

**Université Mentoutri Constantine**

**Faculté des Sciences Exactes**

**Département de Mathématiques**

N° d'ordre :...

N° de série :...

**Mémoire**

**Présenté pour l'obtention du diplôme de MAGISTER en MATEMATIQUE**

**Option**

**Processus stochastique**

**Thème**

**Modélisation empirique des séries  
chronologiques**

**Présenté par**

**DJAOUT ASSIA**

**Devant le jury :**

<b>Président :</b>	<b>DENCHE Mohamed</b>	<b>Professeur</b>	<b>U.M.C</b>
<b>Rapporteur :</b>	<b>DAKHMOUCHE Maghlaoui</b>	<b>M.C</b>	<b>U.M.C</b>
<b>Examineurs :</b>	<b>BOUSHABA Mahmoud</b>	<b>M.C</b>	<b>U.M.C</b>
	<b>BENCHETTAH Azzedine</b>	<b>Professeur</b>	<b>U. Annaba</b>

**Soutenu le : 20 /04/2011**

## REMERCIEMENTS

*Je tiens en tout premier lieu à remercier mon père et ma mère et tous mes frères et mes sœurs pour leur soutien précieux.*

*Je remercie également mon directeur de recherche Monsieur: Meghlaoui Dakhmouche maître de conférences à l'université de Constantine, pour sa fraîcheur d'esprit.*

*Je remercie sincèrement, les membres de jury d'avoir bien voulu accepter de faire partie de la commission d'examineur.*

*J'adresse Mes remerciements aussi, à tous mes amis, mes collègues et tous qui me connaissez.*

*A toute personne qui a participé de près ou de loin pour l'accomplissement de ce modeste travail.*

# Table des matières

<b>Abstract</b>	<b>iii</b>
<b>Résumé</b>	<b>iv</b>
<b>1 Rappels et définitions utiles</b>	<b>4</b>
1.1 Le principe classique d'analyse des séries chronologiques . . . .	4
1.1.1 Estimation du trend en l'absence de saisonnalité . . . .	5
1.1.2 Elimination du trend et de la saisonnalité . . . . .	9
1.2 Rappel sur les processus aléatoires . . . . .	10
1.3 Modèle linéaire général . . . . .	10
1.3.1 Relation entre les poids $\psi$ et les poids $\pi$ . . . . .	12
1.3.2 La fonction génératrice des autocovariances d'un processus linéaire . . . . .	13
1.3.3 La fonction d'autocorrélation partielle . . . . .	14
1.4 Le périodogramme d'un processus . . . . .	16
1.4.1 Le périodogramme d'un processus linéaire stationnaire	19
<b>2 Modèles de Box-Jenkins</b>	<b>20</b>
2.1 Le modèle $AR$ . . . . .	20
2.1.1 Principales caractéristiques d'un modèle $AR(p)$ . . . .	20
2.2 Le modèle $MA$ . . . . .	22
2.2.1 Principales caractéristiques d'un modèle $MA(q)$ . . . .	23
2.3 Le modèle $ARMA(p, q)$ . . . . .	25
2.3.1 Principales caractéristiques d'un modèle $ARMA(p, q)$ .	25
2.4 Causalité et inversibilité des modèles $ARMA$ . . . . .	27
2.5 Modèles linéaires non-stationnaires . . . . .	29
2.5.1 Modèle ARIMA . . . . .	29
2.5.2 Modèle SARIMA . . . . .	29

2.6	Méthode de prévision . . . . .	30
2.6.1	Principe de prévision . . . . .	30
2.6.2	Prévision à l'aide d'un processus <i>ARMA</i> . . . . .	31
<b>3</b>	<b>Identification de modèles pour les séries chronologiques</b>	<b>33</b>
3.1	Position du problème . . . . .	33
3.2	Ajustement d'un modèle et détermination de son ordre . . . . .	34
3.2.1	Méthode d'identification de Box&Jenkins . . . . .	34
3.3	Détermination de l'ordre d'un modèle . . . . .	39
3.3.1	Le critère d'information d'Akaiké . . . . .	40
3.3.2	Le critère d'information de Bayésien . . . . .	43
<b>4</b>	<b>Représentation interne des séries chronologiques</b>	<b>46</b>
4.1	Introduction . . . . .	46
4.2	Modèle état-espace . . . . .	46
4.2.1	Propriétés des représentations internes . . . . .	47
4.3	Représentation état-espace d'un processus stationnaire . . . . .	48
4.4	Représentation état-espace d'un processus <i>ARMA</i> ( $p, q$ ) . . . . .	49
4.5	La fonction d'autocovariance d'un processus <i>ARMA</i> causal . . . . .	52

# Abstract

The objective of this work is to study the technical analysis and the modeling of empirical time series. A time series is a sequence of random variables  $X_{t_1}, X_{t_2}, \dots, X_{t_n}$  observed sequentially in time. These random variables are not independent. This dependence between random variables is measured using the autocorrelations and partial autocorrelations. The functions of autocorrelations and partial autocorrelations of the time series are the essential tools for the identification and adjustment of stochastic models.

In this contribution we focus mainly on linear stochastic models. We synthesized the remarkable properties of *ARMA* models using the techniques of Box and Jenkins and some specificity of the "State-Space" models. It also shows that any *ARMA* model can be expressed as "State-Space". To give more substance to the identification of models, we introduced the concept of inverse autocorrelation function which depends heavily on the specific model type. The information criteria *AIC* and *BIC* are proposed to test the quality of the model chosen. A reformulation of the construction of these criteria and their comparison are proposed.

**Key-Words and Phrases :**

Time series - Stochastic processes - *ARMA* models - Externe representation - State-space representation - *AIC* Criterion - *BIC* Criterion.

# Résumé

L'objectif de ce travail est d'étudier les techniques d'analyse et de modélisation des séries chronologiques empiriques. Une série chronologique est une suite de variables aléatoires  $X_{t_1}, X_{t_2}, \dots, X_{t_n}$  observées séquentiellement dans le temps. Ces variables aléatoires ne sont donc pas indépendantes. Cette dépendance entre les variables aléatoires est mesurée à l'aide des autocorrélations et des autocorrélations partielles. Les fonctions d'autocorrélations et d'autocorrélations partielles des séries chronologiques représentent des outils essentiels pour l'identification et l'ajustement de modèles stochastiques.

Dans ce travail on s'intéresse essentiellement aux modèles stochastiques linéaires ; on a synthétisé les propriétés remarquables des modèles *ARMA* en utilisant les techniques de Box et Jenkins ainsi que certaines spécificités des modèles «Etat-Espace». On montre aussi que tout modèle *ARMA* peut être exprimé sous la forme «Etat-Espace».

Pour donner plus de consistance à l'identification de modèles, on a introduit la notion de fonction d'autocorrélation inverse dont la spécificité dépend étroitement du type de modèle. Les critères d'information *AIC* et *BIC* sont proposés pour éprouver la qualité du modèle choisi. Une reformulation des constructions de ces critères et leur comparaison sont proposées.

**Mots clés :**

Série chronologique - Processus stochastique- Modèles *ARMA*- Représentation externe - Représentation interne - Critère *AIC* - Critère *BIC*.

---

# Introduction

Les séries chronologiques ont commencé à être utilisées dans l'étude des températures et de la pluviométrie. L'étude descriptive de ces séries n'était pas suffisante pour prévoir leurs comportements futurs. Le besoin de prévision de ce type de séries a emmené les scientifiques à construire des modèles mathématiques aptes à répondre à ce besoin.

La majeure partie de la méthodologie statistique classique repose sur des séries d'observations indépendantes. Généralement, ce manque d'indépendance est considéré comme un handicap. Ainsi, un des objectifs d'une bonne expérimentation est d'éliminer cette dépendance. Toutefois, dans le cas de l'analyse des séries chronologiques nous nous occupons de données qui se développent dans le temps et dont chaque observation peut dépendre dans une certaine mesure des observations précédentes. C'est sur cette dépendance que se focalise l'étude des séries chronologiques. Par ailleurs, on peut dire qu'une série chronologique possède une mémoire du passé dans le sens où les valeurs présentes et passées de la série reflètent dans une certaine mesure les valeurs futures.

Une série chronologique est une suite d'observations générées séquentiellement dans le temps. On dit qu'une série chronologique est continue si l'ensemble des instants d'observation est continu, et qu'elle est discrète si l'ensemble de ces instants est discret. Par la suite, on ne considèrera que les séries chronologiques discrètes où les observations sont réalisées en des intervalles de temps équidistants, i.e. d'amplitude  $h$  constante. Ainsi, si nous disposons de  $N$  valeurs successives d'une série chronologique, on note  $X_{t_1}, X_{t_2}, \dots, X_{t_N}$  pour désigner les observations obtenues dans les intervalles de temps  $t_1 = t_0 + h, t_2 = t_0 + 2h, \dots, t_N = t_0 + Nh$ . La valeur  $t_0$  est appelée origine de temps et la valeur  $h$  est appelée unité de temps.  $X_t$  est alors une observation de la série chronologique à l'instant  $t$ . Les séries chronologiques peuvent être obtenues de deux façons : soit par échantillonnage dans un processus aléatoire continu ou par accumulation d'une variable sur une certaine période de temps. La représentation graphique d'une série chronologique nous donne un aperçu sur le comportement de cette série. L'allure d'une série chronologique se matérialise souvent sous forme d'une tendance ou composante déterministe. Elle peut montrer des changements cycliques ou composante saisonnière. Et indépendamment de ces deux composantes, il existe aussi une variabilité due à des causes purement aléatoires. Un phénomène statistique qui se développe dans le temps et qui obéit à des lois de

probabilité est appelé processus stochastique ou processus aléatoire.

Les séries chronologiques qu'on aura à étudier seront considérées comme des réalisations possibles d'un processus stochastique. Il est nécessaire dans ce qui va suivre de distinguer entre le processus stochastique et la série chronologique observée. Une série chronologique  $X_{t_1}, X_{t_2}, \dots, X_{t_N}$  de  $N$  observations successives est considérée comme la réalisation d'un échantillon issu de la population infinie de telles possibilités de réalisation du processus stochastique.

Dans le premier chapitre, on présentera des rappels et des définitions pour nous aider à mieux comprendre l'étude descriptive qui sera exposée dans les chapitres qui suivent. Nous discutons également sur les séries chronologiques déterministes où nous utilisons la tendance et la saisonnalité. Ensuite, nous éliminons ces dernières composantes et nous gardons que la partie aléatoire. Aussi, nous donnons des définitions sur le processus stochastique

En général, Box et Jenkins ([6]) ont proposé une méthode beaucoup plus sophistiquée se concentrant principalement sur la composante aléatoire. L'approche de Box-Jenkins a longtemps été considéré comme l'état de l'art dans la série chronologique. Essentiellement, les modèles *ARIMA* sont des généralisations des modèles linéaires ordinaires de régression, sauf que les "variables explicatives" se composent des observations antérieures de la série chronologique elle-même. Si les données de séries chronologiques s'ajustent effectivement à un modèle *ARIMA* de forme connue, alors les meilleures prévisions possibles pourraient être obtenues par des méthodes de projection classiques. Bien sûr, dans la pratique, même si les données ont été générées par un modèle *ARIMA*, nous ne connaissons pas le nombre de coefficients dans ce modèle ou les valeurs des coefficients. Cela conduit à des problèmes de sélection de modèle et d'estimation de paramètres.

Le premier chapitre est consacré à l'étude des différentes notions de la tendance et de saisonnalité. Des techniques d'élimination de ces composantes sont proposées. Dans un deuxième volet, nous présentons l'étude de modèles linéaires généraux et leurs caractéristiques à savoir la fonction d'autocovariance, la fonction d'autocorrelation, la fonction d'autocorrelation partielle et le périodogramme. Le deuxième chapitre portera sur les modèles de Box-Jenkins et leurs propriétés caractéristiques.

Dans le troisième chapitre on abordera la méthode d'identification de modèles de Box et Jenkins. Cette technique est semi automatique du fait que l'utilisateur intervient en appréciant visuellement les graphes des fonctions d'autocorrélation et des fonctions d'autocorrélation partielle. La confirmation du modèle choisi se fait en utilisant l'un critères d'information, le critère *AIC*



---

ou le critère *BIC*. Le dernier chapitre est consacré à la représentation interne ou la représentation de Kalman-Bucy dans laquelle les séries sont décrites à l'aide de l'équation d'observation :

$$Y_t = G_t X_t + W_t \quad t = 1, 2, \dots$$

et l'équation d'état :

$$X_{t+1} = F_t X_t + V_t \quad t = 1, 2, \dots$$

# Chapitre 1

## Rappels et définitions utiles

### 1.1 Le principe classique d'analyse des séries chronologiques

Les techniques moderne pour l'étude des séries chronologiques ont été initiées par Yule ([32]). Les travaux de Wold ([31]) ont permis de développer une théorie complète des modèles d'autoregressif- moyenne mobile (*ARMA*). Dans les années quarante Wiener ([30]) et Kolmogoroff ([19]) ont résolu le problème de l'estimation des filtres continus et discrets respectivement. Au début des années soixante Kalman([17]) et Kalman et Bucy ([18]) ont étendu les procédures d'estimation de Wiener ([30]) et Kolmogoroff ([19]) aux séries chronologiques non stationnaires en utilisant la représentation état-espace.

L'objectif de l'approche de l'étude des séries chronologiques est de séparer la composante aléatoire de la partie déterministe du processus étudié. Cette étude se compose de deux étapes. La première est appelée "analyse des séries chronologiques" dont l'objectif est de mettre en évidence les caractéristiques du phénomène engendrant la série chronologique. Ceci est réalisé en utilisant simultanément les propriétés des autocorrélations et les propriétés spectrales de la série chronologique. La deuxième étape consiste en l'utilisation de plusieurs types de méthodes telles que la décomposition de la série chronologique, lissage exponentiel ou les modèles autorégressifs-moyennes mobiles. Toutes ces méthodes tendent à isoler le bruit blanc dans les valeurs observées de la série chronologique.

Le principe de décomposition des séries chronologiques est basé sur l'hy-

## 1.1. LE PRINCIPE CLASSIQUE D'ANALYSE DES SÉRIES CHRONOLOGIQUES

---

pothèse que toute série chronologique est constituée de trois composantes principales, la tendance ( $T$ ), la saisonnalité ( $S$ ) et la partie aléatoire ( $Z_t$ ). Soit  $X_t$  le processus observé, si on suppose que la relation entre les trois composantes est additive, on aura alors :

$$X_t = T + S + Z_t$$

La partie aléatoire est bien évidemment la partie la plus intéressante à modéliser pour une série chronologique, on note  $Z_t$  la composante aléatoire, supposée de moyenne nulle, mais possédant en général une structure de corrélation non nulle.

### 1.1.1 Estimation du trend en l'absence de saisonnalité

Supposons, dans un premier temps, que la partie déterministe du modèle soit uniquement composée d'un trend  $T_t$  linéaire :

$$X_t = T_t + Z_t \tag{1.1}$$

Nous allons résumer trois méthodes principales pour estimer  $T_t$ .

#### Estimation paramétrique

Cette méthode consiste à estimer le trend par la méthode des moindres carrés ordinaires. Supposons que l'on observe la série chronologique, il semble naturel d'estimer la composante  $T_t$  de cette série par une fonction linéaire  $\widehat{T}_t = \widehat{a} + \widehat{b}t$  où  $\widehat{a}$  et  $\widehat{b}$  sont des estimateurs des coefficients de la fonction linéaire estimant  $T_t$ . Pour trouver ces estimateurs par la méthode des moindres carrés, il faut minimiser l'erreur quadratique commise en remplaçant  $T_t$  par  $\widehat{T}_t$  dans (1.1).

**Proposition 1.1.1** *Soit  $X_1, \dots, X_N$ , la série chronologique observée. Alors, pour déterminer les coefficients  $\widehat{a}$  et  $\widehat{b}$ , on doit minimiser la quantité :*

$$\sum_{t=1}^N (X_t - a - bt)^2 \tag{1.2}$$

Et, en posant  $\bar{X} = \frac{1}{N} \sum_{t=1}^N X_t$ , la solution de ce problème de minimisation est :

$$\begin{cases} \hat{a} = \frac{4N+2}{N-1} \bar{X} - \frac{6}{N(N-1)} \sum_{t=1}^N tX_t \\ \hat{b} = \frac{6}{N(N-1)} \left( \frac{2}{N(N+1)} \sum_{t=1}^N tX_t - \bar{X} \right) \end{cases}$$

**Preuve** Pour minimiser en  $a$  ;  $b$  l'équation (1.2), il faut dériver cette expression selon  $a$  et  $b$ , ce qui donne, après quelques simplifications et en égalant à zéro, le système :

$$\begin{aligned} (i) \quad & \sum_{t=1}^N X_t - aN - b \sum_{t=1}^N t = 0 \\ (ii) \quad & \sum_{t=1}^N tX_t - a \sum_{t=1}^N t - b \sum_{t=1}^N t^2 = 0 \end{aligned}$$

On peut à présent utiliser les formules usuelles :

$$\sum_{t=1}^N t = \frac{N(N+1)}{2} \quad \text{et} \quad \sum_{t=1}^N t^2 = \frac{N(N+1)(2N+1)}{6}$$

En divisant (i) et (ii) par  $N$ , on obtient :

$$\begin{aligned} \bar{X} - a - \frac{b(N+1)}{2} &= 0 \\ \frac{1}{N} \sum_{t=1}^N tX_t - \frac{a(N+1)}{2} - \frac{(N+1)(2N+1)}{6} b &= 0 \end{aligned}$$

où  $\bar{X}$  est la moyenne empirique de  $\{X_t\}_{t=1}^N$ .

