

# Probabilistic upper bounds for the matrix two-norm

Michiel E. Hochstenbach

Received: date / Accepted: date

**Abstract** We develop probabilistic upper bounds for the matrix two-norm, the largest singular value. These bounds, which are true upper bounds with a user-chosen high probability, are derived with a number of different polynomials that implicitly arise in the Lanczos bidiagonalization process. Since these polynomials are adaptively generated, the bounds typically give very good results. They can be computed efficiently. Together with an approximation that is a guaranteed lower bound, this may result in a small probabilistic interval for the matrix norm of large matrices within a fraction of a second.

**Keywords** Matrix two-norm · probabilistic bound · SVD · singular value problem · singular value decomposition · subspace method · Lanczos bidiagonalization · Lanczos polynomial · Ritz polynomial · condition number · large (sparse) matrix.

**Mathematics Subject Classification (2000)** 65F35, 65F15, 65F50, 15A12, 15A18.

## 1 Introduction

(Golub–Kahan–) Lanczos bidiagonalization [5] (see also, e.g., [6]) is a popular method to approximate singular values of large sparse matrices. Let  $A$  be a real  $m \times n$  matrix with singular value decomposition (SVD)  $A = X\Sigma Y^T$  with singular values

$$0 \leq \sigma_{\min} = \sigma_p \leq \sigma_{p-1} \leq \dots \leq \sigma_2 \leq \sigma_1 = \sigma_{\max} = \|A\|,$$

where  $p := \min\{m, n\}$ , and  $\|\cdot\|$  stands for the 2-norm. Denote the corresponding right singular vectors by  $\mathbf{y}_1, \dots, \mathbf{y}_n$ . Usually, Lanczos bidiagonalization approximates the largest singular values, and, to a lesser extent, the smallest singular values, well. However, the results of the method depend on the choice of the initial vector  $\mathbf{v}_1$ . The obtained approximation to largest singular value  $\sigma_{\max}$  is always

---

M. E. Hochstenbach  
Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB, The Netherlands. [www.win.tue.nl/~hochsten](http://www.win.tue.nl/~hochsten).  
Supported by an NWO Vidi grant.

a lower bound. However, if a poor choice is made for  $\mathbf{v}_1$ , that is, if  $\mathbf{v}_1$  is almost deficient in the direction  $\mathbf{y}_1$ , the true value of  $\|A\|$  may be arbitrarily larger. Often there is no a priori information on  $\mathbf{y}_1$  available. For this reason a random choice for  $\mathbf{v}_1$  is considered relatively safe;  $\mathbf{v}_1$  is usually selected randomly in (industrial) codes.

Using the fact that  $\mathbf{v}_1$  is chosen randomly, we will develop probabilistic bounds for  $\|A\|$ ; i.e., bounds that hold with a user-selected probability  $1 - \varepsilon$ , for  $\varepsilon \ll 1$ . The bounds may be viewed as a side-product or post-processing step of Lanczos bidiagonalization and may be computed efficiently: for large  $A$ , the computational costs are very modest compared to the Lanczos bidiagonalization process itself.

The fact that it is unlikely that a random vector is near-deficient in  $\mathbf{y}_1$  enables us to develop probabilistic inclusion intervals for the matrix norm. Hereby we exploit the fact that the Lanczos polynomials tend to increase rapidly to the right of its largest zero (see Section 2). Therefore, with our new low-cost process as addition to the Lanczos bidiagonalization method, we usually not only get good lower bounds to  $\|A\|$ , but also get sharp upper bounds with a high probability.

Efficient state-of-the-art methods based on Lanczos bidiagonalization use some restart mechanism; see, e.g., [1], [12]. We will not consider restarts in this paper for two main reasons: first, the unrestarted case makes possible the theoretical analysis of Sections 2, 3 and 4; and second, it will turn out that usually a modest number of Lanczos bidiagonalization steps already suffices for quality probabilistic inclusion intervals. We will also assume exact arithmetic; in the experiments in Section 5 we exploit a stable variant with reorthogonalization.

This paper is inspired by [13] and has been organized as follows. Section 2 studies polynomials that are implicitly formed in the Lanczos bidiagonalization process. These are used in Sections 3 and 4 to develop probabilistic upper bounds for the matrix 2-norm. Numerical experiments are presented in Section 5, and a discussion and some conclusions can be found in Section 6.

## 2 Polynomials arising in Lanczos bidiagonalization

Given a vector  $\mathbf{v}_1$  with unit norm, the defining relations of Lanczos bidiagonalization are  $\beta_0 = 0$ ,  $\mathbf{u}_0 = \mathbf{0}$ , and for  $k \geq 1$ :

$$\begin{aligned}\alpha_k \mathbf{u}_k &= A \mathbf{v}_k - \beta_{k-1} \mathbf{u}_{k-1} \\ \beta_k \mathbf{v}_{k+1} &= A^T \mathbf{u}_k - \alpha_k \mathbf{v}_k\end{aligned}\tag{1}$$

where

$$\alpha_j = \mathbf{u}_j^T A \mathbf{v}_j, \quad \beta_j = \mathbf{u}_j^T A \mathbf{v}_{j+1}\tag{2}$$

are nonnegative. After  $k$  steps of the method, these relations can be written in matrix form as

$$\begin{aligned}AV_k &= U_k B_k, \\ A^T U_k &= V_{k+1} \widehat{B}_k^T = V_k B_k^T + \beta_k \mathbf{v}_{k+1} \mathbf{e}_k^T,\end{aligned}$$

where  $\mathbf{e}_k$  is the  $k$ th unit vector, and  $U_k = [\mathbf{u}_1 \cdots \mathbf{u}_k]$  and  $V_k = [\mathbf{v}_1 \cdots \mathbf{v}_k]$  have orthonormal columns spanning the subspaces  $\mathcal{U}_k$  and  $\mathcal{V}_k$ , respectively. The  $k \times k$

matrix

$$B_k = \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ & \ddots & \ddots & & \\ & & \alpha_{k-1} & \beta_{k-1} & \\ & & & & \alpha_k \end{bmatrix}$$

and the  $k \times (k+1)$  matrix  $\widehat{B}_k = [B_k \ \mathbf{0}] + \beta_k \mathbf{e}_k \mathbf{e}_{k+1}^T$  are both upper bidiagonal matrices. We will not consider the rather exceptional situation of a breakdown of the method (a zero  $\alpha_j$  or  $\beta_j$ ) in this paper.

