FILTER FACTORS OF TRUNCATED TLS REGULARIZATION WITH MULTIPLE OBSERVATIONS

Iveta Hnětynková, Praha, Martin Plešinger, Liberec, Praha,
Jana Žáková, Liberec

Received August 18, 2016. First published February 28, 2017.

Abstract. The total least squares (TLS) and truncated TLS (T-TLS) methods are widely known linear data fitting approaches, often used also in the context of very ill-conditioned, rank-deficient, or ill-posed problems. Regularization properties of T-TLS applied to linear approximation problems $Ax \approx b$ were analyzed by Fierro, Golub, Hansen, and O'Leary (1997) through the so-called filter factors allowing to represent the solution in terms of a filtered pseudoinverse of A applied to b. This paper focuses on the situation when multiple observations $b_1, \ldots, b_d$ are available, i.e., the T-TLS method is applied to the problem $AX \approx B$, where $B = [b_1, \ldots, b_d]$ is a matrix. It is proved that the filtering representation of the T-TLS solution can be generalized to this case. The corresponding filter factors are explicitly derived.

Keywords: truncated total least squares; multiple right-hand sides; eigenvalues of rank-d update; ill-posed problem; regularization; filter factors

MSC 2010: 15A18, 65F20, 65F22, 65F30

1. Introduction

In a wide range of applications there is a need to solve linear approximation problems in the form

$$AX \approx B, \quad A \in \mathbb{R}^{m\times n},\ B \in \mathbb{R}^{m\times d},\ X \in \mathbb{R}^{n\times d}. \tag{1.1}$$

This work has been partially supported by the GAČR grant No. GA13-06684S. Iveta Hnětynková is a member of the University Center for Mathematical Modeling, Applied Analysis and Computational Mathematics (Math MAC). The research of Martin Plešinger and Jana Žáková has been supported by the SGS grant of Technical University of Liberec No. 21161/2016.

DOI: 10.21136/AM.2017.0228-16

The matrix A represents a discretized model, the columns of B are observation or measurement (also called data) vectors and the columns of X stand for the unknown solutions.

Since both A and B typically contain errors, the problem (1.1) is usually solved by data fitting approaches looking for some corrections of the observed data or of the model making the problem compatible. Popular methods are ordinary, data, or total least squares methods possibly extended by appropriate constraints; see [18], [12]. When A is ill-conditioned or when the problem (1.1) is ill-posed (meaning that the solution does not depend continuously on the data B), it is necessary to approximate (1.1) by a near problem with better properties; see [8], or [17], Sect. IV.1, p. 85. Such approach is called the regularization, see e.g. [17], [9], [10]. Regularized least squares methods include the truncated singular value decomposition (also called truncated least squares), Tikhonov regularization, truncated (also called regularized) total least squares, and many others, see e.g. [9], [12].

It was shown previously that for problems with $d = 1$ (i.e., single data vector) some of the regularization methods can be interpreted as filtering methods, since the regularized solutions can be written in terms of filtered pseudoinverse of A applied to b, see e.g. [11], Chap. 6, or [4]. The analysis of the corresponding filter factors gives insight into the regularization properties of these methods. However, the case when $d > 1$ (i.e., multiple data) has to our knowledge not been fully addressed.

In the truncated singular value decomposition and Tikhonov regularization, the generalization to $d > 1$ is straightforward, since the filter factors for individual columns of B can be constructed independently. In the paper, we show that this is not true in the truncated total least squares. Thus, we concentrate on the analysis of this method for problems with $d > 1$. We prove that it can also be described as a filtering method by deriving an explicit formula for the underlying filter factors forming a three-way tensor.

Our exposition essentially follows the development in papers [5], [3], [2], and [4] for $d = 1$. We first study spectral properties of rank-d updates of a real symmetric matrix, and the singular value decomposition of a matrix extended by d columns, while generalizing the results presented in [3] and [2]. The work [4] motivates the application in the truncated total least squares regularization of a problem with several observation vectors $b_1, \ldots, b_d$ and the formulation of the filter factors.

For simplicity of derivations, some nonrestrictive assumptions are considered throughout this paper. Let $A^\star B \neq 0$ (otherwise the data vectors are uncorrelated with the model and thus the only reasonable solution is $X = 0$). Let (1.1) be *incompatible*, i.e., $\mathscr{R}(B) \not\subseteq \mathscr{R}(A)$ (otherwise there exists a solution matrix X such that $B - AX = 0$ and no least squares minimization is required), and *overdetermined*, i.e., $m \geqslant n + d$ (otherwise one can add zero rows to A and B). Let B have the *full* *column rank*, i.e., $\mathrm{rank}(B) = d$ (otherwise a right-hand side preprocessing can be applied, see [15], [16]).

The paper is organized as follows. Section 2 summarizes the filter factor representation of least squares based regularized solutions. Section 3 describes the total least squares regularization. Section 4 analyzes the eigenvalues and eigenvectors of specific rank-d updates of symmetric matrices. Section 5 derives the filter factor representation of the total least squares based regularized solutions. Section 6 gives the conclusions.

In the text $M^\star$, M^{-1}, $M^\dagger$ denote the transposition, the inverse, and the Moore-Penrose pseudoinverse of M, respectively, I_m (or just I) denotes the square identity matrix of order m and e_j its jth column, $0_{m,n}$ (or just 0) denotes the $m \times n$ zero matrix. Furthermore, $\mathscr{R}(M)$ denotes the range of M, $\|v\|$, $\|M\|$, $\|M\|_F$ denote the Euclidean norm of a vector v, the spectral and the Frobenius norms of a matrix M, respectively.

