NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2000; **00**:1–11 *Prepared using* nlaauth.cls *[Version: 2000/03/22 v1.0]*

Tikhonov Regularization of Large Symmetric Problems

D. Calvetti[1], L. Reichel[2] and A. Shuibi[3]

[1] *Department of Mathematics, Case Western Reserve University, Cleveland, OH 44106. Email:dxc57@po.cwru.edu.*
[2] *Department of Mathematical Sciences, Kent State University, Kent, OH 44242. E-mail:reichel@math.kent.edu.*
[3] *Department of Mathematical Sciences, Kent State University, Kent, OH 44242. E-mail:ashuibi@math.kent.edu.*

SUMMARY

Many popular solution methods for large discrete ill-posed problems are based on Tikhonov regularization and compute a partial Lanczos bidiagonalization of the matrix. The computational effort required by these methods is not reduced significantly when the matrix of the discrete ill-posed problem, rather than being a general nonsymmetric matrix, is symmetric and possibly indefinite. This paper describes new methods, based on partial Lanczos tridiagonalization of the matrix, that exploit symmetry. Computed examples illustrate that one of these methods can require significantly less computational work than available structure-ignoring schemes. Copyright © 2000 John Wiley & Sons, Ltd.

1. Introduction

This paper is concerned with the computation of an approximate solution of linear systems of equations of the form

$$A\boldsymbol{x} = \boldsymbol{b}, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$ is a large symmetric matrix of ill-determined rank and the right-hand side $\boldsymbol{b} \in \mathbb{R}^n$ is contaminated by an unknown error $\boldsymbol{e} \in \mathbb{R}^n$. In particular, the matrix A has a "cluster" of eigenvalues at the origin and is severely ill-conditioned; it may be singular. Linear systems of equations with such matrices are commonly referred to as linear discrete ill-posed problems. They arise, for instance, when discretizing linear ill-posed problems, such as Fredholm integral equations of the first kind with a smooth kernel. The vector $\boldsymbol{e}$ represents

Contract/grant sponsor: NSF grant; contract/grant number: DMS-0107841.

Contract/grant sponsor: NSF grant; contract/grant number: DMS-0107858.

Contract/grant sponsor: NSF grant ; contract/grant number: DMS-0107858.

Received 17 January 2003
Revised 22 March 2003
Accepted 15 October 2003

the error in the available data. We assume that the error $\boldsymbol{e}$ is unknown, but that an estimate of its norm

$$\delta := \|\boldsymbol{e}\| \tag{2}$$

is available. Here and throughout this paper $\|\cdot\|$ denotes the Euclidean vector norm or the associated induced matrix norm.

Let $\hat{\boldsymbol{b}} := \boldsymbol{b} - \boldsymbol{e}$ denote the error-free right-hand side. We would like to compute a solution, denoted by $\hat{\boldsymbol{x}}$, of the error-free linear system of equations $A\boldsymbol{x} = \hat{\boldsymbol{b}}$ associated with (1). Hence, $\hat{\boldsymbol{x}}$ is the desired approximate solution of (1).

Let $A^\dagger$ denote the Moore-Penrose pseudo-inverse of the matrix A. Due to the error $\boldsymbol{e}$ in the right-hand side $\boldsymbol{b}$ and the severe ill-conditioning of the matrix A, the minimal-norm least-squares solution of (1), given by $A^\dagger\boldsymbol{b}$, in general, is not a meaningful approximation of $\hat{\boldsymbol{x}}$. This difficulty is remedied by replacing the discrete ill-posed problem (1) by a linear system of equations that is less sensitive to the perturbation $\boldsymbol{e}$ in the right-hand side. This replacement is referred to as regularization. The possibly most popular regularization method is due to Tikhonov, which in its simplest form replaces the system (1) by

$$(A^2 + \mu^{-1}I)\boldsymbol{x} = A\boldsymbol{b}, \tag{3}$$

where we have used the fact that A is symmetric. For any positive value of the regularization parameter μ, the Tikhonov equations (3) have the unique solution

$$\boldsymbol{x}^{(\mu)} := (A^2 + \mu^{-1}I)^{-1}A\boldsymbol{b}. \tag{4}$$

The value of $\mu > 0$ determines how sensitive the solution $\boldsymbol{x}^{(\mu)}$ is to the error $\boldsymbol{e}$ in $\boldsymbol{b}$. In the literature $\nu = 1/\mu$ is often used as the regularization parameter, however, for us it is more convenient to consider μ the regularization parameter. We refer to Engl et al. [5] and Hansen [8] for insightful discussions on Tikhonov regularization.

It is easy to verify that

$$\lim_{\mu\to\infty} \boldsymbol{x}^{(\mu)} = A^\dagger\boldsymbol{b}, \qquad \lim_{\mu\searrow 0} \boldsymbol{x}^{(\mu)} = \boldsymbol{0}.$$

