AN ITERATIVE METHOD FOR TIKHONOV REGULARIZATION WITH A GENERAL LINEAR REGULARIZATION OPERATOR

MICHIEL E. HOCHSTENBACH* AND LOTHAR REICHEL†

Dedicated to Charles W. Groetsch.

Abstract. Tikhonov regularization is one of the most popular approaches to solve discrete ill-posed problems with error-contaminated data. A regularization operator and a suitable value of a regularization parameter have to be chosen. This paper describes an iterative method, based on Golub-Kahan bidiagonalization, for solving large-scale Tikhonov minimization problems with a linear regularization operator of general form. The regularization parameter is determined by the discrepancy principle. Computed examples illustrate the performance of the method.

Key words. Discrete ill-posed problem, iterative method, Tikhonov regularization, general linear regularization operator, discrepancy principle.

AMS subject classifications. 65R30, 65R32, 65F10, 65F22.

1. Introduction. We are concerned with the solution of large minimization problems

$$\min_{x\in\mathbb{C}^n} \|Ax-b\|, \qquad A\in\mathbb{C}^{m\times n}, \qquad b\in\mathbb{C}^m, \tag{1.1}$$

where $\|\cdot\|$ denotes the Euclidean vector norm and the matrix A is assumed to have many singular values of different orders of magnitude close to the origin. In particular, the ratio between the largest and smallest singular values is very large and therefore the solution of (1.1) is very sensitive to perturbations in the vector b. Minimization problems with matrices of this kind arise, for instance, from the discretization of ill-posed problems, such as Fredholm integral equations of the first kind. They are commonly referred to as discrete ill-posed problems. For ease of notation, we will assume that $m \geq n$; however, the method described also may be applied, mutatis mutandis, when $m < n$.

The vector b in discrete ill-posed problems (1.1) that arise in science and engineering represents available data and typically is contaminated by a measurement error e, which we also will refer to as "noise." Thus,

$$b = \hat{b} + e, \tag{1.2}$$

where $\hat{b}$ denotes the unknown error-free vector associated with b. We will assume that a bound for the norm of the error,

$$\|e\| \leq \delta, \tag{1.3}$$

is explicitly known and that the linear system of equations associated with the error-free right-hand side,

$$Ax = \hat{b}, \tag{1.4}$$

is consistent.

*Version June 15, 2010. Department of Mathematics and Computing Science, Eindhoven University of Technology, PO Box 513, 5600 MB, The Netherlands, www.win.tue.nl/∼hochsten

†Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA. E-mail: reichel@math.kent.edu

We would like to determine an approximation of the solution of minimal least-squares norm, $\hat{x}$, of the unavailable linear system of equations (1.4). This is accomplished by computing a suitable approximate solution of the available least-squares problem (1.1). Note that the minimal-norm solution of (1.1) generally is a poor approximation of the desired vector $\hat{x}$ due to the error e in b and the severe ill-conditioning of A.

Tikhonov regularization replaces the minimization problem (1.1) by the solution of a penalized least-squares problem

$$\min_{x\in\mathbb{C}^n} \{ \|Ax - b\|^2 + \mu \|Lx\|^2 \} \tag{1.5}$$

with regularization operator $L \in \mathbb{C}^{p\times n}$ and regularization parameter $\mu > 0$. For future reference, we note that the normal equations associated with (1.5) are given by

$$(A^*A + \mu L^*L)\, x = A^*b, \tag{1.6}$$

where A^* and L^* denote the adjoints of A and L, respectively. We will assume that

$$\mathcal{N}(A) \cap \mathcal{N}(L) = \{0\}, \tag{1.7}$$

where $\mathcal{N}$ denotes the null space. Then (1.5) has a unique solution, $x^{(\mu)}$, for all $\mu > 0$. The component of $x^{(\mu)}$ in $\mathcal{N}(L)$ is independent of μ. We remark that for many commonly used regularization operators L, such as approximations of differential operators, the restriction of A to $\mathcal{N}(L)$ is quite well conditioned.

The vectors $x^{(\mu)}$ for $\mu > 0$ are less sensitive to the error e in b than the solution of (1.1). The sensitivity of $x^{(\mu)}$ to e and the difference $x^{(\mu)} - \hat{x}$ depend on both the value of μ and the choice of regularization operator L. We would like $x^{(\mu)} - \hat{x}$ to be small.

The available bound (1.3) for the error e allows us to determine a suitable value of μ by the discrepancy principle, which prescribes that $\mu = \mu(\delta)$ be chosen so that

$$\|Ax^{(\mu)} - b\| = \eta\delta, \tag{1.8}$$

where $\eta > 1$ is a user-specified constant independent of δ. Then

$$\lim_{\delta\searrow 0} x^{(\mu(\delta))} = \hat{x};$$

see, e.g., Engl et al. [10] and Groetsch [12] for proofs in Hilbert space settings. Hence, the numerical solution of (1.5) entails both the determination of a value of μ and the computation of an approximation x_k of the solution $x^{(\mu)}$ of (1.5), such that x_k satisfies (1.8).

We remark that the minimization problem (1.5) with the constraint (1.8) can be formulated as

$$\min_x \|Lx\| \quad \text{such that} \quad \|Ax - b\| = \eta\delta, \tag{1.9}$$

which shows that the regularization parameter μ determined by (1.8) is the Lagrange multiplier for (1.9).

