ON A CONDENSED FORM FOR NORMAL MATRICES UNDER FINITE SEQUENCES OF UNITARY SIMILARITIES

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

Kh. D. Ikramov[1]
Faculty of Computational Mathematics
and Cybernetics
Moscow State University
119899 Moscow
Russia

[1]Research supported by Sonderforschungsbereich 343 *Diskrete Strukturen in der Mathematik*

Summary

It is generally known that any Hermitian matrix can be reduced to a tridiagonal form by a finite sequence of elementary 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. Now the natural question arises, if there exists a condensed form for general normal matrices similar to some extent to the Bunse–Gerstner and Elsner form for the unitary case. In this paper we describe such a form (or, more precisely, a set of such forms) and show that the number of nonzero elements does not exceed $O(n^{3/2})$ where n is the order of the normal matrix given. We outline two approaches to constructing the condensed form. The first approach is a geometrical Lanczos–type one where we use the so–called generalized Krylov sequences. The second, more constructive approach, relies on a sequence of specially chosen Householder transformations.

Our condensed form can be thought of as a variable bandwidth form. It stays therefore invariant under Gaussian elimination without interchanges. We show that, for a normal matrix, this form is also invariant under the QR–algorithm.

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 at 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. It 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. Meanwhile, in the recent paper [Bunse–Gerstner, Elsner] 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.

It is well–known that another way to arrive at the Hessenberg form is some Lanczos–type or, rather, Arnoldi–type process where one is dealing with vector sequences of the form

$$v,\ Av,\ A^2v, \dots, A^{m-1}v \tag{1.1}$$

and their linear spans, the so–called Krylov subspaces

$$\mathcal{K}_m(A,v) = \operatorname{span}\{v, Av, \dots, A^{m-1}v\}. \tag{1.2}$$

Again in this approach normality cannot be exploited since it amounts to a special relation between A and its conjugate A^*, and Krylov subspaces (1.2) do not take A^* into account.

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

$$v, Av, A^*v, A^2v, A^*Av, AA^*v, A^{*2}v, A^3v, \dots \tag{1.3}$$

instead of vectors (1.1). The general description of the sequence (1.3) can be given in terms of **words in two variables** s and t [Horn, Johnson, p. 75]. Such a word is an expression of the form

$$W(s,t) = s^{m_1}\ t^{n_1}\ s^{m_2}\ t^{n_2} \dots\ s^{m_k}\ t^{n_k}, \tag{1.4}$$

with integers $m_1, n_1, \dots, m_k,\ n_k \geq 0$.

The degree of the word $W(s,t)$ is the number $m_1 + n_1 + \dots + m_k + n_k$. If $A \in \mathbb{C}^{n\times n}$ is given then a word in A and A^* is defined as

$$W(A, A^*) = A^{m_1}(A^*)^{n_1}A^{m_2}(A^*)^{n_2}\dots A^{m_k}(A^*)^{n_k}. \tag{1.5}$$

All the vectors of the sequence (1.3) have the form

$$u_k = W_k(A, A^*)v. \tag{1.6}$$

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 $u_k,\ 2^m \leq k < 2^{m+1} - 1$ correspond to all the words of degree m . We call them the **m–th layer** of the sequence (1.3).

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 m–th layer can be written as

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

So there are at most $m+1$ different vectors in this layer.

Definition. The subspace

$$L_m(A, v) = \text{ span}\{ \bigcup_{\text{degree } W \leq m} W(A, A^*)v\} \tag{1.7}$$

is called the **m–th generalized Krylov subspace**. We denote its dimension by ℓ_m , and call $w_m = \ell_m - \ell_{m-1}$ the width of the m–th layer, where we set formally $w_0 = 1$. It is obvious that for a normal matrix A

$$w_m \leq m + 1.$$

The width of the m–th 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 m–th 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., $w_m \leq 1$ for any m , holds for a slightly more general class of normal matrices, namely for the matrices A of the form

$$A = \alpha\ H + \beta\ I,\ \ H = H^*. \tag{1.8}$$

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 nothing else than v , therefore $w_2 \leq 2$. It is easily seen that $w_m \leq 2$ for any m . The same is true for the matrices A of the form

$$A = \alpha\ V + \beta\ I,\ \ VV^* = I. \tag{1.9}$$

In the Lanczos–type procedure below we construct a sequence of orthogonal vectors $v_1, v_2, v_3, \ldots$ from the sequence (1.3). Having found the orthogonal and nonzero vectors $v_1, \ldots, v_m$ we set

$$\widehat{\mathcal{K}}_m = \text{ span}\{v_1, \ldots, v_m\}. \tag{1.10}$$

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

$$\text{orth}_{\mathcal{K}}\ w.$$

We do not require our "generalized Lanczos vectors" $v_1, \ldots, v_m$ to be normalized.

