



Condition Numbers of the Least Squares Problems with Multiple Right-Hand Sides

Lingsheng Meng^a, Bing Zheng^b

^aCollege of Mathematics and Statistics, Northwest Normal University, Lanzhou 730070, PR China

^bSchool of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, PR China

Abstract. In this paper, we investigate the normwise, mixed and componentwise condition numbers of the least squares problem $\min_{X \in \mathbb{R}^{n \times d}} \|AX - B\|_F$, where $A \in \mathbb{R}^{m \times n}$ is a rank-deficient matrix and $B \in \mathbb{R}^{m \times d}$. The closed formulas or upper bounds for these condition numbers are presented, which extend the earlier work for the least squares problem with single right-hand side (i.e. $B \equiv b$ is an m -vector) of several authors. Numerical experiments are given to confirm our results.

1. Introduction

Condition numbers and backward errors play an important role in numerical linear algebra [16]. Condition numbers measure the worst-case magnification in the computed outcome of a small perturbation in the data whereas backward errors can answer the question of how close is the problem actually solved to the one we want to solve. The product of a condition number and backward error provides a first-order of upper bound on the error in a computational solution. In particular, the condition numbers and backward errors of a linear system $Ax = b$ and a linear least squares (LS) problem with single right-hand side $\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$ have been extensively studied in the numerical linear algebra literature, e.g., see [1, 2, 4–6, 11, 12, 14–18, 21, 22].

The authors in [13, 26, 27] studied the backward errors and condition numbers for the following systems with multiple right-hand sides

$$AX = B, \quad A \in \mathbb{R}^{n \times n}, \quad B \in \mathbb{R}^{n \times d}.$$

For the LS problem with multiple right-hand sides

$$\min_{X \in \mathbb{R}^{n \times d}} \|AX - B\|_F, \quad A \in \mathbb{R}^{m \times n}, \quad B \in \mathbb{R}^{m \times d}, \quad (1.1)$$

Sun [20] derived its optimal backward perturbation bounds. In this paper, we will study the conditioning theory of the LS problem (1.1).

2010 *Mathematics Subject Classification.* 65F35, 65F20

Keywords. Least squares, Multiple right-hand sides, Condition number, Moore-Penrose inverse

Received: 01 March 2018; Accepted: 23 May 2019

Communicated by Dragan S. Djordjević

Research supported by the National Natural Science Foundation of China (No. 11701458 and No. 11571004)

Email addresses: menglsh@nwnu.edu.cn (Lingsheng Meng), bzheng@lzu.edu.cn (Bing Zheng)

To develop the conditioning theory of (1.1), a natural way is to transform it into the LS problem with single right-hand side, i.e.,

$$\min_{X \in \mathbb{R}^{n \times d}} \|(I_d \otimes A)\text{vec}(X) - \text{vec}(B)\|_2,$$

by the well-known Kronecker product and “vec” operation at first, then applying the conditioning theory of the LS problem with single right-hand side. But this technique always neglect special structure of the coefficient matrix $I_d \otimes A$, such that the corresponding condition number may be very undesirable. Therefore, it is necessary to study the condition numbers of the LS problem with multiple right-hand sides. Recently, Diao et al. [7] have studied the conditioning theory of $\min_v \|(C \otimes D)v - c\|_2$ when C and D have full column rank. From Corollary 4.1 in [7], we can easily get the closed formulas of the normwise, mixed and componentwise condition numbers for (1.1) when A has full column rank. In this paper, we will present the closed formulas for three kinds of normwise condition numbers and the upper bounds for the mixed and componentwise condition numbers of the LS problem (1.1) when A is a rank deficient matrix.

Throughout the paper, for given positive integers m and n , denote by \mathbb{R}^n the space of n -dimensional real column vectors, by $\mathbb{R}^{m \times n}$ the space of all $m \times n$ real matrices, and by $\|\cdot\|_2$ and $\|\cdot\|_F$ the 2-norm and Frobenius norm of their arguments, respectively. Given a matrix $X = [x_{ij}] \in \mathbb{R}^{m \times n}$, $\|X\|_{\max}$, X^\dagger , X^T denote the max norm, given by $\|X\|_{\max} = \max_{i,j} |x_{ij}|$, the Moore-Penrose inverse and the transpose of X , respectively, and $|X|$ is the matrix whose elements are $|x_{ij}|$. For the matrices $X = [x_{ij}]$, $Y = [y_{ij}] \in \mathbb{R}^{m \times n}$, $X \leq Y$ means $x_{ij} \leq y_{ij}$ for all i, j and we define $\frac{X}{Y} = [z_{ij}] \in \mathbb{R}^{m \times n}$ by

$$z_{ij} = \begin{cases} x_{ij}/y_{ij}, & \text{if } y_{ij} \neq 0, \\ 0, & \text{if } x_{ij} = y_{ij} = 0, \\ \infty, & \text{otherwise.} \end{cases}$$

2. Preliminaries

The operator vec and the Kronecker product will be of particular importance in what follows. The vec operator stacks the columns of the matrix argument into one long vector. For any matrices X and Y , the Kronecker product $X \otimes Y$ is defined by $X \otimes Y = [x_{ij}Y]$. It is enough for our purpose to recall the following properties concerning these operators. A more detailed list of such properties with their proofs can be found, e.g., in [10].

