PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 29, Fasc. 1 (2009), pp. 169–180

KALMAN-TYPE RECURSIONS FOR TIME-VARYING ARMA MODELS AND THEIR IMPLICATION FOR LEAST SQUARES PROCEDURE*

BY

ANTONY GAUTIER (LILLE)

Abstract. This paper is devoted to ARMA models with time-dependent coefficients, including well-known periodic ARMA models. We provide state-space representations and Kalman-type recursions to derive a Wold–Cramér decomposition for the least squares residuals. This decomposition turns out to be very convenient for further developments related to parameter least squares estimation. Some examples are proposed to illustrate the main purpose of these state-space forms.

2000 AMS Mathematics Subject Classification: Primary: 62M10; Secondary: 62M05.

Key words and phrases: Kalman-type recursions; least squares procedure; state-space representations; time-varying ARMA models; Wold–Cramér decomposition.

1. INTRODUCTION

In time series analysis, mounting evidences have been proposed to emphasize that empirical models in many fields such as economics, climatology or engineering are characterized by non-stationarity and parameter instability. Various methods have been devised in order to take non-stationarity into account: they mainly consist of differencing or filtering the original series with the purpose of turning it into a new set of data exhibiting no apparent deviation from stationarity. However, such methods allow for very restricted types of non-stationarities only. They essentially result in seasonal autoregressive integrated moving average (SARIMA) models, which consist in fitting an autoregressive moving average (ARMA) model to the differenced series. This implies that the same ARMA model applies whatever the ‘season’ under consideration. Nevertheless, when considering, for instance, economic daily variables depending on weekdays and on weekends (motorway traffic being an example), it seems natural to hope for different dynamics (and so different forecasting formulae).

* The author would like to thank Christian Francq for inspiration and the referee for a careful reading of the manuscript.

This observation has sparked off an explosion of interest in time series models with time-varying parameters. One notable class of such models is the time-varying ARMA models class, in which parameters can move discretely, for example, between a given number of regimes. In particular, this wide class includes the well-documented periodic ARMA time series models, extensively studied in the statistical literature in the last decade (see e.g. Adams and Goodwin [1], Anderson et al. [2], Basawa and Lund [4], Miamee and Talebi [15], Gautier [12]) and whose parameters switch according to a fixed periodic calendar. More generally, Dahlhaus [8], Azrak and Mélard [3] and Bibi and Francq [5] have dealt with the problem of estimating the whole class of ARMA models with time-dependent coefficients, via quasi-likelihood or least squares techniques.

The main purpose of this paper is therefore to provide easy mathematical methods to investigate further the large sample properties of least squares estimators of general time-varying ARMA models. Since the least squares procedure consists of the minimization of weighted sums of squares of prediction errors (see Godambe and Heyde [13] for a general reference), these methods are based on a state-space representation of the model under study and on a Wold–Cramér decomposition for the residuals. One distinctive feature of this paper is to explore the state-space framework as a statistical tool to address the issue of estimating ARMA models in the presence of time-dependent coefficients. Note however that it is far beyond the scope of this paper to study the asymptotic behaviour of the least squares estimates (to consider this question, see Francq and Gautier [10]).

The remainder of the paper is organized as follows. Section 2 describes the time-varying ARMA model under consideration. Section 3 provides a state-space representation and Kalman-type recursions for the model to obtain a pure time-varying MA representation for the residuals in closed form. In Section 4, the coefficients of the Wold–Cramér decomposition are derived in simple examples. Concluding remarks are gathered in Section 5.

The following notation will be used throughout the paper. The square identity matrix of order m is denoted by $\mathbf{I}_{(m)}$. We denote by $\mathbf{0}_{(m)}$ the 0-vector of size m. Let $\mathbb{N} = \{0, 1, 2, \dots\}$. For seek of simplicity, we indicate the size of any multivariate element into brackets under its notation, e.g. $M_{(m\times n)}$ denotes the matrix M with m rows and n columns. By M' we mean the transpose of any matrix M. We finally denote by $\underline{u}_i$ the i-th vector of the canonical basis of $\mathbb{R}^{p+q}$.

