Condition numbers of the mixed least squares-total least squares problem revisited

Qiaohua Liu\textsuperscript{a}, Qian Zhang\textsuperscript{a} and Dongmei Shen\textsuperscript{b}

\textsuperscript{a}Department of Mathematics, Shanghai University, Shanghai, People’s Republic of China; \textsuperscript{b}School of Statistics and Mathematics, Shanghai Lixin University of Accounting and Finance, Shanghai, People’s Republic of China

**ABSTRACT**

A recent study on the condition numbers of the mixed least squares-total least squares (MTLS) problem is due to Zheng and Yang (Numer Linear Algebra Appl. 2019;26(4):e2239). However, the associated expressions are not compact and the Kronecker-product operations make the computation costly. In this paper, we first present new and alternative closed formula for the first order perturbation estimate and condition numbers of the MTLS solution. Then we reveal the relationship between the new formula and Zheng and Yang’s result. Several new computable formulae and perturbation bounds for the normwise condition number of the MTLS solution are also provided. Finally, mixed and componentwise condition numbers, structured condition numbers are investigated. Through a number of tests, they are shown to be tighter than the normwise condition numbers for sparse and structured problems.

**1. Introduction**

Consider the overdetermined linear system

$$Ax \approx b,$$

where $A \in \mathbb{R}^{m \times n}$ has full column-rank, $b \in \mathbb{R}^m$, and $m \geq n$. In some engineering applications, the data matrix $A$ and the observation vector $b$ are contaminated by some noise, and the total least squares (TLS) model \cite{1,2} is often used to find the best approximations in the Frobenius norm such that

$$\min_{E,f} \|E, f\|_F \quad \text{subject to} \quad (A + E)x = b + f. \quad (1)$$

A vector $x = x_{\text{TLS}}$ satisfying (1) is called a TLS solution. If some rows in $[A, b]$ are free of error, i.e. some rows in $[E, f]$ are set to zero, then the corresponding problem reduces to the total least squares problem with the equality constraint \cite{3,4}. If some columns of $A$ are
known exactly, i.e. some columns in $E$ are set to zero, then the corresponding problem (1) is known as the mixed least squares and total least squares (MTLS) problem [5–7]. It arises in the regression analysis [8], system identification [9] and signal processing [10], etc. The solution to the MTLS problem is denoted by $x_M$. As usual we assume that $Ax = b$ is not consistent, otherwise the best minimizer $[E_*, f_*]$ in (1) is taken to be the zero matrix.

For standard TLS problems with small or medium size, a classical direct solver is based on the singular value decomposition (SVD) [1,2,11–13], where the solution is obtained by normalizing the right singular vector corresponding to the smallest singular value of $[A, b]$. For the solution of the MTLS problem, by the QR factorization of $A$, the problem reduces to a smaller TLS problem that can be solved via the SVD [12]. A method of weighting, due to Liu and Wang [5], interprets the MTLS solution as the limit of the solution to an unconstrained weighted TLS (WTLS) problem, as the positive parameter in the weight matrix tends to zero. This observation allows the application of conventional theories and algorithms for the TLS problem to the WTLS problem, which leads to the Cholesky factorization-based inverse iteration and Rayleigh quotient iteration methods [5] for the solution of the MTLS problem.

The condition number of a problem is fundamental since it measures the worst-case sensitivity of its solution to small perturbations in the input data. The condition numbers of the TLS problem and its extension to the scaled TLS problem have been studied widely, e.g. by Zhou et al. [14], Baboulin and Gratton [15], Li and Jia [16,17] and Wang et al. [18]. The equivalence of the formulae in [14,15,17] for the normwise condition numbers of the TLS solution was revealed by Xie et al. [19]. Mixed and componentwise condition numbers for the TLS and truncated TLS problems were further studied in [20,21], while the condition numbers of multidimensional TLS problems was due to Zheng et al. and Meng et al. [22,23].

The early work on the perturbation analysis of the MTLS problem can be traced back to Paige and Wei [6], Van Huffel and Vandewalle [24] in 1990s. Since then, little attention has been paid to the conditioning of the MTLS problem. Until 2019, by a limit argument on the weighting method [5] for the MTLS problem, Zheng and Yang [25] studied a closed formula for the first order perturbation estimate of the MTLS solution and gave some explicit expressions for the condition number of the MTLS problem. However, no compact form of the normwise condition number was obtained.

In this paper, we revisit the condition numbers of the MTLS problem. With a different approach, we derive another different closed formula for the perturbation estimate and obtain the condition numbers of the MTLS solution. The relation between the new formula and the one in [25] will be revealed. To reduce the computational cost, compact and computable formulae for the condition numbers and their upper bounds are also provided. Moreover, we present structured condition numbers for the MTLS problem and verify the theoretical results through a number of numerical tests.

Before our discussion, some notations are required. $\mathbb{R}^m$ and $\mathbb{R}^{m\times n}$ denote the spaces of $m \times 1$ and $m \times n$ real matrices, respectively. $I_n$, $O_{m\times n}$ and $O_n$ denote the $n \times n$ identity matrix, $m \times n$ and $n \times n$ zero matrices, respectively. If subscripts are ignored, the sizes of identity and zero matrices are clear from the context. $\| \cdot \|_2$, $\| \cdot \|_\infty$ and $\| \cdot \|_F$ denote the 2-norm, $\infty$-norm and Frobenius norm of their arguments, respectively. For a matrix $A$, $|A|$ is a matrix by taking the entrywise absolute value, $A^T$, $a_j$ and $\sigma_j(A)$ denote the transpose, the $j$th column and the $i$-th largest singular value of $A$; vec$(A)$ stacks the columns of $A$ one underneath the other. The Kronecker product of two arbitrary matrices $A$ and $B$ is denoted
by $A \otimes B = [a_i b_j]$ and it has the following properties \[26,27\]:

$$
|A \otimes B| = |A| \otimes |B|, \quad (A \otimes B)^T = A^T \otimes B^T,
$$

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

Let $x = \phi(a)$ be a continuous and Fréchet differentiable mapping from $\mathbb{R}^p$ to $\mathbb{R}^q$. For small perturbations $\delta a$, denote $\delta x = \phi(a + \delta a) - \phi(a)$. According to \[28–31\], the general normwise condition number $\kappa(\phi, a)$, mixed condition number $m(\phi, a)$ and componentwise condition number $c(\phi, a)$ are defined and formulated as follows:

$$
\kappa(\phi, a) := \lim_{\varepsilon \to 0} \sup_{\|\delta a\|_2 \leq \varepsilon \|a\|_2} \frac{\|\delta x\|_2/\|x\|_2}{\|\delta a/\|a\|_2} = \frac{\|\phi'(a)\|_2 \|a\|_2}{\|\phi(a)\|_2},
$$

$$
m(\phi, a) := \lim_{\varepsilon \to 0} \sup_{\|\delta a\|_2 \leq \varepsilon \|a\|_2} \frac{\|\delta x\|_\infty/\|x\|_\infty}{\|\delta a/\|a\|_\infty} = \frac{\|\phi'(a)\|_\infty \|a\|_\infty}{\|\phi(a)\|_\infty},
$$

$$
c(\phi, a) := \lim_{\varepsilon \to 0} \sup_{\|\delta a\|_2 \leq \varepsilon \|a\|_2} \frac{\|\delta x/\|x\|_\infty}{\|\delta a/\|a\|_\infty} = \frac{\|\phi'(a)\|_\infty \|a\|_\infty}{\|\phi(a)\|_\infty},
$$

where $\phi'(a)$ denotes the Fréchet derivative \[28,29\] of $\phi$ at the point $a$. $b/a$ (also denoted by $\frac{b}{a}$ in later notations) is the entrywise division. Note that $\xi/0$ is interpreted as zero if $\xi = 0$ and infinity otherwise. The vectors $a$ and $\phi(a)$ are assumed to have no zero entries throughout this paper.

2. A new perturbation estimate of the MTLS solution

In this section, we first recall some preliminary results about MTLS \[12\], and then provide a new formula for the first order perturbation estimate of the MTLS solution.

