NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2016; **23**:187–204
Published online 12 October 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nla.2020

On the Lanczos and Golub–Kahan reduction methods applied to discrete ill-posed problems

Silvia Gazzola[1], Enyinda Onunwor[2], Lothar Reichel[2] and Giuseppe Rodriguez[3,*,†]

[1]*Dipartimento di Matematica, Università di Padova, via Trieste 63, Padua, Italy*
[2]*Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA*
[3]*Dipartimento di Matematica e Informatica, Università di Cagliari, viale Merello 92, 09123 Cagliari, Italy*

SUMMARY

The symmetric Lanczos method is commonly applied to reduce large-scale symmetric linear discrete ill-posed problems to small ones with a symmetric tridiagonal matrix. We investigate how quickly the nonnegative subdiagonal entries of this matrix decay to zero. Their fast decay to zero suggests that there is little benefit in expressing the solution of the discrete ill-posed problems in terms of the eigenvectors of the matrix compared with using a basis of Lanczos vectors, which are cheaper to compute. Similarly, we show that the solution subspace determined by the LSQR method when applied to the solution of linear discrete ill-posed problems with a nonsymmetric matrix often can be used instead of the solution subspace determined by the singular value decomposition without significant, if any, reduction of the quality of the computed solution.

Received 27 March 2015; Revised 7 August 2015; Accepted 14 September 2015

KEY WORDS: discrete ill-posed problems; Lanczos decomposition; Golub–Kahan bidiagonalization; LSQR; TSVD

1. INTRODUCTION

We are concerned with the solution of linear systems of equations

$$Ax = b \tag{1.1}$$

with a large symmetric matrix $A \in \mathbb{R}^{n\times n}$ whose eigenvalues in magnitude gradually approach zero without a significant gap. Thus, A is very ill conditioned and may be singular. Linear systems of equations with a matrix of this kind are commonly referred to as linear discrete ill-posed problems. They originate, for instance, from the discretization of symmetric ill-posed problems, such as Fredholm integral equations of the first kind with a symmetric kernel. Inconsistent systems (1.1) are treated as least squares problems. Throughout this paper, $\|\cdot\|$ stands for the Euclidean vector norm or the spectral matrix norm. The range of a matrix M is denoted by $\mathcal{R}(M)$.

The right-hand side $b \in \mathbb{R}^n$ in linear discrete ill-posed problems that arise in science and engineering generally represents measured data that are contaminated by an error $\eta \in \mathbb{R}^n$. Thus,

$$b = b_{\text{true}} + \eta, \tag{1.2}$$

where $b_{\text{true}} \in \mathbb{R}^n$ represents the unavailable error-free vector associated with b.

Let $A^\dagger$ denote the Moore–Penrose pseudoinverse of A. We would like to determine an approximation of $x_{\text{true}} := A^\dagger b_{\text{true}}$. Note that the solution of (1.1), given by $x := A^\dagger b = x_{\text{true}} + A^\dagger \eta$,

*Correspondence to: Giuseppe Rodriguez, Dipartimento di Matematica e Informatica, Università di Cagliari, viale Merello 92, 09123 Cagliari, Italy.
†E-mail: rodriguez@unica.it

typically is not a useful approximation of x_{true} because, generally, $\|A^\dagger \eta\| \gg \|x_{\text{true}}\|$. This difficulty is commonly remedied by replacing the linear system (1.1) by a nearby problem, whose solution is less sensitive to the error η in b. The replacement is known as regularization. Among the most popular regularization methods are Tikhonov regularization, partial spectral decomposition, and truncated iteration.

Many solution methods for large-scale problems (1.1) first reduce the system of equations to a problem of small size. The symmetric Lanczos method is a popular reduction method. Application of m steps of this method to A with initial vector b yields a decomposition of the form

$$AV_m = V_{m+1}T_{m+1,m}, \tag{1.3}$$

where $V_{m+1} = [v_1, v_2, \ldots, v_{m+1}] \in \mathbb{R}^{n\times(m+1)}$ has orthonormal columns with $v_1 = b/\|b\|$, $V_m = [v_1, v_2, \ldots, v_m] \in \mathbb{R}^{n\times m}$, and the matrix

$$T_{m+1,m} = \begin{bmatrix} \alpha_1 & \beta_2 & & & & & \\ \beta_2 & \alpha_2 & \beta_3 & & & & \\ & \beta_3 & \alpha_3 & \ddots & & & \\ & & \ddots & \ddots & \beta_{m-1} & & \\ & & & \beta_{m-1} & \alpha_{m-1} & \beta_m & \\ & & & & \beta_m & \alpha_m & \\ & & & & & \beta_{m+1} & \end{bmatrix} \in \mathbb{R}^{(m+1)\times m} \tag{1.4}$$

is tridiagonal with a leading symmetric $m \times m$ submatrix, which we denote by T_m. It follows from the recursions of the Lanczos method that the columns v_j of V_{m+1} can be expressed as

$$v_j = q_{j-1}(A)b, \qquad j = 1, 2, \ldots, m+1,$$

where q_{j-1} is a polynomial of exact degree $j-1$. Consequently, $\{v_j\}_{j=1}^m$ is an orthonormal basis for the Krylov subspace

$$\mathbb{K}_m(A, b) = \text{span}\{b, Ab, \ldots, A^{m-1}b\}. \tag{1.5}$$

The Lanczos method determines the diagonal and subdiagonal elements

$$\alpha_1, \beta_2, \alpha_2, \beta_3, \ldots, \alpha_m, \beta_{m+1}$$

of $T_{m+1,m}$ in order. Generically, the subdiagonal entries β_j are positive, and then the decomposition (1.3) with the stated properties exists. We say that the Lanczos method breaks down at step $j \leqslant m$ if the $\beta_2, \beta_3, \ldots, \beta_j$ are positive and β_{j+1} vanishes. Then a decomposition of the form (1.3) does not exist. However, $\text{span}\{v_i\}_{i=1}^j$ forms an invariant subspace of A and this, generally, is a fortuitous event; see, for example, Saad [1, 2] for discussions on the Lanczos method.

The solution of (1.1) by truncated iteration proceeds by solving

$$\min_{x\in\mathbb{K}_m(A,b)} \|Ax - b\| = \min_{y\in\mathbb{R}^m} \|T_{m+1,m}y - e_1\|b\|\|, \tag{1.6}$$

where the right-hand side is obtained by substituting the decomposition (1.3) into the left-hand side and by exploiting the properties of the matrices involved. Here and in the following, e_j denotes the jth axis vector. Let y_m be a solution of the least squares problem on the right-hand side of (1.6). Then $x_m := V_m y_m$ is a solution of the constrained least squares problem on the left-hand side of (1.6), as well as an approximate solution of (1.1). By choosing m suitably small, the propagation of the error η in b into the computed solution x_m is reduced. This depends on that the condition number of $T_{m+1,m}$, given by $\kappa(T_{m+1,m}) := \|T_{m+1,m}\|\|T_{m+1,m}^\dagger\|$, is an increasing function of m. A large

DOI: 10.1002/nla

condition number indicates that the solution y_m of the right-hand side of (1.6) is sensitive to errors in the data and to roundoff errors introduced during the computations. We will discuss this and other solution methods in the following. For overviews and analyses of solution methods for linear discrete ill-posed problems, we refer to [3, 4].

It is the purpose of the present paper to investigate the structure of the matrix (1.4) obtained by applying the Lanczos method to a symmetric matrix whose eigenvalues 'cluster' at the origin. We will give upper bounds for the size of the subdiagonal entries. These bounds shed light on the solution subspaces generated by the symmetric Lanczos method. In particular, the bounds indicate that the ranges of the matrices V_m essentially contain the span of the $k = k(m)$ eigenvectors of A associated with the k eigenvalues of largest magnitude where $k(m)$ is an increasing function of m and, generally, $k(m) < m$. This observation suggests that it may not be necessary to compute a partial eigendecomposition of A but that it suffices to determine a few Lanczos vectors, which is much cheaper. We will also investigate the solution subspaces determined by application of m steps of Golub–Kahan bidiagonalization to a nonsymmetric matrix A, whose singular values cluster at the origin. We find the solution subspaces determined by m steps of Golub–Kahan bidiagonalization applied to A to essentially contain the spans of the $k = k(m)$ right and left singular vectors of A associated with the k largest singular values, where $k = k(m)$ is an increasing function of m and, generally, $k(m) < m$. This suggests that it may not be necessary to compute singular value or partial singular value decompositions of A but that it suffices to carry out a few steps of Golub–Kahan bidiagonalization, which is much cheaper. The results for the spans of the solution subspaces determined by partial Golub–Kahan bidiagonalization follow from bounds for singular values. These bounds provide an alternative to the bounds recently shown by Gazzola *et al.* [5, 6]. Related bounds also are presented by Novati and Russo [7].

This paper is organized as follows. Section 2 presents our new bounds for the entries β_j of $T_{m+1,m}$ and discusses some implications. Application to the bidiagonal matrices determined by Golub–Kahan bidiagonalization is described in Section 3. A few computed examples are shown in Section 4, and Section 5 contains concluding remarks.

