An eigenvalue problem for derogatory matrices

Tomohiro Suzuki\textsuperscript{a,}\textsuperscript{*}, Toshio Suzuki\textsuperscript{b}

\textsuperscript{a}Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Kofu 400-8511, Japan

\textsuperscript{b}Faculty of Education & Human Sciences, University of Yamanashi, Kofu 400-8510, Japan

Received 15 December 2004

Abstract

A matrix \(A\) is called derogatory if there is more than one Jordan submatrix associated with an eigenvalue \(\lambda\). In this paper, we are concerned with the eigenvalue problem of this type of matrices.

The singularities of the resolvent of \(A : R(z) = (A - zI)^{-1}\) are exactly the eigenvalues of \(A\). Let us consider the Laurent series of \(R\) expanded at \(z = \lambda\) and denote its coefficients \(c_k (-\infty \leq k \leq \infty)\). \(D := c_{-2}\) is the nilpotent operator, that is, there exists the order \(l\) of \(\lambda\) such that \(D^l := c_{-l-1} = 0 (l \geq 1)\). Additionally, for an arbitrary vector \(z\), \(D^{l-1}z\) is an eigenvector of \(\lambda\). Then \(\lambda\) is computed from the corresponding eigenvector \(D^{l-1}z\). In order to estimate the integral representation of \(D^kz\), we apply the trapezoidal rule on the circle enclosing \(\lambda\) but excluding other eigenvalues of \(A\).

It is our result that, so far as related linear equations are solved with necessary precision, the eigenvalues of derogatory matrices can be computed numerically as exactly as we want and so are corresponding (generalized) eigenvectors, too.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Multiple eigenvalue; Jordan canonical form; Generalized eigenvector

1. Introduction

It is well-known that a derogatory and defective matrix \(A\) can be transformed into the Jordan canonical form, and the computation of it is unstable. There are two basic methods [2,3] for determining the Jordan decomposition [1]. There are, however, two difficulties in these methods: the computation of a multiple eigenvalue and the computation of the rank of the matrix (the latter is ill-posed).

Our purpose is to introduce a new computing method for multiple eigenvalue and its eigenvectors of derogatory matrices under the double precision floating point arithmetic. In our method, the eigenvector corresponding to the sought eigenvalue is computed from the solutions of linear equations and the eigenvalue is computed from the obtained eigenvector.

2. Matrix resolvent

The matrix-valued function \(R(z) := (A - zI)^{-1}\) is called the resolvent of the matrix \(A\), where \(I\) is an unit matrix. This function plays an important role in our method.

* Corresponding author.

E-mail address: stomo@yamanashi.ac.jp (T. Suzuki).

0377-0427/ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2005.08.044
Suppose that \( A \) has \( s \) distinct eigenvalues \( \lambda_j \) \((j = 1, \ldots, s)\) and the multiplicity of \( \lambda_i \) is \( m_i \geq 1 \). Let us consider \( D_i^k z \) defined by the following integrals for an arbitrary vector \( z (\neq 0) \).

\[
D_i^k z := -\frac{1}{2\pi i} \int_\Gamma (\zeta - \lambda_i)^k R(\zeta) z d\zeta \quad (k = 0, 1, 2, \ldots),
\]

where the path of integration \( \Gamma \) is a positively oriented closed-curve enclosing \( \lambda_i \) but excluding the other eigenvalues of \( A \). Then the number \( l_i > 1 \) such that

\[
(A - \lambda_i I)D_i^{k-1} z = D_i^k z \neq 0 \quad (k = 1, \ldots, l_i - 1)
\]

for proper \( z \), (2)

\[
(A - \lambda_i I)D_i^{l_i-1} z = D_i^{l_i} z = 0 \quad \text{for all } z,
\]

is called the order of \( \lambda_i \) and \( D_i^{l_i-1} z \) is an eigenvector corresponding to \( \lambda_i \), in addition, \( \{D_i^0 z, D_i^1 z, \ldots, D_i^{l_i-2} z\} \) are generalized eigenvectors (principal vectors) corresponding to \( \lambda_i \) (cf. [4]).

The relation between the order \( l_i \) and the Jordan block associated with \( \lambda_i \) is that \( l_i \) is just equal to the maximum size of the Jordan submatrix (Fig. 1). In other word, we can find a part of the structure of the Jordan block by figuring out \( l_i \).

