

A new Tikhonov regularization method

Martin Fuhry · Lothar Reichel

Received: 27 April 2011 / Accepted: 17 August 2011 /
Published online: 1 September 2011
© Springer Science+Business Media, LLC 2011

Abstract The numerical solution of linear discrete ill-posed problems typically requires regularization, i.e., replacement of the available ill-conditioned problem by a nearby better conditioned one. The most popular regularization methods for problems of small to moderate size, which allow evaluation of the singular value decomposition of the matrix defining the problem, are the truncated singular value decomposition and Tikhonov regularization. The present paper proposes a novel choice of regularization matrix for Tikhonov regularization that bridges the gap between Tikhonov regularization and truncated singular value decomposition. Computed examples illustrate the benefit of the proposed method.

Keywords Ill-posed problem · Tikhonov regularization · Truncated singular value decomposition · Regularization matrix

1 Introduction

Consider the computation of an approximate solution of the minimization problem

$$\min_{x \in \mathbb{R}^n} \|Ax - \mathbf{b}\|, \quad (1.1)$$

where $\|\cdot\|$ denotes the Euclidean vector norm and $A \in \mathbb{R}^{m \times n}$ is a matrix with many singular values of different sizes close to the origin. Minimization

M. Fuhry · L. Reichel (✉)
Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA
e-mail: reichel@math.kent.edu

M. Fuhry
e-mail: martyfuhry@gmail.com

problems (1.1) with a matrix of this kind often are referred to as discrete ill-posed problems. They arise, for example, from the discretization of linear ill-posed problems, such as Fredholm integral equations of the first kind with a smooth kernel. The vector $\mathbf{b} \in \mathbb{R}^m$ in (1.1) represents error-contaminated data. We will for notational simplicity assume that $m \geq n$; however, the methods discussed also can be applied when $m < n$.

Let $\mathbf{e} \in \mathbb{R}^m$ denote the (unknown) error in \mathbf{b} , and let $\hat{\mathbf{b}} \in \mathbb{R}^m$ be the (unknown) error-free vector associated with \mathbf{b} , i.e.,

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

We sometimes will refer to the vector \mathbf{e} as “noise.” The (unavailable) linear system of equations with error-free right-hand side,

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

is assumed to be consistent; however, we do not require the least-squares problem with error-contaminated data \mathbf{b} (1.1) to be consistent.

Let A^\dagger denote the Moore-Penrose pseudoinverse of A . We are interested in computing an approximation of the solution $\hat{\mathbf{x}} = A^\dagger \hat{\mathbf{b}}$ of minimal Euclidean norm of the error-free linear system (1.3) by determining an approximate solution of the error-contaminated least-squares problem (1.1). Note that the solution of (1.1),

$$\check{\mathbf{x}} = A^\dagger \mathbf{b} = A^\dagger (\hat{\mathbf{b}} + \mathbf{e}) = \hat{\mathbf{x}} + A^\dagger \mathbf{e}, \quad (1.4)$$

typically is dominated by the propagated error $A^\dagger \mathbf{e}$ and then is meaningless.

Tikhonov regularization seeks to determine a useful approximation of $\hat{\mathbf{x}}$ by replacing the minimization problem (1.1) by a penalized least-squares problem of the form

$$\min_{\mathbf{x} \in \mathbb{R}^n} \{ \|A\mathbf{x} - \mathbf{b}\|^2 + \|L_\mu \mathbf{x}\|^2 \}, \quad (1.5)$$

where the matrix $L_\mu \in \mathbb{R}^{k \times n}$, $k \leq n$, is referred to as the regularization matrix. The scalar $\mu > 0$ is known as the regularization parameter. The matrix L_μ is commonly chosen to be μI , where I denotes the identity matrix; however, if the desired solution $\hat{\mathbf{x}}$ has particular known properties, then it may be meaningful to let L_μ be a scaled finite difference approximation of a differential operator or a scaled orthogonal projection; see, e.g., [1–3, 5, 7, 10, 11] for examples.

Solving (1.5) requires both the determination of a suitable value of $\mu > 0$ and the computation of the associated solution $\mathbf{x} = \mathbf{x}_\mu$ of the minimization problem. We will assume that a bound for the norm of the error-vector \mathbf{e} is known. Then μ can be determined with the aid of the discrepancy principle; see below for details. The use of the discrepancy principle to determine μ generally requires the solution of (1.5) for several values of μ . When $L_\mu = \mu I$ and the singular value decomposition (SVD) of A is available, the desired value of μ can be computed inexpensively by using a zero-finder, such as Newton’s method.