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 (1.3). If the perpendicular

$$\mathrm{orth}_{\widehat{\mathcal{K}}_m}\; u_{k_m+1}$$

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

$$\mathrm{orth}_{\widehat{\mathcal{K}}_m}\; u_{k_m+2},$$

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

Remark. If the perpendiculars from the vectors u_k of the current layer are all zero, then $\widehat{\mathcal{K}}_m$ is a (common) invariant subspace of A and A^* . In this case, we stop the procedure, or repeat it with a new initial vector $\widehat{v}$ orthogonalized with respect to $\widehat{\mathcal{K}}_m$.

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 (1.10) 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 (1.8).

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 (1.9), to deal with two usual Krylov sequences

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

and

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

switching alternatively from one to another. If, for the current s–th layer, we have

$$A^s v \in \widehat{\mathcal{K}}_m,\ (A^*)^s\ v \in \widehat{\mathcal{K}}_m$$

then $\widehat{\mathcal{K}}_m$ is an A–invariant subspace.

From our construction the following two properties of the generalized Krylov subspaces (1.7) can be derived easily:

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

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

2. For any generalized Lanczos vector v_l belonging to the m–th layer the orthogonality relations

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

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 (1.3) are obtained by applying A and then A^* to the vectors of the s–th 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 $v_1, \ldots, v_m$ have been found which constitute an orthogonal basis of the s–th generalized Krylov subspace. Suppose that the vectors $v_q, \ldots, v_m$ have been constructed by (implicitly) using the s–th layer of the sequence (1.3). 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, s–th 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 (1.3) do not appear directly in Algorithm 2. Nevertheless, by applying A and A^* to the vectors $v_q, \ldots, v_m$ we are still using (1.3) implicitly.

Suppose the orthogonal basis of the whole space $\mathbb{C}^n$ has been constructed consisting of the generalized Lanczos vectors $v_1, \ldots, v_n$. We relate a linear operator $\mathcal{A} : \mathbb{C}^n \Rightarrow \mathbb{C}^n$ to the initial normal matrix A . Denote by B the matrix associated with $\mathcal{A}$ in the Lanczos base $v_1, \ldots, v_n$. It follows from Algorithm 2 that the number N_i of nonzeros in the i–th column b_i of B can be determined by this rule: find 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, s–th and (s+1)–th layers. Therefore,

$$N_i \le w_{s-1} + w_s + w_{s+1}. \tag{1.11}$$

The same bound is valid for the number M_i of nonzeros in the i–th row of B :

$$M_i \leq w_{s-1} + w_s + w_{s+1}. \tag{1.12}$$

If, in Algorithm 2, we only obtain an orthogonal basis $v_1, \ldots, v_p$ of some invariant subspace L of A then bounds (1.11), (1.12) hold for the matrix B_1 of the induced operator $A_{|_L}$ in this basis.

So, B (or B_1) is a variable bandwidth matrix. We establish below a bound for the maximal bandwidth in B . For this purpose we will investigate the behaviour of the numbers w_i in more detail.

We first remind the reader of some useful definitions and facts. The vectors $u_1,\ u_2, \ldots, u_s \in \mathbb{C}^n$ are said to be linear dependent over the linear subspace $L \subset \mathbb{C}^n$ if there exist numbers $\alpha_1,\ a_2, \ldots, \alpha_s$ not all of them zero, such that

$$\alpha_1\ u_1 + \alpha_2\ u_2 + \ldots + \alpha_s\ u_s \in L. \tag{1.13}$$

We might write

$$\alpha_1\ u_1 + \alpha_2\ u_2 + \ldots + \alpha_s\ u_s = 0 \pmod{L} \tag{1.14}$$

instead of (1.13). If (1.14) is possible only with $\alpha_1 = \alpha_2 = \ldots = \alpha_s = 0$ then $u_1,\ u_2, \ldots, u_s$ are linear independent over L .

