

A TRUNCATED PROJECTED SVD METHOD FOR LINEAR DISCRETE ILL-POSED PROBLEMS

S. MORIGI*, L. REICHEL†, AND F. SGALLARI‡

Abstract. Truncated singular value decomposition is a popular solution method for linear discrete ill-posed problems. However, since the singular value decomposition of the matrix is independent of the right-hand side, there are linear discrete ill-posed problems for which this method fails to yield an accurate approximate solution. This paper describes a new approach to incorporating knowledge about properties of the desired solution into the solution process through an initial projection of the linear discrete ill-posed problem. The projected problem is solved by truncated singular value decomposition. Computed examples illustrate that suitably chosen projections can enhance the accuracy of the computed solution.

Key words. ill-posed problem, inverse problem, decomposition, svd, tsvd

1. Introduction. This paper is concerned with the computation of a meaningful approximate solution of linear systems of equations

$$(1.1) \quad A\mathbf{x} = \mathbf{b}, \quad A \in \mathbb{R}^{m \times n}, \quad \mathbf{x} \in \mathbb{R}^n, \quad \mathbf{b} \in \mathbb{R}^m,$$

with a matrix A of ill-determined rank. Such systems often are referred to as linear discrete ill-posed problems. The singular values of A cluster at the origin and this makes the matrix severely ill-conditioned; the matrix may be singular. The right-hand side \mathbf{b} is assumed to be contaminated by an error $\mathbf{e} \in \mathbb{R}^m$, which may stem from discretization or measurement inaccuracies. For notational simplicity, we will assume that $m \geq n$; if $m > n$, then (1.1) is considered a least-squares problem. The method of this paper, suitably modified, also can be applied when $m < n$.

Let $\hat{\mathbf{b}}$ denote the unknown error-free vector associated with \mathbf{b} , i.e.,

$$(1.2) \quad \mathbf{b} = \hat{\mathbf{b}} + \mathbf{e},$$

and assume that the linear system

$$(1.3) \quad A\mathbf{x} = \hat{\mathbf{b}}$$

is consistent; the available linear system (1.1) is not required to be consistent. We would like to determine a solution $\hat{\mathbf{x}}$ of (1.3), e.g., the solution of minimal Euclidean norm. Since the right-hand side $\hat{\mathbf{b}}$ is not available, we seek to determine an approximation of $\hat{\mathbf{x}}$ by computing an approximate solution of the available linear system of equations (1.1). Truncated Singular Value Decomposition (TSVD) is a popular method for determining such an approximate solution.

Introduce the Singular Value Decomposition (SVD)

$$(1.4) \quad A = U\Sigma V^T$$

*Department of Mathematics, University of Bologna, Piazza Porta S. Donato 5, 40127 Bologna, Italy. E-mail: morigi@dm.unibo.it.

†Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA. E-mail: reichel@math.kent.edu. Research supported in part by NSF grant DMS-0107858 and an OBR Research Challenge Grant.

‡CIRAM Department of Mathematics, University of Bologna, Via Saragozza 8, 40123 Bologna, Italy. E-mail: sgallari@dm.unibo.it. Research supported in part by PRIN 2004 grant 2004014411-005.

where $U = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n] \in \mathbb{R}^{m \times n}$, $U^T U = I$, $V = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] \in \mathbb{R}^{n \times n}$, $V^T V = I$, and I denotes the identity matrix. The singular values are the diagonal entries of $\Sigma = \text{diag}[\sigma_1, \sigma_2, \dots, \sigma_n] \in \mathbb{R}^{n \times n}$ and are ordered according to $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_\ell > \sigma_{\ell+1} = \dots = \sigma_n = 0$, where $\ell = \text{rank } A$; see, e.g., [4] for details on the SVD. The TSVD method determines approximate solutions of (1.1) defined by

$$(1.5) \quad \mathbf{x}_k = \sum_{j=1}^k \frac{\mathbf{u}_j^T \mathbf{b}}{\sigma_j} \mathbf{v}_j, \quad k = 1, 2, \dots, \ell.$$

In particular,

$$(1.6) \quad \mathbf{x}_k \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}.$$

Let $\|\cdot\|$ denote the Euclidean vector norm or the associated induced matrix norm. We note for future reference that \mathbf{x}_k satisfies the constrained minimization problem

$$(1.7) \quad \min_{\mathbf{x}} \|\mathbf{x}\| \quad \text{subject to} \quad \|A_k \mathbf{x} - \mathbf{b}\| \text{ minimal},$$

where

$$A_k = \sum_{j=1}^k \sigma_j \mathbf{u}_j \mathbf{v}_j^T$$

is the best rank- k approximation of A with respect to the norm $\|\cdot\|$. The singular values σ_j and the Fourier coefficients $\mathbf{u}_j^T \mathbf{b}$ provide valuable insight into the properties of the linear discrete ill-posed problem (1.1); see, e.g., Hansen [6, 8] for discussions on the application of TSVD to linear discrete ill-posed problems.

Consider the sequence $\eta_k = \|\mathbf{x}_k - \hat{\mathbf{x}}\|$, $k = 1, 2, \dots, \ell$. Generally, the η_k decrease when k increases and k is fairly small, but due to the error \mathbf{e} in the right-hand side \mathbf{b} and the ill-conditioning of A , the η_k typically increase rapidly with k when k is large. Let $k_* \geq 1$ be the smallest index, such that

$$(1.8) \quad \|\mathbf{x}_{k_*} - \hat{\mathbf{x}}\| = \min_{1 \leq k \leq n} \|\mathbf{x}_k - \hat{\mathbf{x}}\|.$$

The index k_* generally is not explicitly known.

Assume that an estimate δ of the norm of the error \mathbf{e} in \mathbf{b} is available and note that the norm of the residual vectors

$$(1.9) \quad \mathbf{r}_k = \mathbf{b} - A\mathbf{x}_k$$

is a decreasing function of k ; we have $\|\mathbf{r}_k\|^2 = \sum_{j=k+1}^n (\mathbf{u}_j^T \mathbf{b})^2$. The discrepancy principle suggests that the smallest integer $k \geq 1$, such that

$$(1.10) \quad \|\mathbf{r}_k\| \leq c\delta$$

be used as an approximation of k_* , where $c > 1$ is a user-supplied constant. We denote this integer by k_{discr} and the associated approximation of $\hat{\mathbf{x}}$ by $\mathbf{x}_{k_{\text{discr}}}$; see, e.g., Hansen [8] for further details on the discrepancy principle.

For many linear discrete ill-posed problems (1.1), the approximate solution $\mathbf{x}_{k_{\text{discr}}}$ furnished by TSVD and the discrepancy principle is a fairly accurate approximation of $\hat{\mathbf{x}}$. However, for some problems, not only $\mathbf{x}_{k_{\text{discr}}}$ but also \mathbf{x}_{k_*} are poor approximations

of $\hat{\mathbf{x}}$. This situation arises when the subspace $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ does not contain an accurate approximation of $\hat{\mathbf{x}}$ for small values of k , cf. (1.6), and the propagated error, due to the error \mathbf{e} in \mathbf{b} , destroys the accuracy in \mathbf{x}_k when k is large.