2. MODEL AND PRELIMINARIES

Consider a time series $(X_t)_{t=1,2,\dots}$ exhibiting changes in regime at known dates. We suppose that there exists a finite number d of regimes which alternate themselves either in a constant periodicity or at irregular time intervals. Denote by s_t the regime corresponding to the index t, so that $s_t = k$ when the time series is in the k-th regime at time t, for $k \in \{1, \dots, d\}$. The sequence (s_t) is assumed to be known as a real-valued sequence of observed changes in regime. The dynamics of X_t in each regime can be described by an ARMA(p, q) equation,

$$X_t - m(s_t) + \sum_{i=1}^{p} a_i(s_t)\{X_{t-i} - m(s_{t-i})\} = \epsilon_t + \sum_{i=1}^{q} b_i(s_t)\epsilon_{t-i} \tag{2.1}$$

for all $t \geqslant 1$, where p, q are minimal orders so that (2.1) is identifiable. Note that, under the additional assumptions made below, $m(s_t)$ can be interpreted as the expectation of X_t. The process (ϵ_t) is a white noise, namely a sequence of uncorrelated random variables with mean 0 and variance 1. The parameter of interest in Model (2.1) is denoted by

$$\theta = \{m(1), \ldots, m(d), a_1(1), \ldots, a_p(d), b_1(1), \ldots, b_q(d)\}'$$

and belongs to an open subset Θ of $\mathbb{R}^{(p+q+1)d}$.

The recursive relation (2.1) requires starting values that we specify as

$$X_t - m(s_t) = \epsilon_t = 0 \quad \text{for } t = 1 - \max\{p, q\}, \ldots, 0.$$

The best one-step predictor in the mean-square sense, that we denote by $\hat{X}_t(\theta) = E_\theta(X_t | X_{t-1}, \ldots, X_1)$, can be computed recursively by

$$\begin{cases} \hat{X}_t(\theta) = m(s_t) \quad \text{for } t = 1 - \max(p, q), \ldots, 0, \\ \hat{X}_k(\theta) = m(s_k) - \sum_{i=1}^{p} a_i(s_k)\{X_{k-i} - m(s_{k-i})\} \\ \qquad + \sum_{i=1}^{q} b_i(s_k)\{X_{k-i} - \hat{X}_{k-i}(\theta)\} \quad \text{for } k = 1, \ldots, t. \end{cases} \tag{2.2}$$

The true value of the parameter θ is denoted by θ_0. So we write

$$\theta_0 = \{m_0(1), \ldots, m_0(d), a_{01}(1), \ldots, a_{0p}(d), b_{01}(1), \ldots, b_{0q}(d)\}'.$$

The prediction errors are finally defined by

$$e_t(\theta) = X_t - \hat{X}_t(\theta). \tag{2.3}$$

Conditions ensuring consistency and asymptotic normality of least squares and quasi-generalized least squares estimators of Model (2.1) subject to stationary and ergodic changes in regime have been given in Francq and Gautier [10]. In the afore-mentioned paper, it has been shown that the residuals (2.3) can be expressed in the following Wold–Cramér decomposition for non-stationary processes (see Cramér [7]):

$$e_t(\theta) = \sum_{i=0}^{t-1} \psi_{t,i}(\theta, \theta_0)\epsilon_{t-i} + c_t(\theta, \theta_0). \tag{2.4}$$

For example, in the time-varying AR(1) case, we have (see Subsection 4.1)

$$\psi_{t,i}(\theta,\theta_0) = (-1)^i \{a_{01}(s_t) - a_1(s_t)\} \prod_{j=0}^{i-2} a_{01}(s_{t-1-j}),$$
$$c_t(\theta,\theta_0) = \{m_0(s_t) - m(s_t)\} + a_1(s_t)\{m_0(s_{t-1}) - m(s_{t-1})\}.$$

Recall that, given the observations $X_1, \dots, X_n$, a least squares estimator of the parameter θ_0 in Model (2.1) is obtained by solving

$$\hat{\theta}_n = \arg\min_{\theta\in\Theta^*} n^{-1} \sum_{t=1}^{n} e_t^2(\theta),$$

where Θ^* is a compact subset of Θ which contains θ_0 (see Cochrane and Orcutt [6]). Accordingly, relation (2.4) turns out to be very convenient to study the large sample properties of the least squares estimators of Model (2.1).

