

N° d'ordre : 01/2010-M/MT

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediene
Faculté de Mathématiques



MEMOIRE

Présenté pour l'obtention du diplôme de Magister

EN : Mathématiques

Spécialité : Modèles Stochastiques de Recherche Opérationnelle

Par: KESSIRA Abderrahim

Thème

Structure probabiliste et inférence dans des
modèles de séries chronologiques à seuils

Soutenu publiquement le 04 / 10 / 2010 devant le jury composé de

M.	BEBBOUCHI Rachid	Professeur à l'USTHB	Président
Mme.	GUERBYENNE Hafida	M.C/A à l'USTHB	Directrice de mémoire
M.	BELBACHIR Hacène	M.C/A à l'USTHB	Examineur
M.	KHALDI Khaled	M.C/A à l'USTHB	Examineur
M.	HAMDI Fayçal	M.C/B à l'USTHB	Invité

Remerciements

Louange à Dieu, le miséricordieux, sans lui rien de tout cela n'aurait pu être.

Je tiens à exprimer ici toutes ma reconnaissance à Mme H. Guerbyenne, ma directrice de mémoire. Je la remercie pour sa contribution à l'aboutissement de ce travail, pour ses précieuses remarques et suggestions, et spécialement pour l'aide morale qu'elle m'a apporté, je la remercie infiniment.

Je tiens ensuite à remercie Mr R. Bebbouchi, pour l'honneur qu'il ma fait en acceptant de présider le jury de ce mémoire.

Messieurs H. Belbachir, F. Hamdi et K. Khaldi, m'ont fait l'honneur d'accepter de faire partie de mon jury de soutenance. Je les en remercie très vivement.

Je tiens enfin à remercie mes parents, mon frère et mes sœurs qui m'ont constamment soutenu durant ce travail.

Table Des Matières

Introduction générale

0 Aperçu sur les modèles non linéaire	1
0.1 Introduction	1
0.2 Autorégression non linear (NLAR)	1
0.3 Principe de threshold et les modèles de threshold	2
0.3.1 La classe des modèles self-exciting threshold autoregressive / moving average (SETAR / MA)	3
0.4 La classe des modèle à seuils sans bruit additif	5
0.4.1 La classe des modèles autorégressifs exponentiels (EAR)	6
0.4.2 La classe des modèles ARMA à coefficients périodiques (PARMA)	7
0.5 la classe des modèles autorégressifs multiplicatifs	7
0.6 La classe des modèles autorégressifs à coefficient aléatoire (RCA)	8
1 Stabilité des chaînes de Markov	10
1.1 Introduction	10
1.2 Noyau de transition et définition d'une chaîne de Markov	10
1.3 Irréductibilité et apériodicité	12
1.4 Chaîne récurrente, chaîne transitoire	16
1.5 Existence des mesures invariantes	18
1.6 Ergodicité	20
1.7 Critères de classification	23
2 Structure probabiliste de modèles à seuils	28
2.1 Modèle autorégressif à seuil TAR(1)	28
2.1.1 Introduction	28
2.1.2 Ergodicité du modèle TAR(1)	29
2.1.3 Modèle TAR(1) général	36
2.1.4 Ergodicité	36

2.2	Modèle moyenne mobile à seuil	42
2.2.1	Introduction	42
2.2.2	Ergodicité des modèles TMA(1)	43
2.2.3	Inversibilité des modèles TMA(1)	46
2.2.4	Extension à des modèles TMA généraux	48
2.3	Modèle ARMA conditionnellement hétéroscédastique à seuils	52
2.3.1	Introduction	52
2.3.2	Le modèle DTARMACH	52
2.3.3	Modèles TARMA $(p; q)$ et TAR (p)	62
3	Estimation dans des modèles autorégressifs à seuils	65
3.1	Modèle TAR(1) simple	65
3.2	Test d'égalité de ϕ_1 et ϕ_2	67
3.3	Quelques simulations	68
3.4	Modèle TAR(1) multiple	73
3.5	Estimation du maximum vraisemblance d'un modèle autorégressif à seuils	75
3.5.1	L'estimation par la méthode du maximum vraisemblance	76
3.5.2	Hypothèses et consistance forte	77
3.5.3	n -consistance de l'estimateur du threshold	78
4	Application et simulations	79
4.1	Introduction	79
4.2	Test pour une non linéarité de type threshold	80
4.2.1	Autorégression arrangée et résidus prédictifs	80
4.2.2	Test de non linéarité	82
4.2.3	Puissance du test	83
4.3	Spécification de la variable threshold	85
4.3.1	Sélection du paramètre de délai	85
4.3.2	Localisation des valeurs des seuils	86
4.4	Procédure de modélisation	91
4.5	Application sur une série réelle	91

Annexe	96
Bibliographie	97

INTRODUCTION GENERALE

Les modèles linéaires introduits depuis plus de sept décénies présentent l'aspect attrayant d'être computationnellement gérables. Mais, il est bien connu que les modèles autorégressifs moyenne mobile de Box et Jenkins (1970) ne réussissent pas à capter certains phénomènes communément observés en pratique. En particulier, il y a l'irréversibilité temporelle exhibée par l'asymétrie entre les périodes d'ascension et de descente dans les données annuelles sur les taches solaires. Motivés par cette déficience des modèles linéaires, des chercheurs ont proposé des modèles non linéaires pour l'analyse de données de séries chronologiques. De nombreuses classes de modèles non linéaires ont, ainsi, vu le jour. Granger et Andersen (1978) ont introduit les modèles bilinéaires. Engle (1982), Bollerslev (1986) ont respectivement développé les modèles conditionnellement hétéroscédastiques (*ARCH*) et conditionnellement hétéroscédastiques généralisés (*GARCH*). Les modèles autorégressifs à coefficients aléatoires (*RCA*) semblent être nés durant la même période (Consilk (1974), Andel (1976), Nicholls et Quinn (1982), etc.), reflétant ainsi les préoccupations des statisticiens toujours soucieux de mieux prendre en charge des données d'observations dans le temps. Hamilton (1989) a introduit les modèles autorégressifs avec changement de régime markovien (*MS - AR*). Les modèles à seuils sont aussi des modèles à changements de régime. Ils ont été introduits par Tong (1977*a*, 1977*b*, 1978). Le changement y est endogène ou lié à une variable de même type. Dans les modèles à changement de régime markovien, la transition y est exogène et dépend d'une chaîne de Markov à espace d'état fini (désignant par exemple la récession ou l'expansion pour une série de données économiques). Le modèle autorégressif à seuils (*TAR*), a la particularité de pouvoir reproduire les phénomènes de cycles limites, d'irréversibilité temporelle, de dépendance entre l'amplitude et la fréquence ainsi que les phénomènes de saut (Tong et Lim (1980)). Ces auteurs ont été rejoints par de nombreux autres chercheurs qui ont saisi l'intérêt d'une telle modélisation puisque celle-ci s'est étendue à d'autres cadres, en particulier, ceux à erreurs conditionnellement hétéroscédastiques.

Les modèles à seuils ont connu un développement important. Ces modèles jouent

un rôle clé dans la modélisation de données cycliques pour les quelles seuls les modèles non linéaires sont appropriés pour en capturer l'évolution. Les modèles non linéaires doivent être capables d'offrir une vision sur la dynamique sous-jacente des données. La théorie déterministe des systèmes dynamiques a été une source d'inspiration pour Tong (Tong (2007)), notamment, les cycles limites, les oscillations non linéaires, les phénomènes de saut, etc.

Les cycles limites représentent l'état stationnaire d'oscillations qui ne dépendent pas des conditions initiales mais dépendent exclusivement des paramètres du système. Ce sont des propriétés intrinsèques. Il existe des cycles limites robustes en ce sens qu'ils sont insensibles à de légères perturbations des paramètres du système et qui sont purement observables. La présence d'un seuil spécifie les modes opérationnels du système et ouvre la voie à la possibilité d'oscillations de type cycle limite.

Intuitivement, une série chronologique est réversible dans le temps si en marchant vers l'avant ou vers l'arrière, la distribution de probabilité ne change pas, i.e. la direction du temps n'a pas d'effet. L'irréversibilité en la négation.

Des exemples de modèles à seuils fondés sur la linéarité par morceaux émanent de nombreux domaines dans lesquels la notion de seuils est dominante : l'ingénierie radio, l'ingénierie marine, l'ingénierie médicale, les servo-systèmes, l'océanographie, la biologie des populations, l'économie, l'hydrologie, les moteurs à vapeur, etc.

L'idée de base d'un modèle à seuils est la linéarisation par morceaux en introduisant une série chronologique indicatrice $\{J_t, t \in \mathbb{Z}\}$. De nombreuses interrogations ont suivi l'application des modèles TAR notamment la détermination de la variable threshold, le paramètre de délai, le nombre de régimes, ...etc. La question la plus importante pour n'importe quel modèle de séries chronologiques est l'existence d'une distribution stationnaire. Le cas des modèles linéaires ARMA est bien connu. Pour les modèles non linéaires, la situation est beaucoup plus compliquée et les résultats qui existent pour le moment sont incomplets. En effet, il est clair qu'un modèle MA à seuils est strictement stationnaire. Pour les

modèles AR non linéaires, quelques résultats généraux existent. On les traite comme étant une chaîne de Markov à espace d'états dans l'espace Euclidien. Un outil puissant, dans ce cas, est le critère de dérive (Drift criterion) initié par Foster (1953) et développé par Tweedie (1975) et Nummelin (1978, 1984). L'idée de base est d'introduire ce que l'on entend par small sets qui jouent le rôle des états d'une chaîne de Markov discrète, ensuite, étudier l'irréductibilité et finalement, vérifier s'il y a un dérive vers le centre de l'espace d'états. De nombreux travaux de Tweedie (1975, 1976) nécessite l'hypothèse que la chaîne de Markov est irréductible. Mais Tweedie (1988) se libère de cette contrainte.

Une notion clé dans l'approche du critère de dérive est la fonction test (g -function), dont le choix n'est pas toujours évident. Cependant, en l'interprétant comme étant la fonction d'énergie 'généralisée', Chan et Tong (1985) ont établi le lien entre la stabilité d'un système déterministe et l'ergodicité d'un système stochastique. D'autres travaux sur le même sujet et pour d'autres modèles à seuils existent, Petrucci et Woolford (1984) ont donné une condition nécessaire et suffisante pour qu'un modèle TAR(1) à deux régimes soit ergodique, Chan *et al* (1985) ont obtenu la condition nécessaire et suffisante pour l'ergodicité d'un TAR(1) à plusieurs régimes, Chan et Tong (1985) ont donné une condition suffisante pour l'ergodicité d'un modèle TAR général. Guégan et Diebolt (1994) ont étudié un modèle β -ARCH pour lequel ils ont examiné les propriétés d'inversibilité d'irréductibilité de récurrence au sens de Harris, d'ergodicité, d'ergodicité géométrique, d'alpha mélange et d'existence des moments. Zakoïan (1994) a considéré un modèle ARCH modifié, à seuils. Il a déterminé les conditions de stationnarité et a considéré les méthodes des moindres carrés et du maximum de vraisemblance pour l'estimation des paramètres du modèle. Liu *et al* (1997) ont obtenu un résultat plus fort que celui de Chan et Tong (1985) et une condition de stationnarité stricte pour un modèle TAR à erreurs conditionnellement hétéroscédastiques. Li et Li (1996) ont étudié un modèle DTARCH pour lequel ils ont développé les techniques d'identification du modèle, d'estimation et de test. Une généralisation de ces résultats est donné pour le modèle DTARMACH dans Ling (1999), il s'est intéressé à la stationnarité et à la finitude des moments pour ce modèle. Ling *et al*

(2007) ont donné des résultats pour l'inversibilité et l'ergodicité du modèle TMA simple et multiple. Plus récemment, Chan et Tong (2010) ont considéré le cas, plus difficile, de l'inversibilité des modèle ARMA non linéaires. Ils ont obtenu une condition suffisante simple pour l'inversibilité locale de ce modèle. Samia et Chan (2010) ont généralisé le modèle à seuils de Tong au modèle plus général GTM (Generalised threshold Model) où ils supposent que la variable réponse appartient à la famille exponentielle et la moyenne conditionnelle est une fonction de régression stochastique linéaire par morceaux. La consistance et la loi de probabilité asymptotique de l'estimateur du maximum de vraisemblance sont déterminées. Pole et Smith (1985) proposent une analyse bayésienne pour des modèles à seuils.

Tjøstheim (1990) a montré que la théorie sur les chaînes de Markov peut être exploitée pour étudier des séries chronologiques non linéaires afin d'obtenir une classification des modèles en stationnaires et non stationnaires. Tsay (1991) propose une méthodologie pour l'analyse de séries non linéaires. Le test de Tsay peut détecter divers types de non linéarité comme la non linéarité de type threshold, la bilinéarité et la non linéarité exponentielle. Cline (2006) propose une méthode pour déterminer si un modèle TAR-ARCH est ergodique et lorsque c'est le cas, de voir quels sont les moments qui existent. Hili (2008) s'intéresse à un modèle TAR à seuils multiples sous des conditions sur les coefficients du modèle qui garantissent la stationnarité, l'existence des moments et la propriété d'alpha mélange du processus. Il établit la consistance et la normalité asymptotique des estimateurs des coefficients du modèle fondés sur la distance de Hellinger. Boucher et Cline (2007, 2009) s'intéressent à la stabilité en terms d'ergodicité V-uniforme ou de transience pour des modèles TAR cycliques. La stabilité découle de l'existence de lois de probabilité stationnaires des chaînes de Markov sous jacentes.

Mais avant de s'engager dans la voie des modèles à seuils, il faut, d'abord, s'assurer que nos données d'observations sont susceptibles d'accepter ce cadre mathématique pour traduire leur évolution. Petrucci et Davies (1986) ont proposé un test portmanteau pour détecter une non linéarité de type autorégression à seuils. Mélard et Roy (1988) ont considéré un modèle ARMA avec seuils qui est une

généralisation des modèles à seuils TAR existants. Ils décrivent une méthode d'estimation des paramètres et modélisent des séries de données réelles. Tsay (1989) a proposé un test de non linéarité de type threshold basé sur les erreurs prédictifs et une procédure de modélisation pour un modèle TAR quelconque. Qian (1998) s'est intéressé à un modèle SETAR (2; p , p). Il a déterminé les propriétés de l'estimateur du maximum de vraisemblance lorsque les erreurs ne sont pas gaussiennes. Ling et Tong (2005) ont proposé un test pour tester la présence de seuils dans un modèle moyenne mobile.

Ce mémoire comporte cinq chapitres.

Dans le chapitre 0 intitulé : '**Aperçu sur les modèles non linéaires**'.

Dans le chapitre 1, intitulé '**Stabilité des chaînes de Markov**', nous introduisons les principaux outils et concepts probabilistes utilisés dans les chapitres suivants. Ce chapitre s'appuie sur l'ouvrage de Meyn et Tweedie (1993).

Le chapitre 2, intitulé '**Structure probabiliste de modèles à seuils**' présente les classes de modèles considérées et établit des conditions d'ergodicité, de stationnarité stricte et de finitude des moments. Nous commençons par considérer le modèle TAR (1) simple, ensuite multiple, après, nous passons au modèle moyenne mobile à seuils TMA(1) simple, puis général, et nous finissons par la classe la plus large, celle des modèle DTARMACH dont nous donnons des exemples de ce modèle.

Le chapitre 3, intitulé '**Estimation dans des modèles autorégressifs à seuils**', est consacré à, l'estimation des moindres carrés pour les modèles TAR(1) simple et multiple à titre d'exemple, ces résultats sont aisément généralisés au cas d'un TAR(p) général. Ainsi, une étude de simulation est menée pour montrer la consistance de ces estimateurs. On étudie ensuite une méthode d'estimation fondée sur le maximum de vraisemblance pour un TAR(p) à deux régimes. Les propriétés asymptotiques des estimateurs obtenus sont dérivées.

Le chapitre 4 intitulé '**Application et simulations**' présente, dans une première partie, une méthode de test, donnée par Tsay (1989). La deuxième partie met en œuvre les méthodes précédentes sur une série de données réelles : le nombre de lynx canadiens pris au piège.

Chapitre 0 : Aperçu sur les modèles non linéaires

0.1 Introduction

Les exigences du terrain ont fait se développer des alternatives à la modélisation linéaire, attrayante tant du point de vue de l'étude probabiliste que de celui de l'analyse statistique, mais limitative quand à la prise en charge de cycles, de dépendance amplitude-fréquence, etc. Ce sont les modèles non linéaires dont nous présentons quelques exemples.

0.2 Autorégression non linear (NLAR)

C'est la classe de modèles qui répondent à la formulation suivante :

$$X_t = \tilde{f}(X_{t-1}, X_{t-2}, \dots, X_{t-k}, \varepsilon_t) \quad t \in \mathbb{Z} \quad (0.1)$$

où \tilde{f} est une fonction $\tilde{f} : \mathbb{R}^{k+1} \rightarrow \mathbb{R}$, avec $\{\varepsilon_t\}$, une séquence de variables aléatoires indépendantes et identiquement distribuées. Typiquement, ε_t est indépendant de X_s , $s < t$. L'équation (0.1) devient une équation aux différences non linéaire d'ordre k , lorsque $\varepsilon_t = 0, \forall t \in \mathbb{Z}$.

Parfois, il convient d'écrire l'équation (0.1) matriciellement par l'introduction d'un vecteur d'état ξ_t , k -dimensionnel et un vecteur de bruit η_t , k -dimensionnel avec :

$$\begin{aligned} \xi_t &= (X_{t-1}, X_{t-2}, \dots, X_{t-k+1})' \\ \eta_t &= (\varepsilon_t, 0, \dots, 0)' \end{aligned}$$

c'est à dire que, (0.1) peut s'écrire comme :

$$\xi_t = \phi(\xi_{t-1}, \eta_t)$$

où $\phi : \mathbb{R}^{2k} \rightarrow \mathbb{R}$ définie par :

$$\begin{aligned}\xi_t^{(1)} &= \tilde{f}(\xi_{t-1}^{(1)}, \dots, \xi_{t-1}^{(k)}, \eta_t^{(1)}) \\ \xi_t^{(2)} &= \xi_{t-1}^{(1)} \\ &\vdots \\ \xi_t^{(k)} &= \xi_{t-1}^{(k-1)}\end{aligned}$$

$\xi_t^{(j)}$ est la $j^{\text{ème}}$ composante de ξ_t .

Si ϕ est telle que :

$$\phi(\xi_{t-1}, \eta_t) = \phi(\xi_{t-1}, 0) + \eta_t \quad (0.2)$$

l'équation (0.2) définit un modèle de bruit additif et elle est équivalente à :

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-k}) + \varepsilon_t \quad (0.3)$$

où $f : \mathbb{R}^k \rightarrow \mathbb{R}$. Fréquemment dans la littérature, l'acronyme NLAR(k) concerne une classe restreinte définie par (0.3).

Aucune forme fonctionnelle n'a été spécifiée pour f et \tilde{f} . En pratique, les formes de ces derniers sont tirées directement des données, cela conduit à ce qu'on appelle 'l'approche non paramétrique de la fonction autorégressive'. Alternative-ment, on peut paramétrer f et \tilde{f} , cela est appelé 'l'approche autorégressive fini paramétrique'.

Notons qu'il est parfois instructif de penser à un modèle NLAR de la forme:

$$X_t = E[X_t | X_{t-1}] + \{Var[X_t | X_{t-1}]\}^{1/2} e_t$$

que l'on peut comparer à un processus de diffusion dans le cas continu.

0.3 Principe de threshold et les modèles de threshold

Le concept de threshold (seuil) est basé sur l'idée de l'approximation locale sur les états, i.e. l'introduction de régimes via seuils. Le principe de threshold, permet d'analyser un système stochastique complexe en le décomposant en sous-systèmes

plus simples. Sous le principe de threshold, on peut regrouper des modèles non linéaires fini-paramétriques.

Soit $\{X_t\}$ une série chronologique k -dimensionnelle. Pour tout t , J_t est une variable aléatoire qui prend les valeurs $\{1, 2, \dots, l\}$. La forme canonique d'un modèle threshold peut s'écrire comme suit:

$$X_t = B^{(J_t)} X_t + A^{(J_t)} X_{t-1} + H^{(J_t)} \varepsilon_t + C^{(J_t)} \quad (0.4)$$

où, pour $J_t = j$, $A^{(j)}$ et $H^{(j)}$ sont les matrices des coefficients d'ordre k (non-aléatoires), $C^{(j)}$ est un vecteur constant de dimension k et $\{\varepsilon_t\}$ une suite de vecteurs aléatoires i.i.d k -dimensionnels de moyenne nulle.

Dans la forme canonique au-dessus, J_t indique le mode du mécanisme dynamique. Le choix de J_t est flexible, pour modéliser un grand nombre de situations. Effectivement, le choix de J_t qui est une fonction de X_t est d'un intérêt évident, du fait que ce choix est relié aux concepts de threshold et de délai.

Il existe quelques cas particuliers du modèle (0.4). Soit $\{r_0, r_1, \dots, r_l\}$, $r_0, r_1, \dots, r_l \in \mathbb{R}$, tels que $r_0 < r_1 < \dots < r_l$, avec : $r_0 = -\infty$ et $r_l = +\infty$, i.e., que l'on définit une partition de \mathbb{R} , $\mathbb{R} = R_1 \cup R_2 \cup \dots \cup R_l$, où $R_i =]r_{i-1}, r_i]$.

0.3.1 La classe des modèles self-exciting threshold autoregressive / moving average (SETAR / MA)

Soit :

$$X_t = (X_{t-1}, X_{t-2}, \dots, X_{t-k+1})'$$

$$A^{(j)} = \left[\begin{array}{cccc|c} a_1^{(j)} & a_2^{(j)} & \dots & a_{k-1}^{(j)} & a_k^{(j)} \\ \hline & & I_{(k-1) \times (k-1)} & & \mathbf{0}_{(k-1) \times 1} \end{array} \right]$$

$$H^{(j)} = \left[\begin{array}{c|c} h_1^{(j)} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right]$$

$$\varepsilon_t = (\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-k+1})'$$

$$C^{(j)} = (a_0^{(j)}, 0, \dots, 0)'$$

et on note $R_{j,d}^{(k)} = \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R} \times R_j \times \mathbb{R} \times \dots \times \mathbb{R}$, le cylindre du produit Cartésien de \mathbb{R} , k fois figurant et de R_j . Pour d fixé, $d \in \{1, 2, \dots, k\}$, avec $J_t = j$, si $X_t \in R_{j,d}^{(k)}$, on a :

$$X_t = a_0^{(j)} + \sum_{i=1}^k a_i^{(j)} x_{t-i} + h_1^{(j)} \varepsilon_t. \quad (0.5)$$

Conditionnellement à $X_{t-d} \in R_j$, $j = 1, 2, \dots, l$, et comme $\{J_t\}$ est maintenant fonction de $\{X_t\}$, on appelle la série univariée $\{X_t\}$ donnée par (0.5), 'un modèle self-exciting threshold autoregressive d'ordre $(l; k, \dots, k)$, et on le note par SETAR $(l; k, \dots, k)$, où k est répété l fois, le graphe ci-dessous montre une réalisation de ce modèle.

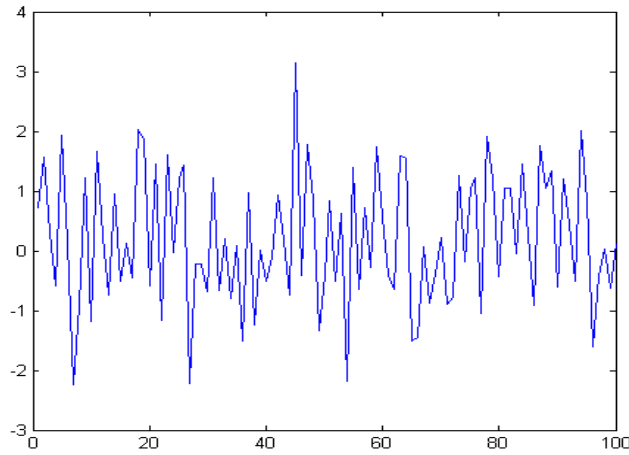


Figure 0.1. $X_t = \begin{cases} 1 + .6X_{t-1} + \varepsilon_t & \text{si } X_{t-1} \leq 0 \\ -1 + .6X_{t-1} + \varepsilon_t & \text{si } X_{t-1} > 0 \end{cases}$ avec $\varepsilon_t \rightarrow N(0, 1)$

Si pour $j = 1, 2, \dots, l$, on a :

$$a_i^{(j)} = 0 \text{ pour } i = k_j + 1, k_j + 2, \dots, k$$

alors, $\{X_t\}$ suit un SETAR $(l; k_1, k_2, \dots, k_l)$. On appelle r_1, \dots, r_{l-1} les paramètres de threshold et d le paramètre de délai.

Notons qu'un SETAR $(1; k)$ est exactement un modèle AR linéaire d'ordre k , si $h_1^{(j)} \varepsilon_t = \varepsilon_t^{(j)}$, la suite $\{\varepsilon_t\}$ devient un processus bruit blanc (hétérogène). Si la première ligne de $H^{(j)}$ est de la forme $(h_1^{(j)}, h_2^{(j)}, \dots, h_k^{(j)})$, $h_k^{(j)} \neq 0$ ($j = 1, 2, \dots, l$), alors, on a le modèle SETAR se généralise à un modèle 'self-exciting threshold autoregressive / moving-average model' (un modèle autorégressif et moyenne mobile à

seuil) d'ordre $(l; k, \dots, k; k-1, \dots, k-1)$, ou SETARMA $(l; k, \dots, k; k-1, \dots, k-1)$, et il prend la forme :

$$X_t = a_0^{(j)} + \sum_{i=1}^k a_i^{(j)} x_{t-i} + \sum_{i=0}^{k-1} h_i^{(j)} \varepsilon_{t-i}$$

Conditionnellement à $X_{t-d} \in R_j$, $j = 1, 2, \dots, l$, SETARMA $(l; k_1, \dots, k_l; k'_1, \dots, k'_l)$, avec k_j et k'_j sont respectivement l'ordre de la partie autorégressive et de la partie moyenne mobile.

0.4 La classe des modèle à seuils sans bruit additif

Un exemple typique prend la forme

$$X_t = a_0^{(J_t)} + a_1^{(J_t)} X_{t-1}$$

où $\{J_t\}$ est une suite de variables aléatoires et J_t est indépendante de X_s , $s < t$. La figure suivante illustre une réalisation d'un modèle de cette classe, qui peut être identifié comme une fractale.

Cette figure représente 100000 réalisations de

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0.25 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

avec la probabilité 0.1,

$$\begin{bmatrix} 0.85 & 0.04 \\ -0.04 & 0.85 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 1.6 \end{bmatrix}$$

avec la probabilité 0.76,

$$\begin{bmatrix} 0.20 & -0.26 \\ 0.26 & 0.22 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0.8 \end{bmatrix}$$

avec la probabilité 0.07,

$$\begin{bmatrix} -0.15 & 0.28 \\ 0.26 & 0.24 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

avec la probabilité 0.07, et

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

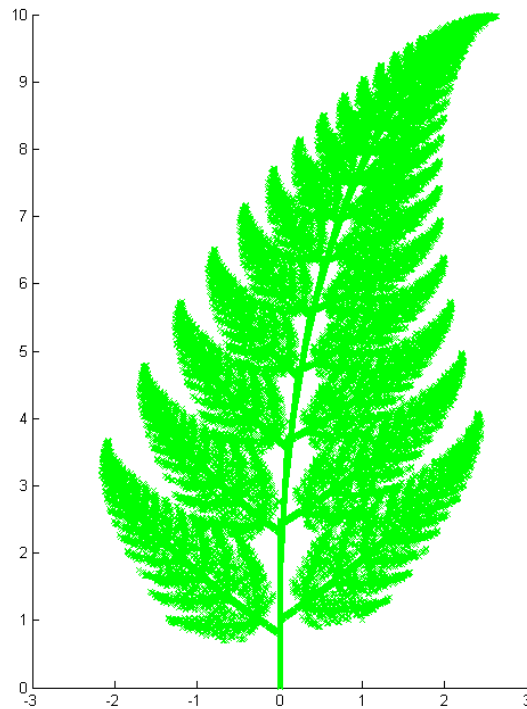


Figure 0.2.

0.4.1 La classe des modèles autorégressifs exponentiels (EAR)

La classe des modèles EAR et ses extensions introduite par Lawrance et Lewis, peut être formellement considérée comme une sous-classe des modèles de threshold avec $\{J_t\}$ indépendante de $\{X_t\}$. On peut illustrer ce point par la représentation d'un EAR(2) qui a la forme d'un threshold

$$X_t = a^{(J_t)} X_{t-1} + b^{(J_t)} X_{t-2} + \varepsilon_t$$

où $\{J_t\}$, est une suite de variables aléatoire i.i.d. telles que :

$$J_t = \begin{cases} 1 & \text{avec la probabilité } 1 - \alpha_2 \\ 2 & \text{avec la probabilité } \alpha_2 \end{cases}$$

Ici, $\{J_t\}$ est indépendante de $\{X_t\}$ et de $\{\varepsilon_t\}$, avec

$$a^{(1)} = \alpha_1 \quad a^{(2)} = 0 \quad b^{(1)} = 0 \quad b^{(2)} = \alpha_2$$

α_1 et α_2 sont des constantes ($0 < \alpha_1, \alpha_2 < 1$).

0.4.2 La classe des modèles ARMA à coefficients périodiques (PARMA)

Cette classe de modèle a été étudiée par Tiao et Grupe (1980), Bentarzi (1995) et bien d'autres, dans un contexte général, et par Jones et Brelsford (1967) dans un contexte météorologique. L'idée essentielle est de permettre aux coefficients des modèles ARMA de varier périodiquement dans le temps. Un exemple simple est

$$X_t = a_0(t) + a_1(t)X_{t-1} + \varepsilon_t$$

où pour $i = 0, 1, 2, \dots$, $a_i(2n) = a_i^{(1)}$ et $a_i(2n+1) = a_i^{(2)}$. Ce type de modèle peut être facilement mis sous une forme de threshold comme suit :

$$X_t = a_0^{(J_t)} + a_1^{(J_t)}X_{t-1} + \varepsilon_t$$

avec la probabilité 1 pour tout t , $J_{2t} = 1$ et $J_{2t+1} = 2$.

0.5 la classe des modèles autorégressifs multiplicatifs

Tous les modèles qu'on a présentés nécessitent un bruit additif. Cela, effectivement n'est pas toujours essentiel. Un exemple de modèle autorégressif non linéaire à bruit non additif, est

$$X_t = \varepsilon_t X_{t-1}^\alpha$$

où $\{\varepsilon_t\}$, est une suite de variables aléatoires positives i.i.d.

0.6 La classe des modèles autorégressifs à coefficient aléatoire (RCA)

L'idée d'un bruit multiplicatif implicite dans la dernière classe de modèles peut être exploitée de telle sorte qu'on peut envisager une nouvelle classe de modèles, en l'occurrence, la classe des modèles autorégressifs à coefficients aléatoires, notée RCA. Une série chronologique $\{X_t\}$ est dite suivre un RCA(k), si X_t satisfait l'équation de la forme :

$$X_t = \sum_{i=1}^k [\beta_i + B_i(t)]X_{t-i} + \varepsilon_t \quad (0.6)$$

où

1. $\{\varepsilon_t\}$ est une suite de variables aléatoires i.i.d. de moyenne nulle et de variance σ^2 ;
2. les β_i , $i = 1, \dots, k$, sont des constantes;
3. posons $B(t) = [B_k(t), \dots, B_1(t)]$, $\{B(t)\}$ forme une suite de vecteurs indépendantes de dimension $1 \times k$ de moyenne nulle et $E[B(t)'B(t)] = C$;
4. $\{B(t)\}$ est indépendant de $\{\varepsilon_t\}$.

Il est clair que l'équation (0.6) peut être étendue au cas multidimensionnel, avec $\{\mathbf{X}_t\}$ une série chronologique vectorielle de dimension p , β_i , $\mathbf{B}_i(t)$ des matrices de dimension $p \times p$, $\{\varepsilon_t\}$ un bruit p -dimensionnel. Les conditions (1) - (4) sont facilement généralisées :

- 1'. $\{\varepsilon_t\}$ est une séquence de vecteurs variables aléatoires i.i.d. de moyenne nulle et de matrice covariance G ;
- 2'. les β_i , $i = 1, \dots, k$, sont des constantes;
- 3'. $\{\mathbf{B}(t)\}$, une séquence de matrices indépendantes de taille $p \times kp$ de moyenne nulle et $E[\mathbf{B}(t)' \otimes \mathbf{B}(t)] = C$, \otimes est l'opérateur du produit de Kronecker;
- 4'. $\{B(t)\}$ est indépendant de $\{\varepsilon_t\}$.

La figure au-dessous illustre une réalisation d'un modèle RCA(1).

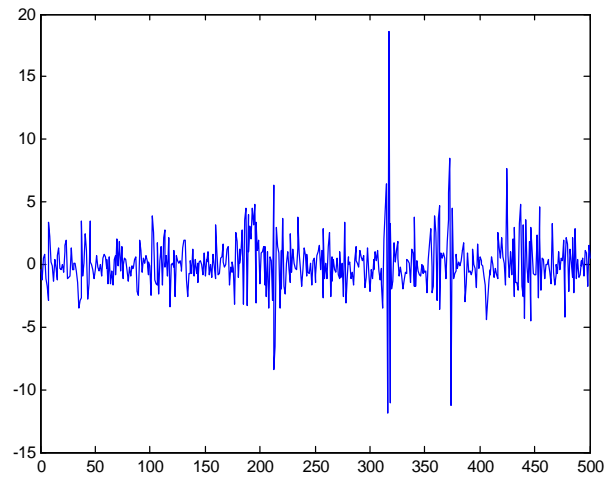


Figure 0.3. $X_t = [-0.1 + B_1(t)]X_{t-1} + \varepsilon_t$, avec $B_1(t) \rightarrow N(0, 0.9^2)$, $\varepsilon_t \rightarrow N(0, 1)$
 $\{B_1(t)\}$ indépendant de $\{\varepsilon_t\}$.