Example 1.1. Define $A = \text{diag}[1, 2^{-1}, 2^{-2}, \dots, 2^{-49}] \in \mathbb{R}^{50 \times 50}$. Thus, the matrices U and V in the singular value decomposition (1.4) of A are identity matrices. Let $\hat{\mathbf{b}} = [1, 2^{-1}, 2^{-2}, \dots, 2^{-49}]^T \in \mathbb{R}^{50}$, $\mathbf{e} = [0, 0, \dots, 0, 2^{-1/2} \cdot 10^{-3}, -2^{-1/2} \cdot 10^{-3}]^T \in \mathbb{R}^{50}$, and let \mathbf{b} be given by (1.2). Then $\hat{\mathbf{x}} = [1, 1, \dots, 1]^T \in \mathbb{R}^{50}$ and $\|\mathbf{e}\| = 1 \cdot 10^{-3}$.

The approximate solution $\mathbf{x}_{k_{\text{discr}}}$ determined by the discrepancy principle (1.10) depends on the choice of the constant c . If $c > 1$ is sufficiently close to one, then the discrepancy principle yields $\mathbf{x}_{49} = [1, \dots, 1, 1 + 2^{47.5} \cdot 10^{-3}, 0]$. Since the penultimate component is about $2 \cdot 10^{11}$, \mathbf{x}_{49} is a poor approximation of $\hat{\mathbf{x}}$. For larger values of c , we obtain approximate solutions \mathbf{x}_k with $1 \leq k \leq 48$, whose first k entries equal one and the remaining $50 - k$ entries vanish. Thus, also in this situation the determined approximate solutions \mathbf{x}_k are poor approximations of $\hat{\mathbf{x}}$. For instance, computations in Matlab with $c = 1.01$ yields the approximate solution \mathbf{x}_{13} with an error $\|\mathbf{x}_{13} - \hat{\mathbf{x}}\| = 6.08$.

The example as presented is finite-dimensional; however, it easily can be extended to infinite dimensions, such that the Picard condition is satisfied: let A be the infinite-dimensional diagonal matrix with j th diagonal entry 2^{1-j} and append zeros to $\hat{\mathbf{b}}$, \mathbf{b} , $\hat{\mathbf{x}}$, and \mathbf{e} , to obtain infinite-dimensional vectors. This extension does not change the numerical values presented above. \square

The difficulty to determine an accurate approximation of $\hat{\mathbf{x}}$ in Example 1.1 stems from the fact that the columns of the matrix V are poorly suited for this purpose. We remark that while the example is quite special, matrices and right-hand sides closely related to the ones of Example 1.1 may arise by application of the SVD to the matrix of linear discrete ill-posed problems. Example 4.1 in Section 4 illustrates that there are less artificial linear discrete ill-posed problems for which TSVD does not provide an accurate approximation of the solution $\hat{\mathbf{x}}$ of (1.3). Indeed, quite a few examples that illustrate shortcomings of the TSVD method can be found in the literature. For instance, Hansen et al. [10] discuss linear discrete ill-posed problems in helioseismology for which TSVD performs poorly. Calvetti et al. [2, Example 4.2] consider a linear discrete ill-posed problem (1.1) with a square matrix A and show that k steps of the GMRES iterative method, for a suitable value of k , determines a more accurate approximation of $\hat{\mathbf{x}}$ in the Krylov subspace $\mathcal{K}_k(A, \mathbf{b}) = \text{span}\{\mathbf{b}, A\mathbf{b}, \dots, A^{k-1}\mathbf{b}\}$ than any approximate solution \mathbf{x}_k given by (1.5). Recently, Eldén [3] presented examples for which, for small values of k , the approximate solution \mathbf{x}_k determined by (1.5) does not approximate $\hat{\mathbf{x}}$ as well as the approximate solution in the k -dimensional Krylov subspace

$$\mathcal{K}_k(A^T A, A^T \mathbf{b}) = \text{span}\{A^T \mathbf{b}, (A^T A)A^T \mathbf{b}, \dots, (A^T A)^{k-1}A^T \mathbf{b}\}$$

determined by the LSQR iterative method.

This paper describes a new approach to circumventing difficulties that can be encountered in applications of TSVD. We first apply orthogonal projections to the matrix A and right-hand side \mathbf{b} and then use TSVD to determine an approximate solution of the projected problem. We refer to this scheme as the Truncated Projected SVD (TPSVD) method. The method is an adaption of the decomposition scheme discussed in [1] to TSVD. The purpose of the projections is to split the solution subspace into two parts, one of which is supplied by the user; the other one is determined by TSVD. The user-supplied subspace makes it possible to incorporate available infor-

mation about the desired solution $\hat{\mathbf{x}}$. For instance, an available rough approximation of $\hat{\mathbf{x}}$ can be supplied to enhance the quality of the computed approximate solution. If $\hat{\mathbf{x}}$ is known to have a steep slope or be of large magnitude, then a subspace can be supplied that allows the representation of such solutions. Several approaches to the selection of a suitable subspace are discussed in Section 2 and computed examples that illustrate these choices are presented in Section 4.

A variety of approaches have been described in the literature for avoiding the difficulties that can be experienced with TSVD. For instance, Hansen et al. [10] propose to replace the minimization problem (1.7) by

$$\min_{\mathbf{x}} \|L\mathbf{x}\| \quad \text{subject to} \quad \|A_k\mathbf{x} - \mathbf{b}\| \text{ minimal},$$

where L is referred to as a regularization operator. Often, L is chosen to be a scaled finite difference approximation of a differential operator, e.g., the bidiagonal

$$(1.11) \quad L = \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \end{bmatrix} \in \mathbb{R}^{(n-1) \times n}$$

or tridiagonal

$$(1.12) \quad L = \begin{bmatrix} -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \\ & & & -1 & 2 & -1 \end{bmatrix} \in \mathbb{R}^{(n-2) \times n}$$

matrices. Computed examples in [10] illustrate that a suitable choice of L can give a better approximation of $\hat{\mathbf{x}}$ than TSVD.

The Generalized SVD (GSVD) of the matrix pair $\{A, L\}$, where L is a regularization operator, is applied in the Truncated GSVD (TGSVD) method; see Hansen [5]. TGSVD with L given by (1.11) or (1.12) gives for many problems better approximations of $\hat{\mathbf{x}}$ than TSVD. However, we remark that the choice of a suitable regularization operator L is not always easy, and the computational effort required to compute the GSVD of the matrix pair $\{A, L\}$ is quite high even for moderately sized matrices; see, e.g., Paige [11] for a discussion on numerical methods. More details on TGSVD can be found in Section 3.

Yet another approach to avoid the shortcomings of TSVD is presented by Hansen et al. [9], who combine SVD with approximation in the ℓ_1 -norm. The aim of this approach is to be able to detect and provide accurate approximations of discontinuities in $\hat{\mathbf{x}}$.

The variety of available solution schemes indicates that there is no best solution method for all linear discrete ill-posed problems. We find the conceptual and computational simplicity of the TPSVD method of the present paper appealing. Computed examples in Section 4 illustrate that TPSVD is able to determine better approximations of $\hat{\mathbf{x}}$ than TSVD also in situations when it is not obvious how to choose a regularization operator L and, therefore, how to apply TGSVD.

