SIAM J. APPL. MATH.
Vol. 15, No. 6, November, 1967
Printed in U.S.A.

WEIGHTING COEFFICIENTS FOR THE PREDICTION OF STATIONARY TIME SERIES FROM THE FINITE PAST*

WILLIAM F. TRENCH†

1. Introduction. This paper presents a method for calculating weighting coefficients for the minimum variance linear prediction of weakly stationary time series on the basis of $n + 1$ consecutive observations. The theoretical solution of this problem is well known, since the coefficients satisfy a linear algebraic system of order $n + 1$; however, the numerical solution of such a system is difficult if n is large. We give a simple algorithm, part of which could be derived from the properties of a certain class of orthogonal polynomials [2]; however, the derivation given here does not require a knowledge of that theory.

2. Formulation of the problem. Let $\{y_m\}$ be a complex-valued weakly stationary time series, with zero mean, unit variance, and correlation sequence $\{\phi_r\}$:

$$E\bar{y}_m y_{r+m} = \phi_r, \qquad \phi_0 = 1. \tag{1}$$

Let the joint distribution of $y_0, y_1, \cdots, y_n$ be of rank $n + 1$, so that the Hermitian matrix

$$T_n = (\phi_{r-s})_{r,s=0}^{n} \tag{2}$$

is positive definite.

We consider the problem of predicting y_{m+k}, $k \geqq 1$, given observed values $y_m, y_{m-1}, \cdots, y_{m-n}$ of a realization of the process. It is known [6, pp. 126–129] that the best prediction based on a linear combination of the observed values is

$$x_{mn}^{(k)} = \sum_{r=0}^{n} \psi_{rn}^{(k)} y_{m-r},$$

where

$$\sum_{s=0}^{n} \phi_{r-s} \psi_{sn}^{(k)} = \phi_{r+k}, \qquad 0 \leqq r \leqq n. \tag{3}$$

This is the best prediction in the sense that, of all linear combinations of $y_m, y_{m-1}, \cdots, y_{m-n}$, it minimizes the variance $E \mid x_{mn}^{(k)} - y_{m+k} \mid^2$. Under the assumption on the rank of the distribution, (3) has a unique solution.

* Received by the editors March 29, 1967.

† Department of Mathematics, Drexel Institute of Technology, Philadelphia, Pennsylvania 19104.

Little work has been devoted to finding explicit expressions for the weighting coefficients in the finite case, or to developing efficient methods for solving (3). For the purely autoregressive case, the complete solution is known explicitly [2, pp. 186–187]. Kozuljaev [3] has solved the prediction problem for the special finite moving average process

$$y_m = (r+1)^{-1/2} \sum_{i=0}^{r} x_{m-i},$$

where $\{x_i\}$ is an uncorrelated zero mean process. In this case,

$$\text{(4)} \qquad \begin{aligned} \phi_i &= 1 - |i/r|, & |i| &\leqq r, \\ \phi_i &= 0, & |i| &> r. \end{aligned}$$

Kozuljaev solved (3) by means of Cramer's rule, which requires considerable calculation of determinants. The author [5] obtained the same results by a simple device (discussed in §4 below) that applies to any process which is the finite moving average of an uncorrelated process. The same device was stated independently for $k = 1$ by Davisson [1].

In [4] the author gave a method for inverting matrices of the form (2) that takes advantage of their simple structure. This paper presents an algorithm for solving (3) directly, without requiring the inversion of (2). In the most general case, the number of multiplications required is proportional to n^2. However, if the generating function

$$\Phi(z) = \sum_{-\infty}^{\infty} \phi_r z^r$$

is rational, the technique simplifies so that the number of multiplications is proportional to n. All processes of the mixed autoregressive, finite moving average type are in this class.

3. Recursion formulas for the general case.

THEOREM 1. *If the joint distribution of* $y_0, y_1, \cdots, y_n$ *is of rank* $n + 1$, *the weighting coefficients satisfy the following recursion formulas*:

$$\text{(5)} \qquad \psi_{00}^{(j)} = \phi_j, \qquad \lambda_0 = 1,$$

and if $1 \leqq m \leqq n$, *then*

$$\text{(6)} \qquad \lambda_m = (1 - |\psi_{m-1,m-1}^{(1)}|^2)\lambda_{m-1},$$

$$\text{(7)} \qquad \psi_{mm}^{(j)} = \lambda_m^{-1}\left(\phi_{m+j} - \sum_{s=0}^{m-1} \psi_{s,m-1}^{(1)} \phi_{m+j-s-1}\right),$$

and

$$\text{(8)} \qquad \psi_{rm}^{(j)} = \psi_{r,m-1}^{(j)} - \psi_{mm}^{(j)} \bar{\psi}_{m-r-1,m-1}^{(1)}, \qquad 0 \leqq r \leqq m-1.$$

Instead of proving Theorem 1 directly, we shall obtain it as a special case of Theorem 2 below. Theorem 1 provides an efficient method for computing $\psi_{0n}^{(k)}, \cdots, \psi_{nn}^{(k)}$, as follows.