Assume that the first $n_1$ columns of $A$ are known exactly. Partition $A = [A_1, A_2]$, where $A_1 \in \mathbb{R}^{m \times n_1}$, $A_2 \in \mathbb{R}^{m \times n_2}$, $n_1 + n_2 = n$. Let $x = [x_1^T, x_2^T]^T$, where $x_1 \in \mathbb{R}^{n_1}, x_2 \in \mathbb{R}^{n_2}$. Then the MTLS problem can be formulated as

$$
\min_{E_2, f} \|E_2, f\|_F \quad \text{subject to} \quad A_1 x_1 + (A_2 + E_2) x_2 = b + f.
$$

One way to solve the MTLS problem is to factorize $[A, b]$ into the QR form first:

$$
Q^T[A_1, A_2, b] = \tilde{R} = \begin{bmatrix}
R_{11} & R_{12} & R_{1b} \\
0 & R_{22} & R_{2b} \\
n_1 & n_2 & 1
\end{bmatrix}_{n_1} m - n_1,
$$

and then solve the reduced TLS problem $R_{22} x_2 \approx R_{2b}$ to obtain $x_2$. The vector $x_1$ is then found from $R_{11} x_1 = R_{1b} - R_{12} x_2$ by backward substitution. According to Golub–Van Loan’s theory for the standard TLS problem \[2\], if

$$
\sigma_{n_2}(R_{22}) > \sigma_{n_2+1}([R_{22}, R_{2b}]) > 0,
$$

then the reduced TLS problem and therefore the MTLS problem have a unique solution.
Lemma 2.1 ([5]): For the MTLS problem (2), let \( \tilde{A} = [A, b], W = \text{diag}(O_n, I_n) \), and \( \tilde{W} = \text{diag}(W, 1) \). If (4) holds, then the MTLS solution \( x_M \) satisfies the following generalized eigenvalue system

\[
\begin{bmatrix}
  A^TA & A^Tb \\
  b^TA & b^Tb
\end{bmatrix}
\begin{bmatrix}
x_M \\
-1
\end{bmatrix} = \tilde{\sigma}_{n+1}^2
\begin{bmatrix}
  W \\
  1
\end{bmatrix}
\begin{bmatrix}
x_M \\
-1
\end{bmatrix},
\]

where

\[
\tilde{\sigma}_{n+1}^2 = \sigma_{n+1}^2([R_{22}, R_{2b}]) = \lambda_{n+1}(\tilde{A}^T\tilde{A}, \tilde{W}) < \lambda_n(A^TA, W) = \sigma_{n}^2(R_{22}).
\]

Here \( \lambda_i(M, N) \) denotes the \( i \)-th largest generalized eigenvalue of the matrix pair \( (M, N) \). The unique solution is given by \( x_M = (A^TA - \tilde{\sigma}_{n+1}^2 W)^{-1}A^Tb \), and it is the minimizer of the following optimization problem

\[
\min_x \frac{\|b - Ax\|_2^2}{1 + x^TWx}
\]

whose optimal value is \( \tilde{\sigma}_{n+1}^2 \).

Consider the mapping \( \varphi \) from \( \mathbb{R}^{m \times (n+1)} \) to \( \mathbb{R}^n \) with

\[
\varphi : [A, b] \mapsto x_M = (A^TA - \tilde{\sigma}_{n+1}^2 W)^{-1}A^Tb.
\]

Let \( a := \text{vec}([A, b]) \). Then

\[
x_M = \varphi([A, b]) = \phi(a).
\]

According to the definitions and formulae of three condition numbers, it is vital to compute the Fréchet derivative of \( \phi(a) \). To this end, let \( \hat{A} = A + \Delta A, \hat{b} = b + \Delta b \) where \( \Delta A \) and \( \Delta b \) denote the perturbations to \( A \) and \( b \), respectively. Then the perturbed MTLS problem has the form

\[
\min_{\hat{E}_2, \hat{f}} \| [\hat{E}_2, \hat{f}] \|_F \quad \text{subject to} \quad \hat{A}_1\hat{x}_1 + (\hat{A}_2 + \hat{E}_2)\hat{x}_2 = \hat{b} + \hat{f}.
\]

In [25], based on a limit argument on the equivalent weighting problem [5] of the MTLS problem, Zheng and Yang proved that

\[
\varphi'([A, b]) = K_{ZY} = \left[-(x^T \otimes D) - (r^T \otimes P^{-1}) \Pi_{m,n}, D\right],
\]

where \( x = x_M \) is the exact MTLS solution, \( P = A^TA - \tilde{\sigma}_{n+1}^2 W \), \( D = P^{-1}(A^T - 2Wxx^TW) \) with \( r = Ax - b \) and \( \tilde{\gamma} = 1 + x^TWx \). \( \Pi_{m,n} \) is a vec-permutation matrix such that vec\((C^T) = \Pi_{m,n}\text{vec}(C)\) holds for any \( m \times n \) arbitrary matrix \( C \). Moreover, they gave the 2-norm estimate of \( K_{ZY} \) as

\[
\|K_{ZY}\|_2 = \gamma \frac{1}{2} \left\| P^{-1} \left( A^TA + \gamma^{-1} \tilde{\sigma}_{n+1}^2 \tilde{\gamma} \left( I_n - 2\frac{Wxx^TW}{\tilde{\gamma}} \right) \right) P^{-1} \right\|_2^{\frac{1}{2}},
\]

where \( \gamma = 1 + \|x\|_2^2 \). Here (10) is a corrected version of [25, Equation (19), Theorem 1], where there is a minor error before finishing the final deduction on page 7.
We will adopt a new technique which is different from that in [25] to derive a new closed formula for the perturbation estimate. It is clear that when \( \| [\Delta A, \Delta b] \|_F \) is sufficiently small, the perturbed MTLS problem has a unique solution \( \hat{x} \) that is real analytic in \( \text{vec}([\Delta A, \Delta b]) \) in some neighborhood of the origin. The following theorem presents the first order Taylor approximation of the solution to the perturbed problem (8), by a similar technique used in [17].

**Theorem 2.1:** For the MTLS problem (2) with the genericity condition (4) or (5), denote the unique solution by \( x_* = x_M \) and define \( r = Ax_* - b, G(x_*) = [x_*^T, -1] \otimes I_m \). If \( [A, b] \) is perturbed to \( [\hat{A}, \hat{b}] := [A + \Delta A, b + \Delta b] \) with \( \| [\Delta A, \Delta b] \|_F \) sufficiently small, then the perturbed problem (8) has a unique MTLS solution \( \hat{x} \) satisfying

\[
\hat{x} = x_* + K \text{vec}([\Delta A, \Delta b]) + \mathcal{O}(\| [\Delta A, \Delta b] \|_F^2),
\]

where

\[
K = \varphi'([A, b]) = -P^{-1} \left( A^T H_0 G(x_*) + [I_n \otimes r^T, \ O_{n \times m}] \right)
\]

with \( P = A^T A - \tilde{\sigma}^2_{n+1} W \) and \( H_0 = I_m - \frac{2rr^T}{\|r\|_2^2} \).

**Proof:** Let \( \varepsilon = \text{vec}([\Delta A, \Delta b]) \) and rewrite \( \hat{x} \) as \( x(\varepsilon) \). Similarly to (6), we can get

\[
x(\varepsilon) = \arg\min_x \frac{\| b + \Delta b - (A + \Delta A)x \|^2_2}{1 + x^T W x}.
\]

It is clear that \( x(0) = x_* \). For sufficiently small \( \varepsilon \), \( x(\varepsilon) \) is real analytic in some neighborhood of the origin, and the Taylor series of \( x(\varepsilon) \) with center the origin are convergent. As a result, to prove (11) it suffices to prove \( \nabla_{\varepsilon} x(0) \), the Jacobian of \( x(\varepsilon) \) at the origin, equals to \( K \). To this end, define the two-variable function

\[
f(x, \varepsilon) = \frac{\| \hat{b} - \hat{A} x \|^2_2}{1 + x^T W x}.
\]

The necessary condition for it to attain the minimum at \( x(\varepsilon) \) is

\[
\nabla_{x, \varepsilon} f(x(\varepsilon), \varepsilon) = 0.
\]

Differentiating (13) with respect to \( \varepsilon \) gives

\[
\nabla^2_{x, \varepsilon} f(x(\varepsilon), \varepsilon) \nabla_{\varepsilon} x(\varepsilon) + \nabla^2_{\varepsilon, \varepsilon} f(x(\varepsilon), \varepsilon) = 0,
\]

from which it follows that

\[
\nabla_{\varepsilon} x(0) = -(\nabla^2_{x, \varepsilon} f(x_*, 0))^{-1} \nabla^2_{\varepsilon, \varepsilon} f(x_*, 0),
\]

provided that \( \nabla^2_{x, \varepsilon} f(x_*, 0) \) is nonsingular.

