KU LEUVEN

RQZ: A rational QZ method for the generalized eigenvalue problem

Daan Camps, Karl Meerbergen, and Raf Vandebril

May 6, 2018

KU Leuven - University of Leuven - Department of Computer Science - NUMA Section

Overview

In this talk we will discuss:

Numerical solution of the generalized eigenvalue problem:

Overview

In this talk we will discuss:

Numerical solution of the generalized eigenvalue problem:

- ♢ Small to medium-sized
- ♢ Regular
- ♢ All eigenvalues required
- ♢ No symmetry assumed

Overview

In this talk we will discuss:

Numerical solution of the generalized eigenvalue problem:

- ◊ Small to medium-sized
- ◊ Regular
- ◊ All eigenvalues required
- ◊ No symmetry assumed

Tools:

Overview

In this talk we will discuss:

Numerical solution of the generalized eigenvalue problem:

- ◊ Small to medium-sized
- ◊ Regular
- ◊ All eigenvalues required
- ◊ No symmetry assumed

Tools:

- ◊ Hessenberg, Hessenberg form
- ◊ Shift & *pole* introduction and swapping
- ◊ Rational Krylov
- ◊ Subspace iteration

Overview

In this talk we will discuss:

Numerical solution of the generalized eigenvalue problem:

- ◊ Small to medium-sized
- ◊ Regular
- ◊ All eigenvalues required
- ◊ No symmetry assumed

Tools:

- ◊ Hessenberg, Hessenberg form
- ◊ Shift & *pole* introduction and swapping
- ◊ Rational Krylov
- ◊ Subspace iteration

Solution:

Overview

Overview

In this talk we will discuss:

Numerical solution of the generalized eigenvalue problem:

- ◊ Small to medium-sized
- ◊ Regular
- ◊ All eigenvalues required
- ◊ No symmetry assumed

Tools:

- ◊ Hessenberg, Hessenberg form
- ◊ Shift & *pole* introduction and swapping
- ◊ Rational Krylov
- ◊ Subspace iteration

Solution:

- ◊ Method of QR-type driven by rational functions

Overview

In this talk we will discuss:

Numerical solution of the generalized eigenvalue problem:

- ♢ Small to medium-sized
- ♢ Regular
- ♢ All eigenvalues required
- ♢ No symmetry assumed

Tools:

- ♢ Hessenberg, Hessenberg form
- ♢ Shift & *pole* introduction and swapping
- ♢ Rational Krylov
- ♢ Subspace iteration

Solution:

- ♢ Method of QR-type driven by rational functions

Generalized eigenvalue problems

- Given A & B: $n \times n$ matrices, either $\mathbb{R}$ or $\mathbb{C}$
- Computation of the triplets $(\alpha, \beta, \boldsymbol{x})$ that satisfy $\beta A\boldsymbol{x} = \alpha B\boldsymbol{x}$
- Procedure:
  1. Reduce the pencil to a *manageable* form
  2. Iterate to generalized Schur form
  3. Recover eigenvectors
- Make use of well-chosen unitary equivalences:

  $(\hat{A}, \hat{B}) = Q^*(A, B)Z$ preserves eigenvalues

Generalized eigenvalue problems

- Given A & B: $n \times n$ matrices, either $\mathbb{R}$ or $\mathbb{C}$
- Computation of the triplets $(\alpha, \beta, \boldsymbol{x})$ that satisfy $\beta A\boldsymbol{x} = \alpha B\boldsymbol{x}$
- Procedure:
  1. Reduce the pencil to a *manageable* form
  2. Iterate to generalized Schur form
  3. Recover eigenvectors
- Make use of well-chosen unitary equivalences:

  $(\hat{A}, \hat{B}) = Q^*(A, B)Z$ preserves eigenvalues

Overview

In this talk we will discuss:

Numerical solution of the generalized eigenvalue problem $\beta A\boldsymbol{x} = \alpha B\boldsymbol{x}$:

- Small to medium-sized
- Regular
- All eigenvalues required
- No symmetry assumed

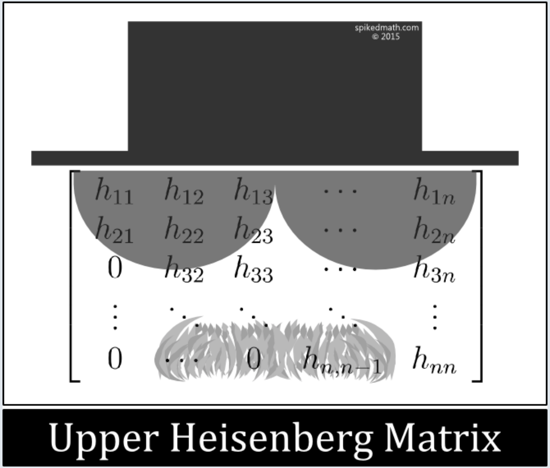
Tools:

- Hessenberg, Hessenberg form
- Shift & *pole* introduction and swapping
- Rational Krylov
- Subspace iteration

Solution:

- Method of QR-type driven by rational functions

Hessenberg, Hessenberg form



source: spikedmath.com/573.html

Hessenberg, Hessenberg form

$$
\underset{A}{\begin{matrix}
\times & \times & \times & \times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times & \times & \times & \times \\
 & \times & \times & \times & \times & \times & \times & \times \\
 & & \times & \times & \times & \times & \times & \times \\
 & & & \times & \times & \times & \times & \times \\
 & & & & \times & \times & \times & \times \\
 & & & & & \times & \times & \times \\
 & & & & & & \times & \times
\end{matrix}}
,
\underset{B}{\begin{matrix}
\times & \times & \times & \times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times & \times & \times & \times \\
 & \times & \times & \times & \times & \times & \times & \times \\
 & & \times & \times & \times & \times & \times & \times \\
 & & & \times & \times & \times & \times & \times \\
 & & & & \times & \times & \times & \times \\
 & & & & & \times & \times & \times \\
 & & & & & & \times & \times
\end{matrix}}
$$

Tools

Hessenberg, Hessenberg form

$$
\underbrace{\begin{matrix}
\times & \times & \times & \times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times & \times & \times & \times \\
 & \times & \times & \times & \times & \times & \times & \times \\
 & & \times & \times & \times & \times & \times & \times \\
 & & & \times & \times & \times & \times & \times \\
 & & & & \times & \times & \times & \times \\
 & & & & & \times & \times & \times \\
 & & & & & & \times & \times
\end{matrix}}_{A}
\quad , \quad
\underbrace{\begin{matrix}
\times & \times & \times & \times & \times & \times & \times & \times \\
\times & \times & \times & \times & \times & \times & \times & \times \\
 & \times & \times & \times & \times & \times & \times & \times \\
 & & \times & \times & \times & \times & \times & \times \\
 & & & \times & \times & \times & \times & \times \\
 & & & & \times & \times & \times & \times \\
 & & & & & \times & \times & \times \\
 & & & & & & \times & \times
\end{matrix}}_{B}
$$