The regularization matrices mentioned above are linear functions of μ . We use the notation L_μ (instead of μL), because we will introduce a regularization matrix that depends on μ in a nonlinear fashion. This regularization matrix may be an attractive substitute for μI when no particular properties of the desired solution \hat{x} are known. Similarly as when the regularization matrix μI is used, the solution of (1.5) with the proposed matrix L_μ easily can be computed when the SVD of A is available. Our new regularization matrix is designed to dampen low frequencies less than the matrix $L_\mu = \mu I$. Numerical examples illustrate the proposed regularization matrix to often yield more accurate approximations x_μ of \hat{x} than the regularization matrix $L_\mu = \mu I$.

Another common regularization method for (1.1) is the truncated SVD (TSVD) method. In this method the smallest singular values of A are set to zero and the minimal-norm solution of the resulting least-squares problem is computed. We determine the truncation index with the discrepancy principle and compare TSVD with Tikhonov regularization.

This paper is organized as follows. Section 2 discusses regularization by the TSVD and Tikhonov methods and introduces our new regularization matrix. Section 3 contains a few computed examples. Concluding remarks and comments on possible extensions can be found in Section 4.

2 Regularization methods

We first introduce the SVD of A and then discuss regularization by the TSVD and Tikhonov methods. The SVD of A is given by

$$A = U \Sigma V^T, \tag{2.1}$$

where $U = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m] \in \mathbb{R}^{m \times m}$ and $V = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] \in \mathbb{R}^{n \times n}$ are orthogonal matrices, and

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

is a (possibly rectangular) diagonal matrix, whose nonnegative diagonal entries σ_j are the singular values of A . They are ordered according to $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$.

Let A be of rank ℓ . Then (2.1) can be expressed as

$$A = \sum_{j=1}^{\ell} \sigma_j \mathbf{u}_j \mathbf{v}_j^T \tag{2.2}$$

with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_\ell > 0$. When the matrix A stems from the discretization of a Fredholm integral equation of the first kind with a smooth kernel, the vectors \mathbf{v}_j and \mathbf{u}_j represent discretizations of singular functions that are defined on the domains of the integral operator and its adjoint, respectively. These singular functions typically oscillate more with increasing index. The representation (2.2) then is a decomposition of A into rank-one matrices $\mathbf{u}_j \mathbf{v}_j^T$

that represent more and more oscillatory components of A with increasing index j .

2.1 Regularization by truncated singular value decomposition

The Moore-Penrose pseudoinverse of A is given by

$$A^\dagger = \sum_{j=1}^{\ell} \sigma_j^{-1} \mathbf{v}_j \mathbf{u}_j^T.$$

The difficulty of solving (1.1) without regularization stems from the fact that the matrix A has “tiny” positive singular values and the computation of the solution (1.4) of (1.1) involves division by these tiny singular values. This results in severe propagation of the error \mathbf{e} in \mathbf{b} and of round-off errors introduced during the calculations into the computed solution of (1.1).

Regularization by the TSVD method overcomes this difficulty by ignoring the tiny singular values of A . Introduce, for $k \leq \ell$, the rank- k approximation of A ,

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

with Moore-Penrose pseudoinverse

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

The TSVD method yields approximate solutions of (1.1) of the form

$$\mathbf{x}_k = A_k^\dagger \mathbf{b} = \sum_{j=1}^k \frac{\mathbf{u}_j^T \mathbf{b}}{\sigma_j} \mathbf{v}_j, \quad k = 1, 2, \dots, \ell. \quad (2.3)$$

It is convenient to use the transformed quantities

$$\tilde{\mathbf{x}}_k = V^T \mathbf{x}_k, \quad \tilde{\mathbf{b}} = [\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_m]^T = U^T \mathbf{b}$$

in the computations. Thus, we compute

$$\tilde{\mathbf{x}}_k = \left[\frac{\tilde{b}_1}{\sigma_1}, \frac{\tilde{b}_2}{\sigma_2}, \dots, \frac{\tilde{b}_k}{\sigma_k}, 0, \dots, 0 \right]^T \quad (2.4)$$

for a suitable value of $1 \leq k \leq \ell$ and then determine the approximate solution $\mathbf{x}_k = V \tilde{\mathbf{x}}_k$ of (1.1).

