QRylov

Rational QZ Algorithm

Raf Vandebril

Raf.Vandebril@cs.kuleuven.be

Dept. of Computer Science, University of Leuven, Belgium

International Conference on
Approximation and Matrix Functions

New view on QR, Krylov, and Chasing

I hope you like it.

Combines/extends work of Watkins, Güttel & Berljafa.

This Lecture

- On computing eigenvalues of a single matrix A.
- I'll link three topics:
  1. Krylov Subspaces
  2. Matrix Structures
  3. QZ algorithms
- I'll do this for three spaces:
  1. Classical Krylov
  2. Extended Krylov
  3. Rational Krylov
- More on this topic
  - Camps, Meerbergen, Vandebril, *A rational QZ algorithm*, arXiv.
  - Pencils (A, B), deflations, properness, tightly packed bulges,...
  - Future work: multi-bulge, multi-pole, real arithmetic, LR variants, ...

Outline

Overview

The classic QR algorithm

- QR algorithm
- Krylov – Hessenberg
- Arnoldi – Implicit Q-Theorem – Uniqueness
- Implicit QR Algorithm: Classic
- Implicit QR Algorithm: Shift Swapping
- Convergence

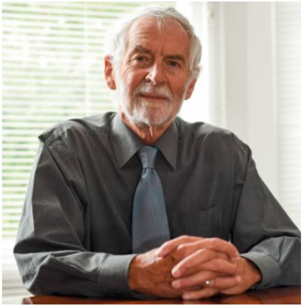
The extended QR algorithm

The rational QR algorithm

Some Conclusions

What about the QR algorithm

- The method of choice for solving 'small dense' eigenvalue problems.
- QR algorithm: 135.000 hits in Google.
- John Francis & Vera Kublanovskaya in the 60's.



- Introduction QR algorithm: 79.600 hits in Google.
- Typical way of introducing QR-type algorithms?

QR Algorithm with shift

$H \xrightarrow{\text{PHASE-II}} T$

QR Algorithm with shift:

$$
\begin{aligned}
&\text{for } k = 1:50\\
&\quad \text{determine shift } \mu_k\\
&\quad [Q,R] = qr(H - \mu_k I);\\
&\quad H = R * Q + \mu_k I;\\
&\text{end}
\end{aligned}
$$

a good choice $\mu_k = h_{n,n}^{(k)}$

```
>> eig(A)'
  11.1055   -3.8556    3.5736    0.1765
```

$a_{i,i-1} \to 0$ with rate $\left|\dfrac{\lambda_i}{\lambda_{i-1}}\right|$

Example:

$$\mu_k = h_{nn}$$

H =

1.0000	-5.0000	-1.5395	-1.2767
-6.0000	8.5556	0.1085	3.6986
0	-5.9182	-2.1689	-1.1428
0	0	-0.1428	3.6133

H1 =

3.3236	-5.5508	4.4353	2.6929
-8.0984	2.9128	-6.1528	-2.4688
0	-2.2250	1.1900	1.0296
0	0	-0.0017	3.5736

H2 =

2.4924	-9.7341	-2.5648	2.3834
-5.9560	5.0131	3.0854	-2.9507
0	-1.4702	-0.0791	0.0767
0	0	-0.0000	3.5736

deflation

References

1. Burden and Faires , Numerical Analysis , 8th Edition
2. Laurene V. Fausett, Applied Numerical Analysis Using MATLAB, 2nd Edition
3. Marco Latini ,The QR Algorithm, Past and Future
4. Wikipedia
5. Dr. Shafiu Jibrin, Numerical Analysis notes.

What about my talk – What is QR

- That's not how I will do it (I build upon Watkins' work).

- Requested: eigenvalues of A.
- Preprocessing, similarity of A to suitable form H.

$$\mathcal{K} \xrightarrow{\text{Subspace Iteration}} \widetilde{\mathcal{K}} = f(H)\mathcal{K}$$

- What will be discussed in this lecture.
  - $f(H)$ determines the convergence (E.g. power method $f(H) = H$).

What about my talk – What is QR

- That's not how I will do it (I build upon Watkins' work).

