

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



N° Réf :.....

Centre Universitaire
Abd Elhafid Boussouf Mila

Institut des Sciences et Technologie

Département de Mathématiques et Informatique

Mémoire préparé en vue de l'obtention du diplôme de Master

EN: Mathématiques

Spécialité : Mathématiques fondamentales et appliquées

Analyse statistique des processus MSAR

Préparé par :

- Kenida Akila
- Chertioua Sihem

Devant le jury

Nom et Prénom de l'enseignant (Grade)

Daoui Amina	M.A.A
Ghezal Ahmed	M.C.B
Bouden Rabeh	M.A.A

Président
Rapporteur
Examineur

Année Universitaire : 2016/2017

Remerciements

Nous voudrions profiter de cette opportunité pour adresser nos profonde gratitude remerciement envers :

Dieu qui nous a donnée la force et le courage pour continuer et a éclaircis notre chemin et pour la réalisation de cette thèse.

Nous aimerions remercier en particulier l'encadreur M^s Ahmed Chezal pour avoir dirigé ce travail, pour son aide, ses conseils, ses encouragements, sa grande disponibilité et son ouverture d'esprit qui nous a aidés à mener bien ce travail.

Nous tenons à exprimer nos plus vifs remerciements à M^{me} Daoui Amina Professeur adjoint à CUM qui nous a fait l'honneur de présider le jury de soutenance.

Le remerciement suivant revient justement à M^s Bouden Rabeh Professeur adjoint à CUM qui a accepté de faire partie du jury de ce travail.

Enfin, nos remerciements à nos très chers parents, frères, sœurs, collègues et amis respectifs qui nous ont en courages et soutenu durant tous nos parcours.

Merci bien!

Plan de thèse

La thèse est organisée de la manière suivante:

L'étude historique des modèles autorégressifs à changements de régimes markoviens fait l'objet du chapitre 1 de cette thèse. Ils ont été introduits il y a 28 ans en économétrie par Hamilton (1989), chaque régime correspondant à un état distinct de l'économie, puis utilisé ensuite dans différents domaines d'applications. Dans un premier paragraphe nous définissons de manière générale les séries chronologiques et ces applications, puis nous proposons différents modèles pour l'étude des séries chronologiques, et parmi eux les modèles autorégressifs à changements de régimes markoviens qui jouent un rôle particulier. Ensuite nous passons à donner l'historique du processus autorégressif puis de Markov switching. En outre nous avons discuté de l'application et les domaines d'utilisations de ce modèle.

Le chapitre 2 est consacré à l'étude de certaines propriétés probabilistes du processus autorégressif à changement de régime markovien. Dans ces modèles, les paramètres peuvent dépendre d'une chaîne de Markov stationnaire non observable à espace d'état fini, discret et homogène. Ainsi, certaines questions fondamentales concernant cette classe de modèles, y compris les conditions nécessaires et suffisantes assurant l'existence de solutions ergodiques stationnaires (dans certains sens), l'existence de moments finis de n'importe quel ordre. En conséquence, on observe que la stationnarité locale du processus n'est ni suffisante ni nécessaire pour obtenir la stationnarité globale. Un certain nombre d'exemples illustratifs sont donnés pour clarifier la théorie et la variété des applications.

L'estimation du modèle MSAR est la recherche de la valeur vraie de ces paramètres. Il y a plusieurs méthodes d'estimation et parmi ces méthodes la méthode de la maximum vraisemblance. Ce problème fait l'objet du chapitre 3.

Dans l'annexe A, nous définissons la chaîne de Markov considérées dans cette thèse, aussi les différentes modes convergences qui sont utilisées dans cette thèse. Dans l'annexe B, nous avons regroupé différents processus stochastique, ces processus sont utilisés au deuxième chapitre et permettent d'étudier les propriétés probabilistes du modèle MSAR. Ensuite nous définissons la stationnarité au sens stricte et au seconde ordre, puis nous définissons la fonction d'auto-covariance et la fonction d'auto-corrélation, ainsi moment d'ordre supérieur.

Enfin, quelques définitions et notation sont fournis dans l'annexe C.

The Title

The Author

The Date

Table des matières

1	Introduction	1
2	Les propriétés probabilistes des processus $MSAR$	5
2.1	Le processus AutoRégressif (AR)	5
2.1.1	Le processus $AR_{(1)}$	5
2.1.2	Le processus $AR_{(p)}$	8
2.2	Le processus $MSAR$	10
2.2.1	Présentation du processus	10
2.2.2	La stationarité et l'existence des moments d'ordre supérieur du $MSAR$	12
3	Estimation des modèle $MSAR$ par le maximum vraisemblance	17
3.1	Modèle et notation	17
4	Appendices	19
4.1	Chaîne de Markov	19
4.1.1	Introduction	19
4.1.2	Probabilité de transition	21
4.1.3	Matrice de transition	21
4.1.4	Classification d'état	22
4.1.5	Exemples	24
4.2	Convergence des suites des variables aléatoires	27
4.2.1	Mode de convergences	28
4.3	Processus stochastique	30
4.3.1	Bruit blanc	30
4.3.2	Le processus causal	30
4.3.3	Le processus MA	31

4.3.4	Ergodicité	31
4.3.5	La stationnarité	32
4.3.6	Autocorrélations	32
4.3.7	Moment d'ordre supérieur	33
4.4	Notions général	37
4.4.1	Le rayon spectral	37
4.4.2	Produit de kronecker	37
4.4.3	La fonction de vraisemblance	38

Chapitre 1

Introduction

Depuis toujours, l'Homme a voulu prédire l'avenir, que ce soit pour prendre de meilleures décisions ou simplement pour satisfaire sa curiosité. Les premières tentatives étaient basées sur l'astrologie ou autres superstitions. Toutefois ces méthodes de prédiction ne peuvent pas être considérées comme rigoureuses aux yeux de la science. Heureusement la science a énormément évolué au cours des dernières décennies ; les méthodes de prévision qui en découlent ne font pas exception. Une méthode de prévision très populaire est basée sur l'étude rigoureuse de séries chronologiques. Cette approche permet de prédire, par exemple, de nombreux phénomènes naturels et financiers. Une série chronologique est constituée de valeurs observées à des intervalles de temps réguliers. Par exemple, les débits annuels sur un cours d'eau ou encore les valeurs mensuelles de titres boursiers sont des séries chronologiques. A la base, l'étude formelle des séries chronologiques consiste à trouver un modèle mathématique qui explique le mieux possible les données observées. A partir de ce modèle, il est possible de faire de la prévision. Cependant, la justesse des prévisions dépend fortement de la qualité du modèle choisi. Il est donc primordial de trouver des modèles qui reflètent le mieux possible la réalité afin de minimiser les erreurs de prévision.

Pour l'étude des séries chronologiques, la théorie est extrêmement bien développée. Par exemple, la méthode de **Box-Jenkins** (**Box & Jenkins**, 1970), basée sur l'unification des modèles dits autorégressifs et à moyenne mobile (*ARMA*), est très bien documentée. On peut se référer aux ouvrages de **Brockwell & Davis** (2002) et **Tsay** (2005) pour un aperçu complet. Des modèles linéaires ont également été introduits. On peut mentionner le modèle autorégressif à changements de régimes markoviens (*MSAR*) qui est composé de deux parties : d'une part, les processus autorégressifs ont été introduits par **George Udny Yule**, Yule utilise le premier modèle autorégressif pour modéliser la série chronologique du nombre de taches solaires plutôt que la méthode du périodogramme de **Schuster**. Un processus autorégressif est un processus où l'on écrit une observation au temps t comme une combinaison linéaire des observations passées plus un certain bruit blanc. D'autre part, les processus à changements de régimes markoviens (**Markov Switching**). La modélisation à changements de régimes markoviens cherche à étudier les changements de

régime intervenus sur une série. L'utilisation de modèles à changements de régimes en macroéconomie s'est largement développée à la suite d'**Hamilton** (1989) qui distingue un régime de récession et d'expansion dans la dynamique de la croissance de l'américain. Ces modèles ont été initialement introduits par **Goldfeld** et **Quandt** (1973) [18]. Et la fonction de vraisemblance qui correspond à ces modèles a été correctement calculée par **Cosslett** et **Lee** (1985). Dans leur étude, ils ont transposé l'algorithme itératif pour chaînes de Markov cachées aux combinaisons markoviennes de modèles linéaires, et après, **Hamilton** (1989) a proposé une méthode d'estimation reposant sur un algorithme de filtrage, qui permet de calculer la vraisemblance du modèle et les probabilités conditionnelles de régime.

Les modèles à changements de régimes markoviens (Markov-Switching Models ou *MSM*) ont été popularisés dans la littérature statistique par **Hamilton** (1989) [24], afin de prendre en compte un certain type de non-stationnarité présente dans de nombreuses séries économiques et financières. Ayant observé que ce type de séries présente souvent des ruptures dans leur moyenne, l'idée originale d'**Hamilton** (1989) [24] fut de modéliser cette non-stationnarité à l'aide d'un processus linéaire par morceaux. En particulier, on suppose que la série observée peut être approchée à l'aide d'un processus autorégressif dont les paramètres évoluent au cours du temps selon une variable non-observable qui suit une chaîne de Markov du premier ordre à K états. Ainsi, **Hamilton** a proposé un modèle univarié $AR_{(4)}$ dont les paramètres sont gouvernés par une chaîne de Markov à deux états pour la variation trimestrielle des Etats-Unis.

Depuis le travail de **Hamilton**, un grand nombre d'études théoriques et empiriques ont été proposées en économétrie sur ce type de modèle, en particulier dans le cas de variables qualitatives (**Grégoir** et **Lenglart**, 1998, 2000).

Ainsi, **Ferrara** et **Anas** (2002) présentent un indicateur économique coïncident permettant de détecter en temps réel les dates d'entrée et sortie de récession pour les Etats-Unis. Cet indicateur est basé sur un modèle à changement de régimes markoviens proposé par **Hamilton** (1989) et appliqué sur différentes séries représentatives du cycle classique américain, choisies de manière adéquate. Les probabilités filtrées obtenues à partir de ces séries sont combinées en tenant compte du risque de faux signaux pour fournir en sortie une probabilité instantané.

Le modèle de changement Markov original se concentre sur le comportement moyen des variables. Ce modèle et ses variantes ont été largement appliquées pour analyser les séries temporelles économiques et financières ; Voir par exemple **Hamilton** (1988, 1989) [27], **Engel** et **Hamilton** (1990) [10], **Lam** (1990), **Garcia** et **Perron** (1996) [16], **Goodwin** (1993) [19], **Diebold** [6], **Lee** et **Weinbach** (1994) [35], **Engel** (1994) [9], **Filardo** (1994) [12], **Ghysels** (1994) [17], **Sola** et **Driffill** (1994) [43], **Kim** et **Yoo** (1995) [32], parmi d'autres. Récemment, ce modèle a également été un choix populaire dans l'étude des cycles économiques de Taiwan ; Voir **Huang**, **Kuan** et **Lin** (1998), **Huang** (1999), **Chen** et **Lin** (2000), **Hsu** et **Kuan** (2001) et **Rau**, **Lin** et **Li** (2001). Etant donné que le modèle de changement de Markov de la moyenne conditionnelle est très efficace, il est naturel de considérer l'incorporation de ce mécanisme de commutation dans des modèles de variance conditionnelle. Une classe leader de modèles de variance conditionnelle est le modèle GARCH (hétérotériorité conditionnelle autorégressive généralisée) introduit par **Engle** (1982) [12] et **Bollerslev** (1986) [3]. **Cai** (1994), **Hamilton** et **Susmel** (1994) [28] et **Gray** (1996) étudient divers

modèles *ARCH* et *GARCH* avec changement de Markov. Ainsi, **Lam** et **Li** (1998) [35] introduisent également Markov switching pour passer au modèle de volatilité stochastique de de **Melino** et **Turnbull** (1990) [40]. Les autres applications financières des modèles de variance conditionnelle de commutation comprennent, entre autres, **Hamilton** et **Lin** (1996) [26], **Dueker** (1997) [8]. **Lin**, **Hung** et **Kuan** (2002) [36] appliquent également ces modèles pour analyser les séries temporelles financières de Taïwan.

Chapitre 2

Les propriétés probabilistes des processus $MSAR$

Dans ce chapitre on étudie certaines propriétés probabilistes du processus autorégressif à changement de régime markovien. Dans ces modèles, les paramètres peuvent dépendre d'une chaîne de Markov stationnaire non observable à espace d'état fini, discrète et homogène. Ainsi, certaines questions fondamentales concernant cette classe de modèles, y compris les conditions nécessaires et suffisantes assurant l'existence de solutions ergodiques stationnaires (dans certains sens), l'existence de moments finis de n'importe quel ordre. En conséquence, on observe que la stationnarité locale du processus n'est ni suffisante ni nécessaire pour obtenir la stationnarité globale. Un certain nombre d'exemples illustratifs sont donnés pour clarifier la théorie et la variété des applications.

2.1 Le processus AutoRégressif (AR)

Un processus autorégressif est un processus aléatoire utilisé pour décrire l'évolution temporelle d'un phénomène statistique suivant des lois probabilité.

2.1.1 Le processus $AR_{(1)}$

Définition 2.1.1 *On dit que le processus X_t est un processus autorégressif d'ordre 1 si on peut écrire*

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \varepsilon_t$$

Où α_0, α_1 sont les paramètres du modèle et ε_t est un bruit blanc *i.i.d.* $(0, \sigma^2)$.

Proposition 2.1.1 *Le paramètre α_1 détermine si le processus $AR_{(1)}$ est stationnaire au sens stricte ou non :*

$$|\alpha_1| : \begin{cases} |\alpha_1| < 1 & \text{Le processus est stationnaire au sens stricte.} \\ |\alpha_1| = 1 & \text{Le processus n'est pas stationnaire.} \\ |\alpha_1| > 1 & \text{Le processus est explosif.} \end{cases}$$

Preuve Soit X_t est un processus autorégressif d'ordre 1 défini par :

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \varepsilon_t$$

On a :

$$\begin{aligned} X_t - \alpha_1 X_{t-1} &= \alpha_0 + \varepsilon_t \\ X_t - \alpha_1 B X_t &= \alpha_0 + \varepsilon_t \\ (1 - \alpha_1 B) X_t &= \alpha_0 + \varepsilon_t \\ X_t &= \left(\frac{1}{1 - \alpha_1 B} \right) (\alpha_0 + \varepsilon_t) = \sum_{n \geq 0} \alpha_1^n B^n (\alpha_0 + \varepsilon_t) \\ &= \sum_{n \geq 0} \alpha_1^n \alpha_0 + \sum_{n \geq 0} \alpha_1^n B^n \varepsilon_t = \sum_{n \geq 0} \alpha_1^n \alpha_0 + \sum_{n \geq 0} \alpha_1^n \varepsilon_{t-n} \\ &= \tilde{\alpha}_0 + \sum_{n \geq 0} \alpha_1^n \varepsilon_{t-n} \end{aligned}$$

Où $\sum_{n \geq 0} \alpha_1^n$ est une série converge alors $|\alpha_1| < 1$, donc le processus est stationnaire au sens stricte.

