ELSEVIER

Statistics & Probability Letters 59 (2002) 37–52

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

Maximum of entropy and extension of covariance matrices for periodically correlated and multivariate processes

Glaysar Castro[a], Valerie Girardin[b,*]

[a] *Escuela de Matematicas, A.P.20513-Caracas 1020A, Venezuela*
[b] *Dept. de Mathématiques, Campus II, Université de Caen, BP5186, Caen 14032, France*

Received January 2001; received in revised form April 2002

Abstract

A one-to-one relationship exists between scalar periodically correlated nonstationary processes and multivariate stationary processes. This fact allows us to transfer results proven for ones to the others. We are interested in a probabilistic approach of results sometimes already known in a different (analytical or numerical) context, in order to simplify, generalize and unify them. We use a probabilistic approach of generalized reflection coefficients to give a constructive condition of extension of partial covariance sequences, achieved by an autoregressive model. We develop a Trench–Zohar recursion for the nonstationary case which leads to an economical algorithm to solve the associated Yule–Walker equations. Shannon and Burg entropies are linked through a Szëgo type theorem. A numerical example is given.

MSC: 15A23; 62M10; 62M15; 62M20; 94A17

Keywords: Periodically correlated processes; Nonstationary processes; Multivariate stationary processes; Maximum entropy; Reflection coefficients; Auto-regressive processes

1. Introduction

Periodically correlated data are often encountered in applications, see Franses (1996) and the references therein: for instance in pollution study (see Tiao and Grupe (1980) for an example in the ozone concentration in towns, with period $d=4$), in river flow forecasting (see McLeod (1993) for a method for determining d), in meteorological studies (see Monin, 1963), in geophysical sciences (see Troutman, 1979), in econometrical time series (see Pagano and Parzen, 1979), etc. Their modeling

* Corresponding author.
E-mail addresses: gcastro@euler.ciens.ucv.ve (G. Castro), girardin@math.unicaen.fr (V. Girardin).

0167-7152/02/$ - see front matter
PII: S0167-7152(02)00159-1

by stationary time series may bring problems (see Tiao and Grupe, 1980), hence the interest of PC processes.

A one-to-one relationship exists between scalar PC processes and multivariate stationary (or MS) processes, see Gladyshev (1961). This fact allows us to transfer results proven for ones to the others. We consider MS processes having a spectral density. It is a Hermitian matrix valued function which may be considered as a spectral density for the dual PC process which, being nonstationary, does not have a natural one.

We are here concerned with the construction of maximum entropy extensions of partial d-variate covariance sequences for PC and MS processes. Such a sequence is a positive definite finite sequence $C=(C(n))_{-N\leqslant n\leqslant N}$ of complex $d\times d$-matrices, i.e., such that the related Toeplitz block matrix (called partial covariance matrix) $C_N=(C(k-l))_{0\leqslant k,l\leqslant N}$ is positive definite. We call extensions of C (or of C_N) indifferently MS or PC processes realizing it or their covariances.

A necessary and sufficient condition for an extension of C to exist is the positive definiteness of the matrix C_N. It has been proven by analytical methods, see Constantinescu (1996), and by band methods, see Gohberg et al. (1989). We will give here a probabilistic constructive proof in which the extension is realized by an auto-regressive model. We will show that its spectral density satisfies a maximum entropy property. The classical Burg entropy can be applied to MS (or by duality to PC) processes X under the form

$$\mathscr{I}[H]=\int_{\Pi}\ln\operatorname{Det}H(\lambda)\,\mathrm{d}\lambda$$

(with $\mathscr{I}[H]=-\infty$, if the integral is not defined). We will call n-step entropy the Shannon entropy of the distribution of the vector $(X(0),\ldots,X(n))$, namely

$$-\int_{\mathbf{R}^n}p_n(x_0,\ldots,x_n)\ln\,p_n(x_0,\ldots,x_n)\,\mathrm{d}x_0\ldots\,\mathrm{d}x_n.$$

For a Gaussian process, the n-step entropy is equal (up to an additive constant) to the determinant of its partial covariance matrix of size n. Divided by n, it converges to the Burg entropy. This result, classical for scalar processes, is shown in Section 5 for MS processes. It provides a link between the two maximization problems. Moreover, both problems have the same solution, the auto-regressive model of order N realizing the required partial covariance sequence. For the n-step entropy, this has been proven by Choi for scalar (1987), for MS (1993) and for general Choi (1998) processes and by Zhang (1997) for PC processes. For the Burg entropy, it has been proven by Burg (see Choi and Cover, 1984) for scalar stationary processes. And it is proven for the MS case in Section 5 below, by a method based on Helson and Lowdenslager (1958) results, similar to the method used by Seghier (1991) for scalar multidimensional processes.

The matrix coefficients of the auto-regressive model can be computed via the Yule–Walker equations using the matrix C_N. It is well-known that for a scalar stationary process, an efficient method for inverting this matrix is the Trench–Zohar algorithm which involves the reflection coefficients of the process X. The reflection coefficients are then defined as the partial autocorrelations (see Brockwell and Davis, 1996), i.e., as the correlations that subsist between $X(k)$ and $X(l)$ after eliminating the linear effect of the intermediate variables. In the MS case, this definition involves the square roots of a matrix, which are not unique. Several methods for choosing a square root have been developed, see Dégerine (1994) and the references therein.

Different generalizations of the reflection coefficients have been developed in the scalar nonstationary and MS cases, see Morf et al. (1978) or Constantinescu (1996) for operator theory approaches and Delsarte et al. (1979) for an analytic approach. We use here a probabilistic approach introduced by Castro (1997). In this approach, each coefficient is replaced by a triangular array of pairs of coefficients linked to the right and left prediction errors. They are defined naturally for PC processes and the duality allows us in Section 3 to compute the solution to the multivariate Yule–Walker equations. The related scalar nonstationary version of the Trench–Zohar algorithm involves no square root matrix. This algorithm has a complexity of order $O(dN^2)$ instead of the order $O(N^3)$ necessary to invert an $N \times N$ matrix in the general case.

This paper is organized as follows. Notation and definitions are given in Section 2. In Section 3, the generalized reflection coefficients are defined and a nonstationary version of the Trench–Zohar recursion is presented, leading to the solution of the Yule–Walker equations. In Section 4, consequences of the recursion are given, as the proof of a sufficient condition for the existence of extensions of a given partial covariance sequence. In Section 5, the maximum Burg entropy extension of a given partial d-variate covariance sequence is shown to be a d-variate auto-regressive process, a Szegö type theorem for the multivariate case is stated and the link between the n-step entropy and the spectral entropy for Gaussian MS processes is established. A numerical example is developed in Section 6.

