On the estimation of periodic ARMA models with uncorrelated but dependent errors

Francq, Christian
Université Lille 3 (EQUIPPE) and CREST
59653 Villeneuve d'Ascq Cedex, France
E-mail: christian.francq@univ-lille3.fr

Roy, Roch
Département de mathématiques et de statistique and CRM, Université de Montréal
CP 6128, succ. Centre-ville, Montréal, Québec, H3C 3J7, Canada
E-mail: royr@dms.umontreal.ca

Saidi, Abdessamad
Research Department, Bank Al-Maghrib, Rabat, Morocco
E-mail: ab.saidi@bkam.ma

1- Introduction

The main goal of this paper is to study the asymptotic properties of least squares (LS) estimation for invertible and causal PARMA models with uncorrelated but dependent errors (weak PARMA). Four different LS estimators are considered: ordinary least squares (OLS), weighted least squares (WLS) for an arbitrary vector of weights, generalized least squares (GLS) in which the weights correspond to the theoretical seasonal variances and quasi-generalized least squares (QLS) where the weights are the estimated seasonal variances. The strong consistency and the asymptotic normality are established for each of them. Obviously, their asymptotic covariance matrices depend on the vector of weights. Our work extends a result of Basawa and Lund (2001) for least squares estimation of PARMA models with independent errors (strong PARMA).

The paper is organized as follows. The definitions of strong and weak PARMA processes are given in Section 2 and situations in which weak PARMA representations naturally arise are presented. In Section 3, we first recall the usual multivariate ARMA representation of a PARMA model. The four LS estimators are described and their asymptotic properties are given in Theorems 1 and 2. The proofs can be found in Francq, Roy and Saidi (2009) (FRS in the sequel). It is seen that the GLS estimators are optimal in the class of WLS estimators when the noise sequence is in a particular class of martingale differences. Arguing as in Francq, Roy and Zakoan (2005), a consistent estimator of the asymptotic covariance matrix of LS estimators under the assumption of a weak noise is proposed. In Section 4, we present an example of weak PARMA models for which the asymptotic covariance matrix of the least squares estimators is given in a close form and is compared to the corresponding matrix under the assumption of a strong noise. The difference can be huge. In FRS, Monte Carlo results are presented. Two different PARMA models with strong and weak noises were used to investigate the size and power of a Wald test based on the proposed consistent estimator of the asymptotic covariance matrix, under the assumption of either a weak or strong noise. The rate of convergence of the estimated asymptotic standard errors is also analysed. Finally, our results were exploited to address the question of day-of-the-week seasonality of four European stock market indices. The space constraint does not allow us to report these numerical results here.

2- Weak and Strong PARMA models

A stochastic process $\{X_t\}$ is called periodically stationary if $\mu_t = \mathbb{E}[X_t]$ and $\gamma_t(h) = \mathbb{E}[X_t X_{t+h}]$,

$h \in \mathbb{Z}$, are both periodic functions in time t with the same period T and $\mathbb{E}\left[X_t^2\right] < +\infty$ for all t. For convenience, the non-periodic notation X_t will be used interchangeably with the periodic notation $X_{nT+\nu}$ which refers to X_t during the season $\nu \in \{1, ..., T\}$ in the cycle n. By definition, a periodic process $\{X_t\}$ follows a periodic (with period T) autoregressive moving average model with the following parameters at season $\nu \in \{1, ..., T\}$: the mean μ_ν, the autoregressive order and coefficients p_ν, $\phi_1(\nu), ..., \phi_{p_\nu}(\nu)$, and the moving average order and coefficients q_ν,$\theta_1(\nu), ..., \theta_{q_\nu}(\nu)$, denoted simply $\text{PARMA}_T(p_1, ..., p_\nu, ..., p_T; q_1, ..., q_\nu, ..., q_T)$, if there exists a periodic white noise sequence $\{\epsilon_t\} = \{\epsilon_{nT+\nu}\}$, *i.e.* $\mathbb{E}[\epsilon_t] = 0$ for all t, $\mathbb{E}[\epsilon_t \epsilon_{t'}] = 0$ for all $t \neq t'$, and $\mathbb{E}\left[\epsilon_{nT+\nu}^2\right] = \sigma_\nu^2 > 0$, such that

$$(1) \quad (X_{nT+\nu} - \mu_\nu) - \sum_{k=1}^{p_\nu} \phi_k(\nu)(X_{nT+\nu-k} - \mu_{\nu-k}) = \epsilon_{nT+\nu} - \sum_{l=1}^{q_\nu} \theta_l(\nu)\epsilon_{nT+\nu-l}.$$