(a) If $k = 1$, start with (5); then compute (6), (7) and (8) for $m = 1, 2, \cdots, n$, with $j = 1$.

(b) If $k > 1$, compute (5) with $j = 1$ and $j = k$. For each m, first compute (6), (7) and (8) with $j = 1$; then compute (7) and (8) with $j = k$. Repeat for $m = 1, 2, \cdots, n$.

This sequence of computations is economical of storage. It is only necessary to retain quantities computed at level $m - 1$ until the computations at level m are complete.

For $k = 1$, this algorithm was derived in [4]. For the reader who wishes to refer to that paper, we point out that $\psi_{rn}^{(1)}$ is denoted there by ψ_{rn}, and λ_n by Δ_{n-1}.

THEOREM 2. *Let $\eta_0, \eta_1, \cdots, \eta_n$ be arbitrary, and for each $m \leqq n$, let $\xi_{0m}, \xi_{1m}, \cdots, \xi_{mm}$ be the solution of*

$$\sum_{s=0}^{m} \phi_{r-s}\xi_{sm} = \eta_r, \qquad 0 \leqq r \leqq m. \tag{9}$$

Then

$$\xi_{00} = \eta_0, \tag{10}$$

and for $1 \leqq m \leqq n$,

$$\xi_{mm} = \lambda_m^{-1}\left(\eta_m - \sum_{s=0}^{m-1} \psi_{s,m-1}^{(1)}\eta_{m-s-1}\right) \tag{11}$$

(where λ_m is defined in Theorem 1) and

$$\xi_{rm} = \xi_{r,m-1} - \xi_{mm}\bar{\psi}_{m-r-1,m-1}^{(1)}, \qquad 0 \leqq r \leqq m - 1. \tag{12}$$

Proof. Since $\phi_0 = 1$, (10) is obvious. Define $\xi_{m,m-1} = 0$. Then, from (9),

$$\sum_{s=0}^{m} \phi_{r-s}(\xi_{sm} - \xi_{s,m-1}) = \delta_{rm}\left(\eta_m - \sum_{s=0}^{m-1} \phi_{m-s}\xi_{s,m-1}\right) \tag{13}$$

for $0 \leqq r \leqq m$. In [4] it is shown that the elements of the last column of T_m^{-1} are

$$b_{rmm} = -\lambda_m^{-1}\bar{\psi}_{m-r-1,m-1}^{(1)}, \qquad 0 \leqq r \leqq m - 1, \tag{14}$$

and

$$b_{mmm} = \lambda_m^{-1}. \tag{15}$$

From (13) and (15),

$$\xi_{mm} = \lambda_m^{-1}\left(\eta_m - \sum_{s=0}^{m-1} \phi_{m-s}\xi_{s,m-1}\right). \tag{16}$$

From (9) with m replaced by $m - 1$,

$$\sum_{s=0}^{m-1} \phi_{m-r-1-s}\xi_{s,m-1} = \eta_{m-r-1}, \quad 0 \leqq r \leqq m - 1.$$

Hence

$$\sum_{r=0}^{m-1} \psi_{r,m-1}^{(1)} \sum_{s=0}^{m-1} \phi_{m-r-1-s}\xi_{s,m-1} = \sum_{r=0}^{m-1} \psi_{r,m-1}^{(1)}\eta_{m-r-1}.$$

By interchanging the order of summation on the left, and noting that, from (3),

$$\sum_{r=0}^{m-1} \phi_{m-r-1-s}\psi_{r,m-1}^{(1)} = \phi_{m-s}, \quad 0 \leqq s \leqq m - 1,$$

we find that

$$\sum_{s=0}^{m-1} \phi_{m-s}\xi_{s,m-1} = \sum_{r=0}^{m-1} \psi_{r,m-1}^{(1)}\eta_{m-r-1}.$$

From this and (16), (11) follows. From (11), (13) and (14), (12) follows, which completes the proof of Theorem 2.