Chapitre 1 : Stabilité des chaînes de Markov

1 Stabilité des chaînes de Markov

1.1 Introduction

De nombreux modèles de séries chronologiques acceptent une représentation markovienne et de là, les résultats connus sur les chaînes de Markov à espace d'états continu peuvent être appliqués pour l'étude de ces modèles. Une approche du problème de la stationnarité figure dans les travaux de Tweedie (1974, 1975, 1983, 1988) et dans l'ouvrage de Meyn et Tweedie (1993). Cette approche est, par exemple, utilisée lorsque la spécification du modèle comporte des effets de seuils excluant l'existence d'une représentation linéaire.

1.2 Noyau de transition et définition d'une chaîne de Markov

La définition suivante introduit la notion de noyau de transition sur un ensemble quelconque \mathcal{X} muni d'une tribue \mathcal{B} .

Définition 1.2.1 *On appelle noyau de transition toute famille $P = \{P(x, B), x \in \mathcal{X}, B \in \mathcal{B}\}$ telle que :*

1. *pour tout $B \in \mathcal{B}$, $x \rightarrow P(x, B)$ est une fonction mesurable de $(\mathcal{X}, \mathcal{B})$ dans $([0, 1], \mathcal{B}([0, 1]))$;*
2. *pour tout $x \in \mathcal{X}$, $P(x, \cdot)$ est une mesure de probabilité sur $(\mathcal{X}, \mathcal{B})$.*

Les $P(x, B)$ sont appelées probabilités de transition (de x vers B).

Etant donnée une mesure de probabilité initiale μ sur $(\mathcal{X}, \mathcal{B})$, la définition d'une chaîne de Markov homogène (X_n) sur $(\mathcal{X}, \mathcal{B})$ repose sur ses distributions finidimensionnelles. Pour tout entier n et tout $(n+1)$ -uplet (B_0, \dots, B_n) de parties de \mathcal{B} , on pose :

$$P_\mu [X_0 \in B_0, \dots, X_n \in B_n] =$$

$$\int_{x_0 \in B_0} \dots \int_{x_{n-1} \in B_{n-1}} \mu(dx_0)P(x_0, dx_1) \dots P(x_{n-1}, B_n) \quad (1.1)$$

pour les passages des lois fini-dimensionnelles aux probabilités d'événements impliquant un nombre quelconque de variable X_n , le théorème de Kolmogorov impose des conditions à l'espace d'état : il s'applique si $(\mathcal{X}, \mathcal{B})$ est : *i*) $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, ou *ii*) un espace dénombrable muni de la tribu de toutes ses parties, ou *iii*) un espace métrique complet possédant une base dénombrable d'ouverts muni de sa tribu borélienne. Nous nous placerons toujours dans l'un de ces cas par la suite. A condition de choisir comme espace Ω l'espace $\mathcal{X}^{\mathbb{N}}$, muni de la tribu produit $\mathcal{B}^{\otimes \mathbb{N}}$, on peut construire un processus $X = (X_n)_{n \in \mathbb{N}}$ et une probabilité P_μ sur $\mathcal{B}^{\mathbb{N}}$ telle que (1.1) soit vraie et pour tout $B \in \mathcal{B}^{\mathbb{N}}$, $P_\mu(B)$ soit la probabilité de l'événement $[X \in B]$. Remarquons en particulier que, si $\{x\} \in \mathcal{B}$ et $\mu(\{x\}) \neq 0$, $P(x, B) = P_\mu[X_1 \in B \mid X_0 = x]$.

Définition 1.2.2 *le processus $X = (X_n)$ dont les lois fini dimensionnelles satisfont (1.1) est appelé chaîne de Markov homogène de noyau de transition P et de mesure initiale μ .*

Etant donnés deux noyaux de transition $P = \{P(x, B)\}$ et $Q = \{Q(x, B)\}$ sur $(\mathcal{X}, \mathcal{B})$, il est utile d'introduire leur produit $PQ = \{[PQ](x, B)\}$ défini par :

$$[PQ](x, B) = \int_{y \in \mathcal{X}} P(x, dy)Q(y, B), \quad \forall x \in \mathcal{X}, \forall B \in \mathcal{B}$$

On vérifie facilement que cette égalité définit un nouveau noyau de transition sur $(\mathcal{X}, \mathcal{B})$. En particulier, on notera P^n le noyau de transition obtenu en effectuant $(n - 1)$ produits de P avec lui même. On a pour $n \geq 1$: $\forall x \in \mathcal{X}, \forall B \in \mathcal{B}$, $P^n(x, B) = \int_{y \in \mathcal{X}} P(x, dy)P^{n-1}(y, B)$, en prenant pour $P^0(y, B)$ la mesure de Dirac en y . On peut vérifier que $P^n(x, B)$ s'interprète comme la probabilité de l'évènement $X_n \in B$ conditionnellement à $X_0 = x$.

Une généralisation utile est la relation de Chapman-Kolmogorov :

$$\forall n \in \mathbb{N}, \forall m, 0 \leq m \leq n, \forall x \in \mathcal{X}, \forall B \in \mathcal{B}$$

$$P^n(x, B) = \int_{y \in \mathcal{X}} P^m(x, dy)P^{n-m}(y, B)$$

On montre cette relation à partir de (1.1) en choisissant $B_i = \mathcal{X}$ pour $1 \leq i \leq n - 1$ et $B_0 = \{x\}$. Cette équation admet une interprétation intéressante : pour atteindre

B en n étapes, partant de x , la chaîne doit nécessairement passer par une valeur y en m étapes, puis tout se passe comme si la chaîne redémarrait de y , pour atteindre B en $n - m$ étapes. Les valeurs prises antérieurement en y n'ont donc plus d'importance pour ce qui est postérieur à l'étape m . Plus généralement, la propriété suivante traduit l'indépendance entre futur et passé conditionnellement au présent :

Propriété 1.2.1 Si $X = (X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov homogène

de mesure initiale μ , pour toute fonction mesurable bornée $h : \Omega \rightarrow \mathbb{R}$, on a

$$E[h(X_{n+1}, X_{n+2}, \dots) \mid X_0, \dots, X_n; X_n = x] = E[h(X_1, X_2, \dots) \mid X_0 = x]$$

[Voir Meyn et Tweedie (1993), proposition 3.4.3]

1.3 Irréductibilité et apériodicité

L'irréductibilité traduit le fait que, partant de n'importe quel point, la chaîne peut atteindre toute partie de mesure positive de l'espace d'état, pour une mesure appropriée. Au préalable, nous définissons pour tout $B \in \mathcal{B}$,

$$\text{- le temps de séjour dans } B : \eta_B = \sum_{n=1}^{\infty} \mathbf{1}_{X_n \in B}$$

où $\mathbf{1}_A$ désigne la variable indicatrice d'un événement A ;

$$\text{- le temps de premier passage par } B : \tau_B = \min\{n \geq 1, X_n \in B\}.$$

La durée moyenne de séjour dans B partant de $x \in \mathcal{X}$ s'obtient par :

$$E[\eta_B \mid X_0 = x] = \sum_{n=1}^{\infty} P^n(x, B) = U(x, B)$$

La famille $\{U(x, B)\}$ a des propriétés similaires à celles d'un noyau de transition mais les $U(x, \cdot)$ ne sont pas des probabilités. On utilisera également la probabilité de premier passage par B partant de x :

$$L(x, B) = P(\tau_B < \infty \mid X_0 = x)$$

Définition 1.3.1 (φ -irréductibilité) On dit que la chaîne $X = (X_n)$ est φ -irréductible s'il existe une mesure φ sur $(\mathcal{X}, \mathcal{B})$ telle que, $\forall B \in \mathcal{B}$ avec $\varphi(B) > 0$, nous ayons $L(x, B) > 0$ pour tout $x \in \mathcal{X}$.

les formulations alternatives suivantes sont équivalentes à la φ – irréductible :

- (i) $\forall x \in X$, quand $\varphi(B) > 0$, $U(x, B) > 0$;
- (ii) $\forall x \in X$, quand $\varphi(B) > 0$, il existe $n > 0$, telle que $P^n(x, B) > 0$.

Il est clair, d'après la définition ci-dessus, que la φ – irréductibilité est quelque peu faible par rapport à l'irréductibilité dans un espace dénombrable, qui exige la communication dans les deux sens. On définit ainsi une extension de la mesure d'irréductibilité, dite "maximale", dont l'existence est assurée par la propriété suivante :

Propriété 1.3.1 *si $(X_n)_{n \in \mathbb{N}}$ est φ – irréductible, il existe une mesure de probabilité ψ sur $(\mathcal{X}, \mathcal{B})$ telle que*

- (i) $(X_n)_{n \in \mathbb{N}}$ est ψ – irréductible;
- (ii) pour toute autre mesure φ' sur $(\mathcal{X}, \mathcal{B})$, (X_n) est φ' – irréductible si et seulement si φ' est absolument continue par rapport à ψ ($\psi \succ \varphi'$)
- (iii) si $\psi(B) = 0$, alors $\psi\{y : L(y, B) > 0\} = 0$.

La mesure ψ est appelée mesure d'irréductibilité maximale.

[Voir Meyn et Tweedie (1993), Proposition 4.2.2]

Notons l'ensemble des événements de ψ – mesure positive par : $B^+ = \{B \in \mathcal{B} : \psi(B) > 0\}$; cet ensemble est défini d'une manière unique du fait que toutes les mesures d'irréductibilité maximale sont équivalentes (Meyn et Tweedie (1993)).

La théorie des chaînes de Markov à espaces d'états quelconques peut être développée par analogie avec celle à espace d'état dénombrable quand \mathcal{X} contient un atome pour la chaîne X .

Définition 1.3.2 (Atomes) *Un ensemble $\alpha \in \mathcal{B}$ est dit "atome" s'il existe une mesure ν sur $(\mathcal{X}, \mathcal{B})$ telle que*

$$P(x, B) = \nu(B), \quad x \in \alpha.$$

Si le chaîne X est ψ – irréductible et $\psi(\alpha) > 0$, on dit que α est un "atome accessible".

Un singleton représente toujours un atome. Cependant, l'intérêt d'utiliser les atomes réside dans les chaînes ψ -irréductibles à espace d'état quelconque. Par une extension convenable de la structure probabiliste de la chaîne, nous pouvons construire artificiellement des ensembles qui ont une structure atomique, permettant ensuite de continuer les analyses en s'inspirant de la théorie des chaînes à espace d'état dénombrable.

Définissons maintenant deux types d'ensembles d'importance majeure : les "small set" et "petite set" qui ont un comportement analogue à celui des atomes.

Définition 1.3.3 (Small sets) *Un ensemble $C \in \mathcal{B}$, est dit small set s'il existe $m > 0$, et une mesure ν_m non triviale sur $(\mathcal{X}, \mathcal{B})$ telle que*

$$\forall x \in C, \forall B \in \mathcal{B}, P^m(x, B) \geq \nu_m(B)$$

et on dit que C est ν_m -small.

Proposition 1.3.1

- (i) *Si $C \in \mathcal{B}$ est ν_n -small, et pour tout $x \in D$ nous avons $P^m(x, C) \geq \delta, (\delta > 0)$, alors D est ν_{n+m} -small, où ν_{n+m} est un multiple de ν_n .*
- (ii) *Si X est ψ -irréductible. alors il existe une suite dénombrable C_i de small sets dans \mathcal{B} telle que $\mathcal{X} = \bigcup_{i=1}^{\infty} C_i$. Les small sets constituent ainsi un recouvrement de l'espace d'état de la chaîne de Markov.*

[Voir Meyn et Tweedie (1993), Proposition 5.2.4]

L'existence des small sets nous permet de montrer, même sur un espace quelconque, que nous avons une rupture périodique finie entre des ensembles cycliques pour la chaîne ψ -irréductible.

Supposons que C est un ensemble ν_M -small, et que $\nu_M(C) > 0$, nous utilisons C et ν_M pour définir un cycle d'une chaîne de Markov ψ -irréductible à espace d'état quelconque. Pour simplifier la notation, on enlève l'indice de ν , alors on a $\forall x \in C, P^M(x, \cdot) \geq \nu(\cdot)$, et $\nu(C) > 0$, ainsi, quand la chaîne démarre de C , elle y revient avec une probabilité positive à l'instant M .

Soit

$$E_C = \{n \geq 1 : C \text{ est } \nu_n\text{-small, avec } \nu_n = \delta_n \nu, \text{ pour un } \delta_n > 0\}$$

l'ensemble des entiers n tels que C est ν_n -small pour une mesure proportionnelle à ν . Notons que, si pour un ensemble $B \subseteq C$, $n, m \in E_C$ on a

$$P^{n+m}(x, B) \geq \int_C P^m(x, dy) P^n(y, B) \geq [\delta_m \delta_n \nu(C)] \nu(B), \quad x \in C;$$

alors E_C est fermé par addition. D'où il existe une "période" naturelle pour l'ensemble C , donnée par le *p.g.c.d* de E_C , cette valeur est en effet, propriété de toute la chaîne, et elle ne dépend pas du choix de C .

Définition 1.3.4 *Soit X une chaîne de Markov φ -irréductible.*

Le plus grand entier d pour lequel X soit d -cyclique est appelé période de la chaîne.

Quand $d = 1$, la chaîne est dite apériodique.

S'il existe un ensemble B , ν_1 -small avec $\nu_1(B) > 0$, la chaîne est dite fortement apériodique.

Les small sets existent toujours dans le cas ψ -irréductible, et ils nous fournissent la plupart des propriétés demandées.

Introduisons maintenant une généralisation des small sets, les petite sets, qui ont des propriétés plus manipulables, spécialement dans une analyse topologique.

Définition 1.3.5 (Petite sets) *On appelle ν_a -petite set tout $C \in \mathcal{B}$ tel que :*

$$K_a(x, B) := \sum_{n=0}^{\infty} a_n P^n(x, B) \geq \nu_a(B), \quad \forall x \in C, \forall B \in \mathcal{B}$$

où $a = (a_n)$ une suite de constantes positives de somme égale à 1 et ν_a est une mesure non triviale sur \mathcal{B} .

Nous pouvons remarquer de cette définition qu'un small set est un petite set si a prend la valeur δ_m pour un certain m . D'où, la propriété d'être small set est plus forte que la propriété d'être petite set.

Remarques 1.3.1 *Soit X une chaîne de Markov ψ -irréductible.*

(i) Si C est ν_a -petite, il existe (b_n) tel que C soit ν_b -petite, avec ν_b mesure maximale d'irréductibilité.

(ii) L'union de deux petite (resp. small) sets est également un petite (resp. small) set.

(iii) Il existe une suite croissante $\{C_i\}$ de ν_c -petite sets, tel que $\mathcal{X} = \bigcup_{i=1}^{\infty} C_i$.

[Voir Meyn et Tweedie (1993), Proposition 5.5.5]

Théorème 1.3.1 *Pour une chaîne irréductible et apériodique, les petite sets sont des small sets.*

[Voir Meyn et Tweedie (1993), Théorème 5.5.7]

1.4 Chaîne récurrente, chaîne transitoire

Lorsque \mathcal{X} est dénombrable et la chaîne (X_n) irréductible, un état est dit transitoire si le temps de séjour moyen dans cet état, partant de ce même état, est fini et il est dit récurrent dans le cas contraire. De plus, tous les états sont de même nature et la chaîne peut être qualifiée de récurrente ou transitoire.

On retrouve la même dichotomie pour un espace d'états quelconque, mais les définitions sont plus délicates. La classification repose sur le (pseudo) noyau de transition $\{U(x, B)\}$, $U(x, B)$ représentant le temps de séjour moyen dans B partant de x .

Définition 1.4.1 *L'ensemble $B \in \mathcal{B}$ est dit*

(i) *récurrent si $U(x, B) = +\infty, \forall x \in B$;*

(ii) *uniformément transitoire si $\exists M < \infty$ t.q $U(x, B) < M, \forall x \in B$;*

(iii) *transitoire s'il existe une famille dénombrable (B_n) d'ensembles uniformément transitoires telle que $B = \bigcup_n B_n$.*

Avec cette définition, un ensemble récurrent peut également être transitoire. Ainsi \mathcal{X} est toujours récurrent et s'il est dénombrable, on peut évidemment l'écrire comme la réunion de tous les singletons qui sont uniformément transitoires et \mathcal{X} est transitoire.

. Le résultat suivant fournit une première dichotomie de l'ensemble des chaînes irréductibles.

Définition 1.4.2 On note R la classe des chaîne récurrentes, définies par :

$$(X_n) \in R \iff U(x, B) = +\infty, \quad \forall x \in \mathcal{X}, \forall B \in \mathcal{B}^+$$

On note T la classe des chaînes transitoires, définie par :

$$(X_n) \in T \iff \exists (B_j)_j, U(x, B_j) \leq M_j < \infty, \quad \forall x \in \mathcal{X}.$$

Théorème 1.4.1 Si (X_n) est une chaîne ψ -irréductible, elle est soit transitoire soit récurrente.

[Voir Meyn et Tweedie (1993), Théorème 8.3.4]

Une notion de récurrence un peu plus forte est celle de récurrence au sens de Harris (Harris-récurrence) fondée sur le calcul de la probabilité que le temps de séjour soit infini.

Définition 1.4.3 Un ensemble B est dit Harris-récurrent si

$$Q(x, B) = P(\eta_B = +\infty \mid X_0 = x) = 1, \quad x \in B.$$

Une chaîne X est dite Harris- récurrente si elle est ψ -irréductible et chaque ensemble dans \mathcal{B}^+ est Harris-récurrent.

Il est clair de cette définition que si un ensemble est Harris-récurrent, alors il est récurrent. En effet, dans la formulation ci-dessus le renforcement de la récurrence vers la récurrence au sens de Harris est explicite car nous passons d'un nombre de visite espéré d'être infini vers l'infinité des visites pour un ensemble presque sûrement.

De manière équivalente, nous pouvons définir une condition pour vérifier si un ensemble est Harris-récurrent, basée seulement sur la probabilité du premier retour $L(x, B)$.

Proposition 1.4.1 On suppose que pour un ensemble $B \in \mathcal{B}$ on a $L(x, B) \equiv 1$, $x \in B$. Alors $Q(x, B) = L(x, B)$ pour tout $x \in \mathcal{X}$, et en particulier B est Harris récurrent.

[Voir Meyn et Tweedie (1993), Proposition 9.1.1]

Le théorème suivant montre l'importance des petites sets.

Théorème 1.4.2 *Si X est une chaîne ψ -irréductible. Alors X est Harris récurrente s'il existe un petite set $C \in \mathcal{B}$ tel que, $\forall x \in \mathcal{X}$, $L(x, C) \equiv 1$.*

[Voir Meyn et Tweedie (1993), Théorème 8.3.4]

1.5 Existence des mesures invariantes

Afin d'obtenir des propriétés de stabilité plus fortes, on est amené à pousser plus loin la classification des chaîne récurrentes. Pour plusieurs raisons, la forme de stabilité la plus forte possible que nous pouvons demander en présence d'une variation persistante est que la distribution de X_n ne change pas avec n . Si cela est le cas, alors, de par la propriété de Markov, les distributions finies dimensionnelles sont invariantes par translation dans le temps. Ceci nous amène à introduire la notion de mesure invariante.

Définition 1.5.1 *Une mesure σ -finie π sur \mathcal{B} tel que*

$$\pi(B) = \int_{\mathcal{X}} \pi(dx)P(x, B), \quad B \in \mathcal{B} \quad [\text{resp.} \quad \pi(B) \geq \int_{\mathcal{X}} \pi(dx)P(x, B)]$$
est dite invariante [resp. sous-invariante].

Si une chaîne X , ψ -irréductible admet une mesure de probabilité invariante, alors X est dite *positive*, sinon elle est dite *nulle*.

Pour une chaîne récurrente, une mesure invariante peut ne pas exister. Lorsqu'une telle mesure existe et est finie, on peut toujours la normaliser pour obtenir une mesure de probabilité invariante. Remarquons qu'étant donnée une mesure de probabilité initiale μ , on a

$$\int \mu(dx)P(x, B) = P_{\mu}[X_1 \in B].$$

Donc si μ est invariante :

$$\begin{aligned}
 \mu(B) &= P_\mu[X_1 \in B] \\
 &= \int [\int \mu(dy)P(y, dx)]P(x, B) \\
 &= \int \mu(dy) \int P(y, dx)P(x, B) \\
 &= \int \mu(dy)P^2(y, B) \\
 &= P_\mu[X_2 \in B]
 \end{aligned}$$

Donc pour tout n , $P_\mu[X_n \in B] = \mu(B)$ ($\forall B \in \mathcal{B}$). En utilisant la propriété de Markov, cela équivaut à la stricte stationnarité de la chaîne : la loi du processus $(X_n, X_{n+1}, \dots, X_{n+k})$ est indépendante de n , pour tout entier k .

Inversement, si (X_n) est strictement stationnaire, X_1 et X_0 ont la même loi, ce qui s'écrit $\forall B \in \mathcal{B}$, $\mu(B) = \int \mu(dx)P(x, B)$, où μ désigne la probabilité initiale. On a donc le résultat suivant.

Propriété 1.5.1 *La chaîne (X_n) est strictement stationnaire si et seulement si sa loi de probabilité initiale est invariante.*

Cette propriété rend cruciale la recherche de mesures invariantes de masse finie. Une autre raison de l'importance des mesures de probabilité invariantes est liée au comportement ergodique (ou de long terme) de la chaîne. Supposons en effet qu'il existe une mesure de probabilité limite π_μ , définie par : $\forall B \in \mathcal{B}$, $P_\mu[X_n \in B] \longrightarrow \pi_\mu(B)$ quand $n \longrightarrow \infty$.

Alors

$$\begin{aligned}
 \pi_\mu(B) &= \lim_{n \rightarrow \infty} \int \mu(dx)P^n(x, B) \\
 &= \lim_{n \rightarrow \infty} \int \mu(dx) \int P^{n-1}(x, dy)P(y, B) \\
 &= \int P(y, B) \lim_{n \rightarrow \infty} \int \mu(dx)P^{n-1}(x, dy) \\
 &= \int P(y, B)\pi_\mu(dy)
 \end{aligned}$$

Donc π_μ est invariante. En particulier, s'il existe une unique mesure de probabilité

invariante, la probabilité limite sera indépendante de μ . On a le résultat suivant :

Théorème 1.5.1 *Si (X_n) est récurrente, elle admet une unique mesure sous-invariante π (à une constante multiplicative près) et cette mesure est invariante. De plus π est équivalente à toute mesure d'irréductibilité maximale ψ .*

[Voir Meyn et Tweedie (1993), Théorèmes 10.4.4 et 10.4.9]

Ce théorème conduit à une nouvelle dichotomie parmi les chaîne récurrentes. On distingue celles admettant une mesure de probabilité invariante des autres. Cette dichotomie s'écrit : $R=P+N$, avec pour $(X_n) \in R$:

$$(X_n) \in P \Leftrightarrow \forall B \in \mathcal{B}, \pi(B) = \int \pi(dx)P(x, B)$$

où π est une mesure de probabilité, et

$$(X_n) \in N \Leftrightarrow \forall B \in \mathcal{B}, \mu(B) \geq \int \mu(dx)P(x, B)$$

où μ est une mesure telle que $\mu(\mathcal{X}) = +\infty$.

1.6 Ergodicité

Jusqu'ici nous avons caractérisé la stabilité d'une chaîne en terme de récurrence. Une autre façon d'envisager la stabilité d'une chaîne consiste à étudier si elle converge vers un régime stationnaire quelque soit son point de départ. Nous avons déjà vu que l'existence d'une mesure de probabilité invariante était une condition nécessaire pour qu'une telle convergence ait lieu. Nous allons voir que pour les chaînes ψ -irréductibles et positives récurrentes la propriété est vérifiée pour une notion forte de la convergence que nous allons préciser.

Rappelons que la norme en variation totale d'une mesure μ est définie par $\|\mu\| = \sup_{f:|f|\leq 1} |\int f d\mu| = \sup_{B \in \mathcal{B}} \mu(B) - \inf_{B \in \mathcal{B}} \mu(B)$, on défini aussi la f -norme par $\|\mu\|_f = \sup_{g:|g|\leq f} |\int g d\mu|$.

Théorème 1.6.1 *Si X est une chaîne de Markov Harris-récurrente, positive et apériodique, alors pour toute mesure initiale μ*

$$\left\| \int \mu(dx)P^n(x, \cdot) - \pi(\cdot) \right\| \longrightarrow 0, \quad n \longrightarrow \infty.$$

[Voir Meyn et Tweedie (1993), Théorèmes 13.3.3]

En prenant pour μ la masse de Dirac en x , un corollaire de ce théorème est que pour toute fonction f bornée et mesurable :

$$\forall x \in \mathcal{X}, \quad \lim_{n \rightarrow \infty} E[f(X_n) | X_0 = x] = \int f d\pi. \quad (1.2)$$

En particulier, pour $f = \mathbf{1}_B$, $B \in \mathcal{B} : P^n(x, B) \rightarrow \pi(B)$, $\forall x \in \mathcal{X}$.

Les chaînes vérifiant les hypothèses du théorème sont dit *ergodique*.

Lorsque f n'est pas supposée bornée, on a le résultat suivant :

Théorème 1.6.2 *Soit X une chaîne de Markov Harris-récurrente, positive de mesure invariante π , fortement apériodique, de mesure d'irréductibilité maximale ψ , et soit $f : \mathcal{X} \rightarrow [1, +\infty[$. Alors*

(i) *si $\int f d\pi = +\infty$, $E[f(X_n) | X_0 = x] \rightarrow +\infty$, $n \rightarrow +\infty$*

(ii) *si $\int f d\pi < +\infty$, alors pour tout $x \in A$, où A est absorbant et tel que $\psi(A^c) = 0$, on a*

$$\|P^n(x, \cdot) - \pi(\cdot)\|_f \rightarrow 0 \quad \text{quand} \quad n \rightarrow +\infty.$$

[Voir Meyn et Tweedie (1993), Théorème 14.3.3]

On obtient en particulier, cette fois-ci pour f non nécessairement bornée (mais sous des hypothèse un peu plus fortes)

$$\forall x \in A, \quad \lim_{n \rightarrow +\infty} E[f(X_n) | X_0 = x] = \int f d\pi. \quad (1.3)$$

Si dans le théorème 1.6.2, la condition (ii) est vraie pour tout x , la chaîne est dite *f-ergodique*. Il est parfois nécessaire de connaître la vitesse de convergence de $\|P^n(x, \cdot) - \pi(\cdot)\|_f$. Sous les hypothèses du théorème précédent, nous dirons que la chaîne est *f-géométriquement ergodique* (ou plus simplement géométriquement ergodique si $f \equiv 1$) s'il existe une constante $r_f > 1$ telle que

$$\sum_{n=1}^{\infty} r_f^n \|P^n(x, \cdot) - \pi(\cdot)\|_f < \infty, \quad \forall x \in \mathcal{X}.$$

Nous donnerons dans la partie suivante un critère assurant une telle propriété.

Remarques 1.6.1 *L'ergodicité géométrique n'assure pas seulement l'existence d'une unique distribution stationnaire pour (X_n) , mais aussi elle assure que les marginales convergent en un temps géométrique vers la distribution stationnaire.*

Enfin (X_n) est dite *uniformément géométriquement ergodique* si

$$\sup_{x \in \mathcal{X}} \| P^n(x, \cdot) - \pi(\cdot) \| \xrightarrow{n \rightarrow +\infty} 0,$$

Le résultat suivant donne des caractérisations équivalentes de l'ergodicité géométrique uniforme.

Théorème 1.6.3 *Sans supposition de la ψ -irréductibilité ou de l'apériodicité, les propriétés suivantes sont équivalentes :*

(i) *X est uniformément ergodique.*

(ii) *Il existe $\rho < 1$ et $R < \infty$ telles que $\| P^n(x, \cdot) - \pi(\cdot) \| \leq R\rho^n$.*

(iii) *Il existe $n \in \mathbb{N}$, $\sup_{x \in \mathcal{X}} \| P^n(x, \cdot) - \pi(\cdot) \| < 1$.*

[Voir Meyn et Tweedie (1993), Théorème 16.2.1]

Une conséquence importante de l'ergodicité uniforme est le lien qui existe avec la notion de mélange.

Théorème 1.6.4 *Si X est uniformément ergodique, il existe $R < +\infty$ et $\rho < 1$ telles que, pour toutes fonctions h et g telles que $|g|, |h| \leq 1, \forall n, k, \forall x \in \mathcal{X}$*

$$| E[g(X_n)h(X_{n+k}) | X_0 = x] - E[g(X_n) | X_0 = x]E[h(X_{n+k}) | X_0 = x] | \leq R\rho^n[1 + \rho^k]$$

la chaîne est dite géométriquement mélangeante.

[Voir Meyn et Tweedie (1993), Théorème 16.1.5]

Aussi, sans avoir besoin de l'irréductibilité on peut donner des critères qui garantissent l'existence d'une mesure de probabilité invariante et la finitude de ses moments.

Définition 1.6.1 *Une chaîne telle que, pour tout ouvert B , la fonction $P(\cdot, B)$ est semi-continue inférieurement est dite Fellerienne ou de Feller.*

Remarques 1.6.2 Dire que la fonction $P(\cdot, B)$ est continue pour tout $B \in \mathcal{B}$ est une contrainte trop forte, par exemple pour les dynamiques avec effets de seuil. Rappelons qu'une fonction h de \mathcal{X} dans \mathbb{R} est dite semi-continue inférieurement si l'ensemble $\{x : h(x) \leq c\}$ (resp. $\{x : h(x) > c\}$) est fermé (resp. ouvert) pour toute constante c

Théorème 1.6.5 Supposons que (X_n) est une chaîne de Feller.

(i) S'il existe pour un ensemble compact $B \in \mathcal{B}$, une fonction g positive et un $\varepsilon > 0$ tels que

$$\int_{B^c} P(x, dy)g(y) \leq g(x) - \varepsilon, \quad x \in B^c,$$

alors, il existe une mesure invariante σ -finie μ pour P avec $0 < \mu(B) < \infty$.

(ii) De plus, si

$$\int_B \mu(dx) \left[\int_{B^c} P(x, dy)g(y) \right] < \infty$$

alors, μ est finie et $\pi = \mu/\mu(X)$ est une mesure de probabilité invariante.

(iii) De plus, si

$$\int_{B^c} P(x, dy)g(y) \leq g(x) - f(x), \quad x \in B^c,$$

alors, μ admet un f -moment fini, i.e. $\int_{\mathcal{X}} \mu(dy)f(y) < \infty$.

Notons que la condition (i) n'est qu'un cas particulier de (iii) quand $f(x) = \varepsilon$. Le théorème 1.6.5 établit que pour un modèle de série chronologique qui admet une représentation Markovienne où (X_n) est la chaîne de Markov sous jacente, il existe une solution strictement stationnaire, et la distribution stationnaire est π si les conditions (i) et (ii) sont vérifiées.

1.7 Critères de classification

Pour des modèles particuliers, il n'est généralement pas facile de vérifier directement les propriétés de chaîne de Markov transitoire, récurrente ou admettant une loi de probabilité invariante. Il existe heureusement des critères s'exprimant simplement à partir des conditions sur les probabilités de transition en une étape.

Ces contraintes sont de deux types : de continuité et de dérive (Drift) moyenne. Pour simplifier la présentation, nous supposons dans cette partie que l'espace d'état X est égal à \mathbb{R}^d , muni de sa topologie usuelle, où d est un entier.

La condition suivante est dite de forte continuité : la fonction $P(., B)$ est continue pour tout $B \in \mathcal{B}$. Cette contrainte se révèle trop forte, par exemple pour les dynamique avec effets de seuil. Rappelons qu'une fonction h de X dans \mathbb{R} est dite semi-continue inférieurement si l'ensemble $\{x : h(x) \leq c\}$ (resp. $\{x : h(x) > c\}$) est fermé (resp. ouvert) pour tout constante c . Une chaîne telle que, pour tout ouvert B , la fonction $P(., B)$ est semi-continue inférieurement est dite *Fellerienne* ou de Feller. On a la définition suivante

Définition 1.7.1 *Si il existe une application $T(.,.)$ telle que*

- (i) $\forall x \in X, \forall B \in \mathcal{B}, P(x, B) \geq T(x, B)$;
- (ii) $T(., B)$ est semi-continue inférieurement pour tout B dans \mathcal{B} ;
- (iii) $\forall x \in X, T(x, .)$ est une mesure non nulle sur $(\mathcal{X}, \mathcal{B})$.

(X_n) est appelée *T-chaîne*.

Une chaîne vérifiant la forte continuité est évidemment une *T-chaîne*. La condition (i) peut être remplacée par des conditions moins fortes, portant sur toutes les probabilités de transition, du type

- (i)' Il existe une suite (a_n) de constantes positives de somme égale à 1 telle que

$$K_a(x, B) = \sum_{n \geq 1} a_n P^n(x, B) \geq T(x, B).$$

T est appelée composante continue de K_a .

Ces contraintes de continuité ne sont pas nécessaires pour la stabilité, mais elles permettent d'obtenir des caractérisations en se limitant aux ensembles compacts.

Théorème 1.7.1 *Si (X_n) est une T-chaîne ψ -irréductible, les compacts sont des petite sets.*

[Voir Meyn et Tweedie (1993), Théorèmes 6.2.5]

L'autre type de condition porte sur l'opérateur de dérive Δ . A toute fonction $V : \mathcal{X} \rightarrow \mathbb{R}^+$, intégrable par rapport aux mesures $P(x, \cdot)$ ($\forall x \in \mathcal{X}$), on associe

$$\Delta V(x) = \int P(x, dy)V(y) - V(x) = E[V(X_{n+1}) - V(X_n) \mid X_n = x]$$

cet opérateur calcule, pour tout point initial x , la valeur moyenne de la variation de $V(x)$ en une étape.

Théorème 1.7.2 *Soit X une chaîne de Markov ψ -irréductible, V une fonction de \mathcal{X} dans \mathbb{R}^+ , intégrable par rapport aux mesures $P(x, \cdot)$ pour tout x , et C un compact. Si la fonction $P(\cdot, B)$ est continue pour tout $B \in \mathcal{B}$, ou si X est une T -chaîne, avec $\Delta V(x) \leq 0$, $\forall x \in C^c$, alors la chaîne X est récurrente.*

[Voir Meyn et Tweedie (1993), Théorème 8.4.3]

En pratique, il s'agit de trouver une telle fonction V dite fonction test ou fonction de Lyapunov, ainsi qu'un compact C permettant de conclure. L'existence d'une loi de probabilité invariante peut être obtenue à partir du résultat suivant.