2. Notation and definitions

A second-order nonstationary scalar unidimensional time series Y is said to be periodically correlated (or PC) if there exists an integer d such that

$$\mathbb{E}Y(t) = \mathbb{E}Y(t+d) \quad \text{and} \quad \mathbb{E}Y(s)Y(t) = \mathbb{E}Y(s+d)Y(t+d), \quad (s,t) \in \mathbb{Z}^2.$$

If $Z_k(t) = Y(k + dt)$ defines the kth component of the d-variate process Z, Gladyshev (1961) proved that Y is PC(d) if and only if Z is weakly stationary. Indeed, for $n, t \in \mathbb{Z}$, $0 \leqslant k,\ l < d$,

$$\mathbb{E}Z_k(t)Z_l(t-n) = \mathbb{E}Y(k+dt)Y(l+d(t-n)) = \mathbb{E}Y(k+dn)Y(l) = \mathbb{E}Z_k(n)Z_l(0).$$

The spectral density of Z is an Hermitian $d \times d$-matrix valued function $H = (h_{kl})$ such that $\mathbb{E}Z_k(t)Z_l(t-n) = \int_{\Pi} h_{kl}(\lambda)e^{in\lambda}\,d\lambda$, for $1 \leqslant k,\ l \leqslant d$ and $\int_{\Pi} \operatorname{Tr} H(\lambda)^* H(\lambda)\,d\lambda < +\infty$, where integration is taken with respect to the Haar measure on the torus Π. The cross-spectral densities of the process h_{kl} for $k \neq l$ are generally complex-valued. The auto-spectral densities h_{kk} are real-valued and nonnegative. Various relations may exist between the cross and auto-spectral densities. See Priestley (1981) for examples and more on multivariate processes.

We consider MS processes realizing a partial d-variate covariance sequence C, i.e., such that

$$\mathbb{E}Z(0)Z(n) = C(n), \quad 0 \leqslant n \leqslant N. \tag{1}$$

Observe that (1) is equivalent to

$$\hat{H}(n) = \left(\int_{\Pi} h_{kl}(\lambda) e^{in\lambda} \, d\lambda \right) = C(n), \quad 0 \leqslant n \leqslant N. \tag{2}$$

We ask H to satisfy a maximum entropy property, hence we are interested in the auto-regressive model of order N realizing the required partial covariance sequence. Following Tiao and Grupe (1980) (see also Franses, 1996), we will say that an MS process Z is an auto-regressive process (or MAR(N)) if it has a representation

$$\sum_{k=0}^{N} A(k) Z(t-k) = \varepsilon(t), \quad t \in \mathbb{Z}, \tag{3}$$

where the coefficients $A(k)$ are $d \times d$ matrices, $A(0)$ is a unit (i.e., with one main diagonal entries) lower triangular matrix and ε is a multivariate white noise process with diagonal covariance matrix Σ_ε.

Note that this representation is equivalent to the more usual representation $\sum_{k=0}^{N} \widetilde{A(k)} Z(t-k) = \tilde{\varepsilon}(t)$, for $t \in \mathbb{Z}$, where $\widetilde{A(0)}$ is the identity matrix I, and $\Sigma_{\tilde{\varepsilon}} = A(0)^{-1} \Sigma_\varepsilon [A(0)^{-1}]^*$ is the modified Cholevski decomposition of the covariance matrix Σ_ε (see Pagano, 1978), and the star denotes conjugate and transpose.

Under suitable assumptions (see Troutman, 1979), Z has a spectral density given by

$$H(\lambda) = P^{-1}(\lambda) \Sigma_\varepsilon [P^{-1}(\lambda)]^* = [I + \tilde{P}(\lambda)]^{-1} \Sigma_{\tilde{\varepsilon}} ([I + \tilde{P}(\lambda)]^{-1})^*, \tag{4}$$

where $P(\lambda) = \sum_{k=0}^{N} A(k) e^{ik\lambda}$ and $\tilde{P}(\lambda) = \sum_{k=1}^{N} \widetilde{A(k)} e^{ik\lambda}$. Reciprocally, any multivariate process with spectral density given by (4) is an MAR process with representation (3).

Similarly, a $PC(d)$ process Y is called a periodic auto-regressive process (or PAR) if it has a representation $Y(t) + \sum_{i=1}^{N_t} \alpha_t(i) Y(t-i) = \varepsilon(t)$, for $t \in \mathbb{Z}$, where $N_t = N_{t+d}$, $\alpha_t(i) = \alpha_{t+d}(i)$ and ε is a white noise process with variance $\sigma_t^2 = \sigma_{t+d}^2$, for $t \in \mathbb{Z}$. This relation can be written as

$$Y(k+ld) + \sum_{i=1}^{N_k} a_k(i) Y(k + d(l-i)) = \varepsilon(k+ld), \quad l \in \mathbb{Z}, \ k = 0, \ldots, d-1, \tag{5}$$

obviously related to relation (3) so that Y is a PAR($d, N_1, \ldots, N_d$) if and only if the dual MS process Z is a MAR(N) with $N = \max_k [(N_k - k)/d] + 1$, where $[\cdot]$ denotes the integer part of a real number, see Pagano (1978).

The matrix coefficients $A(i)$ in (3) can be computed via the Yule–Walker equations with matrix C_N. Since C_N is a positive-definite matrix, it has a unique modified Cholevski factorization $C_N = TDT^*$, where D is a diagonal matrix and T a unit lower triangular matrix, see Horn and Johnson (1985). In the scalar case, the parametric representation of the AR(N) model is $X(t) + \sum_{i=1}^{N} A_i X(t-i) = \varepsilon(t)$, where $(A_1, \ldots, A_N)'$ denotes the first column of T. In the d-variate setting, the covariance matrix C_N can be seen as the partial covariance matrix $\Gamma_{0,N}$ of a scalar PC(d) process and the factorization of its inverse may be interpreted in two different ways: in the d-variate case, the coefficients $A(i)$ of representation (3) are given by the d first columns of T in the triangular factorization of $\Gamma_{0,N}^{-1}$; in the periodic setting, the first d columns of T give the coefficients $\alpha_k(i)$ of the d lines of the representation (5).

3. Generalized reflection coefficients definition

Let $X=(X(t))_{t\in\mathbb{Z}}$ be a scalar nonstationary process defined on a probability space $(\Omega,\mathscr{F},\mathbb{P})$, see (Rozanov, 1967). Assume that $X=\{X(t)\}_{t\in\mathbb{Z}}$ is a basic process, i.e., a basis of the subspace H^X it spans in $L^2(\Omega,\mathscr{F},\mathbb{P})$. Consider the subspaces $H^X_{k,l}=\mathscr{S}\{X(t)\colon k\leqslant t\leqslant l\}$ of H^X spanned by finite sub-families of X. Let $\Pi_{k,l}$ be the orthogonal projector of $L^2(\Omega,\mathscr{F},\mathbb{P})$ onto $H^X_{k,l}$ for the scalar product $\langle P,Q\rangle=\mathbb{E}P\bar{Q}$.

The one-step forward residual given $H^X_{k+1,l}$ and the one-step backward residual given $H^X_{k,l-1}$ are

$$P_{k,l}=X(k)-\Pi_{k+1,l}[X(k)]\quad\text{and}\quad Q_{k,l}=X(l)-\Pi_{k,l-1}[X(l)],\quad k<l\in\mathbb{N},$$

and we set $P_{k,k}=Q_{k,k}=X(k)$. Note that both families $(P_{i,l})_{k\leqslant i\leqslant l}$ and $(Q_{k,j})_{k\leqslant j\leqslant l}$ are basis of $H^X_{k,l}$.

