The Annals of Statistics
1983, Vol. 11, No. 1, 337–340

COVARIANCE MATRICES CHARACTERIZATION BY A SET OF SCALAR PARTIAL AUTOCORRELATION COEFFICIENTS

BY HIDEAKI SAKAI[1]

Kyoto University

It has been shown that the autocovariance matrices of a stationary multivariate time series can be uniquely characterized by a sequence of the normalized partial autocorrelation matrices having singular values less than one.

In this note, we show that the same autocovariance matrices can be also uniquely characterized by a set of sequences of *scalar* partial autocorrelation coefficients whose magnitudes are all less than one.

1. Introduction and summary. It has been shown by Morf, Vieira and Kailath (1978) that the autocovariance matrices of a stationary multivariate time series can be uniquely characterized by a sequence of the normalized partial autocorrelation (PARCOR) matrices, having singular values less than one. This is a nice generalization of the scalar case (Barnoff-Nielsen and Schou, 1973; Burg, 1975; Ramsey, 1974) but it is not an easy task to parametrize the PARCOR matrices satisfying the above constraint.

In this note, using the recent result of Sakai (1982) about circular lattice filtering based on the work of Pagano (1978), we show that the same autocovariance matrices of a stationary d-variates time series can also be characterized uniquely by d sequences of the *scalar* normalized PARCOR coefficients whose magnitudes are all less than one. This may be a more convenient generalization, since now the parametrization becomes quite easy.

2. The circular lattice filtering. Here we give a review of Sakai (1982) for later discussion. Let $\{\mathbf{X}(t)\}$ be a zero-mean real d-variates stationary time series and the scalar process $\{Y(t)\}$ be generated from $\{\mathbf{X}(t)\}$ by

$$Y(j + d(t - 1)) = X_j(t), \tag{1}$$

where $X_j(t)$ is the jth element of $\mathbf{X}(t)$. Then $\{Y(t)\}$ becomes a periodically correlated stationary process of period d (Pagano, 1978).

We denote the autocovariance matrices of $\mathbf{X}(t)$, and the covariances of $Y(t)$ by

$$R_k = E\{X(t)X^T(t - k)\}, \qquad R_{-k} = R_k^T, \tag{2}$$

$$R(s, t) = E\{Y(s)Y(t)\}, \qquad R(t, s) = R(s, t), \tag{3}$$

respectively where "T" denotes the transpose operation. Define the jth order kth channel forward and backward linear prediction errors for $Y(t)$ by

$$\varepsilon(j, k + nd) = Y(k + nd) + \textstyle\sum_{i=1}^{j} \alpha_k(j, i)Y(k + nd - i) \tag{4}$$

$$\eta(j, k + nd) = Y(k + nd - j) + \textstyle\sum_{i=1}^{j} \beta_k(j, j + 1 - i)Y(k + nd - i + 1), \tag{5}$$

respectively. The predictor coefficients $\alpha_k(j, i)$, $\beta_k(j, i)(i = 1, \cdots, j)$ are determined by minimizing $E\{\varepsilon^2(j, k + nd)\}$, $E\{\eta^2(j, k + nd)\}$ with respect to $\alpha_k(j, i)$, $\beta_k(j, i)$, respectively. That is,

---

Received March 1982; revised July 1982.

[1] Department of Applied Mathematics and Physics, Faculty of Engineering, Kyoto University, Kyoto, 606 Japan.

AMS 1970 *subject classification.* Primary 62M10, 62N15, 62M15, 60G10.

