NORTH-HOLLAND

On a Condensed Form for Normal Matrices Under Finite Sequences of Elementary Unitary Similarities

L. Elsner
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
33501 Bielefeld, FRG

and

Kh. D. Ikramov*
Faculty of Computational Mathematics and Cybernetics
Moscow State University
119899 Moscow, Russia

Submitted by Alex Pothen

ABSTRACT

It is generally known that any Hermitian matrix can be reduced to a tridiagonal form by a finite sequence of unitary similarities, namely Householder reflections. Recently A. Bunse-Gerstner and L. Elsner have found a condensed form to which any unitary matrix can be reduced, again by a finite sequence of Householder transformations. This condensed form can be considered as a pentadiagonal or block tridiagonal matrix with some additional zeros inside the band. We describe such a condensed form (or, more precisely, a set of such forms) for general normal matrices, where the number of nonzero elements does not exceed $O(n^{3/2})$, n being the order of the normal matrix given. Two approaches to constructing the condensed form are outlined. The first approach is a geometrical Lanczos-type one where we use the so-called generalized Krylov sequences. The second, more constructive approach is an elimination process using Householder reflections. Our condensed form can be

*Supported by Sonderforschungsbercich 343 *Diskrete Strukturen in der Mathematik*.

LINEAR ALGEBRA AND ITS APPLICATIONS 254:79–98 (1997)

0024-3795/97/$17.00
PII S0024-3795(96)00526-5

thought of as a variable-bandwidth form. An interesting feature of it is that for normal matrices whose spectra lie on algebraic curves of low degree the bandwidth is much smaller.

1. INTRODUCTION

The main incentive of the present paper has been the desire to find for any normal matrix A a condensed form which contains as many zeros as possible and, as opposed to the diagonal form, is reachable by a *finite* sequence of elementary unitary transformations. One can immediately point to such a form, namely the Hessenberg form. The problem with this form is that one does not use normality to obtain a Hessenberg matrix from the initial matrix A. This is even true for the case when A is a unitary matrix, i.e., the reduction to Hessenberg form exploits in no way this special property of A. Unitarity of an upper Hessenberg matrix has however been exploited for the eigenvalue problem. We mention [2] as an example of a whole series of papers on this subject. Meanwhile, in the recent paper [3] the authors have shown how to modify, for a unitary matrix A, the Householder reduction in such a way as to obtain the new condensed form. Compared with the Hessenberg form, it is much more symmetrical in profile and much less populated. In fact, this form can be considered as a pentadiagonal or block tridiagonal matrix with some additional zeros inside the band.

2. A GENERALIZED LANCZOS PROCEDURE

It is well known that another way to arrive at the Hessenberg form is the Arnoldi process, where one is dealing with vector sequences of the form

$$v,\ Av,\ A^2v, \dots,\ A^{m-1}v \tag{1}$$

and their linear spans, the so-called Krylov subspaces

$$\mathscr{K}_m(A, v) = \operatorname{span}\{v,\ Av, \dots,\ A^{m-1}v\}. \tag{2}$$

Again in this approach normality cannot be exploited, since it amounts to a special relation between A and its conjugate A^*, and Krylov subspaces (2) do not take A^* into account.

Below we introduce a generalization of the Arnoldi procedure based on using vectors

$$v,\ Av,\ A^*v,\ A^2v,\ A^*,\ Av,\ AA^*v,\ A^{*2}v,\ A^3v, \dots \tag{3}$$

instead of vectors (1). The general description of the sequence (3) can be given in terms of *words in two variables* s and t [6, p. 75]. Such a word is an expression of the form

$$W(s,t) = s^{m_1}t^{n_1}s^{m_2}t^{n_2}\dots s^{m_k}t^{n_k}, \tag{4}$$

with integers $m_1, n_1, \dots, m_k, n_k \geqslant 0$. The *degree* of the word $W(s,t)$ is the number $m_1 + n_1 + \cdots + m_k + n_k$. We order the words first according to their degree. Words of the same degree are ordered lexicographically. If $A \in C^{m \times n}$ is given, we define

$$W(A, A^*) = A^{m_1}(A^*)^{n_1}A^{m_2}(A^*)^{n_2}\dots A^{m_k}(A^*)^{n_k}. \tag{5}$$

All the vectors of the sequence (3) have the form

$$u_k = W_k(A, A^*)v. \tag{6}$$

The index k refers to the ordering of the words as outlined above. This describes (3) more precisely. Here $u_1 = v$ corresponds to the only word W_1 of degree 0, vectors $u_2 = Av$ and $u_3 = A^*v$ are associated with the words $W_2 = s$ and $W_3 = t$ of degree 1, and so on. Generally, vectors u_k, $2^m \leqslant k < 2^{m+1}$ correspond to all the words of degree m. We call them the *mth layer* of the sequence (3).

For a normal matrix A, the number of different vectors in a layer is drastically reduced. Since A and A^* commute, all the vectors in the mth layer can be written as

$$u = A^{\alpha}(A^*)^{\beta}v, \qquad \alpha + \beta = m.$$

So there are only $m + 1$ vectors in this layer. This case will be dealt with more extensively in the next section.

