Systems & Control Letters 18 (1992) 109-120
North-Holland

Direct block tridiagonalization of single-input single-output systems

Gene Golub

Computer Science Department, Stanford University, Stanford, CA 94305, USA

Bo Kågström

Institute of Information Processing, Umeå University, S-901 87 Umeå, Sweden

Paul Van Dooren

Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Received 18 June 1991
Revised 26 September 1991

Abstract In this paper we derive a direct method for block tridiagonalizing a single-input single-output system triple $\{A, b, c\}$. The method is connected to the nonsymmetric Lanczos procedure developed in [10,2,1] and also leads to canonical representations of such triples

Keywords Lanczos procedure, state space, canonical form; numerical algorithms, direct elimination.

1. Introduction

The Lanczos method is a recursive process for tridiagonalizing a given real or complex matrix A via a similarity transformation T. In numerical linear algebra this is typically used for the computation of the eigenvalues of A by exploiting the resulting tridiagonal form. The recurrence of the unsymmetric Lanczos procedure [12] starts with two (arbitrary) vectors b and c and consists of a biorthogonalization of the Krylov sequences

$$\mathscr{C}_n \triangleq [b \quad Ab \quad A^2b \quad \cdots \quad A^{n-1}b], \qquad \mathscr{O}_n \triangleq \begin{bmatrix} c \\ cA \\ cA^2 \\ \vdots \\ cA^{n-1} \end{bmatrix}, \tag{1}$$

where n is the order of the unsymmetric matrix A. Notice that for notational convenience, b is a column vector and c is a row vector. If n is the dimension of the matrix A, then $\mathscr{C}_n$ and $\mathscr{O}_n$ are also known in systems theory as the *controllability matrix* and *observability matrix*, respectively, of the system triplet $\{A,b, c\}$. Also their product $\mathscr{H}_n = \mathscr{O}_n\mathscr{C}_n$ plays an important role in that area. The elements of that matrix are all moments $h_i = cA^{i-1}b$ of the triplet and along the anti diagonals these elements are easily seen to be equal:

$$\mathscr{H}_n \triangleq \mathscr{O}_n \mathscr{C}_n = \begin{bmatrix} cb & cAb & \cdot\cdot & cA^{n-1}b \\ cAb & & & \\ \cdot & & \cdot^{\cdot^{\cdot}} & \\ & \cdot^{\cdot^{\cdot}} & & \vdots \\ cA^{n-1}b & & \cdots & cA^{2(n-1)}b \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & \cdot & h_n \\ h_2 & & & \\ \vdots & & & \\ & & \cdot^{\cdot^{\cdot}} & \vdots \\ h_n & & \cdots & h_{2n-1} \end{bmatrix}. \tag{2}$$

Such a matrix is called a *Hankel matrix* and its properties with respect to the triplet $\{A, b, c\}$ have been widely studied in the system theory literature [7,5,3]. Recently, the relation between the unsymmetric Lanczos procedure in A with starting vectors b and c, and the system theoretic properties of the triplet $\{A, b, c\}$ has been made explicit in a number of papers in order to explain various forms of breakdown of the Lanczos method. As a result of this, modifications of the tridiagonal reduction process were proposed [10,2,1].

In the present paper we develop a *direct transformation method* leading to a block tridiagonal decomposition as in [2]. Our form, though, also displays the fine structure in the blocks of the tridiagonal form, which has direct similarities with properties of the Hankel matrix $\mathscr{H}_n$. Eventually, from this we also derive several (related) canonical forms for a minimal realization of the triplet $\{A, b, c\}$ under similarity transformation $\{T^{-1}AT, T^{-1}b, cT\}$. All derivations presented here are tied to matrix decompositions of the matrices $\mathscr{O}_n$, $\mathscr{C}_n$ and $\mathscr{H}_n$ and to their rank properties. There are of course strong connections with earlier work on canonical forms (see [3,5,7] and further references therein) but the main emphasis here is on algorithmic and numerical aspects. For those we refer to further discussions in the concluding remarks. Moreover, we feel that the direct elimination procedure given below describes the same results in a more straightforward manner.

2. The unsymmetric Lanczos procedure

Here we briefly recall the unsymmetric Lanczos procedure in the case that no breakdown occurs and relate it to matrix decompositions of $\mathscr{C}_n$, $\mathscr{O}_n$ and $\mathscr{H}_n$. Given two linear independent vectors b and c the biorthogonalization of the Krylov sequences $\mathscr{C}_n$ and $\mathscr{O}_n$ (1) can be expressed as follows. Find L and R such that

$$\hat{A} = L^{t}AR = \begin{bmatrix} z_1 & y_2 & & & 0 \\ x_2 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & y_n \\ 0 & & & x_n & z_n \end{bmatrix}. \tag{3}$$

Here, L and R can be generated from the relations $AR = R\hat{A}$, $A^{t}L = L\hat{A}^{t}$ where the coefficients $\hat{a}_{ij}$ are computed to satisfy the biorthogonalization conditions $L^{t}R = I$ (or a diagonal matrix). The decomposition is derived in the following two lemmas where it is also shown that $\hat{A}$ can be chosen tridiagonal and $L^{t} = R^{-1}$.

