Explicit weighting coefficients for predicting ARMA time series from the finite past *

William F. Trench
Department of Mathematics, Trinity University, San Antonio, TX 78212, United States

Received 20 July 1990

Abstract

Explicit formulas are given for the weighting coefficients in the linear minimum variance predictor of a wide sense stationary autoregressive-moving average time series $k$ steps ahead, given $n + 1$ successive observations of a realization of the process. The formulas involve determinants whose entries are the values of certain polynomials related to the autoregressive part of the process at the zeros of the polynomial that defines the moving average part. The number of observations $n + 1$ enters into the formulas as a parameter, and not so as to increase their complexity as $n$ grows large. Formulas are also given for the variance of the prediction.

Keywords: ARMA time series, prediction, minimum variance, stationary, Toeplitz matrix.

1. Introduction

Let

$$A(z) = \sum_{i=0}^{q} a_i z^i \quad \text{and} \quad C(z) = \sum_{i=0}^{p} c_i z^i$$

be polynomials with complex coefficients, and suppose that

$$a_0 a_q c_0 c_p \neq 0,$$

and $A(z)$ has no zeros in $|z| < R$ for some $R > 1$. Let $\{x_j\}_{-\infty}^{\infty}$ be a real-valued wide sense stationary [9, p.15] uncorrelated time series with zero mean and unit variance, i.e.,

$$E x_j = 0 \quad \text{and} \quad E x_i x_{i+j} = \delta_{0j}, \quad -\infty < j < \infty.$$

Then the stochastic difference equation

$$a_0 y_m + a_1 y_{m-1} + \cdots + a_q y_{m-q} = c_0 x_m + c_1 x_{m-1} + \cdots + c_p x_{m-p}$$

\(1\)

* This work was partially supported by the National Science Foundation under Grant No. DMS-8907939.
defines an autoregressive-moving average (ARMA) time series with zero mean and autocorrelation sequence
\[ \phi_r = E \tilde{y}_t y_{t+r}, \quad -\infty < r < \infty, \quad (2) \]
which is generated by the Laurent expansion
\[ \frac{C(z)C^*(1/z)}{A(z)A^*(1/z)} = \sum_{r=-\infty}^{\infty} \phi_r z^r, \quad \frac{1}{R} < |z| < R, \]
[8, pp.20–23]. (Here and in the following, if \( f(z) \) is a polynomial, then \( f^*(z) \) is the polynomial with coefficients equal to the conjugates of those of \( f(z) \).)

Here we consider the problem of linear least square prediction of \( y_{m+k}, \ k \geq 1 \), given observed values \( y_m, \ldots, y_{m-n} \) of a realization of the process (1). This is a standard problem, and its solution is well understood [8, pp.70–77], [9, pp.100–103]: the matrix \( (\phi_{r-s})_{r,s=0}^{\infty} \) is positive definite, and if
\[ \psi_{0n}, \ldots, \psi_{nn} \]
is the (unique) solution of the system
\[ \sum_{s=0}^{n} \phi_{r-s} \psi_{sn} = \phi_{r+k}, \quad 0 \leq r \leq n, \quad (3) \]
then the quantity
\[ z_{mn}^{(k)} = \sum_{r=0}^{n} \psi_{rn} y_{m-r} \quad (4) \]
is the best estimate of the desired form for \( y_{m+k} \), in that the variance of the error
\[ \sigma_{nk}^2 = E |z_{mn}^{(k)} - y_{m+k}|^2 \]
is less than that which would be obtained from any other linear combination of \( y_m, \ldots, y_{m-n} \).

We are interested here in obtaining explicit formulas for the weighting coefficients \( \psi_{0n}^{(k)}, \ldots, \psi_{nn}^{(k)} \), so we will make no attempt to survey the voluminous existing literature on specially designed numerical methods for solving (3), most of which exploit the Toeplitz structure of (3), but not the fact that \( \{ y_m \} \) is specifically an ARMA process. Numerical methods which do take advantage of the latter were proposed by the author [6] and Newton and Pagano [4]. For the case where \( p = 0 \) (so that (1) is purely autoregressive), explicit formulas for the weighting coefficients are known [1, pp.186–187], and so we assume henceforth that \( p \geq 1 \). Whittle [8, pp.75–76] has shown that in this case the determination of the weighting coefficients can, in principle, be reduced to solving a certain system of \( 2p \) linear equations. For the specific moving average process where (1) reduces to
\[ y_m = (p + 1)^{-1/2} \sum_{i=0}^{p} x_{m-i}, \]
2. The main results

