta^b(k;t \Leftrightarrow 1)$ et $\eta^f(\Leftrightarrow k; \Leftrightarrow t+1)$ pour $k = \Leftrightarrow t+1, \ldots, t \Leftrightarrow 1$. Les nouvelles

variables $X(t)$ et $X(\Leftrightarrow t)$ sont obtenues de la manière suivante :

$$\begin{aligned} X(t) &= X^f(t;\Leftrightarrow t+1) + \varepsilon^f(t;\Leftrightarrow t+1) \\ &= \sum_{k=-t+1}^{t-1} \sigma^f(t;k+1)\beta(t,k)\eta^b(k;t\Leftrightarrow 1) + \sigma^f(t;\Leftrightarrow t+1)Z(t), \\ X(\Leftrightarrow t) &= X^b(\Leftrightarrow t;t) + \varepsilon^b(\Leftrightarrow t;t) \\ &= \sum_{k=-t+1}^{t} \sigma^b(\Leftrightarrow t;k\Leftrightarrow 1)\beta(\Leftrightarrow t,k)\eta^f(k;\Leftrightarrow t+1) + \sigma^b(\Leftrightarrow t;t)Z(\Leftrightarrow t), \end{aligned}$$

où $\sigma^{f2}(t;k)$ et $\sigma^{b2}(\Leftrightarrow t;k)$ sont déterminées selon les relations (1.1) et (1.2) de la Proposition 1.2.1.

Ainsi partant d'un élément $\beta(\cdot,\cdot)$ de $\mathcal{D}_\beta$, il est possible de construire une suite non stationnaire $\{X(t), t \in \mathbb{Z}\!\!\mathbb{Z}\}$ qui admette $\beta(\cdot,\cdot)$ comme fonction d'autocorrélation partielle. Ceci prouve la surjectivité dans l'ensemble $\mathcal{D}_\beta$ de l'application qui à $R(\cdot,\cdot)$ associe $\beta(\cdot,\cdot)$. Les relations de l'algorithme de Levinson-Durbin Généralisé (LDG), décrites dans la section suivante, sont nécessaires pour assurer l'injectivité.

1.2.3 Décomposition de Wold-Cramér

Soit $X^f(t;\Leftrightarrow\infty)$ la projection orthogonale de $X(t)$ sur $\mathcal{M}(\Leftrightarrow\infty;t\Leftrightarrow 1)$, alors $\varepsilon(t) = X(t) \Leftrightarrow X^f(t;\Leftrightarrow\infty)$ est le processus d'innovation (progressive). Nous notons $\sigma_\varepsilon(t)$ sa norme et $\eta(t)$ l'innovation normée, avec pour convention $\eta(t) = \vec{0}$ si $\sigma_\varepsilon(t) = 0$. Un processus $X(\cdot)$ est dit déterminable lorsque $\mathcal{M}(\Leftrightarrow\infty) = \mathcal{M}$. Comme dans le cas stationnaire, ceci équivaut à $\mathcal{M}(\Leftrightarrow\infty;t) = \mathcal{M}(\Leftrightarrow\infty;t\Leftrightarrow 1)$ pour tout t, ou encore $\sigma_\varepsilon^2(t) = 0$ pour tout t. Les processus qui ne satisfont pas la contrainte $\mathcal{M}(\Leftrightarrow\infty) = \mathcal{M}$, sont dits indéterminables. Dans la littérature, on utilise également les termes "(linéairement) singuliers" et "non (linéairement) singuliers" pour désigner les processus respectivement déterminables et indéterminables. La proposition ci-après montre comment la classe des processus indéterminables est facilement caractérisée par la fonction d'autocorrélation partielle. L'orthogonalisation, par le procédé de Gram-Schmidt, des variables $X(t\Leftrightarrow 1)$, $X(t\Leftrightarrow 2)$, ... , engendrant le passé de $X(t)$, montre que les innovations partielles normées rétrogrades $\eta^b(t\Leftrightarrow 1\Leftrightarrow k;t\Leftrightarrow 1)$, $k = 0, 1, \ldots$, forment une base orthonormée de ce passé (les variables éventuellement nulles étant éliminées).

Proposition 1.2.3. *A chaque instant t fixé, la décomposition orthogonale*

de $X(t)$ sur son passé et son innovation est donnée par :

$$X(t) = \varepsilon(t) + \sum_{k=1}^{+\infty} \beta(t, t \Leftrightarrow k)\sigma^f(t; t \Leftrightarrow k+1)\eta^b(t \Leftrightarrow k; t \Leftrightarrow 1). \tag{1.3}$$

De plus la variance de $\varepsilon(t)$ satisfait,

$$\sigma_\varepsilon^2(t) = \beta(t,t) \prod_{k=1}^{+\infty} \left[1 \Leftrightarrow |\beta(t, t \Leftrightarrow k)|^2\right].$$

Démonstration.- On peut déduire de la Proposition 1.2.1, ou directement par projection de $X(t)$ sur la somme orthogonale,

$$\mathcal{M}(t \Leftrightarrow n; t \Leftrightarrow 1) = \mathcal{M}(t \Leftrightarrow n+1; t \Leftrightarrow 1) \oplus \mathcal{L}\left\{\eta^b(t \Leftrightarrow n; t \Leftrightarrow 1)\right\},$$

la relation

$$\varepsilon^f(t; t \Leftrightarrow n) = \varepsilon^f(t; t \Leftrightarrow n+1) \Leftrightarrow \beta(t, t \Leftrightarrow n)\sigma^f(t; t \Leftrightarrow n+1)\eta^b(t \Leftrightarrow n; t \Leftrightarrow 1).$$

Par itération, on obtient

$$\varepsilon^f(t; t \Leftrightarrow n) = X(t) \Leftrightarrow \sum_{k=1}^{n} \beta(t, t \Leftrightarrow k)\sigma^f(t; t \Leftrightarrow k+1)\eta^b(t \Leftrightarrow k; t \Leftrightarrow 1).$$

La projection de $X^f(t; \Leftrightarrow\infty)$ sur $\mathcal{M}(t \Leftrightarrow n; t \Leftrightarrow 1)$ est égale à $X^f(t; t \Leftrightarrow n)$. La limite en moyenne quadratique, $X^f(t; \Leftrightarrow\infty) = l.i.m. X(t; t \Leftrightarrow n)$, donne $\varepsilon^f(t) = l.i.m. \varepsilon^f(t; t \Leftrightarrow n)$ et justifie (1.3). Par suite

$$\sigma_\varepsilon^2(t) = \lim_{n \to +\infty} \sigma^{f2}(t; t \Leftrightarrow n) = \beta(t,t) \prod_{k=1}^{+\infty} \left[1 \Leftrightarrow |\beta(t, t \Leftrightarrow k)|^2\right]. \Diamond$$

On vérifie ainsi qu'à chaque instant t, le présent $X(t)$ se décrit en fonction du passé et de l'innovation $\eta(t)$ à travers les seules valeurs $\beta(t, t \Leftrightarrow k)$, $k \geq 0$, de la fonction d'autocorrélation partielle. En particulier $X(\cdot)$ est un processus indéterminable si et seulement si il existe $t \in \mathbb{Z}$ tel que

$$\beta(t,t) > 0; \qquad |\beta(t, t \Leftrightarrow k)| < 1, \quad k \geq 1; \qquad \sum_{k=1}^{+\infty} |\beta(t, t \Leftrightarrow k)|^2 < +\infty.$$

En effet les contraintes ci-dessus sont équivalentes à $\sigma_\varepsilon^2(t) > 0$.

Parmi les processus indéterminables, on distingue les processus pour lesquels $\mathcal{M}(\Leftrightarrow\infty) = \{\vec{0}\}$. Ces derniers sont dits purement indéterminables. On rencontre également le terme "régulier" pour les désigner. Par ailleurs, l'espace $\mathcal{M}(\Leftrightarrow\infty; t)$ se décompose selon (cf. [Cra61] : Lemme 1, p. 63),

$$\mathcal{M}(\Leftrightarrow\infty; t) = \mathcal{M}_\varepsilon(\Leftrightarrow\infty; t) \oplus \mathcal{M}(\Leftrightarrow\infty),$$

où $\mathcal{M}_\varepsilon(\Leftrightarrow\infty;t) = \overline{\mathcal{L}}\{\varepsilon(s), s \leq t\}$. Ainsi lorsque $X(\cdot)$ est purement indéterminable, on a pour tout t,

$$X(t) = \sum_{k=0}^{+\infty} b(t, t \Leftrightarrow k)\varepsilon(t \Leftrightarrow k), \quad h(t,t) = 1,$$

et

$$R(t,t) = \sum_{k=0}^{+\infty} |b(t, t \Leftrightarrow k)|^2 \, \sigma_\varepsilon^2(t \Leftrightarrow k) < +\infty.$$

Réciproquement soit $X(\cdot)$ un processus défini par la série ci-dessus en fonction d'une suite $\tilde{\varepsilon}(\cdot) = \{\tilde{\varepsilon}(t), t \in \mathbb{Z}\!\!\mathbb{Z}\}$ de variables aléatoires centrées non corrélées. Alors l'espace $\mathcal{M}(\Leftrightarrow\infty;t)$ est inclus dans $\mathcal{M}_{\tilde{\varepsilon}}(\Leftrightarrow\infty;t)$ avec égalité lorsque $\tilde{\varepsilon}(\cdot)$ est le processus d'innovation, et par suite,

$$\bigcap_t \mathcal{M}(\Leftrightarrow\infty;t) \subset \bigcap_t \mathcal{M}_{\tilde{\varepsilon}}(\Leftrightarrow\infty;t) = \left\{\vec{0}\right\}.$$

Ainsi le processus $X(\cdot)$ est purement indéterminable. La décomposition de Wold-Cramér a été établie par Wold [Wol54] dans le cas stationnaire scalaire et étendue par Cramér [Cra61] aux processus vectoriels non stationnaires. Elle indique que tout processus est la somme d'un processus purement indéterminable et d'un déterminable, ces deux processus étant non corrélés entre eux. Plus précisément, nous avons le théorème suivant.

Théorème 1.2.1. *Un processus $X(\cdot)$ non stationnaire se décompose de façon unique sous la forme suivante,*

$$X(t) = U(t) + V(t), \quad t \in \mathbb{Z}\!\!\mathbb{Z},$$

où $U(\cdot)$ et $V(\cdot)$ sont deux processus non corrélés entre eux tels que pour tout $t \in \mathbb{Z}\!\!\mathbb{Z}$, $U(t)$ et $V(t)$ appartiennent à $\mathcal{M}(\Leftrightarrow\infty;t)$. De plus $U(\cdot)$ est un processus purement indéterminable et $V(\cdot)$ est un processus déterminable. D'autre part, le processus $U(\cdot)$ admet la représentation,

$$U(t) = \sum_{k=0}^{+\infty} b(t, t \Leftrightarrow k)\varepsilon(t \Leftrightarrow k), \quad b(t,t) = 1, \quad t \in \mathbb{Z}\!\!\mathbb{Z},$$

où $\varepsilon(t)$ est le processus d'innovation et $\sum_{k=0}^{+\infty} |b(t, t \Leftrightarrow k)|^2 \, \sigma_\varepsilon^2(t \Leftrightarrow k) < +\infty$.

Nous ne donnons pas la démonstration de ce théorème qui est détaillée dans [Cra61] (cf. Théorème 1, p. 63). Notons que pour $t \in \mathbb{Z}\!\!\mathbb{Z}$, les variables $U(t)$ et $V(t)$ sont les projections de $X(t)$ sur $\mathcal{M}_\varepsilon(\Leftrightarrow\infty;t)$ et $\mathcal{M}(\Leftrightarrow\infty)$. Par ailleurs contrairement au cas stationnaire, le processus $X(\cdot)$ peut être purement indéterminable avec cependant $\sigma_\varepsilon^2(t) = 0$ en certains instants t. Ceci provient d'une singularité d'ordre soit fini ($|\beta(t, t \Leftrightarrow k)| = 1$) soit infini ($\sum_{k=1}^{+\infty} |\beta(t, t \Leftrightarrow k)|^2 = +\infty$). Dans ce cas, on peut obtenir l'unicité des

coefficients de la décomposition de $X(t)$ sur son processus d'innovation en posant $b(t, t \Leftrightarrow k) = 0$ lorsque $\varepsilon(t \Leftrightarrow k) = 0$.

Dans le cas stationnaire un processus est dit linéairement singulier d'ordre d lorsque la dimension de $\mathcal{M}$ est finie et égale à d. Il satisfait alors l'équation aux différences d'ordre (minimal) d traduisant la nullité des innovations d'ordre supérieur ou égal à d :

$$\sum_{k=0}^{d} a(k)X(t \Leftrightarrow k) = \sum_{k=0}^{d} a(k)X(t+k) = 0, \quad t \in \mathbb{Z}.$$

Réciproquement s'il existe t et $a(k)$, $k = 0, \ldots, d$, $d+1$ constantes non toutes nulles tels que,

$$\sum_{k=0}^{d} a(k)X(t \Leftrightarrow k) = 0,$$

alors la dimension de $\mathcal{M}$ est inférieure ou égale à d. Notons que cette situation correspond sur le plan de la fonction d'autocorrélation partielle à :

$$|\beta(k)| < 1,\ 1 \leq k < d < +\infty, \quad |\beta(d)| = 1, \quad \beta(k) = 0,\ k > d.$$

Ainsi linéairement singulier d'ordre d implique déterminable. Dans le cas non stationnaire, ces propriétés ne sont plus satisfaites. Lorsque l'espace $\mathcal{M}$ est de dimension d finie, la dimension de $\mathcal{M}(\Leftrightarrow\infty; t)$ peut être égale à d uniquement à partir d'un certain instant t_0. En particulier il peut exister $t \leq t_0$ tel que $\sigma_\varepsilon^2(t) > 0$, et par suite $\beta(t,t) \neq 0$ et $|\beta(t, t \Leftrightarrow k)| < 1$ pour tout $k \in \mathbb{N}^*$. Par conséquent le processus n'est plus nécessairement déterminable. Par ailleurs, on peut avoir $|\beta(t, t \Leftrightarrow k)| = 1$ à un instant t, sans que pour autant la dimension de $\mathcal{M}$ soit finie. Dans la situation non stationnaire, il est donc nécessaire de distinguer le cas où le processus présente des singularités d'ordre fini puisqu'il n'y a plus d'équivalence avec linéairement singulier d'ordre fini. Un processus sera dit *localement déterminable* si il existe $t \in \mathbb{Z}$ tel que $\beta(t,t) = 0$ ou $|\beta(t, t \Leftrightarrow k)| = 1$, pour un retard $k \in \mathbb{N}^*$. Dans le cas contraire, il sera dit *non localement déterminable*. Remarquons que non localement déterminable n'exclut pas le cas $\sigma_\varepsilon^2(t) = 0$ provenant d'une singularité d'ordre infini.

1.3 Algorithme de Levinson-Durbin Généralisé

Lev-Ari et coll. [Lev81] ont étendu les algorithmes de Schur et de Levinson-Durbin aux processus non stationnaires et non localement déterminable (i.e. dont les matrices de covariance sont non dégénérées). Leur algo-

rithme de type Levinson-Durbin détermine, de manière récursive sur l'ordre, les coefficients des innovations partielles normalisées $\eta^f(v;u)$ et $\eta^b(u;v)$ pour $s \leq u \leq v \leq t$, à partir de la fonction d'autocovariance connue sur le domaine $[s,\ldots,t]^2$. Par suite il fournit, de façon non standard, la décomposition de Cholesky de l'inverse de n'importe quelle matrice de covariance non dégénérée. La décomposition de la matrice de covariance elle-même est donnée par l'algorithme de type Schur. Comme dans le cas stationnaire, les paramètres $\beta(u,v)$ sont les éléments clés de ces deux algorithmes. Néanmoins la correspondance biunivoque entre les fonctions $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$ n'est pas évoquée. L'objectif de Lev-Ari et coll. était plutôt d'étendre ces deux algorithmes au cas non stationnaire pour se restreindre ensuite à une classe de processus proches de la stationnarité afin de diminuer leur coût de calcul. L'algorithme de Levinson-Durbin Généralisé (LDG), que nous donnons ci-dessous, a la même structure que celui de [Lev81], mais est étendu au cas localement déterminable. Comme nous l'avons déjà constaté, cette situation, qui n'est pas usuelle pour un processus stationnaire, est plus vraisemblable dans le cas non stationnaire. Par ailleurs, cet algorithme permet de déterminer les coefficients $\beta(u,v)$, pour $(u,v) \in [s,\ldots,t]^2$, connaissant ceux de la fonction $R(\cdot,\cdot)$ sur le même domaine, mais la correspondance inverse s'en déduit immédiatement.

A chaque étape n, l'algorithme détermine (cf. (1.6), (1.7), (1.8) et (1.9) ci-après) les coefficients (pas nécessairement uniques) de la décomposition de $\varepsilon^f(k;k \Leftrightarrow n)$ et $\varepsilon^b(k \Leftrightarrow n;k)$ sur $\mathcal{M}(k \Leftrightarrow n;k)$, soit :

$$\varepsilon^f(k;k \Leftrightarrow n) = \sum_{j=0}^{n} a_k^f(n,j) X(k \Leftrightarrow j), \quad a_k^f(n,0) = 1,$$
$$\varepsilon^b(k \Leftrightarrow n;k) = \sum_{j=0}^{n} a_k^b(n,j) X(k \Leftrightarrow n+j), \quad a_k^b(n,0) = 1.$$

En effet, les quantités $\varepsilon^f(k;k \Leftrightarrow n)$ et $\varepsilon^b(k \Leftrightarrow n;k)$ ainsi construites sont données en fonction de celles obtenues à l'étape précédente, selon les relations résultant de celles de la Proposition 1.2.1,

$$\varepsilon^f(k;k \Leftrightarrow n) = \varepsilon^f(k;k \Leftrightarrow n+1) \Leftrightarrow \beta(k,k \Leftrightarrow n) \frac{\sigma^f(k;k \Leftrightarrow n+1)}{\sigma^b(k \Leftrightarrow n;k \Leftrightarrow 1)} \varepsilon^b(k \Leftrightarrow n;k \Leftrightarrow 1),$$

$$\varepsilon^b(k \Leftrightarrow n;k) = \varepsilon^b(k \Leftrightarrow n;k \Leftrightarrow 1) \Leftrightarrow \beta(k \Leftrightarrow n,k) \frac{\sigma^b(k \Leftrightarrow n;k \Leftrightarrow 1)}{\sigma^f(k;k \Leftrightarrow n+1)} \varepsilon^f(k;k \Leftrightarrow n+1).$$

On montre ainsi, par récurrence, que les variables ci-dessus sont orthogonales à tout élément respectivement de $\mathcal{M}(k \Leftrightarrow n;k \Leftrightarrow 1)$ et de $\mathcal{M}(k \Leftrightarrow n+1;k)$, c'est-à-dire correspondent aux innovations d'ordre n. La définition de la fonction

d'autocorrélation partielle conduit à

$$\beta(k,k\Leftrightarrow n)=\frac{\left\langle \varepsilon^f(k;k\Leftrightarrow n+1),X(k\Leftrightarrow n)\right\rangle}{\sigma^f(k;k\Leftrightarrow n+1)\sigma^b(k\Leftrightarrow n;k\Leftrightarrow 1)}.$$

Si $\sigma^f(k;k\Leftrightarrow n+1)$ ou $\sigma^b(k\Leftrightarrow n;k\Leftrightarrow 1)$ est nul, alors le membre de gauche ainsi que le terme divisé par zéro, dans l'expression ci-dessus, sont nuls. L'égalité (1.4) ci-après est celle qui résulte de cette expression avec la convention $0^{-1}=0$. Les relations de (1.5) déterminent les variances des innovations d'ordre n utiles pour l'étape suivante. Sous forme algorithmique, on obtient la procédure suivante.

Algorithme 1.3.1. (LDG)
Pour $k=s,\dots,t$:

$$R(k,k)=\beta(k,k)=\sigma^{f2}(k;k)=\sigma^{b2}(k;k).$$

Pour $n=1,\dots,t\Leftrightarrow s$:

pour $k=s+n,\dots,t$, *avec les conventions* $\sum_{j=1}^{0}\dots=0$ *et* $0^{-1}=0$:

$$\beta(k,k\Leftrightarrow n)=\frac{R(k,k\Leftrightarrow n)+\sum_{j=1}^{n-1}a_k^f(n\Leftrightarrow 1,j)R(k\Leftrightarrow j,k\Leftrightarrow n)}{\sigma^f(k;k\Leftrightarrow n+1)\sigma^b(k\Leftrightarrow n;k\Leftrightarrow 1)}, \tag{1.4}$$

si $n\neq t\Leftrightarrow s$ *et* $k\geq s+n+1$:

$$\begin{aligned}\sigma^{f2}(k;k-n)&=\left[1-|\beta(k,k-n)|^2\right]\sigma^{f2}(k;k-n+1),\\ \sigma^{b2}(k-n-1;k-1)&=\left[1-|\beta(k-n-1,k-1)|^2\right]\sigma^{b2}(k-n-1;k-2),\end{aligned} \tag{1.5}$$

$$a_k^f(n,n)=\Leftrightarrow\beta(k,k\Leftrightarrow n)\frac{\sigma^f(k;k\Leftrightarrow n+1)}{\sigma^b(k\Leftrightarrow n;k\Leftrightarrow 1)}, \tag{1.6}$$

$$a_k^b(n,n)=\Leftrightarrow\beta(k\Leftrightarrow n,k)\frac{\sigma^b(k\Leftrightarrow n;k\Leftrightarrow 1)}{\sigma^f(k;k\Leftrightarrow n+1)}, \tag{1.7}$$

pour $j=1,\dots,n\Leftrightarrow 1$:

$$a_k^f(n,j)=a_k^f(n\Leftrightarrow 1,j)+a_k^f(n,n)a_{k-1}^b(n\Leftrightarrow 1,n\Leftrightarrow j), \tag{1.8}$$

$$a_k^b(n,j)=a_{k-1}^b(n\Leftrightarrow 1,j)+a_k^b(n,n)a_k^f(n\Leftrightarrow 1,n\Leftrightarrow j). \tag{1.9}$$

La convention $0^{-1}=0$ est nécessaire dans le cas localement déterminable. En effet lorsque $\sigma^b(k\Leftrightarrow n;k\Leftrightarrow 1)=0$, la convention conduit aux relations $a_k^f(n,j)=a_k^f(n\Leftrightarrow 1,j)$, $j=1,\dots,n\Leftrightarrow 1$, $a_k^f(n,n)=0$, cohérentes avec l'égalité $\varepsilon^f(k;k\Leftrightarrow n)=\varepsilon^f(k;k\Leftrightarrow n+1)$ qui résulte de $\varepsilon^b(k\Leftrightarrow n;k\Leftrightarrow 1)=0$. Un

raisonnement analogue pour l'expression (1.9) dans le sens rétrograde finit de justifier cette convention. Remarquons que dans ces expressions, la division par zéro n'a lieu que si la quantité $\beta(k, k \Leftrightarrow n)$ est déjà nulle. En particulier une division par zéro en (1.6) (i.e. $\sigma^b(k \Leftrightarrow n; k \Leftrightarrow 1) = 0$) implique également, en (1.7), que $a_k^b(n, n) = 0$ et par suite, en (1.9), que les premiers coefficients de $\varepsilon^b(k \Leftrightarrow n; k)$ sont ceux de l'innovation de $\varepsilon^b(k \Leftrightarrow n; k \Leftrightarrow 1)$, le dernier étant nul.

Le choix des coefficients $a_\cdot^f(\cdot, \cdot)$ et $a_\cdot^b(\cdot, \cdot)$ est obtenu par le procédé de construction lui-même. Ces coefficients sont uniques lorsque le processus considéré est non localement déterminable. Dans le cas contraire, le choix de l'algorithme peut conduire à des coefficients non nuls pondérant des variables nulles presque sûrement. Supposons que la variable $X(k)$ puisse s'exprimer comme une combinaison linéaire finie de son passé. Soit n le plus petit entier strictement positif tel que $X(k) \in \mathcal{M}(k \Leftrightarrow n; k \Leftrightarrow 1)$. Il est naturel de souhaiter que les coefficients pondérant la variable $X(k \Leftrightarrow n)$ (resp. $X(k)$) soient nuls pour toutes les innovations $\varepsilon^f(v; u)$ (resp. $\varepsilon^b(u; v)$), $s \leq u \leq k \Leftrightarrow n \leq k \leq v \leq t$. Suite à la remarque précédente, ce choix est clairement assuré pour $\varepsilon^f(k; u)$ et $\varepsilon^b(k \Leftrightarrow n; v)$, mais ne l'est pas forcément pour les autres innovations. Néanmoins, ce nouveau jeu de coefficients peut être obtenu à partir de celui qu'impose l'algorithme en remplaçant, dans la situation décrite ci-dessus, la variable $X(k \Leftrightarrow n)$ par $\sum_{j=1}^n a_k^b(n, j) X(k \Leftrightarrow n + j)$ et $X(k)$ par $\sum_{j=1}^n a_k^b(n, j) X(k \Leftrightarrow j)$. Ceci doit être réalisé à partir de la connaissance de l'ensemble des couples (k_i, n_i) vérifiant $|\beta(k_i, k_i \Leftrightarrow n_i)| = 1$. Pour le cas trivial $X(k) = 0$ *p.s.*, équivalent à $\beta(k, k) = 0$, il suffit de poser égaux à zéro tous les coefficients pondérant cette variable.

L'Algorithme *LDG* montre que la relation entre les fonctions $R(\cdot, \cdot)$ et $\beta(\cdot, \cdot)$ est plus forte que la bijection usuelle entre deux ensembles de fonctions au sens de la remarque suivante.

Remarque 1.3.1. *La bijection entre $\beta(\cdot, \cdot)$ et $R(\cdot, \cdot)$ est de nature temporelle au sens où leurs restrictions à $[s, \ldots, t]^2$, c'est-à-dire les paramétrisations de la structure de $\{X(s), \ldots, X(t)\}$, sont en correspondance biunivoque.*

L'extension au cas localement déterminable de l'algorithme de Levinson-Durbin permet de déterminer une décomposition de Cholesky de l'inverse (généralisé) de n'importe quelle matrice hermitienne définie non négative. En effet, rappelons que $R_{s,t} = \{R(s + i, s + j)\}_{i,j=0,\ldots,t-s}$, représente la matrice

de covariance du vecteur aléatoire $[X(s),\dots,X(t)]^T$ et notons

$$\varepsilon^f_{s,t} = \begin{pmatrix} \varepsilon^f(s;s) \\ \varepsilon^f(s+1;s) \\ \vdots \\ \varepsilon^f(t;s) \end{pmatrix} = A^f_{s,t} \begin{pmatrix} X(s) \\ X(s+1) \\ \vdots \\ X(t) \end{pmatrix},$$

où

$$A^f_{s,t} = \begin{pmatrix} 1 & & & \\ a^f_{s+1}(1,1) & 1 & (0) & \\ \vdots & \ddots & \ddots & \\ a^f_t(t \Leftrightarrow s, t \Leftrightarrow s) & \dots & a^f_t(t \Leftrightarrow s, 1) & 1 \end{pmatrix}$$

et

$$E\left\{\varepsilon^f_{s,t}\varepsilon^{f\ *}_{s,t}\right\} = \begin{pmatrix} \sigma^{f2}(s;s) & & & \\ & \sigma^{f2}(s+1;s) & (0) & \\ & (0) & \ddots & \\ & & & \sigma^{f2}(t;s) \end{pmatrix} = \Sigma^{f2}_{s,t},$$

où $*$ désigne le transposé conjugué. Alors une décomposition d'un inverse généralisé $R_{s,t}{}^-$ de la matrice de covariance $R_{s,t}$ est donnée par $R_{s,t}{}^- = A^{f\ *}_{s,t}\Sigma^{f2+}_{s,t}A^f_{s,t}$, où les coefficients des matrices $\Sigma^{f2}_{s,t}$ et $A^f_{s,t}$ sont fournis par l'algorithme. La matrice diagonale $\Sigma^{f2+}_{s,t}$ est l'inverse généralisé de $\Sigma^{f2}_{s,t}$ obtenu en inversant les termes non nuls. Cette décomposition n'est qu'un sous-produit de l'Algorithme 1.3.1 qui détermine en fait tous les coefficients des innovations $\varepsilon^f(v;u)$ et $\varepsilon^b(u;v)$, $s \le u \le v \le t$. Ce procédé, qui permet d'obtenir une décomposition de Cholesky d'une matrice hermitienne définie non négative, n'est pas standard mais rejoint celui proposé par Delsarte et coll. [Del80] dans le cas non localement déterminable.

La décomposition $\Sigma^{f2}_{s,t} = A^f_{s,t}R_{s,t}A^{f\ *}_{s,t}$, implique que les déterminants des matrices $R_{s,t}$ et $\Sigma^{f2}_{s,t}$ sont identiques et par suite

$$\det R_{s,t} = \prod_{k=s}^{t} \sigma^{f2}(k;s),$$

ou encore, en termes d'autocorrélations partielles (cf. Proposition 1.2.1)

$$\det R_{s,t} = \left[\prod_{k=s}^{t} \beta(t,t)\right] \prod_{k=s+1}^{t} \prod_{j=1}^{k-s} \left[1 \Leftrightarrow |\beta(k, k \Leftrightarrow j)|^2\right].$$

Remarquons que dans l'expression de ce déterminant interviennent toutes les autocorrélations partielles $\beta(u,v)$, $s \le u \le v \le t$. Lorsque le processus est

stationnaire, on retrouve la formule

$$\det R_{t-s} = \det R_n = \beta(0)^{n+1} \prod_{j=1}^{n} \left[1 \Leftrightarrow |\beta(j)|^2\right]^{n+1-j}.$$

Une matrice symétrique est définie positive si et seulement si le déterminant de certains mineurs principaux est strictement positif (cf. [GAN60] : théorème 19, p. 337). Ainsi l'expression de $\det R_{s,t}$ et l'Algorithme LDG assurent immédiatement la correspondance biunivoque entre $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$ dans le cas non localement déterminable. C'est l'approche utilisée par Ramsey [RAM74] pour les processus stationnaires non linéairement singuliers d'ordre fini. La condition $\det R_{s,t} \geq 0$, pour tout (s,t), n'est pas suffisante dans le cas localement déterminable. Il est alors nécessaire d'exhiber un processus $X(\cdot)$ admettant la fonction $R(\cdot,\cdot)$, associée à $\beta(\cdot,\cdot)$ par l'Algorithme LDG, comme fonction d'autocovariance. C'est le rôle du procédé constructif que donne la Proposition 1.2.2. La démonstration de Ramsey [RAM74], pour le cas stationnaire, invoque également le processus linéairement singulier, mais comporte une légère erreur dans la preuve de la condition suffisante. Par ailleurs celle de la condition nécessaire, qui utilise l'expression des déterminants, ne s'étend pas au cas non stationnaire.

Notons enfin que cet algorithme permet de tester facilement la positivité d'une matrice $R_{s,t}$ donnée et de décrire l'ensemble des fonctions définies non négatives $R(\cdot,\cdot)$ qui étendent cette série de valeurs. Pour la classe des processus périodiquement corrélés (cf. section 3.1), cet algorithme coïncide avec celui proposé par Sakai [SAK83] dans le cas non localement déterminable et étendu par Pham [PHA92] à la situation générale.

1.4 Modèles autorégressifs

Dans cette section, nous proposons une définition du modèle autorégressif liée à l'approche en termes d'autocorrélations partielles. C'est-à-dire que nous faisons en sorte que, comme dans le cas stationnaire, cette classe de processus soit facilement caractérisée par la fonction $\beta(\cdot,\cdot)$. Cependant la non-stationnarité entraîne la perte de certaines propriétés. Nous rappelons ci-après quelques points remarquables du cas stationnaire pour justifier l'approche retenue dans le cas non stationnaire.

En toute généralité, on peut dire qu'un processus $X(\cdot)$ est autorégressif

d'ordre p lorsqu'il satisfait une équation aux différences stochastique,

$$\sum_{k=0}^{p} \tilde{a}(k)X(t \Leftrightarrow k) = \tilde{\varepsilon}(t), \quad t \in \mathbb{Z}, \quad \tilde{a}(0) = 1, \quad \tilde{a}(p) \neq 0, \tag{1.10}$$

dans laquelle $\tilde{\varepsilon}(\cdot)$ est un bruit blanc de variance $\tilde{\sigma}_\varepsilon^2$. Soit F la mesure spectrale d'un processus stationnaire satisfaisant (1.10), alors on a nécessairement

$$\frac{\tilde{\sigma}_\varepsilon^2}{2\pi} d\lambda = \left|\tilde{\varphi}^b(e^{-i\lambda})\right|^2 dF(\lambda),$$

où $\tilde{\varphi}^b(z) = \sum_{k=0}^{p} \tilde{a}(k)z^k$. Lorsque $\tilde{\varphi}^b(\cdot)$ possède des racines de module 1, la variance $\tilde{\sigma}_\varepsilon^2$ doit être nulle. Le processus est alors linéairement singulier d'ordre $d \leq p$ et satisfait

$$\sum_{k=0}^{d} a(k)X(t \Leftrightarrow k) = 0, \quad t \in \mathbb{Z}, \quad a(0) = 1, \quad a(d) \neq 0,$$

avec $\varphi^b(z) = \prod_{k=1}^{d}(1 \Leftrightarrow z/\tilde{z}_k)$, les racines simples $\tilde{z}_k$ étant prises parmi celles de $\tilde{\varphi}^b(\cdot)$ qui sont de module 1. Notons que ces racines caractérisent uniquement le support de $F(\cdot)$ qui comporte exactement d fréquences distinctes modulo 2π, λ_k, $k = 1, \ldots, d$, définies par $\tilde{z}_k = e^{-i\lambda_k}$. Dans le cas contraire, la solution est unique et sa mesure spectrale, qui est absolument continue, admet la densité

$$f(\lambda) = \frac{\tilde{\sigma}_\varepsilon^2}{2\pi} \left|\tilde{\varphi}^b(e^{-i\lambda})\right|^{-2}.$$

Ainsi l'équation (1.10) admet toujours une solution stationnaire, avec $\tilde{\sigma}_\varepsilon^2 = 0$ lorsque $\tilde{\varphi}^b(\cdot)$ possède des racines situées sur le cercle unité. Par ailleurs soit $X(\cdot)$ une solution de (1.10) stationnaire, alors ce processus admet une autre représentation du même ordre :

$$\sum_{k=0}^{p} a(k)X(t \Leftrightarrow k) = \varepsilon(t), \quad t \in \mathbb{Z}, \quad a(0) = 1, \quad a(p) \neq 0, \tag{1.11}$$

où $\varepsilon(\cdot)$ est le processus d'innovation. Notons que ce résultat est immédiat lorsque $X(\cdot)$ est linéairement singulier puisque dans ce cas $\varepsilon(t) = 0$, pour tout t. Sinon les coefficients de l'équation (1.11) sont obtenus de la façon suivante. Le polynôme $\tilde{\varphi}^b(\cdot)$ est modifié afin qu'il ne possède plus de racine à l'intérieur du disque unité sans changer la densité $f(\cdot)$. En effet soit α une racine de $\tilde{\varphi}^b(\cdot)$ située à l'intérieur du disque. Alors le polynôme

$$\frac{1 \Leftrightarrow \overline{\alpha}z}{1 \Leftrightarrow \frac{1}{\alpha}z} \tilde{\varphi}^b(z),$$

admet $1/\overline{\alpha}$ comme racine à la place de α. L'égalité $|\alpha \Leftrightarrow z| = |1 \Leftrightarrow \overline{\alpha}z|$ sur

$|z| = 1$, implique que

$$\left| \frac{1 \Leftrightarrow \overline{\alpha} z}{1 \Leftrightarrow \frac{1}{\alpha} z} \tilde{\varphi}^b(z) \right| = |\alpha| \left| \tilde{\varphi}^b(z) \right|, \quad |z| = 1.$$

Ainsi en effectuant cette modification pour chaque racine située à l'intérieur du disque, on obtient

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left| \varphi^b(e^{-i\lambda}) \right|^{-2},$$

où $\varphi^b(z) = \sum_{k=0}^{p} a(k) z^k$ est un polynôme dont toutes les racines sont à l'extérieur du disque et σ_ε^2 est la variance $\tilde{\sigma}_\varepsilon^2$ multipliée par le produit des carrés des modules des racines de $\tilde{\varphi}^b(\cdot)$ situées à l'intérieur. Le processus $X(\cdot)$ satisfait alors l'équation (1.11) caractérisée par $\varphi^b(z) \neq 0$ pour $|z| \leq 1$. Notons que dans ce cas, le polynôme $\varphi^f(z) = z^p \varphi^b(1/z)$ est dit stable. Nous montrons ci-après que $\varepsilon(\cdot)$ dans (1.11) est alors le processus d'innovation. Dans le cas où $\varphi^b(z) \neq 0$ pour $|z| \leq 1$, le développement de Laurent de l'inverse de $\varphi^b(\cdot)$ est unilatéral,

$$\frac{1}{\varphi^b(z)} = \sum_{k=0}^{+\infty} b(k) z^k, \quad b(0) = 1,$$

et conduit à

$$X(t) = \sum_{k=0}^{+\infty} b(k) \varepsilon(t \Leftrightarrow k), \quad t \in \mathbb{Z},$$

qui avec (1.11), donne le résultat. Les modèles autorégressifs stationnaires d'ordre p ($AR(p)$) peuvent donc être définis à partir de l'équation (1.11) en imposant que $\varepsilon(\cdot)$ soit le processus d'innovation. Pour assurer l'unicité de cette représentation, nous imposons aussi que l'équation soit d'ordre minimal quand $\sigma_\varepsilon^2 = 0$. L'existence d'une solution est assurée lorsque le polynôme $\varphi^f(\cdot)$ est stable si $\sigma_\varepsilon^2 > 0$, ou sinon possède p racines distinctes sur le cercle unité. Remarquons que pour un processus $AR(p)$ stationnaire, les composantes purement indéterminable et déterminable ne peuvent cohabiter dans la décomposition de Wold-Cramér.

Les autocorrélations partielles $\beta(0), \ldots, \beta(p)$, constituent une paramétrisation du modèle $AR(p)$. De plus un processus stationnaire est $AR(p)$ si et seulement si sa fonction d'autocorrélation partielle satisfait

$$0 < |\beta(p)| \leq 1, \qquad \beta(k) = 0, \quad k > p.$$

Notons que pour montrer que la condition est nécessaire, à partir de l'équation aux différences stochastique, il est important que le bruit blanc soit le processus d'innovation. En effet lorsque $\sigma_\varepsilon^2 > 0$, on a $\beta(k) = 0$ pour $k > p$ si et seulement si pour tout t l'innovation $\varepsilon^f(t; s+1)$ est orthogonale à $\varepsilon^b(s; t \Leftrightarrow 1)$ et donc à $X(s)$ pour $s < t \Leftrightarrow p$. Par ailleurs supposons que $X(\cdot)$ satisfasse

une équation de la forme (1.10), alors

$$\varepsilon^f(t; s+1) = \tilde{\varepsilon}(t) \Leftrightarrow \tilde{\varepsilon}_X(t; s+1),$$

où $\tilde{\varepsilon}_X(t; s+1)$ désigne la projection de $\tilde{\varepsilon}(t)$ sur $\mathcal{M}(s+1; t \Leftrightarrow 1)$. Ainsi la non corrélation entre $\tilde{\varepsilon}(t) \Leftrightarrow \tilde{\varepsilon}_X(t; s+1)$ et $X(s)$ pour tout $s < t \Leftrightarrow p$, équivaut à $\tilde{\varepsilon}(t) \Leftrightarrow \tilde{\varepsilon}_X(t; \Leftrightarrow\infty) \in \mathcal{M}(t \Leftrightarrow p; t)$. Ce résultat est évident lorsque $\tilde{\varepsilon}(\cdot)$ est le processus d'innovation mais ne l'est plus en toute généralité.

Dans le cas où $\sigma_\varepsilon^2 > 0$, les paramètres $\{a(1), \ldots, a(p); \sigma_\varepsilon^2\}$ caractérisent la structure au second ordre du processus $X(\cdot)$ mais sont clairement moins faciles à identifier par rapport aux coefficients $\{\beta(0), \ldots, \beta(p)\}$. De plus l'algorithme de Levinson-Durbin permet de faire le lien entre ces deux paramétrisations. Dans le cas stationnaire, les coefficients des innovations partielles dans le sens rétrograde sont les conjugués de ceux du sens progressif et les variances résiduelles sont identiques dans les deux sens. Les équations (1.6), (1.7), (1.8) et (1.9) de l'Algorithme *LDG* se réduisent alors à : pour $n = 1, \ldots, p$ $a(n, n) = \Leftrightarrow\!\!\!/\beta(n)$ et pour $k = 1, \ldots, n \Leftrightarrow 1$,

$$a(n, k) = a(n \Leftrightarrow 1, k) \Leftrightarrow \beta(n)\overline{a(n \Leftrightarrow 1, n \Leftrightarrow k)}.$$

Les coefficients $a(k)$, $k = 1, \ldots, p$, du modèle sont donnés par $a(p, k)$ et la variance σ_ε^2 par $\beta(0) \prod_{k=1}^{p} [1 \Leftrightarrow |\beta(k)|^2]$. Réciproquement, partant des paramètres $\{a(1), \ldots, a(p); \sigma_\varepsilon^2\}$, les équations ci-dessus peuvent être inversées pour déterminer les autocorrélations partielles associées. On a pour $k = 1, \ldots, p$, $a(p, k) = a(k)$ et $\beta(p) = \Leftrightarrow\!\!\!/ a(p, p)$ puis pour $n = p, \ldots, 2$:

$$\begin{aligned} a(n \Leftrightarrow 1, k) &= \frac{a(n, k) + \beta(n)\overline{a(n, n \Leftrightarrow k)}}{1 \Leftrightarrow |\beta(n)|^2}, \quad 1 \leq k < n \\ \beta(n \Leftrightarrow 1) &= \Leftrightarrow\!\!\!/ a(n \Leftrightarrow 1, n \Leftrightarrow 1). \end{aligned}$$

La variance du processus est donnée par $\beta(0) = \sigma_\varepsilon^2 / \prod_{k=1}^{p} [1 \Leftrightarrow |\beta(k)|^2]$. En particulier, ceci permet de tester facilement s'il existe une solution $AR(p)$ pour un jeu de paramètres $\{a(1), \ldots, a(p); \sigma_\varepsilon^2\}$ donné.

Une façon naturelle pour étendre la définition des modèles $AR(p)$ au cas non stationnaire serait de considérer la représentation (1.10) en faisant dépendre du temps les coefficients ainsi que la variance du bruit blanc. Cependant le problème de l'existence d'une solution est un point délicat. En effet, une approche dans le domaine spectral est clairement impossible. Dans le domaine temporel, on peut citer les travaux de Hallin et coll. [Hal77] basés sur la théorie des équations aux différences linéaires (cf. [Mil68]). Ils donnent des conditions suffisantes d'existence d'une solution, qui s'expriment à travers l'ensemble des filtres et non séparément sur chacun d'eux. De plus lorsque la solution existe, elle est purement indéterminable et $\tilde{\varepsilon}(\cdot)$ est alors

le processus d'innovation. Par ailleurs, il n'est pas évident qu'une solution de l'analogue de l'équation (1.10), lorsqu'elle existe, satisfasse l'analogue de (1.11). Comme nous l'avons vu pour le cas stationnaire, il faudrait montrer que $\tilde{\varepsilon}(t)$, privé de sa projection sur l'ensemble $\mathcal{M}(\Leftrightarrow\infty; t \Leftrightarrow 1)$, appartient à $\mathcal{M}(t \Leftrightarrow p; t \Leftrightarrow 1)$. Néanmoins, afin de conserver la caractérisation de ces modèles par la fonction d'autocorrélation partielle, nous proposons de définir, dans le cas non stationnaire, les modèles $AR(p)$ de la façon suivante.

Définition 1.4.1. *Un processus $X(\cdot)$ est dit autorégressif d'ordre p, noté $AR(p)$, si pour tout $t \in \mathbb{Z}$, il existe des constantes $a_t(k)$, $k = 1, \ldots, p$ telles que*

$$\sum_{k=0}^{p} a_t(k) X(t \Leftrightarrow k) = \varepsilon(t), \quad a_t(0) = 1, \tag{1.12}$$

où $\varepsilon(\cdot)$ est le processus d'innovation, p étant le plus petit entier pour lequel ces relations sont satisfaites.

