

ON MIXED AND COMPONENTWISE CONDITION NUMBERS FOR MOORE–PENROSE INVERSE AND LINEAR LEAST SQUARES PROBLEMS

FELIPE CUCKER, HUAIAN DIAO, AND YIMIN WEI

ABSTRACT. Classical condition numbers are normwise: they measure the size of both input perturbations and output errors using some norms. 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. In this paper, we give explicit expressions, computable from the data, for the mixed and componentwise condition numbers for the computation of the Moore–Penrose inverse as well as for the computation of solutions and residues of linear least squares problems. In both cases the data matrices have full column (row) rank.

1. INTRODUCTION

1.1. General considerations. Condition numbers measure the *worst-case* sensitivity of an input data to *small* perturbations. To the best of our knowledge a general theory of condition numbers was first given by Rice in [12]. Let $\phi : \mathbb{R}^s \rightarrow \mathbb{R}^t$ be a mapping, where \mathbb{R}^s and \mathbb{R}^t are the usual s - and t -dimensional Euclidean spaces equipped with some norms, respectively. If ϕ is continuous and Fréchet differentiable in the neighborhood of $a_0 \in \mathbb{R}^s$, then, according to [12], the *relative normwise condition number* of a_0 is given by

$$\text{cond}(a_0) := \lim_{\varepsilon \rightarrow 0} \sup_{\|\Delta a\| \leq \varepsilon} \left(\frac{\|\phi(a_0 + \Delta a) - \phi(a_0)\|}{\|\phi(a_0)\|} / \frac{\|\Delta a\|}{\|a_0\|} \right) = \frac{\|\phi'(a_0)\| \|a_0\|}{\|\phi(a_0)\|},$$

where $\phi'(a_0)$ is the Fréchet derivative of ϕ at a_0 .

A drawback of condition numbers as defined above is that they ignore the structure of both input and output data with respect to scaling and/or sparsity. When the data is badly scaled or contains many zeros measuring the size of a perturbation in terms of its norm, we are left in the dark concerning the relative size of the perturbation on its small (or zero) entries.

To tackle this drawback, another approach in perturbation theory, known as *componentwise analysis*, has been increasingly considered. To be precise, two different kinds of condition number were studied: first, those measuring the errors in

Received by the editor July 21, 2005 and, in revised form, November 23, 2005.

2000 *Mathematics Subject Classification*. Primary 15A09, 15A12; Secondary 65F35.

Key words and phrases. Condition numbers, componentwise analysis, least squares.

The first author was partially supported by City University SRG grant 7001860.

The third author was partially supported by the National Natural Science Foundation of China under grant 10471027 and Shanghai Education Committee.

©2006 American Mathematical Society
Reverts to public domain 28 years from publication

the output using norms and the input perturbations componentwise, and second, those measuring both the error in the output and the perturbation in the input componentwise. The resulting condition numbers are called *mixed* and *componentwise*, respectively, by Gohberg and Koltracht [6]. We will use this terminology throughout this paper.

By their very nature, condition numbers are defined as limits of suprema. Therefore, their definition does not suggest a way to compute them from the input data. To do so, equivalent explicit expressions are sought or, alternatively, easy to compute and sufficiently sharp upper bounds. This has been done extensively for many problems in linear algebra, mostly for normwise condition numbers.

In this paper we exhibit explicit expressions for mixed and componentwise condition numbers for both Moore–Penrose inverse and linear least squares problems. We also exhibit upper bounds for these condition numbers which are easier to compute for large matrices.

1.2. A brief description of some previous work. Probably the first mixed perturbation analysis was done by Skeel [14]. He performed a mixed perturbation analysis for nonsingular linear systems of equations and a mixed error analysis for Gaussian elimination.

Skeel’s condition number is of mixed type. It is defined using componentwise perturbations on the input data and infinity norm in the solution. In [13], Rohn introduced a new relative condition number measuring both perturbation in the input data and error in the output componentwisely.

They were Gohberg and Koltracht [6] who named Skeel’s condition number *mixed* to distinguish it from componentwise condition numbers such as that in [13]. They also gave explicit expressions for both mixed and componentwise condition numbers, always for square systems of linear equations.

Perturbation theory for rectangular matrices and linear least squares problems existed for quite a while for the normwise case (cf. [15, 18]) and has been further studied in [5, 8, 9, 11].

For the mixed and componentwise settings for the problem of linear least squares, the existing results consisted of bounds for both condition numbers (or first-order perturbation bounds) and unrestricted perturbation bounds. There were no explicit expressions for mixed and componentwise condition numbers. To the best of our knowledge the sharpest such bounds were given in [1, 3, 10].

1.3. Main definitions and results. To define mixed and componentwise condition numbers, the following form of “distance” function will be useful. For any points $a, b \in \mathbb{R}^n$ we define $\frac{a}{b} = (c_1, \dots, c_n)$ with

$$c_i = \begin{cases} a_i/b_i, & \text{if } b_i \neq 0, \\ 0, & \text{if } a_i = b_i = 0, \\ \infty, & \text{otherwise.} \end{cases}$$

Then we define

$$d(a, b) = \left\| \frac{a - b}{b} \right\|_{\infty} = \max_{i=1, \dots, n} \left\{ \frac{|a_i - b_i|}{|b_i|} \right\}.$$

Note that if $d(a, b) < \infty$,

$$d(a, b) = \min\{\nu \geq 0 \mid |a_i - b_i| \leq \nu|b_i| \text{ for } i = 1, \dots, n\}.$$

In the rest of this paper we will only consider pairs (a, b) for which $d(a, b) < \infty$. We can extend the function d to matrices in an obvious manner. We introduce a notation allowing us to do so smoothly. For a matrix $A \in \mathbb{R}^{m \times n}$ we define $\text{vec}(A) \in \mathbb{R}^{mn}$ by $\text{vec}(A) = [a_1^T, \dots, a_n^T]^T$, where $A = [a_1, \dots, a_n]$ with $a_i \in \mathbb{R}^m$, $i = 1, \dots, n$. Then, we define

$$d(A, B) = d(\text{vec}(A), \text{vec}(B)).$$

Note that vec is a homeomorphism between $\mathbb{R}^{m \times n}$ and \mathbb{R}^{mn} . In addition, it transforms norms in the sense that, for all $A \in \mathbb{R}^{m \times n}$,

$$(1.1) \quad \|\text{vec}(A)\|_\infty = \|A\|_{\max}$$

where $\|\cdot\|_{\max}$ is the max norm given by

$$\|A\|_{\max} = \max_{i,j} |A_{ij}|.$$

For $\varepsilon > 0$ we denote $B^0(a, \varepsilon) = \{x \mid d(x, a) \leq \varepsilon\}$. For a partial function $F : \mathbb{R}^p \rightarrow \mathbb{R}^q$ we denote by $\text{Dom}(f)$ its domain of definition.

Definition 1. Let $F : \mathbb{R}^p \rightarrow \mathbb{R}^q$ be a continuous mapping defined on an open set $\text{Dom}(F) \subset \mathbb{R}^p$ such that $0 \notin \text{Dom}(F)$. Let $a \in \text{Dom}(F)$ such that $F(a) \neq 0$.