Remarque 2.1.1 *B est un opérateur retard définit de la manière suivante :*

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

Ainsi que :

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

Ce processus admet les caractéristiques suivantes :

Proposition 2.1.2 *L'espérance¹ d'un processus X_t s'écrit sous la forme suivante :*

$$E \{X_t\} = \frac{\alpha_0}{1 - \alpha_1} \text{ si } |\alpha_1| < 1$$

Preuve Pour déterminer l'espérance d'un processus $AR_{(1)}$, il est à noter que ce processus admet l'écriture suivante :

$$X_t = (1 - \alpha_1 B)^{-1} (\alpha_0 + \varepsilon_t) = (1 + \alpha_1 B + \alpha_1^2 B^2 + \dots) (\alpha_0 + \varepsilon_t)$$

¹L'espérance est nulle si et seulement si $\alpha_0 = 0$.

Donc :

$$\begin{aligned}
E\{X_t\} &= E\{(1 + \alpha_1 B + \alpha_1^2 B^2 + \dots) \alpha_0\} + E\{(1 + \alpha_1 B + \alpha_1^2 B^2 + \dots) \varepsilon_t\} \\
&= E\{(1 + \alpha_1 + \alpha_1^2 + \dots) \alpha_0\} + E\{\varepsilon_t + \alpha_1 \varepsilon_{t-1} + \dots\} \\
&= \frac{\alpha_0}{1 - \alpha_1}
\end{aligned}$$

Car $E\{\varepsilon_t\} = 0$, et $(1 + \alpha_1 + \alpha_1^2 + \dots)$ est la somme infinie d'une suite géométrique de raison $q = \alpha_1$ et de premier terme égale à 1, donc $(1 + \alpha_1 + \alpha_1^2 + \dots) = \frac{1}{1 - \alpha_1}$.

Proposition 2.1.3 *La variance d'un processus X_t s'écrit sous la forme suivante :*

$$Var\{X_t\} = \frac{\sigma^2}{1 - \alpha_1^2} \text{ si } |\alpha_1^2| < 1.$$

Preuve On sait que : $Var\{X_t\} = E\{X_t - E\{X_t\}\}^2 = \gamma(0)$.

Or,

$$\begin{aligned}
X_t - E\{X_t\} &= (1 - \alpha_1 B)^{-1} (\alpha_0 + \varepsilon_t) - \frac{\alpha_0}{1 - \alpha_1} \\
&= \frac{\alpha_0}{1 - \alpha_1} + (\varepsilon_t + \alpha_1 \varepsilon_{t-1} + \alpha_1^2 \varepsilon_{t-2} + \dots) - \frac{\alpha_0}{1 - \alpha_1} \\
&= \sum_{n \geq 0} \alpha_1^n \varepsilon_{t-n}
\end{aligned}$$

D'où $E\{X_t - E\{X_t\}\}^2 = E\{\varepsilon_t + \alpha_1 \varepsilon_{t-1} + \alpha_1^2 \varepsilon_{t-2} + \dots\}^2$, et puisque $E\{\varepsilon_t \varepsilon_s\} = 0$, $\forall t \neq s$, on aura :

$$\begin{aligned}
\gamma(0) &= E\{\varepsilon_t^2\} + \alpha_1^2 E\{\varepsilon_{t-1}^2\} + \alpha_1^4 E\{\varepsilon_{t-2}^2\} + \dots \\
&= (1 + \alpha_1^2 + \alpha_1^4 + \alpha_1^6 + \dots) \sigma^2. \\
&= \frac{\sigma^2}{1 - \alpha_1^2} \text{ si } |\alpha_1^2| < 1
\end{aligned}$$

Proposition 2.1.4 *La fonction d'auto covariance d'un processus X_t s'écrit sous la forme suivante :*

$$Cov\{X_t, X_{t-k}\} = \frac{\alpha_1^k}{1 - \alpha_1^2} \sigma^2 = \alpha_1^k \gamma(0) \text{ si } |\alpha_1^2| < 1$$

Preuve On peut déterminer la fonction d'auto covariance en raisonnant de la même manière. La fonction d'auto covariance est définie par :

$$\gamma(k) = E\{(X_t - E(X_t))(X_{t-k} - E(X_{t-k}))\} \text{ pour } k > 0$$

En utilisant les résultats suivants :

$$\begin{aligned}
X_t - E(X_t) &= \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \alpha_1^2 \varepsilon_{t-2} + \dots \\
X_{t-k} - E(X_{t-k}) &= \varepsilon_{t-k} + \alpha_1 \varepsilon_{t-k-1} + \alpha_1^2 \varepsilon_{t-k-2} + \dots
\end{aligned}$$

On trouve :

$$\begin{aligned}
 \gamma(k) &= E \left\{ (\varepsilon_t + \alpha_1 \varepsilon_{t-1} + \alpha_1^2 \varepsilon_{t-2} + \dots + \alpha_1^k \varepsilon_{t-k} + \alpha_1^{k+1} \varepsilon_{t-k-1} + \dots) \times (\varepsilon_{t-k} + \alpha_1 \varepsilon_{t-k-1} + \alpha_1^2 \varepsilon_{t-k-2} + \dots) \right\} \\
 &= \alpha_1^k E \{ \varepsilon_{t-k}^2 \} + \alpha_1^{k+2} E \{ \varepsilon_{t-k-1}^2 \} + \alpha_1^{k+4} E \{ \varepsilon_{t-k-2}^2 \} + \dots \\
 &= \alpha_1^k (1 + \alpha_1^2 + \alpha_1^4 + \alpha_1^6 + \dots) \sigma^2 \\
 &= \frac{\alpha_1^k}{1 - \alpha_1^2} \sigma^2
 \end{aligned}$$

Corollaire 2.1.1 D'après les propositions (2.1.2), (2.1.3) et (2.1.4) la stationnarité au seconde ordre est vérifiée.

Proposition 2.1.5 La fonction d'autocorrélation d'un processus X_t s'écrit sous la forme suivante :

$$\text{Cor} \{X_t, X_{t-k}\} = \alpha_1^k$$

Exemple 2.1.1 Soit X_t un processus $AR_{(1)}$ défini par :

$$X_t = 0.7X_{t-1} + \varepsilon_t$$

$$X_t = -0.7X_{t-1} + \varepsilon_t.$$

Dans les graphiques suivants, on représente les fonctions d'autocorrélation de ces processus :

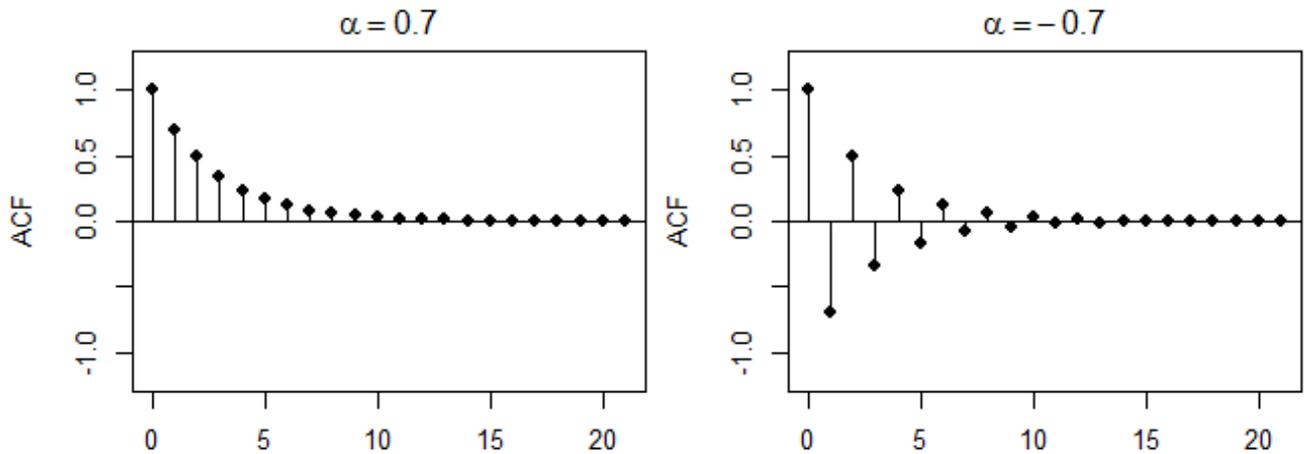


FIGURE 1 : ACF processus $AR_{(1)}$ en fonction de α .

2.1.2 Le processus $AR_{(p)}$

Définition 2.1.2 On dit que le processus X_t est un processus autorégressif d'ordre p si on peut écrire :

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t$$

Ou encore :

$$(1 - \alpha_1 B - \alpha_2 B^2 - \alpha_3 B^3 - \dots - \alpha_p B^p) X_t = \alpha_0 + \varepsilon_t$$

Où $\alpha(B) = (1 - \alpha_1 B - \alpha_2 B^2 - \alpha_3 B^3 - \dots - \alpha_p B^p)$ est un polynôme caractéristique.

Un processus $AR_{(p)}$ est stationnaire si le module des solutions (les racines) de son équation caractéristique est à chaque fois strictement inférieur à 1 en valeur absolue.

Remarque 2.1.2 L'équation caractéristique se trouve très facilement depuis le polynôme caractéristique en substituant à l'opérateur des retards B par la variable x . Pour le processus $(AR_{(p)})$:

$$(1 - \alpha_1 B - \alpha_2 B^2 - \alpha_3 B^3 - \dots - \alpha_p B^p) \text{ devient } (1 - \alpha_1 x - \alpha_2 x^2 - \alpha_3 x^3 - \dots - \alpha_p x^p)$$

On peut montrer que cette condition est équivalente à :

$$\alpha_1 + \alpha_2 + \alpha_3 + \dots + \alpha_p < 1 \text{ et } |\alpha_p| < 1$$

Les différents moments d'un processus autorégressif d'ordre p stationnaire sont :

Proposition 2.1.6 L'espérance d'un processus X_t autorégressif d'ordre p s'écrit sous la forme suivante :

$$E\{X_t\} = \frac{\alpha_0}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_p}$$

Preuve On a :

$$\begin{aligned} E\{X_t\} &= \alpha_0 + \alpha_1 E\{X_{t-1}\} + \alpha_2 E\{X_{t-2}\} + \dots + \alpha_p E\{X_{t-p}\} + E\{\varepsilon_t\} \\ &= \alpha_0 + (\alpha_1 + \alpha_2 + \dots + \alpha_p) E\{X_t\}, \text{ car } E\{X_t\} = E\{X_{t-p}\} \text{ et } E\{\varepsilon_t\} = 0 \\ &= \frac{\alpha_0}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_p} \end{aligned}$$

Proposition 2.1.7 La variance d'un processus X_t autorégressif d'ordre p s'écrit sous la forme suivante :

$$\text{Var}\{X_t\} = \alpha_1 \gamma_1 + \alpha_2 \gamma_2 + \dots + \alpha_p \gamma_p + \sigma^2$$

Proposition 2.1.8 La fonction d'auto covariance d'un processus X_t autorégressif d'ordre p s'écrit sous la forme suivante :

$$\gamma_k = \alpha_1 \gamma_{k-1} + \alpha_2 \gamma_{k-2} + \dots + \alpha_p \gamma_{k-p}, \text{ pour } k = 1, 2, \dots$$

Proposition 2.1.9 La fonction d'autocorrélation d'un processus X_t autorégressif d'ordre p noté ρ_k est donnée sous la forme suivante :

$$\rho_k = \begin{cases} 1 & \text{si } k = 0. \\ \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2} + \dots + \alpha_p \rho_{k-p} & \text{si } k \in \mathbb{Z}^* \end{cases}$$

2.2 Le processus MSAR

Depuis les travaux séminiques de **Hamilton** les modèles de changement de Markov ont suscité un intérêt croissant et deviennent un outil puissant de modélisation et de description des cycles économiques asymétriques (comme proposé initialement par **Hamilton** [25]) et continuent à gagner en popularité en particulier dans les données financières

Ceci est dû à sa plus grande souplesse dans la capture de la persistance et/ou des effets asymétriques sur les chocs de volatilité et leur capacité à modéliser des séries chronologiques qui se caractérisent par certaines caractéristiques, y compris l'excès de kurtosis, d'asymétrie, de rotation ou de changement brusque de régime.

Un *MSM* à temps discret est un processus aléatoire bivarié $((X_t, s_t), t \in \mathbb{Z})$, $\mathbb{Z} := \{0, \pm 1, \pm 2, \dots\}$ telle que : s_t est une chaîne de Markov. Ainsi, les changements dans les régimes peuvent être abrupts, et ils se produisent fréquemment ou parfois selon la probabilité de transition de la chaîne. Cependant, certains modèles linéaires ou non linéaires locaux (c'est-à-dire dans chaque «régime») ont été étudiés afin de capturer le probabiliste et les propriétés statistiques de ces modèles. Par exemple, *MS-ARMA* : **Francq** et **Zakoïban** [14] et **Stelzer** [42], *ARMA MS*—non linéaire et processus bilinéaires : **Lee** [35], **Yao** et **Attali** [45] et **Bibi** et **Aknouche** [1], *MS-GARCH* : **FZ** [15], **Hass** et **al** [29], **Liu** [37] parmi d'autres.

Dans cette thèse, nous proposons alternativement un modèle *MSAR* dans lequel le processus suit localement une représentation linéaire.

2.2.1 Présentation du processus

Un processus autorégressif à changements de régimes à temps discret est un processus aléatoire bivarié

$$((X_t, s_t), t \in \mathbb{Z})$$

$\mathbb{Z} := \{0, \pm 1, \pm 2, \dots\}$ telle que :

- (i) $(s_t, t \in \mathbb{Z})$ est une chaîne de Markov non observable à espace d'état fini, discrète et homogène.
- (ii) La distribution conditionnelle de (X_k) donnée $\{(X_{t-1}, s_k), t \leq k\}$ dépend uniquement de $\{(X_{t-1}, s_k), t \leq k\}$.

Ainsi, les changements dans les régimes peuvent être brusques, et ils se produisent fréquemment ou occasionnellement en fonction de la probabilité de transition de la chaîne.

