Condition Number for Weighted Linear Least Squares Problem and Its Condition Number *

Yimin Wei[⋆] Huaian Diao[†] Sanzheng Qiao[†]

[⋆] Department of Mathematics, Fudan University,
Shanghai, 200433, P.R. of China.
Email: ymwei@fudan.edu.cn

[†] Institute of Mathematics, Fudan University,
Shanghai, 200433, P.R. of China.
Email: hadiao78@yahoo.com

[†] Department of Computing and Software, McMaster University,
Hamilton, Ontario, L8S 4L7, Canada.
Email: qiao@mcmaster.ca

Abstract

In this paper, we investigate the condition numbers for the generalized matrix inversion and the rank deficient linear least squares problem: $\min_x \|Ax - b\|_2$, where A is an m-by-n $(m \geq n)$ rank deficient matrix. We first derive an explicit expression for the condition number in the weighted Frobenius norm $\| [AT, \beta b] \|_F$ of the data A and b, where T is a positive diagonal matrix and β is a positive scalar. We then discuss the sensitivity of the standard 2-norm condition numbers for the generalized matrix inversion and rank deficient least squares and establish relations between the condition numbers and their condition numbers called level-2 condition numbers.

AMS Subject Classifications: 15A12, 65F20.
Keywords: Moore-Penrose inverse, condition number, linear least squares.

1 Introduction

In this paper, we consider a condition number for the linear least squares (LLS) problem

$$\min_x \|Ax - b\|_2,$$

*The first and second author are supported by the National Natural Science Foundation of China under grant 19901006 and Shanghai Education Committee. The third author is partially supported by Natural Science and Engineering Research Council of Canada.

where $A \in R^{m\times n}$ $(m \geq n)$ is a rank deficient matrix. The condition number for the LLS problem with full rank is well studied [3]. In the standard 2-norm analysis, the condition number is defined as

$$\operatorname{cond}(A,b) = \lim_{\epsilon\to 0^+} \sup \left\{ \frac{\|(A+E)^\dagger(b+f) - A^\dagger b\|_2}{\epsilon\, \|A^\dagger b\|_2}, \quad \|E\|_2 \leq \epsilon\|A\|_2,\ \|f\|_2 \leq \epsilon\|b\|_2 \right\},$$

where $A^\dagger$ is the Moore-Penrose inverse of A defined as the unique matrix X satisfying

$$AXA = A, \quad XAX = X, \quad (AX)^{\mathrm{T}} = AX, \quad \text{and} \quad (XA)^{\mathrm{T}} = XA,$$

where A^{T} is the transpose of A [7].

The condition number discussed in this section is based on a general theory of condition introduced by Rice [6]. In the context of the LLS, the problem is viewed as a mapping from a pair (A,b) to the LLS solution $x_{\mathrm{LS}} = A^\dagger b$. The norm of a pair (A,b) in the domain of the mapping is defined by the weighted Frobenius norm:

$$\|[AT \quad \beta b]\|_{\mathrm{F}}, \tag{1.1}$$

where T is positive and diagonal and $\beta > 0$. The weights T and β provide flexibility. Later, we will show that a large diagonal of T allows perturbation on b only and a large β allows perturbation on A only. The norm of a solution x in the image of the mapping is chosen as the Euclidean norm $\|x\|_2$.

In Rice's theory of condition, an absolute δ-condition is first defined by:

$$\mu_\delta = \inf\{\sigma \mid \|[ET \quad \beta f]\|_{\mathrm{F}} \leq \delta \Rightarrow \|(A+E)^\dagger(b+f) - A^\dagger b\|_2 \leq \sigma\delta\}. \tag{1.2}$$

This definition says the image of a δ-neighborhood of a pair (A,b) is contained in a $\sigma\delta$-neighborhood of the solution $A^\dagger b$. So, σ is an upper bound for the magnification of the mapping and μ_δ is the least upper bound. Then, the asymptotic absolute condition number for the weighted LLS problem in the norms chosen above is

$$\mu = \lim_{\delta\to 0} \mu_\delta.$$

The relative condition number is defined by

$$\nu = \frac{\|[AT \quad \beta b]\|_{\mathrm{F}}}{\|A^\dagger b\|_2}\mu.$$

As explained above, similar to the standard condition number, the δ-condition in (1.2) measures the enlargement of the mapping from (A,b) to $A^\dagger b$. What is different from the standard condition number is that the weighted Frobenius norm is used in the domain space of pairs (A,b).

Gratton [4] considered the case when $T = \alpha I$ $(\alpha > 0)$, and A is of full column rank and showed that

$$\mu = \|A^\dagger\|_2 \sqrt{\beta^{-2} + \alpha^{-2}(\|x_{\mathrm{LS}}\|_2^2 + \|A^\dagger\|_2^2\|r\|_2^2)}\,, \tag{1.3}$$

where $r = b - Ax_{\mathrm{LS}}$ is the residual, and when $\alpha = \beta = 1$, the relative condition number

$$\nu = \frac{\|A^\dagger\|_2 \|[A \;\; b]\|_{\mathrm{F}}}{\|x_{\mathrm{LS}}\|_2} \sqrt{\|x_{\mathrm{LS}}\|_2^2 + \|A^\dagger\|_2^2 \|r\|_2^2 + 1}\,.$$

It is also shown in [4] that a large α allows perturbation on b only and a large β allows perturbation on A only.

