Solution of Linear Equations With Rational Toeplitz Matrices*

By Bradley W. Dickinson

Abstract. We associate a sequence of Toeplitz matrices with the rational formal power series \( T(z) \). An algorithm for solving linear equations with a Toeplitz matrix from this sequence is given. The algorithm requires \( O(n) \) operations to solve a set of \( n \) equations, for \( n \) sufficiently large.

1. Introduction. In this paper, we present an algorithm for solving a system of \( N + 1 \) linear equations with special Toeplitz structure:

\[
T_N x = y.
\]

For every \( n > 0 \), let \( T_n = \{ t_{i,j}, 0 \leq i, j \leq n \} \) be a Toeplitz matrix. We assume that \( T_n^{-1} \) exists for \( 0 \leq n \leq N \) and that the formal Laurent series

\[
T(z) = \sum_{k=-\infty}^{\infty} t_k z^k
\]

is a rational function of \( z \). Then, for sufficiently large \( N \), our algorithm requires \( O(N) \) operations to compute the solution to (1). The algorithms of Levinson [5], Bareiss [1], and Zohar [13], which exploit only the Toeplitz structure of \( T_N \), require \( O(N^2) \) operations to solve (1).

This problem is motivated by an important special case of (1) arising in linear least squares estimation theory. When \( T(z) \) is rational and matrices \( T_n \) are symmetric and positive definite for all \( n \), \( T_N \) is the covariance matrix of \( N + 1 \) samples from a wide-sense stationary autoregressive moving-average stochastic process. Trench [10], in a somewhat overlooked paper, outlined an algorithm for solving the linear equations associated with certain estimation problems; his algorithm requires \( O(N) \) operations but, as noted in [10], the details are "tedious to write out" except in the banded case when \( T(z) \) in (2) is a finite series. For the banded case, alternative algorithms can be developed using the result that the (banded) Cholesky factors of \( T_N \) can be obtained in \( O(N) \) operations; see Morf [6] and Rissanen [7].

Coupling Trench's work on inversion of nonsymmetric banded Toeplitz matrices [11] with Zohar's results [13] leads to an efficient algorithm for solution of general banded Toeplitz systems in \( O((p + q)N) \) operations, where \( t_i = 0 \) for \( i > p \) and \( i < -q \)
In this paper we will generalize the solution algorithms of [3] and [13] to a class of "almost-Toeplitz" systems, leading to an algorithm for solving the special Toeplitz systems defined above.

2. Solution of the Toeplitz System. We first present a decomposition of $T_n$ which follows from the rationality of $T(z)$. By separately considering the upper and lower triangular parts of $T_n$ we obtain a representation in terms of banded, triangular Toeplitz matrices. Let

\begin{equation}
T(z) = T_+(z) + T_-(z),
\end{equation}

where by rationality

\begin{align}
T_+(z) &= \frac{t_0}{2} + \sum_{k=1}^{\infty} t_k z^k = c(z) / d(z), \\
T_-(z) &= \frac{t_0}{2} + \sum_{k=1}^{\infty} t_{-k} z^{-k} = \gamma(z) / \delta(z),
\end{align}

and $c(z)$, $d(z)$, $\gamma(z)$ and $\delta(z)$ are polynomials given by

\begin{align}
c(z) &= \sum_{i=0}^{p} c_i z^i, \quad c_p \neq 0, \\
d(z) &= 1 + \sum_{i=1}^{q} d_i z^i, \quad d_q \neq 0, \\
\gamma(z) &= \sum_{i=0}^{r} \gamma_i z^{-i}, \quad \gamma_r \neq 0, \\
\delta(z) &= 1 + \sum_{i=1}^{s} \delta_i z^{-i}, \quad \delta_s \neq 0.
\end{align}

For notational convenience, we write $L_n(w)$ for the lower triangular Toeplitz matrix whose first column is the $(n + 1)$-vector $w$; $U_n(w)$ is the upper triangular Toeplitz matrix whose first row is $w'$, where prime denotes transpose. We define $(n + 1)$-vectors of the coefficients of the polynomials in (5) by

\begin{align}
c_n &= (c_0 c_1 \cdots c_p 0 \cdots 0)', \\
d_n &= (1 d_1 \cdots d_q 0 \cdots 0)', \\
\gamma_n &= (\gamma_0 \gamma_1 \cdots \gamma_r 0 \cdots 0)', \\
\delta_n &= (1 \delta_1 \cdots \delta_s 0 \cdots 0)'.
\end{align}

These vectors are suitably truncated when $n$ is less than $p$, $q$, $r$, or $s$.