Assume that a bound for the norm of the error

$$\|\mathbf{e}\| \leq \varepsilon$$

is available. We then can determine the truncation index k by the discrepancy principle, i.e., we choose k as small as possible so that

$$\|A\mathbf{x}_k - \mathbf{b}\| \leq \eta\varepsilon, \tag{2.5}$$

where $\eta > 1$ is a user-specified constant independent of ε . Thus, the truncation index $k = k_\varepsilon$ depends on ε and generally increases as ε decreases. A proof of the convergence of $\mathbf{x}_{k_\varepsilon}$ to $\hat{\mathbf{x}}$ as $\varepsilon \searrow 0$ in a Hilbert space setting is presented in [5]. It requires the constant $\eta > 1$ in (2.5). In actual computations, we use the representation

$$\|A\mathbf{x}_k - \mathbf{b}\|^2 = \sum_{j=k+1}^m \tilde{b}_j^2$$

to determine k_ε from (2.5). Further details on regularization by TSVD can be found in, e.g., [5, 7].

2.2 Tikhonov regularization with $L_\mu = \mu I$

Substituting (2.1), $\tilde{\mathbf{x}} = V^T \mathbf{x}$, $\tilde{\mathbf{b}} = U^T \mathbf{b}$, and $L_\mu = \mu I$ into (1.5) yields the penalized least-squares problem

$$\min_{\tilde{\mathbf{x}} \in \mathbb{R}^m} \{ \|\Sigma \tilde{\mathbf{x}} - \tilde{\mathbf{b}}\|^2 + \mu^2 \|\tilde{\mathbf{x}}\|^2 \}$$

with solution

$$\tilde{\mathbf{x}}_\mu = (\Sigma^T \Sigma + \mu^2 I)^{-1} \Sigma^T \tilde{\mathbf{b}} \tag{2.6}$$

for $\mu > 0$. The solution of (1.5) is given by $\mathbf{x}_\mu = V \tilde{\mathbf{x}}_\mu$. It satisfies

$$(A^T A + \mu^2 I) \mathbf{x}_\mu = A^T \mathbf{b}. \tag{2.7}$$

The discrepancy principle prescribes that the regularization parameter $\mu > 0$ be determined so that

$$\|A\mathbf{x}_\mu - \mathbf{b}\| = \eta\varepsilon, \tag{2.8}$$

or, equivalently, so that

$$\|\Sigma \tilde{\mathbf{x}}_\mu - \tilde{\mathbf{b}}\|^2 = \eta^2 \varepsilon^2, \tag{2.9}$$

where the constant $\eta > 1$ is independent of ε . This nonlinear equation for μ can be solved, e.g., by Newton’s method. Each evaluation of the left-hand side of (2.9) requires at most $\mathcal{O}(m)$ arithmetic floating point operations. The computational effort needed to determine the desired value of μ therefore is negligible compared with the $\mathcal{O}(mn^2)$ arithmetic floating point operations required for the evaluation of the SVD (2.1) of A . Generally, μ decreases with

ε . Proofs of the convergence $\mathbf{x}_\mu \rightarrow \hat{\mathbf{x}}$ as $\varepsilon \searrow 0$ are provided in [5, 6]. The proofs are in Hilbert space settings and require the constant $\eta > 1$ in (2.8).

2.3 Tikhonov regularization with the new regularization matrix

It follows from (2.6) that Tikhonov regularization with $L_\mu = \mu I$ and $\mu > 0$ dampens all components of $\Sigma^T \tilde{\mathbf{b}}$, i.e., all solution components \mathbf{v}_j of \mathbf{x}_μ . On the other hand, TSVD does not dampen any solution component that is not set to zero; cf. (2.4). It is known that Tikhonov regularization may oversmooth the computed solution when the regularization parameter is well determined; see, e.g., [9] for a recent discussion. We propose to choose a regularization matrix L_μ that provides no damping of solution components \mathbf{v}_j with small index.

