# CALCULATING AN INVARIANT SUBSPACE OF DIAGONALLY DOMINANT MATRICES – PART I

Noah H. Rhee

University of Missouri – Kansas City

**Abstract.** In this paper a modification of the algorithm of Blevins and Stewart for calculating an invariant subspace of diagonally dominant matrices is given. The relation between the algorithm of Blevins and Stewart and our modified algorithm is essentially the same as the relation between the Jacobi method and the Gauss–Seidel method in solving linear systems iteratively.

**1. Introduction.** Let $A$ be a real matrix of order $n$, and write $A$ in the form

$$A = D + E,$$

where

$$D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$$

is the diagonal part of $A$. Following the definition in [1] we shall say that $A$ is diagonally dominant if

$$||E|| = \sigma ||D||,$$

where $\sigma < 1$. Here $||\cdot||$ denotes the Frobenius matrix norm defined by

$$||C||^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}^2,$$

where $C = (c_{ij})$ is any real $m \times n$ matrix. In this paper we are concerned with algorithms for computing an invariant subspace of $A$ when $\sigma << 1$.

The calculation of an invariant subspace of diagonally dominant matrices arises in the problem of refining systems of approximate eigenvectors. [1]

Throughout this paper $\mathbb{R}^n$ denotes real Euclidean space of dimension $n$ and $\mathbb{R}^{m\times n}$ denotes the set of $m \times n$ real matrices. Linear operators on a vector space, as opposed to their matrix representations, are denoted by bold face letters. The superscript $T$ will denote the transpose of a given vector or a matrix.

For the linear operator $\mathbf{T}$ from $\mathbb{R}^{m\times n}$ to $\mathbb{R}^{k\times l}$, we define the spectral norm of $\mathbf{T}$ by

$$||\mathbf{T}||_2 = \sup_{||P||=1} ||\mathbf{T}P||.$$

We note that for any matrices $B$ and $C$ regarded as operators, $||BC|| \leq ||B||_2||C||$ and $||BC|| \leq ||B||||C||_2$, whenever the product $BC$ is defined. We also note that $||B||_2 \leq ||B||$.

In section 2, we briefly summarize the main results of Blevins and Stewart [1]. In section 3, we discuss some further results and give a modified algorithm of Blevins and Stewart. In part 2 of this article, we will give some numerical results which demonstrate the faster convergence of the modified algorithm.

**2. The Main Results of Blevins and Stewart.** Let $A \in \mathbb{R}^{n\times n}$ be a diagonally dominant matrix. There is a natural correspondence between the eigenvectors of $A$ and those of $D$, since $A$ is a diagonally dominant matrix. If $d_i$ is distinct from the other elements of $D$ and if $\sigma$ is sufficiently small, then the Gerschgorin theorem shows that there is a unique eigenvalue of $A$ which is near $d_i$. Since the eigenvector corresponding to the distinct eigenvalue is a continuous function of the entries of the matrix (see [5], p. 67), if the $i$th components of the corresponding eigenvectors of $D$ and $A$ are normalized to unity the other components of the eigenvectors of $A$ must be small. For definiteness we shall compare the eigenvectors corresponding to the diagonal element $d_1$. If we write these eigenvectors of $D$ and $A$ in the forms $(1,0)^T$ and $(1,p^T)^T$, respectively, then $p \in \mathbb{R}^{n-1}$ is small.

The first step is to find the equation satisfied by $p$. Because $(1,p^T)^T$ is an eigenvector of $A$, the vector $A(1,p^T)^T$ lies in the same direction as $(1,p^T)^T$. But the matrix

$$\begin{pmatrix} -p^T \\ I_{n-1} \end{pmatrix}$$

has $n-1$ linearly independent columns, all of which are orthogonal to $(1, p^T)^T$. Here $I_k$ denotes the identity matrix in $\mathbb{R}^{k\times k}$. It follows that

$$\begin{pmatrix} -p^T \\ I_{n-1} \end{pmatrix}^T A \begin{pmatrix} 1 \\ p \end{pmatrix} = 0. \tag{1}$$

If we partition $A$ in the form

$$A = \begin{pmatrix} d_1 & \epsilon_{12}^T \\ \epsilon_{21} & D_2 + E_{22} \end{pmatrix},$$

then (1) becomes

$$Tp = \epsilon_{21} - p\epsilon_{12}^T p, \tag{2}$$

where $T$ is the matrix $T = d_1 I_{n-1} - (D_2 + E_{22})$.

