Journal of Computational Mathematics, Vol.16, No.2, 1998, 121–128.

ON MATRIX UNITARILY INVARIANT NORM CONDITION NUMBER*

Dao-sheng Zheng
(*Department of Mathematics, East China Normal University, Shanghai 200062, China*)

Abstract

In this paper, the unitarily invariant norm $\|\cdot\|$ on $\mathbb{C}^{m\times n}$ is used. We first discuss the problem under what case, a rectangular matrix A has minimum condition number $K(A) = \|A\| \, \|A^+\|$, where A^+ designates the Moore-Penrose inverse of A; and under what condition, a square matrix A has minimum condition number for its eigenproblem? Then we consider the second problem, i.e., optimum of $K(A) = \|A\| \, \|A^{-1}\|_2$ in error estimation.

Key words: Matrix, unitarily invariant norm, condition number

1. Introduction

Since 1984, several chinese mathematicians have obtained many results bout matrix operator norm condition number[11,12,18].

Two kinds matrix condition numbers [9] are :

(1) If $A \in \mathbb{C}^{n\times n}$ is nonsingular, the number $K_\alpha(A) = \|A\|_\alpha \|A^{-1}\|_\alpha$ is called the α-norm condition number of A for its inverse, where $\|\cdot\|_\alpha$ is some matrix norm, such as the 2-norm, *Hölder*-norm, F-norm, etc..

Furthermore, we can generalize the inverse condition number to rectangular matrix case [1], [8], $K(A) = \|A\|_\alpha \|A^+\|_\beta$, and allows $\alpha \neq \beta$.

(2) For a square matrix $A \in \mathbb{C}^{n\times n}$, set

$$V_A = \{X \mid X \in \mathbb{C}^{n\times n},\ X^{-1}AX = J_A, \text{a Jordan form of } A\}. \tag{1.1}$$

Then the number

$$J_\alpha = \inf_{X\in V_A} \{\|X\|_\alpha \|X^{-1}\|_\alpha\} \tag{1.2}$$

is called the α-norm condition number of A for its eigenproblem.

Wilkinson[9] pointed out that a) If matrix A is normal, then $J_2(A) = 1$. b) If A is unitary, then $K_2(A) = 1$.

Zheng[11,12] obtained the necessary and sufficient conditions for minimizing two kinds of p-norm condition numbers $(1 \le p \le \infty)$.

Zheng and Zhao[8] obtained the structures of p-norm isometric matrix $A \in \mathbb{C}^{m\times n}$ and the bounds of $K_p(A) = \|A\|_p \|A^+\|_p$ $(1 \le p \le \infty)$; Wang and Chen obtained the structures of a rectangular matrix A with minimum p-norm condition number $(1 \le p \le \infty,\ p \neq 2)$.

* Received March 11, 1994.

All the above results are concerned with matrix operator norms.

Other results associated with matrix operator norm condition number are given by Yang[10], i.e., the optimum of $K(A) = \|A\| \, \|A^{-1}\|$ in the error estimation of linear equation $Ax = b$ and the process of computing A^{-1}.

In this paper, another important kind matrix norm, the unitarily invariant norm on $\mathbb{C}^{m\times n}$ (UIN) is discussed, and some results associated condition number are obtained.

The rest of the paper is arranged as follows. Section 2 is preliminary. In Section 3, the structures of the rectangular matrices with minimum UIN condition number $K(A) = \|A\| \, \|A^{+}\|$ are discussed. In Section 4, the condition for a square matrix A possesses minimum UIN condition number for its eigenproblem is obtained. Finally, Section 5 is used to describe some results about the optimum of $K(A) = \|A\| \, \|A^{-1}\|_2$ in error estimation, where $\| \cdot \|$ designates a UIN.

