Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 6(9): 9366–9377.
DOI:10.3934/math.2021544
Received: 07 March 2021
Accepted: 11 June 2021
Published: 23 June 2021

Research article

Condition numbers of the minimum norm least squares solution for the least squares problem involving Kronecker products

Lingsheng Meng*and Limin Li

College of Mathematics and Statistics, Northwest Normal University, Lanzhou 730070, China

* **Correspondence:** Email: menglsh@nwnu.edu.cn; Tel: +8615117097863.

Abstract: In this paper, the upper bounds for two kinds of normiwse condition numbers are derived for $\min_x \|(A \otimes B)x - b\|_2$ when the coefficient matrix is of rank deficient. In addition, the upper bounds on the mixed and componentwise condition numbers are also given. Numerical experiments are given to confirm our results.

Keywords: condition number; Kronecker product; least squares; upper bound; componentwise
Mathematics Subject Classification: 65F35

1. Introduction

Consider the least squares (LS) problem involving Kronecker products

$$\min_{x \in \mathbb{R}^{nq}} \|(A \otimes B)x - b\|_2, \tag{1.1}$$

where $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{p\times q}$, $b \in \mathbb{R}^{mp}$, and where the Kronecker product [13] is defined by

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}.$$

The LS problems involving Kronecker products of the type (1.1) arise in many areas of application including signal and image processing, photogrammetry, fast transform algorithms, the Lyapunov approach to stability theory, circuits and systems, and stochastic matrices [8, 19]. Therefore, the LS problems involving Kronecker products have attracted many researchers to study its algorithms [2, 9, 10, 18].

For a given problem, a condition number measures the worst-case sensitivity of its solution to small perturbations in the input data, whereas backward errors reveal the stability of a numerical method. Combined with the backward error, a condition number can provide a first-order upper bound on the error in the computational solution [15]. In [22], the authors gave the upper bounds on the normwise, mixed and componentwise condition numbers for the Kronecker product linear systems $(A \otimes B)x = b$. The level-2 condition numbers for the Kronecker product linear systems have studied by Kang and Xiang [17]. Xiang and Wei [23] derived the explicit expressions of the normwise, mixed and componentwise condition numbers for the Kronecker product linear systems with multiple right-hand sides. For the LS problem involving Kronecker products of (1.1), Diao et al. [7] have studied its conditioning theory when A and B are of full column rank. Chen and Li [5] presented explicit expressions for normwise, mixed and componentwise condition numbers for the weighted least squares problem involving Kronecker when A and B are of full column rank. In this paper, we will present the upper bounds for the normwise, mixed and componentwise condition numbers of the LS problem (1.1) when A or B is of rank deficient.

According to the fact that $\mathrm{rank}(A \otimes B) = \mathrm{rank}(A)\mathrm{rank}(B)$, it follows that $A \otimes B$ is rank deficient when A or B is of rank deficient. In this case, the LS solution to (1.1) always exists but it is nonunique. Therefore the unique minimum norm LS solution $x_{LS} = (A \otimes B)^{\dagger} b = (A^{\dagger} \otimes B^{\dagger})b$ is considered, where $A^{\dagger}$ denotes the Moore-Penrose inverse of A. Moreover, when $A \otimes B$ is a rank deficient matrix, small changes to A or B can produce large changes to $x_{LS} = (A^{\dagger} \otimes B^{\dagger})b$, see Example 1. In other words, a condition number of x_{LS} with respect to rank deficient $A \otimes B$ 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 matrices ΔA or ΔB, i.e. $\Delta A \in \mathcal{S}$ or $\Delta B \in \mathcal{T}$, where

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

and

$$\mathcal{T} = \left\{ \Delta B : R(\Delta B) \subseteq R(B),\ R((\Delta B)^T) \subseteq R(B^T) \right\}.$$

Here $R(A)$ denotes the range of A.

Example 1. Let

$$A = B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ A + \Delta A = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \end{bmatrix},\ b = \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix}^T.$$

By simple computations, we have

$$x_{LS} = (A^{\dagger} \otimes B^{\dagger})b = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^T,\ \widetilde{x}_{LS} = \left((A + \Delta A)^{\dagger} \otimes B^{\dagger}\right)b = \begin{bmatrix} 1 & 0 & \frac{1}{\varepsilon} & 0 \end{bmatrix}^T$$

and

$$\|\widetilde{x}_{LS} - x_{LS}\|_2 = \frac{1}{\varepsilon}.$$

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 $\|X\|_{\max} = \max\limits_{i,j} |x_{ij}|$, the Moore-Penrose inverse and the transpose of X, respectively, and $|X|$ is the matrix whose elements are $|x_{ij}|$. For the matrices $X = [x_{ij}]$, $Y = [y_{ij}] \in \mathbb{R}^{m\times n}$, $X \le Y$ means $x_{ij} \le y_{ij}$ for all i, j and we define $\frac{X}{Y} = [z_{ij}] \in \mathbb{R}^{m\times n}$ by

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

The following lemmas will be used in the later discussion.