2. Least squares based regularization by filtering

We start with a definition of the (ordinary) least squares (LS) problem.

Definition 2.1 (Least squares minimization). Let $AX \approx B$ be the approximation problem (1.1). Then

$$\min_{G \in \mathbb{R}^{m \times d}} \|G\|_F \quad \text{subject to} \quad AX = B + G \tag{2.1}$$

is called the LS minimization problem.

Consider the singular value decomposition (SVD) of A,

$$A = U \Sigma V^\star, \quad \Sigma = \begin{bmatrix} \mathrm{diag}(\sigma_1, \ldots, \sigma_n) \\ 0_{m-n,n} \end{bmatrix}, \; \sigma_1 \geqslant \ldots \geqslant \sigma_n \geqslant 0, \tag{2.2}$$

$U = [u_1, \ldots, u_m]$, $V = [v_1, \ldots, v_n]$, $U^\star = U^{-1}$, $V^\star = V^{-1}$. Denote $r = \mathrm{rank}(A)$, i.e., $\sigma_r > \sigma_{r+1} = \ldots = \sigma_n = 0$.

The standard (minimum Frobenius norm) LS solution of (1.1) can be expressed by the Moore-Penrose pseudoinverse of A

$$X^{\mathrm{LS}} = A^\dagger B.$$

From the definition of the LS problem it is obvious that the corrections for individual columns of $B = [b_1, \ldots, b_d]$ represented by the corresponding columns of G can be determined independently. Thus, employing the SVD of A,

$$x_j^{\mathrm{LS}} \equiv X^{\mathrm{LS}}\, e_j = A^\dagger b_j = \sum_{i=1}^{r} \frac{u_i^\star b_j}{\sigma_i} \cdot v_i, \quad j = 1, \ldots, d, \tag{2.3}$$

i.e., the multiple right-hand side LS problem is equivalent to d single right-hand side LS problems, see e.g. [7], Chap. 5, or [1].

It is well known (see e.g. [11], Chap. 6) that if the problem (1.1) is ill-posed, the components of (2.3) corresponding to large i can be dominated by errors in the data B. This is caused by a combination of properties of B and the presence of a significant number of very small singular values in denominators of (2.3). Thus, many regularization methods are based on the idea to suppress the components corresponding to small σ_i. We mention two popular approaches.

2.1. Truncated SVD (T-SVD). In the truncated SVD (T-SVD), also called truncated LS (T-LS) method, the sum in (2.3) is simply truncated, see e.g. [9], Sect. 5.3. Let t, $t < r$, be the *truncation parameter*. Then the T-SVD regularized solution can be expressed as the *filtered pseudoinverse*

$$x_j^{\text{T-SVD}} = \sum_{i=1}^{t} \frac{u_i^\star b_j}{\sigma_i} \cdot v_i = \sum_{i=1}^{r} f_i \cdot \frac{u_i^\star b_j}{\sigma_i} \cdot v_i, \quad j = 1, \ldots, d, \tag{2.4}$$

where

$$f_1 = \ldots = f_t = 1, \quad f_{t+1} = \ldots = f_r = 0 \tag{2.5}$$

are the corresponding *filter factors*. Note that t can also be understood as a numerical rank of A with respect to the given approximation problem.

2.2. Tikhonov regularization. The Tikhonov regularization (see e.g. [9], Sect. 5.1) tries to minimize the norm of the residual while controlling the norm of the corresponding approximate solution. In particular, it minimizes the functional

$$\min_{X \in \mathbb{R}^{n \times d}} (\|B - AX\|_F^2 + \|LX\|_F^2),$$

where $L \in \mathbb{R}^{p \times n}$ is a given matrix. Similarly as in the LS method, the minimization problem is equivalent to d independent single right-hand side Tikhonov minimization problems

$$\min_{x_j \in \mathbb{R}^n} (\|Ax_j - b_j\|^2 + \|Lx_j\|^2), \quad j = 1, \ldots, d.$$

In the simplest case $L = \lambda I \in \mathbb{R}^{n\times n}$, i.e., the balance between the two norms in the minimization is controlled by the so-called *regularization parameter* $\lambda > 0$. The minimizer can be written as the filtered pseudoinverse

$$x_j^{\text{Tikhonov}} = \sum_{i=1}^{r} f_i \cdot \frac{u_i^\star b_j}{\sigma_i} \cdot v_i, \quad j = 1, \dots, d, \tag{2.6}$$

where

$$f_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}, \quad i = 1, \dots, r \tag{2.7}$$

are the filter factors, see e.g. [11], Chap. 6.

3. Total least squares based regularization by filtering

Contrary to the ordinary LS, in the total least squares (TLS) method we seek for a correction of the right-hand side B and also of the system matrix A, so that the corrected system becomes compatible.

Definition 3.1 (Total least squares minimization)**.** Let $AX \approx B$ be the approximation problem (1.1), then

$$\min_{G\in\mathbb{R}^{m\times d},\, E\in\mathbb{R}^{m\times n}} \|[G, E]\|_F \quad \text{subject to} \quad (A+E)X = B+G \tag{3.1}$$

is called the TLS minimization problem.

We directly see that since the correction E is shared by all right-hand sides in B, the TLS problem with $d > 1$ cannot be equivalently reformulated to d independent TLS problems with individual columns of B as single right-hand sides.

