Core problems in $AX \approx B$

Diana Sima and Sabine Van Huffel

February 10, 2006

Abstract

In this paper, we generalize the core problem concept [Paige and Strakoš, SIMAX, 2006] to multiple right-hand sides linear problems $AX \approx B$. In the single right-hand side case, a core problem $A_{11}x_1 \approx b_1$ may be extracted from $Ax \approx b$. The core problem has the size $(p+1) \times p$, where p is the number of distinct singular values of the matrix A corresponding to left singular subspaces that are not orthogonal to the right-hand side vector; the core matrix A_{11} has as singular values a subset of the distinct nonzero singular values of A. In the multiple right-hand sides case (with d right-hand sides), a core problem of size $(p+d) \times p$ can be extracted, and it exhibits similar properties. In particular, the core subproblem does not suffer from nonuniqueness or nongenericity issues when it is solved in the total least squares sense, even if total least squares problem for the original $AX \approx B$ does.

For the single right-hand side case, practical computation of a core problem can be done with (direct or Lanczos) bidiagonalization; for multiple right-hand sides, we give a detailed *band Lanczos* diagonalization method.

1 Introduction

1.1 The scaled total least squares formulation

In [10, 11], the notion and properties of core problems in linear algebraic systems are introduced and thoroughly analyzed. The context is solving overdetermined nearly compatible linear systems of equations, written in matrix notation as $Ax \approx b$, where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ are given ($m > n$), and x is an unknown vector. The classical methods of least squares, total least squares and data least squares are treated there in an unified manner, under the more general *scaled total least squares (ScTLS)* formulation. The ScTLS solution is the optimal x that solves the minimization

$$\min_{x,\Delta A,\Delta b} \| \begin{bmatrix} \Delta b & \Delta A \end{bmatrix} \|_F^2, \qquad \text{subject to} \quad (A + \Delta A)x\gamma = \gamma b + \Delta b. \tag{1}$$

Here, γ is a positive scaling parameter that allows the corrections on the matrix A to have a different magnitude compared to the corrections on b. The extreme cases when γ goes to zero or to infinity correspond to the least squares and data least squares problems (*i.e.* corrections *only* on b or *only* on A, respectively), while the particular choice $\gamma = 1$ gives the total least squares problem [15].

Unique and minimum norm solution

For finite and nonzero γ, problem (1) becomes the nearest approximation of a matrix with a lower rank matrix, and can be solved using the SVD of the matrix $\begin{bmatrix} \gamma b & A \end{bmatrix}$. In the usual

(generic) case, only the singular vector corresponding to the smallest nonzero singular value is needed, if this last singular value is unique. If the smallest nonzero singular value of $\begin{bmatrix}\gamma b & A\end{bmatrix}$ is multiple, then the minimization (1) does not have unique solution either; then, a convenient and conventional choice is to pick from the corresponding singular subspace a vector that will give the *minimum norm* solution for x [15].

Nongeneric and close-to-nongeneric cases

Another cumbersome case happens when problem (1) does not have a finite optimal solution at all. Examples for this situation can be found in [11] and [15, §3.4]. The explanation is that the corrections ΔA and Δb can get infinitesimally small in magnitude, while the corresponding solution x has some elements that go to infinity. Such a situation is possible when, for instance, the right-hand side vector b is orthogonal to the subspace corresponding to the smallest nonzero singular value of A. This pathological case is a nongeneric case (*i.e.*, it appears in practical data with zero probability). However, the interest in these cases is not only theoretical; in realistic ill-posed problems we encounter *nearly nongeneric* problems, which need specialized methods. The understanding of the nongeneric case provides the background for solving close-to-nongeneric practical problems. (This topic is discussed in [13].)

1.2 Core problems in $Ax \approx b$

The core problem advertised in [11] avoids the nonuniqueness and nongenericity issues, by transforming the original data in A and b to reduced versions A_{11} and b_1, which are smaller in size, but essential for the approximation problem at hand. All redundant and not useful information that was present in the data A and b is eliminated by reducing to A_{11} and b_1.

The suggestion of Paige & Strakoš [11] is to find orthogonal P and Q such that

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

and $\begin{bmatrix} b_1 & A_{11} \end{bmatrix}$ has minimal dimension; then, to solve $A_{11}x_1 \approx b_1$ with the appropriate ScTLS algorithm and to set the final solution as

$$x = Q \begin{bmatrix} x_1 \\ 0 \end{bmatrix},$$

where the zero part comes from setting to zero the solution of the system $A_{22}x_2 \approx 0$.