Dans cette thèse, on propose alternativement un modèle *MSAR* dans lequel le processus suit localement une représentation autorégressive. On dit qu'un processus de \mathbb{R} -valeur $(X_t, t \in \mathbb{Z})$ définie sur un espace de probabilité $(\Omega, \mathfrak{S}, P)$ a une représentation autorégressif à changement de Markov (défini par *MSAR*_(p)), s'il est un solution de l'équation stochastique suivante :

$$X_t = \alpha_0(s_t) + \sum_{i=1}^p \alpha_i(s_t) X_{t-i} + \varepsilon_t(s_t), t \in \mathbb{Z} \quad (2.1)$$

Dans (2.1), $(\varepsilon_t, t \in \mathbb{Z})$ est une séquence indépendante et identiquement distribuée de variables aléatoires définies sur le même espace de probabilité $(\Omega, \mathfrak{S}, P)$, avec :

$$E \{ \log^+ |\varepsilon_t| \} < +\infty \text{ où } \log^+ x = \max \{ 0, \log x \}, x > 0$$

Les fonctions $\alpha_i(\cdot)$ dépendent d'une chaîne de Markov non observée $(s_t, t \in \mathbb{Z})$ sous réserve de l'hypothèse suivante :

- La chaîne de Markov $(s_t, t \in \mathbb{Z})$ est stationnaire, irréductible, apériodique, à espace d'état fini $\mathbb{S} = \{1, 2, \dots, d\}$ (donc ergodique), de matrice de transition à n -étapes :

$$\mathbb{P}^n = \left(p_{ij}^{(n)}, (i, j) \in \mathbb{S} \times \mathbb{S} \right) \text{ où } p_{ij}^{(n)} = P(s_t = j | s_{t-n} = i)$$

avec la matrice de transition en une étape :

$$\mathbb{P} = (p_{ij}, (i, j) \in \mathbb{S} \times \mathbb{S}) \text{ où } p_{ij} = P(s_t = j | s_{t-1} = i)$$

et probabilité initiale stationnaire :

$$\underline{\Pi} = (\pi(1), \dots, \pi(d))', \text{ où } \pi(i) = P(s_t = i), i = 1, \dots, d \text{ telle que : } \underline{\Pi}' = \underline{\Pi}' \mathbb{P}.$$

En outre, on suppose que ε_t et $\{X_{s-1}, s_t, s \leq t\}$ sont indépendantes.

Le but principal de cette thèse est d'étudier certaines propriétés théoriques pour le processus *MSAR*, ainsi et pour le but statistique, il est souvent souhaitable dans la pratique que les procédés de solutions $(X_t, t \in \mathbb{Z})$ pour (2.1) soient stationnaires, ergodiques et satisfaisant $X_t = f(\varepsilon_t, s_t, \varepsilon_{t-1}, s_{t-1}, \dots)$ presque sûrement (*p.s*) où f est une fonction mesurable de \mathbb{R}^∞ à \mathbb{R} .

Certaines notations sont utilisées dans cette thèse :

- $I_{(n)}$ est la matrice identité d'ordre $n \times n$, $\mathbb{I}'_{(n)} := \left(\underbrace{I_{(n)} \dot{\vdots} \dots \dot{\vdots} I_{(n)}}_{d\text{-block}} \right)$ est la matrice d'ordre $n \times nd$ et \mathbb{I}_Δ est la fonction indicatrice de l'ensemble Δ .
- $O_{(k,l)}$ désigne la matrice d'ordre $k \times l$ dont les entrées sont des zéros, pour plus de simplicité, on a : $O_{(k)} := O_{(k,k)}$ et $\underline{O}_{(k)} := \underline{O}_{(k,1)}$, le rayon spectral de la matrice carrée M est noté par $\rho(M)$.
- Soit $\|\cdot\|$ désigne toute norme d'opérateur sur l'ensemble des $m \times n$ et $m \times 1$ matrices, et pour $\gamma \in]0, 1]$, soit $|M|^\gamma := (|m_{ij}|^\gamma)$, puis il est facile de voir que $|\cdot|^\gamma$ est sous-multiplicateur, i.e. $|M_1 M_2|^\gamma \leq |M_1|^\gamma |M_2|^\gamma$, $|MX|^\gamma \leq |M|^\gamma |X|^\gamma$ pour tout vecteur X également sous-additif, i.e. $\left| \sum_i M_i \right|^\gamma \leq \sum_i |M_i|^\gamma$, où l'inégalité $M < N$ désigne la relation élémentaire $m_{ij} < n_{ij}$ pour tout i et j .
- \otimes est le produit de Kronecker des matrices et $M^{\otimes r} = M \otimes M \otimes \dots \otimes M$, r fois. Si $(M(i), i \in \mathbb{N})$ est une matrice séquence d'ordre $n \times n$, on indique pour tout entier l et j , $\prod_{i=l}^j M(i) = M(l) M(l+1) \dots M(j)$ si $l \leq j$ et $l_{(n)}$ d'une autre manière.

– Lorsque les matrices $\underline{M} = (M(i), i \in \mathbb{S})$ sont une suite de matrices non aléatoires, on indique :

$$\mathbb{P}(M) = \begin{pmatrix} p_{11}M(1), & \dots & p_{d1}M(1) \\ \vdots & \dots & \vdots \\ p_{1d}M(d) & \dots & p_{dd}M(d) \end{pmatrix}, \quad \underline{\Pi}(M) = \begin{pmatrix} \pi(1)M(1) \\ \vdots \\ \pi(d)M(d) \end{pmatrix}$$

La prochaine sous-section suivante fournit une représentation de l'espace d'état qui est utilisée pour dériver des conditions suffisantes pour que le processus *MSAR* généré par l'équation (2.1) ait une solution stationnaire unique (au sens fort et faible), causale et ergodique ayant des moments d'ordre supérieur.

2.2.2 La stationnarité et l'existence des moments d'ordre supérieur du *MSAR*

Pour simplicité les notions, on définit le p -vecteur tel que : $\underline{H}' := (1, \underline{Q}'_{(p-1)})$, $\underline{X}'_t := (X_t, \dots, X_{t-p+1})$ et la matrice d'ordre $p \times p$: $(A(s_t), t \in \mathbb{Z})$ tel que : $A(s_t) = \begin{pmatrix} \alpha_1(s_t) & \dots & \alpha_p(s_t) \\ I_{(p-1)} & & \underline{Q}_{(p-1)} \end{pmatrix}_{(p \times p)}$, donc l'équation (2.1) peut s'écrire sous la forme suivante :

$$\underline{X}_t = A(s_t) \underline{X}_{t-1} + \underline{\varepsilon}_t(s_t), \forall t \in \mathbb{Z} \quad (2.2)$$

et $X_t = \underline{H}' \underline{X}_t$ pour tout $t \in \mathbb{Z}$.

Stationnarité au sens stricte

Comme $((s_t, \varepsilon_t), t \in \mathbb{Z})$ est un processus de Markov stationnaire et ergodique avec un espace d'état $\mathbb{S} \times \mathbb{R}$, alors $(A(s_t), t \in \mathbb{Z})$ est aussi un processus stationnaire et ergodique avec $E \{\log^+ \|\underline{\varepsilon}_t(s_t)\|\}$ et $E \{\log^+ \|A(s_t)\|\}$ sont finis. En définissant $A_t(k) = \prod_{j=0}^{k-1} A(s_{t-j})$, il résulte de **Brandt** (Voir **Bougerol** et **Picard**) que la solution unique, causale, bornée en probabilité, strictement stationnaire et ergodique de (2.2) est donnée par :

$$\underline{X}_t = \sum_{k \geq 1} A_t(k) \underline{\varepsilon}_{t-k} + \underline{\varepsilon}_t(s_t), \forall t \in \mathbb{Z} \quad (2.3)$$

Chaque fois que l'exposant de Lyapunov $\gamma_L(A)$, associé à la séquence des matrices aléatoires stationnaires et ergodiques $A = (A(s_t), t \in \mathbb{Z})$ définis par :

$$\gamma_L(A) := \inf E \left\{ \frac{1}{k} \log \|A_t(k)\|, k \geq 1 \right\} \stackrel{p.s.}{=} \lim_{t \rightarrow \infty} \frac{1}{k} \log \|A_t(k)\| \quad (2.4)$$

est strictement négative. La deuxième égalité dans (2.4) peut être justifiée en utilisant le théorème ergodique sous-additif de **Kingman** et l'existence de $\gamma_L(A)$ est toutefois garantie par le fait que $E \{\log^+ \|A(s_t)\|\} \leq E \{\|A(s_t)\|\} < +\infty$.

Théorème 2.2.1 *Considérons le modèle (2.1) avec la représentation vectorielle (2.2), alors :*

$\gamma_L(A) < 0$ est une condition suffisante pour (2.2) admet une solution unique, strictement stationnaire, causale et ergodique, donnée par la série (2.3) qui converge absolument presque sûrement pour tous $t \in \mathbb{Z}$.

Exemple 2.2.1 [Non-nécessité de la stationnarité locale] considérons le modèle MSAR à première ordre :

$$X_t = \alpha_0(s_t) + \alpha_1(s_t) X_{t-1} + \varepsilon_t(s_t), \forall t \in \mathbb{Z}$$

La condition suffisante pour l'existence d'une solution strictement stationnaire se réduit à $\sum_{j=1}^d \pi(i) E \{ \log |\alpha_1(i)| \} < 0$.

Ce exemple montre que la stationnarité locale stricte n'est pas nécessaire, c'est-à-dire l'existence de régimes explosifs ($E \{ \log |\alpha_1(i)| \} > 0$) n'empêche pas les systèmes stationnaire strict.

Proposition 2.2.1 Considérons le modèle $MS - AR_{(p)}$ et on suppose que $E \{ |\varepsilon_0|^\delta \} < \infty$, pour certains $0 < \delta < 1$. Soit $\underline{M}_\delta := \left\{ E \left\{ |A(i)|^\delta \right\}, 1 < i < d \right\}$. Ensuite $\rho(\mathbb{P}(\underline{M}_\delta)) < 1$ implique que $\gamma_L(A) = \gamma_L(M) < 0$ et donc $MS - AR_{(p)}$ a une solution unique, strictement stationnaire et ergodique.

Preuve Comme l'exposant de Lyapunov est indépendante de la norme, en choisissant une norme absolue, i.e. la norme $\|\cdot\|$ telle que : $\|N\|^\delta \leq \| \|N\| \|^{\delta}$ (e.g. $\|N\| = \sum_{i,j} |n_{i,j}|$). Par suite, puisque $\rho(\mathbb{P}(\underline{M}_\delta)) < 1$, il existe $0 < \lambda < 1$ tel que : $\limsup_t \|\mathbb{P}^t(\underline{M}_\delta)\|^{1/t} < \lambda$. D'après l'inégalité de **Jensen** et la sous-multiplication de l'opérateur $|\cdot|^\delta$, on obtient :

$$\begin{aligned} \gamma_L(A)\delta &= \lim_{t \rightarrow +\infty} \frac{1}{t} E \left\{ \log \left\| \prod_{j=0}^{t-1} (A(s_{t-j})) \right\|^\delta \right\} \\ &\leq \lim_{n \rightarrow +\infty} \frac{1}{t} \log E \left\{ \left\| \prod_{j=0}^{t-1} (A(s_{t-j})) \right\|^\delta \right\} \\ &\leq \lim_{n \rightarrow +\infty} \frac{1}{t} \log \left\{ \left\| E \left\{ \prod_{j=0}^{t-1} |A(s_{t-j})|^\delta \right\} \right\| \right\} \\ &\leq \lim_{t \rightarrow +\infty} \sup \log \|\mathbb{P}^t(\underline{M}_\delta)\|^{1/t} < 0 \end{aligned}$$

La stationnarité au seconde ordre

Le problème de trouver des conditions assurant l'existence d'une stationnarité de second ordre de MSAR a été étudié par **Françq et al.**

Théorème 2.2.2 Considérons le modèle (2.2), dans lequel $E \{ \varepsilon_t \} = 0 = E \{ e_t^3 \}$ et $\kappa_4 := E \{ e_t^4 \} < +\infty$.

Si :

$$\lambda_{(2)} := \rho(\mathbb{P}(\underline{A}^{(2)})) < 1 \text{ où } \underline{A}^{(2)} := \left(A^{(2)}(i) = A^{\otimes 2}(i), 1 \leq i \leq d \right) \quad (2.5)$$

Alors l'équation (2.2) a une solution unique, causal, et stationnaire au seconde ordre donnée par :

$$\underline{X}_t = \sum_{k \geq 0} A_t(k) \varepsilon_{t-k} + \varepsilon_t \quad (2.6)$$

En outre, la solution est strictement stationnaire et ergodique.

Preuve Pour vérifier que l'équation (2.6) est bien définie en \mathbb{L}_2 , il suffit de Montre que :

$$\underline{X}_t(k) = \left\{ \prod_{j=0}^{k-1} A(s_{t-j}) \right\} \varepsilon_{t-k} \quad (2.7)$$

converge vers $\underline{Q}_{(p)}$ dans \mathbb{L}_2 quand $k \rightarrow +\infty$. De l'indépendance entre s_t et ε_t , on montre par récurrence que $E \{ \underline{X}_t^{\otimes 2}(k) \} = \mathbb{I}'_{(d^2)} \mathbb{P}^k \left(\underline{A}^{(2)} \right) \underline{\Pi} \left(\underline{\Sigma}^{(2)} \right)$, alors on a pour $k = 2$:

$$\begin{aligned} \underline{X}_t^{\otimes 2}(k) &= (A(s_t) A(s_{t-1}))^{\otimes 2} (\varepsilon_{t-2})^{\otimes 2} \\ &= A^{\otimes 2}(s_t) A^{\otimes 2}(s_{t-1}) \varepsilon_{t-2}^{\otimes 2} \end{aligned}$$

Donc :

$$\begin{aligned} E \{ \underline{X}_t^{\otimes 2}(k) \} &= \sum_{j=1}^d E \{ \underline{X}_t^{\otimes 2}(2) | s_t = j \} \pi(j) \\ &= \sum_{j=1}^d A^{\otimes 2}(j) E \{ A^{\otimes 2}(s_{t-1}) \varepsilon_{t-2}^{\otimes 2} | s_t = j \} \pi(j) \\ &= \sum_{j=1}^d A^{\otimes 2}(j) \sum_{i=1}^d A^{\otimes 2}(i) E \{ \varepsilon_{t-2}^{\otimes 2} | s_{t-1} = i \} \pi(j) P(s_{t-1} = i | s_t = j) \\ &= \sum_{j=1}^d A^{\otimes 2}(j) \sum_{i=1}^d A^{\otimes 2}(i) \sum_{l=1}^d E \{ \varepsilon_{t-2}^{\otimes 2} | s_{t-2} = l \} \pi(j) P(s_{t-2} = l | s_{t-1} = i) P(s_{t-1} = i | s_t = j) \\ &= \sum_{j=1}^d A^{\otimes 2}(j) \sum_{i=1}^d A^{\otimes 2}(i) p_{ij} \sum_{l=1}^d \underline{\Sigma}^{(2)}(l) p_{li} \pi(l) \end{aligned}$$

On trouve que : $E \{ \underline{X}_t^{\otimes 2}(k) \} = \mathbb{I}'_{(s^2)} \mathbb{P}^k \left(\underline{A}^{(2)} \right) \underline{\Pi} \left(\underline{\Sigma}^{(2)} \right)$. Alors, sous la condition (2.5) et par la décomposition de **Jordan**, on obtient :

$$\| \underline{X}_t(k) \|_{\mathbb{L}_2} \leq \left\| \mathbb{I}'_{(s^2)} \mathbb{P}^k \left(\underline{A}^{(2)} \right) \underline{\Pi} \left(\underline{\Sigma}^{(2)} \right) \right\|^{1/2} \leq \left\| \mathbb{I}'_{(s^2)} \right\|^{1/2} \left\| \mathbb{P}^k \left(\underline{A}^{(2)} \right) \right\|^{1/2} \left\| \underline{\Pi} \left(\underline{\Sigma}^{(2)} \right) \right\|^{1/2} \leq K \lambda_{(2)}^{k/2}$$

Donc $\underline{X}_t(k)$ converge vers $\underline{Q}_{(p)}$ dans \mathbb{L}_2 quand $k \rightarrow +\infty$. Par conséquent, l'équation (2.6) est la solution unique de (2.2) qui converge en \mathbb{L}_2 .

