BOUNDS FOR THE LEAST SQUARES RESIDUAL USING SCALED TOTAL LEAST SQUARES

Christopher C. Paige
School of Computer Science, McGill University
Montreal, Quebec, Canada, H3A 2A7
paige@cs.mcgill.ca

Zdeněk Strakoš
Institute of Computer Science, Academy of Sciences of the Czech Republic,
Pod Vodárenskou věží 2, 182 07 Praha 8, Czech Republic
strakos@cs.cas.cz

Abstract The standard approaches to solving overdetermined linear systems $Ax \approx b$ construct minimal corrections to the data to make the corrected system compatible. In ordinary least squares (LS) the correction is restricted to the right hand side b, while in scaled total least squares (Scaled TLS) [10; 7] corrections to both b and A are allowed, and their relative sizes are determined by a real positive parameter γ. As $\gamma \to 0$, the Scaled TLS solution approaches the LS solution. Fundamentals of the Scaled TLS problem are analyzed in our paper [7] and in the contribution in this book entitled *Unifying least squares, total least squares and data least squares*.

This contribution is based on the paper [8]. It presents a theoretical analysis of the relationship between the sizes of the LS and Scaled TLS *corrections* (called the LS and Scaled TLS distances) in terms of γ. We give new upper and lower bounds on the LS distance in terms of the Scaled TLS distance, compare these to existing bounds, and examine the tightness of the new bounds.

This work can be applied to the analysis of iterative methods which minimize the residual norm [9; 6].

Keywords: ordinary least squares, scaled total least squares, singular value decomposition, linear equations, least squares residual.

Introduction

Consider an overdetermined approximate linear system

$$Ax \approx b, \quad A \text{ an } n \text{ by } k \text{ matrix}, \quad b \text{ an } n\text{-vector}, \quad b \notin \mathcal{R}(A), \tag{1}$$

where $\mathcal{R}(M)$ denote the range (column space) of a matrix M. In LS we seek (we use $\|\cdot\|$ to denote the vector 2-norm)

$$\text{LS distance} \quad \equiv \quad \min_{r,x} \|r\| \quad \text{subject to} \quad Ax = b - r. \tag{2}$$

In Scaled TLS, for a given parameter $\gamma > 0$, x, G and r are sought to minimize the Frobenius (F) norm in

$$\text{Scaled TLS distance} \quad \equiv \quad \min_{r,G,x} \|[r, G]\|_F \quad \text{s. t.} \quad (A+G)x\gamma = b\gamma - r. \tag{3}$$

We call the $x = x(\gamma)$ which minimizes this distance the Scaled TLS solution of (3). Here the relative sizes of the corrections G and r in A and $b\gamma$ are determined by the real scaling parameter $\gamma > 0$. As $\gamma \to 0$ the Scaled TLS solution approaches the LS solution. The formulation (3) is studied in detail in [7]. We present an introduction to and refine some results of [7] in our contribution *Unifying least squares, total least squares and data least squares* presented in this book. Here we follow the notation introduced there. In applications γ can have a statistical interpretation, see for example [7, §1], but here we regard γ simply as a variable.

Scaled TLS solutions can be found via the singular value decomposition (SVD). Let $\sigma_{min}(\cdot)$ denote the smallest singular value of a matrix, and let P_k be the orthogonal projector onto the left singular vector subspace of A corresponding to $\sigma_{min}(A)$. The bounds presented here will assume

$$\text{the } n \times (k+1) \text{ matrix } [A, b] \text{ has rank } k+1, \text{ and } \quad P_k b \neq 0. \tag{4}$$

We showed in [7, (3.7)] that this implied

$$0 < \sigma(\gamma) \equiv \sigma_{min}([A, b\gamma]) < \sigma_{min}(A) \quad \text{for all } \gamma > 0. \tag{5}$$

In this case the unique solution of the Scaled TLS problem (3) is (in theory) obtained from scaling the right singular vector of $[A, b\gamma]$ corresponding to $\sigma_{min}([A, b\gamma])$, and the norm of the Scaled TLS correction satisfies, for a given $\gamma > 0$ (see for example [7, (1.9)], or [5, §12.3] when $\gamma = 1$),

$$\text{Scaled TLS distance in (3)} \quad = \quad \sigma_{min}([A, b\gamma]). \tag{6}$$