Notons que nous n'imposons pas la condition trop restrictive $a_t(p) \neq 0$, pour tout t. D'autre part, pour que $\varepsilon(\cdot)$ soit le processus d'innovation dans (1.12), il est nécessaire et suffisant que $\varepsilon(\cdot) = \{\varepsilon(t), t \in \mathbb{Z}\}$ soit une suite de variables aléatoires centrées non corrélées entre elles, telle que pour tout $t \in \mathbb{Z}$, $\varepsilon(t)$ soit non corrélé avec $X(s)$, $s < t$. De plus, la variance $\sigma^2_\varepsilon(t)$ pouvant s'annuler, le processus peut présenter des singularités d'ordre fini. La fonction $\beta(\cdot, \cdot)$ caractérise de façon simple les modèles $AR(p)$ ainsi définis.

Proposition 1.4.1. *Un processus $X(\cdot)$ est autorégressif d'ordre p si et seulement si sa fonction d'autocorrélation partielle $\beta(\cdot, \cdot)$ satisfait*

$$\beta(t, t \Leftrightarrow k) = 0, \quad \forall t \in \mathbb{Z}, \quad \forall k > p; \qquad \exists t \in \mathbb{Z}, \quad \beta(t, t \Leftrightarrow p) \neq 0.$$

Démonstration.- Si $X(\cdot)$ est $AR(p)$, le processus $\varepsilon(\cdot)$ dans (1.12) est le processus d'innovation. Alors $\varepsilon^f(t; t \Leftrightarrow n) = \varepsilon^f(t; t \Leftrightarrow p) = \varepsilon(t)$ pour $n \geq p$ et l'égalité des variances

$$\begin{aligned} \sigma^{f2}(t; t \Leftrightarrow n) &= \beta(t,t) \prod_{k=1}^{n} \left[1 \Leftrightarrow |\beta(t, t \Leftrightarrow k)|^2\right] \\ &= \beta(t,t) \prod_{k=1}^{p} \left[1 \Leftrightarrow |\beta(t, t \Leftrightarrow k)|^2\right] = \sigma^{f2}(t; t \Leftrightarrow p), \end{aligned}$$

montre que $\beta(t, t \Leftrightarrow k) = 0$ pour $k > p$. En effet cela est évident si $\sigma^{f2}(t; t \Leftrightarrow p) > 0$. Sinon cette variance est nulle parce que $\beta(t,t) = 0$ ou parce qu'il existe $k \leq p$ tel que $|\beta(t, t \Leftrightarrow k)| = 1$. Les conventions adoptées font que, dans ces deux situations, $\beta(t, t \Leftrightarrow j)$ sera nul pour $j > p$. Soit maintenant $\tilde{p}$ le plus

grand entier k inférieur ou égal à p pour lequel on ait $\beta(t, t \Leftrightarrow k) \neq 0$ lorsque t parcourt $\mathbb{Z}$. Alors $X(\cdot)$ satisfait (1.12) avec $p = \tilde{p}$ et $\varepsilon(t) = \varepsilon^f(t; t \Leftrightarrow \tilde{p})$, $t \in \mathbb{Z}$. Ceci montre que $p = \tilde{p}$, par définition de l'ordre du modèle, et que l'existence de $t \in \mathbb{Z}$ tel que $\beta(t, t \Leftrightarrow p) \neq 0$ est satisfaite. D'autre part ces derniers points établissent clairement la condition suffisante de la proposition. $\diamondsuit$

La classe des processus autorégressifs ainsi définie est très large. La Proposition 1.2.2 assure l'existence d'un processus $X(\cdot)$ associé à chaque élément de $\mathcal{D}_\beta$ satisfaisant les contraintes de la Proposition 1.4.1 et l'unicité de sa structure est garantie puisqu'elle est caractérisée par $\beta(\cdot, \cdot)$. L'Algorithme *LDG* fournit les paramètres de la représentation (1.12) dont l'unicité (dans le cas localement déterminable) peut toujours être obtenue moyennant certaines conventions (cf. section 1.3). Cependant nous ne sommes pas en mesure de décrire les contraintes sur les coefficients $a_t(k)$, $t \in \mathbb{Z}$, $k = 1, \ldots, p$ et sur les variances résiduelles $\sigma^2_\varepsilon(t)$, $t \in \mathbb{Z}$, pour garantir l'existence au second ordre d'un processus satisfaisant la représentation (1.12). D'autre part, contrairement au cas stationnaire, les paramètres $\{a_t(k), k = 1, \ldots, p; \sigma^2_\varepsilon(t)\}$, $t \in \mathbb{Z}$, ne caractérisent pas la structure du processus $X(\cdot)$ dans le cas non localement déterminable. Soit $X(t) = U(t) + V(t)$, $t \in \mathbb{Z}$, la décomposition de Wold-Cramér de $X(\cdot)$, alors le processus purement indéterminable $U(\cdot)$ satisfait la même représentation que $X(\cdot)$,

$$\sum_{k=0}^{p} a_t(k) U(t \Leftrightarrow k) = \varepsilon(t), \quad t \in \mathbb{Z}, \quad a_t(0) = 1,$$

avec les mêmes coefficients $a_t(k)$, $k = 1, \ldots, p$, et les mêmes variances $\sigma^2_\varepsilon(t)$, $t \in \mathbb{Z}$. Il en est de même pour $V(\cdot)$, mais dans ce cas $\sigma^2_\varepsilon(t)$ est identiquement nul. La fonction d'autocorrélation partielle $\beta_U(\cdot, \cdot)$ associée à $U(\cdot)$ est évidemment différente de celle de $X(\cdot)$, sauf lorsque $V(\cdot)$ n'existe pas. Si l'on exclut les singularités d'ordre fini ($|\beta(t, s)| \neq 1$, $t \neq s$ et $\beta(t, t) \neq 0$), les paramètres du modèle autorégressif sont uniques et sont donnés par l'Algorithme *LDG*. Ils sont donc identiques pour $X(\cdot)$ et $U(\cdot)$. En effet, considérons par exemple le modèle

$$X(t) + aX(t \Leftrightarrow 1) = \varepsilon(t),$$

où la constante a est telle que $|a| < 1$ et $\sigma^2_\varepsilon(t) = 1$, pour tout $t \in \mathbb{Z}$. Alors le processus stationnaire $U(\cdot)$, dont la structure est donnée par $\beta_U(0) = 1/(1 \Leftrightarrow a^2)$, $\beta_U(1) = \Leftrightarrow a$, $\beta_U(k) = 0$, $k > 1$, satisfait l'équation ci-dessus. D'autre part soit $X(\cdot)$, le processus défini par

$$X(t) = U(t) + V(t), \quad t \in \mathbb{Z},$$

où $V(\cdot)$ est un processus non nul, non corrélé avec $U(\cdot)$ et de fonction

d'autocovariance $R_V(\cdot,\cdot)$ telle que

$$\begin{aligned} R_V(t,t) &= a^2 R_V(t \Leftrightarrow 1, t \Leftrightarrow 1), \\ R_V(t, t \Leftrightarrow k) &= \Leftrightarrow a R_V(t \Leftrightarrow 1, t \Leftrightarrow k), \quad k \geq 1. \end{aligned}$$

Alors la fonction d'autocovariance du processus $X(\cdot)$ satisfait,

$$\begin{aligned} R_X(t,t) &= \frac{1}{1 \Leftrightarrow a^2} + R_V(t,t), \\ R_X(t, t \Leftrightarrow k) &= \Leftrightarrow a R_X(t \Leftrightarrow 1, t \Leftrightarrow k), \quad k \geq 1. \end{aligned}$$

L'Algorithme LDG conduit à

$$\begin{aligned} \beta_X(t,t) &= \frac{1}{1 \Leftrightarrow a^2} + R_V(t,t), \\ \beta_X(t, t \Leftrightarrow 1) &= \Leftrightarrow a \frac{\left[\frac{1}{1-a^2} + R_V(t \Leftrightarrow 1, t \Leftrightarrow 1)\right]^{\frac{1}{2}}}{\left[\frac{1}{1-a^2} + a^2 R_V(t \Leftrightarrow 1, t \Leftrightarrow 1)\right]^{\frac{1}{2}}}, \\ \beta_X(t, t \Leftrightarrow k) &= 0, \qquad k > 1. \end{aligned}$$

Il est clair que pour tout $t \in \mathbb{Z}$, $|\beta_X(t, t \Leftrightarrow 1)| < 1$ puisque $|a| < 1$. Ainsi le processus $X(\cdot)$ ne présente pas de singularités d'ordre fini. D'après la Proposition 1.4.1, il est autorégressif d'ordre 1, et les paramètres de la représentation (1.12), qui sont uniques, sont déterminés par

$$\begin{aligned} a_t(k) &= a_t^f(1,1) = a, \\ \sigma_\varepsilon^2(t) &= \left[1 \Leftrightarrow |\beta_X(t, t \Leftrightarrow 1)|^2\right] \beta_X(t,t) = 1. \end{aligned}$$

Les deux processus $X(\cdot)$ et $U(\cdot)$ admettent la même représentation $AR(1)$ mais leurs structures au second ordre sont bien différentes puisque $V(\cdot)$ n'est pas identiquement nul. Remarquons que, pour cet exemple, l'approche de [HAL77] conduit à la solution stationnaire alors que, comme nous venons de le montrer, on peut envisager d'autres solutions.

Une façon naturelle d'introduire les processus autorégressifs dans le cas non stationnaire est de se restreindre à ceux définis pour $t \geq 0$. Dans ce cas on obtient une définition constructive de $X(\cdot)$ en se donnant une condition initiale $\{X(0), \ldots, X(p \Leftrightarrow 1)\}$, un bruit blanc $\{\varepsilon(t), t \geq p\}$ qui lui soit non corrélé, et un jeu de filtres $\{a_t(k), k = 1, \ldots, p\}_{t \geq p}$ quelconque. La paramétrisation comprend la structure de la condition initiale, le jeu de filtres et les variances $\sigma_\varepsilon^2(t)$, $t \geq p$. Le problème de l'existence ne se pose plus et l'étude des paramètres est simplifiée.

Chapitre 2

Aspect spectral

L'analyse temps-fréquence, ou temps-échelle, constitue un domaine d'étude très actuel, en particulier grâce au développement de l'outil "ondelettes". On trouvera une présentation très complète du sujet dans l'ouvrage de Flandrin [Fla93]. Cependant l'accent est mis sur la représentation du signal, le plus souvent déterministe et en temps continu. Nous nous plaçons ici dans l'objectif de décrire la structure de covariance d'un processus aléatoire à temps discret. Le spectre évolutif instantané que nous introduisons est directement associé à la fonction d'autocorrélation partielle décrite précédemment. La notion n'est donc pas transposable au temps continu car l'autocorrélation partielle n'est pas définie dans ce cas. Par ailleurs aucune représentation du processus n'a été jusqu'ici associée à ce nouveau spectre. Notre approche se situe dans le droit fil de l'article de Loynes [Loy68] sur la conception d'un spectre pour les processus non stationnaires. Le spectre que nous proposons est comparable à ceux de Mélard [Mél78] et de Grenier [Gre84] bien que ces derniers soient plus directement reliés à la description temporelle du processus plutôt qu'à celle de sa structure au second ordre.

Le spectre évolutif instantané est défini dans la première section. Il est ensuite analysé à travers les propriétés souhaitées par Loynes. Les comparaisons avec les spectres de Mélard et Grenier font l'objet des sections 3 et 4. Enfin, dans la dernière section, le comportement du spectre évolutif instantané est illustré par quelques exemples notamment dans le cas de la discrétisation du mouvement Brownien et pour les "chirps linéaires".

2.1 Spectre évolutif instantané

Dans le cas stationnaire, la correspondance biunivoque entre la fonction d'autovariance $R(\cdot)$ et la mesure spectrale dF est gérée par la transformée

de Fourier :

$$R(k) = \int_{-\pi}^{\pi} e^{ik\lambda} dF(\lambda), \quad k \in \mathbb{Z},$$

$$F(\lambda) \Leftrightarrow F(\mu) = \lim_{n \to +\infty} \sum_{k=-n}^{+n} \frac{e^{-i\mu k} \Leftrightarrow e^{-i\lambda k}}{2i\pi k} R(k),$$

en tous points de continuité λ et μ de la fonction de répartition F associée à dF.

A ce titre $R(\cdot)$ est une fonction de type positif, ce qui induit un ensemble de contraintes entre les différentes valeurs $R(k)$, $k \in \mathbb{Z}$, difficile à contrôler. C'est la raison pour laquelle la fonction d'autocorrélation partielle $\beta(\cdot)$ est plus attractive. En effet $\beta(\cdot)$ possède comme $R(\cdot)$ la symétrie hermitienne, $\beta(k) = \overline{\beta(\Leftrightarrow k)}$, $k \in \mathbb{Z}$, et doit satisfaire l'une des deux conditions suivantes :

(i) $|\beta(k)| < 1$, $k \in \mathbb{N}^*$,

(ii) $|\beta(k)| < 1$, $1 \leq k < d$, $|\beta(d)| = 1$, $\beta(k) = 0$, $k > d$.

La correspondance entre $\beta(\cdot)$ et $R(\cdot)$ est assurée par l'algorithme de Levinson-Durbin et permet de remonter à la mesure spectrale par la formule d'inversion ci-dessus. Les relations directes entre $\beta(\cdot)$ et dF sont moins classiques, elles sont liées à la théorie des polynômes orthogonaux sur le cercles unité (cf. [Ger60]). Introduisons l'espace $\mathcal{L}^2\{[\Leftrightarrow\pi, \pi], dF\}$ constitué des fonctions à valeurs complexes et de carré intégrable par rapport à la mesure dF. Alors cet espace, muni du produit hermitien,

$$\langle \phi, \psi \rangle_{dF} = \int_{-\pi}^{\pi} \phi(\lambda) \overline{\psi(\lambda)} dF(\lambda),$$

est l'espace de Hilbert engendré par $\left\{e^{in\lambda}, n \in \mathbb{Z}\right\}$. L'orthogonalisation de la suite $\left\{e^{in\lambda}, n \in \mathbb{N}\right\}$ dans $\mathcal{L}^2\{[\Leftrightarrow\pi, \pi], dF\}$, selon le procédé de Gram-Schmidt, définit un système de polynômes orthogonaux $\left\{\varphi_n^f(e^{in\lambda}), n \in \mathbb{N}\right\}$ appelés polynômes de Szegö de première espèce. L'espace $\mathcal{L}^2\{[\Leftrightarrow\pi, \pi], dF\}$ étant isométrique à $\mathcal{M}$ par l'application $\mathcal{J}\{e^{in\lambda}\} = X(n)$, on a $\varepsilon^f(n;0) = \mathcal{J}\{\varphi_n^f(e^{i\lambda})\}$ et par suite

$$\varphi_n^f(z) = \sum_{k=0}^{n} a(n,k) z^{n-k}, \quad n \in \mathbb{N}.$$

Dans le sens rétrograde, on obtient $\varepsilon^b(0;n) = \mathcal{J}\{\varphi_n^b(e^{i\lambda})\}$, où φ_n^b est le polynôme réciproque de φ_n^f :

$$\varphi_n^b(z) = \sum_{k=0}^{n} \overline{a(n,k)} z^k = z^n \overline{\varphi_n^f}(\frac{1}{z}).$$

Le système orthogonal $\left\{\varphi_n^f, n \in I\!N\right\}$ est donc caractérisé en fonction de $\beta(\cdot)$ par la double récurrence, déduite de celle des innovations partielles,

$$\begin{array}{l} \varphi_0^f(z) = \varphi_0^b(z) = 1, \\ \varphi_n^f(z) = z\varphi_{n-1}^f(z) \Leftrightarrow \beta(n)\varphi_{n-1}^b(z), \\ \varphi_n^b(z) = \varphi_{n-1}^b(z) \Leftrightarrow \overline{\beta(n)} z\varphi_{n-1}^f(z), \end{array} \qquad n \in I\!N^*.$$

Les coefficients $\beta(n)$, $n \geq 1$, sont ainsi associés à dF, la variance $\beta(0)$ étant égale à la mesure, selon dF, de $[\Leftrightarrow\pi, \pi]$. La correspondance inverse est basée sur la transformation qui à la mesure dF associe la fonction Carathéodory :

$$G(z) = \int_{-\pi}^{\pi} \frac{e^{i\lambda} + z}{e^{i\lambda} \Leftrightarrow z} dF(\lambda), \quad |z| < 1.$$

En tout point de continuité λ de la fonction de répartition F, on a

$$F(\lambda) = \lim_{r \to 1^-} \int_{-\pi}^{\lambda} \mathcal{R}\left\{G(re^{i\theta})\right\} d\theta,$$

où $\mathcal{R}\{G(z)\}$ désigne la partie réelle de $G(z)$. Par ailleurs la fonction de Carathéodory est donnée, en fonction de $\beta(\cdot)$, par

$$G(z) = \beta(0) \lim_{n \to +\infty} \frac{\psi_n^b(z)}{\varphi_n^b(z)}, \quad |z| < 1,$$

où $\left\{\psi_n^f(e^{in\lambda}), n \in I\!N\right\}$ sont les polynômes associés à la séquence $\{\Leftrightarrow\beta(n), n \in I\!N^*\}$.

Après ce rappel sur le cas stationnaire, nous sommes maintenant en mesure de proposer une définition du spectre pour le cas non stationnaire. L'idée est de décrire à chaque instant t la dépendance entre $X(t)$ et son passé $\{X(t \Leftrightarrow k), k \geq 1\}$, préservant ainsi l'aspect causal, sous forme "stationnaire". En considérant la fonction d'autocovariance $R(t,s)$, il est naturel de retenir la variance de $X(t)$, $R_t(0) = R(t,t)$, et les corrélations avec le passé en posant $R_t(k) = R(t, t \Leftrightarrow k)R_t(0)/\sqrt{R_t(0)R_{t-k}(0)}$, $k \geq 1$. Cependant la fonction $R_t(\cdot)$ ainsi obtenue n'est pas nécessairement de type positif. La même démarche conduite sur la fonction d'autocorrélation partielle $\beta(t,s)$ permet d'aboutir sans difficultés. En effet il suffit de remplacer, dans la démarche précédente, la corrélation entre $X(t)$ et $X(t \Leftrightarrow k)$ par la corrélation partielle entre ces mêmes variables.

Définition 2.1.1. *Soit $\beta(\cdot,\cdot)$ la fonction d'autocorrélation partielle d'un processus $X(\cdot)$. On appelle spectre évolutif instantané de ce processus, la suite $\{dF_t, t \in Z\!\!Z\}$ de mesures sur $[\Leftrightarrow\pi, \pi]$ où, pour t fixé, dF_t est la mesure spectrale associée à la fonction d'autocorrélation partielle $\beta_t(\cdot)$ définie par,*

$$\beta_t(k) = \overline{\beta_t(\Leftrightarrow k)} = \beta(t, t \Leftrightarrow k), \quad k \geq 0.$$

Il est en effet immédiat de constater que les fonctions $\beta_t(\cdot)$ ainsi associées à tout élément de $\mathcal{D}_\beta$ constituent effectivement une suite de fonctions d'autocorrélation partielle de processus stationnaires au second ordre. Lorsque $\beta(\cdot,\cdot)$ présente une singularité d'ordre fini, $\beta(t,t) = 0$ ou $|\beta(t,s)| = 1$, la forme du domaine $\mathcal{D}_\beta$ montre que les mesures composant le spectre seront liées entre elles et ce de façon non descriptible en termes de mesures. En effet on ne sait pas caractériser l'ensemble des mesures dont la fonction d'autocorrélation partielle présente un ou plusieurs zéros isolés. Par contre, si on exclut ce type de singularité, le spectre évolutif instantané peut être constitué de n'importe quelle suite de mesures $\{dF_t, t \in \mathbb{Z}\!\!\mathbb{Z}\}$ non nulles et de support non fini.

Il est évident que deux fonctions d'autocorrélation partielle différentes définissent des spectres évolutifs instantanés différents. La donnée du spectre évolutif instantané d'un processus non stationnaire détermine donc sa fonction d'autocorrélation partielle et d'après la section 1.2, sa structure au second ordre.

On aurait pu définir le spectre évolutif de $X(\cdot)$ à partir de $\beta(t, t+k)$, $k \geq 0$, ce qui ne donnerait pas les mêmes résultats, mais il est plus réaliste de le définir de façon progressive, c'est-à-dire causale, plutôt que rétrograde.

Dans la suite de ce chapitre, lorsque nous considérons plusieurs processus, nous indexons leurs attributs (innovation, fonction d'autocorrélation partielle, etc.) par la lettre qui les désigne.

2.2 Propriétés du spectre évolutif instantané

Dans [Loy68], l'auteur propose deux listes de propriétés souhaitables pour un spectre qui dépend du temps, la première (A1 à A8) est plutôt de nature physique alors que la seconde (B1 à B12) est plutôt de nature mathématique. A partir de ces deux listes, nous présentons les propriétés du spectre évolutif instantané. Chacune d'entre elles sera suivie de sa ou ses références dans la classification de Loynes.

La première propriété est une simple conséquence de la définition du spectre.

Propriété 2.2.1. *Le spectre évolutif instantané est une fonction réelle et positive du temps et de la fréquence* (*Loynes : A1, B1, B8*).

En toute rigueur notre spectre est défini par une famille de mesures. Lorsque celles-ci sont absolument continues, le spectre s'identifie à la famille $\{f_t(\lambda) = dF_t(\lambda)/d\lambda, t \in \mathbb{Z}\!\!\mathbb{Z}\}$ des densités correspondantes. La propriété suivante re-

joint le commentaire qui suit la définition du spectre puisque $\beta(\cdot,\cdot)$ et $R(\cdot,\cdot)$ sont en correspondance biunivoque.

Propriété 2.2.2. *Le spectre évolutif instantané est en bijection avec la covariance du processus* (*Loynes* : *A4, B4*).

La relation $Var\{X(t)\} = \beta_t(0) = \int_{-\pi}^{\pi} dF_t(\lambda)$ donne :

Propriété 2.2.3. *Le spectre évolutif instantané représente une distribution de l'énergie relativement à la fréquence* (*Loynes* : *A2*).

Pour t fixé, la mesure dF_t est la transformée de Fourier de la fonction d'autocovariance d'un processus stationnaire qui admet $\beta_t(\cdot)$ comme fonction d'autocorrélation partielle.

Propriété 2.2.4. *Le spectre s'obtient comme la transformée de Fourier d'une "quantité apparemment sensée"* (*Loynes* : *A8*).

Le processus $X(\cdot)$ est stationnaire au second ordre si et seulement si $\beta(t,s)$ ne dépend que de $(t \Leftrightarrow s)$. Dans ce cas $\beta_t(\cdot)$, qui ne dépend plus de t, est la fonction d'autocorrélation partielle de $X(\cdot)$.

Propriété 2.2.5. *Si $X(\cdot)$ est un processus stationnaire, son spectre se ramène au spectre ordinaire* (*Loynes* : *A5, B5*).

Loynes souhaitait que le spectre d'un processus constitué d'une succession de parties stationnaires soit la succession des spectres stationnaires (Loynes : A6, B6) tout en soulignant l'impossibilité de réaliser de façon exacte une telle condition. Par contre on devrait retrouver les spectres de façon approximative. Pour un processus stationnaire $X(\cdot)$, de fonction d'autocorrélation partielle $\beta(\cdot)$, on désigne par $AR(\leq n)$ le modèle autorégressif d'ordre au plus n caractérisé par $\{\beta(0),\ldots,\beta(n)\}$ et on note dF_n sa mesure spectrale. On sait que le modèle $AR(\leq n)$ constitue une bonne approximation de la structure de $X(\cdot)$. En particulier dF_n et dF coïncident lorsque $X(\cdot)$ est singulier d'ordre d avec $d \leq n$ ou autorégressif d'ordre p avec $p \leq n$. On trouve dans [Ger60] les diverses propriétés de convergence des mesures dF_n vers la mesure dF de $X(\cdot)$. Lorsque $\beta(\cdot)$ est dans ℓ^1, les mesures sont absolument continues et il y a convergence uniforme des densités spectrales $f_n(\lambda) = dF_n(\lambda)/d\lambda$ vers $f(\lambda) = dF(\lambda)/d\lambda$. La vitesse de convergence dépend de celle avec laquelle $\beta(\cdot)$ tend vers zéro (cf. [Ger60] : Théorème 8.5, p. 139),

$$\prod_{k=n+1}^{+\infty} \frac{[1 \Leftrightarrow |\beta(k)|]}{[1+|\beta(k)|]} \leq \frac{f(\lambda)}{f_n(\lambda)} \leq \prod_{k=n+1}^{+\infty} \frac{[1+|\beta(k)|]}{[1 \Leftrightarrow |\beta(k)|]}.$$

Les bornes inférieures (resp. supérieures) étant atteintes en $\lambda = 0$ lorsque $\beta(\cdot)$ est réel constamment négatif (resp. positif), cet encadrement ne peut être amélioré. Si F est absolument continue au voisinage de λ_0 et que sa densité, nulle en λ_0, satisfait

$$|f(\lambda_0 + h) + f(\lambda_0 \Leftrightarrow h)| \leq Ch^\alpha, \quad \alpha > 0,$$

alors f_n vérifie (cf. [GER60] : Théorème 4.12, p. 61),

$$\lim_{n\to+\infty} \inf_{k\geq n} f_k(\lambda_0) = f(\lambda_0) = 0.$$

Lorsque F présente une masse en λ_0, on a $\lim_{n\to+\infty} f_n(\lambda_0) = +\infty$ et la masse en λ_0 est donnée par (cf. [GER60], p. 188),

$$F^d(\lambda_0) = 2\pi \left[\sum_{n=0}^{+\infty} \frac{1}{f_n(\lambda_0)}\right]^{-1}.$$

Proposition 2.2.1. *Soit $X(\cdot)$ un processus constitué d'une succession de parties stationnaires,*

$$X(t) = \begin{cases} X^{(1)}(t) & si\ t \leq 0 \\ X^{(2)}(t) & si\ t > 0 \end{cases},$$

où $X^{(1)}(\cdot)$ et $X^{(2)}(\cdot)$ sont deux processus stationnaires non corrélés. Alors on a

$$dF_t^X = \begin{cases} dF^{(1)} & si\ t \leq 0 \\ dF_{t-1}^{(2)} & si\ t > 0 \end{cases},$$

où $dF^{(1)}$ est la mesure spectrale de $X^{(1)}(\cdot)$ et où $dF_t^{(2)}$ désigne la mesure spectrale du modèle $AR(\leq t)$ sous-jacent à $X^{(2)}(\cdot)$.

Démonstration.- Pour $t \leq 0$ on a $\beta_t^X(\cdot) = \beta^{(1)}(\cdot)$ et pour $t > 0$,

$$\beta_t^X(n) = \begin{cases} \beta^{(2)}(n) & \text{si } 0 \leq n < t \\ 0 & \text{si } n \geq t \end{cases}.$$

Ces égalités sont évidentes pour $t \leq 0$ et pour $0 \leq n < t$. Pour $n \geq t > 0$, il est facile de vérifier que $\varepsilon_X^f(t; t \Leftrightarrow n+1) = \varepsilon_{X^{(1)}}^f(t;1)$ et $\varepsilon_X^b(t \Leftrightarrow n; t \Leftrightarrow 1) = \varepsilon_{X^{(2)}}^b(t \Leftrightarrow n; 0)$. Ces variables étant non corrélées, le résultat est acquis. $\Diamond$

Les propriétés suivantes établissent comment certaines transformations sur le processus $X(\cdot)$ se transposent sur le spectre.

Propriété 2.2.6. *Un décalage en temps du processus se répercute sur le spectre évolutif instantané sous forme d'un décalage temporel de même durée (Loynes : B10) :*

$$Y(t) = X(t+h) \Rightarrow dF_t^Y(\lambda) = dF_{t+h}^X(\lambda).$$

L'égalité $\beta^Y(t, t \Leftrightarrow k) = \beta^X(t+h, t+h \Leftrightarrow k)$, justifie ce résultat.

Propriété 2.2.7. *Une modulation du processus par une exponentielle complexe de pulsation λ_0 décale le spectre d'une quantité λ_0 :*

$$Y(t) = X(t)e^{-i\lambda_0 t} \Rightarrow dF_t^Y(\lambda) = dF_t^X(\lambda + \lambda_0),$$

où l'argument est exprimé modulo 2π (Loynes : B9).

Démonstration.- Les relations entre les erreurs,

$$\begin{array}{l} \varepsilon_Y^f(t; t \Leftrightarrow n+1) = e^{-i\lambda_0 t}\varepsilon_X^f(t; t \Leftrightarrow n+1) \\ \varepsilon_Y^b(t \Leftrightarrow n; t \Leftrightarrow 1) = e^{-i\lambda_0 (t-n)}\varepsilon_X^b(t \Leftrightarrow n; t \Leftrightarrow 1) \end{array} , \quad n \geq 1,$$

montrent que $\beta_t^Y(n) = e^{-i\lambda_0 n}\beta_t^X(n)$, $n \geq 0$. On en déduit, par récurrence, les relations entre les polynômes de Szegö correspondants ainsi que l'égalité de leurs normes :

$$\left\{ \begin{array}{l} \varphi_n^{Yf}(e^{i(\lambda-\lambda_0)}) = e^{-i\lambda_0 n}\varphi_n^{Xf}(e^{i\lambda}) \\ \varphi_n^{Yb}(e^{i(\lambda-\lambda_0)}) = \varphi_n^{Xb}(e^{i\lambda}) \\ \sigma_Y^f(n) = \sigma_X^f(n) \end{array} \right. , \quad n \geq 0,$$

où l'on omet l'indice t pour ne pas alourdir les notations. L'égalité des normes permet de regrouper les caractérisations des mesures,

$$\begin{aligned} \int_{-\pi}^{\pi} \varphi_n^{Xf}(e^{i\lambda})\overline{\varphi_m^{Xf}(e^{i\lambda})}dF_t^X(\lambda) &= \int_{-\pi}^{\pi} \varphi_n^{Yf}(e^{i\lambda})\overline{\varphi_m^{Yf}(e^{i\lambda})}dF_t^Y(\lambda) \\ &= \sigma_X^{f2}(n) \text{ si } n = m, \text{ 0 sinon.} \end{aligned}$$

En utilisant les relations entre les polynômes, on obtient :

$$\begin{aligned} \int_{-\pi}^{\pi} \varphi_n^{Xf}(e^{i\lambda})\overline{\varphi_m^{Xf}(e^{i\lambda})}dF_t^X(\lambda) &= \int_{-\pi}^{\pi} e^{i\lambda_0(n-m)}\varphi_n^{Yf}(e^{i(\lambda-\lambda_0)})\overline{\varphi_m^{Yf}(e^{i(\lambda-\lambda_0)})}dF_t^X(\lambda) \\ &= \int_{-\pi-\lambda_0}^{\pi-\lambda_0} \varphi_n^{Yf}(e^{i\lambda})\overline{\varphi_m^{Yf}(e^{i\lambda})}dF_t^X(\lambda+\lambda_0) \\ &= \int_{-\pi}^{\pi} \varphi_n^{Yf}(e^{i\lambda})\overline{\varphi_m^{Yf}(e^{i\lambda})}dF_t^Y(\lambda). \end{aligned}$$

On a éliminé le terme $e^{i\lambda_0(n-m)}$ car l'intégrale est nulle lorsque m est différent de n. La dernière égalité, considérée pour tous les couples (n, m), équivaut à la relation annoncée entre les deux mesures. $\diamondsuit$

Propriété 2.2.8. *L'effet de la conjugaison et le cas réel se traduisent par (Loynes : B11 c et d) :*

(i) $\forall t \in \mathbb{Z}$, $Y(t) = \overline{X(t)} \Rightarrow \forall t \in \mathbb{Z}$, $dF_t^Y(\lambda) = dF_t^X(\Leftrightarrow\lambda)$

(ii) $\forall t \in \mathbb{Z}$, $X(t) \in \mathbb{R} \Rightarrow \forall t \in \mathbb{Z}$, $dF_t^X(\lambda) = dF_t^X(\Leftrightarrow\lambda)$.

Démonstration.- Lorsque $Y(\cdot) = \overline{X(\cdot)}$, on a $\beta_t^Y(\cdot) = \overline{\beta_t^X(\cdot)}$. Les relations entre les polynômes sont obtenues par récurrence et les normes sont égales :

$$\left\{ \begin{array}{l} \varphi_n^{Yf}(\overline{z}) = \overline{\varphi_n^{Xf}(z)} \\ \varphi_n^{Yb}(\overline{z}) = \overline{\varphi_n^{Xb}(z)} \\ \sigma_Y^f(n) = \sigma_X^f(n) \end{array} \right. , \qquad n \geq 0.$$

Dans ce cas la caractérisation des mesures, compte tenu de l'égalité des normes, conduit à

$$\int_{-\pi}^{\pi} \varphi_n^{Xf}(e^{i\lambda}) \overline{\varphi_m^{Xf}(e^{i\lambda})} dF_t^X(\Leftrightarrow\lambda) = \int_{-\pi}^{\pi} \varphi_n^{Yf}(e^{i\lambda}) \overline{\varphi_m^{Yf}(e^{i\lambda})} dF_t^Y(\lambda),$$

pour tous les couples (n, m). Ceci prouve l'assertion (i). Le cas réel (ii) est une conséquence de (i) mais s'obtient immédiatement en remarquant que $\beta_t^X(\cdot)$ est réel, ce qui équivaut à la symétrie de dF_t^X. $\diamondsuit$

Les assertions a) et b) de la Propriété B11 considérée par Loynes sont liées au retournement du temps. Elle ne sont pas satisfaites ici, sauf dans des situations très particulières. Elles sont équivalentes, en vertu du point (i) de la Propriété 2.2.8, à

(a) $\forall t \in \mathbb{Z}, Y(t) = \overline{X(\Leftrightarrow t)} \Rightarrow \forall t \in \mathbb{Z}, dF_t^Y(\lambda) = dF_{-t}^X(\lambda)$,

(b) $\forall t \in \mathbb{Z}, Y(t) = X(\Leftrightarrow t) \Rightarrow \forall t \in \mathbb{Z}, dF_t^Y(\lambda) = dF_{-t}^X(\Leftrightarrow\lambda)$.

La définition de $Y(\cdot)$ dans (a) se traduit par $\beta_t^Y(k) = \beta_{-t+k}^X(k)$ pour tout $t \in \mathbb{Z}$ et tout $k \geq 0$. La condition $dF_t^Y(\lambda) = dF_{-t}^X(\lambda)$ équivaut à $\beta_t^Y(k) = \beta_{-t}^X(k)$, $k \geq 0$. Ces propriétés seront donc satisfaites si et seulement si $\beta_t^X(k) = \beta_{t+k}^X(k)$ pour tout $t \in \mathbb{Z}$ et tout $k \geq 0$. C'est vrai en particulier dans le cas stationnaire.

Enfin nous terminons ce paragraphe en commentant deux propriétés qui ne sont pas du tout en phase avec le spectre que nous avons proposé :

(i) *Le spectre est une transformation linéaire de la covariance* (*Loynes* : *B2*).

(ii) *Le spectre d'un processus transformé linéairement doit se déduire, de façon simple, de celui du processus original* (*Loynes* : *A3, B3*).

Ces propriétés concernent des opérations sur les processus qui se traduisent facilement en termes de fonctions d'autocovariance. On pense à la somme de deux processus non corrélés dans le cas (i) ou aux opérations de filtrage dans le cas (ii). Il est facile de se convaincre, à l'aide d'exemples simples, que les effets sur le spectre évolutif instantané sont immédiatement très complexes. Le seul point positif est qu'ils sont "calculables" puisque le spectre est en bijection avec la fonction d'autocovariance.

Pour être complet, il faudrait évoquer que le spectre dépend continuement de la covariance (loynes : B12) et considérer le problème de son estimation (loynes : A7, B7). Le premier point demande de préciser le type de continuité, mais ne devrait pas présenter de difficulté compte tenu des relations très claires entre $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$ d'une part et entre $\beta_t(\cdot)$ et dF_t d'autre part. En ce qui concerne l'estimation, il est clair que le problème n'a de sens que pour des classes bien particulières de processus non stationnaires dans la mesure ou l'on observe qu'une seule trajectoire, même sur un temps infini.

2.3 Comparaison avec le spectre évolutif

Dans [MÉL78], après avoir introduit le spectre évolutif, l'auteur présente les propriétés que ce spectre satisfait parmi celles établies par Loynes. Nous proposons de rappeler la démarche qui conduit à définir ce spectre et de le comparer au spectre évolutif instantané à travers les propriétés de Loynes. Notons que ce spectre a également été introduit de façon différente par Tjøstheim [TJØ76].

Considérons un processus $X(\cdot)$ purement indéterminable. Alors la décomposition de Wold-Cramér de ce processus (cf. Théorème 1.2.1) exprime, pour tout t, le développement de $X(t)$ dans la base orthonormée de $\mathcal{M}(\Leftrightarrow\infty;t)$ constituée par les innovations normées $\eta(u)\neq 0$, $u\leq t$:

$$X(t)=\sum_{j=0}^{+\infty} h(t,t\Leftrightarrow j)\eta(t\Leftrightarrow j), \tag{2.1}$$

$(h(t,t\Leftrightarrow j)=b(t,t\Leftrightarrow j)\sigma_\varepsilon(t\Leftrightarrow j))$. On pose $h(t,t\Leftrightarrow j)=0$ lorsque $\eta(t\Leftrightarrow j)=0$ afin de rendre cette décomposition unique. Le processus $X(\cdot)$ admet alors la décomposition de Karhunen sur l'espace de Lebesgue $([\Leftrightarrow\pi,\pi],\mathcal{B},d\lambda)$ suivante :

$$X(t)=\int_{-\pi}^{\pi}\psi_t(\lambda)d\xi(\lambda),$$

où $\psi_t(\lambda)=(2\pi)^{-1/2}\sum_{j=0}^{+\infty}h(t,t\Leftrightarrow j)e^{i\lambda(t-j)}$ et $\xi(\cdot)$ est un processus indexé sur $[\Leftrightarrow\pi,\pi]$, à accroissement orthogonaux, continu à droite en moyenne quadratique, de mesure associée la mesure de Lebesgue, c'est-à-dire telle que

$$E\left\{[\xi(\mu)\Leftrightarrow\xi(\nu)]\overline{[\xi(\mu')\Leftrightarrow\xi(\nu')]}\right\}=\int_{]\nu,\mu]\cap]\nu',\mu']}d\lambda$$

pour tout $\nu\leq\mu$ et $\nu'\leq\mu'$ dans $[\Leftrightarrow\pi,\pi]$. Une telle représentation est analogue à la représentation spectrale d'un processus stationnaire. La double orthogonalité n'est pas conservée car les exponentielles, comme fonctions de décomposition, sont remplacée par $\psi_t(\lambda)$. La représentation de la fonction

d'autocovariance, qui se déduit de celle du processus, est de la forme :

$$R(t,s) = \int_{-\pi}^{\pi} \psi_t(\lambda)\overline{\psi_s(\lambda)}d\lambda.$$

Le spectre évolutif $f_t(\lambda)$ est alors défini de la façon suivante :

$$f_t(\lambda) = |\psi_t(\lambda)|^2 = \frac{1}{2\pi}\left|\sum_{j=0}^{+\infty} h(t,t \Leftrightarrow j)e^{-i\lambda j}\right|^2 .$$

Le spectre évolutif instantané est défini directement à partir de la fonction d'autocorrélation partielle. L'existence d'une représentation du processus $X(\cdot)$ liée à ce spectre est un point difficile qui ne sera pas abordé ici.

Soulignons le fait que le spectre évolutif instantané est défini pour n'importe quel processus alors que celui proposé par Mélard ne l'est que pour les processus purement indéterminables. Pour cette classe particulière de processus, les propriétés établies par Loynes sont dans l'ensemble vérifiées ou infirmées simultanément par ces deux spectres. Le spectre évolutif est une fonction réelle non négative de (t,λ) et coïncide avec la densité spectrale lorsque le processus est stationnaire (Loynes : A1, B8, A5). Il est la transformée de Fourier d'une "quantité apparemment sensée" et satisfait (Loynes : A8, A2)

$$Var\,\{X(t)\} = \int_{-\pi}^{\pi} f_t(\lambda)d\lambda.$$

Les deux spectres se comportent de la même façon vis à vis des transformations élémentaires sur le processus $X(\cdot)$ (Loynes : B9, B10, B11). En ce qui concerne l'effet d'un filtrage (Loynes : A3), Mélard [MÉL78] obtient un résultat approché dans une situation très particulière. Nous avons déjà souligné la difficulté de ce problème dans la section précédente, qui d'ailleurs n'est pas sans rapport avec celui de la représentation du processus.

Une différence importante entre les deux définitions est que, même restreint aux processus purement indéterminables, le spectre évolutif de Mélard n'est pas en bijection avec la fonction d'autocovariance (Loynes : A4). Plus précisément, deux structures au second ordre différentes peuvent donner lieu à un même spectre évolutif. Reprenons le contre-exemple de [MÉL78] : Soit un processus $X(\cdot)$ défini par

$$X(t) = \begin{cases} \varepsilon(t) \Leftrightarrow 0.5\varepsilon(t \Leftrightarrow 1) & t \neq 1 \\ 0.5\varepsilon(1) \Leftrightarrow \varepsilon(0) & t = 1 \end{cases},$$

où $\varepsilon(\cdot) = \{\varepsilon(t), t \in \mathbb{Z}\}$ est un processus constitué de variables aléatoires non corrélées de variance unité. Alors ce processus est le processus d'innovation normé et le spectre évolutif est indépendant du temps,

$$f_t(\lambda) = \frac{1}{2\pi}\left|1 \Leftrightarrow 0.5e^{-i\lambda}\right|^2 .$$

Le processus $X(\cdot)$ présente le même spectre évolutif que celui du processus $Y(\cdot)$ stationnaire $MA(1)$, défini par $Y(t) = \varepsilon(t) \Leftrightarrow 0.5\varepsilon(t \Leftrightarrow 1)$. Une conséquence immédiate de ce contre-exemple est qu'un spectre évolutif indépendant du temps n'assure pas la stationnarité du processus considéré. Rappelons que la correspondance entre la classe des spectres évolutifs instantanés et celle des fonctions d'autocovariance est biunivoque. En particulier, un spectre évolutif instantané indépendant du temps est équivalent à une fonction $\beta_t(\cdot)$ indépendante de t, et donc au fait que le processus considéré est stationnaire.

Le dernier point que nous voulons évoquer concerne les processus stationnaires par morceaux. Soit un processus $X(\cdot)$ défini par

$$X(t) = \begin{cases} X^{(1)}(t) & \text{si } t \leq 0 \\ X^{(2)}(t) & \text{si } t > 0 \end{cases},$$

où $X^{(1)}(\cdot)$ et $X^{(2)}(\cdot)$ sont deux processus stationnaires. Comme pour le spectre évolutif instantané, la fonction $f_t(\cdot)$ coïncide avec la densité spectrale du premier processus lorsque le temps t est négatif. Pour t positif, elle coïncide avec la deuxième densité si les deux processus stationnaires possèdent le même processus d'innovation normé $\eta(\cdot)$. En effet sous cette hypothèse, $\eta(\cdot)$ est aussi le processus d'innovation du processus $X(\cdot)$ et les coefficients de la décomposition de Wold-Cramér de $X(t)$, $t > 0$ (i.e. ceux de la projection de $X(t)$ sur l'espace $\overline{\mathcal{L}}\{\eta(s), s \leq t\}$) sont ceux du processus stationnaire $X^{(2)}(\cdot)$. Ceci n'est pas vrai pour le spectre évolutif instantané car en général $\beta_t(\cdot)$ n'est pas déterminé par $h(t, t \Leftrightarrow \cdot)$. Supposons maintenant que les deux processus stationnaires soient non corrélés. En situation générale il n'existe pas de résultat analogue à l'approximation AR que nous avons obtenue (cf. Proposition 2.2.1). De plus si le processus $X^{(2)}(\cdot)$ est autorégressif d'ordre p, alors pour le spectre de Mélard on retrouve la densité spectrale correspondante au bout d'un temps infini. Ce temps est fini, égal à $p + 1$, pour le spectre évolutif instantané.