It is convenient to define sequences \( \{ \gamma_r \}, \{ \alpha_r \} \) and \( \{ b_r \} \) by their generating functions

\[
C(z)C^*(1/z) = \sum_{r=-p}^{p} \gamma_r z^r,
\]

\[
[A(z)A^*(1/z)]^{-1} = \sum_{r=-\infty}^{\infty} \alpha_r z^r, \quad \frac{1}{R} < |z| < R,
\]  \( (5) \)

and

\[
[A(z)]^{-1} = \sum_{r=-\infty}^{\infty} b_r z^r, \quad |z| < R.
\]  \( (6) \)

Notice that

\[
b_r = 0, \quad r < 0,
\]  \( (7) \)

and that

\[
\phi_r = \sum_{j=-p}^{p} \gamma_j \alpha_{r-j}, \quad -\infty < r < \infty.
\]  \( (8) \)

We will also find it convenient to define

\[
a_i = 0 \quad \text{if } i < 0 \text{ or } i > q, \quad c_i = 0 \quad \text{if } i < 0 \text{ or } i > p, \quad \gamma_i = 0 \quad \text{if } |i| > p,
\]  \( (9) \)

and

\[
\sum_{i}^j = 0 \quad \text{if } i > j.
\]

We need the following definition in order to state our main theorem.

**Definition 1.** Let \( z_1, \ldots, z_L \) be the distinct zeros of the Laurent polynomial

\[
J(z) = C(z)C^*(1/z),
\]

with multiplicities \( m_1, \ldots, m_L \); thus,

\[
L \leq 2p \quad \text{and} \quad m_1 + \cdots + m_L = 2p.
\]

If \( Q(z) \) is an arbitrary Laurent polynomial, define the 2\( p \)-dimensional row vector **generated by** \( Q(z) \) as follows: its first \( m_1 \) entries are \( Q^{(l)}(z_1), 0 \leq l \leq m_1 - 1 \); its next \( m_2 \) entries are \( Q^{(l)}(z_2), 0 \leq l \leq m_2 - 1 \); and so forth. If \( Q_1(z), \ldots, Q_{2p}(z) \) are Laurent polynomials, let

\[
D[Q_1(z), \ldots, Q_{2p}(z)]
\]  \( (10) \)

be the determinant of order 2\( p \) whose \( \mu \)th row, \( 1 \leq \mu \leq 2p \), is generated by \( Q_\mu(z) \). (Notice that \( D[Q_1(z), \ldots, Q_{2p}(z)] \) is not a function of \( z \); rather it is a constant which depends upon the values of \( Q_1(z), \ldots, Q_{2p}(z) \) and their derivatives at the zeros of \( J(z) \).)
If $J(z)$ has $2p$ distinct roots $z_1, \ldots, z_{2p}$ (which occurs if and only if $C(z)$ has distinct roots $\xi_1, \ldots, \xi_p$ such that $\xi_i \xi_j \neq 1$, $1 \leq i, j \leq p$), then

$$D\left[ Q_1(z), \ldots, Q_{2p}(z) \right] = \det\left[ Q_{\mu}(z) \right]_{\mu, \nu = 1}^{2p}.$$ 

The following lemma makes explicit an elementary observation which we use repeatedly below.

**Lemma 2.** Let $\nu$ be a given integer in $\{1, \ldots, 2p\}$. Let $Q_i(z)$, $1 \leq i \leq 2p$, $i \neq \nu$, and $P_1(z), \ldots, P_k(z)$ be given Laurent polynomials, and define

$$P(z) = \sum_{j=1}^{k} d_j P_j(z), \tag{11}$$

where $d_1, \ldots, d_k$ are constants. Then

$$\sum_{j=1}^{k} d_j D\left[ Q_1(z), \ldots, Q_{\mu-1}(z), P_j(z), Q_{\mu+1}(z), \ldots, Q_{2p}(z) \right]$$

$$= D\left[ Q_1(z), \ldots, Q_{\mu-1}(z), P(z), Q_{\mu+1}(z), \ldots, Q_{2p}(z) \right]. \tag{12}$$

**Proof.** All determinants in (12) have identical rows, except for the $\mu$th. Therefore, expanding the determinant on the right in cofactors with respect to the $\mu$th row and invoking (11) yields (12).

