# ASYMPTOTIC INEFFICIENCY OF MEAN-CORRECTION ON PARAMETER ESTIMATION FOR A PERIODIC FIRST-ORDER AUTOREGRESSIVE MODEL[1]

Antony GAUTIER
Laboratoire de Mathématiques Raphaël Salem - UMR 6085 CNRS
Université de Rouen - Faculté des Sciences
Avenue de l'Université - BP 12
76801 Saint Etienne du Rouvray, France
antony.gautier@univ-rouen.fr

Key Words: Periodic ARMA models; Least squares estimators; Mean-correction; Periodic time series.

AMS (2000) Subject Classifications: Primary 62M10; Secondary 62F12, 91B84.

ABSTRACT

A common practice in time series analysis is to fit a centered model to the mean-corrected data set. For stationary autoregressive moving-average (ARMA) processes, as far as the parameter estimation is concerned, fitting an ARMA model without intercepts to the mean-corrected series is asymptotically equivalent to fitting an ARMA model with intercepts to the observed series. We show that, related to the parameter least squares estimation of periodic ARMA models, the second approach can be arbitrarily more efficient than the mean-corrected counterpart. This property is illustrated by means of a periodic first-order autoregressive model. The asymptotic variance of the estimators for both approaches is derived. Moreover, empirical experiments based on simulations investigate the finite sample properties of the estimators.

[1] *Short running head:* Estimation of Periodic ARMA Models.

## 1. INTRODUCTION

Periodic statistical structures of time series have been extensively exhibited in a wide variety of areas ranging, for example, from economics (Parzen and Pagano, 1979) and climatology (Bloomfield et al., 1994) to electrical engineering (Gardner and Franks, 1975) and hydrology (Vecchia, 1985). A class of time series models that is frequently considered for periodic data is the seasonal autoregressive integrated moving-average (SARIMA) models class, popularized through the work of Box and Jenkins (1976) and become standard textbook material (see Brockwell and Davis, 1991). A large literature has, however, already emphasized that the aforementioned SARIMA models, whose coefficients are time-constant, can not properly capture the behaviour of many periodic variables and/or are not realistic for forecasting in general (see Bibi and Francq, 2003).

Accordingly, in spite of their potentially large number of parameters to be estimated and their inadequacy to the convenient asymptotic theory of ergodic stationary processes, periodic autoregressive moving-average (PARMA) models have gained considerable interest in the last decade. The parameters are allowed to vary periodically on time. PARMA models have involved many studies dealing with the parameter least squares (LS) or maximum likelihood (ML) estimation. A short list of some significant contributions on PARMA time series is Pagano (1978), Anderson and Vecchia (1993), McLeod (1994), Adams and Goodwin (1995), Lund and Basawa (2000), Basawa and Lund (2001), Ula and Smadi (2003) and Azrak and Mélard (2004). The reader is referred to Lund and Basawa (1999) or Bibi and Francq (2003) for further discussion on the main differences between SARIMA models and PARMA models. PARMA time series models are examples of periodically correlated processes (whose original source of inspiration was the work of Gladyšhev, 1961, on periodic autocorrelations, see Alpay et al., 2001, for a more recent reference) and have become an appealing tool for handling the possible periodic patterns of real-life data.

In general, most (periodic) time series models are preceded by a preliminary step which consists in removing the (seasonal) sample mean(s) from the series. One can hardly deny that, in (periodic) time series analysis, it is customary to subtract the (seasonal) sample mean(s) of the data from each observation to generate a series to which, in a second time, we fit a zero-mean model. For stationary autoregressive moving-average (ARMA) processes, as far as the parameter estimation is concerned, fitting a zero-mean ARMA model to the mean-corrected series is known to avoid estimating an additional (position) parameter and to be asymptotically equivalent to fitting an ARMA model with intercepts to the observed series. Basically, the asymptotic properties of parameter estimators for stationary ARMA models are invariant to the presence of constants. This property does not hold any more for PARMA models. The main purpose of this paper is to show how the parameter LS estimators of a non-centered periodic first-order autoregressive [PAR(1)] model can be arbitrarily more accurate than the LS estimators of a similar, but mean-corrected PAR(1) model, when the length of the series increases. Due to limited space and for simplicity of presentation, in this paper we restrict the analysis to a daily two-regime PAR(1) model. Intuitively, it seems reasonable to conjecture a similar result for any PARMA time series, whatever the periodicity (for example, the monthly temperature series has seasonal periodicity) and whatever the AR and MA orders. Such a study for general PARMA models requires more technical difficulties and annoying calculations. At this stage, the goal of the paper is merely to provide a simple example of a PARMA time series model that satisfies the expected property on efficiency estimation.

The remainder of the paper proceeds as follows. Section 2 introduces the model under consideration and provides a convenient Markovian representation. This state-space representation is used in Section 3 to derive the asymptotic autocovariance structure of the model. In Section 4, we consider the problem of estimating the model parameters via two different approaches: (i) a 'first-step' (FS) method dealing with the mean-corrected model; (ii) an LS procedure involving position and AR parameters. The asymptotic variance of each estimator is derived. As we will see, the LS estimator can provide substantial gains in terms of asymptotic accuracy compared with the FS estimator. Section 5 is dedicated to numerical illustrations of our theoretical results through Monte Carlo experiments. Section 6 summarizes the main results of the paper and concludes.

## 2. DEFINITION AND MARKOVIAN REPRESENTATION OF THE MODEL

The PAR(1) model we consider is first introduced. Next, we give a Markovian representation of the model which will be useful to study further its asymptotic autocovariance structure.

### 2.1. THE MODEL

We consider a daily time series $(X_t)_{t=0,1,\dots}$ exhibiting periodic changes in regime at known dates. Without loss of generality, let us assume that time $t = 0$ corresponds to Monday. We suppose that there exists a finite number of recurrent regimes which alternate with a constant periodicity. For pragmatic reasons, we consider only two regimes, but this work can be extended to more regimes and the kind of model to be considered could then be relevant for daily time series with different behaviours per day. Applications we have in mind are economic time series. It is common knowledge that a large number of economic variables presents a daily behaviour. For instance, when analyzing macroeconomic or financial time series data, human activity is an important source of daily periodicity into many fields such as electrical consumption, motorway traffic, or stock market indexes, among others. In this paper, we suspect different behaviours on weekdays and on weekends. This hypothesis typically holds, for instance, for the number of daily passengers in a given railway. We make the assumption that Regime 1 corresponds to weekdays and Regime 2 deals with weekends.

We define by $\Delta(1) = \{0, 1, 2, 3, 4, 7, 8, 9, \dots\}$ and $\Delta(2) = \{5, 6, 12, 13, 19, 20, \dots\}$ the set of indices corresponding respectively to the weekday regime and to the weekend regime. The quantity

$$s_t = \mathbb{I}_{\Delta(1)}(t) + 2\mathbb{I}_{\Delta(2)}(t)$$

indicates the regime corresponding to the index $t$, where $\mathbb{I}_{\Delta(k)}(t)$ stands for the indicator function of the set $\Delta(k)$. Because of the periodic structure of the series, we have $s_0 = 1$ and $s_t = s_{t+7j}$ for all $t \geq 0$ and for all $j \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$ such that $j \geq -t/7$. Since we are acquainted with the dates of changes in regime, the sequence $(s_t)_{t=0,1,\dots}$ is known, periodic and purely deterministic. More precisely, $(s_t)_{t=0,1,\dots}$ is composed of the subsequence of regime states $(1,1,1,1,1,2,2)$ that repeats itself in a regular cycle. It is important to outline that the regime switches at regular time intervals and, unlike Francq and Gautier (2004a, 2004b), the observed sequence $(s_t)_{t=0,1,\ldots}$ of regime states is not generated by a stationary ergodic random process.

As previously said, we suppose that the dynamics of $X_t$ in each regime is described by a first-order autoregressive equation. Then we consider the heteroskedastic PAR(1) model

$$\begin{cases} X_0 = m_0(s_0) + \varepsilon_0, \\ X_t - m_0(s_t) = a_0(s_t)\left\{X_{t-1} - m_0(s_{t-1})\right\} + \varepsilon_t, \quad t = 1, 2, \ldots, \end{cases} \tag{2.1}$$

whose unknown vector parameter is denoted by

$$\underline{\theta}_0 = \{\theta_0(1), \theta_0(2), \theta_0(3), \theta_0(4)\}' = \{m_0(1), m_0(2), a_0(1), a_0(2)\}'$$

and belongs to an open subset $\Theta$ of $\mathbb{R}^4$. The innovations $(\varepsilon_t)_{t=0,1,\ldots}$ are a sequence of independent random variables with mean zero and whose variance is allowed to depend on time: we have $\varepsilon_t = \sigma_0(s_t)\eta_t$ where $(\eta_t)_{t=0,1,\ldots}$ is a sequence of independent and identically distributed (iid) random variables with mean zero, variance one, $E(\eta_t^3) = 0$ and $E(\eta_t^4) < \infty$. The components of the nuisance parameter $\underline{\sigma}_0 = \{\sigma_0(1), \sigma_0(2)\}'$ are strictly positive numbers. All the random variables involved in the paper are defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Statistical inference for model (2.1) with $m_0 \equiv 0$ has been studied by Bloomfield et al. (1994), McLeod (1994), Basawa and Lund (2001) or Bibi and Francq (2003), among others.

Iterating (2.1), the following time-dependent MA representation holds

$$X_t - m_0(s_t) = \varepsilon_t + \sum_{i=1}^{t} \phi_{t,i}(\underline{\theta}_0)\varepsilon_{t-i}, \quad t = 0, 1, \ldots, \tag{2.2}$$

where the $\phi_{t,i}(\cdot)$s are known functions from $\mathbb{R}^4$ to $\mathbb{R}$ such that

$$\phi_{t,i}(\underline{\theta}_0) = \prod_{j=0}^{i-1} a_0(s_{t-j}), \quad \forall i \geq 1.$$

In view of (2.2), $\varepsilon_t$ is independent of $X_u$ for $u < t$, and $m_0(s_t)$ can be interpreted as the expectation of $X_t$. It is important to note that, in general, $(X_t)$ is not stationary, even when the initial values are neglected. In particular, for an infinite number of dates $k$, we have $E(X_k) = \cdots = E(X_{k+4}) = m_0(1)$ and $E(X_{k+5}) = E(X_{k+6}) = m_0(2)$. Figure 2.1 displays a realization of length 200 of model (2.1).