(i) The *mixed condition number* of F at a is defined by

$$m(F, a) = \lim_{\varepsilon \rightarrow 0} \sup_{\substack{x \in B^0(a, \varepsilon) \\ x \neq a}} \frac{\|F(x) - F(a)\|_\infty}{\|F(a)\|_\infty} \frac{1}{d(x, a)}.$$

(ii) Suppose $F(a) = (f_1(a), \dots, f_q(a))$ is such that $f_j(a) \neq 0$ for $j = 1, \dots, q$. Then the *componentwise condition number* of F at a is

$$c(F, a) = \lim_{\varepsilon \rightarrow 0} \sup_{\substack{x \in B^0(a, \varepsilon) \\ x \neq a}} \frac{d(F(x), F(a))}{d(x, a)}.$$

In this paper we consider these condition numbers for the *Moore–Penrose inverse* of $A \in \mathbb{R}^{m \times n}$. We recall that this is the unique $n \times m$ matrix A^\dagger satisfying the four matrix equations [2, 17]

$$AA^\dagger A = A, \quad A^\dagger AA^\dagger = A^\dagger, \quad (AA^\dagger)^T = AA^\dagger, \quad (A^\dagger A)^T = A^\dagger A.$$

Here, for a real matrix M , M^T denotes its transpose matrix.

Identifying $\mathbb{R}^{m \times n}$ with \mathbb{R}^{mn} via vec and using (1.1), Definition 1 yields, respectively,

$$m_\dagger(A) := \lim_{\varepsilon \rightarrow 0} \sup_{\|\Delta A/A\|_{\max} \leq \varepsilon} \frac{\|(A + \Delta A)^\dagger - A^\dagger\|_{\max}}{\|A^\dagger\|_{\max}} \frac{1}{\|\Delta A/A\|_{\max}}$$

and

$$c_\dagger(A) := \lim_{\varepsilon \rightarrow 0} \sup_{\|\Delta A/A\|_{\max} \leq \varepsilon} \frac{1}{\|\Delta A/A\|_{\max}} \left\| \frac{(A + \Delta A)^\dagger - A^\dagger}{A^\dagger} \right\|_{\max}.$$

Here $\frac{B}{A}$ is an entrywise division defined by $\frac{B}{A} := \text{vec}^{-1}(\text{vec}(B)/\text{vec}(A))$. Note also that in the definition of $c_\dagger(A)$ we are assuming that A^\dagger has no zero components.

The main result for these condition numbers, Theorem 1, gives explicit expressions for them. Corollary 1 then gives easier-to-compute upper bounds.

In a similar way, one defines, given a full column rank matrix A and a vector b , the condition numbers $m_{ls}(A, b)$, and $c_{ls}(A, b)$ for the computation of the solution x of the least squares problem

$$\min_{v \in \mathbb{R}^n} \|Av - b\|_2,$$

and the condition numbers $m_{res}(A, b)$, and $c_{res}(A, b)$ for the computation of the residue $r = b - Ax$. The main results in Section 4, Theorems 2 and 4, give explicit expressions for them. Easier-to-compute upper bounds are also shown in this section. Theorem 3 also gives sharp bounds for unrestricted (i.e., not necessarily small) perturbations. Section 5 shows similar results for the Moore–Penrose inverse of full row rank matrices and for minimal norm solutions of underdetermined linear systems.

2. PRELIMINARIES

2.1. Kronecker products. If $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$, then the *Kronecker product* $A \otimes B \in \mathbb{R}^{mp \times nq}$ is defined by

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

The following results can be found in [7]

$$(2.1) \quad |A \otimes B| = |A| \otimes |B|,$$

$$(2.2) \quad \text{vec}(AXB) = (B^T \otimes A)\text{vec}(X),$$

where $|A| = (|A_{ij}|)$, A_{ij} is the (i, j) th entry of A .

Also, it is proven in [7] that there exists a matrix $\Pi \in \mathbb{R}^{mn \times mn}$ such that, for all $A \in \mathbb{R}^{m \times n}$,

$$(2.3) \quad \Pi(\text{vec}(A)) = \text{vec}(A^T).$$

The matrix Π is called the *vec-permutation matrix* and can be represented explicitly by

$$(2.4) \quad \Pi = \sum_{i=1}^n \sum_{j=1}^m E_{ij}(m \times n) \otimes E_{ji}(n \times m).$$

Here $E_{ij}(m \times n) = e_i^{(m)}(e_j^{(n)})^T \in \mathbb{R}^{m \times n}$ denotes the (i, j) th elementary matrix and $e_i^{(m)}$ is the vector $[0, \dots, 0, 1, 0, \dots, 0]^T \in \mathbb{R}^m$, the 1 in the i th component (see [7]).

Also it is proved in [11] that for any vector $y \in \mathbb{R}^p$ and matrix $Y \in \mathbb{R}^{p \times q}$,

$$(2.5) \quad (y^T \otimes Y) \Pi = Y \otimes y^T.$$

2.2. Condition numbers and differentiability. The following lemma gives expressions for the mixed and componentwise condition numbers for differentiable functions. In its statement, and in all that follows, if $a \in \mathbb{R}^p$, we denote by $\text{Dg}(a)$ the $p \times p$ diagonal matrix with a_1, \dots, a_p in the diagonal. Recall, for a matrix $A \in \mathbb{R}^{m \times n}$, that its operator norm with respect to $\|\cdot\|_\infty$ satisfies

$$\|A\|_\infty = \max_{i \leq m} \sum_{j=1}^n |A_{ij}|.$$

Lemma 1 ([6]). *Let $F : \mathbb{R}^p \rightarrow \mathbb{R}^q$ be as in Definition 1 and $a \in \text{Dom}(F)$ be such that F is Fréchet differentiable at a . Then,*

- (a) *If $F(a) \neq 0$, then $m(F, a) = \frac{\|DF(a)\text{Dg}(a)\|_\infty}{\|F(a)\|_\infty}$.*
- (b) *If $(F(a))_i \neq 0$ for $i = 1, \dots, q$, then $c(F, a) = \|\text{Dg}(F(a))^{-1}DF(a)\text{Dg}(a)\|_\infty$.* □

We can simplify the expressions in Lemma 1.

Lemma 2. *With the notation above, we have*

- (a) *If $F(a) \neq 0$, then $m(F, a) = \frac{\| |DF(a)| |a| \|_\infty}{\|F(a)\|_\infty}$.*
- (b) *If $(F(a))_i \neq 0$ for $i = 1, \dots, q$, then $c(F, a) = \left\| \frac{|DF(a)| |a|}{|F(a)|} \right\|_\infty$.*

Proof. It readily follows from the fact that for any matrix $A \in \mathbb{R}^{p \times q}$ and diagonal matrix $\text{Dg}(d) \in \mathbb{R}^{q \times q}$ we have

$$\|A\text{Dg}(d)\|_\infty = \| |A\text{Dg}(d)| \|_\infty = \| |A| |\text{Dg}(d)| \|_\infty = \| |A| |\text{Dg}(d)| \mathbf{e}_q \|_\infty = \| |A| |d| \|_\infty.$$

Here $\mathbf{e}_q = (1, 1, \dots, 1) \in \mathbb{R}^q$. □

- Remark 1.*
- (i) Lemma 2 reduces the computation of condition numbers, mostly, to finding explicit expressions for $|DF(a)| |a|$ or, more precisely, matrix expressions for the derivative $DF(a)$. This will therefore be a major concern in the rest of this paper.
 - (ii) To simplify notation, in the rest of this paper we assume that every time we deal with componentwise condition numbers, the computed object has no zero components.

3. MOORE–PENROSE INVERSE

To use Lemma 2 for the Moore–Penrose inverse (of full-rank matrices) we introduce some notation. Consider the set

$$V = \{g \in \mathbb{R}^{mn} \mid g = \text{vec}(G), \text{ with } G \in \mathbb{R}^{m \times n}, \text{rank}(G) = n\}.$$

Note that the set $\{G \in \mathbb{R}^{m \times n} \mid \text{rank}(G) = n\}$ is open in $\mathbb{R}^{m \times n}$ since its complement is the union of the sets $\det(G_s) = 0$, where G_s runs over all $n \times n$ submatrices of G . Moreover, since vec is a homeomorphism between $\mathbb{R}^{m \times n}$ and \mathbb{R}^{mn} , it follows that V is open as well.

