N° d'ordre : **24 / 2003-M / MT**

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

Université des Sciences et de la Technologie Houari Boumediene
Faculté de Mathématiques

Mémoire présenté pour l'obtention du diplôme de Magister

En Mathématiques

Spécialité : Recherche Opérationnelle (Méthodes Stochastiques)

Par : Mr. HAMDI Fayçal

Thème

Etude théorique et algorithmique de la performance des critères d'identification des ARMA périodiques

Soutenu le 16 Décembre 2003, devant le jury composé de :

Mr. M. AÏDER, Maître de Conférences, U.S.T.H.B	Président
Mr. M. BENTARZI, Professeur, U.S.T.H.B	Directeur de thèse
Mr. A. AÏSSANI, Professeur, U.S.T.H.B	Examinateur
Mr. O. ANES, Maître de Conférences, I.N.P.S	Examinateur
Mme. H. GUERBYENNE, Chargée de Cours, U.S.T.H.B	Examinateur

Remerciements

A l'issue de ce travail, je tiens à exprimer toute ma gratitude à l'ensemble des personnes qui ont contribué, chacune à leur manière, à l'accomplissement de cette thèse.

Je voudrait tout d'abord exprimer mes chaleureux et sincères remerciements à Monsieur, le Professeur, Mohamed Bentarzi, pour son encadrement, sa patience et surtout pour tous les conseils avisés qu'il a su me prodiguer tout au long de ce travail. Je lui témoigne toute ma gratitude ainsi que ma reconnaissance la plus vive.

Je voudrais également exprimer mes remerciements et ma reconnaissance à Monsieur, H. Aïder, Maître de Conférences à l'U.S.T.H.B., de m'avoir honorer par son aimable acceptation à présider le Jury de cette thèse.

En toute humilité, je tiens à exprimer ma profonde gratitude à Monsieur, le Professeur A. Aïssani, Doyen de la faculté des sciences à l'U.S.T.H.B., de m'avoir fait l'honneur de faire partie du Jury de cette thèse.

Je voudrais également témoigner ma grande reconnaissance et mes plus vifs remerciements à Monsieur, O. Anes, Maître de Conférences à l'I.N.P.S., de m'avoir fait l'honneur par son acceptation d'appartenir au Jury de cette thèse.

Ma reconnaissance et mes remerciements vont à Madame, H. Guerbyenne, Chargée de Cours à l'U.S.T.H.B., de m'avoir fait l'honneur par son acceptation de participer au Jury de cette thèse.

Je tiens à remercier Monsieur, A. Aknouche, pour son aide et pour les enrichissantes et fructueuses discussions.

Avant de terminer, je voudrais dédier cette thèse à mes parents, à toute ma famille et en particulier à Louisa, qui m'a encouragée dans les périodes de doute, qui m'a aidée et qui m'a apportée soutien.

Enfin, j'exprime toute ma gratitude à mes collègues du Crédit Populaire d'Algérie, qui ont partagé au quotidien mes espoirs et mes inquiétudes, qui m'ont réconforté dans les moments difficiles et avec qui j'ai partagé d'inoubliables instants de détente. Je vous remercie tous chaleureusement : K. Boualem, M. Rabea, O. Hafida et T. Zakia.

Table des matières

Introduction générale

0.1 Introduction

L'évolution de l'économie mondiale vers la moitié du siècle dernier, avait engendré une tendance générale à l'industrialisation et à la production à grande échelle. A cet effet, le développement des séries chronologiques fut une traduction explicite d'une logique économique qui nécessite le contrôle et la prédiction des fluctuations pour une meilleur rentabilisation.

Le théorème de la décomposition de Wold (1938) est la base du développement considérable qu'a connue la classe des modèles linéaires à coefficients constants, dits *autorégressifs moyennes mobiles*, notés ARMA.

L'utilisation de ces modèles ARMA pour représenter et analyser des séries chronologiques est devenue courante depuis les travaux de Box et Jenkins (1970). Ce sont des processus stochastiques adaptés à des fins de prévisions et construits pour fournir une explication du présent par le passé.

La méthodologie de Box et Jenkins (1970), permet de trouver, en plusieurs étapes, un modèle ARMA susceptible de représenter une séries chronologique. Elle n'est en fait que l'application de la méthode scientifique afin d'obtenir le modèle (ARMA) réel générant la série.

Rappelons que la méthode scientifique consiste à formuler des supposition sous forme d'un modèle, à les mettre à l'épreuve et à réviser le modèle en conséquence. Ces étapes sont répétées autant de fois que nécessaire. Une fois le modèle ARMA connu, on peut déterminer mécaniquement les prévisions.

Afin de trouver le meilleur modèle ARMA représentant le processus générant la série, Box et Jenkins (1970) ont proposé une méthodologie en trois étapes permettant le choix du modèle adéquat. En l'occurrence : l'identification du modèle, l'estimation de ses paramètres et les tests d'adéquation.

L'identification, ou la recherche de modèles à partir de données expérimentales, est une préoccupation majeure dans la plupart des disciplines scientifiques. Elle désigne à la fois une démarche scientifique et un ensemble de techniques visant à déterminer des modèles mathématiques capables de reproduire aussi fidèlement que possible le comportement d'un processus physique, chimique, biologique, économique etc.

Dans la littérature des séries chronologiques, nous pouvons distinguer trois catégories de méthodes pour identifier un modèle ARMA stationnaire :

• Les méthodes basées sur la fonction d'autocorrélation, la fonction d'autocorrélation partielle et la fonction d'autocorrélation inverse. Ces méthodes sont fondées sur la comparaison des moments empiriques de la série considérée aux moments théoriques associés aux différentes représentations potentielles.
• Les méthodes basées sur la théorie de l'information. Il s'agit de choisir le modèle, en se basant sur une mesure de l'écart entre la vraie loi (inconnue) et celle du modèle proposé. La mesure habituellement utilisée est l'information de Kullback.
• Les méthodes basées sur l'analyse bayesienne. Il s'agit de choisir le modèle qui a la plus grande probabilité a posteriori. Cette classe de méthodes est très importante, car elle résoud le problème de l'inconsistance des estimateurs obtenus.

En utilisant la fonction d'autocorrélation et la fonction d'autocorrélation partielle Box et Jenkins (1970) ont présenté une méthode d'identification. Cette dernière, n'est pas très utile pour identifier un modèle ARMA mixte lorsque $p \neq 0$ et $q \neq 0$, car la fonction d'autocorrélation d'un processus ARMA(p, q) mixte se comporte comme un mélange de fonctions exponentielles / sinusoïdales amorties après $q - p$ premiers retards et la fonction d'autocorrélation partielle se compose également, d'un mélange de fonctions exponentielles / sinusoïdales atténuées après $p - q$ premiers retards.

Au début des années 70, Akaike a présenté deux critères d'identification connus sous le nom FPE (*Final Prediction Error*) et AIC (*Akaike Information Criterion*). L'inconsistance des estimateurs obtenus a provoquée quelques objections sur ces deux critères. Pour cela le critère BIC (*Bayesian Information Criterion*) et la méthode de Hannan et Quinn (1979) on été proposés. A ce moment, Cleveland a suggéré l'utilisation de la fonction d'autocorrélation inverse. En 1982, Hannan et Rissanen ont utilisé une technique instrumentale de régression en tant que fonction de prévision pour modéliser un processus ARMA. Cette technique a eu comme conséquence, des estimateurs consistants des ordres.

Tout au long des années 70 et 80, plusieurs méthodes d'identification basées sur certaines fonctions de la fonction d'autocorrélation ont été dérivées. Parmi ces méthodes nous pouvons citer la méthode Woodside (1971), la méthode R et S de Gris et *al.* (1978), la méthode du Coin de Beguin et *al.* (1980), les trois méthodes GPAC (*Generalized Partial Autocorrelation*) de Woodward et Gris (1981), de Glasbey (1982) et de Takemura (1984), la méthode ESACF (*Extended Sample Autocorrelation function*) de Tsay et Tiao (1984) et la méthode SCAN de Tsay et Tiao (1985).

Au cours des années 90 d'autres critères de sélection ont été proposé dont la méthode de Choi (1991a), le critère PDC (*Predictive Density Criterion*) de Djuric et Kay (1992) et le critère CIC (*Consistent Information Criterion*) de Ciftcioglu et *al.* (1994).

La plupart des résultats et des méthodes utilisées dans l'analyse des séries chronologiques sont basés sur l'hypothèse de stationnarité faible, du processus qui génère la série temporelle. Lorsque cette hypothèse n'est pas satisfaite, ce qui est fréquent en pratique, ces résultats et ces méthodes ne sont valables que si le processus est stationnaire ou peut être ramené au cas stationnaire par une transformation adéquate : différence ordinaire, différence saisonnière, différence mixte, la famille exponentielle de Box-Cox (1964) ... etc.

En pratique, il est bien connu que de nombreuses séries chronologiques présentent un caractère non régulier qui ne peut être stationnarisé par une transformation adéquate. Par conséquent, le traitement de ce genre de séries en utilisant les modèles classiques $ARIMA$ et $SARIMA$ de Box et Jenkins serait inefficace.

La généralisation du théorème de la décomposition de Wold (1938) aux processus non stationnaires par Cramer (1961), a permis d'élargir la classe des modèles linéaires autoregressifs moyennes mobiles à coefficients constants à la classe des modèles ARMA à coefficients évolutifs dans le temps notés ARMA_t. Dans cette thèse, nous nous intéressons à une classe particulière de ces modèles qui est celle des modèles dont les paramètres sont périodiques dans le temps. Cette classe de modèles permet de représenter les processus du second ordre, dont les fonctions d'autocovariance sont périodiques dans le temps.

Par conséquent, plusieurs séries chronologiques saisonnières rencontrées dans divers domaines tels que l'économie (Franses (1996), Ghysels et Hell (1992), Parzen et Pagano (1979)etc.), l'hydrologie (McLeod (1993), Vecchia (1985), la météorologie ... etc. peuvent être représentées par des modèles ARMA périodiques.

Un autre intérêt de ces modèles périodiques, outre l'étude des phénomènes saisonniers, est dans la relation établie par le théorème de Gladyshev (1961), entre ces modèles et les modèles multivariés autoregressifs moyennes mobiles stationnaires. En effet, ils peuvent être exploités dans l'analyse des séries chronologiques multivariées stationnaires, dans le but de réduire sensiblement, le nombre de paramètres.

Plusieurs travaux de recherche ont été effectués pour étudier les propriétés théoriques des modèles ARMA périodiques. Nous pouvons distinguer deux approches différentes : La première, dite "*period-span-lumping*", se base directement sur le théorème de Gladyshev (1961) et consiste à ramener un processus périodiquement corrélé univarié au processus stationnaire multivarié qui lui est associé et d'étudier, au sens du second ordre, les propriétés de ce dernier. Notons que plusieurs chercheurs ont appliqué cette approche pour étudier différents problèmes liés aux modèles linéaires à coefficient périodiques tels que la propriété de l'inversibilité, parmis lesquels nous citons Pagano (1978), Newton (1982), Cipra (1984). Nous soulignons que la condition d'inversibilité obtenue par ces auteurs n'est que suffisante, ce qui a été montré dans Bentarzi et Hallin (1994).

La deuxième approche, dite "*order-span-lumping*" a été introduite par Bentarzi et Hallin (1994), pour étudier quelques propriétés des modèles périodiques. Cette technique consiste à représenter le modèle univarié (m-varié) moyenne mobile d-périodique d'ordre q par un modèle q-varié (mq-varié) moyenne mobile S-périodique d'ordre 1, où S est en fonction de d et q. Ainsi, le modèle obtenu est exploité pour étudier les propriétés d'un processus moyenne mobile périodique et en particulier la condition nécessaire et suffisante d'inversibilité.

Du fait que les conditions d'inversibilité sont analogues analytiquement aux conditions de causalité. Ula et Smadi (1997) ont exploité cette nouvelle approche de Bentarzi et Hallin (1994), afin de déterminer les conditions de causalité d'un processus autoregressif m-varié d-périodique.

De nombreux travaux de prospection et d'analyse des modèles ARMA périodiques, avec leurs trois volets : identification du modèle, estimation des paramètres et validation des modèles trouvés, ont constitué jusqu'à présent le centre d'intérêt de plusieurs chercheurs dont Pagano (1978), Tiao et Grupe (1980), Andël (1983), Cipra (1985), Vecchia (1985a) et (1985b), Anderson et Vecchia (1993), McLeod (1993), Bentarzi et Hallin (1994), (1996) et (1998), Boshnakov (1996), Franses (1996), Ghysels et al. (1996), Hemis (1999), Bentarzi (2000), Lund et Basawa (1999) et (2001), Bentarzi et Aknouche (2002) et (2003) ... etc. Ces travaux de recherche se sont soldés par de nombreux résultats intéressants.

En ce qui concerne l'estimation des paramètres, nous notons qu'on peut généralement classer les méthodes d'estimation en deux grands groupes selon le type de données sur lesquelles elles se basent :

• Les méthodes *hors-ligne* qui se basent sur des échantillons de taille fixe. Ces méthodes ont connues un développement important. Plusieurs chercheurs se sont penchés sur ce problème, citons par exemple : Pagano (1978) qui a généralisé les estimateurs de Yule Walker au cas périodique, Cleveland et Tiao (1979) qui ont traité le même problème mais avec une approche differente, Tiao et Grup (1980), Salas et al. (1982), Andël (1983), Sakai (1982) qui a amélioré la méthode de Pagano (1978) en proposant une approche récursive vis-à-vis de l'ordre du modèle, Cipra (1985), Vecchia (1985a) qui a proposé un algorithme efficace basé sur le critère du maximum de vraisemblance, Boshnakov (1996) qui a généralisé la méthode de Sakai (1982) et la méthode de Franke (1985).

Ce type de méthodes n'est pas pratique lorsqu'on dispose de données *en-ligne* car à chaque nouvelle donnée introduite, la procédure d'estimation est refaite pour tout le bloc de données, ce qui est coûteux en temps et en espace mémoire.

• Dans le deuxième groupe, on retrouve les méthodes d'estimation *en-ligne* qui sont basées sur des données disponibles progressivement. Ce genre d'estimation a été étudié par Adams et Goodwin (1995), Bentarzi et Aknouche (2002) et (2003).

Quant à l'identification, objet de notre étude, elle est en continuelle expansion, depuis la fin des années 70. Cleveland et Tiao (1976), Sakai (1982), Hamdi (1982), Vecchia (1985b), Mcleod (1992), Bentarzi (2000), etc. avaient pour objectif principal la généralisation des méthodes classiques relatives aux modèles ARMA stationnaires à coefficients constants, au cas des modèles ARMA périodiques.

La plupart de ces travaux visaient à généraliser les méthodes basées sur la théorie de l'information tel que le critère AIC ou celles basée sur l'analyse baysienne, tels que les critères BIC et PDC.

Hamdi (1982) a généralisé la méthode du coin présentée par Beguin et al. (1980) du cas des modèles ARMA stationnaire à coefficients constants au cas des modèles autoregressifs moyennes mobiles à coefficients et ordres dépendants du temps ($\text{ARMA}_t(p_t, q_t)$). Hemis (1999) et Bentarzi (2000) ont généralisé une méthode basée sur l'analyse baysienne (*Critère de la densité prédictive* (PDC)) introduite par Djuric et Kay (1992). Dans le même travail, Hemis (1999) s'est intéressée d'une part à l'estimation, par la technique de Gibbs, des paramètres du modèle autorégressif périodique (PAR) à tendance explicative. D'autre part, elle s'est intéressée à l'étude de l'effet de l'application des estimateurs obtenus sur la performance des critères de sélection de l'ordre, d'un processus autorégressif pur périodique, en introduisant la technique de Gibbs.

L'intérêt de notre travail sera porté sur le problème de l'identification d'un modèle autorégressif moyenne mobile périodique (PARMA). Nous tentons, d'une part, d'étudier et de généraliser quelques méthodes de sélection des ordres d'un modèle ARMA classique au cas d'un modèle ARMA périodique. D'autre part, examiner et évaluer la performance des critères généralisés.

0.2 Apport et présentation de la thèse

Notre travail, dont le thème est "*Etude théorique et algorithmique de la performance des critères d'identification des ARMA périodiques*", est constitué de deux parties. Chaque partie se compose de deux chapitres.

Nous présentons, brièvement, les aspects étudiés.

Première partie :

Identification des modèles ARMA stationnaires

Il apparaît indispensable de rappeler les méthode classique les plus connues dans la littérature des séries chronologiques, puisque l'objectif principal des travaux de recherche effectués dans l'identification des modèles ARMA périodique était la généralisation des méthodes classiques relatives aux modèles ARMA stationnaires à coefficients constants.

Chapitre 1 :

Modèle autorégressif moyenne mobile (ARMA) stationnaire

Ce chapitre est un rappel des définitions et des principaux résultats concernant les modèles ARMA stationnaires. Nous rappelons les conditions de causalité et d'inversibilité, les méthodes d'estimation des paramètres du modèle et les algorithmes les plus connues.

Chapitre 2 :

Identification des modèles ARMA stationnaires

Dans le deuxième chapitre, nous exposons les méthode et les critères les plus connus d'identification d'un modèle ARMA stationnaire. Comme nous nous intéressons à la performance des différents critères, nous faisons une étude de simulation dans le but de :

- Examiner et évaluer la performance de chaque critère.
- Comparer les performances des différents critères.

Deuxière partie :

Identification des modèles ARMA d-périodiques

Chapitre 3 :

Modèle autorégressif moyenne mobile périodique (PARMA)

Le troisième chapitre de notre travail se veut être un aperçu sur l'étude des processus périodiquement corrélés et les modèles ARMA périodiques. Tout d'abord, nous exposons la définition d'un processus périodiquement corrélé, le théorème de Wold-Cramer (1961) ainsi que la définition de la classe des modèles PARMA. Nous étudions ensuite, les propriétés théoriques de cette classe de modèles. En dernier lieu, nous exposons quelques méthodes et algorithmes d'estimation des paramètres d'un modèle PARMA.

Chapitre 4 :

Identification des modèles ARMA d-périodique

Dans le quatrième chapitre, nous nous intéressons au problème de l'identification des modèles autoregressifs moyennes mobiles (PARMA). Ce chapitre est élaboré dans la volonté de présenter quelques généralisation des méthodes classique et d'autre part, effectuer une comparaison entre la performance des ces différentes méthodes. Nous commençons par la généralisation de la méthode du coin de Beguin et al. (1980) du cas stationnaire au cas périodique. Nous exposons la forme des critères AIC, BIC et HQC dans le cas des modèles PARMA. Nous passons en revue par la suite, les travaux de Hemis (1999) et Bentarzi (2000) sur l'identification des modèles autorégressifs périodiques. Nous récapitulons ensuite, la généralisation du critère CIC (*Consistent Information Criterion*) aux cas des modèles PAR.

Nous terminons enfin ce chapitre, par une étude de simulation, en comparant les performances des critères étudiés appliqués à des séries artificielles supposées obtenues à partir des modèles bien spécifiés.

Première partie

Identification des modèles ARMA stationnaires

Chapitre 1

Modèle autoregressif moyenne mobile (ARMA) stationnaire

Introduction

Puisque nous traitons, dans la première partie, l'identification des modèles autorégressifs moyennes mobiles (ARMA), il sera indispensable de rappeler dans un premier chapitre les principaux résultats (sans démonstration) concernant les modèles ARMA stationnaires. Nous rappelons les critères essentiels de base pour l'estimation des paramètres et l'estimation de la fonction d'autocovariance. Nous passerons en revue les algorithmes les plus connus.

Le plan de ce chapitre est le suivant : Nous aurons besoin donc de commencer par la notion de stationnarité (la stationnarité au sens strict et la stationnarité faible) et la décomposition de Wold. Par la suite, nous définissons la classe des processus autoregressifs (AR), moyennes mobiles (MA) et mixte ARMA. Nous étudierons ensuite sous quelles conditions ces processus satisfont l'hypothèse de causalité et d'inversibilité. La cinquième section sera consacrée à l'étude de l'estimation des paramètres d'un modèle ARMA stationnaire. Nous passerons aprés à la dernière section où nous exposerons l'estimation de la fonction d'autocovariance.

1.1 Processus stationnaires

Nous commencerons par un rappel de définition d'un processus stationnaire au sens strict (ou stationnarité forte) pour ensuite étudier les propriétés de la stationnarité du second ordre (ou stationnarité faible). Partant de là nous étudierons des processus stationnaires particuliers que sont les bruits blancs.

1.1.1 Définition d'un processus stationnaire au sens strict

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus stochastique.

Définition (1.1.1) Le processus $\{y_t, t \in \mathbb{Z}\}$ est dit strictement ou fortement stationnaire si quelque soit le n-uplet du temps $t_1 < t_2 < ... < t_n$, tel que $t_i \in \mathbb{Z}$ et pour tout $h \in \mathbb{Z}$ avec $t_i + h \in \mathbb{Z}$, $\forall i, i = 1, ..., n$, la suite $(y_{t_1+h}, y_{t_2+h}, ..., y_{t_n+h})$ à la même distribution que la suite $(y_{t_1}, y_{t_2}, ..., y_{t_n})$, i.e, :

$$F_{y_{t_1+h},...,y_{t_n+h}}(x_1, ..., x_n) = F_{y_{t_1},...,y_{t_n}}(x_1, ..., x_n), \quad \forall x_1, ..., x_n.$$

Une autre façon équivalente de définir la stationnarité forte (ou stationnarité stricte) :

Définition (1.1.2) Un processus est dit stationnaire au sens strict si, pour toutes valeurs $j_1, j_2, ..., j_n$ la distribution conjointe de $(y_{t+j_1}, ..., y_{t+j_n})$ dépend uniquement des intervalles de temps $j_1, j_2, ..., j_n$ et est indépendante de l'instant t.

1.1.2 Stationnarité du second ordre ou stationnaire faible

Dans la pratique, on se limite généralement à requérir la stationnarité du second ordre (ou stationnarité faible) du processus étudié.

Définition (1.1.3) Un processus $\{y_t, t \in \mathbb{Z}\}$ est dit stationnaire du second ordre, ou stationnaire au sens faible, ou stationnaire d'ordre deux si les trois conditions suivantes sont satisfaites :
i) $\forall t \in \mathbb{Z},\ E(y_t^2) < \infty$.
ii) $\forall t \in \mathbb{Z},\ E(y_t) = m$, indépendant de t.
iii)$\forall (t, h) \in \mathbb{Z} \times \mathbb{Z},\ cov(y_t, y_{t+h}) = \gamma(h)$, indépendante de t.

La première condition $E(y_t^2) < \infty$ garantit tout simplement l'existence (ou la convergence) des moments d'ordre deux. La seconde condition $E(y_t) = m$, $\forall t \in \mathbb{Z}$ porte sur les moments d'ordre un et signifie tout simplement que les variables aléatoires y_t doivent avoir la même espérance quelque soit la date t. Autrement dit, l'espérance du processus y_t doit être indépendante du temps. Enfin, la troisième condition, $\gamma(h)$ indépendante de t, porte sur les moments d'ordre deux résumés par la fonction d'autocovariance. Cette condition implique que ces moments doivent être indépendants de la date considérée et ne doivent dépendre que de la différence des instants. Autrement dit la fonction d'autocovariance du processus y_t doit être indépendante du temps. Sous ces trois conditions, il existe alors une ou plusieurs représentations possibles de la série chronologique y_t. Cette hypothèse d'invariance de la fonction d'autocovariance dans le temps permet ainsi de se limiter à une certaine classe de processus.

En résumé, un processus est stationnaire du second ordre si l'ensemble de ses moments sont indépendants du temps. Par la suite le terme stationnaire fera référence à la stationnarité faible.

Quelques précisions doivent être apportées à cette définition. Tout dabord, il convient de noter que la fonction $\gamma(h)$, qui désigne la fonction d'autocovariance du processus $\{y_t, t \in \mathbb{Z}\}$, est symétrique

$$\gamma(h) = \gamma(-h)$$

En second lieu, on note que la troisième condition de cette définition implique en particulier que la variance du processus y_t soit constante et indépendante du temps :

$$var(y_t) = \gamma(0) = \sigma^2 \quad \text{(indépendant de } t\text{)}$$

1.1.3 Processus bruit blanc (*white noise*)

Parmi la classe des processus stationnaire, il existe des processus particuliers que sont les processus bruit blanc (*white noise*). Ces processus sont très souvent utilisés en analyse des séries chronologiques car ils constituent en quelque sorte les briques élémentaires de l'ensemble des processus stochastiques. En effet, nous verrons par la suite que tout processus stationnaire peut s'écrire comme une somme pondérée de bruits blancs (théorème de Wold).

Un bruit blanc est un processus stationnaire à accroissements indépendants. Commençons donc par définir ce qu'est un processus à accroissements indépendants.

Définition (1.1.4) Un processus $\{y_t, t \in \mathbb{Z}\}$ est un processus à accroissements indépendants si pour tout n-uplet du temps $t_1 < t_2 < ... < t_n$, $t_i \in \mathbb{Z}$, les variables aléatoires réelles $y_{t_2} - y_{t_1}, ..., y_{t_n} - y_{t_{n-1}}$ sont indépendantes.

Le processus $\{y_t, t \in \mathbb{Z}\}$ est un processus stationnaire à accroissements indépendants si la loi de probabilité des accroissements $y_{t+h} - y_t$ est indépendante de t, $\forall h \in \mathbb{Z}$. C'est donc une classe particulière de processus stationnaires.

Mais dans la pratique, on retient plus généralement la définition suivante du bruit blanc.

Définition (1.1.5) Un processus $\{y_t, t \in \mathbb{Z}\}$ est un bruit blanc s'il satisfait les deux conditions suivantes $\forall t \in \mathbb{Z}$:

$$E(y_t) = 0$$

$$\gamma(h) = E(y_t y_{t-h}) = \begin{cases} \sigma^2 & si \quad h = 0 \\ 0 & \forall h \neq 0 \end{cases}$$

La première condition signifie tout simplement que l'espérance du processus est indépendante du temps, et de plus qu'elle est nulle. La seconde condition implique bien entendu l'indépendance de la fonction d'autocovariance par rapport au temps (stationnarité). Mais elle implique en outre que les termes d'autocovariances (pour $h \neq 0$) sont tous nuls. Seule la variance est non nulle. Autrement dit, cela signifie que les bruits blancs sont des processus stationnaires particuliers sans mémoire. Le niveau de la série considéré aujourd'hui n'a aucune incidence sur son niveau de demain, tout comme le niveau d'hier n'a aucune incidence sur le niveau d'aujourd'hui.

En outre, on parle de bruit blanc gaussien lorsque la loi de probabilité du processus est elle même gaussienne.

1.2 Théorème de Wold

Le théorème de Wold (1938) est le théorème fondamental de l'analyse des séries chronologiques stationnaires. Nous commencerons par donner l'énoncé de ce théorème, puis nous verrons la décomposition de Wold dans la pratique.

1.2.1 Théorème de la décomposition de Wold

L'énoncé du théorème de Wold est le suivant

Théorème (1.2.1) (*Wold (1938)*)
Tout processus stationnaire du second ordre $\{y_t, t \in \mathbb{Z}\}$ peut être représenté sous la forme :

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + k_t \tag{1.2.1}$$

où les paramètres ψ_j satisfont $\psi_0 = 1$, $\psi_j \in \mathbb{R}$, $\forall j \in \mathbb{N}^*$, $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ et où ε_t est un bruit blanc de moyenne mulle et de variance σ_ε^2. Nous disons que la somme des chocs passés correspond à la composante linéaire stochastique de y_t. Le terme k_t désigne la composante linéaire déterministe choisie telle que $cov(k_t, \varepsilon_{t-j}) = 0$, $\forall j \in \mathbb{Z}$.

D'après le théorème de Wold, si nous omettons la composante déterministe k_t, tout processus stationnaire peut s'écrire comme une somme pondérée infinie convergente, en moyenne quadratique, de chocs passés, ces chocs étant représentés par un bruit blanc de variance finie. L'implication forte de ce théorème est que, si nous connaissons les coefficients (pondérations) ψ_j, $\forall j \in \mathbb{N}$, et si nous connaissons la variance σ_ε^2 du bruit blanc, nous pouvons proposer une représentation de n'importe quel processus stationnaire. Cette représentation est aussi qualifiée de représentation moyenne mobile infinie.

Reste à comprendre ce que peut être cette composante linéaire déterministe k_t. La condition $cov(k_t, \varepsilon_{t-j}) = 0$, implique que ce terme est, par définition, indépendant des chocs. Alors le cas le plus simple est celui d'un processus stationnaire $\{y_t, t \in \mathbb{Z}\}$ d'espérance non nulle, tel que $E(y_t) = m \neq 0$. Puisque le bruit blanc est par définition un processus centré, une somme pondérée de ces chocs est elle-même centrée. Par conséquent, la représentation de Wold du processus $\{y_t, t \in \mathbb{Z}\}$ suppose que nous ajoutons à cette somme pondérée de chocs passés, une composante déterministe qui n'est autre que l'espérance du processus : $k_t = m$. Nous avons donc

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + m$$

et que

$$E(y_t) = E\left[\sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}\right] + m = \sum_{j=0}^{\infty} \psi_j E(\varepsilon_{t-j}) + m = m$$

Dans le théorème de Wold, la condition $\sum_{j=0}^{\infty} \psi_j^2 < \infty$, dite condition de sommabilité des carrés (ou d'intégrabilité des carrés), est particulièrement importante. Elle assure l'existence des moments d'ordre deux du processus $\{y_t, t \in \mathbb{Z}\}$. Sous cette condition, nous disons alors que y_t converge en moyenne quadratique.

Pour comprendre pourquoi cette condition de sommabilité assure l'existence des moments théoriques d'ordre deux, il faut revenir à la définition de la fonction d'autocovariance du processus $\{y_t, t \in \mathbb{Z}\}$ et utiliser la représentation de Wold. Si nous supposons pour simplifier que $E(y_t) = 0$, nous obtenons alors :

$$\gamma(h) = E(y_{t+h} y_t) = E\left(\sum_{j=0}^{\infty} \psi_j \varepsilon_{t+h-j} \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}\right) \qquad \forall h \in \mathbb{Z}$$

Etant donné la définition d'un bruit blanc nous avons $E(\varepsilon_t \varepsilon_{t-j}) = 0$, $\forall j \neq 0$, $E(\varepsilon_t^2) = \sigma_\varepsilon^2$. Par conséquent, puisque l'espérance est un opérateur linéaire, la fonction $\gamma(h)$ peut se réécrire sous la forme :

$$\gamma(h) = \sum_{j=0}^{\infty} \psi_{j+h} \psi_j E(\varepsilon_{t-j}^2) \quad \forall h \in \mathbb{Z}$$

D'où finalement :

$$\gamma(h) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_{j+h} \psi_j \quad \forall h \in \mathbb{Z}$$

Nous pouvons montrer que la condition $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ implique $\sum_{j=0}^{\infty} \psi_{j+h} \psi_j < \infty$, $\forall h \neq 0$. Par conséquent, cette condition suffit à garantir la convergence, et donc l'existence, de l'ensemble des moments d'ordre deux du processus y_t.

1.2.2 Décomposition du Wold dans la pratique

Le théorème de Wold n'est jamais directement applicable, en tant que tel, dans la pratique puisque il suppose que nous exprimons un processus stationnaire comme une somme pondérée infinie de chocs passés. Or, dans la pratique informatique il n'est pas possible de manier des objets mathématiques de dimension infinie.

Pour autant, nous pouvons démontrer qu'il existe toujours pour tout processus stationnaire, un ordre fini, tel que ce processus puisse être approximé par une somme pondérée de chocs passés jusqu'à cet ordre (décomposition de Wold tronqué).

Concrètement, cela signifie que n'importe quelle série stationnaire peut être représentée de façon satisfaisante par une somme pondérée finie de chocs passés. Reste ensuite à estimer les coefficients de pondération (les ψ_j) et la variance du bruit blanc (σ_ε^2). Il est alors possible de proposer une prévision du processus étudié.

1.3 Modèle autorégressif moyenne mobile stationnaire

1.3.1 Définitions

1.3.1.1 Processus MA

La définition générale d'un processus MA est la suivante :

Définition (1.3.1) Le processus $\{y_t, t \in \mathbb{Z}\}$ satisfait une représentation MA d'ordre q, notée MA(q), s'il est solution de l'équation aux différences stochastique suivante :

$$y_t = m + \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

ou encore

$$y_t = m + \Theta(L)\varepsilon_t \tag{1.3.1}$$

où $E(y_t) = m$. Le polynôme $\Theta(L)$ étant défini par $\Theta(L) = -\sum_{j=0}^{q} \theta_j L^j$ où $\forall j < q,\ \theta_j \in \mathbb{R},\ \theta_0 = -1$ et $\theta_q \in \mathbb{R}^*$, avec ε_t est un bruit blanc de moyenne mulle et de variance σ_ε^2.

On reconnaît ici la forme moyenne mobile que nous avions identifié dans la décomposition de Wold, mais qui dans le cas présent est une forme moyenne mobile d'ordre q fini. Par opposition, la décomposition de Wold, définie par l'équation (1.2.1), est aussi qualifiée de MA(∞).

Les différentes conditions sur les paramètres θ_j signifient que dans un MA(q), tous les paramètres du polynômes $\Theta(L)$ peuvent être nuls à l'exception du paramètre correspondant au q^{eme} retard. En effet, si $\theta_j = 0$, $j \geq q$, alors le processus $\{y_t, t \in \mathbb{Z}\}$ satisfait une représentation MA($q-1$). De la même façon que pour la décomposition de Wold, le premier coefficient θ_0 associé au choc contemporain est, par convention, normalisé à l'unité.

Pour un processus MA(q), la constante m correspond à l'espérance du processus puisque les bruits blancs sont par définition centrés $E(\varepsilon_{t-j}) = 0$, $\forall j$:

$$E(y_t) = m + E(\varepsilon_t) - \sum_{j=1}^{q} \theta_j E(\varepsilon_{t-j}) = m$$

1.3.1.2 Processus AR

La définition générale d'un processus AR est la suivante :

Définition (1.3.2) Le processus stationnaire $\{y_t, t \in \mathbb{Z}\}$ satisfait une représentation AR d'ordre p, notée AR(p), s'il est solution de l'équation aux différences stochastique suivante :

$$y_t - \sum_{j=1}^{p} \phi_j y_{t-j} = c + \varepsilon_t$$

ou encore

$$\Phi(L) y_t = c + \varepsilon_t \tag{1.3.2}$$

avec $c \in \mathbb{R}$, $\Phi(L) = -\sum_{i=0}^{p} \phi_j L^j$ où $\forall j < p$, $\phi_j \in \mathbb{R}$, $\phi_0 = -1$ et $\phi_p \in \mathbb{R}^*$, où ε_t est un bruit blanc de moyenne mulle et de variance σ_ε^2.

Nous parlons ici de représentation autorégressive, dans le sens où la variable y_t est déterminée par les valeurs passées $y_{t-1}, y_{t-2}, ..., y_{t-p}$. Les différentes conditions sur les paramètres ϕ_j signifient que dans un AR(p), tous les paramètres du polynômes $\Phi(L)$ peuvent être nuls à l'exception du paramètre correspondant au p^{eme} retard. En effet, si $\phi_j = 0$, $j \geq p$, alors le processus $\{y_t, t \in \mathbb{Z}\}$ satisfait une représentation AR($p-1$). De la même façon que pour la décomposition de Wold, le premier coefficient ϕ_0 associé au choc contemporain est, par convention, normalisé à l'unité.

Contrairement au cas du modèle MA, la constante c de l'équation (1.3.2) ne correspond pas à l'espérance du processus $\{y_t, t \in \mathbb{Z}\}$. Cette espérance, dans le cas d'un processus stationnaire, est déterminée par la relation suivante si $\Phi(1) \neq 0$:

$$E(y_t) = \Phi^{-1}(1) c$$

En effet, reprenons l'écriture du modèle AR(p) donnée par l'équation (1.3.2) :

$$E(y_t) = c + \phi_1 E(y_{t-1}) + \phi_2 E(y_{t-2}) + \cdots + \phi_p E(y_{t-p}) + E(\varepsilon_t)$$

Par définition $E(\varepsilon_t) = 0$. Si nous supposons maintenant que l'espérance de $\{y_t, t \in \mathbb{Z}\}$ ne dépend pas du temps :

$$E(y_{t-j}) = m, \quad \forall j \in \mathbb{Z}$$

Donc, nous avons :

$$\begin{aligned} E(y_t) &= c + \phi_1 E(y_{t-1}) + \phi_2 E(y_{t-2}) + \cdots + \phi_p E(y_{t-p}) \\ &= c + \phi_1 m + \phi_2 m + \cdots + \phi_p m \end{aligned}$$

Ce qui peut se réécrire sous la forme :

$$E(y_t) = m = \frac{c}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}$$

Alors, nous retrouvons bien la formule générale puisque

$$\Phi^{-1}(1) = \left[- \sum_{i=0}^{p} \phi_j \right]^{-1} = \frac{1}{1 - \phi_1 - \phi_2 - \cdots - \phi_p} \quad avec \quad \phi_0 = -1$$

1.3.1.3 Processus mixte ARMA

Naturellement, les processus ARMA se définissent par l'adjonction d'une composante autorégressive AR et d'une composante moyenne mobile MA :

Définition (1.3.3) Le processus stationnaire $\{y_t, t \in \mathbb{Z}\}$ satisfait une représentation ARMA, d'ordre p et q, notée ARMA(p, q), s'il est solution de l'équation aux différences stochastique suivante :

$$y_t - \sum_{i=1}^{p} \phi_i y_{t-i} = c + \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

ou encore

$$\Phi(L) y_t = c + \Theta(L) \varepsilon_t \tag{1.3.3}$$

avec $c \in \mathbb{R}$, $\Theta(L) = -\sum_{j=0}^{q} \theta_j L^j$, $\Phi(L) = -\sum_{i=0}^{p} \phi_j L^j$ où $\forall j < q$, $\theta_j \in \mathbb{R}$, $\forall j < p$, $\phi_j \in \mathbb{R}$, $\phi_0 = \theta_0 = -1$ et $(\phi_p, \theta_q) \in \mathbb{R}^* \times \mathbb{R}^*$, avec ε_t est un bruit blanc de moyenne mulle et de variance σ_ε^2.

Les différentes conditions sur les paramètres ϕ_j et θ_j signifient que dans un ARMA(p, q) tous les paramètres du polynômes $\Phi(L)$ peuvent être nuls à l'exception du paramètre correspondant au p^{eme} retard et que tous les paramètres du polynômes $\Theta(L)$ peuvent être nuls à l'exception du paramètre correspondant au q^{eme} retard. En effet, si $\phi_p = 0$, alors le processus $\{y_t, t \in \mathbb{Z}\}$ satisfait une représentation ARMA$(p-1, q)$. De la même façon, si $\theta_q = 0$, alors le processus $\{y_t, t \in \mathbb{Z}\}$ satisfait une représentation ARMA$(p, q-1)$.

Ainsi, nous constatons que les processus AR et MA ne sont que des cas particuliers des processus ARMA. Un AR(p) correspond à un ARMA$(p, 0)$, de la même façon un MA(q) correspond à un ARMA$(0, q)$.

Tout comme pour les AR, la constante c ne correspond pas dans ce cas à l'espérance du processus $\{y_t, t \in \mathbb{Z}\}$. Cette dernière est donnée par la même relation que celle obtenue dans le cas d'un processus AR :

$$E(y_t) = \Phi^{-1}(1)c$$

1.3.2 Stationnarité, inversibilité et causalité des processus ARMA

Considérons à présent des processus ARMA. La question est alors de savoir sous quelles conditions sur les paramètres du polynôme $\Phi(L)$ ces processus sont ils stationnaires. Nous allons en outre introduire la notion d'inversibilité qui consiste à déterminer s'il existe une représentation AR équivalente pour un processus ARMA. Enfin nous introduisons également la notion de causalité qui consiste à exprimer le processus y_t en fonction des ε_t, présent et passés.

Définition (1.3.4) Un processus $\{y_t, t \in \mathbb{Z}\}$ est inversible si et seulement s'il existe une fonction mesurable h telle que :

$$\varepsilon_t = h(y_t, y_{t-1}, ...), \qquad \text{pour tout } t$$

Définition (1.3.5) Un processus $\{y_t, t \in \mathbb{Z}\}$ est causal si et seulement s'il existe une fonction mesurable g telle que :

$$y_t = g(\varepsilon_t, \varepsilon_{t-1}, ...), \qquad \text{pour tout } t$$

1.3.2.1 Conditions de stationnarité et d'inversibilité

Commençons par énoncer un premier théorème concernant les processus AR :

Théorème (1.3.1)

Un processus $\{y_t, t \in \mathbb{Z}\}$ satisfaisant une représentation AR(p) est toujours inversible. Il est stationnaire lorsque toutes les racines du polynôme $\Phi(L)$, notées $\lambda_j \in C, j \leqslant p$, sont de module strictement supérieur à l'unité.

$$\Phi(\lambda_j) = \sum_{i=1}^{p} \phi_i \lambda_j^i = 0 \Longleftrightarrow \prod_{j=1}^{p} \left(1 - \frac{1}{\lambda_j} L\right) = 0, \quad |\lambda_j| > 1, \ \forall j.$$

Ce théorème indique tout d'abord que tout processus AR peut être inversé et représenté sous la forme d'un processus MA. Le second point, plus fondamental, porte sur la stationnarité du processus. Ce théorème indique que lorsque les racines du polynôme autorégressif sont toutes supérieures à l'unité en module, le processus satisfait les trois conditions de la stationnarité du second ordre.

Un second théorème peut être énoncé pour les processus MA :

Théorème (1.3.2)

Un processus $\{y_t, t \in \mathbb{Z}\}$ satisfaisant une représentation MA(q) est toujours stationnaire. Il est inversible lorsque toutes les racines du polynôme $\Theta(L)$, notées $\beta_j \in C, \forall j \leqslant q$, sont de module strictement supérieur à l'unité.

$$\Theta(\beta_j) = \sum_{i=0}^{q} \theta_i \beta_j^j = 0 \Longleftrightarrow \prod_{j=1}^{q} \left(1 - \frac{1}{\beta_j} L\right) = 0, \quad |\beta_j| > 1, \ \forall j.$$

Ce théorème nous indique tout d'abord que tout processus MA, qui n'est autre qu'une somme pondérée de bruits blancs, est toujours stationnaire. Ce résultat se comprend facilement compte tenu des propriétés des bruits blancs. La seconde partie nous indique qu'un processus y_t défini comme un MA peut être exprimé sous la forme d'un AR si le polynôme associé $\Theta(L)$ est inversible, c'est-à-dire si ses racines sont toutes supérieures à 1 en module.

Bien entendu, il nous reste maintenant à établir les mêmes résultats en ce qui concerne les processus ARMA :

Corollaire (1.3.1)

Les conditions de stationnarité d'un processus ARMA sont déterminées par les racines du polynôme associé à sa partie AR. Les conditions d'inversibilité sont déterminées par les racines du polynôme associé à sa partie MA.

Théorème (1.3.3)

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus ARMA(p, q) défini par $\Phi(L)y_t = \Theta(L)\varepsilon_t$ tel que les pôlynomes $\Phi(.)$ et $\Theta(.)$ d'ordres respectifs p et q n'ont pas de racines communes.

Alors, y_t est inversible si et seulement si les racines de Θ sont de module strictement supérieure à l'unité, i.e.

$$\Theta(z) \neq 0, \quad \forall z \in \mathbb{Z}, \ |z| \leqslant 1.$$

Les coefficients $(\pi_j)_j$ de la représentation

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_j y_{t-j}, \quad t \in \mathbb{Z}.$$

sont alors déterminés par :

$$\Pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\Phi(z)}{\Theta(z)}, \quad |z| \leqslant 1.$$

1.3.2.2 Conditions de causalité des processus ARMA

Théorème (1.3.4)
Soit $\{y_t, t \in \mathbb{Z}\}$ un processus ARMA(p,q) défini par $\Phi(L)y_t = \Theta(L)\varepsilon_t$ tel que les pôlynomes $\Phi(.)$ et $\Theta(.)$ d'ordres respectifs p et q n'ont pas de racines communes.

Alors, y_t est causal si et seulement si les racines de Φ sont de module strictement supérieure à l'unité, i.e.

$$\Phi(z) \neq 0, \quad \forall z \in \mathbb{Z}, \ |z| \leqslant 1.$$

Les coefficients $(\psi_j)_j$ de la représentation

$$\varepsilon_t = \sum_{j=0}^{\infty} \psi_j y_{t-j}, \quad t \in \mathbb{Z}.$$

sont alors déterminés par

$$\Psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\Theta(z)}{\Phi(z)}, \quad |z| \leqslant 1.$$

1.4 Fonction d'autocovariance (ACVF), d'autocorrélation (ACRF) et la fonction d'autocorrélation partielle (PACF)

Définition (1.4.1) La fonction d'autocovariance d'un processus $\{y_t, t \in \mathbb{Z}\}$ centré et stationnaire du second ordre est définie par :

$$\gamma(t,s) = E(y_t y_s) = \gamma(t-s)$$

Nous appelons matrice des autocovariances d'ordre k toute matrice $(k \times k)$ définie par :

$$\Gamma_k = (\gamma(i-j))_{i,j=1,\dots,k} = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(k-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(k-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(k-1) & \gamma(k-2) & \cdots & \gamma(0) \end{pmatrix}.$$

Définition (1.4.2) Nous supposons que $\gamma(0) \neq 0$, c'est-à-dire qu'aucune des variables y_t n'est nulle.

Nous appelons fonction d'autocorrelation, l'application

$$\rho(j) = \frac{\gamma(j)}{\gamma(0)}, \text{ pour } j \in \mathbb{Z}.$$

Nous appelons matrice d'autocorrélation d'ordre k toute matrice $(k \times k)$ définie par :

$$\Sigma_k = (\rho(i-j))_{i,j=1,\dots,k} = \begin{pmatrix} 1 & \rho(1) & \cdots & \rho(k-1) \\ \rho(1) & 1 & \cdots & \rho(k-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(k-1) & \rho(k-2) & \cdots & 1 \end{pmatrix}.$$

Définition (1.4.3) Nous appelons fonction d'autocorrélation partielle d'un processus $\{y_t, t \in \mathbb{Z}\}$ centré et stationnaire du second ordre, la fonction qui à tout $k > 0$, fait correspondre

$$\phi_{k,k} = \frac{\det \Sigma_k^*}{\det \Sigma_k}$$

lorsque Σ_k est réguliere, Σ_k désigne la matrice d'autocorrélation d'ordre k et Σ_k^* est la matrice obtenue en remplaçant la dernière colonne de Σ_k par

$$(\rho(1), \rho(2), ..., \rho(k))^t$$

donc

$$\Sigma_k^* = \begin{pmatrix} \rho(0) & \cdots & \rho(k-2) & \rho(1) \\ \rho(-1) & \cdots & \rho(k-3) & \rho(2) \\ \vdots & & \vdots & \vdots \\ \rho(-k+1) & \cdots & \rho(-1) & \rho(k) \end{pmatrix}.$$

Le coefficient d'autocorrélation partielle, dans le cas stationnaire, peut se définir comme le produit scalaire de deux vecteurs sur le produit de leur norme, donc un cosinus, c'est-à-dire un coefficient de corrélation entre deux variables aléatoires, ou encore le rapport de deux déterminants. (Pour plus de détail voir Hamdi (1982))

Avant de passer à l'étude des méthodes de sélection des ordres d'un modèle ARMA stationnaire, il nous semble utile d'étudier quelques méthodes d'estimation des paramètres du modèle. Nous présentons dans un premier temps, les méthodes des moments où nous aurons besoin de rappeler les équations de Yule-Walker, les équations de Yule-Walker prolongées et le système de Box-Jenkins. Nous détaillerons ensuite quelques algorithmes pour l'estimation récursive des paramètres dans le cas d'un modèle autorégressif pur (AR), moyenne mobile pur (MA) et ARMA mixte. Nous étudierons ensuite la méthode du maximum de vraisemblance. Puis les méthodes des moindres carrés et nous terminons par quelques algorithmes pour estimer les autocovariances en utilisant des méthodes dont la complexité est indépendante de l'observation et ne varie qu'en fonction des ordres du modèle mais en se basant sur la connaissance des paramètres.

1.5 Estimation des paramètres d'un modèle autorégressif moyenne mobile (ARMA)

Les meilleures précisions pour l'estimation des paramètres d'un processus ARMA (p, q) stationnaire s'obtiennent par la méthode du maximum de vraisemblance. Il n'est pas toujours possible de mettre cette méthode en oeuvre car elle suppose, a priori, le processus gaussien. La méthode d'estimation la plus simple est celle de Yule-Walker, elle reduit considérablement le temps de calcul grâce à l'algorithme de Levinson-Durbin. Elle est basée sur la résolution du système linéaire de p équations (dites de Yule-Walker).

La relative simplicité des estimateurs de Yule-Walker dans le cas stationnaire a conduit de nombreux auteurs à étudier leurs propriétés et à développer certaines variations et extensions. Les estimateurs de Yule-Walker étendus $\widehat{\phi}_{k,j}^{(i)}$ avec k, j et i entiers, sont définis par un système d'équations où k représente le nombre d'équations, i un décalage et j varie entre 1 et k. En fonction des différentes valeurs que prennent i et k, ces équations englobent une large classe d'estimateurs. En particulier lorsque $i = 0$ et k est quelconque, on retrouve les estimateurs de Yule-Walker ; dans le cas où $k = p$ les $\widehat{\phi}_{k,j}^{(i)}$ sont alors des estimateurs des paramètres AR du processus.

Les propriétés des estimateurs de Yule-Walker étendus sont assez bien connues dans le cas des processus ARMA stationnaire (Choi (1991a)). Par contre dans le cas instable-stable, les équations de Yule-Walker ne peuvent plus être définies. On ne considère alors que les estimateurs de type moindres carrés qui correspondent aux estimateurs de Yule-Walker dans le cas stationnaire.

1.5.1 Estimation des paramètres par la méthode des moments

1.5.1.1 Estimation des paramètres d'un modèle AR(p)

A. Equations de Yule-Walker (EYW)

Considérons le processus $\{y_t, t \in \mathbb{Z}\}$ autoregressif causal suivant :

$$y_t - \sum_{i=1}^{p} \phi_i \; y_{t-i} = \varepsilon_t. \tag{1.5.1}$$

En multipliant (1.5.1) par y_{t-j} $\;j = 0, ..., p$, et en prenant l'espérance nous obtenons :

$$\gamma(j) - \sum_{i=1}^{p} \phi_i \; \gamma(i-j) = \sigma^2 \delta_{j,0} \;, \qquad j = 0, ..., p \tag{1.5.2}$$

où δ est le symbole de Kronecker. Ou encore sous forme matricielle

$$\begin{cases} \Gamma_p \; \Phi = \gamma_p \\ \sigma^2 = \gamma(0) - \Phi^t \; \gamma_p \end{cases} \tag{1.5.3}$$

où $\Gamma_p = (\gamma(i-j))_{i,j=1,...,p}$ est la matrice $p \times p$ des autocovariances, $\gamma_p = (\gamma(1), ..., \gamma(p))'$ et $\Phi = (\phi_1, ..., \phi_p)'$. Ce système est appelé les équations de Yule-Walker.

Il est bien connu que les paramètres d'un modèle AR(p) satisfont les équations de Yule-Walker prolongées

$$\gamma(j) - \sum_{i=1}^{p} \phi_i \; \gamma(i-j) = 0 \;, \qquad j = p+1, p+2, ...$$

Qui est équivalent à

$$\rho(j) - \sum_{i=1}^{p} \phi_i \; \rho(i-j) = 0 \;, \qquad j = p+1, p+2, ...$$

B. Estimateurs de Yule-Walker

Sur la base d'observations $y_1, ..., y_N$ supposées être une N-réalisation du processus (1.5.1), les estimateurs des pramètres $\phi_1, ..., \phi_p$ et σ^2 peuvent être obtenus en remplaçant les autocovariances théoriques $\gamma(j)$, $j = 0, ..., p$, par leurs estimateurs empiriques

$$\widehat{\gamma}(j) = \widehat{\gamma}(-j) = \frac{1}{N} \sum_{t=1}^{N-j} (\; y_t - \overline{y}) \, (\; y_{t+j} - \overline{y}) \;\;, \qquad j = 0, 1, ..., N-1$$

où $\overline{y} = \frac{1}{N} \sum_{t=1}^{N} \; y_t$. Nous aurons un système qui permet d'obtenir les estimateurs de Yule-Walker $\widehat{\phi}$ et $\widehat{\sigma}^2$.

Ces derniers sont aussi appelés les estimateurs des moments, car ils sont obtenus en remplaçant les moments théoriques par leurs estimations empiriques.

Remarque : Nous appelons l'estimateur $\widehat{\gamma}(j)$ l'ACVF de l'echantillon. Jenkins et watts (1968) ont expliqué pourquoi l'utilisation de cet estimateur de l'ACVF est meilleure qu'un autre.

Nous définissons l'ACRF de l'échantillon comme suit

$$\widehat{\rho}(j) = \frac{\widehat{\gamma}(j)}{\widehat{\gamma}(0)}$$

Généralement, les estimateurs des moments sont moins efficaces que les estimateurs obtenus par des méthodes alternatives telles que la méthode des moindres carrés ou la méthode du maximum de vraisemblance. Cependant, pour un modèle AR(p), il est montré (voir, par exemple, Brokwell et Davis (1988)) que les estimateurs de Yule-Walker possèdent des propriétés asymptotiques équivalentes à celle de l'estimateur du maximum de vraisemblance.

Théorème (1.5.1) (*Convergence des estimateurs de Yule Walker*)
Si $\{y_t, t \in \mathbb{Z}\}$ est un processus autoregressif causal, AR(p) et $\widehat{\Phi}$ est l'estimateur de Yule-Walker du vecteur des paramètres Φ, alors, nous avons :

$$\sqrt{N}(\widehat{\Phi} - \Phi) \rightarrow N(0, \sigma^2 \Gamma_p^{-1}),$$

où Γ_p est la matrice des covariances $(\gamma(i-j))_{i,j=1,\ldots,p}$.

Démonstration : Voir par exemple, Brokwell et Davis (1988), p. 263.

C. Quelques algorithmes pour estimer récursivement les paramètres d'un modèle AR(p)
Puisque les équations de Yule-Walker ont une forme linéaire, nous pouvons utiliser la méthode d'élimination de Gauss pour les résoudre. Pour identifier un processus ARMA, il est nécessaire de résoudre les EYW pour plusieurs paires d'ordres. Par conséquent, la méthode d'élimination de Gauss n'est pas pratique.

Nous définissons $\phi_{k,1}^{(i)}, \ldots, \phi_{k,k}^{(i)}$ comme solutions des EYW suivantes :

$$\rho(j) - \sum_{h=1}^{k} \phi_{k,h}^{(i)} \, \rho(j-h) = 0 \ , \qquad j = i+1, i+2, \ldots, i+k$$

avec k, j et i des entiers, où k,i sont respectivement l'ordre de la partie AR et MA.

Nous définissons aussi la matrice et les vecteurs suivants :

$$\Gamma(k,i) = \begin{pmatrix} \gamma(i) & \gamma(i-1) & \cdots & \gamma(i-k+1) \\ \gamma(i+1) & \gamma(i) & \cdots & \gamma(i-k+2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(i+k-1) & \gamma(i+k-2) & \cdots & \gamma(i) \end{pmatrix};$$

$$\Sigma(k,i) = \frac{1}{\gamma(0)}\Gamma(k,i);$$

$$\gamma(k,i) = \begin{pmatrix} \gamma(i+1) \\ \gamma(i+2) \\ \vdots \\ \gamma(i+k) \end{pmatrix}, \qquad \Phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix}, \qquad \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_q \end{pmatrix},$$

$$\Phi_* = \begin{pmatrix} \phi_0 \\ \phi_1 \\ \vdots \\ \phi_p \end{pmatrix}, \qquad \theta_* = \begin{pmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_q \end{pmatrix}, \qquad \Phi_{k,i} = \begin{pmatrix} \phi_{k,1}^{(i)} \\ \phi_{k,2}^{(i)} \\ \vdots \\ \phi_{k,k}^{(i)} \end{pmatrix}.$$

Avec les notations ci-dessus, les équations de Yule-Walker peuvent être représentées par

$$\Gamma(k,i)\Phi_{k,i} = \gamma(k,i)$$

Nous considérons le modèle AR(p) (de la représentation (1.2.1)) dans ce cas-ci $\{\phi_{k,1}, ..., \phi_{k,k} \setminus k = 1, 2, ...\}$ peut être calculée par l'algorithme de Levinson(1947)-Durbin(1960).

Algorithme de Levinson-Durbin

$1 \stackrel{ere}{=}$ *étape* : Initialisation

Pour $k = 0$, on commence par calculer $\phi_{1,1}$

$$\phi_{1,1} = \rho(1) = \frac{\gamma(1)}{\gamma(0)}.$$

On pose

$$\lambda(1) = 1 - \phi_{1,1}^2.$$

$2 \stackrel{eme}{=}$ *étape* : Construction des $\phi_{k,j}$

Pour $k = 1, 2, ...,$

$$\theta(k) = \rho(k+1) - \phi_{k,1}\rho(k) - \cdots - \phi_{k,k}\ \rho(1)\ ,$$

$$\phi_{k+1,k+1} = \frac{\theta(k)}{\lambda(k)},$$

$$\lambda(k+1) = \lambda(k)(1-\phi^2_{k+1,k+1}).$$

Pour $j = 1, 2, ..., k,$

$$\phi_{k+1,j} = \phi_{k,j} - \phi_{k+1,k+1}\phi_{k,k+1-j} \ . \qquad \square$$

La variance du bruit blanc satisfait

$$\sigma^2 = \gamma(0) - \sum_{k=1}^{p} \phi_k \ \gamma(k).$$

Si nous définissons

$$\sigma_k^2 = \gamma(0) - \sum_{h=1}^{k} \phi_{k,h}, \ \gamma(h) \ , \qquad k = 1, 2, ...,$$

Alors

$$\sigma_k^2 = \gamma(0)\lambda(k) \ , \qquad k = 1, 2, ...$$

Akaike (1973b) a mentionné que l'algorithme de Levinson-Durbin est probablement une des contributions les plus significatives dans le domaine de l'analyse numérique des séries chronologiques. Sa stabilité numérique a été remise en cause par Pagano (1972) et Box et Jenkins (1976, p. 84). Cybenko (1980) a montré que

$$\frac{1}{\min\left\{\prod_{j=1}^{k-1}(1-\phi^2_{j,\ j}), \prod_{j=1}^{k-1}(1-\phi_{j,\ j})\right\}} \le \left\|\Sigma^{-1}(k,0)\right\| \le \prod_{j=1}^{k-1} \frac{1+|\phi_{j,\ j}|}{1-|\phi_{j,\ j}|},$$

où $\|.\|$ est la norme matricielle.

De ces inégalités nous remarquons que l'instabilité numérique n'est pas due à l'algorithme mais à la nature de la matrice d'autocovariance $\Gamma(k, i)$. Elles prouvent que nous ne devrions pas compter sur les solutions des EYW si $\Phi(z) = 0$ admet une racine de module près de 1.

Algorithme de Trench-Zohar

Si nous considèrons le processus mixte ARMA où l'ordre du processus MA est fixé, les EYW constituent un système de Toeplitz. Ainsi, nous pouvons utiliser l'algorithme récursif de Trench-Zohar pour estimer les paramètres du modèle AR.

Nous considérons le cas où $i = q$.

$1 \overset{ere}{=}$ *étape* : Initialisation

$$\begin{aligned}\theta(0,q) &= \rho(q+1),\\ \eta(0,q) &= \rho(q-1),\\ \lambda(0,q) &= \rho(q)\end{aligned}$$

$$\phi_{1,1}^{(q)} = \frac{\theta(0,q)}{\lambda(0,q)},$$
$$\pi_{1,1}^{(q)} = \frac{\eta(0,q)}{\lambda(0,q)},$$
$$\lambda(1,q) = \lambda(0,q)\left\{1 - \phi_{1,1}^{(q)}\pi_{1,1}^{(q)}\right\}.$$

$2 \stackrel{eme}{=}$ *étape* :

Pour $k = 1, 2, ...,$

$$\theta(k,q) = \rho(q+k+1) - \phi_{k,1}^{(q)}\rho(q+k) - \cdots - \phi_{k,k}^{(q)}\rho(q+1),$$
$$\eta(k,q) = \rho(q-k-1) - \pi_{k,1}^{(q)}\rho(q-k) - \cdots - \pi_{k,k}^{(q)}\rho(q-1),$$
$$\phi_{k+1,k+1}^{(q)} = \frac{\theta(k,q)}{\lambda(k,q)},$$
$$\pi_{k+1,k+1}^{(q)} = \frac{\eta(k,q)}{\lambda(k,q)},$$
$$\lambda(k+1,q) = \lambda(k,q)\left\{1 - \phi_{k+1,k+1}^{(q)}\pi_{k+1,k+1}^{(q)}\right\}.$$

Pour $j = 1, ..., k,$

$$\phi_{k+1,\ j}^{(q)} = \phi_{k,\ j}^{(q)} - \phi_{k+1,k+1}^{(q)}\pi_{k,k+1-\ j}^{(q)},$$
$$\pi_{k+1,\ j}^{(q)} = \pi_{k,\ j}^{(q)} - \pi_{k+1,k+1}^{(q)}\phi_{k,k+1-\ j}^{(q)}. \qquad \square$$

Nous remarquons que l'algorithme de Trench-Zohar est une généralisation de Levinson-Durbin.

Choi (1990a, 1991b) a présenté les deux algorithmes suivants, qui sont étroitement liés à l'algorithme de Trench-Zohar et qui sont utiles lorsque ni l'ordre de AR ni l'ordre de MA ne sont connus.

Algorithme de Levinson-Durbin généralisé (Choi (1990))

$1 \stackrel{ere}{=}$ *étape* :

Pour $i = 0$, nous utilisons l'algorithme de Levinson-Durbin pour calculer

$$\left\{\phi_{k,\ j}^{(0)} \setminus k = 1, 2, ...,\ j = 1, ..., k\right\}.$$

$2 \stackrel{eme}{=}$ *étape* :

Pour $i = 1, 2, ...,$

Pour $k = 1, 2, ...,$ nous posons

$$\phi_{k,\ 0}^{(i)} = -1$$

Pour $j = 1, ..., k,$

$$\phi_{k,\ j}^{(i)} = \phi_{k+1,\ j}^{(i-1)} - \frac{\phi_{k+1,\ k+1}^{(i-1)}}{\phi_{k,\ k}^{(i-1)}}\phi_{k,\ j-1}^{(i-1)}. \qquad \square$$

Algorithme de Levinson-Durbin généralisé (Choi (1991))

$1 \stackrel{ere}{=}$ *étape*

Pour $i = 0$, utilisons l'algorithme de Levinson-Durbin pour calculer

$$\left\{\phi^{(0)}_{k,\ j} \setminus k = 1, 2, ...,\ j = 1, ..., k\right\}.$$

$2 \stackrel{eme}{=}$ *étape* :

Pour $i = 1, 2, ...,$

Pour $k = 0$, nous posons

$$\begin{aligned} \lambda(0, i) &= \rho_i, \\ \phi^{(i)}_{1,1} &= \tfrac{\rho_{i+1}}{\rho_i}, \end{aligned}$$

Pour $k = 1, 2, ...,$

$$\begin{aligned} \theta(k, i) &= \rho_{i+k+1} - \phi^{(i)}_{k,1}\rho_{i+k} - \cdots - \phi^{(i)}_{k,k}\rho_{i+1}, \\ \lambda(k, i) &= \lambda(k-1, i)\left\{1 - \frac{\phi^{(i)}_{k,\ k}}{\phi^{(i-1)}_{k,\ k}}\right\}, \\ \phi^{(i)}_{k+1,\ k+1} &= \tfrac{\theta(k,i)}{\lambda(k,i)}. \end{aligned}$$

Pour $j = 1, ..., k,$

$$\phi^{(i)}_{k+1,\ j} = \phi^{(i)}_{k,\ j} + \frac{\phi^{(i-1)}_{k,\ j-1}}{\phi^{(i-1)}_{k,\ k}}\phi^{(i)}_{k+1,\ k+1}\ . \qquad \square$$

Il est clair que l'algorithme de Trench-Zohar est une généralisation de Levinson-Durbin aux processus mixtes ARMA. Du point de vue informatique il est plus efficace que n'importe quel autre algorithme pour résoudre les EYW (pour estimer les paramètres AR), en particulier lorsque les vrais ordres sont inconnus. De plus, il n'existe aucun algorithme qui calcule les séquences artificielles $\left\{\phi^{(i)}_{k,\ j}\right\}$.

