

N° d'ordre : 30 2013-M/MT

République Algérienne Démocratique et Populaire
Ministère De L'Enseignement Supérieur Et De La Recherche
Scientifique
Université Des Sciences Et De La Technologie Houari Boumediene
Faculté Des Mathématiques



Mémoire
Présenté Pour L'Obtention du Dipôme de Magister
En Mathématiques

Spécialité : Statistique Mathématique & Probabilité

Par : Ouicher Fahima

Thème

Matrices de Toeplitz et Estimation en Séries Chronologiques

Soutenu publiquement le 14/07/2013, devant le jury composé de :

Mr. MOHAMED DJEDOUR	Professeur à l'USTHB	Président.
Mme K. DJABALLAH	Maitre de conférences à l'USTHB	Directrice de Mémoire
Mme. H.SAGGOU	Docteur à l'USTHB	Examinatrice
Mme.DJ.SEDDIKI	Docteur à l'USTHB	Examinatrice

Table Des Matières

	Page
Liste Des Tableaux	iii
Liste Des Figures.	iv
Remerciements.	vi
Résumé.	vii
Introduction Générale.	1
Chapitre 1. Les processus stationnaires et les matrices de Toeplitz	3
1.1 Introduction	3
1.2 Notions de base en séries chronologiques	4
1.2.1 La fonction d'autocovariance	4
1.2.2 La fonction d'autocorrélation	4
1.2.3 La stationnarité	5
1.2.4 Le processus Bruit Blanc	6
1.2.5 Opérateur de retard	7
1.3 Présentation des modèles stationnaires autorégressifs	7
1.3.1 Processus autoregressif AR(1).	7
1.3.2 Processus autoregressif AR(2).	8
1.3.3 Le processus autoregressif AR(p)	10
1.3.4 Causalité des modèles autoregressifs	10
1.4 Les matrices de Toeplitz	11
1.4.1 propriétés	12
1.4.2 L'inverse de la matrice de Toeplitz.	13
1.5 Conclusion.	16
Chapitre 2. Estimation des paramètres d'un modèle autoregressif	17
2.1 Introduction	17
2.2 La méthode de Yule-Walker.	18
2.3 La méthode du maximum de vraisemblance.	21

2.3.1	Le principe.	21
2.4	Application la méthode du maximum de vraisemblance au modèle autoregressif AR(1)	24
2.4.1	La méthode utilisant la décomposition de Cholesky	25
2.4.2	La méthode utilisant l'inverse d'une matrice de Toeplitz.	34
2.6	Conclusion.	54
Chapitre 3. Exemples et illustrations numériques.		55
3.1	Introduction et notation.	55
3.2	Modèle et Critères de performance	56
3.2.1	Le biais	56
3.2.2	La racine de L'erreur quadratique moyenne.	57
3.3	Analyse des qualités de l'estimateur MVC	58
3.3.1	Résultats numériques	58
3.3.2	La Distribution asymptotique de l'estimateur du MVC	59
3.3.3	Interprétation des résultats	60
3.4	Analyse des qualités de l'estimateur MVTP.	60
3.4.1	Résultats numériques.	61
3.4.2	La Distribution asymptotique de l'estimateur du MVTP	63
3.4.3	Interprétation des résultats.	63
3.5	Analyse des qualités de l'estimateur EYW	64
3.5.1	Résultats numériques	64
3.5.2	La Distribution asymptotique de l'estimateur du EYW.	66
3.5.3	Interprétation des résultats.	66
3.6	Etude comparative	67
3.6.1	Résultats numériques.	68
3.6.2	Interprétation des résultats.	74
3.7	Conclusion.	74
Conclusion Générale		75
Annexe		77
Bibliography.		81

Liste Des Tableaux

3.1	Les moyennes des estimateurs du MVC des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 10000 répétitions.	58
3.2	Le biais des estimateurs du MVC des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 10000 répétitions.	59
3.3	La racine de l'erreurs quadratiques moyennes des estimateurs du MVC des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 10000 répétitions.	59
3.4	Les moyennes des estimateurs du MVTP des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 100 répétitions	62
3.5	Le biais des estimateurs du MVTP des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 100 répétitions.	62
3.6	La racine de l'erreurs quadratiques moyennes des estimateurs du MVTP des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 100 répétitions.	62
3.7	Les moyennes des estimateurs du EYW des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 10000 répétitions.	65
3.8	Le biais des estimateurs du EYW des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 10000 répétitions.	65
3.9	La racine de l'erreurs quadratiques moyennes des estimateurs du EYW des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 10000 répétitions.	65

Liste Des Figures

3.1	Histogramme de l'estimateur du MVC de paramètre $\Phi = 0.1$ à partir de 10000 réalisations pour $n=50$	60
3.2	Histogramme de l'estimateur du MVC de paramètre $\Phi = 0.1$ à partir de 10000 réalisations pour $n=10000$	60
3.3	Histogramme de l'estimateur du MVTP de paramètre $\Phi = 0.1$ à partir de 100 réalisations pour $n=50$	63
3.4	Histogramme de l'estimateur du MVTP de paramètre $\Phi = 0.1$ à partir de 100 réalisations pour $n=1000$	63
3.5	Histogramme de l'estimateur du EYW de paramètre $\Phi = 0.1$ à partir de 10000 réalisations pour $n=50$	66
3.6	Histogramme de l'estimateur du EYW de paramètre $\Phi = 0.1$ à partir de 10000 réalisations pour $n=10000$	66
3.7	Synthse de biais de l'estimateur $\hat{\Phi}$ obtenus en fonction de n (première cas).	68
3.8	Synthse de biais de l'estimateur $\hat{\sigma}^2$ obtenus en fonction de n (première cas).	68
3.9	Synthse de REQM de l'estimateur $\hat{\Phi}$ obtenus en fonction de n (première cas).	69
3.10	Synthse de REQM de l'estimateur $\hat{\sigma}^2$ obtenus en fonction de n (première cas)	69
3.11	Synthse de biais de l'estimateur $\hat{\Phi}$ obtenus en fonction de n (deuxième cas).	70
3.12	Synthse de biais de l'estimateur $\hat{\sigma}^2$ obtenus en fonction de n (deuxième cas).	70
3.13	Synthse de REQM de l'estimateur $\hat{\Phi}$ obtenus en fonction de n (deuxième cas).	71
3.14	Synthse de REQM de l'estimateur $\hat{\sigma}^2$ obtenus en fonction de n (deuxième cas).	71
3.15	Synthse de biais de l'estimateur $\hat{\Phi}$ obtenus en fonction de n (troisième cas)	72

3.16	Synthse de biais de l'estimateur $\hat{\Phi}$ obtenus en fonction de n (troisième cas)	72
3.17	Synthse de REQM de l'estimateur $\hat{\Phi}$ obtenus en fonction de n (troisième cas)	73
3.18	Synthse de REQM de l'estimateur $\hat{\sigma}^2$ obtenus en fonction de n (troisième cas)	73

Remerciements

Je remercie vivement Madame *Djaballah Khadidja* maitre de conférence à L'USTHB 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'elle m'a prodigué, que j'ai pu mener ce travail à bien. Qu'elle trouve ici ma profonde gratitude.

Je reste très reconnaissant à Monsieur *Djedour Mohamed*, professeur à L'USTHB, pour l'honneur qu'il ma fait en acceptant de présider le jury de ce mémoire.

Je remercie également et profondément Madame *Saggou Hafida*, docteur à L'USTHB, pour l'intérêt qu'elle a bien voulu porter à ce travail et d'avoir accepté de le juger.

Je remercie aussi vivement Madame *Merad Djenat*, docteur à L'USTHB, pour l'intérêt qu'elle a bien voulu porter à ce travail et d'avoir accepté de le juger.

Je remercie aussi ma famille en particulier mon mari et mes parents pour leurs soutiens chaleureux apporté lors de la réalisation de ce travail.

Enfin, je tiens à remercier les collègues de la faculté de mahtématique pour leur soutien.

Résumé

L'objet principal du présent mémoire consiste à étudier les problèmes de l'estimation des paramètres de processus linéaire gaussien autoregressif à partir d'une suite de n observations. L'un des problèmes rencontrés est celui d'exprimer explicitement la fonction de vraisemblance et de trouver la forme analytique de l'estimateur du maximum de vraisemblance. Pour cela, on élabore deux méthodes.

La première, utilisant la décomposition de Cholesky qui nous a permis de trouver la forme analytique de l'estimateur.

La deuxième méthode utilisant l'inverse de la matrice de Toeplitz, basée essentiellement sur le calcul de l'inverse de la matrice variance-covariance et son déterminant.

Et enfin on valide les résultats par des programmes de simulation. Une étude comparative entre les performances des estimateurs du maximum de vraisemblance développés dans ce mémoire et ceux de Yule Walker sera mise en oeuvre.

Mots-clés :

Séries chronologiques- Processus aléatoires- Modèles AR(p)- Matrice de Toeplitz- La fonction de vraisemblance- Estimateur de maximum de vraisemblance- Décomposition de Cholesky- La méthode de Cardon- Normalité asymptotique- Convergence.

Introduction Générale

L'un des axes de recherche fondamental en statistiques est la création de modèles pour expliquer le comportement de phénomènes aléatoires, comprendre le hasard et réduire les incertitudes. Les processus aléatoires ont ainsi fourni des modèles pour analyser de nombreuses données issues de l'économie (1971), la biologie(1960), la théorie du signala (1963), la météorologie(1968) et l'astronomie (1906).

Le travail que nous présentons dans le cadre de cette étude, concerne une classe particulière de processus linéaires stationnaires autorégressifs très employée. Cette classe construite à partir de l'idée que l'observation au temps t s'explique linéairement par les observations précédentes.

Nous nous intéressons principalement au problème d'estimation des paramètres de processus linéaire gaussien autorégressif AR (p) à partir d'une suite de n observations. Diverses méthodes d'estimations des paramètres inconnus peuvent être utilisées. Parmi ces méthodes, Yule Walker et le Maximum de Vraisemblance.

Yulle Walker est une méthode autorégressive consiste à reprendre les équations de Yule-Walker en inversant les relations, on exprime les coefficients en fonction des autocovariances.

Maximum de Vraisemblance est la recherche des valeurs pour ces paramètres qui maximisent la fonction de vraisemblance. Mais cette méthode est délicate car la fonction de vraisemblance est très complexe et n'a pas de dérivée analytique à cause du calcul de l'inverse de la "grosse" matrice variance-covariance et son déterminant.

Dans le présent mémoire, Nous répondrons aux trois questions suivantes :

1. Comment détecter l'expression explicite de la fonction de vraisemblance ?
2. Peut-on trouver la forme analytique de l'estimateur de maximum de vraisemblance des paramètres de processus linéaire gaussien Autorégressif ?
3. Cet estimateur a-t-il des propriétés asymptotiques (la convergence, la normalité asymptotique et l'efficacité) ?

Le premier chapitre est consacré à la présentation des notions de base en séries chronologiques par la suite nous donnons des propriétés mathématiques des modèles autorégressifs, la définition de la matrice de Toeplitz et l'étude de ces propriétés. On a suggéré aussi une méthode de Xiao-GuangLv, Ting-zhu Huang , qui propose l'inverse d'une matrice de Toeplitz.

Dans le seconde chapitre, nous avons abordé l'étude de l'estimation obtenue à partir de la vraisemblance exacte ,premièrement il s'agit de construire la fonction de vraisemblance deuxièmement, il s'agit de trouver l'estimateur qui maximise cette fonction de vraisemblance, mais la maximisation et même le calcul de cette vraisemblance étaient relativement difficile, en particulier à cause du calculs l'inverse de la matrice covariance et son déterminant, surtout lorsque n devient relativement grand.

Pour résoudre ce problème, on utilise deux méthodes :

1. La méthode utilisant la décomposition de Cholesky qui nous a permis de trouver la forme analytique de l'estimateur.
2. La méthode utilisant l'inverse de la matrice de Toeplitz basée essentiellement sur le calcul de l'inverse de la matrice variance-covariance, nous utilisons le théorème de *Xiao-Guang Lv, Ting-zhu Huang*, et avec cette dernière nous obtenons l'expression explicite de la fonction de vraisemblance est la même de celle de Cholesky, mais la maximisation de la vraisemblance étant réalisé par des méthodes numériques.

Dans le dernier chapitre on propose une illustration numérique sur un exemple, à partir d'un échantillon aléatoire simple d'une loi normale multivariée. Par la suite nous étudions les propriétés asymptotique de l'estimateur obtenu de chaque méthode et une étude comparative entre les performances des estimateurs du maximum de vraisemblance et ceux de Yule Walker sera mise en oeuvre.

Chapitre 1

Les processus stationnaires et les matrices de Toeplitz

Sommaire

1.1	Introduction	3
1.2	Notions de base en séries chronologiques	4
1.3	Présentation des modèles stationnaires autorégressifs	7
1.4	Les matrices de Toeplitz	11
1.5	Conclusion	16

1.1 Introduction

Une série chronologique est un processus stochastique $(X_t)_{t \in \mathbb{Z}}$, où t représente le temps. Dans ce chapitre, plusieurs concepts importants liés à l'analyse de séries chronologiques seront abordés. Parmi ceux-ci, on retrouve les notions d'autocorrélation, de stationnarité et de bruit blanc. On énoncera également l'étude d'une famille de processus linéaires stationnaires. Cette classe des processus autorégressifs jouera le rôle prépondérant dans notre modélisation, nous nous intéressons aux propriétés et la forme théorique des différents moments d'un processus AR(p) et aussi on donne la définition de la matrice de Toeplitz. On présentera également la méthode, proposée par Xiao-GuangLv & Ting-Zhu Huang (voir réf[19]), de l'inversion d'une matrice de Toeplitz.

1.2 Notions de base en séries chronologiques

1.2.1 La fonction d'autocovariance (voir réf[13])

Définition 1 La fonction d'autocovariance mesure la covariance entre deux valeurs X_{t+h} et X_t , elle fournit des informations sur l'évaluation de la série et sur les liaisons temporelles pouvant exister entre les différentes composantes de la série. La fonction d'autocovariance $\gamma(h)$ d'un processus stationnaire X_t est définie comme suit :

$$\gamma(h) = \text{cov}(X_{t+h}, X_t) = E[(X_{t+h} - E(X_{t+h}))(X_t - E(X_t))] \quad (1.1)$$

Propriétés :

- $\gamma(0) = \text{cov}(X_t, X_t) = E[(X_t - E(X_t))^2] = \text{Var}(X_t) \geq 0$.
- $|\gamma(h)| \leq \gamma(0)$ (l'inégalité Cauchy-Schwartz).
- $\gamma(h) = \gamma(-h)$ (symétrique, fonction paire).

1.2.2 La fonction d'autocorrélation (voir réf[8])

1.2.2.1 La fonction d'autocorrélation simple

Définition 2 On étudie la "mémoire" d'un processus en calculant son autocorrélation de retard h noté ρ_h

$$\rho_h = \text{corr}(X_{t+h}, X_t) = \frac{\text{cov}(X_{t+h}, X_t)}{\sigma_{X_t} \sigma_{X_{t+h}}} \quad (1.2)$$

qui mesure le lien entre les valeurs du processus 'à deux dates distantes de h . Pour un processus stationnaire, ρ_h prend une forme plus simple :

$$\rho_h = \frac{\gamma(h)}{\sqrt{\gamma(0)}\sqrt{\gamma(0)}} = \frac{\gamma(h)}{\gamma(0)} \quad (1.3)$$

Remarque 3 On peut tracer la courbe $\rho(h) = f(h)$ qui est appelée le corrélogramme.

1.2.2.2 La fonction d'autocorrélation partielle

Définition 4 De même, on définit l'autocorrélation partielle de retard h comme la corrélation entre $(X_t - X_t^*)$ et $(X_{t-h} - X_{t-h}^*)$ où X_t^* désigne la régression de X_t sur les $(h-1)$ valeurs $\{X_{t-1}, X_{t-2}, \dots, X_{t-h+1}\}$:

$$\tau_h = \text{corr}(X_t - X_t^*, X_{t-h} - X_{t-h}^*) = \frac{\text{cov}(X_t - X_t^*, X_{t-h} - X_{t-h}^*)}{\sigma_{(X_t - X_t^*)}\sigma_{(X_{t-h} - X_{t-h}^*)}} \quad (1.4)$$

avec $X_t^* = \sum_{k=1}^{h-1} \alpha_k X_{t-k}$, $X_{t-h}^* = \sum_{k=1}^{h-1} \beta_k X_{t-h-k}$ où les α_k et les β_k sont les coefficients des régressions.

Remarque 5 La quantité (1.4) rend compte de l'intensité de la liaison entre X_t et X_{t-h} en supprimant les liaisons induites par des variables intermédiaires $\{X_{t-1}, X_{t-2}, \dots, X_{t-h+1}\}$. On peut ainsi remarquer que pour tout processus, $\rho_1 = \tau_1$ puisque qu'il n'y a aucune variable intermédiaire entre X_t et X_{t-h} . Comme en régression multiple, l'estimation des τ_h nous permet de mesurer le retard qu'il faut remonter pour trouver une information originale sur X_t .

Remarque 6 En pratique, les ρ_h sont évidemment inconnus mais sur une série $\{x_t, 1 \leq t \leq n\}$ on peut les estimer avec la formule naturelle suivante :

$$\hat{\rho}_h = \frac{\sum_{t=h+1}^n (x_t - \bar{x}) \times (x_{t-h} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2} \quad (1.5)$$

où \bar{x} représente la moyenne des valeurs observées.

1.2.3 La stationnarité (voir réf[18])

La stationnarité est une caractéristique d'une série chronologique qui implique que le comportement de la série ne dépend pas du temps. En particulier, on dit qu'une série $(X_t)_{t \in Z}$ est stable si elle ne comporte pas de tendance saisonnière, ni de tendance à la hausse ou à la baisse. Plus formellement, on distingue deux types de stationnarité, à savoir forte et faible.

1.2.3.1 La stationnarité stricte

Définition 7 On dit qu'un processus aléatoire temporel $(X_t, t \in Z)$ est fortement stationnaire si, pour toute suite d'instantants $\{t_1, t_2, \dots, t_n\}$, il existe un

entier k quelconque tel que la f.d.p. jointe f de $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ est identique à la f.d.p. jointe de $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$:

$$f(X_{t_1}, \dots, X_{t_n}) = f(X_{t_1+h}, \dots, X_{t_n+h}) \quad (1.6)$$

Cette définition signifie que déplacer l'origine du temps de k périodes n'a aucune incidence sur la distribution de probabilités de la v.a. temporelle X_t . Autrement dit, le processus est à l'équilibre. Le problème posé par cette définition est qu'elle est trop stricte, i.e qu'elle est invérifiable de façon pratique. En effet, il faudrait pouvoir calculer une f.d.p jointe pour chaque instant de la série d'observations. Or, à chaque instant on ne dispose que d'une seule donnée d'observation. C'est pourquoi on restreint dans les travaux appliqués la notion de stationnarité à la stationnarité de second ordre encore qualifiée de stationnarité au sens large.

1.2.3.2 La stationnarité au sens large

Définition 8 On dit qu'un processus aléatoire temporel $(X_t, t \in Z)$ est stationnaire de second ordre s'il vérifie les propriétés suivantes :

- $\forall t \in Z, E(X_t^2) < \infty$.
- $\forall t \in Z, E(X_t) = \mu$ indépendant de t .
- $\forall (t, h) \in Z^2, cov(X_{t+h}, X_t) = E[(X_{t+h} - \mu)(X_t - \mu)] = \gamma(h)$.

Les deux premières propriétés signifient que la moyenne et la variance du processus sont indépendantes de l'instant auquel ils sont mesurés. La troisième propriété que la covariance ne dépend que du seul retard k .

1.2.4 Le processus Bruit Blanc (voir réf[20])

Le bruit blanc fait partie de la classe des processus stationnaires. Spécifiquement $(\varepsilon_t)_{t \in Z}$ est un bruit blanc si :

- $E(\varepsilon_t) = 0$.
- $E(\varepsilon_t^2) = \sigma^2 < \infty$.
- $\rho_h = corr(\varepsilon_{t+h}, \varepsilon_t) = 0$, pour tout $h \geq 1$.

Par conséquent, le comportement d'un bruit blanc au temps t n'a aucune incidence sur celui-ci au temps $t + h$. On parle de bruit blanc gaussien lorsque $\varepsilon_t \sim N(0, \sigma^2)$.

1.2.5 Opérateur de retard (voir réf[20])

L'opérateur B décale le processus d'une unité de temps vers le passé, pour un processus stochastique $(X_t)_{t \in \mathbb{Z}}$, l'opérateur de retard, noté B, est défini par :

$$BX_t = X_{t-1} \quad (1.7)$$

1.3 Présentation des modèles stationnaires autorégressifs

Dans l'étude d'une série chronologique, il est naturel de penser que la valeur de la série à la date t peut dépendre des valeurs prises aux dates précédentes $X_t = (X_{t-1}, X_{t-2}, \dots)$. Il n'est généralement pas nécessaire de prendre en compte tout le passé de la série et on peut le plus souvent se limiter à p valeurs :

$$X_t = f(X_{t-1}, X_{t-2}, \dots) + \varepsilon_t \quad (1.8)$$

Où ε_t est un bruit blanc.

1.3.1 Processus autorégressif AR(1) (voir réf[13])

Le processus autorégressif d'ordre 1, noté $AR(1)$, est stationnaire et vérifie l'équation :

$$X_t = c + \Phi X_{t-1} + \varepsilon_t \quad \text{avec} \quad \varepsilon_t \sim N(0, \sigma^2) \quad (1.9)$$

1.3.1.1 Stationnarité ,espérance, variance

La stationnarité implique que :

- si $|\Phi| \geq 1$ processus X_t n'est pas stationnaire
- si $|\Phi| < 1$ le processus X_t est stationnaire.

On sait que le processus est stationnaire donc

$$E(X_t) = E(X_{t-1}) = \mu \quad (a)$$

Alors

$$E(X_t) = c + \Phi E(X_{t-1}) + E(\varepsilon_t) \quad (b)$$

Remplaçons (a) dans (b) on obtient $\mu = c + \Phi\mu$

Donc l'espérance de AR(1) est :

$$E(X_t) = \mu = \frac{c}{1 - \Phi} \quad (1.10)$$

On remplace la valeur de c dans (1.9) on trouve :

$$X_t = \mu(1 - \Phi) + \Phi X_{t-1} + \varepsilon_t$$

$$X_t - \mu = \Phi(X_{t-1} - \mu) + \varepsilon_t$$

$$E(X_t - \mu)^2 = \Phi^2 E[(X_{t-1} - \mu)^2] + E(\varepsilon_t^2) + 2\Phi E[(X_{t-1} - \mu)\varepsilon_t]$$

Or les ε_t ne sont pas corrélés entre eux donc :

$$E(X_t - \mu)^2 = E[(X_{t-1} - \mu)^2] = \gamma_0$$

$$\gamma_0 = \Phi^2 \gamma_0 + 0 + \sigma^2$$

$$\gamma_0 = \frac{\sigma^2}{(1 - \Phi^2)}$$

Alors la variance de AR(1) est :

$$Var(X_t) = \gamma_0 = \frac{\sigma^2}{(1 - \Phi^2)} \quad (1.11)$$

1.3.1.2 Autocorrélation

On a

$$\gamma_h = E[(X_t - \mu)(X_{t-h} - \mu)] = \Phi E[(X_{t-1} - \mu)(X_{t-h} - \mu)] + E[(X_{t-h} - \mu)\varepsilon_t]$$

On sait que $(X_{t-h} - \mu)$ est une fonction linéaire avec $\varepsilon_{t-h}, \varepsilon_{t-h-1}, \varepsilon_{t-h-2}, \dots$

Donc pour $h > 0$:

$E[(X_{t-1} - \mu)(X_{(t-1)-(h-1)} - \mu)] = \gamma_{h-1}$, avec $h - 1$ représentant la période.

