Integral Equations
and Operator Theory
Vol. 8 (1985)

0278-620X/85/060780-25$1.50+0.20/0
© 1985 Birkhäuser Verlag, Basel

LINEAR COMPLEXITY ALGORITHMS FOR SEMISEPARABLE MATRICES†

I. Gohberg,[1] T. Kailath,[2] and I. Koltracht[3]

Linear complexity algorithms are derived for the solution of a linear system of equations with the coefficient matrix represented as a sum of diagonal and semiseparable matrices. LDU-factorization algorithms for such matrices and their inverses are also given. The case in which the solution can be efficiently updated is treated separately.

INTRODUCTION

In the present paper we shall consider $(N+1)\times(N+1)$ matrices

$$R = D+S \tag{0.1}$$

where $D = \mathrm{diag}\{d_k,\ 0 \le k \le N\}$ is a diagonal matrix and S is a semiseparable matrix. By definition a matrix $S = (s_{jm})_{j,m=0}^{N}$ is called semiseparable of order n if

$$s_{jm} = \begin{cases} \sum\limits_{i=1}^{n} g_i(j)h_i(m) & \text{for} \quad 0 \le j < m \le N, \\ \sum\limits_{i=1}^{n} p_i(j)q_i(m) & \text{for} \quad 0 \le m < j \le N, \\ 0 & \text{for} \quad m = j, \end{cases} \tag{0.2}$$

where $g_i = (g_i(j))_{j=0}^{N}$, $h_i = (h_i(j))_{j=0}^{N}$, $p_i = (p_i(j))_{j=0}^{N}$ and

†This work was supported in part by the U.S. Army Research Office, under Contract DAAG29-83-K-0028, and the Air Force Office of Scientific Research, Air Force Systems Command under Contract AF83-0228.

[1]Tel-Aviv University, Dept. of Math. Sci., Israel.

[2]Stanford University, Department of Electrical Engineering, Stanford, CA 94305

[3]Now at University of Calgary, Dept. of Math. and Stat., Calgary, Canada.

$q_i = (q_i(j))_{j=0}^{N}$, $i = 1,\dots,n$ are given vectors in $\mathbb{C}^{N+1}$. In other words the matrix S is composed of the upper triangular part of some matrix of rank at most n, and from the lower triangular part of some other matrix also of a rank at most n. Throughout this paper we consider matrices with nonsingular principal leading minors. Such matrices are called strongly regular. In the last section we indicate however how the obtained results may be applied to nonsingular but not strongly regular matrices.

The following system of $(N+1)\times(N+1)$ matrices plays an important role in our method.

$$R_k = \begin{bmatrix} r_{00} & \cdots & r_{0k} & 0 & \cdots & 0 \\ \vdots & & & & & \\ r_{k0} & \cdots & r_{kk} & 0 & \cdots & 0 \\ r_{k+10} & \cdots & r_{k+1k} & 1 & \cdots & 0 \\ & & & & \ddots & \\ r_{N0} & \cdots & r_{Nk} & 0 & & 1 \end{bmatrix}, \quad k = 0,1,\dots,N. \tag{0.3}$$

In the first section we study solutions of the equations

$$R_k\gamma_k = e_k, \qquad k = 0,1,2,\dots,N \tag{0.4}$$

where e_k is the unit column vector with 1 in the (k+1)-th position and zeros elsewhere. The vectors γ_k, $k = 0,\dots,N$ obey some useful recursions for a sum of a diagonal and semiseparable matrices. Using these recursions we shall prove in Sec. 2 that the triangular factors of the LDU-decomposition of such matrices are also semiseparable of the same order. In Sec. 3 we use the constructive proof of this result to derive an $O(n^2N)$ algorithm for the solution of the equation

$$R\chi = f \tag{0.5}$$

where the matrix R is given by (0.1) and $f \in \mathbb{C}^{N+1}$ is an arbitrary vector. This algorithm derived in Sec. 3 has the drawback that it does not allow an efficient update of the solution, when the matrix R and the right hand side f are increasing in size. In Sec. 4 we show that under certain additional conditions that are satisfied in different applications this difficulty can be avoided. It turns out that if

$$\sum_{i=1}^{n} g_i(k)h_i(k) \neq d_k, \qquad k = 0,\dots,N \tag{0.6}$$

then the upper triangular factor in the UDL factorization of R^{-1} preserves the semiseparability order of the matrix R. We use this result and its constructive proof to derive an alternative algorithm that allows an efficient update of the solution.

We note that sums of diagonal and semiseparable matrices appear in different applications. For example they can be impulse responses of finite dimensional linear (time-invariant or time-variant) systems, or covariance matrices of the response of such systems to a white noise input (see, e.g., Kailath (1980), Chapter 9, Anderson-Moore (1979)). Several matrices differently introduced in the literature can also be recognized as being of this class, e.g., the so-called diagonal innovation matrices (DIM) in Carayanis, Kalouptsidis and Manolakis (1982) (see also Krishna-Morgera (1984)) and the so-called Brownian matrices and their extensions as studied in Picinbono (1981).

For matrices arising from finite-dimensional systems, the integer n in the definition (0.2) can be interpreted as the dimension of the state-vector in a state-space description of the underlying finite-dimensional system. Therefore, we might expect that the "per-step" complexity will depend only upon the size n of the state-vector, and therefore that the complexity of the solution to an N-step problem (leading to N×N matrices) will be O(N). This was demonstrated in the corresponding integral equation context in Kailath (1969), by using the Kalman filter related Riccati equations (to yield $O(n^3N)$ algorithms), and in Kailath (1973), by using the special Chandrasekhar equations for time-invariant systems (to yield $O(n^2N)$ algorithms).

Several results on the inversion of matrices that satisfy the condition (0.6) appear in Gohberg and Kaashoek (1984). These results are based on connections between such matrices and discrete linear state-space systems with boundary conditions.

The results of this paper are applcable in a more general context and do not appeal to any possible underlying state-space system. Our method is based on ideas developed in Gohberg and Koltracht (1985) and Gohberg, Kailath and Koltracht (1985).

