Condition Numbers for a Linear Function of the Solution to the Constrained and Weighted Least Squares Problem and Their Statistical Estimation

Mahvish Samar

Abstract. In this paper, we consider the condition number theory for a linear function of the solution to the constrained and weighted least squares problem. We first present two explicit expressions without Kronecker product of normwise condition number using the classical method for condition numbers. Then, we derive the explicit expression of mixed and componentwise condition numbers by the dual techniques. To estimate these condition numbers with high reliability, we choose the probabilistic spectral norm estimator and the small-sample statistical condition estimation method and devise three algorithms. Numerical experiments are provided to illustrate the obtained results.

1. Introduction

The constrained and weighted linear least squares (CWLS) problem can be stated as follows:

$$(1.1) \qquad \min_{x \in \mathbb{R}^n} (b_2 - A_2 x)^T M_2^{-1} (b_2 - A_2 x) \quad \text{such that } A_1 x = b_1,$$

where $A_1 \in \mathbb{R}^{p \times n}$, $A_2 \in \mathbb{R}^{q \times n}$, $b_1 \in \mathbb{R}^p$, $b_2 \in \mathbb{R}^q$ and $M_2 \in \mathbb{R}^{q \times q}$ is symmetric positive definite. Let

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

with $p + q = m$, and assume $m \geq n \geq p$. Then, from [5, 22], we have that the problem (1.1) has a unique solution if and only if $\text{rank}(A_1) = p$ and $\text{rank}(A) = n$, and the solution x can be stated as follows [20, Lemma 3.1, Remark 3.1]:

(1.2) $$x = (A^{T^\dagger}_{M^{1/2}})^T b,$$

where

(1.3) $$(A^{T^\dagger}_{M^{1/2}})^T = A^\dagger[I - M(PMP)^\dagger]$$

with $M = \text{diag}(0_p, M_2) \in \mathbb{R}^{m \times m}$, $P = I - AA^\dagger$, and $A^\dagger$ denoting the Moore-Penrose inverse of A.

The CWLS problem (1.1) arises in many practical applications of equilibrium systems including optimization, finite elements, structural analysis, and electrical network sand discretization of Stokes flow (see [27, 36]), and hence has been extensively studied in the literature. Some authors investigated its numerical algorithms, stability of algorithms, and perturbation analysis (see e.g., [17–23, 28, 38]). Here, we only introduce some works on perturbation analysis. Wedin [38] first investigated the perturbation theory for CWLS problem. The obtained results were extended by Gulliksson et al. [22]. Based on the perturbation bounds, Gulliksson [17, 18] analyzed the convergence behavior of the iterative refinement procedure and presented the explicit forward error bounds of CWLS problem. In addition, using the general form of the augmented system, Gulliksson et al. [21] also presented the perturbation identities of CWLS problem for both the full-rank and rank-deficient cases. Later, Gulliksson et al. [20] revisited the perturbation theory of CWLS problem and provided the expression of solution in generalized inverse form, i.e., (1.2).

As we know, condition number plays an important role in perturbation theory and error analysis for algorithm, and has been extensively studied (see e.g., [6, 7, 15, 33]). Recently, some scholars [1–3, 8–10, 14, 30, 31] considered the condition numbers of a linear function of the solution to the least squares problem, the total least squares problem, the weighted least squares problem, the indefinite least squares problem and the equality constraint least squares problem. However, to our best knowledge, there is no work on such condition numbers of CWLS problem so far. In this paper, we will consider this problem.

For normwise condition number, two explicit expressions are given by the classical method. For mixed and componentwise condition number, we use the dual techniques. This technique was proposed by Baboulin and Gratton [2] for the mixed and component-wise condition numbers of least squares problem. Later, Diao et al. applied this technique to weighted least squares problem [9] and the total least squares problem [10]. In addition, we consider the statistical estimation of these condition numbers by the probabilistic spectral norm estimator [25] and the small-sample statistical condition estimation (SSCE) method [29]. Three related algorithms are presented. The above results comprise Sections 3, 4, 5, and 6 of this paper, respectively.

2. Preliminaries

In this section, we will discuss some results on the 'vec' operator, Kronecker product, the dual techniques, and the condition numbers.

2.1. Vec operator and Kronecker product

The operator 'vec' is defined as

$$\text{vec}(A) = \left[a_1^T, \ldots, a_n^T\right]^T \in \mathbb{R}^{mn},$$

for $A = [a_1, \ldots, a_n] \in \mathbb{R}^{m \times n}$ with $a_i \in \mathbb{R}^m$ and the Kronecker product between $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$ is defined as $A \otimes B = [a_{ij}B] \in \mathbb{R}^{mp \times nq}$. Some useful results on these two tools are introduced as follows [16, 26]:

$$(2.1) \qquad \qquad \text{vec}(AXB) = (B^T \otimes A)\,\text{vec}(X),$$

$$(2.2) \qquad \qquad \text{vec}(A^T) = \Pi\,\text{vec}(A),$$

$$\|A \otimes B\|_2 = \|A\|_2\|B\|_2,$$

$$(2.3) \qquad \qquad (A \otimes B)^T = (A^T \otimes B^T),$$

$$(2.4) \qquad \qquad (A \otimes B)(C \otimes D) = (AC) \otimes (BD),$$

$$(2.5) \qquad \qquad (y^T \otimes Y)\Pi = (Y \otimes y^T),$$

where X, Y, C and D are matrices of suitable orders, y is a vector and $\Pi \in \mathbb{R}^{mn \times mn}$ is the vec-permutation matrix; see [16] for details.

2.2. Dual techniques and condition numbers

For the Euclidean spaces $\mathbb{S}$ and $\mathbb{Q}$ equipped with the scalar products $\langle \cdot, \cdot \rangle_{\mathbb{S}}$ and $\langle \cdot, \cdot \rangle_{\mathbb{Q}}$, respectively, we denote the corresponding norms as $\|\cdot\|_{\mathbb{S}}$ and $\|\cdot\|_{\mathbb{Q}}$, and let a linear operator $J \colon \mathbb{S} \to \mathbb{Q}$ be well-defined. Thus, the dual norm and the adjoint operator can be defined as follows.

Definition 2.1. The dual norm $\|\cdot\|_{\mathbb{S}^*}$ of the norm $\|\cdot\|_{\mathbb{S}}$ is defined by

$$\|a\|_{\mathbb{S}^*} = \max_{w \neq 0} \frac{\langle a, w \rangle_{\mathbb{S}}}{\|w\|_{\mathbb{S}}},$$

where $a, w \in \mathbb{S}$.

Definition 2.2. The adjoint operator of J, $J^* \colon \mathbb{Q} \to \mathbb{S}$ is defined by

$$\langle y, Ja \rangle_{\mathbb{Q}} = \langle J^* y, a \rangle_{\mathbb{S}},$$

where $a \in \mathbb{S}$ and $y \in \mathbb{Q}$.

For commonly used vector norms and the Frobenius norm, their dual norms are given by

$$\| \cdot \|_{1^*} = \| \cdot \|_\infty, \quad \| \cdot \|_{\infty^*} = \| \cdot \|_1, \quad \| \cdot \|_{2^*} = \| \cdot \|_2, \quad \|A\|_{F^*} = \|A\|_F.$$

For the linear operator J from $\mathbb{S}$ to $\mathbb{Q}$, let $\| \cdot \|_{\mathbb{S},\mathbb{Q}}$ be the operator norm induced by the norms $\| \cdot \|_{\mathbb{S}}$ and $\| \cdot \|_{\mathbb{Q}}$, and for the linear operator from $\mathbb{Q}$ to $\mathbb{S}$, let $\| \cdot \|_{\mathbb{Q}^*,\mathbb{S}^*}$ be the operator norm induced by the dual norms $\| \cdot \|_{\mathbb{S}^*}$ and $\| \cdot \|_{\mathbb{Q}^*}$. We have the following result [2] on these two norms.

Lemma 2.3. *For the linear operator J from $\mathbb{S}$ to $\mathbb{Q}$, we have*

$$\|J\|_{\mathbb{S},\mathbb{Q}} = \|J^*\|_{\mathbb{Q}^*,\mathbb{S}^*}.$$

As noted in [2], it may additionally be desirable to compute $\|J^*\|_{\mathbb{Q}^*,\mathbb{S}^*}$ in place of $\|J\|_{\mathbb{S},\mathbb{Q}}$ when the dimension of the Euclidean space $\mathbb{Q}^*$ is lower than $\mathbb{S}$.

From [33], if ϕ is Fréchet differentiable in the neighborhood of $a \in \mathbb{S}$, then the absolute condition number of ϕ at $a \in \mathbb{S}$ is given by

$$(2.6) \qquad \kappa = \|\mathrm{d}\phi(a)\|_{\mathbb{S},\mathbb{Q}} = \max_{\|z\|_{\mathbb{S}}=1} \|\mathrm{d}\phi(a) \cdot z\|_{\mathbb{Q}},$$

where $\mathrm{d}\phi(a)$ denotes the Fréchet differential of ϕ at a. In view of Lemma 2.3, the following expression of κ can be obtained

$$(2.7) \qquad \kappa = \max_{\|\mathrm{d}a\|_S=1} \|\mathrm{d}\phi(a) \cdot \mathrm{d}a\|_Q = \max_{\|z\|_{\mathbb{Q}^*}=1} \|\mathrm{d}\phi(a)^* \cdot z\|_{\mathbb{S}^*}.$$

If ϕ is nonzero, we have the relative condition number

$$(2.8) \qquad \kappa_n = \kappa \frac{\|a\|_{\mathbb{S}}}{\|\phi(a)\|_{\mathbb{Q}}}.$$

Now, we consider the componentwise metric on a data space $\mathbb{S} = \mathbb{R}^n$. For any input data $a \in \mathbb{R}^n$, we denote by $\mathbb{S}_a$ the subset of all elements $\mathrm{d}a \in \mathbb{R}^n$ satisfying that $\mathrm{d}a_i = 0$ whenever $a_i = 0$, $1 \leq i \leq n$. Thus, we can measure the perturbation $\mathrm{d}a \in \mathbb{S}_a$ of a using the following componentwise norm with respect to a:

$$\|\mathrm{d}a\|_c = \min\{w, |\mathrm{d}a_i| \leq w|a_i|, i = 1, \ldots, n\}.$$

Equivalently, the componentwise norm has the following property

$$(2.9) \qquad \|\mathrm{d}a\|_c = \max\left\{ \frac{|\mathrm{d}a_i|}{|a_i|}, a_i \neq 0 \right\} = \left\| \left(\frac{|\mathrm{d}a_i|}{|a_i|} \right) \right\|_\infty.$$

For this norm, from [10, Equation 2.16], we have the following explicit expression of its dual norm

$$(2.10) \qquad \|\mathrm{d}(a)\|_{c^*} = \|(\|\mathrm{d}a_1\|_{\mathbb{S}^*}, \ldots, \|\mathrm{d}a_n\|_{\mathbb{S}^*})\|_\infty = \|(|\mathrm{d}a_1||a_1|, \ldots, |\mathrm{d}a_n||a_n|)\|_1.$$

Using the above componentwise norm, we can rewrite the condition number κ.

