Journal of Computational Physics 229 (2010) 7477–7487

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

A Rayleigh–Chebyshev procedure for finding the smallest eigenvalues and associated eigenvectors of large sparse Hermitian matrices

Christopher R. Anderson *

UCLA Mathematics Department, P.O. Box 951555, Los Angeles, CA 90095-1555, USA

ARTICLE INFO

Article history:
Received 1 February 2010
Received in revised form 20 May 2010
Accepted 16 June 2010
Available online 23 June 2010

Keywords:
Hermitian matrices
Eigenvalues
Eigenvectors
Inverse iteration
Chebyshev polynomials

ABSTRACT

A procedure is presented for finding a number of the smallest eigenvalues and their associated eigenvectors of large sparse Hermitian matrices. The procedure, a modification of an inverse subspace iteration procedure, uses adaptively determined Chebyshev polynomials to approximate the required application of the inverse operator on the subspace. The method is robust, converges with acceptable rapidity, and can easily handle operators with eigenvalues of multiplicity greater than one. Numerical results are shown that demonstrate the utility of the procedure.

1. Introduction

Inverse subspace iteration [14] is one of the commonly used procedures for finding the smallest eigenvalues and associated eigenvectors of large sparse Hermitian matrices. Inverse iteration is an attractive method because for many problems the computational cost is proportional to the cost of solving linear systems of equations times the number of desired eigenvalues and eigenvectors. General dense matrix procedures for finding eigenvalues and eigenvectors of $N \times N$ symmetric matrices have an $O(N^3)$ computational cost and an $O(N^2)$ storage requirement, so that even for modest sized sparse matrices, one expects inverse iteration to be much more efficient if the solution of the linear systems can be accomplished in less than $O(N^3)$ operations. Moreover, the storage requirements of the method are minimal, the procedure converges rapidly (especially when a good shift is chosen), and as the method is relatively easy to understand, non-experts in computational linear algebra are likely to create successful implementations. This latter fact is important, as libraries of reliable eigensystem routines may not be available for those exploring the use of new computational platforms. Also, for those who work on more established computational platforms, pre-existing routines for the core computational tasks—standard operations in computational linear algebra—are likely to be available so that efficient implementation can be accomplished rather quickly.

Unfortunately, there is a fundamental problem with inverse subspace iteration when it is applied to large sparse matrices. When shifts are chosen for optimal convergence rates, the procedure requires the solution of singular or nearly singular systems of equations. If one uses Gaussian elimination to solve these systems, then this singularity actually helps the computational process [14] (the errors introduced in the solution procedure are in the direction of the desired eigenvectors). However, when standard iterative methods are used to solve the requisite system of equations the convergence rates are

* Tel.: +1 310 825 1298; fax: +1 310 206 2679.
E-mail address: anderson@math.ucla.edu

0021-9991/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2010.06.030

typically reduced or the methods may even fail to converge. Additionally, the development of good preconditioners for these iterative methods is complicated by the singular or nearly singular nature of the systems. Starting with work by [18], and continuing with [5,17,20] this problem has been studied and procedures for avoiding the difficulties caused by the near singular behavior have been formulated. However, if one considers an inverse subspace iteration procedure where the requisite solutions of the linear systems are computed with Krylov subspace based iterative methods [10] such as Conjugate-Gradients (CG) or generalized minimum residual (GMRES), then one recognizes that the approximations to the eigenvectors are ultimately just polynomials in the operator applied to a particular collection of vectors. Essentially, when "good" shifts are used, then the implicit determination of these polynomials by the Krylov iterative method is breaking down. In this paper, the procedure that we present can be viewed as an implementation of an inverse subspace iteration procedure in which polynomial transformations are still used, but the manner in which the polynomials are being determined is different from that associated with a Krylov subspace iterative method. The net result is an efficient method that only requires matrix vector products, can readily take advantage of multi-processor machines, and has very good reliability properties. In particular, the method will always converge if the initial vectors have components in the desired eigenvectors and can handle without difficulty systems with eigenvalues of multiplicity greater than one.