Theorem 1 is a special case of Theorem 2, with $\eta_r = \phi_{r+k}$.

Theorem 1 (with $k = 1$) and Theorem 2 provide an efficient algorithm for computing the solution of a system of the form

$$\sum_{s=0}^{n} \phi_{r-s}\xi_s = \eta_r, \qquad 0 \leqq r \leqq n.$$

4. Prediction of the finite moving average process. If $\{y_m\}$ is a finite moving average of an uncorrelated process, there is an integer p such that $\phi_r = 0$ if $|r| > p$, and the elements of T_m lying outside a diagonal strip vanish. From (3), it can be shown that the definition of $\psi_{rm}^{(k)}$ can be extended for all r so that

$$\sum_{j=-p}^{p} \bar{\phi}_j\psi_{r+j,m}^{(k)} = 0, \qquad -\infty < r < \infty, \tag{17a}$$

$$\psi_{rm}^{(k)} = -\delta_{-k,r}, \qquad -p \leqq r \leqq -1, \tag{17b}$$

$$\psi_{rm}^{(k)} = 0, \qquad m + 1 \leqq r \leqq m + p. \tag{17c}$$

This reduces the calculation of the weighting coefficients to solving a boundary value problem of order $2p$. The author used this fact in [5] for the special case (4), for which the difference equation can be solved explicitly. However, even if it is not convenient to solve (17a)–(17c) explicitly, it is still useful to employ them in the following computational scheme, which provides $\psi_{0n}^{(k)}, \psi_{1n}^{(k)}, \cdots, \psi_{nn}^{(k)}$ and requires a number of multiplications proportional to n, rather than to n^2.

Equation (8) now holds for all r, as can be verified by observing that, with the extended definition, the right member satisfies (17a)–(17c), and coincides with the left member for $-p \leqq r \leqq m + p - 1$, a total of $m + 2p$ values.

A careful examination of (8) shows that if $\psi_{r,m-1}^{(1)}$ is known for $0 \leqq r \leqq p - 1$ and $m - p \leqq r \leqq m - 1$, then $\psi_{rm}^{(1)}$ can be calculated for $0 \leqq r \leqq p - 1$ and $m - p + 1 \leqq r \leqq m$. Similarly, if $k > 1$, and $\psi_{r,m-1}^{(k)}$ is also known for $m - p \leqq r \leqq m - 1$, $\psi_{rm}^{(k)}$ can be calculated for $m - p + 1 \leqq r \leqq m$. To see this, observe that (8) implies that

$$\psi_{mm}^{(k)} = \frac{\psi_{m+p,m-1}^{(k)}}{\bar{\psi}_{-p-1,m-1}^{(1)}}, \tag{18}$$

since $\psi_{m+p,m}^{(k)} = 0$. From (17a) and (17c),

$$\psi_{m+p,m-1}^{(k)} = -\bar{\phi}_p^{-1} \sum_{j=1}^{p} \phi_j \psi_{m-j,m-1}^{(k)};$$

and from (17a) and (17b),

$$\psi_{-p-1,m-1}^{(1)} = \phi_p^{-1} \left(1 - \sum_{j=1}^{p} \bar{\phi}_j \psi_{j-1,m-1}^{(1)}\right). \tag{19}$$

From these last three equations, $\psi_{mm}^{(1)}$ and $\psi_{mm}^{(k)}$ can be computed; the other values of $\psi_{rm}^{(1)}$ and $\psi_{rm}^{(k)}$ mentioned above can then be calculated from (8). In this way it is possible to proceed recursively up to $m = n$. At each step only $2p$ values of $\psi_{rm}^{(1)}$ and (if $k \neq 1$) only p values of $\psi_{rm}^{(k)}$ need be computed. Given $\psi_{rn}^{(k)}$ for $n - p + 1 \leqq r \leqq n$, the remaining coefficients can be obtained recursively from (17a) and (17c):

$$\psi_{rn}^{(k)} = -\phi_p^{-1} \sum_{j=-p+1}^{p} \bar{\phi}_j \psi_{r+p+j,n}^{(k)}, \qquad r = n - p,\ n - p - 1,\ \cdots,\ 0.$$

The computations described in this section are of interest only if $k \leqq p$; since $\phi_r = 0$ if $|r| > p$, (3) and (16) each imply that $\psi_{rn}^{(k)} = 0$ for all r if $k > p$.