1.5.1.2 Estimation des paramètres d'un modèle MA(q)

Nous savons (voir, Box et Jenkins (1976)) que les paramètres et l'ACVF satisfont les équations suivantes :

$$\sum_{i=0}^{p} \phi_i\ \gamma(j-i) = \sigma^2 \sum_{k=0}^{q-j} \psi_k \theta_{j+k}\ , \qquad j = 0, ..., q. \tag{1.5.4}$$

Ici ψ_j est définie par

$$\sum_{k=0}^{j} \phi_k \psi_{j-k} = \begin{cases} \theta_j\ , & j = 0, 1, ..., q, \\ 0\ , & j = q+1, q+2, ..., \end{cases}$$

où on assume que $\phi_i = 0$ pour $i = p+1, p+2, ...$ Si nous posons $\psi(z) = \sum_{l=0}^{\infty} \psi_l\, z^l$, alors $\phi(z)\ \psi(z) = \theta(z)$.

Théoriquement, lorsque l'ACVF est donnée nous pouvons obtenir les paramètres du processus MA et la variance du bruit blanc en résolvant simultanément les équations (1.5.4). Puisque celles-ci sont non-linéaires, nous avons besoin de quelques méthodes itératives pour les résoudre.

Choi (1986) a développé un algorithme itératif afin d'obtenir la solution unique des paramètres MA et la variance du bruit blanc, en utilisant une propriété des matrices triangulaires de Toeplitz. Cet algorithme est simple et économique pour le temps de calcul et pour l'espace mémoire aussi. Mais, sa convergence n'a pas encore été démontrée.

Choi (1987) a présenté la solution de Newton-Raphson des équations non-linéaires. Dans le cas où le processus ARMA (p, q) est stationnaire, il peut être représenté par un modèle MA(∞)

$$y_t = \psi(B)\varepsilon_t\ ,$$

où $\psi(B) = \phi^{-1}(B)\theta(B) = \sum_{j=0}^{\infty} \psi_j B^j$.

Algorithme pour estimer récursivement les paramètres d'un modèle MA(q)

Nous supposons que $\Gamma_{q+1,p+1,0}$ et Ψ soient des matrices de dimension $(q+1) \times (p+1)$ dont l'èlèment (i, j) est donné par

$$(\Gamma_{q+1,p+1,0}\ )_{i,\ j} = \gamma(i-j)$$

$$(\Psi)_{i,\ j} = \begin{cases} 0, & i < j \\ \psi_{i-j}, & si\ non \end{cases}$$

Soit encore Ψ_q une matrice de dimension $(q+1) \times (p+1)$, dont le $(i,\ j)$ èlèment

$$(\Psi_q)_{i,\ j} = \begin{cases} 0, & i > j \\ \psi_{j-i}, & si\ non. \end{cases}$$

L'équation (1.2.4) donne

$$\Gamma_{q+1,p+1,0}\ \Phi_* = \sigma^2 \Psi_q \Psi\ \Phi_*\ .$$

Soient

$c_0 = \sigma,\ c_1 = \sigma\psi_1, ..., c_q = \sigma\psi_q,\ C_q = \sigma\Psi_q,\ C = \sigma\Psi,\ et\ c = (c_0, c_1, ..., c_q)^t$.

Alors, on a

$$C_q C\ \Phi_* = \Gamma_{q+1,p+1,0}\ \Phi_*\ , \tag{1.5.5}$$

que nous pouvons résoudre par l'algorithme suivant :

Algorithme de Newton-Raphson

Pour $j = 0, ..., q$, nous définissons les deux matrices T_j *et* S_j de dimenssion $(q+1) \times (p+1)$ par

$$(T_j)_{r,s} = \begin{cases} c_{j+r-s} & 1 \leq s \leq j+r, 1 \leq r \leq q-j+1 \\ 0 & si\ non, \end{cases}$$
$$(S_j)_{r,s} = \begin{cases} c_{j-r+s} & 1 \leq r \leq q+1, \max\{1, r-j\} \leq s \leq q-j+1 \\ 0 & si\ non \end{cases}$$

Posons
$$U_j = T_j + S_j, \qquad j = 0, ..., q,$$
$$W = (U_0\Phi_*, ..., U_q\Phi_*).$$

Alors la solution de Newton-Raphson de (1.5.5) est obtenue par l'équation récursive

$$c^{(n+1)} = \frac{1}{2}c^{(n)} + (W^{(n)})^{-1}\Gamma_{q+1,p+1,0}\ \phi_*,$$

où l'indice (n) signifie la $n \overset{eme}{=}$ itération. $\square$

Comme toute méthode d'estimation récursive, il y a un grand intérêt à utiliser des valeurs initiales convenables pour ne pas heurter la sensibilité de la méthode. Choi (1987) a prouvé que la convergence, en moyenne quadratique, de l'algorithme de Newton-Raphson, n'est pas assurée et peut dépendre fortement des valeurs initiales.

Pour obtenir des valeurs approximatives des paramètres MA, nous pouvons utiliser le modèle MA(s) avec s assez grand,

$$y_t^{(s)} = \sum_{j=0}^{s} \psi_j^{(s)}\ \varepsilon_{t-j},$$

où $\psi_0^{(s)}, ..., \psi_s^{(s)}$ doivent être déterminés de telle sorte que $cov(y_t^{(s)}, y_{t-j}^{(s)}) = \sigma(j)$ pour $j = 0, ..., s$. Il est connu (voir, par exemple, T. W. Anderson (1971)) que $y_t^{(s)}$ converge vers y_t dans L^2 lorsque $s \to \infty$. Ainsi, nous pouvons choisir un entier s de telle sorte que $\psi_0^{(s)}, ..., \psi_s^{(s)}$ soient proches de $\psi_0, ..., \psi_s$ comme nous souhaitons. Si nous appliquons l'algorithme de Newton-Raphson au modèle MA(s) avec des valeurs initiales

$$c_0^{(0)} = \left\{\gamma(0) + 2\sum_{j=1}^{s}\gamma(j)\right\}^{1/2},$$
$$c_j^{(0)} = \phi_1 c_{j-1}^{(0)} + \cdots + \phi_j c_0^{(0)}, \qquad j = 1, ..., s,$$

où $\phi_{p+1} = \phi_{p+2} = \cdots = 0$, alors les estimateurs resultants convergent aux vraies valeurs de paramètres.

Choi (1987) a proposé un autre ensemble de valeurs initiales

$$
\begin{aligned}
c_0^{(0)} &= \{\gamma(0)\}^{1/2}, \\
c_j^{(0)} &= \phi_1 c_{j-1}^{(0)} + \cdots + \phi_j c_0^{(0)}, \qquad j = 1, ..., q,
\end{aligned}
$$

où $\phi_{p+1} = \phi_{p+2} = \cdots = 0$. Ces valeurs sont choisies de telle sorte que toutes les racines de l'équatrion $\theta^{(0)}(z) = 0$ sont en dehors du cercle d'unité.

1.5.1.3 Estimation des paramètres d'un modèle ARMA(p, q)

Considérons un processus stochastique $\{y_t, t \in \mathbb{Z}\}$ qui admet une représentation autoregressive moyenne mobile d'ordre p et q suivante :

$$
y_t - \sum_{i=1}^{p} \phi_i y_{t-i} = \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \tag{1.5.6}
$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un processus bruit blanc de variance σ^2.

A. Système de Box-Jenkins

Box et Jenkins (1970) ont proposé une généralisation des équations de Yule-Walker.

Supposons maintenant que la série $y_1, ..., y_N$ est une réalisation du modèle (1.5.6). Soit alors, $y_t = \sum_{k=0}^{\infty} \psi_k \ \varepsilon_{t-k}$, $\psi_0 = 1$, la représentation moyenne mobile infinie du modèle (1.5.6) (pour simplifier nous posons $\psi_k = 0$ pour $k < 0$.).

Nous avons

$$
\begin{aligned}
\gamma(k) &= E\left[y_t \ y_{t+k}\right] \\
&= E\left[\sum_{j=0}^{\infty}\sum_{h=0}^{\infty} \psi_j \psi_h \ \varepsilon_{t-j} \ \varepsilon_{t+k-h}\right] \\
&= \sigma^2 \sum_{j=0}^{\infty} \psi_j \ \psi_{j+k}
\end{aligned}
$$

Nous remarquons que $\sigma^2 \psi_j$ coincide avec la $cov(y_{t+j}, \varepsilon_t)$. En multipliant les deux membres de (1.5.6) par y_{t-j} et ε_{t-j} respectivement et en prenant l'espérance nous obtenons un système dit *système de Box et Jenkins*

$$
\left\{
\begin{array}{ll}
\sum_{k=0}^{p} \phi_k \ \gamma(j-k) = \sigma^2 \sum_{k=0}^{q} \theta_k \ \psi_{j-k} \ , & j = 0, ..., p \\
\sum_{k=0}^{p} \phi_k \ \psi_{j-k} = \theta_j \ , \qquad j = 1, ..., q &
\end{array}
\right. \tag{1.5.7}
$$

avec $\phi_0 = \theta_0 = -1$.

B. Estimateurs de Box-Jenkins

Dans le cas où le modèle (1.5.6) est causal et inversible, les paramètres $\phi_1, ..., \phi_p$, et $\theta_1, ..., \theta_q$ peuvent être déterminés à partir de $\gamma(0), ..., \gamma(p), \psi_1, ..., \psi_q$ et σ^2 en résolvant le système (1.5.4). Ainsi en remplaçant cette fois-ci dans le système (1.5.5) les autocovariances $\gamma(0), ..., \gamma(p)$, les coefficients $\psi_1, ..., \psi_q$ et σ^2 par leurs estimateurs empiriques $\widehat{\gamma}(0), ..., \widehat{\gamma}(p), \widehat{\psi}_1, ..., \widehat{\psi}_q$ et $\widehat{\sigma}^2$, nous obtenons les estimateurs des moments $\widehat{\Phi}$ et $\widehat{\theta}$ de Φ et θ respectivement.

$$\begin{cases} \sum_{k=0}^{p} \phi_k \ \widehat{\gamma}(j-k), = \sigma^2 \sum_{k=0}^{q} \theta_k \ \widehat{\psi}_{k-j} \ , \qquad j = 0, ..., p \\ \sum_{k=0}^{p} \phi_k \ \widehat{\psi}_{j-k} = \theta_j \ , \qquad j = 1, ..., q \end{cases} \tag{1.5.8}$$

Franke (1985) a résolu le système de Box-Jenkins en adoptant une approche récursive (pour plus de détaille voir Franke).

Nous passerons ensuite à une autre méthode dite la méthode du maximum de vraisemblance (ML) pour estimer les paramètres d'un modèle ARMA(p, q). Cette méthode nécessite un traitement direct des données.

1.5.2 Estimation par la méthode du maximum de vraisemblance

Supposons que nous voulons estimer les paramètres de ce modèle en se basant sur une réalisation donnée. Alors, ces estimateurs seront choisit de façon à minimiser une fonction critère fixée.

Nous considérons le modèle (1.5.6) avec une hypothèse très précise sur la distribution des $\{\varepsilon_t\}$.

Soit $\{\varepsilon_t, t \in \mathbb{Z}\}$ une suite de variables aléatoires i.i.d, de loi $\mathcal{N}(0, \sigma^2)$. Pour une N-réalisation $y_1, ..., y_N$, le modèle (1.5.6) peut se mettre sous la forme suivante :

$$A \ Y_N = B \ \underline{\varepsilon}_N + C \ I_* \tag{1.5.9}$$

où $Y_N = (\ y_1, ..., y_N)'$, $\underline{\varepsilon}_N = (\varepsilon_1, ..., \varepsilon_N)'$, $I_* = (\ y_0, ..., y_{1-p}, \varepsilon_0, ..., \varepsilon_{1-q})'$. et A et B sont des matrices triangulaires inférieurs de taille $(N \times N)$ à diagonale unité. La matrice C est de taille $(N \times (p+q))$. L'expression exacte de ces matrices est donnée par :

$$A = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ -\phi_1 & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ -\phi_p & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & -\phi_p & \cdots & -\phi_1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ -\theta_1 & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ -\theta_q & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & -\theta_q & \cdots & -\theta_1 & 1 \end{pmatrix}$$

$$C = \begin{pmatrix} \phi_1 & \cdots & \cdots & \phi_p & -\theta_1 & \cdots & \cdots & -\theta_q \\ 1 & \phi_1 & \ddots & \phi_{p-1} & \vdots & -\theta_1 & \ddots & -\theta_{q-1} \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \phi_1 & \vdots & \ddots & \ddots & -\theta_1 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

A. Méthode du maximum de vraisemblance conditionnelle (CML)

Considérons d'abord la fonction de vraisemblance conditionnelle $\mathcal{L}_*(\phi, \theta, \sigma^2)$ du vecteur $\underline{\varepsilon}_N = (\varepsilon_1, ..., \varepsilon_N)'$ sachant que $I_* = 0$. Nous avons alors

$$\mathcal{L}_*(\phi, \theta, \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{N}\varepsilon_t^2\right\}. \tag{1.5.10}$$

Notons par $Y = (\ y_1, ..., y_N)^t$ le vecteur des observations. La matrice Jacobienne de la transformation dans $\mathbb{R}^{\ N}$ qui transforme ε à Y, à l'aide de l'équation (1.5.6), est triangulair inférieure et les éléments diagonaux sont égaux à 1.

La densité de Y sachant que $I_* = 0$ est donc donnée par (1.5.10) où ε_t est exprimé en fonction de y_t au moyen des relations (1.5.6). Nous aurons donc

$$\log \mathcal{L}_*(\phi, \theta, \sigma^2) = -\frac{N}{2}\log(2\pi\sigma^2) - \frac{S_*(\phi, \theta)}{2\sigma^2} \tag{1.5.11}$$

où $S_*(\phi, \theta) = \sum_{t=1}^{N}\varepsilon_t^2$.

Nous avons par définition, les estimateurs du maximum de vraisemblance conditionnelle de ϕ, θ et σ^2 les statistiques $\widetilde{\phi}, \widetilde{\theta}$ et $\widetilde{\sigma}^2$ qui maximisent le log de vraisemblance conditionnelle $\log \mathcal{L}_*(\phi, \theta, \sigma^2)$.

L'équation de vraisemblance correspondante à σ^2 fournit la relation suivante :

$$\sigma^2 = \frac{S_*(\phi, \theta)}{N} \tag{1.5.12}$$

par substitution dans (1.5.9), nous obtenons donc à une constante près

$$-\frac{1}{N}\log(\frac{S_*(\phi, \theta)}{N}) \tag{1.5.13}$$

Ce qui montre que les estimateurs du maximum de vraisemblance conditionnelle de ϕ et θ sont respectivement les statistiques $\widetilde{\phi}$ et $\widetilde{\theta}$, qui minimisent $S_*(\phi, \theta)$. Connaissant $\widetilde{\phi}$ et $\widetilde{\theta}$, nous déterminons l'estimateur $\widetilde{\sigma}^2$ de σ^2 par (1.5.12).

Dans ce travail, les estimateurs de ML seront notés par $\widetilde{\phi}_1, ..., \widetilde{\phi}_p$, $\widetilde{\theta}_1, ..., \widetilde{\theta}_q$et $\widetilde{\sigma}^2$.

B. Méthode du maximum de vraisemblance non-conditionnelle (ML)
La fonction de vraisemblance n'est pas simple, nous pouvons en effet l'écrire sous la forme générale

$$\mathcal{L}(\phi,\theta,\sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}}(\det\Gamma)^{-\frac{N}{2}} \exp\left\{-\frac{1}{2\sigma^2}Y^t\ \Gamma^{-1}\ Y\right\}.$$

où Γ est la matrice des covariances du vecteur Y, qui peut être obtenue à l'aide de (1.5.9). Nous remarquons que σ^2 vérifie

$$\sigma^2 = \frac{1}{N}\ Y^t\ \Gamma^{-1}\ Y \tag{1.5.14}$$

D'où $\log(\mathcal{L}(\phi,\theta,\sigma^2))$ peut s'écrire, à une constante prés, comme suit :

$$\begin{aligned}\log[\mathcal{L}(\phi,\theta,\sigma^2)] &\simeq -N\log(\sigma) - \tfrac{N}{2}\log(\det\Gamma) - \tfrac{1}{2\sigma^2}Y^t\ \Gamma^{-1}\ Y \\ &\simeq -\tfrac{N}{2}\log[\tfrac{1}{N}\ Y^t\ \Gamma^{-1}\ Y] - \tfrac{N}{2}\log(\det\Gamma)\end{aligned}$$

Ce qui montre qu'il suffit de minimiser

$$S(\phi,\theta) = (\det\Gamma)^{\frac{1}{N}}Y^{\ t}\ \Gamma^{-1}\ Y] \tag{1.5.15}$$

Dans le paragraphe suivant nous présentons une troisième méthode pour estimer les paramètres d'un modèle ARMA(p,q) et qui nécessite également un traitement direct des données. Cette méthode ne nécessite pas la supposition que $\{\varepsilon_t\}$ soit gaussien.

1.5.3 Estimation des paramètres par la méthode des moindres carrés

A. Méthode des moindres carrés conditionnelle (CLS)
La méthode des moindres carrés conditionnelle (CLS) est basée sur la somme des carrées conditionnelle, elle consiste à calculer

$$S_*(\phi,\theta) = \sum_{t=1}^{N}\varepsilon_t^2$$

où ε_t est calculée par récurrence, avec une valeur initiale $I_* = 0$. Nous pouvons donc déterminer à partir de (1.5.6) que *pour* $t \geq h \quad (h = \max(p,q))$

$$\varepsilon_t = y_t - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p} + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q}\ , \tag{1.5.16}$$

avec p et q fixés, nous pouvons ainsi, expliciter la valeur de ε_t seulement en fonction de ϕ,θ et y, en substituant les valeurs de chaque membre de la partie gauche de l'équation (1.5.14), par sa valeur récursive. Par la suite nous obtenons

$$\varepsilon_t = f_t(\phi_1,...,\phi_p,\theta_1,...,\theta_q,y_1,...,y_{t-1}) \tag{1.5.17}$$

où $f_t(.)$ est une fonction dépendante du modèle dont sa forme est bien connue dès que nous fixons p et q. La somme des carrés $S_*(\phi, \theta)$ conditionnelle à $I_* = 0$ s'écrit comme suit :

$$S_*(\phi, \theta) = \sum_{t=1}^{N} \varepsilon_t^2 = \sum_{t=1}^{N} f_t(\phi, \theta, y) \tag{1.5.18}$$

Nous remarquons que la fonction $f_t(.)$ n'est pas une fonction quadratique en ϕ et θ, donc il faut recourir à une procédure numérique itérative de régréssion non linéaire ou d'optimisation d'une fonction. La convergence n'est pas assuré, et peut dépendre fortement des valeurs initiales et survenir en un extremum local. Il n'y a généralement pas de problème quand le modèle est bien choisi. Box et Jenkins ont proposé une méthode de prévision rétrospective (*backforecasting*) qui améliore les conditions initiales.

B. Méthode des moindres carrés non-conditionnelle (OLS)

La méthode de Box et Jenkins (1976) consiste à déterminer une valeur de I_* meilleure que 0. Si le modèle (1.5.6) était correct, nous pouvons également écrire :

$$y_t - \phi_1 y_{t+1} - \cdots - \phi_p y_{t+p} = \varepsilon'_t - \theta_1 \varepsilon'_{t+1} - \cdots - \theta_q \varepsilon'_{t+q} \tag{1.5.19}$$

où $\{\varepsilon'_t\}$ sont les innovations en temps inversé. En supposant que $\varepsilon'_{N+k} = 0$ pour $k \geqslant 1$, nous déterminons successivement ε'_N puis ε'_{N-1}.....

De façon similaire à (1.5.17) nous obtenons finalement une estimation $\widehat{I}_*$ de I_* (prévision rétrospective). En considérant maintenant $\widehat{I}_*$ comme valeur conditionnelle de I_*, nous pouvons donc obtenir les ε_t par (1.5.16) seulement en fonction de $\phi, \theta, \widehat{I}_*$ et y . Nous déduisons alors la somme des carrés non conditionnelle.

$$S(\phi, \theta) = \sum_{t=1}^{N} \varepsilon_t^2. \tag{1.5.20}$$

Notons que dans le cas d'un modèle AR pur, les méthodes des moments, du maximum de vraisemblence et des moindres carrés possèdent des propriétés statistiques équivalentes. Dans le cas d'un modèle ARMA mixte, il est montré que les deux dernières méthodes sont plus efficaces.

1.6 Estimation de la fonction d'autocorrélation

Si $\{y_t, t \in \mathbb{Z}\}$ est un processus stationnaire du second ordre, alors nous nous contentons à caractériser y_t par sa moyenne μ et son fonction d'autocovariance $\gamma(.)$. L'estimation de μ, $\gamma(.)$ et de la fonction d'autocorrélation $\rho(.) = \gamma(.)/\gamma(0)$ à partir des observations $y_1, ..., y_N$, joue donc un rôle crucial dans les problèmes de l'inférence et en particulier dans le problème de l'identification du modèle. En cette section nous étudierons l'estimation de la moyenne, la fonction d'autocovariance et la fonction d'autocorrélation qui seront utilisés par la suite dans l'identification des modèles ARMA stationnaires. Puis nous examinerons les propriétés asymptotiques des estimateurs obtenus.

1.6.1 Estimation de la moyenne

Soit $y_1, ..., y_N$ une N-réalisation d'un processus $\{y_t, t \in \mathbb{Z}\}$ stationnaire. L'estimateur le plus simple et le plus naturel de la moyenne μ de y_t est la moyenne de l'échantillon

$$\overline{Y} = \frac{1}{N}\sum_{t=1}^{N} Y_t$$

qui est un esrtimateur sans biais.

Théorème (1.6.1)
Si $\{y_t, t \in \mathbb{Z}\}$ est un processus stationnaire de moyenne μ et de fonction d'autocovariance $\gamma(.)$, alors,

$$var(\overline{y}) = E(\overline{y}-\mu)^2 \underset{N\longrightarrow\infty}{\longrightarrow} 0 \qquad \text{si } \gamma(N) \underset{N\longrightarrow\infty}{\longrightarrow} 0$$

et

$$N \times var(\overline{y}) \underset{N\longrightarrow\infty}{\longrightarrow} \sum_{h=-\infty}^{\infty} \gamma(h) \qquad \text{si } \sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty.$$

Démonstration Voir par exemple, Brockwell et Davis (1988) p.219.

Remarques
1- Si $y_t = \mu + \sum_{j=-\infty}^{+\infty} \psi_j \varepsilon_{t-j}$ où $\sum_{j=-\infty}^{\infty} |\psi_j| < +\infty$. Alors $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$ et

$$N \times var(\bar{y}) \underset{N\longrightarrow\infty}{\longrightarrow} \sum_{h=-\infty}^{\infty} \gamma(h) = 2\pi S(0) = \sigma^2 \left(\sum_{j=-\infty}^{\infty} \psi_j\right)^2.$$

2- Le théorème (1.6.1) montre que si $\gamma(N) \longrightarrow 0$ lorsque $N \longrightarrow \infty$, alors $\overline{y}$ converge en moyenne quadratique (donc en probabilité) vers μ. De plus sous la condition forte $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$ (qui est satisfaite par tout processus ARMA(p,q)) nous avons $N \times var(\overline{y}) \longrightarrow \sum_{h=-\infty}^{\infty} \gamma(h)$, lorsque $N \longrightarrow \infty$ donc la moyenne de l'échantillon $\overline{y}$ est asymptotiquement normalement distribuée avec une moyenne μ et une variance $\frac{1}{N}\sum_{h=-\infty}^{\infty} \gamma(h) = \frac{\sigma^2}{N}\left(\sum_{j=-\infty}^{\infty} \psi_j\right)^2$.

Théorème (1.6.2)

Nous supposons que le processus $\{y_t, t \in \mathbb{Z}\}$ soit de la forme $y_t = \mu + \sum_{j=-\infty}^{+\infty} \psi_j \varepsilon_{t-j}$ où μ est une constante, $\{\varepsilon_t, t \in \mathbb{Z}\}$ une suite de variables aléatoires i.i.d de moyenne nulle, de variance σ^2, $\sum_{j=-\infty}^{\infty} |\psi_j| < +\infty$ et $\sum_{j=-\infty}^{\infty} \psi_j \neq 0$, alors $\overline{y}$ est asymptotiquement normalement distribuée avec une moyenne μ et de variance $\frac{1}{N} \sum_{h=-\infty}^{\infty} \gamma(h) = \frac{\sigma^2}{N} \left(\sum_{j=-\infty}^{\infty} \psi_j \right)^2$.

Démonstration Voir par exemple, Brockwell et Davis (1988) p.225-226.

1.6.2 Estimation de la fonction d'autocorrélation

Soit $y_1, ..., y_N$ une N-réalisation d'un processus stationnaire. Nous pouvons estimer l'ACVF par :

$$\widehat{\gamma}(j) = \widehat{\gamma}(-j) = \frac{1}{N} \sum_{t=1}^{N-j} (\ y_t - \overline{y}) (\ y_{t+j} - \overline{y}) \ , \qquad j = 0, 1, ..., N-1 \qquad (1.6.1)$$

où $\overline{y} = \frac{1}{N} \sum_{t=1}^{N} y_t$. Nous appelons l'estimateur $\widehat{\gamma}(j)$ l'ACVF de l'echantillon.

Nous définissons également l'ACRF de l'échantillon comme suit

$$\widehat{\rho}(j) = \frac{\widehat{\gamma}(j)}{\widehat{\gamma}(0)} \qquad (1.6.2)$$

Théorème (1.6.3) (*Distribution asymptotique de l'ACVF)*

Nous supposons que le processus $\{y_t, t \in \mathbb{Z}\}$ soit de la forme $y_t = \mu + \sum_{j=-\infty}^{+\infty} \psi_j \varepsilon_{t-j}$ où μ est une constante et $\{\varepsilon_t, t \in \mathbb{Z}\}$ une suite de variables aléatoires i.i.d de moyenne nulle, de variance σ^2 et $K_4 < \infty$.

Si $\sum |\psi_j| < +\infty$, alors $\sqrt{N}\ (\widehat{\gamma}(1) - \gamma(1)), ..., \sqrt{N}\ (\widehat{\gamma}(m) - \gamma(m))$ sont asymptotiquement normalement distribuées avec des moyennes nulles et de covariance

$$\lim_{N \longrightarrow +\infty} cov \left\{ \sqrt{N}\ (\widehat{\gamma}(r) - \gamma(r)), \sqrt{N}\ (\widehat{\gamma}(s) - \gamma(s)) \right\}$$
$$= \sum_{j=-\infty}^{+\infty} \{\gamma(j+r)\gamma(j+s) + \gamma(j-r)\gamma(j+s)\} + \frac{K_4}{\sigma^4}\gamma(r)\gamma(s)$$
$$= 4\pi \int_{-\pi}^{\pi} \cos(vr) \cos(vs) S^2(v) dv + \frac{K_4}{\sigma^4}\gamma(r)\gamma(s).$$

où $S(v)$ est la densité spectrale de $\{y_t, t \in \mathbb{Z}\}$.

Démonstration Voir par exemple, Brockwell et Davis (1988) p.220-236.

Théorème (1.6.4) (*Distribution asymptotique de l'ACRF)*

Nous supposons que le processus $\{y_t, t \in \mathbb{Z}\}$ soit de la forme $y_t = \mu + \sum_{j=-\infty}^{+\infty} \psi_j \varepsilon_{t-j}$ où μ est une constante et $\{\varepsilon_t, t \in \mathbb{Z}\}$ une suite de variables aléatoires i.i.d de moyenne nulle et de variance σ^2.

Si $\sum |\psi_j| < +\infty$, $\sum \ |j| \ |\psi_j|^2 < +\infty$

Alors $\sqrt{N} \ (\widehat{\rho}(1) - \rho(1)), ..., \sqrt{N} \ (\widehat{\rho}(N) - \rho(N))$ sont asymptotiquement normalement distribuées avec des moyennes nulles et de covariance

$$\begin{aligned} w_{i,j} &= \lim_{N \longrightarrow +\infty} cov \left\{ \sqrt{N} \ \left(\widehat{\rho}(i) - \rho(i)\right), \sqrt{N} \ \left(\widehat{\rho}(j) - \rho(j)\right) \right\} \\ &= \textstyle\sum_{k=-\infty}^{+\infty} \{ \rho(k+i)\rho(k+j) + \rho(k-i)\rho(k+j) \\ &\qquad + 2\rho(i)\rho(j)\rho^2(k) - 2\rho(i)\rho(k)\rho(k+j) - 2\rho(j)\rho(k)\rho(k+j) \}. \end{aligned}$$

Démonstration Voir par exemple, Brockwell et Davis (1988) p.220-236.

1.6.3 Algorithme récursive pour estimer les autocovariances

Cet algorithme, dû a Wilson (1979), n'est rien d'autre qu'une procédure inverse de l'algorithme de Durbin(1960). Son objectif est de construire, de façon récursive, les autocovariances du modèle à partir de ses paramètres.

Algorithme de Wilson pour estimer les autocovariances d'un modèle AR

Nous supposons que les paramètres $\phi_1, ..., \phi_p$ et σ^2 du modèle AR sont connus.

$1 \overset{ere}{=}$ *étape* : Initialisation

Posons $\sigma_p^2 = \sigma^2$ et $\phi_{p,k} = \phi_k$, $\qquad k = 1, ..., p.$

$2 \overset{eme}{=}$ *étape* : Constitution des $\phi_{k,\ j}$

Pour $k = p-1, ..., 0$

$$\begin{aligned} \alpha_{k+1} &= \phi_{k+1,k+1} \\ \phi_{k,\ j} &= (\phi_{k+1,\ j} - \alpha_{k+1}\phi_{k+1,\ k+1-j})/(1-\alpha_{k+1}^2), \qquad j = 1, ..., k, \qquad (k \geq 1) \\ \sigma_k^2 &= \ \sigma_{k+1}^2/(1-\alpha_{k+1}^2). \end{aligned}$$

$3 \overset{eme}{=}$ *étape* :

Posons $\gamma(0) = \sigma_0^2$

Pour $k = 0, ..., p-1,$

$$\gamma(k+1) = \ \phi_{k,1} \ \gamma(k) + ... + \phi_{k,k} \ \gamma(1) + \sigma_k^2 \ \alpha_{k+1}. \qquad \square$$

Remarque L'algorithme précédent permet de déterminer les autocovariances pour $k = 0, 1, ..., p$, uniquement. Pour $k > p$, nous utiliserons la relation

$$\gamma(k+1) = \quad \phi_1 \; \gamma(k) + ... + \phi_p \; \gamma(k+1-p). \tag{1.6.3}$$

Cet algorithme de construction des $\gamma(0), ..., \gamma(p)$ à partir de $\phi_1, ..., \phi_p$ et σ^2 requiert un nombre d'opérations de l'ordre $O(p^2)$ par contre l'algorithme de McLeod (1975) nécessite un nombre d'opérations d'ordre $O(p^3)$.

Algorithme d'estimation récursive des autocovariances d'un modèle ARMA

Pour un modèle ARMA, le calcul de l'ACVF peut être établit de la même manière que celle du modèle AR, mais nous allons utiliser une approche alternative basée sur la fonction génératrice.

Considèrons le modèle (1.5.6) et supposons que $\phi(z) \neq 0$ pour tout $z \in C$ tel que $|z| \leq 1$. Alors la fonction génératrice des autocovariances $\Gamma(z) = \sum_{k=-\infty}^{+\infty} \gamma(k) \; z^k$ vérifie

$$\Gamma(z) = \sigma^2 \; \theta(z) \; \theta(z^{-1}) / \; \phi(z) \; \phi(z^{-1}) \tag{1.6.4}$$

Cependant, le second membre de cette équation peut être factoriser comme suit :

$$\theta(z) \; \theta(z^{-1}) / \; \phi(z) \; \phi(z^{-1}) = v(z)/\phi(z) + v(z^{-1})/\phi(z^{-1})$$

où $v(z) = \sum_{h=0}^{r} v_h \; z^h$ et $r = \max(p, q)$. Les coefficients de $v(z)$ sont obtenus en résolvant les équations linéaires suivantes

$$\phi(z) \; v(z^{-1}) + \phi(z^{-1}) v(z) = \theta(z) \; \theta(z^{-1})$$

Pour obtenir les $\gamma(k)$, nous dévloppons $v(z)\phi(z) = \sum_{h=0}^{\infty} g_h \; z^h$ par

$$g_k = \phi_1 g_{k-1} + \cdots + \phi_p g_{k-p} + v_k \; , \qquad k = 0, 1, 2, ...$$

En prenant $g_k = 0$ pour $k \leq 0$ et $v_k > r$. Alors nous aurons finalement,

$$\gamma(0) = 2 g_0 \sigma^2 \;\; et \;\; \gamma(k) = g_k \sigma^2 \; pour \;\; k > 0. \tag{1.6.5}$$

Cette approche reduit considérablement le temps de calcul et l'espace mémoire. Cependant, afin de déterminer les v_k il faut résoudre un système de r équations. Wilson (1979) a proposé un algorithme qui requiert un nombre d'opérations de l'ordre $O(p^2)$.

Algorithme de Wilson (1979)

Etant donné le modèle (1.5.6) (nous supposons que $p > q$).

Nous supposons aussi que les paramètres $\phi_1, ..., \phi_p, \theta_1, ..., \theta_q$ et σ^2 soient connus et que nous désirons déterminer les autocovariance $\gamma(1), \gamma(2), ...$ Alors en appliquant la procédure suivante

$1 \overset{ere}{=}$ *étape* : Déteminer les coefficients B_k de $B(z) = \sum_{k=-q}^{q} B_k \ z^k$ par la relation

$$B(z) = \theta(z)\,\theta(z^{-1}) \text{ ou également } B_k = B_{-k} = \sum_{i=0}^{q-k} \theta_i\,\theta_{i+k}\ , \qquad \textit{pour } k = 0, 1, ..., q,$$

soit $B_k = 0$ pour $k = q+1, ..., p$,

$2 \overset{eme}{=}$ *étape* : Le calcul des coefficients $\phi_{i,\ j}$ $\quad i, j = 1, ..., p$

Posez $\phi_{p,\ k} = \phi_k$, $\qquad k = 1, ..., p$

Posez $\alpha_{k+1} = \phi_{k+1,\ k+1}$ pour $k = p-1, ..., 0$

$$\phi_{k,\ j} = (\phi_{k+1,\ j} - \alpha_{k+1}\phi_{k+1,\ k+1-j})/(1 - \alpha_{k+1}^2), \qquad (k \geqslant 1) \quad \textit{pour } j = 1, ..., k,$$

$3 \overset{eme}{=}$ *étape* : Le calcul des coefficients $B_{i,j}$, $i, j = 1, ..., p$,

Pour $k = 1, ..., p$ faire $B_{p,j} = B_j$

$$\delta_{k+1} = B_{k+1,k+1}$$

Pour $j = 1, ..., k$ faire $B_{k,j} = B_{k+1,j} + \delta_{k+1}\phi_{k,k+1-j}$ avec $k \geqslant 1$.

$4 \overset{eme}{=}$ *étape* : Le calcul des coefficients v_j , $j = 1, ..., p$

Posez $v_{0,0} = \frac{1}{2}B_{0,0}$, $v_{k,k+1} = \delta_{k+1}$ $\qquad k = 0, 1, ..., p-1$,

$$\phi_{k,\ j} = (v_{k,\ j} - \alpha_{k+1}v_{k,\ k+1-j})/(1 - \alpha_{k+1}^2),\ j = 0, 1, ..., k+1$$

Poser $v_j = v_{p,\ j}$, $\qquad j = 1, ..., p$,

$5 \overset{eme}{=}$ *étape* : Le calcul des coefficients g_k , $\qquad k = 0, 1, 2, ...$

$$g_k = \phi_1 g_{k-1} + \cdots + \phi_p g_{k-p} + v_k$$
$$g_k = 0 \ \ si \ \ k < 0 \ \ et \ \ v_k = 0 \quad pour \ \ k > r.$$

$6 \overset{eme}{=}$ *étape* : Le calcul des $\gamma(h)$, $\ h = 0, 1, 2, ...$

$$\gamma(0) = 2g_0\sigma^2\ ,$$
$$\gamma(k) = g_k\sigma^2 \ pour \ \ k > 0. \qquad \square$$

Chapitre 2

Identification des modèles ARMA stationnaires

Introduction

Au cours de ces quatre dernières décennies, des progrès considérables ont été connus dans l'analyse de série chronologique. En particulier, de nombreuses techniques utilisant le modèle autorégressif moyenne mobile (ARMA) ont été développées.

L'utilisation de modèle ARMA pour représenter et analyser des séries chronologiques à temps discret est devenue courante depuis les travaux de Box et Jenkins (1970). Ce sont des processus stochastiques adaptés à des fins de prévisions et construits pour fournir une explication du présent par le passé. De ce fait, les processus ARMA(p, q) stationnaires ont été largement étudiés.

L'identification, ou la sélection de des ordres, d'un modèle est considérée comme l'un des problemes majeurs dans l'étude des séries chronologiques. Car elle représente la première étape, selon la méthodologie de Box et Jenkins, dans la modélisation des séries temporelles.

Box et Jenkins (1976) ont utilisé la fonction d'autocorrélation et la fonction d'autocorrélation partielle. Au début des années 70, Akaike a présenté deux méthodes d'identification connues sous le nom de FPE (*Final Prediction Error*) et AIC (*Akaike information Criterion*). Pour éviter le problème de l'inconsistance des estimateurs obtenus, le critère BIC (*Bayesian information Criterion)* et la méthode de Hannan-Quinn (1979) ont été proposés. En même temps, Cleveland a suggéré l'utilisation de la fonction d'autocorrélation inverse. En 1982, Hannan et Rissanen ont utilisé une technique instrumentale de régression en tant que fonctions de prévision pour la modélisation d'un modèle ARMA. Cette technique a eu comme conséquence des estimateurs consistants des ordres.

Tout au long des années 70 et 80, plusieurs méthodes d'identification basées sur quelques fonctions d'autocorrélation ont été étudiées, parmis ces méthodes nous pouvons citer la méthode de Woodside (1971), la méthode R et S de Gris et *al.* (1978), la méthode du coin de Beguin et *al.* (1980), les trois méthodes GPAC (*Generalized Partial autocorrelation*) de Woodward et Gris (1981), de Glasbey (1982) et de Takemura (1984), la méthode ESACF (*Extended sample autocorrelation function*) de Tsay et Tiao (1984) et la méthode SCAN de Tsay et Tiao (1985).

Dans les années 90 trois critères d'identification ont été proposés. Le premier c'est la méthode 3-pattern de Choi (1991a) pour identifier un modèle mixte ARMA à coefficients constants stationnaire en utilisant quelques fonctions d'autocorrélation. Les deux derniers sont des critères pour sélectionner l'ordre d'un modèle autoregressif pur à coefficients constants stationnaire. Le deuxième critère a été élaboré par Djuric et Kay (1992) nommé PDC (*Predictive Density Criterion*) et qui est fondé sur la densité prédictive bayesienne. Le troisième, c'est le CIC (*Consistent Information Criterion*) présenté par Ciftcioglu et *al.* (1994) et qui est fondé sur un test d'hypothèse statistique. Ce critère a comme conséquence un estimateur consistant pour l'ordre d'un modèle AR pur.

Notre objectif, est d'étudier théoriquement et algorithmiquement la performance des critères de sélection des ordres des modèles autoregressifs moyennes mobiles à coefficients constants ARMA(p, q) stationnaire. Ce chapitre présente les méthodes les plus connues. En premier lieu, nous commençons par les méthodes d'identification basées sur la fonction d'autocorrélation et la fonction d'autocorrélation partielle. Nous présenterons, à ce sujet la méthodologie de Box-Jenkins (1976), la méthode du coin de Beguin et *al.* (1980) et la méthode 3-pattern de Choi (1991a). Nous passerons en suite à la méthode de l'erreur de prédiction finale (FPE). Puis, aux méthodes d'identification basées sur la théorie de l'information où nous exposerons le critère le plus connu et qui se base sur une mesure de l'écart entre la vraie loi (inconnue) et celle du modèle proposé. La mesure habituellement utilisée est l'information de Kullback. Dans la quatrième section, nous exposerons une généralisation du critère AIC pour choisir l'ordre d'un modèle AR pur. Ce critère est nommé CIC. Nous présenterons par la suite (sans détail) la classe la plus importante qui est la classe des méthodes d'identification basées sur l'analyse bayesienne. Cette classe est très importante, puisque elle résoud le problème de l'inconsistance des estimateurs obtenus. Pour celà, nous exposerons le critère de l'information de bayes (BIC) et le critère de la densité prédictive (PDC). Nous recapitulerons en suite, dans la sixième section le critère de Hannan-Quinn (HQC). Nous terminerons le chapitre par une étude de simulation, en vérifiant les résultats théoriques à travers une étude de simulation sur des séries artificielles supposées obtenues à partir de modèles bien spécifiés.

2.1 Méthodes d'identification basées sur la fonction d'autocorrélation

Le premier qui avait utilisé l'ACRF et la PCRF pour choisir les ordres des processus AR pur et MA pur c'est Rudra (1952). Il a appliqué la formule de Bartlett pour la variance asymptotique de l'ACRF de l'échantillon. Pour examiner la propriété de troncature de la PACF, il a utilisé la statistique du rapport de vraisemblance

$$\frac{\sqrt{N}\widehat{\phi}_{k,k}}{\sqrt{1-\widehat{\phi}_{k,k}^2}}, \tag{2.1.1}$$

qui est asymptotiquement normalement distribué de moyenne nulle et de variance 1 pour $k = p+1, p+2, ...$

2.1.1 Théorèmes classiques d'identification d'un AR(p) et MA(q)

Théorème (2.1.1)
Soit $\{y_t, t \in \mathbb{Z}\}$ un processus stationnaire et $\{\phi_{k,k} \backslash\ k = 1, 2, ...\}$ sa fonction d'autocorrélation partielle (PACF). Alors y_t est un AR(p), si et seulement si

$$\begin{cases} \phi_{k,k} \neq 0 & si\ k \leq p \\ \phi_{k,k} = 0 & si\ k > p \end{cases}$$

Démonstration Voir par exemple, A. Hamdi (1982).

Théorème (2.1.2)
Soit $\{y_t, t \in \mathbb{Z}\}$ un processus stationnaire de fonction d'autocorrélation $\{\rho(k), k \in \mathbb{Z}\}$. Alors y_t est un MA(q) si et seulement si

$$\begin{cases} \rho(k) \neq 0 & si\ k \leq p \\ \rho(k) = 0 & si\ k > p \end{cases}$$

Démonstration Voir par exemple, A. Hamdi (1982).

Box et Jenkins (1976) ont utilisé la fonction d'autocorrélation et la fonction d'autocorrélation partielle dans l'identification du modèle, nous présentons ci-dessous leurs méthodologie.

2.1.2 Méthode de Box-Jenkins

La méthode d'identification de Box et Jenkins (1976) est fondée sur la comparaison des moments empiriques de la série considérée aux moments théoriques associés aux différentes représentations potentielles. Nous nous concentrerons ici sur les moments d'ordre deux résumés par la fonction d'autocorrélation (ACRF) et la fonction dautocorrélation partielle (PACF).

A- Utilisation de la fonction d'autocorrelation dans l'identification

Nous savons que, pour N assez grand, le biais de ACRF de l'échantillon est

$$E\left(\widehat{\rho}(k)\right) - \rho(k) = -\frac{1}{N}\sum_{j=-\infty}^{+\infty} \rho(j) + o\left(\frac{1}{N}\right). \tag{2.1.2}$$

(voir, par exemple T.W.Anderson(1971)).

La distribution asymptotique de l'ACRF de l'échantillon (le théorème (1.6.2)) prouve que la covariance $cov\left(\widehat{\rho}(r), \widehat{\rho}(s)\right)$ de l'ACRF de l'échantillon est égale à

$$\begin{aligned}\tfrac{1}{N}\textstyle\sum_{j=-\infty}^{+\infty}\{&\rho(j+r)\rho(j+s) + \rho(j-r)\rho(j+s) + 2\rho(r)\rho(s)\rho^2(j)\\ &-2\rho(r)\rho(j)\rho(j+s) - 2\rho(s)\rho(j)\rho(j+r)\} + o\left(\tfrac{1}{N}\right).\end{aligned}$$

En particulier, si le processus est gaussien, les expressions approximatives de la variance et de la covariance deviennent respectivement, d'après Bartlett (1946), comme suit :

$$\begin{aligned}var\left(\widehat{\rho}(k)\right) \simeq \tfrac{1}{N}\textstyle\sum_{j=-\infty}^{+\infty}\{&\rho^2(j) + \rho(j+k)\rho(j-k)\\ &-4\rho(k)\rho(j)\rho(j-k) + 2\rho^2(j)\rho^2(k)\},\end{aligned} \tag{2.1.3}$$

$$cov\left(\widehat{\rho}(r), \widehat{\rho}(s)\right) \simeq \frac{1}{N}\sum_{j=-\infty}^{+\infty}\{\rho(j)\rho(j+r-s) + \rho(j-r)\rho(j+s)\}.$$

D'après le théorème (2.1.2), si $\{y_t, t \in \mathbb{Z}\}$ est un processus MA(q), alors l'ACRF a une troncature au retard q, c'est-à-dire, $\rho(k) = 0 \quad$ pour $k = q+1, q+2....$ Bartlett (1946) a donné l'estimateur de l'écart type de l'ACRF de l'échantillon comme suit :

$$\widehat{\sigma}\left(\widehat{\rho}(k)\right) \simeq \left\{\frac{1}{N}\left(1 + 2\sum_{j=1}^{q}\widehat{\rho}^2(j)\right)\right\}^{1/2}, \qquad k = q+1, q+2, ... \tag{2.1.4}$$

Puisque le biais et l'écart type de l'ACRF de l'échantillon sont d'ordre $o\left(N^{-1}\right)$ et $o\left(N^{-1/2}\right)$ respectivement, alors l'intervalle de confiance approximatif de niveau $(1-\alpha)$ pour $\rho(i)$ est donné par :

$$\left(-q_{\alpha/2}\widehat{\sigma}\left(\widehat{\rho}(i)\right),\ q_{\alpha/2}\widehat{\sigma}\left(\widehat{\rho}(i)\right)\right), \qquad i = q+1, q+2, ...$$

où $q_{\alpha/2}$ est le quantil d'ordre $\alpha/2$ de la loi normale $\mathcal{N}(0,1)$. Puisque l'ACRF d'un AR ou d'un processus mixte ARMA diminue exponentiellement vers 0, par conséquent l'ACRF de l'échantillon évolue de la même manière. De plus, si q^* est le plus petit nombre, tel que

$$\widehat{\rho}(i) \in \left(-q_{\alpha/2}\widehat{\sigma}\left(\widehat{\rho}(i)\right),\ q_{\alpha/2}\widehat{\sigma}\left(\widehat{\rho}(i)\right)\right), \qquad i = q^* + 1, q^* + 2, ... \tag{2.1.5}$$

Alors l'échantillon peut être considéré comme une réalisation d'un modèle MA(q^*).

B- Utilisation de la fonction d 'autocorrelation partielle (PCRF) dans l'identification

Nous passons maitenant à un autre outil de Box-Jenkins pour identifier un modèle ARMA qui est la fonction d'autocorrelation partielle.

Notons par $\phi_{k,k}$ le coefficient d'autocorrélation partielle au retard k du processus $\{y_t, t \in \mathbb{Z}\}$, et $\{\phi_{k,k} \backslash\ k = 1, 2, ...\}$ sa fonction d'autocorrélation partielle (PACF). Le terme *autocorrélation partielle* signifie que $\phi_{k,k}$ est le coefficient de corrélation conditionnel de y_t et y_{t+k} sachant $y_{t+1}, ..., y_{t+k-1}$. Avec plus de précision, nous considèrons l'espace de Hilbert d'un processus stationnaire $\{y_t, t \in \mathbb{Z}\}$ avec l'espérance comme un produit scalaire.

Soit $L_{s,t}$ le sous-espace engendré par $\{y_{s+1}, .., y_{t-1}\}$. Si $\widetilde{y}_t$et $\widetilde{y}_{t+k}$ sont respectivement les projections de y_t et y_{t+k} dans $L_{t,t+k}$, alors $\phi_{k,k}$ est la corrélation de $\widetilde{y}_t - y_t$ et $\widetilde{y}_{t+k}- y_{t+k}$.

D'où

$$\phi_{k,k} = \frac{\langle \widetilde{y}_t - y_t,\ \widetilde{y}_{t+k} - y_{t+k} \rangle}{\|\widetilde{y}_t - y_t\|\ \|\widetilde{y}_{t+k} - y_{t+k}\|}. \tag{2.1.6}$$

Pour la démonstration de cette interprétation voir, par exemple Brockwell-Davis (1987, p 164) ou A. Hamdi (1982, p 141-149).

Pour utiliser la PACF comme outil d'identification du modèle, il est nécessaire d'étudier la distribution asymptotique de l'estiomateur de YW de la PCAF.

Théorème (2.1.3) (*Distribution asymptotique de la PACF*)

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus AR(p). Alors $\sqrt{N}\widehat{\phi}_{p+1,p+1}, \sqrt{N}\widehat{\phi}_{p+2,p+2}, ...$, sont des variables aléatoires asymptotiquement indépendantes de moyennes nulles et de variances 1.

Démonstration Voir par exemple, Brockwell et Davis (1988) p.262-264.

Si le processus fondamental est un modèle AR(p), alors (d'après le théorème (2.1.1)), nous avons $\phi_{p+1,p+1} = \phi_{p+2,p+2} = \cdots = 0$. Puisque le biais de $\left\{\widehat{\phi}_{k,k}\right\}$ est d'ordre $o\left(N^{-1}\right)$, le théorème (2.1.3) nous donne un intervalle de confiance approximatif de niveau $(1 - \alpha)$ pour $\phi_{k,k}$ comme suit :

$$\left(-q_{\alpha/2}\frac{1}{\sqrt{N}},\ q_{\alpha/2}\frac{1}{\sqrt{N}}\right), \qquad k = p + 1, p + 2, ...$$

D'autre part, la PACF d'un processus MA ou d'un ARMA tend exponentiellement vers 0, par conséquent l'estimateur de YW de la PACF a le même comportement. Ainsi, si p^* est le plus petit entier qui satisfait

$$\widehat{\phi}_{k,k} \in \left(-q_{\alpha/2}\frac{1}{\sqrt{N}},\ q_{\alpha/2}\frac{1}{\sqrt{N}}\right), \qquad k = p^*+1, p^*+2, ... \tag{2.1.7}$$

Alors l'échantillon peut être considéré comme une réalisation d'un modèle AR(p^*).

La méthode de Box-Jenkins n'est pas très utile pour identifier un modèle ARMA mixte lorsque $p \neq 0$ et $q \neq 0$. Car la fonction d'autocorrélation d'un modèle ARMA mixte se compose d'un mélange de fonctions exponentielles et/ou sinusoidales atténuées après $q-p$ premiers retards, et la fonction d'autocorrélation partielle se compose également d'un mélange de fonctions exponentielles et/ou de sinusoidales atténuées après $p-q$ premiers retards. Ainsi, même si les graphes de $\{\rho(k)\}$ et $\{\phi_{k,k}\}$ comportent des erreurs, une simple constatation de leurs allures, en général, ne rapporterait pas des valeurs uniques de p et q. La difficulté est grande lorsque on remplace les vraies fonctions par leurs estimateurs (pour plus de détail voir par exemple, Box-Jenkins (1976), p.79 et Choi (1992), p. 31).

De ce fait, la méthode de Box et Jenkins doit être appliquée avec précaution car les autocorrélations successives tendent à être fortement corrélées. D'ailleurs, si N est petit, le biais de l'ACRF de l'échantillon et l'estimateur de YW de la PACF ne sont pas négligeables.

2.1.3 Identification d'un modèle ARMA par la méthode du coin

Avant de présenter la méthode du coin de Beguin et *al.* (1980) pour identifier un modèle autorégressif moyenne mobile stationnaire à coefficients constants, nous donnerons d'abord la définition d'une représentation canonique et minimale d'un processus ARMA. Ensuite nous présenterons une caractérisation d'un processus ARMA par des équations aux différences stochastiques sur sa fonction d'autocorrélation et enfin quelques propriétés du déterminant de corrélation.

A. Représentation canonique et minimale d'un processus ARMA

Considèrons le modèle ARMA(p, q) donné par :

$$y_t - \sum_{i=1}^{p} \phi_i y_{t-i} = \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \tag{2.1.8}$$

où ε_t est un processus bruit blanc, $\phi_p\theta_q \neq 0$ et les $\{\phi_i, 1 < i < p\}$ et $\{\theta_j, 1 < j < q\}$ sont déterminés de façon unique.

Définition (2.1.1)

i) La représentation (2.1.8) est dite canonique si les deux polynomes caractéristiques Φ et Θ associés vérifient les deux conditions suivantes :

1) Φ et Θ n'ont pas de zéro commun,
2) Φ et Θ ont leurs racines à l'extérieur du cercle unité.

ii) Dans la famille d'équations du type (2.1.8), une équation minimale ou représentation minimale est une équation dont la somme du nombre de paramètres dans la partie autoregressive et dans la partie moyenne mobile est la plus petite possible.

iii) Nous disons qu'un processus ARMA est un ARMA(p, q) si son équation canonique fait intervenir $p + q$ paramètres.

Remarque Pour l'équation canonique, nous avons une seule équation ARMA, mais pour la représentation minimale nous avons plusieurs équations avec le même nombre de paramètres.

B. Caractérisation d'un processus ARMA par des équations aux différences stochastique sur sa fonction d'autocorrélation

Définition (2.1.2) Une suite réelle $(X_t)_{t\in\mathbb{Z}}$ satisfait une équation aux différences d'ordre p à partir du rang q, s'il existe p coefficients $\phi_1, \phi_2, ..., \phi_p$, tels que $\phi_p \neq 0$ et

$$X_t - \sum_{j=1}^{p} \phi_j X_{t-j} = \begin{cases} 0 & si \quad t \geq q+1 \\ \alpha & si \quad t = q \end{cases} \qquad où \quad \alpha \neq 0.$$

Notons par $H^2(y_t, t)$ le sous-espace de Hilbert, $H^2(y_t,]-\infty, t])$, engendré par le passé et le présent $(\ldots, y_{t-2}, y_{t-1}, y_t)$ du processus du second ordre $\{y_t, t \in \mathbb{Z}\}$.

Définition (2.1.3) Un processus stichastique, du second ordre, $\{y_t, t \in \mathbb{Z}\}$ est dit purement indéterminable ou encore régulier, si

$$\lim_{t \longrightarrow -\infty} H^2(y_t, t) = \{0\}$$

et purement déterminable ou singulier si

$$\lim_{t \longrightarrow -\infty} H^2(y_t, t) = \lim_{t \longrightarrow +\infty} H^2(y_t, t).$$

Théorème (2.1.4)

Un processus régulier $\{y_t, t \in \mathbb{Z}\}$ est un processus ARMA(p, q) si et seulement si sa fonction d'autocorrélation (ACRF) satisfait une équation aux différences d'ordre minimal p à partir du rang minimal q, i.e, il existe des constantes $\phi_1, \phi_2, ..., \phi_p$ tels que :

$$\rho(j) - \sum_{i=1}^{p} \phi_i \; \rho(j-i) = \begin{cases} 0 & si \quad j = q+1, q+2, ... \\ \alpha & si \quad j = q \end{cases} \tag{2.1.9}$$

où $p > 0, \phi_p \neq 0, q \geq 0$ et q est le plus petit entier satisfaisant (2.1.9).

Démonstration Voir par exemple, A. Hamdi (1982, p. 34)

Proposition (2.1.1)

Si un processus stochastique stationnaire a une représentation (2.1.8) (représentation ARMA(p, q)), alors pour $k = p, p+1, ...$ et $i = q, q+1, ...$ le rang de la matrice d'autocorrélation $\Sigma(k, i)$ est égal à $k - \min(k-p, i-q)$.

C. Déterminant d'autocorrélation

Nous définissons d'abord la matrice d'autocorrélations d'ordre k par :

$$\Sigma(k,i) = \begin{pmatrix} \rho(i) & \rho(i-1) & \cdots & \rho(i-k+1) \\ \rho(i+1) & \rho(i) & \cdots & \rho(i-k+2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(i+k-1) & \rho(i+k-2) & \cdots & \rho(i) \end{pmatrix},$$

$\rho(.)$ étant la fonction d'autocorrélation de y_t.

Nous notons par $\Delta(k, i)$ le déterminant de la matrice $\Sigma(k, i)$, et appelons l'ensemble $\{\Delta(k, i)/k = 1, 2, ..., \; i = 1, 2, ...\}$ l'ensemble Δ du processus ARMA.

Nous présentons dans ce qui suit quelques relations utiles entre les coefficients d'un modèle ARMA(p, q), puis d'autres relations entre les déterminants $\Delta(k, i)$ et les coefficients du modèle.

En multipliant l'équation aux différences (2.1.8) par y_{t-q-h}, en passant à l'espérance et en tenant compte que ε_t est un bruit blanc nous trouvons

$$E\left[y_{t-q-h}\left(y_t - \sum_{i=1}^{p} \phi_i y_{t-i}\right)\right] = E\left[y_{t-q-h}\left(\varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}\right)\right]$$

qui implique

$$\gamma(q+h) - \sum_{i=1}^{p} \phi_i \gamma(q+h-i) = -\sum_{j=1}^{q} \theta_j cov\left(y_{t-q-h}, \varepsilon_{t-j}\right)$$

donc

$$\gamma(q+h) - \sum_{i=1}^{p} \phi_i \gamma(q+h-i) = 0 \qquad pour\ h = 1, 2, ...$$

où encore

$$\rho(q+h) - \sum_{i=1}^{p} \phi_i \rho(q+h-i) = 0 \qquad pour\ h = 1, 2, ... \tag{2.1.10}$$

et

$$\rho(q) - \sum_{i=1}^{p} \phi_i \rho(q-i) = -\theta_q \sigma^2/\gamma(0) \qquad pour\ h = 0 \tag{2.1.11}$$

Proposition (2.1.2)
Les déterminants $\Delta(n,m)$ vérifient les deux propriétés suivantes :

$$\begin{aligned} \Delta(p+n,q) &= -\theta_q \frac{\sigma^2}{\gamma(0)} \Delta(p+n-1,q), \qquad n \geq 1, \\ \Delta(p,q+m) &= (-1)^{p-1} \phi_p \Delta(p,q+m-1), \qquad m \geq 1. \end{aligned}$$

Démonstration
Nous savons que

$$\Delta(k,i) = \begin{vmatrix} \rho(i) & \rho(i-1) & \cdots & \rho(i-k+1) \\ \rho(i+1) & \rho(i) & \cdots & \rho(i-k+2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(i+k-1) & \rho(i+k-2) & \cdots & \rho(i) \end{vmatrix}$$

donc

$$\Delta(p+n,q) = \begin{vmatrix} \rho(q) & \rho(q-1) & \cdots & \rho(q-p-n+1) \\ \rho(q+1) & \rho(q) & \cdots & \rho(q-p-n+2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(q+p+n-1) & \rho(q+p+n-2) & \cdots & \rho(q) \end{vmatrix}$$

Posons $C'_1 = C_1 - \sum_{i=2}^{p} \phi_i C_i$ où C_i est la $i \overset{eme}{=}$ colonne de la matrice $\Sigma(p+n,q)$.
D'après (2.1.10) et (2.1.11), nous trouvons

$$\Delta(p+n,q) = \begin{vmatrix} -\theta_q \frac{\sigma^2}{\gamma(0)} & \rho(q-1) & \cdots & \rho(q-p-n+1) \\ 0 & \rho(q) & \cdots & \rho(q-p-n+2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \rho(q+p+n-2) & \cdots & \rho(q) \end{vmatrix}$$

D'ou

$$\Delta(p+n,q) = -\theta_q \frac{\sigma^2}{\gamma(0)} \Delta(p+n-1,q), \qquad \textit{pour } n \geq 1 \qquad (2.1.12)$$

D'autre part, nous avons

$$\Delta(p,q+m) = \begin{vmatrix} \rho(q+m) & \rho(q+m-1) & \cdots & \rho(q+m-p+1) \\ \rho(q+m+1) & \rho(q+m) & \cdots & \rho(q+m-p+2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(q+m+p-1) & \rho(q+m+p-2) & \cdots & \rho(q+m) \end{vmatrix}$$

$$\begin{aligned} &= (-1)^{p-1} \begin{vmatrix} \rho(q+m+p-1) & \rho(q+m+p-2) & \cdots & \rho(q+m) \\ \rho(q+m) & \rho(q+m-1) & \cdots & \rho(q+m-p+1) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(q+m+p-2) & \rho(q+m+p-3) & \cdots & \rho(q+m-1) \end{vmatrix} \\ &= (-1)^{p-1} \Sigma'(p,q+m) \end{aligned}$$

Posons $L'_p = L_p - \sum_{i=2}^{p} \phi_i L_i$ où L_i est la $i-p+1 \overset{eme}{=}$ ligne de la matrice $\Sigma'(p,q+m)$. D'après (2.1.10) et (2.1.11), nous trouvons

$$(-1)^{p-1} \begin{vmatrix} \phi_p \rho(q+m-1) & \phi_p \rho(q+m-2) & \cdots & \phi_p \rho(q+m-p) \\ \rho(q+m) & \rho(q+m-1) & \cdots & \rho(q+m-p+1) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(q+m+p-2) & \rho(q+m+p-3) & \cdots & \rho(q+m-p+1) \end{vmatrix}$$

c'est-à-dire

$$(-1)^{p-1} \phi_p \begin{vmatrix} \rho(q+m-1) & \rho(q+m-2) & \cdots & \rho(q+m-p) \\ \rho(q+m) & \rho(q+m-1) & \cdots & \rho(q+m-p+1) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(q+m+p-2) & \rho(q+m+p-3) & \cdots & \rho(q+m-p+1) \end{vmatrix}$$

D'où

$$\Delta(p,q+m) = (-1)^{p-1} \phi_p \Delta(p,q+m-1) \qquad \textit{pour } m \geq 1 \qquad (2.1.13)$$

Remarque

Si $\Delta(p,q) = 0$ alors $\forall n,m \in \mathbb{N} \quad \Delta(p+n,q) = 0 \quad \textit{et} \quad \Delta(p,q+m) = 0$.

