Invertible smoothing preconditioners for linear discrete ill-posed problems

D. Calvetti [a,1], L. Reichel [b,2], A. Shuibi [c,2]

[a] *Department of Mathematics, Case Western Reserve University, Cleveland, OH 44106, USA. E-mail:* `dxc57@po.cwru.edu`.

[b] *Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA. E-mail:* `reichel@math.kent.edu`.

[c] *Department of Mathematical Sciences, DePaul University, Chicago, IL 60614, USA. E-mail:* `ashuibi@condor.depaul.edu`.

Abstract

The solution of large linear discrete ill-posed problems by iterative methods has recently received considerable attention. This paper presents invertible smoothing preconditioners which are well suited for use with the GMRES, RRGMRES and LSQR methods.

Key words: Ill-posed problem, iterative method, GMRES, RRGMRES, LSQR, preconditioning

1 Introduction

The present paper is concerned with the solution of linear systems of equations

$$A\boldsymbol{x} = \boldsymbol{b}, \qquad A \in \mathbb{R}^{n\times n}, \qquad \boldsymbol{x}, \boldsymbol{b} \in \mathbb{R}^n, \tag{1}$$

where A is a matrix of ill-determined rank. In particular, the matrix A has many singular values of different orders of magnitude close to the origin, it is ill-conditioned and may be singular. Matrices of ill-determined rank arise from the discretization of linear ill-posed problems, such as Fredholm integral equations of the first kind with a smooth kernel. Linear systems of equations with such a matrix are often referred to as linear discrete ill-posed problems.

[1] Research supported in part by NSF grant DMS-0107841.
[2] Research supported in part by NSF grant DMS-0107858.

The right-hand side vector $\boldsymbol{b}$ in linear discrete ill-posed problems that arise in applications typically is contaminated by an error $\boldsymbol{e} \in \mathbb{R}^n$, which may stem from measurement or discretization errors. Let $\hat{\boldsymbol{b}}$ be the unknown error-free vector associated with $\boldsymbol{b}$, i.e.,

$$\boldsymbol{b} = \hat{\boldsymbol{b}} + \boldsymbol{e}, \tag{2}$$

and assume that the linear system of equations with the unknown error-free right-hand side

$$A\boldsymbol{x} = \hat{\boldsymbol{b}} \tag{3}$$

is consistent. The available linear system (1) is not required to be consistent.

We would like to solve (3) and denote the desired solution by $\hat{\boldsymbol{x}}$. If A is singular, then $\hat{\boldsymbol{x}}$ typically is the least-squares solution of minimal Euclidean norm of (3). Since the right-hand side $\hat{\boldsymbol{b}}$ is not available, we seek to determine an approximation of $\hat{\boldsymbol{x}}$ by computing an approximate solution of the available linear system of equations (1). Let $A^\dagger$ denote the Moore-Penrose pseudoinverse of A. Because of the error $\boldsymbol{e}$ in $\boldsymbol{b}$ and the severe ill-conditioning of the matrix A, the least-squares solution of minimal Euclidean norm of (1), given by $A^\dagger\boldsymbol{b}$, generally is not a meaningful approximation of $\hat{\boldsymbol{x}}$.

A popular approach to determining an approximation of $\hat{\boldsymbol{x}}$ is to apply a few, say k, steps of the conjugate gradient method to the normal equations

$$A^T A\boldsymbol{x} = A^T\boldsymbol{b} \tag{4}$$

associated with (1). We note that when k is too large, the computed iterate $\boldsymbol{x}_k$ is a poor approximation of $\hat{\boldsymbol{x}}$ because it is severely contaminated by propagated errors due to the error $\boldsymbol{e}$ in $\boldsymbol{b}$ and round-off errors introduced during the computation of $\boldsymbol{x}_k$. On the other hand, when k is too small, a better approximation of $\hat{\boldsymbol{x}}$ can be determined by increasing k. Truncated iteration with the conjugate gradient method is analyzed by Hanke [11]. Insightful discussions and computed examples can also be found in [10,14,16].

It is well known that the performance of the conjugate gradient method can be improved by the use of a right smoothing preconditioner; see Hanke [12,13] and Hanke and Hansen [14]. Popular preconditioners for linear discrete ill-posed problems include the A-weighted pseudoinverse

$$M = (I - (A(I - \widetilde{L}^\dagger\widetilde{L}))^\dagger A)\widetilde{L}^\dagger \tag{5}$$

of the bidiagonal matrix

$$\widetilde{L} = \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \end{bmatrix} \in \mathbb{R}^{(n-1)\times n} \tag{6}$$

and of the tridiagonal matrix

$$\widetilde{L} = \begin{bmatrix} -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \end{bmatrix} \in \mathbb{R}^{(n-2)\times n}. \tag{7}$$

The A-weighted pseudoinverse was introduced by Eldén [9] and has subsequently been discussed by Hanke [12,13], Hanke and Hansen [14], and Hansen [16, Section 2.3.2]. Since the matrices (6) and (7) are discrete finite difference operators, the preconditioners (5) approximate integration operators and therefore can be thought of as smoothing operators.

LSQR is an implementation of the conjugate gradient method applied to the normal equation (4) based on partial Lanczos bidiagonalization of A; see, e.g., [18,19]. The Matlab function plsqr in the Regularization Tools program package by Hansen [15] implements the preconditioned LSQR method; roughly, the matrices A and A^T in (4) are replaced by AM and $M^T A^T$, respectively, where M is given by (5) with $\widetilde{L}$ defined by (6) or (7). We remark that application of plsqr does not require the matrix M to be explicitly formed; instead a user provides $\widetilde{L}$ and a matrix $W \in \mathbb{R}^{n\times p}$, whose columns span the null space of $\widetilde{L}$. For instance, when $\widetilde{L}$ is given by (6), we let $W = [1, 1, \ldots, 1]^T$, and when $\widetilde{L}$ is defined by (7), we let W have the columns $[1, 1, \ldots, 1]^T$ and $[1, 2, \ldots, n]^T$.

The GMRES method by Saad and Schultz [20] is one of the most popular iterative methods for the solution of linear systems of equations that arise from the discretization of well-posed problems. An analysis of truncated iteration with the GMRES method applied to linear discrete ill-posed problems of the form (1) is presented in [7], and computed examples reported there and in [4–6] show that the GMRES method, and the closely related Range Restricted GMRES (RRGMRES) method, may yield better approximations of $\hat{\boldsymbol{x}}$ with less computational work than the conjugate gradient method applied to the normal equations (4); see also [3] for a discussion of RRGMRES. We therefore are interested in finding right preconditioners that can be used together with the GMRES and RRGMRES methods. Note that the matrices M defined by (5) with $\widetilde{L}$ given by (6) and (7) cannot be used with the GMRES and RRGMRES methods because they are not square.