Tools

Hessenberg, Hessenberg form

$$
\begin{array}{cccccccc}
\times & \times & \times & \times & \times & \times & \times & \times \\
\color{green}{\times} & \times & \times & \times & \times & \times & \times & \times \\
 & \color{green}{\times} & \times & \times & \times & \times & \times & \times \\
 & & \color{green}{\times} & \times & \times & \times & \times & \times \\
 & & & \color{green}{\times} & \times & \times & \times & \times \\
 & & & & \color{green}{\times} & \times & \times & \times \\
 & & & & & \color{green}{\times} & \times & \times \\
 & & & & & & \color{green}{\times} & \times
\end{array}
\quad , \quad
\begin{array}{cccccccc}
\times & \times & \times & \times & \times & \times & \times & \times \\
\color{red}{\times} & \times & \times & \times & \times & \times & \times & \times \\
 & \color{red}{\times} & \times & \times & \times & \times & \times & \times \\
 & & \color{red}{\times} & \times & \times & \times & \times & \times \\
 & & & \color{red}{\times} & \times & \times & \times & \times \\
 & & & & \color{red}{\times} & \times & \times & \times \\
 & & & & & \color{red}{\times} & \times & \times \\
 & & & & & & \color{red}{\times} & \times
\end{array}
$$

$$A \qquad\qquad\qquad\qquad B$$

$$\text{poles } \Xi = \left(\frac{\color{green}{\times}}{\color{red}{\times}}\right) \subset \bar{\mathbb{C}}$$

Hessenberg, Hessenberg form

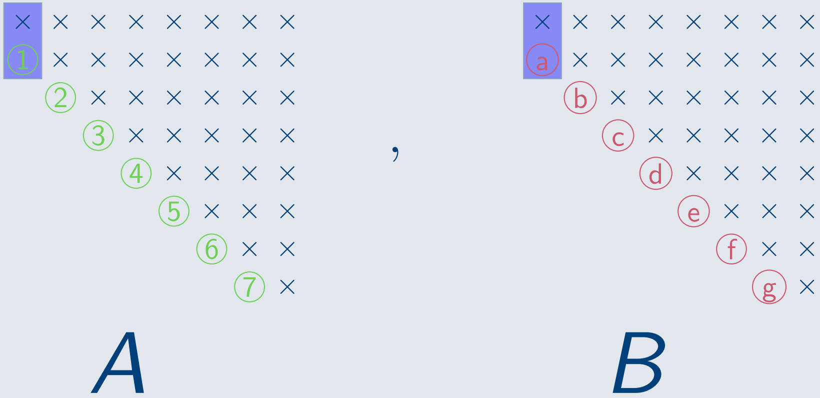
$$
\begin{matrix}
\times & \times & \times & \times & \times & \times & \times & \times \\
\textcircled{1} & \times & \times & \times & \times & \times & \times & \times \\
 & \textcircled{2} & \times & \times & \times & \times & \times & \times \\
 & & \textcircled{3} & \times & \times & \times & \times & \times \\
 & & & \textcircled{4} & \times & \times & \times & \times \\
 & & & & \textcircled{5} & \times & \times & \times \\
 & & & & & \textcircled{6} & \times & \times \\
 & & & & & & \textcircled{7} & \times
\end{matrix}
\quad , \quad
\begin{matrix}
\times & \times & \times & \times & \times & \times & \times & \times \\
\textcircled{a} & \times & \times & \times & \times & \times & \times & \times \\
 & \textcircled{b} & \times & \times & \times & \times & \times & \times \\
 & & \textcircled{c} & \times & \times & \times & \times & \times \\
 & & & \textcircled{d} & \times & \times & \times & \times \\
 & & & & \textcircled{e} & \times & \times & \times \\
 & & & & & \textcircled{f} & \times & \times \\
 & & & & & & \textcircled{g} & \times
\end{matrix}
$$

$$A \qquad\qquad\qquad\qquad B$$

$$\text{poles } \Xi = \left(\frac{\textcircled{1}}{\textcircled{a}}, \frac{\textcircled{2}}{\textcircled{b}}, \ldots\right) \subset \bar{\mathbb{C}}$$

