

Algorithms for Triangular Decomposition of Block Hankel and Toeplitz Matrices with Application to Factoring Positive Matrix Polynomials

By J. Rissanen

Abstract. Algorithms are given for calculating the block triangular factors $A, \hat{A}, B = A^{-1}$ and $\hat{B} = \hat{A}^{-1}$ and the block diagonal factor D in the factorizations $R = AD\hat{A}$ and $BR\hat{B} = D$ of block Hankel and Toeplitz matrices R . The algorithms require $O(p^3n^2)$ operations when R is an $n \times n$ -matrix of $p \times p$ -blocks.

As an application, an iterative method is described for factoring $p \times p$ -matrix valued positive polynomials $R = \sum_{i=-m}^m R_i x^i, R_{-i} = R_i'$, as $\hat{A}(x)\hat{A}'(x^{-1})$, where $\hat{A}(x)$ is outer.

1. Introduction. Let $R = \{R_{ij}\}, i, j = 0, 1, \dots, n$, be a matrix with the entries R_{ij} as $p \times p$ -matrices of real valued elements. Such a matrix is called a (block) Hankel matrix if $R_{ij} = R_{i+j}$, and a (block) Toeplitz matrix if $R_{ij} = R_{j-i}$.

When all the block $(m+1) \times (m+1)$ -sections $\{R_{ij}\}, i, j = 0, 1, \dots, m, m \leq n$, of either a Hankel matrix R or a Toeplitz matrix R are invertible, we shall construct the two factorizations,

$$(1.1) \quad R = AD\hat{A},$$

$$(1.2) \quad BR\hat{B} = D,$$

where A is a lower triangular matrix with $p \times p$ -identity matrices I on the diagonal, \hat{A} is an upper triangular matrix with matrices I on the diagonal, D is an invertible block diagonal matrix, and B and \hat{B} are inverses of A and \hat{A} , respectively. The algorithms for finding the two triangular decompositions require $O(p^3n^2)$ arithmetic operations when R is either a Hankel matrix or a Toeplitz matrix.

Special cases of our algorithms have been derived earlier for different purposes. For $p = 1$ and R a positive definite Toeplitz matrix, Levinson [1] has derived such an algorithm for solving predictor problems, which, in effect, also finds the factors (1.2). A generalization of the same algorithm for $p > 1$ was derived (slightly imprecisely) in [2]. For $p = 1$ and R , a Hankel matrix, an algorithm for finding the factors in (1.1) was derived in [3] by use of the so-called moments and the Lanczos algorithm.

The algorithm for finding the factorization (1.1) for R , a Toeplitz matrix, is believed to be completely new; not even special cases of it seem to have been found before. In the particular case where R is symmetric and positive definite, this algorithm also finds the related Cholesky factorization, $R = \tilde{A}\tilde{A}'$ [6].

Received October 22, 1971, revised April 15, 1972.

AMS (MOS) subject classifications (1970). Primary 65F30; Secondary 65F25, 15A57.

Key words and phrases. Hankel and Toeplitz matrices, triangular decomposition of matrices.

Copyright © 1973, American Mathematical Society

The algorithm for the Cholesky factors has an important application in a numerical solution of the classical spectral factorization problem. This problem, with applications in prediction theory, representation or identification of random processes, and a number of other areas as well, is to factor a $p \times p$ -matrix polynomial $R(x) = \sum_{i=-m}^m R_i x^i$, $R_i = R'_{-i}$, positive for each $|x| = 1$, as follows:

$$(1.3) \quad R(x) = \bar{A}(x)\bar{A}'(x^{-1}),$$

where the factor $\bar{A}(x)$ has the form $\sum_{i=0}^m \bar{A}_i x^i$ and is "outer" [4]. In the present context, $\bar{A}(x)$ being outer merely means that it has the indicated form and is invertible for $|x| \leq 1$. Our algorithm requires an order of magnitude fewer arithmetic operations than the earlier ones in [5] and [7].

