

République Algérienne Démocratique et Populaire
Ministère de l' Enseignement Supérieur
et de la Recherche Scientifique
Université Mohammed Seddik Ben Yahia Jijel
Faculté des Sciences Exactes et Informatique



Mémoire

Présenté pour obtenir le diplôme de Master

En Mathématiques fondamentales

Spécialité : Probabilités et Statistique :

Thème :

Estimation des Modèles Autoregréssifs à Changement de Régime Markovein

Par :

Zeghib Ilham

Stihi Nadjat

Soutenue publiquement le : **25/10/2020** Devant le jury composé de :

Mme.Djeridi Zohra	M.C.B	Université de Jijel	Président
Mme.Sellami Nawel	M.A.A	Université de Jijel	Encadreur
Mme.Ghouil Djoweyda	M.A.A	Université de Jijel	Examineur

Année universitaire : 2019 \ 2020

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

**Remerciement **

En tout premier lieu, nous remercions Dieu le tout puissant, de nous avoir donné la force, la santé pour mener bien notre travail volenté.

Un spécial remerciement à notre encadreur **Mme.Sellami Nawel** pour l'intérêt qu'elle a porté à notre travail et pour la subtilité de ses orientations, ces précieux conseils, sa rigueur scientifique, ces corrections attentives et ses discussions fructueuse.

Mme. Sellami merci également pour votre présence continue et pour tous les efforts et progrès que nous avons accomplis. Ce travail est un témoignage de notre gratitude envers vous. Nous nous souviendrons de votre soutien indéfectible et de votre patience inaltérable. Merci encore une fois.

Nous remercions chaleureusement notre cher enseignant, monsieur **Gherda Mebrouk** pour son aide inestimable, sa disponibilité, ces conseils avisés, ses compétences, sa bienveillance et sa sympathie.

M.Gherda nous vous remercions sincèrement pour votre soutien et votre gentillesse, nous gardons d'excellents souvenirs avec vous.

Nous tenons à remercier les membres des examinatrices **Mme Djeridi Zohra, Mme Ghouil Djo-
weyda** d'avoir accepté de juger ce travail.

Un sincère et honnête merci à nos parents, nos familles et nos proches qui nous ont soutenus et encouragé tout au long de la réalisation de ce mémoire.

Nous tenons également à remercier toutes les personnes qui ont contribué de près ou de loin à la réalisation de ce travail.

**Résumé **

Les caractères non linéaires tels que le changement de dynamique, présentent dans la plupart des phénomènes économiques et financières, dus à l'hétérogénéité entre les observations, font que leur traitement par les modèles linéaires devient inapproprié et le recours aux modèles non linéaires devient indispensable. Le modèle auto-régressif à changement de régime Markovien (MS – AR) est considéré comme l'un des modèles non linéaires qui permet de capter au mieux ces effets non linéaires. La condition nécessaire pour l'efficacité de cette modélisation repose sur la fiabilité de l'estimation des paramètres du modèle ; La méthode d'estimation retenue dans notre étude est basée sur la méthode du maximum de vraisemblance en utilisant l'algorithme EM, à partir des conditions on obtient des estimateurs consistants et asymptotiquement.

Mots clés : Modèle à changement de régime MS–AR, méthode MV, algorithme EM, les chaînes de Markov.

Table des matières

Introduction générale	2
1 Les chaînes de Markov à temps discret	6
1.1 Définitions et exemples :	6
1.2 Caractérisations d'une chaîne de Markov homogène :	8
1.3 La distribution des états d'une chaîne de Markov :	11
1.4 Stationnarité :	12
1.5 Classification des états :	14
1.6 Comportement asymptotique	15
2 Les modèles auto-régressifs à changement de régimes Markoviens MS-AR	17
2.1 Revue littéraire sur les modèles à changement de régimes Markoviens :	18
2.2 Définition et propriétés :	20
2.2.1 Processus auto-régressifs d'ordre p (AR(p)) :	20
2.3 Présentation des modèles MS :	21
2.4 Modèles MS en moyenne-variance :	21
2.5 Les modèles auto-régressifs à changements de régimes Markoviens :	22
2.5.1 Propriétés des modèles MS-AR :	23
3 L'estimation dans les modèles MS-AR	25
3.1 Probabilités filtrées et lissées :	25
3.1.1 Probabilités filtrées :	25
3.1.2 Probabilités lissée :	26
3.2 L'estimation dans les modèles MS-AR :	27
3.2.1 Fonction de vraisemblance :	27

3.2.2	Algorithme espérance-maximisation (<i>EM</i>) :	27
3.3	Estimation des modèles MS(2)-AR(1) via l'algorithme EM :	29
3.4	Exemple d'application : modèle MS-AR en économie :	32
3.4.1	Application à la série du produit national brut (PNB) Français :	32
3.4.2	Formulation par le modèle MS(2)-AR(1) :	33
3.4.3	Estimation du modèle MS(2)-AR(1) :	33
	Conclusion générale	34
	Bibliographie	35

Introduction générale

L'objectif des méthodes statistiques et probabilistes modernes est de concevoir les outils et les bases théoriques qui permettent l'analyse des phénomènes stochastiques complexes que l'on retrouve dans les différents domaines tels que l'économie, la finance, la physique, le traitement d'image, la biologie... Dans ce sens, la modélisation d'une série temporelle, comme étant la trajectoire d'un phénomène aléatoire, vise principalement à analyser et à comprendre le comportement implicite de ce phénomène qui ne peut pas être détecté, en prenant en considération l'effet de l'auto-corrélation. La recherche dans cette discipline est aujourd'hui très avancée et enregistre des résultats prometteurs. Afin de mieux décrire un phénomène aléatoire, il devient nécessaire de construire le modèle qui serait capable de capter ses faits stylisés. Si on étudie les faits récurrents des phénomènes économiques et financiers, on décrira la forte volatilité, le regroupement de volatilité, la leptokurticité, la mémoire longue, l'asymétrie, le changement de dynamique, les queues des épaisses... Ces faits stylisés représentent les grandes causes de la non linéarité des séries économiques et financières ; pour cela le traitement de ces phénomènes par les modèles linéaires, développés principalement par Box et Jenkins devient inapproprié et la recherche des modèles non linéaires qui seraient aptes à prendre en charge ces effets est devenue ainsi indispensable.

Le développement des modèles non linéaires qui répondent à ces faits stylisés, constitue aujourd'hui un centre d'intérêt pour les économétriciens et les statisticiens. On en retrouve trois grandes catégories :

(i) les modèles de la volatilité stochastique tels que les modèles ARCH initialisés par Engel[18] (Autoregressive Conditional Heteroskedasticity) et leur généralisation par les modèles GARCH étudiés par Bollerslev [4]. Ces modèles étudient les problèmes de la forte volatilité et le regroupement de cette dernière (Clustering volatility), et sont présents beaucoup plus dans les séries financières.

(ii) les modèles ARFIMA (Autoregressive Fractionnaly Integrated Moving Average) et ses variantes qui étudient le caractère de mémoire longue.

(iii) les modèles à changement de régimes qui visent tous à comprendre le caractère asymétrique du

changement dans la dynamique du phénomène. Ces derniers regroupent deux types : Les modèles à changement de régimes déterministes introduits par Tong[60], tels que les modèles (Threshold Autoregressive), et les modèles (Self Excited Threshold Autoregressive), qui reposent tous les deux sur la définition d'une variable de seuil connue, et les modèles à changement de régimes aléatoires caractérisés par la méconnaissance du moment du changement de régime.

Les modèles à changement de régimes aléatoires regroupent plusieurs variantes : (i) les Modèles de Markov Cachés (Hidden Markov Models) qui supposent que la séquence d'observations dépend d'une séquence d'état non observée cachée tirée d'une chaîne de Markov, (ii) les modèles de mélange de distributions (Mixtures Models) qui permettent de prendre en compte l'effet de l'hétérogénéité dans la structure des observations en la départageant en sous groupes indépendants entre eux du fait que le processus générateur du changement de régime est un processus indépendant, et enfin, (iii) les modèles auto-régressifs à changement de régimes Markoviens (Markov Switching Autoregressive) qui permettent de capter la majorité des effets non linéaires ainsi que la structure autocorrélée qui existe dans la plupart des séries économiques et financières (Bahar et Hamouri[3], Hamilton[29, 30, 31], Goldfeld et Quandt[25], McCulloch et Tsay [43], Schnatter[55], Gray [27]. . .).

Les modèles auto-régressifs à changement de régimes Markoviens servent à modéliser, les séries qui se comportent d'une manière non linéaire telles que les séries des produits financières dont l'une des caractéristiques est l'alternance entre plusieurs régimes (états) représentant les différents comportements (Les changements brutaux enregistrés dans les séries lors des crises, ... font que cette alternance est aléatoire, d'où la présence d'une série non observée qui doit être traitée par une chaîne de Markov et estimée au même titre que les autres paramètres du modèle).

Dans notre étude, nous traitons le problème de l'estimation des paramètres des modèles non linéaires en raison de leur degré de représentativité des faits stylisés non linéaires qui existent dans les domaines de l'économie et de la finance.

Une grande majorité des utilisateurs des modèles n'accordent pas l'importance à la performance que nécessite la méthode d'estimation des paramètres de ce modèle dont dépend la fiabilité des résultats, tant et si bien que plus les paramètres du modèle sont bien estimés, plus les faits stylisés sont bien détectés, plus le modèle aura un pouvoir prédictif et explicatif fortement représentatif du phénomène étudié qui va assurer une interprétation plus pertinente des résultats. Cependant, plus les paramètres sont mal estimés, plus les faits sont moins bien détectés, ce qui se traduirait par une interprétation erronée et imprécise des résultats obtenus. La structure non linéaire du modèle fait que l'estimation des paramètres du modèle par la maximisation analytique de la fonction de vraisemblance devient difficile, ce qui nécessite le recours à des méthodes itératives qui convergent vers le maximum de cette

dernière. On met alors l'accent sur la méthode la plus utilisée connue sous le nom " l'algorithme EM " (Expectation-Maximisation). Dempster et al [15]. Cet algorithme est utilisé par la plupart des économétriciens notamment pour les modèles en raison de sa simplicité et sa rapidité.

Notre mémoire est subdivisée en trois chapitres structurés comme suit :

Le premier chapitre introduit des notions sur les chaînes de Markov à temps discret qui sont nécessaires à la compréhension des chapitres suivants.

Le deuxième chapitre inclut des modèles auto-régressifs (AR), les modèles auto-régressifs à changement de régimes markoviens (MS-AR) et leurs propriétés nécessaires pour la suite de notre travail.

Dans **le dernier chapitre** nous présentons l'estimation des modèles MS-AR par la méthode (MV) en utilisant l'algorithme EM et nous donnons un exemple de la série PNB (produit national brut).

Les chaînes de Markov à temps discret

Les chaînes de Markov sont des suites des variables aléatoires caractérisées par un aspect particulier de dépendance, qui leur attribue des propriétés singulières et un rôle important en modélisation par exemple en biologie, en physique, en économie ou en recherche opérationnelle. Elles ont été introduites par Andrey Markov en 1906.