2. Preliminaries

Definition 2.1[6,7]**.** *A norm $\|\cdot\|$: $\mathbb{C}^{n\times n} \to \mathbb{R}$ is called unitarily invariant* (*UIN*) *if it satisfies* :

(1) $\|UAV\| = \|A\|$, $\forall A, U, V \in \mathbb{C}^{n\times n}$, *and* $U^H U = V^H V = I_n$.

(2) $\|A\| = \|A\|_2$ *if* $rank(A) = 1$.

Definition 2.2[6,7]**.** *A norm $\Phi : \mathbb{R}^n \to \mathbb{R}$ is called a symmetric guage function* (*SG*) *if it satisfies* :

(1) *For any permutation matrix P,* $\Phi(Px) = \Phi(x)$, $\forall x \in \mathbb{R}^n$.

(2) $\Phi(|x|) = \Phi(x)$, *where* $x = (\xi_1, \cdots, \xi_n)^T$, *and* $|x| = (|\xi_1|, \cdots, |\xi_n|)^T$.

(3) $\Phi(e_1) = 1$, *where e_1 is the first column of* I_n.

The conception of unitarily invariant norm can be generalized to the rectangular matrix case [6], [7, p. 79], and many properties of the UIN can be found in [6] [7] etc..

Lemma 2.1. *Let $\Phi_p : \mathbb{R}^m \to \mathbb{R}$ be a function defined by*

$$\Phi_p(x) = \|x\|_p = \Big(\sum_{i=1}^{m} |\xi_i|^p \Big)^{1/p}, \ (1 \le p \le \infty). \tag{2.1}$$

Then Φ_p is a SG on $\mathbb{R}^m$.

Proof. It is obvious that Φ is the *Hölder* norm on $\mathbb{R}^m$ [5], and satisfies (1) (2) (3) of Definition 2.2. □

If $A \in \mathbb{C}^{k\times l}$, Φ is a SG on $\mathbb{R}^n$, $m = \min\{k.l\} \le n$, $\sigma_1, \cdots, \sigma_m$ are the singular values of A. Then a UIN on $\mathbb{C}^{k\times l}$ may be defined by [6, p. 79]

$$\|A\|_\Phi = \Phi(\sigma_1, \cdots, \sigma_m, 0 \cdots, 0). \tag{2.2}$$

It is easy to see that[6] $\|A\|_{\Phi_0} = \|A\|_2,$ *and* $\|A\|_{\Phi_2} = \|A\|_F$.

Definition 2.3. *If Φ_p is defined by* (2.1), $\|\cdot\|_\Phi$ *is defined by* (2.2). *Then $\|\cdot\|_{\Phi_p}$ is called a pUIN on* $\mathbb{C}^{k\times l}$.

Lemma 2.2. *Suppose $0 \neq A \in \mathbb{C}^{m\times n}$, $\|\cdot\|$ is a UIN family. Then*

$$K(A) = \|A\| \, \|A^{+}\|_2 \ge 1, \ \textit{and } K(cA) = K(A) \ \textit{when } c \neq 0. \tag{2.3}$$

Proof. From [7] [6, p. 80] we know that $\|A\| \geq \|A\|_2$ and $\|A^+\| \geq \|A^+\|_2$. Lemma 2.6 of [8] tells us that $K(A) \geq K_2(A) \geq 1$. From [1] we obtain $(cA)^+ = \frac{1}{c}A^+$, when $c \neq 0$. Thus $K(cA) = K(A)$, when $c \neq 0$. □

Definition 2.4. *A matrix $A \in \mathbb{C}^{m\times n}$ is called 2-norm isometric if it satisfies*

$$\|Ax\|_2 = \|x\|_2,\ \forall x \in \mathbb{C}^n. \tag{2.4}$$

Lemma 2.3[5,11]**.** *A matrix $A \in \mathbb{C}^{m\times n}$ is 2-norm isometric if and only if*

$$A^H A = I_n. \tag{2.5}$$

Lemma 2.4. *For a UIN family, set*

$$L(m,n,r) = \inf_{A\in\mathbb{C}_r^{m\times n}} \{K(A) = \|A\|\ \|A^+\|\}. \tag{2.6}$$

If $r > 0$, then

$$1 \leq L(m,n,r) \leq r^2, \tag{2.7}$$

where $A \in \mathbb{C}_r^{m\times n}$ means m-by-n matrix A has $rank(A) = r$.

