= MATHEMATICS ===

Theorems of the Hoffman–Wielandt Type for the Coneigenvalues of Complex Matrices

Kh. D. Ikramov and Yu. R. Nesterenko

Presented by Academician S.K. Korovin March 29, 2009

Received March 27, 2009

DOI: 10.1134/S106456240904022X

1. The Hoffman–Wielandt theorem (see [1] or [2, Chapter 2, Section 48]) is one of the most useful results in the finite-dimensional perturbation theory of eigenvalues. Here is the formulation of this theorem:

Theorem 1. Let A and B be normal matrices of order n having the eigenvalues $\alpha_1, \alpha_2, ..., \alpha_n$ and $\beta_1, \beta_2, ..., \beta_n$, respectively. Then, there exists a permutation π of the indices 1, 2, ..., n such that

$$\sum_{i=1}^{n} |\alpha_{i} - \beta_{\pi(i)}|^{2} \le ||A - B||_{F}^{2}.$$
 (1)

The symbol $\|\cdot\|_F$ stands for the Euclidean matrix norm (which is also called the Frobenius norm).

This theorem has several extensions in which the normality condition is dropped for one or both matrices.

Theorem 2. Let A be a normal matrix of order n and B be an arbitrary matrix of order n with eigenvalues α_1 , α_2 , ..., α_n and β_1 , β_2 , ..., β_n , respectively. Then, there exists a permutation π of the indices 1, 2, ..., n such that

$$\sum_{i=1}^{n} |\alpha_{i} - \beta_{\pi(i)}|^{2} \le n ||A - B||_{F}^{2}.$$
 (2)

This theorem was proved by Sun (see [3]). The factor n in bound (2) can be replaced by a smaller constant if A is a Hermitian matrix rather than a normal one.

Theorem 3. Let A be a Hermitian matrix of order n and B be an arbitrary matrix of order n with eigenvalues $\alpha_1, \alpha_2, ..., \alpha_n$ and $\beta_1, \beta_2, ..., \beta_n$, respectively. Then, there exists a permutation π of the indices 1, 2, ..., n such that

$$\sum_{i=1}^{n} |\alpha_{i} - \beta_{\pi(i)}|^{2} \le 2 ||A - B||_{F}^{2}.$$
 (3)

Bound (3) was obtained by Kahan (see [4]). Let us now relax the condition on *A*.

Theorem 4. Let A be a diagonalizable matrix of order n and B be an arbitrary matrix of order n with eigenvalues $\alpha_1, \alpha_2, ..., \alpha_n$ and $\beta_1, \beta_2, ..., \beta_n$, respectively. Let X be a nonsingular matrix whose columns are eigenvectors of A. Then, there exists a permutation π of the indices 1, 2, ..., n such that

$$\sum_{i=1}^{n} |\alpha_{i} - \beta_{\pi(i)}|^{2} \le \|X\|_{2}^{2} \|X^{-1}\|_{2}^{2} \|A - B\|_{F}^{2}.$$
(4)

Here, $\operatorname{cond}_2 X = ||X||_2 ||X^{-1}||_2$ is the spectral condition number of X. Inequality (4) can be found in [3].

Our aim in this communication is to present analogs of Theorems 1–4 that concern perturbations of coneigenvalues. Since this concept is only familiar to a narrow circle of persons, we recall the relevant definitions in Section 2. Then, in Section 3 we prove two lemmas absorbing the most difficulties in proving Theorems 5–8. These theorems are stated and justified in Section 4.

2. The eigenvalues of a matrix are scalar invariants of similarity, which is one of the possible actions of the group $GL_n(\mathbb{C})$ on $M_n(\mathbb{C})$ (the space of complex *n*-by-*n* matrices). Here, we are interested in another action called consimilarity. We say that matrices $A, B \in M_n(\mathbb{C})$ are consimilar if there exists a matrix $P \in GL_n(\mathbb{C})$ such that

$$B = PA\overline{P}^{-1}.$$

The bar over the symbol of a matrix means the entrywise conjugation. Unitary congruence is a particular case of consimilarity that corresponds to a unitary matrix *P*.

With every matrix $A \in M_n(\mathbb{C})$, we can associate the set of *n* scalar invariants of consimilarity, which are

Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, 119992 Russia; e-mail: ikramov@cs.msu.su

called the coneigenvalues of *A*. The definition of coneigenvalues as given, say, in [5] relies heavily on special features of the spectra of matrices of the form $B = A\overline{A}$ (or, which is the same, of matrices of the form $C = \overline{A}A$). If λ is a complex eigenvalue of *B*, then $\overline{\lambda}$ is an eigenvalue as well and both have the same multiplicity. The negative eigenvalues of *B* (if any) are necessarily of even algebraic multiplicity.

