

République Algérienne Démocratique et populaire

Ministère de L'Enseignement Supérieur et de la Recherche Scientifique

Université Mentouri Constantine

Faculté des Sciences exacte

Département De Mathématiques

N° d'ordre

N° de série

Mémoire présenté pour l'obtention du Diplôme de Magister en Mathématiques

**MODELES STOCHASTIQUE ET LEUR PREVISION PAR DES
PROCESSUS AUTOREGRESSIFS**

Par

BOULAROUK Yakoub

Spécialité : Mathématique appliquée

Mémoire de Magister soutenu le 15 /06 / 2010 devant le jury

A. L. MARHOUNE	Professeur	Président
Y.YAHIOUCHE	Professeur	Rapporteur
Z. MOHDEB	Professeur	Examineur
F.L. RAHMANI	Maitre de conférences	Examineur

Remerciements

*Je remercie vivement monsieur **Y.YAHIOUCHE** professeur à l'université de constantine, pour avoir accepté de diriger ce travail. C'est, d'ailleurs, grâce à son soutien constant, à ses encouragements et à ses conseils précieux qu'il m'a prodigué, que j'ai pu mener ce travail à bien. Qu'il trouve ici ma profonde gratitude.*

*Je reste très reconnaissant à Monsieur **A.L.MARHOUNE** professeur à l'université de constantine, pour l'honneur qu'il m'a fait en acceptant de présider le jury de ce mémoire.*

*Je remercie aussi vivement monsieur **Z.MOHDEB** professeur à l'université de constantine pour l'intérêt qu'il a bien voulu porter à ce travail et d'avoir accepté de le juger.*

*Je remercie également et profondément monsieur **F.L.RAHMANI** maître de conférence à l'université de constantine pour l'intérêt qu'il a bien voulu porter à ce travail et d'avoir accepté de le juger.*

Je remercie aussi ma famille pour le soutien chaleureux apporté lors de la réalisation de ce travail.

Résumé

L'objet principal de la présente étude consiste en la prévision des futures valeurs des processus stochastiques par des modèles : autorégressifs, moyenne mobile, autorégressifs moyenne mobile et autorégressifs moyenne mobile Intégrés. Ces modèles peuvent comporter certaines propriétés (stationnarité, inversibilité, ...). Pour ce faire, on cherche tout d'abord un modèle adéquat pour la série Chronologique étudiée et on procède à une analyse systématique de celle-ci dans le but de déterminer ses caractéristiques et de voir si celles-ci correspondent bien à celles d'un des modèles mentionnés. En outre, on procède à l'estimation des paramètres du modèle considéré et on teste sa compatibilité avec l'un des modèles autorégressifs, le modèle obtenu est alors utilisé pour le calcul des prévisions des processus stochastiques considérés.

Mots clés

Séries chronologiques – Processus stochastiques – La représentation spectrale d'un processus stochastique – Le modèle ARIMA – Fonction de prévision – Fonction de vraisemblance – Fonction de somme de carrées des erreurs.

Abstract

The main purpose of this study is the prediction of future values of stochastic process by models: auto-regressive, moving average, auto-regressive and moving average, auto-regressive integrated moving average. These models may contain some features (stationarity, invertibility, ...). To do, this we first, sought a suitable model for the series chronological study and underwent a systematic analysis of it in order to determine its characteristics and see if they match those of the models above mentioned. In addition, we will estimate the model parameters considered and test its compatibility with one of the autoregressive models. The resulting model is then used to calculate the predictions of stochastic processes considered.

Key words

Time series – Random process – The spectral representation of a stationary process – ARIMA model – Forecast function – Likelihood function – Square errors function.

ملخص

إن الغرض الرئيسي من هذه الدراسة يتمثل في التنبؤ بالقيم المستقبلية للسياقات العشوائية بواسطة النماذج : ذاتية التقهقر، وسط متحرك، ذاتية التقهقر بوسط متحرك و ذاتية التقهقر بوسط متحرك منضم . هذه النماذج يمكنها أن تتضمن بعض الخواص (إستقرار , قابلية العكس ...). من أجل ذلك نبحت, قبل كل شيء, عن نموذج مناسب للسلسلة الزمنية المدروسة و نجري تحليلا ممنهجا لها بهدف تعيين مميزاتها و النظر فيما إذا كانت هذه الأخيرة تتفق فعلا و مميزات أحد النماذج المذكورة . بالإضافة إلى ذلك نجري تقديرا لوسطاء النموذج المعتبر و نختبر عدم تنافيه مع أحد النماذج ذاتية التقهقر والنموذج المتحصل عليه يستعمل عندئذ لحساب تنبؤات السياقات الإحتمالية المعتبرة.

كلمات مفتاحية

سلاسل زمنية – سياقات عشوائية – التمثيل الطيفي للسياقات العشوائية المستقرة – النماذج ذاتية التقهقر بوسط متحرك منضم – دالة التنبؤ – تابع المعقولية – تابع مجموع مربعات الأخطاء

TABLE DES MATIERES

Introduction	03
Chapitre I : Présentation des modèles linéaires stationnaires et non stationnaires	05
I.1- Définitions et notations	06
I.2- Présentation des processus linéaires stationnaires (AR ,MA ,ARMA).....	10
I.2.1- processus auto-régressifs AR(p).....	10
I.2.2- processus moyenne mobile MA(q).....	10
I.2.3- processus autorégressifs moyenne mobile ARMA(p,q).....	11
I.3- Fonctions d'auto-covariance et d'auto-corrélation	12
I.3.1-fonction d'auto-covariance d'un processus AR(p).....	12
I.3.2-fonction d'auto-covariance d'un processus MA(q).....	13
I.3.3-fonction d'auto-covariance d'un bruit blanc.....	13
I.4-Présentation spectrale du processus ARMA	14
I.4.1- densité spectrale pour un bruit blanc	14
I.4.2- densité spectrale pour les modèles linéaires.....	15
I.4.3- densité spectrale des processus ARMA, AR, MA	15
I.5- Présentation des processus linéaires non stationnaires (ARIMA).....	17
I.5.1- présentation du processus ARIMA	18
I.5.2- relation entre le processus et ses différences.....	18
I.5.3- formes explicites du processus ARIMA.....	19
Chapitre II : Etude de stationnarité, d'inversibilité et de l'effet du bruit additif de l'ARIMA et de l'ARIMA	25
II.1- stationnarité et inversibilité des processus AR ,MA ,ARMA.....	26
II.1.1- stationnarité.....	26
II.1.2- inversibilité.....	29
II.2- Processus ARIMA avec bruit additif.....	31
II.2.1- somme de deux processus moyennes mobiles.....	31
II.2.2- l'effet du bruit additif sur le processus ARIMA	32
Chapitre III : Estimation des paramètres des modèles ARMA et ARIMA	35
III .1- estimation des coefficients de l'ARMA.....	36
III.1.1- estimation par la méthode de maximum du vraisemblance (MMV).....	36
III.1.2- estimation par la méthode des moindres carrées (MMC).....	43
III.2- estimation des ordres p, q	47
III.3- comparaison des deux estimateurs MMV et MMC.....	47

III.4- consistance forte des estimateurs des paramètres de l'ARMA.....	49
III.5- comportement asymptotique des deux estimateurs	52
III.6- test de compatibilité du modèle.....	52
Chapitre IV : Prévission des modèles stochastiques.....	54
IV.1- Choix du prédicteur adéquat	55
IV.2-Prédicteur de l'erreur moyenne quadratique minimum	56
IV.2.1- déviation de l'erreur moyenne quadratique du prédicteur	56
IV.2.2- conformité du prédicteur de L'erreur moyenne quadratique minimale et de l'espérance conditionnelle	57
IV.2.3- procédure de calcul du prédicteur	57
IV.3- Trois formes explicites des futures valeurs du processus ARMA	58
IV.3.1- en fonction d'équation différentielle	58
IV.3.2- en tant que somme pondérée infinie de l'actuel et précédents bruits	59
IV.3.3- en tant que somme pondéré infinie d'observations précédentes et un bruit aléatoire.....	60
IV.4- Trois formes de base pour la prévission de l'ARMA(p,q)	60
IV.4.1- prévisions à partir d'équation différentielle	60
IV.4.2- prévisions en forme intégré de X_t	60
IV.4.3- prévisions en moyenne pondérée des observations précédentes ..	62
IV.5- Prévission de l'ARIMA(p,d,q)	62
IV.6- L'erreur du prédicteur à l'horizon 1	64
IV.6.1- calcul de l'erreur du prédicteur à l'horizon 1	64
IV.6.2- èvolution de la variance d'erreur pour les ARMA	66
IV.7- Intervalle de confiance pour le prédicteur	66
Chapitre V : Mise à jour des prévisions et exemple d'application	68
V.1- Mise à jour des prévisions	69
V.1.1- relation entre les prédicteurs aux temps t et $t + 1$	69
V.1.2- calcul de la valeur du bruit à l'instant $t + 1$	70
V.1.3- calcul des poids	70
V.2- Exemple d'application.....	71
V.2.1- mise au point du programme de calcul des prévisions.....	71
V.2.2- application numérique	74
ANNEXE	86
Programme de calcul	86

Introduction

La prévision des processus stochastiques représente l'un des intérêts majeurs de nos jours (prévision de température, des prix, des indices boursiers, de qualité industriel et de développement des populations...) du fait qu'elle aide à l'établissement de futures stratégies et à la prise de décision pour les buts de planification ...etc.

Dans le but de réaliser des prévisions, plusieurs méthodes ont été envisagées par nombre d'auteurs, notamment, la publication en 1927 par U. Yule d'un article promoteur intitulé : « *the Method of Investigating periodicities in disturbed series with special reference to wolfer's sunspot numbers* » dans lequel il a introduit la notion percutante de littérature des modèles auto-régressifs E. Slutsky publia pour sa part, dans la même année que Yule, un article en langue russe où il a introduit la notion de modèles moyenne mobile; cependant, cet article n'a été traduit en langue anglaise qu'en 1937. Ces propositions ont été par la suite amplifiées par les travaux de G.E.P. Box & G.M. Jenkins sous l'apparition en 1976 d'un livre intitulé : « *Time Series analysis Forecasting and Control* » qui traite, entre autres, les modèles ARIMA relatif aux différences des processus autorégressifs moyenne mobile. Enfin, l'avantage de ces méthodes découle du fait qu'elles soient heuristiques et par conséquent accessibles à un éventail très large d'utilisateurs et praticiens.

Par ailleurs, les modèles de prévisions utilisés dans ce domaine, comportent deux types de représentations :

- linéaire les modèles Auto-régressifs (AR), Moyen mobile (MA), Auto régressifs Moyens mobile (ARMA), Auto-régressifs Moyens mobile intégré (ARIMA)
- non linéaire (modèles auto-régressifs conditionnellement hétéroscédastiques (ARCH), modèles autorégressifs conditionnellement hétéroscédastiques généralisé (GARCH) ...etc.)

Pour notre part, on s'est intéressé, dans ce travail, au développement du premier point. Dès lors, notre contribution se trouve subdivisée en Cinq chapitres qui s'articulent de façon logique et progressive allant du particulier au général et du simple au compliqué, répondant ainsi aux exigences et préoccupations du moment.

Dans le premier chapitre, consacré à la présentation de deux types de processus stochastiques linéaires, à savoir, les modèles linaires stationnaires dans lesquels on trouve le modèle auto-régressif, le modèle moyenne mobile et le modèle auto-régressif moyenne mobile (AR, MA, ARMA) et celui du modèle linéaire non stationnaire où l'on présente le modèle autorégressif moyenne mobile intégrée (ARIMA).

Dans le deuxième chapitre, on s'est consacré à l'étude de la stationnarité, de l'inversibilité et de la proposition de conditions suffisantes pour ces deux propriétés. On s'est intéressé, aussi, à l'étude de processus ARIMA avec bruit additif.

Dans le troisième chapitre, on a abordé l'estimation des paramètres du modèle ARMA en utilisant deux méthodes des plus usuelles, La première étant celle du maximum de vraisemblance (MMV) et la seconde est celle des moindres carrés(MMC). Ensuite, on a procédé à une étude comparative des résultats des deux méthodes et leur comportement asymptotique.

Dans le quatrième chapitre, destiné à la prévision utilisant les processus ARMA et ARIMA, on a utilisé le prédicteur de l'erreur moyenne quadratique minimale, ainsi que la fonction de prévision du modèle ARMA basée sur l'erreur moyenne quadratique minimale. Enfin, on a procédé à la prévision des processus ARIMA (p,d,q) et finissant le chapitre par le calcul de la forme générale de l'erreur du prédicteur à l'horizon l pour le modèle ARMA (p,q) .

Dans le cinquième et dernier chapitre, on a présenté une méthode de mise à jour des prévisions et procéder à la mise au point d'un programme de calcul efficace et performant donnant les prévisions selon la démarche adoptée dans ce mémoire. Pour confirmer les qualités de ce programme, on a proposé une application numérique sur ordinateur à une série chronologique donnant les valeurs du DOW-JONES de janvier 1999 à septembre 2009.

Chapitre I : Présentation des modèles linéaires stationnaires et non stationnaires

I.1- Définitions et notations

On note par :

$\mu_X(t)$ La fonction moyenne du processus X_t .

$\gamma_X(r, s) = cov(X_r, X_s) = E[(X_r - E(X_r))(X_s - E(X_s))]$ La Fonction d'auto-covariance du processus X_t .

$\rho_X(k) = \frac{\gamma_X(k)}{\gamma_X(0)}$ La fonction d'auto-corrélation du processus X_t .

$\tilde{X}_t = X_t - \mu_X(t)$ Le processus centré avec $\mu_X(t) = E(X_t)$.

ε_t Le bruit blanc.

B L'opérateur retard ($B^j X_t = X_{t-j}$).

$L(\beta)$ La fonction de vraisemblance.

$S^2(\cdot)$ La somme des carrés des erreurs.

$\hat{\beta}_n^{MV}$ L'estimateur du maximum de vraisemblance et $\hat{\beta}_n^{MC}$ L'estimateur des moindres carrés pour les coefficients des deux polynômes du processus ARMA pour un échantillon de taille n .

l L'horizon de prévision.

$\hat{X}_t(l)$ La fonction de prévision.

- **Définition 01** : la prévision d'un processus stochastique est la détermination de ce qui se passera prochainement, avec plus au moins de certitude.

- **Définition 02** : une série chronologique (ou série temporelle) est une suite d'observations, d'une certaine grandeur, effectuées à intervalles réguliers, prises séquentiellement dans le temps, débutant à un instant (t) .

- **Définition 03** : un processus stochastique est une famille de variables aléatoires notée $(X_t; t \in T)$ et défini sur un même espace (Ω, F, P) ; ces variables peuvent représenter une certaine grandeur qui varie dans le temps de façon aléatoire dont l'indicateur représente le temps.

- **Définition 04** : un processus (ε_t) i.i.d (identi.iindep.distri), constitue ce qu'on appelle un bruit blanc de moyenne 0 et de variance (σ_ε^2) ssi il est d'espérance nulle $E(\varepsilon_t) = 0$ et de fonction d'auto-covariance définie par :

$$\gamma_\varepsilon(h) = \text{cov}(\varepsilon_{t+h}; \varepsilon_t) = \begin{cases} \sigma_\varepsilon^2 & \text{si } h=0 \\ 0 & \text{sinon} \end{cases}$$

- **Définition 05/ Processus linéaire** : un processus X_t est dit linéaire si on peut l'exprimer en fonction du bruit blanc à travers un filtre linéaire.

Cela dit, on peut représenter les processus linéaires de deux façons équivalentes :

a) Le processus X_t en fonction du bruit blanc

$$X_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \dots$$

$$= \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t-j} = \Psi(B) \varepsilon_t \dots \dots (1.1) \quad \text{où} \quad \Psi(B) = \sum_{j=0}^{+\infty} \psi_j B^j \text{ et } \psi_0 = 1$$

b) Le processus X_t en fonction de son passé avec bruit additif :

$$\begin{aligned} X_t &= \varepsilon_t + \pi_1 X_{t-1} + \pi_2 X_{t-2} + \dots \\ &= \varepsilon_t + \sum_{j=1}^{+\infty} \pi_j X_{t-j} \end{aligned}$$

$$\Leftrightarrow \varepsilon_t = \Pi(B)X_t \dots \dots (1.2) \quad \text{où} \quad \Pi(B) = 1 - \sum_{j=1}^{+\infty} \pi_j B^j$$

- Equivalence des deux représentations (a) et (b) :

Cette équivalence découle du fait que,

$$(a) \wedge (b) \Rightarrow \Psi(B)\Pi(B)X_t = X_t$$

$$\Leftrightarrow \Psi(B)\Pi(B) = 1$$

$$\Leftrightarrow \Psi(B) = [\Pi(B)]^{-1}$$

et par suite,

$$X_t = \Psi(B)\varepsilon_t$$

$$\Leftrightarrow \varepsilon_t = \Psi^{-1}(B)X_t$$

$$\Leftrightarrow \varepsilon_t = \Pi(B)X_t$$

- Définition 06/ Stationnarité stricte : le processus stochastique X_t est dit stationnaire au sens strict si pour toute valeur de n les caractéristiques du n -uple $(X_{t_1}; \dots; X_{t_n})$ sont invariantes par rapport à n'importe quelle translation suivant l'axe du temps, c-à-d, les lois de $(X_{t_1}; \dots; X_{t_n})$ et $(X_{t_1+h}; \dots; X_{t_n+h})$ sont pour tout $(h \in \mathbb{Z})$ identiques.

- **Définition 07/ Stationnarité faible** : le processus stochastique X_t est dit stationnaire au sens faible (où du second ordre) si ses deux moments des premiers ordres sont finis et invariants dans le temps, autrement dit s'ils vérifient les deux propriétés suivantes :

- la moyenne $\mu_X(t) < +\infty$ est indépendante du temps t .

- la covariance $\gamma_X(r,s) = \gamma_X(h)$ est indépendante du temps t
 $\forall h = r - s$ où $(r,s) \in \mathbb{Z}^2$

- **Définition 08 / Causalité** : un processus stochastique X_t est dit causal s'il existe une suite de constantes $\{\psi_j\}_{j \geq 0}$ tel que :

$$\sum_{j=0}^{+\infty} |\psi_j| < \infty \quad \text{et} \quad X_t = \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t-j} \quad \text{où } t = 0, \pm 1; \pm 2; \dots$$

I.2- Présentation des processus linéaires stationnaires (AR, MA, ARMA)

I.2.1- Processus auto-régressifs AR(p)

Soit X_t un processus stochastique de fonction moyenne $\mu_X(t)$. S'il existe un polynôme Φ d'ordre p et un bruit blanc ε_t de moyenne nulle $E(\varepsilon_t)=0$ & de variance $var(\varepsilon_t)=\sigma_\varepsilon^2$ vérifiant $\Phi(B)\tilde{X}_t = \varepsilon_t$, alors X_t est un processus autorégressif.

Autrement dit si $\Phi(B)=1-\sum_{j=1}^p \phi_j B^j$ et $\Phi(B)\tilde{X}_t = \varepsilon_t$ le processus X_t est un AR(p) et il peut être représenté sous forme :

$$\tilde{X}_t = \phi_1 \tilde{X}_{t-1} + \dots + \phi_p \tilde{X}_{t-p} + \varepsilon_t \dots \dots \dots (1.3)$$

Remarque : Cette appellation (auto- régressive) vient du fait que l'expression du processus est une régression multiple de sa valeur courante et de ses (p) valeurs précédentes; c-à-d on peut utiliser ce modèle dans le cas où l'actuelle valeur du processus dépend de ses (p) valeurs précédentes.

I.2.2- Processus moyenne mobile MA(q)

S'il existe un polynôme Θ d'ordre q et un bruit blanc ε_t de moyenne nulle $E(\varepsilon_t)=0$ & de variance $var(\varepsilon_t)=\sigma_\varepsilon^2$ (vérifiant $\hat{X}_t = \Theta(B)\varepsilon_t$) alors X_t forme un processus moyenne mobile.

Autrement dit si $\Theta(B)=\sum_{j=1}^q \theta_j B^j$ et $\hat{X}_t = \Theta(B)\varepsilon_t$ le processus X_t est un MA(q) qui se présente sous forme :

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots \dots \dots + \theta_q \varepsilon_{t-q}$$

$$\Leftrightarrow \tilde{X}_t = X_t - \mu = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots \dots \dots + \theta_q \varepsilon_{t-q} \dots \dots \dots (1.4)$$

On utilise ce modèle dans le cas où les valeurs du processus constituent des perturbations autour d'une valeur moyenne μ .

