An Implicitly Restarted Lanczos Bidiagonalization Method for Computing Smallest Singular Triplets

E. Kokiopoulou, C. Bekas*†*, and E. Gallopoulos*‡*

1 Introduction

We describe the development of a method for the efficient computation of the smallest singular values and corresponding vectors for large sparse matrices [4]. The method combines state-of-the-art techniques that make it a useful computational tool appropriate for large scale computations. The method relies upon Lanczos bidiagonalization (LBD) with partial reorthogonalization [6], enhanced with implicit restarts and harmonic Ritz values. We note that although LBD has been successfully used for the approximation of largest singular values [5], our target in this paper is the computation of the smallest singular values. Thus, in order to design a matrix free method by avoiding shift-and-invert techniques we rely on harmonic Ritz values.

Using LBD for the approximation of the smallest singular values often causes the lengths of the Lanczos bases to become quite large in order to obtain accurate approximations. For that reason, we embed an implicit restarting mechanism in LBD [12], which maintains memory requirements constant at each restart. In order to avoid the explicit inversion of A, we employ a harmonic Ritz value shift strategy [10, 9]. Harmonic Ritz values and vectors have been reported (see e.g. [8]) to be

*Computer Engineering & Informatics Department, University of Patras, Rio 26500, Greece, e-mail: exk@hpclab.ceid.upatras.gr

†—, e-mail: knb@hpclab.ceid.upatras.gr

‡—, e-mail: stratis@hpclab.ceid.upatras.gr

particularly useful for restarting purposes where a certain amount of important information must be included in the subspace for the next restart. Implicit restarting addresses elegantly this problem.

We address the problem of detecting clustered singular values by adopting the ODT deflation scheme [1] [7], which is adapted accordingly so that it can be applied directly on the LBD factorization. Convergence is detected by monitoring the refined residual [3].

2 Description of the method

We next proceed with a brief description of our method.

2.1 Lanczos bidiagonalization

Consider the matrix $A \in \mathbb{C}^{m\times n}$. Then LBD computes the factorizations

$$AV_k = U_{k+1}B_k, \quad \text{or}$$
$$A^*U_{k+1} = V_kB_k^* + \alpha_{k+1}v_{k+1}e_{k+1}^*$$

where the bases $U_{k+1} \in \mathbb{C}^{m\times(k+1)}$ and $V_k \in \mathbb{C}^{n\times k}$ have orthonormal columns and the matrix $B_k \in \mathbb{R}^{(k+1)\times k}$ is lower bidiagonal. Our implementation relies upon the `lanbpro` routine of PROPACK [5], which makes use of a partial reorthogonalization scheme that maintains an acceptable level of orthogonality among Lanczos vectors at low cost.

2.2 Implicitly restarted LBD

We employ the implicit restarting mechanism on LBD via two approaches. Denote $l = k + p$ the maximum dimension of the LBD factorization. The first approach is accomplished by applying the implicitly shifted QR algorithm on matrix $B_lB_l^*$. However, this is unwise from numerical standpoint since the condition number of B_l is squared. Furthermore, we have to recover matrix B_k from $B_kB_k^*$ in every restart, before extending the factorization to size l.

The second approach suggests using Golub-Kahan SVD steps directly on B_l which is widely known as the “bulgechasing process” [2]. This process updates the LBD factorization to $AU_{k+1}^+ = V_k^+B_k^+$, which is what we would have obtained after k steps of LBD with the special starting vector $u_1^+ = (AA^* - \mu^2 I)u_1$ using shift μ. Repeating the above process for p shifts we obtain a bidiagonalization that corresponds to a starting vector which is a polynomial of AA^* on u_1 with zeros at shifts $\mu_i, i = 1, \ldots, p$. This technique of implicit application of polynomial filtering on the starting vector of the method is particularly efficient for approximating a few of the eigenvalues of the matrix at hand.