Any method for solving the nonlinear equation (2) for $p$ is effectively a method for computing the eigenvector corresponding to the eigenvalue near to $d_1$. In theory, this is not difficult to do. In [4] it is shown that under the following condition,

$$||T^{-1}||_2^2 ||\epsilon_{21}|| \, ||\epsilon_{12}|| < \frac{1}{4}, \tag{3}$$

the sequence defined by

$$Tp_{k+1} = \epsilon_{21} - p_k \epsilon_{12}^T p_k, \quad p_0 = 0 \tag{4}$$

converges to a solution of (2). As a practical method, however, this iteration has two drawbacks. First, condition (3) may not be satisfied, that is, $d_1$ may not be sufficiently well separated from the other diagonal elements of $A$. Second, the solution of equation (4) for $p_{k+1}$ is expensive.

Blevins and Stewart disposed of the first problem by showing how to find an invariant subspace corresponding to a cluster of close eigenvalues. Then they disposed of the second problem by using an approximated inverse of $T$.

To explain the idea of Blevins and Stewart we define the following.

Definition 1. Let $\mathfrak{R}(A)$ denote the column space of the matrix $A$.

Definition 2. For a subspace $\Omega \subseteq \mathbb{R}^n$, let $\dim(\Omega)$ denote the dimension of the subspace $\Omega$.

Definition 3. A subspace $\Omega \subseteq \mathbb{R}^n$ is an invariant subspace of $A$ if $A\Omega \subseteq \Omega$.

We note that an eigenvector of $A$ spans an invariant subspace of dimension unity. In order to compute an invariant subspace $\Omega$ of $A$, Blevins and Stewart set up an equation analogous to (2) for a basis for $\Omega$. The following lemma indicates how this was done.

Lemma 4. Let $A \in \mathbb{R}^{n\times n}$ and let $X = [X_1 \;\; X_2]$ be such that $X_1 \in \mathbb{R}^{n\times l}$, $X_2 \in R^{n\times(n-l)}$, $X_1^T X_2 = 0$ and the columns of $X_1(X_2)$ are linearly independent. Then a necessary and sufficient condition that $\mathfrak{R}(X_1)$ be an invariant subspace of $A$ is

$$X_2^T A X_1 = 0. \tag{5}$$

Proof. This lemma is a modification of the Lemma 3.1 in [1]. Since

$$\dim(\mathfrak{R}(X_1)) + \dim(\mathfrak{R}(X_2)) = n$$

and $X_1^T X_2 = 0$, $\mathfrak{R}(X_2)$ is the orthogonal complement of $\mathfrak{R}(X_1)$ in $\mathbb{R}^n$. The remaining proof is identical with the proof given in [1].

For definiteness suppose that the first $l$ diagonal elements of $A$ form a cluster that is well separated from the other diagonal elements, and partition $A$ in the form

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} D_1 + E_{11} & E_{12} \\ E_{21} & D_2 + E_{22} \end{pmatrix},$$

where $D_1 \in \mathbb{R}^{l\times l}$. And they attempted to find a basis for an invariant subspace of $A$ that in some sense corresponds to the matrix $D_1$.

Because $E_{21}$ is presumed small, equation (5) is very nearly satisfied by the matrix

$$X = \begin{pmatrix} I_l & 0 \\ 0 & I_{n-l} \end{pmatrix}.$$

This suggests that we seek $X$ in the form

$$X = [X_1 \;\; X_2] = \begin{pmatrix} I_l & -P^T \\ P & I_{n-l} \end{pmatrix},$$

where $P \in \mathbb{R}^{(n-l)\times l}$ and $||P||$ is small. Note that the columns of $X_1(X_2)$ are linearly independent and $X_1^T X_2 = 0$. Hence, by Lemma 4 the necessary and sufficient condition that $\mathfrak{R}(X_1)$ be an invariant subspace of $A$ is that

$$X_2^T A X_1 = 0.$$

With the partitions of $A$ and $X$, the above equation becomes

$$PA_{11} - A_{22}P = E_{21} - PE_{12}P. \tag{6}$$

Then they defined the linear operator $\mathbf{T}: \mathbb{R}^{(n-l)\times l} \to \mathbb{R}^{(n-l)\times l}$ by

$$\mathbf{T}P = PA_{11} - A_{22}P.$$

So equation (6) becomes

$$\mathbf{T}P = E_{21} - PE_{12}P. \tag{7}$$

Equation (7) is perfectly analogous to equation (2), and it has a small solution under analogous conditions. Since $l$ is unrestricted, one is free to augment $D_1$ until the conditions for the existence of a solution are satisfied, which disposes of the problem of very close diagonal elements.