The TLS problem is significantly more complicated than the ordinary LS problem. It has been studied for a long time, see in particular [6], [18], [21], [20], [14], and recently also [13]. The analysis is based on the SVD of the system matrix A (2.2) and of the *extended matrix* $[B, A]$. Consider the SVD

$$[B, A] = \widehat{U}\widehat{\Sigma}\widehat{V}^\star, \quad \widehat{\Sigma} = \begin{bmatrix} \operatorname{diag}(\widehat{\sigma}_1, \dots, \widehat{\sigma}_{n+d}) \\ 0_{m-(n+d),n+d} \end{bmatrix}, \quad \widehat{\sigma}_1 \geqslant \dots \geqslant \widehat{\sigma}_{n+d} \geqslant 0, \tag{3.2}$$

$\widehat{U} = [\widehat{u}_1, \dots, \widehat{u}_m]$, $\widehat{V} = [\widehat{v}_1, \dots, \widehat{v}_{n+d}]$, $\widehat{U}^\star = \widehat{U}^{-1}$, $\widehat{V}^\star = \widehat{V}^{-1}$. Consider a *truncation parameter* t, $0 \leqslant t \leqslant n$, chosen so that:

(i) $\widehat{\sigma}_{n-t} > \widehat{\sigma}_{(n-t)+1}$ (for $t = n$, we put formally $\widehat{\sigma}_0 = \infty$), and

(ii) the right-upper block in the partitioning

$$\widehat{V} = \begin{bmatrix} \widehat{V}_{11} & \widehat{V}_{12} \\ \underbrace{\widehat{V}_{21}}_{n-t} & \underbrace{\widehat{V}_{22}}_{t+d} \end{bmatrix} \begin{matrix} \}d \\ \}n \end{matrix}, \quad \text{i.e.} \quad \widehat{V}_{12} = \begin{bmatrix} \widehat{v}_{1,(n-t)+1} & \cdots & \widehat{v}_{1,n+d} \\ \vdots & \ddots & \vdots \\ \widehat{v}_{d,(n-t)+1} & \cdots & \widehat{v}_{d,n+d} \end{bmatrix}, \tag{3.3}$$

is of full row rank, i.e., $\operatorname{rank}(\widehat{V}_{12}) = d$ (for $t = n$, we formally consider $\widehat{V}_{11}$ and $\widehat{V}_{21}$ with no columns).

For given A and B, there always exists at least one choice of t ($t = n$), however, there are several options, in general.

The partitioning (3.3) can be used to analyze as well as to solve the TLS problem. For example, if it is possible to set $t = 0$ or if $\widehat{\sigma}_{(n-t)+1} = \widehat{\sigma}_{n+d}$, then the TLS problem has a solution in the form

$$X^{\mathrm{TLS}} = -\widehat{V}_{22}\widehat{V}_{12}^{\dagger}, \tag{3.4}$$

see [18]. However, there is a principal difficulty that the problem (1.1) may not have a TLS solution, even in the simplest case with $d = 1$, see [6]. When $d > 1$, it may also happen that there exists a TLS solution, but it cannot be obtained in the form (3.4), see [14]. For the full solvability analysis of TLS problems we refer to [14]. Note that other orthogonally invariant norms in the TLS definiton (3.1) can be relevant for some problems, see [19].

3.1. Truncated total least squares (T-TLS). Truncated TLS (T-TLS) minimization represents another regularization method for solving (1.1) in case it is ill-posed. Here the idea is to set a threshold $\varepsilon > 0$ such that all smaller singular values of $[B, A]$ are considered redundant and are removed during the regularization process. More precisely, the T-TLS regularized solution is defined as follows, see [4], Sect. 2 or [14], Lemma 6.2.

Definition 3.2 (Truncated total least squares solution). Let $AX \approx B$ be the approximation problem (1.1). Consider $\varepsilon > 0$ such that:

(i) $\widehat{\sigma}_{n-t} > \varepsilon > \widehat{\sigma}_{(n-t)+1}$ holds for some index t, $0 \leqslant t \leqslant n$, and
(ii) $\widehat{V}_{12}$ in the corresponding partitioning (3.3) is of full column rank.

Then

$$X^{\mathrm{T\text{-}TLS}} = -\widehat{V}_{22}\widehat{V}_{12}^{\dagger} \tag{3.5}$$

is called the T-TLS solution of $AX \approx B$.

Note that in real computations the threshold $\varepsilon > 0$ is always chosen such that the above conditions are satisfied. The value of $n - t$ can again be seen as a numerical rank of $[B, A]$, in particular see Step 2 of Algorithm 3.1 [18], Sect. 3.6. Note that the T-TLS solution coincides with the TLS solution of a modified problem, where all $\widehat{\sigma}_l < \varepsilon$ (for $l = (n-t)+1, \ldots, n+d$) are replaced by any number $\widehat{\sigma}$ satisfying $0 < \widehat{\sigma} < \varepsilon$ (see e.g. [14], Lemma 6.2); $\widehat{\sigma}$ then represents the minimal singular value of the modified problem with the multiplicity $d + t$. Consequently, similarly to the TLS, the T-TLS solution of (1.1) with $d > 1$ cannot be obtained directly from d T-TLS solutions of separated single right-hand side problems corresponding to the individual columns of B.

The T-TLS solution for $d = 1$ was analyzed in [4]. It was shown that it can be written in the form of the filtered pseudoinverse

$$x^{\text{T-TLS}} = \sum_{i=1}^{r} f_i \cdot \frac{u_i^\star b}{\sigma_i} \cdot v_i, \tag{3.6}$$

where

$$f_i = \sum_{l=(n-t)+1}^{n+1} \frac{\widehat{v}_{1,l}^2}{\|\widehat{V}_{12}\|_F^2} \cdot \frac{\sigma_i^2}{\sigma_i^2 - \widehat{\sigma}_l^2} \tag{3.7}$$

are the filter factors. Moreover, the partitioning (3.3) for $d = 1$ gives

$$\|\widehat{V}_{12}\|_F^2 = \|\widehat{V}_{12}\|^2 = \sum_{j=1}^{t+1} \widehat{v}_{1,(n-t)+j}^2.$$

For the detailed derivation see [4], Sect. 3.2.

