NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS

Numer. Linear Algebra Appl. 2000; **00**:1–6 *Prepared using* nlaauth.cls *[Version: 2000/03/22 v1.0]*

The Structured Total Least Squares approach for nonlinearly structured matrices[†]

P. Lemmerling[*], S. Van Huffel[1] and B. De Moor[1]

[1] *Department of Electrical Engineering, ESAT-SISTA/COSIC, Katholieke Universiteit Leuven, Kardinaal Mercierlaan 94, 3001 Leuven, Belgium.*

SUMMARY

In this paper an extension of the Structured Total Least Squares (STLS) approach for **nonlinearly structured** matrices is presented in the so-called "Riemannian Singular Value Decomposition" (RiSVD) framework. It is shown that this type of STLS problem can be formulated as a Riemannian

---

[*]Correspondence to: Department of Electrical Engineering, ESAT-SISTA/COSIC, Katholieke Universiteit Leuven, Kardinaal Mercierlaan 94, 3001 Leuven, Belgium. `Email: philippe.lemmerling@esat.kuleuven.ac.be. Phone: +32-16-321796. Fax: +32-16-321970.`

[†]Philippe Lemmerling is supported by a post doctoral KUL scholarschip. Sabine Van Huffel is a Research Associate with the F.W.O. (Fund for Scientific Research - Flanders). Bart De Moor is a Research Associate at the F.W.O. This work is supported by several institutions: the Flemish Government (Research Council K.U.Leuven: Concerted Research Action Mefisto-666, the FWO projects G.0240.99, G.0256.97, and Research Communities: ICCoS and ANMMM, IWT projects: EUREKA 2063-IMPACT, STWW, IDO/99/03), the Belgian State, Prime Minister's Office - OSTC - (IUAP P4-02 (1997-2001) and IUAP P4-24 (1997-2001), Sustainable Mobility Programme - Project MD/01/24 (1997-2000), FWO "Krediet aan navorsers" G.0326.98 and the FWO project G.0200.00), the European Commission (TMR Networks: ALAPEDES and System Identification, Brite/Euram Thematic Network : NICONET, project CHRX-CT97-0160).

Received 7 October 2000

SVD of which the smallest singular value and the corresponding right singular vector have to be determined. A heuristic algorithm is proposed. A simple example of a Vandermonde-type matrix is used to demonstrate the improved accuracy of the obtained parameter estimator when compared to other methods such as Least Squares (LS) or Total Least Squares (TLS).

KEY WORDS: Structured Total Least Squares; parameter estimator; nonlinearly structured matrices

1. Introduction

The Structured Total Least Squares (STLS) approach is an extension of the ordinary Total Least Squares (TLS) [1, 2] approach when structured data matrices are involved. Like the TLS method the STLS method is a parameter estimation method for estimating the parameter vector of a linear system that "explains" the data in the observed data matrix.

The TLS approach solves the following problem:

$$\min_{\Delta S, x} \|\Delta S\|_F^2 \tag{1}$$
$$\text{such that } (A + \Delta A)x = b + \Delta b,$$

where $\|.\|_F$ stands for the Frobeniusnorm, $A, \Delta A \in \mathbb{R}^{m \times n}$, $b, \Delta b \in \mathbb{R}^{m \times 1}$, $S \equiv [A \; b]$ is the observed data matrix and $\Delta S \equiv [\Delta A \; \Delta b]$ is the correction applied to S. The latter approach yields Maximum Likelihood (ML) estimates of x when the "noise part[†]" of S has rows that are independently identically distributed (i.i.d.) with common zero mean vector and common

[†]This means that $S = S_0 + S_n$ where S_0 is the unobserved noiseless data contaminated by the noise in the "noise part" S_n.

covariance matrix that is a (unknown) multiple of the identity matrix I_{n+1}.