Lemma 1.1. [13] *For any matrix $X = (x_{ij}) \in R^{m\times n}$, $Y \in R^{p\times q}$ and $Z \in R^{n\times p}$, we have*

$$|X \otimes Y| = |X| \otimes |Y|, \quad \|X \otimes Y\|_2 = \|X\|_2\|Y\|_2, \quad \|X \otimes Y\|_\infty = \|X\|_\infty\|Y\|_\infty.$$

Furthermore, with $vec(\cdot)$ stacking columns, we have

$$vec(XZY) = (Y^T \otimes X)vec(Z), \; vec(X \otimes Y) = (I_n \otimes K_{qm} \otimes I_p)\left(vec(X) \otimes vec(Y)\right), \; a \otimes c = vec(ca^T),$$

where $a \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $K_{mn} \in \mathbb{R}^{mn\times mn}$ is the permutation matrix defined by

$$K_{mn} = \sum_{i=1}^{m}\sum_{j=1}^{n} E_{ij} \otimes E_{ij}^T.$$

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

Lemma 1.2. [12] *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 1.3. [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*

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

2. Normwise condition numbers

The following weighted Frobenius norm

$$\|(\alpha A, \beta b)\|_F = \sqrt{\alpha^2\|A\|_F^2 + \beta^2\|b\|_2^2}$$

was first used by Gratton for deriving the normwise condition number for the linear least squares problem [14]. Here $\|\diamond\|_F$ denotes the Frobenius norm of a matrix, and $\|\diamond\|_2$ denotes the spectral norm of a matrix or the Euclidean norm of a vector. We will call the latter 2-norm uniformly later in this paper. Subsequently, this kind of norm was used for the partial condition number for the linear least squares problem [1] and the normwise condition number of the truncated singular value solution of a linear ill-posed problem [4]. A more general weighted Frobenius norm $\left\|\begin{bmatrix} AT & \beta b \end{bmatrix}\right\|_F$, where T is a positive diagonal matrix, is sometimes chosen. This is the case, for instance, in [6], which gives the explicit expressions for the normwise, mixed and componentwise condition numbers of the Tikhonov regularization using this norm. In this paper, we use the weighted Frobenius norm and weighted 2-norm which defined by

$$\|(A,B,b)\|_{\mathcal{F}} = \sqrt{\alpha^2\|A\|_F^2 + \beta^2\|B\|_F^2 + \gamma^2\|b\|_2^2},$$

and

$$\|(A,B,b)\|_{\in} = \sqrt{\alpha^2\|A\|_2^2 + \beta^2\|B\|_2^2 + \gamma^2\|b\|_2^2}$$

with $\alpha,\ \beta,\ \gamma > 0$. These norms are very flexible since they allow us to monitor the perturbations on A, B and b. For instance, large values of α enable us to obtain condition number problems where mainly B and b are perturbed.

Consider the perturbed LS problem of (1.1)

$$\min_{x\in\mathbb{R}^{nq}} \|((A+\Delta A)\otimes(B+\Delta B))x - (b+\Delta b)\|_2, \tag{2.1}$$

where ΔA, ΔB and Δb are the perturbations of the input data A, B and b, respectively. The unique minimum norm LS solution to (2.1) is $\widetilde{x}_{LS} = \left((A+\Delta A)^\dagger \otimes (B+\Delta B)^\dagger\right)(b+\Delta b)$. We let the change in the solution be $\Delta x = \widetilde{x}_{LS} - x_{LS}$.

Theorem 2.1. *When A and B are of rank deficient, the condition number*

$$\kappa^{(\mathcal{F})}(A,B,b) = \lim_{\varepsilon\to 0}\sup\left\{\frac{\|\Delta x\|_2}{\varepsilon\|x_{LS}\|_2} : \|(\Delta A,\Delta B,\Delta b)\|_{\mathcal{F}} \le \varepsilon\,\|(A,B,b)\|_{\mathcal{F}},\, \Delta A\in\mathcal{S}, \Delta B\in\mathcal{T}\right\}$$

satisfies

$$\kappa^{(\mathcal{F})}(A,B,b) \le \frac{\left\|\begin{bmatrix}\frac{P}{\alpha} & \frac{Q}{\beta} & -\frac{A^\dagger\otimes B^\dagger}{\gamma}\end{bmatrix}\right\|_2 \|(A,B,b)\|_{\mathcal{F}}}{\|x_{LS}\|_2} := \kappa^{(\mathcal{F})}(A,B,b)^{upper1},$$

where $P = (x_{LS}^T\otimes A^\dagger\otimes B^\dagger)(I_n\otimes K_{qm}\otimes I_p)(I_{mn}\otimes vec(B))$ *and* $Q = (x_{LS}^T\otimes A^\dagger\otimes B^\dagger)(I_n\otimes K_{qm}\otimes I_p)(vec(A)\otimes I_{pq})$.

Proof. When $\|\Delta A\|_2$ is sufficiently small, we may assume that $\|A^\dagger\Delta A\|_2 < 1$. Then, from Lemmas 1.2 and 1.3 with $R(\Delta A)\subseteq R(A)$, $R((\Delta A)^T)\subseteq R(A^T)$, neglecting the second-order terms gives

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

Similarly, we have $(B+\Delta B)^\dagger = B^\dagger - B^\dagger\Delta B B^\dagger$ when $\|\Delta B\|_2$ is sufficiently small and $\Delta B\in\mathcal{T}$. Thus, for small ΔA and ΔB, the linear term in $\Delta x = \left((A+\Delta A)^\dagger\otimes(B+\Delta B)^\dagger\right)(b+\Delta b) - (A^\dagger\otimes B^\dagger)b$ is

$$\begin{aligned}
& (A^\dagger\otimes B^\dagger)\Delta b - \left(A^\dagger\otimes(B^\dagger\Delta B B^\dagger)\right)b - \left((A^\dagger\Delta A A^\dagger)\otimes B^\dagger\right)b \\
= \;& (A^\dagger\otimes B^\dagger)\Delta b - (A^\dagger\otimes B^\dagger)(A\otimes\Delta B)(A^\dagger\otimes B^\dagger)b - (A^\dagger\otimes B^\dagger)(\Delta A\otimes B)(A^\dagger\otimes B^\dagger)b \\
= \;& (A^\dagger\otimes B^\dagger)\Delta b - (A^\dagger\otimes B^\dagger)\,(A\otimes\Delta B + \Delta A\otimes B)\,x_{LS}.
\end{aligned} \tag{2.2}$$