C'est entre autre pour cette raison que Grenier [GRE84] a voulu améliorer le spectre de Mélard en proposant un spectre plus instantané, dans le cas particulier où les processus considérés peuvent être représentés par un modèle plus local qu'une moyenne mobile d'ordre infini. Ce spectre, appelé spectre rationnel, fait l'objet de la section suivante.

2.4 Comparaison avec le spectre rationnel

Dans [GRE84], l'auteur propose trois variantes de spectre rationnel, selon qu'il considère un modèle $ARMA(p, q)$, un modèle d'état observable ou commandable. Il rejette le spectre associé au modèle d'état commandable car

il n'est pas toujours possible de le calculer à partir de la seule connaissance d'un modèle $ARMA$ du processus. C'est pourquoi cette variante ne sera pas traitée.

Le spectre rationnel est défini pour un processus purement indéterminable qui admet la représentation $ARMA(p,q)$

$$\sum_{j=0}^{p} a_j(t \Leftrightarrow j)X(t \Leftrightarrow j) = \sum_{j=0}^{q} b_j(t \Leftrightarrow j)\eta(t \Leftrightarrow j), \tag{2.2}$$

dans laquelle $\eta(\cdot)$ est le processus d'innovation normé. Cette représentation est possible si et seulement si il existe deux entiers naturels p et q, et p fonctions $a_j(t)$, $j = 1, \ldots, p$ tels que les coefficients de la décomposition (2.1) satisfassent :

$$\sum_{j=0}^{\min(n,p)} a_j(t \Leftrightarrow j)h(t \Leftrightarrow j, t \Leftrightarrow n) = 0, \text{ pour } n > q, \text{ avec } a_0(t) = 1.$$

Les paramètres $b_j(t)$, $j = 0, \ldots, q$, sont alors définis par :

$$b_j(t) = \sum_{k=0}^{\min(n,p)} a_k(t + j \Leftrightarrow k)h(t + j \Leftrightarrow k, t).$$

Le processus $X(\cdot)$ admet une représentation sous forme d'état observable équivalente au modèle $ARMA$ (2.2). En effet, introduisons le vecteur $Y(t) = \{Y_j(t), j = 1, \ldots, n\}$, $n = \max(p, q+1)$, défini par $Y_1(t) = X(t)$ et pour $j = 2, \ldots, n$:

$$Y_j(t) = \sum_{k=0}^{j-1} a_k(t+j \Leftrightarrow k \Leftrightarrow 1)X(t+j \Leftrightarrow k \Leftrightarrow 1) \Leftrightarrow \sum_{k=0}^{j-2} b_k(t+j \Leftrightarrow k \Leftrightarrow 1)\eta(t+j \Leftrightarrow k \Leftrightarrow 1),$$

où $a_k(t) = 0$ si $k > p$ et $b_k(t) = 0$ si $k > q$. En remarquant que

$$Y_j(t) = \Leftrightarrow a_j(t \Leftrightarrow 1)Y_1(t \Leftrightarrow 1) + Y_{j+1}(t \Leftrightarrow 1) + b_{j-1}(t \Leftrightarrow 1)\eta(t),$$

on obtient

$$Y(t) = \begin{pmatrix} \Leftrightarrow a_1(t \Leftrightarrow 1) & 1 & & (0) \\ \Leftrightarrow a_2(t \Leftrightarrow 1) & & \ddots & \\ \vdots & (0) & & 1 \\ \Leftrightarrow a_n(t \Leftrightarrow 1) & 0 & \ldots & 0 \end{pmatrix} Y(t \Leftrightarrow 1) + \begin{pmatrix} b_0(t) \\ b_1(t) \\ \vdots \\ b_{n-1}(t) \end{pmatrix} \eta(t), \tag{2.3}$$

et $X(t) = [1, 0, \ldots, 0]^T Y(t)$.

Ces deux représentations conduisent à deux versions possibles pour le spectre de Grenier. Pour t fixé, le spectre rationnel $\rho(t, \lambda)$ du processus $X(\cdot)$ est défini comme la densité spectrale d'un processus "stationnaire" qui admet

la représentation $ARMA$ (2.2) ou la forme observable (2.3) :

$$\rho(t,\lambda) = \frac{1}{2\pi}\left[\frac{B(t,z)B(t,z^{-1})}{A(t,z)A(t,z^{-1})}\right]_{z=e^{i\lambda}}$$

avec

$$A(t,z) = \sum_{j=0}^{p} a_j(t \Leftrightarrow j) z^{-j}, \quad B(t,z) = \sum_{j=0}^{q} b_j(t \Leftrightarrow j) z^{-j},$$

pour le modèle $ARMA(p,q)$ et

$$A(t,z) = \sum_{j=0}^{n} a_j(t \Leftrightarrow 1) z^{-j}, \quad B(t,z) = \sum_{j=0}^{n-1} b_j(t) z^{-j},$$

pour la forme observable.

La constante $(2\pi)^{-1}$ n'apparaît pas dans la définition originale du spectre rationnel. Nous l'avons introduite afin d'obtenir une définition cohérente avec celle de Mélard et celle que nous proposons. Notons que le mot stationnaire ci-dessus est entre guillemets car pour t fixé, rien ne permet d'affirmer que les paramètres $a_j(t \Leftrightarrow j)$, $j = 1, \ldots, p$ et $b_j(t \Leftrightarrow j)$, $j = 0, \ldots, q$, soient ceux d'un modèle $ARMA$ stationnaire. En particulier, le polynôme $A(t,\cdot)$ peut avoir ses racines sur le cercle unité (cf. sous-section 2.5.1).

Nous avons établi à la section 1.4 que le processus $X(\cdot)$ admet une représentation de la forme (2.2) avec $q = 0$ si et seulement si pour chaque t, la fonction $\beta_t(\cdot)$ est nulle au delà de p. Le spectre évolutif instantané d'un tel processus est donc la mesure spectrale d'un processus stationnaire autorégressif. Mais les paramètres $a_j(t \Leftrightarrow j)$, $j = 1, \ldots, p$ ne sont pas nécessairement ceux du modèle stationnaire comme le montre l'exemple suivant. Soit $X(\cdot)$ un processus réel tel que

$$X(t) + a(t \Leftrightarrow 1) X(t \Leftrightarrow 1) = \eta(t),$$

où $\eta(\cdot)$ est le processus d'innovation normé. Sa fonction d'autocovariance vérifie

$$R(t,t) = 1 + a^2(t \Leftrightarrow 1) R(t \Leftrightarrow 1, t \Leftrightarrow 1),$$

$$R(t, t \Leftrightarrow 1) = \Leftrightarrow a(t \Leftrightarrow 1) R(t \Leftrightarrow 1, t \Leftrightarrow 1).$$

L'Algorithme LDG permet de déterminer les premiers coefficients non nuls de la fonction d'autocorrélation partielle :

$$\begin{aligned} \beta_t(0) &= 1 + a^2(t \Leftrightarrow 1) R(t \Leftrightarrow 1, t \Leftrightarrow 1), \\ \beta_t(1) &= \frac{\Leftrightarrow a(t \Leftrightarrow 1) R(t \Leftrightarrow 1, t \Leftrightarrow 1)^{\frac{1}{2}}}{[1 + a^2(t \Leftrightarrow 1) R(t \Leftrightarrow 1, t \Leftrightarrow 1)]^{\frac{1}{2}}}. \end{aligned}$$

Les paramètres du modèle $AR(1)$ stationnaire associé à $\{\beta_t(0), \beta_t(1)\}$ sont

donnés par

$$\alpha_t(1) = \Leftrightarrow\beta_t(1) \quad \text{et} \quad \sigma^2_{\varepsilon,t} = \left[1 \Leftrightarrow \beta^2_t(1)\right] \beta_t(0) = \sigma^2_\varepsilon(t) = 1.$$

Pour t fixé, le spectre évolutif instantané est la mesure spectrale d'un modèle $AR(1)$ stationnaire de paramètres $\{\alpha_t(1); 1\}$. Et les quantités $\alpha_t(1)$ et $a(t \Leftrightarrow 1)$ ne sont pas nécessairement identiques. Pour ce processus particulier, il n'existe pas de différence entre les deux variantes du spectre rationnel : $\rho(t, \lambda)$ est la densité spectrale d'un modèle $AR(1)$ de paramètres $\{a(t \Leftrightarrow 1); 1\}$. Dans le cas où $\alpha_t(1) \neq a(t \Leftrightarrow 1)$, le spectre rationnel et celui que nous proposons sont distincts.

D'autre part, notons que le spectre évolutif instantané d'un processus qui admet une représentation de la forme (2.2) avec $p = 0$, ne s'identifie pas nécessairement à la mesure spectrale d'un modèle moyenne-mobile du même ordre. En effet, considérons un processus $X(\cdot)$ réel non stationnaire tel que

$$X(t) = \eta(t) + b(t \Leftrightarrow 1)\eta(t \Leftrightarrow 1),$$

où $\eta(\cdot) = \{\eta(t), t \in \mathbb{Z}\}$ est une suite de variables aléatoires non corrélées telles que pour tout t, $\sigma^2_\eta(t) = 1$. La fonction $R(\cdot, \cdot)$ satisfait alors

$$\begin{aligned} R(t,t) &= 1 + b^2(t \Leftrightarrow 1), \\ R(t, t \Leftrightarrow 1) &= b(t \Leftrightarrow 1), \\ R(t, t \Leftrightarrow k) &= 0, \quad \forall k \geq 2. \end{aligned}$$

Les coefficients $\beta_t(j)$, $j = 0, 1, 2$, sont déterminés à l'aide de l'Algorithme LDG :

$$\begin{aligned} \beta_t(0) &= 1 + b^2(t \Leftrightarrow 1) \\ \beta_t(1) &= \frac{b(t \Leftrightarrow 1)}{\left\{\left[1 + b^2(t \Leftrightarrow 1)\right]\left[1 + b^2(t \Leftrightarrow 2)\right]\right\}^{\frac{1}{2}}}, \\ \beta_t(2) &= \frac{\Leftrightarrow b(t \Leftrightarrow 1) b(t \Leftrightarrow 2)}{\left\{1 + b^2(t \Leftrightarrow 2)\left[1 + b^2(t \Leftrightarrow 1)\right]\right\}^{\frac{1}{2}} \left\{1 + b^2(t \Leftrightarrow 3)\left[1 + b^2(t \Leftrightarrow 2)\right]\right\}^{\frac{1}{2}}}. \end{aligned}$$

Soit $R_t(\cdot)$ la fonction d'autocovariance d'un processus stationnaire qui admet $dF_t(\cdot)$ comme mesure spectrale, l'algorithme de Levinson-Durbin permet d'évaluer la quantité $R_t(2)$:

$$\begin{aligned} R_t(2) &= \left\{\beta_t(2)\left[1 - \beta^2_t(1)\right] + \beta^2_t(1)\right\} \beta_t(0) \\ &= \frac{b(t-1)}{\left[1 + b^2(t-2)\right]} \left\{ b(t-1) - b(t-2) \frac{\left\{1 + b^2(t-2)\left[1 + b^2(t-1)\right]\right\}^{\frac{1}{2}}}{\left\{1 + b^2(t-3)\left[1 + b^2(t-2)\right]\right\}^{\frac{1}{2}}} \right\}. \end{aligned}$$

Selon le choix des paramètres $b(\cdot)$, la quantité $R_t(2)$ ne s'annule pas. Dans ce cas le spectre évolutif instantané du processus $X(\cdot)$ n'est pas, à chaque instant t, la mesure spectrale d'un modèle $MA(1)$ stationnaire. A fortiori,

ce spectre ne coïncide pas avec celui de Grenier.

Suite aux remarques précédentes, le spectre rationnel et celui que nous proposons sont clairement différents. Nous allons établir ce qui les différencie quant aux propriétés de Loynes. Rappelons que Grenier à introduit son spectre afin d'améliorer celui de Mélard en ce qui concerne les processus stationnaires par morceaux. Des trois spectres, celui de Grenier semble être le mieux adapté à une telle situation. En effet, ce spectre est la succession correspondante des densités spectrales lorsque les processus stationnaires possèdent le même processus d'innovation. Lorsque ces derniers sont non corrélés, le spectre rationnel est la succession des densités spectrales excepté sur une zone transitoire de durée égale à l'ordre $n = \max(p, q+1)$. Notons que ce dernier résultat reste vrai pour notre spectre dans la situation autorégressive, sinon il représente la succession des approximations des mesures spectrales par celles des modèles autorégressifs sous-jacents.

Cette amélioration obtenue par le spectre rationnel se fait au détriment de la Propriété A2 de Loynes qui exige qu'il représente une distribution de l'énergie relativement à la fréquence. Reprenons le contre-exemple exhibé dans [Gre84]. Soit un processus $X(\cdot)$ qui admet une représentation du type (2.2) avec $q = 0$:

$$X(t) + a_1(t \Leftrightarrow 1)X(t \Leftrightarrow 1) = \eta(t),$$

$$a_1(t) = \begin{cases} \Leftrightarrow\alpha & \text{si } t < 0 \text{ ou } t > T \\ \Leftrightarrow\frac{1}{\alpha} & \text{si } t \in [0, T] \end{cases},$$

où α est une constante réelle telle que $|\alpha| < 1$. La variance de la variable $X(t)$ est déterminée par

$$Var\{X(t)\} = \begin{cases} \frac{1}{1-\alpha^2} & t \leq 0 \\ \frac{\alpha^2(\alpha^{-2(t+1)} + \alpha^{-2t} - 1)}{1-\alpha^2} & 1 \leq t \leq T+1 \\ \frac{1+\alpha^{2(t-T-1)}(\alpha^{-2(T+1)} + \alpha^{-2T} - \alpha^2 - 1)}{1-\alpha^2} & t \geq T+2 \end{cases}.$$

Ces dernières expressions montrent que la variance du processus dépend du temps sur les deux intervalles contenus dans $[1, +\infty[$. Par conséquent elle ne peut pas être égale à la quantité $\int_{-\pi}^{\pi} \rho(t, \lambda) d\lambda$ qui est constante sur les domaines $[1, T+1]$ et $[T+2, +\infty[$.

Hormis ces points, le spectre de Grenier conserve les propriétés de celui de Mélard. Les commentaires sur ce spectre sont essentiellement les mêmes que ceux mentionnés dans la section précédente pour le spectre évolutif. En particulier, il n'y a toujours pas de bijection entre la classe des spectres et celle des fonctions d'autocovariance.

2.5 Exemples

Nous illustrons dans cette section le spectre évolutif instantané à travers quelques exemples simples. Lorsque cela est possible, nous le comparons avec le spectre évolutif de Mélard et le spectre rationnel de Grenier.

2.5.1 Discrétisation du mouvement Brownien

Soit un mouvement Brownien standard $\left\{Y(t), t \in I\!R^+\right\}$. Ce processus est centré et sa structure au second ordre est caractérisée par :

$$E\left\{Y(t)Y(s)\right\} = \inf(t,s).$$

En particulier, il est nul presque sûrement à l'instant 0. Nous considérons le processus $X(\cdot)$ à temps discret obtenu par discrétisation de $Y(\cdot)$,

$$X(t) = \begin{cases} 0 \ p.s. & \text{si } t \leq 0 \\ Y(th) & \text{sinon} \end{cases},$$

où h est le pas de discrétisation. La fonction d'autocovariance $R(\cdot,\cdot)$ de $X(\cdot)$ est donnée par $R(t,s) = \sup(0, h\inf(t,s))$. La marche aléatoire définie par

$$\sum_{k=0}^{t} \varepsilon(k), \text{ pour } t > 0, \text{ nulle } p.s. \text{ sinon,}$$

où $\varepsilon(\cdot)$ est un bruit blanc de variance h, possède la même structure au second ordre que le processus $X(\cdot)$. Ce dernier admet ainsi la représentation autorégressive suivante :

$$X(t) \Leftrightarrow X(t \Leftrightarrow 1) = \varepsilon(t),$$

où $\varepsilon(\cdot)$ est le processus d'innovation de variance $\sigma_\varepsilon^2(t) = h$ pour $t > 0$, 0 sinon. La fonction d'autocorrélation partielle $\beta_t(\cdot)$ du processus $X(\cdot)$ est nulle au delà de 1 (cf. Proposition 1.4.1) et ses deux premiers coefficients sont donnés par

$$\begin{aligned} \beta_t(0) &= \sup(0, ht) \\ \beta_t(1) &= \frac{R(t, t \Leftrightarrow 1)}{[R(t \Leftrightarrow 1, t \Leftrightarrow 1)R(t,t)]^{\frac{1}{2}}} = \sqrt{\frac{t \Leftrightarrow 1}{t}} \text{ si } t > 0, \ 0 \text{ sinon.} \end{aligned}$$

Le spectre évolutif instantané de $X(\cdot)$ est nul pour $t \leq 0$ et, pour $t > 0$, est égal à la mesure spectrale du modèle stationnaire $AR(1)$ de paramètres

$$\alpha_t(1) = \Leftrightarrow \beta_t(1) = \Leftrightarrow \sqrt{\frac{t \Leftrightarrow 1}{t}}$$

$$\sigma_{\varepsilon,t}^2 = \left[1 \Leftrightarrow |\beta_t(1)|^2\right] \beta_t(0) = \sigma_\varepsilon^2(t) = h,$$

c'est-à-dire

$$dF_t(\lambda) = \frac{h}{2\pi}\left|1 \Leftrightarrow \sqrt{\frac{t \Leftrightarrow 1}{t}} e^{-i\lambda}\right|^{-2} d\lambda.$$

Notons que le pôle, situé en $\sqrt{(t \Leftrightarrow 1)/t}$, tend vers 1 lorsque t tend vers l'infini et que la mesure spectrale normalisée correspondante, qui ne dépend plus de h, tend alors vers la mesure de Dirac en $\lambda = 0$ avec une densité spectrale qui converge uniformément vers zéro sur tout compact ne contenant pas l'origine.

Le processus $X(\cdot)$ est purement indéterminable et sa décomposition de Wold-Cramér est égale à

$$X(t) = \sum_{j=0}^{t-1} h^{\frac{1}{2}} \eta(t \Leftrightarrow j) \text{ si } t > 0,\ 0 \text{ sinon.}$$

Le spectre évolutif est nul pour $t \leq 0$, sinon il est donné par

$$f_t(\lambda) = \frac{h}{2\pi}\left|\sum_{j=0}^{t-1} e^{-i\lambda j}\right|^2 = \frac{h}{2\pi}\frac{\left|1 \Leftrightarrow e^{-i\lambda t}\right|^2}{\left|1 \Leftrightarrow e^{-i\lambda}\right|^2} \text{ si } \lambda \neq 0,\ \frac{ht^2}{2\pi} \text{ si } \lambda = 0.$$

Il correspond à la mesure spectrale d'une moyenne mobile d'ordre $t \Leftrightarrow 1$ dont tous les zéros sont équirépartis sur le cercle unité : $z = e^{2ik\pi/t}$, $k = 1, \ldots, t \Leftrightarrow 1$ (cf. Figure 2.2). Le spectre normalisé a les mêmes propriétés que le précédent, lorsque t tend vers l'infini, mais la densité spectrale est très irrégulière. Dans ce cas particulier les valeurs de la fonction d'autocovariance $R_t(\cdot)$ associée à $f_t(\lambda)$ coïncident avec les covariances entre $X(t)$ et son passé : $R_t(k) = R(t, t \Leftrightarrow k)$, $k \geq 0$ et $R_t(k) = (t \Leftrightarrow k)h$ si $0 \leq k < t$, 0 si $k \geq t$. Ce n'est plus vrai en terme d'autocorrélations. Ceci montre que le spectre associé à $R_t(\cdot)$, lorsqu'il existe, peut être très différent du spectre évolutif instantané défini par $\beta_t(\cdot)$.

Pour t strictement positif, le processus $X(\cdot)$ admet la représentation $ARMA(0,1)$ de la forme (2.2) avec $a_1(t) = \Leftrightarrow 1$ et $b_0(t) = h^{1/2}$. Les deux versions du spectre rationnel sont identiques, nulles pour t négatif et pour $t > 0$,

$$\rho(t, \lambda) = \frac{h}{2\pi}\left|1 \Leftrightarrow e^{-i\lambda}\right|^{-2}.$$

Ce spectre illustre deux défauts du spectre rationnel : il ne dépend pas du temps, alors que le processus n'est pas stationnaire, et ne fournit pas la variance de $X(t)$ car cette "densité spectrale" n'est pas intégrable. Notons qu'il correspond à la limite de la densité spectrale du spectre évolutif instantané (sans normalisation) lorsque t tend vers l'infini. Les représentations graphiques ci-après illustrent le comportement de ces trois spectres. (Le spectre rationnel et le spectre évolutif instantané apparaissent confondus dans la figure 2.2.)

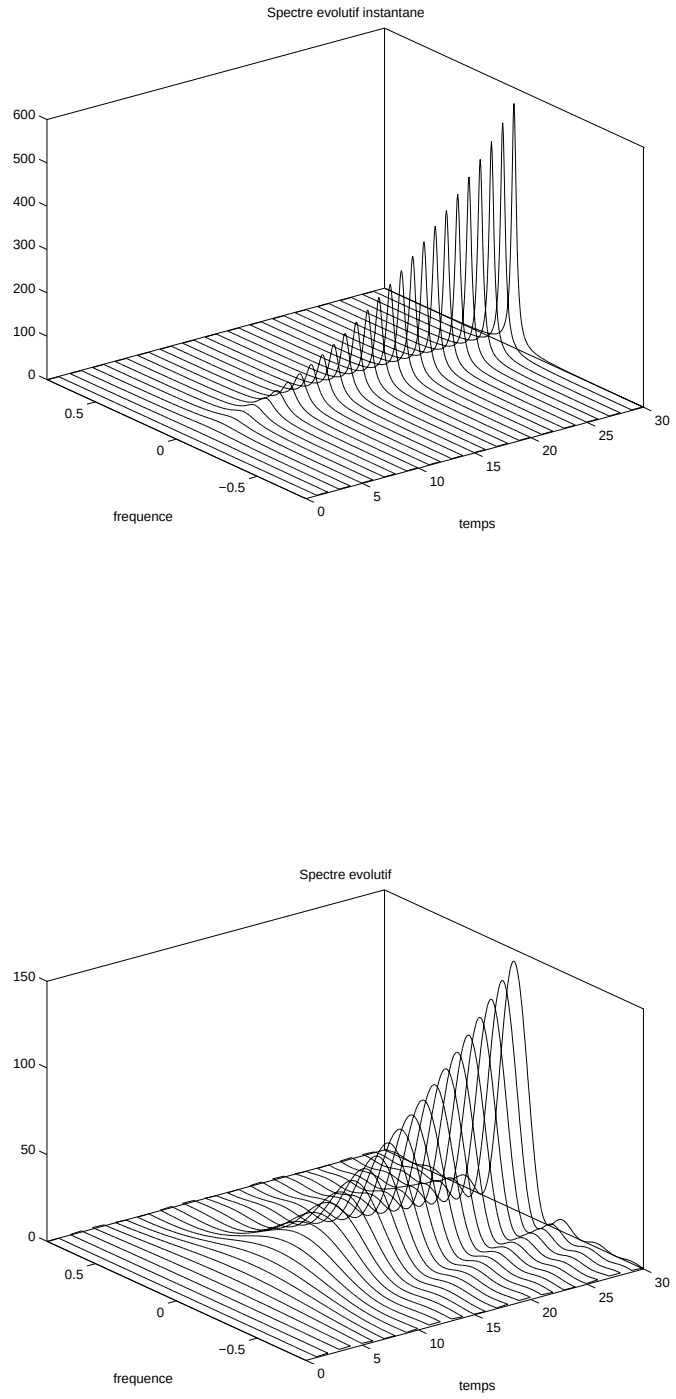


Figure 2.1 : *Représentation en trois dimensions du spectre évolutif instantané et du spectre évolutif,* $t \in [0, \ldots, 30]$, $\lambda \in [\Leftrightarrow 0.8, 0.8]$, $h = 1$.

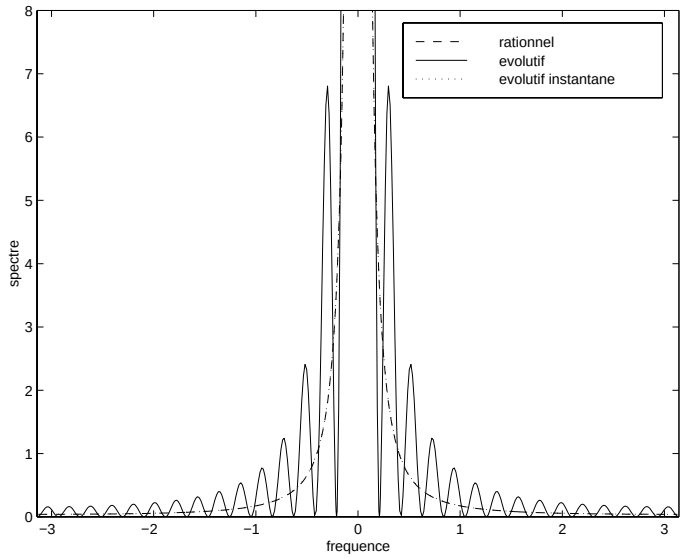


Figure 2.2 : *Représentation du spectre évolutif instantané, du spectre évolutif et du spectre rationnel,* $t = 30$, $\lambda \in [\Leftrightarrow\pi, \pi]$, $h = 1$.

Cet exemple est un cas particulier localement déterminable de la situation générale suivante. Soit $X(\cdot)$ un processus autorégressif défini par

$$\sum_{k=0}^{p} a(k) X(t \Leftrightarrow k) = \varepsilon(t), \quad t \in \mathbb{Z}, \quad a(0) = 1,$$

où $\varepsilon(\cdot)$ est un bruit blanc de variance σ_ε^2 pour $t \geq 0$, nul pour $t < 0$. Le polynôme autorégressif,

$$\varphi^f(z) = \sum_{k=0}^{p} a(k) z^{p-k},$$

est supposé stable (ses racines sont à l'intérieur du cercle unité). Le processus $X(\cdot)$ est nul pour $t < 0$ et asymptotiquement stationnaire. La densité spectrale du régime stationnaire limite est celle du modèle autorégressif associé à $\varphi^f(\cdot)$ et à σ_ε^2 :

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left| \varphi^f(e^{i\lambda}) \right|^{-2}.$$

Par inversion de $\varphi^b(z)$ (cf. section 1.4) on obtient l'expression de $X(\cdot)$ en

fonction du processus d'innovation $\varepsilon(\cdot)$:

$$X(t) = \sum_{k=0}^{+\infty} b(k)\varepsilon(t \Leftrightarrow k), \quad t \in \mathbb{Z}, \quad \sum_{k=0}^{+\infty} b(k)z^k = \frac{1}{\varphi^b(z)}.$$

Le spectre évolutif de Mélard est celui d'une moyenne mobile d'ordre t,

$$f_t(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left| \sum_{k=0}^{t} b(k)e^{-i\lambda k} \right|^2 .$$

Il converge vers $f(\lambda)$ alors que le spectre rationnel de Grenier atteint immédiatement cette limite, $\rho(t,\lambda) = f(\lambda)$ pour $t \geq p$. Notons que la fonction d'autocovariance $R_t(\cdot)$ associée à $f_t(\cdot)$ est encore définie par $R_t(k) = R(t, t \Leftrightarrow k)$, $k \geq 0$ où

$$R_t(k) = \sigma_\varepsilon^2 \sum_{j=0}^{t-k} b(j)b(k+j) \text{ si } 0 \leq k \leq t, \ 0 \text{ si } k > t.$$

Le spectre évolutif instantané est celui d'un modèle autorégressif d'ordre p, dès que t est supérieur ou égal à p, dont la variance du bruit est égale à σ_ε^2 et pour lequel le polynôme autorégressif $\varphi_t^f(\cdot)$ converge vers $\varphi^f(\cdot)$ tout en conservant la propriété de stabilité,

$$dF_t(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left| \varphi_t^f(e^{i\lambda}) \right|^{-2} d\lambda, \quad \varphi_t^f(z) = \sum_{k=0}^{p} a_t(k)z^{p-k}.$$

En effet on a $\varepsilon^f(t; t \Leftrightarrow k) = \varepsilon(t)$ pour $k \geq p$ et $t \geq p$. La stabilité du polynôme $\varphi_t^f(\cdot)$ est garantie par la définition puisqu'il est associé aux autocorrélations partielles $\beta_t(k) = \beta(t, t \Leftrightarrow k)$, $k = 1, \ldots, p$. Sans entrer dans les détails, il est clair que ce spectre est mieux adapté à cette situation que celui de Mélard. Par exemple pour le cas $AR(1)$, $X(t) \Leftrightarrow aX(t \Leftrightarrow 1) = \varepsilon(t)$, $|a| < 1$, on obtient :

$$f_t(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \frac{\left| 1 \Leftrightarrow a^{t+1}e^{-i\lambda(t+1)} \right|^2}{\left| 1 \Leftrightarrow ae^{-i\lambda} \right|^2},$$

$$dF_t(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left| 1 \Leftrightarrow a\sqrt{\frac{1 \Leftrightarrow a^{2t}}{1 \Leftrightarrow a^{2(t+1)}}} e^{-i\lambda} \right|^{-2} d\lambda,$$

avec comme situation limite (spectre de Grenier)

$$\rho(t,\lambda) = f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left| 1 \Leftrightarrow ae^{-i\lambda} \right|^{-2} .$$

2.5.2 Succession de deux autorégressifs stationnaires

Nous illustrons maintenant la situation où $X(\cdot)$ est un processus autorégressif dont les paramètres, associés à deux modèles stationnaires, changent

à l'instant 0. On se place dans le cas de modèles d'ordre un afin d'obtenir des expressions simples pour les différents spectres, facilitant leur comparaison. De façon précise le processus $X(\cdot)$ est défini par

$$X(t) \Leftrightarrow a_t X(t \Leftrightarrow 1) = \varepsilon(t), \quad t \in \mathbb{Z},$$

où $a_t = a_1$ si $t < 0$, a_2 sinon. Le processus $\varepsilon(\cdot)$ est un bruit blanc de variance σ_ε^2 et constitue le processus d'innovation de $X(\cdot)$. Les paramètres a_1 et a_2 sont de module strictement inférieur à 1.

Pour $t < 0$, les trois spectres sont identiques et sont donnés par la densité spectrale du premier modèle :

$$\frac{dF_t(\lambda)}{d\lambda} = f_t(\lambda) = \rho(t,\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left|1 \Leftrightarrow a_1 e^{-i\lambda}\right|^{-2}.$$

Pour $t \geq 0$, seul le spectre de Grenier coïncide immédiatement avec celui du second modèle,

$$\rho(t,\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left|1 \Leftrightarrow a_2 e^{-i\lambda}\right|^{-2}.$$

Les deux autres spectres convergent, lorsque t tend vers l'infini, vers cette limite.

L'expression de $X(t)$, pour $t \geq 0$, en fonction du processus d'innovation $\varepsilon(\cdot)$ est donnée par :

$$X(t) = \sum_{k=0}^{t} a_2^k \varepsilon(t \Leftrightarrow k) + a_2^{t+1} \sum_{k=0}^{+\infty} a_1^k \varepsilon(\Leftrightarrow 1 \Leftrightarrow k).$$

On en déduit

$$\beta_t(0) = \sigma_\varepsilon^2 \left\{ \frac{1}{1 \Leftrightarrow a_2^2} \Leftrightarrow a_2^{t+1} \frac{(a_2^2 \Leftrightarrow a_1^2)}{(1 \Leftrightarrow a_2^2)(1 \Leftrightarrow a_1^2)} \right\},$$

$$\beta_t(1) = a_2 \sqrt{\frac{\beta_{t-1}(0)}{\beta_t(0)}}, \quad \beta_{-1}(0) = \frac{\sigma_\varepsilon^2}{1 \Leftrightarrow a_1^2},$$

$$\frac{dF_t(\lambda)}{d\lambda} = \frac{\sigma_\varepsilon^2}{2\pi} \left|1 \Leftrightarrow \beta_t(1) e^{-i\lambda}\right|^{-2}.$$

Le spectre évolutif instantané est celui d'un modèle $AR(1)$ stationnaire dont le pôle converge vers a_2 lorsque t tend vers l'infini.

Le spectre évolutif de Mélard est donné par

$$\begin{aligned} f_t(\lambda) &= \frac{\sigma_\varepsilon^2}{2\pi} \left| \frac{1}{1 \Leftrightarrow a_2 e^{-i\lambda}} + \frac{(a_1 \Leftrightarrow a_2) a_2^{t+1} e^{-i\lambda(t+2)}}{(1 \Leftrightarrow a_2 e^{-i\lambda})(1 \Leftrightarrow a_1 e^{-i\lambda})} \right|^2, \\ &= \frac{\sigma_\varepsilon^2}{2\pi} \frac{\left| 1 + (a_2 \Leftrightarrow a_1) \sum_{k=1}^{t+1} a_2^{k-1} e^{-i\lambda k} \right|^2}{\left|1 \Leftrightarrow a_1 e^{-i\lambda}\right|^2}. \end{aligned}$$

Il s'agit de la densité d'un modèle $ARMA(1, t+1)$ stationnaire qui converge vers la limite souhaitée, mais de façon beaucoup plus irrégulière. Les figures qui suivent illustrent ce point.

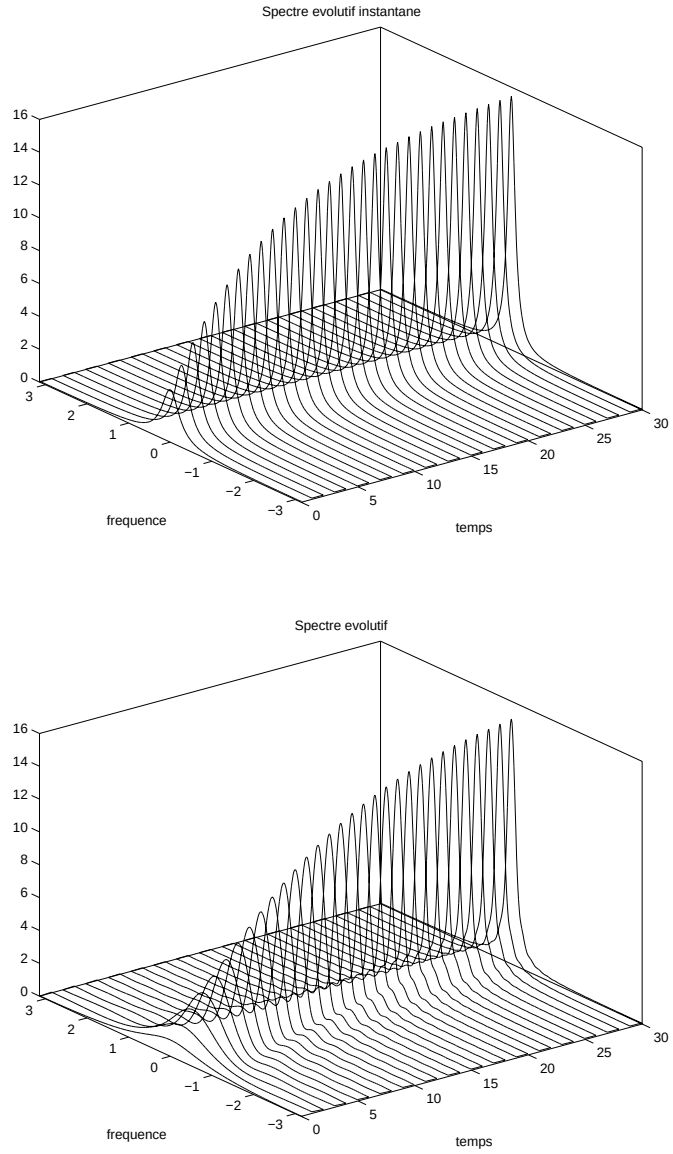


Figure 2.3 : *Représentation en trois dimensions du spectre évolutif instantané et du spectre évolutif,* $t \in [0, \ldots, 30]$, $\lambda \in [\Leftrightarrow\pi, \pi]$, $a_1 = 0.4$, $a_2 = 0.9$, $\sigma_\varepsilon^2 = 1$.

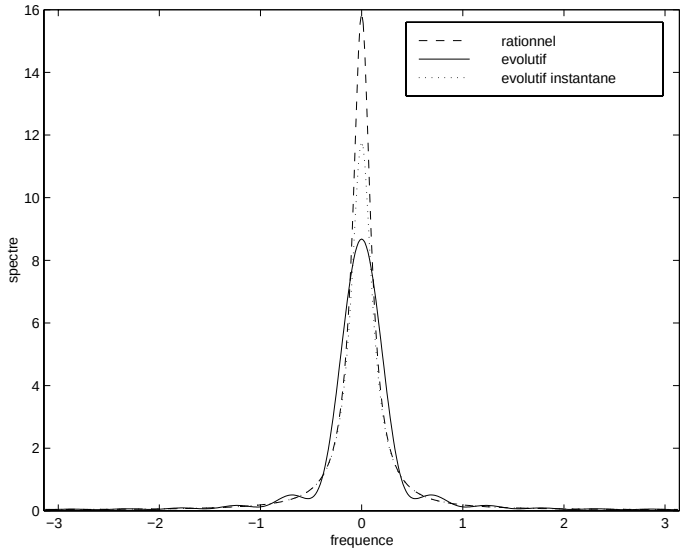


Figure 2.4 : *Représentation du spectre évolutif instantané, du spectre évolutif et du spectre rationnel,* $t = 10$, $\lambda \in [\Leftrightarrow\pi, \pi]$, $a_1 = 0.4$, $a_2 = 0.9$, $\sigma_\varepsilon^2 = 1$.

2.5.3 Processus uniformément modulé

On s'intéresse à un cas de non-stationnarité simple souvent évoqué : les processus uniformément modulés. Ces processus ont été introduits par Priesley [Pri65] comme des processus stationnaires modulés en temps. Il s'agit des processus $Y(\cdot)$ de la forme

$$Y(t) = c(t)X(t), \quad t \in \mathbb{Z},$$

où $X(\cdot)$ est un processus stationnaire et $c(\cdot)$ une fonction à valeurs réelles non négatives. Le spectre évolutif instantané est bien adapté à cette situation. En effet, la fonction d'autocorrélation partielle de $Y(\cdot)$ coïncide avec celle de $X(\cdot)$ en tout point (t, s) où $s \neq t$, et $\beta_t^Y(0) = c^2(t)\beta^X(0)$. On a ainsi

$$dF_t^Y(\lambda) = c^2(t)dF^X(\lambda).$$

Il en est de même pour le spectre évolutif de Mélard, en se restreignant aux processus purement indéterminables, car les coefficients de la décomposition de Wold-Cramér de $Y(\cdot)$ sont donnés par $h^Y(t, t\Leftrightarrow j) = c(t)h^X(j)$. Par contre, ce n'est plus vrai pour le spectre rationnel. Supposons que le processus $X(\cdot)$ admette une représentation $ARMA(p, q)$ de paramètres a_j, $j = 0, \ldots, p$ et b_j, $j = 0, \ldots, q$, et que la fonction $c(\cdot)$ ne s'annule pas. Alors la première

version du spectre rationnel, par exemple, est de la forme

$$\rho(t,\lambda) = c^2(t)\frac{\frac{\sigma_\varepsilon^2}{2\pi}\left|\sum_{j=0}^{q} b_j e^{-i\lambda j}\right|^2}{\left|\sum_{j=0}^{p} \frac{c(t)}{c(t-j)} a_j e^{-i\lambda j}\right|^2}, \quad a_0 = b_0 = 1.$$

Les deux exemples suivant sont consacrés à l'étude du spectre évolutif instantané pour des "chirps linéaires" et leur partie réelle. Ce sont des processus déterminables, $\mathcal{M}(\Leftrightarrow\infty;t) = \mathcal{M}(\Leftrightarrow\infty) = \mathcal{M}$, le spectre évolutif ainsi que le spectre rationnel ne sont donc pas définis pour cette situation.

2.5.4 Chirps linéaire

Considérons un processus $Z(\cdot)$ à temps discret de la forme

$$Z(t) = e^{i\left[\nu_0 t^2 + \lambda_0 t + \Phi\right]}, \quad t \in \mathbb{Z},$$

où ν_0, λ_0 sont des constantes réelles et Φ une variable uniforme sur $]\Leftrightarrow\pi, \pi]$.

A l'origine ces processus sont définis en temps continu. Par échantillonage, ils conservent la même forme avec $\nu_0 = h^2\nu_0'$ et $\lambda_0 = h\lambda_0'$ comme paramètres du modèle discret, où h est le pas de discrétisation et (ν_0', λ_0') la paramétrisation initiale du modèle continu. Dans le plan complexe, $Z(t)$ représente la position d'un mobile qui évolue sur le cercle unité. L'aléa Φ indique sa position à l'instant $t = 0$ et le mouvement est régit par la vitesse de rotation égale à $2\nu_0 t + \lambda_0$. Cette variation linéaire devrait se retrouver sur le spectre si celui-ci constitue une bonne représentation de la fréquence instantanée.

D'après la Propriété 2.2.7, l'effet de la constante λ_0 se traduit par un simple décalage du spectre en fréquence de la quantité $\lambda_0 \mod 2\pi$ appartenant à $]\Leftrightarrow\pi, \pi]$. D'autre part on peut prendre ν_0 dans $]\Leftrightarrow\pi/2, \pi/2]$. Pour $\lambda_0 = 0$ la fonction d'autocovariance de $Z(\cdot)$ satisfait $R(t,s) = e^{i\nu_0(t^2-s^2)}$, ce qui conduit à $\beta_t(0) = 1$ et $\beta_t(1) = e^{i\nu_0(2t-1)}$. Ce dernier coefficient étant de module égal à 1, la fonction $\beta_t(\cdot)$ est nulle au delà de 1. La mesure spectrale dF_t se réduit donc à la mesure de Dirac en la fréquence de $]\Leftrightarrow\pi, \pi]$ égale à $\nu_0(2t\Leftrightarrow 1) \mod 2\pi$. Pour chaque t, le spectre évolutif instantané du processus $Z(\cdot)$ charge la fréquence $\nu_0(2t\Leftrightarrow 1) + \lambda_0 \mod 2\pi$ avec un poids égal à 1. Dans le cas particulier où ν_0 serait de la forme $\pm\pi/T$ pour un entier $T \geq 2$, c'est-à-dire un pas de discrétisation h égal à $\sqrt{\pi/|\nu_0'|T}$ pour la version continue, le processus est périodiquement corrélé de période T. En effet la périodicité du spectre évolutif instantané par rapport à la variable temps se retrouve sur la fonction d'autocorrélation partielle et la Proposition 3.1.1 conduit à

ce résultat. Notons que le processus $Z(\cdot)$ lui-même n'est généralement pas périodique. La Figure 2.5 représente ce spectre dans le plan temps-fréquence. Le temps est porté en abscisse et la fréquence, qui varie entre $\Leftrightarrow\pi$ et π, en ordonnée. Les masses, toutes égales à 1, sont représentées par des ronds.

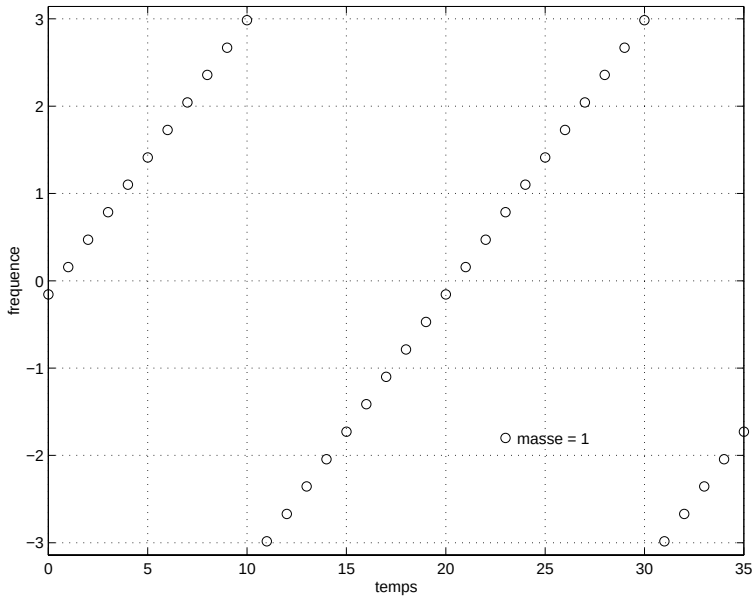


Figure 2.5 : *Représentation temps-fréquence du spectre évolutif instantané,* $\nu_0 = \pi/20$, $\lambda_0 = 0$.