2. THE SYMMETRIC LANCZOS METHOD

This section discusses the convergence of the subdiagonal and diagonal entries of the matrix $T_{m+1,m}$ in (1.3) with increasing dimensions. The proofs use the spectral factorization

$$A = W\Lambda W^T, \tag{2.1}$$

where the matrix $W = [w_1, w_2, \ldots, w_n] \in \mathbb{R}^{n\times n}$ is orthogonal, the superscript ^{T} denotes transposition, and

$$\Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n] \in \mathbb{R}^{n\times n}, \qquad |\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_n| \geqslant 0. \tag{2.2}$$

Theorem 2.1
Let the matrix $A \in \mathbb{R}^{n\times n}$ be symmetric and positive semidefinite, and let its eigenvalues be ordered according to (2.2). Assume that the Lanczos method applied to A with initial vector b does not break down, that is, that n steps of the method can be carried out. Let $\beta_2, \beta_3, \ldots, \beta_{m+1}$ be the subdiagonal entries of the matrix $T_{m+1,m}$ determined by m steps of the Lanczos method; compare with (1.3). Define $\beta_{n+1} := 0$. Then

$$\prod_{j=2}^{m+1} \beta_j \leqslant \prod_{j=1}^{m} \lambda_j, \qquad m = 1, 2, \ldots, n. \tag{2.3}$$

Proof
Introduce the monic polynomial $p_m(t) = \prod_{j=1}^{m}(t - \lambda_j)$ defined by the m largest eigenvalues of A. Using the spectral factorization (2.1), we obtain

DOI: 10.1002/nla

$$\|p_m(A)\| = \|p_m(\Lambda)\| = \max_{m+1\leqslant j\leqslant n} |p_m(\lambda_j)| \leqslant |p_m(0)| = \prod_{j=1}^{m} \lambda_j,$$

where the inequality follows from the fact that all λ_j are nonnegative. Hence,

$$\|p_m(A)b\| \leqslant \|b\| \prod_{j=1}^{m} \lambda_j. \tag{2.4}$$

Application of n steps of the symmetric Lanczos method gives the decomposition $AV_n = V_n T_n$, where $T_n \in \mathbb{R}^{n\times n}$ is symmetric and tridiagonal, and $V_n \in \mathbb{R}^{n\times n}$ is orthogonal with $V_n e_1 = b/\|b\|$. We have

$$p_m(A)b = V_n\, p_m(T_n) V_n^T b = V_n\, p_m(T_n) e_1 \|b\|. \tag{2.5}$$

This relation gives the equality in the following,

$$\|p_m(A)b\| = \|p_m(T_n)e_1\|\|b\| \geqslant \|b\| \prod_{j=2}^{m+1} \beta_j. \tag{2.6}$$

The inequality in (2.6) follows by direct computation. Specifically, one can show by induction on m that

$$\|p_m(T_n)e_1\| \geqslant |e_{m+1}^T p_m(T_n) e_1| = \prod_{j=2}^{m+1} \beta_j.$$

For $m = 1$, the result is trivial. Assume that it is valid for $1 \leqslant m < n$. Then

$$e_{m+2}^T p_{m+1}(T_n) e_1 = e_{m+2}^T (T_n - \lambda_{m+1} I)\, p_m(T_n) e_1 = \phi_{m+2}^T p_m(T_n) e_1,$$

with

$$\phi_{m+2} = \beta_{m+2} e_{m+1} + (\alpha_{m+2} - \lambda_{m+1}) e_{m+2} + \beta_{m+3} e_{m+3}.$$

Because $p_m(T_n)$ is $(2m+1)$-banded, we obtain

$$\phi_{m+2}^T p_m(T_n) e_1 = \beta_{m+2} e_{m+1}^T p_m(T_n) e_1 = \beta_{m+2} \prod_{j=2}^{m+1} \beta_j = \prod_{j=2}^{m+2} \beta_j.$$

Combining (2.4) and (2.6) shows the theorem. □

In practice, the bound (2.3) is often quite sharp; we give a numerical illustration of this in Section 4. Moreover, we can easily derive bounds of the form

$$\beta_{j+1} \leqslant k_j \lambda_j, \quad j = 1, \ldots, n-1, \quad \text{with} \quad k_j := \frac{\prod_{i=1}^{j-1} \lambda_i}{\prod_{i=1}^{j-1} \beta_{i+1}} \geqslant 1\,. \tag{2.7}$$

However, if the bound (2.3) is not sharp, then $k_j \gg 1$, resulting in a meaningless estimate (2.7). A result analogous to Theorem 2.1 for nonsymmetric matrices A, with the Lanczos method replaced by the Arnoldi method, has been shown by Novati and Russo [7] and Gazzola *et al.* [5, 6]. It should be emphasized that the bounds proved in [5–7] are similar to (2.7) but assume b_{true} as starting vector for the Arnoldi algorithm, involve constants whose values are not explicitly known, and are only valid for moderately to severely ill-posed problems; see, for example, Hansen [4] for this classification

DOI: 10.1002/nla

of ill-posed problems. The restriction to symmetric matrices allows us to give a bound that does not explicitly depend on the starting vector of the Lanczos method.

Corollary 2.2
Let $A \in \mathbb{R}^{n\times n}$ be symmetric and positive semidefinite. Assume that the eigenvalues of A 'cluster' at the origin and that the Lanczos method applied to A with initial vector b does not break down. Then both the diagonal and subdiagonal entries α_j and β_j of the tridiagonal Lanczos matrix $T_{m+1,m}$, compare with (1.3), approach zero when j increases and is large enough.

Proof
We first remark that when we let the index j increase, we also may have to increase m in (1.3). The fact that the β_j's approach zero as j increases follows from (2.3) and the clustering of the eigenvalues at zero. We turn to the diagonal entries α_j of the tridiagonal Lanczos matrix T_n. This matrix is similar to A. Therefore, its eigenvalues cluster at the origin, which is the only cluster point. Application of Gershgorin disks to the rows of T_n, using the fact that the off-diagonal entries are 'tiny' for large row numbers, shows that the corresponding diagonal entries are accurate approximations of the eigenvalues of T_n. These entries therefore have to approach zero as j increases. □

We remark that the decrease of the subdiagonal entries β_j of $T_{m+1,m}$ to zero with increasing j follows from the clustering of the eigenvalues of A; it is not necessary that they cluster at the origin. This can be seen by replacing the matrix A in Corollary 2.2 by $A + cI_n$ for some constant $c \in \mathbb{R}$. This substitution adds c to the diagonal entries of $T_{m+1,m}$ in (1.3), and it is employed with a suitable parameter $c > 0$ to perform Lavrentiev-type regularization (cf. [4]).

The assumption in Theorem 2.1 that n steps of the Lanczos method can be carried out simplifies the proof but is not essential. We state the corresponding result when the Lanczos method breaks down at step $k < n$.

Corollary 2.3
Let the matrix $A \in \mathbb{R}^{n\times n}$ be symmetric and positive semidefinite, and let its eigenvalues be ordered according to (2.2). Assume that the Lanczos method applied to A with initial vector b breaks down at step k. Then

$$\prod_{j=2}^{m+1} \beta_j \leqslant \prod_{j=1}^{m} \lambda_j, \qquad m = 1, 2, \ldots, k,$$

where $\beta_{k+1} := 0$.

Proof
Application of k steps of the Lanczos method to the matrix A with initial vector b gives the decomposition

$$AV_k = V_k T_k,$$

where $V_k = [v_1, v_2, \ldots, v_k] \in \mathbb{R}^{n\times k}$ has orthonormal columns with $V_k e_1 = b/\|b\|$ and

$$T_k = \begin{bmatrix} \alpha_1 & \beta_2 & & & & \\ \beta_2 & \alpha_2 & \beta_3 & & & \\ & \beta_3 & \alpha_3 & \ddots & & \\ & & \ddots & \ddots & \beta_{k-1} & \\ & & & \beta_{k-1} & \alpha_{k-1} & \beta_k \\ & & & & \beta_k & \alpha_k \end{bmatrix} \in \mathbb{R}^{k\times k}$$

is symmetric and tridiagonal with positive subdiagonal entries. The inequality (2.4) holds for $m = 1, 2, \ldots, n$; however, the relation (2.5) has to be replaced by

DOI: 10.1002/nla

$$p_m(A)b = V_k\, p_m(T_k)e_1\|b\|, \quad 1 \leqslant m < k\,. \tag{2.8}$$

This relation can be shown by induction on m. Indeed, for $m = 1$, one immediately has

$$p_1(A)b = (A-\lambda_1 I)V_k e_1\|b\| = (AV_k - \lambda_1 V_k)e_1\|b\| = V_k(T_k - \lambda_1 I)e_1\|b\| = V_k\, p_1(T_k)e_1\|b\|\,.$$

Assuming that (2.8) holds for $m < k-1$, for $m+1$, one obtains

$$\begin{aligned} p_{m+1}(A)b &= (A-\lambda_{m+1}I)\,p_m(A)b = (A-\lambda_{m+1}I)V_k\, p_m(T_k)e_1\|b\| \\ &= V_k(T_k - \lambda_{m+1}I)\,p_m(T_k)e_1\|b\| = V_k\, p_{m+1}(T_k)e_1\|b\|. \end{aligned}$$

Analogously to (2.6), we have

$$\|p_m(A)b\| = \|p_m(T_k)e_1\|\|b\| \geqslant \|b\| \prod_{j=2}^{m+1} \beta_j,$$

and the corollary follows. □

We turn to symmetric indefinite matrices. For notational simplicity, we will assume that the Lanczos method does not break down, but this requirement can be relaxed similarly as in Corollary 2.3.

