ISSN 0965-5425, Computational Mathematics and Mathematical Physics, 2009, Vol. 49, No. 2, pp. 203–216. © Pleiades Publishing, Ltd., 2009.
Original Russian Text © M. Ghasemi Kamalvand, Kh.D. Ikramov, 2009, published in Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, 2009, Vol. 49, No. 2, pp. 211–224.

A Method of Congruent Type for Linear Systems with Conjugate-Normal Coefficient Matrices

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 February 1, 2008;
in final form, June 9, 2008

Abstract—Minimal residual methods, such as MINRES and GMRES, are well-known iterative versions of direct procedures for reducing a matrix to special condensed forms. The method of reduction used in these procedures is a sequence of unitary similarity transformations, while the condensed form is a tridiagonal matrix (MINRES) or a Hessenberg matrix (GMRES). The algorithm CSYM proposed in the 1990s for solving systems with complex symmetric matrices was based on the tridiagonal reduction performed via unitary congruences rather than similarities. In this paper, we construct an extension of this algorithm to the entire class of conjugate-normal matrices. (Complex symmetric matrices are a part of this class.) Numerical results are presented. They show that, on many occasions, the proposed algorithm has a superior convergence rate compared to GMRES.

DOI: 10.1134/S0965542509020018

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

1. INTRODUCTION

Minimal residual methods, such as MINRES and GMRES, are well-known iterative versions of direct procedures for reducing a matrix to special condensed forms. The method of reduction used in these procedures is a sequence of unitary similarity transformations, while the condensed form is a tridiagonal matrix (MINRES) or a Hessenberg matrix (GMRES).

A symmetric matrix A having nonreal entries cannot be reduced to tridiagonal form by a unitary similarity transformation. On the other hand, A can be brought to tridiagonal form by a sequence of unitary congruences. This fact was used in [1] to construct the method CSYM for solving systems of linear equations with complex symmetric coefficient matrices.

The authors of [1] claim that CSYM is fundamentally different from Krylov subspace methods. Indeed, the trial subspaces in which CSYM seeks approximate solutions are not Krylov subspaces. However, we show in Section 2 that CSYM can be interpreted as the projection onto $\mathbb{C}^n$ of the conventional Lanczos process performed in a space of double dimension for the Hermitian matrix

$$\hat{A} = \begin{pmatrix} 0 & \overline{A} \\ A & 0 \end{pmatrix} \tag{1}$$

and a special initial vector. (The bar over the symbol of a matrix or a vector means entrywise conjugation.) In Section 4, this interpretation allows us to extend CSYM to systems with conjugate-normal coefficient matrices. Recall that a complex $n \times n$ matrix A is said to be conjugate-normal if

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

In particular, symmetric, skew-symmetric, and unitary matrices are conjugate-normal.

For a conjugate-normal matrix A, matrix (1) is normal in the conventional sense. Consequently, our extension of CSYM is based on the generalized Lanczos process, which was proposed in [2] exactly for normal matrices. A brief description of this process is given in Section 3.

The condensed form of the normal matrix $\hat{A}$ calculated by the generalized Lanczos process is a block tridiagonal matrix whose diagonal blocks have slowly increasing orders. The condensed form of the conjugate-normal matrix A is obtained by projecting the Lanczos process onto $\mathbb{C}^n$ and is also block tridiagonal.

Suppose that $\hat{A}$ satisfies the equation

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

where $f(x, y)$ is a polynomial of degree $k \ll n$. Then, after attaining the value k, the orders of the diagonal blocks stabilize at this value, while the condensed forms of $\hat{A}$ and A become band matrices with the bandwidth depending on k. For the iterative process, this means that the depth of the recursion specifying the next vector does not exceed d.

In Section 5, a detailed discussion of the MINRES-CN2 method is given. This is a variant of our extension of CSYM designed for the case where the polynomial f in formula (2) has the degree $k = 2$. The numerical results obtained with this method and its comparison with GMRES are given in Section 6.

2. METHOD CSYM

Suppose that we have to solve the system of linear equations

$$Ax = b \tag{3}$$

with a complex symmetric matrix A. For a system of a modest order, this can be done by finding a unitary matrix that brings A to tridiagonal form by the congruence transformation

$$Q^{\mathrm{T}}AQ = T. \tag{4}$$

As soon as Q is found, solving system (3) reduces to solving the tridiagonal system of linear equations

$$Ty = c,$$

where

$$y = Q^*x, \quad c = Q^{\mathrm{T}}b.$$

The matrix Q can be constructed as a finite product of Householder reflections or rotations.

The above method is inapplicable to large matrices A. However, the idea of reduction to tridiagonal form is completely reasonable in this case as well. Denote by $q_1, \ldots, q_n$ the columns of Q and assume that T in relation (4) is an irreducible matrix (that is, all the entries $t_{i+1,i}$ on its secondary diagonal are nonzero). Rewrite (4) as

$$AQ = \bar{Q}T$$

and equate the columns with the same indices in this matrix relation. This yields

$$\begin{aligned} Aq_1 &= t_{11}\bar{q}_1 + t_{21}\bar{q}_2, \\ Aq_2 &= t_{12}\bar{q}_1 + t_{22}\bar{q}_2 + t_{32}\bar{q}_3, \\ &\cdots \\ Aq_n &= t_{n-1,n}\bar{q}_{n-1} + t_{nn}\bar{q}_n. \end{aligned} \tag{5}$$

Furthermore,

$$t_{ii} = (Aq_i, \bar{q}_i), \quad i = 1, 2, \ldots, n, \tag{6}$$

$$|t_{21}| = \|Aq_1 - t_{11}\bar{q}_1\|_2, \tag{7}$$

$$|t_{i+1,i}| = \|Aq_i - t_{ii}\bar{q}_i - t_{i-1,i}\bar{q}_{i-1}\|_2, \quad i = 2, 3, \ldots, n-1, \tag{8}$$

$$t_{i,i+1} = t_{i+1,i}, \quad i = 1, 2, \ldots, n-1. \tag{9}$$

Equalities (5)–(9) are very similar to the formulas of the Lanczos method. The main difference is that, instead of the previously calculated vectors q_j, the current product Aq_i is orthogonalized to the conjugate vectors $\bar{q}_j$.

