A projection method for nonlinear eigenvalue problems using contour integrals

Shinnosuke Yokota$^1$ and Tetsuya Sakurai$^{2,3}$

$^1$College of Information Science, University of Tsukuba, Tsukuba 305-8573, Japan
$^2$Department of Computer Science, University of Tsukuba, Tsukuba 305-8573, Japan
$^3$CREST, Japan Science and Technology Agency, Kawaguchi 332-0012, Japan

E-mail yokota@mma.cs.tsukuba.ac.jp

Received June 10, 2012, Accepted September 12, 2012

Abstract

In this paper, we indicate that the Sakurai-Sugiura method with Rayleigh-Ritz projection technique, a numerical method for generalized eigenvalue problems, can be extended to nonlinear eigenvalue problems. The target equation is $T(\lambda)\nu = 0$, where $T$ is a matrix-valued function. The method can extract only the eigenvalues within a Jordan curve $\Gamma$ by converting the original problem to a problem with a smaller dimension. Theoretical validation of the method is discussed, and we describe its application using numerical examples.

Keywords nonlinear eigenvalue problem, contour integral, Rayleigh-Ritz procedure

Research Activity Group Algorithms for Matrix / Eigenvalue Problems and their Applications

1. Introduction

We consider a numerical method using contour integrals to solve nonlinear eigenvalue problems (NEPs). The nonlinear eigenvalue problem involves finding eigenpairs $(\lambda, \nu)$ that satisfy $T(\lambda)\nu = 0$, where the matrix-valued function $T : \Omega \to C^{n \times n}$ is assumed to be holomorphic in some open domain $\Omega \subset C$. NEPs appear in a variety of settings in science and engineering (e.g., modeling drift instabilities in a plasma and modeling a radio-frequency gun cavity [1]).

We herein propose a numerical method using contour integrals to solve NEPs. The method is closely related to the Sakurai-Sugiura method with the Rayleigh-Ritz projection technique (SS-RR) for generalized eigenvalue problems (GEPs) [2] and inherits many of its strong points, including suitability for execution on modern distributed parallel computers.

In this paper, we will further generalize the SS-RR method to NEPs. In the SS-RR method, the original problem is converted to a problem with a smaller dimension. The converted problem is obtained numerically by solving a set of linear equations. These linear equations are derived from the original problem and can form a large system, but they are independent and can be solved in parallel. Moreover, the proposed method is free from the fixed point iterations required in the nonlinear Rayleigh-Ritz iterative method [3], such as the nonlinear Arnoldi algorithm [4] and the nonlinear Jacobi-Davidson method [5].

The Sakurai-Sugiura method with block Hankel matrices (SS-H) which is another numerical method using contour integrals for NEPs has been proposed already in [6]. The SS-RR method uses the same contour integrals of the SS-H method. However, both methods differ in the way they obtain approximate eigenpairs from the subspace. According to [2], in GEPs, the SS-RR method usually gives better numerical results than the SS-H method. We expect that similar results can be obtained in NEPs.

The extension of the SS-RR method for NEPs is discussed from a theoretical point of view using the Keldysh theorem. Numerical examples are also reported, with results that are consistent with the theory.

The remainder of the present paper is organized as follows. In the next section, we introduce the Keldysh theorem for a matrix-valued function $T$, as presented in [7], and an evaluation method for obtaining $T^{-1}$. In Section 3, we show that the contour integrals corresponding to $T^{-1}$ can make it possible to obtain the eigenspace designate. In Section 4, we state the Rayleigh-Ritz procedure [8] and show the algorithm of the SS-RR method. Some numerical examples are shown in Section 5. Finally, conclusions and suggestions for future research are presented in Section 6.

2. Basics of nonlinear eigenvalue problems

Let $T : \Omega \to C^{n \times n}$ be a matrix-valued function that is holomorphic in an open domain $\Omega$. Throughout this paper, we assume that $T$ is regular.