Introduce

$$L_\mu = D_\mu V^T \tag{2.10}$$

with

$$D_\mu^2 = \text{diag} [\max\{\mu^2 - \sigma_1^2, 0\}, \max\{\mu^2 - \sigma_2^2, 0\}, \dots, \max\{\mu^2 - \sigma_n^2, 0\}].$$

Analogously to (2.6), we obtain

$$\tilde{\mathbf{x}}_\mu = (\Sigma^T \Sigma + D_\mu^2)^{-1} \Sigma^T \tilde{\mathbf{b}}. \tag{2.11}$$

If $\sigma_k > \mu \geq \sigma_{k+1}$, then

$$\Sigma^T \Sigma + D_\mu^2 = \text{diag} [\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2, \mu^2, \dots, \mu^2] \in \mathbb{R}^{n \times n}.$$

In particular, if $\mu > 0$, then the above matrix is positive definite and the solution (2.11) exists. The corresponding approximate solution of (1.1) is given by $\mathbf{x}_\mu = V \tilde{\mathbf{x}}_\mu$ and satisfies

$$(A^T A + L_\mu^T L_\mu) \mathbf{x}_\mu = A^T \mathbf{b}; \tag{2.12}$$

cf. (2.7). The value of μ used in (2.11) is the same as in Section 2.2.

In order to avoid severe propagation of the error \mathbf{e} in \mathbf{b} into the computed approximate solution \mathbf{x}_μ , the smallest eigenvalue of the matrix $A^T A + L_\mu^T L_\mu$, which is $\max\{\mu^2, \sigma_n^2\}$, has to be sufficiently large. Moreover, we would like the matrix $L_\mu^T L_\mu$ to be of small norm, because this may help us determine an accurate approximation of $\hat{\mathbf{x}}$. The following result shows the matrix $L_\mu^T L_\mu$ to be optimal in the Frobenius norm, which for a matrix $M \in \mathbb{R}^{n \times n}$ is given by $\|M\|_F = \sqrt{\text{trace}(M^T M)}$.

Theorem 2.1 *Let $M \in \mathbb{R}^{n \times n}$ be a symmetric matrix with spectral factorization $M = V \Lambda V^T$, where $V \in \mathbb{R}^{n \times n}$ is orthonormal and $\Lambda = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_n]$.*

Assume that $\mu \geq \min_{1 \leq j \leq n} \lambda_j$. Let the diagonal matrix $C_\mu = \text{diag}[c_1, c_2, \dots, c_n] \in \mathbb{R}^{n \times n}$ have the entries

$$c_j = \max\{\mu - \lambda_j, 0\}, \quad j = 1, 2, \dots, n.$$

Then the matrix $M + VC_\mu V^T$ has the smallest eigenvalue μ and the distance in the Frobenius norm between M and the closest symmetric matrix with the smallest eigenvalue μ is $\|C_\mu\|_F$.

Proof Let $N \in \mathbb{R}^{n \times n}$ be a symmetric matrix such that all eigenvalues of $H = M + N$ are larger than or equal to μ . Let the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of M and $\gamma_1, \gamma_2, \dots, \gamma_n$ of H be arranged in nondecreasing order. Then by the Wielandt-Hoffman theorem, see, e.g., [12, pp. 104–108], we have that

$$\|N\|_F^2 = \|H - M\|_F^2 \geq \sum_{j=1}^n (\gamma_j - \lambda_j)^2.$$

The right-hand side is minimal when $\gamma_j = \max\{\lambda_j, \mu\}$ for all j . Therefore

$$\|N\|_F^2 \geq \sum_{\lambda_j < \mu} (\mu - \lambda_j)^2.$$

The theorem now follows from the observation that

$$\|VC_\mu V^T\|_F^2 = \|C_\mu\|_F^2 = \sum_{\lambda_j < \mu} (\mu - \lambda_j)^2.$$

□

Corollary 2.2 Let L_μ be defined by (2.10) and assume that μ and σ_1 are strictly positive. Then

$$\|L_\mu\|_F^2 < \|\mu I\|_F^2.$$

Proof We have

$$\|L_\mu\|_F^2 = \|D_\mu\|_F^2 = \sum_{\sigma_j^2 < \mu^2} (\mu^2 - \sigma_j^2) < n\mu^2 = \|\mu I\|_F^2.$$

□

We recall that for Tikhonov regularization with the regularization matrices $L_\mu = \mu I$ or (2.10), the value of the regularization parameter μ is determined by the discrepancy principle based on the regularization matrix $L_\mu = \mu I$ as described in Section 2.2. By Corollary 2.2, the matrix in the regularized normal equations (2.7) with $L_\mu = \mu I$ differs more from $A^T A$ than the matrix in the regularized normal equations (2.12) with the regularization matrix (2.10). This suggests that the solution of (2.12) may be a better approximation of the desired solution \hat{x} than the solution of (2.7). The numerical examples in Section 3 show this, indeed, often to be the case.