Using this analogy with the Lanczos method, we can seek a solution to system (3) by constructing an orthonormal system $q_1, q_2, \ldots$ by formulas (5)–(9). An appropriate choice of the initial vector is

$$q_1 = \frac{1}{\|b\|_2}\bar{b}. \tag{10}$$

Approximate solutions are sought in the trial subspaces

$$\mathcal{L}_k = \operatorname{span}\{q_1, \ldots, q_k\}. \tag{11}$$

In such a process, we expect that an approximate solution of reasonable quality can already be found for $k \ll n$.

The algorithm CSYM, which was proposed in [1], combines recursion (5)–(9) with the principle of orthogonal residual as a device for seeking an approximate solution in the subspace $\mathcal{L}_k$. Thus, CSYM is an analog of the minimal residual method MINRES, which was designed for systems with Hermitian (or real symmetric) matrices. Due to the above-mentioned difference in the orthogonalization methods, trial subspaces (11) used in CSYM are not Krylov subspaces. For instance, for $k = 2m$, we have

$$\mathcal{L}_k = \operatorname{span}\{q_1, \bar{A}\bar{q}_1, \bar{A}Aq_1, (\bar{A}A)\bar{A}\bar{q}_1, \ldots, (\bar{A}A)^{m-1}q_1, (\bar{A}A)^{m-1}\bar{A}\bar{q}_1\}. \tag{12}$$

It follows that $\mathcal{L}_k$ is spanned by the two sequences

$$q_1, \bar{A}Aq_1, \ldots, (\bar{A}A)^{m-1}q_1$$

and

$$\bar{A}\bar{q}_1, (\bar{A}A)\bar{A}\bar{q}_1, \ldots, (\bar{A}A)^{m-1}\bar{A}\bar{q}_1.$$

rather than by a single power sequence as in the Lanczos method. These two sequences are generated by the matrix

$$A_L = \bar{A}A$$

(rather than by A!) and correspond to the initial vectors q_1 and $\bar{A}\bar{q}_1$.

Despite this fact and the opinion stated in [1], CSYM is not fundamentally different from Krylov subspace methods. Indeed, let Hermitian matrix (1) be associated with a symmetric matrix A and the vector

$$v = \begin{pmatrix} q_1 \\ \bar{q}_1 \end{pmatrix} \in \mathbb{C}^{2n} \tag{13}$$

be associated with the initial vector q_1. The power sequence generated by the matrix $\hat{A}$ and by the vector v has the form

$$v, \quad \hat{A}v = \begin{pmatrix} \bar{A}\bar{q}_1 \\ Aq_1 \end{pmatrix}, \quad \hat{A}^2v = \begin{pmatrix} A_Lq_1 \\ \bar{A}_L\bar{q}_1 \end{pmatrix}, \quad \hat{A}^3v = \begin{pmatrix} A_L\bar{A}\bar{q}_1 \\ \bar{A}_LAq_1 \end{pmatrix}, \quad \ldots. \tag{14}$$

It determines the progress of the Lanczos method for $\hat{A}$. At the same time, the upper halves of vectors (14) constitute the sequence

$$q_1, \bar{A}\bar{q}_1, A_Lq_1, A_L\bar{A}\bar{q}_1, \ldots,$$

whose initial segments span the trial subspaces of CSYM. In this sense, CSYM can be regarded as the projection onto $\mathbb{C}^n$ of the Lanczos process performed in a space of double dimension.

The established relation between CSYM and the Lanczos method is used in Section 4 for extending CSYM to the entire class of conjugate-normal matrices. This extension is based on the generalized Lanczos process described in the next section.

3. GENERALIZED LANCZOS PROCESS

Let A be a normal $n \times n$ matrix. For an arbitrary (nonzero) vector $b \in \mathbb{C}^n$, the sequence

$$b, Ab, A^*b, A^2b, AA^*b, A^{*2}b, A^3b, \ldots \tag{15}$$

is called the generalized power sequence generated by the pair (A, b).

It is convenient to think of sequence (15) as consisting of segments of length 1, 2, 3, 4, …, respectively. The kth segment, which is called the kth layer, can be described as the set of vectors $u = W_k(A, A^*)b$, where $W_k(s, t)$ varies over the set of kth degree monomials in (commuting) variables s and t. The symbol $W_0(s, t)$ denotes the empty word; thus, $W_0(A, A^*)b$ is simply the vector b.

The essence of the generalized Lanczos process is the orthogonalization of sequence (15). However, as in the conventional Lanczos method, this power sequence is not constructed explicitly. The details of the orthogonalization process are specified in Section 5, where the method MINRES-CN2 is discussed. Here, we wish to point to the relation between this process and the unitary reduction of a normal matrix to a special condensed form.

For simplicity, we assume that the span of vectors (15) is the entire space $\mathbb{C}^n$. In this case, the vectors q_1, …, q_n produced by the generalized Lanczos process form an orthonormal basis in $\mathbb{C}^n$. We regard A as a linear operator acting in $\mathbb{C}^n$. It can be shown (see [2]) that the matrix of this operator in the basis $q_1, \ldots, q_n$ has the block tridiagonal form

$$\begin{pmatrix} H_{11} & H_{12} & & \\ H_{21} & H_{22} & H_{23} & \\ & H_{32} & H_{33} & \ldots \\ & & \ldots & \end{pmatrix}. \tag{16}$$

The orders of the diagonal blocks $H_{11}, H_{22}, \ldots$ are uniquely determined by sequence (15). To explain this, we call the span of the first $m + 1$ layers the mth generalized Krylov subspace; thus,

$$\mathcal{L}_m(A, b) = \operatorname{span}\{W(A, A^*)b: \deg W \le m\}. \tag{17}$$

The dimension of $\mathcal{L}_m$ is denoted by l_m. The scalar $\omega_m = l_m - l_{m-1}$, $m \ge 1$ is called the width of the mth layer. This scalar is the increment in dimension obtained by adding the vectors in the mth layer to $\mathcal{L}_{m-1}$ and by taking the span of the extended system. By definition, the width of the zeroth layer is one.