Applying the operator vec to (2.2) and using Lemma 1.1, we have

$$\begin{aligned}
& (A^\dagger\otimes B^\dagger)\Delta b - (A^\dagger\otimes B^\dagger)\,(A\otimes\Delta B + \Delta A\otimes B)\,x_{LS} \\
= \;& (A^\dagger\otimes B^\dagger)\Delta b - (x_{LS}^T\otimes A^\dagger\otimes B^\dagger)(\mathrm{vec}(A\otimes\Delta B) + \mathrm{vec}(\Delta A\otimes B))
\end{aligned}$$

$$
\begin{aligned}
&= (A^{\dagger}\otimes B^{\dagger})\Delta b-(x_{LS}^{T}\otimes A^{\dagger}\otimes B^{\dagger})(I_n\otimes K_{qm}\otimes I_p)\,(\mathrm{vec}(A)\otimes \mathrm{vec}(\Delta B)+\mathrm{vec}(\Delta A)\otimes \mathrm{vec}(B))\\
&= (A^{\dagger}\otimes B^{\dagger})\Delta b-(x_{LS}^{T}\otimes A^{\dagger}\otimes B^{\dagger})(I_n\otimes K_{qm}\otimes I_p)((\mathrm{vec}(A)\otimes I_{pq})\mathrm{vec}(\Delta B)\\
&\quad +(I_{mn}\otimes \mathrm{vec}(B))\mathrm{vec}(\Delta A))\\
&= \begin{bmatrix}-\frac{P}{\alpha} & -\frac{Q}{\beta} & \frac{A^{\dagger}\otimes B^{\dagger}}{\gamma}\end{bmatrix}\begin{bmatrix}\alpha\mathrm{vec}(\Delta A)\\ \beta\mathrm{vec}(\Delta B)\\ \gamma\Delta b\end{bmatrix}.
\end{aligned}
$$

Consequently, we have

$$
\begin{aligned}
&\|(A^{\dagger}\otimes B^{\dagger})\Delta b-(A^{\dagger}\otimes B^{\dagger})\,(A\otimes\Delta B+\Delta A\otimes B)\,x_{LS}\|_2\\
\leq\ & \left\|\begin{bmatrix}\frac{P}{\alpha} & \frac{Q}{\beta} & -\frac{A^{\dagger}\otimes B^{\dagger}}{\gamma}\end{bmatrix}\right\|_2 \left\|\begin{bmatrix}\alpha\mathrm{vec}(\Delta A)\\ \beta\mathrm{vec}(\Delta B)\\ \gamma\Delta b\end{bmatrix}\right\|_2\\
=\ & \varepsilon\left\|\begin{bmatrix}\frac{P}{\alpha} & \frac{Q}{\beta} & -\frac{A^{\dagger}\otimes B^{\dagger}}{\gamma}\end{bmatrix}\right\|_2 \|(A,B,b)\|_{\mathcal{F}}\,.
\end{aligned}
$$

Since $(A^{\dagger}\otimes B^{\dagger})\Delta b-(A^{\dagger}\otimes B^{\dagger})\,(A\otimes\Delta B+\Delta A\otimes B)\,x_{LS}$ is the linear term in Δx, Theorem 3.1 holds. □

The formula for $\kappa^{(\mathcal{F})}(A,B,b)^{\text{upper1}}$ in Theorem 2.1 involve the permutation matrix K_{qm} and Kronecker products, which make them inefficient for computation. In order to overcome this shortcoming, the next corollary will provide easily computable upper bound.

Corollary 2.1. *For the estimate $\kappa^{(\mathcal{F})}(A,B,b)^{upper1}$ in Theorem 2.1, we have*

$$
\kappa^{(\mathcal{F})}(A,B,b)^{upper1}\leq \|A^{\dagger}\|_2\|B^{\dagger}\|_2\,\|(A,B,b)\|_{\mathcal{F}}\left(\frac{\|B\|_F}{\alpha}+\frac{\|A\|_F}{\beta}+\frac{1}{\gamma\|x_{LS}\|_2}\right):=\kappa^{(\mathcal{F})}(A,B,b)^{upper2}.
$$

Proof. It follows from Lemma 1.1 that

$$
\begin{aligned}
&\left\|\begin{bmatrix}\frac{P}{\alpha} & \frac{Q}{\beta} & -\frac{A^{\dagger}\otimes B^{\dagger}}{\gamma}\end{bmatrix}\right\|_2 \leq \frac{\|P\|_2}{\alpha}+\frac{\|Q\|_2}{\beta}+\frac{\|A^{\dagger}\otimes B^{\dagger}\|_2}{\gamma}\\
\leq\ & \frac{\|A^{\dagger}\|_2\|B^{\dagger}\|_2\|x_{LS}\|_2\|\mathrm{vec}(B)\|_2}{\alpha}+\frac{\|A^{\dagger}\|_2\|B^{\dagger}\|_2\|x_{LS}\|_2\|\mathrm{vec}(A)\|_2}{\beta}+\frac{\|A^{\dagger}\|_2\|B^{\dagger}\|_2}{\gamma}\\
=\ & \|A^{\dagger}\|_2\|B^{\dagger}\|_2\left(\frac{\|x_{LS}\|_2\|B\|_F}{\alpha}+\frac{\|x_{LS}\|_2\|A\|_F}{\beta}+\frac{1}{\gamma}\right),
\end{aligned}
$$

which together with Theorem 2.1 complete the proof of this corollary. □

The normwise condition number under the weighted 2-norm is given in the following theorem.