SVD form of the core decomposition

A core problem $A_{11}x_1 \approx b_1$ can be obtained from the SVD of $A = U'\Sigma'V'^\top$. Indeed, using Householder rotations and some permutations, it is possible to transform

$$U'^\top \begin{bmatrix} b & AV' \end{bmatrix} \quad \text{into} \quad \left[\begin{array}{c|c|c} c & \Sigma'_1 & 0 \\ \delta & 0 & 0 \\ \hline 0 & 0 & \Sigma'_2 \\ 0 & 0 & 0 \end{array}\right] \begin{array}{l} p \\ 1 \\ n-p \\ m-n-1 \end{array}$$

where all elements of c are nonzero, Σ'_1 contains only distinct nonzero singular values (say p), while Σ'_2 contains the other $n-p$ singular values, including possible multiples and all the zero singular values, if any. The scalar δ is zero when $Ax = b$ is compatible and nonzero otherwise.

Bidiagonal form of the core decomposition

For computational efficiency reasons, a bidiagonal reduction can be used instead of the SVD, *i.e.*,

$$P_1^\top \left[\begin{array}{cc} b & AQ_1 \end{array}\right] = \left[\begin{array}{cc} b_1 & A_{11} \end{array}\right] = \left[\begin{array}{c|cccc} \beta_1 & \alpha_1 & & & \\ & \beta_2 & \alpha_2 & & \\ & & \ddots & \ddots & \\ & & & \beta_p & \alpha_p \\ & & & & (\beta_{p+1}) \end{array}\right]$$

where $\alpha_i\beta_i \neq 0$ and β_{p+1} is zero only for compatible systems. P_1 and Q_1 are the blocks of P and Q in (2) containing the first $p+1$ or p columns, respectively, in these orthogonal matrices.

We can use direct bidiagonalization (with Householder rotations) for small to moderate problem dimensions, or the Lanczos bidiagonalization, suitable for large scale problems [7]. Note that A_{22} need not be bidiagonalized. Thus, the bidiagonalization algorithm is run until a (numerically) zero α_i or β_i is reached.

Paige and Strakoš [11] proved that the bidiagonal reduction leads to a minimally dimensioned A_{11}, where A_{11} has only distinct and nonzero singular values. Moreover, solving the reduced bidiagonal problem $A_{11}x_1 \approx b_1$ with the ScTLS algorithm and transforming back to the full solution $x = Q_1x_1$ actually gives the ScTLS solution (or the minimum norm ScTLS solution in problems with nonunique solution, or the typical solution to a constrained ScTLS problem in nongeneric cases) of $Ax \approx b$, thus, the same solution that is proposed in [15] can be obtained in a possibly more efficient manner.

2 Extension to multiple right-hand sides and to a nullspace formulation

2.1 Scaled total least squares for $AX \approx B$

The ScTLS problem can be extended to multiple right-hand sides problem $AX \approx B$ (where A is $m \times n$, X is $n \times d$, B is $m \times d$, and, for convenience, $m > n + d$), and can be used in cases when one knows that the corrections on B and on A should be scaled at a given ratio with respect to each other:

$$\min_{x,\Delta A,\Delta B} \| \left[\begin{array}{cc} \Delta B & \Delta A \end{array}\right] \|_F^2, \qquad \text{subject to} \quad (A + \Delta A)x\gamma = \gamma B + \Delta B. \tag{3}$$

This type of problems can be solved using direct (linear algebraic) methods, such as the SVD. One of the most general TLS-related formulations that can still be solved with SVD-type computations is the *generalized TLS* [14]. It allows the corrections within each row of $\begin{bmatrix} B & A \end{bmatrix}$ to be differently scaled, but requires that the error on each row is independent from and identically distributed as the error on the other rows; this corresponds to a general (positive semidefinite) covariance matrix – known up to a constant factor – for the errors in each row of $\begin{bmatrix} B & A \end{bmatrix}$. Other more general scalings of the corrections require optimization-based methods instead of numerical linear algebra techniques. See, for instance, the methods proposed for element-wise weighted total least squares [9, Chapter 3].