Lemma 2.4. [2, 9] *Using the above notations and the componentwise norm defined in* (2.9)*, the condition number κ can be expressed by*

$$\kappa = \max_{\|z\|_{\mathbb{Q}^*}=1} \|(\mathrm{d}\phi(a))^* \cdot z\|_{c^*},$$

where $\|\cdot\|_{c^*}$ is given by (2.10)*.*

Consider the following linear function ϕ of the CWLS solution:

(2.11) $\phi \colon \mathbb{R}^{m \times n} \times \mathbb{R}^{m \times m} \times \mathbb{R}^m \to \mathbb{R}^k, \quad (A, M, b) \to \phi(A, M, b) = L^T (A_{M^{1/2}}^{T\dagger})^T b,$

where $L \in \mathbb{R}^{n \times k}$ with $k \leq n$. Then, from [20, Equation 3.6], we know that ϕ is continuously Fréchet differentiable in a neighborhood of (A, M, b). For $\mathrm{d}A \in \mathbb{R}^{m \times n}$, $\mathrm{d}M \in \mathbb{R}^{m \times m}$ and $\mathrm{d}b \in \mathbb{R}^m$, using the chain rules of composition of derivatives, we have

(2.12)
$$\mathrm{d}\phi(A, M, b) \cdot (\mathrm{d}A, \mathrm{d}M, \mathrm{d}b)$$
$$= L^T (A_{M^{1/2}}^{T\dagger})^T M (A_{M^{1/2}}^{T\dagger}) \mathrm{d}A^T r - L^T (A_{M^{1/2}}^{T\dagger})^T \mathrm{d}Ax - L^T (A_{M^{1/2}}^{T\dagger})^T \mathrm{d}Mr + L^T (A_{M^{1/2}}^{T\dagger})^T \mathrm{d}b,$$

where $r = (PMP)^\dagger b$. Applying the operator vec to (2.12) and using (2.1), we have

(2.13)
$$\mathrm{d}\phi(A, M, b) \cdot (\mathrm{d}A, \mathrm{d}M, \mathrm{d}b)$$
$$= \mathrm{vec}(\mathrm{d}\phi(A, M, b) \cdot (\mathrm{d}A, \mathrm{d}M, \mathrm{d}b))$$
$$= \left[(r^T \otimes (L^T (A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger})) \Pi - x^T \otimes (L^T (A_{M^{1/2}}^{T\dagger})^T), -r^T \otimes (L^T (A_{M^{1/2}}^{T\dagger})^T), (L^T (A_{M^{1/2}}^{T\dagger})^T) \right]$$
$$\times \begin{bmatrix} \mathrm{vec}(\mathrm{d}A) \\ \mathrm{vec}(\mathrm{d}M) \\ \mathrm{d}b \end{bmatrix}$$
$$= W \begin{bmatrix} \mathrm{vec}(\mathrm{d}A) \\ \mathrm{vec}(\mathrm{d}M) \\ \mathrm{d}b \end{bmatrix},$$

where

$$W = \left[r^T \otimes (L^T (A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger}) \Pi - x^T \otimes (L^T (A_{M^{1/2}}^{T\dagger})^T), -r^T \otimes (L^T (A_{M^{1/2}}^{T\dagger})^T), (L^T (A_{M^{1/2}}^{T\dagger})^T) \right].$$

Combining (2.13) with (2.6), (2.8), and Lemma 2.4, we can get the normwise or mixed and componentwise condition numbers. Their explicit expressions will be discussed in Sections 3 and 4, respectively.

3. Normwise condition number

In the following theorem, we present two explicit expressions of normwise condition number, which doesn't contain the Kronecker product.

Theorem 3.1. *The normwise condition number for a linear function ϕ of CWLS solution defined in* (2.11) *has the following two explicit expressions:*
(3.1)
$$\kappa_{\phi 1} = \frac{\left\| L^T (A_{M^{1/2}}^{T\dagger})^T \left(M A_{M^{1/2}}^{T\dagger} (A_{M^{1/2}}^{T\dagger})^T M \| r \|_2^2 + (\| r \|_2^2 + \| x \|_2^2 + 1) \right) A_{M^{1/2}}^{T\dagger} L \right\|_2^{1/2} \| [A, M, b] \|_F}{\| L^T x \|_2}$$

and
(3.2)
$$\kappa_{\phi 2} = \frac{\left\| L^T (A_{M^{1/2}}^{T\dagger})^T \left[M A_{M^{1/2}}^{T\dagger} \| r \|_2, \| x \|_2 \left(I_m - \frac{1}{\| r \|_2^2} r r^T \right), \| r \|_2 \left(I_m - \frac{1}{\| r \|_2^2} r r^T \right), I_m \right] \right\|_2 \| [A, M, b] \|_F}{\| L^T x \|_2}.$$

Proof. By (2.8) and (2.13), it is easy to see that the common normwise condition number of ϕ has the following expression

$$(3.3) \qquad\qquad \kappa_n = \frac{\| W \|_2 \| [A, M, b] \|_F}{\| L^T x \|_2}.$$

Note that

$$(3.4) \qquad\qquad \| W \|_2 = \| W W^T \|_2^{1/2} = \| W_1 W_1^T + W_2 W_2^T + W_3 W_3^T \|_2^{1/2},$$

where

$$W_1 = (r^T \otimes (L^T (A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger})) \Pi - x^T \otimes (L^T (A_{M^{1/2}}^{T\dagger})^T),$$
$$W_2 = -r^T \otimes (L^T (A_{M^{1/2}}^{T\dagger})^T) \quad \text{and} \quad W_3 = L^T (A_{M^{1/2}}^{T\dagger})^T.$$

In the following, we compute $W_1 W_1^T$, $W_2 W_2^T$ and $W_3 W_3^T$, respectively, to remove the Kronecker product from (3.3).

Firstly, let

$$W_{11} = (r^T \otimes (L^T (A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger})) \Pi = L^T (A_{M^{1/2}}^{T\dagger})^T ((M A_{M^{1/2}}^{T\dagger}) \otimes r^T),$$
$$W_{12} = -(x^T \otimes (L^T (A_{M^{1/2}}^{T\dagger})^T)) = -L^T (A_{M^{1/2}}^{T\dagger})^T (x^T \otimes I_m),$$

where we have used the result (2.5). Then

$$(3.5) \qquad\qquad W_1 W_1^T = W_{11} W_{11}^T + W_{12} W_{12}^T + W_{11} W_{12}^T + W_{12} W_{11}^T.$$

By (2.3) and (2.4), we get

$$(3.6) \quad \begin{aligned} W_{11} W_{11}^T &= L^T (A_{M^{1/2}}^{T\dagger})^T \left((M A_{M^{1/2}}^{T\dagger}) \otimes r^T \right) \left(((A_{M^{1/2}}^{T\dagger})^T M) \otimes r \right) (A_{M^{1/2}}^{T\dagger} L) \\ &= L^T (A_{M^{1/2}}^{T\dagger})^T \left(M A_{M^{1/2}}^{T\dagger} (A_{M^{1/2}}^{T\dagger})^T M \right) \| r \|_2^2 A_{M^{1/2}}^{T\dagger} L, \end{aligned}$$

$$(3.7) \quad W_{12} W_{12}^T = L^T (A_{M^{1/2}}^{T\dagger})^T (x^T \otimes I_m)(x \otimes I_m)(A_{M^{1/2}}^{T\dagger}) L = L^T (A_{M^{1/2}}^{T\dagger})^T \| x \|_2^2 A_{M^{1/2}}^{T\dagger} L.$$

Considering (1.3) and the fact

$$(3.8) \qquad\qquad (PMP)^\dagger = (PMP)^\dagger M (PMP)^\dagger,$$

which is from [20], we have

$$(3.9) \qquad
\begin{aligned}
r^T A_{M^{1/2}}^{T\dagger} L &= \big(b^T (PMP)^{\dagger^T} A_{M^{1/2}}^{T\dagger} L\big) \\
&= \big(L^T (A_{M^{1/2}}^{T\dagger})^T (PMP)^\dagger b\big)^T \\
&= \big(L^T A^\dagger (I - M(PMP)^\dagger)(PMP)^\dagger b\big)^T \quad \text{(by (1.3))} \\
&= \big(L^T A^\dagger (PMP)^\dagger b - L^T A^\dagger (PMP)^\dagger M (PMP)^\dagger b\big)^T \\
&= \big(L^T A^\dagger (PMP)^\dagger b - L^T A^\dagger (PMP)^\dagger b\big)^T = 0 \quad \text{(by (3.8))}.
\end{aligned}$$

Thus, by (2.3), (2.4) and (3.9), we have

$$(3.10) \qquad
\begin{aligned}
W_{11} W_{12}^T &= -L^T (A_{M^{1/2}}^{T\dagger})^T \big((M A_{M^{1/2}}^{T\dagger}) \otimes r^T\big)(x \otimes I_m)(A_{M^{1/2}}^{T\dagger}) L \quad \text{(by (2.3))} \\
&= -\big(L^T (A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger} x\big) \otimes (r^T A_{M^{1/2}}^{T\dagger} L) \quad \text{(by (2.4))} \\
&= 0 = (W_{12} W_{11}^T)^T \quad \text{(by (3.9))}.
\end{aligned}$$

Substituting (3.6), (3.7), (3.10) into (3.5) implies
$$(3.11)$$
$$W_1 W_1^T = L^T (A_{M^{1/2}}^{T\dagger})^T \big(M A_{M^{1/2}}^{T\dagger} (A_{M^{1/2}}^{T\dagger})^T M\big) \|r\|_2^2 A_{M^{1/2}}^{T\dagger} L + L^T (A_{M^{1/2}}^{T\dagger})^T \|x\|_2^2 A_{M^{1/2}}^{T\dagger} L.$$

Furthermore, using (2.3) and (2.4), we obtain

$$(3.12) \qquad W_2 W_2^T = L^T (A_{M^{1/2}}^{T\dagger})^T (r^T \otimes I_m)(r \otimes I_m) A_{M^{1/2}}^{T\dagger} L = L^T (A_{M^{1/2}}^{T\dagger})^T \|r\|_2^2 A_{M^{1/2}}^{T\dagger} L$$

and

$$(3.13) \qquad\qquad W_3 W_3^T = L^T (A_{M^{1/2}}^{T\dagger})^T A_{M^{1/2}}^{T\dagger} L.$$

Thus, putting (3.11), (3.12) and (3.13) together give

$$(3.14) \qquad W W^T = L^T (A_{M^{1/2}}^{T\dagger})^T \big(M A_{M^{1/2}}^{T\dagger} (A_{M^{1/2}}^{T\dagger})^T M \|r\|_2^2 + (\|r\|_2^2 + \|x\|_2^2 + 1)\big) A_{M^{1/2}}^{T\dagger} L,$$

which combining (3.3) and (3.4) leads to the expression (3.1).

