Trinity University

From the SelectedWorks of William F. Trench

1964

An algorithm for the inversion of finite Toeplitz matrices

William F. Trench, *Trinity University*

Available at: https://works.bepress.com/william_trench/5/

J. Soc. Indust. Appl. Math.
Vol. 12, No. 3, September, 1964
Printed in U.S.A.

AN ALGORITHM FOR THE INVERSION OF FINITE TOEPLITZ MATRICES*

WILLIAM F. TRENCH†

1. Introduction. In this paper we consider the problem of the inversion of positive definite matrices of the form

$$T_n = (\phi_{r-s}) \qquad 0 \leqq r, s \leqq n, n \geqq 0, \tag{1.1}$$

where the sequence $\{\phi_j\}$, $-\infty < j < \infty$, is Hermitian and positive definite. That is

$$\phi_j = \bar{\phi}_{-j}, \tag{1.2}$$

and for any $n \geqq 0$, and nontrivial n-tuple $(y_0, y_1, \cdots, y_{n-1})$,

$$\sum_{r,s=0}^{n-1} \phi_{r-s} y_r \bar{y}_s > 0. \tag{1.3}$$

Matrices of the form (1.1) are called Toeplitz matrices. They play an important role in the theory of discrete random processes. For example, if $\{y_k\}$, $-\infty < k < \infty$, is a real gaussian stationary process with zero mean and variance σ^2 such that, for every n, the joint distribution of $(y_0, y_1, \cdots, y_n)$ is of rank $n+1$, then the autocorrelation sequence

$$\phi_j = \frac{E[y_k \cdot y_{k+j}]}{\sigma^2}$$

possesses properties (1.1), (1.2), and (1.3). In order to find the joint probability density function of $(y_0, y_1, \cdots, y_n)$, or of any $n + 1$ successive variates, it is necessary to invert the matrix T_n.

In this paper, we derive an exact recursive procedure for the numerical inversion of an arbitrary positive definite Toeplitz matrix of *finite* order, which takes full advantage of the strong restrictions placed on its elements by (1.1), (1.2), and (1.3). The number of multiplications required for the inversion of an nth order Toeplitz matrix, using this procedure, is proportional to n^2, rather than to n^3, as in the case of methods which are suitable for arbitrary Hermitian matrices. To the author's knowledge, this inversion algorithm is the first to be specifically designed to take advantage of the peculiar simplicity of the general Toeplitz matrix. In addition, the closing section of the paper is devoted to a statement of an algorithm for the inversion of non-Hermitian matrices of the form (1.1).

* Received by the editors February 22, 1963, and in revised form November 20, 1963.

† Radio Corporation of America, Moorestown, New Jersey. Presently affiliated with Drexel Institute of Technology, Philadelphia, Pa.

Siddiqui [1] and Wise [2] have previously obtained the inverses of Toeplitz matrices, in closed form, in the case where the sequence $\{\phi_j\}$ has a generating function of the form

$$\frac{1}{A(z)A(1/z)} = \sum_{j=-\infty}^{\infty} \phi_j z^j,$$

where $A(z)$ is a polynomial with real coefficients, with no roots on the unit circle. In the theory of stationary time series, it can be shown that $\{\phi_j\}$ is then the autocorrelation sequence of a pure autoregressive process [1]. Also, Calderon, Spitzer, and Widom [3] have considered infinite matrices of the form (1.1), and developed conditions for the existence of the inverse.

2. Computation of the inverse in the general case. Let $B_n = T_n^{-1}$, and b_{rsn}, $0 \leqq r, s \leqq n$, be a typical element of the former. From (1.2) it follows that

$$b_{rsn} = \bar{b}_{srn}.$$

In addition, B_n is symmetric about the secondary diagonal. That is,

$$b_{rsn} = b_{n-s,n-r,n},$$

a property which is called *persymmetry* by Wise [2].

We now derive a recursive procedure for expressing B_{n+1} in terms of B_n. From (1.2) and (1.3), ϕ_0 is real and positive. Hence, we can normalize with $\phi_0 = 1$, so that $b_{000} = 1$.