2.2 The nullspace approximation problem $CY \approx 0$

We consider also the (unscaled) approximation problem in nullspace formulation $CY \approx 0$, with $C \in \mathbb{R}^{m\times(n+d)}$ and $Y \in \mathbb{R}^{(n+d)\times d}$. Note that $AX \approx B$ is a special case of $CY \approx 0$, with $C = \begin{bmatrix} B & A \end{bmatrix}$ and the constraint that Y has as leading $d \times d$ block a nonsingular matrix. The nullspace formulation $CY \approx 0$ is more general, since this constraint is not imposed, but it is necessary to impose a nontriviality condition on Y, such as the constraint that Y is full column rank. As a way of estimating Y, we use the TLS criterion of minimizing $\|\Delta C\|_F^2$ subject to the constraint that the corrected system is consistent, $(C + \Delta C)\, Y = 0$, and an extra nontriviality constraint on Y; for the latter, we set $Y^\top Y = I_d$, which does not restrict the generality.

Obviously, the SVD is a possible key to the nullspace formulation problem, since the problem amounts to finding the nearest low rank approximation of C (with a rank reduction of at least d) and choosing a basis for the nullspace of $C + \Delta C$ in order to compute Y.

Using the SVD of $\begin{bmatrix} \gamma B & A \end{bmatrix}$, or the SVD of C, respectively, the optimal corrected matrices $\begin{bmatrix} \gamma B + \Delta B & A + \Delta A \end{bmatrix}$, and $C + \Delta C$, respectively, are obtained by setting to 0 the smallest d singular values in these decompositions.

2.3 Unique and minimum norm solution

For both the multiple right-hand sides and the nullspace formulations, an intrinsic source of multiple optimal solutions appears whenever the dth smallest singular value σ_n is *multiple* (*i.e.*, $\sigma_n = \sigma_{n+1}$). In this case, the optimal corrections, as well as the optimal solution X (or Y) are nonunique.

To distinguish between all X solutions that give the same minimum residual, the *minimum Frobenius norm* solution for X is typically chosen in the TLS literature [15]. However, the nice closed-form expression for the minimum norm TLS solution $\widehat{X} = (A^\top A - \sigma_{n+1}^2 I)^{-1} A^\top B$ only holds true when the last d singular values of $\begin{bmatrix} B & A \end{bmatrix}$ coincide: $\sigma_{n+1} = \cdots = \sigma_{n+d}$ (see the discussion in [15, §3.3.2]).

Note also that Y is in any case nonunique, since every rotated matrix YS, with S orthogonal, is also an optimal solution for any optimal Y.

2.4 Nongeneric and close-to-nongeneric cases

The nongeneric and close-to-nongeneric cases occur for the $AX \approx B$ formulation when the singular subspace corresponding to the dth smallest singular value (σ_{n+1}) is orthogonal (or nearly orthogonal) to the columns of the B matrix.

The nongenericity concept does not exist for the $CY \approx 0$ formulation, since the source of this type of problems is precisely the partitioning of the data matrix C into A and B.

3 Core problem in $AX \approx B$

For the multiple right-hand sides $AX \approx B$ problem, we consider transformations of the form:

$$P^\top \begin{bmatrix} B & AQ \end{bmatrix} = \left[\begin{array}{c||c|c} B_1 & A_{11} & 0 \\ \hline\hline 0 & 0 & A_{22} \end{array}\right], \tag{4}$$

with P and Q orthogonal matrices and $\begin{bmatrix} B_1 & A_{11} \end{bmatrix}$ of minimal dimensions. Then, the problem $AX \approx B$ becomes separable:

$$A_{11}X_1 \approx B_1 \qquad \text{and} \qquad A_{22}X_2 \approx 0,$$

where the original X is recovered as $X = Q \begin{bmatrix} X_1^\top & X_2^\top \end{bmatrix}^\top$. The second system has as reasonable solution the trivial solution $X_2 = 0$. This choice corresponds to a minimum (Frobenius and 2-) norm X solution.

We argue that a minimally dimensioned A_{11} has, in the case when $AX \approx B$ is incompatible, the size $(p+d) \times p$, where p denotes the number of distinct nonzero singular values of A corresponding to singular subspaces on which B is not orthogonal. To this end, we construct a reduction of the form (4) based on the SVD of A.

3.1 SVD form of the core reduction

Let the full SVD of A be $U'\Sigma'{V'}^\top$, where U' is $m \times m$ orthogonal, V' is $n \times n$ orthogonal, and Σ' is $m \times n$ diagonal matrix. Applying ${U'}^\top$ from the left, we have:

$${U'}^\top \begin{bmatrix} B & AV' \end{bmatrix} = \begin{bmatrix} {U'}^\top B & \Sigma' \end{bmatrix} =: \begin{bmatrix} B' & \Sigma' \end{bmatrix}.$$

Next, B' should be modified – using orthogonal transformations from the left – in a controlled way that introduces zeros, such that a separable decomposition as in (4) is obtained.