The subspace

$$L_m(A, v) = \operatorname{span}\{W(A, A^*)v : \operatorname{degree}(W) \leqslant m\} \tag{7}$$

is called the *mth generalized Krylov subspace*. We denote its dimension by l_m, and call $w_m = l_m - l_{m-1}$ the width of the mth layer. Here we set formally $w_0 = 1$. It is obvious that for normal A one has $w_m \leqslant m + 1$. The width of the mth layer can be less than $m + 1$. We give two examples below.

EXAMPLE 1. For a Hermitian matrix A we have $Av = A^*v$; therefore $w_1 = 1$. Moreover, for any m, all the vectors in the mth layer coincide with the vector $A^m v$, and $w_m = 1$ as long as $A^m v$ is not a linear combination of the previous Krylov vectors. The same conclusion, i.e., $w_m \leqslant 1$ for any m_1 holds for a slightly more general class of normal matrices, namely for the matrices A of the form

$$A = \alpha H + \beta I, \qquad H = H^*. \tag{8}$$

EXAMPLE 2. For a unitary matrix A we have generically $w_1 = 2$. On the other hand, the vector AA^*v in the second layer is just v; therefore $w_2 \leqslant 2$. It is easily seen that $w_m \leqslant 2$ for any m. The same is true for the matrices A of the form

$$A = \alpha V + \beta I, \qquad VV^* = I. \tag{9}$$

In the Lanczos-type procedure below we construct a sequence of orthogonal vectors $v_1, v_2, v_3, \dots$ from the sequence (6). Having found the orthogonal and nonzero vectors $v_1, \dots, v_m$, we set

$$\hat{\mathscr{K}}_m = \operatorname{span}\{v_1, \dots, v_m\}.$$

If Pw is the orthogonal projection of a vector w onto a subspace $\mathscr{K}$, then $w - Pw$ is the perpendicular from w on $\mathscr{K}$ and is denoted by

$$\operatorname{orth}_{\mathscr{K}} w.$$

ALGORITHM 1 (The generalized Lanczos procedure).

1. Choose a random nonzero vector v. Let $v_1 = v$.
2. Assume that orthogonal and nonzero vectors $v_1, \dots, v_m$ have already been found, which constitute an orthogonal basis of the linear span of the first k_m vectors $u_1, \dots, u_{k_m}$ of the sequence (6). If the perpendicular

$$\operatorname{orth}_{\hat{\mathscr{K}}_m} u_{k_m+1}$$

is nonzero, take it as the vector v_{m+1}. Otherwise, try

$$\operatorname{orth}_{\hat{\mathscr{K}}_m} u_{k_m+2},$$

and, if it is nonzero, take it as v_{m+1}, and so on.

Observe that we do not require our *generalized Lanczos vectors* $v_1, \dots, v_m$ to be normalized.

REMARK. If the perpendiculars from the vectors u_k of the current layer are all zero, then $\hat{\mathscr{K}}_m$ is a (common) invariant subspace of A and A^*. In this case, we stop the procedure or repeat it with a new initial vector $\hat{v}$ orthogonalized with respect to $\hat{\mathscr{K}}_m$. We give an illustration of how the procedure above works for two particular classes of normal matrices.

EXAMPLE 3. For a Hermitian matrix A, the subspaces (2) coincide with the usual Krylov subspaces, and Algorithm 1 is essentially the classical Lanczos algorithm. The same assertion holds for all matrices of the form (8).

EXAMPLE 4. As was indicated in Example 2, it is enough, for a unitary matrix A and even for a matrix A of the slightly more general form (9), to deal with two usual Krylov sequences

$$v, Av, A^2v, \dots$$

and

$$A^*v, (A^*)^2v, \dots,$$

switching alternatively from one to the other. If we have

$$A^s v \in \hat{\mathscr{K}}_m, \qquad (A^*)^s v \in \hat{\mathscr{K}}_m$$

for $m = 2s - 1$, then $\hat{\mathscr{K}}_m$ is an A-invariant subspace.

From our construction the following two properties of the generalized Krylov subspaces (7) can be derived easily:

1. For any $x \in L_m$ we have

$$Ax \in L_{m+1}, \qquad A^*x \in L_{m+1}.$$

2. For any generalized Lanczos vector v_l belonging to the mth layer, i.e. $v_l \in L_m \setminus L_{m-1}$, the orthogonality relations

$$Av_l \perp L_{m-2}, \qquad A^*v_l \perp L_{m-2}$$

hold.

Indeed, if $y \in L_{m-2}$ then

$$(Av_l, y) = (v_l, A^* y) = 0$$

and

$$(A^* v_l, y) = (v_l, Ay) = 0,$$

since $A^* y$ and $Ay \in L_{m-1}$ and v_l is orthogonal to L_{m-1}.

The vectors in the $(s+1)$th layer of the sequence (3) are obtained by applying A and then A^* to the vectors of the sth layer. Using this observation and property 2 above, we can restate Algorithm 1 in a more Lanczos-like style.

ALGORITHM 2.