Proof. From Lemma 2.2, $L(m,n,r) \geq 1$ when $r > 0$. Take a particular matrix $A_0 = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{C}_r^{m\times n}$. Then $A_0^+ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{C}_r^{n\times m}$. Assume Φ is a SG satisfies $\Phi(A) = \|A\|$. Then $K(A_0) = \|A_0\|\ \|A_0^+\| = (\Phi(1,\cdots,1,0,\cdots,0))\Phi(1,\cdots,1,0,\cdots,0)) \leq r^2$. So we have $L(m,n,r) \leq r^2$.□

Lemma 2.5[7,pp.321–322]**.** *Suppose $A + E = B \in \mathbb{C}^{n\times n}$, $\|\cdot\|$ is a UIN on $\mathbb{C}^{n\times n}$, $\|A^{-1}\|_2\|E\|_2 < 1$. Then B is nonsingular and*

$$(\|B^{-1} - A^{-1}\|)/(\|A^{-1}\|) \leq K\|E\|_2/(\gamma\|A\|), \tag{2.8}$$

where

$$K = \|A\|\|A^{-1}\|_2,\ \gamma = 1 - K\|E\|_2/\|A\| = 1 - \|A^{-1}\|_2\|E\|_2 > 0. \tag{2.9}$$

Lemma 2.6. *Suppose $A, B \in \mathbb{C}^{n\times n}$. Then there exists a unitary matrix H such that*

$$\|AHB\|_2 = \|A\|_2\|B\|_2. \tag{2.10}$$

Proof. Assume the SVD of A, B are

$$A = U\Sigma_A V \text{ and } B = W\Sigma_B R \tag{2.11}$$

respectively with $\Sigma_A =$diag $(\sigma_1,\cdots,\sigma_n)$ and $\Sigma_B =$diag $(\tau_1,\cdots,\tau_n)$, here $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$, and $\tau_1 \geq \cdots \geq \tau_n$. Then $\|AHB\|_2 = \|\Sigma_A VHW\Sigma_B\|_2$. Set $H = V^H W^H$, we obtain $\|AHB\|_2 = \|\Sigma_A\Sigma_B\|_2 = \|$ diag $(\sigma_1\tau_1,\cdots,\sigma_n\tau_n)\|_2 = \sigma_1\tau_1 = \|A\|_2\|B\|_2$.□

Lemma 2.7. *Suppose $\|\cdot\|$ is a UIN family, $A \in \mathbb{C}_r^{m\times n}$, $r > 0$. Then*

$$1 \leq \frac{\|A\|}{\|A\|_2} \leq r. \tag{2.12}$$

Proof. Using the corresponding SG of $\|\cdot\|$ we obtain $\|A\| = \Phi(\sigma_1,\cdots,\sigma_r,0\cdots,0) \leq r\Phi(\sigma_1,0,\cdots,0) = r\|A\|_2$.□

Lemma 2.8 [7, p. 323]. *Suppose* $\|\cdot\|$ *is a UIN on* $\mathbb{C}^{n\times n}$, $A\in\mathbb{C}_n^{n\times n}$, *x is a soluttion of equation* $Ax=b$, $B=A+E$, $\|A^{-1}\|_2\|E\|_2<1$. *Then* $B\in\mathbb{C}_n^{n\times n}$ *and the solution y of equation* $By=b$ *satisfies*

$$(\|y-x\|_2)/\|x\|_2\leq K\|E\|_2/(\gamma\|A\|). \tag{2.13}$$

Lemma 2.9 [7, pp. 342-343]. *Suppose* $B=A+E\in\mathbb{C}^{n\times n},\triangle=\|A^{-1}\|_2\|E\|_2<1$. *Then B is nonsingular and*