This derivation employs a link between SVDs of the system matrix A and the extended matrix $[b, A]$ shown in [5], [3], and [2]. Here, the SVDs are related to eigen-decompositions of $AA^\star$ and $[b, A][b, A]^\star$, using the fact that $[b, A][b, A]^\star = AA^\star + bb^\star$ can be interpreted as the rank-one update of $AA^\star$.

In the rest of this paper we use a similar technique to extend the filter factor representation of T-TLS to the problems (1.1) with multiple right-hand sides, i.e., for $d > 1$. In order to do this, we first analyze eigenvalues and eigenvectors of positive semidefinite rank-d updates of symmetric matrices. This result is then used to study the SVD of a matrix extended by d columns.

Let $M \in \mathbb{R}^{m\times m}$, $M = M^\star$, be a real symmetric matrix and

$$M = UDU^\star, \quad \text{where } U^\star = U^{-1},\ D = \operatorname{diag}(\delta_1, \ldots, \delta_m) \tag{4.1}$$

its eigendecomposition. The eigenvalues and eigenvectors of general symmetric rank-one updates of M have been studied in [5], Sect. 5 and [3], see also [7], Sect. 8.4.3, pp. 469–471. For simplicity of the exposition we present the derivation for $d = 2$. The generalization to $d > 2$ is straightforward and it is commented on through the text. Let

$$K = M + \widetilde{w}\widetilde{w}^\star + \widetilde{y}\widetilde{y}^\star = M + [\widetilde{w}, \widetilde{y}][\widetilde{w}, \widetilde{y}]^\star \tag{4.2}$$

be a *positive semidefinite* rank-two update of M (note that a general rank-two update is of the form $M \pm \widetilde{w}\widetilde{w}^\star \pm \widetilde{y}\widetilde{y}^\star$). By denoting $C = U^\star K U$, $w = U^\star \widetilde{w}$, $y = U^\star \widetilde{y}$, we get

$$C = D + [w, y][w, y]^\star, \quad \text{where } w = [w_1, \ldots, w_m]^\star,\ y = [y_1, \ldots, y_m]^\star, \tag{4.3}$$

the rank-two update of the diagonal matrix D.

4.1. Eigenvalues of rank-d update. The eigenvalues of K are roots of the characteristic polynomial

$$\chi_K(\lambda) = \chi_C(\lambda) = \det(C - \lambda I) = \det((D - \lambda I) + (ww^\star + yy^\star)). \tag{4.4}$$

The jth column of $(C - \lambda I)$ can be written as a sum of two components c_j^0 and c_j^1 as follows:

$$((D - \lambda I) + (ww^\star + yy^\star))e_j = \underbrace{e_j(\delta_j - \lambda)}_{c_j^0} + \underbrace{[w, y][w_j, y_j]^\star}_{c_j^1}, \quad j = 1, \ldots, m.$$

Using the linearity of determinants in columns, (4.4) becomes the sum of 2^m determinants of matrices formed by putting all possible combinations of the vectors c_j^0 and c_j^1, $j = 1, \ldots, m$, in their columns. In other words,

$$\chi_K(\lambda) = \sum_{\mu=0}^{2^m-1} \det([c_1^{\beta_{1,\mu}}, c_2^{\beta_{2,\mu}}, \ldots, c_m^{\beta_{m,\mu}}]), \quad \text{where } \beta_{j,\mu} \in \{0, 1\}$$

is the jth digit in the binary representation of μ, i.e.,

$$\sum_{j=1}^{m} \beta_{j,\mu} \cdot 2^{j-1} = \mu.$$

Clearly, if the μth determinant contains *more than two* columns c_j^1 originated in the *rank-two* updating matrix $(ww^\star + yy^\star)$, then it is identically equal to zero. Consequently, (4.4) is a sum of determinants of three types of matrices: containing no, one, or two columns of the updating matrix, i.e.,

$$\chi_K(\lambda) = \overbrace{\prod_{l=1}^{m}(\delta_l - \lambda)}^{\chi_M(\lambda) = \det(D - \lambda I)} + \sum_{i=1}^{m}\left((w_i^2 + y_i^2)\prod_{\substack{l=1\\ l\neq i}}^{m}(\delta_l - \lambda)\right) + \sum_{\substack{i,j=1\\ i\neq j}}^{m}\left(\underbrace{((w_i^2 + y_i^2)(w_j^2 + y_j^2) - (w_iw_j + y_iy_j)^2)}_{(w_iy_j - w_jy_i)^2}\prod_{\substack{l=1\\ l\neq i,j}}^{m}(\delta_l - \lambda)\right). \tag{4.5}$$

The first term is the characteristic polynomial of the original matrix M. The second term contains squares of determinats of all 1×1 submatrices of the factor $[w, y]$ of the updating matrix, and the third term contains squares of determinants of all 2×2 submatrices of $[w, y]$. Obviously, in rank-one updates the third term vanishes. In rank-d updates, $\chi_K(\lambda)$ contains at most $d + 1$ analogously structured terms, where the $(j + 1)$st involves squared determinants of all $j \times j$ submatrices of the factor of the updating matrix, $j = 1, \ldots, d$. The following theorem states the result for $d = 2$ in a simpler way, by employing the secular equation.