It is the purpose of the present paper to introduce invertible preconditioners that are closely related to the non-square preconditioners M discussed above. The new preconditioners are described in Section 2. Section 3 comments on the LSQR, GMRES, and RRGMRES iterative methods, which are used in the computed examples presented in Section 4, and discusses the choice of preconditioner. Concluding remarks can be found in Section 5.

We remark that the matrices (6) and (7) are popular regularization operators for Tikhonov regularization, where instead of solving (1), one computes a solution of the minimization problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n}\{\|A\boldsymbol{x}-\boldsymbol{b}\|^2+\mu\|\widetilde{L}\boldsymbol{x}\|^2\}$$

for a positive regularization parameter μ. Here and throughout this paper $\|\cdot\|$ denotes the Euclidean vector norm or the induced matrix norm. The choice of regularization operator $\widetilde{L}$ has received considerable attention; see, e.g., [2,9,10,14,16] and references therein. It is generally beneficial to use a regularization operator $\widetilde{L}$, such that the seminorm

$$\|\boldsymbol{x}\|_{\widetilde{L}}=\|\widetilde{L}\boldsymbol{x}\| \tag{8}$$

is an appropriate measure of the desired solution $\hat{\boldsymbol{x}}$.

2 Invertible smoothing preconditioners

This section describes invertible smoothing preconditioners that are closely related to the noninvertible preconditioners discussed in the previous section. We first consider preconditioners related to the matrix (6). Introduce the invertible upper bidiagonal matrix

$$L=\begin{bmatrix}1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ & & & & \alpha\end{bmatrix}\in\mathbb{R}^{n\times n}, \tag{9}$$

where the last diagonal entry α is a small positive constant. In all computed examples we let $\alpha=1\cdot10^{-8}$. We found the performance of the preconditioner not to be very sensitive to the value of α. Since L is invertible, the expression (5) simplifies to

$$M = L^{-1}. \tag{10}$$

We propose that M be used as a right preconditioner for the GMRES method for certain linear discrete ill-posed problems. Thus, we apply GMRES to the solution of the preconditioned linear system of equations

$$AM\boldsymbol{y} = \boldsymbol{b}.$$

Denote the kth iterate determined by GMRES applied to the solution of this system by $\boldsymbol{y}_k$. Then

$$\boldsymbol{x}_k = M\boldsymbol{y}_k \tag{11}$$

is an approximate solution of (1). More details on the GMRES method are provided in Section 3.

Reflecting the matrix (9) in its diagonal and anti-diagonal yields the invertible lower bidiagonal matrix

$$L = \begin{bmatrix} \alpha & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}. \tag{12}$$

This matrix also will be used to determine a preconditioner M using (10).

We turn to invertible preconditioners that are related to the noninvertible matrix (7). Appending new first and last rows to the matrix (7) yields the symmetric positive definite matrix

$$L = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \in \mathbb{R}^{n \times n}. \tag{13}$$

The associated preconditioner is defined by (10). Halving the last diagonal entry of (13) yields the symmetric positive definite matrix

$$L = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}, \tag{14}$$

and an associated preconditioner defined by (10). The reflection of the matrix (14) in its antidiagonal yields

$$L = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \in \mathbb{R}^{n \times n}, \tag{15}$$

and we also will discuss the associated preconditioner given by (10). Since the matrices L defined by (9), (12) and (13)-(15) are (scaled) finite difference operators, their inverses approximate integration operator. They therefore can be thought of as a smoothers.

The matrices (9) and (12) are bidiagonal and therefore the approximate solution $\boldsymbol{x}_k$ can be computed from the kth iterate $\boldsymbol{y}_k$, cf. (11), by solving

$$L\boldsymbol{x}_k = \boldsymbol{y}_k. \tag{16}$$

The latter requires only $\mathcal{O}(n)$ arithmetic floating point operations. When, instead, the preconditioner M is defined by one of the matrices (13)-(15), we first compute the Cholesky factorization of L and then solve (16) for the approximate solution $\boldsymbol{x}_k$, when the iterate $\boldsymbol{y}_k$ is available. Both the computation of the Cholesky factorization and the solution of (16) can be carried out in only $\mathcal{O}(n)$ arithmetic floating point operations.

The choice of an appropriate preconditioner is not an easy task. Some guidelines that we have found helpful will be discussed in Section 3. For certain problems it may be attractive to use other modifications of the matrices (6) and (7) than those discussed above to define preconditioners (10). We focus on the preconditioners introduced above because of their good performance for certain problems, and because they are inexpensive to apply.

3 Iterative methods

We review some properties of the LSQR, GMRES, and RRGMRES iterative methods used in the numerical experiments reported in the next section. For all iterative methods, we use the initial iterate $\boldsymbol{x}_0 = \mathbf{0}$ for the unpreconditioned linear systems and $\boldsymbol{y}_0 = \mathbf{0}$ for the preconditioned linear systems.

Let L be any one of the invertible matrices (9), (12) or (13)-(15), and introduce

$$\hat{A} = AL^{-1}. \tag{17}$$

The kth iterate $\boldsymbol{y}_k$ determined by the right-preconditioned LSQR method satisfies

$$\|\hat{A}\boldsymbol{y}_k - \boldsymbol{b}\| = \min_{\boldsymbol{y}\in\mathbb{K}_k(\hat{A}^T\hat{A},\hat{A}^T\boldsymbol{b})} \|\hat{A}\boldsymbol{y} - \boldsymbol{b}\|, \qquad \boldsymbol{y}_k \in \mathbb{K}_k(\hat{A}^T\hat{A}, \hat{A}^T\boldsymbol{b}),$$

where

$$\mathbb{K}_k(\hat{A}^T\hat{A}, \hat{A}^T\boldsymbol{b}) = \text{span}\{\hat{A}^T\boldsymbol{b}, (\hat{A}^T\hat{A})\hat{A}^T\boldsymbol{b}, (\hat{A}^T\hat{A})^2\hat{A}^T\boldsymbol{b}, \ldots, (\hat{A}^T\hat{A})^{k-1}\hat{A}^T\boldsymbol{b}\} \tag{18}$$

is a Krylov subspace.

