Zdeněk Strakoš
Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Prague 8
Czech Republic

Time requested: 45
strakos@cs.cas.cz
www.cs.cas.cz/~strakos
tel: +420 266 053 290 , fax: +420 286 585 789
18 years since degree

Core problems in $Ax \approx b$ — analysis of the total least squares problem revisited

Chris Paige
Zdeněk Strakoš

At the last Householder meeting, Chris Paige presented a plenary lecture titled *Bidiagonalization and Approximate Linear Systems* [3]. By suggesting a talk on a complementary topic, we risk the impression that we are going to repeat ourselves. We believe that we have a good reason for taking this risk. The recent results, which complement our understanding from 2002, revise and extend, to our opinion, the analysis of the total least problem essentially started in [1] and then extensively developed in [2] and elsewhere, see [7], [5], [4].

Consider determining x from the real linear system

$$Ax \approx b, \quad A \text{ a nonzero } n \text{ by } k \text{ matrix}, \quad b \text{ a nonzero } n\text{-vector}. \tag{1}$$

It will simplify matters, and nothing will be lost, if we just consider the TLS problem for (1)

$$\text{TLS distance} \equiv \min_{g,E,x} \| [g, E] \|_F \quad \text{subject to} \quad (A + E)x = b + g. \tag{2}$$

Golub and Van Loan [1] analyzed this problem (in a more general setting using weights and scaling parameters, which are for simplicity not considered here). In a nutshell, the constraint in (2) is equivalent to

$$([b, A] + [g, E]) \begin{bmatrix} -1 \\ x \end{bmatrix} = 0.$$

This suggest that the TLS solution is determined by the smallest perturbation of $[b, A]$ which makes it rank defficient. If the right singular vector corresponding to the smallest singular value of $[b, A]$ has nonzero first component, then scaling it so that the first component is -1 gives the TLS solution.

If the smallest singular value of $[b, A]$ is repeated, and if a corresponding right singular vector with nonzero first component can still be found, then the TLS problem lacks a unique solution. Golub and Van Loan [1, pp. 885-886], and later Van Huffel and Vandewalle [2, Thm. 3.7, p. 58], showed how to single out a unique minimum 2-norm TLS solution in this case.

If no right singular vector corresponding to the smallest singular value of $[b, A]$ has nonzero first component, then the analysis and algorithm of Golub and Van Loan cannot be used, and the TLS solution (2) does not exist. Van Huffel and Vandewalle pointed out [2, § 3.4, p. 71] that in this case some directions in the column space of A are not at all correlated with the observation vector b; and in a regression sense, these directions are of no value in "predicting the response" b. To handle this, Van Huffel and Vandewalle define the "nongeneric" TLS problem and solution [2, Defn. 3.2, p. 68] by adding an additional restriction that $[g, E]$ be orthogonal to some right singular vectors of $[b, A]$ with zero first components. In [2, Thm. 3.12, p. 72] it is shown how to obtain such a solution, and the comments following that indicate how to compute the minimum 2-norm solution. The analysis of Van Huffel and Vandewalle is accurate and it covers all possible cases. It is, however, very complicated. Moreover, because it is based on comparison of the SVDs of both $[b, A]$ and A, it identifies the different cases *only after* the SVD decompositions are computed, which has negative effects on the cost and possibly also on the numerical properties of existing algorithms.

We offer a completely different approach to both analysis and computation of the TLS problem based on the concept of a *core problem.*

Suppose, for a moment, that some $[\hat{b}, \hat{A}]$ has the form

$$\left[\begin{array}{c|c} \hat{b} & \hat{A} \end{array}\right] = \left[\begin{array}{c||c|c} b_1 & A_{11} & 0 \\ \hline 0 & 0 & A_{22} \end{array}\right]. \tag{3}$$

Then the approximation problem $\hat{A}\hat{x} \approx \hat{b}$ can be viewed as two *independent* approximation problems

$$A_{11}\, x_1 \approx b_1, \quad A_{22}\, x_2 \approx 0, \quad \hat{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{4}$$

Clearly, A_{22} has no effect on "predicting" b_1, the problem $A_{22}x_2 \approx 0$ has the meaningful solution $x_2 = 0$ and only $A_{11}x_1 \approx b_1$ need be solved. Since the Frobenius norm is unitarily invariant, the previous considerations fully apply (with a proper change of variables) to any $[b, A]$ which can be orthogonally transformed, as in

$$\left[\begin{array}{c||c} \hat{b} & \hat{A} \end{array}\right] = P^T \left[\begin{array}{c||c} b & AQ \end{array}\right], \quad P^{-1} = P^T, \quad Q^{-1} = Q^T, \tag{5}$$

to the form (3).

Paige explained in the lecture corresponding to [3] that for *any* $[b, A]$ the transformation (5) leading to (3) is realized by the Golub and Kahan orthogonal bidiagonalization of $[b, A]$. He suggested, but did not prove, that this gives minimally dimensioned $[b_{11}, A_{11}]$ and maximally dimensioned A_{22} (which may possibly be nonexistent), see [4], [5]. We call minimally dimensioned $[b_{11}, A_{11}]$ a core problem.

Definition 1 *We say $A_{11}x_1 \approx b_1$ is a core problem in $Ax \approx b$ if $[b_1, A_{11}]$ is minimally dimensioned (or A_{22} is maximally dimensioned) subject to (5), (3).*

Summarizing, the Golub and Kahan orthogonal bidiagonalization reveals *for any* $[b, A]$ the hidden structure (5), (3). In our contribution we prove that it indeed gives a core problem for any $[b, A]$. In other words, we show that any unwanted and redundant information which is not useful for finding the solution of (1) is in this way removed to A_{22}, while $[b_1, A_{11}]$ contains only the information which is necessary for the solution process. We prove that the solution $x = Q \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$ constructed from the solution of the core problem is theoretically identical to the minimum 2-norm solution of all formulations of TLS in [2].

The suggested approach is simple and effective. It clearly reveals the hidden structure of useful, irrelevant and redundant information contained in the data $[b, A]$ *before* any SVD is computed. Although many implementation details still need to be worked out, the simplicity of the proposed solution process suggests that it might also have some numerical advantages over the algorithms described in [2]. Furthermore, the concept of core problems is fundamental, and might also be useful in other areas such as handling ill-posed problems.

Although in this abstract we restricted ourselves to the TLS problem (2), our results revise our understanding of both the theory and computations in all forms of linear least squares problems with a single right hand side. The full text of the paper [6] will be submitted elsewhere.

References

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

[2] S. Van Huffel and J. Vandewalle, *The total least squares problem: computational aspects and analysis*, SIAM Publications, Philadelphia PA, 1991.

[3] C.C. Paige and Z Strakoš, *Bidiagonalization and Approximate Linear Systems* (Abstract), XV. Householder Symposium, Peebles, 131-133, 2002.

[4] C.C. Paige and Z Strakoš, *Scaled total least squares fundamentals*, Numerische Mathematik, 91:117-146, 2002.

[5] C.C. Paige and Z Strakoš, *Unifying least squares, total least squares and data least squares.* In: "Total Least Squares and Error-in-Variables Modeling", S. Van Huffel and P. Lemmerling (editors), pp. 25-34, Kluwer Academic Publishers, Dordrecht, 2002.

[6] C.C. Paige and Z Strakoš, *Core problems in $Ax \approx b$* (in preparation), 2004.

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