Numerical computation of the Complex Eigenpair of a Large Sparse Matrix

R. Ọ. Akinọla and A. Spence

23rd March 2016

- Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a large, sparse, nonsymmetric matrix.
- $\mathbf{B} \in \mathbb{R}^{n\times n}$, Symmetric Positive Definite (SPD) matrix.
- We consider the problem of computing the eigenpair $(\mathbf{z}, \lambda)$ from the Generalised Eigenvalue Problem

$$\mathbf{A}\mathbf{z} = \lambda \mathbf{B}\mathbf{z}, \quad \mathbf{z} \in \mathbb{C}^n, \;\; \mathbf{z} \neq \mathbf{0}, \tag{1}$$

  where $\lambda \in \mathbb{C}$ is the eigenvalue.
- We assume that the eigenpair of interest $(\mathbf{z}, \lambda)$ is algebraically simple. The left e-vector $\boldsymbol{\psi}$ is $\ni$

$$\psi^H \mathbf{B}\mathbf{z} \neq 0. \tag{2}$$

By adding the normalisation

$$\mathbf{z}^H \mathbf{B}\mathbf{z} = 1, \tag{3}$$

to (1) & $\mathbf{v} = [\mathbf{z}^T, \lambda]$,

$$\mathbf{F}(\mathbf{v}) = \begin{bmatrix} (\mathbf{A} - \lambda\mathbf{B})\mathbf{z} \\ -\frac{1}{2}\mathbf{z}^H\mathbf{B}\mathbf{z} + \frac{1}{2} \end{bmatrix} = \mathbf{0}. \tag{4}$$

Lemma

$\bar{\mathbf{z}}$ in $\mathbf{z}^H\mathbf{B}\mathbf{z} = \bar{\mathbf{z}}^T\mathbf{B}\mathbf{z}$ is not differentiable.

Proof.

If $z = x + iy$, $\bar{z} = x - iy$, then the Cauchy-Riemann equations are not satisfied. $u(x, y) = x$, $v(x, y) = -y$, then $u_x(x, y) = 1$ and $v_y(x, y) = -1$, whereas the Cauchy-Riemann equations require that $u_x(x, y) = v_y(x, y)$. □

- Newton's method cannot be applied to (4).

- For a **real** eigenpair $(\mathbf{z}, \lambda)$, Newton's method for solving (4) involves the solution of the $(n+1)$ square linear systems

$$\begin{bmatrix} \mathbf{A} - \lambda^{(k)}\mathbf{B} & -\mathbf{B}\mathbf{z}^{(k)} \\ -(\mathbf{B}\mathbf{z}^{(k)})^T & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{z}^{(k)} \\ \Delta\lambda^{(k)} \end{bmatrix} = \begin{bmatrix} -(\mathbf{A} - \lambda^{(k)}\mathbf{B})\mathbf{z}^{(k)} \\ \frac{1}{2}\mathbf{z}^{(k)T}\mathbf{B}\mathbf{z}^{(k)} - \frac{1}{2} \end{bmatrix}, \tag{5}$$

  for real unknowns $\Delta\mathbf{v}^{(k)} = [\Delta\mathbf{z}^{(k)T}, \Delta\lambda^{(k)}]$.
- Update $\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \Delta\mathbf{v}^{(k)}$,for $k = 0, 1, 2, \cdots$.

- Ruhe (1973) added the normalization $\mathbf{c}^H\mathbf{z} = 1$, $\mathbf{c}$ is a fixed complex vector instead of $\mathbf{z}^H\mathbf{B}\mathbf{z} = 1$.
- Parlett and Saad (1987) studied inverse iteration with a fixed complex shift. Numerical examples showed linear convergence to the eigenvalue closest to the shift.
- Tisseur (2001) used the normalization $\mathbf{e}_s^T\mathbf{z} = 1$
- Both Ruhe's and Tisseur's normalizations are differentiable, but we use the natural normalization for the eigenvector.