Turning back to matrix (16), we can say that the order of the diagonal block H_{ii} is ω_{i-1} ($i = 1, 2, \ldots$). For a general normal matrix A, the scalars ω_i increase slowly in accordance with the formula

$$\omega_i = i + 1.$$

It follows (again, see [2]) that the number of nonzero entries in H does not exceed $3\sqrt{2}n^{3/2}$. Note that, if the Arnoldi method (which is a non-Hermitian analog of the Lanczos method) were applied to A, then the normality could not be used and the process would result in a Hessenberg matrix in which the number of nonzero entries is $\approx \frac{1}{2}n^2$.

Suppose that a normal matrix A satisfies the additional relation

$$g(A, A^*) = 0, \tag{18}$$

where $g(x, y)$ is a polynomial of degree $k \ll n$. Then, beginning at $i = k$, the orders of the diagonal blocks H_{ii} in matrix (16) stabilize at k. In this case, H can be regarded as a band matrix whose bandwidth does not exceed $3k$. Condition (18) is equivalent to the requirement that the spectrum of A belong to a plane algebraic curve of degree k.

4. METHOD MINRES-CN FOR CONJUGATE-NORMAL MATRICES

Following Section 2, we associate the matrix $\hat{A}$ with an $n \times n$ matrix A. It is easy to verify that A is conjugate-normal if and only if $\hat{A}$ is normal in the conventional sense.

Fix a unit vector $q_1 \in \mathbb{C}^n$ and define a vector v by formula (13). Assuming that A is conjugate-normal, we construct the generalized power sequence generated by $\hat{A}$ and v. This yields the sequence

$$\begin{gathered} v, \quad \hat{A}v = \begin{pmatrix} \bar{A}\bar{q}_1 \\ Aq_1 \end{pmatrix}, \quad \hat{A}^*v = \begin{pmatrix} A^*\bar{q}_1 \\ A^{\mathrm{T}}q_1 \end{pmatrix}, \quad \hat{A}^2 v = \begin{pmatrix} A_L q_1 \\ \bar{A}_L \bar{q}_1 \end{pmatrix}, \\ \hat{A}\hat{A}^* v = \begin{pmatrix} \bar{A}A^{\mathrm{T}}q_1 \\ AA^*\bar{q}_1 \end{pmatrix}, \quad \hat{A}^{*2} v = \begin{pmatrix} A^*A^{\mathrm{T}}q_1 \\ A^{\mathrm{T}}A^*\bar{q}_1 \end{pmatrix}, \quad \ldots. \end{gathered} \tag{19}$$

The upper halves of these vectors form the sequence

$$q_1, \bar{A}\bar{q}_1, A^*\bar{q}_1, \bar{A}Aq_1, \bar{A}A^{\mathrm{T}}q_1, A^*A^{\mathrm{T}}q_1, \bar{A}A\bar{A}\bar{q}_1, \ldots. \tag{20}$$

Define the layers in this sequence as the projections onto $\mathbb{C}^n$ of the corresponding layers in sequence (19). Thus, the vectors $\bar{A}\bar{q}_1$ and $A^*\bar{q}_1$ form the first layer; $\bar{A}Aq_1$, $\bar{A}A^{\mathrm{T}}q_1$, and $A^*A^{\mathrm{T}}q_1$ form the second layer; etc. Following Section 3, we define the width $\tilde{\omega}_m$ of the mth layer as the increment in the dimension of the span of the vectors in the first i layers when we pass from $i = m - 1$ to $i = m$. It is obvious that

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

For $m = 0$, we set $\tilde{\omega}_0 = 1$.

Suppose that the span of sequence (20) is the entire space $\mathbb{C}^n$. Then, the orthogonalization of this sequence (understood in the same sense as in CSYM) produces an orthonormal basis $q_1, \ldots, q_n$. Define

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

From the coefficients h_{ij} calculated in the course of the orthogonalization, we form the matrix H. It has the block tridiagonal form (16) and the orders of its diagonal blocks are given by the scalars $\tilde{\omega}_i$. Moreover, it holds that

$$AQ = \bar{Q}H; \tag{21}$$

that is, A and H are unitarily congruent.

The block tridiagonal form of H means that, working with the vector Aq, where q belongs to the mth layer, we can perform orthogonalization only with respect to the vectors $\bar{q}_j$, where q_j belongs to either the same layer m or the layers $m - 1$ and $m + 1$. In other words, the depth of the recursion defining the next vector q does not exceed $\tilde{\omega}_{m-1} + \tilde{\omega}_m + \tilde{\omega}_{m+1}$.

If $\hat{A}$ satisfies the equation

$$g(\hat{A}, \hat{A}^*) = 0, \tag{22}$$

where $g(x, y)$ is a polynomial of degree $k \ll n$, then, beginning at $i = k$, the orders of the diagonal blocks H_{ii} stabilize at k.

The method MINRES-CN, which we propose, combines the orthogonalization of sequence (20) (which is not constructed explicitly) with the Galerkin principle as applied to the trial subspaces $\mathcal{L}_k$ (see (11)). As in CSYM, we expect that an approximate solution of reasonable quality can already be found for $k \ll n$; that is, we do not need to construct the entire basis $q_1, \ldots, q_n$ and the entire matrix H.