If the errors ϵ_t are uncorrelated but not necessarily independent, both periodic white noise or weak periodic white noise are used to qualify the error process $\{\epsilon_t\}$ and similarly the terminology "PARMA" or "weak PARMA" is used for the model (1). When the error terms $\sigma_\nu^{-1}\epsilon_{nT+\nu}$ are independent and identically distributed (iid) rather than only uncorrelated, the model (1) is called strong PARMA model and $\{\epsilon_t\}$ is a strong periodic white noise. The process $\{\epsilon_t = \epsilon_{nT+\nu}\}$ can be interpreted as the linear innovation of $\{X_t = X_{nT+\nu}\}$, *i.e.* $\epsilon_t = X_t - \mathbb{E}[X_t|\mathcal{H}_X(t-1)]$ where $\mathcal{H}_X(t)$ is the Hilbert space spanned by $\{X_s, s \leq t\}$.

When the order of both the autoregressive and moving average components are not allowed to vary with season, *i.e.*, when $p_1 = ... = p_T = p$ and $q_1 = ... = q_T = q$ we simply write $\text{PARMA}_T(p; q)$ instead of $\text{PARMA}_T(p, ..., p; q, ..., q)$. The terminologies periodic autoregressive (PAR) model and periodic moving average (PMA) model are respectively used when the moving average orders are null, and when the autoregressive orders are null. If $T = 1$ the process (1) is the usual stationary autoregressive moving average model (ARMA).

Note that (1) is just a model, not a data generating process (DGP). Many DGP can be compatible with a weak PARMA representation. DGP admitting weak PARMA representations can be obtained as i) certain transformations of strong PARMA processes, ii) causal representations of non causal PARMA processes, iii) linear representations of non linear processes, or iv) approximations of the so-called Wold decomposition. Examples are given in FRS. In particular, Roy and Saidi (2008) showed that systematic sampling and temporal aggregation of a strong PARMA process leads in general to a weak PARMA process. They provide a sufficient condition under which the noise of the aggregated process is neither strong nor a martingale difference.

3- Least squares estimation of weak PARMA models

Here, we focus on the asymptotic properties of the least squares estimators of the autoregressive and moving average parameters of the PARMA_T process (1). There is no loss of generality in assuming that $p_1 = ... = p_T = p$ and $q_1 = ... = q_T = q$ by adding coefficients equal to zero (Basawa and Lund, 2000). Furthermore, we suppose that the process is centered, that is $\mu_1 = ... = \mu_T = 0$. We make this assumption to lighten the presentation, but the results stated in this section extend directly for models with constants.

The difference equations (1) can be written in the T-dimensional vector form (Vecchia, 1985)

$$(2) \quad \Phi_0 \mathbf{X}_n - \sum_{k=1}^{p^*} \Phi_k \mathbf{X}_{n-k} = \Theta_0 \boldsymbol{\epsilon}_n - \sum_{l=1}^{q^*} \Theta_l \boldsymbol{\epsilon}_{n-l},$$

where

$$(3) \quad \mathbf{X}_n = (X_{nT+1}, ..., X_{nT+T})', \quad \boldsymbol{\epsilon}_n = (\epsilon_{nT+1}, ..., \epsilon_{nT+T})',$$

$p^* = \lfloor (p-1)/T \rfloor + 1$, $q^* = \lfloor (q-1)/T \rfloor + 1$, the matrix coefficients Φ_k, $k = 0, ..., p^*$ and Θ_l, $l = 0, ..., q^*$, are given in Vecchia (1985). Il particular, Φ_0 and Θ_0 are lower triangular matrices with 1's along the main diagonal.