2. The TPSVD method. Let \mathcal{W} denote a user-chosen subspace of \mathbb{R}^n of (small) dimension ℓ , and let the columns of the matrix $W \in \mathbb{R}^{n \times \ell}$ form an orthonormal basis of \mathcal{W} . Introduce the QR-factorization

$$AW = QR,$$

where $Q \in \mathbb{R}^{m \times \ell}$ has orthonormal columns and $R \in \mathbb{R}^{\ell \times \ell}$ is upper triangular. We will assume that the subspace \mathcal{W} is chosen so that AW is of full rank. Then R is nonsingular. Introduce the orthogonal projectors

$$P_W = WW^T, \quad P_W^\perp = I - P_W, \quad P_Q = QQ^T, \quad P_Q^\perp = I - P_Q,$$

and split the solution \mathbf{x} of the linear system of equations (1.1) according to

$$(2.1) \quad \mathbf{x} = \mathbf{x}' + \mathbf{x}'', \quad \mathbf{x}' = P_W \mathbf{x}, \quad \mathbf{x}'' = P_W^\perp \mathbf{x}.$$

This splitting suggests the decomposition of the system (1.1),

$$(2.2) \quad P_Q A \mathbf{x}' + P_Q A \mathbf{x}'' = P_Q \mathbf{b},$$

$$(2.3) \quad P_Q^\perp A \mathbf{x}'' = P_Q^\perp \mathbf{b},$$

where we have used the fact that $P_Q^\perp A P_W = 0$ in the derivation of (2.3). TSVD is applied to the projected linear system (2.3). Since $P_Q^\perp A P_W^\perp = P_Q^\perp A$, it suffices to compute the SVD of $P_Q^\perp A$.

Having computed an approximate solution $\mathbf{x}_k'' \in \mathbb{R}^n \setminus \mathcal{W}$ of (2.3) by TSVD, where k , for instance, is determined by the discrepancy principle, we compute an approximate solution $\mathbf{x}_k' \in \mathcal{W}$ of (2.2). The latter equation can be expressed as

$$(2.4) \quad R \mathbf{z}' = Q^T (\mathbf{b} - A \mathbf{x}_k'').$$

We compute the solution \mathbf{z}_k' of (2.4), evaluate

$$(2.5) \quad \mathbf{x}_k' = W \mathbf{z}_k',$$

and obtain the approximate solution

$$(2.6) \quad \mathbf{x}_k = \mathbf{x}_k' + \mathbf{x}_k''$$

of (1.1). Note that regularization is only carried out in the subspace orthogonal to \mathcal{W} . Therefore this space should be chosen so that the matrix $R = Q^T A W$ is not very ill-conditioned. For most linear discrete ill-posed problems this condition is satisfied when \mathcal{W} represents smooth functions with few sign changes, such as algebraic or trigonometric polynomials of low degree. The following result, shown in [1], relates the residual errors associated with \mathbf{x}_k'' and \mathbf{x}_k .

THEOREM 2.1. *Let \mathbf{x}_k'' be an approximate solution of (2.3), assume that the linear system of equations (2.4) is solved exactly for \mathbf{z}_k' , and let \mathbf{x}_k' and \mathbf{x}_k be determined by (2.5) and (2.6), respectively. Then*

$$(2.7) \quad \|\mathbf{b} - A \mathbf{x}_k\| = \|P_Q^\perp \mathbf{b} - P_Q^\perp A \mathbf{x}_k''\| = \|\mathbf{b} - A \mathbf{x}_k''\|.$$

Since the system (2.4) is small and typically not very ill-conditioned, we generally are able to determine its solution to high accuracy. The assumption of Theorem 2.1 that this system be solved exactly therefore is reasonable.

Theorem 2.1 shows that the left-hand side of (1.10) can be computed from the residual error of TSVD applied to (2.3). This makes it convenient to use the discrepancy principle or the L-curve criterion for determining a suitable value of k for the approximate solution \mathbf{x}_k computed by the TPSVD method. The computed examples

of Section 4 illustrate the use of the discrepancy principle. However, TPSVD also can be used in conjunction with other criteria for determining a suitable approximate solution \mathbf{x}_k , such as Generalized Cross Validation (GCV). We refer to Hansen [8] for discussions on the L-curve criterion and GCV.

We propose that the subspace \mathcal{W} be chosen so that elements of this space can represent certain desirable known features of $\hat{\mathbf{x}}$, or so that

$$(2.8) \quad \|P_Q^\perp \mathbf{b}\| \ll \|\mathbf{b}\|.$$

This inequality suggests that we choose \mathcal{W} so that $\text{span}(Q)$ contains a significant portion of the error-free right-hand side $\hat{\mathbf{b}}$. We can achieve this by choosing the space \mathcal{W} so that it contains a significant portion of the desired solution $\hat{\mathbf{x}}$. The subspace $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ furnished by TSVD helps to improve the typically rough approximation of $\hat{\mathbf{x}}$ that can be furnished by the vectors in \mathcal{W} .

Example 2.1. Consider the case when W is a unit vector in \mathbb{R}^n . The inequality (2.8) suggests that W be chosen as a solution of the minimization problem

$$\min_{\|W\|=1} \|P_Q^\perp \mathbf{b}\|.$$

When A is invertible, this minimization problem is solved by $W = A^{-1}\mathbf{b}/\|A^{-1}\mathbf{b}\|$, which indicates that it often may be appropriate to choose the subspace $\mathcal{W} = \text{range } W$, so that it contains an accurate approximation of $\hat{\mathbf{x}}$. \square

In many applications of linear discrete ill-posed problems, the general form of the desired solution $\hat{\mathbf{x}}$ is known already before the computation of an approximate solution of (1.1) is begun. For instance, scientists or engineers may know that the solution they seek to determine is of large magnitude, rapidly increasing, or periodic. The framework of the present paper allows incorporation of this information into the solution process. For instance, when the solution is known to be of large magnitude or to increase rapidly and approximately linearly, then using the spaces

$$(2.9) \quad \mathcal{W} = \text{range}[1, 1, \dots, 1]^T$$

and

$$(2.10) \quad \mathcal{W} = \text{range} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ \vdots & \vdots \\ 1 & n \end{bmatrix},$$

respectively, may be beneficial for the accuracy in the computed approximation of $\hat{\mathbf{x}}$.

An alternative approach in these situations is to apply GSVD with a regularization operator L that has the null space (2.9) or (2.10). For instance, the operator (1.11) has the null space (2.9) and the operator (1.12) has the null space (2.10). Often, but not always, these approaches give approximations of $\hat{\mathbf{x}}$ of about the same accuracy. However, the approach of the present paper also can be applied when the choice of regularization operator is not obvious, such as when the desired solution is known to be periodic and can be approximated fairly well by a trigonometric polynomial of low degree, when the solution is known to have jump discontinuities, or in image restoration problems. These situations are discussed in Examples 4.3-4.5 below.