5. METHOD MINRES-CN2

MINRES-CN2 is a specialization of MINRES-CN for the case where the matrix $\hat{A}$ associated with a conjugate-normal matrix A satisfies Eq. (22) with a second-degree polynomial g. It can be shown (see [3]) that, in this case, A is either symmetric, skew-symmetric, or satisfies an equation of the form

$$aA\bar{A} + 2bAA^* + cA^{\mathrm{T}}A^* + dI_n = 0. \tag{23}$$

An equivalent description of this case can be given in terms of scalars known as the coneigenvalues of A (e.g., see [4, Section 3]); namely, all the coneigenvalues of a conjugate-normal matrix A must belong to a plane central curve of the second degree. Its equation has the form

$$\alpha x^2 + \beta y^2 + \gamma = 0. \tag{24}$$

The discussion in the preceding section implies that the condensed form H of A is a block tridiagonal matrix in which H_{11} is a scalar, while all the subsequent diagonal blocks have an order of two. The orthonormal sequence $q_1, q_2, \ldots$ is constructed in the following way:

1. The choice of a unit vector q_1 can be regarded as the first step. If one solves the linear system $Ax = b$, then vector (10) is an appropriate choice.

2. At the second step, the vector Aq_1 is orthogonalized to $\bar{q}_1$; that is,

$$w_2 = Aq_1 - h_{11}\bar{q}_1,$$

where

$$h_{11} = (Aq_1, \bar{q}_1). \tag{25}$$

Now, we set

$$h_{21} = \|w_2\|_2, \quad q_2 = \bar{w}_2/h_{21}.$$

(We assume that $h_{21} \neq 0$; similar assumptions are made at the subsequent steps.)

3. At the third step, the vector $A^{\mathrm{T}}q_1$ is orthogonalized to $\bar{q}_1$ and $\bar{q}_2$; thus,

$$w_3 = A^{\mathrm{T}}q_1 - h_{11}\bar{q}_1 - h_{12}\bar{q}_2.$$

Here, the coefficient

$$h_{11} = (A^{\mathrm{T}}q_1, \bar{q}_1) = (q_1, \bar{A}\bar{q}_1) = (Aq_1, \bar{q}_1)$$

is already known (see (25)), while

$$h_{12} = (A^{\mathrm{T}}q_1, \bar{q}_2).$$

After calculating w_3, we set

$$h_{13} = \|w_3\|_2, \quad q_3 = \bar{w}_3/h_{13}.$$

The second and third steps produce the vectors q_2 and q_3, which, in combination with q_1, form an orthonormal basis in the subspace

$$\mathcal{L}_1 = \mathrm{span}\{q_1, \bar{A}\bar{q}_1, A^*\bar{q}_1\}, \tag{26}$$

spanned by the initial segment of sequence (20). This completes the processing of the first layer in this sequence.

The purpose of several subsequent steps is to orthogonalize the vectors in the second layer of sequence (20) to the already available vectors q_1, q_2, and q_3. There are three vectors in the second layer; however, according to (23), only two of them are linearly independent over subspace (26). Consequently, it suffices to orthogonalize two vectors in the second layer to q_1, q_2, and q_3. Recall that sequence (20) is not constructed explicitly and the vectors in its second layer can be obtained by multiplying A and A^{T} by the vectors q_2 and q_3 in the first layer and by applying the componentwise conjugation to the calculated products. Since only two linearly independent vectors over (26) are required, we can calculate the products Aq_2 and Aq_3. Thus, the fourth and fifth steps are performed as follows.

4. The vector Aq_2 is orthogonalized to $\bar{q}_1$, $\bar{q}_2$, and $\bar{q}_3$; that is,

$$w_4 = Aq_2 - h_{12}\bar{q}_1 - h_{22}\bar{q}_2 - h_{32}\bar{q}_3.$$

Here, h_{12} is the coefficient calculated at the preceding step, while

$$h_{22} = (Aq_2, \bar{q}_2), \quad h_{32} = (Aq_2, \bar{q}_3).$$

After finding w_4, we set

$$h_{42} = \|w_4\|_2, \quad q_4 = \overline{w}_4/h_{42}.$$

5. The vector Aq_3 is orthogonalized to $\bar{q}_1$, $\bar{q}_2$, $\bar{q}_3$, and $\bar{q}_4$; thus,

$$w_5 = Aq_3 - h_{13}\bar{q}_1 - h_{23}\bar{q}_2 - h_{33}\bar{q}_3 - h_{43}\bar{q}_4.$$

Here, only h_{13} is known (after the third step); three remaining coefficients are calculated anew by the formulas

$$h_{23} = (Aq_3, \bar{q}_2), \quad h_{33} = (Aq_3, \bar{q}_3), \quad h_{43} = (Aq_3, \bar{q}_4).$$

After calculating w_5, we set

$$h_{53} = \|w_5\|_2, \quad q_5 = \overline{w}_5/h_{53}.$$

Now, we have found the vectors q_4 and q_5, which, in combination with q_1, q_2, and q_3, form an orthonormal basis in the subspace

$$\mathcal{L}_2 = \operatorname{span}\{q_1, \overline{A}\bar{q}_1, A^*\bar{q}_1, \overline{A}Aq_1, \overline{A}A^{\mathrm{T}}q_1, A^*A^{\mathrm{T}}q_1\}$$

spanned by the zeroth, first, and second layers in sequence (20). Moreover, we have calculated the entries in the first three columns of the condensed form H.