The following theorem constitutes a nonstationary version of the Trench–Zohar recursion for $P_{k,l}$ and $Q_{k,l}$.

Theorem 1. *For any $0\leqslant k<l\leqslant N$, there exists an ordered pair of complex numbers $(r_{k,l},r'_{k,l})$ such that*

$$P_{k,l}=P_{k,l-1}-r_{k,l}Q_{k+1,l}\quad\textit{and}\quad Q_{k,l}=Q_{k+1,l}-r'_{k,l}P_{k,l-1}. \tag{6}$$

The coefficients $r_{k,l}$ and $r'_{k,l}$ are called generalized reflection coefficients.

Proof. Since the process X is basic, we can write

$$\Pi_{k+1,l}[X(k)]=\tilde{X}_{k+1,l-1}-bX(l), \tag{7}$$

where $\tilde{X}_{k+1,l-1}\in H^X_{k+1,l-1}$ and b is a complex number. But $H^X_{k+1,l-1}\subset H^X_{k+1,l}$, so by projection properties $\Pi_{k+1,l-1}\circ\Pi_{k+1,l}=\Pi_{k+1,l-1}$.

Thus, applying $\Pi_{k+1,l-1}$ to (7), we obtain $\Pi_{k+1,l-1}[X(k)]=\tilde{X}_{k+1,l-1}-b\Pi_{k+1,l-1}[X(l)]$. Using this relation in (7) yields $-\Pi_{k+1,l}[X(k)]=-\Pi_{k+1,l-1}[X(k)]+b(X(l)+\Pi_{k+1,l-1}[X(l)])$. Adding $X(k)$ and setting $b=r_{k,l}$ ends the proof of the first part of (6). The proof of the second part is analogous. □

Observe that the recursion follows diagonals: at the nth step, the forward and backward residuals $P_{k,l}$ and $Q_{k,l}$ along the diagonal $l-k=n$ are computed as linear combinations of the corresponding quantities at the $(n-1)$th step. The coefficients $r_{k,l}$ and $r'_{k,l}$ of these linear combinations can be obtained from the residuals of the $(n-1)$th step as stated in the next proposition.

Proposition 1. *For each $(k,l)\in\mathbb{Z}^2$ such that $k<l$, the following relations hold*:

$$r_{k,l}=\langle P_{k,l-1},Q_{k+1,l}\rangle/\|Q_{k+1,l}\|^2\quad\textit{and}\quad r'_{k,l}=\langle Q_{k+1,l},P_{k,l-1}\rangle/\|P_{k,l-1}\|^2, \tag{8}$$

$$r'_{k,l}=\overline{r_{k,l}}\|Q_{k+1,l}\|^2/\|P_{k,l-1}\|^2, \tag{9}$$

$$0\leqslant r_{k,l}r'_{k,l}\leqslant 1. \tag{10}$$

Proof. From the definition, $P_{k,l}$ is orthogonal to $Q_{k+1,l} \in H^X_{k+1,l}$, so that $\langle P_{k,l}, Q_{k+1,l}\rangle = 0$. Taking the inner product by $Q_{k+1,l}$ in the first part of (6) yields $0 = \langle P_{k,l-1}, Q_{k+1,l}\rangle - r_{k,l}\langle Q_{k+1,l}, Q_{k+1,l}\rangle$. Hence the first part of (8) is proven. The proof of the second part is analogous, taking the inner product by $P_{k,l-1}$ in the second part of (6).

Observe that since $P_{k,l}$ is orthogonal to $H^X_{k+1,l}$ and $Q_{k,l}$ is orthogonal to $H^X_{k,l-1}$, the following identities hold:

$$\|P_{k,l}\|^2 = \langle P_{k,l}, P_{k,l}\rangle = \langle P_{k,l}, X(k)\rangle = \langle X(k), P_{k,l}\rangle, \tag{11}$$

$$\|Q_{k,l}\|^2 = \langle Q_{k,l}, Q_{k,l}\rangle = \langle Q_{k,l}, X(l)\rangle = \langle X(l), Q_{k,l}\rangle, \tag{12}$$

$$\langle P_{k,l-1}, Q_{k+1,l}\rangle = \langle P_{k,l-1}, X(l)\rangle = \langle X(k), Q_{k+1,l}\rangle, \tag{13}$$

and relation (9) follows immediately from (8), and (11)–(13).

Finally, we get (10) by a straightforward application of the Cauchy–Shwartz inequality to (8), namely

$$\begin{aligned} r_{k,l}r'_{k,l} &= |\langle P_{k,l-1}, Q_{k+1,l}\rangle|^2 / \|Q_{k+1,l}\|^2 \|P_{k,l-1}\|^2 \\ &\leqslant \|P_{k,l-1}\|^2 \|Q_{k+1,l}\|^2 / \|Q_{k+1,l}\|^2 \|P_{k,l-1}\|^2 = 1. \qquad \square \end{aligned} \tag{14}$$

Note that the square root of (14) gives the definition of partial auto-correlations in the stationary case. Indeed, in this case, $r_{k,l} = r_{k-l,0}$ and $\overline{r_{k,l}} = r'_{k,l}$, since $\|Q_{k+1,l}\| = \|P_{k,l-1}\|$ by (9). Thus, the partial auto-correlation between $X(k)$ and $X(l)$ is naturally defined as the square root of the product $r_{k,l}r'_{k,l}$. In the scalar nonstationary case (so in the MS case too, by duality with the PC case) the backward and forward residuals are not symmetric hence the square root of the product (14) cannot be taken as a generalization of the partial auto-correlations. An ordered pair of coefficients has to be considered, as we have done above.

At each step of the recursion, in the scalar stationary case, only one coefficient has to be computed. Similarly, in the PC(d) case, it is easy to see that $r_{k,l} = r_{k+d,l+d}$, so only d pairs of coefficients have to be computed. This property will allows us in Section 4 to construct economical algorithms for solving the Yule–Walker equations in the periodic case. And the MS case will be obtained by duality.

Let us give now some interesting properties of the reflection coefficients which will be used in Section 4.

First, the covariance kernel $K : \mathbb{N} \times \mathbb{N} \to \mathbb{C}$ of the process X defined by $K(m,n) = \langle X_m, X_n\rangle = \mathbb{E}X(m)\overline{X(n)}$ determines the second-order properties of the process. The value of the pair $(r_{k,l}, r'_{k,l})$ depends only on the values of K in the triangular array $\Delta_{k,l} = \{(m,n)\colon k \leqslant m,n \leqslant l\}$. Conversely, the value $K(k,l)$ may be computed from the triangular array of pairs $\{(r_{m,n}, r'_{m,n})\}_{k\leqslant m,n\leqslant l}$ as shown in the following proposition, which constitutes an aller-retour formula.