Lemma 2.1. *Let T be an invertible transformation. Then*

$$T^{-1}b = \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}, \qquad T^{-1}AT = \begin{bmatrix} \times & \cdots & \cdots & \times \\ x_2 & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ 0 & & x_n & \times \end{bmatrix} \quad \textit{with } x_i \neq 0 \tag{4}$$

if and only if

$$T^{-1}\mathscr{C}_n = \begin{bmatrix} c_1 & \times & \cdots & \times \\ & \ddots & \ddots & \vdots \\ & & \ddots & \times \\ 0 & & & c_n \end{bmatrix} \quad \textit{with } c_i \neq 0. \tag{5}$$

Proof. Defining $\hat{b} = T^{-1}b$ and $\hat{A} = T^{-1}AT$, it immediately follows that

$$T^{-1}\mathscr{C}_n = \hat{\mathscr{C}}_n \triangleq \left[\hat{b},\ \hat{A}\hat{b},\ \hat{A}^2\hat{b}, \ldots,\ \hat{A}^{n-1}\hat{b}\right].$$

One easily shows then recursively that the i-th column of $\hat{\mathscr{C}}_n$ has only i non-zero elements and that $c_i = x_1 x_2 \cdots x_i \neq 0$. Conversely, $\hat{b}$ is the first column of $\hat{\mathscr{C}}_n$ and hence $x_1 = c_1 \neq 0$. Then, one easily shows recursively that column i of $\hat{A}$ has only $i+1$ non-zero elements and that $x_{i+1} = c_{i+1}/c_i \neq 0$. □

From the above lemma it follows that T can be chosen to be unitary. Indeed, consider the QR factorization $\mathscr{C}_n = QR$; then $T^{-1} = Q^*$ triangularizes $\mathscr{C}_n$ and hence $Q^*b = \hat{b}$ and $Q^*AQ = \hat{A}$ must have the required form (if $\mathscr{C}_n$ is invertible).

Lemma 2.2. *Let T be an invertible transformation. Then*

$$T^{-1}AT = \begin{bmatrix} \times & y_2 & & & 0 \\ x_2 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & y_n \\ 0 & & & x_n & \times \end{bmatrix}, \quad T^{-1}b = \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}, \tag{6}$$

$$cT = [\,y_1 0 \cdots \cdots \cdots 0\,] \quad \textit{with } y_i \neq 0,\ x_i \neq 0,$$

if and only if

$$T^{-1}\mathscr{C}_n = \begin{bmatrix} c_1 & \times & \cdots & \times \\ & \ddots & \ddots & \vdots \\ & & \ddots & \times \\ 0 & & & c_n \end{bmatrix} \quad \textit{and} \quad \mathscr{O}_n T = \begin{bmatrix} o_1 & & & 0 \\ \times & \ddots & & \\ \vdots & \ddots & \ddots & \\ \times & \cdots & \times & o_n \end{bmatrix} \tag{7}$$

with $c_i \neq 0$, $o_i \neq 0$.

Proof. Apply the above Lemma on $\{A, b\}$ and $\{A^t, c^t\}$. The tridiagonal form of $\hat{A} = T^{-1}AT$ follows from the fact that it is both upper Hessenberg and lower Hessenberg. □

An immediate corollary of the above two lemmas is now the following.

Corollary 2.1. *Let $\mathscr{H}_n$ be the $n \times n$ Hankel matrix (2) corresponding to the system triplet $\{A, b, c\}$ where n is the dimension of A. Then there exists a transformation to a system $\{T^{-1}AT, T^{-1}b, cT\}$ as in (6) if and only if all the leading principal minors of $\mathscr{H}_n$ are non-zero.*

Proof. If there exist such a system $\{T^{-1}AT, T^{-1}b, cT\}$ then

$$\mathscr{H}_n \triangleq \mathscr{O}_n \mathscr{C}_n = \mathscr{O}_n T \cdot T^{-1} \mathscr{C}_n = \begin{bmatrix} o_1 & & \\ \vdots & \ddots & \\ \times & \cdots & o_n \end{bmatrix} \begin{bmatrix} c_1 & \cdot & \times \\ & \cdot & \vdots \\ 0 & & c_n \end{bmatrix} \tag{8}$$

is a LU decomposition of $\mathscr{H}_n$ without pivoting. Since the o_i and c_i are non-zero, all leading principal minors of $\mathscr{H}_n$ must be non-zero.

Conversely, if this property holds then there exists a LU decomposition of $\mathscr{H}_n$ with nonsingular L and U factors:

$$\mathscr{H}_n = L \cdot U.$$

Since $\mathscr{H}_n$ also equals $\mathscr{O}_n \mathscr{C}_n$ we have that $L^{-1}\mathscr{O}_n$ and $\mathscr{C}_n U^{-1}$ must be each others inverse. Taking $T = \mathscr{C}_n U^{-1}$ and $T^{-1} = L^{-1}\mathscr{O}_n$ yields (7) and hence also (6). □

When $\mathscr{H}_n$ does *not* have *all* leading principal minors non-zero, then one can *not* obtain an unreduced tridiagonal form (6), but instead one can always reduce $\{A, b, c\}$ to a *block* tridiagonal form of the same type as (6). As one would expect this is linked to a *block LU* decomposition of $\mathscr{H}_n$ and the block sizes are related to the index set for which the leading principal minors of $\mathscr{H}_n$ are non-zero. A proof of this based on the Lanczos recursion can be found in [10,2]. In the next section we derive a direct algorithm for this decomposition.

