Golub-Kahan bidiagonalization, Lanczos tridiagonalization and decomposition of data in linear approximation problems.

Zdeněk Strakoš

Institute of Computer Science,
Academy of Sciences of the Czech Republic, Prague

strakos@cs.cas.cz

This talk will present results achieved in collaboration with Christopher C. Paige, McGill University, Montreal, Canada, and Iveta Hnětynková, Charles University, Prague.

Let A be a nonzero n by k real matrix, and b be a nonzero real n-vector. Consider estimating x from the linear approximation problem

$$Ax \approx b\,, \tag{1}$$

where the uninteresting case is for clarity of exposition excluded by the natural assumption $b \not\perp \mathcal{R}(\mathcal{A})$, that is $A^T b \neq 0$. Here we do not primarily deal with A square nonsingular and solving linear algebraic equations. We allow A rectangular of an arbitrary nonzero rank, and assume that the data A, b contain redundant and/or irrelevant information, and are possibly also corrupted by noise.

In a sequence of papers [1, 2, 3] it was proposed to orthogonally transform the the original data b, A into the form

$$P^T \left[\ b \;\|\; AQ\ \right] = \left[\begin{array}{c||c|c} b_1 & A_{11} & 0 \\ \hline 0 & 0 & A_{22} \end{array}\right], \tag{2}$$

where $P^{-1} = P^T,\ Q^{-1} = Q^T, b_1 = \beta_1 e_1$, and A_{11} is a lower bidiagonal matrix with *nonzero bidiagonal elements*. The matrix A_{11} is either square, when (1) is compatible, or rectangular, when (1) is incompatible. The matrix A_{22}, and the corresponding block row and column in (2), can be nonexistent. The original problem is in this way decomposed into the approximation problem

$$A_{11} x_1 \approx b_1\,, \tag{3}$$

and the remaining part $A_{22} x_2 \approx 0$. It was proposed to find x_1 from (3), set $x_2 = 0$, and substitute for the solution of (1)

$$x \equiv Q \left[\begin{array}{c} x_1 \\ 0 \end{array}\right]. \tag{4}$$

The (partial) upper bidiagonalization of $[b, A]$ described above has remarkable properties, see [3, Theorems 2.2, 3.2 and 3.3].

- First, the lower bidiagonal matrix A_{11} with nonzero bidiagonal elements has full column rank and its singular values are simple. Consequently, any zero singular values or repeats that A has must appear in A_{22}.
- Second, A_{11} has minimal dimensions, and A_{22} has maximal dimensions, over all orthogonal transformations giving the block structure in (2), without any additional assumptions on the structure of A_{11} and b_1.

- Finally, all components of $b_1 = \beta_1 e_1$ in the left singular vector subspaces of A_{11}, that is, the first elements of all left singular vectors of A_{11} (multiplied by $\beta_1 \neq 0$), are nonzero.

Alternative proofs based on the relationship between the Golub-Kahan bidiagonalization, the symmetric Lanczos tridiagonalization and properties of Jacobi matrices were given in [6].

In the approach represented by (1)–(4), the data b, A are fundamentally decomposed. The necessary and sufficient information for solving the problem (1) is given by b_1, A_{11}. All irrelevant and repeated information is filtered out to A_{22}. The problem (3) is therefore called a *core problem* within (1).

The core problem formulation can be used to solve least squares, scaled total least squares and data least squares problems. The core problem solutions are identical to the minimal 2-norm solutions of all formulations of the (scaled) total least squares problem with the single right hand side [4]. It gives the minimum norm solution determined by the algorithm of Golub and Van Loan [5], [4, Theorem 3.7, p. 58], if it exists. If such a solution does not exist, then the core problem approach gives the nongeneric minimum norm (scaled) total least squares solution described by Van Huffel and Vandewalle [4, Theorem 3.12, p. 72]. In this way, one simple and efficient approach can be applied to different classes of problems.

In our contribution we will review the theory, mention recent applications of the core problem formulation, and outline the status of investigation of several open questions.

This work has been supported by the National Program of Research "Information Society" under project 1ET400300415, and by the Institutional Research Plan AVOZ10300504.

References

[1] C. C. Paige and Z. Strakoš, Scaled total least squares fundamentals. *Numer. Math.*, 91:117–146, (2002).

[2] C. C. Paige and Z. Strakoš, Unifying least squares, total least squares and data least squares, in "Total Least Squares and Errors-in-Variables Modeling", S. van Huffel and P. Lemmerling, editors, Kluwer Academic Publishers, Dordrecht, pp. 25–34, (2002).

[3] C. C. Paige and Z. Strakoš, Core problems in linear algebraic systems. *SIAM. J. Matrix Anal. Appl.*, 27:261–276, (2006).

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

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

[6] I. Hnětynková and Z. Strakoš, Lanczos tridiagonalization and core problems. *Linear Algebra Appl.*, 421:243-251, (2007).