However, these noise conditions are often violated in signal processing applications in which the observed data matrix is e.g. a Toeplitz matrix constructed from an observed data vector[‡] (see e.g. applications in which linear prediction equations occur). In these frequently occurring cases it can be shown [3] that the STLS approach yields a ML estimate for x:

$$\min_{\Delta S, x} \|\Delta S\|_X^2 \tag{2}$$
$$\text{such that } (A+\Delta A)x = b + \Delta b$$
$$\text{such that } S+\Delta S \text{ has the same structure as } S,$$

with $\|.\|_X$ an appropriate[§] matrix 2−norm.

In recent years many approaches for solving the STLS problem have been proposed [4, 5, 6, 3, 7, 8]. They all start from the same problem formulation (2) but the next step consists of a reformulation of the problem into a constrained optimization problem (the Structured Total Least Norm (STLN) approach [4, 5]), an unconstrained optimization problem (the Constrained Total Least Squares (CTLS) approach [6, 3]) or a nonlinear SVD (the Riemannian SVD (RiSVD) approach [7, 8]). Most of the literature concerns linearly structured matrices (e.g. (block) Hankel matrices, (block) Toeplitz matrices,...). Only recently, an extension towards nonlinearly structured matrices was proposed [9] in the STLN framework.

The major contributions of this paper are the derivation of an extension of the STLS approach for **nonlinearly structured** matrices in the **RiSVD framework** and the development

[‡]As a consequence, entries of the "noise part" of S on the same diagonal are equal.

[§]The right 2−norm has to be chosen if ML properties are desired. E.g. if S is a Toeplitz matrix constructed from an observed data vector s which is perturbed by i.i.d. Gaussian noise, then the following choice should be made: $\|\Delta S\|_X \equiv \|\Delta s\|_2$, with Δs the correction applied to the vector s.

of a **heuristic inverse iteration algorithm** for solving the derived extension. Important applications of nonlinearly structured matrices can be found in signal processing applications. The nonlinear structure described in section 5 is a Vandermonde–type matrix that appears in subspace based parameter estimation algorithms (such as e.g. ESPRIT [10, 11, 12, 13]).

The paper is structured as follows. In section 2 the STLS problem involving nonlinearly structured matrices is formulated and some notation is introduced. The next section derives the RiSVD equations that correspond to the problem stated in section 2. It is shown that the SVD interpretation that can be given to the RiSVD equations derived for linearly structured matrices [7, 8] can also be given to the RiSVD equations derived for nonlinearly structured matrices. A heuristic algorithm for solving the derived RiSVD equations is proposed in section 4. Finally the improved accuracy of the derived parameter estimator is illustrated on a Vandermonde–type matrix that frequently occurs in signal processing applications.

In the remainder of this paper we adopt a Matlab-like [14] notation:

- $A(i,j)$: the entry in the jth column of the ith row of A.
- $A(i,:)$: the ith row of A.
- $A(:,j)$: the jth column of A.
- $A(p:q, r:s)$: the $(q-p+1) \times (s-r+1)$ submatrix of A containing the entries that belong to rows p till q **and** to columns r till s.
- $b(i)$: the entry on the ith row of column vector b.
- $b(p:q)$: the $(q-p+1) \times 1$ subvector of b containing the entries of row p till row q.

2. Problem formulation

In order to formulate the STLS problem (2) for **nonlinearly structured** matrices, we first introduce some notation. Let $T \equiv S + \Delta S$. To indicate the nonlinear structure of a matrix, we make use of matrix functions, indicated by boldface capital letters. E.g.

$$\mathbf{M}(t) : \mathbb{R}^{k\times 1} \to \mathbb{R}^{m\times(n+1)} : t \to \mathbf{M}(t),$$

with e.g.