The paper [8] and the presentation of the bounds in this contribution are greatly simplified by only dealing with problems where (4) holds. The assumption (4) is equivalent to that in [7, (1.10)] plus the restriction $b \notin \mathcal{R}(A)$, which eliminates the theoretically trivial case $b \in \mathcal{R}(A)$. It is sufficient to note here that nearly all practical overdetermined problems will already satisfy (4), but any overdetermined (and incompatible) problem that does not can be reduced to one that does, see [7, §8], and the bounds presented here with this assumption will be applicable to the original problem.

It is known that (see for example [7, (6.3)])

$$\begin{aligned} \lim_{\gamma\to 0} \frac{\text{Scaled TLS distance in (3)}}{\gamma} &= \lim_{\gamma\to 0} \frac{\sigma_{min}([A, b\gamma])}{\gamma} \\ &= \|r\|, \text{ the LS distance in (2)}, \end{aligned} \tag{7}$$

but here we examine the relationship between these distances for *any* $\gamma > 0$. This will bound the rate at which these quantities approach each other for small γ, as well as provide bounds on the LS distance in terms of $\sigma_{min}([A, b\gamma])$, and *vice versa*, for all $\gamma > 0$. It will in general simplify the presentation to assume $\gamma > 0$, since when $\gamma = 0$ is meaningful, the values will be obvious.

Van Huffel and Vandewalle [3] derived several useful bounds for TLS versus LS (the $\gamma = 1$ case). Our results extend some of these to the case of general $\gamma > 0$, as well as provide new bounds.

The contribution is organized as follows. In Section 1 we present our main result, in particular, bounds on the least squares residual norm $\|r\|$ (LS distance) in terms of the scaled total least squares distance $\sigma_{min}([A, b\gamma])$. We show how good these bounds are, and how varying γ gives important insights into the asymptotic relationship between the LS and Scaled TLS distances. In Section 2 we compare our bounds to previous results. In Section 3 we analyze the ratio of the minimal singular values of $[A, b\gamma]$ and A which determines the tightness of the presented bounds.

1. Main result

Our main result relating the LS distance $\|r\|$ to the Scaled TLS distance $\sigma_{min}([A, b\gamma])$ is formulated in the following theorem, see [8, Theorem 4.1 and Corollary 6.1].

Theorem 1 *Given a scalar $\gamma > 0$, and an n by $k+1$ matrix $[A, b]$, use $\sigma(\cdot)$ to denote singular values and $\|\cdot\|$ to denote 2-norms. If r and x*

solve $\min_{r,x} \|r\|$ *subject to* $Ax = b - r$, *and (4) holds, then*

$$0 < \theta(\gamma) \equiv \frac{\sigma_{min}([A, b\gamma])}{\sigma_{max}(A)} \leq \delta(\gamma) \equiv \frac{\sigma_{min}([A, b\gamma])}{\sigma_{min}(A)} < 1, \qquad (8)$$

and we have bounds on the LS residual norm $\|r\|$ *in terms of the Scaled TLS distance* $\sigma_{min}([A, b\gamma])$*:*

$$\lambda_r \equiv \sigma_{min}([A, b\gamma])\{\gamma^{-2} + \|x\|^2\}^{\frac{1}{2}} < \sigma_{min}([A, b\gamma])\,\{\gamma^{-2} + \frac{\|x\|^2}{1-\theta(\gamma)^2}\}^{\frac{1}{2}}$$
$$\leq \|r\| \leq \mu_r \equiv \sigma_{min}([A, b\gamma])\{\gamma^{-2} + \frac{\|x\|^2}{1-\delta(\gamma)^2}\}^{\frac{1}{2}}. \qquad (9)$$