D. Identification d'un modèle ARMA stationnaire par la méthode du coin

L'énoncé du théorème du coin est le suivant :

Théorème du coin (*Beguin et al. (1980)*)
Un processus $\{y_t, t \in \mathbb{Z}\}$ stationnaire régulier est un processus ARMA(p, q) si et seulement si :
$i)$ $\Delta(k, i) = 0 \quad \forall i \geq q + 1 \quad \forall k \geq p + 1.$
$ii)$ $\Delta(k, q) \neq 0 \quad \forall k \geq p.$
$iii)$ $\Delta(p, i) \neq 0 \quad \forall i \geq q.$

Démonstration Nous donnerons la démonstration du théorème du coin dans le cas général (le cas d'un processus ARMA d-périodique) dans le chapitre 4.

Remarques
1- Si le processus $\{y_t, t \in \mathbb{Z}\}$ est un ARMA(p, q) ,alors les déterminants $\Delta(k, i)$ ne sont pas nuls sur les deux bords du coin c'est-à-dire $\Delta(p, q+m) \neq 0, \quad \forall m \geq 1$ et $\Delta(p + n, q) \neq 0, \quad \forall n \geq 1$.
2- Si le processus $\{y_t, t \in \mathbb{Z}\}$ est un ARMA(p, q) ,alors les déterminants $\Delta(k, i)$ sont nuls à l'intérieur du coin c'est à-dire $\Delta(k, i) = 0, \quad \forall i \geq q + 1$ et $\forall k \geq p + 1$.

$k \setminus i$	$\cdots$	$q-1$	q	$q+1$	$q+2$	$\cdots$
$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$p-1$	$\cdots$	$\Delta(p-1, q-1)$	$\Delta(p-1, q)$	$\Delta(p-1, q+1)$	$\Delta(p-1, q+2)$	$\cdots$
p	$\cdots$	$\Delta(p, q-1)$	$\Delta(p, q)$	X	X	$\cdots$
$p+1$	$\cdots$	$\Delta(p+1, q-1)$	X	0	0	$\cdots$
$p+2$	$\cdots$	$\Delta(p+2, q-1)$	X	0	0	$\cdots$
$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	

X : $\Delta(k, i) \neq 0$ sur les bords mais il est nul à l'intérieur du coin.

Tableau (2.1.1)

3- Le processus $\{y_t, t \in \mathbb{Z}\}$ a une représentation minimale MA(q) si et seulement si $\Delta(1, i) = \rho(i)$ pour $i \geq q + 1$ et $\Delta(1, q) = \rho(q) \neq 0$.(ceci est équivalent au théorème (2.1.2)).
4- Le processus $\{y_t, t \in \mathbb{Z}\}$ a une représentation minimale AR(p) si et seulement si $\Delta(k, 1) = 0$ pour $k \geq p+1$ et $\Delta(p, 1) \neq 0$. Ceci est équivalent au théorème (2.1.1)). Car le coefficient d'autocorrélation partielle est donné par

$$\phi_{k,k} = \frac{\det \Sigma_k^*}{\det \Sigma_k}$$

lorsque Σ_k est réguliere, Σ_k désigne la matrice d'autocorrélation d'ordre k et Σ_k^* est la matrice obtenue en remplaçant la dernière colonne de Σ_k par le vecteur $(\rho(1), \rho(2), ..., \rho(k))'$,d'où

$$\phi_{k,k} = \frac{\begin{vmatrix} \rho(0) & \rho(1) & \cdots & \rho(k-2) & \rho(1) \\ \rho(1) & \rho(0) & \cdots & \rho(k-3) & \rho(2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho(k-1) & \rho(k) & \cdots & \rho(1) & \rho(k) \end{vmatrix}}{\begin{vmatrix} \rho(0) & \rho(1) & \cdots & \rho(k-1) \\ \rho(1) & \rho(0) & \cdots & \rho(k-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(k-1) & \rho(k-2) & \cdots & \rho(0) \end{vmatrix}}$$

$$\phi_{k,k} = \frac{(-1)^{k-1}\Delta(k,1)}{\Delta(k,0)}.$$

Si $\Sigma(k,i)$ est une matrice régulière le déterminant $\Delta(k,i)$ peut être représenté par les trois fonctions θ, η et λ définies par :

Pour $k = 1, 2, ..., i = 0, 1, 2, ...$

$$\begin{aligned} \phi(k,i) &= \Sigma^{-1}(k,i)\rho(k,i), \\ \widetilde{\pi}(k,i) &= \Sigma^{-1}(k,i)r(k,i), \\ \theta(k,i) &= \rho(i+k+1) - \left(\widetilde{\phi}(k,i)\right)' \rho(k,i), \\ \eta(k,i) &= \rho(i-k-1) - (\pi(k,i))' r(k,i), \\ \lambda(k,i) &= \rho(i) - (\pi(k,i))' \rho(k,i). \end{aligned}$$

où nous notons par $\widetilde{x} = (x_n, ..., x_1)'$ pour n'importe quel vecteur $x = (x_1, ..., x_n)'$ et

$$\rho(k,i) = \begin{pmatrix} \rho(i+1) \\ \rho(i+2) \\ \vdots \\ \rho(i+k) \end{pmatrix}, \qquad r(k,i) = \begin{pmatrix} \rho(i-k) \\ \rho(i-k+1) \\ \vdots \\ \rho(i-1) \end{pmatrix}.$$

Pour $k = 0$ et $i = 0, 1, ...,$ nous supposons que

$$\begin{aligned} \theta(0,i) &= \rho(i+1), \\ \eta(0,i) &= \rho(i-1), \\ \lambda(0,i) &= \rho(i). \end{aligned}$$

Nous appelons $\{\theta(k,i)\}$, $\{\eta(k,i)\}$ et $\{\lambda(k,i)\}$ les trois fonctions θ, η et λ, respectivement. Nous dénotons les j^{eme} éléments de $\phi(k,i)$ et $\pi(k,i)$ respectivement par $\phi_{k,j}^{(i)}$ et $\pi_{k,j}^{(i)}$ pour tout $j = 1..., k$. En outre, posons $\forall (k,i) \;\; \phi_{k,0}^{(i)} = \pi_{k,0}^{(i)} = -1.$

Proposition (2.1.3)

Si $\Sigma(k,i)$ est une matrice régulière, alors nous avons :

1- $\Delta(k+1,i+1) = (-1)^k\theta(k,i)\Delta(k,i)$.

2- $\Delta(k+1,i) = \lambda(k,i)\Delta(k,i)$.

3- $\Delta(k+1,i-1) = (-1)^k\eta(k,i)\Delta(k,i)$.

Démonstration

1- Nous avons par définition

$$\Delta(k+1,i+1) = \begin{vmatrix} \rho(i+1) & \rho(i) & \cdots & \rho(i-k+1) \\ \rho(i+2) & \rho(i+1) & \cdots & \rho(i-k+2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(i+k+1) & \rho(i+k) & \cdots & \rho(i+1) \end{vmatrix}$$

$$\begin{aligned} \Delta(k+1,i+1) &= (-1)^{-k} \begin{vmatrix} \rho(i+k+1) & \rho(i+k) & \cdots & \rho(i+1) \\ \rho(i+1) & \rho(i) & \ddots & \rho(i-k+1) \\ \rho(i+2) & \rho(i+1) & \ddots & \rho(i-k+2) \\ \vdots & \ddots & \ddots & \vdots \\ \rho(i+k) & \rho(i+k-1) & \cdots & \rho(i) \end{vmatrix} \\ &= (-1)^{-k} \left| \rho(i+k+1)\Delta(k,i) - (v_1,v_2,...,v_k) \begin{pmatrix} \rho(i+1) \\ \rho(i+2) \\ \vdots \\ \rho(i+k) \end{pmatrix} \right| \\ &= (-1)^{-k} \left[\rho(i+k+1)\Delta(k,i) - V'\rho(k,i)\right] \\ &= (-1)^{-k} \left[\rho(i+k+1) - \frac{V'}{\Delta(k,i)}\rho(k,i)\right] \Delta(k,i) \end{aligned}$$

où $V = (v_1, v_2, ..., v_k)^t$ est un vecteur dont les composantes sont les sous déterminants de $\Delta(k+1,i+1)$.

Remarquons que

$$\widetilde{V} = \Delta(k,i)\Sigma^{-1}(k,i)\rho(k,i)$$

d'où

$$\Delta(k+1,i+1) = (-1)^k\theta(k,i)\Delta(k,i) \tag{2.1.14}$$

2- Nous avons

$$\Delta(k+1,i) = \begin{vmatrix} \rho(i) & \rho(i-1) & \cdots & \rho(i-k) \\ \rho(i+1) & \rho(i) & \cdots & \rho(i-k+1) \\ \vdots & \ddots & \ddots & \vdots \\ \rho(i+k) & \rho(i+k-1) & \cdots & \rho(i) \end{vmatrix}$$

$$
\begin{aligned}
\Delta(k+1,i) &= \rho(i)\Delta(k,i) - W'\rho(k,i) \\
&= \left[\rho(i) - \frac{W'}{\Delta(k,i)}\rho(k,i)\right]\Delta(k,i)
\end{aligned}
$$

où $W = (w_1, w_2, ..., w_k)^t$ est un vecteur dont les composantes sont les sous déterminants de $\Delta(k+1,i)$.

Remarquons que

$$\widetilde{W} = \Delta(k,i)\Sigma^{-1}(k,i)r(k,i)$$

d'où

$$\Delta(k+1,i) = \lambda(k,i)\Delta(k,i). \tag{2.1.15}$$

3- Nous avons

$$
\Delta(k+1,i-1) = \begin{vmatrix}
\rho(i-1) & \rho(i-2) & \cdots & \rho(i-k-1) \\
\rho(i) & \rho(i-1) & \cdots & \rho(i-k) \\
\vdots & \vdots & \ddots & \vdots \\
\rho(i+k-1) & \rho(i+k-2) & \cdots & \rho(i-1)
\end{vmatrix}
$$

$$
= (-1)^k \begin{vmatrix}
\rho(i+k-1) & \rho(i-1) & \rho(i-2) & \cdots & \rho(i-k) \\
\rho(i-k) & \rho(i) & \rho(i-1) & \cdots & \rho(i-k+1) \\
\rho(i-k+1) & \rho(i+k-1) & \rho(i+k-1) & \cdots & \rho(i+k-1) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\rho(i-1) & \rho(i+k-1) & \rho(i+k-2) & \cdots & \rho(i)
\end{vmatrix}
$$

$$
\begin{aligned}
&= (-1)^k\left[\rho(i+k-1)\Delta(k,i) - Z'r(k,i)\right] \\
&= (-1)^k\left[\rho(i+k-1) - \frac{Z'}{\Delta(k,i)}r(k,i)\right]\Delta(k,i)
\end{aligned}
$$

où $Z = (z_1, z_2, ..., z_k)'$ est vecteur dont les composantes sont les sous déterminants de $\Delta(k+1, i-1)$.

Remarquons que

$$\widetilde{Z} = \Delta(k,i)\Sigma^{-1}(k,i)r(k,i)$$

d'où

$$\Delta(k+1,i-1) = (-1)^k\eta(k,i)\Delta(k,i). \tag{2.1.16}$$

La proposition (2.1.3) montre que le déterminant Δ peut être calculé par l'algorithme simplifié de Trench-Zohar. Pham (1984) a présenté la relation suivante :

$$\Delta^2(k,i) = \Delta(k,i+1)\Delta(k,i-1) + \Delta(k+1,i)\Delta(k-1,i). \tag{2.1.17}$$

Il a suggeré d'utiliser cette relation pour calculer Δ. Nous pouvons facilement tirer la relation de Pham en utilisant la propriété (2.1.3).

E. Distribution asymptotique de Δ

Pour tester l'hypothèse nulle H_0,

$$H_0 : \Delta(i,k) = 0 \qquad \forall i \geq q+1 \qquad \forall k \geq p+1 \tag{2.1.18}$$

Choi (1990b) a présenté la distribution conjointe asymptotique de l'estimateur du déterminant Δ.

Théorème (2.1.5) (***Distribution asymptotique du déterminant*** Δ)
Soit $\{y_1, ..., y_N\}$ une N-réalisation d'un modèle ARMA (p,q) de la représentation (2.1.8). Alors nous avons :

1- Si $k, j = p, p+1, ...,$ alors $\sqrt{N}\widehat{\Delta}(k+1, q+1)$ et $\sqrt{N}\widehat{\Delta}(j+1, q+1)$ sont asymptotiquement normalement distribuées avec des moyennes nulles et

$$\begin{aligned} a_{k,j} &= \lim_{N \longrightarrow +\infty} cov\left\{\sqrt{N}\widehat{\Delta}(k+1,q+1), \sqrt{N}\widehat{\Delta}(j+1,q+1)\right\} \\ &= (-1)^{k+j}\Delta(k,q)\Delta(j,q)\frac{1}{\gamma^2(0)}\sum_{m=-q}^{q}\gamma_z(m)\gamma_z(m+k-j), \end{aligned}$$

où

$$\gamma_z(m) = \sum_{r=0}^{p}\sum_{s=0}^{p}\phi_r\phi_s\gamma(m+r-s).$$

2- Si $i, j = q, q+1, ...,$ alors $\sqrt{N}\widehat{\Delta}(p+1, i+1)$ et $\sqrt{N}\widehat{\Delta}(p+1, j+1)$ sont asymptotiquement normalement distribuées avec des moyennes nulles et

$$\begin{aligned} b_{k,j} &= \lim_{N \longrightarrow +\infty} cov\left\{\sqrt{N}\widehat{\Delta}(p+1,i+1), \sqrt{N}\widehat{\Delta}(p+1,j+1)\right\} \\ &= \Delta(p,i+1)\Delta(p,j+1)\frac{1}{\gamma^2(0)\phi_p^2}\sum_{m=-q}^{q}\gamma_z(m)\gamma_z(m+i-j). \end{aligned}$$

3- Si $k = p, p+1, ..., i = q, q+1, ...,$ alors $\sqrt{N}\widehat{\Delta}(k+1, q+1)$ et $\sqrt{N}\widehat{\Delta}(p+1, i+1)$ sont asymptotiquement normalement distribuées avec des moyennes nulles et

$$\begin{aligned} c_{k,j} &= \lim_{N \longrightarrow +\infty} cov\left\{\sqrt{N}\widehat{\Delta}(k+1,q+1), \sqrt{N}\widehat{\Delta}(p+1,i+1)\right\} \\ &= (-1)^{k-1}\Delta(k,q)\Delta(p,i+1)\frac{1}{\gamma^2(0)\phi_p}\sum_{m=-q}^{q}\gamma_z(m)\gamma_z(m+k-p-i+q). \end{aligned}$$

Démonstration Voir Choi (1990b).

Jusqu'à présent nous avons vu qu'à partir de la fonction d'autocorrélation et les équations de Yule-Walker, nous pouvons caractériser un modèle ARMA stationnaire par le déterminant de la matrice $\Sigma(k,i)$. En outre, Choi (1991a) a présenté une deuxième caractérisation d'un modèle ARMA stationnaire par les trois fonctions θ, λ et η définies précédemment.

2.1.4 Méthode de 3-Pattern

A. Notations

Nous définissons la fonction $\Phi_k^{(i)}$ et les deux ensembles d'indices $I_{r,s}$, $J_{r,s}$ comme suit :

$$\begin{aligned}
\Phi_k^{(i)}(z) &= -\phi_{k,0}^{(i)} - \phi_{k,1}^{(i)} z - \cdots - \phi_{k,k}^{(i)} z^k, \qquad \phi_{k,0}^{(i)} = -1 \\
I_{r,s} &= \{(k,s)/ \quad k \geqslant r\} \cup \{(r,i)/i \geqslant s\}. \\
J_{r,s} &= \{(k,i)/ \quad k \geqslant r, \quad i \geqslant s\}.
\end{aligned}$$

B. Caractérisation d'un modèle ARMA par les trois fonctions θ, η et λ.

Choi (1991a) a dérivé des caractérisations d'un modèle ARMA stationnaire par les trois fonctions en utilisant les équations de Yule-Walker, qui sont utiles pour l'identification d'un modèle ARMA. Cependant, nous les dériverons plus intuitivement.

1- Nous savons que :

$$\rho(j) - \sum_{h=1}^{k} \phi_{k,h}^{(i)} \, \rho(j-h) = 0 \ , \qquad j = i+1, i+2, ..., i+k$$

qui est équivalent à

$$\phi(k,i) = \Sigma^{-1}(k,i)\rho(k,i) \tag{2.1.19}$$

Nous avons également

$$\theta(k,i) = \rho(i+k+1) - \left(\widetilde{\phi}(k,i)\right)' \rho(k,i) \tag{2.1.20}$$

$$\begin{aligned}
&= \rho(i+k+1) - \left(\phi_{k,k}^{(i)}, \phi_{k,k-1}^{(i)}, ..., \phi_{k,1}^{(i)}\right) \begin{pmatrix} \rho(i+1) \\ \rho(i+2) \\ \vdots \\ \rho(i+k) \end{pmatrix} \\
&= \rho(i+k+1) - \sum_{j=1}^{k} \phi_{k,j}^{(i)} \rho(i+k-j+1).
\end{aligned}$$

D'autre part, nous avons

$$\begin{aligned}
cov\left(y_t, \phi_k^{(i)}(B) y_{t+k+i+1}\right) &= cov\left(y_t, \sum_{j=0}^{k} \phi_{k,j}^{(i)} y_{t+k+i-j+1}\right) \\
&= cov(y_t, y_{t+k+i+1}) - \sum_{j=1}^{k} \phi_{k,j}^{(i)} cov(y_t, y_{t+k+i-j+1}) \\
&= \gamma(k+i+1) - \sum_{j=1}^{k} \phi_{k,j}^{(i)} \gamma(k+i-j+1)
\end{aligned}$$

D'où

$$\gamma(0)\theta(k,i) = cov\left(y_t, \Phi_k^{(i)}(B)y_{t+k+i+1}\right) \tag{2.1.21}$$

Si $(k,i) \in I_{p,q}$, nous avons alors,

$$cov\left(y_t, \Phi(B)y_{t+k+i+1}\right) = cov\left(y_t, \Theta(B)\varepsilon_{t+k+i+1}\right)$$

qui est nulle sous l'hypothèse de stationnarité.

Théorème (2.1.6) (*The θ Pattern*) (*Choi (1991)*)
Un processus stochastique a une représentation (2.1.8) ARMA(p,q) si et seulement si p et q sont les plus petits nombres entiers satisfont

$$\theta(k,i) = 0, \qquad (k,i) \in I_{p,q}.$$

Remarque Nous pouvons facilement prouver que si un processus stochastique a une représentation (2.1.8) ARMA(p,q), alors

$$\theta(k,i) = 0, \qquad (k,i) \in J_{p,q}.$$

D'où

$k \setminus i$	$\cdots$	$q-1$	q	$q+1$	$\cdots$
$\vdots$		$\vdots$	$\vdots$	$\vdots$	
$p-1$	$\cdots$	$\theta(p-1,q-1)$	$\theta(p-1,q)$	$\theta(p-1,q+11)$	$\cdots$
p	$\cdots$	$\theta(p,q-1)$	0	0	$\cdots$
$p+1$	$\cdots$	$\theta(p+1,q-1)$	0	0	$\cdots$
$p+2$	$\cdots$	$\theta(p+2,q-1)$	0	0	$\cdots$
$\vdots$		$\vdots$	$\vdots$	$\vdots$	

Tableau (2.1.2)

2- Nous montrons que

$$\lambda(k,i) = \rho(i) - \left(\widetilde{\phi}(k,i)\right)' r(k,i) \tag{2.1.22}$$

C'est-à-dire que

$$\begin{aligned}
\pi^t(k,i)\rho(k,i) &= \left(\widetilde{\phi}(k,i)\right)' r(k,i) \\
&\Updownarrow \\
(\pi(k,i))'\,\Sigma(k,i)\Sigma^{-1}(k,i)\rho(k,i) &= \left(\widetilde{\phi}(k,i)\right)' \Sigma(k,i)\Sigma^{-1}(k,i)r(k,i) \\
&\Updownarrow \\
(\pi(k,i))'\,\Sigma(k,i)\phi(k,i) &= \left(\widetilde{\phi}(k,i)\right)' \Sigma(k,i)\widetilde{\pi}(k,i)
\end{aligned}$$

Nous avons

$$\begin{aligned}
(\pi(k,i))'\,\Sigma(k,i)\phi(k,i) &= \left(\pi_{k,1}^{(i)},...,\pi_{k,k}^{(i)}\right)\Sigma(k,i)\begin{pmatrix}\phi_{k,1}^{(i)}\\ \vdots\\ \phi_{k,k}^{(i)}\end{pmatrix}\\
&= \sum_{h=1}^{k}\sum_{j=1}^{k}\pi_{k,j}^{(i)}\phi_{k,h}^{(i)}\rho(i+j-h)\\
&= \sum_{j=1}^{k}\sum_{h=1}^{k}\pi_{k,j}^{(i)}\phi_{k,h}^{(i)}\rho(i+j-h)\\
&= \left(\phi_{k,k}^{(i)},...,\phi_{k,1}^{(i)}\right)\Sigma(k,i)\begin{pmatrix}\pi_{k,k}^{(i)}\\ \vdots\\ \pi_{k,1}^{(i)}\end{pmatrix}\\
&= \left(\widetilde{\phi}(k,i)\right)'\Sigma(k,i)\widetilde{\pi}(k,i)
\end{aligned}$$

D'où

$$\begin{aligned}
\lambda(k,i) &= \rho(i)-\left(\widetilde{\phi}(k,i)\right)'r(k,i)\\
&= \rho(i)-\left(\phi_{k,k}^{(i)},...,\phi_{k,1}^{(i)}\right)\begin{pmatrix}\rho(i-k)\\ \vdots\\ \rho(i-1)\end{pmatrix}\\
&= \rho(i)-\sum_{j=1}^{k}\phi_{k,j}^{(i)}\rho(i-j)
\end{aligned}$$

qui implique que

$$\gamma(0)\lambda(k,i)=cov\left(y_t,\Phi_k^{(i)}(B)y_{t+i}\right) \tag{2.1.23}$$

Si $k=p$ et $i>q$, alors

$$\begin{aligned}
\gamma(0)\lambda(p,i) &= cov\left(y_t,\Phi(B)y_{t+i}\right)\\
&= cov\left(y_t,\Theta(B)\varepsilon_{t+i}\right)
\end{aligned}$$

qui est nulle.

Si $k>p$ et $i=q$, alors

$$\begin{aligned}
\gamma(0)\lambda(k,q) &= cov\left(y_t,\Phi(B)y_{t+q}\right)\\
&= cov\left(y_t,\Theta(B)y_{t+q}\right)\\
&= cov\left(y_t,-\theta_q\varepsilon_t\right)\\
&= \theta_0\theta_q\sigma^2.
\end{aligned}$$

Théorème (2.1.7) (***The λ Pattern***) (***Choi (1991)***)

Un processus stochastique a une représentation (2.1.8) ARMA(p, q) si et seulement si p et q sont les plus petits nombres entiers satisfont

$$\begin{aligned}\lambda(p, q+1) &= \lambda(p, q+2) = \cdots = 0, \\ \lambda(p, q) &= \lambda(p+1, q) = \cdots = \frac{\theta_0 \theta_q \sigma^2}{\gamma(0)}.\end{aligned}$$

Si le processus fondamental est d'un modèle autoregressif pur, c'est-à-dire, $q = 0$, alors

$$\lambda(p) = \lambda(p+1) = \cdots = \frac{\sigma^2}{\gamma(0)}.$$

Nous pouvons estimer la variance σ^2 du bruit blans comme suit :

$$\widehat{\sigma}_k^2 = \widehat{\gamma}(0)\widehat{\lambda}(k), \qquad k = p, p+1, ...$$

Remarque

Nous remarquons que si un processus stochastique a une représentation (2.1.8) ARMA(p, q), alors

1. Si $i > q$, alors $\lambda(p, i) = 0$.
2. Si $k \geqslant p$, alors $\lambda(k, q) = \frac{\theta_0 \theta_q \sigma^2}{\gamma(0)}$.
3. Si $s > r \geqslant 0$, alors $\lambda(p+r, q+s) = 0$.

D'où

$k \setminus i$	$\cdots$	q	$q+1$	$q+2$	$\cdots$
$\vdots$		$\vdots$	$\vdots$	$\vdots$	
$p-1$	$\cdots$	$\lambda(p-1, q)$	$\lambda(p-1, q+1)$	$\lambda(p-1, q+2)$	$\cdots$
p	$\cdots$	$\lambda(p, q)$	0	0	$\cdots$
$p+1$	$\cdots$	$\lambda(p+1, q)$	X	0	$\cdots$
$p+2$	$\cdots$	$\lambda(p+2, q)$	X	X	$\cdots$
$\vdots$		$\vdots$	$\vdots$	$\vdots$	

X : $\lambda(i, k)$ est non nulle.

Tableau(2.1.3)

Considérons le modèle ARMA(p, q) associé au modèle de la représentation (2.1.8) suivant :

$$\pi(B^{-1})z_t = \theta(B^{-1})\varepsilon_t, \tag{2.1.24}$$

où

$$\pi(z) = \frac{1}{\phi_p} z^p \phi(z^{-1}) = -\pi_0 - \pi_1 z - \cdots - \pi_p z^p, \qquad \pi_0 = -1.$$

Si les densités spectrales de $\{y_t\}$ et $\{z_t\}$ sont dénotées par $S_y(v)$ et $S_z(v)$, respectivement, alors

$$S_z(v) = \phi_p^2 S_y(v).$$

La relation de Wiener-Khinchine implique que les deux processus ont la même fonction d'autocorrélation (ACRF) $\{\rho(j)\}$. Puisque toutes les racines de $\phi(z) = 0$ sont en dehors du cercle d'unité, toutes les racines de $\pi(z) = 0$ sont à l'intérieur du cercle d'unité. Ainsi, l'ACRF satisfait les équations backward de Yule-Walker

$$\rho(j) = \pi_1\rho(j+1) + \cdots + \pi_p\rho(j+p), \qquad j = q-p+1, q-p+2, ...$$

qui donnent

$$\widetilde{\pi}(p,q) = \Sigma^{-1}(p,q)r(p,q). \tag{2.1.25}$$

Par conséquent, $\pi_{p,j}^{(q)} = \pi_j$ pour $j = 1..., p$. Les équations backward de Yule-Walker impliquent le théorème suivant.

Théorème (2.1.8) (*The* η *Pattern*) (*Choi (1991)*)
Si un processus stochastique a une représentation (2.1.8) ARMA(p,q), alors nous avons

1. $\eta(p-1,i) \neq 0, \qquad i = q+1, q+2, ...$
2. $\eta(p,q+1) = -\lambda(p,q)/\phi_p$.
3. $\eta(p,q+2) = \eta(p,q+3) = \cdots = 0$.

Remarque Nous constatons que si un processus stochastique a une représentation (2.1.8) ARMA(p,q), alors nous avons

1. Si $r \geqslant 0$, alors $\eta(p+r, q+1+r) = -\lambda(p,q)/\phi_p$.
2.Si $r > s \geqslant 0$, alors $\eta(p+r, q+1+s) = \pm\infty$.
3.Si $s > r \geqslant 0$, alors $\eta(p+r, q+1+s) = 0$.

Donc

$k \setminus i$	$\cdots$	$q+1$	$q+2$	$q+3$	$\cdots$
$\vdots$		$\vdots$	$\vdots$	$\vdots$	
$p-1$	$\cdots$	$\eta(p-1,q+1)$	$\eta(p-1,q+2)$	$\eta(p-1,q+3)$	$\cdots$
p	$\cdots$	$\eta(p,q+1)$	0	0	$\cdots$
$p+1$	$\cdots$	∞	$\eta(p+1,q+2)$	0	$\cdots$
$p+2$	$\cdots$	∞	∞	$\eta(p+2,q+3)$	$\cdots$
$\vdots$		$\vdots$	$\vdots$	$\vdots$	

Tableau (2.1.4)

C. Distribution asymptotique de θ

La constatation visuelle des $\widehat{\theta}$, $\widehat{\eta}$ et $\widehat{\lambda}$ peut donner une réponse pour choisir les ordres du processus ARMA. Cependant, pour déterminer statistiquement les ordres en employant les trois fonctions, nous avons besoin de leurs distributions asymptotiques. Choi (1990c) a présenté leurs distributions asymptotiques. Nous présenterons dans le paragraphe suivant seulement la distribution asymptotique de θ.

Théorème (2.1.9) (*Distribution asymptotique de θ*) (*Choi (1990)*)

Soit $y_1, ..., y_N$ une N-réalisation d'un modèle ARMA(p, q) a une représentation (2.1.8). Si $(k, i) \in I_{p,q}$, alors les $\sqrt{N}\widehat{\theta}(k, i)$ sont asymptotiquement normalement distribuées avec des moyennes nulles et covariances

$$\lim_{N\to\infty} cov\left(\sqrt{N}\widehat{\theta}(k,i), \sqrt{N}\widehat{\theta}(j,l)\right) =$$
$$\tfrac{\sigma^2}{\gamma^2(0)} \textstyle\sum_{r=0}^{p} \sum_{s=0}^{q} \sum_{m=0}^{p} \sum_{n=0}^{q} \phi_r\theta_s\phi_m\theta_n\gamma(k+i-j-l-r-s+m+n).$$

Démonstration Voir Choi (1990c).

D. Test statistique

Le théorème (2.1.4) implique que les hypothèses suivantes sont équivalentes.

$$\begin{array}{lcl} H_K & : & \text{les vrais ordres sont } p \text{ et } q. \\ H_E & : & \theta(p, q-1) \neq 0, \theta(p-1, q-1) \neq 0 \text{ et } \theta(p, i) = 0, \text{ pour } i = q, q+1, ... \\ H_J & : & \theta(p-1, q) \neq 0, \theta(p, q-1) \neq 0 \text{ et } \theta(k, q) = 0, \text{ pour } k = p, p+1, ... \end{array}$$

En utilisant la distribution asymptotique du théorème (2.1.9), nous pouvons dériver des statistiques pour tester les hypothèses H_E et H_J. Soient

$$\begin{aligned} h_r(k,i) &= \left(\widehat{\theta}(k,i), \widehat{\theta}(k,i+1), ..., \widehat{\theta}(k,i+r-1)\right)', \\ v_r(k,i) &= \left(\widehat{\theta}(k,i), \widehat{\theta}(k+1,i), ..., \widehat{\theta}(k+r-1,i)\right)', \end{aligned}$$

et soit $W_r(k, i)$ une matrice symétrique de dimension r dont le (α, β) élément est

$$\frac{1}{\widehat{\gamma}^2(0)} \sum_{m=-i}^{i} \hat{\gamma}_z(m)\hat{\gamma}_z(m+\alpha-\beta)$$

où

$$\widehat{\gamma}_z(m) = \sum_{r=0}^{k} \sum_{s=0}^{k} \widehat{\phi}_{k,r}^{(i)} \widehat{\phi}_{k,s}^{(i)} \widehat{\gamma}(m+r-s).$$

Choi (1990d) a défini deux statistiques E et J comme suit :

$$E_r(k,i) = N\left(h_r(k,i)\right)^t W_r^{-1}(k,i) h_r(k,i), \tag{2.1.26}$$

$$J_r(k,i) = N\left(v_r(k,i)\right)^t W_r^{-1}(k,i) v_r(k,i). \tag{2.1.27}$$

Le théorème (2.1.12) nous donne leurs distributions asymptotiques comme suit :

Corollaire (2.1.1) (*Choi (1990)*)

Soit $y_1, ..., y_N$ une N-réalisation d'un modèle ARMA(p,q) a une représentation (2.1.8).

Si $k = p$ et $i = q, q+1...$, alors $E_r(k,i)$ a une distribution asymptotique χ_r^2.

Si $k = p, p+1...$ et $i = q$, alors $J_r(k,i)$ a une distribution asymptotique χ_r^2.

Démonstration Voir Choi (1990d).

Théoriquement, nous devrions employer $E_\infty(k,i)$ et $J_\infty(k,i)$ pour tester les hypothèses H_E et H_J. Cependant, il y a seulement un nombre fini des observations. D'ailleurs, parce que le comportement de $\widehat{\rho}_j$ devient plus erratique lorsque $|j|$ augmente, les distributions asymptotiques de E_r et J_r dans le corollaire (2.1.1) peuvent être différentes de leurs estimateurs empiriques si r est assez grand. Ainsi, nous testerons les hypothèses modifiées suivantes

$$\begin{aligned} H_E(r) &: \theta(p,q-1) \neq 0, \theta(p-1,q) \neq 0 \text{ et } \theta(p,i) = 0, \text{ pour } i = q, ..., q+r-1 \\ H_J(r) &: \theta(p-1,q) \neq 0, \theta(p,q-1) \neq 0 \text{ et } \theta(k,q) = 0, \text{ pour } k = p, ..., p+r-1 \end{aligned}$$

avec r modéré.

Choi (1990d) a conclu sur la base des résultats de simulation qu'il est suffisamment bon pour choisir les ordres d'un modèle ARMA de tester $H_E(3)$ et $H_J(3)$.

En résumé, la détermination des odres d'un processus ARMA en utilisant la méthode 3-patterns et les deux statistiques χ^2 est donnée comme suit.

1°. Nous estimons θ, λ et η. Puis, nous cherchons les paires possibles (k,i) des ordres en se basant sur les théorèmes (2.1.6)-(2.1.11).

2°. Nous choisissons les plus petits k et i parmi toutes les paires possibles tels que :

$$\begin{aligned} E_r(k,i) &< \chi_r^2(1-\alpha), \\ J_r(k,i) &< \chi_r^2(1-\alpha), \end{aligned}$$

où $\chi_r^2(1-\alpha)$ est le quantile d'ordre $(1-\alpha)$ de la loi χ_r^2.

3°. Cous confirmons notre choix, en vérifiant que

$$\begin{aligned} E_r(k,n) &< \chi_r^2(1-\alpha), \quad n = i+1, i+2, ..., \\ J_r(m,i) &< \chi_r^2(1-\alpha), \quad m = k+1, k+2, ..., \end{aligned}$$

et que $E_r(k,i-1)$ et $J_r(k-1,i)$ sont plus grands que 0.

Depuis le début des années 70, quelques méthodes d'identification ont été proposées. Elles doivent choisir les ordres k et i qui minimisent la fonction suivante :

$$P(k,i) = \log \overset{\vee}{\sigma}{}^2_{k,i} + (k+i)\frac{C(N)}{N}, \tag{2.2.1}$$

où $\overset{\vee}{\sigma}{}^2_{k,i}$ est un estimateur de la variance du bruit blanc obtenu en adaptant le modèle ARMA(k,i) aux observations. Puisque $\overset{\vee}{\sigma}{}^2_{k,i}$diminue lorsque les ordres augmentent, donc $\overset{\vee}{\sigma}{}^2_{k,i}$ ne peut pas être un bon critère pour choisir les ordres. Si les ordres augmentent, le biais du modèle estimé diminuera tandis que la variance augmente. Par conséquent, nous devrions faire un compromis entre elles. Pour ce là nous ajoutons le terme de pénalité $(k+i)\frac{C(N)}{N}$ dans le critère de sélection des modèles.

Il est normal d'exiger que les ordres sélectionés augmentent lorsque N augmente. Dans la pratique, aucune borne supérieure des ordres k et i ne pourra être prescrite car le minimum de $P(k,i)$ est s'atteint pour des valeurs tout à fait petites de k et i.

2.2 Méthode de l'erreur de prédiction finale (FPE)

Considérons une N-réalisation $\{y_1, .., y_N\}$ d'un processus AR(p) qui a une représentation comme suit :

$$y_t = \sum_{i=1}^{p} \phi_i \ y_{t-i} + \varepsilon_t.$$

Soit $\{x_t, t \in \mathbb{Z}\}$ un autre processus AR(p) indépendant de $\{y_t, t \in \mathbb{Z}\}$ mais de même structure statistique que $\{y_t, t \in \mathbb{Z}\}$, c'est -à-dire,

$$x_t = \sum_{i=1}^{p} \phi_i \ y_{t-i} + \varepsilon'_t,$$

où le processus bruit blanc $\{\varepsilon'_t, t \in \mathbb{Z}\}$ a la même distribution que $\{\varepsilon_t, t \in \mathbb{Z}\}$ mais les deux processus sont indépendant l'un de l'autre. Alors, le prédicteur de x_{N+1} est donné par :

$$\widehat{x}_{N+1} = \widehat{\phi}_{p,1} x_N + \cdots + \widehat{\phi}_{p,p} x_{N-p+1} \tag{2.2.2}$$

où $\widehat{\phi}_{p,1}, ..., \widehat{\phi}_{p,p}$ sont les estimateurs de Yule-Walker obtenues en adaptant le modèle AR(p) aux observations $\{y_1, .., y_N\}$.

En utilisant le théorème (1.5.1) (la distribution asymptotique des estimateurs de YW) et l'indépendance de $\{x_t, t \in \mathbb{Z}\}$ et $\left\{\widehat{\phi}_{p,1}, ..., \widehat{\phi}_{p,p}\right\}$, Akaike (1969) a prouvé que, pour N assez grand, l'erreur quadratique de prediction est donné par :

$$E\left(\widehat{x}_{N+1}-x_{N+1}\right)^2 \simeq \sigma^2\left(1+\frac{p}{N}\right). \tag{2.2.3}$$

D'autre part, la consistence de l'ACVF de l'échantillon et du théorème (1.5.1) nous donnent, pour N assez grand

$$E\left(\widehat{\sigma}_p^2\right) \simeq \sigma^2\left(1-\frac{p}{N}\right). \tag{2.2.4}$$

Ainsi, $\widehat{\sigma}_p^2\left(1+\frac{p}{N}\right)^{-1}$ est un estimateur moins biaisé que $\widehat{\sigma}_p^2$. D'où, l'estimateur asymptotique correspondant de l'erreur moyenne quadratique de $\hat{x}_{N+1}$est

$$\widehat{\sigma}_p^2\left(1-\frac{p}{N}\right)^{-1}\left(1+\frac{p}{N}\right),$$

qui est asymptotiquement équivalent à $\widehat{\sigma}_p^2\left(1+\frac{2p}{N}\right)$.

Akaike (1969) a défini l'erreur de prédiction finale (FPE) par :

$$FPE(p)=\widehat{\sigma}_p^2\left(1+\frac{2p}{N}\right) \tag{2.2.5}$$

qui est un estimateur asymptotiquement sans biais de l'erreur moyenne quadratique du prédicteur $\hat{x}_{N+1}$. Si la moyenne du processus est inconnue Akaike (1970) a mentionné que le FPE devrait être remplacée par :

$$FPE(p)=\widehat{\sigma}_p^2\left(1+\frac{2(p+1)}{N}\right) \tag{2.2.6}$$

Remarque Puisqu'il est raisonnable de choisir un prédicteur qui réduit au minimum l'erreur moyenne quadratique de prévision, il est signicatif d'estimer l'ordre en réduisant au minimum le FPE(k). Ce dernier s'appellera l'estimateur du minimum de l'erreur de prédiction finale (MFPEE).

$$\widehat{p}=MFPEE=\arg\left\{\min_k\left[\widehat{\sigma}_k^2\left(1+\frac{2(k+1)}{N}\right)\right]\right\}. \tag{2.2.7}$$

Le FPE(k) se compose de deux parties. La première, $\widehat{\sigma}_k^2$ correspond à la variance du meilleur prédicteur linéaire pour un k donné. La deuxième, $\widehat{\sigma}_k^2\frac{2(p+1)}{N}$ qui est due aux écarts statistiques entre $\left\{\widehat{\phi}_{k,1},...,\widehat{\phi}_{k,k}\right\}$ et les vrais paramètres. En général, pour N fixé, la première diminue et la deuxième croit lorsque k augmente. Ainsi, la minimisation du FPE(k) est une technique pour obtenir un ordre optimal, car :

- Si une grande valeur de k est adoptée, alors la FPE(k) tend à être grand.

- Si l'ordre choisi est inférieur au vrai ordre du processus, alors le biais de $\widehat{\sigma}_k^2$ tend à être grand.

Ainsi, la téchnique qui cherche un ordre k qui minimise la FPE(k) équilibre entre la croissance de la FPE(k) et la croissance du biais de $\widehat{\sigma}_k^2$.

Dans la pratique il est impossible de calculer le FPE pour touts les ordres. Habituellement, nous choisissons un nombre entier K comme borne supérieure des ordres possibles. Il est claire que K devrait être assez grand pour contenir le vrai ordre.

Beaucoup d'analystes de serie chronologique comme Akaike (1970), Kashyap (1980) et Hannan(1982) ont montré que la méthode de l'erreur de prédiction finale a une tendance de sur-estimer l'ordre d'un modèle AR.

Théorème (2.2.1)(*Hannan (1982)*)

Si $\{y_t, t \in \mathbb{Z}\}$ est un modèle AR(p), alors

$$\lim_{N \to \infty} \Pr\{FPE(k) > FPE(p)\} = \begin{cases} 1 & k < p \\ \Pr\left\{\chi^2_{k-p} > 2(k-p)\right\}, & k \geq p. \end{cases}$$

où χ_n^2 est une variable aléatoire Chi-deux de n degré de liberté.

Demonstration : Voir Hannan (1982).

2.3 Critère de l'information d'Akaike (AIC)

Dans cette section, nous présenterons en revue, un critère très connu pour la sélection des ordres d'un modèle ARMA stationnaire et qui est basé sur la théorie de l'information. Il s'agit de choisir le modèle en se basant sur une mesure de l'écart entre la vraie lois (inconnue) et celle du modèle proposé. La mesure habituellement utilisée est l'information de Kullback.

Nous choisissons le modèle donnant la plus petite valeur de l'estimation de l'information de Kullback.

2.3.1 Quantité d'information de Kullback-Leibler

Kullback et Leibler (1951) ont proposé une quantité d'information théorique pour mesurer la différence entre deux fonctions de densité dans le sens statistique.

Définition (2.3.1) Soient f et g deux fonctions de densité. Si l'intégral

$$I(f;g) = \int f(x)\log(\frac{f(x)}{g(x)})dx \tag{2.3.1}$$

existe, la quantité $I(f;g)$ s'appelle la quantité d'information de Kullback-Leibler de la densité f par apport à la densité g. Beaucoup de théorèmes statistiques ont prouvé que $I(f;g)$ est une mesure de distinction entre deux fonctions de densité.

2.3.2 Critère de l'information d'Akaike (AIC)

Akaike (1972) a présenté une méthode d'identification en utilisant la quantité d'information de Kullback-Leibler de même que la consistance et la normalité asymptotique des estimateurs obtenus par la méthode du maximum de vraisemblance.

Soit $x_1, ..., x_N$ des variables aléatoires indépendantes et identiquement distribuées de densité $f\left(x/\theta^{\circ}\right)$, où $\theta^{\circ} \in \mathbb{R}^K$ est un paramètre vectoriel fixe mais inconnu. Le problème est d'estimer θ tel que $f(x/\theta)$ sera proche de $f(x/\theta^{\circ})$ au sens statistique. Une des méthodes raisonnables d'estimation doit réduire au minimum la quantité d'information de Kullback-Leibler

$$I(\theta^{\circ}, \theta) = \int f(x/\theta^{\circ}) \log\left(\frac{f(x/\theta^{\circ})}{f(x/\theta)}\right) dx = E_{\theta^{\circ}}\left[\log\left(\frac{f(x/\theta^{\circ})}{f(x/\theta)}\right)\right]. \tag{2.3.2}$$

Soit $\mathbb{R}^k = \{(a_1, ..., a_k, 0, ...0)', a_i \in \mathbb{R}, i = 1, ..., k\}$ un sous-espace de $\mathbb{R}^K$ pour $k = 1, ..., K$ et soit $\widetilde{\theta}_k$ l'estimateur de θ° obtenu par la méthode du maximum de vraisemblance restrictif à R^k. Soit J la matrice $K \times K$ d'information de Fisher. Alors, nous pouvons construire un espace de Hilbert $\left\{\theta, \theta \in R^K\right\}$ avec le produit scalaire et la norme

$$\langle\alpha, \beta\rangle_J = \alpha' J \beta, \qquad \|\alpha\|_J = \sqrt{\langle\alpha, \alpha\rangle_J}$$

Nous savons que le développement de Taylor de $f\left(x/\widetilde{\theta}_k\right)$ au point $\theta = \theta^{\circ}$ donne :

$$I(\theta, \widetilde{\theta}_k) \simeq \frac{1}{2}\left\|\widetilde{\theta}_k - \theta^{\circ}\right\|_J^2 = \frac{1}{2}\left\|\widetilde{\theta}_k - \theta_k^{\circ}\right\|_J^2 + \frac{1}{2}\left\|\theta_k^{\circ} - \theta^{\circ}\right\|_J^2 \tag{2.3.3}$$

où θ_k° est la projection de θ° sur R^k.

La normalité asymptotique des estimateurs ML implique :

$$E\left\{I\left(\theta, \widetilde{\theta}_k\right)\right\} \simeq \frac{1}{2}\left\|\theta_k^{\circ} - \theta^{\circ}\right\|_J^2 + \frac{k}{2N}. \tag{2.3.4}$$

Soit $L(k, K)$ la statistique du rapport de log-vraisemblance pour le test d'hypothèse

$$H_0 : \theta^{\circ} \in \mathbb{R}^k \qquad \textit{contre} \qquad H_1 : \theta^{\circ} \in \mathbb{R}^K - \mathbb{R}^k.$$

Le théorème de Cochran et la normalité asymptotique des estimateurs ML impliquent que la distribution asymptotique de $(-2)L(k, K)$ est une χ^2_{K-k}. De plus,

$$\frac{1}{N}\{(-2)L(k, K) + 2k - K\}$$

est un estimateur asymptotiquement sans-biais de $E\left(I\left(\theta^{\circ}, \widetilde{\theta}_k\right)\right)$ qui est égal à

$$\frac{1}{2N}\left\{2\sum_{t=1}^{N} \log f\left(x_t/\widetilde{\theta}_K\right) - K\right\} + \frac{1}{2N}\left\{-2\sum_{t=1}^{N} \log f\left(x_t/\widetilde{\theta}_k\right) + 2k\right\}.$$

Puisque la première partie est independante de k, alors l'ordre qui minimise l'estimateur de $E\left(I\left(\theta^{\circ}, \widetilde{\theta}_k\right)\right)$ est le même qui minimise

$$AIC^*(k) = -2\sum_{t=1}^{N} \log f\left(x_t/\widetilde{\theta}_k\right) + 2k \tag{2.3.5}$$

Ce dernier s'appelle le critère d'information d'Akaike (AIC) et l'ordre qui minimise l'AIC est appelé l'estimateur du minimum d'AIC (MAICE = *minimum AIC estimate*)

$$\widehat{p} = MAICE = \arg\left\{\min_k \left[-2\sum_{t=1}^{N} \log f\left(x_t/\widetilde{\theta}_k\right) + 2k\right]\right\} \tag{2.3.6}$$

Le MAICE a été utilisé pour choisir les modèles optimaux dans divers domaines de statistique comprenant l'analyse de série chronologique.

Le critère d'information d'Akaike (AIC^*) a été limité à des variables aléatoires indépendantes et identiquement distribuées. Akaike (1973) a commenté que la même procédure peut être étendue pour couvrir le cas des processus de Markov d'ordre fini par l'approche de Billingsley (1961). Après cette idée, Ogata (1980) a dérivé le critère AIC pour des modèles AR. En 1989, Hurvich et Tsai ont présenté un estimateur affine basé sur la quantité d'information de Kullback-Leibler pour des processus AR et ils ont proposé d'utiliser le critère AIC modifié

$$AIC(k) = \log \widetilde{\sigma}_k^2 + \frac{N+k}{N-k-2}. \tag{2.3.7}$$

Le critère AIC pour des processus ARMA est défini par :

$$AIC(k,i) = \log \widetilde{\sigma}_{k,i}^2 + \frac{2(k+i)}{N} \tag{2.3.8}$$

et

$$(\widehat{p}, \widehat{q}) = MAICE = \arg\left\{\min_{(k,i)} \left[\log \widetilde{\sigma}_{k,i}^2 + \frac{2(k+i)}{N}\right]\right\}, \tag{2.3.9}$$

où $\widetilde{\sigma}_{k,i}^2$ est l'estimateur de la variance du bruit blanc obtenu par la méthode du maximum de vraisemblance. L'AIC est un cas particulier de l'AIC* divisé par N.

Notons par p le vrai ordre du processus AR. Nous remarquons que le MFPEE est asymptotiquement équivalent au MAICE car :

$$\log FPE(k) = AIC(k) + o\left(\frac{1}{N}\right). \tag{2.3.10}$$

Shibata (1976) a présenté la distribution asymptotique du MAICE en utilisant la technique de la marche aléatoire.

Lemme (2.3.1)
Soit $Z_1, ..., Z_n$ des variables aléatoires indépendantes et identiquement distribuées et $S_k = Z_1 + ... + Z_k$ $(1 \leq k \leq n)$. Alors les probabilités

$$p_n = \Pr(S_1 > 0, ..., S_n > 0), \quad q_n = \Pr(S_1 \leq 0, ..., S_n \leq 0)$$

peut être représenté comme suit :

$$p_n = \sum_n^* \prod_{i=1}^n \frac{1}{r_i!}\left(\frac{\alpha_i}{i}\right)^{r_i}, \quad q_n = \sum_n^* \prod_{i=1}^n \frac{1}{r_i!}\left(\frac{1-\alpha_i}{i}\right)^{r_i},$$

où $\alpha_i = \Pr(S_i > 0)$ pour $i = 1, ..., n$ et $\sum_n^*$ est la somme sur tous les n-uplet $(r_1, ..., r_n)$ de nombres entiers non-négatifs tel que $r_1 + 2r_2 + ... + nr_n = n$.

Maintenant, nous pouvons présenté la distribution asymptotique de $\widehat{p}$.

Théorème de Shibata
Soit $\{Z_j\}$ une suite de variables aléatoires indépendantes et identiquement distribuées de distribution χ_1^2.
La distribution asymptotique de $\widehat{p}$ est donnée par :

$$\lim_{N\to\infty} \Pr(\widehat{p} = k) = \begin{cases} p_{k-p}p_{K-k} & p \leq k \leq K \\ 0 & \textit{si non} \end{cases}$$

où $p_0 = q_0 = 1$, p_i, q_i sont définis comme dans le lemme précedent et
$\alpha_i = \Pr(\chi_i^2 > 2i)$.

Démonstration Voir Shibata (1976).

Nous remarquons d'après le théorème de Shibata que la distribution asymptotique du MAICE ne dépend pas des coefficients autoregressifs mais de K.
Shibata (1976) a présenté une table des valeurs numériques de la distribution asymptotique dans le cas où $K = 10$. Il a prouvé que la probabilité asymptotique de choisir le vrai ordre est de plus de 70% et celle de choisir $p + 1$ est de plus de 11%.
Hannan (1980) a généralisé le théorème de Shibata au cas d'un modèle ARMA(k, i) pour $k = p$ et $i \geq q$ ou $k \geq p$ et $i = q$.

Théorème (2.3.1) (*Hannan (1980)*)
Soit $\{z_j\}$ une suite de variables aléatoires indépendantes et identiquement distribuées d'une lois χ_1^2 et soit

$$s_j = \textstyle\sum_{l=1}^{j}(z_l - 2),$$
$$\pi(k-p, K-k) = \Pr\{s_1 \leq 0, ..., s_{k-p} \leq 0\}\{s_1 \geq 0, ..., s_{K-k} \geq 0\}.$$

Alors, le MAICE satisfait

$$\lim_{N\to\infty} \Pr\{\widehat{p}=p,\widehat{q}=q\} = \pi(i-q,I-i), \qquad K=p, i\geq q$$
$$\lim_{N\to\infty} \Pr\{\widehat{p}=p,\widehat{q}=q\} = \pi(k-p,K-k), \qquad k\geq p, I=q$$
$$\lim_{N\to\infty} \{\widehat{p}<p \quad ou \quad \widehat{q}<q\} = 0.$$

Démonstration Voir Hannan (1980).

L'inconsistance des estimateurs a formulé quelques objections sur les critères FPE et AIC. Pour éviter ce problème, Ciftcioglu et *al.* (1994) ont présenté une généralisation du critère AIC pour choisir l'ordre d'un modèle AR.

2.4 Critère CIC (*Consistent Information Criterion*)

Soit y_t un modèle autoregressif de deux représentations suivantes :

$$y_t - \sum_{i=1}^{k} \phi_i y_{t-i} = \varepsilon_t, \quad et \quad y_t - \sum_{i=1}^{k+1} \phi'_i y_{t-i} = \varepsilon'_t.$$

où $\{\varepsilon_t, t\in\mathbb{Z}\}, \{\varepsilon'_t, t\in\mathbb{Z}\}$ sont deux processus bruit blanc de moyenne nulle de variance σ_k^2,.σ_{k+1}^2 respectivement.

Sur la base d'observation $y_1, ..., y_N$, les pramètres $\phi_1, ..., \phi_p$ et σ^2 peuvent être obtenus en remplaçant les autocovariances théoriques $\gamma(j)$ par leurs estimateur empiriques. Nous obtenons

$$\begin{aligned}\widehat{\rho}(0) &= \textstyle\sum_{i=1}^{k} \phi_i \widehat{\rho}(i) + \widehat{\sigma}_k^2 \\ &= \textstyle\sum_{i=1}^{k} \phi'_i \widehat{\rho}(i) + \phi'_{k+1}\widehat{\rho}(k+1) + \widehat{\sigma}_{k+1}^2.\end{aligned}$$

où $\widehat{\rho}(.)$ est l'estimateur de la fonction d'autocorrelation et $\widehat{\sigma}_k^2, \widehat{\sigma}_{k+1}^2$désignent les estimateurs de $\sigma_k^2, \sigma_{k+1}^2$respectivement.

D'où

$$\begin{aligned}\widehat{\sigma}_k^2 - \widehat{\sigma}_{k+1}^2 &= \textstyle\sum_{i=1}^{k+1} \phi'_i \widehat{\rho}(i) - \sum_{i=1}^{k} \phi_i \widehat{\rho}(i) \\ &= \textstyle\sum_{i=1}^{k} (\phi'_i - \phi_i)\widehat{\rho}(i) + \phi'_{k+1}\widehat{\rho}(k+1)\end{aligned} \tag{2.4.1}$$

L'algorithme de Levinson-Durbin nous donne l'expression de ϕ'_i en fonction des coefficients ϕ_i comme suit :

$$\phi'_i = \phi_i - \lambda_{k+1}\phi_{k+1-i}, \quad i=1,2,...,k \quad \text{et} \quad \phi'_{k+1} = \lambda_{k+1}$$

où

$$\lambda_{k+1} = \frac{\sum_{i=0}^{k} \phi_i \widehat{\rho}(k+1-i)}{\sum_{i=0}^{k} \phi_i \widehat{\rho}(i)} \quad \text{avec} \quad \phi_0 = -1.$$

Donc,

$$\widehat{\sigma}_k^2 - \widehat{\sigma}_{k+1}^2 = \widehat{\sigma}_k^2 \lambda_{k+1}^2.$$

c'est-à-dire

$$\lambda_{k+1}^2 = \frac{\widehat{\sigma}_k^2 - \widehat{\sigma}_{k+1}^2}{\widehat{\sigma}_k^2} = \phi_{k+1}'^2. \tag{2.4.2}$$

Si le bruit blanc est de loi normale de moyenne nulle et de variance σ^2, alors, $\sqrt{N}\phi'_{k+1}$ est asymptotiquement normale centrée réduite. D'où

$$N\phi_{k+1}'^2 \rightsquigarrow \chi_1^2.$$

Nous supposons que le modèle choisi est un AR(k) tel que :

$$E(\phi'_{k+1}) = 0 \quad et \quad E(\widehat{\sigma}_{k+1}^2) = E(\widehat{\sigma}_k^2) = \sigma^2.$$

Nous proposons le test d'hypothèse suivant :

$$H_0 : \widehat{\sigma}_k^2 = \widehat{\sigma}_{k+1}^2 \quad contre \quad H_1 : \widehat{\sigma}_k^2 > \widehat{\sigma}_{k+1}^2.$$

qui est région critique $\{T > l\}$. C'est-à-dire nous acceptons l'hypotèse H_0 si $T \leq l$.

Pour N assez grand, la statistique

$$T = \frac{Z}{\frac{\widehat{\sigma}_k^2}{\sigma^2}} = (N-k)\frac{\left(\widehat{\sigma}_k^2 - \frac{N-k-1}{N-k}\widehat{\sigma}_{k+1}^2\right)}{\widehat{\sigma}_k^2} \approx (N-k)\frac{\widehat{\sigma}_k^2 - \widehat{\sigma}_{k+1}^2}{\widehat{\sigma}_k^2} \tag{2.4.3}$$

D'autre part

$$(N-k)\frac{\widehat{\sigma}_k^2 - \widehat{\sigma}_{k+1}^2}{\widehat{\sigma}_k^2} = (N-k)\phi_{k+1}'^2 = (1 - \frac{k}{N})N\phi_{k+1}'^2$$

Donc, la statistique T est asymptotiquement de loi chi-deux à 1 degré de liberté. Nous avons

$$\begin{aligned} T \quad &\approx \quad (N-k)\frac{\widehat{\sigma}_k^2 - \widehat{\sigma}_{k+1}^2}{\widehat{\sigma}_k^2} = l \\ &\Longrightarrow \quad N\frac{\widehat{\sigma}_k^2 - \widehat{\sigma}_{k+1}^2}{\widehat{\sigma}_k^2(k+1-k)} - \frac{Nl}{N-k} = 0 \\ &\Longrightarrow \quad N\frac{\triangle \widehat{\sigma}_k^2}{\sigma_k^2 \triangle k} - \frac{N}{N-k}l = 0 \end{aligned}$$

Ainsi le critère CIC est défini par :

$$\frac{\Delta[CIC(k)]}{\Delta k} = N\frac{\frac{\triangle\widehat{\sigma}_k^2}{\widehat{\sigma}_k^2}}{\triangle k} - \frac{N}{N-k}l.$$

Remplaçant le signe de différence par le signe de dérivation

$$\frac{d[CIC(k)]}{dk} = N\frac{\frac{d\widehat{\sigma}_k^2}{\widehat{\sigma}_k^2}}{dk} - \frac{N}{N-k}l.$$

donc nous cherchons l'ordre qui minimise le CIC.

Puisque le niveau de signification du test α

$$\alpha = \int_l^\infty dF_T(t) \quad \text{où } F_T \text{ est la loi de la statistique } T. \tag{2.4.4}$$

Comme α et l sont des fonction de N et de l'ordre du modèle, alors

$$\begin{aligned}
\alpha &= \int_l^\infty dF_T(t) = \frac{k}{N} \\
\Longrightarrow& \frac{d\alpha}{dk} = \frac{d}{dk}\left(\int_l^{+\infty} dF_T(t)\right) = \frac{1}{N} \\
\Longrightarrow& \frac{d\alpha}{dk} = \frac{d}{dk}\left(1 - \int_{-\infty}^{l} dF_T(t)\right) = \frac{1}{N} \\
\Longrightarrow& -f(l)\frac{dl}{dk} = \frac{1}{N}. \text{ où } f \text{ est la densité de } T. \\
\Longrightarrow& -Nf(l)dl = dk.
\end{aligned}$$

En pratique $k \ll N$, alors $N-k$ est considéré comme étant une constante par rapport à k. D'où

$$\begin{aligned}
CIC(k) &= N\log\widehat{\sigma}_k^2 - \frac{N}{N-k}\int l(k)dk \\
&= N\log\widehat{\sigma}_k^2 - \frac{N}{N-k}\int l(k)[-Nf(l(k))dl(k)] \\
&= N\log\widehat{\sigma}_k^2 + \frac{N^2}{N-k}\int l(k)[f(l(k))dl(k)]
\end{aligned}$$

nous trouvons l'expression du critère CIC

$$CIC(k) = N\log\widehat{\sigma}_k^2 - \frac{N}{N-k}k\widetilde{F}(l) \tag{2.4.5}$$

où

$$\widetilde{F}(l) = \frac{\int_l^{+\infty} t dF_T(t)}{\int_l^{+\infty} dF_T(t)} = 1 + \sqrt{\frac{2}{\pi}}l\frac{N}{k}e^{-\frac{1}{2}l}$$

Si nous considérons que $l = l_0$, l_0 est constante par rapport à k, le CIC est donné par

$$CIC(k) = N \log \widehat{\sigma}_k^2 - \frac{N}{N-k} k l_0. \tag{2.4.6}$$

Nous savons que

$$AIC(k) = N \log \widehat{\sigma}_k^2 - 2k.$$

L'égalité entre les deux critère nous donne

$$\frac{N}{N-k} l_0 = 2$$

c'est-à-dire que pour N assez grand $l_0 \simeq 2$. Donc un niveau de signification $\alpha \simeq 0.16$.

Ciftcioglu et *al.* (1994) ont proposé de prendre $\frac{N}{N-k} l_0 \simeq 5.02$ pour l'obtention d'un niveau plus signficatif $\alpha \simeq 0.025$.

2.5 Critère de l'information de Bayes (BIC)

Les méthodes FPE et AIC ont été largement utilisés, dans l'analyse de série chronologique, comme étant des critères objectifs pour sélectionner les ordres du modèle, même s'ils donnent des estimateurs inconsistants des ordres. Habituellement, dans le cas d'un modèle ARMA (p, q), il est raisonnable de supposer que p et q augmentent lorsque la taille de l'échantillon augmente puisque plus de données signifient plus d'information. Ainsi, certains statisticiens indiquent que nous ne devrions pas tenir compte de l'inconsistance des méthodes FPE et AIC dans l'identification d'un modèle ARMA. Cependant, certains sont en désaccord et proposent des nouvelles méthodes d'identification qui ont comme conséquence des estimateurs consistants des vrais ordres. Akaike (1977) a élaboré sous l'hypothèse gaussienne du bruit blanc, le critère BIC (*Bayasian Information Criterion*) pour estimer les ordres (p, q) d'un modèle ARMA. Ce critère permet d'obtenir des estimateurs consistants.

Juste après la publication d'Akaike (1977), quelques statisticiens (Kashyap (1977, 1978 et 1982), Schwarz (1978) par exemple) ont présenté le critère suivant :

$$(-2) \log(ML) + (\text{nombre de paramètres}) \times \log N \tag{2.5.1}$$

où nous choisissons donc, l'ordre qui minimise ce dernier et qui s'appelle MBICE (*minimum BIC estimate*). En particulier pour un processus ARMA, le critère BIC est donné par :

$$BIC(k, i) = \log \widetilde{\sigma}_{k,i}^2 + (k + i) \frac{\log N}{N}, \tag{2.5.2}$$

et

$$(\widehat{p}, \widehat{q}) = MBICE = \arg\left\{\min_{(k,i)}\left[\log \widetilde{\sigma}^2_{k,i} + (k+i)\frac{\log N}{N}\right]\right\}, \qquad (2.5.3)$$

où $\widetilde{\sigma}^2_{k,i}$ est l'estimateur du maximum de vraisemblance de la variance du bruit blanc.