A holomorphic vector-valued function $\nu : \Omega \to C^n$ is called a root function of $T$ at $\lambda \in \Omega$ if

$$T(\lambda)\nu(\lambda) = 0, \quad \nu(\lambda) \neq 0.$$  \hspace{1cm} (1)

Let $\nu$, the order of the zero of $T(z)\nu(z)$ at $z = \lambda$, be a multiplicity of $\nu$. Because $\nu$ is holomorphic, it admits an expansion of the form

$$\nu(z) = \sum_{j=0}^{\infty} (z - \lambda)^j \nu_j, \quad \nu_0 \neq 0.$$
Definition 1  Given a root function (1) of multiplicity \( \nu \), any vector sequence of the form

\[
v_\mu \ldots v_{\mu-1}, \quad 1 \leq \mu \leq \nu
\]

is called a Jordan chain, \( v_\mu \) is an eigenvector, and \( v_1, \ldots, v_{\mu-1} \) are said to be associated vectors for \( v_\mu \) of \( T \) at \( \lambda \).

Let \( n \) be the dimensionality of the eigenspace \( \text{Ker } T(\lambda) \). We assume that for every \( \ell = 1, \ldots, n \), \( v_1^\ell, \ldots, v_{\mu-1}^\ell \) form a Jordan chain corresponding to the eigenvectors and associated vectors of \( T(z) \) at \( \lambda \), i.e.

\[
v^\ell(z) = \sum_{j=1}^{\mu-1} (z - \lambda)^j v^\ell_j, \quad 1 \leq \mu \leq \nu,
\]

where \( \ell = 1, \ldots, n \) and \( v^\ell_0, \ldots, v^\ell_{\mu-1} \) are linearly independent. Then the system

\[
V = (v^\ell_j, j = 0, \ldots, \mu - 1, \ell = 1, \ldots, n)
\]

is called a canonical system of Jordan chains (CSJC) of \( T \) at \( \lambda \). Note that the sum of multiplicities of root functions \( v^1, \ldots, v^n \) is called the algebraic multiplicity of \( \lambda \), and \( n \) is called the geometric multiplicity of \( \lambda \).

Using the notation in the above definition, we state the following theorem regarding all eigenvalues inside \( C \subset \Omega \) (see [7, Corollary 2.8]) for leading our method.

**Theorem 2** Let \( C \subset \Omega \) be a compact set which contains at most finitely many eigenvalues \( \lambda_k, k = 1, \ldots, n(\Omega) \) with corresponding CSJC of \( T \)

\[
V_k = \left( v^{\ell,k}_j, j = 0, \ldots, \mu - 1, \ell = 1, \ldots, n_k \right).
\]

Let

\[
W_k = \left( w^{\ell,k}_j, j = 0, \ldots, \mu - 1, \ell = 1, \ldots, n_k \right)
\]

be the corresponding CSJC's of \( T^H \). Then, there exists a neighborhood \( C \subset U \subset \Omega \) and a holomorphic matrix-valued function \( R : \Omega \to \mathbb{C}^{n \times n} \) such that for all \( z \in U \setminus \{ \lambda_1, \ldots, \lambda_n(C) \} \),

\[
T(z)^{-1} = \sum_{k=1}^{n(\Omega)} \sum_{\ell=1}^{n_k} \sum_{j=1}^{n-1} (z - \lambda_k)^{-j} \sum_{v=0}^{\mu_k - j} v^{\ell,k}_v w^{\ell,k}_v U_k
\]
\[
+ R(z).
\]

\( T^{-1} \) is a holomorphic function such that the eigenvalues are poles by means of Theorem 2. The order of the poles coincides with the maximal length of Jordan chains of the eigenvalues. The singular part of \( T^{-1} \) can be characterized in terms of (generalized) eigenvalues of \( T \) and \( T^{-1} \).