Key words and phrases. Partial autocorrelation coefficients, multivariate stationary processes, circular lattice filtering.

$$\mathbf{R}_k(j)\mathbf{a}_k(j) = (\sigma_k^2(j), 0, \cdots, 0)^T \tag{6}$$

$$\mathbf{R}_k(j)\mathbf{b}_k(j) = (0, \cdots, 0, \tau_k^2(j))^T \tag{7}$$

follow where $\mathbf{a}_k^T(j) = (1, \alpha_k(j, 1), \cdots, \alpha_k(j, j))$, $\mathbf{b}_k^T(j) = (\beta_k(j, j), \cdots, \beta_k(j, 1), 1)$, and the (p, q)-th element of $\mathbf{R}_k(j)$ is $R(k - p + 1, k - q + 1)$, $(1 \le p, q \le j + 1)$. Note that $\mathbf{R}_0(j) = \mathbf{R}_d(j)$.

Then, we have the following efficient algorithm for successively obtaining $\mathbf{a}_k(j)$, $\mathbf{b}_k(j)$, $(j = 0, 1, \cdots)$ (Sakai, 1982).

THEOREM 1. (*A Levinson-Type Circular Recursive Algorithm*) (a) *Initial conditions* $(j = 0)$

$$\sigma_k^2(0) = \tau_k^2(0) = R(k, k), \qquad \Delta_k(0) = R(k, k - 1), \qquad k = 1, \cdots, d. \tag{8}$$

(b) *Order update from* j *to* $j + 1$

(i) *compute*

$$\Delta_k(j) = \sum_{m=0}^{j} R(k - m, k - j - 1)\alpha_k(j, m) \tag{9}$$

$$= \sum_{m=0}^{j} R(k - j - 1 + m, k)\beta_{k-1}(j, m) \tag{10}$$

$$\alpha_k(j + 1, j + 1) = -\Delta_k(j)/\tau_{k-1}^2(j) \tag{11}$$

$$\beta_k(j + 1, j + 1) = -\Delta_k(j)/\sigma_k^2(j) \tag{12}$$

(ii) *update*

$$\alpha_k(j + 1, i) = \alpha_k(j, i) + \alpha_k(j + 1, j + 1)\beta_{k-1}(j, j + 1 - i), \qquad i = 1, \cdots, j, \tag{13}$$

$$\beta_k(j + 1, i) = \beta_{k-1}(j, i) + \beta_k(j + 1, j + 1)\alpha_k(j, j + 1 - i), \qquad i = 1, \cdots, j, \tag{14}$$

$$\sigma_k^2(j + 1) = \sigma_k^2(j)\{1 - \alpha_k(j + 1, j + 1)\beta_k(j + 1, j + 1)\}, \tag{15}$$

$$\tau_k^2(j + 1) = \tau_{k-1}^2(j)\{1 - \alpha_k(j + 1, j + 1)\beta_k(j + 1, j + 1)\} \tag{16}$$

where the subscript $k - 1 = 0$ *is replaced by* d.

We also note that the third condition in (8) must be added to the original version of this result (Sakai, 1982). It is shown there that the stationarity of $\mathbf{X}(t)$ is equivalent to the condition

$$0 \le \alpha_k(j + 1, j + 1)\beta_k(j + 1, j + 1) < 1. \tag{17}$$

As in Morf, Vieira, and Kailath (1978), if we define the normalized PARCOR coefficients by

$$\rho_k(j + 1) = -\frac{\Delta_k(j)}{\sigma_k(j)\tau_{k-1}(j)}, \tag{18}$$

then from (11), (12), and (17), we have $|\rho_k(j + 1)| < 1$. The statistical property of the estimated $\rho_k(j + 1)$ is derived in Sakai (1982) under the assumption that $\rho_k(j + 1) = 0$ for $j \ge p_k$, that is, $\{Y(t)\}$ is a pure periodic autoregressive process.

3. Covariance characterization. We now present the main result of this note.

THEOREM 2. *There is a one-to-one correspondence between a sequence of the autocovariance matrices* $\{R_0, R_1, \cdots, R_N, \cdots\}$ *of d-variate time series and d sequences of* $\{R(k, k), \rho_k(j), (k = 1, \cdots, d; j = 1, 2, \cdots)\}$ *satisfying the condition*

$$R(k, k) > 0, \qquad |\rho_k(j + 1)| < 1. \tag{19}$$

Note that the corresponding constraint in Morf, Vieira, and Kailath (1978) is that R_0 is positive definite and that the normalized PARCOR matrices have singular values less than one.