I.2.3- Processus autorégressif moyenne mobile ARMA (p, q)

Pour mieux représenter le processus stochastique; il est parfois utile de combiner les deux processus : AR(p) (auto-régressif) et MA(q) (moyenne mobile), ce qui donne le processus Autorégressif Moyenne mobile ARMA(p, q).

Donc pour que le processus X_t soit un ARMA(p, q) il faut et il suffit qu'il existe deux polynômes Φ, Θ et un bruit blanc ε_t , de sorte que $\Phi(B)\tilde{X}_t = \Theta(B)\varepsilon_t$

Autrement dit si $\Theta(B) = \sum_{j=1}^q \theta_j B^j$, $\Phi(B) = 1 - \sum_{j=1}^p \phi_j B^j$ et $\tilde{X}_t = \Theta(B)\varepsilon_t$

le processus X_t est un ARMA(p, q) qui se représente sous forme :

$$\tilde{X}_t = \phi_1 \tilde{X}_{t-1} + \phi_2 \tilde{X}_{t-2} + \dots + \phi_p \tilde{X}_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

Si en outre le processus X_t est centré ($\mu=0$) alors :

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

Autrement, on modélise le processus $\tilde{X}_t = X_t - \mu$ au lieu du processus X_t et l'on écrit :

$$\tilde{X}_t = \phi_1 \tilde{X}_{t-1} + \dots + \phi_p \tilde{X}_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \dots \dots \dots (1.5)$$

I.3- Fonctions d'auto-covariance et d'auto-corrélation

Supposons dans ce qui suit, que X_t est un processus stochastique stationnaire centré de fonction d'auto covariance γ_X .

I.3.1- Fonction d'auto covariance d'un processus AR(p)

Si X_t est un processus autorégressif, alors :

$$\begin{aligned} \forall k \geq 1 \quad \gamma_X(k) &= cov(X_{t+k}, X_t) \\ &= E(X_{t+k} \cdot X_t) \\ &= E(X_{t-k} \cdot X_t) \text{ (à cause de la stationnarité du processus)} \end{aligned}$$

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

$$\Rightarrow E(X_{t-k} \cdot X_t) = \phi_1 E(X_{t-k} \cdot X_{t-1}) + \phi_2 E(X_{t-k} \cdot X_{t-2}) + \dots + \phi_p E(X_{t-k} \cdot X_{t-p}) + E(X_{t-k} \cdot \varepsilon_t)$$

$$\Rightarrow \gamma_k = \phi_1 \gamma_X(k-1) + \phi_2 \gamma_X(k-2) + \dots + \phi_p \gamma_X(k-p) \dots \dots \dots (1.6)$$

(relation entre autos-covariances).

en divisant cette dernière relation (1.6) par γ_0 on obtient :

$$\rho_X(k) = \phi_1 \rho_X(k-1) + \phi_2 \rho_X(k-2) + \dots + \phi_p \rho_X(k-p) \dots \dots \dots (1.7)$$

La quantité ϕ_k notée ϕ_{kk} en tant que fonction de k est appelée fonction d'auto-corrélation partielle. Cette fonction constitue alors un outil utile pour reconnaître l'ordre (p) d'un processus AR(p) puisque la fonction d'auto-corrélation est la dernière pondération d'un AR(p).

La quantité ϕ_{kk} définit, alors, l'ordre du processus autorégressif et l'on a :

$$\begin{cases} \phi_{kk} = 0 & \forall k \geq p+1 \\ \phi_{kk} \neq 0 & \text{si : } k = p \end{cases}$$

I.3.2- Fonction d'auto-covariance d'un processus MA(q)

Si X_t est un processus moyenne mobile, alors :

$$\begin{aligned}
 \gamma_X(k) &= \text{cov}(X_{t+k}; X_t) \\
 &= E(X_{t+k} \cdot X_t) \\
 &= E\left[\left(\sum_{j=0}^q \theta_j \varepsilon_{t+k-j}\right)\left(\sum_{i=0}^q \theta_i \varepsilon_{t-i}\right)\right] \\
 &= \sum_{j=0}^q \sum_{i=0}^q \theta_j \theta_i E(\varepsilon_{t-i} \varepsilon_{t+k-j}) \\
 &= \begin{cases} \sigma_\varepsilon^2 \sum_{j=0}^q \theta_j \theta_{j+|k|} & \text{si } k \leq q \\ 0 & \text{sinon} \end{cases} \dots\dots\dots(1.8)
 \end{aligned}$$

Dans le cas où X_t n'est pas centré on le remplace par $\tilde{X}_t = X_t - E(X_t)$ et γ_X sera égale à $\gamma_{\tilde{X}}$.

I.3.3- Fonction d'auto-covariance d'un bruit blanc

A partir de la définition du bruit blanc on écrit sa fonction d'auto-covariance est alors :

$$\gamma_\varepsilon(h) = \begin{cases} \sigma_\varepsilon^2 & \text{si } h = 0 \\ 0 & \text{sinon} \end{cases} \dots\dots\dots(1.9)$$

I.4- Présentation spectrale du processus ARMA

- **Définition/ Densité spectrale** : soit X_t un processus stationnaire de moyenne nulle et de fonction d'auto-covariance $\gamma_X(\cdot)$ satisfaisant $\sum_{h=-\infty}^{+\infty} |\gamma_X(h)| < \infty$, la densité spectrale du processus X_t est alors définie par :

$$f_X(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} e^{-ih\lambda} \gamma_X(h) \quad \text{où } -\infty < \lambda < +\infty \quad \dots\dots\dots(1.10)$$

I.4.1- Densité spectrale pour un bruit blanc

Le bruit blanc étant un processus stationnaire de moyenne nulle alors, sa densité spectrale s'écrit sous forme :

$$\begin{aligned} f_\varepsilon(\lambda) &= \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} e^{-ih\lambda} \gamma_\varepsilon(h) \\ &= \frac{1}{2\pi} \gamma_\varepsilon(0) \\ &= \frac{\sigma_\varepsilon^2}{2\pi} \end{aligned}$$

Proposition I.1 :

Soit X_t un processus stationnaire de moyenne nulle et de densité spectrale $f_X(\lambda)$.

Si $(\psi_j)_{j \geq 0}$ est une suite numérique vérifiant $\sum_{j=0}^{+\infty} |\psi_j| < +\infty$, alors le processus

$Y_t = \sum_{j=-\infty}^{+\infty} \psi_j X_{t-j}$ est stationnaire de moyenne nulle et de densité spectrale :

$$f_Y(\lambda) = \left| \Psi(e^{-i\lambda}) \right|^2 f_X(\lambda) = \Psi(e^{-i\lambda}) \Psi(e^{i\lambda}) f_X(\lambda)$$

I.4.2- Densité spectrale pour les modèles linéaires

Etant donné le modèle linéaire :

$$X_t = \Psi(B)\varepsilon_t \quad \text{avec} \quad \Psi(B) = \sum_{j=-\infty}^{+\infty} \psi_j B^j$$

et compte tenu de la proposition (I.1) la densité du modèle a pour fonction :

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

I.4.3- Densité spectrale des processus ARMA, AR, MA

Dans cette partie supposons que le processus X_t est centré.

A- Le processus ARMA :

Le processus ARMA(p, q) peut s'écrire sous forme :

$$\begin{aligned} \Phi(B)X_t &= \Theta(B)\varepsilon_t \\ \Leftrightarrow X_t &= \Phi^{-1}(B)\Theta(B)\varepsilon_t \quad \text{avec} \quad \Phi(z) \neq 0 \quad \& \quad \forall z : |z| \leq 1 \\ &= \Psi(B)\varepsilon_t; \text{ avec} : \Psi(B) = \Phi^{-1}(B)\Theta(B) = \sum_{j=0}^{+\infty} \psi_j B^j \\ &= \left(\sum_{j=0}^{+\infty} \psi_j B^j \right) \varepsilon_t \end{aligned}$$

il en résulte alors, compte tenu de la proposition (I.1), la densité spectrale du processus ARMA :

$$f_X(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left| \frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})} \right|^2$$

B- Le processus AR :

Le processus AR(p) peut s'écrire sous forme :

$$\begin{aligned}\Phi(B)X_t &= \varepsilon_t \\ \Leftrightarrow X_t &= \Phi^{-1}(B)\varepsilon_t \text{ avec } \Phi(z) \neq 0 \forall z : |z| \leq 1 \\ &= \Psi(B)\varepsilon_t ; \text{ avec } : \Psi(B) = \Phi^{-1}(B) = \sum_{j=0}^{+\infty} \psi_j B^j \\ &= \left(\sum_{j=0}^{+\infty} \psi_j B^j \right) \varepsilon_t\end{aligned}$$

il en résulte alors, compte tenu de la proposition (I.1), la densité spectrale du processus AR :

$$f_X(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \frac{1}{|\Phi(e^{-i\lambda})|^2}$$

C- Le processus MA (q) :

Le processus MA(q) peut s'écrire sous forme :

$$\begin{aligned}X_t &= \Theta(B)\varepsilon_t \\ \Leftrightarrow X_t &= \Theta(B)\varepsilon_t \text{ (sous la condition } \Theta(z) \neq 0 \forall z : |z| \leq 1) \\ &= \Psi(B)\varepsilon_t ; \text{ avec } : \Psi(B) = \Theta(B) \\ &= \sum_{j=0}^q \theta_j B^j = \sum_{j=0}^q \psi_j B^j + \sum_{j=q+1}^{+\infty} \psi_j B^j (\psi_j = \theta_j, \forall j = 0, \dots, q \text{ et } \psi_j = 0 \forall j > q) \\ \Rightarrow X_t &= \left(\sum_{j=0}^{+\infty} \psi_j B^j \right) \varepsilon_t\end{aligned}$$

il en résulte alors, compte tenu de la proposition (I.1), la densité spectrale du processus MA :

$$f_X(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} |\Theta(e^{-i\lambda})|^2$$

I.5- Présentation des processus linéaires non-stationnaires(ARIMA)

On a étudié précédemment, les modèles AR, MA, ARMA qui sont des modèles linéaires stationnaires (les moments d'ordre un et deux sont invariants dans le temps) et qui peuvent donner une bonne représentation de certaines séries chronologiques d'une part. d'autre part, il existe d'autres séries où l'un des deux premiers moments ou les deux à la fois varient dans le temps (série changeant de niveau) ce qui provoque la non-stationnarité, et l'on ne pourra pas utiliser les modèles AR, MA, ARMA pour les représenter.

Une classe des séries chronologiques dont l'un des deux premiers moments ou les deux varient avec le temps, exhibe une homogénéité dans la variation du niveau dans le sens où des parties de la série ont la même moyenne ou la même moyenne avec tendance.

Dans le cas où la non-stationnarité est causée par l'existence des racines unitaires et les différences stationnaires, au lieu de modéliser le processus on modélise ses différences et on extrait cette modélisation.

Ce changement dans le processus crée alors, un nouveau processus (processus des différences). Ainsi, la présentation du processus des différences par un modèle ARMA est devient une présentation du processus initial par le modèle ARIMA.

I.5.1- Présentation du processus ARIMA

S'ils existent deux polynômes Φ_1, Θ d'ordre p, q respectivement, un entier d et un bruit blanc ε_t de moyenne nulle $E(\varepsilon_t) = 0$ et de variance $Var(\varepsilon_t) = \sigma_\varepsilon^2$ vérifiant :

$$\Phi_1(B)(1-B)^d X_t = \Theta(B)\varepsilon_t \dots \dots \dots (1.11)$$

alors X_t est un processus Autorégressif moyen mobile intégré (ARIMA).

Autrement dit, si $\Phi_1(B) = \sum_{j=0}^p \phi_{1,j} B^j$ et $\Theta(B) = \sum_{j=1}^q \theta_j B^j$ avec $\Phi_1(B)(1-B)^d X_t = \Theta(B)\varepsilon_t$ tel que les racines de $\Phi_1(B)$ sont à l'extérieur du disque d'unité, le processus X_t est un ARIMA(p, d, q).

Les différences d'ordre d pour un processus ARIMA(p, d, q) représente alors, de nouveaux processus ARMA (p, q).

$$\Phi(B)\tilde{X}_t = \Theta(B)\varepsilon_t$$

$$\Leftrightarrow \Phi_1(B)(1-B)^d X_t = \Theta(B)\varepsilon_t$$

$$\Leftrightarrow \Phi_1(B)Y_t = \Theta(B)\varepsilon_t \quad \text{Avec} \quad Y_t = (1-B)^d X_t \quad \text{et} \quad \Phi_1(B) \neq 0, \quad \forall B/|B| \leq 1$$

d'où Y_t représente un processus ARMA (p, q) stationnaire.

I.5.2- Relation entre le processus et ses différences

En posant, $Y_t = (1-B)^d X_t \Rightarrow X_t = (1-B)^{-d} Y_t$

et $S = (1-B)^{-1}$

on écrit : $X_t = S^d Y_t \dots \dots \dots (1)$

Le développement fini d'ordre m est alors pour $d=1$:

$$X_t = Y_t + Y_{t-1} + \dots + Y_{m+1} + X_m \dots \dots \dots (1.12)$$

Preuve :

sachant que :

$$S_m = \sum_{j=0}^{m-1} B^j = 1 + B + B^2 + B^3 + \dots + B^{m-1}$$

$$= \frac{1 - B^m}{1 - B}$$

et

$$X_t = \frac{1}{1 - B} Y_t$$

$$\Rightarrow X_t = \frac{(1 - B^{t-m})}{(1 - B)(1 - B^{t-m})} Y_t$$

$$\Rightarrow X_t = \frac{S_{t-m}}{(1 - B^{t-m})} Y_t$$

on trouve :

$$(1 - B^{t-m}) X_t = S_{t-m} Y_t$$

d'où le processus X_t s'écrit en tant que somme de nombre fini de termes d'un processus stationnaire plus une valeur initiale X_m .

I.5.3- Formes explicites du processus ARIMA

En ce qui concerne la présentation de l'ARIMA on peut considérer trois formes explicites pour le modèle général où chacune de ces formes présente un aspect spécial utile dans l'un des problèmes statistiques, où l'actuel valeur de processus peut s'exprimer :

1\ en fonction des précédentes valeurs du processus et l'actuelle valeur du bruit blanc ainsi que ses précédentes valeurs $X_t = H(X_{t-i}, \mathcal{E}_t, \mathcal{E}_{t-j})_{p \geq i \geq 1, q \geq j \geq 1}$.

2\ en fonction de l'actuelle valeur su processus et les précédentes valeurs du bruit blanc $X_t = H(\mathcal{E}_t, \mathcal{E}_{t-i})_{i \geq 1}$.

3\ en fonction des précédentes valeurs du processus et de l'actuelle valeur du bruit $X_t = H(X_{t-i}, \mathcal{E}_t)_{q \geq i \geq 1}$.

1\ Forme d'équation différentielle du modèle :

Dans cette forme on exprime les X_t en fonction des X_{t-j} et ε_{t-k} avec $j \geq 1$ et $k \geq 0$ ce qui donne la représentation suivante :

$$\Phi_1(B)(1-B)^d X_t = \Theta(B)\varepsilon_t \dots \dots \dots (1.13)$$

$$\text{Avec } \Phi_1(B) = \sum_{j=0}^q \phi_{1,j} B^j \quad \text{et} \quad \Theta(B) = \sum_{j=1}^q \theta_j B^j \quad \text{et}$$

$$\Phi(B) = \Phi_1(B)(1-B)^d = \sum_{i=0}^{p+d} \varphi_i B^i$$

ou sous forme plus explicite selon la relation

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_{p+d} X_{t-p-d} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

2\ Forme du processus en fonction du bruit Blanc :

On sait que tout modèle linéaire peut être écrit sous forme :

$$X_t = \varepsilon_t + \sum_{j=1}^{+\infty} \psi_j \varepsilon_{t-j} \dots \dots \dots (1.14)$$

et puisque l'ARIMA est un modèle linéaire on peut alors le représenter sous la forme précédente qui est elle-même celle d'un processus en fonction du bruit blanc.

- Expression générale des poids ψ_j :

On écrivant ,

$$\begin{aligned} \begin{cases} X_t = \Psi(B)\varepsilon_t \\ \Phi(B)X_t = \Theta(B)\varepsilon_t \end{cases} &\Rightarrow \Phi(B)\Psi(B)\varepsilon_t = \Theta(B)\varepsilon_t \\ &\Rightarrow \Phi(B)\Psi(B) = \Theta(B) \\ &\Rightarrow \left(\sum_{j=0}^{p+d} \varphi_j B^j \right) \left(\sum_{k=0}^{+\infty} \psi_k B^k \right) = \sum_{i=0}^q \theta_i B^i \\ &\Rightarrow \sum_{j=0}^{p+d} \sum_{k=0}^{+\infty} \psi_k \varphi_j B^{j+k} = \sum_{i=0}^q \theta_i B^i \end{aligned}$$

et en égalisant les coefficients similaires des deux polynômes terme à terme on trouve :

$$\psi_j = \varphi_1 \psi_{j-1} + \varphi_2 \psi_{j-2} + \dots + \varphi_{p+d} \psi_{j-p-d} - \theta_j$$

$$\text{Sachant que } \psi_0 = 1; \psi_j = 0 \quad \forall j \leq 0 \quad \text{et} \quad \theta_j = 0 \quad \forall j \geq q+1$$

-Forme tronquée du modèle de choc aléatoire

La forme d'équation différentielle du modèle ARIMA étant :

$$\Phi(B)X_t = \Theta(B)\varepsilon_t$$

C'est est une équation différentielle de solution :

$$X_t = C_k(t-k) + I_k(t-k) \dots \dots \dots (1.15)$$

où :

$C_k(t-k)$ est solution générale de l'équation :

$$\Phi(B)C_k(t-k) = 0$$

et où $I_k(t-k)$ est fonction quelconque vérifiant la relation :

$$\Phi(B)I_k(t-k) = \Theta(B)\varepsilon_t$$

$$\Rightarrow I_k(t-k) = \sum_{j=0}^{t-k-1} \psi_j \varepsilon_{t-j} \quad \forall t \geq k+1$$

d'où :

$$X_t = \sum_{j=0}^{t-k-1} \psi_j \varepsilon_{t-j} + C_k(t-k)$$

- La fonction complémentaire en tant qu'espérance conditionnelle :

Une des propriétés de la fonction complémentaire C_k est qu'elle varie en changeant d'origine (k):

$$C_k(t-k) = C_{k-m}(t-k+m) + \sum_{j=k-m+1}^k \psi_{t-j} \varepsilon_j$$

Preuve :

D'après la relation (1.14) on a :

$$X_t = \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t-j} = \sum_{j=0}^{t-k-1} \psi_j \varepsilon_{t-j} + \sum_{j=t-k}^{+\infty} \psi_j \varepsilon_{t-j} \dots\dots\dots(1)$$

or la représentation $\varphi(B)X_t = \theta(B)\varepsilon_t$ est une équation différentielle de solutions :

$$X_t = C_k(t-k) + \sum_{j=0}^{t-k-1} \psi_j \varepsilon_{t-j} \dots\dots (2)$$

d'où à partir d'une comparaison des relations (1) et (2) on a :

$$(1) \text{ et } (2) \Rightarrow C_k(t-k) = \sum_{j=t-k}^{+\infty} \psi_j \varepsilon_{t-j} \dots\dots\dots(3)$$

or, la relation (1.14) peut s'écrire sous forme :

$$X_t = \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t-j} = \sum_{j=0}^{t-k-1} \psi_j \varepsilon_{t-j} + \sum_{j=t-k}^{t-k+m-1} \psi_j \varepsilon_{t-j} + \sum_{j=t-k+m}^{+\infty} \psi_j \varepsilon_{t-j} \dots\dots\dots (4)$$

on tire à partir de (3) et (4) : $C_k(t-k) = C_{k-m}(t-k+m) + \sum_{j=k-m+1}^k \psi_{t-j} \varepsilon_j$ il en résulte que C_k change avec l'origine k .