$$(\|B^{-1}-A^{-1}\|)/\|A^{-1}\|_2\leq\|E\|K/(\|A\|(1-\triangle)), \tag{2.14}$$

where

$$K=\|A^{-1}\|_2\|A\|. \tag{2.15}$$

3. Rectangular Matrix with Minimum pUIN condition number

Theirem 3.1. *Suppose* $\|\cdot\|$ *is a pUIN family. Then (i)* $L(m,n,r)=r^{2/p}$ *when* $r>0$, $1\leq p\leq\infty$. *(ii) When rank* $(A)=r>0$,

$$K(A)=\|A\|\ \|A^{+}\|=r^{2/p}\Leftrightarrow\sigma_1(A)=\cdots=\sigma_r(A)>0. \tag{3.1}$$

Proof. Take

$$A_0=\begin{pmatrix}I_r&0\\0&0\end{pmatrix}\in\mathbb{C}_r^{m\times n}. \tag{3.2}$$

Then $A_0^{+}=\begin{pmatrix}I_r&0\\0&0\end{pmatrix}\in\mathbb{C}_r^{n\times m}$ and

$$K(A_0)=\|A_0\|_{\Phi_p}\|A_0^{+}\|_{\Phi_p}=r^{2/p}\ \textit{when}\ r>0. \tag{3.3}$$

For any $A\in\mathbb{C}_r^{m\times n}$ with its SVD $A=U\begin{pmatrix}\Sigma_r&0\\0&0\end{pmatrix}V^H$. Two possible cases need to be considered.

Case (a) $1\leq p<\infty$. In this case we have

$$\begin{aligned}K^p(A)=\|A\|_{\Phi_p}^p\|A^{+}\|_{\Phi_p}^p&=\left\|\begin{matrix}\Sigma_r&0\\0&0\end{matrix}\right\|_{\Phi_p}^p\left\|\begin{matrix}\Sigma_r^{-1}&0\\0&0\end{matrix}\right\|_{\Phi_p}^p=(\sigma_1^p+\cdots+\sigma_r^p)(\sigma_1^{-p}+\cdots+\sigma_r^{-p})\\&=((\sigma_1^{p/2})^2+\cdots+(\sigma_r^{p/2})^2)((\sigma_1^{-p/2})^2+\cdots+(\sigma_r^{-p/2})^2).\end{aligned}$$

From the Cauchy-Schwartz inequality we see that $K^p(A)\geq r^2$, and equality holds if and only if $\sigma_1=\cdots=\sigma_r$.

Case (b) $p=\infty$. In this case we have $K(A)=\|A\|_{\Phi_\infty}\|A^{+}\|_{\Phi_\infty}=\Phi_\infty(\sigma_1,\cdots,\sigma_r,0\cdots,0)\Phi_\infty(\sigma_1^{-1},\cdots,\sigma_r^{-1},0,\cdots,0)=\sigma_1/\sigma_r\geq1$, and equaliy holds if and only if $\sigma_1=\sigma_r$. Thus Theorem 3.1 is proved.□

From Theorem 3.1, we obtain the following corollaries.

Corollary 3.1. *Suppose* $\|\cdot\|$ *is a pUIN family,* $A\in\mathbb{C}^{m\times n}$. *Then* $K(A)=\|A\|\ \|A^{+}\|=n^{2/p}$ *if and only if*

$$A^HA=cI\ \textit{with a costant}\ c=\|A\|_2^2>0. \tag{3.4}$$

or equivalently

$$K_2(A) = \|A\|_2\|A^+\|_2 = 1. \tag{3.5}$$

Proof. Theorem 3.1 means (3.4) holds if and only if $K(A) = \|A\|\ \|A^+\| = n^{2/p}$. (3.4) means $\frac{1}{\sqrt{c}}A$ is a unitary matrix, and $\left(\frac{1}{\sqrt{c}}A\right)^+ = \frac{1}{\sqrt{c}}A^H$. Thus (3.4) means (3.5) holds.

