

# The Schur algorithm for matrices with Hessenberg displacement structure

G. Heinig and V. Olshevsky

September 23, 2000

## Abstract

A Schur-type algorithm for triangular factorization of matrices  $R$  satisfying a displacement equation  $RF - GR = ZY^T$  with small rank matrices  $Y, Z$  is presented for the case that  $F$  and  $G$  are Hessenberg matrices.

## 1 Introduction

Let  $F, G$  be fixed  $n \times n$  matrices. A  $n \times n$  matrix  $R$  is said to possess an  $(F, G)$ -displacement structure if  $r = \text{rank}(RF - GR)$  is small compared with  $n$ . The integer  $r$  is called  $(F, G)$ -displacement rank of  $R$ . In this sense displacement structure was introduced and investigated in [5]. In earlier papers by T. Kailath and his coinvestigators matrices  $R$  were studied for which the rank of  $R - ZZ^*$  is small, where  $Z$  denotes the forward (or backward) shift. Concerning the later and recent development of theory and applications of matrices with displacement structure we refer to the book [3].

Many important classes of structured matrices have a displacement structure for certain  $F$  and  $G$ , for example Toeplitz, Hankel, Toeplitz-plus-Hankel, Vandermonde, and Cauchy matrices.

There are two important types of algorithms for matrices with displacement structure: Levinson-type and Schur-type. In principle, the Schur-type algorithm produces an LU-factorization of the matrix  $R$  and the Levinson-type a factorization of its inverse. These algorithms are called fast since they have a computational complexity of  $O(n^2)$  compared with  $O(n^3)$  for factorization algorithms of unstructured matrices. Note that both algorithms can also be used for fast solutions of linear systems of equations  $Rx = b$ . From computational point of view the Schur-type algorithm seems to be preferable in most cases, mainly because it is more stable, in general.

In [5] Levinson-type algorithms were presented for the case that  $F$  and  $G$  are Hessenberg matrices. In many papers of T. Kailath and his coauthors Schur-type algorithms were constructed (see [3] and references therein). In all these papers it

is assumed that  $G$  and  $F$  are triangular matrices. Interpreting an argument from [8], it is widely believed that this condition is necessary to achieve a recursive triangular factorization (see, for example, [3], p.9, last paragraph).

The main aim of the present paper is to show that, nevertheless, Schur-type algorithms can be constructed if the matrices  $F$  and  $G$  are not triangular but only Hessenberg. The algorithm will be fast if matrix-vector multiplication by the matrices  $F$  and  $G$  can be done fast, for example if  $F$  and  $G$  are also banded.

Let us note that there are some important classes of matrices for which the matrices  $F$  and  $G$  are not triangular. For example, the important class of Toeplitz-plus-Hankel matrices has a  $(F, G)$ -displacement structure for  $F = G = Z + Z^T$  (see [4]). Chebyshev-Vandermonde and more general polynomial Vandermonde matrices have a  $(F, G)$ -displacement structure where one of the matrices is diagonal and the other is Hessenberg (see [2], [6], [7]). Matrices which are Hankel with respect to a certain basis of orthogonal polynomials have a displacement structure for  $F$  and  $G$  being tridiagonal matrices (see [2], [1]).

The starting point for the construction of the Schur-type algorithm is the well-known fact that the displacement structure is inherited somehow in the Schur complement of a principal section of the matrix  $R$ . This fact was most probably observed for the first time in the thesis [9] by M. Morf. Later it was rediscovered by several authors. We derive this fact in a form which is appropriate for us in Section 2. In Section 3 we discuss how to recover the matrix from its displacement, and in Section 4 we describe the algorithm.

## 2 Displacement structure of Schur complement

If  $R$  has  $(F, G)$ -displacement rank  $r$ , then there exist  $n \times r$  matrices  $Z, Y$  such that

$$RF - GR = ZY^T. \quad (2.1)$$

The matrices  $Y$  and  $Z$  in such a representation are called *generators* of  $R$ .

Let  $R$  be decomposed as

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix},$$

where  $R_{11}$  is square, and let  $F, G, Z$ , and  $Y$  be decomposed accordingly. Then the following is true.