Equivalently,

$$\lambda_\sigma \equiv \|r\|/\{\gamma^{-2} + \frac{\|x\|^2}{1-\delta(\gamma)^2}\}^{\frac{1}{2}} \leq \sigma_{min}([A, b\gamma])$$
$$\leq \|r\|/\{\gamma^{-2} + \frac{\|x\|^2}{1-\theta(\gamma)^2}\}^{\frac{1}{2}} \leq \mu_\sigma \equiv \|r\|/\{\gamma^{-2} + \|x\|^2\}^{\frac{1}{2}}. \quad (10)$$

In addition to that, $\delta(\gamma)$ *is bounded as*

$$\frac{\gamma\|r\|}{\|[A, b\gamma]\|} \leq \delta(\gamma) \leq \frac{\gamma\|r\|}{\sigma_k([A, b\gamma])} \leq \frac{\gamma\|r\|}{\sigma_{min}(A)}. \qquad (11)$$

■

We see that the difference between the upper and the lower bounds in (9) depends on the size of $(1-\delta(\gamma)^2)^{-1}$. If $\delta(\gamma) \ll 1$, then this difference will be very small. Bounds in (11) give us some indication of the size of $\delta(\gamma)$. We see from (11) that if $\gamma\|r\|$ is small compared with $\sigma_k([A, b\gamma])$ then $\delta(\gamma) \ll 1$, but if $\gamma\|r\|$ is not small compared with $\|[A, b\gamma]\|$ then $\delta(\gamma)$ cannot be small. If $[A, b\gamma]$ is well-conditioned in the sense that $\sigma_{min}([A, b\gamma])$ is not too much smaller than $\|[A, b\gamma]\|$, then (11) gives us a very good idea of $\delta(\gamma)$. We will study $\delta(\gamma)$ in more detail in Section 3.

A crucial aspect of Theorem 1 is that it gives both an upper and a lower bound on the minimum residual norm $\|r\|$, or on $\sigma_{min}([A, b\gamma])$, which is the Scaled TLS distance in (3). The weaker lower bound in (9), or upper bound in (10), is sufficient for many uses, and is relatively easy to derive, but the upper bound in (9), or lower bound in (10), is what makes the theorem strong.

The following corollary [8, Corollary 4.2] examines the *tightness* of the bounds (9)–(10), to indicate just how good they can be. In fact it

shows that *all* the relative gaps go to zero (as functions of the scaling parameter γ) at least as fast as $O(\gamma^4)$.

Corollary 1 *Under the same conditions as in Theorem 1, with* $\sigma \equiv \sigma(\gamma) \equiv \sigma_{min}([A, b\gamma])$*, the notation in (9)–(10), and*

$$\begin{aligned} &\eta_r \equiv (\|r\| - \lambda_r)/\|r\|, \quad \eta_\sigma \equiv (\sigma - \lambda_\sigma)/\sigma, \\ &\zeta_r \equiv (\mu_r - \lambda_r)/\|r\|, \quad \zeta_\sigma \equiv (\mu_\sigma - \lambda_\sigma)/\sigma, \end{aligned} \tag{12}$$

we have the following bounds

$$\begin{aligned} &0 < \eta_r \le \zeta_r, \qquad 0 < \eta_\sigma \le \zeta_\sigma, \\ &0 < \zeta_r, \zeta_\sigma < \frac{\gamma^2\|x\|^2}{2 + \gamma^2\|x\|^2} \cdot \frac{\delta(\gamma)^2}{1 - \delta(\gamma)^2} \to 0 \quad as \quad \gamma \to 0, \end{aligned} \tag{13}$$

where the upper bound goes to zero at least as fast as $O(\gamma^4)$. ■

Thus when $\delta(\gamma) \ll 1$, or γ is small, the upper and lower bounds in (9)–(10) are not only very good, but very good in a *relative* sense, which is important for small $\|r\|$ or $\sigma_{min}([A, b\gamma])$. We see Corollary 1 makes precise a nice theoretical observation with practical consequences — small γ ensures very tight bounds (9) on $\|r\|$. In particular, for small γ we see

$$\|r\| \;\approx\; \lambda_r \equiv \sigma_{min}([A, b\gamma])\,\{\gamma^{-2} + \|x\|^2\}^{\frac{1}{2}}, \tag{14}$$

and the relative error is bounded above by $O(\gamma^4)$.