$$\mathbf{M}(t) = \begin{bmatrix} 1 & \dots & 1 & t(n+1) \\ t(1) & \dots & t(n) & t(n+2) \\ t(1)^2 & \dots & t(n)^2 & t(n+3) \\ \vdots & & & \vdots \\ t(1)^{(m-1)} & \dots & t(n)^{(m-1)} & t(n+m) \end{bmatrix},$$

where $t \in \mathbb{R}^{(m+n)\times 1}$ represents an element of the domain of $\mathbf{M}(.)$. Using this notation, it is clear that the STLS problem for nonlinearly structured matrices can be formulated as follows:

$$\begin{aligned} &\min_{t,y} \sum_{i=1}^{k} w(i)(t(i) - s(i))^2 \\ &\text{such that } Ty = 0 \text{ and } y^T y = 1 \\ &\text{such that } S = \mathbf{M}(s) \text{ and } T = \mathbf{M}(t), \end{aligned} \tag{3}$$

with $s, t \in \mathbb{R}^{k\times 1}$. Note that the equality of structure between T and S is imposed by forcing T and S to be the images under $\mathbf{M}(.)$ of respectively t and s. The objective function in (3) is clearly a weighted $2-$norm applied to the difference of the pre-images of T and S. If the matrix function $\mathbf{M}$ is a linear function, then (3) and (2) are equivalent formulations except for the constraint $y^T y = 1$. In (2) the implicit constraint on the parameter vector x can be formulated as $y(n+1) = -1$. As can be seen by comparing both formulations, in general this change in constraint does not change the solution. This is because the constraint on the parameter vector merely excludes the trivial solution $y = 0$. Therefore, only in some pathological cases[¶] for which $y(n+1) = 0$ at the solution of the STLS problem, the above change in constraints does not yield equivalent formulations. To avoid notational clutter, the weights $w(i), i = 1, \ldots, k$ in (3) will be set to 1 from now on, without loss of generality.

3. Derivation

We will now derive a set of Riemnannian SVD equations, by applying the method of Lagrange multipliers to (3). The Lagrangian for the latter problem is:

$$L(t, y, l, \lambda) = \sum_{i=1}^{k} (t(i) - s(i))^2 + l^T \mathbf{M}(t) y + \lambda (1 - y^T y), \tag{4}$$

with $l \in \mathbb{R}^{m \times 1}$ and $\lambda \in \mathbb{R}$ the Lagrange multipliers. At first sight, the nonlinear dependence of $T \equiv \mathbf{M}(t)$ on t makes a derivation similar to the one for linearly structured (see [7, 8]) matrices T impossible. The reason is that at a certain point in the derivation we need to eliminate t out of our equations. However, we can get around this problem in the following way. Make a Taylor series expansion of each entry of $T = \mathbf{M}(t)$ around its corresponding entry in $S = \mathbf{M}(s)$. Symbolically this can be represented as follows:

$$\mathbf{M}(t) = \mathbf{M}(s) + M^{(1)} + M^{(2)} + \ldots,$$

in which the entries of $M^{(j)} \in \mathbb{R}^{m \times (n+1)}$ contain the jth order terms of the Taylor series expansion mentioned before. Let us now define $\hat{H}_2 \in \mathbb{R}^{m \times k}$ as follows

$$(T - S)y \equiv \hat{H}_2 (t - s). \tag{5}$$

¶These cases are referred to as *degenerate* STLS problems [15] but are not considered in this paper.

From the definition it should be clear that the entries of $\hat{H}_2$ are functions of the elements of y, s and $t - s$. If necessary we will indicate this dependence by using the matrix function notation: $\hat{H}_2 = \hat{\mathbf{H}}_{\mathbf{2}}(y, s, t-s)$. Note that if we only retain the first order term of $T = \mathbf{M}(t)$ (i.e. replacing T by $S + M^{(1)}$), the entries of $\hat{H}_2$ are no longer functions of the elements of $t-s$, but only of the elements of y and s. The Lagrangian (4) can now be rewritten as follows:

$$L(t, y, l, \lambda) = \sum_{i=1}^{k} (t(i) - s(i))^2 + l^T(\hat{H}_2(t-s) + Sy) + \lambda(1 - y^T y). \tag{6}$$

