ISSN 0278-6419, Moscow University Computational Mathematics and Cybernetics, 2009, Vol. 33, No. 3, pp. 109–116. © Allerton Press, Inc., 2009.
Original Russian Text © Ghasemi Kamalvand, Kh.D. Ikramov, 2009, published in Vestnik Moskovskogo Universiteta. Vychislitel'naya Matematika i Kibernetika, 2009, No. 3, pp. 5–11.

Low-Rank Perturbations of Normal and Conjugate-Normal Matrices and Their Condensed Forms under Unitary Similarities and Congruences

M. Ghasemi Kamalvand[a] **and Kh. D. Ikramov**[b]

[a] *University of Lorestan, Khorramabad, Islamic Republic of Iran*
e-mail: m_ghasemi98@yahoo.com
[b] *Faculty of Computational Mathematics and Cybernetics,*
Moscow State University, Moscow, 119992 Russia
e-mail: ikramov@cs.msu.su

Received September 1, 2008

Abstract—Two theorems are proved on the condensed forms with respect to unitary similarity and congruence transformations. They provide a theoretical basis for constructing economical iterative methods for systems of linear equations whose matrices are low-rank perturbations of normal and conjugate-normal matrices.

Keywords: normal matrices, conjugate-normal matrices, unitary similarity transformations, unitary congruence transformations, generalized Lanczos process.

DOI: 10.3103/S0278641909030017

1. It is well known that every Hermitian (or real symmetric) matrix can be reduced to tridiagonal form by a unitary (respectively, real orthogonal) similarity transformation. This theoretical fact underlies such popular methods for solving systems of linear equations with Hermitian matrices as the conjugate gradient method, MINRES, and SYMMLQ. Every complex symmetric matrix can be brought to tridiagonal form by a unitary congruence transformation, and this fact is also used in the practical solution of linear systems (for instance, the CSYM algorithm proposed in [1] is worthy of notice).

Let us make our point more exact. The above-mentioned unitary reduction to tridiagonal form is meant as a finite sequence of elementary unitary similarities or congruences or an equivalent of such a sequence (like the Lanczos algorithm or CSYM), which can be described as a finite process that preserves unitary similarity or congruence and employs only arithmetic operations and quadratic radicals. For brevity, any process of this type will be called a *finite orthogonal process*. The finiteness of the procedure is important; otherwise, every Hermitian or symmetric matrix could be brought even to a diagonal form.

A condensed form of an n-by-n matrix A is understood as a matrix that is unitarily similar or congruent to A and has a large number of zero entries. In addition to the above-mentioned tridiagonal form, an example of a condensed form can be the Hessenberg matrix. In this communication, we deal with a specific type of condensed forms, namely, block tridiagonal matrices with square diagonal blocks whose (possibly different) orders are bounded by a number $k \ll n$.

The classical fact on the possibility of unitary tridiagonalization of Hermitian matrices is a particular case of the following result proved in [1].

Theorem 1. *A normal n-by-n matrix A can be brought by a finite orthogonal process to the block tridiagonal form*

$$H = \begin{pmatrix} H_{11} & H_{12} & & \\ H_{21} & H_{22} & H_{23} & \\ & H_{32} & H_{33} & \dots \\ & \dots & \dots & \dots \end{pmatrix}, \tag{1}$$

where the diagonal blocks H_{11}, H_{22}, ... *are square and their orders are generically given by the consecutive integers* 1, 2, ... *If A satisfies the equation*

$$g(A, A^*) = 0, \tag{2}$$

where $g(x, y)$ *is a polynomial of degree* $m \ll n$, *then*, *starting from* $i = m$, *the orders of the diagonal blocks* H_{ii} in matrix (1) *stabilize at the value m.*

An equivalent formulation of condition (2) is the requirement that the entire spectrum of A belong to a plane algebraic curve of degree m (see [2]). For instance, the spectrum of a Hermitian matrix lies on the real axis, that is, curve of degree one; hence, $m = 1$.

An analog of Theorem 1 concerning the reduction to a block tridiagonal form by unitary congruence transformations was proved in [3]. This result holds for conjugate-normal matrices, which are defined by the equality

$$AA^* = \overline{A^*A}$$

and play the same role in the theory of unitary congruences as normal matrices do with respect to unitary similarities.