It is informative to consider the filter factors for the methods discussed. Investigations of regularization methods with the aid of filter factors can be found in Hansen [7] for TSVD and Tikhonov regularization with $L_\mu = \mu I$. Donatelli and Serra-Capizzano [4] use filter factors to study multilevel methods. The unregularized solution $\mathbf{x} = A^\dagger \mathbf{b}$ of (1.1) can be expressed as

$$\mathbf{x} = \sum_{j=1}^{\ell} \frac{\mathbf{u}_j^T \mathbf{b}}{\sigma_j} \mathbf{v}_j.$$

The filter factors show how the components are modified by the regularization method used. For instance, for the TSVD solution \mathbf{x}_k given by (2.3), we have

$$\mathbf{x}_k = \sum_{j=1}^{\ell} \varphi_{k,j}^{(\text{TSVD})} \frac{\mathbf{u}_j^T \mathbf{b}}{\sigma_j} \mathbf{v}_j$$

with the filter factors

$$\varphi_{k,j}^{(\text{TSVD})} = \begin{cases} 1, & 1 \leq j \leq k, \\ 0, & k < j \leq \ell. \end{cases}$$

Similarly, the Tikhonov solution $\mathbf{x}_\mu = V \tilde{\mathbf{x}}_\mu$ defined by (2.6) with $L_\mu = \mu I$ can be written as

$$\mathbf{x}_\mu = \sum_{j=1}^{\ell} \varphi_{\mu,j}^{(\text{Tikhonov})} \frac{\mathbf{u}_j^T \mathbf{b}}{\sigma_j} \mathbf{v}_j$$

with the filter factors

$$\varphi_{\mu,j}^{(\text{Tikhonov})} = \frac{\sigma_j^2}{\sigma_j^2 + \mu^2}, \quad 1 \leq j \leq \ell.$$

Assume that $\sigma_k > \mu \geq \sigma_{k+1}$ and let $\mu > 0$; if $k = n$, then we define $\sigma_{n+1} = 0$. Our new Tikhonov regularization method with L_μ given by (2.10) can be expressed as

$$\mathbf{x}_\mu = \sum_{j=1}^{\ell} \varphi_{\mu,j}^{(\text{new})} \frac{\mathbf{u}_j^T \mathbf{b}}{\sigma_j} \mathbf{v}_j$$

with the filter factors

$$\varphi_{\mu,j}^{(\text{new})} = \begin{cases} 1, & 1 \leq j \leq k, \\ \frac{\sigma_j^2}{\mu^2}, & k < j \leq \ell. \end{cases}$$

Thus, these filter factors are the same as $\varphi_{k,j}^{(\text{TSVD})}$ for $1 \leq j \leq k$, and close to $\varphi_{\mu,j}^{(\text{Tikhonov})}$ for $k < j \leq \ell$.

3 Computed examples

The calculations of this section were carried out using the Python programming language with the open source numerical computation modules NumPy and SciPy. Floating point arithmetic was done with 64 bits, i.e., machine epsilon was about $2.2 \cdot 10^{-16}$. The computed examples are taken from the MATLAB package Regularization Tools [8] and were imported into Python using Pytave, an open source Python module wrapper for Octave.

All examples are obtained by discretizing Fredholm integral equations of the first kind

$$\int_a^b \kappa(s, t)x(t) dt = g(s), \quad c \leq s \leq d, \tag{3.1}$$

with a smooth kernel κ . The discretizations are carried out by Galerkin or Nyström methods and yield linear discrete ill-posed problems (1.1). MATLAB functions in [8] determine discretizations $A \in \mathbb{R}^{n \times n}$ of the integral operators and scaled discrete approximations $\hat{x} \in \mathbb{R}^n$ of the solution x of (3.1). We add an error vector $e \in \mathbb{R}^n$ with normally distributed random entries with zero mean to $\hat{b} = A\hat{x}$ to obtain the vector b in (1.1); cf. (1.2). The vector e is scaled to yield a specified noise level $\|e\|/\|\hat{b}\|$. In particular, $\|e\|$ is available and we can apply the discrepancy principle with $\varepsilon = \|e\|$ to determine the regularization parameter μ in Tikhonov regularization and the truncation index k in TSVD. We let $\eta = 1$ in (2.5) and (2.9) in the computed examples.