Theorem 4.1 (Eigenvalues of rank-two update)**.** *Let $M \in \mathbb{R}^{m\times m}$ be a symmetric matrix with eigenvalues δ_l, $l = 1, \ldots, m$, and let K be its symmetric positive semidefinite rank-two update* (4.2)–(4.3)*. Assume that the spectra of M and K are disjoint. Then the eigenvalues of K are roots of the secular equation*

$$\varphi_K(\lambda) = 1 + \sum_{i=1}^{m}\frac{w_i^2 + y_i^2}{\delta_i - \lambda} + \sum_{\substack{i,j=1\\ i\neq j}}^{m}\frac{\begin{vmatrix} w_i & y_i \\ w_j & y_j \end{vmatrix}^2}{(\delta_i - \lambda)(\delta_j - \lambda)} = 0. \tag{4.6}$$

P r o o f. The secular equation is obtained simply by dividing the characteristic polynomial $\chi_K(\lambda)$ by its first term $\chi_M = \det(D - \lambda I)$. □

Note that in the case when some eigenvalues of M and K coincide (which is easy to verify, since we have the spectrum of M available), we can project the problem onto the subspace orthogonal to the respective eigenspace. For the detailed description of this technique cf. deflation in [3]. For deeper relations between the spectra of M and K and the components of the updating term in the case $d = 1$ we refer to [5], Sect. 5, [3], and [7], Sect. 8.4.3, pp. 469–471.

4.2. Eigenvectors of rank-d update. Let λ_l be the eigenvalues of the updated matrix K. Now we want to determine the corresponding eigenvectors $\widetilde{x}_l$, i.e., to solve

$$K\widetilde{x}_l = (M + \widetilde{w}\widetilde{w}^\star + \widetilde{y}\widetilde{y}^\star)\,\widetilde{x}_l = \widetilde{x}_l\lambda_l, \tag{4.7}$$

or, by denoting $x_l = U^\star\widetilde{x}_l$,

$$Cx_l = (D + ww^\star + yy^\star)\,x_l = x_l\lambda_l. \tag{4.8}$$

The following theorem formulates the result.

Theorem 4.2 (Eigenvectors of rank-two update)**.** *Let $M \in \mathbb{R}^{m\times m}$ be a symmetric matrix, $M = UDU^\star$ its eigendecomposition, and let K be its symmetric positive semidefinite rank-two update* (4.2)–(4.3) *with eigenvalues λ_l, $l = 1,\dots,m$. Assume that the spectra of M and K are disjoint. Denote*

$$D_l = D - \lambda_l I. \tag{4.9}$$

Then the eigenvector $\widetilde{x}_l$ of K corresponding to λ_l has the form

$$\widetilde{x}_l = UD_l^{-1}[w, y]\,p_l = UD_l^{-1}U^\star[\widetilde{w}, \widetilde{y}]\,p_l, \tag{4.10}$$

where $p_l \in \mathbb{R}^2$ is a unit vector.

Proof. Since the spectra of M and K are disjoint, D_l is invertible. Rearranging (4.8) gives

$$\begin{aligned}(D - \lambda_l I)x_l &= -(ww^\star + yy^\star)x_l,\\ x_l &= -D_l^{-1}(w(w^\star x_l) + y(y^\star x_l)),\\ \text{i.e.,}\quad x_l &\in D_l^{-1}\cdot\operatorname{span}\{w, y\} = D_l^{-1}\cdot\operatorname{span}\{U^\star\widetilde{w}, U^\star\widetilde{y}\}.\end{aligned} \tag{4.11}$$

The back-transformation $\widetilde{x}_l = Ux_l$ gives the result. □

The following technical lemma will be useful later.

Lemma 4.3. *Let $D \in \mathbb{R}^{m\times m}$ be a diagonal matrix and $C = D + [w,y][w,y]^\star$ its rank-two update. Let λ_l be an eigenvalue of C such that $D_l = D - \lambda_l I$ is invertible, and let $x_l = D_l^{-1}[w,y]\,p_l$, $\|p_l\| = 1$, be the corresponding eigenvector. Then the vector p_l is the eigenvector of*

$$J = [w,y]^\star D_l^{-1}[w,y] \quad \text{such that} \quad Jp_l = -p_l, \tag{4.12}$$

i.e., J acts like the minus identity on p_l.

Proof. Substituting the formula for x_l into rearranged (4.8) gives

$$\begin{aligned} D_l x_l &= -[w,y][w,y]^\star x_l, \\ \underbrace{D_l D_l^{-1}}_{I}[w,y]\,p_l &= -[w,y]\,\underbrace{[w,y]^\star D_l^{-1}[w,y]}_{J}\,p_l. \end{aligned} \tag{4.13}$$

Consider the eigendecomposition

$$J = Z\Theta Z^\star, \quad \text{where } Z^\star = Z^{-1},\ Z = [z_1, z_2],\ \Theta = \operatorname{diag}(\theta_1, \theta_2).$$

Then (4.13) gives

$$([w,y]Z)\begin{bmatrix} (z_1^\star p_l) \\ (z_2^\star p_l) \end{bmatrix} = [w,y]ZZ^\star p_l = -[w,y]Z\Theta Z^\star p_l = ([w,y]Z)\begin{bmatrix} -\theta_1\,(z_1^\star p_l) \\ -\theta_2\,(z_2^\star p_l) \end{bmatrix}.$$

The linear independence of w and y then implies

$$(z_j^\star p_l) = -\theta_j\,(z_j^\star p_l) \quad \text{for } j = 1,2,$$

i.e., either $(z_j^\star p_l) = 0$ or $\theta_j = -1$. Since $\|p_l\| \neq 0$, there are the following possibilities:

If $(z_1^\star p_l) = 0$, then $(z_2^\star p_l) \neq 0$, $\theta_2 = -1$, so

$$Jp_l = Z\Theta Z^\star p_l = Z\begin{bmatrix} \theta_1\,(z_1^\star p_l) \\ \theta_2\,(z_2^\star p_l) \end{bmatrix} = Z\begin{bmatrix} 0 \\ -1\,(z_2^\star p_l) \end{bmatrix} = -Z\begin{bmatrix} (z_1^\star p_l) \\ (z_2^\star p_l) \end{bmatrix} = -p_l.$$

If $(z_2^\star p_l) = 0$, then the situation is analogous.