$\gamma_h = \Phi \gamma_{h-1}$ au premier ordre et $\gamma_h = \Phi^h \gamma_0$, le coefficient d'autocorrélation vaut donc :

$$\rho_h = \frac{\gamma_h}{\gamma_0} = \Phi^h \quad (1.12)$$

Ce dernier résultat implique que, lorsque $|\Phi| < 1$, les corrélogrammes auront des allures décroissantes amorties.

1.3.2 Processus autorégressif AR(2) (Voir réf[13])

Un processus AR(2) est un processus stationnaire qui vérifie une équation de la forme :

$$X_t = c + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \varepsilon_t \quad (1.13)$$

1.3.2.1 Stationnarité ,espérance, variance

La stationnarité implique que

$$E[X_t] = c + \Phi_1 E(X_{t-1}) + \Phi_2 E(X_{t-2}) + E(\varepsilon_t)$$

$$\mu = c + \Phi_1 \mu + \Phi_2 \mu + 0 \quad \text{donc l'espérance de } AR(2) \text{ est :}$$

$$E[X_t] = \mu = \frac{c}{1 - \Phi_1 - \Phi_2} \quad (1.14)$$

Trouvons γ_0 :

$$\gamma_0 = E[(X_t - \mu)^2]$$

$$\gamma_0 = \Phi_1 E[(X_{t-1} - \mu)(X_t - \mu)] + \Phi_2 E[(X_{t-2} - \mu)(X_t - \mu)] + E[(\varepsilon_t)(X_t - \mu)].$$

Or

$$E[(\varepsilon_t)(X_t - \mu)] = E[(\varepsilon_t)(\Phi_1(X_{t-1} - \mu) + \Phi_2 E(X_{t-2} - \mu) + \varepsilon_t)] = \sigma^2$$

Alors

$$\gamma_0 = \Phi_1 \gamma_1 + \Phi_2 \gamma_2 + \sigma^2 = \rho_1 \Phi_1 \gamma_0 + \rho_2 \Phi_2 \gamma_2$$

$$\gamma_0 = \left[\frac{\Phi_1^2}{1 - \Phi_2} + \frac{\Phi_2 \Phi_1^2}{1 - \Phi_2} + \Phi_2^2 \right] \gamma_0 + \sigma^2$$

Donc la variance de $AR(2)$ est :

$$\gamma_0 = \frac{(1 - \Phi_2) \sigma^2}{(1 + \Phi_2) [(1 - \Phi_2)^2 - \Phi_1^2]} \quad (1.15)$$

1.3.2.2 Autocorrélation

$$X_t = \mu(1 - \Phi_1 - \Phi_2) + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \varepsilon_t$$

$$X_t - \mu = \Phi_1 (X_{t-1} - \mu) + \Phi_2 (X_{t-2} - \mu)$$

On multiplie par $(X_{t-1} - \mu)$ et on prend l'espérance on obtient :

$$\gamma_h = \Phi_1 \gamma_{h-1} + \Phi_2 \gamma_{h-2} \quad \text{pour } h = 1, 2, \dots$$

On divise par γ_0 on obtient l'autocorrélation de $AR(2)$:

$$\rho_h = \Phi_1 \rho_{h-1} + \Phi_2 \rho_{h-2} \quad \text{pour } h = 1, 2, \dots \quad (1.16)$$

$$\begin{aligned} - \text{ Pour } h = 1 \quad \rho_1 &= \Phi_1 + \Phi_2 \rho_1 = \frac{\Phi_1}{1 - \Phi_2} \\ - \text{ Pour } h = 2 \quad \rho_2 &= \Phi_1 \rho_1 + \Phi_2 = \frac{\Phi_1^2 + \Phi_2(1 - \Phi_2)}{1 - \Phi_2} \end{aligned}$$

1.3.3 Le processus autorégressif AR(p) (Voir réf[s])

De façon générale, un processus AR(p) est un processus qui dépend linéairement des p valeurs antérieures :

$$X_t = c + \Phi_1 X_{t-1} + \dots + \Phi_p X_{t-p} + \varepsilon_t \quad (1.17)$$

où ε_t est un bruit blanc.

On peut, sans restriction de généralité, supprimer le terme constant et obtenir un modèle de la forme :

$$X_t = \sum_{i=1}^p \Phi_i X_{t-i} + \varepsilon_t \Leftrightarrow \phi(L) X_t = \varepsilon_t$$

où ϕ est le polynôme de degré p dont les coefficients sont $(1, -\Phi_1, \dots, -\Phi_p)$.

1.3.3.1 Autocorrélation

On montre que les autocorrélations sont solutions des équations *Yule-Walker*

$$\rho_h = \sum_{i=1}^p \Phi_i \rho_{h-i} = 0$$

Comme pour les processus AR(1) et AR(2), on montre que ρ_h décroît exponentiellement

1.3.3.2 Autocorrélation partielle

Dans un tel processus, X_t et X_{t-p+1} sont indépendants conditionnellement aux valeurs intermédiaires $\{X_{t-1}, X_{t-2}, \dots, X_{t-p}\}$ et donc $(h > p) \Leftrightarrow \alpha(h) = 0$.

La valeur à la date t dépend des p dates précédentes et pas des autres. Cette propriété sert à l'identification des modèles et à déterminer l'ordre p d'un processus AR(p) au vu du corrélogramme partiel.

1.3.4 Causalité des modèles autorégressifs (Voir réf[18])

Définition 9 Un processus AR(p) est dit causal lorsqu'il existe une suite de nombres α_k telle que :

$$\sum_{k \in \mathbb{Z}} |\alpha_k| < \infty \quad \text{et} \quad X_t = \sum_{k=0}^{\infty} \alpha_k \varepsilon_{t-k} \quad (1.18)$$

Nous pouvons à présent retenir de notre développement une condition nécessaire et suffisante pour qu'un processus $AR(1)$ soit stationnaire et causal si $|\Phi| < 1$.

Observons également que la situation $|\Phi| = 1$ ne fournit pas de solution stationnaire ce qui s'explique de soi en vue de la forme de l'autocovariance.

La généralisation de ce résultat au cas où le modèle autorégressif est d'ordre $p > 1$ n'est pas directe. Considérons le modèle (1.17) est centré. En introduisant dans cette représentation l'opérateur de retard B , on peut réécrire (1.17) sous la forme :

$$X_t = \varepsilon_t + \sum_{k=0}^{\infty} \Phi_k (B^k X_t) \quad (1.19)$$

Ou encore : $\left(1 - \sum_{k=0}^{\infty} \Phi_k B^k\right) X_t = \varepsilon_t$

C.a.d, on forme l'opérateur $A(B) = \left(1 - \sum_{k=0}^{\infty} \Phi_k B^k\right)$ qui est appliqué à X_t On a donc :

$$A(B)X_t = \varepsilon_t \quad (1.20)$$

Le résultat suivant donne les conditions nécessaires et suffisantes de stationnarité et de causalité des modèles autorégressifs sur le comportement de l'opérateur A .

Proposition 10 *Le processus autorégressif (1.20) est causal et stationnaire si et seulement si le polynôme en $z = A(z) = 1 - \sum_{k=0}^{\infty} \Phi_k z^k$ ("polynôme générateur") est tel que : $A(z) \neq 0$ pour tout $z \in \mathbb{C}$ tel que $|z| \leq 1$. Les coefficients α_k apparaissant dans la représentation (1.18), sont déterminés par : $1 - \sum_{j=0}^{\infty} \alpha_j z^j = \frac{1}{A(z)}$, si $|z| \leq 1$.*

1.4 Les matrices de Toeplitz

Définition 11 (Voir réf[14]) *On appelle matrice de Toeplitz une matrice carrée T_n d'ordre $n \geq 1$ à coefficients réels telle que : $T_n(i,j)$ ne dépend que de $|i - j|$ et $T_n(i,j)$ est définie positive.*

En algèbre linéaire, une matrice de Toeplitz (d'après Otto Toeplitz) ou matrice à diagonales constantes est une matrice dont les coefficients sur une

diagonale descendant de gauche à droite sont les mêmes. Par exemple, la matrice suivante est une matrice de Toeplitz :

$$T = \begin{pmatrix} a_0 & a_{-1} & a_{-2} & \cdots & \cdots & a_{1-n} \\ a_1 & a_0 & a_{-1} & a_{-2} & \ddots & \vdots \\ a_2 & a_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & a_2 & \ddots & \ddots & \ddots & a_{-2} \\ \vdots & \vdots & \ddots & \ddots & \ddots & a_{-1} \\ a_{n-1} & a_{n-2} & \cdots & a_2 & a_1 & a_0 \end{pmatrix}$$

Si l'élément situé à l'intersection des ligne i et colonne j de T est noté $T_{i,j}$ alors on a : $T_{i,j} = a_{i-j}$.

Exemple 12 (Voir réf[5]) *La matrice variance-covariance d'un processus stationnaire discret est une matrice Toeplitz carrée.*

Soit un vecteur d'observation X tel que $X = (X_1, X_2, \dots, X_n)^t$. On définit alors la matrice variance-covariance par :

$$\Gamma_n = E(X.X^t)$$

où X^t représente le transposée de X , c'est-à-dire le vecteur transposé conjugué.

En détaillant la matrice variance-covariance, on obtient immédiatement :

$$\Gamma_n = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-1) \\ \gamma(-1) & \gamma(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \gamma(1) \\ \gamma(1-n) & \cdots & \gamma(-1) & \gamma(0) \end{pmatrix}$$

1.4.1 propriétés (Voir réf[11])

1. La matrice de Toeplitz (variance-covariance) d'un processus stochastique discret stationnaire est hermitienne. Une matrice est dite hermitienne si elle est égale à sa transposée, i.e : $T^t = T$.

2. La matrice de Toeplitz (variance-covariance) d'un processus stationnaire discret est toujours définie non négative (et souvent définie positive).
3. Si une matrice de Toeplitz vérifie de plus $a_i = a_{i+n}$, alors c'est une matrice circulante.

1.4.2 L'inverse de la matrice de Toeplitz (Voir réf[19])

Il est démontré que l'inverse d'une matrice de Toeplitz peut être déterminée par la solution des deux équations standards. La matrice inverse peut être représentée comme une somme de produits de matrices circulantes et supérieures des matrices de Toeplitz triangulaire.

Lemme 13 Soit $T = (a_{i-j})_{i,j=1}^n$ une matrice de Toeplitz d'ordre (n,n) . la matrice T vérifie la formule suivante :

$$K \times T - T \times K = f \times e_n^t - e_1 \times f^t \times J \quad (1.21)$$

Avec

$$K = \begin{pmatrix} 0 & \cdots & \cdots & 1 \\ 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}, \quad J = \begin{pmatrix} & & & 1 \\ & & 1 & \\ & 1 & & \\ 1 & & & \end{pmatrix}, \quad e_1 = \begin{pmatrix} 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

$$e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \quad f = \begin{pmatrix} 0 \\ a_{n-1}-a_{-1} \\ \vdots \\ a_2-a_{-n+2} \\ a_1-a_{-n+1} \end{pmatrix}$$

Théorème 14 Soit $T = (a_{i-j})_{i,j=1}^n$ est une matrice de Toeplitz d'ordre (n,n) . Si chacun des systèmes suivants : $Tx = f$, $Ty = e_1$ a une solution $x = (x_1, x_2, \dots, x_n)^t$, $y = (y_1, y_2, \dots, y_n)^t$. Alors :

1. T est inversible .
2. $T^{-1} = T_1 U_1 + T_2 U_2$.

Où :

$$T_1 = \begin{pmatrix} y_1 & y_n & \cdots & y_2 \\ y_2 & y_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & y_n \\ y_n & \cdots & y_2 & y_1 \end{pmatrix}, \quad U_1 = \begin{pmatrix} 1 & -x_n & \cdots & -x_2 \\ & 1 & \ddots & \vdots \\ & & \ddots & -x_n \\ & & & 1 \end{pmatrix}$$

$$T_2 = \begin{pmatrix} x_1 & x_n & \cdots & x_2 \\ x_2 & x_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_n \\ x_n & \cdots & x_2 & x_1 \end{pmatrix}, \quad U_2 = \begin{pmatrix} 0 & y_n & \cdots & y_2 \\ & 0 & \ddots & \vdots \\ & & \ddots & y_n \\ & & & 0 \end{pmatrix}$$

Preuve

D'après le lemme (13) et $Tx = f$, $Ty = e_1$ nous avons :

$$\begin{aligned} K \times T &= T \times K + f \times e_n^t - e_1 \times f^t \times J \\ &= T \times [K + xe_n^t - yf^t \times J] \end{aligned}$$

Puis

$$\begin{aligned} K^i \times T &= K^{i-1}T \times [K + xe_n^t - yf^t \times J] \\ &= \dots \\ &= T \times [K + xe_n^t - yf^t \times J]^i \end{aligned}$$

Par conséquent :

$$K^i e_1 = K^i \times T \times y = T \times [K + xe_n^t - yf^t \times J]^i \times y$$

On pose

$$t_i = [K + xe_n^t - yf^t \times J]^{i-1} \times y \quad \text{et} \quad \widehat{T} = (t_1, t_2, \dots, t_n)$$

Où

$$T \times t_i = T \times [K + xe_n^t - yf^t \times J]^{i-1} \times y = K^{i-1} \times e_1 = e_i.$$

$$T \times \widehat{T} = T (t_1, t_2, \dots, t_n) = (e_1, e_2, \dots, e_n) = I_n$$

Ainsi, la matrice T est inversible, et l'inverse de T est la matrice

$$\widehat{T} = T^{-1}.$$

Pour (2) : Tout d'abord, il est facile de voir que

$$t_1 = y, \quad t_i = [K + xe_n^t - yf^t \times J] \times t_{i-1} \quad \text{tel que} \quad i \in \{1, 2, \dots, n\}$$

$$t_i = T^{-1}e_i, \quad J \times e_i = e_{n-i+1}$$

$$JTJ = T^t, \quad JJ = I, \quad J^t$$

Alors, pour $i > 1$

$$\begin{aligned} t_i &= K \times t_{i-1} + xe_n^t \times t_{i-1} - yf^t \times J \times t_{i-1} \\ &= K \times t_{i-1} + xe_n^t \times T^{-1}e_{i-1} - yf^t \times J \times T^{-1}e_{i-1} \\ &= K \times t_{i-1} + xe_n^t \times JJT^{-1}Je_{n-i+2} - yf^t \times J \times T^{-t}Je_{i-1} \\ &= K \times t_{i-1} + xe_1^t T^{-t}e_{n-i+2} - yf^t \times J \times T^{-t}e_{n-i+2} \\ &= K \times t_{i-1} + xy^t e_{n-i+2} - yx^t e_{n-i+2} \\ &= K \times t_{i-1} + y_{n-i+2}x - x_{n-i+2}y. \end{aligned}$$

Nous avons donc :

$$t_1 = y, \quad t_2 = Ky + y_n x - x_n y, \dots,$$

$$t_n = K^{n-1}y + K^{n-2}xy_n - K^{n-2}x_n y + \dots + xy_2 - yx_2$$

$$T^{-1} = (t_1, t_2, \dots, t_n)$$

$$T^{-1} = \begin{pmatrix} y, & Ky, & \dots, & K^{n-1}y \end{pmatrix} \times \begin{pmatrix} 1 & -x_n & \dots & -x_2 \\ & 1 & \ddots & \vdots \\ & & \ddots & -x_n \\ & & & 1 \end{pmatrix} +$$

$$\begin{pmatrix} x, & Kx, & \dots, & K^{n-1}x \end{pmatrix} \times \begin{pmatrix} 0 & y_n & \dots & y_2 \\ & 0 & \ddots & \vdots \\ & & \ddots & y_n \\ & & & 0 \end{pmatrix}$$

$$T^{-1} = \begin{pmatrix} y_1 & y_n & \dots & y_2 \\ y_2 & y_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & y_n \\ y_n & \dots & y_2 & y_1 \end{pmatrix} \times \begin{pmatrix} 1 & -x_n & \dots & -x_2 \\ & 1 & \ddots & \vdots \\ & & \ddots & -x_n \\ & & & 1 \end{pmatrix} +$$

$$\begin{pmatrix} x_1 & x_n & \dots & x_2 \\ x_2 & x_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_n \\ x_n & \dots & x_2 & x_1 \end{pmatrix} \times \begin{pmatrix} 0 & y_n & \dots & y_2 \\ & 0 & \ddots & \vdots \\ & & \ddots & y_n \\ & & & 0 \end{pmatrix}$$

Remarque 15 soit la matrice de Toeplitz $T = (a_{i-j})_{i,j=1}^n$ on dit que T est circulante si seulement si les éléments de la matrice $T = (a_{i-j})_{i,j=1}^n$ sont de la forme suivante : $a_i = a_{i-n}$ pour tout $i = 1, \dots, n - 1$. Il est facile de voir que $f = 0$ ainsi $x = T^{-1}f = 0$. De (2) du théorème, nous obtenons.

$$T^{-1} = \begin{pmatrix} y_1 & y_n & \cdots & y_2 \\ y_2 & y_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & y_n \\ y_n & \cdots & y_2 & y_1 \end{pmatrix}$$

Nous concluons que les inverses des matrices de Toeplitz circulantes sont des matrices de Toeplitz circulantes.

1.5 Conclusion

Après avoir défini des notions de base en séries temporelles et présenté une classe particulière de modèles, appelée classe des processus autorégressifs stationnaires, nous permettons de caractériser la structure de corrélation d'une série. Aussi nous terminons ce chapitre en donnant une méthode de Xiao-GuangLv, Ting-zhu Huang, qui propose l'inverse d'une matrice de Toeplitz et nous abordons dans le chapitre suivant l'estimations des paramètres d'un modèle autorégressif à partir d'une suite de n observations.

Chapitre 2

Estimation des paramètres d'un modèle autorégressif

Sommaire

2.1	Introduction.	17
2.2	La méthode de Yule-Walker.	18
2.3	La méthode du Maximum de Vraisemblance	21
2.4	Application la méthode du Maximum de Vraisemblance au modèle autorégressif AR(1)	24
2.7	Conclusion	54

2.1 Introduction

L'estimation des paramètres d'un modèle AR(p) lorsque l'ordre p est supposé connu peut se réaliser par différentes méthodes dans le domaine temporel. Dans ce chapitre nous abordons deux méthodes, celle de Yulle-Walker et de Maximum de Vraisemblance Exacte.

Nous nous intéressons principalement à la méthode du Maximum de Vraisemblance basée sur deux étapes, il s'agit d'abord de construire la fonction de log-vraisemblance, ensuite de trouver l'estimateur du maximum de vraisemblance qui maximise cette fonction de log-vraisemblance, mais la difficulté majeure dans l'évaluation de log-vraisemblance, réside dans la nécessité d'inverser la matrice de variance-covariance (Γ) de taille ($n \times n$), et de calculer son déterminant, qui n'est pas aussi simple, on est alors conduit à utiliser deux méthodes dans le cas où le modèle autorégressif d'ordre 1 :

1. La méthode utilisant la décomposition de Cholesky : *Newbold* a proposé une autre expression de cette vraisemblance plus facile à calculer et qui nous a permis de trouver la forme analytique de l'estimateur du maximum de vraisemblance.
2. La méthode utilisant l'inverse de la matrice de Toeplitz basée sur le théorème de (*Xiao – Guang Lv, Ting – zhu Huang*). Une fois l'inverse calculée, ainsi son déterminant, on l'injecte dans l'expression de la fonction de log vraisemblance afin de calculer les estimateurs des paramètres d'un modèle AR(1).

2.2 La méthode de Yule-Walker

Les équations de Yule-Walker établissent une correspondance directe entre les paramètres du modèle et ses autocovariances. Elles sont utiles pour déterminer la fonction d'autocorrélation ou estimer les paramètres (Voir réf[2],[4]). Considérons tout d'abord un processus autorégressif d'ordre p causal et de moyenne nulle :

$$X_t = \sum_{k=1}^p \Phi_k X_{t-k} + \varepsilon_t, \quad \varepsilon_t \sim N(0; \sigma^2) \quad (2.1)$$

En multipliant chaque membre de cette équation par X_{t-h} , ($h = 0, \dots, p$) et en prenant l'espérance de chaque côté, on obtient, pour le premier membre, la fonction d'autocovariance :

$$E(X_t \cdot X_{t-h}) = \gamma(h) \quad (2.2)$$

et, pour le second membre :

$$\sum_{k=1}^p \Phi_k E(X_{t-k} \cdot X_{t-h}) + E(\varepsilon_t \cdot X_{t-h}) = \begin{cases} \sum_{k=1}^p \Phi_k \gamma(k-h) & \text{si } h = 1, \dots, p \\ \sum_{k=1}^p \Phi_k \gamma(k) + \sigma^2 & \text{si } h = 0 \end{cases} \quad (2.3)$$

car, par l'hypothèse de causalité du processus, on montre que $E(\varepsilon_t \cdot X_{t-h})$ s'annule sauf pour $h = 0$. En égalant (2.2) et (2.3), on a pour $h > 0$, la relation (2.4)

$$\gamma(h) = \begin{cases} \sum_{k=1}^p \gamma(k-h) \Phi_k & \text{si } h = 1, \dots, p \\ \sum_{k=1}^p \gamma(k) \Phi_k + \sigma^2 & \text{si } h = 0 \end{cases} \quad (2.4)$$

Ce système peut s'écrire sous forme matricielle : en notant γ_p le vecteur colonne de longueur p défini par $\gamma_p = (\gamma(1), \gamma(2), \dots, \gamma(p))^t$, en notant Φ le vecteur colonne $(\Phi_1, \Phi_2, \dots, \Phi_p)^t$ et en formant la matrice $(p \times p)$, dont l'élément (i, j) vaut $[\gamma(i-j)]_{i,j=0}^p$ c'est-à-dire :

$$\Gamma_{(p,p)} = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(p-1) \\ \gamma(1) & \gamma(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \gamma(1) \\ \gamma(p-1) & \cdots & \gamma(1) & \gamma(0) \end{pmatrix} \quad (2.5)$$

alors le système (2.4) peut se réécrire matriciellement comme :

$$\begin{cases} \Gamma_p \cdot \Phi = \gamma_p \\ \Phi^t \cdot \gamma_p + \sigma^2 = \gamma(0) \end{cases} \quad (2.6)$$

Ces équations sont appelées équations de *Yule-Walker*. Elles relient les paramètres Φ du processus autorégressif à la fonction d'autocovariance théorique qui définit γ_p et Γ_p . Si on veut estimer les paramètres $\Phi_1, \Phi_2, \dots, \Phi_p$ du modèle, on peut se servir des équations de *Yule-Walker* en utilisant la fonction d'autocovariance empirique $\hat{\gamma}(h)$ comme un estimateur de $\gamma(h)$.