- The rectangular $m \times n$ matrix Σ' ($m > n$) has a zero block below the square diagonal block. The part of B' corresponding to this zero block can be reduced with a QR factorization: $B_Q \begin{bmatrix} B_R \\ 0 \end{bmatrix}$, where B_Q is an orthogonal matrix and B_R is an upper triangular $d \times d$ matrix. The orthogonal matrix from this QR decomposition can be incorporated into the transformations from the left, by multiplying ${U'}^\top$ with $\begin{bmatrix} I & 0 \\ 0 & B_Q \end{bmatrix}$.

- The first column of B' is reduced (as in the one-dimensional case [11]) to a vector where only one nonzero element remains for each (of the, say, p') distinct singular values in Σ' (including the zero singular value, if A is rank-deficient). This can be accomplished with a series of Householder transformations on B' from the left, which will multiply the already existing transformation on the original data B; these will be applied from the right to A as well, such that the diagonal structure of Σ' is kept. For convenience, a permutation of the rows can be performed, such that all nonzero elements corresponding to distinct nonzero singular values of A are grouped together in the first p' elements of the first column of the new B'.

- For the second column of B' (of the new B', in fact), all but $p'+1$ elements can be zeroed, using a Householder rotation without influencing the first column. The price is that the second column will have a nonzero element in a new position. This position is chosen to correspond to a row of Σ' where a copy of a singular value of A resides (for example, a particular choice is the largest multiple singular value for which a nonzero element of the second column of B' exists).

- The reduction of the other columns of B' such that they have at most $p'+i$ nonzero elements (where i is a column index) continues as long as there are still unprocessed columns in B' and there are still multiple singular values left in Σ', which were not already singled out in previous steps.

- At the end, all completely *zero rows* among the first $p'+d$ rows of B', together with the corresponding rows and columns of Σ', will be permuted to the bottom of the decomposition.

This procedure will compute, finally, a decomposition of the form

$$P^\top \begin{bmatrix} B & AQ \end{bmatrix} = \left[\begin{array}{c||cc} B_T & \Sigma_1' & 0 \\ B_R & 0 & 0 \\ \hline\hline 0 & 0 & \Sigma_2' \end{array}\right] =: \left[\begin{array}{c||c|c} B_1 & A_{11} & 0 \\ \hline\hline 0 & 0 & A_{22} \end{array}\right], \tag{5}$$

where B_T is upper trapezoidal with d columns and, say, p rows, B_R is $d \times d$ upper triangular, Σ_1' contains nonzero singular values of A with multiplicities at most d, and Σ_2' contains the zero singular values and the remaining multiple singular values of A.

The following properties of the decomposition (5) are essential for showing that nonuniqueness and nongenericity issues are avoided when solving the core problem $A_{11}X_1 \approx B_1$.

Proposition 1. *1. In the decomposition (5), A_{11} does not have singular values with multiplicity bigger than d.*

2. The number of rows p of B_T equals the number of singular subspaces of A corresponding to distinct singular values, on which B has nonzero projections.

Proof. The first point is obvious by construction. The second point is clear if we look at the particular SVD of A given in (5): the projection of B onto all singular subspaces corresponding to distinct nonzero singular values appears within B_T in (5). By construction, each row of B_T contains nonzero elements. □

As a corollary, we have that the core problem $A_{11}X_1 \approx B_1$ has unique solution (since nonuniqueness happens when the smallest nonzero singular value has multiplicity larger than the number of right-hand sides). Moreover, the problem $A_{11}X_1 \approx B_1$ is generic, because in nongeneric problems the right-hand side is orthogonal to the singular subspace corresponding to the smallest nonzero singular value.

3.2 Banded approximations

The explicit partitioning in $\begin{bmatrix} B & A \end{bmatrix}$ suggests that a *block version* of the Lanczos algorithm can be an appropriate tool; in such a method, the Lanczos process is done with d starting vectors (the d columns of B) at a time. However, instead of a block-Lanczos, an application of band-Lanczos for the reduction to a core problem in the multiple right-hand sides case was proposed in recent talks by Åke Björck [4, 3]. We resketch the banded decomposition and the band-Lanczos method from there, but then we modify it and make it more rigorous. We discuss afterwards the properties of the obtained decomposition.