Dans cet exemple, la constante ν_0 est réelle. L'étude peut se généraliser à une constante complexe : le processus $Z(\cdot)$ est alors de la forme

$$Z(t) = e^{\mu_0 t^2 + i\left[\nu_0 t^2 + \lambda_0 t + \Phi\right]}, \quad t \in \mathbb{Z},$$

où μ_0, ν_0, λ_0 sont trois constantes réelles. Pour t fixé, le spectre de ce processus charge la même fréquence que précédemment, mais avec une masse, égale à $e^{2\mu_0 t^2}$, qui dépend de t.

Ainsi le spectre évolutif instantané des "chirps linéaires" satisfait la représentation idéale puisque son support est porté par les droites de pentes $2\nu_0$. De façon précise, il s'agit de l'ensemble des points de $\mathbb{Z}\times]\Leftrightarrow\pi, \pi]$ de la forme $\{t, \nu_0(2t \Leftrightarrow 1) + \lambda_0 \mod 2\pi\}$, $t \in \mathbb{Z}$.

2.5.5 Partie réelle d'un chirps linéaire

Nous proposons d'étudier maintenant le spectre de la partie réelle du cas précédent. Soit $X(\cdot)$ un processus à valeurs réelles défini par

$$X(t) = \frac{1}{2}\left\{Z(t) + \overline{Z(t)}\right\} = \cos(\nu_0 t^2 + \lambda_0 + \Phi), \quad t \in \mathbb{Z},$$

où ν_0 et λ_0 sont deux constantes réelles et Φ une variable uniforme sur $]\Leftrightarrow\pi, \pi]$. On peut, sans perte de généralité, se ramener au cas où (ν_0, λ_0) appartient à $]\Leftrightarrow\pi/2, \pi/2]\times]\Leftrightarrow\pi, \pi]$. Le processus $X(\cdot)$ est centré et sa fonction d'autocovariance est donnée par

$$R(t,s) = \frac{1}{2}\cos\left[\nu_0(t^2 \Leftrightarrow s^2) + \lambda_0(t \Leftrightarrow s)\right].$$

On a évidemment $R(t,s) = \{R^Z(t,s) + R^{\overline{Z}}(t,s)\}/4$ puisque $Z(\cdot)$ et $\overline{Z}(\cdot)$ sont non corrélés. Par contre la relation $dF_t(\lambda) = \{dF_t^Z(\lambda) + dF_t^{\overline{Z}}(\lambda)\}/4$ n'est pas vérifiée de façon systématique. Elle ne l'est pas en chaque instant qui suit immédiatement l'intersection des droites contenant les supports de $dF_t^Z(\lambda)$ et $dF_t^{\overline{Z}}(\lambda)$. La Figure 2.6 correspond au cas associé à la Figure 2.5. Pour conserver la symétrie évoquant les processus réels, nous avons réparti les masses en π sur le couple $\{\Leftrightarrow\pi, \pi\}$.

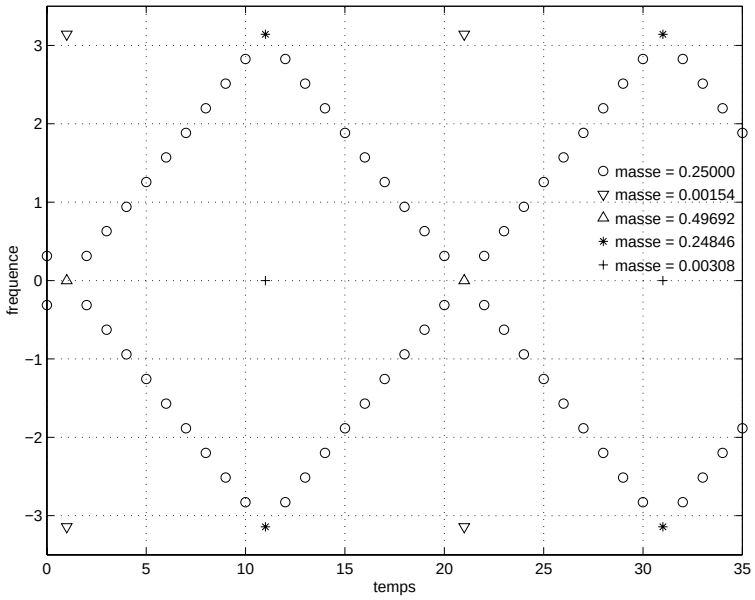


Figure 2.6 : *Représentation temps-fréquence du spectre évolutif instantané,* $\nu_0 = \pi/20$, $\lambda_0 = 0$.

Nous présentons ci-après le calcul du spectre. L'Algorithme LDG permet d'obtenir les premiers coefficients de la fonction $\beta_t(\cdot)$:

$$\begin{aligned}
\beta_t(0) &= \frac{1}{2}, \\
\beta_t(1) &= \cos\left[(2t \Leftrightarrow 1)\nu_0 + \lambda_0\right], \\
\beta_t(2) &= \Leftrightarrow \frac{\sin\left[(2t \Leftrightarrow 3)\nu_0 + \lambda_0\right]\sin\left[(2t \Leftrightarrow 1)\nu_0 + \lambda_0\right]}{\left\{\sin^2\left[(2t \Leftrightarrow 3)\nu_0 + \lambda_0\right]\sin^2\left[(2t \Leftrightarrow 1)\nu_0 + \lambda_0\right]\right\}^{\frac{1}{2}}},
\end{aligned}$$

avec la convention $0^{-1} = 0$. La situation la plus courante est celle pour laquelle (ν_0, λ_0) vérifie $(2t \Leftrightarrow 1)\nu_0 + \lambda_0 \neq k\pi$ pour tout couple (t, k) de $\mathbb{ZZ}^2$. Elle correspond au fait que les droites évoquées plus haut ne se coupent pas en des points d'abscisse entière. Alors $|\beta_t(1)| < 1$ et $|\beta_t(2)| = 1$. Lorsque $\beta_t(2) = \Leftrightarrow 1$, $dF_t(\lambda)$ charge, avec une masse égale à $1/4$ les deux fréquences $\pm[2(t \Leftrightarrow 1)\nu_0 + \lambda_0]$. Lorsque $\beta_t(2) = 1$, $dF_t(\lambda)$ charge les deux fréquences 0 et π avec

$$dF_t(\lambda) = \frac{1}{4}\left\{1 + \cos\left[(2t-1)\nu_0 + \lambda_0\right]\right\}\delta_0(\lambda) + \frac{1}{4}\left\{1 - \cos\left[(2t-1)\nu_0 + \lambda_0\right]\right\}\delta_\pi(\lambda),$$

où $\delta_\theta(\lambda)$ désigne la mesure de Dirac en θ. Ceci se produit à chaque instant t qui suit immédiatement l'intersection des droites contenant les supports de $dF_t^Z(\lambda)$ et $dF_t^{\overline{Z}}(\lambda)$. Dans ce cas le support de $dF_t(\lambda)$ n'est donc plus sur les droites. Il nous reste à considérer la situation d'un couple (ν_0, λ_0) qui vérifie à un instant t la condition $(2t \Leftrightarrow 1)\nu_0 + \lambda_0 = k\pi$. Dans ce cas on a $\beta_t(1) = (\Leftrightarrow 1)^k$. La mesure $dF_t(\lambda)$ charge la fréquence 0 ou π, avec une masse égale à $1/2$, selon que k est pair ou impair. Notons que son support reste sur les droites habituelles. A l'instant $(t+1)$ suivant, on a $|\beta_{t+1}(1)| < 1$ (sauf si $\nu_0 = \pi/2$) et $|\beta_{t+1}(2)| = 0$. Il est donc nécessaire de déterminer $\beta_{t+1}(3)$. Celui-ci vérifie $\beta_{t+1}(3)\sin^2(2\nu_0) = \cos(3k\pi) \Leftrightarrow \cos(k\pi + 2\nu_0)\cos(2\nu_0)$, et par suite $\beta_{t+1}(3) = (\Leftrightarrow 1)^k$. La mesure $dF_{t+1}(\lambda)$ est donnée pour k pair par

$$dF_{t+1}(\lambda) = \frac{1}{2}\left[\frac{1 + \cos(2\nu_0)}{3 \Leftrightarrow \cos(2\nu_0)}\right]\delta_0(\lambda) + \frac{1}{2}\left[\frac{1 \Leftrightarrow \cos(2\nu_0)}{3 \Leftrightarrow \cos(2\nu_0)}\right]\left\{\delta_{-\theta_1}(\lambda) + \delta_{\theta_1}(\lambda)\right\},$$

et pour k impair par

$$dF_{t+1}(\lambda) = \frac{1}{2}\left[\frac{1 + \cos(2\nu_0)}{3 \Leftrightarrow \cos(2\nu_0)}\right]\delta_\pi(\lambda) + \frac{1}{2}\left[\frac{1 \Leftrightarrow \cos(2\nu_0)}{3 \Leftrightarrow \cos(2\nu_0)}\right]\left\{\delta_{-\theta_2}(\lambda) + \delta_{\theta_2}(\lambda)\right\},$$

où θ_2 est la fréquence de $[0, \pi/2]$ déterminée par la relation $\cos\theta_2 = \sin^2\nu_0$ et $\theta_1 = \pi \Leftrightarrow \theta_2$. Dans ce cas le support de $dF_{t+1}(\lambda)$ ne reste pas sur les droites.

Chapitre 3

Processus périodiquement corrélés

Ce chapitre est consacré à l'étude de la paramétrisation de la structure au second ordre des processus périodiquement corrélés. Ces processus ont été introduit par Gladysev [Gla61] comme la classe particulière de processus non stationnaires dont la fonction d'autocovariance est périodique en les deux variables. On obtient la même caractérisation par la fonction d'autocorrélation partielle. Ainsi, comme dans la situation non stationnaire générale, les autocorrélations partielles fournissent une nouvelle paramétrisation facilement identifiable par rapport à la fonction $R(\cdot,\cdot)$. En particulier cette approche permet d'étendre de façon naturelle la méthode du maximum d'entropie à la situation périodique. Dans [Gla61], l'auteur montre que la classe des processus périodiquement corrélés et celle des processus vectoriels stationnaires sont en correspondance biunivoque. Cette relation permet d'envisager aussi une paramétrisation de type vectoriel. Cependant il existe plusieurs extensions de la fonction d'autocorrélation partielle aux processus vectoriels. Nous retenons une fonction matricielle, dite triangulaire, apparaissant très naturellement dans la correspondance avec les processus périodiquement corrélés. Cette fonction, notée $\Pi(\cdot)$, correspond à celle proposée par Pham [Pha92], mais la définition que nous en donnons, est proche de celle de Dégerine [Dég90] par son approche en terme de variables.

La première section est consacrée à la paramétrisation par les autocorrélations partielles scalaires introduites au premier chapitre. Les nouvelles contraintes sur la fonction $\beta(\cdot,\cdot)$ sont obtenues en restreignant le domaine $\mathcal{D}_\beta$ aux fonctions périodiques. La correspondance entre les fonctions $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$ est assurée par l'Algorithme LDG restreint à cette situation. La fonction matricielle $\Pi(\cdot)$ est définie à la deuxième section. Nous montrons que cette fonction caractérise la structure des processus vectoriels station-

naires. La démarche est essentiellement la même que celle utilisée au premier chapitre pour établir la caractérisation par $\beta(\cdot,\cdot)$. Après avoir décrit son domaine de variation, nous présentons un procédé qui permet de construire une séquence d'un processus vectoriel de fonction d'autocorrélation partielle matricielle $\Pi(\cdot)$ connue, prouvant ainsi la surjectivité avec la fonction d'autocovariance $R(\cdot)$. L'injectivité est assurée grâce à un algorithme de type Levinson-Durbin. La correspondance biunivoque entre les paramétrisations en termes d'autocorrélation partielle scalaire $\beta(\cdot,\cdot)$ et vectorielle $\Pi(\cdot)$ fait l'objet de la troisième section. Elle est obtenue en utilisant de façon adéquate l'Algorithme *LDG*. En particulier, elle permet de réaliser la relation entre $\Pi(\cdot)$ et $R(\cdot)$ de façon scalaire. On s'intéresse ensuite, dans la quatrième section, aux processus autorégressifs périodiques ainsi qu'à leur correspondance avec les processus vectoriels stationnaires autorégressifs. Le problème de l'extension des coefficients d'autocovariance d'un processus stationnaire est un problème classique dans le domaine de l'estimation spectrale. Dans une dernière section, nous proposons une nouvelle approche pour traiter ce problème dans le contexte des processus périodiquement corrélés. Pour cela, nous utilisons la paramétrisation par les autocorrélations partielles. Nous étendons ensuite la méthode du maximum d'entropie [Bur75] à la situation périodique et montrons que la solution est autorégressive périodique. Le lien avec le problème de l'extension d'une suite de matrice d'autocovariance d'un processus vectoriel stationnaire est également présenté.

3.1 Fonction d'autocorrélation partielle scalaire

La théorie des processus périodiquement corrélés a débuté avec les travaux de Gladysev [Gla61]. Notons cependant qu'ils ont été introduits quelques années plus tôt sous le nom de processus cyclo-stationnaires dans le contexte de la communication (cf. [Ben58]).

Définition 3.1.1. *Un processus $X(\cdot)$ est dit périodiquement corrélé de période T lorsque sa fonction d'autocovariance est périodique de même période,*

$$R(t+T,s+T) = R(t,s) \text{ pour tout } (t,s) \in \mathbb{Z}^2.$$

En d'autres termes, la structure au second ordre de ces processus est invariante par translation en temps de longueur T. C'est-à-dire que la structure de $\{X(s+T), \ \ldots, X(t+T)\}$ est la même que celle de $\{X(s),\ldots,X(t)\}$. Notons que lorsque $T=1$, on retrouve la classe des processus stationnaires.

Suite à la Remarque 1.3.1, il est clair que la structure des processus périodiquement corrélés peut être caractérisée en termes d'autocorrélations partielles, en remplaçant dans leur définition la fonction $R(\cdot,\cdot)$ par $\beta(\cdot,\cdot)$. Plus précisément, on a la proposition suivante.

Proposition 3.1.1. *Un processus est périodiquement corrélé de période T si et seulement si sa fonction d'autocorrélation partielle satisfait* :

$$\beta(t+T,s+T)=\beta(t,s) \text{ pour tout } (t,s)\in \mathbb{Z}^2.$$

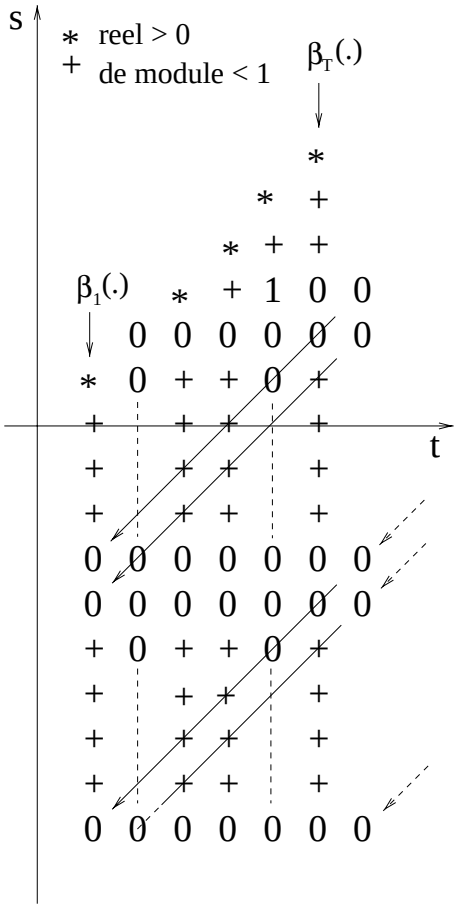


Figure 3.1 : *Illustration du domaine $\mathcal{D}_\beta^T$*

Ainsi, compte tenu des résultats obtenus au premier chapitre, la structure au second ordre de ces processus peut être paramétrisée par T fonctions définies sur $\mathbb{N}$, $\beta_t(n)=\beta(t,t\Leftrightarrow n)$, $t=1,\ldots,T$ qui sont sujettes seulement aux conditions suivantes :

(i) $\beta_t(0) \geq 0$ et $|\beta_t(n)| \leq 1$, $n \geq 1$,

(ii) $\beta_t(0) = 0 \Rightarrow \beta_t(n) = \beta_{(t+n) \bmod T}(n) = 0$, $n \geq 1$,

(iii) $|\beta_t(j)| = 1$, $j > 0 \Rightarrow \beta_t(n) = \beta_{(t+n-j) \bmod T}(n) = 0$, $n > j$,

où $n \bmod T$ est l'entier j de $[1, \ldots, T]$ tel que $n = kT + j$, $k \in \mathbb{Z}$. Nous notons $\mathcal{D}_\beta^T$ l'ensemble des fonctions définies sur $[1, \ldots, T] \times \mathbb{N}$ qui satisfont les contraintes ci-dessus. La Figure 3.1 illustre ce domaine. La périodicité de la fonction d'autocorrélation partielle se traduit par des valeurs de $\beta(\cdot, \cdot)$ invariantes par translation de longueur T et dans la direction de la diagonale. Ainsi cette fonction est caractérisée par ses valeurs, que nous avons notées $\beta_t(n)$, $t = 1, \ldots, T$, $n \in \mathbb{N}$, se trouvant sur une bande de largeur T (cf. Figure 3.1). Par ailleurs la propriété d'invariance par cette translation implique des nouvelles contraintes sur la fonction $\beta(\cdot, \cdot)$ par rapport au cas non stationnaire général. En effet les conditions $\beta_{(t+n) \bmod T}(n) = 0$, $n \geq 1$ du point (ii) et $\beta_{(t+n-j) \bmod T}(n) = 0$, $n > j$ du point (iii) traduisent le fait que les bandes de zéros horizontales dues à des singularités d'ordre fini (cf. Figure 1.2) sont translatées vers le bas de façon périodique.

Toute fonction de $\mathcal{D}_\beta^T$ est la fonction d'autocorrélation partielle d'un processus périodiquement corrélé de période T et inversement. La correspondance avec la fonction d'autocovariance peut alors être réalisée par un algorithme de Levinson-Durbin Périodique (LDP). Cet algorithme correspond à l'Algorithme LDG restreint au cas périodique. En effet, les expressions (1.1) et (1.2) des variances résiduelles en termes de la fonction $\beta(\cdot, \cdot)$ conduisent à

$$\begin{array}{l} \sigma^{f2}(k; k \Leftrightarrow n) = \sigma^{f2}(t; t \Leftrightarrow n) = \sigma_t^{f2}(n), \\ \sigma^{b2}(k \Leftrightarrow n; k) = \sigma^{b2}(t \Leftrightarrow n; t) = \sigma_t^{b2}(n), \end{array} \qquad t = k \bmod T,$$

pour un processus périodiquement corrélé de période T. De plus d'après l'invariance par translation de longueur T de la structure de tels processus, on a dans le cas non localement déterminable,

$$\begin{array}{l} a_k^f(n, j) = a_t^f(n, j), \\ a_k^b(n, j) = a_t^b(n, j), \end{array} \qquad t = k \bmod T.$$

Dans la situation générale, on peut choisir des coefficients pour les innovations partielles, qui satisfassent ces mêmes relations. Notons que par la suite, nous utiliserons indifféremment les notations $\beta(t, t \Leftrightarrow n)$ ou $\beta_t(n)$ et $\sigma(t; t \Leftrightarrow n)$ ou $\sigma_t(n)$. Dans le cas non stationnaire général, on a constaté que les restrictions des fonctions $R(\cdot, \cdot)$ et $\beta(\cdot, \cdot)$ à un domaine de la forme $[s, \ldots, t]^2$ sont en correspondance biunivoque (cf. Remarque 1.3.1). Ici, le caractère périodique de ces deux fonctions permet d'envisager d'autres domaines. Notons $E(p_1, \ldots, p_T)$ les sous-ensembles de $\mathbb{Z}^2$ de la forme $\{(t, s), t = 1, \ldots, T, 0 \leq t \Leftrightarrow s \leq p_t\}$. La correspondance entre les restrictions

de $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$ à un domaine de $\mathbb{Z}^2$ est bijective si et seulement si ce domaine peut se ramener (par périodicité) à un ensemble de la forme de $E(p_1,\ldots,p_T)$ avec $p_{i+1} \leq p_i + 1$, $i = 1,\ldots T$, où $p_{T+1} = p_1$. Ce résultat provient du fait que $\beta(t,s)$ (resp. $R(t,s)$) dépend de l'ensemble des valeurs $R(u,v)$ (resp. $\beta(u,v)$), $s \leq v \leq u \leq t$. Ces ensembles contraints seront notés $E_c(p_1,\ldots,p_T)$. L'Algorithme LDP ci-dessous détermine la correspondance entre $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$ sur un tel domaine.

Algorithme 3.1.1. (LDP)
Pour $k = 1,\ldots,T$ *:*

$$R(k,k) = \beta_k(0) = \sigma_k^{f2}(0) = \sigma_k^{b2}(0).$$

Pour $n = 1,\ldots,\max_{i=1,\ldots,T}(p_i)$, *avec les conventions* $\sum_{j=1}^{0}\ldots = 0$ *et* $0^{-1} = 0$ *:*
pour $k = 1,\ldots,T$, *avec* $n \leq p_k$ *:*

$$\beta_k(n) = \frac{R(k,k \Leftrightarrow n) + \sum_{j=1}^{n-1} a_k^f(n \Leftrightarrow 1, j) R(k \Leftrightarrow j, k \Leftrightarrow n)}{\sigma_k^f(n \Leftrightarrow 1)\sigma_{k-1}^b(n \Leftrightarrow 1)},$$

$$\begin{aligned}
\sigma_k^{f2}(n) &= \left[1 \Leftrightarrow |\beta_k(n)|^2\right] \sigma_k^{f2}(n \Leftrightarrow 1),\\
\sigma_k^{b2}(n) &= \left[1 \Leftrightarrow |\beta_k(n)|^2\right] \sigma_{k-1}^{b2}(n \Leftrightarrow 1),\\
a_k^f(n,n) &= \Leftrightarrow \beta_k(n) \frac{\sigma_k^f(n \Leftrightarrow 1)}{\sigma_{k-1}^b(n \Leftrightarrow 1)},\\
a_k^b(n,n) &= \Leftrightarrow \overline{\beta_k(n)} \frac{\sigma_{k-1}^b(n \Leftrightarrow 1)}{\sigma_k^f(n \Leftrightarrow 1)}
\end{aligned}$$

pour $j = 1,\ldots,n \Leftrightarrow 1$ *:*

$$\begin{aligned}
a_k^f(n,j) &= a_k^f(n \Leftrightarrow 1, j) + a_k^f(n,n) a_{k-1}^b(n \Leftrightarrow 1, n \Leftrightarrow j),\\
a_k^b(n,j) &= a_{k-1}^b(n \Leftrightarrow 1, j) + a_k^b(n,n) a_k^f(n \Leftrightarrow 1, n \Leftrightarrow j),
\end{aligned}$$

où l'indice $k \Leftrightarrow 1 = 0$ *est remplacé par* T.

L'algorithme de type LDP, proposé par Sakai ([Sak82],[Sak83]) dans le cas non localement déterminable, coïncide avec celui-ci pour un domaine $E_c(p_1,\ldots,p_T)$ avec $p_i = p$, $i = 1,\ldots,T$. Par ailleurs, on trouve dans [Pha92] un algorithme LDP qui établit la correspondance entre $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$ dans la situation générale sur un domaine qui peut se ramener à un ensemble $E_c(p_1,\ldots,p_T)$ avec $p_i = pT + i \Leftrightarrow 1$, $i = 1,\ldots,T$. Dans la suite, ces ensembles seront notés $E_c(p,\max)$. Remarquons que la structure de cette procédure est légèrement différente de celle de L'Algorithme 3.1.1. En effet, elle détermine

les quantités $\beta(nT+i,j)$ (resp. $R(nT+i,j)$) pour n variant de 0 à p, $i = 1,\ldots,T$ et $j = i,\ldots,1$ si $n = 0$, $T,\ldots,1$ sinon. Cependant les coefficients des innovations partielles utilisés dans cet algorithme sont les mêmes que ceux obtenus avec L'Algorithme 3.1.1.

3.2 Fonction d'autocorrélation partielle matricielle triangulaire

Sur le plan théorique, les processus périodiquement corrélés sont intéressant, car intimement liés aux processus vectoriels stationnaires. En effet soit $Y(\cdot) = \{Y(t), t \in \mathbb{Z}\}$ le processus vectoriel T-dimensionnel défini à partir d'un processus $X(\cdot)$ par,

$$Y_j(t) = X(j + T(t \Leftrightarrow 1)), \quad j = 1,\ldots,T,\ t \in \mathbb{Z}.$$

Alors ([Gla61] : Théorème 1, p.385) $X(\cdot)$ est périodiquement corrélé de période T si et seulement si le processus vectoriel $Y(\cdot)$ est stationnaire. Nous considérons le processus T-dimensionnel stationnaire $Y(\cdot)$ ainsi associé au processus $X(\cdot)$. La structure au second ordre de $Y(\cdot)$ est déterminée par la fonction d'autocovariance $R(\cdot)$,

$$R(n) = E\left\{Y(t+n)Y(t)^*\right\}, \quad n \in \mathbb{Z}.$$

Les erreurs de prédiction d'ordre n progressive et rétrograde sont notées $e^f(t;n)$ et $e^b(t;n)$ ($e^f(t;0) = e^b(t;0) = Y(t)$) et leurs matrices de covariance , $^{f2}(n)$ et , $^{b2}(n)$.

L'extension de la fonction d'autocorrélation partielle aux processus vectoriels stationnaires est délicate. La difficulté pour définir une matrice d'autocorrélation partielle réside dans le choix de la racine carrée pour normaliser l'autocovariance partielle $\delta(\cdot)$,

$$\delta(0) = R(0), \quad \delta(\Leftrightarrow n)^* = \delta(n) = E\left\{e^f(t;n \Leftrightarrow 1)e^b(t \Leftrightarrow n; n \Leftrightarrow 1)^*\right\}, \quad n > 0.$$

Dégerine [Dég90] donne les conditions que doivent satisfaire les racines carrées des variances résiduelles et leurs inverses pour que la correspondance entre fonction d'autocorrélation partielle et fonction d'autocovariance soit biunivoque. Les fonctions d'autocorrélation partielle, dites triangulaires, ([Dég90], [Pha92]) apparaissent très naturellement dans la correspondance avec les processus périodiquement corrélés. Dans le cas localement déterminable, ces deux définitions sont légèrement différentes. Sur le plan théorique la fonction proposée par Dégerine est la plus attractive. Celle de Pham est plus pratique pour la correspondance avec la fonction $\beta(\cdot,\cdot)$. Nous préférons à l'approche algébrique de Pham, une définition en terme de variables comme dans [Dég90]. Néanmoins la définition que nous proposons est

plus proche de celle de Pham.

3.2.1 Procédé de normalisation et définition

Nous décrivons ci-après le procédé de normalisation des innovations partielles que nous utilisons pour définir la fonction d'autocorrélation partielle matricielle. Soit e un vecteur aléatoire complexe de dimension T, centré et de matrice de covariance $,^2$. Dans le sens progressif, l'orthogonalisation par le procédé de Gram-Schmidt des composantes successives e_j, $j = 1, \ldots, T$, de e définit de nouvelles variables ζ_j, $j = 1, \ldots, T$ et ceci de façon unique si l'on pose $\zeta_j = 0$ lorsque $e_j \in L\{e_{j-1}, \ldots, e_1\}$. Le vecteur ζ est centré et $E\{\zeta\zeta^*\} = J$, où J est une matrice diagonale telle que $J_{j,j} = 0$ si $e_j \in L\{e_{j-1}, \ldots, e_1\}$, 1 sinon. Dans le sens rétrograde, le même procédé est utilisé mais à partir des composantes $e_T, \ldots, e_1$.

Ainsi la fonction d'autocorrélation partielle matricielle triangulaire est obtenue en utilisant une normalisation triangulaire inférieure (resp. supérieure) dans le sens progressif (resp. rétrograde). Plus précisément, les innovations partielles normalisées d'ordre n, définies de façon unique en appliquant les procédés précédents aux variables $e^f(t;n)$ et $e^b(t;n)$, sont notées $\zeta^f(t;n)$ et $\zeta^b(t;n)$. On pose alors la définition suivante.

Définition 3.2.1. *La fonction d'autocorrélation partielle triangulaire de* $Y(\cdot)$, *notée* $\Pi(\cdot)$, *est définie pour tout* $n \in \mathbb{Z}$ *par :*

$$\Pi(0) = R(0),$$
$$\Pi(\Leftrightarrow n)^* = \Pi(n) = E\left\{\zeta^f(t; n \Leftrightarrow 1)\zeta^b(t \Leftrightarrow n; n \Leftrightarrow 1)^*\right\}, \quad n \in \mathbb{N}^*.$$

Par la suite toute normalisation ζ d'un vecteur e sera de type triangulaire défini comme précédemment. On est amené à considérer les relations liant ζ et e. Pour la normalisation triangulaire inférieure, on peut écrire $e = ,_i\zeta$ et $\zeta = ,_i^- e$, où $,_i$ et $,_i^-$ sont des matrices triangulaires inférieures. Nous utiliserons les notations $,_s$ et $,_s^-$ pour les matrices triangulaires supérieures. Dans le cas localement déterminable, les matrices $,_i$ et $,_i^-$ satisfaisant ces relations ne sont pas uniques, bien que ζ le soit par définition. Nous donnons ci-dessous le choix retenu pour ces matrices.

Pour $,_i$, seules les colonnes associées à $\zeta_j = 0$ ne sont pas définies de façon unique. On obtient l'unicité en les posant égales à 0. Ainsi $,_i$ sera l'unique racine triangulaire inférieure de $,^2$ (i.e. $,_i,_i^* = ,^2$) avec des éléments diagonaux positifs et telle que $,_i J = ,_i$. La matrice $,_i$ peut être obtenue à partir d'une décomposition de Cholesky LD^2L^* de $,^2$, où D^2 est une matrice diagonale dont les éléments diagonaux sont positifs ou nuls et L une matrice triangulaire inférieure unitaire (les éléments diagonaux sont égaux à 1). Alors

le produit LD est l'unique racine $,_i$ définie précédemment, et ceci quelque soit la décomposition de Cholesky considérée.

La définition d'un inverse de $,_i$ est un point plus délicat. Soit $,_i^-$ un inverse généralisé de $,_i$ au sens où $,_i,_i^-,_i = ,_i$ (cf. [RAO65] : p. 24). Pour garantir l'écriture $\zeta = ,_i^- e$, il doit satisfaire $,_i^-,_i = J$. Compte tenu de la relation $e = ,_i\zeta$, cette condition est nécessaire et suffisante mais n'assure pas l'unicité. Par la suite $,_i^-$ désignera un élément quelconque de cette classe d'inverses généralisés. Une sous-classe de ces inverses est donnée par les décompositions de Cholesky de $,^2$ de la façon suivante. On désigne par D la racine carrée symétrique de D^2 et par D^+ l'inverse généralisé de D, obtenu en inversant les termes non nuls. Alors la matrice D^+L^{-1} représente un choix possible pour $,_i^-$. Ces inverses sont caractérisés par $J,_i^- = ,_i^-$. Enfin une façon naturelle d'obtenir l'unicité pour $,_i^-$, est de poser égales à 0 les colonnes associées à $\zeta_j = 0$ (i.e. $,_i^- J = ,_i^-$). Les lignes de $,_i^-$ sont alors constituées des coefficients des décompositions de ζ_i sur les variables linéairement indépendantes e_j, $1 \leq j \leq T$ telles que $J_{j,j} = 1$. Cet inverse est donné par la décomposition de Cholesky $(,_i^- J = ,_i^- \Rightarrow J,_i^- = ,_i^-)$ pour laquelle les colonnes de la matrice L associées à $J_{j,j} = 0$ sont celles de l'identité. Cependant nous constaterons par la suite (cf. sous-section 3.2.3) qu'il n'est pas toujours souhaitable d'imposer cette contrainte. En effet dans ce cas, les relations de récurrence pour les inverses des racines des matrices de covariance résiduelles ne sont plus satisfaites.

Nous avons des résultats analogues pour la normalisation triangulaire supérieure. Remarquons que pour ces deux procédés de normalisation, la matrice $,^{-*},^-$ est l'inverse généralisé de $,^2$ au sens de [RAO65].

Dans [PHA92], la fonction d'autocorrélation partielle triangulaire $\tilde{\Pi}(\cdot)$ est présentée pour $n \neq 0$ comme la solution (pas nécessairement unique) de,

$$,^f(n \Leftrightarrow 1)_i\tilde{\Pi}(n),^b(n \Leftrightarrow 1)_s^* = \delta(n).$$

La matrice $\Pi(n)$, qui est alors la solution pour laquelle les lignes et les colonnes de $\tilde{\Pi}(n)$ non définies de façon unique sont nulles, est ensuite utilisée pour décrire les contraintes que doit satisfaire $\tilde{\Pi}(\cdot)$.

La fonction d'autocorrélation partielle triangulaire $\beta(\cdot)_{tr}$ ([DÉG90]) est définie comme précédemment mais en éliminant les composantes des innovations normalisées nulles presque sûrement. Plus précisément, pour $n \geq 1$, $\beta(n)_{tr}$ est la matrice de dimension $r(n) \times r(n)$ $(r(n) \leq T)$ obtenue en supprimant les i-èmes lignes et les j-èmes colonnes de zéros de $\Pi(n)$ qui correspondent à des singularités d'ordre fini, c'est-à-dire $\zeta^f(t; n \Leftrightarrow 1)_i = 0$ $p.s.$ et $\zeta^b(t \Leftrightarrow n; n \Leftrightarrow 1)_j = 0$ $p.s.$

3.2.2 Domaine de variation de $\Pi(\cdot)$

L'avantage de la fonction $\beta(\cdot)_{tr}$ par rapport à $\Pi(\cdot)$ est qu'elle est plus facile à identifier. La matrice $\beta(0)_{tr}$ $(= \Pi(0))$ est définie non négative et $\{\beta(n)_{tr},\ \ n \geq 1\}$ est une séquence de matrices carrées dont les valeurs singulières (valeurs propres de $\beta(n)_{tr}\beta(n)^*_{tr}$) sont inférieures ou égales à 1. L'ordre de $\beta(n+1)_{tr}$ est égal au rang de $I \Leftrightarrow \beta(n)_{tr}\beta(n)^*_{tr}$, l'ordre de $\beta(1)_{tr}$ étant égal au rang de $\beta(0)_{tr}$.

L'identification de la fonction $\Pi(\cdot)$ est un point plus délicat. Notre approche en termes de variables conduit à retrouver de façon différente les contraintes de [PHA92]. Notons $J^f(n)$ et $J^b(n)$ les matrices de covariance des erreurs de prédiction normalisées $\zeta^f(t;n)$ et $\zeta^b(t;n)$. Nous avons le résultat suivant.

Proposition 3.2.1. *Les innovations partielles normalisées satisfont, pour* $n \geq 1$, *les relations de récurrence* $(\zeta^f(t;0) = [\Pi(0)]_i^{1/2-} X(t),\ \zeta^b(t;0) = [\Pi(0)]_s^{1/2-} X(t))$,

$$\zeta^f(t;n) = \left[W^f(n)\right]_i^{\frac{1}{2}-} \left\{\zeta^f(t;n \Leftrightarrow 1) \Leftrightarrow \Pi(n)\zeta^b(t \Leftrightarrow n; n \Leftrightarrow 1)\right\}, \tag{3.1}$$

$$\zeta^b(t;n) = \left[W^b(n)\right]_s^{\frac{1}{2}-} \left\{\zeta^b(t;n \Leftrightarrow 1) \Leftrightarrow \Pi(n)^*\zeta^f(t+n; n \Leftrightarrow 1)\right\}, \tag{3.2}$$

et leurs matrices de covariance sont données pour $n \geq 0$ *par*

$$\begin{aligned} J^f(n) &= \left[W^f(n)\right]_i^{\frac{1}{2}-} \left[W^f(n)\right]_i^{\frac{1}{2}}, \\ J^b(n) &= \left[W^b(n)\right]_s^{\frac{1}{2}-} \left[W^b(n)\right]_s^{\frac{1}{2}}, \end{aligned} \tag{3.3}$$

où $W^f(0) = W^b(0) = \Pi(0)$, *et*

$$\begin{aligned} W^f(n) &= J^f(n \Leftrightarrow 1) \Leftrightarrow \Pi(n)\Pi(n)^*, \\ W^b(n) &= J^b(n \Leftrightarrow 1) \Leftrightarrow \Pi(n)^*\Pi(n). \end{aligned} \tag{3.4}$$

Démonstration.- Supposons que dans (3.1), les variables $\zeta^f(t;n \Leftrightarrow 1)$ et $\zeta^b(t \Leftrightarrow n; n \Leftrightarrow 1)$ soient les erreurs normalisées d'ordre $n \Leftrightarrow 1$ (évident pour $n = 1$). Alors la variable $\zeta^f(t;n \Leftrightarrow 1) \Leftrightarrow \Pi(n)\zeta^b(t \Leftrightarrow n; n \Leftrightarrow 1)$ est l'erreur de prédiction de $\zeta^f(t;n \Leftrightarrow 1)$ sur $\zeta^b(t \Leftrightarrow n; n \Leftrightarrow 1)$, dont la matrice de covariance est $W^f(n)$. L'orthogonalisation, selon le procédé décrit précédemment, des composantes successives de cette erreur donne $\zeta^f(t;n)$, d'où (3.1) et la première relation de (3.3). On utilise un raisonnement analogue dans le sens rétrograde.$\diamondsuit$

En remplaçant dans les relations de (3.4) les matrices $J^f(n \Leftrightarrow 1)$ et $J^b(n \Leftrightarrow 1)$ par l'identité, on obtient deux nouvelles matrices qui seront notées $\tilde{W}^f(n)$ et $\tilde{W}^b(n)$. Dans le sens progressif par exemple, les matrices $W^f(n)$ et $\tilde{W}^f(n)$ sont identiques excepté pour les termes diagonaux tels que $J^f(n \Leftrightarrow 1)_{j,j} = 0$.

Dans ce cas, la j-ème ligne de $\Pi(n)$ est nulle par définition. Par conséquent la j-ème ligne et la j-ème colonne sont composées de 0 pour $W^f(n)$, alors que pour $\tilde{W}^f(n)$ elles coïncident avec celles de l'identité. Ainsi les inverses $W^f(n)_i^{1/2-}$ et $\tilde{W}^f(n)_i^{1/2-}$ ne diffèrent que pour ces colonnes. La relation (3.1) est encore satisfaite avec $\tilde{W}^f(n)_i^{1/2-}$, puisque les composantes de la variable du second membre associées à $J^f(n \Leftrightarrow 1)_{j,j} = 0$ sont nulles presque sûrement. Par contre l'utilisation de $W^f(n)$ dans (3.3) est nécessaire car elle assure que les composantes nulles presque sûrement de $\zeta^f(t;n)$ le restent pour des ordres supérieurs à n. En effet d'après ce qui précède, la matrice $\tilde{J}^f(n)$ associée à $\tilde{W}^f(n)$ selon (3.3) est telle que $\tilde{J}^f(n)_{j,j} = 0$ si $J^f(n)_{j,j} = 0$ et $J^f(n \Leftrightarrow 1)_{j,j} = 1$, 1 sinon. La matrice $\tilde{W}^f(n)$ permet donc de détecter uniquement les singularités d'ordre n. On dispose de résultats analogues dans le sens rétrograde.

La Proposition 3.2.1 montre en particulier que $W^f(n)$ et $W^b(n)$ définis par (3.4) sont des matrices de covariance, et donne la relation entre $J^f(n)$, $J^b(n)$ et la matrice $\Pi(n)$. Ceci implique sur la séquence des matrices d'autocorrélation partielle triangulaire $\{\Pi(n),\, n \geq 0\}$, les contraintes suivantes :

(i) La matrice $\Pi(0)$ est définie non négative et pour $n \geq 1$, les valeurs singulières de $\Pi(n)$ sont inférieures ou égales à 1.

(ii) Soient $J^f(n)$ et $J^b(n)$, $n \geq 0$, les suites de matrices déterminées à partir de $\Pi(n)$, $n \geq 0$, par les relations de (3.3) et (3.4). Alors $J^f(n)$ et $J^b(n)$ sont deux suites décroissantes de matrices (au sens où $J(n \Leftrightarrow 1) \Leftrightarrow J(n)$ est définie non négative) telles que

$$J^f(n \Leftrightarrow 1)\Pi(n)J^b(n \Leftrightarrow 1) = \Pi(n), \quad n \geq 1. \tag{3.5}$$

La décroissance des suites $J^f(n)$ et $J^b(n)$ se traduit par $J(n)J(n \Leftrightarrow 1) = J(n)$. C'est-à-dire que les lignes ou colonnes nulles de $\Pi(n)$, dues aux singularités, le sont aussi pour $\Pi(n+1)$ (cf. 3.5). Nous notons $\mathcal{D}_\Pi^T$, l'ensemble des fonctions $\Pi(\cdot)$ à valeurs matricielles définies sur $\mathbb{Z}$ telles que $\Pi(\Leftrightarrow n) = \Pi(n)^*$, et qui satisfont les contraintes ci-dessus. En effet la proposition suivante montre qu'il s'agit du domaine de variation de $\Pi(\cdot)$.

Proposition 3.2.2. *Soient $\Pi(\cdot)$ un élément de $\mathcal{D}_\Pi^T$ et $\{Z(n), n \in \mathbb{N}\}$ une suite de vecteurs aléatoires de dimension T, non corrélés entre eux, et dont les matrices de covariance sont données par la suite $\{J^f(n), n \in \mathbb{N}\}$ associée à $\Pi(\cdot)$. Alors la séquence $\{Y(n), n \in \mathbb{N}\}$ définie, pour $n = 0, 1, \ldots$, par :*

$$\zeta^f(n;n) = Z(n),$$

pour $k = n, \ldots, 1$:

$$\zeta^f(n; k \Leftrightarrow 1) = \left[W^f(k)\right]_j^{\frac{1}{2}} \zeta^f(n; k) + \Pi(k)\zeta^b(n \Leftrightarrow k; k \Leftrightarrow 1),$$
$$\zeta^b(n \Leftrightarrow k; k) = \left[W^b(k)\right]_s^{\frac{1}{2}-} \left\{\zeta^b(n; k \Leftrightarrow 1) \Leftrightarrow \Pi(k)^* \zeta^f(n; k \Leftrightarrow 1)\right\},$$
$$Y(n) = [\Pi(0)]_i^{\frac{1}{2}} \zeta^f(n; 0),\ \zeta^f(n; 0) = \zeta^b(n; 0),$$

est stationnaire et admet $\Pi(\cdot)$ *comme fonction d'autocorrélation partielle matricielle triangulaire.*

Notons que la démarche pour établir ce résultat est essentiellement la même que pour son homologue du cas non stationnaire (cf. démonstration de la Proposition 1.2.2).