Théorème (2.5.1) (*Hannan (1982)*)
Considérons le modèle ARMA(p, q) causal suivant :

$$y_t - \sum_{i=1}^{p} \phi_i y_{t-i} = \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j},$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est une suite de variables aléatoires non-corrélées de moyenne nulle et de variance $\sigma^2 > 0$. De plus, nous supposons que :

$$E(\varepsilon_t/\mathcal{F}_{t-1}) = 0,\ E(\varepsilon_t^2/\mathcal{F}_{t-1}) = \sigma^2,\ E(\varepsilon_t^4) < \infty$$

où $\mathcal{F}_t$ est le σ-algébre engendré par $\{\varepsilon_t, \varepsilon_{t-1}, ...\}$.

Si $\{\varepsilon_t, t \in \mathbb{Z}\}$ une suite de variables aléatoires i.i.d, alors, le $MBICE$ est fortement consistant.

Si $\{\varepsilon_t, t \in \mathbb{Z}\}$ est une suite de variables aléatoires non-corrélées et $E(|\varepsilon_t|^r) < \infty$ pour $r > 4$, alors, le $MBICE$ est fortement consistant.

Démonstration Voir Hannan (1982).

2.5.1 Approche de Schwarz

Schwarz (1978) a dérivé le critère BIC comme suit :

Soient $x_1, ..., x_N$ des variables aléatoires indépendantes et identiquement distribuées de densité

$$f(x, \theta) = \exp\{\theta^t y(x) - b(\theta)\}.$$

où $\theta \in \Omega$ qui est un sous-ensemble convexe de l'espace euclidien $\mathbb{R}^K$ de dimension K. Les autres modèles ont des espaces paramétriques $\omega_i = \Omega_i \cap \Omega$, où Ω_i est un sous-espace linéaire de $\mathbb{R}^K$ de dimension k_i. Soit α_i la probabilité antérieure que le $i \overset{eme}{=}$ modèle est le vrai. Soit également, μ_i la distribution conditionnelle antérieure de θ sachant que $\theta \in \omega_i$, nous supposons que μ_i est une densité de dimension k_i et qui est localement non nulle dans ω_i.

Théorème (2.5.2) (*Schwarz (1978)*)

Sous les suppositions ci-dessus nous avons :

1° Le modèle qui a la plus grande probabilité postérieure est celui qui maximise

$$S\left(\overline{y}, N, i\right) = \log \int_{\omega_i} \alpha_i \exp\left[\{\theta^t \overline{y} - b(\theta)\} N\right] d\mu_i(\theta), \quad \text{où} \quad \overline{y} = \sum_{t=1}^{N} y(x_t).$$

2° Pour tout i,

$$S(\overline{y}, N, i) = N \times \sup_{\theta \in \omega_i} \left\{\theta^t \overline{y} - b(\theta)\right\} - \frac{k_i}{2} \log N + R,$$

où R est en fonction de N.

Démonstration Voir Schwarz (1978).

Schwarz a proposé d'utiliser seulement les deux premiers termes de $S(\overline{y}, N, i)$ comme critère de sélection. Car

$$N \times \sup_{\theta \in \omega_i} \left\{\theta^t \overline{y} - b(\theta)\right\}$$

est le maximum de log-vraisemblance dans ω_i, $(-2)S(\overline{y}, N, i)$ est asymptotiquement équivalent au BIC. Par conséquent, l'ordre qui a une probabilité a posteriori maximale est asymptotiquement égal au MBICE. Nous remarquons que le résultat final ne dépend pas des probabilités antérieures.

2.5.2 Approche de Kashyap

Kashyap (1977) a présenté une méthode bayésienne pour comparer différents types de structures dynamiques. La règle de décision répond à une grande variété de questions. Il a appliqué cette approche, en 1978 pour le cas autoregressif pur (AR) et en 1982 pour le cas mixte ARMA.

Soit $y = (y_1, ..., y_N)'$ un vecteur d'observation. Nous supposons qu'il y a r hypothèses $H_1, ..., H_r$ tel que le vecteur d'observation vérifie seulement une seule hypothèse. Soit $p_i(y_i, \alpha_i)$ la densité sous l'hypothèse H_i de y_i, où α_i est un paramétre vectoriel de $\omega_i \left(\subset R^{k_i}\right)$ et n'a aucune composante nulle.

Théorème (2.5.3) (*Kashyap (1978)*)

Nous notons par $\widetilde{\alpha}_i$ l'estimateur du maximum de vraisemblance de α_i. Si la densité antérieure $f_i(\alpha_i)$ et la densité $p_i(y_i, \alpha_i)$ satisfont certaines conditions de régularité, alors le log de probabilité postérieure que H_i est la vraie hypothèse sachant y est donnée par :

$$\log \Pr\left(H_i / y\right) = \log p_i\left(y_i; \widetilde{\alpha}_i\right) - \frac{k_i}{2} \log N + O_p(1). \tag{2.5.4}$$

Démonstration Voir Kashyap (1978).

La règle de décision optimale de bayes pour choisir le meilleur modèle d'une série chronologique donnée doit attribuer y à H_k tels que $\Pr(H_k/y)$ soit supérieur ou égale à $\Pr(H_i/y)$ pour $i = 1, ..., r$. Puisque le maximum de vraisemblance est d'ordre $O_p(N)$, $P(H_i/y)$ est asymptotiquement équivalent à

$$\log p_i\left(y_i; \widetilde{\alpha}_i\right) - \frac{k_i}{2} \log N. \tag{2.5.5}$$

Ainsi, ce critère est asymptotiquement équivalent au BIC. De plus, il ne dépend pas des densités antérieures.

2.6 Critère de Hannan et Quinn (HQC)

Nous considérons le modèle ARMA(p, q) causal suivant :

$$y_t - \sum_{i=1}^{p} \phi_i y_{t-i} = \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \tag{2.6.1}$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est une suite de variables aléatoires non-corrélées de moyenne nulle et de variance $\sigma^2 > 0$. De plus, nous supposons que :

$$E(\varepsilon_t/\mathcal{F}_{t-1}) = 0,\ E(\varepsilon_t^2/\mathcal{F}_{t-1}) = \sigma^2,\ E(\varepsilon_t^4) < \infty$$

où $\mathcal{F}_t$ est le σ-algébre engendré par $\{\varepsilon_t, \varepsilon_{t-1}, ...\}$ (ou bien par $\{y_t, y_{t-1}, ...\}$).

Hannan et Quinn (1979) ont présenté un critère d'identification d'un modèle AR pur. Ce critère tend a minimisé

$$\log \widehat{\sigma}_k^2 + k\frac{C(N)}{N} \tag{2.6.2}$$

tels que lorsque $N \to \infty$, l'ordre choisi soit fortement consistant et $\frac{C(N)}{N}$ diminu assez rapidement.

Théorème (2.6.1) (*Hannan et Quinn (1979)*)
Soit $\{y_1, ..., y_N\}$ une réalisation d'un processus AR(p) causal satisfait les conditions ci-dessus. Soit K un entier positif tel que $p < K$. Si

$$\widehat{p} = \arg\left\{\min_{k \le K}\left[\log \widehat{\sigma}_k^2 + 2ck\frac{\log\log N}{N}, \quad c > 1\right]\right\}. \tag{2.6.3}$$

Alors, $\widehat{p}$ est un estimateur de p fortement consistant.

Démonstration Voir Hannan et Quinn (1979).

Nous applons $\widehat{p}$ l'estimateur du minimum du HQC (*minimum HQC estimate*) et on le note par :

$$MHQCE = \arg\left\{\min_{k}\left[\log\widehat{\sigma}_k^2 + 2ck\frac{\log\log N}{N}, \quad c > 1\right]\right\}. \tag{2.6.4}$$

Dans le cas d'un modèle ARMA(p, q) le HQC est donné par :

$$HQC(k, i) = \log\widehat{\sigma}_{k,i}^2 + 2(k+i)c\frac{\log\log N}{N}, c > 1 \tag{2.6.5}$$

et

$$MHQCE = \arg\left\{\min_{(k,i)}\left[\log\widehat{\sigma}_{k,i}^2 + 2(k+i)c\frac{\log\log N}{N}, c > 1\right]\right\}, \tag{2.6.6}$$

Soient I et K deux bornes superieures satisfaisont $p \leq K$ et $q \leq I$ où I et K sont connues a priori.

Théorème (2.6.2) (*Hannan (1982)*)
Considérons le modèle ARMA(p, q) causal suivant :

$$y_t - \sum_{i=1}^{p}\phi_i y_{t-i} = \varepsilon_t - \sum_{j=1}^{q}\theta_j\varepsilon_{t-j},$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées et $E(\varepsilon_t^4) < \infty$. Alors, le $MHQCE = (\widehat{p}, \widehat{q})$ est fortement consistant.

Démonstration Voir Hannan (1982).

Théorème (2.6.3) (*Hannan (1982)*)
Considérons le modèle ARMA(p, q) de la représentation (2.6.1). Soit

$$(\widehat{p}, \widehat{q}) = \arg\left\{\min_{(k,i)}\left[\log\widetilde{\sigma}_{k,i}^2 + (k+i)\frac{C(N)}{N}\right]\right\},$$

où $C(N)$ satistait $\lim_{N\to\infty} C(N) = \infty$, $\lim_{N\to\infty}\frac{C(N)}{N} = 0$. Alors, $(\widehat{p}, \widehat{q})$ est faiblement consistant.

Démonstration Voir Hannan (1982).

2.7 Critère de la densité prédictive (PDC)

Djuric et Kay (1992) ont présenté un critère pour choisir l'ordre d'un modèle AR, ce critère est nomé PDC (*Predictive Density Criterion*),il est fondé sur la densité prédictive bayesienne, pour déterminer l'ordre d'un modèle autoregressif stationnaire.

Cette approche utilise la densité prédictive dans le but de dégager un critère qui minimise la probabilité de l'erreur globale, pour obtenir l'ordre optimal.

Soient $y_1, ..., y_N$ une serie d'observations d'un modèle AR(p) stationnaire representé comme suit

$$y_t = -\sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t. \tag{2.7.1}$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un processus bruit blanc gaussien de moyenne nulle de variance σ^2 et pour tout $|z| \leq 1$: $A(z) = \sum_{i=1}^{p} \phi_i z^i \neq 0$. Le critère PDC défini à partir de la densité prédictive est donné par :

$$PDC(k) = \sum_{t=2}^{N} -\log f(y_t/y_1, ..., y_{t-1}, k) \tag{2.7.2}$$

où $f(y_t/y_1, ..., y_{t-1}, k)$ est la densité prédictive de y_t sachant le passé $y_1, ..., y_{t-1}$. Nous définissons l'ordre optimal par :

$$MPDCE = \arg\left[\min_{0\leq k\leq p}\left\{\sum_{t=2}^{N} -\log f(y_t/y_1, ..., y_{t-1}, k)\right\}\right]. \tag{2.7.3}$$

pour des raisons de calcul nous ne pourrons pas ecrire la densité prédictive lorsque nous n'avons pas le nombre minimum d'observations $t > 2k+1$. Djuric et Kay (1992) ont donné la densité prédictive d'un processus AR stationnaire comme suit :

Proposition (2.7.1) (*Djuric et Kay (1992)*)

La densité prédictive d'un processus AR stationnaire y_t solution de l'équation aux différences stochastique (2.7.1) est donnée pour $t > 2k+1$ par :

$$f(y_t/y_1, ..., y_{t-1}, k) = \begin{cases} \frac{1}{\sqrt{2\pi}} \frac{\Gamma(t/2)}{\Gamma((t-1)/2)} \frac{\left(\sum_{i=1}^{t-1} y_i^2\right)^{\frac{t-1}{2}}}{\left(\sum_{i=1}^{t} y_i^2\right)^{\frac{t}{2}}} & si\ k=0 \\ \frac{1}{\sqrt{2\pi}} \frac{\|H_k(t-2)\|}{\|H_k(t-1)\|} \frac{\Gamma\left(\frac{t-2k}{2}\right)}{\Gamma\left(\frac{t-2k-1}{2}\right)} \times \frac{\left(Y_{k+1,t-1}^t P_k(t-2) Y_{k+1,t-1}\right)^{\frac{t--2k-1}{2}}}{\left(Y_{k+1,t}^t P_k(t-1) Y_{k+1,t}\right)^{\frac{t-2k}{2}}} & si\ k>0 \end{cases} \tag{2.7.4}$$

où

$$H_k(t) = \begin{pmatrix} y_k & y_{k-1} & \cdots & y_1 \\ y_{k+1} & y_k & \cdots & y_2 \\ \vdots & \vdots & \ddots & \vdots \\ y_{t-2} & y_{t-3} & \cdots & y_{t-1-k} \end{pmatrix}, Y_{k,k+t} = \begin{pmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+t} \end{pmatrix}$$

$$P_k^{\perp}(t) = Id - H_k(t) \left\| H_k(t) \right\|^{-2} H_k^t(t).$$

Démonstration : Voir par exemple, Hemis (1999)

Pour $k = 0$, nous trouvons que

$$PDC(0) = J_0 = \sum_{t=2}^{N} - \log f(y_t/y_1, ..., y_{t-1}, k = 0) \tag{2.7.5}$$

Le nombre minimum d'observations est égal à 2 et le seul paramétre est σ^2.

Pour $k = 1$, nous avons deux paramétres inconnus ϕ_1 et σ^2. Pour pouvoir calculer la densité prédictive, il nous faut au moins 4 obsrvations ($t \geq 4$).

Notons

$$J_1' = \sum_{t=2}^{N} - \log f(y_t/y_1, ..., y_{t-1}, k = 1) \tag{2.7.6}$$

Nous ne pouvons pas comparer (2.7.5) et (2.7.6) car les deux expressions sont calculées à partir des densités prédictives pour des nombres d'observations différents. Pour résoudre ce problème, nous écrivons (2.7.6) sous la forme suivante :

$$J_1 = \sum_{t=2}^{3} - \log f(y_t/y_1, ..., y_{t-1}, k = 0) + \sum_{t=4}^{N} - \log f(y_t/y_1, ..., y_{t-1}, k = 1)$$

Pour $k = 2$, nous avons

$$J_2' = \sum_{t=6}^{N} - \log f(y_t/y_1, ..., y_{t-1}, k = 2). \tag{2.7.7}$$

Soit

$$\begin{aligned} J_2 = \textstyle\sum_{t=2}^{3} - \log f(y_t/y_1, ..., y_{t-1}, k = 0) + \sum_{t=4}^{5} - \log f(y_t/y_1, ..., y_{t-1}, k = 1) \\ + \textstyle\sum_{t=6}^{N} - \log f(y_t/y_1, ..., y_{t-1}, k = 2) \end{aligned}$$

Nous procédons de la même manière avec les modèles AR d'ordres superieur. Ainsi nous pouvons calculer l'estimateur MPDCE comme étant la valeur qui minimise J_k. D'où

$$MPDCE = \arg \left\{ \min_{0 \leq k \leq K} J_k \right\}. \tag{2.7.8}$$

où K est un entier positif tel que $p \leq K$.

2.8 Etude de simulation

Dans cette section nous présenterons les résultats de simulation de 9 différents modèles autoregressifs moyennes mobiles. Nous avons simulé 1000 séries de longueur variant de 50 à 200, supposées obtenues à partir de modèles ARMA(p, q) stationnaire. Nous avons fait varier les paramètres de façon à ce que les propriétés de causalité et d'inversibilité soient vérifiées. Pour chaque modèle autoregressif pur, nous avons appliqué les critères FPE, AIC, BIC, HQC et PDC et pour les modèles ARMA mixte les critères AIC, BIC et HQC. L'objet de cette section est de faire une comparaison entre la performance des différents critères de sélection.

Nous avons estimé les paramètres de chaque modèle et pour les différentes tailles par la méthode des moindre carrés. Les sorties de l'estimation contiennent les moyennes et les écart-types des valeurs estimées des paramètres.

Estimation des paramètres du modèle AR(1) par la méthode des moindre carrés.

paramètres / taille	50	100	150	200
$\phi = 0.5$	0.4804 (0.1286)	0.4887 (0.0893)	0.4943 (0.0738)	0.4965 (0.0638)
$\sigma^2 = 1$	0.9817 (0.1969)	0.9982 (0.1422)	0.9913 (0.1218)	0.9967 (0.1032)

Tableau (2.8.1)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1, \ \ \phi = 0.5, \ \ \sigma^2 = 1, \ \ N = 50$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	6.8	7.4	17.3	24.3	22.1	15.5
1	43.0	48.9	70.0	70.3	71.2	82.7
2	14.0	15.1	8.0	4.6	5.2	1.5
3	7.8	7.2	2.4	0.4	0.8	0.3
4	6.7	6.7	1.4	0.3	0.6	0.0
5	4.2	3.7	0.5	0.1	0.1	0.0
6	7.9	6.1	0.2	0.0	0.0	0.0
7	9.6	4.9	0.2	0.0	0.0	0.0

Tableau (2.8.2)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1, \quad \phi = 0.5, \quad \sigma^2 = 1, \quad N = 100$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.6	0.6	1.8	2.1	1.9	0.4
1	53.6	57.1	90.6	94.9	92.4	98.0
2	13.3	13.4	4.8	2.2	4.1	1.4
3	9.1	8.3	2.1	0.7	1.1	0.2
4	5.9	5.9	0.4	0.1	0.4	0.0
5	6.0	5.6	0.1	0.0	0.1	0.0
6	6.5	5.2	0.2	0.0	0.0	0.0
7	5.0	3.9	0.0	0.0	0.0	0.0

Tableau (2.8.3)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1, \quad \phi = 0.5, \quad \sigma^2 = 1, \quad N = 150$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.1	0.2	0.2	0.4	0.2	0.1
1	58.9	61.5	93.6	96.9	93.6	99.0
2	14.1	13.9	4.4	2.5	4.4	0.8
3	7.8	7.3	1.2	0.2	1.2	0.1
4	6.0	5.6	0.4	0.0	0.4	0.0
5	4.1	3.5	0.2	0.0	0.2	0.0
6	5.1	4.8	0.0	0.0	0.0	0.0
7	3.9	3.2	0.0	0.0	0.0	0.0

Tableau (2.8.4)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1, \quad \phi = 0.5, \quad \sigma^2 = 1, \quad N = 200$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0	0.0
1	60.0	62.1	95.2	98.0	94.3	99.6
2	12.6	12.6	3.7	1.9	4.5	0.4
3	7.1	6.8	0.8	0.0	0.8	0.0
4	5.6	5.5	0.2	0.1	0.3	0.0
5	6.2	6.0	0.1	0.0	0.1	0.0
6	4.1	3.4	0.0	0.0	0.0	0.0
7	4.4	3.6	0.0	0.0	0.0	0.0

Tableau (2.8.5)

Estimation des paramètres du modèle AR(1) par la méthode des moindre carrés.

paramètres / taille	50	100	150	200
$\phi = 0.7$	0.6698 (0.1063)	0.6897 (0.0752)	0.6919 (0.0624)	0.6925 (0.0516)
$\sigma^2 = 1$	0.9792 (0.1988)	0.9832 (0.1381)	0.9933 (0.1139)	0.9952 (0.1020)

Tableau (2.8.6)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1, \quad \phi = 0.7, \quad \sigma^2 = 1, \quad N = 50$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	1.2	1.2	2.6	3.9	3.5	1.3
1	48.1	54.2	83.7	90.4	89.2	95.9
2	11.3	12.8	8.4	5.0	5.7	2.7
3	8.9	8.6	2.2	0.4	0.7	0.1
4	5.9	5.9	1.2	0.2	0.4	0.0
5	5.2	3.9	0.9	0.0	0.3	0.0
6	7.1	5.5	0.5	0.1	0.1	0.0
7	12.3	7.9	0.5	0.0	0.1	0.0

Tableau (2.8.7)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1, \quad \phi = 0.7, \quad \sigma^2 = 1, \quad N = 100$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0	0.0
1	56.7	60.1	90.8	94.6	92.4	98.3
2	14.9	14.6	7.5	4.9	6.4	1.6
3	6.3	6.6	1.4	0.4	0.9	0.1
4	6.7	6.6	0.3	0.1	0.3	0.0
5	5.8	4.9	0.0	0.0	0.0	0.0
6	3.6	2.6	0.0	0.0	0.0	0.0
7	6.0	4.6	0.0	0.0	0.0	0.0

Tableau (2.8.8)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1, \ \ \phi = 0.7, \ \ \sigma^2 = 1, \ \ N = 150$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0	0.0
1	59.0	62.2	94.8	97.3	94.8	98.8
2	13.9	14.2	4.4	2.7	4.4	1.2
3	6.4	6.1	0.7	0.0	0.7	0.0
4	5.8	5.3	0.1	0.0	0.1	0.0
5	4.6	3.8	0.0	0.0	0.0	0.0
6	3.8	3.4	0.0	0.0	0.0	0.0
7	6.5	5.0	0.0	0.0	0.0	0.0

Tableau (2.8.9)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1, \ \ \phi = 0.7, \ \ \sigma^2 = 1, \ \ N = 200$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0	0.0
1	61.0	62.6	95.6	97.5	94.7	98.9
2	13.8	13.8	3.8	2.2	4.2	1.0
3	8.4	8.2	0.4	0.2	0.8	0.1
4	5.0	4.9	0.1	0.1	0.2	0.0
5	4.8	4.5	0.0	0.0	0.0	0.0
6	3.2	2.7	0.1	0.0	0.1	0.0
7	3.8	3.3	0.0	0.0	0.0	0.0

Tableau (2.8.10)

Estimation des paramètres du modèle AR(1) par la méthode des moindre carrés.

paramètres / taille	50	100	150	200
$\phi = 0.95$	0.9175 (0.0644)	0.9322 (0.0406)	0.9373 (0.0321)	0.9414 (0.0253)
$\sigma^2 = 1$	0.9715 (0.2020)	0.9860 (0.1369)	0.9981 (0.1174)	0.9966 (0.1008)

Tableau (2.8.11)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1, \quad \phi = 0.95, \quad \sigma^2 = 1, \quad N = 50$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.2	0.1	0.1
1	47.2	53.4	84.6	93.2	91.2	96.7
2	14.8	16.1	9.4	5.3	6.4	3.1
3	7.2	7.2	3.7	0.9	1.5	0.1
4	7.5	6.9	1.2	0.2	0.5	0.0
5	5.7	4.3	0.6	0.1	0.1	0.0
6	8.2	6.3	0.3	0.0	0.0	0.0
7	9.4	5.8	0.2	0.1	0.2	0.0

Tableau (2.8.12)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1, \quad \phi = 0.95, \quad \sigma^2 = 1, \quad N = 100$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0	0.0
1	53.2	56.1	90.9	96.5	93.0	97.8
2	14.5	14.6	6.2	3.1	5.2	2.1
3	7.8	7.6	1.6	0.3	0.9	0.1
4	7.2	6.9	0.9	0.1	0.6	0.0
5	5.9	5.4	0.3	0.0	0.3	0.0
6	4.7	3.4	0.0	0.0	0.0	0.0
7	6.7	6.0	0.1	0.0	0.0	0.0

Tableau (2.8.13)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1, \quad \phi = 0.95, \quad \sigma^2 = 1, \quad N = 150$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0	0.0
1	59.8	62.1	93.6	96.8	93.6	98.0
2	13.9	14.4	5.4	2.8	5.4	2.0
3	6.0	5.6	0.6	0.3	0.6	0.0
4	6.3	6.3	0.2	0.1	0.2	0.0
5	5.3	5.0	0.1	0.0	0.1	0.0
6	5.0	4.1	0.1	0.0	0.1	0.0
7	3.7	2.5	0.0	0.0	0.0	0.0

Tableau (2.8.14)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p=1,\ \ \phi=0.95,\ \ \sigma^2=1,\ \ N=200$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0	0.0
1	62.2	64.2	95.3	97.7	94.8	98.2
2	11.4	10.9	3.3	1.9	3.8	1.7
3	8.0	7.7	1.1	0.3	1.0	0.1
4	5.5	5.7	0.3	0.1	0.3	0.0
5	4.3	3.8	0.0	0.0	0.1	0.0
6	3.4	3.1	0.0	0.0	0.0	0.0
7	5.2	4.6	0.0	0.0	0.0	0.0

Tableau (2.8.15)

Estimation des paramètres du modèle AR(2) par la méthode des moindre carrés.

paramètres / taille	50	100	150	200
$\phi_1=0.777$	0.7601 (0.1293)	0.7679 (0.0876)	0.7722 (0.0710)	0.7797 (0.0618)
$\phi_2=-0.492$	-0.4813 (0.1278)	-0.4872 (0.0923)	-0.4871 (0.0726)	-0.4920 (0.0632)
$\sigma^2=0.9$	0.9592 (0.1993)	0.9811 (0.1437)	0.9862 (0.1149)	0.9948 (0.1003)

Tableau (2.8.16)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p=2,\ \ \phi_1=0.777,\ \ \phi_2=-0.492,\ \ \sigma^2=1,\ \ N=50$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.2	0.4	0.9	0.6	1.4
1	1.2	1.2	4.9	9.6	8.6	15.8
2	51.0	57.8	81.9	84.6	84.2	81.1
3	11.4	12.0	7.1	3.2	4.1	1.7
4	8.4	7.6	3.0	1.0	1.4	0.0
5	7.0	6.4	1.1	0.3	0.5	0.0
6	8.7	6.4	1.1	0.4	0.5	0.0
7	12.3	8.4	0.5	0.0	0.1	0.0

Tableau (2.8.17)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 2,\ \ \phi_1 = 0.777,\ \ \phi_2 = -0.492,\ \ \sigma^2 = 1,\ \ N = 100$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.1	0.4	0.3	1.6
2	56.2	60.5	90.5	95.5	92.1	97.1
3	14.2	13.7	6.8	3.4	5.5	1.0
4	7.5	6.8	2.2	0.7	1.7	0.3
5	7.5	6.5	0.4	0.0	0.4	0.0
6	6.0	5.2	0.0	0.0	0.0	0.0
7	8.6	7.3	0.0	0.0	0.0	0.0

Tableau (2.8.18)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 2,\ \ \phi_1 = 0.777,\ \ \phi_2 = -0.492,\ \ \sigma^2 = 1,\ \ N = 150$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.2
2	57.7	61.0	93.0	97.0	93.0	98.0
3	16.3	15.6	5.5	2.7	5.5	1.7
4	8.7	8.2	1.1	0.3	1.1	0.1
5	5.8	5.5	0.1	0.0	0.1	0.0
6	5.7	5.2	0.2	0.0	0.2	0.0
7	5.8	4.5	0.1	0.0	0.1	0.0

Tableau (2.8.19)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 2,\ \ \phi_1 = 0.777,\ \ \phi_2 = -0.492,\ \ \sigma^2 = 1,\ \ N = 200$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0
2	61.3	63.0	94.7	97.6	93.2	99.0
3	13.5	13.8	4.7	2.2	5.5	1.0
4	7.9	7.7	0.4	0.1	0.9	0.0
5	6.3	5.8	0.1	0.1	0.3	0.0
6	5.4	5.2	0.0	0.0	0.1	0.0
7	5.6	4.5	0.1	0.0	0.0	0.0

Tableau (2.8.20)

Estimation des paramètres du modèle AR(2) par la méthode des moindre carrés.

paramètres / taille	50	100	150	200
$\phi_1 = 0.44$	0.4343 (0.1401)	0.4357 (0.0991)	0.4389 (0.0802)	0.4379 (0.0660)
$\phi_2 = 0.35$	0.3118 (0.1336)	0.3324 (0.1001)	0.3358 (0.0795)	0.3402 (0.0673)
$\sigma^2 = 0.9$	0.8643 (0.1789)	0.8836 (0.1229)	0.8837 (0.1042)	0.8938 (0.0895)

Tableau (2.8.21)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 2, \quad \phi_1 = 0.44, \quad \phi_2 = 0.35, \quad \sigma^2 = 0.9, \quad N = 50$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	4.0	4.4	8.7	13.5	11.9	3.7
1	12.3	14.4	31.2	41.6	39.6	52.7
2	35.6	40.6	49.6	41.0	43.3	41.7
3	12.1	11.9	6.1	3.0	3.8	1.8
4	7.7	7.3	1.4	0.3	0.4	0.1
5	8.5	7.6	1.2	0.1	0.3	0.0
6	8.5	6.6	1.2	0.3	0.5	0.0
7	11.3	7.2	0.6	0.2	0.2	0.0

Tableau (2.8.22)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 2, \quad \phi_1 = 0.44, \quad \phi_2 = 0.35, \quad \sigma^2 = 0.9, \quad N = 100$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0	0.2	0.4	0.4	0.0
1	2.5	2.9	12.0	18.7	14.4	26.4
2	54.7	59.0	79.1	77.0	78.3	72.2
3	12.7	12.5	6.7	3.6	5.3	1.2
4	10.9	10.8	1.2	0.2	1.0	0.2
5	6.6	5.3	0.5	0.1	0.5	0.0
6	5.5	4.1	0.2	0.0	0.0	0.0
7	7.1	5.4	0.1	0.0	0.1	0.0

Tableau (2.8.23)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 2,\ \ \phi_1 = 0.44,\ \ \phi_2 = 0.35,\ \ \sigma^2 = 0.9,\ \ N = 150$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.3	0.3	3.0	5.1	3.0	7.5
2	59.5	61.7	92.2	92.7	92.2	91.4
3	14.1	13.9	3.5	2.1	3.5	1.0
4	8.2	8.1	0.8	0.0	0.8	0.0
5	6.9	6.7	0.5	0.1	0.5	0.1
6	4.8	4.1	0.0	0.0	0.0	0.0
7	6.2	5.2	0.0	0.0	0.0	0.0

Tableau (2.8.24)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 2,\ \ \phi_1 = 0.44,\ \ \phi_2 = 0.35,\ \ \sigma^2 = 0.9,\ \ N = 200$						
p	FPE	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.1	0.1	0.7	1.4	0.7	2.6
2	60.8	62.5	94.1	96.3	93.2	96.0
3	14.1	14.7	3.9	2.0	4.5	1.3
4	8.7	8.0	1.1	0.3	1.3	0.1
5	5.5	5.0	0.1	0.0	0.2	0.0
6	5.5	5.1	0.0	0.0	0.0	0.0
7	5.3	4.6	0.1	0.0	0.1	0.0

Tableau (2.8.25)

Estimation des paramètres du modèle ARMA(1,1) par la méthode des moindre carrés.

paramètres / taille	50	100	150	200
$\phi = 0.2$	0.2329 (0.1863)	0.2151 (0.1314)	0.2120 (0.1035)	0.2118 (0.0874)
$\theta = 0.7$	0.6717 (0.1742)	0.6893 (0.1055)	0.6915 (0.0795)	0.6910 (0.0733)
$\sigma^2 = 1$	0.9653 (0.1969)	0.9892 (0.1410)	0.9923 (0.1199)	0.9985 (0.0982)

Tableau (2.8.26)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p = 1, \quad q = 1, \quad \phi = 0.2, \quad \theta = 0.7, \quad \sigma^2 = 1, \quad N = 50$								
	p/q	1	2	3	4	5	6	7
AIC		9.4	3.4	1.2	1.4	1.2	1.1	0.8
BIC	1	54.1	10.2	2.5	1.4	0.6	0.6	0.3
HQC		75.4	7.5	2.0	0.6	0.3	0.1	0.1
AIC		1.8	1.3	5.1	3.3	2.5	2.2	1.4
BIC	2	7.1	1.1	5.3	2.1	0.5	0.4	0.2
HQC		7.3	0.6	1.0	1.1	0.1	0.1	0.0
AIC		1.5	2.1	2.6	2.9	1.4	1.4	1.2
BIC	3	1.7	2.0	0.5	0.5	0.3	0.3	0.5
HQC		1.0	0.9	0.0	0.0	0.1	0.1	0.1
AIC		0.2	2.6	2.7	1.7	2.7	1.9	2.0
BIC	4	0.7	1.1	0.5	0.2	0.4	0.1	0.1
HQC		0.7	0.4	0.0	0.1	0.0	0.0	0.0
AIC		0.2	2.3	1.9	1.3	1.7	1.7	1.0
BIC	5	0.5	1.4	0.5	0.0	0.1	0.0	0.0
HQC		0.0	0.3	0.0	0.0	0.0	0.0	0.0
AIC		0.1	1.9	1.6	3.0	2.7	2.1	2.3
BIC	6	0.1	0.3	0.1	0.2	0.2	0.1	0.1
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.3	2.3	2.1	3.2	2.7	1.3	1.3
BIC	7	0.2	0.5	0.2	0.2	0.0	0.0	0.0
HQC		0.0	0.1	0.0	0.0	0.0	0.0	0.0

Tableau (2.8.27)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p=1,\ q=1,\ \phi=0.2,\ \theta=0.7,\ \sigma^2=1,\ N=100$								
	p/q	1	2	3	4	5	6	7
AIC		7.5	3.0	1.1	0.7	0.3	0.4	0.6
BIC	1	70.2	8.9	1.7	0.4	0.2	0.0	0.0
HQC		83.8	6.5	0.9	0.2	0.0	0.0	0.0
AIC		1.3	0.7	5.2	3.4	3.0	1.8	1.2
BIC	2	6.9	0.8	3.0	1.1	0.1	0.1	0.0
HQC		5.9	0.7	1.0	0.3	0.0	0.0	0.0
AIC		1.0	1.3	3.7	3.4	1.6	1.4	1.0
BIC	3	1.1	0.8	0.7	0.6	0.0	0.0	0.0
HQC		0.5	0.1	0.0	0.0	0.0	0.0	0.0
AIC		0.4	1.4	2.6	3.3	3.6	2.8	2.4
BIC	4	0.4	0.8	0.6	0.0	0.1	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.2	2.1	1.5	2.1	2.2	3.1	2.5
BIC	5	0.1	0.5	0.1	0.0	0.0	0.0	0.0
HQC		0.1	0.1	0.0	0.0	0.0	0.0	0.0
AIC		0.3	1.8	2.3	2.4	2.0	1.2	4.0
BIC	6	0.2	0.2	0.0	0.0	0.0	0.0	0.0
HQC		0.1	0.1	0.0	0.0	0.0	0.0	0.0
AIC		0.3	0.9	1.7	2.5	2.0	2.2	2.6
BIC	7	0.0	0.0	0.1	0.1	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0

Tableau (2.8.28)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p=1,\ q=1,\ \phi=0.2,\ \theta=0.7,\ \sigma^2=1,\ N=150$								
	p/q	1	2	3	4	5	6	7
AIC		7.7	2.3	1.6	0.8	0.6	0.2	0.4
BIC	1	76.2	8.4	1.7	0.3	0.1	0.1	0.0
HQC		86.6	5.4	0.6	0.2	0.0	0.0	0.0
AIC		2.7	0.7	4.0	3.3	2.3	2.0	1.8
BIC	2	6.3	0.2	1.8	0.7	0.2	0.1	0.1
HQC		5.0	0.1	0.5	0.2	0.0	0.1	0.0
AIC		0.9	0.7	4.9	3.3	2.3	1.3	1.4
BIC	3	1.1	0.3	0.5	0.4	0.0	0.0	0.0
HQC		0.9	0.0	0.0	0.2	0.0	0.0	0.0
AIC		0.6	1.1	1.5	1.8	4.6	3.5	2.0
BIC	4	0.3	0.3	0.2	0.1	0.0	0.1	0.0
HQC		0.1	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.4	1.5	1.9	1.5	3.0	3.7	2.8
BIC	5	0.2	0.0	0.1	0.0	0.0	0.0	0.0
HQC		0.1	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.0	0.5	1.5	1.7	1.9	2.7	4.7
BIC	6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.2	1.4	1.3	2.3	1.9	2.0	2.8
BIC	7	0.0	0.1	0.0	0.1	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0

Tableau (2.8.29)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p=1,\ q=1,\ \phi=0.2,\ \theta=0.7,\ \sigma^2=1,\ N=200$								
	p/q	1	2	3	4	5	6	7
AIC		9.7	3.8	1.3	0.7	0.5	0.3	0.3
BIC	1	80.8	8.2	1.5	0.2	0.0	0.0	0.0
HQC		88.6	5.6	0.8	0.0	0.0	0.0	0.0
AIC		1.6	1.2	3.8	3.8	2.3	1.7	1.1
BIC	2	5.3	0.3	1.0	0.5	0.0	0.0	0.0
HQC		3.9	0.1	0.2	0.2	0.0	0.0	0.0
AIC		0.8	0.5	2.6	2.9	1.8	1.1	1.3
BIC	3	0.7	0.2	0.1	0.0	0.0	0.0	0.0
HQC		0.2	0.0	0.0	0.0	0.0	0.0	0.0
AIC		1.0	0.9	2.2	2.2	3.9	2.8	2.2
BIC	4	0.3	0.2	0.3	0.0	0.0	0.0	0.0
HQC		0.1	0.1	0.1	0.0	0.0	0.0	0.0
AIC		0.7	1.4	1.9	2.0	3.5	3.8	3.1
BIC	5	0.1	0.0	0.0	0.0	0.0	0.1	0.0
HQC		0.1	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.6	1.2	1.5	1.3	1.6	2.5	4.1
BIC	6	0.0	0.0	0.1	0.1	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.7	1.1	1.3	2.1	2.1	1.7	3.5
BIC	7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0

Tableau (2.8.30)

Estimation des paramètres du modèle ARMA(2,1) par la méthode des moindre carrés.

paramètres / taille	50	100	150	200
$\phi_1 = 0.777$	0.7842 (0.1816)	0.7817 (0.1187)	0.7800 (0.0942)	0.7819 (0.0822)
$\phi_2 = -0.492$	-0.5036 (0.1549)	-0.4940 (0.1062)	-0.4926 (0.0880)	-0.4955 (0.0747)
$\theta = 0.5$	0.4771 (0.2027)	0.4884 (0.1343)	0.4930 (0.1064)	0.4899 (0.0857)
$\sigma^2 = 1$	0.9498 (0.2027)	0.9693 (0.1393)	0.9833 (0.1173)	0.9854 (0.0992)

Tableau (2.8.31)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p=2,\ \ q=1,\ \ \phi_1=0.777,\ \ \phi_2=-0.492\ \ \theta=0.5,\ \ \sigma^2=1,\ \ N=50$								
	p/q	1	2	3	4	5	6	7
AIC		0.5	2.1	1.6	0.9	1.2	2.1	1.7
BIC	1	10.3	8.7	4.4	2.5	2.0	1.3	0.6
HQC		25.5	8.4	3.7	1.8	0.5	0.1	0.0
AIC		4.6	1.9	3.3	3.3	2.8	2.7	2.3
BIC	2	32.9	5.4	4.2	3.4	1.5	0.9	0.6
HQC		43.1	4.8	1.8	0.6	0.4	0.1	0.0
AIC		1.0	1.2	1.5	1.3	0.9	2.4	1.3
BIC	3	3.7	1.5	1.4	0.5	0.3	0.4	0.1
HQC		3.8	0.7	0.5	0.0	0.0	0.2	0.1
AIC		0.8	3.8	3.5	2.6	3.0	2.2	2.7
BIC	4	1.4	2.6	1.6	0.3	0.7	0.4	0.0
HQC		0.7	1.1	0.5	0.2	0.0	0.1	0.0
AIC		0.5	3.5	2.9	1.5	0.9	1.5	1.6
BIC	5	0.7	1.4	0.8	0.0	0.0	0.0	0.0
HQC		0.3	0.8	0.0	0.0	0.0	0.0	0.0
AIC		0.4	3.5	2.7	3.0	1.6	2.1	2.2
BIC	6	0.2	1.3	0.0	0.3	0.1	0.0	0.0
HQC		0.1	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.3	2.1	1.8	3.2	2.2	1.6	1.7
BIC	7	0.1	0.7	0.2	0.2	0.1	0.1	0.2
HQC		0.0	0.1	0.0	0.0	0.0	0.0	0.0

Tableau (2.8.32)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p=2,\ q=1,\ \phi_1=0.777,\ \phi_2=-0.492\ \theta=0.5,\ \sigma^2=1,\ N=100$								
	p/q	1	2	3	4	5	6	7
AIC		0.1	1.4	0.7	1.4	1.1	0.3	0.5
BIC	1	3.0	7.2	2.8	1.5	0.8	0.3	0.3
HQC		7.3	7.2	1.6	0.6	0.2	0.1	0.0
AIC		7.8	1.1	1.4	1.6	3.0	2.5	1.9
BIC	2	63.5	1.9	1.5	1.4	0.8	0.4	0.3
HQC		73.8	1.0	1.0	0.4	0.3	0.0	0.0
AIC		2.2	1.4	0.6	1.8	2.4	2.4	1.8
BIC	3	6.1	0.8	0.0	0.5	0.0	0.0	0.0
HQC		4.3	0.3	0.0	0.1	0.0	0.0	0.0
AIC		1.5	1.8	4.0	2.1	2.7	1.9	2.8
BIC	4	1.8	0.7	0.5	0.1	0.0	0.1	0.1
HQC		1.1	0.2	0.0	0.0	0.0	0.0	0.0
AIC		1.0	2.8	3.1	3.3	2.4	1.5	2.2
BIC	5	0.6	1.0	0.5	0.1	0.1	0.0	0.0
HQC		0.4	0.1	0.0	0.0	0.0	0.0	0.0
AIC		0.2	2.1	3.1	2.9	2.1	2.6	2.2
BIC	6	0.1	0.5	0.2	0.0	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.5	1.3	2.0	3.9	3.1	2.4	1.1
BIC	7	0.2	0.0	0.2	0.1	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0

Tableau (2.8.33)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p=2,\ \ q=1,\ \ \phi_1=0.777,\ \ \phi_2=-0.492\ \ \theta=0.5,\ \ \sigma^2=1,\ \ N=150$								
	p/q	1	2	3	4	5	6	7
AIC		0.0	0.5	0.6	0.4	0.8	0.4	0.2
BIC	1	0.5	5.7	2.4	0.7	0.1	0.1	0.0
HQC		1.6	6.0	1.6	0.3	0.1	0.0	0.0
AIC		9.4	1.9	0.9	1.1	2.3	3.0	2.0
BIC	2	75.5	3.1	0.9	0.5	0.6	0.3	0.1
HQC		83.1	1.7	0.7	0.2	0.1	0.0	0.0
AIC		1.5	1.6	0.3	1.4	1.0	1.9	2.3
BIC	3	5.7	0.2	0.0	0.1	0.0	0.0	0.0
HQC		3.4	0.0	0.0	0.0	0.0	0.0	0.0
AIC		1.5	1.6	4.0	2.4	2.5	2.6	3.0
BIC	4	0.8	0.5	0.4	0.1	0.0	0.1	0.0
HQC		0.4	0.2	0.2	0.0	0.0	0.0	0.0
AIC		1.1	2.6	3.1	1.9	1.4	2.4	2.3
BIC	5	0.2	0.6	0.2	0.0	0.0	0.0	0.0
HQC		0.1	0.2	0.0	0.0	0.0	0.0	0.0
AIC		0.1	2.8	2.5	4.0	3.3	2.5	3.6
BIC	6	0.0	0.5	0.0	0.1	0.0	0.0	0.0
HQC		0.0	0.1	0.0	0.0	0.0	0.0	0.0
AIC		0.4	1.8	2.5	3.1	2.9	2.4	2.2
BIC	7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0

Tableau (2.8.34)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p=2,\ q=1,\ \phi_1=0.777,\ \phi_2=-0.492\ \theta=0.5,\ \sigma^2=1,\ N=200$								
	p/q	1	2	3	4	5	6	7
AIC		0.0	0.0	0.7	0.7	0.6	0.5	0.6
BIC	1	0.0	2.6	1.5	0.8	0.0	0.0	0.0
HQC		0.2	3.0	0.9	0.1	0.0	0.0	0.0
AIC		13.4	2.1	0.6	0.7	3.3	2.1	1.6
BIC	2	84.7	2.0	0.4	0.1	0.2	0.1	0.2
HQC		90.2	1.2	0.3	0.0	0.1	0.0	0.0
AIC		2.7	1.1	0.4	0.6	1.1	1.5	0.9
BIC	3	4.1	0.5	0.1	0.1	0.0	0.0	0.0
HQC		3.1	0.0	0.1	0.0	0.0	0.0	0.0
AIC		1.6	1.7	4.5	2.5	1.2	2.8	2.8
BIC	4	1.0	0.3	0.3	0.0	0.1	0.0	0.0
HQC		0.6	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.9	1.3	2.5	3.3	1.5	1.2	2.3
BIC	5	0.5	0.1	0.1	0.0	0.1	0.0	0.0
HQC		0.2	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.7	2.6	3.2	1.8	3.1	2.0	2.3
BIC	6	0.1	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.1	1.4	2.9	4.1	4.6	3.4	2.5
BIC	7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0

Tableau (2.8.35)

Estimation des paramètres du modèle ARMA(1,2) par la méthode des moindre carrés.

paramètres / taille	50	100	150	200
$\phi=0.15$	0.2463 (0.4074)	0.2063 (0.3101)	0.1958 (0.2559)	0.1710 (0.2023)
$\theta_1=0.44$	0.3378 (0.4474)	0.3822 (0.3261)	0.3964 (0.2594)	0.4235 (0.2052)
$\theta_2=0.35$	0.2499 (0.2956)	0.2771 (0.1888)	0.2760 (0.1551)	0.2947 (0.1198)
$\sigma^2=0.9$	0.8465 (0.1818)	0.8787 (0.1278)	0.8848 (0.1079)	0.8886 (0.0927)

Tableau (2.8.36)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p=1,\ q=2,\ \phi=0.15,\ \theta_1=0.44,\ \theta_2=0.35,\ \sigma^2=1,\ N=50$								
	p/q	1	2	3	4	5	6	7
AIC		4.6	4.0	2.4	1.7	0.5	0.8	1.6
BIC	1	30.0	16.7	5.2	1.9	0.8	0.8	0.3
HQC		50.6	19.2	3.2	1.0	0.2	0.1	0.0
AIC		2.7	1.9	2.6	5.4	3.0	3.1	1.2
BIC	2	14.2	3.2	2.9	3.3	1.1	0.9	0.4
HQC		15.4	1.8	1.2	0.8	0.2	0.0	0.0
AIC		1.4	1.0	3.2	1.6	1.7	1.8	1.1
BIC	3	3.2	1.0	1.3	0.4	0.2	0.2	0.1
HQC		2.9	0.3	0.2	0.1	0.1	0.0	0.0
AIC		0.5	2.5	3.6	1.5	2.3	2.6	2.4
BIC	4	1.2	2.5	1.0	0.2	0.3	0.5	0.2
HQC		0.6	0.7	0.3	0.1	0.1	0.0	0.0
AIC		0.3	2.4	2.1	1.2	2.1	1.5	1.4
BIC	5	0.4	1.4	0.4	0.0	0.2	0.2	0.1
HQC		0.2	0.1	0.0	0.0	0.0	0.0	0.0
AIC		0.1	2.4	1.8	1.6	2.0	1.7	2.8
BIC	6	0.1	1.2	0.3	0.3	0.0	0.0	0.0
HQC		0.1	0.2	0.1	0.0	0.0	0.0	0.0
AIC		0.3	2.1	2.8	2.3	2.5	2.2	1.7
BIC	7	0.1	0.4	0.3	0.3	0.1	0.1	0.1
HQC		0.0	0.2	0.0	0.0	0.0	0.0	0.0

Tableau (2.8.37)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p = 1, \quad q = 2, \quad \phi = 0.15, \quad \theta_1 = 0.44, \quad \theta_2 = 0.35, \quad \sigma^2 = 1, \quad N = 100$								
	p/q	1	2	3	4	5	6	7
AIC		2.5	3.7	2.5	1.4	0.4	1.2	0.7
BIC	1	30.2	27.8	6.4	2.0	0.4	0.3	0.1
HQC		46.4	25.9	4.8	0.5	0.1	0.1	0.1
AIC		3.3	1.0	0.9	6.5	3.7	2.7	2.9
BIC	2	18.9	1.0	0.8	2.9	0.6	0.6	0.1
HQC		17.8	0.2	0.3	0.6	0.0	0.1	0.0
AIC		1.1	1.5	2.4	2.2	3.2	1.5	1.3
BIC	3	2.8	0.6	0.7	0.1	0.1	0.0	0.0
HQC		1.8	0.2	0.2	0.1	0.0	0.0	0.0
AIC		0.5	0.6	3.0	2.2	1.8	4.5	2.3
BIC	4	0.8	0.4	0.4	0.1	0.3	0.1	0.0
HQC		0.3	0.3	0.0	0.0	0.0	0.0	0.0
AIC		0.5	0.4	1.2	2.0	3.6	2.6	2.2
BIC	5	0.3	0.0	0.3	0.1	0.1	0.0	0.1
HQC		0.1	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.5	1.4	2.4	2.0	2.4	3.1	1.7
BIC	6	0.2	0.0	0.0	0.1	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.4	1.5	2.4	1.7	2.8	1.0	2.7
BIC	7	0.0	0.2	0.1	0.0	0.0	0.0	0.0
HQC		0.0	0.0	0.1	0.0	0.0	0.0	0.0

Tableau (2.8.38)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p=1,\ q=2,\ \phi=0.15,\ \theta_1=0.44,\ \theta_2=0.35,\ \sigma^2=1,\ N=150$								
	p/q	1	2	3	4	5	6	7
AIC		2.2	6.1	2.7	1.2	0.8	0.1	0.3
BIC	1	28.0	33.0	5.0	0.5	0.2	0.1	0.0
HQC		37.9	33.2	3.1	0.4	0.0	0.0	0.0
AIC		3.0	0.7	1.1	6.5	3.8	2.5	2.7
BIC	2	20.2	0.6	0.5	3.0	0.4	0.5	0.1
HQC		21.1	0.2	0.1	0.7	0.0	0.2	0.0
AIC		1.2	0.5	2.2	2.4	3.0	2.7	1.5
BIC	3	3.5	0.5	1.0	0.2	0.2	0.2	0.0
HQC		2.2	0.1	0.2	0.0	0.0	0.0	0.0
AIC		1.4	0.3	2.8	2.0	2.3	4.3	3.1
BIC	4	0.7	0.1	0.3	0.2	0.0	0.0	0.1
HQC		0.2	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.3	0.2	1.9	2.1	4.1	3.0	3.6
BIC	5	0.2	0.0	0.0	0.2	0.2	0.0	0.1
HQC		0.1	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.0	0.5	2.0	1.3	3.6	2.2	2.8
BIC	6	0.1	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.2	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.0	0.2	1.2	1.0	1.6	2.0	3.0
BIC	7	0.0	0.1	0.0	0.0	0.0	0.0	0.0
HQC		0.0	0.1	0.0	0.0	0.0	0.0	0.0

Tableau (2.8.39)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p = 1,\ q = 2,\ \phi = 0.15,\ \theta_1 = 0.44,\ \theta_2 = 0.35,\ \sigma^2 = 1,\ N = 200$								
	p/q	1	2	3	4	5	6	7
AIC		1.1	5.0	2.9	1.6	0.5	0.5	0.6
BIC	1	21.3	38.4	5.7	1.0	0.0	0.0	0.1
HQC		30.0	37.6	3.3	0.6	0.0	0.0	0.0
AIC		4.9	1.3	0.1	7.3	3.5	3.2	1.9
BIC	2	25.0	0.7	0.1	1.6	0.3	0.1	0.1
HQC		25.0	0.4	0.0	0.9	0.0	0.0	0.0
AIC		0.9	0.7	1.7	2.3	3.2	1.4	1.9
BIC	3	3.3	0.3	0.5	0.3	0.1	0.0	0.0
HQC		1.9	0.1	0.1	0.0	0.0	0.0	0.0
AIC		0.9	0.4	3.2	1.6	1.5	4.6	5.2
BIC	4	0.5	0.0	0.1	0.0	0.1	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.5	0.0	1.8	1.8	2.9	2.7	3.4
BIC	5	0.1	0.1	0.0	0.0	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.5	0.6	1.6	1.1	3.3	2.3	2.7
BIC	6	0.0	0.0	0.0	0.0	0.0	0.0	0.1
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0
AIC		0.3	0.3	1.7	1.1	2.2	1.5	3.8
BIC	7	0.1	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.1	0.0	0.0	0.0	0.0	0.0	0.0

Tableau (2.8.40)

Avant de commenter sur les résultats ci-dessus, nous rappelons qu'une des raison pour ne pas utiliser la méthode du coin pour choisir les ordres d'un modèle autoregressif moyenne mobile, est la difficulté de tester simultanément la nullité du déterminant d'autocorrélation $\Delta(k, i)$ pour $i \geq q+1$ et $k \geq p+1$. Pour plus de clarté, nous proposons quelques résultats de simulation pour le dernier modèle.

Le tableau des $\Delta(p, q)$: D'après le théorème du coin le modèle est un ARMA(4,2)

$p = 1,\ q = 2,\ \phi = 0.15,\ \theta_1 = 0.44,\ \theta_2 = 0.35,\ \sigma^2 = 1,\ N = 100$							
$p\ /\ q$	1	2	3	4	5	6	7
1	-0.006	-0.567	-0.043	0.107	-0.089	0.099	0.240
2	0.567	0.321	0.062	0.008	-0.003	0.031	0.073
3	-0.052	-0.120	-0.034	0.002	0.003	0.012	0.010
4	0.148	0.039	0.023	0.012	0.007	0.004	0.002
5	-0.077	0.015	-0.001	-0.003	0.001	0.000	0.000
6	0.012	0.004	0.002	0.001	0.000	0.000	0.000
7	0.006	-0.001	-0.001	0.000	0.000	0.000	0.000

Le tableau des $\Delta(p,q)$: D'après le théorème du coin le modèle est un ARMA(2,2)

$p=1,\ q=2,\ \phi=0.15,\ \theta_1=0.44,\ \theta_2=0.35,\ \sigma^2=1,\ N=100$							
$p\ /\ q$	1	2	3	4	5	6	7
1	-0.054	-0.233	-0.182	-0.058	0.146	-0.050	-0.040
2	0.236	0.045	0.020	0.030	0.018	0.008	-0.001
3	-0.209	0.011	0.005	-0.009	0.001	-0.002	0.000
4	0.141	0.027	0.007	0.003	-0.001	0.000	0.000
5	0.015	-0.017	-0.006	-0.001	0.000	0.000	0.000
6	0.091	0.014	0.001	0.000	0.000	0.000	0.000
7	-0.040	-0.004	0.000	0.000	0.000	0.000	0.000

Tableau (2.8.42)

Le tableau des $\Delta(p,q)$: D'après le théorème du coin le modèle est un ARMA(1,2)

$p=1,\ q=2,\ \phi=0.15,\ \theta_1=0.44,\ \theta_2=0.35,\ \sigma^2=1,\ N=150$							
$p\ /\ q$	1	2	3	4	5	6	7
1	-0.086	-0.370	-0.072	-0.007	0.029	0.058	0.005
2	0.377	0.130	0.003	0.002	0.001	0.003	0.004
3	-0.149	-0.043	0.004	0.000	0.000	0.000	0.000
4	0.155	0.018	0.002	0.000	0.000	0.000	0.000
5	-0.075	0.001	0.000	0.000	0.000	0.000	0.000
6	0.035	-0.001	0.000	0.000	0.000	0.000	0.000
7	-0.023	-0.002	0.000	0.000	0.000	0.000	0.000

Tableau (2.8.43)

Le tableau des $\Delta(p,q)$: D'après le théorème du coin le modèle est un ARMA(1,2)

$p=1,\ q=2,\ \phi=0.15,\ \theta_1=0.44,\ \theta_2=0.35,\ \sigma^2=1,\ N=200$							
$p\ /\ q$	1	2	3	4	5	6	7
1	-0.153	-0.308	-0.056	-0.009	0.042	-0.002	0.026
2	0.332	0.086	0.000	0.003	0.002	0.001	0.001
3	-0.168	-0.024	0.004	0.000	0.000	0.000	0.000
4	0.145	0.014	0.000	0.000	0.000	0.000	0.000
5	-0.676	-0.007	0.000	0.000	0.000	0.000	0.000
6	0.058	0.005	0.000	0.000	0.000	0.000	0.000
7	-0.019	0.002	0.000	0.000	0.000	0.000	0.000

Tableau (2.8.44)

Nous allons maintenant, après avoir présenté les résultats de simulation dans les tableaux ci-dessus, donner quelques commentaires sur la performance de chaque méthode.

Résultats concernant les modèles autoregressifs pur

• Modèle AR(1) d'équation stochastique $(1 - 0.5L)y_t = \varepsilon_t$: la racine du polynôme caractéristique est égale à 2. Nous remarquons que le critère PDC est le meilleur (il a les plus grandes fréquences) car il estime le vrai ordre dans 99.6% des cas, pour $N = 200$. D'autre part, le critère FPE est le moins performant, avec une estimation de la vraie valeur de l'ordre inférieure à 50% des cas, pour $N = 50$.

• Modèle AR(1) d'équation stochastique $(1 - 0.7L)y_t = \varepsilon_t$: la racine du polynôme caractéristique est égale à 1.429. Nous remarquons que le critère PDC est toujours le meilleur car il estime le vrai ordre dans plus 95% des cas même pour des petites tailles de l'échantillon. Ainsi, le FPE reste le moins performant, avec une estimation de la vraie valeur de l'ordre dans moins de 50% des cas, pour $N = 50$.

• Modèle AR(1) d'équation stochastique $(1 - 0.95L)y_t = \varepsilon_t$: la racine du polynôme caractéristique est égale à 1.053. Nous remarquons que le critère PDC est encore le plus performant par plus de 95%. Nous remarquons également que, les critères HQC et BIC sont des critères consistents, avec une estimation du vrai ordre dans plus de 93% ,84% des cas,respectivement.

• Modèle AR(2) d'équation stochastique $(1 - 0.44L - 0.35L^2)y_t = \varepsilon_t$: les racines du polynôme caractéristique sont -2.432 et 1.179. Nous constatons que pour $N = 50$, tous les critère estiment le vrai ordre dans moins de 50% des cas et que le HQC, CIC et PDC sous-estime le vrais ordre dans plus de 50% des cas. Par contre le FPE et AIC sur-estime le vrai ordre dans plus de 40% des cas. Nous remarquons également, que les critères BIC, HQC, CIC et PDC sont des critères fortement consistants, avec une estimation du vrai ordre dans plus de 93% des cas, pour $N = 200$.

• Modèle AR(2) d'équation stochastique $(1 - 0.777L + 0.492L^2)y_t = \varepsilon_t$: les racines du polynôme caractéristique sont $0.790 \pm 1.187i$. Nous remarquons que les critères FPE et AIC sur-estime le vrai ordre dans plus de 40% des cas. Nous constatons également, que les critères BIC, HQC, CIC et PDC sont des critères fortement consistents, car ils estiment la vraie valeur de l'ordre dans plus de 93% des cas, pour $N = 150$. D'autre part, nous observons que pour une taille assez grande les résultats donnés par les critères PDC, HQC, BIC et CIC sont plus ou moins proches.

Résultats concernant les modèles autoregressifs moyenne mobiles

• Modèle ARMA(1,1) d'équation stochastique $(1 - 0.2L)y_t = (1 - 0.7L)\varepsilon_t$: la racine du polynôme autoregressif est égale à 5 et celle du polynôme moyenne mobile est égale à 1.429. Nous remarquons que le critère HQC est le meilleur car il estime les vrais ordres dans 75% des cas, pour $N = 50$ et 88.6% pour $N = 200$. Cependant, le BIC est moins performant avec 54.1% pour $N = 50$ et

80.8% pour $N = 200$. D'autre part le critère AIC a les plus faibles fréquences, avec une estimation des vrais ordre dans moins de 10% des cas.

• Modèle ARMA(2,1) d'équation stochastique $(1 - 0.777L + 0.492L^2)y_t = (1 - 0.5L)\varepsilon_t$: les racine du polynôme autoregressif sont $0.790 \pm 1.187i$ et celle du polynôme moyenne mobile est égale à 2. Nous constatons cette fois-ci que les trois critères estiment les vrais ordres dans moins de 50% des cas, pour $N = 50$, ils tendent de plus en plus à la sélection des vraies valeurs des ordres du modèle lorsque la taille de l'échantillon augmente.

• Modèle ARMA(1,2) d'équation stochastique $(1 - 0.15L)y_t = (1 - 0.44L - 0.35L^2)\varepsilon_t$: la racine du polynôme autoregressif est égale à 6.667 et les racines du polynôme moyenne mobile sont -2.432 et 1.175. Nous observons dans le cas de ce modèle que les trois critères estiment les vrais ordres dans moins de 40% des cas même lorsque $N = 200$.

En résumé, Nous constatons que les meilleurs résultats sont donnés par les critères PDC, HQC, BIC et CIC dans le cas d'un modèle autoregressif pur et par les critères HQC et BIC dans le cas autoregressif moyenne mobile, ce que nous retrouvons dans la littérature des séries chronologiques.

Nous constatons également, que les critères PDC et HQC donnent toujours les meilleurs résultats même dans le cas des échantillons de petites tailles. Par contre, les critères FPE et AIC ont une tendance à sur-estimer le vrai ordre du modèle, ce qui est prouvé théoriquement (théorème de Sibata (1976) et théorème (2.3.1) de Hannan (1982)). Néanmoins, lorsque la taille est assez grande, les résultats obtenus à partir des critères PDC, HQC, BIC et CIC sont très proches et fortement consistants.

Notons que le temps de calcul et l'espace mémoire sont beaucoup plus grand pour le critère PDC que pour les autres. Par conséquent, d'un point de vue pratique, les critères HQC, BIC et CIC sont meilleurs que le PDC.

Dans le cas d'un modèle mixte, nous remarquons que les critères HQC et BIC restent toujours très proches et que plus la taille de l'échantillon augmente, plus la fréquence de choisir les vraies valeurs des ordres par ces critères augmente. Par conséquent, les estimateurs obtenus par les critères HQC et BIC sont fortement consistants, ce qui est connu dans la littérature des séries chronologiques.

Deuxième partie

Identification des modèles ARMA d-périodiques

Chapitre 3

Modèle autoregressif moyenne mobile périodique (PARMA)

Introduction

Dans la pratique, de nombruses séries chronologiques présentent un caractère non régulier et ne peuvent être rendus stationnaires par des transformations tendancielles, saisonnières ou mixtes. L'utilisation des modèles classiques de Box et Jenkins (1976) pour traiter ce genre de séries serait donc insuffisante et inefficace. Il est claire que l'utilisation des modèles non stationnaires, bien qu'elle soit plus astreignante et délicate, serait plus intéressante car il existe des phénomènes aléatoires qui ne peuvent être représenter par les modèles ARIMA classiques de Box et Jenkins (1976), mais plutôt par des modèles linéaires à coefficients évolutifs dans le temps. En effet, la généralisation du théorème de Wold (1938) par Cramer (1961) aux processus non stationnaire, a permit d'élargir la classe des modèles linéaires autoregressifs moyenne mobile, à coefficients constants (ARMA), à la classe des modèles linéaires autoregressifs moyenne mobile à coefficients dépendant du temps (ARMA_t).

En 1961 Gladyshev a introduit, pour la première fois dans la littérature, la notion des processus périodiquement corrélés. Ce sont des processus du second ordre dont la moyenne et la fonction d'autocorrélation sont des fonctions périodiques dans le temps. L'avantage de cette nouvelle classe de processus est dans la relation, établit par le théorème de Gladyshev (1961), entre ces processus et les processus stationnaires multivariés. En effet, ils peuvent être exploités dans l'analyse des séries chronologiques multivariées stationnaires dans le but de réduire sensiblement le nombre de paramètres.