- $\mathbf{z} = \mathbf{z}_1 + i\mathbf{z}_2$ and $\lambda = \alpha + i\beta$. (1), (4) becomes

$$\mathbf{F}(\mathbf{v}) = \begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_1 + \beta\mathbf{B}\mathbf{z}_2 \\ -\beta\mathbf{B}\mathbf{z}_1 + (\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_2 \\ -\frac{1}{2}(\mathbf{z}_1^T\mathbf{B}\mathbf{z}_1 + \mathbf{z}_2^T\mathbf{B}\mathbf{z}_2) + \frac{1}{2} \end{bmatrix} = \mathbf{0}. \tag{6}$$

The Jacobian, with $\mathbf{v} = [\mathbf{z}_1, \mathbf{z}_2, \alpha, \beta]^T$ is

$$\mathbf{F}_{\mathbf{v}}(\mathbf{v}) = \begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B}) & \beta\mathbf{B} & -\mathbf{B}\mathbf{z}_1 & \mathbf{B}\mathbf{z}_2 \\ -\beta\mathbf{B} & (\mathbf{A} - \alpha\mathbf{B}) & -\mathbf{B}\mathbf{z}_2 & -\mathbf{B}\mathbf{z}_1 \\ -(\mathbf{B}\mathbf{z}_1)^T & -(\mathbf{B}\mathbf{z}_2)^T & 0 & 0 \end{bmatrix}. \tag{7}$$

Let

$$\mathbf{w} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix}, \qquad \mathbf{w}_1 = \begin{bmatrix} \mathbf{z}_2 \\ -\mathbf{z}_1 \end{bmatrix}. \tag{8}$$

We define the real $2n$ by $2n$ matrix $\mathbf{M}$ as

$$\mathbf{M} = \begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A} - \alpha\mathbf{B}) \end{bmatrix}. \tag{9}$$

Also, we form the $2n$ by 2 real matrix

$$\mathbf{N} = \begin{bmatrix} -\mathbf{B}\mathbf{z}_1 & \mathbf{B}\mathbf{z}_2 \\ -\mathbf{B}\mathbf{z}_2 & -\mathbf{B}\mathbf{z}_1 \end{bmatrix} = \begin{bmatrix} -\mathbf{B}_2\mathbf{w} & \mathbf{B}_2\mathbf{w}_1 \end{bmatrix}, \tag{10}$$

where $\mathbf{B}_2 = \begin{bmatrix} \mathbf{B} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \end{bmatrix}$.

Note that because at the root,

$$\begin{bmatrix} (\mathbf{A}-\alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A}-\alpha\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} = \begin{bmatrix} (\mathbf{A}-\alpha\mathbf{B})\mathbf{z}_1 + \beta\mathbf{B}\mathbf{z}_2 \\ (\mathbf{A}-\alpha\mathbf{B})\mathbf{z}_2 - \beta\mathbf{B}\mathbf{z}_1 \end{bmatrix} = \mathbf{0}.$$

In the same vein, we find

$$\begin{bmatrix} (\mathbf{A}-\alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A}-\alpha\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{z}_2 \\ -\mathbf{z}_1 \end{bmatrix} = \mathbf{0}.$$

The Jacobian (7) can be rewritten in partitioned form

$$\mathbf{F}_\mathbf{v} = \begin{bmatrix} \mathbf{M} & -\mathbf{B}_2\mathbf{w} & \mathbf{B}_2\mathbf{w}_1 \\ -(\mathbf{B}_2\mathbf{w})^T & 0 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \mathbf{N} \\ -(\mathbf{B}_2\mathbf{w})^T & \mathbf{0}^T \end{bmatrix}. \tag{11}$$

For all $\psi \in \mathcal{N}(\mathbf{A} - \lambda\mathbf{B})^H \backslash \{\mathbf{0}\}$, we define $\psi = \psi_1 + i\psi_2$, where $\psi_1, \psi_2 \in \mathbb{R}^n$,

$$\begin{aligned}\psi^H(\mathbf{A} - \lambda\mathbf{B}) &= (\psi_1^T - i\psi_2^T)[(\mathbf{A} - \alpha\mathbf{B}) - i\beta\mathbf{B}] \\ &= \psi_1^T(\mathbf{A} - \alpha\mathbf{B}) - \beta\psi_2^T\mathbf{B} \\ &\quad - i[\beta\psi_1^T\mathbf{B} + \psi_2^T(\mathbf{A} - \alpha\mathbf{B})] = \mathbf{0}^T.\end{aligned}$$

$$[\psi_1^T \quad \psi_2^T]\mathbf{M} = [\psi_1^T \quad \psi_2^T]\begin{bmatrix}(\mathbf{A} - \alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A} - \alpha\mathbf{B})\end{bmatrix} = \mathbf{0}^T$$

$[\psi_1^T, \ \psi_2^T]$ is a left nullvector of $\mathbf{M}$. Similarly,

$$[\psi_2^T \quad -\psi_1^T]\mathbf{M} = [\psi_2^T \quad -\psi_1^T]\begin{bmatrix}(\mathbf{A} - \alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A} - \alpha\mathbf{B})\end{bmatrix} = \mathbf{0}^T.$$

$$\mathbf{C} = \begin{bmatrix} \psi_1 & \psi_2 \\ \psi_2 & -\psi_1 \end{bmatrix}. \tag{12}$$