1. Choose a random nonzero vector v. Let $v_1 = v$.
2. Assume that the orthogonal and nonzero vectors $v_1, \dots, v_m$ have been found which constitute an orthogonal basis of the sth generalized Krylov subspace. Suppose that the vector $v_q, \dots, v_m$ have been constructed by (implicitly) using the sth layer of the sequence (3). Then, for each of these vectors in turn, do the following steps:
   (a) Evaluate $w = Av$.
   (b) Orthogonalize w with respect to already accepted Lanczos vectors v_i belonging to the $(s-1)$th, sth, and $(s+1)$th layers.
   (c) If, after step (b), the vector w is nonzero, take it as the most recent Lanczos vector of the $(s+1)$th layer.'
   (d) For the vector v in (a) evaluate $w = A^* v$.
   (e) For the vector w in (d) repeat steps (b) and (c) above.

We see that the vectors in the $(s+1)$th layer of the sequence (3) do not appear directly in Algorithm 2. Nevertheless, by applying A and A^* to the vectors $v_q, \dots, v_m$ we are still using (3) implicitly. Suppose the orthonormal basis of C^n consisting of the generalized Lanczos vectors $v_1, \dots, v_n$ (which we now assume to be normalized) has been constructed. We relate a linear operator $\mathscr{A}: C^n \Rightarrow C^n$ to the initial normal matrix A. Denote by B the matrix associated with $\mathscr{A}$ in the Lanczos basis $v_1, \dots, v_n$. We observe that B is a block tridiagonal matrix with diagonal blocks of sizes w_i, $i = 0, 1, \dots$. It follows from Algorithm 2 that the number N_i of nonzeros in the ith column of B can be determined by this rule: find the index s of the layer to which v_i belongs; then the nonzero elements in b_{*i} can correspond only to the

Lanczos vectors v_k in the $(s-1)$th, sth, and $(s+1)$th layers. Therefore,

$$N_i \leqslant w_{s-1} + w_s + w_{s+1}. \tag{10}$$

The same bound is valid for the number M_i of nonzeros in the ith row of B:

$$M_i \leqslant w_{s-1} + w_s + w_{s+1}. \tag{11}$$

If, in Algorithm 2, we only obtain an orthogonal basis $v_1, \dots, v_p$ of some invariant subspace L of $\mathscr{A}$, then the bounds (10), (11) hold for the matrix B_1 of the induced operator $\mathscr{A}|_{\mathscr{L}}$ in this basis. So B (or B_1) is a variable-band-width matrix.

3. THE NORMAL CASE

We restrict ourselves to the normal case in the following. Here $w_i \leqslant i + 1$, and hence by (10) the number of nonzero elements in each row and column of B is considerably reduced. We illustrate this by exhibiting the pattern of zeros in the general case and the normal case for 8×8 matrices:

$$\begin{pmatrix} * & * & * & 0 & 0 & 0 & 0 & 0 \\ * & * & * & * & * & 0 & 0 & 0 \\ 0 & * & * & * & * & * & * & 0 \\ 0 & * & * & * & * & * & * & * \\ 0 & 0 & * & * & * & * & * & * \\ 0 & 0 & * & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * & * & * \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} * & * & * & 0 & 0 & 0 & 0 & 0 \\ * & * & * & * & * & 0 & 0 & 0 \\ 0 & * & * & * & * & * & 0 & 0 \\ 0 & * & * & * & * & * & * & * \\ 0 & 0 & * & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & * & * & * & * \end{pmatrix}.$$

This is not yet very impressive. However, we will see later that for normal matrices whose eigenvalues lie on a algebraic curve even more zeros appear. For eigenvalues on an ellipse the following pattern arises:

$$\begin{pmatrix} * & * & * & 0 & 0 & 0 & 0 & 0 \\ * & * & * & * & * & 0 & 0 & 0 \\ 0 & * & * & * & * & 0 & 0 & 0 \\ 0 & * & * & * & * & * & * & 0 \\ 0 & 0 & * & * & * & * & * & 0 \\ 0 & 0 & 0 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & * & * & * \end{pmatrix}.$$

We establish below a bound for the maximal bandwidth in B. For this purpose, we will investigate the behavior of the numbers w_i in more detail.

We first remind the reader of some useful definitions and facts. The vectors $u_1, u_2, \dots, u_s \in C^n$ are said to be linearly dependent over the linear subspace $L \subset C^n$ if there exist numbers $\alpha_1, \alpha_2, \dots, \alpha_s$, not all zero, such that

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_s u_s \in L. \tag{12}$$

We might write instead

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_s u_s = 0 \pmod{L}. \tag{13}$$

If this relation is possible only with $\alpha_1 = \alpha_2 = \cdots = \alpha_s = 0$, then $u_1, u_2, \dots, u_s$ are linearly independent over L.

