

Research Article

Behavior of the Correction Equations in the Jacobi–Davidson Method

Yuan Kong  and Yong Fang 

College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590, China

Correspondence should be addressed to Yong Fang; fangyong@sdust.edu.cn

Received 28 May 2019; Accepted 21 July 2019; Published 5 August 2019

Academic Editor: Francesca Vipiana

Copyright © 2019 Yuan Kong and Yong Fang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Jacobi–Davidson iteration method is efficient for computing several eigenpairs of Hermitian matrices. Although the involved correction equation in the Jacobi–Davidson method has many developed variants, the behaviors of them are not clear for us. In this paper, we aim to explore, theoretically, the convergence property of the Jacobi–Davidson method influenced by different types of correction equations. As a by-product, we derive the optimal expansion vector, which imposed a shift-and-invert transform on a vector located in the prescribed subspace, to expand the current subspace.

1. Introduction

The complex n -dimensional vector space and n -by- n matrix space are denoted by \mathbb{C}^n and $\mathbb{C}^{n \times n}$, respectively; and these notations can be straightforwardly carried over to real space. Many iterative methods for computing a few extreme eigenpairs of the standard eigenvalue problem

$$Ax = \lambda x, \text{ with } \|x\| = 1, \quad (1)$$

where $A \in \mathbb{C}^{n \times n}$ is a large, sparse Hermitian matrix, $\lambda \in \mathbb{R}$ is an eigenvalue, and $x \in \mathbb{C}^n$ is the associated eigenvector of matrix A , have been developed, such as the Krylov subspace methods [1, 2], the locally optimal block preconditioned conjugate gradient method [3], the Davidson method [4], and their variants [5–8]. Here, we use $\|\cdot\|$ to denote the Euclidean norm of a vector, which will be defined in the next section.

In the Jacobi–Davidson iteration method, the projection onto the orthogonal complement of the current approximation elegantly avoids the singularity of the shifted matrix and ensures the positive definite of the resulting coefficient matrix when the current approximation is near to the desired eigenvalue. In [9], the authors proposed several

alternative variants of the correction equations and presented some predictions reasonably on the convergence property influenced by these correction equations. Also, in [10], the authors explored the existence and uniqueness on solutions of different correction equations. However, to our best knowledge, there is no analysis to compare the convergence affected by different types of correction equations [11, 12]. In this paper, we aim to analyse, theoretically, the convergence property of the Jacobi–Davidson method influenced by different types of correction equations.

The organization of this paper is as follows. In Section 2, we firstly introduce some notations and present some preliminaries. In Section 3, we then give the main convergence results. And, finally, in Section 5, we end the paper with some concluding remarks.

2. Preliminaries

For convenience of statements, we firstly introduce some notations. In this paper, I_n is used to represent the $n \times n$ identity matrix, and, for short, it can be simplified as I . And, e_i with a suitable dimension is used to denote the i -th column of I . We use $\langle \cdot, \cdot \rangle$ to denote the inner product defined in the complex vector space \mathbb{C}^n , and the

corresponding induced Euclidean norm of either a vector or a matrix is denoted by $\|\cdot\|$. For two given vectors $v, w \in \mathbb{C}^n$, $\cos \phi = \langle v, w \rangle / \|v\| \|w\|$ indicates the angle between v and w , and, thus, the distance between them, with $\|v\| = 1$ and $\|w\| = 1$, can be expressed, equivalently, as $|\sin \angle(v, w)| = \|(I - v v^*)w\|$. In addition, for a given matrix $G \in \mathbb{C}^{n \times n}$, G^T and G^* denote its transpose and conjugate transpose, respectively, and this notation can be straightforwardly carried over to vectors.

The eigenpairs of the Hermitian eigenvalue problem are denoted by $\{(\lambda_i, x_i)\}_{i=1}^n$, with $x_i, i = 1, 2, \dots, n$, being orthonormal, (1). For a given matrix pair (A, B) and a nonzero vector $v \in \mathbb{C}^n$, the function,

$$R(A, B; v) = \frac{v^* A v}{v^* B v}, \quad (2)$$

is defined as the generalized Rayleigh quotient of v associated with the matrix pair (A, B) .