When $\mathscr{H}_n$ does *not* have full rank n then either $\mathscr{O}_n$ or $\mathscr{C}_n$ (or both) are singular and $\{A, b, c\}$ is said to be non minimal. For simplicity we assume in the following sections that the system $\{A, b, c\}$ is *minimal* (i.e. $\mathscr{O}_n$ and $\mathscr{C}_n$ are non singular) and we comment on the general case in the concluding remarks.

3. A direct elimination procedure

We start by reducing our minimal system $\{A, b, c\}$ to the form (4) by a *unitary* transformation $T_1 = Q$. The new system $\{Q^*AQ, Q^*b, cQ\} \triangleq \{\hat{A}, \hat{b}, \hat{c}\}$ has an unreduced Hessenberg matrix $\hat{A}$ since $\mathscr{C}_n$ has full rank ($\{A, b\}$ is controllable). The total flop count for this reduction is $\frac{5}{3}n^3$ for the construction of $\{\hat{A}, \hat{b}, \hat{c}\}$ and an extra n^3 for the construction of Q [11].

Since in this coordinate system $\hat{\mathscr{C}}_n$ is upper triangular, one can only allow uppertriangular transformations to further reduce $\{\hat{A}, \hat{b}, \hat{c}\}$ to (possibly block) tridiagonal form. Suppose in this coordinate system $\hat{c}_{k_1}$ is the *first non-zero element* in the row vector $\hat{c}$ (we call this element y_1 in the future). Then partition the system $\{\hat{A}, \hat{b}, \hat{c}\}$ into the first k_1 columns/rows and the remaining $n - k_1$ ones as in Figure 1. We then show in Figure 2 how to eliminate by a unit upper triangular matrix T_2 all of $\hat{c}_2$ and all of $\hat{A}_{12}$ except for the first row. There we marked the order in which the rows in the system matrices are eliminated and the rows of X_2 are calculated by the numbers 1 to k_1. This order is crucial in order to guarantee that previously created zeros are not destroyed afterwards. The pivot elements that are actually used to zero out the consecutive rows are: y_1 for row 1, x_{k_1} for row 2, up to x_2 for row k_1. The total flop count for the transformation T_2 is $\frac{1}{2}nk_1(n - k_1)$ for updating the system triplet and an additional $nk_1(n - k_1)$ for updating the transformation matrix.

The transformation T_2 has now the following effect on the matrices $\hat{\mathscr{C}}_n$ and $\hat{\mathscr{O}}_n$. Since T_2 is unit upper triangular, $T_2^{-1}\hat{\mathscr{C}}_n$ remains upper triangular (in fact its diagonal elements do not change). The matrix $\hat{\mathscr{O}}_n$, on the other hand, was a full matrix and $\hat{\mathscr{O}}_n T_2$ is now block triangular (see Figure 3) with $o_1 = y_1$ and $o_i = o_{i-1} x_{k_1 - i + 2}$ for $i = 2, \ldots, k_1$.

Redefine $\tilde{A} = T_2^{-1}\hat{A}T_2$ partitioned into blocks $\tilde{A}_{11}$, $\tilde{A}_{12}$, $\tilde{A}_{21}$ and $\tilde{A}_{22}$ as in Figure 2. If we now let the first row of $\tilde{A}_{12}$ play the role of a new $\tilde{c}$ vector, the last column of $\tilde{A}_{21}$ play the role of a new b vector

$$\left[\begin{array}{c|c} & \hat{c} \\ \hline \hat{b} & \hat{A} \end{array}\right] \triangleq \left[\begin{array}{c|c|c} & \hat{c}_1 & \hat{c}_2 \\ \hline \hat{b}_1 & \hat{A}_{11} & \hat{A}_{12} \\ \hline 0 & \hat{A}_{21} & \hat{A}_{22} \end{array}\right] =$$

$$\left[\begin{array}{c|cccc|cccc}
 & & & & y_1 & \times & \cdots & \cdots & \times \\ \hline
x_1 & \times & \cdots & \cdots & \times & \times & \cdots & \cdots & \times \\
 & x_2 & \ddots & & \vdots & \times & \cdots & \cdots & \times \\
 & & \ddots & \ddots & \vdots & \vdots & & & \vdots \\
 & & & x_{k_1} & \times & \times & \cdots & \cdots & \times \\ \hline
 & & & & x_{k_1+1} & \times & \cdots & \cdots & \times \\
 & & & & & x_{k_1+2} & \ddots & & \vdots \\
 & & & & & & \ddots & \ddots & \vdots \\
 & & & & & & & x_n & \times
\end{array}\right]$$

Fig. 1

and $\tilde{A}_{22}$ play the role of a new A matrix, then the above procedure can be repeated. If k_2 is the first non-zero element in that row c (we call this element y_2) and if we define $r_1 = k_1$ and $r_2 = k_1 + k_2$, then a similar step embedded into Figure 2 would yield Figure 4 with an updating transformation

$$T_3 = \left[\begin{array}{c|c|c} I_{k_1} & & \\ \hline & I_{k2} & X_3 \\ \hline & & I_{n-r_2} \end{array}\right]. \tag{9}$$

