



Bounds on Condition Number of Singular Matrix

Zhiping Xiong^a

^a*School of Mathematics and Computational Science, Wuyi University, Jiangmen 529020, P.R.China*

Abstract. For each vector norm $\|x\|_v$, a matrix $A \in C^{m \times n}$ has its operator norm $\|A\|_{\mu v} = \max_{x \neq 0} \frac{\|Ax\|_\mu}{\|x\|_v}$. If A is nonsingular, we can define the condition number of $A \in C^{n \times n}$ as $P(A) = \|A\|_{\mu v} \|A^{-1}\|_{v\mu}$. If A is singular, the condition number of matrix $A \in C^{m \times n}$ may be defined as $P_+(A) = \|A\|_{\mu v} \|A^\dagger\|_{v\mu}$. Let U be the set of the whole self-dual norms. It is shown that for a singular matrix $A \in C^{m \times n}$, there is no finite upper bound of $P_+(A)$, while $\|\cdot\|$ varies on U . On the other hand, it is shown that $\inf_{\|\cdot\| \in U} \|A\|_{\mu v} \|A^\dagger\|_{v\mu} = \frac{\sigma_1(A)}{\sigma_r(A)}$, where $\sigma_1(A)$ and $\sigma_r(A)$ are the largest and smallest nonzero singular values of A , respectively.

1. Introduction

Throughout this paper $C^{m \times n}$ denotes the set of all $m \times n$ matrices over the complex field C and $C_r^{m \times n}$ denotes the set of all $m \times n$ complex matrices with rank r . $O_{m \times n}$ is the $m \times n$ matrix of all zero entries (if no confusion occurs, we will drop the subscript). For a matrix $A \in C^{m \times n}$, A^* and $r(A)$ denote the conjugate transpose and the rank of the matrix A , respectively. Furthermore, let $\|\cdot\|_{\mu v}$ be a operator norm on $C^{m \times n}$, $\|\cdot\|_{v\mu}$ be a operator norm on $C^{n \times m}$, $\|\cdot\|_\mu$ be a vector norm on C^m and $\|\cdot\|_v$ be a vector norm on C^n .

Let $A \in C^{m \times n}$, then the unique matrix $X \in C^{n \times m}$ satisfying the following four Penrose equations [8]:

$$AXA = A, \quad XAX = X, \quad (AX)^* = AX, \quad (XA)^* = XA,$$

is called the Moore-Penrose inverse of A and is denoted by A^\dagger .

For any $A \in C^{n \times n}$ and $k = \text{Ind}(A) = \min\{p : r(A^{p+1}) = r(A^p)\}$, there exists a matrix $X \in C^{n \times n}$ satisfying [4]:

$$XAX = X, \quad XA = AX, \quad A^{k+1}X = A^k,$$

2010 *Mathematics Subject Classification.* Primary 65F15; Secondary 65F10, 65C40

Keywords. Moore-Penrose inverse; Condition number; Singular value decomposition; Self-dual norm

Received: 08 September 2013; Accepted: 19 February 2014

Communicated by Dragana Cvetkovic Ilic

Research supported by the NSFC (Grant No: 11301397) and the Foundation for Distinguished Young Talents in Higher Education of Guangdong, China (Grant No: 2012LYM-0126) and the Basic Theory and Scientific Research of Science and Technology Project of Jiangmen City, China, 2013

Email address: xzpwhere@163.com (Zhiping Xiong)

then X is called the Drazin inverse of A and denoted by A^D . The reader can refer to [1, 4, 8, 10] for basic results on these generalized inverses.