Theorem 2.2. *When A and B are of rank deficient, the condition number*

$$
\kappa^{(\epsilon)}(A,B,b)=\lim_{\varepsilon\to 0}\sup\left\{\frac{\|\Delta x\|_2}{\varepsilon\|x_{LS}\|_2}:\|(\Delta A,\Delta B,\Delta b)\|_{\epsilon}\leq\|(A,B,b)\|_{\epsilon}\,,\Delta A\in\mathcal{S},\Delta B\in\mathcal{T}\right\}
$$

satisfies

$$
\kappa^{(\epsilon)}(A,B,b)\leq\sqrt{\frac{\|A^{\dagger}\|_2^2}{\alpha^2}+\frac{\|B^{\dagger}\|_2^2}{\beta^2}+\frac{\|A^{\dagger}\|_2^2\|B^{\dagger}\|_2^2}{\gamma^2\|x_{LS}\|_2^2}}\,\|(A,B,b)\|_{\epsilon}:=\kappa^{(\epsilon)}(A,B,b)^{upper}.
$$

Proof. By taking the 2-norm of (2.2) and using Lemma 1.1, we obtain

$$\begin{aligned}
& \left\|(A^\dagger \otimes B^\dagger)\Delta b - \left(A^\dagger \otimes (B^\dagger \Delta B B^\dagger)\right) b - \left((A^\dagger \Delta A A^\dagger) \otimes B^\dagger\right) b\right\|_2 \\
= & \left\|(A^\dagger \otimes B^\dagger)\Delta b - \left((A^\dagger A) \otimes (B^\dagger \Delta B) + (A^\dagger \Delta A) \otimes (B^\dagger B)\right) x_{LS}\right\|_2 \\
\leq & \|A^\dagger\|_2 \|B^\dagger\|_2 \|\Delta b\|_2 + \|\Delta A\|_2 \|A^\dagger\|_2 \|x_{LS}\|_2 + \|\Delta B\|_2 \|B^\dagger\|_2 \|x_{LS}\|_2 \\
= & \begin{bmatrix} \frac{\|A^\dagger\|_2 \|x_{LS}\|_2}{\alpha} & \frac{\|B^\dagger\|_2 \|x_{LS}\|_2}{\beta} & \frac{\|A^\dagger\|_2 \|B^\dagger\|_2}{\gamma} \end{bmatrix} \begin{bmatrix} \alpha \|\Delta A\|_2 \\ \beta \|\Delta B\|_2 \\ \gamma \|\Delta b\|_2 \end{bmatrix} \\
\leq & \varepsilon \sqrt{\frac{\|A^\dagger\|_2^2 \|x_{LS}\|_2^2}{\alpha^2} + \frac{\|B^\dagger\|_2^2 \|x_{LS}\|_2^2}{\beta^2} + \frac{\|A^\dagger\|_2^2 \|B^\dagger\|_2^2}{\gamma^2}}\, \|(A, B, b)\|_\in ,
\end{aligned}$$

which together with the definition of $\kappa^{(\in)}(A, B, b)$ complete the proof of this theorem. □

We now give the upper bound estimates of the normwise condition numbers for the LS problem (1.1) under the weighted Frobenius norm and weighted 2-norm when only A and b are perturbed. These results can be proved in the same manners as in Theorems 2.1 and 2.2, and hence, we state them in the following theorem without their proofs.

Theorem 2.3. *When A is of rank deficient, we have*

$$\begin{aligned}
\kappa^{(\mathcal{F})}(A, b) &= \lim_{\varepsilon \to 0} \sup \left\{ \frac{\|\Delta x\|_2}{\varepsilon \|x_{LS}\|_2} : \|(\Delta A, \Delta b)\|_{\mathcal{F}} \leq \varepsilon \|(A, b)\|_{\mathcal{F}}, \Delta A \in \mathcal{S} \right\} \\
&\leq \frac{\left\| \begin{bmatrix} \frac{P}{\alpha} & -\frac{A^\dagger \otimes B^\dagger}{\gamma} \end{bmatrix} \right\|_2 \|(A, b)\|_{\mathcal{F}}}{\|x_{LS}\|_2} := \kappa^{(\mathcal{F})}(A, b)^{upper},
\end{aligned}$$

and

$$\begin{aligned}
\kappa^{(\in)}(A, b) &= \lim_{\varepsilon \to 0} \sup \left\{ \frac{\|\Delta x\|_2}{\varepsilon \|x_{LS}\|_2} : \|(\Delta A, \Delta b)\|_\in \leq \|(A, b)\|_\in , \Delta A \in \mathcal{S} \right\} \\
&\leq \sqrt{\frac{\|A^\dagger\|_2^2}{\alpha^2} + \frac{\|A^\dagger\|_2^2 \|B^\dagger\|_2^2}{\gamma^2 \|x_{LS}\|_2^2}}\, \|(A, b)\|_\in := \kappa^{(\in)}(A, b)^{upper}.
\end{aligned}$$

where $P = (x_{LS}^T \otimes A^\dagger \otimes B^\dagger)(I_n \otimes K_{qm} \otimes I_p)(I_{mn} \otimes vec(B))$ *and*

$$\|(A, b)\|_{\mathcal{F}} = \sqrt{\alpha^2 \|A\|_F^2 + \gamma^2 \|b\|_2^2}, \quad \|(A, b)\|_\in = \sqrt{\alpha^2 \|A\|_2^2 + \gamma^2 \|b\|_2^2}.$$

When $B = 1$, the problem (1.1) reduces to the classical LS problem $\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$. When $B = 1$, it follows from Theorem 2.3 that

$$\kappa^{(\mathcal{F})}(A, b) \leq \frac{\left\| \begin{bmatrix} \frac{x_{LS}^T \otimes A^\dagger}{\alpha} & -\frac{A^\dagger}{\gamma} \end{bmatrix} \right\|_2 \|(A, b)\|_{\mathcal{F}}}{\|x_{LS}\|_2} = \frac{\left\|A^\dagger\right\|_2 \|(A, b)\|_{\mathcal{F}}}{\|x_{LS}\|_2} \sqrt{1 + \|x_{LS}\|_2^2}.$$

The upper bound estimate in the above inequality was proved to be the exact expression of $\kappa^{(\mathcal{F})}(A, b)$ in [21, Corollary 2.2]. Therefore, we conjecture that the upper bound estimates in Theorem 2.1, Theorem 2.2 and Theorem 2.3 are the exact expression of the corresponding normwise condition numbers. This needs to be further studied in the future.