2. Hankel Matrices. We begin with the factorizations (1.1)–(1.2) when R is a (blockwise) $(n + 1) \times (n + 1)$ -matrix $\{R_{i+j}\}$ with all the block $m \times m$ -sections $m \leq n$, nonsingular. Consider the following portion of the matrix equation $BR = D\hat{A} = {}^\Delta P$:

$$(2.1) \quad \begin{pmatrix} I & & & & & \\ & B_{10} & I & & & \\ & \vdots & \ddots & & & \\ & \vdots & & & & \\ B_{k-1,0} & \cdots & I & & & \\ & B_{k0} & \cdots & B_{k,k-1} & I & \\ B_{k+1,0} & \cdots & & & B_{k+1,k} & I \end{pmatrix} \begin{pmatrix} R_0 & R_1 & \cdots & R_k \\ R_1 & R_2 & \cdots & R_{k+1} \\ \vdots & \vdots & & \vdots \\ R_{k-1} & R_k & \cdots & R_{2k-1} \\ R_k & R_{k+1} & \cdots & R_{2k} \\ R_{k+1} & R_{k+2} & \cdots & R_{2k+1} \end{pmatrix} = \begin{pmatrix} P_{00} & P_{01} & \cdots & P_{0k} \\ & \ddots & & \vdots \\ 0 & \cdots & 0 & P_{kk} \\ 0 & & \cdots & 0 \end{pmatrix}.$$

Observe the evident but important fact that these subfactorizations, one for each k , $k \leq n$, form a nested sequence such that a subfactorization for any k' , $k' \leq k$, is also a subfactorization of the one corresponding to k .

Denoting the i th block-row of R as the array (R_i, R_{i+1}, \dots) , we find that the $(k + 1)$ th block-row of B satisfies the equation and, conversely, is determined by it:

$$(2.2) \quad B_{k+1,0}(R_0, \dots, R_k) + \cdots + B_{k+1,k}(R_k, \dots, R_{2k}) + (R_{k+1}, \dots, R_{2k+1}) = 0.$$

By picking the block columns 1 to $k - 1$ from R , we also have

$$(2.3) \quad B_{k,0}(R_1, \dots, R_{k-1}) + \cdots + B_{k,k-1}(R_k, \dots, R_{2k-2}) + (R_{k+1}, \dots, R_{2k-1}) = 0.$$

With the notations,

$$(2.4) \quad \begin{aligned} E_0 &= B_{k+1,0}, \\ E_i &= B_{k+1,i} - B_{k,i-1}, \quad = 1, \dots, k, \end{aligned}$$

we deduce from (2.2) and (2.3), by subtraction, the two equations:

$$(2.5) \quad E_0(R_0, \dots, R_{k-2}) + \dots + E_k(R_k, \dots, R_{2k-2}) = 0, \quad k \geq 2,$$

and

$$(2.6) \quad B_{k+1,0}(R_{k-1}, R_k) + \dots + B_{k+1,k}(R_{2k-1}, R_{2k}) + (R_{2k}, R_{2k+1}) = 0, \quad k \geq 1.$$

Regard Eqs. (2.5) as p homogeneous equations in the p -rows of the block-row (E_0, \dots, E_k) . As the sections $\{R_{i+j}\}$, $i + j \leq 2k$, are nonsingular, we conclude by comparing (2.1) and (2.5) that every row of (E_0, \dots, E_k) must lie in the linear manifold spanned by the $2p$ linearly independent rows in the two block-rows $(B_{k-1,0}, \dots, I, 0)$ and $(B_{k,0}, \dots, I)$ of B . Hence, two matrices C_k and G_{k-1} exist such that

$$(E_0, \dots, E_k) = C_k(B_{k,0}, \dots, I) + G_{k-1}(B_{k-1,0}, \dots, I, 0).$$

The coefficient matrices C_k and G_{k-1} are determined by (2.6) in the following way: First, (2.4) and (2.6) imply

$$E_0(R_{k-1}, R_k) + \dots + E_k(R_{2k-1}, R_{2k}) + (P_{kk}, P_{k,k+1}) = 0.$$

Then, with the preceding equality, we derive the equality

$$C_k(0, P_{kk}) + G_{k-1}(P_{k-1,k-1}, P_{k-1,k}) + (P_{kk}, P_{k,k+1}) = 0$$

which leads to

$$(2.7) \quad \begin{aligned} G_{k-1} &= -D_k D_{k-1}^{-1}, & k \geq 1, \\ C_k &= -(G_{k-1} P_{k-1,k} + P_{k,k+1}) D_k^{-1}, & k \geq 0, \end{aligned}$$