Exemple 2.2.2 Soit le processus :

$$\underline{X}_t = A(s_t) \underline{X}_{t-1} + \varepsilon_t$$

On a :

$$\underline{X}_t^{\otimes 2} = A^{\otimes 2}(s_t)\underline{X}_{t-1}^{\otimes 2} + \varepsilon_t^{\otimes 2} + (A(s_t)\underline{X}_{t-1} \otimes \varepsilon_t + \varepsilon_t \otimes A(s_t)\underline{X}_{t-1})$$

Alors :

$$E \{ \underline{X}_t^{\otimes 2} \} = \sum_{i=1}^d E \{ \underline{X}_t^{\otimes 2} | s_t = i \} \pi(i)$$

On note : $E \{ \underline{X}_t^{\otimes 2} | s_t = i \} = \underline{v}(i)$ donc :

$$E \{ \underline{X}_t^{\otimes 2} \} = \sum_{i=1}^d \pi(i) \underline{v}(i) = \mathbb{I}'_{(d^2)} \underline{\Pi}(V) = (I_d, \dots, I_d) \begin{pmatrix} \pi(1)\underline{v}(1) \\ \vdots \\ \pi(d)\underline{v}(d) \end{pmatrix}$$

D'autre part, pour tout i on a :

$$\begin{aligned} \pi(i)\underline{v}(i) &= \pi(i) E \{ A^{\otimes 2}(i)\underline{X}_{t-1}^{\otimes 2} + \varepsilon_t^{\otimes 2} \} \\ &= \pi(i) A^{\otimes 2}(i) E \{ \underline{X}_{t-1}^{\otimes 2} | s_t = i \} + \pi(i) E \{ \varepsilon_t^{\otimes 2} \} \end{aligned}$$

Par suite :

$$\begin{aligned} E \{ \underline{X}_{t-1}^{\otimes 2} | s_t = i \} &= \sum_{j=1}^d E \{ \underline{X}_{t-1}^{\otimes 2} | s_{t-1} = j \} P(s_{t-1} = j | s_t = i) \\ &= \sum_{j=1}^d E \{ \underline{X}_{t-1}^{\otimes 2} | s_{t-1} = j \} \frac{p_{ji}\pi(j)}{\pi(i)} \end{aligned}$$

Alors :

$$\pi(i)\underline{v}(i) = A^{\otimes 2}(i) \sum_{j=1}^d p_{ji}\pi(j)\underline{v}(j) + \pi(i)\underline{\Sigma}^{(2)}(i)$$

Et on trouve que : $E \{ \underline{X}_t^{\otimes 2} \} = \left(I_{(dp^2)} - \mathbb{P} \left(\underline{A}^{(2)} \right) \right)^{-1} \underline{\Pi} \left(\underline{\Sigma}^{(2)} \right)$.

Existence des moments d'ordre supérieur

Dans cette sous-section, on sera intéressé par des conditions garantissant l'existence de moments d'ordre supérieur pour des processus MSAR strictement stationnaires.

Théorème 2.2.3 *Considérons le modèle (2.2). Pour tout entier positif m , on suppose que $E \{ \varepsilon_t^{m+2} \} < +\infty$ et :*

$$\lambda_{(m)} := \rho(\mathbb{P}(\underline{A}^m)) < 1, \text{ où } \underline{A}^m := (E \{ A_{s_t=i}^{\otimes m}(\varepsilon_t) \}, 1 \leq i \leq d) \quad (2.8)$$

Ensuite, le processus $(\underline{X}_t, t \in \mathbb{Z})$ défini par (2.2) a une solution unique, causal, ergodique et strictement stationnaire donnée par (2.7) et satisfait $E \{ |\underline{X}_t^{\otimes m}| \} < +\infty$.

Preuve On sait que :

$$E \{ \underline{X}_t^{\otimes m}(k) \} = \mathbb{I}'_{(d^m)} \mathbb{P}^k \left(\underline{A}^{(m)} \right) \mathbb{II} \left(\underline{\Sigma}^{(m)} \right)$$

Donc :

$$\| \underline{X}_t^{\otimes m}(k) \|_{\mathbb{L}_m} \leq \| \mathbb{I}'_{(d^m)} \|^{1/m} \| \mathbb{P}^k \left(\underline{A}^{(m)} \right) \|^{1/m} \| \mathbb{II} \left(\underline{\Sigma}^{(m)} \right) \|^{1/m} \quad (2.9)$$

Alors, d'après la condition (2.8) et d'après la décomposition de **Jordan**, $\mathbb{P}^k \left(\underline{A}^{(m)} \right)$ converge vers 0 quand $k \rightarrow +\infty$.

Par conséquent, pour tout t , $\sum_{k \geq 0} \underline{X}_t(k)$ converge vers \underline{X}_t quand $n \rightarrow +\infty$ et le reste est immédiat.

Exemple 2.2.3 Dans le tableau suivant, on résumé les conditions suffisantes pour l'existence de $E \{ X_t^m \}$ dans certains modèles particuliers.

Spécifications	Condition (2.8)	Cas spéciaux ($p = 1$)
Standard ($d = 1$)	$\lambda_{(m)} = \rho \left(E \{ A^{(m)}(\varepsilon_t) \} \right) < 1$	$E \{ \alpha_1^m \} < 1$
Indépendente-switching	$\lambda_{(m)} = \rho \left(E \{ A_{s_t}^{(m)}(\varepsilon_t) \} \right) < 1$	$E \{ \alpha_1^m(s_t) \} < 1$

Chapitre 3

Estimation des modèle $MSAR$ par le maximum vraisemblance

L'estimation du processus général (2.1) est assez difficile même si $d = 1$. Ainsi, dans la littérature, de nombreuses idées ont été établies pour estimer les paramètres de certains processus stationnaires et ergodiques. Les méthodes les plus fréquemment utilisées sont (généralisée) la méthode des moments (G) (MM) et (conditionnelle) la méthode des moindres carrés (C) (MC). Les propriétés asymptotiques des estimations (G) MM et (C) MC ont également été discutées sous certaines restrictions, (voir par exemple **Pham et Tran** (1981) [41], **Guégan et Pham** (1989) [23], **Liu** (1990) [39], **Kim et Billard** (1990) [32], **Grahn** (1995) [20], **Wittwer** (1989) [44] et entre autres).

3.1 Modèle et notation

Dans cette thèse, on va étudier l'estimation de $MSAR_{(p)}$:

$$X_t = \alpha_0(s_t) + \sum_{i=1}^p \alpha_i(s_t) X_{t-i} + \varepsilon_t \quad (3.1)$$

Dans (3.1), le processus $(\varepsilon_t, t \in \mathbb{Z})$ est i.i.d.(0.1), et le nombre des régimes d sont supposés connus et fixé, les paramètres inconnu $\underline{\theta} = (\underline{\alpha}', \underline{p}', \underline{\pi}')$, et ses vraies valeurs noté $\underline{\theta}^0$ appartient à un espace de paramètres $\Theta \subset \mathbb{R}^R$ où $\underline{\alpha} = (\underline{\alpha}_i', 1 \leq i \leq d)'$, $\underline{p} = (\underline{p}_i', 1 \leq i \leq d)'$, $\underline{\pi} = (\pi(1), \dots, \pi(d))'$ avec des projections de coordonnées de vecteurs $\underline{\alpha}_i := (\alpha_0(i), \dots, \alpha_p(i))'$, $\underline{p}_i := (p_{ij}, \dots, p_{id}, i \neq j)'$ (grâce à les contraintes $\sum_{j=1}^d p_{ij} = 1$ pour tout i).

Pour tous entiers a et b , tel que $a \leq b$, soit $\underline{X}_{a:b}$ (resp. $\underline{\chi}_{a:b}$) désigne l'ensemble $\{X_a, X_{a+1}, \dots, X_b\}$

(resp. $\{(X_a, \varepsilon_a), (X_{a+1}, \varepsilon_{a+1}), \dots, (X_b, \varepsilon_b)\}$) avec peut-être $a = -\infty$ dans ce cas on noter \underline{X}_b (resp. $\underline{\chi}_b$). Le problème d'intérêt dans cette section est l'estimation du vecteur de paramètres $\underline{\theta}$ dominant dans (3.1) à partir d'une

séquence observée $\underline{X}_{1:n}$. Pour cela, ont désigné la fonction de densité des observations par : $g_{\underline{\theta}}(\cdot)$ et celui de ε_t par : $f_{\underline{\theta}}(\cdot)$ et ont utilisé $p_{\underline{\theta}}(\cdot, \dots, \cdot)$ pour désigner la densité par rapport à la mesure de probabilité sur $\mathfrak{B}_{\mathbb{S}^n \otimes \mathbb{R}^n}$. La densité conditionnelle correspondante donnée $s_1 = x_1$ et $\underline{\chi}_{1-p_0:0}$ avec $p_0 = p + 1$ est :

$$p_{\underline{\theta}}(\underline{X}_{1:n} | \underline{\chi}_{1-p_0:0}) = \pi(x_1) \left\{ \prod_{t=2}^n p_{x_{t-1}, x_t} \right\} \left\{ \prod_{t=1}^n g_{\underline{\theta}_{x_t}}(X_t | \underline{\chi}_{1-p_0:t-1}) \right\} \quad (3.2)$$

La vraisemblance $L_n(\underline{\theta})$ par rapport à la mesure $\lambda^n \otimes \mu^n$ (où λ Indique la mesure de Lebesgue et μ est le compte sur \mathbb{S}) que on travailler est donné en additionnant sur tous les chemins possibles de la chaîne de Markov la densité conditionnelle $p_{\underline{\theta}}(\underline{X}_{1:n} | \underline{\chi}_{1-p_0:0})$,

$$L_n(\underline{\theta}) = \sum_{(x_1, \dots, x_n) \subset \mathbb{S}^n} \pi(x_1) g_{\underline{\theta}_{x_1}}(X_1 | \underline{\chi}_{1-p_0:0}) \prod_{t=2}^n p_{x_{t-1}, x_t} g_{\underline{\theta}_{x_t}}(X_t | \underline{\chi}_{1-p_0:t-1}) \quad (3.3)$$

qui peut être réécrit en tant que produit de matrices en raison de cette structure simplifiée :

$$L_n(\underline{\theta}) = \underline{1}' \left\{ \prod_{t=2}^n \mathbb{P}(g_{\underline{\theta}}(X_t | \underline{\chi}_{1-p_0:t-1})) \right\} \mathbb{I}(g_{\underline{\theta}}(X_1 | \underline{\chi}_{1-p_0:0}))$$

où $g_{\underline{\theta}}(X_t | \underline{\chi}_{1-p_0:t-1}) = (g_{\underline{\theta}_k}(X_t | \underline{\chi}_{1-p_0:t-1}), 1 \leq k \leq d)$. Un estimateur du maximum de vraisemblance (*EMV*) de $\underline{\theta}$ est définie comme toute solution mesurable $\hat{\underline{\theta}}_n$:

$$\hat{\underline{\theta}}_n = \arg \max_{\underline{\theta} \in \Theta} L_n(\underline{\theta}) \quad (3.4)$$

Remarque 3.1.1 *La consistance et la normalité asymptotique sont étudiées dans les articles suivants (C.f., Bickel et al, Duc et al) pour une discussion plus.*

Chapitre 4

Appendices

Dans ses appendices on donne les définitions, les concepts et les outils fondamentaux qui s'avèrent nécessaires dans notre étude.

4.1 Chaîne de Markov

4.1.1 Introduction

Selon les auteurs, une **chaîne de Markov** est de manière générale et simplifiée, la prédiction du futur, sachant le présent, n'est pas rendue plus précise par des éléments d'information supplémentaires concernant le passé, toute l'information utile pour la prédiction du futur est contenue dans l'état présent du processus. la **chaîne de Markov** portent le nom de leur découvreur, **Andreï Markov**. Quand **Markov** avait introduit son fameux modèle en 1906, il ne s'est pas réoccupé. Il a juste voulez prouver l'indépendance n'est pas nécessaire pour la lois des grands nombres, il a considéré l'exemple d'alternance entre les consonnes et les voyelles qui a été décrit comme une chaîne à deux états presque en même temps, et fidèlement à la tradition française des jeux de probabilistes, **Poincaré** a étudié les **chaînes de Markov** de group fini. **Les physiciens Australiens** et **Thana Ehrenfels** ont proposé en 1907 un modèle de la **chaîne de Markov** qui a beaucoup aidé à clarifier l'issue controversée de l'irréversibilité dynamique. **Sir Francis Galton**, un cousin de Darwin, qui était intéressé par la probabilité de suivre de pairie anglaise, était l'inventeur de processus de branchement. Une généralisation à un espace d'état infini dénombrable a été publiée par **Kolmogorov** en 1936. Une **chaîne de Markov** est une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$ qui permet de modéliser l'évolution dynamique d'un système aléatoire : (X_n) représente l'état du système à l'instant n .

Les applications des **chaînes de Markov** sont très nombreuses : réseaux, génétique de populations, mathématique financières, gestion de stock, algorithmes stochastiques d'optimisation.

Définition 4.1.1 Une suite (X_n) de variables aléatoires à valeurs dans un ensemble au plus dénombrable E est une chaîne de Markov d'espace d'états E si et seulement si pour tout $n \in \mathbb{N}$, pour tout $(i_0, i_1, \dots, i_{n+1})$ dans E telle que :

$$P(X_{n+1} = i_{n+1} | X_n = i_n, \dots, X_0 = i_0) = P(X_{n+1} = i_{n+1} | X_n = i_n)$$

Et voici quelques exemples :

- **La souris dans le labyrinthe** Une souris se déplace dans le labyrinthe. Initialement, elle se trouve dans la (case 1). A chaque minute, elle change de case en choisissant, de manière équiprobable, l'une des cases adjacentes. Dès qu'elle atteint soit la nourriture (case 4), soit sa tanière (case 5), elle y reste. On pose les questions suivantes :
 - Avec quelle probabilité la souris atteint-elle la nourriture plutôt que sa tanière ?
 - Au bout de combien de temps atteint-elle sa tanière ou la nourriture ?