Let x_{comp} denote the computed solution using TSVD or Tikhonov regularization with L_μ defined by (2.10) or with $L_\mu = \mu I$. We are interested in the relative error $\|x_{\text{comp}} - \hat{x}\|/\|\hat{x}\|$ determined by these regularization methods. This error depends on the entries of the error vector e . To gain insight into the average behavior of the solution methods, we report in every example the average of the relative errors in x_{comp} over 1,000 runs for each noise level. We let $n = 200$ in all examples.

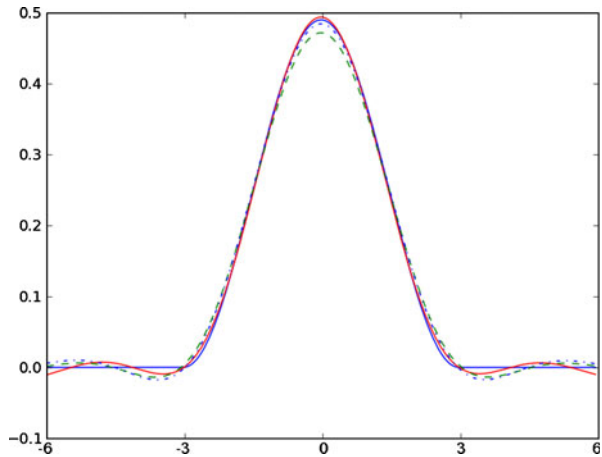
Example 3.1 We first consider the problem phillips from [8]. Let

$$\phi(t) = \begin{cases} 1 + \cos(\frac{\pi t}{3}), & |t| < 3, \\ 0, & |t| \geq 3. \end{cases}$$

The kernel, right-hand side function, and solution of the integral equation (3.1) are given by

$$\begin{aligned} \kappa(s, t) &= \phi(s - t), \quad x(t) = \phi(t), \\ g(s) &= (6 - |s|) \left(1 + \frac{1}{2} \cos\left(\frac{\pi s}{3}\right) \right) + \frac{9}{2\pi} \sin\left(\frac{\pi |s|}{3}\right). \end{aligned}$$

Fig. 1 Example 3.1: the computed solutions determined with L_μ defined by (2.10) (solid red graph), with $L_\mu = \mu I$ (dashed green graph), and with TSVD (dash-dotted blue graph). The vector \mathbf{b} is contaminated by 10% noise. The solid blue graph, which is flat near the ends, displays $\hat{\mathbf{x}}$



and $a = c = -6, b = d = 6$. Figure 1 shows computed solutions determined by TSVD and Tikhonov regularization using the regularization matrices $L_\mu = \mu I$ and (2.10). Table 1 displays the averages of the relative errors in the computed solutions over 1,000 runs for each noise level. Tikhonov regularization with the regularization matrix (2.10) is seen to yield the smallest average errors for all noise levels considered.

Example 3.2 The test problem shaw from [8] is an integral equation (3.1) with

$$\kappa(s, t) = (\cos(s) + \cot(t))^2 \left(\frac{\sin(u)}{u} \right)^2, \quad u = \pi(\sin(s) + \sin(t)),$$

$$x(t) = 2 \exp \left(-6 \left(t - \frac{4}{5} \right)^2 \right) + \exp \left(-2 \left(t + \frac{1}{2} \right)^2 \right).$$

and $a = c = -\pi/2, b = d = \pi/2$. Table 2 shows the average relative errors in the computed solutions over 1,000 runs for each noise level.