3. Numerical integration

In this section, the approximate method for the computation of the contour integrals (1) is described. Let the path of integration \( \Gamma \) be the circle with center \( \lambda \) and radius \( r \) (Fig. 2). We are going to use \( \lambda \) as an approximation of \( \lambda_i \). Dividing the circumference into \( M \) equidistant points \( \mu_j := \lambda + r e^{i\theta_j}, \theta_j := 2\pi j/M \) \((j = 0, \ldots, M - 1)\), we apply the \( M \) points trapezoidal rule for the contour integration. As the result, the numerical integration of (1) is represented as the following expression:

\[
D_i^k z := -\frac{r}{M} \sum_{j=0}^{M-1} (\mu_j - \lambda_i)^k e^{i\theta_j} W_j \quad (k = 0, 1, 2, \ldots),
\]

(4)
where $W_j (j = 0, \ldots, M - 1)$ are the solutions of $M$ linear equations: $W_j := (A - \mu_j I)^{-1} z$. It is important to note that the approximate eigenvector $D^k_z$ is obtained from the linear combination of $W_j$, that is, the accuracy of $D^k_z$ depends on the one of the solution of linear equations.

Let $\lambda'$ be a point near to $\lambda_i$. We consider the following expression which is an approximation of corresponding (generalized) eigenvector.

$$D^k_z := \frac{-r}{M} \sum_{j=0}^{M-1} (\mu_j - \lambda')^k e^{i0_j} W_j \quad (k = 0, 1, 2, \ldots).$$

(5)

We should note that $D^k_z$ can be derived cheaply because it is computed without updating $W_j$ at $\lambda'$, that is, $W_j = (A - (\lambda + re^{i0_j}) I)^{-1} z (j = 0, \ldots, M - 1)$ in (5). In fact we need to solve $M$ linear equations only once at the beginning of an iterative process.

Let $\rho$ be the minimum distance from $\lambda_i$ to the other eigenvalues. Then the accuracy of approximate eigenvectors for $\lambda'$ close enough to $\lambda_i$ are of the following order [6]:

$$\|D^k_z - D^k_{z'}\| = O((r/\rho)^{M-k}).$$

(6)

Let us denote $v' := D^{l-1}_z$ and $\tilde{\lambda}' := v'^* A v'/v'^* v'$, then the following relation holds [6].

$$\tilde{\lambda}' = \lambda_i + (l_i - 1)(\lambda_i - \lambda') + O(|\lambda_i - \lambda'|^2).$$

(7)

This implies that if $l_i$ is resolved, we can apply $(l_i - 1)\lambda_i + \tilde{\lambda}'$ as the 2nd order approximate eigenvalue via $\tilde{\lambda}'$: Rayleigh quotient of $A$ (but $A$ is not Hermitian) for the approximate eigenvector $v'$.

4. Algorithm

The iterative algorithm to obtain the approximate eigenpair $\{\lambda', v'\}$ is expressed in Algorithm 1. We assume that rough approximate eigenvalues are computed in advance. In order to obtain other eigenvectors corresponding to $\lambda'$ the approximate eigenvectors produced from different initial vectors should be orthogonalized.

Algorithm 1.

```
\begin{align*}
\lambda & \leftarrow \text{initial guess} \\
Set M, r and z \\
\text{for } j = 0 \text{ to } M - 1 \text{ do} \\
\theta_j & \leftarrow 2\pi j / M \\
\mu_j & \leftarrow \lambda + re^{i0_j} \\
W_j & \leftarrow (A - \mu_j I)^{-1} z \\
\text{end for} \\
\lambda' & \leftarrow \lambda \\
\text{repeat} \\
\text{for } k = 0 \text{ to sufficiently large number do} \\
D^k_z & \leftarrow (-r/M) \sum_{j=0}^{M-1} (\mu_j - \lambda')^k e^{i0_j} W_j \\
\text{end for} \\
\text{Resolve the order } l \text{ of the eigenvalue} \\
v' & \leftarrow D^{l-1}_z, \quad \tilde{\lambda}' = v'^* A v'/v'^* v' \\
\lambda' & \leftarrow (l_i - 1)\lambda_i + \tilde{\lambda}'/l_i \\
\text{until some condition(s)} \\
\{\lambda', v'\} \text{ is the approximate eigenpair}
\end{align*}
```

In this algorithm, the statements indicated in italics have to be explained. From pre-computed approximations, we can choose the circle with center $\lambda$ and radius $r$ which encloses the sought eigenvalue but excludes the other eigenvalues.
of $A$. According to (6), $M$ is decided in association with a working precision. The vector $z$ is arbitrary except for its size. In our numerical experiment, its elements are random numbers in $[-1,1]$.

