

A Jacobi–Davidson Method for Nonlinear Eigenproblems

Heinrich Voss

Section of Mathematics, Hamburg University of Technology, D – 21071 Hamburg
voss@tu-harburg.de, <http://www.tu-harburg.de/mat/hp/voss>

Abstract. For the nonlinear eigenvalue problem $T(\lambda)x = 0$ we consider a Jacobi–Davidson type iterative projection method. The resulting projected nonlinear eigenvalue problems are solved by inverse iteration. The method is applied to a rational eigenvalue problem governing damped vibrations of a structure.

1 Introduction

In this paper we consider the nonlinear eigenvalue problem

$$T(\lambda)x = 0 \tag{1}$$

where $T(\lambda) \in \mathbb{C}^{n \times n}$ is a family of large and sparse matrices depending on a parameter $\lambda \in D \subset \mathbb{C}$. Problems of this type arise in damped vibrations of structures, vibrations of rotating structures, stability of linear systems with retarded argument, lateral buckling problems or vibrations of fluid-solid structures, to name just a few. As in the linear case $T(\lambda) = \lambda I - A$ a parameter λ is called an eigenvalue of $T(\cdot)$ if problem (1) has a nontrivial solution $x \neq 0$ which is called a corresponding eigenvector.

For linear sparse eigenproblems iterative projection methods such as the Lanczos, Arnoldi or Jacobi–Davidson methods are very efficient. In these approaches one determines approximations to the wanted eigenvalues and corresponding eigenvectors from projections of the large eigenproblem to low-dimensional subspaces which are generated in the course of the algorithm. The small projected eigenproblems are solved by standard techniques.

Similar approaches for general nonlinear eigenproblems were studied in [2], [4], [7], and for symmetric problems allowing maxmin characterizations of the eigenvalues in [1] and [8].

Ruhe in [4] (with further modifications and improvements in [2]) linearized the nonlinear problem (1) by regula falsi and applied an Arnoldi type method to the varying sequence of linear problems thus constructing a sequence of search spaces V_k and Hessenberg matrices H_k which approximate the projection of $T(\sigma)^{-1}T(\lambda_k)$ to V_k . Here λ_k denotes an approximation to the wanted eigenvalue and σ a shift close to that eigenvalue. Then a Ritz vector of H_k corresponding to an eigenvalue of small modulus approximates an eigenvector of the nonlinear

problem (1) from which a new approximation to the corresponding eigenvalue is obtained. Hence, in this approach the two numerical subtasks reducing the large dimension to a low one and solving the projected nonlinear eigenproblem are attacked simultaneously.

In this paper we suggest an iterative projection method for the nonlinear eigenproblem where the two subtasks mentioned in the last paragraph are handled separately. If V_k denotes a subspace of \mathbb{C}^n of small dimension k constructed in the course of the algorithm we solve the projected nonlinear eigenvalue problem $V_k^H T(\lambda) V_k z = 0$ by a dense solver to obtain an approximate eigenvalue λ_k and eigenvector $x_k = V_k z$. After that we expand the space V_k . Similarly as in the Jacobi–Davidson method for linear eigenproblems the expansion direction v_{k+1} of V_k is chosen such that $x_k + \alpha v_{k+1}$ for some $\alpha \in \mathbb{C}$ has a high approximation potential for the eigenvector we are just aiming at. The projection step and the expansion step are repeated alternately until convergence.

Here we consider a method of this type where the search space V_k is expanded by an approximate solution of a correction equation

$$\left(I - \frac{T'(\lambda_k) x_k x_k^H}{x_k^H T'(\lambda_k) x_k} \right) T(\lambda_k) \left(I - \frac{x_k x_k^H}{x_k^H x_k} \right) v = -T(\lambda_k) x_k$$

in a Jacobi–Davidson like manner. In [7] we proposed an expansion of the search space by $v_{k+1} = T(\sigma)^{-1} T(\lambda_k) x_k$ generalizing the residual inverse iteration for dense nonlinear eigenproblems.

