Publ. Math. Debrecen
49 / 3-4 (1996), 211–218

Statistical inference for multidimensional AR processes

By GYULA PAP (Debrecen) and KATALIN VARGA (Debrecen)

Abstract. It is shown that the suitably normalized maximum likelihood estimator of some parameters of multidimensional autoregressive processes with coefficient matrix of a special structure have exactly a normal distribution.

1. Introduction

Consider the 2–dimensional real–valued stationary autoregressive process $X(t)$, $t \geq 0$, given by the stochastic differential equation (SDE)

$$\begin{pmatrix} dX_1(t) \\ dX_2(t) \end{pmatrix} = \begin{pmatrix} -\lambda & -\omega \\ \omega & -\lambda \end{pmatrix} \begin{pmatrix} X_1(t)\,dt \\ X_2(t)\,dt \end{pmatrix} + \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix},$$

where $W(t) = (W_1(t), W_2(t))$, $t \geq 0$, is a standard 2–dimensional Wiener process and $\lambda > 0$, $\omega \in \mathbb{R}$ are unknown parameters. This process is a so–called 2–dimensional Ornstein–Uhlenbeck process.

Now consider the following statistics:

$$s_X^2(t) = \int_0^t (X_1^2(u) + X_2^2(u))du,$$

$$r_X(t) = \int_0^t (X_1(u)\,dX_2(u) - X_2(u)\,dX_1(u)).$$

As it is known the maximum likelihood estimator (MLE) of the parameter ω is given by

$$\widehat{\omega}_X(t) = \frac{r_X(t)}{s_X^2(t)},$$

Mathematics Subject Classification: 62M10.

and the following holds

$$s_X(t)(\widehat{\omega}_X(t)-\omega) \overset{\mathcal{D}}{=} \mathcal{N}(0,1) \qquad \text{for all } t>0,$$

where $\overset{\mathcal{D}}{=}$ denotes equality in distribution. This result was formulated in Arató, Kolgomorov, Sinay [2], and gives not only an asymptotic property but an exact distribution.

We are interested in the multidimensional generalization of the above result. Let $X(t)=(X_1(t),\dots,X_d(t))'$, $t\geq 0$, prime means transposed, be the d–dimensional process given by the stochastic differential equation

$$dX(t)=AX(t)dt+dW(t), \qquad X(0)=0,$$

where $W(t)$, $t\geq 0$, is a standard d–dimensional Wiener process with independent components and A is a $d\times d$ matrix. The following question arises: what type of conditions should be assumed on the matrix A in order that the suitably normalized MLE of its certain entries will have exactly a normal distribution?

G. Pap and M. C. A. van Zuijlen [6] studied d–dimensional processes of the special form

$$dX(t)=\left(-\lambda I_d+\sum\nolimits_{i=1}^{m}\omega_i C_i\right)X(t)\,dt+dW(t), \qquad X(0)=0 \tag{1}$$

where I_d is the $d\times d$ unit matrix, $\lambda,\omega_1,\dots,\omega_m\in\mathbb{R}$ are unknown parameters and $C_1,\dots,C_m$ are fixed $d\times d$ skew–symmetric matrices, i.e., $C_i'=-C_i$, $i=1,\dots,m$. The maximum likelihood estimator of $\omega=(\omega_1,\dots,\omega_m)'$ is given by

$$\widehat{\omega}_X(t)=\sigma_X^{-1}(t)r_X(t),$$

where $\sigma_X(t)$ is the $m\times m$ matrix

$$\sigma_X(t)=\left(\int_0^t\langle C_iX(s),C_jX(s)\rangle ds\right)_{1\leq i,j\leq m},$$

and $r_X(t)$ is the m–dimensional column vector

$$r_X(t)=\left(\int_0^t\langle C_iX(s),dX(s)\rangle\right)'_{1\leq i\leq m}.$$

In [6] it is proved that

$$\sigma_X^{1/2}(t)(\widehat{\omega}_X(t)-\omega) \overset{\mathcal{D}}{=} \mathcal{N}(0,I_m), \qquad \text{for all } t>0, \tag{2}$$

if conditions (C1)–(C3) are satisfied, where

(C1) $C_i' = -C_i$, $i = 1, \dots, m$,

(C2) $(C_iC_j + C_jC_i)C_k = C_k(C_iC_j + C_jC_i)$, $i, j, k = 1, \dots, m$,

(C3) $(C_iC_j + C_jC_i)(C_kC_\ell + C_\ell C_k) \in \mathcal{L}(C_uC_v, 1 \le u, v \le m)$, $i, j, k, \ell = 1, \dots, m$,

where $\mathcal{L}(C_uC_v, 1 \le u, v \le m)$ denotes the linear hull of the matrices $C_uC_v, 1 \le u, v \le m$. The main purpose of this paper is to show that the condition (C3) is superfluous.