The following theorem is our main result.

**Theorem 3.** Suppose

$$n > \max(0, 2q - 2p). \tag{13}$$

Let $D_n$ be the determinant of the form (10) with

$$Q_i(z) = \begin{cases} z^{i-1} A(z), & 1 \leq i \leq p, \\ z^{n+i} A^*(1/z), & p + 1 \leq i \leq 2p. \end{cases} \tag{14}$$

If $1 \leq \nu \leq p$, let $E_{\nu n}(r)$ be the determinant of the form (10) with $Q_i(z)$ as in (14) for $i \neq p + \nu$, and

$$Q_{p+\nu}(z) = z^{n+p-r} \left( \sum_{\mu=0}^{r+p} a_\mu z^{n+\nu} \right) \left( \sum_{r=0}^{n+p-r} a_\nu z^{-r} \right), \quad 0 \leq r \leq n. \tag{15}$$

Then $D_n \neq 0$, and

$$\psi_{n}^{(k)} = -\frac{1}{D_n} \sum_{\nu=0}^{p} b_{k-\nu} E_{\nu n}(r) - \sum_{i=p+r+1}^{q} a_i b_{r+k-i}, \quad 0 \leq r \leq n. \tag{16}$$

Since $T_n$ is positive definite, [7, Theorem 2] implies that $D_n \neq 0$. It is convenient to present the rest of the lengthy proof of this theorem as a series of lemmas. We ask the reader to consider $n$ and $k$ fixed throughout, and to recognize that intermediate quantities defined below (such as $h_{rs}$,
Lemma 4. Let
\[ \Gamma = (\alpha_{s-r})_{r,s=-p}^{n+p}, \]
with \( \{ \alpha_j \} \) as defined in (5); thus \( \Gamma \) is a Toeplitz matrix of order \( n + 2p \), with rows and columns numbered from \(-p\) to \( n + p \) (for convenience below). Then, if (13) holds,
\[ \Gamma^{-1} = (h_{rs})_{r,s=-p}^{n+p}, \]
where
\[ \sum_{s=-p}^{n+p} h_{rs}z^s = H_r(z) = z^rA(1/z)A^*(z), \quad q-p \leq r \leq n + p - q, \]
(18)
Proof. It is easily established that
\[ H_r(z)[A(1/z)A^*(z)]^{-1} = \begin{cases} z^r + O(z^{-p-1}), & -p \leq r \leq q - p - 1, \\ z^r, & q - p \leq r \leq n + p - q, \\ z^r + O(z^{n+p+1}), & n + p - q + 1 \leq r \leq n + p, \end{cases} \]
(19)
where \( O(z^{-p-1}) \) stands for a series in \( 1/z \) starting with the \((p + 1)\)th power, and \( O(z^{n+p+1}) \) stands for a series in \( z \), starting with the \((n + p + 1)\)th power. Therefore
\[ \sum_{j=-p}^{n+p} h_{rj} \alpha_{j-s} = \delta_{rs}, \quad -p \leq r, s \leq n + p, \]
(20)
since the left side here is the coefficient of \( z^s \) in the expansion of the left side of (19). (See (5).) Clearly, (20) is equivalent to (17). \( \square \)

Except for a slight difference in form due to our present peculiar numbering of the rows and columns of \( \Gamma \), Lemma 4 was stated earlier in [6]. (For more on Toeplitz matrices with inverses of this type, see [2].)