These limits, generally, do not furnish meaningful approximations of $\hat{\boldsymbol{x}}$. The determination of a suitable positive finite value of the regularization parameter, such that $\boldsymbol{x}^{(\mu)}$ is a meaningful approximation of $\hat{\boldsymbol{x}}$, is an important part of the solution of the Tikhonov equations (3). Our determination of μ is based on the discrepancy principle. The discrepancy associated with $\boldsymbol{x}^{(\mu)}$ is defined by

$$\boldsymbol{d}^{(\mu)} := \boldsymbol{b} - A\boldsymbol{x}^{(\mu)}. \tag{5}$$

It is easy to show that $\|\boldsymbol{d}^{(\mu)}\|$ is a decreasing function of $\mu > 0$; see, e.g., [2, Theorem 2.1]. The solution $\boldsymbol{x}^{(\mu)}$ of (3) is said to satisfy the discrepancy principle if

$$\|\boldsymbol{d}^{(\mu)}\| = \delta, \tag{6}$$

where δ is defined by (2). We note that $\boldsymbol{x}^{(\mu)} \to A^\dagger\hat{\boldsymbol{b}}$ as $\delta \searrow 0$; see, e.g., [5, Section 4.3]. We denote the value of μ, such that (6) holds, by μ_*. Let $\boldsymbol{b}_0$ denote the orthogonal projection of $\boldsymbol{b}$ onto the null space of A. If A is nonsingular, then $\boldsymbol{b}_0 = \boldsymbol{0}$. We will throughout this paper assume that

$$\|\boldsymbol{b}_0\| < \|\boldsymbol{e}\| < \|\boldsymbol{b}\|. \tag{7}$$

These inequalities secure that the value μ_* is unique and satisfies $0 < \mu_* < \infty$.

In applications, the available value of δ generally is an estimate of $\|\boldsymbol{e}\|$. It is therefore not meaningful to solve equation (6) to very high accuracy. Instead, we will seek to determine a value $\hat{\mu}$, such that

$$\delta \leq \|\boldsymbol{d}^{(\hat{\mu})}\| \leq (1+\varepsilon_\delta)\delta, \tag{8}$$

where the choice of $\varepsilon_\delta > 0$ depends on the accuracy of the estimate δ of $\|\boldsymbol{e}\|$.

There are several methods available for determining an approximation of the vector (4) for a value $\mu = \hat{\mu}$, where $\hat{\mu}$ satisfies (8). These include methods based on partial Lanczos bidiagonalization of A or on solving the Tikhonov equations (3) by the conjugate gradient method; see, e.g., [1, 2, 6] for examples. These methods allow the matrix A to be nonsymmetric and even nonsquare, but they are not able to exploit symmetry in a significant manner when A is symmetric and possibly indefinite.

It is the purpose of the present paper to describe numerical methods that are tailored for the solution of large linear discrete ill-posed problems with a symmetric matrix. The methods discussed compute a partial Lanczos tridiagonalization of A, instead of a partial Lanczos bidiagonalization. A suitable value of the regularization parameter is determined by adapting the technique discussed in [3] to the Tikhonov equations (3).

Large linear symmetric discrete ill-posed problems arise quite frequently in applications. However, besides the method of the present paper, we are only aware of a method discussed in [4] that exploits symmetry. The latter method is based on expanding the Tikhonov solution (4) in terms of Chebyshev polynomials.

This paper is organized as follows. Section 2 introduces functions with a zero at or close to μ_*. These functions are used to determine a value of the regularization parameter that satisfies (8). The computation of such a value and of an associated approximate solution of (3) is discussed in Section 3. We describe symmetry-exploiting Galerkin and minimal residual methods for the solution of the Tikhonov equations (3). Numerical examples are presented in Section 4. These examples illustrate that the symmetry-exploiting minimal residual method may require significantly less computational work than available symmetry-ignoring methods.

2. Auxiliary functions

We introduce functions that are used by the numerical methods for the determination of a suitable value of the regularization parameter and for the computation of an approximate solution of (3). This section reviews and extends results previously discussed in [3].

Let $\boldsymbol{x}^{(\mu)}$ be defined by (4) and introduce the functions

$$f(\mu) := \|A\boldsymbol{x}^{(\mu)} - \boldsymbol{b}\| - \delta, \qquad g(\mu) := \|A\boldsymbol{x}^{(\mu)} - \boldsymbol{b}\|^2 - \delta^2, \qquad \mu > 0. \tag{9}$$

Then $f(\mu)$ and $g(\mu)$ are analytic and decreasing functions of $\mu > 0$. Moreover, $g(\mu)$ is convex; see Propositions 2.1 and 2.2 in [3] for proofs. The assumption (7) secures that $f(\mu)$ and $g(\mu)$ have the unique bounded positive zero μ_*.

Since we will not compute the vectors $\boldsymbol{x}^{(\mu)}$ for $\mu > 0$, we cannot evaluate the functions $f(\mu)$ and $g(\mu)$. Let $\boldsymbol{x}_k^{(\mu)}$ be an available approximation of $\boldsymbol{x}^{(\mu)}$ and introduce the auxiliary functions

$$f_k(\mu) := \|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\| - \delta, \qquad g_k(\mu) := \|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\|^2 - \delta^2, \qquad \mu > 0, \tag{10}$$

whose values can be computed. In Section 3, we determine an approximation $\boldsymbol{x}_k^{(\mu)}$ of $\boldsymbol{x}^{(\mu)}$ by using a k-step partial Lanczos tridiagonalization of A.

We compute an approximation of μ_* by determining approximations of zeros of $f_k(\mu)$ and $g_k(\mu)$. The following theorem establishes when $f(\mu)$ and $f_k(\mu)$ are of the same sign. The theorem uses the residual vector associated with the Tikhonov equations (3),

$$\boldsymbol{r}_k^{(\mu)} := A\boldsymbol{b} - (A^2 + \mu^{-1}I)\boldsymbol{x}_k^{(\mu)}. \tag{11}$$

Theorem 2.1. *Let $f(\mu)$ and $f_k(\mu)$ be defined by (9) and (10), respectively, and let $\boldsymbol{r}_k^{(\mu)}$ be given by (11). If*

$$\frac{\sqrt{\mu}}{2}\|\boldsymbol{r}_k^{(\mu)}\| < |f_k(\mu)|, \tag{12}$$

then $f_k(\mu)$ and $f(\mu)$ are either both positive or both negative.