The paper is organized as follows. Section 2. discusses the expansion of the search space in a Jacobi–Davidson type way. In particular we discuss the approximate solution of the correction equation by a preconditioned Krylov subspace method. Section 3. reviews solvers of dense nonlinear eigenproblems with special emphasis on the fact that nonlinear problems are often small perturbations of linear problems which can be exploited in the solution process. Section 4. contains the Jacobi–Davidson method for nonlinear eigenproblems and Section 5 demonstrates its numerical behavior for a finite element model of a structure.

2 Expanding the Search Space by Jacobi–Davidson

The Jacobi–Davidson method was introduced by Sleijpen and van der Vorst (cf. [6]) for the linear eigenproblem $Ax = \lambda x$, and generalized in a series of papers with different co-authors to general and to polynomial eigenvalue problems (cf. [5]). Its idea is to construct a correction for a given eigenvector approximation x in a subspace orthogonal to x . Namely, if V is the current search space and (θ, u) , $\|u\| = 1$, is a Ritz pair of $Ax = \lambda x$ corresponding to V then V is expanded by a solution t of the so called correction equation

$$(I - uu^H)(A - \theta I)(I - uu^H)t = -(A - \theta I)u, \quad t \perp u.$$

If the correction equation is solved exactly then it is easily seen that the new search space $[V, t]$ contains the vector $\tilde{t} = (A - \theta I)^{-1}u$ obtained by one step

of shifted inverse iteration, and therefore one can expect quadratic (and in the Hermitian case even cubic) convergence.

A natural generalization to the nonlinear eigenproblem (1) which was already suggested in [5] for polynomial eigenvalue problems is the following one: Suppose that the columns of $V \subset \mathbb{C}^k$ form an orthonormal basis of the current search space, and let (u, θ) be a Ritz pair of (1) with respect to V , i.e. $V^H T(\theta) V y = 0$, $u = V y$. Then we consider the correction equation

$$\left(I - \frac{p u^H}{u^H p}\right) T(\theta) \left(I - \frac{u u^H}{u^H u}\right) t = -r, \quad t \perp u \quad (2)$$

where $p := T'(\theta)u$ and $r := T(\theta)u$.

Equation (2) can be rewritten as $T(\theta)t - \alpha p = -r$ where α has to be chosen such that $t \perp u$. Solving for t we obtain

$$t = -u + \alpha T(\theta)^{-1} p = -u + \alpha T(\theta)^{-1} T'(\theta)u,$$

and $u = V y$ yields $\tilde{t} := T(\theta)^{-1} T'(\theta)u \in \text{span}[V, t]$.

Hence, as in the linear case the new search space $\text{span}[V, t]$ contains the vector obtained by one step of inverse iteration with shift θ and initial vector u , and we may expect quadratic or even cubic convergence of the resulting iterative projection method, if the correction equation (2) is solved exactly.

It has been observed by Sleijpen and van der Vorst for linear problems that the correction equation does not have to be solved accurately but fast convergence of the projection method is maintained if the search space is expanded by an approximate solution, and the same holds true for nonlinear problems. For the linear problem they suggested to apply a few steps of a Krylov solver with an appropriate preconditioner.

In the correction equation (2) the operator $T(\theta)$ is restricted to map the subspace u^\perp to $T'(\theta)u^\perp$. Hence, if $K \approx T(\theta)$ is a preconditioner of $T(\theta)$ then a preconditioner for an iterative solver of (2) should be modified correspondingly to

$$\tilde{K} := \left(I - \frac{p u^H}{u^H p}\right) K \left(I - \frac{u u^H}{u^H u}\right).$$

With left-preconditioning equation (2) becomes

$$\tilde{K}^{-1} \tilde{T}(\theta) t = -\tilde{K}^{-1} r, \quad t \perp u. \quad (3)$$

where

$$\tilde{T}(\theta) := \left(I - \frac{p u^H}{u^H p}\right) T(\theta) \left(I - \frac{u u^H}{u^H u}\right).$$

We apply a Krylov solver to equation (3) with initial guess $t = 0$. For the linear case this was already discussed in [6], and the transfer to equation (3) is straightforward.