Proposition 2. *The value $K(k,l)$ of the covariance kernel K is given by*

$$K(k,l) = \|Q_{k+1,l}\|^2 r_{k,l} + \langle X(k) - P_{k,l-1}, X(l)\rangle. \tag{15}$$

Proof. Using (13), we can write $\langle P_{k,l-1}, Q_{k+1,l}\rangle = \langle P_{k,l-1}, X(l)\rangle = \langle X(k), X(l)\rangle - \langle X(k) - P_{k,l-1}, X(l)\rangle$. Putting it in (8) yields $r_{k,l} = [K(k,l) - \langle X(k) - P_{k,l-1}, X(l)\rangle]/\|Q_{k+1,l}\|^2$, which is equivalent to (15). □

This formula will allow us to extend a partial covariance sequence by extending its associated reflection coefficients sequence in Section 4.

Secondly, the denominators in (8) are the prediction errors obtained at the previous step. The next proposition formulas allows one to compute them at the nth step in terms of the generalized reflection coefficients obtained at the previous $(n-1)$ steps.

Proposition 3. *The following identities hold*:

$$\|P_{k,k+n}\|^2 = \|X(k)\|^2 \prod_{j=k+1}^{k+n} (1 - r_{k,j} r'_{k,j}) \quad \textit{and}$$

$$\|Q_{k,k+n}\|^2 = \|X(k+n)\|^2 \prod_{j=k}^{k+n-1} (1 - r_{j,k+n} r'_{j,k+n}). \tag{16}$$

Proof. Set $l = k + n$. From (11) and (13), taking in (6) the scalar product by $X(k)$ and using (8) yields $\|P_{k,l}\|^2 = \|P_{k,l-1}\|^2 - r_{k,l}\langle Q_{k+1,l}, P_{k,l-1}\rangle = \|P_{k,l-1}\|^2(1 - r_{k,l} r'_{k,l})$. And, by repeating n times the same operation, we obtain $\|P_{k,l}\|^2 = \|P_{k,k}\|^2 \prod_{j=k+1}^{l}(1 - r_{k,j} r'_{k,j}) = \|X(k)\|^2 \prod_{j=k+1}^{l}(1 - r_{k,j} r'_{k,j})$, which ends the proof of the first part of (16). The proof of the second part is similar, taking in the second part of (6) the scalar product by $X(l)$. □

We get in this way a factorization of inverse covariance matrices.

Let $\Gamma_{k,l}$ be the covariance matrix of the random vector $(X(k), X(k+1), \ldots, X(l))$ for $0 \leqslant k < l \leqslant N$. Let $\mathbf{q}_{i,j}$ and $\mathbf{p}_{i,j}$, $k \leqslant i < j \leqslant l$ be the $(l-k+1)\times 1$ column vectors of the coefficients of the residuals $Q_{i,j}$ and $P_{i,j}$ in the basis $(X(k), X(k+1), \ldots, X(l))$ of $H^X_{k,l}$. Define the following matrices in terms of their column vectors, $R^P_{k,l} = [\mathbf{p}_{k,l}, \mathbf{p}_{k+1,l}, \ldots, \mathbf{p}_{l,l}]$, and $R^Q_{k,l} = [\mathbf{q}_{k,k}, \mathbf{q}_{k,k+1}, \ldots, \mathbf{q}_{k,l}]$. The matrices $R^P_{k,l}$ are unit lower triangular and the matrices $R^Q_{k,l}$ are unit upper triangular. Therefore,

$$\operatorname{Det} R^Q_{k,l} = \operatorname{Det} R^P_{k,l} = 1. \tag{17}$$

Finally, define the diagonal matrices $D^p_{k,l}$ and $D^q_{k,l}$ as

$$D^P_{k,l} = (R^P_{k,l})^*(\Gamma_{k,l})(R^P_{k,l}) \quad \text{and} \quad D^Q_{k,l} = (R^Q_{k,l})^*(\Gamma_{k,l})(R^Q_{k,l}). \tag{18}$$

Note that $D^P_{k,l}(i,i) = \|P_{i,l}\|^2$ and $D^Q_{k,l}(i,i) = \|Q_{k,i}\|^2$, $k \leqslant i \leqslant l$. Since the process is basic, its covariance matrices $\Gamma_{k,l}$ are invertible. From (18), we obtain the following triangular factorizations for $\Gamma^{-1}_{k,l}$, $k < l$:

$$\Gamma^{-1}_{k,l} = (R^P_{k,l})(D^P_{k,l})^{-1}(R^P_{k,l})^* \quad \text{and} \quad \Gamma^{-1}_{k,l} = (R^Q_{k,l})(D^Q_{k,l})^{-1}(R^Q_{k,l})^*, \tag{19}$$

in which the triangular factors $R^P_{k,l}$ and $R^Q_{k,l}$ are computed recursively following diagonals.

This factorization has also been obtained by Delsarte et al. (1980) by a linear algebraic approach. It is used for solving the Yule–Walker equations, whose solution gives the representation of the auto-regressive process realizing the given partial covariance sequence. In the d-variate stationary case, the coefficients of the MAR(N) model are given by the first d columns of the matrix $R^P_{0,N}$. If $d = 1$, using the Trench–Zohar algorithm to compute $R^P_{0,N}$ leads to an order of $O(N^2)$ operations. If $d > 1$, using the nonstationary Trench–Zohar algorithm for the $PC(d)$ case leads to an order of $O(dN^2)$ operations instead of $O(N^3)$ in the general nonstationary case.

4. Positivity and extension

If an Hermitian matrix is positive definite, it is known to be the covariance matrix of some process. We get the following criterion of positivity from the factorizations above.

Proposition 4. *An $n \times n$ Hermitian matrix is positive definite if and only if its reflection coefficients satisfy for all (k, l) such that $0 \leqslant k < l \leqslant N$,*

$$0 \leqslant r_{k,l} r'_{k,l} < 1. \tag{20}$$

Proof. From (17) and (18), we get $\operatorname{Det} \Gamma_{k,l} = \prod_{k \leqslant i \leqslant l} \|P_{i,l}\|^2$ and $\operatorname{Det} \Gamma_{k,l} = \prod_{k \leqslant i \leqslant l} \|Q_{k,i}\|^2$. Consequently,

$$\operatorname{Det} \Gamma_{k,l} = \|P_{k,l}\|^2 \operatorname{Det} \Gamma_{k+1,l} \quad \text{and} \quad \operatorname{Det} \Gamma_{k,l} = \|Q_{k,l}\|^2 \operatorname{Det} \Gamma_{k,l-1}. \tag{21}$$

Replacing $\|P_{k,l}\|^2$ and $\|Q_{k,l}\|^2$ by their values given by (16) leads to the following determinant formula in terms of the generalized reflection coefficients,

$$\operatorname{Det} \Gamma_{k,l} = \prod_{k \leqslant i < j \leqslant l} \|X(i)\|^2 (1 - r_{i,j} r'_{i,j}). \tag{22}$$

This determinant is positive if and only if relation (20) is satisfied. □

Using the identity (9), relation (20) becomes $0 \leqslant |r_{k,l}|^2 < \|P_{k,l-1}\|^2 / \|Q_{k+1,l}\|^2$. In the stationary or PC(1) case, $\|P_{k,l-1}\| = \|Q_{k+1,l}\|$ and we recognize the classical relation.

The aller-retour formula yields a necessary and sufficient condition for the existence of an extension of a given partial d-variate covariance matrix. We need first the following proposition.