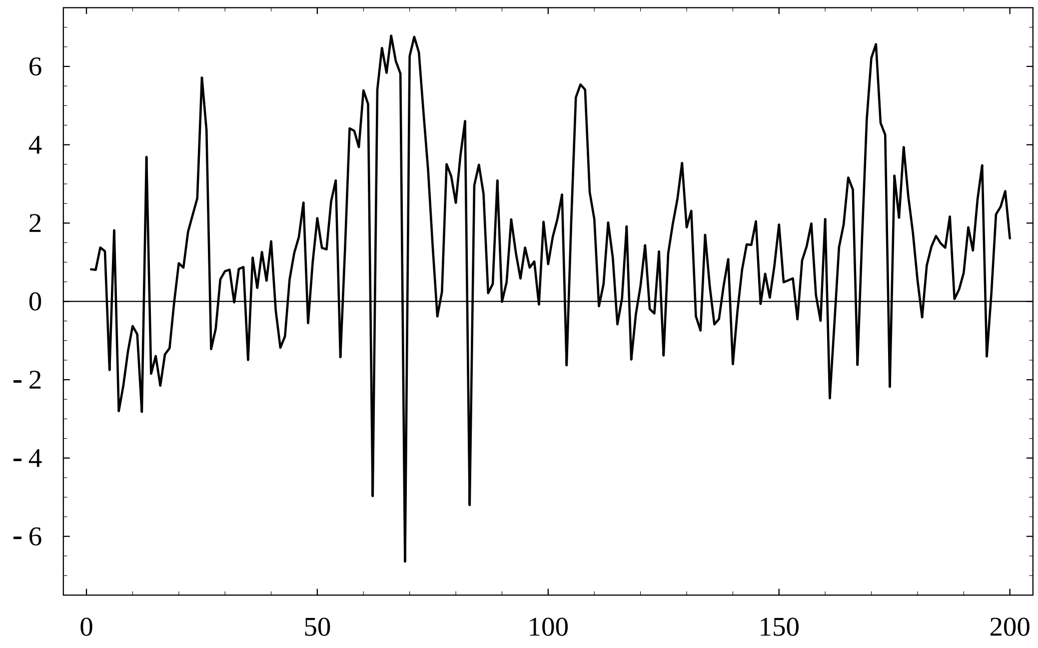


Figure 2.1. A simulation of length 200 for model (2.1) with $\underline{\theta}_0 = (1.0, 0.0, 0.9, -0.9)'$ and $\underline{\sigma}_0 = (1.0, 1.0)'$.

## 2.2. MARKOVIAN REPRESENTATION

With periodic coefficients, it is known that seasons can be embedded into a multivariate process (Tiao and Grupe, 1980). A practical representation of our univariate model (2.1) is the multivariate first-order autoregressive [VAR(1)] form

$$\begin{cases} \underline{Y}_0 = \underline{u}_0, \\ \underline{Y}_t = A\underline{Y}_{t-1} + \underline{u}_t, \quad t = 1, 2, \ldots, \end{cases} \tag{2.3}$$

where

$$
\underline{Y}_t = \begin{pmatrix} X_{7t} - m_0(1) \\ X_{7t-1} - m_0(2) \\ X_{7t-2} - m_0(2) \\ X_{7t-3} - m_0(1) \\ X_{7t-4} - m_0(1) \\ X_{7t-5} - m_0(1) \\ X_{7t-6} - m_0(1) \end{pmatrix}, \quad \underline{Y}_0 = \begin{pmatrix} X_0 - m_0(1) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{u}_0 = \begin{pmatrix} \varepsilon_0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{\varepsilon}_t = \begin{pmatrix} \varepsilon_{7t} \\ \varepsilon_{7t-1} \\ \varepsilon_{7t-2} \\ \varepsilon_{7t-3} \\ \varepsilon_{7t-4} \\ \varepsilon_{7t-5} \\ \varepsilon_{7t-6} \end{pmatrix},
$$

$$
A = \begin{pmatrix} a_0^5(1)a_0^2(2) & 0 & 0 & 0 & 0 & 0 & 0 \\ a_0^4(1)a_0^2(2) & 0 & 0 & 0 & 0 & 0 & 0 \\ a_0^4(1)a_0(2) & 0 & 0 & 0 & 0 & 0 & 0 \\ a_0^4(1) & 0 & 0 & 0 & 0 & 0 & 0 \\ a_0^3(1) & 0 & 0 & 0 & 0 & 0 & 0 \\ a_0^2(1) & 0 & 0 & 0 & 0 & 0 & 0 \\ a_0(1) & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{2.4}
$$

and, for $t \geq 1$, $\underline{u}_t = B\underline{\varepsilon}_t$ with

$$
B = \begin{pmatrix} 1 & a_0(1) & a_0(1)a_0(2) & a_0(1)a_0^2(2) & a_0^2(1)a_0^2(2) & a_0^3(1)a_0^2(2) & a_0^4(1)a_0^2(2) \\ 0 & 1 & a_0(2) & a_0^2(2) & a_0(1)a_0^2(2) & a_0^2(1)a_0^2(2) & a_0^3(1)a_0^2(2) \\ 0 & 0 & 1 & a_0(2) & a_0(1)a_0(2) & a_0^2(1)a_0(2) & a_0^3(1)a_0(2) \\ 0 & 0 & 0 & 1 & a_0(1) & a_0^2(1) & a_0^3(1) \\ 0 & 0 & 0 & 0 & 1 & a_0(1) & a_0^2(1) \\ 0 & 0 & 0 & 0 & 0 & 1 & a_0(1) \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.
$$

Note that $\mathrm{Var}(\underline{\varepsilon}_t)$ is the diagonal matrix $\mathrm{Diag}\{\sigma_0^2(1), \sigma_0^2(2), \sigma_0^2(2), \sigma_0^2(1), \sigma_0^2(1), \sigma_0^2(1), \sigma_0^2(1)\}$. The Markovian state-space representation (2.3) turns out to be a useful tool to study the asymptotic autocovariance structure and the LS estimation of model (2.1).

REMARK 2.1. The conventional notation, where $\underline{Y}_0$ denotes the first cycle, *i.e.*

$$\underline{Y}_0 = \{X_0 - m_0(1), X_1 - m_0(1), \ldots, X_5 - m_0(2), X_6 - m_0(2)\}',$$

and hence $\underline{Y}_t$ can be expressed as

$$\underline{Y}_t = \{X_{7t} - m_0(1), X_{7t+1} - m_0(1), \ldots, X_{7t+4} - m_0(1), X_{7t+5} - m_0(2), X_{7t+6} - m_0(2)\}',$$

can also be used to derive a Markovian representation in the form (2.3). The minor pitfall with this notation is that it involves higher powers for the AR coefficients in the matrix $A$. But one can note that the conventional notation finally gives the same results as the ones related to the second-order structure and to the asymptotic behaviours, respectively presented in Sections 3 and 4.

## 3. SECOND-ORDER STRUCTURE

Arguably, the most fundamental feature of a time series model is its autocovariance structure. In this section, we derive the (asymptotic) autocovariance structure of model (2.1) from its Markovian representation given in (2.3). Let us mention that Shao and Lund (2004) is a recent study of this problem.

### 3.1. AUTOCOVARIANCE FUNCTION

First note that, iterating (2.3), we obtain the MA expansion

$$\underline{Y}_t = \sum_{i=0}^{t} A^i \underline{u}_{t-i}, \quad t = 0, 1, \ldots, \tag{3.1}$$

and, for all $h \geq 0$, we also have

$$\underline{Y}_t = A^h \underline{Y}_{t-h} + \sum_{i=0}^{h-1} A^i \underline{u}_{t-i}, \quad t = 0, 1, \ldots. \tag{3.2}$$

It is worth mentioning that, in view of the special form of the matrix $A$ given by (2.4) and noting $\rho = a_0^5(1)a_0^2(2)$, we have

$$A^i = \rho^{i-1} A \quad \text{ for } i \geq 1.$$

Therefore, from (3.2) and (3.1), since $\underline{Y}_{t-h}$ is independent of $\underline{u}_t, \underline{u}_{t-1}, \ldots, \underline{u}_{t-h+1}$, and $\underline{u}_{t-i}$ is independent of $\underline{u}_{t-j}$ for $i \neq j$, we have for all $t \geq 1$ and $h \geq 0$,

$$\begin{aligned}
\mathrm{Cov}(\underline{Y}_t, \underline{Y}_{t-h}) &= \mathrm{Cov}\left( A^h \underline{Y}_{t-h} + \sum_{i=0}^{h-1} A^i \underline{u}_{t-i}, \underline{Y}_{t-h} \right) \\
&= A^h \mathrm{Var}(\underline{Y}_{t-h}) \\
&= A^h \mathrm{Var}\left( \sum_{i=0}^{t-h} A^i \underline{u}_{t-h-i} \right) \\
&= A^h \mathrm{Var}\left( \underline{u}_{t-h} + \sum_{i=1}^{t-h} \rho^{i-1} A \underline{u}_{t-h-i} \right) \\
&= A^h \left\{ \mathrm{Var}(\underline{u}_{t-h}) + \sum_{i=1}^{t-h} \rho^{2(i-1)} A \mathrm{Var}(\underline{u}_{t-h-i}) A' \right\},
\end{aligned}$$

which finally gives

$$\mathrm{Cov}(\underline{Y}_t, \underline{Y}_{t-h}) = A^h \left\{ \mathrm{Var}(\underline{u}_{t-h}) + \rho^{2(t-h-1)} A \mathrm{Var}(\underline{u}_0) A' + \sum_{i=1}^{t-h-1} \rho^{2(i-1)} A \mathrm{Var}(\underline{u}_{t-h-i}) A' \right\}. \tag{3.3}$$

Letting $\Psi = \mathrm{Var}(\underline{u}_t) = [\psi_{i,j}]_{i,j=1}^{7}$, we have $(\underline{u}_t)_{t=1,2,\ldots}$ iid such that

$$\Psi = B \mathrm{Var}(\underline{\varepsilon}_t) B' \tag{3.4}$$

with diagonal components

$$\begin{aligned}
\psi_{1,1} &= \sigma_0^2(2) a_0^2(1) \left\{ 1 + a_0^2(2) \right\} + \sigma_0^2(1) \left[ 1 + a_0^4(2) \left\{ a_0^2(1) + a_0^4(1) + a_0^6(1) + a_0^8(1) \right\} \right], \\
\psi_{2,2} &= \sigma_0^2(2) \left\{ 1 + a_0^2(2) \right\} + \sigma_0^2(1) a_0^4(2) \left\{ 1 + a_0^2(1) + a_0^4(1) + a_0^6(1) \right\}, \\
\psi_{3,3} &= \sigma_0^2(2) + \sigma_0^2(1) a_0^2(2) \left\{ 1 + a_0^2(1) + a_0^4(1) + a_0^6(1) \right\}, \\
\psi_{4,4} &= \sigma_0^2(1) \left\{ 1 + a_0^2(1) + a_0^4(1) + a_0^6(1) \right\}, \quad \psi_{5,5} = \sigma_0^2(1) \left\{ 1 + a_0^2(1) + a_0^4(1) \right\}, \\
\psi_{6,6} &= \sigma_0^2(1) \left\{ 1 + a_0^2(1) \right\}, \quad \psi_{7,7} = \sigma_0^2(1),
\end{aligned}$$

and off-diagonal elements $\psi_{i,1} = a_0(1)\psi_{i,2}$ for $i = 2, \ldots, 7$, $\psi_{i,2} = a_0(2)\psi_{i,3}$ for $i = 3, \ldots, 7$, $\psi_{i,3} = a_0(2)\psi_{i,4}$ for $i = 4, \ldots, 7$, $\psi_{i,4} = a_0(1)\psi_{i,5}$ for $i = 5, 6, 7$, $\psi_{i,5} = a_0(1)\psi_{i,6}$ for $i = 6, 7$, $\psi_{7,6} = a_0(1)\psi_{7,7}$. The variance matrix of $\underline{u}_0$ is such that its first element $\{\mathrm{Var}(\underline{u}_0)\}_{1,1} = \sigma_0^2(1)$

and its other elements $\{\mathrm{Var}(\underline{u}_0)\}_{i,j}$ are zero for all $(i,j) \neq (1,1)$. Therefore, in view of (3.3) and the variance matrix of the noise, we obtain the autocovariance structure

$$\mathrm{Cov}(\underline{Y}_t, \underline{Y}_{t-h}) = A^h \left\{ \Psi + \rho^{2(t-h-1)} \sigma_0^2(1) AA' + A\Psi A' \sum_{i=1}^{t-h-1} \rho^{2(i-1)} \right\} \tag{3.5}$$

for all $t \geq 1$ and $h \geq 0$. In the particular case $h = 0$, which will be relevant in Section 4, we have

$$\mathrm{Var}(\underline{Y}_t) = \begin{cases} \Psi + \rho^{2(t-1)} \sigma_0^2(1) AA' + A\Psi A' \left\{1 - \rho^{2(t-1)}\right\} (1-\rho^2)^{-1} & \text{if } |\rho| \neq 1, \\ \Psi + \sigma_0^2(1) AA' + (t-1) A\Psi A' & \text{if } |\rho| = 1. \end{cases}$$

We can collect all previous results in a proposition that sums up the autocovariance structure of $(\underline{Y}_t)_{t=1,2,\ldots}$. This proposition notably provides the convergence rate of $\mathrm{Cov}(\underline{Y}_t, \underline{Y}_{t-h})$.