For any matrix $X = (x_{ij}) \in \mathbb{R}^{m \times n}$, $Y \in \mathbb{R}^{p \times q}$ and $Z \in \mathbb{R}^{n \times p}$, we have

$$(X \otimes Y)^T = X^T \otimes Y^T, \quad |X \otimes Y| = |X| \otimes |Y|, \quad \|X \otimes Y\|_2 = \|X\|_2 \|Y\|_2, \tag{2.1}$$

and

$$\text{vec}(XZY) = (Y^T \otimes X)\text{vec}(Z), \quad \text{vec}(X^T) = \Pi_{(m,n)}\text{vec}(X), \tag{2.2}$$

where $\Pi_{(m,n)} \in \mathbb{R}^{mn \times mn}$ is the permutation matrix defined by

$$\Pi_{(m,n)} = \sum_{i=1}^m \sum_{j=1}^n E_{ij} \otimes E_{ij}^T.$$

Here each $E_{ij} \in \mathbb{R}^{m \times n}$ has entry 1 in position (i, j) and all other entries are zero. Furthermore, we have

$$\Pi_{(p,m)}(Y \otimes X) = (X \otimes Y)\Pi_{(n,q)}. \tag{2.3}$$

In addition, the following two lemmas will be used in this paper.

Lemma 2.1. [9] If $E \in \mathbb{R}^{n \times n}$ and $\|E\|_2 < 1$, then $I_n - E$ is nonsingular and

$$(I_n - E)^{-1} = \sum_{k=0}^{\infty} E^k.$$

Lemma 2.2. [3] If $A, \Delta A \in \mathbb{R}^{m \times n}$ satisfy $\|A^\dagger \Delta A\|_2 < 1$, $R(\Delta A) \subseteq R(A)$ and $R((\Delta A)^T) \subseteq R(A^T)$, then

$$(A + \Delta A)^\dagger = (I_n + A^\dagger \Delta A)^{-1} A^\dagger.$$

3. Normwise condition numbers

When A is rank deficient, the LS solution to (1.1) always exists but it is nonunique. Therefore the unique minimum Frobenius norm LS solution $X_{LS} = A^\dagger B$ is considered. Moreover, when A is a rank deficient matrix, small changes to A can produce large changes to $X_{LS} = A^\dagger B$, see [19]. In other words, a condition number of X_{LS} with respect to rank deficient A does not exist or is “infinite”. Hence, in this section, we present the normwise, mixed and componentwise condition numbers of the LS problem (1.1) by restricting changes to the perturbation matrix ΔA of A , i.e. $\Delta A \in \mathcal{S}$, where

$$\mathcal{S} = \{ \Delta A : R(\Delta A) \subseteq R(A), R((\Delta A)^T) \subseteq R(A^T) \},$$

in which $R(A)$ denotes the range of A .

Let $\tilde{A} = A + \Delta A$ and $\tilde{B} = B + \Delta B$, where ΔA and ΔB are the perturbations of the input data A and B , respectively. Consider the perturbed LS problem of (1.1)

$$\min_{X \in \mathbb{R}^{n \times d}} \|\tilde{A}X - \tilde{B}\|_F. \tag{3.1}$$

If $\Delta A \in \mathcal{S}$ and the norm $\|\Delta A\|_2$ is sufficiently small, it follows from Lemma 2.2 that $\text{rank}(\tilde{A}) = \text{rank}(A)$, i.e., \tilde{A} is also rank deficient. Hence the unique minimum Frobenius norm LS solution to (3.1) is $\tilde{X}_{LS} = \tilde{A}^\dagger \tilde{B}$. We let the change in the solution be $\Delta X = \tilde{X}_{LS} - X_{LS}$.

In this section, we present three kinds of normwise condition numbers of (1.1) with respect to different norms. The closed formula for the normwise condition number with respect to the Frobenius norm of the pair (A, B) is given first.