Equation (7) can be solved by an iterative process analogous to (4); however, each iteration requires the solution of an equation of the form $\mathbf{T}P = Q$, which is prohibitively expensive, even when $l = 1$. However, $\mathbf{T}$ can be written $\mathbf{T} = \mathbf{D} + \mathbf{E}$, where

$$\mathbf{D}P = PD_1 - D_2P. \tag{8}$$

and

$$\mathbf{E}P = PE_{11} - E_{22}P. \tag{9}$$

Then they proposed the following iterative process to solve equation (7).

Algorithm 5.

(1) $P_0 = 0$

(2) For $k = 0, 1, \ldots$

$$P_{k+1} = \Phi(P_k),$$

where

$$\Phi(P) = \mathbf{D}^{-1}(E_{21} - PE_{12}P - \mathbf{E}P). \tag{10}$$

Since $E_{11}$ and $E_{22}$ are small, the operator $\mathbf{D}$ is near the operator $\mathbf{T}$. Since $D_1$ and $D_2$ are diagonal matrices, the equation of the form $\mathbf{D}P = Q$ can be easily solved. In fact, the matrix representation of the linear operator $\mathbf{D}$ is a diagonal matrix. This can be seen as follows.

Suppose $\mathbf{D}P = Q$. Let $P = (p_{ij})$ and $Q = (q_{ij})$, $l+1 \le i \le n$, $1 \le j \le l$. Here we shifted the index $i$ by $l$ for notational convenience. From equation (8) we have

$$PD_1 - D_2P = Q.$$

This equation is equivalent to

$$(d_j - d_i)p_{ij} = q_{ij}, \quad l+1 \le i \le n, \ 1 \le j \le l. \tag{11}$$

Note that $q_{ij}$ depends only on $p_{ij}$, $l+1 \le i \le n$, $1 \le j \le l$. This means that the matrix representation of the linear operator $\mathbf{D}$ is a diagonal matrix.

The iteration defined by Algorithm 5 converges to a solution of equation (7) under rather general conditions. They proved the following Theorem.

Theorem 6. Let $\eta = ||E_{12}||$, $\gamma = ||E_{21}||$, $\epsilon = ||\mathbf{E}||_2$, and $\delta = ||\mathbf{D}^{-1}||_2^{-1}$. Then if

$$\delta - \epsilon > 2\sqrt{\eta\gamma}, \tag{12}$$

the sequence $P_k$ defined by Algorithm 5 converges to a solution $P^*$ of equation (7) satisfying

$$||P^*|| \le \frac{2\gamma}{\delta - \epsilon}. \tag{13}$$

Moreover

$$(14) \qquad ||P^* - P_k|| \le \frac{\rho}{1-\rho}||P_k - P_{k-1}||, \quad k = 1, 2, \dots,$$

and

$$(15) \qquad ||P_{k+1} - P_k|| \le \rho||P_k - P_{k-1}||, \quad k = 1, 2, \dots,$$

where

$$(16) \qquad \rho = \frac{\epsilon}{\delta} + \frac{4\eta\gamma}{\delta(\delta - \epsilon)} < 1.$$

To find out eigenvalues and eigenvectors from the given invariant subspace, they also proved the following theorem.