Une autre propriété fournie par la forme tronquée est celle que la fonction complémentaire coïncide avec l'espérance conditionnelle de X_t sachant X_k et l'on a :

$$E\left(\frac{X_t}{X_k}\right) = C_k(t-k)$$

Preuve :

$$\begin{aligned} E\left(\frac{X_t}{X_k}\right) &= E\left(\frac{C_k(t-k) + \sum_{j=0}^{t-k-1} \psi_j \varepsilon_{t-j}}{\sum_{j=0}^{+\infty} \psi_j \varepsilon_{k-j}}\right) \\ &= E\left(\frac{c_k(t-k)}{\{\varepsilon_{k-j}; j \geq 0\}}\right) + \sum_{j=0}^{t-k-1} \psi_j E\left(\frac{\varepsilon_{t-j}}{\{\varepsilon_{k-j}; j \geq 0\}}\right) \\ &= C_k(t-k) \text{ (Car } \varepsilon_t \text{ i.i.d } N(0, \sigma_\varepsilon^2)). \end{aligned}$$

3\ Forme inversée du modèle :

La forme inversée du modèle est l'expression du bruit blanc en fonction des termes du processus qui s'écrit :

$$\varepsilon_t = \pi(B)X_t \dots \dots \dots (1.15) \quad (\Phi(B) \neq 0 \text{ et } \forall z : |z| \leq 1).$$

-L'expression général des $\{\pi_j\}_{j \geq 0}$:

Partant de :

$$\begin{cases} \Phi(B)X_t = \Theta(B)\varepsilon_t \dots \dots \dots (1) \\ \varepsilon_t = \Pi(B)X_t \dots \dots \dots (2) \end{cases}$$

et en portant (2) dans (1) on obtient :

$$\Phi(B)X_t = \Theta(B)\Pi(B)X_t$$

$$\Leftrightarrow \Phi(B) = \Theta(B)\Pi(B)$$

$$\Leftrightarrow 1 - \sum_{j=1}^{p+d} \varphi_j B^j = \left(1 - \sum_{k=1}^q \theta_k B^k\right) \times \left(1 - \sum_{i=1}^{+\infty} \pi_i B^i\right)$$

Ce qui permet d'écrire :

$$\pi_j = \theta_1 \pi_{j-1} + \theta_2 \pi_{j-2} + \dots + \theta_q \pi_{j-q} + \varphi_j \quad \text{pour } j > 0 \quad \text{avec}$$

$$\pi_j = 0 : \forall j < 0 \quad , \quad \varphi_j = 0 : \forall j > p+d$$

or, pour $j > p+d \wedge q$ les poids π_j représentent les solution de l'équation différentielle :

$$\Theta(B)\pi_j = 0$$

Remarque :

Le processus auto-régressif moyenne mobile (ARMA) possède les trois représentations précédentes en remplaçant seulement, les $\varphi_j, j \geq 0$ par les $\phi_j, j \geq 0$.

Chapitre II : Etude de stationnarité, d'inversibilité et de l'effet du bruit additif de l'ARMA & de l'ARIMA

II.1- Stationnarité et inversibilité des processus AR, MA , ARMA

II.1.1- Stationnarité

Un processus X_t est dit stationnaire si X_t possède certaines propriétés statistiques similaires à celle du processus X_{t+h} , $\forall h \in \mathbb{Z}$.

si parmi les propriétés du processus on ne s'intéresse qu'aux moments d'ordre un et deux on aura alors, la stationnarité d'ordre deux ou stationnarité faible.

Proposition II.1 :

Soit Y_t un processus stationnaire au sens faible de moyenne μ_Y nulle et de fonction d'auto-covariance γ_Y .

Si $\sum_{j=-\infty}^{j=+\infty} |\psi_j| < \infty$ alors le processus $X_t = \sum_{j=-\infty}^{+\infty} \psi_j Y_{t-j} = \Psi(B)Y_t$ est stationnaire au sens faible de moyenne μ_X nulle et de fonction d'auto-covariance :

$$\gamma_X(h) = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \psi_k \psi_j \gamma_Y(h+k-j).$$

Proposition II.2 :

Tout processus causal d'un bruit blanc est stationnaire.

Théorème II.1 :

Un processus ARMA(p, q) dans lequel les polynômes Φ et Θ n'ont pas de zéros communs est causal si et seulement si le polynôme $\Phi(B)$ ne possède pas de zéros à l'intérieur du disque unitaire.

Preuve :

Si le processus est un ARMA(p, q) alors on peut l'écrire sous forme :

$$\Phi(B)X_t = \Theta(B)\varepsilon_t \quad \text{où } \Phi(B) = \sum_{j=1}^p \phi_j B^j$$

et l'on note par A_i les racines de $\Phi(B)$ alors, et par suite :

$$\begin{aligned}\Phi(B) &= \prod_{i=1}^p (1 - A_i B) \\ \Rightarrow \Phi^{-1}(B) &= \frac{1}{\prod_{i=1}^p (1 - A_i B)} \\ &= \sum_{j=1}^p \frac{K_i}{(1 - A_i B)}\end{aligned}$$

$$\frac{1}{(1 - A_i B)} = \sum_{k=1}^{+\infty} A_i^k B^k \dots\dots\dots(1)$$

Si $\forall i, (|A_i| > 1)$ alors la série (1) est convergente.

De la représentation précédente du processus on peut extraire la nouvelle représentation :

$$\begin{aligned}\Phi(B)X_t = \Theta(B)\varepsilon_t &\Rightarrow X_t = \Phi^{-1}(B)\Theta(B)\varepsilon_t \\ \text{et} & \quad X_t = \Psi(B)\varepsilon_t\end{aligned}$$

puisque Φ^{-1} est une série convergente et Θ d'ordre fini. Ainsi, Ψ est une série convergente, d'où on conclut que le processus X_t est causal de bruit blanc ε_t .

de plus le processus est stationnaire d'après la proposition (II.2).

1/Le processus auto-régressif AR (p) :

Le processus autorégressifs peut être présenté sous forme :

$$\Phi(B)\tilde{X}_t = \varepsilon_t \Leftrightarrow \Phi(B)\tilde{X}_t = 1.\varepsilon_t$$

qui est celle d'un ARMA avec $\Theta(B) = 1$, autrement dit le AR(p) est un ARMA(p,0).

d'où compte tenu de la proposition (II.2) et du théorème (II.1) le processus X_t est stationnaire si et seulement si le polynôme Φ n'a pas de zéros à l'intérieur du disque d'unité.

2/Le processus moyenne mobile MA(q) :

Le processus moyenne mobile MA (q) peut être présenté sous forme :

$$\begin{aligned}\tilde{X}_t &= \Theta(B)\varepsilon_t = \sum_{j=0}^q \theta_j \varepsilon_{t-j} \\ &= \sum_{j=-\infty}^{+\infty} \psi_j \varepsilon_{t-j} \text{ avec } \psi_j = \begin{cases} \theta_j & \text{si } j = 0; \dots; q \\ 0 & \text{sinon} \end{cases}\end{aligned}$$

$$\Leftrightarrow \sum_{j=-\infty}^{+\infty} \psi_j = \sum_{j=0}^q \theta_j < \infty$$

où ε_t est un processus stationnaire par définition, et la stationnarité du processus X_t résulte de la proposition (II.2), donc le processus moyenne mobile est permanent (stationnaire).

3/ Le processus autorégressif moyenne mobile ARMA(p,q) :

Le processus auto-régressif moyen mobile peut être présenté sous forme :

$$\Phi(B)\tilde{X}_t = \Theta(B)\varepsilon_t$$

où d'après la proposition (II.2) et le théorème (II.1), le processus X_t est stationnaire si et seulement si le polynôme Φ ne possède pas de zéros à l'intérieur du disque d'unité.

II.1.2- Inversibilité :

Théorème II.2 :

Le processus ARMA(p,q) pour lequel Φ et Θ n'ont pas de zéros communs est inversible si et seulement si le polynôme Θ ne possède pas de zéros à l'intérieur du disque d'unité.

Preuve : la démonstration est similaire à celle du Théorème précédent

1/Le processus auto-régressif AR (p) :

Le processus auto-régressif peut être représenté sous forme :

$$\Phi(B)\tilde{X}_t = \varepsilon_t \Leftrightarrow \Phi(B)\tilde{X}_t = 1.\varepsilon_t$$

puisque le AR(p) est un ARMA(p,q) avec $\Theta(B) = 1$ et d'après le théorème (II.2) ce processus est toujours inversible ($\Theta(z) = 1$ et $\Phi(z) \neq 0; \forall z$)

2/Processus moyenne mobiles MA(q) :

Le processus moyenne mobile MA (q) peut être représenté sous forme :

$$\begin{aligned}\tilde{X}_t &= \Theta(B)\varepsilon_t \\ \Leftrightarrow \Phi(B)\tilde{X}_t &= \Theta(B)\varepsilon_t ; \text{avec } \Phi(B) = 1\end{aligned}$$

Alors, MA(q) est un ARMA(0,q)

Donc d'après le théorème (II.2) ce processus est inversible si le polynôme $\Theta(B)$ ne possède pas de zéros à l'intérieur du disque d'unité ($\Theta(B) \neq 0, \forall B$)

3/Processus moyenne mobile intégré IMA(0,q,d) :

Supposons que X_t est un processus stochastique non stationnaire mais, alors que ses différences d'ordre d ($Y_t = (1-B)^d X_t$) le sont.

et si le nouveau processus Y_t (défini précédemment) est représenté par un modèle moyen mobile $MA(q)$ alors le processus initial X_t se trouve représenté par un modèle moyen mobile intégré $IMA(0,d,q)$, et :

$$\left. \begin{array}{l} Y_t = (1-B)^d X_t \\ Y_t = \Theta(B)\varepsilon_t \end{array} \right\} \Rightarrow (1-B)^d X_t = \Theta(B)\varepsilon_t$$

d'où la forme générale du processus moyen mobile intégré :

$$(1-B)^d X_t = \Theta(B)\varepsilon_t \quad \text{où } \Theta(B) \text{ est un polynôme d'ordre } q.$$

Tout comme le modèle moyenne mobile $MA(q)$ pour que le processus moyenne mobile intégré $IMA(0,d,q)$ soit inversible il faut et il suffit que le polynôme $\Theta(B)$ n'y pas de racines à l'intérieur du disque d'unité ce qui assure l'existence de $\Theta^{-1}(B)$, et l'on aura :

$$\varepsilon_t = \Theta^{-1}(B)(1-B)^d X_t$$

Proposition II.3 :

Un processus linéaire X_t est stationnaire lorsque la série $\sum_{j=0}^{+\infty} |\psi_j|$ converge, il est inversible lorsque la série $\sum_{j=0}^{+\infty} |\pi_j|$ converge.

II.2- Processus ARIMA avec bruit additif

II.2.1- Somme de deux processus moyenne mobile

Proposition II.4 :

La somme de deux processus moyenne mobile est un processus moyenne mobile dont l'ordre est le plus grand de ceux des deux processus.

Preuve :

Soit X_t, Y_t deux processus moyenne mobile .

X_t, Y_t étant deux processus moyenne mobile on a, $X_t = \Theta_1(B)\alpha_t$ et $Y_t = \Theta_2(B)\beta_t$

en supposons que α_t et β_t sont deux bruits blancs mutuellement indépendants.

en posant $W_t = X_t + Y_t$.

on écrit $W_t = \Theta_1(B)\alpha_t + \Theta_2(B)\beta_t$

où $\Theta_1(B) = \sum_{j=0}^n \theta_{1,j} B^j$ et $\Theta_2(B) = \sum_{j=0}^m \theta_{2,j} B^j$

or, l'indépendance des deux bruits blancs α_t et β_t implique l'indépendance des deux processus X_t et Y_t il en résulte d'une part : $\gamma_w(h) = \gamma_x(h) + \gamma_y(h)$

et d'autre part :

$$X_t = \sum_{j=0}^m \theta_{1,j} \alpha_{t-j} \Rightarrow \gamma_x(h) = \begin{cases} \sigma_\alpha^2 \sum_{j=0}^m \theta_{1,j} \theta_{1,j+h} & \text{si } h \leq m \\ 0 & \text{sinon} \end{cases} \dots\dots\dots(1)$$

$$Y_t = \sum_{j=0}^n \theta_{2,j} \alpha_{t-j} \Rightarrow \gamma_Y(h) = \begin{cases} \sigma_\beta^2 \sum_{j=0}^n \theta_{2,j} \theta_{2,j+h} & \text{si } h \leq n \\ 0 & \text{sinon} \end{cases} \dots\dots\dots(2)$$

$$(1) \text{ et } (2) \Rightarrow \gamma_W(h) = \begin{cases} \left[\sigma_\alpha^2 \sum_{j=0}^n \theta_{1,j} \theta_{1,j+h} \right] 1_{[0,n]}(h) + \left[\sigma_\beta^2 \sum_{j=0}^m \theta_{1,j} \theta_{1,j+h} \right] 1_{[0,m]} & \text{si } h \leq \sup(n,m) \\ 0 & \text{sinon} \end{cases}$$

Une des principales propriétés du processus moyenne mobile est que sa fonction d'auto-covariance $\gamma(h)$ est nulle pour tout h supérieur à son ordre et l'on conclut que W_t est un MA[$\sup(n,m)$].

II.2.2- L'effet du bruit additif sur le processus ARIMA

Supposons qu'on ne peut pas observer le processus Y_t qui est un ARIMA(p,d,q) mais seulement $X_t = Y_t + \alpha_t$ où α_t est un bruit extérieur corrélé; et qu'on doit définir la nature du processus observé X_t .

Du fait que Y_t est un processus ARIMA(p,d,q) il possède la représentation

$$\Phi(B) \nabla^d Y_t = \Theta(B) \varepsilon_t$$

Et comme il a été mentionné

$$X_t = Y_t + \alpha_t$$

$$\Rightarrow \Phi(B) \nabla^d X_t = \Phi(B) \nabla^d Y_t + \Phi(B) \nabla^d \alpha_t$$

$$\Rightarrow \Phi(B) \nabla^d X_t = \Theta(B) \varepsilon_t + \Phi(B) \nabla^d \alpha_t$$

Dans ce cas on distingue deux cas possible :

1/ α_t est un bruit blanc additif

Par définition on a successivement :

$$\begin{aligned}
 X_t &= Y_t + \alpha_t \\
 \Rightarrow \nabla^d X_t &= \nabla^d Y_t + \nabla^d \alpha_t \\
 \Rightarrow \Phi(B) \nabla^d X_t &= \Phi(B) \nabla^d Y_t + \Phi(B) \nabla^d \alpha_t \\
 \Rightarrow \Phi(B) \nabla^d X_t &= \Theta(B) \varepsilon_t + \Theta'(B) \alpha_t \\
 \Rightarrow \Phi(B) \nabla^d X_t &= \Theta_1(B) \mu_t \text{ (D'après la proposition (II.4))}
 \end{aligned}$$

où $\Theta_1(B)$ est un polynôme d'ordre $Q = \sup(p, q)$ et μ_t un bruit blanc.

(Par conséquent X_t est un processus ARIMA(p, d, Q))

2/ α_t est un bruit corrélé ARMA(p, q)

Si α_t est un ARMA (p, q) alors il existe deux polynômes $\Phi_1(B)$ et $\Theta_1(B)$ et un bruit blanc β_t tel que :

$$\Phi_1(B) \alpha_t = \Theta_1(B) \beta_t, \text{ on a } X_t = Y_t + \alpha_t \text{ avec } \Phi(B) Y_t = \Theta(B) \varepsilon_t$$

et par suite :

$$\begin{aligned}
 X_t &= Y_t + \alpha_t \\
 \Rightarrow \nabla^d X_t &= \nabla^d Y_t + \nabla^d \alpha_t \\
 \Rightarrow \Phi(B) \nabla^d X_t &= \Phi(B) \nabla^d Y_t + \Phi(B) \nabla^d \alpha_t \\
 \Rightarrow \Phi(B) \nabla^d X_t &= \Theta(B) \varepsilon_t + \Phi(B) \nabla^d \alpha_t
 \end{aligned}$$

$$\Rightarrow \Phi_1(B)\Phi(B)\nabla^d X_t = \Phi_1(B)\Theta(B)\varepsilon_t + \Phi(B)\nabla^d \Phi_1(B)\alpha_t$$

$$\Rightarrow \Phi_1(B)\Phi(B)\nabla^d X_t = \Phi_1(B)\Theta(B)\varepsilon_t + \Phi(B)\nabla^d \Theta_1(B)\beta_t$$

$$\Rightarrow \Phi_2(B)X_t = \Theta_1(B)\mu_t$$

donc le processus X_t est un ARIMA, et pour le définir il suffit de déterminer σ_μ^2 et les coefficients des deux polynômes $\Phi_2(B)$ et $\Theta_1(B)$.

Chapitre III: Estimation des paramètres des modèles
ARMA et ARIMA

III.1- Estimation des coefficients de l'ARMA

Après avoir choisi le modèle ARIMA comme modèle de représentation pour notre processus stochastique on estime ses paramètres. Si on suppose que p, d, q sont connues il ne nous reste que l'estimation de $\beta = (\Phi, \Theta, \sigma^2)$.

Pour cela il existe plusieurs méthodes, dans notre travail on utilise deux, la première est celle du maximum de vraisemblance et le deuxième est celle des moindres carrées ([13],[15], [18],[21],[28],[41], [44]).

III.1.1- Estimation par la méthode du maximum de vraisemblance (MMV)

a\ La fonction de vraisemblance

Soit X_t un processus stochastique ARMA.

Considérons :

$X = (X_1, X_2, \dots, X_n)$ un vecteur de n variables aléatoire du processus X_t qui a comme densité conjointe la fonction $f(X/\beta)$ dépendant de $\beta = (\Phi, \Theta, \sigma_\varepsilon^2)$ avec $(\Phi, \Theta) = [(\phi_1, \dots, \phi_p); (\theta_1, \dots, \theta_q)]$ un vecteur de $p+q$ paramètres .

$x = (x_1, x_2, \dots, x_n)$ une observation de ce vecteur.

A partir de l'observation x on estime la valeur $\hat{\beta}$ de β qui rend le modèle compatible aux observations, autrement dit on cherche la valeur de β qui rend $f(\beta/X)$ compatible ; au maximum ; pour l'ensemble d'observations.

La fonction de densité conjointe du vecteur X vue comme fonction de β s'appelle fonction de vraisemblance.

Les fonctions de vraisemblance se diffèrent selon le tableau de données.

Pour un tableau de données où les variables sont indépendantes de loi marginales $f(X_t, \beta)$ la densité jointe du vecteur est le produit des densités marginales :

$$f(X; \beta) = \prod_{t=1}^N f(X_t, \beta) \dots \dots \dots$$

et comme la fonction de vraisemblance est la densité conjointe du vecteur X vue comme fonction de β on a :

$$L\left(\frac{\beta}{X}\right) = L\left(\frac{\beta}{X_1, \dots, X_N}\right) = \prod_{t=1}^N f(X_t, \beta) = g(B) \dots \dots \dots (3.1)$$

$g(B)$ s'appelle la fonction de vraisemblance exacte.

Pour un tableau de données où les variables sont corrélées la fonction de vraisemblance précédente n'est plus valide (à cause de la dépendance des variables), en échange pour calculer la fonction de vraisemblance on écrit la densité jointe sous forme de produit de densités conditionnelle et de fonction de valeurs initiales comme suit :

$$f(X, \beta) = \left(\prod_{t=p+1}^N f\left(\frac{X_t}{X_{t-1}, \dots, X_1}\right) \right) \cdot f(X_p, \dots, X_1) = L(B) \dots \dots \dots (3.2)$$

où $\left(\prod_{t=p+1}^N f\left(\frac{X_t}{X_{t-1}, \dots, X_1}\right) \right)$ s'appelle fonction de vraisemblance conditionnelle.