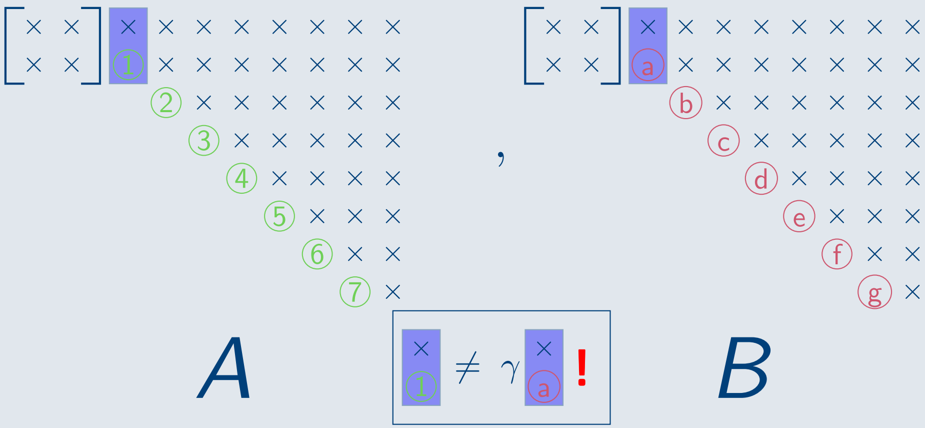
Tools

Introducing a shift



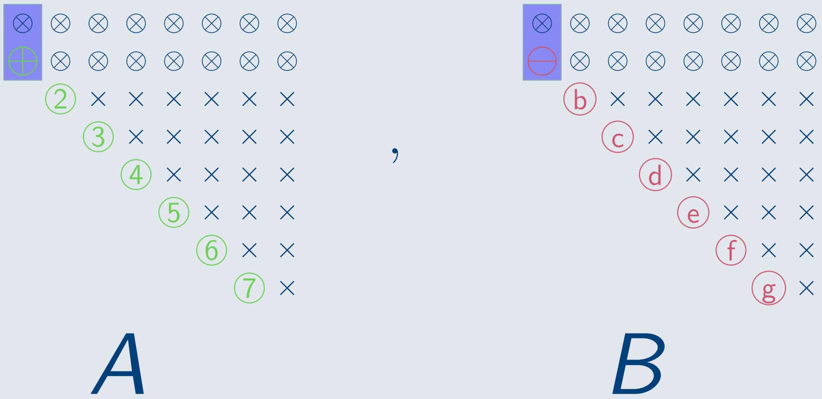
Tools

Introducing a shift



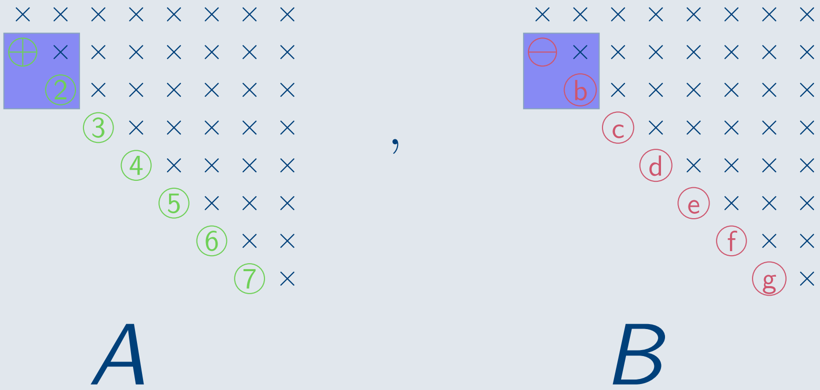
Tools

Introducing a shift



Tools

Swapping poles



Swapping poles

Solve coupled Sylvester equation

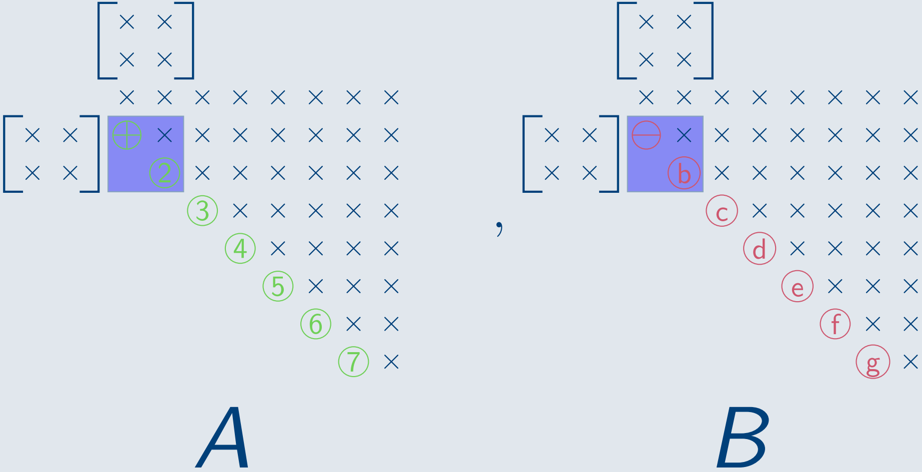
(cfr. reordering Schur form [Kågström and Poromaa])

$$\frac{\oplus}{\ominus} \neq \frac{\textcircled{2}}{\textcircled{b}} \,! \qquad \frac{\oplus}{\ominus} \neq \frac{0}{0} \,! \qquad \frac{\textcircled{2}}{\textcircled{b}} \neq \frac{0}{0} \,!$$

$$\Rightarrow Q^* = \begin{bmatrix} \times & \times \\ \times & \times \end{bmatrix}, Z = \begin{bmatrix} \times & \times \\ \times & \times \end{bmatrix}$$