- Requested: eigenvalues of A.
- Preprocessing, similarity of A to suitable form H.

$$\mathcal{K} \xrightarrow{\text{Subspace Iteration}} \widetilde{\mathcal{K}} = f(H)\mathcal{K}$$

$$H \xrightarrow[\text{Explicit QR}]{} \tilde{H}$$

Implicit QR, bulge chasing

- What will be discussed in this lecture.
  - $f(H)$ determines the convergence (E.g. power method $f(H) = H$).
  - QR-step: pole chasing.

What about my talk – What is QR

- That's not how I will do it (I build upon Watkins' work).

- Requested: eigenvalues of A.
- Preprocessing, similarity of A to suitable form H.

$$\begin{tikzcd}[column sep=huge, row sep=huge]
\mathcal{K} \arrow[r, "\text{Subspace Iteration}"] \arrow[d, leftrightarrow] & \widetilde{\mathcal{K}} = f(H)\mathcal{K} \arrow[d, leftrightarrow] \\
H \arrow[r, "\text{Explicit QR}"'] & \tilde{H}
\end{tikzcd}$$

Implicit QR, bulge chasing

- What will be discussed in this lecture.
  - $f(H)$ determines the convergence (E.g. power method $f(H) = H$).
  - QR-step: pole chasing.
  - Link Krylov $\mathcal{K}$ – Structure of H (uniqueness: implicit Q-theorem).

Outline

Basis for the Krylov subspace

- Given matrix A and a (sufficiently random) starting vector v:

$$\begin{aligned} K(A,v) &= [v, Av, A^2v, A^3v, \ldots, A^{n-1}v], \\ \mathcal{K}(A,v) &= \operatorname{span} K(A,v). \end{aligned}$$

- We desire an orthonormal basis $Q = [q_1, \ldots, q_n]$ for $\mathcal{K}(A,v)$.
- Nestedeness is important:

$$\forall i : \ \operatorname{span}\{q_1, \ldots, q_i\} = \operatorname{span}\{v, Av, \ldots, A^{i-1}v\},$$

  so a gradual construction is required.
- Construct q_i out of the *continuation vector*.
  Options:
  - q_{i+1} constructed out of $A^i v$. (Numerical issues!!)
  - q_{i+1} constructed out of Aq_i. (Standard Arnoldi)
  - q_{i+1} constructed out of Ac for a 'good' $c \in \operatorname{span}\{q_1, \ldots, q_i\}$.

Standard Arnoldi

Classical Gram Schmidt:

1. Normalize v and store as q_1.
2. Make Aq_1 orthogonal to $\text{span}\{q_1\}$, store the rest as r.
   Normalize r and store as q_2.
3. Make Aq_2 orthogonal to $\text{span}\{q_1, q_2\}$, store the rest as r.
   Normalize r and store as q_3.
4. ...

More precisely

- Aq_i orthogonalized against $\text{span}\{q_1, \ldots, q_i\}$ to give rest r.
  Normalize r to get q_{i+1}.
- So Aq_i is a linear combination of $\{q_1, \ldots, q_i, q_{i+1}\}$.

In Matrix Language

- Example: construction of q_3 out of Aq_2.
  Relation for vector Aq_2 depending on q_1, q_2, q_3 $(n = 5)$:

$$Aq_2 = [q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times \\ \times \\ \times \\ 0 \\ 0 \end{bmatrix}.$$

- Matrix relation in detail

$$A[q_1, q_2, q_3, q_4, q_5] = [q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}.$$

- Matrix relation for H Hessenberg:

$$AQ = QH.$$

Arnoldi Process with Different Continuation Vector

Gram Schmidt with different continuation vector.

1. Normalize v and store as q_1.
2. Choose a good vector $c \in \text{span}\{q_1\}$.
   Make Ac orthogonal to $\text{span}\{q_1\}$, store the rest as r.
   Normalize r and store as q_2.
3. Choose a good vector $c \in \text{span}\{q_1, q_2\}$.
   Make Ac orthogonal to $\text{span}\{q_1, q_2\}$, store the rest as r.
   Normalize r and store as q_3.
4. ...

More precisely

- Take a 'good' linear combination $c \in \text{span}\{q_1, \ldots, q_i\}$.
  Ac orthogonalized against $\text{span}\{q_1, \ldots, q_i\}$ to give rest r.
  Normalize r to get q_{i+1}.
- c is linear combination of $\{q_1, \ldots, q_i\}$.
  Ac is a linear combination of $\{q_1, \ldots, q_i, q_{i+1}\}$.