Assume that $n \geqq 0$. We can write

$$T_{n+1} = \begin{bmatrix} 1 & \bar{U}_n^T \\ U_n & T_n \end{bmatrix}, \quad \text{where} \quad U_n = \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_{n+1} \end{bmatrix}$$

and where $\bar{U}_n^T$ is the conjugate transpose of U_n. We can also partition B_{n+1} in the form

$$B_{n+1} = \begin{bmatrix} b_{00,n+1} & \bar{W}_n^T \\ W_n & A_n \end{bmatrix}, \tag{2.1}$$

where $b_{00,n+1}$ is a scalar, W_n is an $(n + 1)$-dimensional column vector, and A_n is an $n+1$ by $n+1$ matrix, all to be determined by means of the equations

$$b_{00,n+1} + \bar{W}_n^T U_n = 1, \tag{2.2}$$

$$b_{00,n+1}\bar{U}_n^T + \bar{W}_n^T T_n = 0, \tag{2.3}$$

$$W_n + A_n U_n = 0, \tag{2.4}$$

$$W_n \bar{U}_n^T + A_n T_n = I. \tag{2.5}$$

Solving (2.3) yields

$$\bar{W}_n^T = -b_{00,n+1}\bar{U}_n^T B_n, \tag{2.6}$$

which can be substituted into (2.2) to yield

$$b_{00,n+1} = \Delta_n^{-1},$$

where

$$\Delta_n = 1 - \bar{U}_n^T B_n U_n .$$

The existence of B_n, and consequently of $b_{00,n+1}$, ensures that $\Delta_n \neq 0$. From (2.6),

$$\bar{W}_n^T = -\Delta_n^{-1}\bar{U}_n^T B_n .$$

Equation (2.5) can be solved to yield

$$A_n = B_n + \Delta_n^{-1} B_n U_n \bar{U}_n^T B_n .$$

In obtaining this solution, we have not used (2.4). This equation is not independent of (2.2), (2.3), and (2.5); but it is consistent with them. Substitution of these results into (2.1) yields

$$B_{n+1} = \begin{bmatrix} \Delta_n^{-1} & -\Delta_n^{-1}\bar{U}_n^T B_n \\ -\Delta_n^{-1} B_n U_n & B_n + \Delta_n^{-1} B_n U_n \bar{U}_n^T B_n \end{bmatrix}, \qquad n \geqq 0. \tag{2.7}$$

This relationship provides a recursive procedure for the calculation of $B_1, B_2, \cdots$. However, it has the disadvantage of requiring the computation of B_r for $r = 0, 1, \cdots, n$ in order to obtain B_{n+1}. Fortunately, it is possible to derive a simpler recursive procedure from (2.7), which we now proceed to do.

Define the vector Ψ_n by

$$\Psi_n = \begin{bmatrix} \psi_{0n} \\ \psi_{1n} \\ \vdots \\ \psi_{nn} \end{bmatrix} = B_n U_n ,$$

or, explicitly

$$\psi_{rn} = \sum_{s=0}^{n} b_{rsn}\phi_{s+1} . \tag{2.8}$$

By inspecting (2.7), we conclude that

$$\begin{aligned}
&\text{(a)} \quad b_{00,n+1} = \Delta_n^{-1}, \\
&\text{(b)} \quad b_{0s,n+1} = -\Delta_n^{-1}\bar{\psi}_{s-1,n}, && 1 \leqq s \leqq n+1, \\
&\text{(c)} \quad b_{r0,n+1} = -\Delta_n^{-1}\psi_{r-1,n}, && 1 \leqq r \leqq n+1, \\
&\text{(d)} \quad b_{rs,n+1} = b_{r-1,s-1,n} + \Delta_n^{-1}\psi_{r-1,n}\bar{\psi}_{s-1,n}, && 1 \leqq r, s \leqq n+1.
\end{aligned} \tag{2.9}$$