Note that (3.14) can be rewritten as

$$(3.15)\ W W^T = L^T (A_{M^{1/2}}^{T\dagger})^T \begin{bmatrix} M A_{M^{1/2}}^{T\dagger} (A_{M^{1/2}}^{T\dagger})^T M \|r\|_2^2 & 0_{n \times m} \\ 0_{m \times n} & (\|x\|_2^2 + \|r\|_2^2 + 1) I_m \end{bmatrix} A_{M^{1/2}}^{T\dagger} L.$$

Further, note that

(3.16)
$$
\begin{bmatrix} MA_{M^{1/2}}^{T\dagger}(A_{M^{1/2}}^{T\dagger})^T M \|r\|_2^2 & 0_{n\times m} \\ 0_{m\times n} & (\|x\|_2^2 + \|r\|_2^2 + 1)I_m \end{bmatrix}
$$
$$
= \begin{bmatrix} MA_{M^{1/2}}^{T\dagger} & 0_{n\times m} \\ 0_{m\times n} & I_m \end{bmatrix} \begin{bmatrix} \|r\|_2^2 & 0_{n\times m} \\ 0_{m\times n} & \|x\|_2^2\left(I_m - \frac{1}{\|r\|_2^2}rr^T\right) + \|r\|_2^2\left(I_m - \frac{1}{\|r\|_2^2}rr^T\right) + I_m \end{bmatrix}
$$
$$
\times \begin{bmatrix} (A_{M^{1/2}}^{T\dagger})^T M & 0_{n\times m} \\ 0_{m\times n} & I_m \end{bmatrix}
$$

and

$$
\|x\|_2^2\left(I_m - \frac{1}{\|r\|_2^2}rr^T\right) + \|r\|_2^2\left(I_m - \frac{1}{\|r\|_2^2}rr^T\right) + I_m
$$

(3.17)
$$
= \begin{bmatrix} \|x\|_2\left(I_m - \frac{1}{\|r\|_2^2}rr^T\right) & \|r\|_2\left(I_m - \frac{1}{\|r\|_2^2}rr^T\right) & I_m \end{bmatrix} \begin{bmatrix} \|x\|_2\left(I_m - \frac{1}{\|r\|_2^2}rr^T\right) \\ \|r\|_2\left(I_m - \frac{1}{\|r\|_2^2}rr^T\right) \\ I_m \end{bmatrix}.
$$

Thus, substituting (3.16) and (3.17) into (3.15), we get

(3.18)
$$
WW^T = L^T (A_{M^{1/2}}^{T\dagger})^T KK^T A_{M^{1/2}}^{T\dagger} L,
$$

where

$$
K = \begin{bmatrix} MA_{M^{1/2}}^{T\dagger}\|r\|_2, \|x\|_2\left(I_m - \frac{1}{\|r\|_2^2}rr^T\right), \|r\|_2\left(I_m - \frac{1}{\|r\|_2^2}rr^T\right), I_m \end{bmatrix}.
$$

Combining (3.3), (3.4), and (3.18), we have the expression (3.2). $\square$

Remark 3.2. Note that the orders of matrices in (3.1), (3.2) and (3.3) are $k\times k$, $k\times(3m+n)$ and $k\times m(n+m+1)$ respectively. So, comparing with the expression (3.3), the expressions of κ_n in Theorem 3.1, i.e., (3.1) and (3.2), require little storage space. Meanwhile, they can be efficiently computed due to the elimination of Kronecker product.

Remark 3.3. Using the weighted QR decomposition [17, Theorem 2.1, Equation 2.2], we can rewrite the expressions of $\kappa_{\phi 1}$ and $\kappa_{\phi 2}$ given in Theorem 3.1. Specifically, given $M = \text{diag}(M_n, M_{m-n})$ and a matrix $A \in \mathbb{R}^{m\times n}$ partitioned as

$$
A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}
$$

with $A_{11} \in \mathbb{R}^{p \times p}$, $A_{12} \in \mathbb{R}^{p \times (n-p)}$, $A_{21} \in \mathbb{R}^{(m-p) \times p}$, $A_{22} \in \mathbb{R}^{(m-p) \times (n-p)}$, then a factorization

$$A = Q \begin{bmatrix} R \\ 0 \end{bmatrix} \Pi^T, \quad M = QMQ^T$$

with Π being a permutation matrix, $R \in \mathbb{R}^{n \times n}$ being upper triangular and nonsingular and $Q \in \mathbb{R}^{m \times m}$ being nonsingular exists if and only if $\mathrm{rank}(A) = n$ and $\mathrm{rank}(A_{11}) = p$. By this decomposition, we have

$$x = \Pi \begin{bmatrix} R^{-1} & 0 \end{bmatrix} Q^{-1} b, \quad r = Q^{-T} \begin{bmatrix} 0 & \\ & M_{m-n}^{-1} \end{bmatrix} Q^{-1} b, \quad (A_{M^{1/2}}^{T\dagger})^T = \Pi \begin{bmatrix} R^{-1} & 0 \end{bmatrix} Q^{-1},$$

$$A_{M^{1/2}}^{T\dagger} = Q^{-T} \begin{bmatrix} R^{-T} \\ 0 \end{bmatrix} \Pi^T, \quad M A_{M^{1/2}}^{T\dagger} (A_{M^{1/2}}^{T\dagger})^T M = QMR^{-T}R^{-1}MQ^T.$$

Thus, putting the above terms into (3.1) and (3.2) lead to

$$\kappa_{\phi 1} = \left\| L^T \Pi R^{-1} Q^{-1} \big((QMR^{-T}R^{-1}MQ^T) \|r\|_2^2 + (\|r\|_2^2 + \|x\|_2^2 + 1) \big) Q^{-T} R^{-T} \Pi^T L \right\|_2^{1/2}$$
$$\times \frac{\|[A, M, b]\|_F}{\|L^T x\|_2}$$

and

$$\kappa_{\phi 2} = \left\| L^T (\Pi R^{-1} Q^{-1}) \left[(QMR^{-T}\Pi^T) \|r\|_2, \|x\|_2 \left(I_m - \frac{1}{\|r\|_2^2} rr^T \right), \|r\|_2 \left(I_m - \frac{1}{\|r\|_2^2} rr^T \right), I_m \right] \right\|_2$$
$$\times \frac{\|[A, M, b]\|_F}{\|L^T x\|_2}.$$

4. Mixed and componentwise condition numbers

We first present an explicit expression of the adjoint operator of the Fréchet derivative $\mathrm{d}\phi(A, M, b)$, which is necessary for deriving the explicit expression of mixed and componentwise condition numbers.

Theorem 4.1. *The adjoint operator of $\mathrm{d}\phi(A, M, b)$, using the scalar products $\langle (A_1, M_1, b_1), (A_2, M_2, b_2) \rangle = \mathrm{trace}(A_1^T A_2) + \mathrm{trace}(M_1^T M_2) + b_1^T b_2$ and $\langle y_1, y_2 \rangle = y_1^T y_2$ defined on $\mathbb{R}^{m \times n} \times \mathbb{R}^{m \times m} \times \mathbb{R}^m$ and $\mathbb{R}^k$ respectively, is*

$$\mathrm{d}\phi^*(y) = \big(r y^T L^T (A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger} - A_{M^{1/2}}^{T\dagger} L y x^T, -A_{M^{1/2}}^{T\dagger} L y r^T, A_{M^{1/2}}^{T\dagger} L y \big)$$

for any $y \in \mathbb{R}^k$.

Proof. Using the scalar product and (2.13), for any $y \in \mathbb{R}^k$, we have

$$\langle y, \mathrm{d}\phi(A, M, b) \cdot (\mathrm{d}A, \mathrm{d}M, \mathrm{d}b)\rangle = y^T (\mathrm{d}\phi(A, M, b) \cdot (\mathrm{d}A, \mathrm{d}M, \mathrm{d}b))$$

(4.1)
$$= y^T W \begin{bmatrix} \mathrm{vec}(\mathrm{d}A) \\ \mathrm{vec}(\mathrm{d}M) \\ \mathrm{d}b \end{bmatrix} = (W^T y)^T \begin{bmatrix} \mathrm{vec}(\mathrm{d}A) \\ \mathrm{vec}(\mathrm{d}M) \\ \mathrm{d}b \end{bmatrix}.$$

Considering the fact $\Pi^{-1} = \Pi^T$, and by (2.1), (2.2), and (2.3), we can obtain

(4.2)

$$W^T y = \begin{bmatrix} \Pi^T \left(r \otimes ((A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger} L)\right) - \left(x \otimes (A_{M^{1/2}}^{T\dagger} L)\right) \\ -r \otimes (A_{M^{1/2}}^{T\dagger} L) \\ A_{M^{1/2}}^{T\dagger} L \end{bmatrix} y \quad \text{(by (2.3))}$$

$$= \begin{bmatrix} \Pi^T \left(r \otimes ((A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger} L)\right) \mathrm{vec}(y) - \left(x \otimes (A_{M^{1/2}}^{T\dagger} L)\right) \mathrm{vec}(y) \\ -\left(r \otimes (A_{M^{1/2}}^{T\dagger} L)\right) \mathrm{vec}(y) \\ A_{M^{1/2}}^{T\dagger} L y \end{bmatrix}$$

$$= \begin{bmatrix} \Pi^{-1}\left(\mathrm{vec}((A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger} L y r^T) - \Pi \, \mathrm{vec}(A_{M^{1/2}}^{T\dagger} L y x^T)\right) \\ \mathrm{vec}(-A_{M^{1/2}}^{T\dagger} L y r^T) \\ A_{M^{1/2}}^{T\dagger} L y \end{bmatrix} \quad \text{(by (2.1))}$$

$$= \begin{bmatrix} \Pi^{-1}\left(\mathrm{vec}((A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger} L y r^T) - \mathrm{vec}((A_{M^{1/2}}^{T\dagger} L y x^T)^T)\right) \\ -\, \mathrm{vec}(A_{M^{1/2}}^{T\dagger} L y r^T) \\ A_{M^{1/2}}^{T\dagger} L y \end{bmatrix} \quad \text{(by (2.2))}$$

$$= \begin{bmatrix} \Pi^{-1} \mathrm{vec}\left(((A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger} L y r^T) - (x y^T L^T (A_{M^{1/2}}^{T\dagger})^T)\right) \\ -\, \mathrm{vec}(A_{M^{1/2}}^{T\dagger} L y r^T) \\ A_{M^{1/2}}^{T\dagger} L y \end{bmatrix}$$

$$= \begin{bmatrix} \mathrm{vec}\left((r y^T L^T (A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger}) - (A_{M^{1/2}}^{T\dagger} L y x^T)\right) \\ -\, \mathrm{vec}(A_{M^{1/2}}^{T\dagger} L y r^T) \\ A_{M^{1/2}}^{T\dagger} L y \end{bmatrix} \quad \text{(by (2.2) and } \Pi^{-1} = \Pi^T).$$