A band decomposition of $\begin{bmatrix} B & A \end{bmatrix}$ has the shape

$$P^\top \begin{bmatrix} B & AQ \end{bmatrix} = \left[\begin{array}{c|c} R & L \\ 0 & 0 \end{array}\right], \tag{6}$$

where R is $(\tilde{p}+d) \times d$ upper triangular, and L is $(\tilde{p}+d) \times \tilde{p}$ banded lower triangular, for some $\tilde{p} \le n$, as in:

$$\left[\begin{array}{c|c} R & L \end{array}\right] = \left[\begin{array}{cc|ccccc} * & * & * & & & & \\ & * & * & * & & & \\ & & * & * & * & & \\ & & & * & * & * & \\ & & & & * & * & \\ & & & & & * & \end{array}\right],$$

where the width of the band equals $d+1$. In this decomposition, zero elements might appear at some point on the band of L; it is desirable to take advantage of this in order to identify a core subproblem. Björck suggests how to use *deflation* steps whenever a zero appears on the outer diagonals of the band. In that case, the band can shrink with one, and the procedure can be repeated until the whole band dies out, leading to a core problem.

To be more rigorous, let us first show how the band decomposition can be obtained with a band Lanczos iterative method, in the spirit of [12]. Denote by $u_1, \ldots, u_n$ the first n columns of the matrix P, by $v_1, \ldots, v_n$ the columns of the matrix Q, and by g_k, h_k the $(d+1)$-dimensional vectors[1] that are obtained at the intersection of the kth row and, respectively, column of $\begin{bmatrix} R & L \end{bmatrix}$ with the band of $\begin{bmatrix} R & L \end{bmatrix}$. Note that the vectors g_k, h_k are not independent, since they share the same elements on the band of $\begin{bmatrix} R & L \end{bmatrix}$. This is more obvious when we write $\begin{bmatrix} R & L \end{bmatrix}$ as:

$$\begin{bmatrix} R & L \end{bmatrix} = \begin{bmatrix} g_1 & & & & \\ g_2 & \ddots & & & \\ & & g_n & & \\ & & g_{n+1} & & \\ & & \vdots & & \\ & & g_{n+d} & & \end{bmatrix} = \begin{bmatrix} h_1 & \cdots & h_d & h_{d+1} & & \\ & & & \ddots & & \\ & & & & h_{n+d-1} & \\ & & & & & h_{n+d} \end{bmatrix}.$$

From (6), two d-term recurrence relationships can be written:

$$A^\top u_k = \begin{bmatrix} v_{k-d} & \ldots & v_k \end{bmatrix} g_k, \tag{7}$$

$$A v_k = \begin{bmatrix} u_k & \ldots & u_{k+d} \end{bmatrix} h_{k+d}. \tag{8}$$

The first relation is well-defined for all $k \geq 1$ if we artificially consider some zero vectors $v_0, v_{-1}, \ldots, v_{-d+1}$. An iterative reduction method based on these recursions and on the conditions that P and Q have orthonormal columns can be expressed with the following steps, that occur at iteration k:

a: $w_1 \leftarrow A^\top u_k - \begin{bmatrix} v_{k-d} & \ldots & v_{k-1} \end{bmatrix} g_k(1:d)$
b: $g_k(d+1) \leftarrow \|w_1\|$
c: $v_k \leftarrow w_1 / \|w_1\|$
d: $h_{k+d}(2:d)^\top = \begin{bmatrix} u_{k+1} & \ldots & u_{k+d-1} \end{bmatrix}^\top A v_k$
e: $w_2 \leftarrow A v_k - \begin{bmatrix} u_k & \ldots & u_{k+d-1} \end{bmatrix} h_{k+d}(1:d)$
f: $h_{k+d}(d+1) \leftarrow \|w_2\|$
g: $u_{k+d} \leftarrow w_2 / \|w_2\|$

Note, however, that the steps c: and g: are not well-defined whenever the norm of the auxiliary vector w_1 or, respectively, the norm of the auxiliary vector w_2 is zero. In fact, when this happens, we need *deflation*; it means that a Krylov subsequence for the right singular subspace or, respectively, for the left singular subspace, has reached a converged vector.

[1] For k smaller than $d+1$, h_k has length k. For k larger than n, g_k has length $n-k$, thus smaller than $d+1$. Note that in this paragraph we let $\tilde{p} = n$.