Théorème 1.7.3 *Sous les hypothèse du théorème 1.7.2 et si*

$$\begin{aligned} \Delta V(x) &\leq -1, & \forall x \in C^c \\ \Delta V(x) &\leq M, & \forall x \in C \end{aligned}$$

où M est une constante, la chaîne (X_n) est récurrente positive.

Dans la première inégalité, la valeur -1 peut être remplacée par n'importe quelle constante strictement négative. Le critère suivant permet d'obtenir l'existence des moments de la probabilité invariante.

Théorème 1.7.4 *Soit (X_n) une chaîne de Markov récurrente positive de loi de probabilité invariante π . Soient V , f et s des fonctions positives telles que*

$$\Delta V(x) \leq -f(x) + s(x), \quad \forall x \in \mathcal{X}.$$

Alors $\int f d\pi \leq \int s d\pi$.

[Voir Meyn et Tweedie (1993), Théorème 14.3.7]

Si dans ce théorème, nous choisissons $s(x) = b\mathbf{1}_C(x)$, où b est une constante positive et $C \in B$, on obtient

$$\int f d\pi \leq b\pi(C) < +\infty.$$

Théorème 1.7.5 *Sous les hypothèses du théorème 1.7.2, si (X_n) est apériodique et s'il existe une fonction $V : \mathcal{X} \rightarrow [1, +\infty[$, et des constantes $\beta > 0$ et $b < +\infty$ et un petite set C , on a pour tout $x \in X$*

$$\Delta V(x) \leq -\beta V(x) + b\mathbf{1}_C(x), \quad (1.4)$$

alors il existe $r > 1$ tel que

$$\sum_{n \geq 0} r^n \|P^n - \pi\|_V < +\infty \quad (1.5)$$

où $\|P^n - \pi\|_V = \sup_{x \in \mathcal{X}} \frac{\|P^n(x, \cdot) - \pi\|_V}{V(x)}$. La chaîne est dite V -uniformément ergodique.

[Voir Meyn et Tweedie (1993), Théorème 16.1.2]

Remarques 1.7.1 (i) *Sous les hypothèse du théorème 1.7.5, on a $\int V d\pi < +\infty$, donc $E_\pi[V(X_n)] < +\infty$, en utilisant le théorème 1.7.4 avec $s(x) = b\mathbf{1}_C(x)$.*

(ii) *Une conséquence immédiate de (1.5) est que*

$$\forall x \in \mathcal{X}, \quad \sum_{n \geq 0} r^n \|P^n(x, \cdot) - \pi(\cdot)\| < RV(x) \quad (1.6)$$

pour R finie. La chaîne est donc géométriquement ergodique.

Il est parfois utile de disposer de critère ne reposant pas sur les probabilités de transition en une étape. Le résultat suivant montre comment adapter ce qui précède.

Théorème 1.7.6 *Sous les hypothèses du théorème 1.7.5, et si on remplace la condition (1.4) par*

$$E[V(X_{n+m})|X_n = x] \leq (1 - \beta)V(x) + b\mathbf{1}_C(x),$$

alors (1.6) est vérifiée. La chaîne est ainsi géométriquement ergodique.

Ce résultat peut être étendu en faisant dépendre de l'état x l'entier m et le réel β .
[Voir Meyn et Tweedie (1993), Théorème 19.1.3]

Chapitre 2 : Structure probabiliste de modèles à seuils

2 Structure probabiliste de modèles à seuils

L'étude de la structure probabiliste de modèles de séries chronologiques est un préalable majeur à l'estimation des paramètres et aux tests d'hypothèses. Dans ce qui suit, nous nous intéressons à l'étude de la structure probabiliste de divers modèles de séries chronologiques à seuils : Les modèles TAR, TMA et DTAR-MACH.

2.1 Modèle autorégressif à seuil TAR(1)

2.1.1 Introduction

Une classe de modèles non linéaires qui est particulièrement utile est celle des modèles autorégressifs à seuil (TAR) introduits par Tong (1978) et Tong et Lim (1980). Plusieurs exemples ont été donnés par ces auteurs qui ont montré que les modèles TAR peuvent donner un meilleur ajustement que les modèles linéaires. En plus, les modèles TAR peuvent refléter un comportement non linéaire sous-jacent (ex. Cycles limites) que les modèles linéaires ne peuvent pas reproduire. Dans Tong et Lim (1980), le problème de l'identification et de l'ajustement ont été considérés, En plus, seulement une condition suffisante a été établie pour l'ergodicité du modèle TAR.

Dans cette section, nous considérons un modèle TAR(1) qui s'écrit sous la forme suivante

$$y_t = \phi_1 y_{t-1}^+ + \phi_2 y_{t-1}^- + \varepsilon_t, \quad t \in \mathbb{N}^* \quad (2.1)$$

avec $y^+ = \max(y, 0)$ et $y^- = \min(y, 0)$. (2.1) peut s'écrit aussi comme

$$y_t = [\phi_1 \mathbf{1}(y_{t-1} > 0) + \phi_2 \mathbf{1}(y_{t-1} \leq 0)] y_{t-1} + \varepsilon_t, \quad t \in \mathbb{N}^* \quad (2.2)$$

dans les deux formes ci-dessus, ϕ_1 et ϕ_2 sont des constantes réelles et $\{\varepsilon_t, t \geq 1\}$ est

une suite de variables aléatoires *i.i.d.*, de moyenne nulle et de densité $f(\cdot)$ strictement positive sur \mathbb{R} .

2.1.2 Ergodicité du modèle TAR(1)

Notons que $\{y_t, t \geq 0\}$, définie par (2.1) est une chaîne de Markov à espace d'état $(\mathbb{R}, \mathcal{B})$, \mathcal{B} est la σ -algèbre de Borel sur \mathbb{R} , la densité de transition est donné par

$$p(x, z) = f(z - \phi_1 x^+ + \phi_2 x^-). \quad (2.3)$$

Si μ est la mesure de Lebesgue sur \mathbb{R} , alors $\{y_t, t \geq 0\}$ est μ -irréductible et aperi-odique (voir Orey (1971)).

Le théorème suivant donne une condition nécessaire et suffisante sur les paramètres ϕ_1 et ϕ_2 pour que $\{y_t\}$ soit ergodique.

Théorème 2.1.1 (Petrucci et Woolford (1984)) *Le processus $\{y_t, t \geq 0\}$, défini par (2.1), est ergodique si et seulement si*

$$\phi_1 < 1, \quad \phi_2 < 1 \quad \text{et} \quad \phi_1 \phi_2 < 1. \quad (2.4)$$

La région d'ergodicité est illustrée ci-dessous.

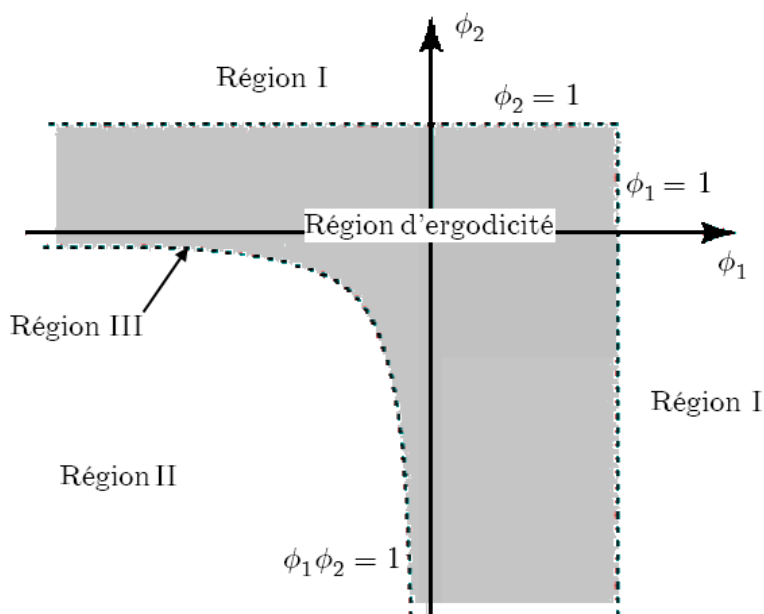


Figure 2.1. Région d'ergodicité

La preuve du théorème 2.1.1 est divisée en quatre lemmes, le premier prouve la suffisance, les autres prouvent la nécessité.

Lemme 2.1.1 *Si ϕ_1 et ϕ_2 satisfont (2.4), alors le processus $\{y_t, t \geq 0\}$ est ergodique.*

Preuve. Définissons la fonction de transition pour la chaîne de Markov $\{y_t\}$ par

$$P(x, A) = \int_A p(x, z) dz, \quad x \in \mathbb{R}, \quad A \in \mathcal{B},$$

il est facile de voir que la loi de transition $\{P(x, \cdot)\}$ est fortement continue (voir Tweedie (1975)). Utilisons le théorème 4.2 de Tweedie (1975), si on arrive à trouver un compact $K \in \mathcal{B}$, de mesure de Lebesgue strictement positive, et une fonction g positive et mesurable sur \mathbb{R} , tels que

$$(i) \int_{\mathbb{R}} p(x, z) g(z) dz \leq g(x) - 1, \quad x \in K^c,$$

$$(ii) \int_{\mathbb{R}} p(x, z) g(z) dz = \lambda(x) \leq R < \infty, \quad x \in K, \quad \text{où } R \text{ est une constante positive.}$$

De (2.4), il est possible de trouver a et b des constantes positives telles que $1 > \phi_1 > -(ba^{-1})$ et $1 > \phi_2 > -(ab^{-1})$. Ainsi, par ce choix de g

$$g(x) = \begin{cases} ax, & x > 0 \\ b|x|, & x \leq 0, \end{cases}$$

il est possible de montrer qu'il existe un $M > 0$ tel que (i) et (ii) soient vérifiées pour $K = [-M, M]$. ■

La preuve de la nécessité de (2.4) est divisée en trois lemmes qui prouvent la non-ergodicité du processus $\{y_t; t \geq 0\}$ pour les valeurs de ϕ_1 et ϕ_2 dans les régions I, II et III respectivement (voir figure 2.1).

Lemme 2.1.2 *Si $\phi_1 \geq 1$ ou $\phi_2 \geq 1$ (Région I), alors le processus $\{y_t\}$ n'est pas ergodique.*

Preuve. Sans restreindre la généralité, on considère le cas $\phi_1 \geq 1$. on suppose d'abord que $\phi_1 > 1$. Alors pour $y_{t-1} > 0$, $E(y_t | y_{t-1}) = \phi_1 y_{t-1}$. Alors pour tout η t.q

$1 < \eta < \phi_1$ et $y_{t-1} > 0$, par l'inégalité de Markov on a

$$P(y_t \leq 2^{-1}(\eta + 1)y_{t-1} \mid y_{t-1}) \leq 2E|\varepsilon_1|[(\eta - 1)y_{t-1}]^{-1}. \quad (2.5)$$

Soit $M > 0$ tel que $c = E|\varepsilon_1|[(\eta - 1)M]^{-1} < 1$, alors, quand $y_1 > M$ on a

$$P(y_2 > 2^{-1}(\eta + 1)y_1 \mid y_1) \geq 1 - c$$

on a aussi

$$P(y_3 > 2^{-1}(\eta + 1)y_2, y_2 > 2^{-1}(\eta + 1)y_1 \mid y_1) \geq (1 - c\beta)(1 - c)$$

où $\beta = 2(\eta + 1)^{-1}$. Si on continue de cette manière, on obtient

$$P(y_{l+1} > 2^{-1}(\eta + 1)y_l, l = 1, \dots, t \mid y_1) \geq \prod_{i=1}^t (1 - c\beta^{i-1}) \geq (1 - c)^{1/(1-\beta)}, \quad \forall t \geq 0.$$

Ainsi, pour tout $y_0 \in \mathbb{R}$,

$$P(y_t \longrightarrow \infty \mid y_0) \geq (1 - c)^{1/(1-\beta)} P(y_1 > M \mid y_0) > 0.$$

alors, $\{y_t\}$ n'est pas ergodique pour $\phi_1 > 1$.

Quand $\phi_1 = 1$, pour $y_1 > 0$, on obtient une marche aléatoire jusqu'au premier instant où on rentre dans le deuxième régime. Cependant, $E(\tau_{]-\infty, 0[} \mid y_1) = \infty$ et puisque $P(y_1 > 0 \mid y_0) > 0, \forall y_0 \in \mathbb{R}$, alors le processus $\{y_t\}$ n'est pas ergodique. (Théorème 7 de Tweedie (1974)). ■

Lemme 2.1.3 *Si $\phi_1 < 0$ et $\phi_1\phi_2 > 1$ (Région II), alors le processus $\{y_t\}$ n'est pas ergodique.*

Preuve. Sans restreindre la généralité, on considère seulement le cas $\phi < -1$ et $\phi_1\phi_2 > 1$. La preuve est semblable à celle du lemme 2.1.2, sauf qu'on montre que la chaîne de Markov $\{y_{2t}; t \geq 0\}$ a la même propriété, pour tout $y_0 \in \mathbb{R}$

$$P(y_{2t} \longrightarrow \infty \mid y_0) > 0. \quad (2.6)$$

En particulier, si $1 < \eta < \phi_1\phi_2$, il existe $M > 0$ tel que $y_{t-2} > M$, alors

$$E(y_t | y_{t-2}) > \eta y_{t-2}, \quad t \geq 2. \quad (2.7)$$

En plus, quand $y_{t-2} > 0$, $E(|y_t - E(y_t | y_{t-2})| | y_{t-2}) \leq \xi < \infty$, $t \geq 0$, choisissons M grand tel que (2.7) soit vérifiée et $2\xi[(\eta - 1)M]^{-1} < 1$, et de façon similaire à la preuve du lemme 2.1.2 on montre que (2.6) est vérifiée. Comme $\{y_{2t}\}$ n'est pas ergodique, alors $\{y_t\}$ est aussi non-ergodique. ■

Lemme 2.1.4 *Si $\phi_1 < 0$ et $\phi_1\phi_2 = 1$ (Région III), alors le processus $\{y_t\}$ n'est pas ergodique.*

Preuve. Aussi, sans restreindre la généralité, on prend $\phi_1 \leq 1$, $-1 \leq \phi_2 < 0$ et $\phi_1\phi_2 = 1$. Alors, pour $y_{t-2} < 0$ et $t \geq 2$

$$\begin{aligned} y_t &= \varepsilon_t + (y_{t-2} + \phi_1\varepsilon_{t-1})\mathbf{1}(y_{t-2} + \phi_1\varepsilon_{t-1} < 0) + \phi_1^{-2}(y_{t-2} + \phi_1\varepsilon_{t-1})\mathbf{1}(y_{t-2} + \phi_1\varepsilon_{t-1} \geq 0) \\ &\leq y_{t-2} + \phi_1\varepsilon_{t-1} + \varepsilon_t \quad p.s. \end{aligned} \quad (2.8)$$

Mais, on remarque pour $t = 3, 5, 7, \dots$, et pour $\gamma_t = \phi_1\varepsilon_{t-1} + \varepsilon_t$ que la suite $\{\gamma_{2j+1}, j \geq 1\}$ est *i.i.d* et de moyenne nulle. On définit la marche aléatoire $\{z_t; t \geq 1\}$ par $z_1 = y_1$ et

$$z_t = z_{t-1} + \gamma_{2t-1}, \quad t > 1.$$

On considère les temps d'arrêt

$$T(y) = \inf\{t > 1 : y_{2t+1} \in]0, +\infty[\}$$

et

$$T(z) = \inf\{t > 1 : z_{t+1} \in]0, +\infty[\}.$$

la relation (2.8) implique que $\{T(z) > n\} \subset \{T(y) > n\}$ quand $y_1 < 0$, comme le processus $\{z_t; t \geq 1\}$ est une marche aléatoire et $E(\gamma_1) = 0$,

$$E(T(y) | y_1) \geq E(T(z) | y_1) = \infty$$

quand $y_1 < 0$. Et donc, le processus n'est pas ergodique. ■

Remarques 2.1.1 *En terme de chaîne de Markov, des lemmes 2.1.1-2.1.4 il résulte que la chaîne $\{y_t\}$ est récurrente positive quand (2.4) est vérifiée, et transitoire à l'intérieur des régions I et II, pour ϕ_1 et ϕ_2 dans la région III, ou sur les frontières de la région I, les lemmes 2.1.2 et 2.1.4 montrent que la chaîne n'est pas récurrente positive, on conjecture qu'elle est récurrente nulle.*

Le théorème 2.1.1 implique l'existence d'une mesure de probabilité invariante pour $\{y_t\}$. En plus, le résultat suivant est obtenu.

Théorème 2.1.2 (Petrucelli et Woolford (1984)) *On suppose $E(|\varepsilon_t|^{2+\xi}) < \infty$, $\xi > 0$. Alors si ϕ_1 et ϕ_2 vérifient (2.4), la mesure de probabilité invariante pour la chaîne $\{y_t\}$ a un moment d'ordre deux fini.*

Preuve. Etant donnée la symétrie qui existe dans la région d'ergodicité, sans restreindre la généralité, on considère seulement le cas où $\phi_1 < \phi_2$. Cela implique que $|\phi_2| < 1$, alors, il existe $k_0 \geq 1$ tel que

$$k_0 = \min\{k \geq 1 : |\phi_1 \phi_2^{k-1}| < 1\}.$$

Notons que pour $\phi_1 \leq 1$ et $-1 < \phi_2 \leq 0$, $k_0 = 2$ par hypothèse, ainsi $k > 2$ seulement si $\phi_1 \leq -1$ et $0 < \phi_2 < 1$. Soit $k^* = k_0 - 1$, $\alpha = \max(|\phi_1|, |\phi_2|)$ et $\eta = 2^{k^*} - 1$. Alors

$$|y_t| < \alpha |\phi_2|^{k^*} |y_{t-k_0}| + |y_{t-k_0}| \sum_{j=1}^{\eta} \delta_j \mathbf{1}(a_j > \beta_j | y_{t-k_0}|) + \sum_{j=1}^{k_0} \gamma_j |\varepsilon_{t-k_0+j}|, \quad (\text{A.1})$$

où $\{\delta_j; 1 \leq j \leq \eta\}$, $\{\beta_j; 1 \leq j \leq \eta\}$ et $\{\gamma_j; 1 \leq j \leq k_0\}$ sont des suites de constantes positives dépendant seulement de ϕ_1 , ϕ_2 et k_0 , pour $1 \leq j \leq \eta$, a_j est une combinaison linéaire de $\{\varepsilon_{t-k_0+l}; 1 \leq l \leq k_0\}$ indépendante de y_{t-k_0} , de moyenne nulle et de moment absolu d'ordre $2 + \xi$ fini. Notons aussi que, par définition, $0 \leq \alpha |\phi_2|^{k^*} < 1$. De (A.1), on a

$$E(|y_t| | y_{t-k_0}) \leq \alpha |\phi_2|^{k^*} |y_{t-k_0}| + M_1, \quad (\text{A.2})$$

où, par l'inégalité de Markov

$$E(|y_{t-k_0}| \mathbf{1}(a_j > \beta_j | y_{t-k_0}) | y_{t-k_0}) \leq \beta_j^{-1} E(|a_j|), \quad 1 \leq j \leq \eta,$$

et

$$E(|\varepsilon_{t-k_0+j}|) \leq A, \quad 1 \leq j \leq k_0$$

On peut prendre

$$M_1 \geq \sum_{j=1}^{k_0} \gamma_j A + \sum_{j=1}^{\eta} \delta_j \beta_j^{-1} E(|a_j|).$$

Cependant, (A. 2) implique que, si $R_1 = \max_{0 \leq l \leq k_0} E(|y_l| | y_0)$, alors

$$E(|y_t| | y_0) \leq [\alpha |\phi_2|^{k^*}]^\mu R_1 + M_1 \sum_{m=0}^{\mu-1} [\alpha |\phi_2|^{k^*}]^m, \quad (\text{A.3})$$

où $\mu = [t/k_0]$, $[\cdot]$ désigne la partie entière. alors, indépendamment de t , il existe une constante $B_1 > 0$ telle que

$$E(|y_t| | y_0) \leq B_1 < \infty, \quad \forall t.$$

Pour n'importe quel entier $\nu \geq 2$, on utilise (A. 1) pour obtenir

$$|y_t| \leq [\alpha |\phi_2|^{k^*}]^\nu |y_{t-k_0}|^\nu + |y_{t-k_0}|^\nu \sum_{j=1}^{\eta} \delta_j^\nu \mathbf{1}(a_j > \beta_j | y_{t-k_0}) + f_1(|y_{t-k_0}|, |\varepsilon_{t-k_0+1}|, \dots, |\varepsilon_t|), \quad (\text{A.4})$$

où f_1 est un un polynôme dont le degré de $|y_{t-k_0}|$ est inférieur ou égal à $\nu - 1$, et on remplace les fonctions indicatrice par 1. Utilisons (A. 4), avec $\nu = 2$, on obtient

$$E(|y_t|^2 | y_{t-k_0}) \leq [\alpha |\phi_2|^{k^*}]^2 |y_{t-k_0}|^2 + f_2(|y_{t-k_0}|) + M_2, \quad (\text{A.5})$$

où f_2 est un polunôme de degré 1, et

$$E(|y_{t-k_0}|^2 \mathbf{1}(a_j > \beta_j | y_{t-k_0}) | y_{t-k_0}) \leq \beta_j^{-2} E(a_j^2), \quad 1 \leq j \leq \eta,$$

On peut choisir

$$M_1 \geq \sum_{j=1}^{\eta} \delta_j^2 \beta_j^{-2} E(a_j^2).$$

De plus,

$$E(f_2(|y_{t-k_0}|) \mid y_0) \leq f_2(B_1) \quad \forall t.$$

Soit $M = M_2 + f_2(B_1)$ et $R_2 = \max_{0 \leq l \leq k_0} E(|y_t|^2 \mid y_0)$,

$$E(|y_t|^2 \mid y_0) \leq [\alpha |\phi_2|^{k^*}]^{2\mu} R_2 + M \sum_{m=0}^{\mu-1} [\alpha |\phi_2|^{k^*}]^{2m}.$$

Comme pour (A.3), indépendamment de t , on peut choisir $B_2 > 0$ tel que

$$E(|y_t| \mid y_0) \leq B_1 < \infty, \quad \forall t. \quad (\text{A.6})$$

Maintenant, choisissons un entier l tel que $l^{-1} < \xi < 1$. Soit $\nu = 2l + 1$ tel que $\nu/l = 2 + l^{-1} < 2 + \xi$. Utilisons (A.4) et le fait que, pour $0 < \alpha < 1$, $|\sum_{i=1}^n c_i|^\alpha \leq \sum_{i=1}^n |c_i|^\alpha$, on obtient

$$|y_t|^\rho \leq [\alpha |\phi_2|^{k^*}]^\rho |y_{t-k_0}|^\rho + |y_{t-k_0}|^\rho \sum_{j=1}^{\eta} \delta_j^\rho \mathbf{1}(a_j > \beta_j \mid y_{t-k_0}) + f_1^{(l)}(|y_{t-k_0}|, |\varepsilon_{t-k_0+1}|, \dots, |\varepsilon_t|),$$

où $\rho = \nu/l$ et $f_1^{(l)}(\cdot)$ est un polynôme en puissance fractionnaire $\{m/l; m = 1, \dots, 2l\}$ pour $|y_{t-k_0}|$ et $\{m/l; m = 1, \dots, 2l + 1\}$ pour $|\varepsilon_t|$. Par conséquent,

$$E(|y_t|^\rho \mid y_{t-k_0}) \leq [\alpha |\phi_2|^{k^*}]^\rho |y_{t-k_0}|^\rho + M_2^{(l)} + f_2^{(l)}(|y_{t-k_0}|)$$

avec $f_2^{(l)}(\cdot)$ est un polynôme en puissance fractionnaire $\{m/l; m = 1, \dots, 2l\}$ pour $|y_{t-k_0}|$, et comme

$$E(|y_{t-k_0}|^\rho \mathbf{1}(a_j > \beta_j \mid y_{t-k_0}) \mid y_{t-k_0}) \leq \beta_j^{-\rho} E(|a_j|^\rho), \quad 1 \leq j \leq \eta,$$

on peut prendre

$$M_2^{(l)} \geq \sum_{j=1}^{\eta} \delta_j^\rho \beta_j^{-\rho} E(|a_j|^\rho).$$

Et comme pour (A.6), on obtient $E(f_2^{(l)}(|y_{t-k_0}|) | y_0) \leq B_2^{(l)} < \infty$, $0 < B_2^{(l)} < \infty$, $\forall t$. Ainsi, avec $M^{(l)} = M_2^{(l)} + B_2^{(l)}$ et $R_\nu^{(l)} = \max_{0 \leq m \leq k_0} E(|y_m|^\rho | y_0)$, on peut majorer $E(|y_t|^\rho | y_0)$ indépendamment de t comme en (A.6).

Utilisons Feller (1971). p. 251-252; par suite de l'ergodicité de $\{y_t\}$, on peut conclure que

$$E(|y_t|^2 | y_0) \longrightarrow E(|y|^2) < \infty$$

où y est une variable aléatoire dont la densité est la densité de probabilité invariante de $\{y_t\}$. ■

Remarques 2.1.2 *Utilisons une méthode similaire dans la preuve du théorème 2.1.2, on peut montrer que si $E(|\varepsilon_t|^{a+\xi}) < \infty$, $\xi > 0$ et $0 \leq a \leq \infty$, alors la mesure de probabilité invariante pour la chaîne $\{y_t\}$ a un moment d'ordre a fini.*

2.1.3 Modèle TAR(1) général

On définit un modèle TAR(1) général à l régimes que l'on note TAR($l; 1, \dots, 1$) par

$$y_t = \phi_0^{(k)} + \phi_1^{(k)} y_{t-1} + \varepsilon_t^{(k)} \quad \text{si } y_{t-1} \in \mathbb{R}_k, \quad (2.9)$$

avec $\mathbb{R}_k =]r_{k-1}, r_k]$, $1 \leq k \leq l$, tel que $-\infty = r_0 < r_1 < \dots < r_l = +\infty$, de manière équivalente, nous pouvons écrire (2.9) comme

$$y_t = \sum_{k=1}^l \mathbf{1}(y_{t-1} \in \mathbb{R}_k) \{ \phi_0^{(k)} + \phi_1^{(k)} y_{t-1} + \varepsilon_t^{(k)} \}, \quad (2.10)$$

avec $\{\phi_i^{(k)}; i = 0, 1; 1 \leq k \leq l\}$ des constantes réelles et pour chaque k , $\{\varepsilon_t^{(k)}; t \geq 1\}$ est une suite de variables aléatoires *i.i.d* centré et de densité $f_k(\cdot)$ strictement positive sur \mathbb{R} , on suppose aussi que $\{\varepsilon_t^k\}$ et $\{\varepsilon_t^j\}$ sont indépendants pour $k \neq j$.

2.1.4 Ergodicité

Notons que, comme pour un modèle TAR(1) simple, $\{y_t; t \geq 1\}$ défini par (2.10), est une chaîne de Markov à espace d'état $(\mathbb{R}, \mathcal{B})$, \mathcal{B} est la σ -algèbre de Borel sur

\mathbb{R} . La densité de transition est donnée par

$$p(x, z) = \sum_{k=1}^l \mathbf{1}(y_{t-1} \in \mathbb{R}_k) f_k(z - \phi_0^{(k)} - \phi_1^{(k)} x). \quad (2.11)$$

Utilisons la définition dans Orey (1971), on note aussi que $\{y_t; t \geq 1\}$ est μ -irréductible et apériodique quand μ est la mesure de Lebesgue sur \mathbb{R} . Cependant, contrairement au cas d'un modèle TAR(1) simple, la loi de transition $\{P(x, \cdot)\}$ correspondant à (2.11), n'est pas nécessairement fortement continue. Alors, nous devons donner le lemme suivant pour permettre ensuite de prouver l'ergodicité.

Lemme 2.1.5 (Tweedie (1975)) *Soit $\{P(x, \cdot)\}$ la loi de transition correspondante à (2.11). Alors, si \mathcal{K} est l'ensemble des compacts dans \mathcal{B} de mesure de Lebesgue positive, alors $0 < \pi(K) < \infty$, $\forall K \in \mathcal{K}$, où $\pi(\cdot)$ est une mesure sous invariante pour $\{y_t\}$.*

Preuve. Soit D l'ensemble des points de discontinuité pour $\{P(x, \cdot)\}$. Alors, par construction, D est fini. Grâce à l'hypothèse d'irréductibilité, il existe une mesure $\pi(\cdot)$ telle que

$$\pi(A) \geq \int_{\mathbb{R}} \pi(dx) P(x, A), \quad A \in \mathcal{B}.$$

Par itération de l'équation ci-dessus, on obtient

$$\beta(1 - \beta)^{-1} \pi(A) \geq \int_{\mathbb{R}} \pi(dx) G_{\beta}(x, A), \quad (2.12)$$

avec $0 < \beta < 1$ et $G_{\beta}(x, A) = \sum_{n=1}^{\infty} \beta^n P^n(x, A)$. Il est facile de voir que $G_{\beta}(x, \cdot)$ est continue pour $x \notin D$ et

$$\lim_{x \uparrow d} G_{\beta}(x, A) > 0, \quad \lim_{x \downarrow d} G_{\beta}(x, A) > 0$$

$\forall d \in D, \mu(A) > 0$. Alors, on a pour tout $K \in \mathcal{K}$,

$$\inf_{x \in K} G_{\beta}(x, A) > 0, \quad \mu(A) > 0.$$

Utilisons (2.12) avec $A \in \mathcal{B}$ tel que $0 < \pi(A) < \infty$, alors

$$\pi(K) \leq \beta(1 - \beta)^{-1} \pi(A) \left[\inf_{x \in K} G_\beta(x, A) \right]^{-1} < \infty.$$

Comme K est de mesure de Lebesgue positive, on a $\pi(K) > 0$. ■

Ce lemme est une extension du lemme 4.1 dans Tweedie (1975) et il est vrai pour des chaînes de Markov plus générales.

Maintenant, on donne les conditions nécessaires et suffisantes sur les paramètres $\{\phi_i^{(k)}; i = 0, 1; 1 \leq k \leq l\}$ pour que le processus $\{y_t\}$ soit ergodique.

Théorème 2.1.3 (Chan et Tong (1985)) *Le processus $\{y_t\}$ défini par (2.11), est ergodique si et seulement si une des conditions suivantes soit vérifiée :*

$$\phi_1^{(1)} < 1, \quad \phi_1^{(l)} < 1, \quad \phi_1^{(1)} \phi_1^{(l)} < 1; \quad (2.13)$$

$$\phi_1^{(1)} = 1, \quad \phi_1^{(l)} < 1, \quad \phi_0^{(1)} > 0; \quad (2.14)$$

$$\phi_1^{(1)} < 1, \quad \phi_1^{(l)} = 1, \quad \phi_0^{(l)} < 0; \quad (2.15)$$

$$\phi_1^{(1)} = 1, \quad \phi_1^{(l)} = 1, \quad \phi_0^{(l)} < 0 < \phi_0^{(1)}; \quad (2.16)$$

$$\phi_1^{(1)} \phi_1^{(l)} = 1, \quad \phi_1^{(1)} < 0, \quad \phi_0^{(l)} + \phi_1^{(l)} \phi_0^{(1)} > 0. \quad (2.17)$$

La preuve du théorème 2.1.3 est dévisée en deux lemmes, le premier lemme vérifie la suffisance, le deuxième prouve la nécessité.

Lemme 2.1.6 (Chan et Tong (1985)) *Si le processus $\{y_t\}$ défini par (2.11) vérifie une des conditions (2.13)-(2.17), alors $\{y_t\}$ est ergodique.*

Preuve. De façon similaire au lemme 2.1.1, nous cherchons un ensemble compact $K \in \mathcal{B}$, de mesure de Lebesgue positive, et une fonction mesurable et positive g sur \mathbb{R} tels que

$$\int_{\mathbb{R}} p(x, z) g(z) dz \leq g(x) - 1, \quad x \notin K \quad (2.18)$$

$$\int_{\mathbb{R}} p(x, z) g(z) dz \leq \lambda(x) \leq R < \infty, \quad x \in K, \text{ pour } R > 0. \quad (2.19)$$

On prouve l'ergodicité de $\{y_t\}$ pour chacune de (2.13)-(2.17) séparément, en indiquant pour chaque cas la fonction g et le compact K pour que la condition

ci-dessus soit vérifiée.

(2.13) : Comme pour le cas de deux régimes, il existe deux constantes positives a et b telles que $1 > \phi_1^{(1)} > -(ba^{-1})$ et $1 > \phi_1^{(l)} > -(ab^{-1})$, et une fonction g définie par

$$g(x) = \begin{cases} ax, & x > 0 \\ b|x|, & x \leq 0. \end{cases}$$

Alors il existe $M > 0$ tel que (2.18) et (2.19) soient vérifiées pour $K = [-M, M]$.

(2.14) : Dans ce cas là, on prend

$$g(x) = \begin{cases} c_1 x, & x > 0 \\ -2[\phi_0^{(1)}]^{-1} x, & x \leq 0 \end{cases}$$

où $c_1 > 2|\phi_1^{(l)}|[\phi_0^{(1)}]^{-1}$. Aussi, il existe $M > 0$ tel que (2.18) et (2.19) soient vérifiées pour $K = [-M, M]$.

(2.15) : Par symétrie, on peut prendre

$$g(x) = \begin{cases} -2[\phi_0^{(l)}]^{-1} x, & x > 0 \\ c_2 x, & x \leq 0 \end{cases}$$

avec $c_2 > -2|\phi_1^{(1)}|[\phi_0^{(l)}]^{-1}$.