Note that for \( \hat{r} = \hat{A} x - \hat{b} \) we have

\[
\frac{1}{2} \nabla_{x} f(x, \varepsilon) = \frac{\hat{r}^T \hat{A}}{1 + x^T W x} - \frac{\| \hat{r} \|^2_2 x^T W}{(1 + x^T W x)^2}.
\]
\[
\frac{1}{2} \nabla_{x,f}(x, \varepsilon) = \frac{\hat{A}^T \hat{A}}{1 + x^T W x} + 4 \frac{\|\hat{r}\|^2 W x x^T W}{(1 + x^T W x)^3} - \frac{\|\hat{r}\|^2 W}{(1 + x^T W x)^2} \\
- \frac{1}{(1 + x^T W x)^2}(2\hat{A}^T \hat{r} x^T W x + 2 W x \hat{r} \hat{A}).
\] 

(16)

In addition, the minimizer \( x_* = x_M \) of \( \min_x f(x,0) \) requires \( \nabla_x f(x_*,0) = 0 \). As a result, (15) and (6) give

\[
A^T r = \frac{\|r\|^2}{1 + x_0^T W x_*} W x_* = \tilde{\sigma}_{n_2+1} W x_*.
\] 

(17)

Substituting (17) into (16) and setting \( \varepsilon = 0 \) and \( x = x_* \), we get

\[
\frac{1}{2} \nabla_{x,f}^2(x_*,0) = \frac{1}{1 + x_0^T W x_*} (A^T A - \tilde{\sigma}_{n_2+1}^2 W).
\] 

(18)

With the QR factorization (3), take \( Y = \begin{bmatrix} I_{n_1} & -R_{11}^{-1} R_{12} \\ 0 & I_{n_2} \end{bmatrix} \). Then the condition (4) ensures the matrix

\[
Y^T (A^T A - \tilde{\sigma}_{n_2+1}^2 W) Y = \begin{bmatrix} R_{11}^T R_{11} & 0 \\ 0 & R_{22}^T R_{22} - \tilde{\sigma}_{n_2+1}^2 I_{n_2} \end{bmatrix}
\]

to be positive definite. Hence, \( A^T A - \tilde{\sigma}_{n_2+1}^2 W \) and \( \nabla_{x,f}^2(x_*,0) \) are positive definite and nonsingular. According to (14), the remaining task is to evaluate \( \nabla_{x,f}^2(x_*,0) \).

Write

\[
\hat{r} = \hat{A}x - \hat{b} = ([x^T, -1] \otimes I_m) \text{vec}(\hat{A}, \hat{b}) = \tilde{G}s,
\]

where \( \tilde{G} = G(x) \) and \( s = \text{vec}(\hat{A}, \hat{b}) = \text{vec}([A, b]) + \varepsilon \). It is obvious that \( \frac{\partial \tilde{G}}{\partial x_i} s = \hat{a}_i, i = 1, \ldots, n, \) and

\[
\frac{1}{2} \nabla_x f(x, \varepsilon) = \frac{1}{1 + x^T W x} \left[ \tilde{s}^T \tilde{G}^T \frac{\partial \tilde{G}}{\partial x_1} \hat{s}, \ldots, \tilde{s}^T \tilde{G}^T \frac{\partial \tilde{G}}{\partial x_n} \hat{s} \right] - \frac{\tilde{s}^T \tilde{G}^T \tilde{G} \hat{s}}{(1 + x^T W x)^2} x^T W,
\]

\[
\frac{1}{2} \nabla_{x,f}^2(x, \varepsilon) = \frac{1}{1 + x^T W x} \begin{bmatrix} \nabla_x (\tilde{s}^T \tilde{G}^T \frac{\partial \tilde{G}}{\partial x_1} \hat{s}) \\ \vdots \\ \nabla_x (\tilde{s}^T \tilde{G}^T \frac{\partial \tilde{G}}{\partial x_n} \hat{s}) \end{bmatrix} - \frac{W}{(1 + x^T W x)^2} \begin{bmatrix} \nabla_x (\tilde{s}^T \tilde{G}^T \tilde{G} \hat{s}) x_1 \\ \vdots \\ \nabla_x (\tilde{s}^T \tilde{G}^T \tilde{G} \hat{s}) x_n \end{bmatrix}.
\]

We therefore conclude that

\[
\frac{1}{2} \nabla_{x,f}^2(x, \varepsilon) = \frac{1}{1 + x^T W x} \left( \hat{A}^T \hat{G} + [I_n \otimes \hat{r}^T, O_{n \times m}] \right) - \frac{2 W x \hat{r} \hat{G}}{1 + (x^T W x)^2}.
\]

Setting \( \varepsilon = 0 \) and \( x(0) = x_M = x_* \) in the above equation, one gets

\[
\nabla_x x(0) = -(\nabla_{x,f}^2(x_*,0))^{-1} \nabla_{x,f}^2(x_*,0)
\]
\[
\begin{align*}
&= (A^TA - \tilde{\sigma}_{n_2+1}^2 W)^{-1} \left( 2Wx^*r^TG(x^*) - A^TG(x^*) - [I_n \otimes r^T, O_{n \times m}] \right) \\
&= (A^TA - \tilde{\sigma}_{n_2+1}^2 W)^{-1} \left( \frac{2A^Tr^TG(x^*) - A^TG(x^*) - [I_n \otimes r^T, O_{n \times m}]}{\|r\|_2^2} \right) \\
&= (A^TA - \tilde{\sigma}_{n_2+1}^2 W)^{-1} \left( -A^TH_0G(x^*) - [I_n \otimes r^T, O_{n \times m}] \right),
\end{align*}
\]
this yields the desired result. 

**Theorem 2.2:** For the first order perturbation estimate (11) of the MTLS solution \( x \), let \( x = x^* \) and the matrices \( K_{ZY} \) and \( K \) be given by (9) and (12), respectively. Then \( K_{ZY} = K \).

**Proof:** Note that for any \( n \times m \) matrix \( M_1 \) we have
\[
M_1G(x) = M_1([x^T, -1] \otimes I_m) = M_1[x_1I_m, \ldots, x_nI_m, -I_m] = [x^T \otimes M_1, -M_1],
\]
and therefore
\[
K = -P^{-1}(A^TH_0G(x) + [I_n, 0_{n \times 1}] \otimes r^T) \\
= [-((x^T \otimes (P^{-1}A^TH_0)), (P^{-1}A^TH_0)] - [P^{-1}(I_n \otimes r^T), O_{n \times m}],
\]
where the matrix \((P^{-1}A^TH_0)\) is exactly the matrix \( D \) in (9) by the relation (17).

For any \( m \times n \) matrix \( Y \),
\[
(r^T \otimes P^{-1})\Pi_{m,n} \text{vec}(Y) = (r^T \otimes P^{-1})\text{vec}(Y^T) = \text{vec}(P^{-1}Y^Tr) \\
= P^{-1}\text{vec}(Y^Tr) = P^{-1}\text{vec}(r^TY) = P^{-1}(I_n \otimes r^T)\text{vec}(Y),
\]
which gives
\[
(r^T \otimes P^{-1})\Pi_{m,n} = P^{-1}(I_n \otimes r^T).
\]
Thus, the assertion in the theorem follows.

**Remark 2.1:** When \( n_1 = 0 \), the MTLS problem (2) reduces to the standard TLS problem (1). Clearly, the first order perturbation estimates characterized by (9) and (11) also become the ones in [15,17], respectively. Theorem 2.2 therefore reveals the equivalence of the first order perturbation estimates in [15,17]. When \( n_2 = 0 \), the MTLS problem reduces to the least squares (LS) problem, and the estimate in (9) is exactly the result from [31]. Hence, Theorem 2.1 unifies the perturbation results for the LS and TLS problems.

### 3. Condition numbers for the MTLS problem

In this section, we first consider the condition numbers of the MTLS problem for general \( A \) and \( b \). Structured condition numbers for the MTLS problem with linear structure will be considered next.
3.1. Compact condition numbers and perturbation analysis

According to Theorem 2.1 and the concept of normwise condition number, the mapping \( \phi \) in (7) satisfies \( \phi'(a) = K \). Thus, the 2-norm relative condition number of the MTLS problem is given by

\[
\kappa(A, b) = \frac{\|K\|_2 \|[A, b]\|_F}{\|x_M\|_2}. \tag{19}
\]

Note that the expression of \( K \) involves Kronecker product operations, which might lead to expensive storage and computational cost. The following theorem provides four different compact expressions for \( \|K\|_2 \), which have not appeared in the literature.