PROPOSITION 3.1. Consider the process $(\underline{Y}_t)_{t=1,2,\ldots}$ defined by (2.3) and let $\rho = a_0^5(1) a_0^2(2)$. For all $h \geq 0$,

i) if $|\rho| \neq 1$, then $\mathrm{Cov}(\underline{Y}_t, \underline{Y}_{t-h}) = A^h \left\{ \Psi + \rho^{2(t-h-1)} \sigma_0^2(1) AA' + A\Psi A' \frac{1-\rho^{2(t-h-1)}}{1-\rho^2} \right\}$,

ii) if $|\rho| = 1$, then $\mathrm{Cov}(\underline{Y}_t, \underline{Y}_{t-h}) = A^h \{ \Psi + \sigma_0^2(1) AA' + (t-h-1) A\Psi A' \}$,

where $\Psi$ is given by (3.4).

REMARK 3.1. In view of (2.4), the matrix $A$ can be written in the form $A = \underline{a}\underline{e}'$, where $\underline{e} = (1, 0, \ldots, 0)'$ and

$$\underline{a} = \{\rho, a_0^4(1) a_0^2(2), a_0^4(1) a_0(2), a_0^4(1), a_0^3(1), a_0^2(1), a_0(1)\}'.$$

We thence have $AA' = \underline{a}\underline{e}'\underline{e}\underline{a}'$ and $A\Psi A' = \underline{a}\underline{e}'\Psi\underline{e}\underline{a}' = \psi_{1,1}\underline{a}\underline{a}'$. Therefore, from (3.5) we can write

$$\mathrm{Cov}(\underline{Y}_t, \underline{Y}_{t-h}) = A^h \left( \Psi + \lambda_{t-h} \underline{a}\underline{a}' \right)$$

with $\lambda_{t-h} = \rho^{2(t-h-1)} \sigma_0^2(1) + \psi_{1,1} \sum_{i=1}^{t-h-1} \rho^{2(i-1)}$, which provides a simpler expression for $\mathrm{Cov}(\underline{Y}_t, \underline{Y}_{t-h})$.

## 3.2. ASYMPTOTIC AUTOCOVARIANCE STRUCTURE

At present, let us deduce the asymptotic autocovariance function of $(\underline{Y}_t)$, defined by $\Gamma_{\underline{Y}}(h) = \lim_{t\to\infty} \mathrm{Cov}(\underline{Y}_t, \underline{Y}_{t-h})$ for all $h \geq 0$. Asymptotically, neglecting the initial conditions, it is well-known that, from (2.3), $(\underline{Y}_t)$ is stable whenever the spectral radius of $A$, noted as $\rho(A)$, is such that $\rho(A) = |\rho| < 1$. Consequently, we suppose that the absolute value of the product of the autoregressive (AR) coefficients over a period is strictly less than unity [this is the causality condition for model (2.1) given by Vecchia, 1985, and Hurd et al., 2002], *i.e.*

$$|\rho| = \left|a_0^5(1)a_0^2(2)\right| < 1. \tag{3.6}$$

Therefore, from Proposition 3.1 i), we obtain the asymptotic lag-zero autocovariance matrix $\Gamma_{\underline{Y}}(0) = [\gamma_{i,j}(0)]_{i,j=1}^7$ by

$$\mathrm{Var}(\underline{Y}_t) \to \Gamma_{\underline{Y}}(0) = \Psi + A\Psi A' \left(1-\rho^2\right)^{-1} \qquad \text{as } t \to \infty. \tag{3.7}$$

As an example, the asymptotic variance of Mondays is

$$\begin{aligned}\gamma_{1,1}(0) \quad &= \quad \left[\left\{1 + a_0^2(1)a_0^4(2) + a_0^4(1)a_0^4(2) + a_0^6(1)a_0^4(2) + a_0^8(1)a_0^4(2)\right\}\sigma_0^2(1)\right.\\ &\qquad \left. + \left\{1 + a_0^2(2)\right\} a_0^2(1)\sigma_0^2(2)\right]\left(1-\rho^2\right)^{-1},\end{aligned}$$

and the off-diagonal elements of $\Gamma_{\underline{Y}}(0)$ are of the form $\gamma_{i,1}(0) = a_0(1)\gamma_{i,2}(0)$ for $i = 2, \ldots, 7$, $\gamma_{i,2}(0) = a_0(2)\gamma_{i,3}(0)$ for $i = 3, \ldots, 7$ etc. In view of the components of the right-hand side in (3.7), the Yule-Walker equations (see Brockwell and Davis, 1991, p. 420) give the expressions of $\Gamma_{\underline{Y}}(h)$ for $h \geq 1$ [we have $\Gamma_{\underline{Y}}(h) = \rho^{h-1}A\Gamma_{\underline{Y}}(0)$].

## 4. PARAMETER ESTIMATION

In this section, we investigate the problem of estimating model (2.1) given a sequence $(X_0, \ldots, X_n)$ of observations. For convenience, we assume that $n+1$ is an integer multiple of 7, so that the sample size consists of $N$ full periods of data. Two different approaches are investigated, and then compared: the first one is based on FS estimators defined below, while the second one consists in a common LS procedure.

### 4.1. A 'FIRST-STEP' APPROACH

For $k = 1, 2$, the FS approach deals with the parameter estimation of model (2.1) using: (i) the average of the $k$-th regime over the $(n+1)$ observations, denoted by $\tilde{m}_n(k)$, to estimate the position parameter $m_0(k)$; (ii) the usual expression of the LS estimator of $a_0(k)$ over the $(n+1)$ mean-corrected data $X_t - \tilde{m}_n(s_t)$, $t = 0, \dots, n$, to estimate the AR coefficient $a_0(k)$.

#### 4.1.1. DEFINITION OF 'FIRST-STEP' ESTIMATORS

In order to estimate the parameter $\underline{\theta}_0$ of model (2.1), we consider in a first time the FS estimator $\tilde{\underline{\theta}}_n = \{\tilde{m}_n(1), \tilde{m}_n(2), \tilde{a}_n(1), \tilde{a}_n(2)\}'$ defined by:

$$\tilde{m}_n(k) = \frac{\sum_{t=0}^{n} X_t \mathbb{I}_k(s_t)}{\sum_{t=0}^{n} \mathbb{I}_k(s_t)} \tag{4.1}$$

and

$$\tilde{a}_n(k) = \frac{\sum_{t=1}^{n} \{X_t - \tilde{m}_n(k)\} \{X_{t-1} - \tilde{m}_n(s_{t-1})\} \mathbb{I}_k(s_t)}{\sum_{t=1}^{n} \{X_{t-1} - \tilde{m}_n(s_{t-1})\}^2 \mathbb{I}_k(s_t)}, \tag{4.2}$$

for $k = 1, 2$. It will be proved, as expected, that the estimator in (4.1) converges to $m_0(k)$. Afterwards, we will show that the estimator in (4.2) has the same asymptotic behaviour as the pseudo-estimator

$$\tilde{a}_n^*(k) = \frac{\sum_{t=1}^{n} \{X_t - m_0(k)\}\{X_{t-1} - m_0(s_{t-1})\}\mathbb{I}_k(s_t)}{\sum_{t=1}^{n} \{X_{t-1} - m_0(s_{t-1})\}^2 \mathbb{I}_k(s_t)}. \tag{4.3}$$

Of course, when the position parameters $m_0(k)$ are unknown, $\tilde{a}_n^*(k)$ can not be computed. Basawa and Lund (2001, Theorem 4.1) and Bibi and Francq (2003, Theorems 2.1–2.2) have shown that, in the case $m_0 \equiv 0$, the estimator in (4.3) is consistent and has a limiting centered normal distribution $\mathcal{N}(0, \Sigma^*)$ related to the parameter LS estimation of model (2.1). We now investigate the asymptotic behaviour of $\tilde{\underline{\theta}}_n$ in the following subsection.

#### 4.1.2. LARGE SAMPLE PROPERTIES

We first focus on the convergence in mean-square (m.s.) of the FS estimator $\tilde{m}_n(k)$. We have the following result.

THEOREM 4.1. Consider model (2.1) for which $\tilde{m}_n(k)$, defined by (4.1), is an estimator of $m_0(k)$. If $|\rho| < 1$, then $\tilde{m}_n(k) \xrightarrow{m.s.} m_0(k)$ as $n \to \infty$, for $k = 1, 2$.