Let $\|x\|_v$ be a vector norm defined on the linear space C^n . Then for a matrix $A \in C^{m \times n}$, we define its operator norm as [9]:

$$\|A\|_{\mu v} = \max_{x \neq 0} \frac{\|Ax\|_{\mu}}{\|x\|_v}, \tag{1}$$

where $x \in C^n$. If $A \in C^{n \times n}$ is nonsingular, we can define the condition number of A as:

$$P(A) = \|A\|_{vv} \|A^{-1}\|_{vv}. \tag{2}$$

Obviously, $\|A\|_{vv} \geq \rho(A)$, $P(A) \geq \rho(A)\rho(A^{-1})$, where $\rho(A)$ denotes the spectral radius of A .

Condition number is a basic concept in numerical algebra and is important in some other fields of numerical analysis, see [3, 5, 7, 9]. The normwise relative condition number measures the sensitivity of matrix inversion and the solution of linear systems. It has attracted considerable attentions and many interesting results have been obtained, see [2, 3, 6, 11, 12].

Let V be the set of the whole norms defined on C^n . In 1984, Huang [6] has shown that for a nonsingular matrix $A \in C^{n \times n}$, there is no finite upper bound of $P(A)$ while $\|\cdot\|_v$ varies on V and there is a further relation between $P(A)$ and $\rho(A)\rho(A^{-1})$:

$$\inf_{\|\cdot\|_v \in V} \|A\|_{vv} \|A^{-1}\|_{vv} = \rho(A)\rho(A^{-1}). \tag{3}$$

Let $A \in C^{n \times n}$ be singular and have Drazin inverse. In 2005, Cui [11] defined another condition number of A as:

$$P_D(A) = \|A\|_{vv} \|A^D\|_{vv}, \tag{4}$$

and shown that

$$\inf_{\|\cdot\|_v \in V} \|A\|_{vv} \|A^D\|_{vv} = \rho(A)\rho(A^D). \tag{5}$$

The Moore-Penrose inverse plays an important role on the theoretical research and numerical computations in the areas of optimization, statistics, ill-posed problem and matrix analysis, see [1, 10]. Actually, for a singular matrix $A \in C^{m \times n}$, $A^{\dagger} \in C^{n \times m}$ is existence and unique. Obviously, when $A \in C^{n \times n}$ is nonsingular, $A^{\dagger} = A^{-1}$. Then for a matrix $A \in C^{m \times n}$, we can define a new condition number:

$$P_{\dagger}(A) = \|A\|_{\mu v} \|A^{\dagger}\|_{v \mu}. \tag{6}$$

Definition 1.1 Let $y \in C^n$ and $\|\cdot\|_v$ is a norm defined on C^n . Then we call $\|\cdot\|_v$ a self-dual norm on C^n , if

$$\|y\|_v = \max_{\substack{x \in C^n \\ \|x\|_v = 1}} |y^* x|.$$

Let $U = \{\|\cdot\|_{\mu\nu}, \mu, \nu\}$ be the set of the whole self-dual norms on $C^{m \times n}$, where $\|\cdot\|_{\mu}$ is a self-dual norm on C^m and $\|\cdot\|_{\nu}$ is a self-dual norm on C^n , the next lemma shows that U is not an empty set. In fact, some well-known operator norms such as $\|\cdot\|_2, \|\cdot\|_{\infty}$ and $\|\cdot\|_F$ are self-dual norm.

Lemma 1.1 Let $y = (y_1, y_2, \dots, y_m)^* \in C^m$ and $\|y\|_2 = \sqrt{\sum_{i=1}^m |y_i|^2}$. Then $\|\cdot\|_2$ is a self-dual norm on C^m .