(2.16) : Le résultat suit de (2.14) et (2.15) avec

$$g(x) = \begin{cases} 2|\phi_0^{(l)}|^{-1} x, & x > 0 \\ -2[\phi_0^{(1)}]^{-1} x, & x \leq 0 \end{cases}$$

(2.17) : Dans ce cas, on considère la chaîne de Markov $\{y_{2t}\}$ avec la loi de transition $\{P^2(x, \cdot)\}$. Prenant

$$g(x) = \begin{cases} ax, & x > 0 \\ -bx, & x \leq 0, \end{cases}$$

avec a et b sont des constantes positives. On obtient pour $x \in \mathbb{R}_j$

$$\begin{aligned} I(x) &= \int_{\mathbb{R}} P^2(x, dz)g(z) \\ &= \sum_{k=1}^l \left\{ a \int_{\beta(k,x)}^{\infty} (u - \beta(k,x)) \int_{R(k,j)} f_k(u - \phi_1^{(k)}a) f_j(a) da du \right. \\ &\quad \left. - b \int_{-\infty}^{\beta(k,x)} (u - \beta(k,x)) \int_{R(k,j)} f_k(u - \phi_1^{(k)}a) f_j(a) da du \right\} \end{aligned}$$

avec

$$\beta(k, x) = -\phi_0^{(k)} - \phi_1^{(k)}\phi_0^{(j)} - \phi_1^{(k)}\phi_1^{(j)}x$$

et

$$R(k, j) = \left\{ z : z + \phi_0^{(j)} + \phi_1^{(j)}x \in \mathbb{R}_k \right\}.$$

Il est clair que pour $x \in \mathbb{R}_j$, $0 < B_j < \infty$, $j = 2, \dots, l-1$, $I(x) \leq B_j$. En plus, il est possible de montrer qu'il existe $M > 0$ tel que

$$\begin{aligned} I(x) &\leq -bx - \frac{b}{2}(\phi_0^{(l)} + \phi_1^{(l)}\phi_0^{(1)}), & x < -M \\ I(x) &\leq ax + \frac{a}{2}(\phi_0^{(1)} + \phi_1^{(1)}\phi_0^{(1)}), & x > -M. \end{aligned}$$

Ainsi, on peut définir a et b tels que (2.18) et (2.19) soient vérifiées pour $K = [-M, M]$. D'où, on peut conclure que $\{y_{2t}\}$ est ergodique, et par suite de l'irréductibilité et de l'apériodicité de $\{y_t\}$ on obtient que $\{y_t\}$ est aussi ergodique. ■

Lemme 2.1.7 (Chan et Tong (1985)) *Si $\{y_t\}$ donné par (2.11), ne satisfait pas une des conditions (2.13)-(2.17), alors le processus $\{y_t\}$ n'est pas ergodique.*

Preuve. On peut distinguer quatre cas :

- (i) $\phi_1^{(1)} > 1$ ou $\phi_1^{(l)} > 1$.
- (ii) $(\phi_1^{(1)} = 1 \text{ et } \phi_0^{(1)} \leq 0)$ ou $(\phi_1^{(l)} = 1 \text{ et } \phi_0^{(l)} \geq 0)$.
- (iii) $\phi_1^{(1)} < 0$, $\phi_1^{(1)}\phi_1^{(l)} > 1$.
- (iv) $\phi_1^{(1)} < 0$, $\phi_1^{(1)}\phi_1^{(l)} = 1$ et $\phi_0^{(1)}\phi_1^{(l)} + \phi_0^{(l)} \leq 0$.

Par une légère modification de la preuve du théorème 2.1.1, les cas (i) – (iii) peuvent être démontrés de façon similaire, pour (iv), on applique le théorème 9.1 (ii) de Tweedie (1976) pour montrer que $\{y_t\}$ n'est pas ergodique. Ainsi, il suffit de trouver une fonction g mesurable et positive, et un compact $K = [-k_1, k_2]$, $k_1, k_2 > 0$, et une constante $B > 0$ tels que

$$\begin{aligned} \int_{\mathbb{R}} p(x, z)g(z)dz &\geq g(x), & x \notin K \\ \int_{\mathbb{R}} p(x, z) |g(z) - g(x)| &\leq B, & x \in B \\ g(x) &> \sup_{z \in K} g(z), & x \notin K. \end{aligned}$$

Ici, on prend

$$g(x) = g_{\alpha\beta}(x) + I_{k,M}(x)$$

avec

$$g_{\alpha\beta}(x) = \begin{cases} ax + \alpha & \text{si } x > 0 \\ -bx + \beta & \text{si } x \leq 0 \end{cases}$$

et

$$I_{k,M}(x) = \begin{cases} k & \text{si } |x| \leq M \\ 0 & \text{sinon} \end{cases}$$

où α, β, k et M sont des constantes strictement positives choisies telles que

$$a\phi_0^{(1)} \geq \beta - \alpha \geq b\phi_0^{(l)}$$

$$M \geq \max(|\phi_0^{(1)}|, |\phi_0^{(l)}|)$$

$$k \geq (a + b) \max(|\phi_0^{(1)}|, |\phi_0^{(l)}|)$$

■

Remarques 2.1.3 Les conditions d'ergodicité du théorème 2.1.3 dépendent seulement des paramètres $\phi_0^{(1)}, \phi_0^{(l)}, \phi_1^{(1)}$ et $\phi_1^{(l)}$, i.e. seulement du comportement du processus dans les régions des régimes extrêmes \mathbb{R}_1 et \mathbb{R}_l .

Théorème 2.1.4 Supposons que pour $r \in \mathbb{N}^*$, $1 \leq k \leq l$, $E(|\varepsilon_t^{(k)}|^r) < \infty$. Alors, si

$\phi_1^{(1)} < 1$, $\phi_1^{(l)} < 1$ et $\phi_1^{(1)}\phi_1^{(l)} < 1$, la mesure de probabilité invariante pour la chaîne $\{y_t\}$ a un moment d'ordre r fini, et la chaîne $\{y_t\}$ est géométriquement ergodique.

Preuve. Choisissons $a, b > 0$ telles que $1 > \phi_1^{(1)} > -(ba^{-1})$ et $1 > \phi_1^{(l)} > -(ab^{-1})$, soit $c > 0$. On définit

$$g(x) = \begin{cases} a^r x^r + c & x > 0 \\ b^r |x|^r + c & x \leq 0. \end{cases}$$

Il est possible de montrer que pour $|x|$ grand, on a

$$\int_{\mathbb{R}} p(x, z)g(z)dz \leq (1 - \varepsilon)g(x), \quad \varepsilon > 0$$

et le résultat suit de Tweedie (1983). ■

2.2 Modèle moyenne mobile à seuil

2.2.1 Introduction

On considère un modèle moyenne mobile à seuil du 1^{er} ordre (TMA(1)), qui génère la série chronologique $\{y_t : t = 0, \pm 1, \dots\}$ selon l'équation stochastique suivante:

$$y_t = [\phi + \psi \mathbf{1}(y_{t-1} \leq r)]\varepsilon_{t-1} + \varepsilon_t \quad (2.20)$$

où ε_t est une suite de variable aléatoire *i.id* de moyenne nulle et de densité $f(x)$. Ling et Tong (2005) ont proposé un test fondé sur le rapport de vraisemblance pour un modèle moyenne mobile linéaire contre un modèle TMA. Cependant, la structure du modèle TMA n'est pas spécifiée. Brockwell et al (1992) ont étudié le modèle autorégressif moyenne mobile à seuil (TARMA) et ils ont obtenu une solution strictement stationnaire et ergodique. Cependant, leur résultat couvre les cas où la partie moyenne mobile de ce modèle ne subit pas un threshold, donc n'inclut pas le TMA comme cas spécial. Liu et Susko (1992) ont prouvé que pour le modèle TMA, il existe toujours une solution strictement stationnaire sans aucune restriction sur les coefficients. Ling (1999) a obtenu le même résultat pour un TARMA, mais sous des conditions restrictives à cause de la présence d'une partie autorégressive. Ainsi les conditions sous lesquelles la solution du modèle TMA est

unique et ergodique est un problème ouvert. Ling et Tong (2005) ont obtenu une condition suffisante pour l'inversibilité d'un modèle TMA général. Cette condition est loin d'être nécessaire quand on la compare avec celle d'un modèle MA linéaire. Ling et al (2007) donne une condition suffisante pour l'ergodicité d'un modèle TMA(1) et des conditions nécessaires et suffisantes pour l'inversibilité du modèle TMA(1).

2.2.2 Ergodicité des modèles TMA(1)

Une méthode standard dans l'étude de l'ergodicité d'un modèle de série chronologique non linéaire est d'écrire le modèle sous la forme d'une chaîne de Markov et ensuite de vérifier le critère donné dans Tweedie (1983) (voir aussi Tjøstheim (1990) et Tweedie (2001)). cette méthode comprend deux parties, la 1^{ère} est de vérifier la ψ -irréductibilité de la chaîne de Markov, la 2^{ème} partie est de vérifier le drift criterion ou critère de dérive.

Le défi reste la vérification de la propriété de l'irréductibilité. Sous certaines conditions, Feigin et Tweedie (1985) ont montré que la chaîne de Markov obtenue à partir d'un RCA est irréductible. Chan et Tong (1985) ont montré l'irréductibilité d'une chaîne de Markov en termes d'une mesure adéquate. Quand on met le modèle TMA(1) sous forme d'une chaîne de Markov (voir Ling (1999)), il n'est pas difficile de voir que ni il s'ajuste au cadre de Chan et Tong (1985) ni satisfait à la condition donnée dans Feigin et Tweedie (1985). Ainsi il semble décidément difficile d'établir l'irréductibilité. Cependant, comme nous le verrons, il n'est pas nécessaire de vérifier l'irréductibilité.

Soit $\{(y_t, \varepsilon_t)\}$, $t = 0, \pm 1, \dots$, définis sur l'espace de probabilité (Ω, F, P) . Ling et al (2007) ont développé, une méthode alternative pour étudier l'ergodicité. D'abord, on définit une suite aléatoire

$$S_n(t) = \begin{cases} \varepsilon_t, & \text{si } n = 0, \\ [\phi + \psi \mathbf{1}(S_{n-1}(t-1) \leq r)]\varepsilon_{t-1} + \varepsilon_t & \text{si } n \geq 1. \end{cases}$$

Lemme 2.2.1 (Ling et al (2007)) *Si $|\psi| \sup_x |xf(x)| < 1$ et $E|\varepsilon_t| < \infty$, alors $\lim_{n \rightarrow \infty} S_n(t)$ existe p.s. $\forall t \in \mathbb{Z}$.*

Preuve. Soit $u_n(t) = [\phi + \psi \mathbf{1}(S_n(t) \leq r)] \varepsilon_t$ et $b = |\psi| \sup_x |xf(x)|$. On a

$$\begin{aligned}
A_n(t) &\equiv |\psi| E [|\mathbf{1}(S_n(t) \leq r) - \mathbf{1}(S_{n-1}(t) \leq r)| |\varepsilon_t|] \\
&= |\psi| E [|\mathbf{1}(u_{n-1}(t-1) + \varepsilon_t \leq r) - \mathbf{1}(u_{n-2}(t-1) + \varepsilon_t \leq r)| |\varepsilon_t|] \\
&= |\psi| E [|\mathbf{1}(\varepsilon_t \leq r - u_{n-1}(t-1)) - \mathbf{1}(\varepsilon_t \leq r - u_{n-2}(t-1))| |\varepsilon_t|] \\
&= |\psi| E \left| \int_{r-u_{n-2}(t-1)}^{r-u_{n-1}(t-1)} |x| f(x) dx \right| \\
&\leq |\psi| \sup_x |xf(x)| E |u_{n-1}(t-1) - u_{n-2}(t-1)| = bA_{n-1}(t-1) \quad (2.21)
\end{aligned}$$

i.e., $A_n(t) \leq bA_{n-1}(t-1) \leq b^2A_{n-2}(t-2) \leq \dots \leq b^{n-1}A_1(t-n+1)$

on sait que

$$\begin{aligned}
A_1(t-n+1) &= |\psi| E [|\mathbf{1}(S_1(t-n+1) \leq r) - \mathbf{1}(S_0(t-n+1) \leq r)| |\varepsilon_{t-n+1}|] \\
&= |\psi| E [|\mathbf{1}(\varepsilon_{t-n+1} \leq r - u_0(t-n)) - \mathbf{1}(\varepsilon_{t-n+1} \leq r)| |\varepsilon_{t-n+1}|] \\
&\leq |\psi| E |\varepsilon_t| \equiv c
\end{aligned}$$

d'où,

$$A_n(t) \leq bA_{n-1}(t-1) \leq b^2A_{n-2}(t-2) \leq \dots \leq b^{n-1}A_1(t-n+1) \leq cb^{n-1} \quad (2.22)$$

Ainsi, on a $E|S_n(t+1) - S_{n-1}(t+1)| = A_n(t) \leq cb^{n-1}$. Utilisant ce résultat et le critère de Cauchy (au sens de L1) nous pouvons montrer que $\lim_{n \rightarrow \infty} S_n(t)$ existe p.s, $\forall t \in \mathbb{Z}$.

Pour cela, nous devons démontrer que $E|S_m(t+1) - S_n(t+1)| \xrightarrow{n \rightarrow \infty} 0$, $n < m$.

$$\begin{aligned}
E|S_m(t+1) - S_n(t+1)| &= E|S_m(t+1) - S_{m-1}(t+1) + S_{m-1}(t+1) - S_{m-2}(t+1) + S_{m-2}(t+1) + \dots \\
&\quad + S_{n+1}(t+1) - S_n(t+1)| \\
&\leq \underbrace{E|S_m(t+1) - S_n(t+1)|}_{A_m(t)} + \underbrace{E|S_m(t+1) - S_n(t+1)|}_{A_{m-1}(t)} + \dots \\
&\quad + \underbrace{E|S_m(t+1) - S_n(t+1)|}_{A_{n+1}(t)} \\
&\leq cb^{m-1} + cb^{m-1} + \dots + cb^n = c \sum_{i=n}^{m-1} b^i = \frac{b^n - b^m}{1-b} \xrightarrow{n \rightarrow \infty} 0, \forall t. \blacksquare
\end{aligned}$$

Ainsi, on peut énoncer le premier résultat comme suit.

Théorème 2.2.1 (Ling et al (2007)) *Sous les hypothèses du Lemme 2.2.1, $\{X_t\}$ défini par $X_t = \lim_{n \rightarrow \infty} S_n(t)$ est l'unique solution strictement stationnaire et ergodique du modèle (2.20).*

Preuve. Puisque $\{\varepsilon_t\}$ est strictement stationnaire et ergodique, sans restreindre la généralité, on suppose qu'on a une transformation T préservant la mesure et ergodique de Ω vers Ω telle que $\varepsilon_t(\omega) = \varepsilon_0(T^t\omega)$. Ainsi $X_t(\omega) = \lim_{n \rightarrow \infty} (S_n(t))(\omega) = \lim_{n \rightarrow \infty} S_n(t, \omega) = \lim_{n \rightarrow \infty} S_n(0, T^t\omega) = X_0(T^t\omega)$, et delà $\{X_t\}$ est strictement stationnaire et ergodique.

Notons que $1(x_n \leq r) \xrightarrow{n \rightarrow \infty} 1(x \leq r)$ si $x \neq r$ pour toute suite $\{x_n\}$ telle que $x_n \xrightarrow{n \rightarrow \infty} x$. Ainsi $1(S_n(t) \leq r) \xrightarrow{n \rightarrow \infty} 1(X_t \leq r)$ lorsque $X_t \neq r$, comme ε_t a une densité, $P(X_t = r) = 0$. Donc si $X_t \neq r \quad \forall t \in \mathbb{Z}$, on a $1(S_n(t) \leq r) \xrightarrow{n \rightarrow \infty} 1(X_t \leq r)$, d'où $\{X_t\}$ est une solution du modèle (2.20).

Pour prouver l'unicité, on suppose qu'il existe une autre solution $\{X'_t\}$, d'une manière similaire à (2.21), on note $u(t) = [\phi + \psi 1(X_t \leq r)]\varepsilon_t$, $u'(t) = [\phi + \psi 1(X'_t \leq r)]\varepsilon_t$ et $b = |\psi| \sup_x |xf(x)|$. On a

$$\begin{aligned} E \left| X_t - X'_t \right| &= |\psi| E \left[\left| \mathbf{1}(X_{t-1} \leq r) - \mathbf{1}(X'_{t-1} \leq r) \right| \varepsilon_{t-1} \right] \\ &= |\psi| E \left[\left| \mathbf{1}(\varepsilon_{t-1} \leq r - u(t-2)) - \mathbf{1}(\varepsilon_{t-1} \leq r - u'(t-2)) \right| \varepsilon_{t-1} \right] \\ &= |\psi| E \left| \int_{r-u'(t-2)}^{r-u(t-2)} |x| f(x) dx \right| \\ &\leq |\psi| \sup_x |xf(x)| E \left| u(t-2) - u'(t-2) \right| \\ &\leq b E \left| X_{t-1} - X'_{t-1} \right| \leq \dots \leq b^n E \left| X_{t-n} - X'_{t-n} \right| \leq cb^n \end{aligned}$$

avec $c = |\psi| E |\varepsilon_t|$. Ainsi $E \left| X_t - X'_t \right| = \lim_{n \rightarrow \infty} b^n c = 0$, alors $X_t = X'_t$ p.s. ■

Il est intéressant de voir que l'ergodicité du modèle (2.20) ne dépend que de ψ , et non de ϕ .

Exemples

- Quand $\varepsilon_t \sim N(0, 1)$, on a $\sup_x |xf(x)| = (\sqrt{2\pi}e)^{-1}$.
- Quand $\varepsilon_t \sim t_n$, on a $\sup_x |xf(x)| = (1 + n^{-1})^{-(n+1)/2} \Gamma\left(\frac{n+1}{2}\right) / [\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)]$.
- Quand ε_t suit une loi double exponentielle de paramètre $\lambda > 0$, on a $\sup_x |xf(x)| =$

$(2e)^{-1}$.

Dans ces cas de figure, ψ appartient à une grande classe de paramètres, et lorsque $\psi = 0$, le modèle (2.20) se réduit à un modèle MA(1) linéaire.

2.2.3 Inversibilité des modèles TMA(1)

Il est bien connu que la condition de stationnarité d'un modèle AR(p) et la condition d'inversibilité d'un modèle MA(p) ont les mêmes restrictions sur les paramètres correspondants.

Sous les mêmes types de restrictions sur les paramètres que pour la stationnarité d'un modèle TAR (p) (Chan et Tong (1985)), Ling et Tong (2005) ont montré que le modèle TMA(p) est inversible. Pour l'ergodicité d'un modèle TAR(1), Petrucelli et Woolford (1984) et Chan et al. (1985) ont établi des résultats plus précis dans ce contexte.

Il est donc intéressant de voir si des conditions similaires suffisent pour l'inversibilité d'un modèle TMA(1).

Soit $A_{t-1} = -[\phi + \psi 1(y_{t-1} \leq r)]$, on a

$$\begin{aligned}
 \varepsilon_t &= y_t + A_{t-1}\varepsilon_{t-1} \\
 &= y_t + A_{t-1}y_{t-1} + A_{t-1}A_{t-2}\varepsilon_{t-2} \\
 &= y_t + A_{t-1}y_{t-1} + A_{t-1}A_{t-2}y_{t-2} + A_{t-1}A_{t-2}A_{t-3}\varepsilon_{t-3} \\
 &\quad \vdots \\
 &= y_t + \sum_{j=1}^J \left(\prod_{i=1}^j A_{t-i} \right) y_{t-j} + \left(\prod_{i=1}^{J+1} A_{t-i} \right) \varepsilon_{t-J+1}
 \end{aligned} \tag{2.23}$$

Comme $\{y_t, t = 0, \pm 1, \dots\}$ est strictement stationnaire et ergodique, alors, en utilisant le critère \sqrt{J} ème de Cauchy, il vient

$$\begin{aligned}
 \lim_{J \rightarrow \infty} \log \left(\prod_{i=1}^J |A_{t-i}| |y_{t-J}| \right)^{1/J} &= \lim_{J \rightarrow \infty} \frac{1}{J} \left[\sum_{i=1}^J \log |A_{t-i}| + \log |y_{t-J}| \right] = E(\log |A_t|) \text{ p.s} \\
 \text{si } E |\log |y_t|| < \infty, \text{ car, si } E |\log |y_t|| < \infty, \text{ alors } |\log |y_t|| \text{ est fini p.s, donc } \lim_{J \rightarrow \infty} \frac{\log |y_t|}{J} &= 0 \\
 \text{p.s, et d'après le théorème ergodique on a } \lim_{J \rightarrow \infty} \frac{1}{J} \sum_{i=1}^J \log |A_{t-i}| &= E(\log |A_t|) \text{ p.s.}
 \end{aligned}$$

Lorsque $E(\log |A_t|) < 0$,

$$\lim_{J \rightarrow \infty} \left(\prod_{i=1}^J |A_{t-i}| |y_{t-J}| \right)^{1/J} = e^{E(\log |A_t|)} \quad p.s.$$

D'où, si pour tout $\omega \in \Omega$ vérifiant l'inégalité précédente, il existe $N(\omega)$ t.q $\forall J \geq N(\omega)$, $\left(\prod_{i=1}^J |A_{t-i}| |y_{t-J}| \right)^{1/J}(\omega) \leq \rho(\omega) < 1$. Ainsi $\left[\sum_{j=1}^J \left(\prod_{i=1}^j A_{t-i} \right) y_{t-j} \right](\omega)$ converge, alors

$$\sum_{j=1}^J \left(\prod_{i=1}^j A_{t-i} \right) y_{t-j} \longrightarrow \sum_{j=1}^{\infty} \left(\prod_{i=1}^j A_{t-i} \right) y_{t-j} \quad (2.24)$$

De la même manière, on peut montrer que le deuxième terme dans (2.23) converge vers 0 *p.s* si $E(\log |A_t|)$. Dans ce cas là, on a

$$\varepsilon_t = y_t + \sum_{j=1}^{\infty} \left(\prod_{i=1}^j A_{t-i} \right) y_{t-j} \quad (2.25)$$

et on dit que le modèle (2.20) est inversible.

Théorème 2.2.2 (Ling et al (2007)) *Soit $\{y_t\}$ l'unique solution strictement stationnaire et ergodique du modèle (2.20), avec $E \log |y_t| < \infty$, alors le modèle (2.20) est inversible si $|\phi|^{1-F_y(r)} |\phi + \psi|^{F_y(r)} < 1$, où $F_y(r)$ est la fonction de répartition de y_t .*

Preuve. Pour que le modèle (2.20) soit inversible, nous devons vérifier que $E \log |y_t| < \infty$ et que $E(\log |A_t|) < 0$.

$$\begin{aligned} E(\log |A_t|) &= E(\log |\phi + \psi \mathbf{1}(y_{t-1} \leq r)|) \\ &= P(y_{t-1}) \log |\phi| + P(y_{t-1} \leq r) \log |\phi + \psi| \\ &= [1 - F_y(r)] \log |\phi| + F_y(r) \log |\phi + \psi| \\ &= \log |\phi|^{1-F_y(r)} + \log |\phi + \psi|^{F_y(r)} \\ &= \log [|\phi|^{1-F_y(r)} |\phi + \psi|^{F_y(r)}] \end{aligned}$$

et si $E(\log |A_t|) < 0$, alors $\log [|\phi|^{1-F_y(r)} |\phi + \psi|^{F_y(r)}] < 0 \Leftrightarrow |\phi|^{1-F_y(r)} |\phi + \psi|^{F_y(r)} < 1$. ■

2.2.4 Extension à des modèles TMA généraux

On considère le modèle moyenne mobile à k seuils MA $(1, k)$:

$$y_t = \left\{ \phi_0 + \sum_{j=1}^k \psi_j \mathbf{1}(r_{j-1} < y_{t-1} \leq r_j) \right\} \varepsilon_{t-1} + \varepsilon_t \quad (2.26)$$

avec $r_0 = -\infty < r_1 < r_2 < \dots < r_k = +\infty$. C'est un cas spécial du modèle SETARMA introduit par Tong (1983). On étendra les théorèmes 2.2.1 et 2.2.2 au modèle (2.26).

D'abord, on définit

$$u_n(t) = \left\{ \phi_0 + \sum_{j=1}^k \psi_j \mathbf{1}(r_{j-1} < S_n(t) \leq r_j) \right\} \varepsilon_t$$

et

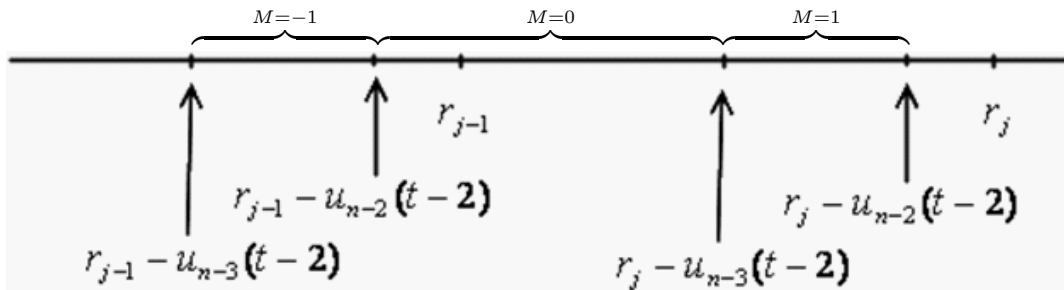
$$S_n(t) = \begin{cases} \varepsilon_t & \text{si } n = 0 \\ u_{n-1}(t-1) + \varepsilon_t & \text{si } n \geq 1 \end{cases}$$

Si pour tout $\omega \in \Omega$ vérifiant $u_{n-3}(t-2) \geq u_{n-2}(t-2)$, on a la décomposition suivante

$$\sum_{j=1}^k \psi_j \{ \mathbf{1}(r_{j-1} < S_{n-1}(t-1) \leq r_j) - \mathbf{1}(r_{j-1} < S_{n-2}(t-2) \leq r_j) \} =$$

$$\sum_{j=1}^k \psi_j \underbrace{ \{ \mathbf{1}(r_{j-1} - u_{n-2}(t-2) < \varepsilon_{t-1} \leq r_j - u_{n-2}(t-2)) - \mathbf{1}(r_{j-1} - u_{n-3}(t-2) < \varepsilon_{t-1} \leq r_j - u_{n-3}(t-2)) \} }_{=M}$$

on peut résumer les valeurs possibles de la quantité M à l'aide du schéma suivant:



Ainsi

$$\sum_{j=1}^k \psi_j \{ \mathbf{1}(r_{j-1} < S_{n-1}(t-1) \leq r_j) - \mathbf{1}(r_{j-1} < S_{n-2}(t-2) \leq r_j) \} =$$

$$\begin{aligned} \sum_{j=1}^k \psi_j \{ & \mathbf{1}(r_j - u_{n-3}(t-2) < \varepsilon_{t-1} \leq r_j - u_{n-2}(t-2)) - \mathbf{1}(r_{j-1} - u_{n-3}(t-2) < \varepsilon_{t-1} \leq r_{j-1} - u_{n-2}(t-2)) \} \\ & = \sum_{j=1}^k (\psi_j + \psi_{j+1}) \mathbf{1}(r_j - u_{n-3}(t-2) < \varepsilon_{t-1} \leq r_j - u_{n-2}(t-2)) \end{aligned}$$

Cela reste vrai aussi pour les $\omega \in \Omega$ vérifiant $u_{n-3}(t-2) \leq u_{n-2}(t-2)$. Alors,

$$\begin{aligned} E |S_n(t) - S_{n-1}(t)| &= E |u_{n-1}(t-1) - u_{n-2}(t-1)| \\ &= E \left| \sum_{j=1}^k (\psi_j + \psi_j) \mathbf{1}(r_j - u_{n-3}(t-2) < \varepsilon_{t-1} \leq r_j - u_{n-2}(t-2)) \right| \\ &\leq \sum_{j=1}^k |\psi_j + \psi_{j+1}| E \left| \int_{r_j - u_{n-3}(t-2)}^{r_j - u_{n-2}(t-2)} xf(x) dx \right| \\ &\leq \underbrace{\sum_{j=1}^k |\psi_j + \psi_{j+1}| \sup_x |xf(x)|}_b E |u_{n-2}(t-2) - u_{n-3}(t-2)| \\ &= bE |S_{n-1}(t-1) - S_{n-2}(t-1)| \end{aligned}$$

on peut montrer comme pour le théorème 2.2.1 que $S_n(t)$ converge *p.s* pour tout t quand $n \rightarrow \infty$ et que $\{X_t\}$ défini par $X_t = \lim_{n \rightarrow \infty} S_n(t)$ est l'unique solution strictement stationnaire et ergodique du modèle (2.26).

Théorème 2.2.3 (Ling et al (2007)) Si $\sum_{j=1}^k |\psi_j + \psi_{j+1}| \sup_x |xf(x)| < 1$, alors $\{X_t\}$ défini par $X_t = \lim_{n \rightarrow \infty} S_n(t)$ est l'unique solution strictement stationnaire et ergodique du modèle (2.26).

Soit $A_{t-1} = - \left[\phi_0 + \sum_{j=1}^k \psi_j \mathbf{1}(r_{j-1} < y_{t-1} \leq r_j) \right]$, le modèle (2.26) est dit inversible si

$$\varepsilon_t = y_t + \sum_{j=1}^{\infty} \left(\prod_{i=1}^j A_{t-i} \right) y_{t-j} \quad p.s$$

Idem

$$\begin{aligned} E(\log |A_t|) &= E \left(\log \left| \phi_0 + \sum_{j=1}^k \psi_j \mathbf{1}(r_{j-1} < y_{t-1} \leq r_j) \right| \right) \\ &= \log |\phi_0 + \psi_1| P(r_0 < y_{t-1} \leq r_1) + \log |\phi_0 + \psi_2| P(r_1 < y_{t-1} \leq r_2) + \dots + \log |\phi_0 + \psi_k| P(r_{k-1} < y_{t-1} \leq r_k) \\ &= \sum_{j=1}^k [F_y(r_j) - F_y(r_{j-1})] \log |\phi_0 + \psi_j| \end{aligned}$$

$$= \log \left[\prod_{j=1}^k |\phi_0 + \psi_j|^{F_y(r_j) - F_y(r_{j-1})} \right]$$

Théorème 2.2.4 *Soit $\{y_t\}$ l'unique solution strictement stationnaire et ergodique du modèle (2.26), avec $E |\log |y_t|| < \infty$, on dit que le modèle (2.26) est inversible si*

$$\prod_{j=1}^k |\phi_0 + \psi_j|^{F_y(r_j) - F_y(r_{j-1})} < 1.$$

Quand $k = 2$, le modèle (2.26) se réduit au modèle (2.20) avec $\phi = \phi_0 + \psi_2$ et $\psi = \psi_1 + \psi_2$, et les résultats dans les théorèmes 2.2.3 et 2.2.4 sont les mêmes que ceux des théorème 2.2.1 et 2.2.2.

Chan *et al.* (1985) ont montré que la stationnarité d'un modèle TAR(1, k) ne dépend que des coefficients des sous modèles dans les deux régimes extrêmes, indépendamment des sous modèles des autres régimes. Par contre, le théorème 2.2.4 montre que les coefficients dans les sous modèles MA(1) autres que ceux des régimes extrêmes jouent aussi un rôle dans l'inversibilité du modèle TMA(1, k).

On considère maintenant le modèle à deux régimes TMA(p, q) :

$$y_t = \sum_{i=1}^p [\phi_j + \psi \mathbf{1}(y_{t-q} \leq r)] \varepsilon_{t-i} + \varepsilon_t \quad (2.27)$$

Quand $q \geq p$, l'inversibilité du modèle (2.27) a été établie par Ling et Tong (2005).

Ici, on étudie seulement la stationnarité et l'ergodicité.

Théorème 2.2.5 (Ling *et al* (2007)) *si $|\psi_p| \sup_x |xf(x)| + \sup_x |f(x)| \times E \left| \sum_{i=1}^{p-1} \psi_i \varepsilon_{t-i} \right| < 1$ quand*

$p = q$ ou $\sup_x |f(x)| \times E \left| \sum_{i=1}^p \psi_i \varepsilon_{t-i} \right| < 1$ quand $p < q$, alors $\{X_t\}$ définie par $X_t = \lim_{n \rightarrow \infty} S_n(t)$ est l'unique solution strictement stationnaire et ergodique du modèle (2.27)

Preuve. On définit

$$S_n(t) = \begin{cases} \varepsilon_t & \text{si } n = 0 \\ \sum_{i=1}^p [\phi_j + \psi \mathbf{1}(S_{n-1}(t-q) \leq r)] \varepsilon_{t-i} + \varepsilon_t & \text{si } n \geq 1 \end{cases}$$

Soit $u_n(t) = \sum_{i=1}^p [\phi_j + \psi \mathbf{1}(S_n(t-q) \leq r)] \varepsilon_{t-i}$, $b = \sup_x |f(x)| \times E \left| \sum_{i=1}^{p-1} \psi_i \varepsilon_{t-i} \right|$

Quand $p = q$, on a

$$\begin{aligned}
 A_n(t) &\equiv E \left[\left| \sum_{i=1}^p \psi_i \varepsilon_{t-i} \right| \left| \mathbf{1}(S_n(t-q) \leq r) - \mathbf{1}(S_{n-1}(t-q) \leq r) \right| \right] \\
 &\leq E \left[\left| \sum_{i=1}^{p-1} \psi_i \varepsilon_{t-i} \right| \left| \mathbf{1}(S_n(t-q) \leq r) - \mathbf{1}(S_{n-1}(t-q) \leq r) \right| \right] + \\
 &\quad E \left[|\psi_p \varepsilon_{t-p}| \left| \mathbf{1}(S_n(t-q) \leq r) - \mathbf{1}(S_{n-1}(t-q) \leq r) \right| \right] \\
 &= E \left[\left| \sum_{i=1}^{p-1} \psi_i \varepsilon_{t-i} \right| \left| \mathbf{1}(r - u_{n-2}(t-q) < \varepsilon_{t-q} < r - u_{n-1}(t-q)) \right| \right] + \\
 &\quad \left| \psi_p \right| E \left[|\varepsilon_{t-p}| \left| \mathbf{1}(r - u_{n-2}(t-q) < \varepsilon_{t-q} < r - u_{n-1}(t-q)) \right| \right] \\
 &= E \left[\left| \sum_{i=1}^{p-1} \psi_i \varepsilon_{t-i} \right| \int_{r-u_{n-2}(t-q)}^{r-u_{n-1}(t-q)} \right] + |\psi_p| E \left[\int_{r-u_{n-2}(t-q)}^{r-u_{n-1}(t-q)} |x| f(x) \right] \\
 &\leq \underbrace{\sup_x |f(x)| \times E \left| \sum_{i=1}^{p-1} \psi_i \varepsilon_{t-i} \right|}_c \times \underbrace{E |u_{n-1}(t-q) - u_{n-2}(t-q)|}_{A_{n-1}(t-q)} + \underbrace{|\psi_p| \sup_x |x f(x)|}_b \times \underbrace{E |u_{n-1}(t-q) - u_{n-2}(t-q)|}_{A_{n-1}(t-q)} \\
 &= (b |\psi_p| + c) A_{n-1}(t-q)
 \end{aligned}$$

d'où

$$A_n(t) \leq (b |\psi_p| + c) A_{n-1}(t-q) \leq (b |\psi_p| + c)^2 A_{n-2}(t-2q) \leq \dots \leq (b |\psi_p| + c)^{n-1} A_1(t-(n-1)q)$$

si $(b |\psi_p| + c) < 1$, alors la suite $A_n(t)$ est de Cauchy, donc elle est convergente.