Conversly, from Theorem 2.2 of [8], $A/\|A\|_2$ is 2-norm isometric when (3.5) holds. Lemma 2.3 tells us (3.4) holds.□

Corollary 3.2. *Suppose* $\|\cdot\|$ *is a* p*UIN family,* $A \in \mathbb{C}_r^{m\times n}$ *and* $0 < r < min\{m, n\}$. *Then*

$$K(A) = \|A\|\ \|A^+\| = r^{2/p} \tag{3.6}$$

if and only if there are two matrices F *and* G *such that*

$$A = FG \tag{3.7}$$

with

$$F \in \mathbb{C}_r^{m\times r} \text{ and } G \in \mathbb{C}_r^{r\times n}, \tag{3.8}$$

and

$$K(F) = \|F\|\ \|F^+\| = r^{2/p},\ K(G) = \|G\|\ \|G^+\| = r^{2/p}. \tag{3.9}$$

Proof. Necessity. Assume the SVD of A is

$$A = U\begin{pmatrix}\Sigma_r & 0\\ 0 & 0\end{pmatrix}V^H = U_1\Sigma_r V_1^H.$$

From Theorem 3.1 we have $\Sigma_r = cI,\ c > 0$. Set $F = cU_1$, and $G = V_1^H$, then (3.7)–(3.9) hold.

Sufficiency. Assume the SVD of F and G are

$$F = W\begin{pmatrix}\tilde{\Sigma}_r\\ 0\end{pmatrix}S^H, G = Q(\hat{\Sigma}_r, 0)Z^H. \tag{3.10}$$

From Theorem 3.1 and (3.9), we have

$$\begin{aligned} A =& FG = W\begin{pmatrix}c_f I_r\\ 0\end{pmatrix}S^H Q(c_g I, 0)Z^H\\ =& W\tilde{S}^H\tilde{Q}\begin{pmatrix}c_f c_g I_r & 0\\ 0 & 0\end{pmatrix}Z^H,\ \tilde{S}^H\tilde{Q} = \begin{pmatrix}S^H Q & 0\\ 0 & I\end{pmatrix} \in \mathbb{C}^{m\times m}. \end{aligned} \tag{3.12}$$

(3.12) means $\sigma_1 = \cdots = \sigma_r = c_f c_g > 0$, and $K(A) = r^{2/p}$.□

4. Square Matrix with Minimum pUIN Condition Number for Its Eigenproblem

Theorem 4.1. *Suppose* $A \in \mathbb{C}^{n\times n}$, $\|\cdot\|$ *is a consistent matrix norm on* $\mathbb{C}^{n\times n}$[5]. *Then there exists a matrix* $\tilde{X} \in V_A$ *such that*

$$K_\alpha(X) = \|\tilde{X}\|_\alpha\|\tilde{X}^{-1}\|_\alpha = J_\alpha(A). \tag{4.1}$$

Notice that if $\|\cdot\|$ is a pUIN, using Theorem 3.1, we can easily prove Theorem 4.1. And a pUIN is a consistent matrix norm.

Proof. For any $\epsilon > 0$, there exists a matrix $X \in V_A$ such that $J_\alpha(A) \le K_\alpha(X) \le J_\alpha(A)+\epsilon$. Without loss of generality, assume $\|X\|_\alpha = 1$. Otherwise take $X' = X/\|X\|_\alpha$, then we have $\|X'\|_\alpha = 1$, $X' \in V_A$. Set $\epsilon_1 > \epsilon_2 > \cdots > \epsilon_k > \cdots$, and $\lim_{k\to\infty} \epsilon_k = 0$. Correspondingly, we obtain a matrix sequence $\{X_k\}$ such that

$$J_\alpha(A) \le K_\alpha(X_k) = \|X_k^{-1}\|_\alpha \le J_\alpha(A) + \epsilon_k. \tag{4.2}$$