On définit un processus qu'il est markovien si son état actuel fournit toute l'information nécessaire pour connaître son évolution future. Donc sa distribution dans le futur étant donné le présent et le passé ne dépend que du présent. Dans un tel contexte, il est intéressant de savoir quelle est la probabilité d'être à une position donnée après un très long moment, ou encore quelle proportion moyenne de temps est passée à cette position sur une longue période de temps. C'est le type de questions qu'on se pose souvent lorsque on est en présence d'une chaîne de Markov.

Ce chapitre présente les principales notions sur les points essentiels que nous utilisons tout au long de ce mémoire.

1.1 Définitions et exemples :

Soit S_0, \dots, S_n, \dots une suite de variables aléatoires définies sur un même espace de probabilité (Ω, A, P) et à valeurs dans E . Soient E un ensemble fini ou dénombrable et $S = \{S_n\}_{n \in \mathbb{N}}$ un processus stochastique à valeurs dans E .

Définition : 1.1.1

On dit que S est une chaîne de Markov si pour tout $n \in \mathbb{N}$, et pour tout $s_0, s_1, \dots, s_{n+1}, \in E$ telle que $P(S_0 = s_0, S_1 = s_1, \dots, S_n = s_n) > 0$, On a :

$$P(S_{n+1} = s_{n+1} / S_n = s_n, S_{n-1} = s_{n-1}, \dots, S_0 = s_0) = P(S_{n+1} = s_{n+1} / S_n = s_n) \quad (1.1)$$

E est appelé espace d'états. La relation (1.1) est appelée markovienne.

Remarque : 1.1.1

La propriété markovienne peut aussi s'écrire pour tout $m, n \in \mathbb{N}$ et $(S_i)_{i \in \mathbb{N}}$ à valeurs dans E , sous la forme :

$$P(S_{m+n} = s_{m+n} / S_m = s_m, S_{m-1} = s_{m-1}, \dots, S_0 = s_0) = P(S_{m+n} = s_{m+n} / S_m = s_m) \quad (1.2)$$

Ainsi, on peut dire d'un processus S qu'il est markovien si son état actuel fournit toute l'information nécessaire pour connaître son évolution future. Sa distribution dans le futur étant donné le présent et le passé ne dépend que du présent. On parle alors d'absence de mémoire.

Définition : 1.1.2

Une chaîne de Markov est dite homogène lorsque la probabilité de transition (1.2) ne dépend pas de n , c'est à dire :

$$P(S_{n+1} = j / S_n = i) = P(S_1 = j / S_0 = i) = p_{ij} \text{ pour tout } n \in \mathbb{N} \text{ et } i, j \in E. \quad (1.3)$$

L'homogénéité d'une chaîne de Markov précise donc que la probabilité de passer l'état i à j reste la même à travers le temps.

Définition : 1.1.3

Soit $S = (S_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène d'espace d'états E . La probabilité de transition de l'état i à l'état j est :

$$p_{ij} = P(S_{n+1} = j / S_n = i) = P(S_1 = j / S_0 = i), \forall n \in \mathbb{N} \quad (1.4)$$

On appelle matrice de transition de la chaîne S , la famille $P = (p_{ij})_{i, j \in E}$, telle que : $0 \leq p_{ij} \leq 1$ et $\sum_{j \in E} p_{ij} = 1, \forall i \in E$. Elle est dite stochastique.

Exemple 1.1.1 :

l'évolution de la météo d'un jour à l'autre. La pluie sur une ville donnée a été, convenablement, caractérisée ou décrite par les chaînes de Markov à deux états :

- L'état 1 représente les jours secs.
- L'état 2 simule les jours pluvieux.

Sous les deux hypothèses suivantes :

$$\begin{cases} P(\text{secs demain} / \text{secs aujourd'hui}) = 0,8 = p_{11}. \\ P(\text{secs demain} / \text{pluie aujourd'hui}) = 0,6 = p_{21}. \end{cases}$$

On définit la variable aléatoire S_n par :

$$S_n = \begin{cases} 1 & \text{secs le jour } n. \\ 2 & \text{pluie le jour } n. \end{cases}$$

La suite $(S_n)_{n \in \mathbb{N}}$ est une chaîne de Markov à temps discret sur les états 1 et 2 dont la matrice de transition P est donnée par :

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$$

On a :

$$\begin{cases} p_{11} + p_{12} = 1 \\ p_{21} + p_{22} = 1 \end{cases} \Rightarrow \begin{cases} p_{12} = 1 - 0.8 = 0.2 \\ p_{22} = 1 - 0.6 = 0.4 \end{cases}$$

Alors :

$$P = \begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix}$$

La figure (1.1) est une représentation graphique de la chaîne de Markov définie en exemple (1.1.1). Les états sont représentés par des cercles numérotés et les probabilités de transitions positives par des flèches.

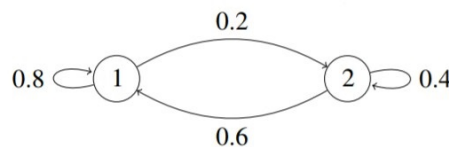


FIGURE 1.1 :

1.2 Caractérisations d'une chaîne de Markov homogène :

Dans ce qui suit, nous considérons que les chaînes de Markov sont homogènes à espace d'états E fini, de matrice de transition P .

Définition : 1.2.1

La matrice de transition à k pas de la chaîne S notée $P^{(k)} = (p_{ij}^{(k)})_{ij}$ pour tout $k, n \in \mathbb{N}$ et $i, j \in E$ est la matrice d'éléments :

$$p_{ij}^{(k)} = P(S_{n+k} = j / S_n = i) = P(S_k = j / S_0 = i). \quad (1.5)$$

La matrice de transition à k pas nous donne la probabilité qu'après un temps k on soit dans l'état j sachant que l'on était initialement dans l'état i .

Proposition : 1.2.1

Quelques propriétés de matrice de transition à k pas :

1. $0 \leq p_{ij}^{(k)} \leq 1$.
2. $\sum_{j \in E} p_{ij}^{(k)} = 1$ pour tout $i \in E$.

Démonstration :

1. La première propriété est immédiate car $p_{ij}^{(k)}$ est une probabilité.
2. Pour la deuxième propriété :

$$\begin{aligned} \sum_{j \in E} p_{ij}^{(k)} &= \sum_{j \in E} P(S_{n+k} = j / S_n = i) \\ &= \sum_{j \in E} \frac{P(S_{n+k} = j, S_n = i)}{P(S_n = i)} \\ &= \frac{P(\bigcup_{j \in E} \{S_{n+k} = j\}, S_n = i)}{P(S_n = i)} \\ &= \frac{P(S_n = i)}{P(S_n = i)} \\ &= 1. \end{aligned}$$

Car les évènements $\{S_{n+k} = j\}_{j \in E}$ sont une partition disjoint de E et donc :

$$P(S_n = i) = \sum_{j \in E} P(S_{n+k} = j, S_n = i)$$

Théorème de Chapman-Kolmogorov : Pour tous état $i, j \in E$, pour tout $n, m \in \mathbb{N}$ et $k \in [0, n]$, on a l'égalité :

$$p_{ij}^{(n+m)} = \sum_{k \in E} p_{ik}^{(n)} p_{kj}^{(m)} \quad (1.6)$$

En notation matricielle, on a :

$$P^{(n+m)} = P^{(n)} P^{(m)} \quad (1.7)$$

Démonstration : Soient k l'état de la chaîne au temps m , on a :

$$\begin{aligned}
 p_{ij}^{(n+m)} &= P(S_{n+m} = j / S_0 = i) \\
 &= P\left(\bigcup_{k \in E} \{S_m = k\}, S_{n+m} = j / S_0 = i\right) \\
 &= P\left(\bigcup_{k \in E} (S_m = k, S_{n+m} = j) / S_0 = i\right) \\
 &= \sum_{k \in E} P(S_m = k, S_{n+m} = j / S_0 = i) \\
 &= \sum_{k \in E} \frac{P(S_m = k, S_{n+m} = j, S_0 = i)}{P(S_0 = i)} \\
 &= \sum_{k \in E} \frac{P(S_{n+m} = j / S_m = k, S_0 = i) \cdot P(S_m = k, S_0 = i)}{P(S_0 = i)} \\
 &= \sum_{k \in E} \frac{P(S_{n+m} = j / S_m = k, S_0 = i) \cdot P(S_m = k / S_0 = i) P(S_0 = i)}{P(S_0 = i)} \\
 &= \sum_{k \in E} P(S_{n+m} = j / S_m = k, S_0 = i) \cdot P(S_m = k / S_0 = i) \\
 &= \sum_{k \in E} \frac{P(S_{n+m} = j, S_m = k / S_0 = i) \cdot P(S_m = k / S_0 = i)}{P(S_m = k / S_0 = i)}
 \end{aligned}$$

$$\left(\text{puisque } P(A \cap B / C) = P(B / C) \cdot P(A / B \cap C) \implies P(A / B \cap C) = \frac{P(A \cap B / C)}{P(B / C)} \right)$$

$$\begin{aligned}
 &= \sum_{k \in E} P(S_{n+m} = j, S_m = k / S_0 = i) \\
 &= \sum_{k \in E} P(S_{n+m} = j / S_m = k, S_0 = i) \cdot P(S_m = k / S_0 = i) \\
 &= \sum_{k \in E} P(S_{n+m} = j / S_m = k) \cdot P(S_m = k / S_0 = i) \\
 &= \sum_{k \in E} p_{kj}^{(n)} \cdot p_{ik}^{(m)} = \sum_{k \in E} p_{ik}^{(m)} \cdot p_{kj}^{(n)}.
 \end{aligned}$$

Exemple de la météo :

Calculer $P^{(3)}$ et $P^{(4)}$.

D'après le Théorème de Chapman-Kolmogorov, on a :

$$P^{(3)} = P^{(1+2)} = P.P^{(2)} = \begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix} \begin{pmatrix} 0.752 & 0.248 \\ 0.744 & 0.256 \end{pmatrix}, P^{(3)} = \begin{pmatrix} 0.752 & 0.248 \\ 0.744 & 0.256 \end{pmatrix}$$

$$P^{(4)} = P^{(1+3)} = P.P^{(3)} = \begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix} \begin{pmatrix} 0.752 & 0.248 \\ 0.744 & 0.256 \end{pmatrix}, P^{(4)} = \begin{pmatrix} 0.75 & 0.24 \\ 0.749 & 0.251 \end{pmatrix}$$

1.3 La distribution des états d'une chaîne de Markov :

La distribution des états des chaînes de Markov après n transitions est notée $\pi_i^{(n)}$. Cette distribution est un vecteur de probabilités contenant la loi de la variable aléatoire S_n .

$$\pi_i^{(n)} = P[S_n = i] \quad \forall i \in E.$$

La distribution $\pi^{(n)}$ dépend de la matrice de transition P mais également de l'état dans lequel le processus a commencé son évolution. De manière générale, cet état est choisi selon une distribution initiale $\pi^{(0)}$.