If $(z_1^\star p_l) \neq 0$ and $(z_2^\star p_l) \neq 0$, then $\theta_1 = \theta_2 = -1$, so $J = -I$, and $Jp_l = -p_l$. □

5. Filter factors of T-TLS in the multiple right-hand side case

Now we use the results from the previous section to study the link between the SVDs of A, see (2.2), and $[B, A]$, see (3.2), by interpreting $[B, A][B, A]^\star$ as a rank-d update of $AA^\star$. Then we give the formula for T-TLS filter factors.

5.1. Relation between SVDs of the system and extended matrices. Consider the symmetric (positive semidefinite) matrix $M = AA^\star \in \mathbb{R}^{m\times m}$. The SVD (2.2) directly gives its eigendecomposition

$$\begin{aligned} M &= AA^\star = U\Sigma\Sigma^\star U^\star = UDU^\star, \quad \text{where} \\ D &= \Sigma\Sigma^\star = \operatorname{diag}(\sigma_1^2, \ldots, \sigma_n^2, 0_{m-n,m-n}). \end{aligned}$$

Consider a symmetric (positive semidefinite) rank-d update K of M, and its transformation

$$\begin{aligned} K &= [B, A][B, A]^\star = AA^\star + BB^\star = M + BB^\star, \\ C &= U^\star K U = U^\star M U + U^\star BB^\star U = D + (U^\star B)(U^\star B)^\star. \end{aligned}$$

Clearly, nonzero eigenvalues of M are squares of nonzero singular values of A, while nonzero eigenvalues of K are squares of nonzero singular values of $[B, A]$. Thus Theorem 4.1 for matrices M and K defined above allows to relate the singular values of $[B, A]$ and A for $d = 2$.

Corollary 5.1. *Let $AX \approx B$ be the approximation problem* (1.1) *with $d = 2$. Consider the SVDs of the system matrix A and of the extended matrix $[B, A]$, see* (2.2) *and* (3.2), *respectively. Then the singular values $\widehat{\sigma}$ of $[B, A]$ are the roots of the secular equation*

$$\psi_{[B,A]}(\widehat{\sigma}) = 1 + \sum_{i=1}^{m} \frac{b_{i,1}^2 + b_{i,2}^2}{\sigma_i^2 - \widehat{\sigma}^2} + \sum_{\substack{i,j=1 \\ i\neq j}}^{m} \frac{\begin{vmatrix} b_{i,1} & b_{i,2} \\ b_{j,1} & b_{j,2} \end{vmatrix}^2}{(\sigma_i^2 - \widehat{\sigma}^2)(\sigma_j^2 - \widehat{\sigma}^2)} = 0. \tag{5.1}$$

In case of a general d, the eigenvalues are again roots of a secular equation of a more complicated form, as we have explained in the previous section. Thus we do not present it here explicitly. Using Theorem 4.2 and Lemma 4.3 and noticing that the eigenvectors of K are the left singular vectors of $[B, A]$, we get the following corollary for singular vectors.

Corollary 5.2. *Let $AX \approx B$ be the approximation problem* (1.1) *with $d = 2$. Consider the SVDs of the system matrix A and of the extended matrix $[B, A]$, see* (2.2) *and* (3.2), *respectively. Let $\widehat{\sigma}_l$ be a nonzero singular value of $[B, A]$. Then the corresponding left singular vector $\widehat{u}_l$ has the form*

$$\widehat{u}_l = \frac{\widehat{u}'_l}{\|\widehat{u}'_l\|}, \quad \widehat{u}'_l = US_l^{-1}U^\star Bp_l, \quad \text{where } S_l = (\Sigma\Sigma^\star - \widehat{\sigma}_l^2 I) \tag{5.2}$$

and $p_l \in \mathbb{R}^2$ is the unit eigenvector of $(U^\star B)^\star S_l^{-1}(U^\star B)$ by Lemma 4.3. The right singular vector $\widehat{v}_l$ can be obtained by the normalization of $[B, A]^\star \widehat{u}'_l$, i.e.,

$$\widehat{v}_l = \frac{\widehat{v}'_l}{\|\widehat{v}'_l\|}, \quad \widehat{v}'_l = [B, A]^\star \widehat{u}'_l = \begin{bmatrix} B^\star US_l^{-1}U^\star B \\ V\Sigma^\star S_l^{-1}U^\star B \end{bmatrix} p_l = \begin{bmatrix} -I \\ V\Sigma^\star S_l^{-1}U^\star B \end{bmatrix} p_l. \tag{5.3}$$

We see that the previous corollary deals with two different normalizations: The singular vectors $\widehat{u}_l$ and $\widehat{v}_l$ are of unit length $\|\widehat{u}_l\| = \|\widehat{v}_l\| = 1$ as usual, but the norms of the auxiliary vectors $\widehat{u}'_l$ and $\widehat{v}'_l$ are given by the unit length vector p_l. Note that in the case of a general d, the structure of both $\widehat{u}_l$ and $\widehat{v}_l$ remains the same as above, with p_l of the form

$$p_l \equiv [p_{1,l}, \ldots, p_{d,l}]^\star \in \mathbb{R}^d, \quad \|p_l\| = 1. \tag{5.4}$$

Since the T-TLS solution is obtained as a product of blocks of $\widehat{V} = [\widehat{v}_1, \ldots, \widehat{v}_{n+d}]$ (see (3.5)), it will be useful to denote

$$\nu_l \equiv \|\widehat{v}'_l\|, \quad l = 1, \ldots, n + d,$$

the norms of the corresponding auxiliary vectors.