Proposition 5. *The reflection coefficients of the finite covariance matrices* $\operatorname{Cov}(Z(t), Z(t+1), \ldots, Z(t+n))$, *for $n \in \mathbb{N}$, of a basic d-variate stationary times series Z satisfy for $0 \leqslant k < l \leqslant n - md$,*

$$r_{k,l} = r_{k+md,l+md} \quad \text{and} \quad r'_{k,l} = r'_{k+md,l+md}. \tag{23}$$

Proof. Let Y be the PC(d) process defined by $Y(k + td) = Z_k(t)$ for $t \in \mathbb{Z}$, $1 \leqslant k \leqslant d$. We know that

$$\langle Y(s), Y(t) \rangle = \langle Y(s + d), Y(t + d) \rangle, \tag{24}$$

so the covariance matrices of $(Z(t), Z(t+1), \ldots, Z(t+n))$ and of $(Y(t+1), Y(t+2), \ldots, Y(t+nd))$ are equal and their reflection coefficients are the same. Let us prove the first part of (23). The proof of the second part is analogous. From (8) and (13), we get

$$r_{k,l} = \langle P_{k,l-1}, Y(l)\rangle / \|Q_{k+1,l}\|^2, \tag{25}$$

$$r_{k+md,l+md} = \langle P_{k+md,l+md-1}, Y(l+md)\rangle / \|Q_{k+md+1,l+md}\|^2. \tag{26}$$

Let us show that for all $0 \leqslant k < d$ and $m \in \mathbb{Z}$,

$$\|Q_{k,l}\|^2 = \|Q_{k+md,l+md}\|^2 \quad \text{and} \quad \langle P_{k,l-1}, Y(l)\rangle = \langle P_{k+md,l+md-1}, Y(l+md)\rangle. \tag{27}$$

Consequently, (25) is equal to (26) and (23) is true.

For proving the first part of (27), let us write $Q_{k,l} = \sum_{i=k}^{l} b_i Y(i)$, with $b_l = 1$. Then, from both relation (12) and property (24), we obtain $\|Q_{k,l}\|^2 = \sum_{i=k}^{l} b_i \langle Y(i), Y(l)\rangle = \sum_{j=k}^{l} b_i \langle Y(i+md), Y(l+md)\rangle = \|Q_{k+md,l+md}\|^2$.

For proving the second part of (27), let us write $P_{k,l} = \sum_{i=k}^{l} a_i Y(i)$, with $a_k = 1$. Let z^n denote the shift operator on H^Y defined by $z^n(Y(k)) = Y(k+n)$ for n and k in $\mathbb{Z}$. From the definition of $P_{k,l}$, we have $z^{md}(P_{k,l}) = z^{md}(Y(k) - \Pi_{k+1,l}[Y(k)]) = z^{md}(Y(k)) - z^{md}(\Pi_{k+1,l}[Y(k)])$. So

$$z^{md}(P_{k,l}) = P_{k+md,l+md} \tag{28}$$

is true if the following relation holds true:

$$z^{md}(\Pi_{k+1,l}[Y(k)]) = \Pi_{k+1+md,l+md}[Y(k+md)]. \tag{29}$$

Let us finish the proof of the second part of (27) first and prove (29) after.

From (28), we have $P_{k+md,l+md} = \sum_{i=k}^{l} a_i Y(i+md)$. Using relation (11) and property (24), we obtain

$$\begin{aligned}\langle P_{k+md,l+md-1}, Y(l+md)\rangle &= \sum_{i=k}^{l} a_i \langle Y(l+md), Y(k+md)\rangle \\ &= \sum_{i=k}^{l} a_i \langle Y(i), Y(l)\rangle = \langle P_{k,l-1}, Y(l)\rangle,\end{aligned}$$

hence the second part of (27).

It remains to prove (29). But $\Pi_{k+1,l}[Y(k)] = \sum_{i=k+1}^{l} \alpha_i Y(i)$, where $(\alpha_i)_{i=k+1}^{l}$ for $k+1 \leqslant s \leqslant l$ satisfy $\langle [Y(k) - \sum_{i=k+1}^{l} \alpha_i Y(j)], Y(s)\rangle = 0$. From property (24), we get

$$\begin{aligned}&\langle Y(k+md), Y(s+md)\rangle - \left\langle \sum_{i=k+1}^{l} \alpha_i Y(i+md), Y(s+md) \right\rangle \\ &\quad = \left\langle \left(Y(k+md) - z^{md}\left[\sum_{i=k+1}^{l} \alpha_i Y(i) \right] \right), Y(s+md) \right\rangle = 0.\end{aligned}$$

Hence, from definition of projection, we have $z^{md}[\sum_{i=k+1}^{l} \alpha_i Y(i)] = \Pi_{k+1+md,l+md}[Y(k+md)]$ and so (29) is demonstrated and the result follows. □

Now, we can state the theorem whose proof constitutes a method of construction of extensions which will be illustrated by a numerical example in Section 6.

Theorem 2. *A necessary and sufficient condition for extending a given partial d-variate covariance matrix C_N is its positive definiteness.*

Proof. The necessary part is obvious. Conversely, if C_N is the covariance matrix of a d-variate random vector $(Z(t), Z(t+1), \ldots, Z(t+N))$, then from Proposition 5, its reflection coefficients satisfy

$$r(k,l) = r(k+dm, l+dm), \quad 0 \leqslant k < l \leqslant nd, \ m \in \mathbb{Z}, \text{ with } l + md \leqslant Nd. \tag{30}$$

The associated triangular arrays of reflection coefficients $(r_{k,l})_{0\leqslant k<l\leqslant n}$ and $(r'_{k,l})_{0\leqslant k<l\leqslant n}$ can be extended in the following way: if $(i,j) \in \mathbb{Z}^2$ is such that $(i,j) = (k+dm, l+dm)$ for some $m \in \mathbb{Z}$ and $0 \leqslant k < l \leqslant n - md$, set $r_{i,j} = r_{k,l}$; elsewhere set $r_{i,j} = 0$. And $r'_{i,j}$ is given by (9). The aller-retour formula (15) gives the covariance kernel value $K(i,j)$ for $i \neq j$. And for $i = j$, it is sufficient to set $K(i,i) = C_{k,k}(0)$ if $i = k + dm$ for some $m \in \mathbb{Z}$ and $0 \leqslant k < d$. □

From (22), we can see that the kernel obtained by this construction maximizes the determinants of the extended covariance matrices. In the next section we will see that it corresponds to the auto-regressive model maximizing the Burg entropy.

5. Burg and *n*-step entropies

Let us determine first the spectral density of the process with maximum of Burg entropy among processes sharing the same partial d-variate covariance sequence. Let $\mathscr{P}^{nm}_{d\times d}$ denote the set of $d \times d$ trigonometric polynomials P with degrees between $n \in \mathbb{Z}$ and $m \in \bar{\mathbb{N}}$, i.e., such that $P(\mathrm{e}^{\mathrm{i}\lambda}) = \sum_{k=n}^{m} A(k)\mathrm{e}^{\mathrm{i}k\lambda}$.