Remarque : 1.3.1

Si l'état initial est connu avec certitude et est égal à i , on a simplement $\pi_i^{(0)} = 1$ et $\pi_j^{(0)} = 0$ pour tout $j \neq i$.

Théorème : 1.3.1

Soit P la matrice de transition d'une chaîne de Markov et $\pi^{(0)}$ la distribution de son état initial.

Pour tout $n \geq 1$, on a :

$$\pi^{(n)} = \pi^{(n-1)}P.$$

et

$$\pi^{(n)} = \pi^{(0)}P^n.$$

Démonstration :

Premièrement, on a pour tout $j \in E$,

$$\begin{aligned} \pi_j^{(1)} &= P[S_1 = j, S_0 = i] = \sum_{i \in E} P[S_1 = j/S_0 = i] P[S_0 = i] \\ &= \sum_{i \in E} p_{ij} \pi_i^{(0)} = \sum_{i \in E} \pi_i^{(0)} p_{ij}. \end{aligned}$$

et

$$\pi^{(1)} = \pi^{(0)}P.$$

La chaîne étant homogène, on obtient immédiatement le premier résultat :

$$\pi^{(n)} = \pi^{(n-1)}P \quad \forall n \geq 1.$$

Pour démontrer le second il suffit de résoudre l'équation de récurrence précédente par substitution.

Exemple 1.3.1 :

Le vecteur propre est composé des solutions du système $:(P - \lambda I_2)\nu$, avec $:\lambda = 1$ et sachant que :

$$\pi_1 + \pi_2 = 1 \text{ et donc : } V_1 + V_2 = 1.$$

Dans le cas présent on a :

$$\left(\left(\begin{pmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right)$$

Avec $:\lambda = 1$ et sous forme de système cela donne :

$$\begin{cases} -0.1\nu_1 + 0.3\nu_2 = 0. \\ 0.1\nu_1 - 0.3\nu_2 = 0. \end{cases}$$

Or on sait que $:\nu_1 = 1 - \nu_2$, et donc $:0.1(1 - \nu_1) - 0.3\nu_2 = 0$ ce qui donne :

$$0.1 - 0.4\nu_2 = 0 \iff \nu_2 = 0.1/0.4 = 0.25 = \pi_2$$

On en déduit alors :

$$\nu_1 = 1 - \nu_2 = 0.75 = \pi_1$$

1.4 Stationnarité :

- S'il existe une mesure $\pi = (\pi)_{i \in E}$ telle que $:\pi P = \pi$, alors π est une mesure stationnaire de S_t . Mesure stationnaire et distribution invariante sont des synonymes.
- Si $\forall i \in E, \pi \geq 0$ et $\sum_i \pi_i = 1$, alors π représente la loi stationnaire de S_t .
- S_t est un processus stationnaire si sa loi initiale, i.e. la loi de S_0 , est stationnaire.

Proposition : 1.4.1

La matrice de transition et la distribution de S_0 déterminent complètement la chaîne S , c'est-à-dire toutes les distributions et les distributions conjointes des variables S_0, \dots, S_n .

Démonstration :

1. Calcul : $P(S_n = j)$

On suppose que : $E = 1, \dots, M$, puisque E est fini alors :

$$\begin{aligned}
 P(S_n = j) &= P(\{S_n = j\} \cap \{S_0 = 1, S_0 = 2, \dots, S_0 = M\}) \\
 &= P(\{S_n = j\}, \bigcup_{i=1}^M \{S_0 = i\}) \\
 &= P(\bigcup_{i=1}^M [\{S_n = j, S_0 = i\}]) \\
 &= \sum_{i=1}^M P(S_n = j, S_0 = i) \\
 &= P(S_n = j, S_0 = 1) + P(S_n = j, S_0 = 2) + \dots + P(S_n = j, S_0 = M) \\
 &= P(S_n = j/S_0 = 1)P(S_0 = 1) + P(S_n = j/S_0 = 2)P(S_0 = 2) + \dots + P(S_n = j/S_0 = M)P(S_0 = M) \\
 &= p_{1j}^{(n)}P(S_0 = 1) + p_{2j}^{(n)}P(S_0 = 2) + \dots + p_{Mj}^{(n)}P(S_0 = M) \\
 &= \pi_1 P_{1j}^{(n)} + \dots + \pi_M P_{Mj}^{(n)}
 \end{aligned}$$

2. Calcul : $P(S_n = i_n, S_{n-1} = i_{n-1}, \dots, S_0 = i_0)$

$$\begin{aligned}
 P(S_n = i_n, S_{n-1} = i_{n-1}, \dots, S_0 = i_0) &= P(S_n = i_n/S_{n-1} = i_{n-1}, \dots, S_0 = i_0)P(S_{n-1} = i_{n-1}, \dots, S_0 = i_0) \\
 &= P(S_n = i_n/S_{n-1} = i_{n-1})P(S_{n-1} = i_{n-1}/S_{n-2} = i_{n-2}, \dots, S_0 = i_0)P(S_{n-2} = i_{n-2}, \dots, S_0 = i_0) \\
 &= p_{i_{n-1}, i_n} p_{i_{n-2}, i_{n-1}} \dots p_{i_0, i_1} P(S_0 = i_0)
 \end{aligned}$$

La dernière égalité est obtenue par récurrence sur l'entier $n \geq 0$, et ce pour tous les états i_0, \dots, i_n .

Exemple 1.4.1 :

Supposons que $P(S_0 = 1) = 0.3$ et $P(S_0 = 2) = 0.7$

Quelle est la probabilité qu'il pleuve le jour 3, $P(S_3 = 2) = ?$

$$\begin{aligned}
 P(S_3 = 2) &= P_{12}^{(3)}P(S_0 = 1) + P_{22}^{(3)}P(S_0 = 2) \\
 &= 0.248x * 0.3 + 0.256 * 0.7 \\
 &= 0.0744 + 0.1792 \\
 &= 0.2536
 \end{aligned}$$

1.5 Classification des états :

Il y a plusieurs type d'états, la classification des états d'une chaîne de Markov joue un rôle essentiel dans l'étude de son comportement à long terme.

Définition : 1.5.1

a) **Un état i est récurrent si :**

$p_{ii} = P(\text{retour } i/\text{depart de } i) = 1$. sinon, ($p_{ii} < 1$). Alors l'état est transient.

b) **Un état récurrent i est récurrent nul si :**

$E(\text{temps de premier retour } i/\text{départ de } i) = +\infty$. Sinon, l'état i est récurrent positif.

c) **La période $d(i)$ d'un état i est définie par**

$$d(i) = \text{pgcd} \left\{ n \geq 1, p_{ii}^{(n)} > 0 \right\}$$

Un état i est périodique si $d(i) > 1$.

Un état i est apériodique si $d(i) = 1$.

d) **Accessibilité d'état :** Un état j est accessible à partir de l'état i si et seulement s'il existe $n \geq 0$ tel que $p_{ij}^{(n)} > 0$.

e) **Communication d'état :** Les états i et j communiquent si l'état i est accessible à partir de l'état j et l'état j est accessible partir de l'état i , ce qui est noté $i \longleftrightarrow j$ i.e : $\exists n \geq 0$ et $m \geq 0$ tels que : $P_{ij}^{(n)} > 0$ et $P_{ji}^{(m)} > 0$.

f) **Irréductibilité d'état :** Une chaine est dite irréductible si tous ses états communiquent ; autrement dit, la chaine forme une seule classe. L'irréductibilité permet un passage de n 'importe quel état de la chaine vers un autre mais pas nécessairement en une seule étape.

g) **Ergodicité d'état :** Un état i est ergodique s'il est apériodique et de récurrence positive. Si $\{S_n\}_{n \in \mathbb{N}}$ est irréductible et que tous ses états sont ergodiques, alors S_n est ergodique. -L'ergodicité de la chaîne de Markov est une propriété importante, elle assure que la distribution de S_n converge vers une distribution invariante quelque soit l'état initial S_0 .

Partition des états :

L'espace des états d'une chaîne de Markov peut être décomposé en classes d'états de même type et de même période. De plus, certains types d'état sont exclus si le nombre d'états dans la classe est fini.

Exemple de météo :

Les états 1 et 2 communiquent

$$\begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix} \iff p_{ij} \geq 0, \forall i, j \in \{1, 2\} = E$$

Définition : 1.5.2

- * Une classe est un sous ensemble de tous les états communiquant entre eux.
- * Une chaîne de Markov est irréductible si elle ne comporte qu'une seule classe.

Exemple de météo :

Les états 1 et 2 communiquent, la chaîne de Markov est irréductible.

1.6 Comportement asymptotique

L'étude du comportement à long terme d'une chaîne de Markov cherche à répondre des questions aussi diverses que :

- La distribution $\pi^{(n)}$ converge-t-elle lorsque $n \rightarrow \infty$?
- Si la distribution $\pi^{(n)}$ converge lorsque $n \rightarrow \infty$ quelle est la limite π^* et cette limite est-elle indépendante de la distribution initiale $\pi^{(0)}$?

Une distribution π est invariante ou stationnaire si

$$\pi = \pi P.$$

Propriétés: 1.6.1

Si $\lim_{n \rightarrow \infty} \pi^{(n)}$ existe, alors la limite est une distribution invariante.

Les deux résultats suivants sont cités sans démonstration.

Théorème : 1.6.1

Une chaîne de Markov possède toujours au moins une distribution invariante.

Théorème : 1.6.2

Une chaîne de Markov possède autant de distributions invariantes linéairement indépendantes que la multiplicité de la valeur propre 1 de sa matrice de transition.

Théorème : 1.6.3

La distribution $\pi^{(n)}$ des états d'une chaîne de Markov converge vers une distribution (invariante) π^* indépendante de la distribution initiale $\pi^{(0)}$ si et seulement si la suite des puissances de la matrice de transition P de la chaîne converge vers une matrice (stochastique) P^* dont toutes les lignes sont égales entre elles. De plus, si tel est le cas chaque ligne de p est égale à π^* .

Remarque : 1.6.1

Si $\pi^* = \lim_{n \rightarrow \infty} \pi^{(n)}$, on parlera de distribution asymptotique, stationnaire ou invariante.

Le théorème suivant résume le comportement asymptotique des chaînes irréductibles et apériodiques.

Théorème : 1.6.4

Soit P la matrice de transition d'une chaîne irréductible apériodique. Les propriétés suivantes sont vérifiées.

- La matrice P^n tend vers une matrice stochastique P^* lorsque n tend vers l'infini.
- les lignes de P^* sont toutes égales entre elles.
- $p_{ij}^* > 0$ pour tout $i, j \in E$
- Pour toute distribution initiale $\pi^{(0)}$,

$$\lim_{n \rightarrow \infty} \pi^{(n)} = \lim_{n \rightarrow \infty} \pi^{(0)} P^n = \pi^*.$$

- π^* est la solution unique du système :

$$\begin{cases} \pi P = \pi \\ \pi 1 = 1 \end{cases}$$

- π^* est égal à n'importe quelle ligne de la matrice P^* .