On construit alors le vecteur colonne $\hat{\gamma}_p = (\hat{\gamma}(1), \hat{\gamma}(2), \dots, \hat{\gamma}(p))^t$ et la matrice $\hat{\Gamma}_p$ construite en substituant γ_p par $\hat{\gamma}_p$ dans (2.6). Les relations trouvées ci-dessus entre Φ, γ_p et Γ_p permettent donc de trouver un estimateur $\hat{\Phi}$ de Φ et un estimateur $\hat{\sigma}^2$ de σ^2 à travers le système de Yule-Walker suivant, à résoudre en $\hat{\Phi}$ et $\hat{\sigma}^2$:

$$\begin{cases} \hat{\Gamma}_p \cdot \hat{\Phi} = \hat{\gamma}_p \\ \hat{\Phi}^t \cdot \hat{\gamma}_p + \hat{\sigma}^2 = \hat{\gamma}(0) \end{cases} \quad (2.7)$$

Pour obtenir un estimateur de Φ il faut donc résoudre la première équation de système (2.7)

a. Soit en inversant la matrice

$$\widehat{\Phi} = \left(\widehat{\Gamma}_p \right)^{-1} \cdot \widehat{\gamma}_p \quad (2.8)$$

Puis en calculant :

$$\widehat{\sigma}^2 = \widehat{\gamma}(0) - \widehat{\Phi}^t \cdot \widehat{\gamma}_p \quad (2.9)$$

b. Ou bien nous signalons déjà ici qu'il existe une méthode pour résoudre ce système d'équations de *Yule-Walker* sans passer par l'inversion de la matrice $\widehat{\Gamma}_p$, en utilisant l'algorithme itératif de *Durbin-Levinson*. Il faut s'assurer que $\widehat{\Gamma}_p$ est une matrice inversible. On montre que c'est toujours le cas si $\widehat{\gamma}(0)$ est strictement positif.

Proposition 16 *Si X_t est un processus stationnaire et causal de la forme (2.1) avec $\varepsilon_t \sim N(0; \sigma^2)$. et si $\widehat{\Phi}$ est l'estimateur de Yule-Walker de Φ , obtenu à partir de n observations (X_1, X_2, \dots, X_n) du processus autorégressif, alors, lorsque $n \rightarrow \infty$: $\sqrt{n} \left(\widehat{\Phi} - \Phi \right) \rightarrow N(0, \sigma^2 \Gamma_p^{-1})$ où Γ_p est la matrice $(p \times p)$ de covariance donnée par (2.5). De plus, pour $n \rightarrow \infty$: $\widehat{\sigma}^2 \rightarrow \sigma^2$*

Exemple 17 *Soit le processus AR(2) de la forme :*

$$\forall t \in Z, X_t - \Phi_1 X_{t-1} - \Phi_2 X_{t-2} = \varepsilon_t. \quad (2.10)$$

On le suppose de plus, centré et causal. En multipliant l'équation (2.10) par X_{t-1} et X_{t-2} puis en prenant l'espérance de ces équations et enfin en divisant par $\gamma(0)$

On obtient les équations de Yule-Walker

$$\begin{pmatrix} \widehat{\rho}(1) \\ \widehat{\rho}(2) \end{pmatrix} = \begin{pmatrix} 1 & \widehat{\rho}(1) \\ \widehat{\rho}(1) & 1 \end{pmatrix} \times \begin{pmatrix} \widehat{\Phi}_1 \\ \widehat{\Phi}_2 \end{pmatrix}$$

On obtient donc la solution suivante

$$\begin{pmatrix} \widehat{\Phi}_1 \\ \widehat{\Phi}_2 \end{pmatrix} = \frac{1}{1 - \widehat{\rho}(1)^2} \begin{pmatrix} \widehat{\rho}(1)(1 - \widehat{\rho}(2)) \\ \widehat{\rho}(2) - \widehat{\rho}(1)^2 \end{pmatrix}$$

Quant à la variance, partant de l'équation (2.10), on obtient :

$$X_t^2 - \Phi_1 X_{t-1} X_t - \Phi_2 X_{t-2} X_t = \varepsilon_t X_t$$

et par passage à l'espérance

$$\gamma(0) - \Phi_1\gamma(1) - \Phi_2\gamma(2) = E(\varepsilon_t X_t)$$

$$\begin{aligned} \text{Mais } E(\varepsilon_t X_t) &= E(\varepsilon_t(\varepsilon_t + \Phi_1 X_{t-1} + \Phi_2 X_{t-2})) \\ &= \sigma^2 + \Phi_1 E(\varepsilon_t X_{t-1}) + \Phi_2 E(\varepsilon_t X_{t-2}) \\ &= \sigma^2 \end{aligned}$$

D'où $\hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\Phi}_1 \hat{\gamma}(1) - \hat{\Phi}_2 \hat{\gamma}(2)$. Comme on a des estimateurs convergents de la fonction d'autocovariance et donc aussi de la fonction d'autocorrélation, on pourra en déduire des estimateurs convergents de Φ_1, Φ_2 et σ^2 .

Sur cet exemple, on voit bien comment dériver les équations de Yule-Walker pour un AR(2).

2.3 La méthode du Maximum de Vraisemblance

L'estimation du Maximum de Vraisemblance est une méthode statistique courante utilisée pour inférer les paramètres de la distribution de probabilité d'un échantillon donné. Cette méthode a été développée par le statisticien *R. A. Fisher* entre 1912 et 1922, (voir réf[1]).

2.3.1 Le principe

Soit une famille paramétrée de distributions de probabilités D_θ dont les éléments sont associés soit à une densité de probabilité connue (distribution continue), soit à une fonction de masse connue (distribution discrète), notée f_θ . On tire un échantillon de n valeurs (x_1, x_2, \dots, x_n) de la distribution, et l'on calcule la densité de probabilité associée aux données observées $f_\theta(x_1, x_2, \dots, x_n, \theta)$. Ceci étant une fonction de θ avec (x_1, x_2, \dots, x_n) fixés, c'est une vraisemblance.

$$F(\theta) = f_\theta(x_1, x_2, \dots, x_n, \theta) \tag{2.11}$$

Lorsque θ n'est pas observable, la méthode du Maximum de Vraisemblance utilise les valeurs de θ qui maximisent $F(\theta)$ estimateur de θ : c'est l'estimateur du maximum de vraisemblance de θ est noté $\hat{\theta}$.

L'estimateur du maximum de vraisemblance peut exister et être unique, ne pas être unique, ou ne pas exister.

Définition 18 Soit X une variable aléatoire réelle, de loi ou bien discrète ou bien continue, dont on veut estimer un paramètre θ . Alors on définit une fonction f telle que :

$$f(x, \theta) = \begin{cases} P_\theta(Y = y) & \text{si } X \text{ est une v.a. discrète} \\ f_\theta(x) & \text{si } X \text{ est une v.a. continue} \end{cases} \quad (2.12)$$

$f_\theta(x)$ représente la densité de X (où θ apparaît) et $P_\theta(Y = y)$ représente une probabilité discrète (où θ apparaît).

On appelle vraisemblance de θ au vu des observations (x_1, x_2, \dots, x_n) d'un échantillon indépendamment et identiquement distribué selon la loi D_θ , le quantité

$$F(x_1, x_2, \dots, x_n, \vartheta) = f(x_1, \theta) \times f(x_2, \theta) \times \dots \times f(x_n, \theta) = \prod_{i=1}^n f(x_i, \theta) \quad (2.13)$$

Définition 19 On appelle estimateur du maximum de vraisemblance, une valeur $\hat{\theta}$ solution du problème de maximisation

$$\hat{\theta} = \text{Arg max } F(x_1, x_2, \dots, x_n, \vartheta) \quad (2.14)$$

On cherche à trouver le maximum de cette vraisemblance pour que les probabilités des réalisations observées soient maximums. Ceci est un problème d'optimisation. On utilise généralement le fait que si F est dérivable (ce qui n'est pas toujours le cas) et si F admet un maximum global en une valeur $\theta = \hat{\theta}$, alors la dérivée première s'annule en $\theta = \hat{\theta}$ et que la dérivée seconde est négative. Réciproquement, si la dérivée première s'annule en $\theta = \hat{\theta}$ et que la dérivée seconde est négative en $\theta = \hat{\theta}$, alors $\theta = \hat{\theta}$ est un maximum local (et non global) de $F(x_1, x_2, \dots, x_n, \theta)$. Il est alors nécessaire de vérifier qu'il s'agit bien d'un maximum global ou bien local.

La vraisemblance étant positive et le logarithme népérien une fonction croissante, il est équivalent et souvent plus simple de maximiser le logarithme népérien de la vraisemblance (le produit se transforme en somme, ce qui est plus simple à dériver). Ainsi en pratique :

a. Cas d'un paramètre unidimensionnel

– La condition nécessaire

$$\frac{\delta F(x_1, x_2, \dots, x_n, \theta)}{\delta \theta} = 0 \quad \text{ou} \quad \frac{\delta \ln F(x_1, x_2, \dots, x_n, \theta)}{\delta \theta} = 0 \quad (2.15)$$

permet de trouver la valeur $\theta = \hat{\theta}$.

– $\theta = \hat{\theta}$ est un maximum local si la condition suffisante est remplie au point critique $\theta = \hat{\theta}$:

$$\frac{\delta^2 F(x_1, x_2, \dots, x_n, \vartheta)}{\delta^2 \theta} < 0 \quad \text{ou} \quad \frac{\delta^2 \ln F(x_1, x_2, \dots, x_n, \theta)}{\delta^2 \theta} < 0 \quad (2.16)$$

b. Cas d'un paramètre multidimensionnel

Dans le cas d'un paramètre multidimensionnel, l'estimateur du maximum de vraisemblance de θ est un vecteur de dimension n (solution des équations de vraisemblance).

– La condition nécessaire

$$\frac{\delta F(x_1, x_2, \dots, x_n, \vartheta)}{\delta \theta_i} = 0 \quad \forall i = 1, \dots, n \quad (2.17)$$

– La matrice des dérivées secondes (Hessienne), est la matrice $H(x; \theta)$ de dimension n d'élément :

$$H_{ij}(x; \theta) = \frac{\delta^2 F(x_1, x_2, \dots, x_n, \theta)}{\delta \theta_j \delta \theta_i} \quad (2.18)$$

est définie négative en $\hat{\theta}$.

Corollaire 20 Une matrice H est définie positive si chacun des mineurs principaux dominants de H est supérieur à 0. Une matrice H est définie négative si ses n mineurs principaux dominant alternent en signe de la manière suivante.

$$a_{11} < 0, \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0, \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} < 0, \text{ ect...}$$

2.4 Application la méthode du Maximum de Vraisemblance au modèle autorégressif AR(1)

Considérons un modèle AR(1) univarié. Nous ferons toujours l'hypothèse que le processus stochastique qui génère la série chronologique est un processus gaussien, écrit habituellement sous la forme :

$$X_t = \Phi X_{t-1} + \varepsilon_t \quad \text{où} \quad \varepsilon_t \sim N(0, \sigma^2) \quad (2.19)$$

Considérons la série chronologique $X = (X_1, X_2, \dots, X_n)^t$ générée par ce processus AR(1). Le problème statistique posé est celui de l'estimation des paramètres

$\theta = (\Phi, \sigma^2)$. Puisque le processus stochastique qui génère la série chronologique est gaussien, la loi de probabilité du vecteur X est normale multivariée, de moyenne nulle et de matrice covariance Γ de dimension (n, n) telle que :

$$\Gamma_{(n,n)} = \begin{pmatrix} E(X_1)^2 & E(X_1.X_2) & \cdots & E(X_1.X_n) \\ E(X_2.X_1) & E(X_2)^2 & \cdots & E(X_2.X_n) \\ \vdots & \vdots & \cdots & \vdots \\ E(X_n.X_1) & E(X_n.X_2) & \cdots & E(X_n)^2 \end{pmatrix} \quad (2.20)$$

Cette hypothèse nous permettra d'écrire simplement la fonction de vraisemblance du modèle qui est la probabilité d'observer $X = (X_1, X_2, \dots, X_n)^t$ sous l'hypothèse gaussienne :

$$f_X(x, \theta) = (2\pi)^{-\frac{n}{2}} (\det \Gamma)^{-\frac{1}{2}} \exp \left[-\frac{1}{2} X^t \Gamma^{-1} X \right] \quad (2.21)$$

La fonction de log-vraisemblance s'écrit :

$$L(\theta) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log(\det \Gamma) - \frac{1}{2} X^t \Gamma^{-1} X \quad (2.22)$$

L'élément général de la matrice variance-covariance écrit sous la forme

$$E(X_t.X_{t-j}) = \frac{\sigma^2 \Phi^j}{1 - \Phi^2} \quad (2.23)$$

La matrice (2.20) devienne :

$$\Gamma = \sigma^2 \times V \quad \text{Où} \quad V = \frac{1}{1 - \Phi^2} \begin{pmatrix} 1 & \Phi & \Phi^2 & \dots & \Phi^{n-1} \\ \Phi & 1 & \Phi & \dots & \Phi^{n-2} \\ \Phi^2 & \Phi & 1 & \dots & \Phi^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi^{n-1} & \Phi^{n-2} & \Phi^{n-3} & \dots & 1 \end{pmatrix}$$

D'autre part :

$$\begin{aligned} (\det \Gamma)^{-1} &= \det (\Gamma^{-1}), \text{ en effet :} \\ \Gamma^{-1} \times \Gamma &= I_d \Rightarrow \det (\Gamma^{-1} \times \Gamma) = \det (I_d) \\ &\Rightarrow \det (\Gamma^{-1}) \times \det (\Gamma) = 1 \\ &\Rightarrow \det (\Gamma^{-1}) = \frac{1}{\det (\Gamma)} = (\det \Gamma)^{-1} \end{aligned}$$

donc la vraisemblance (2.22) s'écrit sous la forme suivante :

$$L(\theta) = -\frac{n}{2} \log(2\pi) + \frac{1}{2} \log(\det \Gamma^{-1}) - \frac{1}{2} X^t \Gamma^{-1} X \quad (2.24)$$

La maximisation et même le calcul de cette vraisemblance sont relativement difficiles; en particulier à cause du calcul l'inverse de la matrice variance-covariance et de son déterminant, surtout lorsque n devient relativement grand. Pour résoudre ce problème on a proposé deux méthodes :

1. *La méthode utilisant la décomposition de Cholesky.*
2. *La méthode utilisant l'inverse de la matrice de Toeplitz.*

2.4.1 La méthode utilisant la décomposition de Cholesky

Comme il semble utopique d'espérer obtenir analytiquement les estimateurs du maximum de vraisemblance exact des paramètres nous utilisons la méthode qui a proposé Newbold (1974) qui a introduit un artifice supplémentaire qui est en fait la décomposition de Cholésky de V sous la forme $V = PP^t$, telle que P est une matrice triangulaire inférieure, à éléments positifs sur la diagonale. Cette dernière transformation facilite à nouveau l'évaluation des formules Γ^{-1} et $\det(\Gamma^{-1})$. Ceci nous permettra non seulement de déterminer la fonction log-vraisemblance de modèle (2.19), mais encore de trouver la forme analytique de ces estimateurs, (Voir réf : [15],[7],[13]).

2.4.1.1 La fonction de vraisemblance d'un modèle autorégressif

AR(1)

La matrice variance - covariance s'écrit sous la forme suivante :

$$\Gamma_{(n,n)} = \sigma^2 \times V \quad \text{et} \quad V_{(n,n)} = \frac{1}{1 - \Phi^2} \begin{pmatrix} 1 & \Phi & \Phi^2 & \dots & \Phi^{n-1} \\ \Phi & 1 & \Phi & \dots & \Phi^{n-2} \\ \Phi^2 & \Phi & 1 & \dots & \Phi^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi^{n-1} & \Phi^{n-2} & \Phi^{n-3} & \dots & 1 \end{pmatrix}$$

Par conséquent

$$\Gamma_{(n,n)} = \sigma^2 \times \frac{1}{1 - \Phi^2} \begin{pmatrix} 1 & \Phi & \Phi^2 & \dots & \Phi^{n-1} \\ \Phi & 1 & \Phi & \dots & \Phi^{n-2} \\ \Phi^2 & \Phi & 1 & \dots & \Phi^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi^{n-1} & \Phi^{n-2} & \Phi^{n-3} & \dots & 1 \end{pmatrix} \quad (2.25)$$

Alors

$$\Gamma_{(n,n)}^{-1} = \frac{(1 - \Phi^2)}{\sigma^2} \begin{pmatrix} 1 & \Phi & \Phi^2 & \dots & \Phi^{n-1} \\ \Phi & 1 & \Phi & \dots & \Phi^{n-2} \\ \Phi^2 & \Phi & 1 & \dots & \Phi^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi^{n-1} & \Phi^{n-2} & \Phi^{n-3} & \dots & 1 \end{pmatrix}^{-1} \quad (2.26)$$

On pose

$$V_{(n,n)}^{-1} = (1 - \Phi^2) \begin{pmatrix} 1 & \Phi & \Phi^2 & \dots & \Phi^{n-1} \\ \Phi & 1 & \Phi & \dots & \Phi^{n-2} \\ \Phi^2 & \Phi & 1 & \dots & \Phi^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi^{n-1} & \Phi^{n-2} & \Phi^{n-3} & \dots & 1 \end{pmatrix}^{-1} \quad (2.27)$$

On a : $V = PP^t$

$$V^{-1} = (PP^t)^{-1} = (P^t)^{-1} \times (P)^{-1}$$

On pose $(P)^{-1} = L$

Alors $V^{-1} = L^t \times L$

Donc

$$\Gamma^{-1} = \sigma^{-2} L^t \times L \quad (2.28)$$

Et d'autre part on a la forme de la matrice L , (Voir réf :[13]).

$$L_{(n,n)} = \begin{pmatrix} \sqrt{1-\Phi^2} & 0 & 0 & \cdots & 0 & 0 \\ -\Phi & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\Phi & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\Phi & 1 \end{pmatrix} \quad (2.29)$$

Il est facile de montrer que $V^{-1} \times V = L^t \times L \times V$

$$I_d = L^t \times L \times V$$

Avec :

$$L \times V = \frac{1}{1-\Phi^2} \begin{pmatrix} \sqrt{1-\Phi^2} & \Phi\sqrt{1-\Phi^2} & \Phi^2\sqrt{1-\Phi^2} & \cdots & \Phi^{n-1}\sqrt{1-\Phi^2} \\ 0 & (1-\Phi^2) & \Phi\sqrt{1-\Phi^2} & \cdots & \Phi^{n-2}\sqrt{1-\Phi^2} \\ 0 & 0 & (1-\Phi^2) & \cdots & \Phi^{n-3}\sqrt{1-\Phi^2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & (1-\Phi^2) \end{pmatrix}$$

Si l'on substitue (2.28) dans (2.24), on retrouve la fonction de log-vraisemblance d'un processus AR(1) gaussien sous la forme suivante :

$$L(\vartheta) = -\frac{n}{2} \log(2\pi) + \frac{1}{2} \log [\det(\sigma^{-2} L^t \times L)] - \frac{1}{2\sigma^2} X^t L^t \times LX \quad (2.30)$$

Pour déterminer la fonction de log vraisemblance, il ne reste plus maintenant qu'à calculer la forme quadratique $\left[-\frac{1}{2\sigma^2} X^t L^t \times LX \right]$ et le déterminant $[\det(\sigma^{-2} L^t \times L)]$

1. Calcul le déterminant :

$$\begin{aligned} \det(\sigma^{-2} L^t \times L) &= (\sigma^{-2})^n \det(L^t \times L) \\ &= \sigma^{-2n} \det(L^t) \times \det(L) \\ &= \sigma^{-2n} [\sqrt{1-\Phi^2}] \times \sqrt{1-\Phi^2} \end{aligned}$$

$$\det(\sigma^{-2}L^t \times L) = \sigma^{-2n}(1 - \Phi^2)$$

Alors :

$$\frac{1}{2} \log [\det(\sigma^{-2}L^t \times L)] = -\frac{n}{2} \log \sigma^2 + \frac{1}{2} \log(1 - \Phi^2) \quad (2.31)$$

2. Calcul la forme quadratique :

On pose :

$$\tilde{X} = L \times X$$

On obtient :

$$\tilde{X} = \begin{pmatrix} \sqrt{1-\Phi^2} & 0 & 0 & \cdots & 0 & 0 \\ -\Phi & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\Phi & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\Phi & 1 \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} (\sqrt{1-\Phi^2})x_1 \\ x_2 - \Phi x_1 \\ x_3 - \Phi x_2 \\ \vdots \\ x_n - \Phi x_{n-1} \end{pmatrix}$$

Et on a :

$$\left(\tilde{X}\right)^t = X^t L^t$$

$$\left(\tilde{X}\right)^t = (LX)^t$$

$$\left(\tilde{X}\right)^t = \left((\sqrt{1-\Phi^2})x_1, x_2 - \Phi x_1, x_3 - \Phi x_2, \dots, x_n - \Phi x_{n-1} \right)$$

Par conséquent :

$$-\frac{1}{2\sigma^2} X^t L^t \times LX = -\frac{1}{2\sigma^2} \left(\tilde{X}\right)^t \times \tilde{X}$$

$$= -\frac{1}{2\sigma^2} \left((\sqrt{1-\Phi^2})x_1, x_2 - \Phi x_1, \dots, x_n - \Phi x_{n-1} \right) \times \begin{pmatrix} (\sqrt{1-\Phi^2})x_1 \\ x_2 - \Phi x_1 \\ x_3 - \Phi x_2 \\ \vdots \\ x_n - \Phi x_{n-1} \end{pmatrix}$$

Ce qui équivaut à :

$$-\frac{1}{2\sigma^2} X^t L^t \times LX = -\frac{1}{2\sigma^2} \left[(1 - \Phi^2) x_1^2 + \sum_{i=2}^n (x_i - \Phi x_{i-1})^2 \right] \quad (2.32)$$

Si l'on substitue (2.31) et (2.32) dans (2.30), la fonction de log-vraisemblance d'un processus AR(1) gaussien est donnée par :

$$L(\vartheta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \frac{1}{2} \log(1-\Phi^2) - \frac{1}{2\sigma^2} \left[(1-\Phi^2) x_1^2 + \sum_{i=2}^n (x_i - \Phi x_{i-1})^2 \right] \quad (2.33)$$

2.4.1.2 Les estimateurs du maximum de vraisemblance des paramètres de modèle autorégressif AR(1)

Dans ce paragraphe nous recherchons quelles sont les valeurs des estimateurs de maximum de vraisemblance des paramètres de processus autorégressif AR(1) qui maximisent le log vraisemblance (2.33)?