As before, $T_3^{-1}T_2^{-1}\mathscr{C}_n$ is still upper triangular and $\hat{\mathscr{O}}_n T_2 T_3$ has the form shown in Figure 5 with $o_{r_1+1} = o_{r_1} \cdot y_2$ and $o_{r_1+i} = o_{r_1+i-1} \cdot x_{r_2-i+2}$ for $i = 2, \ldots, k_2$.

$$T_2 = \left[\begin{array}{c|c} I_{k_1} & X_2 \\ \hline & I_{n-k_1} \end{array}\right] \begin{array}{l} \leftarrow k_1 \\ \vdots \\ \leftarrow 1 \\ \\ \end{array},$$

$$\left[\begin{array}{c|c} & \hat{c}T_2 \\ \hline T_2^{-1}\hat{b} & T_2^{-1}\hat{A}T_2 \end{array}\right] = \left[\begin{array}{c|cccc|cccc}
 & & & & y_1 & 0 & \cdots & \cdots & 0 \\ \hline
x_1 & \times & \cdots & \cdots & \times & \times & \cdots & \cdots & \times \\
 & x_2 & \ddots & & \vdots & 0 & \cdots & \cdots & 0 \\
 & & \ddots & \ddots & \vdots & \vdots & & & \vdots \\
 & & & x_{k_1} & \times & 0 & \cdots & \cdots & 0 \\ \hline
 & & & & x_{k_1+1} & \times & \cdots & \cdots & \times \\
 & & & & & x_{k_1+2} & \ddots & & \vdots \\
 & & & & & & \ddots & \ddots & \vdots \\
 & & & & & & & x_n & \times
\end{array}\right] \begin{array}{l} \leftarrow 1 \\ \\ \leftarrow k_1 \\ \vdots \\ \leftarrow 2 \\ \\ \\ \\ \\ \end{array}$$

Fig 2

$$\hat{O}_n T_2 = \left[\begin{array}{cccc|cccc} 0 & & & o_1 & & & & \\ & & \cdot^{\cdot^{\cdot}} & \times & & & & \\ & \cdot & \cdot^{\cdot^{\cdot}} & \vdots & & 0 & & \\ o_{k_1} & \times & \cdots & \times & & & & \\ \hline \times & \cdots & \cdots & \times & \times & \cdots & \cdots & \times \\ \vdots & & & \cdot & \vdots & & & \vdots \\ \vdots & & & \vdots & \vdots & & & \vdots \\ \times & \cdots & \cdots & \times & \times & \cdots & \cdots & \times \\ \multicolumn{4}{c}{\underbrace{\hphantom{xxxxxxxxxx}}_{k_1}} & \multicolumn{4}{c}{\underbrace{\hphantom{xxxxxxxxxx}}_{n-k_1}} \end{array}\right] \begin{array}{l} \left.\vphantom{\begin{array}{c}0\\0\\0\\0\end{array}}\right\} k_1 \\ \left.\vphantom{\begin{array}{c}0\\0\\0\\0\end{array}}\right\} n-k_1 \\ \vphantom{0} \end{array}$$

Fig 3

$$\left[\begin{array}{c|c} & \hat{c}\,T_2\,T_3 \\ \hline T_3^{-1}T_2^{-1}\hat{b} & T_3^{-1}T_2^{-1}\hat{A}\,T_2\,T_3 \end{array}\right] =$$

$$\left[\begin{array}{c|cccc|cccc|cccc} & & & & y_1 & & & & & & & & \\ \hline x_1 & \times & \cdots & \cdots & \times & & & & y_2 & & & & \\ & x_2 & \ddots & & \vdots & & & & & & & & \\ & & \ddots & \ddots & \vdots & & & & & & & & \\ & & & x_{r_1} & \times & & & & & & & & \\ \hline & & & & x_{r_1+1} & \times & \cdots & \cdots & \times & \times & \cdots & \cdots & \times \\ & & & & & x_{r_1+2} & \ddots & & \vdots & & & & \\ & & & & & & \ddots & \ddots & \vdots & & & & \\ & & & & & & & x_{r_2} & \times & & & & \\ \hline & & & & & & & & x_{r_2+1} & \times & \cdots & \cdots & \times \\ & & & & & & & & & x_{r_2+2} & \ddots & & \cdot \\ & & & & & & & & & & \ddots & \ddots & \vdots \\ & & & & & & & & & & & x_n & \times \end{array}\right]$$

Fig. 4

$$\hat{O}_n T_2 T_3 = \left[\begin{array}{cccc|cccc|cccc} 0 & & & o_1 & & & & & & & & \\ & & \cdot^{\cdot^{\cdot}} & \times & & & & & & & & \\ & \cdot^{\cdot} & \cdot^{\cdot^{\cdot}} & \vdots & & & & & & & & \\ o_{r_1} & \times & \cdots & \times & & & & & & & & \\ \hline \times & \cdots & \cdots & \times & 0 & & & o_{r_1+1} & & & & \\ & & & \vdots & & & \cdot^{\cdot} & \times & & & & \\ \vdots & & & \vdots & & \cdot^{\cdot} & \cdot^{\cdot} & \vdots & & & & \\ \times & & \cdots & \times & o_{r_2} & \times & \cdots & \times & & & & \\ \hline \times & \cdots & \cdots & \times & \times & \cdot & \cdots & \times & \times & \cdots & \cdots & \times \\ \vdots & & & \vdots & \vdots & & & \cdot & \vdots & & & \vdots \\ \vdots & & & \cdot & \vdots & & & \vdots & \vdots & & & \vdots \\ \times & \cdot & \cdots & \times & \times & \cdots & \cdots & \times & \times & \cdot & \cdots & \times \end{array}\right]$$