When solving Fredholm integral equations of the first kind, inexpensive preliminary computations on a coarse grid can give a smooth approximation of the desired solution $\hat{\mathbf{x}}$, and this approximation can be used to define the space \mathcal{W} . It may also be attractive to first solve the Fredholm integral equation on the original (fine) grid and then include a smoothed version of the computed solution in the space \mathcal{W} , possibly in addition to the spaces (2.9) or (2.10). These approaches to choosing \mathcal{W} are illustrated in Examples 4.1 and 4.2 below.

The main advantage of the decomposition method of the present paper, when compared with application of TGSVD with a regularization operators L , is the ease of incorporating known properties of the desired solution into the solution process. Computed examples in Section 4 illustrate that suitable choices of spaces \mathcal{W} can improve the quality of the computed approximate solution significantly compared with TSVD.

We conclude this section by showing that the approximate solutions determined by the method of this paper are invariant under linear transformation by W of the solution of the system (1.1).

THEOREM 2.2. *Consider the linear system of equations*

$$(2.11) \quad A\check{\mathbf{x}} = \check{\mathbf{b}},$$

where $\check{\mathbf{b}} = \mathbf{b} + AW\mathbf{y}$ for some $\mathbf{y} \in \mathbb{R}^\ell$. Let the approximate solution \mathbf{x}_k of (1.1) be determined by (2.6), and let $\check{\mathbf{x}}_k$ be an approximate solution of (2.11) computed analogously. Then

$$\check{\mathbf{x}}_k = \mathbf{x}_k + W\mathbf{y}.$$

Proof. Since $P_Q^\perp \check{\mathbf{b}} = P_Q^\perp \mathbf{b}$, the system (2.3) and therefore the vector \mathbf{x}_k'' are invariant under the transformation of the right-hand side. Moreover, we have $Q^T \check{\mathbf{b}} = Q^T \mathbf{b} + R\mathbf{y}$ and therefore the system analogous to (2.4) for (2.11) can be written as

$$(2.12) \quad R\check{\mathbf{z}}' = Q^T(\mathbf{b} - A\mathbf{x}_k'') + R\mathbf{y}.$$

The solution of (2.12) can be expressed as $\check{\mathbf{z}}'_k = \mathbf{z}'_k + \mathbf{y}$, where \mathbf{z}'_k satisfies (2.4). The approximate solution of (2.11) determined by $\check{\mathbf{z}}'_k$ and \mathbf{x}_k'' is given by $\check{\mathbf{x}}_k = W\check{\mathbf{z}}'_k + \mathbf{x}_k''$, which shows the theorem. \square

Note that the approximate solutions determined by TSVD applied to (1.1) are, in general, not invariant under transformations of the form considered in Theorem 2.2. Problems for which such an invariance is desirable can be solved by the method of the present paper with an appropriate choice of the subspace \mathcal{W} .

Example 2.2. The choice (2.9) of \mathcal{W} yields computed solutions that are invariant under addition of a constant to the solution. Such a transformation is caused by change of origin in a mathematical model. \square

Example 2.3. Consider the finite-dimensional linear discrete ill-posed problem of Example 1.1. TPSVD with \mathcal{W} given by (2.9) yields an accurate approximation \mathbf{x}_1 of $\hat{\mathbf{x}}$ that satisfies the discrepancy principle (1.10), e.g., with $c = 1.01$. We have $\|\mathbf{x}_1 - \hat{\mathbf{x}}\| = 1.57 \cdot 10^{-15}$. The computations were carried out in Matlab. \square

3. The TGSVD method. Let $L \in \mathbb{R}^{p \times n}$ with $1 \leq p \leq n$ be of full rank and such that the null spaces of A and L only intersect trivially. The GSVD of the matrix pair $\{A, L\}$ then is given by

$$(3.1) \quad A = \tilde{U}\tilde{\Sigma}\tilde{Z}^{-1}, \quad L = \tilde{V}[\tilde{M}, 0]\tilde{Z}^{-1},$$

where $\tilde{U} = [\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2, \dots, \tilde{\mathbf{u}}_n] \in \mathbb{R}^{m \times n}$, $\tilde{U}^T \tilde{U} = I$, $\tilde{V} \in \mathbb{R}^{p \times p}$, $\tilde{V}^T \tilde{V} = I$, and $\tilde{Z} = [\tilde{\mathbf{z}}_1, \tilde{\mathbf{z}}_2, \dots, \tilde{\mathbf{z}}_n] \in \mathbb{R}^{n \times n}$ is nonsingular. Moreover,

$$\begin{aligned}\tilde{\Sigma} &= \text{diag}[\tilde{\sigma}_1, \tilde{\sigma}_2, \dots, \tilde{\sigma}_p, 1, 1, \dots, 1] \in \mathbb{R}^{n \times n}, \\ \tilde{M} &= \text{diag}[\tilde{\mu}_1, \tilde{\mu}_2, \dots, \tilde{\mu}_p] \in \mathbb{R}^{p \times p}\end{aligned}$$

with $0 = \tilde{\sigma}_1 = \dots = \tilde{\sigma}_{\tilde{\ell}} < \tilde{\sigma}_{\tilde{\ell}+1} \leq \tilde{\sigma}_{\tilde{\ell}+2} \leq \dots \leq \tilde{\sigma}_p \leq 1$, $1 \geq \tilde{\mu}_1 \geq \tilde{\mu}_2 \geq \dots \geq \tilde{\mu}_p > 0$, and $\tilde{\sigma}_j^2 + \tilde{\mu}_j^2 = 1$ for $1 \leq j \leq p$; see, e.g., Hansen [5] or Golub and Van Loan [4, Section 8.7.3] for details. Hansen [5] describes the TGSVD method, which is based on the factorizations (3.1), for the computation of approximations of $\hat{\mathbf{x}}$. In this method, the smallest generalized singular values $\tilde{\sigma}_j/\tilde{\mu}_j$ are set to zero, i.e., TGSVD determines an approximation of $\hat{\mathbf{x}}$ of the form

$$(3.2) \quad \mathbf{x}_k = \sum_{j=k}^p \frac{\tilde{\mathbf{u}}_j^T \mathbf{b}}{\tilde{\sigma}_j} \tilde{\mathbf{z}}_j + \sum_{j=p+1}^n \tilde{\mathbf{u}}_j^T \mathbf{b} \tilde{\mathbf{z}}_j$$

for some $\tilde{\ell} < k \leq p$. We are using the standard enumeration in GSVD of the $\tilde{\sigma}_j$ and $\tilde{\mu}_j$. The computed approximate solutions (3.2) of (1.3) in Section 4 therefore have index k close to p . We remark that $\text{span}\{\tilde{\mathbf{z}}_{p+1}, \tilde{\mathbf{z}}_{p+2}, \dots, \tilde{\mathbf{z}}_n\}$ is the null space of L . The following analog of Theorem 2.2 holds.

