Using Complex Shifts in Rayleigh Quotient Iteration to Compute Close Eigenvalues

Nils Friess

15th June 2020

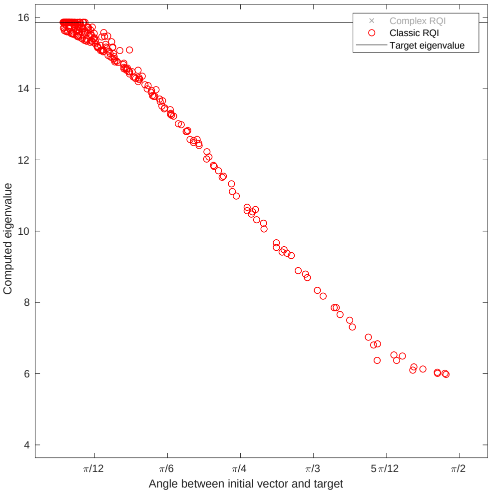
UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

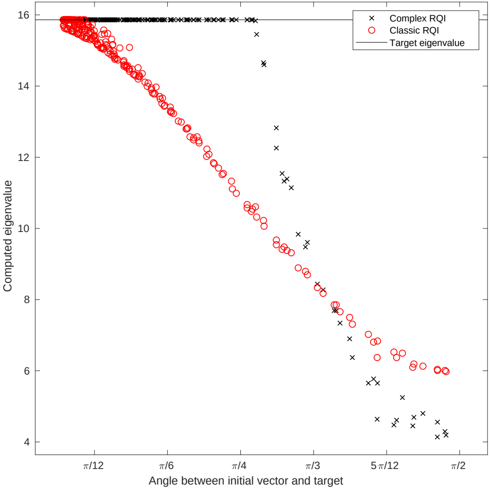
Contents

Contents

Contents

Contents


Complex RQI
Classic RQI
Target eigenvalue
Computed eigenvalue
Angle between initial vector and target
16
14
12
10
8
6
4
π/12
π/6
π/4
π/3
5 π/12
π/2


Complex RQI
Classic RQI
Target eigenvalue
Computed eigenvalue
Angle between initial vector and target
16
14
12
10
8
6
4
π/12
π/6
π/4
π/3
5 π/12
π/2

Compute one eigenpair $(\lambda, \mathbf{v})$ of symmetric matrix $\boldsymbol{A}$, i. e.,

$$\boldsymbol{A}\mathbf{v} = \lambda\mathbf{v}\,.$$

Properties

The eigenvalues $\lambda_1, \ldots, \lambda_n$ of $\boldsymbol{A}$ are real

The eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of $\boldsymbol{A}$ are real

The eigenvectors form an orthogonal basis of $\mathbb{R}^n$

Algorithm: Power method

Input: Nonzero unit vector $\mathbf{x}^{(0)}$

for $k = 1, 2, \ldots$ **until convergence do**

$\mathbf{x}^{(k)} \leftarrow \boldsymbol{A}\mathbf{x}^{(k-1)}$

Normalise $\mathbf{x}^{(k)}$

Converges linearly to $\mathbf{v}_n$ with rate

$$\rho = \frac{|\lambda_{n-1}|}{|\lambda_n|}.$$

Algorithm: Power method

Input: Nonzero unit vector $\mathbf{x}^{(0)}$

for $k = 1, 2, \ldots$ **until convergence do**

  $\mathbf{x}^{(k)} \leftarrow \boldsymbol{A}\mathbf{x}^{(k-1)}$

  Normalise $\mathbf{x}^{(k)}$

Converges linearly to $\mathbf{v}_n$ with rate

$$\rho = \frac{|\lambda_{n-1}|}{|\lambda_n|}.$$

Eigenvalues of $\boldsymbol{A}^{-1}$ are $\dfrac{1}{\lambda_i}$.

Eigenvalues of $\boldsymbol{A} - \mu \boldsymbol{I}$ are $\lambda_i - \mu$.

Eigenvalues of $(\boldsymbol{A} - \mu \boldsymbol{I})^{-1}$ are $\dfrac{1}{\lambda_i - \mu}$.

Eigenvalues of $\boldsymbol{A}^{-1}$ are $\dfrac{1}{\lambda_i}$.