Introduce the bilinear forms

$$\langle f, g \rangle := \mathbf{v}_1^T f(A^T A) g(A^T A) \mathbf{v}_1$$

and

$$[f, g] := \mathbf{v}_1^T A^T f(AA^T) g(AA^T) A \mathbf{v}_1 = \mathbf{v}_1^T f(A^T A) A^T A g(A^T A) \mathbf{v}_1$$

for functions  $f$  and  $g$  that are analytic in a neighborhood of the squares of the singular values of  $A$ . The following result is the starting point for this paper.

**Proposition 1** *The  $\mathbf{u}_k$  and  $\mathbf{v}_k$  can be written as a polynomial of degree  $k-1$  in  $AA^T$ , resp.  $A^T A$ , applied to  $A \mathbf{v}_1$ , resp.  $\mathbf{v}_1$ :*

$$\mathbf{u}_k = p_{k-1}(AA^T) A \mathbf{v}_1, \quad \mathbf{v}_k = q_{k-1}(A^T A) \mathbf{v}_1.$$

The following recurrence relations hold:  $p_{-1}(t) = 0$ ,  $q_0(t) = 1$ , and for  $k \geq 0$ :

$$\begin{aligned} \alpha_{k+1} p_k(t) &= q_k(t) - \beta_k p_{k-1}(t), \\ \beta_{k+1} q_{k+1}(t) &= t p_k(t) - \alpha_{k+1} q_k(t). \end{aligned}$$

Moreover,

$$\begin{aligned} \alpha_k &= \langle p_{k-1}, t q_{k-1} \rangle = [p_{k-1}, q_{k-1}], \\ \beta_k &= \langle p_{k-1}, t q_k \rangle = [p_{k-1}, q_k]. \end{aligned}$$

*Proof* This follows by induction; the recurrence relations follow from substitution into (1). The inner products can be derived from (2).

We now study several useful properties of these Lanczos bidiagonalization polynomials  $p_k$  and  $q_k$  that will be used in the rest of the paper. First, we point out close relations between Lanczos bidiagonalization and two other Lanczos processes. Note that

$$\begin{aligned} A^T A V_k &= A^T U_k B_k \\ &= V_k B_k^T B_k + \beta_k \mathbf{v}_{k+1} \mathbf{e}_k^T B_k \\ &= V_k B_k^T B_k + \alpha_k \beta_k \mathbf{v}_{k+1} \mathbf{e}_k^T \end{aligned} \tag{3}$$

and

$$\begin{aligned} AA^T U_k &= A V_k B_k^T + \beta_k A \mathbf{v}_{k+1} \mathbf{e}_k^T \\ &= U_k B_k B_k^T + \beta_k U_{k+1} B_{k+1} \mathbf{e}_{k+1} \mathbf{e}_k^T \\ &= U_k \widehat{B}_k \widehat{B}_k^T + \alpha_{k+1} \beta_k \mathbf{u}_{k+1} \mathbf{e}_k^T. \end{aligned} \tag{4}$$

We see from these equations that Lanczos bidiagonalization simultaneously performs a Lanczos process on  $A^T A$  with starting vector  $\mathbf{v}_1$ , and a Lanczos process on  $AA^T$  with starting vector  $\mathbf{u}_1 := \alpha_1^{-1} A \mathbf{v}_1$  (the normalized  $A \mathbf{v}_1$ ). The symmetric tridiagonal matrices  $B_k^T B_k$  and  $\widehat{B}_k \widehat{B}_k^T$ , respectively, that arise in the Lanczos methods are decomposed as the product of the bidiagonal matrices that arise in Lanczos bidiagonalization. We use (3) and (4) to characterize the zeros of the polynomials  $p_k$  and  $q_k$ ; see Proposition 2.

Denote the singular values of  $B_k$  by

$$\theta_k^{(k)} \leq \dots \leq \theta_1^{(k)}$$

and the corresponding right singular vectors by  $\mathbf{d}_k^{(k)}, \dots, \mathbf{d}_1^{(k)}$ . We write  $\widehat{\theta}_k^{(k)} \leq \dots \leq \widehat{\theta}_1^{(k)}$  for the singular values of  $\widehat{B}_k$  and  $\widehat{\mathbf{c}}_k^{(k)}, \dots, \widehat{\mathbf{c}}_1^{(k)}$  for its left singular vectors. To avoid a heavy notation we will often omit the superscript  $(k)$  in the sequel. A key aspect of Lanczos bidiagonalization is that often the singular values of both  $B_k$  and  $\widehat{B}_k$  are good approximations to the singular values of  $A$ ; in particular to the largest and (to a lesser extent) to the smallest singular values.

In the next proposition,  $I_k$  stands for the identity of dimension  $k$ .

**Proposition 2** (a) *The zeros of  $q_k$  are exactly  $\theta_1^2, \dots, \theta_k^2$ .*

*This implies that  $q_k(t)$  is a nonzero multiple of  $\det(tI_k - B_k^T B_k)$ .*

(b) *The zeros of  $p_k$  are exactly  $\widehat{\theta}_1^2, \dots, \widehat{\theta}_k^2$ .*

*This implies that  $p_k(t)$  is a nonzero multiple of  $\det(tI_k - \widehat{B}_k \widehat{B}_k^T)$ .*