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

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

On the other hand, if q linear relations are given

$$\begin{aligned} &d_{11}\ u_1 + d_{12}\ u_2 + \ldots + d_{1s}\ u_s = 0 \pmod{L} \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &d_{q1}\ u_1 + d_{q2}\ u_2 + \ldots + d_{qs}\ u_s = 0 \pmod{L} \end{aligned} \tag{1.15}$$

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

$$D = \begin{bmatrix} d_{11} & \ldots & d_{1s} \\ \ldots & \ldots & \ldots \\ d_{q1} & \ldots & d_{qs} \end{bmatrix}$$

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 relation (1.15) permits us to express some q of the vectors $u_1, \ldots, u_s$ as linear combinations (over L) of the rest of $u's$.

We are ready now to prove the following statement.

Theorem 1. For any m ,

$$w_{m+1} \leq w_m + 1. \tag{1.16}$$

Proof. Inequality (1.16) is trivial when all the vectors

$$A^m v,\ A^{m-1}A^*\ v, \dots , (A^*)^m\ v \tag{1.17}$$

which constitute the m–th layer of the sequence (1.3) are linear independent over the $(m-1)$ –th generalized Krylov subspace L_{m-1} . For, in this case, $w_m = m+1$, and w_{m+1} cannot exceed the overall number $m+2$ of vectors in the $(m+1)$ –th layer of (1.3).

Assume now that there exist exactly q linear relations for the vectors (1.17)

$$\begin{aligned} &d_{11}A^m\ v + d_{12}\ A^{m-1}A^*\ v + \dots + d_{1,m+1}(A^*)^m\ v = 0 \ (\mathrm{mod}\ L_{m-1}) \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &d_{q1}\ A^m\ v + d_{q2}\ A^{m-1}A^*\ v + \dots + d_{q,m+1}\ (A^*)^m\ v = 0 \ (\mathrm{mod}\ L_{m-1}) \end{aligned} \tag{1.18}$$

which are independent. Applying A to each of the relations we have

$$\begin{aligned} &d_{11}\ A^{m+1}\ v + d_{12}\ A^m A^*\ v + \dots + d_{1,m+1}\ A(A^*)^m\ v = 0 \ (\mathrm{mod}\ L_m) \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &d_{q1}\ A^{m+1}\ v + d_{q2}\ A^m A^*\ v + \dots + d_{q,m+1}\ A(A^*)^m\ v = 0 \ (\mathrm{mod}\ L_m). \end{aligned} \tag{1.19}$$

These are q independent linear relations for the vectors

$$A^{m+1}\ v,\ A^m A^*\ v, \dots , A(A^*)^m\ v,\ (A^*)^{m+1}\ v \tag{1.20}$$

in the $(m+1)$ –th layer of (1.3). Therefore, no more than $m+2-q$ of these vectors are linear independent over L_m . Now,

$$w_m = m + 1 - q,\ w_{m+1} \leq m + 2 - q = w_m + 1$$

which proves the theorem.

Remark. Suppose now that in the relations (1.18) at least one of the coefficients $d_{1,m+1}, \dots , d_{q,m+1}$ is nonzero. If, for example, $d_{i,m+1} \neq 0$ then, applying A^* to the i–th of relations (1.18) we have

$$d_{i1}\ A^m A^*\ v + \dots + d_{im}\ A(A^*)^m\ v + d_{i,m+1}\ (A^*)^{m+1}\ v = 0. \tag{1.21}$$

This is another linear relation for the vectors (1.20) which is obviously independent of relations (1.19). It follows that

$$w_{m+1} \leq m + 1 - q = w_m. \tag{1.22}$$

Since A and A^* enter symmetrically in all this theory inequality (1.22) holds also in case of at least one of the coefficients $d_{11}, \dots , d_{q1}$ is nonzero.