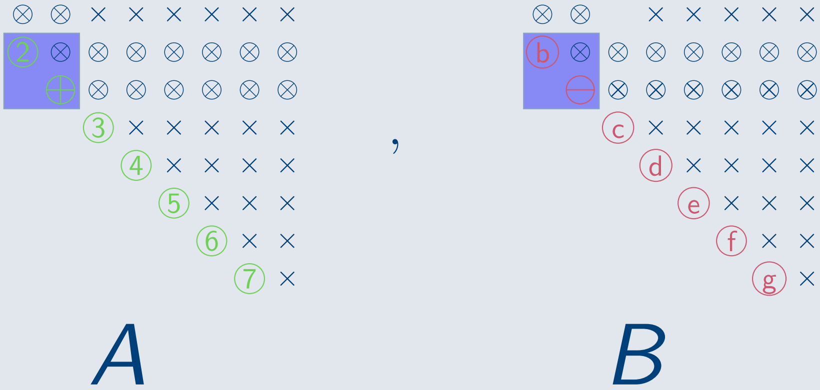
Tools

Swapping poles



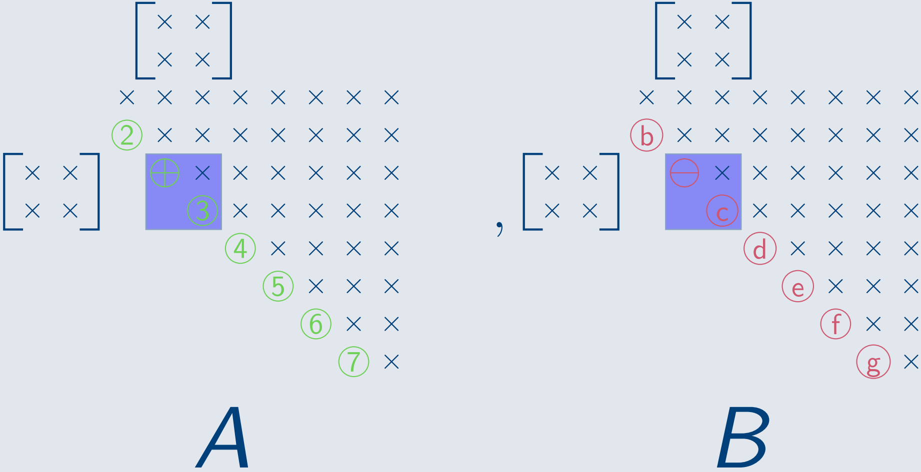
Tools

Swapping poles



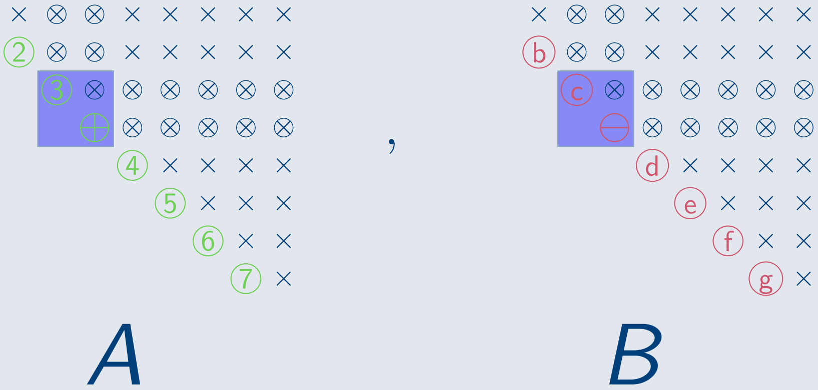
Tools

Swapping poles



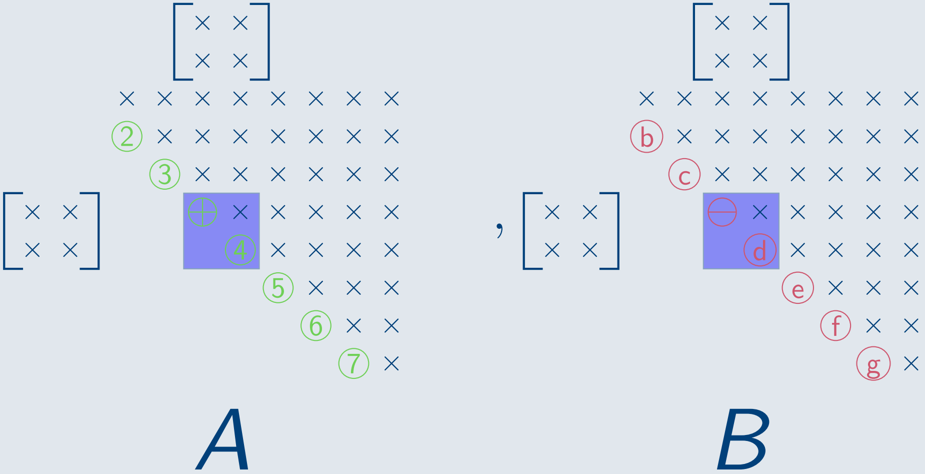
Tools

Swapping poles



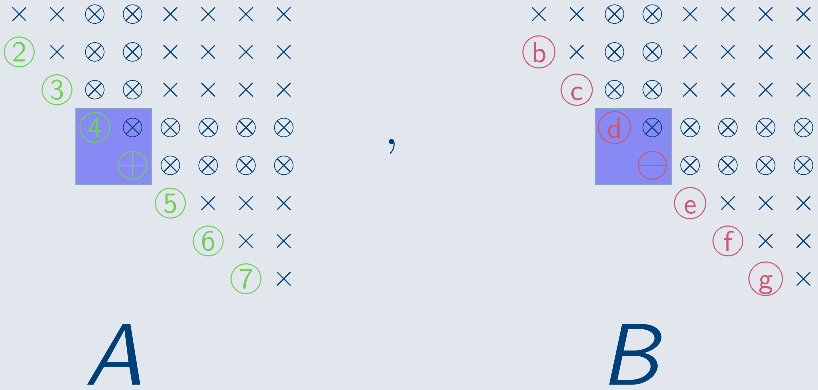
Tools

Swapping poles



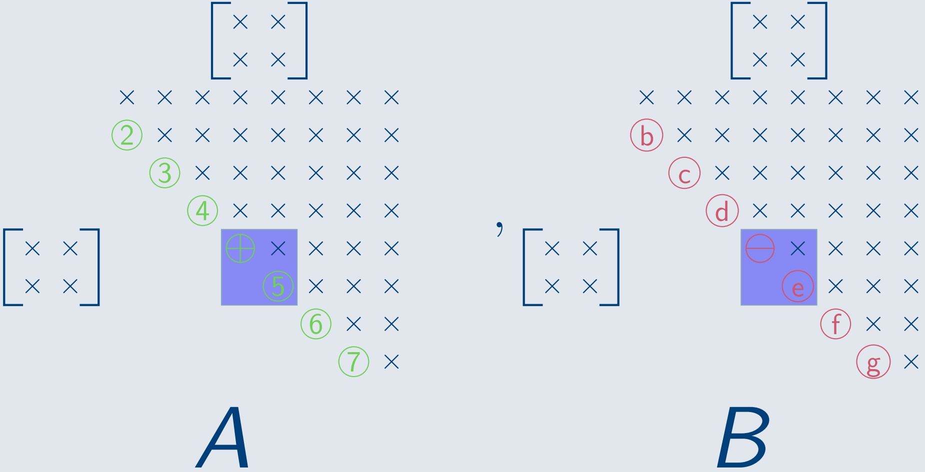
Tools

Swapping poles



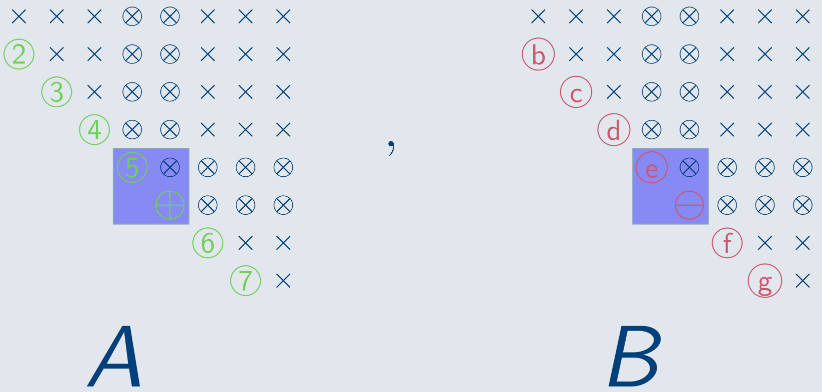
Tools

Swapping poles



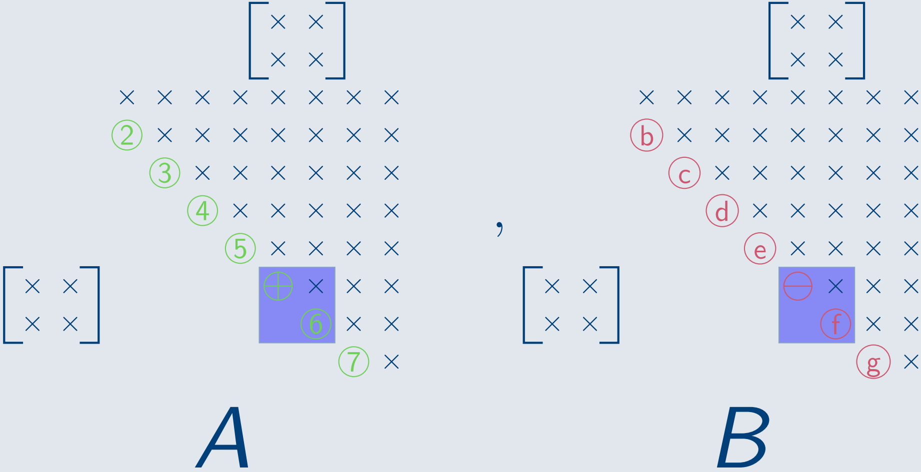
Tools

Swapping poles



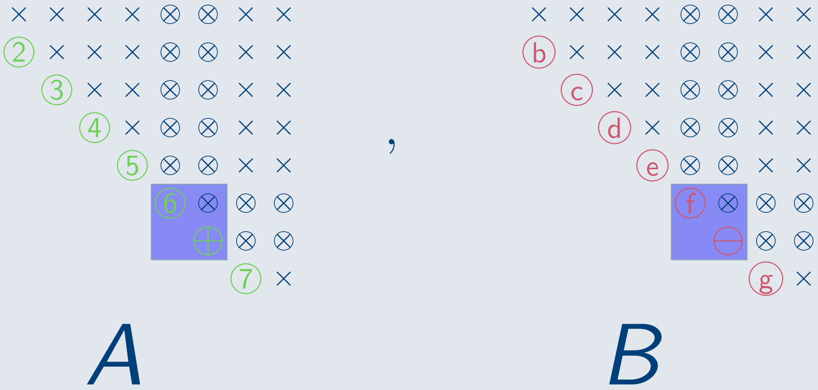
Tools

Swapping poles



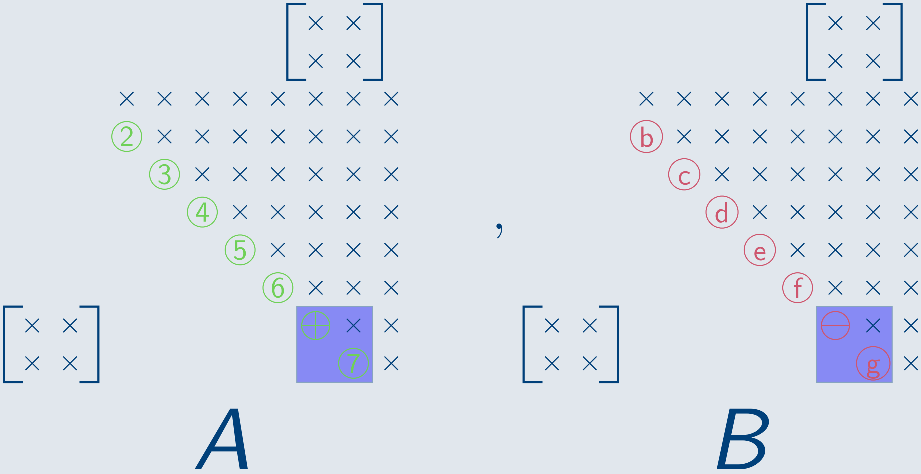
Tools

Swapping poles



Tools

Swapping poles

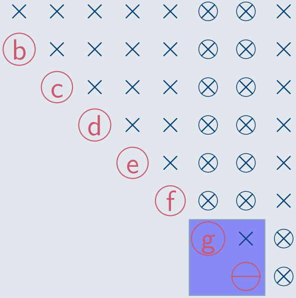


Tools

Swapping poles

A

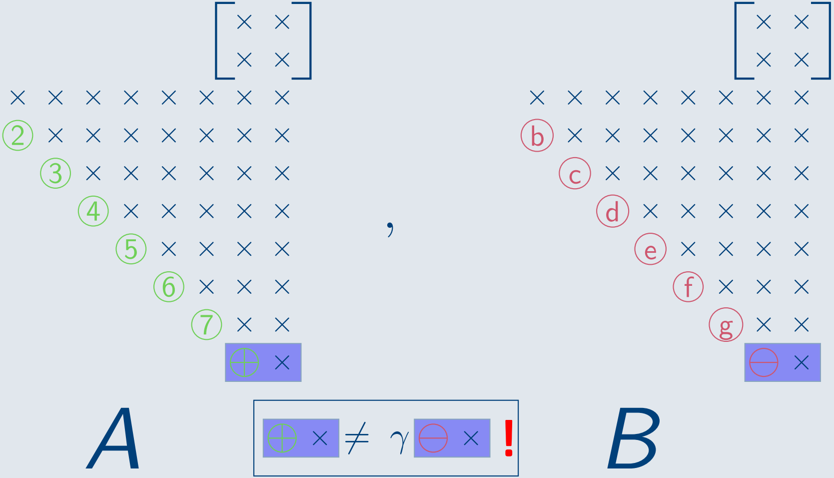
,



B