The property of Hermitian symmetry is evident from these equations. In addition, we have noted above that B_{n+1} is symmetric about the secondary diagonal, so that after some manipulation of indices, we can obtain the following relationships from (2.9):

$$
(2.10)\quad
\begin{aligned}
&\text{(a)}\quad b_{n+1,n+1,n+1} = \Delta_n^{-1}, && \\
&\text{(b)}\quad b_{r,n+1,n+1} = -\Delta_n^{-1}\bar{\psi}_{n-r,n}, && 0 \leqq r \leqq n, \\
&\text{(c)}\quad b_{n+1,s,n+1} = -\Delta_n^{-1}\psi_{n-s,n}, && 0 \leqq s \leqq n, \\
&\text{(d)}\quad b_{rs,n+1} = b_{rsn} + \Delta_n^{-1}\bar{\psi}_{n-r,n}\psi_{n-s,n}, && 0 \leqq r, s \leqq n.
\end{aligned}
$$

By comparing (2.9d) with (2.10d), we can conclude that

$$b_{rsn} = b_{r-1,s-1,n} + \Delta_n^{-1}[\psi_{r-1,n}\bar{\psi}_{s-1,n} - \bar{\psi}_{n-r,n}\psi_{n-s,n}], \qquad 1 \leqq r, s \leqq n.$$

By comparing (2.10a) with (2.9d) for $r = s = n + 1$, we find that

$$\Delta_n^{-1} = b_{nnn} + \Delta_n^{-1} \mid \psi_{nn} \mid^2,$$

so that

$$\Delta_n = (1 - \mid \psi_{nn} \mid^2)\Delta_{n-1}, \qquad n \geqq 0,$$

if we define $\Delta_{-1} = 1$. By comparing (2.9b) with (2.10d), we deduce that

$$b_{0sn} = -\Delta_n^{-1}[\bar{\psi}_{nn}\psi_{n-s,n} + \bar{\psi}_{s-1,n}], \qquad 1 \leqq s \leqq n.$$

To summarize the results to this point, we have shown that Δ_n and the elements of B_n can be expressed completely in terms of the quantities $\Delta_{n-1}, \psi_{0n}, \cdots, \psi_{nn}$ by means of the equations

$$
(2.11)\quad
\begin{aligned}
&\text{(a)}\quad \Delta_n = (1 - \mid \psi_{nn} \mid^2)\Delta_{n-1}, && \\
&\text{(b)}\quad b_{00n} = \Delta_{n-1}^{-1}, && \\
&\text{(c)}\quad b_{0sn} = -\Delta_n^{-1}[\bar{\psi}_{nn}\psi_{n-s,n} + \bar{\psi}_{s-1,n}], && 1 \leqq s \leqq n, \\
&\text{(d)}\quad b_{r0n} = -\Delta_n^{-1}[\psi_{nn}\bar{\psi}_{n-r,n} + \psi_{r-1,n}], && 1 \leqq r \leqq n, \\
&\text{(e)}\quad b_{rsn} = b_{r-1,s-1,n} + \Delta_n^{-1}[\psi_{r-1,n}\bar{\psi}_{s-1,n} - \bar{\psi}_{n-r,n}\psi_{n-s,n}], && \\
& && 1 \leqq r, s \leqq n.
\end{aligned}
$$

In an actual computation, an economy can be effected by means of the following procedure.

(i) Employ (2.11a) and (2.11b) to compute the zeroth row of B_n.

(ii) Employ (2.11e) to compute the rest of the elements lying above both diagonals, or on the diagonals.

(iii) Obtain the remaining elements above or on the secondary diagonal by means of Hermitian symmetry.

(iv) Obtain the elements below the secondary diagonal by means of the symmetry about this diagonal.