Il va de soit que les estimateurs du maximum de vraisemblance sont ici : $\theta = (\Phi, \sigma^2)$, les deux estimateurs sont obtenus donc par résolution du système suivant :

$$\begin{cases} \frac{\partial L(\theta)}{\partial \sigma^2} = 0 \\ \text{et} \\ \frac{\partial L(\theta)}{\partial \Phi} = 0 \end{cases} \quad (2.34)$$

Par dérivation formelle de (2.33) par rapport à σ^2 on obtient :

$$\frac{\partial L(\vartheta)}{\partial \sigma^2} = 0 \iff -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \times \left[(1-\Phi^2) x_1^2 + \sum_{i=2}^n (x_i - \Phi x_{i-1})^2 \right] = 0$$

Dès lors :

$$\hat{\sigma}^2 = \frac{\left[(1-\hat{\Phi}^2) x_1^2 + \sum_{i=2}^n (x_i - \hat{\Phi} x_{i-1})^2 \right]}{n}$$

Par dérivation formelle de (2.33) par rapport à Φ on obtient :

$$\begin{aligned} \frac{\partial L(\vartheta)}{\partial \Phi} = 0 &\iff -\frac{2\Phi}{2(1-\Phi^2)} - \frac{1}{2\sigma^2} \times \left[-2\Phi x_1^2 - \sum_{i=2}^n 2x_{i-1} (x_i - \Phi x_{i-1}) \right] = 0 \\ &\iff -\frac{\Phi}{(1-\Phi^2)} + \frac{1}{2\sigma^2} \times \left[2\Phi x_1^2 + \sum_{i=2}^n 2x_{i-1} (x_i - \Phi x_{i-1}) \right] = 0 \end{aligned}$$

$$\frac{\partial L(\vartheta)}{\partial \Phi} = 0 \iff \frac{\Phi}{(1 - \Phi^2)} = \frac{1}{\sigma^2} \times \left[\Phi x_1^2 + \sum_{i=2}^n x_{i-1} (x_i - \Phi x_{i-1}) \right]$$

Alors le système (2.34) devient :

$$\begin{cases} \hat{\sigma}^2 = \frac{\left[(1 - \hat{\Phi}^2) x_1^2 + \sum_{i=2}^n (x_i - \hat{\Phi} x_{i-1})^2 \right]}{n} \\ \frac{\hat{\Phi}}{(1 - \hat{\Phi}^2)} = \frac{1}{\hat{\sigma}^2} \times \left[\hat{\Phi} x_1^2 + \sum_{i=2}^n x_{i-1} (x_i - \hat{\Phi} x_{i-1}) \right] \end{cases} \quad (2.35)$$

En remplaçant la valeur de $\hat{\sigma}^2$ dans la deuxième équation du système (2.35), on aura :

$$\begin{aligned} \frac{\hat{\Phi}}{(1 - \hat{\Phi}^2)} &= \frac{\hat{\Phi} x_1^2 + \sum_{i=2}^n x_{i-1} x_i - \hat{\Phi} \sum_{i=2}^n x_{i-1}^2}{\frac{\left[(1 - \hat{\Phi}^2) x_1^2 + \sum_{i=2}^n (x_i - \hat{\Phi} x_{i-1})^2 \right]}{n}} \\ \frac{\hat{\Phi}}{n(1 - \hat{\Phi}^2)} &= \frac{\hat{\Phi} x_1^2 + \sum_{i=2}^n x_{i-1} x_i - \hat{\Phi} \sum_{i=2}^n x_{i-1}^2}{-\hat{\Phi}^2 x_1^2 + x_1^2 + \sum_{i=2}^n x_i^2 + \hat{\Phi}^2 \sum_{i=2}^n x_{i-1}^2 - 2\hat{\Phi} \sum_{i=2}^n x_{i-1} x_i} \\ &+ \hat{\Phi}^3 \left[\sum_{i=2}^n x_{i-1}^2 - x_1^2 + n x_1^2 - n \sum_{i=2}^n x_{i-1}^2 \right] + \hat{\Phi}^2 \left[-2 \sum_{i=2}^n x_{i-1} x_i + n \sum_{i=2}^n x_{i-1} x_i \right] + \\ &\hat{\Phi} \left[x_1^2 + \sum_{i=2}^n x_i^2 - n x_1^2 + n \sum_{i=2}^n x_{i-1}^2 \right] - n \sum_{i=2}^n x_{i-1} x_i = 0 \end{aligned}$$

On obtient :

$$\begin{aligned} &\hat{\Phi}^3 (n-1) \left(x_1^2 - \sum_{i=2}^n x_{i-1}^2 \right) + \hat{\Phi}^2 (n-2) \left(\sum_{i=2}^n x_{i-1} x_i \right) + \hat{\Phi} \left((1-n) x_1^2 + \sum_{i=2}^n (x_i^2 + n x_{i-1}^2) \right) \\ &- n \sum_{i=2}^n x_{i-1} x_i = 0 \end{aligned} \quad (2.36)$$

L'équation (2.36) est un polynôme du troisième degré en fonction de $\hat{\Phi}$, pour trouver les racines de ce polynôme, on applique la méthode de Cardon, (Voir réf[17], [3]).

On pose :

$$\begin{cases} a = (n-1) \left(x_1^2 - \sum_{i=2}^n x_{i-1}^2 \right) \\ b = (n-2) \left(\sum_{i=2}^n x_{i-1} x_i \right) \\ c = (1-n) x_1^2 + \sum_{i=2}^n (x_i^2 + n x_{i-1}^2) \\ d = -n \sum_{i=2}^n x_{i-1} x_i \end{cases}$$

Donc l'équation (2.36) devienne :

$$a\widehat{\Phi}^3 + b\widehat{\Phi}^2 + c\widehat{\Phi} + d = 0 \quad (2.37)$$

En remplaçant $\widehat{\Phi}$ par $\widehat{\Phi} = z - \frac{b}{3a}$ on obtient :

$$z^3 + pz + q = 0 \quad (2.38)$$

Avec $p = -\frac{b^2}{3a^2} + \frac{c}{a}$ et $q = \frac{b}{27a} \left(\frac{2b^2}{a^2} - \frac{9c}{a} \right) + \frac{d}{a}$

Maintenant on pose $z = U + V$ et (U, V) sont des complexes, alors l'équation (2.38) devienne :

$$\begin{aligned} (2.38) &\iff (U + V)^3 + p(U + V) + q = 0 \\ (2.38) &\iff U^3 + V^3 + (3UV + p)(U + V) + q = 0 \\ &\iff \begin{cases} 3UV + p = 0 \\ U^3 + V^3 + q = 0 \end{cases} \\ &\iff \begin{cases} UV = -\frac{p}{3} \\ U^3 + V^3 = -q \end{cases} \\ (2.38) &\iff \begin{cases} U^3 V^3 = -\frac{p^3}{27} \\ U^3 + V^3 = -q \end{cases} \end{aligned}$$

Donc U^3, V^3 sont des solutions de l'équation

$$X^2 + qX - \frac{p^3}{27} = 0 \quad (2.39)$$

On calcule le discriminant de l'équation (2.39).

$$\Delta = q^2 - 4 \times 1 \times \left(-\frac{p^3}{27} \right) = q^2 - \frac{4p^3}{27}$$

On a trois cas :

$$\left\{ \begin{array}{l} \text{si } \Delta > 0, U^3 = \frac{-q + \sqrt{\Delta}}{2} \text{ et } V^3 = \frac{-q - \sqrt{\Delta}}{2} \\ \text{si } \Delta < 0, U^3 = \frac{-q + i\sqrt{-\Delta}}{2} \text{ et } V^3 = \frac{-q - i\sqrt{-\Delta}}{2} \\ \text{si } \Delta = 0, U^3 = V^3 = -\frac{q}{2} \end{array} \right. \quad (2.40)$$

Donc :

$$(2.40) \iff \left\{ \begin{array}{l} \text{si } \Delta > 0, U = \sqrt[3]{\frac{-q + \sqrt{\Delta}}{2}} \text{ et } V = \sqrt[3]{\frac{-q - \sqrt{\Delta}}{2}} \\ \text{si } \Delta < 0, U = \sqrt[3]{\frac{-q + i\sqrt{-\Delta}}{2}} \text{ et } V = \sqrt[3]{\frac{-q - i\sqrt{-\Delta}}{2}} \\ \text{si } \Delta = 0, U = V = -\sqrt[3]{\frac{q}{2}} \end{array} \right.$$

Il suffit alors d'associer les trois racines cubiques de U^3 et V^3 deux par deux de façon à obtenir trois couples (U, V) tel que $UV = -\frac{p}{3}$. C'est à dire $(U, V) \in \{(U_0, V_0), (jU_0, j^2V_0), (j^2U_0, jV_0)\}$ tel que $j = \frac{-1 + i\sqrt{3}}{2}$. Puis de reporter les trois couples de valeurs trouvées pour U et V dans l'expression $z = U + V$. Enfin, on revient au premier changement de variable $\hat{\Phi} = z - \frac{b}{3a}$ pour avoir les trois racines de l'équation du troisième degré. Donc l'estimateur du maximum de vraisemblance de paramètre Φ est :

$$\hat{\Phi} = U + V - \frac{b}{3a} \quad (2.41)$$

Remarque 21 On prend que la solution réel, car Φ est un réel, nous distinguerons alors deux cas selon que Δ est positif ou nul . Par exemple si $\Delta > 0$ on prend la couple $(U, V) = (\sqrt[3]{\frac{-q + \sqrt{\Delta}}{2}}, \sqrt[3]{\frac{-q - \sqrt{\Delta}}{2}})$, et si $\Delta = 0$ on prend la couple $(U, V) = (-\sqrt[3]{\frac{q}{2}}, -\sqrt[3]{\frac{q}{2}})$.

Enfin les formes analytique des estimateurs du maximum de vraisemblance exacte de $\theta = (\Phi, \sigma^2)$ sont :

- Si $\Delta > 0$, alors $\hat{\theta} = (\hat{\Phi}, \hat{\sigma}^2)$ s'écrit sous les formes suivantes :

$$\left\{ \begin{array}{l} \hat{\sigma}^2 = \frac{\left[(1 - \hat{\Phi}^2) x_1^2 + \sum_{i=2}^n (x_i - \hat{\Phi} x_{i-1})^2 \right]}{n} \\ \hat{\Phi} = \sqrt{\frac{\sqrt[3]{-\left(\frac{b}{27a} \left(\frac{2b^2}{a^2} - \frac{9c}{a}\right) + \frac{d}{a}\right) + \sqrt{\left(\frac{b}{27a} \left(\frac{2b^2}{a^2} - \frac{9c}{a}\right) + \frac{d}{a}\right)^2 - \frac{4}{27} \left(-\frac{b^2}{3a^2} + \frac{c}{d}\right)^3}}{2}}}{\sqrt[3]{-\left(\frac{b}{27a} \left(\frac{2b^2}{a^2} - \frac{9c}{a}\right) + \frac{d}{a}\right) - \sqrt{\left(\frac{b}{27a} \left(\frac{2b^2}{a^2} - \frac{9c}{a}\right) + \frac{d}{a}\right)^2 - \frac{4}{27} \left(-\frac{b^2}{3a^2} + \frac{c}{d}\right)^3}} - \frac{b}{3a}} \right.} \quad (2.42)$$

Tel que :

$$\left\{ \begin{array}{l} a = (n-1) \left(x_1^2 - \sum_{i=2}^n x_{i-1}^2 \right) \\ b = (n-2) \left(\sum_{i=2}^n x_{i-1} x_i \right) \\ c = (1-n) x_1^2 + \sum_{i=2}^n (x_i^2 + n x_{i-1}^2) \\ d = -n \sum_{i=2}^n x_{i-1} x_i \end{array} \right.$$

- Si $\Delta = 0$, alors $\hat{\theta} = (\hat{\Phi}, \hat{\sigma}^2)$ s'écrit sous les formes suivantes :

$$\left\{ \begin{array}{l} \hat{\sigma}^2 = \frac{\left[(1 - \hat{\Phi}^2) x_1^2 + \sum_{i=2}^n (x_i - \hat{\Phi} x_{i-1})^2 \right]}{n} \\ \hat{\Phi} = -2 \sqrt[3]{\frac{\frac{b}{27a} \left(\frac{2b^2}{a^2} - \frac{9c}{a}\right) + \frac{d}{a}}{2} - \frac{b}{3a}} \end{array} \right. \quad (2.43)$$

Tel que :

$$\begin{cases} a = (n-1) \left(x_1^2 - \sum_{i=2}^n x_{i-1}^2 \right) \\ b = (n-2) \left(\sum_{i=2}^n x_{i-1} x_i \right) \\ c = (1-n) x_1^2 + \sum_{i=2}^n (x_i^2 + n x_{i-1}^2) \\ d = -n \sum_{i=2}^n x_{i-1} x_i \end{cases}$$

2.4.2 La méthode utilisant l'inverse d'une matrice de Toeplitz :

Dans ce paragraphe nous proposons une autre méthode que l'on va appeler la méthode de Toeplitz. On s'intéressera plus particulièrement au calcul de l'inverse de la matrice variance-covariance et nous avons vu que la matrice variance-covariance possède une structure particulière : elle est symétrique et de sous et sur-diagonales identiques donc est une matrice de Toeplitz. Malheureusement l'inverse d'une matrice de Toeplitz n'est pas une matrice de Toeplitz. Cette méthode consiste à effectuer, un changement qui nous a permis d'écrire l'inverse de la matrice variance-covariance sous forme $T^{-1} = T_1 U_1 + T_2 U_2$ telles que T_1, T_2 sont des matrices circulantes et U_1, U_2 sont des matrices triangulaires supérieures.

Nous rappelons que ces matrices d'après l'article de *(Xiao-Guang Lv, Ting-Zhu Huang)*, (voir réf[19]) sont sous les formes suivantes :

$$T_1 = \begin{pmatrix} y_1 & y_n & \cdots & y_2 \\ y_2 & y_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & y_n \\ y_n & \cdots & y_2 & y_1 \end{pmatrix}, \quad U_1 = \begin{pmatrix} 1 & -x_n & \cdots & -x_2 \\ & 1 & \ddots & \vdots \\ & & \ddots & -x_n \\ & & & 1 \end{pmatrix}$$

$$T_2 = \begin{pmatrix} x_1 & x_n & \cdots & x_2 \\ x_2 & x_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_n \\ x_n & \cdots & x_2 & x_1 \end{pmatrix}, \quad U_2 = \begin{pmatrix} 0 & y_n & \cdots & y_2 \\ & 0 & \ddots & \vdots \\ & & \ddots & y_n \\ & & & 0 \end{pmatrix}$$

Telles que :

$x = (x_1, x_2, \dots, x_n)^t$ et $y = (y_1, y_2, \dots, y_n)^t$ sont des solutions des deux équations :

$$\begin{cases} Tx = f \\ Ty = e_1 \end{cases}$$

Avec :

$$T = \begin{pmatrix} a_0 & a_{-1} & a_{-2} & \cdots & \cdots & a_{1-n} \\ a_1 & a_0 & a_{-1} & a_{-2} & \ddots & \vdots \\ a_2 & a_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & a_2 & \ddots & \ddots & \ddots & a_{-2} \\ \vdots & \vdots & \ddots & \ddots & \ddots & a_{-1} \\ a_{n-1} & a_{n-2} & \cdots & a_2 & a_1 & a_0 \end{pmatrix}, e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, f = \begin{pmatrix} 0 \\ a_{n-1} - a_{-1} \\ \vdots \\ a_2 - a_{-n+2} \\ a_1 - a_{-n+1} \end{pmatrix}$$

2.4.2.1 L'inverse de la matrice variance covariance pour un processus autorégressif AR (1)

On a la matrice variance-covariance Γ d'ordre (n,n) pour un processus autorégressif AR(1)

$$\Gamma_{(n,n)} = \sigma^2 \times \frac{1}{1 - \Phi^2} \begin{pmatrix} 1 & \Phi & \Phi^2 & \dots & \Phi^{n-1} \\ \Phi & 1 & \Phi & \dots & \Phi^{n-2} \\ \Phi^2 & \Phi & 1 & \dots & \Phi^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi^{n-1} & \Phi^{n-2} & \Phi^{n-3} & \dots & 1 \end{pmatrix}$$

Alors l'inverse de la matrice variance-covariance est :

$$\Gamma_{(n,n)}^{-1} = \frac{(1 - \Phi^2)}{\sigma^2} \begin{pmatrix} 1 & \Phi & \Phi^2 & \dots & \Phi^{n-1} \\ \Phi & 1 & \Phi & \dots & \Phi^{n-2} \\ \Phi^2 & \Phi & 1 & \dots & \Phi^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi^{n-1} & \Phi^{n-2} & \Phi^{n-3} & \dots & 1 \end{pmatrix}^{-1}$$

On pose

$$T_{(n,n)} = \begin{pmatrix} 1 & \Phi & \Phi^2 & \dots & \Phi^{n-1} \\ \Phi & 1 & \Phi & \dots & \Phi^{n-2} \\ \Phi^2 & \Phi & 1 & \dots & \Phi^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi^{n-1} & \Phi^{n-2} & \Phi^{n-3} & \dots & 1 \end{pmatrix}$$

On obtient :

$$\Gamma^{-1} = \frac{(1 - \Phi^2)}{\sigma^2} T^{-1} \quad (2.44)$$

On remarque que la matrice T est une matrice de Toeplitz, pour calculer l'inverse de cette matrice on applique la méthode qui proposé par $\langle \text{Xiao-Guang Lv, Ting-Zhu Huang} \rangle$.

On a :

$$T^{-1} = T_1 U_1 + T_2 U_2$$

Et pour déterminer les matrices T_1, U_1, T_2, U_2 il suffit de résoudre les deux systèmes suivants :

$$\begin{cases} T \times y = e_1 & (A) \\ T \times x = f & (B) \end{cases} \quad (2.45)$$

Tel que :

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 \\ \Phi^{n-1} - \Phi \\ \vdots \\ \Phi^{n-i} - \Phi^i \\ \vdots \\ \Phi - \Phi^{n-1} \end{pmatrix}, \quad i = 2, \dots, n-2$$

1. Résolution du système (A)

$$(A) \iff T \times y = e$$

$$(A) \iff \begin{pmatrix} 1 & \Phi & \Phi^2 & \dots & \Phi^{n-1} \\ \Phi & 1 & \Phi & \dots & \Phi^{n-2} \\ \Phi^2 & \Phi & 1 & \dots & \Phi^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi^{n-1} & \Phi^{n-2} & \Phi^{n-3} & \dots & 1 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$(A) \iff \begin{cases} y_1 + \Phi y_2 + \Phi^2 y_3 + \dots + \Phi^{n-1} y_n = 1 & \dots & (1) \\ \Phi y_1 + y_2 + \Phi y_3 + \dots + \Phi^{n-2} y_n = 0 & \dots & (2) \\ \Phi^2 y_1 + \Phi y_2 + y_3 + \dots + \Phi^{n-3} y_n = 0 & \dots & (3) \\ \vdots & \vdots & \vdots \\ \Phi^{n-2} y_1 + \Phi^{n-3} y_2 + \dots + y_{n-1} + \Phi y_n = 0 & \dots & (n-1) \\ \Phi^{n-1} y_1 + \Phi^{n-2} y_2 + \dots + \Phi y_{n-1} + y_n = 0 & \dots & (n) \end{cases}$$

- Pour déterminer y_1 :

$$\begin{cases} y_1 + \Phi y_2 + \Phi^2 y_3 + \dots + \Phi^{n-1} y_n = 1 & (1) \\ \Phi \times (\Phi y_1 + y_2 + \Phi y_3 + \dots + \Phi^{n-2} y_n = 0) & (2) \end{cases}$$

$$(1) - \Phi \times (2) \iff y_1 - \Phi^2 y_1 = 1$$

Alors la valeurs de y_1 est :

$$y_1 = \frac{1}{1 - \Phi^2} \quad (2.46)$$

- Pour déterminer y_2 :

$$\begin{cases} \Phi y_1 + y_2 + \Phi y_3 + \dots + \Phi^{n-2} y_n = 0 & (2) \\ \Phi^2 y_1 + \Phi y_2 + y_3 + \dots + \Phi^{n-3} y_n = 0 & (3) \end{cases}$$

$$(2) - \Phi \times (3) \iff \Phi y_1 - \Phi^3 y_1 + y_2 - \Phi^2 y_2 = 0$$

On obtient:

$$y_2 = -\frac{y_1 \times (\Phi - \Phi^3)}{1 - \Phi^2} \quad (2.47)$$

Si l'on substitue (2.46) dans (2.47), on retrouve la valeur de y_2 :

$$y_2 = -\frac{\Phi}{1 - \Phi^2} \quad (2.48)$$

- Pour déterminer y_3 :

$$\begin{cases} \Phi^2 y_1 + \Phi y_2 + y_3 + \dots + \Phi^{n-3} y_n = 0 & (3) \\ \Phi^3 y_1 + \Phi^2 y_2 + \Phi y_3 + y_4 \dots + \Phi^{n-4} y_n = 0 & (4) \end{cases}$$

$$(3) - \Phi \times (4) \iff (\Phi^2 - \Phi^4) y_1 + (\Phi - \Phi^3) y_2 + (1 - \Phi) y_3 = 0$$

On obtient :

$$y_3 = -\frac{(\Phi^2 - \Phi^4) y_1 + (\Phi - \Phi^3) y_2}{(1 - \Phi)} \quad (2.49)$$

Si l'on substitue (2.46) et (2.48) dans (2.49), on retrouve la valeur de y_3 :

$$y_3 = 0 \quad (2.50)$$

• Pour déterminer y_4 :

$$\begin{cases} \Phi^3 y_1 + \Phi^2 y_2 + \Phi y_3 + y_4 + \Phi y_5 \dots + \Phi^{n-4} y_n = 0 & (4) \\ \Phi^4 y_1 + \Phi^3 y_2 + \Phi^2 y_3 + \Phi y_4 + y_5 \dots + \Phi^{n-5} y_n = 0 & (5) \end{cases}$$

$$(4) - \Phi \times (5) \iff (\Phi^3 - \Phi^5) y_1 + (\Phi^2 - \Phi^4) y_2 + (\Phi - \Phi^3) y_3 + (1 - \Phi^2) y_4 = 0$$

On obtient :

$$y_4 = -\frac{(\Phi^3 - \Phi^5) y_1 + (\Phi^2 - \Phi^4) y_2 + (\Phi - \Phi^3) y_3}{(1 - \Phi^2)} \quad (2.51)$$

Si l'on substitue (2.46) et (2.48) et (2.50) dans (2.51), on retrouve y_4 :

$$y_4 = 0 \quad (2.52)$$

On suit le même raisonnement jusqu'à ce que l'on retrouve la valeur de y_n :

$$y_n = 0 \quad (2.53)$$

C'est à dire :

$$\begin{cases} y_1 = \frac{1}{1 - \Phi^2} \\ y_2 = -\frac{\Phi}{1 - \Phi^2} \\ y_i = 0 \text{ pour } i = 3, \dots, n \end{cases}$$

Finalement la solution de système (A) est :

$$y = \left(\frac{1}{1 - \Phi^2}, -\frac{\Phi}{1 - \Phi^2}, 0, \dots, 0 \right)^t \quad (2.54)$$

2. Résolution du système (B)

$$(B) \iff T \times x = f$$

$$\iff \begin{pmatrix} 1 & \Phi & \Phi^2 & \dots & \Phi^{n-1} \\ \Phi & 1 & \Phi & \dots & \Phi^{n-2} \\ \Phi^2 & \Phi & 1 & \dots & \Phi^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Phi^{n-1} & \Phi^{n-2} & \Phi^{n-3} & \dots & 1 \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \Phi^{n-1} - \Phi \\ \Phi^{n-i} - \Phi^i \\ \vdots \\ \Phi - \Phi^{n-1} \end{pmatrix}$$

On obtient :

$$(B) \iff \begin{cases} x_1 + \Phi x_2 + \Phi^2 x_3 + \dots + \Phi^{n-1} x_n = 0 & \dots & (1) \\ \Phi x_1 + x_2 + \Phi x_3 + \dots + \Phi^{n-2} x_n = \Phi^{n-1} - \Phi & \dots & (2) \\ \Phi^2 x_1 + \Phi x_2 + x_3 + \dots + \Phi^{n-3} x_n = \Phi^{n-2} - \Phi^2 & \dots & (3) \\ \vdots & \vdots & \vdots \\ \Phi^{n-2} x_1 + \Phi^{n-3} x_2 + \dots + x_{n-1} + \Phi x_n = \Phi^2 - \Phi^{n-2} & \dots & (n-1) \\ \Phi^{n-1} x_1 + \Phi^{n-2} x_2 + \dots + \Phi x_{n-1} + x_n = \Phi - \Phi^{n-1} & \dots & (n) \end{cases}$$

- On détermine la valeur de x_1 :

$$\begin{cases} x_1 + \Phi x_2 + \Phi^2 x_3 + \dots + \Phi^{n-1} x_n = 0 & (1) \\ \Phi x_1 + x_2 + \Phi x_3 + \dots + \Phi^{n-2} x_n = \Phi^{n-1} - \Phi & (2) \end{cases}$$

$$(1) - \Phi \times (2) \iff x_1 - \Phi^2 x_1 = \Phi^2 - \Phi^n$$

On obtient :

$$x_1 = \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} \quad (2.55)$$

- On détermine la valeur de x_2 :

$$\begin{cases} \Phi x_1 + x_2 + \Phi x_3 + \dots + \Phi^{n-2} x_n = \Phi^{n-1} - \Phi & (2) \\ \Phi \times (\Phi^2 x_1 + \Phi x_2 + x_3 + \dots + \Phi^{n-3} x_n = \Phi^{n-2} - \Phi^2) & (3) \end{cases}$$

$$(2) - \Phi \times (3) \iff x_1 (\Phi - \Phi^3) + x_2 (1 - \Phi^2) = \Phi^3 - \Phi$$

Si on remplace (2.55) dans cette dernière équation on retrouve la valeur

de x_2 sous la forme suivante :

$$x_2 = \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} \quad (2.56)$$

- On détermine la valeur de x_3 :

$$\begin{cases} \Phi^2 x_1 + \Phi x_2 + x_3 + \dots + \Phi^{n-3} x_n = \Phi^{n-2} - \Phi^2 & (3) \\ \Phi^3 x_1 + \Phi^2 x_2 + \Phi x_3 + x_4 + \dots + \Phi^{n-4} x_n = \Phi^{n-3} - \Phi^3 & (4) \end{cases}$$

$$(3) - \Phi \times (4) \iff x_1(\Phi^2 - \Phi^4) + x_2(\Phi - \Phi^3) + (1 - \Phi^2)x_3 = \Phi^4 - \Phi^2$$

Si l'on substitue (2.55) et (2.56) dans cette dernière équation, on obtient la valeur de x_3 :

$$x_3 = 0 \quad (2.57)$$

On suit les même étapes pour trouver les valeurs des x_i tel que $i = 4, \dots, n-1$ on obtient $x_i = 0$ pour $i = 3, \dots, n-1$.