THEOREM 3.1. *Let the regularization operator $L \in \mathbb{R}^{p \times n}$, with $p \leq n$, have the null space $\mathcal{W} = \text{range } W$. Consider the linear system of equations*

$$(3.3) \quad A\check{\mathbf{x}} = \check{\mathbf{b}},$$

where $\check{\mathbf{b}} = \mathbf{b} + AW\mathbf{y}$ for some $\mathbf{y} \in \mathbb{R}^\ell$. Let the approximate solution \mathbf{x}_k of (1.1) be given by (3.2), i.e., \mathbf{x}_k is determined by using the GSVD of the matrix pair $\{A, L\}$. Let $\check{\mathbf{x}}_k$ be an approximate solution of (3.3) computed analogously. Then

$$\check{\mathbf{x}}_k = \mathbf{x}_k + W\mathbf{y}.$$

Proof. The computation of \mathbf{x}_k by TGSVD using the GSVD of the matrix pair $\{A, L\}$ is equivalent computing \mathbf{x}_k by TSVD using the SVD of the matrix AL_A^\dagger , where L_A^\dagger is the A -weighted pseudoinverse of A ; see, e.g., Hansen [8, Sections 2.3 and 3.2] for a definition of L_A^\dagger and a discussions of this connection. The theorem follows from the formulas for the latter approach. \square

4. Numerical examples. This section presents a few computed examples which illustrate the performance of TPSVD. All computations were carried out in Matlab with about 16 significant decimal digits. The right-hand sides in the examples below are contaminated by an error \mathbf{e} of relative norm ε , i.e.,

$$(4.1) \quad \frac{\|\mathbf{e}\|}{\|\hat{\mathbf{b}}\|} = \varepsilon.$$

The entries of \mathbf{e} are normally distributed with zero mean and variance about $\varepsilon/(n\|\hat{\mathbf{b}}\|)$. The constant c in the discrepancy principle (1.10) is set to 1.01 and we let $\delta = \varepsilon\|\hat{\mathbf{b}}\|$ in (1.10).

When solving linear discrete ill-posed problems with a contaminated right-hand side, we would like the subspace \mathcal{W} be such that the component of the desired solution

	TPSVD	TSVD	TGSVD
k_{discr}	2	9	998
$\ \mathbf{x}_{k_{\text{discr}}} - \hat{\mathbf{x}}\ /\ \hat{\mathbf{x}}\ $	$4.88 \cdot 10^{-3}$	$2.34 \cdot 10^{-1}$	$6.66 \cdot 10^{-3}$

TABLE 4.1

Comparison of TPSVD, TSVD, and TGSVD for Example 4.1.

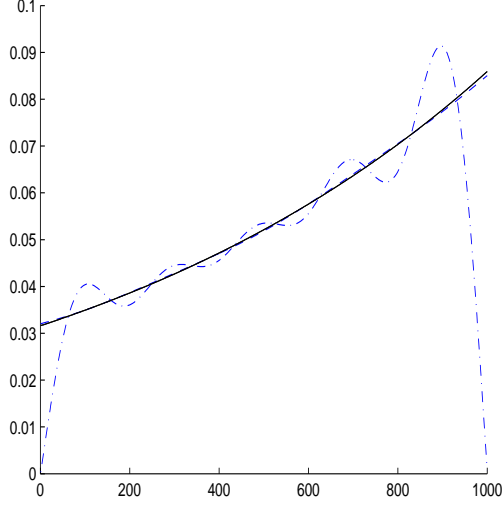


FIG. 4.1. Example 4.1: Exact solution $\hat{\mathbf{x}}$ (continuous curve), approximate solution \mathbf{x}_2 determined by TPSVD using the discrepancy principle (dashed curve), and approximate solution \mathbf{x}_9 determined by TSVD using the discrepancy principle (dash-dotted curve).

$\hat{\mathbf{x}}$ not in \mathcal{W} can be approximated well by a linear combination of the first few columns of the matrix V in (1.4).

Example 4.1. Consider the Fredholm integral equation of the first kind

$$(4.2) \quad \int_0^1 k(s, t)x(t)dt = e^s + (1 - e)s - 1, \quad 0 \leq s \leq 1,$$

where

$$k(s, t) = \begin{cases} s(t - 1), & s < t, \\ t(s - 1), & s \geq t. \end{cases}$$

We discretize the integral equation by a Galerkin method with orthonormal box functions as test and trial functions using the Matlab program `deriv2` from [7]. This program yields a symmetric indefinite matrix $A \in \mathbb{R}^{1000 \times 1000}$ and a scaled discrete approximation $\hat{\mathbf{x}} \in \mathbb{R}^{1000}$ of the solution $x(t) = \exp(t)$ of (4.2). The condition number of A , defined by $\kappa(A) = \|A\|\|A^{-1}\|$, is $1.22 \cdot 10^6$. Figure 4.1 shows $\hat{\mathbf{x}}$ (continuous curve). The error-free right-hand side vector is given by $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$, and the right-hand side vector \mathbf{b} in (1.1) is determined by (1.2) with $\varepsilon = 1 \cdot 10^{-2}$ in (4.1).

We first consider approximants \mathbf{x}_k of $\hat{\mathbf{x}}$ computed by TSVD. The discrepancy principle (1.10) yields $k_{\text{discr}} = 9$. The dash-dotted curve of Figure 4.1 displays \mathbf{x}_9 . The

relative error in \mathbf{x}_9 is seen to be quite large; we have $\|\mathbf{x}_9 - \hat{\mathbf{x}}\|/\|\hat{\mathbf{x}}\| = 2.34 \cdot 10^{-1}$. For comparison, we determine $k_* = 13$ from equation (1.8) and obtain $\|\mathbf{x}_{13} - \hat{\mathbf{x}}\|/\|\hat{\mathbf{x}}\| = 2.06 \cdot 10^{-1}$. Thus, the error in \mathbf{x}_{13} is not much smaller than the error in \mathbf{x}_9 . The low accuracy obtained by TSVD combined with the discrepancy principle therefore does not depend on a failure of the latter, but instead depends on that linear combinations of the first few columns of the matrix V in (1.4) are poorly suited to approximate $\hat{\mathbf{x}}$.

We turn to the TPSVD method. The approximate solution \mathbf{x}_9 determined by TSVD suggests that the desired solution $\hat{\mathbf{x}}$ may be increasing and convex, and that it may be possible to approximate $\hat{\mathbf{x}}$ fairly well by a parabola. We therefore apply TPSVD with

$$(4.3) \quad \mathcal{W} = \text{range} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ \vdots & \vdots & \vdots \\ 1 & n & n^2 \end{bmatrix}.$$

The discrepancy principle (1.10) yields $k_{\text{discr}} = 2$ and the associated approximant \mathbf{x}_2 is displayed in Figure 4.1 (dashed curve). The error $\|\mathbf{x}_2 - \hat{\mathbf{x}}\|/\|\hat{\mathbf{x}}\| = 4.88 \cdot 10^{-3}$ is about 1/48th of the error in the approximate solution \mathbf{x}_9 determined by TSVD.