The reduction to a block tridiagonal form is also possible for certain types of matrices that are not normal or conjugate-normal. Let us write an n-by-n matrix A as

$$A = \hat{H} + \hat{K}, \tag{3}$$

where $\hat{H} = \hat{H}^*, \quad \hat{K} = -\hat{K}^*$.

Theorem 2. *Let A be a matrix of form* (3), *where the skew-Hermitian matrix* $\hat{K}$ *has the rank* $k \geq 1$. *Then*, *A can be brought by a finite orthogonal process to the block tridiagonal form* (1) *in which the orders of the diagonal blocks* H_{ii} *do not exceed* $k + 1$.

Theorem 2 is a restatement of the result proved in [4]. Note that the matrix A in this theorem is not normal if the matrices $\hat{H}$ and $\hat{K}$ do not commute.

The reduction described by Theorem 2 is effected by unitary similarity transformations. However, an analogous fact holds for unitary congruences if one now deals with (skew-symmetric) perturbations of symmetric matrices rather than (skew-Hermitian) perturbations of Hermitian matrices.

Theorem 3. *Let A be an n-by-n matrix written as*

$$A = S + K, \tag{4}$$

where $S = S^T$, $K = -K^T$. *If the skew-symmetric matrix K has rank* $k \geq 1$, *then A can be brought by a finite orthogonal process to the block tridiagonal form* (1) *in which the orders of the diagonal blocks* H_{ii} *do not exceed* $2k + 1$.

Our aim in this paper is to extend Theorems 2 and 3 so that low-rank perturbations could be examined not only for Hermitian and symmetric matrices but for any normal (respectively, conjugate-normal) matrix that can be brought to a block tridiagonal form. Moreover, we allow for general perturbations rather than only skew-Hermitian (respectively, skew-symmetric) ones.

The results obtained in this paper imply that the perturbed matrices can still be brought to block tridiagonal form, albeit the bound on the size of a diagonal block deteriorates proportionally to the rank of the perturbation matrix. These results are stated in Section 4. In Section 2, we recall the construction of the generalized Lanczos process, which underlies the proofs of Theorems 2 and 3. In Section 3, we discuss the relationship between the reductions to condensed forms based on unitary similarities and unitary congruences, respectively. This relationship, as well as the generalized Lanczos process, is used for proving the theorems in Section 4.

2. The Lanczos algorithm is a technique for reducing a Hermitian matrix A to tridiagonal form. The main idea of this algorithm is to orthonormalize the power sequence

$$x, Ax, A^2x, A^3x, \dots,$$

where x is a given or arbitrarily chosen initial vector. Suppose that a linear operator $\mathcal{A}$ acting in $\mathbf{C}^n$ is associated with the matrix A. Then, the matrix of this operator with respect to the orthogonal basis constructed in the Lanczos algorithm is the desired tridiagonal form.

If A is a normal nonhermitian matrix, then one should instead inspect the *generalized power sequence*

$$x,\ Ax,\ A^*x,\ A^2x,\ AA^*x,\ A^{*2}x,\ A^3x,\ \dots. \tag{5}$$

It is convenient to regard sequence (5) as consisting of segments of lengths 1, 2, 3, 4, ..., respectively. The kth segment called the kth layer can be described as the set of vectors $u = W_k(A, A^*)x$, where $W_k(s, t)$ varies over the set of kth degree monomials in the (commuting) variables s and t. The symbol $W_0(s, t)$ denotes the empty word; thus, $W_0(A, A^*)x$ is simply the vector x.

The essence of the *generalized Lanczos process* is the orthonormalization of sequence (5). With this process, we associate the following notation and terminology: the subspace

$$\mathcal{L}_m(A, x) = \operatorname{span}\{W(A, A^*)x : \deg W \le m\} \tag{6}$$

is called the mth *generalized Krylov subspace*. Its dimension is denoted by ℓ_m. The scalar $\omega_m = \ell_m - \ell_{m-1}$ ($m \ge 1$) is called the width of the mth layer. We set $\omega_0 = 1$.