Notice that each eigenvalue $\lambda^{(l)}$ of X_k satisfies[5]

$$|\lambda^{(l)}| \ge \frac{1}{\|X_k^{-1}\|_\alpha} \ge \frac{1}{J_\alpha(A)+\epsilon_k} \ge \frac{1}{J_\alpha+\epsilon_1} = \delta > 0, \tag{4.3}$$

and both $\{X_k^{-1}\}$ and $\{X_k\}$ are bounded. So there exist subsequences of $\{X_k^{-1}\}$ and $\{X_k\}$ such that

$$\lim_{k_i\to\infty} X_{k_i}^{-1} = \tilde{X}^{-1} \text{ and } \lim_{k_i\to\infty} X_k = \tilde{X}. \tag{4.4}$$

From (4.2) and (4.4) we obtain $J_\alpha(A) = K_\alpha(\tilde{X}) = \|\tilde{X}\|_\alpha \|\tilde{X}^{-1}\|_\alpha$.□

Theorem 4.2. Suppoe $A \in \mathbb{C}^{n\times n}$, $\|\cdot\|_\alpha$ is a pUIN on $\mathbb{C}^{n\times n}$. Then

$$J_\alpha(A) = n^{2/p} \tag{4.5}$$

if and only if there exists a unitary matrix U such that

$$U^H AU = J_A, \text{ a Jordan form of } A. \tag{4.6}$$

Proof. Necessity. Since each UIN is a consistent matrix norm[6,7], Theorem 4.1 means there exists a matrix $X \in V_A$ such that $K_\alpha(X) = \|X\|_\alpha \|X^{-1}\|_\alpha = J_\alpha(A) = n^{2/p}$. From Theorem 3.1, X has singular values $\sigma_1 = \cdots = \sigma_n > 0$. Set $U = X/\|X\|_2$, we obtain $U \in V_A$ *and* $U^H U = I_n$, and $U^H AU = X^{-1}AX = J_A$. Sufficiency. From Theorem 3.1 we obtain $K_\alpha(U) = \|U\|_\alpha \|U^+\|_\alpha = n^{2/p}$.□

5. Optimum of $K(A) = \|A\|\|A^{-1}\|_2$ in Error Estimation With Respect to UIN

Theorem 5.1. *Suppose $\|\cdot\|$ is a UIN on $\mathbb{C}^{n\times n}$, $A \in \mathbb{C}^{n\times n}$. If there exists a $\epsilon_0 > 0$ such that when $\|A^{-1}\|_2\|E\|_2 < 1$ and $\|E\| < \epsilon_0$, then E satisfies*

$$\frac{\|A^{-1} - (A+E)^{-1}\|}{\|A^{-1}\|} \le \mu\frac{\|E\|_2}{A}\Big/\Big(1 - \mu\frac{\|E\|_2}{\|A\|}\Big), \tag{5.1}$$

where $\mu > 0$ is independent of E. Then we have

$$K(A) = \|A\|\ \|A^{-1}\|_2 \le \frac{\|A^{-1}\|}{\|A^{-1}\|_2}\mu \le n\mu. \tag{5.2}$$

Proof. From Lemma 2.5 or [5], $A+E$ is nonsingular, and we have $\dfrac{\|A^{-1}-(A+E)^{-1}\|}{\|A^{-1}\|}$ $\leq \dfrac{K}{\gamma}\dfrac{\|E\|_2}{\|A\|}$, $K=\|A\|\ \|A^{-1}\|_2,\ \gamma=1-\|A^{-1}\|_2\|E\|_2$. Using Lemma 2.6, we can find a matrix H with $\|H\|_2=1$ such that $\|A^{-1}HA^{-1}\|_2=\|A^{-1}\|_2^2$.