In this paper, we consider the case when A is rank deficient under the condition that the perturbation E on A satisfies

$$\mathrm{range}(E) \subseteq \mathrm{range}(A) \quad \text{and} \quad \mathrm{range}(E^{\mathrm{T}}) \subseteq \mathrm{range}(A^{\mathrm{T}}), \tag{1.4}$$

where range(E) denotes the column space of E.

The rest of the paper is organized as follows. The absolute and relative condition numbers in the weighted Frobenius norm are given in Section 2. Then, in Section 3, we analyze the sensitivity of the generalized matrix inversion condition number and the rank deficient LLS condition number, called level-2 condition numbers introduced by Higham [5].

2 Condition Numbers

In this section, we present explicit expressions for the absolute and relative condition numbers for the rank deficient LLS problem in the weighted Frobenius norm described in the previous section.

Theorem 2.1 *Suppose the perturbation E in A satisfies the conditions (1.4), the absolute condition number of the rank deficient LLS problem in the weighted Frobenius norm (1.1) on the data A and b and the Euclidean norm on the solution x_{LS} is*

$$\mu = \|A^\dagger\|_2 \sqrt{\beta^{-2} + \|T^{-1} x_{LS}\|_2^2}. \tag{2.1}$$

Proof. From [1], when E is small ($\|A^\dagger\|_2 \|E\|_2 < 1$) and satisfies the conditions (1.4), we have

$$(A + E)^\dagger = (I + A^\dagger E)^{-1} A^\dagger. \tag{2.2}$$

Thus, for small E and f, the linear term in $(A+E)^\dagger (b+f) - A^\dagger b$ is

$$-A^\dagger E A^\dagger b + A^\dagger f = -A^\dagger (E x_{\mathrm{LS}} - f) = -A^\dagger (ET(T^{-1} x_{\mathrm{LS}}) - \beta^{-1}(\beta f)),$$

which implies that

$$\|A^\dagger (E x_{\mathrm{LS}} - f)\|_2^2 = \|A^\dagger (E x_{\mathrm{LS}} - f)\|_{\mathrm{F}}^2 \le \|A^\dagger\|_2^2 (\|ET\|_{\mathrm{F}}^2 \|T^{-1} x_{\mathrm{LS}}\|_2^2 + \beta^{-2} \|\beta f\|_2^2).$$

It then follows that if

$$\|[ET \;\; \beta f]\|_{\mathrm{F}} = \sqrt{\|ET\|_{\mathrm{F}}^2 + \|\beta f\|_2^2} \le \delta,$$

then

$$\|A^\dagger(Ex_{\rm LS}-f)\|_2 \le \delta\|A^\dagger\|_2\sqrt{\|T^{-1}x_{LS}\|_2^2+\beta^{-2}}\,.$$

Since $-A^\dagger(Ex_{\rm LS}-f)$ is the linear term in $(A+E)^\dagger(b+f)-A^\dagger b$, the absolute condition number is bounded above by

$$\mu=\lim_{\delta\to 0}\mu_\delta \le \|A^\dagger\|_2\sqrt{\|T^{-1}x_{LS}\|_2^2+\beta^{-2}}\,. \tag{2.3}$$

In the following we will show that this upper bound is reachable. We will first construct perturbations E_0 and f_0, then show that the linear term $\|A^\dagger(E_0x_{\rm LS}-f_0)\|_2$ equals $\delta\|A^\dagger\|_2\sqrt{\|T^{-1}x_{\rm LS}\|_2^2+\beta^{-2}}$. This proves the theorem since (1.2) says that μ_δ is the minimal upper bound for all perturbations E and f and $\mu=\lim_{\delta\to 0}\mu_\delta$. In particular, let $\mathrm{rank}(A)=r<n$ and u and v be respectively the left and right singular vectors corresponding to the smallest positive singular value σ_r of A, then $\sigma_r^{-1}=\|A^\dagger\|_2$ and

$$A^\dagger u=\|A^\dagger\|_2\,v.$$

Constructing

$$E_0=-\frac{\delta}{\eta}u(T^{-2}x_{\rm LS})^{\rm T} \quad\text{and}\quad f_0=\frac{\delta}{\beta^2\eta}u,$$

where $\eta=\sqrt{\|T^{-1}x_{\rm LS}\|_2^2+\beta^{-2}}$, we have

$$\mathrm{range}(E_0)\subseteq\mathrm{range}(u)\subseteq\mathrm{range}(A),\quad \mathrm{range}(E_0^{\rm T})\subseteq\mathrm{range}(x_{\rm LS})\subseteq\mathrm{range}(A^\dagger)=\mathrm{range}(A^{\rm T}),$$

and

$$\begin{aligned}
\|[E_0T\ \ \beta f_0]\|_{\rm F}^2 &= \|[\frac{\delta}{\eta}u(T^{-1}x_{\rm LS})^{\rm T}\ \ \frac{\delta}{\beta\eta}u]\|_{\rm F}^2\\
&= \frac{\delta^2}{\eta^2}\|u[(T^{-1}x_{\rm LS})^{\rm T}\ \ \beta^{-1}]\|_{\rm F}^2\\
&= \frac{\delta^2}{\eta^2}\|u\|_2^2\eta^2\\
&= \delta^2.
\end{aligned}$$