In Matrix Language

- Example: construction of q_3 out of Ac, for 'good' $c \in \text{span}\{q_1, q_2\}$.
  Relation for Ac depending on q_1, q_2, q_3

$$Ac = [q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times \\ \times \\ \times \\ 0 \\ 0 \end{bmatrix}.$$

- Or, if we write out c:

$$A[q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times \\ \times \\ 0 \\ 0 \\ 0 \end{bmatrix} = [q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times \\ \times \\ \times \\ 0 \\ 0 \end{bmatrix}.$$

In Matrix Language

- Matrix relation in detail

$$
\begin{aligned}
& A[q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix} \\
= \quad & [q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}.
\end{aligned}
$$

- Matrix relation for K upper triangular and H Hessenberg:

$$AQK = QH.$$

- This links to the QZ algorithm: Hessenberg, upper-triangular pencil!

Outline

(Generalized) Eigenvalue Problems

- Eigenvalues of A solutions of $\det(A - \lambda I) = 0$.
- Standard Arnoldi – Classical eigenvalue problem
  - $AQ = QH$.
  - $Q^*AQ = H$ – Hessenberg matrix – single matrix form.
  - Eigenvalues: $\det(A - \lambda I) = \det(H - \lambda I) = 0$.

(Generalized) Eigenvalue Problems

- Eigenvalues of A solutions of $\det(A - \lambda I) = 0$.
- Standard Arnoldi – Classical eigenvalue problem
  - $AQ = QH$.
  - $Q^*AQ = H$ – Hessenberg matrix – single matrix form.
  - Eigenvalues: $\det(A - \lambda I) = \det(H - \lambda I) = 0$.
- Alternative continuation vector – Generalized eigenvalue problem
  - $AQK = QH$.
  - $Q^*AQ = HK^{-1}$ – Hessenberg matrix – single matrix form.
  - Eigenvalues: $\det(A - \lambda I) = \det(HK^{-1} - \lambda I) = \det(H - \lambda K)$.
- We name (H, K) a Hessenberg pair, since HK^{-1} is Hessenberg.
- We will compute eigenvalues of the Hessenberg pair.

What about uniqueness?

The Actors!

We have the relations

$$K(A, v) = [v, Av, \ldots, A^{n-1}v] = QR,$$

and

$$AQK = QH \quad \text{or} \quad Q^*AQ = HK^{-1}.$$

The actors are:

1. The matrix A;
2. unitary structure of Q;
3. starting vector $v = Qe_1$;
4. Krylov space $K(A, v) = \text{span}\{v, Av, A^2v, A^3v, \ldots\}$ (further on: generalizations);
5. Hessenberg structure of H, upper triangular structure of K;
6. contents (the elements) of HK^{-1};
7. contents (the elements) of Q.

Implicit H-Theorem (Arnoldi Process)

Theorem

Given A, v, and a unitary Q ($Qe_1 = v$) such that

$$K(A, v) = [v, Av, \ldots, A^{n-1}v] = QR,$$

is a QR factorization (i.e. Q forms a nested orthogonal basis).

Then we get

$$AQK = QH \quad \text{and} \quad Q^*AQ = HK^{-1},$$

with H Hessenberg and K upper triangular, with

- *(contents) Q essentially unique (uniqueness of QR factorization),*
- *(contents) HK^{-1} essentially unique.*

Implicit Q-Theorem

Theorem

Given A, v ($v = Qe_1$), and a unitary Q such that

$$AQK = QH \quad \textit{and} \quad Q^*AQ = HK^{-1},$$

with H Hessenberg and K upper triangular.

Then we get that

$$K(A, v) = [v, Av, \dots, A^{n-1}v] = QR,$$

is a QR factorization, with

- *(contents) Q essentially unique,*
- *(contents) HK^{-1} essentially unique.*

Classifying the Actors

Fixed constraints for both Theorems

1. The matrix A;
2. unitary structure of Q;
3. starting vector $v = Qe_1$.

Different constraints and different outcomes

Implicit H-Theorem

- $\mathcal{K}(A, v) = \text{span}\{v, Av, A^2v, \ldots\}$;
- $K(A, v) = QR$.

Implicit Q-Theorem

- Hessenberg pair (H, K);
- $AQK = QH$.

Fixed outcomes for both Theorems

1. essentially unique elements of HK^{-1};
2. essentially unique elements of Q.