La solution découle directement de ce système, en vérifiant que  $\hat{a}$  et  $\hat{b}$  sont bien évidemment des minima en considérant les dérivées partielles du second ordre. Des fois, on soupçonne un trend quadratique :

$$\hat{T}_t = \hat{a} + \hat{b}t + \hat{c}t^2 \tag{1.3}$$

Par la même méthode on détermine les estimations de  $\hat{a}$ ,  $\hat{b}$ , et  $\hat{c}$ . La courbe d'équation (1.3) peut alors être utilisée pour prévoir les valeurs futures de la

## 1.1. LE PRINCIPE CLASSIQUE D'ANALYSE DES SÉRIES CHRONOLOGIQUES

---

série,  $X_t$  pour  $t > N$ . Cependant, si on tient compte des résidus  $Z_t = X_t - \widehat{T}_t$ , on peut améliorer cette prévision en tenant compte de la structure de  $Z_t$ . En d'autres termes, si les résidus  $Z_t$  présentent des corrélations entre ses réalisations, on pourra utiliser cette affirmation pour affiner la prévision par le trend ■

### Estimation non paramétrique

Dans certaines situations, il n'est pas facile de trouver le degré du polynôme d'ajustement pour  $T_t$ . Il n'est pas toujours possible d'utiliser la méthode des moindres carrés car le polynôme utilisé au départ pour  $\widehat{T}_t$  n'est parfois ni linéaire ni quadratique. On pourrait utiliser un polynôme avec un degré élevé, mais le nombre de paramètres à estimer serait important et rendrait les calculs fastidieux. Dans cette situation, on a recourt des méthodes non paramétriques de l'estimation du trend, qui ne suppose pas que la tendance soit polynomiale a priori. En effet, supposons que  $T_t$  soit linéaire dans un intervalle  $[t - q; t + q]$ . Dans ce cas, un bon estimateur du trend est donné par

$$\widehat{T}_t = \frac{1}{2q + 1} \sum_{k=-q}^q X_{t+k}$$

On peut à présent calculer  $\widehat{T}_t$  pour chaque valeur de  $t$  en calculant cette moyenne sur les  $2q$  observations autour de  $t$ . On peut aussi voir l'intervalle  $[t - q; t + q]$  comme une "fenêtre" sur les observations, que l'on déplace lorsque  $t$  varie. A chaque valeur de  $t$ , l'estimateur  $\widehat{T}_t$  calcule la moyenne des observations tombant dans cette fenêtre glissante, on dit alors qu'on effectue une estimation par moyenne mobile.

Pour éviter que des problèmes se posent aux bords de la série (lorsque  $t \leq q$ ; ou  $t > N - q$ ), on pose :

$$\begin{cases} X_t = X_1 & \text{pour } t < 1 \\ X_t = X_T & \text{pour } t > N \end{cases}$$

### Méthode des différences itérées

On peut également éliminer le trend sans l'estimer. Définissons tout d'abord l'opérateur de retard comme étant la fonction linéaire  $B$  qui, à toute variable  $X_t$  fait correspondre la variable précédente  $X_{t-1}$  tel que :

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

Cet opérateur peut être utilisé récursivement, pour obtenir l'avant-dernière observation avant  $t$  :

$$B^2(X_t) = B(B(X_t)) = B(X_{t-1}) = X_{t-2}$$

et on peut écrire par induction :

$$B^k(X_t) = X_{t-k}$$

Pour alléger la notation, nous noterons l'application de l'opérateur  $B$  sur  $X_t$  :  $BX_t$ , sans les parenthèses.

On introduit également l'opérateur de différence  $\nabla$  dont l'application à  $X_t$  fournit la différence entre  $X_t$  et la valeur précédente  $X_{t-1}$  de la série :

$$\nabla X_t = X_t - X_{t-1}$$

En vertu de la définition précédente, on a directement :

$$\nabla X_t = X_t - BX_t = (1 - B)X_t$$

Tout comme avec l'opérateur  $B$ , on peut itérer  $\nabla$  pour obtenir :

$$\begin{aligned} \nabla^2 X_t &= \nabla \nabla X_t = (1 - B)[(1 - B)X_t] = (1 - 2B + B^2)X_t \\ &= X_t - 2X_{t-1} + X_{t-2} \end{aligned}$$

Plus généralement, l'opérateur de différence d'ordre  $d$  envoie la valeur  $X_t$  de la série observée sur la différence de  $X_t$  avec la valeur prise par la série au temps  $t - d$ , i.e.

$$\nabla^d X_t = X_t - X_{t-d}$$

et on montre que :

$$\nabla^d X_t = (1 - B^d)X_t$$

Supposons à présent que dans l'équation (1.1) la forme du trend soit un polynôme en  $t$  d'ordre  $k$  :

$$T_t = \sum_{j=0}^k a_j t^j$$

Si on applique à  $T_t$  l'opérateur de retard  $\nabla$  on obtient :

$$\nabla T_t = T_t - T_{t-1} = \sum_{j=0}^k a_j t^j - \sum_{j=0}^k a_j (t-1)^j$$

## 1.1. LE PRINCIPE CLASSIQUE D'ANALYSE DES SÉRIES CHRONOLOGIQUES

---

et, en développant  $(t - 1)^j$  dans le membre de droite, il vient :

$$\sum_{j=0}^k a_j (t - 1)^j = \sum_{j=0}^k a_j t^j + \sum_{j=1}^k b_{j-1} t^{j-1}$$

D'où,

$$\nabla T_t = \sum_{j=1}^k b_{j-1} t^{j-1} = \sum_{i=0}^{k-1} b_i t^i \quad \text{avec } i = j - 1$$

On voit que l'opération à l'aide de  $\nabla$  fait diminuer d'une unité l'ordre du polynôme  $T_t$ . Dès lors, si on applique  $k$  fois l'opérateur  $\nabla$  au polynôme  $T_t$  on obtient  $\nabla^k T_t = k! a_k$  où  $a_k$  est une constante indépendante de  $t$ . Donc, l'opérateur  $\nabla^k$  appliqué à l'équation (1.1), nous donne :

$$\nabla^k X_t = \text{constante} + \nabla^k Z_t$$

En pratique, on ignore le degré du polynôme  $T_t$ . C'est pourquoi on applique l'opérateur  $\nabla$  à la série observée autant de fois qu'il est nécessaire pour éliminer la tendance. Et donc la série obtenue est purement aléatoire. En pratique  $k$  est assez petit ( $k \leq 3$ ).

### 1.1.2 Elimination du trend et de la saisonnalité

Supposons à présent que, en plus d'un trend  $T_t$ , la série comporte une composante saisonnière  $S_t$ . On considère donc le modèle :

$$X_t = T_t + S_t + Z_t \tag{1.4}$$

$S_t$  est une fonction périodique en  $t$  telle que  $S_t = S_{t+d}$  pour un certain  $d > 0$  et  $S_t$  ne comporte pas de composante "trend", et donc :

$$\sum_{j=i}^d S_{t+j} = 0$$

On peut également appliquer la méthode des différences pour éliminer à la fois le trend et la saisonnalité. Etant donné le modèle (1.4), appliquons l'opérateur de différence d'ordre  $d$ ,  $\nabla_d$ , où  $d$  est la période de  $S_t$ , à la série observée  $X_t$  :

$$\begin{aligned} \nabla_d X_t &= X_t - X_{t-d} \\ &= T_t + S_t + Z_t - T_{t-d} - S_{t-d} - Z_{t-d} \\ &= (T_t - T_{t-d}) + (S_t - S_{t-d}) + (Z_t - Z_{t-d}) \end{aligned}$$

où, par définition, le terme  $S_t - S_{t-d}$  est nul.

La série chronologique  $\nabla_d X_t$  est donc désaisonnalisée, mais elle comporte encore un trend  $\nabla_d T_t$ . Pour éliminer la tendance restante, on peut à nouveau utiliser la méthode des différences.

## 1.2 Rappel sur les processus aléatoires

**Définition 1.2.1** *Un processus stochastique  $\{X_t\}_{t \in T}$  est une famille de variables aléatoires définies sur un espace probabilisé  $(\Omega, \mathcal{F}, P)$ .*

**Définition 1.2.2** *Soit  $\{X_t, t \in \mathbb{Z}\}$  un processus tel que  $\text{var}(X_t) < \infty$ . La fonction d'autocovariance  $\gamma(\cdot, \cdot)$  de  $\{X_t\}$  est définie telle que :*

$$\gamma_X(r, s) = \text{cov}(X_r, X_s) = E[(X_r - EX_r)(X_s - EX_s)] \quad r, s \in T$$

**Définition 1.2.3** *Le processus  $\{X_t, t \in \mathbb{Z}\}$  où  $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$ , est dit stationnaire si*

i)  $E|X_t|^2 < \infty$  pour tout  $t \in \mathbb{Z}$

ii)  $EX_t = m$  pour tout  $t \in \mathbb{Z}$

si  $\{X_t, t \in \mathbb{Z}\}$  est stationnaire alors  $\gamma(r, s) = \gamma(r - s, 0)$  pour tout  $r, s \in \mathbb{Z}$  :

Pour tout processus stationnaire  $\gamma_j = \gamma_{-j}, \forall j$ .

**Remarque 1.2.1** *Le processus  $\{X_t, t \in \mathbb{Z}\}$  est dit strictement stationnaire si les lois de distributions jointes de  $(X_{t_1}, X_{t_2}, \dots, X_{t_k})$  et  $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h})$  sont la même pour tous  $t_1, t_2, \dots, t_k, k \in \mathbb{Z}$ .*

**Définition 1.2.4** *Soit  $\{X_t\}$  un processus stationnaire. L'autocorrélation de  $X_t$  est définie telle que :*

$$\rho_X(h) = \gamma_X(h) / \gamma_X(0) = \text{corr}(X_{t+h}, X_t) \quad \text{pour tout } t, h \in \mathbb{Z}$$

## 1.3 Modèle linéaire général

Dans cette section nous allons rappeler quelques définitions et propriétés d'un modèle linéaire général pour un processus aléatoire en nous basant sur la définition de Yule ([32]).



**Définition 1.3.1** *Le processus  $\{Z_t\}$  est appelé bruit blanc, et est noté  $Z_t \sim WN(0, \sigma^2)$ , si les trois propriétés sont vérifiées :*

- i)  $E(Z_t) = 0 \quad \forall t \in \mathbb{Z}$
- ii)  $E(Z_t^2) = \sigma^2$
- iii)  $cov(Z_t, Z_s) = 0 \quad \forall s \neq t$

**Définition 1.3.2** *Soit  $X_t$  un processus aléatoire. L'opérateur  $B$  est appelé opérateur de recul et est défini tel que :*

$$B^j X_t = X_{t-j}$$

*L'opération dual de  $B$ , noté  $F$ , est appelé opérateur d'avance et est défini tel que :*

$$F^j X_t = X_{t+j}$$

*Il est évident que  $F = B^{-1}$ .*

Les processus stochastiques que nous utilisons sont basés sur l'idée de Yule ([32]) qui stipule que tout processus aléatoire est le résultat à travers un filtre linéaire d'un processus de bruits blancs i.e. une somme pondérée de la valeur présente et des valeurs passées de  $\{Z_t\}$ .

Dans la suite nous allons supposer que les processus utilisés sont centrés sans perdre en généralité.

**Définition 1.3.3** *Soit  $\{X_t\}_{t \in \mathbb{Z}}$  un processus aléatoire centré et soit  $\{Z_t\}_{t \in \mathbb{Z}}$  un processus de bruit blanc. Alors  $X_t$  peut être exprimé tel que :*

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j} = \left( \sum_{j=0}^{\infty} \psi_j B^j \right) Z_t \quad (1.5)$$

où  $\psi(B) = \left( \sum_{j=0}^{\infty} \psi_j B^j \right)$  est appelé filtre linéaire ou fonction de transfert.

**Remarque 1.3.1** *Le filtre  $\psi(B)$  est dit stable et le processus  $X_t$  stationnaire si la série  $\sum_{j=0}^{\infty} \psi_j B^j$  est convergente i.e. si  $\sum_{j=0}^{\infty} |\psi_j| < \infty$ . Autrement le processus  $X_t$  est dit non stationnaire.*

**Proposition 1.3.1** *Un processus  $X_t$  centré peut être exprimé comme une combinaison linéaire de valeurs passées de  $X_t$  et de l'innovation présente  $Z_t$  tel que :*

$$X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} + Z_t \quad (1.6)$$

**Preuve** En effet, l'équation (1.5) peut être exprimée telle que :

$$Z_t = X_t - \psi_1 Z_{t-1} - \psi_2 Z_{t-2} - \dots$$

et en remplaçant  $Z_{t-1}$  dans l'équation ci-dessus par :

$$Z_{t-1} = X_{t-1} - \psi_1 Z_{t-2} - \psi_2 Z_{t-3} - \dots$$

On obtient :

$$\begin{aligned} Z_t &= X_t - \psi_1 (X_{t-1} - \psi_1 Z_{t-2} - \psi_2 Z_{t-3} - \dots) - \psi_2 Z_{t-2} - \dots \\ &= X_t - \psi_1 X_{t-1} - \psi_1^2 \psi_2 Z_{t-2} - \psi_1 \psi_3 Z_{t-3} - \dots \end{aligned}$$

Et on procède de la même manière pour  $Z_{t-2}$ , et ainsi de suite pour  $Z_{t-3}$ , etc. Finalement, on aboutit à l'équation (1.6). ■

**Remarque 1.3.2** L'équation (1.6) peut être exprimée telle que :

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} = \left( \sum_{j=0}^{\infty} \pi_j B^j \right) X_t$$

où  $\pi(B) = \left( \sum_{j=0}^{\infty} \pi_j B^j \right)$ .

### 1.3.1 Relation entre les poids $\psi$ et les poids $\pi$

Reprenons l'équation (1.6)

$$\pi(B) X_t = Z_t$$

En multipliant l'équation ci-dessus par la fonction de transfert  $\psi(B)$  on obtient :

$$\psi(B) \pi(B) X_t = \psi(B) Z_t = X_t$$

D'où  $\psi(B) \pi(B) = 1$ , car  $\psi(B) \pi(B) X_t = X_t, \forall X_t$ . Et par conséquent, on déduit la relation entre les poids  $\psi$  et les poids  $\pi$  telle que :

$$\pi(B) = \psi^{-1}(B) \tag{1.7}$$

### 1.3.2 La fonction génératrice des autocovariances d'un processus linéaire

La fonction génératrice des autocovariances d'un processus linéaire est un outil utile pour le calcul de certaines caractéristiques des séries chronologiques. Puisque nous disposons des deux formes d'écriture d'un processus aléatoire, la forme autorégressive infinie et la forme moyenne mobile infinie. Il est intéressant d'établir une relation entre les autocovariances et les poids de ces deux formes.

En effet, si  $\{X_t\}$  est un processus stationnaire ayant pour fonction d'autocovariance  $\gamma(\cdot)$ , alors sa fonction génératrice d'autocovariance est définie par

$$\gamma(B) = \sum_{k=-\infty}^{\infty} \gamma_k B^k \quad (1.8)$$

On peut remarquer que cette fonction existe si la série converge pour tout  $B$ ,  $r < B < r$  où  $r > 1$  ([8]).

Il est souvent facile de calculer la fonction génératrice. Dans ce cas l'autocovariance  $\gamma_k$  peut être déterminée par identification des coefficients de  $B^k$  ou de  $B^{-k}$ . Il est clair qu'un processus  $X_t$  est un bruit blanc si et seulement si sa fonction génératrice d'autocovariance  $\gamma(B)$  est constante pour tout  $B$ . Pour exprimer la fonction génératrice des autocovariances on doit d'abord calculer  $\gamma_k$ . En effet,

$$\gamma_k = E[X_t X_{t+k}] = \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} \psi_j \psi_h E(Z_{t-j} Z_{t+k-h})$$

Par ailleurs, on sait que :

$$E(Z_{t-j} Z_{t+k-h}) = \begin{cases} \sigma_Z^2 & \text{si } h = k + j \\ 0 & \text{si } h \neq k + j \end{cases}$$

D'où

$$\gamma_k = E[X_t X_{t+k}] = \sigma_Z^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}$$

Par suite, en substituant dans la relation (1.8)  $\gamma_k$  par sa valeur, on obtient :

$$\begin{aligned}\gamma(B) &= \sigma_Z^2 \sum_{k=-\infty}^{+\infty} \sum_{j=0}^{\infty} \psi_j \psi_{j+k} B^k \\ &= \sigma_Z^2 \sum_{j=0}^{+\infty} \sum_{k=-j}^{\infty} \psi_j \psi_{j+k} B^k\end{aligned}$$

Sachant que  $\psi_h = 0$  pour  $h < 0$ , et en posant  $h = j + k$  on aura :

$$\gamma(B) = \sigma_Z^2 \sum_{j=0}^{+\infty} \sum_{h=0}^{\infty} \psi_j \psi_h B^{h-j}$$

D'où le résultat

$$\gamma(B) = \sigma_Z^2 \sum_{h=0}^{\infty} \psi_h B^h \sum_{j=0}^{\infty} \psi_j B^{-j}$$

Que l'on peut écrire tel que

$$\gamma(B) = \sigma_Z^2 \psi(B) \psi(B^{-1}) = \sigma_Z^2 \psi(B) \psi(F) \quad (1.9)$$

### 1.3.3 La fonction d'autocorrélation partielle

La fonction d'autocorrélation partielle d'un processus aléatoire, comme sa fonction d'autocorrélation, véhicule une quantité d'information essentielle sur sa structure de dépendance.