At the next two steps, we work with the matrix-vector products Aq_4 and Aq_5. The new fact is that these vectors do not need to be orthogonalized to $\bar{q}_1$. Indeed, we already know that the entries (1,4) and (1,5) of H are zero. This is also obvious from geometric considerations: the scalar product

$$(Aq_4, \bar{q}_1) = (q_4, A^*\bar{q}_1)$$

must vanish because $A^*\bar{q}_1$ belongs to the subspace $\mathcal{L}_1$, while q_4 is orthogonal to this subspace. The same is true of the scalar product $(Aq_5, \bar{q}_1)$.

Thus, the vectors q_6 and q_7 are found from the relations

$$h_{64}\bar{q}_6 = Aq_4 - h_{24}\bar{q}_2 - h_{34}\bar{q}_3 - h_{44}\bar{q}_4 - h_{54}\bar{q}_5$$

and

$$h_{75}\bar{q}_7 = Aq_5 - h_{25}\bar{q}_2 - h_{35}\bar{q}_3 - h_{45}\bar{q}_4 - h_{55}\bar{q}_5 - h_{65}\bar{q}_6.$$

In the last relation, the recursion attains its maximal depth of six.

The subsequent steps are similar. The vector $\bar{q}_{2m}$ $(m \geq 4)$ is calculated by orthogonalizing Aq_{2m-2} to the vectors $\bar{q}_{2m-4}$, $\bar{q}_{2m-3}$, $\bar{q}_{2m-2}$, and $\bar{q}_{2m-1}$. Then, Aq_{2m-1} is orthogonalized to the same four vectors and the found vector $\bar{q}_{2m}$. Normalizing the resulting vector, we obtain $\bar{q}_{2m+1}$.

Now, we discuss the procedure for calculating an approximate solution to the system $Ax = b$. Suppose that $2m + 1$ steps of the above recursion have already been completed. Then, we have the orthonormal vectors $q_1, q_2, \ldots, q_{2m+1}$ and the first $2m - 1$ columns of the block tridiagonal matrix H. Furthermore, the vector relations describing the orthogonalization process can be combined into a single matrix equality

$$AQ_{2m-1} = \overline{Q}_{2m-1}H_{2m-1} + R_{2m-1}. \tag{27}$$

Here,

$$Q_{2m-1} = (q_1q_2\ldots q_{2m-1}),$$

and H_{2m-1} is the order $2m-1$ leading principal submatrix of H. The matrix residual R_{2m-1} has zero columns, except for the last two ones, which have the form

$$r_{2m-2} = h_{2m,2m-2}\bar{q}_{2m} \tag{28}$$

and

$$r_{2m-1} = h_{2m,2m-1}\bar{q}_{2m} + h_{2m+1,2m-1}\bar{q}_{2m+1}. \tag{29}$$

As an approximate solution x_m, we take a vector x in the subspace

$$\mathcal{L}_{m-1} = \operatorname{span}\{q_1, \ldots, q_{2m-1}\}$$

for which the residual

$$r(x) = b - Ax$$

is orthogonal to the subspace

$$\overline{\mathcal{L}}_{m-1} = \operatorname{span}\{\bar{q}_1, \ldots, \bar{q}_{2m-1}\}.$$

In other words, x_m is determined by the conditions

$$Q_{2m-1}^{\mathrm{T}}(b - Ax_m) = 0, \quad x_m \in \mathcal{L}_{m-1}. \tag{30}$$

We seek x_m in the form

$$x_m = \eta_1 q_1 + \ldots + \eta_{2m-1}q_{2m-1} = Q_{2m-1}y_m, \tag{31}$$

where

$$y_m = (\eta_1, \ldots, \eta_{2m-1})^{\mathrm{T}}.$$

Substituting (31) into (30) and using (27), we have

$$0 = Q_{2m-1}^{\mathrm{T}}(b - AQ_{2m-1}y_m) = Q_{2m-1}^{\mathrm{T}}b - H_{2m-1}y_m - Q_{2m-1}^{\mathrm{T}}R_{2m-1}y_m. \tag{32}$$

Relations (10), (28), and (29) imply that

$$Q_{2m-1}^{\mathrm{T}}b = \|b\|_2 e_1^{(2m-1)}, \quad Q_{2m-1}^{\mathrm{T}}R_{2m-1} = 0.$$

Here, $e_1^{(2m-1)}$ is the first coordinate vector of the space $\mathbb{C}^{2m-1}$. Equalities (32) take the form

$$H_{2m-1}y_m = \|b\|_2 e_1^{(2m-1)}. \tag{33}$$

If H_{2m-1} is a nonsingular matrix, then system (33) determines a unique vector y_m and, hence, the desired approximate solution x_m. The block tridiagonal form of this matrix makes it possible to solve (33) using $O(m^2)$ operations.

However, neither x_m nor the (complete) vector y_m have to be calculated as long as the residual $r(x_m)$ is large. The norm of the residual can be found even if only the last two components of y_m are known. Indeed, $r(x_m)$ belongs to the subspace

$$\overline{\mathcal{L}}_m = \operatorname{span}\{\bar{q}_1, \ldots, \bar{q}_{2m-1}, \bar{q}_{2m}, \bar{q}_{2m+1}\}.$$

In view of (30), it has zero components along the vectors $\bar{q}_1, \ldots, \bar{q}_{2m-1}$. The two remaining components corresponding to $\bar{q}_{2m}$ and $\bar{q}_{2m+1}$ are

$$h_{2m,2m-2}\eta_{2m-2} + h_{2m,2m-1}\eta_{2m-1} \quad \text{and} \quad h_{2m+1,2m-1}\eta_{2m-1}.$$

Thus, we have

$$\|r(x_m)\|_2^2 = |h_{2m,2m-2}\eta_{2m-2} + h_{2m,2m-1}\eta_{2m-1}|^2 + |h_{2m+1,2m-1}\eta_{2m-1}|^2. \tag{34}$$

When passing from m to $m+1$, we can update value (34) in the same economical way as in GMRES and MINRES-N2 (see [5, Sections 5, 6]).