Les modèles auto-régressifs à changement de régimes Markoviens MS-AR

Les modèles de changement de régime sont regroupés en deux grandes classes. Les modèles de la **L** première classe considèrent que le changement est caractérisé à travers des variables observables, ceux les plus connus d'entre eux sont les modèles *TAR* (Thershold Autoregressifs) et *SETAR* (Self Excited Autoregressif); quant à la deuxième classe, elle concerne les modèles qui considèrent que le régime ne peut pas être déterministe en lui attribuant le caractère du processus stochastique inobservable, ce qui implique qu'on ne connaît jamais avec certitude le régime en un temps donné; pour cela on assigne les régimes par des probabilités.

Comme leur nom l'indique, les modèles à changement de régimes Markoviens [*Markov Switching Models (MSM)*] appartiennent à la deuxième classe dans laquelle le changement de régime est fait à travers une chaîne de Markov cachée.

Notre objectif est d'étudier un cas particulier des Modèles de changement de régimes Markoviens appelés les modèles *MS-AR* [*Markov Switching Autoregressif*]. Dans ces modèles on prend en considération l'auto-corrélation qui existe dans la variable d'intérêt afin de modéliser les phénomènes qui ont, d'une part, une structure hétérogène dont la procédure consiste à les subdiviser en plusieurs groupes, et d'autre part, nécessitent de prendre en considération l'auto-corrélation qui existe dans la série.

Il existe toutefois d'autres types de modèles à changement de régimes, à savoir :

- ***HMSM (Hidden Markov Switching Models)*** : Ces modèles sont à la base de tous les modèles de Markov cachés [1] qui reposent sur deux séquences emboîtées : une séquence cachée $\{S_t\}_1^T$ décrite par une chaîne de Markov et une séquence observable $\{Y_t\}_1^T$ qui représente un ensemble de variables indépendantes entre elles conditionnellement aux états cachés.
- ***MM (Mixture Models ou Mixture of distributions)*** : Ce sont des modèles de mélange

de distributions [42],[55] qui ont le même principe que les HMM ; la seule différence réside dans la séquence inobservable $\{S_t\}_1^T$ qui est supposée suivre un processus indépendant. Ces modèles permettent principalement le traitement de l'hétérogénéité qui existe dans une population en la subdivisant en sous groupes homogènes. Cette procédure est un type d'apprentissage automatique.

- **TAR et SETAR (Threshold Autoregressifs et Self Excited Autoregressif)** : Ce sont des modèles non linéaires qui reposent sur un changement de régime non aléatoire opéré à travers un seuil fixé.

Dans ce chapitre on étudie le fondement théorique des modèles auto-régressifs à changement de régimes Markoviens $MS-AR$. Après une brève revue des principaux travaux sur ces modèles, on explique leurs principales propriétés.

2.1 Revue littéraire sur les modèles à changement de régimes Markoviens :

Dans cette section on présente les travaux qui ont contribué au développement de l'inférence des modèles à changement de régimes Markoviens (Markov Switching Models MSM).

1. **Quandt et Goldfeld [25]** : Ces auteurs sont les premiers à mettre sous une équation les modèles à changement de régimes [Switching Models]. Ils ont constaté que les modèles traditionnels linéaires ne pouvaient pas être utilisés dans la modélisation des phénomènes aléatoires. Dans leur article, ils ont traité principalement deux types de changement de régimes, à savoir :
 - Changement de régime déterministe (basé sur un indicateur temporel).
 - Changement de régime stochastique (fait à travers une séquence aléatoire mais indépendante).
 Ce n'est que depuis la découverte faite par Hamilton[29], que le modèle est devenu fortement utilisable.
2. **Hamilton[29]** : Hamilton a présenté dans son article une extension du modèle développé par Quandt et Goldfeld [25] selon deux voies :
 - Il a considéré l'effet auto-régressif dans le modèle.
 - Le processus inobservable générateur du régime suit une chaîne de Markov. Implicitement, il constitue un processus dépendant contrairement au modèle développé par Quandt et Goldfeld, au temps discret et dans un espace d'état discret.
 Il a démontré l'utilité des modèles $MS-AR$ dans la modélisation des phénomènes asymétriques

et a mis l'action sur ces modèles pour la modélisation du Cycle des Affaires (Business Cycle) Américain. Plus précisément, Hamilton a pu estimer les datations des points de retournement (turning points) de cycle des affaires et ces estimations étaient réellement proches de celles fixées par le NBER [National Bureau of Economic Research]. Depuis cette découverte, les modèles $MS-AR$ ont été utilisés sur un nombre important de travaux notamment dans le domaine de l'économie et de la finance.

3. **Hamilton** [30] : Par la suite et en 1990, Hamilton a démontré toute la procédure à suivre pour estimer les paramètres du modèle $MS-AR$ ainsi que les probabilités lissées (voir le détail plus loin) où il a utilisé l'algorithme le plus robuste parmi toutes les méthodes d'estimation numérique, appelé algorithme EM afin d'estimer récursivement les paramètres du modèle.
4. **McCulloch, et Tsay** [43] : Ces deux auteurs ont contribué au traitement Bayésien des modèles économétriques en général. Dans leurs travail sur l'estimation Bayésienne des MSM. Ils ont développé l'estimation des coefficients de ces modèles par l'algorithme de Gibbs.
5. **Kim et Nelson** [39] : Ils ont traité l'estimation Bayésienne des modèles à composantes inobservables tels que les modèles d'espace-état et les modèles à changement de régimes Markoviens appliqués à la modélisation des rendements des indices boursiers, des taux d'intérêt ainsi que la modélisation de Cycle des Affaires Américaines.
6. **Franses et Dijk** [21] : Leur ouvrage traite des fondements théoriques des modèles non linéaires tels que les MSM, les réseaux de neurone et les modèles à volatilité stochastique.
7. **Schnatter** [55] : Elle a présenté dans son ouvrage une encyclopédie sur les modèles de mélange de distributions (Mixture Models) ainsi que les MSM qui traitent l'estimation Bayésienne de ces derniers.

En plus des principaux précurseurs du modèle à changement de régimes cités précédemment, d'autres chercheurs ont eu à utiliser ces modèles dans la modélisation des phénomènes dans plusieurs domaines ; on cite principalement :

- [(i)] **Goodwin** [26], **Hamilton** [32], et **Medhioub** [44] : la modélisation des cycles des affaires de différents pays.
- [(ii)] **Moussi** [46] : l'estimation Bayésienne des prix de pétrole à travers les modèles de mélanges de distributions.
- [(iii)] **Dafas** [14] : la modélisation du prix spot du pétrole.
- [(iv)] **Gray** [27] : la modélisation du taux d'intérêt à court terme.
- [(v)] **Weron** [63] : dans la modélisation du prix spot d'électricité.

2.2 Définition et propriétés :

Avant d'entamer les modèles auto-régressifs à changements de régimes markoviens il est utile de faire quelques définitions sur les modèles auto-régressifs classiques.

2.2.1 Processus auto-régressifs d'ordre p ($AR(p)$) :

On appelle processus auto-régressif centré d'ordre p , noté $AR(p)$ un processus $(Y_t)_{t \in T}$ est centré vérifiant une relation du type :

$$\Phi(B)Y_t = \epsilon_t \tag{2.1}$$

Où $(\epsilon_t)_{t \in T}$ est un processus d'innovations qui satisfait les conditions $E(\epsilon_t) = 0$ et $E(\epsilon_t^2) = \sigma_\epsilon^2$ et $\Phi(B)$ est appelé polynôme caractéristique du degré p du processus $(Y_t)_{t \in T}$ défini tel que :

$$\begin{aligned} \Phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \\ &= 1 - \sum_{i=1}^p \phi_i B^i \end{aligned}$$

En outre, l'équation (2.1) peut être écrite sous la forme suivante :

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t. \tag{2.2}$$

Les caractéristiques :

a) Condition de stationnarité : Un processus $AR(p)$ est stationnaire si les racines de son équation caractéristique sont strictement supérieures à 1 en valeur absolue.

b) Les moments : Les différents moments d'un processus stationnaire sont :

$$\begin{aligned} E(Y_t) &= 0 \\ V(Y_t) &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma_\epsilon^2 \\ \gamma(h) = Cov(Y_t, Y_{t-h}) &= \phi_1 \gamma_{h-1} + \phi_2 \gamma_{h-2} + \dots + \phi_p \gamma_{h-p} \end{aligned}$$

c) Auto-corrélation : On montre que les auto - corrélations sont solutions des équations de Yule-walker

$$\rho_h = \sum_{i=1}^{\infty} \phi_i \rho_{h-i} \tag{2.3}$$

Les modèles AR, ARMA permettent de modéliser des processus stationnaires, i.e, dont les caractéristiques (moyenne, variance, etc) ne varient pas avec le temps. Dans certaines situations, ces modèles ne sont pas suffisants, on doit alors utiliser d'autres approches. Les modèles à changement de régime markovien sont classiquement utilisés pour modéliser un processus aléatoire ayant des caractéristiques (ou paramètres) qui changent au cours du temps.

2.3 Présentation des modèles MS :

- Un processus à changement de régime markovien (MS) est un processus bivarié $\{S_t, Y_t\}$, où
- $\{S_t\}_{t \in T}$: désigne une chaîne de Markov cachée du premier ordre à espace d'état fini $E = \{1, \dots, M\}$.
 - $\{Y_t\}_{t \in T}$: est un processus observé avec des valeurs dans un espace \mathbb{R}^d .

Dans la section suivante, on se concentre sur un cas simple des modèles MS populaire en économétrie financière.

2.4 Modèles MS en moyenne-variance :

Soit le processus $(Y_t)_{t \in T}$ suit un modèle à changement de régime markoviens en moyenne-variance, qu'est déterminé par :

$$Y_t = \mu_t(S_t) + \sigma_t(S_t)\epsilon_t \tag{2.4}$$

où $\epsilon_t \sim N(0, 1)$, $\mu_t(S_t)$ et $\sigma_t(S_t)$ dénotent des fonctions mesurables par rapport à une tribu engendrée par la variable aléatoire S_t , où les processus $(\epsilon_t)_t$ et $(S_t)_t$ sont indépendants et S_t une chaîne de Markov homogène, irréductible, apériodique.