Theorem 7. Let $P^*$ be determined as in Theorem 6. Then the eigenvalues of $A$ corresponding to the invariant subspace determined by $P^*$ are the eigenvalues of the matrix $A_{11} + E_{12}P^*$. Moreover, if the columns of $Z$ form a complete set of eigenvectors for $A_{11} + E_{12}P^*$, the corresponding eigenvectors of $A$ are the columns of the matrix

$$\begin{pmatrix} I \\ P^* \end{pmatrix} Z = \begin{pmatrix} Z \\ P^* Z \end{pmatrix}.$$

**3. A Modified Algorithm.** The idea of proving Theorem 6 in [1] is as follows. First, it was shown that all the iterates $P_k$ generated by Algorithm 5 remain in the region defined by (13). Second, it was shown that the function $\Phi$ is a contraction, with constant $\rho$, in that region. Then the result with error bounds (14) and (15) follow from a variant of the contraction mapping theorem.

But we can prove a stronger result by using the standard contraction mapping theorem.

Theorem 8. Contraction Mapping Theorem.

Suppose that $\mathbf{G}: D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ maps a closed set $D_0 \subseteq D$ into itself and that

$$||\mathbf{G}x - \mathbf{G}y|| \le \rho||x - y||, \quad x, y \in D_0,$$

for some $\rho < 1$. Then, for any $x_0 \in D_0$, the sequence generated by

$$x_{k+1} = \mathbf{G}x_k, \quad k = 0, 1, \ldots,$$

converges to the unique fixed point $x^*$ of $\mathbf{G}$ (that is, $\mathbf{G}(x^*) = x^*$), in $D_0$ and

$$||x^* - x_k|| \leq \frac{\rho}{1-\rho}||x_k - x_{k-1}||, \quad k = 1, 2, \ldots .$$

<u>Proof</u>. (See [3], p. 385) . Let

$$\Gamma = \{P \in \mathbb{R}^{(n-l)\times l} : ||P|| \leq \frac{2\gamma}{\delta - \epsilon}\}. \tag{17}$$

By using the above standard Contraction Mapping Theorem, we can improve Theorem 6 as follows.

<u>Theorem 9</u>. Let $\eta = ||E_{12}||$, $\gamma = ||E_{21}||$, $\epsilon = ||\mathbf{E}||_2$, and $\delta = ||\mathbf{D}^{-1}||_2^{-1}$. If

$$\delta - \epsilon > 2\sqrt{\eta\gamma}, \tag{18}$$

then, for any $P_0 \in \Gamma$, the sequence $P_k$ defined by $P_{k+1} = \Phi(P_k)$, $k = 0, 1, \ldots$, converges to the unique solution $P^*$ of equation (7) in $\Gamma$. Moreover

$$||P^* - P_k|| \leq \frac{\rho}{1-\rho}||P_k - P_{k-1}||, \quad k = 1, 2, \ldots, \tag{19}$$

and

$$||P_{k+1} - P_k|| \leq \rho||P_k - P_{k-1}||, \quad k = 1, 2, \ldots, \tag{20}$$

where

$$\rho = \frac{\epsilon}{\delta} + \frac{4\eta\gamma}{\delta(\delta - \epsilon)} < 1. \tag{21}$$

<u>Proof</u>. The fact that $\Phi$ is a contraction in $\Gamma$ has been established in [1], with the contraction constant $\rho$ described in the relation (21). Clearly $\Gamma$ is a closed set in $\mathbb{R}^{(n-l)\times l}$. Now we establish the fact that $\Phi(\Gamma) \subseteq \Gamma$. Suppose that $||P|| \leq 2\gamma/(\delta - \epsilon)$ and

$$||\Phi(P)|| > 2\gamma/(\delta - \epsilon).$$

Then using relation (10) we obtain

$$\frac{2\gamma}{\delta-\epsilon} < ||\Phi(P)|| \leq \frac{1}{\delta}\left(\gamma + \eta||P||^2 + \epsilon||P||\right)$$

$$\leq \frac{1}{\delta}\left[\gamma + \eta\left(\frac{2\gamma}{\delta-\epsilon}\right)^2 + \epsilon\left(\frac{2\gamma}{\delta-\epsilon}\right)\right].$$

Equivalently,

$$2\delta(\delta-\epsilon) < (\delta-\epsilon)^2 + 4\eta\gamma + 2\epsilon(\delta-\epsilon).$$

However, it can be seen easily by using the assumption $\delta-\epsilon > 2\sqrt{\eta\gamma}$ that the right hand side of the above inequality is less than $2\delta(\delta-\epsilon)$. So we obtain a contradiction. It follows that $\Phi(\Gamma) \subseteq \Gamma$. Hence, by the Contraction Mapping Theorem the sequence $\{P_k\}$ converges to the unique fixed point $P^*$ of $\Phi$ in $\Gamma$. This means that $P^*$ is the unique solution of $P^* = \Phi(P^*)$ in $\Gamma$. By using the definition of $\Phi$ in (10) we see that $P^*$ is the unique solution of equation (7) in $\Gamma$.