Theorem 3.1. Let $A \in \mathbb{R}^{m \times n}$ be rank deficient and $B \in \mathbb{R}^{m \times d}$, then the condition number

$$\kappa_1(A, B) = \limsup_{\varepsilon \rightarrow 0} \left\{ \frac{\|\Delta X\|_F}{\varepsilon \|X_{LS}\|_F} : \|[\Delta A \ \Delta B]\|_F \leq \varepsilon \| [A \ B] \|_F, \Delta A \in \mathcal{S} \right\}$$

satisfies

$$\kappa_1(A, B) = \frac{\|A^\dagger\|_2 \| [A \ B] \|_F}{\|X_{LS}\|_F} \sqrt{1 + \|X_{LS}\|_2^2}. \tag{3.2}$$

Proof. When $\|\Delta A\|_2$ is sufficiently small, we may assume that $\|A^\dagger \Delta A\|_2 < 1$. Then, from Lemmas 2.1 and 2.2 with $R(\Delta A) \subseteq R(A)$, $R((\Delta A)^T) \subseteq R(A^T)$, neglecting the second-order terms gives

$$(A + \Delta A)^\dagger = A^\dagger - A^\dagger \Delta A A^\dagger.$$

Thus, for small ΔA and ΔB , the linear term in $\Delta X = (A + \Delta A)^\dagger (B + \Delta B) - A^\dagger B$ is

$$-A^\dagger \Delta A A^\dagger B + A^\dagger \Delta B = -A^\dagger \Delta A X_{LS} + A^\dagger \Delta B, \tag{3.3}$$

which implies that

$$\begin{aligned} \|-A^+\Delta AX_{LS} + A^+\Delta B\|_F &\leq \|A^+\|_2(\|X_{LS}\|_2\|\Delta A\|_F + \|\Delta B\|_F) \\ &= \|A^+\|_2[\|X_{LS}\|_2 \ 1] \begin{bmatrix} \|\Delta A\|_F \\ \|\Delta B\|_F \end{bmatrix} \\ &\leq \|A^+\|_2 \sqrt{1 + \|X_{LS}\|_2^2} \sqrt{\|\Delta A\|_F^2 + \|\Delta B\|_F^2} \\ &\leq \varepsilon \|A^+\|_2 \sqrt{1 + \|X_{LS}\|_2^2} \|[A \ B]\|_F. \end{aligned}$$

Since $-A^+\Delta AX_{LS} + A^+\Delta B$ is the linear term in ΔX , “ \leq ” in (3.2) holds. In the following we will show that this upper bound is reachable.

Let $\text{rank}(A) = r$, u and v be respectively the left and right singular vectors corresponding to the smallest positive singular value σ_r of A , then

$$\|A^+\|_2 = \frac{1}{\sigma_r}, \quad A^+u = \|A^+\|_2 v.$$

Moreover, let w and z be respectively the left and right singular vectors corresponding to the largest singular value of X_{LS} , then

$$X_{LS}z = \|X_{LS}\|_2 w, \quad (X_{LS})^T X_{LS}z = \|X_{LS}\|_2^2 z.$$

Constructing

$$\Delta A = -\frac{\varepsilon \|[A \ B]\|_F}{\sqrt{1 + \|X_{LS}\|_2^2}} uz^T X_{LS}^T, \quad \Delta B = \varepsilon \frac{\|[A \ B]\|_F}{\sqrt{1 + \|X_{LS}\|_2^2}} uz^T,$$

it follows from the fact $\|uv^T\|_F = \|uv^T\|_2 = \|u\|_2\|v\|_2$ (u, v are vectors) that

$$\begin{aligned} \|[\Delta A \ \Delta B]\|_F &= \sqrt{\|\Delta A\|_F^2 + \|\Delta B\|_F^2} \\ &= \varepsilon \frac{\|[A \ B]\|_F}{\sqrt{1 + \|X_{LS}\|_2^2}} \sqrt{\|uz^T X_{LS}^T\|_F^2 + \|uz^T\|_F^2} \\ &= \varepsilon \|[A \ B]\|_F \end{aligned}$$

and

$$R(\Delta A) \subseteq R(u) \subseteq R(A), \quad R((\Delta A)^T) \subseteq R(X_{LS}) \subseteq R(A^+) = R(A^T).$$

With these particular perturbations, we can get

$$\begin{aligned} \|-A^+\Delta AX_{LS} + A^+\Delta B\|_F &= \varepsilon \|[A \ B]\|_F \sqrt{1 + \|X_{LS}\|_2^2} \|A^+uz^T\|_F \\ &= \varepsilon \|[A \ B]\|_F \sqrt{1 + \|X_{LS}\|_2^2} \|A^+\|_2, \end{aligned}$$

giving equality in (3.2). Thus, we obtain (3.2). \square

The normwise condition number when the Frobenius norm is respectively used to measure A and B is given in the following theorem.