The desired representation of $T_n$ follows from the natural isomorphism between the ring of formal power series in $z$ and the ring of semi-infinite (towards the southeast) lower triangular Toeplitz matrices; the coefficient of $z^0$ is associated with the
diagonal element, the coefficient of $z$ with the first subdiagonal element, etc. Polynomials in $z$ correspond to banded matrices, and power series multiplication corresponds to matrix multiplication. Similarly, power series in $z^{-1}$ are naturally isomorphic to semi-infinite (towards the northwest) upper triangular Toeplitz matrices with polynomials in $z^{-1}$ corresponding to banded matrices. In both cases, a power series with a nonzero coefficient of $z^0$ is invertible in the ring; this corresponds to the fact that an invertible triangular Toeplitz matrix has a triangular Toeplitz inverse. (In the finite case, Traub [8] has given an expression for the (Toeplitz) inverse of a triangular Toeplitz matrix.)

Applying the isomorphisms to the power series equations (4a) and (4b) and taking the first $n + 1$ rows and columns of the corresponding matrix products, starting at the northwest and southeast corners, respectively, and combining the lower and upper triangular Toeplitz matrices gives the desired representation of $T_n$.

**Lemma 1.** With the notation defined above, for $n > 0$

\[
T_n = L_n^{-1}(d_n)L_n(c_n) + U_n(\gamma_n)U_n^{-1}(\delta_n).
\]

Since power series multiplication is commutative, we have chosen a convenient ordering of the factors. Now, treating (7) simply as a matrix identity for the class of Toeplitz matrices considered here, we see that $T_n$ can be reduced to a band matrix by cross multiplication, giving

\[
R_n = L_n(d_n)T_nU_n(\delta_n) = L_n(c_n)U_n(\delta_n) + L_n(d_n)U_n(\gamma_n).
\]

Since $T_n$ is Toeplitz, $L_n(d_n)$ is lower triangular, and $U_n(\delta_n)$ is upper triangular, the first equality in (8) shows that for $n > 1$, $R_{n-1}$ is the $n$ by $n$ principal submatrix of $R_n$. However, using the second equation in (8) and the Toeplitz structure of the triangular matrices, we obtain the following important structural property of $R_n$.

**Lemma 2.** For $n > 1$, the matrix $R_n$ defined in (8) satisfies

\[
R_n = \begin{bmatrix}
0 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & R_{n-1}
\end{bmatrix} = c_n\delta_n' + d_n\gamma_n'.
\]

From (6a)-(6d), the nonzero elements of the matrix on the right-hand side of (9) lie in the northwest corner but generally extend beyond the first row and column, so $R_n$ differs from Toeplitz only in its northwest corner. When the degrees $p = r = 0$ so that $c_n' = \gamma_n' = (1 \ 0 \ \cdots \ 0)$, $R_n$ is Toeplitz.

A general theory for the inversion of matrices which can be expressed as sums of products of lower and upper triangular Toeplitz matrices is given by Friedlander et al. [4]. Efficient, recursive algorithms for determining the inverses of such matrices are derived, generalizing the Trench algorithm [9], [12]. The particularly simple form of (9) leads to additional simplifications of the approach in [4]. Furthermore, as in [11], the band structure of $R_n$ may be exploited to reduce the computational complexity by an additional order of magnitude.
We propose a three step approach to solving the original system of equations (1):

(a) Compute \( L_N(d_N)y = \tilde{y} \).

(b) Solve \( R_N \tilde{x} = \tilde{y} \).

(c) Compute \( x = U_N(\delta_N)\tilde{x} \).