Remark 10. The differences between Theorem 6 and Theorem 9 are the following. First, Theorem 9 gives the information that $P^*$ is the unique solution of (7) in $\Gamma$, whereas Theorem 6 does not. Second, in Theorem 6 $P_0$ has to be 0, whereas in Theorem 9, $P_0$ can be any element in $\Gamma$. This fact will be used crucially in our modified algorithm.

Now we are going to suggest a modified algorithm. Recall that ((8), (9))

$$\mathbf{T} = \mathbf{D} + \mathbf{E},$$

where

$$\mathbf{D}P = PD_1 - D_2P,$$

and

$$\mathbf{E}P = PE_{11} - E_{22}P.$$

Now we let

$$\mathbf{T} = \mathbf{L} + \mathbf{D} + \mathbf{U},$$

where

$$\mathbf{L}P = PL_1 - L_2P,$$

and

$$\mathbf{U}P = PU_1 - U_2P.$$

Here $L_1(U_1)$ and $L_2(U_2)$ are strictly lower (upper) triangular parts of $A_{11}$ and $A_{22}$, respectively. We note that $L_1(U_1)$ and $L_2(U_2)$ can also be viewed as the strictly lower (upper) triangular parts of $E_{11}$ and $E_{22}$, respectively.

Then the corresponding algorithm to Algorithm 5, based on the above **LDU** splitting of $\mathbf{T}$, is

Algorithm 10.

(1) $\overline{P}_0 = 0$

(2) For $k = 0, 1, \ldots,$

$$\overline{P}_{k+1} = \Psi(\overline{P}_k),$$

where

$$\Psi(\overline{P}) = (\mathbf{D}+\mathbf{L})^{-1}\left(E_{21} - \overline{P}E_{12}\overline{P} - \mathbf{U}\overline{P}\right). \tag{22}$$

Of course, we determine $\overline{P}_{k+1}$ by solving

$$(\mathbf{D}+\mathbf{L})\overline{P}_{k+1} = E_{21} - \overline{P}_k E_{12}\overline{P}_k - \mathbf{U}\overline{P}_k, \tag{23}$$

which is easy to solve, since $(\mathbf{D}+\mathbf{L})P = P(D_1+L_1) - (D_2+L_2)P$ and $D_i + L_i$ is a lower triangular matrix for $i = 1, 2$. In fact, there is a lower triangular matrix representation of the operator $\mathbf{D}+\mathbf{L}$. This can be seen as follows.

Suppose $(\mathbf{D}+\mathbf{L})P = Q$. Then we have

$$P(D_1+L_1) - (D_2+L_2)P = Q. \tag{24}$$

As before we let $P = (p_{ij})$ and $Q = (q_{ij})$, $l+1 \le i \le n$, $1 \le j \le l$. We also let $L_1 = (l_{ij})$, $l \le i, j \le l$, and $L_2 = (l_{ij})$, $l+1 \le i, j \le n$. If we use the fact that $L_1$ and $L_2$ are strictly lower triangular matrices, equation (24) is equivalent to

$$q_{ij} = (d_j - d_i)p_{ij} - \sum_{k=l+1}^{i-1} l_{ik}p_{kj} + \sum_{k=j+1}^{l} p_{ik}l_{kj}, \quad l+1 \le i \le n, \quad 1 \le j \le l. \tag{25}$$

Here the empty sums are zero.

Now we identify any matrix $V = (v_{ij}) \in \mathbb{R}^{(n-l)\times l}$, $l+1 \le i \le n$, $1 \le j \le l$, with the long vector

$$V = (v_{l+1,l}, v_{l+2,l}, \ldots, v_{n,l}; v_{l+1,l-1}, v_{l+2,l-1}, \ldots, v_{n,l-1}; \ldots; v_{l+1,1}, v_{l+2,1}, \ldots, v_{n,1})^T$$
$$(26) \quad \in \mathbb{R}^{l(n-l)}.$$

That is, we order $v_{ij}$ by the following principles. For any two elements of $v_{i_1 j_1}$ and $v_{i_2 j_2}$ of $V$, $v_{i_1 j_1}$ precedes $v_{i_2 j_2}$ if and only if one of the following conditions holds:

(1) $j_1 > j_2$

(2) $j_1 = j_2$ and $i_1 < i_2$.