- On détermine la valeur de x_n :

Pour déterminer la valeur de x_n , il suffit de remplacer les valeurs des x_i tel que $i = 1, \dots, n-1$ dans la première équation de système (B).

$$\begin{aligned} (1) & \iff x_1 + \Phi x_2 + \Phi^2 x_3 + \dots + \Phi^{n-1} x_n = 0 \\ & \iff \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} + \Phi \times \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} + 0 + \dots + 0 + \Phi^{n-1} x_n = 0 \end{aligned}$$

On obtient :

$$x_n = \frac{\Phi^n - \Phi^{n+2}}{\Phi^{n-1} - \Phi^{n+1}} = \Phi \quad (2.58)$$

Finalement la solution du système (B) est sous la forme suivante :

$$x = \left(\frac{\Phi^2 - \Phi^n}{1 - \Phi^2}, \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2}, 0, \dots, 0, \Phi \right)^t \quad (2.59)$$

Dès qu'on retrouve les solutions de deux systèmes on a la forme des matrices suivantes : T_1, U_1, T_2, U_2 .

Alors :

$$T_1 = \begin{pmatrix} y_1 & y_n & \cdots & y_2 \\ y_2 & y_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & y_n \\ y_n & \cdots & y_2 & y_1 \end{pmatrix} = \begin{pmatrix} \frac{1}{1-\Phi^2} & 0 & \cdots & 0 & -\frac{\Phi}{1-\Phi^2} \\ -\frac{\Phi}{1-\Phi^2} & \frac{1}{1-\Phi^2} & 0 & \cdots & 0 \\ 0 & -\frac{\Phi}{1-\Phi^2} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\frac{\Phi}{1-\Phi^2} & \frac{1}{1-\Phi^2} \end{pmatrix}$$

Ou plus explicitement

$$T_1 = \frac{1}{1-\Phi^2} \begin{pmatrix} 1 & 0 & \cdots & 0 & -\Phi \\ -\Phi & 1 & 0 & \cdots & 0 \\ 0 & -\Phi & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\Phi & 1 \end{pmatrix}$$

$$T_2 = \begin{pmatrix} x_1 & x_n & \cdots & x_2 \\ x_2 & x_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_n \\ x_n & \cdots & x_2 & x_1 \end{pmatrix} = \begin{pmatrix} \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} & \Phi & \cdots & 0 & \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} \\ \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} & \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} & \Phi & \cdots & 0 \\ 0 & \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \Phi \\ \Phi & \cdots & 0 & \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} & \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} \end{pmatrix}$$

$$U_1 = \begin{pmatrix} 1 & -x_n & \cdots & -x_2 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -x_n \\ 0 & \cdots & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\Phi & \cdots & \frac{\Phi - \Phi^{n+1}}{1 - \Phi^2} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\Phi \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

$$U_2 = \begin{pmatrix} 0 & y_n & \cdots & y_2 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & y_n \\ 0 & \cdots & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & -\frac{\Phi}{1 - \Phi^2} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

Ou plus explicitement

$$U_2 = \frac{1}{1-\Phi^2} \begin{pmatrix} 0 & 0 & \cdots & -\Phi \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

Donc on peut calculer la matrice Γ^{-1}

$$\Gamma^{-1} = \frac{(1 - \Phi^2)}{\sigma^2} T^{-1} \quad \text{sachant que } T^{-1} = T_1 U_1 + T_2 U_2$$

Alors :

$$\Gamma^{-1} = \frac{(1 - \Phi^2)}{\sigma^2} [T_1 U_1 + T_2 U_2]$$

$$\Gamma^{-1} = \frac{(1 - \Phi^2)}{\sigma^2} \left[\frac{1}{1 - \Phi^2} \begin{pmatrix} 1 & 0 & \cdots & 0 & -\Phi \\ -\Phi & 1 & 0 & \cdots & 0 \\ 0 & -\Phi & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\Phi & 1 \end{pmatrix} \times \begin{pmatrix} 1 & -\Phi & \cdots & \frac{\Phi - \Phi^{n+1}}{1 - \Phi^2} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\Phi \\ 0 & \cdots & 0 & 1 \end{pmatrix} \right] +$$

$$\frac{(1 - \Phi^2)}{\sigma^2} \left[\frac{1}{1 - \Phi^2} \begin{pmatrix} \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} & \Phi & \cdots & 0 & \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} \\ \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} & \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} & \Phi & \cdots & 0 \\ 0 & \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \Phi \\ \Phi & \cdots & 0 & \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} & \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} \end{pmatrix} \times \begin{pmatrix} 0 & 0 & \cdots & -\Phi \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} \right]$$

On pose

$$R = \begin{pmatrix} 1 & 0 & \cdots & 0 & -\Phi \\ -\Phi & 1 & 0 & \cdots & 0 \\ 0 & -\Phi & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\Phi & 1 \end{pmatrix} \times \begin{pmatrix} 1 & -\Phi & \cdots & \frac{\Phi - \Phi^{n+1}}{1 - \Phi^2} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\Phi \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Et aussi on pose :

$$S = \begin{pmatrix} \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} & \Phi & \cdots & 0 & \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} \\ \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} & \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} & \Phi & \cdots & 0 \\ 0 & \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \Phi \\ \Phi & \cdots & 0 & \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} & \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} \end{pmatrix} \times \begin{pmatrix} 0 & 0 & \cdots & -\Phi \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

Alors l'inverse de la matrice variance-covariance devienne :

$$\Gamma^{-1} = \frac{1}{\sigma^2} [R + S] \quad (2.60)$$

Maintenant nous calculons les deux matrices R et S :

$$R = \begin{pmatrix} 1 & 0 & \cdots & 0 & -\Phi \\ -\Phi & 1 & 0 & \cdots & 0 \\ 0 & -\Phi & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\Phi & 1 \end{pmatrix} \times \begin{pmatrix} 1 & -\Phi & \cdots & \frac{\Phi - \Phi^{n+1}}{1 - \Phi^2} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\Phi \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

On obtient :

$$R_{(n,n)} = \begin{pmatrix} 1 & -\Phi & 0 & 0 & \cdots & 0 & \frac{\Phi^3 - \Phi^{n+1}}{1 - \Phi^2} \\ -\Phi & \Phi^2 + 1 & -\Phi & \ddots & \ddots & \ddots & \frac{\Phi^{n+2} - \Phi^2}{1 - \Phi^2} \\ 0 & -\Phi & \Phi^2 + 1 & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & -\Phi & 0 \\ \vdots & \vdots & \ddots & \ddots & -\Phi & \Phi^2 + 1 & -\Phi \\ 0 & \cdots & \cdots & \cdots & 0 & -\Phi & \Phi^2 + 1 \end{pmatrix} \quad (2.61)$$

$$S = \begin{pmatrix} \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} & \Phi & \cdots & 0 & \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} \\ \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} & \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} & \Phi & \cdots & 0 \\ 0 & \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \Phi \\ \Phi & \cdots & 0 & \frac{\Phi^{n+1} - \Phi}{1 - \Phi^2} & \frac{\Phi^2 - \Phi^n}{1 - \Phi^2} \end{pmatrix} \times \begin{pmatrix} 0 & 0 & \cdots & -\Phi \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

On obtient :

$$S_{(n,n)} = \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 & \frac{\Phi^{n+1} - \Phi^3}{1 - \Phi^2} \\ 0 & 0 & \ddots & \vdots & \vdots & \frac{\Phi^2 - \Phi^{n+2}}{1 - \Phi^2} \\ \vdots & \ddots & \ddots & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & -\Phi^2 \end{pmatrix} \quad (2.62)$$

Si on remplace (2.61) et (2.62) dans (2.60) on retrouve la matrice Γ^{-1} :

$$\Gamma^{-1} = \frac{1}{\sigma^2} \left[\begin{pmatrix} 1 & -\Phi & 0 & \dots & 0 & \frac{\Phi^3 - \Phi^{n+1}}{1 - \Phi^2} \\ -\Phi & \Phi^2 + 1 & -\Phi & \ddots & \ddots & \frac{\Phi^{n+2} - \Phi^2}{1 - \Phi^2} \\ 0 & -\Phi & \Phi^2 + 1 & \ddots & \ddots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -\Phi & 0 \\ \vdots & \vdots & \ddots & -\Phi & \Phi^2 + 1 & -\Phi \\ 0 & \dots & \dots & 0 & -\Phi & \Phi^2 + 1 \end{pmatrix} + \begin{pmatrix} 0 & \dots & 0 & \frac{\Phi^{n+1} - \Phi^3}{1 - \Phi^2} \\ 0 & \vdots & \vdots & \frac{\Phi^2 - \Phi^{n+2}}{1 - \Phi^2} \\ \vdots & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & 0 \\ 0 & \dots & 0 & -\Phi^2 \end{pmatrix} \right]$$

Finalement l'inverse de la matrice variance-covariance d'un modèle autorégressif AR(1) est sous la forme suivante :

$$\Gamma_{(n,n)}^{-1} = \frac{1}{\sigma^2} \begin{pmatrix} 1 & -\Phi & 0 & \dots & 0 & 0 \\ -\Phi & \Phi^2 + 1 & -\Phi & 0 & \vdots & \vdots \\ 0 & -\Phi & \Phi^2 + 1 & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & -\Phi & 0 \\ \vdots & \vdots & \ddots & -\Phi & \Phi^2 + 1 & -\Phi \\ 0 & 0 & \dots & 0 & -\Phi & 1 \end{pmatrix} \quad (2.63)$$

2.4.2.2 La fonction de vraisemblance pour un processus autorégressif AR (1)

La fonction de vraisemblance du modèle autorégressif AR(1) avec la méthode de Toeplitz s'obtient par remplacement de (2.63) dans (2.24)

$$L(\theta) = -\frac{n}{2} \log(2\pi) + \frac{1}{2} \log \left(\det \frac{1}{\sigma^2} \begin{pmatrix} 1 & -\Phi & 0 & \dots & 0 & 0 \\ -\Phi & \Phi^2+1 & -\Phi & 0 & \vdots & \vdots \\ 0 & -\Phi & \Phi^2+1 & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & -\Phi & 0 \\ \vdots & \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & 0 & \dots & 0 & -\Phi & 1 \end{pmatrix} \right) -$$

$$\frac{1}{2\sigma^2} X^t \begin{pmatrix} 1 & -\Phi & 0 & \dots & 0 & 0 \\ -\Phi & \Phi^2+1 & -\Phi & 0 & \vdots & \vdots \\ 0 & -\Phi & \Phi^2+1 & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & -\Phi & 0 \\ \vdots & \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & 0 & \dots & 0 & -\Phi & 1 \end{pmatrix} X \quad (2.64)$$

Pour déterminer l'expression explicite de la fonction de vraisemblance, il suffit de calculer la forme quadratique et le déterminant

1. Calcul le déterminant :

$$\det(\Gamma^{-1}) = \det \frac{1}{\sigma^2} \begin{pmatrix} 1 & -\Phi & 0 & \dots & 0 & 0 \\ -\Phi & \Phi^2+1 & -\Phi & 0 & \vdots & \vdots \\ 0 & -\Phi & \Phi^2+1 & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & -\Phi & 0 \\ \vdots & \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & 0 & \dots & 0 & -\Phi & 1 \end{pmatrix}$$

On pose

$$B_{(n,n)} = \begin{pmatrix} 1 & -\Phi & 0 & \dots & 0 & 0 \\ -\Phi & \Phi^2 + 1 & -\Phi & 0 & \vdots & \vdots \\ 0 & -\Phi & \Phi^2 + 1 & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & -\Phi & 0 \\ \vdots & \vdots & \ddots & -\Phi & \Phi^2 + 1 & -\Phi \\ 0 & 0 & \dots & 0 & -\Phi & 1 \end{pmatrix}$$

On obtient :

$$\det(\Gamma^{-1}) = \left(\frac{1}{\sigma^{2n}}\right) \det(B_{(n,n)}) \quad (2.65)$$

$$\det(B_{(n,n)}) = 1 \cdot \begin{vmatrix} \Phi^2+1 & -\Phi & 0 & \vdots & \vdots \\ -\Phi & \Phi^2+1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & \dots & 0 & -\Phi & 1 \end{vmatrix} \cdot (-(-\Phi)) \begin{vmatrix} -\Phi & -\Phi & 0 & \vdots & \vdots \\ 0 & \Phi^2+1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & \dots & 0 & -\Phi & 1 \end{vmatrix}$$

$$\det(B_{(n,n)}) = \begin{vmatrix} \Phi^2+1 & -\Phi & 0 & \dots & 0 \\ -\Phi & \Phi^2+1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & \dots & 0 & -\Phi & 1 \end{vmatrix} \cdot (-\Phi^2) \begin{vmatrix} \Phi^2+1 & -\Phi & 0 & \dots & 0 \\ -\Phi & \Phi^2+1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & \dots & 0 & -\Phi & 1 \end{vmatrix}$$

$$+\Phi^2 \begin{vmatrix} 0 & -\Phi & 0 & \dots & 0 \\ 0 & \Phi^2+1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & \dots & 0 & -\Phi & 1 \end{vmatrix} \quad (2.66)$$

Tel que :

$$\begin{vmatrix} 0 & -\Phi & 0 & \dots & 0 \\ 0 & \Phi^2 + 1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2 + 1 & -\Phi \\ 0 & \dots & 0 & -\Phi & 1 \end{vmatrix} = 0$$

On pose :

$$\nabla_{(n-1)} = \begin{vmatrix} \Phi^2 + 1 & -\Phi & 0 & \dots & 0 \\ -\Phi & \Phi^2 + 1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2 + 1 & -\Phi \\ 0 & \dots & 0 & -\Phi & 1 \end{vmatrix}$$

Donc la relation (2.65) devienne :

$$\det(B_{(n,n)}) = \nabla_{(n-1)} - \Phi^2 \nabla_{(n-2)}$$

$$\det(B_{(n,n)}) = \nabla_{(n-1)} - \Phi^2 \nabla_{(n-2)} \quad (2.67)$$

et d'autre part on a :

$$\nabla_{(n-1)} = \begin{vmatrix} \Phi^2 + 1 & -\Phi & 0 & \dots & 0 \\ -\Phi & \Phi^2 + 1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & -\Phi & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2 + 1 & -\Phi \\ 0 & \dots & 0 & -\Phi & 1 \end{vmatrix}$$

$$= \begin{vmatrix} \Phi^{2+1} & -\Phi & 0 & \dots & 0 \\ -\Phi & \Phi^{2+1} & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^{2+1} & -\Phi \\ 0 & \dots & 0 & -\Phi & \Phi^{2+1} \end{vmatrix} + \Phi \begin{vmatrix} \Phi^{2+1} & -\Phi & 0 & 0 & \dots & 0 \\ -\Phi & \Phi^{2+1} & -\Phi & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \dots & 0 & -\Phi & \Phi^{2+1} & -\Phi \\ 0 & 0 & \dots & \dots & 0 & -\Phi \end{vmatrix} \quad (2.68)$$

Et on a :

$$\begin{vmatrix} \Phi^{2+1} & -\Phi & 0 & 0 & \dots & 0 \\ -\Phi & \Phi^{2+1} & -\Phi & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \dots & 0 & -\Phi & \Phi^{2+1} & -\Phi \\ 0 & 0 & \dots & \dots & 0 & -\Phi \end{vmatrix} = -\Phi \begin{vmatrix} \Phi^{2+1} & -\Phi & 0 & \dots & 0 \\ -\Phi & \Phi^{2+1} & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^{2+1} & -\Phi \\ 0 & \dots & 0 & -\Phi & \Phi^{2+1} \end{vmatrix} \quad (2.69)$$

Si l'on substitue (2.68) dans (2.67), on retrouve la relation suivante :

$$\nabla_{(n-1)} = \begin{vmatrix} \Phi^{2+1} & -\Phi & 0 & \dots & 0 \\ -\Phi & \Phi^{2+1} & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^{2+1} & -\Phi \\ 0 & \dots & 0 & -\Phi & \Phi^{2+1} \end{vmatrix} - \Phi^2 \begin{vmatrix} \Phi^{2+1} & -\Phi & 0 & \dots & 0 \\ -\Phi & \Phi^{2+1} & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^{2+1} & -\Phi \\ 0 & \dots & 0 & -\Phi & \Phi^{2+1} \end{vmatrix}$$

On pose

$$\tilde{\nabla}_{(n-2)} = \det \begin{pmatrix} \Phi^{2+1} & -\Phi & 0 & \dots & 0 \\ -\Phi & \Phi^{2+1} & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^{2+1} & -\Phi \\ 0 & \dots & 0 & -\Phi & \Phi^{2+1} \end{pmatrix}$$

Ce qui revient à écrire :

$$\nabla_{(n-1)} = \tilde{\nabla}_{(n-2)} - \Phi^2 \tilde{\nabla}_{(n-3)} \quad (2.70)$$

D'après le le remplacement de (2.70) dans (2.67), on obtient :

$$\begin{aligned} \det(B_{(n,n)}) &= \tilde{\nabla}_{(n-2)} - \Phi^2 \tilde{\nabla}_{(n-3)} - \Phi^2 \left[\tilde{\nabla}_{(n-3)} - \Phi^2 \tilde{\nabla}_{(n-4)} \right] \\ \det(B_{(n,n)}) &= \tilde{\nabla}_{(n-2)} - 2\Phi^2 \tilde{\nabla}_{(n-3)} + \Phi^4 \tilde{\nabla}_{(n-4)} \end{aligned} \quad (2.71)$$

Pour déterminer $\det(B_{(n,n)})$, on doit calculer $\tilde{\nabla}(i)$ tel que $i \in \{n-2, n-3, n-4\}$

$$\tilde{\nabla}_{(n-2)} = \begin{vmatrix} \Phi^2+1 & -\Phi & 0 & \dots & 0 \\ -\Phi & \Phi^2+1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & \dots & 0 & -\Phi & \Phi^2+1 \end{vmatrix}$$

$$\tilde{\nabla}_{(n-2)} = (\Phi^2+1) \begin{vmatrix} \Phi^2+1 & -\Phi & 0 & \dots & 0 \\ -\Phi & \Phi^2+1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & \dots & 0 & -\Phi & \Phi^2+1 \end{vmatrix} + \Phi \begin{vmatrix} -\Phi & -\Phi & 0 & \dots & 0 \\ 0 & \Phi^2+1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & \dots & 0 & -\Phi & \Phi^2+1 \end{vmatrix}$$

Et on pose :

$$\det C = \begin{vmatrix} -\Phi & -\Phi & 0 & \dots & 0 \\ 0 & \Phi^2+1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & \dots & 0 & -\Phi & \Phi^2+1 \end{vmatrix}$$

Alors :

$$\det C = -\Phi \begin{vmatrix} -\Phi & -\Phi & 0 & \cdots & 0 \\ 0 & \Phi^2+1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & \cdots & 0 & -\Phi & \Phi^2+1 \end{vmatrix} + \Phi \begin{vmatrix} 0 & -\Phi & 0 & \cdots & 0 \\ 0 & \Phi^2+1 & -\Phi & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\Phi & \Phi^2+1 & -\Phi \\ 0 & \cdots & 0 & -\Phi & 1 \end{vmatrix}$$

Ce qui implique

$$\tilde{\nabla}_{(n-2)} = (\Phi^2 + 1) \tilde{\nabla}_{(n-3)} - \Phi^2 \tilde{\nabla}_{(n-4)} \quad (2.72)$$

On suppose que l'équation (2.72) comme un polynôme de deuxième degré

$$(2.72) \iff \tilde{\nabla}_{(n-2)} - (\Phi^2 + 1) \tilde{\nabla}_{(n-3)} + \Phi^2 \tilde{\nabla}_{(n-4)} = 0$$