When the matrices A and L are of small to moderate size, the minimization problem (1.5) conveniently can be solved with the aid of the generalized singular

value decomposition (GSVD) of the matrix pair $\{A, L\}$; see, e.g., [13]. The present paper is concerned with the situation when A and L are too large to compute their GSVD.

The Tikhonov minimization problem (1.5) is said to be in standard form when L equals the identity operator I. Substituting $y = Lx$ into (1.5) yields the standard-form problem

$$\min_{y\in\mathbb{C}^n} \{ \|AL_A^\dagger y - \bar{b}\|^2 + \mu \|y\|^2 \}, \tag{1.10}$$

with solution $y^{(\mu)}$, where

$$\bar{b} = b - A\bar{x}, \qquad \bar{x} = (A(I - L^\dagger L))^\dagger b, \tag{1.11}$$

and

$$L_A^\dagger = \left(I - (A(I - L^\dagger L))^\dagger A\right) L^\dagger \in \mathbb{R}^{n\times p} \tag{1.12}$$

is the A-weighted generalized inverse of A. Here $L^\dagger$ denotes the Moore-Penrose pseudoinverse of L. The solution of (1.5) is given by

$$x^{(\mu)} = L_A^\dagger y^{(\mu)} + \bar{x};$$

see, e.g., Eldén [9] or Hansen [13, Section 2.3] for details.

An attraction of Tikhonov regularization problems in standard form is that the computations required for determining a suitable value of the regularization parameter, say, by the discrepancy principle or the L-curve, are fairly simple; see [5, 6] for illustrations. However, iterative methods applied to the the solution of the standard form problem (1.10) require matrix-vector product evaluations with the matrices $L_A^\dagger$, $AL_A^\dagger$, and possibly also with $(L_A^\dagger)^*$ and $(AL_A^\dagger)^*$. Only regularization operators L with particular simple structure allow for the efficient evaluation of these matrix-vector products. This includes regularization operators with a small bandwidth, circulant matrices, orthogonal projections, and sparse nonsingular matrices that permit fast LU factorization as well as fast forward and back substitution; see [7, 9, 13, 18, 21, 22] for some examples. Moreover, efficient evaluation of $(A(I - L^\dagger L))^\dagger$ in (1.11) and (1.12) requires that $\mathcal{N}(L)$ be explicitly known and of low dimension.

The method for the solution of (1.5) proposed in Section 2 of this paper can be applied when matrix-vector products with $L_A^\dagger$ and $AL_A^\dagger$ cannot be evaluated efficiently, and when $\mathcal{N}(L)$ is not explicitly known. The method is based on partial Golub-Kahan bidiagonalization of A and requires only matrix-vector product evaluations with the matrices A and L, as well as with their adjoints. This makes the method suitable for the solution of large-scale Tikhonov minimization problems (1.5) with fairly general linear regularization operators L. Section 3 discusses zero-finders for the determination of a value of the regularization parameter, so that (1.8) is approximately satisfied. A few computed examples are presented in Section 4, and concluding remarks can be found in Section 5.

There are not many efficient methods available for the solution of large-scale Tikhonov minimization problems (1.5) with a general linear regularization operator. One of the most interesting such methods is the inner-outer iterative scheme recently proposed by Kilmer et al. [16]. This scheme is inspired by an iterative method due to Zha [23] for computing a partial GSVD of the matrix pair $\{A, L\}$. The scheme [16]

can be expensive for problems that require a large number of inner iterations. Therefore we believe it to be worthwhile to explore alternative approaches. We remark that the "obvious" solution method is to apply the conjugate gradient or preconditioned conjugate gradient method to the normal equations (1.6). However, this approach often is computationally expensive when a suitable value of the regularization parameter μ is not know a priori, because in this situation several systems (1.6) with different values of μ have to be solved. The inner-outer method proposed by Jacobsen et al. [15] requires L^*L to be nonsingular. Many regularization operators of interest do not satisfy this requirement. The recently proposed scheme in [20] can be applied when both A and L are square matrices.

We conclude this section by noting that when the matrix A stems from the discretization of a compact integral operator, discretization implies regularization. When the integral operator is discretized coarsely enough, i.e., when m and n are small, the ratio between the largest and smallest singular values of A may not be very large and we can solve the minimization problem (1.1) in a straightforward manner without Tikhonov regularization. Regularization by discretization is investigated, e.g., by Natterer [19]. The difficulty with this approach is that the appropriate discretization, i.e., a suitable choice of m and n, depends on the error e in b and generally is not known a priori. Tikhonov regularization makes it possible to decouple the discretization from the error in the data b and therefore often is simpler to use.

2. An iterative method. We evaluate an approximate solution of the Tikhonov minimization problem (1.5) by first computing a partial Golub-Kahan bidiagonalization of the matrix A. This yields a Krylov subspace in which the approximate solution is sought. Typically this Krylov subspace contains a fairly accurate approximation of $\hat{x}$. The regularization operator L is projected into this Krylov subspace. The purpose of the regularization operator L is to steer the method towards a suitable approximate solution in the Krylov subspace.