Fig. 5

$$\left[\begin{array}{c|c} & \tilde{c} \\ \hline \tilde{b} & \tilde{A} \end{array}\right] \triangleq \left[\begin{array}{c|c} & \hat{c}T \\ \hline T^{-1}\hat{b} & T^{-1}\hat{A}T \end{array}\right] =$$

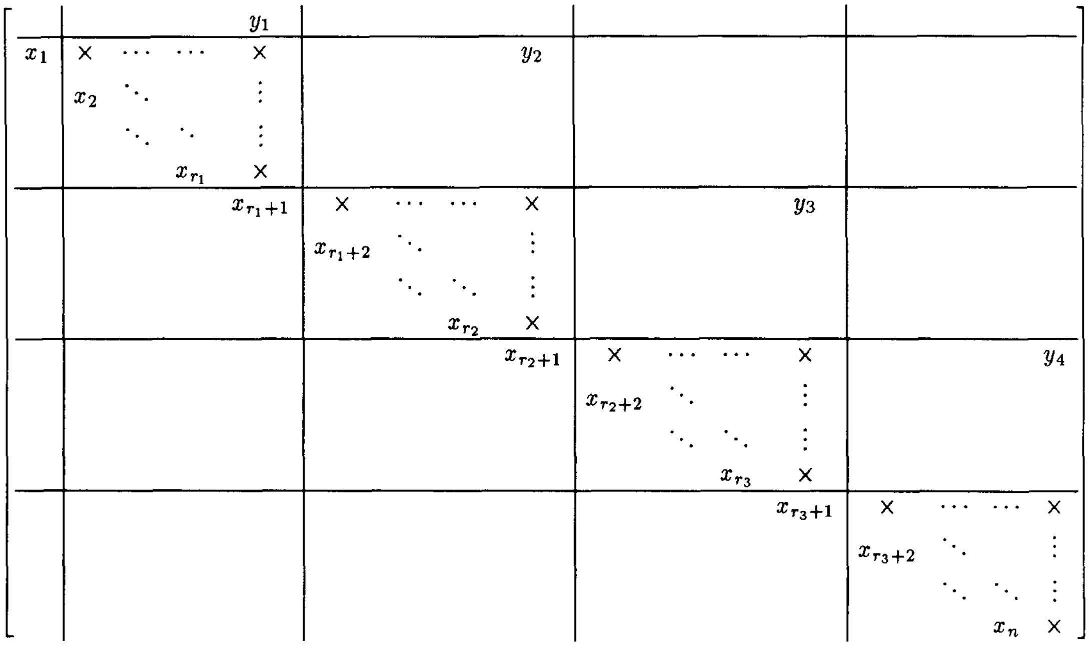

Fig 6

From here on, it is easy to see by induction that the block tridiagonal form shown in Figure 6 is obtained by a product of updating unit upper triangular transformations T_i, which we denote together by the matrix T (we choose $p = 4$ blocks for illustration). Each diagonal block $\tilde{A}_{ii}$ is upper Hessenberg and unreduced. Each off diagonal block has only one non-zero element in the upper right corner. The corresponding $T^{-1}\hat{\mathscr{C}}_n$ is upper triangular and $\hat{\mathscr{O}}_n T$ is block lower triangular. We thus have a block LU factorization of $\mathscr{H}_n$ as illustrated in Figure 7. The 'anti triangular' shape of the $k_i \times k_i$ diagonal blocks of $\hat{\mathscr{O}}_n T$ is easily checked by induction. It also follows from this that the leading principal minors of $\mathscr{H}_n$ are non-zero if and only if those of $\hat{\mathscr{O}}_n T$ are non-zero. Because of the special structure of this block triangular matrix one readily sees that the non-zero leading principal minors are those of dimension

$$r_j = \sum_{i=1}^{j} k_i, \quad j = 1, \ldots, p. \tag{10}$$

The k_i's are thus the 'rank increases' of the nonsingular leading principal submatrices of $\mathscr{H}_n$, as was also observed in [5,10]. Notice that the minimality assumption guarantees that $r_p = n$, which cares for a proper ending of the recursive block reduction.

Remark 3.1. By flipping around the diagonal blocks in Figure 6, i.e. performing a state space transformation

$$T = \operatorname{diag}\left\{ \underbrace{\begin{bmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{bmatrix}}_{k_1}, \ldots, \underbrace{\begin{bmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{bmatrix}}_{k_p} \right\}, \tag{11}$$