Theorem 2.4
Let the eigenvalues $\{\lambda_j\}_{j=1}^n$ of the symmetric matrix $A \in \mathbb{R}^{n\times n}$ be ordered according to (2.2). Assume that the Lanczos method applied to A with initial vector b does not break down. Then

$$\prod_{j=2}^{m+1} \beta_j \leqslant \prod_{j=1}^{m} (|\lambda_{m+1}| + |\lambda_j|), \qquad m = 1, 2, \ldots, n-1. \tag{2.9}$$

Proof
Let $p_m(t)$ be the monic polynomial of the proof of Theorem 2.1. Then just like in that proof

$$\|p_m(A)\| = \|p_m(\Lambda)\| = \max_{m+1\leqslant j\leqslant n} |p_m(\lambda_j)|.$$

It follows from the ordering (2.2) of the eigenvalues that the interval $[-|\lambda_{m+1}|, |\lambda_{m+1}|]$ contains all the eigenvalues $\lambda_{m+1}, \lambda_{m+2}, \ldots, \lambda_n$. Therefore,

$$\max_{m+1\leqslant j\leqslant n} |p_m(\lambda_j)| \leqslant \max_{-|\lambda_{m+1}|\leqslant t\leqslant|\lambda_{m+1}|} |p_m(t)| \leqslant \prod_{k=1}^{m} (|\lambda_{m+1}| + |\lambda_k|).$$

The inequality (2.9) now follows similarly as the proof of the analogous inequality (2.3). □

Assume that the eigenvalues of A cluster at the origin. Then Theorem 2.4 shows that the factors $|\lambda_{m+1}| + |\lambda_k|$ decrease to zero as m and k, with $1 \leqslant k \leqslant m$, increase. Furthermore, the more Lanczos steps are taken, the tighter is the bound for the product of the subdiagonal elements of the matrix $T_{m+1,m}$. Sharper bounds for the product of subdiagonal entries of $T_{m+1.m}$ can be obtained if more information about the spectrum of A is available. For instance, if all but a few eigenvalues of A are known to be nonnegative, then only the factors with the negative eigenvalues have to be modified as in Theorem 2.4, resulting in improved bounds for products of the β_j. Simpler, but cruder, bounds than (2.9) also can be derived. The following is an example.

Corollary 2.5
Let the eigenvalues $\{\lambda_j\}_{j=1}^n$ of the symmetric matrix $A \in \mathbb{R}^{n\times n}$ be ordered according to (2.2). Assume that the Lanczos method applied to A with initial vector b does not break down. Then,

DOI: 10.1002/nla

$$\prod_{j=2}^{m+1} \beta_j \leqslant \prod_{k=1}^{m} (2|\lambda_k|), \qquad m = 1, 2, \ldots, n-1.$$

Proof
The result follows from the observation that $|\lambda_{m+1}| \leqslant |\lambda_k|$ for $1 \leqslant k \leqslant m$. □

Introduce the set of ε-pseudoeigenvectors of $A \in \mathbb{R}^{n\times n}$:

$$\mathbb{V}_\varepsilon := \{x \in \mathbb{R}^n \text{ unit vector} : \exists \lambda \in \mathbb{R} \text{ such that } \|Ax - \lambda x\| \leqslant \varepsilon\}. \tag{2.10}$$

The λ-values associated with ε-pseudoeigenvectors are ε-pseudoeigenvalues of A; see, for example, Trefethen and Embree [8] for an insightful treatment of pseudospectra of matrices and operators.

Substituting the decomposition $A = V_n T_n V_n^T$ into (2.10) and applying Theorem 2.4 show that, for a given $\varepsilon > 0$ and for j sufficiently large, the Lanczos vectors v_j are ε-pseudoeigenvectors of A associated with eigenvalues close to zero. Indeed, by (1.3), we obtain

$$Av_j = AV_m e_j = V_{m+1} T_{m+1,m} e_j = \alpha_j v_j + \beta_j v_{j-1} + \beta_{j+1} v_{j+1}.$$

Because as j increases α_j and β_j approach 0, we can conclude that the v_j are ε-pseudoeigenvectors for j large. Let w_j denote the jth column of the matrix W in (2.1), that is, let w_j be the jth eigenvector of A. Therefore, the space $\text{span}\{w_j\}_{j=1}^k$ is essentially contained in $\text{span}\{v_j\}_{j=1}^m$ for $k = k(m) \leqslant m$ sufficiently small. The notion of 'essentially contained' will be made precise and illustrated in Section 4.

The aforementioned observation about the subspaces $\text{span}\{w_j\}_{j=1}^k$ and $\text{span}\{v_j\}_{j=1}^m$ for $k = k(m)$ has implications for computations. One of the most popular methods for solving linear discrete ill-posed problems is the truncated singular value decomposition (TSVD); see [3, 4]. For symmetric problems, this method simplifies to the truncated spectral decomposition. It is based on expressing an approximate solution of (1.1) as a linear combination of the first few eigenvectors, say $\{w_j\}_{j=1}^k$, of A (cf. (2.1)). The computation of these eigenvectors is more expensive than the determination of the Lanczos vectors $\{v_j\}_{j=1}^m$ for a reasonable $k = k(m) \leqslant m$, because typically several Lanczos decompositions with different initial vectors have to be computed in order to determine the desired eigenvectors; see, for example, Baglama *et al.* [9] and Saad [2] for discussions on methods for computing a few eigenpairs of a large matrix. Because the span of the Lanczos vectors $\{v_j\}_{j=1}^m$ essentially contains the set of the eigenvectors $\{w_j\}_{j=1}^k$, there is, generally, no need to compute the latter. This is illustrated by numerical examples in Section 4.

It is sometimes beneficial to determine an approximate solution of (1.1) in a shifted Krylov subspace

$$\mathbb{K}_m(A, Ab) = \text{span}\{Ab, A^2 b, \ldots, A^m b\} \tag{2.11}$$

instead of in the standard Krylov subspace (1.5). This is discussed and illustrated in [10–12]. Let

$$A\breve{V}_m = \breve{V}_{m+1} \breve{T}_{m+1,m}, \tag{2.12}$$

where $\breve{V}_{m+1} = [\breve{v}_1, \breve{v}_2, \ldots, \breve{v}_{m+1}] \in \mathbb{R}^{n\times(m+1)}$ has orthonormal columns with $\breve{v}_1 = Ab/\|Ab\|$, $\breve{V}_m = [\breve{v}_1, \breve{v}_2, \ldots, \breve{v}_m] \in \mathbb{R}^{n\times m}$, and the tridiagonal matrix $\breve{T}_{m+1,m}$ is of the same form as (1.4). Then, analogously to (1.6), we formally obtain

$$\min_{x\in\mathbb{K}_m(A,Ab)} \|Ax - b\|^2 = \min_{y\in\mathbb{R}^m} \|\breve{T}_{m+1,m} y - \breve{V}_{m+1}^T b\|^2 + \|(I - \breve{V}_{m+1}\breve{V}_{m+1}^T)b\|^2.$$

Let $\breve{y}_m \in \mathbb{R}^m$ denote the solution of the minimization problem on the right-hand side of the last relation. Then $\breve{x}_m := \breve{V}_m \breve{y}_m$ solves the constrained least squares problem on the left-hand side and is an approximate solution of (1.1). Computed examples show the vector $\breve{x}_m$ to

DOI: 10.1002/nla

generally approximate the desired solution x_{true} more accurately than the solution x_m of the minimization problem (1.6). In the computed examples of Section 4, we therefore compute the vectors $\check{x}_1, \check{x}_2, \ldots$.

We note that the analysis in this section is independent of the initial vector b in the Krylov subspace (1.5), except for how this vector affects the occurrence of breakdown. In particular, our analysis carries over to shifted Krylov subspaces of the form (2.11).