Thus in the following, we only describe an algorithm for the second step. Some additional notation will be required. We define for each \( n \geq 0 \) the vector

\[
A_n = (a_{n0} \cdots a_{nn})', \quad a_{nn} = 1
\]

as the solution to the system

\[
D \cdot R_n A_n = (0 \cdots 0 \alpha_n)',
\]

where the scalars \( \{\alpha_n; n \geq 0\} \) are defined recursively below. Similarly, vectors \( P_n \) and \( Q_n \) are defined by

\[
R_n P_n = c_n, \\
R_n Q_n = d_n.
\]

Next we let

\[
\tilde{y}_n = (\tilde{y}_0, \ldots, \tilde{y}_n)'
\]

and define \( X_n \) by

\[
R_n X_n = \tilde{y}_n.
\]

The matrix \( R_n \) has elements \( \{r_{ij}; 0 \leq i, j \leq n\} \). Now we are ready to derive our major result. We proceed in the usual way, obtaining the quantities \( A_{n+1}, P_{n+1}, Q_{n+1}, \) and \( X_{n+1} \) from \( A_n, P_n, Q_n, \) and \( X_n \). Using the structure of \( R_{n+1}, (9) \), we find

\[
R_{n+1} [0 A_n']' = (0 \cdots 0 \alpha_n)' + c_{n+1} e_n + d_{n+1} f_n,
\]

where the first term is an \((n + 1)-\)vector and the scalars \( e_n \) and \( f_n \) are given by

\[
e_n = [0 A_n'] \delta_{n+1},
\]

\[
f_n = [0 A_n'] \gamma_{n+1}.
\]

Since \( R_n \) is a principal submatrix of \( R_{n+1} \), we obtain

\[
R_{n+1} [P_n' 0]' = [c_n' g_n]', \\
R_{n+1} [Q_n' 0]' = [d_n' h_n]', \\
R_{n+1} [X_n' 0]' = [\tilde{y}_n' \Delta_n]'.
\]

where the scalars \( g_n, h_n, \) and \( \Delta_n \) are given by

\[
g_n = [r_{n+1,0} \cdots r_{n+1,n}] P_n, \\
h_n = [r_{n+1,0} \cdots r_{n+1,n}] Q_n.
\]
We let $c_{n+1}$ and $d_{n+1}$ denote the last elements of the vectors $c_n$ and $d_n$, respectively. Then $P_n$ and $Q_n$ can be used to update $A_n$:

$$
A_{n+1} = [0 A'_n]' - [P'_n 0]'e_n - [Q'_n 0]'f_n,
$$

$$
\alpha_{n+1} = \alpha_n + (c_{n+1} - g')e_n + (d_{n+1} - h)f_n.
$$

Now that $A_{n+1}$ is available, it may be used to update the values of $X_n$, $P_n$, and $Q_n$ so that (12), (13), and (15) are satisfied. The required steps are given by

$$
X_{n+1} = [X'_n 0]' - A_{n+1}(\Delta_n - \tilde{y}_{n+1})/\alpha_{n+1},
$$

$$
P_{n+1} = [P'_n 0]' - A_{n+1}(\xi_n - c_{n+1})/\alpha_{n+1},
$$

$$
Q_{n+1} = [Q'_n 0]' - A_{n+1}(h_n - d_{n+1})/\alpha_{n+1}.
$$

This completes the updating calculations.

The initial conditions for the algorithm are quite simple:

$$
A_0 = 1, \quad \alpha_0 = r_{00}, \quad Q_0 = 1/r_{00}, \quad X_0 = y_0/r_{00}, \quad P_0 = c_0/r_{00},
$$

where $c_0$ is obtained from (5a). In verifying the correctness of this algorithm, only the division by $\alpha_{n+1}$ at each stage requires additional justification. Here the assumption that $T_{n}^{-1}$ exists for every $0 \leq n \leq N$ is used. From the first equation of (8), $R_{n}^{-1}$ exists for every $0 \leq n \leq N$ because $U_n$ and $L_n$ are upper triangular matrices. Since $R_n$ is a principal submatrix of $R_{n+1}$, from Eqs. (10) and (11), $\alpha_{n+1} = \det R_{n+1}/\det R_n$; and this justifies the divisions required in the algorithm.