6. NUMERICAL RESULTS

In this section, we discuss some numerical experiments in which MINRES-CN2 was compared with GMRES. In this comparison, the convergence rate was the basic performance criterion. It is well known (see [6]) that, by an appropriate unitary congruence, every conjugate-normal matrix can be transformed into a real block diagonal matrix with diagonal blocks of orders one and two. Furthermore, the 2-by-2 blocks have the form

$$\begin{pmatrix} a_j & -b_j \\ b_j & a_j \end{pmatrix}, \quad b_j \neq 0. \tag{35}$$

It is obvious that such a block is a normal matrix, which implies that the entire block diagonal matrix is normal. Thus, to get an idea of the convergence rate of methods for conjugate-normal matrices, it suffices to examine the systems $Ax = b$ with matrices of the above form (despite the fact that any of these systems can trivially be solved using $O(n)$ arithmetic operations).

Block (35) has complex conjugate eigenvalues

$$a_j \pm ib_j.$$

If $a_j \geq 0$, these scalars are identical to the coneigenvalues of this block. Otherwise, the coneigenvalues are obtained by multiplying the eigenvalues by –1.

The blocks of order one (i.e., the diagonal entries) are eigenvalues of our matrix A, while their moduli are coneigenvalues of this matrix.

Thus, the spectrum of A and its conspectrum do not differ if all the diagonal entries belong to the nonnegative semiaxis.

Recall that our conjugate-normal matrix A must satisfy an additional condition, namely, its conspectrum must lie on a second-degree central curve. This condition was met in the following way: for a chosen curve of type (24), the abscissas a_j were pseudorandom numbers distributed uniformly on an appropriate interval. The ordinates $\pm b_j$ corresponding to these a_j were determined from the equation of the curve. All the matrices in our experiments were well conditioned because the semiaxes of ellipses were not too different from each other and no points with large coordinates were taken on hyperbolas. In addition, no points close to the origin were taken on pairs of lines.

In all of our experiments, the order of systems was 2000. The right-hand sides were generated as pseudorandom vectors with components distributed uniformly on (0, 1). The calculations were performed on a 2 Duo E630 OEM 1.86 GHz PC with core memory of 1024 Mb.

GMRES was represented by the Matlab procedure gmres; for MINRES-CN2, we designed our own Matlab procedure.

To terminate the iteration, we used the condition

$$\|r(x_m)\|_2 < \varepsilon, \tag{36}$$

where ε is a given positive scalar. The value $\|r(x_m)\|_2$ in MINRES-CN2 is calculated as explained in Section 5.

Example 1. The eigenvalues of A are uniformly distributed on the ellipse

$$x^2/16 + y^2/9 = 1.$$

For $\varepsilon = 10^{-6}$, the number of iteration steps in MINRES-CN2 is half as much as in GMRES. This double superiority is somewhat lost when we compare the corresponding times: 0.59 s for MINRES-CN2 against 0.97 s for GMRES. A possible explanation is that our experimental program competed with a well-polished procedure from Matlab.

For both methods, the residual norms are plotted as functions of the iteration step index m in Fig. 1.

In all the plots, $\|b - Ax^m\|$ is shown by $\times$ and $\|b - Ax^{\mathrm{GMRES}}\|$ is shown by $\diamond$.

Example 2. The eigenvalues of A are uniformly distributed on the right half of the ellipse

$$\frac{x^2}{100^2} + \frac{y^2}{9} = 1.$$

By contrast with Example 1, GMRES converges faster than MINRES-CN2 (48 iteration steps and t = 0.58 s against 69 steps and t = 0.81 s; ε is the same as in Example 1). This is a manifestation of a well-known

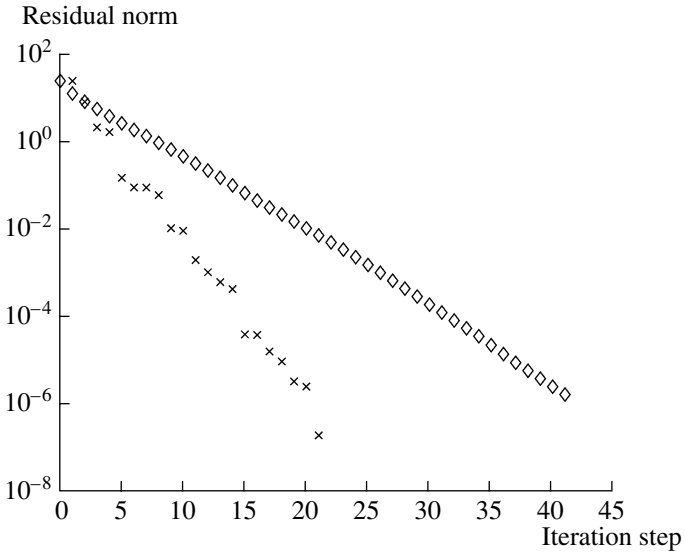

Fig. 1.

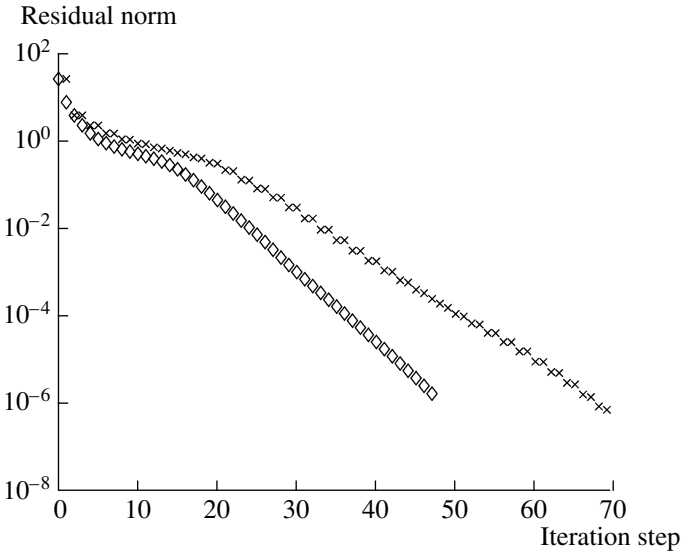

Fig. 2.

fact; namely, the situation where the entire spectrum of a system is contained in a half-plane whose boundary passes through the origin is favorable for GMRES.