Since the operator $\tilde{K}^{-1} \tilde{T}(\theta)$ maps the space u^\perp into itself, and since the initial guess $t = 0$ is an element of u^\perp , all iterates are contained in this space,

and therefore in each step we have to perform one matrix-vector product $y = \tilde{K}^{-1}\tilde{T}(\theta)v$ for some $v \in u^\perp$. To this end we first multiply v by $\tilde{T}(\theta)$ which yields

$$\tilde{y} = \left(I - \frac{pu^H}{u^Hp}\right)T(\theta)v = T(\theta)v - \frac{u^HT(\theta)v}{u^Hp}p,$$

and then we solve $\tilde{K}y = \tilde{y}$, $y \perp u$.

This equation can be rewritten as $Ky - \alpha p = \tilde{y}$, where α is determined from the condition $y \perp u$. Thus, we finally obtain

$$y = K^{-1}\tilde{y} - \frac{u^HK^{-1}\tilde{y}}{u^HK^{-1}p}K^{-1}p \quad (4)$$

which demonstrates that taking into account the projectors in the preconditioner, i.e. using \tilde{K} instead of K , raises the cost of the preconditioned Krylov solver only slightly. To initialize one has to solve the linear system $K\tilde{p} = p$ and to determine the scalar product $\alpha := u^H\tilde{p} = u^HK^{-1}p$. These computations have to be executed just once. Afterwards in each iteration step one has to solve only one linear system $Kw = \tilde{y}$ for w , one has to compute the scalar product $\beta := u^Hw = u^HK^{-1}\tilde{y}$, and to perform one axpy $y = w - (\beta/\alpha)\tilde{y}$ to expand the Krylov space of $\tilde{K}^{-1}\tilde{T}(\theta)$.

3 Solving Projected Nonlinear Eigenproblems

Since the dimensions of the projected eigenproblems are usually small they can be solved by any method for dense nonlinear eigenproblems like inverse iteration or residual inverse iteration.

If $T(\lambda)$ is symmetric or Hermitean such that the eigenvalues are real and can be characterized as minmax values of a Rayleigh functional then the projected problem inherits this property, and the eigenvalues can be determined one after the other by safeguarded iteration. This approach which was discussed for the Jacobi–Davidson method in [1] and for the Arnoldi method in [8] has the advantage that it is most unlikely that the method converges to an eigenvalue that has already been found previously.

In the general case the following strategy is similar to safeguarded iteration. Assume that we want to determine all eigenvalues of problem (1) in the vicinity of a given parameter $\sigma_0 \in D$, and that already $m - 1$ eigenvalues closest to σ_0 have been determined. Assume that $\tilde{\mu}$ is an approximation to the eigenvalue wanted next.

A first order approximation of problem (1) is

$$T(\lambda)x \approx (T(\tilde{\mu}) - \theta T'(\tilde{\mu}))x = 0, \quad \theta = \tilde{\mu} - \lambda. \quad (5)$$

This suggests the method of successive linear problems in Algorithm 1 which was introduced by Ruhe [3], and which converges quadratically.

Of course this method is not appropriate for a sparse problem (1), but in an iterative projection method the dimension of the projected problem which has

Algorithm 1 Method of successive linear problems

- 1: Start with an approximation μ_1 to the m -th eigenvalue of (1)
 - 2: **for** $\ell = 1, 2, \dots$ until convergence **do**
 - 3: solve the linear eigenproblem $T(\mu_\ell)u = \theta T'(\mu_\ell)u$
 - 4: choose the eigenvalue θ such $|\sigma_0 - (\mu_\ell - \theta)|$ is the m -smallest among the eigenvalues
 - 5: $\mu_{\ell+1} = \mu_\ell - \theta$
 - 6: **end for**
-

to be solved in step 3. usually is quite small, and every standard solver for dense eigenproblems applies.