3. GOLUB–KAHAN BIDIAGONALIZATION

A nonsymmetric matrix $A \in \mathbb{R}^{n\times\ell}$ can be reduced to a small bidiagonal matrix by a few steps of Golub–Kahan bidiagonalization; see, for example, [13]. This reduction method is the basis for the popular LSQR algorithm [14] for the solution of least squares problems

$$\min_{x\in\mathbb{R}^\ell} \|Ax - b\|, \tag{3.1}$$

where the vector $b \in \mathbb{R}^n$ can be written as (1.2). We assume for notational simplicity that $1 \leqslant \ell \leqslant n$. Application of $m \ll \ell$ steps of Golub–Kahan bidiagonalization to A with initial vector b gives the decompositions

$$AV'_m = U'_{m+1}B_{m+1,m}, \qquad A^T U'_m = V'_m B_m^T. \tag{3.2}$$

Here, the matrices $U'_{m+1} = [u'_1, u'_2, \ldots, u'_{m+1}] \in \mathbb{R}^{n\times(m+1)}$ and $V'_m = [v'_1, v'_2, \ldots, v'_m] \in \mathbb{R}^{\ell\times m}$ have orthonormal columns, $U'_m = [u'_1, u'_2, \ldots, u'_m]$, and $u'_1 := b/\|b\|$. The columns of V'_m span the Krylov subspace $\mathbb{K}_m(A^TA, A^Tb)$ and

$$B_{m+1,m} = \begin{bmatrix} \alpha_1 & & & & \\ \beta_2 & \alpha_2 & & & \\ & \beta_3 & \alpha_3 & & \\ & & \ddots & \ddots & \\ & & & \beta_m & \alpha_m \\ & & & & \beta_{m+1} \end{bmatrix} \in \mathbb{R}^{(m+1)\times m} \tag{3.3}$$

is lower bidiagonal with leading $m \times m$ submatrix B_m. All entries α_j and β_j are positive. We assume that m is chosen small enough so that the decompositions (3.2) with the stated properties exist. This is the generic situation. Throughout this section, α_j and β_j refer to entries of the matrix (3.3).

The LSQR method applied to the solution of (3.1) solves in step m the minimization problem

$$\min_{x\in\mathbb{K}_m(A^TA,A^Tb)} \|Ax - b\| = \min_{y\in\mathbb{R}^m} \|B_{m+1,m}y - e_1\|b\|\|,$$

where the right-hand side is obtained by substituting (3.2) into the left-hand side. Denote the solution of the right-hand side by y'_m. Then the mth step of LSQR yields the solution $x'_m = V'_m y'_m$ of the left-hand side, which is an approximate solution of (3.1).

Combining the decompositions (3.2) gives

$$A^TAV'_m = V'_{m+1}T'_{m+1,m}, \tag{3.4}$$

where $T'_{m+1,m} = B_{m+1}^T B_{m+1,m}$ is tridiagonal with a symmetric leading $m \times m$ submatrix. The decomposition (3.4) is a Lanczos decomposition, analogous to (1.3), of the symmetric positive semidefinite matrix A^TA. This allows us to apply Theorem 2.1.

Corollary 3.1
Let $A \in \mathbb{R}^{n\times\ell}$ have the singular values $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_\ell \geqslant 0$, and assume that the Golub–Kahan bidiagonalization method applied to A with initial vector b does not break down. Then

DOI: 10.1002/nla

$$\prod_{j=2}^{m+1} \alpha_j \beta_j \leqslant \prod_{j=1}^{m} \sigma_j^2, \qquad m = 1, 2, \ldots, n-1, \tag{3.5}$$

where the α_j and β_j are entries of the bidiagonal matrix (3.3).

Proof
The subdiagonal entries of the matrix $T'_{m+1,m}$ in (3.4) are $\alpha_j \beta_j$, and the eigenvalues of $A^T A$ are σ_j^2. The result therefore follows from Theorem 2.1. □

The aforementioned corollary shows that if the singular values σ_j cluster at zero for large j, then so do the products $\alpha_j \beta_j$ of the entries of the matrix (3.3). Bounds related to (3.5) have been shown by Gazzola *et al.* [6].

Truncated singular value decomposition is a popular solution method for discrete ill-posed problems (3.1) with a nonsymmetric matrix A. Let

$$A = Z \Sigma W^T \tag{3.6}$$

denote the singular value decomposition of A. Thus, $Z \in \mathbb{R}^{n \times n}$ and $W \in \mathbb{R}^{\ell \times \ell}$ are orthogonal matrices, and

$$\Sigma = \operatorname{diag}[\sigma_1, \sigma_2, \ldots, \sigma_\ell] \in \mathbb{R}^{n \times \ell}, \quad \sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > \sigma_{r+1} = \cdots = \sigma_\ell = 0,$$

where r is the rank of A. Substituting (3.6) into (3.1) gives the minimization problem

$$\min_{y \in \mathbb{R}^\ell} \|\Sigma y - Z^T b\|, \qquad y = W^T x.$$

Let the matrices Z_k and W_k be made up of the $1 \leqslant k \leqslant r$ first columns of Z and W, respectively, and let Σ_k denote the leading $k \times k$ principal submatrix of Σ. The TSVD method determines the solution $\check{y}_k \in \mathbb{R}^k$ of

$$\Sigma_k y = Z_k^T b$$

for a suitable value of k. Then $\check{x}_k := W_k \check{y}_k$ is an approximate solution of (3.1). Details and properties of the TSVD solution method are discussed, for example, in [3, 4].

The computation of the decomposition (3.6) is feasible for problems of small to moderate size but expensive for large-scale problems. The computational expense for large-scale problems can be reduced somewhat by computing the matrices Z_k, W_k, and Σ_k instead of Z, W, and Σ; see [15, 16], and references therein for suitable numerical methods. The computation of Z_k, W_k, and Σ_k generally requires that several Golub–Kahan decompositions (3.2) with different initial vectors be evaluated.

Corollary 3.1 indicates why it may not be necessary to compute the matrices Z_k, W_k, and Σ_k. The columns of the matrix W in (3.6) are eigenvectors of $A^T A$, and the σ_j^2 are eigenvalues. It is a consequence of Corollary 3.1, and the fact that the singular values σ_j of A cluster at the origin, that the columns v'_j of the matrix V'_m in (3.2) for small j are accurate approximations of eigenvectors of $A^T A$. This follows from an argument analogous to the discussion in Section 2 based on Corollaries 2.2 and 2.5. Therefore, it generally is not necessary to compute the partial singular value decomposition $\{U_k, V_k, \Sigma_k\}$ of A. Instead, it suffices to determine a partial Golub–Kahan bidiagonalization (3.2), which is cheaper. This is illustrated in the following section.

4. COMPUTED EXAMPLES

To investigate the properties discussed in the previous sections, we applied the symmetric Lanczos and Golub–Kahan bidiagonalization methods to a set of test matrices whose singular values cluster

DOI: 10.1002/nla

at the origin. The numerical experiments were carried out using MATLAB R2014a in double precision arithmetic, that is, with about 15 significant decimal digits.

The symmetric test matrices are listed in Table I and the nonsymmetric ones in Table II. Among the symmetric matrices, one example is negative definite (Deriv2), one is positive definite (Gravity), and the other ones are indefinite. All matrices except one are from the Regularization Tools package [17]. The Lotkin test matrix was generated by the `gallery` function, which is available in the standard MATLAB distribution. All test matrices are of order 200×200 except when explicitly stated otherwise. Figure 1 displays, in logarithmic scale, the values taken by each side of inequalities (2.3), (2.9), and (3.5), with the number of iterations, m, ranging from 1 to the index

Table I. Solution of symmetric linear systems: the errors E_{Lanczos} and E_{TEIG} are optimal for truncated Lanczos iteration and truncated eigenvalue decomposition. The corresponding truncation parameters are denoted by k_{Lanczos} and k_{TEIG}. Three noise levels δ are considered; ℓ denotes the number of Lanczos iterations performed.

Noise	Matrix	ℓ	E_{Lanczos}	k_{Lanczos}	E_{TEIG}	k_{TEIG}
$\delta = 10^{-6}$	Deriv2	200	2.0e-02	49	2.1e-02	199
	Foxgood	24	6.8e-04	6	6.3e-04	6
	Gravity	46	1.2e-03	15	1.2e-03	16
	Phillips	200	5.8e-04	22	5.8e-04	32
	Shaw	19	1.9e-02	10	1.9e-02	10
$\delta = 10^{-4}$	Deriv2	200	1.2e-01	10	1.1e-01	51
	Foxgood	24	1.4e-02	3	4.5e-03	4
	Gravity	45	1.5e-02	7	6.3e-03	12
	Phillips	200	4.8e-03	12	3.9e-03	15
	Shaw	19	4.7e-02	7	3.4e-02	9
$\delta = 10^{-2}$	Deriv2	200	3.1e-01	3	2.3e-01	12
	Foxgood	24	7.7e-02	2	2.9e-02	2
	Gravity	45	8.0e-02	3	3.3e-02	7
	Phillips	200	4.3e-02	6	2.2e-02	8
	Shaw	19	1.2e-01	7	9.1e-02	7

Table II. Solution of nonsymmetric linear systems: the errors E_{LSQR} and E_{TSVD} are optimal for LSQR and truncated singular value decomposition (TSVD). The corresponding truncation parameters are denoted by k_{LSQR} and k_{TSVD}. Three noise levels are considered; ℓ denotes the number of Golub–Kahan iterations performed.