Assume that $u_{k_1}, \dots, u_{k_t}$ is a subset of $u_1, \dots, u_s$ maximal with respect to linear independence over L. Letting, for simplicity, $k_1 = 1, \dots, k_t = t$, we can then represent $u_{t+1}, \dots, u_s$ as linear combinations (over L) of the vectors $u_1, \dots, u_t$:

$$\begin{aligned} u_{t+1} &= c_{t+1,1} u_1 + \cdots + c_{t+1,t} u_t \pmod{L}, \\ &\vdots \\ u_s &= c_{s1} u_1 + \cdots + c_{st} u_t \pmod{L}. \end{aligned}$$

On the other hand, if q linear relations are given,

$$\begin{aligned} d_{11} u_1 + d_{12} u_2 + \cdots + d_{1s} u_s &= 0 \pmod{L}, \\ &\vdots \\ d_{q1} u_1 + d_{q2} u_2 + \cdots + d_{qs} u_s &= 0 \pmod{L}, \end{aligned} \tag{14}$$

and these relations are independent (which means that rank of the matrix

$$\begin{pmatrix} d_{11} & \cdots & d_{1s} \\ \vdots & & \vdots \\ d_{q1} & \cdots & d_{qs} \end{pmatrix}$$

is equal to q), then no more than $s - q$ of the vectors $u_1, \ldots, u_s$ can be independent over L. The reason is that (14) permits us to express q of the vectors $u_1, \ldots, u_s$ as linear combinations (over L) of the rest of u's. We are already now to prove the following statement.

THEOREM 1. *There exists an integer $m_o \geqslant 0$ such that*

$$w_m = m + 1, \qquad m \leqslant m_0, \tag{15}$$

$$w_{m+1} \leqslant w_m, \qquad m \geqslant m_0. \tag{16}$$

Proof. The relation (15) holds as long as all the vectors

$$A^m v, A^{m-1}A^* v, \ldots, (A^*)^m v, \tag{17}$$

which constitute the mth layer of the sequence (3), are linearly independent over the $(m - 1)$th generalized Krylov subspace L_{m-1}. Assume now that there exist exactly q linear relations for the vectors (17),

$$d_{i1} A^m v + d_{i2} A^{m-1}A^* v + \cdots + d_{i,m+1}(A^*)^m v = 0 \pmod{L_{m-1}} \tag{18}$$

for $i = 1, \ldots, q$, which are independent. Applying A to each of the relations, we have

$$d_{i1} A^{m+1} v + d_{i2} A^m A^* v + \cdots + d_{i,m+1} A(A^*)^m v = 0 \pmod{L_m} \tag{19}$$

for $i = 1, \ldots, q$. These are q independent linear relations for the vectors

$$A^{m+1} v, A^m A^* v, \ldots, A(A^*)^m v, (A^*)^{m+1} v \tag{20}$$

in the $(m + 1)$th layer of (3). If in (18) $s \leqslant m + 1$ is the maximal index for which $d_{is} \neq 0$ for some $i \in 1, \ldots, q$, then multiplying the ith equation of (18) by A^* gives a new relation for (20). This is independent of the equations (19), as it is the only one containing explicitly the term $A^{m+1-s}(A^*)^s v$. Therefore, no more than $m + 2 - q - 1 = m + 1 - q = w_m$ vectors of the $(m + 1)$th layer are linearly independent over L_m. This shows (16). Choose now m_0 as the maximal integer for which (15) holds. Then in the next layer there exists at least one linear relation; hence from now on (16) holds. ■

REMARK 1. In Example 7 in Section 4 we give an illustration of the case when $m_0 \ll n$. An easy consequence of Theorem 1 is

COROLLARY 2. *For any* m,

$$w_m < \sqrt{2n}\,, \tag{21}$$

where n *is the order of the normal matrix* A.

Proof. Obviously $w = w_{m_0} = m_0 + 1$ is the maximal member of the sequence $\{w_m\}$. Then, according to the proof of Theorem 1, the vectors of the first m_0 layers are linearly independent and hence

$$\sum_{i=0}^{m_0} (i+1) = \frac{w(w+1)}{2} \leqslant n.$$

This implies (21). ∎

We return now to the matrix B constructed from A by the generalized Lanczos procedure. We call this matrix the *condensed form of the matrix* A. This notion is not quite correct, as B depends also on the starting vector v. If, in Algorithm 2, only a proper invariant subspace of A is found, then the results below hold for the corresponding matrix B_1 of order n_1 smaller than n.

COROLLARY 3. *The number of nonzero entries of each row and column in the condensed form* B *of the normal matrix* A *is bounded by* $3\sqrt{2n}$.

Proof. This assertion is an immediate consequence of the bounds (21), (10), and (11). ∎

We remark however that with some efforts we can improve this bound to $\sqrt{8n-15}$ for $n > 2$.

COROLLARY 4. *The number of nonzeros in the condensed form* B *is bounded by* $\sqrt{18}\,n^{3/2}$.

Proof. This follows immediately from Corollary 3. ∎

REMARK. Of course, there might be only $O(n)$ nonzero elements, for example for Hermitian or unitary matrices. On the other hand, $O(n^{3/2})$ might be quite a realistic bound in a situation when in no layer do we have linear dependencies between the generalized Krylov vectors (3). Indeed, our condensed form contains disjoint square blocks of order $1, 2, \dots, w$ where $w = w_{m_0} = O(\sqrt{n})$. Therefore, the number of nonzeros of B cannot be less than

$$1^2 + 2^2 + \cdots + w^2 = \frac{w(w+1)(2w+1)}{6} = O(n^{3/2}).$$

4. AN ELIMINATION PROCEDURE FOR THE CONDENSED FORM

We are going now to give another algorithm for reducing a general matrix A to the condensed form B described above. In contrast to Algorithms 1 and 2, this one is based on using Householder transformations.