Quand $p < q$, si $c < 1$, alors la suite $A_n(t)$ est convergente. ■

2.3 Modèle ARMA conditionnellement hétéroscédastique à seuils

2.3.1 Introduction

Dans cette section, nous étudions les propriétés probabilistes d'une classe plus large de modèles à seuils qui est la classe des modèles ARMA conditionnellement hétéroscédastique à seuils noté DTARMACH; ce modèle est une généralisation du modèle DTARCH de Li et Li (1996) et inclut plusieurs modèles de séries chronologiques connus comme cas particuliers, comme le modèle GARCH, le modèle ARMA à seuil (TARMA) et le modèle TAR, pour lesquels nous donnerons les conditions de stationnarité et de finitude des moments. Cependant, les difficultés essentielles sont dans l'établissement de l'irréductibilité et dans la construction des fonctions tests pour l'utilisation du critère de Tweedie. Pour cela, pour prouver l'existence de la solution strictement stationnaire pour le modèle DTARMACH, on utilise les résultats de Tweedie (1988) où l'hypothèse de l'irréductibilité a été relaxée.

2.3.2 Le modèle DTARMACH

On définit le modèle ARMA conditionnellement hétéroscédastique à seuils pour le processus $\{y_t\}$ par

$$y_t = \phi_0^{(j)} + \sum_{i=1}^{p_j} \phi_i^{(j)} y_{t-i} + \sum_{i=1}^{q_j} \theta_i^{(j)} \varepsilon_{t-i} + \varepsilon_t, \quad r_{j-1} < y_{t-b} \leq r_j, \quad (2.28)$$

$$\varepsilon_t = z_t h_t^{1/2}, \quad (2.29)$$

$$h_t = \alpha_0^{(k)} + \sum_{i=1}^{u_k} \alpha_i^{(k)} \varepsilon_{t-i}^2 + \sum_{i=1}^{v_k} \beta_i^{(k)} h_{t-i}, \quad a_{k-1} < y_{t-d} \leq a_k, \quad (2.30)$$

où $j = 1, \dots, l_1$, $k = 1, \dots, l_2$, b et d sont des constantes dans \mathbb{N}^* , les paramètres seuils vérifient $-\infty = r_0 < r_1 < \dots < r_{l_1} = \infty$ et $-\infty = a_0 < a_1 < \dots < a_{l_2} = \infty$; les coefficients $\phi_i^{(j)}$, $\theta_i^{(j)}$, $\alpha_i^{(k)}$ et $\beta_i^{(k)}$ sont des constantes réelles; $\alpha_0^{(k)} > 0$, $\alpha_i^{(k)} \geq 0$ et $\beta_i^{(k)} \geq 0$; $\{z_t\}$ est une suite de variables aléatoires *i.i.d* de moyenne nulle et de variance 1. Quand $\theta_i^{(j)} = \beta_i^{(k)} = 0$ pour tout i , j et k , le modèle DTARMACH se réduit à un modèle DTARCH (Li et Li, 1996). Le modèle DTARMACH inclut, comme cas particuliers, plusieurs modèles connus. Parmi ces modèles on a par exemple, les

modèles ARCH, GARCH, TAR, TAR-ARCH, TMA et TARMA et bien d'autres. Pour appliquer le critère de Tweedie, réécrivons (2.28)-(2.30) sous la forme espace état. De (2.29)-(2.30), on a

$$\varepsilon_t^2 = \alpha_0^{(k)} z_t^2 + \sum_{i=1}^{u_k} \alpha_i^{(k)} z_t^2 \varepsilon_{t-i}^2 + \sum_{i=1}^{v_k} \beta_i^{(k)} z_t^2 h_{t-i}, \quad a_{k-1} < y_{t-d} \leq a_k, k = 1, \dots, l_2 \quad (2.31)$$

Posons

$$\begin{aligned} Y_t &= (y_t, \dots, y_{t-p+1}, \varepsilon_t, \dots, \varepsilon_{t-q+1})'_{(p+q) \times 1}, \\ H_t &= (\varepsilon_t^2, \dots, \varepsilon_{t-u+1}^2, h_t, \dots, h_{t-v+1})'_{(u+v) \times 1}, \end{aligned}$$

avec $p = \max_j p_j$, $q = \max_j q_j$, $u = \max_k u_k$ et $v = \max_k v_k$, on note aussi

$$\tilde{\Phi}^{(j)} = \left(\begin{array}{ccc|ccc} \phi_1^{(j)} & \dots & \phi_p^{(j)} & \theta_1^{(j)} & \dots & \theta_q^{(j)} \\ I_{(p-1) \times (p-1)} & & O_{(p-1) \times 1} & & O_{(p-1) \times q} & \\ \hline & & O_{q \times p} & 0 & \dots & 0 \\ & & & I_{(q-1) \times (q-1)} & & O_{(q-1) \times 1} \end{array} \right) \quad (2.32)$$

$$\tilde{\alpha}^{(k)} = \left(\begin{array}{ccc|ccc} \alpha_1^{(k)} z_t^2 & \dots & \alpha_u^{(k)} z_t^2 & \beta_1^{(k)} z_t^2 & \dots & \beta_v^{(k)} z_t^2 \\ I_{(u-1) \times (u-1)} & & O_{(u-1) \times 1} & & O_{(u-1) \times v} & \\ \hline \alpha_1^{(k)} & \dots & \alpha_u^{(k)} & \beta_1^{(k)} & \dots & \beta_v^{(k)} \\ & & O_{(v-1) \times u} & & I_{(v-1) \times (v-1)} & O_{(v-1) \times 1} \end{array} \right) \quad (2.33)$$

où $\phi_i^{(j)} = 0$ si $i > p_j$, $\theta_i^{(j)} = 0$ si $i > q_j$, $\alpha_i^{(k)} = 0$ si $i > u_k$, $\beta_i^{(k)} = 0$ si $i > v_k$, $j = 1, \dots, l_1$, $k = 1, \dots, l_2$, et $I_{r \times r}$ est la matrice identité de taille $r \times r$.

Les équations (2.28)-(2.30) sont équivalentes à l'écriture suivante :

$$Y_t = \sum_{j=1}^{l_1} (\tilde{\Phi}_0^{(j)} + \tilde{\Phi}^{(j)} Y_{t-1}) \mathbf{1}_{(r_{j-1} < y_{t-b} \leq r_j)} + \varepsilon_t \eta, \quad (2.34)$$

$$H_t = \sum_{k=1}^{l_2} (\tilde{\alpha}_{0t}^{(k)} + \tilde{\alpha}_t^{(k)} H_{t-1}) \mathbf{1}_{(a_{k-1} < y_{t-d} \leq a_k)}; \quad (2.35)$$

avec $\tilde{\Phi}_0^{(j)} = (\phi_0^{(j)}, 0, \dots, 0)'_{(p+q) \times 1}$, η est un vecteur de dimension $(p+q) \times 1$ dont toutes

les composantes sont nulles sauf la 1^{ère} et la $(p+1)^{ème}$ qui prennent la valeur 1; $\tilde{\alpha}_{0t}^{(k)}$ est un vecteur de dimension $(u+v) \times 1$ dont toutes les composantes sont nulles sauf la 1^{ère} et la $(u+1)^{ème}$ qui sont $\alpha_0^{(k)} z_t^2$ et $\alpha_0^{(k)}$ respectivement.

Soit

$$X_t = \begin{pmatrix} Y_t \\ H_t \end{pmatrix}, \quad B_{0t}^{(j,k)} = \begin{pmatrix} \tilde{\Phi}_0^{(j)} \\ \tilde{\alpha}_{0t}^{(k)} \end{pmatrix} \text{ et } B_t^{(j,k)} = \begin{pmatrix} \tilde{\Phi}^{(j)} & 0 \\ 0 & \tilde{\alpha}_t^{(k)} \end{pmatrix},$$

en combinant (2.34)-(2.35), on obtient

$$X_t = \sum_{j=1}^{l_1} \sum_{k=1}^{l_2} (B_{0t}^{(j,k)} + B_t^{(j,k)} X_{t-1}) \mathbf{1}_{(X_{t-1} \in \mathbb{R}_{j,b;k;d})} + \tilde{\varepsilon}_t, \quad (2.36)$$

où $\tilde{\varepsilon}_t = (\varepsilon_t \eta', 0, \dots, 0)'_{(p+q+u+v) \times 1}$, et $\mathbb{R}_{j,b;k;d}$ est le produit Cartésien

$$\mathbb{R}_{j,b;k;d} = \mathbb{R}^{b-1} \times]r_{j-1}, r_j] \times \mathbb{R}^{d-b-1} \times]a_{k-1}, a_k] \times \mathbb{R}^{p+q-d} \times \mathbb{R}_+^{u+v},$$

avec $\mathbb{R} =]-\infty, +\infty[$ et $\mathbb{R}_+ = [0, +\infty[$ (il n'y a pas de perte de généralité si on prend $b \leq d \leq p$ du fait que si $d > p$ on a $\phi_i^{(j)} = 0$ pour $i = p+1, \dots, d$).

Il est clair que l'équation (2.36) est équivalente à (2.28)-(2.30) et que $\{X_t\}$ défini par (2.36) est une chaîne de Markov ayant pour espace d'état $\mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v}$. Pour nous assurer que $\{X_t\}$ est une chaîne de Feller, on fait l'hypothèse suivante

Hypothèse 2.3.1 La fonction F sur $\mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v}$ définie par

$$F(x) = \sum_{j=1}^{l_1} \sum_{k=1}^{l_2} [E(B_{0t}^{(j,k)}) + E(B_t^{(j,k)})x] \mathbf{1}_{(x \in \mathbb{R}_{j,b;k;d})}$$

est continue en x .

Théorème 2.3.1 (Ling 1999) *Sous l'hypothèse 2.3.1, si*

$$\sum_{i=1}^p \max_j |\phi_i^{(j)}| < 1 \quad (2.37)$$

et

$$\sum_{i=1}^u \max_k \alpha_i^{(k)} + \sum_{i=1}^v \max_k \beta_i^{(k)} < 1, \quad (2.38)$$

alors, il existe une solution strictement stationnaire $\{y_t, \varepsilon_t\}$ satisfaisant (2.28)-(2.30) et $E_{\pi_1}(|y_t|)$ et $E_{\pi_2}(\varepsilon_t^2)$ sont finis, où π_1 et π_2 sont les distributions stationnaires de $\{y_t\}$ et $\{\varepsilon_t\}$ respectivement.

Remarques 2.3.1

- (i) Quand $\phi_i^{(j)} = \theta_i^{(j)} = \beta_i^{(k)} = 0$ pour tout i, j et k , le modèle DTARMACH se réduit à un modèle ARCH. Dans ce cas là, la condition (2.38) est plus facile à vérifier par rapport à celle de Liu et al. (1997) et, quand $u = 1$, la condition (2.38) se réduit à leur condition, i.e. $\max_k \alpha_1^{(k)} < 1$.
- (ii) Le théorème 4.1 montre seulement l'existence d'une solution strictement stationnaire pour (2.28)-(2.30). Malheureusement, on ne peut pas prouver que cette solution est unique ou ergodique, ce qu'on ne peut donner que pour des cas spéciaux.

Théorème 2.3.2 (Ling (1999)) Supposons que l'hypothèse 3.2.1 est vraie et $E(z_t^{2m})$ est fini pour $m > 0$. Si

$$\sum_{i=1}^p \max_j |\phi_i^{(j)}| < 1 \quad (2.39)$$

et

$$\rho \left[\max_k E(\tilde{\alpha}_t^{(k) \otimes m}) \right] < 1, \quad (2.40)$$

alors (2.28)-(2.30) possède une solution strictement stationnaire $\{y_t, \varepsilon_t\}$, et $E_{\pi_1}(|y_{t-t'_1} y_{t-t'_2} \dots y_{t-t'_m}|)$ et $E_{\pi_2}(|\varepsilon_{t-t''_1} \varepsilon_{t-t''_2} \dots \varepsilon_{t-t''_m}|)$ sont finis, avec $t'_i \in \{0, 1, \dots, p\}$, $t''_i \in \{0, 1, \dots, u\}$, $i = 1, \dots, m$, et π_1 et π_2 les distributions stationnaires de $\{y_t\}$ et $\{\varepsilon_t\}$ respectivement.

Remarques 2.3.2 Dans le théorème 2.3.2, si le modèle DTARMACH se réduit à un modèle DTARCH, alors $\tilde{\alpha}_t^{(k)}$ est remplacée par la matrice suivante :

$$\begin{pmatrix} \alpha_1^{(k)} z_t^2 & \dots & \alpha_u^{(k)} z_t^2 \\ & I_{(u-1) \times (u-1)} & O_{(u-1) \times 1} \end{pmatrix}.$$

Le résultat du théorème 2.3.2, peut être étendu pour permettre à $t'_i > p$ par l'introduction de coefficients nuls $\phi_i^{(j)}$, $i = p + 1, \dots, t'_i$ dans (2.28). Le cas quand $t''_i > u$ peut être

traité de façon analogue.

Afin de rendre la preuve plus transparente, nous donnons les deux lemmes suivants.

Lemme 2.3.1 (Ling (1999)) (i) *Supposons que*

$$\sum_{i=1}^p \max_j |\phi_i^{(j)}| < 1. \quad (2.41)$$

Alors, il existe un vecteur $L > 0$ tel que, $\forall j$, $(I - |\tilde{\Phi}^{(j)\otimes m}|')L > 0$.

(ii) *Supposons que $E(z_t^{2m}) < \infty$ et que*

$$\rho \left[\max_k E(\tilde{\alpha}_t^{(k)\otimes m}) \right] < 1. \quad (2.42)$$

Alors, il existe un vecteur $M > 0$ tel que, $\forall k$, $(I - [E(\tilde{\alpha}_t^{(k)})^{\otimes m}]')M > 0$.

Preuve. Par calcul direct, on obtient

$$\det \left(\lambda - \max_j |\tilde{\Phi}^{(j)}| \right) = \lambda^q \left(\lambda^p - \max_j |\phi_1^{(j)}| \lambda^{p-1} - \dots - \max_j |\phi_p^{(j)}| \right). \quad (2.43)$$

D'après le lemme B.3 (a) (Voir l'annexe), (2.41) et (2.43), on a $\rho(\max_j |\tilde{\Phi}^{(j)}|) < 1$. De plus, d'après le lemme B.2 (Voir l'annexe), $\rho[(\max_j |\tilde{\Phi}^{(j)}|)^{\otimes m}] < 1$, alors $[I - (\max_j |\tilde{\Phi}^{(j)}|')^{\otimes m}]$ est inversible. Du fait que tous les éléments de $(\max_j |\tilde{\Phi}^{(j)}|')^{\otimes m}$ sont positifs, on peut choisir un vecteur $L_1 > 0$ tel que

$$L := \left[I - \left(\max_j |\tilde{\Phi}^{(j)}|' \right)^{\otimes m} \right]^{-1} L_1 = L_1 + \sum_{i=1}^{\infty} \left[\left(\max_j |\tilde{\Phi}^{(j)}|' \right)^{\otimes m} \right]^i L_1 > 0.$$

Par le lemme B.1 (d)-(e) (Voir l'annexe),

$$\max_j |\tilde{\Phi}^{(j)\otimes m}|'L = \max_j (|\tilde{\Phi}^{(j)}|')^{\otimes m} L \leq \left(\max_j |\tilde{\Phi}^{(j)}|' \right)^{\otimes m} L.$$

Alors, $\forall j$

$$\begin{aligned} (I - |\tilde{\Phi}^{(j)\otimes m}|')L &\geq \left(I - \max_j |\tilde{\Phi}^{(j)\otimes m}|' \right) L = \left[I - \max_j (|\tilde{\Phi}^{(j)\otimes m}|') \right] L \\ &\geq \left[I - \left(\max_j |\tilde{\Phi}^{(j)}|' \right)^{\otimes m} \right] L = L_1 > 0, \end{aligned}$$

d'où, (i) est vérifiée. De façon similaire, on peut prouver que (ii) est aussi vérifiée.

■

Lemme 2.3.2 (Ling (1999)) *Supposons qu'il existe deux vecteurs $L > 0$ et $M > 0$ tels que, $\forall j, k$,*

$$(I - |\tilde{\Phi}^{(j)\otimes m}|')L > 0 \quad (2.44)$$

et

$$(I - [E(\tilde{\alpha}_t^{(k)})^{\otimes m}]')M > 0. \quad (2.45)$$

Alors, il existe un ensemble compact

$$A = \left\{ x : \left(\sum_{i=1}^{p+q} |x_i| \right)^m + \left(\sum_{i=1}^{u+v} x_{p+q+i} \right)^m \leq \Delta \right\} \subset \mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v}$$

et $\delta > 0$ tel que, la fonction g définie par

$$g(x) = 1 + |Y^{\otimes m}|'L + (H^{\otimes m})'M \quad (2.46)$$

satisfait

$$E(g(X_t) \mid X_{t-1} = x) \leq g(x) + g_1(x), \quad x \in \mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v}, \quad (2.47)$$

et

$$E(g(X_t) \mid X_{t-1} = x) \leq (1 - \delta)g(x), \quad x \in A^c, \quad (2.48)$$

où $A^c = \mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v} - A$, $x = (Y', H)'$, $Y \in \mathbb{R}^{p+q}$, $H \in \mathbb{R}_+^{u+v}$, x_i est la $i^{\text{ème}}$ composante de x , $\max_{x \in A} g_1(x) < c$, δ et Δ sont des constantes indépendantes de x .

Preuve. la preuve est faite pour $m = 3$, pour d'autres valeurs de m , la technique est analogue. Pour tout $x = (Y', H)' \in \mathbb{R}_{j,b;k,d}$, par un calcul direct, on peut montrer

que

$$\begin{aligned}
(\tilde{\Phi}_0^{(j)} + \tilde{\Phi}^{(j)}Y + \varepsilon_t\eta)^{\otimes 3} &= \tilde{\Phi}^{(j)\otimes 3}Y^{\otimes 3} + [c_1^{(j)}\varepsilon_t + c_2^{(j)}]Y^{\otimes 2} \\
&+ [c_3^{(j)}\varepsilon_t^2 + c_4^{(j)}\varepsilon_t + c_5^{(j)}]Y \\
&+ [c_6^{(j)}\varepsilon_t^3 + c_7^{(j)}\varepsilon_t^2 + c_8^{(j)}\varepsilon_t + c_9^{(j)}], \tag{2.49}
\end{aligned}$$

où $c_1^{(j)}, \dots, c_9^{(j)}$ sont des vecteurs ou des matrices indépendantes de x .

D'après le lemme B.1 (a)-(c) (Voir l'annexe) et l'inégalité triangulaire, il est facile de montrer que

$$\begin{aligned}
E(|\varepsilon_t| \mid X_{t-1} = x) &= E(|z_t h_t^{1/2}| \mid X_{t-1} = x) = E(|z_t|(\alpha_0^{(k)} + H' C^{(k)})^{1/2}) \\
&\leq c_{10}^{(k)}(1 + \tilde{H})^{1/2}, \tag{2.50}
\end{aligned}$$

où $c_{10}^{(k)}$ est une constante positive indépendante de x , $C^{(k)} = (\alpha_1^{(k)}, \dots, \alpha_u^{(k)}, \beta_1^{(k)}, \dots, \beta_v^{(k)})'$, et $\tilde{H} = \sum_{i=1}^{u+v} x_{p+q+i}$. Et de façon similaire, on peut montrer que

$$E(\varepsilon_t^2 \mid X_{t-1} = x) \leq c_{11}^{(k)}(1 + \tilde{H}), \tag{2.51}$$

et

$$E(\varepsilon_t^3 \mid X_{t-1} = x) \leq c_{12}^{(k)}(1 + \tilde{H})^{3/2}, \tag{2.52}$$

où $c_{11}^{(k)}$ et $c_{12}^{(k)}$ sont des constantes positives indépendantes de x .

Soit $\tilde{Y} = \sum_{i=1}^{p+q} |x_i|$, et

$$\begin{aligned}
\lambda_1 &= \max_{j,k} \{ \text{toutes les composantes de } c_{10}^{(k)}|c_1^{(j)}|'L, |c_2^{(j)}|'L, c_{11}^{(k)}|c_3^{(j)}|'L, c_{10}^{(k)}|c_4^{(j)}|'L, \\
&|c_5^{(j)}|'L, c_{12}^{(k)}|c_6^{(j)}|'L, c_{11}^{(k)}|c_7^{(j)}|'L, c_{10}^{(k)}|c_8^{(j)}|'L, |c_9^{(j)}|'L \}.
\end{aligned}$$

De (2.49)-(2.52), on a

$$\begin{aligned}
E[|(\tilde{\Phi}_0^{(j)} + \tilde{\Phi}^{(j)}Y + \varepsilon_t\eta)^{\otimes 3}|'L] &\leq |Y^{\otimes 3}|'|\tilde{\Phi}^{(j)\otimes 3}|'L + |Y^{\otimes 2}|'[(1 + \tilde{H})^{1/2}c_{10}^{(k)}|c_1^{(j)}|'L + |c_2^{(j)}|'L] \\
&\quad + |Y|'[(1 + \tilde{H})c_{11}^{(k)}|c_3^{(j)}|'L + (1 + \tilde{H})^{1/2}c_{10}^{(k)}|c_4^{(j)}|'L + |c_5^{(j)}|'L] \\
&\quad + [(1 + \tilde{H})^{3/2}c_{12}^{(k)}|c_6^{(j)}|'L + (1 + \tilde{H})c_{11}^{(k)}|c_7^{(j)}|'L] \\
&\quad + (1 + \tilde{H})^{1/2}c_{10}^{(k)}|c_8^{(j)}|'L + |c_9^{(j)}|'L \\
&\leq |Y^{\otimes 3}|'|\tilde{\Phi}^{(j)\otimes 3}|'L + \lambda_1\tilde{Y}^2[(1 + \tilde{H})^{1/2} + 1] \\
&\quad + \lambda_1\tilde{Y}[(1 + \tilde{H}) + (1 + \tilde{H})^{1/2} + 1] \\
&\quad + \lambda_1[(1 + \tilde{H})^{3/2} + (1 + \tilde{H}) + (1 + \tilde{H})^{1/2} + 1] \\
&\leq |Y^{\otimes 3}|'|\tilde{\Phi}^{(j)\otimes 3}|'L + 2\lambda_1\tilde{Y}^2(1 + \tilde{H})^{1/2} \\
&\quad + 3\lambda_1\tilde{Y}(1 + \tilde{H}) + 4\lambda_1(1 + \tilde{H})^{3/2}. \tag{2.53}
\end{aligned}$$

De même,

$$\begin{aligned}
E[|(\tilde{\alpha}_{0t}^{(k)} + \tilde{\alpha}_t^{(k)}H)^{\otimes 3}|'M] &= (H^{\otimes 3})'[E(\alpha_t^{(k)\otimes 3})]'M + C_1^{(k)}M + H'(C_2^{(k)})'M + (H^{\otimes 2})'(C_3^{(k)})'M \\
&\leq (H^{\otimes 3})'[E(\alpha_t^{(k)\otimes 3})]'M + \lambda_1(1 + \tilde{H} + \tilde{H}^2), \tag{2.54}
\end{aligned}$$

où $C_1^{(k)}$, $C_2^{(k)}$ et $C_3^{(k)}$ sont des vecteurs ou des matrices constantes dont tous leurs éléments sont positive et independants de x , et $\lambda_2 = \max_k\{\text{toutes les composantes de } C_1^{(k)}M, (C_2^{(k)})'M \text{ et } (C_3^{(k)})'M\}$. D'après (2.53) et (2.54), on a

$$\begin{aligned}
E[g(X_t) | X_{t-1} = x] &= 1 + E[|(\tilde{\Phi}_0^{(j)} + \tilde{\Phi}^{(j)}Y + \varepsilon_t\eta)^{\otimes 3}|'L | X_{t-1} = x] + E[|(\tilde{\alpha}_{0t}^{(k)} + \tilde{\alpha}_t^{(k)}H)^{\otimes 3}|'M] \\
&\leq 1 + |Y^{\otimes 3}|'|\tilde{\Phi}^{(j)\otimes 3}|'L + (H^{\otimes 3})'[E(\alpha_t^{(k)\otimes 3})]'M + g_1(x) \\
&= 1 + |Y^{\otimes 3}|'L + (H^{\otimes 3})'M - |Y^{\otimes 3}|'L^{(j)} - (H^{\otimes 3})'M^{(k)} + g_1(x) \\
&= g(x) \left[1 - \frac{|Y^{\otimes 3}|'L^{(j)} + (H^{\otimes 3})'M^{(k)}}{g(x)} + \frac{g_1(x)}{g(x)} \right], \tag{2.55}
\end{aligned}$$

où $g_1(x) = 2\lambda_1\tilde{Y}^2(1 + \tilde{H})^{1/2} + 3\lambda_1\tilde{Y}(1 + \tilde{H}) + 4\lambda_1(1 + \tilde{H})^{3/2} + \lambda_1(1 + \tilde{H} + \tilde{H}^2)$, $L^{(j)} = (I - |\tilde{\Phi}^{(j)\otimes 3}|'L)$, et $M^{(k)} = (I - [E(\alpha_t^{(k)\otimes 3})]'M)$.

Soit

$$\begin{aligned}
A &= \{x : \tilde{Y}^3 + \tilde{H}^3 \leq \Delta, x \in \mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v}\}, \\
\lambda_3 &= \min_{j,k} \{\text{toutes les composantes de } L^{(j)} \text{ et } M^{(k)}\}, \\
\lambda_4 &= \max\{\text{toutes les composantes de } L \text{ et } M\}, \\
\lambda_5 &= \min\{\text{toutes les composantes de } L \text{ et } M\}.
\end{aligned}$$

Il est clair que l'ensemble A est un compact sur $\mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v}$. D'après (2.44) et (2.45), $L^{(j)}$ et $M^{(k)}$ sont des vecteurs positifs pour tout j et k , même pour L et M , alors λ_3 , λ_4 et λ_5 sont positives. D'après (2.55), on a

$$E[g(X_t) \mid X_{t-1} = x] \leq g(x) - g_1(x), \quad x \in \mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v},$$

où $\max_{x \in A} g_1(x) < c(\Delta)$, $c(\Delta)$ est une constante indépendante de x . Alors (2.47) est vérifiée.

Prenons $\Delta > \max\{1/\lambda_4, 1\}$. Quand $x \in A^c$,

$$\lambda_5 \Delta < \lambda_5(\tilde{Y}^3 + \tilde{H}^3) \leq g(x) \leq 1 + \lambda_4(\tilde{Y}^3 + \tilde{H}^3) \leq 2\lambda_4(\tilde{Y}^3 + \tilde{H}^3).$$

Ainsi,

$$\frac{|Y^{\otimes 3}|'L^{(j)} + (H^{\otimes 3})'M^{(k)}}{g(x)} \geq \frac{\lambda_3(\tilde{Y}^3 + \tilde{H}^3)}{2\lambda_4(\tilde{Y}^3 + \tilde{H}^3)} = \frac{\lambda_3}{2\lambda_4}, \quad (2.56)$$

et de plus, comme $\tilde{Y}^2, \tilde{H}^2 \leq (\tilde{Y}^3 + \tilde{H}^3)^{2/3}$, et $1 + \tilde{H} \leq 2(\tilde{Y}^3 + \tilde{H}^3)^{1/3}$, on a

$$\frac{g_1(x)}{g(x)} \leq \frac{g_1(x)}{\lambda_5(\tilde{Y}^3 + \tilde{H}^3)} \leq \frac{C_1(\tilde{Y}^3 + \tilde{H}^3)^{5/6}}{\lambda_5(\tilde{Y}^3 + \tilde{H}^3)} \leq \frac{C}{\sqrt[6]{\Delta}}, \quad (2.57)$$

où C_1 et C sont des constantes positives indépendantes de x et Δ . D'après (2.56) et (2.57), on a

$$E[g(X_t) \mid X_{t-1} = x] \leq g(x) \left(1 - \frac{\lambda_3}{2\lambda_4} + \frac{C}{\sqrt[6]{\Delta}}\right).$$

Si $0 < \delta < \lambda_3/2\lambda_4$ et $\Delta > \max\{1, 1/\lambda_4, C^6/(\lambda_3/2\lambda_4 - \delta)^6\}$, on a

$$E[g(X_t) \mid X_{t-1} = x] \leq g(x)(1 - \delta), \quad x \in A^c,$$

ainsi, (2.48) est vérifiée. ■

Par un calcul direct, on peut montrer que l'équation caractéristique de $\max_k \{E(\tilde{\alpha}_t^{(k)})\}$ est

$$\lambda^{u+v} - \lambda^v \sum_{i=1}^u \max_k \alpha_i^{(k)} \lambda^{u-i} - \lambda^u \sum_{i=1}^v \max_k \beta_i^{(k)} \lambda^{v-i} = 0.$$

En appliquant le lemme B.3 (a) et la condition (2.38) du théorème 2.3.1, il vient

$$\rho \left[\max_k E(\tilde{\alpha}_t^{(k)}) \right] < 1.$$

Ainsi, quand $m = 1$, théorème 2.3.2 se réduit au théorème 2.3.1. Dans ce qui suit, on donne seulement la preuve du théorème 2.3.2.

Preuve. On donne la preuve pour $m = 3$ seulement comme dans le lemme 2.3.2. D'abord, par l'hypothèse 2.3.1, pour chaque fonction g continue et bornée sur $\mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v}$, $E[g(X_t) \mid X_{t-1} = x]$ est continue en x , i.e. $\{X_t\}$ est une chaîne de Feller. Maintenant, d'après (2.39)-(2.40), lemme 2.3.1 et lemme 2.3.2, on sait qu'il existe deux vecteurs L et $M > 0$, un ensemble compact $A = \{x : \tilde{Y}^3 + \tilde{H}^3 \leq \Delta, x \in \mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v}\}$, et $\delta > 0$ tel que, la fonction définie par

$$g(x) = 1 + |Y^{\otimes 3}|'L + (H^{\otimes 3})'M$$

satisfait

$$E(g(X_t) \mid X_{t-1} = x) \leq g(x) + g_1(x), \quad x \in \mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v}, \quad (2.58)$$

et

$$E(g(X_t) \mid X_{t-1} = x) \leq (1 - \delta)g(x), \quad x \in A^c, \quad (2.59)$$

où $\max_{x \in A} g_1(x) < c$, et $\Delta (> 0)$ et δ sont des constantes indépendantes de x .

D'après (2.59), comme $g(x) \geq 1$, $E(g(X_t) \mid X_{t-1} = x) \leq g(x) - \delta$. D'après le théorème 1.6.5 (i), il existe μ une mesure σ -finie invariante pour P avec $0 < \mu(A) < \infty$.

Notons $C = \max\{\text{toutes les composantes de } L \text{ et } M\}$. D'après (2.58), quand $x \in A$,

$$E(g(X_t) \mid X_{t-1} = x) \leq 1 + C[\tilde{Y}^3 + \tilde{H}^3] + g_1(x) \leq \Delta_1 < \infty,$$

où Δ_1 est une constante positive indépendante de x . Alors,

$$\int_A \mu(dx) \left\{ \int_{A^c} P(x, dy) g(y) \right\} \leq \int_A \mu(dx) E(g(X_t) | X_{t-1} = x) \leq \Delta_1 \mu(A) < \infty.$$

D'après le théorème 1.6.5 (ii), $\{X_t\}$ possède une mesure μ invariante et finie, alors $\pi = \mu/\mu(\mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v})$ est une mesure de probabilité invariante pour $\{X_t\}$, i.e. il existe donc une solution strictement stationnaire satisfaisant (2.36).

Soit $f(x)$ la fonction définie sur $\mathbb{R}^{p+q} \times \mathbb{R}_+^{u+v}$ par

$$f(x) = C_1 \delta [\tilde{Y}^3 + \tilde{H}^3].$$

où $C_1 = \min\{\text{toutes les composantes de } L \text{ et } M\}$. D'après (2.59), il est clair que, quand $x \in A^c$,

$$\int_{A^c} P(x, dy) g(y) \leq E[g(X_t) | X_{t-1} = x] \leq g(x) - \delta g(x) \leq g(x) - f(x).$$

Ainsi, d'après le théorème 1.6.5 (iii), on a

$$E_\pi(f(X_t)) = C_1 \delta E \left[\left(\sum_{i=1}^{p+q} |x_{i,t}| \right)^3 + \left(\sum_{i=1}^{u+v} x_{p+q+i,t} \right)^3 \right] < \infty,$$

où π est la distribution stationnaire de $\{X_t\}$, et $x_{i,t}$ est la $i^{\text{ème}}$ composante de X_t . Ainsi,

$$E_{\pi_1}(|y_{t-t'_1} y_{t-t'_2} y_{t-t'_3}|) < \infty \text{ et } E_{\pi_2}(|\varepsilon_{t-t'_1} \varepsilon_{t-t'_2} \varepsilon_{t-t'_3}|) < \infty,$$

où π_1 et π_2 sont les distributions stationnaires de $\{y_t\}$ et $\{\varepsilon_t\}$ respectivement. ■

2.3.3 Modèles TARMA (p, q) et TAR (p)

Dans cette partie, on considère deux cas spéciaux du modèle DTARMACH qui sont les modèle AR à seuils TAR et ARMA à seuils TARMA.