Taking the derivatives of L w.r.t. t, y, l and λ and setting these equal to 0 we obtain the following equations:

$$2(t-s) + \hat{H}_2^T l = 0 \tag{7}$$

$$T^T l = \lambda y \tag{8}$$

$$\hat{H}_2(t-s) + Sy = 0 \text{ or } Ty = 0 \tag{9}$$

$$y^T y = 1. \tag{10}$$

Apparently, the introduction of $\hat{H}_2$ and the corresponding reformulation of the Lagrangian (4) into (6), allows us to avoid the difficulty related to the nonlinear dependence of T on t. It should however be noted that (7) contains an approximation. The exact Lagrange multiplier equation (7) should be

$$2(t-s) + \hat{H}_2^T l + \xi = 0$$

where $\xi \in \mathbb{R}^{k \times 1}$ and

$$\xi(i) = l^T \frac{\delta \hat{H}_2}{\delta t(i)} (t-s).$$

Thus by omitting this vector from (7), we assume that the entries of $\hat{H}_2$ do not depend on the elements of $t-s$ (since then $\xi = 0$). As mentioned before this corresponds to an approximation of $T = \mathbf{M}(t)$ by its first order approximation $S+M^{(1)}$. The latter means that the approximation performed in (7) is valid as long as $t-s$ is small enough to justify the replacement of T by its first order approximation.

We now continue our derivation using the four Lagrange multiplier equations (7)-(10). Multiplying the left and right hand side of (8) with y^T and combining this with (9) and (10) we find that $\lambda = 0$. Substituting (7) into (9) we obtain

$$0.5\hat{D}_2 l = Sy, \tag{11}$$

where $\hat{D}_2 \in \mathbb{R}^{m\times m}$ is defined by $\hat{D}_2 \equiv \hat{H}_2\hat{H}_2^T$. As a consequence the entries of $\hat{D}_2$ are functions of the elements of y, s and $t-s$. If necessary we will indicate this dependence by using the matrix function notation: $\hat{D}_2 = \hat{\mathbf{D}}_\mathbf{2}(y, s, t-s)$. Since the entries of $\hat{H}_2$ are linear in the elements of y (see (5)) it is always possible to define the matrix $\hat{H}_1 \in \mathbb{R}^{(n+1)\times k}$ such that

$$\hat{H}_1^T y \equiv \hat{H}_2^T l, \tag{12}$$

with $\hat{H}_1 \in \mathbb{R}^{(n+1)\times k}$ a matrix which entries depend on the elements of l, s and $t-s$. If necessary we will indicate this dependence by using the matrix function notation: $\hat{H}_1 = \hat{\mathbf{H}}_\mathbf{1}(l, s, t-s)$. Multiplying both sides of equation (5) with l^T to the left and combining this with (12) we obtain:

$$y^T(T-S)^T l = y^T\hat{H}_1(t-s).$$

Since the latter equation is valid for all y, we have:

$$(T-S)^T l = \hat{H}_1(t-s). \tag{13}$$

Combining (13) and (8) we obtain:

$$-S^T l = \hat{H}_1(t-s). \tag{14}$$

Substituting (7) in (14) yields

$$S^T l = 0.5\hat{H}_1\hat{H}_2^T l. \tag{15}$$

Substituting (12) into (15) we obtain

$$S^T l = 0.5\hat{D}_1 y \tag{16}$$

where $\hat{D}_1 \in \mathbb{R}^{(n+1)\times(n+1)}$ is defined by $\hat{D}_1 \equiv \hat{H}_1\hat{H}_1^T$. As a consequence the entries of $\hat{D}_1$ are functions of the elements of l, s and $t-s$. Again if necessary this dependence can be shown explicitly by using the matrix function notation: $\hat{D}_1 = \hat{\mathbf{D}}_\mathbf{1}(l, s, t-s)$. Absorbing irrelevant constants in l we have derived (see (11), (16) and (10)) the following set of Riemannian SVD equations:

$$\begin{aligned} Sy &= \hat{D}_2 l \\ S^T l &= \hat{D}_1 y \\ y^T y &= 1. \end{aligned} \tag{17}$$