As relation (1.21) shows the condition at the beginning of this remark remains fulfilled at the next layer. Therefore, if linear relation of the type (1.18) first appears at the m_0 –th layer, and $A^{m_0}v$ (or $(A^*)^{m_0}v$) enters this relation with a nonzero coefficient then for all subsequent layers

$$w_m \leq w_{m_0},\ m = m_0 + 1,\ m_0 + 2, \dots . \tag{1.23}$$

Theorem 2. For any m,

$$w_m \le \sqrt{2n}, \tag{1.24}$$

where n is the order of the matrix A.

Proof. Let $w = w_{m_0}$ be the maximal member of the sequence $\{w_m\}$. Then, according to Theorem 1, there exists an index m_1 such that

$$m_1 < m_0 \ \text{ and } \ w_{m_1} = w - 1,$$

there exists an index m_2 such that

$$m_2 < m_1 \ \text{ and } \ w_{m_2} = w - 2,$$

and so on. Finally, we can find an index m_{w-1} such that

$$m_{w-1} < m_{w-2} \ \text{ and } \ w_{m_{w-1}} = 1.$$

For example, one can let $m_{w-1} = 0$. Now we have

$$1 + 2 + \ldots + w - 1 + w = \frac{w(w+1)}{2} = w_{m_{w-1}} + \ldots + w_{m_1} + w_{m_0} \le \sum w = n.$$

So,

$$w^2 \le w(w+1) \le 2n$$

from which the bound (1.24) follows.

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. 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 maximal bandwidth in the condensed form B of the normal matrix A is bounded by $3\sqrt{2n}$.

Proof. This assertion is a simple consequence of the bounds given in (1.24),(1.11) and (1.12).

Corollary 4. The number of elements inside the envelope of the condensed form B (i.e., **the profile** of B; see definition of the profile in [George, Liu]) is bounded by $O(n^{3/2})$.

Proof. The envelope of B is contained in the band with the constant bandwidth $3\sqrt{2n}$. This gives us the upper bound in Corollary 4. Of course, the true profile might only be $O(n)$ as is the case, for example, with Hermitian or unitary matrices. On the other hand, $O(n^{3/2})$ might be quite a realistical bound in a situation when in no layer we have linear dependencies between the generalized Krylov vectors (1.3). Indeed, in this case our envelope contains disjoint square blocks of order $1, 2, \ldots, w$ where $w = w_{\max} = O(\sqrt{n})$. Therefore, the profile of B cannot be less than

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

Remark. Essentially the same reasoning as in the proof above leads us to the conclusion that as much as $O(n^2)$ arithmetical operations might be required when the Gaussian elimination is performed for B. Indeed, the elimination should be applied, in particular, to disjoint square blocks of order $1, 2, \ldots, w$, and these blocks must be considered as dense matrices even if they contain some zeros. So, the overall work cannot be less than

$$A(1) + A(2) + \ldots + A(w), \tag{1.25}$$

and it is well–known that the Gaussian elimination for a dense $k \times k$ matrix costs $A(k) \approx \frac{k^3}{3}$ flops (where a flop is understood as one floating point multiplication + one addition). Therefore, the sum (1.25) equals to $O(w^4) = O(n^2)$.

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

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

Algorithm 3.

1. Let $\ell_0 = 2$
2. For $k = 1, 2, \ldots$, if k is odd then go to Step 3; otherwise, go to Step 4.
3. If $k = 2m - 1$ then check the entries $\ell_{k-1} + 1$ through n in the column $a_{*,m}$ of the current matrix A. If all these entries are zero let $\ell_k = \ell_{k-1}, k \to k + 1$ and go to Step 2. Otherwise, find the Householder transformation $P_k = P(\ell_{k-1}, n, a_{*,m})$ and perform the similarity

   $$A \to P_k A P_k. \tag{1.26}$$

   Let $\ell_k = \ell_{k-1} + 1$. If $\ell_k = n$ go to Step 5; otherwise, $k \to k + 1$, go to Step 2.
4. If $k = 2m$ then check the entries $\ell_{k-1} + 1$ through n in the row $a_{m,*}$ of the current matrix A. If all these entries are zero let $\ell_k = \ell_{k-1}, k \to k+1$ and go to Step 2. Otherwise, find the Householder transformation $P_k = P(\ell_{k-1}, n, a_{m,*})$ and perform the similarity (1.26). Let $\ell_k = \ell_{k-1} + 1$. If $\ell_k = n$ go to Step 5; otherwise, $k \to k + 1$, go to Step 2.
5. Stop the procedure.