3. 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 [20]. 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 [11]. We adopt their terminology and define the mixed and componentwise condition numbers for the LS problem (1.1) are defined as follows:

$$m(A,B,b)=\lim_{\varepsilon\to 0}\sup_{\substack{|\Delta A|\le\varepsilon|A|,\ |\Delta B|\le\varepsilon|B|\\ |\Delta b|\le\varepsilon|b|,\ \Delta A\in\mathcal{S},\ \Delta B\in\mathcal{T}}}\frac{\|\Delta x\|_\infty}{\varepsilon\|x_{LS}\|_\infty}$$

and

$$c(A,B,b)=\lim_{\varepsilon\to 0}\sup_{\substack{|\Delta A|\le\varepsilon|A|,\ |\Delta B|\le\varepsilon|B|\\ |\Delta b|\le\varepsilon|b|,\ \Delta A\in\mathcal{S},\ \Delta B\in\mathcal{T}}}\frac{1}{\varepsilon}\left\|\frac{\Delta x}{x_{LS}}\right\|_\infty .$$

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

The following theorem gives the upper bounds for the mixed and componentwise condition numbers of the LS problem (1.1).

Theorem 3.1. *When A and B are of rank deficient, we have*

$$\begin{aligned} m(A,B,b) &\le \frac{\left\||P|vec(|A|)+|Q|vec(|B|)+|A^\dagger\otimes B^\dagger||b|\right\|_\infty}{\|x_{LS}\|_\infty} := m(A,B,b)^{upper1}\\ &\le \frac{\left\|2\left((|A^\dagger||A|)\otimes(|B^\dagger||B|)\right)|x_{LS}|+|A^\dagger\otimes B^\dagger||b|\right\|_\infty}{\|x_{LS}\|_\infty} := m(A,B,b)^{upper2}\end{aligned}$$

and

$$\begin{aligned} c(A,B,b) &\le \left\|\frac{|P|vec(|A|)+|Q|vec(|B|)+|A^\dagger\otimes B^\dagger||b|}{|x_{LS}|}\right\|_\infty := c(A,B,b)^{upper1}\\ &\le \left\|\frac{2\left((|A^\dagger||A|)\otimes(|B^\dagger||B|)\right)|x_{LS}|+|A^\dagger\otimes B^\dagger||b|}{|x_{LS}|}\right\|_\infty := c(A,B,b)^{upper2},\end{aligned}$$

where $P=(x_{LS}^T\otimes A^\dagger\otimes B^\dagger)(I_n\otimes K_{qm}\otimes I_p)(I_{mn}\otimes vec(B))$ *and* $Q=(x_{LS}^T\otimes A^\dagger\otimes B^\dagger)(I_n\otimes K_{qm}\otimes I_p)(vec(A)\otimes I_{pq})$.

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

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

where $D_A=\mathrm{diag}(\mathrm{vec}(A))$. Similarly, we have $\mathrm{vec}(\Delta B)=D_B D_B^\dagger\mathrm{vec}(\Delta B)$ and $\Delta b=D_b D_b^\dagger\Delta b$ with $D_B=\mathrm{diag}(\mathrm{vec}(B))$ and $D_b=\mathrm{diag}(b)$. Thus the linear term $(A^\dagger\otimes B^\dagger)\Delta b-(A^\dagger\otimes B^\dagger)(A\otimes\Delta B+\Delta A\otimes B)\,x_{LS}$ of Δx can be rewritten as

$$(A^\dagger\otimes B^\dagger)\Delta b-(A^\dagger\otimes B^\dagger)(A\otimes\Delta B+\Delta A\otimes B)\,x_{LS}$$

$$
\begin{aligned}
&= -P\mathrm{vec}(\Delta A) - Q\mathrm{vec}(\Delta B) + (A^\dagger \otimes B^\dagger)\Delta b \\
&= -PD_A D_A^\dagger \mathrm{vec}(\Delta A) - QD_B D_B^\dagger \mathrm{vec}(\Delta B) + (A^\dagger \otimes B^\dagger)D_b D_b^\dagger \Delta b \\
&= \begin{bmatrix} -PD_A & -QD_B & (A^\dagger \otimes B^\dagger)D_b \end{bmatrix} \begin{bmatrix} D_A^\dagger \mathrm{vec}(\Delta A) \\ D_B^\dagger \mathrm{vec}(\Delta B) \\ D_b^\dagger \Delta b \end{bmatrix}. \qquad (3.1)
\end{aligned}
$$

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

$$
\begin{aligned}
&\|(A^\dagger \otimes B^\dagger)\Delta b - (A^\dagger \otimes B^\dagger)(A \otimes \Delta B + \Delta A \otimes B)\, x_{LS}\|_\infty \\
\leq\ & \varepsilon \left\| \left[-PD_A \quad -QD_B \quad (A^\dagger \otimes B^\dagger)D_b \right] \right\|_\infty .
\end{aligned}
$$

Since $(A^\dagger \otimes B^\dagger)\Delta b - (A^\dagger \otimes B^\dagger)(A \otimes \Delta B + \Delta A \otimes B)\, x_{LS}$ is the linear term of Δx, $m(A, B, b)$ is bounded by

$$
\begin{aligned}
m(A, B, b) &\leq \frac{\left\| \left[-PD_A \quad -QD_B \quad (A^\dagger \otimes B^\dagger)D_b \right] \right\|_\infty}{\|x_{LS}\|_\infty} \\
&= \frac{\left\| \left[|-PD_A| \quad |-QD_B| \quad |(A^\dagger \otimes B^\dagger)D_b| \right] e \right\|_\infty}{\|x_{LS}\|_\infty} \\
&= \frac{\left\| |P|\mathrm{vec}(|A|) + |Q|\mathrm{vec}(|B|) + |A^\dagger \otimes B^\dagger||b| \right\|_\infty}{\|x_{LS}\|_\infty},
\end{aligned}
$$

where e is an $mn + mp + pq$ dimensional vector with all entries equal to one.