Putting (4.2) into (4.1) implies

$$\langle y, \mathrm{d}\phi(A, M, b) \cdot (\mathrm{d}A, \mathrm{d}M, \mathrm{d}b)\rangle$$
$$= \mathrm{vec}((r y^T L^T (A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger}) - (A_{M^{1/2}}^{T\dagger} L y x^T))^T \mathrm{vec}(\mathrm{d}A)$$
$$- \mathrm{vec}(A_{M^{1/2}}^{T\dagger} L y r^T)^T \mathrm{vec}(\mathrm{d}M) + (A_{M^{1/2}}^{T\dagger} L y)^T \mathrm{d}b,$$

which together with the fact that, for the matrices A_1 and A_2 of the same orders, $\operatorname{vec}(A_1^T)\operatorname{vec}(A_2) = \operatorname{trace}(A_1^T A_2)$, gives

$$\langle y, \mathrm{d}\phi(A, M, b) \cdot (\mathrm{d}A, \mathrm{d}M, \mathrm{d}b)\rangle$$

$$= \operatorname{trace}(((ry^T L^T (A_{M^{1/2}}^{T^\dagger})^T M A_{M^{1/2}}^{T^\dagger}) - (A_{M^{1/2}}^{T^\dagger} Lyx^T))^T \mathrm{d}A - (A_{M^{1/2}}^{T^\dagger} Lyr^T)^T \mathrm{d}M)$$

$$\quad + (A_{M^{1/2}}^{T^\dagger} Ly)^T \mathrm{d}b$$

$$= \langle (ry^T L^T (A_{M^{1/2}}^{T^\dagger})^T M A_{M^{1/2}}^{T^\dagger} - A_{M^{1/2}}^{T^\dagger} Lyx^T, -A_{M^{1/2}}^{T^\dagger} Lyr^T, A_{M^{1/2}}^{T^\dagger} Ly), (\mathrm{d}A, \mathrm{d}M, \mathrm{d}b)\rangle$$

$$= \langle \mathrm{d}\phi^*(A, M, b) \cdot y, (\mathrm{d}A, \mathrm{d}M, \mathrm{d}b)\rangle,$$

from which, we have the desired result. $\qquad\square$

Now, we give an explicit expression of the condition number κ (2.7) for the linear function ϕ of the CWLS solution.

Theorem 4.2. *The condition number (2.7) for the linear function ϕ of the CWLS solution is expressed by*

$$\kappa = \max_{\|y\|_{\mathbb{Q}^*}=1} \left\| [VD_A, HD_M, BD_b]^T Ly \right\|_1 = \left\| [VD_A, HD_M, BD_b]^T L \right\|_{\mathbb{Q}^*,1},$$

where

$$(4.3) \qquad \begin{aligned} V &= (A_{M^{1/2}}^{T^\dagger})^T (M A_{M^{1/2}}^{T^\dagger} \otimes r^T) - (A_{M^{1/2}}^{T^\dagger})^T (x^T \otimes I_m), \\ H &= (A_{M^{1/2}}^{T^\dagger})^T (r^T \otimes I_m), \quad B = (A_{M^{1/2}}^{T^\dagger})^T, \end{aligned}$$

and D_X denotes the diagonal matrix $\operatorname{diag}(\operatorname{vec}(X))$ for any matrix X.

Proof. Let $\mathrm{d}a_{ij}$, $\mathrm{d}m_{ij}$ and $\mathrm{d}b_i$ be the entries of $\mathrm{d}A$, $\mathrm{d}M$ and $\mathrm{d}b$, respectively. Thus, using (2.10), we have

$$\|(\mathrm{d}A, \mathrm{d}M, \mathrm{d}b)\|_{c^*} = \sum_{i,j} |\mathrm{d}a_{ij}||a_{ij}| + \sum_{i,j} |\mathrm{d}m_{ij}||m_{ij}| + \sum_i |\mathrm{d}b_i||b_i|.$$

By Theorem 4.1, we derive that

$$\|\mathrm{d}\phi^*(y)\|_{c^*} = \sum_{i,j=1}^{m,n} |a_{ij}| \left| \left(ry^T L^T (A_{M^{1/2}}^{T^\dagger})^T M A_{M^{1/2}}^{T^\dagger} - A_{M^{1/2}}^{T^\dagger} Lyx^T\right)_{ij} \right|$$

$$\quad + \sum_{i,j=1}^{m,m} |m_{ij}| \left| (A_{M^{1/2}}^{T^\dagger} Lyr^T)_{ij} \right| + \sum_{i=1}^{m} |b_i| \left| (A_{M^{1/2}}^{T^\dagger} Ly)_i \right|$$

$$= \sum_{i,j=1}^{m,n} |a_{ij}| \left| r_i \left((A_{M^{1/2}}^{T^\dagger})^T M A_{M^{1/2}}^{T^\dagger} e_j\right)^T - x_j \left((A_{M^{1/2}}^{T^\dagger})^T e_i\right)^T Ly \right|$$

$$\quad + \sum_{i,j=1}^{m,m} |m_{ij}| \left| r_j \left((A_{M^{1/2}}^{T^\dagger})^T e_i\right)^T Ly \right| + \sum_{i=1}^{m} |b_i| \left| \left((A_{M^{1/2}}^{T^\dagger})^T e_i\right)^T Ly \right|,$$

where r_i is the ith component of r. Noting (4.3), it can be verified that $\left(r_i((A_{M^{1/2}}^{T\dagger})^T M\right.$ $A_{M^{1/2}}^{T\dagger})e_j - x_j(A_{M^{1/2}}^{T\dagger})^T e_i)$ is the $(m(j-1)+i)$th column of the $n \times (mn)$ matrix V and $r_j(A_{M^{1/2}}^{T\dagger})^T e_i$ is the $(m(j-1)+i)$th column of the $m \times (mm)$ matrix H. Thus, the above expression equals

$$\left\|\begin{bmatrix} D_A V^T Ly \\ D_M H^T Ly \\ D_b B^T Ly \end{bmatrix}\right\|_1 = \left\|[VD_A, HD_M, BD_b]^T Ly\right\|_1.$$

Then, by Lemma 2.4, we obtain the desired result.   □

By Theorem 4.2, we can find the explicit expression of the mixed condition number for the linear function ϕ of the CWLS solution easily.

Corollary 4.3. *Using the above notations, when the infinity norm is chosen as the norm in the solution space $\mathbb{Q}$, we obtain*

$$\kappa_\infty = \left\||L^T V| \operatorname{vec}(|A|) + |L^T H| \operatorname{vec}(|M|) + |L^T B||b|\right\|_\infty.$$

When the infinity norm is chosen as the norm in the solution space $\mathbb{R}^n$, the corresponding mixed condition number is given by

$$(4.4) \qquad \kappa_m = \frac{\left\||L^T V| \operatorname{vec}(|A|) + |L^T H| \operatorname{vec}(|M|) + |L^T B||b|\right\|_\infty}{\|L^T x\|_\infty}.$$

Now, we give an alternative expression of κ_m, which doesn't contain the Kronecker product.

Corollary 4.4. *Using the above notations, we get*

$$(4.5) \qquad \begin{aligned} \kappa_m &= \frac{\left\| \sum_{j=1}^n |L^T((A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger} e_j r^T - x_j(A_{M^{1/2}}^{T\dagger})^T)||A(:,j)|\right\|_\infty}{\|L^T x\|_\infty} \\ &\quad + \frac{\sum_{j=1}^m |L^T r_j(A_{M^{1/2}}^{T\dagger})^T||M(:,j)| + |L^T(A_{M^{1/2}}^{T\dagger})^T||b|\right\|_\infty}{\|L^T x\|_\infty}, \end{aligned}$$

where $A(:,j)$ is the jth column of A.

Proof. Partitioning $V = [V_1, \ldots, V_n]$ and $H = [H_1, \ldots, H_m]$, where each V_j, $1 \le j \le n$, is an $n \times m$ matrix and each H_j, $1 \le j \le m$, is an $m \times m$ matrix, we get

$$(4.6) \qquad \begin{aligned} \kappa_m &= \frac{\left\||L^T V| \operatorname{vec}(|A|) + |L^T H| \operatorname{vec}(|M|) + |L^T B||b|\right\|_\infty}{\|L^T x\|_\infty} \\ &= \frac{\left\||L^T V_j||A(:,j)| + |L^T H_j||M(:,j)| + |L^T B||b|\right\|_\infty}{\|L^T x\|_\infty}. \end{aligned}$$

Recall that $\left(r_i((A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger})e_j - x_j(A_{M^{1/2}}^{T\dagger})^T e_i\right)$ is the $(m(j-1)+i)$th column of V. Then, we have

$$V_j = ((A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger} e_j r^T - x_j(A_{M^{1/2}}^{T\dagger})^T)e_i.$$

Similarly, by the fact that $r_j(A_{M^{1/2}}^{T\dagger})^T e_i$ is the $(m(j-1)+i)$th column of H, we get

$$H_j = r_j(A_{M^{1/2}}^{T\dagger})^T e_i.$$

Putting V_j and H_j into (4.6) leads to (4.5). $\qquad\square$

By Theorem 4.2, we can also get the expression of the componentwise condition number for the linear function ϕ of the CWLS solution easily.

Corollary 4.5. *Considering the componentwise norm defined by*

$$\|y\|_c = \min\{w, |y_i| \le w|(Lx)_i|, i = 1, \ldots, k\} = \max\{|y_i|/|(Lx)_i|, i = 1, \ldots, k\}$$

in the solution space, we have the following three expressions for the componentwise condition number for the linear function ϕ of the CWLS solution

$$
\begin{aligned}
\kappa_c &= \left\| D_{L^T x}^{-1} L^T [VD_A, HD_M, BD_b] \right\|_\infty \\
&= \left\| |D_{L^T x}^{-1}| (|L^T V| \operatorname{vec}(|A|) + |L^T H| \operatorname{vec}(|M|) + |L^T B||b|) \right\|_\infty \\
&= \left\| \sum_{j=1}^n |D_{L^T x}^{-1} L^T V_j||A(:,j)| + \sum_{j=1}^m |D_{L^T x}^{-1} L^T H_j||M(:,j)| + |D_{L^T x}^{-1} L^T B||b| \right\|_\infty .
\end{aligned}
$$
(4.7)

In the following, we give upper bounds for κ_m and κ_c.