We need to introduce some notation (e.g. [3]). For any $A \in C^{n \times n}$ its columns and rows are denoted by the corresponding small letter as $a_{*1}, a_{*2}, \dots, a_{*n}$ and $a_{1*}, a_{2*}, \dots, a_{n*}$, respectively. For $1 \leqslant j < n$ we denote by $P(j, z)$ the Householder transformation which eliminates the entries $j+1$ through n in the vector z and does not change the first $j-1$ entries of z.

ALGORITHM 3.

1. Let $l_0 = 2$, $k = 1$. As long as $k \leqslant n$, do the following.
2. Let $m = (k+1)/2$. Check the entries $l_{k-1} + 1$ through n in the column a_{*m} of the current matrix A. If all these entries are zero, let $l_k = l_{k-1}$, $k \to k+1$, and go to step 3. Otherwise, find the Householder transformation $P_k = P(l_{k-1}, a_{*m})$ and perform the similarity
$$A \to P_k A P_k. \tag{22}$$
Let $l_k = l_{k-1} + 1$. If $l_k = n$, go to step 4; otherwise, $k \to k+1$ and go to step 3.
3. Let $m = k/2$. Check the entries $l_{k-1} + 1$ through n in the row a_{m*} of the current matrix A. If all these entries are zero, let $l_k = l_{k-1}$, $k \to k + 1$, and go to step 2. Otherwise, find the Householder transformation $P_k = P(l_{k-1}, a_{m*})$ and perform the similarity (22). Let $l_k = l_{k-1} + 1$. If $l_k = n$, go to step 4; otherwise, $k \to k+1$ and go to step 2.
4. Stop the procedure.

It can be shown formally (see e.g. [9] and [4]) that the resulting matrix is up to scaling the condensed form B determined by Algorithm 2 with the first column of P_1 as starting vector. We will not use this result here, but explain the behavior of Algorithm 3 directly.

This algorithm applied to a general matrix A will generically eliminate $n - 2$ entries in the first stage, $n - 3$ in the second stage, and so on, so that it stops after $n - 2$ stages, having eliminated $(n - 1)(n - 2)/2$ elements. This is completely analogous to the transformation to Hessenberg form.

In this form, however, normality of A has a strong effect on the algorithm. This is not obvious. We therefore hasten to show how normality does influence it. We do that at first for the particular classes (8) and (9), and then for some related classes, and also for general normal matrices.

Example 5. Let A at the beginning of Algorithm 3 be a matrix of the form (8). After the first stage of the algorithm we have $l_1 = 3$ and the current column a_{*1} has zero entries 3 through n. Then (8) shows that the current row a_{1*} has also zero entries 3 through n. Therefore, we let $l_2 = 3$, and do no similarity at the second stage of the procedure. At the third stage, we eliminate entries 4 through n in the column a_{*2}, and let $l_3 = 4$. Again by (8) we have that in the current row a_{2*} entries 4 through n are already zero. So we let $l_4 = 4$ and pass to the fifth stage, and so on. The result is a tridiagonal matrix.

Example 6. Let A at the entry of Algorithm 3 be a matrix of the form (9). Then the first three stages run according to the general prescription with $l_1 = 3$, $l_2 = 4$, $l_3 = 5$. Now, if A is just unitary, then already after the second stage the appearance of additional zeros could be registered. Indeed, in the column a_{*1} entries 3 through n are zero. Setting the reducible case aside, we can count on entry 2 being nonzero. Since, after the second stage, columns 4 through n have the first entry zero, their orthogonality with a_{*1} means that the second entry in each of these columns is also zero. In fact, this is exactly what has been mentioned on p. 753 of [3]. It follows that, excepting the assignment $l_4 = 5$, we can omit the fourth stage. But the fifth stage can also be omitted, again with the exception of the assignment $l_5 = 5$. This follows from the orthogonality of a_{1*} and rows 5 through n, assuming again that the $(1, 3)$ entry is not zero. After that, the sixth stage and the seventh one go along the general rules, and so on. All that is said above is also true for a matrix A of the form (9), since such a matrix essentially differs only by a diagonal shift from a purely unitary one.

EXAMPLE 7. Consider the plane quadric of the form

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 = 1. \tag{23}$$

Here $a_{11}, a_{12}, a_{22} \in R$, and the curve (23) is a hyperbola if

$$a_{11}a_{22} - a_{12}^2 < 0,$$

and an ellipse if

$$a_{11} > 0, \qquad a_{11}a_{22} - a_{12}^2 > 0.$$

Letting

$$x = \frac{z + \bar{z}}{2}, \qquad y = \frac{z - \bar{z}}{2i},$$

we can obtain the complex form of Equation (23):

$$cz^2 + \bar{c}\bar{z}^2 + 2dz\bar{z} = 4. \tag{24}$$

Here

$$c = a_{11} - a_{22} - (2a_{12})i, \qquad d = a_{11} + a_{22}.$$

Now, we can define a class of normal matrices by the relations

$$AA^* = A^*A, \qquad cA^2 + \bar{c}A^{*2} + 2dAA^* = 4I. \tag{25}$$

These are exactly normal matrices with eigenvalues belonging to the curve (24). Unitary matrices are just a particular case of this definition corresponding to $c = 0$, $d = 2$ [i.e., $a_{11} = a_{22} = 1$, $a_{12} = 0$ in (23)].