1. First Recursions for Semiseparable Matrices

In this section we deduce some basic equalities which will be used for recursive solution of linear systems and which in the semiseparable case already give an $O(N^2)$ solution algorithm.

Let $R = (r_{jm})_{j,m=0}^{N}$ be an $(N+1)\times(N+1)$ matrix with complex entries. Let R_k^ℓ denote the principal leading minor $(r_{jm})_{j,m=0}^{k}$. We will assume that all principal leading minors R_k^ℓ, $k = 0,\dots,N$ are invertible. Such matrices will be called strongly regular. For the matrix R we introduce the following system of $(N+1)\times(N+1)$ matrices

$$R_k = \begin{bmatrix} r_{00} & \cdots & r_{0k} & 0 & \cdots & 0 \\ \vdots & & & & & \\ r_{k0} & \cdots & r_{k,k} & 0 & \cdots & 0 \\ r_{k+10} & \cdots & r_{k+1,k} & 1 & \cdots & 0 \\ & & & & \ddots & \\ r_{N0} & \cdots & r_{Nk} & 0 & \cdots & 1 \end{bmatrix}, \quad k = 0,\dots,N. \tag{1.1}$$

It is clear that R_k is invertible if and only if R_k^ℓ is invertible. An important role in this section will be played by vectors γ_k that satisfy the equations

$$R_k\gamma_k = e_k, \qquad k = 0,\dots,N \tag{1.2}$$

where e_k is the unit column vector with 1 on the (k+1)-th position and zeros elsewhere. If the vectors γ_k are known we can recursively find the solution of the equation

$$R\chi = f \tag{1.3}$$

in the following way. First remark that the solution of the equation $R_0\chi_0 = f$ is given by

$$\chi_0 = \begin{bmatrix} f(0)r_{00}^{-1} \\ f(1)-f(0)r_{10}r_{00}^{-1} \\ \vdots \\ f(N)-f(0)r_{N0}r_{00}^{-1} \end{bmatrix}, \tag{1.4}$$

where f(k) denotes the (k+1)-th entry of the vector f. Suppose now that the solution of the equation $R_{k-1}\chi_{k-1} = f$ is known. Then it is easy to see that the vector

$$\chi_k = \begin{bmatrix} \chi_{k-1}(0) \\ \vdots \\ \chi_{k-1}(k-1) \\ 0 \\ \chi_{k-1}(k+1) \\ \vdots \\ \chi_{k-1}(N) \end{bmatrix} + \chi_{k-1}(k)\gamma_k \tag{1.5}$$

is the solution of the equation $R_k\chi_k = f$. Continuing this recursion we get the last vector χ_N which obviously is the solution of (1.3).

Now let $R = D+S$ be a strongly regular sum of a diagonal matrix $D = \text{diag}\{d_k,\ 0 \le k \le N\}$ and a semiseparable matrix S. A matrix $S = (s_{jm})_{j,m=0}^{N}$ is called semiseparable of order n if for some vectors g_i, h_i, p_i, $q_i \in \mathbb{C}^{N+1}$, $i = 1,\ldots,n$

$$s_{jm} = \begin{cases} \sum_{i=1}^{n} g_i(j)h_i(m) & \text{for} \quad 0 \le j < m \le N \\ \sum_{i=1}^{n} p_i(j)q_i(m) & \text{for} \quad 0 \le m < j \le N . \\ 0 & \text{for} \quad m = j \end{cases} \tag{1.6}$$

Applying R_k to the unit vector e_k we get

$$R_k e_k = d_k e_k + \sum_{i=1}^{n} h_i(k) \begin{bmatrix} g_i(0) \\ \vdots \\ g_i(k-1) \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \sum_{i=1}^{n} q_i(k) \begin{bmatrix} 0 \\ \vdots \\ 0 \\ p_i(k+1) \\ \vdots \\ p_i(N) \end{bmatrix} . \tag{1.7}$$

It is now convenient to introduce some notations. Given n column vectors $\alpha_1,\ldots,\alpha_n \in \mathbb{C}^{N+1}$, we denote by α the $(N+1)\times n$ matrix

$$\alpha = [\alpha_1,\ldots,\alpha_n] \tag{1.8}$$

and by $\alpha(k)$ the k-th row of this matrix:

$$\alpha(k) = [\alpha_1(k),\ldots,\alpha_n(k)]. \tag{1.9}$$

We also introduce projection matrices

$$V_k = \begin{bmatrix} 0 & & & & & \\ & \ddots & & & 0 & \\ & & 0 & & & \\ & & & 1 & & \\ & 0 & & & \ddots & \\ & & & & & 1 \end{bmatrix} \tag{1.10}$$

with first k+1 zeros on the diagonal and projection matrices

$$W_k = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ & & & 0 & & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \tag{1.11}$$

with zero on the (k+1)-th position on the diagonal.

Using these notations and denoting by T the transpose we can rewrite (1.7) after a simple transformation as

$$R_k e_k = [d_k - h(k) g^T(k)] e_k + g h^T(k) + V_k p q^T(k) - V_k g h^T(k).$$

Now, applying R_k^{-1} to both sides of this equality and using the fact that $R_k^{-1} V_k = V_k$, we get

$$e_k = (d_k - h(k) g^T(k)) \gamma_k + \varphi_k h^T(k) + V_k p q^T(k) - V_k g h^T(k) \tag{1.12}$$

where $\varphi = [\varphi_k^1, \ldots, \varphi_k^n]$ with φ_k^i defined as solutions of the equations

$$R_k \varphi_k^i = g_i, \qquad i = 1, \ldots, n. \tag{1.13}$$

Using the formula (1.7) we can write

$$\varphi_k^i = W_{k-1} \varphi_{k-1}^i + \varphi_{k-1}^i(k) \gamma_k, \qquad i = 1, \ldots, n. \tag{1.14}$$