Recall that in the definition of $c(A, B, b)$, we assume that x_{LS} has no zero entries. Hence, it follows from (3.1) and the assumption $|\Delta A| \leq \varepsilon|A|$, $|\Delta B| \leq \varepsilon|B|$ and $|\Delta b| \leq \varepsilon|b|$ that

$$
\begin{aligned}
&\left\| \frac{(A^\dagger \otimes B^\dagger)\Delta b - (A^\dagger \otimes B^\dagger)(A \otimes \Delta B + \Delta A \otimes B)\, x_{LS}}{x_{LS}} \right\|_\infty \\
=\ & \left\| D_{x_{LS}}^{-1} \left((A^\dagger \otimes B^\dagger)\Delta b - (A^\dagger \otimes B^\dagger)(A \otimes \Delta B + \Delta A \otimes B)\, x_{LS} \right) \right\|_\infty \\
\leq\ & \left\| D_{x_{LS}}^{-1} \left[-PD_A \quad -QD_B \quad (A^\dagger \otimes B^\dagger)D_b \right] \right\|_\infty \left\| \begin{bmatrix} D_A^\dagger \mathrm{vec}(\Delta A) \\ D_B^\dagger \mathrm{vec}(\Delta B) \\ D_b^\dagger \Delta b \end{bmatrix} \right\|_\infty \\
\leq\ & \varepsilon \left\| D_{x_{LS}}^{-1} \left[-PD_A \quad -QD_B \quad (A^\dagger \otimes B^\dagger)D_b \right] \right\|_\infty ,
\end{aligned}
$$

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

$$
\begin{aligned}
c(A, B, b) &\leq \left\| D_{x_{LS}}^{-1} \left[-PD_A \quad -QD_B \quad (A^\dagger \otimes B^\dagger)D_b \right] \right\|_\infty \\
&= \left\| \left| D_{x_{LS}}^{-1} \right| \left(|P|\mathrm{vec}(|A|) + |Q|\mathrm{vec}(|B|) + |A^\dagger \otimes B^\dagger||b| \right) \right\|_\infty \\
&= \left\| \frac{|P|\mathrm{vec}(|A|) + |Q|\mathrm{vec}(|B|) + |A^\dagger \otimes B^\dagger||b|}{|x_{LS}|} \right\|_\infty .
\end{aligned}
$$

Using Lemma 1.1, we can get

$$
|P|\mathrm{vec}(|A|) \leq (|x_{LS}^T| \otimes |A^\dagger| \otimes |B^\dagger|)(I_n \otimes K_{qm} \otimes I_p)(I_{mn} \otimes \mathrm{vec}(|B|))\mathrm{vec}\left((\mathrm{vec}(|A|))^T\right)
$$

$$\begin{aligned} &= (|x_{LS}^T| \otimes |A^\dagger| \otimes |B^\dagger|)(I_n \otimes K_{qm} \otimes I_p)\text{vec}\left(\text{vec}(|B|)(\text{vec}(|A|))^T\right) \\ &= (|x_{LS}^T| \otimes |A^\dagger| \otimes |B^\dagger|)(I_n \otimes K_{qm} \otimes I_p)(\text{vec}(|A|) \otimes \text{vec}(|B|)) \\ &= (|x_{LS}^T| \otimes |A^\dagger| \otimes |B^\dagger|)\text{vec}(|A| \otimes |B|) \\ &= (|A^\dagger| \otimes |B^\dagger|)(|A| \otimes |B|)|x_{LS}| = \left((|A^\dagger||A|) \otimes (|B^\dagger||B|)\right)|x_{LS}|. \end{aligned}$$

Similarly, we can deduce that

$$|Q|\text{vec}(|B|) \leq \left((|A^\dagger||A|) \otimes (|B^\dagger||B|)\right)|x_{LS}|.$$

Because of the monotonicity property of the infinity norm, the upper bounds $m(A, B, b)^{\text{upper2}}$ and $c(A, B, b)^{\text{upper2}}$ can be obtained by applying the aforementioned two inequalities to $m(A, B, b)^{\text{upper1}}$ and $c(A, B, b)^{\text{upper1}}$ and by using the matrix norm triangular inequality. □

The following theorem gives the upper bound estimates of the mixed and componentwise condition numbers for the LS problem (1.1) when only A and b are perturbed, which can be proved in the same way as Theorem 3.1.