$$\mathcal{H}_n = \hat{\mathcal{O}}_n T \cdot T^{-1} \hat{\mathcal{C}}_n =$$

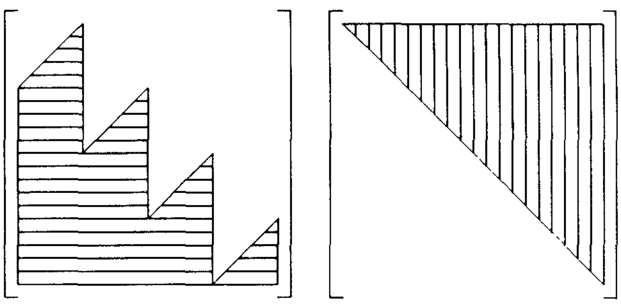

Fig 7

one obtains its dual form, where the role of $\tilde{b}$ and $\tilde{c}$ is interchanged and $\tilde{A}$ is lower Hessenberg. The 'profile' (i.e. the pattern of zero and non-zero elements) of Figure 6 is then just its transpose. Also $\hat{\mathcal{O}}_n T$ is then lower triangular and $T^{-1}\hat{\mathcal{C}}_n$ upper block triangular. Their 'profile' is then each others transpose again. This is used in the various canonical forms presented in the next section

Remark 3.2. If we define $n_0 = n$, $n_i = n_{i-1} - k_i = n - r_i$ then the total operation count for updating the system $\{\hat{A}, \hat{b}, \hat{c}\}$ to $\{\tilde{A}, \tilde{b}, \tilde{c}\}$ is

$$\text{flops} = \sum_{i=1}^{p-1} \frac{n_{i-1} k_i n_i}{2} \tag{12}$$

and twice this amount for accumulating T. The above flop count is in fact maximized when all $k_i = 1$ (the tridiagonal case!) and the flop count is then only $\frac{1}{6}n^3$. Notice that this is much less than the preliminary reduction to $\{\hat{A}, \hat{b}, \hat{c}\}$.

4. Canonical form

In this section we still assume $\{A, b, c\}$ to be minimal. It was shown above that the block tridiagonal form of Figure 6 is intimately connected to the *block LU* decomposition of $\mathcal{H}_n$ shown in Figure 7. What are the degrees of freedom left over for these profiles? From the properties of LU decompositions [4] it is well known that the only degrees of freedom left over in a decomposition with this latter profile is a block diagonal matrix:

$$T_{\text{up}} = \begin{bmatrix} U_1 & & & \\ & U_2 & & \\ & & \ddots & \\ & & & U_p \end{bmatrix} \tag{13}$$

where each block U_i is $k_i \times k_i$ and upper triangular. It is readily seen that the off diagonal blocks in Figure 6 become $U_{i-1}^{-1}\tilde{A}_{i-1,i}U_i$ and $U_i^{-1}\tilde{A}_{i,i-1}U_{i-1}$ respectively, and they both have still only one non-zero element in the top right corner. The diagonal blocks become $U_i^{-1}\tilde{A}_{i,i}U_i$ and their profile can be affected by this, as is shown in the following lemma.

Lemma 4.1 [12]. *An unreduced upper Hessenberg matrix can always be transformed by an upper triangular similarity $U^{-1}HU$ to either one of the two forms*:

$$H_1 = \begin{bmatrix} \alpha_1 & \cdot & \cdot & \cdots & \alpha_r \\ 1 & & & & \\ & \ddots & & & \\ & & & 1 & \end{bmatrix}, \quad H_2 = \begin{bmatrix} & & & \alpha_r \\ 1 & & & \vdots \\ & \ddots & & \vdots \\ & & 1 & \alpha_1 \end{bmatrix}, \tag{14}$$

where the α_i determine the characteristic polynomial of H.

$$\left[\begin{array}{c|c} & cT \\ \hline T^{-1}b & T^{-1}AT \end{array}\right] = \left[\begin{array}{c|ccc|ccc|ccc}
 & & \beta_1 & & & & & & & \\ \hline
1 & \alpha_1 & \cdots\;\cdots & \alpha_{r_1} & & & \beta_2 & & & \\
 & 1 & & & & & & & & \\
 & & \ddots & & & & & & & \\
 & & & 1 & & & & & & \\ \hline
 & & & 1 & \alpha_{r_1+1} & \cdots\;\cdots & \alpha_{r_2} & & & \beta_3 \\
 & & & & 1 & & & & & \\
 & & & & & \ddots & & & & \\
 & & & & & & 1 & & & \\ \hline
 & & & & & & 1 & \alpha_{r_2+1} & \cdots\;\cdots & \alpha_{r_3} \\
 & & & & & & & 1 & & \\
 & & & & & & & & \ddots & \\
 & & & & & & & & & 1
\end{array}\right]$$

Fig 8.

These are also known as *companion* forms of the matrix H. They both determine U completely up to a scalar factor d, which obviously cancels out in $U^{-1}HU$. This factor d in turn can be used to make either the elements of $\tilde{c}$ and $\tilde{A}_{i-1,i}$ equal to 1, or those of $\tilde{b}$ and $\tilde{A}_{i,i-1}$ equal to 1. In each of those two cases p elements are involved which can be put equal to 1 by an appropriate choice of the p scalar factors d_i, still free to choose in the U_i factors of (13). Depending on these choices we now present two types of canonical forms of a minimal system $\{A, b, c\}$ under state space transformations T. We first choose to put the non-zero elements of $\tilde{b}$ and $\tilde{A}_{i,i-1}$ equal to 1. We use $p = 3$ blocks for illustration in Figure 8.