Theorem. *Let $X(t)$, $t \ge 0$, be the process given by* (1). *Let us suppose that the conditions* (C1) *and* (C2) *are satisfied. Then* (2) *holds.*

In Section 2 some preparatory lemmas are given. We prove the Theorem in Section 3. Section 4 contains some special cases. It should be remarked that we consider only processes $X(t)$, $t \ge 0$, with initial value $X(0) = 0$, but the results can be extended for processes with random initial value $X(0) = \xi$ having absolutely continuous distribution which does not depend on the parameter ω, as in [6]. This extension of the results cover the stationary solution of the SDE (1).

2. Preliminaries

We shall make use of the following explicit formula which is a special case of Lemma 11.6 in [4].

Lemma 1. *Consider a standard d–dimensional Wiener process $W(t)$, $t \ge 0$. For all $t \ge 0$ let $B(t)$ and $Q(t)$ be $d \times d$ matrices such that $Q(t)$ is symmetric, positive semidefinite and*

$$\operatorname{Tr} \int_0^T (B(t)B'(t) + Q(t))\, dt < \infty. \tag{3}$$

Then

$$\mathbb{E} \exp\left\{ -\int_0^T \left(\int_0^t B(s)\, dW(s) \right)' Q(t) \left(\int_0^t B(s)\, dW(s) \right) dt \right\} = \exp\left\{ \frac{1}{2} \operatorname{Tr} \int_0^T B(t)B'(t)\Gamma(t)\, dt \right\},$$

where $\Gamma(t)$, $t \ge 0$, are negative semidefinite matrices determined by the Riccati differential equation

$$\dot{\Gamma}(t) = 2Q(t) - \Gamma(t)B(t)B'(t)\Gamma(t), \qquad \Gamma(T) = 0.$$

Let us denote the cone of the symmetric, positive semidefinite $d \times d$ matrices by $\mathcal{C}_d$. We shall also use that the distribution of a symmetric, positive semidefinite $d \times d$ random matrix is uniquely determined by the value of its Laplace transform on the cone $\mathcal{C}_d$.

Lemma 2. *If σ is a random matrix with $\sigma' = \sigma$ and $\sigma \geq 0$ then the distribution of σ is uniquely determined by the Laplace transform $\psi : \mathcal{C}_d \to (0, \infty)$ given by*

$$\psi(\alpha) := \mathbb{E} \exp\{-\operatorname{Tr}(\alpha'\sigma)\} = \mathbb{E} \exp\left\{ -\sum_{i=1}^{d} \sum_{j=1}^{d} \alpha_{ij}\sigma_{ij} \right\}, \qquad \alpha \in \mathcal{C}_d.$$

Proof. First we prove that for $\alpha \in \mathcal{C}_d$ we have $\operatorname{Tr}(\alpha'\sigma) \geq 0$. It is well known that there is a matrix $\beta \in \mathcal{C}_d$ such that $\alpha = \beta^2 = \beta'\beta$. The matrix $\beta\sigma\beta'$ is again symmetric and positive definite since

$$\langle \beta\sigma\beta' x, x \rangle = \langle \sigma(\beta' x), (\beta' x) \rangle \geq 0, \qquad x \in \mathbb{R}^d.$$

Hence, indeed

$$\operatorname{Tr}(\alpha'\sigma) = \operatorname{Tr}(\beta'\beta\sigma) = \operatorname{Tr}(\beta\sigma\beta') \geq 0.$$

For fixed $k \in \{1, \ldots, d\}$ let us consider the matrix $\alpha^{(k)} \in \mathcal{C}_d$ with entries

$$\alpha_{ij}^{(k)} = \begin{cases} 1 & \text{if } i = j = k, \\ 0 & \text{else.} \end{cases}$$

Then $\operatorname{Tr}\left((\alpha^{(k)})'\sigma\right) = \sigma_{kk}$.

For fixed $k, \ell \in \{1, \ldots, d\}$, $k \neq \ell$, let us consider the matrix $\alpha^{(k\ell)} \in \mathcal{C}_d$ with entries

$$\alpha_{ij}^{(k\ell)} = \begin{cases} 1 & \text{if } i, j \in \{k, \ell\}, \\ 0 & \text{else.} \end{cases}$$

Then $\operatorname{Tr}\left((\alpha^{(k\ell)})'\sigma\right) = \sigma_{kk} + 2\sigma_{k\ell} + \sigma_{\ell\ell}$.