**Theorem 3.1:** With the notations in Theorem 2.1, the absolute condition number \( \kappa = \|K\|_2 \) of the MTLS solution \( x \) has the following equivalent forms:

\[
\kappa_1 = \|P^{-1}(\gamma A^T A - A^T r x^T - x r^T A + \|r\|_2^2 I_n)P^{-1}\|_2^{1/2}, \tag{20}
\]

\[
\kappa_2 = \gamma^{1/2} \left\| \left( A^T A + \gamma^{-1} \tilde{\sigma}_{n_2+1}^2 \tilde{\gamma} \left( I_n - \frac{Wxx^T + xx^T W}{\tilde{\gamma}} \right) \right) P^{-1} \right\|_2^{1/2}, \tag{21}
\]

\[
\kappa_3 = \bigg\| P^{-1} \left[ A^T, \|x\|_2 A^T - \|x\|_2 \tilde{\sigma}_{n_2+1}^2 \frac{Wxr^T}{\|r\|_2^2}, \|r\|_2^2 I_n - \tilde{\sigma}_{n_2+1}^2 Wxx^T \right] \bigg\|_2, \tag{22}
\]

\[
\kappa_4 = \bigg\| P^{-1} \left[ (1 + \beta)A^T - \beta \tilde{\sigma}_{n_2+1}^2 \frac{Wxr^T}{\|r\|_2^2}, \|r\|_2^2 I_n - \tilde{\sigma}_{n_2+1}^2 Wxx^T \right] \bigg\|_2, \tag{23}
\]

where \( W = \text{diag}(O_{n_1}, I_{n_2}) \), \( \gamma = 1 + \|x\|_2^2 \), \( \tilde{\gamma} = 1 + x^T W x \), and \( \beta = -1 \pm \sqrt{1 + \|x\|_2^2} \).

**Proof:** It is clear that

\[
\kappa = \|K\|_2 = \|KK^T\|_2^{1/2}.
\]

By setting \( \tilde{G} = G(x) \), \( \Gamma = [I_n, 0_{n \times 1}] \otimes r^T \) in (12), it is easy to show that

\[
\tilde{G}\tilde{G}^T = 1 + \|x\|_2^2 = \gamma, \quad \Gamma\Gamma^T = \|r\|_2^2 I_n, \quad \tilde{G}\Gamma^T = rx^T. \tag{24}
\]

Since the Householder matrix \( H_0 \) involved in \( K \) satisfies \( H_0 r = -r \), we get

\[
KK^T = P^{-1} \left( A^T H_0 \tilde{G} + \Gamma \right) \left( \tilde{G}^T H_0^T A + \Gamma^T \right) P^{-1} = P^{-1}(\gamma A^T A - A^T r x^T - x r^T A + \|r\|_2^2 I_n)P^{-1}.
\]

The relations (20) and (21) then follow from (17). Furthermore,

\[
KK^T = P^{-1} \begin{bmatrix} A^T, & I_n \end{bmatrix} \begin{bmatrix} \gamma I_m & -r x^T \\ -x r^T & \|r\|_2^2 I_n \end{bmatrix} \begin{bmatrix} A \\ I_n \end{bmatrix} P^{-1}. \tag{25}
\]

The middle matrix in (25) has the factorizations

\[
\begin{bmatrix} \gamma I_m & -r x^T \\ -x r^T & \|r\|_2^2 I_n \end{bmatrix} = U \begin{bmatrix} I_m + \|x\|_2^2 P_0 \\ O \end{bmatrix} \begin{bmatrix} O \\ \|r\|_2^2 I_n \end{bmatrix} U^T = UD_1 D_1^T U^T, \tag{26}
\]
where $i = 1, 2$, $P_0 = I_m - \frac{1}{\|r\|_2^2}rr^T$, $D_1 = \text{diag}([I_m, \|x\|_2P_0, \|r\|_2I_n])$, and

$$D_2 = \begin{bmatrix} I_m + \beta P_0 & O \\ \|r\|_2^2 I_n & O \end{bmatrix}, \quad U = \begin{bmatrix} I_m - \frac{1}{\|r\|_2^2}rx^T \\ O \end{bmatrix}. $$

Let $Z_i = [A^T, I_n]UD_i$. Then

$$\|K\|_2 = \|KK^T\|_2^{1/2} = \|P^{-1}Z_iZ_i^TP^{-1}\|_2^{1/2} = \|P^{-1}Z_i\|_2. $$

By the fact in (17), we have the expressions for $\|P^{-1}Z_1\|_2$ and $\|P^{-1}Z_2\|_2$ in (22)–(23), respectively.

**Theorem 3.2:** With the notations in Theorem 2.1, let $x$ and $\hat{x}$ be the solutions to the original and perturbed MTLS problems, respectively. Then $\Delta x = \hat{x} - x$ satisfies

$$\frac{\|\Delta x\|_2}{\|x\|_2} \lesssim \kappa_b \frac{\|\Delta b\|_2}{\|b\|_2} + \kappa_A \frac{\|\Delta A\|_2}{\|A\|_2}, \tag{27}$$

where $\kappa_b = \frac{\|b\|_2}{\|x\|_2} \|(A^TA - \tilde{\sigma}_{n_{2+1}}^2W)^{-1}A\|_2$, and

$$\kappa_A = \frac{\|A\|_2}{\|x\|_2} \left( \|r\|_2 \left\| (A^TA - \tilde{\sigma}_{n_{2+1}}^2W)^{-1} \right\|_2 + \|x\|_2 \left\| (A^TA - \tilde{\sigma}_{n_{2+1}}^2W)^{-1}A \right\|_2 \right)$$

with $r = Ax - b$.

**Proof:** In (11)-(12), by setting $x_* = x$, we have

$$G(x_*)\text{vec}([\Delta A, \Delta b]) = \Delta Ax - \Delta b,$$

$$[I_n \otimes r^T, O_{n \times m}]\text{vec}([\Delta A, \Delta b]) = \text{vec} \left( r^T [\Delta A, \Delta b] \begin{bmatrix} I_n \\ 0 \end{bmatrix} \right) = \Delta A^Tr.$$

Thus,

$$\|\Delta x\|_2 \lesssim \|P^{-1}A\|_2(\|\Delta A\|_2\|x\|_2 + \|\Delta b\|_2) + \|P^{-1}\|_2\|r\|_2\|\Delta A\|_2,$$

from which the result in (27) follows.

**Remark 3.1:** There is a slight difference between (21) and (10). We reexamine the rationality of (10) and use a different approach from that in [25] to evaluate $\|K_{ZY}\|_2 = \max_{\|y\|_2 = 1} \|K_{ZY}y\|_2$ in (9). The new estimate of $\|K_{ZY}\|_2$ is deduced as follows:

$$\|K_{ZY}\|_2^2 = \max_{\|y\|_2 = 1} \|K_{ZY}^T\text{vec}(y)\|_2^2 = \max_{\|y\|_2 = 1} \left\| \begin{bmatrix} -[\text{vec}(D^Tyx^T) + \Pi_{m,n}^T \text{vec}(P^{-1}yr^T)] \\ D^Ty \end{bmatrix} \right\|_2^2$$

$$= \max_{\|y\|_2 = 1} \left( \|\text{vec}(D^Tyx^T) + \text{vec}(P^{-1}yr^T)\|_2^2 + \|D^Ty\|_2^2 \right)$$

$$= \max_{\|y\|_2 = 1} \left( \|\text{vec}(xy^TD + P^{-1}yr^T)\|_2^2 + \|D^Ty\|_2^2 \right)$$
\[
\begin{align*}
&= \max_{\|y\|_2=1} \left( \|xy^TD + P^{-1}yr^T\|_F^2 + \|DT^Ty\|_2^2 \right) \\
&= \max_{\|y\|_2=1} \left( \text{tr}[(xy^TD + P^{-1}yr^T)(xy^TD + P^{-1}yr^T)^T] + \|DT^Ty\|_2^2 \right) \\
&= \max_{\|y\|_2=1} y^T \left( 1 + \|x\|_2^2 \right) DD^T + Drx^TP^{-1} + P^{-1}xr^TD^T + P^{-2}\|r\|_2^2 \right)y, \\
&= \|1 + \|x\|_2^2 \|DD^T + Drx^TP^{-1} + P^{-1}xr^TD^T + P^{-2}\|r\|_2^2, \\
\end{align*}
\]

where \( D = P^{-1}A^TH_0 \) by the relation (17) and
\[
DD^T = P^{-1}A^TAP^{-1}, \quad Dr = -P^{-1}A^Tr = -\tilde{\sigma}_{n_2+1}^2P^{-1}Wx.
\]

This gives
\[
\|K_{ZY}\|_2 \geq P^{-1}\left( \gamma A^TA - \tilde{\sigma}_{n_2+1}^2Wxx^TW + \|r\|_2^2I_n \right)P^{-1}\|1/2, \\
\]

where \( \gamma = 1 + \|x\|_2^2 \) and \( \|r\|_2^2 = \tilde{\sigma}_{n_2+1}^2(1 + x^TWx) \) according to (17). Consequently, \( \|K_{ZY}\|_2 \) is equivalent to the estimate (21). We will provide numerical experiments to compare our estimates with the one in [25] to illustrate the rationality of our estimates and the incorrectness of the norm estimate in (10).