Now define the mapping $\phi : V \rightarrow \mathbb{R}^{mn}$ by $\phi(\text{vec}(G)) = \text{vec}(G^\dagger)$. By definition we have,

$$m_\dagger(A) = m(\phi; \text{vec}(A)), \quad \text{and} \quad c_\dagger(A) = c(\phi; \text{vec}(A)).$$

To make use of the above we would like to have an explicit expression for the derivative $D\phi$. Lemma 4 below exhibits such an expression. Its proof uses the following well-known result.

Lemma 3. ([2, 16, 17]) *Let $A \in \mathbb{R}^{m \times n}$ and suppose $\{A_k\}$ is a sequence of $m \times n$ matrices satisfying $\lim_{k \rightarrow \infty} A_k = A$. A necessary and sufficient condition for $\lim_{k \rightarrow \infty} A_k^\dagger = A^\dagger$ is*

$$\text{rank}(A_k) = \text{rank}(A)$$

for sufficiently large k . □

Lemma 4. *The mapping ϕ is continuous and Fréchet differentiable at a for all $a \in V$. Moreover, it has the matrix expression $D\phi(a) = M(A)$ where*

$$M(A) = \left[- \left(A^{\dagger T} \otimes A^{\dagger} \right) + \left((I_m - AA^{\dagger}) \otimes (A^T A)^{-1} \right) \Pi \right].$$

Here I_m denotes the $m \times m$ identity matrix.

Proof. The continuity of ϕ on V follows immediately from Lemma 3.

The following equation is well known (see [16, Page 150, eqn. 3.35]):

$$(A + \Delta A)^{\dagger} - A^{\dagger} = -A^{\dagger} \Delta A A^{\dagger} + (A^T A)^{-1} (\Delta A)^T (I_m - AA^{\dagger}) + \mathcal{O}(\|\Delta A\|^2).$$

Omitting the second-order terms, using the vec function, and denoting $a = \text{vec}(A)$ and $\delta a = \text{vec}(\Delta A)$, we have

$$\begin{aligned} \phi(a + \delta a) - \phi(a) &\approx \text{vec} \left(-A^{\dagger} \Delta A A^{\dagger} + (A^T A)^{-1} (\Delta A)^T (I_m - AA^{\dagger}) \right) \\ &= - \left(A^{\dagger T} \otimes A^{\dagger} \right) \text{vec}(\Delta A) + \left((I_m - AA^{\dagger}) \otimes (A^T A)^{-1} \right) \text{vec}(\Delta A^T) \\ &= \left[- \left(A^{\dagger T} \otimes A^{\dagger} \right) + \left((I_m - AA^{\dagger}) \otimes (A^T A)^{-1} \right) \Pi \right] \delta a, \end{aligned}$$

where the second line follows from (2.2) and Π is the vec -permutation matrix defined by (2.3).

So the Fréchet derivative of ϕ at a is given by $D\phi(a) = M(A)$. □

The main result in this section is the following theorem. It provides explicit expressions for the condition numbers we defined for the Moore–Penrose inverse computation.

Theorem 1. *Let $A \in \mathbb{R}^{m \times n}$ be such that $\text{rank}(A) = n$. Then*

- (a) $m_{\dagger}(A) = \frac{\| |M(A)| \text{vec}(|A|) \|_{\infty}}{\| \text{vec}(A^{\dagger}) \|_{\infty}},$
- (b) $c_{\dagger}(A) = \left\| \frac{|M(A)| \text{vec}(|A|)}{\text{vec}(A^{\dagger})} \right\|_{\infty}.$

Proof. By Lemmas 2 and 4,

$$m_{\dagger}(A) = m(\phi; a) = \frac{\| |D\phi(a)| |a| \|}{\| \phi(a) \|_{\infty}} = \frac{\| |M(A)| \text{vec}(|A|) \|_{\infty}}{\| \text{vec}(A^{\dagger}) \|_{\infty}}$$

and

$$c_{\dagger}(A) = c(\phi; a) = \left\| \frac{|M(A)| |a|}{| \phi(a) |} \right\|_{\infty} = \left\| \frac{|M(A)| \text{vec}(|A|)}{\text{vec}(|A^{\dagger}|)} \right\|_{\infty}. \quad \square$$

Theorem 1 gives explicit expressions for the condition numbers $m_{\dagger}(A)$ and $c_{\dagger}(A)$. While these expressions are sharp they may not be easy to compute by their dependence on the (large) matrix Π and the need to compute Kronecker products. The next corollary gives easier to compute upper bounds for these condition numbers.

Lemma 5. For any matrices M, N, P, Q, R , and S with dimensions making the following well defined

$$[M \otimes N + (P \otimes Q) \Pi] \text{vec}(R),$$

$$\frac{[M \otimes N + (P \otimes Q) \Pi] \text{vec}(R)}{S},$$

$$NRM^T \text{ and } QR^T P^T,$$

we have

$$\left\| \frac{[M \otimes N + (P \otimes Q) \Pi] \text{vec}(|R|)}{|S|} \right\|_{\infty} \leq \left\| \text{vec}(|N||R||M|^T + |Q||R|^T|P|^T) \right\|_{\infty}$$

and

$$\left\| \frac{[M \otimes N + (P \otimes Q) \Pi] \text{vec}(|R|)}{|S|} \right\|_{\infty} \leq \left\| \frac{\text{vec}(|N||R||M|^T + |Q||R|^T|P|^T)}{|S|} \right\|_{\infty}.$$

Proof. It is easy to see that

$$\begin{aligned} \left| [M \otimes N + (P \otimes Q) \Pi] \text{vec}(|R|) \right| &\leq [|M| \otimes |N| + (|P| \otimes |Q|) \Pi] \text{vec}(|R|) \\ &= (|M| \otimes |N|) \text{vec}(|R|) + (|P| \otimes |Q|) \text{vec}(|R|^T) \\ &= \text{vec}(|N||R||M|^T + |Q||R|^T|P|^T), \end{aligned}$$

the first line by (2.1) and the last by (2.2). Taking norms (and dividing by $|S|$ before doing so for the second inequality in the statement) proves the lemma. \square

Corollary 1. In the hypothesis of Theorem 1 we have

$$\begin{aligned} \text{(a) } m_{\dagger}(A) &\leq \frac{\left\| |A^{\dagger}| |A| |A^{\dagger}| + |(A^T A)^{-1}| |A^T| |I_m - AA^{\dagger}| \right\|_{\max}}{\|A^{\dagger}\|_{\max}}, \\ \text{(b) } c_{\dagger}(A) &\leq \left\| \frac{|A^{\dagger}| |A| |A^{\dagger}| + |(A^T A)^{-1}| |A^T| |I_m - AA^{\dagger}|}{A^{\dagger}} \right\|_{\max}. \end{aligned}$$