It is known that the norm of successive iterates is monotonically increasing, i.e., $\|\boldsymbol{y}_k\| \geq \|\boldsymbol{y}_{k-1}\|$ for all $k \geq 1$, see [17] for a proof, and it follows from (16) that the computed approximate solutions $\boldsymbol{x}_k$ of (1) satisfy

$$\|L\boldsymbol{x}_k\| \geq \|L\boldsymbol{x}_{k-1}\|, \qquad k \geq 1. \tag{19}$$

This inequality suggests that the preconditioner M should be chosen so that the L-norm

$$\|\boldsymbol{x}\|_L = \|L\boldsymbol{x}\| \tag{20}$$

is a natural measure for the size of the desired solution $\hat{\boldsymbol{x}}$ of the problem at hand, where M and L are related by (10).

It is interesting to compare the L-norm defined by an invertible preconditioner to the seminorm (8) associated with a related noninvertible preconditioner. We obtain for L and $\widetilde{L}$ defined by (9) and (6), respectively, that

$$\|L\boldsymbol{x}\|^2 = \|\widetilde{L}\boldsymbol{x}\|^2 + \alpha^2 x_n^2,$$

where $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T$. Similarly, when L is given by (12), we have

$$\|L\boldsymbol{x}\|^2 = \|\widetilde{L}\boldsymbol{x}\|^2 + \alpha^2 x_1^2.$$

Since α^2 is "tiny," the difference between $\|L\boldsymbol{x}\|$ and $\|\widetilde{L}\boldsymbol{x}\|$ typically is negligible for either one of the invertible preconditioners (9) and (12). We therefore expect the invertible preconditioners defined by (9) and (12) to perform well when the seminorm (8) is an appropriate measure of the size of the computed solution, i.e., when the regularization operator $\widetilde{L}$ is suitable for Tikhonov regularization. The performance of the preconditioners determined by (9) and (12) is illustrated in Section 4.

Let as usual $\hat{\boldsymbol{x}}$ denote the desired solution of (3). Numerical experiments indicate that it often is beneficial to choose the invertible preconditioner M, such that most of the entries of the vector $L\hat{\boldsymbol{x}}$ are of about the same magnitude, where L is given by (10). For instance, when the entries of $\hat{\boldsymbol{x}}$ grow or decrease about linearly with their index number, the matrices L given by (9) or (12) often define suitable preconditioners, because they approximate the first difference operator (6). For these matrices, all entries of $L\hat{\boldsymbol{x}}$, except for the first or last one, are of about the same size.

We turn to the matrix $\widetilde{L}$ given by (7) and the associated invertible matrices L. Letting L be defined by (13), (14), and (15) yields, in order,

$$\|L\boldsymbol{x}\|^2 = \|\widetilde{L}\boldsymbol{x}\|^2 + (2x_1 - x_2)^2 + (2x_n - x_{n-1})^2, \tag{21}$$
$$\|L\boldsymbol{x}\|^2 = \|\widetilde{L}\boldsymbol{x}\|^2 + (2x_1 - x_2)^2 + (x_n - x_{n-1})^2, \tag{22}$$
$$\|L\boldsymbol{x}\|^2 = \|\widetilde{L}\boldsymbol{x}\|^2 + (x_1 - x_2)^2 + (2x_n - x_{n-1})^2. \tag{23}$$

The rows added to the matrix $\widetilde{L}$ to obtain L generally should be chosen so that $|\|L\hat{\boldsymbol{x}}\|^2 - \|\widetilde{L}\hat{\boldsymbol{x}}\|^2|$ is small for the desired solution $\hat{\boldsymbol{x}}$. For instance, if $\hat{\boldsymbol{x}}$ is the discretization of a solution $x(t)$ of a Fredholm integral equation of the first kind on an interval, then the rows added to $\widetilde{L}$ to obtain L should be chosen depending on the behavior of $x(t)$ at the boundary points of the interval.

Example 1 *Let $\hat{\boldsymbol{x}}$ be a discretization of* $\sin(t)$ *on a uniform mesh in the open interval* $(0, \pi)$*, i.e.,*

$$\hat{\boldsymbol{x}} = [\sin(h), \sin(2h), \ldots, \sin(\pi - h)]^T \in \mathbb{R}^n, \qquad h = \frac{\pi}{n+1}. \tag{24}$$

Let L be defined by (13). A typical entry of $\widetilde{L}\hat{\boldsymbol{x}}$ is $\mathcal{O}(h^2)$, while the first and last entries of $L\hat{\boldsymbol{x}}$ are $\mathcal{O}(h^3)$. Thus, when h is small, the first and last entries of $L\hat{\boldsymbol{x}}$ are smaller than many of the other entries. Moreover, $|\|L\hat{\boldsymbol{x}}\|^2 - \|\widetilde{L}\hat{\boldsymbol{x}}\|^2|$ is bounded by $\mathcal{O}(h^6)$, showing that the L-norm (20) and the seminorm (8) are close for the vector (24) when h is small.

Note that when L is given by (14) or (15), we have $\|L\hat{\boldsymbol{x}}\|^2 - \|\widetilde{L}\hat{\boldsymbol{x}}\|^2 = \mathcal{O}(h^2)$. This suggests that if $\widetilde{L}$ is a suitable regularization operator for Tikhonov regularization of (1), then the preconditioner determined by (13) may be more appropriate than the preconditioners defined by (14) and (15) when the desired solution is of the form (24). Computed examples that illustrate this can *be found in Section 4.*

Example 2 *Let* $\hat{\boldsymbol{x}}$ *be a discretization of* $\sin(t)$ *on a uniform mesh on the interval* $(0, \pi/2]$, *i.e.,*

$$\hat{\boldsymbol{x}} = [\sin(h), \sin(2h), \ldots, \sin(\frac{\pi}{2})]^T \in \mathbb{R}^n, \qquad h = \frac{\pi}{2n}. \tag{25}$$

Let L *be defined by (14). Then it follows from (22) that* $\|L\hat{\boldsymbol{x}}\|^2 - \|\widetilde{L}\hat{\boldsymbol{x}}\|^2 = \mathcal{O}(h^4)$. *We note that when* L *is given by (13) the corresponding difference is* $\mathcal{O}(1)$. *This suggests that the preconditioner determined by (14) may be more appropriate than the preconditioner defined by (13) for problems with a solution of the form (25); see Section 4 for a computed example.*

Example 3 *The effectiveness of the preconditioner defined by (13) can be improved in Example 2 if the linear function* $\ell(t) = \frac{2}{\pi}t$ *is subtracted from* $\sin(t)$. *The vector corresponding to (25) is given by*

$$\hat{\boldsymbol{x}} = [\sin(h) - \frac{2h}{\pi}, \sin(2h) - \frac{4h}{\pi}, \ldots, 0]^T \in \mathbb{R}^n, \qquad h = \frac{\pi}{2n}. \tag{26}$$