Ce chapitre est consacré à l'étude d'une classe particulière des modèles ARMA_t et qui est souvent très utile pour la modélisation des séries saisonnières est celle dont les paramètres sont périodiques dans le temps. Cette classe a connue dans ces dernières années des progrés considérables (voir par exemple, Pagano (1978), Clevland et Tiao (1979), Tiao et Grupe (1980), Vecchia (1985a et 1985b), Vecchia et Ballerini (1992), Anderson et Vicchia (1993), Adams et Goodwin (1995), Bentarzi et Hallin (1994), (1996) et (1998), Benterzi (1998), Franses et Klock (1995)).

Nous commencerons ce chapitre par définir la notion de processus périodiquement corrélés (processus périodique au sens strict et processus périodiquement corrélé) et la généralisation du théorème de Wold (1938) par Cramer (1961) aux processus non stationnaire. Puis nous allons définir la classe des processus autoregressifs d-périodique (PAR_d), moyennes mobiles d-périodique (PMA_d) et les processus ARMA d-périodique. Nous étudierons ensuite, les propriétés théoriques des modèle ARMA périodique. La cinquième section sera consacrée à l'étude de l'estimation des paramètres d'un modèle autoregressif moyenne mobile d-périodique. Nous exposerons à la fin de ce chapitre l'estimation de la fonction d'autocovariance d'un processus périodiquement corrélé.

3.1 Processus périodiquement corrélés

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus du second ordre de moyenne $\mu_t = E(y_t)$, $t \in \mathbb{Z}$ et de fonction d'autocovariance

$$\gamma(t,s) = cov(y_t, y_s), \quad t, s \in \mathbb{Z}$$

3.1.1 Définition d'un processus périodique au sens strict

Définition (3.1.1) Le processus $\{y_t, t \in \mathbb{Z}\}$ est dit strictement périodique (ou périodique au sens fort) s'il existe un entier d positif tel que pour tout entier positif k et pour toute suite $t_1, t_2, ..., t_k$, la distribution conjointe de $(y_{t_1}, y_{t_2}, ..., y_{t_k})$ est la même que celle de $(y_{t_1+d}, y_{t_2+d}, ..., y_{t_k+d})$.

Le plus petit entier d positif vérifiant cette proprieté est appelé période du processus.

3.1.2 Définition d'un processus périodiquement corrélé

Définition (3.1.2) Le processus $\{y_t, t \in \mathbb{Z}\}$ est dit périodiquement corrélé (ou périodique au sens faible) s'il existe un entier d positif tel que :

$$\mu_{t+d\tau} = \mu_t, \quad \forall t, \tau \in \mathbb{Z}$$
$$\gamma(t + d\tau, s + d\tau) = \gamma(y_t, y_s), \quad t, s, \tau \in \mathbb{Z}$$

c'est-à-dire, si la moyenne et la fonction d'autocovariance sont périodiques dans le temps.

3.1.3 Processus bruit blanc périodique

Les processus bruits blancs stationnaires sont souvent utilisés dans l'analyse des séries chronologiques pour générer d'autres processus stationnaires. De manière analogue les processus bruits blancs d-périodiques constituent la classe élémentaire de l'ensemble des processus périodiquement corrélés. En effet, nous verrons par la suite que tout processus du second ordre peut s'écrire comme une somme pondérée de bruits blancs plus un processus purement déterminable (théorème de Wold-Cramer).

Définition (3.1.3)
Nous dirons que le processus $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un processus bruit blanc périodique s'il vérifi les propriétés suivantes :

i) Le processus $\{\varepsilon_t, t \in \mathbb{Z}\}$ est centré, c'est-à-dire que $E(\varepsilon_t) = 0$ pour tout $t \in \mathbb{Z}$.

ii) Sa variance est d-périodique, c'est-à-dire que $\sigma^2_{i+d\tau} = \sigma^2_i$ pour tout $i = 1, ..., d, \quad t \in \mathbb{Z}$.

iii) Le processus $\{\varepsilon_t, t \in \mathbb{Z}\}$ est non corrélé, c'est-à-dire $E(\varepsilon_t \varepsilon_s) = 0$, pour tout t et s tel que $t \neq s$.

Nous remarquons que les processus bruits blancs d-périodique sont tout simplement des processus bruits blancs de variance d-périodique dans le temps.

3.2 Théorème de Wold-Cramer

Les modèles linéaires non stationnaires à coefficients dépendants du temps ont connu un développement très important dans ces dernières décennies et son souvent utiliser malgré leur complexité par rapport aux modèles linéaires stationnaires. Ce développement est dû à la généralisation du théorème de Wold (1938) par Cramer (1961) de la décomposition des processus stationnaires aux processus non stationnaires. Cette généralisation a comme conséquence l'extension de la classe des modèles linéaires ARMA à coefficients constants aux modèles linéaires ARMA non-stationnaire, à coefficients dépendants du temps, dont les modéles autoregressifs moyennes mobiles périodiquement corrélés sont des cas particuliers.

3.2.1 Théorème de la décomposition de Wold-Cramer

L'énoncé du théorème de Wold-Cramer est le suivant :

Théorème (3.2.1) (***Théorème de la décomposition de Wold-Cramer***)
Tous processus stochastique, du second ordre $\{y_t, t \in \mathbb{Z}\}$ possède une décomposition linéaire unique donnée par :

$$y_t = U_t + V_t \tag{3.2.1}$$

tel que :

i) Les processus U_t et V_t appartiennent au sous-espace de Hilbert $H^2(y_t, t)$.

ii) Ces processus sont orthogonaux. De plus, le processus U_t est la composante linéaire déterministe (U_t est un processus singulier) et V_t désigne la composante linéaire stochastique du processus y_t (V_t est un processus régulier) et peut être représentée par une somme pondérée infinie unique et convergente (en moyenne quadratique) de la forme

$$V_t = \sum_{s=-\infty}^{t} a_{t,s}\varepsilon_s \tag{3.2.2}$$

où ε_t est un processus bruit blanc. Le processus $\{a_{t,t}\varepsilon_t, t \in \mathbb{Z}\}$ est alors dit innovation du processus et le processus $\{\varepsilon_t, t \in \mathbb{Z}\}$ est l'innovation normée.

3.2.2 Décomposition de Wold-Cramer dans la pratique

Nous remarquons, d'après le théorème de Wold-Cramer (1961), que tout processus du second ordre peut être exprimer comme étant une somme pondérée infinie convergente, en moyenne quadratique, de chocs passés. Pour toute cette classe de processus, la décomposition de Wold-Cramer est donc une première représentation possible. Toute-fois, cette représentation n'est pas pratique.

En effet, nous verrons dans le dernier chapitre, que lorsque nous cherchons à identifier une série chronologique, nous appliquons toujours un *principe de parcimonie*, i.e., nous adoptons en priorité la représentation nécessitant l'estimation du minimum de paramètres. Or par définition, si l'on devait appliquer la décomposition de Wold-Cramer, cela supposerait que nous estimons une infinité de paramètres (les $a_{t,s}$ et la variance σ_t^2). Donc, dans la pratique, il convient de rechercher un ordre fini, tel que cette série chronologique puisse être approximée par une somme pondérée de chocs passés jusqu'à cet ordre.

3.3 Modèle autoregressif moyenne mobile périodique (PARMA)

Parmis les représentations les plus utilisées, pour représenter les processus du second ordre d-périodique, figurent les représentations autoregressive moyenne mobile d-périodique (PARMA). Cette représentation consiste en l'adjonction d'une composante autorégressive d-périodique d'ordre fini (PAR) et d'une composante moyenne mobile d-périodique d'ordre fini (PMA).

Nous allons donc commencer par définir la classe des processus PAR, PMA et PARMA. Nous étudierons ensuite sous quelles conditions ces processus satisfont l'hypothèse de causalité et d'inversibilité.

Nous supposons sans perdre de généralité que $\mu_t = 0, t \in \mathbb{Z}$.

3.3.1 Définitions

3.3.1.1 Processus moyenne mobile périodique (PMA)

Définition (3.3.1) Le processus périodiquement corrélé $\{y_t, t \in \mathbb{Z}\}$ d-periodique admet une représentation moyenne mobile périodique d'ordre q_t, noté $\mathrm{MA}_d(p_t)$, s'il est solution de l'équation aux difference stochastique suivante :

$$y_t = \varepsilon_t - \sum_{j=1}^{q_t} \theta_{t,j}\varepsilon_{t-j} \tag{3.3.1}$$

ou encore

$$y_t = \Theta_t(L)\varepsilon_t$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un bruit blanc périodique de variance σ_t^2, les paramètres $\theta_{i,j},\ j = 1, ..., q_t$ et q_t sont des fonctions d-périodiques, en t et le polynôme $\Theta_t(L)$ est égal à $1 - \sum_{i=1}^{q_t} \theta_{t,i} L^i$.

3.3.1.2 Processus autoregressif périodique (PAR)

Définition (3.3.2) Le processus périodiquement corrélé $\{y_t, t \in \mathbb{Z}\}$ d-periodique admet une représentation autoregressive périodique d'ordre p_t, noté $\mathrm{PAR}_d(p_t)$, s'il est solution de l'équation aux difference stochastique suivante :

$$y_t - \sum_{j=1}^{p_t} \phi_{t,j} y_{t-j} = \varepsilon_t \tag{3.3.2}$$

ou encore

$$\Phi_t(L)y_t = \varepsilon_t$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un bruit blanc périodique de variance σ_t^2, les paramètres $\phi_{t,j},\ j = 1, ..., p_t$, et p_t sont des fonctions d-périodiques, en t et le polynôme $\Phi_t(B)$ est égal à $1 - \sum_{i=1}^{p_t} \phi_{t,i} L^i$.

3.3.1.3 Processus autoregressif moyenne mobile périodique (PARMA)

Définition (3.3.3) Le processus périodiquement corrélé $\{y_t, t \in \mathbb{Z}\}$ d-periodique admet une représentation autoregressive moyenne mobile périodique d'ordre (p_t, q_t), noté $\mathrm{PARMA}_d(p_t, q_t)$, s'il est solution de l'équation aux difference stochastique suivante :

$$y_t - \sum_{j=1}^{p_t} \phi_{t,j} y_{t-j} = \varepsilon_t - \sum_{j=1}^{q_t} \theta_{t,j}\varepsilon_{t-j} \tag{3.3.3}$$

ou encore

$$\Phi_t(L)y_t = \Theta_t(L)\varepsilon_t$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un bruit blanc périodique de variance σ_t^2, les paramètres $\phi_{t,j},\ j = 1, ..., p_t,\quad \theta_{t,j},\ j = 1, ..., q_t,\ p_t$ et q_t sont des fonctions d-périodiques, en t et les polynômes $\Phi_t(B) = 1 - \sum_{i=1}^{p_t} \phi_{t,i} L^i$ et $\Theta_t(B) = 1 - \sum_{i=1}^{q_t} \theta_{t,i} L^i$ ne possèdent pas de zéros communs.

Posons $p = \max_{1 \leq t \leq d} p_t$ et $q = \max_{1 \leq t \leq d} q_t$ alors, (3.3.3) peut s'écrire comme suit

$$y_t - \sum_{j=1}^{p} \phi_{t,j} y_{t-j} = \varepsilon_t - \sum_{j=1}^{q} \theta_{t,j} \varepsilon_{t-j}. \tag{3.3.4}$$

3.3.2 Inversibilité et causalité des processus ARMA d-périodique

3.3.2.1 Définitions

Définition (3.3.4) Le modèle (3.3.3) est dit causal s'il admet une solution de la forme

$$y_t = \sum_{j=0}^{\infty} \psi_{t,j} \varepsilon_{t-j} \tag{3.3.5}$$

où la series infinie est convergente en noyenne quadratique. L'expression précédente est dite représentation de Wold-Cramer du processus $\{y_t, t \in \mathbb{Z}\}$.

Définition (3.3.5) Le modèle (3.3.3) est dit inversible s'il existe des fonctions $\pi_{t,j}$ telles que :

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_{t,j} y_{t-j} \tag{3.3.6}$$

où la series infinie est convergente, en noyenne quadratique. L'expression précédente est dite formule de déconvolution.

Pour étudier ces deux propriétés dans le cas des modèles PARMA nous pouvons employer deux approches. La première se base sur le théorème de Gladyshev (1961). Cette approche consiste à ramener un modèle périodique univarié à un modèle multivarié stationnaire et d'étudier les propriétés de ce dernier. La deuxième, c'elle de Bentarzi et Hallin (1994), ramène le modèle moyenne mobile périodique (autoregressif périodique) à un modèle moyenne mobile périodique (autoregressif périodique) d'ordre 1.

3.3.2.2 Approche de "*period-span-lumping*"

A. Modèle ARMA multivarié associé à un modèle PARMA univarié

Théorème de Gladychev

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus centré périodiquement corrélé de période d. Pour tout $t \in \mathbb{Z}$, l'entier $i \in \{1, ..., d\}$ et l'entier τ tel que $t = i + d\tau$ nous définissons le processus d-varié $X_\tau = (X_{\tau,1}, X_{\tau,2}, ..., X_{\tau,d})'$ où $X_{\tau,i} = y_{i+d\tau}$.

Alors, $\{y_t, t \in \mathbb{Z}\}$ est périodiquement corrélé si et seulement si le processus multivarié correspondant est stationnaire.

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus univarié périodiquement corrélé qui admet une représentation autoregressive moyenne mobile d-périodique d'ordre (p_t, q_t), de la forme suivante :

$$y_t - \sum_{j=1}^{p_t} \phi_{t,j} y_{t-j} = \varepsilon_t - \sum_{j=1}^{q_t} \theta_{t,j} \varepsilon_{t-j}.$$

Soient $\Phi(0)$, $\Theta(0)$, $\Phi(k)$, $\Theta(k)$ des matrices carrées de dimensions $d \times d$ définies comme suit :

$$(\Phi(0))_{i,j} = \begin{cases} 1 & i = j \\ 0 & i < j \\ \phi_{i,i-j} & i > j \end{cases}$$

$$(\Theta(0))_{i,j} = \begin{cases} 0 & i < j \\ \theta_{i,i-j} & i \geq j \end{cases}$$

et

$$(\Phi(k))_{i,j} = \phi_{i,kd+i-j}, \qquad 1 \leq k \leq p^*,$$

$$(\Theta(k))_{i,j} = \theta_{i,kd+i-j}, \qquad 1 \leq k \leq p^*.$$

où $\phi_{i,n} = 0 \quad$ pour $n > p_i$ et $\theta_{i,n} = 0 \quad$ pour $n > q_i$.

Proposition (3.3.1)

Le modèle d-varié autoregressif moyenne mobile stationnaire associé au modèle univarié autoregressif moyenne mobile d'ordre (p_t, q_t) d-périodique ci-dessus, est donné par l'équation aux différences stochastique suivante :

$$\sum_{j=0}^{p^*} \Phi(j) X_{\tau-j} = \sum_{j=0}^{q^*} \Theta(j) \eta_{\tau-j} \tag{3.3.7}$$

où $\eta_\tau = (\eta_{\tau,1}, \eta_{\tau,2}, ..., \eta_{\tau,d})'$ avec $\eta_{\tau,i} = \varepsilon_{i+d\tau}, \quad i = 1, ..., d$ est un processus bruit blanc d-varié dont la matrice de covariance est diagonale d'éléments σ_i^2,
$i = 1, ..., d$. Le processus d-varié $\{X_\tau, \tau \in \mathbb{Z}\}$ est défini par :
$X_\tau = (X_{\tau,1}, X_{\tau,2}, ..., X_{\tau,d})'$ avec $X_{\tau,i} = y_{i+d\tau}, i = 1, ..., d$ et l'ordre (p^*, q^*) est donné par :

$$p^* = \max_{1 \leq i \leq d} \left(\frac{p_i - i}{d} \right) + 1, q^* = \max_{1 \leq i \leq d} \left(\frac{q_i - i}{d} \right) + 1.$$

Les matrices $\Phi(k)$, $\Theta(k), k = 0, 1, ..., p^*$ et $r = 0, 1, ..., q^*$ sont comme ci-dessus.

Démonstration Voir par exemple, Bentarzi (1995) p.26-28.

B. Condition de causalité d'un modèle ARMA multivarié

Soit le modèle multivarié à coefficients constants ARMA(p^*, q^*) de la représentation ci-dessous

$$\sum_{j=0}^{p^*} A_j X_{t-j} = \sum_{j=0}^{q^*} B_j \varepsilon_{t-j} \tag{3.3.8}$$

où ε_t est un processus multivarié non corrélé et les matrices A_t et B_t sont des matrices carrées réelles.

Le modèle de représentation ci-dessus est causal si et seulement si les racines de l'équation caractéristique

$$\det\left(\sum_{j=0}^{p^*} A_j z^{p-j}\right) = 0, \quad z \in \mathbb{C} \tag{3.3.9}$$

sont à l'intérieur du cercle unité.

C. Condition d'inversibilité d'un modèle ARMA multivarié

Le modèle multivarié à coefficients constants ARMA(p^*, q^*) de la représentation

$$\sum_{j=0}^{p^*} A_j X_{t-j} = \sum_{j=0}^{q^*} B_j \varepsilon_{t-j}$$

où ε_t est un processus multivarié non corrélé et les matrices A_t et B_t sont des matrices carrées réelles est inversible si et seulement si les racines de l'équation caractéristique

$$\det\left(\sum_{j=0}^{q^*} B_j z^{q-j}\right) = 0, \quad z \in \mathbb{C} \tag{3.3.10}$$

sont à l'intérieur du cercle unité.

D. Condition de causalité d'un modèle AR périodique

Le modèle univarié AR(p_t) d-périodique

$$y_t - \sum_{j=1}^{p_t} \phi_{t,j} y_{t-j} = \varepsilon_t$$

est causal si et seulement si le modèle d-varié autoregressif à coefficients constants qui lui correspond

$$\sum_{j=0}^{p^*} \Phi(j) X_{\tau-j} = \eta_\tau$$

est causal c'est -à-dire, si seulement si les racines de l'équation caractéristique $\det\left(\sum_{j=0}^{p^*}\Phi(j)z^{p-j}\right)=0, \quad z\in\mathbb{C}$ sont à l'intérieur du cercle unité.

E. Condition d'inversibilité d'un modèle MA périodique

Le modèle univarié MA(q_t) d-périodique

$$y_t=\varepsilon_t-\sum_{j=1}^{q_t}\theta_{t,j}\varepsilon_{t-j}$$

est inversible si et seulement si le modèle d-varié moyenne mobile à coefficients constants associé est inversible c'est-à-dire, les racines de l'équation caractéristique $\det\left(\sum_{j=0}^{q^*}\Theta(j)z^{j}\right)=0, \quad z\in\mathbb{C}$ sont à l'intérieur du cercle unité.

3.3.2.3 Approche de "*order-span lumping*"

Nous présentons maintenant une autre approche, dite *order-span lumping* , introduite par Bentarzi et Hallin (1994), pour étudier des propriétés des modèles périodiques. Cette technique consiste a représenté le modèle univarié (m-varié) moyenne mobile d-périodique d'ordre q par un modèle q (mq) varié moyenne mobile S-périodique d'ordre 1, où S est en fonction de d et q. Ainsi, le modèle obtenu est exploité pour étudier les propriétés d'un processus moyenne mobile périodique et en particulier les conditions d'inversibilité. Ula et Smadi (1997) ont employé cette approche pour établir une condition nécessaire et suffisante pour qu'un AR périodique soit causal.

A. Modèle multivarié moyenne mobile périodique d'ordre 1 associé à un modèle multivarié moyenne mobile périodique d'ordre q

Soit $\{y_t, t\in\mathbb{Z}\}$ un processus m-varié moyenne mobile d'ordre q et d-périodique

$$y_t=\sum_{j=0}^{q}\theta_{t,j}\varepsilon_{t-j} \tag{3.3.11}$$

où $\{\varepsilon_t, t\in\mathbb{Z}\}$ est un bruit blanc périodique de variance σ_t^2 les paramètres $\theta_{t,j},\ j=1,...,q$ sont des matrices $m\times m$.

Bentarzi et Hallin (1994) ont défini le processus mq-varié suivant :

$$\begin{aligned}Y_T &= \left(y'_{qT},y'_{qT-1}...,y'_{qT-q+1}\right)',\\ \eta_T &= \left(\varepsilon'_{qT},\varepsilon'_{qT-1}...,\varepsilon'_{qT-q+1}\right)'.\end{aligned}$$

Nous remarquons que le modèle (3.3.11), peut être représenté sous la forme suivante :

$$Y_T=\Theta_{T,0}\eta_T+\Theta_{T,1}\eta_{T-1} \tag{3.3.12}$$

où les matrices $(mq \times mq)$ $\Theta_{T,0}, \Theta_{T,1}$ sont données en fonction des coefficients $\theta_{t,j}, \quad j = 0, ..., q_t$ du (3.3.11) comme suit :

$$\Theta_{T,0} = \begin{pmatrix} \theta_{qT,0} & \theta_{qT,1} & \cdots & \theta_{qT,i-1} & \cdots & \theta_{qT,q-1} \\ 0_m & \theta_{qT-1,0} & \theta_{qT-1,1} & \vdots & \cdots & \theta_{qT-1,q-2} \\ \vdots & 0_m & \ddots & \vdots & \ddots & \vdots \\ 0_m & \vdots & \ddots & \theta_{qT-i+1,0} & \cdots & \theta_{qT-i+1,q-1} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_m & \cdots & \cdots & \cdots & \cdots & \theta_{qT-q+1,q} \end{pmatrix}$$
$$\Theta_{T,1} = \begin{pmatrix} \theta_{qT,q} & 0_m & \cdots & 0_m & \cdots & 0_m \\ \theta_{qT-1,q-1} & \theta_{qT-1,q} & 0_m & \vdots & \cdots & 0_m \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \theta_{qT-i+1,q-i+1} & \vdots & \ddots & \theta_{qT-i+1,q} & \cdots & 0_m \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \theta_{qT-1,0} & \cdots & \cdots & \cdots & \cdots & \theta_{qT-q+1,q} \end{pmatrix} \tag{3.3.13}$$

Le modèle (3.3.12) est un modèle mq-varié moyenne mobile S-périodique d'ordre 1, où S $(S > 1)$ est le plus petit entier tel que qS est le plus petit multiple commun de q et d.

Bentarzi et Hallin (1994) ont montré que :

Proposition (3.3.2) (*Bentarzi et Hallin (1994)*)

Le modèle m-varié moyenne mobile d-périodique d'ordre q est inversible si et seulement si les racines de l'équation caractéristique $|\Psi - Iz| = 0$ sont à l'intérieur du cercle unité, où Ψ est une matrice carrée de dimension $(mq \times mq)$ donnée par :

$$\Psi = \Theta_{S,0}^{-1}\Theta_{S,1}\Theta_{S-1,0}^{-1}\Theta_{S-1,0}...\Theta_{2,0}^{-1}\Theta_{2,1}\Theta_{1,0}^{-1}\Theta_{1,1} \tag{3.3.14}$$

où les matrices $\Theta_{j,0}$ et $\Theta_{j,1}$, $j = 1, ..., S$ sont données par (3.3.13).

Démonstration Voir Bentarzi et Hallin (1994).

Proposition (3.3.3) (*Bentarzi (1995)*)

Le modèle univarié moyenne mobile d-périodique d'ordre 1

$$y_t = \theta_{t,0}\varepsilon_t + \theta_{t,1}\varepsilon_{t-1}$$

est inversible si et seulement si $\left|\frac{\theta_{1,1}\theta_{2,1}...\theta_{d,1}}{\theta_{1,0}\theta_{2,0}...\theta_{d,0}}\right| < 1$.

Démonstration Voir Bentarzi (1995) p.108-110

Par analogie, Ula et Smadi (1997) ont exploité la technique introduite par Bentarzi et Hallin (1994) pour obtenir une condition suffisante pour la stationnarité des modèles autorégressifs moyenne mobiles.

B. Modèle AR multivarié périodique d'ordre 1 associé à un modèle AR multivarié périodique d'ordre p

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus périodiquement corrélé de période d satisfaisant le modèle m-varié autoregressif périodique de période d et d'ordre p et :

$$y_t - \sum_{j=1}^{p} \phi_{t,j} y_{t-j} = \varepsilon_t \tag{3.3.15}$$

Ula et Smadi (1997) ont défini le processus suivant :

$$\begin{aligned} Y_T &= \left(y'_{pT}, y'_{pT-1} ..., y'_{pT-p+1}\right)', \\ \eta_T &= \left(\varepsilon'_{pT}, \varepsilon'_{pT-1} ..., \varepsilon'_{pT-p+1}\right)'. \end{aligned}$$

Le modèle (3.3.15) s'écrira donc sous la forme suivante :

$$\Phi_{T,0} Y_T - \Phi_{T,1} Y_{T-1} = \eta_T \tag{3.3.16}$$

où les matrices $\Theta_{T,0}$ et $\Theta_{T,1}$ de dimension $(mp \times mp)$ sont données en fonction des coefficients périodiques $\phi_{t,j}, \quad j = 1, ..., p$ comme suit :

$$\Phi_{T,0} = \begin{pmatrix} I_m & -\phi_{pT,1} & \cdots & -\phi_{pT,i-1} & \cdots & -\phi_{pT,p-1} \\ 0_m & I_m & -\phi_{pT-1,1} & \vdots & \cdots & \phi_{pT-1,p-2} \\ \vdots & 0_m & \ddots & \vdots & \ddots & \vdots \\ 0_m & \vdots & \ddots & I_m & \cdots & \phi_{pT-i+1,p-1} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_m & \cdots & \cdots & \cdots & 0_m & I_m \end{pmatrix}$$

$$\Phi_{T,1} = \begin{pmatrix} \phi_{pT,p} & 0_m & \cdots & 0_m & \cdots & 0_m \\ \phi_{pT-1,p-1} & \phi_{pT-1,p} & 0_m & \vdots & \cdots & 0_m \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \phi_{pT-i+1,p-i+1} & \vdots & \ddots & \phi_{pT-i+1,p} & \cdots & 0_m \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \phi_{pT-1,0} & \cdots & \cdots & \cdots & \cdots & \phi_{pT-p+1,p} \end{pmatrix} \tag{3.3.17}$$

Proposition (3.3.4) (*Ula et Smadi (1997)*)

Le modèle m-varié autoregressif d-périodique d'ordre q est stationnaire si et seulement si les racines du polynôme $|Iz-\Omega|=0$ sont à l'intérieur du cercle unité, où Ω est la matrice carrée de dimension $(mp\times mp)$ donnée par :

$$\Omega=\Phi_{S,0}^{-1}\Phi_{S,1}\Phi_{S-1,0}^{-1}\Phi_{S-1,1}...\Phi_{2,0}^{-1}\Phi_{2,1}\Phi_{1,0}^{-1}\Phi_{1,1} \tag{3.3.18}$$

où les matrices $\Phi_{j,0}$ et $\Phi_{j,1}$, $j=1,...,S$ sont données par (3.3.17).

Démonstration Voir Ula et Smadi (1997).

3.4 Fonction d'autocovariance et la fonction d'autocorrélation dans le cas périodiques

3.4.1 Définitions

Définition (3.4.1) Soit $\{y_t,t\in\mathbb{Z}\}$ un processus périodiquement corrélé de période d , de moyenne nulle et de fonction d'autocovariance $\gamma(.,.)$. Pour tout entier $t\in\mathbb{Z}, i=1,...,d$ et $\tau\in\mathbb{Z}$ tel que $t=i+d\tau$, alors la fonction d'autocovariance pour l'horizon h et relative à la $i\overset{eme}{=}$ période, notée $\gamma_h^{(i)}, i=1,...,d$ est définie par :

$$\gamma_h^{(i)}=\gamma(t,t-h)=\gamma(i,i-h),\quad i=1,...,d,\ h\in\mathbb{N}\ et\ t\in\mathbb{Z}.$$

Il est facile de vérifier que

$$\gamma_h^{(i+d\tau)}=\gamma_h^{(i)}\quad et\quad \gamma_{-h}^{(i)}=\gamma_h^{(i+h)}.$$

Définition (3.4.2) Nous appelons matrice d'autocovariances d'ordre k et relative à la i^{eme} période d'un processus périodiquement corrélé d-périodique toute matrice $(k\times k)$ définie par :

$$\Gamma^{(i)}(k,h)=\begin{pmatrix} \gamma_{-h}^{(i-1)} & \gamma_{-h+1}^{(i-2)} & \cdots & \gamma_{-h+k-1}^{(i-k)} \\ \gamma_{-h-1}^{(i-1)} & \gamma_{-h}^{(i-2)} & \ddots & \gamma_{-h+k-2}^{(i-k)} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{-h-k+1}^{(i-1)} & \gamma_{-h-k+2}^{(i-2)} & \cdots & \gamma_{-h}^{(i-k)} \end{pmatrix},\quad \text{où } i=1,...,d,\quad h,k\in N.$$

Définition (3.4.3) Nous appelons matrice d'autocorrélation d'ordre k et relative à la i^{eme} période d'un processus périodiquement corrélé d-périodique toute matrice $(k\times k)$ définie par :

$$\Sigma^{(i)}(k,h)=\begin{pmatrix} \rho_{-h}^{(i-1)} & \rho_{-h+1}^{(i-2)} & \cdots & \rho_{-h+k-1}^{(i-k)} \\ \rho_{-h-1}^{(i-1)} & \rho_{-h}^{(i-2)} & \ddots & \rho_{-h+k-2}^{(i-k)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{-h-k+1}^{(i-1)} & \rho_{-h-k+2}^{(i-2)} & \cdots & \rho_{-h}^{(i-k)} \end{pmatrix}, \quad i=1,...,d, \quad h,k\in\mathbb{N}.$$

où $\rho_h^{(i+d\tau)}=\rho_h^{(i)}=\frac{\gamma_h^{(i)}}{\gamma_0^{(i)}}$ est la fonction d'autocorrélation pour l'horizon h et relative à la $i \overset{eme}{=}$ période du processus y_t

3.4.2 Relations entre les autocovariances d'un processus multivarié stationnaire et les autocovariances d'un processus périodiquement corrélé

Rapplons que d'après le théorème de Gladychev (1961) le processus d-varié $\{X_\tau, \tau\in\mathbb{Z}\}$ défini par :

$$\begin{aligned} X_\tau &= (X_{\tau,1}, X_{\tau,2}, ..., X_{\tau,d})' \\ &= (y_{1+d\tau}, y_{2+d\tau}, ..., y_{k+d\tau})' \end{aligned}$$

est stationnaire si et seulement si le processus $\{y_t, t\in\mathbb{Z}\}$ est périodiquement corrélé de période d.

Soit $\Gamma(\tau_1,\tau_2)$ la matrice de variance-covariance de X_{τ_1} et X_{τ_2} définie par :

$$\Gamma(\tau_1,\tau_2)=\left(\gamma_{i,j}(\tau_1,\tau_2)\right)_{i,j=1,...,d} \qquad \tau_1,\tau_2\in Z$$

où

$$\begin{aligned} \gamma_{i,j}(\tau_1,\tau_2) &= cov\left(X_{\tau_1,i}, X_{\tau_2,j}\right) \quad i,j=1,...,d \\ &= cov\left(y_{i+d\tau_1}, y_{j+d\tau_2}\right) \quad i,j=1,...,d. \end{aligned}$$

Nous avons alors,

$$\gamma_{i,j}(\tau_1,\tau_2)=\gamma_{i-j+d(\tau_1-\tau_2)}^{(i)} \quad i,j=1,...,d.$$

Le processus d-varié X_τ étant stationnaire, sa matrice d'autocovariance à l'horizon h est donnée par :

$$\Gamma(h)=\Gamma(\tau,\tau-h)=\left(\gamma_{i-j+dh}^{(i)}\right)_{i,j=1,...,d}.$$

3.5 Estimation des paramètres d'un modèle autoregressif moyenne mobile périodique

Dans cette section, nous présenterons les principaux résultats concernant l'estimation des paramètres d'un modèle autoregressif moyenne mobile d-périodique. Nous passerons en revue quelques algorithmes d'estimation.

Les méthodes d'estimation sont généralement classées en deux grandes classes : les méthodes *hors-ligne* et les méthodes *en-ligne*. Dans la première classe, il s'agit d'une série de taille fixe, alors que dans la deuxième les données sont progressivement disponibles.

Comme il a été déjà signalé dans le cas stationnaire, les méthodes d'estimation *hors-ligne* peuvent être classées en deux groupes : les méthodes utilisant directement les données et les méthodes utilisant des transformations des données. Nous commencerons, par la première classe en exposant les principaux résultats concernant l'estimation, par la méthode des moments, des paramètres d'un processus autoregressif pur d-périodique d'ordre constant p. Nous récapitulerons ensuite, dans la même sous section la méthode des moments pour un modèle ARMA mixte d-périodique. Nous terminerons cette sous section par un exposé de l'algorithme de Boshnakov (1996). La deuxième sous section sera consacrée à l'estimation par la méthode du maximum de vraisemblance. Nous exposerons à ce sujet l'algorithme de Vecchia (1985a). Nous terminerons cette classe par la présentation d'une deuxième méthode basée sur l'utilisation directe des données qui est la méthode des moindres carrés.

Parmis les méthodes d'estimation *en-ligne*, nous présenterons deux algorithmes. Le premier est l'algorithme des moindres carrés récursifs périodique (PRLS), relatif aux modèles autoregressifs purs. Le second qui n'est autre qu'une généralisation du premier aux cas des modèles autoregressifs moyennes mobiles, c'est l'algorithme *Periodic Recursive Maximum Likelihood* (PRML).

3.5.1 Estimation des paramètres par la méthode des moments

3.5.1.1 Estimation des paramètres d'un modèle AR d-périodiques

A. Equations de Yule-walker périodiques Soit $\{y_t, t \in \mathbb{Z}\}$ un processus autoregressif d-périodique d'ordre p_t et de représentation (3.3.2), qui peut s'écrire aussi sous la forme suivante

$$y_{k+d\tau} - \sum_{j=1}^{p_t} \phi_{k+d\tau,j} y_{k+d\tau-j} = \varepsilon_{k+d\tau}, \quad k = 1, ..., d \tag{3.5.1}$$

Multiplions les membres de l'équation (3.5.1) par $y_{t-\nu}$ (ou également par $y_{k+d\tau-\nu}$, $k = 1, ..., d$). Pour ν positif et ε_t supposée non corrélée avec $y_{t-\nu}$, en prenant l'espérance des deux nembre de l'équation, nous obtenons le résultat suivant :

$$\gamma(k,k-\nu)-\sum_{j=1}^{p_k}\phi_{i,j}\gamma(k-j,k-\nu)=\delta_{\nu 0}\sigma_k^2,\quad \nu\geq 0. \tag{3.5.2}$$

où $\gamma(.,.)$ est la fonction d'autocovariance du processus et $\delta_{\nu 0}$ désigne le symbole de Kronecker.

L'estimation des paramètres d'un modèle ARMA d-périodique par la méthode des moments, est semblable à l'estimation des modèles ARMA stationnaire, basée sur les équations de Yule-Walker périodiques. Le principe de ces méthodes consiste à remplacer, dans les EYW périodiques, les autocovariances théoriques par leurs estimations empiriques.

Supposons que nous disposons d'une série de taille dN considérée comme réalisation d'un modèle PARMA gaussien.

Soit

$$\mu_i=\frac{1}{N}\sum_{\tau=0}^{N-1}y_{i+d\tau}$$

un estmateur sans biais de la moyenne relative à la i^{eme} période.

Dans ce cas l'estimateur, par la méthode des moments, de la fonction d'autocovariance $\gamma_h^{(i)}$, pour la i^{eme} période et l'horizon h, est la fonction d'autocovariance empirique notée par $\widetilde{\gamma}_h^{(i)}$ et définie par

$$\widetilde{\gamma}_h^{(i)}=\frac{1}{N}\sum_{\tau=0}^{N-1}\left(y_{i+d\tau}-\mu_i\right)\left(y_{i+d\tau-h}-\mu_{i-h}\right)$$

où $y_{i+d\tau-l}=0$ pour $i+d\tau-l<1$ ou pour $i+d\tau-l>dN$.

B. Estimateurs de Yule-Walker dans le cas périodique Considérons le processus $\{y_{i+d\tau}, i=1,...,d,\tau\in\mathbb{Z}\}$ supposé être représenté par le modèle suivant :

$$y_{i+d\tau}-\sum_{j=1}^{p}\phi_{i,j}y_{i+d\tau-j}=\varepsilon_{i+d\tau},\quad i=1,...,d;\tau\in\mathbb{Z} \tag{3.5.3}$$

Notons par $\Phi_i, i=1,...,d$ le p_i-vecteur des paramètres, relatif à la i^{eme} période, définit par $(\phi_{i,1},\phi_{i,2},...,\phi_{i,p})'$. En multipliant (3.5.3) par $y_{i+d\tau-h}$, $h=1,...,p$, et en prenant l'espérance mathématique des deux membres, alors nous obtenons pour chaque i, l'ensemble des $p+1$ équations suivantes :

$$\gamma_h^{(i)}-\sum_{j=1}^{p}\phi_{i,j}\gamma_{h-j}^{(i-j)}=\delta_{h0}\sigma_i^2,\quad i=1,...,d$$

ou encore, sous la forme matricielle

$$\begin{cases} \Gamma_i \Phi_i = \gamma_i \\ \quad \sigma_i^2 = \gamma_0^{(i)} - \Phi_i' \gamma_i \end{cases} \tag{3.5.4}$$

avec $(\Gamma_i)_{k,j=1,\dots,p} = \gamma_{j-k}^{(i-k)}$ et $\gamma_i = \left(\gamma_1^{(i)}, \gamma_2^{(i)}, \dots, \gamma_p^{(i)}\right)'$.

Les estimateurs $\widetilde{\Phi}_i$ et $\widetilde{\sigma}_i^2$ par la méthode des moments de Φ_i et σ_i^2 sont obtenus à partir de (3.5.4) en remplaçant les autocovariances théoriques par leurs estimateurs empiriques et en résolvant le système d'équations suivant :

$$\begin{cases} \widetilde{\Gamma}_i \widetilde{\Phi}_i = \widetilde{\gamma}_i \\ \quad \widetilde{\sigma}_i^2 = \widetilde{\gamma}_0^{(i)} - \widetilde{\Phi}_i' \widetilde{\gamma}_i \end{cases} \tag{3.5.5}$$

Nous avons maintenant besoin de déterminer les propriétés de ces estimateurs

Théorème (3.5.1) (*Pagano (1978)*)
Si $y_1, y_2, \dots, y_{dN}$ une dN-réalisation d'un modèle autoregressif d-périodique d'ordre constant p. Alors, Les estimateurs $\widetilde{\Phi}$ et $\widetilde{\sigma}^2$ définis par (3.5.5) sont convergent presque sûrement. De plus, $\sqrt{N}(\widetilde{\Phi}_i - \Phi_i), \quad i = 1, \dots, d$ est asymptotiquement normalement distribuée avec une moyenne nulle et de matrice de covariance F^{-1}, où F est le bloc approprié de la matrice d'information de Fisher.

Démonstration Voir Pagano (1978).

3.5.1.2 Estimation des paramètres d'un modèle PARMA

Lorsque nous rajoutons une composante moyenne mobile c'est-à-dire en considérant le modèle PARMA, l'estimation par la méthode des moments devient plus délicate. En effet, pour les paramètres moyenne mobile périodiques, nous aurons besoin de résoudre des systèmes non linéaires très compliqués. Cependant, les estimateurs obtenus par la méthode des moments serons considérés seulement comme valeurs initiales pour les méthodes itératives telles que la méthode du maximum de vraisemblence ou des moindres carrés. En effet, cette classe de méthodes, comme nous verrons par la suite, requiert une routine d'optimisation numérique et par conséquent un ensemble d'itérations nécessaire jusqu'à l'obtention de la convergence voulue. Il y'a donc un grand intérêt à utiliser des valeurs initiales convenables pour ne pas altérer la sensibilité de la méthode.

D'après Vecchia (1985a), la valeur $\widetilde{\Phi}_i$ obtenue par la résolution du système

$$\widetilde{\Gamma}_{i,q} \widetilde{\Phi}_i = \widetilde{\gamma}_{i,q}$$

où $\left(\widetilde{\Gamma}_{i,q}\right)_{k,j=1,\dots,p} = \widetilde{\gamma}_{j-k-q}^{(i-k-q)}$ et $\widetilde{\gamma}_{i,q} = \left(\widetilde{\gamma}_{q+1}^{(i)}, \widetilde{\gamma}_{q+2}^{(i)}, \dots, \widetilde{\gamma}_{q+p}^{(i)}\right)'$ semble être une valeur initiale convenable.

3.5.1.3 Estmation des moments par des méthodes récursives

A. Prédiction des processus périodiquement corrélés

Soit $L^2(\Omega, A, P)$ l'espace de Hilbert des variables aléatoires de carrés intégrables, muni du produit scalaire $\langle x, y\rangle = E(xy)$. Soit $\{y_t, t \in \mathbb{Z}\}$ un processus du second ordre. Notons par $M_{t,r}, t \in \mathbb{Z},\ r \in \mathbb{N}^*$, le sous-espace de Hilbert engendré par $y_{t-r}, y_{t-r+1}, ..., y_{t-1}$. Alors, le prédicteur linéaire (ou le meilleur prédicteur linéaire au sens des moindres carrés) $\widehat{y}_t$ de y_t sachant $y_{t-r}, y_{t-r+1}, ..., y_{t-1}$ est défini, d'après le théorème de projection, comme la projection orthogonale (au sens de L^2) de y_t sur le sous-espace $M_{t,r}$, c'est-à-dire

$$\widehat{y}_t = P_{M_{t,r}} y_t.$$

A1. Equation de prédiction prospective

Puisque $\widehat{y}_t \in M_{t,r}$, il existe donc des coefficients $\phi_{t,j}^{(r)}, j = 1, ..., r$ tels que

$$\widehat{y}_t = \sum_{j=1}^{r} \phi_{t,j}^{(r)} y_{t-j}$$

où les coefficients $\phi_{t,j}^{(r)}, j = 1, ..., r$ satisfont les équations de prédiction, c'est-à-dire

$$\langle y_t - \widehat{y}_t, y_{t-j}\rangle = 0 \text{ pour tout } j = 1, ..., r$$

ou encore

$$\sum_{k=1}^{r} \phi_{t,k}^{(r)} \langle y_{t-k}, y_{t-j}\rangle = \langle y_t, y_{t-j}\rangle\,, \quad j = 1, ..., r$$

d'où

$$\sum_{k=1}^{r} \phi_{t,k}^{(r)} \gamma(t-k, t-j) = \gamma(t, t-j), \quad j = 1, ..., r$$

où $\gamma(.,.)$ est la fonction d'autocovariance du processus y_t. La dernière équation peut se mettre sous la forme matricielle suivante :

$$\Gamma_{t,r} \Phi_{t,r} = \gamma_{t,r},$$

où $\Phi_{t,r} = \left(\phi_{t,1}^{(r)}, \phi_{t,2}^{(r)}, ..., \phi_{t,r}^{(r)}\right)'$, $\gamma_{t,r} = (\gamma(t, t-1), \gamma(t, t-2), ..., \gamma(t, t-r))'$ et $\Gamma_{t,r} = (\Gamma_{t,r}(j,k))_{j,k=1,...,r}$, $\Gamma_{t,r}(j,k) = \gamma(t-k, t-j)$

De plus si la matrice $\Gamma_{t,r}$ est régulière, alors les coefficients du meilleur prédicteur linéaire de y_t peuvent être définis par l'équation suivante :

$$\Phi_{t,r} = \Gamma_{t,r}^{-1} \gamma_{t,r}.$$

A2. Equation de prédiction rétrospective

Notons par $\widetilde{y}_t^{(r)}$ la meilleure projection de y_t sur le sous espace $M_{t+r+1,r}$ qui est engendré par $y_{t+1}, y_{t+2}, ..., y_{t+r}$. Comme $\widetilde{y}_t \in M_{t+r+1,r}$, il existe donc des coefficients $\beta_{t,j}^{(r)}, j = 1, ..., r$ tel que

$$\widetilde{y}_t = \sum_{j=1}^{r} \beta_{t,j}^{(r)} y_{t+j}$$

où les coefficients $\beta_{t,j}^{(r)}, j = 1, ..., r$ satisfont les équations de prédiction, c'est-à-dire

$$\langle y_t - \widetilde{y}_t, y_{t+j} \rangle = 0 \text{ pour tout } j = 1, ..., r$$

ou encore

$$\sum_{k=1}^{r} \beta_{t,k}^{(r)} \langle y_{t-k}, y_{t+j} \rangle = \langle y_t, y_{t+j} \rangle, \quad j = 1, ..., r$$

d'où

$$\sum_{k=1}^{r} \beta_{t,k}^{(r)} \gamma(t-k, t+j) = \gamma(t, t+j), \quad j = 1, ..., r$$

où $\gamma(.,.)$ est la fonction d'autocovariance du processus y_t.

Remarque

- Si le processus $\{y_t, t \in \mathbb{Z}\}$ est périodiquement corrélé de période d, alors la fonction d'autocovariance est d-périodique. Donc nous aurons pour tout $t, \tau \in \mathbb{Z}$ et $i = 1, ..., d$ tel que $t = i + d\tau$:

$$\Gamma_{i+d\tau,r} = \Gamma_{i,r} \text{ et } \gamma_{i+d\tau,r}. = \gamma_{i,r}, \quad i = 1, ..., d; \tau \in \mathbb{Z}.$$

- Si en plus la matrice $\Gamma_{i,r}$ est régulière, alors les coefficients de prédictions seront d-périodiques et ils sont donnés par l'équation suivante :

$$\sum_{k=1}^{r} \phi_{t,k}^{(r)} \gamma_{j-k}^{(i-k)} = \gamma_j^{(i)}, \quad j = 1, ..., r; i = 1, ..., d.$$

B. Estimation des moments par des méthodes récursives

Dans ce paragraphe, nous présenterons une méthode récursive, relativement à l'ordre du modèle, pour l'estimation des paramètres d'un modèle autoregressif moyenne mobile d-périodique. L'idée principale de cette méthode est de déterminer progressivement les coefficients du meilleur prédicteur de y_t sachant $y_{t-1}, \cdots, y_{t-p}$, $\varepsilon_{t-1}, \cdots, \varepsilon_{t-q}$.

Soit alors,

$$\widehat{y}_t^{(p,q)} = E\left(y_t/y_{t-1},\cdots,y_{t-p},\varepsilon_{t-1},\cdots,\varepsilon_{t-q}\right) = \sum_{j=1}^{p}\phi_{t,j}^{(p,q)}y_{t-j} + \sum_{j=1}^{q}\theta_{t,j}^{(p,q)}\varepsilon_{t-j}$$

$$\varepsilon_t^{(p,q)} = y_t - \widehat{y}_t^{(p,q)}, \quad \sigma_i^{2(p,q)} = \left\|y_t - \widehat{y}_t^{(p,q)}\right\|^2 = E\left(\varepsilon_t^{(p,q)}\right)^2.$$

$$\widetilde{y}_t^{(p,q)} = E\left(y_t/y_{t+1},\cdots,y_{t+p},\varepsilon_{t+p-q+1},\cdots,\varepsilon_{t+p}\right)$$
$$= \sum_{j=1}^{p}\beta_{t,j}^{(p,q)}y_{t+j} + \sum_{j=1}^{q}\delta_{t,j}^{(p,q)}\varepsilon_{t+p-q+j}$$

$$a_t^{(p,q)} = y_t - \widetilde{y}_t^{(p,q)}, \quad \nu_i^{2(p,q)} = \left\|y_t - \widetilde{y}_t^{(p,q)}\right\|^2 = E\left(a_t^{(p,q)}\right)^2.$$

Boshnakov (1996) a présenté un algorithme pour calculer récursivement les paramètres d'un modèle autoregressif moyenne mobile périodique. C'est une généralisation de l'algorithme de Levinson-Durbin (1960) pour un modèle autoregressif stationnaire, l'algorithme de Sakai (1982) pour un modèle autoregressif périodique et l'algorithme de Franke (1985) pour un modèle ARMA stationnaire.

Considérons le modèle autoregressif moyenne mobile d-périodique suivant :

$$y_{i+d\tau} = \sum_{j=1}^{p}\phi_{i,j}y_{i+d\tau-j} + \varepsilon_{i+d\tau} - \sum_{j=1}^{q}\theta_{i,j}\varepsilon_{i+d\tau-j}, \quad i=1,...,d; \tau\in\mathbb{Z}$$

Nous supposons que le modèle est causal. C'est-à-dire

$$y_{i+d\tau} = \varepsilon_{i+d\tau} + \sum_{j=1}^{\infty}h_{i,j}y_{i+d\tau-j}$$

où la series infinie est convergente, en moyenne quadratique. Alors le schéma récursif pour calculer $\phi_{i,j}$ et $\theta_{i,k}$ à partir de $\gamma_h^{(i)}$, $i=1,...,d,\ j=1,...,P,\ k=1,...,Q$ est donné par l'algorithme suivant :

Algorithme de Boshnakov (1996)

$1 \stackrel{ere}{=}$ *étape* : Pour $p=q=0$

$$\sigma_i^{2(0,0)} = \nu_i^{2(0,0)} = \gamma_0^{(i)}, i=1,...,d$$

$2 \stackrel{eme}{=}$ *étape* : Pour $p=0,\ q\geq 1$

$$\begin{aligned}
\theta_{i,k}^{(0,q)} &= h_{i,k}, \quad k=1,...,q\\
\sigma_i^{2(0,q)} &= \sigma_i^{2(0,q-1)} - \left(h_{i,q}\right)^2\sigma_{i-q}^2\\
\delta_{i,k}^{(0,q)} &= h_{i,q-k}, \quad k=1,...,q\\
\nu_i^{2(0,q)} &= \nu_i^{2(0,q-1)} - \left(h_{i,q-1}\right)^2\sigma_{i-q+1}^2
\end{aligned}$$

$3\ \underset{=}{^{eme}}$ *étape* : Pour $p \geq 1,\ q = 0$

$$\begin{aligned}
\phi_{i,p}^{(p,0)} &= \left(\gamma_p^{(i)} - \sum_{j=1}^{p-1} \beta_{i-p,j}^{(p-1,0)} \gamma_{p-j}^{(i)}\right) \Big/ \nu_{i-p}^{2(p-1,0)} \\
\phi_{i,j}^{(p,0)} &= \phi_{i,j}^{(p-1,0)} - \phi_{i,p}^{(p,0)} \beta_{i-p,p-j}^{(p-1,0)}, \quad j = 1, ..., p-1 \\
\sigma_i^{2(p,0)} &= \sigma_i^{2(p-1,0)} - \left(\phi_{i,p}^{(p,0)}\right)^2 \nu_{i-p}^{2(p-1,0)} \\
\beta_{i,p}^{(p,0)} &= \left(\gamma_{-p}^{(i)} - \sum_{j=1}^{p-1} \phi_{i+p,j}^{(p-1,0)} \gamma_{j-p}^{(i)}\right) \Big/ \sigma_{i+p}^{2(p-1,0)} \\
\beta_{i,j}^{(p,0)} &= \beta_{i,j}^{(p-1,0)} - \beta_{i,p}^{(p,0)} \phi_{i+p,p-j}^{(p-1,0)}, \quad j = 1, ..., p-1 \\
\nu_i^{2(p,0)} &= \left(1 - \phi_{i+p,p}^{(p,0)} \beta_{i,p}^{(p,0)}\right) \sigma_i^{2(p-1,0)}
\end{aligned}$$

$4\ \underset{=}{^{eme}}$ *étape* : Pour $p \geq 1,\ q \geq 1$

$$\begin{aligned}
\phi_{i,p}^{(p,q)} &= \frac{\left(\gamma_p^{(i)} - \sum_{j=1}^{p-1} \beta_{i-p,j}^{(p-1,q)} \gamma_{p-j}^{(i)} - \sum_{j=1}^{q} \delta_{i-p,j}^{(p-1,q)} \sigma_{i-1-q+j}^2 h_{i,q+1-j}\right)}{\nu_{i-p}^{2(p-1,q)}} \\
\phi_{i,j}^{(p,q)} &= \phi_{i,j}^{(p-1,q)} - \phi_{i,p}^{(p,q)} \beta_{i-p,p-j}^{(p-1,q)}, \quad j = 1, ..., p-1 \\
\theta_{i,k}^{(p,q)} &= \theta_{i,k}^{(p-1,q)} - \phi_{i,p}^{(p,q)} \delta_{i-p,q+1-k}^{(p-1,q)}, \quad k = 1, ..., q \\
\sigma_i^{2(p,q)} &= \sigma_i^{2(p-1,q)} - \left(\phi_{i,p}^{(p,q)}\right)^2 \nu_{i-p}^{2(p-1,q)} \\
\beta_{i,p}^{(p,q)} &= \frac{\left(\gamma_p^{(i+p)} - \sum_{j=1}^{p-1} \phi_{i+p,j}^{(p-1,q-1)} \gamma_{p-j}^{(i+p-j)} - \sum_{j=1}^{q-1} \delta_{i+p,j}^{(p-1,q-1)} \sigma_{i+p-j}^2 h_{i,j-p}\right)}{\left(\sigma_{i+p}^{2(p-1,q-1)} - \sigma_{i+p}^2\right)} \\
\beta_{i,j}^{(p,q)} &= \beta_{i,j}^{(p-1,q-1)} - \beta_{i,p}^{(p,q)} \phi_{i+p,p-j}^{(p-1,q-1)}, \quad j = 1, ..., p-1 \\
\delta_{i,k}^{(p,q)} &= \delta_{i,k}^{(p-1,q-1)} - \beta_{i,p}^{(p,q)} \theta_{i+p,q-k}^{(p-1,q-1)}, \quad k = 1, ..., q-1 \\
\delta_{i,q}^{(p,q)} &= -\beta_{i,p}^{(p,q)} \\
\nu_i^{2(p,q)} &= \nu_i^{2(p-1,q-1)} - \left(\beta_{i,p}^{(p,q)}\right)^2 \left(\sigma_{i+p}^{2(p-1,q-1)} - \sigma_{i+p}^2\right)
\end{aligned}$$

$5\ \underset{=}{^{eme}}$ *étape* : Pour $p \geq 1,\ q \geq 1$ et $\sigma_{i+p}^{2(p-1,q-1)} - \sigma_{i+p}^2 \neq 0$

$$\beta_{i,p}^{(p,q)} = \left(\phi_{i+p,p}^{(p,q-1)} \nu_i^{2(p-1,q-1)}\right) \Big/ \left(\sigma_{i+p}^{2(p-1,q-1)} - \sigma_{i+p}^2\right)$$

Les équations ci-dessus seront appliqées pour $p = 1, ..., P,\ q = 1, ..., Q$ et $i = 1, ..., d$. Les paramètres seront obtenus à partir de

$$\phi_{i,j} = \phi_{i,j}^{(P,Q)},\ \theta_{i,k} = \theta_{i,k}^{(P,Q)} \text{et } \sigma_i^2 = \sigma_i^{2(P,Q)} \text{ pour tout } i = 1, ..., d. \qquad \square$$

3.5.2 Estimation par la méthode du maximum de vraisemblance

Supposons maintenant que nous voulons estimer les paramètres d'un modèle PARMA, en se basant sur une réalisation donnée. Alors, ces estimateurs seront choisis de façon à minimiser une fonction critère fixée.

Nous considérons le modèle (3.3.3) avec une hypothèse très précise sur la distribution du bruit blanc. Soit $\{\varepsilon_t, t \in \mathbb{Z}\}$ une suite de variables aléatoires i.i.d, de loi $\mathcal{N}(0, \sigma_t^2)$. Pour une dN-réalisation $y_1, ..., y_{dN}$, le modèle (3.3.3) peut se mettre sous la forme suivante :

$$A\ Y_N = B\ \underline{\varepsilon}_N + C\ I_* \tag{3.5.6}$$

où $Y_N = (\ y_1, ..., y_{dN})'$, $\underline{\varepsilon}_N = (\varepsilon_1, ..., \varepsilon_{dN})'$, $I_* = (\ y_0, ..., y_{1-p}, \varepsilon_0, ..., \varepsilon_{1-q})'$. et A et B sont des matrices triangulaires inférieurs de taille $(dN \times dN)$ à diagonale unité. La matrice C est de taille $(dN \times (p+q))$. L'expression exacte de ces matrices est donnée par :

$$\text{A=}\begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ -\phi_{2,1} & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ -\phi_{p,p-1} & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & -\phi_{p,p-1} & \cdots & -\phi_{p,1} & 1 \end{pmatrix}, \quad \text{B=}\begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ -\theta_{2,1} & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ -\theta_{q,q-1} & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & -\theta_{q,q-1} & \cdots & -\theta_{q,1} & 1 \end{pmatrix}$$

$$\text{C=}\begin{pmatrix} -\phi_{1,p} & \cdots & \cdots & -\phi_{1,1} & \theta_{1,q} & \cdots & \cdots & \theta_{1,1} \\ 1 & -\phi_{2,p} & \ddots & -\phi_{2,2} & 0 & \theta_{2,q} & \ddots & \theta_{2,2} \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\phi_{p,p} & \vdots & \ddots & \ddots & \theta_{p,p} \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

Pour plus de clarté considérons la série $y_1, ..., y_{dN}$ supposée issue d'un modèle $\mathrm{MA}_2(1)$ suivant :

$$\begin{cases} y_{1+2\tau} = \varepsilon_{1+2\tau} - \theta_1 \varepsilon_{2\tau}, \\ y_{2+2\tau} = \varepsilon_{2+2\tau} - \theta_2 \varepsilon_{1+2\tau}. \end{cases} \tag{3.5.7}$$

où $\theta = (\theta_1, \theta_2)'$ est le vecteur des paramètres à estimer.

3.5.2.1 la méthode du maximum de vraisenblance conditionnelle pour les modèles PARMA

Soit $\mathcal{L}_*(\theta_1, \theta_2)$ la fonction de vraisemblance conditionnelle à $\varepsilon_0 = 0$. Nous savons que la densité conjointe du $\varepsilon = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_{2N})$ est égale à

$$f(\varepsilon) = \left(\frac{1}{\sqrt{2\pi}}\right)^{2N} \left(\prod_{t=1}^{2N} \sigma_t^2\right)^{-1/2} \exp\left(-\frac{1}{2}\sum_{t=1}^{2N} \frac{\varepsilon_t^2}{\sigma_t^2}\right),$$

ou encore

$$f(\varepsilon)=\left(\frac{1}{\sqrt{2\pi}}\right)^{2N}\sigma_1^{-N}\sigma_2^{-N}\exp\left\{-\frac{1}{2}\left(\sum_{\tau=0}^{N-1}\frac{\varepsilon_{1+2\tau}^2}{\sigma_1^2}+\sum_{\tau=0}^{N-1}\frac{\varepsilon_{2+2\tau}^2}{\sigma_2^2}\right)\right\}. \qquad (3.5.8)$$

Supposons que $y=(y_1,y_2,...,y_{2N})$ est une $2N$-réalisation du modèle. La matrice Jacobienne de la transformation de $\mathbb{R}^{2N}$ qui applique y sur ε à l'aide de (3.5.6) est triangulaire inférieure et les éléments diagonaux sont égaux à 1. La fonction de densité de y sachant $\varepsilon_0=0$ est donc donnée par (3.5.8). Nous aurons donc

$$\begin{aligned}\log\mathcal{L}_*(\theta_1,\theta_2,\sigma_1^2,\sigma_2^2)&=N\times\log(2\pi)\\&-\tfrac{N}{2}\log\sigma_1^2-\tfrac{N}{2}\log\sigma_2^2-\tfrac{1}{2}\left(\textstyle\sum_{\tau=0}^{N-1}\frac{\varepsilon_{1+2\tau}^2}{\sigma_1^2}+\sum_{\tau=0}^{N-1}\frac{\varepsilon_{2+2\tau}^2}{\sigma_2^2}\right)\end{aligned} \qquad (3.5.9)$$

cette dérnière nous donne

$$\left\{\begin{array}{l}\frac{\partial\log\mathcal{L}_*(\theta_1,\theta_2)}{\partial\sigma_1^2}=-\frac{N}{2\sigma_1^2}+\frac{1}{2}\left(\sum_{\tau=0}^{N-1}\frac{\varepsilon_{1+2\tau}^2}{\sigma_1^4}\right)\\\frac{\partial\log\mathcal{L}_*(\theta_1,\theta_2)}{\partial\sigma_2^2}=-\frac{N}{2\sigma_2^2}+\frac{1}{2}\left(\sum_{\tau=0}^{N-1}\frac{\varepsilon_{2+2\tau}^2}{\sigma_2^4}\right)\end{array}\right.$$

d'où

$$\left\{\begin{array}{l}\frac{\partial\log\mathcal{L}_*(\theta_1,\theta_2)}{\partial\sigma_1^2}=0\\\frac{\partial\log\mathcal{L}_*(\theta_1,\theta_2)}{\partial\sigma_2^2}=0\end{array}\right.\Longrightarrow\left\{\begin{array}{l}\sigma_1^2=\frac{1}{N}\left(\sum_{\tau=0}^{N-1}\varepsilon_{1+2\tau}^2\right)\\\sigma_2^2=\frac{1}{N}\left(\sum_{\tau=0}^{N-1}\varepsilon_{2+2\tau}^2\right)\end{array}\right. \qquad (3.5.10)$$

en remplaçant ce résultat dans (3.5.9), nous obtenons à une constante près

$$-\frac{N}{2}\log\left(\sum_{\tau=0}^{N-1}\varepsilon_{1+2\tau}^2\right)-\frac{N}{2}\log\left(\sum_{\tau=0}^{N-1}\varepsilon_{2+2\tau}^2\right) \qquad (3.5.11)$$

Ce qui montre que l'estimateur du maximum de vraisemblance $\left(\widehat{\theta}_1,\widehat{\theta}_2\right)$ de (θ_1,θ_2) sera choisi de façon à minimiser (3.5.11). Sachant $\left(\widehat{\theta}_1,\widehat{\theta}_2\right)$ nous pouvons déterminer l'estimateur $(\widehat{\sigma}_1^2,\widehat{\sigma}_2^2)$ de (σ_1^2,σ_2^2) par (3.5.10).

Il est facile de prouver pour un modèle $\mathrm{ARMA}_d(p,q)$ que

$$\widehat{\sigma}_i^2=\frac{1}{N}\left(\sum_{\tau=0}^{N-1}\varepsilon_{i+d\tau}^2\right),\quad i=1,...,d$$

Ainsi, en remplaçant dans la vraisemblance nous trouverons que l'estimateur du maximum de vraisemblance sera choisi de façon à minimiser

$$\log\left(\prod_{i=1}^{d}\sum_{\tau=0}^{N-1}\varepsilon_{i+d\tau}^2\right).$$

3.5.2.2 la méthode du maximum de vraisenblance non conditionnelle pour les modèles PARMA

Nous savons que

$$A\ Y_N = B\ \varepsilon_N + C\ I_*$$

qui implique

$$Y_N = A^{-1}\ B\ \varepsilon_N + A^{-1}\ C\ I_*$$

d'où

$$cov\,(Y_N) = A^{-1}\ B\ cov\,(\varepsilon_N)\ \ B'\ \ \left(A^{-1}\right)' + A^{-1}\ C\ cov\,(I_*)\ \ C'\ \ \left(A^{-1}\right)' \tag{3.5.12}$$

Or $cov(\varepsilon) = Diag(\sigma_1^2, ..., \sigma_{dN}^2)I_{dN} = D_{1,dN}^2 I_{dN}$, où I_{dN} est la matrice identité d'ordre dN et $D_{i,j}^k$ est la matrice diagonale d'éléments $\sigma_i^k, \sigma_{i+1}^k, ..., \sigma_j^k$. Pour plus de commodité, notons Γ la matrice de covariance du vecteur $Diag(\sigma_1^{-1}, ..., \sigma_{dN}^{-1})Y_N$, alors nous obtiendrons la matrice de covariance de Y_N qui est égale à $D_{1,dN}\Gamma D_{1,dN}^{'}$.