Proof. Using Theorem 1 and Lemma 5 with $M = -A^{\dagger T}$, $N = A^{\dagger}$, $P = I_m - AA^{\dagger}$, $Q = (A^T A)^{-1}$, $R = A$, and $S = \text{vec}(A^{\dagger})$, we obtain

$$\begin{aligned} m_{\dagger}(A) &\leq \frac{\left\| \text{vec}(|A^{\dagger}| |A| |A^{\dagger}| + |(A^T A)^{-1}| |A^T| |I_m - AA^{\dagger}|) \right\|_{\infty}}{\|\text{vec}(A^{\dagger})\|_{\infty}} \\ &= \frac{\left\| |A^{\dagger}| |A| |A^{\dagger}| + |(A^T A)^{-1}| |A^T| |I_m - AA^{\dagger}| \right\|_{\max}}{\|A^{\dagger}\|_{\max}} \end{aligned}$$

and

$$\begin{aligned} c_{\dagger}(A) &\leq \left\| \frac{\text{vec}(|A^{\dagger}| |A| |A^{\dagger}| + |(A^T A)^{-1}| |A^T| |I_m - AA^{\dagger}|)}{\text{vec}(A^{\dagger})} \right\|_{\infty} \\ &= \left\| \frac{|A^{\dagger}| |A| |A^{\dagger}| + |(A^T A)^{-1}| |A^T| |I_m - AA^{\dagger}|}{|A^{\dagger}|} \right\|_{\max}. \quad \square \end{aligned}$$

4. LINEAR LEAST SQUARES PROBLEMS

We consider linear least squares problems (LLS)

$$(4.1) \quad \min_{v \in \mathbb{R}^n} \|Av - b\|_2,$$

where $A \in \mathbb{R}^{m \times n}$, $\text{rank}(A) = n$, and $b \in \mathbb{R}^m$. Since A , as a linear map, is injective there is a unique minimizer x for (4.1).

This minimizer satisfies

$$A^T Ax = A^T b,$$

and therefore, since $A^\dagger = (A^T A)^{-1} A^T$ for full column rank matrices A ,

$$x = A^\dagger b = (A^T A)^{-1} A^T b.$$

Let x be as above, $\Delta b \in \mathbb{R}^m$, and $\Delta A \in \mathbb{R}^{m \times n}$ such that $\text{rank}(A + \Delta A) = n$. Consider the problem

$$(4.2) \quad \min_{w \in \mathbb{R}^n} \|(A + \Delta A)w - (b + \Delta b)\|_2.$$

Then there is a unique minimizer y and letting $\Delta x := y - x$, we have

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

The mixed and componentwise condition numbers for LLS are defined as follows:

$$m_{\text{ls}}(A, b) := \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|\Delta A| \leq \varepsilon |A| \\ |\Delta b| \leq \varepsilon |b|}} \frac{\|\Delta x\|_\infty}{\varepsilon \|x\|_\infty},$$

$$c_{\text{ls}}(A, b) := \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|\Delta A| \leq \varepsilon |A| \\ |\Delta b| \leq \varepsilon |b|}} \frac{1}{\varepsilon} \left\| \frac{\Delta x}{x} \right\|_\infty.$$

Just as in the previous section, to comfortably make use of Lemma 1, we define the mapping $\psi : V \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ by

$$\psi(g, f) := (\text{vec}^{-1} g)^\dagger f.$$

Note that $m_{\text{ls}}(A, b) = m(\psi; a, b)$ and $c_{\text{ls}}(A, b) = c(\psi; a, b)$.

Lemma 6. *The set $V \times \mathbb{R}^m$ is open and ψ is a continuous mapping on $V \times \mathbb{R}^m$. In addition, for all $(a, b) \in V \times \mathbb{R}^m$, ψ is Fréchet differentiable at (a, b) and has the matrix expression $D\psi(a, b) = [H(A, b), A^\dagger]$, where*

$$H(A, b) = - (x^T \otimes A^\dagger) + (A^T A)^{-1} \otimes r^T.$$

Proof. The first statement is trivial. We next proceed with the claimed equality. To do so, note that

$$(4.3) \quad \begin{aligned} \Delta x &\approx -A^\dagger \Delta A x + (A^T A)^{-1} (\Delta A)^T r + A^\dagger \Delta b \\ &= \left[- (x^T \otimes A^\dagger) + (r^T \otimes (A^T A)^{-1}) \Pi, A^\dagger \right] \begin{bmatrix} \delta a \\ \Delta b \end{bmatrix} \\ &= \left[- (x^T \otimes A^\dagger) + (A^T A)^{-1} \otimes r^T, A^\dagger \right] \begin{bmatrix} \delta a \\ \Delta b \end{bmatrix}, \end{aligned}$$

where the first line is shown in [11, (1.4)], the second follows from (2.2) and the fact that vec applied to a vector yields the vector itself, and the last line follows from (2.5).

From (4.3) we can rewrite

$$\psi(a + \Delta a, b + \Delta b) - \psi(a, b) \approx [-(x^T \otimes A^\dagger) + (A^T A)^{-1} \otimes r^T, A^\dagger] \begin{bmatrix} \delta a \\ \Delta b \end{bmatrix}.$$

Then the Fréchet derivative of ψ at (A, b) is

$$D\psi(a, b) = [-(x^T \otimes A^\dagger) + (A^T A)^{-1} \otimes r^T, A^\dagger]. \quad \square$$

We can now give expressions for the mixed and componentwise condition number of LLS. Recall, for the first we assume $x \neq 0$ and for the second $x_i \neq 0$ for $i = 1, \dots, n$.

Theorem 2. *Let $A \in \mathbb{R}^{m \times n}$, $\text{rank}(A) = n$, and $b \in \mathbb{R}^m$. We have*

$$m_{\text{ls}}(A, b) = \frac{\| |H(A, b)| \text{vec}(|A|) + |A^\dagger| |b| \|_\infty}{\|x\|_\infty},$$

$$c_{\text{ls}}(A, b) = \left\| \frac{|H(A, b)| \text{vec}(|A|) + |A^\dagger| |b|}{x} \right\|_\infty.$$

Furthermore, if $r = 0$ (i.e., for consistent linear systems $Ax = b$), we have

$$m_{\text{ls}}(A, b) = \frac{\| |A^\dagger| |A| |x| + |A^\dagger| |b| \|_\infty}{\|x\|_\infty},$$

$$c_{\text{ls}}(A, b) = \left\| \frac{|A^\dagger| |A| |x| + |A^\dagger| |b|}{x} \right\|_\infty.$$

Proof. By Lemmas 2 and 6

$$m_{\text{ls}}(A, b) = \frac{\left\| |D\psi(a, b)| \begin{bmatrix} |a| \\ |b| \end{bmatrix} \right\|_\infty}{\|x\|_\infty} = \frac{\left\| [|H(A, b)|, |A^\dagger|] \begin{bmatrix} |a| \\ |b| \end{bmatrix} \right\|_\infty}{\|x\|_\infty}$$

$$= \frac{\| |H(A, b)| \text{vec}(|A|) + |A^\dagger| |b| \|_\infty}{\|x\|_\infty}$$

and

$$c_{\text{ls}}(A, b) = \left\| \frac{|D\psi(a, b)| \begin{bmatrix} |a| \\ |b| \end{bmatrix}}{|x|} \right\|_\infty = \left\| \frac{|H(A, b)| \text{vec}(|A|) + |A^\dagger| |b|}{|x|} \right\|_\infty.$$

For consistent linear systems replace r by 0 in $H(A, b)$ to obtain

$$m_{\text{ls}}(A, b) = \frac{\| |A^\dagger| |A| |x| + |A^\dagger| |b| \|_\infty}{\|x\|_\infty} \quad \text{and} \quad c_{\text{ls}}(A, b) = \left\| \frac{|A^\dagger| |A| |x| + |A^\dagger| |b|}{|x|} \right\|_\infty. \quad \square$$

Remark 2. When $n = m$ the consistent case of Theorem 2 recovers the known expressions [10] for mixed and componentwise condition numbers for linear systems.