No use of the banded structure of $R_n$ has yet been made; the algorithm of Theorem 1 applies to any matrix $R_N$ having the structure in (9) and with $R_{n}^{-1}$ defined for each $n$. This includes some Toeplitz matrices, for example. With $\gamma_n = c_n = (1 0 \cdots 0)'$, $f_n = 0$ in (17b) for all $n$ and $Q_n$ in (13) is not required so the algorithm reduces to the Levinson-Trench-Zohar algorithm [13]. To exploit the banded nature of $R_N$, we make a minor assumption that $\rho = \max(p, q)$ is the lower bandwidth of $R_N$; that is we assume $r_{p+j,j} = 0$ and $r_{p+k,j} = 0$ for $k > j$. This is not a limitation because of (5), (6) and (8)

$$
r_{p+j,j} = \left\{ \begin{array}{ll}
\gamma_0 d_q = t_0 d_q/2 \neq 0 & \text{for } q > p, \\
c_p \neq 0 & \text{for } p > q, \\
c_p + \gamma_0 d_q & \text{for } p = q,
\end{array} \right.
$$

so this condition can be assured by modifying the fraction of the constant term to that which is assigned to $T(z)$ in (4b) if necessary. Some observations now follow directly:

(a) Computing (17a) and (17b) requires only the first $s$ and $r$ components of $A_n$, respectively. Let $\sigma = \max(s, r)$; $\sigma$ will ordinarily be the upper bandwidth of $R_N$.

(b) Only the last $\rho$ elements of $P_n$, $Q_n$, and $X_n$ are needed to compute (19a)–(19c).

(c) Consequently, in (20), (22b), (22c) only the first $\sigma$ and last $\rho$ elements of
$A_{n+1}$, $Q_{n+1}$, and $P_{n+1}$ need to be computed for $n$ larger than $\rho + \sigma$. In (22a) only the last $\rho$ elements of $X_{n+1}$ need to be computed. When $n + 1$ reaches $N$, the remaining elements of $X_N$ are computed by back substitution.

We define

$$X_N = (X_{N0} \cdots X_{NN})'.$$

Then for $N - \rho > j > 0$ we take

$$X_N^j = (1/r_{j+\rho,i}) \left( \tilde{V}_{j+\rho} - \sum_{i=j+1}^{j+\rho+\sigma} r_{j+\rho,i} X_{Ni} \right),$$

where $X_{Ni} = 0$ for $i > N$.

Together with the algorithms of Theorem 1, these modifications provide an algorithm for solving $R_N \tilde{X} = \tilde{Y}$; as discussed earlier, this is the only nontrivial step in the solution of (1) when $T_N$ is rational. An operation count (of multiplications) shows that solution of (1) requires $(10\rho + 5\sigma + 6)N + O((\rho + \sigma)^2)$ operations. Notice that because $R_N$ is Toeplitz except in its upper $(\rho + 1)$ by $(\sigma + 1)$ corner, all of its elements can be computed in $O((\rho + \sigma)^2)$ operations. This is still true if $T(z)$ is given in factored form

$$T(z) = (b(z)/d(z))(\beta(z)/\delta(z))$$

as is often the case in applications such as the linear estimation problems considered by Trench [10].

3. Discussion. Our algorithm differs from Trench's [10] in the following way. By extracting triangular Toeplitz factors of known form from $T_N$, namely $L_N^{-1}(d_N)$ and $L_N^{-1}(\delta_N)$, we are left with a banded nearly-Toeplitz system to solve. It appears that Trench removes nearly-Toeplitz factors from $T_N$ in order to be left with a banded Toeplitz system to solve. His motivation for so doing was the availability of an efficient algorithm for such systems. We have shown that a very similar algorithm can be used to solve the banded nearly-Toeplitz system.

If the rational power series $T(z)$ converges for some annulus centered on the origin in the complex plane, then subject to some minor assumptions, the existence of $T_n^{-1}$ for $0 \leq n \leq N$ can be expressed as a constraint on the poles and zeros of $T(z)$. The additional assumptions are that with $T(z) = N(z)/D(z)$ for relatively prime polynomials $N(z)$ and $D(z)$, $N(0) \neq 0$ and $N(z)$ has distinct zeros. Under these circumstances, Day [2] gives an explicit formula for the determinant of $T_n$ in terms of the zeros of $N(z)$ and $D(z)$, and a nonzero determinant is equivalent to the invertibility of $T_n$.

Department of Electrical Engineering and Computer Science
Princeton University
Princeton, New Jersey 08540