Theorem 3.2. *When A is of rank deficient, we have*

$$\begin{aligned} m(A, b) &= \lim_{\varepsilon \to 0} \sup_{\substack{|\Delta A| \leq \varepsilon |A|, |\Delta b| \leq \varepsilon |b| \\ \Delta A \in \mathcal{S}}} \frac{\|\Delta x\|_\infty}{\varepsilon \|x_{LS}\|_\infty} \\ &\leq \frac{\left\| |P| vec(|A|) + |A^\dagger \otimes B^\dagger||b| \right\|_\infty}{\|x_{LS}\|_\infty} \leq \frac{\left\| \left((|A^\dagger||A|) \otimes (|B^\dagger||B|)\right)|x_{LS}| + |A^\dagger \otimes B^\dagger||b| \right\|_\infty}{\|x_{LS}\|_\infty} \end{aligned}$$

and

$$\begin{aligned} c(A, b) &= \lim_{\varepsilon \to 0} \sup_{\substack{|\Delta A| \leq \varepsilon |A|, \ |\Delta b| \leq \varepsilon |b| \\ \Delta A \in \mathcal{S}}} \frac{1}{\varepsilon} \left\| \frac{\Delta x}{x_{LS}} \right\|_\infty \\ &\leq \left\| \frac{|P| vec(|A|) + |A^\dagger \otimes B^\dagger||b|}{|x_{LS}|} \right\|_\infty \leq \left\| \frac{\left((|A^\dagger||A|) \otimes (|B^\dagger||B|)\right)|x_{LS}| + |A^\dagger \otimes B^\dagger||b|}{|x_{LS}|} \right\|_\infty, \end{aligned}$$

where $P = (x_{LS}^T \otimes A^\dagger \otimes B^\dagger)(I_n \otimes K_{qm} \otimes I_p)(I_{mn} \otimes vec(B))$.

When $B = 1$, it follows from Theorem 3.2 that

$$m(A, b) \leq \frac{\left\| |x_{LS}^T \otimes A^\dagger| \text{vec}(|A|) + |A^\dagger||b| \right\|_\infty}{\|x_{LS}\|_\infty} = \frac{\left\| |A^\dagger||A||x_{LS}| + |A^\dagger||b| \right\|_\infty}{\|x_{LS}\|_\infty}$$

and

$$c(A, b) \leq \left\| \frac{|x_{LS}^T \otimes A^\dagger| \text{vec}(|A|) + |A^\dagger||b|}{|x_{LS}|} \right\|_\infty = \left\| \frac{|A^\dagger||A||x_{LS}| + |A^\dagger||b|}{|x_{LS}|} \right\|_\infty,$$

which have been obtained in [16].

4. Numerical experiments

We consider the LS problem (1.1) with

$$A = \begin{bmatrix} 9\times 10^i & 0 & 0 \\ 0 & 2 & 2 \\ 3 & 0 & 0 \end{bmatrix},\ B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 3 & 0 & 0 \end{bmatrix},\ b = \begin{bmatrix} 100 \\ 100 \\ \vdots \\ 100 \end{bmatrix} \in \mathbb{R}^9,\ i = 0, 2, 4.$$

We first compare $\kappa^{(\mathcal{F})}(A,B,b)^{\text{upper1}}$, $\kappa^{(\mathcal{F})}(A,B,b)^{\text{upper2}}$, $\kappa^{(\epsilon)}(A,B,b)^{\text{upper2}}$ with the upper bounds of the mixed and componentwise condition numbers given in Theorem 3.1. 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 upper bounds of the normwise condition numbers become larger and larger, while, comparatively, the upper bounds of 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.

Table 1. Comparison of condition numbers.

	$i=0$	$i=2$	$i=4$
$\kappa^{(\mathcal{F})}(A,B,b)^{\text{upper1}}$	1.4398×10^2	4.6210×10^2	4.3838×10^4
$\kappa^{(\mathcal{F})}(A,B,b)^{\text{upper2}}$	5.3253×10^2	1.0724×10^5	1.0126×10^9
$\kappa^{(\epsilon)}(A,B,b)^{\text{upper}}$	3.0347×10^3	8.6573×10^5	8.2129×10^9
$m(A,B,b)^{\text{upper1}}$	3.0000	3.0000	3.0000
$m(A,B,b)^{\text{upper2}}$	3.0000	3.0000	3.0000
$c(A,B,b)^{\text{upper1}}$	3.0000	3.0000	3.0000
$c(A,B,b)^{\text{upper2}}$	3.0000	3.0000	3.0000

For $i=0$, suppose the perturbations are $\Delta A = 10^{-j}\times A$, $\Delta B = 10^{-j}\times B$ and $\Delta B = 10^{-j}\times \text{rand}(9,1)$, where rand$(\cdot)$ is the MATLAB function. Obviously, $\Delta A \in \mathcal{S}$ and $\Delta B \in \mathcal{T}$. Define $\varepsilon_1 = \frac{\|(\Delta A,\Delta B,\Delta b)\|_{\mathcal{F}}}{\|(A,B,b)\|_{\mathcal{F}}}$, $\varepsilon_2 = \frac{\|(\Delta A,\Delta B,\Delta b)\|_{\epsilon}}{\|(A,B,b)\|_{\epsilon}}$ and $\varepsilon_3 = \max\{\varepsilon : |\Delta A| \le \varepsilon|A|,\ |\Delta B| \le \varepsilon|B|\ |\Delta b| \le \varepsilon|b|\}$, it follows from the definitions of $\kappa^{(\mathcal{F})}(A,B,b)$, $\kappa^{(\epsilon)}(A,B,b)$, $m(A,B,b)$ and $c(A,B,b)$ that

$$\frac{\|\Delta x\|_2}{\|x_{LS}\|_2} \le \varepsilon_1\kappa^{(\mathcal{F})}(A,B,b)^{\text{upper1}} + O(\varepsilon_1^2),\ \frac{\|\Delta x\|_2}{\|x_{LS}\|_2} \le \varepsilon_2\kappa^{(\epsilon)}(A,B,b)^{\text{upper}} + O(\varepsilon_2^2)$$

and

$$\frac{\|\Delta x\|_\infty}{\|x_{LS}\|_\infty} \le \varepsilon_3 m(A,B,b)^{\text{upper1}} + O(\varepsilon_3^2),\ \left\|\frac{\Delta x}{x_{LS}}\right\|_\infty \le \varepsilon_3 c(A,B,b)^{\text{upper1}} + O(\varepsilon_3^2)$$

for small ε_1, ε_2 and ε_3. As shown in Table 2, the error bounds given by the upper bounds of the condition numbers are at most two order of magnitude larger than the actual errors. This illustrates that, the estimates $\kappa^{(\mathcal{F})}(A, B, b)^{\text{upper1}}$, $\kappa^{(\epsilon)}(A, B, b)^{\text{upper}}$, $m(A, B, b)^{\text{upper1}}$ and $c(A, B, b)^{\text{upper1}}$ can estimate their corresponding condition numbers well.