Let $\sigma(B) = {\lambda_1, \lambda_2, ..., \lambda_n}$ be the spectrum of *B*.

Definition. The coneigenvalues of $A \in M_n(\mathbb{C})$ are *n* scalars $\mu_1, \mu_2, ..., \mu_n$ obtained as follows:

(1) If $\lambda_i \in \sigma(B)$ does not lie on the negative real semiaxis, then the corresponding coneigenvalue μ_i is defined as the square root of λ_i with a nonnegative real part:

$$\mu_i = \lambda_i^{1/2}, \quad \operatorname{Re}\mu_i \ge 0.$$

The multiplicity of μ_i is set equal to that of λ_i .

(2) With a real negative $\lambda_i \in \sigma(B)$, we associate two conjugate purely imaginary coneigenvalues

$$\mu_i = \pm \lambda_i^{1/2}.$$

The multiplicity of each coneigenvalue is set equal to half the multiplicity of λ_i .

Definition. A matrix $A \in M_n(\mathbb{C})$ is said to be conjugate-normal if

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

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

3. With a matrix $A \in M_n(\mathbb{C})$, we associate the matrix

$$\hat{A} = \left(\begin{array}{c} 0 & A \\ \overline{A} & 0 \end{array}\right)$$

of double order.

Lemma 1. If $\mu_1, \mu_2, ..., \mu_n$ are the coneigenvalues of $A \in M_n(\mathbb{C})$, then the scalars

$$\mu_1, \mu_2, \ldots, \mu_n, -\mu_1, -\mu_2, \ldots, -\mu_n$$

constitute the spectrum of \hat{A} .

This lemma is proved in [5].

Let $\alpha = \{\alpha_1, \alpha_2, ..., \alpha_m\}$ and $\beta = \{\beta_1, \beta_2, ..., \beta_m\}$ be two point sets in the complex plane. With each permutation

$$\pi = \left(\begin{array}{ccc} 1 & 2 & \dots & m \\ \pi(1) & \pi(2) & \dots & \pi(m) \end{array}\right) \in S_m$$

we associate a nonnegative scalar ρ_{π} defined by the formula

 $\rho_{\pi}^{2} = |\alpha_{1} - \beta_{\pi(1)}|^{2} + |\alpha_{2} - \beta_{\pi(2)}|^{2} + \dots + |\alpha_{m} - \beta_{\pi(m)}|^{2}.$

DOKLADY MATHEMATICS Vol. 80 No. 1 2009

The least such scalar ρ_{π} over all the permutations in S_m can be interpreted as the distance between the sets α and β .

Lemma 2. Let m = 2n be an even integer. Assume that both sets $\alpha = {\alpha_1, \alpha_2, ..., \alpha_m}$ and $\beta = {\beta_1, \beta_2, ..., \beta_m}$ are symmetric about the imaginary axis and contain no purely imaginary numbers. Then, for each permutation $\pi \in S_m$, there exists a permutation $\sigma \in S_m$ such that

$$\operatorname{Re}\alpha_{i}\operatorname{Re}\beta_{\sigma(i)} > 0, \quad i = 1, 2, ..., m$$
(5)

and

$$\rho_{\sigma} \le \rho_{\pi}. \tag{6}$$

Proof. The symbol α_i (β_j) stands for the point symmetric to α_i^s (respectively, β_j^s). By assumption, $\alpha_i^s \in \alpha$ and $\beta_j^s \in \beta$ for all *i* and *j*.

Construct a bipartite graph Γ on the vertices α_1 , α_2 , ..., α_m and β_1 , β_2 , ..., β_m in accordance with the permutation π ; thus, its edges are the pairs

$$(\alpha_i, \beta_{\pi(i)}), \quad i = 1, 2, ..., m.$$
 (7)

If no edge intersects the imaginary axis, then there is nothing to prove. In this case, the permutation σ can be identified with π .

Suppose that there are edges in Γ that intersect the imaginary axis. For each edge (7), we supplement Γ with the edge

$$(\alpha_i^s, \beta_{\pi(i)}^s), \quad i = 1, 2, ..., m.$$
 (8)

This doubles the degree of each vertex (which, up to this moment, was one) and can convert Γ into a multigraph $\tilde{\Gamma}$. For instance, the above process doubles each pair of symmetric edges that connect the vertices α_i and $\beta_{\pi(i)}$ belonging to the same half-plane. Another result of the transition from Γ to $\tilde{\Gamma}$ is the doubling of ρ^2 .