In [6], Sleijpen and Van der Vorst proposed the efficient Jacobi–Davidson iteration method for computing a few eigenpairs of the Hermitian eigenvalue problem (1). In each iteration step, say, the k -th, of the Jacobi–Davidson method, the so-called correction equation,

$$(I - u_1 u_1^*)(A - \theta I)(I - u_1 u_1^*)t = -r, \quad \text{for } t \perp u_1, \quad (3)$$

with $r = (A - \theta I)u_1$ being the residual with respect to the current unit approximation u_1 and its Rayleigh quotient $\theta = u_1^* A u_1$, requires to be solved to expand the current projection subspace.

It should be noticed that the projection onto the orthogonal complement of u_1 guarantees positive definite and well-conditioning of the coefficient matrix of the correction equation in (3). In addition, the Jacobi–Davidson iteration method attains, elegantly, the cubic convergence rate locally [6, 13–15].

In [9], Genseberger and Sleijpen proposed a more restrictive alternative correction equation:

$$(I - U U^*)(A - \theta I)(I - U U^*)t = -r, \quad \text{for } t \perp U, \quad (4)$$

where U is an n -by- m matrix of which the columns expand an orthonormal basis of the current projection subspace, say, subspace \mathcal{U} .

Although the projection in (4) is more complicated and expensive than that in (3), projecting onto the orthogonal complement of the current approximate subspace may further improve the conditioning of the correction equation, especially in the case of clustering eigenvalues. Obviously, if U is chosen as $U = [u_1]$, with u_1 being the current Ritz vector, the correction equation in (4) is reduced to that in (3).

Generally speaking, the solutions of the correction equations in (3) and (4) do not lead to the same expansion of the projection subspace, and, as emphasized in [9], it is unclear which of them is more effective in actual computations. Therefore, different convergence behaviors of the Jacobi–Davidson method are expected.

In the Jacobi–Davidson iteration method, once we obtain the current search subspace \mathcal{U} , we firstly project matrix

A onto it and obtain the corresponding projected matrix $T = U^* A U$; we then compute the eigenpairs $\{(\theta_i, s_i)\}_{i=1}^m$ of T with smaller size, that is, $T s_i = \theta_i s_i$ with $\|s_i\| = 1$; and, finally, use the Ritz pairs $\{(\theta_i, U s_i)\}_{i=1}^m$ to approximate the desired eigenpairs of matrix A , see [6] for more details. Note that the Ritz vectors $u_i = U s_i, i = 1, 2, \dots, m$, can also span subspace \mathcal{U} , and, thus, in the subsequent discussions, we assign $U = (u_1, u_2, \dots, u_m)$ in the correction equation in (4).

3. The Main Results

In this section, we exhibit the main result which depicts different convergence behaviors of the Jacobi–Davidson method affected by different types of correction equations (3) and (4).

If the correction equations (3) and (4) are solved exactly, we, frequently, obtain the solutions:

$$t_o = -u_1 + \frac{(A - \theta I)^{-1} u_1}{u_1^* (A - \theta I)^{-1} u_1}, \quad (5)$$

$$t_g = -u_1 + (A - \theta I)^{-1} U (U^* (A - \theta I)^{-1} U)^{-1} U^* u_1,$$

respectively. Thus, expanding the current search subspace \mathcal{U} by t_o and t_g in the above equations is equivalent to

$$\begin{aligned} t_o &= (A - \theta I)^{-1} U y_o, \\ t_g &= (A - \theta I)^{-1} U y_g, \end{aligned} \quad (6)$$

where $y_o = e_1$, $y_g = (U^* (A - \theta I)^{-1} U)^{-1} e_1$, and $e_1 = (1, 0, \dots, 0) \in \mathbb{R}^m$, as u_1 is included in the current search subspace.

Next, we first present a result which depicts the distance $|\sin \angle(x_1, t_y)|$ between $t_y = (A - \theta I)^{-1} U y$ and the desired eigenvector x_1 .