Démonstration.- On suppose avoir construit la séquence stationnaire $\{Y(t), t = 0, \ldots, n \Leftrightarrow 1\}$ en fonction des variables $Z(t)$, $t = 0, \ldots, n \Leftrightarrow 1$, de telle sorte que sa fonction d'autocorrélation partielle triangulaire coïncide avec $\Pi(\cdot)$ sur le domaine $[0, \ldots, n \Leftrightarrow 1]$. De plus on dispose de la base $\left\{\zeta^b(n \Leftrightarrow k; k \Leftrightarrow 1), k = 1, \ldots, n\right\}$ de l'espace $L\left\{Y(t), t = 0, \ldots, n \Leftrightarrow 1\right\}$. Notons que ces hypothèses n'engagent à rien pour $n = 0$, mais qu'elles sont aussi vérifiées pour $n = 1$ après avoir effectué le premier pas de récurrence :

$$\zeta^f(0; 0) = \zeta^b(0; 0) = Z(0); \quad Y(0) = [\Pi(0)]_i^{\frac{1}{2}} \zeta^f(0; 0).$$

L'algorithme, à travers la première relation de récurrence, définit la variable $Y(n)$ sous la forme,

$$Y(n) = \sum_{k=1}^{n} W_0^{k-1} \Pi(k) \zeta^b(n \Leftrightarrow k; k \Leftrightarrow 1) + W_0^n \zeta^f(n; n),$$

où $W_j^k = \prod_{l=j}^{k} \left[W^f(l)\right]_i^{\frac{1}{2}}$, $j = 0, \ldots, k$, et $\zeta^f(n; n) = Z(n)$ est non corrélé avec les éléments de $L\left\{Y(t), t = 0, \ldots, n \Leftrightarrow 1\right\}$. Notons (cf. (3.7)) que les quantités W_0^k, $k \geq 0$, sont les racines , $^f(k)_i$. Les éléments $\zeta^b(n \Leftrightarrow k; k \Leftrightarrow 1)$, $k = 1, \ldots, n$, et $\zeta^f(n \Leftrightarrow 1; n \Leftrightarrow 1)$ étant parfaitement définis, on a pour $1 \leq j \leq n \Leftrightarrow 1$,

$$\begin{aligned}
E\left\{Y(n)Y(n-j)^*\right\} &= \sum_{k=1}^{n-1} W_0^{k-1} \Pi(k) E\left\{\zeta^b(n-k; k-1) Y(n-j)^*\right\}, \\
&= \sum_{k=1}^{n-1} W_0^{k-1} \Pi(k) E\left\{\zeta^b(n-1-k; k-1) Y(n-1-j)^*\right\}, \\
&= E\left\{[Y(n-1) - W_0^{n-1} \zeta^f(n-1; n-1)] Y(n-1-j)^*\right\}, \\
&= E\left\{Y(n-1)Y(n-1-j)^*\right\}.
\end{aligned}$$

La matrice de covariance de $Y(n)$ est $\Pi(0)$ car

$$E\left\{Y(n)Y(n)^*\right\} = \sum_{k=1}^{n} W_0^{k-1}\Pi(k)\Pi(k)^*[W_0^{k-1}]^* + W_0^n[W_0^n]^*,$$

avec

$W_0^k[W_0^k]^* = W_0^{k-1}W^f(k)[W_0^{k-1}]^* = W_0^{k-1}[W_0^{k-1}]^* \Leftrightarrow W_0^{k-1}\Pi(k)\Pi(k)^*[W_0^{k-1}]^*$.
La séquence $\{Y(0), \ldots, Y(n)\}$ est donc stationnaire. Compte tenu du choix de la base de $L\left\{Y(t), t = 0, \ldots, n \Leftrightarrow 1\right\}$, on a pour $j = n \Leftrightarrow 1, \ldots, 0$, ($W_j^k = I_T$ pour $k < j$)

$$\zeta^f(n;j) = \sum_{k=j+1}^{n} W_{j+1}^{k-1}\Pi(k)\zeta^b(n \Leftrightarrow k; k \Leftrightarrow 1) + W_{j+1}^n\zeta^f(n;n),$$

$$\begin{aligned} E\left\{\zeta^f(n;j)\zeta^f(n;j)^*\right\} &= \sum_{k=j+1}^{n} W_{j+1}^{k-1}\Pi(k)\Pi(k)^*[W_{j+1}^{k-1}]^* + W_{j+1}^n[W_{j+1}^n]^* \\ &= \sum_{k=j+1}^{n-1} W_{j+1}^{k-1}\Pi(k)\Pi(k)^*[W_{j+1}^{k-1}]^* + W_{j+1}^{n-1}[W_{j+1}^{n-1}]^* \\ &= \ldots = J^f(j), \end{aligned}$$

$$\left\{\zeta^f(n;j)\zeta^b(n \Leftrightarrow j \Leftrightarrow 1; j)^*\right\} = \Pi(j+1)J^b(j) = \Pi(j+1),$$

Ainsi $J^f(j)$ est la matrice de covariance de $\zeta^f(n;j)$ et $\Pi(j+1)$ est l'autocorrélation partielle triangulaire entre $Y(n)$ et $Y(n \Leftrightarrow j \Leftrightarrow 1)$. La première partie de l'hypothèse de récurrence à l'étape n est acquise. Par conséquent la deuxième relation de récurrence de l'algorithme construit effectivement la nouvelle base $\left\{\zeta^b(n \Leftrightarrow k; k), k = 0, \ \ldots, n\right\}$ de l'espace $L\left\{Y(t), t = 0, \ \ldots, n\right\}$. $\diamondsuit$

La proposition ci-dessus prouve la surjectivité dans l'ensemble $\mathcal{D}_\Pi^T$ de l'application qui à $R(\cdot)$ associe $\Pi(\cdot)$. Les relations de l'algorithme de type Levinson-Durbin, décrites dans la sous-section suivante, sont nécessaires pour assurer l'injectivité.

3.2.3 Algorithme de type Levinson-Durbin

L'algorithme de Levinson-Durbin a été généralisé aux processus vectoriels stationnaires réguliers par Whilttle [Whi63]. Cet algorithme s'étend au cas localement déterminable en utilisant les inverses généralisés (cf. [Dég88], [Pha92]). Ces procedures établissent la correspondance entre la fonction d'autocovariance $R(\cdot)$ et la fonction d'autocovariance partielle $\delta(\cdot)$. Nous proposons un algorithme de Levinson-Durbin Généralisé Vectoriel ($LDGV$) qui a la même structure que les précédents mais avec les matrices $\Pi(n)$ comme

paramètres clés au lieu de $\delta(n)$. C'est une conséquence directe de la proposition suivante. Cette dernière résulte du Théorème 1 de [DÉG90] avec $,^{-*}, ^{-}$ comme choix pour l'inverse généralisé de $,^{2}$.

Proposition 3.2.3. *Les innovations partielles satisfont, pour $n \in \mathbb{N}^*$ et $t \in \mathbb{Z}$, les relations de récurrence,*

$$\begin{aligned} e^f(t;n) &= e^f(t;n \Leftrightarrow 1) \Leftrightarrow, {}^f(n \Leftrightarrow 1)_i \Pi(n), {}^b(n \Leftrightarrow 1)_s^- e^b(t \Leftrightarrow n; n \Leftrightarrow 1), \\ e^b(t;n) &= e^b(t;n \Leftrightarrow 1) \Leftrightarrow, {}^b(n \Leftrightarrow 1)_s \Pi(n)^*, {}^f(n \Leftrightarrow 1)_i^- e^f(t+n; n \Leftrightarrow 1), \end{aligned} \tag{3.6}$$

et les matrices de covariance résiduelles sont données par,

$$\begin{aligned} , {}^{f2}(n) &= , {}^f(n \Leftrightarrow 1)_i \tilde{W}^f(n), {}^f(n \Leftrightarrow 1)_i^*, \\ , {}^{b2}(n) &= , {}^b(n \Leftrightarrow 1)_s \tilde{W}^b(n), {}^b(n \Leftrightarrow 1)_s^*. \end{aligned}$$

Les racines des matrices de covariance résiduelles à l'ordre n s'expriment en fonction de celles à l'ordre $n \Leftrightarrow 1$ et de $\Pi(n)$ à travers les relations suivantes,

$$\begin{aligned} , {}^f(n)_i &= , {}^f(n \Leftrightarrow 1)_i \tilde{W}^f(n)_i^{\frac{1}{2}}, \\ , {}^b(n)_s &= , {}^b(n \Leftrightarrow 1)_s \tilde{W}^f(n)_s^{\frac{1}{2}}. \end{aligned} \tag{3.7}$$

En effet la première égalité de (3.7) est satisfaite lorsque, dans le second membre, les colonnes ayant un zéro sur la diagonale sont nulles. Or le i-ème terme diagonal du second membre est nul si et seulement si celui de $\tilde{W}^f(n)_i^{1/2}$ ou celui de $, {}^f(n \Leftrightarrow 1)_i$ est lui-même égal à zéro. Dans le premier cas la i-ème colonne de $\tilde{W}^f(n)_i^{1/2}$ est nulle et donc celle de $, {}^f(n)_i$ aussi. Dans le second cas, on a vu que la i-ème ligne et la i-ème colonne de $\tilde{W}^f(n)$ sont celles de l'identité. Compte tenu du procédé de normalisation, il en est de même pour sa racine, d'où la première égalité. Un raisonnement analogue conduit au résultat dans le sens rétrograde. Remarquons que dans l'expression des matrices de covariance résiduelles de la Proposition 3.2.3, les matrices $\tilde{W}^f(n)$ et $\tilde{W}^b(n)$ peuvent être remplacer par $W^f(n)$ et $W^b(n)$. C'est aussi le cas pour (3.7) puisque $W^f(n)_i^{1/2-} J^f(n) = W^f(n)_i^{1/2-}$ et $W^b(n)_s^{1/2-} J^b(n) = W^b(n)_s^{1/2-}$. Cependant d'un point de vue calculatoire, il est préférable d'utiliser $\tilde{W}^f(n)$ et $\tilde{W}^b(n)$.

Les relations (3.1) et (3.2) de la Proposition 3.2.1 et celles de (3.6) de la Proposition 3.2.3 conduisent à,

$$\begin{aligned} \zeta^f(t;n) &= \tilde{W}^f(n)_i^{\frac{1}{2}-}, {}^f(n \Leftrightarrow 1)_i^- e^f(t;n), \\ \zeta^b(t;n) &= \tilde{W}^b(n)_s^{\frac{1}{2}-}, {}^b(n \Leftrightarrow 1)_s^- e^b(t;n). \end{aligned}$$

Ainsi les inverses de $, {}^f(n)_i$ et de $, {}^b(n)_s$ peuvent être obtenus de façon récursive sur l'ordre par,

$$\begin{aligned} , {}^f(n)_i^- &= \tilde{W}^f(n)_i^{\frac{1}{2}-}, {}^f(n \Leftrightarrow 1)_i^-, \\ , {}^b(n)_s^- &= \tilde{W}^b(n)_s^{\frac{1}{2}-}, {}^b(n \Leftrightarrow 1)_s^-. \end{aligned} \tag{3.8}$$

Comme pour (3.7), on peut aussi utiliser $W^f(n)$ et $W^b(n)$. Remarquons que si $W^f(n)_i^{1/2-}$ et $,^f(n \Leftrightarrow 1)_i^-$ sont des inverses tels que $J,^- = ,^-$, alors il en est de même pour $,^f(n)_i^-$. Dans (3.8), l'inverse $\tilde{W}^f(n)_i^{1/2-}$ doit satisfaire $,^-J = ,^-$, pour obtenir le même résultat. Par contre la contrainte $,^-J = ,^-$ n'est pas récursive sur l'ordre au sens où si la matrice $,^f(n \Leftrightarrow 1)_i^-$ satisfait cette condition alors ceci n'est plus nécessairement vrai pour $,^f(n)_i^-$.

Comme dans le cas non stationnaire, l'algorithme détermine, à l'aide des relations de récurrence (3.6) les coefficients matriciels (pas nécessairement uniques) de la décomposition de $e^f(k;n)$ et $e^b(k \Leftrightarrow n;n)$ sur $L\{Y(s), s = k \Leftrightarrow n, \ldots, k\}$, soit:

$$e^f(k;n) = \sum_{j=0}^{n} A^f(n,j) Y(k \Leftrightarrow j), \quad A^f(n,0) = I_T,$$

$$e^b(k \Leftrightarrow n;n) = \sum_{j=0}^{n} A^b(n,j) Y(k \Leftrightarrow n+j), \quad A^b(n,0) = I_T.$$

Algorithme 3.2.1. *(LDGV)*

$$R(0) = \Pi(0),$$

$$,^f(0)_i = \Pi(0)_i^{\frac{1}{2}}, \quad ,^f(0)_i^- = \Pi(0)_i^{\frac{1}{2}-}$$

$$,^f(0)_s = \Pi(0)_s^{\frac{1}{2}}, \quad ,^b(0)_s^- = \Pi(0)_s^{\frac{1}{2}-}$$

Pour $n = 1, 2, \ldots,$

$$\Pi(n) = ,^f(n \Leftrightarrow 1)_i^- \{R(n) + \sum_{j=1}^{n-1} A^f(n \Leftrightarrow 1, j) R(n \Leftrightarrow j)\},\ ^b(n \Leftrightarrow 1)_s^{-*}.$$

Mettre à jour les coefficients matriciels des innovations partielles d'ordre n,

$$A^f(n,n) = \Leftrightarrow,\ ^f(n \Leftrightarrow 1)_i \Pi(n),\ ^b(n \Leftrightarrow 1)_s^-,$$

$$A^b(n,n) = \Leftrightarrow,\ ^b(n \Leftrightarrow 1)_s \Pi(n)^*,\ ^f(n \Leftrightarrow 1)_i^-,$$

et pour $j = 1, \ldots, n \Leftrightarrow 1$*:*

$$A^f(n,j) = A^f(n \Leftrightarrow 1, j) + A^f(n,n) A^b(n \Leftrightarrow 1, n \Leftrightarrow j),$$

$$A^b(n,j) = A^b(n \Leftrightarrow 1, j) + A^b(n,n) A^f(n \Leftrightarrow 1, n \Leftrightarrow j).$$

Calculer les racines et les inverses de

$$\tilde{W}^f(n) = I_T \Leftrightarrow \Pi(n)\Pi(n)^*, \quad \tilde{W}^b(n) = I_T \Leftrightarrow \Pi(n)^*\Pi(n).$$

Mettre à jour les racines et les inverses des matrices de covariance résiduelles d'ordre n,

$$,^f(n)_i = ,^f(n \Leftrightarrow 1)_i \tilde{W}^f(n)_i^{\frac{1}{2}}, \quad ,^f(n)_i^- = \tilde{W}^f(n)_i^{\frac{1}{2}-},\ ^f(n \Leftrightarrow 1)_i^-,$$

$$, {}^b(n)_s = , {}^b(n \Leftrightarrow 1)_s \tilde{W}^b(n)_s^{\frac{1}{2}}, \quad , {}^b(n)_s^- = \tilde{W}^b(n)_s^{\frac{1}{2}-}, {}^b(n \Leftrightarrow 1)_s^-.$$

On retrouve la même situation que dans le cas scalaire en ce qui concerne les singularités. Les coefficients des filtres ne sont pas uniques mais choisis par l'algorithme.

Dans le cas non localement déterminable, Morf et coll. [MOR78a] ont proposé une extension de la fonction d'autocorrélation partielle matricielle. Cette fonction est obtenue avec une normalisation triangulaire inférieure dans les deux sens progréssif et rétrograde. Ils donnent également une version normalisée de l'algorithme de type Levinson qui consiste à utiliser les coefficients des innovations normalisées. Les relations de récurrences (3.1) et (3.2) permettent de donner la version normalisée de l'algorithme *LDGV*. Cette procédure a la même structure que celle de Morf et coll., mais les racines et leurs inverses du sens rétrograde sont remplacés par des matrices triangulaires supérieures. De plus, elle est aussi valide dans le cas localement déterminable.

3.3 Correspondance entre $\beta(\cdot,\cdot)$ et $\Pi(\cdot)$

La correspondance entre les fonctions $\beta(\cdot,\cdot)$ et $\Pi(\cdot)$ est biunivoque puisque ce sont deux paramétrisations de la même structure via l'association entre $X(\cdot)$ et $Y(\cdot)$. Nous donnons ci-après les relations de passage entre ces deux paramétrisations.

Compte tenu que $R(0)_{i,j} = R(i,j)$, $i,j = 1,\dots T$, la relation entre $\Pi(0)$ $(= R(0))$ et $\beta(i,j)$, $i,j = 1\dots T$, est donnée par l'Algorithme *LDG*. Pour $n \geq 1$, la correspondance entre $\Pi(n)$ et $\beta(nT+i,j)$, $i,j = 1,\dots T$, est encore assurée par cet algorithme. En effet, le procédé de normalisation conduit à $\Pi(n)_{i,j} = \left\langle \eta^f(nT+i;T+1), \eta^b(j;nT) \right\rangle$. Ainsi, sauf pour $(i,j) = (1,T)$, on n'a donc pas directement $\Pi(n)_{i,j} = \beta(nT+i,j)$, mais la relation cherchée est obtenue grâce au résultat suivant.

Proposition 3.3.1. *Pour $n \geq 1$ et pour $(i,j) \in [1,\dots T]^2$, la corrélation partielle entre $\eta^f(nT+i;T+1)$ et $\eta^b(j;nT)$ dans l'ensemble*

$$\{\eta^b(k;nT), k = j,\dots,T; \eta^f(nT+k;T+1), k = 1,\dots,i\},$$

est égale à $\rho(X(nT+i), X(j)/X(j),\dots,X(nT+i))$, c'est-à-dire à $\beta(nT+i,j)$.

Démonstration.- Afin de faciliter la compréhension de cette démonstration, on pose

$$U^{(n)}(i) = \eta^b(i;nT), \quad U^{(n)}(T+i) = \eta^f(nT+i;T+1), \quad i = 1,\dots,T.$$

Les ensembles $\mathcal{L}\{X(s),\ldots,X(t)\}$ et $\mathcal{L}\left\{U^{(n)}(s),\ldots,U^{(n)}(t)\right\}$ sont notés $\mathcal{M}_X(s;t)$ et $\mathcal{M}_U(s;t)$ respectivement avec $\mathcal{M}(s;t)=\emptyset$ lorsque $s>t$. Soient $\varepsilon_U^f(T+i;j+1)$ et $\varepsilon_U^b(j;T+i\Leftrightarrow 1)$ les erreurs de prédiction de $U^{(n)}(T+i)$ et de $U^{(n)}(j)$ sur $\mathcal{M}_U(j+1;T+i\Leftrightarrow 1)$, $i,j=1,\ldots,T$. La proposition provient du fait que,

$$\begin{aligned}\sigma^f(nT+i;T+1)\varepsilon_U^f(T+i;j+1) &= \varepsilon^f(nT+i;j+1),\\ \sigma^b(j;nT)\varepsilon_U^b(j;T+i\Leftrightarrow 1) &= \varepsilon^b(j;nT+i\Leftrightarrow 1).\end{aligned}$$

En effet, les variables $\varepsilon_U^f(T+i;j+1)$ et $\varepsilon^f(nT+i;j+1)$ (resp. $\varepsilon_U^b(j;T+i\Leftrightarrow 1)$ $\varepsilon^b(j;nT+i\Leftrightarrow 1)$) sont nulles presque sûrement lorsque $\sigma^f(nT+i;T+1)=0$ (resp. $\sigma^b(j;nT)=0$). Par suite, les erreurs de prédiction normalisées associées à $\varepsilon_U^f(T+i;j+1)$ et $\varepsilon_U^b(j;T+i\Leftrightarrow 1)$ sont égales à $\eta^f(nT+i;j+1)$ et $\eta^b(j;T+i\Leftrightarrow 1)$. Montrons maintenant que les relations ci-dessus sont bien satisfaites. Ce résultat est immédiat lorsque dans la première égalité, $j=T$ et dans la seconde, $i=1$. Par ailleurs la variable $\varepsilon_U^f(T+i;j+1)$ appartient par définition à l'ensemble $\mathcal{M}_U(j+1;T+i)$ qui est inclus dans $\mathcal{M}_X(j+1;nT+i)$ et comme $\sigma^f(nT+i;T+1)U^{(n)}(T+i)=\varepsilon^f(nT+i;T+1)$, on a

$$\sigma^f(nT+i;T+1)\varepsilon_U^f(T+i;j+1)=X(nT+i)+X^f(nT+i;j+1),$$

avec $X^f(nT+i;j+1)\in\mathcal{M}_X(j+1;nT+i\Leftrightarrow 1)$. La variable ci-dessus sera caractéristique de l'innovation $\varepsilon^f(nT+i;j+1)$ si l'on montre qu'elle est orthogonale aux éléments de $\mathcal{M}_X(j+1;nT+i\Leftrightarrow 1)$. Or tout élément de ce dernier peut se décomposer en une somme de deux termes, l'un appartenant à $\mathcal{M}_U(j+1;T+i\Leftrightarrow 1)$ et l'autre dans $\mathcal{M}_X(T+1;nT)$. L'erreur $\varepsilon_U^f(T+i;j+1)$ est par définition, orthogonale à $\mathcal{M}_U(j+1;T+i\Leftrightarrow 1)$ et l'espace $\mathcal{M}_U(j+1;T+i)$, auquel elle appartient, est orthogonal à $\mathcal{M}_X(T+1;nT)$. La variable $\sigma^f(nT+i;T+1)\varepsilon_U^f(T+i;j+1)$ est donc bien l'innovation $\varepsilon^f(nT+i;j+1)$. On procède de façon analogue pour le sens rétrograde.$\diamondsuit$

On note $R^{(n)}$ la matrice de covariance du vecteur $[\zeta^b(t\Leftrightarrow n;n\Leftrightarrow 1)^*,\zeta^f(t;n\Leftrightarrow 1)^*]^*$, soit

$$R^{(n)}=\begin{pmatrix} J^b(n\Leftrightarrow 1) & \Pi(n)^* \\ \Pi(n) & J^f(n\Leftrightarrow 1)\end{pmatrix}.$$

Cette matrice définit la restriction à $[1,\ldots,2T]^2$ d'une structure d'autocovariance non stationnaire, $R^{(n)}(t,s)=R^{(n)}_{t,s}$. D'après la proposition précédente, les autocorrélations partielles $\beta^{(n)}(t,s)$ associées à cette structure satisfont

$$\beta^{(n)}(T+i,j)=\beta(nT+i,j),\quad i,j=1,\ldots,T.$$

Par ailleurs les matrices $J^b(n\Leftrightarrow 1)$ et $J^f(n\Leftrightarrow 1)$ étant des matrices diagonales, on a pour $i,j=1,\ldots,T$,

$$\beta^{(n)}(i,j)=J^b(n\Leftrightarrow 1)_{i,j},\quad \beta^{(n)}(T+i,T+j)=J^f(n\Leftrightarrow 1)_{i,j}.$$

Ainsi l'Algorithme LDG qui donne la correspondance entre $R^{(n)}$ et $\beta^{(n)}(\cdot,\cdot)$, établit celle entre $\Pi(n)$ et $\beta(nT+i,j)$, $i,j=1,\ldots,T$. Remarquons que les matrices $J^b(n \Leftrightarrow 1)$ et $J^f(n \Leftrightarrow 1)$ sont utilisées dans cette correspondance. Cependant, ces matrices peuvent être remplacées dans $R^{(n)}$ par l'identité. Ce qui équivaut à poser $\beta^{(n)}(i,i)=1$, $i=1,\ldots,2T$. Dans ce cas, les égalités $\Pi(n)_{i,l}=\Pi(n)_{j,l}=0$, conduisent aux égalités $\beta(nT+l,j)=\beta(nT+i,j)=0$, $l=1,\ldots,T$, qui assurent que la fonction $\beta(\cdot,\cdot)$ reste dans $\mathcal{D}_\beta^T$, et inversement. Ainsi la bijection entre $\beta(\cdot,\cdot)$ et $\Pi(\cdot)$ est donnée par l'intermédiaire de celles entre $\Pi(n)$ et $\{\beta(nT+i,j), i,j=1,\ldots,T\}$, n appartenant à $\mathbb{N}$. Notons que l'algorithme de Pham [PHA92], qui détermine la correspondance entre $\tilde{\Pi}(n)$ et $\beta(nT+i,j)$, correspond à l'Algorithme LDG restreint à la situation précédente dans laquelle les matrices $J^b(n \Leftrightarrow 1)$ et $J^f(n \Leftrightarrow 1)$ sont remplacées par l'identité.

On dispose donc d'une autre approche pour donner la relation entre $R(\cdot)$ et $\Pi(\cdot)$. En effet, on peut montrer que l'application qui à $\beta(\cdot,\cdot)$ associe $\Pi(\cdot)$ établit effectivement une bijection entre les ensembles $\mathcal{D}_\beta^T$ et $\mathcal{D}_\Pi^T$. L'intérêt de cette approche est que la relation entre $R(\cdot)$ et $\Pi(\cdot)$ est réalisée de façon scalaire.

La définition de $\beta(n)_{tr}$ en termes de variables relatives au processus $X(\cdot)$ est plus lourde que celle de $\Pi(n)$. Comme les composantes des innovations normalisées nulles presque sûrement sont éliminées, le (i,j)-ème terme de $\beta(n)_{tr}$ est donné par $\left\langle \eta^f(nT+\tilde{\imath};T+1), \eta^b(\tilde{\jmath};nT) \right\rangle$, avec $(\tilde{\imath},\tilde{\jmath})$ éventuellement différent de (i,j). Pour obtenir la correspondance entre $\beta(nT+i,j)$, $i,j=1,\ldots T$, et $\beta(n)_{tr}$, il faudrait connaître pour chaque n, les applications qui à i et j variant de 1 à $r(n)$, associent $\tilde{\imath}$ et $\tilde{\jmath}$ dans $[1,\ldots,T]$. Ce qui revient en fait à se donner les matrices $J^f(n \Leftrightarrow 1)$ et $J^b(n \Leftrightarrow 1)$. La correspondance de $\beta(nT+i,j)$, $i,j=1,\ldots T$, avec $\beta(n)_{tr}$ ne peut donc pas être réalisée indépendemment de n, contrairement à celle avec $\Pi(n)$.

3.4 Modèles autorégressifs périodiques

Les modèles autorégressifs périodiques (ARP) ont été introduits par Jones et coll. [JON67]. Ces modèles sont définis comme à la section 1.4 mais avec des paramètres périodiques dûs à la périodicité de leur structure. Cependant, il n'est pas fait l'hypothèse que $\varepsilon(\cdot)$ soit le processus d'innovation. Pour les mêmes raisons que celles évoquées à la section 1.4, nous retenons la définition suivante.

Définition 3.4.1. *Un processus $X(\cdot)$ est dit autorégressif périodique de période T et d'ordre $(p_1,\ldots,p_T)$, noté $ARP(p_1,\ldots,p_T)$, si il est périodiquement corrélé de période T et si pour tout $t \in \mathbb{Z}$, il existe des constantes*

$a_t(k)$, $k = 1, \ldots, p_t$, *telles que*

$$\sum_{k=0}^{p_t} a_t(k) X(t \Leftrightarrow k) = \varepsilon(t), \quad a_t(0) = 1, \quad a_t(p_t) \neq 0, \tag{3.9}$$

où $\varepsilon(\cdot)$ *est le processus d'innovation et* $a_{t+T}(k) = a_t(k)$, $k = 1, \ldots, p_t$.

Contrairement à [JON67], nous n'indiquons pas dans la définition que la variance du processus $\varepsilon(\cdot)$ satisfait $\sigma^2_{t+T} = \sigma^2_t$, car cette contrainte est automatiquement vérifiée lorsque $\varepsilon(\cdot)$ est le processus d'innovation. Par ailleurs, dans [JON67], il n'est pas précisé si la variance σ^2_t peut s'annuler. Nous ne faisons également aucune hypothèse sur σ^2_t, permettant ainsi au processus $X(\cdot)$ d'être localement déterminable. Comme dans la situation générale non stationnaire, l'unicité d'une telle représentation est assurée dans le cas non localement déterminable et peut être obtenue moyennant des conventions sinon. Par contre, concernant l'existence, on peut donner une condition suffisante sur les paramètres du modèle grâce à la correspondance de ces modèles avec les processus vectoriels autorégressifs stationnaires. Ce problème sera évoqué juste après la description de cette correspondance. L'analogue de la définition précédente pour les modèles autorégressifs vectoriels stationnaires est la suivante.

Définition 3.4.2. *Un processus* $Y(\cdot)$ *est dit autorégressif stationnaire d'ordre* p, *noté* $VAR(p)$, *si il est stationnaire et admet la représentation,*

$$\sum_{k=0}^{p} A(k) Y(t \Leftrightarrow k) = e(t), \quad A(0) = I_T, \tag{3.10}$$

où $e(\cdot) = \{e(t), t \in \mathbb{Z}\}$ *est une suite de vecteurs aléatoires non corrélés, centrés, de matrice de covariance* $, \frac{2}{e}$, *telle que* $e(t)$ *soit non corrélé avec* $Y(s)$, $s < t$.

Dans [PAG78], les représentations (3.9) et (3.10) sont considérées dans le cas non localement déterminable (i.e $\sigma^2_t > 0$, $t = 1, \ldots, T$, ou de façon équivalente $, \frac{2}{e}$ non singulière) et sans exiger que la variable $\varepsilon(t)$ (resp. $e(t)$) soit non corrélée avec le passé de $X(t)$ (resp. $Y(t)$). Ces deux représentations sont équivalentes ([PAG78] : Théorème 1, p. 1311). En effet soit $X(\cdot)$ un processus qui admet la représentation (3.9). Alors le processus T-dimensionnel $Y(\cdot)$ associé à $X(\cdot)$, satisfait

$$LY(t) + \sum_{k=0}^{p} A'(k) Y(t \Leftrightarrow k) = e'(t),$$

où $p = \max_j [(p_j \Leftrightarrow j)/T] + 1$ et $[x]$ désigne la partie entière de x. La variable $e'(t)$ est donnée par le vecteur $[\varepsilon(1 + (t \Leftrightarrow 1)T)^*, \ldots, \varepsilon(T + (t \Leftrightarrow 1)T)^*]^*$ de

matrice de covariance $D^2 = \text{diag}\{\sigma_1^2, \ldots, \sigma_T^2\}$, les matrices L et $A'(\cdot)$ de taille $T \times T$ sont définies par,

$$\begin{array}{l} L_{ij} = a_i(i \Leftrightarrow j) \text{ pour } j \leq i, \text{ 0 sinon} \\ A'(k)_{ij} = a_i(kT + i \Leftrightarrow j), \ k = 1, \ldots, p \end{array} \qquad i, j = 1, \ldots, T.$$

Par conséquent $Y(\cdot)$ satisfait (3.10) avec $A(k) = L^{-1}A'(k)$, $k = 1, \ldots, p$, et $,_e^2 = L^{-1}D^2L^{-1*}$. Réciproquement, l'unique décomposition de Cholesky $L^{-1}D^2L^{-1*}$ de $,_e^2$ donne le passage entre la représentation (3.10) et (3.9). Ce résultat s'étend sans difficulté au cas localement déterminable. Les coefficients de L et $A'(\cdot)$ ne sont plus définis de façon unique et certains éléments de D^2 sont nuls. De plus, l'égalité $e(t) = L^{-1}e'(t)$ montre que $\varepsilon(\cdot)$ est le processus d'innovation de $X(\cdot)$ si et seulement si $e(\cdot)$ est celui de $Y(\cdot)$. On a donc l'équivalence entre les Définitions 3.4.1 et 3.4.2, soit le résultat suivant.

Théorème 3.4.1. *Le processus $X(\cdot)$ est ARP de période T et d'ordre $(p_1, \ldots, p_T)$ si et seulement si le processus T-dimensionnel $Y(\cdot)$ associé, est $VAR(p)$ avec $p = \max_j [(p_j \Leftrightarrow j)/T] + 1$.*

Remarquons que, selon la même démarche, on étend cette correspondance aux innovations partielles $e^f(t;n)$ et $\varepsilon^f(i + (t \Leftrightarrow 1)T; 1 + (t \Leftrightarrow n \Leftrightarrow 1)T)$, $i = 1, \ldots, T$. Plus précisément, on a

$$\begin{array}{ll} ,^{f2}(n) = [L_n^f]^{-1} D_n^{f2} [L_n^f]^{-1*}, & n \in I\!N, \\ A^f(n,k) = [L_n^f]^{-1} A'^f(n,k), & k = 1, \ldots, p, \end{array} \tag{3.11}$$

avec $D_n^{f2} = \text{diag}\{\sigma_1^{f2}(nT), \sigma_2^{f2}(nT+1), \ldots, \sigma_T^{f2}((n+1)T \Leftrightarrow 1)\}$, et

$$\begin{array}{l} \{L_n^f\}_{ij} = a_i^f(nT \Leftrightarrow 1 + i, i \Leftrightarrow j) \text{ si } j \leq i, \text{ 0 sinon} \\ \{A'^f(n,k)\}_{ij} = a_i^f(nT \Leftrightarrow 1 + i, kT + i \Leftrightarrow j), \ k = 1, \ldots, n \end{array} \qquad i, j = 1, \ldots, T.$$

Le produit $[L_n^f]^{-1} D_n^f$ fournit la racine $,^f(n)_i$ et $[D_n^f]^+ L_n^f$ est un candidat possible pour son inverse. L'inverse contraint $,^- J = ,^-$ est donné par le jeu de coefficients qui ne pondèrent pas la variable $X(i + (t \Leftrightarrow 1)T)$ lorsque $\varepsilon^f(i + (t \Leftrightarrow 1)T; 1 + (t \Leftrightarrow n \Leftrightarrow 1)T) = 0$ *p.s.* Rappelons que ce nouveau jeu de coefficients peut être obtenu à partir de celui qu'impose l'Algorithme *LDP* (cf. section 1.3). Dans le sens rétrograde, on dispose de relations analogues avec une matrice triangulaire supérieure unitaire L_n^b.

Il est clair que les paramètres du modèle *ARP* ne sont pas arbitraires. D'après le théorème précédent, l'existence d'un processus *ARP*, associé à un jeu de paramètres, se ramène au problème analogue du cas vectoriel stationnaire. Une condition suffisante (et nécessaire dans le cas non localement déterminable) de cette existence, est que le polynôme $\det[\sum_{k=0}^p A(k) z^k]$ n'ait pas de racines de module inférieur ou égal à 1 (cf. [HAN70]). Cette condition ne se traduit pas de façon simple sur les coefficients du modèle *ARP*. En

particulier, les filtres $a_i(\cdot)$, $i = 1, \ldots, T$, ne sont pas nécessairement stables au sens usuel. L'identification de ces paramètres est donc un point délicat.

C'est pourquoi, comme dans la situation plus générale, nous préférons retenir la caractérisation de ces modèles par leur fonction d'autocorrélation partielle. En effet d'après la Proposition 1.4.1, il est clair qu'un processus $X(\cdot)$ est ARP de période T et d'ordre $(p_1, \ldots, p_T)$ si et seulement si sa fonction d'autocorrélation partielle satisfait pour $t = 1, \ldots, T$,

$$\beta_t(p_t) \neq 0, \quad \text{et} \quad \beta_t(k) = 0, \quad k > p_t.$$

Les paramètres (pas nécessairement uniques) du modèle sont alors donnés par l'Algorithme LDP. De façon analogue, les modèles $VAR(p)$ sont caractérisés par,

$$\Pi(p) \neq 0_T, \quad \Pi(k) = 0_T, \quad k > p.$$

Soulignons le fait que la fonction $\Pi(\cdot)$ ne permet pas d'identifier directement les ordres $p_1, \ldots, p_T$. Par exemple lorsque l'entier n est tel que $1 < nT + i \Leftrightarrow p_i \leq T$, les coefficients $\Pi(n)_{i,j}$, $j < nT + i \Leftrightarrow p_i$, ne sont pas nécessairement nuls.

Nous donnons ci-après l'analogue des équations de Yule-Walker pour les processus périodiquement corrélés. Soit $X(\cdot)$ un processus $ARP(p_1, \ldots, p_T)$, alors sa fonction d'autocovariance satisfait, pour $t \in \mathbb{Z}$, $s \leq t$,

$$R(t, s) + \sum_{k=0}^{p_t} a_t(k) R(t \Leftrightarrow k, s) = \sigma_t^2 \delta_{t,s},$$

où $\delta_{t,s} = 1$ si $t = s$, 0 sinon. Notons que ces équations peuvent comporter moins d'inconnues que celles associées à la représentation vectorielle. En effet, désignons par $ARP(p, \max)$, les modèles ARP dont les ordres sont donnés par $p_i = pT + i \Leftrightarrow 1$, $i = 1, \ldots, T$. Ces modèles sont les seuls pour lesquels le nombre de paramètres est égal à celui du processus vectoriel associé $(\sum_{i=1}^{T} p_i + T = pT^2 + T(T+1)/2)$. Pour les autres processus ARP, le nombre de paramètres est strictement inférieur et cette différence peut être très importante. Par exemple, une diminution des ordres d'un $ARP(p, \max)$, en conservant l'un d'eux, n'entraine aucune modification sur l'ordre du modèle vectoriel et donc sur le nombre de ses paramètres. Pour la plupart des situations, il semble donc plus intéressant de considérer le modèle ARP.

D'après ce qui précède, les modèles ARP peuvent être paramétrisés de façon équivalente par les premiers coefficients soit d'autocovariance soit d'autocorrélation partielle. Notons cependant une différence fondamentale entre ces deux approches. En effet partant des quantités $\beta_t(k)$, $t = 1, \ldots, T$, $k \leq p_t$, il existe toujours un modèle $ARP(p_1, \ldots, p_T)$ dont la structure soit donnée par ces coefficients du moment que ces derniers satisfont les contraintes de $\mathcal{D}_\beta^T$. Pour cela il suffit de prolonger la fonction $\beta(\cdot, \cdot)$ à zéro

au delà de l'ordre du modèle puisque ce prolongement respecte les contraintes de $\mathcal{D}_\beta^T$. Contrairement au cas stationnaire, ceci n'est plus vrai en terme d'autocovariance. La raison est que le prolongement de $R(\cdot,\cdot)$ par les équations de Yule-Walker ne conserve pas toujours le caractère positif de cette fonction. En effet soit $E(p_1,\ldots,p_T)$ le domaine associé. S'il s'agit d'un domaine de bijection $E_c(p_1,\ldots,p_T)$, le modèle ARP est parfaitement défini, c'est celui dont la fonction d'autocorrélation partielle correspond à $R(\cdot,\cdot)$ sur ce domaine et s'annule à l'extérieur. Dans la situation générale, on peut vérifier par construction l'existence d'un modèle ARP de la façon suivante. Soient $E_c(p'_1,\ldots,p'_T)$ le plus petit domaine contraint contenant $E(p_1,\ldots,p_T)$, $\tilde{R}(\cdot,\cdot)$ et $\tilde{\beta}(\cdot,\cdot)$, les fonctions associées à l'éventuelle solution. Les valeurs de $\tilde{R}(\cdot,\cdot)$ sont données par $R(\cdot,\cdot)$ sur $E(p_1,\ldots,p_T)$ et la fonction $\tilde{\beta}(\cdot,\cdot)$ doit être nulle sur $E_c(p'_1,\ldots,p'_T)\backslash E(p_1,\ldots,p_T)$. Alors il est possible de déterminer les valeurs manquantes de $\tilde{R}(\cdot,\cdot)$ et de $\tilde{\beta}(\cdot,\cdot)$ sur $E_c(p'_1,\ldots,p'_T)$. En effet on calcule, par l'Algorithme LDP, $\tilde{\beta}(t,s)$ lorsque $(t,s)\in E(p_1,\ldots,p_T)$, $\tilde{R}(t,s)$ sinon. Ainsi le modèle $ARP(p_1,\ldots,p_T)$ existe si et seulement si $\tilde{\beta}(\cdot,\cdot)$ est un élément de $\mathcal{D}_\beta^T$, garantissant ainsi le caractère positif de la fonction $\tilde{R}(\cdot,\cdot)$.

3.5 Extension par la méthode du maximum d'entropie

L'extension d'une suite de matrices d'autocovariance $R_0,\ldots,R_n$, est un problème classique dans le domaine de l'estimation spectrale. La méthode du maximum d'entropie a été proposée par Burg [Bur75] dans le cas régulier et étendue à certaines situations localement déterminables par Inouye [Ino84]. La solution équivaut à ajuster le modèle $VAR(n)$ dont la structure coïncide avec $R_0,\ldots,R_n$. Les paramètres de ce modèle sont solution des équations de Yule-Walker et peuvent être obtenus en utilisant l'Algorithme $LDGV$. Dans le cadre périodiquement corrélé, la formulation du problème d'extension n'est pas unique. Elle dépend du domaine sur lequel la fonction $R(\cdot,\cdot)$ est donnée.

3.5.1 Le problème d'extension

Parmi ces domaines, l'ensemble $[1,\ldots,N]^2$ semble le plus naturel et correspond à celui du vectoriel lorsque $N=(n+1)T$. De façon générale, considérons un ensemble de la forme $E_c(p_1,\ldots,p_T)$ introduit dans la première section de ce chapitre. Soient $R(t,s)$, $(t,s)\in E_c(p_1,\ldots,p_T)$, les valeurs d'une fonction périodique de période T. C'est-à-dire que la fonction

$R(\cdot,\cdot)$ est connue en tout point de la forme $(kT+i, kT+i \Leftrightarrow j)$, $k \in \mathbb{Z}$, $i = 1, \ldots, T$, $j = 0, \ldots, p_i$. Nous souhaitons déterminer les conditions pour que ces valeurs soient celles d'une fonction d'autocovariance et lorsque ces conditions sont remplies, la façon de déterminer l'ensemble des extensions d'une telle fonction. Lorsqu'on veut résoudre ce problème directement en termes d'autocovariances, on se heurte à deux difficultés. En premier lieu, le domaine $E_c(p_1, \ldots, p_T)$ doit être de la forme $[1, \ldots, N]^2$ pour pouvoir considérer la matrice $R_{1,N}$ (cf. sous-section 1.2.1). En second lieu, on retrouve la difficulté pour montrer que cette matrice est définie non négative. Par contre l'utilisation de la paramétrisation par les autocorrélations partielles permet de résoudre ce problème de manière élégante. En effet l'Algorithme LDP fournit les coefficients $\beta_t(n)$, $t = 1, \ldots, T$, $n = 0, \ldots, p_t$, à partir des valeurs de $R(\cdot,\cdot)$ sur $E_c(p_1, \ldots, p_T)$. Ainsi ces dernières valeurs seront celles prises par une fonction d'autocovariance si les quantités $\beta_t(n)$ satisfont les contraintes de $\mathcal{D}_\beta^T$ puisque la correspondance entre $R(\cdot,\cdot)$ et $\beta(\cdot,\cdot)$ est biunivoque sur un tel domaine. Notons cependant que dans le cas localement déterminable, il faut prendre soin de vérifier que les contraintes sur $R(\cdot,\cdot)$ dues aux singularités sont bien satisfaites. En effet, si on considère par exemple la situation $|\beta_i(k)| = 1$, l'Algorithme LDP donne

$$\sigma_i^f(k+n) = \sigma^b_{(i+n) \bmod T}(k+n) = 0, \quad n \geq 0,$$

et la convention $0^{-1} = 0$ conduit à

$$\beta_i(k+n+1) = \beta_{(i+n+1) \bmod T}(k+n+1) = 0, \quad n \geq 0.$$

Ainsi les coefficients $\beta_t(n)$ satisfont les contraintes de $\mathcal{D}_\beta^T$ dues aux singularités. Dans ce cas, les valeurs $R(t,s)$, $(t,s) \in E_c(p_1, \ldots, p_T)$, seront bien celles associées aux coefficients $\beta_t(n)$, si elles satisfont :

$$R(l, l \Leftrightarrow k \Leftrightarrow n \Leftrightarrow 1) + \sum_{j=1}^{k+n} a_l^f(k+n, j) R(l \Leftrightarrow j, l \Leftrightarrow k \Leftrightarrow n \Leftrightarrow 1) = 0,$$

pour $l = i, (i+n+1) \bmod T$. Quand les coefficients $R(t,s)$ sont ceux d'une fonction d'autocovariance, l'ensemble des extensions d'une telle fonction est alors décrit à travers les extensions des autocorrélations partielles $\beta_t(n)$ restant dans $\mathcal{D}_\beta^T$.

