Journal of Computational and Applied Mathematics 311 (2017) 513–521

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

On convergence of iterative projection methods for symmetric eigenvalue problems

Kensuke Aishima

Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan

ARTICLE INFO

Article history:
Received 23 June 2015
Received in revised form 28 December 2015

Keywords:
Iterative methods for eigenvalue problems
Global convergence
Rayleigh–Ritz procedure
Restarting
Preconditioning

ABSTRACT

We prove global convergence of particular iterative projection methods using the so-called shift-and-invert technique for solving symmetric generalized eigenvalue problems. In particular, we aim to provide a variant of the convergence theorem obtained by Crouzeix, Philippe, and Sadkane for the generalized Davidson method. Our result covers the Jacobi–Davidson and the rational Krylov methods with restarting and preconditioning that are important techniques for modern eigensolvers. More specifically, we prove that the Ritz pairs converge to exact eigenpairs, even though they are not necessarily the target eigenpairs. We would like to emphasize that our proof is not a routine consideration of Crouzeix, Philippe, and Sadkane. To complete the proof, we discover a key lemma, which leads to a very simple convergence proof, resulting in a new theorem similar to that of Crouzeix, Philippe, and Sadkane.

1. Introduction

In this study, we focus on convergence theory for iterative projection methods for finding smallest eigenvalues of the generalized eigenvalue problem

$$Ax = \lambda Bx, \quad A, B \in \mathbb{R}^{n\times n}, \tag{1}$$

where A is symmetric, and B is symmetric positive definite. Solving such eigenvalue problems is important in many scientific and engineering applications. For example, (1) arises from discretization of a self-adjoint operator for elliptic partial differential equations. Some discretization methods are associated with projection to a finite-dimensional subspace.

For large matrix eigenvalue problems, projection methods are also effective, as in the case of the discretization of infinite dimensional operators. Nearly all the effective projection methods for generalized symmetric eigenvalue problems are based on the Rayleigh–Ritz procedure for a subspace of the Euclidean space $\mathbb{R}^n$ [1–5]. The best-known method is the Lanczos method using a Krylov subspace [6]. In recent times, the rational Krylov method has also attracted much attention [7–12]. See [13–17] for recent developments regarding Krylov subspace methods. Furthermore, the Davidson [18] and Jacobi–Davidson methods [19] are familiar. Moreover, a generalized Davidson method exists [20–22], which can be viewed as a general framework that includes the Lanczos, Davidson, and Jacobi–Davidson methods. Furthermore, the steepest descent and conjugate gradient methods for minimizing the Rayleigh-quotient

$$\rho(x) := \frac{x^{\mathrm{T}}Ax}{x^{\mathrm{T}}Bx} \tag{2}$$

to obtain the eigenvalue often use the Rayleigh–Ritz procedure [23,24]. Moreover, some contour integral methods [25,26] can be viewed as belonging to the class of the Rayleigh–Ritz procedure.

E-mail address: kensuke_aishima@mist.i.u-tokyo.ac.jp.

http://dx.doi.org/10.1016/j.cam.2016.08.035
0377-0427/

This study discusses theoretical global convergence properties for such methods. In 1994, Crouzeix, Philippe, and Sadkane derived a global convergence theorem [27, Theorem 2.1] for the restarted generalized Davidson method that covers the restarted block Lanczos method, whereas Sorensen proved global convergence of the restarted Lanczos in 1992 [28, Theorem 5.9]. Furthermore, in 2015, Sorensen's result was extended to general situations [29, Theorems 3 and 4] to cover modern sophisticated restart strategies [16,17].

The steepest descent (PSD) and locally optimal conjugate gradient methods (LOPCG) with preconditioning for the Rayleigh-quotient are also effective with convergence rates that can be derived for an appropriate initial guess and preconditioning [23,30–32]. In recent times, it was shown in [33] that global and asymptotic convergence of the basic conjugate gradient (CG) method applies to nonlinear Hermitian eigenvalue problems. From another standpoint, PSD and LOPCG can be viewed as a sort of restarted generalized Davidson method. In 2003, Ovtchinnikov presented convergence estimates from this viewpoint in [22]. Although such an approach has derived sharper convergence estimates than that of the generalized Davidson method with suitable initial guesses for more than a decade [34,35], they do not show global convergence properties of the Jacobi–Davidson method for any initial guess, which cannot be covered by Crouzeix, Philippe, and Sadkane [27, Theorem 2.1]. Global convergence of the restarted Jacobi–Davidson method was proved in [29, Theorem 6] for the first time in 2015. In particular, [29, Theorem 6] shows that the smallest Ritz value converges to an exact eigenvalue, although not necessarily to the smallest one, in the same manner as [27, Theorem 2.1]. This global convergence property is also similar to the Rayleigh-quotient iteration [4, Theorem 4.9.1].