L'estimation de la fonction de vraisemblance semble être compliquée, nous pouvons en effet l'écrire sous la forme générale

$$\mathcal{L}(\phi, \theta, Y_N) = \left(\frac{1}{\sqrt{2\pi}}\right)^{2N} \left(\prod_{t=1}^{dN} \sigma_t^2 \det \Gamma\right)^{-1/2} \exp\left\{-\frac{1}{2}\left(Y_N' D_{1,dN}^{-1} \Gamma^{-1} D_{1,dN}^{-1} Y_N\right)\right\}. \tag{3.5.13}$$

D'où

$$\begin{aligned}\log \mathcal{L}(\phi, \theta, Y_N) = -\left(\tfrac{dN}{2}\right) \log(2\pi) - \tfrac{1}{2}\log(\det \Gamma) \\ -\tfrac{N}{2}\textstyle\sum_{t=1}^{d} \log \sigma_t^2 - \tfrac{1}{2}\left(Y_N' D_{1,dN}^{-1}\Gamma^{-1} D_{1,dN}^{-1} Y_N\right)\end{aligned} \tag{3.5.14}$$

En explicitant l'équation normale par rapport à σ_t^2 et en remplaçant dans l'équation (3.5.14) ceci fournit l'estimateur du maximum de vraisemblance.

3.5.2.3 Algorithme de Vecchia (1985)

A. Fonction de vraisemblance lorsque $p = 0$

Lorsque $p = 0$, nous avons :

$$y_{i+d\tau} = \varepsilon_{i+d\tau} - \sum_{j=1}^{q} \theta_{i,j}\varepsilon_{i+d\tau-j}, \quad i = 1, ..., d; \tau \in \mathbb{Z} \tag{3.5.15}$$

Soient

$$\begin{aligned} y_m &= \left(y_{m+1}, y_{m+2}, ..., y_{dN}\right)' \\ \varepsilon_m &= \left(\varepsilon_{m+1}, \varepsilon_{m+2}, ..., \varepsilon_{dN}\right)' \end{aligned} \quad avec\ 0 \leq m \leq dN - 1$$

De la représentation (3.5.15), nous aurons :

$$y_m = L_{\theta,m}\varepsilon_m - M_{\theta,m}\varepsilon_*,$$

où $\varepsilon_* = (\varepsilon_{m+1-q}, \varepsilon_{m+2-q}, ..., \varepsilon_m)^{'}$, $L_{\theta,m}$ est une matrice $(dN-m)\times(dN-m)$ triangulaire inférieure définie par :

$$(L_{\theta,m})_{i,j} = \begin{cases} 1 & si\ i=j \\ 0 & si\ i<j \\ -\theta_{m+i,i-j} & si\ i>j \end{cases}$$

et $M_{\theta,m}$ est une matrice $(dN-m)\times q$ définie par :

$$(M_{\theta,m})_{i,j} = \begin{cases} \theta_{m+i,q+i-j} & si\ i<q \\ 0 & si\ i>q \end{cases}$$

Vecchia (1985a) a donné l'expression de la fonction de vraisemblance du vecteur y_m par :

$$\mathcal{L}(\theta,\sigma/y_m) = \left(\tfrac{1}{\sqrt{2\pi}}\right)^{(dN-m)} |D_{m+1-q,dN}|^{-\frac{1}{2}} |A_{\theta,m}|^{-\frac{1}{2}} \times \exp\left\{-\tfrac{1}{2}\left[\hat{\varepsilon}_*^{'} D_{m+1-q,m}^{-1}\hat{\varepsilon}_* + \hat{\varepsilon}_m^{'} D_{m+1,dN}^{-1}\hat{\varepsilon}_m\right]\right\}, \quad (3.5.16)$$

où

$$\begin{aligned} D_{i,j} &= diag\left\{\sigma_i^2, \sigma_{i+1}^2, ..., \sigma_j^2\right\}, \\ A_{\theta,m} &= D_{m+1-q,m}^{-1} + F_{\theta,m}^{'} D_{m+1,dN}^{-1} F_{\theta,m}, \\ F_{\theta,m} &= -L_{\theta,m}^{-1} M_{\theta,m}, \\ \widehat{\varepsilon}_* &= E(\varepsilon_*/y_m) \text{ et } \widehat{\varepsilon}_m = E(\varepsilon_m/y_m) \end{aligned} \quad (3.5.17)$$

B. Approximation de la fonction de vraisemblance pour $p>0$

Pour $p>0$ nous ne pouvons pas calculer la fonction de vraisemblance exacte en utilisant la méthode précédente. Cependant une transformation adéquate permet d'obtenir une bonne aproximation de celle-ci.

Soit la transformation suivante :

$$w_{p+j} = y_{p+j} - \sum_{k=1}^{p} \theta_{p+j,k} y_{p+j-k}, \quad j = 1,2,...,dN-p \quad (3.5.18)$$

et

$$w_p = (w_{p+1}, w_{p+2}, ..., w_{dN})^{'}.$$

Cette transformation permet de passer à une série chronologique $\{w_t, t\in\mathbb{Z}\}$ d'un modèle $\mathrm{MA}_d(q)$ en considérant $y_1, y_2, ..., y_p$ fixes. La transformation de y_p à w_p a un Jacobien égal à 1.

Comme la série $\{w_t, t \in \mathbb{Z}\}$ est d'un modèle $\mathrm{MA}_d(q)$, nous avons donc

$$\begin{aligned}-2\log\left[\mathcal{L}(\phi,\theta,\sigma/w_p)\right] = dN\log(2\pi) + \log|D_{p+1-q,dN}| + \log|A_{\theta,p}| \\ +\widehat{\varepsilon}^{'}_{*}D^{-1}_{p+1-q,p}\widehat{\varepsilon}_{*} + \widehat{\varepsilon}^{'}_{p}D^{-1}_{p+1,dN}\widehat{\varepsilon}_{p},\end{aligned} \tag{3.5.19}$$

où $D_{i,j}, A_{\theta,p}$ et $F_{\theta,p}$ sont donnees par (3.5.17)

$$\widehat{\varepsilon}_{*} = (\widehat{\varepsilon}_{p+1-q}, \widehat{\varepsilon}_{p+2-q}, ..., \widehat{\varepsilon}_{p}) = A^{-1}_{\theta,p}F^{'}_{\theta,p}D^{-1}_{p+1}L^{-1}_{\theta,p}w_p.$$

et

$$\widehat{\varepsilon}_{p+j} = w_{p+j} + \sum_{k=1}^{p}\theta_{p+j,k}\widehat{\varepsilon}_{p+j-k}, \quad j = 1, 2, ..., dN - p.$$

C. Estimation du maximum de vraisemblance par l'algorithme de Vecchia

Algorithme de Vecchia (1985)

$1 \overset{ere}{=}$ *étape* : Nous commencons par des estimations initiales $\widehat{\phi}_0$ et $\widehat{\theta}_0$ obtenues par la méthode des moments.

$2 \overset{eme}{=}$ *étape* : Nous Posons $\widehat{\varepsilon}_j = 0$, pour $p+1-q \leq j \leq p$ et nous calculons $\widehat{\varepsilon}_j$, $p+1 \leq j \leq dN$ par la formule

$$\widehat{\varepsilon}_{p+j} = w_{p+j} + \sum_{k=1}^{p}\theta_{p+j,k}\widehat{\varepsilon}_{p+j-k}.$$

$3 \overset{eme}{=}$ *étape* : Nous calculons les estimateurs des variances

$$\widehat{\sigma}_i^2 = (N - K_2)^{-1}\sum_{n=1}^{N-K_1}\widehat{\varepsilon}^2_{dK_1+i+d(n-1)},$$

où $\widehat{\varepsilon}_j = 0$ pour $dK_1 < j \leq p - q$, K_1 est le plus grand entier satisfaisant $dK_1 < p + 1 - q$ et K_2 est le plus grand entier satisfaisant $dK_2 < p + 1 + d$.

$4 \overset{eme}{=}$ *étape* : Nous utilisons une routine d'optimisation non linéaire pour déterminer $\widehat{\phi}$ et $\widehat{\theta}$ minimisant $-2\log\left[\mathcal{L}(\phi,\theta,\widehat{\sigma}/w_p)\right]$, où $\widehat{\sigma} = \left(\widehat{\sigma}_1^2, \widehat{\sigma}_2^2, ..., \widehat{\sigma}_d^2\right)^{'}$.

$5 \overset{eme}{=}$ *étape* : Nous calculons $\widehat{\varepsilon}_j$, pour $p + 1 - q \leq j \leq dN$ comme suit :

$$\widehat{\varepsilon}_{*} = (\widehat{\varepsilon}_{p+1-q}, \widehat{\varepsilon}_{p+2-q}, ..., \widehat{\varepsilon}_{p}) = A^{-1}_{\theta,p}F^{'}_{\theta,p}D^{-1}_{p+1}L^{-1}_{\theta,p}w_p.$$

et

$$\widehat{\varepsilon}_{p+j} = w_{p+j} + \sum_{k=1}^{p}\theta_{p+j,k}\widehat{\varepsilon}_{p+j-k}, \quad j = 1, 2, ..., dN - p.$$

avec $\sigma = \widehat{\sigma}$, $\phi = \widehat{\phi}$ et $\theta = \widehat{\theta}$.

$6 \overset{eme}{=}$ *étape* : Nous répétons les étapes 3-5 jusqu'à l'obtention de précision voulue.

3.5.3 Estimation des paramètres par la méthode des moindres carrés

3.5.3.1 Méthode des moindres carrés conditionnelle dans le cas périodique

Cette méthode consiste à choisir un estimateur des paramètres du modèle minimisant la somme des carrés. Contrairement au cas d'un ARMA stationnaire les membres de la somme seront pondérés par la variance de l'erreur. Nous supposons que $I_* = \varepsilon_0$ est fixé, soit par exemple égal à 0. De (5.3.7), nous avons alors

$$\begin{aligned}
\varepsilon_1 &= y_1 \\
\varepsilon_2 &= y_2 + \theta_2 y_1 \\
\varepsilon_2 &= y_3 + \theta_1 y_2 + \theta_1 \theta_2 y_1
\end{aligned}$$

et ainsi de suite

$$\begin{aligned}
\varepsilon_{1+2\tau} &= \sum_{j=0}^{2\tau} \theta_1^{[(j+1)/2]} \theta_2^{[j/2]} y_{1+2\tau-j} \\
\varepsilon_{2+2\tau} &= \sum_{j=0}^{2\tau} \theta_1^{[j/2]} \theta_2^{[(j+1)/2]} y_{2+2\tau-j}
\end{aligned}$$

où $[x]$ désigne la partie entière de x.

Nous avons pu expliciter la valeur de ε_t en fonction de $(\theta_1, \theta_2, \sigma_1^2, \sigma_2^2)$ et de y seulement. Par la suite, sachant que $I_* = 0$, la somme des carrés s'écrit comme suit :

$$\begin{aligned}
S_*(\theta_1, \theta_2, \sigma_1^2, \sigma_2^2) &= \sum_{t=1}^{dN} \frac{\varepsilon_t^2}{\sigma_t^2} = \sum_{i=1}^{2} \sum_{\tau=0}^{N-1} \frac{\varepsilon_{i+d\tau}^2(\theta_1, \theta_2, \sigma_1^2, \sigma_2^2)}{\sigma_{i+d\tau}^2} \\
&= \sum_{\tau=0}^{N-1} \frac{\varepsilon_{1+d\tau}^2(\theta_1, \theta_2, \sigma_1^2, \sigma_2^2)}{\sigma_1^2} + \sum_{\tau=0}^{N-1} \frac{\varepsilon_{2+d\tau}^2(\theta_1, \theta_2, \sigma_1^2, \sigma_2^2)}{\sigma_2^2}
\end{aligned}$$

qui est égale à

$$\sum_{\tau=0}^{N-1} \frac{\left(\sum_{j=0}^{2\tau} \theta_1^{[(j+1)/2]} \theta_2^{[j/2]} y_{1+2\tau-j}\right)^2}{\sigma_1^2} + \sum_{\tau=0}^{N-1} \frac{\left(\sum_{j=0}^{2\tau} \theta_1^{[j/2]} \theta_2^{[(j+1)/2]} y_{2+2\tau-j}\right)^2}{\sigma_2^2} \tag{3.5.20}$$

Nous pouvons remarquer que la fonction S_* n'est pas une fonction quadratique en $(\theta_1, \theta_2, \sigma_1^2, \sigma_2^2)$. L'équation normale permettant d'estimer $(\theta_1, \theta_2, \sigma_1^2, \sigma_2^2)$ n'est donc pas linéaire.

Notons qu'en pratique nous ne pouvons pas toujours évaluer la fonction S_* explicitement en fonction des paramètres et des observations. Mais la valeur de l'erreur ε_t sera déduite à partir d'un schéma récursif.

3.5.3.2 Méthode des moindres carrés non conditionnelle dans le cas périodique

Cette méthode se diffère de la première par le fait que nous devons déterminer une valeur de $I_* = \varepsilon_0$ meilleur que 0. Si le modèle (3.3.3) est correct, nous pouvons écrire également

$$\begin{cases} y_{1+2\tau} = \varepsilon'_{1+2\tau} - \theta_1 \varepsilon'_{2+2\tau} \\ y_{2+2\tau} = \varepsilon'_{2+2\tau} - \theta_2 \varepsilon'_{3+2\tau} \end{cases} \tag{3.5.21}$$

où les ε'_t sont les innovations en temps inversé. En supposant que $\varepsilon'_{k+2N} = 0$ pour $k \geq 1$, nous déterminons successivement $\varepsilon'_{2N} = y_{2N}$ puis $\varepsilon'_{2N-1} = y_{2N-1} + \theta_1 y_{2N}$.

De façon similaire au (3.5.20) nous obtenons

$$\begin{array}{rcl} \varepsilon'_{2N-2} & = & y_{2N-2} + \theta_2 y_{2N-1} + \theta_1 \theta_2 y_{2N} \\ & & \vdots \\ \varepsilon'_1 & = & \displaystyle\sum_{j=1}^{2N} \theta_1^{[j/2]} \theta_2^{[(j-1)/2]} y_j \end{array}$$

ce qui fournit finalement la prédiction rétrospéctive

$$\widehat{y}_0 = -\theta_2 \varepsilon'_1 = \varepsilon'_1 = -\sum_{j=1}^{2N} \theta_1^{[j/2]} \theta_2^{[(j+1)/2]} y_j$$

En prenant maintenant $\varepsilon_0 = \widehat{y}_0$ nous pouvons obtenir les ε_t par (2.5.20). Nous déduisons alors la somme des carrés non conditionnelle comme suit

$$\begin{aligned} S(\theta_1, \theta_2, \sigma_1^2, \sigma_2^2) = \sum_{t=1}^{dN} \frac{\varepsilon_t^2}{\sigma_t^2} &= \sum_{\tau=0}^{N-1} \sum_{j=0}^{2\tau} \frac{\theta_1^{[(j+1)/2]} \theta_2^{[j/2]} y_{1+2\tau-j}}{\sigma_1^2} \\ &+ \sum_{\tau=0}^{N-1} \sum_{j=0}^{2\tau} \frac{\theta_1^{[j/2]} \theta_2^{[(j+1)/2]} y_{2+2\tau-j}}{\sigma_2^2} - \sum_{t=1}^{dN} \theta_1^{[j/2]} \theta_2^{[(j+1)/2]} y_j. \end{aligned} \tag{3.5.22}$$

Nous présenterons (sans détail), dans ce qui suit, deux algorithmes d'estimation *en-ligne*. Le premier c'est l'algorithme des moindres carrés récursifs périodique (PRLS) relatif aux modèles autoregressifs purs. Le second qui n'est autre qu'une généralisation du premier, aux cas des modèles autoregressifs moyennes mobiles, est l'algorithme *Periodic Recursive Maximum Likelihood* (PRML).

3.5.4 Estimation *en-ligne* des modèles PARMA

3.5.4.1 Algorithme des moindres carrés récursif périodique (PRLS)

Cet algorithme est dérivé, par Bentarzi et Aknouche (2002), de la méthode usuelle des moindres carres, mais sous une forme recursive.

Considérons le modèle causal autoregressif d-périodique d'ordre constant p suivant :

$$y_{i+d\tau} = \sum_{j=1}^{p} \phi_{i,j} y_{i+d\tau-j} + \varepsilon_{i+d\tau} \quad i = 1, ..., d, \quad \tau \in \mathbb{Z} \tag{3.5.23}$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un bruit blanc périodique de variance σ_t^2. Soit $\beta = \left(\phi_1^{'}, \phi_2^{'}, ..., \phi_d^{'}\right)^{'}$ le vecteur de dimension $pd\times 1$ des paramètres autoregressifs, où $\phi_i = (\phi_{i,1}, \phi_{i,2}, ..., \phi_{i,p})^{'}$, $i = 1, 2, ..., d$.

Nous définissons le vecteur $\varphi_{i+d\tau}$ de dimemsion $pd \times 1$ comme suit :

$$\varphi_{i+d\tau}(j) \begin{cases} y_{i+d\tau-(j-(i-1)p)} & si\ j = (i-1)p+1, ..., ip \\ 0 & si\ non \end{cases}$$

Les equations récursives de l'algorithme PRLS pour estimer les paramètres d'un modèle AR d-périodique sont données par Bentarzi et Aknouche comme suit :

Algorithme de Bentarzi et Aknouche (2002)

$$\begin{aligned}
&\widehat{\varepsilon}_{i+d\tau} = y_{i+d\tau} - \widehat{\beta}^{'}_{i+d\tau-1}\varphi_{i+d\tau}, \\
&\widehat{\sigma}_j^{2(i,0)} = \begin{cases} \widehat{\varepsilon}_j^2 & si \quad 1 \leq j \leq i \\ 0 & si \quad\ j \geq i \end{cases} \\
&\widehat{\sigma}_j^{2(i,\tau)} = \begin{cases} \frac{\tau}{\tau+1}\widehat{\sigma}_j^{2(i,\tau-1)} + \frac{\tau}{\tau+1}\widehat{\varepsilon}^2_{j+d\tau} & si \quad 1 \leq j \leq i \\ \frac{\tau-1}{\tau}\widehat{\sigma}_j^{2(i,\tau-1)} + \frac{1}{\tau}\widehat{\varepsilon}^2_{j+d(\tau-1)} & si \quad\ j \geq i \end{cases} \\
&L_{i+d\tau} = \left(P_{i+d\tau-1}\varphi_{i+d\tau}\right) / \left(\widehat{\sigma}_i^2 + \varphi^{'}_{i+d\tau} L_{i+d\tau-1}\varphi_{i+d\tau}\right), \\
&\widehat{\beta}_{i+d\tau} = \widehat{\beta}_{i+d\tau-1} + L_{i+d\tau}\widehat{\varepsilon}_{i+d\tau}, \\
&P_{i+d\tau} = P_{i+d\tau-1} - L_{i+d\tau}\varphi^{'}_{i+d\tau}P_{i+d\tau-1}. \qquad \square
\end{aligned}$$

3.5.4.2 Algorithme Periodic Recursive Maximum Likelihood (PRML)

Cet algorithme n'est qu'une généralisation de l'algorithme PRLS, aux cas des modèles autoregressifs moyennes mobiles.

Considérons le modèle causal autoregressif moyenne mobile d-périodique d'ordre constant (p, q) suivant :

$$y_{i+d\tau} = \sum_{j=1}^{p} \phi_{i,j} y_{i+d\tau-j} + \varepsilon_{i+d\tau} - \sum_{j=1}^{q} \theta_{i,j}\varepsilon_{i+d\tau-j}, \quad i = 1, ..., d; \tau \in \mathbb{Z} \tag{3.5.24}$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un bruit blanc périodique de variance σ_t^2.

Soit $\beta = \left(\phi_1^{'}, \theta_1^{'}, \phi_2^{'}, \theta_2^{'}, ..., \phi_d^{'}, \theta_d^{'}\right)^{'}$ le vecteur de dimension $d(p+q) \times 1$ des paramètres, où $\phi_i = (\phi_{i,1}, \phi_{i,2}, ..., \phi_{i,p})^{'}$, $\theta_i = (\theta_{i,1}, \theta_{i,2}, ..., \theta_{i,q})^{'}$, $i = 1, 2, ..., d$. Nous définissons le vecteur $\varphi_{i+d\tau}$ de dimemsion $d(p+q) \times 1$ comme suit :

$$\varphi_{i+d\tau}(j) \begin{cases} y_{i+d\tau-(j-(i-1)(p+q))} & si \quad j = (i-1)(p+q)+1, ..., i(p+q)-q \\ -\varepsilon_{i+d\tau-(j-((i-1)(p+q)+p))} & si \quad j = (i-1)(p+q)+1+p, ..., i(p+q) \\ 0 & si\ non \end{cases}$$

Les equations récursives de l'algorithme PRML pour estimer les paramètres d'un modèle ARMA d-périodique sont données par Bentarzi et Aknouche (2003) comme suit :

Algorithme de Bentarzi et Aknouche (2003)

$$\begin{aligned}
&\widehat{\varepsilon}_{i+d\tau} = y_{i+d\tau} - \widehat{\beta}^{'}_{i+d\tau-1}\varphi_{i+d\tau}, \\
&\widehat{\sigma}_j^{2(i,0)} = \begin{cases} \widehat{\varepsilon}_j^2 & si \quad 1 \leq j \leq i \\ 0 & si \quad j \geq i \end{cases} \\
&\widehat{\sigma}_j^{2(i,\tau)} = \begin{cases} \frac{\tau}{\tau+1}\widehat{\sigma}_j^{2(i,\tau-1)} + \frac{\tau}{\tau+1}\widehat{\varepsilon}^2_{j+d\tau} & si \quad 1 \leq j \leq i \\ \frac{\tau-1}{\tau}\widehat{\sigma}_j^{2(i,\tau-1)} + \frac{1}{\tau}\widehat{\varepsilon}^2_{j+d(\tau-1)} & si \quad j \geq i \end{cases} \\
&L_{i+d\tau} = \left(P_{i+d\tau-1}\psi_{i+d\tau}\right) / \left(\widehat{\sigma}_i^2 + \psi^{'}_{i+d\tau}L_{i+d\tau-1}\psi_{i+d\tau}\right), \\
&\widehat{\beta}_{i+d\tau} = \widehat{\beta}_{i+d\tau-1} + L_{i+d\tau}\widehat{\varepsilon}_{i+d\tau}, \\
&\psi_{i+d\tau+1} = \textstyle\sum_{k=1}^{q} \widehat{\theta}_{i,k}^{(i+d\tau-1)}\psi_{i+d\tau-k+1} + \varphi_{i+d\tau}, \\
&P_{i+d\tau} = P_{i+d\tau-1} - L_{i+d\tau}\psi^{'}_{i+d\tau}P_{i+d\tau-1}
\end{aligned}$$

où

$$\widehat{\beta}_{i+d\tau} = \left(\widehat{\phi}_{1,1}^{(i+d\tau)}, ..., \widehat{\phi}_{1,p}^{(i+d\tau)}, \widehat{\theta}_{1,1}^{(i+d\tau)}, ..., \widehat{\theta}_{1,q}^{(i+d\tau)}, ..., \widehat{\phi}_{d,1}^{(i+d\tau)}, ..., \widehat{\phi}_{d,p}^{(i+d\tau)}, \widehat{\theta}_{d,1}^{(i+d\tau)}, ..., \widehat{\theta}_{d,q}^{(i+d\tau)}\right).$$

3.6 Estimation de la fonction d'autocorrélation d'un processus périodiquement corrélé

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus périodiquement corrélé de période d , de moyenne μ_t et de fonction d'autocovariance $\gamma(.,.)$ d-périodique. Pour tout entier $t \in \mathbb{Z}$, $i = 1, ..., d$ et $\tau \in \mathbb{Z}$ tel que $t = i + d\tau$, soit $\gamma_h^{(i)}$ sa fonction d'autocovariance pour l'horizon h et relative à la $i \overset{eme}{=}$ période.

L'estimation de μ_t, $\left\{\gamma_h^{(i)}, i = 1, ..., d\right\}$ et de fonction d'autocorrelations relative à la i^{eme} période $\left\{\rho_h^{(i)}, i = 1, ..., d\right\}$ à partir des observations $y_1, ..., y_{dN}$, joue donc un rôle très important dans les problèmes de l'inférence et en particulier dans le problème de l'identification du modèle (voir le dernier chapitre). Dans cette section, nous étudierons l'estimation de la moyenne, de la fonction d'autocovariance $\gamma_h^{(i)}$ et la fonction d'autocorrélation $\rho_h^{(i)}$ qui seront utilisés dans la méthode du coin pour identifier les modèles ARMA périodique. Puis nous examinerons les propriétés asymptotiques des estimateurs obtenus.

3.6.1 Estimation de la moyenne

Soit $y_1, ..., y_{dN}$ une dN-réalisation d'un processus $\{y_t, t \in \mathbb{Z}\}$périodiquement corrélé d-périodique. L'estimateur le plus simple et le plus naturel de la moyenne μ_t de y_t est la moyenne de l'échantillon

$$\overline{y} = \frac{1}{dN} \sum_{t=1}^{dN} y_t = \frac{1}{dN} \sum_{i=1}^{d} \sum_{\tau=0}^{N-1} y_{i+d\tau}, \tag{3.6.1}$$

qui est un estimateur sans biais. Nous remarquons que la moyenne de la i^{eme} période μ_i est estimé par

$$\overline{y}_i = \frac{1}{N} \sum_{\tau=0}^{N-1} y_{i+d\tau}. \tag{3.6.2}$$

La consistance de l'estimateur $\overline{y}_i$ sous certaines conditions modérées sur $\gamma_h^{(i)}$ peut être établie facilement en appliquant le théorème (1.6.1) sur le processus $\{y_{i+d\tau}, \tau \in \mathbb{Z}\}$.

Proposition (3.6.1)

Si $\{y_t, t \in \mathbb{Z}\}$ est un processus périodiquement corrélé d-périodique de moyenne μ_t et de fonction d'autocovariances relative à la i^{eme} période $\left\{\gamma_h^{(i)}, i = 1, ..., d\right\}$, alors,

$$E\left[(\overline{y}_i - \mu_i)^2\right] \underset{N \longrightarrow \infty}{\longrightarrow} 0 \qquad \text{si } \gamma_N^{(i)} \underset{N \longrightarrow \infty}{\longrightarrow} 0, \; i = 1, ..., d$$

et

$$N \times E\left[(\overline{y}_i - \mu_i)^2\right] \underset{N \longrightarrow \infty}{\longrightarrow} \sum_{h=-\infty}^{\infty} \gamma_h^{(i)} \qquad \text{si } \sum_{h=-\infty}^{\infty} \left|\gamma_h^{(i)}\right| < \infty, \; i = 1, ..., d.$$

Démonstration En appliquant le théorème de Gladyshev (1961) et le théorème (1.6.1) nous retrouvons le résultat.

Remarque Le théorème (3.6.1) montre que si $\gamma_N^{(i)} \longrightarrow 0$, $i = 1, ..., d$ lorsque $N \longrightarrow \infty$, alors $\overline{y}_i$ converge en moyenne quadratique (donc en probabilité) vers μ_i. De plus, sous la condition forte $\sum_{h=-\infty}^{\infty} \left|\gamma_h^{(i)}\right| < \infty$ (qui est satisfaite par tout processus $\text{ARMA}_d(p, q)$) nous avons $N \times E\left[(\overline{y}_i - \mu_i)^2\right] \underset{N \longrightarrow \infty}{\longrightarrow} \sum_{h=-\infty}^{\infty} \gamma_h^{(i)}$, lorsque $N \longrightarrow \infty$ donc la moyenne de la i^{eme} période de l'échantillon $\bar{y}_i$ est asymptotiquement normalement distribuée.

Proposition (3.6.2)

Nous supposons que le processus $\{y_t, t \in \mathbb{Z}\}$ est causal. Alors $\overline{y}_i$ est asymptotiquement normalement distribuée avec une moyenne μ_i et de variance $\frac{1}{N} \sum_{h=-\infty}^{\infty} \gamma_h^{(i)}$.

Démonstration En appliquant le théorème Gladyshev (1961) et le théorème (1.6.2) nous retrouvons le résultat.

3.6.2 Estimation de la fonction d'autocorrélation

Soit $y_1, ..., y_{dN}$ une dN-réalisation d'un processus périodiquement corrélé, d-périodique. Pour plus de simplicité, nous supposerons dans tout le reste de cette sous-section que $d = 2$. Comme dans le cas univarié et multivarié, l'estimateur naturel de la fonction d'autocovariance $\gamma_h^{(i)} = E\left[(y_{i+d\tau} - \mu_i)(y_{i+d\tau-h} - \mu_{i-h})\right]$ est

$$\widehat{\gamma}_h^{(i)} = \frac{1}{N} \sum_{\tau=0}^{N--1} \left(y_{i+d\tau} - \overline{y}\right)\left(y_{i+d\tau-h} - \overline{y}\right). \tag{3.6.3}$$

où $y_{i+d\tau-l} = 0$ pour $i + d\tau - l < 1$ ou pour $i + d\tau - l > dN$.

Remarquons que

$$cov(y_{i+d\tau}, y_{j+d\tau-h}) = cov(y_{i+d\tau}, y_{i+d\tau-(i-j+h)}) = \gamma_{i-j+h}^{(i)}$$

Cette fonction d'autocovariance peut être estimer par :

$$\widehat{\gamma}_{i-j+h}^{(i)} = \frac{1}{N} \sum_{\tau=0}^{N--1} \left(y_{i+d\tau} - \overline{y}\right)\left(y_{j+d\tau-h} - \overline{y}\right), \tag{3.6.4}$$

Si $i = j$, (3.6.4) est réduit à $\overset{\wedge}{\gamma}_h^{(i)}$.

Nous estimons également la fonction d'autocorrélation entre $y_{i+d\tau}$ et $y_{j+d\tau-h}$ par :

$$\widehat{\gamma}_{i-j+h}^{(i)} = \widehat{\gamma}_{i-j+h}^{(i)} \Big/ \widehat{\gamma}_0^{(i)}, h \in \mathbb{Z}. \tag{3.6.5}$$

Nous montrons d'abord la consistance faible de l'estimateur $\widehat{\gamma}_h^{(i)}$ (et par conséquent $\widehat{\rho}_h^{(i)}$) pour des moyennes mobiles périodique d'ordre infini. Puis nous calculerons la distribution asymptotique de $\widehat{\gamma}_h^{(i)}$ et de $\widehat{\rho}_h^{(i)}$ dans quelques cas spéciaux.

Théorème (3.6.1)

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus 2-périodique causal. C'est-à-dire $y_t = \mu_t + \sum_{k=0}^{+\infty} \psi_{t,k}\varepsilon_{t-k}$ où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un bruit blanc périodique de variance σ_t^2 et $\{\psi_{t,k}, t \in \mathbb{Z}; k \geq 0\}$ est une suite telle que $\sum_{k=0}^{\infty} |\psi_{t,k}| < +\infty, \forall t \in \mathbb{Z}$.

Alors, pour tout h fixe $h \geq 0$ et $i = 1, 2$

$$\widehat{\gamma}_h^{(i)} \xrightarrow{\text{Pr}} \gamma_h^{(i)} \qquad \text{lorque } N \longrightarrow \infty \text{ et } \widehat{\rho}_h^{(i)} \xrightarrow{\text{Pr}} \rho_h^{(i)} \qquad \text{lorque } N \longrightarrow \infty.$$

Démonstration En appliquant le théorème Gladyshev (1961) et le théorème (11.2.1) dans Brockwell et Davis (1988) p. 408, nous retrouverons le résultat

En général, la dérivation de la distribution asymptotique de la fonction d'autocorrélation de l'échantillon dans le cas d-périodique est tout à fait compliquée même pour des moyennes mobiles d-périodique. Un cas spécial important se présente lorsque le processus y_t est un processus causal autoregressif gaussien d-périodique. La distribution asymptotique de $\widehat{\gamma}_h^{(i)}$ pour un tel processus est donnée par le théorème suivant :

Théorème (3.6.2) (*Pagano (1978)*)

Si y_t est un processus causal autoregressif gaussien d-périodique, alors

i) $\widehat{\gamma}(i,j) \longrightarrow \gamma(i,j)$ presque sûrement lorsque $N \longrightarrow \infty$.

ii) $E\left((\widehat{\gamma}(i,j) - \gamma(i,j))^2\right) \longrightarrow 0$ lorsque $N \longrightarrow \infty$.

iii) $N \times cov\left(\widehat{\gamma}(i,j), \widehat{\gamma}(k,l)\right) \longrightarrow \sum_{h=-\infty}^{\infty} \begin{bmatrix} \gamma(i,k+dh)\gamma(j,l+dh)+ \\ \gamma(i,l+dh)\gamma(j,k+hd) \end{bmatrix}$

lorsque $N \longrightarrow \infty$.

Démonstration Voir Pagano (1978).

Théorème (3.6.3)

Si y_t est un processus causal autoregressif gaussien d-périodique. Les variables aléatoires$\sqrt{N}\left(\widehat{\gamma}(i,j) - \gamma(i,j)\right), \quad 1 \leq i \leq d, \quad 0 \leq j \leq q$, où q est un entier positif fixe, sont asymptotiquement normalement distribuées avec des moyennes nulles et de covariance donnée dans le théorème (3.6.2).

Démonstration Voir par exemple, Lund et Basawa (1999).

Chapitre 4

Identification des modèles ARMA d-périodiques

Introduction

Depuis la fin des années 70, l'identification des modèles ARMA périodiques est en continuelle expansion. L'objectif principal des travaux de recherche effectués par Cliveland et Tiao (1976), Sakai (1982), Vecchia (1985b), Mcleod (1992), Bentarzi (2000), ...etc. était la généralisation des méthodes classiques relatives aux modèles ARMA stationnaires à coefficients constants, au cas des modèles ARMA périodiques.

La plupart de ces travaux visaient à généraliser les méthodes basées sur la théorie de l'information tel que le critère AIC ou celles basée sur l'analyse baysien, tels que les critères BIC et PDC.

Hemis (1999) et Bentarzi (2000) ont généralisé une méthode basée sur l'analyse baysien (*Critère de la densité prédictive* (PDC)) introduite par Djuric et Kay (1992). Dans ce même travail, Hemis (1999) s'est intéressée d'une part à l'estimation, par la technique de Gibbs, des paramètres du modèle autorégressif périodique (PAR) à tendance explicative. D'autre part, elle s'est intéressée à l'étude de l'effet de l'application des estimateurs obtenus sur la performance des critères de sélection de l'ordre d'un processus autorégressif pur, en introduisant la technique de Gibbs.

Notre objectif dans ce chapitre est d'une part d'étudier théoriquement les critères de sélection des modèles PARMA d-périodiques et d'autre part, effectuer une comparaison entre la performance des ces différents critères. Nous commencerons ce chapitre par la généralisation de la méthode du coin de Beguin et *al.* (1980) du cas stationnaire au cas périodique. Nous présenterons à ce sujet une caractérisation d'un processus ARMA périodique par des équations aux différences stochastique sur sa fonction d'autocorrélation. Nous passerons ensuite, à l'exposer de quelques relations utiles entre les coefficients du modèle (les équations de

Yule-Walker périodiques). Dans la même sous-section, nous exposerons d'autres relations entre les déterminants d'autocorrélation et les coefficients du modèle. Nous allons par la suite répondre à la question suivante : Sous quelles conditions sur les déterminants d'autocorrélation, un processus ARMA d-périodique admet une représentation minimale $\text{ARMA}_d(p,q)$? Nous terminerons cette section par l'énoncé du théorème du coin, dans le cas des modèles ARMA périodiques. Dans la deuxième section, nous exposerons les critères AIC, BIC et HQC dans le cas des modèles autorégressifs moyennes mobiles d-périodiques. Nous passerons en revue par la suite, les travaux de Hemis (1999) et Bentarzi (2000) sur l'identification des modèles autorégressifs périodiquement corrélés. Nous exposerons les principaux résultats obtenus. Nous récapitulerons ensuite, la généralisation du critère CIC (*Consistent Information Criterion*) aux cas des modèles autorégressifs purs d-périodiques. Nous terminerons enfin ce chapitre, par une étude de simulation, en comparant les performances des critères étudiés appliqués à des séries artificielles supposées obtenues à partir des modèles bien spécifiés.

4.1 Identification d'un modèle PARMA par la méthode du coin

4.1.1 Caractérisation d'un processus PARMA par des équations aux différences stochastique sur sa fonction d'autocorrélation

Rappelons qu'un processus $\{y_t, t \in \mathbb{Z}\}$ est dit autoregressif moyenne mobile périodiquement corrélé d-périodique, s'il est solution de l'équation aux différences stochastique suivante :

$$y_t - \sum_{j=1}^{p_t} \phi_{t,j} y_{t-j} = \varepsilon_t - \sum_{j=1}^{q_t} \theta_{t,j} \varepsilon_{t-j}, \quad t \in \mathbb{Z} \tag{4.1.1}$$

où encore

$$\Phi_t(L) y_t = \Theta_t(L) \varepsilon_t, \quad t \in \mathbb{Z}$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un bruit blanc périodique de variance σ_t^2, les paramètres $\phi_{t,j},\ j = 1, ..., p_t,\quad \theta_{t,j},\ j = 1, ..., q_t,\ p_t$ et q_t sont des fonctions d-périodiques et les polynômes $\Phi_t(B) = 1 - \sum_{i=1}^{p_t} \phi_{t,i} L^i$ et $\Theta_t(B) = 1 - \sum_{i=1}^{q_t} \theta_{t,i} L^i$ ne possèdent pas de zéros communs.

Proposition (4.1.1)
Si $\{y_t, t \in \mathbb{Z}\}$ est un processus autoregressif moyenne mobile périodiquement corrélé d-périodique, alors sa fonction d'autocorrélation $\rho_h^{(i)}$ pour l'horizon h et

relative à la i^{eme} période $i = 1, ..., d, \quad h \in \mathbb{N}$, vérifie une équation aux différences d'ordre p_i à partir du rang q_i notée (p_i, q_i)

$$\rho_{-h}^{(i)} - \sum_{j=1}^{p_i} \phi_{i,j} \rho_{-h+j}^{(i-j)} = \begin{cases} 0 & si \quad h \geq q_i + 1 \\ -\theta_{i,q_i} \frac{\sigma_{i-q_i}^2}{\gamma_0^{(i)}} & si \quad h = q_i \end{cases}$$

ceci $\forall i = 1, ..., d$.

Démonstration

1- Montrons la première relation :

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus autoregressif moyenne mobile périodiquement corrélé d-périodique solution de l'équation aux différences stochastique (4.1.1). Puisque ε_t est un processus bruit blanc, nous avons :

$$cov(\varepsilon_{t-j}, y_{t-q_t-k}) = 0 \quad pour \quad k \geq 1, \ j = 0, 1, ..., q_t$$

Multiplions l'équation (4.1.1) par y_{t-h} et passons à l'espérance, nous obtenons alors,

$$\begin{aligned} E\left[\left(y_t - \textstyle\sum_{j=1}^{p_t} \phi_{t,j} y_{t-j}\right) y_{t-h}\right] &= E\left[\left(\varepsilon_t - \textstyle\sum_{j=1}^{q_t} \theta_{t,j} \varepsilon_{t-j}\right) y_{t-h}\right] \\ E\left[\left(y_{i+d\tau} - \textstyle\sum_{j=1}^{p_i} \phi_{i,j} y_{i+d\tau-j}\right) y_{i+d\tau-h}\right] &= E\left[\left(\varepsilon_{i+d\tau} - \textstyle\sum_{j=1}^{q_i} \theta_{i,j} \varepsilon_{i+d\tau-j}\right) y_{i+d\tau-h}\right] \end{aligned}$$

D'où

$$\rho_{-h}^{(i)} - \sum_{j=1}^{p_i} \phi_{i,j} \rho_{-h+j}^{(i-j)} = 0 \quad pour \quad h \geq q_i + 1 \tag{4.1.2}$$

2- Montrons la seconde relation

Multiplions (4.1.1) par y_{t-q_t} et passons à l'espérance nous obtenons

$$E\left[\left(y_{i+d\tau} - \sum_{j=1}^{p_i} \phi_{i,j} y_{i+d\tau-j}\right) y_{i+d\tau-q_i}\right] = E\left[\left(\varepsilon_{i+d\tau} - \sum_{j=1}^{q_i} \theta_{i,j} \varepsilon_{i+d\tau-j}\right) y_{i+d\tau-q_i}\right]$$

d'où

$$\rho_{-q_i}^{(i)} - \sum_{j=1}^{p_i} \phi_{ij} \rho_{-q_i+j}^{(i-j)} = -\theta_{i,q_i} \frac{\sigma_{i-q_i}^2}{\gamma_0^{(i)}} \neq 0 \tag{4.1.3}$$

Proposition (4.1.2)

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus stochastique, $\{\varepsilon_t, t \in \mathbb{Z}\}$ son processus d'innovation et $\left\{\rho_h^{(i)} \Big/ \ h \in N, \ i = 1, ..., d\right\}$ sa fonction d'autocorrélation vérifiant une équation aux différences (p_i, q_i)

$$\rho_{-h}^{(i)} - \sum_{j=1}^{p_i} \phi_{i,j} \rho_{-h+j}^{(i-j)} = \begin{cases} 0 & si \quad h \geq q_i + 1 \\ -\theta_{i,q_i} \frac{\sigma_{i-q_i}^2}{\gamma_0^{(i)}} & si \quad h = q_i \end{cases}$$

Alors, le processus $\{y_t, t \in \mathbb{Z}\}$ est un PARMA défini par :

$$y_{i+d\tau} - \sum_{j=1}^{p_i} \phi_{i,j} y_{i+d\tau-j} = \varepsilon_{i+d\tau} - \sum_{j=1}^{q_i} \theta_{i,j} \varepsilon_{i+d\tau-j}, \qquad i \in \mathbb{N}, \tau \in \mathbb{Z},$$

où $\phi_{i,j}, j = 1, ..., p_i$ sont les fonctions d-périodique de l'équation aux différences et les $\theta_{i,j}, j = 1, ..., q_i$ sont aussi d-périodique et déterminés d'une façon unique.

Démonstration

Posons

$$A_t = y_t - \sum_{j=1}^{p_t} \phi_{t,j} y_{t-j} - \varepsilon_t + \sum_{j=1}^{q_t} \theta_{t,j} \varepsilon_{t-j} \tag{4.1.4}$$

où les fonctions $\phi_{t,j}$ sont celles de l'équation aux différences et les $\theta_{t,j}$ sont des suites de fonctions réelles.

Premièrement, nous montrons que $A_t \in H^2(y_t, t)$: Nous avons

$$\begin{aligned} A_t &= y_t - \textstyle\sum_{j=1}^{p_t} \phi_{t,j} y_{t-j} - \varepsilon_t + \sum_{j=1}^{q_t} \theta_{t,j} \varepsilon_{t-j} \\ &= (y_t - \varepsilon_t) - \textstyle\sum_{j=1}^{p_t} \phi_{t,j} y_{t-j} + \sum_{j=1}^{q_t} \theta_{t,j} \varepsilon_{t-j} \\ &= \pi_t(y_t) - \textstyle\sum_{j=1}^{p_t} \phi_{t,j} y_{t-j} + \sum_{j=1}^{q_t} \theta_{t,j} \varepsilon_{t-j} \end{aligned}$$

où $\pi_t(y_t)$ est la projection de y_t sur l'espace de Hilbert $H^2(y_t, t)$.

Nous remarquons que chaque terme de cette somme appartient à $H^2(y_t, t)$, donc $A_t \in H^2(y_t, t)$.

Deuxièment, nous montrons que $A_t = 0$ c'est-à-dire que $A_t \perp H^2(y_t, t)$ ou bien

$$\langle A_t, y_{t-h} \rangle = 0, h \geq 1 \tag{4.1.5}$$

a) Cette relation est vérifiée indépendamment du choix des fonctions $\theta_{t,j}$ pour $h \geq q_t + 1$, vu l'équation aux différences et le fait que ε_t soit un processus d'innovation.

b) Nous supposons que $\theta_{t,j} = 0$ si $\sigma^2_{t-j} = 0$, donc la famille $\theta_{t,j}$ est déterminée de proche en proche, en écrivant les équations (4.1.5) pour $1 \leq h < q_t + 1$:

Pour $h = q_t$ nous avons :

$$\begin{aligned} \langle A_t, y_{t-q_t} \rangle &= \langle (y_t - \varepsilon_t), y_{t-q_t} \rangle - \textstyle\sum_{j=1}^{p_t} \phi_{t,j} \langle y_{t-j}, y_{t-q_t} \rangle + \sum_{j=1}^{q_t} \theta_{t,j} \langle \varepsilon_{t-j}, y_{t-q_t} \rangle \\ &= \gamma^{(i)}_{-q_i} - \textstyle\sum_{j=1}^{p_i} \phi_{i,j} \gamma^{(i-j)}_{-q_i+j} + \theta_{i,q_i} \sigma^2_{i-q_i} \end{aligned}$$

Soit

$$-\theta_{i,q_i} \sigma^2_{i-q_i} = \gamma^{(i)}_{-q_i} - \sum_{j=1}^{p_i} \phi_{i,j} \gamma^{(i-j)}_{-q_i+j}$$

Comme le second mombre de cette équation est non nul, θ_{i,q_i} est déterminé d'une façon unique. Ce choix de $-\theta_{i,q_i}$ implique que

$$\langle A_t, y_{t-q_t}\rangle = 0.$$

La même chose pour les autres coefficients $\theta_{i,j}$ avec $1 \leq j < q_i$.

$$\begin{aligned}\langle A_t, y_{t-j}\rangle &= \langle (y_t - \varepsilon_t), y_{t-j}\rangle - \textstyle\sum_{k=1}^{p_t} \phi_{t,k}\langle y_{t-k}, y_{t-j}\rangle + \sum_{k=1}^{q_t} \theta_{t,k}\langle \varepsilon_{t-k}, y_{t-j}\rangle \\ &= \gamma_{-j}^{(i)} - \textstyle\sum_{k=1}^{p_i} \phi_{i,j}\gamma_{-j+k}^{(i-k)} + \theta_{i,j}\sigma_{i-j}^2 + \sum_{k=j+1}^{q_t} \theta_{t,k}\langle \varepsilon_{i+d\tau-k}, y_{i+d\tau-j}\rangle\end{aligned}$$

Soit

$$-\theta_{i,j}\sigma_{i-j}^2 = \gamma_{-j}^{(i)} - \sum_{k=1}^{p_i} \phi_{i,j}\gamma_{-j+k}^{(i-k)} + \sum_{k=j+1}^{q_t} \theta_{t,k}\langle \varepsilon_{i+d\tau-k}, y_{i+d\tau-j}\rangle$$

Les coefficients sont également déterminés d'une façon unique. D'où

$$\langle A_t, y_{t-q_t+1}\rangle = 0, ..., \langle A_t, y_{t-q_t+j}\rangle = 0 \quad et \quad \langle A_t, y_{t-1}\rangle = 0$$

Donc, la connaisance des $\phi_{i,j}$ et ce choix des $\theta_{i,j}$ montrent que y_t est un processus PARMA qui vérifi l'équation aux différences stochastique suivante :

$$y_{i+d\tau} - \sum_{j=1}^{p_i} \phi_{i,j} y_{i+d\tau-j} = \varepsilon_{i+d\tau} - \sum_{j=1}^{q_i} \theta_{i,j}\varepsilon_{i+d\tau-j}, \qquad i \in \mathbb{N}, \tau \in \mathbb{Z}$$

Remarque Soit $\{y_t, t \in \mathbb{Z}\}$ un processus stochastique, $\{\varepsilon_t, t \in \mathbb{Z}\}$ son processus d'innovation et $\left\{ \rho_h^{(i)} \Big/ \ h \in N,\ i = 1, ..., d \right\}$ sa fonction d'autocorrélation vérifiant une équation aux différences (p_i, q_i)

$$\rho_{-h}^{(i)} - \sum_{j=1}^{p_i} \phi_{i,j}\rho_{-h+j}^{(i-j)} = 0 \quad pour \quad h \geq q_i + 1$$

Alors, y_t est un PARMA défini par

$$y_{i+d\tau} - \sum_{j=1}^{p_i} \phi_{i,j} y_{i+d\tau-j} = \varepsilon_{i+d\tau} - \sum_{j=1}^{q_i^{'}} \theta_{i,j}\varepsilon_{i+d\tau-j}, \qquad i \in N, \tau \in Z$$

où q_i' est le plus grand entier h pour lequel

$$\rho_{-h}^{(i)} - \sum_{j=1}^{p_i} \phi_{i,j}\rho_{-h+j}^{(i-j)} \neq 0$$

et nous somme alors ramené au théorème précédent.

Théorème de caractérisation

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus stochastique, $\{\varepsilon_t, t \in \mathbb{Z}\}$ son processus d'innovation.

Le processus y_t est un autoregressif moyenne mobile périodiquement corrélé d-perirdique si et seulement si pour chaque période $i = 1, ..., d$, sa fonction d'autocorrélation $\left\{\rho_h^{(i)} \Big/ h \in N,\ i = 1, ..., d\right\}$ vérifie une équation aux différences (p_i, q_i)

$$\rho_{-h}^{(i)} - \sum_{j=1}^{p_i} \phi_{i,j} \rho_{-h+j}^{(i-j)} = \begin{cases} 0 & si \quad h \geq q_i + 1 \\ -\theta_{i,q_i} \frac{\sigma_{i-q_i}^2}{\gamma_0^{(i)}} & si \quad h = q_i \end{cases}$$

Démonstration La condition nécéssaire est vérifiée d'après la proposition (4.1.1) et la condition suffisante est aussi vérifiée d'aprè la proposition (4.1.2).

4.1.2 Déterminant d'autocorrélation

Rappelons que la matrice d'autocorrélations d'ordre k et relative à la i^{eme} période d'un processus périodiquement corrélé est définie par :

$$\Sigma^{(i)}(k,h) = \begin{pmatrix} \rho_{-h}^{(i-1)} & \rho_{-h+1}^{(i-2)} & \cdots & \rho_{-h+k-1}^{(i-k)} \\ \rho_{-h-1}^{(i-1)} & \rho_{-h}^{(i-2)} & \cdots & \rho_{-h+k-2}^{(i-k)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{-h-k+1}^{(i-1)} & \rho_{-h-k+2}^{(i-1)} & \cdots & \rho_{-h}^{(i-k)} \end{pmatrix},$$

où $\rho_h^{(i)} = \frac{\gamma_h^{(i)}}{\gamma_0^{(i)}}$ est la fonction d'autocorrélation du processus $\{y_t,\ t = i + d\tau \in \mathbb{Z}\}$ pour l'horizon h et relative à la i^{eme} période $i = 1, ..., d, \quad h, k \in \mathbb{N}$ et $\tau \in \mathbb{Z}$. Nous notons par $\Delta^{(i)}(k,h)$ le déterminant de la matrice $\Sigma^{(i)}(k,h)$, et appelons l'ensemble $\left\{\Delta^{(i)}(k,h)/k = 1, 2, ...h = 1, 2, ...\right\}$ l'ensemble $\Delta^{(i)}$ relative à la $i \overset{eme}{=}$ période du processus y_t.

Nous présenterons dans ce qui suit, quelques relations utiles, entres les déterminants $\Delta^{(i)}(k,h)$ et les coefficients d'un modèle $\mathrm{ARMA}_d(p_t, q_t)$.

Considèrons le modèle PARMA de la representation (4.1.1). Celle-ci peut s'écrire comme suit :

$$y_{i+d\tau} - \sum_{j=1}^{p_i} \phi_{i,j} y_{i+d\tau-j} = \varepsilon_{i+d\tau} - \sum_{j=1}^{q_i} \theta_{i,j} \varepsilon_{i+d\tau-j}, \quad i = 1, ..., d, \quad \tau \in \mathbb{Z} \quad (4.1.6)$$

Proposition (4.1.3)

Les déterminants $\Delta^{(i)}(n,m)$ vérifiont les deux propriétés suivantes

$$\Delta^{(i+1)}(p_i + n, q_i) = -\theta_{i,q_i} \frac{\sigma_{i-q_i}^2}{\gamma_0^{(i)}} \Delta^{(i)}(p_i + n - 1, q_i) \qquad n \geq 1$$

$$\Delta^{(i+1)}(p_i, q_i + m) = (-1)^{p_i - 1} \phi_{i,p_i} \Delta^{(i)}(p_i, q_i + m - 1) \qquad m \geq 1$$

Démonstration

Nous avons par définition

$$\Delta^{(i+1)}(k,h)=\begin{vmatrix}\rho^{(i)}_{-h} & \rho^{(i-1)}_{-h+1} & \cdots & \rho^{(i-k+1)}_{-h+k-1}\\ \rho^{(i)}_{-h-1} & \rho^{(i-1)}_{-h} & \cdots & \rho^{(i-k+1)}_{-h+k-2}\\ \vdots & \vdots & \ddots & \vdots\\ \rho^{(i)}_{-h-k+1} & \rho^{(i-1)}_{-h-k+2} & \cdots & \rho^{(i-k+1)}_{-h}\end{vmatrix}$$

donc

$$\Delta^{(i+1)}(p_i+n,q_i)=\begin{vmatrix}\rho^{(i)}_{-q_i} & \rho^{(i-1)}_{-q_i+1} & \cdots & \rho^{(i-p_i-n+1)}_{-q_i+p_i+n-1}\\ \rho^{(i)}_{-q_i-1} & \rho^{(i-1)}_{-q_i} & \cdots & \rho^{(i-p_i-n+1)}_{-q_i+p_i+n-2}\\ \vdots & \vdots & \ddots & \vdots\\ \rho^{(i)}_{-q_i-p_i-n+1} & \rho^{(i-1)}_{-q_i-p_i-n+2} & \cdots & \rho^{(i-p_i-n+1)}_{-q_i}\end{vmatrix}$$

Posons $C'_1 = C_1 - \sum_{j=2}^{p_i}\phi_{i,j}C_j$ où C_j est la j^{eme} colonne de la matrice $\Sigma^{(i+1)}(p_i+n,q_i)$. D'après la relation (4.1.2) et (4.1.3), nous trouvons

$$\Delta^{(i+1)}(p_i+n,q_i)=\begin{vmatrix}-\theta_{i,q_i}\frac{\sigma^2_{i-q_i}}{\gamma_0^{(i)}} & \rho^{(i-1)}_{q_i-1} & \cdots & \rho^{(i-p_i+n)}_{q_i-k+1}\\ 0 & \rho^{(i-1)}_{q_i} & \cdots & \rho^{(i-p_i+n)}_{q_i-p_i+n+2}\\ \vdots & \vdots & \ddots & \vdots\\ 0 & \rho^{(i-1)}_{q_i+p_i+n-2} & \cdots & \rho^{(i-p_i+n)}_{q_i}\end{vmatrix}$$

D'ou

$$\Delta^{(i+1)}(p_i+n,q_i) = -\theta_{i,q_i}\frac{\sigma^2_{i-q_i}}{\gamma_0^{(i)}}\Delta^{(i)}(p_i+n-1,q_i) \qquad \textit{pour} \quad n\geq 1 \qquad (4.1.7)$$

D'autre part, nous avons

$$\Delta^{(i+1)}(p_i,q_i+m)=\begin{vmatrix}\rho^{(i)}_{-q_i-m} & \rho^{(i-1)}_{-q_i-m+1} & \cdots & \rho^{(i-p_i+1)}_{-q_i-m+p_i-1}\\ \rho^{(i)}_{-q_i-m-1} & \rho^{(i-1)}_{-q_i-m} & \cdots & \rho^{(i-p_i+1)}_{-q_i-m+p_i-2}\\ \vdots & \vdots & \ddots & \vdots\\ \rho^{(i)}_{-q_i-m-p_i+1} & \rho^{(i-1)}_{-q_i-m-p_i+2} & \cdots & \rho^{(i-p_i+1)}_{-q_i-m}\end{vmatrix}$$

$$=(-1)^{p_i-1}\begin{vmatrix}\rho^{(i)}_{-q_i-m-p_i+1} & \rho^{(i-1)}_{-q_i-m-p_i+2} & \cdots & \rho^{(i-p_i+1)}_{-q_i-m}\\ \rho^{(i)}_{-q_i-m} & \rho^{(i-1)}_{-q_i-m+1} & \cdots & \rho^{(i-p_i+1)}_{-q_i-m+p_i-1}\\ \vdots & \vdots & \ddots & \vdots\\ \rho^{(i)}_{-q_i-m-p_i+2} & \rho^{(i-1)}_{-q_i-m-p_i+3} & \cdots & \rho^{(i-p_i+1)}_{-q_i-m+1}\end{vmatrix}$$

$$=(-1)^{p_i-1}\Sigma'^{(i+1)}(p_i,q_i+m)$$

Posons $L'_{p_i} = L_{p_i} - \sum_{j=2}^{p_i} \phi_{i,j} L_j$ où L_j est la $j - p_i + 1 \overset{eme}{=}$ ligne de la matrice $\Sigma^{(i+1)\prime}(p_i, q_i + m)$. D'après (4.1.2) et (4.1.3), nous trouvons

$$\Delta^{(i+1)}(p_i, q_i + m) = (-1)^{p_i - 1} \begin{vmatrix} \phi_{i,p_i}\rho^{(i-1)}_{-q_i-m+1} & \phi_{i,p_i}\rho^{(i-2)}_{-q_i-m+2} & \cdots & \phi_{i,p_i}\rho^{(i-p_i)}_{-q_i-m+p_i} \\ \rho^{(i-1)}_{-q_i-m} & \rho^{(i-2)}_{-q_i-m+1} & \cdots & \rho^{(i-p_i)}_{-q_i-m+p_i-1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i-1)}_{-q_i-m-p_i+2} & \rho^{(i-2)}_{-q_i-m-p_i+3} & \cdots & \rho^{(i-p_i)}_{-q_i-m+1} \end{vmatrix}$$

c'est-à-dire

$$\Delta^{(i+1)}(p_i, q_i + m) = (-1)^{p_i - 1}\phi_{i,p_i} \begin{vmatrix} \rho^{(i-1)}_{-q_i-m+1} & \rho^{(i-2)}_{-q_i-m+2} & \cdots & \rho^{(i-p_i)}_{-q_i-m+p_i} \\ \rho^{(i-1)}_{-q_i-m} & \rho^{(i-2)}_{-q_i-m+1} & \cdots & \rho^{(i-p_i)}_{-q_i-m+p_i-1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i-1)}_{-q_i-m-p_i+2} & \rho^{(i-2)}_{-q_i-m-p_i+3} & \cdots & \rho^{(i-p_i)}_{-q_i-m+1} \end{vmatrix}$$

D'où

$$\Delta^{(i+1)}(p_i, q_i + m) = (-1)^{p_i - 1}\phi_{i,p_i}\Delta^{(i)}(p_i, q_i + m - 1) \qquad \textit{pour} \quad m \geq 1 \quad (4.1.8)$$

Remarque
Si $\Delta^{(i+1)}(p_i, q_i) = 0$, alors $\forall n, m \in \mathbb{N} \quad \Delta^{(i)}(p_i+n, q_i) = 0$ *et* $\Delta^{(i)}(p_i, q_i+m) = 0$.