Theorem 3.2. *Let $A \in \mathbb{R}^{m \times n}$ be rank deficient and $B \in \mathbb{R}^{m \times d}$, then the condition number*

$$\kappa_2(A, B) = \limsup_{\varepsilon \rightarrow 0} \left\{ \frac{\|\Delta X\|_F}{\varepsilon \|X_{LS}\|_F} : \|\Delta A\|_F \leq \varepsilon \|A\|_F, \|\Delta B\|_F \leq \varepsilon \|B\|_F, \Delta A \in \mathcal{S} \right\}$$

satisfies

$$\kappa_2(A, B) = \frac{\|A^+\|_2}{\|X_{LS}\|_F} (\|A\|_F \|X_{LS}\|_2 + \|B\|_F). \tag{3.4}$$

Proof. By taking the Frobenius norm of (3.3), we obtain

$$\| -A^\dagger \Delta A X_{LS} + A^\dagger \Delta B \|_F \leq \varepsilon \|A^\dagger\|_2 (\|X_{LS}\|_2 \|A\|_F + \|B\|_F),$$

giving “ \leq ” in (3.4). Letting

$$\Delta A = -\frac{\varepsilon \|A\|_F}{\|X_{LS}\|_2} u z^T X_{LS}^T, \quad \Delta B = \varepsilon \|B\|_F u z^T$$

with u is the left singular vector corresponding to the smallest positive singular value σ_r of A and z is the right singular vectors corresponding to the largest singular value of X_{LS} , it is easy to check that

$$\|\Delta A\|_F = \varepsilon \|A\|_F, \quad \|\Delta B\|_F = \varepsilon \|B\|_F$$

and

$$R(\Delta A) \subseteq R(A), \quad R((\Delta A)^T) \subseteq R(A^T).$$

Hence, we have

$$\| -A^\dagger \Delta A X_{LS} + A^\dagger \Delta B \|_F = \varepsilon \|A^\dagger\|_2 (\|X_{LS}\|_2 \|A\|_F + \|B\|_F),$$

showing that equality is possible in (3.4). \square

Remark 1. Note that $\|\Delta A\|_F \leq \varepsilon \|A\|_F$ and $\|\Delta B\|_F \leq \varepsilon \|B\|_F$ imply

$$\|[\Delta A \ \Delta B]\|_F \leq \| [A \ B] \|_F,$$

hence it follows from the definitions of $\kappa_1(A, B)$ and $\kappa_2(A, B)$ that

$$\kappa_2(A, B) \leq \kappa_1(A, B).$$

It follows from (3.2) and (3.4) that

$$\begin{aligned} \kappa_2(A, B) &\leq \frac{\|A^\dagger\|_2}{\|X_{LS}\|_F} \left\| \begin{bmatrix} \|X_{LS}\|_2 & 1 \end{bmatrix} \right\|_F \left\| \begin{bmatrix} \|A\|_F \\ \|B\|_F \end{bmatrix} \right\|_F \\ &\leq \frac{\|A^\dagger\|_2}{\|X_{LS}\|_F} \sqrt{1 + \|X_{LS}\|_2^2} \sqrt{\|A\|_F^2 + \|B\|_F^2} \\ &= \frac{\|A^\dagger\|_2 \| [A \ B] \|_F}{\|X_{LS}\|_F} \sqrt{1 + \|X_{LS}\|_2^2} \\ &= \kappa_1(A, B), \end{aligned}$$

which also illustrates this fact.

Next theorem describes the characterization of the normwise condition number for the LS problems when 2-norm of matrix is used in Theorem 3.2. The proof of Theorem 3.3 is similar to that of Theorem 3.2 and so is omitted here.

Theorem 3.3. Let $A \in \mathbb{R}^{m \times n}$ be rank deficient and $B \in \mathbb{R}^{m \times d}$, then we have

$$\begin{aligned} \kappa_3(A, B) &= \limsup_{\varepsilon \rightarrow 0} \left\{ \frac{\|\Delta X\|_2}{\varepsilon \|X_{LS}\|_2} : \|\Delta A\|_2 \leq \varepsilon \|A\|_2, \|\Delta B\|_2 \leq \varepsilon \|B\|_2, \Delta A \in \mathcal{S} \right\} \\ &= \|A^\dagger\|_2 \|A\|_2 + \frac{\|A^\dagger\|_2 \|B\|_2}{\|X_{LS}\|_2}. \end{aligned} \tag{3.5}$$

When $d = 1$, i.e., $B \equiv b$, it follows from Theorems 3.1-3.3 that

Corollary 3.4. [23, 25] Let $A \in \mathbb{R}^{m \times n}$ be rank deficient and $b \in \mathbb{R}^m$. Then we have

$$\kappa_1(A, b) = \frac{\|A^\dagger\|_2 \|[A \ b]\|_F}{\|x_{LS}\|_2} \sqrt{1 + \|x_{LS}\|_2^2},$$

$$\kappa_2(A, b) = \|A^\dagger\|_2 \|A\|_F + \frac{\|A^\dagger\|_2 \|b\|_2}{\|x_{LS}\|_2}$$

and

$$\kappa_3(A, b) = \|A^\dagger\|_2 \|A\|_2 + \frac{\|A^\dagger\|_2 \|b\|_2}{\|x_{LS}\|_2}.$$

where $x_{LS} = A^\dagger b$.