We now derive recursion formulas for the computation of the vectors $\Psi_1, \Psi_2, \cdots$, which do not require the calculation of $B_1, B_2, \cdots$. From (2.7), with n replaced by $n-1$,

$$B_n = \begin{bmatrix} \Delta_{n-1}^{-1} & -\Delta_{n-1}^{-1}\bar{\Psi}_{n-1}^T \\ -\Delta_{n-1}^{-1}\Psi_{n-1} & B_{n-1} + \Delta_{n-1}^{-1}\Psi_{n-1}\bar{\Psi}_{n-1}^T \end{bmatrix}, \qquad n \geqq 0. \tag{2.12}$$

Define

$$V_n = \begin{bmatrix} \bar{\phi}_{n+1} \\ \bar{\phi}_n \\ \vdots \\ \bar{\phi}_1 \end{bmatrix}, \qquad n \geqq 0.$$

From (2.12)

$$B_n V_n = \begin{bmatrix} \Delta_{n-1}^{-1}(\bar{\phi}_{n+1} - \bar{\Psi}_{n-1}^T V_{n-1}) \\ B_{n-1} V_{n-1} - \Delta_{n-1}^{-1}(\bar{\phi}_{n+1} - \bar{\Psi}_{n-1}^T V_{n-1})\Psi_{n-1} \end{bmatrix}. \tag{2.13}$$

The rth component of $B_n V_n$ is given by

$$\begin{aligned}(B_n V_n)_r &= \sum_{s=0}^{n} b_{rsn}\bar{\phi}_{n+1-s} = \sum_{s=0}^{n} b_{r,n-s,n}\bar{\phi}_{s+1} \\ &= \sum_{s=0}^{n} b_{s,n-r,n}\bar{\phi}_{s+1} = \sum_{s=0}^{n} \bar{b}_{n-r,s,n}\bar{\phi}_{s+1} = \bar{\psi}_{n-r,n},\end{aligned}$$

where we obtain the last equality by inspection of (2.8). By taking conjugates on both sides of (2.13), writing the result in component form, and using the last identity with r replaced by $n-r$, we find that

$$\text{(2.14)} \quad \begin{aligned} &\text{(a)} \quad \psi_{rn} = \psi_{r,n-1} - q_n\bar{\psi}_{n-r-1,n-1} \qquad 0 \leqq r \leqq n-1, \\ &\text{(b)} \quad \psi_{nn} = q_n, \end{aligned}$$

where

$$q_n = \Delta_{n-1}^{-1}(\phi_{n+1} - \Psi_{n-1}^T\bar{V}_{n-1}) = \Delta_{n-1}^{-1}\left(\phi_{n+1} - \sum_{s=0}^{n-1}\psi_{s,n-1}\phi_{n-s}\right), \tag{2.15}$$

and, recalling (2.11a),

$$\Delta_n = (1 - |q_n|^2)\Delta_{n-1}. \tag{2.16}$$

Equations (2.14), (2.15), and (2.16) provide a recursive means for the calculation of $\Delta_n, \psi_{0n}, \cdots, \psi_{nn}$ in terms of $\Delta_{n-1}, \psi_{0,n-1}, \cdots, \psi_{n-1,n-1}$ and $\phi_1, \phi_2, \cdots, \phi_{n+1}$, for $n \geqq 1$. For starting conditions,

$$\psi_{00} = \phi_1, \qquad \Delta_{-1} = 1. \tag{2.17}$$

Thus, if it is required to obtain B_n for some positive n, it is only necessary to use (2.14) through (2.17) for $0 \leqq r \leqq n$, and then B_n can be obtained from (2.11). However, the relationships expressed by (2.11) are still somewhat awkward, because they seem to indicate that the inverse of T_n depends upon ϕ_{n+1}, by virtue of the dependence of $\psi_{0n}, \cdots, \psi_{nn}$ on the latter. We will now derive a new set of relationships, similar to (2.11), which does not suffer from this defect. First we define, for every n, $\psi_{-1,n} = -1$ and $\psi_{r,n} = 0$ if $r > n$. With this extended definition, we write (2.14) as

$$\psi_{rn} = \psi_{r,n-1} - \psi_{nn}\bar{\psi}_{n-r-1,n-1}, \qquad r \geqq -1. \tag{2.18}$$