For this form $T^{-1}\mathscr{C}_n$ is *unit upper triangular* and $\mathscr{O}_nT$ is block lower triangular with *anti diagonal* blocks of size $k_i \times k_i$. A variant of this form is to choose the non-zero elements of $\tilde{c}$ and $\tilde{A}_{i-1,i}$ equal to 1. The resulting form is easily checked to be almost identical to Figure 8, except that the β_i elements now occur below the block diagonal. Then $T^{-1}\mathscr{C}_n$ is no more identity on the diagonal but $\mathscr{O}_nT$ has *unit anti diagonal* blocks. Both these variants are connected to decompositions of $\mathscr{H}_n$ of the type described in Figure 9. These are connected to the work of [2] as discussed in the next section.

Another type of canonical form uses the second companion form of Lemma 4.1. Again we use $p = 3$ for illustration in Figure 10. Now $T^{-1}\mathscr{C}_n$ is *block upper triangular with identity* on each diagonal block and $\mathscr{O}_nT$ is block lower triangular with anti triangular blocks of size $k_i \times k_i$. We thus have a decomposition of $\mathscr{H}_n$ of the type shown in Figure 11. The variant here of choosing upper block diagonal elements equal to 1 maintains this profile but the identity matrices in U become diagonal matrices while the anti triangular matrices in L have now 1's on the anti diagonal. For each of these four forms there is of course yet a dual form where the diagonal blocks $\tilde{A}_{ii}$ are lower Hessenberg.

$\mathcal{H}_n = L \cdot U =$

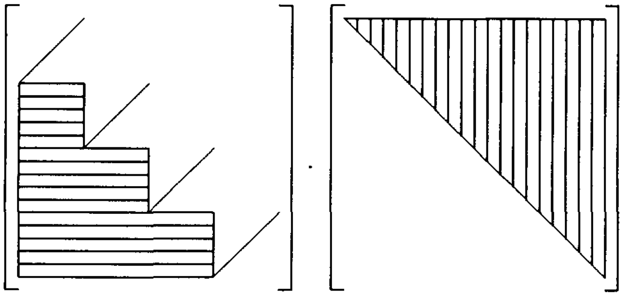

Fig 9.

$$\left[\begin{array}{c|c} & cT \\ \hline T^{-1}b & T^{-1}AT \end{array}\right] = \left[\begin{array}{c|ccc|ccc|ccc}
 & & & \beta_1 & & & & & & \\ \hline
1 & & & \alpha_{r_1} & & & \beta_2 & & & \\
 & 1 & & & & & & & & \\
 & & \ddots & \cdot & & & & & & \\
 & & 1 & \alpha_1 & & & & & & \\ \hline
 & & & 1 & & & \alpha_{r_2} & & & \beta_3 \\
 & & & & 1 & & \vdots & & & \\
 & & & & & \ddots & \cdot & & & \\
 & & & & & 1 & \alpha_{r_1+1} & & & \\ \hline
 & & & & & & 1 & & & \alpha_{r_3} \\
 & & & & & & & 1 & & \vdots \\
 & & & & & & & & \ddots & \vdots \\
 & & & & & & & & 1 & \alpha_{r_2+1}
\end{array}\right]$$

Fig 10

$\mathcal{H}_n = L \quad U =$

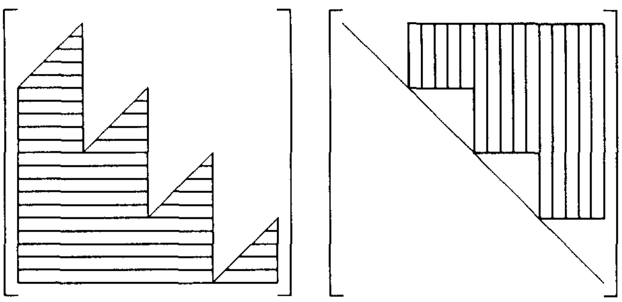

Fig 11

From the above discussions it is readily seen that all these forms can be derived from one another by just moving around block triangular factors and scalings between the L and U factors of $\mathcal{H}_n$ in Figures 9 and 11. These forms are therefore essentially identical. They indeed all have the *same* nonzero elements $\{\alpha_i, i = 1, \ldots, n\}$ and $\{\beta_i, i = 1, \ldots, p\}$. Notice that the form in the work of [10] is in fact related to Figures 10 and 11.

5. Concluding remarks

The approach presented in this paper assumed everywhere that the system $\{A, b, c\}$ is minimal. In a sense this is not a restriction since controllability and observability are implicitly checked by the algorithm. Indeed, the reduction to upper Hessenberg form (Figure 1) would detect the uncontrollable subsystem (when $x_i = 0$) and one would only continue with the controllable subsystem $\{A_c, b_c, c_c\}$. The recurrence (Figures 2-4-6) applied to that subsystem would detect its observable subsystem $\{A_{co}, b_{co}, c_{co}\}$ (when some $y_i = 0$) and one would stop the procedure there. The links with $\mathcal{H}_n$ and the canonical forms are of course still valid for the minimal subsystem $\{A_{co}, b_{co}, c_{co}\}$ of $\{A, b, c\}$. The canonical forms presented in the previous section are not really new since they are essentially to be found in the work e.g. of [7,5,1]. The novelty here is the simplicity of their derivation from elementary matrix operations and of their mutual connection.