5.2. Filter factors. Comparing the partitioning (3.3) of $\widehat{V}$ with (5.3), we see that the lth columns of $[\widehat{V}_{11}, \widehat{V}_{12}]$ and $[\widehat{V}_{21}, \widehat{V}_{22}]$ are given by

$$[\widehat{V}_{11}, \widehat{V}_{12}]\, e_l = -\frac{1}{\nu_l} \cdot p_l$$

and

$$\begin{aligned} [\widehat{V}_{21}, \widehat{V}_{22}]\, e_l &= \frac{1}{\nu_l} \cdot V\Sigma^\star S_l^{-1}U^\star Bp_l \\ &= \frac{1}{\nu_l} \cdot \sum_{i=1}^{r} \frac{\sigma_i^2}{\sigma_i^2 - \widehat{\sigma}_l^2} \cdot \frac{u_i^\star Bp_l}{\sigma_i} \cdot v_i \\ &= \frac{1}{\nu_l} \cdot \sum_{i=1}^{r} \sum_{j=1}^{d} \frac{\sigma_i^2}{\sigma_i^2 - \widehat{\sigma}_l^2} \cdot p_{j,l} \cdot \frac{u_i^\star b_j}{\sigma_i} \cdot v_i. \end{aligned}$$

Since $\widehat{V}_{12}$ is of full row rank d, the T-TLS solution (3.5) can be written as

$$X^{\text{T-TLS}} = -\widehat{V}_{22}\widehat{V}_{12}^{\dagger} = (\widehat{V}_{22}\,W)\,\underbrace{(-W^{-1}\widehat{V}_{12}^{\star}(\widehat{V}_{12}\widehat{V}_{12}^{\star})^{-1})}_{\Omega}, \tag{5.5}$$

where

$$W \equiv \operatorname{diag}(\nu_{(n-t)+1}, \ldots, \nu_{n+d})$$

allows us to accumulate the normalization coefficients ν_l in only one factor of (5.5) while yielding a result as similar to the orginal one (see [4], Theorem 3.6, p. 1229) as possible. Thus the (l,k)th entry of $\Omega \in \mathbb{R}^{(t+d)\times d}$ is

$$\omega_{l,k} = -\frac{1}{\nu_{(n-t)+l}} \cdot e_l^{\star}\widehat{V}_{12}^{\star}(\widehat{V}_{12}\widehat{V}_{12}^{\star})^{-1}e_k = p_{(n-t)+l}^{\star}\,(\nu_{(n-t)+l}^2\,\widehat{V}_{12}\widehat{V}_{12}^{\star})^{-1}e_k\,.$$

The kth column of the T-TLS solution is then a linear combination of the columns of $\widehat{V}_{22}\,W$ with the coefficients $\omega_{l,k}$. The next theorem summarizes this result and shows how the T-TLS solution can be expressed in terms of a filtered pseudoinverse of A applied to the columns of B.

Theorem 5.3 (Filter factors of T-TLS regularization)**.** *Let* $AX \approx B$ *be the approximation problem* (1.1)*. Consider the SVDs of the system matrix* A *and of the extended matrix* $[B, A]$*, see* (2.2) *and* (3.2)*, respectively. Let* $X^{\text{T-TLS}} \in \mathbb{R}^{n\times d}$ *be its* T-TLS *solution* (3.5)*. The inverse-mapping of* b_j*, the* j*th column of the right-hand side matrix* B*, onto* $x_k^{\text{T-TLS}}$*, the* k*th column of the solution matrix* $X^{\text{T-TLS}}$*, is given by*

$$x_k^{\text{T-TLS}} = \sum_{i=1}^{r}\sum_{j=1}^{d}\underbrace{\left(\sum_{l=(n-t)+1}^{n+d}\omega_{l-(n-t),k}\cdot\frac{\sigma_i^2}{\sigma_i^2-\widehat{\sigma}_l^2}\cdot p_{j,l}\right)}_{f_{i,j,k}}\frac{u_i^{\star}b_j}{\sigma_i}\cdot v_i, \tag{5.6}$$

where $p_{j,l}$ *are the coefficients of the unit vector* p_l (5.4)*, and* $\omega_{l-(n-t),k}$ *are the entries of the matrix* Ω *given by* (5.5)*.*

The proof follows directly from the discussion above. We see that the filter factors $f_{i,j,k}$ in fact form a three-way tensor of the size $r \times d \times d$. The behavior of filter factors requires further research. In particular, the structure of $p_{j,l}$ as well as sizes of $\omega_{l-(n-t),k}$ have to be analyzed. On the other hand, we can conclude that the structure of the filter factors $f_{i,j,k}$ is essentially the same as that of the factors (3.7) for the single right-hand side T-TLS method. In particular, for $d = 1$ we obtain $p_{1,l} = 1$ (see (4.11)), $\widehat{v}_{1,l} = -\nu_l^{-1} p_{1,l} = -\nu_l^{-1}$ (see (5.3)), and

$$\omega_{l-(n-t),k} = \omega_{l-(n-t),1} = -\nu_l^{-1} \cdot \widehat{v}_{1,l} \, \|\widehat{V}_{12}\|_F^{-2} = \widehat{v}_{1,l}^2 \, \|\widehat{V}_{12}\|_F^{-2}.$$

Substituting this into (5.6) gives back the original formula (3.6).