4. Mixed and componentwise condition numbers

The normwise condition number measures both the input and output data errors by norms. Norms can tell us the overall size of a perturbation but not how that size is distributed among the elements it perturbs, and this information can be important when the data is badly scaled or contains many zeros [18]. To take into account the relative of each data component, and, in particular, a possible data sparseness, componentwise condition numbers have been increasingly considered. These are mostly of two kinds: mixed and componentwise. The terminologies of mixed and componentwise condition numbers may be first used by Gohberg and Koltracht [8]. We adopt their terminology and define the mixed and componentwise condition numbers for the LS problem (1.1) are defined as follows:

$$m(A, B) = \lim_{\epsilon \rightarrow 0} \sup_{\substack{|\Delta A| \leq \epsilon |A| \\ |\Delta B| \leq \epsilon |B| \\ \Delta A \in S}} \frac{\|\Delta X\|_{\max}}{\epsilon \|X_{LS}\|_{\max}}$$

and

$$c(A, B) = \lim_{\epsilon \rightarrow 0} \sup_{\substack{|\Delta A| \leq \epsilon |A| \\ |\Delta B| \leq \epsilon |B| \\ \Delta A \in S}} \frac{1}{\epsilon} \left\| \frac{\Delta X}{X_{LS}} \right\|_{\max}.$$

We assume that $X_{LS} \neq 0$ for $m(A, B)$ and X_{LS} has no zero entries for $c(A, B)$.

The following theorem gives the upper bounds for the mixed and componentwise condition numbers of the LS problem.

Theorem 4.1. Let $A \in \mathbb{R}^{m \times n}$ be rank deficient and $B \in \mathbb{R}^{m \times d}$. Then we have

$$m(A, B) \leq \frac{\| |A^\dagger| |A| |X_{LS}| + |A^\dagger| |B| \|_{\max}}{\|X_{LS}\|_{\max}} := \bar{m}(A, B) \tag{4.1}$$

and

$$c(A, B) \leq \left\| \frac{|A^\dagger| |A| |X_{LS}| + |A^\dagger| |B|}{X_{LS}} \right\|_{\max} := \bar{m}(A, B). \tag{4.2}$$

Proof. According to $|\Delta A| \leq \epsilon |A|$, we know that the zero elements of A are not permitted to be perturbed. Therefore,

$$\text{vec}(\Delta A) = D_A D_A^\dagger \text{vec}(\Delta A),$$

where $D_A = \text{diag}(\text{vec}(A))$. Similarly, we have $\text{vec}(\Delta B) = D_B D_B^\dagger \text{vec}(\Delta B)$ with $D_B = \text{diag}(\text{vec}(B))$. Thus the linear term $-A^\dagger \Delta A X_{LS} + A^\dagger \Delta B$ of ΔX can be rewritten as

$$\begin{aligned} & \text{vec}(-A^\dagger \Delta A X_{LS} + A^\dagger \Delta B) \\ &= -(X_{LS}^T \otimes A^\dagger) \text{vec}(\Delta A) + (I_d \otimes A^\dagger) \text{vec}(\Delta B) \\ &= \begin{bmatrix} -X_{LS}^T \otimes A^\dagger & I_d \otimes A^\dagger \end{bmatrix} \begin{bmatrix} D_A & 0 \\ 0 & D_B \end{bmatrix} \begin{bmatrix} D_A^\dagger \text{vec}(\Delta A) \\ D_B^\dagger \text{vec}(\Delta B) \end{bmatrix}. \end{aligned} \tag{4.3}$$

Taking the infinity norm and using the assumption $|\Delta A| \leq \varepsilon|A|$ and $|\Delta B| \leq \varepsilon|B|$, we have

$$\begin{aligned} \|-A^\dagger \Delta A X_{LS} + A^\dagger \Delta B\|_{\max} &= \|\text{vec}(-A^\dagger \Delta A X_{LS} + A^\dagger \Delta B)\|_\infty \\ &\leq \varepsilon \left\| \begin{bmatrix} (-X_{LS}^T \otimes A^\dagger) D_A & (I_d \otimes A^\dagger) D_B \end{bmatrix} \right\|_\infty. \end{aligned}$$