Proof. According to the Definition 1.1, we only need to show

$$\|y\|_2 = \max_{\substack{x \in C^m \\ \|x\|_2=1}} |y^*x|. \tag{7}$$

Let $x = (x_1, x_2, \dots, x_m)^* \in C^m$ and $\|x\|_2 = \sqrt{\sum_{i=1}^m |x_i|^2} = 1$. Then

$$\begin{aligned} |y^*x| &= |(y_1^*, y_2^*, \dots, y_m^*) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}| = |y_1^*x_1 + y_2^*x_2 + \dots + y_m^*x_m| \\ &\leq |y_1^*x_1| + |y_2^*x_2| + \dots + |y_m^*x_m| \leq |y_1^*||x_1| + |y_2^*||x_2| + \dots + |y_m^*||x_m| \\ &\leq \sqrt{\sum_{i=1}^m |y_i|^2} \sqrt{\sum_{i=1}^m |x_i|^2}. \end{aligned} \tag{8}$$

Since x varies on C^m , it follows that

$$\max_{\substack{x \in C^m \\ \|x\|_2=1}} |y^*x| \leq \|y\|_2. \tag{9}$$

On the other hand, taking $x' = \frac{y}{\|y\|_2}$, we have $x' \in C^m$ and $\|x'\|_2 = 1$. Then

$$|y^*x| = \left| \frac{y^*y}{\|y\|_2} \right| = \frac{|y^*y|}{\|y\|_2} = \|y\|_2. \tag{10}$$

Combining (7), (8), (9) with (10), we proved lemma 1.1. \square

Let $U = \{\|\cdot\|_{\mu\nu}, \mu, \nu\}$ be the set of the whole self-dual norms and $A \in C^{m \times n}$. In this article we will show that there is no finite upper bound of $P_+(A)$ while $\|\cdot\|_{\mu}, \|\cdot\|_{\nu}$ vary on U and

$$\inf_{\|\cdot\|_{\mu} \in U, \|\cdot\|_{\nu} \in U} \|A\|_{\mu\nu} \|A^{\dagger}\|_{\nu\mu} = \frac{\sigma_1(A)}{\sigma_r(A)}, \tag{11}$$

where $\sigma_1(A)$ and $\sigma_r(A)$ are the largest and smallest nonzero singular values of A , respectively.

In order to get the main result of this paper, we need the following lemma, which will be used in this paper.

Lemma 1.2 [1, 10] Let $A \in C_r^{m \times n}$, then there exist unitary matrices $P \in C^{m \times m}$ and $Q \in C^{n \times n}$ such that

$$A = P \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q^*, \tag{12}$$

where $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, $\sigma_i = \sqrt{\lambda_i}$, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$ are the nonzero eigenvalues of A^*A , and $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ are the nonzero singular values of A . The Moore-Penrose inverse of A is

$$A^{\dagger} = Q \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} P^*, \tag{13}$$

where $\Sigma^{-1} = \text{diag}(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_r})$.

According to the formula (17), we get the conclusion that when $|s| \rightarrow 0$, $\|A\|_{\mu(Q_s)\nu(Q_s)}$ has no finite upper bound.

On the other hand, for any norm, the following statement holds [9],

$$\frac{1}{\sigma_r^2} = \rho((A^\dagger)^* A^\dagger) \leq \|(A^\dagger)^* A^\dagger\|_{\mu\mu} \leq \|(A^\dagger)^*\|_{\mu\nu} \|A^\dagger\|_{\nu\mu}, \tag{18}$$

where σ_r is the smallest nonzero singular value of A .

From the formula (1), we have

$$\|A^\dagger\|_{\nu\mu} = \max_{\substack{x \in C^m \\ \|x\|_\mu=1}} \|A^\dagger x\|_\nu, \tag{19}$$

where $\|\cdot\|_\mu$ on C^m and $\|\cdot\|_\nu$ on C^n are the self-dual norms of the set U .