**Remark 3.2:** In (12), the matrix \( K \) is of size \( n \times m(n+1) \), while the associated matrices in (21)–(23) are respectively of size \( n \times n, n \times (2m+n), n \times (m+n) \), which are more economical in storage. From the aspect of computational efficiency, the advantages of (21) over (20), (23) over (22) are obvious since they need less matrix-matrix multiplications. However, as pointed out in [15,32], the explicit formulations of matrix cross product \( A^TA \) and \( P^{-1} \) in (21) are not expected. The formula in (23) is preferred in avoiding the matrix cross product. In terms of (3), its calculation can be implemented by utilizing the intermediate results from solving the MTLS problem. For example, the inverse of \( P \) can be expressed as
\[
P^{-1} = \begin{bmatrix}
(R_{11}^T R_{11})^{-1} + R_{11}^{-1}R_{12}S^{-1}R_{12}^T R_{11}^{-T} & -R_{11}^{-1}R_{12}S^{-1} \\
-S^{-1}R_{12}^T R_{11}^{-T} & S^{-1}
\end{bmatrix},
\]

where \( S^{-1} = (R_{22}^T R_{22} - \tilde{\sigma}_{n_2+1}^2I)^{-1} \) can be an intermediate result from the solution of the TLS problem \( R_{22}x_2 \approx R_{2b} \), say, in the small or medium-sized MTLS problem, the SVD of \( [R_{22}, R_{2b}] \) is available and \( S^{-1} \) can be computed cheaply based on the SVD and the result in [33, Lemma 2]; and when solving the large MTLS problems by the Rayleigh quotient and preconditioned conjugate gradient (RQI-PCG) [34] method, an approximation of \( \tilde{\sigma}_{n_2+1} \) is available and the linear system \( (R_{22}^T R_{22} - \tilde{\sigma}_{n_2+1}^2I)w = f \) can be efficiently solved based on the preconditioned conjugate gradient method. The main cost in each iteration step only involves the solution of two triangular linear systems. \( P^{-1} \) based on triangular linear systems can therefore be efficiently computed and good numerical stability is also preserved [32, Chapter 8].

**Remark 3.3:** In the case that \( n_1 = 0 \), the Kronecker-product-free expression in (21) reduces to the compact formula for the normwise condition number of the TLS problem.
In [18, Theorem 3], Wang et al. proved that for the TLS problem, the absolute condition number takes the form

\[ \bar{\kappa}_3 = \|K\|_2 = \|P^{-1} A^T, \|x\|_2 A^T \left( I_m - \frac{1}{\|r\|_2^2} r r^T \right), \|r\|_2 \left( I_n - \frac{1}{\|r\|_2^2} A^T r x^T \right) \|. \]

where \( P = A^T A - \sigma_{n+1} (A, b) I_n \), and \( r = Ax - b \) with \( x \) being the TLS solution. Combined with the equality in (17) for \( n_1 = 0 \), the above estimate is just a special case of (22).

For the perturbation bound of the MTLS problem, the result in (27) is equivalent to Zheng and Yang’s result in [25, Theorem 5]. It also includes the one for the perturbation bound of the standard TLS problem as a special case, see [19, Theorem 3.1].

According to the definitions of the mixed and componentwise condition numbers, they can be formulated as

\[ m(A, b) = \frac{\|K\| \|\text{vec}(A, b)\|_\infty}{\|x\|_\infty}, \quad c(A, b) = \frac{\|K\| \|\text{vec}(A, b)\|_\infty}{\|x\|_\infty}. \]

Note that \( K = K_{ZY} \) by Theorem 2.2. In practical computations, the upper bounds below are Kronecker-product free and can be alternatives to improve the computational efficiency. The proof of the theorem is straightforward.

**Theorem 3.3:** With the notations in Theorem 2.1, the mixed and componentwise condition numbers of the MTLS problem are bounded as

\[ m(A, b) \leq \frac{\|P^{-1} A^T H_0|(|A\|_x + |b|) + |P^{-1} A^T r|\|_\infty}{\|x\|_\infty}, \quad (29) \]

\[ c(A, b) \leq \frac{\|P^{-1} A^T H_0|(|A\|_x + |b|) + |P^{-1} A^T r|\|_\infty}{|x|}. \quad (30) \]

**3.2. Structured condition numbers**

If the matrix \( A \) in the MTLS problem lies in a linear subspace \( \mathcal{S} \) which consists of a class of structured matrices, then any matrix in \( \mathcal{S} \) can be represented by a linear combination of linearly independent matrices \( S_1, S_2, \ldots, S_q \in \mathcal{S} \), i.e. \( A = \sum_{i=1}^{q} \alpha_i S_i \), and

\[ \text{vec}(A) = \sum_{i=1}^{q} \alpha_i \text{vec}(S_i) = \Phi^\text{struct}_A \alpha, \]

where \( \Phi^\text{struct}_A = [\text{vec}(S_1), \text{vec}(S_2), \ldots, \text{vec}(S_q)] \), and \( \alpha = [\alpha_1, \alpha_2, \ldots, \alpha_q]^T \).

For simplicity, in this paper we always assume that each element of \( A \in \mathcal{S} \) depends on a single component of \( \alpha \). Several kinds of structured matrices fall into this category, such as Toeplitz, Hankel, and circulant matrices. By the statement in [17, Theorem 4.1], \( \Phi^\text{struct}_A \) has orthogonal columns with full column rank and at most one nonzero entry in each row.
Note that
\[ \text{vec}([A, b]) = \Phi_{A,b}^{\text{struct}} s := \begin{bmatrix} \Phi_{A}^{\text{struct}} & O \\ O & I_m \end{bmatrix} \begin{bmatrix} \alpha \\ b \end{bmatrix}. \]

For the perturbed MTLS problem, the perturbation \([\Delta A, \Delta b]\) is assumed to have the same structure as that of \([A, b]\), i.e. \(\text{vec}([\Delta A, \Delta b]) = \Phi_{A,b}^{\text{struct}} \epsilon\) for some \(\epsilon \in \mathbb{R}^{d+m}\).

Define the mapping \(\phi\) from \(\mathbb{R}^{q+m}\) to \(\mathbb{R}^n\) such that
\[ \phi(\begin{bmatrix} \alpha \\ b \end{bmatrix}) = x_M = (A^T A - \tilde{\sigma}^2_{n+1} W)^{-1} A^T b. \]
Based on (11), the first order perturbation result becomes
\[ \Delta x = K \Phi_{A,b}^{\text{struct}} \epsilon + O(\|\epsilon\|_2^2). \]

In view of the concept of condition numbers, the structured normwise, mixed and componentwise condition numbers of the MTLS solution take the following forms
\[
\kappa_{\text{struct}}(\alpha, b) = \|K \Phi_{A,b}^{\text{struct}}\|_2 \|\begin{bmatrix} \alpha^T \\ b^T \end{bmatrix}\|_2 \frac{\|x_M\|_2}{\|x_M\|_\infty},
\]
\[
m_{\text{struct}}(\alpha, b) = \|K \Phi_{A,b}^{\text{struct}}\|_\infty \|\begin{bmatrix} |\alpha|^T \\ |b|^T \end{bmatrix}\|_\infty \frac{\|x_M\|_\infty}{\|x_M\|_\infty},
\]
\[
c_{\text{struct}}(\alpha, b) = \|K \Phi_{A,b}^{\text{struct}}\|_\infty \|\begin{bmatrix} |\alpha|^T \\ |b|^T \end{bmatrix}\|_\infty \frac{\|x_M\|_\infty}{\|x_M\|_\infty},
\]
(31)

where
\[ K \Phi_{A,b}^{\text{struct}} = -P^{-1} \left( A^T H_0 [x^T \otimes I_m, -I_m] + [I_n \otimes r^T, O_{n \times m}] \right) \begin{bmatrix} \Phi_{A}^{\text{struct}} & O \\ O & I_m \end{bmatrix}, \]
which is Kronecker-product-free and can be computed more efficiently with less storage.

4. Numerical experiments

In this part, we first present numerical experiments to verify the usefulness of our first order perturbation result, and then compare three types of condition numbers via several examples. All experiments are coded by MATLAB R2012b with machine precision \(2.22 \times 10^{-16}\) and done on a laptop with Intel Core i5-5200U CPU @ 2.20GHz and the memory is 4 GB.