On a : $\forall i, j P = (S_t = j / S_{t-1} = i, S_{t-2}, \dots, S_0) = P(S_t = j / S_{t-1} = i) = p_{ij}$. Où la matrice de transition $P = (p_{ij})_{1 \leq i, j \leq M}$ est stochastique, $0 \leq p_{ij} \leq 1$ et $\sum_{j=1}^M P_{ij} = 1$. Autrement dit, Y_t / S_t suit une loi normale de moyenne $\mu(S_t)$ et d'écart type $\sigma(S_t)$ ($Y_t / S_t \sim N(\mu(S_t), \sigma(S_t))$) alors la densité de probabilité conditionnelle de Y_t est :

$$f(Y_t / S_t = i, \theta) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(\frac{-(y_t - \mu_i)^2}{2\sigma_i^2}\right) \quad i = \{1, \dots, M\}.$$

Dans ce cas le vecteur des paramètres du modèle est : $\theta = (\{\mu_i, \sigma_i^2\})$, $i = \{1, \dots, M\}$.

2.5 Les modèles auto-régressifs à changements de régimes Markoviens :

On note par $(Y_t)_{t \in T}$ la série observable qui représente la variable d'intérêt. Le modèle auto-régressif à changement de régimes Markovien *MS-AR* considère que le comportement de la série Y_t dépend d'un processus inobservable $(S_t)_{t \in T}$ exogène générateur du régime à la date t qui est représentée par une chaîne de Markov cachée [47],[1].

Définition : 2.5.1

On définit le modèle auto-régressif à changement de régimes Markoviens *MS(M) - AR(p)* à travers un processus bivarié $(Y_t, S_t)_{t \in T}$ tel que :

- $\{S_t\}_1^T$: est un processus inobservable défini comme une chaîne de Markov homogène, elle même définie dans un espace d'état fini et discret $E = \{1, \dots, M\}$, dont est M le nombre de régimes de matrice de transition $P = (p_{ij})_{i,j \in E}$ telle que :

$$p_{ij} = P(S_t = j / S_{t-1} = i)$$

- $\{Y_t\}_1^T$: est un processus auto-régressif linéaire défini par :

$$\forall t \in T; Y_t = \sum_{i=1}^p \phi_i(S_t) Y_{t-i} + \epsilon_t \tag{2.5}$$

où $(\epsilon_t)_{t \in T}$ est un processus strictement stationnaire, les fonctions $\phi_i(S_t)$ dépendent d'une chaîne $(S_t)_{t \in T}$ de Markov homogène, irréductible, apériodique à espace d'état fini $E = \{1, \dots, M\}$ et indépendante de $(\epsilon_t)_{t \in T}$.

Le modèle (2.4) contient comme cas particuliers plusieurs cas des modèles intéressants ayant été étudiés dans la littérature en effet : Pour la commodité et l'élégance des résultats ,nous réécrivons le modèle (2.5) sous la forme vectorielle :

$$\bar{Y}_t = \Phi(S_t) \bar{Y}_{t-1} + \epsilon_t \tag{2.6}$$

avec

$$\bar{Y}_t = \begin{pmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p-1} \end{pmatrix} \Phi(S_t) = \begin{pmatrix} \phi_1(S_t) & \phi_2(S_t) & \cdots & \phi_p(S_t) & 0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \tag{2.7}$$

et

$$\varepsilon_t = \begin{pmatrix} \epsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

où \bar{Y}_t est un vecteur de dimension $(p + 1 * 1)$ et $\Phi(S_t)$ est une matrice de dimension $(p + 1 * p + 1)$ et ε_t est un vecteur de dimension $(p + 1 * 1)$. Afin d'étudier les propriétés probabilistes du modèle (2.5), nous faisons les hypothèses suivantes :

- H1) $E(\log^+ |\epsilon_t|) < +\infty$ où $\log^+(y) = \max \{ \log(y), 0 \}, y > 0$.
- H2) nous supposons que ϵ_t et $(Y_t, S_t)_{t \in T}$ sont indépendants.
- H3) La chaîne $(S_t)_{t \in T}$ est supposée stationnaire, irréductible et apériodique de matrice de transition : $P = (p_{ij})_{i,j \in E}$ où $p_{ij} = P(S_t = j / S_{t-1} = i)$ pour $i, j \in E$.

De distribution stationnaire initiale :

$$\pi = (\pi(1); \pi(2), \dots, \pi(M)) \text{ avec } \pi(i) = P(S_t = i), i = \{1, \dots, M\}.$$

Exemple 2.5.1 -MS(2)-AR(1) : Soit $(Y_t)_{t \in T}$ une suite de variable aléatoire suit un modèle MS(2)-AR(1) qu' est un modèle du premier ordre ($p = 1$) et avec deux régimes ($M = 2$) si elle admet la représentation :

$$Y_t = \phi_1(S_t)Y_{t-1} + \epsilon_t$$

où bien :

$$Y_t = \begin{cases} \phi_1(1)Y_{t-1} + \epsilon_{(t)}. \\ \phi_2(2)Y_{t-1} + \epsilon_{(t)}. \end{cases}$$

où $\epsilon_t \sim N(0, 1)$, $(\phi_1(1); \phi_1(2))$ les paramètres du processus pouvant générer de tels observations.

La matrice de transition $P = (p_{ij})_{1 \leq i, j \leq 2}$ qui détermine l'évolution de S_t est :

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \text{ telle que : } \begin{cases} p_{11} + p_{12} = 1 \\ p_{21} + p_{22} = 1 \end{cases} \quad (2.8)$$

2.5.1 Propriétés des modèles MS-AR :

Les propriétés théoriques de ces modèles (MS-AR) est plus complexes, pour cela on met quelques théorèmes qui sont démontrées dans Brandt [6] et Pierre Ailliot [1].

Stationnarité stricte et l'inversibilité et la causalité des modèles MS-AR

Des conditions garantissant l'existence et l'unicité d'une solution stationnaire stricte pour les modèles

MS-AR peuvent être trouvées dans Brandt [6] et Bougerol et al[5] lorsque le processus $\{\Phi_t, \varepsilon_t\}$ est strictement stationnaire et ergodique, il est démontré que si les deux conditions suivantes sont vérifiées :

- (a) $E[\log^+(\|\varepsilon_t\|)] < \infty$ et $E[\log^+(\|\Phi_t\|)] < \infty$ avec $\log^+(y) = \max(\log(y), 0)$.
- (b) L'exposant de Lyapunov noté $\gamma(\Phi)$ associé à $\Phi(t)$ est défini par :

$$\gamma(\Phi) = \inf \left\{ \frac{1}{t} E[\log(\|\Phi_t, \Phi_{t-1}, \dots, \Phi_1\|)] \quad / t \geq 1 \right\}$$

où $\|\cdot\|$ désigne une norme sur l'ensemble des matrices $(p+1 * p+1)$ et $(p * 1)$, la norme d'une matrice $\Phi(S_t) = (a_{ij})$ est définie par : $\|\Phi(S_t)\| = \max_i \left\{ \sum_{j=1}^q |a_{ij}| \right\}$.

L'existence de $\gamma(\Phi)$ est garantie par la finitude des quantités :

$$E[\log^+(\|\varepsilon_t\|)] < \infty \text{ et } E[\log^+(\|\Phi_t\|)] < \infty \text{ avec } \log^+(y) = \max(\log(y), 0), y \geq 0.$$

Théorème : 2.5.1

Soit $(Y_t)_{t \in T}$ le processus défini par (2.5) et $(\bar{Y}_t)_{t \in T}$ sa représentation vectorielle (2.6), alors

1. $\gamma(\Phi) < 0$ est une condition suffisante pour que (2.6) admet une unique solution causale, strictement stationnaire, inversible, ergodique donnée par :

$$\forall t \in T : \bar{Y}_t = \sum_{k=1}^{\infty} \bar{Y}_t(k) \varepsilon_{t-k} + \varepsilon_t \tag{2.9}$$

où la série (2.9) converge absolument p.s.

2. Inversement, si la représentation (2.6) admet une solution strictement stationnaire, alors $\gamma(\Phi) < 0$.

Chapitre 3

L'estimation dans les modèles MS-AR

fin d'estimer les paramètres du modèle MS-AR présente dans le chapitre précédent, on peut **A** procéder selon deux approches :

- L'approche une de vraisemblance où on estime les paramètres du modèle par la méthode du Maximum de Vraisemblances (MV) ou par l'une de ces dérivée. Dans notre cas, on a choisi l'algorithme EM comme une technique d'optimisation itérative qui donne des estimations du vecteur ϑ qui converge localement vers l'estimateur de MV [30].
- L'approche deux : L'estimation Bayésienne des paramètres du vecteur ϑ . Etant face à un problème d'estimation sous la présence des données inobservables représentées par la séquence cachée $\{S_t\}_1^T$, voir Gelfand et Smith[23], Tanner et Wong[58] en utilisant MCMC qui permet d'estimer les paramètres sous la présence des données manquantes ou inobservables.

Dans ce chapitre on décrit une démarche méthodique sur la calibration des modèles MS-AR par l'algorithme EM.

3.1 Probabilités filtrées et lissées :

3.1.1 Probabilités filtrées :

Notation :

$$\Psi_t = (y_1, \dots, y_T).$$

$$\underline{S}_t = (S_1, \dots, S_T).$$

Le filtre d'Hamilton est le calcul de la probabilité $P(S_t | \Psi_t, \theta)$, dite probabilité filtrée, qui nous procure une inférence du régime au temps t étant donné les rendements observés jusqu'à ce moment.

Cette inférence est donc obtenue lors du calcul de la log-vraisemblance du modèle, sans effort addi-

tionnel. La première étape à calculer :

$$f(y_t, S_t/\Psi_{t-1}) = f(y_t/S_t, \Psi_{t-1})P(S_t/\Psi_{t-1})$$

et

$$\begin{aligned} f(y_t/\Psi_{t-1}) &= \sum_{i=1}^M f(y_t, S_t = i/\Psi_{t-1}) \\ f(y_t/\Psi_{t-1}) &= \sum_{i=1}^M f(y_t/S_t = i, \Psi_{t-1})P(S_t/\Psi_{t-1}) \end{aligned}$$

telle que :

$$\begin{aligned} P(S_t = j/\Psi_t) &= P(S_t = j/y_t, \Psi_{t-1}) \\ &= \frac{f(y_t, S_t = j/\Psi_{t-1})}{f(y_t/\Psi_{t-1})} \\ &= \frac{f(y_t, S_t = j/\Psi_{t-1})}{\sum_{i=1}^M f(y_t, S_t = i/\Psi_{t-1})} \\ &= \frac{f(y_t/S_t = j, \Psi_{t-1})P(S_t = j/\Psi_{t-1})}{\sum_{i=1}^M f(y_t/S_t = i, \Psi_{t-1})P(S_t = i/\Psi_{t-1})} \end{aligned} \quad (3.1)$$

3.1.2 Probabilités lissée :

Une seconde inférence d'intérêt est la probabilité $P(S_t/\Psi_t, \theta)$, dite probabilité lissée, qu'est basée sur l'échantillon complet des données. Pour calculer les probabilités lissées, il faut faire suivre le filtre d'Hamilton. Il consiste à calculer les probabilités $P(S_t/\Psi_t, \theta)$, récursivement de $t = T-1$ à 1 en utilisant les deux relations suivantes :

$$\begin{aligned} P(S_t = i, S_{t+1} = j/\Psi_t, \theta) &= P(S_{t+1} = j/\Psi_t, \theta)P(S_t = i, S_{t+1} = j/\Psi_t, \theta) \\ &= P(S_{t+1} = j/\Psi_t, \theta) \frac{P(S_t = i, S_{t+1} = j, \Psi_t, \theta)}{P(S_{t+1} = j/\Psi_t, \theta)} \\ &= \frac{P(S_{t+1} = j/\Psi_t, \theta)P(S_t = i/\Psi_t, \theta)p_{ij}}{P(S_{t+1} = j/\Psi_t, \theta)} \quad i, j = 1 \dots M \\ P(S_t = i/\Psi_t, \theta) &= \sum_{i=1}^M P(S_t = i, S_{t+1} = j/\Psi_t, \theta) \end{aligned}$$

Remarque : 3.1.1

Les probabilités lissées sont plus précises que les probabilités filtrées puisque les premières sont basées sur plus d'information que les deuxièmes.