Proof. Introduce the spectral factorization

$$A = W\Lambda W^T, \tag{13}$$

where $\Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \cdots, \lambda_n]$ and $W \in \mathbb{R}^{n\times n}$ is orthogonal. Using this factorization, it is easy to show that

$$\|A(A^2 + \mu^{-1}I)^{-1}\| \le \frac{\sqrt{\mu}}{2}, \qquad \mu > 0. \tag{14}$$

It follows from the definitions of $f(\mu)$ and $f_k(\mu)$, and from (14), that

$$\begin{aligned}
|f_k(\mu) - f(\mu)| &\le \|A(\boldsymbol{x}_k^{(\mu)} - \boldsymbol{x}^{(\mu)})\| = \|A(\boldsymbol{x}_k^{(\mu)} - (A^2 + \mu^{-1}I)^{-1}A\boldsymbol{b})\| \\
&= \|A(A^2 + \mu^{-1}I)^{-1}((A^2 + \mu^{-1}I)\boldsymbol{x}_k^{(\mu)} - A\boldsymbol{b})\| \\
&\le \|A(A^2 + \mu^{-1}I)^{-1}\|\|\boldsymbol{r}_k^{(\mu)}\| \le \frac{\sqrt{\mu}}{2}\|\boldsymbol{r}_k^{(\mu)}\|.
\end{aligned}$$

It is now straightforward to verify that if (12) holds, then $f_k(\mu)$ and $f(\mu)$ are either both positive or both negative. □

The next theorem yields a bound on the distance between $g(\mu)$ and $g_k(\mu)$ in terms of the norm of the residual vector (11).

Theorem 2.2. *Let $g(\mu)$ and $g_k(\mu)$ defined (9) and (10), respectively, and let $\boldsymbol{r}_k^{(\mu)}$ be given by (11). Then*

$$|g_k(\mu) - g(\mu)| \le \frac{\sqrt{\mu}}{2}\|\boldsymbol{r}_k^{(\mu)}\| \left(\frac{\sqrt{\mu}}{2}\|\boldsymbol{r}_k^{(\mu)}\| + 2(f_k(\mu) + \delta)\right). \tag{15}$$

Proof. We have

$$\begin{aligned}
|g_k(\mu) - g(\mu)| &= \left|\|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\|^2 - \|A\boldsymbol{x}^{(\mu)} - \boldsymbol{b}\|^2\right| \\
&= \left(\|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\| + \|A\boldsymbol{x}^{(\mu)} - \boldsymbol{b}\|\right)\left|\|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\| - \|A\boldsymbol{x}^{(\mu)} - \boldsymbol{b}\|\right| \\
&= \left(\|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\| + \|A\boldsymbol{x}^{(\mu)} - \boldsymbol{b}\|\right)|f_k(\mu) - f(\mu)| \\
&\le \left(\|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\| + \|A\boldsymbol{x}^{(\mu)} - \boldsymbol{b}\|\right)\frac{\sqrt{\mu}}{2}\|\boldsymbol{r}_k^{(\mu)}\|,
\end{aligned}$$

where the last inequality follows from the proof of Theorem 2.1. Moreover,

$$\begin{aligned}
\|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\| + \|A\boldsymbol{x}^{(\mu)} - \boldsymbol{b}\| &= 2\|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\| - \Big(\|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\| - \|A\boldsymbol{x}^{(\mu)} - \boldsymbol{b}\|\Big) \\
&\leq 2\|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\| + \Big|\|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\| - \|A\boldsymbol{x}^{(\mu)} - \boldsymbol{b}\|\Big| \\
&= 2(f_k(\mu) + \delta) + |f_k(\mu) - f(\mu)| \\
&\leq 2(f_k(\mu) + \delta) + \frac{\sqrt{\mu}}{2}\|\boldsymbol{r}_k^{(\mu)}\|.
\end{aligned}$$

This shows the theorem. □

We will determine an approximation $\boldsymbol{x}_k^{(\mu)}$ of $\boldsymbol{x}^{(\mu)}$ by applying k steps of the Lanczos process to the matrix A with initial vector $\boldsymbol{b}$. The following result indicates how large k should be chosen.

Corollary 2.3. *Assume that* $\mu = \mu_*$. *Then*

$$\frac{\sqrt{\mu}}{2}\|\boldsymbol{r}_k^{(\mu)}\| \leq \varepsilon_\delta \delta \tag{16}$$

implies

$$\|A\boldsymbol{x}_k^{(\mu)} - \boldsymbol{b}\| \leq (1 + \varepsilon_\delta)\delta. \tag{17}$$

Proof. Let $\mu = \mu_*$. Then $f(\mu) = 0$ and it follows from the proof of Theorem 2.1 that

$$|f_k(\mu)| \leq \frac{\sqrt{\mu}}{2}\|\boldsymbol{r}_k^{(\mu)}\|.$$

This shows (17). □

3. Solution methods based on Lanczos tridiagonalization

We describe a Galerkin method and a minimal residual method for computing approximate solutions of the Tikhonov equations (3). Both methods are based on partial Lanczos tridiagonalization of the matrix A.