Now, observe that the condition (2), implies

$$\psi^H \mathbf{Bz} = [\psi_1^T \mathbf{Bz}_1 + \psi_2^T \mathbf{Bz}_2] + i[\psi_1^T \mathbf{Bz}_2 - \psi_2^T \mathbf{Bz}_1] \neq 0.$$

Theorem

Assume that the eigenpair $(\mathbf{z}, \lambda)$ of the pencil $(\mathbf{A}, \mathbf{B})$ is algebraically simple. If $\mathbf{z}_1$ and $\mathbf{z}_2$ are nonzero vectors, then $\phi = \{\tau[\mathbf{z}_2^T, -\mathbf{z}_1^T, 0, 0], \tau \in \mathbb{R}\}$ is the eigenspace corresponding to the zero eigenvalue of $\mathbf{F_v}(\mathbf{v})$ at the root.

Proof

Post-multiply $\mathbf{F_v}(\mathbf{v})$ by the unknown nonzero vector $\boldsymbol{\phi} = [\mathbf{p}', \mathbf{q}']^T$ and $\mathbf{H} = \mathbf{C}^T\mathbf{N}$

$$\begin{bmatrix} \mathbf{M} & \mathbf{N} \\ -(\mathbf{B}_2\mathbf{w})^T & \mathbf{0}^T \end{bmatrix} \begin{bmatrix} \mathbf{p}' \\ \mathbf{q}' \end{bmatrix} = \mathbf{0}.$$

$$\mathbf{M}\mathbf{p}' + \mathbf{N}\mathbf{q}' = \mathbf{0} \tag{13}$$

$$\mathbf{w}^T\mathbf{B}_2\mathbf{p}' = 0. \tag{14}$$

$$\begin{aligned} \mathbf{H} &= \begin{bmatrix} \psi_1^T & \psi_2^T \\ \psi_2^T & -\psi_1^T \end{bmatrix} \begin{bmatrix} -\mathbf{B}\mathbf{z}_1 & \mathbf{B}\mathbf{z}_2 \\ -\mathbf{B}\mathbf{z}_2 & -\mathbf{B}\mathbf{z}_1 \end{bmatrix} \\ &= \begin{bmatrix} -(\psi_1^T\mathbf{B}\mathbf{z}_1 + \psi_2^T\mathbf{B}\mathbf{z}_2) & \psi_1^T\mathbf{B}\mathbf{z}_2 - \psi_2^T\mathbf{B}\mathbf{z}_1 \\ \psi_1^T\mathbf{B}\mathbf{z}_2 - \psi_2^T\mathbf{B}\mathbf{z}_1 & (\psi_1^T\mathbf{B}\mathbf{z}_1 + \psi_2^T\mathbf{B}\mathbf{z}_2) \end{bmatrix}. \end{aligned}$$

$$\mathbf{C}^T\mathbf{M}\mathbf{p}' + \mathbf{C}^T\mathbf{N}\mathbf{q}' = \mathbf{0}. \tag{15}$$

But, $\mathbf{C}^T\mathbf{M} = \mathbf{0}^T$. Consequently, $\mathbf{C}^T\mathbf{N}\mathbf{q}' = \mathbf{0}$, or

$$\begin{aligned}
\mathbf{H}\mathbf{q}' &= \mathbf{C}^T\mathbf{N}\mathbf{q}' \\
&= \begin{bmatrix} -(\psi_1^T\mathbf{B}\mathbf{z}_1 + \psi_2^T\mathbf{B}\mathbf{z}_2) & \psi_1^T\mathbf{B}\mathbf{z}_2 - \psi_2^T\mathbf{B}\mathbf{z}_1 \\ \psi_1^T\mathbf{B}\mathbf{z}_2 - \psi_2^T\mathbf{B}\mathbf{z}_1 & (\psi_1^T\mathbf{B}\mathbf{z}_1 + \psi_2^T\mathbf{B}\mathbf{z}_2) \end{bmatrix} \mathbf{q}' \\
&= \mathbf{0}.
\end{aligned}$$

Now, $\det \mathbf{H} \neq 0$. $\mathbf{H}$ is nonsingular. Thus, $\mathbf{q}' = \mathbf{0}$. Equation (13) now becomes $\mathbf{M}\mathbf{p}' = \mathbf{0}$, meaning that $\mathbf{p}' \in \mathcal{N}(\mathbf{M})$, $\mathbf{p}' = \mu\mathbf{w} + \tau\mathbf{w}_1$. From (14),

$$0 = \mathbf{w}^T\mathbf{B}_2\mathbf{p}' = \mu\mathbf{w}^T\mathbf{B}_2\mathbf{w} + \tau\mathbf{w}^T\mathbf{B}_2\mathbf{w}_1.$$