Theorem 3. *Let C be a partial covariance sequence. If $H^* : \Pi \to \mathscr{M}^+_{d\times d}$ is such that*

$$H^*(\lambda) = [I + P_m(\lambda)]^{-1} \Sigma ([I + P_m(\lambda)]^{-1})^*$$

is positive definite for all λ, with $\Sigma \in \mathscr{M}^+_{d\times d}$ and $P_m \in \mathscr{P}^{1N}_{d\times d}$ such that

$$\det \Sigma = \|I + P_m\|_C^2 = \inf_{\mathscr{P}^{1N}_{d\times d}} \int \mathrm{tr}[(I + P(\lambda))^*(I + P(\lambda)F_C(\lambda))]\,\mathrm{d}\lambda,$$

where $F_C(\lambda) = \sum_{n=-N}^{N} C(n)\mathrm{e}^{\mathrm{i}n\lambda}$, then $\mathscr{I}[H^*] = \max_{H \text{ s. t. } (2)} \mathscr{I}[H]$.

Proof. Let $H:\Pi \to \mathscr{M}^{+}_{d\times d}$ be element of $L^2(\Pi,\mathscr{B}(\Pi),\lambda)$.

We have $\exp \mathscr{I}[H] = \exp \int_\Pi \ln \operatorname{Det} H(\lambda)\,\mathrm{d}\lambda = \exp \int_\Pi \operatorname{Tr} \ln H(\lambda)\,\mathrm{d}\lambda$, and Theorem 8 of Helson and Lowdenslager (1958) states that

$$\exp \int_\Pi \operatorname{Tr} \ln H(\lambda)\,\mathrm{d}\lambda = \inf \int_\Pi \operatorname{Tr}\left[(J+P(\lambda))^*(J+P(\lambda))H(\lambda)\right]\mathrm{d}\lambda,$$

where the infimum is taken over the matrices $J \in \mathscr{M}^{+}_{d\times d}$ such that $\operatorname{Det} J = 1$ and $P \in \mathscr{P}^{1+\infty}_{d\times d}$.

Now, H may be written $H = F + G$, with $F \in \mathscr{P}^{-NN}_{d\times d}$ and $G \in L^2$ orthogonal to $\mathscr{P}^{-NN}_{d\times d}$ for the scalar product $\langle A,B\rangle = \int_\Pi \operatorname{Tr} A(\lambda)^* B(\lambda)\,\mathrm{d}\lambda$. If H satisfies (2), then $F = F_C$. If $P \in \mathscr{P}^{1N}_{d\times d}$, then $(I+P)^*(I+P) \in \mathscr{P}^{0N}_{d\times d}$, hence

$$\begin{aligned}\int_\Pi \operatorname{Tr}\left[(I+P(\lambda))^*(I+P(\lambda))H(\lambda)\right]\mathrm{d}\lambda &= \langle (I+P)^*(I+P), H\rangle \\ &= \langle (I+P)^*(I+P), F_C\rangle = \|I+P\|_C^2,\end{aligned}$$

so that $\exp \mathscr{I}[H] \leqslant \inf_{\mathscr{P}^{1N}_{d\times d}} \|I+P\|_C^2$, which we will now show to be exactly $\exp \mathscr{I}[H^*]$.

Indeed, $\ln \operatorname{Det} H = \ln \operatorname{Det} \Sigma - \ln \operatorname{Det}(I+P)(I+P)^*$, and we may write $\operatorname{Det}(I+P)(I+P)^* = 1+Q$ where Q is a trigonometric polynomial with no constant term. Thus, by Theorem 3 of Helson and Lowdenslager (1958), we know that $\int_\Pi \ln(1+Q(\lambda))\,\mathrm{d}\lambda = \ln \int (1+Q(\lambda))\,\mathrm{d}\lambda = 0$. Hence $\int_\Pi \ln \operatorname{Det} H(\lambda)\,\mathrm{d}\lambda = \ln \operatorname{Det} \Sigma = \ln \|I+P_m\|_C^2$, and the result is proven. □

The function H^* appears to be the spectral density of a MAR(N) (or PAR) process as given in (4), which means that the maximum of Burg entropy is obtained for the spectral density of a MAR process, or, if PC processes are considered, for the spectral density of a PAR process. Note that, since ln Det is a strictly concave function on the set of Hermitian positive definite matrices, H^* is unique.

Let us now establish a link between n-step entropy and Burg entropy. The n-step entropy of Gaussian MS processes is known (see Choi, 1993) to be equal to

$$\frac{nd}{2}\ln(2\pi \mathrm{e}) + \frac{1}{2}\ln \operatorname{Det} \Gamma_n. \tag{31}$$

Since the n-step entropy of any process is less than the n-step entropy of the Gaussian process with the same partial covariance matrix, its maximization is equivalent to the maximization of the determinants of covariance matrices of size n among those sharing the same first block of size N, namely C_N. The determinant formula (22) allows us to establish a criteria for the maximization of the determinant in terms of the reflection coefficients.

Proposition 6. *Let $\{(r_{k,l}, r'_{k,l})\}_{0\leqslant k<l\leqslant N}$ be the reflection coefficients of a given partial covariance sequence C.*

The n-step maximum entropy of a PC(d) *process* (*or d-variate stationary process*) *realizing C is obtained for a process with the reflection coefficients $r_{i,j} = r_{k,l}$ if $(i,j) = (k+dm, l+dm)$ for some $m \in \mathbb{Z}$ and $0 \leqslant k < l \leqslant N$, and $r_{i,j} = 0$ elsewhere for $0 \leqslant i < j \leqslant n$.*

Proof. The coefficients $r_{i,j}$ defined above have the necessary and sufficient property (30) for being the reflection coefficients of a PC(d) process Y. The determinant of the partial covariance matrix Γ_n of Y is given by (22). We extend thus the given reflection coefficients in order to have the maximum of null coefficients while preserving the PC(d) structure. Hence the associated process maximizes the determinant of the matrices with the same first $N \times N$ block C_N, and so Y maximizes the n-step entropy among processes realizing the given partial covariance sequence C. The result for MS processes is obtained by duality. □

It is well-known that for scalar stationary processes $(\ln \operatorname{Det} \Gamma_n)/n$ converges to the Burg entropy of the process. In order to state the same result for the d-variate case, we need the following Szegö type limit theorem. Observe that in the d-variate case Γ_n is defined only if $n = Nd$ for some integer N.

Theorem 4. *Let $(\Gamma_{dN})_{N\in\mathbb{N}}$ be the covariance matrix sequence of a given d-variate process Z. Then*

$$\det \Gamma_{dN}/\det \Gamma_{d(N-1)} \to \exp \int_{\Pi} \ln \det H(\lambda)\, \mathrm{d}\lambda, \quad N \to +\infty.$$

Proof. Let σ_N^2 be the variance of the error of the best linear prediction of $Z(t)$ knowing the finite past $Z(t-1),\ldots,Z(t-N)$. By projection properties, $\sigma_N^2 \to \sigma^2$, where σ^2 is the variance of the innovation of Z, that is the error of the best linear prediction of $Z(t)$ knowing the infinite past. Helson and Lowdenslager (1958) proved that $\sigma^2 = \exp \int_{\Pi} \ln \operatorname{Det} H(\lambda)\, \mathrm{d}\lambda$. Hence, it is sufficient to show that $\sigma_N^2 = \operatorname{Det} \Gamma_{dN}/\operatorname{Det} \Gamma_{d(N-1)}$.