Note that the last entry of (26) vanishes and the first entry is small when h *is. We remark that if (25) is the desired solution of a linear system of equations (3), then (26) is the solution of a system of equations with a modified right-hand side. This is illustrated in Example 5 below.*

Let L *be defined by (13). The first and last entries of the vector* $L\hat{\boldsymbol{x}}$ *are* $\mathcal{O}(h^3)$ *and* $\mathcal{O}(h)$, *respectively, and it follows that* $\|L\hat{\boldsymbol{x}}\|^2 - \|\widetilde{L}\hat{\boldsymbol{x}}\|^2 = \mathcal{O}(h^2)$. *For* h *small, this difference is much smaller than the lower bound of* $\mathcal{O}(1)$ *for the vector (25) of Example 2. We remark that the linear term in (26) does not affect the entries of* $\widetilde{L}\hat{\boldsymbol{x}}$, *which therefore are* $\mathcal{O}(h^2)$ *as in Example 2.*

When applying the preconditioned LSQR method with any one of the invertible preconditioners introduced to the solution of large linear systems of equations, the dominating computational effort for computing the kth approximate solution $\boldsymbol{x}_k$ is the evaluation of $2k+1$ matrix-vector products with the matrices A or A^T.[3] We note that when applying a noninvertible preconditioner defined by a matrix $\widetilde{L}$ with a null space of dimension p, the computation of $\boldsymbol{x}_k$ requires $p+1$ additional matrix-vector product evaluations for the computation of AW and the initial residual vector $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$, where the columns of W contain a basis of the null space of $\widetilde{L}$; see Section 1.[4] For details on

[3] This operation count is for the implementation plsqr in [15], which follows the description in [18]. One matrix-vector product evaluation can be saved when certain computed auxiliary quantities are not required.

[4] The vector $\boldsymbol{x}_0$ is determined by plsqr and is generally nonvanishing when $\widetilde{L}$ has a nontrivial null space.

preconditioning with a noninvertible preconditioner, we refer to Björck and Eldén [2] or Hansen [16, Section 2.3].

We turn to the preconditioned GMRES method. The kth iterate $\boldsymbol{y}_k$ determined by GMRES when applied to the solution of (1), using a right invertible preconditioner M, satisfies

$$\|\hat{A}\boldsymbol{y}_k - \boldsymbol{b}\| = \min_{\boldsymbol{y}\in\mathbb{K}_k(\hat{A},\boldsymbol{b})} \|\hat{A}\boldsymbol{y} - \boldsymbol{b}\|, \qquad \boldsymbol{y}_k \in \mathbb{K}_k(\hat{A}, \boldsymbol{b}),$$

where $\hat{A}$ is defined by (17), M and L are related by (10), and

$$\mathbb{K}_k(\hat{A}, \boldsymbol{b}) = \text{span}\{\boldsymbol{b}, \hat{A}\boldsymbol{b}, \hat{A}^2\boldsymbol{b}, \ldots, \hat{A}^{k-1}\boldsymbol{b}\}. \tag{27}$$

The kth approximate solution $\boldsymbol{x}_k$ of (1) is given by (11).

The norm of $\boldsymbol{y}_k = L\boldsymbol{x}_k$ is not guaranteed to increase monotonically with k. Nevertheless, for many problems it does, and the guidelines for the choice of M for preconditioned LSQR apply to preconditioned GMRES. When the matrix A is large, the dominating work for computing $\boldsymbol{x}_k$ is the evaluation of k matrix-vector products with the matrix A, provided that $\boldsymbol{x}_k = M\boldsymbol{y}_k$ can be computed fairly rapidly. The latter is the case for the preconditioners considered in the present paper.

Preconditioned RRGMRES [5] differs from preconditioned GMRES only in that the Krylov subspaces $\mathbb{K}_k(\hat{A}, \boldsymbol{b})$ used by the latter method are replaced by $\mathbb{K}_k(\hat{A}, \hat{A}\boldsymbol{b})$, $k = 1, 2, 3, \ldots$. Thus, the kth iterate $\boldsymbol{y}_k$ determined by preconditioned RRGMRES satisfies

$$\|\hat{A}\boldsymbol{y}_k - \boldsymbol{b}\| = \min_{\boldsymbol{y}\in\mathbb{K}_k(\hat{A},\hat{A}\boldsymbol{b})} \|\hat{A}\boldsymbol{y} - \boldsymbol{b}\|, \qquad \boldsymbol{y}_k \in \mathbb{K}_k(\hat{A}, \hat{A}\boldsymbol{b}).$$

Preconditioned RRGMRES can give a smaller error in the kth iterate than preconditioned GMRES. This is illustrated in Section 4. Suitable preconditioners for GMRES often are appropriate for RRGMRES as well. For large matrices A, the dominating computational work required to compute $\boldsymbol{x}_k$ is the evaluation of $k + 1$ matrix-vector products with the matrix A, provided that $\boldsymbol{x}_k$ can be determined rapidly from $\boldsymbol{y}_k$.

4 Numerical examples

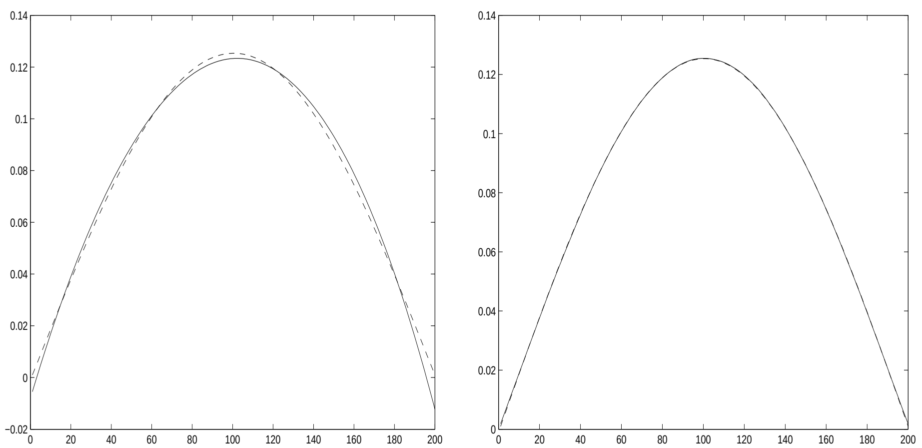
This section presents a few computed examples that illustrate the performance of the different preconditioners discussed. The examples also show that preconditioned GMRES and RRGMRES may determine approximate solutions of higher quality than preconditioned LSQR.

