Authors' response to the referees' reports

Paper title: Left vs right representations for solving weighted low rank approximation problems
Authors: Ivan Markovsky and Sabine Van Huffel

August 15, 2006

We thank the referees for their relevant and useful comments.

In this document, we quote in **bold face** statements from the reports. Our replies follow in ordinary print.

Answer to referee #1

It is mentioned that this [I/O representation] is not a generally valid parameterization but the potential problems in applications where this is (nearly) invalid are not discussed. What are the numerical properties in cases when the square blocks assumed nonsingular are singular or at least ill-conditioned? Can you detect such cases and how do you resolve the problem?

Lack of existence of an I/O representation in solving TLS problems corresponds to what are called nongeneric TLS problems [VV88, VV91, PS06]. In the TLS literature, however, the focus is on solving approximately a linear system of equations $AX \approx B$, so that the I/O representation is a part of the problem formulation and is a priori fixed. As a consequence, the methods for solving nongeneric TLS problems *reformulate* the TLS problem in such a way that the reformulated problem becomes solvable. The solution of the reformulated problem is then defined to be a solution of the nongeneric TLS problem.

Our objective is to solve a low-rank approximation problem that (at least in the unweighted case $W = I$) always has a solution. Therefore, it is no longer justified to change the problem formulation. In our case, a nongeneric TLS problem is caused by a poor choice of an I/O representation that (contrary to the situation treated in the TLS literature) is at our disposal to revise by selecting in a different way the columns of the data matrix that form A. Therefore, we can avoid nongeneric TLS problems by choosing a different I/O representation that has a solution (or, better, leads to a well conditioned computational problem). We believe that the best approach to avoid nongeneric TLS problems in doing low-rank approximation is to choose the right parameterization.

Section 4 is revised accordingly and Note 1 is replaced by a note explaining the link to nongeneric total least squares problems and the core problem of [PS06]. Now the A and B matrices in the $AX_{\mathrm{r}} = B$ and $X_\ell A = B$ representations are defined as submatrices of the matrix $\Pi_\ell D \Pi_{\mathrm{r}}$, where $\Pi_\ell \in \mathbb{R}^{m\times m}$ and $\Pi_{\mathrm{r}} \in \mathbb{R}^{n\times n}$ are permutation matrices. The permutations Π_ℓ and Π_{r} can *always* be chosen, so that the parameters X_ℓ and X_{r} exist.

Of course, the problem now shifts to how to select the permutations Π_ℓ and Π_{r} in an optimal (from the point of view of well conditioned computations) way. This is the famous subset selection problem. Numerically it is not straightforward to implement and needs a test for singularity of submatrices. The same issue occurs in the core problem formulation. A topic of current research is to find methods that can reliably detect nongeneric total least squares problems.

WLRA problems with structured weighting matrices also appears in factor analysis; see, e.g., ... dealing with a diagonal weighting.

The reference is relevant and was included in the paper. As mentioned in the end of Section 6, page 491, however, the proposed cris-cross method does not necessarily converge to a minimum point of the WLRA cost function.

Section 6 does not appear to be well worked out and lack motivations.

We agree. The numerical implementation deserves special attention and will be treated elsewhere.

Approximating the general weighting matrix by a Kronecker product is sometimes used in applications. The authors do not seem to be aware of the fact that this problem can be converted to a rank-one approximation problem ...

Indeed, we were not aware of this reference and thank the referee for pointing it to us. The paper was revised accordingly. (See page 2 of the revised version.)

The authors propose an approach approximating only the diagonal part of the full weighting matrix. The motivation behind this approximation is not given and unclear, in particular in view of the two references given above.

For diagonal weight matrix W, our approximation coincides with the approximation described in [VP93]. The procedure in [VP93], however, applies for the general case, so in the revised version of the paper we refer to [VP93].

In general, it is also difficult to argue that using an approximation of the general weighting matrix is better than using, e.g., $W = I$, the identity weighting. The sensitivity of the solution wrt to certain errors in the weighting may be large.

This is true. However, simulation results show that the heuristic “the better the weight matrix W is approximated, the better (in terms of the original cost function) suboptimal solution is computed” is true. Thus we propose to use the best possible approximation of W, namely the one computed by the method of [VP93].

The formulations of the first two minimization problem on page 5 are bit unclear. Why do you include d-hat in the minimization? A similar remark applies on page 8.

$\hat{d}$ is an optimization variable. Of course, in (WLRA_P) and (WLRA_L) it can be eliminated trivially by substitution, but we believe that the problem formulation is clearer with $\hat{d}$ present, because it suggests that the corresponding quantity is an approximation of d.

Also on page 8, there seems to be some misplaced hats.

Corrected.

Do you need the last "reshape" in your matlab code on page 9?

It is not needed and is removed.

Answer to referee #2

The main step of the proposed method consists in the replacement of the difficult nonconvex original problem (WLRA) by one of four substantially simpler problems (WLRA"). This should be stated in form of a theorem with appropriate assumptions ensuring the existence of (I/O) reprsentations.

The statement about the equivalence of (WLRA) and (WLRA') is implicitly contained in Theorems 3 and 4. In the revised version of the paper, we have reformulated Theorems 3 and 4, so that the statement is made explicit.

We prefer to state the conditions ensuring the existence of an I/O representation in Section 4, Theorem 6, where the notion of an I/O representation is introduced.

The user would also expect some hints what to do in the "nongeneric case".

This comment is pointed out also by the first referee. Please refer to the answer to the first question of referee 1.

From this point of view Note 1 should be suitably reformulated.

Note 1 is deleted. Please refer to the answer to the first question of referee 1.

The conjecture, concerning the choice of the "correct" representation (supported by the simulation example) seems me too daring.

Two major factors determining the difficulty of unconstrained optimization problems are:

1. the type of the cost function (nonconvex vs. convex, least squares vs. general nonlinear, etc), and
2. the number of optimization variables.

Irrespective of the parameterization, the WTLS problem is nonconvex and there is no reason to believe that the left or right I/O parameterization leads to an "easier" or more "difficult" nonconvex cost function. The number of parameters in those two representations, however, is different and this is the basis for the conjecture: from two optimization problems with the same type of cost function, we choose the one that has fewer optimization variables.

In the equality constraint of (1WTLS) it should probably be X1 without hat.

Yes, we have corrected this typo.

On page 3, line 7 it should probably be "advantageous".

Corrected.

References

[PS06] C. Paige and Z. Strakos. Core problems in linear algebraic systems. *SIAM J. Matrix Anal. Appl.*, 2006.

[VP93] C. Van Loan and N. Pitsianis. Approximation with Kronecker products. In M. Moonen and G. Golub, editors, *Linear Algebra for Large Scale and Real Time Applications*, pages 293–314. Kluwer Publications, 1993.

[VV88] S. Van Huffel and J. Vandewalle. Analysis and solution of the nongeneric total least squares problem. *SIAM J. Matrix Anal. Appl.*, 9:360–372, 1988.

[VV91] S. Van Huffel and J. Vandewalle. *The total least squares problem: Computational aspects and analysis*. SIAM, Philadelphia, 1991.