On peut essayer de répondre à ces questions en construisant un arbre d'écrivant les chemins possibles. Par exemple, il est clair que la souris se retrouve dans sa tanière au bout d'une minute avec probabilité $1/3$. Sinon, elle passe soit dans la case 2, soit dans la case 3, et depuis chacune de ces cases elle a une chance sur deux de trouver la nourriture. Il y a donc une probabilité de $1/6$ que la souris trouve la nourriture au bout de deux minutes.

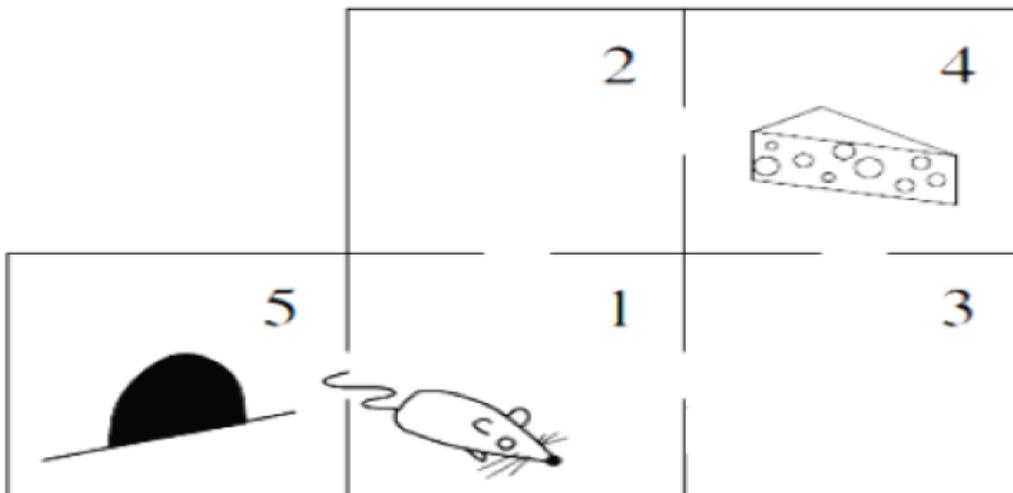


FIGURE 1 :La souris dans le labyrinthe.

- **jeu Pile ou Face** On lance une pièce régulière (pile=1, face=0) jusqu'à ce qu'on obtienne pour la première

fois soit 11 (deux piles consécutifs) soit 01 (face suivi de pile). A partir sur 11 alors que B parie sur 01.

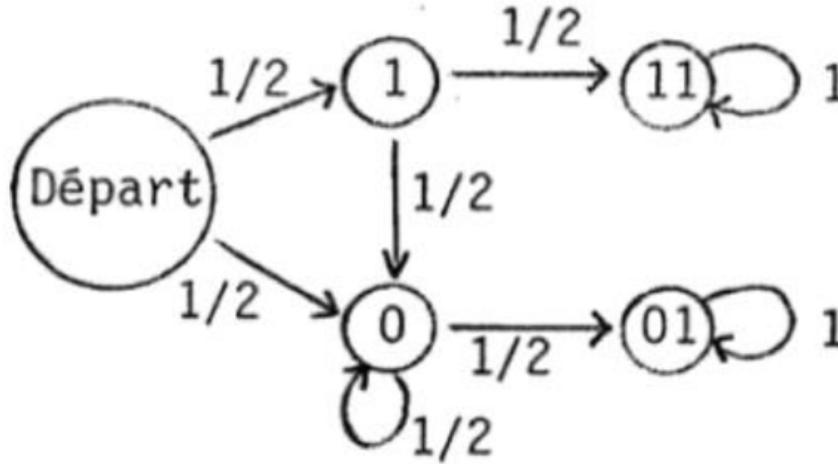


FIGURE 2 :Jeu pile ou face

4.1.2 Probabilité de transition

Définition 4.1.2 Soit $(X_n)_{n \in \mathbb{N}}$ une Chaîne de Markov à espace d'états E , $\forall i, j \in E$, la probabilité p_{ij} est appelée probabilité de transition de l'état i à l'état j tel que :

$$p_{ij} = P(X_n = j | X_{n-1} = i)$$

4.1.3 Matrice de transition

Définition 4.1.3 Soit $(X_n)_{n \in \mathbb{N}}$ chaîne de Markov de matrice de transition \mathbb{P} , alors \mathbb{P} vérifie :

- Pour tout $i, j \in E$ on a $p_{i,j} \in [0, 1]$
- Pour tout $i, j \in E$ on a $\sum_{j \in E} p_{i,j} = 1$.

Remarque 4.1.1 Une matrice de transition d'une chaîne de Markov est une matrice stochastique.

Quelque exemple sur la matrice de transition

$$\mathbb{P}_1 = \begin{pmatrix} 0,55 & 0,45 \\ 0,8 & 0,2 \end{pmatrix}, \quad \mathbb{P}_2 = \begin{pmatrix} 0,75 & 0,1 & 0,15 \\ 0,4 & 0,4 & 0,2 \\ 0,6 & 0,1 & 0,3 \end{pmatrix}, \quad \mathbb{P}_3 = \begin{pmatrix} 0,25 & 0,65 \\ 0,8 & 0,2 \end{pmatrix}.$$

Remarque 4.1.2 On peut aussi traduire ces informations à l'aide d'un graphe probabiliste.

Probabilité initial

Définition 4.1.4 Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov de matrice de transition \mathbb{P} et $\underline{\Pi}$ le vecteur-ligne décrivant l'état initial du système étudié telle que :

$$\underline{\Pi} = (\pi(1), \dots, \pi(d))', \text{ où } \pi(i) = P(s_t = i), i = 1, \dots, d$$

telle que :

$$\underline{\Pi}' = \underline{\Pi}' \mathbb{P}$$

4.1.4 Classification d'état

Etats irréductibles

Définition 4.1.5 L'état j est accessible depuis l'état i si et seulement s'il existe :

$$n \geq 0 \text{ telle que } P_{ij}^{(n)} > 0$$

Définition 4.1.6 Les états i et j communiquent si et seulement s'il existe :

$$n \geq 0 \text{ et } m \geq 0 : \text{ telle que } P_{ij}^{(n)} > 0 \text{ et } P_{ji}^{(m)} > 0$$

Remarque 4.1.3 Les états i et j communiquent s'ils sont accessibles l'un à partir de l'autre.

Exemple 4.1.1 Considérons une chaîne de Markov à valeurs dans $E = \{a, b, c, d, e\}$ et dont la matrice et le graphe de transition sont données par :

$$\mathbb{P} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

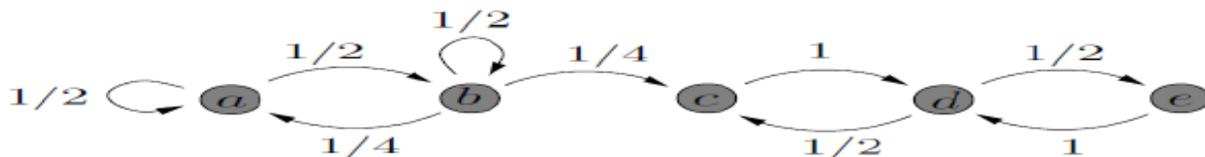


FIGURE 4 : Graphe probabiliste.

La chaîne comporte deux classes irréductibles : $\{a, b\}$ et $\{c, d, e\}$.

Etats récurrents et transients

notation 4.1.1 Pour tout état i , désignons par T_i le temps d'atteinte de l'état i à partir de l'instant 1, autrement dit :

$$T_i = \inf \{n \geq 1, X_n = i\}$$

ce temps d'atteinte est un temps d'arrêt de la chaîne.

Définition 4.1.7 On dit que l'état i est récurrent si, partant de l'état i , la probabilité que la chaîne de Markov retourne à l'état i en un temps fini est égale à 1

$$P(T_i < +\infty | X_0 = i) = 1$$

Remarque 4.1.4 L'état i est dit transient dans le cas contraire i.e. : $P(T_i < +\infty | X_0 = i) > 0$.

Exemple 4.1.2 Reprenons le cas de la chaîne de Markov à valeurs dans $E = \{a, b, c, d, e\}$ définie dans le paragraphe précédent. L'état b est transient en effet, l'état c est accessible depuis b mais pas l'inverse. Autrement dit, en allant en c depuis b , on est sûr de ne jamais y revenir. Et même cas pour l'état a . Les états c, d et e seront quant à eux récurrents.

Etat absorbant

Définition 4.1.8 Soit \mathbb{P} une matrice de transition à un système étudier :

$$\mathbb{P} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

L'état $\{4\}$ est absorbant.

Etat périodicité

Définition 4.1.9 Soit $i \in I$. On appelle période de i , et on note $d(i)$ le PGCD de tous les entiers $n \geq 1$ pour les quels $p_{ii}^{(n)} > 0$, $d(i) = \text{pgcd}(n \geq 1, p_{ii}^{(n)} > 0)$,

- Si $d(i) = d \geq 2$ on dit que est i périodique de période d .
- Si $d(i) = 1$, on dit i que est apériodique.

Remarque 4.1.5 Une chaîne apériodique est une chaîne de dont tous les états sont apériodiques.

Exemple 4.1.3 On considère 5 points équités partis sur un cercle. Un promeneur saute chaque instant, d'un point à l'un de ses voisins avec la probabilité $1/2$ pour chaque voisin. L'ensemble des états est : $E = \{1, 2, 3, 4, 5\}$. On a $p_{i,i+1} = p_{i,i-1} = \frac{1}{2}$ pour $i \in \{2, 3, 4\}$, et $p_{1,2} = p_{1,5} = \frac{1}{2}$ et $p_{5,1} = p_{5,4} = \frac{1}{2}$ ce qui donne la matrice :

$$\mathbb{P} = \begin{pmatrix} 0 & 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 & 0 \end{pmatrix}$$

La chaîne est irréductible (tous les états communiquent). Comme tous les états sont dans la même classe, leur période est la même. On a :

$$d(1) = p \operatorname{gcd} \left\{ n \geq 1; p_{1,1}^{(n)} > 0 \right\} p_{1,1}^{(2)} = p_{1,2} p_{2,1} = \frac{1}{4} > 0, p_{1,1}^{(5)} \geq p_{1,2} p_{2,3} p_{3,4} p_{4,5} p_{5,1} = \frac{1}{2^5} > 0$$

Or le seul diviseur commun à 2 et à 5 est 1. On a donc $d = d(1) = 1$ cette chaîne est donc apériodique.

4.1.5 Exemples

Exercice 1 Dans un certain pays, il ne fait jamais beau deux jours de suite. Si un jour il fait beau, le lendemain il peut neiger ou pleuvoir avec autant de chances. Si un jour il pleut ou il neige, il y a une chance sur deux qu'il y ait changement de temps le lendemain, et s'il y a changement, il y a une chance sur deux que ce soit pour du beau temps.

- Former, à partir de cela, une chaîne de Markov et en déterminer sa matrice de transition.
- Si un jour il fait beau, quel est le temps le plus probable pour le sur lendemain ?
- Si on suppose que l'on a que deux états (beau temps et mauvais temps), déterminer la matrice de transition de la nouvelle chaîne ainsi obtenue.

Solution 4.1.1 On a l'ensemble des états suivants $E = \{BT, PL, N\}$ et le temps pour un jour ne dépend que du temps du jour précédent, indépendamment de la période de l'année également. On a donc bien une chaîne de Markov, de matrice de transition :

$$\mathbb{P} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}$$

Pour le temps du surlendemain, il faut déterminer P^2 . Mais seule la première ligne de P^2 nous intéresse car on veut

déterminer les probabilités à partir d'un jour de beau temps. On a :

$$\begin{aligned}
 P_{BT,BT}^{(2)} &= P_{BT,BT} \times P_{BT,BT} \times P_{BT,PL} \times P_{PL,BT} + P_{BT,N} \times P_{N,BT} \\
 &= 0 \times 0 + \frac{1}{2} \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{4} = \frac{1}{4} \\
 P_{BT,P}^{(2)} &= P_{BT,BT} \times P_{BT,P} + P_{BT,PL} \times P_{PL,PL} + P_{BT,N} \times P_{N,BT} \\
 &= 0 \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{4} = \frac{3}{8} \\
 P_{BT,N}^{(2)} &= P_{BT,BT} \times P_{BT,N} + P_{BT,PL} \times P_{PL,N} + P_{BT,N} \times P_{N,N} \\
 &= 0 \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{2} = \frac{3}{8}
 \end{aligned}$$

Ainsi, si un jour il fait beau, le temps le plus probable pour le sur lendemain est la pluie ou la neige. On suppose maintenant que $E' = \{BT, MT\}$, ce qui est possible puisque la pluie et la neige se comportent de la même façon pour ce qui est des transitions. Donc $\mathbb{Q} = \begin{pmatrix} 0 & 1 \\ 1/4 & 3/4 \end{pmatrix}$.

Exercice 2 On considère sur $E = \{1, \dots, 6\}$ la matrice de transition :

$$\mathbb{P} = \begin{pmatrix} p_{1,1} & 0 & 1/2 & 0 & 0 & 0 \\ 0 & P_{2,2} & 0 & 3/4 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & P_{4,4} & 0 & 0 \\ 1/4 & 0 & 1/2 & 0 & P_{5,5} & 0 \\ 0 & 0 & 0 & 0 & 0 & P_{6,6} \end{pmatrix},$$

- Compléter la matrice \mathbb{P} pour en faire une matrice de transition.
- Représenter le graphe de la chaîne de Markov.
- Quels sont les états récurrents et transitoires de cette chaîne ?.

