Journal of Computational Mathematics
Vol.36, No.5, 2018, 739–760.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1703-m2016-0580

CONVERGENCE ANALYSIS OF A LOCALLY ACCELERATED PRECONDITIONED STEEPEST DESCENT METHOD FOR HERMITIAN-DEFINITE GENERALIZED EIGENVALUE PROBLEMS*

Yunfeng Cai
LMAM & School of Mathematical Sciences, Peking University, Beijing 100871, China
Email: yfcai@math.pku.edu.cn
Zhaojun Bai
Department of Computer Science and Department of Mathematics, University of California, Davis, CA 95616, USA
Email : bai@cs.ucdavis.edu
John E. Pask
Physics Division, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA
Email: pask1@llnl.gov
N. Sukumar
Department of Civil and Environmental Engineering, University of California, Davis, CA 95616, USA
Email: nsukumar@ucdavis.edu

Abstract

By extending the classical analysis techniques due to Samokish, Faddeev and Faddeeva, and Longsine and McCormick among others, we prove the convergence of the preconditioned steepest descent with implicit deflation (PSD-id) method for solving Hermitian-definite generalized eigenvalue problems. Furthermore, we derive a nonasymptotic estimate of the rate of convergence of the PSD-id method. We show that with a proper choice of the shift, the indefinite shift-and-invert preconditioner is a locally accelerated preconditioner, and is asymptotically optimal which leads to superlinear convergence Numerical examples are presented to verify the theoretical results on the convergence behavior of the PSD-id method for solving ill-conditioned Hermitian-definite generalized eigenvalue problems arising from electronic structure calculations. While rigorous and full-scale convergence proofs of preconditioned block steepest descent methods in practical use still largely eludes us, we believe the theoretical results presented in this paper shed light on an improved understanding of the convergence behavior of these block methods.

Mathematics subject classification: 65F08, 65F15, 65Z05, 15A12.
Key words: Eigenvalue problem, Steepest descent method, Preconditioning, Superlinear convergence.

1. Introduction

We consider the Hermitian-definite generalized eigenvalue problem

$$Hu = \lambda Su, \tag{1.1}$$

where H and S are n-by-n Hermitian matrices and S is positive-definite. The scalar λ and nonzero vector u satisfying (1.1) are called *eigenvalue* and *eigenvector*, respectively. The pair

* Received March 21, 2016 / Revised version received November 28, 2016 / Accepted March 27, 2017 / Published online June 22, 2018 /

(λ, u) is called an *eigenpair*. All eigenvalues of (1.1) are known to be real. Our task is to compute few smallest eigenvalues and the corresponding eigenvectors. We are particularly interested in solving the eigenvalue problem (1.1), where the matrices H and S are large and sparse, and there is *no obvious* gap between the eigenvalues of interest and the rest. Furthermore, S is nearly singular and H and S share a near-nullspace. It is called an ill-conditioned generalized eigenvalue problem in [5], a term we will adopt in this paper. The ill-conditioned generalized eigenvalue problem is considered to be an extremely challenging problem.[1)]

Beside examples such as those cited in [5], the ill-conditioned eigenvalue problem (1.1) arises from the discretization of enriched Galerkin methods. The partition-of-unity finite element (PUFE) method [14], which falls within the class of enriched Galerkin methods, is a promising approach in quantum-mechanical materials calculations, see [3] and references therein. In the PUFE method, physics-based basis functions are added to the classical finite element (polynomial basis) approximation, which affords the method improved accuracy at reduced costs versus existing techniques. However, due to near linear-dependence between the polynomial and enriched basis functions, the system matrices that stem from such methods are ill-conditioned, and share a large common near-nullspace. Furthermore, there is in general no clear gap between the eigenvalues that will be sought and the rest. Another example of the ill-conditioned eigenvalue problem (1.1) arises from modeling protein dynamics using normal-mode analysis [2, 10, 11, 17].

In this paper, we focus on a preconditioned steepest descent with implicit deflation method, PSD-id method in short, to solve the eigenvalue problems (1.1). The basic idea of the PSD-id method is simple. Denote all the eigenpairs of (1.1) by (λ_1, u_1), (λ_2, u_2), $\ldots$, (λ_n, u_n), and the eigenvalue and eigenvector matrices by $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and $U = [u_1\ u_2\ \cdots\ u_n]$, respectively. Assume that the eigenvalues $\{\lambda_i\}$ are in an ascending order $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$. The following variational principles are well-known, see [27, p.99] for example:

$$\lambda_i = \min_{U_{i-1}^H S z = 0} \rho(z) \quad \text{and} \quad u_i = \operatorname*{argmin}_{U_{i-1}^H S z = 0} \rho(z), \tag{1.2}$$

where $U_{i-1} = [u_1\ u_2\ \cdots\ u_{i-1}]$ and $\rho(z)$ is the Rayleigh quotient

$$\rho(z) = \frac{z^H H z}{z^H S z}. \tag{1.3}$$

On assuming that U_{i-1} is known, one can find the ith eigenpair by minimizing the Rayleigh quotient $\rho(z)$ with z being S-orthogonal against U_{i-1} under the algorithmic framework of the preconditioned steepest descent minimization.

The idea of computing the algebraically largest eigenvalue and its corresponding eigenvector of (1.1) (with $B = I$) using the steepest descent (SD) method dates back to early 1950s [7] and [4, Chap.7]. In [13], block steepest descent (BSD) methods are proposed to compute several eigenpairs simultaneously. The preconditioned steepest descent (PSD) method was introduced around late 1950s [24, 25]. The block PSD (BPSD) methods have appeared in the literature, see [1, 16] and references therein. Like the PSD method, the PSD-id method studied in this paper computes one eigenpair at a time. To compute the ith eigenpair, the search subspace of PSD-id is *implicitly* orthogonalized against the previously computed $i-1$ eigenvectors. The preconditioner at each iteration of PSD-id is flexible (i.e., could change at every iteration) and can be indefinite, instead of being fixed and positive definite as in [1, 16, 25].

[1)] W. Kahan, Refining the general symmetric definite eigenproblem, poster presentation at Householder Symposium XVIII 2011, available http://www.cs.berkeley.edu/∼wkahan/HHXVIII.pdf