Journal of Computational and Applied Mathematics 236 (2012) 2179–2185

Contents lists available at SciVerse ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Combining approximate solutions for linear discrete ill-posed problems

Michiel E. Hochstenbach [a,*], Lothar Reichel [b]

[a] Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB, The Netherlands

[b] Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA

ARTICLE INFO

Article history:
Received 23 November 2010

Keywords:
Ill-posed problem
Linear combination
Solution norm constraint
TSVD
Tikhonov regularization
Discrepancy principle

ABSTRACT

Linear discrete ill-posed problems of small to medium size are commonly solved by first computing the singular value decomposition of the matrix and then determining an approximate solution by one of several available numerical methods, such as the truncated singular value decomposition or Tikhonov regularization. The determination of an approximate solution is relatively inexpensive once the singular value decomposition is available. This paper proposes to compute several approximate solutions by standard methods and then extract a new candidate solution from the linear subspace spanned by the available approximate solutions. We also describe how the method may be used for large-scale problems.

1. Introduction

We are concerned with the numerical solution of linear least-squares problems

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|A\boldsymbol{x}-\boldsymbol{b}\| \tag{1}$$

with a matrix $A \in \mathbb{R}^{m\times n}$ with many singular values of different orders of magnitude close to the origin. Throughout this paper, $\|\cdot\|$ denotes the Euclidean vector norm. The "clustering" of singular values at zero makes the matrix A severely ill-conditioned; in particular, A may be singular. Least-squares problems with a matrix with many singular values of different sizes close to the origin are commonly referred to as discrete ill-posed problems because they arise, for instance, from the discretization of ill-posed problems such as Fredholm integral equations of the first kind. The vector $\boldsymbol{b} \in \mathbb{R}^m$ in discrete ill-posed problems (1) that arise in applications represents measured data and, therefore is typically contaminated by an error $\boldsymbol{e} \in \mathbb{R}^m$. For notational simplicity, we will assume that $m \geq n$; however, the solution methods discussed also can be applied, after minor modifications, when $m < n$.

Let $\hat{\boldsymbol{b}} \in \mathbb{R}^m$ denote the unknown error-free right-hand side vector associated with $\boldsymbol{b}$, i.e.,

$$\boldsymbol{b} = \hat{\boldsymbol{b}} + \boldsymbol{e}.$$

We assume the linear system of equations with the unavailable error-free right-hand side,

$$A\boldsymbol{x} = \hat{\boldsymbol{b}}, \tag{2}$$

to be consistent, and we would like to determine an accurate approximation of its solution $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ of minimal Euclidean norm by computing a suitable approximate solution of the available least-squares problem (1). We remark that, due to the error $\boldsymbol{e}$ in $\boldsymbol{b}$ and the ill-conditioning of A, the straightforward solution of (1) generally does not give a meaningful approximation of $\hat{\boldsymbol{x}}$.

* Corresponding author.
E-mail address: reichel@math.kent.edu (L. Reichel).
URL: http://www.win.tue.nl/~hochsten (M.E. Hochstenbach).

0377-0427/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2011.09.040

Discrete ill-posed problems (1) of small to moderate size are often solved by first computing the singular value decomposition (SVD),

$$A = U\Sigma V^T, \tag{3}$$

where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal matrices and

$$\Sigma = \mathrm{diag}[\sigma_1, \sigma_2, \ldots, \sigma_n] \in \mathbb{R}^{m\times n}.$$

The superscript ^{T} denotes transposition and the singular values are ordered according to

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

Availability of the singular value decomposition makes it possible to compute approximations of $\hat{\boldsymbol{x}}$, e.g., by Tikhonov regularization or truncated singular value decomposition (TSVD), in a simple manner. The computationally most demanding part of the solution process is the determination of the SVD. Usually the SVD is applied to compute only one approximation of $\hat{\boldsymbol{x}}$; see, e.g., [1,2] for discussions and illustrations. We propose to first apply the SVD to determine several approximations, say $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p$, of $\hat{\boldsymbol{x}}$ and then to extract a new approximation of $\hat{\boldsymbol{x}}$ from the available approximations. The extraction is carried out by forming a suitable linear combination of $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p$. Numerical examples in Section 4 illustrate the benefit of this approach. We remark that for small to moderate values of p, the computational effort to determine the approximate solutions $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p$ of (1) is negligible in comparison with the arithmetic work required to evaluate the factorization (3) of A.