Substituting (1.14) into (1.12) we get

$$e_k = (d_k - h(k) g^T(k) + h(k) \varphi_{k-1}^T(k)) \gamma_k + W_k \varphi_{k-1} h^T(k) +$$
$$+ V_k p q^T(k) - V_k g h^T(k).$$

Computing the (k+1)-th entries of the vectors on both sides of this equation, we get

$$1 = (d_k - h(k) g^T(k) + h(k) \varphi_{k-1}^T(k)) \gamma_k(k) \tag{1.15}$$

and hence

$$\gamma_k = (d_k - h(k) g^T(k) + h(k) \varphi_{k-1}^T(k))^{-1} [-W_k \varphi_{k-1} h^T(k)$$
$$- V_k p q^T(k) + V_k g h^T(k)]. \tag{1.16}$$

This formula, together with (1.14) gives us an $O(N^2)$ algorithm for the computation of the vectors γ_k, $k = 0,\ldots,N$. Indeed, using (1.4) we can easily find γ_0 and $\varphi_0^1,\ldots,\varphi_0^n$. Now, given $\varphi_{k-1}^1,\ldots,\varphi_{k-1}^n$ we can find γ_k via (1.16) and then $\varphi_k^1,\ldots,\varphi_k^n$ via (1.14).

Finally, we remark that similar recursions can be derived for the solutions of the equations

$$(R^T)_k\omega_k = e_k, \qquad k = 0,\ldots,N,$$

where R^T is the transpose of R.

2. Linear Complexity Algorithms for LDU Factorization

In this section we derive an $O(n^2N)$ algorithm for the triangular factorization of a strongly regular sum of a diagonal matrix and a semiseparable matrix of order n. The derivation is based on the recursion (1.16) and on the fact that the triangular factors of R also have the semiseparable structure.

First let us show that the vectors γ_k and ω_k, $k = 0,\ldots,N$ determine the LDU-factorization of R. Since R is strongly regular, it admits a unique representation

$$R = L\mathcal{D}U,$$

where L is a lower triangular matrix with ones on the diagonal, U is upper triangular with ones on the diagonal and $\mathcal{D}$ is a diagonal matrix. Let us consider the equation

$$RU^{-1} = L\mathcal{D}. \tag{2.1}$$

Since

$$R_k^{\ell}\begin{bmatrix} \gamma_k(0) \\ \vdots \\ \\ \gamma_k(k) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \qquad k = 0,\ldots,N,$$

we conclude that the k-th column of U^{-1} has the form $\gamma_k^{-1}(k)[\gamma_k(0),\ldots,\gamma_k(k),0,\ldots,0]^T$, and hence

$$\delta_k = \gamma_k^{-1}(k), \qquad k = 0,\ldots,N. \tag{2.2}$$

It follows from the definition of R_k and γ_k that

$$\begin{bmatrix} r_{k+1,0} & \cdots & r_{k+1,k} \\ \vdots & & \\ r_{N0} & \cdots & r_{Nk} \end{bmatrix} \begin{bmatrix} \gamma_k(0) \\ \vdots \\ \gamma_k(k) \end{bmatrix} = -\begin{bmatrix} \gamma_k(k+1) \\ \vdots \\ \gamma_k(N) \end{bmatrix}$$

while from (2.1) we have

$$\frac{1}{\gamma_k(k)} \begin{bmatrix} r_{k+1,0} & \cdots & r_{k+1,k} \\ \vdots & & \\ r_{N,0} & \cdots & r_{Nk} \end{bmatrix} \begin{bmatrix} \gamma_k(0) \\ \\ \gamma_{k(k)} \end{bmatrix} = \delta_k \begin{bmatrix} \ell_{k+1,k} \\ \vdots \\ \ell_{N,k} \end{bmatrix}.$$

Therefore $[\ell_{k+1,k} \cdots \ell_{Nk}]^T = -[\gamma_{k(k+1)} \cdots \gamma_{k(n)}]^T$ and we can write

$$L = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ -\gamma_0(1) & 1 & \ldots & 0 \\ -\gamma_0(2) & -\gamma_1(2) & \ldots & 0 \\ & & \ldots & \\ -\gamma_0(N) & -\gamma_1(N) & \ldots & 1 \end{bmatrix}. \tag{2.3}$$

Similarly, comparing the equations

$$R^T(L^T)^{-1} = U^T\mathcal{D} \quad \text{and} \quad (R^T)_k\omega_k = e_k$$

we get

$$U = \begin{bmatrix} 1 & -\omega_0(1) & -\omega_0(2) & \ldots & -\omega_0(N) \\ 0 & 1 & -\omega_1(2) & \ldots & -\omega_1(N) \\ & & & \ldots & \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}. \tag{2.4}$$

We now show that triangular factors always preserve the order of semiseparability of the original matrix.

THEOREM 2.1. *Let* $R = D+S$ *be a strongly regular sum of a diagonal matrix and a semiseparable matrix of order* n. *Then* R *admits the* LDU-*factorization*

$$R = (I+S_L)\mathcal{D}(I+S_U), \tag{2.5}$$

with S_L *lower triangular and* S_U *upper triangular semiseparable matrices each of order* n.

PROOF. Let $g_i, h_i, q_i, p_i \in \mathbb{C}^{N+1}$, $i = 1,\dots,n$ be the vectors that determine the matrix S and, for $k = 0,\dots,N$ let

$$R_k\gamma_k = e_k$$
$$R_k\varphi_k^i = g_i, \qquad i = 1,\dots,n.$$

First let us solve these equations for $k = 0$. Recalling (1.4) we have

$$\gamma_0 = \begin{bmatrix} d_0^{-1} \\ -p(1)q^T(0)d_0^{-1} \\ \\ -p(N)q^T(0)d_0^{-1} \end{bmatrix}, \quad \varphi_0 = \begin{bmatrix} g(0)d_0^{-1} \\ g(1) \quad -p(1)q^T(0)g(0)d_0^{-1} \\ \vdots \\ g(N) \quad -p(N)q^T(0)g(0)d_0^{-1} \end{bmatrix} \tag{2.6}$$

where φ_0 is an $(N+1)\times n$ matrix: $\varphi_0 = [\varphi_0^1,\dots,\varphi_0^n]$. It will be convenient to introduce for $k = 0,\dots,N$ the following $(N-k)\times(N+1)$ projection matrices

$$T_k = \begin{bmatrix} 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 \end{bmatrix}$$

with first $k+1$ zero columns.