When $\delta(\gamma) < 1$, [3, Thm. 2.7] showed (for $\gamma = 1$) the closed form TLS solution $x\gamma = x(\gamma)\gamma$ of (3) is

$$x(\gamma)\gamma = \{A^TA - \sigma^2_{min}([A, b\gamma])I\}^{-1}A^Tb\gamma,$$

and with $r_{ScaledTLS} \equiv b\gamma - Ax(\gamma)\gamma$, [3, (6.19)] showed (for $\gamma = 1$)

$$\|r_{ScaledTLS}\| = \sigma_{min}([A, b\gamma])(1 + \|x(\gamma)\gamma\|^2)^{\frac{1}{2}}. \tag{15}$$

Relation (14) can be seen to give an analogue of this for the LS solution: since $r\gamma = b\gamma - Ax\gamma$ in (2), the bounds (9), (11) and (13) show a strong relationship between $\gamma\|r\|$ and $\sigma_{min}([A, b\gamma])$ for small $\delta(\gamma)$, γ, $\|r\|$ or $\|x\|/(1 - \delta(\gamma)^2)$:

$$\gamma\|r\| \;\approx\; \sigma_{min}([A, b\gamma])\,\{1 + \gamma^2\|x\|^2\}^{\frac{1}{2}}. \tag{16}$$

The assumption $P_kb \neq 0$ in (4) is not necessary for proving the bounds (9)–(10). From the proof of Theorem 1 in [8] it is clear that these bounds

only require $\delta(\gamma) < 1$. However $\delta(\gamma) < 1$ does not guarantee $P_k b \neq 0$. When $P_k b = 0$, $\|r\|$ contains no information whatsoever about $\sigma_{min}(A)$, while the bounds do. By assuming $P_k b \neq 0$ we avoid this inconsistency. Moreover, we will consider various values of the parameter γ, and so we prefer the theorem's assumption to be independent of γ.

We end this section by a comment on possible consequences of Theorem 1 for understanding methods for large Scaled TLS problems. For small $\delta(\gamma)$, γ, $\|r\|$ or $\|x\|^2/(1-\delta(\gamma)^2)$, (10) with (11) and (13) show

$$\sigma_{k+1}^2([A, b\gamma]) \approx \frac{\gamma^2\|r\|^2}{1+\gamma^2\|x\|^2} = \|[A, b\gamma]\begin{pmatrix} -x\gamma \\ 1 \end{pmatrix}\|^2 / \|\begin{pmatrix} -x\gamma \\ 1 \end{pmatrix}\|^2;$$

so the Scaled TLS distance is well approximated using the Rayleigh quotient corresponding to the unique LS solution of $Ax\gamma = b\gamma - r\gamma$. As pointed out by Åke Björck in a personal communication, this may help to explain the behaviour of algorithms proposed in [1].

2. Comparison with previous bounds

The best previously published bounds relating LS and TLS distances appear to be those of Van Huffel and Vandewalle [3]. The relevant bounds of that reference, and a new bound, can be derived from (9), and we present them as a corollary (cf. [8, Corollary 5.1]).

Corollary 2 *Under the same conditions and assumptions as in Theorem 1, with* $\sigma(\gamma) \equiv \sigma_{min}([A, b\gamma])$, $\delta(\gamma) \equiv \sigma_{min}([A, b\gamma])/\sigma_{min}(A)$,

$$\begin{aligned} \frac{\sigma_{min}([A, b\gamma])}{\gamma} &\leq \frac{\sigma_{min}([A, b\gamma])}{\gamma}\{1 - \frac{\sigma_{min}^2([A, b\gamma])}{\|A\|^2} + \frac{\|b\|^2\gamma^2}{\|A\|^2}\}^{\frac{1}{2}} \\ &\leq \|r\| \leq \frac{\sigma_{min}([A, b\gamma])}{\gamma}\{1 - \delta(\gamma)^2 + \frac{\|b\|^2\gamma^2}{\sigma_{min}(A)^2}\}^{\frac{1}{2}}. \end{aligned} \tag{17}$$