3.2 L'estimation dans les modèles MS-AR :

De nombreux algorithmes, permettant le calcul numérique des estimateurs du maximum de vraisemblance (noté EMV) dans les modèles à variables cachées, ont été proposés dans la littérature. Nous avons choisi principalement à savoir l'algorithme EM. L'objectif de ce paragraphe est de décrire précisément comment nous avons utilisé cet algorithme en pratique pour estimer les paramètres du modèle $MS(2) - AR(1)$ introduits au chapitre 2.

3.2.1 Fonction de vraisemblance :

Soit (y_1, \dots, y_T) une observation du processus $Y_{1:T}$, tel que $\theta \in \Theta$, l'ensemble des paramètres inconnus (i.e la probabilité initiale π , la matrice de transition P et les paramètres $\phi_i, i = 1, \dots, p$ du modèle). On cherche à estimer θ à partir des observations disponibles $\{y_1, \dots, y_T\}$ en cherchant la valeur des paramètres qui rend la fonction de vraisemblance $L(\theta) = P(\Psi_t; \theta)$ maximale. On a

$$P(\Psi_t; \theta) = \sum_{S_1, \dots, S_T \in \{1, \dots, M\}} P(\Psi_t, \underline{S}_t; \theta)$$

avec $P(\Psi_t, \underline{S}_t; \theta)$ la vraisemblance des données complètes (noté $L_c(\theta)$) qui vérifie, en utilisant la formule de Bayes et les propriétés d'indépendance conditionnelle du modèle.

$$\begin{aligned} L_c(\theta) &= P(\Psi_t, \underline{S}_t; \theta) \\ &= f(y_1, S_1/\theta) \prod_{t=2}^T f(y_t, S_t/\Psi_{t-1}, \underline{S}_{t-1}; \theta) \\ &= f(y_1, S_1/\theta) P(S_1) \prod_{t=2}^T f(y_t/S_t, \Psi_{t-1}, \underline{S}_{t-1}; \theta) P(S_t/\Psi_{t-1}, \underline{S}_{t-1}; \theta) \\ &= f(y_1/S_1; \theta) P(S_1) \prod_{t=2}^T f(y_t/S_t; \theta) P(S_t/S_{t-1}; \theta) \end{aligned}$$

Cependant (3.1) est peu utile en pratique pour calculer numériquement la fonction de vraisemblance (et donc l'estimateur du maximum de vraisemblance) car elle fait intervenir la somme de M^T termes, chacun des termes étant un produit de T termes.

La méthode usuelle pour maximiser la vraisemblance consiste alors à utiliser un algorithme EM.

3.2.2 Algorithme espérance-maximisation (EM) :

L'algorithme EM est une méthode qui permet d'approcher l'estimateur du maximum de vraisemblance (EMV) quand les données sont incomplètes (cas d'une variable cachée par exemple) ou quand

une partie des données est manquante (cas d'une censure par exemple). Il a été introduit initialement dans le cadre des modèles *CMC* par Baum et al (1970), puis étendu par la suite aux modèles à variables cachés plus généraux par Dempeter et al[15] et aux modèles *MS-AR* par Hamilton[29]. Il s'agit d'un algorithme itératif pour trouver des estimateurs du maximum de vraisemblance locaux de la fonction de vraisemblance des observations, lorsque chaque observation contient une partie cachée. Ainsi, on suppose que chaque donnée est un couple de type (S, Y) où Y est sa partie observable et S sa partie cachée, nous supposons connue d'une manière explicite la forme de la fonction densité jointe $P(S, Y; \theta)$ où θ est l'ensemble des paramètres du modèle à estimer.

-Principe de l'algorithme EM :

Nous disposons d'observations i.i.d $Y = (y_1, \dots, y_T)$ de fonction de vraisemblance notée $P(Y/\theta)$ où $\log P(Y/\theta)$ qu'est impossible à maximiser. On considère des données cachées $S = (S_1, \dots, S_M)$ dont la connaissance rendrait possible la maximisation de la vraisemblance des données complètes, $\log P(Y, S/\theta)$. L'estimateur est naturellement $E_{S/Y, \theta^{(m)}}[\log P(Y, S/\theta)]$. Et on maximise enfin cette vraisemblance estimée, nous détermine la nouvelle valeur du paramètre. Il s'agit d'un algorithme itératif partant d'une valeur initiale $\theta^{(0)} \in \Theta$ des paramètres. A chaque itération de l'algorithme, il y a deux étapes, à savoir l'étape *E* (Expectation) et l'étape *M* (Maximisation). Nous décrivons ci-dessous plus précisément ces deux étapes dans le cas des modèles *MS-AR*. $\theta^{(m)}$ désigne la valeur des paramètres après m itérations.

Étape E (Expectation) : On évalue l'espérance selon les données observées et la valeur des paramètres déterminées à l'itération précédente ce qui se calcule comme suit :

$$\Psi(\theta/\theta^{(0)}) = E_{S/Y, \theta^{(0)}}[\log P(Y, S/\theta)]$$

Étape M(Maximisation) : Cette étape procède à la maximisation de la vraisemblance en utilisant l'estimation des données inconnues effectuées à l'étape précédente, et met à jour la valeur des paramètres pour la prochaine itération. Ainsi, le passage de l'itération (m) à l'itération $(m + 1)$ de l'algorithme consiste à déterminer :

$$\hat{\theta}^{(m+1)} = \arg \max_{\theta} \{E_{S/Y, \hat{\theta}^{(m)}}[\log P(Y, S/\theta)]\}$$

Ces deux étapes sont répétées jusqu'à ce que la différence entre la fonction de vraisemblance de l'itération $(m + 1)$ et celle de l'itération (m) ne change pas. On résume ces étapes comme suit :

1. Choisir $\theta^{(0)}$ d'une manière arbitraire.

2. $m = 0$: calculer
 - Étape E : $\Psi(\theta/\theta^{(0)})$
 - Étape M : $\hat{\theta}^{(1)} = \arg \max_{\theta} \left\{ E_{S/Y, \hat{\theta}^{(0)}} [\log P(Y, S/\theta)] \right\}$
3. Si la valeur de $\hat{\theta}^{(m+1)}$ n'est pas convergente vers la valeur de $\hat{\theta}^{(m)}$ faire :

$$\Psi(\theta/\theta^{(m)}) = E_{S/Y, \theta^{(m)}} [\log P(X, S/\theta)]$$

- Étape M : l'étape de maximisation par rapport à $\theta \in \Theta$

$$\hat{\theta}^{(m)} = \arg \max_{\theta} \left\{ \Psi(\theta/\theta^{(m)}) \right\}$$

4. Si $\theta^{(m)} = \theta^{(m+1)}$, on arrête.
5. Fin

Convergence de l'algorithme EM :

Wu[64] a discuté en détail la convergence de l'algorithme (EM). Il donne des conditions de régularité qui rendent tous les points limites de l'algorithme (EM) des points stationnaires de la vraisemblance.

Ces conditions de régularité sont :

- (i) $\Theta_{\hat{\theta}} = \{\hat{\theta} \in \Theta, L_n(\hat{\theta}) \geq L_n(\hat{\theta}')\}$ est compact pour tout $L_n(\hat{\theta}') > -\infty$.
- (ii) $L_n(\hat{\theta})$ est continue dans Θ^d et différentiable à l'intérieur de Θ^d .
- (iii) $Q(\hat{\theta}, \hat{\theta}', X)$ est continu par rapport à $\hat{\theta}$ et $\hat{\theta}'$.

3.3 Estimation des modèles MS(2)-AR(1) via l'algorithme EM :

Supposons qu'on a $\{y_1, \dots, y_T\}$ une observation du processus aléatoire $(Y_t)_{t \in T}$ suit un modèle MS(2)-AR(1). Soit $S = (S_1, \dots, S_T)$ la réalisation du régime non observé de tout l'échantillon.

Les paramètres inconnus sont : $\phi_1(S(t)) = (\phi_1(1), \phi_1(2))$ et $p_{11}, p_{12}, p_{21}, p_{22}$ et π_1, π_2 puisque $p_{11} + p_{12} = 1, p_{21} + p_{22} = 1$ et $\pi_1 + \pi_2 = 1$. Donc θ le paramètre inconnu est appartient à un espace de $\Theta \in \mathbb{R}^5$. La vraie valeur de θ sera notée θ_0 . Finalement, on note la densité des innovations $(\epsilon_t)_t$ par $f(\cdot)$. Alors, la distribution conditionnelle jointe de (Y, S) si S était observée sachant Y_1 et S_1 sont

données par :

$$\begin{aligned}
 f(\Psi_t, \underline{S}_t/\theta) &= f(y_1, S_1/\theta) \prod_{t=2}^T f(y_t, S_t/\Psi_{t-1}, \underline{S}_{t-1}; \theta) \\
 &= f(y_1, S_1/\theta) P(S_1) \prod_{t=2}^T f(y_t/S_t, \Psi_{t-1}, \underline{S}_{t-1}; \theta) P(S_t/\Psi_{t-1}, \underline{S}_{t-1}; \theta) \\
 &= f(y_1/S_1; \theta) P(S_1) \prod_{t=2}^T f(y_t/S_t; \theta) P(S_t/S_{t-1}; \theta)
 \end{aligned}$$

où $(\epsilon_{\phi_1(S_t)}(t))_{t \in T}$ est le processus déterminé par $\epsilon_{\phi_1(S_t)}(t) = Y_t - \phi_1(S_t)Y_{t-1}$. En choisissant la loi normale comme loi de distribution de (ϵ_t) , nous obtenons :

$$f(\epsilon_{\phi_1(S_t)}(t)) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(Y_t - \phi_1(S_t)Y_{t-1})^2}{2}\right)$$