$\mathbf{w}^T\mathbf{B}_2\mathbf{w}_1 = 0$ and $\mathbf{w}^T\mathbf{B}_2\mathbf{w} \neq 0$, $\mu = 0$ and so $\mathbf{p}' = \tau\mathbf{w}_1$. Hence, for all

$\tau \in \mathbb{R}\backslash\{0\}$, $\mathbf{p}' = [\tau\mathbf{z}_2, -\tau\mathbf{z}_1]^T \in \mathcal{N}(\mathbf{M})$ also satisfies equation (14). $\boldsymbol{\phi} = \tau[\mathbf{z}_2, -\mathbf{z}_1, 0, 0]^T$ as the only nonzero nullvector of $\mathbf{F}_\mathbf{v}(\mathbf{v})$.

Corollary

If the eigenpair $(\mathbf{z}, \lambda)$ of $(\mathbf{A}, \mathbf{B})$ is algebraically simple, then the Jacobian $\mathbf{F}_\mathbf{v}(\mathbf{v})$ in (11) is of full rank at the root.

Proof.

The theorem above guarantees the existence of a single nonzero nullvector of $\mathbf{F}_\mathbf{v}(\mathbf{v})$ at the root, then $\operatorname{rank}\left(\mathbf{F}_\mathbf{v}(\mathbf{v})\right) = 2n + 1$. Therefore, the Jacobian (7) is of full rank at the root. (Dimension theorem). □

Eigenpair Computation using Gauss-Newton's method

- $\mathbf{A}$, $\mathbf{B}$, $\mathbf{v}^{(0)} = [\mathbf{z}_1^{(0)}, \mathbf{z}_2^{(0)}, \alpha^{(0)}, \beta^{(0)}]^T$, $k_{\max}$ and tol.
- for $k = 0, 1, 2, \ldots,$ until convergence
  - Find the reduced QR factorisation of $\mathbf{F}_{\mathbf{v}}(\mathbf{v}^{(k)})^T = \mathbf{QR}$.
  - Solve $\mathbf{R}^T\mathbf{g}^{(k)} = -\mathbf{F}(\mathbf{v}^{(k)})$ for $\mathbf{g}^{(k)}$ in (7).
  - Compute $\Delta\mathbf{v}^{(k)} = \mathbf{Qg}^{(k)}$ for $\Delta\mathbf{v}^{(k)}$ using (6).
  - Update $\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \Delta\mathbf{v}^{(k)}$.
  - $\mathbf{v}^{(k_{\max})}$.

The stopping condition for the algorithm above is

$$\|\Delta\mathbf{v}^{(k)}\| \leq tol.$$

Consider the 200 by 200 matrix $\mathbf{A}$ `bwm200.mtx` from the matrix market library. It is the discretised Jacobian of the Brusselator wave model for a chemical reaction. The resulting eigenvalue problem with $\mathbf{B} = \mathbf{I}$ was also studied in Parlett & Saad and we are interested in finding the rightmost eigenvalue of $\mathbf{A}$ which is closest to the imaginary axis and its corresponding eigenvector.

$\alpha^{(0)} = 0.0$, $\beta^{(0)} = 2.5$ in line with Parlett & Saad and took $\mathbf{z}_1^{(0)} = \frac{\mathbf{1}}{2\|\mathbf{1}\|}$ and $\mathbf{z}_2^{(0)} = \frac{\mathbf{1}\sqrt{3}}{2\|\mathbf{1}\|}$, where **1** is the vector of all ones.

k	$\alpha^{(k)}$	$\beta^{(k)}$	$\Vert\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)}\Vert$	$\Vert\boldsymbol{\lambda}^{(k+1)} - \boldsymbol{\lambda}^{(k)}\Vert$	$\Vert\Delta\mathbf{v}^{(k)}\Vert$	$\Vert\mathbf{F}(\mathbf{v}^{(k)})\Vert$
0	0.00	2.50	3.8e+00	7.8e-01	3.9e+00	3.6e+01
1	2.34e-1	1.75	1.8e+00	2.2e-01	1.8e+00	7.8e+00
2	1.18e-1	1.94	8.1e-01	1.4e-01	8.2e-01	1.7e+00
3	4.47e-2	2.06	2.5e-01	7.0e-02	2.6e-01	3.4e-01
4	8.82e-3	2.12	3.1e-02	1.7e-02	3.5e-02	3.7e-02
5	2.48e-4	2.13	4.8e-04	5.2e-04	7.1e-04	7.1e-04
6	1.80e-5	2.13	1.2e-07	2.5e-07	2.8e-07	2.8e-07
7	1.81e-5	2.13	2.1e-14	2.9e-14	3.6e-14	6.0e-14