Noise	Matrix	ℓ	E_{LSQR}	k_{LSQR}	E_{TSVD}	k_{TSVD}
$\delta = 10^{-6}$	Baart	10	5.1e-02	6	5.1e-02	6
	Heat	196	5.3e-03	54	5.4e-03	74
	Lotkin	18	3.1e-01	10	3.1e-01	10
	Tomo	195	7.6e-03	195	7.6e-03	195
	Wing	7	3.3e-01	5	3.3e-01	5
$\delta = 10^{-4}$	Baart	10	7.7e-02	5	7.7e-02	5
	Heat	196	1.5e-02	26	1.5e-02	37
	Lotkin	18	4.3e-01	7	4.3e-01	7
	Tomo	195	2.1e-02	195	2.3e-02	195
	Wing	7	4.5e-01	4	4.5e-01	4
$\delta = 10^{-2}$	Baart	10	1.5e-01	3	1.5e-01	3
	Heat	196	9.4e-02	13	9.8e-02	21
	Lotkin	18	4.5e-01	3	4.5e-01	3
	Tomo	195	1.9e-01	48	2.0e-01	180
	Wing	7	6.0e-01	2	6.0e-01	2

DOI: 10.1002/nla

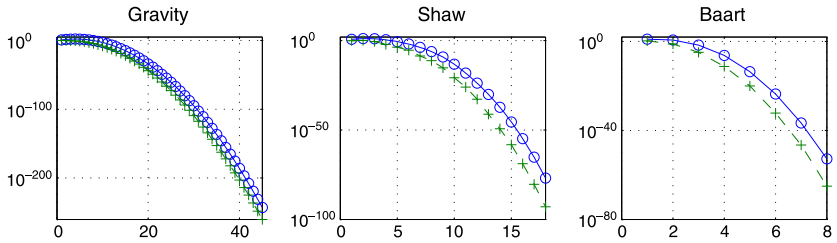

Figure 1. Behavior of the bounds (2.3) (left), (2.9) (center), and (3.5) (right), with respect to the iteration index m. The first test matrix is symmetric positive definite; the second is symmetric undefinite; the third is unsymmetric. The left-hand side of each inequality is represented by crosses and the right-hand side by circles.

corresponding either to a breakdown of the algorithm or to the last nonzero value of both inequality sides. The graphs show that the bounds provided by Theorems 2.1 and 2.4, and by Corollary 3.1, are quite sharp.

Now we illustrate that subspaces $\mathcal{R}(\breve{V}_k)$ generated by the Lanczos method (2.12) essentially contain subspaces of eigenvectors of A associated with the eigenvalues of largest magnitude. We also discuss the convergence of the largest eigenvalues of the matrices $\breve{T}_k$ in (2.12) to eigenvalues of A of largest magnitude. Here, $\breve{T}_k \in \mathbb{R}^{k\times k}$ is the matrix obtained by neglecting the last row of the matrix $\breve{T}_{k+1,k} \in \mathbb{R}^{(k+1)\times k}$ defined by (2.12) with m replaced by k. The Lanczos method is applied for n steps or until breakdown occurs, that is, a subdiagonal element of $\breve{T}_k$ is smaller than 10^{-12}; ℓ denotes the number of steps performed by the method. The initial column of the matrices $\breve{V}_k$ is $Ab_{\text{true}}/\|Ab_{\text{true}}\|$.

Let $\{\breve{\lambda}_i^{(k)}\}_{i=1}^{k}$ denote the eigenvalues of the matrix $\breve{T}_k$. We compare the eigenvalues $\breve{\lambda}_i^{(k)}$ of largest magnitude to the corresponding eigenvalues λ_i of the matrix A. All eigenvalues are ordered according to decreasing magnitude. For each Lanczos step k, we compute the relative difference

$$R_{\lambda,k} := \max_{i=1,2,\ldots,\lceil \frac{k}{3} \rceil} \frac{|\breve{\lambda}_i^{(k)} - \lambda_i|}{|\lambda_i|}. \tag{4.1}$$

Thus, we evaluate the maximum relative difference over the $\lceil \frac{k}{3} \rceil$ eigenvalues of largest modulus; $\lceil q \rceil$ denotes the integer closest to $q \in \mathbb{R}$. The graphs for $R_{\lambda,k}$, for $k = 1, 2, \ldots, \ell$, are displayed in the left column of Figure 2 for each of the five symmetric test matrices.

We turn to a comparison of subspaces. For each k, let $\breve{T}_k = \breve{W}_k \breve{\Lambda}_k \breve{W}_k^T$ be the spectral factorization of $\breve{T}_k$, where

$$\breve{\Lambda}_k = \operatorname{diag}\left[\breve{\lambda}_1^{(k)}, \breve{\lambda}_2^{(k)}, \ldots, \breve{\lambda}_k^{(k)}\right], \quad \breve{W}_k = \left[\breve{w}_1^{(k)}, \breve{w}_2^{(k)}, \ldots, \breve{w}_k^{(k)}\right],$$

and introduce the matrix $V_{k,i} = [v_1^{(k)}, v_2^{(k)}, \ldots, v_i^{(k)}]$ consisting of the first i columns of $V_k \breve{W}_k$. The columns of $V_{k,i}$ are the Ritz vectors of A associated with the i Ritz values of largest magnitude, $\breve{\lambda}_1^{(k)}, \breve{\lambda}_2^{(k)}, \ldots, \breve{\lambda}_i^{(k)}$. Partition the matrix containing the eigenvectors of A, compare with (2.1), according to $W = [W_i^{(1)} \; W_{n-i}^{(2)}]$, where $W_i^{(1)} \in \mathbb{R}^{n\times i}$ contains the first i eigenvectors and $W_{n-i}^{(2)} \in \mathbb{R}^{n\times(n-i)}$ the remaining ones. The columns of $W_i^{(1)}$ and $W_{n-i}^{(2)}$ span orthogonal subspaces. We compute, for $k = 1, 2, \ldots, \ell$, the quantities

$$R_{w,k} := \max_{i=1,2,\ldots,\lceil \frac{k}{3} \rceil} \|V_{k,i}^T W_{n-i}^{(2)}\|. \tag{4.2}$$

The norm $\|V_{k,i}^T W_{n-i}^{(2)}\|$ measures the distance between the subspaces $\mathcal{R}(V_{k,i})$ and $\mathcal{R}(W_i^{(1)})$; see, for example, [13]. Thus, $R_{w,k}$ is small when $\operatorname{span}\{w_j\}_{j=1}^{\lceil \frac{k}{3} \rceil}$ is approximately contained in $\operatorname{span}\{v_j^{(k)}\}_{j=1}^{k}$, that is, when the solution subspace generated by the Lanczos vectors essentially

DOI: 10.1002/nla

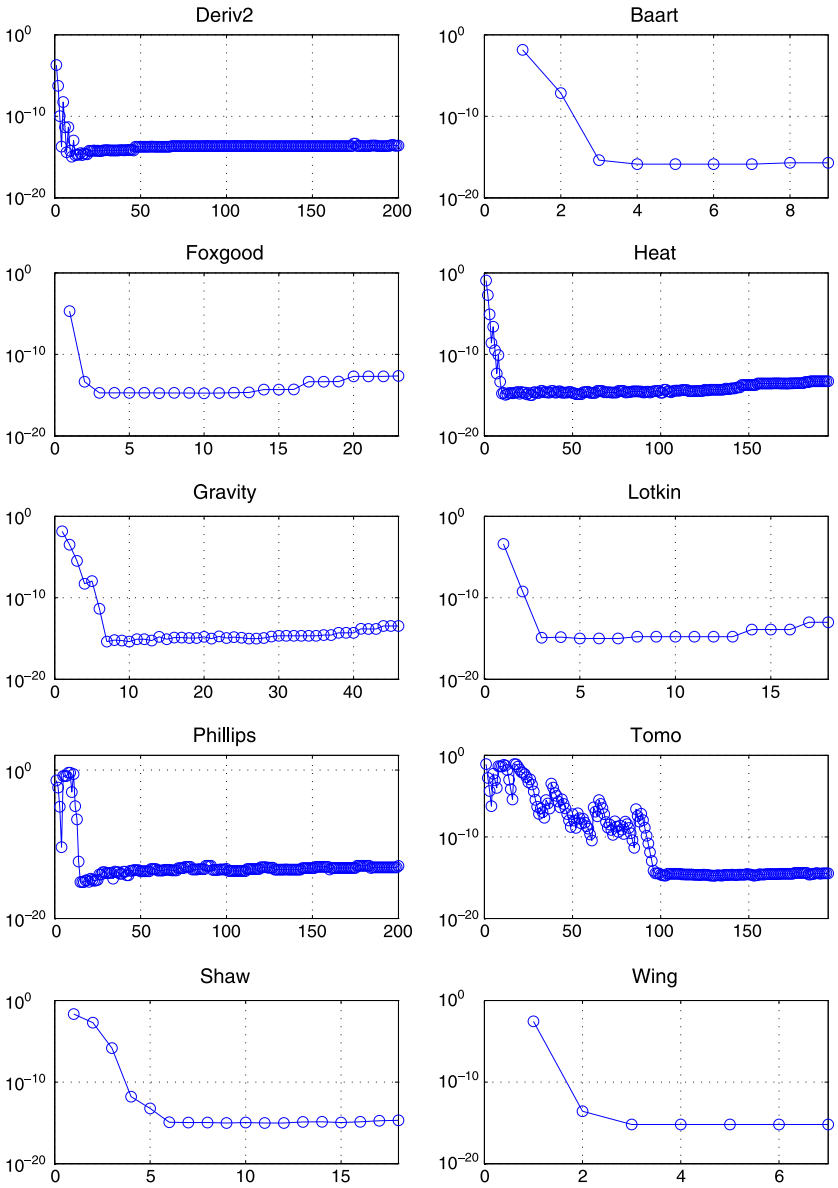

Figure 2. The graphs in the left column display the relative error $R_{\lambda,k}$ between the eigenvalues of the symmetric test problems and the corresponding Ritz values generated by the Lanczos process. The right column shows the behavior of $R_{\sigma,k}$ for the unsymmetric problems; see (4.1) and (4.3).

contains the space generated by the first $\lceil \frac{k}{3} \rceil$ eigenvectors. The graphs in the left column of Figure 3 shows $R_{w,k}$, for $k = 1, 2, \ldots, \ell$, for the symmetric test matrices. The distances between subspaces $\| V_{k,i}^T W_{n-i}^{(2)} \|$ are displayed in Figure 4, for $k = 10$ and $i = 1, \ldots, k$, for two symmetric test matrices.