We can guess roughly the (algebraic) multiplicity of the sought eigenvalue from pre-computed approximations. We should execute the computation of $k+1$’s vectors $D^k_z$ up to at least this multiplicity until the order $l$ of $\lambda'$ is decided by some method.

One of the most important problems of our method is “how to resolve the order of the eigenvalue?”. According to the chain relations (2) and (3), $D^k_z$ is a zero vector, in a theoretical sense. Therefore we should have to find $k$ such that $\|D^k_z\| \approx 0$ ($\| \cdot \|$ indicates Euclidean norm). There are some practical methods to check it proposed by authors [5]. As we obtain a better approximation of $\lambda_i$, we can see clearly the boundary number $l_i$ by observing the sequence of $\{\|D^k_z\|, k = 0, 1, \ldots\}$.

5. Numerical example

In this section, we show some numerical examples which express the behaviour of our method well. All the programs below are performed by MATLAB (Ver. 7).

Let us consider the following matrix $A_1$. To represent the elements exactly in floating point number, $A_1$ is made by multiplying the matrix of [1, Example 3] by 100.


The eigenvalues of $A_1$ are $\{\lambda_1, \lambda_2, \lambda_3\} = \{-100, -200, 700\}$, where $\lambda_1$ and $\lambda_2$ are simple eigenvalues and $\lambda_3$ is the eigenvalue of multiplicity $m_3 = 6$. Using the MATLAB function (eig($A_1$)), we can obtain the approximations: $\{-200.0011, -100.0011, 700.002 + 0.00004i, 700.002 - 0.4e - 5i, 700.001 + 0.4e - 5i, 700.001 - 0.4e - 5i, 699.998, 699.996\}$, where $0_k$ means $k$ consecutive $0$s. From this approximations, we set the initial guess $\lambda = 700.02$ and the parameters $M = 16$ and $r = 50.0$. In order to reduce the condition number of $(A_1 - \mu I)$, the radius $r$ of the path of integration was enlarged. Though $M$ might seem to be too small as the number of partitions of the path of integration, this is applicable since $(r/\rho)^M \approx 5.42e - 20$. Then the maximum value of the condition number of $(A_1 - \mu I)$ is $3.74e + 3$.

The $l$-finding criterion used in this experiment is that find $k$ such that $\|D^k_z\|^2/\|D^k_z\|^2 < \epsilon = 1.0e - 3$. After five iterations, we can obtain the following magnitude of norm of approximate eigenvectors:

$$\|D^0_z\| = 1.76,$$
$$\|D^1_z\| = 99.2,$$
$$\|D^2_z\| = 64.1e + 4,$$
$$\|D^3_z\| = 4.28e - 7,$$
$$\|D^4_z\| = 7.15e - 5,$$
$$\|D^5_z\| = 5.81e - 3.$$ 

Note the wide gap between $\|D^2_z\|$ and $\|D^3_z\|$. We can resolve the order of $\lambda_3$ is $l_3 = 3$, and therefore, there is at least one eigenvector other than $D^2_z$. The approximate eigenvalue is $\lambda' = 699.9911 - 3.02e - 13$. 


If the number of eigenvectors corresponding to \( \lambda \) is greater than 1, the sequences \( \{ \mathcal{D}_{z_j}^k \}, \ j = 1, 2, \ldots \) derived from independent initial vectors \( z_j \) yield the same number of linearly independent eigenvectors. In order to resolve the structure of the remaining Jordan block(s) of size three associated with \( \lambda_3 \), we should orthogonalize \( \{ \mathcal{D}_{z_j}^k \}, \ j = 2, 3, \ldots \) for each \( k \). In this experience, there are two linearly independent vectors of order three. As the result we reason the Jordan submatrix corresponding to \( \lambda_3 \) consists of two blocks of size three.