3. MAIN RESULTS

State-space representations and the associated Kalman recursions have had a profound impact on time series analysis and many related areas. The techniques were originally developed in connection with the control of linear systems (see Hannan and Deistler [14]). Many time series models can be formulated as special cases of the general state-space models. The Kalman recursions, which play a key role in the estimation of state-space models, provide a unified approach to prediction and estimation for all processes that can be given in a state-space form. In this section, a state-space representation of Model (2.1) and Kalman-type recursions are therefore derived. Next we apply them to obtain the time-varying MA representation (2.4) in closed form.

3.1. State-space representation. Let us introduce the following vectors:

$$\underset{(p+q)\times 1}{\underline{X}_t} = \underline{X}_t(\theta) = \begin{pmatrix} X_t - m(s_t) \\ X_{t-1} - m(s_{t-1}) \\ \vdots \\ X_{t-p+1} - m(s_{t-p+1}) \\ \epsilon_t \\ \epsilon_{t-1} \\ \vdots \\ \epsilon_{t-q+1} \end{pmatrix}, \qquad \underset{(p+q)\times 1}{\underline{\epsilon}_t} = \begin{pmatrix} \epsilon_t \\ 0 \\ \vdots \\ 0 \\ \epsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

and the following square matrix:

$$\underset{(p+q)\times(p+q)}{\Phi_t(\theta)}$$

$$= \begin{pmatrix} -a_1(s_t) & -a_2(s_t) & \cdots & -a_p(s_t) & b_1(s_t) & b_2(s_t) & \cdots & b_q(s_t) \\ 1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots & & & \vdots \\ \vdots & \ddots & 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

whose $(p+1,p)$-entry equals zero. We have implicitly assumed that $p \geqslant 1$ and $q \geqslant 1$, without loss of generality because the $a_i(\cdot)$'s and $b_i(\cdot)$'s can be equal to zero. The following result is then straightforward.

PROPOSITION 3.1. *Consider Model* (2.1) *and the above-mentioned notation* $\underline{X}_t$, $\underline{\epsilon}_t$ *and* $\Phi_t(\theta)$. *We have the following state-space form*:

$$\underline{X}_t = \underline{\epsilon}_t + \Phi_t(\theta)\underline{X}_{t-1}, \quad t = 1, 2, \ldots, \tag{3.1}$$

where $\underline{X}_0 = \mathbf{0}_{(p+q)}$.

3.2. Wold–Cramér decomposition for the residuals. We now concentrate on the way to obtain the Wold–Cramér representation (2.4) for the prediction errors of least squares estimation, via Kalman-type recursions. Consider the following square matrices:

$$\underset{(p+q)\times(p+q)}{J} = \begin{pmatrix} 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & & & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0 \\ -1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

where the -1 value is on the $(p+1)$-st line, and

$$\underset{(p+q)\times(p+q)}{K} = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & & & & & & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & & & & & & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix},$$

where both 1's are respectively on the first and $(p+1)$-st lines. The main result is now presented.

PROPOSITION 3.2. *Consider Model* (2.1) *and let the least squares residuals be defined by* (2.3). *Equation* (2.4) *holds with* $\psi_{t,0}(\theta,\theta_0)=1$,

$$\psi_{t,i}(\theta,\theta_0) = \underline{u}'_{p+1}\Big[\sum_{k=0}^{i}\Big\{\prod_{j=0}^{k-1} J\Phi_{t-j}(\theta)\Big\}K\Big\{\prod_{j=0}^{i-k-1}\Phi_{t-k-j}(\theta_0)\Big\}\Big]K\underline{u}_1 \tag{3.2}$$

for $i=1,\dots,t-1$, *and*

$$c_t(\theta,\theta_0) = \sum_{i=0}^{t-1}\underline{u}'_{p+1}\Big\{\prod_{j=0}^{i-1} J\Phi_{t-j}(\theta)\Big\}K\underline{u}_1\,\{m_0(s_{t-i})-m(s_{t-i})\}. \tag{3.3}$$