This paper is organized as follows. Section 2 discusses the details of our approach to form a new linear combination of available approximations of $\hat{\boldsymbol{x}}$. Some methods to determine approximations of $\hat{\boldsymbol{x}}$ using the SVD of A are reviewed in Section 3. Here, we assume that a bound

$$\|\boldsymbol{e}\| \leq \varepsilon \tag{4}$$

is available. This bound makes it possible to use the discrepancy principle when determining approximations of $\hat{\boldsymbol{x}}$. Computed examples are presented in Section 4, and a conclusion and comments on how to extend the approach of this paper to large-scale problems can be found in Section 5.

2. A linear combination approach

Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p$ denote computed approximations of the desired minimal-norm solution $\hat{\boldsymbol{x}}$ of the error-free linear system of Eq. (2). Numerical methods based on the SVD for computing these approximations are described in Section 3. Let

$$m = \min_{i=1,2,\ldots,p} \|\boldsymbol{x}_i\|, \qquad M = \max_{i=1,2,\ldots,p} \|\boldsymbol{x}_i\|.$$

Introduce the linear space

$$\mathcal{W} = \mathrm{span}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p\} \tag{5}$$

and let the columns of $W \in \mathbb{R}^{n\times p}$ form an orthonormal basis for $\mathcal{W}$. The number of approximate solutions, p, typically is fairly small. In the computed examples of Section 4, we let $p = 3$.

We describe an approach for extracting a new approximation $\widetilde{\boldsymbol{x}}$ of $\hat{\boldsymbol{x}}$ from $\mathcal{W}$. Thus,

$$\widetilde{\boldsymbol{x}} = W\widetilde{\boldsymbol{y}} \tag{6}$$

for a certain vector $\widetilde{\boldsymbol{y}} \in \mathbb{R}^p$. We would like to choose $\widetilde{\boldsymbol{y}}$ so that the residual norm $\|\boldsymbol{b} - AW\widetilde{\boldsymbol{y}}\|$ is small. The residual norm is minimized by $\widetilde{\boldsymbol{y}} = (AW)^{\dagger}\boldsymbol{b}$, where $(AW)^{\dagger}$ denotes the Moore–Penrose pseudoinverse of the matrix AW. However, this vector $\widetilde{\boldsymbol{y}}$ may be of (much) larger norm than the desired vector $\hat{\boldsymbol{x}}$, i.e.,

$$\|\widetilde{\boldsymbol{x}}\| = \|\widetilde{\boldsymbol{y}}\| = \|(AW)^{\dagger}\boldsymbol{b}\| > \|\hat{\boldsymbol{x}}\|.$$

This is usually undesirable, and experiments suggest that this generally renders solutions of (much) worse quality than the approach of the present paper. We propose to impose constraints on $\|\widetilde{\boldsymbol{y}}\| = \|\widetilde{\boldsymbol{x}}\|$. For instance, we may require

$$m \leq \|\widetilde{\boldsymbol{y}}\| \leq M$$

for certain constants m and M. The following result shows that under a weak condition it suffices to only consider the upper bound

$$\|\widetilde{\boldsymbol{y}}\| = M. \tag{7}$$

Proposition 2.1. *Consider the constrained least-squares problem*

$$\min_{m\leq\|\boldsymbol{y}\|\leq M} \|\boldsymbol{b} - AW\boldsymbol{y}\|, \tag{8}$$

and assume that $M \leq \|(AW)^{\dagger}\boldsymbol{b}\|$, where $(AW)^{\dagger}$ denotes the Moore–Penrose pseudoinverse of the matrix AW. Then the solution $\widetilde{\boldsymbol{y}}$ of (8) satisfies (7).

Proof. Consider the constrained least-squares problem

$$\min_{\|\boldsymbol{y}\|=\Delta} \|\boldsymbol{b} - AW\boldsymbol{y}\|, \tag{9}$$

and assume that $\Delta \leq \|(AW)^{\dagger}\boldsymbol{b}\|$. Then using Lagrange multipliers, one can show that the solution $\boldsymbol{y}$ of (9) satisfies

$$(W^T A^T AW + \mu I)\boldsymbol{y} = W^T A^T \boldsymbol{b} \tag{10}$$

for some constant $\mu \geq 0$. Here and throughout this paper, I denotes the identity matrix of appropriate order. It can be established, e.g., by using the SVD of the matrix AW, that the norm of the solution $\boldsymbol{y} = \boldsymbol{y}_\mu$ of (10) is a monotonically decreasing function of μ with

$$\lim_{\mu \searrow 0} \|\boldsymbol{y}_\mu\| = \|(AW)^{\dagger}\boldsymbol{b}\|, \qquad \lim_{\mu \to \infty} \|\boldsymbol{y}_\mu\| = 0.$$

Moreover, the norm of the residual error $\|\boldsymbol{b} - AW\boldsymbol{y}_\mu\|$ is monotonically increasing with μ. The proposition follows from these observations. $\square$

Generally, we would like to choose $\Delta = M \approx \|\hat{\boldsymbol{x}}\|$ in (8) and (9). We will return to the choice of M in Section 3.