For this example, TGSVD with the 4-diagonal regularization operator

$$(4.4) \quad L = \begin{bmatrix} -1 & 3 & -3 & 1 & & \\ & -1 & 3 & -3 & 1 & \\ & & \ddots & \ddots & \ddots & \ddots \\ & & & -1 & 3 & -3 & 1 \end{bmatrix} \in \mathbb{R}^{(n-3) \times n},$$

which is a scaled approximation of a 3rd derivative operator with null space is (4.3), performs almost as well as TPSVD. The discrepancy principle (1.10) yields $k_{\text{discr}} = 998$ and we obtain the error $\|\mathbf{x}_{998} - \hat{\mathbf{x}}\|/\|\hat{\mathbf{x}}\| = 6.66 \cdot 10^{-3}$, which is about 36% larger than the corresponding error for TPSVD. Table 4.1 provides a summary of the errors in the computed solutions. This example illustrates that preliminary computations with TSVD can provide valuable information about the solution that can be used to determine a suitable space for TPSVD or a suitable regularization operator L for TGSVD. \square

Example 4.2. This example differs from Example 4.1 only in how knowledge of the desired solution $\hat{\mathbf{x}}$ is incorporated into the solution process. Thus, let the linear system of equations (1.1) be the same as in Example 4.1 and determine the set \mathcal{W} by solving a small linear system of equations defined by a coarse discretization of the integral equation (4.2). Specifically, we discretize (4.2) using a Galerkin method with only four box functions as test and trial functions. This gives the matrix $A_s \in \mathbb{R}^{4 \times 4}$. We determine an associated right-hand side $\mathbf{b}_s \in \mathbb{R}^4$ by projecting the available noisy right-hand side $\mathbf{b} \in \mathbb{R}^{1000}$ into \mathbb{R}^4 . The small linear system of equations

$$(4.5) \quad A_s \mathbf{x} = \mathbf{b}_s$$

obtained in this manner is solved by a direct solution method without regularization. Since the matrix A_s in (4.5) is not very ill-conditioned, $\kappa(A_s) = 18$, regularization is not required. Prolongation of the solution $\mathbf{x}_s \in \mathbb{R}^4$ by piecewise linear interpolation, using the Matlab function `interp1`, yields the approximation $\mathbf{x}_{\text{prol}} \in \mathbb{R}^{1000}$ with error

	TPSVD	TSVD	TGSVD
k_{discr}	13	49	789
$\ \mathbf{x}_{k_{\text{discr}}} - \hat{\mathbf{x}}\ /\ \hat{\mathbf{x}}\ $	$1.86 \cdot 10^{-2}$	$6.49 \cdot 10^{-2}$	$2.26 \cdot 10^{-2}$

TABLE 4.2

Comparison of TPSVD, TSVD, and TGSVD for Example 4.3.

$\|\mathbf{x}_{\text{prol}} - \hat{\mathbf{x}}\| = 2.13 \cdot 10^{-2}$. TPSVD with $\mathcal{W} = \text{span}\{\mathbf{x}_{\text{prol}}\}$, and using the discrepancy principle, gives the approximate solution \mathbf{x}_2 of (1.1) with error $\|\mathbf{x}_2 - \hat{\mathbf{x}}\| = 1.11 \cdot 10^{-2}$. Thus, \mathbf{x}_2 is a better approximation of $\hat{\mathbf{x}}$ than all approximations furnished by TSVD and TGSVD. For comparison, we also note that for this choice of \mathcal{W} , $k_* = 3$ and $\|\mathbf{x}_3 - \hat{\mathbf{x}}\| = 1.08 \cdot 10^{-2}$. \square

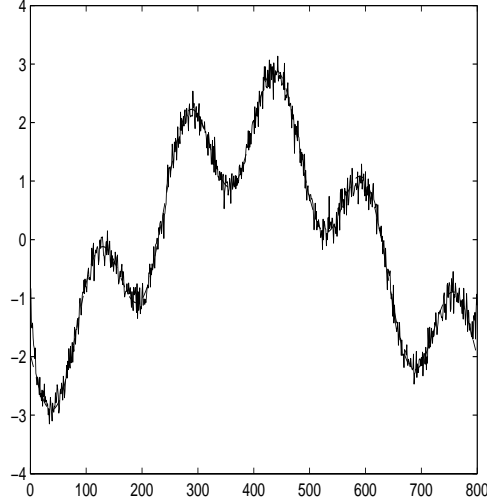


FIG. 4.2. Example 4.3: Original signal $\hat{\mathbf{x}}$ (dashed curve) and the associated corrupted signal (solid curve) with $\varepsilon = 1 \cdot 10^{-1}$ in (4.1).

Example 4.3. We consider the reconstruction of a discrete signal $\hat{\mathbf{x}} \in \mathbb{R}^{800}$ defined by the discretization of

$$x(t) = \sin(5t) + 2 \cos(t), \quad t \in [-\pi, \pi],$$

on a uniform grid with $n = 800$ points. Specifically, we seek to determine an approximation of $\hat{\mathbf{x}}$ from the vector $\mathbf{b} \in \mathbb{R}^{800}$, an available noisy (due to the transmission process) and slightly smoothed (due to the signal capturing method) version of $\hat{\mathbf{x}}$. Let $A \in \mathbb{R}^{800 \times 800}$ be a Toeplitz matrix, such that $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ models convolution of $\hat{\mathbf{x}}$ with a Gaussian with mean zero and variance 0.4. The matrix A is numerically singular; we have $\kappa(A) = 9 \cdot 10^{16}$. The vector $\hat{\mathbf{b}}$ represents a smoothed signal associated with $\hat{\mathbf{x}}$. The available signal \mathbf{b} is obtained by adding noise $\mathbf{e} \in \mathbb{R}^{800}$ to $\hat{\mathbf{b}}$, cf. (1.2), with $\varepsilon = 1 \cdot 10^{-1}$ in (4.1). Figure 4.2 shows the original (unavailable) signal $\hat{\mathbf{x}}$ (dashed curve) and the available corrupted signal \mathbf{b} (continuous curve). The former is barely visible.

