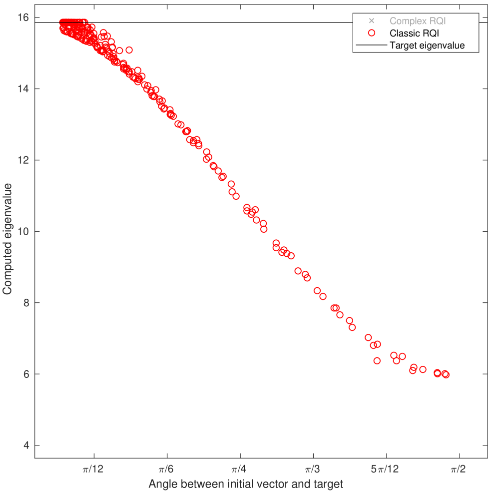
Using Complex Shifts in Rayleigh Quotient Iteration to Compute Close Eigenvalues

Nils Friess

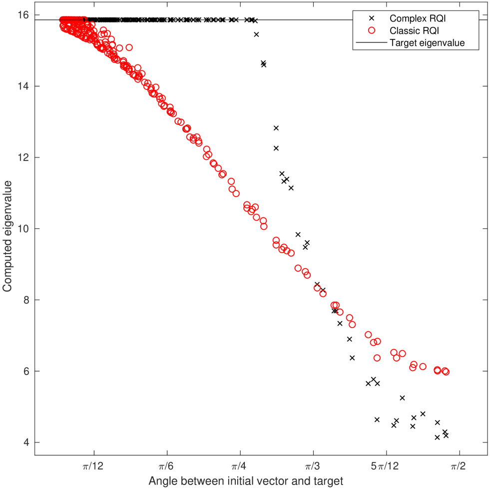
17th June 2020

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Contents



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


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
5π/12
π/2

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

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

Theorem

- *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 $k = 0, 1, \ldots$ **until convergence do**

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

Normalise $\mathbf{x}^{(k+1)}$

Theorem

- *The eigenvalues of* $\boldsymbol{A}^{-1}$ *are* $\dfrac{1}{\lambda_i}$.
- *The eigenvalues of* $\boldsymbol{A} - \mu \boldsymbol{I}$ *are* $\lambda_i - \mu$.
- *The 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 $k = 0, 1, \ldots$ **until convergence do**

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

Normalise $\mathbf{x}^{(k+1)}$

Rayleigh Quotient

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

Rayleigh Quotient

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

Eigenvector approximation from eigenvalue approximation?

- Shifted Inverse Iteration

Eigenvalue approximation from eigenvector approximation?

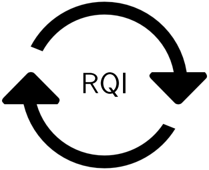
- Rayleigh Quotient

Eigenvector approximation from eigenvalue approximation?

- Shifted Inverse Iteration

Eigenvalue approximation from eigenvector approximation?

- Rayleigh Quotient



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 $k = 0, 1, \ldots$ **until convergence do**

Solve $\left(\boldsymbol{A} - \mu^{(k)}\boldsymbol{I}\right)\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$ for $\mathbf{x}^{(k+1)}$

Normalise $\mathbf{x}^{(k+1)}$

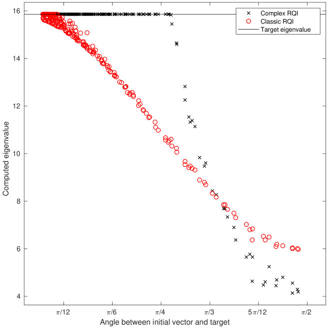
$\mu^{(k+1)} \leftarrow \left(\mathbf{x}^{(k+1)}\right)^{\mathsf{T}} \boldsymbol{A}\mathbf{x}^{(k+1)}$

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

- Cubic convergence
- Converges for almost all starting vectors

- Cubic convergence
- Converges for almost all starting vectors

Convergence behaviour can be erratic



Observation

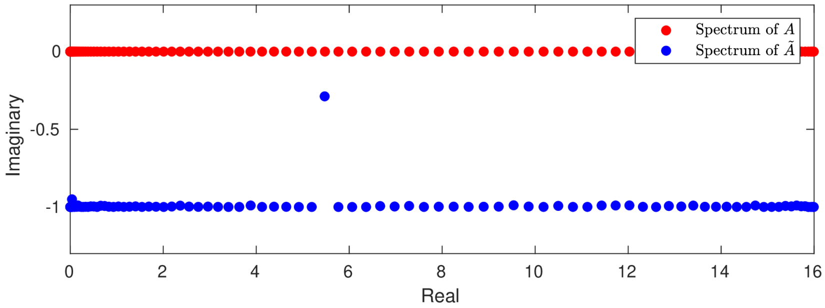
Convergence depends on initial shift but **not** initial vector