Let $l = \tau u$ with $u^T u = 1$. We then have

$$\hat{\mathbf{D}}_\mathbf{1}(l, s, t-s) = \tau^2\hat{\mathbf{D}}_\mathbf{1}(u, s, t-s),$$

because the entries of $\hat{D}_1$ are quadratic[||] functions of the elements of l. To avoid notational clutter, we agree that from this point on we assume the symbol $\hat{D}_1$ represents $\hat{\mathbf{D}}_\mathbf{1}(u, s, t-s)$. As already mentioned before, the only purpose of the constraint $y^T y = 1$ is to exclude the trivial solution $y = 0$, without excluding any possible direction of y. This is also the reason why the Lagrange multiplier λ associated with the constraint $y^T y = 1$ turned out to be 0. Therefore

[||]Note that this is due to the fact that the entries of $\hat{H}_1$ are linear functions of the elements of l, see (12).

it is allowed** to change the constraint $y^T y = 1$ into $y^T \Xi y = 1$ where $\Xi \in \mathbb{R}^{(n+1)\times(n+1)}$ is a positive semidefinite matrix. Since the matrix $\hat{D}_1$ is positive semidefinite by construction (see the definition of $\hat{D}_1$ in (16)) the set of equations (17) is equivalent to the following set of Riemannian SVD equations:

$$Sy = \hat{D}_2 u\tau \tag{18}$$

$$S^T u = \hat{D}_1 y\tau \tag{19}$$

$$u^T u = 1 \tag{20}$$

$$y^T \hat{D}_1 y = 1. \tag{21}$$

Note the equivalence with the classical SVD equations. The difference consists mainly in the introduction of the matrices $\hat{D}_1$ and $\hat{D}_2$ whose entries depend on the entries of respectively $u, s, t-s$ and $y, s, t-s$. Furthermore, the difference with the Riemannian SVD equations derived for linearly structured [7, 8] matrices, is the fact that $\hat{D}_1$ and $\hat{D}_2$ *not only* depend on the entries of respectively u and y. Summarizing, we have obtained a set of Riemannian SVD equations containing left and right singular vectors u and y (corresponding to a singular value τ) that determines a stationary point for the STLS problem (3).

Multiplying the left and right-hand side of (18) with $S^T \hat{D}_2^{-1}$ and using (19) we obtain

$$S^T \hat{D}_2^{-1} S y = \hat{D}_1 y \tau^2.$$

In combination with (21) this becomes

$$y^T S^T \hat{D}_2^{-1} S y = \tau^2. \tag{22}$$

In order to give an interpretation to the Riemannian singular value τ we will derive a new

**As mentioned before, we do not consider degenerate STLS problems.

formula for the objective function in (3) which will also be necessary in the heuristic algorithm described further on. Combining (7) with (9) we find that

$$l = 2(\hat{H}_2\hat{H}_2^T)^{-1}Sy.$$

Substituting the latter in (7) yields

$$t - s = -\hat{H}_2^T(\hat{H}_2\hat{H}_2^T)^{-1}Sy. \tag{23}$$

Thus using the definition of $\hat{D}_2$ we find:

$$(t-s)^T(t-s) = y^T S^T \hat{D}_2^{-1} Sy. \tag{24}$$

Comparing (24) with (22), we see that the nonlinear extension of STLS can be formulated as follows:

> Find the triplet (u, τ, y) corresponding to the smallest τ, that solves the Riemannian SVD equations (18)-(21).

Note that to arrive at (23) and the interpretation for τ we assumed $\hat{D}_2$ to be invertible. It is however still possible to give the same interpretation for τ when $\hat{D}_2$ is rank deficient. We could e.g. normalize l in the following way: $u = \frac{l}{\tau} = \frac{l}{(l^T\hat{D}_2 l)^{0.5}}$ and then the interpretation follows right away at the expense of a more difficult constraint on u, namely $u^T\hat{D}_2 u = 1$. Formula (23) can be changed into

$$t - s = -\hat{H}_2^T(\hat{H}_2\hat{H}_2^T)^{\dagger}Sy, \tag{25}$$

where $M^{\dagger}$ denotes the pseudo-inverse of M . The fact that $\hat{D}_2$ is not invertible, means that the system of equations

$$0.5\hat{D}_2 l = Sy$$

is possibly insolvable. Since the latter equation corresponds to one of the Riemannian SVD equations we derived, it is obvious that in the algorithm based on these Riemannian SVD equations the system has to be solvable at convergence. In that case the pseudo-inverse is merely a way to select the smallest norm solution. As a matter of fact, the example we describe in section 5 contains a rank deficient $\hat{D}_2$.