*Proof.* From (2.2) and (4.1), it is immediate that $\tilde{m}_n(k)$ is unbiased. Then, to prove Theorem 4.1, it remains to show $\mathrm{Var}\{\tilde{m}_n(k)\} \to 0$ as $n \to \infty$ (see Brockwell and Davis, 1991, Proposition 6.2.4, p. 203). For $k = 1, 2$, it can be written

$$\sum_{t=0}^{n} \{X_t - m_0(k)\}\, \mathbb{I}_k(s_t) = \sum_{t=0}^{N} \underline{c}_k' \underline{Y}_t, \tag{4.4}$$

where $\underline{Y}_N = \{0, X_{7N-1} - m_0(2), X_{7N-2} - m_0(2), X_{7N-3} - m_0(1), \ldots, X_{7N-6} - m_0(1)\}'$, $\underline{c}_1 = (1, 0, 0, 1, 1, 1, 1)'$ and $\underline{c}_2 = (0, 1, 1, 0, 0, 0, 0)'$. Denote

$$n_k = \sum_{t=0}^{n} \mathbb{I}_k(s_t) \tag{4.5}$$

the number of dates in Regime $k$ (we have $n_1 = 5N$ and $n_2 = 2N$). Therefore, from (4.1) and (4.4), we have

$$\begin{aligned}
\xi^2_{k,n_k} &= \mathrm{Var}\left\{\sqrt{n_k}\tilde{m}_n(k)\right\} \\
&= \mathrm{Var}\left[\frac{1}{\sqrt{n_k}} \sum_{t=0}^{n} \{X_t - m_0(k)\}\, \mathbb{I}_k(s_t)\right] \\
&= \frac{1}{n_k} \underline{c}_k' \mathrm{Var}\left(\sum_{t=0}^{N} \underline{Y}_t\right) \underline{c}_k \\
&= \frac{1}{n_k} \underline{c}_k' \left\{ \sum_{t=0}^{N} \mathrm{Var}\left(\underline{Y}_t\right) + \sum_{\substack{i,j=0 \\ i \neq j}}^{N} \mathrm{Cov}\left(\underline{Y}_i, \underline{Y}_j\right) \right\} \underline{c}_k.
\end{aligned}$$

Asymptotically, under (3.6), $\underline{Y}_t$ is stable. We thence obtain

$$\begin{aligned}
\xi^2_k &= \lim_{n_k \to \infty} \xi^2_{k,n_k} \\
&= \frac{1}{5\mathbb{I}_1(k) + 2\mathbb{I}_2(k)} \underline{c}_k' \left[\Gamma_{\underline{Y}}(0) + \sum_{i=1}^{\infty} \left\{\Gamma_{\underline{Y}}(i) + \Gamma'_{\underline{Y}}(i)\right\}\right] \underline{c}_k \\
&= \frac{1}{5\mathbb{I}_1(k) + 2\mathbb{I}_2(k)} \underline{c}_k' \left[\Gamma_{\underline{Y}}(0) + \sum_{i=1}^{\infty} \left\{A^i \Gamma_{\underline{Y}}(0) + \Gamma_{\underline{Y}}(0)(A^i)'\right\}\right] \underline{c}_k \\
&= \frac{1}{5\mathbb{I}_1(k) + 2\mathbb{I}_2(k)} \underline{c}_k' \left[\Gamma_{\underline{Y}}(0) + \sum_{i=1}^{\infty} \left\{\rho^{i-1} A \Gamma_{\underline{Y}}(0) + \Gamma_{\underline{Y}}(0) \rho^{i-1} A'\right\}\right] \underline{c}_k,
\end{aligned}$$

which finally gives

$$\xi_k^2 = \frac{1}{5\mathbb{I}_1(k) + 2\mathbb{I}_2(k)} \underline{c}_k' \left[ \Gamma_{\underline{Y}}(0) + (1-\rho)^{-1} \left\{ A\Gamma_{\underline{Y}}(0) + \Gamma_{\underline{Y}}(0)A' \right\} \right] \underline{c}_k. \tag{4.6}$$

The right-hand side of (4.6) is obtained from Section 3 as follows,

$$\begin{aligned} \xi_1^2 \;=\; & 7\left\{25(\rho-1)^2\right\}^{-1} \Big[ \{5 + 2a_0(1)(4 + a_0^2(2)) + a_0^4(1)(3 + 2a_0^2(2))^2 \\ & +2a_0^3(1)(5 + 4a_0^2(2) + a_0^4(2)) + a_0^2(1)(10 + 4a_0^2(2) + a_0^4(2)) \\ & +2a_0^7(1)(1 + a_0^2(2) + 3a_0^4(2)) + 2a_0^5(1)(3 + 4a_0^2(2) + 3a_0^4(2)) \\ & +a_0^8(1)(1 + 4a_0^4(2)) + a_0^6(1)(4 + 4a_0^2(2) + 7a_0^4(2))\} \, \sigma_0^2(1) \\ & +\sigma_0^2(2)a_0^2(1) \left\{1 + a_0^2(2)\right\} \left\{1 + a_0(1) + a_0^2(1) + a_0^3(1) + a_0^4(1)\right\}^2 \Big], \end{aligned} \tag{4.7}$$

and

$$\begin{aligned} \xi_2^2 \;=\; & 7\left\{4(\rho-1)^2\right\}^{-1} \Big[ \left\{1 + a_0^2(1) + a_0^4(1) + a_0^6(1) + a_0^8(1)\right\} \\ & \left\{2a_0^3(2)\sigma_0^2(1) + a_0^4(2)\sigma_0^2(2) + a_0^2(2)\sigma_0^2(1)\right\} + 2\sigma_0^2(2) \\ & +2\left\{1 + a_0^5(1)\right\} a_0(2)\sigma_0^2(2) + \left\{1 + a_0^{10}(1)\right\} a_0^2(2)\sigma_0^2(2) \Big]. \end{aligned} \tag{4.8}$$

Finally, from (4.1), (4.6), (4.7), (4.8) and under (3.6), we obtain

$$\lim_{n\to\infty} \mathrm{Var}\left\{\tilde{m}_n(k)\right\} = \lim_{n_k\to\infty} \frac{\xi_{k,n_k}^2}{n_k} = 0,$$

which completes the proof of Theorem 4.1. $\square$

The following theorem shows that $\tilde{m}_n(k)$ is a root-$n$ consistent estimator of $m_0(k)$, and gives its limiting distribution.

THEOREM 4.2. Under the assumptions of Theorem 4.1, we have

$$\sqrt{n}\left\{\tilde{m}_n(k) - m_0(k)\right\} \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}(0, \xi_k^2 \pi_k^{-1}) \qquad \text{as } n \to \infty,$$

where $\xi_k^2$ is given by (4.7)–(4.8), $\pi_k = n_k/n$, and $n_k$ is given by (4.5).

*Proof.* The whole notations introduced in the proof of Theorem 4.1 still hold. To prove Theorem 4.2, we use the Lindeberg central limit theorem for triangular arrays of random variables (see Billingsley, 1995, Theorem 27.2, p. 359). Note that $(n\pi_k)^{1/2} = (n_k)^{1/2}$. First, in view of (2.2), there exist some coefficients $\alpha_{\ell,n}^{(k)}$ such that

$$\sum_{t=0}^{n} \{X_t - m_0(k)\} \mathbb{I}_k(s_t) = \sum_{t=0}^{n} \left\{ \sum_{i=0}^{t} \phi_{t,i}(\underline{\theta}_0)\varepsilon_{t-i} \right\} \mathbb{I}_k(s_t) = \sum_{\ell=0}^{n} \alpha_{\ell,n}^{(k)} \varepsilon_\ell. \tag{4.9}$$

For example, $\alpha_{\ell,n}^{(1)} = \mathbb{I}_1(s_\ell) + \sum_{t'=\ell+1}^{n} a_0(s_{t'}) \cdots a_0(s_{\ell+1}) \mathbb{I}_1(s_{t'})$. Therefore, from (4.1) and (4.9), we have

$$\sqrt{n_k}\frac{\tilde{m}_n(k) - m_0(k)}{\xi_k} = \sum_{\ell=0}^{n} \frac{\alpha_{\ell,n}^{(k)}\varepsilon_\ell}{\sqrt{n_k}\xi_{k,n_k}} \times \frac{\xi_{k,n_k}}{\xi_k}$$

where $\lim_{n_k\to\infty} \xi_{k,n_k}/\xi_k = 1$ in view of (4.6), and $\sum_{\ell=0}^{n} \alpha_{\ell,n}^{(k)} \varepsilon_\ell (\sqrt{n_k}\xi_{k,n_k})^{-1}$ denotes the sum of some independent, but not identically distributed random variables. The collection made by the elements of this sum is called a triangular array of random variables. It suffices now to check the Lindeberg condition. For all $\zeta > 0$, since $\varepsilon_t = \sigma_0(s_t)\eta_t$ with $(\eta_t)$ iid, we have

$$\lim_{n\to\infty} \sum_{\ell=0}^{n} \int_{\left\{\left|\frac{\alpha_{\ell,n}^{(k)}\varepsilon_\ell}{\sqrt{n_k}\xi_{k,n_k}}\right| \geq \zeta\right\}} \left\{ \frac{\alpha_{\ell,n}^{(k)}\varepsilon_\ell}{\sqrt{n_k}\xi_{k,n_k}} \right\}^2 d\mathbb{P} = \lim_{n\to\infty} \sum_{\ell=0}^{n} \frac{\left\{\alpha_{\ell,n}^{(k)}\right\}^2 \sigma_0^2(s_\ell)}{n_k \xi_{k,n_k}^2} \int_{\left\{|\eta_1| \geq \left|\frac{\sqrt{n_k}\xi_{k,n_k}\zeta}{\alpha_{\ell,n}^{(k)}\sigma_0(s_\ell)}\right|\right\}} \eta_1^2 d\mathbb{P},$$

where

$$\sum_{\ell=0}^{n} \frac{\left\{\alpha_{\ell,n}^{(k)}\right\}^2 \sigma_0^2(s_\ell)}{n_k \xi_{k,n_k}^2} = 1 \tag{4.10}$$

from (4.9) and (4.6). Besides, there exists $0 < \alpha^{(k)} < \infty$ such that $|\alpha_{\ell,n}^{(k)}| \leq \alpha^{(k)}$ [because of the condition (3.6)], and so

$$\left\{ |\eta_1| \geq \left| \frac{\sqrt{n_k}\xi_{k,n_k}\zeta}{\alpha_{\ell,n}^{(k)}\sigma_0(s_\ell)} \right| \right\} \subseteq \left\{ |\eta_1| \geq \frac{\sqrt{n_k}|\xi_{k,n_k}|\zeta}{\alpha^{(k)} \max_{k=1,\ldots,d} |\sigma_0(k)|} \right\}.$$

Therefore

$$\lim_{n\to\infty} \sum_{\ell=0}^{n} \frac{\left\{\alpha_{\ell,n}^{(k)}\right\}^2 \sigma_0^2(s_\ell)}{n_k \xi_{k,n_k}^2} \int_{\left\{|\eta_1| \geq \left|\frac{\sqrt{n_k}\xi_{k,n_k}\zeta}{\alpha_{\ell,n}^{(k)}\sigma_0(s_\ell)}\right|\right\}} \eta_1^2 d\mathbb{P} \leq \lim_{n\to\infty} \int_{\left\{|\eta_1| \geq \frac{\sqrt{n_k}|\xi_{k,n_k}|\zeta}{\alpha^{(k)} \max_{k=1,\ldots,d}|\sigma_0(k)|}\right\}} \eta_1^2 d\mathbb{P}, \tag{4.11}$$

and

$$\left\{ |\eta_1| \geq \frac{\sqrt{n_k}|\xi_{k,n_k}|\zeta}{\alpha^{(k)} \max_{k=1,\ldots,d} |\sigma_0(k)|} \right\} \to \emptyset \quad \text{as } n_k \to \infty. \tag{4.12}$$

Moreover,

$$\int_{\Omega} \eta_1^2 d\mathbb{P} < \infty. \tag{4.13}$$

Finally, in view of (4.10), (4.11), (4.12) and (4.13), the Lebesgue's dominated convergence theorem implies

$$\lim_{n\to\infty} \int_{\left\{|\eta_1| \geq \frac{\sqrt{n_k}|\xi_{k,n_k}|\zeta}{\alpha^{(k)} \max_{k=1,\ldots,d}|\sigma_0(k)|}\right\}} \eta_1^2 d\mathbb{P} = 0$$

for $\zeta > 0$, which satisfies the Lindeberg condition. $\square$

We now turn to the estimator $\tilde{a}_n(k)$ defined by (4.2). We have the following result.

THEOREM 4.3. Consider model (2.1) for which $\tilde{a}_n(k)$, defined by (4.2), is an estimator of $a_0(k)$. Let $\tilde{a}_n^*(k)$, defined by (4.3), be a pseudo-estimator of $a_0(k)$. If $|\rho| < 1$, then

$$\sqrt{n}\left\{\tilde{a}_n(k) - \tilde{a}_n^*(k)\right\} \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \to \infty,$$

for $k = 1, 2$.