The matrix $\Gamma_{dN} = \Gamma_{0,dN}$ may be seen as the partial covariance matrix of a PC(d) process. Thus, by the first part of the factorization formula (19), its inverse may be written as $\Gamma_{0,dN}^{-1} = (R_{0,dN}^P)(D_{0,dN}^P)^{-1}(R_{0,dN}^P)^*$. This factorization is unique and $\operatorname{Det} \Gamma_{d-1,dN} = \operatorname{Det} \Gamma_{0,d(N-1)}$ by stationarity. Seeing the matrix as the covariance matrix of a d-variate stationary process, we have by the second part of (21),

$$\begin{aligned}\sigma_N^2 &= \|P_{0,dN}\|^2 \|P_{1,dN}\|^2 \cdots \|P_{d-1,dN}\|^2 \\ &= \frac{\operatorname{Det} \Gamma_{0,dN}}{\operatorname{Det} \Gamma_{1,dN}} \frac{\operatorname{Det} \Gamma_{1,dN}}{\operatorname{Det} \Gamma_{2,dN}} \cdots \frac{\operatorname{Det} \Gamma_{d-2,dN}}{\operatorname{Det} \Gamma_{d-1,dN}} = \frac{\operatorname{Det} \Gamma_{0,dN}}{\operatorname{Det} \Gamma_{d-1,dN}} = \frac{\operatorname{Det} \Gamma_{0,dN}}{\operatorname{Det} \Gamma_{0,d(N-1)}}.\end{aligned}$$

And the result follows. □

Corollary 1. *Up to an additive constant, the n-step entropy of a Gaussian* MS *process divided by n converges to its Burg entropy.*

Proof. If Z is Gaussian, its n-step entropy can be written as in (31). And applying Cesáro's lemma to the sequence $(\ln[\operatorname{Det} \Gamma_{d(N+1)}/\operatorname{Det} \Gamma_{dN}])_{N\in\mathbb{N}}$ yields the result. □

Note that another quantity converges to the Burg entropy too. Indeed, by the scalar case, considering the components of Z as scalar time series, we know that $1/n \ln \Gamma_n(k,k) \to \int_\Pi \ln h_{kk}(\lambda)\,\mathrm{d}\lambda$, where h_{kk} is the auto-spectral density of Z_k, for $k = 1,\ldots,d$. Since $H(\lambda)$ is positive definite for any $\lambda \in \Pi$, we know that $\int_\Pi \ln \mathrm{Det}\, H(\lambda)\,\mathrm{d}\lambda = \sum_{k=1}^{d} \int_\Pi \ln h_{kk}(\lambda)\,\mathrm{d}\lambda$, which means that

$$\sum_{k=1}^{d} \frac{1}{n} \ln \Gamma_n(k,k) \to \int_\Pi \ln \mathrm{Det}\, H(\lambda)\,\mathrm{d}\lambda.$$

6. Numerical illustration

For mere illustration purpose, let us take the following partial covariance matrix of order 2 of a PC(2) process.

$$C_2 = \begin{pmatrix} 1 & 0.5 & 0.3 & 0.1 & 0.07 & 0.003 \\ 0.5 & 1 & 0.25 & 0.45 & 0.3 & 0.002 \\ 0.3 & 0.25 & 1 & 0.5 & 0.3 & 0.1 \\ 0.1 & 0.45 & 0.5 & 1 & 0.25 & 0.45 \\ 0.07 & 0.3 & 0.3 & 0.25 & 1 & 0.5 \\ 0.003 & 0.002 & 0.1 & 0.45 & 0.5 & 1 \end{pmatrix}.$$

Its reflection coefficient arrays are

$$R_{0,2} = \begin{pmatrix} 0.5 & 0.1867 & -0.3096 & -0.1347 & 0.3586 \\ & 0.25 & 0.4333 & 0.2026 & -0.4815 \\ & & 0.5 & 0.1867 & -0.3096 \\ & & & 0.25 & 0.4333 \\ & & & & 0.5 \end{pmatrix} \quad \text{and}$$

$$R'_{0,2} = \begin{pmatrix} 0.5 & 0.2333 & -0.2751 & -0.1755 & 0.2683 \\ & 0.25 & 0.3467 & 0.228 & -0.3695 \\ & & 0.5 & 0.2333 & -0.2751 \\ & & & 0.25 & 03467 \\ & & & & 0.5 \end{pmatrix},$$

where the null coefficients do not appear.

We obtain the following completions of the reflection coefficients arrays of order 4 by the method described in the proof of Theorem 2,

$$R_{0,4} = \begin{pmatrix} 0.5 & 0.1867 & -0.3096 & -0.1347 & 0.3586 & & & & \\ & 0.25 & 0.4333 & 0.2026 & -0.4815 & & & & \\ & & 0.5 & 0.1867 & -0.3096 & -0.1347 & 0.3586 & & \\ & & & 0.25 & 0.4333 & 0.2026 & -0.4815 & & \\ & & & & 0.5 & 0.1867 & -0.3096 & -0.1347 & 0.3586 \\ & & & & & 0.25 & 0.4333 & 0.2026 & -0.4815 \\ & & & & & & 0.5 & 0.1867 & -0.3096 \\ & & & & & & & 0.25 & 0.4333 \\ & & & & & & & & 0.5 \end{pmatrix}$$

and

$$R'_{0,4} = \begin{pmatrix} 0.5 & 0.2333 & -0.2751 & -0.1755 & 0.2683 & & & & \\ & 0.25 & 0.3467 & 0.228 & -0.3695 & & & & \\ & & 0.5 & 0.2333 & -0.2751 & -0.1755 & 0.2683 & & \\ & & & 0.25 & 0.3467 & 0.228 & -0.3695 & & \\ & & & & 0.5 & 0.2333 & -0.2751 & -0.1755 & 0.2683 \\ & & & & & 0.25 & 0.3467 & 0.228 & -0.3695 \\ & & & & & & 0.5 & 0.2333 & -0.2751 \\ & & & & & & & 0.25 & 0.3467 \\ & & & & & & & & 0.5 \end{pmatrix}.$$