**Définition 1.3.4** *La fonction d'autocorrélation partielle (pacf)  $\alpha(k)$  d'un processus stationnaire est définie par*

$$\alpha(1) = \text{corr}(X_2, X_1) = \rho(1)$$

et

$$\alpha(k) = \text{corr}(X_{k+1} - P_{\overline{sp}\{1, X_2, \dots, X_k\}} X_{k+1}, X_1 - P_{\overline{sp}\{1, X_2, \dots, X_k\}} X_1)$$

où les projections  $P_{\overline{sp}\{1, X_2, \dots, X_k\}}$  est défini tel que :

$$P_{\overline{sp}\{1, X_2, \dots, X_k\}} X_{k+1} = \sum_{i=0}^k \alpha_i X_i \quad X_0 = 1$$

où  $\alpha_0, \dots, \alpha_n$  sont une base du sous-espace  $\overline{sp}\{1, X_2, \dots, X_k\}$ .

Il est possible de définir l'autocorrélation partielle autrement. Soit  $\{X_t\}$  un processus stationnaire de moyenne nulle et de fonction d'autocovariance  $\gamma(\cdot)$  tel que  $\gamma(h) \rightarrow 0$  quand  $h \rightarrow \infty$ . Supposons que  $\phi_{kj}$ ,  $j = 1, \dots, k$ ,  $k = 1, 2, \dots$  sont les coefficients de la régression de  $X_{k+1}$  sur  $X_1, \dots, X_k$ , i.e.

$$P_{\overline{sp}\{1, X_2, \dots, X_k\}} X_{k+1} = \sum_{j=1}^k \phi_{kj} X_{k+1-j}$$

Alors à partir de l'équation

$$\langle X_{k+1} - P_{\overline{sp}\{1, X_2, \dots, X_k\}} X_{k+1}, X_j \rangle = 0, \quad j = 1, \dots, k$$

On obtient la fonction d'autocorrélation pour des  $X_j$  pour  $j = 1, 2, \dots, k$  :

$$\rho_j = \phi_{k1} \rho_{j-1} + \dots + \phi_{k(k-1)} \rho_{j-k+1} + \phi_{kk} \rho_{j-k} \quad j = 1, 2, \dots, k \quad (1.10)$$

Le système d'équations(1.10) est le système de Yule-Walker ([32], [?]) que l'on peut écrire sous la forme matricielle telle que :

$$P_k \phi_k = \rho^{(k)} \quad (1.11)$$

où

$$P_k = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & 1 \end{bmatrix}, \quad \phi_k = \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} \quad \text{et} \quad \rho^{(k)} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix}$$

A l'aide de la relation (1.11) on peut calculer les  $\phi_{kk}$  tel que

$$\phi_{11} = \rho_1$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

**Définition 1.3.5** L'autocorrélation partielle  $\alpha(k)$  de  $\{X_t\}$  à l'horizon  $k$  est

$$\alpha(k) = \phi_{kk} \quad k \geq 1$$

où  $\phi_{kk}$  est déterminé par l'équation (1.11)

## 1.4 Le périodogramme d'un processus

Soit  $X_1, X_2, \dots, X_n$  une observation d'un processus aléatoire  $X_t$ . Le périodogramme est généralement utilisé comme estimateur de la densité spectrale du processus  $X_t$ . Il est souvent utilisé comme outil d'identification d'un modèle stochastique pour une série chronologique.

Pour illustrer le calcul du périodogramme on va supposer que le nombre  $n$  d'observations est impair. Il a été établi qu'un processus  $X_t$  peut être modélisé sous forme d'une série de Fourier telle que :

$$X_t = a_0 + \sum_{i=1}^q (a_i c_{it} + b_i s_{it}) + Z_t$$

où  $c_{it} = \cos(2\pi\lambda_i t)$ ,  $s_{it} = \sin(2\pi\lambda_i t)$  et  $\lambda_i = \frac{i}{N}$  est le  $i^{\text{ème}}$  harmonique de la fréquence fondamentale  $\frac{1}{N}$ .

Les estimateurs des moindres carrés des coefficients  $a_0$  et  $(a_i, b_i)$ ,  $i = 1, 2, \dots, q$  sont déterminés tels que :

$$a_0 = \bar{X} \tag{1.12}$$

$$\begin{cases} a_i = \frac{2}{N} \sum_{t=1}^N X_t c_{it} \\ b_i = \frac{2}{N} \sum_{t=1}^N X_t s_{it} \end{cases} \quad i = 1, 2, \dots, q \tag{1.13}$$

**Définition 1.4.1** *Le périodogramme est l'ensemble des  $q = \frac{(N-1)}{2}$  valeurs  $I_N(\lambda_i)$ ,  $i = 1, 2, \dots, q$  définies telles que :*

$$I_N(\lambda_i) = \frac{N}{2} (a_i^2 + b_i^2) \tag{1.14}$$

où  $I_N(\lambda_i)$  est appelé l'intensité de la fréquence  $\lambda_i$ .

**Remarque 1.4.1** *Si  $N$  est pair, i.e.  $N = 2q$ , les équations (1.12), (1.13) et (1.14) restent valables pour  $i = 1, 2, \dots, (q-1)$  mais*

$$a_q = \frac{1}{N} \sum_{t=1}^N (-1)^t X_t$$

$$b_q = 0$$

#### 1.4. LE PÉRIODOGRAMME D'UN PROCESSUS

---

De plus

$$I_N(\lambda_q) = I(0, 5) = Na_q^2$$

Il est clair que la fréquence la plus haute est 0.5 cycle par intervalle du temps. La définition du periodogramme peut être reformulée telle que :

$$I_N(\lambda) = \frac{N}{2} (a_\lambda^2 + b_\lambda^2) \quad \text{où} \quad 0 \leq \lambda \leq \frac{1}{2} \quad (1.15)$$

Alors,  $I_N(\lambda)$  est appelée spectre empirique.

Le spectre empirique ainsi défini fournit un point de départ pour la théorie de l'analyse spectrale des séries chronologiques.

**Proposition 1.4.1** *Le spectre empirique  $I_N(\lambda)$  et la fonction d'autocovariance  $c_k$  sont liés par la relation :*

$$I_N(\lambda) = 2 \left[ c_0 + 2 \sum_{k=1}^{N-1} c_k \cos(2\pi\lambda k) \right] \quad 0 \leq \lambda \leq \frac{1}{2}$$

**Preuve** En effet, en écrivant la relation (1.15) telle que :

$$I_N(\lambda) = \frac{N}{2} (a_\lambda - ib_\lambda)(a_\lambda + ib_\lambda) = \frac{N}{2} d_\lambda d_\lambda^*$$

où  $d_\lambda^*$  est le conjugué de  $d_\lambda$ .

Par ailleurs d'après ([6])

$$d_\lambda = \frac{2}{N} \sum_{t=1}^N X_t [\cos(2\pi\lambda t) - i \sin(2\pi\lambda t)] = \frac{2}{N} \sum_{t=1}^N X_t e^{-i2\pi\lambda t}$$

Donc, d'après (1.13) on a :

$$d_\lambda = \frac{2}{N} \sum_{t=0}^N (X_t - \bar{X}) e^{-i2\pi\lambda t}$$

Par conséquent :

$$I_N(\lambda) = \frac{2}{N} \sum_{t=0}^N \sum_{t'=0}^N (X_t - \bar{X})(X_{t'} - \bar{X}) e^{-i2\pi\lambda(t-t')}$$

En mettant  $k = t - t'$  et

$$c_k = \frac{1}{N} \sum_{t=0}^{N-k} (X_t - \bar{X}) (X_{t+k} - \bar{X})$$

Donc, on obtient

$$I_N(\lambda) = 2 \left[ c_0 + 2 \sum_{k=0}^{N-1} c_k \cos(2\pi\lambda k) \right] \quad 0 \leq \lambda \leq \frac{1}{2}$$

Donc le périodogramme est la transformée de Fourier de la fonction d'auto-covariance empirique. ■

Par ailleurs, nous pouvons calculer la valeur moyenne de  $I_N(\lambda)$  telle que :

$$E[I_N(\lambda)] = 2 \left[ E(c_0) + 2 \sum_{k=0}^{N-1} E[c_k] \cos(2\pi\lambda k) \right] \quad (1.16)$$

où

$$\lim_{N \rightarrow \infty} E[c_k] = \gamma_k$$

En prenant la limite de (1.16) quand  $N$  tend vers l'infini, le spectre  $I(\lambda)$  sera défini par :

$$I(\lambda) = \lim_{N \rightarrow \infty} E[I_N(\lambda)] = 2 \left[ \gamma_0 + 2 \sum_{k=0}^{\infty} \gamma_k \cos(2\pi\lambda(k)) \right] \quad 0 \leq \lambda \leq \frac{1}{2}$$

**Remarque 1.4.2** *Il est intéressant de noter que :*

$$|I(\lambda)| \leq 2 \left[ |\gamma_0| + 2 \sum_{k=1}^{\infty} |\gamma_k| |\cos(2\pi\lambda k)| \right] \leq 2 \left( |\gamma_0| + 2 \sum_{k=1}^{\infty} |\gamma_k| \right)$$

où

$$\gamma_0 = \sigma_X^2 = \int_0^{\frac{1}{2}} I(\lambda) d\lambda$$

**Définition 1.4.2** *La densité spectrale d'un processus est donnée par :*

$$f(\lambda) = \frac{I(\lambda)}{\sigma_X^2} = 2 \left[ 1 + 2 \sum_{k=1}^{\infty} \lambda_k \cos(2\pi\lambda k) \right], \quad 0 \leq \lambda \leq \frac{1}{2}$$



**Remarque 1.4.3** *Du fait que la densité spectrale  $f(\lambda)$  est une densité de probabilité, alors elle vérifie la propriété suivante :*

$$\int_0^{\frac{1}{2}} f(\lambda) d\lambda = 1$$

### 1.4.1 Le périodogramme d'un processus linéaire stationnaire

On suppose que  $B = e^{-i2\pi\lambda}$ , dans l'équation de fonction génératrice d'autocovariance (1.9), nous obtenons le périodogramme d'un processus linéaire :

$$\begin{aligned} I(\lambda) &= 2\sigma_Z^2 \boldsymbol{\psi}(e^{-i2\pi\lambda}) \boldsymbol{\psi}(e^{-i2\pi\lambda}) & (1.17) \\ &= 2\sigma_Z^2 |\boldsymbol{\psi}(e^{-i2\pi\lambda})|^2 & 0 \leq \lambda \leq \frac{1}{2} \end{aligned}$$

Dans ce qui suit, on va étudier des cas particuliers de processus linéaires généraux. On pourra utiliser ces propositions pour trouver leur fonction d'autocovariance et leur densité spectrale.

# Chapitre 2

## Modèles de Box-Jenkins

Les modèles de Box et Jenkins sont des modèles linéaires qui peuvent être stationnaires ou non stationnaires. Les modèles stationnaires sont regroupés dans la famille des modèles *ARMA*, qui est composée des modèles Autorégressifs (*AR*), des modèles en Moyennes Mobiles (*MA*) et des modèles mixtes Autorégressifs-Moyennes-Mobiles (*ARMA*).

### 2.1 Le modèle *AR*

**Définition 2.1.1** Soit  $X_t$  un processus stochastique stationnaire, on dit que  $X_t$  est un processus autorégressif d'ordre  $p$  noté  $AR(p)$  le processus qui vérifie l'équation suivante :

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + Z_t$$

où  $Z_t \sim WN(0, \sigma^2)$

**Remarque 2.1.1** Le polynôme  $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$  est le polynôme caractéristique du processus autorégressif. Le modèle s'écrit souvent  $\phi(B) X_t = Z_t$ .

#### 2.1.1 Principales caractéristiques d'un modèle $AR(p)$

**La fonction d'autocorrélation** La fonction d'autocovariance d'un processus autorégressif  $AR(p)$  est donnée par l'équation de récurrence suivante :

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \quad k = 1, 2, \dots, p \quad (2.1)$$

En divisant l'équation (2.1) par  $\gamma_0$ , on obtient la fonction d'autocorrélation satisfaisante à l'équation de récurrence suivante :

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p} \quad k = 1, 2, \dots, p \quad (2.2)$$

L'équation (2.2) peut être exprimée alors telle que :

$$\phi(B) \rho_k = 0$$

Les équations (2.2),  $k = 1, 2, \dots, p$  sont aussi appelées équations de Yule-Walker([32], [29]).

Par ailleurs, on sait que la solution générale de l'équation (2.2) est telle que :

$$\rho_k = A_1 G_1^k + A_2 G_2^k + \cdots + A_p G_p^k \quad (2.3)$$

où  $G_i^k, i = 1, 2, \dots, p$  sont les racines du polynôme caractéristique de l'équation de récurrence (2.2).

De plus, le polynôme caractéristique de l'équation de récurrence (2.2) est le polynôme dual du polynôme  $\phi(B)$ . Or, pour un processus stationnaire les racines de  $\phi(B)$  sont à l'extérieur du cercle unité, donc les racines du polynôme caractéristique de l'équation (2.3) sont à l'intérieur du cercle unité. Finalement, en remarquant que  $\max \{G_1, G_2, \dots, G_p\} < 1$ , il vient que :

$$|\rho_k| \leq (A \max \{G_1, G_2, \dots, G_p\})^k$$

où  $A = A_1 + A_2 + \cdots + A_p$  est une constante.

**Remarque 2.1.2** *Nous constatons que la fonction d'autocorrélation  $\rho_k$  d'un processus  $AR(p)$  décroît exponentiellement vers zéro. Cette décroissance exponentielle est lisse si  $\rho_k$  est positive et oscillante si le signe de  $\rho_k$  alterne.*

**La fonction d'autocorrélation partielle** Considérons l'équation (1.10) pour  $k = 1, 2, \dots, k$ . On obtient un système d'équations linéaires en  $\rho_1, \rho_2, \dots, \rho_k$  de coefficients  $\phi_{k1}, \phi_{k2}, \dots, \phi_{kk}$ . Alors, la fonction d'autocorrélation partielle  $\alpha(k)$  d'un processus  $AR(p)$  est définie telle que  $\alpha(k) = \phi_{kk}$  où  $\phi_{kk}$  est le dernier coefficient de la régression de  $X_{k+1}$  sur  $X_k, X_{k-1}, \dots, X_1$  et déterminé à l'aide de l'équation (1.11).

Pour un processus  $AR(p)$  :

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_p X_{t-p} = Z_t \quad Z_t \sim WN(0, \sigma^2)$$

Nous avons pour  $k > p$ ,

$$P_{\overline{sp}\{X_2, \dots, X_k\}} X_{k+1} = \sum_{j=1}^p \phi_j X_{k+1-j} \quad (2.4)$$

Donc si  $Y \in \overline{sp}\{X_2, \dots, X_k\}$  alors par la causalité  $Y \in \overline{sp}\{Z_j, j \leq k\}$  tel que

$$\langle X_{k+1} - \sum_{j=1}^p \phi_j X_{k+1-j}, Y \rangle = \langle Z_{k+1}, Y \rangle = 0$$

Pour  $k > p$  nous allons calculer à partir de (2.4) :

$$\begin{aligned} \alpha(k) &= \text{Corr} \left( X_{k+1} - \sum_{j=1}^p \phi_j X_{k+1-j}, X_1 - P_{\overline{sp}\{X_2, \dots, X_k\}} X_1 \right) \\ &= \text{Corr} (Z_{k+1}, X_1 - P_{\overline{sp}\{X_2, \dots, X_k\}} X_1) = 0 \end{aligned}$$

Pour  $k \leq p$  les valeurs de  $\alpha(k)$  sont calculées d'après la définition (1.3.5).

**Remarque 2.1.3** *Nous constatons que la fonction d'autocorrélation partielle  $\alpha(k)$  d'un processus  $AR(p)$  est égale à zéro quand  $k > p$ . Autrement dit la fonction d'autocorrélation partielle d'un processus autorégressif chute à zéro au delà de  $p$ .*

**Le périodogramme** Pour le processus  $AR(p)$ ,  $\psi(B) = \phi^{-1}(B)$  et  $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$  et en utilisant le périodogramme du processus linéaire(1.17), alors on obtient l'équation suivante :

$$I(\lambda) = \frac{2\sigma_Z^2}{[1 - \phi_1 e^{-i2\pi\lambda} - \phi_2 e^{-i4\pi\lambda} - \dots - \phi_p e^{-2\pi p\lambda}]^2} \quad 0 \leq \lambda \leq \frac{1}{2}$$

## 2.2 Le modèle $MA$

**Définition 2.2.1** *Soit  $X_t$  un processus stochastique stationnaire, on dit que  $X_t$  est un processus moyenne mobile d'ordre  $q$  noté  $MA(q)$  le processus qui vérifie l'équation suivante :*

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \dots + \theta_q Z_{t-q} \quad (2.5)$$

où  $Z_t \sim WN(0, \sigma^2)$ .

**Remarque 2.2.1** Un tel processus est donc linéaire pour la suite sommable :

$$\psi_j = 0, \quad j < 0, \quad \psi_0 = 1, \quad \psi_1 = \theta_1, \dots, \psi_q = \theta_q, \quad \psi_j = 0, \quad j > q.$$

On écrit habituellement  $X_t = \boldsymbol{\theta}(B) Z_t$  où  $\boldsymbol{\theta}(B)$  est le polynôme

$$\boldsymbol{\theta}(B) = 1 + \theta_1 B + \dots + \theta_q B^q$$

et  $B$  est l'opérateur de décalage arrière (appelé aussi opérateur retard). Le polynôme  $\boldsymbol{\theta}(B)$  est appelé polynôme caractéristique du processus moyenne mobile.

Il est clair qu'un processus moyenne mobile est causal ; cependant, il n'est pas toujours inversible.