Throughout this paper e_j denotes the jth axis vector of appropriate dimension. Application of k steps of Golub-Kahan bidiagonalization to the matrix A with initial vector b yields the matrices $U_{k+1} \in \mathbb{C}^{m\times(k+1)}$ and $V_k \in \mathbb{C}^{n\times k}$ with orthonormal columns, and the bidiagonal matrix $\tilde{C}_k \in \mathbb{C}^{(k+1)\times k}$, such that

$$AV_k = U_{k+1}\tilde{C}_k, \qquad A^*U_k = V_kC_k, \qquad U_{k+1}e_1 = b/\|b\|, \tag{2.1}$$

where $U_k \in \mathbb{C}^{m\times k}$ is made up of the k first columns of U_{k+1}, $C_k \in \mathbb{C}^{k\times k}$ consists of the first k rows of $\tilde{C}_k$, and the columns of V_k span the Krylov subspace

$$\mathcal{K}_k(A^*A, A^*b) = \text{span}\{A^*b, (A^*A)A^*b, \ldots, (A^*A)^{k-1}A^*b\}; \tag{2.2}$$

see, e.g., Björck [3] for details.

We will use the QR factorization

$$LV_k = Q_kR_k, \tag{2.3}$$

where $Q_k \in \mathbb{C}^{p\times k}$ has orthonormal columns and $R_k \in \mathbb{C}^{k\times k}$ is upper triangular. In applications of interest $k \ll p$.

The computational effort to determine the decompositions (2.1) when m and n are large is dominated by the k matrix-vector product evaluations required with each one of the matrices A and A^*. The matrix L generally is very sparse. Therefore, the computational effort needed to evaluate LV_k typically is much smaller than the effort required for the evaluation of k matrix-vector products with A.

We require the computed kth approximation, x_k, of the solution of (1.5) to live in the Krylov subspace (2.2). It can be expressed as $x_k = V_k y_k$ for some vector $y_k \in \mathbb{C}^k$. Substituting $x = V_k y$ into (1.5) and using the properties (2.1) and (2.3) yields the reduced minimization problem

$$\min_{y\in\mathbb{C}^k} \left\| \begin{bmatrix} \tilde{C}_k \\ \sqrt{\mu}\, R_k \end{bmatrix} y - \begin{bmatrix} e_1\|b\| \\ 0 \end{bmatrix} \right\|. \tag{2.4}$$

Since the subspace dimension k typically is quite small, this least-squares problem can be solved efficiently by a direct method. For instance, we may first transform the matrix in (2.4) into upper triangular form by application of a judiciously chosen sequence of Givens rotations. Due to the assumption (1.7), the solution $y_k^{(\mu)}$ of (2.4) is unique for all $\mu > 0$.

We determine μ by requiring that $y_k = y_k^{(\mu)}$ satisfies

$$\|\tilde{C}_k y_k - e_1\|b\|\,\| = \eta\delta. \tag{2.5}$$

Let μ_k denote the solution μ of (2.5). The computation of μ_k requires that a sequence of least-squares problems (2.4) with different μ-values be solved. More details on the computation of μ_k and y_k are provided in Section 3.

PROPOSITION 2.1. *Let μ_k solve* (2.5)*, denote by $y_k = y_k^{(\mu_k)}$ the associated solution of* (2.4)*, and let $x_k = V_k y_k$ be the corresponding approximate solution of* (1.5)*. Then*

$$\|Ax_k - b\| = \eta\delta.$$

Proof. We have

$$\|Ax_k - b\| = \|AV_k y_k - b\| = \|\tilde{C}_k y_k - e_1\|b\|\,\|.$$

The proposition now follows from (2.5). □

When increasing the number of bidiagonalization steps k, the QR factorization of LV_k, see (2.3), has to be updated. Formulas for updating a QR factorization are described by Daniel et al. [8]; see also [11, Section 12.5]. Note that only the upper triangular matrices R_k, $k = 1, 2, \ldots$, are required, but not the associated matrices Q_k with orthonormal columns.

This paper focuses on the determination of a suitable regularization parameter μ for Tikhonov regularization. However, the number of bidiagonalization steps, k, also may be regarded a regularization parameter. It restricts the (sub)space in which the computed approximation of $\hat{x}$ is sought to k dimensions. We comment on the choice of k further in Sections 3 and 4.

The null space $\mathcal{N}(L)$ can be important for achieving an accurate approximation $x^{(\mu)}$ of $\hat{x}$ by Tikhonov regularization (1.5). Since the component of $x^{(\mu)}$ in $\mathcal{N}(L)$ is independent of $\mu > 0$, we may choose L so that $\mathcal{N}(L)$ represents important known features of the desired solution $\hat{x}$. However, the reduced regularization operator R_k in (2.4) typically is nonsingular also when L has a nontrivial null space. We now describe a splitting of the minimization problem (1.1), such that Tikhonov regularization is not applied to the solution component in $\mathcal{R}(W)$, where $W \in \mathbb{C}^{n\times\ell}$ is a user-supplied matrix and $\mathcal{R}(W)$ denotes its range. This splitting has previously been applied in iterative and direct methods for ill-posed problems in [2, 4, 17].