*Proof* From (3) it may be checked that the pairs  $(\theta_j^2, V_k \mathbf{d}_j)$ ,  $j = 1, \dots, k$ , satisfy the Galerkin condition

$$A^T A V_k \mathbf{d}_j - \theta_j^2 V_k \mathbf{d}_j \perp \mathcal{V}_k.$$

Since  $V_k \mathbf{d}_j \in \mathcal{V}_k$ , we can write

$$V_k \mathbf{d}_j = s_j(A^T A) \mathbf{v}_1 \tag{5}$$

for a polynomial  $s_j = s_j^{(k)}$  of degree at most  $k-1$ . For all  $j = 1, \dots, k$ , we have that  $(A^T A - \theta_j^2 I) V_k \mathbf{d}_j$  is in  $\mathcal{V}_{k+1}$  but is orthogonal to  $\mathcal{V}_k$ . Therefore, these vectors have to be nonzero multiples of the vector  $\mathbf{v}_{k+1} = q_k(A^T A) \mathbf{v}_1$ . Hence,  $q_k(t)$  should contain all factors  $(t - \theta_j^2)$ , and therefore is a nonzero multiple of

$$\mu(t) = (t - \theta_1^2) \dots (t - \theta_k^2).$$

Part (b) follows in a similar manner starting with the Galerkin condition

$$AA^T U_k \widehat{\mathbf{c}}_j - \widehat{\theta}_j^2 U_k \widehat{\mathbf{c}}_j \perp \mathcal{U}_k$$

for the pairs  $(\widehat{\theta}_j^2, U_k \widehat{\mathbf{c}}_j)$ . Since  $U_k \widehat{\mathbf{c}}_j \in \mathcal{U}_k$ , we can write

$$U_k \widehat{\mathbf{c}}_j = r_j(AA^T) A \mathbf{v}_1 \tag{6}$$

for a polynomial  $r_j = r_j^{(k)}$  of degree at most  $k-1$ . For all  $j = 1, \dots, k$ , we have that  $(AA^T - \widehat{\theta}_j^2 I) U_k \widehat{\mathbf{c}}_j$  is in  $\mathcal{U}_{k+1}$  but is orthogonal to  $\mathcal{U}_k$ . Therefore, these vectors have

to be nonzero multiples of the vector  $\mathbf{u}_{k+1} = p_k(AA^T)A\mathbf{v}_1$ . Hence,  $p_k(t)$  should contain all factors  $(t - \widehat{\theta}_j^2)$ , and therefore is a nonzero multiple of

$$\widehat{\mu}(t) = (t - \widehat{\theta}_1^2) \cdots (t - \widehat{\theta}_k^2);$$

cf. also the discussion in [11, p. 266–267].

**Corollary 3**

$$\begin{aligned} V_k \mathbf{d}_j &= \nu_j(A^T A) \mathbf{v}_1 / \|\nu_j(A^T A) \mathbf{v}_1\| \quad (j = 1, \dots, k), \\ \|\mu(A^T A) \mathbf{v}_1\| &= \min \|\omega(A^T A) \mathbf{v}_1\|, \\ U_k \widehat{\mathbf{c}}_j &= \widehat{\nu}_j(AA^T) A \mathbf{v}_1 / \|\widehat{\nu}_j(AA^T) A \mathbf{v}_1\|, \quad (j = 1, \dots, k), \\ \|\widehat{\mu}(AA^T) A \mathbf{v}_1\| &= \min \|\omega(AA^T) A \mathbf{v}_1\|, \end{aligned}$$

where  $\nu_j(t) = \mu(t)/(t - \theta_j^2)$ ,  $\widehat{\nu}_j(t) = \widehat{\mu}(t)/(t - \widehat{\theta}_j^2)$ , and the minimum is taken over all monic polynomials  $\omega$  of degree  $k$ .

*Proof* This follows from the proof of the previous proposition; cf. also [11, p. 266].

The following result will be used for an efficient numerical procedure in the next section.

**Proposition 4** *The polynomials  $p_k$  and  $q_k$  have positive leading coefficients and increase strictly monotonically to the right of their largest zeros  $\widehat{\theta}_1^2$  and  $\theta_1^2$ , respectively.*

*Proof* This follows from Proposition 2 and the fact that  $p_k$  and  $q_k$  are polynomials of degree  $k$ .

**Proposition 5** *For  $1 \leq j \leq k$  the convergence to the largest singular values is monotonic:*

$$\theta_j^{(k)} \leq \widehat{\theta}_j^{(k)} \leq \theta_j^{(k+1)} \leq \sigma_j.$$

*Proof* This follows from the fact that  $\widehat{B}_k$  is the matrix  $B_k$  expanded with an extra  $(k+1)$ st column. Likewise,  $B_{k+1}$  is  $\widehat{B}_k$  expanded with an extra  $(k+1)$ st row. Now apply [8, (3.3.17)], see also [7, Theorem 4.3].

Taking  $j = 1$  in Proposition 5, this implies that the largest singular values of  $B_k$  and  $\widehat{B}_k$  are guaranteed lower bounds for  $\|A\|$  of increasing quality. Furthermore, the polynomials  $p_k$  and  $q_k$  will be used for probabilistic bounds for the matrix norm in the next section.

### 3 Probabilistic bounds for the matrix norm

We will now develop probabilistic bounds for  $\|A\|$  ( $= \sigma_1 = \sigma_{\max}$ ), making use of the fact that the polynomials  $p_k$  and  $q_k$  tend to increase rapidly to the right of their largest zeros  $\widehat{\theta}_1$  and  $\theta_1$ , respectively. Let

$$\mathbf{v}_1 = \sum_{j=1}^n \gamma_j \mathbf{y}_j$$

be the decomposition of the starting vector  $\mathbf{v}_1$  with respect to the right singular vectors.

**Lemma 6** *We have  $p_k(\sigma_1^2) > 0$  and  $q_k(\sigma_1^2) > 0$ .*

*Proof* This follows from the combination of Propositions 4 and 5.

We now arrive at the main argument. From

$$1 = \|\mathbf{v}_{k+1}\|^2 = \|q_k(A^T A)\mathbf{v}_1\|^2 = \sum_{j=1}^n \gamma_j^2 q_k(\sigma_j^2)^2$$

and  $q_k(\sigma_1^2) > 0$  (see Lemma 6) it follows that

$$1 \geq |\gamma_1| q_k(\sigma_1^2).$$

If  $\gamma_1$  would be known, this estimate would provide an upper bound  $\sigma_{\text{up}}$  for  $\|A\| = \sigma_{\text{max}}$ : let  $\sigma_{\text{up}}$  be the largest zero of

$$f_1(t) = q_k(t^2) - 1/|\gamma_1|. \quad (7)$$

One may check that this number  $\sigma_{\text{up}}$  exists and is larger than  $\theta_1 = \theta_1^{(k)}$ ; it may for instance be determined numerically efficiently by bisection on the interval  $[\theta_1^{(k)}, \|A\|_F]$  which is guaranteed to contain  $\sigma_{\text{max}}$ . (Note that  $\sigma_{\text{up}}$  might incidentally even be larger than  $\|A\|_F$  for small  $k$ ; in this case we proceed with a larger  $k$ , as the information is not useful.)