Corollary 4.6. *The mixed and componentwise condition numbers for the linear function ϕ of the CWLS solution can be bounded as*

$$
\begin{aligned}
\kappa_m \le \kappa_m^{upp} &= \frac{\left\| L^T (A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger} D_{|A^T||r|} \right\|_\infty}{\|L^T x\|_\infty} + \frac{\left\| L^T (A_{M^{1/2}}^{T\dagger})^T D_{|A||x|} \right\|_\infty}{\|L^T x\|_\infty} \\
&\quad + \frac{\left\| L^T (A_{M^{1/2}}^{T\dagger})^T D_{|M||r|} \right\|_\infty}{\|L^T x\|_\infty} + \frac{\left\| L^T (A_{M^{1/2}}^{T\dagger})^T D_{|b|} \right\|_\infty}{\|L^T x\|_\infty},
\end{aligned}
$$

$$
\begin{aligned}
\kappa_c \le \kappa_c^{upp} &= \left\| D_{L^T x}^{-1} L^T (A_{M^{1/2}}^{T\dagger})^T M A_{M^{1/2}}^{T\dagger} D_{|A^T||r|} \right\|_\infty + \left\| D_{L^T x}^{-1} L^T (A_{M^{1/2}}^{T\dagger})^T D_{|A||x|} \right\|_\infty \\
&\quad + \left\| D_{L^T x}^{-1} L^T (A_{M^{1/2}}^{T\dagger})^T D_{|M||r|} \right\|_\infty + \left\| D_{L^T x}^{-1} L^T (A_{M^{1/2}}^{T\dagger})^T D_{|b|} \right\|_\infty.
\end{aligned}
$$

Proof. Firstly, note that for any matrix $N \in \mathbb{R}^{p \times q}$ and diagonal matrix $D_v \in \mathbb{R}^{q \times q}$,

(4.8) $\|ND_v\|_\infty = \||ND_v|\|_\infty = \||N||D_v|\|_\infty = \||N||D_v|e\|_\infty = \||N||D_v|\|_\infty = \||N||v|\|_\infty,$

where $e = [1, \ldots, 1] \in \mathbb{R}^q$. Using (2.1), (2.2), (2.5), (4.3), (4.8), and the triangle inequality, we have

$$
\begin{aligned}
\kappa_m &\leq \frac{\||L^T V| \operatorname{vec}(|A|)\|_\infty}{\|L^T x\|_\infty} + \frac{\||L^T H| \operatorname{vec}(|M|)\|_\infty}{\|L^T x\|_\infty} + \frac{\||L^T B||b|\|_\infty}{\|L^T x\|_\infty} \\
&\leq \frac{\||L^T (A_{M^{1/2}}^{T^\dagger})^T (M A_{M^{1/2}}^{T^\dagger} \otimes r^T)| \operatorname{vec}(|A|)\|_\infty}{\|L^T x\|_\infty} + \frac{\||L^T (A_{M^{1/2}}^{T^\dagger})^T (x^T \otimes I_m)| \operatorname{vec}(|A|)\|_\infty}{\|L^T x\|_\infty} \\
&\quad + \frac{\||L^T (A_{M^{1/2}}^{T^\dagger})^T (r^T \otimes I_m)| \operatorname{vec}(|M|)\|_\infty}{\|L^T x\|_\infty} + \frac{\||L^T (A_{M^{1/2}}^{T^\dagger})^T||b|\|_\infty}{\|L^T x\|_\infty} \quad \text{(by (4.3))} \\
&\leq \frac{\||L^T (A_{M^{1/2}}^{T^\dagger})^T M A_{M^{1/2}}^{T^\dagger}|(|r^T| \otimes I_n)\Pi \operatorname{vec}(|A|)\|_\infty}{\|L^T x\|_\infty} \\
&\quad + \frac{\||L^T (A_{M^{1/2}}^{T^\dagger})^T (|x^T| \otimes I_m)| \operatorname{vec}(|A|)\|_\infty}{\|L^T x\|_\infty} + \frac{\||L^T (A_{M^{1/2}}^{T^\dagger})^T|(|r^T| \otimes I_m) \operatorname{vec}(|M|)\|_\infty}{\|L^T x\|_\infty} \\
&\quad + \frac{\||L^T (A_{M^{1/2}}^{T^\dagger})^T||b|\|_\infty}{\|L^T x\|_\infty} \quad \text{(by (2.5))} \\
&= \frac{\||L^T (A_{M^{1/2}}^{T^\dagger})^T M A_{M^{1/2}}^{T^\dagger}||A^T||r|\|_\infty}{\|L^T x\|_\infty} + \frac{\||L^T (A_{M^{1/2}}^{T^\dagger})^T||A||x|\|_\infty}{\|L^T x\|_\infty} \\
&\quad + \frac{\||L^T (A_{M^{1/2}}^{T^\dagger})^T||M||r|\|_\infty}{\|L^T x\|_\infty} + \frac{\||L^T (A_{M^{1/2}}^{T^\dagger})^T||b|\|_\infty}{\|L^T x\|_\infty} \quad \text{(by (2.1) and (2.2))} \\
&= \frac{\|L^T (A_{M^{1/2}}^{T^\dagger})^T M A_{M^{1/2}}^{T^\dagger} D_{|A^T||r|}\|_\infty}{\|L^T x\|_\infty} + \frac{\|L^T (A_{M^{1/2}}^{T^\dagger})^T D_{|A||x|}\|_\infty}{\|L^T x\|_\infty} \\
&\quad + \frac{\|L^T (A_{M^{1/2}}^{T^\dagger})^T D_{|M||r|}\|_\infty}{\|L^T x\|_\infty} + \frac{\|L^T (A_{M^{1/2}}^{T^\dagger})^T D_{|b|}\|_\infty}{\|L^T x\|_\infty} = \kappa_m^{upp} \quad \text{(by (4.8))}.
\end{aligned}
$$

Similarly, we can prove $\kappa_c \leq \kappa_c^{upp}$. $\square$

5. Statistical condition estimates

In this part, we focus on estimating the normwise, mixed and componentwise condition numbers for the linear function ϕ of the CWLS solution.

5.1. Estimating normwise condition number

We use two algorithms to estimate the normwise condition number. The first one, outlined in Algorithm 5.1, is from [25] and has been applied to estimate the normwise condition number for matrix equations [32,37], equality constrained linear least squares problem [30], indefinite least squares problem [31] and K-weighted pseudoinverse $L_K^\dagger$ [35]. The second one, outlined in Algorithm 5.2, is based on the SSCE method [29] and has been used for some least squares problems [4,30,31].

Algorithm 5.1 Probabilistic condition estimator

Input: ϵ, d (d is the dimension of Krylov space and usually determined by the algorithm itself) and the matrix

$$\widehat{K} = L^T (A_{M^{1/2}}^{T\dagger})^T \left[M A_{M^{1/2}}^{T\dagger} \|r\|_2, \|x\|_2 \left(I_m - \frac{1}{\|r\|_2^2} rr^T \right), \|r\|_2 \left(I_m - \frac{1}{\|r\|_2^2} rr^T \right), I_m \right].$$

Output: Probabilistic spectral norm estimator of the normwise condition number (3.2): κ_{pn}.

1. Choose a starting vector v_0 uniformly and randomly from the unit t-sphere S_{t-1} with $t = 3m + n$.

2. Compute $\|\widehat{K}\|_2 \in [\alpha_1, \alpha_2]$ by probabilistic spectral norm estimator.

   (a) Find $\delta \in [0, \pi/2]$ by t, ϵ and

$$\int_0^{\arcsin(\delta)} \cos^{t-2}(q)\, dq = \frac{\epsilon}{2} \int_0^1 q^{(t-3)/2} (1-q)^{-1/2}\, dq \left(\epsilon \int_0^{\pi/2} \cos^{t-2}(q)\, dq \right).$$

   (b) **for** $j = 1, \ldots, d$

   (c) $u = \widehat{K} v_j$

   (d) **if** $j > 1$

   (e) $u = u - \beta_{j-1} u_{j-1}$

   (f) $u = u - [u_1, \ldots, u_{j-1}](u^T [u_1, \ldots, u_{j-1}])^T$

   (g) **end**

   (h) $\tau_j = \|u\|_2$

   (i) $u_j = u/\tau_j$

   (j) $v = \widehat{K}^T u$

   (k) $v = v - \tau_j v_j$

   (l) $v = v - [v_1, \ldots, v_{j-1}](v^T [v_1, \ldots, v_{j-1}])^T$

   (m) $\beta_j = \|v\|_2$

   (n) $v_{j+1} = v/\beta_j$

   (o) **end**

   (p) Find the largest singular value α_1 of B_d and an upper bidiagonal matrix with τ_i on the diagonal and β_i on upper subdiagonal.

   (q) Find the probabilistic upper bound α_2 for $\|\widehat{K}\|_2$ with probability $\geq 1-\epsilon$ by a Lanczos polynomial (see [25]).

3. Estimate the normwise condition number (3.2) by

$$\kappa_{pn} = \frac{\kappa_p \|[A, M, b]\|_F}{\|L^T x\|_2} \quad \text{with} \quad \kappa_p = \sqrt{\frac{\alpha_1 + \alpha_2}{2}}.$$

5.2. Estimating mixed and componentwise condition numbers

To estimate mixed and componentwise condition numbers, we need the following SSCE method, which is from [29] and has been applied to many problems (see e.g., [11–13, 31, 32, 34, 37]).

Algorithm 5.2 SSCE method for the normwise condition number of the CWLS solution

Input: Sample size k, $L = I_n$ and the matrix

$$\widehat{W} = (A_{M^{1/2}}^{T^\dagger})^T \big(M A_{M^{1/2}}^{T^\dagger} (A_{M^{1/2}}^{T^\dagger})^T M \|r\|_2^2 + (\|r\|_2^2 + \|x\|_2^2 + 1) \big) A_{M^{1/2}}^{T^\dagger}.$$

Output: SSCE estimates of the normwise condition number of CWLS solution (3.1): κ_{sn}.

1. Generate k vectors $z_1, \ldots, z_k$ uniformly and randomly from the unit n-sphere S_{n-1} and set $Z = [z_1, \ldots, z_k]$.

