Statistics and Probability Letters 100 (2015) 192–202

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Consistency of quasi-maximum likelihood estimator for Markov-switching bilinear time series models

Abdelouahab Bibi *, Ahmed Ghezal

Département de mathématiques, Université de Constantine (1), Algeria

ARTICLE INFO

Article history:
Received 13 October 2014
Received in revised form 8 February 2015
Accepted 10 February 2015
Available online 25 February 2015

MSC:
primary 62M10
secondary 62M05

Keywords:
Markov-switching bilinear processes
Stationarity
Invertibility
Quasi-maximum likelihood
Strong consistency

ABSTRACT

In this paper, we consider the class of Markov switching bilinear processes (*MS–BL*) that offer remarkably rich dynamics and may be considered as an alternative to model non Gaussian data which exhibit structural changes. In these models, the parameters are allowed to depend upon a latent time-homogeneous Markov chain with finite state space. Analysis based on models with time-varying coefficients has long suffered from the lack of an asymptotic theory except in very restrictive cases. So, first, some basic issues concerning this class of models including sufficient conditions ensuring the existence of stationarity (in strict sense) and ergodic solutions are given. Second, we illustrate the fundamental problems linked with *MS–BL* models, i.e., parameters estimation by considering a maximum likelihood (*ML*) approach. So, we provide the detail on the asymptotic properties of *ML*, in particular, we discuss conditions for its strong consistency.

1. Introduction

Markov-switching (*MS*) time series models have received recently a growing interest in macroeconomics because of their ability to adequately describe various observed time series subjected to change in regime. Flexibility is one of the main advantages of such models which becomes an appealing tool for the modeling of business cycles as originally proposed by Hamilton (1989) and continue to gain popularity especially in financial time series which exhibits structural changes.

There are various different ways to model time series $(X_t, t \in \mathbb{Z})$, $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ which is allowed to change through a finite number of regimes according an (un)observable Markov chain $(s_t, t \in \mathbb{Z})$ sometimes called "regime" with a finite state $\mathbb{S} = \{1, \ldots, d\}$ corresponding to the regime numbers. The most popular model for fitting such datasets is given by

$$X_t = h_{s_t}\left(X_{t-i}, e_{t-j}, 0 < i \leq P, 0 \leq j \leq Q\right) \tag{1.1}$$

for some measurable function $h_{.}$ and some innovation process $(e_t, t \in \mathbb{Z})$. In other words, given s_t, X_t is formed by a regression on $X_{t-1}, \ldots, X_{t-P}$ and on $e_t, \ldots, e_{t-Q}$ with h_{s_t} the regression function. These regression functions may be linear as well as non-linear were investigated in order to capture the probabilistic and statistical properties of such models. For

* Corresponding author.
E-mail addresses: a.bibi@umc.edu.dz (A. Bibi), a.ghezal@centre-univ-mila.dz (A. Ghezal).

http://dx.doi.org/10.1016/j.spl.2015.02.010
0167-7152/

instance, *MS–ARMA*: Francq and Zakoïan (2001), *MS*-nonlinear *ARMA*: Lee (2005) and Yao and Attali (2000), *MS–GARCH*: Francq and Zakoïan (2005) among others. As important subclass of models bound with these regression functions is the so-called discrete-time bilinear process $(X_t, t \in \mathbb{Z})$ generated by the following stochastic difference equation

$$X_t = a_0 (s_t) + \sum_{i=1}^{p} a_i(s_t)X_{t-i} + \sum_{j=0}^{q} b_j(s_t)e_{t-j} + \sum_{j=1}^{Q}\sum_{i=1}^{P} c_{ij}(s_t)X_{t-i}e_{t-j}, \tag{1.2}$$

denoted by $MS–BL(p, q, P, Q)$. In (1.2), the innovation process $(e_t, t \in \mathbb{Z})$ is supposed to be an independent and identically distributed (*i.i.d.*) defined on the same probability space that $(X_t, t \in \mathbb{Z})$ with $E\{e_t\} = 0$ and $E\left\{\log^+ |e_t|\right\} < +\infty$ where for $x > 0$, $\log^+ x = \max(\log x, 0)$. The functions $a_i(s_t)$, $b_j(s_t)$ and $c_{ij}(s_t)$ depend upon a latent first order Markov chain $(s_t, t \in \mathbb{Z})$ defined on a measurable space $(\Omega, \Im, P)$ with finite state-space $\mathbb{S} = \{1, \ldots, d\}$ that controls the dynamics of the process $(X_t, t \in \mathbb{Z})$. So, given $s_t = k \in \mathbb{S}$, X_t satisfies a general bilinear model with parameters $a_i(k)$, $b_j(k)$ and $c_{ij}(k)$. Throughout the paper, $(s_t, t \in \mathbb{Z})$ subject to the following assumption:

Assumption 1. The Markov chain $(s_t, t \in \mathbb{Z})$ is irreducible, aperiodic (and hence stationary and ergodic), n-step transition probabilities matrix $\mathbb{P}^n = \left(p_{ij}^{(n)}, (i, j) \in \mathbb{S} \times \mathbb{S}\right)$ where $p_{ij}^{(n)} = P(s_t = j|s_{t-n} = i)$ with one-step transition probability matrix $\mathbb{P} := \left(p_{ij}\right)_{(i,j)\in\mathbb{S}\times\mathbb{S}}$ where $p_{ij} := p_{ij}^{(1)} = P(s_t = j|s_{t-1} = i)$ for $i, j \in \mathbb{S}$, and stationary distribution $\underline{\pi} = (\pi(1), \ldots, \pi(d))'$ where $\pi(i) = P(s_0 = i)$, $i = 1, \ldots, d$. In addition, we shall assume that e_t and $\{(X_{s-1}, s_t), s \le t\}$ are independent.