Lemma 5. The sequence \( \{ w_r \}_{-p}^{n+p} \) satisfies the difference equation
\[ \sum_{j=-p}^{p} \gamma_j w_{r-j} = 0, \quad 0 \leq r \leq n, \]
(21)
and the boundary conditions
\[ \sum_{j=0}^{q} \tilde{a}_j w_{j-i} = -b_{k-i}, \quad 1 \leq i \leq p, \]
(22)
and
\[ \sum_{j=0}^{q} a_j w_{n+i-j} = 0, \quad 1 \leq i \leq p, \] (23)
if and only if
\[ w_r = \sum_{s=0}^{n} \alpha_{r-s} \psi_{sn}^{(k)} - \alpha_{r+k}, \quad -p \leq r \leq n + p. \] (24)

**Proof.** If \( \{ w_r \} \) is defined by (24), then (3) and (8) imply (21). To verify (22) and (23), recall that
\[ \sum_{l=0}^{q} \tilde{a}_l \alpha_{r+l} = b_r \] (25)
and that
\[ \sum_{l=0}^{q} a_l \alpha_{r-l} = 0, \quad \nu > 0, \]
because of (5)–(7).

For necessity, suppose that \( \{ w_r \} \) satisfies (21)–(23), and define \( \xi_{-p}, \ldots, \xi_{n+p} \) as the solution of the system
\[ \sum_{s=-p}^{n+p} \alpha_{r-s} \xi_s = w_r + \alpha_{r+k}, \quad -p \leq r \leq n + p. \] (26)

We will show that
\[ \xi_r = 0 \quad \text{if} \quad -p \leq r \leq -1 \quad \text{or} \quad n + 1 \leq r \leq n + p. \] (27)
This implies that (26) is equivalent to
\[ w_r = \sum_{s=0}^{n} \alpha_{r-s} \xi_s - \alpha_{r+k}, \quad -p \leq r \leq n + p. \] (28)

Substituting this into (21) and recalling (8) shows that
\[ \sum_{s=0}^{n} \phi_{r-s} \xi_s = \phi_{r+k}, \quad 0 \leq r \leq n. \]
Comparing this with (3), we conclude that
\[ \xi_r = \psi_{rn}^{(k)}, \quad 0 \leq r \leq n, \] (29)
since \( (\phi_{r-s})_{n,s=0}^{n} \) is positive definite. Since this and (28) imply (24), the proof of necessity reduces to establishing (27).

To this end we use Lemma 4 to solve (26) for \( \xi_{-p}, \ldots, \xi_{n+p}: \)
\[ \xi_r = \sum_{s=-p}^{n+p} h_{rs} w_s + \sum_{s=-p}^{n+p} h_{rs} \alpha_{r+k}, \quad -p \leq r \leq n + p. \] (30)
As observed in the proof of Lemma 4, the second term on the right of (30) is the coefficient of $z^{-k}$ in the expansion of the left side of (19); hence, routine manipulations using (5), (6) and (18) yield

$$\sum_{s=-p}^{n+p} h_{rs} \alpha_{s+k} = \begin{cases} \delta_{-k,r} - \sum_{i=p+r+1}^{q} a_i b_{r+k-i}, & -p \leq r \leq q - p - 1, \\ \delta_{-k,r}, & q - p \leq r \leq n + p. \end{cases}$$

(31)

Now define

$$W(z) = \sum_{s=-p}^{n+p} w_s z^s,$$

and notice that the first term on the left of (30) is the constant term in the expansion of

$$G_r(z) = H_r(z) W(1/z).$$

(32)

If $n + 1 \leq r \leq n + p$, it is convenient to write (32) as

$$G_r(z) = \left( \sum_{i=-n-1}^{n+p-r} \tilde{a}_i z^{i+r} \right) A(1/z) W(1/z).$$

(33)

(Recall (9).) Because of (23), the coefficient of $z^j$ in the expansion of $A(1/z) W(1/z)$ is zero for $-n - p \leq j \leq -n - 1$; therefore, (33) shows that $G_r(z)$ has zero constant term if $n + 1 \leq r \leq n + p$. Consequently,

$$\sum_{s=-p}^{n+p} h_{rs} w_s = 0, \quad n + 1 \leq r \leq n + p.$$

This and (31) imply that $\xi_r = 0$ if $n + 1 \leq r \leq n + p$.