4. Algorithm

Following the interpretation of τ, we propose an inverse iteration algorithm for solving the Riemannian SVD equations (18)-(21). For notational convenience we use $\Delta s \equiv t - s$. Furthermore, the superscript $[k]$ denotes an entity at the kth iteration.

Nonlinear STLS Algorithm

Input: extended data matrix $S \equiv \mathbf{M}(s)$ (note the structure that has to be preserved is contained in the definition of the matrix function)

Output: y, t

Step 1: Initialization: $k = 0$, $\Delta s^{[0]} = 0$, $[Q\ R] = qr(S)$,

$Q_1 = Q(:, 1 : n+1)$, $Q_2 = Q(:, n+2 : m)$, $R_1 = R(1 : n+1, :)$,

$[U\ \Sigma\ V] = svd(S)$, $u^{[0]} = U(:, n+1)$, $y^{[0]} = V(:, n+1)$,

$\tau^{[0]} = \Sigma(n+1, n+1)$

Step 2: $z^{[k+1]} = R_1^{-T} \hat{D}_1^{[k]} y^{[k]}$

Step 3: $w^{[k+1]} = -(Q_2^T \hat{D}_2^{[k]} Q_2)^{-1} Q_2^T \hat{D}_2^{[k]} Q_1 z^{[k+1]}$

Step 4: $u^{[k+1]} = Q_1 z^{[k+1]} + Q_2 w^{[k+1]}$

Step 5: $y^{[k+1]} = R_1^{-1} Q_1^T \hat{D}_2^{[k+1]} u^{[k+1]} \tau^{[k]}$

Step 6: $u^{[k+1]} = u^{[k+1]}/\|u^{[k+1]}\|_2$

Step 7: $\Delta s^{[k+1]} = -\hat{H}_2^{[k+1]^T}(\hat{D}_2^{[k+1]})^{\dagger} S y^{[k+1]}$

Step 8: $\gamma^{[k+1]} = (y^{[k+1]})^T \hat{D}_1^{[k+1]} y^{[k+1]}$

Step 9: $y^{[k+1]} = y^{[k+1]}/(\gamma^{[k+1]})^{0.5}$

Step 10: $\tau^{[k+1]} = (y^{[k+1]})^T S^T u^{[k+1]}$

Step 11: $t^{[k+1]} = s + \Delta s^{[k+1]}$

Step 12: if $rank(M(t^{[k+1]})) < n + 1$ then stop

else goto Step 2

Note that $[Q\ R] = qr(S)$ is a shorthand notation indicating that $Q \in \mathbb{R}^{m\times m}$ and $R \in \mathbb{R}^{m\times(n+1)}$ form the QR decomposition of $S = QR$, $[U\ \Sigma\ V] = svd(S)$ is a shorthand notation indicating that the columns of $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{(n+1)\times(n+1)}$ contain the left and right singular vectors corresponding to the singular values on the diagonal of $\Sigma \in \mathbb{R}^{m\times(n+1)}$. Furthermore, entities that are calculated using other entities always use the most recently calculated entities. E.g. $\hat{D}_2^{[k+1]}$ in Step 7 is calculated using $y^{[k+1]}$ and $\Delta s^{[k]}$.