Application of k steps of the Lanczos process with initial vector $\boldsymbol{b}$ to the symmetric matrix A yields the partial Lanczos tridiagonalization

$$AV_k = V_{k+1}T_{k+1,k}, \tag{18}$$

where $V_{k+1} \in \mathbb{R}^{n\times(k+1)}$, $V_{k+1}^T V_{k+1} = I_{k+1}$, $V_{k+1}\boldsymbol{e}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|$, $V_k \in \mathbb{R}^{n\times k}$ consists of the k first columns of V_{k+1}, and

$$T_{k+1,k} := \begin{bmatrix} \alpha_0 & \beta_1 & & & & 0 \\ \beta_1 & \alpha_1 & \beta_2 & & & \\ & & \ddots & \ddots & \ddots & \\ & & & \beta_{k-2} & \alpha_{k-2} & \beta_{k-1} \\ & & & & \beta_{k-1} & \alpha_{k-1} \\ 0 & & & & & \beta_k \end{bmatrix} \in \mathbb{R}^{(k+1)\times k}$$

is tridiagonal with positive sub- and super-diagonal entries. Throughout this paper I_j denotes the identity matrix of order j and $\boldsymbol{e}_j$ the jth axis vector of appropriate dimension. We tacitly assume that k is small enough, so that the decomposition (18) with the stated properties exists. If the entry β_k vanishes, then the numerical methods to be described simplify. We will not dwell on this further, since generically β_k is positive. The columns of V_k span the Krylov subspace

$$\mathcal{K}_k(A,\boldsymbol{b}) := \mathrm{span}\{\boldsymbol{b}, A\boldsymbol{b}, \cdots, A^{k-1}\boldsymbol{b}\}. \tag{19}$$

The Galerkin method determines, for a fixed value of μ, an approximate solution $\boldsymbol{x}_k^{(\mu)}$ of (3) in the Krylov subspace (19). We express the Galerkin solution as a linear combination of the columns of V_k. This gives the Galerkin equations

$$V_k^T(A^2+\mu^{-1}I)V_k\boldsymbol{y} = V_k^T A\boldsymbol{b}, \tag{20}$$

which, using (18), can be expressed as

$$(T_{k+1,k}^T T_{k+1,k} + \mu^{-1}I_k)\boldsymbol{y} = T_{k+1,k}^T \boldsymbol{e}_1\|\boldsymbol{b}\|.$$

The latter equations are the normal equations for the least-squares problem

$$\min_{\boldsymbol{y}\in\mathbb{R}^k} \left\| \begin{bmatrix} T_{k+1,k} \\ \frac{1}{\sqrt{\mu}} I_k \end{bmatrix} \boldsymbol{y} - \boldsymbol{e}_1\|\boldsymbol{b}\| \right\|. \tag{21}$$

We solve the least-squares problem in $\mathcal{O}(k)$ arithmetic floating point operations by first computing the QR-factorization of the matrix using a judiciously chosen sequence of Givens rotations. Denote the computed solution of (21) by $\boldsymbol{y}_k^{(\mu)}$. Then $\boldsymbol{x}_k^{(\mu)} = V_k\boldsymbol{y}_k^{(\mu)}$ is the solution of the Galerkin equations.

The minimal residual method for the Tikhonov equations (3) solves the minimization problem

$$\min_{\boldsymbol{x}\in\mathcal{K}_k(A,\boldsymbol{b})} \|(A^2+\mu^{-1}I)\boldsymbol{x} - A\boldsymbol{b}\| = \min_{\boldsymbol{y}\in\mathbb{R}^k} \|(A^2+\mu^{-1}I)V_k\boldsymbol{y} - A\boldsymbol{b}\|. \tag{22}$$

Using the Lanczos decomposition (18), the right-hand side simplifies to

$$\min_{\boldsymbol{y}\in\mathbb{R}^k} \left\|(T_{k+2,k+1}T_{k+1,k} + \mu^{-1}I_{k+2,k})\boldsymbol{y} - T_{k+2,k+1}\boldsymbol{e}_1\|\boldsymbol{b}\|\right\|, \tag{23}$$

where $T_{k+2,k+1} \in \mathbb{R}^{(k+2)\times(k+1)}$ is the tridiagonal matrix obtained by applying $k+1$ steps of the Lanczos process with initial vector $\boldsymbol{b}$ to the matrix A, and $I_{k+2,k}$ denotes the leading $(k+2)\times k$ submatrix of I_{k+2}.

The minimization problem (23) can be solved efficiently without forming the matrix $T_{k+2,k+1}T_{k+1,k} + \mu^{-1}I_{k+2,k}$ as follows. Write the matrix in the form

$$T_{k+2,k+1}T_{k+1,k} + \mu^{-1}I_{k+2,k} = \begin{bmatrix} T_{k+1,k}^T T_{k+1,k} + \mu^{-1}I_k \\ W_k^T \end{bmatrix}, \tag{24}$$

where $W_k = [\boldsymbol{w}_k^{(1)}, \boldsymbol{w}_k^{(2)}] \in \mathbb{R}^{k\times 2}$ with $\boldsymbol{w}_k^{(1)} := \beta_k(T_k + \alpha_k I_k)\boldsymbol{e}_k$ and $\boldsymbol{w}_k^{(2)} := \beta_k\beta_{k+1}\boldsymbol{e}_k$. Here β_{k+1} is the last subdiagonal element of $T_{k+2,k+1}$. Factor the right-hand side of (24) according to

$$\begin{bmatrix} T_{k+1,k}^T T_{k+1,k} + \mu^{-1}I_k \\ W_k^T \end{bmatrix} = \begin{bmatrix} I_k \\ \hat{W}_k^T \end{bmatrix} \left(T_{k+1,k}^T T_{k+1,k} + \mu^{-1}I_k\right), \tag{25}$$