Table 1 Example 3.1: average relative errors in the computed solutions for the phillips test problem for several noise levels

Noise level %	Tikhonov regularization		TSVD
	L defined by (2.10)	$L = \mu I$	
10.0	$2.39 \cdot 10^{-2}$	$5.12 \cdot 10^{-2}$	$4.27 \cdot 10^{-2}$
5.0	$2.29 \cdot 10^{-2}$	$3.52 \cdot 10^{-2}$	$2.49 \cdot 10^{-2}$
1.0	$1.69 \cdot 10^{-2}$	$2.00 \cdot 10^{-2}$	$2.41 \cdot 10^{-2}$
0.1	$6.04 \cdot 10^{-3}$	$8.75 \cdot 10^{-3}$	$9.75 \cdot 10^{-3}$

Table 2 Example 3.2: average relative errors in the computed solutions for the shaw test problem for several noise levels

Noise level %	Tikhonov regularization		TSVD
	L defined by (2.10)	$L = \mu I$	
10.0	$1.60 \cdot 10^{-1}$	$1.70 \cdot 10^{-1}$	$1.60 \cdot 10^{-1}$
5.0	$1.51 \cdot 10^{-1}$	$1.56 \cdot 10^{-1}$	$1.51 \cdot 10^{-1}$
1.0	$8.43 \cdot 10^{-2}$	$1.10 \cdot 10^{-1}$	$8.69 \cdot 10^{-2}$
0.1	$4.68 \cdot 10^{-2}$	$4.91 \cdot 10^{-2}$	$4.76 \cdot 10^{-2}$

Tikhonov regularization with L_μ defined by (2.10) yields the smallest relative errors.

Example 3.3 We consider the problem ilaplace from [8], which is a discretization of an inverse Laplace transform with

$$\kappa(s, t) = \exp(-st), \quad x(t) = \exp(-t/2), \quad g(s) = \frac{2}{2s + 1},$$

and $a = c = 0, b = d = \infty$. Table 3 displays the average relative errors in the computed solutions for each noise level. Tikhonov regularization with L_μ defined by (2.10) yields the smallest relative errors over 1,000 runs for each noise level.

Example 3.4 This test problem uses the code deriv2 in [8]. The kernel, solution, and right-hand side of (3.1) are given by

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

$$x(t) = t,$$

$$g(s) = \frac{s^3 - s}{6},$$

and $a = c = 0, b = d = 1$. Thus, the kernel k is the Green’s function for the second derivative. Table 4 shows the average relative errors in the

Table 3 Example 3.3: average relative errors in the computed solutions for the ilaplace test problem for several noise levels

Noise level %	Tikhonov regularization		TSVD
	L defined by (2.10)	$L = \mu I$	
10.0	$2.04 \cdot 10^{-1}$	$2.15 \cdot 10^{-1}$	$2.14 \cdot 10^{-1}$
5.0	$1.92 \cdot 10^{-1}$	$2.02 \cdot 10^{-1}$	$1.99 \cdot 10^{-1}$
1.0	$1.72 \cdot 10^{-1}$	$1.78 \cdot 10^{-1}$	$1.76 \cdot 10^{-1}$
0.1	$1.46 \cdot 10^{-1}$	$1.50 \cdot 10^{-1}$	$1.48 \cdot 10^{-1}$

Table 4 Example 3.4: average relative errors in the computed solutions for the deriv2 test problem for several errors

Noise level %	Tikhonov regularization		TSVD
	L defined by (2.10)	$L = \mu I$	
10.0	$3.16 \cdot 10^{-1}$	$3.47 \cdot 10^{-1}$	$3.30 \cdot 10^{-1}$
5.0	$2.84 \cdot 10^{-1}$	$3.10 \cdot 10^{-1}$	$3.01 \cdot 10^{-1}$
1.0	$2.19 \cdot 10^{-1}$	$2.39 \cdot 10^{-1}$	$2.42 \cdot 10^{-1}$
0.1	$1.51 \cdot 10^{-1}$	$1.64 \cdot 10^{-1}$	$1.72 \cdot 10^{-1}$

computed solutions over 1,000 runs for several noise levels. Tikhonov regularization with the regularization matrix (2.10) gives the smallest average errors.