Since we generally do not know (an estimate to)  $\gamma_1$  in practice, we are interested in the probability that  $|\gamma_1|$  is smaller than a given (small) constant. A small  $|\gamma_1|$  corresponds to an unlucky choice of an initial vector: in this case  $\mathbf{v}_1$  is almost orthogonal to  $\mathbf{y}_1$ . The following lemma states a suitable result and enables us to establish *probabilistic bounds*, i.e., bounds that hold with a certain (user-defined, high) probability. The proof uses the fact that if  $\mathbf{v}_1$  has been chosen randomly with respect to the uniform distribution over the unit sphere  $S^{n-1}$  in  $\mathbb{R}^n$ , then, as a result,  $(\gamma_1, \dots, \gamma_n)$  is also random in  $S^{n-1}$ . It is easy to construct this random vector (Matlab code: `v1=randn(n,1)`; `v1=v1/norm(v1)`); see, e.g., [9, p. 1116].

**Lemma 7** *Assume that the starting vector  $\mathbf{v}_1$  has been chosen randomly with respect to the uniform distribution over the unit sphere  $S^{n-1}$  and let  $\delta \in [0, 1]$ . Then*

$$P(|\gamma_1| \leq \delta) = 2 G\left(\frac{n-1}{2}, \frac{1}{2}\right)^{-1} \cdot \int_0^{\arcsin(\delta)} \cos^{n-2}(t) dt,$$

where  $G$  denotes Euler's Beta function:  $G(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt$ .

*Proof* See [13, Lemma 3.1].

If we would like to have an upper bound for  $\|A\|$  that is correct with probability at least  $1 - \varepsilon$ , then we first determine the value of  $\delta$  for which

$$\int_0^{\arcsin(\delta)} \cos^{n-2}(t) dt = \frac{\varepsilon}{2} G\left(\frac{n-1}{2}, \frac{1}{2}\right) \quad \left( = \varepsilon \int_0^{\pi/2} \cos^{n-2}(t) dt \right) \quad (8)$$

holds, e.g., by bisection on the interval  $[0, \frac{\pi}{2}]$ . The integrals in (8) may be computed using an appropriate quadrature formula.

Moreover, for small  $\varepsilon$ , which is our main interest, the behavior of  $\delta$  as a function of  $\varepsilon$  is roughly  $\delta = \delta(\varepsilon) \approx \varepsilon \cdot \frac{1}{2} G(\frac{n-1}{2}, \frac{1}{2})$  as is proven in the next result. As an example, we mention that for  $n = 1000$  and  $\varepsilon = 0.01$ , the true and estimated value for  $\delta$  with Proposition 8 differ only  $\approx 2.6 \cdot 10^{-5}$  relatively.

**Proposition 8** *Given  $0 < \varepsilon \ll 1$ , let  $\delta = \delta(\varepsilon)$  satisfy (8). Then*

$$\delta'(0) = \lim_{\varepsilon \rightarrow 0} \frac{\delta(\varepsilon)}{\varepsilon} = \frac{1}{2} G(\frac{n-1}{2}, \frac{1}{2}).$$

*Proof* First note that  $\arcsin(\delta) = \delta + \mathcal{O}(\delta^3)$  for  $\delta \rightarrow 0$ . Let  $F(\delta(\varepsilon)) = \int_0^{\delta(\varepsilon)} \cos^{n-2}(t) dt$ . Then

$$\lim_{\varepsilon \rightarrow 0} \frac{F(\delta(\varepsilon)) - F(0)}{\varepsilon} = \cos^{n-2}(0) \cdot \delta'(0) = \frac{1}{2} G(\frac{n-1}{2}, \frac{1}{2}),$$

which proves the statement.

When we replace  $|\gamma_1|$  in (7) by the value  $\delta$  computed from (8) and determine the zero  $\sigma_{\text{up}} > \theta_1^{(k)}$ , this  $\sigma_{\text{up}}$  is an upper bound for the largest singular value  $\sigma_{\text{max}}$  of  $A$  with probability at least  $1 - \varepsilon$ , which we call a probabilistic upper bound. This zero may be computed efficiently, since the evaluation of  $p_k$  and  $q_k$  may be carried out via the recurrence relations as in Proposition 1. (Note that a loop is often preferable over a recursion for a fast implementation.)

A similar line of reasoning can also be followed for the  $p_k$  polynomials: from

$$1 = \|\mathbf{u}_{k+1}\|^2 = \|p_k(AA^T)A\mathbf{v}_1\|^2 = \sum_{j=1}^n \gamma_j^2 \sigma_j^2 p_k(\sigma_j^2)^2$$

it follows that (using Lemma 6)

$$1 \geq |\gamma_1| \sigma_1 p_k(\sigma_1^2).$$

Again, if  $\gamma_1$  would be known, the largest zero of

$$f_2(t) = t p_k(t^2) - 1/|\gamma_1|$$

would yield an upper bound  $\sigma_{\text{up}}$  for  $\sigma_{\text{max}}$ ; where we replace the unknown  $\gamma_1$  by  $\delta$ . Hence we have proved the following theorem.