Since $-A^\dagger \Delta A X_{LS} + A^\dagger \Delta B$ is the linear term of ΔX , $m(A, B)$ is bounded above by

$$\begin{aligned} m(A, B) &\leq \frac{\left\| \begin{bmatrix} (-X_{LS}^T \otimes A^\dagger) D_A & (I_d \otimes A^\dagger) D_B \end{bmatrix} \right\|_\infty}{\|X_{LS}\|_{\max}} \\ &= \frac{\left\| \begin{bmatrix} |(-X_{LS}^T \otimes A^\dagger) D_A| & |(I_d \otimes A^\dagger) D_B| \end{bmatrix} e \right\|_\infty}{\|X_{LS}\|_{\max}} \\ &= \frac{\left\| |X_{LS}^T \otimes A^\dagger| \text{vec}(|A|) + |I_d \otimes A^\dagger| \text{vec}(|B|) \right\|_\infty}{\|X_{LS}\|_{\max}} \\ &= \frac{\left\| \text{vec}(|A^\dagger| |A| |X_{LS}| + |A^\dagger| |B|) \right\|_\infty}{\|X_{LS}\|_{\max}} \\ &= \frac{\left\| |A^\dagger| |A| |X_{LS}| + |A^\dagger| |B| \right\|_{\max}}{\|X_{LS}\|_{\max}}, \end{aligned}$$

where e is an $m(n + d)$ dimensional vector with all entries equal to one.

Recall that in the definition of $c(A, B)$, we assume that X_{LS} has no zero entries. Hence, it follows from (4.3) and the assumption $|\Delta A| \leq \varepsilon|A|$, $|\Delta B| \leq \varepsilon|B|$ that

$$\begin{aligned} & \left\| \frac{-A^\dagger \Delta A X_{LS} + A^\dagger \Delta B}{X_{LS}} \right\|_{\max} = \left\| \frac{\text{vec}(-A^\dagger \Delta A X_{LS} + A^\dagger \Delta B)}{\text{vec}(X_{LS})} \right\|_\infty \\ &= \|D_{X_{LS}}^{-1} \text{vec}(-A^\dagger \Delta A X_{LS} + A^\dagger \Delta B)\|_\infty \\ &\leq \left\| D_{X_{LS}}^{-1} \begin{bmatrix} (-X_{LS}^T \otimes A^\dagger) D_A & (I_d \otimes A^\dagger) D_B \end{bmatrix} \right\|_\infty \left\| \begin{bmatrix} D_A^\dagger \text{vec}(\Delta A) \\ D_B^\dagger \text{vec}(\Delta B) \end{bmatrix} \right\|_\infty \\ &\leq \varepsilon \left\| D_{X_{LS}}^{-1} \begin{bmatrix} (-X_{LS}^T \otimes A^\dagger) D_A & (I_d \otimes A^\dagger) D_B \end{bmatrix} \right\|_\infty, \end{aligned}$$

where $D_{X_{LS}} = \text{diag}(\text{vec}(X_{LS}))$. Hence, we have

$$\begin{aligned} c(A, B) &\leq \left\| D_{X_{LS}}^{-1} \begin{bmatrix} (-X_{LS}^T \otimes A^\dagger) D_A & (I_d \otimes A^\dagger) D_B \end{bmatrix} \right\|_\infty \\ &= \left\| D_{X_{LS}}^{-1} \left(|X_{LS}^T \otimes A^\dagger| \text{vec}(|A|) + |I_d \otimes A^\dagger| \text{vec}(|B|) \right) \right\|_\infty \\ &= \left\| \frac{\text{vec}(|A^\dagger| |A| |X_{LS}| + |A^\dagger| |B|)}{\text{vec}(|X_{LS}|)} \right\|_\infty \\ &= \left\| \frac{|A^\dagger| |A| |X_{LS}| + |A^\dagger| |B|}{X_{LS}} \right\|_{\max}. \end{aligned}$$

The proof of the theorem is now completed. \square

Remark 2. Theorem 4.1 gives upper bounds for the mixed and componentwise condition numbers. In fact, for some special matrices, (4.1) is attainable. For any A and B satisfying $A = |A|$, $A^\dagger = |A^\dagger|$ and $B = |B|$, let $\Delta A = -\varepsilon A$ and $\Delta B = \varepsilon B$. We can get

$$|\Delta A| = \varepsilon|A|, |\Delta B| = \varepsilon|B|, R(\Delta A) \subseteq R(A), \text{ and } R((\Delta A)^T) \subseteq R(A^T).$$

For these particular matrices ΔA and ΔB , we have

$$\begin{aligned} m(A, B) &\geq \limsup_{\varepsilon \rightarrow 0} \frac{\|(A + \Delta A)^\dagger(B + \Delta B) - A^\dagger B\|_{\max}}{\varepsilon \|X_{LS}\|_{\max}} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{2}{1 - \varepsilon} = 2. \end{aligned}$$

On the other hand, since $A = |A|$, $A^\dagger = |A^\dagger|$ and $B = |B|$,

$$\frac{\| |A^\dagger| \|A\|_{X_{LS}} + |A^\dagger| \|B\|_{\max} \|_{\max}}{\|X_{LS}\|_{\max}} = \frac{\|A^\dagger A A^\dagger B + A^\dagger B\|_{\max}}{\|A^\dagger B\|_{\max}} = 2.$$

Similarly, we can prove that (4.2) is also attainable for some special matrices.