Let us see now how Algorithm 3 proceeds when applied to a general normal matrix A. The first four stages run according to the general description, and $l_1 = 3$, $l_2 = 4$, $l_3 = 5$, $l_4 = 6$. It is easily seen that, after the fourth stage is completed, the column a_{*1} is orthogonal to columns 6 to n. We assume that the entry $a_{1,3}$ is nonzero. Now, by normality the row a_{1*} is also orthogonal to rows 6 to n, so entries 6 to n of the column a_{*3} are zero and we can skip the fifth stage, letting only $l_5 = 6$. Stage 9 is the next one in which due to normality no elimination takes place. Here the column a_{*2} is orthogonal to columns 9 to n; by normality the same holds for the corresponding rows, and, assuming $a_{2,5} \neq 0$, we get that entries 9 to n of column 5 are zero. Hence this elimination step is skipped, leaving only $l_9 = 9$.

Similarly, in stages 11, 15, 17, 19, 23 no elimination takes place provided that the entries $a_{3,6}, a_{4,8}, a_{5,9}, a_{6,10}, a_{7,12}$ are nonzero. The vanishing of one of those entries indicates that $w_m \leq m$ (and hence by Theorem 1 also $w_k \leq m$ for all $k > m$) where m is the index of the current layer. Let us explain this: Denote by $\mathscr{L}_k$ the linear span of the first k coordinate vectors $e_1, \dots, e_k$ in C^n. Obviously $Ae_1 \in \mathscr{L}_2$. Now $a_{1,3} = 0$ means $A^*e_1 \in \mathscr{L}_2$ also, so that $w_1 \leqslant 1$. Similarly $a_{2,5} = 0$ gives $A^*e_2 \in \mathscr{L}_4$. Hence $A^*Ae_1 \in A^*\mathscr{L}_2 \subset \mathscr{L}_4$, which shows $\dim \operatorname{span}\{e_1, Ae_1, A^*e_1, A^2e_1, A^*Ae_1\} \leqslant 4$. This implies $w_2 \leqslant 2$.

Let us now discuss the behavior of Algorithm 3 for matrices of the class (25). Taking into consideration the form the matrix A assumes after the first stages, we have

$$Ae_1 \in \mathscr{L}_2, \qquad A^*e_1 \in \mathscr{L}_3, \qquad A^2e_1 \in A\mathscr{L}_2 \subset \mathscr{L}_4, \qquad AA^*e_1 \in A\mathscr{L}_3 \subset \mathscr{L}_5. \tag{26}$$

Now we are able to prove that entries 6 through n of the third row of A are zero. Assuming the (1, 3) entry is nonzero (otherwise $w_1 \leqslant 1$ and the statement is true anyhow), we can write

$$e_3 = \alpha A^*e_1 + \beta e_1 + \gamma e_2 \tag{27}$$

with $\alpha \neq 0$. We multiply this relation by A^*. This leads to

$$A^*e_3 = \alpha A^{*2}e_1 + \beta A^*e_1 + \gamma A^*e_2 \in \mathscr{L}_5. \tag{28}$$

In the last implication we have used that by (25) and $c \neq 0$

$$A^{*2}e_1 = \rho A^2e_1 + \sigma AA^*e_1 + \tau e_1 \in \mathscr{L}_5 \tag{29}$$

holds for some ρ, σ, and τ. By (28) our claim is proved.

This fact could easily be predicted. Indeed, Equation (25) implies the linear dependence (over L_1) of the vectors in the second layer of the sequence (3), which means that $w_2 = 2$. Recall that this equality implies

$$w_m \leqslant 2 \qquad \text{for all} \quad m \geqslant 2.$$

Therefore, the condensed form B of a matrix A from the class (25) is a band matrix with the bandwidth $\leqslant 7$.

5. NORMAL NEARLY HERMITIAN MATRICES

Another example where A and A^* satisfy a quadratic equation is the following. Suppose that a given real normal matrix A has exactly one pair of complex conjugate eigenvalues $a \pm ib$, all the other eigenvalues being real. Then the pair of lines

$$(x - a)y = 0$$

contains the spectrum of A. The complex form of this equation is

$$z^2 - \bar{z}^2 = 2a(z - \bar{z}),$$

which implies that A itself satisfies the matrix relation

$$A^2 - A^{*2} = 2a(A - A^*).$$

This means that

$$w_2 \leqslant 2$$

and, according to Theorem 1,

$$w_m \leqslant 2, \qquad m > 2.$$

Similar considerations apply to the case of r complex conjugate pairs of eigenvalues, where r is not too large.

In fact, in the examples of this section we have not just stabilization of bandwidth, but a real decrease. The following theorem implies that the condensed form is finally tridiagonal. More precisely, $w_m \leqslant 1$ for $m > 2r$.