Using these matrices we get from (2.6)

$$T_0\varphi_0 = T_0g - T_0pq^T(0)g(0)/d_0$$

and

$$T_0\gamma_0 = -T_0pq^T(0)/d_0.$$

Introducing the notations

$$x_0 = -q^T(0)/d_0, \quad y_0 = -q^T(0)g(0)/d_0 \tag{2.7}$$

we have

$$T_0\gamma_0 = T_0px_0$$
$$T_0\varphi_0 = T_0g + T_0py_0.$$

We now show that a similar representation holds for all $k = 0,\dots,N-1$. Assume that for some column vector $x_{k-1} \in \mathbb{C}^n$ and for some $n\times n$-matrix y_{k-1} the following equalities hold:

$$T_{k-1}\gamma_{k-1} = T_{k-1}px_{k-1}$$
$$T_{k-1}\varphi_{k-1} = T_{k-1}g + T_{k-1}py_{k-1}.$$

Since, in particular, for the first entries of the latter equality we have

$$\varphi_{k-1}(k) = g(k)+p(k)y_{k-1}$$

it follows from (1.15) that

$$\gamma_k(k) = (d_k+p(k)y_{k-1}h^T(k))^{-1}. \tag{2.8}$$

Now it follows from (1.14) and (1.16) that if we chose

$$x_k = -\gamma_k(k)[y_{k-1}h^T(k)+q^T(k)], \tag{2.9}$$

$$y_k = y_{k-1} + x_k[g(k)+p(k)y_{k-1}] \tag{2.10}$$

then

$$T_k\gamma_k = T_kpx_k, \tag{2.11}$$

$$T_k\varphi_k = T_kg + T_kpy_k. \tag{2.12}$$

Let us now consider solutions of the equations

$$R_k^T\psi_k^i = h_i, \quad i = 1,\dots,n, \ k = 0,\dots,N$$
$$R_k^T\omega_k = e_k, \qquad\qquad k = 0,\dots,N.$$

As above we can show that there exist column vectors $u_k \in \mathbb{C}^n$ and $n\times n$ matrices v_k such that

$$T_k \omega_k = T_k h u_k, \quad k = 0,\dots,N-1 \tag{2.13}$$

$$T_k \psi_k = T_k q + T_k h v_k, \quad k = 0,\dots,N-1. \tag{2.14}$$

and they can be found recursively starting with

$$u_0 = -g^T(0)/d_0, \quad v_0 = -g^T(0)q(0)/d_0 \tag{2.15}$$

and for $k = 1,\dots,N-1$

$$u_k = -\gamma_k(k)[v_{k-1}p^T(k)+g^T(k)] \tag{2.16}$$

and

$$v_k = v_{k-1} + u_k[q(k)+h(k)v_{k-1}]. \tag{2.17}$$

It follows now from (2.3), (2.4), (2.8), (2.11) and (2.13) that

$$S_L = \begin{bmatrix} 0 & 0 & \dots & 0 \\ -p(1)x_0 & 0 & \dots & 0 \\ -p(2)x_0 & -p(2)x_1 & \dots & 0 \\ \vdots & \vdots & & \\ -p(N)x_0 & -p(N)x_1 & \dots & 0 \end{bmatrix}, \tag{2.18}$$

$$S_U = \begin{bmatrix} 0 & -h(1)u_0 & -h(2)u_0 & \dots & -h(N)u_0 \\ 0 & 0 & -h(2)u_1 & \dots & -h(N)u_1 \\ & & & \dots & \\ 0 & 0 & 0 & & 0 \end{bmatrix} \tag{2.19}$$

and

$$\mathcal{D} = \mathrm{diag}\{\delta_k,\ 0 \le k \le N\}$$

with

$$\delta_0 = 1, \ \delta_k = d_k + p(k)y_{k-1}h^T(k), \quad k = 1,\dots,N. \tag{2.20}$$

Since x_j, u_j, $j = 0,\dots,N-1$ are column vectors of length n and $p(j)$, $q(j)$, $j = 1,\dots,N$ are row vectors of length n, it follows from the definition that S_L and S_U have semiseparability order n. The Theorem is proved. □

For the special case of covariance matrices this result was first proved in Gevers and Kailath (1973).

We remark that the above proof is constructive and in fact gives an efficient algorithm for the computation of the LDU-factorization of R. Note first that substituting (2.9) into (2.10) we get

$$y_k = y_{k-1} - \gamma_k(k)[y_{k-1}h^T(k)+q^T(k)][g(k)+p(k)y_{k-1}].$$

Similarly

$$v_k = v_{k-1} - \omega_k(k)[v_{k-1}p^T(k)+g^T(k)][q(k)+h(k)v_{k-1}].$$

Since $\gamma_k(k) = \omega_k(k)$ and $v_0 = y_0^T$, we conclude that

$$v_k = y_k^T, \quad \text{for} \quad k = 0,\dots,N.$$

ALGORITHM 2.1

Let the matrix R satisfy the conditions of Theorem 2.1. Then LDU-factorization of R can be found as follows:

1. Start with x_0, y_0 and u_0 defined by (2.7) and (2.15).
2. Compute recursively, for $k = 1,\dots,N-1$ x_k, y_k and u_k via (2.9), (2.10) and (2.16).
3. The LDU-factorization of R is now given by (2.18), (2.19).

This algorithm will take $(4n^2+2n)N$ multiplications/divisions and $(4n^2+3n)N$ additions/subtractions.