The solution of (9) with the constraint (7) can be computed efficiently with the aid of the QR factorization

$$AW = QR,$$

where $Q \in \mathbb{R}^{m\times p}$ has orthonormal columns and $R \in \mathbb{R}^{p\times p}$ is upper triangular. Substituting this factorization into (10) yields

$$(R^T R + \mu I)\boldsymbol{y} = R^T Q^T \boldsymbol{b}.$$

These are the normal equations associated with the least-squares problem

$$\min_{\boldsymbol{y}\in\mathbb{R}^p} \left\| \begin{bmatrix} R \\ \sqrt{\mu} I \end{bmatrix} \boldsymbol{y} - \begin{bmatrix} Q^T \boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix} \right\|.$$

We solve this least-squares problem for a sequence of μ-values and apply Newton's method to determine a value of μ that yields a solution $\boldsymbol{y} = \boldsymbol{y}_\mu$ that satisfies (7); see, e.g., [3] for details on these computations.

The following results shed some light on when the minimization problem (8) may yield an improved approximate solution of (1).

Proposition 2.2. *Let $\boldsymbol{x} \in \mathbb{R}^n$ be a given approximation of $\hat{\boldsymbol{x}}$. Assume that there is a vector $\boldsymbol{w} \in \mathcal{W}$ such that $(A\boldsymbol{x} - \boldsymbol{b})^T A\boldsymbol{w} \neq \boldsymbol{0}$. Then there is a vector $\delta\boldsymbol{x} \in \mathcal{W}$ with*

$$\|A(\boldsymbol{x} + \delta\boldsymbol{x}) - \boldsymbol{b}\| < \|A\boldsymbol{x} - \boldsymbol{b}\|.$$

Proof. The result follows from

$$\|A(\boldsymbol{x} + \boldsymbol{w}) - \boldsymbol{b}\|^2 = \|A\boldsymbol{x} - \boldsymbol{b}\|^2 + 2\,(A\boldsymbol{x} - \boldsymbol{b})^T A\boldsymbol{w} + \|A\boldsymbol{w}\|^2$$

and by letting $\delta\boldsymbol{x}$ be a sufficiently small multiple of $\boldsymbol{w}$. $\square$

Proposition 2.3. *Let $\boldsymbol{x} \in \mathbb{R}^n$ satisfy $\|\boldsymbol{x}\| = M$ for some constant M. Assume that there is a vector $\boldsymbol{w} \in \mathcal{W}$ such that*

$$(A\boldsymbol{x} - \boldsymbol{b})^T A\boldsymbol{w} < 0,$$

$$-\gamma \leq \frac{\boldsymbol{x}^T \boldsymbol{w}}{\boldsymbol{w}^T \boldsymbol{w}} < 0, \tag{11}$$

for some $\gamma > 0$ sufficiently small. Then there is a vector $\delta\boldsymbol{x} \in \mathcal{W}$ with

$$\|A(\boldsymbol{x} + \delta\boldsymbol{x}) - \boldsymbol{b}\| < \|A\boldsymbol{x} - \boldsymbol{b}\|, \tag{12}$$

$$\|\boldsymbol{x} + \delta\boldsymbol{x}\| = M. \tag{13}$$

Proof. Let $\delta\boldsymbol{x} = \alpha\boldsymbol{w}$. Then by the proof of Proposition 2.2, the inequality (12) holds for all constants $\alpha > 0$ sufficiently small. We obtain from (13) that

$$M^2 = \|\boldsymbol{x} + \alpha\boldsymbol{w}\|^2 = M^2 + 2\alpha\boldsymbol{x}^T\boldsymbol{w} + \alpha^2\boldsymbol{w}^T\boldsymbol{w}.$$

It follows that $\alpha = -2\boldsymbol{x}^T\boldsymbol{w}/\boldsymbol{w}^T\boldsymbol{w}$. In view of (11), we have $\alpha \leq 2\gamma$. Therefore, by choosing γ sufficiently small, we can secure that (12) holds. $\square$

3. The choice of the search space

First, we review the discrepancy principle and some solutions methods for (1) based on the SVD (3) of A. These methods can be used to determine the search space $\mathcal{W}$. Other approaches to determine suitable components in $\mathcal{W}$ are also discussed.

A vector $\boldsymbol{x}$ is said to satisfy the discrepancy principle if

$$\|A\boldsymbol{x} - \boldsymbol{b}\| \leq \eta\varepsilon, \tag{14}$$

where ε is the error bound (4) and $\eta > 1$ is a user-specified constant. The discrepancy principle is commonly used to determine the truncation index in the truncated SVD method or the regularization parameter in Tikhonov regularization; see below.