On pose $R = \tilde{\nabla}_{(n-3)}$, on obtient :

$$R^2 - (\Phi^2 + 1)R + \Phi^2 = 0$$

On calcule le discriminant de ce polynôme

$$\begin{aligned} \Delta &= (\Phi^2 + 1)^2 - 4 \times \Phi^2 \\ \Delta &= (1 - \Phi)^2 (1 + \Phi)^2 \end{aligned}$$

le discriminant Δ est positif donc il existe deux racines

$$R_1 = 1 \text{ et } R_2 = \Phi^2$$

Et d'autre part on a :

$$\tilde{\nabla}_{(n-2)} = \alpha R_1^{(n-2)} + \beta R_2^{(n-2)} \iff \tilde{\nabla}_{(n-2)} = \alpha + \beta \Phi^{2(n-2)}$$

Il suffit de trouver les valeurs de α et β :

$$\begin{cases} \cdot \text{ si } n = 3 : \tilde{\nabla}_{(1)} = \alpha + \beta \Phi^2 \\ \cdot \text{ si } n = 4 : \tilde{\nabla}_{(2)} = \alpha + \beta \Phi^4 \end{cases} \quad (2.73)$$

Où :

$$\begin{aligned} \tilde{\nabla}_{(1)} &= 1 + \Phi^2 \\ \tilde{\nabla}_{(2)} &= \begin{vmatrix} 1 + \Phi^2 & -\Phi \\ -\Phi & 1 + \Phi^2 \end{vmatrix} = (1 + \Phi^2)^2 - \Phi^2 \\ \tilde{\nabla}_{(2)} &= 1 + \Phi^2 + \Phi^4 \end{aligned}$$

Alors on obtient :

$$(2.72) \iff \begin{cases} 1 + \Phi^2 = \alpha + \beta\Phi^2 & (a) \\ 1 + \Phi^2 + \Phi^4 = \alpha + \beta\Phi^4 & (b) \end{cases}$$

$$(a) - (b) \iff \beta = \frac{-\Phi^2}{1 - \Phi^2}$$

On remplace la valeur de β dans l'équation (a), on obtient :

$$\alpha = \frac{1}{1 - \Phi^2}$$

Ce qui implique que :

$$\begin{aligned} \tilde{\nabla}_{(n-2)} &= \alpha + \beta\Phi^{2(n-2)} \\ &= \frac{1}{1 - \Phi^2} - \frac{\Phi^2 \times \Phi^{2(n-2)}}{1 - \Phi^2} \\ \tilde{\nabla}_{(n-2)} &= \frac{1}{1 - \Phi^2} (1 - \Phi^{2(n-2)+2}) \end{aligned}$$

Ce qui équivaut à

$$\tilde{\nabla}_{(i)} = \frac{1}{1 - \Phi^2} (1 - \Phi^{2(i)+2}) \text{ tel que } i \in \{n - 2, n - 3, n - 4\}$$

Si l'on substitue les valeurs de $\tilde{\nabla}_{(i)}$ dans l'équation (2.71) on retrouve le déterminant de la matrice B :

$$\begin{aligned} (2.71) \iff \det(B_{(n,n)}) &= \tilde{\nabla}_{(n-2)} - 2\Phi^2\tilde{\nabla}_{(n-3)} + \Phi^4\tilde{\nabla}_{(n-4)} \\ \iff \det(B_{(n,n)}) &= \frac{1}{1 - \Phi^2} (1 - \Phi^{2n-2}) - \frac{2\Phi^2}{1 - \Phi^2} (1 - \Phi^{2n-4}) + \frac{\Phi^4}{1 - \Phi^2} (1 - \Phi^{2n-6}) \end{aligned}$$

D'après les calculs on obtient :

$$\det(B_{(n,n)}) = (1 - \Phi^2) \quad (2.74)$$

Finalement le déterminant de l'inverse de la matrice variance-covariance $\Gamma_{(n,n)}^{-1}$ est donné par :

$$\det(\Gamma^{-1}) = \left(\frac{1}{\sigma^{2n}}\right) \det(B_{(n,n)})$$

Alors on a :

$$\det(\Gamma^{-1}) = \left(\frac{1}{\sigma^{2n}}\right) (1 - \Phi^2) \quad (2.75)$$

2. Calcul la forme quadratique :

$$\begin{aligned}
X^t \Gamma^{-1} X &= \frac{1}{\sigma^2} (x_1, x_2, \dots, x_n)^t \times \begin{pmatrix} 1 & -\Phi & 0 & \dots & \dots & 0 \\ -\Phi & \Phi^2+1 & -\Phi & \ddots & & \vdots \\ 0 & -\Phi & \Phi^2+1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & -\Phi & 0 \\ \vdots & & 0 & -\Phi & \Phi^2+1 & -\Phi \\ 0 & \dots & \dots & 0 & -\Phi & 1 \end{pmatrix} \\
&\times \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\
&= \frac{1}{\sigma^2} \left((x_1 - \Phi x_2), \dots, (-\Phi x_{i-2} + (\Phi^2 + 1)x_{i-1} - \Phi x_i), \dots, (-\Phi x_{n-1} + x_n) \right)^t \\
&\times \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
X^t \Gamma^{-1} X &= \frac{1}{\sigma^2} [(x_1 - \Phi x_2)x_1 + (-\Phi x_1 + (\Phi^2 + 1)x_2 - \Phi x_3)x_2 + \dots \\
&+ (-\Phi x_{n-2} + (\Phi^2 + 1)x_{n-1} - \Phi x_n)x_{n-1} + (-\Phi x_{n-1} + x_n)x_n]
\end{aligned}$$

$$\begin{aligned}
X^t \Gamma^{-1} X &= \frac{1}{\sigma^2} [\sum_{i=1}^{n-2} ((-\Phi x_i + (\Phi^2 + 1)x_{i+1} - \Phi x_{i+2})x_{i+1}) + x_1^2 + x_n^2 \\
&- \Phi (x_1 x_2 + x_{n-1} x_n)] \tag{2.76}
\end{aligned}$$

Si l'on substitue (2.75) et (2.76) dans (2.64), l'expression explicite de la fonction de log-vraisemblance d'un processus AR(1) gaussien est donnée par :

$$\begin{aligned}
L(\theta) &= -\frac{n}{2} \log(2\pi) + \frac{1}{2} \log\left[\frac{1}{\sigma^{2n}} (1 - \Phi^2)\right] - \frac{1}{2\sigma^2} [x_1^2 + x_n^2 - \Phi (x_1 x_2 + x_{n-1} x_n) + \\
&\sum_{i=1}^{n-2} ((-\Phi x_i + (\Phi^2 + 1)x_{i+1} - \Phi x_{i+2})x_{i+1})] \tag{2.77}
\end{aligned}$$

Ou plus explicitement

$$\begin{aligned}
(2.77) \iff L(\theta) &= -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \log(1 - \Phi^2) - \frac{1}{2\sigma^2} [x_1^2 + x_n^2 - \\
\Phi x_1 x_2 - \Phi x_{n-1} x_n + &-\Phi \sum_{i=1}^{n-2} x_i x_{i+1} + (\Phi^2 + 1) \sum_{i=1}^{n-2} x_{i+1}^2 - \Phi \sum_{i=1}^{n-2} x_{i+2} x_{i+1}]
\end{aligned}$$

$$(2.77) \iff L(\theta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \log(1 - \Phi^2) - \frac{1}{2\sigma^2} [(x_1^2 + x_n^2 + \sum_{i=1}^{n-2} x_{i+1}^2) + (-\Phi \sum_{i=1}^{n-2} x_i x_{i+1} - \Phi x_{n-1} x_n) + (-\Phi \sum_{i=1}^{n-2} x_{i+2} x_{i+1} - \Phi x_1 x_2) + \Phi^2 \sum_{i=1}^{n-2} x_{i+1}^2]$$

$$(2.77) \iff L(\theta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \log(1 - \Phi^2) - \frac{1}{2\sigma^2} [\sum_{i=0}^{n-1} x_{i+1}^2 - \Phi \sum_{i=1}^{n-1} x_i x_{i+1} - \Phi \sum_{i=0}^{n-2} x_{i+2} x_{i+1} + \Phi^2 \sum_{i=1}^{n-2} x_{i+1}^2]$$

$$(2.77) \iff L(\theta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \log(1 - \Phi^2) - \frac{1}{2\sigma^2} [\sum_{i=1}^n x_i^2 - 2\Phi \sum_{i=1}^{n-1} x_i x_{i+1} + \Phi^2 \sum_{i=1}^{n-2} x_{i+1}^2]$$

$$(2.77) \iff L(\theta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \log(1 - \Phi^2) - \frac{1}{2\sigma^2} [x_1^2 + \sum_{i=2}^n x_i^2 - 2\Phi \sum_{i=2}^n x_i x_{i-1} + \Phi^2 x_1^2 - \Phi^2 x_1^2 + \Phi^2 \sum_{i=1}^{n-2} x_{i+1}^2]$$

$$(2.77) \iff L(\theta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \log(1 - \Phi^2) - \frac{1}{2\sigma^2} [\sum_{i=2}^n x_i^2 - 2\Phi \sum_{i=2}^n x_i x_{i-1} + x_1^2 - \Phi^2 x_1^2 + \Phi^2 \sum_{i=0}^{n-2} x_{i+1}^2]$$

$$(2.77) \iff L(\theta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \log(1 - \Phi^2) - \frac{1}{2\sigma^2} [x_1^2 (1 - \Phi^2) + \sum_{i=2}^n x_i^2 - 2\Phi \sum_{i=2}^n x_i x_{i-1} + \Phi^2 \sum_{i=2}^n x_{i-1}^2]$$

$$(2.77) \iff L(\theta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \log(1 - \Phi^2) - \frac{1}{2\sigma^2} [x_1^2 (1 - \Phi^2) + \sum_{i=2}^n (x_i - \Phi x_{i-1})^2]$$

Finalement la fonction de log-vraisemblance d'un processus AR(1) gaussien est donnée par :

$$L(\vartheta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \frac{1}{2} \log(1 - \Phi^2) - \frac{1}{2\sigma^2} \left[(1 - \Phi^2) x_1^2 + \sum_{i=2}^n (x_i - \Phi x_{i-1})^2 \right] \quad (2.78)$$

2.4.2.3 Les estimateurs du maximum de vraisemblance des paramètres de modèle autorégressif AR(1)

Nous remarquons que l'expression explicite de la fonction de vraisemblance est la même que celle de Cholesky, mais la maximisation de la vraisemblance étant réalisée par des méthodes numériques.

Remarque 22 *Les résultats que nous avons obtenus concernent seulement le modèle autorégressif AR (1). dans le cas d'un modèle AR(p) nous avons rencontré des difficultés pour obtenir les mêmes résultats mais (Hamilton, J.D) a utilisé une autre façon pour donner l'expression explicite de la fonction de vraisemblance pour un modèle AR(p), (Voir Annexe).*

2.5 Conclusion

Ce chapitre a fourni un traitement plus détaillé sur l'estimation du maximum de vraisemblance des paramètres d'un processus autorégressif AR(1). Après de longs et complexes calculs, nous avons donné l'expression de la fonction de vraisemblance et on a réussi à trouver la forme analytique des estimateurs de maximum de vraisemblance d'un processus autorégressif AR(1). Dans le prochain chapitre, on confirmera les résultats analytiques par une étude de simulation des propriétés des estimateurs obtenus.

Chapitre 3

Exemples et illustrations numériques

Sommaire

2.1	Introduction et notations.	55
2.2	Modèle et Critères de performance	56
2.3	Analyse des qualités de l'estimateur MVC.	58
2.4	Analyse des qualités de l'estimateur MVTP.	61
2.5	Analyse des qualités de l'estimateur EYW.	64
2.6	Etude comparative.	67
2.7	Conclusion	74

3.1 Introduction et notations

Ce chapitre est consacré à l'étude par simulation des propriétés asymptotiques à savoir la convergence, la normalité asymptotique et l'efficacité des estimateurs obtenus. Toutes les méthodes d'estimation ont été implantées à l'aide du logiciel Matlab, exploité sous Windows . Nous présentons tout d'abord le modèle et les critères de performance utilisés. Nous considérons en suite les qualités de cet estimateur. Puis nous étudions les propriétés asymptotique de l'estimateur obtenu de chaque méthode. Enfin nous effectuons une comparaison avec les estimateurs obtenus, en analysant les gains de performance dus à l'estimation des différentes méthodes, mais avant de profiler dans ce chapitre nous donnons brièvement quelques notations que nous allons utiliser durant notre travail.

- **MVC** : La méthode de Maximum de Vraisemblance utilisant la décomposition de Cholésky.
- **MVTP** : La méthode de Maximum de Vraisemblance utilisant l'inverse de la matrice de Toeplitz.
- **EYW** : La méthode des équations de Yull-Walker.
- **REQM** : Racine de l'erreur quadratique moyenne.
- **B** : Biais.

3.2 Modèle et Critères de performance

Le modèle simulé est un modèle autorégressif d'ordre 1 suivant l'équation :

$$X_t = \Phi X_{t-1} + \varepsilon_t \text{ où } \varepsilon_t \sim N(0, \sigma^2). \quad (3.1)$$

Ce modèle est réalisé à partir du générateur de nombres aléatoires issu de la version 7.10.0 du logiciel MATLAB, les simulations sont faites pour les différentes valeurs du paramètre autorégressif $\Phi = (0.1, -0.5, 0.9)$ avec différentes valeurs de la variance de l'innovation $\sigma^2 = (1, 0.5)$. Pour valider les méthodes d'estimation qui sont présentées au Chapitre 3, on a d'abord voulu vérifier si les paramètres Φ et σ^2 du processus autorégressif, sont bien estimés. Pour ce faire, la racine de l'erreur quadratique moyenne et le biais de chacun des paramètres ont été estimés à partir de r répétitions pour un échantillon de taille n générés du modèle autorégressif d'ordre un. On rappelle que si $\hat{\theta}_n$ est un estimateur de θ , alors le biais et la racine carré de l'erreur quadratique moyenne empirique sont donnés comme suit :

3.2.1 Le biais (voir réf[9])

On définit le biais de la façon suivante :

$$B(\hat{\theta}_n) = E(\hat{\theta}_n) - \theta. \quad (3.2)$$

On dit qu'un estimateur est sans biais si $E(\hat{\theta}_n) = \theta$ et un estimateur sera asymptotiquement non biaisé si $\lim_{n \rightarrow \infty} B(\hat{\theta}_n) = 0$.

3.2.2 La racine de L'erreur quadratique moyenne (voir réf[9])

Elle permet de définir la qualité de l'estimateur au second degré.

$$REQM(\hat{\theta}_n) = \sqrt{EQM(\hat{\theta}_n)} = \sqrt{E\left(\left(\hat{\theta}_n - \theta\right)^2\right)} \quad (3.3)$$

Si $\lim_{n \rightarrow \infty} (REQM) = 0$ alors l'estimateur est asymptotiquement constant.

Et le lien entre les trois mesures est :

$$REQM(\hat{\theta}_n) = \sqrt{Var(\hat{\theta}_n) + B(\hat{\theta}_n)^2} \quad (3.4)$$

Remarque 23 Quand les deux expressions explicites pour $REQM(\hat{\theta})$, $B(\hat{\theta})$ ne sont pas disponibles, comme c'est le cas ici, on les estime par :

$$B(\hat{\theta}) = \frac{1}{r} \sum_{i=1}^r \hat{\theta}_n^i - \theta \quad (3.5)$$

$$\widehat{REQM}(\hat{\theta}) = \frac{1}{r} \sum_{i=1}^r \left(\hat{\theta}_n^i - \theta\right)^2 \quad (3.6)$$

où $\hat{\theta}_n^{(1)}, \dots, \hat{\theta}_n^{(r)}$ sont les estimateurs obtenus des différentes réalisations.

3.3 Analyse des qualités de l'estimateur MVC

Pour évaluer la performance de la méthode du Maximum de Vraisemblance utilisant la décomposition de Cholésky (*MVC*) décrite au Chapitre 3, on estimera la moyenne, le Biais et la racine de l'erreur quadratique moyenne pour différentes tailles d'échantillon. Et aussi on représente la distribution asymptotique de l'estimateur du paramètre autorégressif selon l'histogramme pour deux tailles d'échantillon différentes.

3.3.1 Résultats numériques

Les Tableaux (3.1), (3.2), (3.3) présentent les résultats de la moyenne, le biais et La racine de l'erreurs quadratiques moyennes estimés empiriquement de l'estimateur $\hat{\theta} = (\hat{\Phi}, \hat{\sigma}^2)$ à partir de r=10000 réalisation selon la méthode de (*MVC*) et les tailles d'échantillons considérées furent n = 50, 100, 1000, 10000. Et dans les figures (3.1), (3.2) nous présentons la distribution de l'estimateur du MVC pour deux tailles différentes n=50 et n=10000 à partir de 10000 réalisation.

		n=50		n=100		n=1000		n=10000	
Φ	σ^2	$E(\hat{\Phi})$	$E(\hat{\sigma}^2)$	$E(\hat{\Phi})$	$E(\hat{\sigma}^2)$	$E(\hat{\Phi})$	$E(\hat{\sigma}^2)$	$E(\hat{\Phi})$	$E(\hat{\sigma}^2)$
0.1		0.0944	0.9824	0.0992	0.9874	0.0996	0.9988	0.1000	0.9997
-0.5	1	-0.4806	0.9785	-0.4894	0.9900	-0.4987	0.9992	-0.4997	1.0000
0.9		0.8676	0.9905	0.8830	0.9922	0.8982	0.9987	0.8998	0.9999
0.1		0.0992	0.4912	0.0971	0.4949	0.1005	0.4993	0.1001	0.5000
-0.5	0.5	-0.4818	0.4915	-0.4900	0.4950	-0.4991	0.4994	-0.4998	0.4999
0.9		0.8671	0.4942	0.8828	0.4946	0.8980	0.4994	0.8999	0.4999

Table 3.1 : Les moyennes des estimateurs du MVC des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 10000 répétitions.

		n=50		n=100		n=1000		n=10000	
Φ	σ^2	$B(\widehat{\Phi})$	$B(\widehat{\sigma}^2)$	$B(\widehat{\Phi})$	$B(\widehat{\sigma}^2)$	$B(\widehat{\Phi})$	$B(\widehat{\sigma}^2)$	$B(\widehat{\Phi})$	$B(\widehat{\sigma}^2)$
0.1		-0.0056	-0.0176	-0.0007	-0.0126	-0.0003	-0.0012	0	-0.0002
-0.5	1	0.0194	-0.0215	0.0106	-0.0100	0.0013	-0.0008	0.0003	0
0.9		-0.0324	-0.0095	-0.0170	-0.0078	-0.0018	-0.0013	-0.0001	-0.0001
0.1		-0.0008	-0.0088	-0.0029	-0.0051	0.0005	-0.0007	0.0001	0
-0.5	0.5	0.0182	-0.0085	0.0100	-0.0050	-0.0009	-0.0005	0.0002	0
0.9		-0.0329	-0.0058	-0.0172	-0.0054	-0.0020	-0.0005	-0.0001	-0.0001

Table 3.2 : Le biais des estimateurs du MVC des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 10000 répétitions.

		n=50		n=100		n=1000		n=10000	
		REQM		REQM		REQM		REQM	
Φ	σ^2	$\widehat{\Phi}$	$\widehat{\sigma}^2$	$\widehat{\Phi}$	$\widehat{\sigma}^2$	$\widehat{\Phi}$	$\widehat{\sigma}^2$	$\widehat{\Phi}$	$\widehat{\sigma}^2$
0.1		0.1413	0.1996	0.0997	0.1433	0.0311	0.0447	0.0099	0.0143
-0.5	1	0.1254	0.1974	0.0888	0.1407	0.0276	0.0440	0.0087	0.0139
0.9		0.0841	0.2011	0.0533	0.1400	0.0144	0.0445	0.0043	0.0140
0.1		0.1399	0.0985	0.0996	0.0702	0.0315	0.0224	0.0099	0.0070
-0.5	0.5	0.1269	0.0993	0.0871	0.0705	0.0275	0.0224	0.0086	0.0070
0.9		0.0844	0.1005	0.0528	0.0704	0.0143	0.0223	0.0043	0.0071

Table 3.3 : La racine de l'erreur quadratique moyenne des estimateurs du MVC des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 10000 répétitions.

3.3.2 La Distribution asymptotique de l'estimateur du MVC

La normalité asymptotique de l'estimateur du (MVC) sera analysée à travers l'histogramme de la distribution de l'estimateur de paramètre autorégressif. On prend le cas où le paramètre autorégressif est faible ($\Phi = 0.1$) et la variance de l'innovation est égale à 1. Les figures (3.1), (3.2) présentent la distribution de l'estimateur du (MVC) pour deux tailles différentes $n=50$ et $n=10000$ à partir de 10000 réalisations.

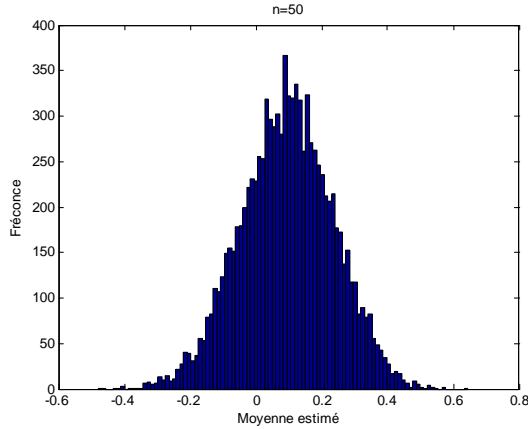


Figure 3.1 : Histogramme de l'estimateur du MVC de paramètre $\Phi = 0.1$ a partir de 10000 réalisations pour $n=50$.

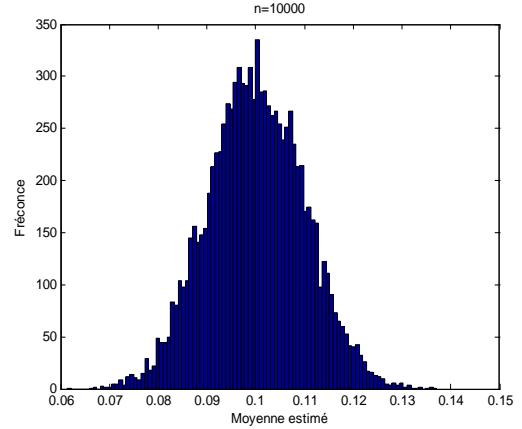


Figure 3.2 : Histogramme de l'estimateur du MVC de paramètre $\Phi = 0.1$ a partir de 10000 réalisations pour $n=10000$.

3.3.3 Interprétation des résultats

A la vue des résultats des tableaux (3.1), (3.2), (3.3) on constate les points suivants :

- On remarque que le Biais de $\hat{\Phi}$ et $\hat{\sigma}^2$ sont petits dans tous les cas considérés. Plus n augmente le Biais de $\hat{\sigma}^2$ et de $\hat{\Phi}$ diminue jusqu'à son annulation, donc on peut dire que l'estimateur du (MVC) est asymptotiquement sans biais.
- Les REQM sont généralement petites dans tous les cas considérés.
- Les estimateurs des paramètres Φ et σ^2 performant bien pour différents cas considérés. De plus, ces estimateurs deviennent davantage précis à mesure que la taille des échantillons augmente, car les REQM et le biais sont d'autant plus faibles que le nombre d'observations est grand.
- Egalement, on constate que la valeur du paramètre autorégressif Φ ne semble pas exercer une grande influence sur la précision de $\hat{\Phi}$. Ce n'est pas le cas de la variance : en effet, les REQM de $\hat{\sigma}^2$ sont beaucoup plus grandes pour $\sigma^2 = 1$ que pour $\sigma^2 = 0.5$.