Now, for E_0 and f_0, the linear term

$$\begin{aligned}
&-A^\dagger E_0x_{\rm LS}+A^\dagger f_0\\
=\ &\frac{\delta}{\eta}A^\dagger u(T^{-2}x_{\rm LS})^{\rm T}x_{\rm LS}+\frac{\delta}{\beta^2\eta}A^\dagger u\\
=\ &\frac{\delta}{\eta}A^\dagger u(\|T^{-1}x_{\rm LS}\|_2^2+\beta^{-2})\\
=\ &\delta\eta\|A^\dagger\|_2\,v,
\end{aligned}$$

which implies that

$$\|A^\dagger E_0 x_{\rm LS} - A^\dagger f_0\|_2 = \delta \|A^\dagger\|_2 \sqrt{\|T^{-1}x_{\rm LS}\|_2^2 + \beta^{-2}}\,.$$

This completes the proof of the theorem. □

Corollary 2.1 *Taking $T = I$ and $\beta = 1$ in the condition number μ of Theorem 2.1 gives the case where both A and b are equally perturbed. By letting $T = \alpha I$, where $\alpha > 0$, and $\alpha \to \infty$ ($\beta \to \infty$), no perturbation on the matrix A (on the right-hand side b) is permitted.*

Proof. The perturbations E and f must satisfy $\|[ET \quad \beta f]\|_{\rm F} \le \delta$. Therefore, $T = \alpha I$ and $\alpha \to \infty$ imply $E = 0$, that is no perturbation on A. Similarly, $\beta \to \infty$ implies no perturbation on b. □

Using the definition of the relative condition number ν we can get the following formula.

Corollary 2.2 *When the perturbation E in A satisfies the conditions (1.4), the relative condition number with equal perturbations on A and b ($T = I$ and $\beta = 1$) is*

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

Note that comparing (2.1) with (1.3) in the full rank case, the residual term $\|A^\dagger\|_2^2 \|r\|_2^2$ is missing in our μ in (2.1). The reason is that when the perturbation E satisfies the condition (1.4), $E^{\rm T} r = 0$.

3 Condition Number Sensitivity

In practice, the computed condition number is the exact condition number for a perturbed problem. How sensitive is the condition number to the perturbation on the data? Demmel [2] introduced the concept of condition numbers of the condition numbers, called level-2 condition numbers by Higham [5]. Demmel [2] showed that for certain problems, the level-2 condition number is the condition number up to a constant factor. In this section, we show that the level-2 condition numbers for the generalized matrix inversion and rank deficient least squares are in the same magnitude order of their corresponding condition numbers.

We begin with the definitions of the standard condition numbers. The condition number for the generalized matrix inversion is defined by

$$\operatorname{cond}(A) = \lim_{\epsilon \to 0^+} \sup \left\{ \frac{\|(A+E)^\dagger - A^\dagger\|_2}{\epsilon \|A^\dagger\|_2}, \quad \|E\|_2 \le \epsilon \|A\|_2 \right\}, \tag{3.1}$$

where E satisfies the conditions (1.4). The standard condition number for the least squares is

$$\operatorname{cond}(A, b) = \lim_{\epsilon \to 0^+} \sup \left\{ \frac{\|(A+E)^\dagger (b+f) - A^\dagger b\|_2}{\epsilon\, \|A^\dagger b\|_2}, \quad \|E\|_2 \le \epsilon \|A\|_2,\ \|f\|_2 \le \epsilon \|b\|_2 \right\}, \tag{3.2}$$

where E satisfies the conditions (1.4).

Wei and Wang [8, Corollaries 2.1, 3.1] derived

$$\mathrm{cond}(A) = \|A^\dagger\|_2 \|A\|_2, \quad \text{and} \quad \mathrm{cond}(A,b) = \|A^\dagger\|_2\|A\|_2 + \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2}. \tag{3.3}$$

The computed condition number $\mathrm{cond}(A)$ can be regarded as the exact condition number $\mathrm{cond}(A+E)$ for some small perturbation E. Thus, we define the level-2 condition number for the generalized matrix inversion:

$$\mathrm{cond}^{[2]}(A) = \lim_{\epsilon \to 0^+} \sup \left\{ \frac{|\mathrm{cond}(A+E) - \mathrm{cond}(A)|}{\epsilon\, \mathrm{cond}(A)}, \quad \|E\|_2 \le \epsilon \|A\|_2 \right\}, \tag{3.4}$$

where E satisfies the conditions (1.4).