Now, we transform $\tilde{\Gamma}$ as follows. For each quadruple α_i , α_i^s , $\beta_{\pi(i)}$, $\beta_{\pi(i)}^s$, we verify whether the edge (α_i , $\beta_{\pi(i)}$) (and, hence, the symmetric edge (α_i^s , $\beta_{\pi(i)}^s$)) intersects the imaginary axis. If an intersection occurs, then the above edges are replaced by (α_i , $\beta_{\pi(i)}^s$) and (α_i^s , $\beta_{\pi(i)}$).

Both vertices of each new edge belong to the same half-plane. Geometrically, this transformation means that the diagonals of the quadrangle with the vertices α_i ,

 α_i^s , $\beta_{\pi(i)}$, $\beta_{\pi(i)}^s$ (in the nondegenerate case, this is a rectangle or a right-angular trapezoid) are replaced by its lateral sides.

It is obvious that

$$|\alpha_{i} - \beta_{\pi(i)}|^{2} + |\alpha_{i}^{s} - \beta_{\pi(i)}^{s}|^{2} < |\alpha_{i} - \beta_{\pi(i)}^{s}|^{2} + |\alpha_{i}^{s} - \beta_{\pi(i)}|^{2}.$$

Having performed the above operations for each quadruple of vertices whose incident edges intersect the imaginary axis, we obtain a new multigraph $\hat{\Gamma}$ with the following properties:

1. As before, the degree of each vertex is two.

2. No edge in $\hat{\Gamma}$ intersects the imaginary axis.

3. $\hat{\Gamma}$ is symmetric about the imaginary axis.

4. The value of ρ^2 is less for $\hat{\Gamma}$ compared to $\tilde{\Gamma}$.

Now, we transform $\hat{\Gamma}$ into a bipartite graph $\overline{\Gamma}$ by deleting half of its edges according to the following rule:

Let $\hat{\Gamma}_k$ be a connected subgraph of $\hat{\Gamma}$. The properties of the latter imply that $\hat{\Gamma}$ contains a connected subgraph $\hat{\Gamma}_k^s$ symmetric to $\hat{\Gamma}_k$. Both these subgraphs do not intersect the imaginary axis.

 $\hat{\Gamma}_k$ has an even number of vertices, and the degree of each vertex is two. Geometrically, $\hat{\Gamma}_k$ is a polygon Mwith an even number of vertices. Each edge of this polygon connects a vertex belonging to α with a vertex belonging to β . Consequently, when we traverse M in either of the two possible directions, we go through the vertices belonging alternatively to α and β . The same is

true of the subgraph $\hat{\Gamma}_k^s$.

Moving around $\hat{\Gamma}_k$ in the chosen direction, we delete each second edge. Thus, exactly half of the edges are preserved. This results in a bipartite graph $\overline{\Gamma}_k$.

In the subgraph $\hat{\Gamma}_k^s$, we delete the edges that are symmetric to those in $\overline{\Gamma}_k$. The resulting bipartite graph is denoted by $\overline{\Gamma}_k^s$. Note that this graph is not symmetric to $\overline{\Gamma}_k$.

It is easy to see that the total contribution to ρ^2 of all the edges deleted from $\hat{\Gamma}_k$ and $\hat{\Gamma}_k^s$ is equal to the total contribution of the remaining edges, that is, to the contribution made by the edges of $\overline{\Gamma}_k$ and $\overline{\Gamma}_k^s$.

Having performed the above operations for each pair of symmetric connected subgraphs of $\hat{\Gamma}$, we arrive at the required bipartite graph $\overline{\Gamma}$. This graph is the union of all the bipartite graphs of the types $\overline{\Gamma}_k$ and $\overline{\Gamma}_k^s$. Moreover, the transition from $\hat{\Gamma}$ to $\overline{\Gamma}$ halves the value of ρ^2 .

The graph $\overline{\Gamma}$ determines the desired permutation σ . Indeed, no edge of this graph intersects the imaginary axis; moreover,

$$\rho_{\sigma}^{2} = \rho(\bar{\Gamma})^{2} = \frac{1}{2}\rho(\hat{\Gamma})^{2} < \frac{1}{2}\rho(\tilde{\Gamma})^{2} = \rho(\Gamma)^{2} = \rho_{\pi}^{2}.$$
(9)

4. The following proposition is an analog of the Hoffman–Wielandt theorem with the eigenvalues replaced by coneigenvalues.