When $d = 1$, i.e., $B \equiv b$, it follows from Theorem 4.1 that

Corollary 4.2. [24] Let $A \in \mathbb{R}^{m \times n}$ be rank deficient and $b \in \mathbb{R}^m$. Then we have

$$m(A, b) \leq \frac{\| |A^\dagger| \|A\|_{X_{LS}} + |A^\dagger| \|b\|_{\infty} \|_{\infty}}{\|x_{LS}\|_{\infty}}$$

and

$$c(A, b) \leq \left\| \frac{|A^\dagger| \|A\|_{X_{LS}} + |A^\dagger| \|b\|_{\infty}}{x_{LS}} \right\|_{\infty}.$$

5. Numerical experiments

We consider the LS problem (1.1) with

$$A = \begin{bmatrix} 9 \times 10^i & 0 & 0 \\ 0 & 2 & 2 \\ 3 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \\ 4 & 4 \end{bmatrix}, i = 0, 2, 4, 6.$$

We first compare $\kappa_1(A, B)$, $\kappa_2(A, B)$, $\kappa_3(A, B)$ with the upper bounds of the mixed and componentwise condition numbers given in Theorem 4.1. Thus, upon computations in MATLAB R2015b with precision 2.2204×10^{-16} , we get the results listed in Table 1. From Table 1, we find that as the (1, 1)-element of A increases, the normwise condition numbers become larger and larger, while, comparatively, the mixed and componentwise condition numbers have no change. This is mainly because the mixed and componentwise condition numbers notice the structure of the coefficient matrix A with respect to scaling, but the normwise condition numbers ignore it.

Now we show the tightness of the upper bound estimates on the mixed and componentwise condition numbers provided in Theorem 4.1. For $i = 0$, suppose the perturbations are $\Delta A = 10^{-j} \times A$ and $\Delta B = 10^{-j+1} \times \text{rand}(4, 2)$, where $\text{rand}(\cdot)$ is the MATLAB function. Obviously, $\Delta A \in \mathcal{S} = \{ \Delta A : R(\Delta A) \subseteq R(A), R((\Delta A)^T) \subseteq R(A^T) \}$. Define $\varepsilon_1 = \min\{ \varepsilon : |\Delta A| \leq \varepsilon|A|, |\Delta B| \leq \varepsilon|B| \}$, we list the computed results in Table 2. As shown in Table 2, the error bounds given by the upper bounds of the condition numbers in Theorem 4.1 are at most one order of magnitude larger than the actual errors. This illustrates that, as the estimates of their corresponding condition numbers, the upper bounds in Theorem 4.1 are tight.

Table 1: Comparison of condition numbers

| | $\kappa_1(A, B)$ | $\kappa_2(A, B)$ | $\kappa_3(A, B)$ | $\bar{m}(A, B)$ | $\bar{c}(A, B)$ |
|---------|----------------------|----------------------|----------------------|-----------------|-----------------|
| $i = 0$ | 4.6969 | 4.6698 | 4.5076 | 2.0000 | 2.0000 |
| $i = 2$ | 335.6361 | 286.1393 | 286.1375 | 2.0000 | 2.0000 |
| $i = 4$ | 3.3562×10^4 | 2.8462×10^4 | 2.8462×10^4 | 2.0000 | 2.0000 |
| $i = 6$ | 3.3562×10^6 | 2.8461×10^6 | 2.8461×10^6 | 2.0000 | 2.0000 |

Table 2: Comparisons of our estimated errors with the exact errors

| j | 6 | 8 | 10 | 12 |
|---|-------------------------|-------------------------|--------------------------|--------------------------|
| $\ \Delta X\ _{\max}/\ X_{LS}\ _{\max}$ | 1.0145×10^{-6} | 1.0323×10^{-8} | 1.7363×10^{-10} | 1.5721×10^{-12} |
| $\varepsilon_1 \bar{m}(A, b)$ | 1.3897×10^{-5} | 8.7749×10^{-8} | 1.4187×10^{-9} | 1.7818×10^{-11} |
| $\ \Delta X/X_{LS}\ _{\max}$ | 4.0578×10^{-6} | 2.4696×10^{-8} | 3.0069×10^{-10} | 4.3665×10^{-12} |
| $\varepsilon_1 \bar{c}(A, b)$ | 1.3897×10^{-5} | 8.7749×10^{-8} | 1.4187×10^{-9} | 1.7818×10^{-11} |

Acknowledgement

I'm grateful to the handing editor and referees for their helpful comments and suggestions.