3. Generation of an eigenspace using contour integrals

Let \( U \in \mathbb{C}^{n \times L} \) be a matrix with \( L \leq n \). We compute the contour integrals for a simple closed curve \( \Gamma \subset \Omega \),

\[
S_k = \frac{1}{2\pi i} \int_{\Gamma} g_k(z) T(z)^{-1} U \, dz, \quad k \in \mathbb{N},
\]

where \( g_k \) are polynomial functions of the \( \kappa \)-th degree. Let \( T \) have no eigenvalues on the contour \( \Gamma \). In this case, we obtain the following relation by the residue theorem and (4).

\[
S_k = \sum_{k=1}^{n(\Gamma)} \sum_{\ell=1}^{n_k} \sum_{j=1}^{n-1} \sum_{v=0}^{\mu_k - j} g_k^\ell_j T(\lambda_k) \sum_{v=0}^{\mu_k - j} v^{\ell,k}_v w^{\ell,k}_v U_k
\]
\[
+ U_k (\Gamma).
\]

In numerical calculations, we evaluated the contour integrals using a numerical integral, refer to [9]. The approximations for \( S_k \) are represented as following,

\[
\hat{S}_k = \sum_{p=1}^{N} w_p g_p(z_p) T(z_p)^{-1} U_k
\]

where \( z_p, p = 1, \ldots, N \) are integral points and \( w_p, p = 1, \ldots, N \) are weights corresponding to \( z_p \). From (4) we obtain

\[
\hat{S}_k = \sum_{k=1}^{n(\Omega)} \sum_{\ell=1}^{n_k} \sum_{j=1}^{n-1} \sum_{v=0}^{\mu_k - j} f(j,k)(\lambda_k) v^{\ell,k}_v w^{\ell,k}_v U_k
\]
\[
+ B_{N,k},
\]

where

\[
f(j,k)(\lambda) = \sum_{p=1}^{N} w_p g_p(z_p) (z_p - \lambda)^j / (z_p - \lambda)
\]

Let \( \epsilon \) be sufficiently small. Then by the residue theorem, there exists \( N_0 \) such that

\[
N > N_0 \Rightarrow \| B_{N,k} \| < \epsilon.
\]

Hereinafter, we assume that we obtain a \( N \) that satisfies the following relation:

\[
B_{N,k} = O_{n,L}(\epsilon).
\]

Let \( \lambda_1, \ldots, \lambda_n(\Gamma) \) be all eigenvalues of \( T \) satisfy \( \| f(j,k) \| \geq \epsilon \). Then, we obtain

\[
\hat{S}_k = \sum_{k=1}^{n(\Omega)} \sum_{\ell=1}^{n_k} \sum_{j=1}^{n-1} \sum_{v=0}^{\mu_k - j} f(j,k)(\lambda_k) v^{\ell,k}_v w^{\ell,k}_v U_k
\]
\[
+ \hat{U}_k.
\]

Theorem 3  Let the assumption of the following relation be satisfied.

\[
\text{span}(\hat{S}) = \text{span}(\hat{V}),
\]

where

\[
\hat{S} = (\hat{S}_0, \ldots, \hat{S}_{K-1}) = (\hat{s}_1, \ldots, \hat{s}_{KL}), \quad K \in \mathbb{N}^+,
\]
\[
\hat{V} = \left( v^{\ell,k}_j, j = 0, \ldots, \mu_k - 1, \ell = 1, \ldots, n_k \right),
\]

\[
k = 1, \ldots, n(\Gamma).
\]

Suppose that \( \{ \lambda_1, \ldots, \lambda_n(\Gamma) \} \) contain all eigenvalues of \( T \) inside \( \Gamma \). Denote \( \hat{V} \) the approximate eigenspace cor-
responding to all eigenvalues of $T$ inside $\Gamma$; then for $V$, the following relation holds:

$$V \subseteq \text{span}(\hat{S}).$$ (9)

Note that, satisfying (8), it is sufficient to satisfy the following conditions.

(i) $\sum_{k=1}^{\mu} \hat{f}_{j,k}(\lambda)\hat{v}_{i,k}^j\hat{w}_{\ell,k}^j \hat{U}^j \neq 0$

(ii) $K(\text{rank } U) \geq \text{rank } \hat{V}$

Theorem 3 denotes that the subspace contains the eigenspace designate can be generated with contour integrals.