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

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

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

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



Algorithm: Complex RQI (first version)

for $k = 0, 1, \ldots$ **until convergence do**

  Solve $\left(\boldsymbol{A} - i\gamma\left(\boldsymbol{I} - \mathbf{u}\mathbf{u}^{\mathsf{T}}\right) - \mu^{(k)}\boldsymbol{I}\right)\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$

  Normalise $\mathbf{x}^{(k+1)}$

  $\mu^{(k+1)} \leftarrow \left(\mathbf{x}^{(k+1)}\right)^{*}\boldsymbol{A}\mathbf{x}^{(k+1)}$

Algorithm: Complex RQI (first version)

for $k = 0, 1, \ldots$ **until convergence do**

- Solve $\left(\boldsymbol{A} - i\gamma^{(k)}\left(\boldsymbol{I} - \mathbf{x}^{(k)}(\mathbf{x}^{(k)})^*\right) - \mu^{(k)}\boldsymbol{I}\right)\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$
- Normalise $\mathbf{x}^{(k+1)}$
- $\mu^{(k+1)} \leftarrow (\mathbf{x}^{(k+1)})^* \boldsymbol{A}\mathbf{x}^{(k+1)}$
- $\gamma^{(k+1)} \leftarrow \|(\boldsymbol{A} - \mu^{(k+1)}\boldsymbol{I})\mathbf{x}^{(k+1)}\|$

Theorem

It suffices to use

$$\tilde{\boldsymbol{A}} = \boldsymbol{A} - i\gamma^{(k)}\boldsymbol{I}$$

instead of

$$\tilde{\boldsymbol{A}} = \boldsymbol{A} - i\gamma^{(k)}\left(\boldsymbol{I} - \mathbf{x}^{(k)}\left(\mathbf{x}^{(k)}\right)^{*}\right).$$

Algorithm: Complex Rayleigh Quotient Iteration

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

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

$\gamma^{(0)} \leftarrow \|(\boldsymbol{A} - \mu^{(0)}\boldsymbol{I})\mathbf{x}^{(0)}\|$

$\sigma^{(0)} \leftarrow \mu^{(0)} + i\gamma^{(0)}$

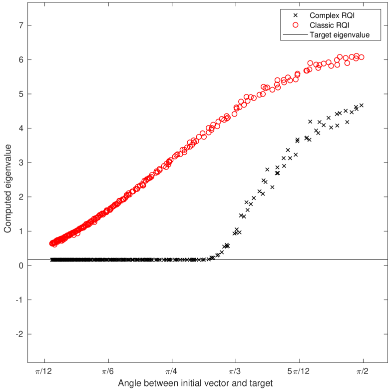
for $k = 0, 1, \ldots$ **until convergence do**

- Solve $(\boldsymbol{A} - \sigma^{(k)}\boldsymbol{I})\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$
- Normalise $\mathbf{x}^{(k+1)}$
- $\mu^{(k+1)} \leftarrow \left(\mathbf{x}^{(k+1)}\right)^{*} \boldsymbol{A}\mathbf{x}^{(k+1)}$
- $\gamma^{(k+1)} \leftarrow \|(\boldsymbol{A} - \mu^{(k+1)}\boldsymbol{I})\mathbf{x}^{(k+1)}\|$
- $\sigma^{(k+1)} \leftarrow \mu^{(k+1)} + i\gamma^{(k+1)}$

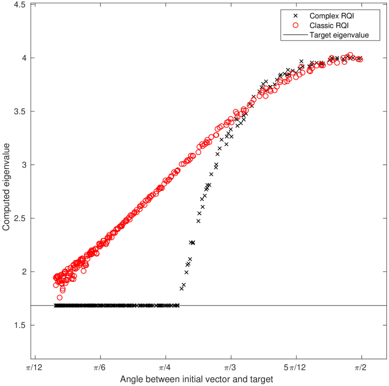
$\mathbf{x} \leftarrow \mathrm{Re}(\mathbf{x}^{(k+1)})$