On dit que le processus $\{y_t\}$ suit un modèle ARMA à seuils, noté TARMA (p, q) ,

s'il est solution de l'équation stochastique suivante

$$y_t = \phi_0^{(j)} + \sum_{i=1}^p \phi_i^{(j)} y_{t-i} + \sum_{i=1}^q \theta_i^{(j)} \varepsilon_{t-i} + \varepsilon_t, \quad r_{j-1} < y_{t-b} \leq r_j, \quad (2.60)$$

où $j = 1, \dots, l$, les paramètres seuils satisfont $-\infty = r_0 < r_1 < \dots < r_l = \infty$, b est le paramètre de délai, les coefficients $\phi_i^{(j)}$ et $\theta_i^{(j)}$ sont des constantes réelles, et $\{\varepsilon_t\}$ est une suite de variables aléatoire *i.i.d.* Un modèle TAR(p) est modèle TARMA($p, 0$)

$$y_t = \phi_0^{(j)} + \sum_{i=1}^p \phi_i^{(j)} y_{t-i} + \varepsilon_t, \quad r_{j-1} < y_{t-b} \leq r_j. \quad (2.61)$$

Comme dans (2.34), nous pouvons réécrire les équations (2.60)-(2.61) comme suit

$$Y_t = \sum_{j=1}^l (\tilde{\Phi}_0^{(j)} + \tilde{\Phi}^{(j)} Y_{t-1}) \mathbf{1}_{(Y_{t-1} \in \mathbb{R}_{j,b})} + \varepsilon_t \eta, \quad (2.62)$$

où $\tilde{\Phi}_0^{(j)}$, $\tilde{\Phi}^{(j)}$ et η pour (2.60) sont définies comme dans (2.34), et pour (2.61)

$$\tilde{\Phi}^{(j)} = \begin{pmatrix} \phi_1^{(j)} & \dots & \phi_p^{(j)} \\ & I_{(p-1) \times (p-1)} & O_{(p-1) \times 1} \end{pmatrix}$$

$\tilde{\Phi}_0^{(j)} = (\phi_0^{(j)}, 0, \dots, 0)'_{p \times 1}$, η est un vecteur de dimension $p \times 1$ dont toutes les composantes sont nulles sauf la 1^{ère}, $\mathbb{R}_{j,b}$ est le produit cartésien

$$\mathbb{R}_{j,b} = \mathbb{R}^{b-1} \times]r_{j-1}, r_j] \times \mathbb{R}^{p-b} \times \mathbb{R}^q \text{ pour (2.60);}$$

$$\mathbb{R}_{j,b} = \mathbb{R}^{b-1} \times]r_{j-1}, r_j] \times \mathbb{R}^{p-b} \text{ pour (2.61)}$$

Hypothèse 2.3.2 la fonction F définie sur $\mathbb{R}^{p+q}(\mathbb{R}^p)$ par

$$F(x) = \sum_{k=1}^l [\tilde{\Phi}_0^{(k)} + \tilde{\Phi}^{(k)} x] \mathbf{1}_{(x \in \mathbb{R}_{k,b})}$$

est continue en x .

D'après les théorèmes 2.3.1-2.3.2, on a le résultat suivant.

Théorème 2.3.3 (Ling (1999)) *Supposons que l'hypothèse 2.3.2 est vraie et que $E(|\varepsilon_t|^m)$*

est finie. Si

$$\sum_{i=1}^p \max_j |\phi_i^{(j)}| < 1 \quad (2.63)$$

alors (2.60) et (2.61) possèdent des solutions strictements stationnaires, avec $E_\pi(|y_{t-t_1} y_{t-t_2} \dots y_{t-t_m}|)$, où $t_i \in \{0, 1, \dots, p\}$, π est la distribution stationnaire de $\{y_t\}$.

Remarques 2.3.3 Le théorème ci-dessus sous l'hypothèse 2.3.2 montre que, la partie moyenne mobile n'influe pas sur la stationnarité et la finitude des moments pour le modèle ARMA à seuils, i.e. que la contribution dans la stationnarité et la finitude des moments de la partie moyenne mobile peut être ignorée.

Pour le modèle TAR(p), on donne une condition suffisante d'ergodicité géométrique suivante.

Théorème 2.3.4 (Ling (1999)) Supposons que $\{\varepsilon_t\}$ possède une distribution absolument continue, semi-continue inférieurement et strictement positive sur \mathbb{R} . Si $E(|\varepsilon_t|^m) < \infty$ et $\sum_{i=1}^p \max_j |\phi_i^{(j)}| < 1$, alors, $\{y_t\}$ est géométriquement ergodique, et $E_\pi(|y_{t-t_1} y_{t-t_2} \dots y_{t-t_m}|)$ est finie, $t_i \in \{0, 1, \dots, p\}$, $i = 1, \dots, m$, et π est la distribution stationnaire de $\{y_t\}$.

Preuve. La preuve est similaire aux théorèmes 2.3.1 et 2.3.2, en utilisant la condition (b) (Voir l'annexe), le théorème 1.7.5 et la remarque 1.7.1 (ii). ■

Chapitre 3 : Estimation dans des modèles autorégressifs à seuils

3 Estimation dans des modèles autorégressifs à seuils

3.1 Modèle TAR(1) simple

On suppose que la suite $\{\varepsilon_t\}$ possède un moment d'ordre $2+\xi$ fini, $\xi > 0$, et alors la distribution stationnaire de $\{y_t\}$ possède un moment d'ordre 2 fini. σ^2 représente la variance de l'erreur, nous noterons par la suite $(y_t^+)^2$ et $(y_t^-)^2$ par y_t^{+2} et y_t^{-2} respectivement. On fait intervenir une transformation de Householder pour les n_i observations relatives au $i^{\text{ème}}$ régime, les estimateurs des moindres carrés, pour ϕ_1 et ϕ_2 , que l'on peut obtenir aisément, sont donnés par

$$\widehat{\phi}_1 = \frac{\sum_{t=1}^n y_t y_{t-1}^+}{\sum_{t=1}^n y_t^{+2}} \quad (3.1)$$

$$\widehat{\phi}_2 = \frac{\sum_{t=1}^n y_t y_{t-1}^-}{\sum_{t=1}^n y_t^{-2}} \quad (3.2)$$

donc la détermination de $\widehat{\phi}_1$ et $\widehat{\phi}_2$ est effectuée de façon indépendante pour chacun des régimes. L'estimateur naturel pour σ^2 est

$$\widehat{\sigma}^2 = n^{-1} \sum_{t=1}^n \left(y_t - \widehat{\phi}_1 y_{t-1}^+ - \widehat{\phi}_2 y_{t-1}^- \right)^2. \quad (3.3)$$

Notons que $\widehat{\phi}_1$, $\widehat{\phi}_2$ et $\widehat{\sigma}^2$ sont aussi les estimateurs du maximum vraisemblance pour ϕ_1, ϕ_2 et σ^2 respectivement sous l'hypothèse de normalité des erreurs.

Les deux théorèmes suivants établissent la consistance et la normalité asymptotique pour les estimateurs dans (3.1)-(3.3) quand le processus $\{y_t\}$ est ergodique.

Théorème 3.1.1 *Si ϕ_1 et ϕ_2 vérifient (2.4), alors $\widehat{\phi}_1$, $\widehat{\phi}_2$ et $\widehat{\sigma}^2$ sont des estimateurs consistants pour ϕ_1, ϕ_2 et σ^2 respectivement.*

Preuve. Notons que (3.1) et (3.2) peuvent être réécrits comme

$$\widehat{\phi}_1 = \phi_1 + \sum_{t=1}^n y_{t-1}^+ \varepsilon_t \Big/ \sum_{t=1}^n y_{t-1}^{+2} \quad (3.4)$$

$$\widehat{\phi}_2 = \phi_2 + \sum_{t=1}^n y_{t-1}^- \varepsilon_t \Big/ \sum_{t=1}^n y_{t-1}^{-2}. \quad (3.5)$$

Puisque la chaîne de Markov $\{y_t\}$ est ergodique, elle est donc fortement mélangeante, cela implique que $\{y_t^{+k}\}$ est aussi fortement mélangeant, $\forall k \in \mathbb{Z}$, $\{y_{t-1}^+ \varepsilon_t\}$ est encore fortement mélangeant, d'où, ce dernier est aussi ergodique. En utilisant le théorème 2 de Hannan (1970) (p.203), on obtient quand $n \rightarrow \infty$

$$n^{-1} \sum_{t=1}^n y_t^{+2} \rightarrow E(y_t^{+2}) \quad p.s.$$

et

$$n^{-1} \sum_{t=1}^n y_t^+ \varepsilon_t \rightarrow E(y^+) E(\varepsilon_1) = 0 \quad p.s.$$

ainsi, $\widehat{\phi}_1 \rightarrow \phi_1$ *p.s.* quand $n \rightarrow \infty$. De façon similaire, on a $\widehat{\phi}_2 \rightarrow \phi_2$ *p.s.* quand $n \rightarrow \infty$.

De (3.3), on obtient

$$\widehat{\sigma}^2 = n^{-1} \sum_{t=1}^n \varepsilon_t^2 - (\widehat{\phi}_1 - \phi_1) n^{-1} \sum_{t=1}^n y_{t-1}^+ \varepsilon_t - (\widehat{\phi}_2 - \phi_2) n^{-1} \sum_{t=1}^n y_{t-1}^- \varepsilon_t$$

d'après le résultat ci-dessus, on a $\widehat{\sigma}^2 \rightarrow \sigma^2$ *p.s.* quand $n \rightarrow \infty$. ■

Théorème 3.1.2 (Petrucci et Woolford (1984)) *Si ϕ_1 et ϕ_2 vérifient (2.4), alors*

$$\lim_{n \rightarrow \infty} P((nE(y^{+2}))^{1/2}(\widehat{\phi}_1 - \phi_1) \leq x\sigma, (nE(y^{-2}))^{1/2}(\widehat{\phi}_2 - \phi_2) \leq y\sigma) = \Phi(x)\Phi(y),$$

où $\Phi(\cdot)$ est la fonction de répartition de la loi normale standard.

Preuve. Soit $\xi_1, \xi_2 \in \mathbb{R}$, on considère

$$\xi_1 (nE(y^{+2}))^{1/2} (\widehat{\phi}_1 - \phi_1) + \xi_2 (nE(y^{-2}))^{1/2} (\widehat{\phi}_2 - \phi_2).$$

qui est asymptotiquement équivalent à

$$n^{-1/2} \sum_{t=1}^n (\psi_1 y_{t-1}^+ \varepsilon_t + \psi_2 y_{t-1}^- \varepsilon_t), \quad (3.6)$$

avec $\psi_1 = \xi_1 / E(y^{+2})^{1/2}$ et $\psi_2 = \xi_2 / E(y^{-2})^{1/2}$. Cependant,

$$\{\psi_1 y_{t-1}^+ \varepsilon_t + \psi_2 y_{t-1}^- \varepsilon_t, t \geq 1\}$$

est une suite de différence de martingale et satisfait les conditions du théorème 23.1 dans Billingsley (1968). Ainsi, (3.6) converge en loi vers une loi normale $N(0, \sigma_0^2)$ où $\sigma_0^2 = \sigma^2(\xi_1^2 + \xi_2^2)$. d'où le résultat. ■

3.2 Test d'égalité de ϕ_1 et ϕ_2

On veut tester l'hypothèse nulle

$$H_0: \phi_1 = \phi_2 = \phi$$

contre l'alternative

$$H_1: \phi_1 \neq \phi_2.$$

On suppose que $\{\varepsilon_t\}$ est distribuée normalement, la région de rejet pour le test du rapport de vraisemblance est donné par

$$\lambda = [\hat{\sigma}^2 / \hat{\sigma}_R^2]^{(n-1)/2} < c, \quad c > 0, \quad (3.7)$$

avec

$$\hat{\sigma}_R^2 = n^{-1} \sum_{t=1}^n (y_t - \hat{\phi} y_{t-1})^2 \quad \text{où} \quad \hat{\phi} = \frac{\sum_{t=1}^n y_t y_{t-1}}{\sum_{t=1}^n y_{t-1}^2}$$

et $\hat{\sigma}^2$ est donné par (3.3).

D'après (3.7), on a

$$\begin{aligned}
 -2 \ln \lambda &= -(n-1) \ln \left[1 - (\hat{\phi} - \phi)^2 \sum_{t=1}^n y_{t-1}^2 / \sum_{t=1}^n \varepsilon_t^2 \right] \\
 &\quad - (n-1) \ln \left[1 - (\hat{\phi}_1 - \phi_1)^2 \sum_{t=1}^n y_{t-1}^{+2} / \sum_{t=1}^n \varepsilon_t^2 - (\hat{\phi}_2 - \phi_2)^2 \sum_{t=1}^n y_{t-1}^{-2} / \sum_{t=1}^n \varepsilon_t^2 \right].
 \end{aligned} \tag{3.8}$$

Ainsi, d'après le théorème 3.1.2, sous H_0

$$[\sqrt{n}(\hat{\phi}_1 - \phi_1), \sqrt{n}(\hat{\phi}_2 - \phi_2)] \xrightarrow[n \rightarrow \infty]{\varphi} [Y_1, Y_2]$$

où, Y_1 et Y_2 sont des variables aléatoires *i.i.d.* et de loi $N(0, 2\sigma^2/E(y^2))$.

Cela implique que

$$\sqrt{n}(\hat{\phi} - \phi) = \sqrt{n}(\hat{\phi}_1 - \phi_1) \sum_{t=1}^n y_{t-1}^{+2} / \sum_{t=1}^n y_{t-1}^2 + \sqrt{n}(\hat{\phi}_2 - \phi_2) \sum_{t=1}^n y_{t-1}^{-2} / \sum_{t=1}^n y_{t-1}^2 \xrightarrow{\varphi} \frac{Y_1 + Y_2}{2}.$$

D'après (3.8), et en utilisant un développement de Taylor à l'ordre un, $-2 \ln \lambda$ est asymptotiquement équivalent, sous H_0 , à

$$n(\hat{\phi}_1 - \phi_1)^2 \sum_{t=1}^n y_{t-1}^{+2} / \sum_{t=1}^n \varepsilon_t^2 + n(\hat{\phi}_2 - \phi_2)^2 \sum_{t=1}^n y_{t-1}^{-2} / \sum_{t=1}^n \varepsilon_t^2 - n(\hat{\phi} - \phi)^2 \sum_{t=1}^n y_{t-1}^2 / \sum_{t=1}^n \varepsilon_t^2. \tag{3.9}$$

qui converge en loi vers $(Y_1 - Y_2)E(y^2)/4\sigma^2$ qui suit une χ^2 à un degré de liberté. Alors, $P(-2 \ln \lambda \leq x)$ converge en loi vers une χ_1^2 quand $n \rightarrow \infty$, sous l'hypothèse nulle.

3.3 Quelques simulations

Les simulations du modèle (2.1) sont effectuées pour déterminer les propriétés des estimateurs et le test d'hypothèse pour les échantillons de petite et grande tailles. Dans toutes les simulations, la distribution de $\{\varepsilon_t\}$ est $N(0, 1)$. Pour ϕ_1 et ϕ_2 dans les régions I et II, la nature transitoire du processus était explicite.

Des résultats de simulation sélectionnés sont montrés dans les tableaux ci-dessous; les valeurs sélectionnées pour les paramètres ϕ_1 et ϕ_2 sont dans la région d'ergodicité et la région III. Pour chaque couple ϕ_1, ϕ_2 1000 simulations ont été faites, d'abord

avec 100 observations et ensuite avec 1000 observations. Chaque échantillon généré a été utilisé pour calculer les estimateurs de ϕ_1 , ϕ_2 et σ^2 via (3.1)-(3.3) et $-2 \ln \lambda$ par (3.7). Les quantités calculées dans les tableaux sont :

$M(\widehat{\phi}_j)$: La moyenne des 1000 simulations pour $\widehat{\phi}_j$, $j = 1, 2$.

$$M(\widehat{\phi}_j) = (1000)^{-1} \sum_{i=1}^{1000} \widehat{\phi}_j^i, \quad j = 1, 2.$$

$SE(\widehat{\phi}_j)$: L'erreur standard des 1000 simulations pour $\widehat{\phi}_j$, $j = 1, 2$.

$$SE(\widehat{\phi}_j) = \left[(1000)^{-1} \sum_{i=1}^{1000} (\widehat{\phi}_j^i - M(\widehat{\phi}_j))^2 \right]^{1/2}, \quad j = 1, 2.$$

$MESE(\widehat{\phi}_j)$: Moyenne de l'erreur standard des 1000 simulations pour $\widehat{\phi}_j$, $j = 1, 2$.

$$MESE(\widehat{\phi}_j) = (1000)^{-1} \sum_{i=1}^{1000} ESE(\widehat{\phi}_j^i), \quad j = 1, 2, \text{ avec } ESE(\widehat{\phi}_1) = \widehat{\sigma}_i \left(\sum_{t=1}^n y_{t-1, i}^{+2} \right)^{-1/2}$$

$$ESE(\widehat{\phi}_2) = \widehat{\sigma}_i \left(\sum_{t=1}^n y_{t-1, i}^{-2} \right)^{-1/2}, \quad j = 1, 2.$$

$SESE(\widehat{\phi}_j)$: L'erreur standard de l'erreur standard des 1000 simulations pour $\widehat{\phi}_j$, $j = 1, 2$.

$$SESE(\widehat{\phi}_j) = \left[(1000)^{-1} \sum_{i=1}^{1000} (ESE(\widehat{\phi}_j^i) - MESE(\widehat{\phi}_j))^2 \right]^{1/2}, \quad j = 1, 2.$$

$M(\widehat{\sigma}^2)$: La moyenne des 1000 simulations pour $\widehat{\sigma}^2$.

$$M(\widehat{\sigma}^2) = (1000)^{-1} \sum_{i=1}^{1000} \widehat{\sigma}_i^2.$$

$SE(\widehat{\sigma}^2)$: L'erreur standard des 1000 simulations pour σ^2 .

$$SE(\widehat{\sigma}^2) = \left[(1000)^{-1} \sum_{i=1}^{1000} (\widehat{\sigma}_i^2 - M(\widehat{\sigma}^2))^2 \right]^{1/2}.$$

$P01(P05)$: La probabilité de rejeter H_0 , *i.e.* lorsque la quantité $-2 \ln \lambda$ dépasse la

valeur $\chi_{\hat{\phi}_1, 0.01}^2(\chi_{\hat{\phi}_1, 0.05}^2)$.

En observant les deux tableaux ci-dessous, nous pouvons tirer quelques propriétés:

1. En général, $\hat{\phi}_1$ et $\hat{\phi}_2$ exhibaient plus de performance quand ϕ_1 et ϕ_2 sont tous les deux négatifs, cela est, peut être, justifié, par l'occurrence des observations autant négatives que positives. Aussi, nous pouvons remarquer la consistance de $\hat{\phi}_1$ et $\hat{\phi}_2$, au niveau de $M(\cdot)$ et $SE(\cdot)$ quand le nombre d'observations augmente.
2. Dans les trois derniers cas dans les deux tableaux où ϕ_1 et ϕ_2 sont proches de l'hyperbole dans la région d'ergodicité et sur cette hyperbole, on remarque la performance de $\hat{\phi}_1$ et $\hat{\phi}_2$ même avec peu d'observations.
3. En moyenne, $MESE(\cdot)$ tend à sous estimer l'erreur standard correspondant $SE(\cdot)$.
4. $\hat{\sigma}^2$ montre une consistance dans l'estimation et semble in affecté par la performance des deux estimateurs $\hat{\phi}_1$ et $\hat{\phi}_2$.
5. Le test proposé a montré sa puissance de détecter l'égalité de deux paramètres, ce que l'on peut voir même pour une taille d'échantillon modérée.

Résultats des 1000 simulations pour la taille d'échantillon 100

ϕ_1	ϕ_2	$M(\hat{\phi}_1)$	$SE(\hat{\phi}_1)$	$M(\hat{\phi}_2)$	$SE(\hat{\phi}_2)$	$MESE(\hat{\phi}_1)$	$SESE(\hat{\phi}_1)$	$MESE(\hat{\phi}_2)$	$SESE(\hat{\phi}_2)$	$M(\hat{\sigma}^2)$	$SE(\hat{\sigma}^2)$	$P01$	$P05$
0.9	0.5	0.863	0.077	0.431	0.239	0.060	0.019	0.178	0.084	0.979	0.140	0.30	0.61
0.9	-0.5	0.878	0.057	-0.555	0.314	0.049	0.011	0.302	0.094	0.979	0.138	0.98	1.00
0.9	-10.0	0.899	0.022	-10.003	0.640	0.021	0.005	0.578	0.231	0.988	0.140	1.00	1.00
0.1	0.5	0.073	0.170	0.478	0.113	0.163	0.025	0.111	0.016	0.983	0.141	0.25	0.51
0.1	-0.5	0.093	0.120	-0.509	0.171	0.116	0.012	0.163	0.020	0.982	0.140	0.65	0.84
0.1	-10.0	0.101	0.021	-9.992	0.207	0.020	0.003	0.203	0.029	0.978	0.143	1.00	1.00
-0.1	0.5	-0.123	0.173	0.478	0.112	0.177	0.027	0.106	0.014	0.987	0.146	0.62	0.84
-0.1	-0.5	-0.111	0.124	-0.503	0.148	0.121	0.012	0.149	0.017	0.990	0.140	0.29	0.52
-0.1	-10.0	-0.097	0.006	-9.994	0.051	0.004	0.002	0.042	0.022	0.975	0.135	1.00	1.00
-0.9	0.5	-0.925	0.223	0.486	0.092	0.211	0.034	0.088	0.009	0.978	0.141	1.00	1.00
-0.9	-0.5	-0.896	0.115	-0.491	0.097	0.114	0.014	0.094	0.011	0.986	0.134	0.54	0.77
0.9	0.9	0.824	0.174	0.833	0.152	0.087	0.066	0.084	0.073	0.976	0.145	0.02	0.08
0.5	0.5	0.469	0.140	0.466	0.134	0.125	0.023	0.125	0.022	0.979	0.136	0.01	0.06
-0.5	-0.5	-0.496	0.127	-0.507	0.123	0.121	0.013	0.122	0.013	0.983	0.136	0.01	0.06
-0.9	-0.9	-0.883	0.071	-0.881	0.070	0.066	0.013	0.066	0.013	0.982	0.139	0.01	0.05
-0.2	-5.0	-0.193	0.013	-4.994	0.048	0.009	0.004	0.042	0.022	0.984	0.141	1.00	1.00
-0.2	-4.9	-0.193	0.014	-4.892	0.057	0.010	0.004	0.048	0.021	0.981	0.147	1.00	1.00
-0.19	-5.0	-0.183	0.014	-4.994	0.065	0.012	0.004	0.057	0.021	0.974	0.143	1.00	1.00

Résultats des 1000 simulations pour la taille d'échantillon 1000

ϕ_1	ϕ_2	$M(\hat{\phi}_1)$	$SE(\hat{\phi}_1)$	$M(\hat{\phi}_2)$	$SE(\hat{\phi}_2)$	$MESE(\hat{\phi}_1)$	$MESE(\hat{\phi}_2)$	$SESE(\hat{\phi}_1)$	$SESE(\hat{\phi}_2)$	$M(\hat{\sigma}^2)$	$SE(\hat{\sigma}^2)$	$P01$	$P05$
0.9	0.5	0.896	0.017	0.494	0.051	0.017	0.052	0.002	0.052	0.005	0.044	1.00	1.00
0.9	-0.5	0.898	0.014	-0.505	0.089	0.015	0.092	0.001	0.092	0.008	0.046	1.00	1.00
0.9	-10.0	0.900	0.007	-10.002	0.181	0.007	0.173	0.001	0.173	0.018	0.043	1.00	1.00
0.1	0.5	0.098	0.052	0.498	0.034	0.051	0.035	0.002	0.035	0.002	0.046	1.00	1.00
0.1	-0.5	0.099	0.038	-0.501	0.050	0.037	0.052	0.001	0.052	0.002	0.044	1.00	1.00
0.1	-10.0	0.100	0.006	-10.000	0.064	0.006	0.064	0.000	0.064	0.003	0.047	1.00	1.00
-0.1	0.5	-0.101	0.057	0.497	0.031	0.055	0.033	0.002	0.033	0.001	0.047	1.00	1.00
-0.1	-0.5	-0.099	0.039	-0.501	0.045	0.038	0.047	0.001	0.047	0.002	0.044	1.00	1.00
-0.1	-10.0	-0.100	0.001	-10.000	0.005	0.000	0.004	0.000	0.004	0.002	0.045	1.00	1.00
-0.9	0.5	-0.902	0.066	0.499	0.027	0.066	0.028	0.003	0.028	0.001	0.044	1.00	1.00
-0.9	-0.5	-0.900	0.037	-0.500	0.029	0.036	0.030	0.001	0.030	0.001	0.045	1.00	1.00
0.9	0.9	0.895	0.021	0.896	0.022	0.020	0.020	0.003	0.020	0.003	0.044	0.01	0.05
0.5	0.5	0.497	0.039	0.496	0.039	0.039	0.039	0.002	0.039	0.002	0.046	0.01	0.04
-0.5	-0.5	-0.499	0.038	-0.498	0.039	0.039	0.039	0.001	0.039	0.001	0.045	0.01	0.06
-0.9	-0.9	-0.899	0.020	-0.898	0.020	0.020	0.020	0.001	0.020	0.001	0.044	0.01	0.06
-0.2	-5.0	-0.199	0.001	-4.999	0.005	0.001	0.004	0.000	0.004	0.002	0.045	1.00	1.00
-0.2	-4.9	-1.999	0.002	-4.899	0.010	0.002	0.010	0.000	0.010	0.002	0.045	1.00	1.00
-0.19	-5.0	-0.189	0.003	-4.999	0.015	0.003	0.014	0.000	0.014	0.002	0.046	1.00	1.00

3.4 Modèle TAR(1) multiple

Dans ce qui suit, on suppose que

A1 : le processus $\{y_t\}$ est ergodique et sa distribution stationnaire possède un moment d'ordre deux fini.

A2 : $E(\varepsilon_t^{(k)}) = \sigma_{(k)}^2 < \infty$, $1 \leq k \leq l$.

On note que si $\phi_1^{(1)} < 1$, $\phi_1^{(l)} < 1$, $\phi_1^{(1)}\phi_1^{(l)} < 1$ et $\sigma_{(k)}^2 < \infty$, $1 \leq k \leq l$, alors, d'après le théorème 2.1.4, la distribution stationnaire de $\{y_t\}$ possède un moment d'ordre deux fini. Soit $J(k)$ l'ensemble des entiers naturels tels que $\{0 \leq t \leq n-1 : y_t \in \mathbb{R}_k\}$ et $n(k)$ sa cardinalité pour $1 \leq k \leq l$.

Supposons que les seuils r_k , $1 \leq k \leq l$, sont connus, les estimateurs des moindres carrés pour les paramètres $\{\phi_i^{(k)}\}$ sont donnés par

$$\widehat{\phi}_1^{(k)} = \left[\sum_{t \in J(k)} y_t y_{t+1} - \sum_{t \in J(k)} y_t \sum_{t \in J(k)} y_{t+1} / n(k) \right] / nS_{(k)}^2 \quad (3.10)$$

$$\widehat{\phi}_0^{(k)} = n(k)^{-1} \left[\sum_{t \in J(k)} y_{t+1} - \widehat{\phi}_1^{(k)} \sum_{t \in J(k)} y_t \right] \quad (3.11)$$

avec

$$nS_{(k)}^2 = \sum_{t \in J(k)} y_t^2 - \left(\sum_{t \in J(k)} y_t \right)^2 / n(k)$$

et l'estimateur naturel correspondant à $\sigma_{(k)}^2$ est donné par

$$\widehat{\sigma}_{(k)}^2 = n(k)^{-1} \sum_{t \in J(k)} \left(y_{t+1} - \widehat{\phi}_0^{(k)} - \widehat{\phi}_1^{(k)} y_t \right)^2. \quad (3.12)$$

Les deux théorèmes suivants établissent la forte consistance et la normalité asymptotique des estimateurs (3.10)-(3.12) quand le processus $\{y_t\}$ est ergodique.

Théorème 3.4.1 (Chan et Tong (1985)) *Sous les hypothèses A1 et A2, $\widehat{\phi}_i^{(k)}$ et $\widehat{\sigma}_{(k)}^2$, $i = 0, 1$; $1 \leq k \leq l$, sont des estimateurs fortement consistant pour $\phi_i^{(k)}$ et $\sigma_{(k)}^2$, $i = 0, 1$; $1 \leq k \leq l$.*

Preuve. De façon analogue à la preuve de la consistance forte du modèle TAR(1)

simple, on réécrit les estimateurs (3.10) et (3.11) comme

$$\begin{aligned}\widehat{\phi}_1^{(k)} &= \phi_1^{(k)} + \left[\sum_{t \in J(k)} y_t \varepsilon_{t+1}^{(k)} - \sum_{t \in J(k)} y_t \sum_{t \in J(k)} \varepsilon_{t+1}^{(k)} / n(k) \right] / nS_{(k)}^2 \\ \widehat{\phi}_0^{(k)} &= \phi_0^{(k)} + n(k)^{-1} \left[\sum_{t \in J(k)} y_{t+1}^2 \sum_{t \in J(k)} \varepsilon_{t+1}^{(k)} - \sum_{t \in J(k)} y_t \sum_{t \in J(k)} y_t \varepsilon_{t+1}^{(k)} \right] / nS_{(k)}^2\end{aligned}$$

comme pour la preuve du théorème 3.1.1, on peut montrer que

$$\widehat{\phi}_i^{(k)} \longrightarrow \phi_i^{(k)} \quad p.s. \text{ quand } n \longrightarrow \infty, \quad i = 0, 1; \quad 1 \leq k \leq l.$$

La même chose pour $\widehat{\sigma}_{(k)}^2$, réécrivons (3.12) comme

$$\begin{aligned}\widehat{\sigma}_{(k)}^2 &= n(k)^{-1} \sum_{t \in J(k)} \{ (\varepsilon_{t+1}^{(k)})^2 + (\widehat{\phi}_0^{(k)} - \phi_0^{(k)})^2 + (\widehat{\phi}_1^{(k)} - \phi_1^{(k)})^2 y_t^2 \\ &\quad - 2\varepsilon_{t+1}^{(k)} (\widehat{\phi}_0^{(k)} - \phi_0^{(k)}) - 2\varepsilon_{t+1}^{(k)} (\widehat{\phi}_1^{(k)} - \phi_1^{(k)}) y_t + 2(\widehat{\phi}_0^{(k)} - \phi_0^{(k)}) (\widehat{\phi}_1^{(k)} - \phi_1^{(k)}) y_t \}\end{aligned}$$

par application du résultat précédent, il est clair que

$$\widehat{\sigma}_{(k)}^2 \longrightarrow \sigma_{(k)}^2 \quad p.s. \text{ quand } n \longrightarrow \infty, \quad 1 \leq k \leq l.$$

■

pour le théorème suivant, nous utilisons les notations suivantes

$$\gamma^{(k)} = E((y_t^{(k)})^2) - E^2(y_t^{(k)}) / \pi(\mathbb{R}_k), \quad 1 \leq k \leq l.$$

avec $n(k) \longrightarrow \pi(\mathbb{R}_k)$ quand $n \longrightarrow \infty$. Pour $n = 1, 2, \dots$, soit $\Phi(n)$ un vecteur de taille $2l \times 1$, dont ses $(2k-1)^{\text{ème}}$ et $(2k)^{\text{ème}}$ composantes sont

$$\begin{aligned}[n\gamma^{(k)}\pi(\mathbb{R}_k)/\sigma_{(k)}^2 E((y_t^{(k)})^2)]^{1/2} (\widehat{\phi}_0^{(k)} - \phi_0^{(k)}) \\ [n\gamma^{(k)}/\sigma_{(k)}^2]^{1/2} (\widehat{\phi}_1^{(k)} - \phi_1^{(k)}).\end{aligned}$$

respectivement pour $k = 1, \dots, l$. Soit D la matrice par blocs de taille $2l \times 2l$, avec

les matrices $D_{2 \times 2}^{(k)}$ sur la diagonale et des zéros ailleurs, avec

$$D^{(k)} = \begin{bmatrix} 1 & -E(y_t^{(k)})/[E((y_t^{(k)})^2)\pi(\mathbb{R}_k)]^{1/2} \\ -E(y_t^{(k)})/[E((y_t^{(k)})^2)\pi(\mathbb{R}_k)]^{1/2} & 1 \end{bmatrix}.$$

Théorème 3.4.2 *Sous les hypothèses A1 et A2, quand $n \rightarrow \infty$, $\Phi(n)$ converge en loi vers une loi normale multidimensionnelle $N(\mathbf{0}, D)$.*

3.5 Estimation du maximum vraisemblance d'un modèle autorégressif à seuils

Pour un modèle SETAR à un seul paramètre threshold, Chan (1993) a obtenu la consistance et la distribution asymptotique de l'estimateur des moindres carrés pour la vraie valeur du paramètre inconnu. Dans cette section, on dérive des propriétés similaires pour les estimateurs du maximum vraisemblance du même modèle SETAR(2; p, p) sous des conditions de régularité sur la densité des erreurs, non nécessairement gaussienne. On définit un modèle SETAR(2; p, p) par :

$$y_t = h(\mathbf{y}_{t-1}, \theta) + \varepsilon_t, \quad t \geq 1, \quad (3.13)$$

pour $\theta = (\theta'_1, \theta'_2, r, d)' \in \mathbb{R}^{2p+3} \times \{1, 2, \dots, p\}$, où $\mathbf{y}_{t-1} = (y_{t-1}, y_{t-2}, \dots, y_{t-p})'$, $\theta_j = (\phi_0^{(j)}, \phi_1^{(j)}, \dots, \phi_p^{(j)})' \in \mathbb{R}^{p+1}$, $j = 1, 2$, et pour $\mathbf{x} \in \mathbb{R}^p$,

$$h(\mathbf{x}, \theta) = (\phi_0^{(1)} + \sum_{i=1}^p \phi_i^{(1)} x_i) \mathbf{1}(x_d \leq r) + (\phi_0^{(2)} + \sum_{i=1}^p \phi_i^{(2)} x_i) \mathbf{1}(x_d > r).$$

Les erreurs $\{\varepsilon_t\}$ dans (3.13) sont des variables aléatoires *i.i.d* de moyenne nulle et de variance constante. Dans ce qui suit, on suppose que la série définie par le modèle (3.13) est stationnaire et ergodique. Pour le cas d'un modèle de threshold à erreurs gaussiennes, Tong (1983) a construit un estimateur donné sur la méthode du maximum vraisemblance pour les paramètres inconnus en utilisant le critère d'information d'Akaike (Akaike (1973)). Pour r inconnu, Chan et al. (1985) ont obtenu la normalité asymptotique et la consistance des estimateurs des moindres carrés pour le paramètre $\theta_c = (\theta'_1, \theta'_2)'$ sous des conditions de régularité et pour

$p = 1$. Cependant, en pratique, le paramètre threshold r est inconnu et il prend un nombre infini de valeurs dans \mathbb{R} . Dans ce cas là, Petrucci (1986) a prouvé que l'estimateur des moindres carrés de θ est fortement consistant pour un modèle SETAR(2; 1, 1). Contrairement au modèle AR linéaire, la fonction de vraisemblance n'est pas continue en r en général. Ainsi, la méthode ordinaire pour le calcul de l'estimateur du maximum vraisemblance ne peut pas être adoptée.