The examples are discrete linear ill-posed problems from the Regularization Tools package by Hansen [15]. In all examples, the desired solution $\hat{\boldsymbol{x}}$ is available, which we use to compute the error-free right-hand side $\hat{\boldsymbol{b}}$ in (3). The error vector $\boldsymbol{e}$ has normally distributed entries with zero mean. The variance of the entries is chosen so that $\boldsymbol{e}$ achieves the noise level $\|\boldsymbol{e}\|/\|\hat{\boldsymbol{b}}\| = 1 \cdot 10^{-3}$. All computations were carried out in Matlab with machine epsilon $2 \cdot 10^{-16}$.

We are interested in how good approximations $\boldsymbol{x}_k$ of $\hat{\boldsymbol{x}}$ the GMRES, RRGMRES and LSQR methods can determine with and without preconditioners. We therefore report the minimal error

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}_{k^*}\| = \min_{k\geq 0} \|\hat{\boldsymbol{x}} - \boldsymbol{x}_k\|$$

and the index $k^* \geq 0$ for which this error is achieved. The number of matrix-vector product evaluations with A or A^T required by the different methods to compute $\boldsymbol{x}_{k^*}$ is also tabulated. The iterative methods reorthogonalize the Krylov subspace bases.



(a) (b)

Fig. 1. Example 4: (a) Approximate solution $\boldsymbol{x}_3$ computed by unpreconditioned RRGMRES (continuous curve) and exact solution $\hat{\boldsymbol{x}}$ (dashed curve). (b) Approximate solution $\boldsymbol{x}_3$ computed by GMRES with an invertible preconditioner determined by (13) (continuous curve) and exact solution $\hat{\boldsymbol{x}}$ (dashed curve).

Example 4 *Consider the Fredholm integral equation of the first kind*

$$\int_0^{\pi} \exp(s\cos(t))x(t)dt = 2\frac{\sinh(s)}{s}, \qquad 0 \leq s \leq \frac{\pi}{2}, \tag{28}$$

which is discussed by Baart [1]. We discretize the integral equation by a Galerkin method with orthonormal box functions using the Matlab program baart *in the Regularization Tools package [15] to obtain a linear system of equations (1) with a nonsymmetric matrix $A \in \mathbb{R}^{200\times 200}$ and a scaled discrete approximation $\hat{\boldsymbol{x}}$ of the solution $x(t) = \sin(t)$, $0 \leq t \leq \pi$, of (28). The matrix A*

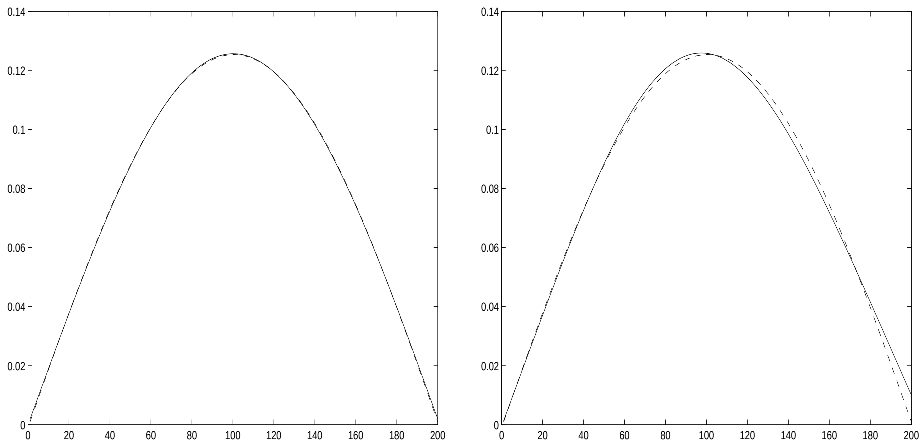


Fig. 2. Example 4: (a) Approximate solution $\boldsymbol{x}_3$ computed by LSQR with an invertible preconditioner defined by (13) (continuous curve) and exact solution $\hat{\boldsymbol{x}}$ (dashed curve). (b) Approximate solution $\boldsymbol{x}_1$ determined by LSQR with a noninvertible preconditioner determined by (7) (continuous curve) and exact solution $\hat{\boldsymbol{x}}$ (dashed curve).

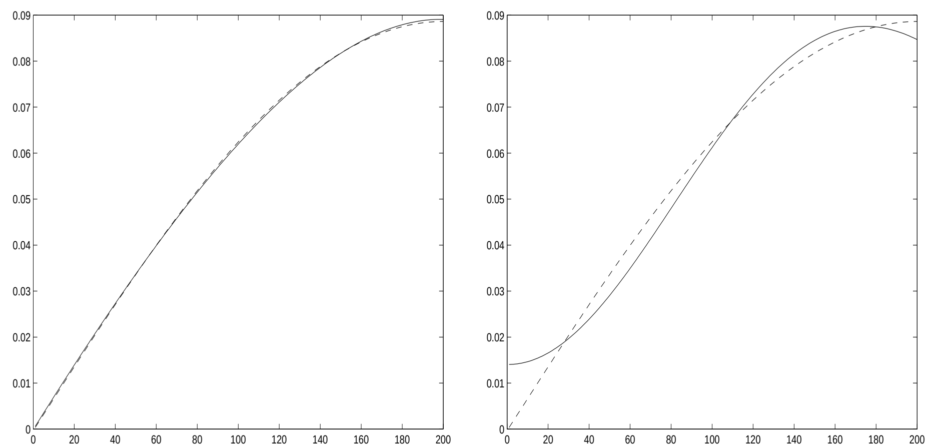
method	preconditioner defined by	smallest error $\|\hat{\boldsymbol{x}} - \boldsymbol{x}_{k^*}\|$	iterations k^*	matrix-vector products
GMRES	-	$6.0 \cdot 10^{-2}$	3	3
RRGMRES	-	$4.5 \cdot 10^{-2}$	3	4
LSQR	-	$1.5 \cdot 10^{-1}$	4	9
GMRES	(13)	$3.1 \cdot 10^{-3}$	3	3
RRGMRES	(13)	$3.3 \cdot 10^{-3}$	3	4
LSQR	(13)	$4.9 \cdot 10^{-3}$	3	7
GMRES	(15)	$5.8 \cdot 10^{-2}$	4	4
RRGMRES	(15)	$6.4 \cdot 10^{-2}$	4	5
LSQR	(15)	$7.1 \cdot 10^{-2}$	4	9
LSQR	(7)	$3.5 \cdot 10^{-2}$	1	6

Table 1
Example 4: Comparison of iterative methods and preconditioners.

is of ill-determined rank; it has condition number $\kappa(A) = 5.2 \cdot 10^{18}$, *where* $\kappa(A) = \|A\|\|A^{-1}\|$. *Let* $\hat{\boldsymbol{b}} = A\hat{\boldsymbol{x}}$ *and define the right-hand side of (1) by (2).*