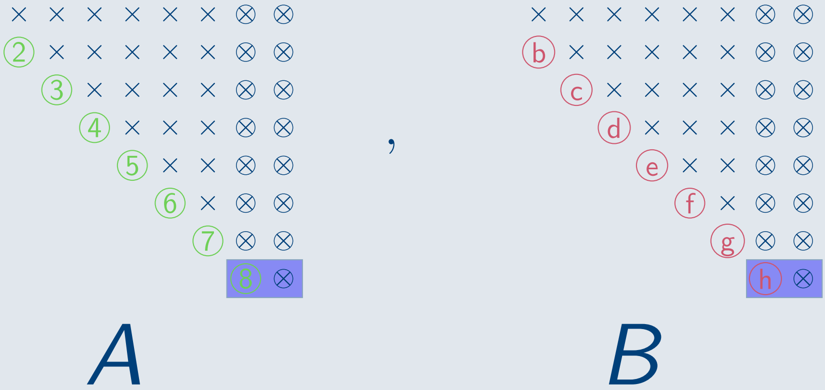
Tools

Introducing a pole



Tools

Introducing a pole



The algorithm in a nutshell:

The algorithm in a nutshell:

1. Introduce shift at the top

RQZ algorithm

The algorithm in a nutshell:

1. Introduce shift at the top
2. Swap it all the way down

RQZ algorithm

The algorithm in a nutshell:

1. Introduce shift at the top
2. Swap it all the way down
3. Introduce pole at the end

The algorithm in a nutshell:

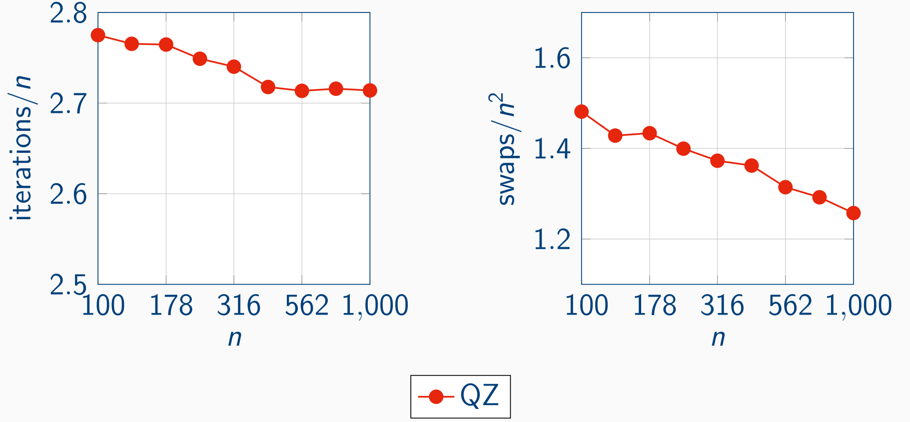
1. Introduce shift at the top
2. Swap it all the way down
3. Introduce pole at the end

Poles at ∞ **(**$\times = 0$**)** $\rightarrow$ **QZ method**: Bulge exchange interpretation [Watkins]

Caution: shift $\notin \Xi$ to avoid slower convergence

Numerical example

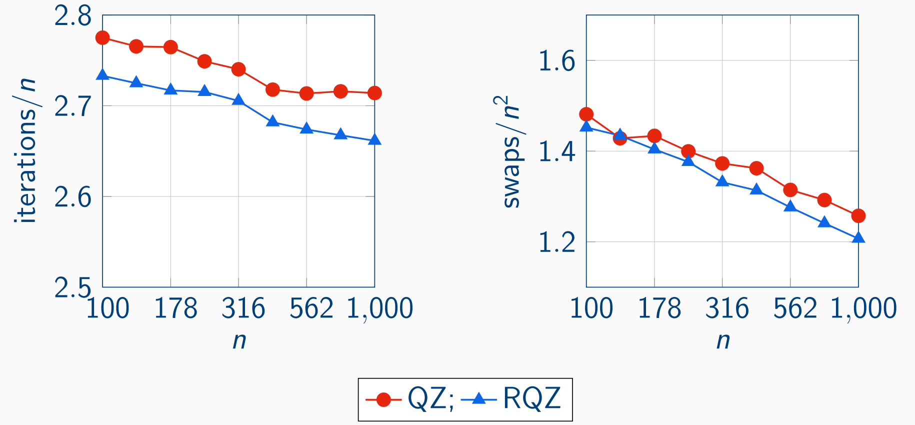
Is it worth it?