where $D_k = P_{kk}$. Define $G_i = 0$, if i is negative, and $P_{ij} = 0$ and $B_{ij} = 0$, if i or j are negative. Then, we get the first recursion:

$$(2.8) \quad \begin{aligned} B_{k+1,i} &= B_{k,i-1} + C_k B_{ki} + G_{k-1} B_{k-1,i}, & k \geq 0, i = 0, \dots, k+1, \\ B_{00} &= I, \end{aligned}$$

which implies $B_{ii} = I$ for $i \geq 0$; we put $B_{ij} = P_{ji} = 0$ for $j > i$. Rather than calculate $P_{k,k+1}$ and D_k from (2.1) as sums, it is better to obtain them as by-products from a recursion for the elements P_{ij} as follows.

Equations (2.1) with (2.8) lead to

$$(2.9) \quad \begin{aligned} P_{jk} &= \sum_{i=0}^j B_{ji} R_{k+i} \\ &= P_{j-1,k+1} + C_{j-1} P_{j-1,k} + G_{j-2} P_{j-2,k}, & k > 0, j = 1, \dots, k, \\ P_{0k} &= R_k, & k \geq 0, \\ D_k &= P_{kk}. \end{aligned}$$

From P , the matrix \hat{A} can be recovered by

$$(2.10) \quad \hat{A} = D^{-1} P.$$

Equations (2.7), (2.9), and (2.10) generalize those in [3]. Equations (2.7)–(2.10) give both B and \hat{A} in $O(2n^2(p^3 + p^2))$ operations.

By transposing Eqs. (1.1) and (1.2) we have

$$(2.11) \quad R' = \hat{A}' D' A', \quad \hat{B}' R' B' = D'.$$

Therefore, in order to get the factors \hat{B}' and A' , we just replace R_i by R'_i in all the preceding equations.

3. Toeplitz Matrices. Consider again the following portion of the matrix equation $BR = D\hat{A} = {}^\Delta P$, where all the block $k \times k$ -sections of the Toeplitz matrix R are nonsingular:

$$(3.1) \quad \begin{pmatrix} I & & & & \\ B_{10} & I & & & \\ \vdots & & & & \\ B_{k0} & \cdots & B_{k,k-1} & I & \\ B_{k+1,0} & \cdots & & B_{k+1,k} & I \end{pmatrix} \begin{pmatrix} R_0 & R_1 & \cdots & R_k \\ R_{-1} & R_0 & R_1 & R_{k-1} \\ \vdots & \vdots & \vdots & \vdots \\ R_{-k} & \cdots & & R_0 \\ R_{-k-1} & \cdots & & R_{-1} \end{pmatrix} = \begin{pmatrix} P_{00} & P_{01} & \cdots & P_{0k} \\ & P_{11} & \cdots & P_{1k} \\ & & \ddots & \vdots \\ 0 & \cdots & 0 & P_{kk} \\ 0 & & \cdots & 0 \end{pmatrix}.$$

We can then deduce the two equations, which, conversely, determine the block-rows of B :

$$(3.2) \quad \begin{aligned} B_{k+1,0}(R_0, \cdots, R_k) + B_{k+1,1}(R_{-1}, R_0, \cdots, R_{k-1}) \\ + \cdots + (R_{-k-1}, \cdots, R_{-1}) = 0, \\ B_{k0}(R_0, \cdots, R_{k-1}) + B_{k1}(R_{-1}, R_0, \cdots, R_{k-2}) \\ + \cdots + (R_{-k}, \cdots, R_{-1}) = 0. \end{aligned}$$

With the notations (2.4), we further deduce from (3.2) (by solving the second equation for (R_{-k}, \cdots, R_{-1}) and substituting the result in the first:

$$(3.3) \quad E_0(R_1, \cdots, R_k) + E_1(R_0, \cdots, R_{k-1}) + \cdots + E_k(R_{-k+1}, \cdots, R_0) = 0, \\ B_{k+1,0}R_0 + B_{k+1,1}R_{-1} + \cdots + R_{-k-1} = 0.$$

Our plan is to express the first block-row (R_1, \cdots, R_k) in the form $Y_1(R_0, \cdots, R_{k-1}) + \cdots + Y_k(R_{-k+1}, \cdots, R_0)$ where the Y_i 's do not involve the $B_{k+1,i}$'s which we are looking for; for, once this is done, we can deduce from (3.3) an equation of the form

$$(3.4) \quad X_1(R_0, \cdots, R_{k-1}) + \cdots + X_k(R_{-k+1}, \cdots, R_0) = 0,$$

which has only the solution $X_i = 0$ for $i = 1, \cdots, n$. This is true since the blockwise $k \times k$ -submatrix of R with the indicated rows is nonsingular by assumption.