3. The Main Algorithm

In this section we consider the solution of the equation

$$R\chi = f \tag{3.1}$$

where R is a strongly regular sum of a diagonal matrix and a semi-separable matrix of order n. This solution can be found in $O(n^2N)$ arithmetic operations in the following way. First compute the LDU factorization of R via the Algorithm 2.1,

$$R = (I+S_L)\mathcal{D}(I+S_U).$$

Then find the solution of the equation

$$(I+S_L)a = f$$

and finally solve the equation

$$(I+S_U)\chi = \mathcal{D}^{-1}a.$$

Let us now show how to solve linear systems of equations with triangular semiseparable matrices. Let

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ -p(1)x_0 & 1 & & \vdots \\ \vdots & \vdots & & \\ -p(N)x_0 & -p(N)x_1 & & 1 \end{bmatrix} \begin{bmatrix} a(0) \\ \vdots \\ a(N) \end{bmatrix} = \begin{bmatrix} f(0) \\ \vdots \\ f(N) \end{bmatrix}.$$

Obviously $a(0) = f(0)$. Suppose that $a(0),\ldots,a(k-1)$ are known. Then

$$a(k) - \sum_{j=0}^{k-1} p(k)x_j a(j) = f(k).$$

Let us introduce column vectors $z_k \in \mathbb{C}^n$ via

$$z_0 = 0, \quad z_k = \sum_{j=0}^{k-1} a(j)x_j, \quad k = 1,\ldots,N.$$

Then clearly

$$z_k = z_{k-1} + a(k-1)x_{k-1}, \quad k = 1,\ldots,N \tag{3.2}$$

and

$$a(k) = p(k)z_k + f(k), \; k = 1,\ldots,N. \tag{3.3}$$

The computation of the entries of the vector a via (3.2) and (3.4) will take 2nN multiplications and 2nN additions. In a similar way we can solve the linear system with the upper triangular semiseparable matrix

$$\begin{bmatrix} 1 & -h(1)u_0 & -h(2)u_0 & \cdots & -h(N)u_0 \\ 0 & 1 & -h(2)u_1 & \cdots & -h(N)u_1 \\ & & & \cdots & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \chi(0) \\ \vdots \\ \chi(N) \end{bmatrix} = \begin{bmatrix} \gamma_0(0)a(0) \\ \vdots \\ \gamma_N(N)a(N) \end{bmatrix}.$$

First we set

$$\chi(N) = \gamma_N(N)a(N), \quad t(N) = 0$$

and then compute recursively for $k = 1,\dots,N$

$$t(N-k) = t(N-k+1)+\chi(N-k+1)h(N-k+1),$$

$$\chi(N-k) = t(N-k)u_{N-k} + \gamma_{N-k}(N-k)a(N-k).$$

The vectors $t(k)$, $k = 0,\dots,N$ are here row vectors of length n.

ALGORITHM 3.1

Let $R = D+S$ be a strongly regular sum of a diagonal matrix $D = \mathrm{diag}\{d_k,\ 0 \le k \le N\}$ and a semiseparable matrix of order n $S = \{s_{jm}\}_{j,m=0}^{N}$, where

$$s_{jm} = \begin{cases} \sum_{i=1}^{n} g_i(j)h_i(m) & \text{for} \quad 0 \le j < m \le N \\ \sum_{i=1}^{n} p_i(j)q_i(m) & \text{for} \quad 0 \le m < j \le N \ . \\ 0 & \text{for} \quad j = m \end{cases}$$

Then one can find the solution of the equation

$$R\chi = f \tag{3.4}$$

in the following way.

1. Start with

$$\gamma_0(0) = \omega_0(0) = 1/d_0 \tag{3.5}$$

$$x_0 = -\gamma_0(0)q^T(0) \tag{3.6}$$

$$u_0 = -\omega_0(0)g^T(0) \tag{3.7}$$

$$v_0^T = y_0 = -\gamma_0(0)q^T(0)g(0) \tag{3.8}$$

and compute recursively for $k = 1,2,\dots,N$

$$c_k = y_{k-1}h^T(k), \quad \Delta_k = v_{k-1}p^T(k) \tag{3.9}$$

$$\gamma_k(k) = \omega_k(k) = (d_k+p(k)c_k)^{-1} \tag{3.10}$$

$$x_k = -\gamma_k(k)[c_k+q^T(k)] \tag{3.11}$$

$$u_k = -\omega_k(k)[\Delta_k+g^T(k)] \tag{3.12}$$

$$v_k^T = y_k = y_{k-1}+x_k[g(k)+p(k)y_{k-1}]. \tag{3.13}$$

2. Start with

$$a(0) = f(0) \tag{3.14}$$

and compute recursively for $k = 1,\ldots,N$

$$z_k = z_{k-1}+a(k-1)x_k, \tag{3.15}$$

$$a(k) = p(k)z_k+f(k). \tag{3.16}$$

3. Start with

$$\chi(N) = \gamma_N(N)a(N), \quad t(N) = 0 \tag{3.17}$$

and compute recursively for $k = 1,\ldots,N$

$$t(N-k) = t(N-k+1)+\chi(N-k+1)h(N-k+1) \tag{3.18}$$

$$\chi(N-k) = t(N-k)u_{N-k}+\gamma_{N-k}(N-k)a(N-k). \tag{3.19}$$

This algorithm will take $(4n^2+6n+1)N$ multiplications/divisions and $(4n^2+8n+2)N$ additions/subtractions.

We remark that here x_k, u_k, c_k, d_k and z_k are column vectors of length n, $p(k)$, $q(k)$, $g(k)$, $h(k)$ and $t(k)$ are row vectors of length n, y_k are $n\times n$ matrices and $\gamma_k(k)$, $f(k)$, $a(k)$ and $\chi(k)$ are complex numbers.