Eigenvalues of $\boldsymbol{A} - \mu \boldsymbol{I}$ are $\lambda_i - \mu$.

Eigenvalues of $(\boldsymbol{A} - \mu \boldsymbol{I})^{-1}$ are $\dfrac{1}{\lambda_i - \mu}$.

Eigenvalues of $\boldsymbol{A}^{-1}$ are $\dfrac{1}{\lambda_i}$.

Eigenvalues of $\boldsymbol{A} - \mu \boldsymbol{I}$ are $\lambda_i - \mu$.

Eigenvalues of $(\boldsymbol{A} - \mu \boldsymbol{I})^{-1}$ are $\dfrac{1}{\lambda_i - \mu}$.

Algorithm: Shifted Inverse Iteration

Input: Nonzero unit vector $\mathbf{x}^{(0)}$

for $k = 1, 2, \ldots$ **until convergence do**

Solve $(\boldsymbol{A} - \sigma \boldsymbol{I})\mathbf{x}^{(k)} = \mathbf{x}^{(k-1)}$

Normalise $\mathbf{x}^{(k)}$

Converges linearly to with rate

$$\rho = \frac{|\mu_1 - \sigma|}{|\mu_2 - \sigma|}.$$

Algorithm: Shifted Inverse Iteration

Input: Nonzero unit vector $\mathbf{x}^{(0)}$

for $k = 1, 2, \ldots$ **until convergence do**

- Solve $(\boldsymbol{A} - \sigma \boldsymbol{I})\mathbf{x}^{(k)} = \mathbf{x}^{(k-1)}$
- Normalise $\mathbf{x}^{(k)}$

Converges linearly to with rate

$$\rho = \frac{|\mu_1 - \sigma|}{|\mu_2 - \sigma|}\,.$$

Eigenvector approximation from eigenvalue approximation?

Eigenvalue approximation from eigenvector approximation?

Shifted Inverse Iteration

Rayleigh Quotient

$$\mathcal{R}_A(\mathbf{x}) = \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}}$$

Eigenvector approximation from eigenvalue approximation?

Shifted Inverse Iteration

Eigenvalue approximation from eigenvector approximation?

Rayleigh Quotient

$$\mathcal{R}_A(\mathbf{x}) = \frac{\mathbf{x}^\top A\mathbf{x}}{\mathbf{x}^\top \mathbf{x}}$$

Eigenvector approximation from eigenvalue approximation?

Shifted Inverse Iteration

Eigenvalue approximation from eigenvector approximation?

Rayleigh Quotient

$$\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) = \frac{\mathbf{x}^\mathsf{T} \boldsymbol{A} \mathbf{x}}{\mathbf{x}^\mathsf{T} \mathbf{x}}$$

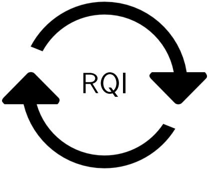
Eigenvector approximation from eigenvalue approximation?

Shifted Inverse Iteration

Eigenvalue approximation from eigenvector approximation?

Rayleigh Quotient

$$\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) = \frac{\mathbf{x}^\mathsf{T} \boldsymbol{A} \mathbf{x}}{\mathbf{x}^\mathsf{T} \mathbf{x}}$$



$$\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) = \arg\min_{\mu \in \mathbb{C}} \|\boldsymbol{A}\mathbf{x} - \mu\mathbf{x}\|$$

Eigenpair $(\lambda, \mathbf{v})$: $$|\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) - \lambda| = \mathcal{O}\left(\|\mathbf{x} - \mathbf{v}\|^2\right)$$

Algorithm: Rayleigh Quotient Iteration

Input: Nonzero unit vector $\mathbf{x}^{(0)}$

$\mu^{(0)} \leftarrow \left(\mathbf{x}^{(0)}\right)^{\mathsf{T}} \boldsymbol{A}\mathbf{x}^{(0)}$

for $k = 1, 2, \ldots$ **until convergence do**

- Solve $\left(\boldsymbol{A} - \mu^{(k)}\boldsymbol{I}\right)\mathbf{y}^{(k)} = \mathbf{x}^{(k-1)}$ for $\mathbf{y}^{(k)}$
- $\mathbf{x}^{(k)} \leftarrow \mathbf{y}^{(k)} / \left\|\mathbf{y}^{(k)}\right\|$
- $\mu^{(k)} \leftarrow \left(\mathbf{x}^{(k)}\right)^{\mathsf{T}} \boldsymbol{A}\mathbf{x}^{(k)}$