A few comments on the left-hand side graphs of Figures 2 and 3 are in order. The left graphs of Figure 3 show that the span of the first $\lceil \frac{k}{3} \rceil$ eigenvectors of A are numerically contained in the span of the first k Lanczos vectors already for quite small values of k. We remark that this is not true if we compare the spaces spanned by the first k eigenvectors of A and by its first k Lanczos vectors. Graphs that compare the span of the first $\lceil \frac{k}{2} \rceil$ eigenvectors of A with the span of the first k Lanczos vectors look similar to the graphs shown but display slower convergence; see graphs in the left column of Figure 5. Thus, k has to be larger in order for the first $\lceil \frac{k}{2} \rceil$ eigenvectors of A

DOI: 10.1002/nla

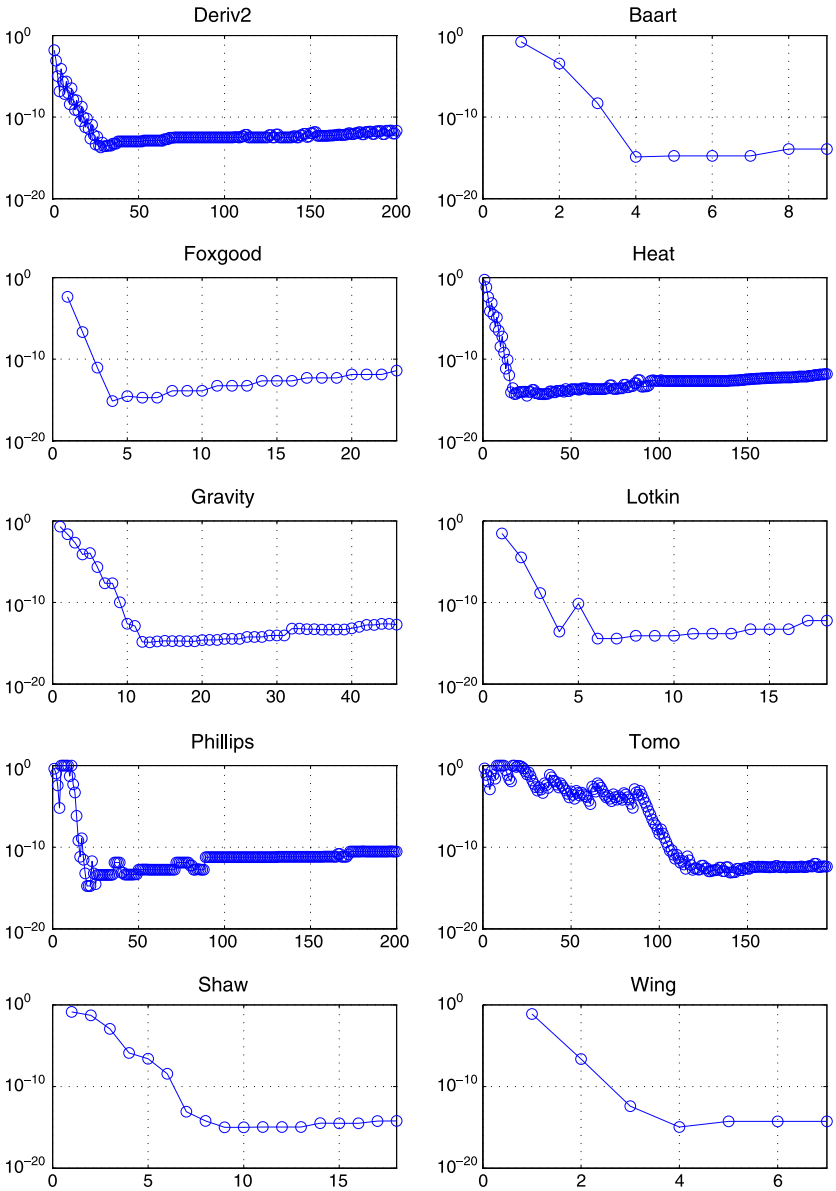

Figure 3. Distance between the subspace spanned by the first $\lceil k/3 \rceil$ eigenvectors (resp. singular vectors) of the symmetric (resp. nonsymmetric) test problems and the subspace spanned by the corresponding Lanczos (resp. Golub–Kahan) vectors; see (4.2) and (4.4).

to be numerically in the span of the first k Lanczos vectors. Figure 2 shows excellent agreement between the first $\lceil \frac{k}{3} \rceil$ Ritz values of A and the corresponding eigenvalues already for small k. The convergence of the first $\lceil \frac{k}{2} \rceil$ Ritz values to the corresponding eigenvalues is somewhat slower than the convergence displayed.

We use the Lanczos decomposition (2.12) in our illustrations because this decomposition gives approximate solutions of (1.1) and (3.1) of higher quality than the decompositions (1.3). Analogs of Figures 2 and 3 based on the Lanczos decomposition (1.3) look essentially the same as the figures shown. Finally, we remark that the Lanczos decomposition (2.12) is computed with reorthogonalization. Without reorthogonalization, the convergence illustrated by Figures 2 and 3 does not hold.

We turn to nonsymmetric matrices A. The Lanczos method is replaced by the Golub–Kahan method (3.2) and the spectral factorization by the singular value decomposition (3.6). The index ℓ

DOI: 10.1002/nla

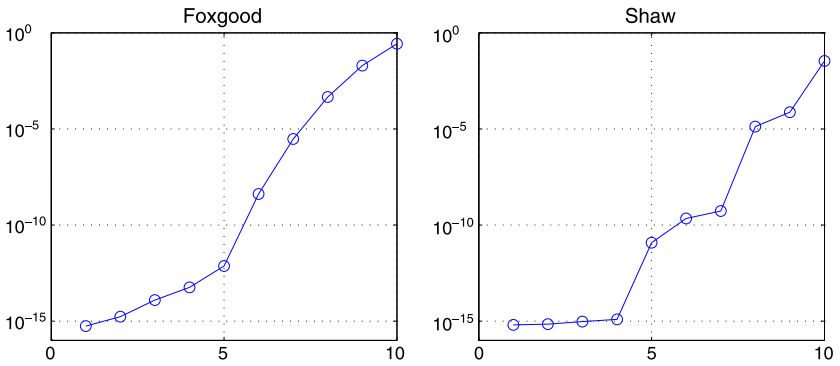

Figure 4. Distance $\|V_{k,i}^T W_{n-i}^{(2)}\|$, $i = 1, 2, \ldots, k$, between the subspace spanned by the first i eigenvectors of the Foxgood (left) and Shaw (right) matrices, and the subspace spanned by the corresponding i Ritz vectors at iteration $k = 10$.

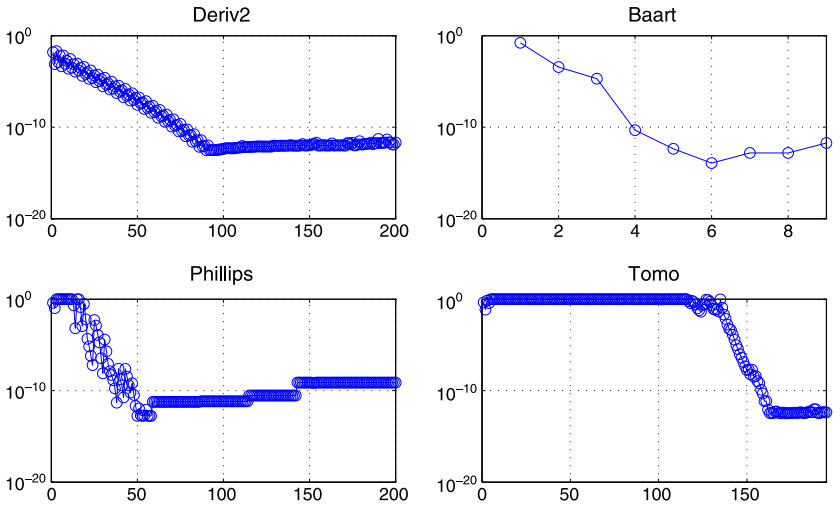

Figure 5. Distance between the subspace spanned by the first $\lceil k/2 \rceil$ eigenvectors (resp. singular vectors) of selected symmetric (resp. nonsymmetric) test problems and the subspace spanned by the corresponding Lanczos (resp. Golub–Kahan) vectors. The index ℓ ranges from 1 to either the dimension of the matrix ($n = 200$) or to the iteration where there is a breakdown in the factorization process.

denotes either the order n of the matrix or the step at which an element of the bidiagonal projected matrix $B_{\ell+1,\ell}$ is less than 10^{-12}, that is, the step at which a breakdown happens. The graphs in the right-hand side column of Figure 2 show the relative differences

$$R_{\sigma,k} := \max_{i=1,2,\ldots,\lceil \frac{k}{3} \rceil} \frac{|\breve{\sigma}_i^{(k)} - \sigma_i|}{|\sigma_i|} \tag{4.3}$$

between the singular values $\{\breve{\sigma}_i^{(k)}\}_{i=1}^k$ of $B_{k+1,k}$ and those of A. The graphs are similar to those in the left-hand side column, except for the example Tomo, which displays slow convergence; this behavior is probably linked to the fact that the Tomo test problem is much less ill conditioned than the other test problems.