$L(\beta)$ est alors fonction de vraisemblance exacte.

b) La fonction de vraisemblance pour un ARIMA :

Supposons qu'on a $N = n + d$ observations du processus X_t représenté par un modèle ARIMA(p, d, q)

De ces N observations on peut extraire n observations de $W_t = \nabla^d X_t$ où l'estimation des paramètres $\{\beta = (\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q, \sigma_\varepsilon^2)\}$ sera équivalente à l'estimation des paramètres de W_t .

b.1\ La fonction de vraisemblance inconditionnelle :

Soit W_t le processus présenté par le modèle ARMA (p, q) , de moyenne $\mu=0$, qui s'écrit sous forme :

$$W_t = \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

$$\Rightarrow \varepsilon_t = W_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Et en supposant que les variables du vecteur $\varepsilon = (\varepsilon_i)_{i=1, \dots, n}$ i.i.d selon une loi normale centré et d'écart type σ_ε , la densité conjointe de ce vecteur sera la densité d'un vecteur gaussien définie par :

$$f_{(\varepsilon_1, \dots, \varepsilon_n)} = \frac{1}{(2\pi\sigma_\varepsilon^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma_\varepsilon^2} \sum_{i=1}^n \varepsilon_i^2}$$

Posons en outre,

$$W = (W_1, \dots, W_n), \varepsilon = (\varepsilon_1, \dots, \varepsilon_n), e_* = (W_{1-p}, \dots, W_0, \varepsilon_{1-q}, \dots, \varepsilon_0) = (W_*, \varepsilon_{1-q}, \dots, \varepsilon_0)$$

$$L_\phi = \begin{bmatrix} 1 & -\phi_1 & -\phi_2 & \dots & \dots & -\phi_p & 0 & 0 \\ 0 & 1 & -\phi_2 & \dots & \dots & -\phi_p & 0 & 0 \\ 0 & 0 & 1 & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & 1 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 & 1 & -\phi_p & \dots & \dots \\ \dots & \dots & \dots & \dots & 0 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, L_\theta = \begin{bmatrix} 1 & -\theta_1 & -\theta_2 & \dots & \dots & -\theta_q & 0 & 0 \\ 0 & 1 & -\theta_2 & \dots & \dots & -\theta_q & 0 & 0 \\ 0 & 0 & 1 & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & 1 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 & 1 & -\theta_q & \dots & \dots \\ \dots & \dots & \dots & \dots & 0 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$F = \begin{bmatrix} A_p & B_q \\ 0 & 0 \end{bmatrix}, A_p = \begin{bmatrix} \phi_p & \phi_{p-1} & \dots & \dots & \dots & \phi_1 \\ 0 & \phi_p & \phi_{p-1} & \dots & \dots & \phi_2 \\ \dots & 0 & \dots & \dots & \dots & \dots \\ \dots & 0 & 0 & \dots & \phi_{p-1} & \dots \\ \dots & \dots & \dots & 0 & \phi_p & \phi_{p-1} \\ 0 & 0 & 0 & 0 & 0 & \phi_p \end{bmatrix} \text{ et } B_q = \begin{bmatrix} \theta_q & \theta_{q-1} & \dots & \dots & \dots & \theta_1 \\ 0 & \theta_q & \dots & \dots & \dots & \theta_2 \\ \dots & 0 & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 & \theta_q & \theta_{q-1} \\ 0 & 0 & 0 & 0 & 0 & \theta_q \end{bmatrix}$$

on voit clairement que la forme d'équation différentielle du processus ARMA (relation (1.5)) est équivalente au système :

$$L_\Phi W = L_\Theta \varepsilon + F e_* \dots \dots \dots (3.3)$$

$$\Rightarrow \varepsilon = L_\Theta^{-1} (L_\Phi W - F e_*) \dots \dots \dots (3.4)$$

ε et e_* étant deux vecteurs indépendants de lois gaussiennes, on supposant σ_ε^2 connue

on a :

$$f\left(\varepsilon, e_* / \sigma_\varepsilon^2\right) = f\left(\varepsilon / \sigma_\varepsilon^2\right) f\left(e_* / \sigma_\varepsilon^2\right)$$

Si Ω la fonction d'auto covariance du vecteur e_* , alors :

$$\begin{aligned} f\left(\varepsilon, e_* / \sigma_\varepsilon^2\right) &= \frac{1}{(2\pi\sigma_\varepsilon^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma_\varepsilon^2}(\varepsilon'\varepsilon)} \frac{1}{(2\pi\sigma_\varepsilon^2)^{\frac{p+q}{2}} |\Omega|^{\frac{1}{2}}} e^{-\frac{1}{2\sigma_\varepsilon^2}(e_*'\Omega^{-1}e_*)} \\ &= \frac{1}{(2\pi\sigma_\varepsilon^2)^{\frac{n+p+q}{2}} |\Omega|^{\frac{1}{2}}} e^{-\frac{1}{2\sigma_\varepsilon^2}(\varepsilon'\varepsilon + e_*'\Omega^{-1}e_*)} \end{aligned}$$

considérons la transformation $J : J(\varepsilon, e_*) = (W, e_*); \varepsilon = L_\Theta^{-1}(L_\Phi W - F e_*)$

le jacobien de J étant unitaire alors,

$$f\left(W, e_* / \Phi, \Theta, \sigma^2\right) = \frac{1}{(2\pi\sigma^2)^{\frac{n+p+q}{2}} |\Omega|^{\frac{1}{2}}} e^{-\frac{1}{2\sigma^2}S(\Phi, \Theta, e_*)}$$

où $S(\Phi, \Theta, e_*) = (L_\Phi X - F e_*)' L_\Theta^{-1} L_\Theta (L_\Phi X - F e_*) + e_*' \Omega^{-1} e_*$

$$= S(\Phi, \Theta) + (e_* - \hat{e}_*)' D (e_* - \hat{e}_*)$$

Avec : $S(\Phi, \Theta) = S(\Phi, \Theta, \hat{e}_*) = \hat{e}_*' \hat{e}_* + \hat{e}_*' \Omega^{-1} \hat{e}_*$

$\hat{e}_* = E(e_* / X, \Phi, \Theta) = D^{-1} F' L_\Theta^{-1} L_\Theta^{-1} L_\Phi X; D = \Omega^{-1} + F' L_\Theta^{-1} L_\Theta^{-1} F, \Omega \sigma_\varepsilon^2 = cov(e_* e_*')$

d'où,

$$\begin{aligned}
 f\left(\frac{W, e_*}{\Phi, \Theta, \sigma^2}\right) &= \frac{1}{(2\pi\sigma^2)^{\frac{n+p+q}{2}} |\Omega|^{\frac{1}{2}}} e^{\frac{-1}{2\sigma^2}[S(\Phi, \Theta) + (e_* - \hat{e}_*)' \Omega^{-1} (e_* - \hat{e}_*)]} \\
 &= \frac{1}{(2\pi\sigma_\varepsilon^2)^{\frac{n}{2}} |\Omega|^{\frac{1}{2}} |D|^{\frac{1}{2}}} e^{-\frac{1}{2\sigma_\varepsilon^2} S(\Phi, \Theta)} \cdot \frac{|D|^{\frac{1}{2}}}{(2\pi\sigma_\varepsilon^2)^{\frac{p+q}{2}}} e^{-\frac{1}{2\sigma_\varepsilon^2} (e_* - \hat{e}_*)' \Omega^{-1} (e_* - \hat{e}_*)} \\
 &= f\left(\frac{W}{\Phi, \Theta, \sigma_\varepsilon^2}\right) \cdot f\left(\frac{e_*}{\Phi, \Theta, \sigma_\varepsilon^2}\right) \\
 \text{ou encore: } L(\beta) &= f\left(\frac{W}{\Phi, \Theta, \sigma_\varepsilon^2}\right) = \frac{1}{(2\pi\sigma_\varepsilon^2)^{\frac{n}{2}} |\Omega|^{\frac{1}{2}} |D|^{\frac{1}{2}}} e^{-\frac{1}{2\sigma_\varepsilon^2} S(\Phi, \Theta)} \dots\dots\dots (3.5)
 \end{aligned}$$

de ce qui précède la fonction de vraisemblance exacte est de la forme:

$$L(\beta) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}} |\Omega|^{\frac{1}{2}} |D|^{\frac{1}{2}}} e^{-\frac{1}{2\sigma_\varepsilon^2} S(\Phi, \Theta)}$$

et le logarithme de la fonction de vraisemblance exacte (inconditionnelle) pour le processus ARMA aura pour expression :

$$\Rightarrow \ln[L(\beta)] = f(\Phi, \Theta) - n \ln(\sigma) - \frac{S(\Phi, \Theta)}{2\sigma_\varepsilon^2} \dots\dots\dots (3.6)$$

Par la méthode précédente on a calculé la fonction de vraisemblance exacte utilisant un changement de variable en se basant sur l'hypothèse d'indépendance des bruits, mais on peut la calculer utilisant une autre méthode et d'autres hypothèses comme suit :

dans un processus stochastique les variables aléatoires ne sont pas indépendantes, donc pour un n -uple $W = (W_1, \dots, W_n)$ on peut définir la fonction de vraisemblance exacte par :

$$f(W/\beta) = \left(\prod_{t=1}^n f\left(\frac{W_t}{W_{t-1}, \dots, W_1, \beta}\right) \right)$$

posons :

$$\begin{aligned} \hat{W}_t &= E\left(\frac{W_t}{W_{t-1}, \dots, W_1, \beta}\right) \\ \sigma_\varepsilon^2 v_t &= E(W_t - \hat{W}_t)^2 \\ \det(M_n^{(p,q)}) &= v_1 \dots v_n \end{aligned}$$

$M_n^{(p,q)}$ étant la matrice d'auto-covariance du n -uplet W .

Si on suppose que le processus W_t est de loi normale, on aura :

$$f\left(\frac{W_t}{W_{t-1}, \dots, W_1, \beta}\right) = \frac{1}{[2\pi\sigma_\varepsilon^2 v_t]^{\frac{1}{2}}} e^{-\frac{1}{2\sigma_\varepsilon^2} \frac{(W_t - \hat{W}_t)^2}{v_t}}$$

et par suite :

$$f(W/\beta) = \prod_{t=1}^n \frac{1}{[2\pi\sigma_\varepsilon^2 v_t]^{\frac{1}{2}}} e^{-\frac{1}{2\sigma_\varepsilon^2} \sum_{t=1}^n \frac{(W_t - \hat{W}_t)^2}{v_t}}$$

de cette dernière relation on conclut que :

$$L(\Phi, \Theta, \sigma_\varepsilon^2) = \prod_{t=1}^n \frac{1}{[2\pi\sigma_\varepsilon^2 v_t]^{\frac{1}{2}}} e^{-\frac{1}{2\sigma_\varepsilon^2} \sum_{t=1}^n \frac{(W_t - \hat{W}_t)^2}{v_t}}$$

$$\begin{aligned} \Rightarrow l(\beta) = \ln[L(\Phi, \Theta, \sigma_\varepsilon^2)] &= \frac{n}{2} \ln(2\pi) - \frac{n}{2} \sum_{t=1}^n \sigma_\varepsilon^2 - \frac{1}{n} \sum_{t=1}^n \ln(v_t) - \frac{1}{2\sigma_\varepsilon^2} S^2(\Phi, \Theta) \\ &= \frac{n}{2} \ln(2\pi) - \frac{n}{2} \sum_{t=1}^n \sigma_\varepsilon^2 - \frac{1}{n} \ln[\det(M_n^{(p,q)})] - \frac{1}{2\sigma_\varepsilon^2} S^2(\Phi, \Theta) \dots (3.7) \end{aligned}$$

avec : $S^2(\Phi, \Theta) = \sum_{t=1}^n \frac{(W_t - \hat{W}_t)^2}{v_t}$

b.2\ La fonction de vraisemblance conditionnelle :

Prenons à nouveau le processus W_t défini précédemment par :

$$W_t = \varepsilon_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

$$\Leftrightarrow \varepsilon_t = W_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

On ne peut pas calculer les $(\varepsilon_t)_{t=1, \dots, n}$ directement à partir de l'équation différentielle car cela nécessite des valeurs initiales. Pour cette raison supposons qu'on a p valeurs initiales du processus $W_* = (W_0, W_{-1}, \dots, W_{1-p})$ et q valeurs du bruit blanc $\varepsilon_* = (\varepsilon_0, \dots, \varepsilon_{1-q})$ dans ce cas les $(\varepsilon_t)_{t=1, \dots, n}$ peuvent être calculer directement de la relation ci-dessus conditionnellement aux valeurs initiales.

Ainsi, pour n'importe quel choix des paramètres et des valeurs initiales, on peut calculer successivement un ensemble de valeurs des $\varepsilon_t \left(\frac{(\Phi, \Theta)}{W, W_*, \varepsilon_*} \right)_{t=0, 1, \dots, n}$

et par conséquent, la loi conjointe des $(\varepsilon_i)_{i=1, \dots, n}$ qui sont des variables identiquement équidistribués de loi gaussienne centré et d'écart type σ_ε est :

$$f_{(\varepsilon_1, \dots, \varepsilon_n)} = \frac{1}{(2\pi\sigma_\varepsilon^2)^n} e^{-\frac{1}{2\sigma_\varepsilon^2} \sum_{t=1}^n \varepsilon_t^2}$$

Or, pour une série chronologique donnée de W_t le logarithme de la fonction de vraisemblance conditionnée par le choix de W_* , s'écrit comme suit :

$$\begin{aligned} \ln [L(\Phi, \Theta, \sigma_\varepsilon^2)] &= -\frac{n}{2} \ln(2\pi) - n \ln(\sigma_\varepsilon) - \frac{1}{2\sigma_\varepsilon^2} \sum_{t=1}^n \varepsilon_t^2 \left(\frac{\Phi, \Theta}{W_*, \varepsilon_*, W} \right) \\ &= -\frac{n}{2} \ln(2\pi) - n \ln(\sigma_\varepsilon) - \frac{1}{2\sigma_\varepsilon^2} S^2(\Phi, \Theta) \dots \dots \dots (3.8) \end{aligned}$$

c\ L'estimateur du Maximum de vraisemblance

L'estimateur du maximum de vraisemblance est la valeur de $\beta = (\Phi, \Theta, \sigma_\varepsilon^2)$ qui maximise la fonction de vraisemblance.

Or, la fonction de vraisemblance en tant que fonction de Φ, Θ n'est pas une forme quadratique ; c'est d'ailleurs; pour cette raison qu'on utilise des méthodes itératives ([13],[48]) afin de la maximiser par rapport à Φ, Θ . Autrement dit la détermination de $\hat{\Phi}, \hat{\Theta}$.

Une fois $\hat{\Phi}, \hat{\Theta}$ trouvés (ces derniers ne dépendent pas de σ_ε^2), la valeur de $\hat{\sigma}_\varepsilon^2$ peut être trouvée en maximisant $l_* = \ln(L)$ par rapport à σ_ε^2 .

$$\begin{aligned} \ln[L(\Phi, \Theta, \sigma_\varepsilon^2)] &= -\frac{n}{2} \ln(2\pi) - n \ln(\sigma_\varepsilon) - \frac{1}{2\sigma_\varepsilon^2} S^2(\Phi, \Theta) \\ \Rightarrow \frac{\partial l_*}{\partial \sigma_\varepsilon^2} &= -\frac{n}{2\sigma_\varepsilon^2} + \frac{S^2(\hat{\Phi}, \hat{\Theta})}{2\sigma_\varepsilon^4} \\ \Rightarrow \hat{\sigma}_\varepsilon^2 &= \frac{S^2(\hat{\Phi}, \hat{\Theta})}{n} \dots\dots\dots(3.9) \end{aligned}$$

III.1.2- Estimation par la méthode des moindres carrés (MMC)

a\ La somme de carrés des erreurs :

Supposons que notre modèle s'écrit sous forme :

$$W_i = \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \varepsilon_i \dots\dots\dots(3.10)$$

où les $x_{ij}, j = 1, \dots, n$ sont des observations obtenues lors d'une expérience, les $\beta_j, j = 1, \dots, n$ sont des paramètres inconnus qu'on compte estimer à partir de tableau de données, et les $\varepsilon_i, i = 1, \dots, n$ sont des erreurs identiquement équi-distribuées et non corrélées de moyenne nulle et de variance σ_ε^2 .

la relation (3.10) s'écrit sous forme matricielle comme suit :

$$\begin{bmatrix} W_1 \\ W_2 \\ \cdot \\ \cdot \\ W_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{1,2} & \cdot & \cdot & x_{1,k} \\ x_{21} & x_{2,2} & \cdot & \cdot & x_{2,k} \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ x_{n1} & x_{n2} & & & x_{n,k} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_n \end{bmatrix} \Leftrightarrow W = X\beta + \varepsilon \dots \dots (3.11)$$

où la matrice X est de rang k .

La somme de carré des erreurs est :

$$S^2(\beta) = \varepsilon' \varepsilon = (W - X\beta)'(W - X\beta)$$

Dans le but de calculer le minimum de S^2 on décompose $(W - X\beta)$ comme suit :

$$(W - X\beta) = (W - X\hat{\beta}) - X(\beta - \hat{\beta})$$

et la valeur de $\hat{\beta}$ qui minimise S^2 est $\hat{\beta} = (X'X)^{-1}X'W \dots \dots \dots (3.12)$

ce qui permet d'écrire :

$$S^2(\beta) = S^2(\hat{\beta}) + (\beta - \hat{\beta})' X' X (\beta - \hat{\beta}) \dots \dots \dots (3.13)$$

b) La somme de carrés des erreurs pour un ARMA :

Comme pour le calcul de la fonction de vraisemblance supposons qu'on a $N = n + d$ observations du processus X_t et que ce dernier est généré par un modèle ARIMA(p, d, q)

$$\Phi(B)\nabla^d X_t = \Theta(B)\varepsilon_t$$

de ces N observations on peut extraire n observations de W_t ($W_t = \nabla^d X_t$) et l'estimation des paramètres de X_t [$\beta = (\Phi, \Theta, \mu)$] sera équivalente à l'estimation des paramètres de W_t .

b.1\ La somme inconditionnelle de carrés des erreurs :

Tant que le processus W_t admet une présentation par le modèle ARMA on a :

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad \text{pour } t=1, \dots, n$$

cette dernière relation peut s'écrire sous forme matricielle comme suit :

$$L_\phi W = L_\theta \varepsilon + F e_* \dots \dots \dots (3.14)$$

où $W = (W_1, \dots, W_n)$, $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$, $e_* = (W_{1-p}, \dots, W_0, \varepsilon_{1-q}, \dots, \varepsilon_0)$

$$L_\phi = \begin{bmatrix} 1 & -\phi_1 & -\phi_2 & \dots & \dots & -\phi_p & 0 & 0 \\ 0 & 1 & -\phi_2 & \dots & \dots & -\phi_p & 0 & 0 \\ 0 & 0 & 1 & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & 1 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 & 1 & -\phi_p & \dots & \dots \\ \dots & \dots & \dots & \dots & 0 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad L_\theta = \begin{bmatrix} 1 & -\theta_1 & -\theta_2 & \dots & \dots & -\theta_q & 0 & 0 \\ 0 & 1 & -\theta_2 & \dots & \dots & -\theta_q & 0 & 0 \\ 0 & 0 & 1 & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & 1 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 & 1 & -\theta_q & \dots & \dots \\ \dots & \dots & \dots & \dots & 0 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$F = \begin{bmatrix} A_p & B_q \\ 0 & 0 \end{bmatrix}, \quad A_p = \begin{bmatrix} \phi_p & \phi_{p-1} & \dots & \dots & \dots & \phi_1 \\ 0 & \phi_p & \phi_{p-1} & \dots & \dots & \phi_2 \\ \dots & 0 & \dots & \dots & \dots & \dots \\ \dots & 0 & 0 & \dots & \phi_{p-1} & \dots \\ \dots & \dots & \dots & 0 & \phi_p & \phi_{p-1} \\ 0 & 0 & 0 & 0 & 0 & \phi_p \end{bmatrix} \quad \text{et } B_q = \begin{bmatrix} \theta_q & \theta_{q-1} & \dots & \dots & \dots & \theta_1 \\ 0 & \theta_q & \dots & \dots & \dots & \theta_2 \\ \dots & 0 & \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 & \theta_q & \theta_{q-1} \\ 0 & 0 & 0 & 0 & 0 & \theta_q \end{bmatrix}$$

on remarque dans la forme matricielle que W dépend de e_* qui est un vecteur des valeurs initial.