Outline

Algorithmic Idea: Equivalences on Pencil

- Single step: Hessenberg pair (H, K) to new Hessenberg pair $(\tilde{H}, \tilde{K})$.
- In the end: the pair has to become upper triangular revealing the eigenvalues.
- Operating on the pencil (H, K) via equivalences

$$
\begin{array}{rcl}
(H, K) & \sim & HK^{-1} \\
(\tilde{H}, \tilde{K}) = Q^*(H, K)Z & \sim & \tilde{H}\tilde{K}^{-1} = Q^* HZ\, Z^* K^{-1} Q.
\end{array}
$$

- Implicit Q-theorem states:
  - Given first column of $Qe_1 = v$, given $A = HK^{-1}$,
  - imposing $(\tilde{H}, \tilde{K})$ to be a Hessenberg pair,
  - then the outcome $\tilde{H}\tilde{K}^{-1}$ is fixed.
- Implicit algorithm:
  - Construct Q and Z on the fly,
  - satisfying the constraints of the implicit Q-theorem.

Algorithmic Idea: Initializing

- Introduce perturbation via Q_0, with $Q_0 e_1 = Q e_1$
- Restore structure via Q_i, so that $Q = Q_0 Q_1 Q_2 Q_3$ (for our case $n = 5$). Q_1, Q_2, and Q_3 do not touch the first column anymore!
- More on Q_0: keep things simple (single shift, single bulge,...).
  - Structure of Q_0:

$$Q_0 = \qquad = \begin{bmatrix} \times & \times & & & \\ \times & \times & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix}.$$

  - We choose a particular, good $\mu \in \mathbb{C}$:

$$(HK^{-1} - \mu I)e_1 \approx Q_0 e_1.$$

  - The symbol $\approx$ means equal up to a scalar.

Initialization

- $Q_0 e_1 \approx (HK^{-1} - \mu I)e_1$.
- Q_0 is a rotator acting on rows 1 and 2.

$$Q_0^* H$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}$$

$$Q_0^* K$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}$$

Initialization

- $Q_0 e_1 \approx (HK^{-1} - \mu I)e_1$.
- Q_0 is a rotator acting on rows 1 and 2.

$$Q_0^* H \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}$$

$$Q_0^* K \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ \otimes & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}$$

Structure mismatch in row 2.

Chasing Step 1(a)

- Construct Z_1 to annihilate the right bulge.
- Z_1 works on columns 1 and 2.

$$Q_0^* H Z_1 \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}$$

$$Q_0^* K Z_1 \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ \otimes & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}$$

Chasing Step 1(a)

- Construct Z_1 to annihilate the right bulge.
- Z_1 works on columns 1 and 2.

$$Q_0^* H Z_1 \quad \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \otimes & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}$$

$$Q_0^* K Z_1 \quad \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}$$

Chasing Step 1(b)

- Construct Q_1 to annihilate the left bulge.
- Q_1 works on rows 2 and 3 (does not destroy $Q_0 e_1$).

$$Q_1^* Q_0^* H Z_1$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \otimes & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}$$

$$Q_1^* Q_0^* K Z_1$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}$$

Chasing Step 1(b)

- Construct Q_1 to annihilate the left bulge.
- Q_1 works on rows 2 and 3 (does not destroy $Q_0 e_1$)..

$$Q_1^* Q_0^* H Z_1 \quad \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}$$

$$Q_1^* Q_0^* K Z_1 \quad \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \otimes & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}$$

Structure mismatch in row 3.

Chasing Step 2

- Construct Z_2 to annihilate the right bulge (columns 2 and 3).
- Construct Q_2 to annihilate the left bulge (rows 3 and 4).

$$Q_2^* Q_1^* Q_0^* H Z_1 Z_2$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}$$

$$Q_2^* Q_1^* Q_0^* K Z_1 Z_2$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \otimes & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}$$

Chasing Step 2

- Construct Z_2 to annihilate the right bulge (columns 2 and 3).
- Construct Q_2 to annihilate the left bulge (rows 3 and 4).

$$Q_2^* Q_1^* Q_0^* H Z_1 Z_2 \qquad Q_2^* Q_1^* Q_0^* K Z_1 Z_2$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix} \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & \otimes & \times & \times \\ & & & & \times \end{bmatrix}$$

Structure mismatch in row 4.