Using the classical result on the Laplace transform of a random vector with nonnegative coordinates we know that the joint distribution of the random variables

$$\{\sigma_{kk} : 1 \leq k \leq d\} \cup \{\sigma_{kk} + 2\sigma_{k\ell} + \sigma_{\ell\ell} : 1 \leq k < \ell \leq d\}, \tag{4}$$

is uniquely determined by the Laplace transform

$$\varphi(s_k, 1 \leq k \leq d;\, s_{k\ell}, 1 \leq k < \ell \leq d) := \mathbb{E} \exp\left\{ -\sum_{k=1}^{d} s_k \sigma_{kk} - \sum_{1 \leq k < \ell \leq d} s_{k\ell}\sigma_{k\ell} \right\}, \quad s_k, s_{k\ell} \geq 0.$$

Clearly

$$
\begin{aligned}
&\varphi(s_k, 1 \le k \le d;\ s_{k\ell}, 1 \le k < \ell \le d) \\
&\qquad = \mathbb{E} \exp\left\{ - \sum_{k=1}^{d} s_k \operatorname{Tr}\left((\alpha^{(k)})' \sigma \right) - \sum_{1 \le k < \ell \le d} s_{k\ell} \operatorname{Tr}\left((\alpha^{(k\ell)})' \sigma \right) \right\} \\
&\qquad = \mathbb{E} \exp\{ - \operatorname{Tr}(\alpha' \sigma) \} = \psi(\alpha),
\end{aligned}
$$

where

$$
\alpha = \sum_{k=1}^{d} s_k \alpha^{(k)} + \sum_{1 \le k < \ell \le d} s_{k\ell} \alpha^{(k\ell)} \in \mathcal{C}_d .
$$

Consequently the joint distribution of the random variables in (4) is uniquely determined by the Laplace transform $\psi : \mathcal{C}_d \to (0, \infty)$ of the random matrix σ, hence, the joint distribution of the entries of the matrix σ is also uniquely determined by $\psi : \mathcal{C}_d \to (0, \infty)$ since there is a one-to-one correspondence between the entries of σ and the random variables in (4). $\square$

3. Proof of the Theorem

The proof can be carried out as in [6]. We have to show only that for all $T > 0$ the distribution of the symmetric, positive semidefinite random matrix $\sigma_X(T)$ does not depend on the parameter $\omega = (\omega_1, \dots, \omega_m)'$. Using Lemma 2 it is sufficient to show that the Laplace transform

$$
\Psi_T(\alpha) = \mathbb{E} \exp\left\{ - \sum_{i,j=1}^{m} \alpha_{i,j} \int_0^T \langle C_i X(t), C_j X(t) \rangle \, dt \right\}, \qquad \alpha \in \mathcal{C}_d,
$$

does not depend on the parameter ω. Using the notation

$$
C := \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_{ij} C_i' C_j ,
$$

we have

$$
\Psi_T(\alpha) = \mathbb{E} \exp\left\{ - \int_0^T X'(t) C X(t) \, dt \right\}.
$$

Next we show that C is a symmetric, positive semidefinite matrix. We use again that there exists a matrix $\beta \in \mathcal{C}_d$ such that $\alpha = \beta^2 = \beta'\beta$, hence $\alpha_{ij} = \sum_{k=1}^d \beta_{ki}\beta_{kj}$. We have

$$\langle Cx, x\rangle = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^d \beta_{ki}\beta_{kj} \langle C_i' C_j x, x\rangle = \sum_{k=1}^d \left| \sum_{i=1}^m \beta_{ki} C_i x \right|^2 \geq 0,$$

thus $C \in \mathcal{C}_d$, indeed.

Let

$$A = -\lambda I_d + \sum\nolimits_{i=1}^m \omega_i C_i.$$

It is known that the solution $X(t)$, $t \geq 0$, of the SDE (1) can be represented in the form

$$X(t) = \int_0^t e^{(t-s)A} dW(s).$$

Consequently,

$$\int_0^T X'(t) C X(t)\, dt = \int_0^T \left(\int_0^t e^{-sA} dW(s) \right)' e^{tA'} C e^{tA} \left(\int_0^t e^{-sA} dW(s) \right) dt.$$