Example 3.5 In the above examples the regularization parameter μ for Tikhonov regularization and the truncation index k for TSVD are determined with the aid of the discrepancy principle. The present example compares the performance of the methods when the optimal values of μ and the truncation index are used, i.e., we use the values that give the most accurate approximations of $\hat{\mathbf{x}}$. These values of μ and k generally are not available when solving discrete ill-posed problems. This example illustrates that the superior performance of Tikhonov regularization with the regularization matrix (2.10) in Examples 3.1–3.4 does not depend on that the discrepancy principle was used to compute μ and the truncation index k in these examples. Table 5 shows average relative errors in the computed solutions for Tikhonov regularization using the regularization matrices $L_\mu = \mu I$ and (2.10), as well as for TSVD, for the test problem of Example 3.1. Tikhonov regularization with the regularization matrix (2.10) is seen to yield the smallest average relative errors for all noise levels.

The above example suggests that Tikhonov regularization with the regularization matrix (2.10) may be attractive also when the regularization parameter μ is determined by methods other than the discrepancy principle, such as by extrapolation [1–3], generalized cross validation, or the L-curve [7].

Table 5 Example 3.5: average relative errors in the computed solutions for the phillips test problem with optimal regularization parameter μ for several noise levels

Noise level %	Tikhonov regularization		TSVD
	L defined by (2.10)	$L = \mu I$	
10.0	$2.33 \cdot 10^{-2}$	$4.39 \cdot 10^{-2}$	$4.27 \cdot 10^{-2}$
5.0	$2.16 \cdot 10^{-2}$	$3.17 \cdot 10^{-2}$	$2.49 \cdot 10^{-2}$
1.0	$1.57 \cdot 10^{-2}$	$1.92 \cdot 10^{-2}$	$2.39 \cdot 10^{-2}$
0.1	$5.47 \cdot 10^{-3}$	$8.19 \cdot 10^{-3}$	$9.92 \cdot 10^{-3}$

4 Conclusion and extensions

In all examples of Section 3, as well as in many other computed examples, Tikhonov regularization with the regularization matrix (2.10) yields a smaller average error in the computed approximate solutions than Tikhonov regularization with the regularization matrix $L_\mu = \mu I$ and TSVD. The regularization matrix (2.10) therefore can be attractive to use when the SVD of the matrix A in (1.1) can be computed. We remark that the regularization method of the present paper can be applied to penalized least-squares problems (1.5) with a fairly general (linear) regularization matrix L_μ after the least-squares problem has been transformed into standard form. Transformation methods are discussed, e.g., in [7, Sections 2.3.1 and 2.3.2] and [11]. The regularization method can be applied to large-scale problems after these have been reduced to small or moderate size by a Krylov subspace method.

Acknowledgement We would like to thank a referee for suggestions that improved the presentation.

References

1. Brezinski, C., Redivo-Zaglia, M., Rodriguez, G., Seatzu, S.: Extrapolation techniques for ill-conditioned linear systems. *Numer. Math.* **81**, 1–29 (1998)
2. Brezinski, C., Rodriguez, G., Seatzu, S.: Error estimates for linear systems with applications to regularization. *Numer. Algor.* **49**, 85–104 (2008)
3. Brezinski, C., Rodriguez, G., Seatzu, S.: Error estimates for the regularization of least squares problems. *Numer. Algor.* **51**, 61–76 (2009)
4. Donatelli, M., Serra-Capizzano, S.: Filter factor analysis of an iterative multilevel regularization method. *Electron. Trans. Numer. Anal.* **29**, 163–177 (2008)
5. Engl, H.W., Hanke, M., Neubauer, A.: *Regularization of Inverse Problems*. Kluwer, Dordrecht (1996)
6. Groetsch, C.W.: *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*. Pitman, Boston (1984)
7. Hansen, P.C.: *Rank-Deficient and Discrete Ill-Posed Problems*. SIAM, Philadelphia (1998)
8. Hansen, P.C.: Regularization tools version 4.0 for Matlab 7.3. *Numer. Algor.* **46**, 189–194 (2007)
9. Klann, E., Ramlau, R.: Regularization by fractional filter methods and data smoothing. *Inverse Probl.* **24**, 025018 (2008)
10. Morigi, S., Reichel, L., Sgallari, F.: A truncated projected SVD method for linear discrete ill-posed problems. *Numer. Algor.* **43**, 197–213 (2006)
11. Reichel, L., Ye, Q.: Simple square smoothing regularization operators. *Electron. Trans. Numer. Anal.* **33**, 63–83 (2009)
12. Wilkinson, J.H.: *The Algebraic Eigenvalue Problem*. Oxford University Press, Oxford (1965)