Corollary 2. *We have the following bounds:*

$$m_{\text{ls}}(A, b) \leq m_{\text{ls}}^{\text{upper}} := \frac{\| |A^\dagger| |A| |x| + |(A^T A)^{-1}| |A^T| |r| + |A^\dagger| |b| \|_\infty}{\|x\|_\infty},$$

$$c_{\text{ls}}(A, b) \leq c_{\text{ls}}^{\text{upper}} := \left\| \frac{|A^\dagger| |A| |x| + |(A^T A)^{-1}| |A^T| |r| + |A^\dagger| |b|}{|x|} \right\|_\infty.$$

Proof. Using Theorem 2, equality (2.5), and Lemma 5, we have

$$\begin{aligned}
 m_{\text{ls}}(A, b) &= \frac{\| |[-(x^T \otimes A^\dagger) + (r^T \otimes (A^T A)^{-1}) \Pi] \text{vec}(|A|) + |A^\dagger| |b| \|_\infty}{\|x\|_\infty} \\
 &\leq \frac{\| |A^\dagger| |A| |x| + |(A^T A)^{-1}| |A^T| |r| + |A^\dagger| |b| \|_\infty}{\|x\|_\infty}
 \end{aligned}$$

and

$$\begin{aligned}
 c_{\text{ls}}(A, b) &= \left\| \frac{|[(x^T \otimes A^\dagger) - (r^T \otimes (A^T A)^{-1}) \Pi] \text{vec}(|A|) + |A^\dagger| |b|}{|x|} \right\|_\infty \\
 &\leq \left\| \frac{|A^\dagger| |A| |x| + |(A^T A)^{-1}| |A^T| |r| + |A^\dagger| |b|}{|x|} \right\|_\infty. \quad \square
 \end{aligned}$$

Remark 3. In [3], Björck derived a bound for $m_{\text{ls}}(A, b)$. Namely, he proved that

$$m_{\text{ls}}(A, b) \leq \mathbf{M}(A, b) := \frac{\| |A^\dagger| |A| |x| + |A^\dagger| |b| \|_\infty + \left\| (A^T A)^{-1} |A^T| |r| \right\|_\infty}{\|x\|_\infty}.$$

We note that, by the triangle inequality, $m_{\text{ls}}^{\text{upper}} \leq \mathbf{M}(A, b)$.

Condition numbers bound the *worst-case* sensitivity of an input data only to *small* perturbations. If ε is the size of the perturbation, a term $\mathcal{O}(\varepsilon^2)$ is neglected and therefore, the bound only holds for ε small enough. One says that condition numbers are *first-order* bounds for these sensitivities. Occasionally, one is interested in bounds for *unrestricted* perturbations. The following result exhibits such unrestricted perturbation bounds for LLS.

Theorem 3. *Let $A, \Delta A \in \mathbb{R}^{m \times n}$ satisfy $\text{rank}(A) = \text{rank}(A + \Delta A) = n$. Let $\Delta b \in \mathbb{R}^m$ and x, y be the solutions of (4.1) and (4.2), respectively. If for some $E \in \mathbb{R}^{m \times n}$ and $f \in \mathbb{R}^m$ we have $|\Delta A| \leq \varepsilon E$ and $|\Delta b| \leq \varepsilon f$, then*

$$\begin{aligned}
 (4.4) \quad \frac{\|y - x\|_\infty}{\|x\|_\infty} &\leq \varepsilon \frac{\| |(y^T \otimes A^\dagger) - (A^T A)^{-1} \otimes s^T | \text{vec}(E) + |A^\dagger| f \|_\infty}{\|x\|_\infty}, \\
 \frac{\|s - r\|_\infty}{\|r\|_\infty} &\leq \varepsilon \frac{\| |y^T \otimes (I_m - AA^\dagger) + A^{\dagger T} \otimes s^T | \text{vec}(E) + |I_m - AA^\dagger| f \|_\infty}{\|r\|_\infty},
 \end{aligned}$$

where $s = b + \Delta b - (A + \Delta A)y$. Both inequalities are sharp.

To prove Theorem 3 we need some preparation.

Lemma 7. *Let $A, \Delta A, b, \Delta b, x, y, E$, and f be as in the hypothesis of Theorem 3. Then*

$$(4.5) \quad y - x = [-(y^T \otimes A^\dagger) + (A^T A)^{-1} \otimes s^T, A^\dagger] \begin{bmatrix} \delta a \\ \Delta b \end{bmatrix},$$

$$(4.6) \quad s - r = [-(y^T \otimes (I - AA^\dagger)) - (A^\dagger A)^T \otimes s^T, I_m - AA^\dagger] \begin{bmatrix} \delta a \\ \Delta b \end{bmatrix},$$

where $s = b + \Delta b - (A + \Delta A)y$.

Proof. In [10, Theorem 4.2] it is shown that

$$\begin{aligned} y - x &= -A^\dagger \Delta A y + (A^T A)^{-1} (\Delta A)^T s + A^\dagger \Delta b, \\ s - r &= -(I_m - AA^\dagger) \Delta A y - (A^\dagger)^T \Delta A^T s + (I_m - AA^\dagger) \Delta b. \end{aligned}$$

Then

$$\begin{aligned} y - x &= \text{vec} \left(-A^\dagger \Delta A y + (A^T A)^{-1} (\Delta A)^T s \right) + A^\dagger \Delta b \\ &= -(y \otimes A^\dagger) \text{vec}(\Delta A) + (s \otimes (A^T A)^{-1}) \text{vec}(\Delta A^T) + A^\dagger \Delta b \\ &= \left[-(y^T \otimes A^\dagger) + (A^T A)^{-1} \otimes s^T, A^\dagger \right] \begin{bmatrix} \delta a \\ \Delta b \end{bmatrix} \end{aligned}$$

and similarly for the second equality. □

For the proof of the following lemma, we use an idea from [6].

Lemma 8. *Let $A \in \mathbb{R}^{m \times n}$, $D = \text{Dg}(d) \in \mathbb{R}^{n \times n}$, and $v \in \mathbb{R}^n$ satisfying $|v_i| \leq \varepsilon |d_i|$ for $i = 1, \dots, n$. Then*

$$\left\| AD \frac{v}{d} \right\|_\infty \leq \varepsilon \|AD\|_\infty,$$

and v can be chosen to make the upper bound attainable.

Proof. The first inequality is easy to prove. For the second statement, let $i \leq m$ such that

$$\|AD\|_\infty = \sum_{k=1}^n |a_{ik} d_k|.$$

Define v by $v_k = \varepsilon \text{sgn}(a_{ik} d_k) d_k$, where sgn is the sign function, i.e., $\text{sgn}(x) = 1$ if $x \geq 0$ and $\text{sgn}(x) = -1$ otherwise. Then

$$\frac{v}{d} = \varepsilon \begin{bmatrix} \text{sgn}(A_{i1} d_1) \\ \vdots \\ \text{sgn}(A_{in} d_n) \end{bmatrix} \quad \text{and} \quad \left\| \frac{v}{d} \right\|_\infty = \varepsilon$$

and the i th row of $AD \frac{v}{d}$ is given by

$$\left(AD \frac{v}{d} \right)_i = \varepsilon [A_{i1} d_1 \text{sgn}(A_{i1} d_1), \dots, A_{in} d_n \text{sgn}(A_{in} d_n)] = \varepsilon [|A_{i1} d_1|, \dots, |A_{in} d_n|].$$

Therefore

$$\left\| \left(AD \frac{v}{d} \right)_i \right\|_\infty = \varepsilon \|AD\|_\infty. \quad \square$$

Proof of Theorem 3. By Lemma 7, writing $e = \text{vec}(E)$,

$$\begin{aligned} y - x &= \left[-(y^T \otimes A^\dagger) + (A^\dagger A^{\dagger T}) \otimes s^T, A^\dagger \right] \begin{bmatrix} \delta a \\ \Delta b \end{bmatrix} \\ &= \left[-(y^T \otimes A^\dagger) + (A^\dagger A^{\dagger T}) \otimes s^T, A^\dagger \right] \begin{bmatrix} \text{Dg}(e) & 0 \\ 0 & \text{Dg}(f) \end{bmatrix} \begin{bmatrix} \delta a \\ \frac{\delta b}{f} \end{bmatrix}. \end{aligned}$$

Taking norms and using Lemma 8, we obtain

$$\begin{aligned} \|y - x\|_\infty &\leq \varepsilon \left\| \left[-(y^T \otimes A^\dagger) + (A^\dagger A^{\dagger T}) \otimes s^T, A^\dagger \right] \begin{bmatrix} \text{Dg}(e) & 0 \\ 0 & \text{Dg}(f) \end{bmatrix} \right\|_\infty \\ &= \varepsilon \left\| \left[(y^T \otimes A^\dagger) - (A^T A)^{-1} \otimes s^T \right] \text{vec}(E) + |A^\dagger| f \right\|_\infty, \end{aligned}$$

the second line as in Lemma 2. This proves inequality (4.4). Sharpness follows from Lemma 8.