According to the formula (1) and Definition 1.1, we have

$$\|A^\dagger\|_{\nu\mu} = \max_{\substack{x \in C^m \\ \|x\|_\mu=1}} \|A^\dagger x\|_\nu = \max_{\substack{x \in C^m \\ \|x\|_\mu=1}} \max_{\substack{y \in C^n \\ \|y\|_\nu=1}} |x^* (A^\dagger)^* y|, \tag{20}$$

and

$$\|(A^\dagger)^*\|_{\mu\nu} = \max_{\substack{y \in C^n \\ \|y\|_\nu=1}} \|(A^\dagger)^* y\|_\mu = \max_{\substack{y \in C^n \\ \|y\|_\nu=1}} \max_{\substack{x \in C^m \\ \|x\|_\mu=1}} |y^* A^\dagger x| = \|A^\dagger\|_{\nu\mu}. \tag{21}$$

Combining (18), (19), (20) with (21), we have

$$\frac{1}{\sigma_r} \leq \|A^\dagger\|_{\nu\mu}. \tag{22}$$

By (17) and (22), we have the conclusion that

$$P_+(A) = \|A\|_{\mu\nu} \|A^\dagger\|_{\nu\mu}$$

has no finite upper bound, while $\|\cdot\|$ varies on the self-dual norms set U . \square

Theorem 2.2 Let $A = [a_{ij}] \in C_r^{m \times n}$ be a nonzero matrix. Then the condition number of A : $P_+(A) = \|A\|_{\mu\nu} \|A^\dagger\|_{\nu\mu} \geq \frac{\sigma_1(A)}{\sigma_r(A)}$, while $\|\cdot\|$ varies on U .

Proof. From [9], we know that σ_1^2 is the spectral radius of A^*A (i.e. $\sigma_1^2 = \rho(A^*A)$). Thus

$$\sigma_1^2 = \rho(A^*A) \leq \|A^*A\|_{\nu\nu} \leq \|A^*\|_{\nu\mu} \|A\|_{\mu\nu}. \tag{23}$$

By (1), we have

$$\|A\|_{\mu\nu} = \max_{\|x\|_\nu=1} \|Ax\|_\mu.$$

Since $\|\cdot\|_\mu$ and $\|\cdot\|_\nu$ are self-dual norms, then according to Definition 1.1, we have

$$\|A\|_{\mu\nu} = \max_{\substack{x \in C^n \\ \|x\|_\nu=1}} \|Ax\|_\mu = \max_{\substack{x \in C^n \\ \|x\|_\nu=1}} \max_{\substack{y \in C^m \\ \|y\|_\mu=1}} |x^* A^* y|. \tag{24}$$

and

$$\|A^*\|_{\nu\mu} = \max_{\substack{y \in C^m \\ \|y\|_\mu=1}} \|A^* y\|_\nu = \max_{\substack{y \in C^m \\ \|y\|_\mu=1}} \max_{\substack{x \in C^n \\ \|x\|_\nu=1}} |y^* Ax| = \|A\|_{\mu\nu}. \tag{25}$$

Combining (23), (24) with (25) yields $\sigma_1^2 \leq \|A\|_{\mu\nu}^2$, that is

$$\sigma_1 \leq \|A\|_{\mu\nu}. \tag{26}$$

By analogy with above proof, we have

$$\frac{1}{\sigma_r} \leq \|A^\dagger\|_{\nu\mu}. \tag{27}$$

From the formulas (26) and (27), we obtain the conclusion that

$$P_+(A) = \|A\|_{\mu\nu}\|A^\dagger\|_{\nu\mu} \geq \frac{\sigma_1(A)}{\sigma_r(A)},$$

holds while $\|\cdot\|$ varies on U . \square

In the above part, we shown that for a singular matrix A there is no finite upper bound of $P_+(A)$, while $\|\cdot\|$ varies on U . In the following theorem we will show that

$$\inf_{\|\cdot\| \in U} P_+(A) = \frac{\sigma_1(A)}{\sigma_r(A)}. \tag{28}$$

Theorem 2.3 Let $A = [a_{ij}] \in C_r^{m \times n}$ be a nonzero matrix. Then

$$\inf_{\|\cdot\| \in U} P_+(A) = \inf_{\|\cdot\|_\mu \in U, \|\cdot\|_\nu \in U} \|A\|_{\mu\nu}\|A^\dagger\|_{\nu\mu} = \frac{\sigma_1(A)}{\sigma_r(A)}. \tag{29}$$

Proof. Let $P \in C^{m \times m}$, $Q \in C^{n \times n}$ be two unitary matrices, such that

$$P^*AQ = \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix}, \tag{30}$$

where $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$ and $\sigma_1 \geq \dots \geq \sigma_r > 0$. Then from Lemma 1.2, we have

$$A^\dagger = Q \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} P^*,$$

where $\Sigma^{-1} = \text{diag}(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_r})$.