Chasing Step 3

- Construct Z_3 to annihilate the right bulge (columns 3 and 4).
- Construct Q_3 to annihilate the left bulge (rows 4 and 5).

$$Q_3^* Q_2^* Q_1^* Q_0^* H Z_1 Z_2 Z_3$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}$$

$$Q_3^* Q_2^* Q_1^* Q_0^* K Z_1 Z_2 Z_3$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & \otimes & \times \end{bmatrix}$$

Structure mismatch in row 5.

Finalization

- Construct Z_4 to annihilate the right bulge.
- Z_4 operates on columns 4 and 5.

$$Q_3^* Q_2^* Q_1^* Q_0^* H Z_1 Z_2 Z_3 Z_4 \qquad Q_3^* Q_2^* Q_1^* Q_0^* K Z_1 Z_2 Z_3 Z_4$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix} \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}$$

Desired structure obtained: implicit QZ step executed.

Outline

Swapping Diagonal Elements

- Consider the upper triangular pencil

$$\left(\begin{bmatrix} \alpha_1 & \times \\ & \alpha_2 \end{bmatrix}, \begin{bmatrix} \beta_1 & \times \\ & \beta_2 \end{bmatrix} \right).$$

- Suppose $\frac{\alpha_1}{\beta_1} \neq \frac{\alpha_2}{\beta_2}$,
- then we can construct Q and Z such that

$$Q^* \left(\begin{bmatrix} \alpha_1 & \times \\ & \alpha_2 \end{bmatrix}, \begin{bmatrix} \beta_1 & \times \\ & \beta_2 \end{bmatrix} \right) Z = \left(\begin{bmatrix} \tilde{\alpha}_2 & \times \\ & \tilde{\alpha}_1 \end{bmatrix}, \begin{bmatrix} \tilde{\beta}_2 & \times \\ & \tilde{\beta}_1 \end{bmatrix} \right),$$

- with

$$\frac{\alpha_1}{\beta_1} = \frac{\tilde{\alpha}_1}{\tilde{\beta}_1} \qquad \frac{\alpha_2}{\beta_2} = \frac{\tilde{\alpha}_2}{\tilde{\beta}_2}.$$

- The ratios of the diagonal elements are swapped.
(Cfr. reordering eigenvalues in the Schur form.)

Alternative Notation

- Simplified notation:

$$Q^* \left(\begin{bmatrix} ① & \times \\ & ② \end{bmatrix}, \begin{bmatrix} ❶ & \times \\ & ❷ \end{bmatrix} \right) Z = \left(\begin{bmatrix} ② & \times \\ & ① \end{bmatrix}, \begin{bmatrix} ❷ & \times \\ & ❶ \end{bmatrix} \right).$$

- The elements left (①,❶) and right (①,❶) differ.
  The elements left (②,❷) and right (②,❷) differ.

- But, the ratios remain the same. For both left and right:

$$\frac{①}{❶} = \alpha \quad \text{and} \quad \frac{②}{❷} = \beta,$$

  α and β could be 0, $\in \mathbb{C}$, or ∞.

- The ratios, i.e. eigenvalues, have been swapped.
  (Cfr. bulge pencils, as introduced by Watkins.)

Reconsidering the Initializing

- More on Q_0:
  - We take a particular choice, for a chosen μ:
$$Q_0 e_1 \approx (HK^{-1} - \mu I) e_1.$$
  - Since $Ke_1 \approx e_1$ (K is upper triangular):
$$Q_0 e_1 \approx (H - \mu K) e_1.$$
  - Rewriting
$$e_1 \approx (Q_0^* H - \mu Q_0^* K) e_1.$$
  - Considering the second element of the vector we get
$$(Q_0^* H)_{21} - \mu (Q_0^* K)_{21} = 0.$$
  - Or we get as ratios of the subdiagonal elements:
$$\frac{(Q_0^* H)_{21}}{(Q_0^* K)_{21}} = \mu.$$

Initialization: Introducing a Shift

- Given

$$\left(\begin{bmatrix} \times & \times \\ \text{①} & \times \\ & \text{②} \end{bmatrix}, \begin{bmatrix} \times & \times \\ \text{❶} & \times \\ & \text{❷} \end{bmatrix}\right)$$

- There exists Q_0 (with $Q_0 e_1 = (H - \mu K)e_1$), such that

$$Q_0^* \left(\begin{bmatrix} \times & \times \\ \text{①} & \times \\ & \text{②} \end{bmatrix}, \begin{bmatrix} \times & \times \\ \text{❶} & \times \\ & \text{❷} \end{bmatrix}\right) = \left(\begin{bmatrix} \times & \times \\ \mu & \times \\ & \text{②} \end{bmatrix}, \begin{bmatrix} \times & \times \\ \mu & \times \\ & \text{❷} \end{bmatrix}\right)$$

- Elements ① and ❶ are replaced by μ and μ.
- We have new elements whose ratio satisfies

$$\frac{\mu}{\mu} = \mu.$$

- Comment: in fact we do not need $(H - \mu K)e_1$, Q_0 can be computed directly.

