Journal of Computational Mathematics, Vol.9, No.4, 1991, 360–368.

EXTENSIONS OF THE KANTOROVICH INEQUALITY AND THE BAUER-FIKE INEQUALITY*1)

Sun Ji-guang
(*Computing Center, Academia Sinica, Beijing, China*)

Abstract

This paper proves a Kantorovich-type inequality on the matrix of the type

$$\frac{1}{2}\left(Q_1^H A Q_1 Q_1^H A^{-1} Q_1 + Q_1^H A^{-1} Q_1 Q_1^H A Q_1\right),$$

where A is an $n \times n$ positive definite Hermitian matrix and Q_1 is an $n \times m$ matrix with rank $(Q_1) = m$. The result is applied to get an extension of the Bauer-Fike inequality on condition numbers of similarities that block diagonalized matrices.

Let $A \in I\!R^{n\times n}$ (the set of complex $n \times n$ matrices), and let z_j, w_j be right and left eigenvectors of A corresponding to the eigenvalue λ_j, i.e.,

$$Az_j = \lambda_j z_j, \quad w_j^H A = \lambda_j w_j^H.$$

Define

$$s_j \equiv \cos\theta(z_j, w_j) = \frac{|w_j^H z_j|}{\|z_j\|_2 \|w_j\|_2},$$

where $\theta(z_j, w_j)$ denotes the angle between the one dimensional linear subspaces $\mathcal{R}(z_j)$ and $\mathcal{R}(w_j)$ spanned by z_j and w_j, respectively. Moreover, suppose that $Z, W \in \mathbb{C}^{n\times n}$ satisfy

$$W^H Z = I, \quad W^H AZ = \text{ diag } (\lambda_1, \cdots, \lambda_n), \tag{0.1}$$

and let

$$\kappa_2(A) \equiv \inf \|Z\|_2 \|Z^{-1}\|_2, \tag{0.2}$$

where $\|\cdot\|_2$ denotes the spectral norm, and the infimum is taken with respect to both matrices Z and W satisfying (0.1).

It is well known that if λ_j is a simple eigenvalue of A, then s_j is uniquely determined. Bauer and Fike [1] and Wilkinson [9] proved that the quantities s_j and $\kappa_2(A)$ give some measures of the sensitivity of the eigenvalues to perturbations of the elements of A, so s_j and $\kappa_2(A)$ are called condition numbers of the eigenvalues of A.

The condition numbers s_j and $\kappa_2(A)$ are related by the Bauer-Fike inequality[1]

$$\frac{1}{s_j} \le \frac{1}{2}\left(\kappa_2(A) + \frac{1}{\kappa_2(A)}\right). \tag{0.3}$$

This paper will give an extension of (0.3).

Suppose that $\mathcal{X}_1, \cdots, \mathcal{X}_r$ are linear subspaces of $\mathbb{C}^n$, and

$$\mathbb{C}^n = \mathcal{X}_1 \oplus \cdots \oplus \mathcal{X}_r, \quad \dim(\mathcal{X}_j) = m_j \quad \forall j. \tag{0.4}$$

* Received August 31, 1988.
1) The Project Supported by National Natural Science Foundation of China.

Let[3]

$$\mathcal{Y}_j = \bigcap_{\substack{k=1 \\ k \neq j}}^{r} \mathcal{X}_k^{\perp}, \quad j = 1, \cdots, r, \tag{0.5}$$

where $\mathcal{X}_k^{\perp}$ denotes the orthogonal complement subspace of $\mathcal{X}_k$ in $\mathbb{C}^n$. Obviously,

$$\dim(\mathcal{Y}_j) = m_j \quad \forall j, \quad \mathcal{Y}_1 \oplus \cdots \oplus \mathcal{Y}_r = \mathbb{C}^n. \tag{0.6}$$

Take matrices X_j, Y_j so that the columns of X_j, Y_j form orthonormal bases of $\mathcal{X}_j, \mathcal{Y}_j$, respectively. Since $(X_1, \cdots, X_r)$ and $(Y_1, \cdots, Y_r)$ are nonsingular $n \times n$ matrices, and

$$(Y_1, \cdots, Y_r)^H (X_1, \cdots, X_r) = \operatorname{diag}(Y_1^H X_1, \cdots, Y_r^H X_r),$$

the matrices $Y_j^H X_j$ are nonsingular. Define

$$\Theta(X_j, Y_j) \equiv \arccos(X_j^H Y_j Y_j^H X_j)^{\frac{1}{2}} > 0 \tag{0.7}$$

and

$$S_j \equiv \left\| [\cos \Theta(X_j, Y_j)]^{-1} \right\|^{-1}, \tag{0.8}$$

where $\|\cdot\|$ is any unitarily invariant norm, and $\Theta > 0 (\geq 0)$ denotes that Θ is a positive definite (semidefinite) Hermitian matrix. Especially, S_j will be written as $S_j^{(2)}$ or $S_j^{(F)}$ if we take the spectral norm $\|\cdot\|_2$ or the Frobenius norm $\|\cdot\|_F$ in (0.8), respectively.

The author [7] has proved that if $\mathcal{X}_j$ is an invariant right subspace of A corresponding to the semisimple eigenvalue λ_j of multiplicity m_j, then the quantity S_j^{-1} gives a measure of the sensitivity of the eigenvalue λ_j to perturbations of the elements of A.

The symbol $\mathcal{R}(\cdot)$ stands for the column space. Let

$$\mathcal{Z} = \{ Z \in \mathbb{C}^{n \times n} : Z = (Z_1, \cdots, Z_r), \quad Z_j \in \mathbb{C}^{n \times m_j}, \quad \mathcal{R}(Z_j) = \mathcal{X}_j \}, \tag{0.9}$$

and let

$$\kappa_2 \equiv \inf_{Z \in \mathcal{Z}} \|Z\|_2 \|Z^{-1}\|_2. \tag{0.10}$$

The Bauer-Fike inequality (0.3) has been extended by the author in the form ([7, Theorem 3.1])

$$\frac{1}{S_j^{(F)}} \leq \frac{\sqrt{m_j}}{2} (\kappa_2 + \frac{1}{\kappa_2}). \tag{0.11}$$

In this paper we shall give the following extension of (0.3).

Theorem 1. *Let $\mathcal{X}_1, \cdots, \mathcal{X}_r$ be linear subspaces of $\mathbb{C}^n$ satisfying (0.4). Let $\mathcal{Y}_1, \cdots, \mathcal{Y}_r$ be defined by (0.5), S_j by (0.8), and κ_2 by (0.10). Then*

$$\frac{1}{S_j^{(2)}} \leq \frac{1}{2} (\kappa_2 + \frac{1}{\kappa_2}), \quad j = 1, \cdots, r. \tag{0.12}$$

We can prove that inequalities (0.12) are equivalent to a result of Demmel [2] (a proof of the equivalence will be given in Appendix). We shall prove inequalities (0.12) by using a Kantorovich-type inequality stated in the following theorem.

Theorem 2. *Let $A \in \mathbb{C}^{n \times n}$ be any positive definite Hermitian matrix with the eigenvalues $\{\omega_j\}$ satisfying*

$$0 < l \leq \omega_n \leq \cdots \leq \omega_1 \leq L. \tag{0.13}$$