For proof, we note first from (1) that R_k's are expressed in terms of $R(s, t)$'s by

$$R_0 = \begin{bmatrix} R(1,1) & R(1,2) & \cdots & R(1,d) \\ R(2,1) & R(2,2) & \cdots & R(2,d) \\ \vdots & \vdots & & \vdots \\ R(d,1) & R(d,2) & \cdots & R(d,d) \end{bmatrix},$$

(20)

$$R_1 = \begin{bmatrix} R(1,1-d) & R(1,2-d) & \cdots & R(1,0) \\ R(2,1-d) & R(2,2-d) & \cdots & R(2,0) \\ \vdots & \vdots & & \vdots \\ R(d,1-d) & R(d,2-d) & \cdots & R(d,0) \end{bmatrix}, \cdots.$$

Thus, given a truncated sequence $\{R_0, R_1, \cdots, R_N\}$ of the autocovariance matrices, the algorithm of Theorem 1 yields d sequences of $R(k, k)$, $\rho_k(j)$, $(k = 1, \cdots, d; j = 1, \cdots, p_k)$ satisfying (19) where $p_k = Nd + k - 1$, since from (20) we have $R(k, k) > 0$, $R(k, k-1)$ and can start the algorithm by (8), and from (20) we see that the largest order that can be defined for the kth channel must satisfy $k - j = 1 - Nd$.

Defining $\tilde{\alpha}_k(j+1, i) = \alpha_k(j+1, i)/\sigma_k(j+1)$, $\tilde{\beta}_k(j+1, i) = \beta_k(j+1, i)/\tau_k(j+1)$, $(i = 1, \cdots, j+1)$ and noting from (15), (17), and (18) the equalities

$$\sigma_k(j+1)/\tau_k(j+1) = \sigma_k(j)/\tau_{k-1}(j) \tag{21}$$

$$\sigma_k(j+1)/\sigma_k(j) = \tau_k(j+1)/\tau_{k-1}(j) = \sqrt{1 - \rho_k^2(j+1)}, \tag{22}$$

we obtain a normalized Levinson-type algorithm as

$$\begin{aligned} -\sigma_k(j)\tau_{k-1}(j)\rho_k(j+1) \\ &= R(k, k-j-1) + \sigma_k(j) \textstyle\sum_{m=1}^{j} R(k-m, k-j-1)\tilde{\alpha}_k(j, m) \\ &= R(k-j-1, k) + \tau_{k-1}(j) \textstyle\sum_{m=1}^{j} R(k-j-1+m, k)\tilde{\beta}_{k-1}(j, m) \end{aligned} \tag{23}$$

$$\tilde{\alpha}_k(j+1, j+1) = \rho_k(j+1)/\tau_k(j+1) \tag{25}$$

$$\tilde{\beta}_k(j+1, j+1) = \rho_k(j+1)/\sigma_k(j+1) \tag{26}$$

where for $i = 1, \cdots, j$

$$\tilde{\alpha}_k(j+1, i) = \{\tilde{\alpha}_k(j, i) + \rho_k(j+1)\tilde{\beta}_{k-1}(j, j+1-i)\}/\sqrt{1 - \rho_k^2(j+1)} \tag{27}$$

$$\tilde{\beta}_k(j+1, i) = \{\tilde{\beta}_{k-1}(j, i) + \rho_k(j+1)\tilde{\alpha}_k(j, j+1-i)\}/\sqrt{1 - \rho_k^2(j+1)} \quad i = 1, \cdots, j \tag{28}$$

with the initial conditions

$$\sigma_k^2(0) = \tau_k^2(0) = R(k, k), \; -\sigma_k(0)\tau_{k-1}(0)\rho_k(1) = R(k, k-1). \tag{9'}$$