*Proof.* Using (4.2), (4.3), and the elementary relation $ab^{-1} - cd^{-1} = (a-c)b^{-1} + c(b^{-1} - d^{-1})$, we can write

$$\sqrt{n}\left\{\tilde{a}_n(k) - \tilde{a}_n^*(k)\right\} = c_n(k) + d_n(k) \tag{4.14}$$

where

$$c_n(k) = c_n^{(1)}(k) c_n^{(2)}(k) \tag{4.15}$$

with

$$\begin{aligned} c_n^{(1)}(k) &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \mathbb{I}_k(s_t) \left[ X_t \left\{ m_0(s_{t-1}) - \tilde{m}_n(s_{t-1}) \right\} + X_{t-1} \left\{ m_0(k) - \tilde{m}_n(k) \right\} \right. \\ & \quad \left. + \tilde{m}_n(s_{t-1}) \left\{ \tilde{m}_n(k) - m_0(k) \right\} + m_0(k) \left\{ \tilde{m}_n(s_{t-1}) - m_0(s_{t-1}) \right\} \right], \\ c_n^{(2)}(k) &= \left[ \frac{1}{n} \sum_{t=1}^{n} \left\{ X_{t-1} - \tilde{m}_n(s_{t-1}) \right\}^2 \mathbb{I}_k(s_t) \right]^{-1}, \end{aligned}$$

and

$$d_n(k) = d_n^{(1)}(k) d_n^{(2)}(k)$$

with

$$
\begin{aligned}
d_n^{(1)}(k) &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{X_t - m_0(k)\}\{X_{t-1} - m_0(s_{t-1})\} \mathbb{I}_k(s_t), \\
d_n^{(2)}(k) &= \left[\frac{1}{n} \sum_{t=1}^{n} \{X_{t-1} - \tilde{m}_n(s_{t-1})\}^2 \mathbb{I}_k(s_t)\right]^{-1} - \left[\frac{1}{n} \sum_{t=1}^{n} \{X_{t-1} - m_0(s_{t-1})\}^2 \mathbb{I}_k(s_t)\right]^{-1}.
\end{aligned}
$$

In view of Theorem 4.1, the differences of the form $m_0(k) - \tilde{m}_n(k)$ converge in probability to 0 as $n \to \infty$. We thence can deduce that

$$
c_n^{(1)}(k) = o_{\mathbb{P}}(1). \tag{4.16}
$$

Moreover, we have

$$
\frac{1}{n} \sum_{t=1}^{n} \{X_{t-1} - \tilde{m}_n(s_{t-1})\}^2 \xrightarrow[n\to\infty]{} \sum_{k=1}^{2} \pi_k E\{X_t \mathbb{I}_k(s_t) - \tilde{m}_n(k)\}^2 > 0,
$$

where the last inequality results from Subsection 3.2. This entails that there exists $\delta > 0$ such that

$$
c_n^{(2)}(k) \to \delta \quad \text{as } n \to \infty. \tag{4.17}
$$

Therefore, in view of (4.16) and (4.17), $c_n(k)$ defined by (4.15) is such that $c_n(k) = o_{\mathbb{P}}(1)$. Similar arguments show that $d_n(k) = o_{\mathbb{P}}(1)$. In view of (4.14), we finally obtain the result claimed in Theorem 4.3. $\square$

REMARK 4.1. As a consequence to Theorem 4.3, the variable $\sqrt{n}\{\tilde{a}_n(k) - a_0(k)\}$ has a limiting normal distribution with asymptotic covariance matrix $\Sigma^*$ given, for example, by Bibi and Francq (2003, Subsection 3.1). This comment can be straightforwardly deduced by the fact that the variable $X_t - m_0(s_t)$ satisfies model (2.1) with $m_0(\cdot) = 0$.

### 4.2. A LEAST SQUARES PROCEDURE

In this subsection, we investigate the problem of estimating model (2.1) via an LS procedure. The method is based on the minimization of a sum of squares of residuals (see, *e.g.*, Godambe and Heyde, 1987, for a general reference). Note that the direct study of the

LS estimator of the parameter $\underline{\theta}_0$ of model (2.1) is not obvious because the model is not stationary.

### 4.2.1. DEFINITION OF LEAST SQUARES ESTIMATORS

Let $\underline{\theta} = \{\theta(1), \theta(2), \theta(3), \theta(4)\}' = \{m(1), m(2), a(1), a(2)\}'$ belong to $\Theta$ and let $\Theta^*$ be a compact subset of $\Theta$ which contains a neighborhood of $\underline{\theta}_0$. Given a sequence $(X_0, \ldots, X_n)$ of observations, we define an LS estimator as any measurable solution of

$$\underline{\hat{\theta}}_n = \arg\min_{\underline{\theta} \in \Theta^*} Q_n(\underline{\theta}), \quad Q_n(\underline{\theta}) = \frac{1}{2n} \sum_{t=1}^{n} e_t^2(\underline{\theta}), \tag{4.18}$$

where the prediction errors are given by

$$e_t(\underline{\theta}) = X_t - m(s_t) - a(s_t)\{X_{t-1} - m(s_{t-1})\}. \tag{4.19}$$

From (4.19), the LS estimator $\underline{\hat{\theta}}_n = \{\hat{m}_n(1), \hat{m}_n(2), \hat{a}_n(1), \hat{a}_n(2)\}'$, defined by (4.18), is clearly non-linear with respect to the parameters. Therefore, unlike the LS estimator in the mean-corrected case, it is worth noting that it seems difficult to give $\underline{\hat{\theta}}_n$ in an explicit form like (4.3). Just notice that $\underline{\hat{\theta}}_n$ can be obtained by estimating the VAR(1) model (2.3) subject to linear and non-linear constraints (coming from the special form of $A$). In practice, the computation of $\underline{\hat{\theta}}_n$ requires numerical optimization routines, such as gradient, quasi-Newton or Simplex methods. In Section 5, the optimization method which is used for the Monte Carlo experiments is the 'Downhill Simplex Method' (see Press et al., 1986, for a precise reference). An initial value of $\underline{\theta}$ must be specified: clearly it is possible to take $0 \in \mathbb{R}^4$ as initial value for the LS procedure. The FS estimator $\underline{\tilde{\theta}}_n$ also seems to be relevant to initialize the LS procedure.

### 4.2.2. ASYMPTOTIC BEHAVIOUR

This part presents a result on consistency and asymptotic normality of the LS estimators $\underline{\hat{\theta}}_n$. The asymptotic covariance matrix is derived. For any $\underline{\theta} \in \Theta$, we consider the following matrices

$$I(\underline{\theta}) = \lim_{n\to\infty} \mathrm{Var}\left\{\sqrt{n}\frac{\partial}{\partial \underline{\theta}} Q_n(\underline{\theta})\right\} \quad \text{and} \quad J(\underline{\theta}) \overset{a.s.}{=} \lim_{n\to\infty} \frac{\partial^2}{\partial \underline{\theta} \partial \underline{\theta}'} Q_n(\underline{\theta}).$$

For simplicity, we will omit the notation $\underline{\theta}$ in all quantities taken at the true value $\underline{\theta}_0$. Hence $I = I(\underline{\theta}_0)$, $J = J(\underline{\theta}_0)$, and $e_t = e_t(\underline{\theta}_0)$.

THEOREM 4.4. Consider model (2.1) and let $\hat{\underline{\theta}}_n$, defined by (4.18), be an LS estimator of $\underline{\theta}_0$. If $|\rho| < 1$, then

$$\hat{\underline{\theta}}_n \xrightarrow{a.s.} \underline{\theta}_0 \quad \text{and} \quad \sqrt{n}(\hat{\underline{\theta}}_n - \underline{\theta}_0) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}(0, \Sigma) \quad \text{as } n \to \infty,$$

with limiting covariance matrix $\Sigma = J^{-1} I J^{-1}$.

*Proof.* This result comes from a straightforward adaptation of the proofs of Theorems 1–2 in Bibi and Francq (2003). $\square$

It remains to compute the asymptotic covariance matrix $\Sigma$ involved in Theorem 4.4. First note that $e_t = \sigma_0(s_t)\eta_t$, and from (4.19),

$$\frac{\partial e_t}{\partial \underline{\theta}} = \begin{pmatrix} -\mathbb{I}_1(s_t) + a_0(s_t)\mathbb{I}_1(s_{t-1}) \\ -\mathbb{I}_2(s_t) + a_0(s_t)\mathbb{I}_2(s_{t-1}) \\ -\mathbb{I}_1(s_t)\left\{X_{t-1} - m_0(s_{t-1})\right\} \\ -\mathbb{I}_2(s_t)\left\{X_{t-1} - m_0(s_{t-1})\right\} \end{pmatrix}.$$

We take an interest to the information matrix $I$. Pointing out that each day of the week can be clearly identified by a product of five indicator functions of the form $\mathbb{I}_k(s_t)$ (for instance, we can say that the date $t$ corresponds either to Monday if $\mathbb{I}_1(s_t)\mathbb{I}_2(s_{t-1})\mathbb{I}_2(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_1(s_{t-4}) = 1$, or to Sunday if $\mathbb{I}_2(s_t)\mathbb{I}_2(s_{t-1})\mathbb{I}_1(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_1(s_{t-4}) = 1$ etc.) and since $e_t = \sigma_0(s_t)\eta_t$ with

$(\eta_t)$ iid, mean zero, variance one, and $\eta_t$ is independent of $X_u$ for $u < t$, we have

$$
\begin{aligned}
I &= \lim_{n\to\infty} \mathrm{Var}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} e_t \frac{\partial e_t}{\partial \underline{\theta}}\right) \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \mathrm{Var}\left(e_t \frac{\partial e_t}{\partial \underline{\theta}}\right) \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E\left(e_t^2 \frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\right) \\
&= \frac{1}{7}\left[\sigma_0^2(1)E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_1(s_t)\mathbb{I}_2(s_{t-1})\mathbb{I}_2(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_1(s_{t-4})\right\}\right. \\
&\quad +\sigma_0^2(1)E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_1(s_t)\mathbb{I}_1(s_{t-1})\mathbb{I}_2(s_{t-2})\mathbb{I}_2(s_{t-3})\mathbb{I}_1(s_{t-4})\right\} \\
&\quad +\sigma_0^2(1)E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_1(s_t)\mathbb{I}_1(s_{t-1})\mathbb{I}_1(s_{t-2})\mathbb{I}_2(s_{t-3})\mathbb{I}_2(s_{t-4})\right\} \\
&\quad +\sigma_0^2(1)E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_1(s_t)\mathbb{I}_1(s_{t-1})\mathbb{I}_1(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_2(s_{t-4})\right\} \\
&\quad +\sigma_0^2(1)E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_1(s_t)\mathbb{I}_1(s_{t-1})\mathbb{I}_1(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_1(s_{t-4})\right\} \\
&\quad +\sigma_0^2(2)E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_2(s_t)\mathbb{I}_1(s_{t-1})\mathbb{I}_1(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_1(s_{t-4})\right\} \\
&\quad \left.+\sigma_0^2(2)E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_2(s_t)\mathbb{I}_2(s_{t-1})\mathbb{I}_1(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_1(s_{t-4})\right\}\right].
\end{aligned}
$$