Set $E=\epsilon H$, we obtain $\|A^{-1}HA^{-1}\|_2=\epsilon\|A^{-1}\|_2^2$. When $\|A^{-1}\|_2\|E\|_2<1$, we have[5]

$$(A+E)^{-1}=(I+A^{-1}E)A^{-1}=\sum_{k=0}^{\infty}(-A^{-1}E)^kA^{-1}. \tag{5.3}$$

$$\|A^{-1}-(A+E)^{-1}\|=\Big\|\sum_{k=1}^{\infty}(-A^{-1}E)^kA^{-1}\Big\|\geq\|A^{-1}EA^{-1}\|_2-\|A^{-1}\|_2^3\|E\|_2^2$$
$$\cdot\sum_{k=0}^{\infty}\|(A^{-1}E)^k\|_2=\epsilon\|A^{-1}\|_2^2-\epsilon^2\|A^{-1}\|_2^3\sum_{k=0}^{\infty}\|(A^{-1}E)^k\|_2.$$

$$K(A)=\|A\|\ \|A^{-1}\|_2=\frac{\|A\|\ \|A^{-1}\|_2^2}{\|A^{-1}\|_2}=\frac{\|A\|}{\|E\|_2}\frac{\|A^{-1}EA^{-1}\|_2}{\|A^{-1}\|_2} \tag{5.4}$$

$$\leq\frac{\|A\|}{\|E\|_2}\frac{\|A^{-1}-(A+E)^{-1}\|+\|A^{-1}\|_2^3\|E\|_2^2\sum_{k=0}^{\infty}\|A^{-1}E\|_2^k}{\|A^{-1}\|}\frac{\|A^{-1}\|}{\|A^{-1}\|_2}.$$

Let $\epsilon=\|E\|_2\to 0$, we obtain $K(A)\leq\mu\dfrac{\|A^{-1}\|}{\|A^{-1}\|_2}$. From Lemma 2.7, $\|A^{-1}\|\|A^{-1}\|_2\leq n$.□

Theorem 5.2. *Suppose $\|\cdot\|$ is a UIN on $\mathbb{C}^{n\times n}$, $A\in\mathbb{C}_n^{n\times n}$, $B=A+E$ and $\|A^{-1}\|_2\|E\|_2<1$. If there is a $\epsilon_0>0$ such that when $\|E\|_2<\epsilon_0$, the solutions x,y of equations $Ax=b$, $Bx=b$ satisfy*

$$\frac{\|x-y\|_2}{\|x\|_2}\leq\frac{\frac{\delta\|E\|_2}{\|A\|}}{1-\frac{\delta\|E\|_2}{\|A\|}}, \tag{5.5}$$

where δ is independent of E. Then

$$K(A)=\|A\|\ \|A^{-1}\|_2\leq\delta. \tag{5.6}$$

Proof. From Lemma 2.8, we obtain

$$\frac{\|x-y\|_2}{\|x\|_2}\leq K\frac{\|E\|_2}{\|A\|}/(1-K\frac{\|E\|_2}{\|A\|})=\frac{K}{\gamma}\frac{\|E\|_2}{\|A\|}.$$

So (5.6) means that $K(A)=\|A\|\ \|A^{-1}\|_2=K$ is optimum in error estimate equation (5.5).

From (5.3) we obtain

$$x-y=[A^{-1}-(A+E)^{-1}]b=x-\sum_{k=0}^{\infty}(-A^{-1}E)^kx$$

$$= -\sum_{k=1}^{\infty}(-A^{-1}E)^k x = A^{-1}Ex - (A^{-1}E)^2\sum_{k=0}^{\infty}(-A^{-1}E)^k x. \tag{5.7}$$