Sequence (5) is certainly not constructed explicitly (just as the conventional power sequence is not constructed explicitly in the classical Lanczos algorithm). An implicit construction of this sequence and its orthogonalization are performed as follows. Suppose that we have already found an orthonormal basis $q_1, q_2, \dots, q_{\ell_m}$ in the subspace $\mathcal{L}_m(A, x)$; moreover, the last vectors $q_{\ell_{m-1}+1}, \dots, q_{\ell_m}$ in this basis have been obtained using the vectors in the mth layer of sequence (5). Now, in a certain order, we construct the vectors $Aq_{\ell_{m-1}+1}, \dots, Aq_{\ell_m}, A^*q_{\ell_{m-1}+1}, \dots, A^*q_{\ell_m}$, which are then orthogonalized to the current orthonormal system.

From this description, we can easily derive the following properties of the generalized Krylov subspaces (see [2, Section 2]):

(1) If $x \in \mathcal{L}_m$, then

$$Ax \in \mathcal{L}_{m+1}, \quad A^*x \in \mathcal{L}_{m+1}.$$

(2) If $q_\ell \in \mathcal{L}_m \backslash \mathcal{L}_{m-1}$, then

$$Aq_\ell \perp \mathcal{L}_{m-2}, \quad A^*q_\ell \perp \mathcal{L}_{m-2}.$$

As before, we associate with A the linear operator $\mathcal{A}$ acting in the n-dimensional space. Suppose that the application of the generalized Lanczos process to A and the initial vector x produces the orthonormal basis $q_1, \dots, q_n$. Then, properties (1) and (2) imply that the matrix of $\mathcal{A}$ with respect to this basis has block tridiagonal form (1). Furthermore, the orders n_i of the diagonal blocks H_{ii} are determined by the scalars ω_i; namely,

$$n_i = \omega_{i-1}, \quad i = 1, 2, \dots.$$

In particular, n_1 is always one.

The idea of the generalized Lanczos process is also applicable to a nonnormal matrix A. An important distinction from the normal case is that A and A^* do not commute any longer. Therefore, the kth layer of the generalized power sequence should now be defined as the set of vectors $u = W_k(A, A^*)x$, where $W_k(s, t)$ is an arbitrary kth degree monomial in the *noncommuting* variables s and t. As a result, the bound

$$\omega_i \le i + 1, \quad i = 0, 1, 2, \dots,$$

which is valid for all the normal matrices, is replaced by the inequality

$$\omega_i \le 2^i, \quad i = 0, 1, 2, \dots.$$

If this inequality realistically describes the situation with a specific matrix A, then the use of the generalized Lanczos process for such a matrix is hardly reasonable. However, for some classes of nonnormal matrices, the scalars ω_i can be bounded by a small constant for all i. Theorem 2 specifies one of such classes.

3. The block tridiagonal matrix (1) constructed in the preceding section using the generalized Lanczos process is unitarily similar to the original matrix A. In this section, we discuss the reduction to condensed forms based on unitary congruences.

We begin with the case of a conjugate-normal matrix A. With such a matrix, we associate the matrix

$$\hat{A} = \begin{pmatrix} 0 & \bar{A} \\ A & 0 \end{pmatrix} \tag{7}$$

of the double order. (The bar over the symbol of a matrix or a vector means entry-wise conjugation.) It is easy to verify that matrix (7) is normal in the conventional sense.

We fix a nonzero vector $x \in \mathbf{C}^n$ with which we associate the vector

$$v = \begin{pmatrix} x \\ \bar{x} \end{pmatrix} \tag{8}$$

of the double dimension. Consider the generalized power sequence generated by $\hat{A}$ and vector (8):

$$\begin{gathered} v, \hat{A}v = \begin{pmatrix} \bar{A}\bar{x} \\ Ax \end{pmatrix}, \quad \hat{A}^*v = \begin{pmatrix} A^*\bar{x} \\ A^Tx \end{pmatrix}, \quad \hat{A}^2v = \begin{pmatrix} A_Lx \\ A_R\bar{x} \end{pmatrix}, \\ \hat{A}\hat{A}^*v = \begin{pmatrix} \bar{A}A^Tx \\ AA^*\bar{x} \end{pmatrix}, \quad \hat{A}^{*2}v = \begin{pmatrix} A^*A^Tx \\ A^TA^*\bar{x} \end{pmatrix}, \quad \ldots. \end{gathered} \tag{9}$$

The symbols A_L and A_R stand for the matrices $\bar{A}A$ and $A\bar{A}$, respectively.