Data: 9 random matrix pairs, $n \in [100, 1000]$, reduced to H-T, averaged over 10 runs

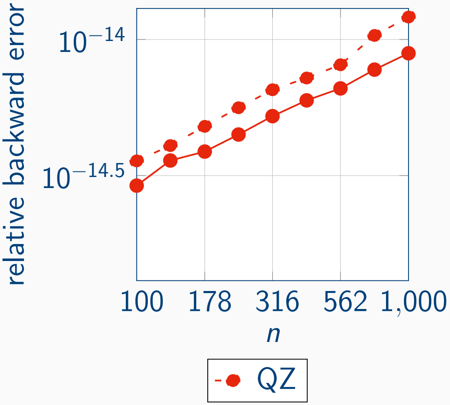
Numerical example

Is it worth it?



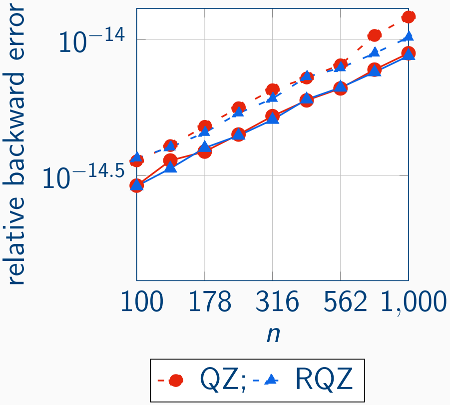
Data: 9 random matrix pairs, $n \in [100, 1000]$, reduced to H-T, averaged over 10 runs

Is it worth it?



Data: 9 random matrix pairs, $n \in [100, 1000]$, reduced to H-T, averaged over 10 runs

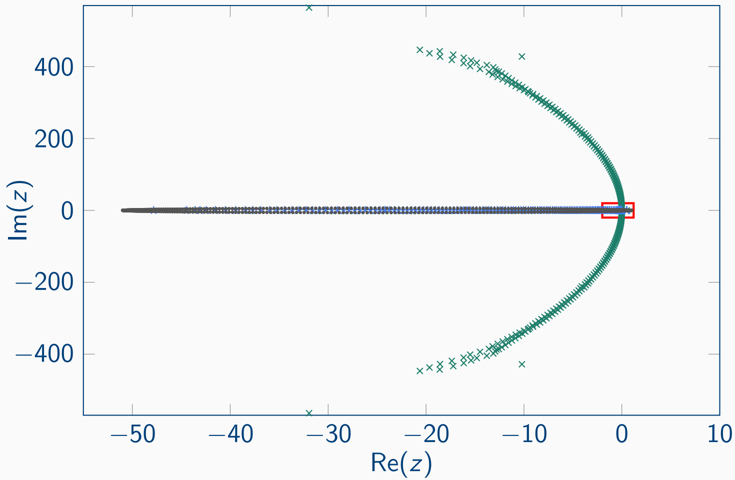
Is it worth it?



Data: 9 random matrix pairs, $n \in [100, 1000]$, reduced to H-T, averaged over 10 runs

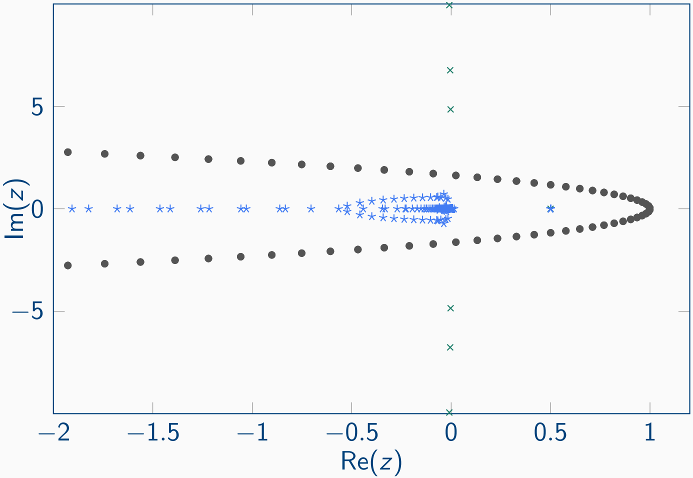
Numerical example 2: Reduction to Hessenberg, Hessenberg

Data: MHD matrix pair from MatrixMarket, $n = 1280$



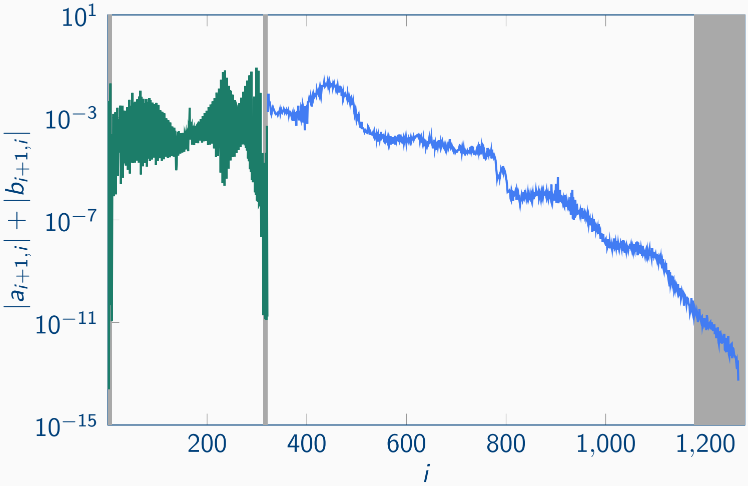
Numerical example 2: Reduction to Hessenberg, Hessenberg

Data: MHD matrix pair from MatrixMarket, $n = 1280$



Numerical example 2: Reduction to Hessenberg, Hessenberg

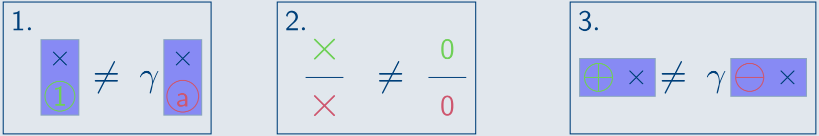
Data: MHD matrix pair from MatrixMarket, $n = 1280$



A word on the theory

Definition: Properness

The Hessenberg, Hessenberg pair (A, B) is called *proper* (or *irreducible*) if:



A word on the theory

Why and how does RQZ work?