5. Numerical example

In this section we describe an example consisting of a nonlinearly structured overdetermined system of equations and determine a solution using the nonlinear extension of the RiSVD framework. We start from a Vandermonde system of equations:

$$\begin{bmatrix} 1 & 1 \\ s(1) & s(2) \\ s(1)^2 & s(2)^2 \\ s(1)^3 & s(2)^3 \end{bmatrix} \begin{bmatrix} x(1) \\ x(2) \end{bmatrix} \approx \begin{bmatrix} b(1) \\ b(2) \\ b(3) \\ b(4) \end{bmatrix}. \tag{26}$$

This type of system occurs frequently in subspace based estimation algorithms: after the so-called signal poles have been determined, the linear parameters (amplitudes and phases) are determined solving a system of type (26) (see e.g. [16]). Note that the relation between the vector y in formulation (3) and the vector x in the previously mentioned example is $x = -y(1:n)/y(n+1)$, with $n = 2$ in this case. We assume that the right-hand side $b \in \mathbb{R}^{4\times 1}$ is noise free and that the vector $s \in \mathbb{R}^{2\times 1}$ contains noisy measurements. In order to obtain a solution $x \in \mathbb{R}^{2\times 1}$ we have to solve (3) with

$$T = \mathbf{M}(t) = \begin{bmatrix} 1 & 1 & b(1) \\ t(1) & t(2) & b(2) \\ t(1)^2 & t(2)^2 & b(3) \\ t(1)^3 & t(2)^3 & b(4) \end{bmatrix}. \tag{27}$$

To give an idea on how the involved matrices depend on the different variables of the problem, we give an overview of the most important matrices used in the derivation and in the algorithm:

$$\hat{H}_1 = \begin{bmatrix} \begin{array}{c} u(2)+u(3)(2s(1)+\Delta s(1))+ \\ u(4)(3s(1)^2+3s(1)\Delta s(1)+\Delta s(1)^2) \end{array} & 0 \\ 0 & \begin{array}{c} u(2)+u(3)(2s(2)+\Delta s(2))+ \\ u(4)(3s(2)^2+3s(2)\Delta s(2)+\Delta s(2)^2) \end{array} \\ 0 & 0 \end{bmatrix} \quad (28)$$

$$\hat{H}_2 = \begin{bmatrix} 0 & 0 \\ y(1) & y(2) \\ (2s(1)+\Delta s(1))y(1) & (2s(2)+\Delta s(2))y(2) \\ (3s(1)^2+3s(1)\Delta s(1)+\Delta s(1)^2)y(1) & (3s(2)^2+3s(2)\Delta s(2)+\Delta s(2)^2)y(2) \end{bmatrix}. \quad (29)$$

To test the algorithm, we first generate a consistent Vandermonde system, i.e. we start from a $t_{ex} \in \mathbb{R}^{2\times 1}$ and an $y_{ex} \in \mathbb{R}^{3\times 1}$ such that $T(t_{ex})y_{ex} = 0$ with $\mathbf{M}(t)$ as in (27), $t_{ex} = [1.5\ 2.52]^T$, $x_{ex} = [1\ 2]^T$ and $y_{ex} = [x_{ex}^T\ \ -1]^T/\|[x_{ex}^T\ \ -1]^T\|_2$. The consistency is obtained by setting $b = \mathbf{M}(t_{ex})(:, 1:2)x_{ex}$. Let $w \in \mathbb{R}^{2\times 1}$ be a noise vector, the elements of which are i.i.d. Gaussian noise of standard deviation 0.1. We will now do a Monte-Carlo simulation using 500 noise realizations. In each run we determine the solution $x_{NSTLS} = -(y(1:2))/y(3)$ of problem (3) with $s = t_{ex} + w$ and $\mathbf{M}(t)$ as in (27), using the algorithm described in section 4. For each of the 500 noise realizations, the algorithm converged, meaning that problem (3) was solved. For reasons of comparison we also calculate for each run the LS (x_{LS}) and the TLS solution of the overdetermined Vandermonde system of equations. For the three methods, the relative accuracy of the solution $acc_{meth} \equiv \frac{\|x_{meth}-x_{ex}\|_2}{\|x_{ex}\|_2}$, $meth$ =LS, TLS, NSTLS is averaged over the 500 runs. We obtain the following results:

$$acc_{LS} = 1.934e-2, acc_{TLS} = 1.935e-2, acc_{NSTLS} = 1.233e-13.$$

It is thus obvious that using the nonlinear STLS algorithm we obtain far better results. In the result presented here it seems that the algorithm not only finds a rank deficient approximation (see stop criterion of the algorithm in section 4) $\mathbf{M}(t)$ of $\mathbf{M}(s)$ having the same nonlinear structure (i.e. the algorithm solves problem (3)), but under the given noise circumstances, it turns out that the nonlinear STLS algorithm also yields a far better statistical accuracy than the LS or TLS estimate. This seems an obvious result as one expects intuitively that problem (3) yields ML results when s is perturbed by i.i.d. Gaussian noise. However, a thorough analysis of the statistical properties of the derived estimator is the subject of further research.