The upper halves of vectors (9) form the sequence

$$x, \bar{A}\bar{x}, A^*\bar{x}, \bar{A}Ax, \bar{A}A^Tx, A^*A^Tx, \bar{A}A\bar{A}\bar{x}, \ldots. \tag{10}$$

We split this sequence into layers defining the kth layer as the set of vectors corresponding to the kth layer of sequence (9). Similarly, with the mth subspace $\mathcal{L}_m(\hat{A}, v)$, we associate the subspace $\tilde{\mathcal{L}}_m(A, x)$ formed of the upper halves of the vectors $z \in \mathcal{L}_m(\hat{A}, v)$. The dimension of $\tilde{\mathcal{L}}_m(A, x)$ is denoted by $\tilde{\ell}_m$, and the scalar $\tilde{\omega}_m = \tilde{\ell}_m - \tilde{\ell}_{m-1}$ $(m \ge 1)$ is called the width of the mth layer in (10). It is obvious that

$$\tilde{\omega}_m \le \omega_m, \quad m = 1, 2, \ldots, \tag{11}$$

where ω_m is the width of the mth layer in (9).

Now, we perform an unusual orthogonalization of sequence (10). By analogy with the generalized Lanczos algorithm, we can describe this process as follows. Suppose that we have already found an orthonormal basis $q_1, q_2, \ldots, q_{\tilde{\ell}_m}$ in the subspace $\tilde{\mathcal{L}}_m(A, x)$; moreover, the last vectors $q_{\tilde{\ell}_{m-1}+1}, \ldots, q_{\tilde{\ell}_m}$ in this basis have been obtained using the vectors in the mth layer of sequence (10). Then, in a certain order, we construct the vectors $Aq_{\tilde{\ell}_{m-1}+1}, \ldots, Aq_{\tilde{\ell}_m}, A^Tq_{\tilde{\ell}_{m-1}+1}, \ldots, A^Tq_{\tilde{\ell}_m}$. Each of these vectors is orthogonalized to the current orthonormal system formed of the conjugate vectors $\bar{q}_1, \bar{q}_2, \ldots$. If the complete orthogonalization results in a nonzero vector, this vector is normalized and, after applying conjugation once more, becomes the new vector q_j.

It can be shown that the subspaces $\tilde{\mathcal{L}}_m$ have similar properties to those of the generalized Krylov subspaces (see properties (1) and (2) in Section 2):

(1) If $x \in \tilde{\mathcal{L}}_m$, then

$$Ax \in \overline{\tilde{\mathcal{L}}}_{m+1}, \quad A^Tx \in \overline{\tilde{\mathcal{L}}}_{m+1}.$$

(2) If $q_\ell \in \tilde{\mathcal{L}}_m \backslash \tilde{\mathcal{L}}_{m-1}$, then

$$Aq_\ell \perp \overline{\tilde{\mathcal{L}}}_{m-2}, \quad A^T q_\ell \perp \overline{\tilde{\mathcal{L}}}_{m-2}.$$

The symbol $\overline{\mathcal{M}}$ as applied to a subspace $\mathcal{M}$ denotes the subspace

$$\overline{\mathcal{M}} = \{x \in \mathbf{C} | \bar{x} \in \mathcal{M}\}.$$

Suppose that the application of the above process to the conjugate-normal matrix A and the initial vector x produces an orthonormal basis $q_1, \ldots, q_n$ in $\mathbf{C}^n$. Define the unitary matrix

$$Q = (q_1 q_2 \ldots q_n).$$

Then, properties (1) and (2) stated in this section imply the matrix equality

$$AQ = \bar{Q}H, \tag{12}$$

where H is a matrix of form (1) in which the orders of the diagonal blocks are determined by the scalars $\tilde{\omega}_i$. Rewriting (12) as

$$Q^T A Q = H,$$

we conclude that A and H are unitarily congruent.