Similarly, we define the level-2 condition number for the least squares:

$$\begin{aligned} & \mathrm{cond}^{[2]}(A,b) \\ = \; & \lim_{\epsilon \to 0^+} \sup \left\{ \frac{|\mathrm{cond}(A+E, b+f) - \mathrm{cond}(A,b)|}{\epsilon\, \mathrm{cond}(A,b)}, \; \|E\|_2 \le \epsilon\|A\|_2, \|f\|_2 \le \epsilon \|b\|_2 \right\}, \end{aligned} \tag{3.5}$$

where E satisfies the conditions (1.4) and $b \in \mathrm{range}(A)$.

In the following, we will show that under certain conditions on the perturbations E and f, $\mathrm{cond}^{[2]}(A)$ and $\mathrm{cond}^{[2]}(A,b)$ are the same as $\mathrm{cond}(A)$ and $\mathrm{cond}(A,b)$ respectively up to constant factors.

Before deriving the level-2 condition numbers, we present a useful bound for $\|(A+E)^\dagger\|_2$.

Lemma 3.1 *Under the conditions (1.4),*

$$\|(A+E)^\dagger\|_2 = \sup\left\{ \|A^\dagger\|_2 (1 + \epsilon\, \mathrm{cond}(A)) + O(\epsilon^2), \quad \|E\|_2 \le \epsilon \|A\|_2 \right\}$$

for small $\epsilon > 0$.

Proof. From (2.2), under the conditions (1.4) and $\|E\|_2 \le \epsilon\|A\|_2$,

$$\|(A+E)^\dagger\|_2 = \|A^\dagger - A^\dagger E A^\dagger\|_2 + O(\epsilon^2)$$

for small $\epsilon > 0$.

On the one hand, since $\|E\|_2 \le \epsilon \|A\|_2$,

$$\|A^\dagger - A^\dagger E A^\dagger\|_2 \le \|A^\dagger\|_2 (1 + \epsilon \|A\|_2 \|A^\dagger\|_2) = \|A^\dagger\|_2(1 + \epsilon\, \mathrm{cond}(A)).$$

On the other hand, we construct an E_0 such that $\|A^\dagger - A^\dagger E_0 A^\dagger\|_2 \ge \|A^\dagger\|_2(1 + \epsilon\, \mathrm{cond}(A))$. Let u be the rth left singular vector of A, where $r = \mathrm{rank}(A)$, then

$$\|A^\dagger u\|_2 = \sigma_r^{-1} = \|A^\dagger\|_2.$$

Defining
$$v = -A^{\dagger}u/\|A^{\dagger}\|_2,$$
we have $\|v\|_2 = 1$, $v \in \mathrm{range}(A^{\dagger}) = \mathrm{range}(A^{\mathrm{T}})$, and
$$v^{\mathrm{T}}A^{\dagger}u = -\|A^{\dagger}u\|_2^2/\|A^{\dagger}\|_2 = -\|A^{\dagger}\|_2.$$
Now, we construct
$$E_0 = \epsilon\|A\|_2 uv^{\mathrm{T}},$$
then $E_0A^{\dagger}u = -\epsilon\|A\|_2\|A^{\dagger}\|_2 u$. Also, it can be verified that $\|E_0\|_2 = \epsilon\|A\|_2$, and E_0 satisfies the conditions (1.4), since $\mathrm{range}(E_0) \subseteq \mathrm{range}(u) \subseteq \mathrm{range}(A)$ and $\mathrm{range}(E_0^{\mathrm{T}}) \subseteq \mathrm{range}(v) \subseteq \mathrm{range}(A^{\mathrm{T}})$. Finally, applying $E_0A^{\dagger}u = -\epsilon\|A\|_2\|A^{\dagger}\|_2 u$ and $\|A^{\dagger}u\|_2 = \|A^{\dagger}\|_2$, we get
$$\begin{aligned} \|A^{\dagger} - A^{\dagger}E_0A^{\dagger}\|_2 &\geq \|(A^{\dagger} - A^{\dagger}E_0A^{\dagger})u\|_2 \\ &= \|A^{\dagger}u + \epsilon\|A\|_2\|A^{\dagger}\|_2 A^{\dagger}u\|_2 \\ &= \|A^{\dagger}\|_2(1 + \epsilon\,\mathrm{cond}(A)), \end{aligned}$$
which completes the proof. □

The following theorem shows that the level-2 condition number for the generalized matrix inversion is about the same as the condition number.

Theorem 3.1 *Under the conditions (1.4), the difference between the level-2 condition number* $\mathrm{cond}^{[2]}(A)$ *and the condition number* $\mathrm{cond}(A)$ *is bounded by:*
$$\left|\mathrm{cond}^{[2]}(A) - \mathrm{cond}(A)\right| \leq 1.$$

Proof. Following the first equation in (3.3), we consider
$$\mathrm{cond}(A+E) = \|A+E\|_2\|(A+E)^{\dagger}\|_2.$$
The inequality $\|E\|_2 \leq \epsilon\|A\|_2$ implies that $\|A+E\|_2 \leq (1+\epsilon)\|A\|_2$. Using Lemma 3.1, we have
$$\begin{aligned} \mathrm{cond}(A+E) &= \|A+E\|_2\|(A+E)^{\dagger}\|_2 \\ &\leq (1+\epsilon)\|A\|_2(\|A^{\dagger}\|_2(1+\epsilon\,\mathrm{cond}(A)) + O(\epsilon^2)) \\ &= \mathrm{cond}(A)(1+\epsilon\,\mathrm{cond}(A)+\epsilon) + O(\epsilon^2). \end{aligned} \tag{3.6}$$
It then follows that
$$\frac{\mathrm{cond}(A+E) - \mathrm{cond}(A)}{\epsilon\,\mathrm{cond}(A)} \leq \mathrm{cond}(A) + 1 + O(\epsilon). \tag{3.7}$$

