Referee report on the PhD Thesis

The Total Least Squares Problem and Reduction of Data in $AX \approx B$

written by Ing. martin Plešinger

Faculty of Mechatronics and Interdisciplinary Engineering Studies,
Technical University of Liberec, Liberec
&
Institute of Computer Science,
Academy of Sciences of the Czech Republic, Prague

Summary of the thesis

Linear approximation problems of the form

$$AX \approx B, \qquad A \in \mathbb{R}^{m\times n}, \quad X \in \mathbb{R}^{n\times d}, \quad B \in \mathbb{R}^{m\times d},$$

where X is unknown, arise in many applications ranging from for instance medical image deblurring to astronomical observations. The classical *least squares* problem with multiple right-hand side solves

$$\min_{X,G} \|G\|_F \qquad \text{subject to} \qquad AX = B + G,$$

where $\|\cdot\|_F$ denotes the Frobenius norm and it allows for correction of only the right-hand side. The *total least squares* (TLS) problem takes also into account possible measurement errors in the data matrix A; the TLS problem is defined as the minimization problem

$$\min_{X,E,G} \|[G|E]\|_F \qquad \text{subject to} \qquad (A+E)X = B + G.$$

A theory for solvability of TLS problems with a single right-hand side ($d = 1$) has been developed since the seventies and reached a considerable degree of maturity with the so-called *core problem* concept introduced by Paige and Strakoš. The core problem within a TLS problem is a subproblem of minimal dimension that contains all necessary and sufficient information to solve the problem. It allows transparent classification of the given type of TLS problem and, in addition, it implies a computationally inexpensive way to separate redundant data. For the multiple right-hand side case some analysis of the types of TLS problems that can arise has been done, in particular in the book of Van Huffel and Vandewalle. Algorithms to compute generalized core problems for multiple right-hand sides have been proposed by Björck and Sima. However, a thorough theoretical foundation for core problems in the multiple right-hand side case, reaching from a proper definition to the question whether they can play a similarly fundamental role as in the single right-hand side case, has not been laid so far.

The main subject of this PhD thesis is analysis of total least squares problems in the multiple right-hand side case with emphasis on possible exploitation of a generalized core problem concept. The thesis starts with a concise introductory overview of existing theory for the single right-hand side case. Results on existence and uniqueness of a TLS solution are presented and the core problem concept is explained. The first main part of the thesis (Part II) is concerned with classification of the types of TLS problems that can arise in the multiple right-hand side case. Based on the work of Van Huffel and Vandewalle dating back from the beginning of the nineties, a comprehensive and logical overview of the individual types is given. Here it must be mentioned that the classification is significantly more complicated than in the single right-hand side case. Thanks to an unusual amount of clarifying figures and tables and text passages recalling relevant previous information, the reader gets hardly lost in the description of the many situations that may arise. The presented classification is original and it extends the classical work of Van Huffel and Vandewalle considerably.

The probably most important part of the thesis (Part III) focuses on a core problem concept for the multiple right-hand side case. In analogy with the single right-hand side case, two forms of a generalized core problem are described. In Chapter 4 it is shown how a core problem can be computed with the help of the SVD's of the data matrix A and of parts of the right-hand sides B. This is a rather straightforward extension of the so-called *SVD-form of the core problem* with a single right-hand side. The extension of the so-called *banded form of the core problem* is less trivial and forms the subject of Chapter 5. Here the core problem is obtained via a band generalization of Golub-Kahan's bidiagonalization algorithm which was proposed by Bj örck and Sima. After a clear description of the band algorithm, the remaining of the chapter contains a theory that is used to prove that this band form shares some important properties with the SVD-form of the core problem. This to my knowledge new theory introduces *wedge-shaped* matrices, matrices who's bandwidth decay gradually when moving down the main diagonal. Wedge-shaped matrices can be seen as generalizations of the class of tridiagonal matrices that is called Jacobi matrices. It is shown that some well-known spectral properties of Jacobi matrices can be generalized to wedge-shaped matrices and these properties are neatly used to prove the desired results. As before, the new results are presented in a clear way and it is a pleasure to read this very elegant theoretical part of the thesis.