2. Orthonormalize these vectors using the QR factorization $[Z, \sim] = QR(Z)$.

3. For $i = 1, \ldots, k$, compute κ_i by
$$\kappa_i = \frac{\sigma_i \|[A, M, b]\|_F}{\|L^T x\|_2},$$
where
$$\sigma_i = \sqrt{z_i^T (A_{M^{1/2}}^{T^\dagger})^T \big(M A_{M^{1/2}}^{T^\dagger} (A_{M^{1/2}}^{T^\dagger})^T M \|r\|_2^2 + (\|r\|_2^2 + \|x\|_2^2 + 1) \big) A_{M^{1/2}}^{T^\dagger} z_i}.$$

4. Approximate ω_k and ω_n by
$$\omega_k \approx \sqrt{\frac{2}{\pi(k - 1/2)}}.$$

5. Estimate the normwise condition number (3.1) by
$$\kappa_{sn} = \frac{\omega_k}{\omega_n} \sqrt{\sum_{i=1}^{k} \kappa_i^2}.$$

---

6. Numerical experiments

In this section, we provide some numerical experiments for comparing different condition numbers for the linear function ϕ of the CWLS solution and testing the above algorithms. All the numerical experiments are performed in Matlab 2016a and on a PC with Intel i3-4005U CPU 1.70 GHz and 4.00 GB RAM.

Example 6.1. Similar to [17], we generate the following matrices and vectors to compare the condition numbers for the linear function ϕ of the CWLS solution

$$A = \begin{bmatrix} 1 & 1 & 5 & 4 \\ 1 & 2 & 4 & 2 \\ 1 & 3 & 3 & \eta \\ 1 & 1 & 6 & 1 \\ 1 & 6 & 10 & 2 \end{bmatrix}, \quad M = \begin{bmatrix} \mu & 0 & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 & 0 \\ 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$x = (-12, 1, 3, 3)^T$, $d_1 = [-1, -1, 1, 1, 0]^T$, $d_2 = [1, -2, 1, 0, 0]^T$, $r = d_1 + (5/\eta - 1)d_2$ and

Algorithm 5.3 SSCE method for the mixed and componentwise condition numbers of the CWLS solution: κ_m and κ_c

Input: Sample size k and matrix W in (2.13).

Output: SSCE estimates of mixed and componentwise condition numbers of CWLS solution: κ_{sm}, κ_{sc}.

1. Let $t = m(n + m + 1)$. Generate k random vectors $[z_1, \ldots, z_k] \to Z$ uniformly and randomly from the unit t-sphere S_{t-1} and set $Z = [z_1, \ldots, z_k]$.

2. Orthonormalize these vectors using the QR factorization $[Z, \sim] = QR(Z)$.

3. Compute $u_i = Cz_i$, and estimate the mixed and componentwise condition numbers in (4.4) and (4.7) by

$$\kappa_{sm} = \frac{\|\kappa_{sce}\|_\infty}{\|L^T x\|_\infty}, \quad \kappa_{sc} = \left\| \frac{\kappa_{sce}}{L^T x} \right\|_\infty,$$

where $\kappa_{sce} = \frac{\omega_k}{\omega_t} \left| \sum_{i=1}^k |u_i|^2 \right|^{1/2}$, and the power and square root operation are performed on each entry of u_i, $i = 1, \ldots, k$.

$b = Mr + Ax$. For the matrix L, we choose

$$L_0 = I_4, \quad L_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{4 \times 2}, \quad L_2 = (0, 0, 0, 1)^T \in \mathbb{R}^{4 \times 1}.$$

μ	η	L	κ_n	κ_m	κ_m^u	κ_c	κ_c^u
		L_0	7.3665e+01	7.6557e+00	1.0253e+02	2.7004e+01	4.4736e+01
10^{-2}	10^{-2}	L_1	6.9928e+01	7.6557e+00	1.0253e+02	2.7004e+01	4.4736e+01
		L_2	8.7455e+01	1.0318e+01	3.9297e+01	1.0318e+01	1.4998e+01
		L_0	8.1647e+01	8.5187e+00	1.1625e+02	4.1044e+01	6.3822e+01
10^{-6}	10^{-5}	L_1	7.8353e+01	8.5187e+00	1.1625e+02	4.1044e+01	6.3822e+01
		L_2	1.1475e+02	1.2753e+01	1.1625e+02	1.2753e+01	1.7902e+01
		L_0	3.0092e+02	2.4906e+01	3.1442e+02	1.3049e+02	1.6137e+02
10^{-8}	10^0	L_1	3.1275e+02	2.4906e+01	3.1442e+02	1.3049e+02	1.6137e+02
		L_2	5.0293e+02	4.4666e+01	1.4266e+02	4.4666e+01	5.5333e+01
		L_0	1.4772e+05	7.0741e+02	2.7120e+03	2.7402e+02	2.0695e+03
10^3	10^2	L_1	1.6936e+06	7.0741e+02	2.7120e+03	2.7402e+02	2.0695e+03
		L_2	6.0398e+04	1.4430e+02	2.8094e+02	1.4430e+02	1.9042e+02

Table 6.1: Comparisons of condition numbers and their upper bounds.

The numerical results are shown in Table 6.1 for various values of μ and η. We also do experiments based on the above matrices except that the original diagonal matrix M is replaced by a non-diagonal matrix $Q^T MQ$, where Q is a random orthogonal matrix

734 Mahvish Samar

obtained from the QR decomposition of a random matrix. The numerical results are shown in Table 6.2. From Tables 6.1 and 6.2, we can find that the mixed and componentwise condition numbers are smaller than the normwise ones.

μ	η	L	κ_n	κ_m	κ_m^u	κ_c	κ_c^u
		L_0	8.6830e+01	1.6773e+01	3.4759e+01	2.5478e+01	3.7475e+01
10^{-3}	10^0	L_1	9.0468e+01	1.6773e+01	3.4759e+01	2.5478e+01	3.7475e+01
		L_2	7.1700e+01	1.4028e+01	4.3596e+01	1.4028e+01	4.7282e+01
		L_0	1.4752e+02	1.8521e+01	4.7302e+01	4.4140e+01	5.5838e+01
10^{-4}	10^1	L_1	1.5330e+02	1.8521e+01	4.7302e+01	4.4140e+01	5.5838e+01
		L_2	6.0053e+01	8.9989e+00	2.2840e+01	8.9989e+00	4.7607e+01
		L_0	1.2672e+03	8.5359e+01	5.9791e+02	9.0859e+01	9.5931e+01
10^{-1}	10^{-2}	L_1	1.3239e+03	8.5359e+01	5.9791e+02	9.0859e+01	9.5931e+01
		L_2	1.1700e+03	5.6547e+01	1.2414e+02	5.6547e+01	4.7282e+02
		L_0	1.0408e+04	3.4238e+01	5.1495e+02	4.1085e+02	6.1495e+02
10^2	10^3	L_1	1.0867e+04	3.4238e+01	5.1495e+02	4.1085e+02	6.1495e+02
		L_2	7.0671e+01	2.2765e+00	4.6186e+00	2.2765e+00	8.6085e+00

Table 6.2: Comparisons of condition numbers and their upper bounds.

Example 6.2. In this example, we choose $L = I_n$, and generate the matrices A and M as follows:

$$A = QDU, \quad M = Q \begin{bmatrix} M_n & 0 \\ 0 & I_{m-n} \end{bmatrix} Q^T,$$

where $Q \in \mathbb{R}^{m \times m}$, $U \in \mathbb{R}^{n \times n}$ are the random orthogonal matrix, $D \in \mathbb{R}^{m \times n}$ is a diagonal matrix with diagonal elements distributed exponentially from κ^{-1} to 1 and $M_n = \text{diag}(\mu)$. In addition, we set the solution x to be $x = (1, 2^2, \ldots, n^2)$ and $b = Mr + Ax$ with r being random vector of 2-norm ρ, i.e., $\rho = \|r\|_2$.

In the specific experiments, we set $m = 50$, $n = 30$, and $p = 20$. For each pair of κ, μ and ρ, 1000 CWLS problems are generated to test the performance.

For Algorithm 5.1, we choose the parameters: $\delta = 0.01$ and $\epsilon = 0.001$. For Algorithms 5.2 and 5.3, we set $k = 2$. By varying the κ, μ and ρ, we have the numerical results on the ratios defined as follows:

$$r_p := \kappa_{pn}/\kappa_n, \quad r_s := \kappa_{sn}/\kappa_n, \quad r_m := \kappa_{sm}/\kappa_m, \quad r_c := \kappa_{sc}/\kappa_c.$$

The mean and variance of these ratios are reported in Tables 6.3 and 6.4, respectively. We also plot these numerical results and CPU time in Figures 6.1 and 6.2. The time ratios are defined by

$$t_p := \frac{t_1}{t}, \quad t_s := \frac{t_2}{t}, \quad t_m := \frac{t_3}{t}, \quad t_c := \frac{t_4}{t},$$

κ		10^1		10^2		10^3	
$\rho = 10^1$		Mean	Variance	Mean	Variance	Mean	Variance
$\mu = 10^{-4}$	r_p	1.0001e+00	3.8382e-12	1.0000e+00	3.9226e-12	1.0001e+00	3.6789e-12
	r_s	2.9550e+00	9.3512e-02	2.9546e+00	9.0331e-02	2.9744e+00	9.8811e-02
$\mu = 10^{-5}$	r_p	1.0000e+00	1.7705e-06	1.0000e+00	2.1614e-11	1.0000e+00	2.4665e-08
	r_s	2.1751e+00	1.3160e-01	2.1524e+00	1.2955e-01	2.1796e+00	1.3633e-01
$\mu = 10^{-6}$	r_p	1.0000e+00	3.0639e-08	10000e+00	3.3239e-08	1.0000e+00	1.4221e-08
	r_s	1.8324e+00	1.3171e-01	1.8338e+00	1.5551e-01	1.8297e+00	1.5833e-01
$\rho = 10^2$		Mean	Variance	Mean	Variance	Mean	Variance
$\mu = 10^{-4}$	r_p	1.0000e+00	3.8878e-12	1.0002e+00	3.8482e-12	1.0002e+00	3.9344e-12
	r_s	2.9536e+00	9.4841e-02	2.9637e+00	9.4633e-02	2.9712e+00	9.4020e-02
$\mu = 10^{-5}$	r_p	1.0000e+00	2.0674e-11	1.0000e+00	2.1240e-11	1.0002e+00	9.6383e-07
	r_s	2.1757e+00	1.3311e-01	2.1495e+00	1.3233e-01	2.1716e+00	1.2692e-01
$\mu = 10^{-6}$	r_p	1.0000e+00	8.9658e-09	1.0000e+00	1.6412e-08	1.0002e+00	1.4220e-08
	r_s	1.8237e+00	1.5151e-01	1.8188e+00	1.5864e-01	1.8337e+00	1.5080e-01
$\rho = 10^3$		Mean	Variance	Mean	Variance	Mean	Variance
$\mu = 10^{-4}$	r_p	1.0000e+00	4.4062e-12	1.0000e+00	4.1030e-12	1.0001e+00	4.0542e-12
	r_s	2.9913e+00	8.6931e-02	3.0217e+00	9.0505e-02	3.0081e+00	9.0301e-02
$\mu = 10^{-5}$	r_p	1.0000e+00	2.1453e-11	1.0000e+00	2.1129e-11	1.0000e+00	2.0967e-11
	r_s	2.1821e+00	1.2442e-01	2.1629e+00	1.2383e-01	2.1621e+00	1.1901e-01
$\mu = 10^{-6}$	r_p	1.0000e+00	3.1310e-08	1.0000e+00	1.7005e-08	1.0000e+00	2.5774e-08
	r_s	1.8255e+00	1.4654e-01	1.8453e+00	1.4585e-01	1.8322e+00	1.4621e-01