3.5.2 La méthode du maximum d'entropie

Dans cette sous-section, nous proposons d'étendre la méthode du maximum d'entropie à la situation précédente. Dans le cas stationnaire (scalaire ou vectoriel) et lorsque la structure ne présente pas de singularité d'ordre fini, cette méthode a été proposée par Burg [Bur75]. Elle consiste à choisir la

solution du problème d'extension des coefficients $R_0, \ldots, R_n$, pour laquelle le déterminant de la matrice de covariance du processus d'innovation est maximum. En effet, la méthode sélectionne la densité spectrale P qui maximise la quantité

$$\int_{-\pi}^{\pi} \ln \det P(\lambda) d\lambda,$$

sous les contraintes,

$$\int_{-\pi}^{\pi} e^{ik\lambda} P(\lambda) d\lambda = R_k, \quad k = \Leftrightarrow n, \ldots, n.$$

Par ailleurs, on a (cf. [WIE57] : Théorème 7.10, p. 145)

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \det 2\pi P(\lambda) d\lambda = \ln \det , \frac{2}{e},$$

où $, \frac{2}{e}$ est la matrice de covariance du processus d'innovation associé à la structure définie par P. Notons que les covariances résiduelles $, ^{f2}(n)$, $n \geq 0$, constituent une suite de matrices croissantes dont la limite est donnée par $, \frac{2}{e} = \lim_{n \to +\infty} , ^{f2}(n)$. D'après la relation de (3.11) et les propriétés du déterminant, on a

$$\det , \frac{2}{e} = \lim_{n \to +\infty} \det D_n^{f2} = \prod_{k=1}^{T} \sigma_k^2(\infty),$$

où $\sigma_k^2(\infty)$ est la variance du processus d'innovation associé à la structure périodique à un instant t tel que $t \mod T = k$. Ainsi dans le cadre périodique, la méthode du maximum d'entropie consiste à choisir parmi les structures qui sont données par $R(\cdot, \cdot)$ sur $E_c(n, \max)$, celle pour laquelle le produit

$$\prod_{k=1}^{T} \sigma_k^2(\infty),$$

est maximum. Il est clair que l'on peut étendre cette méthode sans difficulté à un domaine quelconque de la forme $E_c(p_1, \ldots, p_T)$. Dans le cas localement déterminable, il existe $(t, s) \in E_c(p_1, \ldots, p_T)$ tel que $|\beta_t(t \Leftrightarrow s)| = 1$ et par suite $\sigma_t^2(\infty) = 0$ pour n'importe quelle solution du problème d'extension. C'est pourquoi on propose d'étendre la méthode du maximum d'entropie de la façon suivante. Parmi les solutions du problème d'extension, cette méthode consiste à retenir celle pour laquelle la quantité

$$\prod_{k \in S} \sigma_k^2(\infty), \quad S = \left\{ k \in [1, \ldots, T], \sigma_k^{f2}(p_k) \neq 0 \right\},$$

est maximum. Ainsi on retrouve le résultat analogue à la méthode classique, à savoir que la solution est ARP d'ordre au plus $(p_1, \ldots, p_T)$. En effet d'après la Proposition 1.2.3 la variance $\sigma_k^2(\infty)$ du processus d'innovation pour un

processus périodiquement corrélé est donnée par

$$\sigma_k^2(\infty) = \beta_k(0) \prod_{j=0}^{+\infty} \left[1 \Leftrightarrow |\beta_k(j)|^2\right].$$

Ainsi cette méthode est équivalente à maximiser séparément chaque variance $\sigma_k^2(\infty)$, $k \in S$. De plus pour un tel entier k on a

$$\sigma_k^2(\infty) \leq \beta_k(0) \prod_{j=0}^{p_k} \left[1 \Leftrightarrow |\beta_k(j)|^2\right],$$

avec égalité si et seulement si $\beta_k(j) = 0$ pour $j \geq p_k$.

Nous montrons ci-après que notre méthode coïncide également avec l'extension proposée par Inouye [Ino84] dans le cas vectoriel stationnaire. L'auteur se place dans la situation localement déterminable où il existe un processus purement indéterminable associé à la solution du problème d'extension des matrices $R_0, \ldots, R_n$. Il utilise un critère spectral et montre que la solution maximisant l'entropie est celle pour laquelle la matrice de covariance du processus d'innovation $,_e^2$ satisfait

$$,_e^2 \geq \tilde{},_e^2,$$

où cette inégalité signifie que la matrice $,_e^2 \Leftrightarrow \tilde{},_e^2$ est définie non négative et où $\tilde{},_e^2$ est la matrice de covariance du processus d'innovation associée à n'importe quelle autre solution. Cependant nous avons constaté une erreur dans la preuve de cette équivalence sans pour autant pouvoir affirmer qu'elle soit fausse. Par ailleurs, le critère portant sur les matrices de covariances équivaut à ajuster le modèle $VAR(n)$ dont la structure coïncide avec $R_0, \ldots, R_n$. La solution que nous proposons est telle que $,_e^2 = ,^{f2}(n)$ et par conséquent est identique à celle de Inouye.

3.5.3 Fonctions de cyclo-corrélation

On observe dans [Alp95] une approche différente pour résoudre un problème d'extension analogue à celui évoqué plus haut, mais formulé en termes de fonctions de cyclo-corrélation et considéré uniquement dans le cas non localement déterminable. La structure au second ordre d'un processus périodiquement corrélé de période T est de la forme [Gla61],

$$R(t+n,t) = \sum_{k=0}^{T-1} B_k(n) e^{\frac{2\pi i k t}{T}}, \quad (t,n) \in \mathbb{Z}^2.$$

Les fonctions $B_k(\cdot)$, $k = 0, \ldots, T \Leftrightarrow 1$, sont appelées fonctions de cyclo-corrélation. Inversement, ces fonctions sont données par, $k = 0, \ldots, T \Leftrightarrow 1$,

$$B_k(n) = \frac{1}{T} \sum_{t=0}^{T-1} R(t+n, t) e^{-\frac{2\pi i k t}{T}}, \quad n \in \mathbb{Z}.$$

Pour des raisons pratiques, la définition des fonctions $B_k(\cdot)$, $k = 0, \ldots, T \Leftrightarrow 1$, est prolongée à n'importe quel entier k par l'égalité $B_{k+T}(\cdot) = B_k(\cdot)$. Dans [ALP95], les auteurs considèrent l'extension de telles fonctions à partir des données $B_k(n)$, $k = 0, \ldots, T \Leftrightarrow 1$, $0 \leq n \leq N$. Ces quantités représentent les premières valeurs prises par des fonctions de cyclo-corrélation si et seulement si [GLA61] la séquence des matrices $B(n)_{l,j=0,\ldots,T-1}$, $n = 0, \ldots, N$, définies par

$$B(n)_{l,j} = B_{j-l}(n) e^{-\frac{2\pi i l n}{T}}, \tag{3.12}$$

est celle des premières matrices d'autocovariance d'un processus T-dimensionnel stationnaire. L'ensemble des extensions de telles fonctions est alors décrit par l'intermédiaire de celles de $B(n)$, $n = 0, \ldots, N$, respectant la forme de (3.12). Cet ensemble est caractérisé analytiquement grâce aux polynômes de Szegö matriciels (cf. [DEL79]). Cette procédure consiste donc à traiter un problème d'extension d'une séquence particulière de matrice d'autocovariance. Les auteurs retrouvent le modèle ARP comme solution du critère du maximum d'entropie dans ce cas particulier. Cette approche est clairement plus délicate que celle que nous proposons. Par ailleurs, ce problème est équivalent à celui considéré à la sous-section 3.5.1 avec un domaine $E_c(p_1, \ldots, p_T)$ tel que $p_i = N$, $i = 1, \ldots, T$. L'utilisation des autocorrélations partielles permet de retrouver leur solution de façon immédiate et de l'étendre au cas localement déterminable.

Chapitre 4

Estimation des modèles ARP

Ce chapitre est consacré à l'estimation de la structure au second ordre des modèles ARP. On trouve essentiellement trois méthodes. Celle de Pagano [Pag78] constitue une extension de la méthode de Yule-Walker. Celles de Boshnakov [Bos94] et de Sakai [Sak82] sont deux extensions différentes de la méthode de Burg [Bur75]. Nous proposons d'étendre deux autres méthodes aux processus périodiquement corrélés. La première est celle des AutoCorrélations Partielles Empiriques ($ACPE$) introduite par Dégerine dans le cas stationnaire scalaire ([Dég86], [Dég93]). La seconde procédure est celle des Rapports d'Energies Résiduelles (RER) présentée par Dickinson également dans la situation scalaire [Dic78]. Compte tenu de l'équivalence entre les modèles ARP et VAR, il est intéressant de considérer également les méthodes vectorielles, c'est-à-dire qui portent sur l'estimation de la structure du processus VAR. Comme pour la situation périodique, ce sont des extensions des méthodes classiques du cas stationnaire scalaire. Elles sont comparables dans leur conception à celles périodiques. On trouve les homologues de la méthode de Yule-Walker et RER [Dic79]. Celle d'$ACPE$ a été aussi présentée dans le cadre vectoriel (cf. [Dég92], [Pha92], [Dég94]). Pour les techniques de Burg, la méthode de Nuttall [Nut76] (cf. aussi [Str77]) correspond à celle de Boshnakov [Bos94], et celle de Morf et coll. [Mor78b] à celle de Sakai [Sak82].

Dans la première section, nous présentons les méthodes scalaires existantes et donnons la façon d'étendre celles d'$ACPE$ et RER à la situation périodique. Les procédures de type Burg et $ACPE$ sont regroupées dans une même méthodologie permettant de facilement les comparer sur le plan de la conception. Les approches vectorielles font l'objet de la deuxième section. Pour chacune d'entre elles, on étudie l'équivalence avec son homologue périodique afin de vérifier si la structure périodique résultante est la même. Dans la dernière section, les méthodes sont comparées par simulation.

4.1 Les approches scalaires

Il s'agit d'estimer les paramètres d'un modèle ARP de période T et d'ordre $(p_1, \ldots, p_T)$ à valeurs complexes, au vu d'une séquence de longueur m, $X(1), \ldots, X(m)$. On se place dans la situation où l'ordre ainsi que la période du modèle sont connus. Les méthodes vont consister à estimer soit les premiers coefficients de la fonction $R(\cdot,\cdot)$ pour Yule-Walker, soit ceux de $\beta(\cdot,\cdot)$ pour les deux extensions de type Burg, $ACPE$ et RER. L'estimation des paramètres du modèle est alors obtenue grâce à l'Algorithme LDP (cf. section 3.4).

4.1.1 Méthode de Yule-Walker

Comme dans le cas stationnaire, on introduit la fonction d'autocovariance empirique $\hat{R}_m(\cdot,\cdot)$, définie pour tout $u, v \in \mathbb{Z}$, par,

$$\hat{R}_m(v,u) = \frac{1}{\left[\frac{m-1}{T}\right]+1} \sum_{k=-\infty}^{+\infty} X(v+kT)\overline{X(u+kT)},$$

en posant $X(t) = 0$ pour $t \notin \{1, \ldots, m\}$. Notons que la quantité $\left[\frac{m-1}{T}\right]+1$ est le nombre maximum de produits non nuls intervenant dans les sommes ci-dessus. Ainsi seuls les estimateurs des variances $R(t,t)$ pour t tel que $t \bmod T = 1, \ldots, m \bmod T$, seront sans biais. On retient cet estimateur biaisé afin que la fonction d'autocovariance empirique satisfasse les contraintes de la fonction $R(\cdot,\cdot)$, c'est-à-dire soit de type positif. En effet pour tout $u, v \in \mathbb{Z}$, $u \leq v$, on a

$$\begin{bmatrix} \hat{R}_m(u,u) & \hat{R}_m(u,u+1) & \ldots & \hat{R}_m(u,v) \\ \hat{R}_m(u+1,u) & \hat{R}_m(u+1,u+1) & \ldots & \hat{R}_m(u+1,v) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}_m(v,u) & \hat{R}_m(v,u+1) & \ldots & \hat{R}_m(v,v) \end{bmatrix} =$$

$$\frac{1}{\left[\frac{m-1}{T}\right]+1} \sum_{k=-\infty}^{+\infty} \begin{pmatrix} X(u+kT) \\ X(u+1+kT) \\ \vdots \\ X(v+kT) \end{pmatrix} \begin{pmatrix} \overline{X(u+kT)} & \overline{X(u+1+kT)} & \ldots & \overline{X(v+kT)} \end{pmatrix},$$

La méthode de Yule-Walker a été étendue aux processus périodiquement corrélés par Pagano [PAG78]. Elle consiste à ajuster le modèle $ARP(p_1, \ldots, p_T)$, dont la structure est donnée par $\hat{R}_m(\cdot,\cdot)$ sur $E(p_1, \ldots, p_T)$. Contrairement au cas stationnaire, l'existence du modèle ARP n'est pas toujours assurée sauf lorsque

$$p_{i+1} \leq p_i + 1, \quad i = 1, \ldots, T.$$

En effet la structure définie par les premiers coefficients de $\hat{R}_m(\cdot,\cdot)$ n'est pas toujours celle d'un modèle $ARP(p_1,\ldots,p_T)$. On a constaté à la section 3.4 que selon l'ordre du modèle, il n'existe pas nécessairement une solution ARP dont la structure coïncide avec ces premiers coefficients. Notons que le procédé introduit dans cette section pour vérifier l'existence du modèle ARP, permet simultanément d'obtenir les estimations des coefficients d'autocorrélation partielle et des paramètres du modèle lorsqu'il est défini.

4.1.2 Autocorrélations partielles empiriques

Dans cette sous-section, nous étendons aux processus périodiquement corrélés la méthode $ACPE$ introduite par Dégerine ([DÉG86], [DÉG93]). Cette méthode est basée sur une analyse géométrique naturelle des observations. En effet, pour estimer les coefficients $\beta(k)$, $k=1,\ldots,p$, on considère toute les séquences de longueur $k+1$ dont on dispose. Ce qui amène à introduire le sous-espace vectoriel de C^{m-k} engendré par les vecteurs

$$\begin{array}{cccccc} \vec{X}_{m-k}(1) & \vec{X}_{m-k}(2) & & \ldots & \vec{X}_{m-k}(k) & \vec{X}_{m-k}(k+1) \\ \downarrow & \downarrow & & & \downarrow & \downarrow \\ \left[\begin{array}{cccc} X(1) & X(2) & \ldots & X(k) & X(k+1) \\ X(2) & X(3) & \ldots & X(k+1) & X(k+2) \\ \vdots & \vdots & & \vdots & \vdots \\ X(m \Leftrightarrow k) & X(m \Leftrightarrow k+1) & \ldots & X(m \Leftrightarrow 1) & X(m) \end{array}\right], \end{array}$$

muni du produit hermitien empirique usuel $(\langle\cdot,\cdot\rangle_e)$ de C^{m-k}. Les vecteurs $\vec{X}_{m-k}(j)$, $j=1,\ldots,k+1$, sont de bons représentants des variables $X(j)$, $j=1,\ldots,k+1$, au sens où l'espérance de $\langle\vec{X}_{m-k}(i),\vec{X}_{m-k}(j)\rangle_e$ est donnée par la covariance entre $X(i)$ et $X(j)$. L'estimateur du coefficient $\beta(k)$ est alors obtenu en remplaçant dans sa définition théorique les variables $X(j)$, $j=1,\ldots,k+1$, par leur représentants et le produit hermitien $\langle\cdot,\cdot\rangle$ par $\langle\cdot,\cdot\rangle_e$.

Dans la situation périodique, afin d'obtenir des représentants des variables $X(s),\ldots,X(t)$, on considère les séquences de longueur $t \Leftrightarrow s+1$ translatées en temps de T plutôt que 1. Rappelons que les séquences $\{X(s),\ldots,X(t)\}$ et $\{X(s+T),\ldots,X(t+T)\}$ possèdent la même structure. Plus précisément, on introduit la matrice

$$X^m_{s,t} = \left[\begin{array}{cccc} X(s) & X(s+1) & \ldots & X(t-1) & X(t) \\ X(s+T) & X(s+1+T) & \ldots & X(t-1+T) & X(t+T) \\ \vdots & \vdots & & \vdots & \vdots \\ X(s+m_tT) & X(s+1+m_tT) & \ldots & X(t-1+m_tT) & X(t+m_tT) \end{array}\right],$$

avec $m_t = \left[\frac{m-t}{T}\right]$, $t \Leftrightarrow s = 1, \ldots, p_t \leq m_t$, et $s \in [1, \ldots, T]$. Notons que $X(t + m_t T)$ est la dernière variable dont on dispose pour un temps qui a le même reste que t dans la division entière par T. Les colonnes de $X_{s,t}^m$ sont notées $\vec{X}_{m_t+1}(u)$ $u = s, \ldots, t$. Le sous-espace vectoriel de C^{m_t+1} engendré par ces vecteurs est muni du produit hermitien empirique usuel

$$\left\langle \vec{X}_{m_t+1}(u), \vec{X}_{m_t+1}(v) \right\rangle_e = \frac{1}{m_t+1} \sum_{j=0}^{m_t} X(u+jT)\overline{X(v+jT)}.$$

La structure obtenue reproduit, en moyenne, celle de la séquence $\{X(s), \ldots, X(t)\}$ au sens attendu,

$$E\left\{ \left\langle \vec{X}_{m_t+1}(u), \vec{X}_{m_t+1}(v) \right\rangle_e \right\} = \langle X(u), X(v) \rangle = R(u,v).$$

Ainsi, le coefficient d'autocorrélation partielle empirique $\hat{\beta}_{acpe}(t,s)$, est la "corrélation partielle" entre $\vec{X}_{m_t+1}(s)$ et $\vec{X}_{m_t+1}(t)$ dans l'ensemble

$$\left\{ \vec{X}_{m_t+1}(s), \ldots, \vec{X}_{m_t+1}(t) \right\},$$

selon cette analogie. Nous utiliserons également $\rho_e(\cdot,\cdot/\cdot)$ pour désigner la "corrélation partielle empirique". De façon plus précise, on note $\vec{\varepsilon}^{\,f}_{m_t+1}(t; s+1)$ et $\vec{\varepsilon}^{\,b}_{m_t+1}(s; t \Leftrightarrow 1)$ les erreurs de prédiction obtenues selon le critère des moindres carrés,

$$\vec{\varepsilon}^{\,f}_{m_t+1}(t; s+1) = \sum_{j=0}^{t-s-1} \overline{a}_t^f(t \Leftrightarrow s \Leftrightarrow 1, j) \vec{X}_{m_t+1}(t \Leftrightarrow j), \quad \overline{a}_t^f(t \Leftrightarrow s \Leftrightarrow 1, 0) = 1,$$

$$\vec{\varepsilon}^{\,b}_{m_t+1}(s; t \Leftrightarrow 1) = \sum_{j=0}^{t-s-1} \overline{a}_{t-1}^b(t \Leftrightarrow s \Leftrightarrow 1, j) \vec{X}_{m_t+1}(s+j), \quad \overline{a}_{t-1}^b(t \Leftrightarrow s \Leftrightarrow 1, 0) = 1.$$

Soulignons le fait que la barre est employée ici pour désigner les coefficients définies par le critère des moindres carrés. Par la suite, nous utiliserons cette notation lorsque cela n'est pas ambigu. Ces erreurs de prédiction sont caractérisées par, $j = 1, \ldots, t \Leftrightarrow s \Leftrightarrow 1$,

$$\left\langle \vec{\varepsilon}^{\,f}_{m_t+1}(t; s+1), \vec{X}_{m_t+1}(t \Leftrightarrow j) \right\rangle_e = \left\langle \vec{\varepsilon}^{\,b}_{m_t+1}(s; t \Leftrightarrow 1), \vec{X}_{m_t+1}(s+j) \right\rangle_e = 0.$$

On définit, avec la convention $0^{-1} = 0$, les erreurs normalisées associées,

$$\vec{\eta}^{\,f}_{m_t+1}(t; s+1) = \frac{\vec{\varepsilon}^{\,f}_{m_t+1}(t; s+1)}{\overline{\sigma}^f_{m_t+1}(t; s+1)}, \quad \overline{\sigma}^f_{m_t+1}(t; s+1) = \left\| \vec{\varepsilon}^{\,f}_{m_t+1}(t; s+1) \right\|_e,$$

$$\vec{\eta}^{\,b}_{m_t+1}(s; t \Leftrightarrow 1) = \frac{\vec{\varepsilon}^{\,b}_{m_t+1}(s; t \Leftrightarrow 1)}{\overline{\sigma}^b_{m_t+1}(s; t \Leftrightarrow 1)}, \quad \overline{\sigma}^b_{m_t+1}(s; t \Leftrightarrow 1) = \left\| \vec{\varepsilon}^{\,b}_{m_t+1}(s; t \Leftrightarrow 1) \right\|_e.$$

Les estimateurs des coefficients d'autocorrélation partielle $\hat{\beta}_{acpe}(\cdot,\cdot)$ sont don-

nés, pour $s = 1, \ldots, T$, $0 < t \Leftrightarrow s \leq p_t$, par :

$$\begin{aligned} \hat{\beta}_{acpe}(t,s) &= \left\langle \vec{\eta}^{f}_{m_t+1}(t; s+1), \vec{\eta}^{b}_{m_t+1}(s; t \Leftrightarrow 1) \right\rangle_e , \\ \hat{\beta}_{acpe}(s,s) &= \left\| \vec{X}_{m_s+1}(s) \right\|^2_e . \end{aligned}$$

Dans le cas scalaire stationnaire ($T = 1$), les coefficients d'autocorrélation partielle, ainsi estimés, correspondent à ceux de la version non symétrisée de Dégerine présentée au début de cette sous-section. Dans la situation périodiquement corrélée, la mise en œuvre de cette méthode nécessite a priori autant de décompositions de Cholesky que de coefficients à calculer (soit $\sum_{i=1}^{T} p_i$). Cependant lorsque $m = NT$, les quantités $\hat{\beta}_{acpe}(t,s)$ apparaissent dans un algorithme de calcul rapide proposé par Pham [PHA92] pour déterminer les matrices d'autocorrélation partielle empiriques triangulaires. Dans ce cas, on dispose d'un algorithme rapide pour les déterminer.

4.1.3 Rapports d'énergies résiduelles

Pour les processus stationnaires [DIC78], la méthode RER représente une approximation d'$ACPE$. Les coefficients $\beta(k)$, $0 \leq k \leq p$, sont estimés en utilisant le même sous-espace engendré par $\vec{X}_{m-p}(s)$, $s = 1, \ldots, p$, au lieu des sous-espaces de C^{m-k} successifs. Il nous semble intéressant de considérer cette méthode dans la situation périodiquement corrélée puisqu'elle constitue, pour les processus stationnaires, une bonne approximation et nécessite moins de calculs.

Dans la situation périodique, on introduit les entiers t_i, $i = 1, \ldots, T$, tels que $t_i = kT + i$, $k \in \mathbb{N}$, et $0 < t_i \Leftrightarrow p_i \leq T$. Les estimateurs $\hat{\beta}_{rer}(t_i, s)$, $i = 1, \ldots, T$, $0 < t_i \Leftrightarrow s \leq p_i \leq m_{t_i}$, sont donnés par

$$\rho_e(\vec{X}_{m_{t_i}+1}(s), \vec{X}_{m_{t_i}+1}(t_i) / \vec{X}_{m_{t_i}+1}(s), \ldots, \vec{X}_{m_{t_i}+1}(t_i))$$

Avec les notations introduites précédemment, on obtient pour $i = 1, \ldots, T$,

$$\begin{aligned} &\hat{\beta}_{rer}(t_i, s) = \left\langle \vec{\eta}^{f}_{m_{t_i}+1}(t_i; s+1), \vec{\eta}^{b}_{m_{t_i}+1}(s; t_i \Leftrightarrow 1) \right\rangle_e , \quad 0 < t_i \Leftrightarrow s \leq p_i, \\ &\hat{\beta}_{rer}(t_i, t_i) = \left\| \vec{X}_{m_{t_i}+1}(t_i) \right\|^2_e . \end{aligned}$$

Les coefficients $\hat{\beta}_{rer}(t_i, s)$ sont distincts de $\hat{\beta}_{acpe}(t_i, s)$ excepté pour les $(T + p_i \Leftrightarrow t_i + 1)$ derniers. Notons que cette méthode nécessite au plus T décompositions de Cholesky. En effet lorsque $m_{t_i} = m_{t_{i'}}$, les coefficients $\hat{\beta}_{rer}(t_i, \cdot)$ et $\hat{\beta}_{rer}(t_{i'}, \cdot)$ peuvent être déterminés à partir de la décomposition de Cholesky d'une même matrice. Cette procédure est plus facile à mettre en œuvre que $ACPE$, mais n'apporte pas nécessairement un gain en rapidité de calcul selon l'ordre du modèle ARP.

4.1.4 Méthodes de type Burg

Dans le cas scalaire stationnaire réel, la procédure de Burg [BUR75] consiste à estimer les coefficients $\beta(k)$ en minimisant, de manière récursive sur l'ordre, la somme des carrés des erreurs de prédiction progressive et rétrograde d'ordre k. Remarquons qu'une extension directe de cette méthode conduirait au problème de minimisation suivant,

$$\min_{a_t^f(t-s,\cdot),\ a_t^b(t-s,\cdot)} \left\{ \left\| \sum_{j=0}^{t-s} a_t^f(t-s,j)\vec{X}_{m_t+1}(t-j) \right\|_e^2 + \left\| \sum_{j=0}^{t-s} a_t^b(t-s,j)\vec{X}_{m_t+1}(s+j) \right\|_e^2 \right\},$$

sous les contraintes

$$a_t^f(t-s,t-s) = -\beta(t,s)\frac{\sigma^f(t;s+1)}{\sigma^b(s;t-1)}, \quad a_t^b(t-s,t-s) = -\beta(s,t)\frac{\sigma^b(s;t-1)}{\sigma^f(t;s+1)},$$

pour $j = 1, \ldots, t \Leftrightarrow s \Leftrightarrow 1$:

$$a_t^f(t-s,j) = \hat{a}_t^f(t-s-1,j) + a_t^f(t-s,t-s)\hat{a}_{t-1}^b(t-s-1,t-s-j),$$

$$a_t^b(t-s,j) = \hat{a}_{t-1}^b(t-s-1,j) + a_t^b(t-s,t-s)\hat{a}_t^f(t-s-1,t-s-j).$$

Ce procédé, conduisant parfois à des coefficients d'autocorrélation partielle en module plus grands que 1, est à rejeter.

On trouve deux méthodes de type Burg différentes, celle de Sakai [SAK82] et celle de Boshnakov [BOS94]. Ces procédures sont plutôt basées sur la minimisation des erreurs de la forme,

$$\left[1 \Leftrightarrow |\beta(t,s)|^2\right]^{\frac{1}{2}} \eta^f(t;s), \quad \left[1 \Leftrightarrow |\beta(t,s)|^2\right]^{\frac{1}{2}} \eta^b(s;t).$$

Rappelons que ces variables satisfont, pour $s < t$, les relations suivantes (cf. Proprosition 1.2.1),

$$\left[1 \Leftrightarrow |\beta(t,s)|^2\right]^{\frac{1}{2}} \eta^f(t;s) = \eta^f(t;s+1) \Leftrightarrow \beta(t,s)\eta^b(s;t \Leftrightarrow 1),$$

$$\left[1 \Leftrightarrow |\beta(t,s)|^2\right]^{\frac{1}{2}} \eta^b(s;t) = \eta^b(s;t \Leftrightarrow 1) \Leftrightarrow \beta(s,t)\eta^f(t;s+1).$$

La valeur qui réalise le minimum dans,

$$\min_{\beta} \left\{ \left\| \eta^f(t;s+1) \Leftrightarrow \beta\eta^b(s;t \Leftrightarrow 1) \right\|^2 + \left\| \eta^b(s;t \Leftrightarrow 1) \Leftrightarrow \overline{\beta}\eta^f(t;s+1) \right\|^2 \right\},$$

est donnée par

$$\frac{2\left\langle \eta^f(t;s+1), \eta^b(s;t \Leftrightarrow 1) \right\rangle}{\left\| \eta^f(t;s+1) \right\|^2 + \left\| \eta^b(s;t \Leftrightarrow 1) \right\|^2},$$

qui est égal à

$$\left\langle \eta^f(t;s+1), \eta^b(s;t \Leftrightarrow 1) \right\rangle = \beta(t,s),$$

puisque les innovations $\eta^f(t;s+1)$ et $\eta^b(s;t \Leftrightarrow 1)$ sont normées. Sur le plan empirique ce critère est appliqué de façon récursive sur l'ordre, les quantités $\hat{\eta}^f(t;s+1)$ et $\hat{\eta}^b(s;t \Leftrightarrow 1)$ étant définies à partir des estimations obtenues aux étapes précédentes. La différence entre la méthode de Sakai et celle de Bosnhakov résulte d'un choix différent pour l'estimation des variances résiduelles dans la définition des erreurs $\hat{\eta}^f(t;s+1)$ et $\hat{\eta}^b(s;t \Leftrightarrow 1)$.

Dans [Bos94], le rapport des variances résiduelles $\sigma^f(t;s)/\sigma^b(s;t)$ est supposé connu égal à $\hat{q}(t;s) = \hat{\sigma}^f(t;s+1)/\hat{\sigma}^b(s;t \Leftrightarrow 1)$. Ainsi la quantité $\hat{\beta}_{b1}(t,s)$, $0 < t \Leftrightarrow s \leq p_t$, est la valeur qui réalise le minimum dans,

$$\min_{\beta} \left\{ \left\| \hat{\varepsilon}^f_{m_t+1} \Leftrightarrow \beta \hat{q}(t;s) \hat{\varepsilon}^b_{m_t+1} \right\|_e^2 + \left\| \hat{q}(t;s) \hat{\varepsilon}^b_{m_t+1} \Leftrightarrow \overline{\beta} \hat{\varepsilon}^f_{m_t+1} \right\|_e^2 \right\},$$

où les erreurs de prédiction $\hat{\varepsilon}^f_{m_t+1} = \hat{\varepsilon}^f_{m_t+1}(t;s+1)$ et $\hat{\varepsilon}^b_{m_t+1} = \hat{\varepsilon}^b_{m_t+1}(s;t \Leftrightarrow 1)$ sont déterminées à partir des coefficients $\hat{\beta}(u,v)$, $(u,v) \in [s,\ldots,t]^2 \setminus (t,s)$, estimés aux étapes précédentes. On a alors,

$$\hat{\beta}_{b1}(t,s) = \frac{2 \left\langle \hat{\varepsilon}^f_{m_t+1}(t;s+1), \hat{\varepsilon}^b_{m_t+1}(s;t \Leftrightarrow 1) \right\rangle_e}{\frac{1}{\hat{q}(t;s)} \left\| \hat{\varepsilon}^f_{m_t+1}(t;s+1) \right\|_e^2 + \hat{q}(t;s) \left\| \hat{\varepsilon}^b_{m_t+1}(s;t \Leftrightarrow 1) \right\|_e^2}.$$

La différence avec la méthode de Sakai provient de l'absence de contraintes sur les variances résiduelles. Le problème de minimisation est le suivant,

$$\min_{\beta} \left\{ \left\| \tilde{\eta}^f_{m_t+1} \Leftrightarrow \beta \tilde{\eta}^b_{m_t+1} \right\|_e^2 + \left\| \tilde{\eta}^b_{m_t+1} \Leftrightarrow \overline{\beta} \tilde{\eta}^f_{m_t+1} \right\|_e^2 \right\},$$

où $\tilde{\eta}^f_{m_t+1} = \tilde{\eta}^f_{m_t+1}(t;s+1)$ et $\tilde{\eta}^b_{m_t+1} = \tilde{\eta}^b_{m_t+1}(s;t \Leftrightarrow 1)$ sont les erreurs $\hat{\varepsilon}^f_{m_t+1}(t;s+1)$ et $\hat{\varepsilon}^b_{m_t+1}(s;t \Leftrightarrow 1)$ définies précédemment, normalisées avec la norme associée au produit hermitien empirique. Dans ce cas, les coefficients $\beta(t,s)$, $0 < t \Leftrightarrow s \leq p_t$, sont estimés comme des corrélations,

$$\hat{\beta}_{b2}(t,s) = \left\langle \tilde{\eta}^f_{m_t+1}(t;s+1), \tilde{\eta}^b_{m_t+1}(s;t \Leftrightarrow 1) \right\rangle_e.$$

Pour ces deux procédures, les premiers coefficients $\hat{\beta}(t,t)$, $i = 1,\ldots,T$, sont donnés par $\left\| \vec{X}_{m_t+1}(t) \right\|_e^2$. Lorsque $T = 1$, la méthode de Boshnakov coïncide avec celle de Burg, alors que celle de Sakai conduit à des estimateurs différents.

4.1.5 Comparaison des méthodes

Dans cette sous-section, les deux méthodes de type Burg et $ACPE$ sont regroupées dans une même méthodologie permettant ainsi de les comparer facilement. On pose pour $t = 1,\ldots,T$, $\hat{\beta}(t,t) = \left\| \vec{X}_{m_t+1}(t) \right\|_e^2$ et pour $s \in$

$[1, \ldots, T]$, $0 < t \Leftrightarrow s \leq p_t$, l'estimateur $\hat{\beta}(t,s)$ est obtenu par une procédure en deux étapes de la façon suivante.

(i) Choisir les coefficients $\tilde{a}_t^f(t \Leftrightarrow s \Leftrightarrow 1, \cdot)$ et $\tilde{a}_{t-1}^b(t \Leftrightarrow s \Leftrightarrow 1, \cdot)$ des filtres donnant les erreurs de prédictions d'ordre $t \Leftrightarrow s \Leftrightarrow 1$,

$$\tilde{\varepsilon}_{m_t+1}^f(t; s+1) = \sum_{j=0}^{t-s-1} \tilde{a}_t^f(t \Leftrightarrow s \Leftrightarrow 1, j) \vec{X}_{m_t+1}(t \Leftrightarrow j),$$

$$\tilde{\varepsilon}_{m_t+1}^b(s; t \Leftrightarrow 1) = \sum_{j=0}^{t-s-1} \tilde{a}_{t-1}^b(t \Leftrightarrow s \Leftrightarrow 1, j) \vec{X}_{m_t+1}(s+j),$$

auxquelles on associe les éléments empiriques

$$\tilde{\sigma}_{m_t+1}^{f2}(t; s+1) = \left\| \tilde{\varepsilon}_{m_t+1}^f(t; s+1) \right\|_e^2, \quad \tilde{\sigma}_{m_t+1}^{b2}(s; t \Leftrightarrow 1) = \left\| \tilde{\varepsilon}_{m_t+1}^b(s; t \Leftrightarrow 1) \right\|_e^2,$$

$$\tilde{\delta}_{m_t+1}(t,s) = \left\langle \tilde{\varepsilon}_{m_t+1}^f(t; s+1), \tilde{\varepsilon}_{m_t+1}^b(s; t \Leftrightarrow 1) \right\rangle_e.$$

(ii) Choisir les estimateurs $\tilde{\sigma}^{f2}(t; s+1)$ et $\tilde{\sigma}^{b2}(s; t \Leftrightarrow 1)$ des variances résiduelles.

Alors $\hat{\beta}(t,s)$ est défini par

$$\hat{\beta}(t,s) = \frac{2\tilde{\delta}_{m_t+1}(t,s)}{\frac{\tilde{\sigma}^b(s;t-1)}{\tilde{\sigma}^f(t;s+1)} \tilde{\sigma}_{m_t+1}^{f2}(t; s+1) + \frac{\tilde{\sigma}^f(t;s+1)}{\tilde{\sigma}^b(s;t-1)} \tilde{\sigma}_{m_t+1}^{b2}(s; t \Leftrightarrow 1)}.$$

Remarquons que l'estimateur $\hat{\beta}(t,s)$ est obtenu en appliquant le critère utilisé pour les méthodes de Burg. Lorsqu'on choisit à l'étape (ii) les variances empiriques définies en (i), il est alors donné comme une corrélation.

Les extensions de type Burg prennent à l'étape (i), les filtres associés aux estimateurs $\hat{\beta}(u,v)$, $v \leq u$ avec $(u,v) \in [s, \ldots, t]^2 \setminus (t,s)$, obtenus aux étapes précédentes. Puis, pour l'étape (ii), Boshnakov [BOS94] utilise les variances résiduelles associées à ces estimateurs alors que Sakai [SAK82] retient celles empiriques définies en (i). La méthode des autocorrélations partielles empiriques consiste à utiliser les filtres donnés par le critère des moindres carrés en (i) et les variances résiduelles empiriques pour (ii).

Pour les méthodes de Burg, on peut déjà noter que les contraintes dans la construction récursive des filtres des erreurs de prédiction constituent a priori un handicap par rapport à $ACPE$ ou même RER. Concernant la comparaison entre ces deux approches, on constate que les estimations des coefficients d'autocorrélation partielle proposées par Boshnakov sont de la forme

$$\frac{2 \left\langle \hat{\varepsilon}^f, \hat{\varepsilon}^b \right\rangle_e}{\frac{1}{\hat{q}} \left\| \hat{\varepsilon}^f \right\|_e^2 + \hat{q} \left\| \hat{\varepsilon}^b \right\|_e^2},$$

et ceux de Sakai,

$$\frac{\left\langle \hat{\varepsilon}^f, \hat{\varepsilon}^b \right\rangle_e}{\|\hat{\varepsilon}^f\|_e \times \|\hat{\varepsilon}^b\|_e}.$$

Par conséquent, si à une étape les filtres des erreurs de prédiction sont les mêmes pour les deux méthodes, alors l'estimation de Sakai sera plus grande en module que celle de Boshnakov. Ceci laisse prévoir que la méthode de Sakai pourrait donner de meilleurs résultats pour les situations proches du cas localement déterminable. Notons cependant que les différences obtenues lors des simulations (cf. section 4.3) ne semblent pas être significatives.

Par ailleurs, lorsque les données proviennent d'une série localement déterminable, les coefficient $\overline{a}_t^f(t \Leftrightarrow s \Leftrightarrow 1, \cdot)$ et $\overline{a}_{t-1}^b(t \Leftrightarrow s \Leftrightarrow 1, \cdot)$ ne sont pas nécessairement uniques, mais les erreurs de prédiction $\vec{\varepsilon}_{m_t+1}^f(t; s+1)$ et $\overleftarrow{\varepsilon}_{m_t+1}^b(s; t \Leftrightarrow 1)$ sont bien définies par le critère des moindres carrés. Dans ce cas les singularités d'ordre fini sont estimées presque sûrement par $ACPE$ ou RER puisque les filtres sont ceux donnés par le critère des moindres carrés. Cependant les filtres traduisant les singularités ne sont généralement pas cohérents avec $\hat{\beta}(\cdot, \cdot)$ dont l'estimation devrait être remise en cause. Les trois autres méthodes, qui n'ont pas été introduites dans la situation localement déterminable, fonctionnent encore mais sans garantir que la structure estimée soit celle d'une série localement déterminable.

Dans la situation stationnaire, toutes les méthodes excepté RER, sont récursives sur l'ordre. Cette propriété est intéressante lorsque l'ordre du modèle est inconnu. En effet une méthode est récursive sur l'ordre du modèle si une augmentation de ce dernier ne modifie pas la valeur des premiers coefficients de $\hat{\beta}(\cdot)$. Dans la situation périodique, $ACPE$ est la seule méthode qui garde en toute généralité cette propriété. Pour cette dernière, les estimateurs des autocorrélations partielles sont parfaitement définis au delà de l'ordre du modèle. Ils sont simplement tronqués à zéro lorsqu'on se donne un ordre. Pour les méthodes de type Burg, cette propriété n'est pas toujours satisfaite selon l'ordre du modèle considéré et la façon dont il est augmenté. Ceci provient du fait que l'estimation de $\beta(t, s)$ dépend, à travers les filtres des erreurs de prédiction, de toute celles de $\beta(u, v)$, $v \leq u$ avec $(u, v) \in [s, \ldots, t]^2 \setminus (t, s)$. En particulier une modification de la valeur de $\hat{\beta}(t, s)$, quantité qui intervient dans le calcul de $\hat{a}_t^f(t \Leftrightarrow s, \cdot)$ mais aussi de $\hat{a}_t^b(t \Leftrightarrow s, \cdot)$, change celle de $\hat{\beta}(t, s \Leftrightarrow 1)$ mais aussi de $\hat{\beta}(t+1, s)$. Ainsi lorsque p_i est augmenté de 1, l'estimation de $\hat{\beta}(i+1, i \Leftrightarrow p_i \Leftrightarrow 1) = \hat{\beta}_{i+1}(p_i + 2)$ est à recalculer si dans le modèle précédent $p_{i+1} > p_i + 1$. Le même type d'argument montre que la méthode de Yule-Walker ne conserve pas cette propriété pour les processus périodiquement corrélés.

4.2 Les approches vectorielles

Étant donnée l'association $\{X(\cdot), Y(\cdot)\}$, on dispose de la séquence $Y(1)$, $\ldots, Y(N)$, lorsque le processus est observé sur une longueur $m = NT$. Ainsi les méthodes d'estimation autorégressives vectorielles donnent lieu à des méthodes scalaires et inversement.

La différence fondamentale entre les approches vectorielle et scalaire est que la structure périodique résultant d'une méthode vectorielle est celle d'un processus $ARP(p, max)$. Parmi tous les processus stationnaires $VAR(p)$, les méthodes vectorielles estiment donc qu'une sous-classe de ces modèles. Ce problème est contourné avec les méthodes scalaires puisqu'elles permettent d'estimer un modèle ARP de n'importe quel ordre. Notons cependant qu'elles nécessitent plus d'information a priori. En effet, l'ordre du modèle ARP est constitué de T paramètres $p_1, \ldots, p_T$, alors que celui du modèle VAR est donné par un seul paramètre, p. Néanmoins lorsque l'ordre du modèle est inconnu, un critère portant sur la structure périodique (cf. [SAK82]) semble préférable au cas vectoriel, celui-ci pouvant introduire une surparamétrisation. Afin de comparer les deux approches vectorielle et scalaire, nous considérons par la suite les modèles $ARP(p, max)$.

Dégerine [DÉG94] regroupe dans un cadre général la plupart des méthodes d'estimation vectorielles utilisant les matrices d'autocorrélation partielle. On s'intéresse plus particulièrement aux procédures faisant intervenir les matrices d'autocorrélation partielle triangulaires définies au troisième chapitre. Le schéma, proposé par Dégerine [DÉG94], se réduit alors à l'analogue de celui qu'on propose pour l'approche scalaire. Il est constitué du choix des filtres des erreurs de prédiction et de celui des estimateurs des matrices de covariances résiduelles. L'estimateur de $\Pi(k)$ est donné par une équation de Lyapounov provenant d'un problème de minimisation analogue au cas périodique. Dans l'analogie entre méthodologies vectorielle et scalaire, la généralisation de la technique de Burg, proposée par Nuttall [NUT76] (cf. aussi [STR77]) correspond à celle de Boshnakov [BOS94] et celle de Morf et coll. [MOR78b] à celle de Sakai [SAK82]. Cependant ces méthodes ne sont pas équivalentes. Par contre, l'extension d'$ACPE$ au cas périodique équivaut à celle définie par les matrices d'autocorrélation partielle empiriques triangulaires ([DÉG92], [PHA92], [DÉG94]) résultant d'ailleurs des mêmes types de choix dans les deux étapes. Ceci est également vrai pour la méthode de Pagano [PAG78] et l'extension de RER, qui n'entrent pas dans le cadre méthodologique scalaire. Remarquons que le choix de la normalisation pour les méthodes vectorielles a une influence pour les estimations de type $ACPE$, RER ou Morf et coll. A l'origine, pour ces deux dernières méthodes, les estimations étaient obtenues avec une normalisation triangulaire inférieure dans

les deux sens progressif et rétrograde.

Nous présentons ci-après les différentes méthodes vectorielles évoquées et décrivons, dans la mesure du possible, les méthodes d'estimation scalaires qui en résultent.

4.2.1 Méthode de Yule-Walker

Cette méthode consiste à ajuster le modèle $VAR(p)$ dont la structure est donnée pour $u = 0, \ldots, p$, par les covariances empiriques biaisées,

$$\hat{R}_N(u) = \frac{1}{N} \sum_{k=1}^{N-u} Y(k+u)Y(k)^*.$$

Le (i,j)-ème élément de la matrice $\hat{R}_N(u)$, qui est donné par

$$\hat{R}_N(u)_{i,j} = \frac{1}{N} \sum_{k=1}^{N-u} X(i + (k+u \Leftrightarrow 1)T)\overline{X(j + (k \Leftrightarrow 1)T)},$$

est égal à $\hat{R}_m(i + uT, j)$. Compte tenu de la correspondance entre modèles VAR et ARP (cf. section 3.4), la structure périodique résultante est celle de l'$ARP(p, max)$ dont la structure est donnée, pour $t = 1, \ldots, T$, $0 \leq t \Leftrightarrow s \leq p_t$, par $\hat{R}_m(t, t \Leftrightarrow s)$. Il y a donc équivalence entre la méthode de Yule-Walker vectorielle et celle périodique.

4.2.2 Matrices d'autocorrélation partielle empiriques triangulaires

De façon générale, la méthode $ACPE$ est définie dans [Dég92] (cf. aussi [Dég94]) pour un procédé de normalisation quelconque. Un algorithme de calcul, par une approche scalaire, des matrices d'$ACPE$ triangulaires est donné dans [Pha92].