■

When $\gamma = 1$ the weaker lower bound and the upper bound in (17) are the equivalents for our situation of (6.34) and (6.35) in [3]. The stronger lower bound seems new. A slightly weaker upper bound was derived in [2, (2.3)]. Experimental results presented, e.g., in [8] demonstrate that our bounds in (9) can be significantly better than those in (17). The relationship of these bounds is, however, intricate. While (17) was in [8, Corollary 5.1] derived from (9), it is not *always* true that the latter is tighter. When $\delta(\gamma) \approx 1$ and $\|r\| \approx \|b\|$, it is possible for the upper bound in (17) to be smaller than that in (9). But in this case

$\sigma_{min}([A, b\gamma]) \approx \sigma_{min}(A)$, and then the upper bound in (17) becomes the trivial $\|r\| \lesssim \|b\|$. Summarizing, when the upper bound in (17) is tighter than the upper bound in (9), the former becomes trivial and the latter is irrelevant.

The bounds (17) and (9) differ because the easily available $\|x\|$ in (9) was replaced by its upper and lower bounds to obtain (17). But there is another reason (9) is preferable to (17). The latter bounds require knowledge of $\sigma_{min}(A)$, as well as $\sigma_{min}([A, b\gamma])$. Admittedly (8) shows we also need these to know $\delta(\gamma)$ exactly, but, assuming that (4) holds, we know $\delta(\gamma) < 1$, and is bounded away from 1 always (see Theorem 2 in the following section). In fact there are situations where we know $\delta(\gamma) \ll 1$. Thus (9) is not only simpler and often significantly stronger than (17), it is more easily applicable.

3. Tightness parameter

The results presented above show the crucial role of the parameter $\delta(\gamma) = \sigma_{min}([A, b\gamma])/\sigma_{min}(A)$. It represents a ratio of the smallest singular value of the matrix appended by a column (here $[A, b\gamma]$) to the smallest singular value of the original matrix (here A). Though the definition is simple, the nature of $\delta(\gamma)$ is very subtle and its behaviour is very complicated.

Let the $n \times k$ matrix A have rank k and singular values σ_i with singular value decomposition (SVD)

$$A = U_A \Sigma V^T, \quad \Sigma \equiv \mathrm{diag}(\sigma_1, \ldots, \sigma_k), \quad \sigma_1 \geq \ldots \geq \sigma_k > 0. \tag{18}$$

Here U_A is $n \times k$ matrix, $U_A^T U_A = I_k$, Σ is $k \times k$, and $k \times k$ V is orthogonal. Let

$$a \equiv (\alpha_1, \ldots, \alpha_k)^T \equiv [u_1, \ldots, u_k]^T b = U_A^T b. \tag{19}$$

The elements of a are the components of the vector of observations b in the directions of the left singular vectors of the data matrix A.

Assume (4) holds. Then using the notation in (18)–(19), $0 < \sigma(\gamma) < \sigma_k \equiv \sigma_{min}(A)$ holds for all $\gamma > 0$, and the Scaled TLS distance in (3) is $\sigma(\gamma) \equiv \sigma_{min}([A, b\gamma])$, which is the smallest positive solution of

$$0 = \psi_k(\sigma(\gamma), \gamma) = \gamma^2 \|r\|^2 - \sigma(\gamma)^2 - \gamma^2 \sigma(\gamma)^2 \sum_{i=1}^{k} \frac{|\alpha_i|^2}{\sigma_i^2 - \sigma(\gamma)^2}. \tag{20}$$

Moreover, if (4) holds and $\gamma > 0$, then $0 < \delta(\gamma) < 1$, and $\delta(\gamma)$ increases as γ increases, and decreases as γ decreases, strictly monotonically. This was derived in [7, §4]. With $\gamma = 1$, (20) was derived in [4], see also [3, Thm. 2.7, & (6.36)]. These latter derivations assumed the weaker

condition $\sigma_{min}([A,b]) < \sigma_{min}(A)$, and so do not generalize to Scaled TLS for all $\gamma > 0$, see [7].