Now suppose that $-p \leq r \leq -1$. Then it is convenient to write

$$G_r(z) = \left( \sum_{i=0}^{r+p} a_i z^{r-i} \right) A^*(z) W(1/z).$$

(34)

Because of (22),

$$A^*(z) W(1/z) = - \sum_{i=1}^{p} b_{k-i} z^i + \cdots,$$

where "\cdots" stands for a sum of nonpositive powers of $z$. From this and (34), the constant term in $G_r(z)$ is

$$\sum_{s=-p}^{n+p} h_{rs} w_s = - \sum_{i=0}^{r+p} a_i b_{r+k-i}.$$

This, (30) and (31) imply that

$$\xi_r = \delta_{-k,r} - \sum_{i=0}^{q} a_i b_{r+k-i} = 0$$

for $-p \leq r \leq -1$ (see (6)).
Lemma 6. For \( v = 1, \ldots, p \), define \( D_{\nu \nu}(r) \) for \(-\infty < r < \infty\) as
\[
D_{\nu \nu}(r) = D[Q_1(z), \ldots, Q_{2p}(z)],
\]
where \( Q_i(z) \) is as in (14) if \( i \neq p + v \), while
\[
Q_{p+v}(z) = z^{n+p-r}.
\]
Then
\[
\sum_{j=-p}^{p} \gamma_j D_{\nu \nu}(r-j) = 0, \quad -\infty < r < \infty, \tag{35}
\]
and
\[
\sum_{j=0}^{q} \tilde{a}_j D_{\nu \nu}(j-i) = \delta_{iv} D_{\nu \nu}, \quad 1 \leq i \leq p, \tag{36}
\]
and
\[
\sum_{j=0}^{q} a_j D_{\nu \nu}(n+i-j) = 0, \quad 1 \leq i \leq p. \tag{37}
\]

Proof. From Lemma 2, the left sides of the last three equations can be regarded as determinants with all rows except the \((p+v)\)th the same as those of \( D_{\nu \nu} \), and \((p+v)\)th rows generated respectively by
\[
z^{n+p-j}J(z) \quad \text{for (35),} \tag{38}
\]
\[
z^{n+p+i}A^*(1/z) \quad \text{for (36),} \tag{39}
\]
and
\[
z^{p-i}A(z) \quad \text{for (37).} \tag{40}
\]
From Definition 1, the row generated by (38) consists entirely of zeros, which implies (35). The polynomial in (39) is the same as the generator of row \( p + i \) of \( D_{\nu \nu} \); hence, the determinant on the left of (36) is zero unless \( i = v \), in which case it equals \( D_{\nu \nu} \). This proves (36). The polynomial in (40) is the same as the generator of row \( p - i + 1 \) of \( D_{\nu \nu} \); hence, the determinant on the left of (37) has two identical rows, which implies (37). \( \square \)

Proof of Theorem 3. Lemma 6 easily implies that the unique solution of (21)–(23) is
\[
w_r = -\frac{1}{D_{\nu \nu}} \sum_{p-1}^{p} b_{k-r} D_{\nu \nu}(r), \quad -p \leq r \leq n + p. \tag{41}
\]
Since the polynomial on the right of (15) is \( z^{n+p}H_r(1/z) \) (see (18)), simply comparing the definitions of \( E_{\nu \nu}(r) \) and \( D_{\nu \nu}(r) \) and invoking (41) shows that
\[
\sum_{s=-p}^{n+p} h_{rs} w_s = -\frac{1}{D_{\nu \nu}} \sum_{p-1}^{p} b_{k-r} E_{\nu \nu}(r).
\]
This and (29)–(31) imply (16). \( \square \)
3. Remarks on implementation

The generator of row \( p + \nu \) of \( E_{\nu q}(r) \) reduces to
\[
Q_{p+\nu}(z) = z^{n+p-\nu}A(z)A^* (1/z)
\]
if \( q - p \leq r \leq n + p - q \). (See (9) and (15).) Therefore,
\[
\sum_{j=-p}^{p} \gamma_j E_{\nu q}(r - j) = 0, \quad q \leq r \leq n - q.
\]
(42)
since, from Lemma 2, the quantity on the left can be interpreted as a determinant of the type (10) with the \((p + \nu)\)th row generated by \( z^{n+p-\nu}J(z)A(z)A^*(1/z) \), which means that it consists entirely of zeros, from Definition 1. Therefore, if any \( 2p \) successive values of \( \{ E_{\nu q}(r)\}_{r=q \leq}^{n-q} \) are known, then the rest can be computed recursively from (42). This observation also applies to the first sum in (16). Whether recursive computation based on (42) is stable would presumably depend upon the location of the zeros of \( J(z) \).