From equation (25) we note that $q_{ij}$ depends only on $p_{ik}$ for which $j \le k \le l$ and $p_{kj}$ for which $l+1 \le k \le i$. Hence, according to our ordering $q_{ij}$ depends only on $p_{ij}$ and all the elements that precede it. This means that if we identify $\mathbb{R}^{(n-l)\times l}$ to $\mathbb{R}^{l(n-l)}$ with the above ordering, the matrix representation of $\mathbf{D}+\mathbf{L}$ is a lower triangular matrix in $\mathbb{R}^{l(n-l)\times l(n-l)}$. So equation (23) is easy to solve. Henceforth we shall use the ordering defined by relation (26).

We note that the cost of calculating $\overline{P}_{k+1}$ from $\overline{P}_k$ by using Algorithm 10 is exactly the same as calculating $P_{k+1}$ from $P_k$ using Algorithm 5 except calculating $\overline{P}_1$ from $\overline{P}_0 = 0$, which is more expensive than calculating $P_1$ from $P_0 = 0$.

The relation between Algorithm 5 and Algorithm 10 is essentially the same as the relation between the Jacobi method and the Gauss–Seidel method in solving linear systems iteratively. It is well known that in many applications the Gauss–Seidel method is faster than the Jacobi method (see [2], p. 324). In part 2 we will see numerically that Algorithm 10 is faster than Algorithm 5.

However, to establish a rigorous convergence criterion of Algorithm 10 we need to compute $\overline{\delta} = ||(\mathbf{D}+\mathbf{L})^{-1}||_2^{-1}$, which is prohibitively expensive to calculate. So we suggest the following algorithm.

Algorithm 11.

(1) $\overline{P}_0 = 0$

(2) For $k = 0, 1, \ldots,$

$\overline{P}_{k+1} = \Psi(\overline{P}_k)$.

If $||\overline{P}_{k+1}|| > 2\gamma/(\delta - \epsilon)$ or $||\overline{P}_{k+1} - \overline{P}_k|| > \rho||\overline{P}_k - \overline{P}_{k-1}||$ $(k \ge 1)$,

go to (3);

otherwise

go to (2).

(3) For $j = k, k+1, \ldots,$

$\overline{P}_{j+1} = \Phi(\overline{P}_j)$.

Remark 11. In step (2) the condition $||\overline{P}_{k+1}|| > 2\gamma/(\delta - \epsilon)$ means that $\{\overline{P}_k\}_{k=0}^{\infty}$ does not converge to $P^*$, and the condition $||\overline{P}_{k+1} - \overline{P}_k|| > \rho||\overline{P}_k - \overline{P}_{k-1}||$ $(k \geq 1)$ means that $\{P_k\}$ generated by Algorithm 5 will converge to $P^*$ faster than $\{\overline{P}_k\}$ generated by Algorithm 10.

Remark 12. The convergence of the sequence of $\{\overline{P}_k\}_{k=0}^{\infty}$, generated by Algorithm 11, is guaranteed by step (3) together with Theorem 9. So Algorithm 11 enjoys possible faster convergence of Algorithm 10 as well as the guaranteed convergence of Algorithm 5.

## *References*


1. M. M. Blevins and G. W. Stewart, "Calculating the Eigenvectors of Diagonally Dominant Matrices," *Journal of the Association for Computing Machinery,* 21 (1974), 261–271.

2. W. W. Hager, *Applied Numerical Linear Algebra,* Prentice Hall, Englewood Cliffs, N.J., 1988.

3. J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables,* Academic Press, New York, 1970.

4. G. W. Stewart, "Error Bounds for Invariant Subspaces of Closed Operators," *SIAM J. Numer. Anal.,* 8 (1971), 196–208.

5. J. H. Wilkinson, *The Algebraic Eigenvalue Problem,* Oxford Univ. Press, London/New York, 1965.