where $\hat{W}_k = [\hat{\boldsymbol{w}}_k^{(1)}, \hat{\boldsymbol{w}}_k^{(2)}] \in \mathbb{R}^{k\times 2}$. The columns of $\hat{W}_k$ satisfy

$$(T_{k+1,k}^T T_{k+1,k} + \mu^{-1} I_k)\hat{\boldsymbol{w}}_k^{(i)} = \boldsymbol{w}_k^{(i)}, \qquad i = 1, 2,$$

and we compute them by solving the least-squares problems

$$\min_{\hat{\boldsymbol{w}}\in\mathbb{R}^k} \left\| \begin{bmatrix} T_{k+1,k} \\ \frac{1}{\sqrt{\mu}} I_k \end{bmatrix} \hat{\boldsymbol{w}} - \sqrt{\mu} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{w}_k^{(i)} \end{bmatrix} \right\|, \qquad \boldsymbol{0} \in \mathbb{R}^{k+1}, \qquad i = 1, 2. \tag{26}$$

These minimization problems can be solved in $\mathcal{O}(k)$ arithmetic operations similarly as (21).

Using the factorization (25) and letting

$$(T_{k+1,k}^T T_{k+1,k} + \mu^{-1} I_k)\boldsymbol{y} = \boldsymbol{z}, \tag{27}$$

we can express the minimization problem (23) as

$$\min_{\boldsymbol{z}\in\mathbb{R}^k} \left\| \begin{bmatrix} I_k \\ \hat{W}_k^T \end{bmatrix} \boldsymbol{z} - T_{k+2,k+1}\boldsymbol{e}_1 \|\boldsymbol{b}\| \right\|, \tag{28}$$

and solve this problem in $\mathcal{O}(k)$ arithmetic floating point operations by computing the QR-factorization of the matrix using Givens rotations. Having computed the solution $\boldsymbol{z}_k^{(\mu)}$ of (28), we determine the solution $\boldsymbol{y}_k^{(\mu)}$ of (27) with the right-hand side $\boldsymbol{z} = \boldsymbol{z}_k^{(\mu)}$ by solving a least-squares problem analogous to (26). Finally, the solution of the minimization problem (22) is given by $\boldsymbol{x}_k^{(\mu)} = V_k \boldsymbol{y}_k^{(\mu)}$.

In summary, for each value of μ, the Galerkin equations (20) as well as the minimization problem (22) can be solved in $\mathcal{O}(k)$ arithmetic floating point operations. The approximation of the Tikhonov solution (4) determined by either one of these methods is expressed as $\boldsymbol{x}_k^{(\mu)} = V_k \boldsymbol{y}_k^{(\mu)}$. We note that the Lanczos decomposition (18) allows us to evaluate the functions $f_k(\mu)$ and $g_k(\mu)$ defined by (10) inexpensively according to

$$f_k(\mu) = \left\| T_{k+1,k}\boldsymbol{y}_k^{(\mu)} - \boldsymbol{e}_1\|\boldsymbol{b}\| \right\| - \delta, \qquad g_k(\mu) = \left\| T_{k+1,k}\boldsymbol{y}_k^{(\mu)} - \boldsymbol{e}_1\|\boldsymbol{b}\| \right\|^2 - \delta^2.$$

We turn to the determination of μ, k and an approximate solution $\boldsymbol{x}_k^{(\mu)}$ of the Tikhonov equations (3), such that (17) holds. The linear system (1) is assumed to be so large, that the dominating work of the solution method is the computation of the partial Lanczos tridiagonalization (18). The solution method seeks to keep the number of steps k of the Lanczos process small.

Our numerical method determines a sequence of values of $\mu_0, \mu_1, \mu_2, \ldots$, of the reqularization parameter μ before terminating. We would like these values to satisfy

$$\mu_j \leq \mu_*, \qquad j = 0, 1, 2, \ldots\ , \tag{29}$$

where the value μ_* is characterized by the requirement that equation (6) holds for $\mu = \mu_*$. The inequalities (29) avoid that we have to solve intermediate underregularized Tikhonov equations (3) with $\mu = \mu_j \gg \mu_*$. Underregularized Tikhonov equations can be severely ill-conditioned and their solution may require the number of Lanczos steps k to be larger than the solution of the Tikhonov equations with $\mu = \mu_*$. Thus, the solution of intermediate underregularized Tikhonov equations may increase the computational work unnecessarily.

We choose the initial value $\mu_0 = 0$ and apply a Newton step to the equation $g(\mu) = 0$ to determine the next value $\mu_1 := -g(0)/g'(0)$, where we define

$$g(0) := \lim_{\mu \searrow 0} g(\mu) = \|\boldsymbol{b}\|^2 - \delta^2, \qquad g'(0) := \lim_{\mu \searrow 0} g'(\mu) = -2\|A\boldsymbol{b}\|^2.$$

The vector $A\boldsymbol{b}$ is computed during the determination of the partial Lanczos tridiagonalization (18) of A with $k = 1$. Thus, the computation of $g(0)$ and $g'(0)$ does not require the evaluation of additional matrix-vector products with the matrix A. Since $g(\mu)$ is decreasing and convex, we have $\mu_1 \leq \mu_*$.

Having determined μ_1 and the partial Lanczos tridiagonalization (18) with $k = 1$ as described, we check whether $|f_1(\mu_1)| \leq \varepsilon_\delta \delta$, because in this case the vector $\boldsymbol{x}_1^{(\mu_1)}$ satisfies (17) and is an acceptable approximate solution of (1).

