

Partial spectrum of large hermitean matrices

Chris Johnson and A. D. Kennedy*

The University of Edinburgh, UK

E-mail: chrisj@epcc.ed.ac.uk, adk@ph.ed.ac.uk

We present a new variant of a Krylov space algorithm for finding the eigenpairs of a large hermitean matrix where the eigenvalues lie in a specified low density part of the spectrum. The method uses selective re-orthogonalization and re-starting to find each eigenpair once. We present theoretical bounds on the convergence rate, and show that these work well in practice for the hermitean Wilson Dirac operator. We have implemented the method in Chroma, and we show that it is significantly faster than the Ritz method currently available.

The XXIX International Symposium on Lattice Field Theory, Lattice2011

July 11-16, 2011

Lake Tahoe, U.S.A.

*Speaker.

Contents

1. Introduction	2
2. The Algorithm	2
3. Theoretical Bounds	3

1. Introduction

We have developed a new algorithm based on the LANSO method of Parlett and Scott [1] for finding eigenpairs of a Hermitean matrix A near large voids in its spectrum, and we have implemented it in Chroma [2]. Such an eigensolver is useful for low-mode averaging, accelerating the application of rational approximations to the overlap operator, and for many other applications. We also provide a theoretical analysis leading to bounds that explain the rapid convergence of eigenvalues near large voids in the spectrum, or more generally in regions of low spectral density, and show that these bounds are applicable in practice for the QCD Wilson–Dirac operator.

2. The Algorithm

For an $n \times n$ Hermitean matrix A an orthonormal basis for the Krylov subspace

$$\mathcal{K}_j(A, \mathbf{q}) = \text{span}(\mathbf{q}, A\mathbf{q}, A^2\mathbf{q}, A^3\mathbf{q}, \dots, A^{j-1}\mathbf{q})$$

may be built iteratively from a starting vector \mathbf{q} using Gram–Schmidt orthonormalization. In this basis A itself is Hessenberg, in fact as A is Hermitean it has a tridiagonal form T ; thus \mathcal{K} can be constructed using a three-term recurrence rather than orthogonalizing with respect to all previous Krylov vectors,

$$A\mathbf{q}_j = \beta_j\mathbf{q}_{j+1} + \alpha_j\mathbf{q}_j + \beta_{j-1}\mathbf{q}_{j-1},$$

where α and β are the diagonal and off-diagonal elements of T .

From time-to-time we partially diagonalize T using the QR method with suitably chosen shifts in order to find the Ritz pairs of A . The Ritz pairs are the eigenpairs of the tridiagonal matrix T restricted to the Krylov subspace. The MaxMin theorem shows that extremal eigenvalues of A are approximated by the extremal Ritz values.

A well-known problem is that although the implicit orthogonality saves us a lot of work it is soon spoiled by rounding errors leading to spurious copies of eigenvalues. Specifically, eigenvalues which have converged reappear as they are not being explicitly suppressed by reorthogonalization. We therefore orthogonalize with respect to those converged eigenvectors that are within the part of the spectrum that we are interested in.

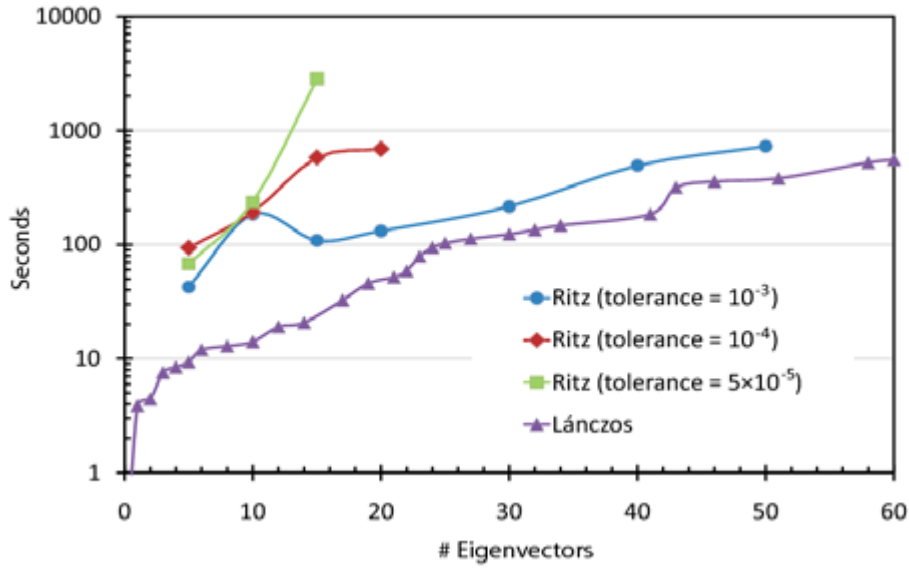


Figure 1: Timing comparisons of the Ritz method and our Lánczos method for the Wilson–Dirac operator on a typical $24^3 \times 48$ lattice. A range of different tolerances for the residual norm $\|(A - \theta)\mathbf{y}\|$ are presented for the Ritz method. In the Lánczos case the accuracy of all the eigenpairs increases with the number of iterations, and is always at least as good as the best Ritz values for the same number of iterations.

We use the following three-term recurrence relation

$$\mathbf{y} \cdot \mathbf{q}_{n+1} \leq \tau_{n+1} = \frac{\max_{\theta \in \Sigma} |\theta - \alpha_n| \tau_n + |\beta_{n-1}| \tau_{n-1}}{|\beta_n|}$$

where the Ritz vector \mathbf{y} belongs to a Ritz value θ lying in that part of the spectrum Σ in which we are interested. This indicates when we need to reorthogonalize: in practice such reorthogonalization is the most costly part of the calculation.