Theorem 5. Let A and B be conjugate-normal n-by-n matrices with coneigenvalues $\alpha_1, \alpha_2, ..., \alpha_n$ and $\beta_1, \beta_2, ..., \beta_n$, respectively. Then, there exists a permutation τ of the indices 1, 2, ..., n such that

$$\sum_{i=1}^{n} |\alpha_i - \beta_{\tau(i)}|^2 \le 2 ||A - B||_F^2.$$
(10)

Proof. Suppose that neither *A* nor *B* has purely imaginary coneigenvalues. With *A* and *B*, we associate the matrices

$$\hat{A} = \begin{pmatrix} 0 & A \\ \overline{A} & 0 \end{pmatrix}$$
 and $\hat{B} = \begin{pmatrix} 0 & B \\ \overline{B} & 0 \end{pmatrix}$. (11)

For conjugate-normal A and B, these matrices are normal in the conventional sense. By Lemma 1, matrices (11) have the spectra

$$\{\gamma_{1}, \gamma_{2}, ..., \gamma_{m}\}\$$

$$= \{\alpha_{1}, \alpha_{2}, ..., \alpha_{n}, -\alpha_{1}, -\alpha_{2}, ..., -\alpha_{n}\},\$$

$$\{\delta_{1}, \delta_{2}, ..., \delta_{m}\}\$$

$$= \{\beta_{1}, \beta_{2}, ..., \beta_{n}, -\beta_{1}, -\beta_{2}, ..., -\beta_{n}\},\$$

respectively, where m = 2n. Using the Hoffman–Wielandt theorem, we find a permutation π such that

$$\sum_{i=1}^{m} |\gamma_{i} - \delta_{\pi(i)}|^{2} \le \|\hat{A} - \hat{B}\|_{F}^{2}$$

If σ is the permutation described in Lemma 2, then we have

$$\sum_{i=1}^{m} |\gamma_i - \delta_{\sigma(i)}|^2 \le \|\hat{A} - \hat{B}\|_F^2 = 2\|A - B\|_F^2.$$
(12)

The scalars γ_i and δ_j belonging to the right half-plane are the coneigenvalues of *A* and *B*, respectively. Distinguishing them among the items of the sum in inequality (12), we obtain the desired permutation τ .

Suppose that either *A* or *B* or both have some purely imaginary coneigenvalues. Then, for all sufficiently small $\varepsilon > 0$, they can be replaced by perturbed matrices A_{ε} and B_{ε} with the following properties:

(i) A_{ε} and B_{ε} remain conjugate-normal;

(ii)
$$||A - A_{\varepsilon}||_F^2 = \varepsilon^2$$
 and $||B - B_{\varepsilon}||_F^2 = \varepsilon^2$

(iii) neither A_{ε} nor B_{ε} has purely imaginary coneigenvalues.

DOKLADY MATHEMATICS Vol. 80 No. 1 2009

The existence of the required matrices A_{ε} and B_{ε} follows from the description of the canonical forms of conjugate-normal matrices with respect to unitary congruences as given, for instance, in [5].

Although inequality (10) holds for A_{ε} and B_{ε} , the corresponding permutation τ depends on ε . Since the number of permutations of *n* symbols is finite, we can extract a stationary subsequence from the sequence τ_{ε} as $\varepsilon \rightarrow +0$. This subsequence determines the desired permutation $\tau = \tau(0)$. The theorem is proved.

Now, we consider an analog of Theorem 2.

Theorem 6. Let A be a conjugate-normal matrix of order n and B be an arbitrary matrix of order n with coneigenvalues $\alpha_1, \alpha_2, ..., \alpha_n$ and $\beta_1, \beta_2, ..., \beta_n$, respectively. Then, there exists a permutation τ of the indices 1, 2, ..., n such that

$$\sum_{i=1}^{n} |\alpha_{i} - \beta_{\tau(i)}|^{2} \le 4n ||A - B||_{F}^{2}.$$
 (13)

The proof goes along the same lines as in the preceding theorem. The distinctions are as follows:

1. As before, we associate with A and B the matrices \hat{A} and \hat{B} . \hat{A} remains a normal matrix; however, \hat{B} should now be regarded as an arbitrary matrix.

2. Instead of the Hoffman-Wielandt bound, we use Sun's bound (2) in inequality (12) bearing in mind that

 \hat{A} and \hat{B} have the order m = 2n.

3. The fact that a matrix *B* having some purely imaginary coneigenvalues can be replaced by an arbitrarily close matrix B_{ε} that has no purely imaginary coneigenvalues is now substantiated by the theorem on the canonical form of complex matrices with respect to consimilarities. This theorem, which is an analog of the classical theorem on the Jordan form, can be found, for instance, in [6].