Example 4.1: Consider the estimation of the parameters in a transfer function model [35], given in its errors-in-variables form
\[
C(q^{-1}) y_0(t) = B(q^{-1}) u_0(t),
\]
\[
u(t) = u_0(t) + \Delta u(t),
\]
\[
y(t) = y_0(t) + \Delta y(t),
\]
where \(u_0(t)\) and \(y_0(t)\) are unmeasurable noise-free inputs and outputs, \(\Delta u(t)\) and \(\Delta y(t)\) represent all stochastic disturbances to the inputs and outputs, respectively; \(A(q^{-1})\) and \(B(q^{-1})\) are polynomials taking the form
\[
B(q^{-1}) = b_1 q^{-1} + \cdots + b_n q^{-n_1},
\]
\[ C(q^{-1}) = 1 + c_1 q^{-1} + \cdots + c_{n_2} q^{-n_2}, \]

where \( q^{-1} \) is a backward shift operator such that \( q^{-1} y(t) = y(t-1) \). In order to estimate the parameters in the transfer function, we need to solve the approximate system \( \phi(t)x \approx y(t) \), for \( t = 1, 2, \ldots, N \) and

\[
\phi(t) = [u(t-1), \ldots, u(t-n_1), y(t-1), \ldots, -y(t-n_2)]^T, \\
x = [b_1, \ldots, b_{n_1}, c_1, \ldots, c_{n_2}]^T,
\]

in which the entries \( u(j), y(j) \) with \( j \leq 0 \) are set to zero. This produces the approximate linear system \( Ax \approx z \), where

\[ A = [\phi(k_0+1), \ldots, \phi(k_0+m)]^T, \quad z = [y(k_0+1), \ldots, y(k_0+m)]^T. \]

Here \( k_0 + 1(\geq 1) \) and \( m \) are the starting point and the number of chosen samples for parameter estimation, respectively. In [10], Van Huffel and Vandewalle proposed the TLS model to solve the linear equation when the inputs \( u(t) \) and outputs \( y(t) \) are affected by the noise. In our test, we assume that the inputs \( u(t) \) are noise-free, and the outputs \( y(t) \) are affected by white noise with zero mean and variance 0.01. To estimate the parameters, the MTLS model is necessary in this case.

Take \( m = 30, n = 20, n_1 = 10 \) and \( k_0 = \max(n_1, n_2) = 10 \). Generate entrywise perturbation

\[ [\Delta A, \Delta b] = \epsilon \cdot \text{rand}(m, n+1) \odot [A, b], \]

where \( \odot \) denotes the entrywise multiplication. Denote

\[ \eta_{\Delta x}^{ZY} = \| \Delta x - K_0^{ZY} \text{vec}([\Delta A, \Delta b]) \|_2, \quad \eta_{\Delta x}^{new} = \| \Delta x - K^0 \text{vec}([\Delta A, \Delta b]) \|_2, \]

where \( K_0^{ZY} \) and \( K^0 \) represent the matrices \( K_{ZY} \) and \( K \) in (9) and (12), respectively. For different \( \epsilon \) and random perturbations, in Table 1, we compare the first order perturbation estimates in (9) and (12). The absolute normwise condition number estimates in (10), (21) and (23) are also displayed in the table.

In Table 1, we observe that for different parameter \( \epsilon \) and perturbations, \( \eta_{\Delta x}^{ZY} \) and \( \eta_{\Delta x}^{new} \) are all of the magnitude \( O(\epsilon^2) \), indicating the rationality of the first order perturbation estimates in (9) and (12). Among the five methods for evaluating the absolute normwise condition number \( \| K \|_2 \) or \( \| K_{ZY} \|_2 \), we find that four methods give the same value while the estimate via (10) does not match the true value of \( \| K_{ZY} \|_2 \) in (9). This illustrates the correctness of our theoretical results.

### Table 1. Comparisons of the absolute normwise condition number with different forms for the perturbed MTLS problem.

| \( \epsilon \)   | \( \| \Delta x \|_2 \) | \( \eta_{\Delta x}^{ZY} \) | \( \eta_{\Delta x}^{new} \) | \( \| K_0^{ZY} \|_2 \) (9) | \( \| K_{ZY} \|_2 \) (10) | \( \| K^0 \|_2 \) (12) | \( K_2 \) (21) | \( K_4 \) (23) |
|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-------------|-------------|
| 1e−2           | 3.63e−2         | 2.93e−4         | 2.93e−4         | 18.50           | 19.80           | 18.50           | 18.50       | 18.50       |
| 1e−4           | 1.77e−3         | 2.00e−7         | 2.00e−7         | 11.43           | 16.46           | 11.43           | 11.43       | 11.43       |
| 1e−6           | 1.95e−5         | 7.89e−11        | 7.89e−11        | 50.51           | 56.79           | 50.51           | 50.51       | 50.51       |
| 1e−8           | 2.20e−8         | 5.41e−15        | 5.41e−15        | 13.56           | 14.13           | 13.56           | 13.56       | 13.56       |
Example 4.2: In this example, we compare the relative error of the MTLS solution with the estimated upper bounds, based on the normwise condition number and the perturbation bound in (27) as well. Firstly, we construct the random MTLS problems as follows. The input data $A$ and $b$ are generated according to the QR factorization (3), where $Q$ is a random orthogonal matrix, and

$$[R_{11}, R_{12}, R_{1b}] = \text{triu}((\text{qr}(\text{rand}(n_1, n + 1))).$$

Here $\text{qr}(\cdot)$ and $\text{triu}(\cdot)$ are Matlab commands to produce the QR factorization and the upper triangular part of a matrix, respectively. With random unit vectors $y \in \mathbb{R}^{m-n_1}$ and $z \in \mathbb{R}^{n_2+1}$, set $Y = I - 2yy^T$, $Z = I - 2zz^T$, $D = \text{diag}(n_2, n_2 - 1, \ldots, 1, 1 - e_p)$ with a given parameter $e_p$. Similarly to [15], the subblock matrix $[R_{22}, R_{2b}]$ is generated by

$$[R_{22}, R_{2b}] = \text{triu}(\text{qr}(Y \left[ \begin{array}{c} D \\ O \end{array} \right] Z^T)).$$

By the interlacing property [1] of singular values, we get

$$0 \leq \sigma_{n_2}(R_{22}) - \sigma_{n_2+1}([R_{22}, R_{2b}]) \leq \sigma_{n_2}([R_{22}, R_{2b}]) - \sigma_{n_2+1}([R_{22}, R_{2b}]) = e_p,$$

where the quantity $\sigma_{n_2}(R_{22}) - \sigma_{n_2+1}([R_{22}, R_{2b}])$ measures the distance of our problem to nongenericity. By varying $e_p, m, n$ and $n_1$, we can generate different MTLS problems. With small values of $e_p$, it is possible to study the utility of the condition numbers.

Consider the perturbation as in Example 4.1 with $\epsilon = 10^{-10}$. In Table 2, we compare the efficiency of computing the normwise condition numbers with different forms for the MTLS problem. Let $\kappa_0 = \|K^0\|_2$ be computed via formula (12) and

$$\varepsilon_1 = \frac{\|\Delta A, \Delta b\|_F}{\|A, b\|_F}, \quad \tilde{\kappa}_i = \frac{\kappa_i \|[A, b]\|_F}{\|x\|_2}, \quad i = 0, 4,$$

where $\tilde{\kappa}_i$ is the relative normwise condition number, with the matrix inverse $P^{-1}$ in $\kappa_i$ being computed via the formula (28) and [33, Lemma 2].

Note that the forward error of the MTLS solution can be bounded by $\varepsilon_1 \tilde{\kappa}_0$. These bounds and actual relative forward errors for different problems are listed in Table 2, where the CPU time in seconds for computing the corresponding upper bounds is displayed. It is observed that the computation of $\tilde{\kappa}_0$ is less efficient due to the large size of $K$, while the computations of the bound (27) and the upper bounds based on $\tilde{\kappa}_4$ are efficient, among which the bound (27) is as sharp as $\varepsilon_1 \tilde{\kappa}_0$ and $\varepsilon_1 \tilde{\kappa}_4$, and they are about three orders of magnitude larger than actual forward errors.

Table 2. Comparisons of the upper bounds with forward errors for perturbed MTLS problems.

| $m = 300$ | $n = 200$ | $\frac{\|\Delta x\|_2}{\|x\|_2}$ | $\varepsilon_1 \tilde{\kappa}_0$ (Time) | $\varepsilon_1 \tilde{\kappa}_4$ (Time) | (27) (Time) |
|-----------|-----------|------------------|-----------------|-----------------|----------------|
| $e_p = 0.9$ | $n_1 = 60$ | 1.14e−9 | 3.90e−7 (14.23) | 3.90e−7 (7.14) | 4.45e−7 (7.14) |
| $n_1 = 120$ | 1.28e−9 | 1.41e−6 (15.32) | 1.41e−6 (7.69) | 1.69e−6 (7.69) |
| $n_1 = 180$ | 3.89e−9 | 1.11e−5 (14.01) | 1.11e−5 (7.04) | 1.28e−5 (7.04) |
| $e_p = 0.0009$ | $n_1 = 60$ | 9.26e−8 | 5.89e−5 (14.46) | 5.89e−5 (7.29) | 8.98e−5 (7.29) |
| $n_1 = 120$ | 3.19e−7 | 1.52e−4 (13.92) | 1.52e−4 (6.99) | 1.64e−4 (7.02) |
| $n_1 = 180$ | 4.92e−8 | 3.13e−5 (14.01) | 3.13e−5 (7.05) | 3.27e−5 (7.05) |
Example 4.3: Consider the intercept model arising in the ‘errors in variables’ regression model [8]: \[ \alpha + x_1 c_1 + \cdots + x_n c_n = b, \]
where \( b \) and \( c_i \) are observed \( m \times 1 \) vectors, and \( \alpha \in \mathbb{R}^m \) is the intercept vector of the linear model. The model gives rise to the overdetermined set of equations
\[
[1_m, C] \begin{bmatrix} \alpha \\ x \end{bmatrix} = b,
\]
where \( 1_m = [1, \ldots, 1]^T \), and \( C = [c_1, c_2, \ldots, c_n] \in \mathbb{R}^{m \times n} \). Note that the first column of the left-hand side matrix is known exactly. If \( C \) and \( b \) are contaminated by noise, then we get the MTLS problem.