Popular techniques for the solution of (1) include:

(a) The truncated SVD method using the discrepancy principle; this method uses the singular value decomposition (3) to determine the approximate solution

$$\boldsymbol{x}_{\rm tsvd} = \sum_{j=1}^{k} \frac{\boldsymbol{u}_j^T \boldsymbol{b}}{\sigma_j} \boldsymbol{v}_j \tag{15}$$

of (1). The truncation index k is chosen as small as possible so that $\boldsymbol{x}_{\rm tsvd}$ satisfies the discrepancy principle (14). Thus, k is such that

$$\sum_{j=k+1}^{n} (\boldsymbol{u}_j^T \boldsymbol{b})^2 \leq (\eta\varepsilon)^2 \leq \sum_{j=k}^{n} (\boldsymbol{u}_j^T \boldsymbol{b})^2.$$

Properties of this method are discussed in, e.g., [1].

(b) Tikhonov regularization using the discrepancy principle: Tikhonov regularization in its simplest form replaces the solution of (1) by the solution of

$$(A^T A + \mu I)\boldsymbol{x} = A^T \boldsymbol{b} \tag{16}$$

for a suitable value of the regularization parameter $\mu > 0$. Denote the solution by $\boldsymbol{x}_\mu$. Substituting the singular value decomposition (3) into (16) shows that

$$\boldsymbol{x}_\mu = \sum_{j=1}^{n} \frac{\sigma_j}{\sigma_j^2 + \mu} (\boldsymbol{u}_j^T \boldsymbol{b}) \boldsymbol{v}_j. \tag{17}$$

The parameter μ is commonly chosen as large as possible so that $\boldsymbol{x}_\mu$ satisfies (14), so that

$$\|\boldsymbol{b} - A\boldsymbol{x}_\mu\|^2 = \sum_{j=1}^{n} \frac{\mu^2}{(\sigma_j^2 + \mu)^2} (\boldsymbol{u}_j^T \boldsymbol{b})^2 = (\eta\varepsilon)^2;$$

see, e.g., [1,4] for properties of this method. The desired value of μ can be determined, e.g., with the aid of Newton's method.

(c) Tikhonov regularization using the quasi-optimality criterion: the regularization parameter $\mu > 0$ in the Tikhonov Eq. (16) is determined so that the solution, which is of the form (17), minimizes $\mu \to \|\mu\, \boldsymbol{x}'(\mu)\|$. This criterion can be applied when no bound (4) for the norm of the error in $\boldsymbol{b}$ is available. Properties of the quasi-optimality criterion have recently been discussed in [5].

The methods mentioned are used to determine a search space in the computed examples of Section 4; each method yields an approximate solution, the span of which defines a search space $\mathcal{W}$; cf. (5). We would like to stress the fact that other solution methods for (1) can also be used to determine components of $\mathcal{W}$. These include the regularized total least-squares method, modified TSVD methods using enriched solution subspaces, generalized singular value decomposition methods, and methods that impose upper or lower bounds on the computed solution or on the norm of the computed solution. Large-scale problems can be handled by applying the approach of the present paper to the reduced problems obtained by Krylov subspace methods; see, e.g., [6,3,7–10] and references therein for discussions on a variety of the mentioned methods. Other selection criteria for the regularization parameter in Tikhonov regularization (16), such as the L-curve, also can be applied to determine candidate solutions for inclusion in $\mathcal{W}$.

Assume that the approximate solutions $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p$ of (1) have been determined by p different methods. We then propose to define the parameter M in (7) by

$$M = \max_{i=1,2,\ldots,p} \|\boldsymbol{x}_i\|.$$

This generally allows the computed solution to be larger than the smallest one(s) of the candidate solutions $\boldsymbol{x}_i$; moreover, it may be viewed as a natural choice in the light of Proposition 2.1. The computed examples of the following section show that this approach often yields a better approximation of $\hat{\boldsymbol{x}}$ than any one of the candidate solutions $\boldsymbol{x}_i$.

Table 1
Qualities of the TSVD and Tikhonov solution (both matching the discrepancy principle), the quasi-optimal solution, and the linear combination technique, for $n = 100$ examples with 0.1% error and $\eta = 1.1$. The last column shows the ρ-value of the linear combination solution. Each column represents the average over 1000 different error vectors.