6. Conclusions

In this paper we developed an extension of the STLS problem for nonlinearly structured matrices in the RiSVD framework. It is shown that –as for linearly structured matrices– this problem can be reformulated as a Riemannian SVD of which the smallest singular value and the corresponding right singular vector have to be determined.

An inversion iteration-type heuristic algorithm for solving the Riemannian SVD equations is proposed. Using an example of a Vandermonde–type system, it is shown that the algorithm converges. A preliminary statistical analysis shows that the derived estimator outperforms the LS and the TLS estimators for this example.

REFERENCES

1. G. H. Golub and C.F. Van Loan. An analysis of the total least squares problem. *SIAM J. Numer. Anal.*, 1980.

2. S. Van Huffel and J. Vandewalle. *The Total Least Squares Problem: computational aspects and analysis*, volume 9. SIAM, Philadelphia, 1991.
3. T. J. Abatzoglou, J. M. Mendel, and G. A. Harada. The constrained total least squares technique and its applications to harmonic superresolution. *IEEE Transactions on Signal Processing*, 39:1070–1086, 1991.
4. J.B. Rosen, H. Park, and J. Glick. Total least norm formulation and solution for structured problems. *SIAM Journal on Matrix Anal. Appl.*, 17(1):110–128, 1996.
5. S. Van Huffel, H. Park, and J. B. Rosen. Formulation and solution of structured total least norm problems for parameter estimation. *IEEE Trans. on Signal Processing*, 44(10):2464–2474, 1996.
6. T. J. Abatzoglou and J. M. Mendel. Constrained total least squares. In *Proceedings of IEEE International Conf. on Acoustics, Speech & Signal Processing*, pages 1485–1488, Dallas, 1987.
7. B. De Moor. Structured total least squares and l_2 approximation problems. *Linear Algebra and its Applications, Special Issue on Numerical Linear Algebra Methods in Control, Signals and Systems, (Van Dooren, Ammar, Nichols, and Mehrmann, eds.)*, 188-189:163–207, July 1993.
8. B. De Moor. Total least squares for affinely structured matrices and the noisy realization problem. *IEEE Transactions on Signal Processing*, 42:3104–3113, November 1994.
9. J. B. Rosen, H. Park, and J. Glick. Structured total least norm for nonlinear problems. *SIAM Journal on Matrix Analysis and Applications*, 20(1):14–30, 1999.
10. A. Paulraj, R. Roy, and T. Kailath. A subspace rotation approach to signal parameter estimation. *Proceedings IEEE*, pages 1044–1045, July 1986.
11. R. Roy and T. Kailath. Esprit–estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech , Signal Processing*, 37(7):984–995, July 1989.
12. A.J van der Veen, E. F. Deprettere, and A. L. Swindlehurst. Subspace based signal analysis using singular value decompostion. *Proceedings IEEE*, 81:1277–1308, September 1993.
13. D.V. Bhaskar Rao and K.S. Arun. Model based processing of signals: A state space approach. *Proceedings IEEE*, 80, February 1992.
14. The MathWorks Inc. *MATLAB User's Guide. High-Performance Numeric Computation and Visualization Software*, August 1992.
15. P. Lemmerling. *Structured total least squares: analysis,algorithms and applications.* PhD thesis, Dept. of Elect. Eng., ESAT-SISTA/COSIC, Katholieke Universiteit Leuven, May 1999.
16. S. Van Huffel, C. Decanniere, H. Chen, and P. Van Hecke. Algorithm for time-domain nmr data fitting based on total least squares. *J. Magn. Reson.*, A110:228–237, 1994.