On the other hand, $\|E\|_2 \leq \epsilon\|A\|_2$ also implies that $\|A+E\|_2 \geq (1-\epsilon)\|A\|_2$. Again, from Lemma 3.1, there exists an E_0 such that
$$\begin{aligned} \mathrm{cond}(A+E_0) &= \|A+E_0\|_2\|(A+E_0)^{\dagger}\|_2 \\ &\geq (1-\epsilon)\|A\|_2(\|A^{\dagger}\|_2(1+\epsilon\,\mathrm{cond}(A)) + O(\epsilon^2)) \\ &= \mathrm{cond}(A)(1+\epsilon\,\mathrm{cond}(A)-\epsilon) + O(\epsilon^2), \end{aligned} \tag{3.8}$$

which implies

$$\frac{\mathrm{cond}(A+E_0)-\mathrm{cond}(A)}{\epsilon\,\mathrm{cond}(A)} \geq \mathrm{cond}(A)-1+O(\epsilon). \tag{3.9}$$

Combining (3.7) and (3.9) proves the theorem. □

Next, we present the relations between the level-2 condition number $\mathrm{cond}^{[2]}(A,b)$ for the least squares and the condition number $\mathrm{cond}(A,b)$.

Theorem 3.2 *Under the conditions (1.4), the level-2 condition number* $\mathrm{cond}^{[2]}(A,b)$ *for the least squares defined in (3.5) is bounded by:*

$$\frac{\mathrm{cond}(A,b)}{(1+\gamma)^2}-\frac{1}{1+\gamma} \leq \mathrm{cond}^{[2]}(A,b) \leq 2\,\mathrm{cond}(A,b),$$

where $\gamma=\|b\|_2/\|AA^\dagger b\|_2$, *the secant of the angle between* b *and the projection* $AA^\dagger b$.

Proof. Following the second equation in (3.3), we first consider

$$\mathrm{cond}(A+E,b+f)=\mathrm{cond}(A+E)+\frac{\|(A+E)^\dagger\|_2\|b+f\|_2}{\|(A+E)^\dagger(b+f)\|_2}.$$

Applying $\|b+f\|_2 \leq (1+\epsilon)\|b\|_2$ and Lemma 3.1, we have

$$\|(A+E)^\dagger\|_2\|b+f\|_2 \leq \|A^\dagger\|_2\|b\|_2(1+\epsilon\,\mathrm{cond}(A)+\epsilon)+O(\epsilon^2).$$

Using the definition (3.2), we get

$$\begin{aligned}
\frac{1}{\|(A+E)^\dagger(b+f)\|_2} &\leq \frac{1}{\|A^\dagger b\|_2-\|(A+E)^\dagger(b+f)-A^\dagger b\|_2} \\
&= \|A^\dagger b\|_2^{-1}\frac{1}{1-\frac{\|(A+E)^\dagger(b+f)-A^\dagger b\|_2}{\|A^\dagger b\|_2}} \\
&= \frac{1}{\|A^\dagger b\|_2}\left(1+\frac{\|(A+E)^\dagger(b+f)-A^\dagger b\|_2}{\|A^\dagger b\|_2}\right)+O(\epsilon^2) \\
&\leq \frac{1}{\|A^\dagger b\|_2}(1+\epsilon\,\mathrm{cond}(A,b))+O(\epsilon^2).
\end{aligned}$$

Consequently,

$$\frac{\|(A+E)^\dagger\|_2\|b+f\|_2}{\|(A+E)^\dagger(b+f)\|_2} \leq \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2}(1+\epsilon\,\mathrm{cond}(A,b)+\epsilon\,\mathrm{cond}(A)+\epsilon)+O(\epsilon^2).$$

Thus, from (3.6),

$$\begin{aligned}
\mathrm{cond}(A+E,b+f) &\leq \mathrm{cond}(A)(1+\epsilon\,\mathrm{cond}(A)+\epsilon) \\
&+ \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2}(1+\epsilon\,\mathrm{cond}(A,b)+\epsilon\,\mathrm{cond}(A)+\epsilon)+O(\epsilon^2).
\end{aligned}$$

It then follows from the second equation in (3.3) that

$$\begin{aligned}
& \operatorname{cond}(A+E, b+f) - \operatorname{cond}(A, b) \\
\leq\ & \epsilon \operatorname{cond}(A)(\operatorname{cond}(A)+1) + \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2}\,\epsilon(\operatorname{cond}(A,b)+\operatorname{cond}(A)+1) + O(\epsilon^2) \\
\leq\ & \epsilon \operatorname{cond}(A,b)(\operatorname{cond}(A,b)+\operatorname{cond}(A)+1) + O(\epsilon^2).
\end{aligned}$$