As a second test matrix we take the following matrix:

\[
A_2 = \begin{bmatrix}
0 & I_4 \\
-B_2 & -B_1
\end{bmatrix},
\]

\[
B_1 = \begin{bmatrix}
3\alpha & -(1 + \alpha^2 + 2\beta^2) & \alpha(1 + 2\beta^2) & -\beta^2(\alpha^2 + \beta^2) \\
2 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 \\
0 & 0 & 2 & 0
\end{bmatrix},
\]

\[
B_2 = \begin{bmatrix}
-1 + 2\alpha^2 & \alpha - \alpha(\alpha^2 + 2\beta^2) & 2\alpha^2\beta^2 & -\alpha\beta^2(\alpha^2 + \beta^2) \\
2\alpha & -(\alpha^2 + 2\beta^2) & 2\alpha^2\beta^2 & -\beta^2(\alpha^2 + \beta^2) \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0
\end{bmatrix},
\]

and \( \alpha = 10^{-4}, \beta = 1 + \alpha \). \( I_k \) is an unit matrix of size \( k \). \( A_2 \) is constructed to resemble the matrix of a very ill-conditioned stability problem [3]. The eigenvalues of \( A_2 \), \( \{ \pm i, \pm (1 + \alpha)i, -\alpha \pm (1 + \alpha)i, 0, -\alpha \} \), are all simple, but are grouped into three clusters.

We cannot differentiate very close eigenvalues from a multiple one using the floating point arithmetic. Similarly, a multiple eigenvalue seems to be close ones under the influence of round off errors. In [3], the authors treated the close eigenvalues \( \{ i, (1 + \alpha)i, -\alpha + (1 + \alpha)i \} \) as a multiple one and discussed the structure of its Jordan block. The conclusion is that the mean value of the three eigenvalues is treated as the multiple one and its Jordan submatrix consists of two blocks of size two and one. In contrast, the Jordan structure of one block of size three \((l = 3, \text{geometric multiplicity is one})\) is said to be not well determined.

For this test matrix, we set the initial guess \( \hat{\lambda} = 1.02i \) and the parameters \( M = 24, r = 0.125 \) so that the path of integration encloses the three eigenvalues around \( i \). Then the maximum value of the condition number of \( (A_2 - \mu_j I) \) is 2.55e+03.

After four iterations, it almost converged to yield the following magnitude of norms of approximate eigenvectors:

\[
\| \mathcal{D}_{z_1}^0 \| = 6.01,
\]

\[
\| \mathcal{D}_{z_1}^1 \| = 4.23,
\]

\[
\| \mathcal{D}_{z_1}^2 \| = 2.32e - 8,
\]

\[
\| \mathcal{D}_{z_1}^3 \| = 4.59e - 9,
\]

\[
\| \mathcal{D}_{z_1}^4 \| = 1.56e - 12,
\]

\[
\| \mathcal{D}_{z_1}^5 \| = 1.55e - 16.
\]

The norms above are obtained by treating the clustered eigenvalues as the multiple one. We can resolve as a temporary result that the order of \( \lambda_3 \) is \( l = 2 \) using the same criterion of the first example. The approximate eigenvalue is \( \hat{\lambda}' = -1.39e - 5 + 1.000040i \). The invariant subspace corresponding to the three simple eigenvalues seems to be that of the multiple one with the Jordan submatrix of two blocks of size two and one.

Remark 1. This is formally the same result with that of [3]. But according to our theory, \( \| \mathcal{D}_{z_1}^2 \| \) is not small enough to be a result for a multiple eigenvalue. We should conclude that these three eigenvalues are not a multiple one but their eigenvectors only seem to be those of a defective and derogatory Jordan block. Because such a problem is not aimed by this paper, precise discussions on it will be left for the future.
6. Conclusion

We enumerate important items to conclude this paper.

- The (generalized) eigenvectors are expressed by the contour integral of the resolvent \((A - zI)^{-1}\) of the original matrix.
- The eigenvectors are approximated by the linear combination of the solutions of linear equations: \(W_j = (A - \mu_j I)^{-1}z\) each of which is not ill-posed.
- Determining the order of the eigenvalue by some method, we can obtain the eigenvalue from the corresponding eigenvector. We should propose a more effective method to determine the order of the eigenvalue.
- The structure of more than one Jordan submatrix corresponding to an eigenvalue can be determined by choosing linearly independent sequence of \(\mathbb{D}_{z_p}^{k}\)'s among these from independent initial vector \(z_p\)'s.

References