Quite often the nonlinear eigenvalue problem under consideration is a (small) perturbation of a linear eigenvalue problem. As a numerical example we will consider a finite element model of a vibrating structure with nonproportional damping. Using a viscoelastic constitutive relation to describe the behavior of a material in the equations of motions yields a rational eigenvalue problem for the case of free vibrations. A finite element model obtains the form

$$T(\omega) := \left(\omega^2 M + K - \sum_{j=1}^J \frac{1}{1 + b_j \omega} \Delta K_j \right) x = 0. \quad (6)$$

If the damping is not too large the eigenmodes of the damped and the undamped problem do not differ very much although the eigenvalues do. Therefore, in step 3. of Algorithm 2 it is reasonable to determine an eigenvector y of the undamped and projected problem $(\omega^2 V^H M V + V^H K V)y = 0$ corresponding to the m -smallest eigenvalue ω_m^2 , determine an approximate eigenvalue $\tilde{\omega}$ of the nonlinear projected problem from the complex equation $y^H V^H T(\omega) V y = 0$ or $e^H V^H T(\sigma)^{-1} T(\omega) V y = 0$, and correct it by (residual) inverse iteration.

4 Jacobi–Davidson Method for Nonlinear Eigenproblems

A template for the Jacobi–Davidson method for the nonlinear eigenvalue problem (1) is given in Algorithm 2.

Remarks on some of its steps are in order:

1. In V preinformation about the wanted eigenvectors (which may be gained from previous solutions of similar problems) can be introduced into the method.
If we are interested in eigenvalues close to a given parameter σ and no information on eigenvectors is at hand we can start the Jacobi–Davidson method with an orthogonal basis V of an invariant subspace of the linear eigenproblem $T(\sigma)u = \theta u$ (or $T(\sigma)u = \theta T'(\sigma)u$) corresponding to eigenvalues θ which are small in modulus.
8. As the subspaces expand in the course of the algorithm the increasing storage and the computational cost for solving the projected eigenproblems may

Algorithm 2 Nonlinear Jacobi–Davidson method

- 1: Start with an initial basis V , $V^H V = I$; $m = 1$
- 2: determine a preconditioner $K \approx T(\sigma)^{-1}$, σ close to the first wanted eigenvalue
- 3: **while** $m \leq$ number of the wanted eigenvalues **do**
- 4: compute an approximation to the m -th wanted eigenvalue λ_m and corresponding eigenvector x_m of the projected problem $V^H T(\lambda) V x = 0$
- 5: determine the Ritz vector $u = V x_m$ and the residual $r = T(\lambda_m) u$
- 6: **if** $\|r\|/\|u\| < \epsilon$ **then**
- 7: accept approximate eigenpair (λ_m, u) ; increase $m = m + 1$;
- 8: reduce the search space V if necessary
- 9: choose an approximation (λ_m, u) to the next eigenpair
- 10: compute the residual $r = T(\lambda_m) u$;
- 11: **end if**
- 12: Find an approximate solution of the correction equation

$$\left(I - \frac{T'(\lambda_m) u u^H}{u^H T'(\lambda_m) u}\right) T(\sigma) \left(I - \frac{u u^H}{u^H u}\right) t = -r \tag{7}$$

- 13: orthogonalize $t = t - V V^H t$, $v = t/\|t\|$, and expand the subspace $V = [V, v]$
 - 14: determine a new preconditioner $K \approx T(\lambda_m)$ if necessary
 - 15: update the projected problem
 - 16: **end while**
-

make it necessary to restart the algorithm and to purge some of the basis vectors. Since a restart destroys information on the eigenvectors and particularly the one the method is just aiming at we restart only if an eigenvector has just converged. A reasonable search space after restart is the space spanned by the already converged eigenvectors (or a space slightly larger).