### 2.2.1 Principales caractéristiques d'un modèle $MA(q)$

**La fonction d'autocorrélation** Le calcul de la fonction d'autocovariance d'un processus  $MA(q)$  est obtenu tel que :

$$\gamma_k = E(X_t X_{t-k}) = E[(Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q})(Z_{t-k} + \theta_1 Z_{t-k-1} + \dots + \theta_q Z_{t-k-q})]$$

et donc

$$\gamma_k = \begin{cases} (\theta_k + \theta_1 \theta_{k+1} + \theta_2 \theta_{k+2} + \dots + \theta_{q+k} \theta_q) \sigma_Z^2 & k = 1, 2, \dots, q \\ 0 & k > q \end{cases}$$

La variance du processus  $X_t$  est donnée par la formule suivante :

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_Z^2$$

Par conséquent, la fonction d'autocorrélation est définie telle que :

$$\rho_k = \begin{cases} \frac{\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q+k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} & k = 1, 2, \dots, q \\ 0 & k > q \end{cases} \quad (2.6)$$

**Remarque 2.2.2** Nous constatons que la fonction d'autocorrélation  $\rho_k$  d'un processus  $MA(q)$  est égale à zéro quand  $k > q$ . Autrement dit la fonction d'autocorrélation d'un processus moyenne mobile chute à zéro au delà de  $q$ .

**La fonction d'autocorrélation partielle** Contrairement aux équations linéaires de Yule-Walker([32], [29]) (1.10), les  $q$  équations obtenues pour le processus  $MA(q)$  sont non linéaires. En effet, pour un processus moyenne mobile  $MA(1)$  défini tel que :

$$X_t = Z_t + \theta Z_{t-1}, \quad |\theta| < 1, \quad Z_t \sim WN(0, \sigma^2)$$

L'autocorrélation partielle d'ordre 1 est telle que :

$$\alpha(1) = \rho(1) = \frac{\theta}{(1 + \theta^2)}$$

De plus, on sait que  $P_{\overline{sp}\{X_2\}}X_3 = \frac{\theta}{(1+\theta^2)}X_2 = P_{\overline{sp}\{X_2\}}X_1$ , alors le calcul de  $\alpha(2)$  nous donne :

$$\alpha(2) = \text{Corr} \left( X_3 - \theta(1 + \theta^2)^{-1} X_2, X_1 - \theta(1 + \theta^2)^{-1} X_2 \right) = \frac{-\theta^2}{(1 + \theta^2 + \theta^4)}$$

Et ainsi de suite,  $\alpha(k)$  sera telle que :

$$\alpha(k) = -\frac{(-\theta)^k (1 - \theta^2)}{1 - \theta^{2(k+1)}}$$

**Remarque 2.2.3** *Nous constatons que la fonction d'autocorrélation partielle  $\alpha(k)$  d'un processus  $MA(q)$  décroît exponentiellement vers zéro.*

**Le périodogramme** Pour le processus  $MA(q)$  on a :

$$\psi(B) = \theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

Donc, en utilisant l'équation (1.17) le périodogramme d'un processus  $MA(q)$  est obtenu tel que :

$$I(\lambda) = 2\sigma_Z^2 \left| 1 - \theta_1 e^{-i2\pi\lambda} - \theta_2 e^{-i4\pi\lambda} - \dots - \theta_q e^{-i2\pi q\lambda} \right|^2 \quad 0 \leq \lambda \leq \frac{1}{2}$$

## 2.3 Le modèle ARMA(p, q)

**Définition 2.3.1** On dit qu'un processus stationnaire et centré  $X_t$  est un processus ARMA(p, q) si pour chaque  $t$  il vérifie la relation :

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \quad (2.7)$$

où  $Z_t \sim WN(0, \sigma^2)$ .

**Remarque 2.3.1** En utilisant l'opérateur  $B$ , l'équation (2.7) peut être exprimée telle que :

$$\phi(B) X_t = \theta(B) Z_t$$

où  $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$  et  $\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$  sont appelés polynômes caractéristiques respectifs des parties autorégressives et moyennes mobiles du processus.

### 2.3.1 Principales caractéristiques d'un modèle ARMA(p, q)

**La fonction d'autocorrélation** En multipliant les deux membres de l'équation (2.7) par  $X_{t-k}$  nous constatons que la fonction d'autocovariance satisfait l'équation de différences :

$$\gamma_k = \phi_1 \gamma_{k-1} + \cdots + \phi_p \gamma_{k-p} + \gamma_{XZ}(k) - \theta_1 \gamma_{XZ}(k-1) - \cdots - \theta_q \gamma_{XZ}(k-q)$$

où  $\gamma_{XZ}(k)$  est la fonction de covariance de  $X_t$  et de  $Z_t$  définie par  $\gamma_{XZ}(k) = E[X_{t-k} Z_t]$ . Comme  $X_{t-k}$  dépend seulement des innovations avant l'instant  $t-k$  par la représentation en moyenne mobile infinie :

$$X_{t-k} = \psi(B) Z_{t-k} = \sum_{j=0}^{\infty} \psi_j Z_{t-k-j}$$

Il suit

$$\gamma_{XZ}(k) = \begin{cases} 0 & k < 0 \\ \varphi_k \sigma_Z^2 & k \geq 0 \end{cases}$$

Dorénavant l'équation précédente pour  $\gamma_k$  peut être exprimée comme suit :

$$\gamma_k = \phi_1 \gamma_{k-1} + \cdots + \phi_p \gamma_{k-p} - \sigma_Z^2 (\theta_k \psi_0 + \theta_{k+1} \psi_1 + \cdots + \theta_q \varphi_{q-k}) \quad (2.8)$$

Avec le paramètre  $\theta_0 = -1$ .

Nous voyons que (2.8) implique :

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p} \quad k \geq q + 1$$

La variance égale à :

$$\gamma_0 = \phi_1 \gamma_1 + \cdots + \phi_p \gamma_p + \sigma_Z^2 (1 - \theta_1 \psi_1 - \cdots - \theta_q \varphi_q) \quad (2.9)$$

Et dorénavant :

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p} \quad k \geq q + 1$$

Ou bien

$$\phi(B) \rho_k = 0 \quad k \geq q + 1$$

Par exemple le processus  $ARMA(1, 1)$  :

$$X_t - \phi_1 X_{t-1} = Z_t - \theta_1 Z_{t-1}$$

telle que :

$$(1 - \phi_1 B) X_t = (1 - \theta_1 B) Z_t$$

A partir de l'équation (2.8) et (2.9) nous obtenons :

$$\begin{aligned} \gamma_0 &= \phi_1 \gamma_1 + \sigma_Z^2 (1 - \theta_1 \psi_1) \\ \gamma_1 &= \phi_1 \gamma_0 - \theta_1 \sigma_Z^2 \\ \gamma_k &= \phi_1 \gamma_{k-1} \quad k \geq 2 \end{aligned}$$

Avec  $\psi_1 = \phi_1 - \theta_1$ .

Par conséquent, la résolution des deux premières équations en  $\gamma_0$  et  $\gamma_1$  ci-dessus, la fonction d'autocovariance du processus est déterminée telle que :

$$\begin{aligned} \gamma_0 &= \frac{1 + \theta_1^2 - 2\phi_1 \theta_1}{1 - \phi_1^2} \sigma_Z^2 \\ \gamma_1 &= \frac{(1 - \phi_1 \theta_1)(\phi_1 - \theta_1)}{1 - \phi_1^2} \sigma_Z^2 \\ \gamma_k &= \phi_1 \gamma_{k-1} \quad k \geq 2 \end{aligned}$$

Il est clair que  $\rho_k$  vérifie l'équation de récurrence  $\rho_k = \phi_1 \rho_{k-1}$ ,  $k \geq 2$ , alors  $\rho_k = \phi_1^{k-1} \rho_1$ ,  $k > 1$ . Ainsi, la fonction d'autocorrélation décroît exponentiellement quand  $k \rightarrow \infty$ . Cette décroissance exponentielle est lisse si  $\phi_1$  est positif et oscillante si  $\phi_1$  est négatif. En effet, le signe de  $\rho_1$  est déterminé par le signe de  $(\phi_1 - \theta_1)$ .



**Remarque 2.3.2** La fonction d'autocorrélation  $\rho_k$  d'un processus ARMA( $p, q$ ) décroît exponentiellement vers zéro.

**La fonction d'autocorrélation partielle** Un processus ARMA( $p, q$ ) peut être exprimé tel que :

$$X_t = \boldsymbol{\theta}(B) \boldsymbol{\phi}^{-1}(B) Z_t$$

où  $\boldsymbol{\phi}^{-1}(B)$  est un polynôme de  $B$ .

Ainsi, la fonction d'autocorrélation partielle d'un processus mixte est infinie en étendue. Elle se comporte finalement comme la fonction d'autocorrélation partielle d'un processus moyenne mobile pure, selon l'ordre de la moyenne mobile et les valeurs des paramètres qu'il contient.

**Remarque 2.3.3** La fonction d'autocorrélation partielle  $\alpha(k)$  d'un processus ARMA( $p, q$ ) décroît exponentiellement vers zéro.

**Le périodogramme** On utilisant l'équation (1.17), le périodogramme du processus mixte est :

$$\begin{aligned} I(\lambda) &= 2\sigma_Z^2 \frac{|\boldsymbol{\theta}(e^{-i2\pi\lambda})|^2}{|\boldsymbol{\phi}(e\lambda)|^2} \\ &= 2\sigma_Z^2 \frac{|1 - \theta_1 e^{-i2\pi\lambda} - \dots - \theta_q e^{-i2\pi q\lambda}|^2}{|1 - \phi_1 e^{-i2\pi\lambda} - \dots - \phi_p e^{-i2\pi p\lambda}|^2} \quad 0 \leq \lambda \leq \frac{1}{2} \end{aligned}$$

## 2.4 Causalité et inversibilité des modèles ARMA

**Définition 2.4.1** Un processus ARMA( $p, q$ ) défini par l'équation  $\boldsymbol{\phi}(B) X_t = \boldsymbol{\theta}(B) Z_t$  est dit causal s'il existe une suite de constantes  $\{\xi_i\}$  telle que  $\sum_{j=0}^{\infty} |\xi_j| < \infty$  et  $X_t$  peut s'exprimer tel que :

$$X_t = \sum_{i=0}^{\infty} \xi_i Z_{t-i} \quad t = 0, \pm 1, \dots$$

**Remarque 2.4.1** La définition de la causalité d'un modèle n'est pas une propriété du processus  $\{X_t\}$  seul, mais plutôt elle exprime le rapport entre les deux processus  $\{X_t\}$  et  $\{Z_t\}$  apparaissant dans l'équation du modèle ARMA. Il a été établi par Yule ([32]) qu'un  $\{X_t\}$  est causal s'il est le résultat à travers un filtre linéaire causal d'un processus bruit blanc  $\{Z_t\}$ .

**Définition 2.4.2** Un processus  $ARMA(p, q)$  défini par l'équation  $\phi(B)X_t = \theta(B)Z_t$  est dit inversible s'il existe une suite de constantes  $\{\nu_j\}$  telle que  $\sum_{j=0}^{\infty} |\nu_j| < \infty$  et  $Z_t$  peut s'exprimer tel que :

$$Z_t = \sum_{j=0}^{\infty} \nu_j X_{t-j}, \quad t = 0, \pm 1, \dots$$

**Théorème 2.4.1** Un processus  $ARMA(p, q)$  est causal (resp. inversible) i.e. qu'il peut être représenté sous la forme  $X_t = \sum_{i=0}^{\infty} \xi_i Z_{t-i}$  où  $\sum_{i=0}^{\infty} \xi_i^2 < \infty$  ( $Z_t = \sum_{j=0}^{\infty} \nu_j X_{t-j}$  où  $\sum_{j=0}^{\infty} \nu_j^2 < \infty$ ), ssi toutes les racines de son polynôme caractéristique  $\phi(B)$  (resp.  $\theta(B)$ ) sont à l'extérieur du cercle unité. Les coefficients  $\xi_i$  (resp.  $\nu_j$ ) sont dans ce cas les coefficients de la série de Taylor de  $\xi(B) = \frac{\theta(B)}{\phi(B)}$  (resp.  $\nu(B) = \frac{\phi(B)}{\theta(B)}$ ).

**Théorème 2.4.2** a) Pour un processus  $ARMA(p, q)$  avec toutes les racines du polynôme caractéristique  $\phi(B)$  à l'extérieur du cercle unité, les coefficients  $\xi_i = \sigma^{-2} EX_t Z_{t-i}$  de la représentation causale  $X_t = \sum_{i=0}^{\infty} \xi_i Z_{t-i}$  satisfont à l'équation de récurrence suivante :

$$\xi_k = \begin{cases} 1 & \text{si } k = 0 \\ \theta_k + \sum_{i=1}^{\min(k,q)} \phi(i) \xi(k-i) & \text{si } 1 \leq k \leq q \\ \sum_{i=1}^{\min(k,q)} \phi(i) \xi(k-i) & \text{si } k > q \end{cases}$$

b) Pour un processus  $ARMA(p, q)$  avec toute les racines du polynôme caractéristique  $\theta(B)$  à l'extérieur du cercle unité, les coefficients  $\nu_j = \sigma^{-2} EX_t Z_{t-j}$  de la représentation inverse  $Z_t = \sum_{j=0}^{\infty} \nu_j X_{t-j}$  satisfont à l'équation de récurrence suivante :

$$\nu_k = \begin{cases} 1 & \text{si } k = 0 \\ -\phi_k + \sum_{j=1}^{\min(k,p)} \theta(j) \nu(k-j) & \text{si } 1 \leq k \leq p \\ \sum_{j=1}^{\min(k,p)} \theta(j) \nu(k-j) & \text{si } k > p \end{cases}$$

## 2.5 Modèles linéaires non-stationnaires

### 2.5.1 Modèle ARIMA

La classe des modèles non stationnaires est très utile dans le domaine de l'économétrie et les mathématiques financières. Ce genre de modèle permet de représenter des processus présentant un certain type de non-stationnarité que l'on peut éliminer par différentiation d'ordre  $d$ .

**Définition 2.5.1** *On dit qu'un processus  $X_t$  est un ARIMA (Autoregressive Integrated Moving Average) d'ordre  $(p, d, q)$  si  $\nabla^d X_t$  un processus ARMA( $p, q$ ). On modélise alors le processus  $X_t$  sous la forme :*

$$\phi(B)(1-B)^d X_t = \phi(B) \nabla^d X_t = \theta(B) Z_t$$

où le polynôme  $\phi(B)$  est de degré  $p$  et le polynôme  $\theta(B)$  est de degré  $q$ .  
On écrit que la série  $X_t$  suit un processus ARIMA( $p, d, q$ ).

**Remarque 2.5.1** *Les polynômes  $\phi(B)$  et  $\theta(B)$  sont de degrés  $p$  et  $q$  respectivement et  $\phi(B) \neq 0$  pour  $|B| \leq 1$ , alors que le polynôme  $\Phi(B)$  admet une racine multiple  $B = 1$ .*

### 2.5.2 Modèle SARIMA

Nous avons déjà introduit l'importante classe des modèles ARMA pour représenter un processus stationnaire. Les modèles SARIMA sont une généralisation de cette classe qui incorpore une vaste gamme de processus non stationnaires.

**Définition 2.5.2** *Soient  $d$  et  $D$  des entiers non-négatifs, alors un processus  $\{X_t\}$  est dit processus SARIMA si  $Y_t = (1-B)^d (1-B^s)^D X_t$  est un processus ARMA de la forme :*

$$\phi(B) \Phi(B^s) Y_t = \theta(B) \Theta(B^s) Z_t$$

où  $Z_t \sim WN(0, \sigma^2)$ .

**Remarque 2.5.2** *L'opérateur  $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$  est l'opérateur générateur d'un AR( $p$ ), L'opérateur  $\theta(B) = 1 + \theta_1 B + \dots + \theta_p B^q$  est l'opérateur générateur d'un MA( $q$ ) et où, pour la saisonnalité  $Y_t - Y_{t-s}$ ,*

l'opérateur  $\Phi(B) = 1 - \Phi_1 B - \dots - \Phi_P B^P$  est l'opérateur générateur d'un AR(P), et l'opérateur  $\Theta(B) = 1 + \Theta_1 B + \dots + \Theta_Q B^Q$  est l'opérateur générateur d'un MA(Q). On note que le processus  $\{Y_t\}$  est causal si et seulement si  $\phi(B) \neq 0$  et  $\Phi(B) \neq 0$  pour  $|B| \leq 1$ . Les valeurs de  $d$  et  $D$  sont en général inférieures à 1, alors que les valeurs de  $p$  et  $q$  sont généralement inférieures à 2.