Table 1 compares preconditioned and unpreconditioned GMRES, RRGMRES and LSQR. The invertible preconditioner $M = L^{-1}$ *defined by the tridiagonal matrix (13) gives approximate solutions with the smallest errors. GMRES* *and RRGMRES can be seen to require less work and give higher accuracy than LSQR. The table also shows the performance of the invertible preconditioner defined by the tridiagonal matrix (15). Each one of the iterative methods requires more iterations and yields worse approximations of* $\hat{\boldsymbol{x}}$ *when the preconditioner is determined by (15) than by (13). When the preconditioner defined by (15) is replaced by the preconditioner determined by (14), the number of iterations required for computing the most accurate approximate solution with each iterative method remains the same, but the best approximation of* $\hat{\boldsymbol{x}}$ *achieved by each method deteriorates. We therefore have not tabulated the latter results. These numerical experiments with preconditioners based on different completions of the rectangular matrix (7) illustrate the importance of how this completion is carried out. We remark that the noninvertible preconditioner determined by (7) gives lower accuracy than the invertible preconditioner defined by (13) when used with LSQR.*

Figure 1(a) shows the computed solution that best approximates $\hat{\boldsymbol{x}}$ *when no preconditioner is used, as well as* $\hat{\boldsymbol{x}}$*. Similarly, Figure 1(b) displays the best approximation of* $\hat{\boldsymbol{x}}$ *determined by a preconditioned method. Figure 2(a) and (b) show the best computed approximations of* $\hat{\boldsymbol{x}}$ *determined by LSQR with preconditioners defined by (13) and (7), respectively.*



(a) (b)

Fig. 3. Example 5: (a) Approximate solution $\boldsymbol{x}_3$ computed by unpreconditioned RRGMRES (continuous curve) and exact solution $\hat{\boldsymbol{x}}$ (dashed curve). (b) Approximate solution $\boldsymbol{x}_3$ computed by unpreconditioned LSQR (continuous curve) and exact solution $\hat{\boldsymbol{x}}$ (dashed curve).

Example 5 *This example consider the Fredholm integral equation obtained by halving the ranges of* s *and* t *in (28), i.e., we consider*

$$\int_0^{\pi/2} \exp(s\cos(t))x(t)dt = 2\frac{\sinh(s)}{s}, \qquad 0 \leq s \leq \frac{\pi}{4}. \tag{29}$$

We generate a discretization of (29) as follows. First determine the discretiza-

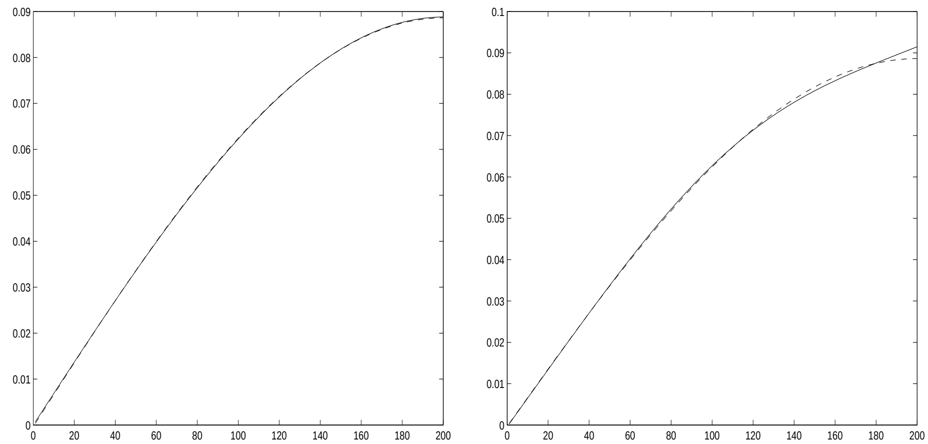
method	preconditioner defined by	smallest error $\|\hat{\boldsymbol{x}} - \boldsymbol{x}_{k^*}\|$	iterations k^*	matrix-vector products
GMRES	-	$1.2 \cdot 10^{-1}$	2	2
RRGMRES	-	$4.8 \cdot 10^{-3}$	3	4
LSQR	-	$5.4 \cdot 10^{-2}$	3	7
GMRES	SoLF	$8.1 \cdot 10^{-2}$	1	1
RRGMRES	SoLF	$4.8 \cdot 10^{-3}$	3	4
LSQR	SoLF	$1.9 \cdot 10^{-2}$	3	7
GMRES	(14)	$5.3 \cdot 10^{-3}$	3	3
RRGMRES	(14)	$4.7 \cdot 10^{-3}$	3	4
LSQR	(14)	$2.0 \cdot 10^{-3}$	2	5
GMRES	(13)	$3.9 \cdot 10^{-1}$	3	3
RRGMRES	(13)	$3.2 \cdot 10^{-1}$	3	4
LSQR	(13)	$3.9 \cdot 10^{-1}$	3	7
GMRES	SoLF+(13)	$4.5 \cdot 10^{-3}$	2	2
RRGMRES	SoLF+(13)	$4.6 \cdot 10^{-3}$	2	3
LSQR	SoLF+(13)	$3.7 \cdot 10^{-3}$	2	5
LSQR	(7)	$9.3 \cdot 10^{-3}$	1	6

Table 2
Example 5: Comparison of iterative methods and preconditioners.

tion $A'\boldsymbol{x}' = \boldsymbol{b}'$ *with* $A' \in \mathbb{R}^{400\times 400}$ *and* $\boldsymbol{x}' \in \mathbb{R}^{400}$ *of (28) by using the code* baart *from [15]. Then let* A *be the leading* 200×200 *principal submatrix of* A' *and let* $\hat{\boldsymbol{x}}$ *consist of the first* 200 *entries of* $\boldsymbol{x}'$. *The error-free right-hand side in (3) is determined by* $\hat{\boldsymbol{b}} = A\hat{\boldsymbol{x}}$. *The matrix* A *so defined is of ill-determined rank; it has condition number* $\kappa(A) = 9.6 \cdot 10^{17}$. *The vector* $\hat{\boldsymbol{x}}$ *is a scaled approximation of the solution* $x(t) = \sin(t)$, $0 \leq t \leq \pi/2$, *of (29). An important feature of this example is that, differently from Example 4,* $x(t)$ *does not vanishing at the right-hand side endpoint of the interval of integration.*

Table 2 compares preconditioned and unpreconditioned GMRES, RRGMRES and LSQR. The invertible preconditioner defined by the tridiagonal matrix (14) gives approximate solutions with among the smallest errors, while the preconditioner defined by (13) does not perform as well. This is the opposite of the situation in Example 4.