If it is desired to predict \( y_{m+k} \) for many values of \( k \), then it would be sensible to compute the intermediate quantities
\[
E^{(r)}_{mn} = \frac{1}{D_n} \sum_{r=0}^{n} E_{\nu q}(r) y_{m-r}, \quad 1 \leq \nu \leq p.
\]
(43)
just once, and then compute \( z_{mn}^{(k)} \) for the various values of \( k \) by means of the formula
\[
z_{mn}^{(k)} = - \sum_{\nu=1}^{p} b_{k-\nu} E_{mn}^{(\nu)} - \sum_{r=0}^{q} \left( \sum_{i=p+r+1}^{q} a_{r+k-i} \right) y_{m-r},
\]
which follows from (4) and (16).

4. Variance of the estimate

In this section we find the variance
\[
\sigma_{nk}^2 = E y_{m+k} - z_{mn}^{(k)}.
\]
(44)
The known formula \([9, \text{p.102}]\)
\[
\sigma_{nk}^2 = \phi_0 - \sum_{s=0}^{n} \phi_{k-s} \psi_{mn}^{(k)}
\]
(45)
follows from straightforward manipulations based on (2) and (3); however, the presence of the summation with respect to \( s \) makes (45) of limited usefulness if \( n \) is large. The following theorem provides a more tractable formula.

**Theorem 7.** Let \( E_{\nu q}(r), 1 \leq \nu \leq p, \ r = 1, 2, \ldots \), be the determinant of the form (10) with \( Q_i(z) \) as in (14) if \( i \neq p + \nu \) and
\[
Q_{p+\nu}(z) = z^{n+2p+r}A^*(1/z).
\]
Let
\[ \Delta_r = \sum_{j=0}^{r} b_j \gamma_{j+p-r}, \]
and suppose that
\[ n \geq \max(0, 2q - 2p, q - 1). \]
Then the variance (44) is given by
\[ \sigma_{nk}^2 = \sum_{i=0}^{k-1} \Delta_i b_{i-p} - \frac{1}{D} \sum_{p=1}^{p} b_{k-p} \sum_{i=0}^{k-1} \Delta_i F_{i,n}(k-l). \]

**Proof.** Let
\[ e_r = \phi_{r+k} - \sum_{s=0}^{n} \phi_{r-s} \psi_{sn}^{(k)}, \]
and note that
\[ e_r = 0, \quad 0 \leq r \leq n, \]
because of (3), and, from (45),
\[ e_{-k} = \sigma_{nk}^2. \]
We will derive a triangular linear system of \( k \) equations in \( e_{-1}, \ldots, e_{-k} \), and solve it to obtain a tractable formula for \( \sigma_{nk}^2 \).
From (8), we can rewrite (49) as
\[ e_r = \sum_{j=-p}^{p} \gamma_j \left( \alpha_{r+k-j} - \sum_{s=0}^{n} \alpha_{r-s} \psi_{sn}^{(k)} \right). \]
With \( w_r \) as defined in (41) for all \( r \),
\[ \sum_{j=-p}^{p} \gamma_j w_{r-j} = 0, \quad -\infty < r < \infty; \]
therefore, we can rewrite (52) as
\[ e_r = \sum_{j=-p}^{p} \gamma_j \eta_{r-j}, \]
where
\[ \eta_r = \alpha_{r+k} - \sum_{s=0}^{n} \alpha_{r-s} \psi_{sn}^{(k)} + w_r. \]
From (24),
\[ \eta_r = 0, \quad -p \leq r \leq n + p. \]
To eliminate the sum with respect to \( s \) in (52), we introduce the quantities
\[ K_\mu = \sum_{l=0}^{q} \tilde{a}_l e_{l-\mu}, \quad 1 \leq \mu \leq k. \]
Substituting (53) into (56) and invoking (55) shows that

$$K_\mu = \sum_{j=p-\mu+1}^{p} \gamma_j \sum_{l=0}^{q} a_l \eta_{l-j-\mu}, \quad 1 \leq \mu \leq k,$$