Remark 1.1. The innovation process $(e_t, t \in \mathbb{Z})$ may be relaxed to a strictly stationary and heteroscedastic process, i.e., $e_t = \sqrt{h_t}\eta_t$ in which $(\eta_t, t \in \mathbb{Z})$ is an *i.i.d.* (0, 1) and h_t is the conditional variance of $(e_t, t \in \mathbb{Z})$. For instance, $(h_t, t \in \mathbb{Z})$ may be modeled as an $MS - (G)\,ARCH$ model.

The $MS–BL\,(p, q, P, Q)$ encompasses many commonly used models in the literature, indeed:

(i) Standard $BL\,(p, q, P, Q)$ models: These models are obtained by assuming constant the functions $a_i(.)$, $b_j(.)$ and $c_{ij}(.)$ in (1.2) or equivalently by assuming that the chain $(s_t, t \in \mathbb{Z})$ has a single regime (e.g., Granger and Anderson, 1978).

(ii) Hidden-Markov models (*HMM*): In contrast with *MSM*, *HMM* are characterized by the fact that given $s_t = k$, $(X_t, t \in \mathbb{Z})$ is a sequence of conditionally independent random variables with the conditional distribution of each X_t depending on the corresponding state k. For instance, the equation $X_t = a_0(s_t) + b_0(s_t)\,e_t$ with $b_0(s_t) \neq 0$ obtained by setting $a_i(.) = b_j(.) = c_{ij}(.) = 0$ for all $1 \le i, j \le \mathrm{Max}\,\{p, q, P, Q\}$ in (1.2) defines an *HMM* model (e.g., Francq and Roussignol, 1997).

(iii) Markov-switching *ARMA* models (*MS–ARMA*): These models are obtained by setting $c_{ij}(.) = 0$ for all i and j in (1.2) (e.g., Francq and Zakoïan, 2001, Yang, 2000 and Xie, 2009).

(iv) Some classes of $MS–GARCH\,(p, q)$: (e.g., Abramson and Cohen, 2007, Francq and Zakoïan, 2005). (See also Kristensen, 2009 for the building of $GARCH\,(p, q)$ models as special case of $BL\,(p, q, P, Q)$.)

(v) Independent-switching $BL\,(p, q, P, Q)$: In this specification, analyzed by Aknouche and Rabehi (2010) in the bilinear models, with $(s_t, t \in \mathbb{Z})$ is an *i.i.d.* process.

Before we proceed, we need to introduce some algebraic notations.

1.1. Algebraic notations

- $I_{(n)}$ is the $n \times n$ identity matrix, $O_{(k,l)}$ denotes the matrix of order $k \times l$ whose entries are zeros, for simplicity we set $O_{(k)} := O_{(k,k)}$ and $\underline{O}_{(k)} := O_{(k,1)}$.
- The spectral radius of square matrix M is noted $\rho(M)$.
- $\|.\|$ denotes any induced matrix norm on the set of $m \times n$ and $m \times 1$ matrices, and for $\gamma \in]0, 1]$, let $|M|^\gamma := \left(\left|m_{ij}\right|^\gamma\right)$, then it is easy to see that the operator $|.|^\gamma$ is submultiplicative, i.e., $|M_1M_2|^\gamma \le |M_1|^\gamma |M_2|^\gamma$, $\left|M\underline{X}\right|^\gamma \le |M|^\gamma \left|\underline{X}\right|^\gamma$ for any appropriate vector $\underline{X}$ and $\left|\sum_i M_i\right|^\gamma \le \sum_i |M_i|^\gamma$ where the inequality $M \le N$ denotes the elementwise relation $m_{ij} \le n_{ij}$ for all i and j.
- For any set of non random matrices $C := \{C(i), i \in \mathbb{S}\}$, we shall note

$$\mathbb{P}(C) = \left(p_{ij}C'(j)\right)'_{1\le i,j\le d} \quad \text{and} \quad \underline{\pi}(C) = \left(\pi(1)C'(1) \vdots \ldots \vdots \pi(d)C'(d)\right)'.$$

An overview of the paper is as follows. In the next section, we first give sufficient conditions that ensure the existence of strictly, second-order stationary, ergodic and $\sigma((e_r, s_r), r \le t)$-measurable solution of (1.2) and other probabilistic properties such as the invertibility and the existence of moments for some finite order. The conditions are shown to reduce