In view of the asymptotic autocovariance structure of model (2.1) obtained in Subsection 3.2, simple computations give the previous expectations. For example, we have

$$
E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_1(s_t)\mathbb{I}_2(s_{t-1})\mathbb{I}_2(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_1(s_{t-4})\right\} = \begin{pmatrix} 1 & -a_0(1) & 0 & 0 \\ -a_0(1) & a_0^2(1) & 0 & 0 \\ 0 & 0 & \gamma_{2,2}(0) & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

for Mondays, and

$$E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_2(s_t)\mathbb{I}_2(s_{t-1})\mathbb{I}_1(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_1(s_{t-4})\right\} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \{a_0(2)-1\}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \gamma_{3,3}(0) \end{pmatrix}$$

for Sundays. Denote by $M_{i,j}$ the element of the $i$-th row and $j$-th column of a matrix $M$. Then we deduce the following expression for matrix $I$

$$I = \frac{1}{7}\begin{pmatrix} I^{(11)} & 0 \\ 0 & I^{(22)} \end{pmatrix}, \tag{4.20}$$

where the elements of the symmetric $(2 \times 2)$-matrix $I^{(11)}$ are

$$\begin{aligned}
I^{(11)}_{1,1} &= \sigma_0^2(2)a_0^2(2) + \sigma_0^2(1)\left[1 + 4\{a_0(1)-1\}^2\right], \\
I^{(11)}_{2,2} &= \sigma_0^2(1)a_0^2(1) + \sigma_0^2(2)\left[1 + \{a_0(2)-1\}^2\right], \\
I^{(11)}_{1,2} &= -\left\{\sigma_0^2(1)a_0(1) + \sigma_0^2(2)a_0(2)\right\},
\end{aligned}$$

and the $(2 \times 2)$-matrix $I^{(22)}$ is diagonal with

$$\begin{aligned}
I^{(22)}_{1,1} &= \sigma_0^2(1)\left(1-\rho^2\right)^{-1}\Big[\sigma_0^2(1)\left\{4 + a_0^4(2) + 3a_0^2(1) + 2a_0^4(1) + a_0^6(1) + 2a_0^2(1)a_0^4(2)\right. \\
&\quad \left. + 3a_0^4(1)a_0^4(2) + 4a_0^6(1)a_0^4(2) + 5a_0^8(1)a_0^4(2)\right\} \\
&\quad + \sigma_0^2(2)\left\{1 + a_0^2(1) + a_0^4(1) + a_0^6(1) + a_0^8(1)\right\}\left\{1 + a_0^2(2)\right\}\Big], \\
I^{(22)}_{2,2} &= \sigma_0^2(2)\left(1-\rho^2\right)^{-1}\Big[\sigma_0^2(1)\left\{1 + a_0^2(1) + a_0^4(1) + a_0^6(1) + a_0^8(1)\right\}\left\{1 + a_0^2(2)\right\} \\
&\quad + \sigma_0^2(2)\left[1 + a_0^{10}(1)\left\{1 + 2a_0^2(2)\right\}\right]\Big].
\end{aligned}$$

We now concentrate on the information matrix $J$. From (4.18) and following similar arguments to those given to compute the matrix $I$, we have

$$\begin{aligned}
J &\overset{a.s.}{=} \lim_{n\to\infty} n^{-1}\sum_{t=1}^{n}\left(\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'} + e_t\frac{\partial^2 e_t}{\partial\underline{\theta}\partial\underline{\theta}'}\right)\\
&= E\left(\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\right)\\
&= \frac{1}{7}\left[E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_1(s_t)\mathbb{I}_2(s_{t-1})\mathbb{I}_2(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_1(s_{t-4})\right\}\right.\\
&\quad +E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_1(s_t)\mathbb{I}_1(s_{t-1})\mathbb{I}_2(s_{t-2})\mathbb{I}_2(s_{t-3})\mathbb{I}_1(s_{t-4})\right\}\\
&\quad +E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_1(s_t)\mathbb{I}_1(s_{t-1})\mathbb{I}_1(s_{t-2})\mathbb{I}_2(s_{t-3})\mathbb{I}_2(s_{t-4})\right\}\\
&\quad +E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_1(s_t)\mathbb{I}_1(s_{t-1})\mathbb{I}_1(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_2(s_{t-4})\right\}\\
&\quad +E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_1(s_t)\mathbb{I}_1(s_{t-1})\mathbb{I}_1(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_1(s_{t-4})\right\}\\
&\quad +E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_2(s_t)\mathbb{I}_1(s_{t-1})\mathbb{I}_1(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_1(s_{t-4})\right\}\\
&\quad \left.+E\left\{\frac{\partial e_t}{\partial \underline{\theta}}\frac{\partial e_t}{\partial \underline{\theta}'}\mathbb{I}_2(s_t)\mathbb{I}_2(s_{t-1})\mathbb{I}_1(s_{t-2})\mathbb{I}_1(s_{t-3})\mathbb{I}_1(s_{t-4})\right\}\right]
\end{aligned}$$

since $E(e_t) = 0$ and

$$\frac{\partial^2 e_t}{\partial\underline{\theta}\partial\underline{\theta}'} = \begin{pmatrix} 0 & 0 & \mathbb{I}_1(s_t)\mathbb{I}_1(s_{t-1}) & \mathbb{I}_2(s_t)\mathbb{I}_1(s_{t-1}) \\ 0 & 0 & \mathbb{I}_1(s_t)\mathbb{I}_2(s_{t-1}) & \mathbb{I}_2(s_t)\mathbb{I}_2(s_{t-1}) \\ \mathbb{I}_1(s_t)\mathbb{I}_1(s_{t-1}) & \mathbb{I}_1(s_t)\mathbb{I}_2(s_{t-1}) & 0 & 0 \\ \mathbb{I}_2(s_t)\mathbb{I}_1(s_{t-1}) & \mathbb{I}_2(s_t)\mathbb{I}_2(s_{t-1}) & 0 & 0 \end{pmatrix},$$

which is independent of $\eta_t$. The expectations above have been already computed for the matrix $I$. Therefore, $J$ is obtained in the form

$$J = \frac{1}{7}\begin{pmatrix} J^{(11)} & 0 \\ 0 & J^{(22)} \end{pmatrix}, \tag{4.21}$$

where

$$\begin{aligned}
J^{(11)}_{1,1} &= 1 + a_0^2(2) + 4\{a_0(1) - 1\}^2, \quad J^{(11)}_{2,2} = 1 + a_0^2(1) + \{a_0(2) - 1\}^2,\\
J^{(11)}_{1,2} &= J^{(11)}_{2,1} = -\{a_0(1) + a_0(2)\},\\
J^{(22)}_{1,1} &= \sigma_0^{-2}(1)I^{(22)}_{1,1}, \quad J^{(22)}_{2,2} = \sigma_0^{-2}(2)I^{(22)}_{2,2}, \quad J^{(22)}_{1,2} = J^{(22)}_{2,1} = 0.
\end{aligned}$$

Finally, in view of (4.20) and (4.21), we obtain $\Sigma = J^{-1}IJ^{-1}$ in the form

$$\Sigma = \begin{pmatrix} \Sigma^{(11)} & 0 \\ 0 & \Sigma^{(22)} \end{pmatrix},$$

with computable analytic expressions for $\Sigma_{i,j}^{(11)}$ (that we do not report hereafter for reasons of space), and

$$\begin{aligned}
\Sigma_{1,1}^{(22)} &= 7\left(1-\rho^2\right)\sigma_0^2(1)\left[\left\{4+a_0^4(2)+5a_0^8(1)a_0^4(2)+a_0^2(1)\left\{3+2a_0^4(2)\right\}\right.\right.\\
&\quad \left.+a_0^4(1)\left\{2+3a_0^4(2)\right\}+a_0^6(1)\left\{1+4a_0^4(2)\right\}\right\}\sigma_0^2(1)\\
&\quad \left.+\left\{1+a_0^2(1)+a_0^4(1)+a_0^6(1)+a_0^8(1)\right\}\left\{1+a_0^2(2)\right\}\sigma_0^2(2)\right]^{-1},\\
\Sigma_{2,2}^{(22)} &= 7\left(1-\rho^2\right)\sigma_0^2(2)\left[\left\{1+a_0^2(1)+a_0^4(1)+a_0^6(1)+a_0^8(1)\right\}\left\{1+a_0^2(2)\right\}\sigma_0^2(1)\right.\\
&\quad \left.+\left\{1+a_0^{10}(1)\left\{1+2a_0^2(2)\right\}\right\}\sigma_0^2(2)\right]^{-1},\\
\Sigma_{1,2}^{(22)} &= \Sigma_{2,1}^{(22)} \;=\; 0.
\end{aligned}$$

4.3. COMPARISON OF THE ASYMPTOTIC ACCURACY OF BOTH METHODS

The following result is an immediate consequence of Theorems 4.2–4.4.

THEOREM 4.5. Under the assumptions and with the notations of Theorems 4.2–4.4, we have $\Sigma^{(22)} = \Sigma^*$. Moreover, if $\sigma_0(1) = \sigma_0(2)$, then

$$\Sigma_{1,1}^{(11)} \leq \xi_1^2\pi_1^{-1} \quad \text{and} \quad \Sigma_{2,2}^{(11)} \leq \xi_2^2\pi_2^{-1}.$$

Theorem 4.5 shows, in a first part, that the LS and the FS estimates are equivalent related to their asymptotic accuracy concerning the estimation of the AR coefficients, whatever the value of the nuisance parameters $\sigma_0(\cdot)$. In a second part, Theorem 4.5 focuses on the homoskedastic case [*i.e.* $\sigma_0(1) = \sigma_0(2)$] and states that the gain in terms of asymptotic accuracy for the LS estimate compared with the FS estimate is concentrated on the position parameters $m_0(\cdot)$. The efficient gain can be significant [see Figures 4.1–4.2 for the case $\underline{\sigma}_0 = (1.0, 1.0)'$]. For instance, if $\underline{\theta}_0 = (1.0, 0.0, 0.9, -0.9)'$ and $\underline{\sigma}_0 = (1.0, 1.0)'$, then we have

$\Sigma^{(11)}_{1,1} = 3.7838$ and $\xi_1^2\pi_1^{-1} = 100.439$, $\Sigma^{(11)}_{2,2} = 1.2915$ and $\xi_2^2\pi_2^{-1} = 1.6545$. Let us mention that, when $m_0(1) = m_0(2)$ and $a_0(1) = a_0(2)$ (case with only one regime), the LS and FS methods are asymptotically equivalent: we have $\Sigma^{(11)}_{k,k} = \xi_k^2\pi_k^{-1}$, as expected. Moreover, the inequalities involved in Theorem 4.5 become strict when $a_0(1) \neq a_0(2)$.