Let the matrix $W \in \mathbb{C}^{n\times\ell}$ have orthonormal columns and introduce the QR factorization $AW = \breve{Q}\breve{R}$, where $\breve{Q} \in \mathbb{C}^{n\times\ell}$ has orthonormal columns and $\breve{R} \in \mathbb{C}^{\ell\times\ell}$ is

upper triangular. We may assume that W is chosen so that $\breve{R}$ is nonsingular. Define the orthogonal projectors

$$P_W = WW^*, \qquad P_W^\perp = I - WW^*, \qquad P_{\breve{Q}} = \breve{Q}\breve{Q}^*, \qquad P_{\breve{Q}}^\perp = I - \breve{Q}\breve{Q}^*.$$

Then (1.1) can be written as

$$\begin{aligned}\min_{x\in\mathbb{C}^n} \|Ax-b\|^2 &= \min_{x\in\mathbb{C}^n} \{\|P_{\breve{Q}}Ax - P_{\breve{Q}}b\|^2 + \|P_{\breve{Q}}^\perp Ax - P_{\breve{Q}}^\perp b\|^2\} \\ &= \min_{x\in\mathbb{C}^n} \{\|P_{\breve{Q}}AP_W x - (P_{\breve{Q}}b - P_{\breve{Q}}AP_W^\perp x)\|^2 + \|P_{\breve{Q}}^\perp AP_W^\perp x - P_{\breve{Q}}^\perp b\|^2\},\end{aligned}$$

where we have used that $I = P_W + P_W^\perp$ and $P_{\breve{Q}}^\perp AP_W = O$. Let $y = W^*x$. Then

$$\|P_{\breve{Q}}AP_W x - (P_{\breve{Q}}b - P_{\breve{Q}}AP_W^\perp x)\| = \|Ry - (\breve{Q}^*b - \breve{Q}^*AP_W^\perp x)\|. \tag{2.6}$$

Since R is nonsingular, we may for any $P_W^\perp x$ choose y so that the expression in the right-hand side of (2.6) vanishes. This choice of y shows that

$$\min_{x\in\mathbb{C}^n} \|Ax-b\| = \min_{x\in\mathbb{C}^n} \|P_{\breve{Q}}^\perp AP_W^\perp x - P_{\breve{Q}}^\perp b\|. \tag{2.7}$$

We solve the projected problem in the right-hand side of (2.7) by the method of the present paper and then determine y so that the right-hand side of (2.6) vanishes. Since $P_{\breve{Q}}^\perp AP_W^\perp = P_{\breve{Q}}^\perp A$, we may omit the projector $P_W^\perp$ in the projected problem. This splitting is applied in Examples 4.2 and 4.3 below. Generally, the number of columns, ℓ, of the matrix W is quite small, say, $\ell \leq 3$.

3. Determining the regularization parameter. This section discusses the computation of $\mu = \mu_k$ so that $y_k = y_k^{(\mu_k)}$ satisfies (2.5). Introduce the function

$$\phi(\nu) = \|\tilde{C}_k y_k - e_1\|b\|\,\|^2, \qquad \mu = 1/\nu, \qquad 0 < \nu < \infty, \tag{3.1}$$

where $y_k = y_k^{(\mu)}$ is the solution of (2.4). Then equation (2.5) can be expressed as

$$\phi(\nu) = \eta^2\delta^2. \tag{3.2}$$

We first describe an approach that can be applied when the matrix R_k in (2.3) is not ill-conditioned and $\tilde{C}_k^* e_1 \neq 0$. These conditions typically are satisfied. For instance, the latter condition holds when $A^*b \neq 0$. The following proposition is formulated in terms of the QR factorization

$$\tilde{C}_k = \tilde{Q}_k \tilde{R}_k,$$

where $\tilde{Q}_k \in \mathbb{C}^{(k+1)\times k}$ has orthonormal columns and $\tilde{R}_k \in \mathbb{C}^{k\times k}$ is upper triangular.

PROPOSITION 3.1. *Assume that the matrix R_k in (2.3) is nonsingular and that $\tilde{C}_k^* e_1 \neq 0$. Let $\hat{R} = \tilde{R}_k R_k^{-1}$. Then the function (3.1) can be expressed as*

$$\phi(\nu) = \|b\|^2 e_1^* \tilde{Q}_k (\nu\hat{R}\hat{R}^* + I)^{-2} \tilde{Q}_k^* e_1 + \|b\|^2 e_1^*(I - \tilde{Q}_k\tilde{Q}_k^*)e_1. \tag{3.3}$$

Consequently, $\phi(\nu)$ is strictly decreasing and convex, and equation (3.2) has a unique solution $0 < \nu_k < \infty$, provided that

$$\|P_{\mathcal{N}(\tilde{C}_k^*)} e_1\|\|b\| < \eta\delta < \|b\|, \tag{3.4}$$

where $P_{\mathcal{N}(\tilde{C}_k^*)}$ denotes the orthogonal projector onto $\mathcal{N}(\tilde{C}_k^*)$.