Let D_ε and D'_ε be two diagonal matrices as follow:

$$D_\varepsilon = \begin{pmatrix} 1 & & & & \\ & \varepsilon & & & \\ & & \ddots & & \\ & & & \varepsilon^{n-1} & \\ & & & & \varepsilon^{n-1} \end{pmatrix}, \quad D'_\varepsilon = \begin{pmatrix} 1 & \varepsilon & & & \\ & \varepsilon & \varepsilon^2 & & \\ & & \ddots & \ddots & \\ & & & \varepsilon^{m-2} & \varepsilon^{m-1} \\ & & & & \varepsilon^{m-1} \end{pmatrix},$$

where ε is a positive real number.

Suppose $x \in C^n$ and $y \in C^m$, we define

$$\|x\|_{\nu(D_\varepsilon)} = \|D_\varepsilon^{-1}Q^*x\|_\infty \quad \text{and} \quad \|y\|_{\mu(D'_\varepsilon)} = \|D'_\varepsilon^{-1}P^*y\|_\infty.$$

Corresponding, we have

$$\|A\|_{\mu(D'_\varepsilon)\nu(D_\varepsilon)} = \max_{x \neq 0} \frac{\|Ax\|_{\mu(D'_\varepsilon)}}{\|x\|_{\nu(D_\varepsilon)}} = \max_{x \neq 0} \frac{\|D'_\varepsilon^{-1}P^*Ax\|_\infty}{\|D_\varepsilon^{-1}Q^*x\|_\infty} = \max_{z=D_\varepsilon^{-1}Q^*x \neq 0} \frac{\|D'_\varepsilon^{-1}P^*AQD_\varepsilon z\|_\infty}{\|z\|_\infty} = \|D'_\varepsilon^{-1}P^*AQD_\varepsilon\|_\infty \tag{31}$$

and

$$\|A^\dagger\|_{v(D_\varepsilon)\mu(D'_\varepsilon)} = \|D_\varepsilon^{-1}Q^*A^\dagger PD'_\varepsilon\|_\infty. \tag{32}$$

According to the above proof, we obtain

$$D'_\varepsilon^{-1}P^*AQD_\varepsilon = \begin{pmatrix} 1 & \varepsilon & & & \\ \varepsilon & \varepsilon^2 & & & \\ & & \ddots & & \\ & & & \varepsilon^{m-2} & \varepsilon^{m-1} \\ & & & & \varepsilon^{m-1} \end{pmatrix}^{-1} \begin{pmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} 1 & & & & \\ \varepsilon & & & & \\ & \ddots & & & \\ & & \varepsilon^{n-1} & & \end{pmatrix} = \begin{pmatrix} H & O \\ O & O \end{pmatrix}, \tag{33}$$

where $H = \begin{pmatrix} \sigma_1 & -\varepsilon\sigma_2 & \cdots & \cdots & O(\varepsilon) \\ & \sigma_2 & -\varepsilon\sigma_3 & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \sigma_{r-1} & -\varepsilon\sigma_r \\ & & & & \sigma_r \end{pmatrix}$, and