When $\boldsymbol{x}_1^{(\mu_1)}$ is not acceptable, we let $j := 1$ and increase k. Corollary 2.3 suggests that k be increased until

$$\frac{\sqrt{\mu_j}}{2}\|\boldsymbol{r}_k^{(\mu_j)}\| \leq \varepsilon_\delta \delta. \tag{30}$$

For each value of k, we check whether

$$|f_k(\mu_j)| \leq \varepsilon_\delta \delta, \tag{31}$$

because this inequality implies that $\boldsymbol{x}_k^{(\mu_j)}$ is an acceptable approximate solution of (1) and the computations can be terminated.

The norm of the residual vector $\boldsymbol{r}_k^{(\mu_j)}$ depends on whether $\boldsymbol{x}_k^{(\mu_j)}$ is computed by solving the Galerkin equations (20) or the minimization problem (22). The latter approach gives a residual vector of smaller or equal norm as the former. Therefore, the inequality (30) for many problems is satisfied for a smaller value of k when the minimization problem (22) is solved than when the Galerkin equations (20) are solved. This is illustrated by a numerical example in Section 4. We therefore favor the computation of $\boldsymbol{x}_k^{(\mu_j)}$ by solving the minimization problem (22).

When (30) is satisfied, but (31) is not, we replace μ_j by a larger value. The new value, denoted by μ_{j+1}, is determined by the secant method. This method requires two initial values μ_j and μ_{j-1}. The function $g(\mu)$ is decreasing and convex and the secant method, when applied to the solution of $g(\mu) = 0$, would give monotonic and superlinear convergence. In particular, the computed approximations μ_j, $j = 0, 1, 2, \ldots$, of μ_* would all satisfy $\mu_j \leq \mu_*$.

However, as already noted the function $g(\mu)$ is expensive to evaluate. We therefore apply the secant method to determine a zero of the function $g_k(\mu)$ instead. In view of Theorem 2.2, the difference $g_k(\mu) - g(\mu)$ is small when k is sufficiently large. Having determined the new approximation μ_{j+1} of μ_* in the manner described, we seek to satisfy the inequality (30), with j replaced by $j + 1$, by increasing the value of k if so required.

4. Numerical examples

This section presents two examples. The first one compares the method based on solving the minimization problem (22) with the symmetry-ignoring method for the solution of the Tikhonov equations (3) proposed in [2]. The second example compares the solution of the

Galerkin equations (20) and the solution of the minimization problem (22). All computations were carried out in Matlab on a personal computer with about 16 significant digits. We let $\varepsilon_\delta := 1$ in all examples. This corresponds to letting $\eta := 2$ in the method advocated in [2].

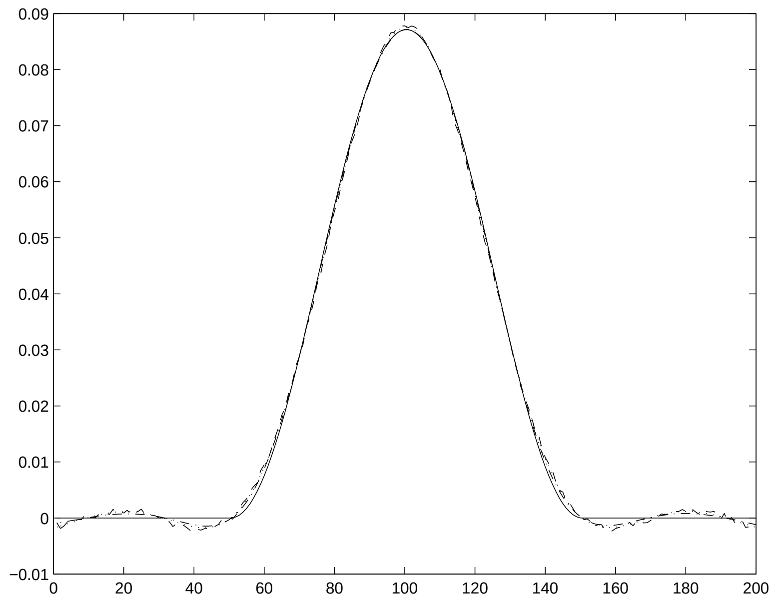

Figure 1. Example 4.1: Solution $\hat{\boldsymbol{x}}$ of error-free linear system (continuous curve), computed approximate solution $\boldsymbol{x}_4^{(\mu_4)}$ determined by the minimal residual method of the present paper (dashed curve), Tikhonov solution $\boldsymbol{x}^{(\mu_4)}$ (dash-dotted curve) and approximate solution $\widetilde{\boldsymbol{x}}$ determined by the method in [2] (dotted curve).

Example 4.1. We consider the test problem phillips from the Regularization Tools package by Hansen [7]. This problem is a Fredholm integral equation of the first kind

$$\int_{-6}^{6} k(s,t)x(t)dt = b(s), \qquad -6 \le s \le 6, \tag{32}$$

with solution

$$x(t) := \begin{cases} 1+\cos(\frac{\pi}{3}t), & |t| < 3, \\ 0, & \text{otherwise}, \end{cases}$$

kernel $k(s,t) := x(s-t)$ and right-hand side function

$$b(s) := (6-|s|)(1+\frac{1}{2}\cos(\frac{\pi}{3}s)) + \frac{9}{2\pi}\sin(\frac{\pi}{3}|s|).$$

We apply Matlab code provided in [7] to discretize (32) by a Galerkin method with orthonormal box functions and obtain the symmetric indefinite matrix $\bar{A} \in \mathbb{R}^{200\times 200}$ and the right-hand side $\bar{\boldsymbol{b}} \in \mathbb{R}^{200}$. The matrix $\bar{A}$ has many tiny singular values and its condition number $\kappa(\bar{A}) := \|\bar{A}\|\|\bar{A}^{-1}\|$ is $4.2 \cdot 10^7$.