Choose \( b \) to be a random vector. For the matrix \( C \), we consider the following two cases:

(a) Take \( m = 6, n = 4 \) and
\[
C = \begin{bmatrix}
\delta & 0 & 0 & 0 \\
0 & \delta & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & \delta
\end{bmatrix},
\]
where \( \delta \) is a tiny positive parameter.

Set \( A = [1_m, C] \), and in the perturbed MTLS problem, the first column of \( A \) is not perturbed, while the matrix \( C \) is perturbed according to its structure, i.e.
\[
[\Delta C, \Delta b] = 10^{-10} \cdot \text{rand}(m, n + 1) \odot [C, b].
\]

Define
\[
\varepsilon_2 = \min \{ \varepsilon : |\Delta A| \leq \varepsilon |A|, |\Delta b| \leq \varepsilon |b| \}.
\]

In Table 3, we compare the actual relative forward errors \( \frac{\|\Delta x\|_2}{\|x\|_2}, \frac{\|\Delta x\|_\infty}{\|x\|_\infty} \) and \( \frac{\|\Delta x\|_\infty}{\|x\|_\infty} \) with the estimated bounds \( \varepsilon_1 \kappa_4, \varepsilon_2 m^u, \varepsilon_2 c^u \), respectively, in which \( m^u, c^u \) are the upper bounds given by (29)-(30). The tabulated results show that the normwise condition number multiplied by backward error is far from the actual relative error of the solution when \( \delta \) decreases, while the bounds based on mixed and componentwise condition numbers can estimate the forward error much more tightly. Note that the upper bounds \( m^u, c^u \) have less computational cost and are sharp estimates of \( m, c \), respectively, and hence they can be alternatives for evaluating the forward error of the solution.

### Table 3. Comparison of upper bounds with forward errors for perturbed MTLS problems.

| \( \delta \) | \( n_1 \) | \( \frac{\|\Delta x\|_2}{\|x\|_2} \) | \( \varepsilon_1 \kappa_4 \) | \( \frac{\|\Delta x\|_\infty}{\|x\|_\infty} \) | \( \varepsilon_2 m^u \) | \( \varepsilon_2 c^u \) | \( \frac{\|\Delta x\|_\infty}{\|x\|_\infty} \) |
|------|------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| \( 10^{-2} \) | 1    | 9.68e−11        | 1.52e−8         | 9.34e−11        | 3.64e−9         | 3.89e−9         | 3.32e−10        | 1.25e−8         | 1.38e−8         |
|       | 3    | 5.59e−11        | 2.63e−8         | 5.59e−11        | 2.61e−10        | 2.82e−10        | 7.41e−11        | 1.88e−9         | 2.50e−9         |
| \( 10^{-4} \) | 1    | 2.09e−10        | 7.55e−6         | 2.06e−10        | 1.85e−9         | 4.47e−9         | 2.96e−10        | 2.65e−9         | 6.23e−9         |
|       | 3    | 3.42e−11        | 2.85e−6         | 2.98e−11        | 4.58e−10        | 5.33e−10        | 9.27e−11        | 1.42e−9         | 1.66e−9         |
| \( 10^{-6} \) | 1    | 1.09e−10        | 2.73e−4         | 9.93e−11        | 9.01e−10        | 9.75e−10        | 4.02e−10        | 2.79e−9         | 3.08e−9         |
|       | 3    | 1.00e−11        | 5.91e−4         | 7.69e−12        | 1.06e−9         | 1.24e−9         | 2.14e−10        | 1.81e−8         | 2.12e−8         |
(b) Take \( C = T \) to be a large \( m \times (m - 2\omega) \) Toeplitz matrix, whose first column is given by \( t_{i,1} = i \) for \( i = 1, 2, \ldots, 2\omega + 1 \), and zero otherwise. Entries in the first row are given by \( t_{1,j} = t_{1,1} \) if \( j = 1 \), and zero otherwise.

For the coefficient matrix \( A = [1_m, T] \) in the intercept model, it has a special structure with

\[
A = S_1 + t_{1,1}S_2 + \cdots + t_{2\omega+1,1}S_{2\omega+2},
\]

where \( S_1 = [1_m, O_{m \times (m - 2\omega)}], S_i = [0_{m \times 1}, \hat{S}_{i-1}] \) for \( i = 2, \ldots, 2\omega + 2 \) and \( \hat{S}_1 = \begin{bmatrix} I_{m-2\omega} \\ O_{2\omega \times (m-2\omega)} \end{bmatrix} \), \( \hat{S}_i = Y_0\tilde{S}_{i-1} \) with \( Y_0 \) being a lower shift matrix of order \( m \). Note that (32) with \( q = 2\omega + 2 \) yields

\[
[S_1x, S_2x, \ldots, S_{2\omega+2}x] = [x_11_m, Tx],
\]

\[
[S_1^T r, S_2^T r, \ldots, S_{2\omega+2}^T r] = \text{diag}((1_m^T r), H_r),
\]

where \( T_x \) is an \( m \times (m - 2\omega) \) lower Toeplitz matrix with \( x_2, \ldots, x_{m-2\omega+1}, 0, \ldots, 0 \) in its first column, and \( H_r \) is an \( (m - 2\omega) \times (2\omega + 1) \) anti-upper Hankel matrix with \( r_1, \ldots, r_{m-2\omega} \) in its first column and \( r_{m-2\omega}, \ldots, r_m \) in its last row.

Let \( b = \lambda \hat{b} \) with \( \hat{b} \) being a random vector, and the random entrywise perturbations of \( t_1 \) and \( b \) satisfy

\[
\Delta t_1 = 10^{-10} \cdot \text{randn}(m, 1) \odot t_1, \quad \Delta b = 10^{-10} \cdot \text{randn}(m, 1) \odot b.
\]

The matrix \( T \) is perturbed to \( \hat{T} \) with the same structure as \( T \).

To evaluate upper bounds via structured condition numbers (denoted by \( \kappa^s, m^s, c^s \), resp.) in (31)-(32), denote

\[
\epsilon_1^s = \frac{\| [\Delta a^T, \Delta b^T] \|_2}{\| [a^T, b^T] \|_2}, \quad \epsilon_2^s = \min \{ \epsilon : |\Delta a| \leq \epsilon |a|, |\Delta b| \leq \epsilon |b| \},
\]

in which \( a = [1, t_{1,1}, \ldots, t_{2\omega+1,1}]^T \). In Table 4, we take \( n_1 = 1, \omega = 8 \) and \( m = 500, 1000 \) to compare general condition numbers with structured ones. For general mixed and componentwise condition numbers, we use \( \epsilon_2 m^u, \epsilon_2 c^u \) instead to evaluate the upper bounds of \( \frac{\| \Delta x \|_\infty}{\| x \|_\infty} \) or \( \frac{\| \Delta x \|_\infty}{\| x \|_\infty} \), since for large values of \( m \), the storage and computation of matrix \( K \) in (12) might exceed the maximal memory of the computer.