Proof. The representation (3.3) follows from

$$I - \hat{R}(\hat{R}^*\hat{R} + \nu^{-1}I)^{-1}\hat{R}^* = (\nu\hat{R}\hat{R}^* + I)^{-1}$$

and shows that ϕ is decreasing and convex. Moreover, it follows from (3.3) that

$$\lim_{\nu \searrow 0} \phi(\nu) = \|b\|^2.$$

Since the function ϕ is decreasing, the upper bound of (3.4) has to be satisfied in order for equation (3.2) to have a positive solution. The lower bound of (3.4) corresponds to $\mu = 0$ in (2.4). Therefore,

$$\lim_{\nu \to \infty} \phi(\nu) = \min_{y \in \mathbb{C}^k} \|\tilde{C}_k y - e_1 \|b\| \,\|^2 = \|P_{\mathcal{N}(\tilde{C}_k^*)} e_1\|^2 \|b\|^2.$$

Since ϕ is decreasing, the lower bound of (3.4) has to be satisfied in order for equation (3.2) to have a bounded solution. Therefore, when the bounds (3.4) hold, equation (3.2) has a unique bounded solution. □

We remark that $\|P_{\mathcal{N}(\tilde{C}_k^*)} e_1\|$ is decreasing when k is increasing. This follows from the observations that $I = P_{\mathcal{N}(\tilde{C}_k^*)} + P_{\mathcal{R}(\tilde{C}_k)}$, $P_{\mathcal{N}(\tilde{C}_k^*)}$ and $P_{\mathcal{R}(\tilde{C}_k)}$ are orthogonal, and $\|P_{\mathcal{R}(\tilde{C}_k)} e_1\|$ is increasing with k. Therefore, to satisfy the left-hand side inequality in (3.4), k has to be sufficiently large. In actual computations, it generally suffices to choose k fairly small. This is illustrated in Section 4.

Using (3.3), the function $\phi(\nu)$ can be evaluated by solving a least-squares problem related to (2.4). The derivative $\phi'(\nu)$ can be computed by solving another least-squares problem with the same matrix. This allows for efficient implementation of Newton's method for the solution of (3.2); see, e.g., [6].

When the matrix R_k is ill-conditioned, the GSVD of the matrix pair $\{\tilde{C}_k, R_k\}$ can be used. Substituting the GSVD into (2.4) and (3.1) gives a simple expression for the evaluation of $\phi(\nu)$. However, each increase of k requires the computation of the GSVD of a new matrix pair $\{\tilde{C}_k, R_k\}$. Thus, typically GSVDs of several matrix pairs have to be computed and the computational effort is larger than if the approach of Proposition 3.1 is used.

4. Numerical examples. The right-hand sides in the examples below are contaminated by an error e with normally distributed entries with zero mean. The entries are scaled to correspond to a specified relative error,

$$\varepsilon = \|e\|/\|\hat{b}\|. \tag{4.1}$$

The constant η in the discrepancy principle (1.8) is set to 1.1 in all examples, and we let $\delta = \varepsilon \|\hat{b}\|$ in (1.8).

EXAMPLE 4.1. We discretize the integral equation

$$\int_0^\pi \exp(s\cos(t))x(t)\,dt = 2\,\frac{\sinh(s)}{s}, \qquad 0 \le s \le \frac{\pi}{2},$$

discussed by Baart [1] by a Galerkin method with piecewise constant test and trial functions using the MATLAB code baart from [14]. This yields the nonsymmetric matrix $A \in \mathbb{R}^{1000\times 1000}$ of ill-determined rank. The code also furnishes the "exact"

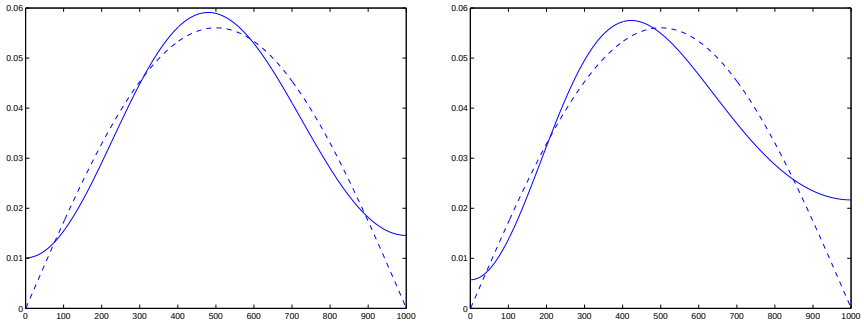


FIG. 4.1. *Example 4.1: (a) Computed approximate solution with tridiagonal regularization operator* (4.2) *(continuous graph) and solution* $\hat{x}$ *of the error-free problem* (1.4) *(dashed graph), (b) computed approximate solution with regularization operator* $L = I$ *(continuous graph) and solution* $\hat{x}$ *of the error-free problem* (1.4) *(dashed graph).*

solution $\hat{x}$, which represents a scaled sine function. We determine the error-free right-hand side of (1.4) as $\hat{b} = A\hat{x}$. The associated contaminated vector b in (1.1) is obtained by adding 0.1% normally distributed zero mean "noise" e to $\hat{b}$; cf. (1.2). Thus, $\varepsilon = 1 \cdot 10^{-3}$ in (4.1).