A partir de cette forme matricielle on trouve :

$$\begin{aligned} \varepsilon &= L_\theta^{-1}(L_\phi W - F e_*) \Rightarrow \varepsilon = L_\theta^{-1} L_\phi W - L_\theta^{-1} F e_* \\ &\Rightarrow \varepsilon = L_\theta^{-1} L_\phi W - L_\theta^{-1} F \hat{e}_* - L_\theta^{-1} F e_* + L_\theta^{-1} F \hat{e}_* \\ &\Rightarrow \varepsilon = L_\theta^{-1}(L_\phi W - F \hat{e}_*) - L_\theta^{-1} F(e_* - \hat{e}_*) \end{aligned}$$

utilisant la théorie des moindres carrés on montre que :

$$S^2(\Phi, \Theta, e_*) = \mathcal{E}' \mathcal{E} = [L_{\Theta}^{-1}(L_{\Phi} W - F \hat{e}_*) - L_{\Theta}^{-1} F (e_* - \hat{e}_*)] [L_{\Theta}^{-1}(L_{\Phi} W - F \hat{e}_*) - L_{\Theta}^{-1} F (e_* - \hat{e}_*)]' \\ = \hat{\mathcal{E}}' \hat{\mathcal{E}} + \hat{e}_* \Omega^{-1} \hat{e}_* + (e_* - \hat{e}_*)' D (e_* - \hat{e}_*)$$

avec $D = \Omega^{-1} + F' L_{\Theta}^{-1} L_{\Theta}^{-1} F$ et $\hat{e}_* = E\left(\frac{e_*}{W}, \Phi, \Theta\right) = D^{-1} F' L_{\Theta}^{-1} L_{\Theta}^{-1} L_{\Phi} W$

dans le cas où on ne possède pas de valeurs initiales e_* on les estime par \hat{e}_* définie ci-dessus ce qui donne la fonction des moindres carrés inconditionnelle définie par :

$$S^2(\Phi, \Theta) = S^2(\Phi, \Theta, \hat{e}_*) = \hat{\mathcal{E}}' \hat{\mathcal{E}} + \hat{e}_* \Omega^{-1} \hat{e}_* \dots \dots \dots (3.16)$$

b.2\ La somme conditionnelle de carrés des erreurs :

Si on a un vecteur de valeurs initiales on calcule les \mathcal{E}_t de façon récursive à partir de la relation :

$$\mathcal{E}_t = W_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p} - \theta_1 \mathcal{E}_{t-1} - \theta_2 \mathcal{E}_{t-2} - \dots - \theta_q \mathcal{E}_{t-q}$$

et S^2 sera :

$$S^2\left(\frac{\Phi, \Theta}{e_*, W}\right) = \sum_{t=1}^n \mathcal{E}_t^2\left(\frac{\Phi, \Theta}{e_*, W}\right)$$

c\ L'estimateur des moindres carrés :

L'estimateur des moindres carrés est la valeur de $\beta = (\Phi, \Theta, \sigma_{\mathcal{E}}^2)$ qui minimise la fonction des moindres carrés S^2 . Pour les processus ARMA la fonction de somme des carrés des erreurs vue comme fonction de (Φ, Θ) n'est pas une forme quadratique, pour cette raison on utilise des méthodes itératives ([13]) afin de la minimiser.

Une fois $\hat{\Phi}, \hat{\Theta}$ trouvés (ces derniers ne dépendent pas de $\sigma_{\mathcal{E}}^2$), on calcule la valeur de

$\hat{\sigma}_{\mathcal{E}}^2$ par la relation :

$$\hat{\sigma}_{\mathcal{E}}^2 = \frac{S^2(\hat{\Phi}, \hat{\Theta})}{n - p - q}$$

III.2- Estimation des ordres p, q

Précédemment les paramètres p, q sont supposées connues mais réellement on ne connaît pas leur vraie valeurs, donc il est bien évident souhaitable de trouver des valeurs proches \hat{p}, \hat{q} afin de définir le modèle optimal pour la représentation de processus générés, pour cela plusieurs critères ont été suggérés parmi les critères proposés le plus performant ([35],[47]) est le critère d'information d'Akaike corrigé (AICC)

dans ce critère Hurvich et Tsai ont proposé une nouvelle statistique définie par :

$$AICC = -2\ln[L(\hat{\beta})] + \frac{2n(p+q+1)}{n-(p+q)-2} \dots\dots\dots (3.19)$$

et \hat{p}, \hat{q} sont les deux valeurs de p, q qui minimise la AICC.

Remarque : on ce qui concerne l'ordre d , on différencie le processus jusqu'à qu'on obtient un processus stationnaire est d sera le nombre de différenciation.

III.3- Comparaison des deux estimateurs (MMV) et (MMC)

Rappelons que l'estimateur du maximum de vraisemblance est la valeur $\hat{\beta}_{MV} = (\hat{\Phi}, \hat{\Theta}, \hat{\sigma}^2)$ de β qui maximise $L(B)$ et l'estimateur du moindres carrés est la valeur $\hat{\beta}_{MC} = (\hat{\Phi}, \hat{\Theta}, \hat{\sigma}^2)$ de β qui minimise $S^2(B)$.

Soit W_t un processus ARMA(p, q) causal et inversible.

du fait que le processus W_t est un ARMA sa densité est donnée par :

$$f_w(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \left| \frac{\Theta(e^{-i\lambda})}{\Phi(e^{-i\lambda})} \right|^2 = \frac{\sigma_\varepsilon^2}{2\pi} g(\lambda, \beta_0)$$

cela sous l'hypothèse que le polynôme Φ ne possède pas de zéros à l'intérieur du disque d'unité (la causalité).

Proposition III.1 :

Soit $C = \{\beta \in \mathfrak{R}^{p+q} / \Phi(z)\Theta(z) \neq 0 \text{ si } |z| \leq 1; \Phi \text{ et } \Theta \text{ n'ont pas de zéros communs}\}$

Si $\beta \in C$ alors :

$$\begin{aligned} 1 \setminus \ln[\det(M_n^{(p,q)})] &> 0 \\ 2 \setminus \ln[\det(M_n^{(p,q)})] &\xrightarrow{n \rightarrow +\infty} 0 \end{aligned}$$

Preuve :

Supposons que le processus W_t est un ARMA(p,q) de fonction de vraisemblance exacte

$$L(\beta) \text{ alors : } l(\beta) = \ln[L(\Phi, \Theta, \sigma^2)] = \frac{n}{2} \ln(2\pi) - \frac{n}{2} \sum_{t=1}^n \sigma_\varepsilon^2 - \frac{1}{n} \ln[\det(M_n^{(p,q)})] - \frac{1}{2\sigma_\varepsilon^2} \sum_{t=1}^n \frac{(W_t - \hat{W}_t)^2}{v_t}$$

$$\det(M_n^{(p,q)}) = v_1 \dots v_n \quad \text{et} \quad v_t = E(W_t - \hat{W}_t)^2 : \forall t = 1, \dots, n$$

On a la forme inversé de l'ARMA :

$$\varepsilon_t = W_t + \sum_{j=1}^{+\infty} \pi_j W_{t-j}$$

$$\Rightarrow \sigma_\varepsilon^2 = \text{var}(\varepsilon_{t+1}) = E\left(W_{t+1} + \sum_{j=1}^{+\infty} \pi_j W_{t+1-j}\right)^2$$

$$\leq \sigma_\varepsilon^2 v_t$$

$$\leq E\left(W_{t+1} + \sum_{j=1}^t \pi_j W_{t+1-j}\right)^2$$

$$\leq E\left(W_{t+1} - \sum_{j=t+1}^{+\infty} \pi_j W_{t+1-j}\right)^2$$

$$= 1 + E\left(\sum_{j=t+1}^{+\infty} \pi_j W_{t+1-j}\right)^2 \leq 1 + E\left(\sum_{j=t+1}^{+\infty} |\pi_j|^2\right) \gamma_w(0)$$

d'une part $v_t \geq 1$ et $v_1 \geq 1$ ce qui implique que :

$$\det(M_n^{(p,q)}) = v_1 \dots v_n > 1 \Leftrightarrow \ln(\det(M_n^{(p,q)})) > 0$$

d'autre part :

$$\left(\sum_{j=t+1}^{+\infty} |\pi_j|^2 \right) \gamma_w(0) \xrightarrow{t \rightarrow +\infty} 0 \text{ et } v_t \xrightarrow{t \rightarrow +\infty} 1$$

donc on conclut que :

$$0 \leq \frac{1}{n} \ln[\det(M_n^{(p,q)})] = \frac{1}{n} \sum_{t=1}^n \ln(v_t) \xrightarrow{n \rightarrow +\infty} 0 \Leftrightarrow \frac{1}{n} \ln[\det(M_n^{(p,q)})] > 0$$

De la relation (3.7) et la proposition (III.1) il résulte que l'estimateur du maximum de vraisemblance et l'estimateur des moindres carrés coïncides quand n tends vers l'infini autrement dit ils ont les mêmes propriétés asymptotique

III.4- Consistance forte des estimateurs des paramètres de l'ARMA

Soit $C = \{ \beta \in \mathfrak{R}^{p+q} / \Phi(z)\Theta(z) \neq 0 \text{ si } |z| \leq 1; \Phi \text{ et } \Theta \text{ n'ont pas de zéros communs} \}$

Proposition III.2 :

Pour un vecteur fixé dans C on a :

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{g(\lambda, \beta_0)}{g(\lambda, \beta)} d\lambda > 1 : \forall \beta \in \bar{C} \text{ tel que } \beta \neq \beta_0$$

Proposition III.3 :

Il existe un ensemble B dans le quel pour n'importe quelle suite convergente $(\beta_n)_{n \in \mathbb{N}}$ $(\beta_n \xrightarrow{ps} \beta \in C)$ on a les deux possibilités suivantes :

$$1 \backslash \text{si } \beta \in C \quad S_n^2(\beta_n) \xrightarrow{ps} \frac{\sigma_0^2}{2\pi} \int_{-\pi}^{\pi} \frac{g(\lambda, \beta_0)}{g(\lambda, \beta)} d\lambda$$

$$2 \backslash \text{si } \beta \in \partial C \quad \lim_{n \rightarrow +\infty} \inf S_n^2(\beta_n) \stackrel{ps}{\geq} \frac{\sigma_0^2}{2\pi} \int_{-\pi}^{\pi} \frac{g(\lambda, \beta_0)}{g(\lambda, \beta)} d\lambda$$

Théorème III.1 :

Soit $\hat{\beta}_n^{MV}$ l'estimateur du maximum de vraisemblance et $\hat{\beta}_n^{MC}$ l'estimateur des moindres carrés pour un processus ARMA(p,q) de vrai paramètre $\beta_0 \in C$ et de variance σ_0^2 donc :

$$1 \setminus \hat{\beta}_n^{MV} \xrightarrow{ps} \beta_0 \text{ et } S_n^2(\hat{\beta}_n^{MV}) \xrightarrow{ps} \sigma_0^2$$

$$2 \setminus \hat{\beta}_n^{MC} \xrightarrow{ps} \beta_0 \text{ et } S_n^2(\hat{\beta}_n^{MC}) \xrightarrow{ps} \sigma_0^2$$

Preuve :

1\ L'estimateur du maximum de vraisemblance

a\ $\hat{\beta}_n^{MV} \xrightarrow{ps} \beta_0$?

Supposons que $\hat{\beta}_n^{MV}$ ne converge pas vers β_0 , sous l'hypothèse de la compacité de \bar{C} il existe une sous-suite $\hat{\beta}_{n_k}^{MV} \rightarrow \beta \in \bar{C}, (\beta \neq \beta_0)$

D'après les propositions (III.1) et (III.3) on' a :

$$\liminf_{k \rightarrow +\infty} S_{n_k}^2(\hat{\beta}_{n_k}^{MC}) \stackrel{(3.b)}{\geq} \frac{\sigma_0^2}{2\pi} \int_{-\pi}^{\pi} \frac{g(\lambda, \beta_0)}{g(\lambda, \beta)} d\lambda \stackrel{(1)}{\geq} \sigma_0^2 \dots\dots (1)$$

D'après la proposition (III.3) et la définition de $\hat{\beta}_n^{MC}$

$$\limsup_{n \rightarrow +\infty} S_n^2(\hat{\beta}_n^{MC}) \stackrel{\text{par définition}}{\leq} \lim_{n \rightarrow +\infty} S_n^2(\beta_0) \leq \sigma_0^2 \dots\dots (2)$$

$$(1) \text{ et } (2) \Rightarrow \liminf_{k \rightarrow +\infty} S_{n_k}^2(\hat{\beta}_{n_k}^{MC}) \geq \limsup_{n \rightarrow +\infty} S_n^2(\hat{\beta}_n^{MC})$$

ce qui est impossible donc $\hat{\beta}_n^{MC} \xrightarrow{ps} \beta_0$.

$$b) S_n^2(\hat{\beta}_n^{MV}) \xrightarrow{ps} \sigma_0^2 ?$$

Selon la proposition (III.3) partie (1) Il existe un ensemble B sur dans le quel pour n'importe quelle suite $(\beta_n)_{n \in \mathbb{N}}, \beta_n \xrightarrow{ps} \beta \in C$ on a si $\beta \in C$ alors :

$$S_n^2(\beta_n) \xrightarrow{ps} \frac{\sigma_0^2}{2\pi} \int_{-\pi}^{\pi} \frac{g(\lambda, \beta_0)}{g(\lambda, \beta)} d\lambda$$

dans le cas présent la suite est $(\hat{\beta}_n^{MV})_{n \in \mathbb{N}}$ et elle converge vers β_0 donc

$$S_n^2(\beta_n) \xrightarrow{ps} \frac{\sigma_0^2}{2\pi} \int_{-\pi}^{\pi} \frac{g(\lambda, \beta_0)}{g(\lambda, \beta_0)} d\lambda = \sigma_0^2$$

2\ L'estimateur des moindres carrées

$$a) \hat{\beta}_n^{MC} \xrightarrow{ps} \beta_0 ?$$

Supposons $\hat{\beta}_n^{MC}$ que ne converge pas vers β_0 .

sous la condition de compacité il existe une sous-suite $\hat{\beta}_{n_k}^{MC} \rightarrow \beta \in \bar{C}, (\beta \neq \beta_0)$

$$\begin{aligned} \ln(\sigma_0^2) &\leq \liminf \ln S_n^2(\hat{\beta}_{MC}) \\ &\leq \liminf \ln l(\hat{\beta}_{MC}) \\ &\leq \limsup \ln l(\hat{\beta}) \\ &\leq \limsup \ln l(\beta_0) \\ &\leq \limsup_{n \rightarrow +\infty} \left[\frac{1}{n} \ln(\det M_n^{(p,q)}(\beta_0)) + \ln(S_n^2(\beta_0)) \right] \\ &= \ln(\sigma_0^2) \\ &= \ln[l(\beta_0)] \end{aligned}$$

ce qui résulte que $\hat{\beta}_{n_k}^{MC} \rightarrow \beta_0$

Ce dernier résultat contredit la supposition donc elle est fausse d'où $\hat{\beta}_n^{MC} \xrightarrow{ps} \beta_0$

$$b) S_n^2(\hat{\beta}_n^{MC}) \xrightarrow{ps} \sigma_0^2 \quad (\text{Même démonstration que celle de } \hat{\beta}_n^{MV}).$$

III.5- Comportement asymptotique des deux estimateurs (MMV) et (MMC)

Théorème III.2 :

Soit $\hat{\beta}_n^{MV}$ l'estimateur du maximum de vraisemblance et $\hat{\beta}_n^{MC}$ l'estimateur de moindres carrés pour un processus ARMA(p,q) de vrai paramètre $\beta_0 \in \mathbb{C}$ et de variance σ_0^2 , donc :

$$\hat{\beta}_n^{MV} \text{ est un } N\left(0, \frac{1}{nW(\beta_0)}\right)$$

$$\hat{\beta}_n^{MC} \text{ est un } N\left(0, \frac{1}{nW(\beta_0)}\right)$$

où :

$$W(\beta_0) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\frac{\partial [\ln g(\lambda(\beta_0))] }{\partial \beta} \right) \left(\frac{\partial \ln [g(\lambda, \beta_0)]}{\partial \beta} \right)' d\lambda$$

Preuve : (cf : [8],[42])

III.6- Test de compatibilité du modèle

Une fois les paramètres du modèle sont estimés on procède à un test de vérification pour le modèle estimé (reproduit le modèle qui a engendré les données). Pour cela les résidus obtenus à partir du modèle estimé sont utilisés pour vérifier s'ils se comportent comme des erreurs bruit blanc à l'aide d'un test "portemanteau" (test global qui permet de tester l'hypothèse d'indépendance des résidus). Le test le plus usuel est le test proposé par Box-Pierce :

Considérons la statistique Q définie par :

$$Q = T \sum_{h=1}^s \hat{\rho}^2(h) \dots \dots \dots (3.21)$$

$\hat{\rho}^2(h)$ étant la corrélation empirique entre les résidus, distants de h.

Sous l'hypothèse d'indépendance, cette statistique suit asymptotiquement une loi de χ^2 avec $s - p - q$ degrés de liberté. Si l'hypothèse n'est pas rejetée on dira que le modèle estimé est une bonne approximation du vrai modèle. Dans le cas contraire on rejettera le modèle estimé et on procédera à identifier un nouveau modèle.

Le problème avec la statistique de Box-Pierce est qu'elle est de mauvaises performances pour les petits échantillons.

En échange on utilise le test de Ljung-Box qui est performant même pour les échantillons de petites tailles.

la statistique de Ljung-Box est définie par :

$$Q = T(T + 2) \sum_{h=1}^s \frac{\hat{\rho}^2(h)}{T - h} \dots\dots\dots(3.22)$$

T : le nombre d'observations

s : nombre de résidus auxquels on test l'auto corrélations.

$\hat{\rho}^2(h)$: coefficients d'auto corrélations (pour le décalage h)

Q : test de portemanteau statistique.

Si la valeur de Q pour la série chronologique (l'échantillon) dépasse la valeur critique d'une distribution Chi-deux avec s degré de liberté alors au moins une valeur de $\hat{\rho}^2$ est différente de zéro au degré significative précisé, ce qui mène à rejeter l'hypothèse d'indépendance des résidus et par conséquent le rejet de modèle, sinon le modèle constitue une bonne représentation pour le processus stochastique.

Chapitre IV : Pr vision des mod les stochastiques

IV.1- Choix du prédicteur adéquat

On ce qui concerne la prévision on est censé de connaître une future valeur désirée la plus proche possible de la vraie valeur $X_{t+l}; l > 0$ au moment où on est à l'instant t .

on cherche donc à prévoir la valeur que va prendre le processus à l'instant $t+l$, dans cela on s'appuie sur une observation $x = (x_1, x_2, \dots, x_t)$ des variables X_1, X_2, \dots, X_t .

dans le cadre de la prévision de l'ARMA on choisit de prendre comme prédicteur une combinaison linéaire de ces $X_i; 1 \leq i \leq t$.

$$\text{Autrement dit : } \hat{X}_t(l) = a_1 X_1 + a_2 X_2 + \dots + a_t X_t \quad \dots \dots \dots (4.1)$$

ce qui est équivalent à une combinaison linéaire des $\varepsilon_{t-j}; j \geq 0$ du fait que pour toutes m , X_m s'écrit comme combinaison linéaire des $\varepsilon_{t-j}; j \geq 0$.

D'après la relation (1.14) on a :

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

ce qui permet d'écrire :

$$\begin{aligned} \hat{X}_t(l) &= \sum_{k=0}^{+\infty} \hat{\psi}_k \varepsilon_{t-k} \\ &= \sum_{j=l}^{+\infty} \hat{\psi}_j \varepsilon_{t+l-j} \dots \dots \dots (4.2) \end{aligned}$$

dans le cadre choisit il est bien naturel de prendre celle de ces combinaisons linéaires qui est la plus proche de la variable aléatoire X_{t+l} . Dans notre travail on prend la combinaison linéaire qui minimise l'erreur moyenne quadratique définie par :

$$E[X_{t+l} - \hat{X}_t(l)]^2 = E(X_{t+l} - \sum_{j=l}^{+\infty} \hat{\psi}_j X_{t+l-j})^2 \dots \dots \dots (4.3)$$

et pour définir le prédicteur il suffit de calculer les valeurs des $\hat{\psi}_j, j \geq l$.

IV.2- Prédicteur de l'erreur moyenne quadratique minimum

IV.2.1- Déviation de l'erreur moyenne quadratique du prédicteur

Supposons qu'on est à l'instant t et qu'on veut estimer une future valeur X_{t+l} du processus X_t par $\hat{X}_t(l)$ sachant que ce dernier s'écrit sous la forme d'une combinaison linéaire des $\varepsilon_{t-j}; j \geq 0$.