**Lemma 2.1** *If  $R$  satisfies (2.1) and  $R_{11}$  is nonsingular, then the Schur complement  $\tilde{R} = R_{22} - R_{21}R_{11}^{-1}R_{12}$  satisfies*

$$\tilde{R}\tilde{F} - \tilde{G}\tilde{R} = \tilde{Z}\tilde{Y}^T,$$

where

$$\begin{aligned}\tilde{Z} &= Z_2 - QZ_1, & \tilde{Y}^T &= Y_2^T - Y_1^T P \\ \tilde{F} &= F_{22} - F_{21}P, & \tilde{G} &= G_{22} - QG_{12} \\ P &= R_{11}^{-1}R_{12}, & Q &= R_{21}R_{11}^{-1}.\end{aligned}$$

*Proof.* From the “magic wand” Schur complement formula

$$R = \begin{bmatrix} I & 0 \\ Q & I \end{bmatrix} \begin{bmatrix} R_{11} & 0 \\ 0 & \tilde{R} \end{bmatrix} \begin{bmatrix} I & P \\ 0 & I \end{bmatrix}$$

we conclude

$$\begin{bmatrix} R_{11} & 0 \\ 0 & \tilde{R} \end{bmatrix} F^\# - G^\# \begin{bmatrix} R_{11} & 0 \\ 0 & \tilde{R} \end{bmatrix} = \begin{bmatrix} I & 0 \\ -Q & I \end{bmatrix} ZY^T \begin{bmatrix} I & -P \\ 0 & I \end{bmatrix}, \quad (2.2)$$

where

$$F^\# = \begin{bmatrix} I & P \\ 0 & I \end{bmatrix} F \begin{bmatrix} I & -P \\ 0 & I \end{bmatrix}, \quad G^\# = \begin{bmatrix} I & 0 \\ -Q & I \end{bmatrix} G \begin{bmatrix} I & 0 \\ Q & I \end{bmatrix}.$$

Note that the right lower block of the  $2 \times 2$  block matrices of  $F^\#$  and  $G^\#$  are  $\tilde{F}$  and  $\tilde{G}$ , respectively. Using this fact, we obtain the displacement formula for  $\tilde{R}$  taking the right lower blocks of both sides of (2.2).  $\square$

### 3 Reconstruction of $R$

In order to construct a fast algorithm for LU-factorization of  $R$  it is necessary to reconstruct the matrix from its displacement  $YZ^T$  and possibly a few more parameters. In the general case there is no fast way to do this. Therefore we restrict ourselves to a special case: We assume that  $F = [f_{ij}]$  is upper and  $G = [g_{ij}]$  is lower Hessenberg, i.e.  $f_{ij} = g_{ji} = 0$  if  $i - j > 1$ . Furthermore we assume that

$$f_{i+1,i} \neq 0 \quad \text{and} \quad g_{i,i+1} \neq 0 \quad (i = 1, \dots, n-1). \quad (3.1)$$

**Lemma 3.1** *Let  $R$  satisfy (2.1), where  $F$  is upper and  $G$  lower Hessenberg. Then the entries of the matrix  $R$  can be reconstructed recursively from the generators  $Z$  and  $Y$  and the first row and column of  $R$  with the help of the formulas*

$$r_{\cdot, k+1} = \frac{1}{f_{k+1,k}} (Gr_{\cdot, k} + Zy_k - \sum_{j=1}^k f_{jk} r_{\cdot, j}), \quad (k = 1, \dots, n-1)$$

where  $y_k$  denotes the  $k$ th column of  $Y^T$ , or

$$r_{k+1, \cdot} = \frac{1}{g_{k,k+1}} (r_{k, \cdot} F - z_k^T Y^T - \sum_{j=1}^k g_{kj} r_{j, \cdot}). \quad (k = 1, \dots, n-1).$$

We will use only the special case  $k = 1$  of Lemma 3.1, which are the formulas

$$r_{\cdot, 2} = \frac{1}{f_{21}} ((G - f_{11}I)r_{\cdot, 1} + Zy_1), \quad r_{2, \cdot} = \frac{1}{g_{12}} (r_{1, \cdot}(F - g_{11}I) + z_1^T Y^T). \quad (3.2)$$

## 4 The algorithm

We assume now that the matrix  $R$  satisfying (2.1) is strongly nonsingular. In this case the matrix admits a unique LDU-factorization

$$R = LDU, \quad (4.3)$$

where  $L$  is lower and  $U$  upper triangular, and  $D$  is diagonal. Instead of working on the matrix  $R$  we will work on its generators. In addition we will use also its first row and column. We present now an algorithm that constructs recursively the generators and first rows and columns of the matrices  $R^{(k)} = [r_{ij}^{(k)}]_{i,j=1}^{n+1-k}$  ( $k = 1, \dots, n$ ), where  $R^{(1)} = R$  and  $R^{(k+1)}$  is the Schur complement of the  $(1, 1)$ -entry  $r_{11}^{(k)}$  of  $R^{(k)}$ . The factors in (4.3) are then easily constructed. In fact, the  $k$ th column  $l_k$  of  $L$  and the  $k$ th row  $u_k$  of  $U$  are given by

$$l_k = \frac{1}{r_{11}^{(k)}} \begin{bmatrix} 0 \\ r_{\cdot,1}^{(k)} \end{bmatrix}, \quad u_k = \frac{1}{r_{11}^{(k)}} \begin{bmatrix} r_{1,\cdot}^{(k)} & 0 \end{bmatrix}.$$

Here “0” stand for a zero vector of appropriate length. The diagonal factor is given by  $D = \text{diag} [r_{11}^{(k)}]_{k=1}^n$ .

