Available online at www.sciencedirect.com
ScienceDirect
C. R. Acad. Sci. Paris, Ser. I 345 (2007) 293–296
COMPTES RENDUS MATHEMATIQUE
http://france.elsevier.com/direct/CRASS1/

Numerical Analysis

Convergence analysis of the Jacobi–Davidson method applied to a generalized eigenproblem

Grace Hechme

Laboratoire POEMS, UMR 2706 CNRS/ENSTA/INRIA, École nationale supérieure de techniques avancées, 32, boulevard Victor, 75739 Paris cedex 15, France

Received 14 September 2006; accepted after revision 18 July 2007

Available online 21 August 2007

Presented by Olivier Pironneau

Abstract

In this Note we consider the Jacobi–Davidson method applied to a nonsymmetric generalized eigenproblem. We analyze the convergence behavior of the method when the linear systems involved, known as the correction equations, are solved approximately. Our analysis also exhibits quadratic convergence when the corrections are solved exactly. ***To cite this article: G. Hechme, C. R. Acad. Sci. Paris, Ser. I 345 (2007).***

Résumé

Sur la convergence de la méthode de Jacobi–Davidson appliquée à un problème aux valeurs propres généralisé. Dans cette Note, la méthode de Jacobi–Davidson appliquée à un problème aux valeurs propres généralisé non symétrique est considérée. Nous analysons la convergence de la méthode quand les systèmes linéaires mis en jeu, plus connus sous le nom d'équations de correction, sont résolus approximativement. Notre analyse montre également la convergence quadratique de la méthode pour des solutions exactes de la correction. ***Pour citer cet article : G. Hechme, C. R. Acad. Sci. Paris, Ser. I 345 (2007).***

1. Introduction

We consider the computation of the eigenvalue with largest real part and the associated eigenvector of a generalized eigenproblem of order $n+m$

$$Ax = \lambda Bx, \tag{1}$$

where $B = \left(\begin{smallmatrix} I_n & 0 \\ 0 & 0 \end{smallmatrix}\right)$.

Such an eigenproblem arises in particular, in the linear stability analysis of differential and algebraic system of equations (DAE). For instance, after spatial discretization, hydrodynamic models [2] lead to a DAE of the following form:

E-mail address: grace.hechme@ensta.fr.

1631-073X/$ – see front matter
doi:10.1016/j.crma.2007.07.003

$$\begin{cases} \dfrac{\mathrm{d}v}{\mathrm{d}t} = F(v) - Gp, \quad G^{\mathrm{T}} v = 0; \\ v(0) = v_0, \quad p(0) = p_0. \end{cases}$$

In this framework,

$$A = \begin{pmatrix} J & -G \\ G^{\mathrm{T}} & 0 \end{pmatrix}, \qquad x = \begin{pmatrix} v \\ p \end{pmatrix},$$

where $v \in \mathbb{R}^n$ and $p \in \mathbb{R}^m$ are the velocity and the pressure, F is a quadratic function in v and $G \in \mathbb{R}^{n\times m}$ is a matrix of rank m. The matrix $J \in \mathbb{R}^{n\times n}$ is the Jacobian matrix of F computed at a steady state of the DAE.

Several approaches for the solution of the eigenproblem (1) can be used, mainly, the Shift-and-Invert Arnoldi method [3] and the Jacobi–Davidson method [4]. The first requires a judicious choice of the approximation of the sought after eigenpair (λ, x), as well as the exact solution of a linear system of order $n+m$. The latter can be prohibitive when $n+m$ is large. Such drawbacks do not occur with the Jacobi–Davidson method (JD). The method constructs a sequence (λ_k, x_k) such that $x_k^* B x_k \neq 0$ and $\lambda_k = \frac{x_k^* A x_k}{x_k^* B x_k}$, that converges to (λ, x) under suitable assumptions. Each iteration k requires the approximate solution of the equation below, known as the correction equation [4]

$$\begin{cases} \left(I - \dfrac{B x_k x_k^*}{x_k^* B x_k}\right)(A - \lambda_k B) t_k \;=\; -(A - \lambda_k B) x_k, \\ x_k^* t_k = 0. \end{cases} \tag{2}$$

The approximation x_{k+1} is then constructed as a linear combination of x_1 and of the vectors $t_1, \ldots, t_k$. The existence and uniqueness of the system's solution is analyzed hereafter.

In the sequel we give the main results on the convergence analysis of the Jacobi–Davidson method which is proved in detail in [1]. The convergence is based on an estimation of the angle between the eigenvector x and the particular combination $x_k + t_k$. In order to carry out the analysis, we investigate the method through one iteration. For ease of presentation, we set $\hat{x} = x_k$ and $\hat{\lambda} = \lambda_k$.