If we drop the requirement that A be conjugate-normal, the associate matrix (7) will not longer be normal. However, all the above constructions and definitions are still applicable with the distinction that the width of the ith layer in sequence (9) generically increases to 2^i. Inequality (11) relating the width of a layer in (9) and the width of the corresponding layer in (10) remains also valid.

4. We first examine low-rank perturbations of normal matrices.

Theorem 4. *Let N be a normal n-by-n matrix such that the generalized Lanczos process with an arbitrary initial vector brings N to a block tridiagonal form in which the orders of the diagonal blocks do not exceed the scalar ω_0. Then, for every matrix R of rank $k \ll n$, the matrix*

$$A = N + R \tag{13}$$

can be brought to a block tridiagonal form in which the orders of the diagonal blocks do not exceed $\hat{\omega} = (2k+1)\omega_0$.

Proof. We write the perturbation matrix R as a sum of k dyads; that is,

$$R = x_1 y_1^* + \ldots + x_k y_k^*, \quad x_1, y_1, \ldots, x_k, \quad y_k \in \mathbf{C}^n.$$

Taking an arbitrary nonzero vector $v \in \mathbf{C}^n$, we construct the sequence of subspaces $\mathcal{M}_k$ according to the following rules:

$$\begin{gathered} \mathcal{M}_0 = \operatorname{span}\{v\}, \quad \mathcal{M}_1 = \operatorname{span}\{vNv, N^*v, x_1, \ldots, x_k, y_1, \ldots, y_k\}, \\ \mathcal{M}_{k+1} = \operatorname{span}\{\mathcal{M}_k, N\mathcal{M}_k, N^*\mathcal{M}_k\}, \quad k = 1, 2, \ldots. \end{gathered} \tag{14}$$

The right-hand side of (14) should be understood as the span of all the vectors in the indicated subspaces.

We show that the inclusions

$$A\mathcal{M}_k \subset \mathcal{M}_{k+1} \tag{15}$$

and

$$A^*\mathcal{M}_k \subset \mathcal{M}_{k+1} \tag{16}$$

hold for all k. If $k = 0$, then we have

$$\begin{gathered} Av = Nv + \alpha_1 x_1 + \ldots + \alpha_k x_k \in \operatorname{span}\{Nv, x_1, \ldots, x_k\}, \\ A^*v = N^*v + \beta_1 y_1 + \ldots + \beta_k y_k \in \operatorname{span}\{N^*v, y_1, \ldots, y_k\}, \end{gathered}$$

where

$$\alpha_j = y_j^* v, \quad \beta_j = x_j^* v, \quad j = 1, 2, \ldots, k.$$

Thus, inclusions (15) and (16) hold for $k = 0$.

If $k > 0$ and $z \in \mathcal{M}_k$, we have

$$Az = Nz + \gamma_1 x_1 + \ldots + \gamma_k x_k \in \text{span}\{N\mathcal{M}_k, x_1, \ldots, x_k\},$$
$$A^*z = N^*z + \delta_1 y_1 + \ldots + \delta_k y_k \in \text{span}\{N^*\mathcal{M}_k, y_1, \ldots, y_k\},$$

where

$$\gamma_j = y_j^* z, \quad \delta_j = x_j^* z, \quad j = 1, 2, \ldots, k.$$

Since $x_1, y_1, \ldots, x_k, y_k \in \mathcal{M}_1 \subset \mathcal{M}_k$ $(k > 1)$, it holds that

$$A\mathcal{M}_k \subset \text{span}\{\mathcal{M}_k, N\mathcal{M}_k\} \subset \mathcal{M}_{k+1}$$

and

$$A^*\mathcal{M}_k \subset \text{span}\{\mathcal{M}_k, N^*\mathcal{M}_k\} \subset \mathcal{M}_{k+1}.$$

Now, we apply an orthonormalization process similar to the generalized Lanczos algorithm to the chain of subspaces

$$\mathcal{M}_0 \subset \mathcal{M}_1 \subset \mathcal{M}_2 \subset \ldots. \tag{17}$$

Suppose that this process results in an orthonormal basis $q_1, \ldots, q_n$ in $\mathbf{C}^n$. Split this basis into the parts (layers) that correspond to the quotient spaces

$$\mathcal{K}_j = \mathcal{M}_j / \mathcal{M}_{j-1}, \quad j = 1, 2, \ldots. \tag{18}$$