**Theorem 9** *Assume that we have carried out  $k$  steps of Lanczos bidiagonalization with starting vector  $\mathbf{v}_1$  which has been chosen randomly with respect to the uniform distribution over  $S^{n-1}$ , and let  $\varepsilon \in (0, 1)$ . Then the largest zero of the polynomials*

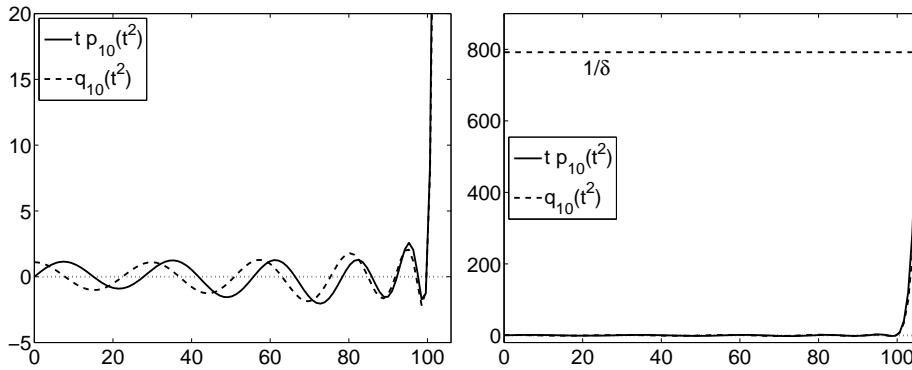
$$f_1(t) = q_k(t^2) - 1/\delta \tag{9}$$

$$f_2(t) = t p_k(t^2) - 1/\delta \tag{10}$$

*with  $\delta$  given by (8), are upper bounds for  $\|A\|$  with probability at least  $1 - \varepsilon$ .*

In Figure 1 we give an idea of the behavior of the polynomials  $p$  and  $q$ . For  $A = \text{diag}(1 : 100)$ , we carry out 10 steps of Lanczos bidiagonalization with a random starting vector.

We take  $\varepsilon = 0.01$ , then it follows from (8) that  $1/\delta \approx 792$ . The largest singular value of  $B_{10}$  is  $\theta_1 \approx 99.83$ , while that of  $\hat{B}_{10}$  is  $\hat{\theta}_1 \approx 99.86$ . Determining the  $t > \theta_1$  for which  $q_{10}(t^2) = 1/\delta$  gives the probabilistic bound  $\sigma_{\text{up}} \approx 105.87$  which is correct with probability at least 99%. Likewise,  $t p_{10}(t^2) = 1/\delta$  yields  $\sigma_{\text{up}} \approx 105.35$ . We refer to Section 5 for many more numerical experiments.



**Fig. 1** The Lanczos polynomials  $q_{10}(t^2)$  and  $t p_{10}(t^2)$  after 10 steps of Lanczos bidiagonalization, with  $\varepsilon = 0.01$ . Their largest zeros determine guaranteed lower bounds for  $\|A\|$ . The intersection points with the line  $1/\delta$  determine upper bounds for  $\|A\|$  with probability at least 99%. The only difference between the two figures is the scale on the vertical axis.

#### 4 Ritz polynomials

In Section 2 we have also introduced, in addition to the ‘‘Lanczos’’ polynomials  $p_k$  and  $q_k$ , the ‘‘Ritz’’ polynomials  $r_j = r_j^{(k)}$  and  $s_j = s_j^{(k)}$ , for  $j = 1, \dots, k$ ; see (5) and (6). These polynomials are associated with the approximate right and left singular vectors  $V_k \mathbf{d}_j$  and  $U_k \widehat{\mathbf{c}}_j$ , which are sometimes called Ritz vectors in the context of eigenvalue problems; we will use the same terminology in this situation. We will now exploit the polynomials  $r_1$  and  $s_1$  corresponding to the largest approximate singular vectors (that is, the approximate left and right singular vectors corresponding to the largest approximate singular values  $\widehat{\theta}_1^{(k)}$  and  $\theta_1^{(k)}$ , respectively). The following result is similar to Proposition 4.

**Proposition 10** *The polynomials  $r_1$  and  $s_1$  have positive leading coefficients and increase strictly monotonically to the right of their largest zeros  $\widehat{\theta}_2^2$  and  $\theta_2^2$ , respectively.*

*Proof* This follows from Corollary 3 and the fact that  $r_1$  and  $s_1$  are polynomials of degree  $k - 1$ .

Recall from (5) that  $V_k \mathbf{d}_1 = s_1(A^T A) \mathbf{v}_1$  is the approximation to the right singular vector corresponding to the largest singular value  $\theta_1$  of  $B_k$ , which is an approximation (more precisely, a lower bound) for  $\|A\|$ . Since

$$\theta_1^2 = \|AV_k \mathbf{d}_1\|^2 = \sum_{j=1}^n \gamma_j^2 \sigma_j^2 s_1(\sigma_j^2)^2$$

we derive

$$\theta_1 \geq |\gamma_1| \sigma_1 s_1(\sigma_1^2).$$

Analogously, since from (6) we have  $U_k \widehat{\mathbf{c}}_1 = r_1(AA^T) A \mathbf{v}_1$ , we get

$$\widehat{\theta}_1^2 = \|A^T U_k \widehat{\mathbf{c}}_1\|^2 = \sum_{j=1}^n \gamma_j^2 \sigma_j^4 r_1(\sigma_j^2)^2$$



so

$$\widehat{\theta}_1 \geq |\gamma_1| \sigma_1^2 r_1(\sigma_1^2).$$

The next result follows in a similar way as Theorem 9.

**Theorem 11** *Assume that the starting vector  $\mathbf{v}_1$  has been chosen randomly with respect to the uniform distribution over  $S^{n-1}$  and let  $\varepsilon \in (0, 1)$ . Then the largest zero of the polynomials*

$$f_3(t) = t s_1(t^2) - \theta_1/\delta \tag{11}$$

$$f_4(t) = t^2 r_1(t^2) - \widehat{\theta}_1/\delta \tag{12}$$

with  $\delta$  given by (8), are upper bounds for  $\|A\|$  with probability at least  $1 - \varepsilon$ .

**Remark** In [13], Chebyshev polynomials of the first kind were also studied. These polynomials on a given interval have the property that their absolute value is at most 1 on this interval and that they tend to sharply increase outside this interval. Nevertheless, experience in [13] shows that the Lanczos and Ritz polynomials, which are implicitly generated and “adapted” to the problem at hand, naturally tend to give better probabilistic bounds than “fixed” Chebyshev polynomials that only use partial information, such as the approximations  $\theta_1$  and  $\theta_k$  to the largest, respectively smallest singular value. Therefore, we do not study this type of polynomial in this paper.