A look at the second equation in (3.2) suggests how the stated goal is achieved: define $R^* = \{R_{ij}^*\}$ with $R_{ij}^* = R_{i-j}$; i.e., replace in (3.1) R_i by R_{-i} . R^* is seen to be equivalent to R , so that the factorization (1.2), $B^*R^*\hat{B}^* = D^*$, exists. Then, as in (3.2),

$$(3.5) \quad B_{k_0}^*(R_0, \dots, R_{-k+1}) + B_{k_1}^*(R_1, \dots, R_{-k+2}) + \dots + (R_k, \dots, R_1) = 0,$$

where the $B_{k_i}^*$'s play the role of the Y_i 's as planned. By solving (3.5) for (R_k, \dots, R_1) and substituting the result in (3.3), we get

$$(E_1 - E_0 B_{k,k-1}^*)(R_0, \dots, R_{k-1}) + \dots + (E_k - E_0 B_{k_0}^*)(R_{-k+1}, \dots, R_0) = 0,$$

which, since (3.4) has only the trivial solution, holds if and only if

$$(3.6) \quad E_i - E_0 B_{k,k-i}^* = 0, \quad i = 1, \dots, k.$$

The second equation in (3.3) determines E_0 as follows: First, (3.3) with (2.4) implies

$$E_0 R_0 + \dots + E_k R_{-k} + \sum_{i=0}^k B_{k_i} R_{-1-i} = 0.$$

Then, with (3.6), we get

$$(3.7) \quad \begin{aligned} E_0 &= C_k = -F_k D_k^{*-1}, & k &\geq 0, \\ F_k &= \sum_{i=0}^k B_{k_i} R_{-i-1}, & k &\geq 0, \\ D_k^* &= \sum_{i=0}^k B_{k_i}^* R_{i-k}, & k &\geq 0. \end{aligned}$$

By defining $B_{i,i} = B_{i_j}^* = 0$ for i or j negative or $i < j$, and $B_{i,i} = B_{i_i}^* = I$ for $i \geq 0$, we deduce from (2.4) and (3.6):

$$(3.8) \quad B_{k+1,i} = B_{k,i-1} + C_k B_{k,k-i}^*, \quad i = 0, \dots, k, k \geq 0.$$

Similarly, we obtain

$$(3.9) \quad B_{k+1,i}^* = B_{k,i-1}^* + C_k^* B_{k,k-i}, \quad i = 0, \dots, k, k \geq 0,$$

where

$$(3.10) \quad \begin{aligned} C_k^* &= -F_k^* D_k^{-1}, & k &\geq 0, \\ F_k^* &= \sum_{i=0}^k B_{k_i}^* R_{i+1}, & k &\geq 0, \\ D_k &= \sum_{i=0}^k B_{k_i} R_{k-i}. \end{aligned}$$

Further, D_k and D_k^* satisfy by (3.7) and (3.9) the recursions

$$(3.11) \quad \begin{aligned} D_{k+1} &= D_k + C_k F_k^*, & k &\geq 0, \\ D_0 &= R_0, \end{aligned}$$

$$(3.12) \quad \begin{aligned} D_{k+1}^* &= D_k^* + C_k^* F_k, & k &\geq 0, \\ D_0^* &= R_0. \end{aligned}$$