Theorem 1. Consider $U = (u_1, u_2, \dots, u_m)$ whose columns $u_i, i = 1, 2, \dots, m$, are Ritz vectors of A , and $\theta = u_1^* A u_1$, the Rayleigh quotient of u_1 . Assume that u_i admits the orthogonal direct sum decomposition:

$$u_i = c_{1i} x_1 + c_{2i} x_2 + \dots + c_{mi} x_m,$$

$$\text{with } \sum_{k=1}^m |c_{ki}|^2 = 1, \quad i = 1, 2, \dots, m, \quad (7)$$

and let $t_y = (A - \theta I)^{-1} U y$, then $y_e = C^* e_1$ solves the following minimum problem:

$$\min_{y \in \mathbb{C}^{m \times 1}} |\sin \angle(x_1, t_y)| = \min_{y \in \mathbb{C}^{m \times 1}} \frac{\|(I - x_1 x_1^*)(A - \theta I)^{-1} U y\|}{\|(A - \theta I)^{-1} U y\|}, \quad (8)$$

where $C = (c_{ij}) \in \mathbb{C}^{n \times m}$ with its elements being c_{ij} and $e_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^n$.

Proof. Based on the orthogonal direct sum decomposition (7), the matrix U can be represented as $U = X C$ with $X = (x_1, x_2, \dots, x_m)$, which indicates that

$$(A - \theta I)^{-1}U = (A - \theta I)^{-1}XC = X\Sigma^{-1}C, \quad (9)$$

where $\Sigma = \text{diag}(\lambda_1 - \theta, \lambda_2 - \theta, \dots, \lambda_n - \theta)$. Then, it follows from $(I - x_1 x_1^*)X = (0, x_2, \dots, x_n)$ that

$$\begin{aligned} & \|(I - x_1 x_1^*)(A - \theta I)^{-1}Uy\|^2 \\ &= y^* C^* \Sigma^{-1} \begin{pmatrix} 0 \\ x_2^* \\ \vdots \\ x_n^* \end{pmatrix} (0, x_2, \dots, x_n) \Sigma^{-1} C y \\ &= y^* C^* \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & (\lambda_2 - \theta)^{-2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\lambda_n - \theta)^{-2} \end{pmatrix} C y, \\ & \|(A - \theta I)^{-1}Uy\|^2 = y^* C^* \Sigma^{-2} C y. \end{aligned} \quad (10)$$

According to the property of the generalized Rayleigh quotient, we know that vector y such that $Cy = e_1$ minimizes

$$|\sin \angle(x_1, t_y)| = \frac{\|(I - x_1 x_1^*)(A - \theta I)^{-1}Uy\|^2}{\|(A - \theta I)^{-1}Uy\|^2}, \quad (11)$$

with the minimum being zero, and, thus, we obtain the least squares solution $y_e = C^* e_1$ uniquely.

In [16], the author considered the problem of optimal expansion of subspaces for eigenvector approximations. Specifically, starting from a prescribed subspace, the author derived the optimal resulting vector, obtained from multiplying matrix A with a vector located in the prescribed subspace, to expand the current prescribed subspace. Likewise, Theorem 1 provides the optimal vector, obtained by imposing a shift-and-invert transform $(A - \theta I)^{-1}$ on a vector located in the prescribed subspace, to expand the current subspace. \square

Theorem 2. Assume that the current search subspace spanned by U obtained by the Jacobi–Davidson method satisfies

$$|C_1^T y_g| > |c_{11}| \|y_g\|, \quad (12)$$

and then, we have the following estimate

$$|\cos \angle(y_g, y_e)| > |\cos \angle(y_o, y_e)|, \quad (13)$$

where y_e , y_o , and y_g are defined as above and $C_1^T = (c_{11}, c_{12}, \dots, c_{1m})$ is the first row of matrix C .

Proof. It follows from the orthogonal direct sum decomposition in (7) that

$$y_g = (U^* (A - \theta I)^{-1} U)^{-1} e_1 = (C^* \Sigma^{-1} C)^{-1} e_1. \quad (14)$$

As a result, under assumption (12) and by straightforward computations, we obtain

TABLE 1: Numerical results for Example 1.