Proof. First, one can write (2.2) as

$$\begin{aligned} \underline{\hat{X}}_{t|t-1} &= \Phi_t(\theta)\underline{\hat{X}}_{t-1|t-1}, \\ \underline{\hat{X}}_{t|t} &= J\underline{\hat{X}}_{t|t-1}+\underline{x}_t, \quad t=1,2,\dots, \end{aligned} \tag{3.4}$$

where $\underline{\hat{X}}_{0|0}=\mathbf{0}_{(p+q)}$,

$$\underset{(p+q)\times 1}{\underline{\hat{X}}_{t|t-1}} = \begin{pmatrix} \hat{X}_t - m(s_t) \\ X_{t-1}-m(s_{t-1}) \\ \vdots \\ X_{t-p+1}-m(s_{t-p+1}) \\ 0 \\ e_{t-1}(\theta) \\ \vdots \\ e_{t-q+1}(\theta) \end{pmatrix}, \quad \underset{(p+q)\times 1}{\underline{\hat{X}}_{t|t}} = \begin{pmatrix} X_t - m(s_t) \\ X_{t-1}-m(s_{t-1}) \\ \vdots \\ X_{t-p+1}-m(s_{t-p+1}) \\ e_t(\theta) \\ e_{t-1}(\theta) \\ \vdots \\ e_{t-q+1}(\theta) \end{pmatrix},$$

and

$$\underset{(p+q)\times 1}{\underline{x}_t} = \begin{pmatrix} X_t - m(s_t) \\ 0 \\ \vdots \\ 0 \\ X_t - m(s_t) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

From (3.4), we obtain

$$\hat{\underline{X}}_{t|t} = \underline{x}_t + J\Phi_t(\theta)\hat{\underline{X}}_{t-1|t-1} = \underline{x}_t + \sum_{k=1}^{t-1} \Big\{ \prod_{j=0}^{k-1} J\Phi_{t-j}(\theta) \Big\} \underline{x}_{t-k}. \tag{3.5}$$

Since

$$\underline{X}_t(\theta_0) = \underline{\epsilon}_t + \sum_{k=1}^{t-1} \Big\{ \prod_{j=0}^{k-1} \Phi_{t-j}(\theta_0) \Big\} \underline{\epsilon}_{t-k}$$

and

$$\underline{x}_t = K\underline{X}_t = K\underline{X}_t(\theta_0) + K\left\{\underline{X}_t - \underline{X}_t(\theta_0)\right\},$$

we obtain from (3.5)

$$\begin{aligned} \hat{\underline{X}}_{t|t} = &\sum_{i=0}^{t-1} \sum_{k=0}^{i} \Big\{ \prod_{j=0}^{k-1} J\Phi_{t-j}(\theta) \Big\} K \Big\{ \prod_{j=0}^{i-k-1} \Phi_{t-k-j}(\theta_0) \Big\} \underline{\epsilon}_{t-i} \\ &+ \sum_{i=0}^{t-1} \Big\{ \prod_{j=0}^{i-1} J\Phi_{t-j}(\theta) \Big\} K \left\{ \underline{X}_{t-i}(\theta) - \underline{X}_{t-i}(\theta_0) \right\}, \end{aligned} \tag{3.6}$$

where $\prod_{j=0}^{k-1} J\Phi_{t-j}(\theta) = \mathbf{I}_{(p+q)}$ when $k = 0$ and $\prod_{j=0}^{i-k-1} \Phi_{t-k-j}(\theta_0) = \mathbf{I}_{(p+q)}$ when $k = i$. Note that

$$K\left\{\underline{X}_t(\theta) - \underline{X}_t(\theta_0)\right\} = \begin{pmatrix} m_0(s_t) - m(s_t) \\ 0 \\ \vdots \\ 0 \\ m_0(s_t) - m(s_t) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Finally, the proof is completed when taking the $(p+1)$-st component of $\hat{\underline{X}}_{t|t}$ given by (3.6). ■