Our bounds containing the factor $(1-\delta(\gamma)^2)^{-1}$ would be useless if $\delta(\gamma) = 1$ and of limited value when $\delta(\gamma) \approx 1$. The following theorem ([8, Theorem 3.1]) shows that when (4) holds, $\delta(\gamma)$ is *bounded away from unity* for all γ, giving an upper bound on $(1-\delta(\gamma)^2)^{-1}$. It is important that these bounds exist, but remember they are worst case bounds, and give no indication of the sizes of $\delta(\gamma)$ or $(1-\delta(\gamma)^2)^{-1}$ for the values of γ we will usually be interested in.

Theorem 2 *With the notation and assumptions of (18)–(20), let $n \times k$ A have singular values $\sigma_1 \geq \ldots \geq \sigma_j > \sigma_{j+1} = \ldots = \sigma_k > 0$. Then since (4) holds,*

$$\|P_k b\|^2 = \sum_{i=j+1}^{k} |\alpha_i|^2 > 0, \tag{21}$$

$$\delta(\gamma)^2 \equiv \frac{\sigma_{min}^2([A,b\gamma])}{\sigma_k^2} \leq \frac{\|r\|^2}{\|P_k b\|^2 + \|r\|^2} < 1 \quad \textit{for all } \gamma \geq 0, \tag{22}$$

$$(1-\delta(\gamma)^2)^{-1} \leq 1 + \|r\|^2/\|P_k b\|^2 \quad \textit{for all } \gamma \geq 0, \tag{23}$$

where P_k is described just before (4). ■

This shows that when (4) holds, $\delta(\gamma)$ is bounded away from unity, so $\sigma_{min}([A,b\gamma])$ is bounded away from $\sigma_{min}(A)$, for all γ.

The inequality (22) has a useful explanatory purpose. We cannot have $\delta(\gamma) \approx 1$ unless $P_k b$, the projection of b onto the left singular vector subspace of A corresponding to $\sigma_{min}(A)$, is very small compared to r. It is straightforward to show that replacing A by

$$\tilde{A} = A - \sum_{i=j+1}^{k} u_i \sigma_{min}(A) v_i^T$$

in (2) increases the square of the LS residual by $\|P_k b\|^2$, thus giving a small relative change when $P_k b$ is small compared to r. This confirms that the criterion (4) (see also [7, (1.10]) is exactly what is needed. When $P_k b = 0$ the smallest singular value $\sigma_{min}(A)$ has no influence to the solution of the LS problem and should be eliminated from our considerations. When $P_k b$ is small, elimination of $\sigma_{min}(A)$ (replacing of A by $\tilde{A}$) has little effect on the LS solution.

We will finish this contribution by a short note illustrating the conceptual and technical complications which arise when the assumption (4)

is not used. First we must analyze when $\delta(\gamma) = 1$. The necessary and sufficient conditions for $\delta(\gamma) = 1$ were given in [7, Theorem 3.1]. Here we will explain the main idea in relation to the secular equation (20). Let $n \times k$ A have singular values $\sigma_1 \geq \ldots \geq \sigma_j > \sigma_{j+1} = \ldots = \sigma_k > 0$. When $\delta(\gamma) = 1$, b has no components in the left singular vector subspace of A corresponding to $\sigma_{min}(A)$, $P_k b = 0$, $\alpha_{j+1} = \ldots = \alpha_k = 0$ and the matrix with the appended column $[A, b\gamma]$ has $k - j$ singular values equal to $\sigma_{min}(A)$. The singular values of $[A, b\gamma]$ different from those of A are solutions $\sigma(\gamma)$ of the deflated secular equation, see [11, Ch2, §47, pp. 103-104],

$$0 = \psi_j(\sigma(\gamma), \gamma) = \gamma^2 \|r\|^2 - \sigma(\gamma)^2 - \gamma^2 \sigma(\gamma)^2 \sum_{i=1}^{j} \frac{|\alpha_i|^2}{\sigma_i^2 - \sigma(\gamma)^2}, \qquad (24)$$

where the summation term is ignored if all singular values of A are equal. Note that $\psi_j(0, \gamma) > 0$, so that $\delta(\gamma) = 1$ requires that $\psi_j(\sigma_k, \gamma) \geq 0$ (if $\psi_j(\sigma_k, \gamma) < 0$, then the deflated secular equation (24) must have a positive solution σ less than $\sigma_{min}(A)$ which contradicts the condition $\delta(\gamma) = 1$).