Now we review the SS-H method [6] for NEPs whose eigenvalues are simple.

Let two matrices $H_{KL}$, $H_{KL}^\zeta$ be

$$H_{KL} = (\hat{M}_{i,j+1})_{j=1}^K, \quad H_{KL}^\zeta = (\hat{M}_{i,j-1})_{j=1}^K,$$

where $\hat{M}_k = U^H\hat{S}_k$, $U \in \mathbb{C}^{n \times L}$, $k = 1, \ldots, 2K - 1$. We compute the numerical rank $\hat{m}$ of $H_{KL}$ by the singular value decomposition and set following matrices,

$$H_{m} = H_{KL}(1 : \hat{m}, 1 : \hat{m}), \quad H_{m}^\zeta = H_{KL}^\zeta(1 : \hat{m}, 1 : \hat{m}).$$

Then, we can obtain the approximate eigenvalues inside $\Gamma$ of $T$ by solving the following GEPs,

$$H_{m}^\zeta x = \hat{\lambda}H_{m}x,$$

and the approximate eigenvectors by computing

$$\hat{v} = \hat{S}(1 : \hat{m})x,$$

where $\hat{S}(1 : \hat{m}) = (\hat{s}_1 \ldots \hat{s}_\hat{m})$.

Additionally, we refer to another integral method proposed by Beyn in [7]. The method differs only slightly from the SS-H method in using the following matrices instead of $H_{KL}$ and $H_{KL}^\zeta$.

$$P_{KL} = (\hat{S}_{i,j+1})^K_{i,j=1}, \quad P_{KL}^\zeta = (\hat{S}_{i,j-1})^K_{i,j=1}.$$

However, Beyn’s method can also be applied to problems whose eigenvalues are not simple because of its theoretical approach, the Keldysh theorem. The SS-H method can be extended by the same idea.

In our research, we use the same theory for generating a subspace. Nevertheless, we present a different approach, a projection method, to obtain the eigenvalues from the subspace in the next section.

4. A projection method for nonlinear eigenvalue problems

From a mathematical point of view, the nonlinear Rayleigh–Ritz iterative method for solving NEPs is an extension of the Rayleigh–Ritz subspace projection technique for solving linear eigenvalue problems. The nonlinear Rayleigh–Ritz method may be summarized as follows.

Rayleigh–Ritz method

(a) Select a subspace $K$ that satisfies $V \subseteq K$.

(b) Compute approximate eigenpairs $(\hat{\lambda}, \hat{v}_0)$ with generalized eigenvectors $\hat{v}_1, \ldots, \hat{v}_{\mu-1}$ where $\hat{\lambda}$ inside $\Gamma$ satisfies the Galerkin condition:

$$\hat{v}_j \in K, \quad (T(\hat{\lambda})\hat{v}(\hat{\lambda}))^{(j)} \perp K,$$ (10)

where $j = 0, \ldots, \mu - 1$ and $\hat{v}(z) = \sum_{j=0}^{\mu-1}(z - \hat{\lambda})^j\hat{v}_j$ and $(T(\hat{\lambda})\hat{v}(\hat{\lambda}))^{(j)}$ is the $j$–the derivation.

(c) Select approximate eigenpairs that satisfy proper conditions.

At step (b), let $Q$ be an orthonormal basis of $K$; then $v_j = Qx_j$ for $m$-vector $x_j$, where $m$ is the dimension of the subspace $K$ (or rank $V$). Hence, Step (b) is equivalent to determining eigenpairs $(\hat{\lambda}, x_0)$ with generalized eigenvectors $x_1, \ldots, x_{\mu-1}$ of the reduced NEP

$$(T(\hat{\lambda})x(\hat{\lambda}))^{(j)} = 0,$$ (11)

for $j = 0, \ldots, \mu - 1$, where $\hat{T} = Q^HTQ : \Omega \subset \mathbb{C} \rightarrow \mathbb{C}^{m \times m}$ is a holomorphic matrix-valued function, $x(z) = \sum_{j=0}^{\mu-1}(z - \hat{\lambda})^jx_j$. The values $\hat{\lambda}$ referred to as Ritz values, and $x_0, \ldots, x_{\mu-1}$ are the corresponding Ritz vectors.