From (2), we can in principle deduce the properties of weak PARMA parameter estimation from existing results on parameter estimation of a vector ARMA model under general assumptions on the white noise process including dependence cases. Here, we have preferred to work in the univariate PARMA setting for various reasons. First, results obtained directly in terms of the univariate PARMA representation are more directly usable because fewer parameters are involved and their estimation is easier. For example, in a PARMA$_T(1,1)$ model, there are $2T$ AR and MA parameters to estimate. The corresponding VARMA(1,1) representation contains $4T^2$ parameters among which $2T(2T-1)$ are constrained to 1 or 0. The difference in the number of parameters to be estimated is huge, specially with monthly or weekly data ($T = 12, 52$). Second, the VARMA representation (2) is not standard, the matrices Φ_0 and Θ_0 are not in general the identity matrix.

The covariance matrix of the T-dimensional white noise $\boldsymbol{\epsilon}_n$ is $\Sigma_\epsilon = \mathrm{Diag}(\sigma_1^2, ..., \sigma_T^2) > 0$. Denote by B the lag operator such that $B^h \mathbf{X}_n = \mathbf{X}_{n-h}$. Equation (2) can be written as

$$(4) \quad \Phi(B)\mathbf{X}_n = \Theta(B)\boldsymbol{\epsilon}_n$$

where $\Phi(z) = \Phi_0 - \Phi_1 z - ... - \Phi_{p^*} z^{p^*}$ and $\Theta(z) = \Theta_0 - \Theta_1 z - ... - \Theta_{q^*} z^{q^*}$ are the matrix polynomials of the vectorial autoregressive moving average representation. It is important to note that the lag operator B operates on the cycle index n. When it acts on the time index $t = nT + \nu$ of the periodic process $\{Z_t\}$, it gives $B^k Z_{nT+\nu} = Z_{(n-k)T+\nu}$.

Hereafter, we impose the following conditions on the process $\{\mathbf{X}_n\}$.

(A1) The PARMA process $\{X_{nT+\nu}\}$ is causal and invertible, in the sense that, the roots of $\det\{\Phi(z)\}$ and of $\det\{\Theta(z)\}$ are greater than one in modulus. Furthermore, we assume that the VARMA model (2) is identifiable.

For notation, let $\vec{\phi}(\nu) = (\phi_1(\nu), ..., \phi_p(\nu))'$ and $\vec{\theta}(\nu) = (\theta_1(\nu), ..., \theta_q(\nu))'$ respectively denote the vectors of autoregressive and moving average parameters for season ν. The $T(p+q)$-dimensional collection of all PARMA parameters is denoted by

$$\vec{\alpha} := \left(\vec{\phi}(1)', ..., \vec{\phi}(T)', \vec{\theta}(1)', ..., \vec{\theta}(T)'\right)'.$$

The white noise variances $\vec{\sigma}^2 = \left(\sigma_1^2, ..., \sigma_T^2\right)'$ will be treated as nuisance parameters.

Let $X_1, ..., X_{NT}$ be a data sample from the causal and invertible PARMA model (1) with the true parameter value $\vec{\alpha} = \vec{\alpha}_0$ and $\vec{\sigma}^2 = \vec{\sigma}_0^2$. The sample contains N full periods of data which are indexed from 0 to $N-1$. Indeed, when $0 \le n \le N-1$ and $1 \le \nu \le T$, $nT + \nu$ goes from 1 to NT. It is understood that $\vec{\alpha}_0$ belongs to the parameter space

$$\Omega = \left\{ \vec{\alpha} = \left(\vec{\phi}(1)', ..., \vec{\phi}(T)', \vec{\theta}(1)', ..., \vec{\theta}(T)'\right)' \in \mathbb{R}^{T(p+q)} \text{ such that (A1) is verified} \right\}.$$

For $\vec{\alpha} \in \Omega$, let $\epsilon_{nT+\nu}(\vec{\alpha})$ be the periodically second-order stationary solution of

$$(5) \quad \epsilon_{nT+\nu}(\vec{\alpha}) = X_{nT+\nu} - \sum_{i=1}^{p} \phi_i(\nu) X_{nT+\nu-i} + \sum_{j=1}^{q} \theta_j(\nu) \epsilon_{nT+\nu-j}(\vec{\alpha}).$$