Krylov subspaces

1. Krylov subspace: $\mathcal{K}_i(M, \boldsymbol{v}) = \mathcal{R}(\boldsymbol{v}, M\boldsymbol{v}, \ldots, M^{i-1}\boldsymbol{v})$

A word on the theory

Why and how does RQZ work?

Krylov subspaces

1. Krylov subspace: $\mathcal{K}_i(M, \boldsymbol{v}) = \mathcal{R}(\boldsymbol{v}, M\boldsymbol{v}, \ldots, M^{i-1}\boldsymbol{v})$
2. rational Krylov subspace: $\mathcal{K}_i^{\text{rat}}(M, \boldsymbol{v}, \Xi = (\xi_1, \ldots, \xi_{i-1})) = q_\Xi(M)^{-1}\, \mathcal{K}_i(M, \boldsymbol{v})$

A word on the theory

Why and how does RQZ work?

Krylov subspaces

1. Krylov subspace: $\mathcal{K}_i(M, \boldsymbol{v}) = \mathcal{R}(\boldsymbol{v}, M\boldsymbol{v}, \dots, M^{i-1}\boldsymbol{v})$
2. rational Krylov subspace: $\mathcal{K}_i^{\text{rat}}(M, \boldsymbol{v}, \Xi = (\xi_1, \dots, \xi_{i-1})) = q_\Xi(M)^{-1}\, \mathcal{K}_i(M, \boldsymbol{v})$

Theorem

If (A, B) is a proper Hessenberg pair with poles $(\xi_1, \dots, \xi_{n-1})$ then for $i = 1, \dots, n-1$:

$$\mathcal{K}_i^{\text{rat}}(AB^{-1}, \boldsymbol{e}_1, (\xi_1, \dots, \xi_{i-1})) = \mathcal{K}_i^{\text{rat}}(B^{-1}A, \boldsymbol{e}_1, (\xi_2, \dots, \xi_i)) = \mathcal{R}(\boldsymbol{e}_1, \dots, \boldsymbol{e}_i) = \mathcal{E}_i$$

A word on the theory

Why and how does RQZ work?

Theorem: Implicit Q (and Z)

Given a pair (A, B), the matrices Q and Z that transform it to proper Hessenberg form,

$$(\hat{A}, \hat{B}) = Q^* (A, B) Z,$$

are determined *essentially unique* if $Q\boldsymbol{e}_1$ and the poles are fixed.

A word on the theory

Why and how does RQZ work?

Nested subspace iteration

An RQZ step with shift ϱ on a pencil with poles $(\xi_1, \dots, \xi_{n-1})$ and new pole ξ_n performs nested subspace iteration for $i = 1, \dots, n-1$ accelerated by

$$\mathcal{R}(\boldsymbol{q}_1, \dots, \boldsymbol{q}_i) = (A - \varrho B)(A - \xi_i B)^{-1} \mathcal{E}_i$$
$$\mathcal{R}(\boldsymbol{z}_1, \dots, \boldsymbol{z}_i) = (A - \xi_{i+1} B)^{-1} (A - \varrho B) \mathcal{E}_i.$$

A word on the theory

Why and how does RQZ work?

Nested subspace iteration

An RQZ step with shift ϱ on a pencil with poles $(\xi_1, \ldots, \xi_{n-1})$ and new pole ξ_n performs nested subspace iteration for $i = 1, \ldots, n{-}1$ accelerated by

$$\mathcal{R}(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_i) = (A - \varrho B)(A - \xi_i B)^{-1}\mathcal{E}_i$$
$$\mathcal{R}(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_i) = (A - \xi_{i+1} B)^{-1}(A - \varrho B)\mathcal{E}_i.$$

What does this mean?

- QR step with shift ϱ on entire space

A word on the theory

Why and how does RQZ work?

Nested subspace iteration

An RQZ step with shift ϱ on a pencil with poles $(\xi_1, \ldots, \xi_{n-1})$ and new pole ξ_n performs nested subspace iteration for $i = 1, \ldots, n-1$ accelerated by

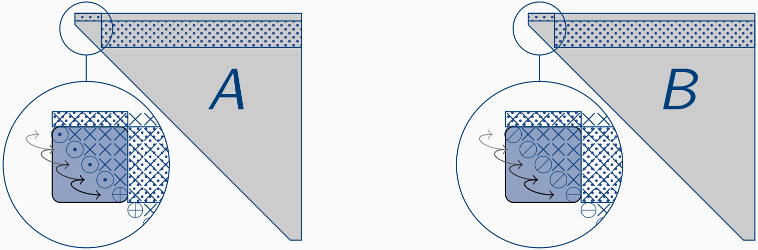
$$\mathcal{R}(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_i) = (A - \varrho B)(A - \xi_i B)^{-1}\mathcal{E}_i$$
$$\mathcal{R}(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_i) = (A - \xi_{i+1} B)^{-1}(A - \varrho B)\mathcal{E}_i.$$

What does this mean?

- QR step with shift ϱ on entire space
- RQ steps with tightly packed shifts Ξ on selected subspaces

Further extensions

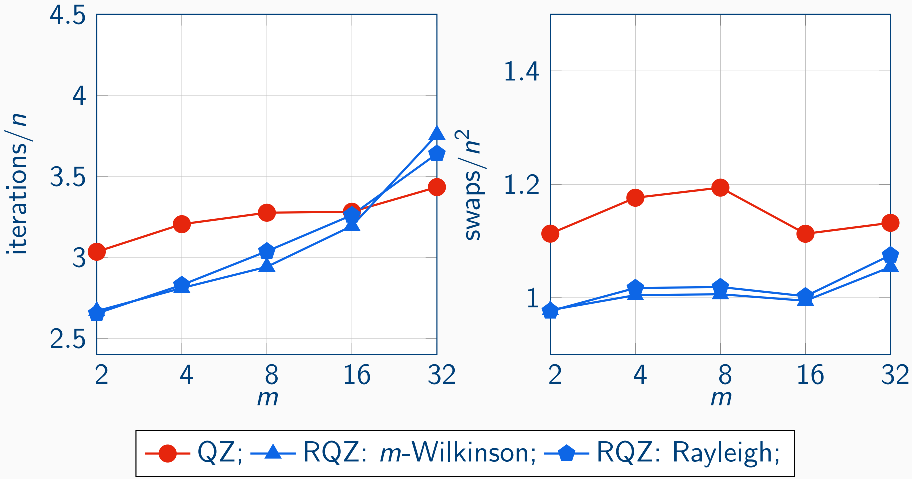
Tightly packed shifts



→ More cache efficient implementations (Level 3 BLAS)

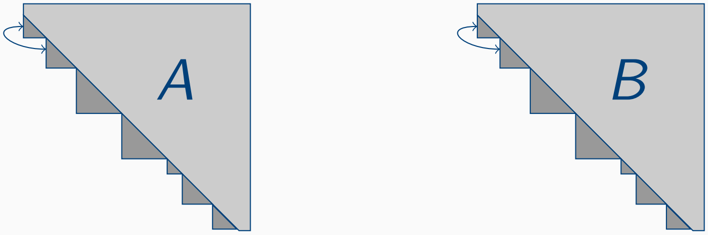
Further extensions

Tightly packed shifts



Further extensions

Block Hessenberg



$\rightarrow$ complex conjugate shifts and poles in real arithmetic for real pencils

Further extensions

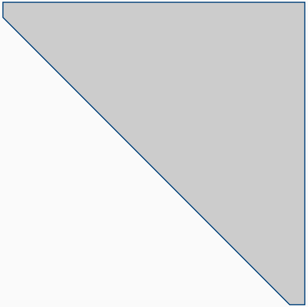
Aggressive Early deflation

The performance of the QR algorithm can be significantly improved by an aggressive early deflation technique ([Braman, Byers and Mathias]) and similar techniques have been developed for the QZ method ([Kågström and Kressner]).