From these matrices, we get the extension of the covariance matrix by the aller-retour formula of Proposition 2,

$$\Gamma_{0,4} = \begin{pmatrix} 1 & 0.5 & 0.3 & 0.1 & 0.07 & 0.003 & -0.0004 & 0.0035 & -0.0115 & 0.0132 \\ 0.5 & 1 & 0.25 & 0.45 & 0.3 & 0.002 & 0.1873 & -0.1732 & 0.004 & 0.1099 \\ 0.3 & 0.25 & 1 & 0.5 & 0.3 & 0.1 & 0.07 & 0.003 & -0.0004 & 0.0035 \\ 0.1 & 0.45 & 0.5 & 1 & 0.25 & 0.45 & 0.3 & 0.002 & 0.1873 & 0.1732 \\ 0.07 & 0.3 & 0.3 & 0.25 & 1 & 0.5 & 0.3 & 0.1 & 0.07 & 0.003 \\ 0.003 & 0.002 & 0.1 & 0.45 & 0.5 & 1 & 0.25 & 0.45 & 0.3 & 0.002 \\ -0.004 & 0.1873 & 0.07 & 0.3 & 0.3 & 0.25 & 1 & 0.5 & 0.3 & 0.1 \\ 0.0035 & -0.1732 & 0.003 & 0.002 & 0.1 & 0.45 & 0.5 & 1 & 0.25 & 0.45 \\ 0.0115 & 0.004 & -0.0004 & 0.1873 & 0.07 & 0.3 & 0.3 & 0.25 & 1 & 0.5 \\ 0.0132 & -0.1099 & 0.0035 & -0.1732 & 0.003 & 0.002 & 0.1 & 0.45 & 0.5 & 1 \end{pmatrix}.$$

The first factorization in (19) takes here the form $\Gamma_{0,2}^{-1} = R_{0,2}^P (D_{0,2}^P)^{-1} (R_{0,2}^P)^*$, where

$$R_{0,2}^P = \begin{pmatrix} 1 & & & & & \\ 0.7239 & 1 & & & & \\ -0.4455 & 0.1464 & 1 & & & \\ 0.5279 & -0.6333 & -0.5607 & 1 & & \\ 0.3282 & -0.4264 & -0.3146 & -0.0333 & 1 & \\ -0.3586 & 0.4815 & 0.3096 & -0.4333 & -0.5 & 1 \end{pmatrix}$$

and $D_{0,2}^P$ is the diagonal matrix $(0.5791, 0.6247, 0.6562, 0.7967, 0.7500, 1)$.

The coefficients of the two first columns of $R_{0,2}^P$ are the coefficients of the MAR(2) process Z or of the PAR(2) process Y which maximizes the entropy among the processes with the same covariance matrix C_2, and $\Gamma_{0,4}$ is its covariance matrix of order 4. Representation (3) for Z takes here the form

$$\begin{pmatrix} 1 & 0 \\ 0.7239 & 1 \end{pmatrix} \begin{pmatrix} Z_0(t) \\ Z_1(t) \end{pmatrix} + \begin{pmatrix} -0.4455 & 0.1464 \\ 0.5279 & -0.6333 \end{pmatrix} \begin{pmatrix} Z_0(t-1) \\ Z_1(t-1) \end{pmatrix} + \begin{pmatrix} 0.3282 & -0.4264 \\ -0.3586 & 0.4815 \end{pmatrix} \begin{pmatrix} Z_0(t-2) \\ Z_1(t-2) \end{pmatrix} = \begin{pmatrix} \varepsilon_0(t) \\ \varepsilon_1(t) \end{pmatrix}$$

and representation (5) for Y is

$$Y(2t) + 0.1464Y(2t-1) - 0.4455Y(2t-2) - 0.4264Y(2t-3) + 0.3282Y(2t-4) = \varepsilon(2t),$$

$$Y(2t+1) + 0.7239Y(2t) - 0.6333Y(2t-1) + 0.5279Y(2t-2) + 0.4815Y(2t-3) - 0.3586Y(2t-4) = \varepsilon(2t+1).$$

References

Brockwell, P., Davis, D., 1996. Introduction to Time Series and Forecasting, Springer Texts in Statistics. Springer, New York.

Castro, G., 1997. Coefficients de réflexion généralisés, Extension de covariances mutidimensionnelles et autres applications. Thèse Univ. Paris-Sud, France.

Choi, B.-S., 1987. A proof of Burg's theorem, in: Maximum Entropy and Bayesian Spectral Analysis and Estimation Problems, Proc. 3rd Workshop, Laramie/Wyo., Eds C.R. Smith and G.J. Erickson, Fundam. Theor. Phys., 21, 75–84.

Choi, B.-S., 1993. Multivariate maximum entropy spectrum. J. Multivariate Anal. 46, 56–60.

Choi, B.-S., 1998. An information-theoretic derivation of the 2-D maximum entropy spectrum. IEEE Signal Proc. Lett. 5, 271–272.

Choi, B.-S., Cover, T., 1984. An information-theoretic proof of Burg's maximum entropy spectrum. Proc. IEEE 72, 1094–1095.

Constantinescu, T., 1996. Schur Parameters, Factorization and Dilation Problems. Operator Theory: Advances and Applications, Vol. 82. Birkhauser, Basel.

Dégerine, S., 1994. Sample partial autocorrelation function of a multivariate time series. J. Multivariate Anal. 50, 294–313.

Delsarte, P., Genin, Y., Kamp, Y., 1979. Schur parametrization of positive definite block-Toeplitz systems. SIAM J. Appl. Math. 36, 34–45.

Delsarte, P., Genin, Y., Kamp, Y., 1980. A method of matrix inverse triangular decomposition based on contiguous principal submatrices. Linear Algebra Appl. 31, 199–212.

Franses, P.H., 1996. Periodicity and Stochastic Trends in Economic Time Series, Advanced Texts in Econometrics. Oxford University Press, Oxford.

Gladyshev, E., 1961. Periodically correlated random sequences. Sov. Math. Dokl. 2, 385–388.

Gohberg, I., Kraashoek, M., Woerdeman, H., 1989. The band method for positive and contractive extension problems. Integral Equations Oper. Theory 22, 109–155.

Helson, H., Lowdenslager, D., 1958. Prediction theory and Fourier series in several variables. Acta Math. 99, 165–212.

Horn, R., Johnson, C., 1985. Matrix Analysis. Cambridge University Press, Cambridge.

McLeod, A., 1993. Parsimony, model adequacy and periodic correlation in time series forecasting. Int. Statist. Rev. 61, 387–393.

Monin, A., 1963. Stationary and periodic time series in the general circulation of the atmosphere. In: Rosenblatt (Ed.), Proceedings of the Symposium on Time Series Analysis. Wiley, New York, pp. 144–151.

Morf, M., Vieira, A., Kailath, T., 1978. Covariance characterization by partial autocorrelation matrices. Ann. Statist. 6, 643–648.

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

Pagano, M., Parzen, E., 1979. An approach to modeling seasonally stationary time series. J. Econom. 9, 137–153.

Priestley, M., 1981. Spectral Analysis and Time Series, Vol. 2, Academic Press, New York.

Rozanov, Y., 1967. Stationary Random Processes. Holden-Day, San Francisco.

Seghier, A., 1991. Extension de fonctions de type positif et entropie associée. Cas multidimensionnel. Ann. Inst. H. Poincaré 8, 651–675.

Tiao, G., Grupe, M., 1980. Hidden periodic auto-regressive-moving average models in time series data. Biometrika 67, 365–373.

Troutman, B., 1979. Some results in periodic autoregression. Biometrika 66, 219–228.

Zhang, H., 1997. Maximum entropy modeling of periodically correlated processes. IEEE Trans. Inform. Theory 43, 2033–2035.