Pour $k = 0, \ldots, p < N/2$, on considère les matrices de taille $(N \Leftrightarrow k) \times (k+1)T$,

$$Y_k^N = \begin{bmatrix} Y(1)^T & Y(2)^T & \ldots & Y(k)^T & Y(k+1)^T \\ Y(2)^T & Y(3)^T & \ldots & Y(k+1)^T & Y(k+2)^T \\ \vdots & \vdots & & \vdots & \vdots \\ Y(N-k)^T & Y(N-k+1)^T & \ldots & Y(N-1)^T & Y(N)^T \end{bmatrix}.$$

Les blocs colonnes de taille $T \times (N \Leftrightarrow k)$ de Y_k^N sont notés $\vec{Y}_{N-k}(u)^T$, $u = 1, \ldots, k+1$. Le sous-espace vectoriel engendré par les matrices $\vec{Y}_{N-k}(u)$,

$u = 1, \ldots, k+1$, est muni du produit intérieur,

$$\left[\vec{Y}_{N-k}(u), \vec{Y}_{N-k}(v)\right] = \frac{1}{N \Leftrightarrow k} \vec{Y}_{N-k}(u) \vec{Y}_{N-k}(v)^*.$$

Les matrices $\vec{Y}_{N-k}(1), \ldots, \vec{Y}_{N-k}(k+1)$ sont des représentants de $Y(1), \ldots, Y(k+1)$, au sens où

$$E\left\{\left[\vec{Y}_{N-k}(u), \vec{Y}_{N-k}(v)\right]\right\} = E\left\{Y(u)Y(v)^*\right\} = R(u \Leftrightarrow v).$$

Les erreurs de prédiction $\vec{e}^{\,f}_{N-k}(k+1; k \Leftrightarrow 1)$ et $\vec{e}^{\,b}_{N-k}(1; k \Leftrightarrow 1)$ sont obtenues selon le critère des moindres carrés ($\vec{e}^{\,f}_{N-1}(2;0) = \vec{Y}_{N-1}(2)$, $\vec{e}^{\,b}_{N-1}(1;0) = \vec{Y}_{N-1}(1)$),

$$\vec{e}^{\,f}_{N-k}(k+1; k \Leftrightarrow 1) = \sum_{j=0}^{k-1} \overline{A}^f(k \Leftrightarrow 1, j) \vec{Y}_{N-k}(k+1 \Leftrightarrow j), \quad \overline{A}^f(k \Leftrightarrow 1, 0) = I_T,$$

$$\vec{e}^{\,b}_{N-k}(1; k \Leftrightarrow 1) = \sum_{j=0}^{k-1} \overline{A}^b(k \Leftrightarrow 1, j) \vec{Y}_{N-k}(1+j), \quad \overline{A}^b(k \Leftrightarrow 1, 0) = I_T.$$

Elles sont caractérisées par, $u = 2, \ldots, k$,

$$\left[\vec{e}^{\,f}_{N-k}(k+1; k \Leftrightarrow 1), \vec{Y}_{N-k}(u)\right] = \left[\vec{e}^{\,b}_{N-k}(1; k \Leftrightarrow 1), \vec{Y}_{N-k}(u)\right] = 0.$$

Les erreurs normalisées associées sont définies par,

$$\vec{\zeta}^{\,f}_{N-k}(k+1; k \Leftrightarrow 1) = \overline{,}^{\,-f}_{N-k}(k \Leftrightarrow 1)^{-}_{i} \vec{e}^{\,f}_{N-k}(k+1; k \Leftrightarrow 1),$$

$$\vec{\zeta}^{\,b}_{N-k}(1; k \Leftrightarrow 1) = \overline{,}^{\,-b}_{N-k}(k \Leftrightarrow 1)^{-}_{s} \vec{e}^{\,b}_{N-k}(1; k \Leftrightarrow 1),$$

où

$$\overline{,}^{\,f2}_{N-k}(k \Leftrightarrow 1) = \left[\vec{e}^{\,f}_{N-k}(k+1; k \Leftrightarrow 1), \vec{e}^{\,f}_{N-k}(k+1; k \Leftrightarrow 1)\right],$$

$$\overline{,}^{\,b2}_{N-k}(k \Leftrightarrow 1) = \left[\vec{e}^{\,b}_{N-k}(1; k \Leftrightarrow 1), \vec{e}^{\,b}_{N-k}(1; k \Leftrightarrow 1)\right].$$

Alors les matrices d'$ACPE$ triangulaires sont données par $\hat{\Pi}_{ACPE}(0) = \left[\vec{Y}_N(1), \vec{Y}_N(1)\right]$, et pour $k = 1, \ldots, p$,

$$\hat{\Pi}_{ACPE}(k) = \left[\vec{\zeta}^{\,f}_{N-k}(k+1; k \Leftrightarrow 1), \vec{\zeta}^{\,b}_{N-k}(1; k \Leftrightarrow 1)\right].$$

La méthode $ACPE$ vectorielle est équivalente à celle introduite dans le cadre périodique lorsqu'il s'agit d'un modèle $ARP(p, max)$. De façon plus précise, nous avons la proposition suivante.

Proposition 4.2.1. *Soit $\hat{\beta}(\cdot,\cdot)$ la fonction d'autocorrélation partielle du processus périodique associé au modèle $VAR(p)$ dont la structure est donnée par $\hat{\Pi}_{ACPE}(k)$, $k = 0, \ldots, p$. Alors $\hat{\beta}(\cdot,\cdot)$ coïncide avec $\hat{\beta}_{acpe}(\cdot,\cdot)$ sur $E(pT, pT+1, \ldots, pT+T \Leftrightarrow 1)$.*

Avant d'établir la preuve de cette proposition, nous donnons ci-dessous une remarque afin d'en faciliter la compréhension. Ce résultat est trivial compte tenu de l'approche géométrique utilisée dans le premier chapitre.

Remarque 4.2.1. *Soient U_i, $i = 1, \ldots, n$, n vecteurs de C^d et $\beta(i,j)$ les coefficients associés à $R(i,j) = \langle U_i, U_j \rangle_e$, $i,j = 1, \ldots n$, via l'Algorithme LDG. Alors les quantités $\beta(i,j)$ sont aussi les "corrélations partielles empiriques",*

$$\beta(i,j) = \rho_e(U_i, U_j/U_j, \ldots, U_i).$$

Démonstration.- Dans un premier temps, l'estimateur de $\Pi(\cdot)$ est exprimé en fonction des erreurs de prédiction associées au processus $X(\cdot)$. La matrice Y_k^N est égale à $X_{1,(k+1)T}^{NT}$, et les représentants $\vec{Y}_{N-k}(u)$, $u = 1, \ldots, k+1$, sont donnés par,

$$\vec{Y}_{N-k}(u) = \begin{bmatrix} \vec{X}_{N-k}((u \Leftrightarrow 1)T + 1)^T \\ \vdots \\ \vec{X}_{N-k}((u \Leftrightarrow 1)T + T)^T \end{bmatrix}.$$

Le (i,j)-ème élément du produit intérieur entre $\vec{Y}_{N-k}(u)$ et $\vec{Y}_{N-k}(v)$ satisfait,

$$\left[\vec{Y}_{N-k}(u), \vec{Y}_{N-k}(v)\right]_{i,j} = \left\langle \vec{X}_{N-k}((u \Leftrightarrow 1)T + i), \vec{X}_{N-k}((v \Leftrightarrow 1)T + j) \right\rangle_e.$$

Suite à ce qui précède et d'après la caractérisation des erreurs de prédiction vectorielles et périodiques, on a

$$\vec{e}_{N-k}^{f}(k+1; k \Leftrightarrow 1) = \begin{bmatrix} \vec{\varepsilon}_{N-k}^{f}(kT+1; T+1)^T \\ \cdots \\ \vec{\varepsilon}_{N-k}^{f}(kT+T; T+1)^T \end{bmatrix},$$

$$\vec{e}_{N-k}^{b}(1; k \Leftrightarrow 1) = \begin{bmatrix} \vec{\varepsilon}_{N-k}^{b}(1; kT)^T \\ \cdots \\ \vec{\varepsilon}_{N-k}^{b}(T; kT)^T \end{bmatrix}.$$

Par ailleurs le procédé de normalisation conduit, dans le cas régulier, à

$$\left[\vec{\zeta}_{N-k}^{f}(k+1; k \Leftrightarrow 1), \vec{\zeta}_{N-k}^{f}(k+1; k \Leftrightarrow 1)\right] = I_T, \tag{4.1}$$

$$\left[\vec{\zeta}_{N-k}^{b}(1; k \Leftrightarrow 1), \vec{\zeta}_{N-k}^{b}(1; k \Leftrightarrow 1)\right] = I_T. \tag{4.2}$$

Dans le cas localement déterminable, les matrices identités sont remplacées par $\hat{J}^f(k \Leftrightarrow 1)$ et $\hat{J}^b(k \Leftrightarrow 1)$. Notons $A_{i\cdot}$, la i-ème ligne d'une matrice A. D'après l'égalité (4.1), les vecteurs $\vec{\zeta}_{N-k}^{f}(k+1; k \Leftrightarrow 1)_{i\cdot}^T$, $i = 1, \ldots, T$, sont obtenus par l'orthogonalisation pour $\langle \cdot, \cdot \rangle_e$, selon le procédé de Gram-schmidt, des variables successives $\vec{e}_{N-k}^{f}(k+1; k \Leftrightarrow 1)_{i\cdot}^T$. On a donc $\vec{\zeta}_{N-k}^{f}(k+1; k \Leftrightarrow 1)_{i\cdot}^T = \vec{\eta}_{N-k}^{f}(kT+i; T+1)$, $i = 1, \ldots, T$. Un raisonnement analogue dans le sens

rétrograde conduit à $\vec{\zeta}^{b}_{N-k}(1; k \Leftrightarrow 1)^{T}_{i.} = \vec{\eta}^{b}_{N-k}(i; kT)$. Comme dans le cadre théorique, la matrice $\hat{\Pi}_{ACPE}(k)$ s'exprime donc simplement en terme de variables associées à la structure périodique,

$$\hat{\Pi}_{ACPE}(0)_{i,j} = \left\langle \vec{X}_N(i), \vec{X}_N(j) \right\rangle_e, \quad i,j = 1, \ldots, T,$$

$$\hat{\Pi}_{ACPE}(k)_{i,j} = \left\langle \vec{\eta}^{f}_{N-k}(kT+i; T+1), \vec{\eta}^{b}_{N-k}(j; kT) \right\rangle_e, \quad k > 0. \tag{4.3}$$

Pour $k = 0$, la correspondance entre $\hat{\Pi}_{ACPE}(0)$ et $\hat{\beta}(i,j)$, $i,j = 1, \ldots, T$, est obtenue (cf. section 3.3) avec l'Algorithme *LDG*. la Remarque 4.2.1 avec $U_i = X_{N-1}(i)$, $i = 1, \ldots, T$, montre que $\hat{\beta}(i,j) = \hat{\beta}_{acpe}(i,j)$, $i,j = 1, \ldots, T$. Pour $k = 1, \ldots, p$, et pour $i = 1, \ldots, T$, on pose

$$U^{(k)}_i = \vec{\eta}^{b}_{N-k}(i; kT), \quad U^{(k)}_{T+i} = \vec{\eta}^{f}_{N-k}(kT+i; T+1).$$

D'après les égalités (4.1),(4.2) et (4.3) les coefficients $\hat{R}^{(k)}(i,j)$ que nous avons introduits à la section 3.3 sont égaux à $\langle U^{(k)}_i, U^{(k)}_j \rangle_e$. Rappelons que $\hat{\beta}(kT+i, j)$ est donné par $\hat{\beta}^{(k)}(T+i, j)$, où $\hat{\beta}^{(k)}(\cdot,\cdot)$ est la fonction associée à $\hat{R}^{(k)}(\cdot,\cdot)$ par l'Algorithme *LDG*. Ainsi la Remarque 4.2.1 appliquée à $U^{(k)}_i$, $i = 1, \ldots, 2T$, montre que $\hat{\beta}^{(k)}(T+i, j)$, est la "corrélation partielle" entre $\vec{\eta}^{f}_{N-k}(kT+i; T+1)$ et $\vec{\eta}^{b}_{N-k}(j; kT)$ dans l'ensemble

$$\{\vec{\eta}^{b}_{N-k}(n; kT), n = j, \ldots, T; \vec{\eta}^{f}_{N-k}(kT+n; T+1), n = 1 \ldots, i\}.$$

De plus, on dispose de l'analogue empirique de la Proposition 3.3.1 en remplaçant dans sa démonstration $\langle \cdot, \cdot \rangle$ par $\langle \cdot, \cdot \rangle_e$ et les éléments théoriques par leurs représentants empiriques. La "corrélation partielle" précédente est donc aussi celle donnée par

$$\rho_e(\vec{X}_{N-k}(kT+i), \vec{X}_{N-k}(j) / \vec{X}_{N-k}(j), \ldots, \vec{X}_{N-k}(kT+i)),$$

soit $\hat{\beta}_{acpe}(kT+i, j)$. $\diamondsuit$

4.2.3 Rapports d'énergies résiduelles

Cette méthode a été proposée par Dickinson [Dic79]. Comme dans le cas scalaire, les estimateurs de $\Pi(k)$, $k = 0, \ldots, p$, sont déterminés à partir de la même matrice Y^N_p au lieu des matrices successives Y^N_k. Dans [Dic79], $\hat{\Pi}_{RER}(\cdot)$ est obtenu en utilisant une normalisation triangulaire inférieure dans les deux sens progressif et rétrograde. Nous proposons une version différente en accord avec la définition introduite au troisième chapitre et directement liée à la méthode du cas périodique. Notons que la structure de $Y(\cdot)$, ainsi estimée, est différente de celle résultant de [Dic79].

Avec les notations introduites précédemment, l'estimateur de $\Pi(k)$ est

défini par,

$$\hat{\Pi}_{RER}(0) = \left[\vec{Y}_{N-p}(p+1), \vec{Y}_{N-p}(p+1)\right],$$
$$\hat{\Pi}_{RER}(k) = \left[\vec{\zeta}^{f}_{N-p}(p+1; k \Leftrightarrow 1), \vec{\zeta}^{b}_{N-p}(p+1 \Leftrightarrow k; k \Leftrightarrow 1)\right], \quad 0 < k \leq p.$$

En utilisant les éléments de la preuve de la Proposition 4.2.1, on montre que la fonction $\hat{\beta}(\cdot,\cdot)$ associée à $\hat{\Pi}_{RER}(\cdot)$ est donnée par, $i,j = 1,\ldots,T$, $k = 1,\ldots,p$,

$$\hat{\beta}(i,i) = \left\|\vec{X}_{N-p}(pT+i)\right\|_e^2,$$
$$\hat{\beta}(i,j) = \left\langle \vec{\eta}^{f}_{N-p}(pT+i; pT+j+1), \vec{\eta}^{b}_{N-p}(pT+j; pT+i-1)\right\rangle_e, \; j < i,$$
$$\hat{\beta}(kT+i,j) = \left\langle \vec{\eta}^{f}_{N-p}(pT+i; (p-k)T+j+1), \vec{\eta}^{b}_{N-p}((p-k)T+j; pT+i-1)\right\rangle_e.$$

Par ailleurs pour un modèle $ARP(p,max)$, les entiers t_i, définis à la sous-section 4.1.3, sont donnés par $pT+i$ car $p_i = pT+i \Leftrightarrow 1$. La fonction $\hat{\beta}(\cdot,\cdot)$ coïncide donc avec $\hat{\beta}_{rer}(\cdot,\cdot)$ lorsque le modèle considéré est $ARP(p,max)$.

4.2.4 Extensions de la méthode de Burg

On trouve deux généralisations différentes de la méthode de Burg au cas vectoriel non localement déterminable. La première a été proposée indépendemment par Nuttall [Nut76] et Strand [Str77]. Cette méthode est comparable dans sa conception à celle de Boshnakov [Bos94]. La seconde, comparable à celle de Sakai [Sak82], a été introduite par Morf et coll. [Mor78b]. Comme pour la situation scalaire, la matrice $\Pi(k)$ est estimée de façon récursive sur l'ordre, à partir des erreurs de prédiction définies par les estimations obtenues aux étapes précédentes. Remarquons que pour ces deux méthodes, la récurrence commence avec $\hat{\Pi}(0) = \hat{\Pi}_{ACPE}(0)$. D'autre part, les coefficients des filtres ainsi que les variances empiriques scalaires d'$ACPE$ satisfont les contraintes de Burg sur la plage des erreurs faisant intervenir $\hat{\Pi}(0)$. Ils coïncident donc avec ceux des méthodes de Burg périodiques qui sont associés aux estimations obtenues aux étapes précédentes. Dans ce cas, les estimateurs $\hat{\beta}(i,j)$ $i,j = 1,\ldots,T$, sont identiques pour les méthodes de Sakai et $ACPE$. De plus l'estimateur de Boshnakov se ramène au coefficient de Sakai lorsque les variances empiriques sont égales à celles estimées. Les premiers coefficients $\beta(i,j)$, $i,j = 1,\ldots,T$, sont donc estimés de la même façon par les méthodes de Boshnakov, Sakai ou $ACPE$ et sont ceux associés à $\hat{\Pi}_{ACPE}(0)$ par l'Algorithme LDG. Ainsi pour l'estimation de $\Pi(0)$, les deux méthodes de Burg vectorielles sont équivalentes à celles périodiques qui leurs sont respectivement comparables. Néanmoins cette équivalence n'est plus vraie pour les ordres suivants. Pour chacune de ces deux méthodes, après les

avoir présentées, nous donnons les raisons pour lesquelles nous n'avons plus cette correspondance.

Méthode de Nuttall

Dans [NUT76] ou [STR77], la méthode porte sur l'estimation de l'autocovariance partielle. Afin de souligner les similitudes avec la technique de Boshnakov, nous présentons la méthode d'estimation pour $\Pi(\cdot)$ qui en résulte. Remarquons que cette procédure ne dépend pas du choix de la normalisation. On introduit les erreurs de prédictions d'ordre $k \Leftrightarrow 1$,

$$\hat{e}^f_{N-k}(k+1; k \Leftrightarrow 1) = \sum_{j=0}^{k-1} \hat{A}^f(k \Leftrightarrow 1, j) \vec{Y}_{N-k}(k+1 \Leftrightarrow j),$$

$$\hat{e}^b_{N-k}(1; k \Leftrightarrow 1) = \sum_{j=0}^{k-1} \hat{A}^b(k \Leftrightarrow 1, j) \vec{Y}_{N-k}(1+j),$$

où $\hat{A}^f(k \Leftrightarrow 1, \cdot)$ et $\hat{A}^b(k \Leftrightarrow 1, \cdot)$ sont les coefficients des innovations partielles d'ordre $k \Leftrightarrow 1$ associées aux estimations $\hat{\Pi}(0), \ldots, \hat{\Pi}(k \Leftrightarrow 1)$, obtenues aux étapes précédentes. Alors $\hat{\Pi}_{B1}(k)$, $k = 1, \ldots, p$, est la matrice Π qui réalise le minimum dans,

$$\min_{\Pi} tr \left\{ \left[\hat{\zeta}^f_{N-k} - \Pi \hat{\zeta}^b_{N-k}, \hat{\zeta}^f_{N-k} - \Pi \hat{\zeta}^b_{N-k} \right] + \left[\hat{\zeta}^b_{N-k} - \Pi^* \hat{\zeta}^f_{N-k}, \hat{\zeta}^b_{N-k} - \Pi^* \hat{\zeta}^f_{N-k} \right] \right\},$$

où tr désigne la trace, $\hat{\zeta}^f_{N-k}$, $\hat{\zeta}^b_{N-k}$ sont les représentants des innovations partielles normalisées,

$$\hat{\zeta}^f_{N-k}(k+1; k \Leftrightarrow 1) = \hat{}^f(k \Leftrightarrow 1)^-_i \hat{e}^f_{N-k}(k+1; k \Leftrightarrow 1),$$
$$\hat{\zeta}^b_{N-k}(1; k \Leftrightarrow 1) = \hat{}^b(k \Leftrightarrow 1)^-_s \hat{e}^b_{N-k}(1; k \Leftrightarrow 1).$$

L'estimateur $\hat{\Pi}_{B1}(k)$ est alors la solution de l'équation de Lyapounov,

$$A\Pi(k) + \Pi(k)B = 2C, \tag{4.4}$$

avec

$$A = \left[\hat{\zeta}^f_{N-k}(k+1; k \Leftrightarrow 1), \hat{\zeta}^f_{N-k}(k+1; k \Leftrightarrow 1) \right],$$

$$B = \left[\hat{\zeta}^b_{N-k}(1; k \Leftrightarrow 1), \hat{\zeta}^b_{N-k}(1; k \Leftrightarrow 1) \right],$$

$$C = \left[\hat{\zeta}^f_{N-k}(k+1; k \Leftrightarrow 1), \hat{\zeta}^b_{N-k}(1; k \Leftrightarrow 1) \right].$$

Pour cette situation, il semble qu'il soit impossible de décrire simplement la méthode périodique correspondante. Néanmoins on montre que cette dernière est différente de celle de Bosnhakov. Cette différence provient de

la façon dont est déterminée l'estimation de $\Pi(k)$. En effet, supposons que les estimations de $\Pi(i)$, $i = 0, \ldots, k \Leftrightarrow 1$, soient identiques pour les deux méthodes et plaçons nous à la k-ième étape. D'après les relations de (3.11) et leur homologue dans le sens rétrograde, on a

$$\hat{e}^f_{N-k}(k+1; k \Leftrightarrow 1) = [\hat{L}^f_{k-1}]^{-1} \begin{bmatrix} \hat{\varepsilon}^f_{N-k}(kT+1; T+1)^T \\ \vdots \\ \hat{\varepsilon}^f_{N-k}(kT+T; T+1)^T \end{bmatrix}, \tag{4.5}$$

$$\hat{e}^b_{N-k}(1; k \Leftrightarrow 1) = [\hat{L}^b_{k-1}]^{-1} \begin{bmatrix} \hat{\varepsilon}^b_{N-k}(1; kT)^T \\ \vdots \\ \hat{\varepsilon}^b_{N-k}(T; kT)^T \end{bmatrix}, \tag{4.6}$$

où $\hat{\varepsilon}^f_{N-k}$, $\hat{\varepsilon}^b_{N-k}$ correspondent aux erreurs de prédiction introduites à la sous-section 4.1.4. Les inverses des racines des matrices de covariance résiduelles estimées (uniques dans le cas non localement déterminable) sont donnés par

$$\hat{,}^f(k \Leftrightarrow 1)^-_i = [\hat{D}^f_{k-1}]^+ \hat{L}^f_{k-1}, \quad \hat{,}^b(k \Leftrightarrow 1)^-_s = [\hat{D}^b_{k-1}]^+ \hat{L}^b_{k-1},$$

et par suite

$$\hat{\zeta}^f_{N-k}(k+1; k \Leftrightarrow 1) = \begin{bmatrix} \hat{\eta}^f_{N-k}(kT+1; T+1)^T \\ \vdots \\ \hat{\eta}^f_{N-k}(kT+T; T+1)^T \end{bmatrix},$$

$$\hat{\zeta}^b_{N-k}(1; k \Leftrightarrow 1) = \begin{bmatrix} \hat{\eta}^b_{N-k}(1; kT)^T \\ \vdots \\ \hat{\eta}^b_{N-k}(T; kT)^T \end{bmatrix},$$

où $\hat{\eta}^f_{N-k}(\cdot,\cdot)$, $\hat{\eta}^b_{N-k}(\cdot,\cdot)$ sont les erreurs de prédiction "normalisées" par les écarts-types résiduels $\hat{\sigma}^f(\cdot,\cdot)$, $\hat{\sigma}^b(\cdot,\cdot)$. Afin de faciliter la lecture, nous modifions légèrement les notations des erreurs de prédiction en omettant la référence à l'ordre k. On dispose donc de deux matrices $\hat{\zeta}^f$ et $\hat{\zeta}^b$ dont les i-èmes lignes transposées sont notées $\hat{\eta}^f_{i,T+1}$ et $\hat{\eta}^b_{j,0}$. Dans [Nut76], la matrice $\Pi(k)$ est estimée comme le coefficient de liaison linéaire, introduit dans [Dég94], entre $\hat{\zeta}^f$ et $\hat{\zeta}^b$. C'est-à-dire comme la solution de l'équation de Lyapounov (4.4). Dans [Bos94], l'approche scalaire conduit à définir de façon non globale l'estimateur $\hat{\Pi}(k)$. En effet, cette méthode se ramène à la procédure suivante.

(i) Pour $i = 1, \ldots, T$, $j = T, \ldots, 1$, on calcul les coefficients $\hat{\beta}_{i,j}$ comme,

$$\hat{\beta}_{i,j} = \frac{2 \left\langle \hat{\eta}^f_{i,j+1}, \hat{\eta}^b_{j,i-1} \right\rangle_e}{\|\hat{\eta}^f_{i,j+1}\|^2_e + \|\hat{\eta}^b_{j,i-1}\|^2_e}.$$

Lorsque $(i,j) \neq (T,1)$ les nouvelles variables $\hat{\eta}^f_{i,j}$ et $\hat{\eta}^b_{j,i}$, nécessaires pour les étapes $(i, j \Leftrightarrow 1)$ et $(i+1, j)$, sont déterminées par

$$\hat{\eta}^f_{i,j} = \left[1 \Leftrightarrow |\hat{\beta}_{i,j}|^2\right]^{-\frac{1}{2}} \left\{\hat{\eta}^f_{i,j+1} \Leftrightarrow \hat{\beta}_{i,j}\hat{\eta}^b_{j,i-1}\right\},$$

$$\hat{\eta}^b_{j,i} = \left[1 \Leftrightarrow |\hat{\beta}_{i,j}|^2\right]^{-\frac{1}{2}} \left\{\hat{\eta}^b_{j,i-1} \Leftrightarrow \overline{\hat{\beta}_{i,j}}\hat{\eta}^f_{i,j+1}\right\}.$$

(ii) L'estimateur $\hat{\Pi}(k)$ est la matrice d'autocorrélation partielle triangulaire associée à $\hat{\beta}(kT+i,j) = \hat{\beta}_{i,j}$, $i,j = 1,\ldots,T$, selon la procédure de la section 3.3.

Ainsi la méthode de Boshnakov définit un estimateur $\hat{\Pi}(k)$ à partir des mêmes quantités que dans [NUT76], mais de façon non équivalente. En effet pour $T=2$ par exemple, l'équation de Lyapounov (4.4) est équivalente au système suivant,

$$\begin{pmatrix} A_{1,1}+B_{1,1} & A_{1,2} & B_{2,1} & 0 \\ B_{1,2} & A_{1,1}+B_{2,2} & 0 & A_{1,2} \\ A_{2,1} & 0 & A_{2,2}+B_{1,1} & B_{2,1} \\ 0 & A_{2,1} & B_{1,2} & A_{2,2}+B_{2,2} \end{pmatrix} \begin{pmatrix} \Pi_{1,1} \\ \Pi_{1,2} \\ \Pi_{2,1} \\ \Pi_{2,2} \end{pmatrix} = 2 \begin{pmatrix} C_{1,1} \\ C_{1,2} \\ C_{2,1} \\ C_{2,2} \end{pmatrix}.$$

Or d'après le point (ii) de la procédure précédente, le coefficient $\hat{\Pi}(k)_{1,2}$ est égal à $\hat{\beta}(2k+1,2)$ et d'après (i), il est donc déterminé par $2C_{1,2}/(A_{1,1}+B_{2,2})$. Contrairement à la méthode de Nuttall, il n'est pas solution du système précédent car $(0, 1/(A_{1,1}+B_{2,2}), 0, 0)$ n'est pas la deuxième ligne de l'inverse de la matrice de ce système. Ceci montre clairement que la méthode de Nuttall et Strand n'est pas équivalente à celle de Boshnakov.

Méthode de Morf et coll.

Nous présentons ci-après la méthode de Morf et coll. [MOR78b]. Comme pour la méthode *RER*, nous proposons une version différente de l'originale, mais en accord avec la définition de $\Pi(\cdot)$. Le critère utilisé est alors le suivant,

$$tr\left\{\left[\tilde{\zeta}^f_{N-k} \Leftrightarrow \Pi\tilde{\zeta}^b_{N-k}, \tilde{\zeta}^f_{N-k} \Leftrightarrow \Pi\tilde{\zeta}^b_{N-k}\right] + \left[\tilde{\zeta}^b_{N-k} \Leftrightarrow \Pi^*\tilde{\zeta}^f_{N-k}, \tilde{\zeta}^b_{N-k} \Leftrightarrow \Pi^*\tilde{\zeta}^f_{N-k}\right]\right\},$$

où les représentants des innovations partielles normalisées, $\tilde{\zeta}^f_{N-k}$ et $\tilde{\zeta}^b_{N-k}$ sont donnés par,

$$\tilde{\zeta}^f_{N-k}(k+1; k \Leftrightarrow 1) = \tilde{,}^f_{N-k}(k \Leftrightarrow 1)^-_i \hat{e}^f_{N-k}(k+1; k \Leftrightarrow 1),$$

$$\tilde{\zeta}^b_{N-k}(1; k \Leftrightarrow 1) = \tilde{,}^b_{N-k}(k \Leftrightarrow 1)^-_s \hat{e}^b_{N-k}(1; k \Leftrightarrow 1).$$

avec

$$\tilde{,}^{f2}_{N-k}(k \Leftrightarrow 1) = \left[\hat{e}^f_{N-k}(k+1; k \Leftrightarrow 1), \hat{e}^f_{N-k}(k+1; k \Leftrightarrow 1)\right],$$

$$\tilde{,}^{b2}_{N-k}(k \Leftrightarrow 1) = \left[\hat{e}^b_{N-k}(1; k \Leftrightarrow 1), \hat{e}^b_{N-k}(1; k \Leftrightarrow 1)\right].$$

Les matrices de covariance empiriques de ces représentants étant égales à l'identité, l'équation de Lyapounov est éliminée et les matrices $\hat{\Pi}_{B2}(k)$, $k = 1, \ldots, p$, sont déterminées comme des coefficients de corrélation,

$$\hat{\Pi}_{B2}(k) = \left[\tilde{\zeta}^f_{N-k}(k+1; k \Leftrightarrow 1), \tilde{\zeta}^b_{N-k}(1; k \Leftrightarrow 1)\right].$$

Comme précédemment, la différence entre cette méthode et celle de Sakai réside dans l'estimation de la matrice $\Pi(k)$ ou de façon équivalente de celles des coefficients $\beta(kT+i, j)$, $i, j = 1, \ldots, T$. On propose de traiter ce point de manière différente car pour la méthode de Morf et coll., il est possible de décrire directement la procédure d'estimation des quantités $\beta(kT+i, j)$. En effet à l'étape k, on dispose des erreurs de prédiction périodiques $\hat{\varepsilon}^f_{i,T+1} = \hat{\varepsilon}^f_{N-k}(kT+i; T+1)$ et $\hat{\varepsilon}^b_{i,0} = \hat{\varepsilon}^b_{N-k}(i; kT)$. Alors la structure périodique résultant de la méthode de Morf et coll. est obtenue de la façon suivante.

Proposition 4.2.2. *Pour $k = 1, \ldots, p$, soient $\hat{\beta}(kT+i, j)$, $i, j = 1, \ldots, T$, les coefficients d'autocorrélation partielle associés à $\hat{\Pi}_{B2}(k)$. Alors ces coefficients sont donnés par,*

$$\rho_e(\hat{\varepsilon}^f_{i,T+1}, \hat{\varepsilon}^b_{j,0} / \hat{\varepsilon}^b_{n,0}, n = j, \ldots, T; \hat{\varepsilon}^f_{n,T+1}, n = 1, \ldots, i).$$

Démonstration.- Comme pour $ACPE$, la Remarque 4.2.1 avec pour $i = 1, \ldots, T$,

$$U^{(k)}_i = \tilde{\zeta}^b_{N-k}(1; k \Leftrightarrow 1)^T_{i\cdot}, \quad U^{(k)}_{T+i} = \tilde{\zeta}^f_{N-k}(k+1; k \Leftrightarrow 1)^T_{i\cdot},$$

conduit à

$$\hat{\beta}(kT+i, j) = \rho_e(U^{(k)}_{T+i}, U^{(k)}_j / U^{(k)}_j, \ldots, U^{(k)}_{T+i}).$$

Par ailleurs les vecteurs $U^{(k)}_{T+i}$, $i = 1, \ldots, T$ sont obtenus par l'orthogonalisation pour $\langle \cdot, \cdot \rangle_e$ selon le procédé de Gram-schmidt, des variables successives $\hat{e}^f_{N-k}(k+1; k \Leftrightarrow 1)^T_{i\cdot}$. La relation (4.5) où $[\hat{L}^f_{k-1}]^{-1}$ est une matrice triangulaire inférieure unitaire, implique qu'ils peuvent être aussi déterminés selon le même procédé mais à partir des variables $\hat{\varepsilon}^f_{1,T+1}, \ldots, \hat{\varepsilon}^f_{T,T+1}$. On obtient un résultat similaire pour $U^{(k)}_i$, $i = 1, \ldots, T$, grâce à la relation (4.6). L'analogue empirique de la Proposition 3.3.1 (appliquée avec $n = 1$) finit de montrer cette proposition. $\diamondsuit$

La procédure de Morf et coll. n'est donc pas équivalente à celle de Sakai. En effet, rappelons que pour cette dernière les coefficients $\hat{\beta}(kT+i, j)$, $i, j = 1, \ldots, T$, sont déterminés selon la procédure suivante.

Pour $i = 1, \ldots, T$, $\hat{\sigma}^{f2}_{i,T+1} = \hat{\sigma}^{f2}(kT+i; T+1)$ et $\hat{\sigma}^{b2}_{i,0} = \hat{\sigma}^{b2}(i; kT)$.

Pour $i = 1, \ldots, T$, pour $j = T, \ldots, 1$,

$$\hat{\beta}_{b2}(kT+i,j) = \frac{\left\langle \hat{\varepsilon}^f_{i,j+1}, \hat{\varepsilon}^b_{j,i-1} \right\rangle_e}{\|\hat{\varepsilon}^f_{i,j+1}\|_e \times \|\hat{\varepsilon}^b_{j,i-1}\|_e},$$

lorsque $(i,j) \neq (T,1)$,

$$\hat{\varepsilon}^f_{i,j} = \hat{\varepsilon}^f_{i,j+1} \Leftrightarrow \hat{\beta}_{b2}(kT+i,j)\frac{\hat{\sigma}^f_{i,j+1}}{\hat{\sigma}^b_{j,i-1}}\hat{\varepsilon}^b_{j,i-1},$$

$$\hat{\varepsilon}^b_{j,i} = \hat{\varepsilon}^b_{j,i-1} \Leftrightarrow \overline{\hat{\beta}_{b2}(kT+i,j)}\frac{\hat{\sigma}^b_{j,i-1}}{\hat{\sigma}^f_{i,j+1}}\hat{\varepsilon}^f_{i,j+1},$$

$$\hat{\sigma}^{f2}_{i,j} = \left[1 \Leftrightarrow |\hat{\beta}_{b2}(kT+i,j)|^2\right]\hat{\sigma}^{f2}_{i,j+1}, \quad \hat{\sigma}^{b2}_{j,i} = \left[1 \Leftrightarrow |\hat{\beta}_{b2}(kT+i,j)|^2\right]\hat{\sigma}^{b2}_{j,i-1}.$$

Ainsi, sauf pour $(i,j) = (1,T)$, les quantités $\hat{\beta}(kT+i,j)$ et $\hat{\beta}_{b2}(kT+i,j)$ sont différentes.

4.3 Résultats de simulation

Dans cette section, les différentes méthodes d'estimation sont analysées par simulation. Les facteurs déterminants sont la proximité du modèle à la singularité, la longueur de la séquence et le type de paramètres étudiés. Pour nos comparaisons, nous avons retenu quatre modèles M1, M2, M3 et M4. Le modèle M1 est un $ARP(4,4)$ régulier, les autocorrélations partielles n'excédant pas 0.85 en module. On considère deux modèles proches de la singularité, un $ARP(6,6,6)$ pour M2 et un $ARP(3,max)$ pour M4. Le modèle M3 est un $ARP(7,6)$ régulier comme M1, mais avec des valeurs pour les coefficients des filtres plus élevées. La méthode $ACPE$ est d'abord comparée avec celle de Yule-Walker pour des séries courtes issues du modèle M1. On retrouve le défaut de la méthode de Yule-Walker dû au biais. Comme dans le cas stationnaire, on constate que $ACPE$ et RER conduisent à des résultats similaires. Ces deux dernières méthodes sont alors comparées avec les deux extensions de Burg scalaires. Les résultats obtenus pour M2 montrent le mauvais comportement des méthodes de Burg pour des séries courtes issues d'un modèle proche de la singularité. Le modèle M3 illustre le fait que ces méthodes sont également sensibles à l'ordre de grandeur des coefficients des filtres. Nous considérons ensuite le modèle M4 qui permet de comparer les approches scalaires et celles vectorielles. Ceci concerne essentiellement les procédures de type Burg puisque les deux approches sont équivalentes pour les autres méthodes lorsqu'il s'agit d'un modèle $ARP(p,max)$. Une étude sur

le comportement asymptotique pour le modèle M4 indique que les méthodes $ACPE$ et RER semblent préférables même sur des séries de longueur moyenne.

Dans la première sous-section, nous décrivons la façon dont sont obtenus les résultats de simulation. La comparaison des différentes méthodes fait l'objet de la deuxième sous-section.

4.3.1 Présentation des résultats

Sans perte de généralité, on se place dans le cas réel avec $m = NT$. Les essais sont réalisés à partir d'un fichier de 1 000 000 de nombres pseudo-aléatoires de loi normale centrée réduite, pour lequel on a observé une moyenne de 0.00097 et un écart-type de 0.99996. Les méthodes sont comparées à travers l'estimation des paramètres $\beta_t(n)$,

$$\rho_t(n) = R(t, t \Leftrightarrow n)/\sqrt{R(t,t)R(t \Leftrightarrow n, t \Leftrightarrow n)},$$

$R(t,t)$, $a_t(n)$, et σ_t^2, $t = 1, \ldots, T$, $n = 1, \ldots, p_t$. Remarquons que nous avons préféré étudier séparément la variance du processus $R(t,t)$ et les autocorrélations $\rho_t(n)$, plutôt que les coefficients d'autocovariances. Ceci afin de faciliter l'analyse puisque, excepté pour RER, les différentes méthodes estiment de la même façon la variance du processus. Pour chaque modèle, les méthodes concernées sont appliquées sur des séquences de longueur m et avec un nombre de répétitions n_r donnés. Les résultats détaillés de ces simulations sont donnés dans des tables en Annexe. La valeur des paramètres du modèle considéré se trouve dans la première colonne intitulée "MODÈLE". Pour chaque estimateurs, nous donnons dans la colonne correspondante, sa valeur moyenne (sur l'ensemble des n_r répétitions) suivi de la racine carré de l'écart quadratique moyen (EQM). Les résultats sont présentés période par période (i.e. pour t fixé). Pour les paramètres $\beta_t(\cdot)$, $\rho_t(\cdot)$ et $a_t(\cdot)$, nous mentionnons par une ligne intitulée "Synthèse", un résumé des résultats obtenus sur la période considérée. Par exemple pour un vecteur de paramètres $\theta(\cdot)$ de dimension d, ($d = p_t$ pour la t-ième période) le biais moyen et la valeur moyenne de la racine de l'EQM sont résumés par,

$$\left\{\frac{1}{d}\sum_{i=1}^{d}\left[\text{biais}[\hat{\theta}_i(k)]\right]^2\right\}^{\frac{1}{2}}, \qquad \left\{\frac{1}{d}\sum_{i=1}^{d}\text{EQM}[\hat{\theta}_i(k)]\right\}^{\frac{1}{2}}.$$

l'EQM permet d'apprécier de façon globale la différence entre deux méthodes résultant d'écarts dûs au biais et à la variance. Nous n'avons pas indiqué la variance afin de ne pas surcharger les tableaux, mais des commentaires préciserons, de façon qualitative, la part qui lui revient dans les différences constatées entre les méthodes. Par ailleurs pour chaque type de paramètres,

nous donnons sur une ligne intitulée "Synt. globale", une mesure globale du biais et de la racine de l'EQM sur l'ensemble des périodes obtenue selon le même principe : $d = \sum_{t=1}^{T} p_t$ pour $\beta.(\cdot)$, $\rho.(\cdot)$, $a.(\cdot)$, et $d = T$ pour $R(\cdot,\cdot)$, $\sigma_.^2$. Enfin des tableaux de synthèse et des graphiques sont utilisés dans les commentaires ci-dessous.

4.3.2 Étude comparative

Yule-Walker

La méthode de Yule-Walker [PAG78] est comparée avec $ACPE$, pour des séries courtes ($m = 50, 100$) issues du modèle M1. On retrouve le défaut classique dû au biais particulièrement sensible pour des séries courtes. La Table I de l'Annexe montre que les deux méthodes diffèrent peu pour l'estimation de $\rho(\cdot,\cdot)$. Pour les autres paramètres, les résultats des synthèses globales obtenues pour $m = 50$ sont rappelés dans le Tableau 4.1 ci-dessous. On observe que le biais et l'EQM sont plus élevés pour Yule-Walker et en particulier pour les paramètres du modèle (i.e. les filtres et les variances résiduelles). Ici seul le biais est à l'origine de la différence entre les deux méthodes (les variances sont proches). Ceci semble cohérent puisque cette méthode introduit essentiellement un biais sur les autocovariances. Pour cet exemple, les autres méthodes conduisent à des résultats comparables à ceux d'$ACPE$. On a pu remarquer que l'écart entre les deux méthodes s'accentue pour un modèle proche de la singularité même pour des séries longues.

Tableau 4.1 : *Comparaison entre Yule-Walker et ACPE, M1, $m = 50$.*

Paramètre	$\beta(\cdot,\cdot)$		$a.(\cdot)$		$\sigma_.^2$	
	biais	$(\text{EQM})^{\frac{1}{2}}$	biais	$(\text{EQM})^{\frac{1}{2}}$	biais	$(\text{EQM})^{\frac{1}{2}}$
Yule-Walker	0.136	0.211	0.251	0.388	0.172	0.261
$ACPE$	0.044	0.161	0.040	0.261	0.034	0.066

Le modèle M1 fait parti des modèles pour lesquels la méthode de Yule-Walker fournit toujours une solution (cf. sous-section 4.1.1). Soit M1' le modèle $ARP(4,0)$ dont la fonction d'autocorrélation partielle coïncide avec celle de M1 sur $E(4,0)$. Pour une série de longueur 50 (resp. 100), on a observé 26.7 (resp. 28.5) pourcentage d'échec sur 2000 répétitions.

ACPE et RER

Les résultats obtenus par $ACPE$ sont très satisfaisant sur l'ensemble des paramètres. Comme dans le cas stationnaire, la méthode RER en

constitue une bonne approximation (cf. Annexe : Tables II à VIII). Elle représente donc une solution de remplacement puisqu'elle est plus facile à mettre en œuvre. Cependant d'un point de vue calculatoire, elle n'apporte pas nécessairement quelque chose car elle peut nécessiter T décompositions de Cholesky selon l'ordre du modèle d'une part et on dispose de l'algorithme de calcul rapide de [PHA92] pour $ACPE$ d'autre part.

Extensions de Burg scalaires

On observe les mêmes phénomènes que pour la situation stationnaire concernant les modèles proches de la singularité et pour des séries courtes. Les méthodes de Burg ([SAK82], [BOS94]) sont d'abord comparées avec $ACPE$ et RER pour des séries de longueur $m = 105$ issues du modèle M2 proche de la singularité. On constate que les méthodes de Burg, qui se comportent de façon analogue, sont bien moins bonnes excepté pour l'estimation de $\rho(\cdot,\cdot)$ pour laquelle les quatre méthodes sont équivalentes (cf. Annexe : Table II). Pour une vue d'ensemble, nous rappelons ci-dessous les synthèses globales pour cet exemple.