$$D_\varepsilon^{-1}Q^*A^\dagger PD'_\varepsilon = \begin{pmatrix} 1 & & & & \\ & \varepsilon^{-1} & & & \\ & & \ddots & & \\ & & & \varepsilon^{1-n} & \end{pmatrix} \begin{pmatrix} \frac{1}{\sigma_1} & & & & \\ & \ddots & & & \\ & & \frac{1}{\sigma_r} & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} 1 & \varepsilon & & & \\ \varepsilon & \varepsilon^2 & & & \\ & & \ddots & & \\ & & & \varepsilon^{m-2} & \varepsilon^{m-1} \\ & & & & \varepsilon^{m-1} \end{pmatrix} = \begin{pmatrix} T & O \\ O & O \end{pmatrix}, \tag{34}$$

where $T = \begin{pmatrix} \frac{1}{\sigma_1} & \frac{\varepsilon}{\sigma_1} & & & \\ & \frac{1}{\sigma_2} & \frac{\varepsilon}{\sigma_2} & & \\ & & \ddots & \ddots & \\ & & & \frac{1}{\sigma_r} & \frac{\varepsilon}{\sigma_r} \end{pmatrix}$.

Combining the formulas (31), (32), (33) with (34), we have

$$\lim_{\varepsilon \rightarrow 0} \|A\|_{\mu(D'_\varepsilon)v(D_\varepsilon)} \|A^\dagger\|_{v(D_\varepsilon)\mu(D'_\varepsilon)} = \lim_{\varepsilon \rightarrow 0} (\sigma_1 + O(\varepsilon)) \left(\frac{1}{\sigma_r} + \varepsilon \frac{1}{\sigma_r} \right) = \frac{\sigma_1}{\sigma_r}. \tag{35}$$

That is, there exists some self-dual norms such that

$$\inf_{\|\cdot\| \in \mathcal{U}} P_+(A) = \inf_{\|\cdot\| \in \mathcal{U}} \|A\|_{\mu\nu} \|A^\dagger\|_{v\mu} = \frac{\sigma_1(A)}{\sigma_r(A)}. \quad \square \tag{36}$$

Acknowledgements

The author would like to thank Professor Dragana Cvetkovic Ilic and the anonymous referees for their very detailed comments and constructive suggestions, which greatly improved the presentation of this paper.

References

- [1] A. Ben-Israel, T. N. E. Greville, *Generalized Inverses: Theory and Applications*, Wiley-Interscience, 1974; 2nd Edition, Springer-Verlag, New York, 2002.
- [2] X. K. Cui, Bounds on condition number of a singular matrix and its applications, *Appl. Math. Comput.*, 162 (2005) 81-93.
- [3] H. A. Diao, Y. M. Wei, On forbenius normwise condition numbers for Moore-Penrose inverse and linear least squares problems, *Numer. Linear Algebra Appl.*, 14 (2007) 603-610.
- [4] M. P. Drazin, Pseudo inverses in associate rings and semigroups, *Amer. Math. Mon.*, 65 (1985) 506-514.
- [5] G. R. Goldstein, J. A. Goldstein, The best generalized inverse, *J. Math. Anal. Appl.*, 252 (2000)91-101.
- [6] H. C. Huang, Bounds on condition number of a matrix, *J. Comput. math.*, 2(4)(1984)356-360.
- [7] D. J. Higham, Condition numbers and their condition numbers, *SIAM J. Appl. Math.*, 24 (1973)315-323.
- [8] R. Penrose, A generalised inverse for matrices, *Proc. Cambridge Philos. Soc.*, 52 (1955) 406-413.
- [9] J. G. Sun, *Perturbation analysis for matrix*, Second ed., Science Press, Beijing, 2001.
- [10] G. Wang, Y. Wei, S. Qiao, *Generalized Inverses: Theory and Computations*, Science Press, Beijing, 2004.
- [11] Y. Wei, G. Wang, The perturbation theory for the Drazin inverse and its applications, *Linear Algebra Appl.*, 258 (1997)179-186.
- [12] Y. Wei, On the perturbation of the group inverse and oblique projection, *Appl. Math. Comput.*, 98 (1999)29-40.