Note that, almost surely, $\epsilon_{nT+\nu}(\vec{\alpha}_0) = \epsilon_{nT+\nu}$ for all $n \in \mathbb{Z}$ and $\nu \in \{1, ..., T\}$. Moreover, $\epsilon_{nT+\nu}(\vec{\alpha})$ can be approximated by $e_{nT+\nu}(\vec{\alpha})$ which is also determined recursively in t via a truncated version of (5)

$$(6) \quad e_{nT+\nu}(\vec{\alpha}) = X_{nT+\nu} - \sum_{i=1}^{p} \phi_i(\nu) X_{nT+\nu-i} + \sum_{j=1}^{q} \theta_j(\nu) e_{nT+\nu-j}(\vec{\alpha}),$$

where the unknown starting values are set to zero: $e_0(\vec{\alpha}) = ... = e_{1-q}(\vec{\alpha}) = X_0 = ... = X_{1-p} = 0$.

Let $\delta > 0$ be a constant chosen such that $\vec{\alpha}_0$ belongs to the interior of the compact set

$$\Omega_\delta = \left\{ \vec{\alpha} \in \mathbb{R}^{T(p+q)} | \text{the zeros of } \det\{\Phi(z)\} \text{ and those of } \det\{\Theta(z)\} \text{ have modulus} \geq 1+\delta \right\}.$$

The random variable $\hat{\vec{\alpha}}_{\text{OLS}}$ is called the ordinary least squares (OLS) estimator of $\vec{\alpha}$ if it satisfies, almost surely,

$$(7) \quad S_N(\hat{\vec{\alpha}}_{\text{OLS}}) = \min_{\vec{\alpha} \in \Omega_\delta} S_N(\vec{\alpha}) \qquad \text{where} \qquad S_N(\vec{\alpha}) = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{\nu=1}^{T} e^2_{nT+\nu}(\vec{\alpha}).$$

Because of the presence of heteroscedastic innovations, the OLS estimator might be inefficient. We will see that, for some vectors of weights $\vec{\omega}^2 = \left(\omega_1^2, ..., \omega_T^2\right)'$, the OLS estimator is asymptotically outperformed by the weighted least squares (WLS) estimator $\hat{\vec{\alpha}}_{\text{WLS}} = \hat{\vec{\alpha}}^{\vec{\omega}^2}_{\text{WLS}}$ defined by

$$(8) \quad Q_N^{\vec{\omega}^2}(\hat{\vec{\alpha}}_{\text{WLS}}) = \min_{\vec{\alpha} \in \Omega_\delta} Q_N^{\vec{\omega}^2}(\vec{\alpha}) \qquad \text{where} \qquad Q_N^{\vec{\omega}^2}(\vec{\alpha}) = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{\nu=1}^{T} \omega_\nu^{-2} e^2_{nT+\nu}(\vec{\alpha}).$$

We will also see that an optimal WLS estimator is the generalized least squares (GLS) estimator

$$(9) \quad \hat{\vec{\alpha}}_{\text{GLS}} = \hat{\vec{\alpha}}^{\vec{\sigma}_0^2}_{\text{WLS}}.$$

The GLS estimator assumes that $\vec{\sigma}_0^2$ is known. In practice, this parameter has also to be estimated. Given any consistent estimator $\hat{\vec{\sigma}}^2$ of $\vec{\sigma}_0^2$, a quasi-generalized least squares (QLS) estimator of $\vec{\alpha}_0$ is defined by

$$(10) \quad \hat{\vec{\alpha}}_{\text{QLS}} = \hat{\vec{\alpha}}^{\hat{\vec{\sigma}}^2}_{\text{WLS}}.$$

One possible consistent estimator of σ_ν^2 is $\hat{\sigma}_\nu^2 = \frac{1}{N}\sum_{n=0}^{N-1} e^2_{nT+\nu}(\hat{\vec{\alpha}}_{\text{OLS}})$. To establish the consistency of the least squares estimators, an additional assumption is needed.

(A2) The T-dimensional white noise $\left\{\boldsymbol{\epsilon}_n = (\epsilon_{nT+1}, ..., \epsilon_{nT+\nu}, ..., \epsilon_{nT+T})', n \in \mathbb{Z}\right\}$ in (4) is strictly stationary and ergodic.