The second inequality can be proved similarly. □

In what follows we assume that $b \notin \mathcal{R}(A)$, where $\mathcal{R}(A)$ denotes the range of A . That is, the residual vector $r \neq 0$. For componentwise results, we also assume that $r_i \neq 0$ for $i = 1, \dots, m$. Define the mixed and componentwise condition numbers for r as

$$m_{\text{res}}(A, b) := \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|\Delta A| \leq \varepsilon |A| \\ |\Delta b| \leq \varepsilon |b|}} \frac{\|\Delta r\|_\infty}{\varepsilon \|r\|_\infty},$$

$$c_{\text{res}}(A, b) := \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|\Delta A| \leq \varepsilon |A| \\ |\Delta b| \leq \varepsilon |b|}} \frac{1}{\varepsilon} \left\| \frac{\Delta r}{r} \right\|_\infty.$$

Define the function $\Phi : V \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ by

$$\Phi(g, f) := (I_m - (\text{vec}^{-1}g)(\text{vec}^{-1}g)^\dagger)f.$$

Then $m_{\text{res}}(A, b) = m(\Phi; a, b)$ and $c_{\text{res}}(A, b) = c(\Phi; a, b)$.

Lemma 9. *The function Φ is continuous. Moreover, for all $(a, b) \in V \times \mathbb{R}^m$, it is Fréchet differentiable at (a, b) and has the matrix expression $D\Phi(a, b) = [Q(A, b), I_m - AA^\dagger]$ where*

$$Q(A, b) = - (x^T \otimes (I_m - AA^\dagger)) + (A^\dagger)^T \otimes r^T.$$

Proof. Let $(\delta a, \Delta b)$ be a perturbation of (a, b) and $\Delta A = \text{vec}^{-1}(\delta a)$. It is easy to see that, omitting second- (and higher-) order terms, the perturbed residual vector $r + \Delta r$ satisfies

$$\begin{aligned} r + \Delta r &= b + \Delta b - (A + \Delta A)(A + \Delta A)^\dagger(b + \Delta b) \\ &\approx r - (I_m - AA^\dagger)\Delta Ax - (A^\dagger)^T(\Delta A)^T r + (I_m - AA^\dagger)\Delta b, \end{aligned}$$

i.e., using (2.2),

$$\Delta r \approx \left[- (x^T \otimes (I_m - AA^\dagger)) + (A^\dagger)^T \otimes r^T, I_m - AA^\dagger \right] \begin{bmatrix} \delta a \\ \Delta b \end{bmatrix}.$$

So

$$D\Phi(a, b) = \left[- (x^T \otimes (I_m - AA^\dagger)) + (A^\dagger)^T \otimes r^T, I_m - AA^\dagger \right]. \quad \square$$

Theorem 4. *With the notation above, the mixed and componentwise condition numbers for r satisfy*

$$m_{\text{res}}(A, b) = \frac{\| |Q(A, b)| \text{vec}(|A|) + |I_m - AA^\dagger| |b| \|_\infty}{\|r\|_\infty},$$

$$c_{\text{res}}(A, b) = \left\| \frac{|Q(A, b)| \text{vec}(|A|) + |I_m - AA^\dagger| |b|}{r} \right\|_\infty.$$

Proof. From Lemmas 2 and 9, we have that

$$\begin{aligned} m_{\text{res}}(A, b) = m(\Phi; a, b) &= \frac{\left\| \left| D\Phi(a, b) \right| \begin{bmatrix} |a| \\ |b| \end{bmatrix} \right\|_{\infty}}{\|\Phi(a, b)\|_{\infty}} \\ &= \frac{\left\| |Q(A, b)| \text{vec}(|A|) + |I_m - AA^{\dagger}| |b| \right\|_{\infty}}{\|r\|_{\infty}} \end{aligned}$$

and

$$\begin{aligned} c_{\text{res}}(A, b) = c(\Phi; a, b) &= \left\| \frac{\left| D\Phi(a, b) \right| \begin{bmatrix} |a| \\ |b| \end{bmatrix}}{\Phi(a, b)} \right\|_{\infty} \\ &= \left\| \frac{\left| |Q(A, b)| \text{vec}(|A|) + |I_m - AA^{\dagger}| |b| \right|}{r} \right\|_{\infty}. \quad \square \end{aligned}$$

The next corollary gives (easier to compute, no occurrences of \otimes) upper bounds for the residual vector mixed and componentwise condition numbers.

Corollary 3. *Let $A \in \mathbb{R}^{m \times n}$ satisfy $\text{rank}(A) = n$. Let $b \in \mathbb{R}^m$, $b \notin \mathcal{R}(A)$, x be the solution of (4.1) and $r = b - Ax$. Then*

$$\begin{aligned} m_{\text{res}}(A, b) &\leq m_{\text{res}}^{\text{upper}}(A, b) \\ &:= \frac{\left\| |I_m - AA^{\dagger}| |A| |x| + |(A^{\dagger})^T| |A|^T |r| + |I_m - AA^{\dagger}| |b| \right\|_{\infty}}{\|r\|_{\infty}}, \\ c_{\text{res}}(A, b) &\leq c_{\text{res}}^{\text{upper}}(A, b) \\ &:= \left\| \frac{|I_m - AA^{\dagger}| |A| |x| + |(A^{\dagger})^T| |A|^T |r| + |I_m - AA^{\dagger}| |b|}{|r|} \right\|_{\infty}. \end{aligned}$$

Proof. Using Theorem 4, equality (2.5) and Lemma 5, we obtain

$$\begin{aligned} m_{\text{res}}(A, b) &= \frac{\left\| [x^T \otimes (I_m - AA^{\dagger}) - (r^T \otimes (A^{\dagger})^T) \Pi] \text{vec}(|A|) + |I_m - AA^{\dagger}| |b| \right\|_{\infty}}{\|r\|_{\infty}} \\ &\leq \frac{\left\| |I_m - AA^{\dagger}| |A| |x| + |(A^{\dagger})^T| |A|^T |r| + |I_m - AA^{\dagger}| |b| \right\|_{\infty}}{\|r\|_{\infty}} \end{aligned}$$

and

$$\begin{aligned} c_{\text{res}}(A, b) &= \left\| \frac{\left| [x^T \otimes (I_m - AA^{\dagger}) - (r^T \otimes (A^{\dagger})^T) \Pi] \text{vec}(|A|) + |I_m - AA^{\dagger}| |b| \right|}{|r|} \right\|_{\infty} \\ &\leq \left\| \frac{|I_m - AA^{\dagger}| |A| |x| + |(A^{\dagger})^T| |A|^T |r| + |I_m - AA^{\dagger}| |b|}{|r|} \right\|_{\infty}. \quad \square \end{aligned}$$

5. FULL ROW RANK AND UNDERDETERMINED SYSTEMS

Suppose $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = m$. Then the Moore–Penrose inverse of A can be written as

$$A^{\dagger} = A^T (AA^T)^{-1}.$$

Let

$$U = \{g \in \mathbb{R}^{mn} \mid g = \text{vec}(G), \text{ with } G \in \mathbb{R}^{m \times n}, \text{rank}(G) = m\}.$$

Now define the mapping $\Psi : U \rightarrow \mathbb{R}^{mn}$ given by $\Psi(\text{vec}(G)) = \text{vec}(G^\dagger)$.