Conversely, given d sequences of $\{R(k, k), \rho_k(j), (k = 1, \cdots, d; j = 1, \cdots, p_k)\}$ satisfying (19), we can generate $R(k, k-j)(j = 1, \cdots, p_k)$ in the following way. First, use (9′) to obtain $R(k, k-1)$, and use (22), and (25)–(28) successively to compute $\tilde{\alpha}_k(j, i)$, $\tilde{\beta}_k(j, i)$, $(i = 1, \cdots, j)$ from $\rho_k(1), \cdots, \rho_k(j)$. Then, from (23) or (24), we obtain $R(k, k - j - 1)$ by using previous $R(k, k-i)$, $\tilde{\alpha}_k(j, i)$ or $\tilde{\beta}_{k-1}(j, i)$, $(i = 1, \cdots, j)$, and $\rho_k(j+1)$. We feel that the use of (24) is more appropriate, since it consists of kth channel $R(k, k-i)$, $(k-1)$th channel $\tilde{\beta}_{k-1}(j, i)$ and $\tau_{k-1}(j)$ while the use of (23) requires kth to $(k-j)$th channel covariances which ultimately spread to whole channels, showing the inappropriateness to parallel processing. Anyway, we can obtain the kth rows of $R_0, R_1, \cdots, R_N$ each from right to left, except R_0 for which only the lower triangular elements are required, where N is an integer satisfying $\max_k(p_k - k + 1)/d \le N < \max_k(p_k - k + 1)/d + 1$ and we extend $R(k, k-j)$ for $p_k < j \leq Nd + k - 1$ by putting $\rho_k(j) = 0$. This completes the proof of Theorem 2.

A reviewer has pointed out that the result in this note is implicit in Delosme and Morf (1980) and in Lev-Ari and Kailath (1981) which treat the covariance characterization problem of general nonstationary processes. However, it seems to the author that further argument is needed to deduce the present result from the above two papers. It is also stressed by the reviewer that the covariance characterization is better described by the Schur-type algorithms (Lev-Ari and Kailath, 1981) rather than the Levinson-type algorithms. Actually we can develop such an algorithm for our case but do not present it here.

Acknowledgment. The author wishes to express his sincere thanks to Prof. H. Tokumaru of Kyoto University for his useful advice and Y. Iiguni for pointing out several errors in the earlier manuscript.

REFERENCES

BARNOFF-NIELSEN, O. and SCHOU, G. (1973). On the parameterization of autoregressive models by partial autocorrelations. *J. Multivariate Anal.* **3** 408–419.

BURG, J. P. (1975). Maximum entropy spectral analysis. Ph.D. dissertation, Geophysics Dept., Stanford Univ.

DELOSME, J.-M. and MORF, M. (1980). A tree classification of algorithms for Toeplitz and related equations including generalized Levinson and doubling type algorithms. *Proc. 19th IEEE Conf. on Decision and Control* 42–46.

LEVI-ARI, H. and KAILATH, T. (1981). Schur and Levinson algorithms for nonstationary processes. *Proc.* 1981 *IEEE Int. Conf. on Acoust., Speech, Signal Processing* 860–864.

MORF, M., VIEIRA, A. and KAILATH, T. (1978). Covariance characterization by partial autocorrelation matrices. *Ann. Statist.* **6** 643–648.

PAGANO, M. (1978). Periodic and multiple autoregressions. *Ann. Statist.* **6** 1310–1317.

RAMSEY, F. (1974). Characterization of the partial autocorrelation function. *Ann. Statist.* **2** 1296–1301.

SAKAI, H. (1982). Circular lattice filtering using Pagano's method. *IEEE Trans. Acoust., Speech, Signal Processing* **ASSP-30** 279–287.

DEPARTMENT OF APPLIED MATHEMATICS
AND PHYSICS
KYOTO UNIVERSITY
KYOTO 606, JAPAN.