Dans la section suivante, Qian (1998) propose une méthode d'estimation fondée sur le maximum vraisemblance pour le paramètre $\theta = (\theta'_c, r, d)'$. Ensuite, cet auteur donne des résultats sur la consistance forte et la n -consistance et la normalité asymptotique de $\hat{\theta}_n = (\hat{\theta}'_{cn}, \hat{r}_n, \hat{d}_n)'$ estimateur du vrai paramètre θ .

3.5.1 L'estimation par la méthode du maximum vraisemblance

On commence par la définition des estimateurs du maximum vraisemblance des paramètres inconnus sous-jacent θ dans le modèle (3.13). On suppose que le vrai paramètre θ est un point intérieur de l'espace des paramètres $\mathbb{R}^{2p+2} \times \bar{\mathbb{R}} \times \{1, 2, \dots, p\}$ ($\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$). Il existe un sous ensemble K de \mathbb{R}^{2p+2} tel que θ soit un point intérieur dans $K \times \bar{\mathbb{R}} \times \{1, 2, \dots, p\}$.

Notons $\Omega = K \times \bar{\mathbb{R}} \times \{1, 2, \dots, p\}$, alors Ω est un ensemble compact. Soit $\vartheta = (\alpha', \beta', s, q)'$ un point de Ω . Soit $\mathbf{y}_t = (y_t, \dots, y_{t-p+1})$, ainsi $\{\mathbf{y}_t\}$ est une chaîne de Markov. Soit $g_\vartheta(\mathbf{y}_0)$ la densité initiale de \mathbf{y}_0 sous ϑ , f est la fonction densité de ε_t , alors la densité de transition en une étape, partant de \mathbf{y}_0 , est $f(y_t - h(\mathbf{y}_{t-1}, \vartheta))$, $t \geq 1$. Si on observe $(\mathbf{y}_0, \dots, \mathbf{y}_n)$, alors, la fonction de vraisemblance sous ϑ est $\prod_{t=1}^n f(y_t - h(\mathbf{y}_{t-1}, \vartheta))g_\vartheta(\mathbf{y}_0)$. Soit $\hat{\theta}_n = (\hat{\theta}'_{cn}, \hat{r}_n, \hat{d}_n)'$ une fonction mesurable de $(\mathbf{y}_0, \dots, \mathbf{y}_n)$ de \mathbb{R}^{n+p+1} vers Ω telle que $\hat{\theta}_n$ maximise la fonction de vraisemblance conditionnelle

$$L_n(\vartheta) := \prod_{t=1}^n f(y_t - h(\mathbf{y}_{t-1}, \vartheta)) \quad \text{sur } \Omega.$$

Notons $\vartheta = (\vartheta'_c, s, q)'$, $\theta = (\theta'_c, r, d)'$. L'algorithme de maximisation est donné par les étapes suivantes :

Etape 1. Pour s fixé dans $\bar{\mathbb{R}}$, $q \in \{1, 2, \dots, p\}$, notons $L_{nsq}(\vartheta_c) = L_n(\vartheta_c, s, q) = L_n(\vartheta)$.

Soit $\vartheta_{cn}(s, q) \in K$ tel que

$$\vartheta_{cn}(s, q) = \arg \max_{\vartheta_c \in K} L_{nsq}(\vartheta_c).$$

Etape 2. Considérons ensuite la fonctions vraisemblance $(s, q) \rightarrow L_n(\vartheta_{cn}(s, q), s, q)$. Notons que $L_n(\vartheta_{cn}(s, q), s, q)$ a un nombre fini de valeurs possibles. Soit \hat{r}_n, \hat{d}_n les plus petites valeurs satisfaisant :

$$L_n(\vartheta_{cn}(\hat{r}_n, \hat{d}_n), \hat{r}_n, \hat{d}_n) = \max_{s \in \mathbb{R}, q \in \{1, \dots, p\}} L_n(\vartheta_{cn}(s, q), s, q),$$

on remplace ensuite \hat{r}_n, \hat{d}_n dans $\vartheta_{cn}(s, q)$ pour avoir $\hat{\theta}_{cn} = \vartheta_{cn}(\hat{r}_n, \hat{d}_n)$. Ainsi, $\hat{\theta}_n = (\hat{\theta}'_{cn}, \hat{r}_n, \hat{d}_n)'$ est un estimateur du maximum vraisemblance de θ . Ainsi, d'après la définition de $\hat{\theta}_{cn}, \hat{r}_n$ et \hat{d}_n , on a

$$L_n(\hat{\theta}'_{cn}, \hat{r}_n, \hat{d}_n) = L_n(\vartheta_{cn}(\hat{r}_n, \hat{d}_n), \hat{r}_n, \hat{d}_n) \geq L_n(\vartheta_{cn}(s, q), s, q) \geq L_n(\vartheta)$$

alors

$$L_n(\hat{\theta}'_{cn}, \hat{r}_n, \hat{d}_n) = \sup_{\vartheta \in \Omega} L_n(\vartheta).$$

Cela signifie que $\hat{\theta}_n = (\hat{\theta}'_{cn}, \hat{r}_n, \hat{d}_n)'$ est un MLE de θ .

3.5.2 Hypothèses et consistance forte

On commence cette partie par les hypothèses sur la densité f des erreurs ε_t et le processus sous-jacent.

(C₁) f est absolument continue et strictement positive sur tout \mathbb{R} . de \dot{f} , soit $\varphi = \dot{f}/f$ et $I(f) = \int \varphi^2(x)f(x)dx < \infty$.

(C₂) φ est Lip(1).

(C₃) φ est différentiable et la dérivée $\dot{\varphi}$ est Lip(1).

(C₄) $E|\varepsilon_t|^4 < \infty$.

Pour vérifier la n -consistance et la distribution limite de l'estimateur du threshold, nous avons besoin de considérer les hypothèses suivantes :

(M₁) Le point $r \in \mathbb{R}$ est un point de discontinuité pour h , i.e. il existe $X^* = (1, y_0, \dots, y_{1-p})'$ tel que $(\theta_1 - \theta_2)'X^* \neq 0$, $y_{1-d} = r$.

(M₂) $\{y_t\}$ admet une unique mesure π telle que $\exists K, \rho < 1, \forall x \in \mathbb{R}^p, \forall k \geq 1, \|P^k(x, \cdot) - \pi(\cdot)\| \leq K\rho^k(1+|x|)$, avec $\|\cdot\|$ et $|\cdot|$ représentent la norme de la variation totale et la norme Euclidienne respectivement.

Remarques 3.5.1 *Les densités de la loi normale, de la loi logistique et la t-densité (Student) à m d.l.l, $m > 4$ satisfassent les conditions (C₁) – (C₄). De l'équation invariante $g_\theta(x) = \int f(x - h(y, \theta))g_\theta(y)dy, x \in \mathbb{R}^p$, la condition (C₁) implique que g_θ est bornée sur les compacts. La finitude de $I(f)$ implique la finitude de f (Koul, 1992, p. 52). Lip(1) et la différentiabilité de φ impliquent la bornitude de $\dot{\varphi}$.*

Maintenant, on donne le résultat important suivant.

Théorème 3.5.1 (Qian (1999)) *Supposons que $\{y_t\}$ dans le modèle (3.13) est stationnaire et ergodique, si les conditions (C₁) et (C₂) sont vérifiées. Alors*

$$\hat{\theta}_n \longrightarrow \theta \quad p.s \text{ quand } n \longrightarrow \infty \text{ (sous } \theta \text{)}.$$

[Voir L. Qian (1998)]

3.5.3 n -consistance de l'estimateur du threshold

La discontinuité de h en r va nous donner un résultat plus fort pour l'estimateur \hat{r}_n . i.e. qui est la n -consistance de \hat{r}_n .

Théorème 3.5.2 (Qian (1998)) *Supposons que les conditions (C₁)–(C₄), (M₁) et (M₂) sont vérifiées. Alors*

$$\lim_{n \rightarrow \infty} |n(\hat{r}_n - r)| = 0.$$

[Voir L. Quian (1998)]

Chapitre 4 : Application et simulations

4 Application et simulations

4.1 Introduction

Le modèle autorégressif à seuils proposé par Tong (1987) est un des modèles non linéaires qui a la particularité de pouvoir reproduire les phénomènes de cycles limites, d'irréversibilité temporelle, de dépendance amplitude-fréquence et du phénomène de saut (Voir Tong et Lim (1980)). La plupart des motivations qui existent pour ce modèle concernent les cycles limites des séries temporelles cycliques et, en effet, ce modèle est capable de reproduire des cycles limites asymétriques. Cependant, les modèles autorégressifs à seuils n'ont pas reçu beaucoup d'attention dans les applications, jusqu'à Tsay (1989). Cela est dû à (a) l'absence d'une procédure de modélisation (b) l'incapacité d'identifier la variable threshold et d'estimer ses valeurs. On présente dans cette partie une procédure, celle de Tsay (1989) pour la construction d'un modèle autorégressif à seuils. Une statistique du test proposé est présentée pour tester la non linéarité et pour spécifier la variable threshold basée sur les résidus prédictifs, des outils graphiques supplémentaires pour identifier le nombre et localiser les thresholds potentiels. Finalement, cette statistique est utilisée pour construire le modèle à seuils. La statistique et ses propriétés sont dérivées grâce à une régression linéaire récursive arrangée; on évalue ses performances pour de petits échantillons en faisant une étude de simulation et pour des séries réelles connues dans la littérature.

Rappelons qu'un processus $\{y_t\}$ est dit suivre un modèle autorégressif à seuils, noté TAR($l; p, d$) s'il est solution de l'équation stochastique suivante

$$y_t = \phi_0^{(j)} + \sum_{i=1}^p \phi_i^{(j)} y_{t-i} + \varepsilon_t^{(j)}, \quad r_{j-1} \leq y_{t-d} < r_j \quad (4.1)$$

où $j = 1, \dots, l$ et $d \in \mathbb{N}^*$, $-\infty = r_0 < r_1 < \dots < r_l = \infty$, $\forall j$, $\{\varepsilon_t^{(j)}\}$ est une suite de

différences de martingale satisfaisant

$$\begin{aligned} E(\varepsilon_t^{(j)} \mid \mathcal{F}_{t-1}) &= 0, \\ \sup_t E(|\varepsilon_t^{(j)}|^\delta \mid \mathcal{F}_{t-1}) &< \infty \text{ p.s pour } \delta > 2, \end{aligned} \quad (4.2)$$

où \mathcal{F}_{t-1} est la σ -algèbre engendrée par $\{\varepsilon_{t-i}^{(j)} \mid i = 1, 2, \dots; j = 1, \dots, l\}$. Un tel processus partitionne l'espace Euclidien unidimensionnel en l régimes et suit dans chacun de ces régime un modèle autorégressif. Ce modèle a été proposé comme une alternative à des modèles linéaires pour décrire les phénomènes mentionnés ci-dessus. Néanmoins, le modèle TAR présente des difficultés techniques. Tong et Lim (1980) on proposé une procédure de test et d'estimation, mais celle ci nécessite des étapes de calcul intensif, et ils ne disposaient pas d'une statistique pour tester le besoin de recourir à un modèle à seuils pour une série de données d'observations. Alors, Tsay a proposé un test de non linéarité et une procédure de modélisation en quatres étapes. Ces étapes peuvent être utilisées de façon itérative si le nombre de régimes est grand ou le degré de non linéarité est faible.

4.2 Test pour une non linéarité de type threshold

Le test proposé est celui donné par Tsay (1989), qui est relié au test portmanteau de non linéarité de Petruccilli et Davies (1986), du fait qu'il est ainsi basé sur la régression arrangée et les résidus prédictifs. Néanmoins, les deux tests sont différents dans le sens où les résidus prédictifs sont exploités. Le test proposé est une combinaison des tests de non linéarité de Keenan (1985), Tsay (1986) et de Petruccilli et Davies (1986).

4.2.1 Autorégression arrangée et résidus prédictifs

Une autorégression pour un $AR(p)$ avec n observations est donné par

$$y_t = (1 \ y_{t-1} \ \dots \ y_{t-p})\beta + \varepsilon_t \quad t = p + 1, \dots, n$$

où β est le vecteur des coefficients de dimension $(p+1)$, ε_t est le bruit. $(y_t, 1, y_{t-1}, \dots, y_{t-p})$ est appelé un *cas* pour le modèle $AR(p)$. Alors, l'autorégression arrangée est une

autorégression avec des cas arrangés, basée sur un régresseur particulier. Pour le modèle TAR dans (1.1), l'autorégression arrangée devient utile si elle est arrangée par rapport à la variable threshold. Pour illustration, on considère les cas $l = 2$, pour un TAR(2; p, d) à n observations, les valeurs possibles pour la variable threshold sont $\{y_h, \dots, y_{n-d}\}$, où $h = \max\{1, p+1-d\}$. Soit π_i l'indice de temps pour la $i^{\text{ème}}$ plus petite observation de $\{y_h, \dots, y_{n-d}\}$. Réécrivons le modèle comme suit

$$y_{\pi_i+d} = \begin{cases} \phi_0^{(1)} + \sum_{\nu=1}^p \phi_\nu^{(1)} y_{\pi_i+d-\nu} + \varepsilon_{\pi_i+d}^{(1)} & \text{si } i \leq s \\ \phi_0^{(2)} + \sum_{\nu=1}^p \phi_\nu^{(2)} y_{\pi_i+d-\nu} + \varepsilon_{\pi_i+d}^{(2)} & \text{si } i > s \end{cases} \quad (4.3)$$

où s satisfait $y_{\pi_s} < r_1 \leq y_{\pi_{s+1}}$. Cette autorégression est utile pour le modèle TAR, parce qu'elle sépare effectivement les deux régimes; la séparation ne nécessite pas la connaissance de la valeur threshold r_1 , seulement le nombre d'observations dans chaque régime dépend de r_1 , et si on considère que r_1 est connue, les estimateurs des paramètres sont facilement obtenus. Comme la valeur de threshold est inconnue, cependant, on peut procéder séquentiellement, les estimateurs des moindres carrés $\hat{\phi}_\nu^{(1)}$ sont consistants pour $\phi_\nu^{(1)}$ s'il y a un nombre suffisamment grand d'observations dans le premier régime. Dans ce cas là, les résidus prédictifs sont asymptotiquement un bruit blanc et sont orthogonaux aux régresseurs $\{y_{\pi_i+d-\nu} \mid \nu = 1, \dots, p\}$. Quand i atteint s ou le dépasse, le résidu prédictif pour l'observation $y_{\pi_{s+1}+d}$ est biaisé parce que le modèle change à l'instant $\pi_{s+1} + d$. Ainsi, l'orthogonalité entre les résidus prédictifs et les régresseurs n'existe plus dès que l'autorégression récursive atteint les observations du deuxième régime et les estimateurs $\hat{\phi}_\nu^{(1)}$ dévient vers $\phi_\nu^{(2)}$, ceux du deuxième régime car on utilise les observations générés par ce dernier (Voir figure ci-dessus). Notons que la connaissance de la vraie valeur de r_1 n'est pas exigée; on demande seulement l'existence d'un threshold non trivial. Pour l'autorégression arrangée (4.3), soit $\hat{\beta}_m$ l'estimateur des moindres carrés pour le vecteur des paramètres β basé sur les m premiers cas, P_m la matrice inverse de $X'X$ associé, et x_{m+1} le vecteur des régresseurs pour l'observation à introduire dans l'autorégression. Alors, l'estimateur des moindres carrés se calcul à l'aide de

$$\begin{aligned}\widehat{\beta}_{m+1} &= \widehat{\beta}_m + K_{m+1}[y_{\pi_{m+1+d}} - x'_{m+1}\widehat{\beta}_m] \\ D_{m+1} &= 1 + x'_{m+1}P_mx'_{m+1} \\ K_{m+1} &= P_mx_{m+1}/D_{m+1}\end{aligned}$$

où

$$P_{m+1} = \left(I - P_m \frac{x_{m+1}x'_{m+1}}{D_{m+1}} \right) P_m$$

(Voir Ertel et Fowlkes 1976; Goodwin et Payne 1977), et les résidus prédictifs et les résidus prédictifs standardisé sont donnés par

$$\widehat{\varepsilon}_{\pi_{m+1+d}} = y_{\pi_{m+1+d}} - x'_{m+1}\widehat{\beta}_m \quad (4.4)$$

et

$$\widehat{e}_{\pi_{m+1+d}} = \widehat{\varepsilon}_{\pi_{m+1+d}} / \sqrt{D_{m+1}}. \quad (4.5)$$

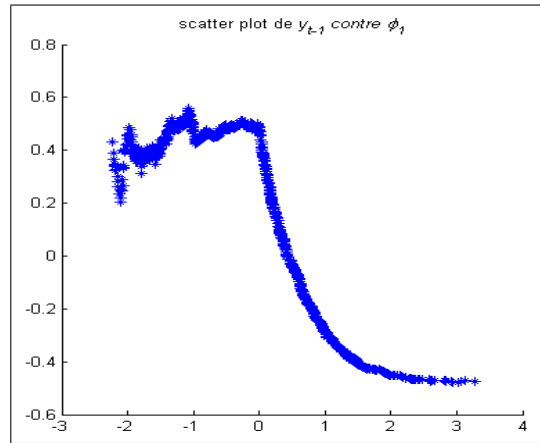


Figure 4.1. Scatter plot de y_{t-1} contre $\widehat{\phi}_1$ pour l'autorégression arrangée du

$$\text{modèle } y_t = \begin{cases} 1 + .5y_{t-1} & \text{si } y_{t-1} \leq 0 \\ -1 - .5y_{t-1} & \text{si } y_{t-1} > 0 \end{cases}$$

4.2.2 Test de non linéarité

Maintenant, on donne le test de non linéarité proposé. Pour p et d fixés, le nombre effectif d'observations utilisé dans l'autorégression arrangée est $n - d - h + 1$. Sup-

posons que l'autorégression commence avec b observations, alors on a $n-d-b-h+1$ résidus prédictifs. Faisons l'autorégression suivante

$$\widehat{e}_{\pi_i+d} = \omega_0 + \sum_{\nu=1}^p \omega_\nu y_{\pi_i+d-\nu} + a_{\pi_i+d}, \quad \text{pour } i = b+1, \dots, n-d-h+1, \quad (4.6)$$

et on calcule la statistique F associée

$$\widehat{F}(p, d) = \frac{(\sum \widehat{e}_t^2 - \sum \widehat{a}_t^2)/(p+1)}{\sum \widehat{a}_t^2/(n-d-b-p-h)}, \quad (4.7)$$

où la sommation se fait sur toutes les observations dans (4.6) et \widehat{a}_t sont les résidus des moindres carrés de (4.6). L'argument (p, d) dans \widehat{F} est utilisé pour montrer la dépendance du ratio F en p et d .

Théorème 4.2.1 *Supposons que y_t est un processus AR linéaire stationnaire d'ordre p . Alors y_t suit le modèle (4.1) avec $l = 1$. Pour n grand, la statistique $\widehat{F}(p, d)$ définie par (4.7) suit asymptotiquement la distribution de Fisher F à $p+1$ et $n-d-b-p-h$ d.l.l, de plus $(p+1)\widehat{F}(p, d)$ est asymptotiquement une χ^2 à $(p+1)$ d.l.*

Ce théorème peut être prouvé en utilisant une technique pareille à celle de Tsay (1986, théorème 1) et Keenan (1985, Lemme 3.1). Ils utilisent la propriété de consistance des estimateurs de moindres carrés d'un modèle AR linéaire et le théorème centrale limite des martingales de Bilingsley (1961). Notons que la distribution asymptotique de la statistique $\widehat{F}(p, d)$ reste vraie si les résidus prédictifs standardisés \widehat{e}_t sont remplacés par les résidus prédictifs ordinaires $\widehat{\varepsilon}_t$ de (4.4). Néanmoins, les résidus prédictifs standardisés semblent d'être préférables quand la taille d'échantillon est courte. Pour les grandes tailles d'échantillons, les résidus prédictifs ordinaires peuvent économiser le temps de calcul.

4.2.3 Puissance du test

Pour tester la puissance de test on teste d'abord la puissance de $\widehat{F}(p, d)$ pour la détection de la non linéarité à 1% et 5% pour 1000 réalisations d'un TAR(2; 1, 1) avec $(\phi_0^{(1)}, \phi_1^{(1)}, \phi_0^{(2)}, r_1, \sigma_1^2, \sigma_2^2) = (1, 0.5, 1, 1, 1, 1)$ et $\phi_1^{(2)}$ est donné dans la table ci-dessous. Les tailles d'échantillons utilisées sont 50, 100 et 1000, pour

chaque simulation d'un échantillon on génère $n + 200$ observations et on garde les n dernières réalisations pour réduire l'effet de la valeur initiale. Dans le test, $p = 1$ et $d = 1$ et $b = (n/10) + p$. il est clair de la table ci-dessous que la statistique \widehat{F} est très puissante pour la détection de non linéarité sauf dans le cas de $\phi_1^{(2)} = 0$ pour les tailles 50 et 100. Pour le modèle linéaire, i.e. pour $\phi_1^{(2)} = 0.5$, l'erreur de 1^{er} espèce n'est pas grande.

Table 1. Fréquence de rejet du modèle linéaire basé sur 1000 réalisation d'un TAR(2; 1, 1)

$\phi_1^{(2)}$	$n = 50$		$n = 100$		$n = 1000$	
	1%	5%	1%	5%	1%	5%
-2	1000	1000	1000	1000	1000	1000
-1	894	972	1000	1000	1000	1000
-0.5	351	654	866	976	1000	1000
0	28	161	92	306	991	1000
0.5	9	62	3	38	8	39

Les tableaux suivants reproduisent les simulations de la table précédente en changeant seulement les termes constants $\phi_0^{(j)}$ et la valeur du seuil, et en prenant $\phi_0^{(j)} = 0$ pour $j = 1, 2$ et $r_1 = 0$ pour la table 2 et $\phi_0^{(1)} = 1$, $\phi_0^{(2)} = -1$ et $r_1 = 0$ pour la table 3. Du fait que le terme constant est lié au niveau du processus, il est intéressant de voir son effet sur le test.

Table 2. Fréquence de rejet du modèle linéaire basé sur 1000 réalisation d'un TAR(2; 1, 1)

$\phi_1^{(2)}$	$n = 50$		$n = 100$		$n = 1000$	
	1%	5%	1%	5%	1%	5%
-2	446	694	942	986	1000	1000
-1	147	349	577	792	1000	1000
-0.5	46	175	264	484	1000	1000
0	13	64	51	196	915	971
0.5	16	62	19	120	31	132

Table 3. Fréquence de rejet du modèle linéaire basé
sur 1000 réalisation d'un TAR(2; 1, 1)

$\phi_1^{(2)}$	$n = 50$		$n = 100$		$n = 1000$	
	1%	5%	1%	5%	1%	5%
-2	1000	1000	1000	1000	1000	1000
-1	952	992	1000	1000	1000	1000
-0.5	552	830	974	992	1000	1000
0	83	261	391	716	1000	1000
0.5	4	56	39	184	1000	1000

Nous remarquons des tableaux 2 et 3 quand on les compare aux résultats de la table 1 que le changement en niveau influe sur la puissance du test (mais celui ci reste efficace) spécialement dans le cas de petits échantillons, ce qui n'apparait pas pour le cas de $n = 1000$ observations.

4.3 Spécification de la variable threshold

4.3.1 Sélection du paramètre de délai

La difficulté majeure dans la modélisation d'un modèle TAR est la spécification de la variable threshold qui joue un rôle clé dans la nature de ce modèle non linéaire. Tong et Lim (1980) ont utilisé le critère d'information d'Akaike (Akaike 1974) pour la sélection de d après avoir choisi tous les autres paramètres. Tsay propose une procédure qui donne d avant d'avoir localisé les valeurs des seuils, cette méthode est motivée par la performance de la statistique F , pour p fixé; l'estimateur du paramètre de délai noté d_p est donné par

$$\widehat{F}(p, d_p) = \max_{\nu \in S} \{\widehat{F}(p, \nu)\}, \quad (4.8)$$

l'indice p signifie que l'estimation de d peut dépendre de p , et S est un ensemble d'entiers positifs présélectionné de valeurs possibles de d . Pour la simplicité, toutes les statistiques du test $\widehat{F}(p, \nu)$ de (4.8) sont supposées avoir les mêmes degrés de liberté; cela est possible si on fait une sélection appropriée pour b dans la récursion.

Notons que le choix de d_p est quelque peu heuristique, il est basé sur l'idée que si le modèle TAR est le vrai modèle, alors on prend le paramètre de délai qui donne la plus grande valeur significative dans le test de non linéarité. La table suivante donne un exemple de série réelle connue dans la littérature, qui est la série du Lynx Canadien \log_{10} transformée.

Table 4. Test de non linéarité pour une série réelle

Les valeurs de d								
1	2	3	4	5	6	7	8	9
Les valeurs de \hat{F} pour la série : $p = 9$								
2.168	3.162	2.163	1.321	0.696	1.433	1.486	0.968	0.637

Ici l'ensemble des délais possibles est $S = \{1, \dots, p\}$ avec $p = 9$, et $F_{10, 93} = 1.9345$ on obtient $d_9 = 2$, notons qu'on trouve le même résultat pour p de 2 à 9, ce qui nous laisse dire que la sélection de d est stable par rapport à p . En général, d_p peut varier avec p , ainsi, on peut démarrer avec un ordre p raisonnable, en se basant sur une statistique d'identification comme la fonction d'autocorrélation partielle, et ensuite, raffiner l'ordre si c'est nécessaire.

4.3.2 Localisation des valeurs des seuils

Pour un modèle TAR, le choix des valeurs des seuils doit se faire avec précaution, pour voir cela, on prend le cas de $l = 2$, la valeur threshold vérifie $y_{\pi_s} < r_1 \leq y_{\pi_{s+1}}$, ainsi, n'importe quelle valeur dans l'intervalle $[y_{\pi_s}, y_{\pi_{s+1}}]$ est aussi bonne pour estimer r_1 , car elle produit les mêmes résultats de l'ajustement pour le modèle. Alors, faire une sélection parmi une infinité de valeurs possibles reste un problème ouvert.

En général, on peut donner une méthode pour localiser les valeurs des seuils, cette méthode est basé sur les scatterplots de plusieurs statistiques contre la variable threshold, les statistique utilisées montrent un aspect spécial pour le modèle TAR. Bien que les graphes ne soient pas des tests formels, ils nous fournissent des informations utiles pour localiser les thresholds. Les graphes utilisés sont (a) le scatterplot des résidus prédictifs standardisés ou les résidus prédictifs ordinaires

contre y_{t-d_p} , (b) le scatterplot du coefficient estimé $\phi_{d,m}^{(1)}$ de (4.3) contre y_{t-d_p} , et (d) les scatterplots des t ratio des coefficients estimés $\phi_{1,m}^{(1)}$ et $\phi_{0,m}^{(1)}$ contre y_{t-d_p} . Dans le cadre d'une autorégression arrangée, le modèle TAR connaît plusieurs changements à chaque fois il passe par une valeur threshold r_j , les exemples suivants montrent des réalisations de ces graphes pour des série de 1000 observations ($\sigma^{(j)} = 1$).

D'après les simulations montrées ci-dessous et d'autres non illustrées, on peut donner quelques remarques intéressantes sur les scatterplots, le graphe du t ratio de $\phi_{0,m}^{(1)}$ est important, car il désigne le changement dans le niveau du processus, la même chose pour le scatterplot du terme $\phi_{d,m}^{(1)}$ contre y_{t-d} , car ce dernier représente la relation entre y_t et y_{t-d} . L'utilité des scatterplots ne se limite pas au cas d'une seule valeur de threshold. Finalement, comme les y_{t-d} ne sont pas équi-espacés, il est parfois utile dans les scatterplot d'enlever quelques points qui ont des valeurs relativement grandes pour y_{t-d} . Ainsi les $b = (n/10) + p$ derniers points dans les scatterplots sont enlevés.

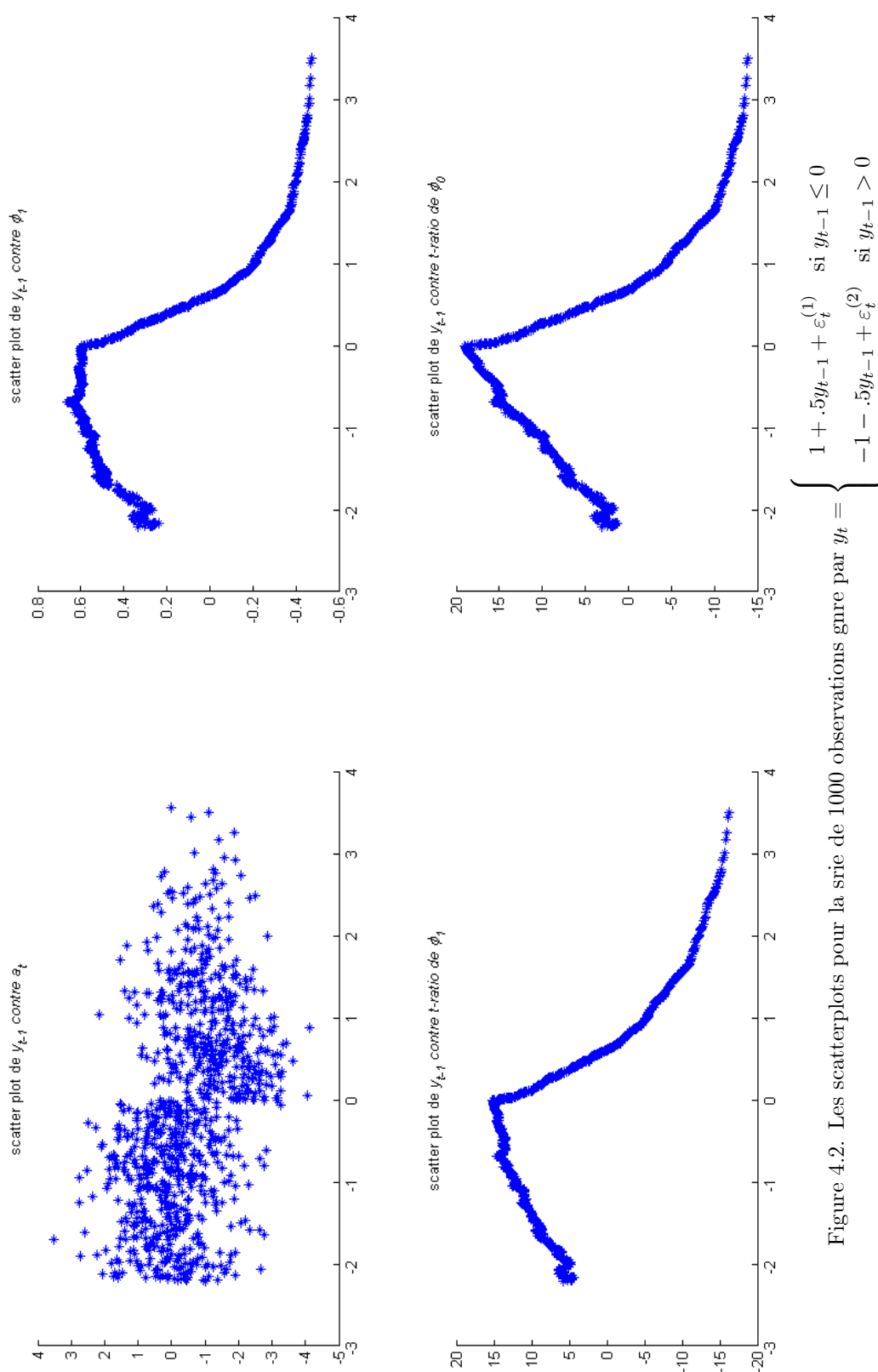
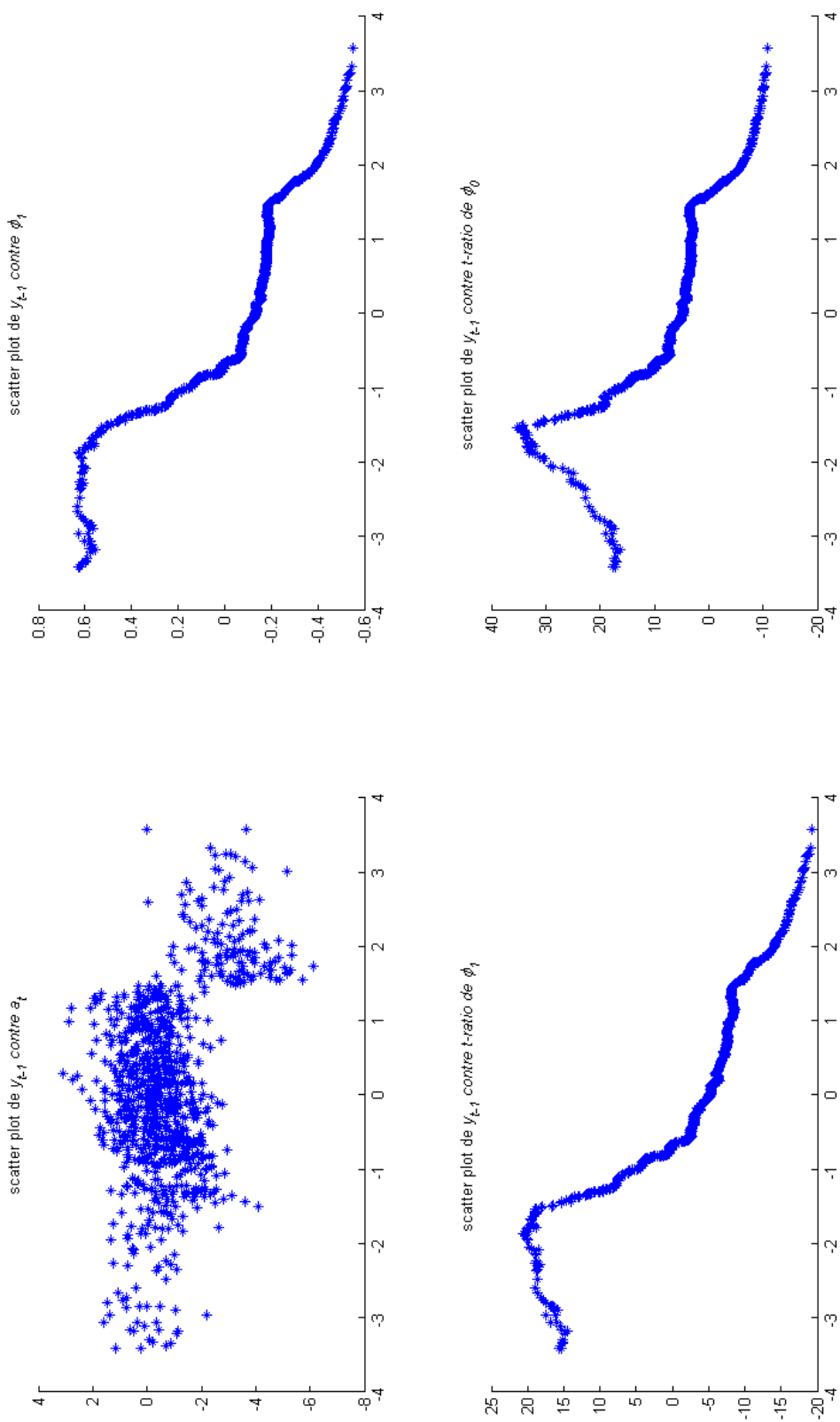