The dimensions of subspaces (18) play the same role as the scalars ω_i do in the generalized Lanczos process; namely, they determine the sizes of block rows and columns in the matrix H corresponding to the underlying linear operator $\mathcal{A}$ in the basis $q_1, \ldots, q_n$. Now, inclusion (15) means that at most one subdiagonal block is nonzero in each block column of H. This is the block adjacent to the diagonal block. Similarly, inclusion (16) implies that, in each block row of H, only one superdiagonal block can be nonzero, namely, the block adjacent to the diagonal block. In other words, H is a block tridiagonal matrix, and it remains to estimate the sizes of its diagonal blocks.

According to formula (14), the subspace $\mathcal{K}_j$ is determined by the jth layer of the generalized power sequence generated by N and v and by the $(j-1)$th layers of the sequences generated by N and the vectors $x_1, \ldots, x_k, y_1, \ldots, y_k$. By the hypotheses of the theorem, the widths of all the layers in each of these sequences do not exceed ω_0. Consequently, the dimension of each subspace $\mathcal{K}_j$ is bounded by the scalar $\hat{\omega} = (2k+1)\omega_0$.

Assume that, for some k, sequence (17) stabilizes; that is,

$$\mathcal{M}_k = \mathcal{M}_{k+1} = \mathcal{M}_{k+2} = \ldots.$$

If an orthonormal basis in $\mathbf{C}^n$ is not yet obtained, then $\mathcal{M}_k$ is a nontrivial common invariant subspace of A and A^*. In this case, one should act exactly as in the conventional Lanczos algorithm; namely, choose a nonzero vector that is orthogonal to $\mathcal{M}_k$ and perform the same actions as those performed above with the vector v. As a result, the current orthonormal system will be extended by new vectors. It is possible that, to obtain a complete basis in $\mathbf{C}^n$, this procedure has to be repeated. Suppose that the desired basis is finally constructed. Then, the operator $\mathcal{A}$ has a block diagonal matrix H with respect to this basis. The argument given above applies to each diagonal block, which proves the theorem in this case as well.

Now, we examine low-rank perturbations of conjugate-normal matrices.

Theorem 5. *Let N be a conjugate-normal n-by-n matrix such that the process described in Section* 3 *and started with an arbitrary initial vector brings N to a block tridiagonal form in which the orders of the diagonal blocks do not exceed the scalar ω_0. Then, for every matrix R of rank $k \ll n$, the matrix*

$$A = N + R \tag{19}$$

can be brought to a block tridiagonal form in which the orders of the diagonal blocks do not exceed $\breve{\omega} = 2(4k+1)\omega_0$.

Proof. Instead of A, we consider the corresponding matrix $\hat{A}$. Substituting expression (19) into (7), we obtain

$$\hat{A} = \begin{pmatrix} 0 & \bar{N} \\ N & 0 \end{pmatrix} + \begin{pmatrix} 0 & \bar{R} \\ R & 0 \end{pmatrix} = \hat{N} + \hat{R}. \tag{20}$$

The matrix $\hat{N}$ is normal, while $\hat{R}$ is obviously of rank $2k$.

Suppose that the process described in Section 3 and started with the initial vector q_1 generates an orthonormal basis in $\mathbf{C}^n$ in which N has a block tridiagonal form. By assumption, the orders of the diagonal blocks in this form do not exceed the scalar ω_0.

Let us apply the generalized Lanczos process to the matrix $\hat{N}$ and the initial vector

$$v = \begin{pmatrix} q_1 \\ \bar{q}_1 \end{pmatrix}. \tag{21}$$

Then, n orthonormal vectors in $\mathbf{C}^{2n}$ will be obtained. Being split into layers, these vectors satisfy the relations

$$\tilde{\omega}_i \le \omega_i \le 2\tilde{\omega}_i \le 2\omega_0, \quad i = 1, 2, \dots. \tag{22}$$

Here, $\tilde{\omega}_i$ are the orders of the diagonal blocks in the block tridiagonal form of N, while ω_i are the widths of the layers in the generalized power sequence generated by the pair $(\hat{N}, v)$.