Method	Index	m			
		2^5	2^6	2^7	2^8
JD	IT	85	94	166	175
	CPU	1.315	8.124	60.881	249.865
	ERR	$4.62E-11$	$1.36E-10$	$9.18E-10$	$9.31E-10$
JDV	IT	71	85	91	94
	CPU	0.936	6.349	26.837	120.923
	ERR	$1.18E-11$	$1.56E-10$	$6.11E-10$	$5.53E-09$

$$|\cos \angle(y_g, y_e)| > |\cos \angle(y_o, y_e)|, \quad (15)$$

which completes the proof. \square

Theorem 2 indicates that vector y_g approximates the least squares solution y_e better than that of y_o under the assumption in (12), which straightforwardly results in that the exact solution of (4) approximates the desired eigenvector x_1 better than that of the exact solution of (3).

4. Numerical Results

In this section, we use one simple example to examine the numerical behavior of the Jacobi–Davidson method affected by different type of correction equations. We simplify the Jacobi–Davidson method equipped with (3) as the JD method and that equipped with (4) as the JDV method.

We would like to compare the two methods JD and JDV in aspects of the number of iteration steps (“IT”) and the computing time in seconds (“CPU”). In addition, for the last approximation $\theta^{(k_o)}$, we use the absolute error (“ERR”) of the approximation to the desired eigenvalue, namely, $\text{ERR} = |\theta^{(k_o)} - \lambda_1|$, to check whether the two methods converge to a same target eigenvalue. The two methods are terminated once their residual norms could achieve the stopping criterion $(\|r^{(k)}\|/\|r^{(k)}\|) \leq 10^{-8}$, where $r^{(k)} = (A - \theta^{(k)} I)x^{(k)}$ is the residual, $x^{(k)}$ and $\theta^{(k)}$ are the approximate eigenvector, and the Rayleigh quotient of $x^{(k)}$, respectively.

Example 1. [1, 5] Let us consider the following partial differential equation:

$$-\frac{\partial}{\partial x} \left(\mu(x, y) \frac{\partial u}{\partial x} \right) - \frac{\partial}{\partial y} \left(\nu(x, y) \frac{\partial u}{\partial y} \right) = \lambda u, \quad (16)$$

defined on the domain of $[0, 1] \times [0, 1]$ imposed with the Dirichlet boundary conditions. We could acquire the eigenvalue problem (1) by discretization of the partial differential equation (16) by five-point finite difference scheme on an $m \times m$ grid. Here, $\mu(x, y)$ and $\nu(x, y)$ are chosen with $\mu(x, y) = \nu(x, y) = -e^{xy}/(1 + x + y)$, and the dimension of the resulting matrix is $n = m^2$.

We use Table 1 to list the numerical results of Example 1. From this table, we observe that the JDV method behaves more efficient and robust than the JD method, and that the JDV method outperforms the JD method in both terms of iteration steps and CPU time, although the two methods converge to the desired eigenpair globally.

5. Concluding Remarks

In this paper, we analyse, theoretically, the behaviors of the Jacobi–Davidson method affected by correction equations (3) and (4), respectively. Admittedly, the comparisons may be impractically useful for unknowns of the eigenvectors, and, thus, presenting their comparisons in actual computations may be a challenging and interesting problem.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

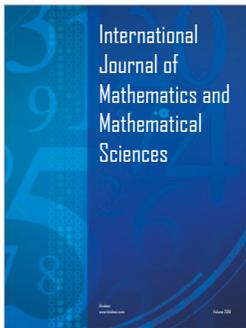
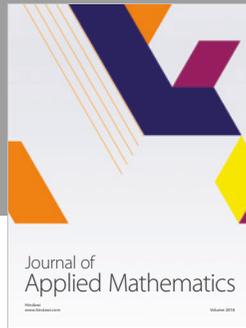
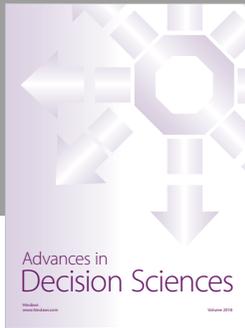
The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 61602188), Shandong Provincial Natural Science Foundation (Grant no. ZR2019QD018), Scientific Research Foundation of Shandong University of Science and Technology for Recruited Talents (Grant nos. 2017RCJJ068 and 2017RCJJ069).