ALGORITHM SCHUR-HESSSENBERG

**Initialization:**

$$\begin{aligned} r_{\cdot,1}^{(1)} &= r_{\cdot,1}, & r_{1,\cdot}^{(1)} &= r_{1,\cdot}, \\ Z^{(1)} &= Z, & Y^{(1)} &= Y, & F^{(1)} &= F, & G^{(1)} &= G. \end{aligned}$$

**Recursion:**

For  $k = 1, \dots, n-1$ , compute

1. the  $k$ th entry of  $D$ , the  $k$ th column of  $L$  and  $k$ th row of  $U$  by

$$d^{(k)} = r_{11}^{(k)}, \quad l^{(k)} = \frac{1}{d^{(k)}} r_{\cdot,1}^{(k)}, \quad u^{(k)} = \frac{1}{d^{(k)}} r_{1,\cdot}^{(k)}.$$

2. the second column/row of  $R^{(k)}$  by

$$\begin{aligned} r_{2,\cdot}^{(k)} &= \frac{1}{f_{21}^{(k)}} ((G^{(k)} - f_{11}^{(k)} I) r_{1,\cdot}^{(k)} + Z^{(k)} y_1^{(k)}), \\ r_{\cdot,2}^{(k)} &= \frac{1}{g_{12}^{(k)}} ((F^{(k)} - g_{11}^{(k)} I) r_{\cdot,1}^{(k)} + z_1^{(k)T} Y^{(k)T}). \end{aligned}$$

3. the first column/row of  $R^{(k+1)}$  by

$$r_{1,\cdot}^{(k+1)} = r^{(k)'} - r_{12}^{(k)} l^{(k)'}$$

Here the prime means that the first component is cutted.

4. the new generators by

$$Z^{(k+1)} = Z^{(k)'} - l^{(k)} z_1^{(k)T}$$

5. the new displacement operators

$$F^{(k+1)} = F^{(k)''} - f_{21}^{(k)} e_1 u^{(k)'}, \quad G^{(k+1)} = G^{(k)''} - g_{12}^{(k)} l^{(k)'} e_1^T.$$

The algorithm has in general complexity  $O(n^3)$  in view of the matrix-vector multiplication in Step 2. The complexity reduces if this multiplication can be carried out with less amount. This is the case if  $F$  and  $G$  are tridiagonal matrices. This is our main result.

**Theorem 4.1** *If  $F$  and  $G$  are banded Hessenberg matrices and  $R$  is strongly nonsingular then the Schur-Hessenberg algorithm computes the triangular factorization of  $R$  with complexity  $O(n^2)$ .*

## References

- [1] L. Gemignani, A fast algorithm for generalized Hankel matrices arising in finite-moment problems, *Linear Algebra Appl.* 267 (1997), 41-52.
- [2] G. Heinig, W. Hoppe, K. Rost, Structured matrices in interpolation and approximation problems, *Wiss. Zeitschr. d. TU Karl-Marx-Stadt* 31, 2 (1989), 196-202.
- [3] T. Kailath, A. Sayed, *Fast reliable algorithms for matrices with structure*, SIAM 1999.
- [4] G. Heinig, P. Jankowski and K. Rost, Fast algorithms for Toeplitz-plus-Hankel matrices. *Numerische Mathematik* 52: 628-665 (1988).
- [5] G. Heinig, K. Rost, *Algebraic Methods for Toeplitz-like matrices and operators*, Akademie-Verlag Berlin and Birkhäuser Basel, Boston, Stuttgart, 1984.
- [6] T. Kailath, V. Olshevsky, Displacement structure approach to Chebyshev-Vandermonde matrices, *Integral Equ. and Operator Theory* 22: 65-92 (1995).
- [7] ???
- [8] H. Lev-Ari and T. Kailath, Triangular factorization of structured Hermitian matrices, *Operator Theory: Advances and Appl.* 18: 301-324 (1986).
- [9] M. Morf, Fast algorithms for multivariable systems, Dissertation, Stanford University, Dept. of Electrical Engineering, August 1974.