A Jacobi–Davidson like method for nonlinear eigenvalue problems based on singularity theory

Kathrin Schreiber1,* and Hubert Schwetlick2,**

1 Technische Universität Berlin, Institut für Mathematik, Straße des 17. Juni 136, 10623 Berlin, Germany.
2 Technische Universität Dresden, Institut für Numerische Mathematik, 01062 Dresden, Germany.

We present a Jacobi–Davidson like correction formula for left and right eigenvector approximations for non-Hermitian nonlinear eigenvalue problems. It exploits techniques from singularity theory for characterizing singular points of nonlinear equations. Unlike standard nonlinear Jacobi-Davidson, the correction formula does not contain derivative information and works with orthogonal projectors only. Moreover, the basic method is modified in that the new eigenvalue approximation is taken as a nonlinear Rayleigh functional obtained as root of a certain scalar nonlinear equation the existence of which—as well as a first order perturbation expansion—is shown.

1 Generalized Jacobi–Davidson

Consider the nonlinear eigenvalue problem

$$T(\lambda)x = 0$$

(1)

where $T(\cdot) : \mathbb{C} \to \mathbb{C}^{n \times n}$ is a sufficiently smooth matrix-valued mapping. We seek an eigentriplet $(\lambda_*, x_*, y_*)$, where $\lambda_*$ is a simple eigenvalue and $x_*, y_*$ is the right and left eigenvector, resp. A well-known way for solving this problem is given by Newton’s method applied to an extended system, where a normalization condition, for instance $w^H x - 1 = 0$ with $0 \neq w \in \mathbb{C}^n$, is added to (1). The Newton step then looks as follows

$$
\begin{bmatrix}
T(\lambda) & \dot{T}(\lambda)u_H & 0 \\
0 & u_H^H & 0 \\
\end{bmatrix}
\begin{bmatrix}
u_+ - u \\
\lambda_* - \lambda \\
\end{bmatrix}
= -
\begin{bmatrix}
T(\lambda)u \\
u^H u_H - 1 \\
\end{bmatrix}
\iff
T(\lambda)u_+ + (\lambda_* - \lambda)\dot{T}(\lambda)u = 0 \quad \text{with } u_H u_+ = 1,
$$

(2)

see [1]. If $\|u\| = 1$ and $w = u$, then $\lambda_* - \lambda$ can be eliminated from (2) which yields the nonlinear Jacobi–Davidson correction equation, cf. [2],

$$A_{JD}s := \left(I - \frac{\dot{T}(\lambda)u_H}{u_H^H T(\lambda)u}\right)T(\lambda)(I - uu^H)s = -r_u, \quad u^H s = 0
$$

(3)

for the correction $s \equiv u_+ - u$, where we assumed, that $T(\theta)u := r_u \perp u$, i.e., $u^H T(\theta)u = 0$, which is guaranteed by a Rayleigh–Ritz step beforehand. If used for non-Hermitian problems, JD converges quadratically in case $x^H T(\lambda_*)x \neq 0$.

Now, let $A_{JD}$ denote the linear operator in (3) that maps $s \perp u$ to $A_{JD}s \perp u$, which is one-to-one for sufficiently good $(\lambda, u, v)$, and consider it at the solution $(\lambda_*, x_*, y_*)$. Then one can prove, cf. [3],

$$\frac{1}{\|y^H T(\lambda_*)x_*\|} \cdot \frac{|x^H T(\lambda_*)x_*|}{\|T(\lambda_*)\|} \leq \|A_{JD}^{-1}\| \leq \frac{\|\dot{T}(\lambda_*)\|}{\|y^H T(\lambda_*)x_*\|} \cdot \|T(\lambda_*)\|.
$$

Hence, for ill-conditioned eigenvalues, $A_{JD}$ will have a large inverse. The reason is the right bordering vector $\dot{T}(\lambda)u$ in (2).

Since $y_* \perp \text{im } T(\lambda_*) = (\ker T(\lambda_*)^H)^\perp$ would be the optimal bordering, we replace $\dot{T}(\lambda)u$ with an approximation $v$ to $y_*$. The same ansatz is done in singularity theory, cf. [4]. The augmented system, that we work with, is then

$$
\begin{bmatrix}
T(\lambda) & v \\
u^H & 0 \\
\end{bmatrix}
\begin{bmatrix}
x \\
\mu \\
\end{bmatrix}
= \begin{bmatrix}
0 \\
1 \\
\end{bmatrix}
\iff
T(\lambda)x + v\mu = 0 \quad u^H x = 1.
$$

(4)

It defines $\mu = \mu(\lambda)$ as function of $\lambda$, and one Newton step with respect to $\lambda$ is performed. The corresponding algorithm and analysis is provided in [3], quadratic convergence for simple eigenvalues is shown. System (4) can be written in terms of projected matrices and yields $(I - uu^H)T(\lambda)(I - uu^H)s = -T(\lambda)u$, with $u^H s = 0$, where a dual equation, which has to be solved simultaneously in order to obtain the correction for the left eigenvector, comes from the dual system of (4).