Thus, we get

$$\frac{\operatorname{cond}(A+E, b+f) - \operatorname{cond}(A, b)}{\epsilon \operatorname{cond}(A, b)} \leq \operatorname{cond}(A,b) + \operatorname{cond}(A) + 1 + O(\epsilon). \tag{3.10}$$

Now, we derive a lower bound for $\operatorname{cond}(A+E, b+f)$. The condition $\|f\|_2 \leq \epsilon\|b\|_2$ implies $\|b+f\|_2 \geq (1-\epsilon)\|b\|_2$. Moreover,

$$\|(A+E)^\dagger\|_2 \geq \|A^\dagger\|_2 - \epsilon \operatorname{cond}(A)\|A^\dagger\|_2 + O(\epsilon^2), \tag{3.11}$$

since, from (2.2),

$$\begin{aligned}
\|A^\dagger\|_2 &= \|A^\dagger - A^\dagger E A^\dagger + A^\dagger E A^\dagger\|_2 \\
&\leq \|(I + A^\dagger E)^{-1} A^\dagger\|_2 + \epsilon\|A^\dagger\|_2^2\|A\|_2 + O(\epsilon^2) \\
&= \|(A+E)^\dagger\|_2 + \epsilon \operatorname{cond}(A)\|A^\dagger\|_2 + O(\epsilon^2).
\end{aligned}$$

Thus,

$$\begin{aligned}
\|(A+E)^\dagger\|_2\|b+f\|_2 &\geq (\|A^\dagger\|_2 - \epsilon \operatorname{cond}(A)\|A^\dagger\|_2 + O(\epsilon^2))(1-\epsilon)\|b\|_2 \\
&= \|A^\dagger\|_2\|b\|_2(1 - \epsilon \operatorname{cond}(A) - \epsilon) + O(\epsilon^2).
\end{aligned}$$

Applying the definition (3.2), we get

$$\begin{aligned}
\frac{1}{\|(A+E)^\dagger(b+f)\|_2} &\geq \frac{1}{\|A^\dagger b\|_2 + \|(A+E)^\dagger(b+f) - A^\dagger b\|_2} \\
&= \frac{1}{\|A^\dagger b\|_2}\left(1 - \frac{\|(A+E)^\dagger(b+f) - A^\dagger b\|_2}{\|A^\dagger b\|_2}\right) + O(\epsilon^2) \\
&\geq \frac{1}{\|A^\dagger b\|_2}(1 - \epsilon \operatorname{cond}(A,b)) + O(\epsilon^2).
\end{aligned}$$

We then have

$$\frac{\|(A+E)^\dagger\|_2\|b+f\|_2}{\|(A+E)^\dagger(b+f)\|_2} \geq \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2}(1 - \epsilon \operatorname{cond}(A) - \epsilon \operatorname{cond}(A,b) - \epsilon) + O(\epsilon^2).$$

From (3.11) and $\|A+E\|_2 \geq (1-\epsilon)\|A\|_2$,

$$\begin{aligned}
\|(A+E)^\dagger\|_2\|A+E\|_2 &\geq (\|A^\dagger\|_2 - \epsilon \operatorname{cond}(A)\|A^\dagger\|_2)(1-\epsilon)\|A\|_2 + O(\epsilon^2) \\
&= \operatorname{cond}(A)(1 - \epsilon \operatorname{cond}(A) - \epsilon) + O(\epsilon^2).
\end{aligned}$$

We then get a lower bound for $\mathrm{cond}(A+E, b+f)$:

$$\begin{aligned} & \mathrm{cond}(A+E, b+f) \\ \geq\ & \mathrm{cond}(A)(1-\epsilon\,\mathrm{cond}(A)-\epsilon) + \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2}(1-\epsilon\,\mathrm{cond}(A)-\epsilon\,\mathrm{cond}(A,b)-\epsilon)+O(\epsilon^2), \end{aligned}$$

which leads to

$$\begin{aligned} & \mathrm{cond}(A+E, b+f) - \mathrm{cond}(A,b) \\ =\ & \mathrm{cond}(A+E, b+f) - \mathrm{cond}(A) - \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2} \\ \geq\ & -\epsilon\left(\mathrm{cond}(A) + \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2}\right)(\mathrm{cond}(A)+\mathrm{cond}(A,b)+1) + \epsilon\,\mathrm{cond}(A)\mathrm{cond}(A,b) + O(\epsilon^2) \\ =\ & -\epsilon\,\mathrm{cond}(A,b)(\mathrm{cond}(A)+\mathrm{cond}(A,b)+1) + \epsilon\,\mathrm{cond}(A)\mathrm{cond}(A,b) + O(\epsilon^2) \\ \geq\ & -\epsilon\,\mathrm{cond}(A,b)(\mathrm{cond}(A)+\mathrm{cond}(A,b)+1) + O(\epsilon^2). \end{aligned}$$