We compare approximations of $\hat{x}$ determined with the tridiagonal regularization operator

$$L = \begin{bmatrix} -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots & \ddots \\ & & -1 & 2 & -1 \end{bmatrix} \in \mathbb{R}^{(n-2)\times n}, \tag{4.2}$$

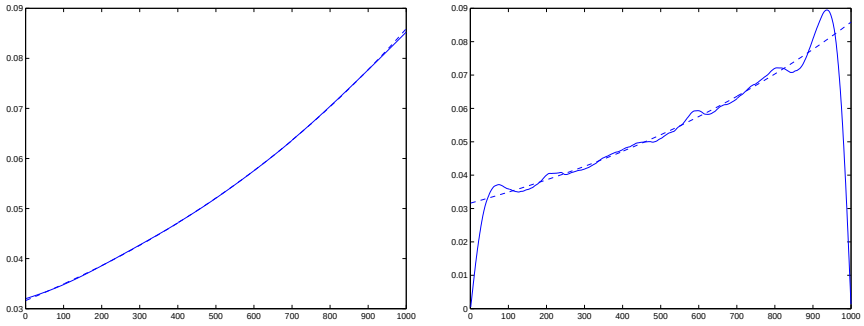
which is a scaled approximation of a second derivative operator, with approximations obtained with $L = I$. The number of bidiagonalization steps k in (2.1) has to be large enough so that (2.5) can be satisfied. In the present example, we let $k = 5$. The computed approximate solution x_5 has relative error $\|x_5 - \hat{x}\|/\|\hat{x}\| = 1.6 \cdot 10^{-1}$ when $L = I$, and relative error $\|x_5 - \hat{x}\|/\|\hat{x}\| = 1.0 \cdot 10^{-1}$ when L is given by (4.2). Thus, the former choice yields an increase of 60% of the error in the computed approximate solution. The computed approximate solutions are displayed in Figure 4.1.

We remark that the Krylov subspace $\mathcal{K}_5(A^*A, A^*b)$ contains a fairly accurate approximation $V_5V_5^*\hat{x}$ of $\hat{x}$. We have $\|\hat{x} - V_5V_5^*\hat{x}\|/\|\hat{x}\| = 5.3 \cdot 10^{-2}$. The purpose of the regularization operator L is to help determine an accurate approximation of $V_5V_5^*\hat{x}$. The present example shows the regularization operator (4.2) to yield a better approximation of $\hat{x}$ than $L = I$.

We note that the particular operator (4.2) allows the application of the A-weighted generalized inverse $L_A^\dagger$ of L; cf. (1.12). The purpose of this example is to show that an improvement of the quality of the computed approximate solution also can be achieved without applying $L_A^\dagger$.

The small dimension $k = 5$ of the solution would appear to contributes significantly to the regularization of the present problem. However, $k = 10$ bidiagonalization steps yields the computed approximate solutions x_{10} with $\|x_{10} - \hat{x}\|/\|\hat{x}\| = 1.6 \cdot 10^{-1}$ for $L = I$ and x_{10} with $\|x_{10} - \hat{x}\|/\|\hat{x}\| = 1.0 \cdot 10^{-1}$ for L defined by (4.2). Thus, the difference in the quality of the computed approximate solutions for $k = 5$ and $k = 10$

is negligible. We conclude that the subspace dimension, when larger or equal to 5, only has a minor influence on the computed solutions in this example. □



(a) (b)

FIG. 4.2. *Example 4.2: (a) Computed approximate solution with tridiagonal regularization operator* (4.2) *and splitting of the problem based on* $\mathcal{R}(W)$ *with* W *given by* (4.3) *(continuous graph) and solution* $\hat{x}$ *of the error-free problem* (1.4) *(dashed graph), (b) computed approximate solution with regularization operator* $L = I$ *without splitting (continuous graph) and solution* $\hat{x}$ *of the error-free problem* (1.4) *(dashed graph).*

EXAMPLE 4.2. Consider the Fredholm integral equation of the first kind

$$\int_0^1 k(s,t)x(t)\,dt = \exp(s) + (1-e)s + 1, \qquad 0 \le s \le 1,$$

where

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

We discretize the integral equation by a Galerkin method with orthonormal box functions as test and trial functions using the MATLAB program deriv2 from Regularization Tools [14] and obtain the symmetric indefinite matrix $A \in \mathbb{R}^{1000\times 1000}$ and the solution $\hat{x}$ of the error-free linear system (1.4). The vector $\hat{x}$ is a scaled discrete approximation of the exponential function. The error-free right-hand side $\hat{b}$ of (1.4) and the associated noise-contaminated vector b are determined similarly as in Example 4.1. In particular, $\varepsilon = 1 \cdot 10^{-3}$ in (4.1).

We first compute an approximate solution x_{10} with $L = I$ and 10 bidiagonalization steps. Figure 4.2(b) displays x_{10}. The relative error $\|x_{10} - \hat{x}\|/\|\hat{x}\| = 1.7 \cdot 10^{-1}$ is fairly large. Our first attempt to reduce this error by instead using the regularization operator (4.2) was not successful; we obtained, again with 10 bidiagonalization steps, an approximate solution with the larger relative error $1.8 \cdot 10^{-1}$.