5. The mixed moving average, autoregressive case. The method of the previous section can be generalized to the case where the generating function

$$\Phi(z) = \sum_{-\infty}^{\infty} \phi_r z^r$$

is a rational function. If $\{x_r\}$ is an uncorrelated sequence and $\{y_m\}$ is the solution of

$$(20)\qquad y_m + a_1 y_{m-1} + \cdots + a_q y_{m-q} = c_0 x_m + c_1 x_{m-1} + \cdots + c_p x_{m-p}, \qquad -\infty < m < \infty,$$

then $\{y_m\}$ is weakly stationary, provided the roots of

$$A(z) = \sum_0^q a_i z^i, \qquad a_0 = 1,$$

lie outside the unit circle. The autocorrelation sequence is generated by

$$(21)\qquad \phi(z) = \sigma^{-2}\, C(z)C^*(1/z)/[A(z)A^*(1/z)],$$

where σ is chosen so that $\phi_0 = 1$, and

$$C(z) = \sum_0^p c_r z^r.$$

The coefficients a_q, c_0, and c_p are assumed to be nonzero. (If f is a polynomial, then f^* is the polynomial whose coefficients are the conjugates of those of f.)

For the process (20), $\psi_{rn}^{(k)}$ can be calculated with a number of multiplications proportional to n. The algorithm is readily derivable from Theorems 3 and 4 below.

Define $\{\gamma_r\}$, $\{\alpha_r\}$ and $\{b_r\}$ by their generating functions:

$$(22)\qquad C(z)C^*(1/z) = \sum_{-p}^{p} \gamma_r z^r,$$

$$(23)\qquad [A(z)A^*(1/z)]^{-1} = \sum_{-\infty}^{\infty} \alpha_r z^r$$

and

$$[A(z)]^{-1} = \sum_0^\infty b_r z^r.$$

THEOREM 3. *Let $\{w_{rm}^{(k)}\}$ be the solution of the difference equation*

$$(24)\qquad \sum_{j=-p}^{p} \bar{\gamma}_j w_{r+j,m}^{(k)} = 0, \qquad -\infty < r < \infty,$$

that satisfies the boundary conditions

$$(25)\qquad \sum_{j=0}^{q} \bar{a}_j w_{j-i,m}^{(k)} = -b_{k-i}, \qquad 1 \leqq i \leqq p,$$

and

$$\sum_{j=0}^{q} a_j w_{m+i-j,m}^{(k)} = 0, \qquad 1 \leqq i \leqq p. \tag{26}$$

Then if $m \geqq q - p$,

$$w_{rm}^{(k)} = \sum_{s=0}^{m} \alpha_{r-s} \psi_{sm}^{(k)} - \alpha_{r+k}, \qquad -p \leqq r \leqq m + p. \tag{27}$$

Proof. It is only necessary to verify that the right side of (27) satisfies (24), (25) and (26), because there is only one sequence which does so. From (21), (22) and (23),

$$\phi_r = \sigma^{-2} \sum_{j=-p}^{p} \gamma_j \alpha_{r-j} .$$

Using this, it can be shown that the right side of (27) satisfies (24) for $0 \leqq r \leqq m$, by substituting in (3) and regrouping terms. It also satisfies (25) and (26), as can be verified by using the relations between $\{\alpha_r\}$, $\{a_r\}$, and $\{b_r\}$ that are implied by their generating functions. This completes the proof of Theorem 3.

The system (24), (25) and (26) reduces to (17) if $A(z) \equiv 1$, in which case $w_{rm}^{(k)} = \psi_{rm}^{(k)}$.

Theorem 3 reduces the calculation of $\psi_{rm}^{(k)}$, $0 \leqq r \leqq m$, to solving a difference equation of order $2p$, for if $w_{rm}^{(k)}$ is known for $0 \leqq r \leqq m$, the weighting coefficients can be obtained by solving (27), with $0 \leqq r \leqq m$. This is easy to do, because the matrix $T_m = (\alpha_{r-s})$, $0 \leqq r, s \leqq m$ is invertible by inspection if $m \geqq 2q$. If the (r, s)th element of T_m^{-1} is denoted by b_{rsm}, then

$$\sum_{s=0}^{m} b_{rsm} z^r = \begin{cases} z^r \left(\sum_0^r a_i z^{-i} \right) A^*(z), & 0 \leqq r \leqq q - 1, \\ z^r A(1/z) A^*(z), & q \leqq r \leqq m - q, \\ z^r A(1/z) \left(\sum_0^{m-r} \bar{a}_i z^i \right), & m - q + 1 \leqq r \leqq m. \end{cases} \tag{28}$$