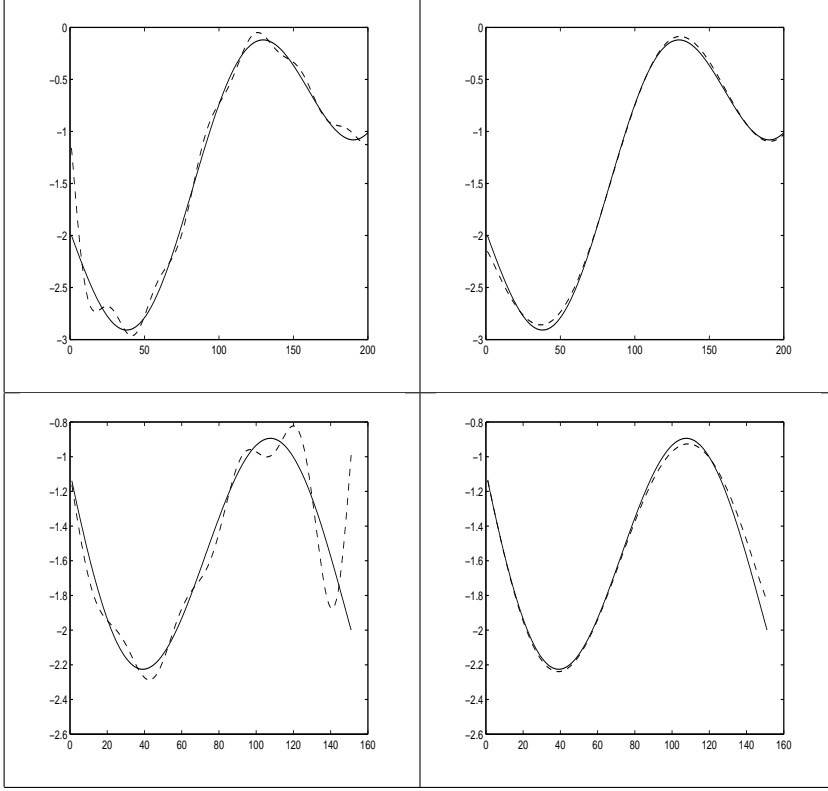


FIG. 4.3. *Example 4.3: The top graphs show the first 200 components of the exact solution $\hat{\mathbf{x}}$ (solid curves), the corresponding components of the approximation \mathbf{x}_{49} determined TSVD using the discrepancy principle (dashed curve, left-hand side graph), and the corresponding components of the approximation \mathbf{x}_{13} determined TPSVD using the discrepancy principle (dashed curve, right-hand side graph). The bottom graphs show the last 150 entries of the exact solution $\hat{\mathbf{x}}$ (solid curves), the corresponding components of \mathbf{x}_{49} determined by TSVD (dashed curve, left-hand side graph), and the corresponding components of \mathbf{x}_{13} determined by TPSVD (dashed curve, right-hand side graph).*

Assume that we know that the sought signal is periodic and of low frequency. It is then natural to apply TPSVD with

$$(4.6) \quad \mathcal{W} = \text{range} \begin{bmatrix} 1 & \sin(-\pi) & \cos(-\pi) \\ 1 & \sin(-\pi + \frac{2\pi}{n-1}) & \cos(-\pi + \frac{2\pi}{n-1}) \\ \vdots & \vdots & \vdots \\ 1 & \sin(\pi) & \cos(\pi) \end{bmatrix} \subset \mathbb{R}^{800 \times 3}.$$

The discrepancy principle (1.10) yields $k_{\text{discr}} = 13$ and we obtain the relative error $\|\mathbf{x}_{13} - \hat{\mathbf{x}}\|/\|\hat{\mathbf{x}}\| = 1.86 \cdot 10^{-2}$. Two different parts of the reconstructed signal \mathbf{x}_{13} are displayed with dashed curves on the right-hand side graphs of Figure 4.3; the top graph shows the first 200 entries of \mathbf{x}_{13} and the bottom graph the last 150 entries. The corresponding parts of the original signal $\hat{\mathbf{x}}$ are depicted with continuous curves.

We now consider approximations \mathbf{x}_k of \mathbf{x} determined by TSVD and obtain $k_{\text{discr}} = 49$ and the relative error $\|\mathbf{x}_{49} - \hat{\mathbf{x}}\|/\|\hat{\mathbf{x}}\| = 6.49 \cdot 10^{-2}$. The dashed curves in the left-hand side graphs of Figure 4.3 display the first 200 (top graph) and last 150 (bottom graph) components of the computed approximation \mathbf{x}_{49} . Comparison of the right-

	TPSVD	TSVD
k_{discr}	782	871
$\ \mathbf{x}_{k_{\text{discr}}} - \hat{\mathbf{x}}\ /\ \hat{\mathbf{x}}\ $	$4.23 \cdot 10^{-2}$	$4.35 \cdot 10^{-2}$

TABLE 4.3

Comparison of TPSVD and TSVD for Example 4.4.

hand side and left-hand side graphs of Figure 4.3 shows the reconstruction obtained by TPSVD using the discrepancy principle to be significantly more accurate than the reconstruction determined by TSVD.

We also apply TGSVD; however, the choice of a suitable regularization operator L is not obvious. Using the operator (4.4) gives a fairly good approximation of $\hat{\mathbf{x}}$. The null space of L is of the same dimension as the space (4.6). The discrepancy principle yields $k_{\text{discr}} = 789$ and the relative error $\|\mathbf{x}_{789} - \hat{\mathbf{x}}\|/\|\hat{\mathbf{x}}\| = 2.26 \cdot 10^{-2}$. Thus, TGSVD furnishes a better approximation of $\hat{\mathbf{x}}$ than TSVD, but worse than TPSVD. The relative errors achieved by the different approaches are summarized in Table 4.2. \square

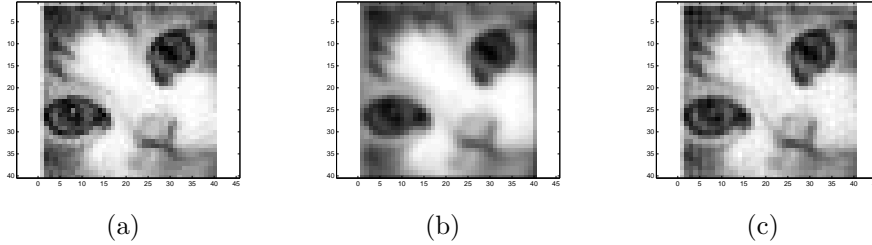


FIG. 4.4. Example 4.4: (a) unavailable original image, (b) available contaminated image, (c) image restored by TPSVD.

Example 4.4. We consider the restoration of a discrete image that has been contaminated by blur and noise. The “original” noise- and blur-free image, which is displayed by Figure 4.4(a), is assumed to be unavailable. It consists of 40×40 pixels, whose values are stored column-wise in the vector $\hat{\mathbf{x}} \in \mathbb{R}^{1600}$. The Matlab function `blur` from [7] with parameters *band* = 3 and *sigma* = 0.7 is applied to generate a blurring operator $A \in \mathbb{R}^{1600 \times 1600}$ that models spatially invariant Gaussian blur. The vector $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ represents a blurred version of the original image $\hat{\mathbf{x}}$. An error $\mathbf{e} \in \mathbb{R}^{1600}$ of relative norm $\varepsilon = 1 \cdot 10^{-2}$, cf. (4.1), which models “noise,” is added to $\hat{\mathbf{b}}$ to obtain the right-hand side \mathbf{b} of (1.1). The latter vector represents the available contaminated version of $\hat{\mathbf{x}}$. Figure 4.4(b) shows the image represented by \mathbf{b} .