2. Properties of the correction equation

Consider $(\hat{\lambda}, \hat{x})$ the current approximation of a given finite eigenpair (λ, x) of eigenproblem (1). We seek to compute a correction (ν, t) of $(\hat{\lambda}, \hat{x})$, with t orthogonal to $\hat{x}$, such that

$$\alpha x = \hat{x} + t, \lambda = \hat{\lambda} + \nu \quad \text{with } \alpha \neq 0.$$

In other words,

$$A(\hat{x} + t) = (\hat{\lambda} + \nu) B (\hat{x} + t). \tag{3}$$

Assume that $\|\hat{x}\|_2 = \|x\|_2 = 1$, and $\hat{x}^* B x \neq 0$. Equality (3) enables us to write:

$$\nu = \frac{\hat{x}^*(A - \hat{\lambda} B)(\hat{x} + t)}{\hat{x}^* B (\hat{x} + t)} = \frac{\hat{x}^*(A - \hat{\lambda} B) t}{\hat{x}^* B (\hat{x} + t)}.$$

Substituting ν by its value in (3) and multiplying on the left by the projection operator $(\frac{I - B\hat{x}\hat{x}^*}{\hat{x}^* B \hat{x}})$ yields the following nonlinear equation

$$\left(I - \frac{B\hat{x}\hat{x}^*}{\hat{x}^* B \hat{x}}\right)(A - \hat{\lambda} B) t = -(A - \hat{\lambda} B)\hat{x} + \frac{\hat{x}^*(A - \hat{\lambda} B) t (\hat{x}^* B \hat{x} B t - \hat{x}^* B t B \hat{x})}{\alpha (\hat{x}^* B \hat{x})(\hat{x}^* B x)}. \tag{4}$$

Consider $\{\hat{x}\}^\perp$ and $\{B\hat{x}\}^\perp$ the orthogonal complement subspaces of $\hat{x}$ and $B\hat{x}$, U and V the matrices whose columns span orthonormal bases of $\{\hat{x}\}^\perp$ and $\{B\hat{x}\}^\perp$ respectively. Let $\mathcal{T}$ be the linear operator defined as follows:

$$\mathcal{T} : \{\hat{x}\}^\perp \longrightarrow \{\hat{x}\}^\perp, \\ t \longrightarrow \left(I - \frac{B\hat{x}\hat{x}^*}{\hat{x}^* B \hat{x}}\right)(A - \hat{\lambda} B) t.$$

Denote by $\Lambda\{A, B\}$ the set of eigenvalues of the matrix pair (A, B), by t_{quad} the solution of Eq. (4). In the following, we give conditions for existence and uniqueness of t_{quad}:

Theorem 2.1. *If*

$$\hat{\lambda} \notin \Lambda\{V^{\mathrm{T}}AU, V^{\mathrm{T}}BU\} \tag{5}$$

then the operator $\mathcal{T}$ is nonsingular. Furthermore, if

$$\frac{\|(A-\hat{\lambda}B)\hat{x}\|_2\|A-\hat{\lambda}B\|_2}{\alpha|\hat{x}^*B\hat{x}||\hat{x}^*Bx|} \leqslant \frac{\sigma_{\min}^2(\mathcal{T})}{8} \tag{6}$$

then there exists a unique vector $t_{\text{quad}} \in \{\hat{x}\}^{\perp}$ that solves (4)*. In addition, t_{quad} satisfies*

$$\|t_{\text{quad}}\|_2 = \tan\theta_{(x,\hat{x})} \quad \text{and} \quad \|t_{\text{quad}}\|_2 \leqslant 2\frac{\|(A-\hat{\lambda}B)\hat{x}\|_2}{\sigma_{\min}(\mathcal{T})},$$

where $\theta_{(x,\hat{x})}$ is the angle between x and $\hat{x}$, and $\sigma_{\min}(\mathcal{T})$ denotes the smallest singular value of $\mathcal{T}$.

It is straightforward to see that if the last term of (4) is omitted, we obtain a linear equation which we recognize as the Jacobi–Davidson correction equation (2). Denote by t_{lin} the solution of (2), the norm of the difference between t_{quad} and t_{lin} represents the error due to the linearization of (4). An upper bound is given in the proposition below:

Proposition 2.2. *Under hypothesis* (5) *and* (6)*, we have that*:

$$\|t_{\text{quad}} - t_{\text{lin}}\|_2 \leqslant \frac{2\|A-\hat{\lambda}B\|_2\|t_{\text{quad}}\|_2^2}{\alpha|\hat{x}^*B\hat{x}||\hat{x}^*Bx|\sigma_{\min}(\mathcal{T})}.$$

Theorem 2.1 shows, in particular, that when the angle between the eigenvector x and the approximation $\hat{x}$ is small, the vector t_{quad} has small components. Hence, according to Proposition 2.2, the error resulting from the linearization of (4) is small.