In connection with recent developments of the (inexact) Rayleigh-quotient iteration [36–42], the (inexact) rational Krylov method using the so-called shift-and-invert technique or the Cayley transform has recently been thoroughly investigated [7,8,10,11]. The shift-and-invert or Cayley transform techniques can be viewed as a sort of preconditioning, which has been also thoroughly investigated to accelerate general iterative projection methods [43–45].

With such a background of studies, we would like to construct a general framework concerning global convergence for iterative projection methods. In particular, we aim to extend the convergence proof for the Jacobi–Davidson [29, Theorem 6] to more general methods in the same manner as that of Crouzeix, Philippe, and Sadkane [27, Theorem 2.1]. Our result covers the rational Krylov method with a restart strategy and preconditioning. We would like to emphasize that our proof is not a routine consideration of [27, Theorem 2.1]. We need a key lemma (Lemma 1) to complete the convergence proof. This key lemma leads to a very simple convergence proof, though the convergence rate cannot be derived from it.

This paper is organized as follows. Section 2 is devoted to descriptions of the Rayleigh–Ritz procedure with restart strategy in an abstract form and the convergence proof by Crouzeix, Philippe, and Sadkane [27] for the generalized Davidson method for solving standard symmetric eigenvalue problems. This theorem is extended to generalized symmetric eigenvalue problems, and our goal is clarified in Section 3. In Section 4, we derive a new convergence theorem including the Jacobi–Davidson and rational Krylov methods with restarting and preconditioning in an abstract form.

Notation. Throughout this study, $A \in \mathbb{R}^{n\times n}$ is symmetric, $B \in \mathbb{R}^{n\times n}$ is symmetric positive definite, and the generalized eigenvalues for (A, B) are $\lambda_1 \le \cdots \le \lambda_n$. Furthermore, $X_i := [x_1, \ldots, x_i]$ the matrix whose jth column is the corresponding eigenvector x_j to λ_j for any $j = 1, \ldots, i$, normalized as $X_i^{\mathrm{T}} B X_i = I$, where I is the identity matrix. For any $V \in \mathbb{R}^{n\times i}$, let $\mathrm{span}\{V\}$ be the subspace spanned by the columns of V. Moreover, k is the number of desired smallest eigenvalues. For any vector $v \in \mathbb{R}^n$, let $\|v\|$ be $\sqrt{v^{\mathrm{T}} v}$ and $\|v\|_B$ be $\sqrt{v^{\mathrm{T}} B v}$.

2. Rayleigh–Ritz procedure and convergence theory [27]

We describe the Rayleigh–Ritz procedure with restart strategy in an abstract form as follows.

Algorithm 1 A framework of iterative projection methods with restarting for computing the k smallest eigenvalues of $Ax = \lambda Bx$.

Input: $A, B \in \mathbb{R}^{n\times n}$ and $V^{(0)} = [v_1^{(0)}, \ldots, v_{m_0}^{(0)}] \in \mathbb{R}^{n\times m_0}$

for $\ell := 0, 1, \ldots,$ **do**
    compute $A^{(\ell)} = V^{(\ell)\mathrm{T}} A V^{(\ell)}$, $B^{(\ell)} = V^{(\ell)\mathrm{T}} B V^{(\ell)}$
    compute the k smallest eigenvalues for $(A^{(\ell)}, B^{(\ell)})$: $\lambda_1^{(\ell)} \le \cdots \le \lambda_k^{(\ell)}$
    compute the corresponding Ritz vectors $x_1^{(\ell)}, \ldots, x_k^{(\ell)}$
    compute $V^{(\ell+1)} := [v_1^{(\ell+1)}, \ldots, v_{m_{\ell+1}}^{(\ell+1)}]$, where $\mathrm{span}\{V^{(\ell+1)}\} \ni x_i^{(\ell)}$ for $i = 1, \ldots, k$
end for

For example, the restarted Lanczos method corresponds to the situation where $V^{(\ell)}$ for any $\ell \in \mathbb{N}$ is a Krylov subspace. The relationship of Algorithm 1 and modern solvers to generalized symmetric eigenvalue problems is discussed in the next section. In general, most iterative projection methods are related to Algorithm 1, and are of one of two types, i.e., the residual based and the Rayleigh-quotient iteration methods as shown in Table 1 in the next section. In this section, we first discuss global convergence of the above algorithm for standard eigenvalue problems, for which $B = I$.