12. The correction equation can be solved by a preconditioned Krylov solver, e.g.
13. The first two statements represent the classical Gram–Schmidt process. It is advisable to repeat this orthogonalization step once if the norm of t is reduced by more than a modest factor, say $\|t\|/\|v_{\text{old}}\| < 0.25$, e.g.
14. We solved the correction equation (7) by a few steps of preconditioned GMRES where we kept the preconditioner for a couple of eigenvalues. We terminated the solver of (7) in the k -th outer iteration for the m -th eigenvalue if the residual was reduced by at least $\tau_k = 2^{-k}$, and we allowed at most 10 steps of the solver. If the required accuracy τ_k was not met after at most 5 iteration steps we updated the preconditioner. However, we allowed at most one update for every eigenvalue λ_m .

5 Numerical Experiments

To test the Jacobi–Davidson method we consider the rational eigenvalue problem (6) governing damped vibrations of a column

$$\{(x, y, z) : 0.8 < \sqrt{x^2 + y^2} < 1, 0 < z < 5\}$$

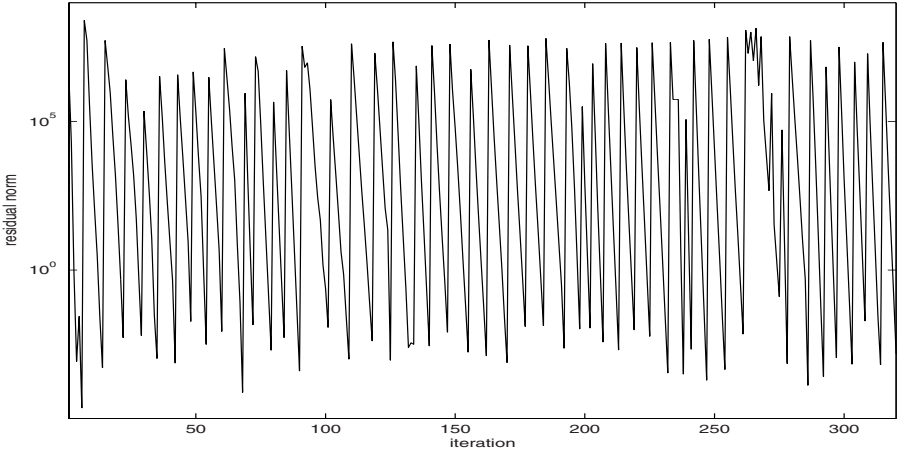


Fig. 1. Convergence history without restarts

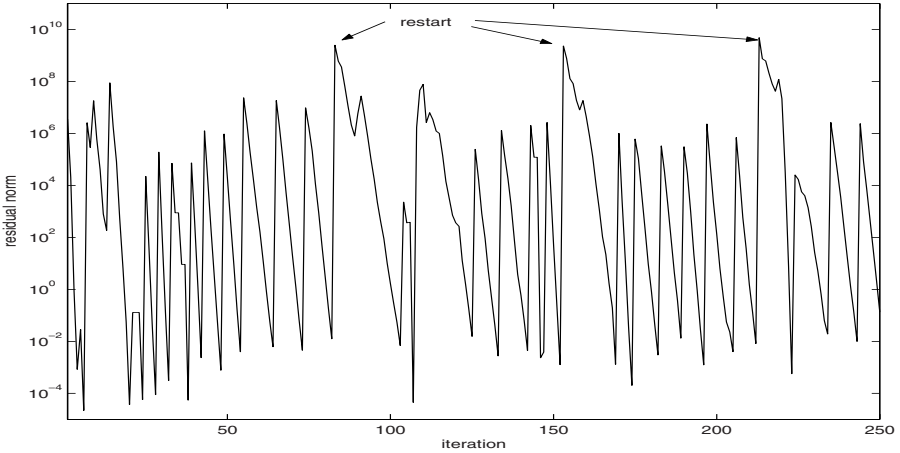


Fig. 2. Convergence history with restarts (first 250 iterations)