Example. Consider the sequence $\{y_j\}$ obtained from the uncorrelated sequence $\{x_j\}$ by

$$y_j - \lambda y_{j-1} = x_j - x_{j-1} ,$$

where $-1 < \lambda < 1$. The difference equation (17) reduces to

$$w_{r+1,m}^{(k)} - 2w_{rm}^{(k)} + w_{r-1,m}^{(k)} = 0, \qquad -\infty < r < \infty,$$

$$w_{-1,n}^{(k)} - \lambda w_{0n}^{(k)} = -\lambda^{k-1},$$

$$w_{m+1,m}^{(k)} - \lambda w_{mm}^{(k)} = 0.$$

Thus, $w_{rm}^{(k)}$ is a first degree polynomial in r; the coefficients can be determined from the boundary conditions. The result is

$$w_{rm}^{(k)} = \frac{-\lambda^{k-1}[(m-r)(1-\lambda)+1]}{(1-\lambda)[m(1-\lambda)+2]}.$$

T_m^{-1} can be obtained from (28) with $A(z) = 1 - \lambda z$; the final result is

$$\psi_{rm}^{(k)} = (1-\lambda)^2 w_{rm}^{(k)}.$$

Even if it is not convenient to solve the difference equation explicitly, Theorem 3 and (28) can be combined with the following theorem, whose proof we omit, to devise an efficient algorithm for calculating the weighting coefficients. The required number of multiplications is proportional to n.

THEOREM 4. *For any* $k \geqq 1$,

$$w_{rm}^{(k)} = w_{r,m-1}^{(k)} - \psi_{mm}^{(k)} \bar{w}_{m-r-1,m-1}, \qquad -\infty < r < \infty,$$

and

$$\psi_{mm}^{(k)} = \left(\sum_{j=0}^{q} a_j w_{m+p-j,m-1}^{(k)}\right)\left(\sum_{j=0}^{q} a_j \bar{w}_{j-p-1,m-1}^{(1)}\right)^{-1}. \tag{29}$$

The algorithm is constructed in a way analogous to that of §3; the details are tedious to write out for the general case, but straightforward for any particular $A(z)$ and $C(z)$.

6. Variance of the estimates. Let

$$\sigma_{nk}^2 = E\left|y_{m+k} - \sum_{r=0}^{n} \psi_{rn}^{(k)} y_{m-r}\right|^2.$$

The known result,

$$\sigma_{nk}^2 = 1 - \sum_{r=0}^{n} \bar{\psi}_{rn}^{(k)} \phi_{k+r},$$

follows from (1) and (3). In particular, (19) implies that

$$\bar{\psi}_{-p-1,m-1}^{(1)} = \phi_p^{-1} \sigma_{m-1,1}^2$$

for the moving average case. This does not vanish, because the process is nondeterministic; hence, the division in (18) is legitimate. For the mixed moving average, autoregressive case, it can be shown that

$$\sum_{j=0}^{q} a_j \bar{w}_{j-p-1,m-1}^{(1)} = \sigma_{m-1,1}^2 / \bar{\gamma}_p,$$

which does not vanish; hence, the division in (29) is legitimate.

REFERENCES

[1] L. D. DAVISSON, *Prediction of time series from finite past*, this Journal, 13 (1965), pp. 819–826.

[2] U. GRENANDER AND G. SZEGÖ, *Toeplitz Forms and Their Applications*, University of California Press, Berkeley and Los Angeles, 1958.

[3] P. A. KOZULJAEV, *On the theory of extrapolation of stationary sequences*, Selected Transl. Math. Statist. and Prob., vol. 3, American Mathematical Society, Providence, 1962, pp. 91–129.

[4] W. F. TRENCH, *An algorithm for the inversion of finite Toeplitz matrices*, this Journal, 12 (1964), pp. 515–522.

[5] ———, *On the extrapolation of a special class of stationary time series*, Ann. Math. Statist., 36 (1965), pp. 1426–1432.

[6] A. M. YAGLOM, *An Introduction to the Theory of Stationary Random Functions*, rev. ed., Prentice-Hall, Englewood Cliffs, New Jersey, 1962.