## 2.6 Méthode de prévision

### 2.6.1 Principe de prévision

Soit  $\{X_t\}$  un processus aléatoire du second ordre défini sur  $(\Omega, P)$ . Le problème le plus couramment rencontré est la prévision des valeurs futures de ce processus. Prédire une valeur future  $X_{N+k}$  de  $\{X_t\}$  au vu d'une observation  $X = (X_1, X_2, \dots, X_N)'$ , revient à chercher une fonction déterministe  $f : \mathbb{C}^k \rightarrow \mathbb{C}$  telle que  $\widehat{X}_{N+k} = f(X_1, X_2, \dots, X_N)$ . La prédiction optimale au sens des moindres carrés consiste à choisir  $f$  de telle sorte que la moyenne du carré de l'erreur de prévision  $E \left[ \left| X_{N+k} - \widehat{X}_{N+k} \right|^2 \right]$  soit minimale. Dans l'espace de Hilbert  $\mathcal{H}_N = [X_1, X_2, \dots, X_N]$  engendré par  $\{X_1, X_2, \dots, X_N\}$ , ce problème classique admet une solution unique  $\widehat{X}_{N+k} = E^{\mathcal{H}_N}(X_{N+k}) = E(X_{N+k}/X_1, X_2, \dots, X_N) = f(X_1, X_2, \dots, X_N)$ . En d'autres termes,  $\widehat{X}_{N+k}$  est l'espérance conditionnelle de  $X_{N+k}$  lorsque l'on observe  $X_1, X_2, \dots, X_N$ . En pratique la fonction  $f$  est difficile à déterminer, soit parce que la loi jointe de  $(X_1, X_2, \dots, X_N, X_{N+k})$  est mal connue, soit parce qu'elle ne se prête pas aux calculs explicites. Ceci nous amène à restreindre la classe des estimateurs considérés. Le principe reste toujours le même, i.e. on cherchera à minimiser  $E \left[ \left| X_{N+k} - \widehat{X}_{N+k} \right|^2 \right]$  où  $\widehat{X}_{N+k} = f(X_1, X_2, \dots, X_N)$ . Mais la fonction  $f$  sera déterminée parmi les fonctions linéaires. L'unique solution est évidemment donnée par  $\widehat{X}_{N+k} = p_{roj}(X_{N+k})$  où  $p_{roj}$  est la projection orthogonale sur le sous-espace vectoriel  $\mathcal{H}_N$  engendré par  $X_1, X_2, \dots, X_N$  dans  $L^2(\Omega, P)$ .

Le principe de la prédiction linéaire revient à considérer que l'information fournie par l'observation  $X$  du processus ne dépend que de l'espace  $\mathcal{H}_N$ , sous-espace vectoriel fermé de  $L^2(\Omega, P)$  engendré par les  $X$ . Le sous-espace  $\mathcal{H}_N$  est souvent appelé l'enveloppe linéaire de  $X$ . Cela équivaut à ne considérer comme accessibles au calcul que les combinaisons linéaires des

$X_t, t = 1, 2, \dots, N$  observés.

### 2.6.2 Prédiction à l'aide d'un processus *ARMA*

Soit  $X_1, X_2, \dots, X_N$  une observation d'un processus  $\{X_t\}_{t \in \mathbb{Z}}$  obéissant à un modèle  $ARMA(p, q)$ . Rappelons d'abord les notations suivantes :

- L'instant  $N$  est appelé origine de prédiction
- La constante  $k > 0$  est l'horizon (lag) de la prédiction
- $\widehat{X}_N(k) = \widehat{X}_{N+k} = E^{\mathcal{H}_N}(X_{N+k})$  est la prédiction à l'horizon  $k$  de  $X_{N+k}$ .
- $e_N(k) = X_{N+k} - \widehat{X}_N(k)$  est l'erreur de prédiction et  $E^{\mathcal{H}_N}[e_N(k)] = 0$

Pour calculer une prédiction  $\widehat{X}_N(k)$  de la valeur future  $X_{N+k}$  de ce processus nous utiliserons les formes autorégressives infinies ou moyennes mobiles infinies, i.e.

$$X_{N+k} = \sum_{i=0}^{\infty} \pi_{Ni}^{(k)} X_{N+k-i} + Z_{t+k} \quad (2.10)$$

ou bien

$$X_{N+k} = \sum_{i=0}^{\infty} \psi_{Ni}^{(k)} Z_{t+k-i}$$

où les poids  $\pi_{Ni}^{(k)}$  sont les coefficients du polynôme  $\theta^{-1}(B)\phi(B)$  et les poids  $\psi_{Ni}^{(k)}$  sont les coefficients du polynôme  $\theta(B)\phi^{-1}(B)$ .

Sachant que  $\widehat{X}_N(k) = E^{\mathcal{H}_N}(X_{N+k})$ , il est clair que  $\widehat{X}_N(k)$  est une combinaison linéaire de valeurs passées et présentes du processus. En effet, en passant aux prévisions dans l'équation (2.10) On aura :

$$\widehat{X}_N(k) = \sum_{i=0}^{\infty} \pi_{Ni}^{(k)} [X_{N+k-i}] + [Z_{N+k}] \quad (2.11)$$

où

$$[X_N(k)] = \widehat{X}_N(k) = E^{\mathcal{H}_N}(X_{N+k}) \text{ et } [Z_{N+k}] = \widehat{Z}_N(k)$$

Toutefois, la propriété d'inversibilité du modèle *ARMA* assure que les poids  $\pi$  dans (2.11) forment une série convergente. Par conséquent, pour le calcul d'une prédiction à un lag donné, la dépendance à  $X_{N-i}$  pour  $i > k$  peut être ignorée. Dans la pratique, les poids  $\pi$  d'habitude décroissent assez rapidement, de sorte que n'importe quelle forme du modèle est employée dans le

calcul, seule une longueur moyenne de la série  $X_N, X_{N-1}, \dots, X_{N-k}$  est nécessaire pour calculer une précision suffisante. Les méthodes que nous évoquons sont modifiées facilement pour calculer les prévisions exactes de l'échantillon fini,  $E_t [X_N(k) / X_N, X_{N-1}, \dots, X_{N-k}]$ , basé sur la longueur fini des données  $X_N, X_{N-1}, \dots, X_{N-k}$ .

On peut calculer les espérances conditionnelles qui se produisent dans l'équation (2.11) que si  $i$  est un entier non-négatif on a :

$$\begin{aligned} [X_{N-i}] &= X_{N-i} & i \geq 0 \\ [X_{N+i}] &= \widehat{X}_N(i) & i \geq 0 \\ [Z_{N-i}] &= Z_{N-i} = X_{N-i} - \widehat{X}_{N-i-1}(1) & i \geq 0 \\ [Z_{N+i}] &= 0 & i \geq 0 \end{aligned}$$

**Proposition 2.6.1** *Dans le cas d'un modèle ARMA( $p, d, q$ ), la meilleure "prévision linéaire" d'origine  $N$ ,  $\widehat{X}_t(k)$ , satisfait la relation :*

$$\widehat{X}_N(k) = E^{\mathcal{H}_N}(X_{N+k}) = \sum_{i=1}^{p+d} \varphi_i \widehat{X}_N(k-i) + \sum_{i=1}^q \theta_i \widehat{Z}_{N+k-i} \quad (2.12)$$

Les inconnues  $\widehat{Z}_{N-i}$ ,  $i \geq 0$  peuvent être enlevées en utilisant la représentation inverse "π" du bruit en fonction de la série, ou en utilisant  $\widehat{Z}_N = Y_N - \widehat{Y}_{N-1}(1)$ . Une estimation arbitraire de  $Z_0, Z_{-1}, \dots$ , ou de  $X_0, X_{-1}, \dots$  sera nécessaire.

**Remarque 2.6.1** *Pour  $k > q$ , la formule (2.12) est exactement la récurrence homogène de Yule-Walker  $\varphi(B)\widehat{X}_N(k) = 0$ , donc la prévision sera donnée par la solution de cette équation qui passe par les  $p+q$  points pivots  $X_{t+q}, \dots, X_{t+q-p-d+1}$ , ajustés en tenant compte du bruit  $Z_t$ , comme indiqué en (2.12).*

# Chapitre 3

## Identification de modèles pour les séries chronologiques

### 3.1 Position du problème

Tout phénomène aléatoire ne peut être observé que ponctuellement sous forme d'une série chronologique  $X_1, X_2, \dots, X_N$ . Le problème posé est d'identifier le type de modèle qui régit ce phénomène. L'idée de base est de supposer que le modèle recherché est du type *ARIMA*. Ensuite, après avoir stationnariser le processus par différentiation, on utilisera les principales propriétés des modèles *ARMA* pour restreindre la recherche aux sous classes des modèles  $AR(p)$ ,  $MA(q)$  ou  $ARMA(p, q)$ .

En effet, il a été établi que la fonction d'autocovariance  $\gamma(k)$  ou bien la fonction d'autocorrélation  $\rho(k)$  d'un processus stationnaire  $MA(q)$  s'annule dès que  $k > q$  alors que pour un processus stationnaire  $AR(p)$  elle décroît exponentiellement quand  $k \rightarrow \infty$ . D'autre part, la fonction d'autocorrélation partielle  $\alpha(k)$  d'un processus  $MA(q)$  décroît exponentiellement vers zéro quand  $k \rightarrow \infty$  alors que la fonction d'autocorrélation partielle d'un processus  $AR(p)$  chute à zéro dès que  $k > p$ . Par contre, pour un processus  $ARMA(p, q)$  la fonction d'autocorrélation  $\rho(k)$  et la fonction d'autocorrélation partielle  $\alpha(k)$  décroissent en même temps vers zéro quand  $k \rightarrow \infty$ . Donc, nous allons essayer de déterminer l'ordre et le type de modèle qui représente le mieux la série chronologique en étudiant le comportement des fonctions  $\rho(k)$  et  $\alpha(k)$ .

## 3.2 Ajustement d'un modèle et détermination de son ordre

Les techniques d'identification sont fondamentalement empiriques. Le terme 'identification' désigne en général le choix d'une classe de modèles dans un ensemble plus vaste. En pratique l'identification concerne le choix du nombre et de signification des paramètres qui seront ensuite estimés. Après avoir utilisé les techniques de vraisemblance pour faire de l'estimation, nous allons les utiliser pour l'identification.

### 3.2.1 Méthode d'identification de Box&Jenkins

Supposons que l'on observe une série chronologique  $X_1, X_2, \dots, X_N$ . Pour ajuster un modèle stochastique *ARMA* au phénomène observé, on procède comme suit : on estime simultanément la fonction d'autocorrélation  $\rho(k)$  de la série chronologique  $X_1, X_2, \dots, X_N$  par  $\hat{\rho}(k)$ , et sa fonction d'autocorrélation partielle  $\alpha(k)$  par  $\hat{\alpha}(k)$ . Puis on étudie leurs graphes appelés respectivement corrélogramme et corrélogramme partiel.

- Si  $\hat{\rho}(k)$  décroît vers zéro et si  $\hat{\alpha}(k)$  chute à zéro à partir d'un certain rang (i.e. les valeurs de  $\hat{\alpha}(k)$  sont négligeables à partir d'un certain rang), alors on opte pour un modèle autorégressif. Pour décider de l'ordre de ce modèle *AR* on se base sur le résultat du test de l'hypothèse  $H_0 : \hat{\alpha}(k) = 0, \text{ pour } k > p$  contre  $H_1 = \overline{H_0}$ . Alors, si  $H_0$  est acceptée le rang du modèle *AR* est égal à  $p$ . Pour illustrer cette procédure nous allons considérer l'exemple de la série chronologique C reprise dans l'appendice([6]).

La figure 3.1 est une représentation de la série C



### 3.2. AJUSTEMENT D'UN MODÈLE ET DÉTERMINATION DE SON ORDRE

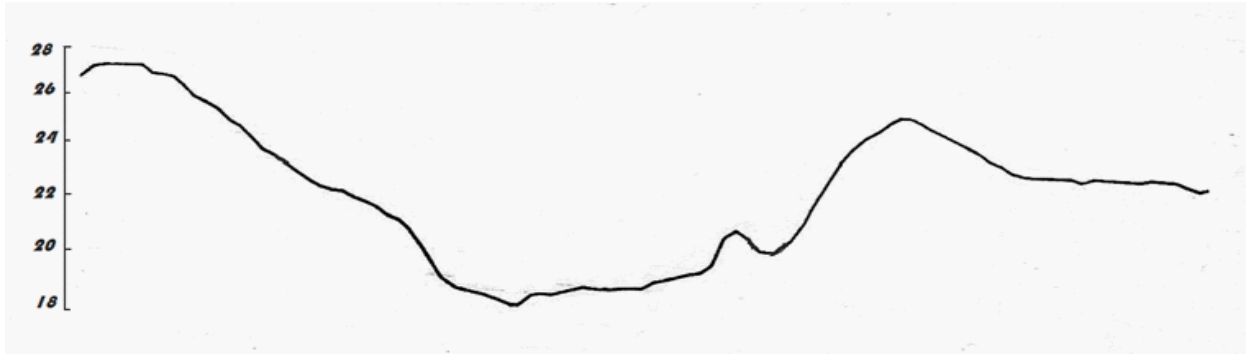


Figure 3.1 : série C : Température incontrôlée d'un processus chimique : lectures à chaque minute

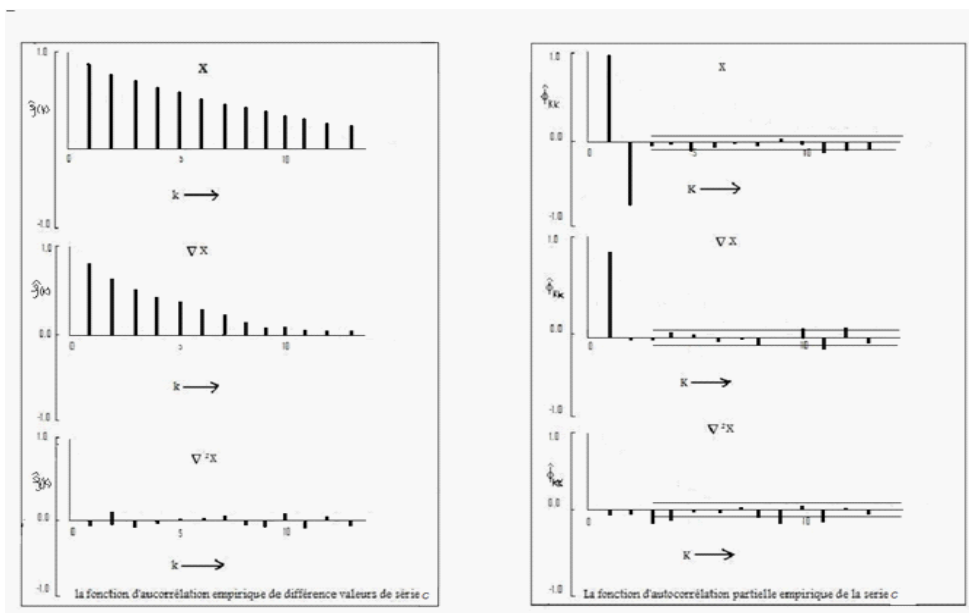


Figure 3.2

La figure (3.2) représente la fonction  $\hat{\rho}(k)$  et la fonction  $\hat{\alpha}(k)$ . On remarque dans ce graphe que les valeurs de la fonction  $\hat{\rho}(k)$  et les valeurs de la fonction  $\hat{\alpha}(k)$  s'approchent de 0 après l'horizon 2, donc la série  $X$  représente le modèle  $AR(2)$ . en appliquant une première différenciation  $\nabla X$ , on remarque dans le deuxième graphe de la fonction  $\hat{\rho}(k)$  de  $\nabla X$  que  $\hat{\rho}(k)$  décroît

exponentiellement vers 0. Par contre, la fonction  $\hat{\alpha}(k)$  chute à 0 après l'horizon 1, on dit que la série représente le modèle  $ARI(1, 1, 0)$ . L'application d'une deuxième différenciation  $\nabla^2 X$  montre que les valeurs de la fonction  $\hat{\rho}(k)$  et la fonction  $\hat{\alpha}(k)$  s'approche de 0, la série représente le modèle bruit incorréllé  $IMA(0, 2, 0)$ .

- Si  $\hat{\alpha}(k)$  décroît vers zéro et si  $\hat{\rho}(k)$  chute à zéro à partir d'un certain rang (i.e. les valeurs de  $\hat{\rho}(k)$  sont négligeables à partir d'un certain rang), alors on opte pour un modèle en moyenne mobile. Pour décider de l'ordre de ce modèle  $MA$  on se base comme dans le premier sur le résultat du test de l'hypothèse  $H_0 : \hat{\rho}(k) = 0, \text{ pour } k > q$  contre  $H_1 = \overline{H}_0$ . Alors, si  $H_0$  est acceptée le rang du modèle  $MA$  est égal à  $q$ . Pour illustrer cette procédure nous allons considérer l'exemple de la série chronologique A reprise dans l'appendice ([6]).