## 5 Numerical experiments

First, we give a pseudocode for Lanczos bidiagonalization with reorthogonalization and the computation of the probabilistic bounds.

---

**Algorithm:** Lanczos bidiagonalization method with probabilistic upper bounds.

**Input:** Matrix  $A$ , random starting vector  $\mathbf{v}_1$ , probability level  $\varepsilon$ , Krylov dimension  $k$ .

**Output:** A lower bound approximation  $\widehat{\theta}_1$  to  $\|A\|$  and a probabilistic upper bound  $\sigma_{\text{up}}$ , where  $\|A\| \leq \sigma_{\text{up}}$  holds with probability at least  $1 - \varepsilon$ .

---

```

1: Determine  $\delta$  from  $n$  and  $\varepsilon$ , see (8)
2: for  $j = 1, \dots, k$ 
3:    $\mathbf{u} = A\mathbf{v}_j$ 
4:   if  $j > 1$ 
5:      $\mathbf{u} = \mathbf{u} - \beta_{j-1}\mathbf{u}_{j-1}$ 
6:      $\mathbf{u} = \mathbf{u} - U_{j-1}(\mathbf{u}^T U_{j-1})^T$ 
7:   end
8:    $\alpha_j = \|\mathbf{u}\|$ 
9:    $\mathbf{u}_j = \mathbf{u} / \alpha_j$ 
10:   $\mathbf{v} = A^T \mathbf{u}$ 
11:   $\mathbf{v} = \mathbf{v} - \alpha_j \mathbf{v}_j$ 
12:   $\mathbf{v} = \mathbf{v} - V_j(\mathbf{v}^T V_j)^T$ 
13:   $\beta_j = \|\mathbf{v}\|$ 
14:   $\mathbf{v}_{j+1} = \mathbf{v} / \beta_j$ 
15: end
16: Determine largest singular value  $\hat{\theta}_1$  of  $\hat{B}_k$ 
17: Determine probabilistic upper bound  $\sigma_{\text{up}}$  for  $\|A\|$  with probability  $\geq 1 - \varepsilon$ 
    using  $f_2$  (see (10))

```

---

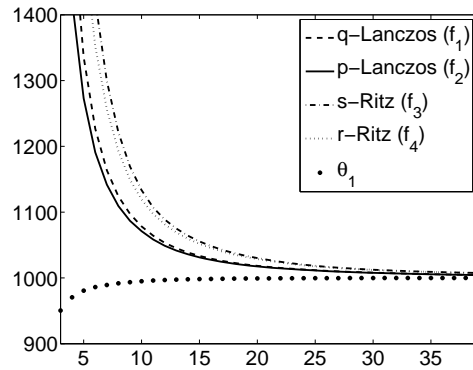
A few remarks about the algorithm: lines 6 and 12 implement reorthogonalization in a computationally efficient way. (Although reorthogonalization turned out to be unnecessary in the experiments, we still recommend it to ensure stability.) The probabilistic bounds may be computed in line 16, but also if desired after lines 8 or 13. We propose to use polynomial  $f_2$  (see (10) and below for the motivation). As explained earlier in the paper, breakdowns as well as restarts are not included.

**Experiment 1** To get an idea of the behavior of the probabilistic bounds, we first take  $n = 1000$ ,  $A = \text{diag}(1:1000)$ ,  $\varepsilon = 0.01$ , and a random  $\mathbf{v}_1$  on  $S^{n-1}$  as explained before Lemma 7; see Figure 2. Indicated are as a function of the iteration number  $k$ :

- the largest singular values  $\hat{\theta}_1^{(k)}$  of the bidiagonal  $k \times (k+1)$  matrices  $\hat{B}_k$ , which are guaranteed lower bounds for  $\|A\|$  (dots);
- the probabilistic upper bounds based on the polynomials  $f_1$  using the Lanczos polynomials  $q_k$  (see (9), dashed);
- the probabilistic upper bounds based on the polynomials  $f_2$  using the Lanczos polynomials  $p_k$  (see (10), solid);
- the probabilistic upper bounds based on the polynomials  $f_3$  using the Ritz polynomials  $s_1^{(k)}$  (see (5) and (11), dash-dotted); and
- the probabilistic upper bounds based on the polynomials  $f_4$  using the Ritz polynomials  $r_1^{(k)}$  (see (6) and (12), dotted).

As may be seen and expected, the Lanczos polynomials  $p_k$  and  $q_k$  (degree  $k$ , largest zero  $\hat{\theta}_1$  and  $\theta_1$ , respectively) yield better bounds than the Ritz polynomials  $r_1$  and  $s_1$  (degree  $k-1$ , largest zero  $\hat{\theta}_2$  and  $\theta_2$ , respectively; recall that  $\hat{\theta}_1 \geq \hat{\theta}_2$  and  $\theta_1 \geq \theta_2$ ). Comparing the two Lanczos polynomials,  $f_2$  with degree  $2k+1$  gives better results than the polynomial  $f_1$  with degree  $2k$ ; note also that the largest zero  $\hat{\theta}_1$  of  $p_k$  is not smaller than the largest zero  $\theta_1$  of  $q_k$ .

We see that for rather modest  $k$  we already obtain reasonably sharp guaranteed lower bounds and probabilistic upper bounds. Based on this experience, we will



**Fig. 2** Ritz values  $\hat{\theta}_1$  and probabilistic upper bounds ( $\varepsilon = 0.01$ ) for the matrix norm of  $A = \text{diag}(1:1000)$ .

only consider the lower bounds  $\hat{\theta}_1^{(k)}$  in the following experiments (note that  $\hat{\theta}_1^{(k)} \geq \theta_1^{(k)}$ ), and the probabilistic upper bounds derived from the polynomials  $f_2$ , based on the Lanczos polynomials  $p_k$ , as these tend to be sharper than those obtained with the other polynomials.