which is clamped at its bottom $z = 0$. The instantaneous Young's modulus is set to $E = 2.06 \times 10^{11}$, the instantaneous Poisson's rate is $\nu = 0.3$, and the density is set to $\rho = 7800$. For the nonproportional damping we use in addition the following parameters, $\Delta\nu = 0.27$, and $\Delta E = 6 \times 10^{10}$ for $0 < x < 2.5$, and $\Delta E = 5 \times 10^{10}$ for $2.5 < x < 5$. The relaxation constant is set to $b = 10^{-4}$.

Discretizing this problem by linear Lagrangian elements we obtained the rational eigenproblem (6) of dimension 11892, and the bandwidth of the stiffness matrix K was after reducing it by reverse Cuthill–McKee algorithm still 665. For symmetry reasons we determined only eigenvalues with negative imaginary part, and we computed 50 of them one after another with decreasing imaginary part.

The nonlinear projected eigenproblems were solved by inverse iteration with an initial guess obtained from the corresponding undamped projected problem as explained at the end of Section 3.

The experiments were run under MATLAB 6.5 on a Pentium 4 processor with 2 GHz and 1 GB RAM. We preconditioned by the LU factorization of $T(\sigma)$, and terminated the iteration if the norm of the residual was less than 10^{-2} .

Starting with an eigenvector of the linear eigenproblem $Kx = \lambda Mx$ corresponding to the smallest eigenvalue the algorithm without restarts needed 320 iteration steps, i.e. an average of 6.4 iterations per eigenvalue, to approximate all 50 eigenvalues (including double eigenvalues) with maximal negative imaginary part. To solve the correction equations a total of 651 GMRES steps were needed, and 6 updates of the preconditioner were necessary. Fig. 1. contains the convergence history.

Restarting the Jacobi–Davidson process if the dimension of the search space exceeded 80 the method needed 7 restarts. Again all 50 eigenvalues were found by the method requiring 422 iterations, 840 GMRES steps, and 16 updates of the preconditioner. The convergence history in Fig. 2. looks very similar to the one without restarts, however, after a restart the speed of convergence was reduced considerably. After a restart an average of 17.1 iterations was necessary to gather enough information about the search space and to make the method converge, whereas for the other iteration steps the average number of steps for convergence was 7.0.

References

1. T. Betcke and H. Voss. A Jacobi–Davidson–type projection method for nonlinear eigenvalue problems. *Future Generation Computer Systems*, 20(3):363 – 372, 2004.
2. P. Hager. *Eigenfrequency Analysis. FE-Adaptivity and a Nonlinear Eigenvalue Problem*. PhD thesis, Chalmers University of Technology, Göteborg, 2001.
3. A. Ruhe. Algorithms for the nonlinear eigenvalue problem. *SIAM J. Numer. Anal.*, 10:674 – 689, 1973.
4. A. Ruhe. A rational Krylov algorithm for nonlinear matrix eigenvalue problems. *Zapiski Nauchnyh Seminarov POMI*, 268:176 – 180, 2000.
5. G.L. Sleijpen, G.L. Booten, D.R. Fokkema, and H.A. van der Vorst. Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems. *BIT*, 36:595 – 633, 1996.
6. G.L. Sleijpen and H.A. van der Vorst. A Jacobi-Davidson iteration method for linear eigenvalue problems. *SIAM J. Matr. Anal. Appl.*, 17:401 – 425, 1996.
7. H. Voss. An Arnoldi method for nonlinear eigenvalue problems. Technical Report 56, Section of Mathematics, Hamburg University of Technology, 2002. To appear in BIT Numerical Mathematics.
8. H. Voss. An Arnoldi method for nonlinear symmetric eigenvalue problems. In *Online Proceedings of the SIAM Conference on Applied Linear Algebra, Williamsburg*, <http://www.siam.org/meetings/laa03/>, 2003.