In the next section we review inverse subspace iteration and present results demonstrating the typical problems that arise when one solves the requisite linear systems using a standard implementation of a Conjugate-Gradient method. In the third section we describe our procedure in general terms and then provide the details required for its implementation. In the last section we present computational results. In addition to results of a computation of the eigenvalues of a discretization of a 2D elliptic PDE we present results on the computation of the smallest eigenvalues and their associated eigenvectors of a large $(10^6 \times 10^6)$ matrix arising from a Full Configuration Interaction (FCI) treatment of a multi-particle Schroedinger equation. In this discussion we also describe the multi-processor implementation that was required to obtain the computational results.

The procedure presented in this paper can be viewed or understood in many different ways. For example, in addition to seeing it as a variant of inverse subspace iteration, it can also be viewed as a variant of general subspace iteration with Chebyshev acceleration [14,19]. Our presentation was selected in an effort to motivate and describe the procedure in terms that would be more readily understood by anyone familiar with inverse iteration—a topic typically covered in undergraduate numerical analysis classes. The computation of the smallest eigenvalues of large sparse symmetric matrices is of considerable interest in quantum chemistry and quantum physics applications, and there are other types of methods that are commonly used. In particular, there are procedures [6,11,13,21,22,24] based on Davidson's method [9] and procedures [3,8,7,15,16,23] based on the Lanczos method. Of particular interest are the procedures described in [24], where adaptive Chebyshev filtering is used in conjunction with Davidson's method and [3] where Chebyshev filtering is used in conjunction with a Lanczos procedure. Direct comparisons have not been carried out between these other procedures and the one presented here, but we conjecture that the computational efficiency of the Rayleigh–Chebyshev procedure should be close to that of a block Lanczos procedure with re-orthogonalization.

In the following, when the task of finding the eigenvalues and eigenvectors of a matrix A is discussed, we will also refer to A as a linear operator. We do this to emphasize that the method presented here only requires the evaluation of Av for input vectors v—an explicit matrix representation of A is not required for the procedure.

2. Difficulties with inverse subspace iteration with shift

Inverse subspace iteration with a shift is a power method using the operator $(A-\sigma)^{-1}$ applied to a collection of vectors S_v that are orthogonalized at each step. The value σ is the "shift". Approximations to the eigenvalues and eigenvectors closest to σ are obtained from the projection of the operator on S_v, e.g. by a Rayleigh–Ritz procedure. The method in its simplest form is

Inverse subspace iteration with shift

Given an initial collection of M vectors, S_0,

(a) Compute $V_v = (A-\sigma I)^{-1} S_{v-1}$ by solving $(A-\sigma I)\ v_v^i = s_{v-1}^i$. (1)
(b) Orthonormalize $V_v = Q_v R_v$ by modified Gram–Schmidt.
(c) Form $H_v = Q_v^T A Q_v$.
(d) Diagonalize $H_v = G_v^T \Theta_v G_v$.
(e) Form $S_v = Q_v G_v$ the Rayleigh–Ritz approximations to the eigenvectors and test for the convergence of approximate eigenvalues $\theta_i^v,\ i = 1 \ldots M$. If convergence has not been obtained, repeat iteration starting with (a) and $v = v + 1$.

When the solution of the linear systems $(A-\sigma)$ required to apply $(A-\sigma)^{-1}$ can be successfully computed, then the diagonal elements $\theta_i^{(v)}$ of Θ_v converge to the eigenvalues of A closest to σ. If one seeks the smallest eigenvalues of A and desires that $(A-\sigma)$ be positive definite to facilitate the iterative solution of the requisite linear equations, then σ should be chosen to be less than the smallest eigenvalue of A. (For a more efficient implementation of the algorithm that avoids the cost associated with step (*c*) see [14].)

If λ_i are the eigenvalues of A, then for a fixed value of σ the standard convergence results [14], show that the approximate eigenvalues $\theta_i^{(v)}$ satisfy