$\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) = \frac{\mathbf{x}^{\mathsf{T}}\boldsymbol{A}\mathbf{x}}{\mathbf{x}^{\mathsf{T}}\mathbf{x}}$

Converges locally cubically

- Correct digits triple

Converges for almost all starting vectors

- Impossible to fail in practice

Converges locally cubically

- Correct digits triple

Converges for almost all starting vectors

- Impossible to fail in practice

Converges locally cubically

- Correct digits triple

Converges for almost all starting vectors

- Impossible to fail in practice

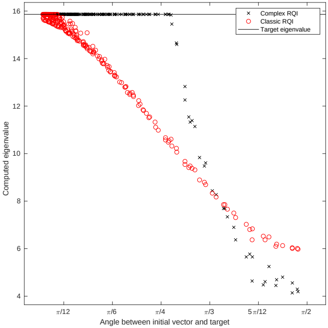
Converges locally cubically

- Correct digits triple

Converges for almost all starting vectors

- Impossible to fail in practice

Convergence behaviour can be erratic



Observation

Convergence depends on initial shift but **not** initial vector

Idea: Perturb original problem to increase distance between eigenvalues.

Suppose $\mathbf{u} \approx \mathbf{v}_k$.

$$\boldsymbol{A} \quad \longrightarrow \quad \boldsymbol{A} - i\gamma(\boldsymbol{I} - \mathbf{u}\mathbf{u}^\mathsf{T})$$

Idea: Perturb original problem to increase distance between eigenvalues.

Suppose $\mathbf{u} \approx \mathbf{v}_k$.

$$\boldsymbol{A} \quad \longrightarrow \quad \boldsymbol{A} - i\gamma(\boldsymbol{I} - \mathbf{u}\mathbf{u}^\mathsf{T})$$

Suppose $\mathbf{u} \approx \mathbf{v}_k$. Set $\tilde{\boldsymbol{A}} = \boldsymbol{A} - i\gamma(\boldsymbol{I} - \mathbf{u}\mathbf{u}^\mathsf{T})$.

Vector $\mathbf{x}$ that is almost parallel to $\mathbf{u}$:

$$\tilde{\boldsymbol{A}}\mathbf{x} = \boldsymbol{A}\mathbf{x} - i\gamma(\mathbf{x} - \mathbf{u}\,\underbrace{\mathbf{u}^\mathsf{T}\mathbf{x}}_{\approx 1}) \approx \boldsymbol{A}\mathbf{x}\,.$$

Vector $\mathbf{y}$ that is almost perpendicular to $\mathbf{u}$:

$$\tilde{\boldsymbol{A}}\mathbf{y} = \boldsymbol{A}\mathbf{y} - i\gamma(\mathbf{y} - \mathbf{u}\,\underbrace{\mathbf{u}^\mathsf{T}\mathbf{y}}_{\approx 0}) \approx (\boldsymbol{A} - i\gamma\boldsymbol{I})\mathbf{y}\,.$$

Suppose $\mathbf{u} \approx \mathbf{v}_k$. Set $\tilde{\boldsymbol{A}} = \boldsymbol{A} - i\gamma(\boldsymbol{I} - \mathbf{u}\mathbf{u}^\mathsf{T})$.

Vector $\mathbf{x}$ that is almost parallel to $\mathbf{u}$:

$$\tilde{\boldsymbol{A}}\mathbf{x} = \boldsymbol{A}\mathbf{x} - i\gamma(\mathbf{x} - \mathbf{u}\underbrace{\mathbf{u}^\mathsf{T}\mathbf{x}}_{\approx 1}) \approx \boldsymbol{A}\mathbf{x}\,.$$

Vector $\mathbf{y}$ that is almost perpendicular to $\mathbf{u}$:

$$\tilde{\boldsymbol{A}}\mathbf{y} = \boldsymbol{A}\mathbf{y} - i\gamma(\mathbf{y} - \mathbf{u}\underbrace{\mathbf{u}^\mathsf{T}\mathbf{y}}_{\approx 0}) \approx (\boldsymbol{A} - i\gamma\boldsymbol{I})\mathbf{y}\,.$$