Solution 4.1.2 On a :

- Pour tout $i, j \in E$, $p_{i,j} \in [0, 1]$.
- Pour tout $i \in E$ on a : $\sum_{j \in E} p_{i,j} = 1$,

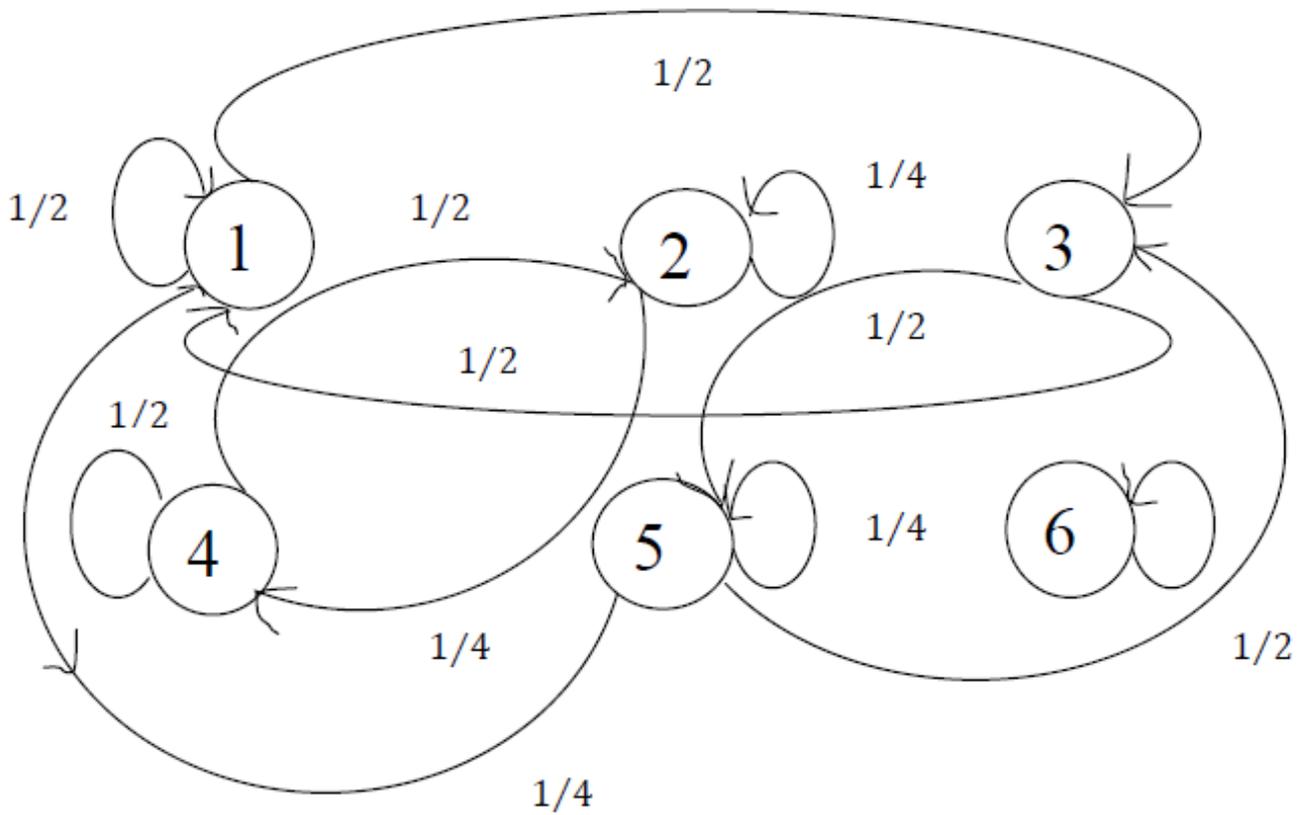
$$\begin{aligned}
 p_{1,1} &= 1 - (p_{1,2} + p_{1,3} + p_{1,4} + p_{1,5} + p_{1,6}) \\
 &= 1 - (0 + 1/2 + 0 + 0 + 0) = 1/2
 \end{aligned}$$

Même cas pour ; $p_{2,2}$, $p_{3,3}$, $p_{4,4}$, $p_{5,5}$

Donc :

$$\mathbb{P} = \begin{pmatrix} 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 3/4 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 1/4 & 0 & 1/2 & 0 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Le graphe de la chaîne de Markov est comme suite :



L'état E est transient en effet, l'état $\{1\}$ est accessible depuis $\{5\}$ mais pas l'inverse. Autrement dit, en 1 allant en depuis 5, on est sûr de ne jamais y revenir. La chaîne comporte 3 classes récurrentes : $\{1, 3\}$, $\{2, 4\}$, $\{3, 5\}$

Exercice 3 Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov sur $\{1, 2, 3\}$ de matrice de transition :

$$\mathbb{P} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 2/3 & 1/3 \\ p & 1-p & 0 \end{pmatrix},$$

Calculer $P = (X_3 = 1 | X_0 = 1)$ dans les cas suivants : $p = 1/6$ et $p = 1/12$.

Solution 4.1.3 On a :

$$p_{11}^{(3)} = p(X_3 = 1 | X_0 = 1)$$

Et on a pour $p = 1/6$ et $n = 1$, $P_{11}^{(1)} = 0$.

Pour $n = 2$

$$\mathbb{P}^{(2)} = \mathbb{P}^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 2/3 & 1/3 \\ 1/6 & 5/6 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 2/3 & 1/3 \\ 1/6 & 5/6 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 1/18 & 13/18 & 4/18 \\ 0 & 13/18 & 5/18 \end{pmatrix}$$

Donc $P_{11}^{(2)} = 0$.

Pour $n = 3$

$$\mathbb{P}^{(3)} = \mathbb{P}^3 = \mathbb{P}^2 \mathbb{P} = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 1/18 & 13/18 & 4/18 \\ 0 & 13/18 & 5/18 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 2/3 & 1/3 \\ 1/6 & 5/6 & 0 \end{pmatrix} = \begin{pmatrix} 1/18 & 13/18 & 4/18 \\ 1/26 & 39/54 & 13/54 \\ 5/108 & 77/108 & 13/54 \end{pmatrix}$$

Donc $P_{11}^{(3)} = 0$.

Remarque 4.1.6 De même manière on calcule $P = (X_3 = 1 | X_0 = 1)$ pour $p = 1/12$.

4.2 Convergence des suites des variables aléatoires

Inégalité de Markov

Soit X une variable aléatoire à valeurs non-négatives et l'espérance $E\{X\}$ est finie, pour tout réel $a > 0$, on a :

$$P(X \geq a) \leq \frac{E\{X\}}{a}$$

Preuve La démonstration qui suit s'applique à une variable continue de densité f :

$$E\{X\} = \int_0^{+\infty} xf(x)dx = \int_0^a xf(x)dx + \int_a^{+\infty} xf(x)dx \geq \int_a^{+\infty} xf(x)dx \geq \int_0^a af(x)dx = a \int_0^a f(x)dx = aP(X \geq a)$$

Lemme de Borel-Cantelli

Soit $(X_n)_{n \in \mathbb{N}}$ une famille d'événements, on note :

$$\lim_{n \rightarrow +\infty} \sup_{n \in \mathbb{N}} (X_n) = \bigcap_{n \geq 0} \left(\bigcup_{j \geq n} X_j \right)$$

$$\lim_{n \rightarrow +\infty} \inf_{n \in \mathbb{N}} (X_n) = \bigcup_{n \geq 0} \left(\bigcap_{j \geq 0} X_j \right)$$

Dans un espace probabilisé (Ω, E, P) , considérons une suite $(X_n)_{n \in \mathbb{N}}$ d'éléments de X .

1. Si $\sum_{n \geq 0} P(X_n) < +\infty$ alors $P(\lim_{n \rightarrow +\infty} \sup_{n \in \mathbb{N}} (X_n)) = 0$.
2. Si $\sum_{n \geq 0} P(X_n) = +\infty$ alors $P(\lim_{n \rightarrow +\infty} \sup_{n \in \mathbb{N}} (X_n)) = 1$.

Preuve Soit (X_n) une suite d'événements dans un espace de probabilité et supposons que la somme des probabilités de la (X_n) est fini. On suppose que :

$$\sum_{n \geq 0} P(X_n) < +\infty$$

et pour tout $i \in \mathbb{N}$, on a :

$$\lim_{n \rightarrow +\infty} \sup_{n \in \mathbb{N}} (X_n) = \bigcap_{n \geq 0} \bigcup_{j \geq n} X_j \subset \bigcup_{j \geq i} X_j$$

Donc :

$$P\left(\lim_{n \rightarrow +\infty} \sup_{n \in \mathbb{N}} (X_n)\right) \leq P\left(\bigcup_{j \geq i} X_j\right) \leq \sum_{j \geq i} P(X_j)$$

Ce dernier terme tends vers 0 quand j tends vers $+\infty$ car c'est le reste d'une série convergente donc : $P(\lim_{n \rightarrow +\infty} \sup_{n \in \mathbb{N}} (X_n)) = 1$.

4.2.1 Mode de convergences

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires et soit X une variable aléatoire. On définit les modes de convergence suivants :

Convergence presque sûre

Définition 4.2.1 On dit que la suite X_n converge vers X presque sûre si :

$$P\left(\left\{\omega \in E \mid \lim_{n \rightarrow +\infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

Dans ce cas on note : $X_n \xrightarrow{p.s.} X$.

Exemple 4.2.1 Soit $(Y_n)_{n \in \mathbb{N}}$ une suite des variables aléatoires et $X_n = Y_1 + \dots + Y_n$. On montre que $\frac{X_n}{n^2}$ converge presque sûre vers 0. En effet l'ensemble $E = \{0 \leq |X_n| \leq n, \text{ pour tout } n\}$ et telle que $P(E) = 1$. Donc pour $\omega \in E$ on a que $0 \leq \left|\frac{X_n(\omega)}{n^2}\right| \leq \frac{1}{n}$ ce qu'implique que :

$$\lim_{n \rightarrow +\infty} \frac{X_n(\omega)}{n^2} = 0, \text{ Pour } \omega \in E$$

et donc que :

$$P\left(\lim_{n \rightarrow +\infty} \frac{X_n(\omega)}{n^2} = 0\right) = 1$$

qui montre la convergence presque sûre.

Propriétés sur la convergence presque sûre

Proposition 4.2.1 Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires réelles

- Si pour tout $\varepsilon > 0$, $\sum_{n=0}^{+\infty} P(|X_n - X| \geq \varepsilon) < +\infty$ alors $X_n \xrightarrow{p,s} X$.
- Si les variables aléatoires X_n sont indépendantes alors $X_n \xrightarrow{p,s} 0$ ssi pour tout $\varepsilon > 0$, on a $\sum_{n=0}^{+\infty} P(|X_n| \geq \varepsilon) < +\infty$.

Proposition 4.2.2 Soit f est une application continue de \mathbb{R}^2 dans \mathbb{R} , $(X_n)_{n \in \mathbb{N}}$ et $(Y_n)_{n \in \mathbb{N}}$ deux suites de variables aléatoires réelles convergeant presque sûre respectivement vers les v.a.r X et Y , alors la suite de v.a.r $(f(X, Y))_n$ converge presque sûre vers la v. a. r $f(X, Y)$.

Proposition 4.2.3 Soit (X_n, X) et (Y_n, Y) deux suites de variables aléatoires réelles alors :

- $(X_n, Y_n) \xrightarrow{p,s} (X, Y) \Leftrightarrow (X_n \xrightarrow{p,s} X, Y_n \xrightarrow{p,s} Y)$
- Si $X_n \xrightarrow{p,s} X$ et $Y_n \xrightarrow{p,s} Y$ alors :
 - $X_n + Y_n \xrightarrow{p,s} X + Y$
 - $X_n Y_n \xrightarrow{p,s} Y X$
 - $\frac{X_n}{Y_n} \rightarrow \frac{X}{Y}$ avec $Y_n \neq 0$ et (presque sûre).

Convergence en probabilité

Définition 4.2.2 Soit X_n une suite des variables aléatoires, on dit que X_n converge en probabilité vers 0 et on note $X_n \xrightarrow{P} 0$ si :

$$\forall \varepsilon > 0 : \lim_{n \rightarrow +\infty} P(|X_n| > \varepsilon) = 0$$

Définition 4.2.3 Soit X_n une suite des variables aléatoires, on dit que X_n est borné de probabilité, si pour tout $\varepsilon > 0$, il existe M_ε tel que :

$$P(|X_n| > M_\varepsilon) < \varepsilon, \text{ pour tout } n$$

Nous pouvons examiner l'idée de la limite de probabilité de la manière alternative suivante :

On dit que X_n est borné de probabilité pour tout $\varepsilon > 0$, il existe un ensemble $F \in \mathcal{F}$ et un nombre M_ε tel que pour tout $\omega \in F$, $|X_n(\omega)| < M_\varepsilon \forall n$, et $P(F^c) < \varepsilon$.

Pour s'habituer à ces idées, prouvons le fait de base suivante : si $X_n \xrightarrow{P} 0$, alors X_n est borné de probabilité.

C'est si $X_n = o_p(1)$, alors $X_n = O_p(1)$. Cela correspond à l'idée de base de l'analyse réelle qui indique qu'une séquence convergente est délimitée.

L'argument est le suivant ; supposer $X_n \xrightarrow{P} 0$. Soit $\varepsilon > 0$; on sait que $P(|X_n| > 1) \rightarrow 0$, par conséquent, il y a n_0 tel que :

$$P(|X_n| > 1) < \varepsilon, \text{ pour tout } n \geq n_0.$$

Maintenant, on choisit M_0 suffisamment grand pour que $P(|X_i| > M_0) < \varepsilon$ pour tout $i = 1, \dots, n_0 - 1$. Ensuite, on a pour $M = (1, M_0)$ que $P(|X_n| > M) < \varepsilon$, pour tout n .

Qui complète la preuve que $X_n = O_p(1)$.

4.3 Processus stochastique

Définition 4.3.1 *un processus stochastique est une famille $(X_t)_{t \in T}$ des variables aléatoires indexés par le temps t , les mots processus et stochastique signifient respectivement fonction et aléatoire, alors qu'une variable aléatoire X associée à chaque $\omega \in \Omega$ une réalisation $X(\omega)$, un processus stochastique $(X_t)_{t \in T}$ associé à chaque ω une fonction $(X_t)_{t \in T}$:*

$$\begin{aligned} T &\longrightarrow E \\ t &\longmapsto X_t(\omega) \end{aligned}$$

$tq : T$ l'ensemble des temps.

Remarque 4.3.1 *Lorsque l'ensemble temps T est au plus dénombrable (par exemple $T = \mathbb{N}$), on parle de processus stochastique à temps discret, lorsqu'il est continue (i.e. $T = [0; t_0]$ ou $T = \mathbb{R}_+$) on parle processus stochastique à temps continue.*

4.3.1 Bruit blanc

Définition 4.3.2 *(Faible bruit blanc) Un Bruit Blanc est une suite de v.a.r. noté par $(\varepsilon_n)_{n \in \mathbb{N}}$ telle que :*

- (i) $E(\varepsilon_t) = 0, \forall t \in \mathbb{Z}$.
- (ii) $Var(\varepsilon_t) = \sigma^2, \forall t \in \mathbb{Z}$.
- (iii) $Cov(\varepsilon_t, \varepsilon_{t+h}) = 0, \forall t, h \in \mathbb{Z}, h \neq 0$.