Table 6.3: The efficiency of statistical condition estimates by Algorithms 5.1 and 5.2.
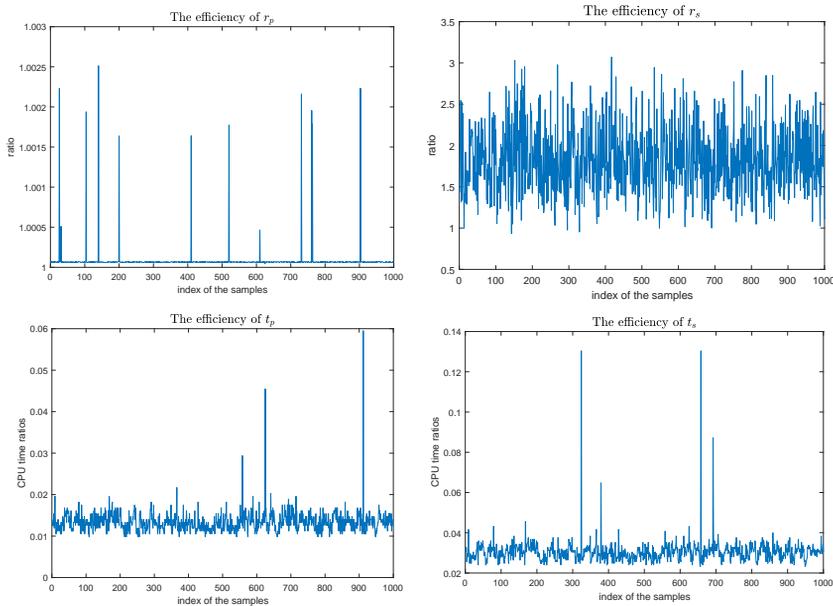

Figure 6.1: Efficiency of condition estimators and CPU time of Algorithms 5.1 and 5.2.

κ		10^1		10^2		10^3	
$\rho = 10^1$		Mean	Variance	Mean	Variance	Mean	Variance
$\mu = 10^{-4}$	r_m	1.3807e+00	7.7220e-02	1.3782e+00	8.0501e-02	1.3835e+00	8.1402e-02
	r_c	0.8322e+00	1.9441e-01	0.8370e+00	1.7952e-01	0.8615e+00	1.8543e-01
$\mu = 10^{-5}$	r_m	1.6976e+00	2.0242e-01	1.7073e+00	1.9431e-01	1.7035e+00	1.9312e-01
	r_c	1.1760e+00	3.6573e-01	1.1450e+00	3.5104e-01	1.1554e+00	3.4043e-01
$\mu = 10^{-6}$	r_m	1.8818e+00	2.9632e-01	1.9308e+00	3.5793e-01	1.8706e+00	3.2762e-01
	r_c	1.4282e+00	5.0721e-01	1.3615e+00	5.0001e-01	1.3465e+00	4.5590e-01
$\rho = 10^2$		Mean	Variance	Mean	Variance	Mean	Variance
$\mu = 10^{-4}$	r_m	1.3251e+00	7.3202e-02	1.3079e+00	7.4221e-02	1.3362e+00	8.0021e-02
	r_c	0.8071e+00	1.7441e-01	0.7872e+00	1.6101e-01	0.8185e+00	1.6173e-01
$\mu = 10^{-5}$	r_m	1.5878e+00	1.6023e-01	1.6128e+00	1.9144e-01	1.6067e+00	1.8671e-01
	r_c	1.0771e+00	2.9923e-01	1.0986e+00	3.3625e-01	1.0676e+00	2.9664e-01
$\mu = 10^{-6}$	r_m	1.7930e+00	3.2042e-01	1.7706e+00	2.7282e-01	1.7888e+00	2.8513e-01
	r_c	1.2866e+00	4.1963e-01	1.2776e+00	4.1824e-01	1.2987e+00	4.5624e-01
$\rho = 10^3$		Mean	Variance	Mean	Variance	Mean	Variance
$\mu = 10^{-4}$	r_m	1.0986e+00	4.2912e-02	1.0981e+00	3.8142e-02	1.0983e+00	3.9241e-02
	r_c	1.0574e+00	8.5532e-02	1.0081e+00	9.1812e-02	1.0058e+00	9.6831e-02
$\mu = 10^{-5}$	r_m	1.0657e+00	7.9813e-02	1.0823e+00	8.0414e-02	1.0730e+00	7.3033e-02
	r_c	1.0073e+00	1.4053e-01	1.0043e+00	1.4354e-01	1.0721e+00	1.3933e-01
$\mu = 10^{-6}$	r_m	1.1000e+00	1.0652e-01	1.1132e+00	1.1690e-01	1.1244e+00	1.2063e-01
	r_c	1.0795e+00	1.6881e-01	1.0080e+00	1.8233e-01	1.0156e+00	1.7393e-01

Table 6.4: The efficiency of statistical condition estimates by Algorithm 5.3.
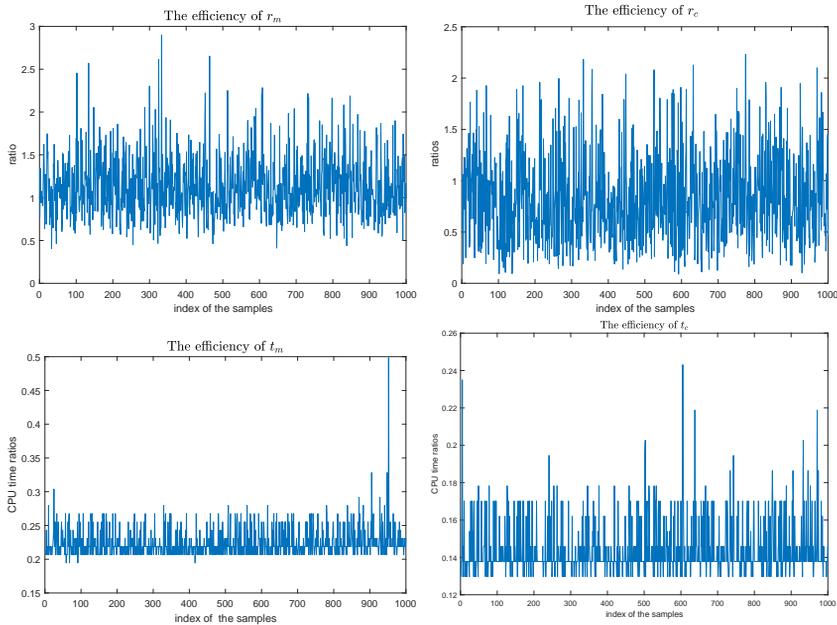

Figure 6.2: Efficiency of condition estimators and CPU time of Algorithm 5.3.

where t is the CPU time of computing the CWLS solution x by the weighted QR decomposition [17] and t_1, t_2, t_3 and t_4 are the CPU time of Algorithms 5.1, 5.2 and 5.3. These results suggest that these three algorithms are very effective and reliable in estimating condition numbers.

In the rest of this section, we test the effectiveness of over-estimation ratios of Algorithms 5.1, 5.2 and 5.3. For the perturbations, we generate them as

$$\Delta A = \varepsilon_1 \times (E \odot A), \quad \Delta M = \varepsilon_1 \times (F \odot M), \quad \Delta b = \varepsilon_1 \times (g \odot b),$$

where $\varepsilon_1 = 10^{-8}$ and E, F and g are random matrices whose entries are uniformly distributed in the open interval $(-1, 1)$ and $\odot$ denotes componentwise product for two matrices. When the perturbations are small enough, we denote the unique solution by $x + \Delta x$ of the following perturbed CWLS problem:

$$\min_{x \in \mathbb{R}^n} ((b_2 + \Delta b_2) - (A_2 + \Delta A_2)(x + \Delta x))^T (M_2 + \Delta M_2)^{-1}$$
$$\times ((b_2 + \Delta b_2) - (A_2 + \Delta A_2)(x + \Delta x))$$
$$\text{such that} \quad (A_1 + \Delta A_1)(x + \Delta x) = (b_1 + \Delta b_1),$$

where $\Delta A_1 \in \mathbb{R}^{p \times n}$, $\Delta A_2 \in \mathbb{R}^{q \times n}$, $\Delta b_1 \in \mathbb{R}^p$, $\Delta b_2 \in \mathbb{R}^q$, $\Delta M_2 \in \mathbb{R}^{q \times q}$ and

$$\Delta A = \begin{bmatrix} \Delta A_1 \\ \Delta A_2 \end{bmatrix}, \quad \Delta b = \begin{bmatrix} \Delta b_1 \\ \Delta b_2 \end{bmatrix}.$$

To measure the effectiveness of the estimators, we define the over-estimation ratios

$$r_p^{over} := \frac{\kappa_{pn} \cdot \varepsilon_1}{\|\Delta x\|_2 / \|x\|_2}, \qquad r_s^{over} := \frac{\kappa_{sn} \cdot \varepsilon_1}{\|\Delta x\|_2 / \|x\|_2},$$
$$r_m^{over} := \frac{\kappa_{sm} \cdot \varepsilon_1}{\|\Delta x\|_\infty / \|x\|_\infty}, \qquad r_c^{over} := \frac{\kappa_{sc} \cdot \varepsilon_1}{\|\Delta x / x\|_\infty}.$$