We use \mathbf{b} as available a priori information about the desired image; i.e., we apply TPSVD with $\mathcal{W} = \text{span } \mathbf{b}$. The discrepancy principle yields the vector \mathbf{x}_{782} . Figure 4.4(c) shows the “restored image” represented by \mathbf{x}_{782} . TSVD is applied similarly, i.e., the discrepancy principle yields \mathbf{x}_{871} . Table 4.3 summarizes the computed results. \square

Example 4.5. We consider the reconstruction of a piecewise constant signal $\hat{\mathbf{x}} \in$

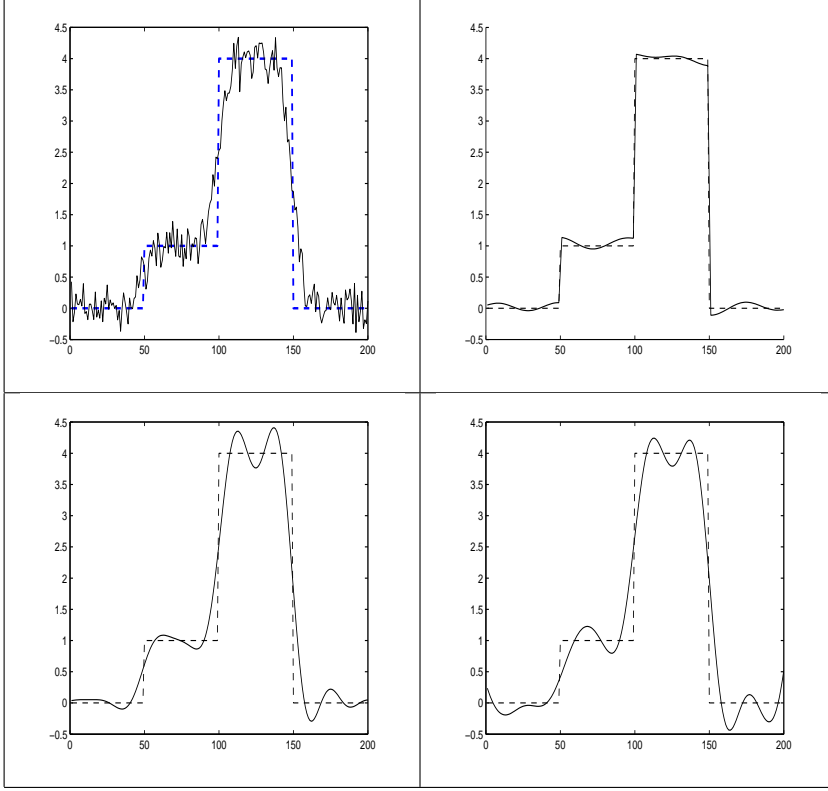


FIG. 4.5. Example 4.5: The top left-hand side graph shows the desired signal $\hat{\mathbf{x}}$ (dashed curve) and the available corrupted signal (solid curve) with relative noise $\varepsilon = 1 \cdot 10^{-1}$. The continuous curves top-right, bottom-right, and bottom-left display, in order, the reconstructed signals determined by TPSVD, TSVD, and TGSVD using the discrepancy principle.

	TPSVD	TSVD	TGSVD
k_{discr}	7	15	186
$\ \mathbf{x}_{k_{\text{discr}}} - \hat{\mathbf{x}}\ / \ \hat{\mathbf{x}}\ $	$8.72 \cdot 10^{-2}$	$1.97 \cdot 10^{-1}$	$2.08 \cdot 10^{-1}$

TABLE 4.4

Comparison of TPSVD, TSVD, and TGSVD for Example 4.5.

\mathbb{R}^{200} defined by the discretization of

$$x(t) = \begin{cases} 0, & t \in [0, 1/4], \\ 1, & t \in (1/4, 1/2], \\ 4, & t \in (1/2, 3/4], \\ 0. & t \in (3/4, 1], \end{cases}$$

on a uniform grid with $n = 200$ points starting from an available vector $\mathbf{b} \in \mathbb{R}^{200}$ contaminated by smear and noise, as in Example 4.3. We suppose that the locations of the discontinuities of the signal are captured by the signal acquisition equipment and therefore are known. These lead us to choose \mathcal{W} as a 3-dimensional space of piecewise constant functions with possible discontinuities at $1/4$, $1/2$, and $3/4$. The operator A is of the same kind as in Example 4.3 and the error $\mathbf{e} \in \mathbb{R}^{200}$ in \mathbf{b} is of relative norm $\varepsilon = 0.1$; cf. (4.1).

Table 4.4 shows the relative errors $\|\mathbf{x}_k - \hat{\mathbf{x}}\|/\|\hat{\mathbf{x}}\|$ achieved by TSVD, TPSVD, and TGSVD, using the discrepancy principle. The latter method is used with the regularization operator (4.4), which has a null space of the same dimension as \mathcal{W} . TPSVD can be seen to yield the best approximation of the desired signal $\hat{\mathbf{x}}$. The graphs in Figure 4.5 illustrate the advantage of TPSVD over both TGSVD and TSVD. The graphs display the exact and computed solution as functions of the grid points.

The conditions of the present example are fairly special. The purpose of the example is to illustrate the ease with which auxiliary information about the desired solution can be incorporated in the solution process. \square

5. Conclusion. The projections in TPSVD provide a novel way to incorporate knowledge about the behavior of $\hat{\mathbf{x}}$, the desired solution of the noise-free problem (1.3), into the solution process. Numerical examples show that TPSVD can give better approximations of $\hat{\mathbf{x}}$ than standard TSVD. TPSVD uses auxiliary information about $\hat{\mathbf{x}}$ in a different way than TGSVD and can be more natural to use in certain situations, such as in Examples 4.2-4.5. TPSVD and TGSVD modify standard TSVD for the solution of linear discrete ill-posed problems. Which one of these methods should be used depends on the problem at hand and on the auxiliary information available.

Acknowledgment. We would like to thank Jing Li and a referee for comments.

REFERENCES

- [1] J. Baglama and L. Reichel, Decomposition methods for large linear discrete ill-posed problems, *J. Comput. Appl. Math.*, in press.
- [2] D. Calvetti, B. Lewis, and L. Reichel, GMRES, L-curves and discrete ill-posed problems, *BIT*, 42 (2002), pp. 44–65.
- [3] L. Eldén, Partial least squares vs. Lanczos bidiagonalization I: analysis of a projection method for multiple regression, *Comp. Stat. Data Anal.*, 46 (2004), pp. 11–31.
- [4] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.
- [5] P. C. Hansen, Regularization, GSVD and truncated GSVD, *BIT*, 29 (1989), pp. 491–504.
- [6] P. C. Hansen, Truncated SVD solutions to discrete ill-posed problems with ill-determined numerical rank, *SIAM J. Sci. Stat. Comput.*, 11 (1990), pp. 503–518.
- [7] P. C. Hansen, Regularization tools: a Matlab package for analysis and solution of discrete ill-posed problems, *Numer. Algorithms*, 6 (1994), pp. 1–35.
- [8] P. C. Hansen, *Rank Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1998.
- [9] P. C. Hansen, M. Jacobsen, J. M. Rasmussen, and H. Sørensen, The PP-TSVD algorithm for image restoration problems, in *Methods and Applications of Inversion*, eds. P. C. Hansen, B. H. Jacobsen, and K. Mosegaard, *Lecture Notes in Earth Sciences # 92*, Springer, Berlin, 2000, pp. 171–202.
- [10] P. C. Hansen, T. Sekii, and H. Shibahashi, The modified truncated SVD method for regularization in general form, *SIAM J. Sci. Stat. Comput.*, 13 (1992), pp. 1142–1150.
- [11] C. C. Paige, Computing the generalized singular value decomposition, *SIAM J. Sci. Stat. Comput.*, 7 (1986), pp. 1126–1146.