If in addition, we now define $b_{rsn} = 0$ if r or s is negative, we can rewrite (2.11) more compactly as

$$b_{rsn} = b_{r-1,s-1,n} + \Delta_n^{-1}[\psi_{r-1,n}\bar{\psi}_{s-1,n} - \bar{\psi}_{n-r,n}\psi_{n-s,n}], \qquad 0 \leqq r, s \leqq n. \tag{2.19}$$

Finally, this can be rewritten in the form

$$b_{rsn} = b_{r-1,s-1,n} + \Delta_{n-1}^{-1}[\psi_{r-1,n-1}\bar{\psi}_{s-1,n-1} - \bar{\psi}_{n-r,n-1}\psi_{n-s,n-1}], \tag{2.20}$$

in which none of the terms are dependent on ϕ_{n+1}. The equivalence of (2.19) and (2.20) can be verified by substituting from (2.18) into (2.19), and applying (2.11a).

For the reader's convenience, we will now collect the formulas which define the algorithm, assuming that we wish to invert T_n.

$$\psi_{00} = \phi_1, \qquad \Delta_{-1} = 1.$$

$$\Delta_{m-1} = (1 - |\psi_{m-1,m-1}|^2)\Delta_{m-2}.$$

$$\psi_{mm} = \Delta_{m-1}^{-1}\Big(\phi_{m+1} - \sum_{s=0}^{m-1}\psi_{s,m-1}\phi_{m-s}\Big).$$

$$\psi_{rm} = \psi_{r,m-1} - \psi_{mm}\bar{\psi}_{m-r-1,m-1}, \qquad 0 \leqq r \leqq m-1.$$

The last three formulas are used for $1 \leqq m \leqq n-1$. To obtain B_n, compute as follows.

$$b_{00n} = \Delta_{n-1}^{-1}.$$

$$b_{r0n} = -\Delta_{n-1}^{-1}\psi_{r-1,n-1}, \qquad 1 \leqq r \leqq n.$$

$$b_{rsn} = b_{r-1,s-1,n} + \Delta_{n-1}^{-1}(\psi_{r-1,n-1}\bar{\psi}_{s-1,n-1} - \bar{\psi}_{n-r,n-1}\psi_{n-s,n-1}).$$

The last two formulas can be used to compute the elements which lie in the triangular section of B_n which is bounded by the zeroth column and the two diagonals (including points on the boundary). The rest of the elements can be obtained from these by exploiting the Hermitian symmetry about the principal diagonal, and the symmetry about the secondary diagonal.

It is also of interest to compute the determinant of T_n, which we denote by $\| T_n \|$. In (2.7), take determinants, multiply the first row of $\| B_{n+1} \|$ by ψ_{rn}, and add the result to the $(r+1)$th row, $0 \leqq r \leqq n$, to obtain

$$\| B_{n+1} \| = \begin{Vmatrix} \Delta_n^{-1} & -\Delta_n^{-1}\bar{U}_n^T B_n \\ 0 & B_n \end{Vmatrix} = \Delta_n^{-1} \| B_n \|, \qquad n \geqq 0,$$

so that

$$\| T_{n+1} \| = \prod_{r=0}^{n} \Delta_r \qquad n \geqq 0;$$

$$\| T_0 \| = 1.$$

In addition to their importance in the computation of the inverse matrix B_n, the vectors $\{\Psi_n\}$ play an important role in the problem of linear prediction of a stationary time series. A detailed discussion of this problem is given by Yaglom [4, pp. 97–103]. The result of interest can be described as follows. Let $\{y_k\}$ be a complex stationary time series with autocorrelation sequence

$$\phi_j = E[\bar{y}_k y_{j+k}].$$

Then, if we wish to predict the value of y_{k+1}, given $y_k, y_{k-1}, \cdots, y_{k-n}$, it can be shown that the best linear estimator of y_{k+1} is given by

$$\tilde{y}_{k+1} = \sum_{r=0}^{n} \psi_{rn} y_{k-r}, \tag{2.21}$$

in the sense that of all linear combinations of $y_k, y_{k-1}, \cdots, y_{k-n}$, (2.21) is the one for which the variance $E(| \tilde{y}_{k+1} - y_{k+1} |^2)$ is minimized.