REMARK 3.1. It is worth mentioning that equations (3.2)–(3.3) show that the coefficient $\psi_{t,i}(\theta, \theta_0)$ only depends on a finite number of past values of the regimes, more precisely on $s_t, \ldots, s_{t-i+1}$, whereas the intercept $c_t(\theta, \theta_0)$ generally depends on all the past values of the regimes. Note also that $e_t(\theta_0) = \epsilon_t$, $c_t(\theta_0, \theta_0) = 0$ and $\psi_{t,i}(\theta_0, \theta_0) = 0$ for all $i \geqslant 1$.

4. EXAMPLES

In this section, we show that the expressions of the $\psi_{t,i}(\theta, \theta_0)$'s and $c_t(\theta, \theta_0)$'s given by (3.2) and (3.3), respectively, can be simplified in particular cases.

4.1. AR(1) model. First consider the simple example of Model (2.1) with $(p, q) = (1, 0)$. The vectorial representations are however used for $p = q = 1$. We have (3.1) with

$$\Phi_t(\theta) = \begin{pmatrix} -a_1(s_t) & 0 \\ 0 & 0 \end{pmatrix}, \quad \underline{X}_t = \begin{pmatrix} X_t - m(s_t) \\ \epsilon_t \end{pmatrix}, \quad \underline{\epsilon}_t = \begin{pmatrix} \epsilon_t \\ \epsilon_t \end{pmatrix},$$

and (3.4)–(3.6) with

$$J = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \quad \underline{x}_t = \begin{pmatrix} X_t - m(s_t) \\ X_t - m(s_t) \end{pmatrix}, \quad K = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

Note that

$$\prod_{j=0}^{k-1} J\Phi_{t-j}(\theta) = 0 \quad \text{for } k \geqslant 2.$$

Therefore, we obtain (3.2) with

$$\psi_{t,1}(\theta, \theta_0) = \underline{u}_2' \left\{ K\Phi_t(\theta_0) + J\Phi_t(\theta)K \right\} K\underline{u}_1 = -a_{01}(s_t) + a_1(s_t),$$

$$\begin{aligned} \psi_{t,i}(\theta, \theta_0) &= \underline{u}_2' \left[\sum_{k=0}^{1} \Big\{ \prod_{j=0}^{k-1} J\Phi_{t-j}(\theta) \Big\} K \begin{pmatrix} \prod_{j=0}^{i-k-1} -a_{01}(s_{t-k-j}) & 0 \\ 0 & 0 \end{pmatrix} \right] K\underline{u}_1 \\ &= (-1)^i \prod_{j=0}^{i-1} a_{01}(s_{t-j}) + (-1)^{i-1} a_1(s_t) \prod_{j=0}^{i-2} a_{01}(s_{t-1-j}) \\ &= (-1)^i \left\{ a_{01}(s_t) - a_1(s_t) \right\} \prod_{j=0}^{i-2} a_{01}(s_{t-1-j}) \end{aligned}$$

for $i \geqslant 1$ with usual conventions. Moreover, we have the intercept

$$\begin{aligned} c_t(\theta, \theta_0) &= \underline{u}_2' K\underline{u}_1 \left\{ m_0(s_t) - m(s_t) \right\} + \underline{u}_2' J\Phi_t(\theta) K\underline{u}_1 \left\{ m_0(s_{t-1}) - m(s_{t-1}) \right\} \\ &= \left\{ m_0(s_t) - m(s_t) \right\} + a_1(s_t) \left\{ m_0(s_{t-1}) - m(s_{t-1}) \right\}. \end{aligned}$$