Il s'en suit que, la vraisemblance marginale de $Y, L_T(Y, \theta)$ sera simplement la somme de $P(Y, S; \theta)$ sur toutes les valeurs possibles de (S_1, \dots, S_T) :

$$\begin{aligned}
 f(\Psi_t, \underline{S}_t/\theta) &= [P(S_1 = 2)f(y_1/S_1 = 2; \theta)\pi_1 + P(S_1 = 1)f(y_1/S_1 = 1; \theta)(1 - \pi_1)] \\
 &\quad \prod_{t=2}^T P(S_t = 2, S_{t-1} = 2)f(y_t/S_t = 2; \theta)p_{22} \\
 &\quad + P(S_t = 1, S_{t-1} = 2)f(y_t/S_t = 1; \theta)(1 - p_{22}) \\
 &\quad + P(S_t = 2, S_{t-1} = 1)f(y_t/S_t = 2; \theta)(1 - p_{11}) \\
 &\quad + P(S_t = 1, S_{t-1} = 1)f(y_t/S_t = 1; \theta)p_{11}
 \end{aligned}$$

telle que :

$$\begin{aligned}
 \log P(\Psi_t, \underline{S}_t/\theta) &= P(S_1 = 2)[\log f(y_1/S_1 = 2; \theta) + \log \pi_1] \\
 &\quad + P(S_1 = 1)[\log f(y_1/S_1 = 1; \theta) + \log(1 - \pi_1)] \\
 &\quad + \sum_{t=2}^T P(S_t = 2) \log f(y_t/S_t = 2; \theta) + P(S_t = 1) \log f(y_t/S_t = 1; \theta) \\
 &\quad + P(S_t = 2, S_{t-1} = 2) \log(p_{22}) + P(S_t = 1, S_{t-1} = 2) \log(1 - p_{22}) \\
 &\quad + P(S_t = 2, S_{t-1} = 1) \log(1 - p_{11}) + P(S_t = 1, S_{t-1} = 1) \log(p_{11})
 \end{aligned}$$

alors

$$\log P(\Psi_t/\theta) = \log\left(\sum_{S_1=1}^2 \sum_{S_2=1}^2 \dots \sum_{S_T=1}^2 P(\Psi_t, \underline{S}_t/\theta)\right)$$

Comme les données non observées S sont présentes dans la vraisemblance, l'estimation paramétrique des modèles MS(2)-AR(1) ne peut pas être réalisée par la méthode MV directement mais plutôt à

l'aide de l'algorithme EM. On détermine l'expression exacte de $\Psi(\theta^{(m+1)}/\theta^{(m)})$, elle est donnée par :

$$\begin{aligned}
 \Psi(\theta^{(m+1)}/\theta^{(m)}) &= E[\log f(\Psi_t, \underline{S}_t/\theta^{(m)})] \\
 &= \pi_1^{(m)}[\log f(y_1/S_1 = 2; \theta) + \log \pi_1^{(m)}] \\
 &+ (1 - \pi_1^{(m)})[\log f(y_1/S_1 = 1; \theta^{(m)}) + \log((1 - \pi_1^{(m)}))] \\
 &+ \sum_{t=2}^T \{P(S_t = 2/\Psi_t; \theta^{(m)})\log f(y_t/S_t = 2; \theta^{(m)})\} \\
 &+ P(S_t = 1/\Psi_t; \theta^{(m)})\log f(y_t/S_t = 1; \theta^{(m)}) \\
 &+ P(S_t = 2, S_{t-1} = 2/\Psi_t; \theta^{(m)})\log(p_{22}) \\
 &+ P(S_t = 1, S_{t-1} = 2/\Psi_t; \theta^{(m)})\log(1 - p_{22}) \\
 &+ P(S_t = 2, S_{t-1} = 1/\Psi_t; \theta^{(m)})\log(1 - p_{11}) \\
 &+ P(S_t = 1, S_{t-1} = 1/\Psi_t; \theta^{(m)})\log(p_{11})
 \end{aligned}$$

Dans la seconde étape (M-step) de nouveaux estimateurs MV du paramètre, basés sur les inférences retenues dans l'étape (E-step), sont calculés. Les 2 étapes sont répétées jusqu'à ce que le maximum de vraisemblance soit atteint Une description détaillée de l'algorithme dans le cas du modèle MS(2)-AR(1) . **Etape M (M-step)** : Dans la seconde étape de l'algorithme EM, un nouvel et meilleur estimateur du maximum de vraisemblance $\theta^{(m+1)}$ est calculé. Hamilton (1990) a donné les résultats de la maximisation de l'expression (3.1) en choisissant $\theta^{(m+1)}$ satisfaisant les égalités suivantes :

$$\frac{\partial \mathbf{E}[\log f(\Psi_t, \underline{S}_T/\theta^{(m)})]}{\partial \phi_i^{(m+1)}}$$

et

$$\phi_1^{(m+1)}(k) = \frac{\sum_{t=2}^T y_t y_{t-1} P(S_t = k/\Psi_{t-1}, \phi_1^{(m)})}{\sum_{t=2}^T P(S_t = k/\Psi_t, \phi_1^{(m)})}, k = 1, 2$$

3.4 Exemple d'application : modèle MS-AR en économie :

Les modèles à changement de régime Markovien ont connu un fort développement depuis 1989. Hamilton propose un modèle non linéaire mais stationnaire du produit national brut PNB américain. Il révèle également l'impact de ce nouveau modèle sur la croissance à long terme et l'attention portée à l'histoire du cycle économique.

Nous présentons un exemple de la série française PNB (présentée dans la publication Ingénierie économique et financière du Dauphin Mestre IEF 272), car elle constitue une chaîne centrale dans les études de macroéconomie et un bon exemple de l'application du modèle MS-AR à l'économie .

3.4.1 Application à la série du produit national brut (PNB) Français :

En économie, le PNB ou Produit national brut mesure la production sur une période donnée, tel que le PNB est égal au PIB (Produit intérieur brut).

La série choisie d'utiliser pour cette étude est le taux de croissance du PIB français depuis le premier trimestre de l'année 1980 jusqu'à juillet 2013. Pour l'étude du taux de croissance du PIB français, on a 135 valeurs trimestrielles (qui proviennent du site de l'OCDE). Nous allons au cours de cette étude estimer cette variable à l'aide d'un modèle auto-régressif à changement de régimes Markoviens MS-AR.

Ci-dessous la représentation graphique de la série.

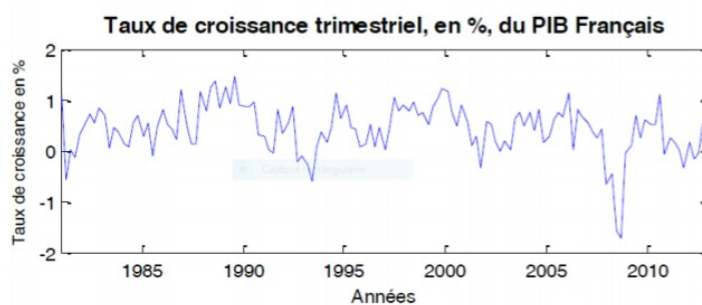


FIGURE 3.1:

A l'aide de la figure(3.1), nous pouvons remarquer que l'allure générale du taux de croissance du PIB français semble stationnaire et que durant la période que nous étudions, la France a connu trois périodes majeures de récession. La première période de récession correspond à la récession des années 1992-1993 qui est due à des politiques monétaires excessivement restrictives. La deuxième période

correspond à la période de récession suite à la crise économique due à la crise des supprimes qui a affecté l'économie mondiale en 2008. La dernière récession qui a affectée la France fut la récession due à la crise de la dette en zone Euro.

Remarque : 3.4.1

Notre série est bien stationnaire, nous pouvons continuer l'étude du taux de croissance du PIB français.

3.4.2 Formulation par le modèle MS(2)-AR(1) :

Considérons $y_t = \log(PIB)_t$; à long terme, le taux de croissance est constant c'est-à-dire que : $E(y_t) = E(\log PIB_t) = \mu$, n'est pas affecté par un changement d'état. A l'inverse, un changement d'état modifie le niveau du PIB. Donc on suppose que y_t suit un modèle auto-régressif à changement de régime markoviens d'ordre 1 avec 2 régimes MS(2)-AR(1) présentés par l'équation suivante :

$$y_t = \phi(S_t)y_{t-1} + \epsilon_t \text{ ou } \epsilon_t \sim N(0, 1).$$

3.4.3 Estimation du modèle MS(2)-AR(1) :

Dans le but d'évaluer les performances de la procédure d'estimation du modèle $MS(2) - AR(1)$, basée sur l'algorithme EM, étudiée dans le chapitre précédent où :

$$y_t = \phi(S_t)y_{t-1} + \epsilon_t$$

où $(\epsilon_t)_{t \in T}$ un bruit blanc gaussien $N(0, 1)$ et $(S_t)_{t \in T}$ est une chaîne de Markov du premier ordre stationnaire, irréductible et homogène. Les simulations ont été réalisées pour différents ensembles des valeurs initiales $\theta^{(0)}$.

Les moyennes empiriques et les écart-types des estimateurs (EM) des paramètres sont donnés dans le Tableau- 1- pour différentes tailles $n = 500, 1000$ et 2000 et pour les paramètres $\phi(1) = 0.15, \phi(2) = 0.25, \sigma^2 = 1, p_{11} = 0.75, p_{22} = 0.8$ et $\pi_1 = 0.44$. Comme attendu, les résultats indiquent que quelque soit le choix des valeurs initiales la méthode d'estimation fonctionne parfaitement (en effet, les résultats pour différents valeurs initiales sont très proches des résultats du Tableau- 1-. Les moyennes empiriques sont très proches des vraies valeurs pour toutes les tailles d'échantillons et les écart-types sont très petits. De plus, l'augmentation de la taille de l'échantillon améliore les résultats en terme de réduction de biais.)

Tableau- 1- : Moyennes et écart-types des estimations des paramètres avec :

$$\phi(1) = 0.15, \phi(2) = 0.25, \sigma^2 = 1,$$

$$p_{11} = 0.75, p_{22} = 0.8, \pi_1 = 0.44.$$

Expérience 1 : n=500 Expérience 2 : n=1000 Expérience 3 : n=2000

	Mean	RMSE		Mean	RMSE		Mean	RMSE
$\widehat{\phi}(1)$	0.1527	0.1563	$\widehat{\phi}(1)$	0.1534	0.1094	$\widehat{\phi}(1)$	0.1499	0.0490
$\widehat{\phi}(2)$	0.2510	0.0042	$\widehat{\phi}(2)$	0.2493	0.0025	$\widehat{\phi}(2)$	0.2500	0.0015
p_{11}	0.7514	0.0350	p_{11}	0.7482	0.0263	p_{11}	0.7502	0.0179
p_{22}	0.8016	0.0362	p_{22}	0.8014	0.0256	p_{22}	0.7987	0.0180
π_1	0.4382	0.0291	π_1	0.4415	0.0282	π_1	0.4407	0.0161

Conclusion générale

La contribution principale de ce travail a été l'extension de l'algorithme d'expectation–maximisation (EM) initié par Hamilton [29] pour couvrir le modèle auto-régressif à changement de régime markovien (*MS-AR*). La classe des modèles auto-régressifs a été bien étudiée dans la littérature, mais la difficulté dans notre cas est que les coefficients dépendent d'une chaîne de Markov non observable et homogène avec un espace d'états fini. Basé sur des travaux antérieurs à savoir, ceux de [17], [29], et [64] nous avons montré que l'algorithme EM converge.