Equations (3.8)–(3.12) describe an algorithm for finding the factor B in (1.2). Again, \hat{B}' can be obtained from the same formulas by replacing R_i by R'_{-i} . Hence, in particular, if R is a symmetric Toeplitz matrix and $R_i = R'_{-i}$, then $\hat{B}' = B$. But even another simplification results: as was shown by J. P. Burg (unpublished),

$$(3.13) \quad F_k^* = F'_k \quad (\text{if } R_{-i} = R'_i).$$

To see this, observe that with (3.2) and (3.5) we have

$$(3.14) \quad F_k^* = (I, B_{k,k-1}^*, \dots, B_{k0}^*, 0) \begin{pmatrix} R_0 & R_1 & \cdots & R_{k+1} \\ R'_1 & R_0 & \cdots & R_k \\ \vdots & & & \vdots \\ R'_{k+1} & \cdots & R_0 \end{pmatrix} \begin{pmatrix} 0 \\ B'_{k0} \\ \vdots \\ B'_{k,k-1} \\ I \end{pmatrix} = F'_k.$$

We point out that it is possible to derive a recurrence equation also for F_k and F_k^* , as we shall show shortly. However, these recurrence relations save calculations only when both factorizations (1.1) and (1.2) are determined.

We shall now turn to the problem of finding the factors in (1.1). The situation turns out to be different from the case with Hankel matrices in that the algorithms for determining the factorizations (1.1) and (1.2), respectively, are independent from each other. Beginning with (3.1), we have

$$(3.15) \quad P_{i,k+1} = \sum_{j=0}^i B_{ij} R_{k+1-j}, \quad 1 \leq i \leq k+1, \\ P_{0k} = R_k, \quad k \geq 0.$$

An application of (3.8) to this expression leads to

$$P_{i,k+1} = C_{i-1} R_{k+1} + R_{k+1-i} \\ + \sum_{j=1}^{i-1} B_{i-1,j-1} R_{k+1-j} + C_{i-1} \sum_{j=1}^{i-1} B_{i-1,i-1-j}^* R_{k+1-j}, \quad k \geq 0.$$

By adding R_{k+1-i} to the first sum and writing

$$(3.16) \quad Q_{ik} = \sum_{j=0}^i B_{ij}^* R_{k-i+j+1}, \quad i = 0, \dots, k, k \geq 0,$$

for the second sum joined with R_{k+1} (with $i-1$ replaced by i), we obtain

$$(3.17) \quad P_{i+1,k+1} = P_{ik} + C_i Q_{ik}, \quad k \geq 0, i = 0, \dots, k, \\ P_{0k} = R_k, \quad k \geq 0.$$

An application of (3.9) to (3.16) gives in turn the recursion for the Q_{ik} 's:

$$(3.18) \quad Q_{i+1,k} = Q_{ik} + C_i^* P_{ik}, \quad k \geq 1, i = 0, \dots, k-1, \\ Q_{0k} = R_{k+1}, \quad k \geq 0.$$

Now observe from (3.16) and (3.15):

$$(3.19) \quad \begin{aligned} Q_{kk} &= F_k^*, & k \geq 0, \\ P_{kk} &= D_k, & k \geq 0. \end{aligned}$$

In order to also get F_k recursively, replace, in (3.15)–(3.16), R_i by R_{-i} and B_{ij} by B_{ij}^* . Then,

$$(3.20) \quad \begin{aligned} P_{i+1,k+1}^* &= P_{ik}^* + C_i^* Q_{ik}^*, & k \geq 0, i = 0, \dots, k, \\ P_{0k}^* &= R_{-k}, & k \geq 0, \end{aligned}$$

and

$$(3.21) \quad \begin{aligned} Q_{i+1,k}^* &= Q_{ik}^* + C_i P_{ik}^*, & k \geq 1, i = 0, \dots, k-1, \\ Q_{0k}^* &= R_{-k-1}, & k \geq 0. \end{aligned}$$

Further,

$$(3.22) \quad \begin{aligned} Q_{kk}^* &= F_k, & k \geq 0, \\ P_{kk}^* &= D_k^*, & k \geq 0. \end{aligned}$$

Finally, we now can express C_i and C_i^* as follows:

$$(3.23) \quad \begin{aligned} C_k &= -Q_{kk}^* P_{kk}^{*-1}, & k \geq 0, \\ C_k^* &= -Q_{kk} P_{kk}^{-1}, & k \geq 0. \end{aligned}$$

With

$$(3.24) \quad \hat{A} = D^{-1}P,$$

Eqs. (3.17)–(3.23) describe an algorithm for the factor \hat{A} in (1.1). By replacing R_i by R_{-i} , the same algorithm also finds the factor A' . In particular, when R is a symmetric Toeplitz matrix, $\hat{A}' = A$, and (3.14) gives

$$(3.25) \quad Q_{kk}^* = Q'_{kk}, \text{ (if } R_{-i} = R'_i \text{)}.$$

These recurrence relations require $O(4n^2(p^3 + p^2))$ arithmetic operations to determine the factorization (1.1), and only $O(5n^2(p^3 + p^2))$ to find both (1.1) and (1.2) in case $p > 1$. In case $p = 1$, they require only $O(2n^2)$ for (1.1) and $O(3n^2)$ for both (1.1) and (1.2).