The next chapter starts with showing that the common properties of banded and SVD-form (proved before) imply minimal dimensions of the considered generalized core problems. This leads to a precise and justified definition of a core problem in the multiple right-hand side case. The rest of the chapter investigates the newly defined core problem with respect to the classification of TLS problems introduced earlier and with respect to its relation to the core problem for the single right-hand side case. In this context, the investigation reveals two interesting and unexpected aspects of the multiple right-hand side case. Firstly, a core problem may be decomposable in sub-core problems and, secondly, a core problem needs not have, in general, a TLS solution. It is shown on examples how decomposable core problems can arise and it is explained why they do not arise with a single right-hand side. Also examples of core problems without TLS solution are presented.

In the last part of the thesis some results of experiments are reported. One experiment demonstrates possible numerical difficulties with detecting the core problem; the second experiment shows how the algorithm for reduction to the banded form of the core problem can be used to solve an ill-posed problem with a right-hand side disturbed by noise. This is a new hybrid strategy. Both experiments address the single right-hand side case. The last chapter formulates some conclusions and future work that follow from the results of the previous chapters.

Evaluation

The present thesis is convincing for several reasons. First of all it contains new results, mainly in Chapters 5 and 6, which show the ability of the author to solve highly abstract and theoretical problems. In Chapter 5 the newly developed and beautiful theory can be successfully applied to formulate a consistent definition of a core problem. The outcome of Chapter 6 poses perhaps more questions than it answers, but I find this chapter an even more convincing achievement than the previous one. It is sometimes more difficult to give a thorough explanation of why something is *not* behaving in the expected way than of why things work fine. In short, it must be concluded that the core problem extended to the multiple right-hand side case lacks some of the important properties it has in the single right-hand side case. The author is aware of the somehow problematic nature of the new core problem and formulates in his conclusions a conjecture that might open the door to deeper understanding. Still the results of Chapter 6 seem very noteworthy to me, not in the last place for the practical user who will try to reduce his data by core problem detection anyway. In the context of practical usage, it would perhaps be interesting to explain why the thesis refers constantly to computation of TLS solutions with the so-called *classical TLS algorithm*, Algorithm 3.1 on page 54. As follows from the text, the algorithm has the shortcoming that it does not compute, for problems of the class $\mathcal{F}_2$, a TLS solution but a non-generic solution. Moreover, Algorithm 3.1 produces different solutions for a decomposable core problem depending on whether the core is treated decomposed or not. Is Algorithm 3.1 the only algorithm that can be used in practice ? Is it possible to comment on whether a modification of the algorithm, based on (3.37) rather than on (3.24), is feasible ? What problems can arise when adapting the algorithm to circumvent non-unique solutions for decomposable core problems ?

An appealing quality of the thesis is its impressive conciseness and completeness, combined with a truly luxurious presentation. Any questions occurring during a first reading seem to be immediately answered in the text, either by words or through clarifying figures. The English used is in general very good, with the exception of some constantly returning bohemisms (denote **by**, to have **a** unique solution) which, however, do not prevent from understanding the meaning of the text. Also the number of typos in the text is remarkably low.

In comparison with the previous chapters, Chapters 7 and 8 make a less exemplary impression, not only because they are shorter. It is a bit of a pity that Chapter 7 treats a TLS problem with single right-hand side where the whole thesis treats the multiple right-hand side case. The hybrid method in Chapter 8 may be interesting but its relation with TLS problems is quite remote.

Conclusion

The thesis addresses a very recent and rather fundamental question relevant when solving TLS problems with multiple right-hand side. The results obtained while attempting (and partially succeeding) to give the answer are treated in a very systematic and scientific way and form a solid basis for possible future publications. The results demonstrate clearly that the author of the thesis is qualified for independent scientific work. I strongly recommend awarding the title Ph.D. to Ing. Martin Plešinger.

Prague, 30.05.2008 Jurjen Duintjer Tebbens