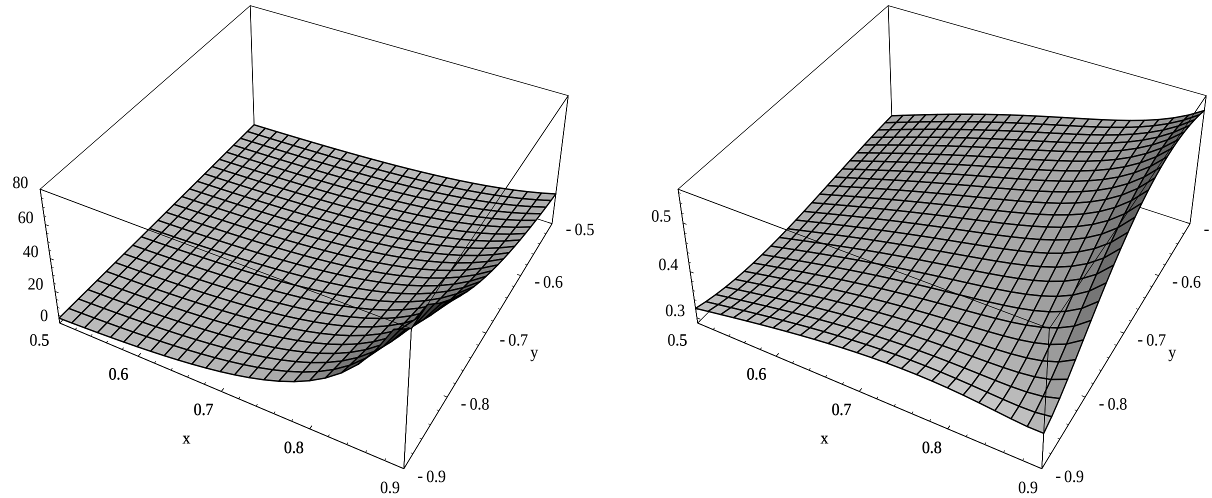


Figure 4.1: in the left graph, for model (2.1) with $\underline{\sigma}_0 = (1.0, 1.0)'$, difference between $\xi_1^2\pi_1^{-1}$ [the asymptotic variance of $\tilde{m}_n(1)$] and $\Sigma^{(11)}_{1,1}$ [the asymptotic variance of $\hat{m}_n(1)$] as function of $a_0(1)$ (varying from 0.5 to 0.9 on the x-axis) and $a_0(2)$ (varying from -0.9 to -0.5 on the y-axis). In the right graph, as in the previous one, but for the asymptotic variance of $\tilde{m}_n(2)$ and $\hat{m}_n(2)$.

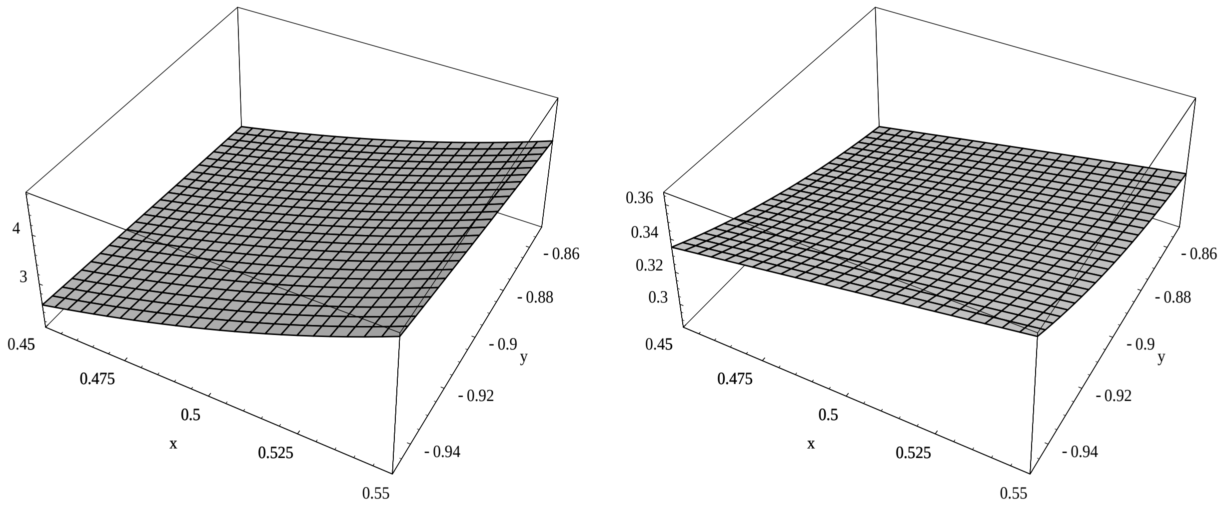


Figure 4.2: as in Figure 4.1, but with $a_0(1)$ varying from 0.45 to 0.55 on the x-axis and $a_0(2)$ varying from -0.95 to -0.85 on the y-axis.

REMARK 4.2. Such a result on the position parameters estimation does not hold any more when $\sigma_0(1) \neq \sigma_0(2)$. In that case, as in the time-constant coefficients setting, the LS estimator has no optimal properties any more. To test for heteroskedasticity, one shall derive the joint law of $\{\hat{\sigma}_n(1), \hat{\sigma}_n(2)\}'$, where $\hat{\sigma}_n(k)$ is a consistent estimator of $\sigma_0(k)$ given by

$$\hat{\sigma}_n^2(k) = n_k^{-1} \sum_{t=1}^{n} e_t^2(\underline{\hat{\theta}}_n) \mathbb{I}_k(s_t)$$

It would also be of particular interest to consider quasi-generalized LS (QGLS)- or ML-type estimators that insert the nuisance parameter in their criterion. The reader is referred to Francq and Gautier (2004a) for further developments on this topic.

As for all asymptotic results, the validity of our theorems for approximating the distribution of the estimators in small samples can be legitimately questioned. We propose, in the next section, some Monte Carlo experiments aimed to illustrate the performance of the LS estimator compared with the performance of the FS estimator in finite samples.

## 5. NUMERICAL ILLUSTRATIONS

To gauge the proposed estimation methods, this section numerically investigates the finite sample properties of previous estimators using Monte Carlo experiments. We carried out simulations for model (2.1) with $\underline{\sigma}_0 = (1.0, 1.0)'$, initial conditions $X_{-1} - m(s_{-1}) = \eta_{-1} = 0$, and $(\eta_t)$ is an iid $\mathcal{N}(0, 1)$ process. We successively generated 5,000 independent trajectories of size $n = 100$, $n = 500$ and $n = 10,000$ of the model. For each trajectory, the parameters $m(1)$, $m(2)$, $a(1)$ and $a(2)$ have been estimated using the two methods involved in the paper. In Tables 1–3 below corresponding respectively to $n = 100$, $n = 500$ and $n = 10,000$, we report the simulation results for the 5,000 replications. Each table is split into two sub-tables according to specific values of $\underline{\theta}_0$ chosen for the simulation experiments. Within Tables 1–3, the column 'param.' indicates the parameter to be estimated; the column 'meth.' gives the estimation method which has been used (either FS or LS); the column 'bias' indicates the average of the difference between the true value and the estimate over the 5,000 replications; the column 'st. dev.' gives the standard deviations of the difference between the true value and the estimate over the 5,000 replications; the column 'RMSE' contains the root mean squared errors (RMSE) over the 5,000 simulations; the column 'tr. val. (AV)' contains the value of the asymptotic variances computed from the asymptotic theory presented in Theorems 4.2–4.4; the column '$n\times$MSE' reports the computation of $n$ times the mean squared errors (MSE) over the 5,000 replications; finally, the column 'mean (AV)' corresponds to the average of the estimate of the asymptotic variance over the 5,000 simulations.

Overall in this study, in accordance with the asymptotic theory, Tables 1–3 point out the superiority of the LS estimator over the FS estimator for finite sample sizes. Indeed, the LS estimator of $m(k)$ is much more accurate than the FS estimator, and the two estimators are equivalent for the estimation of $a(k)$. When the sample size increases, the quality of the inference improves rapidly. It is seen that $\hat{m}_n(k)$ is always more efficient than $\tilde{m}_n(k)$, and that the efficiency of $\hat{m}_n(k)$ versus $\tilde{m}_n(k)$ increases as $n$ increases. As expected, the bias of our estimates is more and more zero since the sample size increases. When $n = 500$ and

$n = 10,000$, the true value of the asymptotic variances is close to $n$ times the MSE. This is not necessarily true for $n = 100$ (for instance, for Design 1, $\Sigma^*_{1,1} = 0.266$ whereas the MSE of $\tilde{a}_n(1)$ multiplied by $n$ is 1.390), which reflects the fact that, in that case, our $n$ is not large enough for the asymptotic formulae in Theorems 4.2–4.3 to be entirely reliable. This means that the asymptotic theory should be used with caution for moderate sample sizes and can be successfully applied to finite samples of reasonable sizes. Design 2 has been notably conducted with $m_0(1) = m_0(2) = 0.0$, and similar experiments, not reported here, have also been conducted with other values for $\underline{\theta}_0$, especially when $m_0(1)$ and $m_0(2)$ are extremely large, and/or very close to zero. It can be mentioned that this practical issue does not affect the behavior of our estimates. To conclude, these simple Monte Carlo simulation experiments clearly emphasize that, even though PARMA models with periodic intercepts have a larger number of parameters to be estimated compared to mean-corrected PARMA models, the LS approach provides better estimations than the FS method. This observation is particularly of interest to time series practitioners.

REMARK 5.1. Despite the main result of the paper (stated in Theorem 4.5), it is worth noting that the FS estimates are reasonably good estimates in so far as they are consistent and asymptotically normally distributed (see Theorems 4.1–4.3). For example, by the numerical study, the bias of the estimator $\tilde{m}_n(1)$ is about 0.7% or 0.5% of the true value when $n = 500$. Even if the LS estimators are more accurate in terms of asymptotic variance, they could not be written in an explicit form and need numerical optimization routines, and so there is a computation cost when considering LS estimates instead of FS estimates. However, it is always of interest for the time series analyst to obtain the more accurate estimation method as possible, especially when considering statistical inference on the model under study (e.g., for the construction of asymptotic confidence intervals and hypothesis tests).