Problem	Tikh (di.pr.)	TSVD (di.pr.)	Tikh (quasi)	Lin.comb.	ρ
baart	$1.59 \cdot 10^{-1}$	$1.66 \cdot 10^{-1}$	$1.43 \cdot 10^{-1}$	$1.34 \cdot 10^{-1}$	−0.44
deriv2-1	$1.87 \cdot 10^{-1}$	$2.05 \cdot 10^{-1}$	$1.92 \cdot 10^{-1}$	$1.73 \cdot 10^{-1}$	−0.82
deriv2-2	$1.80 \cdot 10^{-1}$	$1.96 \cdot 10^{-1}$	$1.85 \cdot 10^{-1}$	$1.66 \cdot 10^{-1}$	−0.91
deriv2-3	$1.94 \cdot 10^{-2}$	$2.51 \cdot 10^{-2}$	$1.88 \cdot 10^{-2}$	$1.87 \cdot 10^{-2}$	−0.0046
foxgood	$2.26 \cdot 10^{-2}$	$3.11 \cdot 10^{-2}$	$1.86 \cdot 10^{-2}$	$1.29 \cdot 10^{-2}$	−0.45
gravity	$2.06 \cdot 10^{-2}$	$2.75 \cdot 10^{-2}$	$1.78 \cdot 10^{-2}$	$1.65 \cdot 10^{-2}$	−0.13
heat	$4.62 \cdot 10^{-2}$	$5.84 \cdot 10^{-2}$	$4.31 \cdot 10^{-2}$	$4.44 \cdot 10^{-2}$	0.089
ilaplace	$1.20 \cdot 10^{-1}$	$1.26 \cdot 10^{-1}$	$1.10 \cdot 10^{-1}$	$1.06 \cdot 10^{-1}$	−0.26
phillips	$1.36 \cdot 10^{-2}$	$1.90 \cdot 10^{-2}$	$1.60 \cdot 10^{-2}$	$1.23 \cdot 10^{-2}$	−0.26
shaw	$6.33 \cdot 10^{-2}$	$4.91 \cdot 10^{-2}$	$5.66 \cdot 10^{-2}$	$5.81 \cdot 10^{-2}$	0.64

Table 2
Same as Table 1, but now with 1% error, $\eta = 1.1$.

Problem	Tikh (di.pr.)	TSVD (di.pr)	Tikh (quasi)	Lin.comb.	ρ
baart	$2.21 \cdot 10^{-1}$	$1.69 \cdot 10^{-1}$	$1.74 \cdot 10^{-1}$	$1.72 \cdot 10^{-1}$	0.046
deriv2-1	$2.87 \cdot 10^{-1}$	$3.10 \cdot 10^{-1}$	$7.34 \cdot 10^{-1}$	$2.92 \cdot 10^{-1}$	0.011
deriv2-2	$2.77 \cdot 10^{-1}$	$2.99 \cdot 10^{-1}$	$3.77 \cdot 10^{-1}$	$2.57 \cdot 10^{-1}$	−0.19
deriv2-3	$4.89 \cdot 10^{-2}$	$4.89 \cdot 10^{-2}$	$4.46 \cdot 10^{-2}$	$4.69 \cdot 10^{-2}$	0.53
foxgood	$4.69 \cdot 10^{-2}$	$3.21 \cdot 10^{-2}$	$3.17 \cdot 10^{-2}$	$3.73 \cdot 10^{-2}$	0.37
gravity	$4.54 \cdot 10^{-2}$	$6.15 \cdot 10^{-2}$	$4.02 \cdot 10^{-2}$	$3.80 \cdot 10^{-2}$	−0.1
heat	$1.40 \cdot 10^{-1}$	$1.71 \cdot 10^{-1}$	$1.23 \cdot 10^{-1}$	$1.20 \cdot 10^{-1}$	−0.072
ilaplace	$1.59 \cdot 10^{-1}$	$1.67 \cdot 10^{-1}$	$1.48 \cdot 10^{-1}$	$1.42 \cdot 10^{-1}$	−0.3
phillips	$2.98 \cdot 10^{-2}$	$2.58 \cdot 10^{-2}$	$2.87 \cdot 10^{-2}$	$4.02 \cdot 10^{-2}$	3.6
shaw	$1.55 \cdot 10^{-1}$	$1.70 \cdot 10^{-1}$	$1.51 \cdot 10^{-1}$	$1.28 \cdot 10^{-1}$	−1.1

4. Numerical examples

Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p$ be approximate solutions of (1) and define the qualities (relative errors)

$$q_i = \frac{\|\hat{\boldsymbol{x}} - \boldsymbol{x}_i\|}{\|\hat{\boldsymbol{x}}\|}, \quad i = 1, 2, \ldots, p.$$