* e-mail: schreibe@math.tu-berlin.de, Phone: +49 30 314 25741, Fax: +49 30 314 79706
** e-mail: Hubert.Schwetlick@tu-dresden.de, Phone: +49 351 4633 4267, Fax: +49 351 4633 4268
We remark, that the second system in S3 includes the adjoint operator of the first one, such that the same preconditioner can be used. Moreover, both projectors are orthogonal and no multiplication with $T(\lambda)$ is needed.

The norm of the inverse operator at the solution does not depend on the term $y^HT(\lambda_s)x_s$, nor on $x^HT(\lambda_s)x_s$. Thus it is especially well-suited for ill-conditioned eigenvalues, or the case when $x^HT(\lambda_s)x_s = 0$. For numerical experiments see [3].

## 2 Nonlinear Rayleigh functionals

Step 1 of the method above uses some sort of higher-dimensional Rayleigh functional, the theoretical properties of the one-dimensional one shall be illuminated here.

Rayleigh functionals for nonlinear eigenvalue problems are known since 1955 [5], though in the then following literature only Hermitian problems with sets of real eigenvalues are discussed, and authors presume, that a Rayleigh functional exists. Following Ostrowski’s [6] two-sided Rayleigh quotient, we define the generalized or two-sided nonlinear Rayleigh functional $p : (u, v) \in K(x_s, \varepsilon) \times K(y_s, \varepsilon) \rightarrow p(u, v) \in \mathbb{C}$ by

\[
(T(p(u, v))u, v) = 0,
\]

see [7], with $(x, y) := y^HTx$ and $K(z, \varepsilon) = \{ w \in \mathbb{C}^n : \angle(\text{span}\{w\}, \text{span}\{z\}) \leq \varepsilon \}$. The proofs of the following Theorem and Lemma are given in [7].

**Theorem 2.1** Suppose $T(\lambda)$ is holomorphic on the open disc $S(\lambda, \tau_0)$, or if real, suppose $\dot{T}(\lambda)$ is there Lipschitz continuous, and that $\lambda_s$ is simple with $x_s, y_s$ as corresponding eigenvectors of unit norm. Then there are constants $0 < \tau_0 \leq \tau_s$, $0 < \varepsilon_0 < \pi/2$ such that, for all $(u, v) \in K(x_s, \varepsilon_0) \times K(y_s, \varepsilon_0)$, there exists a unique $p = p(u, v) \in S_0 := S(\lambda_s, \tau_0)$ with $(T(p(u, v)), u) = 0$. Moreover, one has

\[
|p(u, v) - \lambda_s| \leq \frac{3}{8} \frac{\|T(\lambda_s)\|}{|y_sT(\lambda_s)x_s|} \cdot \tan(\angle(\text{span}\{u\}, \text{span}\{x_s\})) \cdot \tan(\angle(\text{span}\{v\}, \text{span}\{y_s\})).
\]

We also proved a first order perturbation expansion of the complex Rayleigh functional $p(u, v)$ in [7], where there is a Taylor expansion in the real case but not in the complex one.

**Lemma 2.2** Let the assumptions of Theorem 2.1 be satisfied, and let $\tau_0, \varepsilon_0$ be as specified in the Theorem. Then there exists $\delta_0 > 0$ with the following properties: For all $(u, v) \in K(x_s, \varepsilon_0/2) \times K(y_s, \varepsilon_0/2)$ and all $(s, t)$ with $\|s\| \leq \delta_u \|u\|$, $\|t\| \leq \delta_v \|v\|$, $0 \leq \delta_u, \delta_v \leq \delta_0$, we have $(u + s, v + t) \in K(x_s, \varepsilon_0) \times K(y_s, \varepsilon_0)$, and $p(u + s, v + t)$ exists and is unique in $S(\lambda_s, \tau_0)$. Moreover, we have

\[
p(u + s, v + t) = p(u, v) - \frac{v^HT(p(u, v))s + u^HT(p(u, v))u}{v^HT(p(u, v))u} - g(s, t) \quad \text{with} \quad |g(s, t)| \leq K(\delta_u + \delta_v)^2.
\]

Although we cannot assume that $p(u, v)$ is differentiable (since even if defined as generalized Rayleigh quotient $p(u, v) = (v^HAu)/v^Hu$ for a matrix $A$, it is not), Lemma 2.2 provides a first order perturbation bound, which applied to $u = x_s, v = y_s$ gives $p(x_s + s, y_s + t) - \lambda_s = \rho(s, t)$, which reflects what stationarity would be in the real (differentiable) case.

Both results can be applied to the standard nonlinear Rayleigh functional defined by equation $(T(p(u))u, u) = 0$ when $T(\lambda)$ is Hermitian, hence $y_s = x_s$, by setting $v = u$, cf. [7].

## References