The direct matrix derivation presented here also makes sense because of some advantage over the Lanczos type recurrence of [10,2,6]. We therefore discuss the differences below. The complexity of the direct approach is appealing. In the worst case, $\frac{5}{3}n^3 + \frac{1}{6}n^3 = \frac{11}{6}n^3$ flops are needed for calculating $\{\hat{A}, \hat{b}, \hat{c}\}$ in the block form of Figure 6 or even in any of the canonical forms. The calculation of the transformation T itself takes $n^3 + \frac{1}{3}n^3 = \frac{4}{3}n^3$ flops. The Lanczos procedure on the other hand, requires $2n$ matrix products Ab_i and $A^tc_i^t$, but if A is sparse this may be relatively cheap, say $2mn^2$, where m is the number of non-zero elements in one row/column of A. For sparse systems the Lanczos procedure also has the advantage of not requiring to store A as such, whereas the direct method would require this. The biorthogonalization requires in total $2n^3$ flops, but this includes the construction of T and T^{-1}. If full reorthogonalization is *not* requested then this figure can be replaced by $10kn^2$ where k is the average block size of the diagonal blocks. The numerical stability of the direct approach is, normally superior to that of the Lanczos procedure. In the first stage of the direct approach, only unitary transformations are performed. The rank tests there involve the off diagonal elements x_i of the Hessenberg matrix (4) which are computed from orthogonal similarity transformations. In the Lanczos procedure the rank tests involve scalars or block matrices, which are computed from non-unitary transformations of the original $\{A, b, c\}$ triplet. In the direct approach, similar disadvantages are encountered in the second stage as well. The upper triangular transformation eliminating the elements above the diagonal of $\hat{A}$, may be very ill-conditioned. When using a threshold δ, instead of checking for non-zero elements, one can bound the condition number of the updating triangular transformations T_i by some function of δ. This is far less obvious in the Lanczos procedure, where the condition number of the matrices T and T^{-1}, constructed during the process, is far more difficult to monitor.

A final disadvantage of the Lanczos procedure is the multiplication with the A and A^t matrices, which cause instability when the eigenvalues of A have a large spread (when A is 'stiff'). This is e.g. one of the reasons why reorthogonalization is needed in the Lanczos procedure. This drawback is typical for the Lanczos procedure and is *not* encountered in the direct approach. Similar remarks already hold for the first step of our approach, which constructs the Hessenberg form $\{\hat{A}, \hat{b}, \hat{c}\}$ in (9), from which the controllable subsystem can be determined [11]. Any approach based on the Krylov sequence A^ib may suffer from instabilities which do not occur in direct methods and the Lanczos-based methods are therefore to be avoided when controllability is to be detected [9]. But as mentioned in [9,8], trying to transform a system to a canonical form is in general not recommended. For this reason one could stop the reduction procedure in this paper after the block tridiagonal form of Figure 6. Proceeding beyond that involves diagonal blocks in companion form (14), which is known to be delicate [8]. The dense block tridiagonal form is therefore a good compromise when one is only interested in finding the block structure of the Hankel matrix in Figure 7. All this does not preclude possible instabilities in the block tridiagonalization. But the big advantage here is that one can easily check when instabilities may occur. The pivots y_i and x_i are precisely the elements to be inverted in subsequent steps of the algorithm, and putting a threshold on those will keep future transformations bounded. In short, we feel that the direct approach is probably the most sensible and flexible way to proceed towards one of the forms described in this paper.

References

[1] D. Boley, S Elhay, G. Golub and M Gutknecht, The Lanczos algorithm and reconstructing indefinite weights, Intern Report, Stanford University (1990)

[2] D. Boley and G. Golub, The nonsymmetric Lanczos algorithm and controllability, Intern. Report, Stanford University (1990)

[3] P. Fuhrmann, On the partial realization problem and the recursive inversion of Hankel and Toeplitz matrices, in R Brualdi et al., Eds , *Linear Algebra and its Role in Linear System Theory*, Special Issue of Contemporary Mathematics Vol 47 (Amer Math Soc., Providence, RI, 1985)

[4] G. Golub and C. Van Loan, *Matrix Computations*, 2nd Ed (Johns Hopkins Univ Press, Baltimore, MD, 1989).

[5] W. Gragg, A. Lindquist, On the partial realization problem, *Linear Algebra Appl* **50** (1983) 277–319.

[6] M Gutknecht, A completed theory for the Lanczos algorithm, submitted for publication (1989).

[7] R Kalman, On partial realizations, transfer functions and canonical forms, *Acta Polytech Scand* **31** (1979) 9–32

[8] C Kenney and A Laub, Controllability and stability radii for companion form systems, *Math Control Signals Systems* **1** (1988) 239–256

[9] C Paige, Properties of numerical algorithms related to computing controllability, *IEEE Trans Automat Control* **26** (1981) 130–138

[10] B Parlett, Reduction to tridiagonal form and minimal realizations, submitted for publication (1989)

[11] P Van Dooren and M Verhaegen, On the use of unitary state space transformations, in R Brualdi et al., Eds , *Linear Algebra and its Role in Linear System Theory*, Special Issue of Contemporary Mathematics Vol 47 (Amer Math Soc , Providence, RI, 1985) pp 447–463

[12] J Wilkinson, *The Algebraic Eigenvalue Problem* (Oxford Univ Press, Oxford, 1965)