THEOREM 2. *Let A be normal,*

$$A = S + K, \qquad S = S^*, \qquad K = -K^*, \qquad SK = KS, \tag{30}$$

and

$$k = \operatorname{rank} K < \frac{n-1}{2}.$$

Then $w_m \leqslant 1$ for $m > k$, so that the condensed form of A has a tridiagonal tail.

Proof. Instead of studying linear dependences of the generalized Krylov sequence

$$v,\ Av,\ A^*v,\ A^2v,\ AA^*v,\ A^{*2}v,\ A^3v, \dots,$$

we do the same for the matrix sequence

$$I,\ A,\ A^*,\ A^2,\ AA^*,\ A^{*2},\ A^3, \dots$$

and observe that for $m = 0, 1, \dots$ the linear span of the matrices in the mth layer is also the span of the matrices

$$S^m, S^{m-1}K, \dots, SK^{m-1}, K^m.$$

This replacement of the vector sequence by the matrix sequence is justified by the fact that for any v in whose eigenvector expansion no coefficient vanishes, linear dependence between terms of the vector sequence is equivalent to linear dependence between the corresponding terms of the matrix sequence. This shows also that for normal A important parameters of the condensed forms, namely the sizes of the diagonal blocks w_i, do not generically depend on the initial vector v.

As S and K can be simultaneously diagonalized, we can easily see that the span of all products S^rK^s, $s > 0$, $r \geqslant 0$ is a linear subspace $\mathscr{L}$ of $C^{n,n}$ of dimension at most k. It can be less than k if eigenvalues coincide. The span of all those products in the first i layers is denoted by

$$M_i = \operatorname{span}\{S^{t-s}K^s, 0 < s \leqslant t \leqslant i\}$$

and can be generated by the recursion

$$M_0 = \{0\}, \qquad M_i = \operatorname{span}\{M_{i-1}, KM_{i-1}, S^{i-1}K\}, \quad i > 0.$$

We claim

$$M_i = M_{i+1} \quad \Rightarrow \quad M_{i+1} = M_{i+2}. \tag{31}$$

Obviously $M_i = M_{i+1}$ if and only if $KM_i \subset M_i$ and $S^iK \in M_i$. But then

$$KM_{i+1} = KM_i \subset M_i$$

and also

$$S^{i+1}K = SS^iK \in SM_i \subset M_{i+1} = M_i.$$

Hence

$$M_{i+2} = \operatorname{span}\{M_{i+1}, KM_{i+1}, S^{i+1}K\} \subset M_i,$$

and (31) follows.

If we define now p to be the smallest natural number for which $M_p = M_{p+1}$, then by the previous result $p \leqslant k$. Each of the further layers gives at most one new independent matrix, namely the corresponding power of S. Hence $w_m \leqslant 1$ for $m \geqslant p + 1$. The dimension of $L_m(A, v)$ is not more than (and is generically equal to) the dimension of

$$\operatorname{span}\{S^{t-s}K^s, 0 \leqslant s \leqslant t \leqslant p\} = \operatorname{span}\{M_p, I, S, \ldots, S^p\}$$

and the latter is at most $k + p + 1 < n$, so that there are 1-by-1 diagonal blocks at the right lower corner of the condensed form. ■

Let us illustrate Theorem 2 in the case of a real normal matrix with r pairs of conjugate complex eigenvalues. If $r = 1$ and $a \pm ib$ are the two nonreal eigenvalues, then K is a skew-Hermitian matrix commuting with A with eigenvalues $\pm ib$ and $n - 2$ zero eigenvalues, hence of rank 2. In the second layer $SK = aK$, so that K, K^2 span the space $\mathscr{L}$. This shows $w_1 = w_2 = 2$, and $w_m \leqslant 1$ for $m > 2$.

For $r = 2$ and nonreal eigenvalues $a_j \pm ib_j$, $j = 1, 2$, with $a_1 \neq a_2$, $b_1 \neq b_2$ S and K are given in a suitable basis by

$$S = \operatorname{diag}(a_1, a_1, a_2, a_2, \lambda_5, \dots, \lambda_n), \quad iK = \operatorname{diag}(b_1, -b_1, b_2, -b_2, 0, \dots, 0)$$

From this it is easy to see that the matrices K, SK, K^2, SK^2 form a basis of $\mathscr{S}$. Hence we have $w_0 = 1$, $w_1 = 2$, $w_2 = 3$, $w_3 = 2$, $w_m \leqslant 1$ for $m > 3$.

We mention finally that the considerations of this section originated from a question posed to one of us by K. Veselić. Let H be a real normal unreduced upper Hessenberg matrix. If H has only one pair of conjugate complex eigenvalues, can this fact be recognized through the form of H? We have just seen that our condensed form reacts by shrinking its bandwidth to 3. On the contrary, the Hessenberg form need not to convert itself into a matrix with reduced upper bandsize. This can be shown by counterexamples for any dimension n.

6. CONCLUDING REMARKS

REMARK 1. In Theorem 2 we have seen that the condensed form of a Hermitian matrix S perturbed by a very special low-rank matrix K [see (30)] does not differ very much from the tridiagonal form, which is the condensed form of S. The problem of more general perturbations, still of low rank, will be treated among others in a subsequent paper.