Without loss of generality, we order the approximations according to increasing quality,

$$q_1 \le q_2 \le \cdots \le q_p.$$

Let $\widetilde{q}$ denote the relative error of the approximate solution $\widetilde{\boldsymbol{x}}$ defined by (6). We define the following indicator of the quality of $\widetilde{\boldsymbol{x}}$,

$$\rho = \frac{\widetilde{q} - q_1}{q_p - q_1}.$$

The parameter ρ is a convenient measure with:

- $\rho < 0$ indicating that $\widetilde{\boldsymbol{x}}$ is a better approximation of $\hat{\boldsymbol{x}}$ than any one of the approximate solutions $\boldsymbol{x}_i$, $i = 1, 2, \ldots, p$;
- $\rho = 0$ indicating that $\widetilde{\boldsymbol{x}}$ approximates $\hat{\boldsymbol{x}}$ as accurately as the best of the approximate solutions $\boldsymbol{x}_i$;
- $\rho = 1$ indicating that $\widetilde{\boldsymbol{x}}$ approximates $\hat{\boldsymbol{x}}$ as well as the worst of the approximate solutions $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p$;
- $\rho > 1$ indicating that all of the approximate solutions $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p$ approximate $\hat{\boldsymbol{x}}$ more accurately than $\widetilde{\boldsymbol{x}}$.

Table 1 presents results for several test examples from [11] of dimension $n = 100$ with 0.1% error in $\boldsymbol{b}$. The search space $\mathcal{W}$ is spanned by three standard approximate solutions of (1) computed by TSVD (15) using the discrepancy principle to determine the truncation index, by Tikhonov regularization using the discrepancy principle to determine the regularization parameter, and by Tikhonov regularization using the quasi-optimal criterion to define the regularization parameter; see Section 3. Once the SVD has been computed, the solutions obtained with these three methods can be evaluated quite rapidly. However, we remark that any space spanned by approximate solutions may be used.

Each column of Table 1 represents the average over 1000 different error vectors $\boldsymbol{e}$ with normally distributed random entries with zero mean. The table shows that, on average the approximate solution $\widetilde{\boldsymbol{x}}$ often yields better approximations of $\hat{\boldsymbol{x}}$ than the best of the approximate solutions determined by the original three methods. For two of the problems, the quality of $\widetilde{\boldsymbol{x}}$ is between the best and the worst of first computed approximate solutions.

For the results of Tables 2–6, we vary the error level and the parameter η in (14). We see that the linear combination approach frequently gives a new approximate solution that improves on the three basis solutions, in particular for low error levels and/or larger η-values.

Table 3
Same as Table 1, but now with 10% error, $\eta = 1.1$.

Problem	Tikh (di.pr.)	TSVD (di.pr.)	Tikh (quasi)	Lin.comb.	ρ
baart	$3.76 \cdot 10^{-1}$	$3.47 \cdot 10^{-1}$	$3.20 \cdot 10^{-1}$	$2.88 \cdot 10^{-1}$	-0.57
deriv2-1	$4.45 \cdot 10^{-1}$	$4.75 \cdot 10^{-1}$	$9.37 \cdot 10^{-1}$	$4.05 \cdot 10^{-1}$	-0.081
deriv2-2	$4.32 \cdot 10^{-1}$	$4.55 \cdot 10^{-1}$	$8.61 \cdot 10^{-1}$	$3.90 \cdot 10^{-1}$	-0.098
deriv2-3	$1.14 \cdot 10^{-1}$	$1.21 \cdot 10^{-1}$	$1.05 \cdot 10^{-1}$	$1.11 \cdot 10^{-1}$	0.4
foxgood	$2.22 \cdot 10^{-1}$	$2.76 \cdot 10^{-1}$	$8.83 \cdot 10^{-2}$	$9.35 \cdot 10^{-2}$	0.028
gravity	$1.25 \cdot 10^{-1}$	$1.66 \cdot 10^{-1}$	$1.65 \cdot 10^{-1}$	$1.65 \cdot 10^{-1}$	0.98
heat	$4.10 \cdot 10^{-1}$	$4.37 \cdot 10^{-1}$	$5.59 \cdot 10^{-1}$	$3.42 \cdot 10^{-1}$	-0.46
ilaplace	$2.22 \cdot 10^{-1}$	$2.36 \cdot 10^{-1}$	$2.04 \cdot 10^{-1}$	$1.97 \cdot 10^{-1}$	-0.21
phillips	$1.08 \cdot 10^{-1}$	$1.13 \cdot 10^{-1}$	$1.40 \cdot 10^{-1}$	$1.40 \cdot 10^{-1}$	1
shaw	$2.31 \cdot 10^{-1}$	$2.73 \cdot 10^{-1}$	$1.84 \cdot 10^{-1}$	$1.83 \cdot 10^{-1}$	-0.014

Table 4
Same as Table 1, but now with 0.1% error, $\eta = 1.2$.