4. Efficient Updating of the Solution

In many applications it is required to update the solution of the system $R\chi = f$ when the matrix R and the right hand side f are increasing in size. Because of the back substitution in step 3 each update via Algorithm 3.1 will require O(N) arithmetic operations. There is however a special case in which this back substitution can be avoided, namely when

$$d_k - \sum_{i=1}^{n} g_i(k)h_i(k) \neq 0, \quad k = 0,\ldots,N-1. \tag{4.1}$$

This situation occurs in many applications. For example, the matrix D and the matrix S completed on the diagonal by

$$s_{kk} = \sum_{i=1}^{n} g_i(k)h_i(k), \qquad k = 0,\ldots,N-1$$

can be covariance matrices and hence positive definite (Gevers and Kailath (1973), Kailath, Morf and Sidhu (1973)). The condition (4.1) also holds for matrices I+S arising from the discretization of Fredholm integral equations of the second kind with semi-separable kernels (Gohberg, Koltracht (1985)).

We shall show that the condition (4.1) guarantees that the upper triangular factor in the UDL-factorization of R^{-1} preserves the semiseparability order of the original matrix R. This will lead to an efficiently updatable $O(n^2N)$ algorithm for the solution of the system $R\chi = f$.

THEOREM 4.1. *Let* $R = D+S$ *be a strongly regular sum of a diagonal matrix* $D = \mathrm{diag}\{d_k,\ 0 \le k \le N\}$ *and a semiseparable matrix* $S = \{s_{jm}\}_{j,m=0}^{N}$ *of order* n

$$s_{jm} = \begin{cases} \sum_{i=1}^{n} g_i(j)h_i(m) & \text{for} \quad 0 < j < m \le N \\ \sum_{i=1}^{n} p_i(j)q_i(m) & \text{for} \quad 0 \le m < j \le N \\ 0 & \text{for} \quad j = m \end{cases}$$

such that

$$\sum_{i=1}^{n} h_i(k)g_i(k) \neq d_k, \quad k = 0,\ldots,N.$$

Then the upper triangular factor U *in the UDL factorization of* R^{-1} *admits the representation*

$$U = I+S_U$$

with S_U *upper triangular semiseparable matrix of order* n.

PROOF. First we remind the well known fact that the solutions of the equations

$$R_k\gamma_k = e_k, \qquad k = 0,\dots,N$$

determine the upper triangular factor in the UDL-factorization of R^{-1} in the following way,

$$U = \begin{bmatrix} 1 & \gamma_1(0)/\gamma_1(1) & \dots & \gamma_N(0)/\gamma_N(N) \\ 0 & 1 & \dots & \gamma_N(1)/\gamma_N(N) \\ & & \dots & \\ 0 & 0 & \dots & 1 \end{bmatrix}. \tag{4.2}$$

Let us again consider solutions of the equations

$$R_k\varphi_k^i = g_i, \quad i = 1,\dots,n, \; k = 0,\dots,N.$$

It follows from (1.4) that

$$\varphi_0(0) = [\varphi_0^1(0),\dots,\varphi_0^n(0)] = g(0)/d_0. \tag{4.3}$$

Denoting $L_0 = I_n$ and $g(0)/d_0 = \lambda(0)$ we can rewrite (4.3) in the form

$$\varphi_0(0) = \lambda(0)L_0.$$

Next we shall show that for $k = 1,\dots,N$ there exist column vectors $\ell_k \in \mathbb{C}^n$ and invertible $n\times n$ matrices L_k such that

$$\begin{bmatrix} \gamma_k(0) \\ \vdots \\ \gamma_k(k-1) \end{bmatrix} = \begin{bmatrix} \lambda(0) \\ \vdots \\ \lambda(k-1) \end{bmatrix}\ell_k, \quad \begin{bmatrix} \varphi_k(0) \\ \vdots \\ \varphi_k(k) \end{bmatrix} = \begin{bmatrix} \lambda(0) \\ \vdots \\ \lambda(k) \end{bmatrix}L_k. \tag{4.4}$$

Suppose that such a representation holds for $k-1$. Let us choose

$$\ell_k = -\gamma_k(k)L_{k-1}h^T(k) \tag{4.5}$$

and

$$L_k = L_{k-1}[I_n-\gamma_k(k)h^T(k)[g(k)+p(k)y_{k-1}]]. \tag{4.6}$$

Since

$$\gamma_k(k) = (d_k-h(k)g^T(k)+h(k)[g(k)+p(k)y_{k-1}]^T)^{-1}$$

and since by the assumption $d_k-h(k)g^T(k) \neq 0$, the matrix

$I_n - \gamma_k(k)h^T(k)[g(k)+p(k)y_{k-1}]$ is invertible and hence L_k is invertible. Now, choosing

$$\lambda(k) = \gamma_k(k)[g(k)+p(k)y_{k-1}]L_k^{-1} , \tag{4.7}$$

we get from (1.14) and (1.16) the representation (4.4). It follows from from (4.2) and (4.4) that $U = I+S_U$, where

$$S_U = \begin{bmatrix} 0 & \lambda(0)\ell_1/\gamma_1(1) & \cdots & \lambda(0)\ell_N/\gamma_N(N) \\ 0 & 0 & \cdots & \lambda(1)\ell_N/\gamma_N(N) \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Since $\lambda(j)$, $j = 0,\ldots,N-1$ are row vectors of length n and $\ell_j/\gamma_j(j)$, $j = 1,\ldots,N$ are column vectors of length n, it follows from the definition that S_U is a semiseparable matrix of order n. □

The following example shows that if the condition (4.1) in Theorem 4.1 is not met, then the upper triangular factor of R^{-1} may not preserve the semiseparability order of R. First we remark that a 3×3 matrix

$$U = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}$$

is semiseparable of order 1 if and only if $b \neq 0$ or $b = 0$ and $ac = 0$. Now let

$$R = \begin{bmatrix} 1 & \alpha & \beta \\ \alpha & \alpha & \beta \\ \beta & \beta & \beta \end{bmatrix} \qquad \alpha \neq 0,1,\ \alpha \neq \beta,\ \beta \neq 0.$$

R is apparently semiseparable of order 1. However, the triangular factors of the inverse given by

$$R^{-1} = \begin{bmatrix} 1 & -1/\alpha & 0 \\ 0 & 1 & -\beta/\alpha \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/(\alpha-\alpha^2) & 0 \\ 0 & 0 & \alpha/\beta(\alpha-\beta) \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ -1/\alpha & 1 & 0 \\ 0 & -\beta/\alpha & 1 \end{bmatrix}$$

have semiseparability order 2.