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

j	10	11	12
$\|\|\Delta x\|\|_2/\|\|x_{LS}\|\|_2$	1.9936×10^{-10}	1.9933×10^{-11}	1.9961×10^{-12}
$\varepsilon_1\kappa^{(\mathcal{F})}(A, B, b)^{\text{upper1}}$	5.2458×10^{-10}	5.2301×10^{-11}	5.2224×10^{-12}
$\varepsilon_2\kappa^{(\epsilon)}(A, B, b)^{\text{upper}}$	1.0291×10^{-8}	1.0256×10^{-9}	1.0238×10^{-10}
$\|\|\Delta x\|\|_\infty/\|\|x_{LS}\|\|_\infty$	1.9936×10^{-10}	1.9941×10^{-11}	1.9956×10^{-12}
$\varepsilon_3 m(A, B, b)^{\text{upper1}}$	3.0000×10^{-10}	3.0000×10^{-11}	3.0000×10^{-12}
$\left\|\frac{\Delta x}{x_{LS}}\right\|_\infty$	1.9963×10^{-10}	1.9941×10^{-11}	1.9971×10^{-12}
$\varepsilon_3 c(A, B, b)^{\text{upper1}}$	3.0000×10^{-10}	3.0000×10^{-11}	3.0000×10^{-12}

Acknowledgments

This work was supported by the Foundation for Distinguished Young Scholars of Gansu Province (Grant No.20JR5RA540).

Conflict of interest

The authors declare there is no conflict of interest.

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. A. Barrlund, Efficient solution of constrained least squares problems with Kronecker product structure, *SIAM J. Matrix Anal. Appl.,* **19** (1994), 154–160.
3. A. Ben-Israel, On error bounds for generalized inverses, *SIAM J. Numer. Anal.,* **3** (1966), 585–592.
4. E. H. Bergou, S. Gratton, J. Tshimanga, The exact condition number of the truncated singular value solution of a linear ill-posed problem, *SIAM J. Matrix Anal. Appl.,* **35** (2014), 1073–1085.

5. T. Chen, W. Li, On condition numbers for the weighted Moore-Penrose inverse and the weighted least squares problem involving Kronecker products, *East Asian J. Appl. Math.,* **4** (2014), 1–20.
6. D. Chu, L. Lin, R. C. E. Tan, Y. Wei, Condition numbers and perturbation analysis for the Tikhonov regularization of discrete ill-posed problems, *Numer. Linear Algebra Appl.,* **18** (2011), 87–103.
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. D. W. Fausett, C. T. Fulton, Large least squares problems involving Kronecker products, *SIAM J. Matrix Anal. Appl.,* **15** (1994), 219–227.
9. D. W. Fausett, C. T. Fulton, H. Hashish, Improved parallel QR method for large least squares problems involving Kronecker products, *J. Comput. Appl. Math.,* **78** (1997), 63–78.
10. C. T. Fulton, L. Wu, Parallel algorithms for large least squares problems involving Kronecker products, *Nonlinear Anal. Theory Methods Appl.,* **30** (1997), 5033–5040.
11. I. Gohberg, I. Koltracht, Mixed, componentwise, and structured condition numbers, *SIAM J. Matrix Anal. Appl.,* **14** (1993), 688–704.
12. G. H. Golub, C. F. Van Loan, *Matrix computations,* 4th ed., Johns Hopkins University Press, Baltimore, 2013.
13. A. Graham, *Kronecker products and matrix calculus with application,* Wiley, New York, 1981.
14. S. Gratton, On the condition number of linear least squares problems in a weighted Frobenius norm, *BIT Numer. Math.,* **36** (1996), 523–530.
15. N. J. Higham, *Accuracy and stability of numerical algorithms,* 2nd ed., SIAM, Philadelphia, 2002.
16. W. Kang, H. Xiang, Condition numbers with their condition numbers for the weighted Moore-Penrose inverse and the weighted least squares solution, *J. Appl. Math. Comput.,* **22** (2006), 95–112.
17. W. Kang, H. Xiang, Level-2 condition numbers for least-squares solution of Kronecker product linear systems, *Int. J. Comput. Math.,* **85** (2008), 827–841.
18. A. Marco, J. J. Martínez, R. Viaña, Least squares problems involving generalized Kronecker products and application to bivariate polynomial regression, *Numer. Algorithms,* **82** (2019), 21–39.
19. P. A. Regalia, S. K. Mitra, Kronecker products, unitary matrices and signal processing applications, *SIAM Rev.,* **31** (1989), 586–613.
20. J. Rohn, New condition numbers for matrices and linear systems, *Computing,* **41** (1989), 167–169.
21. Y. Wei, H. Diao, S. Qiao, Condition number for weighted linear least squares problem, *J. Comput. Math.,* **25** (2007), 561–572.
22. H. Xiang, H. Diao, Y. Wei, On perturbation bounds of Kronecker product linear systems and their level-2 condition numbers, *J. Comput. Appl. Math.,* **183** (2005), 210–231.
23. 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.

AIMS Press