Il faut noté que lorsque l'ordre du processus autoregressif moyenne mobile périodiquement corrélé est périodique dans le temps, il suffit de considérer sans perdre de généralité, le processus comme étant un processus autoregressif moyenne mobile d-périodique d'ordre (p, q) constant, tel que $p = \max\limits_{1\leq i\leq d} p_i$ et $q = \max\limits_{1\leq i\leq d} q_i$. Donc, le modèle (4.1.6) peut être écrit sous la forme suivante :

$$y_{i+d\tau} - \sum_{j=1}^{p} \phi_{i,j} y_{i+d\tau-j} = \varepsilon_{i+d\tau} - \sum_{j=1}^{q} \theta_{i,j}\varepsilon_{i+d\tau-j}, \quad i = 1, ..., d, \quad \tau \in \mathbb{Z} \quad (4.1.9)$$

Proposition (4.1.4)
Les déterminants $\Delta^{(i)}(n, m)$ vérifiont les deux propriétés suivantes :

$$\Delta^{(i+1)}(p, q) = -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma^{(i)}_0}\Delta^{(i)}(p - 1, q) + (-1)^{p-1}\phi_{i,p}\Delta^{(i)}(p, q - 1)$$

$$\begin{aligned}\Delta^{(i+1)}(p - 1, q) = &-\theta_{i,q}\tfrac{\sigma^2_{i-q}}{\gamma^{(i)}_0}\Delta^{(i)}(p - 2, q) \\ &+(-1)^{p-2}\phi_{i,p-1}\Delta^{(i)}(p - 1, q - 1) + \phi_{i,p}\Delta,\end{aligned}$$

où

$$\Delta = \begin{vmatrix} \rho^{(i-p)}_{-q+p} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+2)}_{-q+p-2} \\ \rho^{(i-p)}_{-q+p+1} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+2)}_{-q+p-3} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i-p)}_{-q+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+2)}_{-q} \end{vmatrix}.$$

Démonstration

1) Nous avons par définition

$$\Delta^{(i+1)}(p,q) = \begin{vmatrix} \rho^{(i)}_{-q} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+1)}_{-q+p-1} \\ \rho^{(i)}_{-q-1} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+1)}_{-q+p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i)}_{-q-p+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+1)}_{-q} \end{vmatrix}$$

et nous savons que

$$\begin{aligned} \rho^{(i)}_{-h} - \textstyle\sum_{j=1}^{p} \phi_{i,j}\rho^{(i-j)}_{-h+j} &= \rho^{(i)}_{-h} - \phi_{i,1}\rho^{(i-1)}_{-h+1} - \phi_{i,2}\rho^{(i-2)}_{-h+2} - \cdots - \phi_{i,p}\rho^{(i-p)}_{-h+p} \\ &= \begin{cases} -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma^{(i)}_0} & si \quad h = q \\ 0 & si \quad h \geq q+1 \end{cases} \end{aligned}$$

donc

$$\rho^{(i)}_{-q} - \phi_{i,1}\rho^{(i-1)}_{-q+1} - \phi_{i,2}\rho^{(i-2)}_{-q+2} - \cdots - \phi_{i,p-1}\rho^{(i-p+1)}_{-q+p-1} = -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma^{(i)}_0} + \phi_{i,p}\rho^{(i-p)}_{-q+p}$$

et pour $h \geq q+1$

$$\rho^{(i)}_{-h} - \phi_{i,1}\rho^{(i-1)}_{-h+1} - \phi_{i,2}\rho^{(i-2)}_{-h+2} - \cdots - \phi_{i,p-1}\rho^{(i-p+1)}_{-h+p-1} = \phi_{i,p}\rho^{(i-p)}_{-h+p}$$

Posons $C'_1 = C_1 - \sum_{j=2}^{p} \phi_{i,j}C_j$ où C_j est la j ^{eme} colonne de la matrice $\Sigma^{(i+1)}(p,q)$. D'après (4.1.2) et (4.1.3), nous trouvons

$$\Delta^{(i+1)}(p,q) = \begin{vmatrix} -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma^{(i)}_0} + \phi_{i,p}\rho^{(i-p)}_{-q+p} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+1)}_{-q+p-1} \\ \phi_{i,p}\rho^{(i-p)}_{-q-1+p} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+1)}_{-q+p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{i,p}\rho^{(i-p)}_{-q+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+1)}_{-q} \end{vmatrix}$$

$$= -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma^{(i)}_0}\Delta^{(i)}(p-1,q) + \begin{vmatrix} \phi_{i,p}\rho^{(i-p)}_{-q+p} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+1)}_{-q+p-1} \\ \phi_{i,p}\rho^{(i-p)}_{-q-1+p} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+1)}_{-q+p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{i,p}\rho^{(i-p)}_{-q+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+1)}_{-q} \end{vmatrix}$$

$$\Delta^{(i+1)}(p,q) = -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma_0^{(i)}}\Delta^{(i)}(p-1,q) + \phi_{i,p}\begin{vmatrix} \rho^{(i-p)}_{-q+p} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+1)}_{-q+p-1} \\ \rho^{(i-p)}_{-q-1+p} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+1)}_{-q+p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i-p)}_{-q+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+1)}_{-q} \end{vmatrix}$$

$$= -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma_0^{(i)}}\Delta^{(i)}(p-1,q) + (-1)^{p-1}\phi_{i,p}\begin{vmatrix} \rho^{(i-1)}_{-q+1} & \rho^{(i-2)}_{-q+2} & \cdots & \rho^{(i-p)}_{-q+p} \\ \rho^{(i-1)}_{-q} & \rho^{(i-2)}_{-q+1} & \cdots & \rho^{(i-p)}_{-q+p-1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i-1)}_{-q-p+2} & \rho^{(i-2)}_{-q-p+3} & \cdots & \rho^{(i-p)}_{-q-1} \end{vmatrix}$$

D'où

$$\Delta^{(i+1)}(p,q) = -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma_0^{(i)}}\Delta^{(i)}(p-1,q) + (-1)^{p-1}\phi_{i,p}\Delta^{(i)}(p,q-1) \qquad (4.1.10)$$

2) Nous savons que

$$\Delta^{(i+1)}(p,q) = \begin{vmatrix} \rho^{(i)}_{-q} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+1)}_{-q+p-1} \\ \rho^{(i)}_{-q-1} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+1)}_{-q+p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i)}_{-q-p+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+1)}_{-q} \end{vmatrix}$$

De même nous avons

$$\Delta^{(i+1)}(p-1,q) = \begin{vmatrix} \rho^{(i)}_{-q} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+2)}_{-q+p-2} \\ \rho^{(i)}_{-q-1} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+2)}_{-q+p-3} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i)}_{-q-p+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+2)}_{-q} \end{vmatrix}$$

Sachant que

$$\rho^{(i)}_{-q} - \phi_{i,1}\rho^{(i-1)}_{-q+1} - \cdots - \phi_{i,p-1}\rho^{(i-p+2)}_{-q+p-2} = -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma_0^{(i)}} + \phi_{i,p-1}\rho^{(i-p+1)}_{-q+p-1} + \phi_{i,p}\rho^{(i-p)}_{-q+p}$$

et pour $h \geq q+1$

$$\rho^{(i)}_{-h} - \phi_{i,1}\rho^{(i-1)}_{-h+1} - \cdots - \phi_{i,p-1}\rho^{(i-p+2)}_{-h+p-2} = \phi_{i,p-1}\rho^{(i-p+1)}_{-h+p-1} + \phi_{i,p}\rho^{(i-p)}_{-h+p}$$

Posons $C_1' = C_1 - \sum_{j=2}^{p} \phi_{i,j} C_j$ où C_j est la j ^{eme} colonne de la matrice $\Sigma^{(i+1)}(p-1,q)$. Alors,

$\Delta^{(i+1)}(p-1,q) =$

$$\begin{vmatrix} -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma^{(i)}_0} + \phi_{i,p-1}\rho^{(i-p+1)}_{-q+p-1} + \phi_{i,p}\rho^{(i-p)}_{-q+p} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+2)}_{-q+p-2} \\ \phi_{i,p-1}\rho^{(i-p+1)}_{-q+p-2} + \phi_{i,p}\rho^{(i-p)}_{-q+p+1} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+2)}_{-q+p-3} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{i,p-1}\rho^{(i-p+1)}_{-q} + \phi_{i,p}\rho^{(i-p)}_{-q+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+2)}_{-q} \end{vmatrix}$$

$$\Delta^{(i+1)}(p-1,q) = -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma^{(i)}_0}\Delta^{(i+1)}(p-2,q)$$

$$+ \begin{vmatrix} \phi_{i,p-1}\rho^{(i-p+1)}_{-q+p-1} + \phi_{i,p}\rho^{(i-p)}_{-q+p} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+2)}_{-q+p-2} \\ \phi_{i,p-1}\rho^{(i-p+1)}_{-q+p-2} + \phi_{i,p}\rho^{(i-p)}_{-q+p+1} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+2)}_{-q+p-3} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{i,p-1}\rho^{(i-p+1)}_{-q} + \phi_{i,p}\rho^{(i-p)}_{-q+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+2)}_{-q} \end{vmatrix}$$

$$\Delta^{(i+1)}(p-1,q) = -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma^{(i)}_0}\Delta^{(i+1)}(p-2,q)$$

$$+\phi_{i,p-1} \begin{vmatrix} \rho^{(i-p+1)}_{-q+p-1} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+2)}_{-q+p-2} \\ \rho^{(i-p+1)}_{-q+p-2} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+2)}_{-q+p-3} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i-p+1)}_{-q} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+2)}_{-q} \end{vmatrix}$$

$$+\phi_{i,p} \begin{vmatrix} \rho^{(i-p)}_{-q+p} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+2)}_{-q+p-2} \\ \rho^{(i-p)}_{-q+p+1} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+2)}_{-q+p-3} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i-p)}_{-q+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+2)}_{-q} \end{vmatrix}$$

$$\Delta^{(i+1)}(p-1,q) = -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma_0^{(i)}}\Delta^{(i+1)}(p-2,q)$$

$$+(-1)^{p-2}\phi_{i,p-1}\begin{vmatrix} \rho^{(i-1)}_{-q+1} & \rho^{(i-2)}_{-q+2} & \cdots & \rho^{(i-p+1)}_{-q+p-1} \\ \rho^{(i-1)}_{-q} & \rho^{(i-2)}_{-q+1} & \cdots & \rho^{(i-p+1)}_{-q+p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i-1)}_{-q-p+2} & \rho^{(i-2)}_{-q-p+3} & \cdots & \rho^{(i-p+1)}_{-q} \end{vmatrix}$$

$$+\phi_{i,p}\begin{vmatrix} \rho^{(i-p)}_{-q+p} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+2)}_{-q+p-2} \\ \rho^{(i-p)}_{-q+p+1} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+2)}_{-q+p-3} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i-p)}_{-q+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+2)}_{-q} \end{vmatrix}$$

D'où

$$\begin{aligned}\Delta^{(i+1)}(p-1,q) &= -\theta_{i,q}\frac{\sigma^2_{i-q}}{\gamma_0^{(i)}}\Delta^{(i)}(p-2,q) \\ &\quad +(-1)^{p-2}\phi_{i,p-1}\Delta^{(i)}(p-1,q-1)+\phi_{i,p}\Delta\end{aligned} \tag{4.1.11}$$

où

$$\Delta = \begin{vmatrix} \rho^{(i-p)}_{-q+p} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+2)}_{-q+p-2} \\ \rho^{(i-p)}_{-q+p+1} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+2)}_{-q+p-3} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i-p)}_{-q+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+2)}_{-q} \end{vmatrix}$$

Théorème (4.1.1)

Si $\{y_t, t\in\mathbb{Z}\}$ est un processus autoregressif moyenne mobile périodiquement corrélé d-périodique (PARMA), vérifiant l'équation aux différences stochastique

$$y_{i+d\tau} - \sum_{j=1}^{p_i}\phi_{i,j}y_{i+d\tau-j} = \varepsilon_{i+d\tau} - \sum_{j=1}^{q_i}\theta_{i,j}\varepsilon_{i+d\tau-j}, \qquad i=1,...,d, \tau\in\mathbb{Z}$$

Alors, $\Delta^{(i+1)}(p_i+n, q_i+m)=0, \quad pour \quad n\geq 1, m\geq 1$

Démonstration

Soit $\{y_t, t\in\mathbb{Z}\}$ un processus PARMA, d'après le théorème de caractérisation nous avons

$$\rho^{(i)}_{-h} - \sum_{j=1}^{p_i}\phi_{i,j}\rho^{(i-j)}_{-h+j} = 0 \quad pour \quad h\geq q_i+1$$

Nous avons aussi par définition

$$\Delta^{(i+1)}(k,h)=\begin{vmatrix}\rho_{-h}^{(i)} & \rho_{-h+1}^{(i-1)} & \cdots & \rho_{-h+k-1}^{(i-k+1)}\\ \rho_{-h-1}^{(i)} & \rho_{-h}^{(i-1)} & \cdots & \rho_{-h+k-2}^{(i-k+1)}\\ \vdots & \vdots & \ddots & \vdots\\ \rho_{-h-k+1}^{(i)} & \rho_{-h-k+2}^{(i-1)} & \cdots & \rho_{-h}^{(i-k+1)}\end{vmatrix}$$

donc

$$\Delta^{(i+1)}(p_i+n,q_i+m)=\begin{vmatrix}\rho_{-q_i-m}^{(i)} & \rho_{-q_i-m+1}^{(i-1)} & \cdots & \rho_{-q_i-m+p_i+n-1}^{(i-p_i-n+1)}\\ \rho_{-q_i-m-1}^{(i)} & \rho_{-q_i-m}^{(i-1)} & \cdots & \rho_{-q_i-m+p_i+n-2}^{(i-p_i-n+1)}\\ \vdots & \vdots & \ddots & \vdots\\ \rho_{-q_i-m-p_i-n+1}^{(i)} & \rho_{-q_i-m-p_i-n+2}^{(i-1)} & \cdots & \rho_{-q_i-m}^{(i-p_i-n+1)}\end{vmatrix}$$

Pour $m \geq 1$ et $n \geq 1$, nous remarquons que chaque colonne du déterminant précédant est une combinaison linéaire des autres d'après la relation (4.1.2), donc le déterminant $\Delta^{(i+1)}(p_i+n,q_i+m)$ est nul.

4.1.3 Représentations minimales d'un modèle ARMA$_d(p_t,q_t)$

Nous proposons les définitions suivantes de la minimalité dans la famille des équations aux différences $\{(p_{t,j},q_{t,j})\}_{j\in J}$.

4.1.3.1 Définitions

Définition (4.1.1) Nous disons que l'équation aux différences est d'ordre minimal p_t à partir du rang minimal q_t si :

$$\begin{cases}\text{en posant } \alpha=\inf(p_{t,j}+q_{t,j}) \text{ et } J_0=\{j\ /\ p_{t,j}+q_{t,j}=\alpha\}\\ q_t=\inf\limits_{j\in J_0} q_{t,j} \text{ et } p_t=\alpha-q_t\end{cases}$$

Remarque La dernière définition signifie qu'il ne peut pas exister p'_t et q'_t vérifiant :

$$\begin{cases}p'_t+q'_t<p_t+q_t\\ \quad ou\\ p'_t+q'_t<p_t+q_t \text{ et } q'_t<q_t\end{cases}$$

Définition (4.1.2)

i) La fonction d'autocorrélation $\left\{\rho_h^{(i)}/h\in N, i=1,...,d\right\}$ admet une équation aux différences d'ordre minimal p_t à partir du rang minimal q_t si :

$$\begin{cases}\text{en posant } \alpha=\inf(p_{t,j}+q_{t,j}) \text{ et } J_0=\{j/p_{t,j}+q_{t,j}=\alpha\}\\ q_t=\inf\{q_{t,j}\}_{j\in J_0} \text{ et } p_t=\alpha-q_t\end{cases}$$

$ii)$ Nous disons que $\{y_t, t \in \mathbb{Z}\}$ a une représentation minimale d'ordre p_t et q_t si l'équation aux différences sur sa fonction d'autocorrélation est d'ordre minimal p_t à partir du rang minimal q_t.

Définition (4.1.3) Nous disons que $\{y_t, t \in \mathbb{Z}\}$ est un processus $\text{ARMA}_d(p_t, q_t)$, s'il a une représentation minimale d'ordre p_t et q_t.

Dans la suite de cette section, nous nous proposerons d'étudier le rapport entre la minimalité et la nullité du déterminant d'autocorrélation $\Delta^{(i)}(p_i, q_i)$, afin de donner la méthode du coin pour identifier les processus PARMA. Nous étudierons sans perdre de généralité, le cas où les ordres sont constants.

4.1.3.2 Caractérisation de la représentation minimale d'un modèle ARMA périodiquement corrélé d'ordres constants par les déterminants $\Delta^{(i)}$

Corollaire (4.1.1)

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus autoregressif moyenne mobile périodiquement corrélé d-périodique d'ordres constants p et q tels que : $\forall i = 1, ..., d$, $\Delta^{(i)}(p, q) \neq 0$.

Alors, le processus $\{y_t, t \in \mathbb{Z}\}$ est un processus $\text{ARMA}_d(p, q)$. En plus, il a une représentation minimale unique.

Démonstration

$i)$ Montrons par l'absurde que y_t a une représentation minimale :

Supposons que la représentation de $\{y_t, t \in \mathbb{Z}\}$ n'est pas minimale pour tout t, donc il existe $t_0 \in \mathbb{Z}$ ($t_0 = i_0 + d\tau_0$ où $i_0 \in \{1, ..., d\}$ et $\tau_0 \in \mathbb{Z}$) le processus y_t admet une représentation d'ordre p'_{i_0} et q'_{i_0} avec $p'_{i_0} < p$ ou $q'_{i_0} < q$. Nous allons montrer que la seul condition $p'_{i_0} < p$ ou $q'_{i_0} < q$ contre dit l'hyposèse $\Delta^{(i)}(p, q) \neq 0 \quad \forall i = 1, ..., d$.

Nous distinguons deux cas :

$1^{\underline{er}}$ cas $p'_{i_0} < p$:

Si $p'_{i_0} < p$ deux cas sont possibles pour q'_{i_0}

$a)$ Si $q'_{i_0} < q$, donc $\exists n, m \geq 1$ tel que

$$p'_{i_0} + n = p \quad et \quad q'_{i_0} + m = q$$

et nous avons alors d'après le théorème (4.1.1)

$$\Delta^{(i_0+1)}(p, q) = \Delta^{(i_0+1)}(p'_{i_0} + n, q'_{i_0} + m) = 0$$

et nous sommes alors en contradiction avec l'hypothèse.

$b)$ Si $q'_{i_0} \geq q$, donc $\exists n, m, r \geq 1$ tel que

$$q'_{i_0} + m = q + r \quad et \quad p'_{i_0} + n = p.$$

et nous avons alors,

$$\Delta^{(i_0+1)}(p, q+r) = \Delta^{(i_0+1)}(p'_{i_0} + n, q'_{i_0} + m) = 0$$

mais d'après la proposition (4.1.3)

$$\Delta^{(i_0+1)}(p, q+r) = (-1)^{r(p-1)} \phi_{i_0,p} \phi_{i_0-1,p} \cdots \phi_{i_0-r+1,p} \Delta^{(i_0-r+1)}(p, q) \neq 0$$

d'où la contradiction.

$2^{\underline{eme}}$ cas $q'_{i_0} < q$:

Démonstration analogue à celle du premier cas.

$ii)$ Montrons que y_t a une représentation minimale unique : D'après (a), y_t a une représentation minimale d'ordre (p, q) et $\forall i = 1, ..., d$, $\Delta^{(i)}(p, q) \neq 0$. Nous voulons montrer que cette écriture minimale est unique.

Considérons alors une écriture minimale d'un processus PARMA

$$y_t - \sum_{j=1}^{p} \phi_{t,j} y_{t-j} = \varepsilon_t - \sum_{j=1}^{q} \theta_{t,j} \varepsilon_{t-j}.$$

ce qui permet d'écrire

$$\begin{pmatrix} \rho^{(i)}_{-q-1} \\ \rho^{(i)}_{-q-2} \\ \vdots \\ \rho^{(i)}_{-q-p} \end{pmatrix} = \Delta^{(i)}(p, q) \begin{pmatrix} \phi_{i,1} \\ \phi_{i,2} \\ \vdots \\ \phi_{i,p} \end{pmatrix}$$

Comme $\Delta^{(i)}(p, q) \neq 0$, nous avons donc,

$$\begin{pmatrix} \phi_{i,1} \\ \phi_{i,2} \\ \vdots \\ \phi_{i,p} \end{pmatrix} = \left[\Delta^{(i)}(p, q)\right]^{-1} \begin{pmatrix} \rho^{(i)}_{-q-1} \\ \rho^{(i)}_{-q-2} \\ \vdots \\ \rho^{(i)}_{-q-p} \end{pmatrix}$$

c'est-à-dire que la famille $\{\phi_{i,1}, \phi_{i,2}, ..., \phi_{i,p}\}$ de cette écriture minimale est unique. De plus, nous savons que les $\theta_{i,j}, j = 1, ..., q$ correspondants sont également uniques (voir la proposition (4.1.2)), donc l'écriture précédente est unique.

Corollaire (4.1.2)

Si $\{y_t, t \in \mathbb{Z}\}$ est un processus autoregressif moyenne mobile périodiquement corrélé d-périodique d'odres constants p et q.

Alors, $\Delta^{(i)}(p, q) \neq 0 \quad \forall i = 1, ..., d.$

Démonstration Nous raisonnons par l'absurde et nous utilisons le Corollaire précédent.

Du Corollaire (4.1.1) et (4.1.2) nous pouvons tirer une condition nécessaire et suffisante pour qu'un processus PARMA d'ordres constants (p, q) soit un processus $\text{ARMA}_d(p, q)$ et qu'il admet une représentation minimale unique.

Théorème (4.1.2)
Un processus autoregressif moyenne mobile périodiquement corrélé d-périodique d'ordres constants p et q est un processus $\text{ARMA}_d(p, q)$ si et seulement si le déterminant $\Delta^{(i)}(p, q) \neq 0 \quad \forall i = 1, ..., d$, dont la représentation minimale est unique.

Démonstration
La condition nécessaire est verifiée par le Corollaire (4.1.1).
La condition suffisante est verifiée par le Corollaire (4.1.2).

Proposition (4.1.5)
Soit $\{y_t, t \in \mathbb{Z}\}$ un processus PARMA d'ordre constant p et q et tel que :
1) $\Delta^{(i)}(p, q) = 0 \quad \forall i = 1, ..., d,$
2) $\Delta^{(i)}(p - 1, q) \neq 0 \quad \forall i = 1, ..., d$
Alors,
a) La représentation de y_t n'est pas minimale pour tout t,
b) $\Delta^{(i+1)}(p - 1, q + m) \neq 0, \quad m \geq 1,$
c) y_t a une représentation minimale d'ordre $p'_i = p - 1$ et $q'_i < q$ quelque soit $t \in \mathbb{Z}$,

$$y_t - \sum_{j=1}^{p-1} a_{t,j} y_{t-j} = \varepsilon_t - \sum_{j=1}^{q'} \theta'_{t,j} \varepsilon_{t-j}.$$

avec $a_{t,p-1} \neq 0$ et les coefficients sont déterminés d'une façon unique.

Démonstration
a) Déjà établi, d'après le théorème (4.1.2).
b) $1^{\underline{er}}$ cas Si $p > 1$
Comme $\Delta^{(i)}(p - 1, q) \neq 0, \quad \forall i = 1, ..., d$, ceci implique que $\Sigma^{(i)}(p - 1, q)$ est une matrice régulière. Nous pouvons alors, définir $p - 1$ coefficients comme suit :

$$\begin{pmatrix} a_{i,1} \\ a_{i,2} \\ \vdots \\ a_{i,p-1} \end{pmatrix} = \left[\Sigma^{(i)}(p - 1, q)\right]^{-1} \begin{pmatrix} \rho^{(i)}_{-q-1} \\ \rho^{(i)}_{-q-2} \\ \vdots \\ \rho^{(i)}_{-q-p+1} \end{pmatrix}$$

C'est-à-dire

$$\begin{pmatrix} \rho^{(i)}_{-q-1} \\ \rho^{(i)}_{-q-2} \\ \vdots \\ \rho^{(i)}_{-q-p+1} \end{pmatrix} = \Sigma^{(i)}(p-1,q) \begin{pmatrix} a_{i,1} \\ a_{i,2} \\ \vdots \\ a_{i,p-1} \end{pmatrix}$$

qui implique

$$\rho^{(i)}_{-h} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-h+j} = 0, \qquad q+1 \leq h \leq q+p-1$$

Montrons que $a_{i,p-1} \neq 0, \quad \forall i = 1, ..., d.$
Nous avons

$$\Delta^{(i+1)}(p,q) = \begin{vmatrix} \rho^{(i)}_{-q} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+1)}_{-q+p-1} \\ \rho^{(i)}_{-q-1} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+1)}_{-q+p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i)}_{-q-p+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+1)}_{-q} \end{vmatrix}$$

$$\Delta^{(i+1)}(p,q) = \begin{vmatrix} \rho^{(i)}_{-q} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-q+j} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+1)}_{-q+p-1} \\ 0 & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+1)}_{-q+p-2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+1)}_{-q} \end{vmatrix}$$

$$= \left[\rho^{(i)}_{-q} - \textstyle\sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-q+j}\right] \Delta^{(i)}(p-1,q)$$

Donc

$$\Delta^{(i+1)}(p,q) = \left[\rho^{(i)}_{-q} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-q+j}\right] \Delta^{(i)}(p-1,q) = 0$$

ceci implique que

$$\rho^{(i)}_{-q} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-q+j} = 0$$

et d'après la proposition (4.1.4) nous obtenons

$$\Delta^{(i+1)}(p-1,q) = (-1)^{p-2} a_{i,p-1}\Delta^{(i)}(p-1,q-1) \neq 0$$

D'où

$$a_{i,p-1} \neq 0$$

Pour montrer que $\Delta^{(i+1)}(p-1,q+m)\neq 0$, pour $m\geq 1$, nous calculons simultanément $\Delta^{(i+1)}(p-1,q+m)$ et $\Delta^{(i+1)}(p,q+m)$ et utilisons le raisonnement par récurrence.

Nous voulons montrer que la relation

$$\rho_{-h}^{(i)}-\sum_{j=1}^{p-1}a_{i,j}\rho_{-h+j}^{(i-j)}=0$$

est vraie pour tout $m\geq q$. Nous avons déjà montrer qu'elle est vraie pour $q+1\leq m\leq q+p-1$, montrons qu'elle l'est encore pour $m=p+q$.

Nous avons

$$\Delta^{(i+1)}(p-1,q+1)=(-1)^{p-2}a_{i,p-1}\Delta^{(i)}(p-1,q)\neq 0.$$

comme $\Delta^{(i)}(p,q)=0,\quad \forall i=1,...,d$, alors d'après la proposition (4.3.1)

$$\begin{aligned}\Delta^{(i+1)}(p,q+1) &= 0\\ &= \left[\rho_{-q-p}^{(i)}-\sum_{j=1}^{p-1}a_{i,j}\rho_{-q-p+j}^{(i-j)}\right]\Delta^{(i)}(p-1,q)\end{aligned}$$

d'où

$$\left[\rho_{-q-p}^{(i)}-\sum_{j=1}^{p-1}a_{i,j}\rho_{-q-p+j}^{(i-j)}\right]=0$$

Pour montrer que la relation

$$\rho_{-h}^{(i)}-\sum_{j=1}^{p-1}a_{i,j}\rho_{-h+j}^{(i-j)}=0$$

est encore vraie pour $m=q+p+1$, nous calculons

$$\Delta^{(i+1)}(p,q+2)=\left[\rho_{-q-p-1}^{(i)}-\sum_{j=1}^{p-1}a_{i,j}\rho_{-q-p-1+j}^{(i-j)}\right]\Delta^{(i)}(p-1,q+1)$$

donc

$$\rho_{-q-p-1}^{(i)}-\sum_{j=1}^{p-1}a_{i,j}\rho_{-q-p-1+j}^{(i-j)}=0$$

Par récurrence alors :

$$\Delta^{(i+1)}(p,q+m)=\left[\rho_{-q-m-p+1}^{(i)}-\sum_{j=1}^{p-1}a_{i,j}\rho_{-q-m-p+1+j}^{(i-j)}\right]\Delta^{(i)}(p-1,q+m-1)=0$$

Pour avoir
$$\rho^{(i)}_{-q-m-p+1} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-q-m-p+1+j} = 0$$
il suffit de montrer que $\Delta^{(i)}(p-1, q+m-1) \neq 0$. Pour celà nous appliquons aussi un raisonnement par récurrence :
$$\Delta^{(i+1)}(p-1, q+m-1) = (-1)^{p-2} a_{i,p-1}\Delta^{(i)}(p-1, q+m-1) \neq 0.$$
Il en résulte alors que
$$\rho^{(i)}_{-h} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-h+j} = 0, \qquad \forall h \geq q$$
et y_t est un processus PARMA d'ordre $p'_i = p-1$ et $q'_i < q$.

Les coefficients $a_{i,1}, a_{i,2}, ..., a_{i,p-1}$ sont déterminés d'une façon unique et nous avons
$$y_t - \sum_{j=1}^{p-1} a_{t,j} y_{t-j} = \varepsilon_t - \sum_{j=1}^{q'} \theta'_{t,j}\varepsilon_{t-j}, \qquad q' < q.$$

$2^{\underline{eme}}$ cas $p = 1$. Le théorème devient comme suit :

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus PARMA d'ordre constant $p = 1$ et q et tel que $\Delta^{(i)}(1, q) = 0 \quad \forall i = 1, ..., d$. Dans ce cas $\Delta^{(i+1)}(1, q+m) = \rho^{(i)}_{-q-m} = 0, \quad m \geq 0$.

Si $\Delta^{(i+1)}(1, q-1) = \rho^{(i)}_{-q+1} \neq 0$, alors y_t est un PMA d'ordre $q-1$.

Généralement, si nous notons par q' le plus grand entier vérifiant $q' < q$ et $\Delta^{(i+1)}(1, q') \neq 0$, alors y_t est un PMA(q').

Proposition (4.1.6)

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus PARMA d'ordres constants p, q et tel que :

1) $\Delta^{(i)}(p, q) = 0 \quad \forall i = 1, ..., d$,

2) $\Delta^{(i)}(p, q-1) \neq 0 \quad \forall i = 1, ..., d$

Alors,

a) La représentation de y_t n'est pas minimale pour tout t,

b) $\Delta^{(i+1)}(p+n, q-1) \neq 0, \quad n \geq 1$,

c) y_t a une représentation minimale d'ordre $q'_i = q-1$ et $p'_i < p$ quelque soit $i = 1, ..., d$.

Démonstration La démonsration de cette proposition est analogue a celle de la proposition précédente.

Proposition (4.1.7)

Si $\{y_t, t \in \mathbb{Z}\}$ est un processus PARMA d'ordres constants p et q, les deux propositions suivantes sont incompatibles

1) $\exists i_0 \in \{1, ..., d\} \quad \Delta^{(i_0+1)}(p+1, q) \neq 0.$
2) $\forall i \in \{1, ..., d\} \quad \Delta^{(i)}(p, q) = 0.$

Démonstration

Supposons que $\forall i \in \{1, ..., d\} \quad \Delta^{(i)}(p, q) = 0.$

Nous savons (voir proposition (4.1.3)) que

$$\Delta^{(i+1)}(p+n, q) = -\theta_{i,q_i} \frac{\sigma^2_{i-q}}{\gamma^{(i)}_0} \Delta^{(i)}(p+n-1, q) \qquad n \geq 1$$

Soit i_0 un entier positif tel que $i_0 \in \{1, ..., d\}$, alors,

$$\Delta^{(i_0+1)}(p+1, q) = -\theta_{i_0,q} \frac{\sigma^2_{i_0-q}}{\gamma^{(i_0)}_0} \Delta^{(i_0)}(p, q) = 0$$

ce qui contre dit l'hypothèse.

Proposition (4.1.8)

Si $\{y_t, t \in \mathbb{Z}\}$ est un processus PARMA d'ordre constant p et q, les deux propositions suivantes sont incompatibles

1) $\exists i_0 \in \{1, ..., d\} \quad \Delta^{(i_0+1)}(p, q+1) \neq 0.$
2) $\forall i \in \{1, ..., d\} \quad \Delta^{(i)}(p, q) = 0.$

Démonstration

La démonstration de cette proposition est analogue à celle de la proposition précédente.

Proposition (4.1.9)

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus PARMA d'ordres constants p et q et solution de l'équation aux différences stochastique (4.1.9), $\phi_{i,p}\theta_{i,q} \neq 0$, tel que nous avons : quelque soit $i = 1, ..., d$.

1) $\Delta^{(i)}(p, q) = 0,$
2) $\Delta^{(i)}(p-1, q) \neq 0,$
3) $\Delta^{(i)}(p, q-1) = 0.$

Alors, $\Delta^{(i)}(p+n, q-1) = 0, \quad \forall n \geq 0, \forall i = 1, ..., d.$

Démonstration

Les deux hypothèses (1) et (2) montrent d'après la proposition (4.1.5) que $\Delta^{(i)}(p-1, q+m) \neq 0 \quad m \geq 0.$

Comme $\Delta^{(i)}(p-1,q) \neq 0, \quad \forall i = 1,...,d$, ceci implique que $\Sigma^{(i)}(p-1,q)$ est une matrice régulière. Nous pouvons alors, définir $p-1$ coefficients comme suit :

$$\begin{pmatrix} a_{i,1} \\ a_{i,2} \\ \vdots \\ a_{i,p-1} \end{pmatrix} = \left[\Sigma^{(i)}(p-1,q)\right]^{-1} \begin{pmatrix} \rho^{(i)}_{-q-1} \\ \rho^{(i)}_{-q-2} \\ \vdots \\ \rho^{(i)}_{-q-p+1} \end{pmatrix}$$

C'est-à-dire

$$\begin{pmatrix} \rho^{(i)}_{-q-1} \\ \rho^{(i)}_{-q-2} \\ \vdots \\ \rho^{(i)}_{-q-p+1} \end{pmatrix} = \Sigma^{(i)}(p-1,q) \begin{pmatrix} a_{i,1} \\ a_{i,2} \\ \vdots \\ a_{i,p-1} \end{pmatrix}$$

qui implique

$$\rho^{(i)}_{-h} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-h+j} = 0, \qquad q+1 \leq h \leq q+p-1$$

Nous avons aussi

$$\Delta^{(i+1)}(p,q) = \begin{vmatrix} \rho^{(i)}_{-q} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+1)}_{-q+p-1} \\ \rho^{(i)}_{-q-1} & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+1)}_{-q+p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i)}_{-q-p+1} & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+1)}_{-q} \end{vmatrix}$$

$$= \begin{vmatrix} \rho^{(i)}_{-q} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-q+j} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p+1)}_{-q+p-1} \\ 0 & \rho^{(i-1)}_{-q} & \cdots & \rho^{(i-p+1)}_{-q+p-2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \rho^{(i-1)}_{-q-p+2} & \cdots & \rho^{(i-p+1)}_{-q} \end{vmatrix}$$

$$= \left[\rho^{(i)}_{-q} - \textstyle\sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-q+j}\right] \Delta^{(i)}(p-1,q) = 0$$

d'où

$$\rho^{(i)}_{-q} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-q+j} = 0$$

Puisque nous avons $\forall i = 1,...,d \quad \Delta^{(i)}(p-1,q) \neq 0$, alors d'aprés la proposition (4.1.8) $\Delta^{(i)}(p-1,q-1) \neq 0$. Nous avons aussi $\Delta^{(i+1)}(p,q-1) = 0$, donc,

$$\rho^{(i)}_{-q+1} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-q+1+j} = 0$$

Il en résulte que

$$\Delta^{(i+1)}(p+n,q-1) = \begin{vmatrix} \rho^{(i)}_{-q+1} & \rho^{(i-1)}_{-q+2} & \cdots & \rho^{(i-p-n+1)}_{-q+p+n} \\ \rho^{(i)}_{-q} & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p-n+1)}_{-q+p+n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{(i)}_{-q-p-n+2} & \rho^{(i-1)}_{-q-p-n+3} & \cdots & \rho^{(i-p-n+1)}_{-q+1} \end{vmatrix}$$

$$= \begin{vmatrix} 0 & \rho^{(i-1)}_{-q+2} & \cdots & \rho^{(i-p-n+1)}_{-q+p+n} \\ 0 & \rho^{(i-1)}_{-q+1} & \cdots & \rho^{(i-p-n+1)}_{-q+p+n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \rho^{(i-1)}_{-q-p-n+3} & \cdots & \rho^{(i-p-n+1)}_{-q+1} \end{vmatrix} = 0, \qquad \forall n \geq 0.$$

Le problème qui se pose maintenant est le suivant : que peut-on dire de $\Delta^{(i)}(p-1,q-2)$ sachant les hypothèse de la proposition précédente ?

Deux cas peuvent se présenter :

$1 \overset{er}{=}$ cas :

$$\rho^{(i)}_{-q+2} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-q+2+j} \neq 0$$

Alors, $\{y_t, t \in \mathbb{Z}\}$ est un modèle $\text{ARMA}_d(p-1,q-2)$ puisque nous savons que

$$\rho^{(i)}_{-h} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-h+j} = 0 \quad h \geq q-1$$

$2 \overset{eme}{=}$ cas :

$$\rho^{(i)}_{-q+2} - \sum_{j=1}^{p-1} a_{i,j}\rho^{(i-j)}_{-q+2+j} = 0$$

Alors, $\{y_t, t \in \mathbb{Z}\}$ est un modèle $\text{ARMA}_d(p-1,q')$ avec $q' < q-2$.

Proposition (4.1.10) Soit $\{y_t, t \in \mathbb{Z}\}$ est un processus PARMA d'ordre constant p et q, $\phi_{t,p}\theta_{t,q} \neq 0$ et tel que nous avons quelque soit $i = 1, ..., d$:

1) $\Delta^{(i)}(p,q) = 0$,
2) $\Delta^{(i)}(p,q-1) \neq 0$,
3) $\Delta^{(i)}(p-1,q) = 0$.

Alors, $\Delta^{(i)}(p-1,q+m) = 0 \quad m \geq 0$.

Démonstration La démonstration de cette proposition est analogue à celle de la proposition précédente.

Remarque Avec les mêmes hypothèses de la proposition précédente, si $\Delta^{(i)}(p-1,q) \neq 0$, alors, $\Delta^{(i)}(p-1,q+m) \neq 0$ et $\{y_t, t \in \mathbb{Z}\}$ est un modèle PARMA d'ordre minimal $(p-1, q-1)$.

4.1.4 Théorème du coin pour les modèles PARMA

L'énoncé du théorème du coin pour les modèles PARMA est le suivant

Théorème du coin
Soit $\{y_t, t \in \mathbb{Z}\}$ un processus autoregressif moyenne mobile périodiquement corrélé d-périodique d'ordres constants

Si y_t vérifi les conditions suivantes quelque soit $i = 1, ..., d$

1) $\Delta^{(i)}(p+n, q+m) = 0 \quad n, m \geq 1$,

2) $\Delta^{(i)}(p+n, q) \neq 0$ et $\Delta^{(i)}(p, q+m) \neq 0 \quad n, m \geq 0$.

Alors,

i) y_t est un modèle $\text{ARMA}_d(p, q)$,

ii) y_t a une écriture minimale unique.

Démonstration Les hypothèses (1) et (2) montrent d'après les propositions (4.1.5) et (4.1.6) que $\{y_t, t \in \mathbb{Z}\}$ est un modèle PARMA(p, q).

Comme $\Delta^{(i)}(p, q) \neq 0$, alors, d'après le Corollaire (4.1.1) $\{y_t, t \in \mathbb{Z}\}$ a une représentation minimale et une écriture minimale unique.

4.2 Identification d'un modèle PARMA par les critères AIC, BIC et HQC.

Cette section sera consacrée à l'identification des modèles autoregressifs moyennes mobiles d-périodique par les critères AIC, BIC et HQC. Avant de présenter ces critères dans le cas périodique, nous avons besoin de caclculer $\widehat{\beta}$ et $\widehat{\sigma}_i^2$, $i = 1, ..., d$, les valeurs qui maximisent $\mathcal{L}(\beta, \sigma^2)$.

Supposons que la série $y_1, ..., y_{dN}$ est obtenue à partir d'un modèle causal et inversible autoregressif moyenne mobile d-périodique d'ordre constant (p, q) suivant :

$$y_t = \sum_{j=1}^{p} \phi_{t,j} y_{t-j} + \varepsilon_t - \sum_{j=1}^{q} \theta_{t,j} \varepsilon_{t-j}, \quad t \in \mathbb{Z} \tag{4.2.1}$$

ou encore

$$y_{i+d\tau} = \sum_{j=1}^{p} \phi_{i,j} y_{i+d\tau-j} + \varepsilon_{i+d\tau} - \sum_{j=1}^{q} \theta_{i,j} \varepsilon_{i+d\tau-j}, \quad i = 1, ..., d \quad \tau \in \mathbb{Z} \tag{4.2.2}$$

où $\{y_t, t \in \mathbb{Z}\}$ est un bruit blanc périodique de loi $\mathcal{N}(0, \sigma_t^2)$. Soit $\beta = \left(\phi_1^{'}, \theta_1^{'}, \phi_2^{'}, \theta_2^{'}, ..., \phi_d^{'}, \theta_2^{'}\right)^{'}$ le vecteur de dimension $d(p+q) \times 1$ des paramètres, où $\phi_i = (\phi_{i,1}, \phi_{i,2}, ..., \phi_{i,p})^{'}, \theta_i = (\theta_{i,1}, \theta_{i,2}, ..., \theta_{i,q})^{'}$, $i = 1, 2, ..., d$. Notons par $\sigma^2 = (\sigma_1^2, \sigma_2^2, ..., \sigma_d^2)$.

Soit $\mathcal{L}(\beta, \sigma^2)$ la fonction de vraisemblance du modèle (4.2.2), qui est donnée par :

$$\begin{aligned}\mathcal{L}(\beta, \sigma^2) &= \left(\tfrac{1}{\sqrt{2\pi}}\right)^{dN} \left(\prod_{t=1}^{dN} V_t\right)^{-1/2} \exp\left(-\tfrac{1}{2}\sum_{t=1}^{dN} \tfrac{\left(y_t - \hat{y}_t\right)^2}{V_t}\right) \\ &= \frac{\left(\prod_{\tau=0}^{N-1}\prod_{i=1}^{d} V_{i+d\tau}\right)^{-1/2}}{(2\pi)^{dN/2}} \exp\left(-\tfrac{1}{2}\sum_{\tau=0}^{N-1}\sum_{i=1}^{d} \tfrac{\left(y_{i+d\tau} - \hat{y}_{i+d\tau}\right)^2}{V_{i+d\tau}}\right)\end{aligned} \tag{4.2.3}$$

où $\widehat{y}_t = E\left(y_t / y_1, ..., y_{t-1}\right)$ est le meilleur prédicteur linéaire de y_t sachant $y_1, ..., y_{t-1}$ et $V_t = E\left[\left(y_t^2 - \widehat{y}_t^2\right)^2 \big/ y_1, ..., y_{t-1}\right]$ (pour plus de détail voir Brockwell et Davis (1988) p. 254-256). L'estimateur du maximum de vraisemblance $\left(\widehat{\beta}, \widehat{\sigma}^2\right)$ de (β, σ^2) sera choisit de façon à maximiser (4.2.3).

Puisque $\{y_t, t \in \mathbb{Z}\}$ est un processus gaussien le théorème de projection prouve que la variance conditionelle est égale à la variance non conditionelle $V_t = E\left[\left(y_t^2 - \widehat{y}_t^2\right)^2\right]$. Lund et Basawa (2001) ont montré comment calculer $\widehat{y}_t$ et V_t rapidement avec un algorithme récursif par rapport à t. Comme dans le cas stationnaire, $\widehat{y}_t$ est en fonction de $\widehat{\beta}$ seulement mais V_t est en fonction de $\widehat{\beta}$ et $\widehat{\sigma}^2$.

D'autre part, Lund et Basawa (2001) ont prouvé que $V_{i+d\tau}\left(\widehat{\beta}, \widehat{\sigma}^2\right) \longrightarrow \sigma_i^2$ lorsque $\tau \longrightarrow \infty$ pour tout $i = 1, ..., d$. Anderson et al. (1999) ont montré que

$$E\left[\left(y_{i+d\tau} - \widehat{y}_{i+d\tau}(\widehat{\beta}) - \varepsilon_{i+d\tau}\left(\widehat{\beta}\right)\right)^2\right] \longrightarrow 0$$

lorsque $\tau \longrightarrow \infty$ pour tout $i = 1, ..., d$. Puisque y_t est inversible donc, $\varepsilon_{i+d\tau}\left(\widehat{\beta}\right)$ est une combinaison linéaire de $y_{i+d\tau}, y_{i+d\tau-1}, ...$ Cependant, $\varepsilon_{i+d\tau}\left(\widehat{\beta}\right)$ peut être rapproché par

$$\varepsilon_{i+d\tau}^{*}\left(\widehat{\beta}\right) = \sum_{j=0}^{i+d\tau-1} \pi_{i,j} y_{i+d\tau-j}.$$

En remplaçant les approximations σ_i^2 pour $V_{i+d\tau}\left(\widehat{\beta}, \widehat{\sigma}^2\right)$, $\varepsilon_{i+d\tau}\left(\widehat{\beta}\right)$ pour $y_{i+d\tau} - \widehat{y}_{i+d\tau}(\widehat{\beta})$ et $\varepsilon_{i+d\tau}^{*}\left(\widehat{\beta}\right)$ pour $\varepsilon_{i+d\tau}\left(\widehat{\beta}\right)$ dans (4.2.3), nous trouvons une fonction de vraisemblance aproximative notée $\mathcal{L}_*(\beta, \sigma^2)$

$$\mathcal{L}_*(\beta, \sigma^2) = \left(\frac{1}{\sqrt{2\pi}}\right)^{dN} \left(\prod_{i=1}^{d} \sigma_i^2\right)^{-N/2} \exp\left(-\frac{1}{2}\sum_{\tau=0}^{N-1}\sum_{i=1}^{d} \frac{\varepsilon_{i+d\tau}^{*2}(\beta)}{\sigma_i^2}\right) \tag{4.2.4}$$

d'où

$$\log\left(\mathcal{L}_*(\beta,\sigma^2)\right) = \sum_{i=1}^{d}\left\{-\frac{1}{2}\sum_{\tau=0}^{N-1}\frac{\varepsilon_{i+d\tau}^{*2}(\beta)}{\sigma_i^2} - \frac{N}{2}\log(2\pi) - \frac{N}{2}\log\left(\sigma_i^2\right)\right\} \quad (4.2.5)$$

Si nous dérivons (4.2.5) par rapport à β et σ_i^2, nous retrouvons les équations suivantes :

$$\sum_{i=1}^{d}\sum_{\tau=0}^{N-1}\sigma_i^{-2}\varepsilon_{i+d\tau}^{*2}(\beta)\left(\frac{\partial\varepsilon_{i+d\tau}^{*2}(\beta)}{\partial\beta}\right) = 0 \quad (4.2.6)$$

et

$$\sigma_i^2 - \frac{1}{N}\sum_{\tau=0}^{N-1}\varepsilon_{i+d\tau}^{*2}(\beta) = 0, \quad i = 1,...,d. \quad (4.2.7)$$

Soit $\widehat{\beta}$ et $\widehat{\sigma}_i^2$, $i = 1,...,d$, les valeurs qui maximisent $\mathcal{L}(\beta,\sigma^2)$ et soit $\widehat{\beta}^*$ et $\widehat{\sigma}_i^{2*}$, $i = 1,...,d$, les valeurs qui maximisent $\mathcal{L}_*(\beta,\sigma^2)$. Les équations (4.2.6) et (4.2.7) donnent :

$$\sum_{i=1}^{d}\sum_{\tau=0}^{N-1}\frac{\varepsilon_{i+d\tau}^{*}\left(\widehat{\beta}^*\right)\left(\left.\frac{\partial\varepsilon_{i+d\tau}^{*}(\beta)}{\partial\beta}\right|_{\beta=\widehat{\beta}^*}\right)}{N^{-1}\sum_{\tau=0}^{N-1}\varepsilon_{i+d\tau}^{*2}\left(\widehat{\beta}^*\right)} = 0 \quad (4.2.8)$$

Le calcul de $\widehat{\sigma}_i^{2*}$ est très facile à partir de (4.2.7) lorsque le vecteur $\widehat{\beta}^*$ est connu

$$\widehat{\sigma}_i^{2*} - \frac{1}{N}\sum_{\tau=0}^{N-1}\varepsilon_{i+d\tau}^{*2}\left(\widehat{\beta}^*\right) = 0, \quad i = 1,...,d. \quad (4.2.9)$$

D'où

$$\begin{aligned}
-2\log\left(\mathcal{L}_*(\widehat{\beta}^*,\widehat{\sigma}^{2*})\right) &= -2\log(ML) \\
&= N\textstyle\sum_{i=1}^{d} 1 + \log(2\pi) + \log(\widehat{\sigma}_i^{2*}) \\
&= N\textstyle\sum_{i=1}^{d}\log(\widehat{\sigma}_i^{2*}) + \text{constante}.
\end{aligned}$$

Par définition nous avons :

$$AIC = -2\log(ML) + 2(\text{nombre de paramètres}).$$

donc, pour un modèle ARMA d-périodique

$$AIC(p,q) = N\sum_{i=1}^{d}\log\left(\widehat{\sigma}_i^{2*}(p,q)\right) + 2(d(p+q)),$$

où $\widehat{\sigma}_i^{2*}(p,q)$ est l'estimateur du maximum de vraisemblance de la variance du bruit blanc relatif à la i^{eme} période obtenu en adaptant le modèle PARMA(p,q) aux observations.

Donc,

$$(\widehat{p},\widehat{q}) = MAICE = \arg\left\{\min_{p,q}\left[\sum_{i=1}^{d} AIC_i(p,q)\right]\right\}, \tag{4.2.10}$$

où

$$AIC_i(p,q) = N\log\left(\widehat{\sigma}_i^{2*}(p,q)\right) + 2(p+q).$$

Nous savons également que

$$BIC = -2\log(ML) + (\text{nombre de paramètres})\log(N).$$

alors, pour un modèle ARMA d-périodique

$$BIC(p,q) = N\sum_{i=1}^{d}\log\left(\widehat{\sigma}_i^{2*}(p,q)\right) + (d(p+q))\log(N),$$

où $\widehat{\sigma}_i^{2*}(p,q)$ est l'estimateur du maximum de vraisemblance de la variance du bruit blanc relatif à la i^{eme} période.

D'où

$$(\widehat{p},\widehat{q}) = MBICE = \arg\left\{\min_{p,q}\left[\sum_{i=1}^{d} BIC_i(p,q)\right]\right\}, \tag{4.2.11}$$

où

$$BIC_i(p,q) = N\log\left(\widehat{\sigma}_i^{2*}(p,q)\right) + (p+q)\log(N).$$

D'une façon analogue à celle du AIC et BIC, le critère HQC est défini pour un modèle ARMA d-périodique comme suit :

$$HQC(p,q) = N\sum_{i=1}^{d}\log\left(\widehat{\sigma}_i^{2*}(p,q)\right) + (d(p+q))c\log\left(\log(N)\right), \quad c>1,$$

où $\widehat{\sigma}_i^{2*}(p,q)$ est l'estimateur du maximum de vraisemblance de la variance du bruit blanc relatif à la i^{eme} période.

Enfin, nous avons

$$(\widehat{p},\widehat{q}) = MHQCE = \arg\left\{\min_{p,q}\left[\sum_{i=1}^{d} HQC_i(p,q)\right]\right\}, \tag{4.2.12}$$

où

$$HQC_i(p,q) = N\log\left(\widehat{\sigma}_i^{2*}(p,q)\right) + (p+q)c\log\left(\log(N)\right), \quad c>1.$$

4.3 Identification d'un modèle PAR par le critère PDC

Hemis (1999) et Bentarzi (2000) ont étudié le problème de la sélection de l'ordre des modèles autorégressifs périodiques. Ils ont généralisé le critère de la densité prédictive bayésienne classique PDC (*Prediction Density Criterion*), qui a été dériver par Djuric et Kay (1992), pour choisir l'ordre d'un modèle autorégressif stationnaire.

Soit $\{y_t, t \in \mathbb{Z}\}$ un processus autoregressif d'ordre p_t à coefficients périodiques de période d et de la représentation suivante :

$$y_t - \sum_{j=1}^{k_t} \phi_{t,j} y_{t-j} = \varepsilon_t, \qquad t \in \mathbb{Z} \tag{4.3.1}$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un bruit blanc périodique de moyenne nulle et de variance σ_t^2. Les paramètres $\phi_{t,j},\ j = 1, ..., k_t,\ k_t$ et σ_t^2 sont des fonctions d-périodiques.

Il faut noter que lorsque l'ordre du processus autoregressif est périodique dans le temps, il suffit de considérer sans perdre de généralité, le processus comme étant un processus autoregressif d-périodique d'ordre k constant, tel que $k = \max_{j=1,...,d} k_j$. Le modèle (4.3.1) peut être écrit sous la forme suivante :

$$y_{i+d\tau} - \sum_{j=1}^{k} \phi_{i,j}^{(k)} y_{i+d\tau-j} = \varepsilon_{i+dt}, \qquad t = i + d\tau,\ i = 1, ..., d \text{ et } \tau \in \mathbb{Z} \tag{4.3.2}$$

Pour calculer la densité prédictive bayésienne, nous avons besoin de quelques notations et hypothèses.

Soient $\Phi^{(k)} = \left(\Phi_1'^{(k)}, \Phi_2'^{(k)}, ..., \Phi_d'^{(k)}\right)'$ le vecteur des paramétres autoregressifs de dimension $kd \times 1$ du modèle (4.3.2), où le vecteur $\Phi_i^{(k)} = \left(\phi_{i,1}^{(k)}, \phi_{i,2}^{(k)}, ..., \phi_{i,k}^{(k)}\right)'$, $i = 1, 2, ..., d$, et le paramètre inconnu $\theta^{(k)} = \left(\Phi'^{(k)}, \alpha'^{(k)}\right)'$, avec $\alpha^{(k)} = (\sigma_1^{(k)}, \sigma_2^{(k)}, ..., \sigma_d^{(k)})'$.

En concernant la distribution du processus d'innovation et les paramètres du modèle autorégressif périodique (4.3.1), nous faisons les hypothèses suivantes sous lesquelles nous dériverons le critère de la densité prédictive bayésienne généralisé.

H1. La distribution du processus innovation ε_t est supposé gaussienne de moyenne nulle et de variance périodique σ_t^2.

H2. Les paramètres autorégressifs $\phi_{i,1}^{(k)}$ satisfont la condition nécessaire et suffisante pour que le processus autorégressif périodique soit causal.

H3. La loi antérieure de $\theta^{(k)}$ sachant k, est choisie toutes les fois que l'approche Bayesienne est adoptée, pour être une densité antérieure non-informative , i.e.

$$f\left(\Phi^{(k)}, \alpha^{(k)} / k\right) \propto \prod_{i=1}^{d} \frac{1}{\sigma_i^{(k)}}, \quad \Phi^{(k)} \in \mathbb{R}^{dk}, \alpha \in \mathbb{R}^{*d} \ et\ k \in \mathbb{N}$$

Sous ces trois hypothèses, nous pouvons énoncé le résultat de Bentarzi (2000) sur la densité prédictive d'un modèle autorégressif périodique.

Proposition (4.1.1) (*Bentarzi (2000)*)
La densité prédictive d'un processus périodiquement corrélé y_t, solution de l'équation aux différences stochastique d-périodique (4.3.1) est donnée pour
$t - k = J_0 + md$, $(1 \leq J_0 \leq d-1)$ par :

1- Pour $k = 0$

Pour $J_0 = 0$ et $m \geq 2$

$$f\left(y_t/\underline{y_{t-1}}; k=0\right) = \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m-1}{2}\right)} \times \frac{\left(\left(Y^{(k)}_{k+d,m-2,d}\right)^t Y^{(k)}_{k+d,m-1,d}\right)^{\frac{m-1}{2}}}{\left(\left(Y^{(k)}_{k+d,m-1,d}\right)^t Y^{(k)}_{k+d,m-1,d}\right)^{\frac{m}{2}}} \tag{4.3.3}$$

Pour $1 \leq J_0 \leq d-1$ et $m \geq 1$

$$f\left(y_t/\underline{y_{t-1}}; k=0\right) = \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)} \times \frac{\left(\left(Y^{(k)}_{k+J_0,m-1,d}\right)^t Y^{(k)}_{k+J_0,m-1,d}\right)^{\frac{m}{2}}}{\left(\left(Y^{(k)}_{k+J_0,m,d}\right)^t Y^{(k)}_{k+J_0,m,d}\right)^{\frac{m+1}{2}}} \tag{4.3.4}$$

2- Pour $k \in \mathbb{N}^*$

Pour $J_0 = 0$ et $m \geq k+2$

$$f\left(y_t/\underline{y_{t-1}}; k>0\right) = \frac{1}{\sqrt{\pi}} \frac{|W_{k+d}(k)(t-2)|}{|W_{k+d}(k)(t-1)|} \frac{\Gamma\left(\frac{m-k}{2}\right)}{\Gamma\left(\frac{m-k-1}{2}\right)} \times \frac{\left(\left(Y^{(k)}_{k+d,m-2,d}\right)^t V^{(k)}_{k+d}(m-2) Y^{(k)}_{k+d,m-2,d}\right)^{\frac{m-k-1}{2}}}{\left(\left(Y^{(k)}_{k+d,m-1,d}\right)^t V^{(k)}_{k+d}(m-1) Y^{(k)}_{k+d,m-1,d}\right)^{\frac{m-k}{2}}} \tag{4.3.5}$$

Pour $1 \leq J_0 \leq d-1$ et $m \geq k+1$

$$f(y_t/\underline{y_{t-1}}; k>0) = \frac{1}{\sqrt{\pi}} \frac{|W_{k+J_0}(k)(t-1)|}{|W_{k+J_0}(k)(t)|} \frac{\Gamma\left(\frac{m-k+1}{2}\right)}{\Gamma\left(\frac{m-k}{2}\right)} \times \frac{\left(\left(Y^{(k)}_{k+J_0,m-1,d}\right)^t V^{(k)}_{k+J_0}(m-1) Y^{(k)}_{k+J_0,m-1,d}\right)^{\frac{m-k-1}{2}}}{\left(\left(Y^{(k)}_{k+J_0,m,d}\right)^t V^{(k)}_{k+J_0}(m-2) Y^{(k)}_{k+J_0,m,d}\right)^{\frac{m-k}{2}}} \tag{4.3.6}$$

où

$$\begin{aligned}
Y_{k+j,m,d}^{(k)} &= \left(y_{k+j}^{(k)}, y_{k+j+d}^{(k)}, ..., y_{k+j+md}^{(k)}\right)' \quad \textit{pour } 1 \leq j \leq d \\
H_{k+j}^{(k)}(m) &= \begin{pmatrix} y_{k+j,0}^{(k)} \\ y_{k+j,1}^{(k)} \\ \vdots \\ y_{k+j,m}^{(k)} \end{pmatrix} = \begin{pmatrix} y_{k+j-1}^{(k)} & y_{k+j-2}^{(k)} & \cdots & y_{j}^{(k)} \\ y_{k+j+d-1}^{(k)} & y_{k+j+d-2}^{(k)} & \cdots & y_{j+d}^{(k)} \\ \vdots & \ddots & \ddots & \vdots \\ y_{k+j+md-1}^{(k)} & y_{k+j+md-2}^{(k)} & \cdots & y_{j+md}^{(k)} \end{pmatrix}
\end{aligned}$$

avec

$$\begin{aligned}
Y_{k+j,\tau}^{(k)} &= \left(y_{k+j-1+d\tau}^{(k)}, y_{k+j-2+d\tau}^{(k)}, ..., y_{j+d\tau}^{(k)}\right)' \quad \textit{pour } \tau = 0, 1, 2, ..., m-1 \\
V_{k+j}^{(k)}(m) &= \left\{ I - H_{k+j}^{(k)}(m) \left(W_{k+j}^{(k)}(m)\right)^{-1} H_{k+j}^{(k)\prime}(m) \right\}, \\
W_{k+j}^{(k)}(m) &= H_{k+j}^{(k)\prime}(m) H_{k+j}^{(k)}(m), \quad \textit{pour } 1 \leq j \leq d \quad \textit{et } m = \left[\frac{t-k}{2}\right].
\end{aligned}$$

et $i = t - d[t/d]$, c'est-à-dire $t = i + d\tau, \quad i = 1, 2, ..., d$ et $t \in \mathbb{Z}$.

Démonstration Voir Bentarzi (2000).

Remarque Nous remarquons que les résultats donnés par cette proposition donnent, dans le cas stationnaire (c'est-à-dire dans le cas $d = 1$), la même densité prédictive $f\left(y_t/\underline{y_{t-1}}; k\right)$ donnée par Djuric et Kay (1992).

Dans le cas des modèles autoregressifs périodiquement corrélés, le critère de la densité prédictive bayésienne établi par Hemis (1999) et Bentarzi (2000), est le même critère PDC de Djuric et Kay (1992). Plus de précision, l'estimateur optimal de l'ordre d'un modèle PAR est donné par

$$\widehat{p} = \arg\left\{ \min_{0 \leq k \leq K} \left(\sum_{t=2}^{N} -\log f\left(y_t/\underline{y_{t-1}}; k\right)\right)\right\}. \tag{4.3.7}$$

où N est la taille de la série chronologique observée qui constitue une réalisation finie du processus stationnaire satisfaisant le modèle autorégressif fondamental d'ordre k et K est un entier positif assez grand.

Pour l'application du critère de la densité prédictive bayésienne, nous devons employer le même nombre d'observations. Pour résoudre ce problème Bentarzi (2000) a utilisé la même technique de Djuric et Kay (1992) dans le cas stationnaire. En effet, en procédant de la même manière, nous pouvons exprimer récursivement les DK_k, pour $k > 2$, comme suit :

Pour $k = 0$

$$DK_k = -\sum_{t=2d}^{N} \log f\left(y_t/\underline{y_{t-1}}; k\right) \tag{4.3.8}$$

Pour $k > 0$

$$DK_k = \sum_{r=0}^{k-1} \sum_{t=2(r+1)d}^{2(r+2)d-1} -\log f\left(y_t/\underline{y_{t-1}}; r\right) + \sum_{t=2(k+1)d}^{N} -\log f\left(y_t/\underline{y_{t-1}}; k\right). \tag{4.3.9}$$

D'où les DK_k sont calculées de même nombre d'observations, $(N - 2d + 1)$. Par conséquent, elles sont proportionnées pour la comparaison. L'estimation de l'ordre est donnée par la valeur de k qui minimise le PDC, c'est-à-dire,

$$\widehat{p} = arg\left\{\min_{0\leq k\leq K} DK_k\right\}. \tag{4.3.10}$$

En se basant sur quelques résultats de simulation Hemis (1999) et Bentarzi (2000) ont conclu que le critère de la densité prédictive (PDC) vérifi une bonne performance dans la sélection de l'ordre des processus autoregressifs périodiques. Ainsi, l'approche utilisée pour dégager ce critère à partir de la densité prédictive bayesienne donne de très bons résultats.

4.4 Identification d'un modèle PAR par le critère CIC

Afin d'élaborer le critère CIC dans le cas des modèles autoregressifs périodiques, Hemis (1999) a utilisé le théorème de Cladyshev (1961), qui consiste à représenter un processus d-périodique ($d \geq 2$) par un certain processus d-varié stationnaire d'ordre p^*. Ainsi, le critère CIC de Ciftcioglu et *al.* (1994) pour des processus autoregressifs d-variés.

La statistique utilisée par Ciftcioglu et *al.* (1994) pour tirer le critère CIC dans le cas d-varié est donnée par :

$$T_d = \frac{\left|\widehat{\Sigma}_k\right| - \left|\widehat{\Sigma}_{k+1}\right|}{\left|\widehat{\Sigma}_k\right|} \frac{(N - d^2 k)}{d^2} \tag{4.4.1}$$

où $\widehat{\Sigma}_k$ est le déterminant de la matrice de covariance du bruit blanc et k a été remplacer par d^2k car : Pour un modèle autoregressif d'ordre k le nombre de coefficients est égal à k, cependant pour un modèle autoregressif d-varié chacun des nouveaux coefficients est une matrice $d \times d$. Danc le nombre de coefficients est égal à d^2k. En plus le terme d^2 dans le dénominateur de l'équation (4.4.1) surgit parce que la différence de χ^2-variables aléatoires a d^2 degré de liberté au lieu 1 pour le cas univarié.

Alors, pour accepter l'hpothèse H_0, il faut que la statistique T_d vérifier la condition suivante :

$$T_d = \frac{\left|\widehat{\Sigma}_k\right| - \left|\widehat{\Sigma}_{k+1}\right|}{\left|\widehat{\Sigma}_k\right|} \frac{(N-d^2k)}{d^2} = \frac{\left|\widehat{\Sigma}_k\right| - \left|\widehat{\Sigma}_{k+1}\right|}{\left|\widehat{\Sigma}_k\right|} \frac{N}{d^2}(1-\frac{d^2k}{N}) \leq l_d$$

Le niveau de signification du test dans le cas d-varié est égal à

$$\alpha = \int_{l_d}^{\infty} dF_{T_d}(t) = \frac{d^2k}{N}, \tag{4.4.2}$$

où F_{T_d} est la loi de la statistique T_d et l_d est le niveau de décision.

Enfin, l'expression du critère CIC pour les processus d-varié autoregressif est donnée comme suit :

$$CIC(k) = N \log \left|\widehat{\Sigma}_k\right| + \frac{N}{(N-d^2k)} \times d^2k \times \widetilde{F}_d \tag{4.4.3}$$

où

$$\widetilde{F}_d(l_d) = 1 + \sqrt{\frac{2}{\pi}\frac{N}{d^2k}}\sqrt{l_d}e^{-\frac{1}{2}l_d}.$$

4.5 Etude de simulation

Pour étudier la performance des différents critères de sélection d'un modèle PARMA d-périodique, exposés précédemment, nous proposons d'appliquer ces méthodes d'identification sur des données artificielles issues d'un modèle PARMA d-périodique bien choisi.

Nous avons simulé 1000 séries de longueur variant de 50 à 200, supposées obtenues à partir de différents modèles PARMA d-périodique. Nous avons fait varier les paramètres de façon à ce que les propriétés de causalité et d'inversibilité soient vérifiées. Nous avons choisi des différentes périodes en tenant compte des réalités pratiques. Pour chaque modèle autoregressif pur, nous avons appliqué les critères AIC, CIC, BIC, HQC et PDC. Alors que pour les modèles ARMA mixte d-périodique nous avons appliqué les critères AIC, BIC et HQC.

Nous nous sommes intéressés à la fréquence de choix d'ordres dans 1000 répétitions pour les différents critères. L'objectif de cette section est de faire une comparaison entre la performance des différents critères de sélection.