Définition 4.3.3 *(Fort bruit blanc) Il convient de noter qu'aucune hypothèse d'indépendance n'est faite Dans la définition du faible bruit blanc. Les variables à différentes dates ne sont pas corrélées et La distinction est particulièrement cruciale pour les séries chronologiques financières. Il est parfois nécessaire de Hypothèse (iii) par l'hypothèse la plus forte (iii') les variables ε_n et ε_{n+h} sont indépendantes et identiquement distribuées. Le processus ε_n est alors dit être un fort bruit blanc.*

4.3.2 Le processus causal

Définition 4.3.4 *On dit que le processus X_t est un processus causal si on peut écrire :*

$$X_t = \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-h} \text{ avec } \sum_{i=0}^{\infty} |\phi_1^i| < \infty.$$

Parfois, lorsque l'on parle d'un processus causal, on dit que celui-ci a une représentation MA.

4.3.3 Le processus MA

Définition 4.3.5 *Un processus est dit moyenne mobile d'ordre q , noté par MA_q , s'il admet l'écriture suivante :*

$$X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_q X_{t-q} + \varepsilon_t \text{ où } \varepsilon_t \sim BB(0,1).$$

Ce processus est toujours stationnaire, mais il est inversible que lorsque les racines (en module) du polynôme caractéristique associé sont toutes dehors du cercle unité. C'est-à-dire :

$$X_t \text{ est inversible} \Leftrightarrow (1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - \cdots - \phi_q L^q)$$

4.3.4 Ergodicité

On dit qu'une suite stationnaire est ergodique si elle satisfait la loi forte des grands nombres.

Définition 4.3.6 (*Processus stationnaire ergodique*) *Un processus strictement stationnaire $(Z_t)_{t \in \mathbb{Z}}$, à valeurs réelles, est dit ergodique si et seulement si, pour tout borélien B et tout entier k ,*

$$n^{-1} \sum_{t=1}^n 1_B(Z_t, Z_{t+1}, \dots, Z_{t+k}) \rightarrow P\{(Z_1, \dots, Z_{1+k}) \in B\}$$

avec probabilité 1².

Certaines transformations de suites ergodiques restent ergodiques.

Définition 4.3.7 *On dit que processus stochastique $\{X(t), t \in T\}$ est ergodique si tout caractéristique du processus peut être obtenue, avec une probabilité de 1, à partir d'une seule réalisation $X(t, s)$ du processus. Un processus stochastique peut être, en particulier, ergodique par rapport à la moyenne, à la fonction de répartition, à la fonction d'autocorrélation, ect.*

On dit qu'une suite stationnaire est ergodique si elle satisfait la loi forte des grands nombres.

Théorème 4.3.1 *Si $(Z_t)_{t \in \mathbb{Z}}$ est une suite strictement stationnaire ergodique et si $(Y_t)_{t \in \mathbb{Z}}$ est définie par :*

$$Y_t = f(\dots, Z_{t-1}, Z_t, Z_{t+1}, \dots)$$

où f est une fonction mesurable de \mathbb{R}^∞ dans \mathbb{R} , alors $(Y_t)_{t \in \mathbb{Z}}$ est également une suite strictement stationnaire ergodique.

En particulier, si $(X_t)_{t \in \mathbb{Z}}$ est la solution stationnaire non anticipative de l'équation $AR(1)$

$$X_t = \alpha_1 X_{t-1} + \varepsilon_t, |\alpha_1| < 1$$

alors le théorème montre que $(X_t)_{t \in \mathbb{Z}}$, $(X_{t-1} \varepsilon_t)_{t \in \mathbb{Z}}$ et $(X_{t-1}^2)_{t \in \mathbb{Z}}$ sont des suites stationnaires et ergodiques.

4.3.5 La stationnarité

Sommairement, la stationnarité traduit la capacité d'un processus à ne pas dépendre de l'indice temporel. Ce dernier est dès lors entièrement décrit par sa loi stationnaire qui, par définition, n'évolue plus au cours du temps. On comprend ainsi qu'une telle propriété est certes d'intérêt pratique considérable, mais possède également un fort impact théorique puisqu'on la retrouve comme hypothèse à la base de nombreux résultats. On distingue généralement la stationnarité au sens strict de la stationnarité au sens faible. Pour les définir, considérons un processus X_t .

Définition 4.3.8 *On dit que le processus (X_t) est stationnaire au sens strict (ou fortement stationnaire) si la loi de $\{X_1, \dots, X_k\}$ est la même que la loi de $\{X_{1+h}, \dots, X_{k+h}\}$ pour tout $k \in \mathbb{N}, h \in \mathbb{Z}$.*

La stationnarité stricte est une hypothèse de travail très forte, nécessairement délicate à vérifier en pratique lorsque le processus n'est pas gaussien. C'est pourquoi l'on a introduit une notion de stationnarité moins contraignante.

Définition 4.3.9 *le processus (X_t) est dit stationnaire du second ordre (ou faiblement stationnaire) si :*

- $E(X_t) = m, \forall t \in \mathbb{Z}$.
- $Var(X_t) < \infty, \forall t \in \mathbb{Z}$.
- $Cov(X_t, X_{t-h}) = \gamma(h), \forall h, t \in \mathbb{Z}$.

On dit aussi que le processus est « stationnaire au second ordre », en relation avec la stabilisation de sa variance. C'est à cette propriété de stationnarité que nous ferons implicitement référence par la suite. Notons que la stationnarité stricte implique bien entendu la stationnarité faible. L'exemple le plus trivial de processus stationnaire est le bruit blanc.

4.3.6 Autocorrélations

L'une des principales motivations de la modélisation chronologique d'un événement aléatoire est sa structure de corrélation temporelle, en d'autres termes son niveau d'autocorrélation. Il existe, entre autres, deux outils permettant d'évaluer l'autocorrélation d'une série chronologique.

La fonction d'auto covariance

Définition 4.3.10 On dit que $\{\gamma(h)\}_{h \in \mathbb{Z}}$ une fonction d'auto covariance si elle mesure la covariance entre une variable et cette même variable à des dates différentes, pour un délai h .

$$\gamma(h) = Cov(X_t, X_{t-h}) = E[(X_t - E(X_t))(X_{t-h} - E(X_{t-h}))]$$

Ainsi :

$$\gamma(0) = Var(X_t) = E[(X_t - E(X_t))^2] = \sigma_X^2$$

La fonction d'autocorrélation

Définition 4.3.11 La fonction d'autocorrélation est définie par :

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

avec $\rho(0) = 1$ et $|\rho(h)| < 1$ donc $(|\gamma(h)| < \gamma(0))$.

L'équivalent empirique de la fonction d'autocorrélation, noté $\hat{\rho}(h)$ est obtenu à partir de l'estimateur suivant pour l'autocovariance $\hat{\gamma}(h)$ à l'ordre h :

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{j=1}^{n-h} (X_j - \bar{X})(X_{j+h} - \bar{X}) = \hat{\gamma}(-h)$$

Où

$$\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$$

Désigne la moyenne empirique.

On définit de manière analogue la fonction d'autocorrélation empirique par :

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \text{ pour } |h| < n$$

4.3.7 Moment d'ordre supérieur**Moment**

Définition 4.3.12 Pour tous entier n non nul, on définit le moment d'ordre n d'une variable aléatoire X prenant les valeurs $\{x_k; k \in \mathbb{k}\}$ comme l'espérance, si elle existe, de la variable aléatoire X^n :

$$m_n(X) = \mathbf{E}(X^n)$$

C'est-à-dire :

$$m_n(X) = \mathbf{E}(X^n) = \sum_{k \in \mathbb{k}} x_k^n p\{X = x_k\}$$

Pour une variable aléatoire discrète.

Et :

$$m_n(X) = \mathbf{E}(X^n) = \int_{-\infty}^{+\infty} x^n f(x) dx$$

Pour une variable aléatoire continue.

Remarque 4.3.2 *Le moment d'ordre 0 d'une variable aléatoire existe toujours et vaut*

$$m_0(X) = \mathbf{E}(X^0) = \mathbf{E}(1) = 1$$

le moment d'ordre 1, s'il existe, est l'espérance de X

Théorème 4.3.2 *Soit X une variable aléatoire réelle (discrète ou à densité), et soit $n \geq 2$. Si X admet un moment d'ordre $n \geq 2$, alors X admet des moments d'ordre 1, 2, ..., n.*

Variables centrées réduites

Si X est une variable admettant une espérance, alors la variable

$$Y = X - \mathbf{E}(X)$$

est une variable aléatoire **centrée**, c'est-à-dire d'espérance nulle. Si de plus Y admet une variance strictement positive (admet un moment d'ordre 2), on définit la variable aléatoire

$$Z = \frac{X - \mathbf{E}(X)}{\sigma_X} = \frac{X - m(X)}{\sigma_X} = \frac{Y}{\sigma_X}$$

qu'on appelle **variable aléatoire centrée réduite** associée à X. Son espérance est nulle (variable **centrée**) et sa variance est égale à 1 (variable **réduite**). Si f est la densité de X, alors celle de Z est

$$f(x) = \sigma \cdot f(\sigma(x + m))$$

Moments centrés

Soit X une variable aléatoire (discrète ou à densité), admettant une espérance. On dit que X admet **un moment centré d'ordre n** si la variable aléatoire centrée $Y = X - \mathbf{E}(X)$ admet un moment d'ordre n . Le moment centré d'ordre n de X est alors :

$$\mu_n(X) = m_n(Y) = \mathbf{E}(Y^n) = \mathbf{E}[(X - \mathbf{E}(X))^n]$$

Définition 4.3.13 Soit X une variable aléatoire (discrète ou à densité). Alors X admet un moment centré d'ordre n si et seulement si elle admet un moment d'ordre n .

Variance et écart-type

La variance d'une variable aléatoire $V(X)$ est l'espérance mathématique du carré de l'écart à l'espérance mathématique c'est un paramètre de dispersion qui correspond au moment centré d'ordre 2 de la variable aléatoire X .

Définition 4.3.14 On dit qu'une variable aléatoire **admet une variance** si elle admet un moment d'ordre 2. Dans ce cas, on définit sa **variance** comme son moment centré d'ordre 2 :

$$V(X) = \mu_2(X) = \mathbf{E}[(X - \mathbf{E}(X))^2]$$

La variance d'une variable aléatoire discrète La variance d'une variable aléatoire réelle discrète X est une valeur numérique permettant de quantifier la dispersion des valeurs possibles de X par rapport à sa moyenne (son espérance)

X est un variable aléatoire discrète de loi de probabilité $(X_i; p_i)$ définie sur un nombre fini (n) d'événements élémentaires alors la variance est égale à :

$$V(X) = \sum_{i=1}^n (X_i - E(X))^2 P_i$$

La variance d'une variable aléatoire continue X est une variable aléatoire continue donnée par sa densité de probabilité alors la variance de X et le nombre réel positive tel que :

$$V(X) = \int_{-\infty}^{+\infty} (X - E(X))^2 f(x) dx$$

Propriété de la variance

Proposition 4.3.1 *Si X est une variable aléatoire admettant une variance, alors X^2 admet une espérance et*

$$\begin{aligned}\mathbf{V}(X) &= \mathbf{E} [(X - \mathbf{E}(X))^2] = \mathbf{E} [X^2 + [\mathbf{E}(X)]^2 - 2X\mathbf{E}(X)] = \mathbf{E}(X^2) - 2\mathbf{E}(X)^2 + \mathbf{E}(X)^2 \\ &= \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \mu_2(X) - \mu_1(X)^2\end{aligned}$$

(La formule de Koing-Hygens)

Par ailleurs nous admettons que la variance d'une somme de deux variables aléatoires indépendantes est égale à la somme des variances des variables :

$$\mathbf{V}(X + Y) = \mathbf{V}(X) + \mathbf{V}(Y)$$

(car X et Y sont indépendantes)

Si les variables ne sont pas indépendantes :

$$\mathbf{V}(X + Y) = \mathbf{V}(X) + \mathbf{V}(Y) + 2cov(X, Y)$$

Avec :

$$cov(X, Y) = \mathbf{E}([X - \mathbf{E}(X)][Y - \mathbf{E}(Y)])$$

Proposition 4.3.2 *Soient X une variable aléatoire (discrète ou à densité) admettant une variance Soient a et b des réels. Alors $aX + b$ admet une variance,*

$$\mathbf{V}(aX + b) = a^2\mathbf{V}(X)$$

Remarque 4.3.3 (Stricte positivité de la variance) *Remarquons bien si X est absolument continue et admet une variance, alors $\mathbf{V}(X) > 0$; en effet, on peut écrire, en notant $m = E(X)$,*

$$\mathbf{V}(X) = \mathbf{E} [(X - m)^2] = \int_{-\infty}^{+\infty} (x - m)^2 f(x) dx$$

Cette intégrale ne peut être nulle que si f est presque partout nulle, ce qui est exclu.

Ecart-Type La racine carrée de $Var(X)$ est appelée l'**écart-type** de X , qui se note σ .

$$\sigma = \sqrt{Var(X)}$$

Les variables aléatoires discrètes sont souvent réparties en catégories selon le type de leur loi. Les sections suivantes présentent quelques-uns de ces types.

4.4 Notions général

4.4.1 Le rayon spectral

Définition 4.4.1 Si A est un *endomorphisme sur un espace de Banach complexe* E , on appelle **rayon spectral** de A , et on note $\rho(A)$, en dimension finie, pour un endomorphisme de **valeurs propres complexes** $\lambda_1, \lambda_2, \dots, \lambda_n$, le rayon spectral est égal à $\max_i |\lambda_i|$.

Le théorème de Gelfand nous dit que le rayon spectral $\rho(A)$ d'un endomorphisme A est donné par la formule $\rho(A) = \lim_{n \rightarrow +\infty} \|A^n\|^{1/n}$.

4.4.2 Produit de Kronecker

En mathématiques le produit de Kronecker est une opération portant sur les matrices. Il s'agit d'un cas particulier du produit tensoriel. Il est ainsi dénommé en hommage au mathématicien allemand Leopold Kronecker.