Figure 4.2. Les scatterplots pour la srie de 1000 observations gre par y_t



$$y_t = \begin{cases} 3 + .5y_{t-1} + \varepsilon_t^{(1)} & \text{si } y_{t-1} \leq -1.5 \\ -0.3y_{t-1} + \varepsilon_t^{(2)} & \text{si } -1.5 < y_{t-1} \leq 1.5 \\ -3 - .5y_{t-1} + \varepsilon_t^{(3)} & \text{si } y_{t-1} > 1.5 \end{cases}$$

Figure 4.3. Les scatterplots pour la srie de 1000 observations gure par $y_t =$

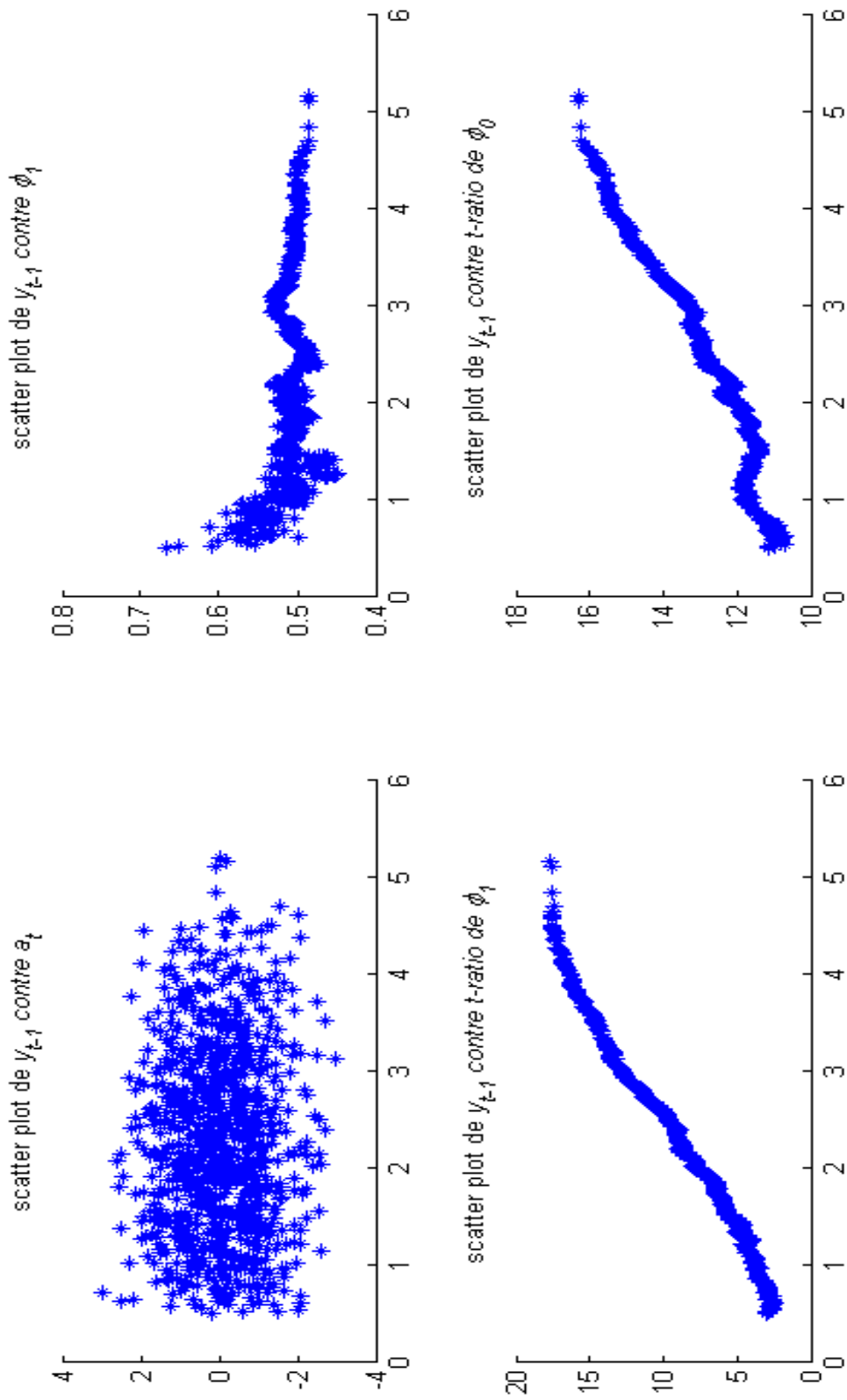


Figure 4.4. Les scatterplots pour la srie de 1000 observations gnr par $y_t = 1 + 0.5y_{t-1} + \varepsilon_t$

4.4 Procédure de modélisation

Utilisons les résultats précédents, on peut proposer une procédure de modélisation d'un modèle TAR qui comprend les quatre étapes suivantes

Etape 1. Choisir le retard p et l'ensemble S des valeurs possibles de d .

Etape 2. Faire l'autorégression arrangée pour le retard p donné et chaque élément d de S , ensuite tester la non linéarité à l'aide de $\widehat{F}(p, d)$. Si le processus est non linéaire, sélectionner d_p qui donne le $\max_{\nu \in S} \widehat{F}(p, \nu)$.

Etape 3. Pour p et d_p donnés, localiser les valeurs des seuils à l'aide des scatterplots de la section précédente.

Etape 4. Raffiner l'ordre et les valeurs des thresholds si c'est nécessaire, dans chaque régime à l'aide des techniques d'autorégression linéaire.

Le choix de p dans l'étape 1 peut être fait en se basant sur la fonction d'autocorrélation partielle (PACF) ou sur un critère d'information comme le critère d'information d'Akaike. Tsay s'est basé sur la fonction d'autocorrélation partielle pour deux raisons (a) le PACF fournit souvent une valeur raisonnable pour p et n'impose aucune pénalisation, (b) le critère d'information AIC est parfois trompeur si le modèle est en effet non linéaire, du fait qu'il est utilisé pour trouver le meilleur modèle et tend à pénaliser les modèles à retard p élevé.

4.5 Application sur une série réelle

Maintenant, on applique le procédure et les statistiques sur une exemple réel qui est la série annuelle du Lynx Canadien (1821-1934, soit 114 observations) transformée par logarithme (en base 10) qui a été étudiée extensivement par plusieurs auteurs (Voir en particulier, Campbell et Walker 1977, Tong 1977a, et Bhansali 1979), cette série présente les aspects suivants qui caractérisent le modèle TAR

- (i) Cyclicité approximative de 10 ans à amplitude variable;
- (ii) La période est ascendante d'un minimum local vers le maximum local suivant et excède la période en descendant d'un maximum local vers le minimum local suivant, ce qui montre l'irréversibilité dans le temps dans la série.

(iii) La série présente un cycle limite qui apparaît dans le graphe du scatterplot entre y_t et y_{t-1} , et entre y_t et y_{t-2} , nous remarquons une trajectoire en spirale, de l'extérieur vers l'intérieur, une trajectoire circulaire laissant une région intérieure non traversée, ce qui montre l'existence d'un cycle limite.

Et d'après le test de non linéarité, on voit la nécessité de l'application du modèle TAR pour ce type de données.

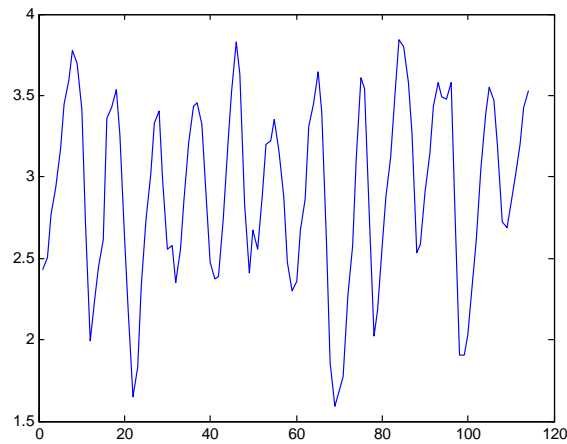


Figure 4.5. Graphe de la série Lynx Canadien Log_{10} transformée (1821-1934).

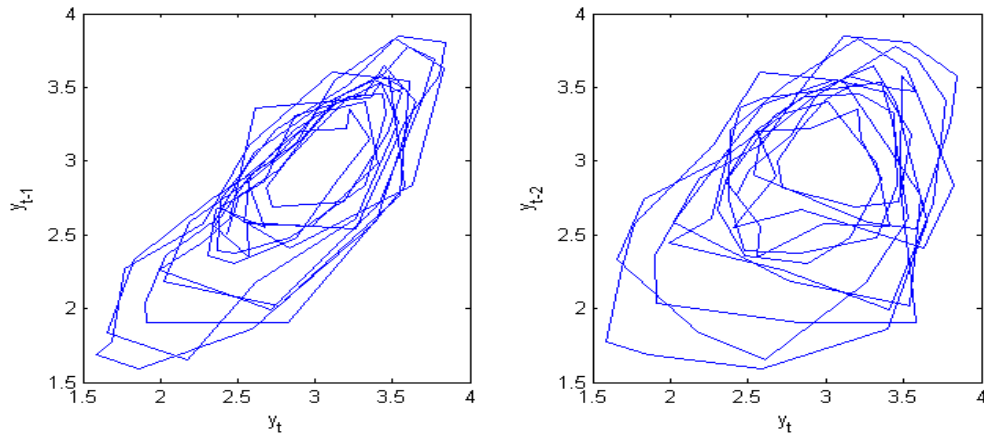


Figure 4.6. scatterplot entre y_t et y_{t-1} , et entre y_t et y_{t-2}

Comme le nombre d'observations est de 114 données, on commence avec $p = 3$ et $S = \{1, 2, 3\}$ à l'étape 1 de la procédure proposée, les statistiques F du test de non linéarité sont 6.89, 11.12 et 3.18 respectivement. Alors $p = 3$ et $d = 2$ sont pris à titre d'essai. pour spécifier les seuils à l'étape 4, on trace les scatterplot suivants

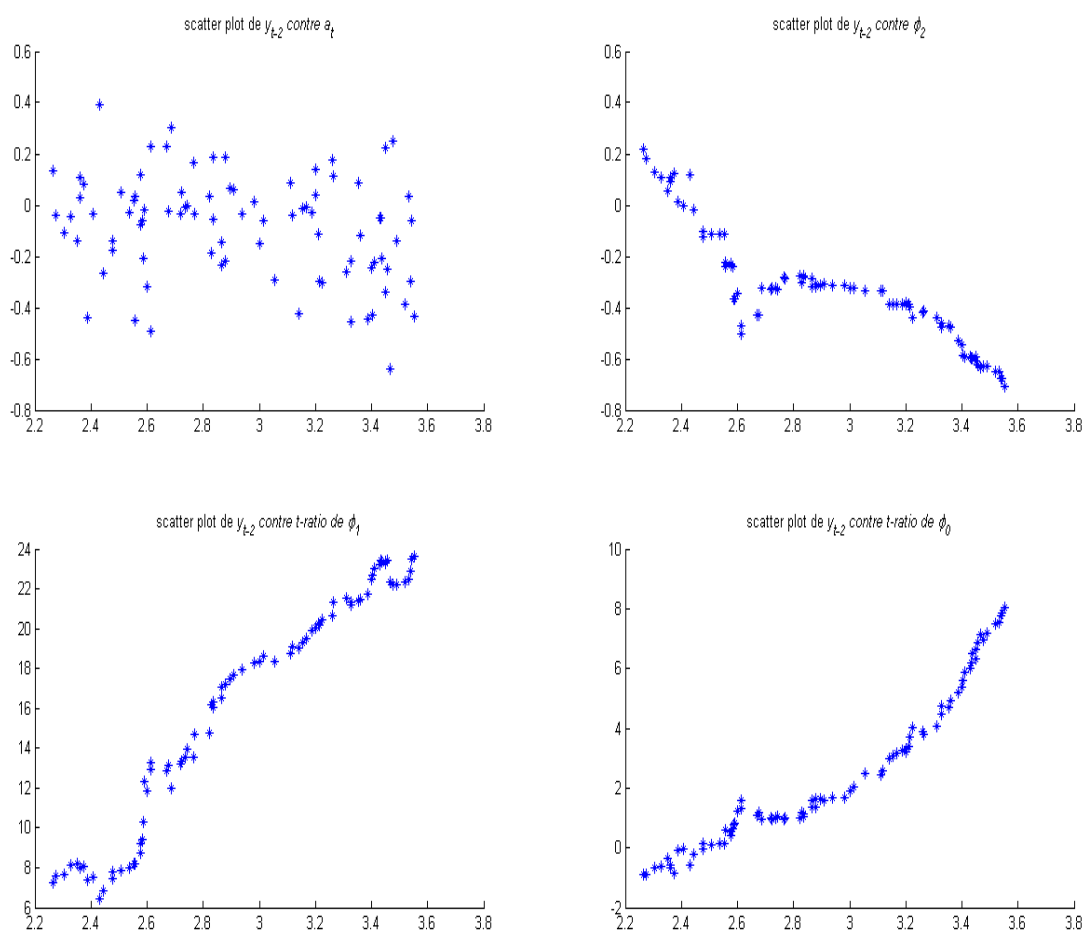


Figure 4.7. Scatterplot de y_{t-2} vs a_t (Haut à gauche), y_{t-2} vs $\phi_{2,m}^{(1)}$ (Haut à droite), y_{t-2} vs t ratio de $\phi_{1,m}^{(1)}$ (Bas à gauche), y_{t-2} vs t ratio de $\phi_{0,m}^{(1)}$ (Bas à droite),

D'après les résultats des scatterplots on remarque plusieurs signes de présence de seuils susceptibles d'être les valeurs thresholds que l'on cherche. On remarque par exemple le changement autour des valeurs $y_{t-2} = 2.6$ et $y_{t-2} = 3.1$ dans le scatterplot de y_{t-2} vs a_t , et autour de la valeur $y_{t-2} = 2.6$ dans le scatterplot de y_{t-2} vs $\phi_{2,m}^{(1)}$ et autour des valeurs $y_{t-2} = 2.4$ et $y_{t-2} = 2.6$, pour le dernier scatterplot le changement observé se trouve dans la zone de nullité du paramètre, alors on ne tient pas compte de ce résultat. Ainsi toutes ces valeurs et les valeurs proches de ces dernières vont être testées lors de l'estimation du modèle TAR, en se basant sur le critère d'information d'Akaike pour le choix du modèle. On trouve que le modèle TAR à trois régimes avec les valeurs thresholds $r_1 = 2.373$ et $r_2 = 3.154$ et les ordres 1,7 et 2 respectivement est le modèle qui résulte en la plus petite valeur

d'AIC = -335.92 qui s'écrit sous la forme suivante (t ratio entre parenthèses)

$$\begin{aligned}
 y_t &= 0.083 + 1.096y_{t-1} + \varepsilon_t^{(1)} \\
 &\quad (0.457) \quad (13.602) \\
 &= 0.628 + 0.961y_{t-1} - 0.114y_{t-2} + 0.226y_{t-3} - 0.613y_{t-4} + 0.475y_{t-5} - 0.393y_{t-6} + 0.284y_{t-7} + \varepsilon_t^{(2)} \\
 &\quad (1.224) \quad (7.475) \quad (-0.500) \quad (1.129) \quad (-3.310) \quad (2.463) \quad (-2.174) \quad (2.489) \\
 &= 2.319 + 1.529y_{t-1} - 1.265y_{t-2} + \varepsilon_t^{(3)} \\
 &\quad (3.351) \quad (14.190) \quad (-5.788)
 \end{aligned}$$

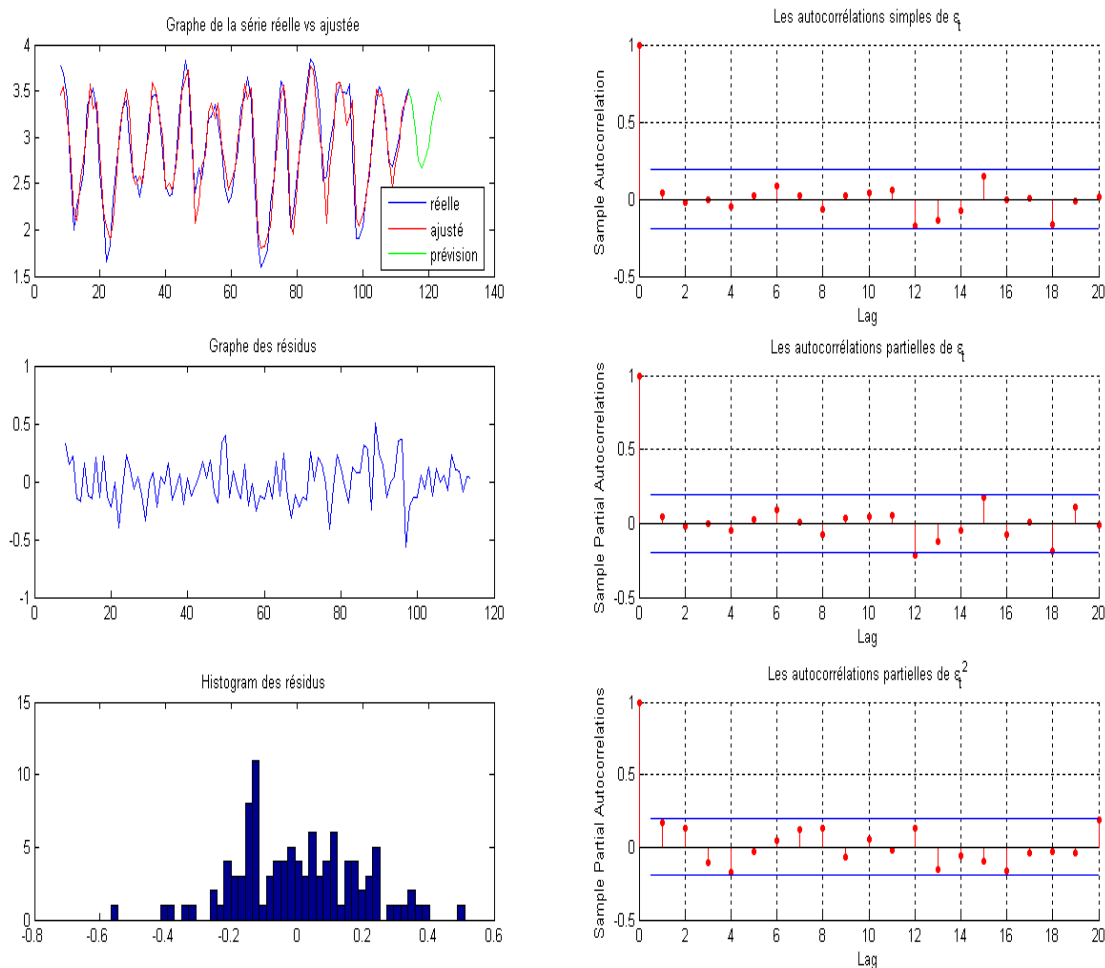


Figure 4.8.

Les écart-types des résidus son $\sigma_1 = 0.13$, $\sigma_2 = 0.161$ et $\sigma_3 = 0.235$. Le diagnostic des ACF et PACF des résidus du modèle ne suggère aucune inadéquation, la même chose pour les PACF des résidus aux carrés.

Tong dans sa modélisation de la série sur le Lynx a proposé un modèle TAR(2; 8, 3), donc à deux régimes où sa valeur AIC = -326.76, En comparant les deux modèles, il est intéressant de noter que la valeur de son threshold est 3.116 qui est proche de notre deuxième threshold, aussi, les études de Haggan et al. (1984) de la série Lynx \log_{10} transformé indiquent l'existence d'un threshold possible autour de 2.2 (pour y_{t-1} plutôt que y_{t-2}). Ce qui semble agréer avec nos résultats.

Annexe

\otimes est le produit de Kronecker, $A^{\otimes n} = A \otimes A \otimes \dots \otimes A$ (n fois).

$\rho(A)$ est le rayon spectral de la matrice carée A , i.e. $\max\{|\text{valeurs propres}| \text{ de } A\}$.

Un vecteur $M > 0$ (≥ 0) signifie que chaque composante de M est strictement positive (positive).

Pour une matrice $A = (a_{ij})$, $|A| = (|a_{ij}|)$.

Pour une suite de matrices, $A^{(k)} = (a_{ij}^{(k)})$, $\max_k |A^{(k)}| := (\max_k |a_{ij}^{(k)}|)$.

Pour les matrices $A = (a_{ij})$ et $B = (b_{ij})$, $A \geq$ (ou \leq) B signifie que $a_{ij} \geq$ (ou \leq) b_{ij} pour tout i, j .

Lemme B.1 Supposons que A, B et A_i sont des matrices de taille $m \times n$, et M un vecteur de dimension n . Alors

- (a) $|AM| \leq |A||M|$
- (b) Si $M \geq 0$ et $A \geq B$, alors $AM \geq BM$
- (c) $|\sum A_i| \leq \sum |A_i|$
- (d) $\max_i (A_i^{\otimes k}) \leq (\max_i |A_i|)^{\otimes k}$
- (e) $[\max_i |A_i^{\otimes k}|]' = \max_i |A_i^{\otimes k}|' = \max_i (|A_i'|^{\otimes k})$.

Lemme B.2 Supposons que A est une matrice de taille $k \times k$, alors $\rho(A^{\otimes n}) \leq 1$ ssi $\rho(A) \leq 1$.

Lemme B.3

- (a) Supposons que $|a_1| + |a_2| + \dots + |a_n| < 1$. Alors toutes les racines de l'équation $x^n - a_1 x^{n-1} - \dots - a_n = 0$ sont à l'intérieur du cercle unité.
- (b) Supposons que toutes les racines de l'équation $x^n - a_1 x^{n-1} - \dots - a_n = 0$ sont à l'intérieur du cercle unité. Alors $a_1 + a_2 + \dots + a_n < 1$, où a_1, a_2, \dots, a_n sont des nombres réels.

References

- [1] Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle. 2nd Int. Symp. on Inf. Theory (B. N. Petrov et F. Csaki, pp. 267-281. Budapest : Akademiai Kiado).
- [2] Akaike, H. (1978) On the threshold of a time series model. *Statistician*. 27. 215-235.
- [3] Akaike, H. (1979), A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. *Biometrika*, 66, 237-242.
- [4] An, H.Z. et Chen, S.G. (1997) A note on the ergodicity of non-linear autoregressive model. *Statst. Probab. Lett.* 34. 365-372.
- [5] An, H.Z. et Huang, F.C. (1996) The geometrical ergodicity of non-linear autoregressive models. *Statist. Sinica*. 6. 943-956.
- [6] Anděl, I. Netuka et Zvára, K. (1984) On threshold autoregressive processes. *Kybernetika*. 20. 89-106.
- [7] Bartlett, M. S. (1966) *Stochastic Processes*, 2nd ed. Cambridge University Press.
- [8] Beneš, V. E. (1967) Existence of finite invariant measures for Markov processes. *Proc. Amer. Math. Soc.* 18. 1058-1061.
- [9] Billingsley, P. (1968) *Convergence of Probability Measures*. Wiley, New York.
- [10] Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*. 31. 307-327.
- [11] Bollerslev, T., Chou, R. Y., et Kroner, K. F. (1992) ARCH modelling in finance. *J. Econometrics*. 52. 5-59.
- [12] Brockwell, J.P. et Davis. R. A. (1991) *Time series : Theory and Methods*. Springer-Verlag. 2nd edition. New York.

- [13] Brockwell, J.P., Liu. J. et Tweedie, R. L. (1992) On the existence of stationary threshold autoregressive moving average processes. *J. Time Ser. Anal.* 13, 95-107.
- [14] Bulmer, M. G. (1974) A statistical analysis of 10 year cycle in Canada. *J. Anim. Ecol.*, 43, 701-718.
- [15] Campbell, M. J. et Walker, A. M. (1977) A survey of statistical work on the MacKenzie River series of annual Canadian lynx trappings for the years 1821-1934 and a new analysis. *J. R. Statist. Soc. A.* 140, 411-431.
- [16] Chen, R. et Tsay, R.S. (1991) On the ergodicity of TAR(1) processes. *Ann. Appl. Probab.* 1. 613-634.
- [17] Chan, K. S., Petrucci, J. D., Tong, H., et Woolford, S. W. (1985) A multiple threshold Ar(1) model. *J. Appl. Prob.* 22, 267-269.
- [18] Chan, K. S., Tong, H. (1985) On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equation. *Adv. Appl. Prob.* 17, 666-678.
- [19] Chan, K. S., Tong, H. (1986) On estimating thresholds in autoregressive models. *J. Time Ser. Anal.* 7. 197-190.
- [20] Chan, K. S., Tong, H. (2010) A note on the invertibility of non linear ARMA models. *J. Stat. Plan. Inf. Lecture note.*
- [21] Cline, D.B.H, et Pu, H.H. (1999) Stability of non-linear AR(1) time series with delay. *Stoch. Proc. Appl.* 82. 307-333.
- [22] Cline, D.B.H, et Pu, H.H. (2004) Stability and Lyapunov exponent of the threshold AR-ARCH models. *Ann. Appl. Probab.* 14. 1920-1949.
- [23] Duffo, M. (1990) *Méthodes récursives aléatoires.* Masson.
- [24] Engle, R. F. (1982) Autoregressive conditional heteroskedasticity with estimates of variance of U.K. inflation. *Econometrica.* 50. 987-1008.

- [25] Feller, W. (1971) An Introduction to Probability Theory and its Applications, Vol.2. Wiley, New York.
- [26] Foster, F. G. (1953) On the stochastic matrices associated with certain queuing processus. *Ann. Math. Statist.* 24, 355-360.
- [27] Guegan, D. et Diebolt, J. (1994) Probabilistic properties of the β -ARCH model. *Statist. Sinica.* 4. 71-89.
- [28] Hamman, V. et Ozaki, T. (1980) Amplitude-dependent exponential AR model fitting for non-linear random vibrations. Dans Anderson(Ed.), O.D. *Time Series.* North-Holland. Amesterdam. 57-71.
- [29] Keenan, D. M. (1985) A Tukey non-additivity type test for time series non-linearity. *Biometrika.* 72. 39-44.
- [30] Klimko, L. A. et Nelson, P. I. (1978) On conditionnal least squares estimation for stochastic processus. *Ann. Statist.* 6, 629-642.
- [31] Lai, T. L. et Wei, C. Z. (1982) Least squares estimates in stochastic regression models with applications to identification and control of dynamic systemes. *Annals of Statistics.* 10. 154-166.
- [32] Liedscher, E. (2005) Towards a unified approach for proving geometric ergodicity and mixing properties of non-linear autoregressive processes. *J. Time. Ser. Anal.* 26. 669-689.
- [33] Ling, S. (1999) On probability properties of a double threshold ARMA conditionnal heteroskedasticity model. *J. Appl. Probab.* 36. 688-705.
- [34] Ling, S. et Tong, H. (2005) Testing a linear moving average model against threshold moving average models. *Ann. Statist.* 33. 2529-2552.
- [35] Ling, S., Tong, H. et Li, D. (2007) Ergodicity and inversibility of threshold moving average models. *Bernoulli.* 13. 161-168.
- [36] Li, C.W. et Li, W.K. (1996) On a double threshold autoregressive conditional heteroskedastic time series model. *J. Appl. Econometrics.* 11. 253-274.

- [37] Liu, J. et Susko, E. (1992) On strict stationarity and ergodicity of a non-linear ARAMA model. *J. Appl. Probab.* 29. 363-373.
- [38] Liu, J., Li, W.K. et Li, C.W. (1997) On threshold autoregression with conditional heteroscedastic variances. *J. Statist. Plann. Inference.* 62. 279-300.
- [39] Meyn, P. S. et Tweedie, R. L. (1993) *Markov chains and stochastic stability.* Springer-Verlag. London.
- [40] Mélard, G. et Roy, R. (1988) Modèles de séries chronologiques avec seuils. *Rev. Statist. Appl.* XXXVI(4). 5-24.
- [41] Milhøj, A. (1985) The moment structure of ARCH processes. *Scand. J. Statist.* 12. 281-292.
- [42] Minorsky, N. (1962) *Non-linear Oscillations.* New York : Van Nostrand.
- [43] Nummelin, E. (1984) *General irreducible Markov chains and non-negative operators.* Cambridge University Press, Cambridge.
- [44] Orey, S. (1971) *Limit Theorems for Markov Chain Transition Probabilities.* Van Nostrand Reinhold, New York.
- [45] Petrucci, J. D. et Woolford, S. W. (1984) Athreshold AR(1) model. *J. Appl. Prob.* 21, 270-286.
- [46] Petrucci, J. D. (1986) On the consistency of least squares estimators for a threshold AR(1) model. *J. Time Ser. Anal.* 7. 269-278.
- [47] Petrucci, J. D. et Davies, N. (1986) A portmanteau test for self-exciting threshold autoregressive-type npn-linearity in time series. *Biometrika.* 73. 687-694.
- [48] Pflug, Ch.G, (1983) The limiting log-likelihood processes for discontinuous density families. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 64. 15-35.
- [49] Pole, A. et Smith, A.F.M. (1985) A bayesian analysis of some threshold switching models. *J. Econ.* 29. 97-119.

- [50] Prestley, M. B. (1980) State-dependent models : a general approach to non-linear time series analysis. *J. Time Series Analysis* 1, 47-71.
- [51] Qian. L. (1998) On maximum likelihood estimators for a threshold autoregression. *J. Statist. Plan. Inf.* 75. 21-46.
- [52] Samia, I. N., et Chan, K. S. (2009) Maximum likelihood estimation of generalised threshold model. Lecture note.
- [53] Tjøstheim, D. (1990) Non-linear time series and Markov chains. *Adv. Appl. Probab.* 22. 587-611.
- [54] Tong, H. (1977a) Some comments on the Canadian lynx data. *J. R. Statist. Soc., A*, 140, 432-436, 448-468.
- [55] Tong, H. (1977b) Discussion of a paper by A. J. Lawrance and N. T. Kottegoda. *J. R. Statist. Soc. A*, 140, 34-35.
- [56] Tong, H. (1978) On threshold model. In *Pattern Recognition and Signal Processing*, ed. C. H. Chen, Sijhoof and Nordhoff, Alphen aan den Rijn, The Netherlands.
- [57] Tong, H. et Lim, K. S. (1980) Threshold autoregression, limit cycles and cyclical data. *J. R. Statist. Soc. B*42, 245-292.
- [58] Tong, H. (1983) *Threshold Models in Non Linear Time series Analysis*. Lecture Notes in Statistics 21, Springer-Verlag, New York.
- [59] Tong, H. (2007) Birth of th threshold time series model. *Statist. Sinica.* 17, (1), 8-14.
- [60] Tong, H. (2010) Threshold models in time series analysis-30 years on. Lecture note.
- [61] Trasvirta, T. et Lukkonen, R (1985) Choosing between linear and threshold autoregressive models. Dans Anderson(Ed.), *O.D. Time Series Analysis: Theory and Practice*. 7. Northe-Holland. Amesterdam. 129-137.
- [62] Tsay, R. S. (1986) Non-linearity test for time series. *Biometrika.* 73. 461-466.

- [63] Tsay, R. S. (1989) Testing and modeling threshold autoregressive processes. J. Amer. Statist. Assoc. 84. 231-240.
- [64] Tsay, R. S. (1998) Testing and modeling multivariate threshold models. J. Amer. Statist. Assoc. 93. 1188-1202.
- [65] Tweedie, R. L. (1974) R -theory for Markov chains on a general state space I: solidarity properties and R -recurrent chains. Ann. Prob. 2, 840-864.
- [66] Tweedie, R. L. (1974) R -theory for Markov chains on a general state space II: r -subinvariant measures for r -transient chains, Ann. Prob. 2, 865-878.
- [67] Tweedie, R. L. (1975) Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space. stoch. Proc. Appl. 3, 385-403.
- [68] Tweedie, R. L. (1976) Criteria for classifying general Markov chains. Adv. Appl. Prob. 8, 737-771.
- [69] Tweedie, R. L. (1983) The existence of moments for stationary Markov chains. J. Appl. Prob. 20, 191-196.
- [70] Tweedie, R. L. (1988) Invariant measure for Markov chains with no irreducibility assumptions. J. Appl. Prob. 25A, 275-285.
- [71] Tweedie, R. L. (2001) Drift conditions and invariant measures for Markov chains. Stoch. Process. Appl. 92. 345-354.