**Experiment 2** We now experiment with some common SVD test matrices of relatively small size, available either from the MatrixMarket [10] or from SVDPACK<sup>1</sup>, to be able to compare with the exact  $\|A\|$ . In Table 1 we compare the performances of Matlab’s `normest`, a power method on  $A^T A$  (third column), and Lanczos bidiagonalization (fourth column), where we allow 20 iterations in both cases, that is, 20 matrix-vector products (MVs) with  $A$  and 20 MVs with  $A^T$ . As expected, Lanczos bidiagonalization always gives better results and sometimes much better results. The reason for this is that the estimation of `normest` is based on  $\|A^T \mathbf{w}\|/\|\mathbf{w}\|$ , where  $\mathbf{w} = (A^T A)^{19} \mathbf{Av}_1$ , while Lanczos bidiagonalization maximizes the same norm over all vectors  $\mathbf{w}$  in the Krylov space

$$\mathcal{K}_{20}(A^T A, \mathbf{Av}_1) := \text{span}(\mathbf{Av}_1, (A^T A)\mathbf{Av}_1, \dots, (A^T A)^{19}\mathbf{Av}_1)$$

In addition, we give the error  $\sigma_{\text{up}} - \|A\|$ , where we have computed probabilistic upper bounds  $\sigma_{\text{up}}$  for  $\|A\|$  using  $f_2$  (see (10)) with  $\varepsilon = 0.01$ , i.e., which are correct with probability at least 99%. We see from Table 1 that the overestimation of the probabilistic upper bounds is always smaller than the underestimation of `normest`; sometimes even much smaller.

**Experiment 3** Next, we consider the  $11390 \times 1265$  term-by-document matrix `hypatia`<sup>2</sup>, a term-by-document matrix with 109056 nonzeros. The computation of a few of the largest singular triplets is commonly asked for such a matrix. These determine a low-rank approximation of the matrix, and the angles between the search vectors and the columns of the computed low-rank approximation are used for informational retrieval; see [2] and references. After 10 steps of Lanczos bidiagonalization applied to this matrix we get  $\hat{\theta}_1 \approx 342.2469$  while the upper bound

<sup>1</sup> [www.netlib.org/svdpack/](http://www.netlib.org/svdpack/)

<sup>2</sup> Available via [www.math.uri.edu/~jbaglama/](http://www.math.uri.edu/~jbaglama/)

**Table 1** For several SVD test matrices: `normest`: the error  $\|A\| - \nu$ , where  $\nu$  is the approximation obtained with 20 steps of the power method on  $A^T A$  as implemented in Matlab's `normest`; `bidiag`: the error  $\|A\| - \hat{\theta}_1$ , where  $\hat{\theta}_1$  is the approximation acquired with 20 steps of Lanczos bidiagonalization; `bdprob`: the error  $\sigma_{\text{up}} - \|A\|$ , where the probabilistic upper bound  $\sigma_{\text{up}}$ , computed after 20 steps of Lanczos bidiagonalization, is a true upper bound for  $\|A\|$  with probability at least 99%.

Matrix	Size	normest	bidiag	bdprob
abb313	313 × 176	$1.20 \cdot 10^{-3}$	$8.88 \cdot 10^{-15}$	$8.25 \cdot 10^{-7}$
hor131	434 × 434	$7.07 \cdot 10^{-2}$	$1.11 \cdot 10^{-16}$	$1.15 \cdot 10^{-10}$
pores3	532 × 532	$1.67 \cdot 10^3$	$1.41 \cdot 10^{-3}$	$3.12 \cdot 10^2$
sherman1	1000 × 1000	$3.48 \cdot 10^{-2}$	$7.30 \cdot 10^{-8}$	$6.64 \cdot 10^{-3}$
illc1033	1033 × 320	$3.15 \cdot 10^{-2}$	$7.25 \cdot 10^{-8}$	$3.83 \cdot 10^{-3}$
well1033	1033 × 320	$4.42 \cdot 10^{-2}$	$4.80 \cdot 10^{-12}$	$4.48 \cdot 10^{-5}$
well1850	1850 × 712	$1.52 \cdot 10^{-3}$	$2.68 \cdot 10^{-10}$	$1.13 \cdot 10^{-3}$
amoco	1436 × 330	$1.67 \cdot 10^{-5}$	$3.20 \cdot 10^{-14}$	$2.49 \cdot 10^{-14}$
apple1	3206 × 44	$4.77 \cdot 10^{-7}$	$3.33 \cdot 10^{-15}$	$1.29 \cdot 10^{-12}$
apple2	1472 × 294	$5.59 \cdot 10^{-7}$	$2.22 \cdot 10^{-15}$	$2.66 \cdot 10^{-15}$
jagmesh8	1141 × 1141	$1.00 \cdot 10^{-1}$	$2.06 \cdot 10^{-3}$	$5.62 \cdot 10^{-2}$
fidap004	1601 × 1601	$4.81 \cdot 10^{-3}$	$5.28 \cdot 10^{-8}$	$2.33 \cdot 10^{-3}$
illc1850	1850 × 712	$2.16 \cdot 10^{-2}$	$5.99 \cdot 10^{-7}$	$5.76 \cdot 10^{-3}$
west0479	479 × 479	$9.23 \cdot 10^2$	$5.82 \cdot 10^{-12}$	$1.16 \cdot 10^{-10}$
west2021	2021 × 2021	$7.23 \cdot 10^2$	$2.30 \cdot 10^1$	$1.75 \cdot 10^2$
diag(1:1000)	1000 × 1000	$1.10 \cdot 10^1$	$7.10 \cdot 10^{-1}$	$1.24 \cdot 10^1$
rand-1/2	1000 × 1000	$4.04 \cdot 10^{-1}$	$3.14 \cdot 10^{-2}$	$7.45 \cdot 10^{-2}$
randn	1000 × 1000	$1.39 \cdot 10^0$	$1.25 \cdot 10^{-2}$	$7.86 \cdot 10^{-1}$
rand(-1,0,1)	1000 × 1000	$4.02 \cdot 10^{-1}$	$1.21 \cdot 10^{-1}$	$6.69 \cdot 10^{-1}$
triu(randn)	1000 × 1000	$6.15 \cdot 10^{-1}$	$1.07 \cdot 10^{-1}$	$5.74 \cdot 10^{-1}$

with probability at least 99% is  $\hat{\theta}_1 + 2.43 \cdot 10^{-5}$ , leaving just a small interval for  $\|A\|$ . The upper bound with probability at least 99.9% is  $\hat{\theta}_1 + 2.43 \cdot 10^{-4}$ ; therefore, we may have confidence in the value of  $\hat{\theta}_1$ .