The proof of Theorem 4.1 is constructive and gives an efficient algorithm for the computation of the vectors γ_k for $k = 0,\dots,N$. Having these vectors we can find recursively the solution of the equation $R\chi = f$ using the method described in Sec. 1 in (1.4) and (1.5).

Let R satisfy the conditions of Theorem 4.1 and let

$$R_k\chi_k = f, \quad k = 0,\dots,N.$$

As in the proofs of Theorems 2.1 and 4.1 we can show that for $k = 0,\dots,N$ there exist column vectors $A_k, C_k \in \mathbb{C}^n$ and complex numbers $B(k)$ such that

$$\begin{bmatrix} \chi_k(0) \\ \vdots \\ \chi_k(k) \end{bmatrix} = \begin{bmatrix} \lambda(0) \\ \vdots \\ \lambda(k) \end{bmatrix} A_k + \begin{bmatrix} B(0) \\ \vdots \\ B(k) \end{bmatrix}$$

and

$$\begin{bmatrix} \chi_k(k+1) \\ \vdots \\ \chi_k(N) \end{bmatrix} = \begin{bmatrix} p(k+1) \\ \vdots \\ p(N) \end{bmatrix} C_k + \begin{bmatrix} f(k+1) \\ \vdots \\ f(N) \end{bmatrix}.$$

Moreover, A_k, C_k and $B(k)$ can be determined recursively, starting with

$$A_0 = 0, \; B(0) = f(0)\gamma_0(0), \; C_0 = -\gamma_0(0)q^T(0) \tag{4.8}$$

via

$$A_k = A_{k-1} + [h(k)C_{k-1}+f(k)]\ell_k \tag{4.9}$$

$$B(k) = \gamma_k(k)[h(k)C_{k-1}+f(k)] - \lambda(k)A_k \tag{4.10}$$

$$C_k = C_{k-1} + [h(k)C_{k-1}+f(k)]x_k \tag{4.11}$$

where x_k are defined in (2.7) and (2.9).

Let us also remark that we can avoid the inversion of the matrices L_k in (4.7) by computing their inverses recursively via

$$L_k^{-1} = [I_n+h^T(k)[g(k)+p(k)y_{k-1}]]L_{k-1}^{-1}. \tag{4.12}$$

ALGORITHM 4.1

Let $R = D+S$ be a strongly regular sum of a diagonal matrix D and a semiseparable matrix S of order n such that

$$d_k \neq g(k)h^T(k), \quad k = 0,\dots,N-1.$$

Then one can find the solution of the equation

$$R\chi = f \tag{4.13}$$

in the following way.

1. Use part 1 of the Algorithm 3.1 to find x_k, y_k for $k = 0,\dots,N$.
2. Starting with $\lambda(0)$ and L_0 defined by (4.3) compute recursively for $k = 1,\dots,N$, ℓ_k, L_k, $\lambda(k)$ and L_k^{-1} via (4.5), (4.6), (4.12) and (4.7).
3. Starting with A_0, C_0 and $B(0)$ defined in (4.8) compute recursively for $k = 1,\dots,N$, A_k, C_k and $B(k)$ via (4.9)-(4.11).
4. The solution of the equation (4.13) is given now by

$$\chi = \begin{bmatrix} \lambda(0) \\ \vdots \\ \lambda(N) \end{bmatrix} A_N + \begin{bmatrix} B(0) \\ \vdots \\ B(N) \end{bmatrix}.$$

This algorithm will take $(9n^2+7n)N$ multiplications/divisions and $(9n^2+5n)N$ additions/subtractions.

Suppose now that we have computed the solution χ_N and the matrix R is increased in size by one while preserving the order of semiseparability. Since x_N, y_N, ℓ_N, L_N, L_N^{-1}, $\lambda(N)$, A_N, C_N, $B(N)$ are known, we can find

$$\gamma_{N+1}(N+1) = (d_{N+1}+p(N+1)y_N h^T(N+1))^{-1}$$

$$x_{N+1} = -\gamma_{N+1}(N+1)[y_N h^T(N+1)+q^T(N+1)]$$

$$y_{N+1} = y_N + x_{N+1}[g(N+1)+p(N+1)y_N]$$

$$\ell_{N+1} = -\gamma_{N+1}(N+1)L_N h^T(N+1)$$

$$L_{N+1} = L_N[I_n-\gamma_{N+1}(N+1)h^T(N+1)[g(N+1)+p(N+1)y_N]]$$

$$L_{N+1}^{-1} = [I_n + h^T(N+1)[g(N+1)+p(N+1)y_N]L_N^{-1}$$

$$\lambda(N+1) = \gamma_{N+1}(N+1)[g(N+1)+p(N+1)y_N]L_{N+1}^{-1}$$

$$A_{N+1} = A_N + [h(N+1)C_N+f(N+1)]C_{N+1}$$

$$B(N+1) = \gamma_{N+1}(N+1)[h(N+1)C_N+f(N+1)]-\lambda(N+1)A_{N+1}$$

$$C_{N+1} = C_N + [h(N+1)C_N+f(N+1)]X_{N+1}.$$

The new solution X_{N+1} is now given by

$$X_{N+1} = \begin{bmatrix} \lambda(0) \\ \vdots \\ \lambda(N+1) \end{bmatrix} A_{N+1} + \begin{bmatrix} B(0) \\ \vdots \\ B(N+1) \end{bmatrix}.$$

An easy calculation shows that this update will take $9n^2+7n$ multiplications and $9n^2+5n$ additions/subtractions. The same one update via Algorithm 3.1 would require $2nN+4n^2+4n$ multiplications/divisions and $2nN+4n^2+6n$ additions/subtractions.