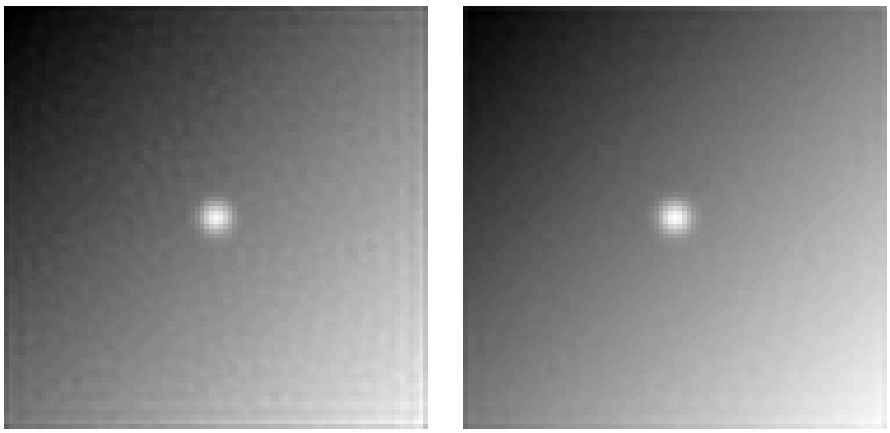
A more accurate approximation of $\hat{x}$ can be computed by splitting the problem as described in Example 2.1. Let the columns of $W \in \mathbb{R}^{n\times 3}$ form an orthonormal basis for the subspace

$$\mathcal{W} = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix}, \begin{bmatrix} 1 \\ 2^2 \\ \vdots \\ n^2 \end{bmatrix} \right\}. \tag{4.3}$$

The columns of W can represent quadratic growth of the solution. The component of the computed solution in $\mathcal{R}(W)$ is not affected by regularization. We solve the minimization problem

$$\min_{x\in\mathbb{C}^n} \|P_{\bar{Q}}^{\perp} Ax - P_{\bar{Q}}^{\perp} b\|$$

by Tikhonov regularization. The regularization operator (4.2) and 5 bidiagonalization steps yield the approximate solution x_5 shown in Figure 4.2(a) with relative error $\|x_5 - \hat{x}\|/\|\hat{x}\| = 2.4 \cdot 10^{-3}$. If we instead use the regularization operator $L = I$, then we obtain an approximate solution with relative error $3.7 \cdot 10^{-3}$. Thus, splitting the problem based on the matrix (4.3) and using the regularization operator (4.2) yields the most accurate approximation of $\hat{x}$. □



(a) (b)

FIG. 4.3. *Example 4.3: (a) Restoration obtained with L defined by* (4.5)*, (b) restoration obtained with L defined by* (4.5) *and splitting determined by* (4.3).

EXAMPLE 4.3. We consider a 91×91-pixel test image, which shows the superposition of a Gaussian and a linear function. The available image is contaminated by blur and 0.5% noise. The pixel values of this image, ordered column wise, determines the vector $b \in \mathbb{R}^{8281}$. The blurring operator is represented by the symmetric block Toeplitz matrix with Toeplitz blocks,

$$A = (2\pi\sigma^2)^{-1} T \otimes T, \tag{4.4}$$

where T is a 91×91 symmetric banded Toeplitz matrix, whose first row is given by `[exp(-((0:band-1).^2)/(2*sigma^2)); zeros(1,n-band)]`. The parameter `band` is the half-bandwidth of the matrix T. The parameter σ controls the effective width of the underlying Gaussian point spread function

$$h(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right),$$

which models blurring. We let `band` $= 16$ and $\sigma = 1.5$. The matrix A so obtained is numerically singular.

Following Kilmer et al. [16], we use the regularization operator

$$L = \begin{bmatrix} I & \otimes & L_1 \\ L_1 & \otimes & I \end{bmatrix}, \qquad L_1 = \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \end{bmatrix} \in \mathbb{R}^{90\times 91}, \tag{4.5}$$

where L_1 is a discrete approximation of the first derivative operator on a regular grid. The A-weighted generalized inverse of L, which is of size 16380×8281, is unwieldy to use in an iterative method. However, the method of the present paper is quite easy to apply when L is of the form (4.5).

Figure 4.3(a) shows the restoration x_{20} obtained with $k = 20$ bidiagonalization steps and the regularization operator (4.5). It has relative error $\|x_{20} - \hat{x}\|/\|\hat{x}\| = 2.36\cdot 10^{-2}$. The best approximation of $\hat{x}$ in $\mathcal{K}_{20}(A^*A, A^*b)$ has relative error $2.28\cdot 10^{-2}$. Thus, the approximation x_{20} of $\hat{x}$ is close to the best possible in the Krylov subspace. Nevertheless, the background "ringing" is not pleasing. We first seek to determine a more accurate restoration by increasing the dimension of the Krylov subspace. Increasing the number of bidiagonalization steps to 30 and 40, indeed, yields more accurate restorations; we have $\|x_{30} - \hat{x}\|/\|\hat{x}\| = 2.31 \cdot 10^{-2}$ and $\|x_{40} - \hat{x}\|/\|\hat{x}\| = 2.30 \cdot 10^{-2}$. However, the restorations x_{30} and x_{40} are visually indistinguishable from x_{20}.