Theorem 1 *Under Assumptions (A1) and (A2), and for any* $\vec{\omega}^2 = (\omega_1^2, \ldots, \omega_T^2) > 0$, *where the inequality applies element-wise, we have for the LS estimators respectively defined by (7), (8), (9) and (10),*

$$\hat{\vec{\alpha}}_{\text{OLS}} \to \vec{\alpha}_0, \quad \hat{\vec{\alpha}}^{\vec{\omega}^2}_{\text{WLS}} \to \vec{\alpha}_0, \quad \hat{\vec{\alpha}}_{\text{GLS}} \to \vec{\alpha}_0, \quad \hat{\vec{\alpha}}_{\text{QLS}} \to \vec{\alpha}_0, \qquad \textit{almost surely as } N \to \infty.$$

See FRS for the proof. Let $||\mathbf{Z}||_r = \left[\mathbb{E}||\mathbf{Z}||^r\right]^{1/r}$ where $||.||$ stands for the Euclidean norm of a vector. In addition to Assumptions (A1) and (A2), we need the following assumption to establish the asymptotic normality.

(A3) The T-variate stationary process $\{\mathbf{X}_n\}$ is such that for some $\tau > 0$, $||\mathbf{X}_n||_{4+2\tau} < \infty$ and $\sum_{h=0}^{\infty} [\underline{\alpha}_{\mathbf{X}}(h)]^{\frac{\tau}{2+\tau}} < \infty$ where the $\underline{\alpha}_{\mathbf{X}}(h)$'s are the strong mixing coefficients of $\{\mathbf{X}_n, n \in \mathbb{Z}\}$.

Theorem 2 *Under the assumptions of Theorem 1 and (A3), as* $N \to \infty$,

$$\sqrt{N}(\hat{\vec{\alpha}}_{\text{LS}} - \vec{\alpha}_0) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{V}_{\text{LS}})$$

where the subscript LS *stands for* OLS, WLS, GLS *or* QLS, *and where*

$$\mathbf{V}_{\mathrm{OLS}} = \mathbf{V}\left(\vec{\alpha}_0, (1,\ldots,1)'\right), \quad \mathbf{V}_{\mathrm{WLS}} = \mathbf{V}(\vec{\alpha}_0, \vec{\omega}^2), \quad \mathbf{V}_{\mathrm{GLS}} = \mathbf{V}_{\mathrm{QLS}} = \mathbf{V}(\vec{\alpha}_0, \vec{\sigma}_0^2),$$

with

$$(11)\quad \mathbf{V}(\vec{\alpha}_0, \vec{\omega}^2) = \left(\mathbf{J}(\vec{\alpha}_0, \vec{\omega}^2)\right)^{-1} \mathbf{I}(\vec{\alpha}_0, \vec{\omega}^2) \left(\mathbf{J}(\vec{\alpha}_0, \vec{\omega}^2)\right)^{-1},$$

$$\mathbf{I}(\vec{\alpha}_0, \vec{\omega}^2) = \sum_{\nu=1}^{T}\sum_{\nu'=1}^{T} \omega_\nu^{-2}\omega_{\nu'}^{-2} \sum_{k=-\infty}^{\infty} \mathbb{E}\left[\left(\epsilon_\nu(\vec{\alpha}_0)\left(\frac{\partial \epsilon_\nu(\vec{\alpha})}{\partial \vec{\alpha}}\right)_{\vec{\alpha}=\vec{\alpha}_0}\right)\left(\epsilon_{kT+\nu'}(\vec{\alpha}_0)\left(\frac{\partial \epsilon_{kT+\nu'}(\vec{\alpha})}{\partial \vec{\alpha}}\right)_{\vec{\alpha}=\vec{\alpha}_0}\right)'\right]$$

and

$$\mathbf{J}(\vec{\alpha}_0, \vec{\omega}^2) = \sum_{\nu=1}^{T} \omega_\nu^{-2} \mathbb{E}\left[\left(\frac{\partial \epsilon_\nu(\vec{\alpha})}{\partial \vec{\alpha}}\right)_{\vec{\alpha}=\vec{\alpha}_0} \left(\frac{\partial \epsilon_\nu(\vec{\alpha})}{\partial \vec{\alpha}}\right)'_{\vec{\alpha}=\vec{\alpha}_0}\right].$$

The proof of this theorem and of the following remarks are given in FRS.