It is interesting to note that for the particular choice of $\gamma = \sigma_k / \|r\|$, the condition $\psi_j(\sigma_k, \gamma) \geq 0$ is equivalent to $\alpha_1 = \ldots = \alpha_j = 0$, i.e. $U_A^T b = 0$ and $r = b$. In the other words, $\delta(\gamma) < 1$ for $\gamma < \sigma_k / \|b\|$ (the last column of the matrix $[A, b\gamma]$ has for $\gamma < \sigma_k / \|b\|$ norm less than $\sigma_{min}(A)$), and for the choice $\gamma_b \equiv \sigma_k / \|b\|$ the condition $\delta(\gamma_b) = 1$ is equivalent to the fact that in the LS problem (2) the LS solution $x = 0$ is trivial and $r = b$. When this particular γ_b is used with (17), we obtain (see also [6, Section 2])

$$\delta(\gamma_b)\|b\| \leq \|r\| \leq \delta(\gamma_b)\|b\| \{2 - \delta(\gamma_b)^2\}^{\frac{1}{2}}. \qquad (25)$$

The results presented here have been successfully applied outside the Errors-in-Variables Modeling field for analysis of convergence and numerical stability of Krylov subspace methods, see [9], [6].

4. Conclusion

Summarizing, our contribution (which is based on [8]) shows new bounds for the LS residual norm $\|r\| = \min_x \|b - Ax\|$ in terms of the Scaled TLS distance $\sigma_{\min}([A, b\gamma])$, and presents several important corollaries describing the tightness of the bounds and their dependence on the parameter γ. The bounds were seen to be very good when $\sigma_{\min}([A, b\gamma])$ was sufficiently smaller than $\sigma_{\min}(A)$. When $\sigma_{\min}([A, b\gamma]) \approx \sigma_{\min}(A)$, it

is shown that the smallest singular value $\sigma_{\min}(A)$ and its singular vectors did not play a significant role in the solution of the LS problem.

Acknowledgments

This work was supported by NSERC of Canada Grant OGP0009236 and by the GA AS CR under grant A2030801. Part of this work was performed while Zdenek Strakos was visiting Emory University, Atlanta, GA, U.S.A.

References

[1] Å. Björck, P. Heggerness, and P. Matstoms, *Methods for large scale total least squares problems*, SIAM J. Matrix Anal. Appl., 22:413–442, 2000.

[2] A. Greenbaum, M. Rozložník and Z. Strakoš. *Numerical behavior of the modified Gram-Schmidt GMRES implementation.* BIT, 37(3):706–719, 1997.

[3] S. Van Huffel and J. Vandewalle. *The Total Least Squares Problem: Computational Aspects and Analysis*, SIAM Publications, Philadelphia PA, 1991.

[4] G. H. Golub and C. F. Van Loan. *An analysis of the total least squares problem*, SIAM J. Numer. Anal., 17:883–893, 1980.

[5] G. H. Golub and C. F. Van Loan. *Matrix Computations*, The Johns Hopkins University Press, Baltimore MD, third ed., 1996.

[6] J. Liesen, M. Rozložník and Z. Strakoš. *On Convergence and Implementation of Residual Minimizing Krylov Subspace Methods*, to appear in SIAM J. Sci. Comput.

[7] C. C. Paige and Z. Strakoš. *Scaled total least squares fundamentals*, to appear in Numerische Mathematik.

[8] C. C. Paige and Z. Strakoš. *Bounds for the least squares distance using scaled total least squares*, to appear in Numerische Mathematik.

[9] C. C. Paige and Z. Strakoš. *Residual and backward error bounds in minimum residual Krylov subspace methods*, submitted to SIAM J. Sci. Comput., October 2000.

[10] B. D. Rao. *Unified treatment of LS, TLS and truncated SVD methods using a weighted TLS framework*, In: S. Van Huffel (editor), *Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modelling*, pp. 11–20, SIAM Publications, Philadelphia PA, 1997.

[11] J. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford England, 1965.