Newton's method in complex arithmetic

- $\mathbf{A}, \mathbf{v}^{(0)} = [\mathbf{z}_1^{(0)}, \mathbf{z}_2^{(0)}, \alpha^{(0)}, \beta^{(0)}]^T, k_{\max}$ & tol.
- For $k = 0, 1, 2, \ldots,$ until convergence
- Compute the LU factorisation of

$$\begin{bmatrix} \mathbf{A} - \lambda^{(k)}\mathbf{I} & -\mathbf{z}^{(k)} \\ -(\mathbf{z}^{(k)})^H & 0 \end{bmatrix}.$$

- $$\mathbf{d}^{(k)} = -\begin{bmatrix} (\mathbf{A} - \lambda^{(k)}\mathbf{I})\mathbf{z}^{(k)} \\ -\frac{1}{2}\mathbf{z}^{(k)H}\mathbf{z}^{(k)} + \frac{1}{2} \end{bmatrix}.$$
- Solve $L\mathbf{y}^{(k)} = \mathbf{d}^{(k)}$ for $\mathbf{y}^{(k)}$.
- Solve $U\Delta\mathbf{v}^{(k)} = \mathbf{y}^{(k)}$ for $\Delta\mathbf{v}^{(k)}$.
- Update $\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \Delta\mathbf{v}^{(k)}$.

Newton's method in Complex Arithmetic

k	$\alpha^{(k)} + i\beta^{(k)}$	$\|\mathbf{z}^{(k+1)} - \mathbf{z}^{(k)}\|$	$\|\lambda^{(k+1)} - \lambda^{(k)}\|$	$\|\Delta\mathbf{v}^{(k)}\|$	$\|\mathbf{F}(\mathbf{v}^{(k)})\|$
0	0.00+2.50i	3.8e+00	7.8e-01	3.9e+00	3.6e+01
1	2.34e-1+1.75i	1.8e+00	2.2e-01	1.8e+00	7.8e+00
2	1.18e-1+1.94i	8.1e-01	1.4e-01	8.2e-01	1.7e+00
3	4.47e-2+2.06i	2.5e-01	7.0e-02	2.6e-01	3.4e-01
4	8.82e-3+2.12i	3.1e-02	1.7e-02	3.5e-02	3.7e-02
5	2.48e-4+2.13i	4.8e-04	5.2e-04	7.1e-04	7.1e-04
6	1.80e-5+2.13i	1.2e-07	2.5e-07	2.8e-07	2.8e-07
7	1.81e-5+2.13i	1.1e-14	3.7e-14	3.8e-14	6.3e-14

- Quadratic convergence.
- No need to worry about the choice of $\mathbf{c}$.
- Strange: If we neglect differentiability, we still obtained quadratic convergence for $\mathbf{B} = \mathbf{I}$.

Acknowledgements

- Everyone present
- University of Bath for Funding during my Ph. D.

G. W. Stewart.
Matrix Algorithms, volume II: Eigensystems.
SIAM, 2001.

E. Kreyszig.
Advanced Engineering Mathematics.
John Wiley & Sons, Inc., New York, eighth edition, 1999.

A. Ruhe.
Algorithms for the Nonlinear Eigenvalue Problem.
SIAM J. Matrix Anal. Appl., 10(4):674–689, 1973.

B. N. Parlett and Y. Saad.
Complex Shift and Invert Strategies for Real Matrices.
Lin. Alg. Appl., pages 575–595, 1987.

- K. Meerbergen, and D. Roose.
  Matrix Transformations for Computing Rightmost Eigenvalues of Large Sparse Non-Symmetric Eigenvalue Problems.
  IMA Journal of Numerical Analysis, 16:297–346, 1996.
- F. Tisseur.
  Newton's Method in Floating Point Arithmetic and Iterative Refinement of Generalized Eigenvalue Problems.
  SIAM J. Matrix Anal. Appl., 22:1038–1057, 2001.
- P. Deuflhard.
  Newton Methods for Nonlinear Problems, chapter 4, pages 174–175. Springer, 2004.
- B. Boisvert, R. Pozo, K. Remington, B. Miller, and R. Lipman. Matrix Market.