4.2. MA(1) model. Now, we consider the MA(1) case, i.e. Model (2.1) with $(p,q)=(0,1)$. Using the notation of the AR(1) case, equations (3.1)–(3.6) hold with

$$\Phi_t(\theta)=\begin{pmatrix} 0 & b_1(s_t) \\ 0 & 0 \end{pmatrix}.$$

Note that $\prod_{j=0}^{i-k-1}\Phi_{t-k-j}(\theta_0)=0$ for $k\leqslant i-2$. Therefore we have (2.4) with

$$\begin{aligned}
\psi_{t,i}(\theta,\theta_0) &= \underline{u}_2'\big[\big\{\prod_{j=0}^{i-1}J\Phi_{t-j}(\theta)\big\}K+\big\{\prod_{j=0}^{i-2}J\Phi_{t-j}(\theta)\big\}K\Phi_{t-i+1}(\theta_0)\big]K\underline{u}_1 \\
&= \underline{u}_2'\left[(-1)^i\begin{pmatrix} 0 & 0 \\ 0 & \prod_{j=0}^{i-1}b_1(s_{t-j}) \end{pmatrix}K\right. \\
&\quad \left.+(-1)^{i-1}\begin{pmatrix} 0 & 0 \\ 0 & \prod_{j=0}^{i-2}b_1(s_{t-j}) \end{pmatrix}K\Phi_{t-i+1}(\theta_0)\right]K\underline{u}_1 \\
&= (-1)^i\prod_{j=0}^{i-1}b_1(s_{t-j})+(-1)^{i-1}\prod_{j=0}^{i-2}b_1(s_{t-j})b_{01}(s_{t-i+1}) \\
&= (-1)^i\prod_{j=0}^{i-2}b_1(s_{t-j})\left\{b_1(s_{t-i+1})-b_{01}(s_{t-i+1})\right\}
\end{aligned}$$

for $i\geqslant 1$. The intercept is equal to

$$\begin{aligned}
c_t(\theta,\theta_0) &= \sum_{i=0}^{t-1}(-1)^i\underline{u}_2'\begin{pmatrix} 0 & 0 \\ 0 & \prod_{j=0}^{i-1}b_1(s_{t-j}) \end{pmatrix}K\underline{u}_1\left\{m_0(s_{t-i})-m(s_{t-i})\right\} \\
&= \sum_{i=0}^{t-1}(-1)^i\big\{\prod_{j=0}^{i-1}b_1(s_{t-j})\big\}\left\{m_0(s_{t-i})-m(s_{t-i})\right\}.
\end{aligned}$$

4.3. ARMA(1,1) model. Now, consider an ARMA(1,1) model, i.e. Model (2.1) with $(p,q)=(1,1)$. Maintaining our notation, equations (3.1)–(3.6) hold with

$$\Phi_t(\theta)=\begin{pmatrix} -a_1(s_t) & b_1(s_t) \\ 0 & 0 \end{pmatrix}.$$

Note that

$$\begin{aligned}
&\prod_{j=0}^{k-1}J\Phi_{t-j}(\theta) \\
&\quad=\begin{pmatrix} 0 & 0 \\ (-1)^{k-1}\big\{\prod_{j=0}^{k-2}b_1(s_{t-j})\big\}a_1(s_{t-k+1}) & (-1)^k\prod_{j=0}^{k-1}b_1(s_{t-j}) \end{pmatrix}
\end{aligned}$$

for $k\geqslant 1$, and

$$\prod_{j=0}^{i-k-1} \Phi_{t-k-j}(\theta_0) =$$

$$\begin{pmatrix} (-1)^{i-k}\prod_{j=0}^{i-k-1} a_{01}(s_{t-k-j}) & (-1)^{i-k-1}\big\{ \prod_{j=0}^{i-k-2} a_{01}(s_{t-k-j})\big\} b_{01}(s_{t-i+1}) \\ 0 & 0 \end{pmatrix}$$