Bibliographie

- [1] **Ailliot.P**,2004,Modèle auto-régressifs à changements de régimes Markovien,Application aux séries temporelles de vent, Thèse de Doctorat, Université de Rennes 1. France.
- [2] **Aragon. Y**, (2011), Séries temporelles avec R, Springer.
- [3] **Bhar.R. et Hamori.S** (2004), Hidden Markov Models Application financial economics, Kluwer Academic Publishers.
- [4] **Bollerslev.T**(1986),Generalized Autoregressive Conditional Heteroskedasticity,Journal of Econometrics ,vol 31,307-327.
- [5] **Bougerol. P.Picard, N.**(1992). Strict stationarity of generalized autoregressive processus. Ann. Probab. 20, 1714-1730.
- [6] **Brandt, A.** (1986) L'équation stochastique $Y_n + 1 = A_n Y_n + B_n$ avec des coefficients fixes. Adv. En appl. Probab. 18, pages 211-220.
- [7] **Breman, P.** (1999), Chaînes de Markov, Gibbs Fields, Simulation de Monte-Carlo et Files d'attente, Springer-Verlag New York, Inc., p. 52-75.
- [8] **Brémaud.P**,(1998), Markov chains Gibbs fields, Monte Carlo Simulation and queues, Springer,
- [9] **Celeux.G et Soromenho.G.** (1993). An Entropy criterion for assessing the number of clusters in mixture model. INRIA.France.
- [10] **Chang JK, Charles RN** (1999) .L'économie américaine devient-elle plus stable? Approche bayésienne basée sur un modèle de Markov-commutation du cycle économique. La revue économique et statistique, 81 (4) : 608-616
- [11] **Chen.S.W et Lin.J.L** (1999). Modeling Business Cycle in Taiwan with Time-Varying Markov-Switching Models.
- [12] **Chib.S** (1996), Calculating posterior distributions and modal estimates in Markov mixture models, Journal of Econometrics, Vol 75,79-97.

-
- [13] **Ching, WK, Michael. K. Ng.** (2006) Chaînes de Markov : Modèles, algorithmes et applications, Springer Science-Business Media, Inc, pp115.
- [14] **Dafas.P.A.** (2004). Estimating the parameters of a mean-reverting Markov-switching jump-diffusion model for crude oil spot prices.
- [15] **Dempster.A.P et Laide.N.M et Rubin.D.B** (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm, Journal of the Royal Statistical Society, vol 39,N :1.1-38.
- [16] **De Micheaux.P.L et Droulihet.R et Liquet.B.** (2011). Le logiciel R métrise de langage. Springer.
- [17] **Douc, E .R. Moulines, T. Ryden.** (2004). Propriétés asymptotiques de l'estimateur du maximum de vraisemblance dans les modèles autorégressifs avec régime de Markov, The Annals of Statistics 32 (5) 2254–2304.
- [18] **Engel.R.F.** (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance. Econometrica, Vol 50, No 4, 987-1007.
- [19] **Ephraim et Merhav.**(2002), Les processus de Markov cachés, IEEE TRANSACTIONS ON THEORY THEORY, 48 (6) p. 1518-1569.
- [20] **Franqc. C, M. Roussignol,** (1998) Ergodicité des processus autorégressifs avec commutation de Markov et cohérence de l'estimateur du maximum de vraisemblance, Statistiques 27 (2) 1–38.
- [21] **Franses.P.H et Dijk.D.V.** (2003). Non-Linear time series models in empirical finance. CAMBRIDGE UNIVERSITY PRESS.
- [22] **Frédéric .S** .(2015). L'algorithme EM : une courte représentation .CNRS, UMR 5199 PACEA, avenue de faculté, Bat. B8.
- [23] **Gelfand.A.E et Smith.A.F.M.**(1989). Sampling based approaches to calculating marginal densities. Journal of the American Statistical Association, vol 85, No 410, 398-409.
- [24] **George Yin. G., Zhang. Q.**(2005) Chaînes de Markov en temps discret : méthodes et applications à deux échelles de temps, Springer Science-Business Media, Inc, p. 7-71.
- [25] **Goldfeld.S.M et Quandt.R.E.** (1973). The estimation of structural shifts by switching regressions. Econometric Program Research. PRINCETON UNIVERSITY. New Jersey.
- [26] **Goodwin.T.H.** (1993). Business-Cycle Analysis with a Markov-Switching Model. Journal of Business Economic Statistics Vol 11, N : 3 , 331-339.
- [27] **Gray.F.S,** (1996), Modeling the Conditional Distribution of Interest Rates as a RegimeSwitching Process, Journal of Financial Economics 42, 27-62.
-

-
- [28] **Haggstrom, O.** (2002) Chaînes de Markov finies et applications algorithmiques, Cambridge University Press, p. 7-39
- [29] **Hamilton.J.D**, (1989),A new approach to the economic analysis of nonstationary time series and the business cycle *Econometrica*, vol 57, N : 2, 357-384.
- [30] **Hamilton.J.D**, (1990), Analysis of time series subject to change in regime, *Journal of Econometrics*, 45, 39-70.
- [31] **Hamilton.J.D**, (1994),Time series analysis. PRINCETON UNIVERSITY PRESS. NEW JERSEY.
- [32] **Hamilton.J.D.** (2011). Calling recession in real time. *International Journal of Forecasting*, 27, 1006–1026.
- [33] **Henneke.J.S et Rachev.S.T et Fabozzi.F.J**, (2006),MCMC Based Estimation of Markov Switching ARMA-GARCH Models.
- [34] **HM Krolzig.** (1997) Autorégressions sur vecteur de commutation de Markov. Modélisation, inférences statistiques et applications à l'analyse du cycle économique. Notes de cours en économie et systèmes mathématiques, Springer-Verlag, Berlin.
- [35] **James D. Hamilton.** (2005) .Régime des modes de fonctionnement. Département des affaires économiques, 0508, Université de Californie à San Diego, La Jolla,vers92093-0508.
- [36] **Jean F. D, Benjamin J.** (2005). Springer Berlin Heidelberg à New York.
- [37] **Jeff.W** (1983), On the convergence proprieties of the EM algorithm., *The Annals of Statistics*, Vol 11, N : 1,95-103.
- [38] **Kim, C. J** (1994) ,Dynamic linear models with Markov-switching, *Journal of econometrics*, vol 60, 1- 22.
- [39] **Kim.C.J et Nelson.C.R.** (a1999). State-Space Models with Regime Switching. The MIT Press. Cambridge. UK.
- [40] **Kulkarni.VG.** 2011 .Introduction à la modélisation et à l'analyse de systèmes stochastiques. Deuxième édition. Springer Science + Business Media, LLC
- [41] **Malik, S. et Pitt, MK.**(2011). Filtres à particules pour la vraisemblance continue évaluation et maximisation. *Journal of Econometrics* , 165 (2) : 190-209.
- [42] **McLachlan.G.** (2000). Finite Mixture Models. Wiley and Son.USA.
- [43] **McCulloch.R.E et Tsay.R.S** (1994), Statistical analysis of economic time series via Markov Switching Models, University of Chicago.
-

-
- [44] **Medhioub. I**, (2005). Asymétrie des cycles économiques et changement de régime : cas de la Tunisie, Ecole Supérieure de Commerce de Sfax, Tunisie.
- [45] **Michael PC, Hans MK**.(1997). Comparaison de la performance prévue des modèles à commutation de Markov et à modèles autorégressifs à seuil du PNB américain. England, UK, numéro 499.
- [46] **Moussi.O**. (2008). Modélisation de l'évolution des prix de pétrole : Estimation Bayésienne. Thèse de Doctorat. Institut National de la Planification et de la Statistique. Algérie.
- [47] **Olteaunu.M**, (2006), Modèles à changement de régime : Application aux données financières, Thèse de Doctorat, Université Paris-1-Panthéon-Sorbonne, France.
- [48] **Pierre A**. (2004) Modèles autorégressifs à changements de régimes markoviens. Application aux séries temporelles de vent. Thèse de doctorat, université de Rennes.
- [49] **Pierre A**. (2006) Quelques résultats théoriques sur des modèles autorégressifs à commutation de markov avec des innovations gamma. comptes Rendus de l'Académie des Sciences de Paris, 343 (4) : 271-274 .
- [50] **Piger.J** (2007), Econometrics : Models of Regime Changes, University of Oregon.
- [51] **Rabiner.L.R**. (1988). A tutorial on hidden Markov models and selected applications in speech recognition. IEEE, vol 77, NO 2, 257-286.
- [52] **Rocio A, Maximo C, Manuel R**. (2017) Inférence sur les probabilités filtrées et lissées dans les modèles autorégressifs à commutation de Markov. ECO2015-65637-P.
- [53] **Sabin, L**.(2014) .Processus stochastiques · cours et exercices corrigés. Ellipses Édition
- [54] **Schafer.J.L**, (2000), Analysis of incomplete multivariate data, Chapman and Hall, US
- [55] **Schnatter.S.F**. (2006). Finite Mixture and Markov Switching Models. Springer. USA.
- [56] **Smith. A et Naik.P.A et Tsai.C.L**. (2005). Markov-Switching Model Selection Using Kullback-Leibler Divergence. Department of Agricultural and Resource Economics University of California.
- [57] **Stroock, DW**. (2005) Introduction aux processus de Markov, Springer-Verlag Berlin Heidelberg, pp 23-52
- [58] **Tanner.M.A et Wong.H.W**.(1987). The calculation of posterior distribution by Data augmentation. Journal of the American Statistical Association, vol 82, No 398, 528-540.
- [59] **Timmermann, A**. (2000). Moments de modèles de Markov de commutation. Journal of Econometrics, 96 (1), 75-111.
-

- [60] **Tong, H.** (1978). On a threshold model Pattern Recognition and Signal Processing, Amsterdam : Sijthoff and Noordhoff.
- [61] **Touche.N** ; (2010), Etude de quelques modèles de séries chronologiques à changement de régimes Markoviens, Mémoire de Magister, Département de Recherche Opérationnelle, USTHB.
- [62] **Wang.P**, (2003), Financial Econometrics, Routledge.
- [63] **Weron.R.**(2006). Modeling and Forecasting Electricity Loads and Prices. Wiley Sons.
- [64] **Wu, C.-FJ** .(1983). Sur les propriétés de convergence de l'algorithme EM. The Annals of Statistics , 11 (1) : 95-103.
- [65] **Yao JF, JG Attali.** (2000). Sur la stabilité des processus AR non linéaires avec commutation de Markov, Avancées en probabilité appliquée, 32 394–407.
- [66] **Zucchini. W., L. MacDonald.** (2009) Modèle de Markov caché pour les séries chronologiques, introduction à l'aide de R. Chapman et Hall.