Table 1. Characteristics of the empirical distribution of FS and LS estimates of model (2.1) with $\underline{\sigma}_0 = (1.0, 1.0)'$, for the sample size $n = 100$. The number of replications is 5,000.

| param. | meth. | bias | st. dev. | RMSE | tr. val. (AV) | $n\times$MSE | mean (AV) |
|---|---|---|---|---|---|---|---|
| Design 1: $m_0(1) = 1.0$, $m_0(2) = 0.0$, $a_0(1) = 0.9$, $a_0(2) = -0.9$ | | | | | | | |
| $m(1)$ | FS | 0.0101 | 0.9226 | 0.9226 | 100.439 | 85.119 | 85.117 |
| | LS | 0.0041 | 0.2052 | 0.2052 | 3.7838 | 4.211 | 4.211 |
| $m(2)$ | FS | -0.0002 | 0.1307 | 0.1307 | 1.6545 | 1.708 | 1.708 |
| | LS | -0.0012 | 0.1173 | 0.1173 | 1.2915 | 1.376 | 1.375 |
| $a(1)$ | FS | 0.0794 | 0.0872 | 0.1179 | 0.266 | 1.390 | 0.761 |
| | LS | 0.0215 | 0.0636 | 0.0671 | 0.266 | 0.450 | 0.404 |
| $a(2)$ | FS | -0.0162 | 0.0959 | 0.0972 | 0.665 | 0.945 | 0.919 |
| | LS | -0.0049 | 0.0943 | 0.0944 | 0.665 | 0.891 | 0.888 |
| Design 2: $m_0(1) = 0.0$, $m_0(2) = 0.0$, $a_0(1) = 0.5$, $a_0(2) = -0.9$ | | | | | | | |
| $m(1)$ | FS | 0.0024 | 0.2322 | 0.2322 | 5.692 | 5.392 | 5.390 |
| | LS | -0.0003 | 0.1640 | 0.1640 | 2.521 | 2.690 | 2.688 |
| $m(2)$ | FS | 0.0003 | 0.1360 | 0.1360 | 1.778 | 1.850 | 1.851 |
| | LS | 0.0005 | 0.1245 | 0.1245 | 1.457 | 1.550 | 1.550 |
| $a(1)$ | FS | 0.0329 | 0.0966 | 0.1021 | 0.827 | 1.042 | 0.934 |
| | LS | 0.0188 | 0.0981 | 0.0999 | 0.827 | 0.998 | 0.962 |
| $a(2)$ | FS | -0.0176 | 0.1516 | 0.1526 | 2.049 | 2.329 | 2.298 |
| | LS | 0.0000 | 0.1543 | 0.1543 | 2.049 | 2.381 | 2.381 |

Table 2. As Table 1, but for $n = 500$.

| param. | meth. | bias | st. dev. | RMSE | tr. val. (AV) | $n\times$MSE | mean (AV) |
|---|---|---|---|---|---|---|---|
| Design 1: $m_0(1) = 1.0$, $m_0(2) = 0.0$, $a_0(1) = 0.9$, $a_0(2) = -0.9$ | | | | | | | |
| $m(1)$ | FS | 0.0074 | 0.4495 | 0.4495 | 100.439 | 101.025 | 101.003 |
| | LS | 0.0012 | 0.0865 | 0.0865 | 3.7838 | 3.741 | 3.738 |
| $m(2)$ | FS | 0.0008 | 0.0571 | 0.0571 | 1.6545 | 1.630 | 1.633 |
| | LS | -0.0002 | 0.0508 | 0.0508 | 1.2915 | 1.290 | 1.288 |
| $a(1)$ | FS | 0.0149 | 0.0274 | 0.0311 | 0.266 | 1.484 | 0.373 |
| | LS | 0.0036 | 0.0242 | 0.0244 | 0.266 | 0.298 | 0.292 |
| $a(2)$ | FS | -0.0039 | 0.0369 | 0.0371 | 0.665 | 0.688 | 0.682 |
| | LS | -0.0012 | 0.0368 | 0.0368 | 0.665 | 0.677 | 0.677 |
| Design 2: $m_0(1) = 0.0$, $m_0(2) = 0.0$, $a_0(1) = 0.5$, $a_0(2) = -0.9$ | | | | | | | |
| $m(1)$ | FS | -0.0005 | 0.1057 | 0.1057 | 5.692 | 5.586 | 5.584 |
| | LS | -0.0005 | 0.0722 | 0.0722 | 2.521 | 2.606 | 2.609 |
| $m(2)$ | FS | 0.0001 | 0.0600 | 0.0600 | 1.778 | 1.800 | 1.803 |
| | LS | 0.0003 | 0.0540 | 0.0540 | 1.457 | 1.458 | 1.460 |
| $a(1)$ | FS | 0.0063 | 0.0416 | 0.0421 | 0.827 | 0.886 | 0.868 |
| | LS | 0.0003 | 0.0417 | 0.0419 | 0.827 | 0.878 | 0.870 |
| $a(2)$ | FS | -0.0039 | 0.0601 | 0.0644 | 2.049 | 2.074 | 2.066 |
| | LS | -0.0005 | 0.0643 | 0.0643 | 2.049 | 2.074 | 2.068 |

Table 3. As Table 1, but for $n = 10,000$.

| param. | meth. | bias | st. dev. | RMSE | tr. val. (AV) | $n\times$MSE | mean (AV) |
|---|---|---|---|---|---|---|---|
| Design 1: $m_0(1) = 1.0$, $m_0(2) = 0.0$, $a_0(1) = 0.9$, $a_0(2) = -0.9$ | | | | | | | |
| $m(1)$ | FS | 0.0021 | 0.0997 | 0.0997 | 100.439 | 99.401 | 99.472 |
| | LS | 0.0004 | 0.0195 | 0.0195 | 3.7838 | 3.803 | 3.818 |
| $m(2)$ | FS | 0.0002 | 0.0126 | 0.0126 | 1.6545 | 1.588 | 1.592 |
| | LS | 0.0002 | 0.0112 | 0.0112 | 1.2915 | 1.254 | 1.259 |
| $a(1)$ | FS | 0.0007 | 0.0052 | 0.0053 | 0.266 | 0.281 | 0.272 |
| | LS | 0.0002 | 0.0052 | 0.0052 | 0.266 | 0.270 | 0.268 |
| $a(2)$ | FS | -0.0002 | 0.0083 | 0.0083 | 0.665 | 0.689 | 0.685 |
| | LS | -0.0001 | 0.0083 | 0.0083 | 0.665 | 0.689 | 0.683 |
| Design 2: $m_0(1) = 0.0$, $m_0(2) = 0.0$, $a_0(1) = 0.5$, $a_0(2) = -0.9$ | | | | | | | |
| $m(1)$ | FS | -0.0001 | 0.0242 | 0.0242 | 5.692 | 5.856 | 5.845 |
| | LS | -0.0003 | 0.0160 | 0.0160 | 2.521 | 2.560 | 2.560 |
| $m(2)$ | FS | 0.0000 | 0.0134 | 0.0134 | 1.778 | 1.796 | 1.788 |
| | LS | -0.0001 | 0.0121 | 0.0121 | 1.457 | 1.464 | 1.468 |
| $a(1)$ | FS | 0.0002 | 0.0090 | 0.0090 | 0.827 | 0.810 | 0.805 |
| | LS | 0.0000 | 0.0090 | 0.0090 | 0.827 | 0.810 | 0.805 |
| $a(2)$ | FS | -0.0004 | 0.0146 | 0.0146 | 2.049 | 2.132 | 2.129 |
| | LS | -0.0002 | 0.0146 | 0.0146 | 2.049 | 2.132 | 2.129 |

## 6. CONCLUSION

This paper gives theoretical and numerical results aimed to justify that the common practice in time series analysis of fitting a zero-mean model to mean-corrected data sets can be joined to a significant lack of asymptotic accuracy for the parameter LS estimates of PARMA models. Including the process expectation in the model modifies the asymptotic results: the LS estimator remains consistent and asymptotically normal, but for an homoskedastic case, the asymptotic variance can be more accurate than the one in the mean-corrected approach.

We provided explicit analytical expressions for the asymptotic variance of the estimators and conducted Monte Carlo experiments to examine the finite sample inference. The numerical study proposed in this paper showed that the mean-correction could be tragic on the accuracy of the position parameters estimates. As claimed before, extending these results to QGLS- or ML-type estimators (to support for heteroskedasticity) represents an appealing continuation of this work, surely leading to very interesting results and challenging methodological issues, that merits further attention.

REFERENCES

Adams, G. J., Goodwin, G. C. (1995). Parameter estimation for periodic ARMA models. *J. Time Ser. Anal.* 16:127–145.

Alpay, D., Chevreuil, A., Loubaton, Ph. (2001). An extension problem for discrete-time periodically correlated stochastic processes. *J. Time Ser. Anal.* 22:1–11.

Anderson, P. L., Vecchia, A. V. (1993). Asymptotic results for periodic autoregressive moving-average processes. *J. Time Ser. Anal.* 14:1–18.

Azrak, R., Mélard, G. (2004). Asymptotic properties of quasi-likelihood estimators for ARMA models with time-dependent coefficients. To appear in *Stat. Inference Stoch. Process.*

Basawa, I. V., Lund, R. B. (2001). Large sample properties of parameter estimates for periodic ARMA models. *J. Time Ser. Anal.* 22:651–663.

Bibi, A., Francq, C. (2003) Consistent and asymptotically normal estimators for cyclically time-dependent linear models. *Ann. Inst. Statist. Math.* 55:41–68.

Billingsley, P. (1995). *Probability and Measure.* New York: John Wiley and Sons.

Bloomfield, P., Hurd, H. L., Lund, R. B. (1994). Periodic correlation in stratospheric ozone data. *J. Time Ser. Anal.* 15:127–150.

Box, G. E. P., Jenkins, G. M. (1976). *Time Series Analysis. Forecasting and Control.* New

Jersey: Prentice Hall, Englewood Cliffs.

Brockwell, P. J., Davis, R. A. (1991). *Time Series: Theory and Methods.* New York: Springer-Verlag.

Francq, C., Gautier, A. (2004a). Large sample properties of parameter least squares estimates for time-varying ARMA models. *J. Time Ser. Anal.* 25:765–783.

Francq, C., Gautier, A. (2004b). Estimation of time-varying ARMA models with Markovian changes in regime, *Statist. Probab. Lett.* 70:243–251.

Gardner, W., Franks, L. E. (1975). Characterization of cyclostationary random signal processes. *IEEE Trans. Inform. Theory* 21:4–14.

Gladyšhev, E. G. (1961). Periodically correlated random sequences. *Soviet Math.* 2:385–388.

Godambe, V. P., Heyde, C. C. (1987). Quasi-likelihood and optimal estimation. *International Statistical Review* 55:231–244.

Hurd, H. L., Makagon, A., Miamee, A.G. (2002). On AR(1) models with periodic and almost periodic coefficients. *Stochastic Process. Appl.* 100:167–185.

Lund, R. B., Basawa, I. V. (1999). Modeling and inference for periodically correlated time series. In *Asymptotics, Nonparametrics, and Time Series*, Ed. S. Ghosh, pp. 37–62, New York: Marcel Dekker.

Lund, R. B., Basawa, I. V. (2000). Recursive prediction and likelihood evaluation for periodic ARMA models. *J. Time Ser. Anal.* 21:75–93.

McLeod, A. I. (1994). Diagnostic checking of periodic autoregression models with application. *J. Time Ser. Anal.* 15:221–233.

Pagano, M. (1978). On periodic and multiple autoregressions. *Ann. Statist.* 6:1310–1317.

Parzen, E., Pagano M. (1979). An approach to modeling seasonally stationary time series.

*J. Econometrics* 9:137–153.

Press, W. H., Flannery, B. P., Teukolsky, S. A., Vetterling, W. T. (1986). *Numerical Recipes.* Cambridge: Cambridge University Press.

Shao, Q., Lund, R. B. (2004). Computation and characterization of autocorrelations and partial autocorrelations in periodic ARMA models. *J. Time Ser. Anal.* 25:359–372.

Tiao, G. C., Grupe, M. R. (1980). Hidden periodic autoregressive-moving average models in time series data. *Biometrika* 67:365–373.

Ula, T. A., Smadi, A. A. (2003). Identification on periodic moving-average models. *Comm. Statist. Theory Methods* 32:2465–2475.

Vecchia, A. V. (1985). Periodic autoregressive-moving average (PARMA) modeling with applications to water resources. *Water Resour. Bull.* 21:721–730.

*ACKNOWLEDGMENTS. The author is grateful for the constructive comments and suggestions for improvements made by the referee. The author also expresses his most sincere thanks and acknowledgments to Christian Francq and Jean-Michel Zakoïan for their helpful discussions.