Original

- Original matrix: all elements ❶ = ❷ = … = 0
- Hence ①/❶ = ②/❷ = … = ∞.

$$H \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ ① & \times & \times & \times & \times \\ & ② & \times & \times & \times \\ & & ③ & \times & \times \\ & & & ④ & \times \end{bmatrix}$$

$$K \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & 0 & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

Initialization

- Q_0 constructed to introduce the shift μ.
- Q_0 acts on rows 1 and 2 and introduces the shift μ.

$$Q_0^* H$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \mu & \times & \times & \times & \times \\ & 2 & \times & \times & \times \\ & & 3 & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

$$Q_0^* K$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \mu & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & 0 & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

Shift present in row 2.

Initialization

- Q_0 constructed to introduce the shift μ.
- Q_0 acts on rows 1 and 2 and introduces the shift μ.

$Q_0^* H$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \mu & \times & \times & \times & \times \\ & 2 & \times & \times & \times \\ & & 3 & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

$Q_0^* K$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \mu & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & 0 & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

Red block subjected to swapping!

Chase step 1

- The shift μ moved from row 2 to row 3.
- Q_1 and Z_1 execute swap.

$$Q_1^* Q_0^* H Z_1$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & \mu & \times & \times & \times \\ & & 3 & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

$$Q_1^* Q_0^* K Z_1$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & \mu & \times & \times & \times \\ & & 0 & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

Shift located in row 3, red block needs swapping!

Chase step 2

- The shift μ moved from row 3 to row 4.
- Q_2 and Z_2 execute swap.

$$Q_2^* Q_1^* Q_0^* H Z_1 Z_2$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & \mu & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

$$Q_2^* Q_1^* Q_0^* K Z_1 Z_2$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & \mu & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

Shift located in row 4, red block needs swapping!

Chase step 3

- The shift μ moved from row 4 to row 5.
- Q_3 and Z_3 execute swap.

$$Q_3^* Q_2^* Q_1^* Q_0^* H Z_1 Z_2 Z_3$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & 4 & \times & \times \\ & & & \mu & \times \end{bmatrix}$$

$$Q_3^* Q_2^* Q_1^* Q_0^* K Z_1 Z_2 Z_3$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & 0 & \times & \times \\ & & & \mu & \times \end{bmatrix}$$

Shift located in row 5!

Finalization

- Remove the shift and restore the Hessenberg structure.
- Z_4 operates on columns 4 and 5, and removes the shift.

$$Q_3^* Q_2^* Q_1^* Q_0^* H Z_1 Z_2 Z_3 Z_4$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & 4 & \times & \times \\ & & & 5 & \times \end{bmatrix}$$

$$Q_3^* Q_2^* Q_1^* Q_0^* K Z_1 Z_2 Z_3 Z_4$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & 0 & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

Shift is removed!

Outline

Convergence Analysis

- Convergence is subspace iteration with a basis transformation.
- Subspace iteration determined by the initialization:

$$Q_0 e_1 = (H - \mu K)e_1 = (HK^{-1} - \mu I)e_1.$$

- Subspace iteration driven by $(HK^{-1} - \mu I)$.
- Convergence: lower right corner of $\tilde{H}\tilde{K}^{-1}$ gets pushed to μ.

Outline

What about extended Krylov

- Extended Krylov: 8.250 hits in Google.
- V. Druskin & L. Knizhnerman.

- Jagels, Reichel, Simoncini, ...
- Applications
  - matrix functions,
  - matrix equations.