As κ_{pn}, κ_{sn}, κ_{sm} and κ_{sc} are the outputs from Algorithms 5.1, 5.2 and 5.3. Typically, the ratios in $(0.1, 10)$ are acceptable [24, Chapter 19]. Recall the CPU time of over-estimation ratios are defined by

$$t_p^{over} := \frac{t_5}{t}, \quad t_s^{over} := \frac{t_6}{t}, \quad t_m^{over} := \frac{t_7}{t}, \quad t_c^{over} := \frac{t_8}{t},$$

where t is the CPU time of computing the CWLS solution x by the weighted QR decomposition [17] and t_5, t_6, t_7 and t_8 are the CPU time of over-estimation ratios for Algorithms 5.1, 5.2 and 5.3. From Figures 6.3 and 6.4, we can see that r_p^{over} and r_s^{over} may seriously overestimate the true relative normwise error and the mixed and componentwise condition estimations r_m^{over} and r_c^{over} are effective. However, the CPU time of Algorithms 5.1 and 5.2 are smaller than Algorithm 5.3.
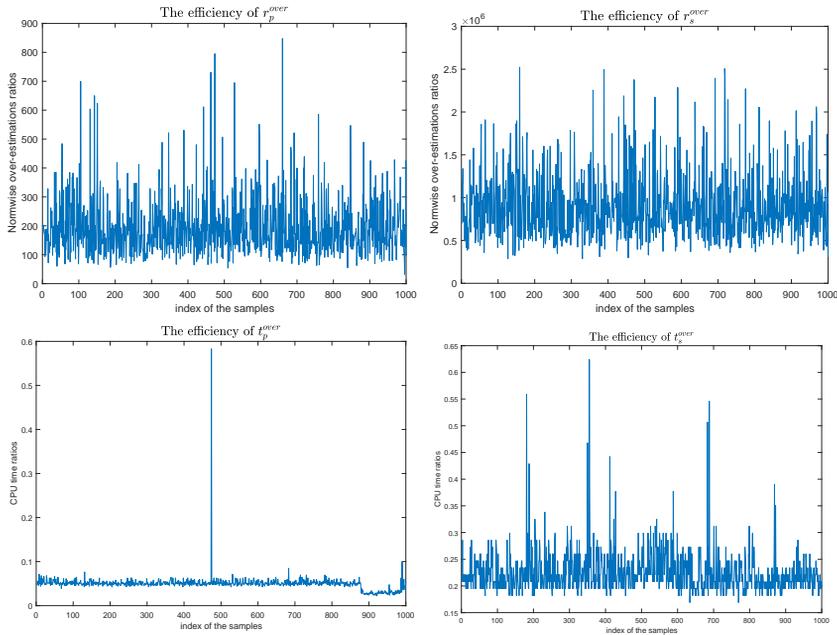
Figure 6.3: Efficiency of over-estimation ratios and CPU time of Algorithms 5.1 and 5.2.
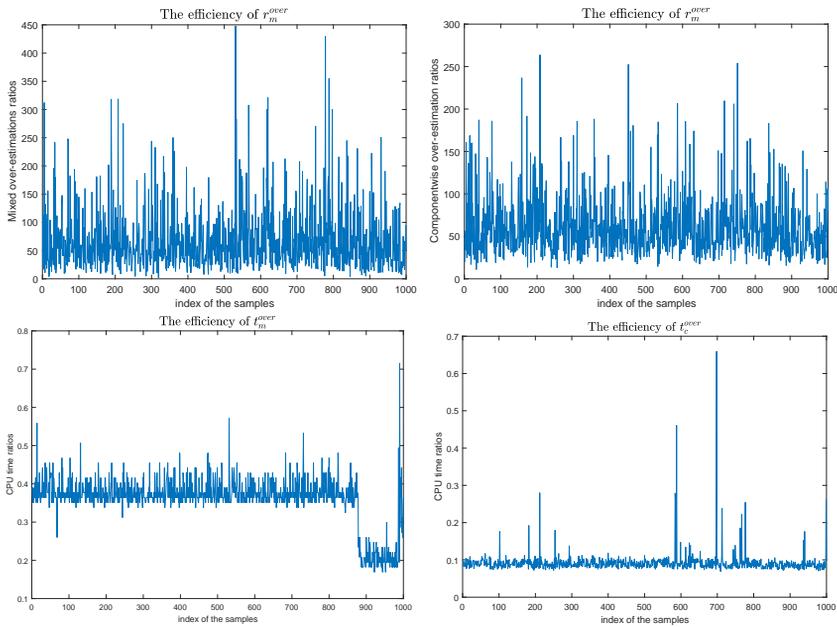

Figure 6.4: Efficiency of over-estimation ratios and CPU time of Algorithm 5.3.

References

[1] M. Arioli, M. Baboulin and S. Gratton, *A partial condition number for linear least squares problems*, SIAM J. Matrix Anal. Appl. **29** (2007), no. 2, 413–433.

[2] M. Baboulin and 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), no. 1, 3–19.

[3] _____, *A contribution to the conditioning of the total least-squares problem*, SIAM J. Matrix Anal. Appl. **32** (2011), no. 3, 685–699.

[4] M. Baboulin, S. Gratton, R. Lacroix and A. J. Laub, *Statistical estimates for the conditioning of linear least squares problems*, in: *Parallel Processing and Applied Mathematics, Part I*, 124–133, Lecture Notes in Comput. Sci. **8384**, Springer, Heidelberg, 2014.

[5] Å. Björck, *Numerical Methods for Least Squares Problems*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996.

[6] P. Bürgisser and F. Cucker, *Condition: The geometry of numerical algorithms*, Grundlehren der Mathematischen Wissenschaften **349**, Springer, Heidelberg, 2013.

[7] F. Cucker, H. Diao and Y. Wei, *On mixed and componentwise condition numbers for Moore-Penrose inverse and linear least squares problems*, Math. Comp. **76** (2007), no. 258, 947–963.

[8] H.-A. Diao, *Condition numbers for a linear function of the solution of the linear least squares problem with equality constraints*, J. Comput. Appl. Math. **344** (2018), 640–656.

[9] H.-A. Diao, L. Liang and S. Qiao, *A condition analysis of the weighted linear least squares problem using dual norms*, Linear Multilinear Algebra **66** (2018), no. 6, 1085–1103.

[10] H.-A. Diao and Y. Sun, *Mixed and componentwise condition numbers for a linear function of the solution of the total least squares problem*, Linear Algebra Appl. **544** (2018), 1–29.

[11] H.-A. Diao, Y. Wei and S. Qiao, *Structured condition numbers of structured Tikhonov regularization problem and their estimations*, J. Comput. Appl. Math. **308** (2016), 276–300.

[12] H.-A. Diao, Y. Wei and P. Xie, *Small sample statistical condition estimation for the total least squares problem*, Numer. Algorithms **75** (2017), no. 2, 435–455.

[13] H.-A. Diao, H. Xiang and Y. Wei, *Mixed, componentwise condition numbers and small sample statistical condition estimation of Sylvester equations*, Numer. Linear Algebra Appl. **19** (2012), no. 4, 639–654.

[14] H.-A. Diao and T.-Y. Zhou, *Backward error and condition number analysis for the indefinite linear least squares problem*, Int. J. Comput. Math. **96** (2019), no. 8, 1603–1622.

[15] I. Gohberg and I. Koltracht, *Mixed, componentwise, and structured condition numbers*, SIAM J. Matrix Anal. Appl. **14** (1993), no. 3, 688–704.

[16] A. Graham, *Kronecker Products and Matrix Calculus: With applications*, Ellis Horwood Series in Mathematics and its Applications, Ellis Horwood, Chichester, New York, 1981.

[17] M. Gulliksson, *Iterative refinement for constrained and weighted linear least squares*, BIT **34** (1994), no. 2, 239–253.

[18] _____, *Backward error analysis for the constrained and weighted linear least squares problem when using the weighted QR factorization*, SIAM J. Matrix Anal. Appl. **16** (1995), no. 2, 675–687.

[19] _____, *On the modified Gram-Schmidt algorithm for weighted and constrained linear least squares problems*, BIT **35** (1995), no. 4, 453–468.

[20] M. Gulliksson, X.-Q. Jin and Y.-M. Wei, *Perturbation bounds for constrained and weighted least squares problems*, Linear Algebra Appl. **349** (2002), 221–232.

[21] M. Gulliksson, I. Söderkvist and P.-Å. Wedin, *Algorithms for constrained and weighted nonlinear least squares*, SIAM J. Optim. **7** (1997), no. 1, 208–224.

[22] M. Gulliksson and P.-Å. Wedin, *Modifying the QR-decomposition to constrained and weighted linear least squares*, SIAM J. Matrix Anal. Appl. **13** (1992), no. 4, 1298–1313.

[23] _____, *Perturbation theory for generalized and constrained linear least squares*, Numer. Linear Algebra Appl. **7** (2000), no. 4, 181–195.

[24] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, Second edition, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.

[25] M. E. Hochstenbach, *Probabilistic upper bounds for the matrix two-norm*, J. Sci. Comput. **57** (2013), no. 3, 464–476.

[26] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[27] P. D. Hough and S. A. Vavasis, *Complete orthogonal decomposition for weighted least squares*, SIAM J. Matrix Anal. Appl. **18** (1997), no. 2, 369–392.

[28] X.-Q. Jin, *A preconditioner for constrained and weighted least squares problems with Toeplitz structure*, BIT **36** (1996), no. 1, 101–109.

[29] C. S. Kenney and A. J. Laub, *Small-sample statistical condition estimates for general matrix functions*, SIAM J. Sci. Comput. **15** (1994), no. 1, 36–61.

[30] H. Li and S. Wang, *Partial condition number for the equality constrained linear least squares problem*, Calcolo **54** (2017), no. 4, 1121–1146.

[31] _____, *On the partial condition numbers for the indefinite least squares problem*, Appl. Numer. Math. **123** (2018), 200–220.

[32] H. Li, S. Wang and C. Zheng, *Perturbation analysis for the periodic generalized coupled Sylvester equation*, Int. J. Comput. Math. **94** (2017), no. 10, 2011–2026.

[33] J. R. Rice, *A theory of condition*, SIAM J. Numer. Anal. **3** (1966), no. 2, 287–310.

[34] M. Samar, A. Farooq and C.-l. Mu, *Structured condition numbers and statistical condition estimation for the LDU factorization*, Appl. Math. J. Chinese Univ. Ser. B **35** (2020), no. 3, 332–348.

[35] M. Samar, H. Li and Y. Wei, *Condition numbers for the K-weighted pseudoinverse $L_K^\dagger$ and their statistical estimation*, Accepted in Linear Multilinear Algebra, (2019), 19 pp. https://doi.org/10.1080/03081087.2019.1618235

[36] G. Strang, *A framework for equilibrium equations*, SIAM Rev. **30** (1988), no. 2, 283–297.

[37] S. Wang, H. Yang and H. Li, *Condition numbers for the nonlinear matrix equation and their statistical estimation*, Linear Algebra Appl. **482** (2015), 221–240.

[38] P.-Å. Wedin, *Perturbation theory and condition numbers for generalized and constrained linear least squares problem*, Technical Report UMINF 125.85, Inst. of Info. Proc., University of Umeå, Sweden, 1985.

Mahvish Samar
Department of Mathematics, Shantou University, Shantou 515063, China
E-mail address: mahvishsamar@hotmail.com