(57)

if \( n \geq q - 1 \), as assumed in (47). From (54),

$$\sum_{l=0}^{q} a_l \eta_{l-r} = \sum_{l=0}^{q} a_l a_{l-r+k} - \sum_{s=0}^{n} \left( \sum_{l=0}^{q} a_l \alpha_{l-r-s} \right) \psi_{sn}^{(k)} + \sum_{l=0}^{q} a_l w_{l-r}.$$  

(58)

Because of (25), (58) can be rewritten as

$$\sum_{l=0}^{q} a_l \eta_{l-r} = b_{k-r} - \sum_{s=0}^{n} b_{-r-s} \psi_{sn}^{(k)} + \sum_{l=0}^{q} a_l w_{l-r}.$$  

(59)

Since \( b_s = 0 \) if \( t < 0 \), the first sum on the right of (59) vanishes if \( r > 0 \). Therefore, since \( j+\mu \geq p+1 > 0 \) for the terms remaining in (57), we conclude that

$$K_\mu = \sum_{j=p-\mu+1}^{p} \gamma_j \Gamma_{j+\mu}, \quad 1 \leq \mu \leq k,$$

(60)

where

$$\Gamma_r = b_{k-r} + \sum_{l=0}^{q} a_l w_{l-r}.$$  

(61)

Having established (60) let us look again at (56), which can be rewritten as

$$K_\mu = \sum_{l=0}^{\mu-1} a_l e_{l-\mu} = \sum_{\nu=1}^{\mu} a_{\mu-\nu} e_{-\nu}, \quad 1 \leq \mu \leq k,$$

(62)

because of (50). This system in \( e_{-1}, \ldots, e_{-k} \) has the triangular Toeplitz matrix

$$\left( a_{\mu-\nu} \right)_{\mu, \nu=1}^{k}.$$  

(63)

It follows easily from (6) that the inverse of (63) is the triangular Toeplitz matrix

$$\left( b_{\mu-\nu} \right)_{\mu, \nu=1}^{k},$$

and therefore we can solve (62) to obtain

$$e_{-k} = \sum_{\nu=1}^{k} b_{k-r} K_\nu.$$  

This, (51) and (60) imply that

$$\sigma_{nk}^2 = \sum_{\nu=1}^{k} b_{k-r} \sum_{j=p-\nu+1}^{p} \gamma_j \Gamma_{j+\nu} = \sum_{\nu=1}^{k} \sum_{m=1}^{\nu} \gamma_{p+m-\nu} \Gamma_{p+1} \Gamma_{m}.$$  

(64)
Changing the order of summation and then reversing the first sum in the result enables us to rewrite (64) as

\[
\sigma_{nk}^2 = \sum_{m=1}^{k} \Delta_{k-m} \Gamma_{p+m} = \sum_{m=1}^{k} \Delta_k b_k \sigma_{p-m} + \sum_{m=1}^{k} \Delta_k \sum_{j=0}^{q} \alpha_j w_{j-p-m},
\]

(65)

(See (46) and (61).) To obtain (48) from (65), let \( l = k - m \) and recall (41), Lemma 2 and the definitions of \( D_{kn}(r) \) and \( F_{kn}(r) \). □

In particular, setting \( k = 1 \) in (65) and recalling that \( a_0 = 1/b_0 \) yields

\[
\sigma_{n1}^2 = \tilde{a}_0^{-1} \gamma_p \sum_{j=0}^{q} \tilde{a}_j w_{j-p-1},
\]

a result previously stated in [6]. On setting \( k = 1 \) in (48) and recalling (46) with \( r = 0 \), we obtain

\[
\sigma_{n1}^2 = -D_n^{-1} |b_0|^2 \gamma_p F_{1n}(1).
\]

Since \( F_{1n} = (-1)^p D_{n+1} \) (recall the definition of \( F_{kn}(m) \) in the statement of Theorem 7), this can be rewritten as

\[
\sigma_{n1}^2 = (-1)^p |a_0|^{-2} \gamma_p \left( \frac{D_{n+1}}{D_n} \right).
\]

References