Starting with the one-dimensional subspace $\mathcal{M}_0 = \mathrm{span}\{v\}$, we construct the same sequence of the subspaces $\mathcal{M}_k$ for $\hat{A}$ as in the proof of Theorem 4. Taking into account the relation $\mathrm{rank}\,\hat{R} = 2k$ and using inequalities (22), we conclude that the dimensions of quotient spaces (18) are bounded by the scalar

$$\breve{\omega} = 2(4k+1)\omega_0. \tag{23}$$

Moreover, the dimension of the maximal subspace among $\mathcal{M}_k$ is at least n because these subspaces contain the generalized Krylov subspaces for the pair $(\hat{N}, v)$.

Now, we project the subspaces

$$\mathcal{M}_0 \subset \mathcal{M}_1 \subset \mathcal{M}_2 \subset \dots$$

on $\mathbf{C}^n$. In other words, we take the upper half q in each base vector

$$v = \begin{pmatrix} q \\ \bar{q} \end{pmatrix}$$

of these subspaces. The resulting set $\breve{q}_1, \dots, \breve{q}_n$ of vectors in $\mathbf{C}^n$ can be split into layers in a natural way. Since the width of any layer cannot increase in the process of projecting, scalar (23) is still an upper bound for the width of the layer in the system $\breve{q}_1, \dots, \breve{q}_n$.

Suppose that the vectors $\breve{q}_1, \dots, \breve{q}_n$ are linearly independent. Then, without loss of generality, $\breve{q}_1, \dots, \breve{q}_n$ can be regarded as orthonormal vectors. The structure of the subspaces $\mathcal{M}_k$ implies that A has a block tridiagonal form with respect to the basis $\breve{q}_1, \dots, \breve{q}_n$, and the orders of the diagonal blocks in this form do not exceed the scalar $\breve{\omega}$. In this case, the theorem is proved.

It remains to consider the case when projecting on $\mathbf{C}^n$ reduces the dimension; that is, an orthonormal system

$$\breve{q}_1, \dots, \breve{q}_m, \quad m < n, \tag{24}$$

rather than a basis is obtained. The span of this system is a common coninvariant subspace of A and A^T.

We choose a (unit) vector $\tilde{q}_1$ that is orthogonal to each vector in system (24) and perform the same operations as those performed above for the vector q_1. Typically, this process converts system (24) into an orthonormal basis

$$\breve{q}_1, \ldots, \breve{q}_m, \quad \tilde{q}_1, \ldots, \tilde{q}_l, \quad l + m = n$$

in $\mathbf{C}^n$. In this basis, A has a block tridiagonal form, and the above bound $\breve{\omega}$ is valid for the orders of the diagonal blocks. A distinction from the preceding case is that we now have the direct sum of two block tridiagonal submatrices.

It may happen that the system $\breve{q}_1, \ldots, \breve{q}_m, \tilde{q}_1, \ldots, \tilde{q}_l$ is not yet a basis. Then, we choose a vector $\tilde{\tilde{q}}_1$ that is orthogonal to this system and repeat the above construction. In this way, we ultimately obtain an orthonormal basis in $\mathbf{C}^n$ in which A has the desired form.

Remark. The results of Theorems 4 and 5 are of practical interest for solving systems of linear equations only when the scalars ω_0 and k are fairly small.

REFERENCES

1. A. Bunse-Gerstner and R. Stöver, "On a Conjugate Gradient-Type Method for Solving Complex Symmetric Linear Systems," Linear Algebra Appl. **287,** 105–123 (1999).
2. L. Elsner and Kh. D. Ikramov, "On a Condensed Form for Normal Matrices under Finite Sequences of Elementary Unitary Similarities," Linear Algebra Appl. **254,** 79–98 (1997).
3. Kh. D. Ikramov, "On Reducing Complex Matrices to Condensed Forms," Mat. Zametki **82,** 550–559 (2007).
4. M. Dana and Kh. D. Ikramov, "More on the Solution of Systems of Linear Equations with the Matrices That Are Low-Rank Perturbations of Hermitian Matrices," Zap. Nauchn. Seminarov POMI **334,** 68–77 (2005).