Let the entries of the error vector $\boldsymbol{e} \in \mathbb{R}^{200}$ be normally distributed with zero mean and be normalized so that

$$\frac{\|\boldsymbol{e}\|}{\|\bar{\boldsymbol{b}}\|} = 1 \cdot 10^{-3}. \tag{33}$$

The matrix A and the contaminated right-hand side $\boldsymbol{b}$ in (1) are obtained by scaling the matrix $\bar{A}$ and the contaminated right-hand $\bar{\boldsymbol{b}}+\boldsymbol{e}$ to satisfy the conditions $\|\boldsymbol{b}\| = 1$ and $\|A\boldsymbol{b}\| = 2$. The purpose of this scaling is to make our results comparable with those presented in [2].

The minimal residual method of the present paper, which solves (22), required 4 steps of the Lanczos process applied to the matrix A with initial vector $\boldsymbol{b}$ before determining the value $\mu_4 = 2.5 \cdot 10^2$ and the approximate solution $\boldsymbol{x}_4^{(\mu_4)}$ of (1). The latter satisfies (8). Thus, the method computed the partial Lanczos tridiagonalization (18) with $k = 4$ and evaluated 4 matrix-vector products with the matrix A.

The Matlab program phillips determines a vector, denoted by $\hat{\boldsymbol{x}}$, that is a scaled discretization of the solution $x(t)$ of (32). We obtain

$$\|\boldsymbol{x}_4^{(\mu_4)} - \hat{\boldsymbol{x}}\| = 1.6 \cdot 10^{-2}, \quad \|\boldsymbol{x}^{(\mu_4)} - \hat{\boldsymbol{x}}\| = 9.8 \cdot 10^{-3}, \quad \|\boldsymbol{x}_4^{(\mu_4)} - \boldsymbol{x}^{(\mu_4)}\| = 8.4 \cdot 10^{-3},$$

where $\boldsymbol{x}^{(\mu_4)}$ is the solution (4) of the Tikhonov equations (3) with $\mu = \mu_4$.

For comparison, we note that the method presented in [2] required the evaluation of 8 matrix-vector products with the matrix A to determine an approximate solution $\widetilde{\boldsymbol{x}}$ of (1) that satisfies (8). We have

$$\|\widetilde{\boldsymbol{x}} - \hat{\boldsymbol{x}}\| = 1.3 \cdot 10^{-2}, \quad \|\widetilde{\boldsymbol{x}} - \boldsymbol{x}^{(\mu_4)}\| = 3.9 \cdot 10^{-3}, \quad \|\widetilde{\boldsymbol{x}} - \boldsymbol{x}_4^{(\mu_4)}\| = 6.0 \cdot 10^{-3}.$$

Thus, the computed approximate solutions $\boldsymbol{x}_4^{(\mu_4)}$ and $\widetilde{\boldsymbol{x}}$ of (1) are of about the same distance form $\hat{\boldsymbol{x}}$, but the determination of $\boldsymbol{x}_4^{(\mu_4)}$ required only half the number of matrix-vector product evaluations with A. The vectors $\hat{\boldsymbol{x}}$, $\boldsymbol{x}_4^{(\mu_4)}$, $\boldsymbol{x}^{(\mu_4)}$ and $\widetilde{\boldsymbol{x}}$ are displayed in Figure 1. □

In the above example the symmetry-exploiting minimal residual method, which solves (22), requires half the number of matrix-vector product evaluations than the symmetry-ignoring scheme described in [2]. Although this has been observed in several numerical experiments, there is no guarantee that the reduction in matrix-vector product evaluations is 50% because the methods use different criteria to determine how many matrix-vector products to evaluate and to select a value of the regularization parameter.

Example 4.2. Consider the blur- and noise-free image shown in Figure 2(a). This image is made available by Nagy [9]. It is represented by 256×256 pixels. The pixel values range from 0 to 255 and represent gray levels. They are stored row-wise in the vector $\hat{\boldsymbol{x}} \in \mathbb{R}^{256^2}$. We assume that only a contaminated version of this image is available, and we would like to determine an approximation of $\hat{\boldsymbol{x}}$. Let the matrix $A \in \mathbb{R}^{256^2 \times 256^2}$ represent a discretized blurring operator; it is the Kronecker product of a symmetric Toeplitz matrix $T = [t_{ij}] \in \mathbb{R}^{256 \times 256}$ with itself, where

$$t_{ij} = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(i-j)^2}{2\sigma^2}\right), & |i-j| \le \rho, \\ 0, & \text{otherwise}, \end{cases} \tag{34}$$

with $\sigma := 3.5$ and $\rho := 12\sigma$. The matrix A is of ill-determined rank. The blurred, but noise-free, image is given by $A\hat{\boldsymbol{x}}$. Let the error vector $\boldsymbol{e}$ have normally distributed random entries with zero mean and be normalized so that (33) holds. The vector $\boldsymbol{e}$ represents noise in the available image $\boldsymbol{b} := A\hat{\boldsymbol{x}} + \boldsymbol{e}$, which is displayed in Figure 2(b). The value of δ is given by (2).