Let Z and W be the orthogonal matrices in the singular value decomposition (3.6) of A, and partition these matrices similarly as we did for symmetric matrices A, that is, $W = [W_i^{(1)}\ W_{n-i}^{(2)}]$ and $Z = [Z_i^{(1)}\ Z_{n-i}^{(2)}]$, where the submatrices $W_i^{(1)}, Z_i^{(1)}$ contain the first i singular vectors and $W_{n-i}^{(2)}, Z_{n-i}^{(2)}$ the remaining $n - i$ ones. To investigate the convergence of subspaces, we substitute

DOI: 10.1002/nla

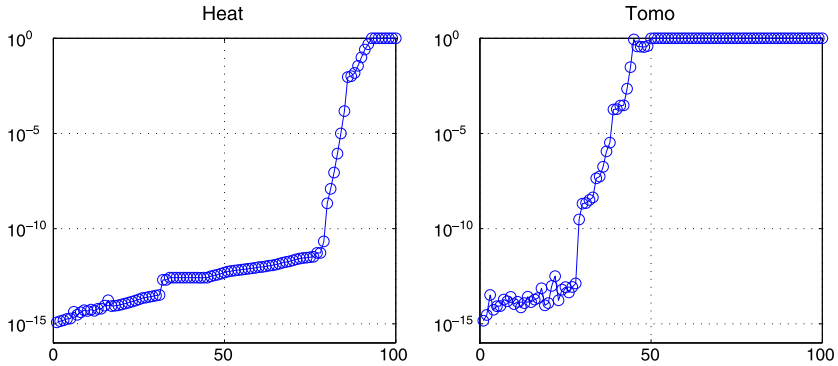

Figure 6. Distance $\max\{\|V_{k,i}^T W_{n-i}^{(2)}\|, \|U_{k,i}^T Z_{n-i}^{(2)}\|\}$, $i = 1, 2, \ldots, k$, between the subspace spanned by the first i singular vectors of the Heat (left) and Tomo (right) matrices and the subspace spanned by the corresponding i Golub–Kahan vectors at iteration $k = 100$.

the singular value decomposition $B_{k+1,k} = \breve{Z}_{k+1} \breve{\Sigma}_{k+1,k} \breve{W}_k^T$ into (3.2) and consider

$$R_{(z,w),k} := \max_{i=1,2,\ldots,\lceil \frac{k}{3} \rceil} \left\{ \|V_{k,i}^T W_{n-i}^{(2)}\|, \|U_{k,i}^T Z_{n-i}^{(2)}\| \right\}, \tag{4.4}$$

where $V_{k,i}$ and $U_{k,i}$ are made up of the first i columns of $V_k \breve{W}_k$ and $U_{k+1} \breve{Z}_{k+1}$, respectively. Then $R_{(z,w),k}$ measures the distance between subspaces determined by the singular vectors of A and those defined by vectors computed with the Golub–Kahan method.

The quantities $R_{(z,w),k}$ are displayed, for $k = 1, 2, \ldots, \ell$, in the right column of Figure 3. Figure 5 depicts graphs for the quantities $R_{w,k}$ and $R_{(z,w),k}$ with the maximum computed over the first $\lceil k/2 \rceil$ vectors for four test problems. Figure 6 shows the value taken by $\max\{\|V_{k,i}^T W_{n-i}^{(2)}\|, \|U_{k,i}^T Z_{n-i}^{(2)}\|\}$, $i = 1, 2, \ldots, 100$, which express the distance between spaces spanned by singular vectors and vectors determined by the Golub–Kahan method.

We now compare the performances of different regularization methods. The test problems from [17] define both a matrix A and a solution x_{true}; the solution of the Lotkin example is the same as for the Shaw example. The error-free data vector is defined by $b_{\text{true}} := Ax_{\text{true}}$, and the contaminated data vector is given by (1.2) with

$$\eta := \widehat{\eta} \, \|b_{\text{true}}\| \frac{\delta}{\sqrt{n}},$$

where the random vector $\widehat{\eta} \in \mathbb{R}^n$ models Gaussian noise with mean zero and variance one and δ is a chosen noise level. In our experiments, we let $\delta = 10^{-6}, 10^{-4}, 10^{-2}$.

We measure the accuracy attainable by each regularization method by the relative error

$$E_{\text{method}} = \frac{\|x_{k_{\text{method}}} - x_{\text{true}}\|}{\|x_{\text{true}}\|} = \min_{k=1,2,\ldots,\ell} \frac{\|x_k - x_{\text{true}}\|}{\|x_{\text{true}}\|}, \tag{4.5}$$

which is obtained by choosing the value $k = k_{\text{method}}$ that minimizes the error for the method under consideration.

Table I reports the result obtained by comparing truncated Lanczos iteration (1.6) to truncated eigenvalue solution for symmetric test problems. The minimal errors (4.5) obtained by applying the Lanczos method and the truncated eigenvalue decomposition method, denoted by E_{Lanczos} and E_{TEIG}, are reported in the fourth and sixth columns, respectively. The truncation parameters that produce the minimal errors are listed in the fifth and seventh columns. The third column shows how many Lanczos iterations were executed; an entry smaller than 200 indicates that breakdown occurred. Both errors and truncation parameters values are averages over 20 realization of the

DOI: 10.1002/nla

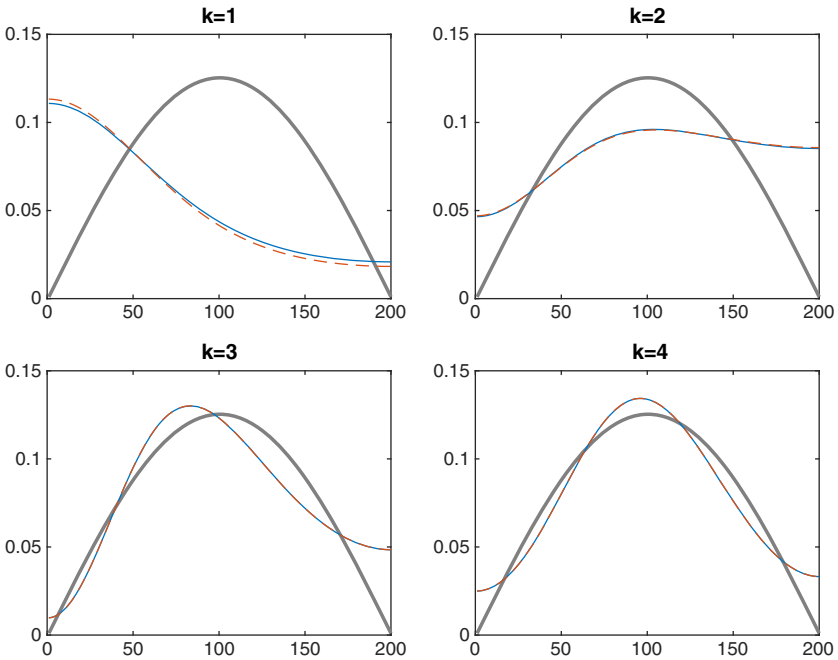

Figure 7. The first four LSQR solutions to the Baart test problem (thin lines) are compared with the corresponding truncated singular value decomposition (TSVD) solutions (dashed lines) and with the exact solution (thick line). The size of the problem is $n = 200$; the noise level is $\delta = 10^{-4}$. The thin and dashed lines are very close.

random noise. Three noise levels δ were considered. The results in Table I suggest that, for the test problems considered, the truncated Lanczos projection method is able to produce solutions essentially equivalent to those obtained by truncated eigenvalue decomposition, with a cheaper algorithm, as the number of iterations required is sometimes far less than the number of singular values required.

Table II reports results obtained for nonsymmetric linear discrete ill-posed problems (3.1). Here, the LSQR method is compared with TSVD. The table confirms the conclusions deduced from Table I.

Figure 7 displays the first four regularized solutions produced by the LSQR and TSVD methods when applied to solve the Baart test problem with a noise-contaminated vector b. The noise level is $\delta = 10^{-4}$. The approximate solutions determined by the LSQR and TSVD methods can be seen to approach each other when the number of iterations k or the truncation parameter k is increased from one to four.