In Algorithm 1, we take exact deflation into account. This algorithm is more similar to the band reductions proposed in [5, 2, 6, 1] than to the one in [12]; it is based on recursions that are a bit modified compared to (7-8):

$$A^\top \begin{bmatrix} u_1 & \dots & u_k \end{bmatrix} = \begin{bmatrix} v_1 & \dots & v_{k-i_c} \end{bmatrix} {L^{(k)}}^\top, \tag{9}$$

$$A \begin{bmatrix} v_1 & \dots & v_{k-i_c} \end{bmatrix} = \begin{bmatrix} u_1 & \dots & u_k \end{bmatrix} L^{(k)} + [\underbrace{0 \ \dots \ 0}_{k-d_c-i_c} \ \widehat{u}_{k+1} \ \dots \ \widehat{u}_{k+d_c}], \tag{10}$$

where $L^{(k)}$ denotes $L(1:k, 1:k-i_c)$, the $k \times (k-i_c)$ banded lower trapezoidal matrix that is obtained up to the kth iteration. The band of $L^{(k)}$ has, at the beginning, a height equal to d, but, during iterations, it may shrink and it may even deviate from the main diagonal. Thus, we introduce an integer d_c to denote the 'current' bandwidth before iteration k, and an integer i_c to denote the number of zero elements between the main diagonal and the band. (See Figure 1). This explains why we do not work with an $L^{(k)}$ of size $k \times k$,

$$L^{(6)} = \begin{bmatrix} * & & & & \\ * & * & & & \\ * & * & * & & \\ * & * & * & * & \\ & * & * & * & * \\ & & 0 & * & * \end{bmatrix}, \quad L^{(7)} = \begin{bmatrix} * & & & & \\ * & * & & & \\ * & * & * & & \\ * & * & * & * & \\ & * & * & * & * \\ & & 0 & * & * \\ & & & 0 & * \end{bmatrix}, \quad \begin{bmatrix} L^{(8)} & 0 \end{bmatrix} = \begin{bmatrix} * & & & & \\ * & * & & & \\ * & * & * & & \\ * & * & * & * & \\ & * & * & * & * \\ & & 0 & * & * \\ & & & 0 & 0 \end{bmatrix}$$

Figure 1: Example of the shape evolution of the banded lower trapezoidal matrix $L^{(k)}$ during iterations. Starting with a bandwidth $d_c = d = 4$, a first zero at $(6, 3)$ shrinks d_c to 3; the zero is carried along to $(7, 4)$ and $(8, 5)$. A second zero at $(8, 8)$ implies that d_c becomes 2, and i_c increases from 0 to 1, marking a completely zero last column.

which would have the last i_c columns equal to zero. The hat vectors $\widehat{u}_{k+1}, \dots, \widehat{u}_{k+d_c}$ are candidate vectors that are used in order to avoid having to compute $u_{k+1}, \dots, u_{k+d}$ before the kth iteration, as it is the case in (8).

Comparing the recurrence relationship (10) with what we have at step $k-1$,

$$A \begin{bmatrix} v_1 & \dots & v_{k-1-i_c} \end{bmatrix} = \begin{bmatrix} u_1 & \dots & u_{k-1} \end{bmatrix} L^{(k-1)} + [\underbrace{0 \ \dots \ 0}_{k-1-d_c-i_c} \ \widehat{u}_k \ \dots \ \widehat{u}_{k-1+d_c}], \tag{11}$$

we can extract the work that needs to be performed to fulfill the kth iteration:

- First, if $\|\widehat{u}_k\| \neq 0$, compute u_k:

$$L_{k,k-d_c-i_c} \leftarrow \|\widehat{u}_k\|, \qquad u_k \leftarrow \widehat{u}_k / L_{k,k-d_c-i_c}. \tag{12}$$

- Then, compute the nonzero elements on the last row of $L^{(k)}$ by multiplying (10) and (11) with $u_k^\top$ from the left and exploiting the orthonormality of $\{u_1, \dots, u_k\}$:

$$L_{k,j-d_c-i_c} \leftarrow u_k^\top \widehat{u}_j, \qquad \text{for } j = k+1, \dots, k+d_c.$$

- With these values plugged in, update $\widehat{u}_{k+1}, \dots, \widehat{u}_{k+d_c-i_c}$ from (10):

$$\widehat{u}_j \leftarrow \widehat{u}_j - u_k L_{k,j-d_c-i_c}, \qquad \text{for } j = k+1, \dots, k+d_c.$$

- Now, if v_{k-i_c} was not previously computed (*e.g.*, when $i_c = 0$), it is possible, from (9), to single out:

$$\widehat{v}_{k-i_c} = A^\top u_k - \sum_{j=k+1}^{k+d_c-i_c-1} v_{j-d_c} L_{k,j-d_c},$$

- If $\|\widehat{v}_{k-i_c}\| \neq 0$, set

$$L_{k,k-i_c} \leftarrow \|\widehat{v}_{k-i_c}\|, \qquad v_{k-i_c} \leftarrow \widehat{v}_{k-i_c}/L_{k,k-i_c}.$$

- Finally, from (10), we initialize $\widehat{u}_{k+d_c}$:

$$\widehat{u}_{k+d_c} \leftarrow Av_{k-i_c} - u_k L_{k,k-i_c}. \tag{13}$$

As announced, Algorithm 1 shows how deflation is used in order to avoid any divisions by zero that can occur in (12) or (13). Each of the two possible kinds of deflation steps decrease the bandwidth d_c by one, and the second one (for (13)) increases i_c by one, as well; this corresponds to cutting a last zero column from the matrix L.