Therefore,

$$\frac{\mathrm{cond}(A+E, b+f) - \mathrm{cond}(A,b)}{\epsilon\,\mathrm{cond}(A,b)} \geq -(\mathrm{cond}(A)+\mathrm{cond}(A,b)+1)+O(\epsilon). \tag{3.12}$$

Combining (3.10) and (3.12), we get

$$\frac{|\mathrm{cond}(A+E, b+f) - \mathrm{cond}(A,b)|}{\epsilon\,\mathrm{cond}(A,b)} \leq \mathrm{cond}(A)+\mathrm{cond}(A,b)+1+O(\epsilon).$$

Finally,

$$\mathrm{cond}^{[2]}(A,b) \leq \mathrm{cond}(A)+\mathrm{cond}(A,b)+1 \leq 2\,\mathrm{cond}(A,b),$$

since

$$\mathrm{cond}(A,b) = \|A^\dagger\|_2\|A\|_2 + \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2} \geq \mathrm{cond}(A)+1.$$

In the following, we derive a lower bound for $\mathrm{cond}^{[2]}(A,b)$ defined in (3.5). Using (3.3), we have

$$\begin{aligned} & \mathrm{cond}(A+E, b+f) - \mathrm{cond}(A,b) \\ =\ & \mathrm{cond}(A+E) + \frac{\|(A+E)^\dagger\|_2\|b+f\|_2}{\|(A+E)^\dagger(b+f)\|_2} - \mathrm{cond}(A) - \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2}. \end{aligned}$$

Let $f=0$, then

$$\frac{\|(A+E)^\dagger\|_2\|b+f\|_2}{\|(A+E)^\dagger(b+f)\|_2} - \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2} = \frac{\|(A+E)^\dagger\|_2\|A^\dagger b\|_2 - \|A^\dagger\|_2\|(A+E)^\dagger b\|_2}{\|(A+E)^\dagger b\|_2\|A^\dagger b\|_2}\|b\|_2.$$

From Lemma 3.1, for any $\epsilon > 0$, we can find an E_0 such that

$$\|(A+E_0)^\dagger\|_2 \geq \|A^\dagger\|_2(1+\epsilon\,\mathrm{cond}(A)) + O(\epsilon^2)$$

and, from (2.2), we have

$$\|(A+E)^\dagger b\|_2 = \|(I+A^\dagger E)^{-1}A^\dagger b\|_2 \leq (1+\epsilon\,\mathrm{cond}(A))\|A^\dagger b\|_2 + O(\epsilon^2).$$

Thus

$$\|(A+E_0)^\dagger\|_2\|A^\dagger b\|_2 - \|A^\dagger\|_2\|(A+E_0)^\dagger b\|_2 \geq O(\epsilon^2),$$

which means that for $E = E_0$ and $f = 0$,

$$\frac{\|(A+E)^\dagger\|_2\|b+f\|_2}{\|(A+E)^\dagger(b+f)\|_2} - \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2} \geq O(\epsilon^2).$$

Hence, for $E = E_0$ and $f = 0$,

$$\mathrm{cond}(A+E, b+f) - \mathrm{cond}(A,b) \geq \mathrm{cond}(A+E) - \mathrm{cond}(A) + O(\epsilon^2). \tag{3.13}$$

On the other hand, let $E = E_0$ be given in Lemma 3.1, then, from (3.9),

$$\mathrm{cond}(A+E_0) - \mathrm{cond}(A) \geq \epsilon\,\mathrm{cond}(A)(\mathrm{cond}(A)-1) + O(\epsilon^2). \tag{3.14}$$

From (3.3) and the inequality

$$\|A^\dagger b\|_2 \geq \frac{\|AA^\dagger b\|_2}{\|A\|_2},$$

we have

$$\begin{aligned}
\mathrm{cond}(A,b) &= \mathrm{cond}(A) + \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2} \\
&\leq \mathrm{cond}(A) + \frac{\|A\|_2\|A^\dagger\|_2\|b\|_2}{\|AA^\dagger b\|_2} \\
&= \mathrm{cond}(A)\left(1 + \frac{\|b\|_2}{\|AA^\dagger b\|_2}\right).
\end{aligned}$$

Consequently, for $E = E_0$ and $f = 0$, using (3.13) and (3.14), we get

$$\begin{aligned}
&\frac{\mathrm{cond}(A+E,b+f) - \mathrm{cond}(A,b)}{\epsilon\,\mathrm{cond}(A,b)} \\
\geq\ & \frac{\mathrm{cond}(A+E) - \mathrm{cond}(A)}{\epsilon\,\mathrm{cond}(A,b)} + O(\epsilon) \\
\geq\ & \frac{(\mathrm{cond}(A)-1)\mathrm{cond}(A)}{\mathrm{cond}(A,b)} + O(\epsilon) \\
\geq\ & \frac{\mathrm{cond}(A)-1}{1+\|b\|_2/\|AA^\dagger b\|_2} + O(\epsilon) \\
\geq\ & \frac{\mathrm{cond}(A,b)}{(1+\|b\|_2/\|AA^\dagger b\|_2)^2} - \frac{1}{1+\|b\|_2/\|AA^\dagger b\|_2} + O(\epsilon).
\end{aligned}$$