It appears that normality is used in no way in this algorithm. We therefore hasten to show how the normality does influence it. We do that at first for the particular classes (1.8) and (1.9), and then for some related classes, and also for general normal matrices.

Example 5. Let A at the entry of Algorithm 3 be a matrix of the form (1.8). After the first stage of the algorithm we have $\ell_1 = 3$ and the current column $a_{*,1}$ has zero entries 3 through n. Then (1.8) shows that the current row $a_{1,*}$ has also zero entries 3 through n. Therefore, we let $\ell_2 = 3$, 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 $\ell_3 = 4$. Again owing to (1.8) we have that in the current row $a_{*,2}$ entries 4 through n are already zeros. So, we let $\ell_4 = 4$, and pass to the fifth stage, and so on.

Example 6. Let A at the entry of Algorithm 3 be a matrix of the form (1.9). Then the first three stages run according to the general prescription with $\ell_1 = 3,\ \ell_2 = 4,\ \ell_3 = 5$. Now, if A is just orthogonal then already after the second stage the appearance of additional zeros could be registered. Indeed, in the column $a_{*,1}$ the entries 3 through n are zero; letting the reducible case aside, we can assume the entry 2 being nonzero. Since, after the second stage, the 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 in [Bunse–Gerstner, Elsner, p. 753]. It follows that, excepting the assignment $\ell_4 = 5$, we can omit the fourth stage. But the fifth stage can also be omitted, again with the exception of the assignment $\ell_5 = 5$! This follows from the orthogonality of $a_{1,*}$ and the rows 5 through n assuming again that the entry (1,3) is not zero. After that, the sixth stage and the seventh one go along the general rules, and so on. All what is said above is also true for a matrix A of the form (1.9) since such a matrix essentially differs only by a diagonal shift from a purely orthogonal one.

Example 7. Consider the plane quadric of the form

$$a_{11}\ x^2 + 2\ a_{12}\ xy + a_{22}\ y^2 = 1. \tag{1.27}$$

Here $a_{11}, a_{12}, a_{22} \in \mathbb{R}$, and the curve (1.27) is a hyperbola if

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} < 0, \tag{1.28}$$

and an ellipse if

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0. \tag{1.29}$$

Letting

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

we can obtain the complex form of equation (1.27):

$$cz^2 + \overline{cz}^2 + 2dz\overline{z} = 4. \tag{1.30}$$

Here

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

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

$$AA^* = A^*A, \;\; cA^2 + \overline{c}A^{*2} + 2dAA^* = 4I. \tag{1.31}$$

These are the normal matrices with eigenvalues belonging to the curve (1.30). These matrices could be called elliptical if conditions (1.29) hold, and hyperbolical, in case of (1.28) holds. 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 (1.27)).

Let us see how Algorithm 3 proceeds when applied to a general normal matrix A, and also to a matrix A of the class (1.31) with $c \neq 0$.
Now the first four stages run according to the general description, and $\ell_1 = 3, \; \ell_2 = 4, \; \ell_3 = 5, \; \ell_4 = 6$. It is easily seen that, after the fourth stage is completed, the column $a_{*,1}$ is orthogonal to the columns 6 to n. We can assume that the entry (1,3) is nonzero; otherwise, beginning with the third stage, Algorithm 3 would switch on the tridiagonal, Hermitian like course. Now, since the row $a_{1,*}$ should be also orthogonal to the rows 6 to n, the entries 6 to n of the column $a_{*,3}$ are zero and we can skip the fifth stage, letting only $\ell_5 = 6$. The sixth stage will then again go in accordance with the general rules, and the entries 7 to n of the third row will be zeroed.