Algorithm 1 Band Lanczos reduction of A with exact deflation

1: $k \leftarrow 0$, $d_c \leftarrow d$, $i_c \leftarrow 0$
2: 2a: compute the QR factorization of B: $B = Q_B R_B$
2b: initialize $\widehat{u}_1, \ldots, \widehat{u}_d$ with the first d columns of the orthogonal matrix Q_B
3: **repeat**
4: $k \leftarrow k + 1$
5: **if** $\|\widehat{u}_k\| = 0$ **then** deflate:
5a: set $L_{j,j-d_c-i_c} = 0$ for all $j > k$
5b: $d_c \leftarrow d_c - 1$
5c: **for** $j = k, \ldots, k + d_c - 1$, set $\widehat{u}_j \leftarrow \widehat{u}_{j+1}$
5d: go to 5:
6: 6a: $L_{k,k-d_c-i_c} \leftarrow \|\widehat{u}_k\|$
6b: $u_k \leftarrow \widehat{u}_k / L_{k,k-d_c}$
6c: **for** $k = k + 1, \ldots, k + d_c - i_c$
$L_{k,j-d_c-i_c} \leftarrow u_k^\top \widehat{u}_j, \qquad \widehat{u}_j \leftarrow \widehat{u}_j - u_k L_{k,j-d_c-i_c},$
7: **if** v_{k-i_c} not computed **then** set $\widehat{v}_{k-i_c} \leftarrow A^\top u_k$
else go to 11:
8: **if** $\|\widehat{v}_{k-i_c}\| = 0$ **then** deflate:
8a: set $L_{j,j-i_c} = 0$ for all $j > k$
8b: $i_c \leftarrow i_c + 1$ and $d_c \leftarrow d_c - 1$
8c: go to 11:
9: 9a: $L_{k,k-i_c} \leftarrow \|\widehat{v}_{k-i_c}\|, \qquad v_{k-i_c} \leftarrow \widehat{v}_{k-i_c}/L_{k,k-i_c}$
9b: **for** $j = 1, \ldots, i_c$
$\widehat{v}_{k+j} \leftarrow \widehat{v}_{k+j} - v_k L_{k,k-i_c+j}$
10: $\widehat{u}_{k+d_c} \leftarrow Av_{k-i_c} - u_k L_{k,k-i_c}$
11: **until** $k = n + 1$ or $d_c = 0$.

Algorithm 1 is equivalent to the band Lanczos reduction for square symmetric matrices [6] applied simultaneously to $A^\top A$ and $AA^\top$, with starting vectors the columns of $A^\top B$ and B, respectively, or one applied to the larger symmetric $\left[\begin{smallmatrix} 0 & A \\ A^\top & 0 \end{smallmatrix}\right]$, with starting block $\left[\begin{smallmatrix} B \\ 0 \end{smallmatrix}\right]$. For the symmetric version in [6], it is known that after d exact deflations have occurred, then, as in the case $d = 1$, the Lanczos vectors span an invariant subspace and all the eigenvalues of the banded matrix are also eigenvalues of the original matrix. In our case, this translates to the certainty that after d deflation steps occur, the singular values of the partial banded matrix L are a subset of the singular values of A, and the singular values of $\begin{bmatrix} R & L \end{bmatrix}$ are a subset of the singular values of $\begin{bmatrix} B & A \end{bmatrix}$.

It is known that in exact arithmetic, a Krylov subspace generated by a single starting vector can provide only one copy of an eigenvalue. Finding multiplicities requires a block Krylov subspace. Each deflation step in Algorithm 1 corresponds to the convergence of a singular value. This means that when the algorithm terminates, the banded reduction has at most d singular values with multiplicity greater than 1. Denote the final number of columns of L by p. Then, P has the size $(p+d) \times p$. If the SVD of L is $U_L \Sigma_L V_L^\top$, we conjecture that $U_L^\top R$ does not have any completely zero rows (still to be proven!). This means that in fact p is the number of singular values of A with multiplicity smaller than $d+1$ for which the corresponding left singular subspaces are not orthogonal onto B. This reason is our evidence that the band reduction is in fact a core decomposition for multiple right-hand sides problems.