References

- [1] M. Arioli, M. Baboulin, S. Gratton, *A partial condition number for linear least squares problems*, SIAM J. Matrix Anal. Appl. 29 (2007) 413–433.
- [2] M. Baboulin, S. Gratton, *Using dual techniques to derive componentwise and mixed condition numbers for a linear function of a linear least squares solution*, BIT 49 (2009) 3–19.
- [3] A. Ben-Israel, *On error bounds for generalized inverses*, SIAM J. Numer. Anal. 3 (1966) 585–592.
- [4] Å. Björck, *Component-wise perturbation analysis and error bounds for linear least squares solutions*, BIT 31 (1991) 238–244.
- [5] F. Cucker, H. Diao, Y. Wei, *On mixed and componentwise condition numbers for Moore-Penrose inverse and linear least squares problems*, Math. Comp. 76 (2007) 947–963.
- [6] H. Diao, Y. Wei, *On Frobenius normwise condition numbers for Moore-Penrose inverse and linear least-squares problems*, Numer. Linear Algebra Appl. 14 (2007) 603–610.
- [7] H. Diao, W. Wang, Y. Wei, S. Qiao, *On condition numbers for Moore-Penrose inverse and linear least squares problem involving Kronecker products*, Numer. Linear Algebra Appl. 20 (2013) 44–59.
- [8] I. Gohberg, I. Koltracht, *Mixed, componentwise, and structured condition numbers*, SIAM J. Matrix Anal. Appl. 14 (1993) 688–704.
- [9] G.H. Golub, C.F. Van Loan, *Matrix Computations*, 4th ed., Johns Hopkins University Press, Baltimore, 2013.
- [10] A. Graham, *Kronecker Products and Matrix Calculus with Application*, Wiley, New York, 1981.
- [11] S. Gratton, *On the condition number of linear least squares problems in a weighted Frobenius norm*, BIT 36 (1996) 523–530.
- [12] M. Gu, *Backward perturbation bounds for linear least squares problems*, SIAM J. Matrix Anal. Appl. 20 (1998) 363–372.
- [13] D.J. Higham, N.J. Higham, *Componentwise perturbation theory for linear systems with multiple right-hand sides*, Linear Algebra Appl. 174 (1992) 111–129.
- [14] D.J. Higham, N.J. Higham, *Backward error and condition of structured linear systems*, SIAM J. Matrix Anal. Appl. 13 (1992) 162–175.
- [15] D.J. Higham, *Condition numbers and their condition numbers*, Linear Algebra Appl. 214 (1995) 193–213.
- [16] N.J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, 2002.
- [17] A.N. Malyshev, *A unified theory of conditioning for linear least squares and Tikhonov regularization solutions*, SIAM J. Matrix Anal. Appl. 24 (2003) 1186–1196.

- [18] J. Rohn, *New condition numbers for matrices and linear systems*, *Computing* 41 (1989) 167–169.
- [19] G.W. Stewart, *On the perturbation of the pseudo-inverse, projections, and linear squares problems*, *SIAM Rev.* 19 (1977) 634–662.
- [20] J.G. Sun, *Optimal backward perturbation bounds for linear least squares problem with multiple right-hand sides*, *IMA J. Numer. Anal.* 16 (1996) 1–11.
- [21] J.G. Sun, Z. Sun, *Optimal backward perturbation bounds for underdetermined systems*, *SIAM J. Matrix Anal. Appl.* 18 (1997) 393–402.
- [22] B. Walden, R. Karlson, J.G. Sun, *Optimal backward perturbation bounds for the linear least squares problem*, *Numer. Linear Algebra Appl.* 2 (1995) 271–286.
- [23] Y. Wei, D. Wang, *Condition numbers and perturbation of the weighted Moore-Penrose inverse and weighted linear least squares problem*, *Appl. Math. Comput.* 145 (2003) 45–58.
- [24] Y. Wei, W. Xu, S. Qiao, H. Diao, *Componentwise condition numbers for generalized matrix inversion and linear least squares*, *Numer. Math. J. Chinese Univ. (English Ser.)* 14 (2005) 277–286.
- [25] Y. Wei, H. Diao, S. Qiao, *Condition number for weighted linear least squares problem*, *J. Comput. Math.* 25 (2007) 561–572.
- [26] X. Yang, H. Dai, Q. He, *Condition numbers and backward perturbation bound for linear matrix equations*, *Numer. Linear Algebra Appl.* 18 (2011) 155–165.
- [27] H. Zhang, H. Xiang, Y. Wei, *Condition numbers for linear systems and Kronecker product linear systems with multiple right-hand sides*, *Int. J. Comput. Math.* 84 (2007) 1805–1817.