**Experiment 4** Finally, we take the  $23560 \times 23560$  matrix `af23560` [10], with 460598 nonzeros, arising in computational fluid dynamics. Ten steps of Lanczos bidiagonalization applied to this matrix yields  $\hat{\theta}_1 \approx 645.7$ . The probabilistic upper bound with  $\varepsilon = 0.01$  (probability at least 99%) is  $\sigma_{\text{up}} \approx 646.8$ , while  $\varepsilon = 0.001$  leads to  $\sigma_{\text{up}} \approx 652.0$ . We may therefore conclude that  $\|A\|$  is in the interval  $[\hat{\theta}_1, 1.01\hat{\theta}_1]$  with probability at least 99.9%. This small probabilistic interval (the lower bound  $\hat{\theta}_1$  as well as the probabilistic upper bound) is obtained in about 0.15 second on a laptop with processor speed about  $7 \cdot 10^9$  flops/sec. This clearly indicates the usefulness of the developed probabilistic bounds for large (sparse) matrices.

## 6 Discussion and conclusions

We have developed probabilistic upper bounds for the matrix norm. The bounds may be efficiently computed during or after the Lanczos bidiagonalization process. As we have seen from the experiments, Lanczos bidiagonalization with the probabilistic bounds may give very good results and may be superior to the power method on  $A^T A$ , as for instance implemented in Matlab's function `normest`, using

the same number of MVs.<sup>3</sup> We have proposed various functions  $f_1$ ,  $f_2$ ,  $f_3$ , and  $f_4$ ; for reasons described in Experiment 1 we advocate the use of  $f_2$  (see (10)).

Multiple runs of the method may also be combined to increase the reliability of the estimates. If  $\mathbf{v}_1$  and  $\hat{\mathbf{v}}_1$  are two independently chosen random initial vectors leading to probabilistic upper bounds  $\sigma_{\text{up}}^{(1)}$  and  $\sigma_{\text{up}}^{(2)}$  with probability at least  $1 - \varepsilon$ , then  $\max\{\sigma_{\text{up}}^{(1)}, \sigma_{\text{up}}^{(2)}\}$  is an upper bound with probability at least  $1 - \varepsilon^2$ .

As many other iterative subspace methods, the proposed method is matrix-free, which means that  $A$  need not be known explicitly, as long as  $A\mathbf{v}$  and  $A^T\mathbf{u}$  can be computed for arbitrary vectors  $\mathbf{v}$  and  $\mathbf{u}$  of appropriate sizes.

It would be very desirable to be able to develop probabilistic upper bounds for the condition number  $\kappa(A) = \|A\| \|A^{-1}\|$ . Unfortunately, the polynomials generated in Lanczos bidiagonalization are not useful for this, as they do not increase near the origin; in fact the polynomials are either even or odd. The Lanczos bidiagonalization process only provides guaranteed upper bounds ( $\hat{\theta}_k$  or  $\theta_k$ ) for  $\sigma_{\min}$ . Indeed, finding a lower bound for the smallest singular value is known to be difficult; see, e.g., [3] and references. (Note that the results in [4] are based on expensive matrix factorizations.) In the context of Lanczos bidiagonalization, the best available ‘‘probabilistic estimate’’ for  $\kappa(A)$  might be  $\sigma_{\text{up}}/\hat{\theta}_k$ , where  $\hat{\theta}_k$  is the smallest singular value of  $\hat{B}_k$  and  $\sigma_{\text{up}}$  is the probabilistic upper bound of  $f_2$ . However, we note that since the approximation  $\hat{\theta}_k \approx \sigma_{\min}$  might be arbitrarily poor, this is not a bound of any type. Indeed, experiments with the matrices of Table 1 sometimes gave disappointing results (such as underestimation by a factor 1000). Further progress in reliable and inexpensive estimation of the matrix condition number would be very welcome.

## References

1. J. BAGLAMA AND L. REICHEL, *Augmented implicitly restarted Lanczos bidiagonalization methods*, SIAM J. Sci. Comput., 27 (2005), pp. 19–42.
2. ———, *Restarted block Lanczos bidiagonalization methods*, Numer. Algorithms, 43 (2006), pp. 251–272 (2007).
3. C. H. BISCHOF, *Incremental condition estimation*, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 312–322.
4. C. H. BISCHOF, J. G. LEWIS, AND D. J. PIERCE, *Incremental condition estimation for sparse matrices*, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 644–659.
5. G. H. GOLUB AND W. KAHAN, *Calculating the singular values and pseudo-inverse of a matrix*, J. Soc. Indust. Appl. Math. Ser. B Numer. Anal., 2 (1965), pp. 205–224.
6. G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, The John Hopkins University Press, Baltimore, London, 3rd ed., 1996.
7. M. E. HOCHSTENBACH, *A Jacobi–Davidson type SVD method*, SIAM J. Sci. Comput., 23 (2001), pp. 606–628.
8. R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
9. J. KUCZYŃSKI AND H. WOŹNIAKOWSKI, *Estimating the largest eigenvalue by the power and Lanczos algorithms with a random start*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 1094–1122.
10. *The Matrix Market*. <http://math.nist.gov/MatrixMarket>, a repository for test matrices.
11. B. N. PARLETT, *The Symmetric Eigenvalue Problem*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998.

<sup>3</sup> We hereby would like to make a case for the replacement of `normest` in Matlab by a procedure based on Lanczos bidiagonalization.

- 
12. M. STOLL, *A Krylov–Schur approach to the truncated SVD*, *Linear Algebra Appl.*, 436 (2012), pp. 2795–2806.
  13. J. L. M. VAN DORSELAER, M. E. HOCHSTENBACH, AND H. A. VAN DER VORST, *Computing probabilistic bounds for extreme eigenvalues of symmetric matrices with the Lanczos method*, *SIAM J. Matrix Anal. Appl.*, 22 (2000), pp. 837–852.