- Nous remarquons d'après les figures (3.1), (3.2) Lorsque la taille d'échantillon n augmente, la distribution de l'estimateur $\hat{\Phi}$ se rapproche de plus en plus de la courbe en cloche de la loi normale.

Finalement ces résultats montrent également que la méthode d'estimation de (MVC) est efficace et donne d'excellents résultats, car l'estimateur de (MVC) admet des propriétés asymptotiques comme la convergence, la normalité asymptotiques et il est asymptotiquement sans biais.

3.4 Analyse des qualités de l'estimateur MVTP

Plusieurs simulations ont été réalisées pour étudier le comportement de l'estimateur du maximum de vraisemblance utilisant l'inverse de la matrice de Toeplitz ($MVTP$), ces simulations sont faites pour les différentes valeurs du paramètre autorégressif et la variance de l'innovation tel que $\Phi = (0.1, -0.5, 0.9)$ avec $\sigma^2 = (1, 0.5)$. Nous utilisons 100 réalisations pour plusieurs valeurs de la taille n de l'échantillon, l'estimateur $\hat{\theta} = (\hat{\Phi}, \hat{\sigma}^2)$ est calculé pour chaque réalisation et les résultats statistiques (la moyenne, le biais, REMQ, la distribution de l'estimateur $\hat{\Phi}$) sont donnés ci-dessous.

3.4.1 Résultats numériques

Les Tableaux (3.4), (3.5) et (3.6) présentent les résultats de la moyenne, le biais et REMQ de l'estimateur du MVTP à partir de 100 réalisation pour les tailles d'échantillon $n = 50$, $n = 100$ et $n = 1000$, respectivement, et les figures (3.3) et (3.4) donne la distribution de l'estimateur du MVTP pour deux tailles différentes $n=50$ et $n=1000$ à partir de 100 réalisations.

		n=50		n=100		n=1000	
Φ	σ^2	$E(\widehat{\Phi})$	$E(\widehat{\sigma}^2)$	$E(\widehat{\Phi})$	$E(\widehat{\sigma}^2)$	$E(\widehat{\Phi})$	$E(\widehat{\sigma}^2)$
0.1		0.0856	0.9739	0.0945	0.9891	0.1071	0.9968
-0.5	1	-0.4795	0.9837	-0.4875	0.9971	-0.4938	0.9968
0.9		0.8688	0.9836	0.8879	0.9896	0.8985	0.9967
0.1		0.1058	0.7023	0.0940	0.7016	0.1042	0.5911
-0.5	0.5	-0.4673	0.6896	-0.4887	0.6967	-0.4944	0.0735
0.9		0.8669	0.7127	0.8862	0.6944	0.8870	0.5744

Table 3.4 : Les moyennes des estimateurs du MVTP des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 100 répétitions.

		n=50		n=100		n=1000	
Φ	σ^2	$B(\widehat{\Phi})$	$B(\widehat{\sigma}^2)$	$B(\widehat{\Phi})$	$B(\widehat{\sigma}^2)$	$B(\widehat{\Phi})$	$B(\widehat{\sigma}^2)$
0.1		-0.0144	-0.0261	-0.0055	-0.0109	0.0071	-0.0032
-0.5	1	0.0205	-0.0163	0.0125	-0.0029	0.0062	-0.0032
0.9		-0.0312	-0.0164	-0.0121	-0.0104	-0.0015	-0.0033
0.1		0.0058	0.2023	-0.0060	0.2016	0.0042	0.0911
-0.5	0.5	0.0327	0.1896	0.0113	0.1967	0.0056	0.0735
0.9		-0.0331	0.2127	-0.0138	0.1944	-0.0013	0.0744

Table 3.5 : Le biais des estimateurs du MVTP des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 100 répétitions.

		n=50		n=100		n=1000	
		REQM		REQM		REQM	
Φ	σ^2	$\widehat{\Phi}$	$\widehat{\sigma}^2$	$\widehat{\Phi}$	$\widehat{\sigma}^2$	$\widehat{\Phi}$	$\widehat{\sigma}^2$
0.1		0.1412	0.1017	0.1073	0.0793	0.0320	0.0247
-0.5	1	0.1117	0.1016	0.0824	0.0736	0.0278	0.0247
0.9		0.0820	0.1006	0.0447	0.0656	0.0145	0.0247
0.1		0.1287	0.2128	0.0966	0.2058	0.0221	0.1963
-0.5	0.5	0.1375	0.2026	0.0873	0.2032	0.0250	0.1811
0.9		0.0754	0.2241	0.0529	0.2008	0.0165	0.1801

Table 3.6 : La racine de l'erreur quadratique moyenne des estimateurs du MVTP des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 100 répétitions.

3.4.2 La Distribution asymptotique de l'estimateur du MVTP

Nous présentons l'histogramme de la distribution de l'estimateur de paramètre autorégressif dans le cas où $\Phi = 0.1$ pour les tailles d'échantillon 50, 1000 à partir de 100 réalisations.

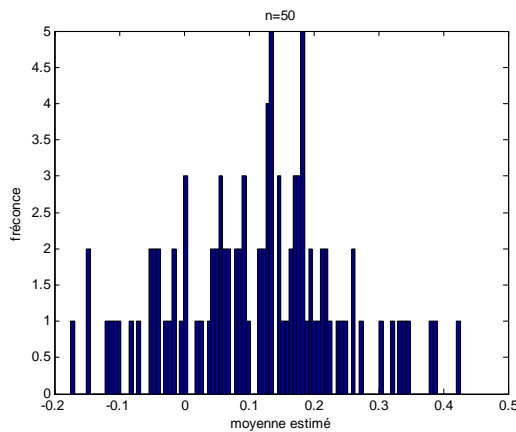


Figure 3.3 : Histogramme de l'estimateur du MVTP de paramètre $\Phi = 0.1$ à partir de 100 réalisations pour $n=50$.

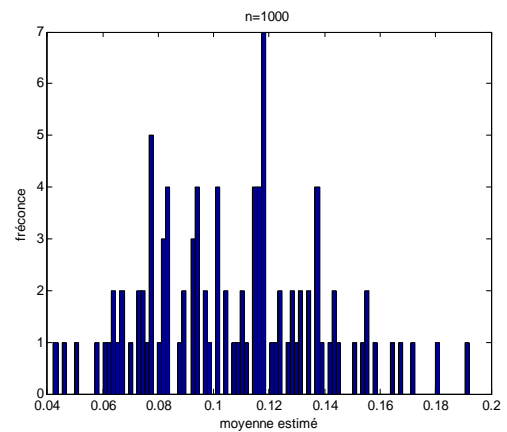


Figure 3.4 : Histogramme de l'estimateur du MVTP de paramètre $\Phi = 0.1$ à partir de 100 réalisations pour $n=1000$.

3.4.3 Interprétation des résultats

Selon les résultats des Tableaux (3.4), (3.5),(3.6) on voit que :

- les estimateurs sont proches des vraies valeurs et cela pour les divers cas considérés et différentes tailles d'échantillon n , et plus n augmente plus les estimateurs s'améliorent et s'approchent de la vraie valeur des paramètres, et plus la valeur du biais diminue.
- Les REQM sont généralement très petites. Egalement, on voit que la performance (REQM) des estimateurs a diminuer lorsque la taille d'échantillon

n augmente, donc on peut dire que l'estimateur du MVTP est asymptotiquement consistant.

- la valeur du paramètre autorégressif Φ ne semble pas exercer une grande influence sur la précision de $\hat{\Phi}$. Ce n'est pas le cas de la variance : en effet, la REQM de $\hat{\sigma}^2$ sont beaucoup plus grandes pour $\sigma^2 = 0.5$ que pour $\sigma^2 = 1$.
- D'après les figures (3.3), (3.4) on remarque que Plus la taille d'échantillon n augmente, la distribution de l'estimateur de paramètre autorégressif $\hat{\Phi}$ se rapproche de plus en plus de la courbe en cloche de la loi normale, alors l'estimateur du MVTP de paramètre autorégressif Φ est asymptotiquement distribué selon une loi normale.

Ces simulations intensives confirment que la méthode du MVTP donne des bons résultats, les propriétés asymptotiques générales sont vérifiées soit la convergence et la normalité asymptotique et aussi l'estimateur du MVTP est asymptotiques sans biais, mais l'inconvénient c'est qu'elle exige beaucoup de temps pour l'exécution à cause du calcul de déterminant de la matrice Γ^{-1} .

3.5 Analyse des qualités de l'estimateur EYW

La méthode de Yulle-Walker est une méthode d'estimation autorégressive qui est la plus simple et la plus rapide, pour analyser le comportement de l'estimateur (EYW), on a généré 10000 réalisations pour plusieurs valeurs de la taille n de l'échantillon avec différentes valeurs de Φ et σ^2 .

3.5.1 Résultats numériques

Dans les tableaux(3.7), (3.8) et (3.9) nous donnons les résultats de la moyenne, le biais et REQM de l'estimateur du MVTP à partir de 10000 réalisation pour les tailles d'échantillon $n = 50$, $n = 100$ et $n = 1000$, $n=10000$ respectivement, et les figures 3.3 et 3.4 donne la distribution de l'estimateur du EYW pour deux tailles différentes $n=50$ et $n=1000$ à partir de 10000 réalisations.

		n=50		n=100		n=1000		n=10000	
Φ	σ^2	$E(\hat{\Phi})$	$E(\hat{\sigma}^2)$	$E(\hat{\Phi})$	$E(\hat{\sigma}^2)$	$E(\hat{\Phi})$	$E(\hat{\sigma}^2)$	$E(\hat{\Phi})$	$E(\hat{\sigma}^2)$
0.1		0.0965	0.9792	0.0978	0.9889	0.0994	0.9996	0.1000	1.0010
-0.5	1	-0.4715	1.0111	-0.4847	1.0066	-0.4985	0.9986	0.4999	1.0000
0.9		0.8470	1.0656	0.8733	1.0313	0.8971	1.0034	0.8997	1.0002
0.1		0.095	0.4880	0.0982	0.4959	0.1018	0.4970	0.1000	0.5000
-0.5	0.5	-0.4709	0.4930	-0.4853	0.4972	-0.4985	0.4998	-0.4998	0.5000
0.9		0.8700	0.6977	0.8722	0.5154	0.8973	0.5015	0.8998	0.5000

Table 3.7 : Les moyennes des estimateurs du EYW des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 10000 répétitions.

		n=50		n=100		n=1000		n=10000	
Φ	σ^2	$B(\hat{\Phi})$	$B(\hat{\sigma}^2)$	$B(\hat{\Phi})$	$B(\hat{\sigma}^2)$	$B(\hat{\Phi})$	$B(\hat{\sigma}^2)$	$B(\hat{\Phi})$	$B(\hat{\sigma}^2)$
0.1		-0.0035	-0.0208	-0.0022	-0.0111	-0.0006	-0.0004	0	0.001
-0.5	1	0.0285	0.0111	0.0153	0.0066	0.0015	-0.0014	0.0001	0
0.9		-0.0530	0.0656	-0.0267	0.0313	-0.0029	0.0034	-0.0003	0.0002
0.1		-0.0050	-0.0102	-0.0018	-0.0041	0.0018	-0.0030	0	0
-0.5	0.5	0.0291	-0.0070	0.0147	-0.0028	0.0015	-0.0002	0.0002	0
0.9		-0.0300	0.1977	-0.0278	0.0154	-0.0027	0.0015	-0.0002	0

Table 3.8 : Le biais des estimateurs du EYW des paramètres du modèle autorégressif d'ordre un, estimées à l'aide de 10000 répétitions.

		n=50		n=100		n=1000		n=10000	
		REQM		REQM		REQM		REQM	
Φ	σ^2	$\hat{\Phi}$	$\hat{\sigma}^2$	$\hat{\Phi}$	$\hat{\sigma}^2$	$\hat{\Phi}$	$\hat{\sigma}^2$	$\hat{\Phi}$	$\hat{\sigma}^2$
0.1		0.1364	0.1993	0.0982	0.1411	0.0318	0.0450	0.0100	0.0142
-0.5	1	0.1245	0.2022	0.0887	0.1433	0.0274	0.0447	0.0086	0.0146
0.9		0.0963	0.2573	0.0585	0.1615	0.0142	0.0450	0.0044	0.0142
0.1		0.1380	0.1004	0.0989	0.0701	0.0317	0.0223	0.0099	0.0070
-0.5	0.5	0.1274	0.0989	0.0879	0.0707	0.0275	0.0220	0.0087	0.0071
0.9		0.0802	0.2099	0.0590	0.0818	0.0143	0.0228	0.0044	0.0070

Table 3.9 : La racine de l'erreurs quadratiques moyennes des estimateurs du EYW du modèle AR(1), estimées à l'aide de 10000 répétitions.

3.5.2 La Distribution Asymptotique de l'Estimateur

EYW

La normalité asymptotique de l'estimateur du EYW sera analysée à travers l'histogramme de la distribution de l'estimateur de paramètre autorégressif. Les figures (3.5), (3.6) présentent la distribution de l'estimateur du EYW dans le cas $\Phi = 0.1$ pour deux tailles différentes $n=50$ et $n=10000$ à partir de 10000 réalisations.

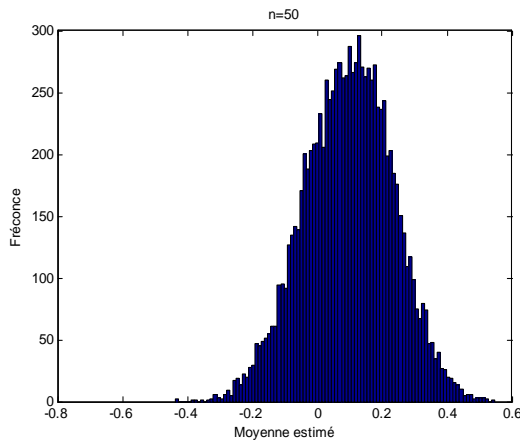


Figure 3.5 : Histogramme de l'estimateur du EYW de paramètre $\Phi = 0.1$ à partir de 10000 réalisations pour $n=50$.

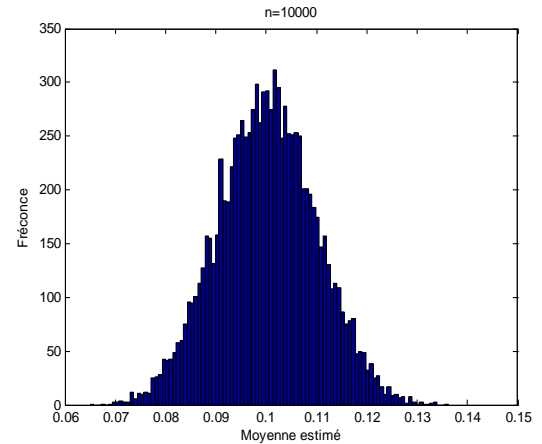


Figure 3.6 : Histogramme de l'estimateur du EYW de paramètre $\Phi = 0.1$ à partir de 10000 réalisations pour $n=10000$.

3.5.3 Interprétation des résultats

D'après les résultats des Tableaux (3.7), (3.8), (3.9) on remarque que :

- Le biais de $\hat{\Phi}$ et $\hat{\sigma}^2$ sont petites dans tous les cas considérés . On remarque aussi que les estimateurs deviennent davantage précis à mesure que n augmente en plus le biais se diminue jusqu'à s'annule. donc on peut dire que l'estimateur de EYW est asymptotiquement sans biais .
- Les REQM sont généralement très petites dans tout les cas considérés et plus n augmente plus la valeur REQM diminue.

- D'après les figures (3.5), (3.6) on remarque que Plus la taille d'échantillon n augmente, la distribution de l'estimateur de paramètre autorégressif $\hat{\Phi}$ se rapproche de plus en plus de la courbe en cloche de la loi normale, alors l'estimateur du EYW de paramètre autorégressif Φ est asymptotiquement distribué selon une loi normale.

Ces résultats montrent que l'estimateur de EYW est : asymptotiquement sans biais, converge vers la vraie valeur, et l'estimateur du paramètre autoregressif est asymptotiquement distribué selon la loi normale.

3.6 Etude comparative

L'objectif de ce paragraphe est d'analyser par simulation l'importance de chacun des points qui différencient les procédures MVC, MVTP, EYW. Pour se faire on s'intéresse à les valeurs du REQM et ceux de biais quand le nombre d'observations augmente et on rappelle que la racine de l'erreur quadratique moyenne et le biais sont très utiles pour comparer plusieurs estimateurs. Si les deux estimateurs à comparer sont biaisés ou sans biais, l'estimateur le plus efficace est simplement celui qui a l'erreur quadratique moyenne la plus faible, (Voir réf[12]).

Les valeurs du REQM et ceux de biais sont calculés pour différentes valeurs de $\Phi = (0.1, -0.5, 0.9)$ avec $\sigma^2 = 1$, à partir de 100 réalisations pour $n=25$ jusqu'à 1000. Les résultats sont donnés sous forme de graphique.

3.6.1 Résultats numériques

3.6.1.1 Première cas : synthèse de biais et de REQM pour $\Phi = 0.1$

avec $\sigma^2 = 1$

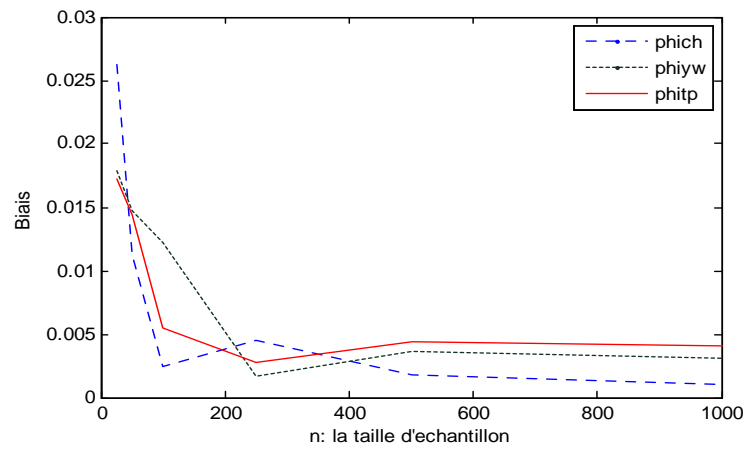


Figure 3.7 : Synthèse de biais de l'estimateur $\hat{\Phi}$ obtenu en fonction de n

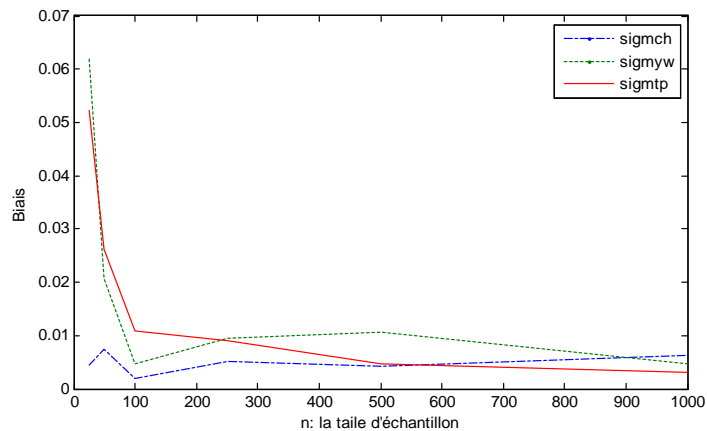


Figure 3.8 : Synthèse de biais de l'estimateur $\hat{\sigma}^2$ obtenu en fonction de n

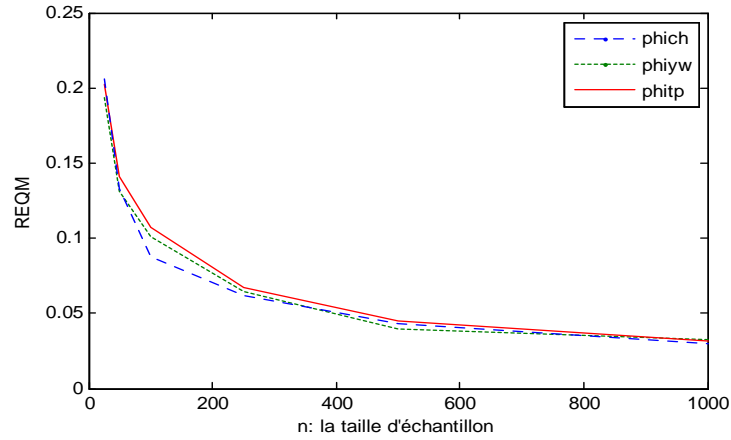


Figure 3.9 : Synthse de REQM de l'estimateur $\hat{\Phi}$ obtenu en fonction de n

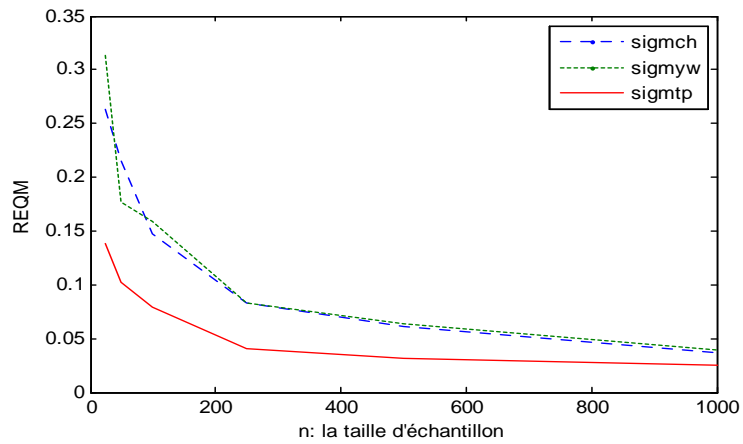


Figure 3.10 : Synthse de REQM de l'estimateur $\hat{\sigma}^2$ obtenu en fonction de n

3.6.1.2 Deuxième cas : synthèse de biais et de REQM $\Phi = -0.5$
avec $\sigma^2 = 1$

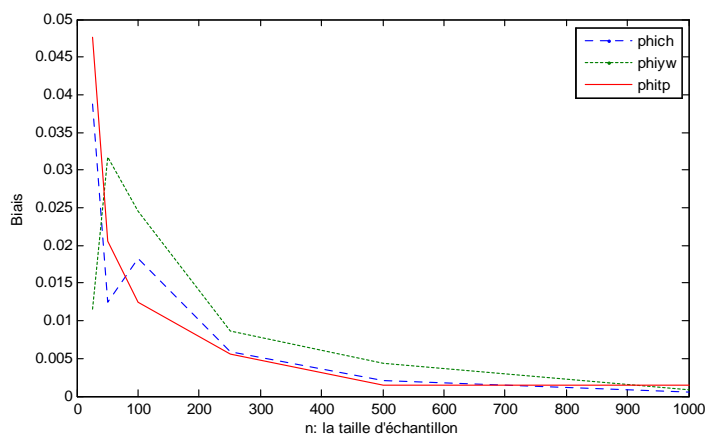


Figure 11 : Synthèse de biais de l'estimateur $\hat{\Phi}$ obtenu en fonction de n

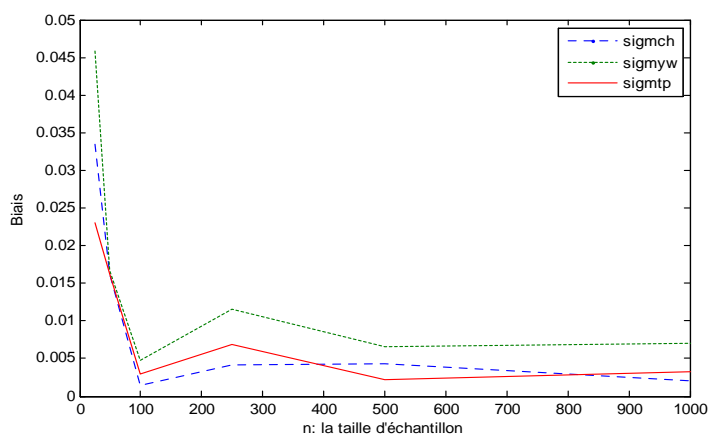


Figure 12 : Synthèse de biais de l'estimateur $\hat{\sigma}^2$ obtenu en fonction de n

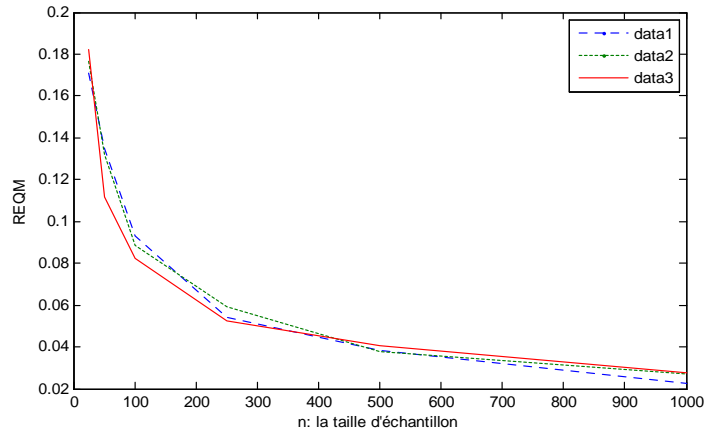


Figure 13 : Synthse de REQM de l'estimateur $\hat{\Phi}$ obtenu en fonction de n.