2.3 Harmonic Ritz values

Harmonic Ritz values of A are equal to the reciprocal of the Ritz values of A^{-1} and have been proposed for the approximation of the smallest in magnitude eigenvalues of A [11], [9]. Since we have applied the implicit restarting mechanism on AA^* we compute the harmonic Ritz values produced by the oblique projection scheme,

$$AA^*\tilde{u}_{l+1} - \tilde{\theta}_{l+1}\tilde{u}_{l+1} \perp AA^*\mathcal{U}_{l+1},$$

where $\tilde{\theta}_j$ and $\tilde{u} \in \mathbb{C}^{m\times 1}$ is the corresponding harmonic Ritz value and vector respectively. It can be proven, after some algebraic manipulation, that the harmonic Ritz values are derived by the singular values of the $(l+2)\times(l+1)$ lower bidiagonal matrix B_{l+1}. We use the largest harmonic Ritz values as shifts in the implicit restarted LBD.

2.4 Refined residual

It has been observed that when a Ritz value has converged, the corresponding Ritz vector may have not converged [13, Sec. 4.3]. This pitfall can be addressed by substituting Ritz vector for refined Ritz vectors [3]. The refined Ritz vectors are designed to minimize the 2-norm of the residual with respect to the subspace involved. Concerning the LBD, the computation of the refined residual and singular vectors $\tilde{u} = U_{k+1}c,\ c \in \mathbb{C}^{(k+1)\times 1}$ and $\tilde{v} = V_k d,\ d \in \mathbb{C}^{k\times 1}$ results to the solution of the joint minimization problem

$$\min_{\tilde{u},\tilde{v}} \| \left[\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} - \tilde{\sigma} I_{m+n} \right] \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} \|_2$$

2.5 Deflation

When the smallest approximate singular triplet has converged we should deflate it in order to proceed to the computation of the next singular triplet. R. Lehoucq and D. Sorensen have proposed an efficient deflation scheme in [7], that is called orthogonal deflating transformation (ODT for short). It can be proven that this scheme can be applied directly on the LBD factorization in order to deflate the converged singular triplet.

Bibliography

[1] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst. *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*. SIAM, Philadelphia, 2000.

[2] Å. Björck, E. Grimme, and P. Van Dooren. An implicit bidiagonalization algorithm for ill-posed systems. *BIT*, 34:510–534, 1994.

[3] Z. Jia. Refined Iterative algorithms based on Arnoldi's process for large unsymmetric eigenproblems. *Linear Algebra Appl.*, 259:1–23, 1997.

[4] E. Kokiopoulou, C. Bekas, and E. Gallopoulos. Computing smallest singular triplets with implicitly restarted Lanczos bidiagonalization. Technical report, Computer Eng. & Informatics Dept., HPCLAB-TR-17-02-03, 2003. Submitted for publication.

[5] R. M. Larsen. PROPACK: A software package for the symmetric eigenvalue problem and singular value problems on Lanczos and Lanczos bidiagonalization with partial reorthogonalization. See URL http://soi.stanford.edu/ rmunk/PROPACK/.

[6] R. M. Larsen. *Lanczos bidiagonalization with partial reorthogonalization*. PhD thesis, Dept. Computer Science, University of Aarhus, DK-8000 Aarhus C, Denmark, Oct. 1998.

[7] R.B. Lehoucq and D. C. Sorensen. Deflation techniques for an implicitly restarted Arnoldi iteration. *SIAM J. Matrix Anal. Appl.*, 17:789–821, 1996.

[8] R. B. Morgan. Implicitly Restarted `GMRES` and Arnoldi methods for nonsymmetric systems of equations. *SIAM J. Matrix Analy. Appl.*, 21(4):1112–1135, March 2000.

[9] R.B. Morgan. Computing Interior Eigenvalues of Large Matrices. *Linear Algebra Appl.*, 154/156:289–309, 1991.

[10] C.C. Paige, B.N. Parlett, and H.A. van der Vorst. Approximate solutions and eigenvalue bounds from Krylov subspaces. *Num. Lin. Algebra with Appl.*, 2(2):115–134, 1995.

[11] C.C. Paige, B.N. Parlett, and H.A. van der Vorst. Approximate solutions and eigenvalue bounds from Krylov subspaces. *Num. Lin. Algebra with Appl.*, 2(2):115–134, 1995.

[12] D.C. Sorensen. Implicitly restarted Arnoldi/Lanczos methods for large scale eigenvalue problems. In D.E. Keyes, A. Sameh, and V. Venkatakrishnan, editors, *Parallel Numerical Algorithms*, pages 119–166. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.

[13] G. W. Stewart. *Matrix Algorithms. Volume II: Eigensystems.* SIAM, Philadelphia, 2001.