for $k \leqslant i-1$. Therefore, we obtain (2.4) with

$$\begin{aligned}
\psi_{t,i}(\theta,\theta_0) &= (-1)^{i-1}\Big\{ \prod_{j=0}^{i-2} a_{01}(s_{t-j})\Big\} \{b_{01}(s_{t-i+1}) - a_{01}(s_{t-i+1})\} \\
&\quad + \sum_{k=1}^{i-1} (-1)^{i}\Big\{ \prod_{j=0}^{k-2} b_1(s_{t-j})\Big\} \{a_1(s_{t-k+1}) - b_1(s_{t-k+1})\} \\
&\quad \times \Big\{ \prod_{j=0}^{i-k-2} a_{01}(s_{t-k-j})\Big\} \{b_{01}(s_{t-i+1}) - a_{01}(s_{t-i+1})\} \\
&\quad + (-1)^{i-1}\Big\{ \prod_{j=0}^{i-2} b_1(s_{t-j})\Big\} \{a_1(s_{t-i+1}) - b_1(s_{t-i+1})\} \\
&= \sum_{k=1}^{i-1} (-1)^{i}\Big\{ \prod_{j=0}^{k-2} b_1(s_{t-j})\Big\} \{a_1(s_{t-k+1}) - a_{01}(s_{t-k+1})\} \\
&\quad \times \Big\{ \prod_{j=0}^{i-k-2} a_{01}(s_{t-k-j})\Big\} \{b_{01}(s_{t-i+1}) - a_{01}(s_{t-i+1})\} \\
&\quad + (-1)^{i-1}\Big\{ \prod_{j=0}^{i-2} b_1(s_{t-j})\Big\} \{a_1(s_{t-i+1}) - b_1(s_{t-i+1}) \\
&\quad + b_{01}(s_{t-i+1}) - a_{01}(s_{t-i+1})\}
\end{aligned}$$

for $i \geqslant 1$. In this example, the intercept equals

$$\begin{aligned}
c_t(\theta,\theta_0) &= \underline{u}_2' K \underline{u}_1 \{m_0(s_t) - m(s_t)\} \\
&\quad + \sum_{i=1}^{t-1} \underline{u}_2'\Big\{ \prod_{j=0}^{i-1} J\Phi_{t-j}(\theta)\Big\} K \underline{u}_1 \{m_0(s_{t-i}) - m(s_{t-i})\} \\
&= m_0(s_t) - m(s_t) + \{a_1(s_t) - b_1(s_t)\}\{m_0(s_{t-1}) - m(s_{t-1})\} \\
&\quad + \sum_{i=2}^{t-1} (-1)^{i-1}\Big\{ \prod_{j=0}^{i-2} b_1(s_{t-j})\Big\} \{a_1(s_{t-i+1}) - b_1(s_{t-i+1})\} \\
&\quad \times \{m_0(s_{t-i}) - m(s_{t-i})\}.
\end{aligned}$$

4.4. AR(2) model. Finally, let us have a look at Model (2.1) with $(p,q) = (2,0)$. In this case, equations (3.1)–(3.6) hold with

$$\Phi_t(\theta) = \begin{pmatrix} -a_1(s_t) & -a_2(s_t) & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\underline{X}_t = \begin{pmatrix} X_t - m(s_t) \\ X_{t-1} - m(s_{t-1}) \\ \epsilon_t \end{pmatrix}, \quad \underline{\epsilon}_t = \begin{pmatrix} \epsilon_t \\ 0 \\ \epsilon_t \end{pmatrix},$$

and

$$J = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \underline{x}_t = \begin{pmatrix} X_t - m(s_t) \\ 0 \\ X_t - m(s_t) \end{pmatrix}, \quad K = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Now,

$$\prod_{j=0}^{k-1} J\Phi_{t-j}(\theta) = 0 \quad \text{when } k \geqslant 3.$$

So we have (2.4) with

$$\begin{aligned}\psi_{t,i}(\theta,\theta_0) &= \underline{u}_3' \Big[\sum_{k=0}^{2} \Big\{ \prod_{j=0}^{k-1} J\Phi_{t-j}(\theta) \Big\} K \Big\{ \prod_{j=0}^{i-k-1} \Phi_{t-k-j}(\theta_0) \Big\} \Big] K \underline{u}_1 \\ &= \underline{u}_3' \left\{ K\Phi_t(\theta_0)\Phi_{t-1}(\theta_0) + J\Phi_t(\theta)K\Phi_{t-1}(\theta_0) \right. \\ &\quad \left. + J\Phi_t(\theta)J\Phi_{t-1}(\theta)K \right\} \prod_{j=0}^{i-3} \Phi_{t-2-j}(\theta_0) K \underline{u}_1.\end{aligned}$$