6. Conclusions

In this paper, we have studied the symmetric positive semidefinite rank-2 update of a real symmetric matrix. We have derived the formula for its eigenvectors, and described its eigenvalues as roots of a particular secular equation. We have explained how these results can be generalized to $d > 2$. It has been proved that the T-TLS solution can be expressed as a filtered pseudoinverse of A applied to B, with filter factors given in a tensor form. This generalizes the results obtained previously for $d = 1$. Further analysis of the behavior of the filter factors can help to understand regularization properties of the T-TLS in the future. Such study is however beyond the scope of this paper.

Acknowledgements. We wish to thank Miro Rozložník for his valuable comments which led to improvements of our manuscript.

References

[1] *Å. Björck*: Numerical Methods for Least Squares Problems. Society for Industrial and Applied Mathematics, Philadelphia, 1996. zbl MR doi
[2] *J. R. Bunch, C. P. Nielsen*: Updating the singular value decomposition. Numer. Math. *31* (1978), 111–129. zbl MR doi
[3] *J. R. Bunch, C. P. Nielsen, D. C. Sorensen*: Rank-one modification of the symmetric eigenproblem. Numer. Math. *31* (1978), 31–48. zbl MR doi
[4] *R. D. Fierro, G. H. Golub, P. C. Hansen, D. P. O'Leary*: Regularization by truncated total least squares. SIAM J. Sci. Comput. *18* (1997), 1223–1241. zbl MR doi
[5] *G. H. Golub*: Some modified matrix eigenvalue problems. SIAM Rev. *15* (1973), 318–334. zbl MR doi
[6] *G. H. Golub, C. F. Van Loan*: An analysis of the total least squares problem. SIAM J. Numer. Anal. *17* (1980), 883–893. zbl MR doi
[7] *G. H. Golub, C. F. Van Loan*: Matrix Computations. Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, 2013. zbl MR
[8] *J. Hadamard*: Le problème de Cauchy et les équations aux dérivées partielles linéaires hyperboliques. Hermann, Paris, 1932. zbl
[9] *P. C. Hansen*: Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion. SIAM Monographs on Mathematical Modeling and Computation 4, Society for Industrial and Applied Mathematics, Philadelphia, 1998. zbl MR doi
[10] *P. C. Hansen*: Discrete Inverse Problems. Insight and Algorithms. Fundamentals of Algorithms 7, Society for Industrial and Applied Mathematics, Philadelphia, 2010. zbl MR doi

[11] *P. C. Hansen, J. G. Nagy, D. P. O'Leary*: Deblurring Images. Matrices, Spectra, and Filtering. Fundamentals of Algorithms 3, Society for Industrial and Applied Mathematics, Philadelphia, 2006. zbl MR doi

[12] *P. C. Hansen, V. Pereyra, G. Scherer*: Least Squares Data Fitting with Applications, Johns Hopkins University Press, Baltimore. 2013. zbl MR

[13] *I. Hnětynková, M. Plešinger, D. M. Sima*: Solvability of the core problem with multiple right-hand sides in the TLS sense. SIAM J. Matrix Anal. Appl. *37* (2016), 861–876. zbl MR doi

[14] *I. Hnětynková, M. Plešinger, D. M. Sima, Z. Strakoš, S. Van Huffel*: The total least squares problem in $AX \approx B$: a new classification with the relationship to the classical works. SIAM J. Matrix Anal. Appl. *32* (2011), 748–770. zbl MR doi

[15] *I. Hnětynková, M. Plešinger, Z. Strakoš*: The core problem within a linear approximation problem $AX \approx B$ with multiple right-hand sides. SIAM J. Matrix Anal. Appl. *34* (2013), 917–931. zbl MR doi

[16] *I. Hnětynková, M. Plešinger, Z. Strakoš*: Band generalization of the Golub-Kahan bidiagonalization, generalized Jacobi matrices, and the core problem. SIAM J. Matrix Anal. Appl. *36* (2015), 417–434. zbl MR doi

[17] *F. Natterer*: The Mathematics of Computerized Tomography, B. G. Teubner, Stuttgart; John Wiley & Sons, Chichester. 1986. zbl MR doi

[18] *S. Van Huffel, J. Vandewalle*: The Total Least Squares Problem: Computational Aspects and Analysis. Frontiers in Applied Mathematics 9, Society for Industrial and Applied Mathematics, Philadelphia, 1991. zbl MR doi

[19] *X.-F. Wang*: Total least squares problem with the arbitrary unitarily invariant norms. Linear Multilinear Algebra *65* (2017), 438–456. zbl MR doi

[20] *M. Wei*: Algebraic relations between the total least squares and least squares problems with more than one solution. Numer. Math. *62* (1992), 123–148. zbl MR doi

[21] *M. Wei*: The analysis for the total least squares problem with more than one solution. SIAM J. Matrix Anal. Appl. *13* (1992), 746–763. zbl MR doi

Authors' addresses: *Iveta Hnětynková*, Charles University, Faculty of Mathematics and Physics, Ke Karlovu 3, 121 16 Praha 2, Czech Republic, e-mail: iveta.hnetynkova@mff.cuni.cz; *Martin Plešinger* (corresponding author), Technical University of Liberec, Department of Mathematics, Studentská 1402/2, 461 17 Liberec, Czech Republic and Institute of Computer Science of the Czech Academy of Sciences, Pod Vodárenskou věží 2, 182 07 Praha 8, Czech Republic, e-mail: martin.plesinger@tul.cz; *Jana Žáková*, Technical University of Liberec, Department of Mathematics, Studentská 1402/2, 461 17 Liberec, Czech Republic, e-mail: jana.zakova@tul.cz.