In case R is a symmetric positive definite matrix, the factorization

$$(3.26) \quad R = \bar{A} \bar{A}'$$

is called a Cholesky factorization [6]. Such a factorization is obtained from (1.1) by first factoring D as

$$(3.27) \quad D = GG',$$

where G may be taken to be lower triangular with positive elements on the diagonal, and by then putting

$$(3.28) \quad \bar{A} = AG.$$

Therefore, with this addition, our algorithm calculates even the Cholesky factorizations of symmetric positive definite Toeplitz matrices.

As was shown in [5] for $p = 1$ and in [7] for $p \geq 1$, the Cholesky factorization has a close connection with the classical spectral factorization problem: Let $R = \{R_{ij}\}$, $i, j = 0, 1, \dots$, $R_{ij} = R_{j-i}$ and $R_i = 0$ for $i > m$, be the infinite symmetric positive definite block Toeplitz matrix obtained from the coefficients of a $p \times p$ -matrix polynomial $R(x) = \sum_{i=-m}^m R_i x^i$, $R_i = R_i^L$, which is positive definite for $|x| = 1$. Then, as proved in [7], the rows of the Cholesky factor \bar{A} of R converge:

$$(\bar{A}_{n,n-m}, \dots, \bar{A}_{nn}) \rightarrow (\bar{A}_m, \dots, \bar{A}_0) \quad \text{as } n \rightarrow \infty.$$

The limits \bar{A}_i define the outer factor $\bar{A}(x) = \sum_{i=0}^m \bar{A}_i x^i$ of the spectral factorization (1.3) of $R(x)$.

In this way, the algorithm (3.17)–(3.23) with (3.27)–(3.28), where now $P_{ik} = 0$ for $k - i > m$ and $Q_{ik} = 0$ for $k - i \geq m$, provides a new method for calculating the spectral factors with any desired accuracy. These formulas require $O(p^3 m)$ operations for each block-row of \bar{A} as compared with $O(p^3 m^2)$ with the earlier algorithms in [5] and [7] which basically are Gauss-type elimination schemes. Finally, in this case, our algorithm (just as that in [5] and [7]) is also numerically stable as can be shown by standard stability arguments.

IBM Corporation
 Monterey & Cottle Roads
 San Jose, California 95114

1. N. LEVINSON, "The Wiener RMS error criterion in filter design and prediction," Appendix B of N. Wiener's book *Extrapolation, Interpolation, and Smoothing Stationary Time Series with Engineering Applications*, Wiley, New York, 1949, pp. 129–148.
2. R. A. WIGGINS & E. A. ROBINSON, "Recursive solution to the multichannel filtering problem," *J. Geophys. Res.*, v. 70, 1965, pp. 1885–1891. MR 32 #589.
3. J. L. PHILLIPS, "The triangular decomposition of Hankel matrices," *Math Comp.*, v. 25, 1971, pp. 599–602.
4. A. DEVINATZ, "The factorization of operator valued functions," *Ann. of Math.*, (2), v. 73, 1961, pp. 458–495. MR 23 #A3997.
5. F. L. BAUER, "Ein direktes Iterationsverfahren zur Hurwitz-Zerlegung eines Polynoms," *Arch. Elek. Übertr.*, v. 9, 1955, pp. 285–290. MR 17, 900.
6. J. H. WILKINSON & C. REINSCH, *Handbook for Automatic Computation*. Vol. II. *Linear Algebra*, Springer-Verlag, New York, 1971.
7. J. RISSANEN & L. BARBOSA, "Properties of infinite covariance matrices and stability of optimum predictors," *Information Sci.*, v. 1, 1969, pp. 221–236. MR 39 #5032.