The tabulated results in Table 4 show that for evaluating the forward error \( \frac{\| \Delta x \|_2}{\| x \|_2} \) of large and structured MTLS problems, the unstructured and structured normwise condition numbers both become worse when the matrix \( [A, \lambda \hat{b}] \) becomes badly scaled, while the estimates based on structured normwise condition numbers are tighter than those based on unstructured ones. In evaluating \( \frac{\| \Delta x \|_\infty}{\| x \|_\infty} \) or \( \frac{\| \Delta x \|_\infty}{\| x \|_\infty} \), the estimates based on structured mixed and componentwise condition numbers are sharper than unstructured ones, and they are only one order of magnitude larger than corresponding relative forward errors. The results indicate that structured mixed and componentwise condition numbers are more robust and preferred for badly scaled and structured problems.
Table 4. Comparison of upper bounds with forward errors for perturbed MTLS problems.

\[
\begin{array}{cccccccccccc}
\lambda & \frac{\Delta x}{\|x\|_2} & \varepsilon_1 \kappa_4 & \varepsilon_1 \kappa^4 & \frac{\Delta x}{\|x\|_\infty} & \varepsilon_2 m^u & \varepsilon_2 m^f & \|\frac{\Delta x}{x}\|_\infty & \varepsilon_2 c^d & \varepsilon_2 c^f \\
\hline
10^{-2} & 4.6e-9 & 9.6e-7 & 4.3e-8 & 4.4e-9 & 1.0e-7 & 2.5e-8 & 4.4e-6 & 1.1e-4 & 2.8e-5 \\
10^{-4} & 5.2e-9 & 8.9e-5 & 4.1e-6 & 6.2e-9 & 8.7e-8 & 2.7e-8 & 1.5e-6 & 2.4e-5 & 1.2e-5 \\
10^{-6} & 5.7e-9 & 8.6e-3 & 3.9e-4 & 5.6e-9 & 1.1e-7 & 2.7e-8 & 4.5e-7 & 1.1e-5 & 3.8e-6 \\
\end{array}
\]

\[
\begin{array}{cccccccccccc}
\lambda & \frac{\Delta x}{\|x\|_2} & \varepsilon_1 \kappa_4 & \varepsilon_1 \kappa^4 & \frac{\Delta x}{\|x\|_\infty} & \varepsilon_2 m^u & \varepsilon_2 m^f & \|\frac{\Delta x}{x}\|_\infty & \varepsilon_2 c^d & \varepsilon_2 c^f \\
\hline
10^{-2} & 5.2e-9 & 1.4e-6 & 4.2e-8 & 4.0e-9 & 1.9e-7 & 3.9e-8 & 1.9e-6 & 6.4e-5 & 2.7e-5 \\
10^{-4} & 1.2e-8 & 1.0e-4 & 3.3e-6 & 1.2e-8 & 2.4e-7 & 5.9e-8 & 4.6e-6 & 1.7e-4 & 6.0e-5 \\
10^{-6} & 1.5e-8 & 1.3e-2 & 4.2e-4 & 9.7e-9 & 1.5e-7 & 3.6e-8 & 8.4e-6 & 1.5e-4 & 4.3e-5 \\
\end{array}
\]

5. Conclusion

In this paper, we first present a new closed formula for the first order perturbation estimate of the MTLS solution, and then reveal the close relation of our new perturbation result to the one in [25]. Based on the new formula, normwise, mixed and componentwise condition numbers and corresponding structured condition numbers of the MTLS problem are also derived. Matrix Kronecker product operations involved in the expressions make the computation costly, and hence we propose to simplify the expressions to improve the computational efficiency. For the normwise condition number, we show that it can be transformed into several compact forms, and the perturbation bound is also an alternative. From a number of numerical experiments, we have seen that the new forms and bounds for the normwise condition number have great computational efficiency and require less storage. For structured, sparse and badly-scaled MTLS problems, the sparse pattern and the magnitude of the entries are better utilized in (structured) mixed and componentwise condition numbers than the normwise condition number. It is more suitable to adopt (structured) mixed and componentwise condition numbers to measure the conditioning of badly scaled and structured MTLS problems.

Acknowledgments

The authors would like to thank the handling editor Ren-Cang Li and two anonymous referees for their valuable comments and suggestions which led to great improvements to this paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research is partially supported by the National Natural Science Foundation of China (11001167).

References

[1] Björck Å. Numerical methods for least squares problems. Philadelphia: SIAM; 1996.
[2] Golub GH, Van Loan CF. An analysis of the total least squares problem. SIAM J Numer Anal. 1980;17(6):883–893.
[3] Liu QH, Chen CP, Zhang Q. Perturbation analysis for total least squares problems with linear equality constraint. Appl Numer Math. 2021;161:69–81.
[4] Liu QH, Jin SF, Yao L, et al. The revisited total least squares problems with linear equality constraint. Appl Numer Math. 2020;152:275–284.
[5] Liu QH, Wang MH. On the weighting method for mixed least squares-total least squares problems. Numer Linear Algebra Appl. 2017;24(5):e2094.
[6] Paige CC, Wei MS. Analysis of the generalized total least squares problem $AX \approx B$ when some columns of $A$ are free of error. Numer Math. 1993;65(1):177–202.
[7] Yan SJ, Fan JY. The solution set of the mixed LS-TLS problem. Int J Comput Math. 2001;77(4):545–561.
[8] Gleser LJ. Estimation in a multivariate "errors in variables" regression model: large sample results. Ann Statist. 1981;9:24–44.
[9] Stoica P, Söderström T. Bias correction in least squares identification. Int J Control. 1982;35(3):449–457.
[10] Van Huffel S, Vandewalle J. Comparison of total least squares and instrumental variable methods for parameter estimation of transfer function models. Int J Control. 1989;50:1039–1056.
[11] Golub GH, Van Loan CF. Matrix computation. 4th ed. Baltimore: Johns Hopkins University Press; 2013.
[12] Van Huffel S, Vandewalle J. The total least-squares problem: computational aspects and analysis. In: Frontier in applied mathematics. Philadelphia: SIAM; 1991.
[13] Xie PP, Xiang H, Wei YM. Randomized algorithms for total least squares problems. Numer Linear Algebra Appl. 2018;26:e2219.
[14] Zhou LM, Lin LJ, Wei YM, et al. Perturbation analysis and condition numbers of scaled total least squares problems. Numer Algorithms. 2009;51(3):381–399.
[15] Baboulin M, Gratton S. A contribution to the conditioning of the total least squares problem. SIAM J Matrix Anal Appl. 2011;32(3):685–699.
[16] Jia ZX, Li BY. On the condition number of the total least squares problem. Numer Math. 2013;125(1):61–87.
[17] Li BY, Jia ZX. Some results on condition numbers of the scaled total least squares problem. Linear Algebra Appl. 2011;435(3):674–686.
[18] Wang SX, Li HY, Yang H. A note on the condition number of the scaled total least squares problem. Calcolo. 2018;55:46.
[19] Xie PP, Xiang H, Wei YM. A contribution to perturbation analysis for total least squares problems. Numer Algorithms. 2017;75(2):381–395.
[20] Diao HA, Sun Y. Mixed and componentwise condition numbers for a linear function of the solution of the total least squares problem. Linear Algebra Appl. 2018;544:1–29.
[21] Meng QL, Diao HA, Bai ZJ. Condition numbers for the truncated total least squares problem and their estimations. Numer Linear Algebra Appl. 2021;28(5):e2369.
[22] Meng LS, Zheng B, Wei YM. Condition numbers of the multi-dimensional total least squares problems having more than one solution. Numer Algorithms. 2020;84:887–908.
[23] Zheng B, Meng LS, Wei YM. Condition numbers of the multidimensional total least squares problem. SIAM J Matrix Anal Appl. 2017;38(3):924–948.
[24] Van Huffel S, Vandewalle J. Analysis and properties of the generalized total least squares problem $AX \approx B$ when some or all columns in $A$ are subject to error. SIAM J Matrix Anal Appl. 1989;10:294–315.
[25] Zheng B, Yang ZS. Perturbation analysis for mixed least squares-total least squares problems. Numer Linear Algebra Appl. 2019;26:e2239.
[26] Langville AN, Stewart WJ. The Kronecker product and stochastic automata networks. J Comput Appl Math. 2004;167:429–447.
[27] Graham A. Kronecker products and matrix calculus with application. New York: Wiley; 1981.
[28] Rice JR. A theory of condition. SIAM J Numer Anal. 1966;3:287–310.
[29] Geurts AJ. A contribution to the theory of condition. Numer Math. 1982;39:85–96.
[30] Gohberg I, Koltracht I. Mixed, componentwise, and structured condition numbers. SIAM J Matrix Anal Appl. 1993;14:688–704.
[31] Cucker F, Diao HA, Wei YM. On mixed and componentwise condition numbers for Moore-Penrose inverse and linear least squares problems. Math Comp. 2007;76(258):947–963.

[32] Higham NJ. Accuracy and stability of numerical algorithms. 2nd ed. Philadelphia: SIAM; 2002.

[33] Diao HA, Wei YM, Xie PP. Small sample statistical condition estimation for the total least squares problem. Numer Algorithms. 2017;75(2):1–21.

[34] Björck Å, Heggernes P, Matstoms P. Methods for large scale total least squares problems. SIAM J Matrix Anal Appl. 2000;22(2):413–429.

[35] Young P. The instrumental variable method: a practical approach to identification and parameter estimation. 7th IFAC/IFORS Symposium on Identification and System Parameter Estimation. 1985;18(5):1–15.