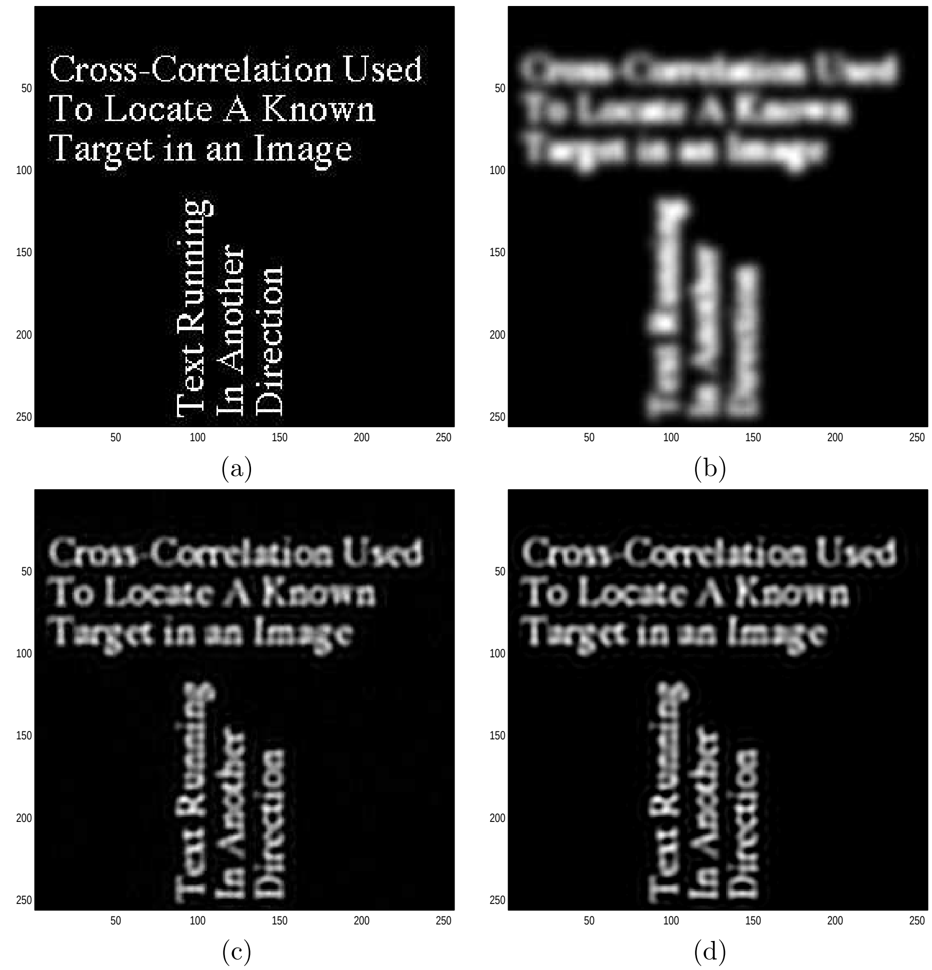

Figure 2. Example 4.2: (a) Blur- and noise-free image, (b) available image contaminated by blur and noise, (c) image restored by solving the minimization problem (23) with $k = 23$, (d) image restored by solving the Galerkin equations (20) with $k = 45$.

The minimal residual method, which solves (22), determined the Lanczos decomposition (18) with $k = 23$, before exiting with $\mu_{23} = 5.0 \cdot 10^4$ and the approximate solution $\boldsymbol{x}_{23}^{(\mu_{23})}$ of (1). The latter satisfies (8). The vector $\boldsymbol{x}_{23}^{(\mu_{23})}$ represents the restored image shown in Figure 2(c).

When instead an approximate solution of the Tikhonov equations was determined by solving the Galerkin equations (20), a partial Lanczos tridiagonalization (18) with $k = 45$ was required. Specifically, this approach determined the value $\mu_{45} = 3.2 \cdot 10^4$ of the regularization parameter and the vector $\boldsymbol{x}_{45}^{(\mu_{45})}$, which satisfies (8). The restored image that corresponds to the approximate solution $\boldsymbol{x}_{45}^{(\mu_{45})}$ of (1) is shown in Figure 2(d).

We note that the vectors $\boldsymbol{x}_{23}^{(\mu_{23})}$ and $\boldsymbol{x}_{45}^{(\mu_{45})}$ represent restored images of about the same quality, however, the number of matrix-vector product evaluations required by the method based on the solution of the minimization problem (22) is about half the number needed by the method based on the solution of the Galerkin equations (20). $\Box$

REFERENCES

1. Å. Björck, Numerical Methods for Least Squares Problems, SIAM, Philadelphia, 1996.
2. D. Calvetti and L. Reichel, Tikhonov regularization of large linear problems, BIT, to appear.
3. D. Calvetti, L. Reichel, F. Sgallari, and G. Spaletta, A regularizing Lanczos iteration method for underdetermined linear systems, J. Comput. Appl. Math., 115 (2000), pp. 101–120.
4. D. Calvetti, L. Reichel, and Q. Zhang, Iterative solution methods for large linear discrete ill-posed problems, Applied and Computational Control, Signals and Circuits, 1 (1999), pp. 313–367.
5. H. W. Engl, M. Hanke, and A. Neubauer, Regularization of Inverse Problems, Kluwer, Dordrecht, 1996.
6. A. Frommer and P. Maass, Fast CG-based methods for Tikhonov-Phillips regularization, SIAM J. Sci. Comput., 20 (1999), pp. 1831–1850.
7. P. C. Hansen, Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems, Numer. Algorithms, 6 (1994), pp. 1–35. Software is available in Netlib at the web site http://www.netlib.org.
8. P. C. Hansen, Rank Deficient and Discrete Ill-Posed Problems, SIAM, Philadelphia, 1998.
9. J. G. Nagy, RestoreTools: An object oriented Matlab package for image restoration. Available at the web site `http://www.mathcs.emory.edu/~nagy/`.