Remark 1. If $\hat{\lambda} \in \Lambda\{V^{\mathrm{T}}AU, V^{\mathrm{T}}BU\}$, the eigenvalue λ is possibly multiple and a block version of Jacobi–Davidson should be used.

3. Convergence estimates

In general, JD is applied to large sparse eigenproblems, and the exact solution of (2) can be extremely expensive to obtain numerically. Therefore, an approximation $t_{\text{app}} \in \{\hat{x}\}^{\perp}$ of t_{lin} is computed. We assume that

$$\|t_{\text{lin}} - t_{\text{app}}\|_2 \leqslant \beta\|t_{\text{lin}}\|_2 \quad \text{with } 0 < \beta < 1. \tag{7}$$

An iterative solver such as GMRES combined with an ad hoc preconditioner can be expected to give an approximate solution satisfying the assumption above (see [1]).

The next theorems provide some insight about the evolution of the approximation computed by JD with respect to the exact eigenvector x.

Theorem 3.1. *Under assumptions* (5)*,* (6) *and* (7) *we have*

$$\sin\theta_{(x,\hat{x}+t_{\text{app}})} \leqslant \left((1+\beta)\frac{2\|A-\hat{\lambda}B\|_2\tan\theta_{(x,\hat{x})}}{\alpha|\hat{x}^*B\hat{x}||\hat{x}^*Bx|\sigma_{\min}(\mathcal{T})} + \beta\right)\sin\theta_{(x,\hat{x})}. \tag{8}$$

Thus, if $\sigma_{\min}(\mathcal{T})$ is large and $\theta_{(x,\hat{x})}$ is small, then

$$\sin\theta_{(x,\hat{x}+t_{\text{app}})} = \mathcal{O}(\sin^2\theta_{(x,\hat{x})}) + \mathcal{O}(\beta\sin\theta_{(x,\hat{x})}).$$

In particular, if (2) is solved exactly, then $\beta = 0$ and we clearly see that quadratic convergence occurs.

A variant of Theorem 3.1 can be derived using a weaker assumption on the approximate solution t_{app}. We simply consider that t_{app} is computed by an iterative method with a tolerance γ, i.e.

$$\|\mathcal{T}(t_{\text{app}}) + (A-\hat{\lambda}B)\hat{x}\|_2 \leqslant \gamma\|(A-\hat{\lambda}B)\hat{x}\|_2. \tag{9}$$

Then we have the following theorem:

Theorem 3.2. *Under assumptions* (5), (6) *and* (9) *we have*

$$\sin\theta_{(x,\hat{x}+t_{\mathrm{app}})} \leqslant \frac{1}{\sigma_{\min}(\mathcal{T})}\left(\frac{2\|A-\hat{\lambda}B\|_2}{\alpha|\hat{x}^*B\hat{x}||\hat{x}^*Bx|}\tan\theta_{(x,\hat{x})} + \gamma\left(\|A-\lambda B\|_2 + \frac{\|A-\hat{\lambda}B\|_2}{\alpha|\hat{x}^*Bx|}\right)\right)\sin\theta_{(x,\hat{x})}. \tag{10}$$

This theorem shows that faster convergence can be achieved by taking a small tolerance γ. For instance, quadratic convergence is obtained with $\gamma = \delta\|(A-\hat{\lambda}B)\hat{x}\|_2$, where δ is a small scalar.

Remark 2. If the Euclidean inner product is replaced by the semi-inner product induced by the matrix B, a second version of the Jacobi–Davidson method can be developed, for which, similar convergence results hold [1].

Acknowledgement

The author would like to express her gratitude to Prof. M. Sadkane for his fruitful remarks.

References

[1] G. Hechme, Theoretical and numerical spectral analysis of some linearized hydrodynamic models, Ph.D. thesis, Universit é de Bretagne Occidentale, Brest, December 2005 (in French); http://www.ensta.fr/~hechme/new/these.html.
[2] R. Peyret, T.D. Taylor, Computational Methods for Fluid Flow, Springer Series in Computational Physics, Springer-Verlag, New York, 1983.
[3] Y. Saad, Numerical Methods for Large Eigenvalue Problems, Algorithms and Architectures for Advanced Scientific Computing, Manchester University Press, Manchester, UK, 1992.
[4] G.L. Sleijpen, A.G. Booten, D.R. Fokkema, H.A. Van Der Vorst, Jacobi–Davidson type methods for generalized eigenproblems and polynomial eigenproblems, BIT 36 (3) (1996) 595–633.