Lemma 10. *The mapping Ψ is continuous. Moreover, for all $a \in U$, it is Fréchet differentiable at a and has the matrix expression $D\Psi(a) = N(A)$ where*

$$N(A) = \left[- \left(A^{\dagger T} \otimes A^\dagger \right) + \left((AA^T)^{-1} \otimes (I_n - A^\dagger A) \right) \Pi \right].$$

Proof. The continuity of ϕ on V follows immediately from Lemma 3.

The following equality is well known (see [16, p. 150, eqn. 3.35]):

$$(A + \Delta A)^\dagger - A^\dagger = -A^\dagger \Delta A A^\dagger + (I_n - A^\dagger A)(\Delta A)^T (AA^T)^{-1} + \mathcal{O}(\|\Delta A\|^2).$$

Omitting the second-order term and using the vec function, we have

$$\begin{aligned} \Psi(a + \delta a) - \Psi(a) &= \text{vec} \left(-A^\dagger \Delta A A^\dagger + (I_n - A^\dagger A)(\Delta A)^T (AA^T)^{-1} \right) \\ &= - \left(A^{\dagger T} \otimes A^\dagger \right) \text{vec}(\Delta A) + \left((AA^T)^{-1} \otimes (I_n - A^\dagger A) \right) \text{vec}(\Delta A^T) \\ &= \left[- \left(A^{\dagger T} \otimes A^\dagger \right) + \left((AA^T)^{-1} \otimes (I_n - A^\dagger A) \right) \Pi \right] \text{vec}(\Delta A), \end{aligned}$$

where Π is the vec-permutation matrix defined by (2.3).

So the Fréchet derivative of ϕ at a is given by

$$D\Psi(a) = \left[- \left(A^{\dagger T} \otimes A^\dagger \right) + \left((AA^T)^{-1} \otimes (I_n - A^\dagger A) \right) \Pi \right]. \quad \square$$

The next result exhibits expressions for the condition numbers of A (for the Moore–Penrose inverse).

Theorem 5. *Let $A \in \mathbb{R}^{m \times n}$ be such that $\text{rank}(A) = m$. Then*

$$\begin{aligned} m_\dagger(A) &= \frac{\| |N(A)| \text{vec}(|A|) \|_\infty}{\| \text{vec}(A^\dagger) \|_\infty}, \\ c_\dagger(A) &= \left\| \frac{|N(A)| \text{vec}(|A|)}{\text{vec}(A^\dagger)} \right\|_\infty. \end{aligned}$$

Proof. From Lemmas 2 and 10 we get

$$m_\dagger(A) = m(\Psi; a) = \frac{\| |D\Psi(a)| |a| \|_\infty}{\| \Psi(a) \|_\infty} = \frac{\| |N(A)| \text{vec}(|A|) \|_\infty}{\| \text{vec}(A^\dagger) \|_\infty}$$

and

$$c_\dagger(A) = c(\Psi; a) = \left\| \frac{|D\Psi(a)| |a|}{\Psi(a)} \right\|_\infty = \left\| \frac{|N(A)| \text{vec}(|A|)}{\text{vec}(A^\dagger)} \right\|_\infty. \quad \square$$

Corollary 4. *Let $A \in \mathbb{R}^{m \times n}$ be such that $\text{rank}(A) = m$. Then*

$$\begin{aligned} m_\dagger(A) &\leq \frac{\| |A^\dagger| |A| |A^\dagger| + |I_n - A^\dagger A| |A^T| \|(AA^T)^{-1}\|_{\max}}{\|A^\dagger\|_{\max}}, \\ c_\dagger(A) &\leq \left\| \frac{|A^\dagger| |A| |A^\dagger| + |I_n - A^\dagger A| |A^T| \|(AA^T)^{-1}\|_{\max}}{A^\dagger} \right\|_{\max}. \end{aligned}$$

Proof. Using Theorem 5 and Lemma 5, we have

$$\begin{aligned} m_{\dagger}(A) &\leq \frac{\|\text{vec}(|A^{\dagger}||A||A^{\dagger}| + |I_n - A^{\dagger}A||A^T|(AA^T)^{-1})\|_{\infty}}{\|\text{vec}(A^{\dagger})\|_{\infty}} \\ &= \frac{\||A^{\dagger}||A||A^{\dagger}| + |I_n - A^{\dagger}A||A^T|(AA^T)^{-1}\|_{\max}}{\|A^{\dagger}\|_{\max}} \end{aligned}$$

and

$$\begin{aligned} c_{\dagger}(A) &\leq \left\| \frac{\text{vec}(|A^{\dagger}||A||A^{\dagger}| + |I_n - A^{\dagger}A||A^T|(AA^T)^{-1})}{\text{vec}(A^{\dagger})} \right\|_{\infty} \\ &= \left\| \frac{|A^{\dagger}||A||A^{\dagger}| + |I_n - A^{\dagger}A||A^T|(AA^T)^{-1}}{A^{\dagger}} \right\|_{\max}. \quad \square \end{aligned}$$

For underdetermined systems

$$Av = b,$$

where $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = m$ and $b \in \mathbb{R}^m$, the set of solutions is an affine subspace of \mathbb{R}^n with the dimension of $\mathcal{N}(A)$, the nullspace of A . It contains a unique point x minimizing the 2-norm. It is well known that this solution is $x = A^{\dagger}b$. Consider the problem of, given A and b , finding x . This problem induces condition numbers $m_{\min}(A, b)$ and $c_{\min}(A, b)$. The following result exhibits expressions for these condition numbers.

We define the function $\chi : U \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ by

$$\chi(g, f) := (\text{vec}^{-1}g)^{\dagger}f.$$

Then $m_{\min}(A, b) = m(\chi; a, b)$ and $c_{\min}(A, b) = c(\chi; a, b)$. Also, we assume that $x \neq 0$ and, for componentwise results, that $x_i \neq 0$ for $i = 1, \dots, n$.