The Tomo test problems arises from the discretization of a 2D tomography problem. Its numerical solution displays some interesting features. It is clear from Table II that, when the noise level is large, LSQR produces an approximate solution after k_{LSQR} steps that is of essentially the same quality as approximate solutions determined by TSVD with a truncation parameter k_{TSVD} that is much larger than k_{LSQR}. To better understand this behavior, we consider an image of size 15×15 pixels. This gives rise to a minimization problem (3.1) with a matrix $A \in \mathbb{R}^{225\times 225}$. Figure 8 shows the relative errors E_{LSQR} and E_{TSVD} as functions of the parameter k. LSQR is seen to give much faster convergence to x_{true}. The best attainable approximate solutions by LSQR and TSVD are displayed in Figure 9. The LSQR method yields the best approximation of x_{true} at step $k_{\mathrm{LSQR}} = 66$. The upper right plot displays this computed solution; the image x_{true} is shown in the upper left plot. For comparison, the lower left plot of Figure 9 shows the TSVD solution for $k = 66$. This restoration is seen to be of poor quality. The best approximate solution determined by TSVD has truncation index $k_{TSVD} = 216$; it can be seen to be of about the same quality as the best LSQR solution. The results for nonsymmetric problems agree with the ones presented by Hanke [18].

DOI: 10.1002/nla

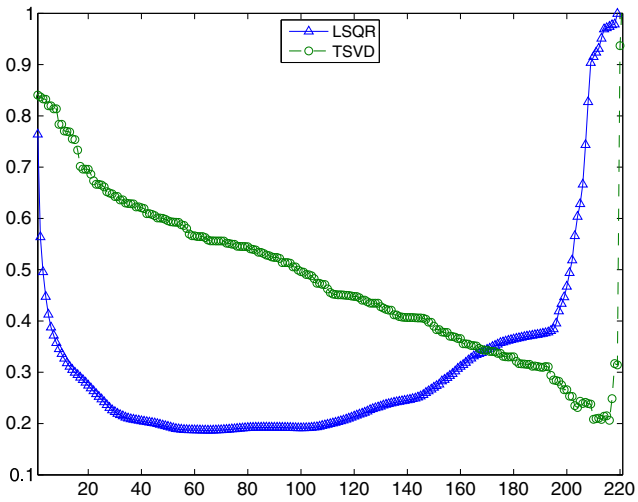

Figure 8. Convergence history for the LSQR and truncated singular value decomposition (TSVD) solutions to the Tomo example of size $n = 225$, with noise level $\delta = 10^{-2}$. The error E_{LSQR} has a minimum at $k = 66$, while E_{TSVD} is minimal for $k = 215$.

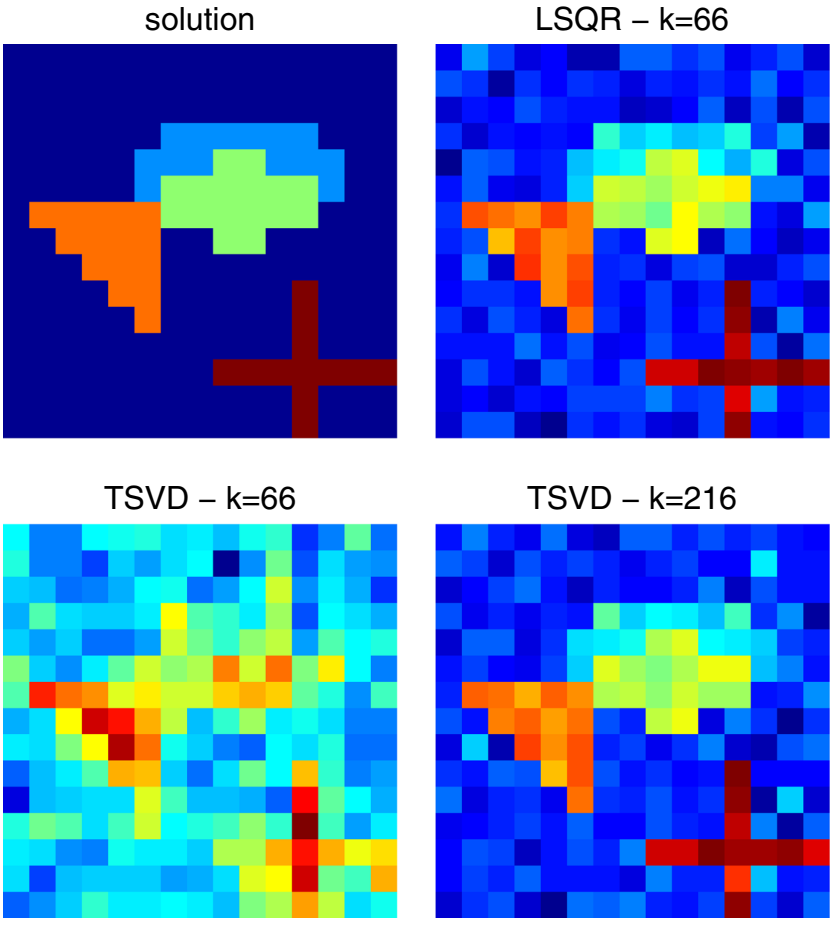

Figure 9. Solution by LSQR and truncated singular value decomposition (TSVD) to the Tomo example of size $n = 225$, with noise level $\delta = 10^{-2}$: exact solution (top left), optimal LSQR solution (top right), TSVD solution corresponding to the same truncation parameter (bottom left), and optimal TSVD solution (bottom right).

DOI: 10.1002/nla

5. CONCLUSION

This paper shows that the largest eigenvalues (in magnitude) of symmetric matrices and the largest singular values of nonsymmetric matrices that are defined by linear discrete ill-posed problems are well approximated by the corresponding eigenvalues and singular values of projected problems determined by a few steps of the Lanczos or Golub–Kahan bidiagonalization methods, respectively. This is similar for the corresponding eigenvectors and singular vectors, and suggests that it often suffices to use a partial Lanczos decomposition or a partial Golub–Kahan bidiagonalization, which are cheaper to compute than partial spectral or singular value decompositions, to determine a solution. Computed examples provide illustrations.

ACKNOWLEDGEMENTS

The work of S. G. was partially supported by the 'Giovani Ricercatori 2014' GNCS-INdAM project. The work of L. R. was supported by the University of Cagliari RAS Visiting Professor Program and by NSF grant DMS-1115385. The work of G. R. was partially supported by INdAM-GNCS.

REFERENCES

1. Saad Y. *Iterative Methods for Sparse Linear Systems* (2nd edn). SIAM: Philadelphia, 2003.
2. Saad Y. *Numerical Methods for Large Eigenvalue Problems,* Revised Edition. SIAM: Philadelphia, 2011.
3. Engl HW, Hanke M, Neubauer A. *Regularization of Inverse Problems*. Kluwer: Dordrecht, 1996.
4. Hansen PC. *Rank-deficient and Discrete Ill-posed Problems*. SIAM: Philadelphia, 1998.
5. Gazzola S, Novati P, Russo MR. Embedded techniques for choosing the parameter in Tikhonov regularization. *Numerical Linear Algebra with Applications* 2014; **21**:796–812.
6. Gazzola S, Novati P, Russo MR. On Krylov projection methods and Tikhonov regularization method. *Electronic Transactions on Numerical Analysis* 2015; **44**:83–123.
7. Novati P, Russo MR. A GCV based Arnoldi–Tikhonov regularization method. *BIT Numerical Mathematics* 2014; **54**:501–521.
8. Trefethen LN, Embree M. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton University Press: Princeton, 2005.
9. Baglama J, Calvetti D, Reichel L. IRBL: an implicitly restarted block Lanczos method for large-scale Hermitian eigenproblems. *SIAM Journal on Scientific Computing* 2003; **24**:1650–1677.
10. Calvetti D, Lewis B, Reichel L. On the choice of subspace for iterative methods for linear discrete ill-posed problems. *International Journal of Applied Mathematics and Computer Science* 2001; **11**:1069–1092.
11. Dykes L, Marcellán F, Reichel L. The structure of iterative methods for symmetric linear discrete ill-posed problems. *BIT Numerical Mathematics* 2014; **54**:129–145.
12. Neuman A, Reichel L, Sadok H. Implementations of range restricted iterative methods for linear discrete ill-posed problems. *Linear Algebra and Its Applications* 2012; **436**:3974–3990.
13. Golub GH, Van Loan CF. *Matrix Computations* (4th ed.) Johns Hopkins University Press: Baltimore, 2013.
14. Paige CC, Saunders MA. An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software* 1982; **8**:43–71.
15. Baglama J, Reichel L. Augmented implicitly restarted Lanczos bidiagonalization methods. *SIAM Journal on Scientific Computing* 2005; **27**:19–42.
16. Baglama J, Reichel L. An implicitly restarted block Lanczos bidiagonalization method using Leja shifts. *BIT Numerical Mathematics* 2013; **53**:285–310.
17. Hansen PC. Regularization tools version 4.0 for Matlab 7.3. *Numerical Algorithms* 2007; **46**:189–194.
18. Hanke M. On Lanczos based methods for the regularization of discrete ill-posed problems. *BIT Numerical Mathematics* 2001; **41**:1008–1018.

DOI: 10.1002/nla