A better approach to obtain a more accurate restoration is to split the problem using the matrix (4.3), similarly as in Example 4.2. This splitting, the regularization operator (4.5), and 20 bidiagonalization steps, give the restoration x_{20} with relative error $\|x_{20} - \hat{x}\|/\|\hat{x}\| = 1.26 \cdot 10^{-2}$. The restoration is depicted in Figure 4.3(b). It displays much less "ringing" than the restoration of Figure 4.3(a). We remark that the approximation of $\hat{x}$ obtained with the splitting determined by (4.3), 20 bidiagonalization steps, and $L = I$ is less accurate than the one shown in Figure 4.3(b). □

Example 4.3 illustrates that the splitting described in Example 2.1 can be beneficial for image restoration. However, the success of the particular splitting used in the above example depends on the image. For instance, this splitting does not improve the restoration of the image of the following example.

EXAMPLE 4.4. We apply the regularization operator (4.5) to restore the 91×91-pixel image groetsch, which has been contaminated by blur defined by (4.4) and by 0.1% noise. The contaminated image is shown in Figure 4.4(a). The restored image, x_{50}, determined with 50 bidiagonalization steps and the regularization operator (4.5) is displayed in Figure 4.4(b). It has relative error $\|x_{50} - \hat{x}\|/\|\hat{x}\| = 6.91 \cdot 10^{-2}$. □

5. Conclusion. We have presented a new iterative method for the solution of Tikhonov-regularized large-scale discrete ill-posed problems, which allows the linear regularization operator L to be of general form; neither the GSVD of the matrix pair $\{A, L\}$ nor the A-weighted generalized inverse of L are required. Only the evaluation of matrix-vector products with the matrices A, A^*, and L is demanded. The regularization parameter is determined during the iterations.

Acknowledgment. The authors would like to thank the referees for helpful comments. LR would like to thank MH for making an enjoyable visit to TU/e possible, during which work on the present paper was initiated.

(a) (b)

FIG. 4.4. *Example 4.4: (a) Blur- and noise-contaminated image, (b) restored image.*

REFERENCES

[1] M. L. Baart, The use of auto-correlation for pseudo-rank determination in noisy ill-conditioned least-squares problems, IMA J. Numer. Anal., 2 (1982), pp. 241–247.

[2] J. Baglama and L. Reichel, Decomposition methods for large linear discrete ill-posed problems, J. Comput. Appl. Math., 198 (2007), pp. 332–342.

[3] Å. Björck, Numerical Methods for Least Squares Problems, SIAM, Philadelphia, 1996.

[4] A. Bunse-Gerstner, V. Guerra-Ones, and H. Madrid de La Vega, An improved preconditioned LSQR for discrete ill-posed problems, Math. Comput. Simulation, 73 (2006), pp. 65–75.

[5] D. Calvetti, P. C. Hansen, and L. Reichel, L-curve curvature bounds via Lanczos bidiagonalization, Electron. Trans. Numer. Anal., 14 (2002), pp. 20–35.

[6] D. Calvetti and L. Reichel, Tikhonov regularization of large linear problems, BIT, 43 (2003), pp. 263–283.

[7] D. Calvetti, L. Reichel, and A. Shuibi, Invertible smoothing preconditioners for linear discrete ill-posed problems, Appl. Numer. Math., 54 (2005), pp. 135–149.

[8] J. W. Daniel, W. B. Gragg, L. Kaufman, and G. W. Stewart, Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization, Math. Comp., 30 (1976), pp. 772–795.

[9] L. Eldén, A weighted pseudoinverse, generalized singular values, and constrained least squares problems, BIT, 22 (1982), pp. 487–501.

[10] H. W. Engl, M. Hanke, and A. Neubauer, Regularization of Inverse Problems, Kluwer, Dordrecht, 1996.

[11] G. H. Golub and C. F. Van Loan, Matrix Computations, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.

[12] C. W. Groetsch, The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind, Pitman, Boston, 1984.

[13] P. C. Hansen, Rank-Deficient and Discrete Ill-Posed Problems, SIAM, Philadelphia, 1998.

[14] P. C. Hansen, Regularization tools version 4.0 for Matlab 7.3, Numer. Algorithms, 46 (2007), pp. 189–194.

[15] M. Jacobsen, P. C. Hansen, and M. A. Saunders, Subspace preconditioned LSQR for discrete ill-posed problems, BIT, 43 (2003), pp. 975–989.

[16] M. E. Kilmer, P. C. Hansen, and M. I. Español, A projection-based approach to general-form Tikhonov regularization, SIAM J. Sci. Comput., 29 (2007), pp. 315–330.

[17] S. Morigi, L. Reichel, and F. Sgallari, A truncated projected SVD method for linear discrete ill-posed problems, Numer. Algorithms, 43 (2006), pp. 197–213.

[18] S. Morigi, L. Reichel, and F. Sgallari, Orthogonal projection regularization operators, Numer. Algorithms, 44 (2007), pp. 99–114.

[19] F. Natterer, Regularisierung schlecht gestellter Probleme durch Projektionsverfahren, Numer. Math., 28 (1977), pp. 329–341.

[20] L. Reichel, F. Sgallari, and Q. Ye, Tikhonov regularization based on generalized Krylov subspace methods, submitted for publication.
[21] L. Reichel and Q. Ye, Simple square smoothing regularization operators, Electron. Trans. Numer. Anal., 33 (2009), pp. 63–83.
[22] G. Wahba, Spline Models for Observational Data, SIAM, Philadelphia, 1990.
[23] H. Zha, Computing the generalized singular values/vectors of large sparse or structured matrix pairs, Numer. Math., 72 (1996), pp. 391–417.