5. Concluding Remarks

As in the proof of Theorem 4.1 we can show that the condition

$$\sum_{i=1}^{n} p_i(k)q_i(k) \neq d_k, \quad k = 0,\ldots,N$$

implies that the lower triangular factor of R^{-1} preserves the semiseparability order of R. Thus the following result is true.

THEOREM 5.1. *Let* $R = D+S$ *be a strongly regular sum of a diagonal matrix* $D = \mathrm{diag}\{d_k,\ 0 \le k \le N\}$ *and a semiseparable matrix* S *of order* n, *such that*

$$\sum_{i=1}^{n} h_i(k)g_i(k) \neq d_k, \quad k = 0,\ldots,N$$

$$\sum_{i=1}^{n} p_i(k)q_i(k) \neq d_k, \quad k = 0,\ldots,N.$$

Then R^{-1} *admits the UDL-factorization*

$$R^{-1} = (I+S_U)\mathcal{D}^{-1}(I+S_L)$$

with S_U *upper triangular and* S_L *lower triangular semiseparable matrices each of order* n.

Moreover, the lower triangular factor of R^{-1} can be efficiently computed via recursive formulas similar to (3.6), (3.7), (3.13)

1. Start with

$$\mu_0 = p(0)/d_0, \quad M_0 = I_n.$$

2. Compute recursively for $k = 1,\ldots,N$

$$m_k = -\omega_k(k)M_{k-1}p^T(k)$$

$$M_k = M_{k-1}[I_n-\omega_k(k)p^T(k)][q(k)+h(k)y_{k-1}^T]$$

$$M_k^{-1} = [I_n+p^T(k)[q(k)+h(k)y_{k-1}^T]]M_{k-1}^{-1}$$

$$\mu_k = \omega_k(k)[q(k)+h(k)y_{k-1}^T]M^{-1}(k).$$

3. The lower triangular factor of R^{-1} is given by

$$S_L = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ \mu_0 m_1/\omega_1(1) & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ \mu_0 m_N/\omega_N(N) & \mu_1 m_N/\omega_N(N) & \ldots & 0 \end{bmatrix}.$$

We assume here that the numbers $\omega_k(k)$ and the matrices y_k^T are computed in the course of the Algorithm 2.1.

Regarding the complexity of the derived algorithms, we remark that many calculations there can be performed simultaneously. Therefore the coefficients of n^2N in the evaluation of the complexity of these algorithms can be substantially reduced by the use of parallel processors. In particular the complexity of the Algorithm 5.1 will become $4n^2N$ multiplications/divisions and $4n^2N$ additions/subtractions.

Algorithms 2.1 and 4.1 can be applied only when the matrix R is strongly regular. One way to deal with nonsingular but not strongly regular matrices is to consider instead of the system of equation

$$R\chi = f$$

the equivalent so-called normal system

$$R^*R\chi = R^*f,$$

where R^* is the adjoint. It is easy to see that if R is a sum of a diagonal matrix and a semiseparable matrix of order n, then R^*R is a strongly regular matrix that is a sum of a diagonal matrix and a semiseparable matrix of order at most 4n. However this procedure will square the conditioning of the original problem and more direct methods might be better.

REFERENCES

1. Anderson, B.D. and Moore, J.B., "Optimal Filtering", Prentice-Hall, Inc., Englewood Cliffs, N.J., 1979.

2. Carayanis, G., Kalouptsidis, N. and Manolakis, D., "Fast Recursive Algorithms for a Class of Linear Equations", IEEE Trans. ASSP, Vol. ASSP-30, no. 2, pp. 227-239, 1982.

3. Gevers, M. and Kailath, T., "An Innovations Approach to Least-Squares Estimation, Part VI: Discrete-Time Innovations Representations and Recursive Estimation", IEEE Trans. on Autom. Contr., Vol. AC-18, 6, pp. 588-600, Dec. 1973.

4. Gohberg, I. and Kaashoek, M., "Time Varying Linear Systems with Boundary Conditions and Integral Operators, I: The Transfer Operator and its Properties, Int. E. Op. Theory, Vol. 7, no. 3, pp. 325-391, 1984.

5. Gohberg, I., Kailath, T. and Koltracht, I., "Linear Systems of Equations with Recursive Structure", Lin. Alg. and Appl. 1985, (to appear).

6. Gohberg, I., Kailath, T. and Koltracht, I., "Fast Matrix Factorization", preliminary report, 1984.

7. Gohberg, I. and Koltracht, I., "Numerical Solutions of Integral Equations, Fast Algorithms and Krein-Sobolev Equations", Numerische Mathematik, 1985 (to appear).

8. Kailath, T., "Fredholm Resolvents, Wiener-Hopf Equations, and Riccati Differential Equations", IEEE Trans. on Inform. Theory, IT-15, no. 6, pp. 665-672, November 1969.

9. Kailath, T., "Some New Algorithms for Recursive Estimation in Constant Linear Systems", IEEE Trans. on Inform. Theory, Vol. 19, no. 6, pp. 750-760, 1973.

10. Kailath, T., "Linear Systems", Prentice-Hall, Inc., Englewood Cliffs, N.J., 1980.

11. Kailath, T., Morf, M. and Sidhu, G.S., "Some New Algorithms for Recursive Estimation in Constant Linear Discrete-Time Systems", Proc. Seventh Princeton Conf. on Inform. Sci. & Systs., Princeton, N.J., pp. 344-352, March 1973.

12. Krishna, H. and Morgera, S., "Linear Complexity Fast Algorithms for a Class of Linear Equations", ICASSP, San Diego, CA, 1984.

13. Picinbono, B., "Fast Algorithms for Brownian Matrices", IEEE Trans. ASSP, Vol. ASSP-31, pp. 512-514, April 1983.

I. Gohberg
Department of Mathematical Sciences
Tel Aviv University
Tel Aviv, Israel

T. Kailath
Department of Electrical Engineering
Stanford University
Stanford, CA, U.S.A. 94305

I. Koltracht
Department of Mathematics and Statistics
University of Calgary
Calgary, AB, Canada T2N 1N4

Submitted: September 3, 1985