Eigenvectors and eigenvalues of $\tilde{\boldsymbol{A}}$:

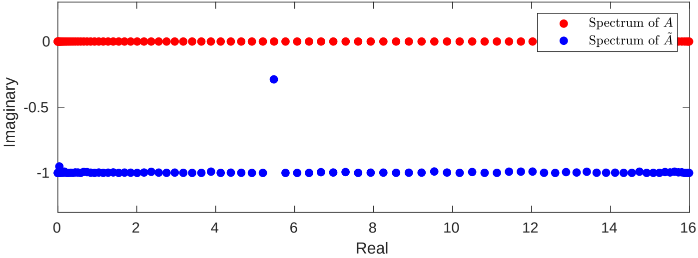
$$\tilde{\boldsymbol{A}}\mathbf{v}_k \approx \boldsymbol{A}\mathbf{v}_k = \lambda_k \mathbf{v}_k$$

$$\tilde{\boldsymbol{A}}\mathbf{v}_j \approx (\boldsymbol{A} - i\gamma \boldsymbol{I})\mathbf{v}_j = (\lambda_j - i\gamma)\mathbf{v}_j$$

Eigenvectors and eigenvalues of $\tilde{\boldsymbol{A}}$:

$$\tilde{\boldsymbol{A}}\mathbf{v}_k \approx \boldsymbol{A}\mathbf{v}_k = \lambda_k \mathbf{v}_k$$

$$\tilde{\boldsymbol{A}}\mathbf{v}_j \approx (\boldsymbol{A} - i\gamma \boldsymbol{I})\mathbf{v}_j = (\lambda_j - i\gamma)\mathbf{v}_j$$



Write $\tilde{\boldsymbol{A}} = \tilde{\boldsymbol{A}}_{(0)} + \tilde{\boldsymbol{A}}_{(1)}$ with

$$\tilde{\boldsymbol{A}}_{(0)} := \boldsymbol{A} - i\gamma(\boldsymbol{I} - \mathbf{v}_k \mathbf{v}_k^\mathsf{T})$$

$$\tilde{\boldsymbol{A}}_{(1)} := i\gamma(\mathbf{u}\mathbf{u}^\mathsf{T} - \mathbf{v}_k \mathbf{v}_k^\mathsf{T})$$

Theorem

$$\lambda_j(\tilde{\boldsymbol{A}}) = \lambda_j(\boldsymbol{A}) + i\gamma\left(\langle \mathbf{v}_j, \mathbf{u}\rangle^2 - 1\right) + \mathcal{O}\left(\left\|\tilde{\boldsymbol{A}}_{(1)}\right\|^2\right),$$

where

$$\left\|\tilde{\boldsymbol{A}}_{(1)}\right\| = \gamma\sqrt{1 - \langle \mathbf{u}, \mathbf{v}_k\rangle^2}.$$

Write $\tilde{\boldsymbol{A}} = \tilde{\boldsymbol{A}}_{(0)} + \tilde{\boldsymbol{A}}_{(1)}$ with

$$\tilde{\boldsymbol{A}}_{(0)} := \boldsymbol{A} - i\gamma(\boldsymbol{I} - \mathbf{v}_k \mathbf{v}_k^\mathsf{T})$$

$$\tilde{\boldsymbol{A}}_{(1)} := i\gamma(\mathbf{u}\mathbf{u}^\mathsf{T} - \mathbf{v}_k \mathbf{v}_k^\mathsf{T})$$

Theorem

$$\lambda_j(\tilde{\boldsymbol{A}}) = \lambda_j(\boldsymbol{A}) + i\gamma\left(\langle \mathbf{v}_j, \mathbf{u}\rangle^2 - 1\right) + \mathcal{O}\left(\left\|\tilde{\boldsymbol{A}}_{(1)}\right\|^2\right),$$

where

$$\left\|\tilde{\boldsymbol{A}}_{(1)}\right\| = \gamma\sqrt{1 - \langle \mathbf{u}, \mathbf{v}_k\rangle^2}\,.$$