Problem	Tikh (di.pr.)	TSVD (di.pr.)	Tikh (quasi)	Lin.comb.	ρ
baart	$1.62 \cdot 10^{-1}$	$1.66 \cdot 10^{-1}$	$1.43 \cdot 10^{-1}$	$1.34 \cdot 10^{-1}$	-0.44
deriv2-1	$1.94 \cdot 10^{-1}$	$2.14 \cdot 10^{-1}$	$1.92 \cdot 10^{-1}$	$1.76 \cdot 10^{-1}$	-0.74
deriv2-2	$1.87 \cdot 10^{-1}$	$2.05 \cdot 10^{-1}$	$1.85 \cdot 10^{-1}$	$1.69 \cdot 10^{-1}$	-0.76
deriv2-3	$2.10 \cdot 10^{-2}$	$2.69 \cdot 10^{-2}$	$1.88 \cdot 10^{-2}$	$1.78 \cdot 10^{-2}$	-0.12
foxgood	$2.50 \cdot 10^{-2}$	$3.11 \cdot 10^{-2}$	$1.86 \cdot 10^{-2}$	$1.29 \cdot 10^{-2}$	-0.46
gravity	$2.25 \cdot 10^{-2}$	$2.83 \cdot 10^{-2}$	$1.78 \cdot 10^{-2}$	$1.65 \cdot 10^{-2}$	-0.12
heat	$4.90 \cdot 10^{-2}$	$6.48 \cdot 10^{-2}$	$4.31 \cdot 10^{-2}$	$4.32 \cdot 10^{-2}$	0.008
ilaplace	$1.25 \cdot 10^{-1}$	$1.27 \cdot 10^{-1}$	$1.10 \cdot 10^{-1}$	$1.06 \cdot 10^{-1}$	-0.24
phillips	$1.48 \cdot 10^{-2}$	$2.45 \cdot 10^{-2}$	$1.60 \cdot 10^{-2}$	$1.15 \cdot 10^{-2}$	-0.33
shaw	$7.23 \cdot 10^{-2}$	$4.91 \cdot 10^{-2}$	$5.66 \cdot 10^{-2}$	$5.77 \cdot 10^{-2}$	0.37

Table 5
Same as Table 1, but now with 0.1% error, $\eta = 1.5$.

Problem	Tikh (di.pr.)	TSVD (di.pr.)	Tikh (quasi)	Lin.comb.	ρ
baart	$1.67 \cdot 10^{-1}$	$1.66 \cdot 10^{-1}$	$1.43 \cdot 10^{-1}$	$1.34 \cdot 10^{-1}$	-0.42
deriv2-1	$2.10 \cdot 10^{-1}$	$2.32 \cdot 10^{-1}$	$1.92 \cdot 10^{-1}$	$1.84 \cdot 10^{-1}$	-0.19
deriv2-2	$2.03 \cdot 10^{-1}$	$2.23 \cdot 10^{-1}$	$1.85 \cdot 10^{-1}$	$1.77 \cdot 10^{-1}$	-0.2
deriv2-3	$2.48 \cdot 10^{-2}$	$2.69 \cdot 10^{-2}$	$1.88 \cdot 10^{-2}$	$1.76 \cdot 10^{-2}$	-0.15
foxgood	$2.81 \cdot 10^{-2}$	$3.11 \cdot 10^{-2}$	$1.86 \cdot 10^{-2}$	$1.29 \cdot 10^{-2}$	-0.46
gravity	$2.59 \cdot 10^{-2}$	$3.61 \cdot 10^{-2}$	$1.78 \cdot 10^{-2}$	$1.73 \cdot 10^{-2}$	-0.026
heat	$5.70 \cdot 10^{-2}$	$8.52 \cdot 10^{-2}$	$4.31 \cdot 10^{-2}$	$4.25 \cdot 10^{-2}$	-0.015
ilaplace	$1.33 \cdot 10^{-1}$	$1.41 \cdot 10^{-1}$	$1.10 \cdot 10^{-1}$	$1.09 \cdot 10^{-1}$	-0.035
phillips	$1.70 \cdot 10^{-2}$	$2.47 \cdot 10^{-2}$	$1.60 \cdot 10^{-2}$	$1.13 \cdot 10^{-2}$	-0.53
shaw	$9.38 \cdot 10^{-2}$	$1.01 \cdot 10^{-1}$	$5.66 \cdot 10^{-2}$	$5.65 \cdot 10^{-2}$	-0.001

Table 6
Same as Table 1, but now with 0.1% error, $\eta = 2$.