$\mathbf{x} \leftarrow \mathbf{x}/\|\mathbf{x}\|$

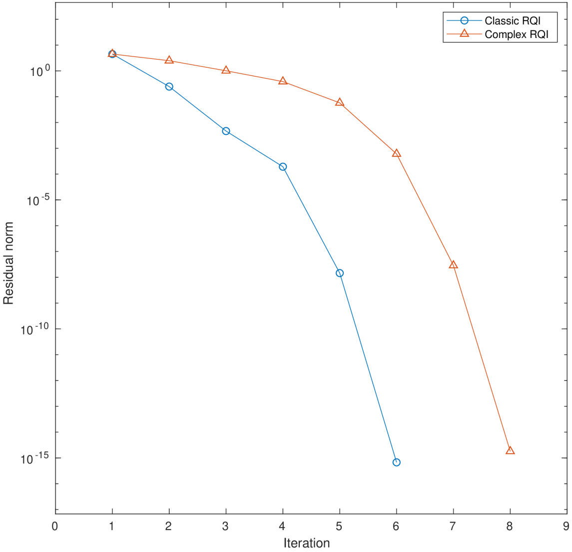
$\mu \leftarrow \mathbf{x}^{\mathsf{T}}\boldsymbol{A}\mathbf{x}$

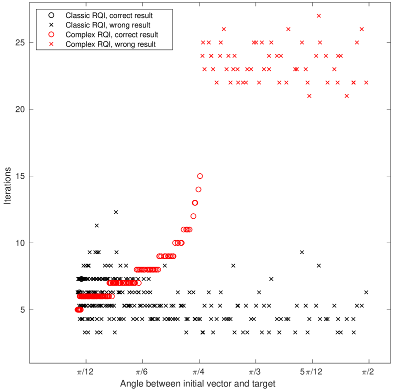


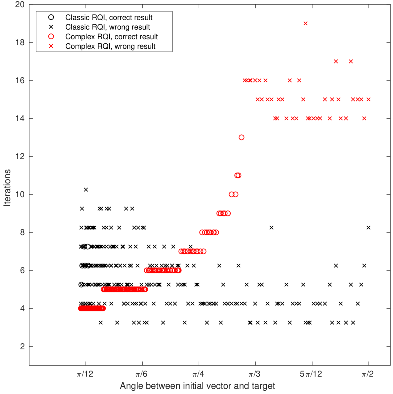
Complex RQI
Classic RQI
Target eigenvalue
Computed eigenvalue
Angle between initial vector and target
7
6
5
4
3
2
1
0
-1
-2
π/12
π/6
π/4
π/3
5π/12
π/2



Complex RQI
Classic RQI
Target eigenvalue
Computed eigenvalue
Angle between initial vector and target
4.5
4
3.5
3
2.5
2
1.5
π/12
π/6
π/4
π/3
5π/12
π/2


Classic RQI
Complex RQI
Residual norm
10^{0}
10^{-5}
10^{-10}
10^{-15}
0 1 2 3 4 5 6 7 8 9
Iteration


Classic RQI, correct result
Classic RQI, wrong result
Complex RQI, correct result
Complex RQI, wrong result
Iterations
5
10
15
20
25
π/12
π/6
π/4
π/3
5π/12
π/2
Angle between initial vector and target


Classic RQI, correct result
Classic RQI, wrong result
Complex RQI, correct result
Complex RQI, wrong result
Iterations
2
4
6
8
10
12
14
16
18
20
π/12
π/6
π/4
π/3
5π/12
π/2
Angle between initial vector and target

Summary

- Unpredictability of classic RQI
- Complex RQI succeeds for error angles below 45°
- More iterations than classic RQI but still fast

- Theoretical analysis
- Other test problems

Summary

- Unpredictability of classic RQI
- Complex RQI succeeds for error angles below 45°
- More iterations than classic RQI but still fast

- Theoretical analysis
- Other test problems

Write $\tilde{\boldsymbol{A}} = \boldsymbol{A} - i\gamma(\boldsymbol{I} - \mathbf{u}\mathbf{u}^\mathsf{T}) = \tilde{\boldsymbol{A}}_{(0)} + \tilde{\boldsymbol{A}}_{(1)}$ 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 $\tilde{\boldsymbol{A}} = \boldsymbol{A} - i\gamma(\boldsymbol{I} - \mathbf{u}\mathbf{u}^{\mathsf{T}}) = \tilde{\boldsymbol{A}}_{(0)} + \tilde{\boldsymbol{A}}_{(1)}$ 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

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