Définition 4.4.2 Soient A une matrice de taille $m \times n$ et B une matrice de taille $p \times q$. Leur produit tensoriel est la matrice $A \otimes B$ de taille mp par nq , définie par blocs successifs de taille $p \times q$, le bloc d'indice i, j valant $(a)_{ij} \cdot B$. En d'autres termes :

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}$$

$$\text{Exemple 4.4.1} \quad \begin{pmatrix} 1 & 3 & 2 \\ 1 & 0 & 0 \\ 1 & 2 & 2 \end{pmatrix} \otimes \begin{pmatrix} 0 & 5 \\ 5 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1. \begin{pmatrix} 0 & 5 \\ 5 & 0 \\ 1 & 1 \end{pmatrix} & 3. \begin{pmatrix} 0 & 5 \\ 5 & 0 \\ 1 & 1 \end{pmatrix} & 2. \begin{pmatrix} 0 & 5 \\ 5 & 0 \\ 1 & 1 \end{pmatrix} \\ 1. \begin{pmatrix} 0 & 5 \\ 5 & 0 \\ 1 & 1 \end{pmatrix} & 0. \begin{pmatrix} 0 & 5 \\ 5 & 0 \\ 1 & 1 \end{pmatrix} & 0. \begin{pmatrix} 0 & 5 \\ 5 & 0 \\ 1 & 1 \end{pmatrix} \\ 1. \begin{pmatrix} 0 & 5 \\ 5 & 0 \\ 1 & 1 \end{pmatrix} & 2. \begin{pmatrix} 0 & 5 \\ 5 & 0 \\ 1 & 1 \end{pmatrix} & 2. \begin{pmatrix} 0 & 5 \\ 5 & 0 \\ 1 & 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 & 5 & 0 & 15 & 0 & 10 \\ 5 & 0 & 0 & 15 & 10 & 0 \\ 1 & 1 & 3 & 3 & 2 & 2 \\ 0 & 5 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 10 & 0 & 10 \\ 5 & 0 & 10 & 0 & 10 & 0 \\ 1 & 1 & 2 & 2 & 2 & 2 \end{pmatrix}.$$

Propriétés 4.4.1 *Le produit de Kronecker est bilinéaire et associatif : sous réserve de compatibilité des tailles pour A , B et C , on a les équations suivantes :*

$$A \otimes (B + \lambda.C) = (A \otimes B) + \lambda(A \otimes C)$$

$$(A + \lambda.B) \otimes C = (A \otimes C) + \lambda(B \otimes C)$$

$$A \otimes (B \otimes C) = (A \otimes B) \otimes C$$

Le produit de Kronecker n'est pas commutatif ; cependant pour toutes A et B il existe deux matrices de permutation P et Q telles que $A \otimes B = P(B \otimes A)Q$

Si de plus A et B ont la même taille, alors $A \otimes B$ et $B \otimes A$ sont équivalentes par permutation sur les vecteurs de la base :

$$A \otimes B = P^{-1}(B \otimes A)P = {}^t P(B \otimes A)Q$$

où P est une matrice de permutation.

Propriétés sur le produit de Kronecker :

On a la propriété suivante qui mélange les aspects liés au produit matriciel usuel et au produit de Kronecker :

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

On peut en déduire que $(A \otimes B)$ est inversible si et seulement si A et B sont inversibles, auquel cas : $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

On a la propriété suivante sur la transposée : ${}^t(A \otimes B) = {}^t A \otimes {}^t B$.

4.4.3 La fonction de vraisemblance

La fonction de vraisemblance notée $L(x_1, \dots, x_n | \theta_1, \dots, \theta_k)$ est une fonction de probabilités conditionnelles qui décrit les valeurs x_j d'une loi statistique en fonction des paramètres θ_i . Elle s'exprime à partir de la fonction de densité $f(x|\theta)$ par

$$L(x_1, \dots, x_n | \theta) = \prod_{i=1}^n f(x_i; \theta)$$

avec

$$f(x_i; \theta) = \begin{cases} f_{\theta}(x) & \text{si } X \text{ est une variable aléatoire continue.} \\ P_{\theta}(X = x) & \text{si } X \text{ est une variable aléatoire discrète.} \end{cases}$$

L'estimateur de maximum de vraisemblance

Définition 4.4.3 *L'estimation de maximum de vraisemblance de θ est la valeur $\hat{\theta}_n$ de θ qui rend maximale la fonction de vraisemblance $L(x_1, \dots, x_n | \theta)$.*

L'estimateur de maximum de vraisemblance (EMV) de θ est la variable aléatoire correspondante

Donc $\hat{\theta}_n$ sera en général calculé en maximisant la log-vraisemblance :

$$\hat{\theta}_n = \arg \max_{\theta} \ln L(x_1, \dots, x_n | \theta).$$

Bibliographie

- [1] Bibi.A, and A. Aknouche, *Stationnarité et b-méange des processus bilinéaires généraux à changement de régime markovien*, Sciences Paris Ser. I 348(3 – 4) (2010). 185 – 188.
- [2] Bibi.A, and A.Ghazel, *On the Markov-switching bilinear processes : stationarity, higher-order moments and β -mixing*, Stochastics An International Journal of Probability and Stochastic Processes : formerly Stochastics and Stochastics Reports. Published online : 04 Jun 2015. 3 – 15.
- [3] Bollerslev, T. (1986). *Generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics, 31, 307 – 327.
- [4] Bougerol.P, and N. Picard, *Strict stationarity of generalized autoregressive processes*, Ann.Prob. 20(4) (1992). 1714 – 1730.
- [5] Box, G. E. P. & JENKINS, G. M. (1970). *Times series analysis*. Forecasting and control. San Francisco, Calif. : Holden-Day.
- [6] Brandt.A, *The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients*, Adv. Appl.Prob. 18(1) (1986). 211 – 220.
- [7] Brockwell, P.J. and Davis, R.A. (1987), *Time Series : Theory and Methods*, Springer-Verlag, New York.
- [8] Chen, S.-W. and J.-L. Lin (2000a). *Modeling business cycles in Taiwan with time-varying Markov-switching models*, Academia Economic Papers, 28, 17 – 42.
- [9] Diebold, F.X., J.-H. Lee and G.C. Weinbach (1994). *Regime switching with time-varying transition probabilities*, in C. Hargreaves (ed.) *Nonstationary Time Series Analysis and Cointegration*. 283 – 302.
- [10] Duc, R; E. Moulines et T. Rydèn (2004). *Asymptotic properties of the maximum likelihood estimator in autoregressive models with Markov regime*. The Annals of Statistics. Vol. 32, No. 5, 2254 – 2304.
- [11] Dueker, M.J. (1997). *Markov switching in GARCH processes and mean-reverting stock market volatility*, Journal of Business & Economic Statistics, 15, 26 – 34.
- [12] Engel, C. (1994). *Can the Markov switching model forecast exchange rates ?*, Journal of International Economics, 36, 151 – 165.

- [13] Engel, C. and J.D. Hamilton (1990). *Long swings in the dollar : Are they in the data and do markets know it ?*, American Economic Review, 80, 689 – 713.
- [14] Engle, R. (1982). *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation*, Econometrica, 50, 987 – 1007.
- [15] Filardo, A.J. (1994). *Business-cycle phases and their transitional dynamics*, Journal of Business & Economic Statistics, 12, 299 – 308.
- [16] Francq, C. and J.M. Zakoian, *L^2 -structures of standard and switching-regime GARCH models*, Stoch. Process. Appl. 115(2005).1557 – 1582.
- [17] Francq, C. and J.M. Zakoian, *Stationarity of multivariate Markov-switching ARMA models*, J. Econ. 102(2) (2001).339 – 364.
- [18] Francq, C. Zakoian, J.-M., 2005. *L^2 -structures of standard and switching-regime GARCH models*. Stochastic Process. Appl. 115, 1557 – 1582.
- [19] Garcia, R. and P. Perron (1996). *An analysis of the real interest rate under regime shifts*, Review of Economics and Statistics, 78, 111 – 125.
- [20] Ghysels, E. (1994). *On the periodic structure of the business cycle*, Journal of Business & Economic Statistics, 12, 289 – 298.
- [21] Goldfeld, S.M. and R.E. Quandt (1973). *A Markov model for switching regressions*, Journal of Econometrics, 1, 3 – 16.
- [22] Goodwin, T.H. (1993). *Business-cycle analysis with a Markov switching model*, Journal of Business & Economic Statistics, 11, 331 – 339.
- [23] Grahn, T., 1995. *A conditional least squares approach to bilinear time series estimation*. J. Time Ser. Anal. 16, 509 – 529.
- [24] Granger, C.W.J., Anderson, A., 1978. *An Introduction to Bilinear Time Series Models*. Vandenhoeck and Ruprecht, Gottingen.
- [25] Granger, C.W.J, and A. Anderson, *An introduction to bilinear time series models*, Vandenhoeck and Ruprecht, Gottingen, 1978.
- [26] Guégan, D., Pham, D.T., 1989. *A note on the estimation of the parameters of the diagonal bilinear model by the method of least squares*. Scand. J. Stat. 16,129 – 136.
- [27] Gray, S.F. (1996). *Modeling the conditional distribution of interest rates as a regime- switching process*, Journal of Financial Economics, 42, 27 – 62.
- [28] Hamilton, J.D. (1989). *A new approach to the economic analysis of nonstationary time series and the business cycle*, Econometrica, 57, 357 – 384.

- [29] Hamilton.J.D, *A new approach to the economic analysis of nonstationary time series and the business cycle*, Econometrica 57(2) (1989).357 – 384.
- [30] Hamilton, J.D. and G. Lin (1996). *Stock market volatility and the business cycle*, Journal of Applied Econometrics, 11, 573 – 593.
- [31] Hamilton, J.D. (1988). *Rational-expectations econometric analysis of changes in regimes : An investigation of the term structure of interest rates*, Journal of Economic Dynamics and Control, 12, 385 – 423.
- [32] Hamilton, J.D. and R. Susmel (1994). *Autoregressive conditional heteroscedasticity and changes in regime*, Journal of Econometrics, 64, 307 – 333.
- [33] Haas.M, S. Mittnik, and M.S. Paoella, *A new approach to Markov-switching GARCH models*,J. Financ. Econ. 2(4)(2004).493 – 530.
- [34] Hsu, S.-H. and C.-M. Kuan (2001). Identifying Taiwan's business cycles in 1990 : *An application of the bivariate Markov switching model and Gibbs sampling (in Chinese)*, Journal of Social Sciences and Philosophy, 13, 515 – 540.
- [35] Huang, C.H. (1999). Phases and characteristics of Taiwan business cycles : *A Markov switching analysis*, Taiwan Economic Review, 27, 185 – 214.
- [36] Huang, Y.-L., C.-M. Kuan, and K.S. Lin (1998). *Identifying the turning points of business cycles and forecasting real GNP growth rates in Taiwan (in Chinese)*, Taiwan Economic Review, 26, 431 – 457.
- [37] Jacquier, E., N.G. Polson, and P. Rossi (1994). *Bayesian analysis of stochastic volatility models (with discussion)*, Journal of Business & Economic Statistics, 12, 371 – 417.
- [38] kenida.A,and M.kedida,and CH.CHEkai,Chaine de Markov.6 – 21,24 – 27.
- [39] Kim, W.Y., Billard, L., 1990. *Asymptotic properties for the first-order bilinear time series model*. Comm. Statist. Theory Methods 19,1171 – 1183.
- [40] Kingman.J.F.C, *The ergodic theory of subadditive stochastic processes*, J. Roy. Statist. Soc.Ser. B 30 (1968).499–510.
- [41] Lam, P.S. (1990). *The Hamilton model with a general autoregressive component* Journal of Monetary Economics, 26, 409 – 432.
- [42] Lee.O, *Probabilistic properties of a nonlinear ARMA process with Markov switching*,Commun. Stat. Theory Methods 34(1) (2005).193 – 204.
- [43] Lin, C.-C., M.-W. Hung, and C.-M. Kuan (2002). The dynamic behavior of short term interest rates in Taiwan : *An application of the regime switching model (in Chinese)*, Academia Economic Papers, 30, 29 – 55, 2002.
- [44] Liu.J.C, *Stationarity of a Markov-switching GARCH model*, J. Financ. Econ. 4(4) (2006).573 – 593

- [45] Liu, J., 1988. *On the general bilinear time series model*. J. Appl. Probab. 25, 553 – 564.
- [46] Liu, J., 1990. *Estimation for some bilinear time series*. Stoch. Models 6, 649 – 665.
- [47] Melino, A. and S.M. Turnbull (1990). *Pricing foreign currency options with stochastic volatility*, Journal of Econometrics, 45, 239 – 265.
- [48] Pham, D.T., Tran, L.T., 1981. *On the first-order bilinear time series model*. J. Appl. Probab. 18, 617 – 627.
- [49] Rau, H.-H., H.-W. Lin, and M.-Y. Li (2001). *Examining Taiwan's business cycle via two-period MS Models (in Chinese)*, Academia Economic Papers, forthcoming.
- [50] Stelzer.R. , *On Markov-switching ARMA processes—stationarity, existence of moments, and geometric ergodicity*,Econ. Theory 25(1) (2009).43 – 62.
- [51] Sola, M. and J. Driffill (1994). *Testing the term structure of interest rates using a stationary vector autoregression with regime switching*, Journal of Economic Dynamics and Control, 18, 601 – 628.
- [52] Tsay, R.S. (1987), "*Conditional heteroskedasticity time series analysis*", Journal of the American Statistical Association, 82, 590 – 604.
- [53] Wittwer, G., 1989. *Some remarks on bilinear time series models*. Statistics 20, 521 – 529.
- [54] Yao.J.-F., and J.-G. Attali, *On stability of nonlinear AR processes with Markov switching*, Adv.Appl. Prob. 32(2) (2000).394 – 407.

Conclusion

Dans cette thèse, nous avons analysé la structure probabiliste des modèles MSAR. Tout d'abord, des conditions suffisantes pour l'existence des solutions strictement stationnaires sont données pour le modèle MSAR général. Ensuite, nous avons prouvé l'existence de la stationnarité au second ordre et l'existence des moments d'ordre supérieur. Par ailleurs, nous étudions la question de l'estimation des paramètres pour un processus MSAR stationnaire et ergodique dans lequel nous permettons aux coefficients de varier selon une chaîne de Markov homogène et non observable avec un espace d'état fini.

Résumé.

Grâce à un large potentiel d'application couvrant notamment l'économie et la finance, plusieurs modèles linéaires à coefficients dépendant du temps sont récemment apparus dans la littérature statistique des séries temporelles. Dans cette thèse, nous considérons le modèle autorégressif à changements de régimes markoviens. Nous étudions certaines propriétés probabilistes des modèles, ces modèles, les paramètres peuvent dépendre d'une chaîne de Markov stationnaire non observable à espace d'état fini, discret et homogène. Nous donnons des conditions assurant l'existence des solutions ergodiques stationnaires (dans certains sens), ainsi l'existence des moments d'ordre supérieur finis. Le problème de l'estimation des paramètres est également traité par une approche basée sur la fonction de vraisemblance.

Mots-clés.

Modèle autorégressifs à changements de régimes markoviens, stationnarité, causalité, ergodique, moments d'ordre supérieur, estimation par la méthode de la maximum vraisemblance.

Abstract.

Thanks to their possible application to a wide variety of fields including economics and finance, many linear time series models with time-dependent coefficients have recently paid attention in the statistical literature. This paper is devoted to the Markov switching autoregressive time series models. Conditions ensuring the existence of a unique, strictly stationary, causal and ergodic solution and the existence of higher-order moments are given. The problem of estimating the parameters is also investigated through an approach based on maximum likelihood estimation.

Keywords.

Markov switching autoregressive models, stationary, causal, ergodic higher-order moments, maximum likelihood estimation.