3. Generalization to the non-Hermitian case. The algorithm of §2 can be generalized to the case where T_n is of the form (1.1) but not necessarily Hermitian. The algorithm described here can be employed to invert T_n if each of the matrices $T_0, T_1, \cdots, T_n$ is nonsingular. The derivation of the algorithm follows essentially the same lines as in the Hermitian case and will be omitted.

For the non-Hermitian case we consider, for each m, two $(m+1)$-dimensional vectors, Ψ_m and H_m, defined by

$$T_m \Psi_m = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{m+1} \end{bmatrix} \quad \text{and} \quad T_m^T \mathrm{H}_m = \begin{bmatrix} \phi_{-1} \\ \phi_{-2} \\ \vdots \\ \phi_{-(m+1)} \end{bmatrix}.$$

The components of these two sequences of vectors satisfy the recursion formulas

$$\psi_{rm} = \psi_{r,m-1} - \psi_{mm}\eta_{m-r-1,m-1}, \qquad 0 \leqq r \leqq m-1,$$

$$\eta_{rm} = \eta_{r,m-1} - \eta_{mm}\psi_{m-r-1,m-1}, \qquad 0 \leqq r \leqq m-1,$$

$$\psi_{mm} = \Delta_{m-1}^{-1}\left[\phi_{m+1} - \sum_{s=0}^{m-1}\psi_{s,m-1}\phi_{m-s}\right],$$

$$\eta_{mm} = \Delta_{m-1}^{-1}\left[\phi_{-(m+1)} - \sum_{s=0}^{m-1}\eta_{s,m-1}\phi_{-(m-s)}\right],$$

$$\Delta_{m-1} = (1 - \psi_{m-1,m-1}\eta_{m-1,m-1})\Delta_{m-2},$$

with starting conditions

$$\psi_{00} = \phi_1, \qquad \eta_{00} = \phi_{-1}, \qquad \Delta_{-1} = 1.$$

We have assumed, without loss of generality, that $\phi_0 = 1$.

In order to compute the inverse of T_n, these computations are carried out for $m = 1, \cdots, n-1$, and the elements of B_n are given by

$$b_{00n} = \Delta_{n-1}^{-1},$$

$$b_{r0n} = -\Delta_{n-1}^{-1}\psi_{r-1,n-1}, \qquad 1 \leqq r \leqq n,$$

$$b_{0sn} = -\Delta_{n-1}^{-1}\eta_{s-1,n-1}, \qquad 1 \leqq s \leqq n,$$

$$b_{rsn} = b_{r-1,s-1,n} + \Delta_{n-1}^{-1}[\psi_{r-1,n-1}\eta_{s-1,n-1} - \eta_{n-r,n-1}\psi_{n-s,n-1}], \quad 1 \leqq r,\ s \leqq n.$$

It is only necessary to use the last formula for $r + s \leqq n$, since the remaining elements can be obtained from the relationship $b_{rsn} = b_{n-s,n-r,n}$ (reflection about the secondary diagonal).

4. Acknowledgment. I extend my thanks to Dr. Perry Scheinok for his helpful comments on the presentation of this work. I also thank the referee for drawing my attention to the connection between the vectors $\{\Psi_n\}$ and the prediction problem.

REFERENCES

[1] M. Siddiqui, *On the inversion of the sample covariance matrix in a stationary autoregressive process*, Ann. Math. Statist., 29(1958), pp. 585–588.

[2] J. Wise, *The autocorrelation function and the spectral density function*, Biometrika, 42(1955), pp. 151–159.

[3] A. Calderõn, F. Spitzer, and H. Widom, *Inversion of Toeplitz matrices*, Illinois J. Math., 3(1959), pp. 490–498.

[4] A. M. Yaglom, *An Introduction to the Theory of Stationary Random Functions*, revised English edition, (transl., R. A. Silverman), Prentice-Hall, Englewood Cliffs, New Jersey, 1962.