Problem	Tikh (di.pr.)	TSVD (di.pr)	Tikh (quasi)	Lin.comb.	ρ
baart	$1.73 \cdot 10^{-1}$	$1.66 \cdot 10^{-1}$	$1.43 \cdot 10^{-1}$	$1.34 \cdot 10^{-1}$	-0.33
deriv2-1	$2.28 \cdot 10^{-1}$	$2.52 \cdot 10^{-1}$	$1.92 \cdot 10^{-1}$	$1.91 \cdot 10^{-1}$	-0.019
deriv2-2	$2.20 \cdot 10^{-1}$	$2.33 \cdot 10^{-1}$	$1.85 \cdot 10^{-1}$	$1.81 \cdot 10^{-1}$	-0.066
deriv2-3	$2.95 \cdot 10^{-2}$	$4.39 \cdot 10^{-2}$	$1.88 \cdot 10^{-2}$	$1.86 \cdot 10^{-2}$	-0.0075
foxgood	$3.11 \cdot 10^{-2}$	$3.11 \cdot 10^{-2}$	$1.86 \cdot 10^{-2}$	$1.29 \cdot 10^{-2}$	-0.46
gravity	$2.97 \cdot 10^{-2}$	$4.00 \cdot 10^{-2}$	$1.78 \cdot 10^{-2}$	$1.76 \cdot 10^{-2}$	-0.008
heat	$6.84 \cdot 10^{-2}$	$9.62 \cdot 10^{-2}$	$4.31 \cdot 10^{-2}$	$4.29 \cdot 10^{-2}$	-0.0036
ilaplace	$1.39 \cdot 10^{-1}$	$1.45 \cdot 10^{-1}$	$1.10 \cdot 10^{-1}$	$1.10 \cdot 10^{-1}$	-0.014
phillips	$1.93 \cdot 10^{-2}$	$2.47 \cdot 10^{-2}$	$1.60 \cdot 10^{-2}$	$1.13 \cdot 10^{-2}$	-0.54
shaw	$1.18 \cdot 10^{-1}$	$1.23 \cdot 10^{-1}$	$5.66 \cdot 10^{-2}$	$5.63 \cdot 10^{-2}$	-0.0041

5. Conclusion

The evaluation of several approximate solutions $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p$ of (1) is inexpensive when the SVD of the matrix A is available. The computed examples illustrate that the "linear combination" approximate solution extracted from $\mathcal{W} = \text{span}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p\}$ in many cases furnishes a better approximation of the desired solution $\hat{\boldsymbol{x}}$ of the unavailable error-free system than any of the approximate solutions $\boldsymbol{x}_i$. The proposed scheme provides an inexpensive approach to determine an improved solution from a set of available approximate solutions.

Large-scale problems can be treated by first projecting them, e.g., by a Krylov subspace method, to a problem of small size and then proceeding as described in the present paper to obtain several approximate solutions of this small problem. A new solution can be extracted as described in Sections 2 and 3, and then be projected back into the high-dimensional solution (sub)space. This yields an approximate solution of the original (large) problem. Finally, we note that approximate solutions of (1) also can be determined by methods that do not require the evaluation of an SVD.

References

[1] H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Kluwer, Dordrecht, 1996.
[2] P.C. Hansen, Rank-Deficient and Discrete Ill-Posed Problems, SIAM, Philadelphia, 1998.
[3] D. Calvetti, L. Reichel, Tikhonov regularization with a solution constraint, SIAM J. Sci. Comput. 26 (2004) 224–239.
[4] C.W. Groetsch, The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind, Pitman, Boston, 1984.
[5] F. Bauer, S. Kindermann, Recent results on the quasi-optimality principle, J. Inverse Ill-Posed Probl. 17 (2009) 5–18.
[6] D. Calvetti, B. Lewis, L. Reichel, F. Sgallari, Tikhonov regularization with nonnegativity constraint, Electron. Trans. Numer. Anal. 18 (2004) 153–173.
[7] M.E. Hochstenbach, L. Reichel, Subspace-restricted singular value decompositions for linear discrete ill-posed problems, J. Comput. Appl. Math. 235 (2010) 1053–1064.
[8] M.E. Hochstenbach, N. McNinch, L. Reichel, Discrete ill-posed least-squares problems with a solution norm constraint, Lin. Alg. Appl. (in press).
[9] J. Lampe, H. Voss, A fast algorithm for solving regularized total least squares problems, Electron. Trans. Numer. Anal. 31 (2008) 12–24.
[10] S. Morigi, L. Reichel, F. Sgallari, A truncated projected SVD method for linear discrete ill-posed problems, Numer. Algorithms 43 (2006) 197–213.
[11] P.C. Hansen, Regularization tools version 4.0 for Matlab 7.3, Numer. Algorithms 46 (2007) 189–194.