3.3 Computing optimal solutions and corrections from a core problem

We assume that a core reduction (or a truncated core reduction) is readily computed, *e.g.*, from an SVD-based method or from a Lanczos bi- (or banded) diagonalization method. For the single right-hand side problem $Ax \approx b$, this decomposition is given in (2); for the multiple right-hand side problem $Ax \approx B$, it is given in (4). In order to compute a ScTLS solution for the original problem, we first need to solve in the ScTLS sense the core problem. Then, it is straightforward to arrive at the minimum norm ScTLS solution of the original problem, projecting back the core solution to the original subspace.

In [11], the authors show how one can efficiently solve a core problem in bidiagonal form in each of the ScTLS special cases. For the multiple right-hand sides case, the essential computations for solving the core problem in banded form are the solution of a linear system or the SVD of a matrix with banded structure. Efficient and stable algorithms are available (see, *e.g.*, [8] for the SVD computation of a banded lower triangular matrix).

4 Conclusions

We addressed the approximate linear modeling problems gathered under the roof of the scaled total least squares formulation, and extended the core problems to multiple right-hand sides systems. We discussed that in the multiple right-hand sides case (with, say, d right-hand sides), the core matrix A_{11} has size $(p+d) \times p$ and its singular values are a subset of the distinct singular values of A plus some of their multiples (if any). None of the singular values of the core part has multiplicity bigger than d; moreover, the singular values of A_{11} correspond to left singular subspaces of A that are not orthogonal to the right-hand side B. We propose an SVD form and a banded form of the core system. We provide details of a band Lanczos diagonalization method.

References

[1] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, editors. *Templates for the Solution of Algebraic Eigenvalue Problems: a Practical Guide*. Software, Environments, and Tools. SIAM, Philadelphia, PA, 2000.

[2] Z. Bai and R. W. Freund. A band symmetric Lanczos process based on coupled recurrences with applications. Technical report numerical analysis manuscript, Bell Laboratories, Murray Hill, NJ, USA, 1998.

[3] Å. Björck. A band-Lanczos generalization of bidiagonal decomposition. http://www.mai.liu.se/~akbjo/umea05.pdf, November 2005.

[4] Å. Björck. Bidiagonal decomposition and least squares. http://wwwmaths.anu.edu.au/events/sy2005/odatalks/canb05.pdf, September 2005.

[5] P. Feldmann and R. W. Freund. Reduced-order modeling of large linear subcircuits via a block Lanczos algorithm. In *Proceedings of the* 32*nd Design Automation Conference*, New York, 1995. ACM.

[6] R. W. Freund and P. Feldmann. Reduced-order modeling of large linear passive multiterminal circuits using matrix-padé approximation. In *Proceedings of the Design, Automation and Test in Europe Conference*, pages 530–537, Los Alamitos, CA, 1998. IEEE Computer Society Press.

[7] G. H. Golub and W. Kahan. Calculating the singular values and pseudo-inverse of a matrix. *SIAM Journal on Numerical Analysis*, 2(2):205–224, 1965.

[8] G. H. Golub, F. T. Luk, and M. L. Overton. A block Lanczos method for computing the singular values and corresponding singular vectors of a matrix. *ACM Transactions on Mathematical Software (TOMS)*, 7(2):149–169, 1981.

[9] I. Markovsky, J. C. Willems, B. De Moor, and S. Van Huffel. *Exact and Approximate Modeling of Linear Systems: A Behavioral Approach*. SIAM, To be published 2006.

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

[11] C. C. Paige and Z. Strakoš. Core problems in linear algebraic systems. *SIAM Journal on Matrix Analysis and Applications*, 27(3):861–875, 2006.

[12] A. Ruhe. Implementation aspects of band Lanczos algorithms for computation of eigenvalues of large sparse symmetric matrices. *Mathematics of Computation*, 1979.

[13] D. M. Sima and S. Van Huffel. Level choice in truncation methods for linear models. Manuscript, 2006.

[14] S. Van Huffel and J. Vandewalle. Analysis and properties of the generalized total least squares problem $AX \approx B$ when some or all columns in A are subject to error. *SIAM Journal on Matrix Analysis and Applications*, 10(3):294–315, 1989.

[15] S. Van Huffel and J. Vandewalle. *The Total Least Squares Problem: Computational Aspects and Analysis*, volume 9 of *Frontiers in Applied Mathematics*. SIAM, Philadelphia, 1991.