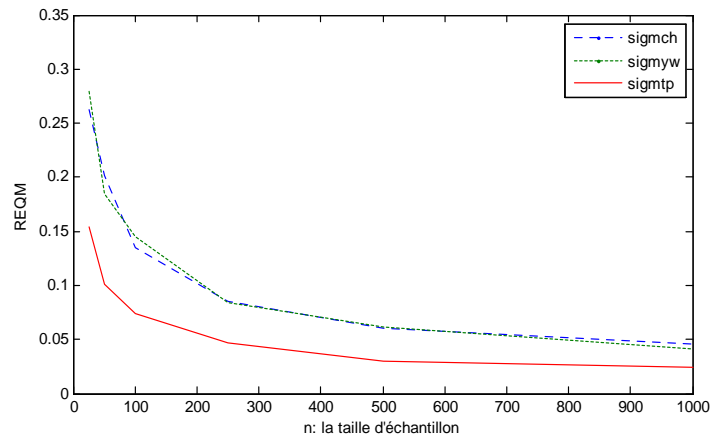


Figure 14 : Synthse de REQM de l'estimateur $\hat{\sigma}^2$ obtenu en fonction de n.

3.6.1.3 Troisième cas : synthèse de biais et de REQM pour $\Phi = 0.9$
avec $\sigma^2 = 1$

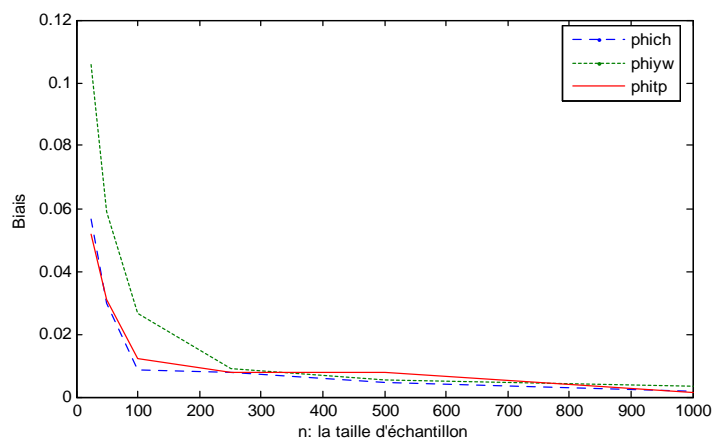


Figure 15 : Synthèse de biais de l'estimateur $\hat{\Phi}$ obtenu en fonction de n

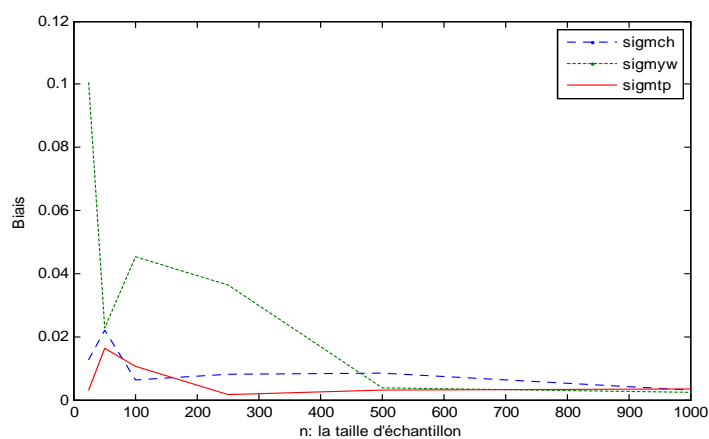


Figure 16 : Synthèse de biais de l'estimateur $\hat{\sigma}^2$ obtenu en fonction de n .

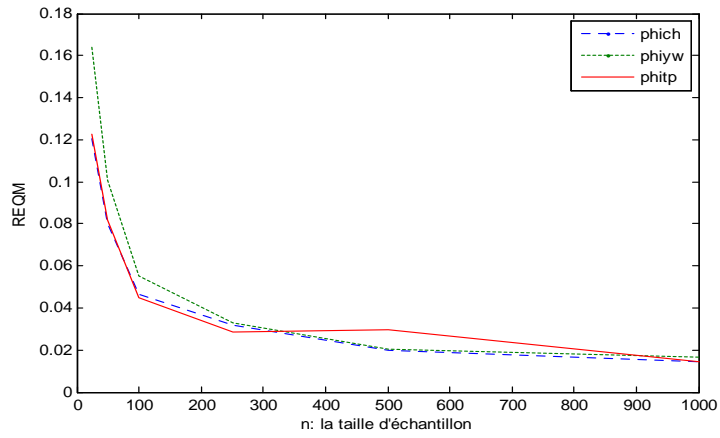


Figure 17 : Synthse de REQM de l'estimateur $\hat{\Phi}$ obtenu en fonction de n .

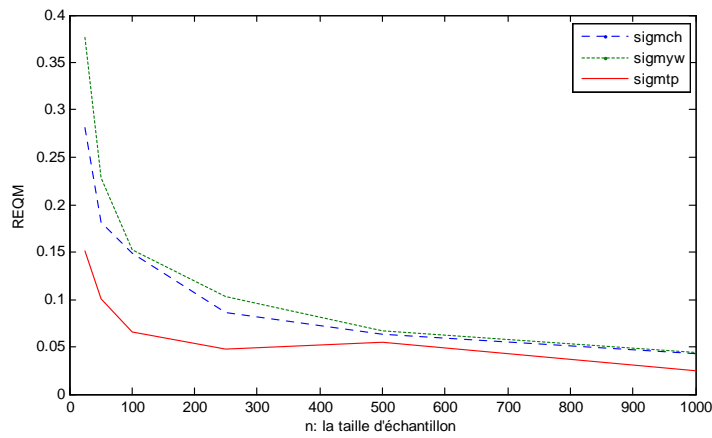


Figure 18 : Synthse de REQM de l'estimateur $\hat{\sigma}^2$ obtenu en fonction de n .

3.6.2 Interprétation des résultats

D'après les figures des synthèses du biais et de la REQM on remarque que :

- Pour les trois cas, elles confirment la convergence de l'estimateur obtenu de chaque méthode. On constate aussi que le biais devient négligeable à partir de $n \geq 500$, en particulier par rapport à la REQM qui décroît régulièrement.
- L'estimateur de la variance de l'innovation de MVTP est plus efficace que l'estimateur de MVC et de EYW par contre l'estimateur de paramètre autorégressif de MVC et de MVTP leurs REQM sont presque équivalents et ces deux dernières sont inférieures que la REQM de EYW.
- Les biais des estimateurs des paramètres (Φ, σ^2) de EYW sont plus grand que ceux de MVC et de MVTP et ils sont presque équitables quand $n \geq 800$.

3.7 Conclusion

Dans ce chapitre nous avons pu montrer des bonnes propriétés asymptotiques des estimateurs obtenus de chaque méthodes, telles que la convergence et la normalité asymptotique. Notons par ailleurs que l'estimation des paramètres d'un modèle autorégressif AR(1) utilisant la méthode MVTP donne de bons résultats, l'inconvénient est que son mise en œuvre complexe et lourde en temps de calcul. La méthode MVC donne presque les mêmes estimations que MVTP et elle n'exige pas beaucoup de temps d'exécution, par contre la méthode de yule-walker est plus simple et la plus rapide mais également moins performante dans le cas de séries courtes.

Conclusion Générale

Au cours de ce mémoire, nous avons présenté une étude basée sur l'estimation des paramètres de processus linéaire gaussien autorégressif AR(p) à partir d'une suite de n observations, et notre intérêt a été focalisé sur la méthode de Maximum de Vraisemblance.

Les résultats de ce travail se résument principalement en les points suivants :

i) Dans le premier chapitre on présente une méthode proposée par *Xiao-GuangLv, Ting-zhu Huang* qui nous a permis de calculer l'inverse de la matrice de covariance d'un modèle autorégressif d'ordre 1

ii) Dans le deuxième chapitre nous avons considéré l'estimation du maximum de vraisemblance pour le modèle autorégressif, nous avons construit les expressions explicites de la fonction de vraisemblance pour un modèle AR(1) et nous avons détecté l'EMV exact dans le cas d'un modèle autorégressif d'ordre 1.

Ce qui concerne le modèle AR(1) notre point de départ c'est d'utiliser deux méthodes :

1. La méthode de Maximum de Vraisemblance utilisant la décomposition de Cholésky qui nous a permis de trouver la forme analytique de l'EMV.
2. La méthode de Maximum de Vraisemblance utilisant l'inverse de la matrice de Toeplitz basée essentiellement sur le calcul de l'inverse de la matrice variance-covariance et son déterminant, mais avec cette dernière nous obtenons l'expression explicite de la fonction de vraisemblance est la même de celle de Cholesky, mais la maximisation de la vraisemblance étant réalisé par des méthodes numériques.

iii) Dans le troisième chapitre nous avons validée les propriétés asymptotiques de l'estimateur obtenu de chaque méthode par une simulation numérique, nous effectuons une comparaison avec la méthode de Yule-Walker.

Nous constatons que les deux méthodes MVC et MVTP donnent presque les mêmes estimations et des résultats très satisfaisants, mais l'inconvénient

de la méthode MVTP demande beaucoup de temps d'exécution de programme par rapport à MVC quelle est basé sur des formules explicites du l'EMV et facilement programmables pour des applications potentielles. La comparaison entre les méthodes (MVC, MVTP, EYW) montre la supériorité de la méthode de MVC notamment en terme de temps d'exécution par rapport à la méthode de MVTP et ces deux dernières sont bien meilleurs que la méthode de Yule-Walker dans le cas de séries courtes .

Nous espérons qu'à travers ce mémoire nous avons pu répondre à la problématique posée et que cette étude a donné les réponses espérées.

Annexe

Application la méthode du maximum de vraisemblance au modèle autorégressif AR(p)

L'évaluation de la fonction de vraisemblance du modèle autorégressif d'ordre p on se basant sur la relation unissant les fonctions de densité jointe, marginale et conditionnelle, (voir réf[13]).

Soit un processus AR(p) gaussien :

$$X_t = \Phi_1 X_{t-1} + \dots + \Phi_p X_{t-p} + \varepsilon_t \quad \text{où } \varepsilon_t \sim N(0, \sigma^2) \quad (1)$$

Le vecteur contenant les paramètres de la population à estimer est :

$$\theta = (\Phi_1, \Phi_2, \dots, \Phi_p, \sigma^2). \quad (2)$$

Définissons X_p comme le vecteur de dimension $(p \times 1)$ qui contient les p premières observations de l'échantillon (X_1, X_2, \dots, X_p) . De plus, nous savons que le processus AR(p) est centré. Définissons $\sigma^2 V_p$ comme la matrice variance-covariance de (X_1, X_2, \dots, X_p) :

$$\sigma^2 V_p = \begin{pmatrix} E(X_1)^2 & E(X_1.X_2) & \dots & E(X_1.X_p) \\ E(X_2.X_1) & E(X_2)^2 & \dots & E(X_2.X_p) \\ \vdots & \vdots & \dots & \vdots \\ E(X_p.X_1) & E(X_p.X_2) & \dots & E(X_p)^2 \end{pmatrix}$$

$$\sigma^2 V_p = \begin{pmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(p-1) \\ \gamma(1) & \gamma(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \gamma(1) \\ \gamma(p-1) & \dots & \gamma(1) & \gamma(0) \end{pmatrix} \quad (3)$$

Où γ_j est l'auto-covariance d'ordre j pour un processus AR(p).

La densité des p premières observations est celle d'une variable qui suit une loi $N(0, \sigma^2 V_p)$ est :

$$f_{X_p}(x_1, x_2, \dots, x_p, \theta) = (2\pi)^{-\frac{p}{2}} (\det(\sigma^2 V_p))^{-\frac{1}{2}} \exp \left[-\frac{1}{2\sigma^2} X_p^t V_p^{-1} X_p \right] \quad (4)$$

Pour le reste des observations dans l'échantillon, $(X_{p+1}, X_{p+2}, \dots, X_n)$, il est possible d'utiliser la décomposition de l'erreur de prévision. Conditionnelle aux $t-1$ premières observations, l'observation t est gaussienne et de moyenne égale $\Phi_1 x_{t-1} + \Phi_2 x_{t-2} + \dots + \Phi_p x_{t-p}$ et de variance σ^2 . Seuls les p observations les plus récentes sont prises en considération pour cette distribution. Alors, pour $t > p$:

$$f_{X_t|X_{t-1}, \dots, X_1}(x_t | x_{t-1}, \dots, x_1, \theta) = f_{X_t|X_{t-1}, \dots, X_{t-p}}(x_t | x_{t-1}, \dots, x_{t-p}, \theta)$$

$$f_{X_t|X_{t-1}, \dots, X_1}(x_t | x_{t-1}, \dots, x_1, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(x_t - \Phi_1 x_{t-1} - \Phi_2 x_{t-2} - \dots - \Phi_p x_{t-p})^2}{2\sigma^2} \right] \quad (5)$$

La fonction de vraisemblance pour la totalité de l'échantillon $X = (X_1, X_2, \dots, X_n)^t$ s'écrit :

$$f_X(x_n, x_{n-1}, \dots, x_1, \theta) = f_{X_p}(x_1, x_2, \dots, x_p, \theta) \times \prod_{t=p+1}^n f_{X_t|X_{t-1}, \dots, X_{t-p}}(x_t | x_{t-1}, \dots, x_{t-p}, \theta)$$

Et la log-vraisemblance est alors :

$$L(\theta) = \log f_X(x_n, x_{n-1}, \dots, x_1, \theta)$$

$$L(\theta) = -\frac{p}{2} \log(2\pi) - \frac{p}{2} \log(\sigma^2) + \frac{1}{2} \log(\det V_p^{-1}) - \frac{1}{2\sigma^2} X_p^t V_p^{-1} X_p - \frac{n-p}{2} \log(2\pi) - \frac{n-p}{2} \log(\sigma^2) - \sum_{t=1+p}^n \frac{(x_t - \Phi_1 x_{t-1} - \Phi_2 x_{t-2} - \dots - \Phi_p x_{t-p})^2}{2\sigma^2}$$

Dés lors :

$$L(\theta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) + \frac{1}{2} \log(\det V_p^{-1}) - \frac{1}{2\sigma^2} X_p^t V_p^{-1} X_p - \sum_{t=1+p}^n \frac{(x_t - \Phi_1 x_{t-1} - \Phi_2 x_{t-2} - \dots - \Phi_p x_{t-p})^2}{2\sigma^2} \quad (6)$$

L'évaluation de (6) nécessite inversant la matrice V_p d'ordre $(p \times p)$, on note l'élément de V_p^{-1} de la colonne j et de ligne i par $v_{ij}(p)$.

Galbraith montré que, (voir réf[10]) :

$$v_{ij}(p) = \left[\sum_{k=0}^{i-1} \Phi_k \Phi_{k+j-i} - \sum_{k=p+1-j}^{p+i-j} \Phi_k \Phi_{k+j-i} \right] \text{ pour } 1 \preceq i \preceq j \preceq p \quad (7)$$

où $\Phi_0 = -1$. Valeurs de $v_{ij}(p)$ pour $i \succ j$ on peut le déduire du fait que V_p^{-1} est symétrique ($v_{ij}(p) = v_{ji}(p)$).

Exemple 24 Pour un processus $AR(1)$, V_p^{-1} est un scalaire, dont la valeur est trouvé en prenant $i = j = p = 1$:

$$V_1^{-1} = \left[\sum_{k=0}^0 \Phi_k \Phi_k - \sum_{k=1}^1 \Phi_k \Phi_k \right] = (\Phi_0^2 - \Phi_1^2) = (1 - \Phi_1^2). \text{ ainsi la variance de } AR(1) \text{ est } \gamma_0 = \sigma^2 V = \frac{\sigma^2}{(1 - \Phi_1^2)}.$$

si on substitue la valeur de V_1^{-1} dans (6) on obtient la fonction de log vraisemblance de $AR(1)$ est la même de celle de Cholésky.

$$L(\theta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \frac{1}{2} \log(1 - \Phi^2) - \frac{1}{2\sigma^2} \left[(1 - \Phi^2) x_1^2 + \sum_{i=2}^n (x_i - \Phi x_{i-1})^2 \right]$$

Exemple 25 pour un processus $AR(2)$ la relation (7) implique :

$$V_2^{-1} = \begin{pmatrix} (1 - \Phi_2^2) & -(\Phi_1 + \Phi_1 \Phi_2) \\ -(\Phi_1 + \Phi_1 \Phi_2) & (1 - \Phi_2^2) \end{pmatrix}$$

On calcule facilement $\det(V_2^{-1})$ et $X_2^t V_2^{-1} X_2$

$$\det(V_2^{-1}) = \det \left((1 + \Phi_2) \begin{pmatrix} (1 - \Phi_2) & -\Phi_1 \\ -\Phi_1 & (1 - \Phi_2) \end{pmatrix} \right) = (1 + \Phi_2)^2 [(1 - \Phi_2)^2 - \Phi_1^2]$$

$$X_2^t V_2^{-1} X_2 = (x_1, x_2) (1 + \Phi_2) \begin{pmatrix} (1 - \Phi_2) & -\Phi_1 \\ -\Phi_1 & (1 - \Phi_2) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= (1 + \Phi_2) [(1 - \Phi_2) x_1^2 - 2\Phi_1 x_1 x_2 + (1 - \Phi_2) x_2^2]$$

La fonction de log vraisemblance exacte de processus $AR(2)$ est sous la forme suivante :

$$\begin{aligned}
 L(\theta) &= -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) + \frac{1}{2} \log((1 + \Phi_2)^2 [(1 - \Phi_2)^2 - \Phi_1^2]) \\
 &\quad - \left(\frac{(1 + \Phi_2)}{2\sigma^2} \right) ((1 - \Phi_2) x_1^2 - 2\Phi_1 x_1 x_2 + (1 - \Phi_2) x_2^2) \\
 &\quad - \sum_{t=1+p}^n \frac{(x_t - \Phi_1 x_{t-1} - \Phi_2 x_{t-2})^2}{2\sigma^2}
 \end{aligned}$$

Remarque 26 *Le log-vraisemblance du modèle $AR(p)$ est une fonction non linéaire compliquée de $\theta = (\Phi_1, \Phi_2, \dots, \Phi_p, \sigma^2)$. Il est impossible de trouver la forme analytique des estimateurs de maximum de vraisemblance . Alors la maximisation de la vraisemblance étant réalisé par des méthodes numériques (par exemple : par la méthode de Newton-Raphson, la méthode de score).*

Bibliographie

- [1] Aldrich, J. (1990). R.A. Fisher and the making of maximum likelihood 1912-1922, Statistical Science.
- [2] Azencott, R, Dacunha-Castelle, D, (1986). *Series of irregular observations*, Springer Verlag, New York.
- [3] Bourbaki, N. (2006). *Elements d'histoire des mathématiques*. Springer.
- [4] Brockwell, P.J. and Davis, R.A. *Introduction to Time Series and Forecasting*. Second Edition.
- [5] Cappe, O. Charbit, M. Moulines, E (2007). *Introduction aux séries temporelles*.
- [6] Charpentier, A. (2005). *Introduction à la théorie des processus en temps discret Modèles ARIMA et méthode Box & Jenkins*.
- [7] Davidson, R., MacKinnon, J.G. (1993). *Estimation and Inference in Econometrics*. New York, Oxford University Press.
- [8] Daudin, J.J, et Duby, C et Trécourt, P. (1996). *Analyse de Séries Chronologiques*. INA-PG, Mathématiques.
- [9] Deneire, L. (2008). *Estimation - Détection*. Notes de cours.
- [10] Galbraith, R.F, and Galbraith, J.I. (1974). *On the Inverses of Some Patterned Matrices Arising in the Theory of Stationary Time Series*. Journal of Applied Probability 11 pp63-71.
- [11] Gentle, J.E. *Matrix algebra*.
- [12] Greene, W.H. (2005). *Econométrie*. Paris, Pearson Education.
- [13] Hamilton, J.D (1954). *Time series analysis*.
- [14] Heinig, G et Rost, K. (1984). *Algebraic methods for Toeplitz-like matrices and operators*. Birkhäuser Verlag, Bâle.

- [15] Kestemont,R.M. *Sur l'évaluation de la fonction de vraisemblance du modèle ARMA stationnaire*. Université catholique de Louvain.
- [16] Lavis,D.A, and Southern, B.W. (1996). *The inverse of asymmetric banded Toeplitz matrix*.
- [17] Perrin,D. *La méthode de cardon et les imaginaires*.
- [18] Rainer von Sachs, Sébastien Van Bellegem. *Séries chronologiques*. Institut de statistique. Université catholique de Louvain.
- [19] Xiao-GuangLv and Ting-Zhu Huang. (2007). *A note on inversion of Toeplitz matrices*. University of Electronic and Technology of China.
- [20] Girard,Y. (2011).*Séries chronologiques à une et plusieurs variables, synthèse des méthodes classiques et modèles à base de copules"*. Mémoire présenté à l'université du Québec.