We will show that Lemma 1 can be applied with $B(t) = e^{-tA}$ and $Q(t) = e^{tA'} C e^{tA}$. Clearly the conditions (C1) and (C2) imply

$$B(t)B'(t) = e^{2\lambda t} I_d$$

and $AC = CA$, hence

$$Q(t) = C e^{tA'} e^{tA} = e^{-2\lambda t} C,$$

and we conclude the validity of the condition (3). Applying Lemma 1 and using the above formulae we obtain

$$\Psi_T(\alpha) = \exp\left\{ \frac{1}{2} \operatorname{Tr} \int_0^T e^{2\lambda t} \Gamma(t)\, dt \right\}, \qquad \alpha \in \mathcal{C}_d,$$

where $\Gamma(t)$, $t \geq 0$, is defined by

$$\dot{\Gamma}(t) = 2e^{-2\lambda t} C - e^{2\lambda t} \Gamma^2(t), \qquad \Gamma(T) = 0.$$

Consequently the Laplace transform Ψ_T does not depend on the parameter ω and the proof is completed. $\square$

4. Special cases

We give some application of the Theorem.

Corollary 1. *Consider the d-dimensional process $X(t)$, $t \geq 0$, given by*

$$dX(t) = \left(-\lambda I + \sum\nolimits_{i=1}^{m} \omega_i C_i\right) X(t)\,dt + dW(t), \qquad X(0) = 0,$$

where

$C_i' = -C_i$, $i = 1, \ldots, m$,

$C_i C_j = -C_j C_i$, $1 \leq i < j \leq m$.

Then the maximum likelihood estimators of the parameters $\omega_1, \ldots, \omega_m$ are given by

$$\widehat{\omega}_X^{(i)}(t) = \frac{r_X^{(i)}(t)}{\left(s_X^{(i)}(t)\right)^2},$$

where

$$r_X^{(i)}(t) = \int_0^t \langle C_i X(s), dX(s)\rangle, \qquad \left(s_X^{(i)}(t)\right)^2 = \int_0^t |C_i X(s)|^2 ds,$$

and

$$\left(s_X^{(1)}(t)\left(\widehat{\omega}_X^{(1)} - \omega_1\right), \ldots, s_X^{(m)}(t)\left(\widehat{\omega}_X^{(m)} - \omega_m\right)\right) \overset{\mathcal{D}}{=} \mathcal{N}(0, I_m),$$

for all $t > 0$.

Corollary 2. *Consider the d-dimensional process $X(t)$, $t \geq 0$, given by*

$$dX(t) = (-\lambda I + \omega C) X(t)\,dt + dW(t), \qquad X(0) = 0,$$

where $C' = -C$.

Then the maximum likelihood estimator of the parameter ω is

$$\widehat{\omega}_X(t) = \frac{r_X(t)}{s_X^2(t)},$$

where

$$r_X(t) = \int_0^t \langle C X(s), dX(s)\rangle, \qquad s_X^2(t) = \int_0^t |C X(s)|^2 ds,$$

and

$$s_X(t)\left(\widehat{\omega}_X(t) - \omega\right) \overset{\mathcal{D}}{=} \mathcal{N}(0, 1), \qquad \textit{for all } t > 0.$$

References

[1] M. Arató, Linear stochastic systems with constant coefficients. A statictical approach. (Lecture Notes in Control and Inf., Vol. 45, 309 pp.), *Springer-Verlag, Berlin*, 1982. (in Russian, Nauka, Moscow, 1989)

[2] M. Arató, A. N. Kolmogorov and Ya. G. Sinay, Estimation of the parameters of a complex stationary Gaussian Markov Process, *Dokl. Akad. Nauk SSSR* **146** (1962), 747–750.

[3] I. Fazekas, Maximum likelihood estimation of parameters of multidimensional stationary AR processes, *Computers Math. Appl.* **27** (1994), 19–24.

[4] R. S. Liptser and A. N. Shiryayev, Statistics of Random Processes I, II, *Springer-Verlag*, *New York*, 1977.

[5] G. Pap, On the distribution of estimates of parameters of multidimensional stationary AR processes, *Computers Math. Appl.* **27** (1994), 1–8.

[6] G. Pap and M. C. A. van Zuijlen, Parameter estimation with exact distribution for multidimensional Ornstein–Uhlenbeck process, *Report 9340, Dept. of Math. Catholic Univ. Nijmegen,* (1993); (to appear in J. Multivariate Anal.).

GYULA PAP
INSTITUTE OF MATHEMATICS AND INFORMATICS
LAJOS KOSSUTH UNIVERSITY
H–4010 DEBRECEN
HUNGARY

KATALIN VARGA
INSTITUTE OF MATHEMATICS AND INFORMATICS
LAJOS KOSSUTH UNIVERSITY
H–4010 DEBRECEN
HUNGARY

(Received August 24, 1994; revised November 1, 1995)