La figure 3.3 est une représentation de la série A

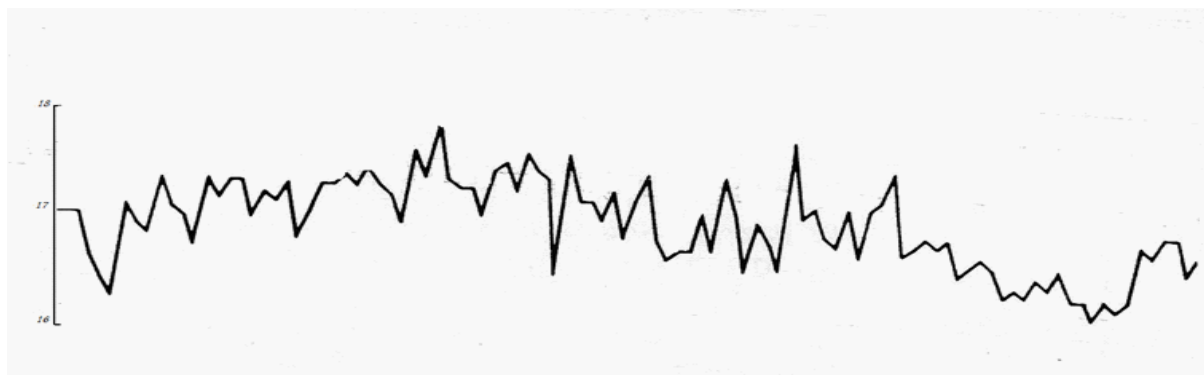


Figure 3.3 : Série A : Concentration incontrôlée processus chimique : lectures de deux heures :

### 3.2. AJUSTEMENT D'UN MODÈLE ET DÉTERMINATION DE SON ORDRE

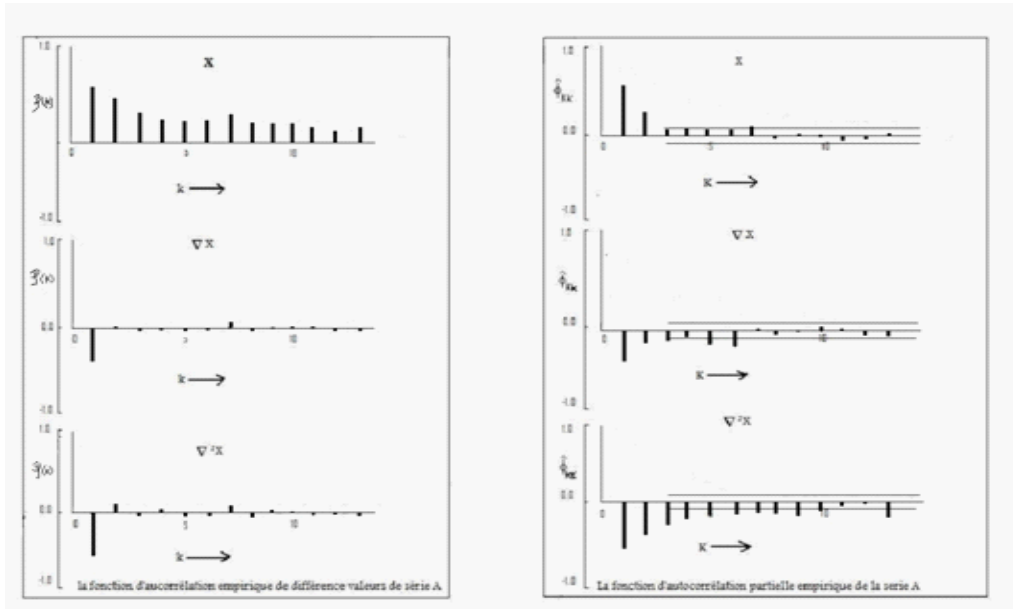


Figure 3.4

La figure (3.4) représente le graphe de la fonction d'autocorrélation empirique et la fonction d'autocorrélation partielle empirique de la série  $A$ , on remarque dans cette figure que la fonction d'autocorrélation empirique varie d'une manière irrégulière ce qui veut dire que cette série est non stationnaire. La même remarque peut être faite pour la fonction d'autocorrélation partielle empirique. Pour que la série  $A$  devienne stationnaire, il faut lui appliquer une première différenciation  $\nabla X$ . Ainsi, les valeurs de la fonction d'autocorrélation empirique de  $\nabla X$  s'approche de 0 après l'horizon 1, mais la fonction d'autocorrélation partielle empirique décroît d'une manière irrégulière donc on peut appliquer une deuxième différenciation  $\nabla^2 X$ . la fonction  $\hat{\rho}(k)$  de  $\nabla^2 X$  s'approche de 0 après l'horizon 1, et la fonction  $\hat{\alpha}(k)$  décroît exponentiellement vers 0 donc la série  $A$  représente le modèle  $IMA(0, 2, 1)$ .

- Si  $\hat{\alpha}(k)$  décroît vers zéro et si  $\hat{\rho}(k)$  décroît vers zéro (i.e. les valeurs de  $\hat{\rho}(k)$  sont négligeables) à partir d'un certain rang alors on opte pour un modèle mixte  $ARMA$ . Pour estimer l'ordre de ce modèle  $ARMA$  Pour illustrer cette procédure nous allons considérer l'exemple de la série chronologique  $F$  reprise dans l'appendice ([6]).

La figure 3.5 est une représentation de la série  $F$

CHAPITRE 3. IDENTIFICATION DE MODÈLES POUR LES SÉRIES  
CHRONOLOGIQUES

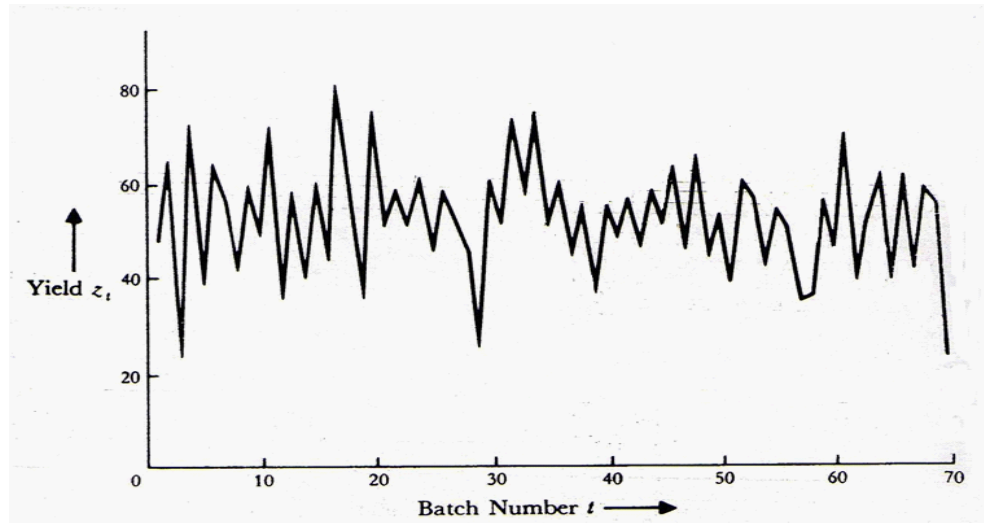


Figure 3.5 : Série F : les rendements de 70 lots consécutifs à partir d'un processus chimique

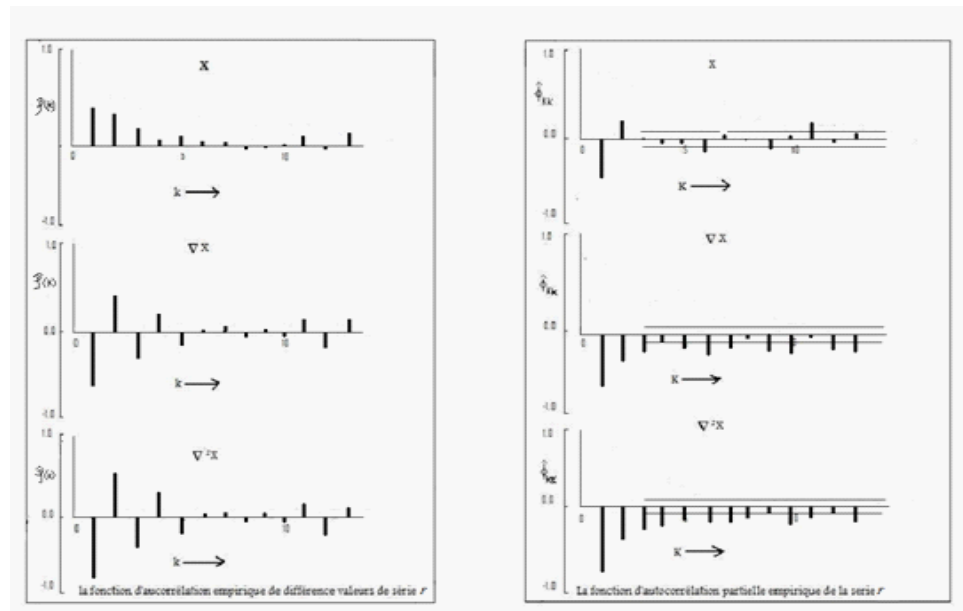


Figure 3.6

### 3.3. DÉTERMINATION DE L'ORDRE D'UN MODÈLE

---

La figure (3.6) représente les graphes de la fonction  $\hat{\rho}(k)$  et la fonction  $\hat{\alpha}(k)$  de la série  $F$ . On voit que la fonction  $\hat{\rho}(k)$  de  $X$  décroît irrégulièrement et on trouve la même remarque pour la fonction  $\hat{\alpha}(k)$ . Ainsi, on doit appliquer une première différenciation  $\nabla X$ . Dans le graphe de la fonction  $\hat{\rho}(k)$  de  $\nabla X$ , on remarque qu'elle décroît (oscillante) vers 0. Par contre, la fonction  $\hat{\alpha}(k)$  de  $\nabla X$  décroît d'une manière irrégulière. On applique une deuxième différenciation  $\nabla^2 X$ . Nous constatons que la fonction  $\hat{\rho}(k)$  de  $\nabla^2 X$  décroît (oscillante) vers 0 et la fonction  $\hat{\alpha}(k)$  de  $\nabla^2 X$  décroît également d'une manière exponentielle vers 0 donc la série  $F$  représente le modèle  $ARIMA(1, 2, 1)$ .

D'où, pour un processus  $AR(p)$  le  $ACF$  et  $PACF$  théorique sont comme suit :

$$\begin{aligned} ACF(i) &\neq 0 \quad \forall i \\ PACF(i) &\neq 0 \quad \forall i = 1, \dots, p \\ PACF(i) &= 0 \quad \forall i > p \end{aligned}$$

Pour un processus  $MA(q)$  le  $ACF$  et  $PACF$  théorique sont comme suit :

$$\begin{aligned} ACF(i) &\neq 0 \quad \forall i = 1, \dots, q \\ ACF(i) &= 0 \quad \forall i > q \\ PACF(i) &\neq 0 \quad \forall i \end{aligned}$$

Les processus mixtes de type  $ARMA$  peuvent présenter des graphes d'autocorrélation et d'autocorrélation partielle plus complexes à interpréter et nécessitent plusieurs itérations de type identification-estimation-diagnostic. Pour choisir l'ordre d'un modèle mixte  $ARMA$ , la double observation de l'acf et de la pacf seront nécessaires. D'ailleurs, remarquons de nouveau qu'en pratique il est toujours judicieux de regarder ces deux fonctions, l'acf et la pacf empirique, ensemble pour pouvoir ajuster le modèle  $AR$  et  $MA$ . L'acf et la pacf de modèle  $ARMA$  décroissent tous deux lentement.

### 3.3 Détermination de l'ordre d'un modèle

Dans ce paragraphe, on considère des structures statistiques  $(\Omega, H, P_\theta)_{\theta \in \Theta}$  dans lesquelles les probabilités  $P_\theta$  sont absolument continues par rapport à la mesure de Lebesgue. Si le vecteur  $\theta$  appartient à un sous-espace de  $\mathbb{R}^M$  de dimension  $k \leq M$ ,  $k$  est appelé ordre du modèle. Le problème de la détermination de l'ordre  $k$  du modèle sera alors un problème de décision statistique basé sur un échantillon de taille  $N$ . Le principe du critère d'Akaike consiste à choisir pour un ordre du modèle la valeur de  $k$  qui minimise une certaine fonction de risque  $R(\theta, \hat{\theta})$ .

Le critère d'information d'Akaiké (*AIC*) et le critère d'information bayésienne (*BIC*) sont les plus utilisés pour choisir le meilleur d'entre deux modèles plausibles. Un modèle est meilleur qu'un autre si son *AIC* ou son *BIC* est le plus petit. Le critère d'information d'Akaiké est basé sur la distance de Kullback-Leiber, alors que le *BIC* est basé sur la vraisemblance intégrée dans la théorie bayésienne.

### 3.3.1 Le critère d'information d'Akaiké

Nous allons résumer dans ce paragraphe les travaux d'Akaiké ([1]). Ce dernier avait établi un critère d'information pour la détermination de l'ordre d'un modèle ajusté à une série chronologique. Ce critère est basé essentiellement sur l'information de Kulback-Leiber. En plaçant le problème dans le cadre de la théorie de la décision, l'auteur définit une fonction de perte  $W(\theta, \hat{\theta})$  telle que :

$$W(\theta, \hat{\theta}) = (-2) \int f(x; \theta) \log \left( \frac{f(x; \hat{\theta})}{f(x; \theta)} \right) dx$$

où  $-\int f(x; \theta) \log \left( \frac{f(x; \hat{\theta})}{f(x; \theta)} \right) dx$  désigne l'information de Kulback-Leiber([20]) Mais, la fonction de perte est aléatoire et donc ne peut être maîtrisée, alors il minimise la fonction de risque définie telle que :

$$R(\theta, \hat{\theta}) = E(W(\theta, \hat{\theta}))$$

où l'espérance est calculée par rapport à la distribution de  $\hat{\theta}$ .

En effet, supposons qu'on dispose de  $N$  réalisations indépendantes  $x_1, x_2, \dots, x_N$  d'une variable aléatoire  $X$  alors la moyenne empirique du logarithme du rapport des vraisemblances  $\frac{1}{N} \sum_{i=1}^N \log \left( \frac{f(x_i; \hat{\theta})}{f(x_i; \theta)} \right)$  est un estimateur consistant de la perte  $W(\theta, \hat{\theta})$  ([1]). Il est tout à fait naturel de s'attendre à ce que, pour au moins une valeur de  $N$  assez grande, la valeur de  $\hat{\theta}$  maximise  $\frac{1}{N} \sum_{i=1}^N \log \left( \frac{f(x_i; \hat{\theta})}{f(x_i; \theta)} \right)$  et ainsi minimise la perte  $W(\theta, \hat{\theta})$ .

### 3.3. DÉTERMINATION DE L'ORDRE D'UN MODÈLE

---

Le problème de la minimisation de  $W(\theta, \hat{\theta})$  ou de  $R(\theta, \hat{\theta})$  peut être résolu en combinant les deux techniques principales de la statistique inférentielle, à savoir l'estimation de paramètres et les tests d'hypothèses statistiques, en un seul problème de décision. Au lieu de considérer une seule estimation du paramètre  $\theta$ , on détermine les estimations de ce paramètre correspondant à différentes restrictions possibles de la distribution d'échantillonnage à des sous-espaces. On résout alors ce problème de décision multiple à l'aide de techniques de tests multiples ou de tests séquentiels.

Le problème d'identification de modèle est souvent formulé comme un problème de choix de  $f(x; \theta^{(k)})$  ( $k = 1, \dots, L$ ) basé sur  $N$  observations de  $X$ , où  $\theta^{(k)}$  est la restriction de  $\theta$  à un sous-espace de dimension  $k$ , i.e.  $\theta^{(k)} = (\theta_1, \theta_2, \dots, \theta_k, 0, \dots, 0)'$ . Et  $k$  est appelé ordre du modèle.

Akaïké propose l'utilisation de la statistique :

$$W_L^{(k)} = -\frac{2}{N} \sum_{i=1}^N \log \left( \frac{f(x_i; \hat{\theta}^{(k)})}{f(x_i; \hat{\theta}_L)} \right) \quad (3.1)$$

comme un estimateur de  $W(\theta, \hat{\theta}^{(k)})$ .

Nous remarquons par ailleurs que la statistique  $NW_L^{(k)}$  est une statistique de test du rapport des log-vraisemblances classiques, distribuée asymptotiquement suivant un khi-deux à  $L - k$  degrés de liberté quand le vrai paramètre appartient à l'espace de  $\theta^{(k)}$ . Définissons alors

$$W(\theta, \hat{\theta}^{(k)}) = \inf_{\theta^{(k)}} W(\theta, \theta^{(k)})$$

où  $\theta$  est le vrai paramètre de la distribution.

Alors, on s'attend à ce que  $W_L^{(k)}$  converge et

$$W_L^{(k)} \rightarrow W(\theta, \hat{\theta}^{(k)})$$

Supposons que  $W(\theta, \theta^{(k)})$  est suffisamment régulière en  $\theta$  et que  $W(\theta, \theta^{(k)}) > 0$  pour  $\theta \neq \theta^{(k)}$ . Alors la quantité  $W(\theta, \theta^{(k)})$  peut être remplacé par l'expression :

$$W_2(\theta, \hat{\theta}^{(k)}) = \sum_{l=1}^L \sum_{m=1}^L (\hat{\theta}_l^{(k)} - \theta_l) (\hat{\theta}_m^{(k)} - \theta_m) C(l, m)(\theta)$$

où  $C(l, m)(\theta)$  est le  $(l, m)$ ème élément de la matrice d'information de Fisher défini tel que :

$$C(l, m)(\theta) = - \int \left( \frac{\partial^2 \log f}{\partial \theta_l \partial \theta_m} \right) f_{\theta} dx.$$

Posons  $C(l, m) = C(l, m)(\theta)$  et notons par  $\|\theta\|_c$  la norme dans l'espace de  $\theta$  définie telle que :

$$\|\theta\|_c^2 = \sum_{l=1}^L \sum_{m=1}^L \theta_l \theta_m C(l, m)$$

On a alors :

$$W_2(\theta, \hat{\theta}^{(k)}) = \left\| \hat{\theta}^{(k)} - \theta \right\|_c^2 \quad (3.2)$$

Par ailleurs, redéfinissons  $\theta^{(k)}$  par la relation :

$$\left\| \hat{\theta}^{(k)} - \theta \right\|_c^2 = \min_{\theta^{(k)}} \left\| \theta^{(k)} - \theta \right\|_c^2 \quad (3.3)$$

Des relations (3.2), (3.3) on déduit que :

$$W_2(\theta, \hat{\theta}^{(k)}) = \left\| \theta^{(k)} - \theta \right\|_c^2 + \left\| \hat{\theta}^{(k)} - \theta^{(k)} \right\|_c^2.$$

On considère que, quand  $N$  augmente,  $N$  et  $k$  sont choisis de sorte que l'expression  $\sqrt{N}(\theta_m^{(k)} - \theta_m)$ , ( $m = 1, 2, \dots, k$ ) soit bornée.