Now, we turn to an analog of the Kahan theorem (see Theorem 3).

Theorem 7. Let A be a symmetric matrix of order n and B be an arbitrary matrix of order n with coneigenvalues $\alpha_1, \alpha_2, ..., \alpha_n$ and $\beta_1, \beta_2, ..., \beta_m$ respectively. Then, there exists a permutation τ of the indices 1, 2, ..., n such that

$$\sum_{i=1}^{n} |\alpha_i - \beta_{\tau(i)}|^2 \le 4 ||A - B||_F^2.$$
(14)

The proof uses the same scheme as in two preceding theorems. A new circumstance is that, for a symmetric

A, the matrix \hat{A} is Hermitian. Therefore, we use Kahan's bound (3) in inequality (12). The only purely imaginary coneigenvalue that the symmetric matrix Acan have is zero. We can get rid of this zero coneigenvalue by shifting the principal diagonal at an arbitrarily small ε .

DOKLADY MATHEMATICS Vol. 80 No. 1 2009

We precede our last theorem by a definition and another lemma.

Definition. A matrix $A \in M_n(\mathbb{C})$ is said to be condiagonalizable if there exists a nonsingular matrix *X* such that

$$D = X^{-1}A\overline{X} \tag{15}$$

is a diagonal matrix.

This definition of a condiagonalizable matrix is taken from [7, Section 4.6].

Lemma 3. Let $A \in M_n(\mathbb{C})$ be a condiagonalizable matrix, and let a nonsingular matrix X bring A to a diagonal matrix D by the consimilarity transformation (15).

Then, the matrix \hat{A} associated with A is diagonalizable

in the conventional sense, and the matrix Y bringing A to diagonal form by similarity can be chosen so that

$$\operatorname{cond}_2 Y = \operatorname{cond}_2 X. \tag{16}$$

Proof. We set

$$Z = X \otimes \overline{X}.$$

It is easy to verify that the similarity transformation

$$\hat{A} \to Z^{-1} \hat{A} Z$$

results in the Hermitian matrix

$$\hat{D} = \left(\begin{array}{c} 0 & D \\ \overline{D} & 0 \end{array}\right).$$

Moreover, we have

$$\operatorname{cond}_2 Z = \operatorname{cond}_2 X.$$

Let Q be a unitary matrix of order 2n that transforms \hat{D} to diagonal form by similarity. Then, the matrix

$$Y = ZQ$$

is composed of the eigenvectors of \hat{A} . Furthermore, we have

$$\operatorname{cond}_2 Y = \operatorname{cond}_2 Z = \operatorname{cond}_2 X.$$

The lemma is proved.

Theorem 8. Let A be a condiagonalizable matrix of order n and B be an arbitrary matrix of order n with coneigenvalues $\alpha_1, \alpha_2, ..., \alpha_n$ and $\beta_1, \beta_2, ..., \beta_n$, respectively. Let X be a nonsingular matrix that brings A to diagonal form by the consimilarity transformation (15). Then, there exists a permutation τ of the indices 1, 2, ..., n such that

$$\sum_{i=1}^{n} |\alpha_{i} - \beta_{\tau(i)}|^{2} \leq 2 \|X\|_{2}^{2} \|X^{-1}\|_{2}^{2} \|A - B\|_{F}^{2}.$$
(17)

The proof is based on the same scheme as in Theorems 5–7. By Lemma 3, \hat{A} is a diagonalizable matrix, and the eigenvector matrix Y can be chosen so as to sat-

isfy condition (16). As for \hat{B} , it should be regarded as a general matrix. In inequality (12), we use the Sun bound from Theorem 4 bearing in mind equality (16). Finally, if A has purely imaginary coneigenvalues (namely, the zero coneigenvalue), it can be replaced by the matrix

$$A_{\varepsilon} = X D_{\varepsilon} \overline{X}^{-1},$$

where D_{ε} is a nonsingular diagonal matrix arbitrarily close to D.

REFERENCES

- 1. A. J. Hoffman and H. W. Wielandt, Duke Math. J. 20, 37–39 (1953).
- 2. J. H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon, Oxford, 1965).
- 3. J.-G. Sun, Linear Algebra Appl. 246, 215-223 (1996).
- 4. W. M. Kahan, Proc. Am. Math. Soc. 48, 11–17 (1975).
- Kh. D. Ikramov, Zap. Nauchn. Semin. POMI 334, 111– 120 (2006).
- Y. P. Hong and R. A. Horn, Linear Algebra Appl. 102, 143–168 (1988).
- 7. R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985).