Note that the nonlinear Rayleigh–Ritz iterative method needs to iterate steps (a) and (b) until a proper subspace is obtained.

Consider now the SS-RR method, the solver for large-scale linear eigenvalue problems. By Theorem 3, coverage of the SS-RR method can be extended to NEPs using span($\hat{S}$) as a subspace $K$ at step (a). Furthermore, the SS-RR method requires no iterations to obtain $K$. Eventually, we select the approximate eigenpairs whose $\hat{\lambda}$ satisfies $\hat{\lambda} \in \Gamma$ at step (c).

We show the algorithm for the SS-RR method for NEPs below.

Algorithm(SS-RR)

Input: $U \in \mathbb{C}^{n \times L}$, $N$, $K$, $L$, $\delta$

Output: $\hat{\lambda}_k$, $\hat{v}_{j,k}^j$, $j = 0, \ldots, \mu$, $k = 1, \ldots, n(\Gamma)$

1. Set $z_p$, $w_p$, $p = 1, \ldots, N$

2. Compute $T(z)^{-1}U$, $p = 1, \ldots, N$

3. Compute $\hat{S}_k$, $k = 0, \ldots, K - 1$ by (5)

4. Construct $\hat{S} = (\hat{S}_0 \cdots \hat{S}_{K-1})$.

5. Construct $\hat{S}_m = \hat{S}(1 : \hat{m})$, where $\hat{m}$ is the numerical rank whose tolerance is $\delta$ of $\hat{S}$.

6. Construct the orthonormal basis $Q$ from $\hat{S}_m$

7. Form $\hat{T} = Q^HTQ$

8. Compute the eigenvalues $\hat{\lambda}_k$ inside $\Gamma$ and the corresponding Jordan chains $x_{j,k}^j$ of $\hat{T}$, where $j = 0, \ldots, \mu$, $k = 1, \ldots, n(\Gamma)$

9. Set $\hat{v}_{j,k}^j = Qx_{j,k}^j$, $j = 0, \ldots, \mu$, $k = 1, \ldots, n(\Gamma)$

According to [9], we can set $z_p$ and $w_p$ for any numerical integrals. The dimensionality of the generated subspace depends on the numerical rank of $\hat{S}$.
5. Numerical example

In this section, we confirm the validity of the proposed method using NEPs. The algorithm is implemented in MATLAB 7.12. We choose $U$ as a random matrix. We use $N$-point trapezoidal rule for numerical integrals. Let the functions $g_n(z)$ be $z^n$ where $n = 0, \ldots, K - 1$. In case of solving a polynomial eigenvalue problem (QEP), we use the MATLAB command `polyeig`, which is a linearization method for QEPs using ZQ algorithm [10], for projection matrix. In other cases, we use Beyn’s method for the projected matrix. The computed eigenvectors are normalized so that $\|v\|_2 = 1$. We use the residual $\|T(\lambda)v\|/\|T(\lambda)\|$ for the benchmark of the method. As a consequence of Theorem 3 and Galerkin iteration, the proposed method does not need fixed point iterations.

Other numerical methods for NEPs using contour integrals have already been proposed in [6, 7]. However, the SS-RR method use a different way, the Rayleigh–Ritz procedure, to obtain eigenvalues from a subspace. Through examples, the SS-RR method gives good numerical results. The choice of a large $N$ or $L$ results in smaller residuals for any methods using contour integrals. On the other hand, smaller values of $N$ and $L$ are preferable regarding computational cost. The SS-RR method allows for the use of larger $K$, leading to better residuals even for smaller $N$ and $L$ values.

Error analysis for the proposed method and the estimation of suitable parameters remain as topics for future research.

### References