Let $B = (x, 0, \cdots, 0) \in \mathbb{C}^{n\times n}$. For any $C \in \mathbb{C}^{n\times n}$ we have $\|A^{-1}CB\| = \|A^{-1}CB\|_2$. From Lemma 2.6, there exists a matrix H such that $\|H\|_2 = 1$ and $\|A^{-1}Hx\|_2 = \|A^{-1}HB\|_2 = \|A^{-1}HB\| = \|A^{-1}\|_2\|B\|_2 = \|A^{-1}\|_2\|x\|_2$. Take $E = \epsilon H$, $\|E\|_2 = \epsilon$. From (5.7) we obtain $\|x-y\|_2 \geq \|A^{-1}Ex\|_2 - \|A^{-1}\|_2^2\|E\|_2^2\|x\|_2\sum_{k=0}^{\infty}\|A^{-1}E\|_2^k$. Hence

$$\begin{aligned} K(A) = \|A^{-1}\|_2\|A\| &= \frac{\|A\|}{\|E\|_2}\frac{\|A^{-1}Ex\|_2}{\|x\|_2} \\ &\leq \frac{\|A\|}{\|E\|_2}\frac{\|x-y\|_2 + \|A^{-1}\|_2^2\|E\|_2^2\|x\|_2\sum_{k=0}^{\infty}\|A^{-1}E\|_2^k}{\|x\|_2} \\ &\leq \delta\Big/\Big(1-\delta\frac{\|E\|_2}{\|A\|}\Big) + \epsilon(\|A^{-1}\|_2^2\sum_{k=0}^{\infty}\|A^{-1}E\|_2^k). \end{aligned} \tag{5.8}$$

Thus we obtain $K(A) \leq \lim_{\epsilon\to 0}\Big(\frac{\delta}{1-\delta\|E\|_2/\|A\|} + \epsilon\Big(\|A^{-1}\|_2^2\sum_{k=0}^{\infty}\|A^{-1}E\|_2^k\Big)\Big) = \delta.\square$

Notice that Lemma 2.9 enables us to prove another theorem analogue to Theorem 5.1.

References

[1] Ben-Israel, T. Greville, Generalized Inverses: Theory and Applications, John Wiley, New York, 1974.

[2] D.H. Chen, Some conclusions on the condition numbers of a matrix, *J. East China Norm, Univ. Natur. Sci.*(in Chinese), 3(1986), 11–18.

[3] G.L. Chen, A further study of two classes of condition numbers, *J. East China Norm, Univ. Natur. Sci.*(in Chinese), 1(1988), 13–15.

[4] J.X. Kuang, G.L. Chen, On the necessary and sufficient condition for minimizing the k-condition number of an operator, *Numer. Math. J. Chinese Univ.*(in Chinese), 4(1987), 376–379.

[5] G.W. Stewart, Introduction to Matrix Computations, Academic Press, New York, 1973.

[6] G.W. Stewart, J.G. Sun, Matrix Perturbation Theory, Academic Press, New York, 1990.

[7] G.J. Sun, Matrix Perturbation Analysis (in Chinese), Academic Press, Beijing, 1987.

[8] C. Wang, D.S. Zheng, G.L. Chen, S.Q. Zhao, Structures of p-isometric matrices and rectangular matrices with minimum p-norm condition number, *Linear Alg. Appl.*, 184(1993), 261–278.

[9] J.H. Wilkinson, The Algebraic Eigevalue Problem, Clarendon, Oxford, England, 1965.

[10] C.M. Yang, Optimum of $K(A) = \|A\|\,\|A^{-1}\|$ in error estimate, *J. East China Norm. Univ. Natur. Sci.*(in Chinese), 3(1986), 35–38.

[11] D.S. Zheng, Necessary and sufficent conditions for the spectral codition numbers of matrix to be equal to one matrix norm, *J. East China Norm. Univ. Natur. Sci.*(in Chinese), 3(1984), 29–32.

[12] D.S. Zheng, Bounds on the matrix norm of a rectangular matrix and the p-condition number of a square matrix, *J. East China Norm. Univ. Natur. Sci.*(in Chinese), 3(1986) 27–34.