This idea is based on the LANSO algorithm [1] but as we are only interested in *part* of the spectrum we only maintain orthogonality with respect to eigenvectors belonging to that part.

We have implemented our algorithm in Chroma [2], and the results of our comparison of it with the Chroma implementation of the Ritz method [3] which requires explicit Gram–Schmidt orthogonalization are shown in Figure 1.

The other well-known problem with the Lánczos algorithm is that in exact arithmetic we only get one eigenvector from each degenerate eigenspace. However, with finite precision all eigenvectors will eventually grow out of rounding errors. Nevertheless it is often advantageous to restart with a starting vector orthogonal to all good eigenvectors as this limits the size of the Krylov space.

3. Theoretical Bounds

We are interested in eigenvalues near to a large void in the spectrum. Such eigenvalues are the largest eigenvalues of a shifted-and-squared matrix $A' = -(A - \Lambda)^2$ whose Krylov subspace \mathcal{K}' is contained in that for the original matrix A with the same number of applications of A . We have

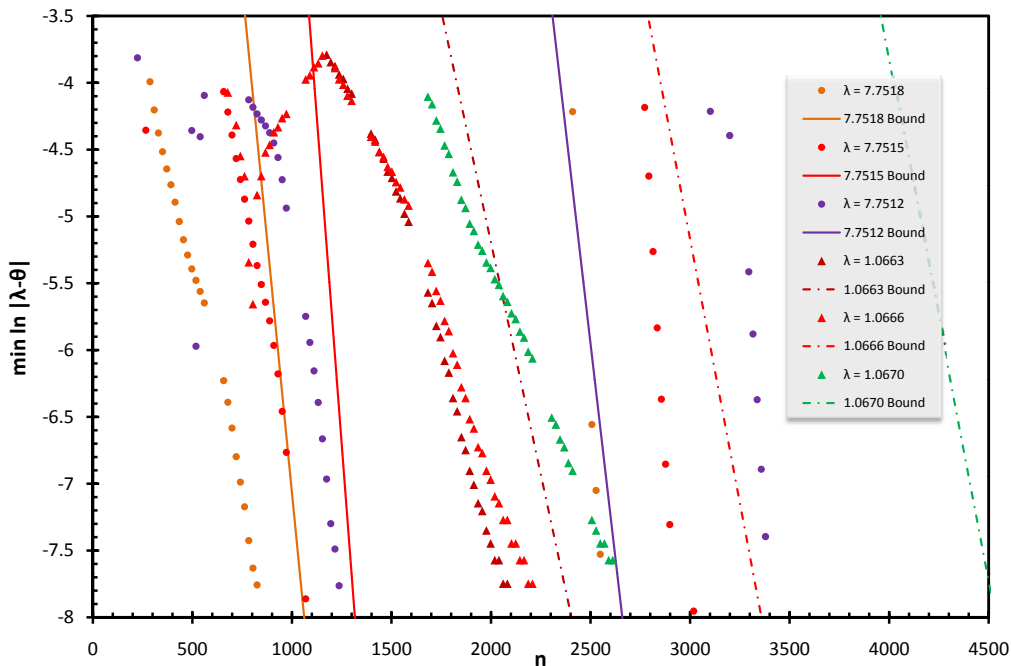


Figure 2: Convergence of the Ritz values θ to the eigenvalues λ for the Wilson–Dirac operator on a typical $24^3 \times 48$ lattice. The lines show our generalisation of the Kaniel–Paige–Saad bounds. Note the appearance of orthogonal nearly degenerate eigenvectors from rounding errors.

extended the Kaniel–Paige–Saad [4, 5] analysis for near-extremal eigenvalues to the case of Ritz pairs in low density regions by showing that the bounds on the convergence of the near-extremal eigenpairs of A' in \mathcal{H}' lead to useful bounds on their convergence in \mathcal{H} . Indeed, the bounds corresponding to the best possible choices for the shift Λ hold without our having to know the values of Λ explicitly. The details of this will be given elsewhere [6]. Figure 2 shows that these bounds can be fairly tight.

Acknowledgments

We gratefully acknowledge the support of the Centre for Numerical Algorithms and Intelligent Software (EPSRC EP/G036136/1) and the Scottish Funding Council in the preparation of this work. We would also like to thank Balint J3o for his assistance with Chroma.

References

- [1] B. N. Parlett and D. S. Scott. The L3nczos algorithm with selective orthogonalization. *Mathematics of Computation*, 33(145):217–238, January 1979.
- [2] Robert G. Edwards and Balint J3o. The Chroma software system for lattice QCD. *Nucl. Phys. Proc. Suppl.*, 140:832, 2005.
- [3] B. Bunk. Computing the lowest eigenvalues of the fermion matrix by subspace iterations. *Nucl. Phys. Proc. Suppl.*, 53:987–989, 1997.

- [4] S. Kaniel. Estimates for some computational techniques in linear algebra. *Math. Comp.*, pages 369–378, 1966.
- [5] Y. Saad. On the rates of convergence of the Lánczos and the block-Lánczos methods. *SIAM J. Numer. Anal.*, 17(5):687–706, 1980.
- [6] Chris Johnson and A.D. Kennedy. Bounds on the convergence of Ritz values from Krylov subspaces to interior eigenvalues of Hermitean matrices. 2011. arXiv:Math.NA/1110.3034.