REMARK 2. In conclusion, we show that the condensed form B of a normal matrix A is invariant under the QR algorithm (see e.g. [7] or [11]) given that no shift coincides with an exact eigenvalue of A. This result is folklore; see e.g., [3, p. 764], where however no proof is given. We give here a short proof. Denote by B_1 the matrix obtained from B through one step of some version of the QR algorithm. Then

$$B_1 = RBR^{-1} = Q^*BQ, \tag{32}$$

where Q and R are a unitary and a right triangular factor, respectively, in the QR decomposition of some polynomial function in B,

$$f(B) = QR.$$

For example, $f(A) = A - \tau I$ for the single-step QR algorithm with τ as a shift, and $f(A) = A^2 - (\tau_1 + \tau_2)A + \tau_1\tau_2 I$ for the double-step QR algorithm with the shifts τ_1 and τ_2.

Notice that the left equality in (32) shows that the lower envelope of B_1 (for a definition of the envelope in the symmetric case see [6]) is the same as that for B. Now, for any function φ, (32) implies

$$\varphi(B_1) = R\varphi(B)R^{-1} = Q^*\varphi(B)Q. \tag{33}$$

It is well known that the normality of B amounts to B^* being a polynomial in B; see e.g., [8, p. 110]. Therefore

$$B^* = \varphi_0(B)$$

for some polynomial φ_0, and since B_1 is unitarily similar to B, we have also

$$B_1^* = \varphi_0(B_1).$$

Using (33) with $\varphi = \varphi_0$, we obtain

$$B_1^* = RB^*R^{-1}.$$

This means that the upper envelope of B_1 (i.e., the lower envelope of B_1^*) is the same as that for B.

This result does not contradict the fact that nonsymmetric tridiagonal matrices do not stay tridiagonal under the QR algorithm. Also, it implies that normal band matrices are invariant under this algorithm. For other forms of symmetric matrices invariant under the QR algorithm we refer to [1].

REMARK 3. Finally, we would like to clarify the relation our results have to those by Vojevodin [10] and Faber and Manteuffel [5]. In terms of our paper the problem these authors treat is as follows: Describe the matrices that can be brought, through a unitary similarity, to Hessenberg form with few nonzero diagonals in the upper part. Their answer is: these matrices are normal with the spectrum on a line, if a single nonzero upper diagonal is allowed (Faber and Manteuffel), and normal with A^* being a polynomial of degree s in A, if s nonzero upper diagonals are allowed (Vojevodin). In contrast to this, we search for a condensed form more symmetric relative to the main diagonal. Bandform with bandwidth > 3 is one of the possibilities. Such a form is reachable even if A^* is not a polynomial of low degree in A. An example is unitary matrices, which by [3] can be reduced to pentadiagonal form. On the other hand, for an $n \times n$ unitary matrix U with simple spectrum, no polynomial in the equation

$$U^* = U^{-1} = f(U)$$

can have a degree lower than $n - 1$.

REMARK 4. There have been some numerical experiments with Algorithms 2 and 3. In the Diploma thesis [9] by Stefanie Krause it was shown that Algorithm 3 applied to normal matrices did work quite satisfactorily, also finding the small bandform, when the eigenvalues were constructed to lie on an ellipse. However, no proof of stability has been given. As was to be expected, Algorithm 2 required the use of reorthogonalizations and hence turned out to be numerically inferior to the elimination algorithm.

REFERENCES

1 P. Arbenz and G. H. Golub, Matrix shapes invariant under the symmetric QR algorithm, *Numer. Linear Algebra Appl.* 2:87–93 (1995).
2 G. S. Ammar, W. B. Gragg, and L. Reichel, On the eigenproblem for orthogonal matrices, in *Proceedings of the 25th IEEE Conference on Decision and Control*, Athens, 1986.
3 A. Bunse-Gerstner and L. Elsner, Schur parameter pencils for the solution of the unitary eigenproblem, *Linear Algebra Appl.* 154–156:741–778 (1991).
4 L. Elsner and Kh. D. Ikramov, On nonnormal matrices that can be reduced to band form under finite sequences of elementary unitary similarities, Preprint 96-008, SFB 343, Bielefeld.
5 V. Faber and T. Manteuffel, Necessary and sufficient conditions for the existence of a conjugate gradient method, *SIAM J. Numer. Anal.* 21:352–362 (1984).
6 J. A. George and J. W. Liu, *Computer Solution of Large Sparse Positive Definite Linear Equations*, Prentice-Hall, 1981.
7 G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed., Johns Hopkins U.P., 1989.
8 R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge U.S., 1985.
9 S. Krause, Transformationen normaler Matrizen auf eine verdichtete Bandgestalt, Diplomarbeit, Fakultät für Mathematik, Bielefeld, 1996.
10 V. V. Vojevodin, The problem of non-self-adjoint extensions of the conjugate gradient method is closed (in Russian), *Zh. Vychisl. Mat. Mat. Fiz.* 23(2):447–479 (1983).
11 D. S. Watkins, *Fundamentals of Matrix Computations*, Wiley, 1991.

Received 12 September 1995; final manuscript accepted 25 September 1996