The graphs of the residual norms are shown in Fig. 2.

Example 3. The spectrum of A is uniformly distributed on the segments of both halves of the hyperbola

$$x^2/4 - y^2/9 = 1, \tag{37}$$

corresponding to x for which

$$4 < |x| < 6.$$

The behavior of both methods is similar to what we have seen in the case of the whole ellipse: for $\varepsilon = 10^{-8}$, GMRES needs 97 steps and the time $t = 2.6$ s, while MINRES-CN2 requires 59 steps and the time 1.4 s.

The graphs of the residual norms are shown in Fig. 3.

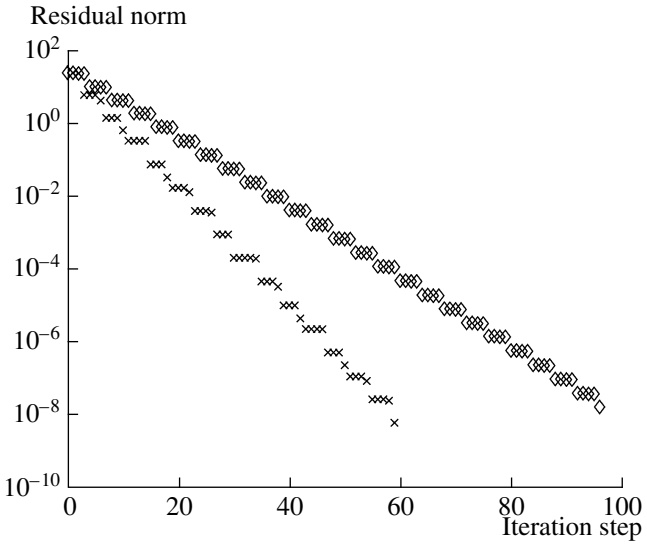

Fig. 3.

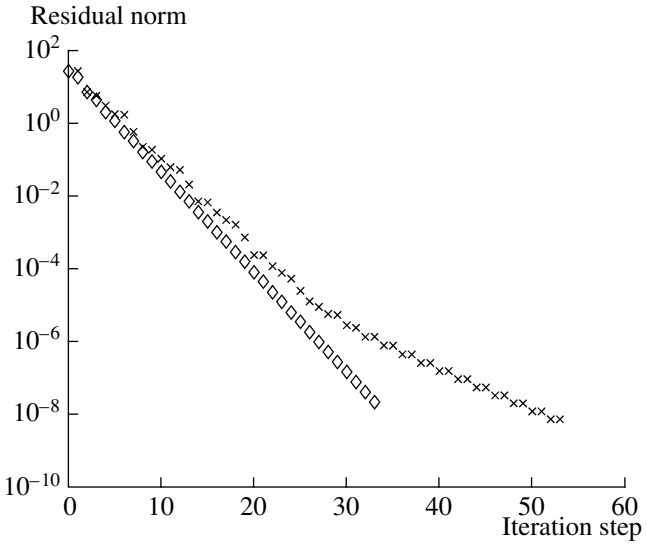

Fig. 4.

Example 4. The spectrum of A is uniformly distributed on the segments of the right half of hyperbola (37) corresponding to x in the interval (2, 3); $\varepsilon = 10^{-8}$. Here, we observe the same phenomenon as in Example 2 (and the explanation is the same as in that example): GMRES converges faster than MINRES-CN2 (34 iteration steps and $t = 0.52$ s against 53 steps and $t = 0.69$ s).

The graphs of the residual norms are shown in Fig. 4.

Example 5. The spectrum of A is uniformly distributed on the segments of the pair of lines

$$y = \pm\frac{3}{4}x,$$

corresponding to x with

$$8/3 < |x| < 20/3.$$

We set $\varepsilon = 10^{-5}$. In this experiment, MINRES-CN2 demonstrated a significantly better performance than GMRES: 71 iteration steps and $t = 1.61$ s against 157 steps and $t = 4.78$ s.

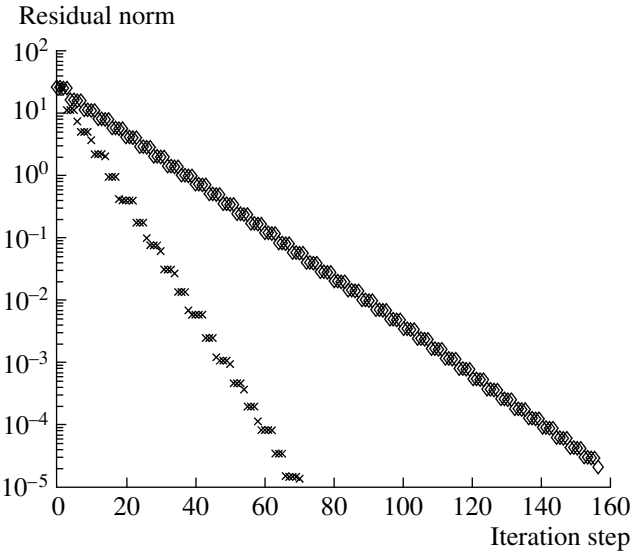

Fig. 5.