Notons par V l'erreur moyenne quadratique de l'estimateur $\hat{X}_t(l)$.

$$\begin{aligned} V &= E\left(X_{t+l} - \hat{X}_t(l)\right)^2 \\ &= E\left[\left(\sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j}\right) + \left(\sum_{j=l}^{+\infty} (\psi_j - \hat{\psi}_j) \varepsilon_{t+l-j}\right)\right]^2 \\ &= \sigma_\varepsilon^2 \sum_{j=0}^{l-1} \psi_j^2 + \sigma_\varepsilon^2 \sum_{j=l}^{+\infty} (\psi_j - \hat{\psi}_j)^2 \dots\dots\dots(4.4) \end{aligned}$$

de (4.4) le minimum de la fonction d'erreur moyenne quadratique (V) est atteint quand $\psi_j = \hat{\psi}_j; \forall j \geq l$.

de ce dernier résultat on déduit que :

$$\hat{X}_t(l) = \sum_{j=l}^{+\infty} \psi_j \varepsilon_{t+l-j} \dots\dots\dots(4.5)$$

IV.2.2- Conformité du prédicteur de L'erreur moyenne quadratique minimale et de l'espérance conditionnelle

Proposition IV.1 :

Si les $\varepsilon_{t-j}; j \geq 0$ sont des variables aléatoires indépendantes, le prédicteur de l'erreur moyenne quadratique minimal coïncide avec l'espérance conditionnelle de la future valeur du processus X_{t+l} sachant ses précédentes valeurs $X = (X_{t-j})_{j \geq 0}$

$$\hat{X}_t(l) = E\left(\frac{X_{t+l}}{X_{t-j}; j \geq 0}\right)$$

Preuve :

On' a :

$$\begin{aligned} X_{t+l} &= \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t+l-j} \quad \text{et} \quad \hat{X}_t(l) = \sum_{k=l}^{+\infty} \psi_k \varepsilon_{t+l-k} = \sum_{j=0}^{+\infty} \psi_{j+l} \varepsilon_{t-j} \\ E\left(\frac{X_{t+l}}{X_{t-j}; j \geq 0}\right) &= E\left(\frac{\sum_{j=0}^{+\infty} \psi_j \varepsilon_{t+l-j}}{\varepsilon_{t-j}; j \geq 0}\right) \\ &= E\left(\frac{\sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} + \sum_{j=l}^{+\infty} \psi_j \varepsilon_{t+l-j}}{\varepsilon_{t-j}; j \geq 0}\right) \\ &= E\left(\frac{\sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j}}{\varepsilon_{t-j}; j \geq 0}\right) + E\left(\frac{\sum_{j=l}^{+\infty} \psi_j \varepsilon_{t+l-j}}{\varepsilon_{t-j}; j \geq 0}\right) \\ &= E\left(\frac{\sum_{j=l}^{+\infty} \psi_j \varepsilon_{t+l-j}}{\varepsilon_{t-j}; j \geq 0}\right) = \sum_{j=l}^{+\infty} \psi_j \varepsilon_{t+l-j} = \hat{X}_t(l) \end{aligned}$$

par conséquent dans le cas où le processus $\varepsilon_{t-j}; j \geq 0$ est un bruit blanc, le prédicteur de l'erreur moyenne quadratique minimal à l'horizon l coïncide avec l'espérance conditionnelle de X_{t+l} sachant $X_{t-j}; j \geq 0$.

IV.2.3- Procédure de calcul du prédicteur

Pour obtenir le prédicteur $\hat{X}_t(l)$ de X_{t+l} on suit les étapes suivantes :

1\ On écrit le modèle de X_{t+l} sous une de ses trois formes.

2\ On garde les $X_{t-j}; j \geq 0$ qui s'est produits avant l'instant t tel qu'ils sont.

3\ Les $X_{t+j}; j \geq 0$ qui s'est produits après l'instant t seront remplacés par $\hat{X}_t(j)$

4\ Les $\varepsilon_{t-j}; j \geq 0$ qui s'est produits avant l'instant t sont disponibles à partir de la relation

$$\varepsilon_{t+1} = X_{t+1} - \hat{X}_t(1)$$

5\ On remplace les $\varepsilon_{t+j}; j \geq 0$ qui s'est produits après l'instant t par zéro.

IV.3- Trois formes explicites des futures valeurs du processus ARMA

On a vu précédemment que le processus ARMA se représente sous trois formes explicites ; de la même façon on va représenter ses futures valeurs dont chacune des trois formes est utile dans un des problèmes de la prévision.

Pour la présentation des futures valeurs de l'ARMA on utilise les mêmes relations de présentation de l'actuel valeur du processus avec le remplacement de l'instant t par $t+l$.

Dans le premier chapitre on a vu que le processus ARMA se représente sous trois formes possibles qui sont :

1\ La forme d'équation différentielle

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

2\ La forme de processus en fonction de bruit blanc

$$X_t = \varepsilon_t + \sum_{j=1}^{+\infty} \psi_j \varepsilon_{t-j} \quad \text{avec} \quad \psi_j = \phi_1 \psi_{j-1} + \phi_2 \psi_{j-2} + \dots + \phi_p \psi_{j-p} \quad \text{sachant que} \quad \psi_0 = 1,$$

$$\psi_j = 0 \forall j \leq 0 \quad \text{et} \quad \theta_j = 0 \forall j \geq q+1.$$

3\ Le bruit en fonction des processus (La forme inversé)

$$\varepsilon_t = \pi(B)X_t = X_t + \pi_1 X_{t-1} + \pi_2 X_{t-2} + \dots$$

Avec :

$$\pi_j = \theta_1 \pi_{j-1} + \theta_2 \pi_{j-2} + \dots + \theta_q \pi_{j-q} + \phi_j \quad \forall j > 0 \quad \text{sachant que } \pi_j = 0 : \forall j > 0$$

$$\phi_j = 0 \quad \forall j > p$$

De ces trois formes on tire trois formes pour représenter les futures valeurs qui sont :

IV.3.1- En fonction d'équation différentielle

$$X_{t+l} = \phi_1 X_{t+l-1} + \phi_2 X_{t+l-2} + \dots + \phi_p X_{t+l-p} + \varepsilon_{t+l} + \theta_1 \varepsilon_{t+l-1} + \dots + \theta_q \varepsilon_{t+l-q} \dots (4.7)$$

IV.3.2- En tant que somme pondérée infinie de l'actuel et précédents bruits

$$\begin{aligned} X_{t+l} &= \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t+l-j} / \psi_0 = 1 \\ &= \sum_{j=0}^{t+l-k-1} \psi_j \varepsilon_{t+l-j} + C_k(t+l-k) \quad [\text{La forme tronquée}] \\ &= \sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} + C_t(l) \dots (4.8) \quad (\text{Pour } t = k) \end{aligned}$$

on a démontré précédemment que la fonction complémentaire change quand on change d'origine.

IV.3.3- En tant que somme pondéré infinie d'observations précédentes et un bruit aléatoire

$$\varepsilon_{t+l} = \pi(B)X_{t+l} = X_{t+l} + \pi_1 X_{t+l-1} + \pi_2 X_{t+l-2} + \dots (4.9)$$

IV.4- Trois formes de base pour la prévision de l'ARMA(p,q)

On a vu dans la proposition (IV.1) que sous la condition d'indépendance des bruits blancs le prédicteur de l'erreur moyenne quadratique minimal $\hat{X}_t(l)$ coïncide avec l'espérance conditionnelle de X_{t+l} sachant $X_{t-j}; j \geq 0$; ce qui permet de le définir à partir des autres formes de l'ARMA :

IV.4.1- Prévisions à partir d'équation différentielle

On 'a :

$$\hat{X}_t(l) = E\left(\frac{X_{t+l}}{X_{t-j}; j \geq 0}\right)$$

$$= \begin{cases} \phi_1 \hat{X}_t(l-1) + \dots + \phi_p \hat{X}_t(l-p) - \left(\sum_{j=1}^q \theta_j \varepsilon_{t+l-j}\right) & \text{si } l \leq q \\ \phi_1 \hat{X}_t(l-1) + \dots + \phi_p \hat{X}_t(l-p) & \text{si } l > q \end{cases} \dots\dots\dots(4.10)$$

IV.4.2- Prévisions en forme intégrée de X_t

$$X_{t+l} = \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t+l-j} \Rightarrow \hat{X}_t(l) = E\left(\frac{X_{t+l}}{X_{t-j}; j \geq 0}\right)$$

$$\Rightarrow \hat{X}_t(l) = E\left(\frac{X_{t+l}}{\varepsilon_{t-j}; j \geq 0}\right)$$

$$\Rightarrow \hat{X}_t(l) = E\left(\frac{\sum_{i=0}^{+\infty} \psi_i \varepsilon_{t+l-i}}{\varepsilon_{t-j}; j \geq 0}\right)$$

$$\Rightarrow \hat{X}_t(l) = \sum_{i=0}^{+\infty} \psi_i E\left(\frac{\varepsilon_{t+l-i}}{\varepsilon_{t-j}; j \geq 0}\right)$$

$$\Rightarrow \hat{X}_t(l) = \sum_{i=l}^{+\infty} \psi_i \varepsilon_{t+l-i}$$

Comme on peut le calculer à partir de la forme tronquée comme suit :

$$\begin{aligned}
 X_{t+l} &= \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t+l-j} & / \psi_0 = 1 \\
 &= \sum_{j=0}^{t+l-k-1} \psi_j \varepsilon_{t+l-j} + C_k(t+l-k) \\
 &= \sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} + C_t(l) \text{ (Quand } t = k \text{)}
 \end{aligned}$$

ce qui implique que :

$$\begin{aligned}
 \hat{X}_t(l) &= E \left(\frac{\sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} + C_t(l)}{\varepsilon_{t-i}; i \geq 0} \right) \\
 &= \sum_{j=0}^{l-1} \psi_j E \left(\frac{\varepsilon_{t+l-j}}{\varepsilon_{t-i}; i \geq 0} \right) + E \left(\frac{C_t(l)}{\varepsilon_{t-i}; i \geq 0} \right) \\
 &= E \left(\frac{C_t(l)}{\varepsilon_{t-i}; i \geq 0} \right) \\
 &= C_t(l) \dots \dots \dots (4.12)
 \end{aligned}$$

où $C_t(l)$ représente la fonction complémentaire à l'origine t .

IV.4.3- Prévisions en moyenne pondérée d'observations précédentes

$$\begin{aligned}
 \varepsilon_{t+l} &= X_{t+l} - \sum_{j=1}^{+\infty} \pi_j X_{t+l-j} \\
 \Rightarrow X_{t+l} &= \varepsilon_{t+l} + \sum_{j=1}^{+\infty} \pi_j X_{t+l-j}
 \end{aligned}$$

$$\begin{aligned} \Rightarrow \hat{X}_t(l) &= E \left(\frac{\varepsilon_{t+l} + \sum_{j=1}^{+\infty} \pi_j X_{t+l-j}}{X_{t-i}; i \geq 0} \right) \\ &= E \left(\frac{\varepsilon_{t+l}}{\varepsilon_{t-i}; i \geq 0} \right) + \sum_{j=1}^{+\infty} \pi_j E \left(\frac{X_{t+l-j}}{X_{t-i}; i \geq 0} \right) \\ &= \sum_{j=1}^{+\infty} \pi_j \hat{X}_t(l-j) \dots \dots \dots (4.13) \end{aligned}$$

sous la condition d'inversement de l'ARIMA (4.13) est une série convergente.

Dans la pratique on ne prend en considération que les $X_{t-j}; k \geq j \geq 0$ du fait que la corrélation entre X_{t+l} et $X_{t-j}; j \geq k+1$ peut-être négligé.

IV.5- Prédiction de l'ARIMA(p,d,q)

Par définition de processus ARIMA(p,d,q) ses différences d'ordre d sont un ARMA qui est un processus stationnaire, donc pour prévoir le ARIMA on prévois le ARMA puis on extrait les prédictions de l'ARIMA à partir de celles de l'ARMA.

Supposons que le processus X_t est un ARIMA(p,d,q) et que $W_t = (1-B)^d X_t$ est un ARMA(p,q) stationnaire .

Le prédicteur de l'ARIMA(p,d,q) est donné par la relation :

$$\hat{X}_t(l) = \begin{cases} \hat{W}_t(l) - \sum_{k=1}^l C_d^k (-1)^k X_{t-k+1} + \sum_{k=l+1}^d C_d^k (-1)^k \hat{X}_t(l-k) & \text{si } l \leq d \\ \hat{W}_t(l) - \sum_{k=1}^d C_d^k (-1)^k \hat{X}_t(l-k) & \text{si } l > d \end{cases} \dots\dots\dots(4.14)$$

Preuve :

$$\begin{aligned} W_t &= (1-B)^d X_t \\ &= [(1-B)^d X_t - X_t] + X_t \\ \Rightarrow X_t &= W_t - [(1-B)^d X_t - X_t] \\ &= W_t - \left[\sum_{k=0}^d C_d^k (-B)^k X_t - X_t \right] \\ &= W_t - \left[C_0^d X_t + \sum_{k=1}^d C_d^k (-B)^k X_t - X_t \right] \\ &= W_t - \sum_{k=1}^d C_d^k (-B)^k X_t \\ &= W_t - \sum_{k=1}^d C_d^k (-1)^k X_{t-k} \\ \Rightarrow X_{t+l} &= W_{t+l} - \sum_{k=1}^d C_d^k (-1)^k X_{t-k+l} \end{aligned}$$

$$\Rightarrow \hat{X}_t(l) = \begin{cases} \hat{W}_t(l) - \sum_{k=1}^l C_d^k (-1)^k X_{t-k+1} + \sum_{k=l+1}^d C_d^k (-1)^k \hat{X}_t(l-k) & \text{si } l \leq d \\ \hat{W}_t(l) - \sum_{k=1}^d C_d^k (-1)^k \hat{X}_t(l-k) & \text{si } l > d \end{cases}$$

IV.6- L'erreur du prédicteur à l'horizon l

IV.6.1- Calcul de l'erreur du prédicteur à l'horizon l

On reprend la forme des futures valeurs en fonction des bruits blancs

$$X_{t+l} = \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t+l-j}$$

On a vu que le prédicteur est :

$$\hat{X}_t(l) = \sum_{j=l}^{+\infty} \psi_j \varepsilon_{t+l-j}$$

des deux dernières formules il s'ensuit :

$$\begin{aligned} X_{t+l} &= \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t+l-j} \\ &= \sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} + \sum_{j=l}^{+\infty} \psi_j \varepsilon_{t+l-j} \\ &= \sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} + \hat{X}_t(l) \end{aligned}$$

d'où si on note par $e_t(l)$ l'erreur de prédicteur à l'horizon l on écrit :

$$\begin{aligned} e_t(l) &= X_{t+l} - \hat{X}_t(l) \\ &= \sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} \dots\dots\dots(4.15) \end{aligned}$$

Espérance et variance de l'erreur de prédicteur

a) Espérance

$$\begin{aligned}e_t(l) &= \sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} \\ \Rightarrow E[e_t(l)] &= E\left[\sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} \right] \\ &= \sum_{j=0}^{l-1} \psi_j E(\varepsilon_{t+l-j}) \\ &= 0 \dots \dots \dots (4.16)\end{aligned}$$

b) Variance

$$\begin{aligned}e_t(l) &= \sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} \\ \Rightarrow \text{var}[e_t(l)] &= E[e_t(l) - E(e_t(l))]^2 \\ &= E[e_t(l)]^2 \\ &= E\left[\sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j} \right]^2 \\ &= \sum_{j=0}^{l-1} \psi_j^2 E(\varepsilon_{t+l-j}^2) \\ &= \sigma_\varepsilon^2 \sum_{j=0}^{l-1} \psi_j^2 \dots \dots \dots (4.17)\end{aligned}$$

IV.6.2- Évolution de la variance d'erreur pour les ARMA

On voit clairement que la variance d'erreur de prévision croît avec l et que la valeur de la variance d'erreur de prédictor obtenu pour $l = 1$ est la plus petite. Lorsque $l \rightarrow \infty$, la variance d'erreur tend vers :

$$\begin{aligned} VAR_{\infty} &= \lim_{l \rightarrow +\infty} \sigma_{\varepsilon}^2 (1 + \psi_1^2 + \dots + \psi_{l-1}^2) \\ &= \sigma_{\varepsilon}^2 \left(1 + \sum_{j=1}^{+\infty} \psi_j^2 \right) \dots\dots\dots (4.18) \end{aligned}$$

ce qui est bien normal puisque, étant donné que le prédictor tend vers zéro quand l'horizon tend vers l'infini, l'erreur de prévision finit par ressembler à la variance de processus (pour un processus moyenne mobile, cette valeur est atteinte dès que l'horizon dépasse l'ordre).

Pour résumer :

Proposition IV.2 :

La variance d'erreur de prévision à horizon l dans un processus ARMA croît depuis la variance du bruit blanc (valeur prise pour $l = 1$) jusqu'à la variance du processus lui-même.

IV.7- Intervalles de confiance pour le prédictor

En Supposons maintenant que le bruit est gaussien, Autrement dit les ε_t sont des variables aléatoires gaussiennes, centrées, indépendantes et de variance σ_{ε}^2 . Il résulte que toutes les variables considérées ici (des combinaisons linéaires finies où infinies de ces ε_t) sont aussi gaussiennes.

Notamment l'erreur de prévision à l'horizon l est :

$$e_t(l) = X_{t+l} - \hat{X}_t(l) = \sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j}$$

et elle est gaussienne centrée de variance :

$$VAR_l = \sigma^2(1 + \psi_1^2 + \dots + \psi_{l-1}^2)$$

d'où :

$$\frac{X_{t+l} - \hat{X}_t(l)}{\sigma_\varepsilon^2(1 + \sum_{j=0}^{l-1} \psi_j^2)} \approx N(0, \sigma_\varepsilon^2)$$

il en résulte que l'intervalle de confiance est :

$$\left[\hat{X}_t(l) - u_{\frac{\alpha}{2}} \left(1 + \sum_{j=0}^{l-1} \psi_j^2\right)^{\frac{1}{2}} s_\varepsilon ; \hat{X}_t(l) + u_{\frac{\alpha}{2}} \left(1 + \sum_{j=0}^{l-1} \psi_j^2\right)^{\frac{1}{2}} s_\varepsilon \right]$$

à une probabilité $1 - \alpha$ de contenir X_{t+l} .

Cet intervalle est donc un intervalle de confiance (de niveau $1 - \alpha$) pour X_{t+l} , et on remarque, comme on s'y attendait, que la largeur de cet intervalle croît avec l .

s_ε^2 est un estimateur sans biais de σ_ε^2

Chapitre V : Mise à jour des prévisions et exemple d'application

V.1- Mise à jour des prévisions

La prévision des futures valeurs du processus X_t , à l'origine de temps t est donnée par $\hat{X}_t(l)$ qui est le prédicteur de l'erreur moyenne quadratique minimal, mais dès que X_{t+1} devient disponible on procède à une mise à jour de prédicteur (c.-à-d. : recherche de $\hat{X}_{t+1}(l)$) ce qui nécessite une relation entre le prédicteur à l'instant t et le prédicteur à l'instant $t+1$.

V.1.1- Relation entre les prédicteurs aux temps t et $t+1$

L'expression générale du prédicteur est :

$$\hat{X}_t(l) = \sum_{j=l}^{+\infty} \psi_j \varepsilon_{t+l-j}$$

ce qui permet d'écrire :

$$\begin{cases} \hat{X}_t(l+1) = \sum_{j=l+1}^{+\infty} \psi_j \varepsilon_{t+l-j+1} \\ \hat{X}_{t+1}(l) = \sum_{j=l}^{+\infty} \psi_j \varepsilon_{t+l-j+1} \end{cases}$$

$$\Rightarrow \hat{X}_{t+1}(l) - \hat{X}_t(l+1) = \sum_{j=l}^{+\infty} \psi_j \varepsilon_{t+l-j+1} - \sum_{j=l+1}^{+\infty} \psi_j \varepsilon_{t+l-j+1}$$

$$= \psi_l \varepsilon_{t+1}$$

donc :

$$\hat{X}_{t+1}(l) = \hat{X}_t(l+1) + \psi_l \varepsilon_{t+1} \dots \dots \dots (5.1)$$

d'après cette dernière relation on a besoin de la valeur de ε_{t+1} et les poids de la forme du processus en fonction de bruit blanc.