Alors :

$$NW_L^{(k)} = N \left\| \hat{\theta} - \theta \right\|_c^2 - N \left\| \hat{\theta}^{(k)} - \theta^{(k)} \right\|_c^2$$

Par une simple manipulation, on déduit :

$$NW_L^{(k)} = N \left\| \theta^{(k)} - \theta \right\|_c^2 + N \left\| \hat{\theta} - \theta \right\|_c^2 - N \left\| \hat{\theta}^{(k)} - \theta^{(k)} \right\|_c^2 - 2N \left\langle \hat{\theta} - \theta, \hat{\theta}^{(k)} - \theta \right\rangle_c.$$

où  $\langle, \rangle_c$  représente le produit scalaire défini par  $C(l, m)$ .

Soit  $C^{-1}$  l'inverse de la matrice d'information de Fisher. Supposons que la distribution de  $\sqrt{N}(\hat{\theta} - \theta)$  tend vers la distribution  $N(0, C^{-1})$ . L'auteur montre que géométriquement  $(\hat{\theta}^{(k)} - \theta^{(k)})$  est approximativement la projection de  $(\hat{\theta} - \theta)$  sur l'espace des  $\theta^{(k)}$ . De ce résultat on peut montrer que les statistiques  $N \left\| \hat{\theta} - \theta \right\|_c^2 - N \left\| \hat{\theta}^{(k)} - \theta^{(k)} \right\|_c^2$  et  $N \left\| \hat{\theta}^{(k)} - \theta^{(k)} \right\|_c^2$  sont



### 3.3. DÉTERMINATION DE L'ORDRE D'UN MODÈLE

---

distribuées asymptotiquement suivant  $\chi_{(L-k)}^2$  et  $\chi_{(k)}^2$  respectivement. Finalement, et d'après ce qui précède  $W_L^{(k)}$  peut ne pas être un bon estimateur de  $W_2(\theta, \hat{\theta}^{(k)})$  et  $r(\hat{\theta}, \hat{\theta}^{(k)}) = N^{-1} (NW_L^{(k)} + 2k - L)$  peut servir comme un estimateur utile de  $EW_2(\theta, \hat{\theta}^{(k)})$ .

Il est intéressant de remarquer que dans les situations pratiques,  $L$  est peut être très grand même infini, et de ce fait peut ne pas être défini clairement. Mais du fait que nous nous intéressons uniquement à  $\hat{\theta}^{(k)}$  qui minimise  $r(\hat{\theta}, \hat{\theta}^{(k)})$ , il suffit de calculer :

$$\nu^{(k)} = NW_L^{(k)} + 2k$$

ou bien

$$AIC = \lambda^{(k)} = -2 \sum_{i=1}^N \log f(x_i, \hat{\theta}^{(k)}) + 2k \quad (3.4)$$

et le  $\hat{\theta}^{(k)}$  cherche à minimiser la quantité  $AIC$ .

#### 3.3.2 Le critère d'information de Bayésien

Considérons  $X = (X_1, X_2, \dots, X_n)'$  un échantillon aléatoire de distribution inconnue  $f_\theta$ . Le problème est d'estimer la distribution  $f_\theta$ . On se donne une collection finie de modèles statistiques  $\{M_1, M_2, \dots, M_m\}$ . A chaque modèle  $M_i$  correspond une densité  $g_{M_i}$  de paramètre  $\theta_i$  de dimension  $K_i$ ,  $M_i \in \{M_1, M_2, \dots, M_m\}$ .  $\Theta_i$  est un espace de dimension  $K_i$  auquel appartient  $\theta_i$ . Le  $BIC$  cherche à déterminer le modèle  $M_i$  qui maximise la probabilité à posteriori  $P(M_i/X)$  :

$$M_{BIC} = \arg \max P(M_i/X) \quad (3.5)$$

Le  $BIC$  cherche à sélectionner le modèle le plus vraisemblable au vue des données.

D'après la formule de bayes on a :

$$P(M_i/X) = \frac{P(X/M_i) P(M_i)}{P(X)} \quad (3.6)$$

En supposant que la loi à priori des modèles  $M_i$  est non informative, i.e.  $P(M_1) = P(M_2) = \dots = P(M_m)$ , i.e. aucun modèle n'est privilégié, et

d'après (3.5) et (3.6) la recherche du meilleur modèle ne nécessite que le calcul de la distribution  $P(X/M_i)$  que l'on obtient tel que :

$$P(X/M_i) = \int_{\Theta_i} P(X, \theta_i/M_i) d\theta_i = \int g_{M_i}(X, \theta_i) P(\theta_i/M_i) d\theta_i$$

où  $g_{M_i}(X, \theta_i)$  est la vraisemblance du n-échantillon correspondant au modèle  $M_i$  de paramètre  $\theta_i$  :

$$g_{M_i}(X, \theta_i) = P(X/\theta_i, M_i) = L(\theta_i, M_i)$$

On réécrit cet intégrale sous la forme :

$$P(X/M_i) = \int e^{g(\theta_i)} d\theta_i, \text{ où } g(\theta_i) = \log(g_{M_i}(X, \theta_i) P(\theta_i/M_i)) \quad (3.7)$$

où  $P(X/M_i)$  est appelée vraisemblance intégrée pour le modèle  $M_i$ .

Le calcul de cette vraisemblance est rarement possible. Mais peut approcher  $P(X/M_i)$  en utilisant l'approximation de Laplace ([21]). Ainsi, l'équation (3.7) peut s'exprimer telle que :

$$P(X/M_i) = e^{g(\theta_i^*)} \left( \frac{2\pi}{N} \right) |A_{\theta_i^*}| + \mathcal{O}(N^{-1})$$

où  $\theta_i^* = \arg \max_{\theta_i \in \Theta_i} L_n(\theta_i)$  et  $A_{\theta_i^*}$  l'opposé de la matrice hessienne des secondes dérivées partielles de la fonction  $L_n(\theta_i) : \mathbb{R}^d \rightarrow \mathbb{R}$  en  $\theta_i$  et  $L_N(\theta_i) = \frac{g(\theta_i)}{N} = \frac{1}{N} \sum_{K=1}^N \log(g_{M_i}(X_K, \theta_i)) + \frac{\log P(\theta_i/M_i)}{N}$  :

$$A_{\theta_i^*} = - \left[ \frac{\partial^2 L_n(\theta_i)}{\partial \theta_i^j \partial \theta_i^l} \right]_{j,l} \Big|_{\theta_i = \theta_i^*}$$

$\theta_i^*$  peut être remplacée par l'estimateur du maximum de vraisemblance  $\hat{\theta}_i$  :

$$\hat{\theta}_i = \arg \max \frac{1}{N} g_{M_i}(X, \theta_i)$$

et  $A_{\theta_i^*}$  par  $I_{\hat{\theta}_i}$  où  $I_{\hat{\theta}_i}$  est la matrice d'information de Fisher pour une observation définie par :

$$I_{\hat{\theta}_i} = -E \left[ \frac{\partial^2 \log(g_{M_i}(X_1, \theta_i))}{\partial \theta_i^j \partial \theta_i^l} \right]_{j,l} \Big|_{\theta_i = \hat{\theta}_i}$$

### 3.3. DÉTERMINATION DE L'ORDRE D'UN MODÈLE

---

Nous obtenons alors :

$$\begin{aligned}\log(P(X/M_i)) &= \log\left(g_{M_i}(X, \hat{\theta}_i)\right) - \frac{K_i}{2} \log(N) + \log\left(P(\hat{\theta}_i/M_i)\right) + \frac{K_i}{2} \log(2\pi) \\ &\quad - \frac{1}{2} \log(|I_{\hat{\theta}_i}|) + \mathcal{O}\left(N^{-\frac{1}{2}}\right)\end{aligned}$$

D'où :

$$\log(P(X/M_i)) \approx \log\left(g_{M_i}(X, \hat{\theta}_i)\right) - \frac{K_i}{2} \log(N)$$

Le *BIC* pour un modèle  $M_i$  est défini par :

$$BIC_i = -2 \log\left(g_{M_i}(X, \hat{\theta}_i)\right) + K_i \log(N)$$

Par conséquent, le meilleur modèle pour ce critère est défini tel que :

$$M_{BIC} = \arg \min_{M_i} BIC_i$$

Dans la pratique, le critère *BIC* appelé aussi critère d'information de Schwartz (*SIC*) est donné sous la forme simple suivante :

$$BIC = \left( \frac{-2 \ln L + k \ln(N)}{N} \right)$$

$N$  = Le nombre d'observations

$k$  = Le nombre de paramètres qui doivent être estimés

$L$  = Le maximum de la fonction de vraisemblance du modèle estimé

# Chapitre 4

## Représentation interne des séries chronologiques

### 4.1 Introduction

Une classe générale de modèles réveillant beaucoup d'intérêt actuel est les modèles état-espace. Ils ont été à l'origine développés par les ingénieurs de contrôle, particulièrement pour les applications concrétistes des systèmes de navigation. Pourtant, ils ont aussi été trouvés utiles dans beaucoup de types de problème de série chronologique, comme les prévisions à court terme. Ce chapitre présente des modèles état-espace pour l'analyse de série chronologique. Aussi bien qu'en décrivant le filtre de Kalman, qui est une méthode générale importante pour manipuler des modèles état-espace. Essentiellement, La filtration de Kalman est une méthode pour le traitement de signal qui donne des estimations optimales de l'état actuel d'un système dynamique. Il se compose d'un ensemble d'équations pour recueillir l'estimant de l'état actuel d'un système et pour trouver des variances de ces estimations.

### 4.2 Modèle état-espace

**Définition 4.2.1** *Soit  $Y_t$  un vecteur  $(n \times 1)$  de variables observées à l'instant  $t$ . Une classe riche de modèles dynamiques pour  $Y_t$  peut être décrite en termes d'un vecteur  $(r \times 1)$  probablement non observé,  $X_t$  connu comme le vecteur d'état. Le système suivant donne à la représentation état-espace de*

la dynamique de  $Y_t$  l'équation suivante :

$$Y_t = G_t X_t + W_t \quad t = 1, 2, \dots, \quad (4.1)$$

où

$$X_{t+1} = F_t X_t + V_t \quad t = 1, 2, \dots, \quad (4.2)$$

**Remarque 4.2.1** Avant d'aller plus loin, on présente les hypothèses à utiliser dans l'analyse des équations d'état (4.2) et l'équation d'observation (4.1).

$F_1, F_2, \dots$  est un ordre indiqué ( $v \times v$ ) matrice.

$G_1, G_2, \dots$  est un ordre indiqué ( $w \times v$ ) matrice.

$\{X_t, (V_t', W_t')', t = 1, 2, \dots\}$  est un ordre orthogonal de vecteurs aléatoires avec le deuxième moments fini.

### 4.2.1 Propriétés des représentations internes

**Définition 4.2.2** Soit  $\mathbf{H} = (\mathcal{H}_t, t \in \mathbb{Z})$  une famille croissante d'espaces gaussiens. On dira que cette famille est admissible si  $Y_t$  est adapté à  $\mathbf{H}$ . De plus, on dira que le processus gaussien centré  $(X_t, W_t)$  stationnairement corrélé avec  $Y_t$ , est une  $\mathbf{H}$ -représentation d'ordre  $q$  de  $Y_t$  si et seulement si :

(i)  $(X_t, t \in \mathbb{Z})$  est un processus à valeurs dans  $\mathbb{R}^q$ , markovien par rapport à  $\mathbf{H}$ .

(ii)  $(W_t, t \in \mathbb{Z})$  est un processus à valeurs dans  $\mathbb{R}^p$ , adapté et innovant par rapport à  $\mathbf{H}$ .

(iii) il existe une matrice  $H$  de dimension  $(p, q)$  telle que :

$$Y_t = H X_{t-1} + W_t \quad \forall t \in \mathbb{Z}$$

Par ailleurs, on dira qu'une représentation de  $Y_t$  est d'ordre minimal, noté  $O(Y_t)$ , s'il n'existe pas de représentation d'ordre strictement inférieur.

**Proposition 4.2.1** Le processus  $W_t$  est un bruit blanc et le vecteur  $(X_t \ Y_t)'$  est markovien par rapport à  $\mathbf{H}$ .

**Preuve** Nous rappelons qu'un bruit blanc est par définition un processus gaussien totalement "décorrélé". D'après le point (ii) de la définition (4.2.2), on a :

$$E(W_{t+s} W_s) = E\{E^{\mathcal{H}_s}(W_{t+s}) W_s\} = 0 \quad \forall t \geq 1$$

Donc,  $W$  est un bruit blanc.

On déduit aussi à partir du point (iii) de la définition (4.2.2) que :

$$W_t = Y_t - E^{\mathcal{H}_{t-1}} Y_t \quad (4.3)$$

Réciproquement, si  $W_t$  est défini par la relation (4.3) il est clair que c'est un bruit blanc adapté et innovant par rapport à  $\mathbf{H}$ .

Par ailleurs, il est évident que  $(X_t \ Y_t)'$  est adapté à  $\mathbf{H}$ . De plus, si  $F$  désigne la matrice de transition de  $X_t$ , on a :

$$E^{\mathcal{H}_{t-1}} X_t = F X_{t-1} \quad (4.4)$$

Par ailleurs,  $W_t$  étant innovant par rapport à  $\mathbf{H}$ , il vient :

$$E^{\mathcal{H}_{t-1}} Y_t = H X_{t-1} \quad (4.5)$$

D'après les relations (4.4) et (4.5) on peut affirmer que le vecteur  $(X_t \ Y_t)'$  est markovien par rapport à  $\mathbf{H}$ , de matrice de transition  $\mathbf{F} = \begin{pmatrix} F & 0 \\ H & 0 \end{pmatrix}$ . On obtient alors le modèle :

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \mathbf{F} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} V_t \\ W_t \end{bmatrix}$$

Ou bien de façon explicite :

$$\begin{cases} X_t = F X_{t-1} + V_t \\ Y_t = H X_{t-1} + W_t \end{cases}$$

On retrouve le modèle ou la représentation état-espace de Kalman-Bucy ([18]).

■

**Remarque 4.2.2** *On déduit de ce qui précède que :*

$$E^{\mathcal{H}_t} (Y_{t+k}) = H F^{k-1} X_{t-1} \quad \forall k \geq 1$$

### 4.3 Représentation état-espace d'un processus stationnaire

**Définition 4.3.1** *Considérons les équations d'observation et d'état suivantes :*

$$Y_t = G X_t + W_t \quad t = 0, \pm 1, \dots$$

4.4. REPRÉSENTATION ÉTAT-ESPACE D'UN PROCESSUS  
ARMA(P, Q)

---

$$X_{t+1} = FX_t + V_t \quad t = 0, \pm 1, \dots \quad (4.6)$$

où  $F$  et  $G$  sont  $v \times v$  et  $w \times w$  matrices respectivement,  $\{V_t\} \sim WN(0, Q)$ ,  $\{W_t\} \sim WN(0, s)$ ,  $E(V_t W_t) = s$  pour tout  $t$  et  $V_t \perp W_t$  pour tout  $t \leq t$ . L'équation d'état (4.6) est dite stable (ou causale) si la matrice  $F$  a toutes ses valeurs dans l'intérieur du cercle d'unité ou équivalent si  $\det(I - F_z) \neq 0$  pour tout  $z \in \mathbb{C}$  tel que  $|z| \leq 1$ . On dit alors aussi la matrice  $F$  stable.

**Remarque 4.3.1** Dans le cas stable l'équation (4.6) a la solution stationnaire unique donnée par :

$$X_t = \sum_{j=0}^{\infty} F^j V_{t-j-1} \quad (4.7)$$

La transmission de l'ordre d'observation :

$$Y_t = W_t + \sum_{j=0}^{\infty} G F^j V_{t-j-1}$$

est alors stationnaire.