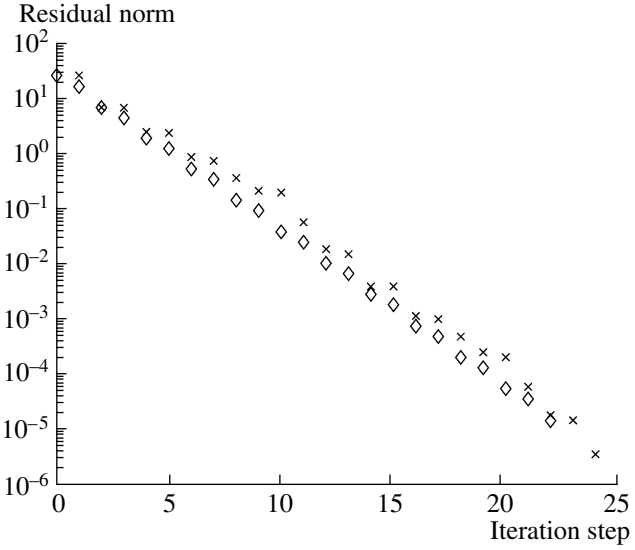

Fig. 6.

The graphs of the residual norms are shown in Fig. 5.

Example 6. For the same pair of lines as in the preceding example, the eigenvalues of A were chosen only on the right rays, that is, for x in the interval (8/3, 20/3); ε is again 10^{-5}. The convergence rate was about the same for both methods: 25 iteration steps and $t = 0.66$ s for MINRES-CN2 against 23 steps and $t = 0.62$ s for GMRES.

The graphs of the residual norms are shown in Fig. 6.

Example 7. One-half of the eigenvalues of A are real and are uniformly distributed on the intervals

$$2 < |x| < 5.$$

The remaining eigenvalues are purely imaginary and are uniformly distributed on the intervals

$$2 < |y| < 5. \tag{38}$$

For $\varepsilon = 10^{-5}$, the situation is very similar to that observed in Example 5: MINRES-CN2 needs 71 steps and $t = 1.72$ s, while GMRES requires 169 steps and the time 5.34 s.

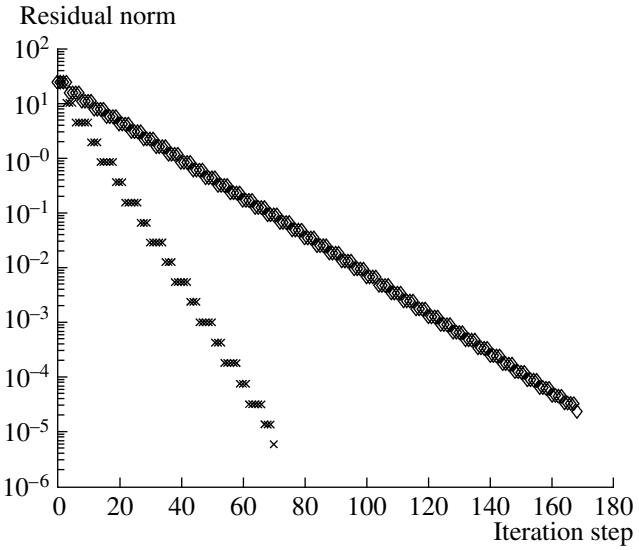

Fig. 7.

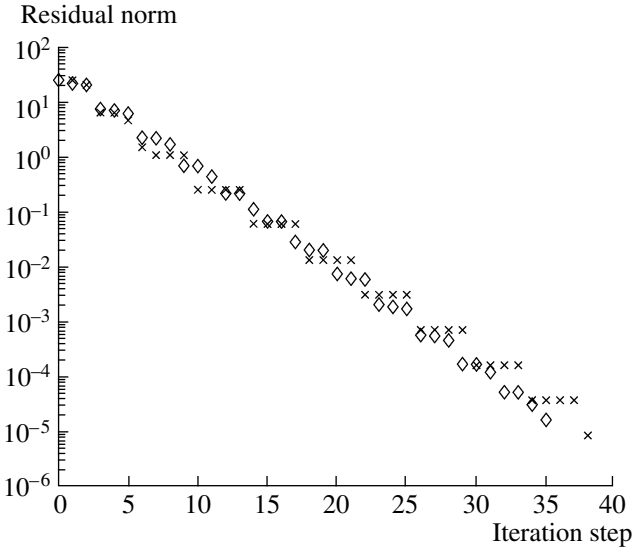

Fig. 8.

The graphs of the residual norms are shown in Fig. 7.

Example 8. Here, we use the same pair of lines $xy = 0$ as in the preceding example. One-half of the eigenvalues are purely imaginary and are uniformly distributed on the intervals

$$5 < |y| < 8.$$

The remaining eigenvalues are positive and are uniformly distributed on the interval (5, 8). The value of ε remains as before. The convergence of both methods is the same as in Example 6: they require the same time (0.97 s) and about the same number of iterations steps (39 steps for MINRES-CN2 against 36 steps for GMRES).

The graphs of the residual norms are shown in Fig. 8.

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, "Conjugate-Normal Matrices and Matrix Equations in A, $\bar{A}$, and A^{T}," Dokl. Akad. Nauk **412**, 305–307 (2007) [Dokl. Math. **75**, 55–57 (2007)].
4. Kh. D. Ikramov, "On Tridiagonal Conjugate-Normal Matrices," Zh. Vychisl. Mat. Mat. Fiz. **47**, 179–185 (2007) [Comput. Math. Math. Phys. **47**, 173–179 (2007)].
5. M. Dana, A. G. Zykov, and Kh. D. Ikramov, "A Minimal Residual Method for a Special Class of Linear Systems with Normal Coefficients Matrices," Zh. Vychisl. Mat. Mat. Fiz. **45**, 1928–1937 (2005) [Comput. Math. Math. Phys. **45**, 1854–1863 (2005)].
6. H. Fassbender and Kh. D. Ikramov, "Some Observations on the Youla Form and Conjugate-Normal Matrices," Linear Algebra Appl. **422**, 29–38 (2007).