

CORE PROBLEM THEORY  
IN LINEAR APPROXIMATION PROBLEMS

Zdeněk Strakoš

*Institute of Computer Science, Academy of Sciences of the Czech Republic,  
Pod Vodárenskou věží 2, 182 07 Prague 8, Czech Republic,  
and  
Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Prague 8  
email: strakos@cs.cas.cz.*

A joint work with Chris Paige, McGill University, Montreal.

**Abstract**

Consider determining  $x$  from the real linear system

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

We consider the TLS problem for (1)

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

Golub and Van Loan analyzed this problem in 1980. In a nutshell, the constraint in (2) is equivalent to

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

This suggests that the TLS solution is determined by the smallest perturbation of  $[b, A]$  which makes it rank deficient. 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.

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 in 1991 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 in 1991 the “nongeneric”

TLS problem and solution by adding an additional restriction that  $[g, E]$  be orthogonal to some right singular vectors of  $[b, A]$  with zero first components. The analysis of Van Huffel and Vandewalle is accurate and it covers all possible cases. However, since 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.

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

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

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

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

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

$$[\hat{b} \parallel \hat{A}] = P^T [b \parallel AQ], \quad P^{-1} = P^T, \quad Q^{-1} = Q^T, \quad (5)$$

to the form (3).

For *any*  $[b, A]$  the transformation (5) leading to (3) is realized by the Golub and Kahan orthogonal bidiagonalization of  $[b, A]$ . This gives minimally dimensioned  $[b_{11}, A_{11}]$  and maximally dimensioned  $A_{22}$  (which may possibly be non-existent). We call minimally dimensioned  $[b_{11}, A_{11}]$  a *core problem*.

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 previously used formulations of TLS.