Assume that the values of the solution $x(t)$ *at the endpoints of the interval of integration are known. We then can subtract the linear function* $\ell(t) = \frac{\pi}{2}t$



(a) (b)

Fig. 4. Example 5: (a) Approximate solution $\boldsymbol{x}_2$ computed by LSQR with an invertible preconditioner defined by (14) (continuous curve) and exact solution $\hat{\boldsymbol{x}}$ (dashed curve). (b) Approximate solution $\boldsymbol{x}_1$ determined by LSQR with a noninvertible preconditioner determined by (7) (continuous curve) and exact solution $\hat{\boldsymbol{x}}$ (dashed curve).

from the solution $x(t)$ of (29) and solve the discretized integral equation with a suitably modified right-hand side for an approximation of $x(t) - \ell(t)$. This function vanishes at the endpoints of the interval of integration. This approach is in Table 2 referred to as "SoLF" for Subtraction of Linear Function. The table illustrates the performance of SoLF without preconditioner and with a preconditioner based on (13). SoLF combined with the preconditioner based on (13) can be seen to give among the highest accuracy with all iterative methods.

Table 2 illustrates the importance of completing the matrix (7) in a proper way to determine suitable invertible preconditioners. Moreover, the table indicates that removal of a linear function to change the boundary behavior of the solution of (1) can improve the quality of the computed solution.

Figure 3(a) shows the most accurate approximate solution computed without a preconditioner. It is determined by RRGMRES. For comparison Figure 3(b) displays the most accurate computed solution determined by unpreconditioned LSQR. Figure 4(a) and (b) show computed approximate solutions determined by preconditioned LSQR with preconditioners defined by (14) and (7), respectively.

Example 6 *We consider the Fredholm integral equation of the first kind*

$$\int_0^1 k(s,t)x(t)dt = \frac{1}{6}(s^3 - s), \qquad 0 \le s \le 1, \tag{30}$$

method	preconditioner defined by	smallest error $\|\hat{\boldsymbol{x}} - \boldsymbol{x}_{k^*}\|$	iterations k^*	matrix-vector products
GMRES	-	$1.1 \cdot 10^{-1}$	6	6
RRGMRES	-	$9.0 \cdot 10^{-2}$	10	11
LSQR	-	$9.1 \cdot 10^{-2}$	12	25
GMRES	(9)	$3.8 \cdot 10^{-3}$	8	8
RRGMRES	(9)	$5.4 \cdot 10^{-3}$	12	13
LSQR	(9)	$6.8 \cdot 10^{-3}$	8	17
GMRES	(12)	$3.4 \cdot 10^{-3}$	8	8
RRGMRES	(12)	$5.2 \cdot 10^{-3}$	11	12
LSQR	(12)	$6.8 \cdot 10^{-3}$	8	17
LSQR	(6)	$6.8 \cdot 10^{-3}$	7	17

Table 3
Example 6: Comparison of iterative methods and preconditioners.

where

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

This equation is discussed, e.g., by Delves and Mohamed [8, p. 315]. 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 [15]. The program yields a symmetric indefinite matrix* $A \in \mathbb{R}^{200\times 200}$ *and a scaled discrete approximation* $\hat{\boldsymbol{x}} \in \mathbb{R}^{200}$ *of the solution* $x(t) = t$ *of (30). The error-free right-hand side vector is given by* $\hat{\boldsymbol{b}} = A\hat{\boldsymbol{x}}$*, and the right-hand side vector* $\boldsymbol{b}$ *in (1) is determined by (2).*

Table 3 displays the performance of the iterative methods. We use preconditioners based on the invertible bidiagonal matrices (9) and (12) for all iterative methods and a preconditioner based on the noninvertible bidiagonal matrix (6) with LSQR. GMRES with preconditioner determined by (12) is seen to yield the highest accuracy.

The preconditioners discussed in this paper may increase the condition number. For the present example, we have $\kappa(A) = 4.9 \cdot 10^4$ *and* $\kappa(\hat{A}) = 1.2 \cdot 10^{14}$*, where* $\hat{A}$ *is given by (17) and* L *by (12). The purpose of the preconditioners is to change the Krylov subspaces, cf. (18) and (27), and to provide smoothing via (11).*

Figure 5(a) shows the best computed approximation of $\hat{\boldsymbol{x}}$ *determined without a preconditioner, and Figure 5(c) displays the best computed approximation*