V.1.2- Calcul de la valeur du bruit à l'instant $t + 1$

L'erreur du prédicteur à l'horizon l est :

$$e_t(l) = X_{t+l} - \hat{X}_t(l) \\ = \sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j}$$

et pour $l = 1 / e_t(1) = \varepsilon_{t+1}$ donc $\varepsilon_{t+1} = X_{t+1} - \hat{X}_t(1) \dots \dots (5.2)$

V.1.3- Calcul des poids

Dans la mise à jour des prévisions on utilise les poids de la forme linéaire du processus et pour les calculer on a:

$$\begin{cases} \Phi(B)X_t = \Theta(B)\varepsilon_t \\ X_t = \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t-j} = \Psi(B)\varepsilon_t \end{cases}$$

$$\Rightarrow \Phi(B)\Psi(B)\varepsilon_t = \Theta(B)\varepsilon_t$$

$$\Leftrightarrow \Phi(B)\Psi(B) = \Theta(B) \dots \dots \dots (1)$$

de la relation (1) et par égalisation des coefficients de $B^j; j \geq 0$ on trouve :

$$\psi_j = \varphi_1 \psi_{j-1} + \dots \dots \dots + \varphi_{p+d} \psi_{j-p-d} - \theta_j / j \leq p + d - 1 \wedge \theta_j = 0 \forall j > q$$

$$\text{et } \psi_j = \varphi_1 \psi_{j-1} + \dots \dots \dots + \varphi_{p+d} \psi_{j-p-d} / j > p + d - 1 \wedge q$$

donc par récurrences on calcule les poids l'un après l'autre.

de la relation (5.1) dès que X_{t+1} est disponible on peut mettre l'estimateur à jour.

Remarque : Dans le cas où on a pré visionné les $\{X_{t+j}; j = 1; \dots; L\}$ on ne peut maitre à jour que les $\{X_{t+j}; j = 1; \dots; L-1\}$ prévisions, mais on ce qui concerne $\hat{X}_{t+1}(l)$ on le détermine à partir de l'équation différentielle.

V.2- Exemple d'application

V.2.1- Mise au point du programme de calcul des prévisions

Comme il a été présenté dans les chapitres de ce mémoire pour avoir les prévisions utilisant les modèles ARIMA (La méthode de Box-Jenkins) on est censé de suivre plusieurs étapes qui nécessitent des calculs parfois long ou difficile, afin d'éviter cela on a constitué utilisant le langage R un programme informatique standard qui calcule les prévisions automatiquement une fois qu'on lui donne l'ordre de différentiation (d).

R est un langage et environnement pour le calcul et graphiques statistiques. Il s'agit d'un Projet GNU qui est semblable au langage S et de l'environnement qui a été développé chez Bell Laboratoires (anciennement AT & T, désormais Lucent Technologies) par John Chambers et ses collègues. R peut être considéré comme une application différente de S. Il y a quelques différences importantes, mais trop de code écrit pour S s'exécute sans modification conformément à la règle

R offre une grande variété de statistiques (modélisation linéaire et non linéaire, tests statistiques classiques, analyse de séries chronologiques, la classification, le regroupement, ...) et techniques graphiques et il est très extensible. Le langage S est souvent le véhicule de choix pour la recherche dans la méthodologie statistique et R offre une voie Open Source pour la participation à cette activité.

Une des forces de R est la facilité avec laquelle bien conçue, à la publication des parcelles de qualité peuvent être produites, y compris les symboles mathématiques et les formules si nécessaire. Un grand soin a été pris en charge par défaut pour les choix de conception en images de mineurs, mais l'utilisateur conserve le contrôle intégral.

R est disponible comme logiciel libre sous les termes de la Free Software Foundations GNU General Public License sous forme de code source. Il compile et fonctionne sur une large variété de plates-formes UNIX et des systèmes similaires (y compris Free BSD et Linux), Windows et MacOS.

L'environnement R

R est une suite intégrée de services logiciels pour la manipulation des données, le calcul et l'affichage graphique. Il comprend

- une gestion efficace des données et de stockage,
- une suite d'opérateurs pour les calculs sur les tableaux, dans des matrices particulières,
- un vaste et cohérent, collection d'outils intégrés pour l'analyse intermédiaire des données,
- graphique installations d'analyse de données et afficher soit l'écran ou sur papier, et
- bien développé, le langage de programmation simple et efficace qui comprend les conditionnels, boucles, définie par l'utilisateur des fonctions récursives et d'entrée et des installations de production.

Le terme "environnement" est destinée à le caractériser comme un entièrement planifié et cohérent, plutôt qu'une accumulation progressive d'outils très spécifiques et rigides, comme c'est souvent le cas pour les logiciels d'autres données d'analyse.

R, comme S, est conçu autour d'un langage informatique vrai, et il permet aux utilisateurs d'ajouter des fonctionnalités supplémentaires en définissant de nouvelles fonctions. Une grande partie du système est lui-même écrit dans le dialecte R de S, ce qui le rend facile pour les utilisateurs de suivre les choix algorithmiques effectués. Pour les tâches de calcul intensif, C, C++ et Fortran peuvent être liés et appelés lors de l'exécution. Les utilisateurs avancés peuvent écrire du code C pour manipuler des objets R directement.

Beaucoup d'utilisateurs pensent de R comme système de statistique. Nous préférons penser de celui-ci d'un environnement dans lequel les techniques statistiques sont mises en œuvre. R peut être étendu (facilement) par l'intermédiaire de *forfaits*. Il ya environ huit paquets fournis avec la distribution de R et beaucoup d'autres sont disponibles à travers la famille CRAN de sites Internet couvrant un éventail très large de la statistique moderne.

R Latex a son propre format semblable à la documentation, qui est utilisée pour fournir une documentation complète, tant en ligne dans un certain nombre de formats et sur papier.

Dans notre travail l'avantage de ce logiciel est qu'il contient quelques fonctions prédéfinies comme la fonction `arima0` pour estimer les coefficients du modèle et la fonction `diff` qui calcule les différences d'ordre un (1) pour le processus, utilisant ces fonctions et quelque d'autre on a construit un programme qui calcule ce qui suit :

- 1 .Les ordres du processus `p,q`
- 2 .Les différences d'ordre `d`. `Q(d)`
- 3 .Les coefficients du processus ARMA(`p,q`). `arima0`
- 4 .Les résidus du modèle considéré. `bruit()`
- 5 .Les prévisions des différences d'ordre `d`. `pred()`
- 6 .Les prévisions de processus original. `pred.X()`
- 7 .Les poids de la forme de processus en fonction des bruits blanc `poids.arma()`
- 8 .La mise à jour pour les prévisions des différences. `mise.a.jour()`
- 9 .La mise à jour pour les prévisions du processus original. `pred.X.j()`
10. Les erreurs de prédictions pour le processus original. `erreur.pred()`

Remarque : Le programme est dans l'annexe

V.2.2- Application numérique

Pour l'application numérique on a choisit comme série chronologique le DOW-JONES qui est un indice boursier Américain.

Tableau de données :

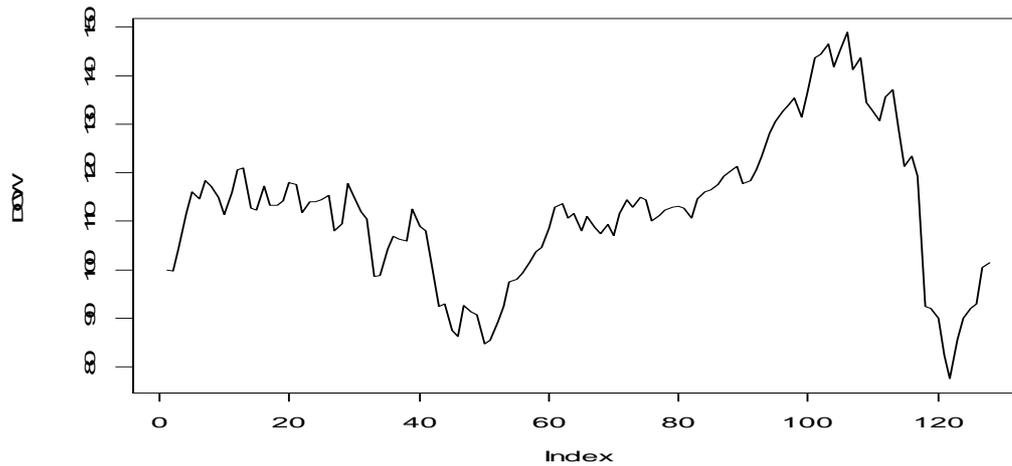
Le tableau de données ci-dessus présente la variation de DOW-JONES dès janvier 1999 jusqu'à septembre 2009.

Janvier 1999	100,00	Mai 2001	117,85
Février 1999	99,82	Juin 2001	115,31
Mars 1999	104,45	juillet 2001	111,91
Avril 1999	111,54	Août 2001	110,46
Mai 1999	116,09	Septembre 2001	98,53
Juin 1999	114,63	Octobre 2001	98,75
Juillet 1999	118,40	Novembre 2001	104,17
Août 1999	117,11	Décembre 2001	106,90
Septembre 1999	114,92	Janvier 2002	106,25
Octobre 1999	111,34	Février 2002	105,93
Novembre 1999	115,86	Mars 2002	112,41
Décembre 1999	120,51	Avril 2002	108,86
Janvier 2000	121,04	Mai 2002	107,96
Février 2000	112,73	Juin 2002	101,66
Mars 2000	112,27	Juillet 2002	92,48
Avril 2000	117,15	Août 2002	93,01
Mai 2000	113,18	Septembre 2002	87,65
Juin 2000	113,33	Octobre 2002	86,19
Juillet 2000	114,14	Novembre 2002	92,53
Août 2000	117,96	Décembre 2002	91,28
Septembre 2000	117,60	Janvier 2003	90,75
Octobre 2000	111,81	Février 2003	84,77
Novembre 2000	114,10	Mars 2003	85,44
Décembre 2000	114,07	Avril 2003	89,23
Janvier 2001	114,38	Mai 2003	92,34
Février 2001	115,40	Juin 2003	97,43
Mars 2001	107,96	Juillet 2003	98,00
Avril 2001	109,55	Août 2003	99,43

Septembre 2003	101,62	Octobre 2006	128,12
Octobre 2003	103,69	Novembre 2006	130,56
Novembre 2003	104,55	Décembre 2006	132,54
Décembre 2003	108,51	Janvier 2007	133,98
Janvier 2004	112,86	Février 2007	135,35
Février 2004	113,55	Mars 2007	131,39
Mars 2004	110,56	Avril 2007	136,49
Avril 2004	111,58	Mai 2007	143,63
Mai 2004	108,04	Juin 2007	144,36
Juin 2004	111,02	Juillet 2007	146,43
Juillet 2004	108,78	Août 2007	141,79
Août 2004	107,44	Septembre 2007	145,08
Septembre 2004	109,31	Octobre 2007	148,87
Octobre 2004	107,11	Novembre 2007	141,17
Novembre 2004	111,55	Décembre 2007	143,65
Décembre 2004	114,37	Janvier 2008	134,41
Janvier 2005	112,88	Février 2008	132,97
Février 2005	114,88	Mars 2008	130,67
Mars 2005	114,29	Avril 2008	135,54
Avril 2005	110,12	Mai 2008	137,05
Mai 2005	111,21	Juin 2008 (r)	129,12
Juin 2005	112,30	Juillet 2008	121,24
Juillet 2005	112,81	Août 2008	123,48
Août 2005	113,03	Septembre 2008	119,23
Septembre 2005	112,75	Octobre 2008	98,28
Octobre 2005	110,57	Novembre 2008	92,31
Novembre 2005	114,64	Décembre 2008	91,99
Décembre 2005	115,98	Janvier 2009	90,05
Janvier 2006	116,40	Février 2009	82,44
Février 2006	117,57	Mars 2009	77,49
Mars 2006	119,35	Avril 2009	85,63
Avril 2006	120,26	Mai 2009	89,88
Mai 2006	121,35	Juin 2009	92,02
Juin 2006	117,78	Juillet 2009	92,95
Juillet 2006	118,25	Août 2009	100,40
Août 2006	120,56	Septembre 2009	101,48
Septembre 2006	123,48		

On sait que le processus ARIMA fournit une bonne représentation pour n'importe quel processus stationnaire ou de différence stationnaire, donc avant de choisir le modèle ARIMA on doit d'abord voir s'il est stationnaire ou pas.

Présentation graphique de la trajectoire des données brutes :



Commentaire : à partir de la représentation graphique des données brutes il se voit clairement que le processus n'est pas stationnaire, pour cela on calcule les différences d'ordre 1.

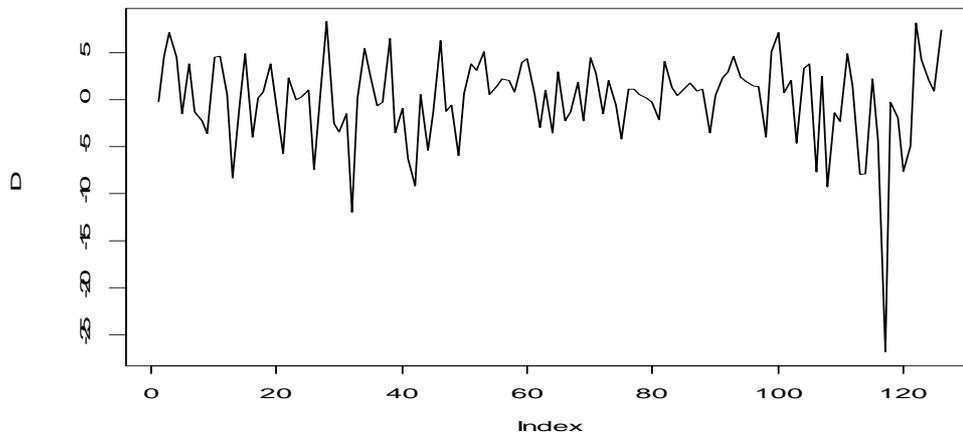
Tableau de données :

t	ε_t	t	ε_t
1	-0,18	13	-8,31
2	4,63	14	-0,46
3	7,09	15	4,88
4	4,55	16	-3,97
5	-1,46	17	0,15
6	3,77	18	0,81
7	-1,29	19	3,82
8	-2,19	20	-0,36
9	-3,58	21	-5,79
10	4,52	22	2,29
11	4,65	23	-0,03
12	0,53	24	0,31

t	ε_t	t	ε_t
25	1,02	58	0,86
26	-7,44	59	3,96
27	1,59	60	4,35
28	8,3	61	0,69
29	-2,54	62	-2,99
30	-3,4	63	1,02
31	-1,45	64	-3,54
32	-11,93	65	2,98
33	0,22	66	-2,24
34	5,42	67	-1,34
35	2,73	68	1,87
36	-0,65	69	-2,2
37	-0,32	70	4,44
38	6,48	71	2,82
39	-3,55	72	-1,49
40	-0,9	73	2
41	-6,3	74	-0,59
42	-9,18	75	-4,17
43	0,53	76	1,09
44	-5,36	77	1,09
45	-1,46	78	0,51
46	6,34	79	0,22
47	-1,25	80	-0,28
48	-0,53	81	-2,18
49	-5,98	82	4,07
50	0,67	83	1,34
51	3,79	84	0,42
52	3,11	85	1,17
53	5,09	86	1,78
54	0,57	87	0,91
55	1,43	88	1,09
56	2,19	89	-3,57
57	2,07	90	0,47

t	ε_t	t	ε_t
91	2,31	109	4,87
92	2,92	110	1,51
93	4,64	111	-7,93
94	2,44	112	-7,88
95	1,98	113	2,24
96	1,44	114	-4,25
97	1,37	115	-20,95
98	-3,96	116	-5,97
99	5,1	117	-0,32
100	7,14	118	-1,94
101	0,73	119	-7,61
102	2,07	120	-4,95
103	-4,64	121	8,14
104	-7,7	122	4,25
105	2,48	123	2,14
106	-9,24	124	0,93
107	-1,44	125	7,45
108	-2,3	126	1,08

Présentation graphique de la trajectoire des différences d'ordre un (1) :



Commentaire : d'après la Présentation graphique de trajectoire des différences d'ordre un à part une seule volatilité on observe une stationnarité, donc on peut représenter le DOW-JONES par le modèle ARIMA(p,1,q) qui prend la forme:

$$(1 - \phi_1 B)(1 - B)X_t = (1 + \theta_1 B)\varepsilon_t$$

$$\Leftrightarrow (1 - \phi_1 B)Y_t = (1 + \theta_1 B)\varepsilon_t \text{ (avec } Y_t = (1 - B)X_t \text{)}$$

$$\Leftrightarrow Y_t = \phi_1 Y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

Calcul des coefficients du processus ARMA(p, q)

Pour le calcul des coefficients de modèle on utilise la fonction `arima0` avec laquelle on a la possibilité d'utiliser deux méthodes d'estimation des paramètres qui sont :

1. La méthode de maximum de vraisemblance par la commande :

```
arima0(Q(d),order=c(p,0,q),method = c("MC"))
```

Pour notre tableau de données les coefficients de modèle sont :

$$\phi_1 = -0.7806 \text{ et } \theta_1 = 0.9831$$

2. La méthode des moindres carrés par la commande :

```
arima0(Q(d),order=c(p,0,q),method = c("CSS"))
```

Pour notre tableau de données les coefficients de modèle sont :

$$\phi_1 = -0.8096 \text{ et } \theta_1 = 0.9667$$

Calcul des erreurs de modèle (bruits)

Utilisant la fonction `bruit()` défini dans le programme on trouve les résidus autrement dit les erreurs de modèle par rapport à la série chronologique (DOW-JONES)

Dans notre tableau de données (Tableau des différences) on a 126 observations ce qui résulte 126 erreurs, dans le tableau suivant on ne présente que les 20 dernières.

Erreurs de modèle utilisant $\hat{\Phi}_{MC}, \hat{\Theta}_{MC}$:

t	ε_t	t	ε_t	t	ε_t
107	6.73313	113	21.42926	120	6.14341
108	-0.96094	114	-27.37414	121	-12.45967
109	-14.54472	115	22.28386	122	14.02447
110	14.62875	116	-32.1721	123	-13.31544
111	-23.9227	117	5.995695	124	21.60252
112	26.8313	118	-9.556975	125	-16.51341
113	-26.291	119	-14.49452	126	25.14671

Erreurs de modèle utilisant $\hat{\Phi}_{MV}, \hat{\Theta}_{MV}$:

t	ε_t	t	ε_t	t	ε_t
107	-4.788251	113	18.954873	120	6.148551
108	-10.54357	114	-25.33536	121	-12.50891
109	10.861087	115	20.956276	122	14.496683
110	-20.19024	116	-31.00294	123	-13.86544
111	23.594461	117	5.30687	124	22.125059
112	-23.4806	118	-8.854707	125	-17.50299
113	18.954873	119	-14.24949	126	26.327321

Calcul des prévisions :

Pour le calcul numérique des prévisions on utilise la fonction $\text{pred}()$ pour le processus de différences et $\text{pred.X}()$ pour le processus original.

On ce qui concerne la formule mathématique de la fonction de prévision des différences d'ordre un (1) elle est donnée par :

$$\begin{cases} \hat{Y}_t(1) = -0.7950Y_t + 0.9954\varepsilon_t \\ \hat{Y}_t(l) = -0.7950\hat{Y}_t(l-1); \forall l \geq 2 \end{cases}$$

A partir des prédictions du processus Y_t et de la relation $X_t = Y_t - X_{t-1}$ qui est équivalente à $Y_t = X_t + X_{t-1}$ on calcule les prédictions de processus X_t utilisant la relation :

$$\begin{cases} \hat{X}_t(1) = \hat{Y}_t(1) + X_t \\ \hat{X}_t(i) = \hat{Y}_t(i) + \hat{X}_t(i-1) \quad \forall i \geq 2 \end{cases}$$

par l'utilisation des deux fonctions mentionnées ci-dessus on obtient le tableau suivant :

Tableau des prévisions des différences d'ordre un (1) et de DOW JONES

1\ Utilisant $\hat{\Phi}_{MC}, \hat{\Theta}_{MC}$

L'horizon de prédiction (l)	Y(Les différences d'orde 1 du processus X)	X(DOW JONES)
1	-16.71713	83.68287
2	13.533709	97.21658
3	-10.9565	86.26008
4	8.870069	95.13015
5	-7.180954	87.94919
6	5.813494	93.76269
7	-4.706438	89.05625
8	3.810197	92.86645
9	-3.084626	89.78182
10	2.497225	92.27904

2\ Utilisant $\hat{\Phi}_{MV}, \hat{\Theta}_{MV}$

L'horizon de prédiction (l)	Y(Les différences d'ordre 1 du processus X)	X(DOW JONES)
1	-17.93277	82.46723
2	13.99872	96.46595
3	-10.92771	85.53824
4	8.5304173	94.06866
5	-6.659034	87.40962
6	5.1981911	92.60781
7	-4.057824	88.54999
8	3.1676282	91.71762
9	-2.472721	89.24489
10	1.9302616	91.17516

Calcul des poids : Pour la mise à jour des prévisions une on fois une qu'on est une nouvelle donnée nécessite les poids de la forme causal du modèle, on obtient les poids utilisant la fonction `poids.arma()` défini dans le programme et on trouve :

1\ Utilisant $\hat{\Phi}_{MC}, \hat{\Theta}_{MC}$

l	Le poids ψ_l	l	Le poids ψ_l
1	-1.7763145	6	0.2815338
2	0.6554057	7	-0.227922
3	-0.5305976	8	0.1845189
4	0.4295566	9	-0.149381
5	-0.3477567	10	0.1209347

2\ Utilisant $\hat{\Phi}_{MV}, \hat{\Theta}_{MV}$

l	Le poids ψ_l	l	Le poids ψ_l
1	-1.763699	6	0.2262799
2	0.6093713	7	-0.176639
3	-0.475689	8	0.1378884
4	0.3713333	9	-0.107639
5	-0.289871	10	0.0840253

Mise à jour de prédicteur

Dans le calcul numérique on utilise les fonctions **mise.a.jour()** pour calculer les mises à jour des prévisions pour les différences et **pred.X.j()** pour calculer les mises à jour des prévisions de processus original.