## 4.4 Représentation état-espace d'un processus ARMA(p, q)

La représentation état-espace n'est pas unique. Nous donnerons deux représentations pour le processus ARMA(p, q). Pour le premier cas, nous considérons le processus ARMA(p, q) défini tel que :

$$\phi(B) Y_t = \theta(B) Z_t \quad t = 0, \pm 1, \dots \quad (4.8)$$

où  $\{Z_t\} \sim WN(0, \sigma^2)$  et  $\phi(B) \neq 0$  pour  $|B| \leq 1$ , soit :

$$r = \max(p, q + 1) \quad \phi_j = 0 \text{ pour } j > p \quad \theta_j = 0 \text{ pour } j > q \text{ et } \theta_0 = 1$$

C'est clair qu'à partir de (4.8), nous pouvons écrire :

$$Y_t = [\theta_{r-1} \theta_{r-2} \dots \theta_0] X_t \quad (4.9)$$

où

$$X_t = \begin{bmatrix} X_{t-r+1} \\ X_{t-r+2} \\ \vdots \\ X_t \end{bmatrix} \quad (4.10)$$

et

$$\phi(B) X_t = Z_t \quad t = 0, \pm 1, \dots \quad (4.11)$$

Donc

$$X_{t+1} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ \phi_r & \phi_{r-1} & \phi_{r-2} & \dots & \phi_1 \end{bmatrix} X_t + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \phi_1 \end{bmatrix} Z_{t+1} \quad (4.12)$$

Les équations (4.9) et (4.10) sont respectivement l'équation d'observation et l'équation d'état. L'hypothèse de causalité implique que (4.10) a une solution stationnaire unique qui détermine un ordre stationnaire par l'équation d'observation (4.9). Il est facile de vérifier que l'ordre de cette équation satisfait les équations *ARMA* (4.8) et coïncide avec leur solution stationnaire unique. Pour la deuxième représentation, nous considérons le processus  $Y_t$  un processus *ARMA*( $p, q$ ) défini par l'équation (4.8). Nous allons maintenant établir une représentation état-espace dimensionnelle baissée qui est dérivée de la représentation précédente :

$$m = \max(p, q) \quad \text{et} \quad \phi_j = 0 \quad \text{pour} \quad j > p$$

Alors :

$$Y_t = [1 \ 0 \ 0 \ \dots \ 0] X_t + Z_t, \quad t = 0, 1, \dots, \quad (4.13)$$

où  $X_t$  est la solution unique stationnaire de :

$$X_{t+1} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ \phi_m & \phi_{m-1} & \phi_{m-2} & \dots & \phi_1 \end{bmatrix} X_t + \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{m-1} \\ \psi_{m-2} \end{bmatrix} Z_t, \quad t = 0, 1, \dots, \quad (4.14)$$



4.4. REPRÉSENTATION ÉTAT-ESPACE D'UN PROCESSUS  
ARMA(P, Q)

---

où  $\psi_1, \psi_2, \dots, \psi_m$  sont les coefficients de  $B, B^2, \dots, B^m$ , où  $\psi(B) = \frac{\theta(B)}{\phi(B)}$ ,  $|B| \leq 1$  (si  $m = 1$  les coefficients de  $X_t$  dans (4.13) et (4.12) sont 1 et  $\phi_1$  respectivement)

**Preuve** Le résultat

$$GF^{-1}H = \psi_i, \quad i = 1, \dots, m, \quad (4.15)$$

et

$$F^m - \phi_1 F^{m-1} - \dots - \phi_m I = 0 \quad (4.16)$$

À partir de (4.13) et de (4.14), nous avons :

$$\begin{aligned} Y_t &= GX_t + Z_t \\ Y_{t+1} &= GFX_t + GFZ_t + Z_{t+1} \\ &\vdots \\ Y_{t+m} &= GF^m X_t + GF^{m-1} H Z_t + \dots + Z_{t+m} \end{aligned}$$

Ces équations, au même temps avec (4.15) et (4.16), impliquent que :

$$Y_{t+m} - \phi_1 Y_{t+m-1} - \dots - \phi_m Y_{t+1} = \begin{bmatrix} -\phi_m & -\phi_{m-1} & \dots & -\phi_1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \psi_1 & 1 & 0 & \dots & 0 \\ \psi_2 & \psi_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi_m & \psi_{m-1} & \psi_1 & \dots & 1 \end{bmatrix} \begin{bmatrix} Z_t \\ Z_{t+1} \\ Z_{t+2} \\ \vdots \\ Z_{t+m} \end{bmatrix}$$

Si  $\psi_1, \psi_2, \dots, \psi_m$  sont les coefficients de  $B, B^2, \dots, B^m$ , où  $\psi(B) = \frac{\theta(B)}{\phi(B)}$ ,  $|B| \leq 1$  i.e.  $\psi_1 = \phi_1 \psi_{j-1} + \phi_1 \psi_{j-1} + \dots + \phi_j + \theta_j$ , nous pouvons conclure que  $Y_t$  satisfait l'équation ARMA :

$$Y_{t+m} - \phi_1 Y_{t+m-1} - \dots - \phi_m Y_{t+1} = \begin{bmatrix} \theta_m & \theta_{m-1} & \dots & \theta_1 & 1 \end{bmatrix} \begin{bmatrix} Z_t \\ Z_{t+1} \\ Z_{t+2} \\ \vdots \\ Z_{t+m} \end{bmatrix}$$

Ainsi le processus stationnaire  $Y_t$  défini par (4.13), (4.14) satisfait  $\phi(B) Y_t = \theta(B) Z_t$ ,  $t = 0, 1, \dots$ , et donc coïncide avec la solution unique stationnaire de ces équations. ■

## 4.5 La fonction d'autocovariance d'un processus *ARMA* causal

Si  $Y_t$  est le processus *ARMA* causal défini par (4.8), à partir de la deuxième représentation du processus *ARMA* et l'équation (4.7), nous aurons :

$$Y_t = GX_t + Z_t$$

où

$$G = [ 1 \ 0 \ 0 \ \dots \ 0 ]$$

et

$$X_t = HZ_{t-1} + FHZ_{t-2} + F^2HZ_{t-3} + \dots$$

Avec la matrice carrée  $F$  et le vecteur column  $H$  comme dans (4.14), nous obtenons immédiatement les équations :

$$Y_t = Z_t + \sum_{j=1}^{\infty} GF^{j-1}HZ_{t-j} \quad (4.17)$$

Par suite :

$$\gamma_Y(k) = \begin{cases} \sigma^2 \left[ 1 + \sum_{j=1}^{\infty} GF^{j-1}HH'F^{j-1}G' \right] & \text{si } k = 0 \\ \sigma^2 \left[ GF^{|k|-1}H + \sum_{j=1}^{\infty} GF^{j-1}HH'F^{|k|+j-1}G' \right] & k \neq 0 \end{cases}$$

Les coefficients  $\psi_j$  dans la représentation  $Y_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$  peuvent être lus dans (4.17) comme :

$$\psi_j = \begin{cases} 1 & \text{si } j = 0 \\ GF^{j-1}H & \text{si } j \geq 1 \end{cases}$$

où  $\psi_j$  converge géométriquement vers zéro comme  $j \rightarrow \infty$ .

4.5. LA FONCTION D'AUTOCOVARIANCE D'UN PROCESSUS  
ARMA CAUSAL

## Appendice

1	17.0	34	17.4	67	17.0	100	16.9	133	16.9	166	16.9
2	16.6	35	17.4	68	16.9	101	16.5	134	16.9	167	16.9
3	16.3	36	17.1	69	17.3	102	17.2	135	17.0	168	17.0
4	16.1	37	17.6	70	16.8	103	16.4	136	16.5	169	16.7
5	17.1	38	17.7	71	17.3	104	17.0	137	16.7	170	16.9
6	16.9	39	17.4	72	17.4	105	17.0	138	16.8	171	17.3
7	16.8	40	17.8	73	17.7	106	16.7	139	16.7	172	17.8
8	17.4	41	17.6	74	16.8	107	16.2	140	16.7	173	17.8
9	17.1	42	17.5	75	16.9	108	16.6	141	16.6	174	17.6
10	17.0	43	16.5	76	17.0	109	16.9	142	16.5	175	17.5
11	16.7	44	17.8	77	16.9	110	16.5	143	17.0	176	17.0
12	17.4	45	17.3	78	17.0	111	16.6	144	16.7	177	16.9
13	17.2	46	17.3	79	16.6	112	16.6	145	16.7	178	17.1
14	17.4	47	17.1	80	16.7	113	17.0	146	16.9	179	17.2
15	17.4	48	17.4	81	16.8	114	17.1	147	17.4	180	17.4
16	17.0	49	16.9	82	16.7	115	17.1	148	17.1	181	17.5
17	17.3	50	17.3	83	16.4	116	16.7	149	17.0	182	17.9
18	17.2	51	17.6	84	16.5	117	16.8	150	16.8	183	17.0
19	17.4	52	16.9	85	16.4	118	16.3	151	17.2	184	17.0
20	16.8	53	16.7	86	16.6	119	16.6	152	17.2	185	17.0
21	17.1	54	16.8	87	16.5	120	16.8	153	17.4	186	17.2
22	17.4	55	16.8	88	16.7	121	16.9	154	17.2	187	17.3
23	17.4	56	17.2	89	16.4	122	17.1	155	16.9	188	17.4
24	17.5	57	16.8	90	16.4	123	16.8	156	16.8	189	17.4
25	17.4	58	17.6	91	16.2	124	17.0	157	17.0	190	17.0
26	17.6	59	17.2	92	16.4	125	17.2	158	17.4	191	18.0
27	17.4	60	16.6	93	16.3	126	17.3	159	17.2	192	18.2
28	17.3	61	17.1	94	16.4	127	17.2	160	17.2	193	17.6
29	17.0	62	16.9	95	17.0	128	17.3	161	17.1	194	17.8
30	17.8	63	16.6	96	16.9	129	17.2	162	17.1	195	17.7
31	17.5	64	18.0	97	17.1	130	17.2	163	17.0	196	17.2
32	18.1	65	17.2	98	17.1	131	17.5	164	17.4	197	17.4
33	17.5	66	17.3	99	16.7	132	16.9	165	17.2		

Series A : Chemical process concentration  
readings : every 2hours

Ce tableau a été utilisé pour la construction du graphe de la figure 3.3 page 36.

26.6	22.7	19.6	25.6	24.4	23.7	20.8	24.9	24.1	23.3
27.0	22.6	19.6	25.8	24.4	23.6	20.8	24.9	24.1	23.2
27.1	22.4	19.6	26.1	24.5	23.4	20.8	24.9	24.0	23.3
27.1	22.2	19.6	26.3	24.5	23.2	20.9	25.1	24.0	23.3
27.1	22.0	19.7	26.3	24.4	23.0	20.8	25.0	24.0	23.2
27.1	21.8	19.9	26.2	24.3	22.8	20.8	25.0	23.9	23.1
26.9	21.4	20.0	26.0	24.2	22.6	20.7	25.0	23.8	22.9
26.8	20.9	20.1	25.8	24.2	22.4	20.7	25.0	23.8	22.8
26.7	20.3	20.2	25.6	24.0	22.0	20.8	24.9	23.7	22.6
26.4	19.7	20.3	25.4	23.9	21.6	20.9	24.8	23.7	22.4
26.0	19.4	20.6	25.2	23.7	21.3	21.2	24.7	23.6	22.2
25.8	19.3	21.6	24.9	23.6	21.2	21.4	24.6	23.7	21.8
25.6	19.2	21.9	24.7	23.5	21.2	21.7	24.5	23.6	21.3
25.2	19.1	21.7	24.5	23.5	21.1	21.8	24.5	23.6	20.8
25.0	19.0	21.3	24.4	23.5	21.0	21.9	24.5	23.6	20.2
24.6	18.9	21.2	24.4	23.5	20.9	22.2	24.5	23.5	19.7
24.2	18.9	21.4	24.4	23.5	21.0	22.5	24.5	23.5	19.3
24.0	19.2	21.7	24.4	23.7	21.0	22.8	24.5	23.4	19.1
23.7	19.3	22.2	24.4	23.8	21.1	23.1	24.5	23.3	19.0
23.4	19.3	23.0	24.3	23.8	21.2	23.4	24.4	23.3	
23.1	19.4	23.8	24.4	23.9	21.1	23.8	24.4	23.3	
22.9	19.5	24.6	24.4	23.9	20.9	24.1	24.2	23.4	
22.8	19.6	25.1	24.4	23.8	20.8	24.6	24.2	23.4	

Series C : Chemical process temperature readings :  
every minute

Ce tableau a été utilisé pour la construction du graphe de la figure 3.1 page 35.

*CHAPITRE 4. REPRÉSENTATION INTERNE DES SÉRIES  
CHRONOLOGIQUES*

---

47	44	50	62	68
64	80	71	44	38
23	55	56	64	50
71	37	74	43	60
38	74	50	52	39
64	51	58	38	59
55	57	45	59	40
41	50	54	55	57
59	60	36	41	54
48	45	54	53	23
71	57	48	49	
35	50	55	34	
57	45	45	35	
40	25	57	54	
58	59	50	45	

Series F : Yields from a batch chemical process

Ce tableau a été utilisé pour la construction du graphe de la figure 3.5 page 38.

# Bibliographie

- [1] Akaiké. H., "Information theory and an extension of the maximum likelihood principle", Institute of statistical Mathematics 4-6-7 Minami-Azabu, Minato-ru, Tokyo, Japan 106, 2nd International Symposium on Information Theory.
- [2] Andrew R. Liddle, "Information criteria for astrophysical model selection", 28 June 2008, Astronomy centre, university of sussex, Brighton BN1 9QH, United Kingdom 235-244.
- [3] Avram. F, "Séries temporelles : régression, modélisation ARIMA(p,d,q), et modélisation espace- état", 29 juin 2006.
- [4] Azencot R. et Dacunha-castelle D. "Séries d'observations irrégulières. Modélisation et prévision". Masson,
- [5] Rapacchi. B, Dur, dur "Les séries chronologiques", Centre Interuniversitaire de Calcul de Grenoble, 18 août 1993.
- [6] Box G.E.P. and Jenkins G.M. "Time series analysis, forecasting and control", holden day, San francisco, 1976.
- [7] Brian D. Ripley, Michael Brady, and Stephen M. Smith, "Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data Mark W. Woolrich", Received November 16, 2000 ; published online October 24, 2001
- [8] Brockwell.J.Davis.A. "Time series : Theory and methods". Springer in statistics. Second edition, 1979.
- [9] Dachmouche. M, "Modélisation des séries chronologiques : Méthode de Box&Jenkins", Département de Mathématique, Faculté des sciences, Université Mentouri Constantine, Les 1ères journées d'informatique pour l'entreprise, 04-06 Mars 2002, p 252-261.
- [10] Dakhmouche. M. "Statistique des processus lineaire : test et estimation". Thèse de doctorat. P : 4-29. 2006.

- [11] Diggle Peter. "Time series : a biostatistical introduction",1990.
- [12] Dobre. L, Alexandru. A, "Modelling unemployment rate using Box-Jenkins procedure", Journal of applied, Quantitative methods, Summer 2008.
- [13] Ferland.R, "MAT8181 Séries chronologiques", 9 février 2009.
- [14] Francis M. Mutua,"The use of the Akaike Information Criterion in the identification of an optimum flood frequency model Department of Meteorology", University of Nairobi, Kenya, June 1994, Hydrological sciences- journal- des science hydrological, 39,38, June 1994.
- [15] Hamilton J.D." Time series analysis". Princeton university press, Princeton, 1994.
- [16] Ojo. J.F. "On the performance and estimation of spectral and bispectral analysis of time series data". Asian journal of mathematics and statistics 1(3) : 150-158,2008.
- [17] Kalman, R. E. , "A new approach to linear filtering and prediction problems theory," J. Basic Engrg., 1961.
- [18] Kalman, R. E., and R. S. Bucy, "New results in linear filtering and prediction theory," J. Basic Engrg., 1961.
- [19] Kolmogoroff, A., "stationary sequences in Hilbert space," Bull. Math. Univ. Moscow, 1941.
- [20] Kullback, S. and Leibler, R. A., "On information and sufficiency". Ann. Math. Statist. 22 (1951) 79-86.
- [21] Lebarbier. E, Tristan Mary-Huard, "Le critère BIC : fondements théoriques et interprétation"
- [22] Le Gland. F, IRISA, "Introduction au Filtrage en Temps Discret Filtrage de Kalman et Modèles de Markov Cachés "/ INRIA Rennes 2007-08.
- [23] Makridakis. S, "A survey of time series", Int. Stat. Rev., Vol.44, No. 1, 1976, pp, 29-70. Longman Group Ltd. Printed in Great Britain.
- [24] McQuarrie, A. D. R., et Tsai, C. - L., "Régression et choix de modèle de série chronologique". Monde scientifique.1998.
- [25] Meyler. A, Kenny. G and Quinn. T, "Forecasting irish inflation using ARIMA models", Aidan Meyler. 1998, Technical paper.
- [26] Schwarz, G.« Estimant la dimension d'un modèle ». Annales des statistiques 6 (2) : 461-464.1978.

- [27] Shuhua Hu, " Akaike Information Criterion", March 15, 2007, Center for research in scienc computation , North carolina state university raleigh NC.
- [28] Von Sachs. M, Sébastien Van Bellegem, "Stat 2414 Séries chronologiques", Institut de statistique. Université de louvain, Voie de Roman Pays,20, B-1348 Louvain-la-Neuve, Belgium, 4 Edition, 26septembre 2005.
- [29] Walker, A.M. (1964), "Asymptotic properties of least squares estimates of the parameters of the spectrum of a stationary non-deterministic time series", J. Aust. Math. Soc., 4, 363-384.
- [30] Wiener, N., "Extrapolation, Interpolation and Smoothing of Stationary Time Series", Wiley, New York, 1949.
- [31] Wold, H. O., "A study in the analysis of stationary time series", Almqvist & Wiksell, Uppsala, Sweden, 1938 (2nd. ed. 1954).
- [32] Yule, G. U., "On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot nunspot numbers",1927.

## ملخص

الهدف من هذا العمل هو دراسة التحليل ونمذجة السلسلة الزمنية التجريبية. السلسلة الزمنية هي عبارة عن سلسلة من المتغيرات العشوائية  $X_{t_1}, X_{t_2}, \dots, X_{t_n}$  لاحظناها بالتسلسل في الوقت. إذا هذه المتغيرات العشوائية غير مستقلة. ويقاس هذا الارتباط بين المتغيرات العشوائية باستخدام الارتباطات الذاتية والارتباطات الذاتية الجزئية. توابع الارتباطات الذاتية والارتباطات الذاتية الجزئية من السلسلة الزمنية تمثل أدوات أساسية لتحديد وتعديل نماذج عشوائية.

في هذا العمل نركز أساسا على النماذج العشوائية الخطية ; لخصنا خصائص بارزة من النماذج ARMA باستخدام تقنيات Box et Jenkins وكذا بعض النماذج النوعية المتمثلة في "état-espace". كما يمكننا أن نعبر عن النماذج ARMA بصيغة « état-espace »

لإعطاء مزيد من التماسك لتحديد النماذج, أدخلنا مفهوم تابع الارتباط الذاتي العكسي, الذي يعتمد اعتمادا كبيرا على نوع نموذج معين. معايير المعلومات AIC و BIC تم اقتراحهم لتبيين نوعية النموذج المختار. كما يتم عرض إعادة صياغة التركيب الخاص بهذه المعايير و مقارناتها.