Defining $\gamma = \|b\|_2/\|AA^\dagger b\|_2$, which is the secant of the angle between b and the projection $AA^\dagger b$, we claim that

$$\frac{\mathrm{cond}(A,b)}{(1+\gamma)^2} - \frac{1}{1+\gamma} \geq 0, \quad \text{equivalently} \quad \mathrm{cond}(A,b) \geq 1+\gamma.$$

Indeed, using

$$\mathrm{cond}(A) \geq 1, \qquad \|A^\dagger b\|_2 \leq \|A^\dagger\|_2 \|AA^\dagger b\|_2,$$

and the second equation in (3.3), we get

$$\mathrm{cond}(A,b) = \mathrm{cond}(A) + \frac{\|A^\dagger\|_2\|b\|_2}{\|A^\dagger b\|_2} \geq 1 + \frac{\|b\|_2}{\|AA^\dagger b\|_2} = 1+\gamma.$$

Thus, from the definition (3.5) of $\mathrm{cond}^{[2]}(A,b)$, we have the lower bound:

$$\mathrm{cond}^{[2]}(A,b) \geq \frac{\mathrm{cond}(A,b)}{(1+\gamma)^2} - \frac{1}{1+\gamma}. \qquad \Box$$

When the residual $\|r\|_2$ is small, $\|b\|_2 \approx \|AA^\dagger b\|_2$ since $\|b\|_2^2 = \|r\|_2^2 + \|AA^\dagger b\|_2^2$, then the lower bound is approximately

$$\mathrm{cond}^{[2]}(A,b) \geq \frac{\mathrm{cond}(A,b)}{4} - \frac{1}{2}.$$

This theorem shows that the level-2 condition number $\mathrm{cond}^{[2]}(A,b)$ is almost the same as the condition number $\mathrm{cond}(A,b)$, up to a small constant.

Finally we note that in practice, the computed condition number $\mathrm{cond}(A) = \|A\|_2\|A^\dagger\|_2$ for the generalized matrix inversion of A is actually $\|A+E_1\|_2\|(A+E_2)^\dagger\|_2$, where the perturbations E_1 and E_2 may be different. We can generalize the definition (3.4) to

$$\overline{\mathrm{cond}^{[2]}}(A) = \lim_{\epsilon\to 0^+} \sup \left\{ \frac{\|A+E_1\|_2\|(A+E_2)^\dagger\|_2 - \mathrm{cond}(A)}{\epsilon\,\mathrm{cond}(A)}, \quad \|E_1\|_2, \|E_2\|_2 \leq \epsilon\|A\|_2 \right\}.$$

We can show that $\overline{\mathrm{cond}^{[2]}(A)}$, as $\mathrm{cond}^{[2]}(A)$ in Lemma 3.1, is also essentially same as $\mathrm{cond}(A)$. Specifically, let $E_1 = \epsilon A$ and $E_2 = E_0$ given in Lemma 3.1, then, from Lemma 3.1, we get

$$\begin{aligned} \|A+E_1\|_2\|(A+E_2)^\dagger\|_2 &= (1+\epsilon)\|A\|_2\|A^\dagger\|_2(1+\epsilon\,\mathrm{cond}(A)) + O(\epsilon^2) \\ &= \mathrm{cond}(A) + \epsilon\,\mathrm{cond}(A)(1+\mathrm{cond}(A)) + O(\epsilon^2). \end{aligned}$$

It then follows that

$$\overline{\mathrm{cond}^{[2]}}(A) = \mathrm{cond}(A) + 1.$$

Conclusion In this paper, we first present explicit expressions for the absolution and relative condition numbers for the rank deficient least squares problems in the weighted Frobenius norm. Because the problem is rank deficient, we impose the conditions (1.4) on the perturbation matrix E. As a consequence, our condition numbers are independent of the residual. We then analyze the level-2 condition numbers for the generalized matrix inversion and rank deficient least squares problem in 2-norm. We show that the level-2 condition numbers are essentially the same as their corresponding condition numbers.

References

[1] A. Ben-Israel, On error bounds for generalized inverses, *SIAM J. Numer. Anal.*, 3(1966) 585–592.

[2] J.W. Demmel, On condition numbers and the distance the nearest ill-posed problem, Numer. Math., 51 (1987), 251-289.

[3] G.H. Golub and C.F. Van Loan, *Matrix Computations,* 3rd Edition, Johns Hopkins University, Baltimore, 1996.

[4] S. Gratton, On the condition number for linear least squares problems in a weighted Frobenius norm, BIT 36(1996), 523-530.

[5] D.J. Higham, Condition numbers and their condition numbers. Linear Algebra Appl., 214 (1995), 193-213.

[6] J. R. Rice, A theory of condition, SIAM J. Numer. Anal., 3 (1966), 287-310.

[7] G. Wang, Y. Wei and S. Qiao, *Generalized Inverses: Theory and Computations,* Science Press, Beijing, 2004.

[8] Y. Wei and D. Wang, Condition numbers and perturbation of the weighted Moore-Penrose inverse and weighted linear least squares problem. Appl. Math. Comput., 145 (2003), 45-58.