All what is said in the previous paragraph except the last sentence is also valid for a matrix A of class (1.31). But now we can say more. Denote by $\mathcal{L}_k$ the linear span of the first k coordinate vectors in $\mathbb{C}^n$. Taking into consideration the form the matrix A assumes after the first stages, we have

$$A\, e_1 \in \mathcal{L}_2, \;\; A^*e_1 \in \mathcal{L}_3, \;\; A^2e_1 \in A\, \mathcal{L}_2 \subset \mathcal{L}_4, \;\; AA^*e_1 \in \; A\mathcal{L}_3 \subset \mathcal{L}_5. \tag{1.32}$$

Now we are able to prove that the entries 6 through n of the third row of A are zeroes. Since the entry (1,3) is nonzero we can write

$$e_3 = \alpha \; A^*e_1 + \beta \; e_1 + \gamma \; e_2 \tag{1.33}$$

with $\alpha \neq 0$. From (1.33) we derive

$$\begin{aligned} (A^*e_3, e_i) &= (e_3, Ae_i) = \alpha(A^*e_1, Ae_i) + \beta(e_1, Ae_i) + \gamma(e_2, Ae_i) \\ &= \alpha(A^*e_1, Ae_i) = \alpha(A^{*2}e_1, e_i), \;\; \text{if} \;\; i \geq 6, \end{aligned} \tag{1.34}$$

since the entries 6 through n of the first and second row are zeroes. Using (1.31), we have

$$(A^{*2}e_1, e_i) = \frac{1}{4}\overline{c}^{\,-1}(e_1, e_i) - c\overline{c}^{\,-1}(A^2e_1, e_i) - 2d\,\overline{c}^{\,-1}(AA^*e_1, e_i). \tag{1.35}$$

It follows from (1.32) that

$$\begin{aligned} &(A^2e_1, e_i) = 0, \;\; \text{if} \;\; i \geq 5, \\ &(AA^*e_1, e_i) = 0, \;\; \text{if} \;\; i \geq 6. \end{aligned}$$

So, with $i \geq 6$, all the terms in the right–hand side of (1.35) are zeroes, and

$$(A^* e_3, e_i) = 0, \ i \geq 6.$$

This fact could easily be predicted. Indeed, equation (1.31) implies the linear dependence (over L_1) the vectors in the second layer of the sequence (1.3) which means that $w_2 = 2$. Recall that this equality implies

$$w_m \leq 2 \ \text{ for all } \ m \geq 2.$$

Therefore, the condensed form B of a matrix A from the class (1.31) is a band matrix with the bandwidth ≤ 6 .

Example 8. Suppose that a given real normal matrix A has exactly one pair of complex conjugate eigenvalues, all the other eigenvalues being real. Then, for some real a , the pair of lines

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

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

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

which implies that the A itself satisfies the matrix relation

$$A^2 - A^{T2} = a(A - A^T).$$

This means that

$$w_2 \leq 2$$

and, according to Remark after Theorem 1,

$$w_m \leq 2, \ m > 2.$$

This example can be generalized allowing for k pairs of complex conjugate eigenvalues. In particular, if $k = 2$ then we can place the spectrum of A on the product of three lines

$$(x - a_1)(x - a_2)\ y = 0$$

for some $a_1, \ a_2 \in \mathbb{R}$. This gives $w_3 \leq 3$ and $w_m \leq 3, \ m > 3$.

In conclusion, we show that the **condensed** form B of a **normal** matrix A is invariant under the QR–algorithm given that no shift coincides with an exact eigenvalue of A . Denote by B_1 the matrix obtained from B through one step of some version of the QR–algorithm. Then

$$B_1 = RBR^{-1} = Q^* BQ \tag{1.36}$$

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

$$f(A) = QR.$$

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 middle expression for B_1 in (1.36) shows that the lower envelope of B_1 is the same as the one for B . Now, for any function φ , equality (1.36) implies

$$\varphi(B_1) = R\,\varphi(B)R^{-1} = Q^*\varphi(B)Q. \tag{1.37}$$

It is well–known that the normality of B amounts to B^* being a polynomial in B , see e.g. [Horn-Johnson, 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 (1.37) with $\varphi = \varphi_0$, we obtain

$$B_1^* = R\,B^*\,R^{-1}.$$

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

References

1. A. Bunse–Gerstner and L. Elsner, *Schur parameter pencils for the solution of the unitary eigenproblem*, Linear Algebra Appl., **154–156** (1991), 741–778.

2. R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

3. J.A. George and J.W. Liu, *Computer Solution of Large Sparse Positive Definite Linear Equations*, Prentice Hall, 1981.