On ce qui concerne la formule mathématique utiliser dans la mise à jour des prévisions on a les deux relations suivantes :

$$\left\{ \begin{array}{l} \varepsilon_{t+1} = X_{t+1} - \hat{X}_t(1) \dots \dots \dots (1) \\ et \\ \hat{X}_{t+1}(l) = X_t(l+1) + \psi_l \varepsilon_{t+1} \dots \dots \dots (2) \end{array} \right.$$

L'application des deux fonctions précédentes (**mise.a.jour(l)**, **pred.X.j(l)**) pour l=15 nous donne les tableaux suivants :

1\ Utilisant $\hat{\Phi}_{MC}, \hat{\Theta}_{MC}$

L'horizon de prédiction (l)	Y(Les différences d'ordre 1 du processus X)	X(DOW JONES)
1	12.093924	113.5739
2	-10.26959	103.3043
3	8.016673	111.321
4	-6.257994	105.063
5	4.88513	109.9481
6	-3.813442	106.1347
7	2.976858	109.1116
8	-2.323802	106.7878
9	1.814012	108.6018
10	-1.416058	107.1857

2\ Utilisant $\hat{\Phi}_{MV}, \hat{\Theta}_{MV}$

L'horizon de prédiction (l)	Y(Les différences d'ordre 1 du processus X)	X(DOW JONES)
1	11.615289	113.0953
2	-10.24866	102.8466
3	8.297024	111.1436
4	-6.717032	104.4266
5	5.437917	109.8645
6	-4.402381	106.1347
7	3.564041	105.4622
8	-2.885346	109.0262
9	2.335893	106.1408
10	-1.891072	108.4767

L'évolution de l'erreur du prédicteur

1\ Utilisant $\hat{\Phi}_{MC}, \hat{\Theta}_{MC}$

L'horizon l	L'erreur $e_t(l)$
1	1.429557
2	1.71109
3	1.895609
4	2.016544
5	2.095805
6	2.147754
7	2.181801
8	2.204116
9	2.218741
10	2.228326

2\ Utilisant $\hat{\Phi}_{MV}, \hat{\Theta}_{MV}$

L'horizon l	L'erreur $e_l(l)$
1	4.110636
2	4.481969
3	4.708249
4	4.846137
5	4.930162
6	4.981365
7	5.012566
8	5.03158
9	5.043166
10	5.050226

ANNEXE

Programme de calcul

```
DOW<- c(100.00,99.82,104.45,111.54,116.09,114.63,118.40,117.11,114.92,  
  
111.34,115.86,120.51,121.04,112.73,112.27,117.15,113.18,113.33,114.14,117.96,117.60,11  
1.81,114.10,114.07,114.38,115.40,107.96,109.55,117.85,115.31,111.91,110.46,98.53,98.75,  
104.17,106.90,106.25,105.93,112.41,108.86,107.96,101.66,92.48,93.01,87.6,86.19,92.53,91  
.28,90.75,84.77,85.44,89.23,92.34,97.43,98.00,99.43,101.62,103.69,104.54,108.51,112.86,1  
13.55,110.56,111.58,108.04,111.02,108.78,107.44,109.31,107.11,111.55,114.37,112.88,114  
.88,114.29,110.12,111.21,112.30,112.81,113.03,112.75,110.57,114.64,115.98,116.40,117.5  
7,119.35,120.26,121.35,117.78,118.25,120.56,123.48,128.12,130.56,132.54,133.98,135.35,  
131.39,136.49,143.63,144.36,146.43,141.79,145.08,148.87,141.17,143.65,134.41,132.97,13  
0.67,135.54,137.05,129.12,121.24,123.48,119.23,92.31,91.99,90.05,82.44,77.49,85.63,89.8  
8,92.02,92.95,100.40,101.48)
```

```
d<-1
```

```
Q=function(i){
```

```
if(d==0){
```

```
A<-DOW
```

```
A}
```

```
else{
```

```
A<-DOW[1:length(DOW)-1]
```

```

re=rep(1:i)

for(i in 1:i)A=diff(A)

A}

A}

D<-Q(d)

p<-1

q<-1

model<-arima0(D,order=c(p,0,q),method = c("ML"))

z<-model$coef

ar.coef=z[1:p]

ma.coef=function(s){

re=rep(1,s)

for(i in 1:s)re[i]=z[p+i]

re}

a=function(t){

```

```
re=rep(1,p)

if (t>p)

for(i in 1:p) re[i]=D[t-i]

re}
```

```
bruit=function(n){

re=rep(1,n)

for(i in 1:p)re[i]=0

for(i in p+1:n){

g=function(i){

if(i<p+2){

l<-q-i+1

a<-data.frame(0,1:l)[,1]

b<-c(re[1:i-1])

f<-c(a,b)

f}

else{
```

```
s<-i-1

if(i<q){

b<-c(re[1:s])

ac<-q-i+1

a<-data.frame(0,1:ac)[,1]

f<-c(b,a)

f}

else{

t<-i-q

f<-c(re[t:s])

f}

f}

f}

k=function(i){

if (i<q+1){

b<-c(re[1:i])

s<-q-i+1
```

```

a<-data.frame(0,1:s)

if (i==1){

f<-as.matrix(data.frame(c(t(a[1])))

f}

if (i>1){

b<-c(re[1:i-1])

f<-as.matrix(data.frame(c(t(a[1]),b)))

f}

}

if(i==q+1){

f<-as.matrix(data.frame(c(re[1:q])))

}

if(i>q+1){

a<-as.matrix(data.frame(c(re[i-q+1:i-1])))

f<-a[1:q]

f}

f}

```

```

l<-p-q

if(l<0){

re[i]=re[i]=D[i]-(z[1:p]**a(i-1)+ma.coef(q)**g(i))

re}

else{

re[i]=re[i]=D[i]-(z[1:p]**a(i-1)+ma.coef(q)**k(i))

re}

re}

re}

re}

pred=function(l){

pr=rep(1,l)

a1<-length(D)-1

a2<-length(D)-q

pr[1]=z[1:p]**a(length(D))+ma.coef(q)**bruit(a1)[a2:a1]

kf=function(i){

if(i-p<0){

```

```

a3<-i-1

s<-a(length(D)+1)[p:1]

pr=z[1:a3]%%pr[a3:1]+z[p:i]%%s[i:p]

pr}

else{

if(i-p==0){

a3<-i-1

a4<-p-i+1

pr=z[1:a3]%%pr[a3:1]+z[p]%%a(length(D)+1)[1]

pr}

else{

if(i-p>0){

a3<-i-1

a8<-i-p

pr=z[1:p]%%pr[a3:a8]

pr}

pr}

```

```

pr}

pr}

kg=function(i){

if (q-i > 0){

a9<-length(DOW)+i-q

a10<-length(D)-p

a12<-length(D)

a11<-length(D)-q+i

pr=ma.coef(q)[i:q]%%bruit(a10)[a11:a12]

pr}

else{

if (q-i==0){

a6<-p+q

a9<-length(D)+i-q

pr=ma.coef(q)[q]*bruit(a9)[a9]

pr}

else{

```

```
if (q-i < 0){
```

```
pr=0
```

```
pr}
```

```
pr}
```

```
pr}
```

```
pr}
```

```
for(i in 2:l){
```

```
pr[i]=kg(i)+kf(i)
```

```
pr}
```

```
pr}
```

```
pred.X=function(l){
```

```
s=rep(1,l)
```

```
if(d==1){
```

```
s[1]=pred(l)[1]+DOW[length(DOW)-1]
```

```
for(i in 2:l){
```

```
h<-i-1
```

```

s[i]=pred(l)[i]+s[h]

s}

s}

else{

if(d==2){

s[1]=pred(l)[1]+2*DOW[length(DOW)-1]-DOW[length(DOW)-2]

s[2]=pred(l)[2]+2*s[1]-DOW[length(DOW)-1]

for(i in 3:l){

h<-i-1

f<-i-2

s[i]=pred(l)[i]+2*s[h]-s[f]

s}

s}

s}

s}

}

poids.ar=function(l){

re=rep(1,l)

```

```

re[1]=z[1]

for(i in 2:l){

if(i<p){

v<-c(re[i-1:1],1)

re[i]=z[1:i]%%v

re}

else{

if(i==p){

v<-c(re[i-1:1],1)

re[i]=z[1:i]%%v

re}

else{

if(i>p){

n<-i-p

m<-i-1

re[i]=z[1:p]%%re[m:n]

re}

```

```
re}
```

```
re}
```

```
re}
```

```
re}
```

```
poids.arma=function(l){
```

```
poids=rep(1:l)
```

```
r<-ma.coef(q)
```

```
if(q==0){
```

```
for(i in 1:l){
```

```
poids[i]=poids.ar(1)[i]
```

```
poids}
```

```
poids}
```

```
else{
```

```
if(q!=0){
```

```
for(i in 1:l){
```

```
if(i<q){
```

```
poids[i]=poids.ar(1)[i]-r[i]
```

```
}
```

```
else{
```

```
if(i==q){
```

```
poids[i]=poids.ar(1)[i]-r[i]
```

```
}
```

```
else{
```

```
if(i>q){
```

```
poids[i]=poids.ar(1)[i]
```

```
}
```

```
}
```

```
}
```

```
}
```

```
poids}
```

```
poids}
```

```
poids}
```

```
mise.a.jour=function(l){
```

```

ar=rep(1,1)

a<-DOW[length(DOW)]-DOW[length(DOW)-1]

for(i in 2:1){

S<-i-1

ar[i]=pred(1)[i]+(a*poids.arma(1)[S])

ar}

ar[2:1]}

pred.X.j=function(l){

s=rep(1,1)

if(d==1){

s[1]=mise.a.jour(1)[1]+DOW[length(DOW)]

for(i in 2:1){

h<-i-1

s[i]=mise.a.jour(1)[i]+s[h]

s}

s}

```

```

else{

if(d==2){

s[1]=mise.a.jour(l)[1]+2*DOW[length(DOW)-1]-DOW[length(DOW)-2]

s[2]=mise.a.jour(l)[2]+2*s[1]-DOW[length(DOW)-1]

for(i in 3:l){

h<-i-1

f<-i-2

s[i]=mise.a.jour(l)[i]+2*s[h]-s[f]

s}

s}

s}

s[1:l-1]}

```

```

s.c.p=function(l){

```

```

re=rep(1,l)

```

```

re[1]=1

```

```
for(i in 2:l){  
  
h<-i-1  
  
re[i]=re[h]+(poids.arma(l)[i])^2  
  
re}  
  
re}  
  
erreur.pred<-(length(D)/(length(D)-1)*var(bruit(127)[1:length(D)]))*(s.c.p(10))
```

Bibliographies

La présente bibliographie ne prétend pas l'exhaustivité et ne constitue qu'une sélection d'ouvrages et d'articles susceptibles d'aider le lecteur à une meilleure compréhension de notre travail.

Ouvrages

- [1] Christiaan Heij, Paul de Boer, Philip Hans Franses, Teun Kloek, Herman K. van Dijk : *Econometric Methods with Applications in Business and Economics, Oxford university* , (First published 2004).
- [2] Jianqing Fan and Qiwei Yao: *Non linear Time Series; Nonparametric and Parametric Methods, Springer-Verlag, New York, (2003).*
- [3] Clements. M.P. and Hendry. D.F : *A Companion to Economic Forecasting. Oxford: Blackwell, (2002).*
- [4] Burnham. K.P. and Anderson. D.R : *Model Selection and Multi-model Inference : A Practical Information Theoretic Approach, New York: Springer, (2002).*
- [5] Clements. M.P. and Hendry. D.F : *Forecasting Economic Time Series, Cambridge: Cambridge University Press, (1998).*
- [6] Peter J. Brockwell Richard A. Davis : *Introduction to Time Series and Forecasting, Springer-Verlag New York Inc, (1996).*
- [7] Gouriéroux. C. and Monfort. A : *Statistics and Econometric Models, vols. 1 and 2. Cambridge: Cambridge University Press, (1995).*
- [8] James D. Hamilton : *Times series analysis ,Princeton university press, (1994).*
- [9] T. Söderström : *Discrete-time Stochastic Systems ; Estimation and Control, Prentice Hall, (1994).*
- [10] Choi. B : *ARMA Model Identification, New York: Springer, (1992).*
- [11] Montgomery, Johnson and Gardiner : *Forecasting and Time series Analysis, McGraw-Hil, (1990).*
- [12] McCullagh. P. and Nelder. J.A : *Generalized Linear Models-2nd ed, New York : Chapman & Hall, (1989).*

- [13] Peter J. Brockwell and Richard A. Davis : Time series Theory and methods, *Springer-Verlag : New York Inc*, (1987).
- [14] Wayne A. Fuller : Introduction to Statistical Time Series , *John Wiley*, (1976).
- [15] Box. G.E.P. and Jenkins. G.M : Time Series Analysis; Forecasting and Control, 2nd edn, *San Francisco: Holden-Day*, (1976).
- [16] Box. G.E.P. and Tiao. G.C : Comparison of forecast and actuality, *Applied Statistics* 64, 195–200, (1976).
- [17] Chatfield. C : The Analysis of Time Series; Theory and Practice, *New York: Wiley*, (1975).

Articles

- [18] Nora Muler, Daniel Pena and Victor J. Yohai : Robust Estimation for ARMA models, *The Annals of Statistics*, Vol. 37, No. 2, 816–840, (2009).
- [19] Jean-Marie Dufour : Forecasting of stationary and ARIMA process, *McGill University*, (February 2008).
- [20] U. Grenander and M. Miller : Pattern Theory, From Representation to Inference, *Oxford University Press*, (2007).
- [21] Muler. N., Pena, D. and Yohai. V.J : Robust estimates for ARMA models, Technical report, (2007).
- [22] Yue. L, Chen. X. R : Asymptotic normality of quasi maximum likelihood estimate in generalized linear models, *Chinese Annals of Mathematics, Ser. B*, 26: 467-474, (2005).
- [23] Yin. C.M, Zhao. L.C & WEI. C.D : Strong consistency of maximum quasi-likelihood estimation in generalized linear models, *Science in China, Ser. A*, 1009-1014, (2005).
- [24] Fortes. M, Ninot. G & Delignières. D : The Auto-Regressive Integrated Moving Average procedures: Interest for APA study. *Adapted Physical Activity Quarterly*, 22,221-236, (2005).
- [25] Yue. L, Chen. X.R : Rate of strong consistency of quasi maximum likelihood estimate in generalized linear models, *Science in China, Ser. A*, 47(6): 882-893, (2004).
- [26] Diebold. F.X : Elements of Forecasting, Mason, OH: Thomson South-Western , (2004).

- [27] Javier Contier : ARIMA Models to Predict Next-Day Electricity Price , *IEEE Transactions on Power Systems Vol 18; No. 3*, (August, 2003).
- [28] Jean-Marie Durfour : Estimation de modèle ARMA par la méthode du maximum de vraisemblance, *Université de Montréal*, (2002).
- [29] E. Weiss : Forecasting commodity prices using ARIMA, *Technical Analysis of Stocks & Commodities, Vol. 18, No.1, pp.18-19*, (2000).
- [30] Didier Delignières : Séries temporelles – Modèles ARIMA, *Séminaire EA "Sport – Performance – Santé"*, (Mars 2000).
- [31] Chang. Y.I : Strong consistency of maximum quasi-likelihood estimate in generalized linear models via a last time, *Statist- Probab- Letters, 45: 237-246*, (1999).
- [32] Harvey. D.I, Leybourne. S.J. and Newbold. P : Testing the equality of prediction mean squared errors, *International Journal of Forecasting*, (1997).
- [33] F.Borgard and D.Guégan : Etudes des séries chronologiques à temps discret. Comparaison de logiciels, *Revue de statistiques appliquées, Tome 44, n°4, p.59-80*, (1996).
- [34] O. Torrès : Some Exact Inference Procedures for Stationary and Nonstationary Autoregressive Processes, *Proceedings of the Business and Economic Statistics Section of the American Statistical Association, Washington, DC, 135-140*, (1994).
- [35] Davidson. R. and MacKinnon. J.G : Estimation and Inference in Econometrics, *New York: Oxford University Press*, (1993).
- [36] Chatfield. C : Calculating Interval Forecasts (with discussion), *Journal of Business & Economic Statistics*, 11, 121-144, (1993).
- [37] Granger. C.W.J : Forecasting in Business and Economics, *Academic Press*, (1989).
- [38] Granger & Newbold : Forecasting economic time series, *Academic Press*, (1986).
- [39] Fahrmeir. L., Kanfmann. H : Consistency and asymptotic normality of the maximum likelihood estimator in generalized linear models, *Ann- Statist*, 13: 342-368, (1985).
- [40] Fahrmeir. L, Kanfmann. H : Consistency and asymptotic normality of the maximum likelihood estimator in generalized linear models, *Ann-Statist*, 13: 342-368, (1985).
- [41] Jones. R.H : Maximum likelihood fitting of ARMA models to time series with miss observations, *Technometrics* 22, 389-395, (1980).

- [42] Harvey. A.C., Phillips. G.D : Maximum likelihood estimation of regression models with autoregressive-moving average disturbances, *Biometrika* 66, 49-58, (1979).
- [43] Hannan. E.J. and Quinn. B : The determination of the order of an autoregression, *Journal of the Royal Statistical Society B* 41, 190–1, (1979).
- [44] Ljung, G.M. and Box. G.E.P : The likelihood function of stationary autoregressive-moving average models, *Biometrika* 66, 2, 265—270, (1979).
- [45] CRAIG F. ANSLEY : An algorithm for the exact likelihood of a mixed autoregressive-moving average process, *Biometrika* 66: 59-65, (1979).
- [46] A. C. HARVEY and G. D. A. PHILLIPS : Maximum likelihood estimation of regression models with autoregressive-moving average disturbances , *Biometrika* 66: 49-58, (1979).
- [47] Shibata. R : Selection of the order of an autoregressive model by Akaike's information criterion, *Biometrika* 71, 117–26, (1976).
- [48] NEWBOLD. P : The Exact Likelihood Function for a Mixed Autoregressive-Moving Average Process, *Biometrika*, 61, 423–426, (1974).
- [49] Akaike. H : Maximum likelihood identification of Gaussian autoregressive moving average models, *Biometrika* 60, 255-265, (1973).
- [50] Kendall. M.G : On autoregressive time series, *Biometrika*. 33, 105-122, (1944).

Conclusion

La présente étude nous a permis de développer, essentiellement, la prévision des modèles stochastiques par des processus auto-régressifs (AR, MA, ARMA, ARIMA), ce qui nous a amené à de résultats beaucoup plus élaborés et accessibles que ceux obtenus déjà dans ce domaine en conjuguant, bien entendu, simplicité et rigueur mathématique.

Ceci découle, entre autres, de la mise au point d'un programme, de calcul efficace et performant.

Ainsi l'importance de la présente étude se trouve illustrée au chapitre V par la mise à jour des prévisions et un exemple d'application numérique.

Par ailleurs, ce travail ne prétend pas d'être une fin en soi, mais ouvre des perspectives remarquables sur la recherche future dans ce domaine, notamment, dans le cas des processus stochastiques non linéaires.