Tableau 4.2 : *Comparaison entre Burg, RER et ACPE, M2, m=105.*

Paramètre	$\beta(\cdot,\cdot)$		$a.(\cdot)$		$10^2 \times \sigma_.^2$	
	biais	$(\text{EQM})^{\frac{1}{2}}$	biais	$(\text{EQM})^{\frac{1}{2}}$	biais	$(\text{EQM})^{\frac{1}{2}}$
Sakai	0.124	0.208	0.523	0.738	4.098	6.097
Boshnakov	0.125	0.208	0.531	0.747	4.153	6.164
RER	0.066	0.156	0.017	0.251	0.499	0.799
$ACPE$	0.065	0.154	0.017	0.256	0.475	0.799

De façon détaillée, la Figure 4.1 ci-après représente le biais et la racine de l'EQM pour les estimateurs des autocorrélations partielles. Les paramètres $\beta_1(1), \ldots, \beta_1(p_1), \beta_2(1), \ldots, \beta_T(p_T)$ sont numérotés de 1 à $\sum_{i=1}^{T} p_i$. Ces numéros sont reportés en abscisse et pour chacun d'entre eux, la valeur du biais ou de la racine de l'EQM est donnée en ordonnée. Les différentes périodes sont séparées par des tracés verticaux en pointillés. Le défaut de conception des méthodes de Burg, dû aux contraintes dans la construction récursive des filtres, est bien illustré par ces deux graphes : sur chaque début de période, ces procédures se comportent comme les deux autres (les estimateurs $\hat{\beta}_i(k)$, $k = 0, \ldots, i \Leftrightarrow 1$, sont identiques à ceux d'$ACPE$), puis la différence s'accentue en fin de chaque période.

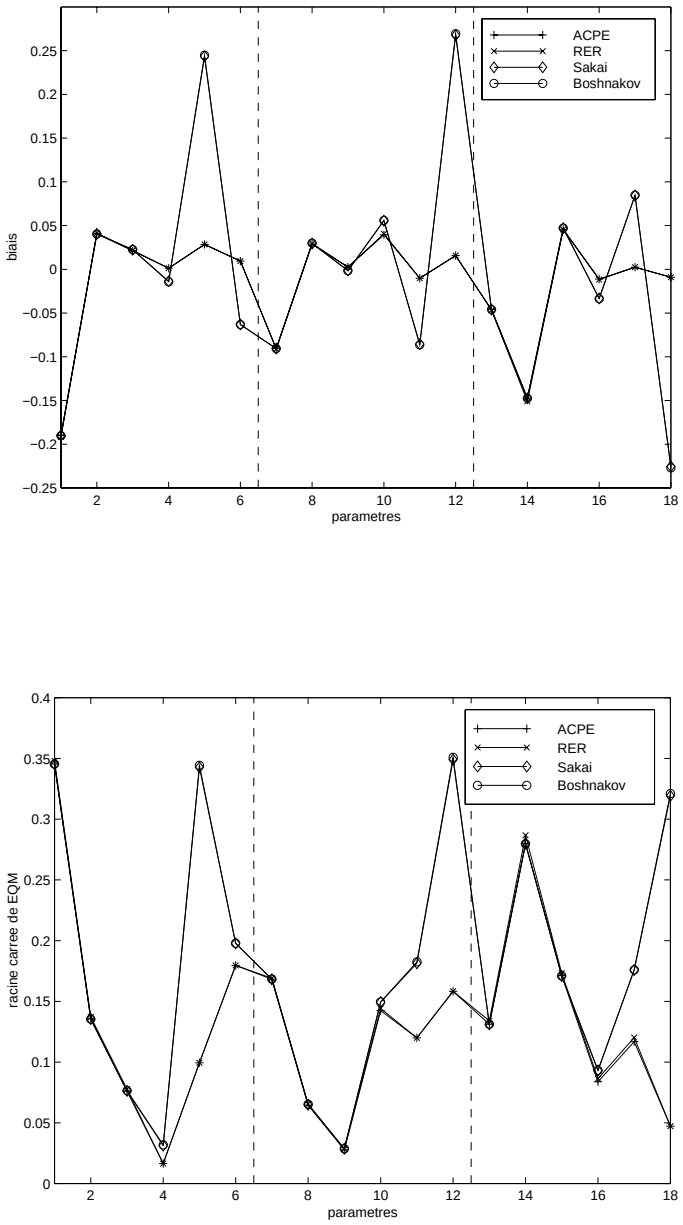


Figure 4.1 : *Biais et racine carrée de l'EQM des estimateurs des autocorrélations partielles, M2,* $m = 105$, $n_r = 2000$.

En effet rappelons que, contrairement aux deux autres méthodes, la valeur de $\hat{\beta}(t,s)$ dépend, au travers des filtres, de toutes les estimations $\hat{\beta}(u,v)$, $(u,v) \in [s,\ldots,t]^2 \setminus (t,s)$, obtenues aux étapes précédentes.

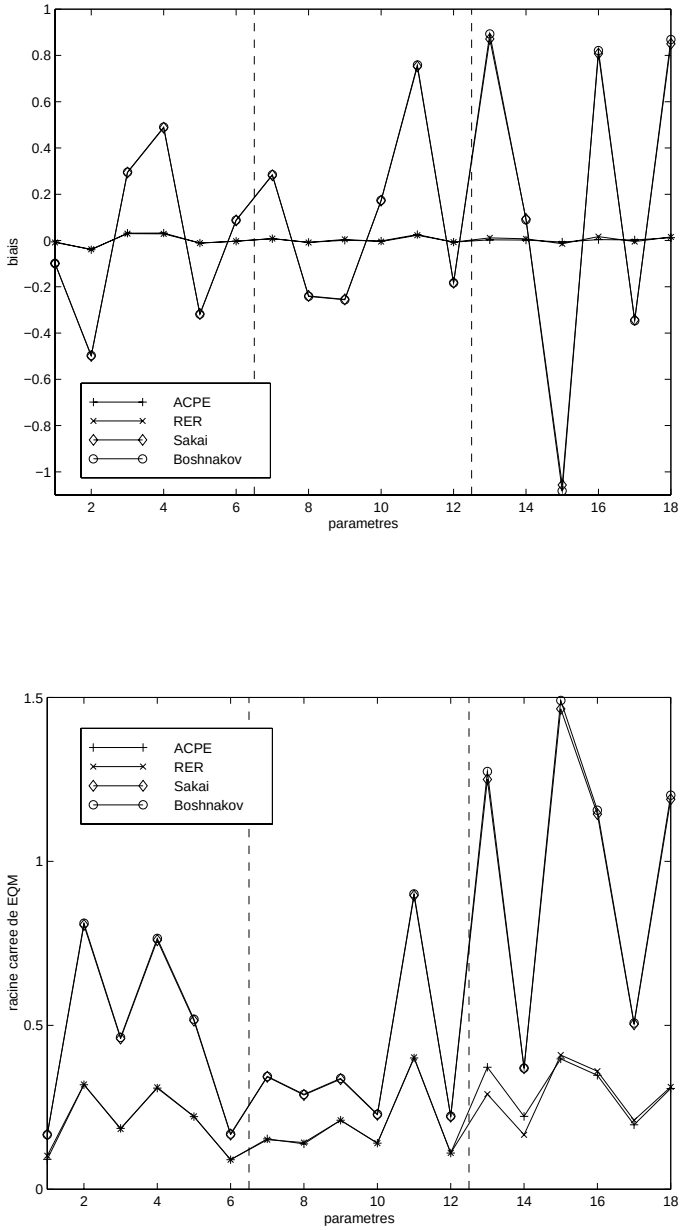


Figure 4.2 : *Biais et racine carrée de l'EQM des estimateurs des filtres, M2,* $m = 105$, $n_r = 2000$.

Ainsi les erreurs sur les coefficients $\beta_i(k)$ en module proche de 1 (cf. $\beta_1(4)$ par exemple) se répercutent sur ceux d'ordre supérieur non seulement de la même période mais aussi des autres. En particulier, ceci peut conduire à

de très mauvaises estimations pour des coefficients même s'ils ne sont pas proche de 1 en module (cf. $\beta_2(6)$) et laisse prévoir que ce défaut influe d'autant plus que l'ordre du modèle est grand. Le fait que les méthodes de Burg sous-estiment les coefficients $\beta_i(k)$ en module proche de 1 apparaît clairement dans le cas de $\sigma_\cdot^2$ où le biais correspond à une surestimation des variances résiduelles de chaque période (cf. Annexe : Table II). La Figure 4.2, à la page précédente, est l'analogue de la Figure 4.1 avec $a_1(1), \ldots, a_1(p_1)$, $a_2(1), \ldots, a_T(p_T)$, comme paramètres. Elle montre que les méthodes de Burg sont encore moins bonnes dans l'estimation des filtres, en particulier lorsque la valeur des coefficients est plus grandes (cf. dernière période). Notons que l'écart de l'EQM constaté entre ces méthodes, provient de l'augmentation du biais ainsi que celle de la variance. On observe les mêmes phénomènes pour une série deux fois plus longue, avec une légère amélioration dans l'estimation des autocorrélations partielles (cf. Annexe : Table III).

Sur l'exemple précédent, on a constaté que les méthodes de Burg semblaient sensibles à l'ordre de grandeur des coefficients des filtres. Dans la situation périodique, on peut trouver des modèles pour lesquels l'ordre de grandeur de ces coefficients est élevé sans que pour autant les modèles soient proches de la singularité. Par exemple, les modules des autocorrélations partielles de M3 ne dépassent pas 0.85, alors que les coefficients des filtres atteignent 10.853 (cf. Annexe : Table IV). Pour des séries courtes issues de ce modèle ($m = 70$), comme précédemment les quatre méthodes sont comparables pour l'estimation de $\rho(\cdot, \cdot)$. Pour les autres paramètres, les synthèses globales sont rappelées dans le tableau suivant.

Tableau 4.3 : *Comparaison entre Burg, RER et ACPE, M3, m=70.*

Paramètre	$\beta(\cdot,\cdot)$		$a_\cdot(\cdot)$		$\sigma_\cdot^2$	
	biais	$(\text{EQM})^{\frac{1}{2}}$	biais	$(\text{EQM})^{\frac{1}{2}}$	biais	$(\text{EQM})^{\frac{1}{2}}$
Sakai	0.063	0.170	0.651	0.950	0.031	0,072
Boshnakov	0.064	0.170	0.670	0.967	0.033	0.075
RER	0.046	0.165	0.097	0.514	0.024	0.035
ACPE	0.043	0.160	0.098	0.532	0.022	0.034

On observe une petite différence entre les procédures de Burg et les deux autres dans l'estimation des autocorrélations partielles, excepté pour les derniers coefficients de chaque période (cf. Figure 4.3). En particulier le biais et la racine de l'EQM de $\hat{\beta}_1(7)$ sont multipliés, pour les méthodes de Burg, par 5 et 2 respectivement. Cet écart se creuse pour le filtre de la première période (cf. Figure 4.4), où se trouvent les coefficients de grandes

valeurs. Par contre les quatre méthodes sont comparables sur la deuxième période.

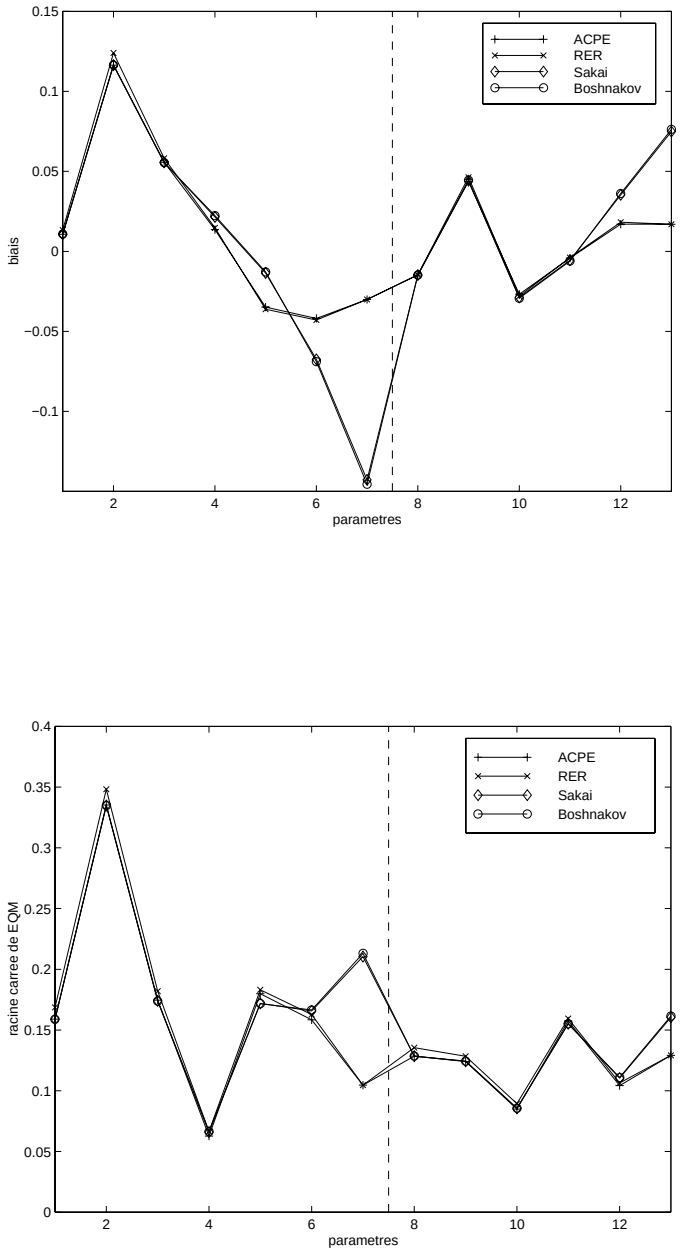


Figure 4.3 : *Biais et racine carrée de l'EQM des estimateurs des autocorrélations partielles, M3,* $m = 70$, $n_r = 2000$.

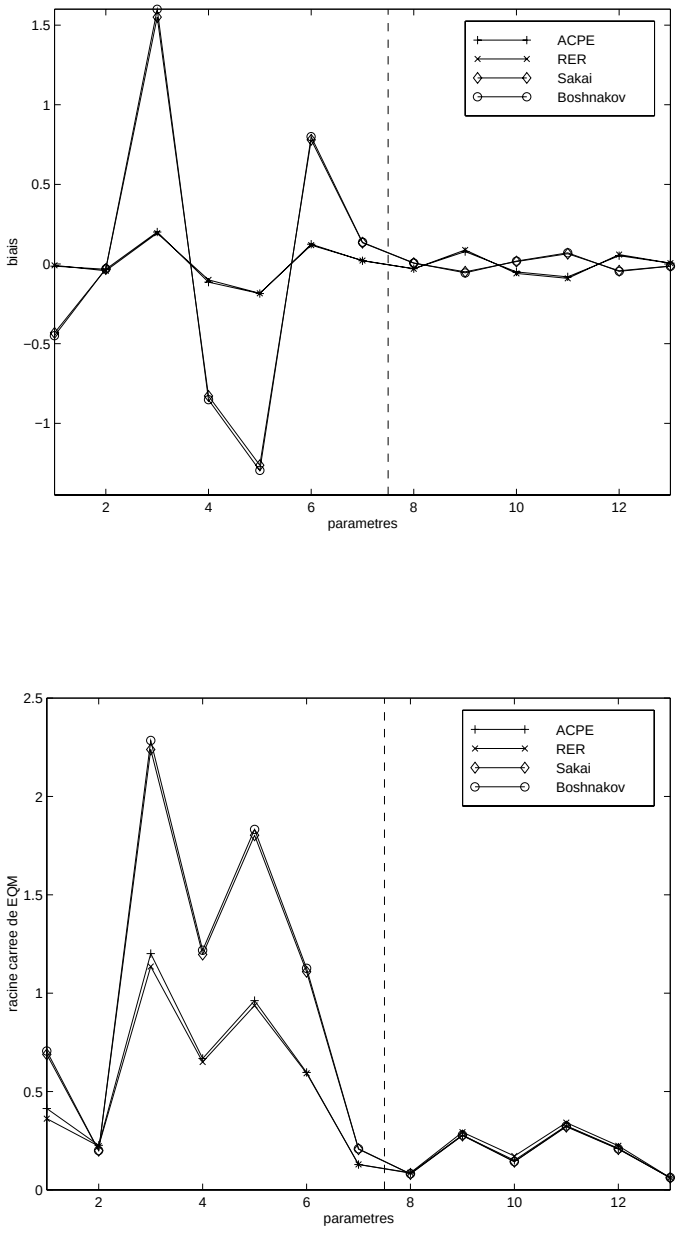


Figure 4.4 : *Biais et racine carrée de l'EQM des estimateurs des filtres, M3,* $m = 70$, $n_r = 2000$.

Remarquons que l'augmentation de l'EQM résulte de celle du biais et de celle de la variance pour les autocorrélations partielles et les filtres. Par contre, le biais des estimateurs des variances résiduelles est du même ordre

de grandeur mais avec une variance plus forte pour ceux de Burg (cf. Annexe : Table IV). Pour des séries deux fois plus longues (cf. Annexe : Table V), il existe encore une différence non négligeable dans l'estimation du filtre de la première période. Sinon les méthodes sont comparables.

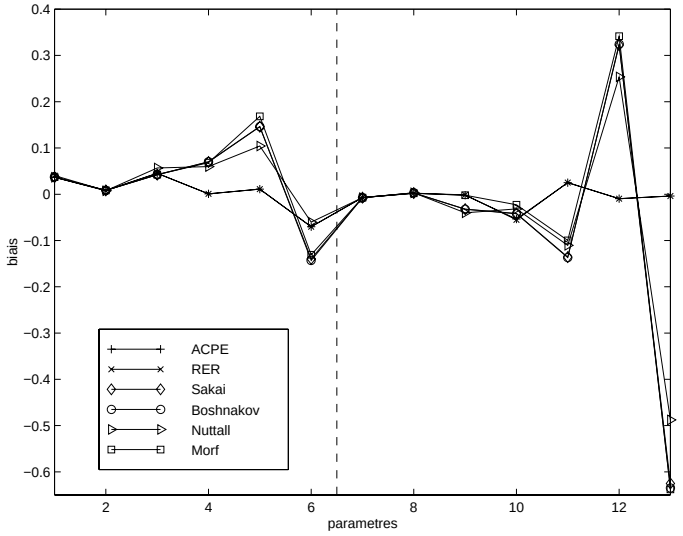
Extensions de Burg vectorielles

Les méthodes de Burg liées aux processus vectoriels stationnaires ([NUT 76], [MOR78b]) sont ici comparées à leurs homologues périodiques ([BOS94], [SAK82]). Afin de faire ressortir les éventuelles différences, nous considérons des séries courtes ($m = 70, 140$) issues du modèle M4, qui comme M2 est proche de la singularité. Les résultats obtenus par $ACPE$ et RER sont également présentés (cf. Annexe : Tables VI et VII). Notons que les méthodes vectorielles du même type conduisent à des résultats identiques puisqu'il s'agit d'un modèle $ARP(p, max)$. Concernant la comparaison des méthodes de Burg avec $ACPE$ et RER, on observe les mêmes comportements que pour M2, avec des différences plus marquées (cf. Tableau 4.4, Figures 4.5 et 4.6).

Tableau 4.4 : *Comparaison entre Burg, RER et ACPE, M4, m=70.*

Paramètre	$\beta(\cdot,\cdot)$		$a.(\cdot)$		$\sigma^2_\cdot$	
	biais	$(\text{EQM})^{\frac{1}{2}}$	biais	$(\text{EQM})^{\frac{1}{2}}$	biais	$(\text{EQM})^{\frac{1}{2}}$
Sakai	0.208	0.283	2.123	3.195	0.121	0.212
Boshnakov	0.211	0.284	2.242	3.220	0.124	0.214
Nuttall	0.162	0.241	1.511	2.479	0.079	0.152
Morf et coll.	0.213	0.288	1.700	2.955	0.116	0.202
RER	0.031	0.134	0.118	1.925	0.018	0.029
$ACPE$	0.031	0.131	0.135	1.880	0.017	0.028

Ceci peut s'expliquer par le fait que le modèle M4 possède des autocorrélations partielles en module très proche de 1 ($|\beta_2(2)| = |\beta_2(7)| = 0.99$) et que la valeur des coefficients des filtres associés peut être très élevée ($a_1(3) = 22.013$). En particulier les méthodes de Burg conduisent à des résultats très médiocres pour le coefficient $\beta_2(7)$ qui se trouve en fin de période (cf. Annexe : Tables VI et VII). C'est aussi le cas pour les filtres et on retrouve la tendance à surestimer les variances résiduelles pour les méthodes de Burg. Concernant la comparaison des méthodes de Burg entre elles, la procédure de Nuttall améliore légèrement les résultats sur l'ensemble des paramètres, exceptés pour $\rho(\cdot,\cdot)$ et $R(\cdot,\cdot)$ où les méthodes sont comparables. Pour la procédure de Morf et coll., on constate une légère amélioration uniquement sur le filtre de la première période (cf. Figure 4.6).



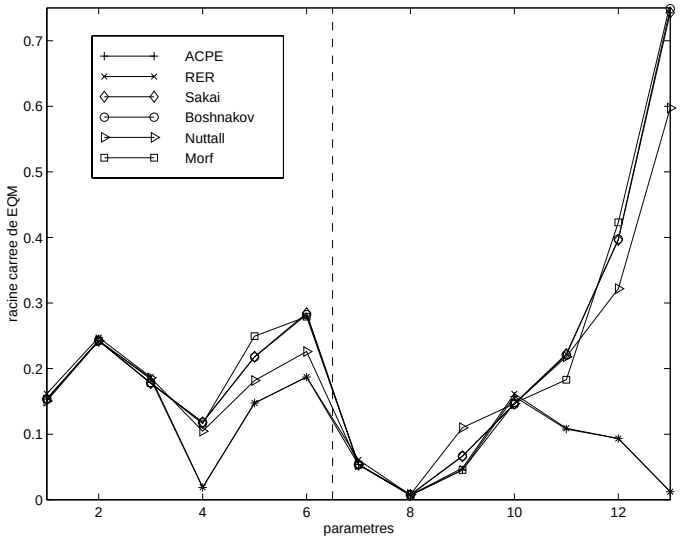
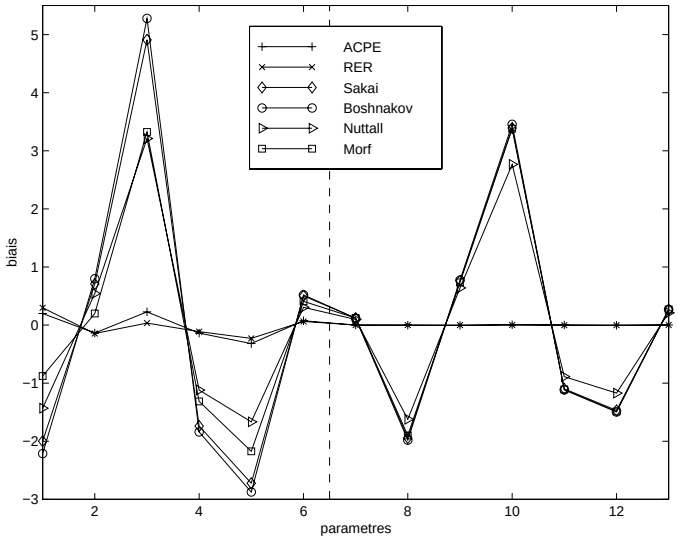


Figure 4.5 : *Biais et racine carrée de l'EQM des estimateurs des autocorrélations partielles, M4,* $m = 70$, $n_r = 2000$.



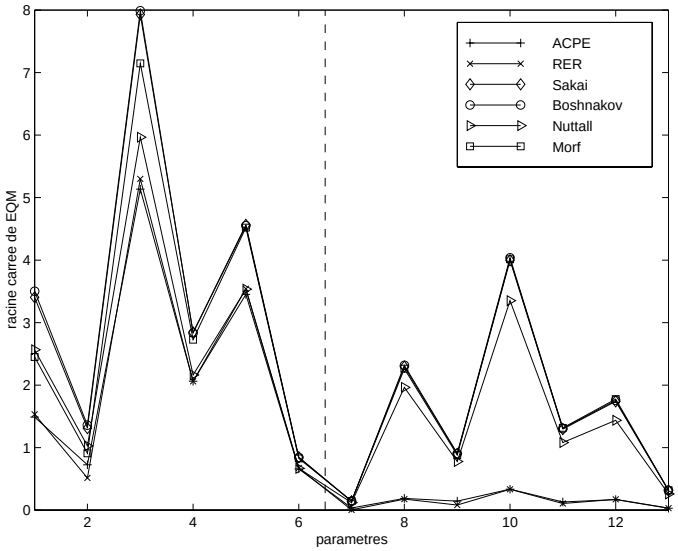
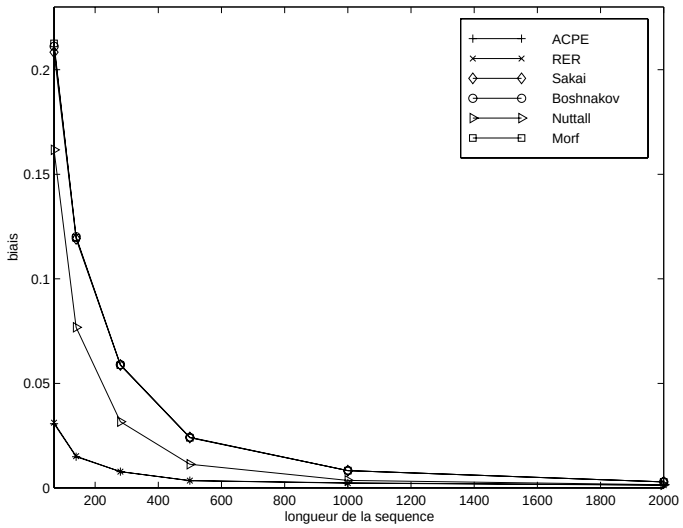


Figure 4.6 : *Biais et racine carrée de l'EQM des estimateurs des filtres, M4,* $m = 70$, $n_r = 2000$.



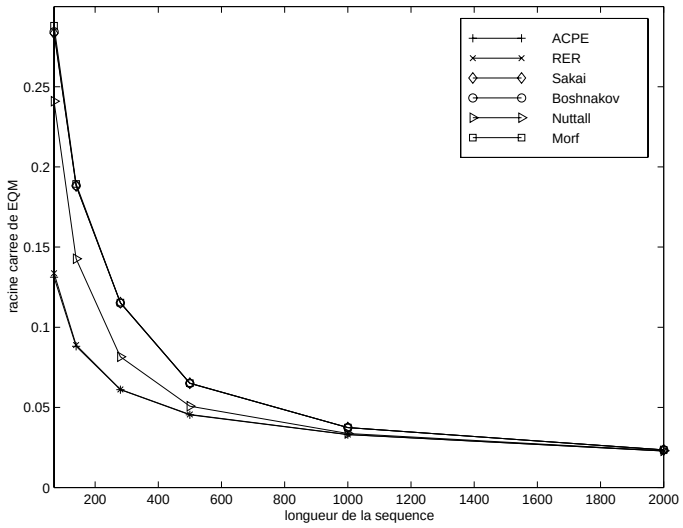
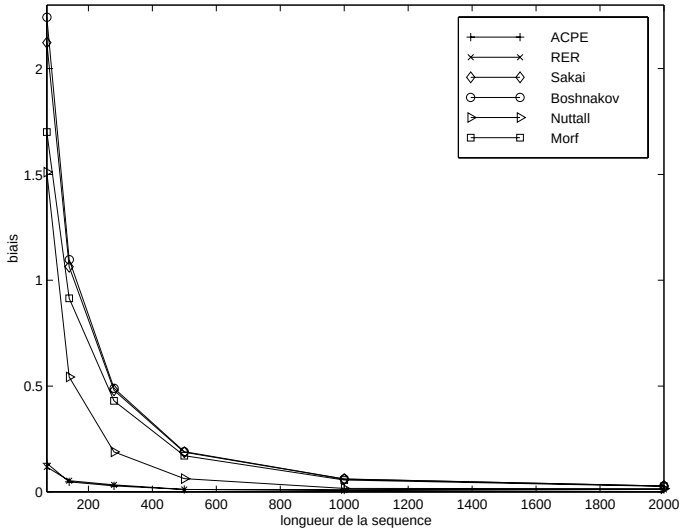


Figure 4.7 : *Synthèse globale du biais et de l'EQM pour les estimateurs des autocorrélations partielles, M4.*



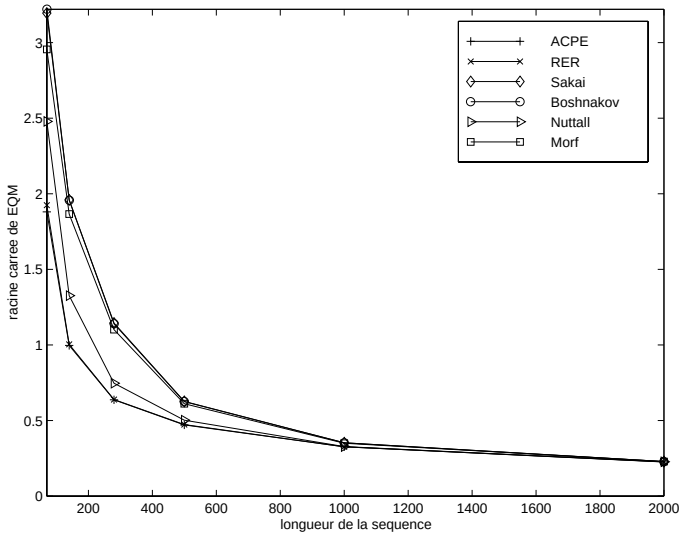
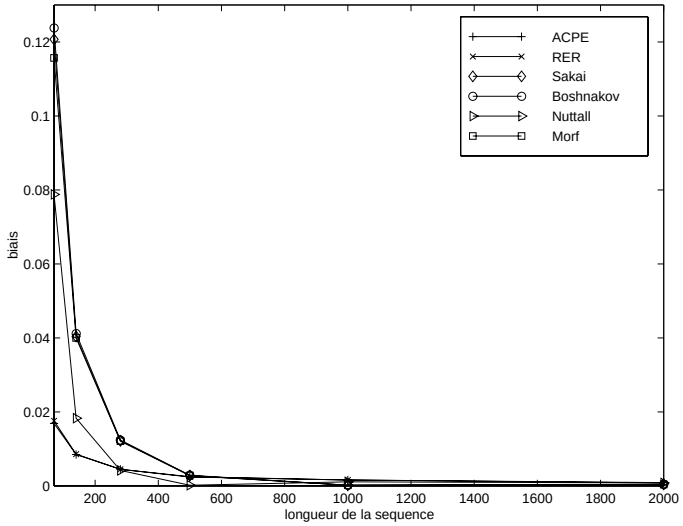


Figure 4.8 : *Synthèse globale du biais et de l'EQM pour les estimateurs des filtres, M4.*



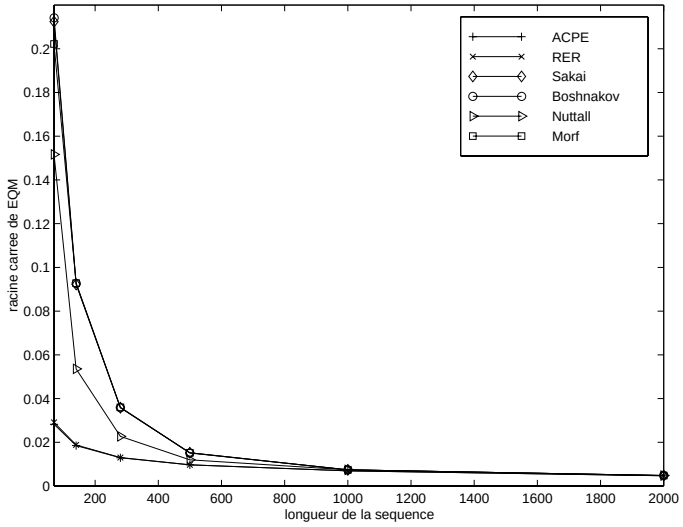


Figure 4.9 : *Synthèse globale du biais et de l'EQM pour les variances résiduelles, M4.*

Ceci confirme que le problème des méthodes de Burg provient essentiellement des contraintes dans la construction récursive des filtres. Rappelons que les procédures de Burg vectorielles introduisent moins de contraintes que celles

scalaires (cf. sous-section 4.2.4). En effet à chaque étape k, les contraintes des approches vectorielles qui portent sur les erreurs progressive et rétrograde, se traduisent d'un point de vue scalaire sur $2 \times T$ erreurs. Pour les méthodes scalaires à la même étape, les filtres intermédiaires associés aux innovations partielles $\varepsilon^f(kT+i; j+1)$ et $\varepsilon^b(j; kT+i \Leftrightarrow 1)$, $i, j = 1, \ldots, T$ sont aussi contraints, soit au total $2 \times T^2$. Ainsi plus la période T est grande, plus les approches scalaires introduisent de contraintes par rapport à celles vectorielles. Ces dernières semblent donc préférables. Notons cependant que d'un point de vue calculatoire, elles sont plus coûteuses et plus difficiles à mettre en œuvre, comme par exemple la résolution de l'équation de Lyapounov pour [NUT76]. De plus, rappelons que les méthodes vectorielles ne permettent d'estimer que les modèles $ARP(p, max)$.

Les Figures 4.7, 4.8 et 4.9 ci-avant, représentent l'évolution de la synthèse globale du biais et de l'EQM pour le modèle M4 en fonction de la longueur de la séquence. Nous considérons les résultats obtenus par les estimateurs des quatre méthodes de Burg, $ACPE$, et RER. Ils sont donnés dans la Table VIII en Annexe. Les longueurs des séquences considérées sont égales successivement à $m = 70$, 140, 280, 500 avec $n_r = 2000$, $m = 1000$ avec $n_r = 1000$, et $m = 2000$ avec $n_r = 500$. Nous ne représentons pas les graphes associés à $\rho_\cdot(\cdot)$ et $R(\cdot, \cdot)$ car pour ces paramètres les méthodes conduisent à des résultats similaires. Ces Figures montrent que la méthode de Nuttall se détache très nettement des autres procédures de Burg. Cependant ses performances restent bien inférieures à celle d'$ACPE$ et de RER. La méthode de Morf et coll. quant à elle, demeure très proche de celles scalaires. Ceci peut s'expliquer par le fait qu'elle n'est pas très éloignée dans sa conception de celle de Sakai lorsque T est petit (cf. sous-section 4.2.4). Par ailleurs, on constate le bon comportement d'$ACPE$ et de RER. En effet il faut attendre une série de longueur 1000 pour que les méthodes deviennent comparables. Ceci indique que dans certaines situations, l'utilisation d'$ACPE$ ou RER est préférable même sur des séries de taille moyenne.

Annexe

Bibliographie

[Alp95] D. Alpay, A. Chevreuil, P. Loubaton, *Extension des cyclocorrélations et estimation spectrale au sens du maximum d'entropy pour les processus aléatoires périodiquement corrélés.* 15 ème Colloque GRETSI-Juan-Les-Pins: 33-36, 1995.

[Ben58] W. R. Bennett, *Statistics of regenerative digital transmission.* Bell System Tech. J. **37**: 1501-1542, 1958.

[Bos94] G. Boshnakov, *Periodically correlated sequences: some properties and recursions.* Research Report 1, Division of Quality Technology and Statistics, Luleo University, Sweden, 1994.

[Box94] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, *Time Series Analysis, Fore-casting and Control.* Third Edition. Prentice Hall, Englewood Cliffs, New Jersey, 1994.

[Bur75] J. P. Burg, *Maximum entropy spectral analysis.* Ph. D. dissertation, Dep. Geop. Standford Uni., CA, 1975.

[Cra61] H. Cramér, *On some classes of nonstationary stochastic processes.* Proceeding of 4th Berkeley Symposium on Mathematics, Statistic and Probability, University of California Press **2**: 57-78, 1961.

[Dég86] S. Dégerine, *Autocorrélation partielle empirique et applications.* Rapport de Recherche, TIM3, RR 627-M-, 1986.

[Dég88] S. Dégerine, *Fonction d'autocorrélation partielle et estimation autorrégressive dans le domaine temporel.* Thèse de Doctorat d'État, Université J.Fourier, Grenoble, 1988.

[Dég90] S. Dégerine, *Canonical partial autocorrelation function of a multivariate time series.* Annals of Statistics, **18**(2): 961-971, 1990.

[Dég92] S. Dégerine, *Autocorrélation partielle empirique d'une série chronologique vectorielle.* Rapport de Recherche, LMC (URA 397), RR 896-M-, 1992.

[Dég93] S. Dégerine, *Sample partial autocorrelation function.* IEEE Transactions on Signal Processing, **41**(1): 403-407, 1993.

[Dég94] S. Dégerine, *Sample partial autocorrelation function of a multivariate time series.* Journal of Multivariate Analysis, **50**(2): 294-313, 1994.

[Dég96a] S. Dégerine, S. Lambert, *Fonction d'autocorrélation partielle des processus à temps discret non stationnaires et spectre évolutif associé.* Rapport Technique, LMC (URA 397), RT 156, 1996.

[Dég96b] S. Dégerine, S. Lambert, *Evolutive instantaneous spectrum associated with partial autocorrelation function for nonstationary time series.* Proceeding TFTS-96: 457-460, 1996.

[Del79] PH. Delsarte, Y. Genin, Y. Kamp, *Schur parametrization of definite block-Toeplitz system.* SIAM Journal of Applied Mathematics, **36**(1): 34-45, 1979.

[Del80] PH. Delsarte, Y. Genin, Y. Kamp, *A method of matrix inverse triangular decomposition based on contiguous principal submatrices.* Linear Algebra and its Applications, **31**: 199-212, 1980.

[Des90] F. Desbouvries, C. Gueguen, *A Gohberg-Semencul formula for linear time-varyng systems.* Signal Processing V : Theories and Applications, L. Torres, E. Masgrau and M. A. Lagunas eds, Elsevier Science Publishers, B. V., 1990.

[Des93] F. Desbouvries, *Recursive least-squares lattices and trigonometry in the spherical triangle.* ICASSP, **III**: 404-407, 1993.

[Dic78] B. W. Dickinson, *Autoregressive estimation using residual energies ratios.* IEEE Transactions on Information Theory **24**(4): 503-506, 1978.

[Dic79] B. W. Dickinson, *Estimation of partial correlation matrices using*

Cholesky decomposition. IEEE Transactions on Automatic Control, **24**(2): 302-305, 1979.

[Dur60] J. Durbin, *The fitting of time-series models.* Rev. Inst. Internat Statist., (28): 233-244, 1960.

[Fla93] P. Flandrin, *Temps-fréquence.* Traité des Nouvelles Technologies, Séries Traitement du Signal, Hermes, 1993.

[Gan60] F. R. Gantmacher, *Matrix theory Vol.I.* Chelsea Publishing Company New-York, N.Y., 1960.

[Ger60] Y. L. Geronimus, *Polynomials orthogonal on a circle and interval.* Pergamon Press, 1960.

[Gla61] E. G. Gladysev, *Periodically random sequences.* Soviet Mathematics, **2**: 385-388, 1961.

[Gre84] Y. Grenier, *Modélisation de signaux non stationnaires.* Thèse de Doctorat d'État, Paris Sud, 1984.

[Hal77] M. Hallin, G. Mélard, *Indéterminabilité pure et inversibilité des processus autorégressifs-moyenne mobile à coefficients dépendant du temps.* Cahier du CERO, **19**(3-4): 385-392, 1977.

[Han70] E. J. Hannan, *Multivariate time series.* Wiley, 1970.

[Ino84] Y. Inouye, *Maximum entropy spectral estimation for regular time series of degenerate rank.* IEEE Transactions on Acoustics, Speech, and Signal Processing **32**(4): 733-740, 1984.

[Jon67] R. H. Jones, W. M. Brelsford, *Time series with periodic structure.* Biometrika **54**(3-4): 403-408, 1967.

[Kai82] T. Kailath, *Time-variant and time-invariant lattice filters for nonstationary processes.* Outils et modèles mathématiques pour l'automatique, l'analyse de systèmes et le traitement du signal, Paris, éds du CNRS: 417-464, 1982.

[Kay83] S. M. Kay, *Recursive maximum likelihood estimation of autoregressive processes.* IEEE Transactions on Acoustics, Speech, and Signal Pro-

cessing **31**(1): 56-65, 1983.

[LAM96] S. LAMBERT, *Extension of autocovariance coefficients sequence for periodically correlated random processes by using the partial autocorrelation function.* Proceeding EUSIPCO 96: 503-506, 1996.

[LAM97a] S. LAMBERT, *Fonction d'autocorrélation partielle empirique des processus périodiquement corrélés.* Rapport Technique, LMC (URA 397), RT 174, 1997.

[LAM97b] S. LAMBERT, *Sur l'estimation autorégressive des processus périodiquement corrélés.* 16 ème Colloque GRETSI-Grenoble : 503-506, 1997.

[LEV47] N. LEVINSON, *The Wiener RMS (root-mean-square) error criterion in filter design and prediction.* Journal of Mathematics Physic, (25): 261-278, 1947.

[LEV81] H. LEV-ARI, T. KAILATH, *Schur and Levinson algorithms for non-stationary processes.* Proc. IEEE Int. Conf. Acoustics, Speech Signal Process. (Atantla, GA, mar. 1981): 860-864, 1981.

[LOY68] R. M. LOYNES, *On the concept of the spectrum for non-stationary processes.* Journal Royal Statistics Society, **B**(30): 1-30, 1968.

[MÉL78] G. MÉLARD, *Propriétés du spectre évolutif d'un processus non stationnaire.* Ann. Inst. Henri Poincaré, **XIV**(4): 411-424, 1978.

[MIL68] K. S. MILLER, *Linear difference equations.* New York : Benjamin, 1968.

[MOR78a] M. A. MORF, A. VIEIRA, T. KAILATH *Covariance characterization by autocorrelation matrices.* Annals of Statistics, **6**(3): 643-648, 1978.

[MOR78b] M. A. MORF, A. VIEIRA, D. LEE, T. KAILATH *Recursive multichannel maximum entropy spectral estimation.* IEEE Transactions on Geoscience Electronics, **16**(2): 85-94, 1978.

[NUT76] A. H. NUTTALL, *Positive definite spectral estimate and stable correlation recursion for multivariate linear predictive spectral analysis.* NUSC Tech. Doc. 5729, Naval Underwater System Center, New London, CT, Nov. 14, 1976.

[PAG78] M. PAGANO, *On periodic and multiple autoregressions.* Annals of Statistics, **6**(6): 1310-1317, 1978.

[PHA88] D. T. PHAM, *Maximum likelihood estimation of autoregressive model by relaxation on the reflection coefficients.* IEEE Transactions on Acoustics, Speech, and Signal Processing **38**(1): 175-177, 1988.

[PHA92] D. T. PHAM, *Fast scalar algorithms for multivariate partial correlations and their empirical counterparts.* Rapport de Recherche, LMC (URA 397), RR 903-M-, 1992.

[PRI65] M. B. PRIESTLEY, *Evolutionary spectra and non-stationary processes.* Journal Royal Statistics Society, **B**(27): 204-237, 1965.

[QUE49] M. H. QUENOUILLE, *Approximate tests of correlation in time series.* Journal Royal Statistics Society, **B**(11): 68-84, 1949.

[RAM74] F. L. RAMSEY, *Caracterization of the partial autocorrelation function.* Annals of Statistics, **2**(6): 1296-1301, 1974.

[RAO65] C. R. RAO, *Linear statistical inference and its applications.* Wiley, New York, 1965.

[SAK82] H. SAKAI, *Circular lattice filtering using Pagano's method.* IEEE Transactions on Acoustics, Speech, and Signal Processing **30**(2): 279-287, 1982.

[SAK83] H. SAKAI, *Covariance matrices characterization by a set of scalar partial autocorrelation coefficients.* Annals of Statistics, **11**(1): 337-340, 1983.

[STR77] O. N. STRAND, *Multichannel complex maximum entropy (autoregressive) analysis.* IEEE Transactions on Automatic Control, **22**(4): 634-640, 1977.

[TJØ76] D. T. TJØSTHEIM, *Spectral generating operators for non-stationary processes.* Adv. Appl. Prob. **8**(4): 831-846, 1976.

[WHI63] P. WHITTLE, *On the fitting of multivariate autoregression and approximate canonical factorization of a spectral density matrix.* Biometrika

50(1-2): 129-134, 1963.

[Wie57] N. Wiener, P. Masani, *The prediction theory of multivariate stochastic processes, I.* Acta. Math., **98**: 111-150, 1957.

[Wol54] H. Wold, *A study in the analysis of stationary time series.* Stockholm, Almqvist & Wiksell, 1954.

[Yul07] G. U. Yule, *On the theory of correlation for any number of variables, treated by a new system of notation.* Royal Society Proceeding, **79**(A): 85-96, 1907.