Further extensions

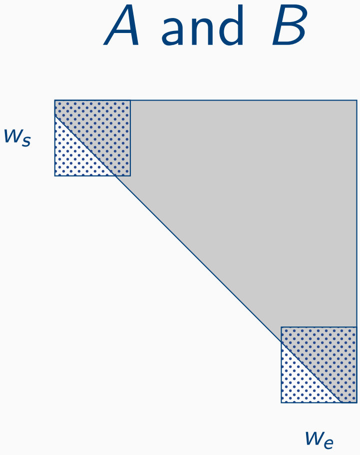
Aggressive Early deflation

A and B



Further extensions

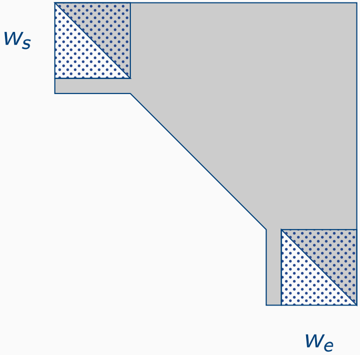
Aggressive Early deflation



$n \gg w_e > w_s$

Aggressive Early deflation

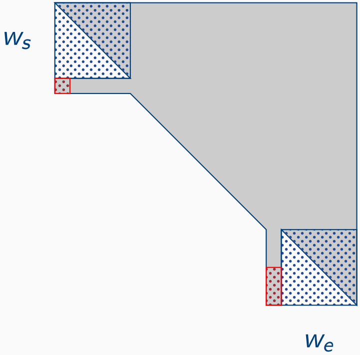
A and B



$$n \gg w_e > w_s$$

Aggressive Early deflation

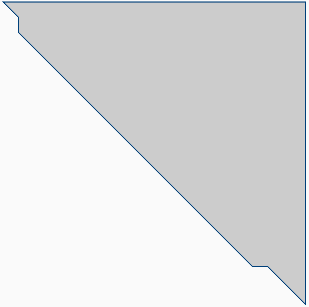
A and B



$$n \gg w_e > w_s$$

Aggressive Early deflation

A and B



Further extensions

Standard eigenvalue problems

- RQZ applies equivalence transformations on the pencil:

$$(\hat{A}, \hat{B}) = Q^*(A, B)Z$$

- Consequently we have two similarity transformations:

$$\hat{A}\hat{B}^{-1} = Q^*AB^{-1}Q \quad \text{and} \quad \hat{B}^{-1}\hat{A} = Z^*B^{-1}AZ$$

Further extensions

Standard eigenvalue problems

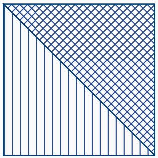
- RQZ applies equivalence transformations on the pencil:

$$(\hat{A}, \hat{B}) = Q^*(A, B)Z$$

- Consequently we have two similarity transformations:

$$\hat{A}\hat{B}^{-1} = Q^* A B^{-1} Q \quad \text{and} \quad \hat{B}^{-1}\hat{A} = Z^* B^{-1} A Z$$

$$\hat{A}\hat{B}^{-1} - diag(\alpha, \xi_1, \ldots, \xi_{n-1}) =$$



Further extensions

Standard eigenvalue problems

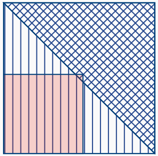
- RQZ applies equivalence transformations on the pencil:

$$(\hat{A}, \hat{B}) = Q^*(A, B)Z$$

- Consequently we have two similarity transformations:

$$\hat{A}\hat{B}^{-1} = Q^* A B^{-1} Q \quad \text{and} \quad \hat{B}^{-1}\hat{A} = Z^* B^{-1} A Z$$

$$\hat{A}\hat{B}^{-1} - diag(\alpha, \xi_1, \ldots, \xi_{n-1}) =$$



Further extensions

Standard eigenvalue problems

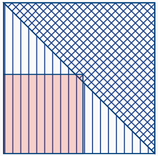
- RQZ applies equivalence transformations on the pencil:

$$(\hat{A}, \hat{B}) = Q^*(A, B)Z$$

- Consequently we have two similarity transformations:

$$\hat{A}\hat{B}^{-1} = Q^* A B^{-1} Q \quad \text{and} \quad \hat{B}^{-1}\hat{A} = Z^* B^{-1} A Z$$

$$\hat{A}\hat{B}^{-1} - diag(\alpha, \xi_1, \ldots, \xi_{n-1}) =$$



Extended Hessenberg + diagonal = rational Hessenberg

→ my connection with this mini-symposium

Extensions

Rational Krylov method

The connection between RQZ and the rational Krylov method can be used:

- to compute the Ritz values from the Hessenberg, Hessenberg recurrence pencil
- to filter and restart the rational Krylov method

Extensions

Rational Krylov method

$$A\, V_{k+1}\, \underline{G}_k = B\, V_{k+1}\, \underline{H}_k$$

with:

- $\mathcal{R}(V_{k+1}) = \mathcal{K}^{\text{rat}}_{k+1}(AB^{-1}, \boldsymbol{v}, \Xi_{1:k})$
- $(\underline{H}_k, \underline{G}_k)$ the Hessenberg, Hessenberg recurrence pencil

Applying an RQZ step with shift ϱ, we get $\mathcal{K}^{\text{rat}}_{k}(AB^{-1}, \hat{\boldsymbol{v}}, \Xi_{2:k})$ with:

$$\hat{\boldsymbol{v}} = (A - \xi_1\, B)^{-1}(A - \varrho\, B)\, \boldsymbol{v}$$

Conclusions

1. RQZ is a generalization of QZ
2. Implicit rational subspace iteration is promising
3. New shift and pole strategies can be a powerful tool to compute invariant subspaces

Further reading:

arXiv:1802.04094

`http://numa.cs.kuleuven.be/software/rqz/`

Conclusions

1. RQZ is a generalization of QZ
2. Implicit rational subspace iteration is promising
3. New shift and pole strategies can be a powerful tool to compute invariant subspaces

Further reading:

arXiv:1802.04094

`http://numa.cs.kuleuven.be/software/rqz/`

Thank you for your attention!