References

- [1] C.-Q. Miao, “Filtered Krylov-like sequence method for symmetric eigenvalue problems,” *Numerical Algorithms*, pp. 1–17, 2018.
- [2] B. N. Parlett, *The Symmetric Eigenvalue Problem*, SIAM, Philadelphia, PA, USA, 1998.
- [3] A. V. Knyazev, “Toward the optimal preconditioned eigensolver: locally optimal block preconditioned conjugate gradient method,” *SIAM Journal on Scientific Computing*, vol. 23, no. 2, pp. 517–541, 2001.
- [4] E. R. Davidson, “The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices,” *Journal of Computational Physics*, vol. 17, no. 1, pp. 87–94, 1975.
- [5] C.-Q. Miao, “Computing eigenpairs in augmented Krylov subspace produced by Jacobi-Davidson correction equation,” *Journal of Computational and Applied Mathematics*, vol. 343, pp. 363–372, 2018.
- [6] G. L. G. Sleijpen and H. A. Van der Vorst, “A Jacobi-Davidson iteration method for linear eigenvalue problems,” *SIAM Journal on Matrix Analysis and Applications*, vol. 17, no. 2, pp. 401–425, 1996.
- [7] B.-S. Song and Y. Kong, “Solution to PSPACE-complete problem using P systems with active membranes with time-freeness,” *Mathematical Problems in Engineering*, vol. 2019, Article ID 5793234, 8 pages, 2019.
- [8] Y. Zhou, “Studies on Jacobi-Davidson, Rayleigh quotient iteration, inverse iteration generalized Davidson and Newton updates,” *Numerical Linear Algebra with Applications*, vol. 13, no. 8, pp. 621–642, 2006.
- [9] M. Genseberger and G. L. G. Sleijpen, “Alternative correction equations in the Jacobi-Davidson method,” *Numerical Linear Algebra with Applications*, vol. 6, no. 3, pp. 235–253, 1999.
- [10] G. Wu and H.-K. Pang, “On the correction equation of the Jacobi-Davidson method,” *Linear Algebra and Its Applications*, vol. 522, pp. 51–70, 2017.
- [11] Y. Lin, H. Dong, and Y. Fang, “N-soliton solutions for the NLS-like equation and perturbation theory based on the Riemann-Hilbert problem,” *Symmetry*, vol. 11, no. 6, p. 826, 2019.
- [12] Y.-X. Lin, Y. Fang, and H.-H. Dong, “Prolongation structures and N-soliton solutions for a new nonlinear Schrodinger-type equation via Riemann-Hilbert approach,” *Mathematical Problems in Engineering*, vol. 2019, Article ID 4058041, 10 pages, 2019.
- [13] Z.-Z. Bai and C.-Q. Miao, “On local quadratic convergence of inexact simplified Jacobi-Davidson method,” *Linear Algebra and Its Applications*, vol. 520, pp. 215–241, 2017.
- [14] Z.-Z. Bai and C.-Q. Miao, “On local quadratic convergence of inexact simplified Jacobi-Davidson method for interior eigenpairs of Hermitian eigenproblems,” *Applied Mathematics Letters*, vol. 72, pp. 23–28, 2017.
- [15] Z.-Z. Bai, C.-Q. Miao, and S. Jian, “On multistep Rayleigh quotient iterations for Hermitian eigenvalue problems,” *Computers & Mathematics with Applications*, vol. 77, no. 9, pp. 2396–2406, 2019.
- [16] Q. Ye, “Optimal expansion of subspaces for eigenvector approximations,” *Linear Algebra and Its Applications*, vol. 428, no. 4, pp. 911–918, 2008.



Hindawi

Submit your manuscripts at
www.hindawi.com