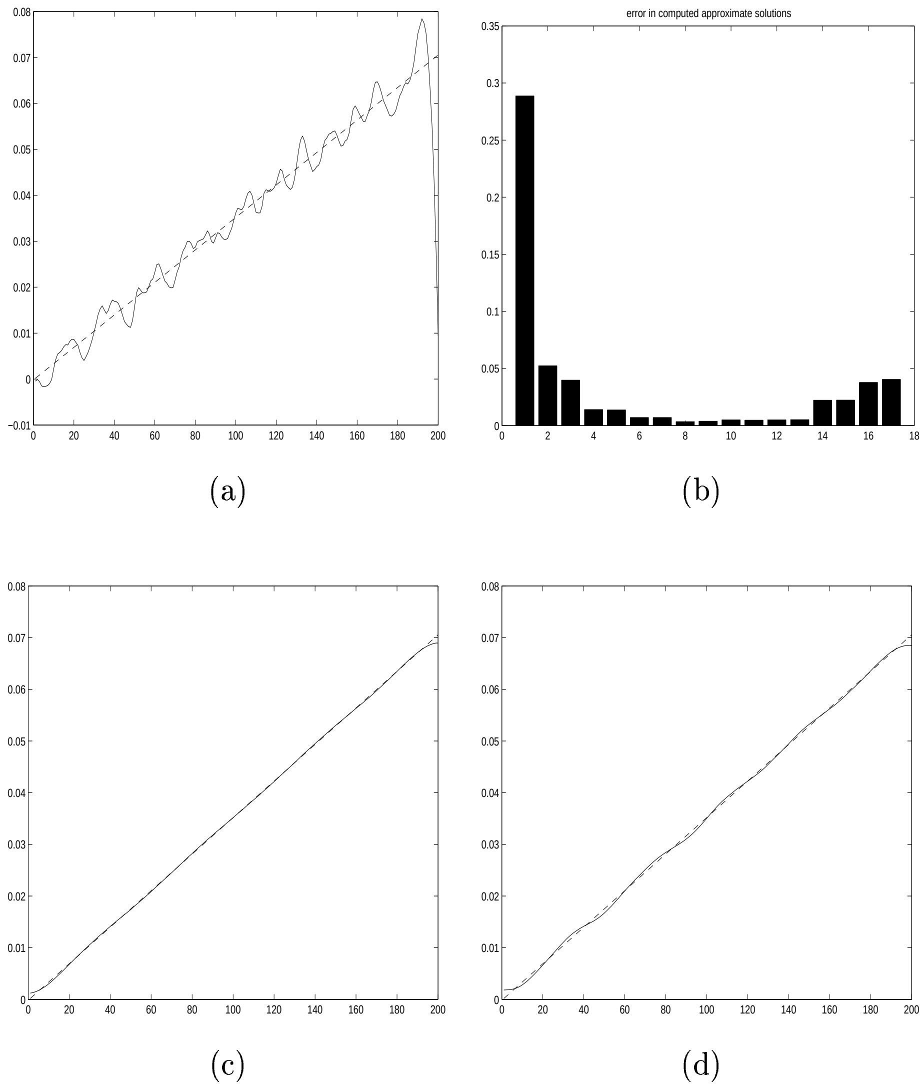


Fig. 5. Example 6: (a) Approximate solution $\boldsymbol{x}_{10}$ computed by unpreconditioned RRGMRES (continuous curve) and exact solution $\hat{\boldsymbol{x}}$ (dashed curve). (b) Errors $\|\hat{\boldsymbol{x}} - \boldsymbol{x}_k\|$ for $1 \leq k \leq 17$ in approximate solutions $\boldsymbol{x}_k$ computed by GMRES with preconditioner defined by (12). (c) Approximate solution $\boldsymbol{x}_8$ computed by GMRES with a preconditioner determined by (12) (continuous curve) and exact solution $\hat{\boldsymbol{x}}$ (dashed curve). (d) Approximate solution $\boldsymbol{x}_7$ computed by LSQR with preconditioner determined by (6) (continuous curve) and exact solution $\hat{\boldsymbol{x}}$ (dashed curve).

of $\hat{\boldsymbol{x}}$ *determined with a preconditioner. The latter approximate solution was computed by GMRES with a preconditioner defined by (12). For comparison, Figure 5(d) shows the best approximation of* $\hat{\boldsymbol{x}}$ *determined by LSQR with a preconditioner based on (6). Figure 5(b) shows the error history for preconditioned GMRES used to compute the approximate solution* $\boldsymbol{x}_8$ *displayed in Figure 5(c). The error norms* $\|\hat{\boldsymbol{x}} - \boldsymbol{x}_k\|$ *displayed in Figure 5(b) increase with* k *for* $k \geq 8$*. This illustrates the need to terminate the iterations before propagated errors due to the error* $\boldsymbol{e}$ *in* $\boldsymbol{b}$ *and round-off errors destroy the accuracy in the computed approximate solutions* $\boldsymbol{x}_k$*.*

5 Conclusion and future work

The paper describes ways to construct invertible preconditioners from noninvertible regularization operators for Tikhonov regularization. The numerical examples show that suitably defined invertible preconditioners can perform well not only with GMRES and RRGMRES, but also with LSQR. Moreover, preconditioned GMRES and RRGMRES are seen to yield computed solutions that approximate the solution $\hat{\boldsymbol{x}}$ of the error-free system (3) as well or better than the best iterates determined by preconditioned LSQR with fewer matrix-vector product evaluations. We are presently studying theoretical properties of invertible preconditioners as well as investigating the use of other classes of invertible preconditioners than those discussed in this paper.

Acknowledgements. We would like to thank Martin Hanke and Mike Saunders for comments and Per Christian Hansen for suggestions that improved the presentation.

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) 241–247.

[2] Å. Björck, L. Eldén, Methods in numerical linear algebra for ill-posed problems, Report, Department of Mathematics, Linköping University, Linköping, Sweden, 1979.

[3] D. Calvetti, B. Lewis, L. Reichel, GMRES-type methods for inconsistent systems, Linear Algebra Appl. 316 (2000) 157–169.

[4] D. Calvetti, B. Lewis, L. Reichel, Restoration of images with spatially variant blur by the GMRES method, in Advanced Signal Processing Algorithms, Architectures, and Implementations X, ed. F. T. Luk, Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE), vol. 4116, The International Society for Optical Engineering, Bellingham, WA, 2000, pp. 364–374.

[5] D. Calvetti, B. Lewis, L. Reichel, On the choice of subspace for iterative methods for linear discrete ill-posed problems, Int. J. Appl. Math. Comput. Sci. 11 (2001) 1069–1092.

[6] D. Calvetti, B. Lewis, L. Reichel, GMRES, L-curves and discrete ill-posed problems, BIT 42 (2002) 44–65.

[7] D. Calvetti, B. Lewis, L. Reichel, On the regularizing properties of the GMRES method, Numer. Math. 91 (2002) 605–625.

[8] L. M. Delves, J. L. Mohamed, Computational Methods for Integral Equations, Cambridge University Press, Cambridge, 1985.

[9] L. Eldén, A weighted pseudoinverse, generalized singular values, and constraint least squares problems, BIT 22 (1982) 487–502.

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

[11] M. Hanke, Conjugate Gradient Type Methods for Ill-Posed Problems, Longman, Harlow, 1995.

[12] M. Hanke, Regularization with differential operators: An iterative approach, Numer. Funct. Anal. Optim. 13 (1992) 523–540.

[13] M. Hanke, Iterative solution of underdetermined linear systems by transformation to standard form, in Proceedings Numerical Methods in Theory and Practice, Department of Mathematics, University of West Bohemia, Plzen, Czech Republic, 1993, pp. 55–63.

[14] M. Hanke, P. C. Hansen, Regularization methods for large-scale problems, Surv. Math. Ind. 3 (1993) 253–315.

[15] P. C. Hansen, Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems, Numer. Algorithms 6 (1994) 1–35.

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

[17] M. R. Hestenes, E. Stiefel, Methods of conjugate gradients for solving linear systems, J. Res. Nat. Bur. Standards 49 (1952) 409–436.

[18] C. C. Paige, M. A. Saunders, LSQR: An algorithm for sparse linear equations and sparse least squares, ACM Trans. Math. Software 8 (1982) 43–71.

[19] C. C. Paige, M. A. Saunders, Algorithm 583 LSQR: Sparse linear equations and least squares problems, ACM Trans. Math. Software 8 (1982) 195–209.

[20] Y. Saad and M. H. Schultz, GMRES: A generalized minimal residual method for solving nonsymmetric linear systems, SIAM J. Sci. Stat. Comput. 7 (1986) 856–869.