Lemma 11. *The function χ is continuous. Moreover, for all (a, b) in $U \times \mathbb{R}^m$, it is Fréchet differentiable at (a, b) and $D\chi(a, b) = [P(A, b), A^{\dagger}]$, where*

$$P(A, b) = -x^T \otimes A^{\dagger} + (I_n - A^{\dagger}A) \otimes (x^T A^{\dagger}).$$

Proof. We have

$$\begin{aligned} \Delta x &\approx -A^{\dagger}\Delta Ax + (I_n - A^{\dagger}A)(\Delta A)^T(A^{\dagger})^T x + A^{\dagger}\Delta b \\ &= \text{vec}(-A^{\dagger}\Delta Ax + (I_n - A^{\dagger}A)(\Delta A)^T(A^{\dagger})^T x + A^{\dagger}\Delta b) \\ &= [-x^T \otimes A^{\dagger} + (x^T A^{\dagger}) \otimes (I_n - A^{\dagger}A)\Pi] \text{vec}(\Delta A) + A^{\dagger}\Delta b \\ &= [-x^T \otimes A^{\dagger} + (x^T A^{\dagger}) \otimes (I_n - A^{\dagger}A)\Pi, A^{\dagger}] \begin{bmatrix} \delta a \\ \Delta b \end{bmatrix} \\ &= [-x^T \otimes A^{\dagger} + (I_n - A^{\dagger}A) \otimes (x^T A^{\dagger}), A^{\dagger}] \begin{bmatrix} \delta a \\ \Delta b \end{bmatrix}, \end{aligned}$$

where the first line is proved in [4, Theorem 2.1]. □

Theorem 6. *Let $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = m$ and $b \in \mathbb{R}^m$. Then*

$$\begin{aligned} m_{\min}(A, b) &= \frac{\||P(A, b)| \text{vec}(|A|) + |A^{\dagger}||b|\|_{\infty}}{\|x\|_{\infty}}, \\ c_{\min}(A, b) &= \left\| \frac{|P(A, b)| \text{vec}(|A|) + |A^{\dagger}||b|}{x} \right\|_{\infty}. \end{aligned}$$

Proof. From Lemmas 2 and 11, we have

$$m_{\min}(A, b) = m(\chi; a, b) = \frac{\left\| |D\chi(a, b)| \begin{bmatrix} |a| \\ |b| \end{bmatrix} \right\|_{\infty}}{\|\chi(a, b)\|_{\infty}} = \frac{\| |P(A, b)| \text{vec}(|A|) + |A^{\dagger}| |b| \|_{\infty}}{\|x\|_{\infty}}$$

and

$$c_{\dagger}(A) = c(\chi; a, b) = \left\| \frac{|D\chi(a, b)| \begin{bmatrix} |a| \\ |b| \end{bmatrix}}{\chi(a, b)} \right\|_{\infty} = \left\| \frac{|P(A, b)| \text{vec}(|A|) + |A^{\dagger}| |b|}{x} \right\|_{\infty}. \quad \square$$

Corollary 5. *Let $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = m$ and $b \in \mathbb{R}^m$. Then*

$$m_{\min}(A, b) \leq m_{\min}^{\text{upper}}(A, b) := \frac{\| |A^{\dagger}| |A| |x| + |I_n - A^{\dagger} A| |A^T| |A^{\dagger T} x| + |A^{\dagger}| |b| \|_{\infty}}{\|x\|_{\infty}},$$

$$c_{\min}(A, b) \leq c_{\min}^{\text{upper}}(A, b) := \left\| \frac{|A^{\dagger}| |A| |x| + |I_n - A^{\dagger} A| |A^T| |A^{\dagger T} x| + |A^{\dagger}| |b|}{x} \right\|_{\infty}.$$

Proof. Using Theorem 6, (2.5), and Lemma 5,

$$m_{\min}(A, b) = \frac{\| |[(x^T \otimes A^{\dagger}) - ((x^T A^{\dagger}) \otimes (I_n - A^{\dagger} A)) \Pi] \text{vec}(|A|) + |A^{\dagger}| |b| \|_{\infty}}{\|x\|_{\infty}}$$

$$\leq \frac{\| |A^{\dagger}| |A| |x| + |I_n - A^{\dagger} A| |A^T| |A^{\dagger T} x| + |A^{\dagger}| |b| \|_{\infty}}{\|x\|_{\infty}}$$

and

$$c_{\min}(A, b) = \left\| \frac{|[(x^T \otimes A^{\dagger}) - ((x^T A^{\dagger}) \otimes (I_n - A^{\dagger} A)) \Pi] \text{vec}(|A|) + |A^{\dagger}| |b|}{x} \right\|_{\infty}$$

$$\leq \left\| \frac{|A^{\dagger}| |A| |x| + |I_n - A^{\dagger} A| |A^T| |A^{\dagger T} x| + |A^{\dagger}| |b|}{x} \right\|_{\infty}. \quad \square$$

REFERENCES

1. M. Arioli, I.S. Duff and P.P.M. de Rijk, *An augmented system approach to sparse least-squares problems*, Numer. Math. 55(1989), pp. 667-684. MR1005065 (90g:65048)
2. A. Ben-Israel and T.N.E. Greville, *Generalized Inverses: Theory and Applications*, 2nd Edition, Springer Verlag, New York, 2003. MR1987382 (2004b:15008)
3. Å. Björck, *Component-wise perturbation analysis and error bounds for linear least squares solutions*, BIT, 31(1991), pp. 238-244. MR1112220 (92i:65079)
4. J. Demmel and N. Higham, *Improved error bounds for underdetermined system solvers*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1-14. MR1199540 (94c:65050)
5. A.J. Geurts, *A Contribution to the theory of condition*, Numer. Math., 39(1982), pp. 85-96. MR0664538 (83g:65046)
6. I. Gohberg and I. Koltracht, *Mixed, componentwise, and structured condition numbers*, SIAM J. Matrix Anal. Appl., 14(1993), pp. 688-704. MR1227773 (94j:65062)
7. A. Graham, *Kronecker Products and Matrix Calculus with Application*, Wiley, New York, 1981. MR0640865 (83g:15001)
8. S. Gratton, *On the condition number of linear least squares problems in a weighted Frobenius norm*, BIT, 36(1996), no.3, pp. 523-530. MR1410095 (97h:65050)
9. J.F. Grcar, *Optimal sensitivity analysis of linear least squares*, Lawrence Berkeley National Laboratory, Report LBNL-52434, 2003.

10. N.J. Higham, *A survey of componentwise perturbation theory in numerical linear algebra*, Proceedings of Symposia in Applied Mathematics, Vol.48, 1994, pp. 49-77. MR1314843 (96a:65065)
11. A.N. Malyshev, *A unified theory of conditioning for linear least squares and Tikhonov regularization solutions*, SIAM J. Matrix Anal. Appl., 24(2003), no.4, pp. 1186-1196. MR2003329 (2004f:65052)
12. J.R. Rice, *A theory of condition*, SIAM J. Numer. Anal., 3(1966), pp. 217-232. MR0211576 (35:2454)
13. J. Rohn, *New condition numbers for matrices and linear systems*, Computing, 41(1989), pp. 167-169. MR0981682 (90a:65104)
14. R.D. Skeel, *Scaling for numerical stability in Gaussian elimination*, J. Assoc. Comput. Mach., 26(1979), No.3, pp. 817-526. MR0535268 (80e:65051)
15. G.W. Stewart, *On the perturbation of pseudo-inverses, projections and linear least squares problems*, SIAM Rev., 19 (1977), pp. 634-662. MR0461871 (57:1854)
16. G.W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990. MR1061154 (92a:65017)
17. G. Wang, Y. Wei and S. Qiao, *Generalized Inverses: Theory and Computations*, Science Press, Beijing/New York, 2004.
18. P.Å. Wedin, *Perturbation theory for pseudo-inverses*, BIT, 13(1973), pp. 217-232.

DEPARTMENT OF MATHEMATICS, CITY UNIVERSITY OF HONG KONG, 83 TAT CHEE AVENUE,
KOWLOON TONG, HONG KONG, P.R. OF CHINA
E-mail address: macucker@math.cityu.edu.hk

DEPARTMENT OF MATHEMATICS, CITY UNIVERSITY OF HONG KONG, 83 TAT CHEE AVENUE,
KOWLOON TONG, HONG KONG, P.R. OF CHINA
E-mail address: 50007445@student.cityu.edu.hk

SCHOOL OF MATHEMATICAL SCIENCES, FUDAN UNIVERSITY, SHANGHAI 200433 AND KEY LAB-
ORATORY OF MATHEMATICS FOR NONLINEAR SCIENCES (FUDAN UNIVERSITY), MINISTRY OF EDU-
CATION, P.R. OF CHINA
E-mail address: ymwei@fudan.edu.cn