Pour l'estimation des paramètres dans le cas des modèles autoregressifs, nous avons utilisé l'algorithme de Boshnakov (1996) et pour les modèles PARMA l'algorithme *Periodic Recursive Maximum Likelihood* (PRML) de Bentarzi et Aknouche (2003). Les sorties de l'estimation contiennent les moyennes et les écart types des valeurs estimés des paramètres.

Estimation des paramètres du modèle $PAR_2(1)$ par l'algorithme de Boshnakov

paramètres / taille	50	100	150	200
$\phi_{1,1} = 0.5$	0.4532 (0.0907)	0.4763 (0.0623)	0.4859 (0.0487)	0.4896 (0.0409)
$\sigma_1^2 = 1$	1.0662 (0.3209)	1.0376 (0.2208)	1.0313 (0.1744)	1.0157 (0.1433)
$\phi_{1,2} = 1.4$	1.3817 (0.1417)	1.3931 (0.0970)	1.3943 (0.0733)	1.3951 (0.0638)
$\sigma_2^2 = 1$	0.9686 (0.2740)	0.9709 (0.1974)	0.9820 (0.1589)	0.9919 (0.1417)

Tableau (4.5.1)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1,\ \ d = 2,\ \ N = 50$ $\phi_{1,1} = 0.5,\ \ \sigma_1^2 = 1,\ \ \phi_{1,2} = 1.4,\ \ \sigma_2^2 = 1$					
p	*AIC*	*BIC*	*HQC*	*CIC*	*PDC*
0	0.0	0.0	0.0	0.0	0.0
1	44.8	82.2	94.5	90.6	95.1
2	22.0	13.6	4.9	8.5	3.8
3	6.7	1.4	0.4	0.5	0.7
4	10.5	2.0	0.2	0.3	0.4
5	3.2	0.5	0.0	0.0	0.0
6	9.2	0.3	0.0	0.1	0.0
7	3.6	0.0	0.0	0.0	0.0

Tableau (4.5.2)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1,\ \ d = 2,\ \ N = 100$ $\phi_{1,1} = 0.5,\ \ \sigma_1^2 = 1,\ \ \phi_{1,2} = 1.4,\ \ \sigma_2^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	45.8	87.3	94.8	90.9	98.2
2	20.7	10.5	4.9	8.2	1.5
3	5.3	1.2	0.2	0.7	0.3
4	11.0	0.8	0.1	0.2	0.0
5	4.3	0.1	0.0	0.0	0.0
6	8.4	0.1	0.0	0.0	0.0
7	4.5	0.0	0.0	0.0	0.0

Tableau (4.5.3)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1,\ \ d = 2,\ \ N = 150$ $\phi_{1,1} = 0.5,\ \ \sigma_1^2 = 1,\ \ \phi_{1,2} = 1.4,\ \ \sigma_2^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	48.2	89.1	96.4	91.4	98.7
2	22.0	9.7	3.6	7.1	1.3
3	5.4	0.6	0.0	0.9	0.0
4	10.8	0.5	0.0	0.5	0.0
5	3.0	0.0	0.0	0.0	0.0
6	6.9	0.1	0.0	0.1	0.0
7	3.7	0.0	0.0	0.0	0.0

Tableau (4.5.4)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1,\ \ d = 2,\ \ N = 200$ $\phi_{1,1} = 0.5,\ \ \sigma_1^2 = 1,\ \ \phi_{1,2} = 1.4,\ \ \sigma_2^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	49.0	92.8	97.3	91.9	99.3
2	21.3	6.8	2.6	7.1	0.7
3	5.4	0.2	0.1	0.6	0.0
4	10.2	0.2	0.0	0.2	0.0
5	4.3	0.0	0.0	0.2	0.0
6	5.6	0.0	0.0	0.0	0.0
7	4.2	0.0	0.0	0.0	0.0

Tableau (4.5.5)

Estimation des paramètres du modèle $PAR_3(1)$ par l'algorithme de Boshnakov

paramètres / taille	60	120	150	180
$\phi_{1,1} = -0.9$	-0.8437 (0.2421)	-0.8815 (0.1546)	-0.8736 (0.1420)	-0.8834 (0.1261)
$\sigma_1^2 = 1$	1.0301 (0.3333)	1.0147 (0.2302)	1.0111 (0.2013)	1.0073 (0.1839)
$\phi_{1,2} = 0.4$	0.4037 (0.1754)	0.3993 (0.1201)	0.4049 (0.1036)	0.4026 (0.0963)
$\sigma_2^2 = 1$	0.9562 (0.3145)	0.9760 (0.2117)	0.9768 (0.1991)	0.9759 (0.1801)
$\phi_{1,3} = 0.2$	0.1979 (0.2057)	0.1957 (0.1452)	0.1932 (0.1272)	0.1971 (0.1152)
$\sigma_3^2 = 1$	0.9528 (0.3052)	0.9926 (0.2308)	0.9765 (0.1984)	0.9933 (0.1788)

Tableau (4.5.6)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1,\ d = 3,\ N = 60$ $\phi_{1,1} = -0.9,\ \sigma_1^2 = 1,\ \phi_{1,2} = 0.4,\ \sigma_2^2 = 1,\ \phi_{1,3} = 0.2,\ \sigma_3^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.6	2.3	1.2	4.6
1	32.4	84.2	95.3	91.4	89.1
2	4.3	4.9	1.9	3.0	4.9
3	13.3	6.4	0.4	2.3	0.9
4	14.1	2.9	0.1	1.8	0.4
5	3.2	0.2	0.0	0.0	0.1
6	15.1	0.4	0.0	0.2	0.0
7	17.6	0.4	0.0	0.1	0.0

Tableau (4.5.7)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1,\ d = 3,\ N = 120$ $\phi_{1,1} = -0.9,\ \sigma_1^2 = 1,\ \phi_{1,2} = 0.4,\ \sigma_2^2 = 1,\ \phi_{1,3} = 0.2,\ \sigma_3^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	33.7	93.4	99.0	93.3	97.5
2	3.4	2.3	0.4	2.6	2.3
3	13.3	2.8	0.4	2.5	0.2
4	11.9	1.1	0.2	1.5	0.0
5	3.9	0.1	0.0	0.0	0.0
6	11.1	0.3	0.0	0.1	0.0
7	19.7	0.0	0.0	0.0	0.0

Tableau (4.5.8)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p=1,\ \ d=3,\ \ N=150$ $\phi_{1,1}=-0.9,\ \ \sigma_1^2=1,\ \ \phi_{1,2}=0.4,\ \ \sigma_2^2=1,\ \ \phi_{1,3}=0.2,\ \ \sigma_3^2=1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	33.8	94.0	98.5	93.3	97.9
2	6.0	1.8	0.8	1.9	1.9
3	12.6	3.3	0.7	3.7	0.2
4	14.2	0.6	0.0	0.9	0.0
5	5.6	0.0	0.0	0.0	0.0
6	11.1	0.2	0.0	0.1	0.0
7	16.7	0.1	0.0	0.1	0.0

Tableau (4.5.9)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p=1,\ \ d=3,\ \ N=180$ $\phi_{1,1}=-0.9,\ \ \sigma_1^2=1,\ \ \phi_{1,2}=0.4,\ \ \sigma_2^2=1,\ \ \phi_{1,3}=0.2,\ \ \sigma_3^2=1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	35.5	94.1	99.0	93.7	98.7
2	4.7	2.3	0.5	1.8	1.3
3	13.3	2.7	0.4	3.5	0.0
4	13.5	0.7	0.1	0.7	0.0
5	2.7	0.1	0.0	0.0	0.0
6	12.4	0.1	0.0	0.2	0.0
7	17.9	0.0	0.0	0.1	0.0

Tableau (4.5.10)

Estimation des paramètres du modèle $\text{PAR}_4(1)$ par l'algorithme de Boshnakov

paramètres / taille	80	120	160	200
$\phi_{1,1} = 1.5$	1.4183 (0.1553)	1.4366 (0.1189)	1.4518 (0.0989)	1.4663 (0.0844)
$\sigma_1^2 = 1$	1.7698 (0.9097)	1.5394 (0.6577)	1.4144 (0.5189)	1.3148 (0.4038)
$\phi_{1,2} = -0.5$	-0.4832 (0.0784)	-0.4913 (0.0607)	-0.4936 (0.0544)	-0.4942 (0.0467)
$\sigma_2^2 = 1$	0.9556 (0.2938)	0.9667 (0.2536)	0.9829 (0.2205)	0.9784 (0.1889)
$\phi_{1,3} = 1.2$	1.1864 (0.1351)	1.1976 (0.1022)	1.1949 (0.0869)	1.1984 (0.0768)
$\sigma_3^2 = 1$	0.9370 (0.3023)	0.9630 (0.2581)	0.9829 (0.2215)	0.9746 (0.1899)
$\phi_{1,4} = 0.7$	0.6889 (0.1025)	0.6888 (0.0779)	0.6938 (0.0680)	0.6934 (0.0600)
$\sigma_4^2 = 1$	0.9481 (0.3215)	0.9607 (0.2576)	0.9742 (0.2178)	0.9731 (0.2032)

Tableau (4.5.11)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1,\ \ d = 4,\ \ N = 80$ $\phi_{1,1} = 1.5,\ \sigma_1^2 = 1,\ \phi_{1,2} = -0.5,\ \sigma_2^2 = 1,\ \phi_{1,3} = 1.2,\ \sigma_3^2 = 1,\ \phi_{1,4} = 0.7,\ \sigma_4^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	17.8	75.2	93.9	82.5	96.5
2	8.7	12.4	5.1	10.5	2.2
3	2.9	1.5	0.1	1.1	1.1
4	12.5	4.4	0.7	2.9	0.0
5	24.1	4.8	0.1	2.5	0.2
6	24.4	1.3	0.1	0.5	0.0
7	9.6	0.4	0.0	0.0	0.0

Tableau (4.5.12)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p=1,\ d=4,\ N=120$ $\phi_{1,1}=1.5,\ \sigma_1^2=1,\ \phi_{1,2}=-0.5,\ \sigma_2^2=1,\ \phi_{1,3}=1.2,\ \sigma_3^2=1,\ \phi_{1,4}=0.7,\ \sigma_4^2=1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	17.6	83.1	96.1	85.1	98.3
2	11.0	9.5	3.4	8.8	1.5
3	1.3	1.0	0.1	1.0	0.2
4	13.0	3.3	0.3	2.5	0.0
5	24.9	2.8	0.1	1.9	0.0
6	23.7	0.2	0.0	0.6	0.0
7	8.5	0.1	0.0	0.1	0.0

Tableau (4.5.13)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p=1,\ d=4,\ N=160$ $\phi_{1,1}=1.5,\ \sigma_1^2=1,\ \phi_{1,2}=-0.5,\ \sigma_2^2=1,\ \phi_{1,3}=1.2,\ \sigma_3^2=1,\ \phi_{1,4}=0.7,\ \sigma_4^2=1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	19.6	88.8	96.6	86.5	98.9
2	10.3	7.2	3.0	7.7	1.1
3	1.9	0.5	0.0	0.6	0.0
4	12.4	1.8	0.3	2.4	0.0
5	22.3	1.2	0.1	2.4	0.0
6	23.6	0.5	0.0	0.4	0.0
7	9.9	0.0	0.0	0.0	0.0

Tableau (4.5.14)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p=1,\ d=4,\ N=200$ $\phi_{1,1}=1.5,\ \sigma_1^2=1,\ \phi_{1,2}=-0.5,\ \sigma_2^2=1,\ \phi_{1,3}=1.2,\ \sigma_3^2=1,\ \phi_{1,4}=0.7,\ \sigma_4^2=1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	19.8	89.2	96.8	86.8	99.3
2	10.4	8.0	2.9	8.8	0.7
3	1.3	0.1	0.0	0.5	0.0
4	13.7	1.4	0.2	1.8	0.0
5	25.2	1.1	0.1	1.6	0.0
6	20.7	0.2	0.0	0.5	0.0
7	8.9	0.0	0.0	0.0	0.0

Tableau (4.5.15)

Estimation des paramètres du modèle $PAR_7(1)$ par l'algorithme de Boshnakov

paramètres / taille	70	140	210	280
$\phi_{1,1} = 1.1$	0.9916 (0.2284)	1.0380 (0.1367)	1.0654 (0.1071)	1.0710 (0.0889)
$\sigma_1^2 = 1$	1.7377 (1.0333)	1.3933 (0.5429)	1.2349 (0.3739)	1.1760 (0.3184)
$\phi_{1,2} = -0.9$	-0.8969 (0.1495)	-0.8949 (0.0989)	-0.8952 (0.0776)	-0.8997 (0.0689)
$\sigma_2^2 = 1$	0.9258 (0.4391)	0.9343 (0.3102)	0.9654 (0.2610)	0.9634 (0.2155)
$\phi_{1,3} = 0.7$	0.6985 (0.1515)	0.6977 (0.1039)	0.6992 (0.0811)	0.7017 (0.0680)
$\sigma_3^2 = 1$	0.9324 (0.4358)	0.9481 (0.3096)	0.9805 (0.2609)	0.9746 (0.2222)
$\phi_{1,4} = 1.2$	1.1985 (0.1801)	1.2007 (0.1213)	1.2016 (0.1005)	1.1965 (0.0853)
$\sigma_4^2 = 1$	0.9145 (0.4336)	0.9225 (0.3060)	0.9588 (0.2522)	0.9764 (0.2243)
$\phi_{1,5} = 0.2$	0.1958 (0.1373)	0.1957 (0.0927)	0.1938 (0.0738)	0.2002 (0.0639)
$\sigma_5^2 = 1$	0.8788 (0.3957)	0.9444 (0.3048)	0.9809 (0.2577)	0.9611 (0.2246)
$\phi_{1,6} = -1.4$	-1.3874 (0.3242)	-1.3932 (0.2171)	-1.4011 (0.1692)	-1.4010 (0.1490)
$\sigma_6^2 = 1$	0.8924 (0.4264)	0.9706 (0.3025)	0.9634 (0.2547)	0.9725 (0.2146)
$\phi_{1,7} = 0.9$	0.8945 (0.1973)	0.8932 (0.1316)	0.8948 (0.1003)	0.9010 (0.0900)
$\sigma_7^2 = 1$	0.8832 (0.4325)	0.9566 (0.3037)	0.9765 (0.2590)	0.9766 (0.2211)

Tableau (4.5.16)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1,\ \ d = 7,\ \ N = 140$ $\phi_{1,1} = 1.1,\ \sigma_1^2 = 1,\ \phi_{1,2} = -0.9,\ \sigma_2^2 = 1,\ \phi_{1,3} = 0.7,\ \sigma_3^2 = 1,\ \phi_{1,4} = 1.2,\ \sigma_4^2 = 1$ $\phi_{1,5} = 0.2,\ \sigma_5^2 = 1,\ \phi_{1,6} = -1.4,\ \sigma_6^2 = 1\phi_{1,7} = 0.9,\ \sigma_7^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	0.3	33.8	72.6	35.2	98.3
2	3.0	28.4	20.7	28.4	1.7
3	8.4	21.9	6.1	21.5	0.0
4	13.3	8.8	0.4	8.4	0.0
5	8.6	3.0	0.2	2.7	0.0
6	1.6	0.5	0.0	0.5	0.0
7	64.8	3.6	0.0	3.3	0.0

Tableau (4.5.17)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 1,\ \ d = 7,\ \ N = 210$ $\phi_{1,1} = 1.1,\ \sigma_1^2 = 1,\ \phi_{1,2} = -0.9,\ \sigma_2^2 = 1,\ \phi_{1,3} = 0.7,\ \sigma_3^2 = 1,\ \phi_{1,4} = 1.2,\ \sigma_4^2 = 1$ $\phi_{1,5} = 0.2,\ \sigma_5^2 = 1,\ \phi_{1,6} = -1.4,\ \sigma_6^2 = 1\phi_{1,7} = 0.9,\ \sigma_7^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	0.7	45.9	78.8	41.3	99.7
2	4.0	26.7	15.0	26.5	0.3
3	7.0	16.8	5.4	18.3	0.0
4	11.8	6.2	0.6	7.6	0.0
5	5.6	2.2	0.1	2.9	0.0
6	0.7	0.2	0.0	0.3	0.0
7	70.2	2.0	0.1	3.1	0.0

Tableau (4.5.18)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p=1,\ d=7,\ N=280$ $\phi_{1,1}=1.1,\ \sigma_1^2=1,\ \phi_{1,2}=-0.9,\ \sigma_2^2=1,\ \phi_{1,3}=0.7,\ \sigma_3^2=1,\ \phi_{1,4}=1.2,\ \sigma_4^2=1$ $\phi_{1,5}=0.2,\ \sigma_5^2=1,\ \phi_{1,6}=-1.4,\ \sigma_6^2=1\phi_{1,7}=0.9,\ \sigma_7^2=1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	0.9	51.2	80.4	97.6	99.9
2	4.6	23.0	14.8	2.4	0.1
3	7.9	16.2	4.0	0.0	0.0
4	11.1	6.6	0.7	0.0	0.0
5	6.7	1.3	0.0	0.0	0.0
6	0.5	0.1	0.0	0.0	0.0
7	68.3	1.6	0.1	0.0	0.0

Tableau (4.5.19)

Estimation des paramètres du modèle $\text{PAR}_2(2)$ par l'algorithme de Boshnakov

paramètres / taille	50	100	150	180
$\phi_{1,1}=0.777$	0.7344 (0.2039)	0.7610 (0.1371)	0.7625 (0.1125)	0.7671 (0.0933)
$\phi_{2,1}=-0.492$	-0.4605 (0.1972)	-0.4847 (0.1329)	-0.4818 (0.1109)	-0.4829 (0.0943)
$\sigma_1^2=1$	0.8270 (0.3103)	0.8770 (0.1898)	0.8973 (0.1522)	0.9068 (0.1182)
$\phi_{1,2}=0.44$	0.4735 (0.1874)	0.4494 (0.1281)	0.4505 (0.1034)	0.4436 (0.0879)
$\phi_{2,2}=0.35$	0.2913 (0.1729)	0.3225 (0.1237)	0.3292 (0.0983)	0.3378 (0.0861)
$\sigma_2^2=1$	0.8933 (0.2498)	0.9319 (0.1736)	0.9440 (0.1352)	0.9461 (0.1190)

Tableau (4.5.20)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 2,\ d = 2,\ N = 50$ $\phi_{1,1} = 0.777,\ \phi_{2,1} = -0.492,\ \sigma_1^2 = 1,\ \phi_{1,2} = 0.44,\ \phi_{2,2} = 0.35,\ \sigma_2^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.1	0.3	1.3	0.6	1.3
1	6.9	15.1	25.2	21.5	35.8
2	55.1	73.3	69.3	72.8	56.9
3	10.0	5.7	2.2	2.3	5.1
4	9.6	2.5	0.5	1.6	0.5
5	6.7	0.9	0.7	0.5	0.2
6	6.9	1.7	0.6	0.1	0.1
7	4.7	0.5	0.2	0.6	0.1

Tableau (4.5.21)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 2,\ d = 2,\ N = 100$ $\phi_{1,1} = 0.777,\ \phi_{2,1} = -0.492,\ \sigma_1^2 = 1,\ \phi_{1,2} = 0.44,\ \phi_{2,2} = 0.35,\ \sigma_2^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	0.4	1.5	3.5	3.3	10.2
2	58.2	93.2	94.8	92.4	86.0
3	9.0	3.3	1.2	2.6	3.2
4	12.3	1.2	0.2	0.9	0.5
5	4.8	0.5	0.1	0.3	0.1
6	9.9	0.0	0.0	0.4	0.0
7	5.4	0.3	0.2	0.1	0.0

Tableau (4.5.22)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 2,\ d = 2,\ N = 150$ $\phi_{1,1} = 0.777,\ \phi_{2,1} = -0.492,\ \sigma_1^2 = 1,\ \phi_{1,2} = 0.44,\ \phi_{2,2} = 0.35,\ \sigma_2^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	0.1	0.2	0.4	0.3	2.2
2	58.5	94.8	98.0	95.4	94.9
3	8.2	2.8	1.3	2.3	2.8
4	13.2	1.5	0.2	1.6	0.1
5	5.0	0.4	0.1	0.2	0.0
6	9.8	0.2	0.0	0.1	0.0
7	5.2	0.1	0.0	0.1	0.0

Tableau (4.5.23)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p=2,\ d=2,\ N=200$ $\phi_{1,1}=0.777,\ \phi_{2,1}=-0.492,\ \sigma_1^2=1,\ \phi_{1,2}=0.44,\ \phi_{2,2}=0.35,\ \sigma_2^2=1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.6
2	59.7	96.4	99.3	96.0	97.3
3	8.7	3.0	0.5	2.9	2.1
4	12.2	0.5	0.2	1.0	0.0
5	4.4	0.0	0.0	0.0	0.0
6	9.4	0.1	0.0	0.1	0.0
7	5.6	0.0	0.0	0.0	0.0

Tableau (4.5.24)

Estimation des paramètres du modèle $\mathrm{PAR}_4(2)$ par l'algorithme de Boshnakov

paramètres / taille	80	120	160	200
$\phi_{1,1}=-0.5$	-0.4786 (0.1533)	-0.4904 (0.1251)	-0.4814 (0.1057)	-0.4900 (0.0975)
$\phi_{2,1}=-0.75$	-0.7216 (0.4207)	-0.7425 (0.3381)	-0.7189 (0.2900)	-0.7343 (0.2578)
$\sigma_1^2=1$	1.2089 (0.3946)	1.1871 (0.3188)	1.1881 (0.2856)	1.1783 (0.2384)
$\phi_{1,2}=1.9$	1.6990 (0.2951)	1.7616 (0.2253)	1.7889 (0.1862)	1.8055 (0.1661)
$\phi_{2,2}=0.84$	0.7511 (0.1145)	0.7816 (0.0890)	0.7935 (0.0721)	0.8005 (0.0644)
$\sigma_2^2=1$	6.5730 (2.2912)	6.6492 (1.8708)	6.7286 (1.767)	6.7070 (1.4457)
$\phi_{1,3}=-0.2$	-0.1934 (0.0896)	-0.1942 (0.0741)	-0.1973 (0.0634)	-0.2018 (0.0566)
$\phi_{2,3}=-0.48$	-0.4886 (0.1652)	-0.4887 (0.1370)	-0.4876 (0.1173)	-0.4831 (0.1059)
$\sigma_3^2=1$	0.9003 (0.3012)	0.9163 (0.2397)	0.9583 (0.2173)	0.9660 (0.1887)
$\phi_{1,4}=-2$	-1.9922 (0.2019)	-1.9960 (0.1595)	-1.9978 (0.1341)	-1.9996 (0.1243)
$\phi_{2,4}=0.48$	0.4817 (0.1027)	0.4862 (0.0804)	0.4811 (0.0719)	0.4797 (0.0639)
$\sigma_4^2=1$	0.9034 (0.3153)	0.9325 (0.2351)	0.9461 (0.2102)	0.9643 (0.2041)

Tableau (4.5.25)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p=2,\quad d=4,\quad N=80$					
$\phi_{1,1}=-0.5,\ \phi_{2,1}=-0.75,\ \sigma_1^2=1,\ \phi_{1,2}=1.9,\ \phi_{2,2}=0.84,\ \sigma_2^2=1$					
$\phi_{1,3}=-0.2,\ \phi_{2,3}=-0.48,\ \sigma_3^2=1,\ \phi_{1,4}=-2,\ \phi_{2,4}=0.48,\ \sigma_4^2=1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.2	1.0	0.1	0.0
2	48.7	89.0	96.7	94.4	94.4
3	3.5	0.9	0.3	0.9	3.8
4	18.1	6.5	1.3	2.6	1.3
5	15.7	2.4	0.5	1.8	0.3
6	8.9	0.7	0.2	0.2	0.1
7	5.1	0.3	0.0	0.0	0.1

Tableau (4.5.26)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p=2,\quad d=4,\quad N=120$					
$\phi_{1,1}=-0.5,\ \phi_{2,1}=-0.75,\ \sigma_1^2=1,\ \phi_{1,2}=1.9,\ \phi_{2,2}=0.84,\ \sigma_2^2=1$					
$\phi_{1,3}=-0.2,\ \phi_{2,3}=-0.48,\ \sigma_3^2=1,\ \phi_{1,4}=-2,\ \phi_{2,4}=0.48,\ \sigma_4^2=1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.1	0.0	0.0
2	48.7	93.8	99.2	95.2	97.0
3	3.3	0.9	0.3	0.5	2.6
4	21.2	3.8	0.3	2.9	0.4
5	16.0	1.4	0.1	1.1	0.0
6	7.0	0.1	0.0	0.3	0.0
7	3.8	0.0	0.0	0.0	0.0

Tableau (4.5.27)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 2, \ d = 4, \ N = 160$ $\phi_{1,1} = -0.5, \ \phi_{2,1} = -0.75, \ \sigma_1^2 = 1, \ \phi_{1,2} = 1.9, \ \phi_{2,2} = 0.84, \ \sigma_2^2 = 1$ $\phi_{1,3} = -0.2, \ \phi_{2,2} = -0.48, \ \sigma_3^2 = 1, \ \phi_{1,4} = -2, \ \phi_{2,2} = 0.48, \ \sigma_4^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0
2	49.6	95.4	99.7	95.2	98.9
3	3.2	0.9	0.1	0.2	0.8
4	23.1	2.3	0.2	3.3	0.3
5	14.1	1.4	0.0	1.3	0.0
6	6.7	0.0	0.0	0.0	0.0
7	3.3	0.0	0.0	0.0	0.0

Tableau (4.5.28)

Fréquences de choix d'ordre dans 1000 répétitions pour les différents critères.

$p = 2, \ d = 4, \ N = 200$ $\phi_{1,1} = -0.5, \ \phi_{2,1} = -0.75, \ \sigma_1^2 = 1, \ \phi_{1,2} = 1.9, \ \phi_{2,2} = 0.84, \ \sigma_2^2 = 1$ $\phi_{1,3} = -0.2, \ \phi_{2,3} = -0.48, \ \sigma_3^2 = 1, \ \phi_{1,4} = -2, \ \phi_{2,4} = 0.48, \ \sigma_4^2 = 1$					
p	AIC	BIC	HQC	CIC	PDC
0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0
2	52.6	97.8	99.5	96.4	99.0
3	2.5	0.2	0.0	0.8	10
4	22.9	1.4	0.3	2.3	0.0
5	15.1	0.6	0.2	0.5	0.0
6	4.5	0.0	0.0	0.0	0.0
7	2.4	0.0	0.0	0.0	0.0

Tableau (4.5.29)

Estimation des paramètres du modèle PARMA$_2$(1,1) par l'algorithme PRML

paramètres / taille	50	100	150	200
$\phi_{1,1} = 0.5$	0.5328 (0.1060)	0.5181 (0.0752)	0.5165 (0.0545)	0.5102 (0.0480)
$\theta_{1,1} = 0.9$	0.7972 (0.2090)	0.8485 (0.1452)	0.8716 (0.1076)	0.8761 (0.0885)
$\sigma_1^2 = 1$	0.9662 (0.3029)	0.9924 (0.2025)	0.9823 (0.1682)	0.9936 (0.1447)
$\phi_{1,2} = 1.4$	1.3767 (0.1749)	1.3794 (0.1085)	1.3856 (0.0931)	1.3867 (0.0797)
$\theta_{1,2} = -0.7$	-0.5638 (0.2393)	-0.6362 (0.1582)	-0.6413 (0.1246)	-0.6647 (0.1017)
$\sigma_2^2 = 1$	0.9883 (0.3045)	0.9967 (0.2116)	0.9939 (0.1633)	0.9906 (0.1470)

Tableau (4.5.30)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p = 1,\ \ q = 1,\ \ d = 2,\ \ N = 50$

$\phi_{1,1} = 0.5,\ \ \theta_{1,1} = 0.9,\ \ \sigma_1^2 = 1,\ \ \phi_{1,2} = 1.4,\ \ \theta_{1,2} = -0.7,\ \ \sigma_2^2 = 1$

	p / q	1	2	3	4	5	6	7
AIC	1	15.8	4.2	1.8	1.3	0.7	0.7	0.5
BIC		49.8	9.4	2.5	0.9	0.1	0.2	0.2
HQC		70.5	9.4	1.6	0.0	0.1	0.0	0.1
AIC	2	5.5	6.5	3.3	1.6	0.9	0.6	1.4
BIC		8.7	5.8	2.4	0.9	0.1	0.4	0.3
HQC		7.7	2.6	0.8	0.2	0.0	0.1	0.0
AIC	3	4.2	1.9	0.5	0.5	1.1	0.8	1.0
BIC		3.9	1.7	0.1	0.1	0.2	0.3	0.1
HQC		2.7	0.9	0.1	0.0	0.1	0.1	0.0
AIC	4	2.9	1.7	0.8	0.9	1.0	1.0	1.2
BIC		2.2	0.6	0.3	0.0	0.1	0.0	0.1
HQC		0.9	0.1	0.1	0.0	0.0	0.0	0.0
AIC	5	3.6	0.8	0.9	0.6	0.8	1.0	1.5
BIC		2.3	0.4	0.1	0.0	0.3	0.2	0.1
HQC		0.9	0.0	0.1	0.0	0.0	0.0	0.1
AIC	6	2.9	1.0	0.9	0.8	0.9	1.0	2.8
BIC		0.7	0.5	0.2	0.0	0.1	0.0	0.4
HQC		0.3	0.2	0.0	0.0	0.0	0.0	0.0
AIC	7	4.6	2.0	1.1	1.4	1.1	2.2	3.8
BIC		1.2	0.7	0.2	0.1	0.4	0.2	0.5
HQC		0.3	0.0	0.0	0.0	0.0	0.0	0.0

Tableau (4.5.31)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p=1,\ q=1,\ d=2,\ N=100$ $\phi_{1,1}=0.5,\ \theta_{1,1}=0.9,\ \sigma_1^2=1$ $\phi_{1,2}=1.4,\ \theta_{1,2}=-0.7,\ \sigma_2^2=1$								
	p / q	1	2	3	4	5	6	7
AIC	1	25.0	6.9	2.3	0.9	0.6	0.2	0.2
BIC		65.4	10.1	1.5	0.2	0.0	0.0	0.0
HQC		75.5	8.2	0.8	0.0	0.0	0.0	0.0
AIC	2	2.5	9.8	3.2	1.4	1.3	1.0	0.8
BIC		4.9	6.7	1.3	0.4	0.1	0.2	0.1
HQC		4.0	4.3	0.6	0.1	0.0	0.0	0.0
AIC	3	2.6	3.2	2.4	0.6	1.1	1.3	0.6
BIC		2.1	1.1	0.6	0.1	0.1	0.0	0.0
HQC		1.4	0.5	0.2	0.0	0.0	0.0	0.0
AIC	4	2.4	1.5	1.2	0.6	0.8	0.6	0.7
BIC		1.4	0.6	0.2	0.0	0.0	0.0	0.0
HQC		1.1	0.3	0.0	0.0	0.0	0.0	0.0
AIC	5	3.3	1.2	0.8	0.7	0.4	0.4	0.4
BIC		1.1	0.1	0.0	0.1	0.0	0.0	0.0
HQC		0.4	0.0	0.0	0.0	0.0	0.0	0.0
AIC	6	4.2	0.8	0.8	0.3	1.0	0.5	0.6
BIC		1.1	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.3	0.0	0.0	0.0	0.0	0.0	0.0
AIC	7	4.0	1.6	0.6	0.8	0.6	0.2	1.1
BIC		0.3	0.2	0.0	0.0	0.0	0.0	0.0
HQC		0.2	0.1	0.0	0.0	0.0	0.0	0.0

Tableau (4.5.32)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p=1,\ \ q=1,\ \ d=2,\ \ N=150$ $\phi_{1,1}=0.5,\ \ \theta_{1,1}=0.9,\ \ \sigma_1^2=1$ $\phi_{1,2}=1.4,\ \ \theta_{1,2}=-0.7,\ \ \sigma_2^2=1$								
	p / q	1	2	3	4	5	6	7
AIC		25.8	7.3	2.2	1.0	0.3	0.6	0.3
BIC	1	67.1	9.8	1.7	0.3	0.0	0.0	0.0
HQC		77.8	9.0	0.8	0.1	0.0	0.0	0.0
AIC		2.5	12.7	3.1	1.9	1.5	1.2	0.8
BIC	2	3.3	9.5	1.6	0.4	0.0	0.1	0.1
HQC		2.2	7.1	0.4	0.1	0.0	0.0	0.1
AIC		1.0	2.6	2.7	0.8	0.9	0.7	0.6
BIC	3	1.2	0.7	0.8	0.0	0.0	0.0	0.0
HQC		0.8	0.3	0.0	0.0	0.0	0.0	0.0
AIC		3.1	1.4	1.0	0.6	0.4	0.5	0.8
BIC	4	1.1	0.3	0.0	0.0	0.0	0.0	0.0
HQC		0.5	0.1	0.0	0.0	0.0	0.0	0.0
AIC		1.9	1.1	1.5	0.8	0.2	0.6	0.3
BIC	5	0.9	0.1	0.0	0.0	0.0	0.0	0.0
HQC		0.5	0.0	0.0	0.0	0.0	0.0	0.0
AIC		2.7	1.4	0.4	0.5	0.4	0.3	0.6
BIC	6	0.8	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.2	0.0	0.0	0.0	0.0	0.0	0.0
AIC		3.7	1.7	1.0	0.8	0.7	0.3	0.8
BIC	7	0.2	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0

Tableau (4.5.33)

Estimation des paramètres du modèle $\text{PARMA}_2(2,1)$ par l'algorithme PRML

paramètres / taille	50	100	150	200
$\phi_{11} = 0.7$	0.7211 (0.3881)	0.7073 (0.2074)	0.7103 (0.1657)	0.6917 (0.1360)
$\phi_{21} = -0.5$	-0.5338 (0.1579)	-0.5078 (0.1118)	-0.5060 (0.0849)	-0.5044 (0.0716)
$\theta_{11} = -0.5$	-0.4763 (0.4198)	-0.4893 (0.2323)	-0.4924 (0.1864)	-0.5140 (0.1546)
$\sigma_1^2 = 1$	0.9408 (0.3087)	0.9525 (0.2040)	0.9659 (0.1668)	0.9848 (0.1408)
$\phi_{12} = 0.4$	0.3429 (0.2268)	0.3810 (0.1587)	0.3915 (0.1354)	0.3861 (0.1091)
$\phi_{22} = 0.3$	0.3279 (0.2290)	0.3137 (0.1608)	0.3033 (0.1346)	0.3095 (0.1067)
$\theta_{12} = 0.7$	0.5570 (0.3007)	0.6331 (0.2058)	0.6548 (0.1672)	0.6603 (0.1467)
$\sigma_2^2 = 1$	0.9154 (0.2960)	0.9519 (0.1920)	0.9707 (0.1629)	0.9694 (0.1399)

Tableau (4.5.34)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p = 2, \; q = 1, \; d = 2, \; N = 50$ $\phi_{1,1} = 0.7, \; \phi_{21} = -0.5, \; \theta_{1,1} = -0.5, \; \sigma_1^2 = 1$ $\phi_{1,2} = 0.4, \; \phi_{22} = 0.3, \; \theta_{1,2} = 0.7, \; \sigma_2^2 = 1$								
	p / q	1	2	3	4	5	6	7
AIC		4.1	4.5	2.5	1.4	0.8	0.6	0.6
BIC	1	18.6	8.4	3.1	1.1	0.3	0.1	0.1
HQC		29.3	9.9	2.2	0.4	0.1	0.0	0.1
AIC		15.5	3.2	1.5	1.9	1.6	1.2	1.1
BIC	2	33.4	4.2	2.1	0.6	0.6	0.5	0.3
HQC		39.3	4.0	1.0	0.2	0.2	0.1	0.0
AIC		6.3	2.4	1.2	1.3	1.2	1.0	1.2
BIC	3	8.3	2.3	1.0	0.4	0.2	0.2	0.1
HQC		7.0	1.3	0.1	0.1	0.0	0.0	0.0
AIC		3.5	1.2	1.2	0.7	0.9	1.1	0.9
BIC	4	3.3	0.7	0.5	0.2	0.3	0.2	0.2
HQC		1.9	0.2	0.0	0.1	0.1	0.0	0.0
AIC		3.2	0.9	0.9	1.0	1.3	1.1	1.9
BIC	5	2.2	0.2	0.4	0.1	0.2	0.1	0.2
HQC		1.2	0.1	0.0	0.0	0.1	0.0	0.0
AIC		3.0	1.5	0.9	0.7	1.3	1.7	1.4
BIC	6	1.2	0.6	0.0	0.0	0.3	0.3	0.0
HQC		0.1	0.2	0.0	0.0	0.0	0.1	0.0
AIC		4.1	2.0	1.0	1.3	1.5	1.4	3.3
BIC	7	1.4	0.7	0.2	0.3	0.2	0.1	0.0
HQC		0.5	0.1	0.0	0.0	0.0	0.0	0.0

Tableau (4.5.35)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p = 2, \quad q = 1, \quad d = 2, \quad N = 100$

$\phi_{1,1} = 0.7, \quad \phi_{21} = -0.5, \quad \theta_{1,1} = -0.5, \quad \sigma_1^2 = 1$

$\phi_{1,2} = 0.4, \quad \phi_{22} = 0.3, \quad \theta_{1,2} = 0.7, \quad \sigma_2^2 = 1$

	p / q	1	2	3	4	5	6	7
AIC		1.4	4.6	2.5	1.0	0.3	0.3	0.2
BIC	1	8.3	8.4	1.6	0.2	0.1	0.1	0.0
HQC		13.1	8.3	1.0	0.0	0.1	0.0	0.0
AIC		26.4	5.1	3.2	1.5	1.4	0.8	0.7
BIC	2	60.0	4.5	1.4	0.3	0.3	0.0	0.0
HQC		66.5	2.3	0.8	0.0	0.2	0.0	0.0
AIC		6.6	3.4	1.6	1.5	1.1	1.1	0.7
BIC	3	7.9	1.8	0.2	0.2	0.1	0.1	0.0
HQC		5.8	0.5	0.1	0.0	0.0	0.0	0.0
AIC		4.0	1.5	0.6	1.4	0.5	0.2	0.8
BIC	4	2.3	0.2	0.1	0.0	0.0	0.0	0.0
HQC		0.8	0.1	0.1	0.0	0.0	0.0	0.0
AIC		3.7	0.9	0.6	0.9	0.4	1.1	0.7
BIC	5	0.9	0.0	0.0	0.1	0.0	0.0	0.0
HQC		0.2	0.0	0.0	0.0	0.0	0.0	0.0
AIC		4.1	1.2	0.8	0.3	0.5	0.6	0.9
BIC	6	0.7	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.1	0.0	0.0	0.0	0.0	0.0	0.0
AIC		3.5	1.9	0.9	0.4	1.1	0.7	0.7
BIC	7	0.1	0.1	0.0	0.0	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0

Tableau (4.5.36)

Fréquences de choix des ordres dans 1000 répétitions pour les différents critères.

$p=2,\ q=1,\ d=2,\ N=150$ $\phi_{1,1}=0.7,\ \phi_{21}=-0.5,\ \theta_{1,1}=-0.5,\ \sigma_1^2=1$ $\phi_{1,2}=0.4,\ \phi_{22}=0.3,\ \theta_{1,2}=0.7,\ \sigma_2^2=1$								
	p / q	1	2	3	4	5	6	7
AIC		0.1	3.4	0.9	1.6	0.9	0.4	0.2
BIC	1	1.6	7.7	1.0	0.5	0.0	0.0	0.0
HQC		3.7	7.9	0.5	0.2	0.0	0.0	0.0
AIC		30.4	5.9	2.3	1.8	0.8	1.4	1.0
BIC	2	72.0	4.6	0.5	0.4	0.0	0.0	0.0
HQC		78.5	3.1	0.1	0.1	0.0	0.0	0.0
AIC		7.9	4.1	1.2	1.1	0.3	1.2	0.7
BIC	3	5.3	1.7	0.0	0.1	0.0	0.0	0.0
HQC		3.3	0.7	0.0	0.1	0.0	0.0	0.0
AIC		4.5	1.9	0.9	0.9	0.5	0.6	0.6
BIC	4	2.3	0.4	0.0	0.1	0.0	0.0	0.0
HQC		1.2	0.0	0.0	0.0	0.0	0.0	0.0
AIC		4.8	1.4	0.5	0.4	0.6	0.6	0.3
BIC	5	1.2	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.4	0.0	0.0	0.0	0.0	0.0	0.0
AIC		2.2	0.4	0.4	0.6	0.3	0.7	0.7
BIC	6	0.1	0.0	0.0	0.0	0.0	0.0	0.0
HQC		0.0	0.0	0.0	0.0	0.0	0.0	0.0
AIC		3.3	1.6	0.9	0.7	0.4	1.1	0.6
BIC	7	0.4	0.1	0.0	0.0	0.0	0.0	0.0
HQC		0.2	0.0	0.0	0.0	0.0	0.0	0.0

Tableau (4.5.37)

Du point de vu informatique, il est très difficile d'utiliser le théoème du coin pour identifier un modèle autoregressif moyenne mobile d-périodique. Car, il n'est pas facile de tester l'hypothèse

$$H_0 : \forall i = 1, ..., d, \quad \Delta^{(i)}(k, h), \quad \text{pour } k \geq q + 1 \text{ et } h \geq p + 1,$$

ce qui est similaire au cas stationnaire. Pour celà, nous présenterons quelques résultats de simulation pour le dernier modèle.

D'après le théorème du coin le modèle est un $\mathrm{ARMA}_2(3,1)$

$p = 2, \; q = 1, \; d = 2, \; N = 50$							
$\phi_{1,1} = 0.7, \; \phi_{21} = -0.5, \; \theta_{1,1} = -0.5, \; \sigma_1^2 = 1$							
$\phi_{1,2} = 0.4, \; \phi_{22} = 0.3, \; \theta_{1,2} = 0.7, \; \sigma_2^2 = 1$							
Le tableau des $\Delta^{(1)}(p, q)$							
p / q	1	2	3	4	5	6	7
1	0.661	0.669	0.236	0.289	0.129	0.178	0.140
2	−0.271	−0.133	−0.005	−0.044	0.036	0.025	0.006
3	−0.032	0.023	−0.005	0.007	−0.003	0.003	−0.001
4	0.001	0.009	−0.001	0.001	0.000	0.000	0.000
5	0.003	0.001	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Le tableau des $\Delta^{(2)}(p, q)$							
p / q	1	2	3	4	5	6	7
1	0.601	0.013	−0.005	−0.155	0.068	0.194	0.167
2	0.385	0.012	0.103	−0.061	−0.047	0.013	−0.007
3	−0.168	0.039	−0.037	0.009	−0.009	0.003	−0.002
4	0.022	0.002	0.005	0.000	0.000	0.000	0.000
5	0.005	0.001	0.000	0.000	0.000	0.000	0.000
6	0.001	0.000	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Tableau (4.5.38)

D'après le théorème du coin le modèle est un $ARMA_2(3,1)$

$p = 2, \quad q = 1, \quad d = 2, \quad N = 100$

$\phi_{1,1} = 0.7, \quad \phi_{21} = -0.5, \quad \theta_{1,1} = -0.5, \quad \sigma_1^2 = 1$

$\phi_{1,2} = 0.4, \quad \phi_{22} = 0.3, \quad \theta_{1,2} = 0.7, \quad \sigma_2^2 = 1$

Le tableau des $\Delta^{(1)}(p,q)$

p / q	1	2	3	4	5	6	7
1	0.585	0.712	0.314	0.351	0.247	0.119	0.029
2	−0.374	−0.228	0.057	0.014	0.005	0.008	0.005
3	0.002	−0.032	−0.011	0.001	−0.003	0.001	−0.001
4	0.021	0.017	−0.002	0.001	0.000	0.000	0.000
5	0.011	0.001	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Le tableau des $\Delta^{(2)}(p,q)$

p / q	1	2	3	4	5	6	7
1	0.578	−0.065	0.108	0.117	0.076	0.085	0.026
2	0.403	−0.110	−0.049	0.017	−0.011	0.004	0.000
3	−0.134	0.009	0.004	0.003	0.000	0.000	0.000
4	0.008	0.001	0.001	0.000	0.000	0.000	0.000
5	0.004	0.000	0.000	0.000	0.000	0.000	0.000
6	0.002	0.000	0.000	0.000	0.000	0.000	0.000
7	0.001	0.000	0.000	0.000	0.000	0.000	0.000

Tableau (4.5.39)

D'après le théorème du coin le modèle est un $\mathrm{ARMA}_2(2,1)$

$p=2,\ q=1,\ d=2,\ N=100$							
$\phi_{1,1}=0.7,\ \phi_{21}=-0.5,\ \theta_{1,1}=-0.5,\ \sigma_1^2=1$							
$\phi_{1,2}=0.4,\ \phi_{22}=0.3,\ \theta_{1,2}=0.7,\ \sigma_2^2=1$							
Le tableau des $\Delta^{(1)}(p,q)$							
$p\ /\ q$	1	2	3	4	5	6	7
1	0.435	0.578	0.029	0.125	0.021	0.027	0.010
2	-0.363	-0.209	0.039	0.014	-0.002	-0.001	-0.002
3	-0.085	-0.027	0.002	0.001	0.000	0.000	0.000
4	0.044	0.001	0.000	0.000	0.000	0.000	0.000
5	0.006	0.000	0.000	0.000	0.000	0.000	0.000
6	0.004	0.000	0.000	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Le tableau des $\Delta^{(2)}(p,q)$							
$p\ /\ q$	1	2	3	4	5	6	7
1	0.493	-0.337	-0.112	0.093	0.034	-0.028	0.015
2	0.552	-0.146	-0.057	0.011	0.004	-0.001	-0.001
3	-0.189	0.017	0.000	0.000	0.000	0.000	0.000
4	-0.022	0.002	0.001	0.000	0.000	0.000	0.000
5	0.016	0.000	0.000	0.000	0.000	0.000	0.000
6	0.002	0.000	0.000	0.000	0.000	0.000	0.000
7	-0.001	0.000	0.000	0.000	0.000	0.000	0.000

Tableau (4.5.40)

Après avoir présenté les résultats de simulation dans les tableaux ci-dessus, nous donnerons dans ce qui suit quelques commentaires sur la performance de chaque critère d'identification.

Résultats relatifs aux modèles autoregressifs pur d-périodique

• Modèle $\mathrm{AR}_2(1)$ Nous remarquons que le critère PDC est le plus performant car il estime le vrai ordre dans 95.1% des cas, pour $N = 50$. D'autre part, le critère AIC est le moins performant, avec une estimation de la vraie valeur de l'ordre inférieure à 50% des cas, même quand la taille de l'échantillon égale à 200.

• Modèle $\mathrm{AR}_3(1)$ Nous constatons cette fois-ci que le critère PDC est moins performant que le HQC. Toutefois, le AIC reste le moins performant de tout les critères étudiés, avec une estimation du vrai ordre inférieure à 36%.

• Modèle $\mathrm{AR}_4(1)$ Nous remarquons que plus la taile de l'échantillon augmente plus les critères PDC, HQC, BIC et CIC se raprochent et deviennent plus consistants. Le critère AIC demeure encore le moins fiable avec une éstimation inférieure à 20% pour $N = 200$.

• Modèle $AR_7(1)$ Nous notons que le critère le plus performant est le PDC, vient ensuite le HQC puis le CIC, le BIC et enfins le AIC, avec les fréquences du choix du vrai ordre respéctives : 98.3%, 72.6%, 35.2%, 33.8% et 3%, pour $N = 140$. Quand la taille de l'échantillon est de 280, les critères les plus performants sont le PDC et le CIC avec plus de 97% de vrais ordre éstimés

• Modèle $AR_2(2)$ Pour $N = 50$, nous constatons que le BIC est le plus performant avec une estimation du vrai ordre dans 73.3%. Nous constatons également que le critère AIC sur-estime le vrai ordre dans 37.9% des cas. D'autre part, les critères PDC, HQC, CIC et BIC sous-estiment la vraie valeur de l'ordre dans 37.1%, 26.5%, 22.1% et 11.3% des cas respectivement. Quand la taille de l'échantillon est de 200, les critères BIC, PDC, HQC et CIC sont très proches et fortement consistants.

• Modèle $AR_4(2)$ Quand la taille de l'échantillon est de 80, les critères PDC, HQC et CIC ont une performance similaire. Pour $N = 200$, le PDC, HQC, CIC ainsi que le BIC sont fortement consistants et très proches. Par contre, l'estimation par le AIC n'excède pas 53%.

Résultats relatifs aux modèles autoregressifs moyennes mobiles d-périodique

• Modèle $ARMA_2(1,1)$: Nous constatons que le critère HQC est le plus performant car il estime les vrais ordres dans 70.5% des cas, pour $N = 50$ et 77.8% pour $N = 150$. Cependant, le BIC est moins performant avec 49.8% pour $N = 50$ et 67.1% pour $N = 150$. D'autre part, le critère AIC a les plus faibles fréquences, avec une estimation des vrais ordres dans moins de 26% des cas.

• Modèle $ARMA_2(2,1)$: Nous remarquons cette fois-ci que les trois critères estiment les vrais ordres dans moins de 40% des cas, pour $N = 50$. Nous remarquons également que les critères BIC et HQC tendent de plus en plus à la sélection des vraies valeurs des ordres du modèle lorsque la taille de l'échantillon augmente. En effet, ils estiment les vrais ordres dans plus de 60% et de 72% des cas, quand la taille est de 100 et 150 respectivement. Alors que l'estimation par le critère AIC n'excède pas les 31% quand la taille de l'échantillon est de 150.

En résumé, les critères PDC, HQC, BIC et CIC donnent les meilleurs résultats dans le cas d'un modèle autoregressif pur d-périodique, alors que le HQC et le BIC sont les plus performants dans le cas autoregressif moyenne mobile d-périodique, ce qui est similaire aux résultats obtenus dans le cas stationnaire.

Pour des échantillons de petites tailles, nous constatons également, que les critères PDC et HQC donnent toujours les meilleurs résultats. Par contre, le critère AIC sur-estime le vrai ordre du modèle, ce qui correspond aux résultats obtenus dans le cas classique. Cependant, lorsque la taille est assez grande, les résultats obtenus à partir des critères PDC, HQC, BIC et CIC sont très proches et fortement consistants (se qui a été démontré par Hemis (1999) et Bentarzi

(2000) en particulier pour le PDC).

De même que pour le cas classique, le temps de calcul ainsi que l'espace mémoire sont beaucoup plus grand dans le cas périodique pour le critère PDC que pour les autres. Il est donc plus judicieux, en pratique, d'utiliser les critères HQC, BIC et CIC que le PDC.

Dans le cas d'un modèle ARMA mixte d-périodique les fréquences de choisir les vraies valeurs des ordres par les critères HQC et BIC, augmente avec l'augmentation de la taille de l'échantillon, se qui démontre la consistance des estimateurs obtenus par ces critères. Ces résultats sont similaire a ceux obtenus dans le cas des modèles ARMA classique.

Chapitre 5

Conclusion

L'identification d'un modèle est considérée comme l'un des problème majeurs dans l'analyse des séries chronologiques car elle représente une étape très importante dans la modélisation.

Le but de notre travail était d'étudier le problème de l'identification d'un modèle autorégressif moyenne mobile. Nous nous sommes intéressés, en particulier, au problème de la sélection des ordres d'un modèle ARMA périodique.

Dans notre travail, nous avons présenté une synthèse des critères de sélection des modèles ARMA dans le cas stationnaire et leurs extension dans le cas périodique.

Nous avons présenté également, les propriétés théoriques des estimateurs obtenus par ces critères.

Nous avons fait une étude de simulation pour les deux cas stationnaire et périodique, dans le but :

- Examiner et évaluer la performance de chaque critère.
- Comparer les performances des différents critères.

Nous avons pu constater à partir de plusieurs simulations réalisées (1000 simulations) pour des modèles ARMA stationnaire et périodiques que les critères HQC et BIC sont les plus performants dans le cas mixte tant stationnaire que périodique.

Nous constatons également, que la fréquence de choisir les vrais ordres par les critères HQC et BIC augmente proportionnellement avec la taille de l'échantillon. Par conséquent, les estimateurs obtenus par ces critères sont fortement consistants, ce qui a été prouvé par plusieurs travaux de recherche.

Nous avons noté qu'il est difficile d'utiliser le théorème du Coin que ce soit dans le cas stationnaire ou périodique pour identifier un modèle ARMA mixte en raison de la difficulté de tester simultanément l'hypothèse $\Delta^{(i)}(k,h)=0$ pour $k \geq p+1$ et $h \geq q+1$, pour tout $i = 1, ..., d$.

Alors que lorsque l'ordre moyenne mobile est nul, nous remarquons que les meilleurs résultats sont donné par les critères PDC, HQC, BIC et CIC dans le cas stationnaire et périodique.

Nous remarquons également, que les propriétés statistiques (la convergence et la consistance) des estimateurs obtenus par les critères PDC, HQC, BIC et CIC sont conservés dans le cas périodique. Cependant, le critère AIC à tendance à surestimer la vraie valeur de l'ordre du modèle dans le cas stationnaire et périodique.

Bien que performante, l'identification d'un modèle autoregressif par le critère PDC est plus coûteuse, mais elle demeure très utile pour l'estimation des paramètres par la méthode de léchantillonnage de Gibbs.

Nous Notons également que malgré sa bonne performance, le critère PDC na pas pu être généralisé pour identifier un ARMA mixte stationnaire ou périodique.

Enfin, il apparaît de notre étude que les critères généralisés au cas périodique conserves les mêmes performances du cas classique.

Bibliographie

[1] G. J. Adams and G. C. Goodwin (1995), Parameter Estimation for Periodic ARMA Models, J. Time series analysis, vol 16. No 2, p. 127-145.

[2] H. Akaike (1969), Fitting autoregressive models for prediction, Ann. Inst. Statist. Math., 21, p. 243-247.

[3] H. Akaike (1970), Statistical predictor identification, Ann. Inst. Statist. Math., 22, p. 203-217.

[4] H. Akaike (1972), Use of an information theoretic quantity for statistical model identification, Proc. 5th Hawail Inter. Conference on System Sciences, p. 249-250.

[5] H. Akaike (1973a), Maximum likelihood identification of Gaussian autoregressive moving average models, Biometrika, 60, p. 255-265.

[6] H. Akaike (1973b), Block Toeplitz matrix inversion, SIAM J. App. Math., 24, p. 234-241.

[7] H. Akaike (1977), An objective use Bayesian models, Ann. Statist. Math., 29, p. 9-20.

[8] A. Aknouche (2000), Estimation en-ligne et hors-ligne des Modèles ARMA Périodiques, Thèse de Magister, Institut de Mathématique, USTHB.

[9] T. W. Anderson (1971), The Statistical Analysis of Time Series, Wiley, New York.

[10] P. L. Anderson, M. M. Meerschaert and A. V. Vecchia (1999), Innovations Algorithm for Periodically Stationary Time Series, Stochastic Processes and their Applications, 83, p. 149-169.

[11] Z. D. Bai, K. Subramanyam, and L. C. Zhao (1988), On Determination of the Order of an Autoregressive Model, J. Multivariate Analysis 27, p. 40-52.

[12] J. M. Beguin, C. Gourieroux and A. Monfort (1980), Identification of a mixed autoregressive moving average process : The corner method, in Time Series, O. D. Anderson, Ed., North-Holland, Amesterdam, p. 423-436.

[13] M. Bentarzi (1995), Modèles de Séries Chronologiques à Cofficients Périodiques, Thèse de Doctorat, Institut de Mathématique, USTHB.

[14] M. Bentarzi (1998), Model-Building Problem of Periodically Correlated m-Variate Moving Average Processes, J. Multivariate Analysis 66, p. 1-21.

[15] M. Bentarzi (2000), Predictive Density Order Selection Criterion of Periodic Autoregressve, A paraître.

[16] M. Bentarzi and A. Aknouche (2002), On-line Estimation Algorithm for Autoregressive Models, A paraître.

[17] M. Bentarzi and A. Aknouche (2003), On-line Estimation Algorithm for Autoregressive Moving Average Models, A paraître.

[18] M. Bentarzi and M. Hallin (1994), On The invertibility of Periodic Moving Average Models, J. Time series analysis, vol 15. No 3, p. 263-268.

[19] M. Bentarzi and M. Hallin (1996), Locally Optimal Tests Against Periodic Autoregression, Parametric and Nonparametric Approaches, Econometric Theory, 12, p. 88-112.

[20] G. N. Boshnakov (1996), Recursive Computation of the Parameters of Periodic Autoregressive Moving Average Processes, J. Time Series Analysis, Vol. 17, No 4, p. 333-349.

[21] G. E. P. Box and G. M. Jenkins (1976), Time Series Analysis, Forecasting and Control, Holden-Day, Revsed Edition, San Francisco.

[22] P. J. Brockwell and R. A. Davis (1988), Time Series : Theory and Methods, Springer Verlag. New York.

[23] B. S. Choi (1986), An algorithm for solving the extended Yule-Walker equations of an autoregressive moving average time series, IEEE Trans. Information Theory, IT-32, p. 417-419.

[24] B. S. Choi (1987), A newton-Raphson solution for MA parameters of mixed autoregressive moving average process, J. Korean Statist. Soc., 16, p. 1-9.

[25] B. S. Choi (1990a), The asymptotic distribution of the extended Yule-Walker estimates of a mixed ARMA process, Technical Report 140, Department of Statistics, University of California, Santa Barbara, CA 93106, October 14.

[26] B. S. Choi (1990b), The asymptotic joint distribution of the correlation determinant estimates of an ARMA process, Technical Report 147, Department of Statistics, University of California, Santa Barbara, CA 93106, November 6,1990. Also, appear in Commun. Statist., A20, p. 2823-2835.

[27] B. S. Choi (1990c), The asymptotic distributions of the θ, λ and η functions for identifying a mixed ARMA process, Technical Report 152, Department of Statistics, University of California, Santa Barbara, CA 93106, November 27.

[28] B. S. Choi (1990d), Two chi-square statistics for determining orders of an ARMA process, Technical Report 156, Department of Statistics, University of California, Santa Barbara, CA 93106, December 31.

[29] B. S. Choi (1991a), On the asymptotic distribution of the generalized partial autocorrelation function in ARMA processes, J. Time Series Anal., 12, p. 193-205.

[30] B. S. Choi (1991b), An algorithm for Hannan and Rissanen's ARMA modeling method, Technical Report 178, Department of Statistics, University of California, Santa Barbara, CA 93106, June 25.

[31] B. S. Choi (1992), ARMA Model Identification, Srinpger Verlag.

[32] O. Ciftcioglu, J. E. Hoogenboom, and Hugo van Dam (1994), A Consitent Estimator for the Model Order of an Autoregressive Process, IEEE Tranaction on Signal Processing, vol. 42, No. 6, p. 1471-1476.

[33] T. Cipra (1985), Periodic Moving Average Process, Aplikace Matematiky 30, p. 218-229.

[34] W. P. Cleveland and G. C. Tiao (1979), Modeling Seasonal Time Series, Revue Economic Appliquee 32, p. 107-129.

[35] G. Cybenko (1980), The numerical stability of the Levinson-Durbin algorithme for Toepliz systems of equations, SIAM J. Scientific Statist. Comput., 1, p. 303-319.

[36] P. M. Djuric and S. M. Kay (1992), Order Selection of Autoregressive Models, IEEE Tranaction on Signal Processing, vol. 40, No. 11, p. 2829-2833.

[37] J. Franke (1985), A Levinson-Durbin recursion for autoregressive moving average processes, Biometrika, 72, 3, p. 573-581.

[38] P. H. Franses (1996), Periodicity and Stochastic Trends in Economic Times Series, Oxford University Press.

[39] A. Hamdi (1982), Identification des processus ARMA à coefficients dépendant du temps, Thèse de Doctorat, Université Paul Sabatier, Toulouse.

[40] E. J. Hannan (1980), The estimation of the order of an ARMA process, Ann. Statist., 8, p. 1071-1081.

[41] E. J. Hannan (1982), Testing for autocorrelation and Akaike's criterion, in Essays in Statistical Science, Special volume 19A of J. App. Prob., The Applied Probability Trust, Sheffield, p. 403-412.

[42] E. J. Hannan and B. G. Quinn (1979), The Determination of the Order of an Autoregression, J. R. Statist. Sco. B, No 2, p. 190-195.

[43] R. Hemis (1999), Sélection de l'Ordre des Modèles Autoregressifs Périodiques, Approche Classique et Approche Bayesienne, Thèse de Magister, Institut de Mathématique, USTHB.

[44] R. L. Kashyap (1977), A Bayesian comparison of different classes of dynamic models using empirical data, IEEE Trans. Automatic Control, AC 22, p. 715-727.

[45] R. L. Kashyap (1978), Optimal Feature Selection and Decision Rules in Classification Problems with Time Series, IEEE Trans, Information Theory, Vol. IT 25, No 3, May.

[46] R. L. Kashyap (1980), Inconsistency of the AIC Rule for Setimating the Order of Autoregressive Models, IEEE Trans, Automatic Control, Vol. AC 25, No. 5, October.

[47] R. L. Kashyap (1982), Optimal choice of AR and MA parts in autoregressive moving average models, IEEE Trans. Pattern Analysis and Machine Intelligence, PAMI 4, p. 99-104.

[48] R. Lund and I. V. Basawa (1999), Modeling and Inference for Periodically Correlated Time Series, In Asymptotics Nonparametrics and Time Series, p. 37-62.

[49] R. Lund and I. V. Basawa (2001), Large Sample Properties of Parameter Estimates for Periodic ARMA Models,, J. Time series analysis, vol 22. No 6, p. 651-663.

[50] A. I. MacLeod (1975), The Derivation of the theorical autocovariances function of autoregressive moving average time series, App. Statist. 24, p. 255-256.

[51] A. I. McLeod (1992), Parsimony, Model Adequacy and Periodic Correlation in Time Series Forecasting, international Statistical Review, Vol. 61, No. 3, p. 387-393.

[52] A. I. McLeod (1995), Diagnostic Checking Periodic Autoregression Models with application, J. Time series analysis, Vol. 15, No 2, p. 221-233.

[53] M. Pagano (1978), On Periodic and Multiple Autoregressions, Annal. Statist., Vol. 6, No. 6, p. 1310-1317.

[54] H. Sakai (1982), Circular Lattice Filtering Using Pagano's Method, IEEE Trans, Acoustics, Speech, Signal Processing, Vol. Assp 30, No. 2, Avril.

[55] G. Schwarz (1978), Estimating the Dimension of a Model, Ann. Statist., Vol. 6, No 2, p. 461-464.

[56] R. Shibata (1976), Selection of the order of an Autoregressive Model by Akaike's Information criterion, Biometrika, Vol. 63, No. 1, p. 117-126.

[57] T. A. Ula and A. A. Smadi (1997), Periodic stationarity Conditions for Periodic Autoregressive Moving Average Processes as Eigenvalue Problems, Water Resources Research, Vol. 33, No. 8, p. 1929-1934.

[58] A. V. Vecchia (1985a), Maximum Likelihood Estimation for Periodic Autoregressive Moving Average Models, Technometrics, Vol. 27, No. 4, p. 375-384.

[59] A. V. Vecchia (1985b), Periodic Autoregressive moving average (PARMA) modeling with application to water resources. Water Resources Bulletin, October, Vol. 21, No. 5, p. 721-730.

[60] G. T. Wilson (1979), Some Efficient Computational Procedures for High Order ARMA Models, J. Statist., Coomput., Simul., Vol 8, p. 301-309.