Remark 1 In the periodic AR case with $\vec{\mu} = (\mu_1, \ldots, \mu_T)' = \vec{0}$, the OLS and WLS estimators coincide.

Remark 2 In the strong PARMA setting, we obtain that

$$\mathbf{I}(\vec{\alpha}_0, \vec{\sigma}_0^2) = \sum_{\nu=1}^{T} \sigma_{0\nu}^{-2} \mathbb{E}\left[\left(\frac{\partial \epsilon_\nu(\vec{\alpha})}{\partial \vec{\alpha}}\right)_{\vec{\alpha}=\vec{\alpha}_0} \left(\frac{\partial \epsilon_\nu(\vec{\alpha})}{\partial \vec{\alpha}}\right)'_{\vec{\alpha}=\vec{\alpha}_0}\right] = \mathbf{J}(\vec{\alpha}_0, \vec{\sigma}_0^2).$$

This implies that the asymptotic covariance matrix of the QLS estimators for a PARMA model with independent errors is

$$(12)\quad \mathbf{V}_{\mathrm{QLS}} = \left(\mathbf{J}(\vec{\alpha}_0, \vec{\sigma}_0^2)\right)^{-1}.$$

This result was obtained by Basawa and Lund (2001).

Remark 3 If $\{\epsilon_t\}$ is a martingale difference such that $\mathbb{E}\left[\epsilon_t^2 \middle| \mathcal{F}_{t-1}\right] = \mathbb{E}\left[\epsilon_t^2\right]$ where $\mathcal{F}_t$ is the σ-field spanned by $\{\epsilon_s, s \le t\}$, the QLS estimator is an optimal LS estimator in the sense that $\mathbf{V}_{\mathrm{WLS}} - \mathbf{V}_{\mathrm{QLS}}$ is a semi-definite positive matrix.

Remark 4 It can be shown that $\mathbf{J}(\vec{\alpha}_0, \vec{\omega}^2)$ is consistently estimated by the empirical mean

$$\hat{\mathbf{J}}(\vec{\alpha}_0, \vec{\omega}^2) = \sum_{\nu=1}^{T} \omega_\nu^{-2} \frac{1}{N} \sum_{n=0}^{N-1} \left[\left(\frac{\partial e_{nT+\nu}(\vec{\alpha})}{\partial \vec{\alpha}}\right)_{\vec{\alpha}=\hat{\vec{\alpha}}_{\mathrm{LS}}} \left(\frac{\partial e_{nT+\nu}(\vec{\alpha})}{\partial \vec{\alpha}}\right)'_{\vec{\alpha}=\hat{\vec{\alpha}}_{\mathrm{LS}}}\right].$$

Note that the matrix $(2\pi)^{-1}\mathbf{I}(\vec{\alpha}_0, \vec{\omega}^2)$ is the spectral density at frequency zero of the process

$$\Upsilon_n = \sum_{\nu=1}^{T} \omega_\nu^{-2} \epsilon_{nT+\nu}(\vec{\alpha}_0) \left(\frac{\partial \epsilon_{nT+\nu}(\vec{\alpha})}{\partial \vec{\alpha}}\right)_{\vec{\alpha}=\vec{\alpha}_0}.$$

A similar approach was followed in Francq, Roy and Zakoan (2005). For the numerical illustrations presented in FRS, we used a VAR spectral estimator consisting in: i) fitting VAR(p) models for $p = 0, \ldots, p_{\max}$ to the series $\hat{\Upsilon}_n$, $n = 0, \ldots, N-1$, where $\hat{\Upsilon}_n$ is obtained by replacing $\epsilon_{nT+\nu}(\vec{\alpha}_0))$ and its derivatives by $e_{nT+\nu}(\hat{\vec{\alpha}}_{\mathrm{LS}})$ and its derivatives in Υ_n; ii) selecting the order p which minimizes an information criteria and approximating $\mathbf{I}(\vec{\alpha}_0, \vec{\omega}^2)$ by (2π) times the spectral density at frequency zero of the estimated VAR(p) model.