Here, the intercept is

$$\begin{aligned}c_t(\theta,\theta_0) &= \underline{u}_3' K \underline{u}_1 \left\{ m_0(s_t) - m(s_t) \right\} + \underline{u}_3' J\Phi_t(\theta) K \underline{u}_1 \left\{ m_0(s_{t-1}) - m(s_{t-1}) \right\} \\ &\quad + \underline{u}_3' J\Phi_t(\theta) J\Phi_{t-1}(\theta) K \underline{u}_1 \left\{ m_0(s_{t-2}) - m(s_{t-2}) \right\} \\ &= m_0(s_t) - m(s_t) + a_1(s_t) \left\{ m_0(s_{t-1}) - m(s_{t-1}) \right\} \\ &\quad + a_2(s_t) \left\{ m_0(s_{t-2}) - m(s_{t-2}) \right\}.\end{aligned}$$

5. CONCLUDING REMARKS

This paper can be appreciated as a preliminary technical step before the least squares estimation of general time-varying ARMA models. We have investigated the state-space framework of such non-stationary time series models to obtain a convenient representation for the prediction errors. We have provided a simple methodology that allows to derive the coefficients of the Wold–Cramér decomposition for the residuals which play a crucial role in the least squares minimization procedure. Of course, one can investigate this question using a straightforward univariate approach but this way seems rather tedious. The method presented in the paper is conversely based on the multivariate area and represents an appealing alternative to univariate calculations.

REFERENCES

[1] G. J. Adams and G. C. Goodwin, *Parameter estimation for periodic ARMA models*, J. Time Ser. Anal. 16 (1995), pp. 127–146.

[2] P. L. Anderson, M. M. Meerschaert and A. V. Vecchia, *Innovations algorithm for periodically stationary time series*, Stochastic Process. Appl. 83 (1999), pp. 149–169.

[3] R. Azrak and G. Mélard, *Asymptotic properties of quasi-maximum likelihood estimators for ARMA models with time-dependent coefficients*, Stat. Inference Stoch. Process. 9 (2006), pp. 279–330.

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

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

[6] D. Cochrane and D. H. Orcutt, *Applications of least squares regression to relationships containing autocorrelated error terms*, J. Amer. Statist. Assoc. 44 (1949), pp. 32–61.

[7] H. Cramér, *On some classes of nonstationary stochastic processes*, Proc. 4-th Berkeley Symp. Math. Statist. Prob. 2 (1961), pp. 57–78.

[8] R. Dahlhaus, *Fitting time series models to nonstationary processes*, Ann. Statist. 25 (1997), pp. 1–37.

[9] P. de Jong and J. Penzer, *The ARMA model in state space form*, Statist. Probab. Lett. 70 (2004), pp. 119–125.

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

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

[12] A. Gautier, *Asymptotic inefficiency of mean-correction on parameter estimation for a periodic first-order autoregressive model*, Comm. Statist. Theory Methods 35 (2006), pp. 2083–2106.

[13] V. P. Godambe and C. C. Heyde, *Quasi-likelihood and optimal estimation*, Internat. Statist. Rev. 55 (1987), pp. 231–244.

[14] E. J. Hannan and M. Deistler, *The Statistical Theory of Linear Systems*, Wiley, New York 1988.

[15] A. G. Miamee and S. Talebi, *On PC solutions of PARMA*(p,q) *models*, Probab. Math. Statist. 25 (2005), pp. 279–288.

Université Lille 3
Laboratoire EQUIPPE
BP 60149
59653 Villeneuve d'Ascq Cedex, France
E-mail: antony.gautier@univ-lille3.fr

Received on 17.4.2008;
revised version on 9.10.2008