4- Example of covariance matrix calculations

Consider the following m-dependent periodic white noise with $T = 2$:

$$\epsilon_{nT+\nu} = \sigma_\nu \prod_{j=0}^{m} \xi_{nT+\nu-j}, \qquad \nu = 1, 2,$$

and assume that the iid sequence $\{\xi_t\}$ is such that $\mathbb{E}[\xi_t] = 0$, $\mathbb{E}\left[\xi_t^2\right] = 1$ and $\mathbb{E}\left[\xi_t^4\right] = \kappa < \infty$.

From a realization $X_t = \epsilon_t, t = 1, ..., NT$, of that weak white noise, suppose that a statistician fits the following $\text{PAR}_2(1)$ model:

$$\begin{cases} X_{nT+1} - \phi_1 X_{nT} = \epsilon_{nT+1}, \\ X_{nT+2} - \phi_2 X_{nT+1} = \epsilon_{nT+2}. \end{cases}$$

The true parameter values are $\phi_1 = \phi_2 = 0$ and $\vec{\sigma}^2 = \vec{\sigma}_0^2$. According to Theorem 2, some moment calculations show that the asymptotic covariance matrix of the QLS estimator of $\sqrt{N}\left(\hat{\phi}_1, \hat{\phi}_2\right)'$ is given by

$$\mathbf{V}_{QLS}^{(w)} = \left(\mathbf{J}(\vec{\alpha}_0, \vec{\sigma}_0^2)\right)^{-1} \mathbf{I}(\vec{\alpha}_0, \vec{\sigma}_0^2) \left(\mathbf{J}(\vec{\alpha}_0, \vec{\sigma}_0^2)\right)^{-1} = \kappa^m \begin{pmatrix} \frac{\sigma_{01}^2}{\sigma_{02}^2} & 0 \\ 0 & \frac{\sigma_{02}^2}{\sigma_{01}^2} \end{pmatrix}.$$

On the other hand from (12), the corresponding asymptotic covariance matrix under the assumption of a strong noise is equal to

$$\mathbf{V}_{QLS}^{(s)} = \left(\mathbf{J}(\vec{\alpha}_0, \vec{\sigma}_0^2)\right)^{-1} = \begin{pmatrix} \frac{\sigma_{01}^2}{\sigma_{02}^2} & 0 \\ 0 & \frac{\sigma_{02}^2}{\sigma_{01}^2} \end{pmatrix}.$$

It is clear that $\mathbf{V}_{QLS}^{(w)}$ and $\mathbf{V}_{QLS}^{(s)}$ can be very different. For example, if the iid sequence $\{\xi_t\}$ is $\mathcal{N}(0,1)$, $\mathbb{E}\left[\xi_t^4\right] = 3$, and the discrepancy between the two matrices is important even for small m. It may lead the statistician to wrongly reject the hypothesis that $\phi_1 = \phi_2 = 0$ if he does not take into account the dependence of the errors ϵ_t.

REFERENCES

Basawa, I. V. and Lund, R. (2001). Large sample properties of parameter estimates for periodic ARMA models. *Journal of Time Series Analysis* 22, 651-663.

Francq, C., Roy, R. and Saidi, A. (2009). Asymptotic properties of weighted least squares estimation in weak PARMA models. Technical Report No. 3292, Centre de recherches mathématiques, Université de Montréal and MPRA Paper No. 28721 (http://mpra.ub.uni-muenchen.de/28721/). To appear in *Journal of Time Series Analysis.*

Francq, C., Roy, R. and Zakoan, J. M. (2005). Diagnostic checking in ARMA models with uncorrelated errors. *Journal of the American Statistical Association* 100, 532–544.

Roy, R. and Saidi A. (2008). Temporal aggregation and systematic sampling in PARMA processes. *Computational Statistics and Data analysis* 52, 4287-4304.

Vecchia, A. V. (1985). Periodic autoregressive-moving average modeling with applications to water resources. *Water Resources Bulletin* 21, 721-730.