Basis for the extended Krylov subspace

- Given matrix A and a starting vector v:

$$\begin{aligned} K_E(A,v) &= [v, A^{-1}v, Av, A^{-2}v, A^2v, \ldots]. \\ \mathcal{K}_E(A,v) &= \operatorname{span} K_E(A,v). \end{aligned}$$

- The order of positive and negative powers can be chosen freely.
  We can have a succession of negative powers, followed by positive powers,...
- We desire an orthonormal basis $Q = [q_1, \ldots, q_n]$ for $\mathcal{K}_E(A,v)$.
- Construct q_{i+1} out of the *continuation vector* $c \in \operatorname{span}\{q_1, \ldots, q_i\}$.
  Two possibilities for the Gram Schmidt procedure:
  - Positive power: q_{i+1} constructed out of Ac.
  - Negative power: q_{i+1} constructed out of $A^{-1}c$.

In Matrix Language

More precisely for a positive power (same as before)

- Take good $c \in \text{span}\{q_1, \dots, q_i\}$.
  Ac orthogonalized against $\text{span}\{q_1, \dots, q_i\}$ to give rest r.
  Normalize r to get q_{i+1}.
- So $Ac \in \text{span}\{q_1, \dots, q_i, q_{i+1}\}$.

Example: construction of q_3 out of Ac, $c \in \text{span}\{q_1, q_2\}$.

- Relation for vector Ac depending on q_1, q_2, q_3 ($n = 5$):

$$Ac = A[q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times \\ \times \\ 0 \\ 0 \\ 0 \end{bmatrix} = [q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times \\ \times \\ \times \\ 0 \\ 0 \end{bmatrix}.$$

In matrix Language

More precisely for a negative power

- Take good $c \in \operatorname{span}\{q_1, \ldots, q_i\}$.
  $A^{-1}c$ orthogonalized against $\operatorname{span}\{q_1, \ldots, q_i\}$ to give rest r.
  Normalize r to get q_{i+1}.
- So $A^{-1}c \in \operatorname{span}\{q_1, \ldots, q_i, q_{i+1}\}$.
- Or $c \in \operatorname{span}\{Aq_1, \ldots, Aq_{i+1}\}$.

Example on the construction of q_4.

- $c \in \operatorname{span}\{q_1, q_2, q_3\}$.
  $c \in \operatorname{span}\{Aq_1, Aq_2, Aq_3, Aq_4\}$.

$$A[q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times \\ \times \\ \times \\ \times \\ 0 \end{bmatrix} = c = [q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times \\ \times \\ \times \\ 0 \\ 0 \end{bmatrix}.$$

In matrix Language

- Matrix relation in detail

$$
\begin{aligned}
A[q_1, q_2, q_3, q_4, q_5] & \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & \times & \times & \times \\ & & & & \times \end{bmatrix} \\
= \quad [q_1, q_2, q_3, q_4, q_5] & \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & & \times & \times \\ & & & \times & \times \end{bmatrix},
\end{aligned}
$$

- Positive power: Hessenberg right, upper triangular left.

In matrix Language

- Matrix relation in detail

$$A[q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & \times & \times & \times \\ & & & & \times \end{bmatrix}$$

$$= \quad [q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & & \times & \times \\ & & & \times & \times \end{bmatrix},$$

- Negative power: Hessenberg left, upper triangular right.

In matrix Language

- Matrix relation in detail

$$
\begin{aligned}
& A[q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & \times & \times & \times \\ & & & 0 & \times \end{bmatrix} \\
= \quad & [q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & 0 & \times & \times \\ & & & \times & \times \end{bmatrix}.
\end{aligned}
$$

- Zero on the left OR on the right!
- We name (H, K) and extended Hessenberg pair. We have

$$AQK = QH.$$

Outline

Implicit Theorems

Fixed constraints for both Theorems

1. The matrix A;
2. unitary structure of Q;
3. starting vector $v = Qe_1$.

Different constraints and different outcomes

Implicit H-Theorem

- $\mathcal{K}_E(A, v) = \text{span}\{v, A^{-1}v, Av, A^{-2}v, \ldots\}$;
- $K_E(A, v) = QR$.

Implicit Q-Theorem

- Extended pair (H, K);
- $AQK = QH$.

Fixed outcomes for both Theorems

1. essentially unique elements of HK^{-1};
2. essentially unique elements of Q.

Outline

Original

- We have ❶ $= 0$ or ⓘ $= 0$, for $i = 1, \ldots, 4$.
- Hence ⓘ/❶ equals ∞ or 0.

$$H \quad \begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & ② & \times & \times & \times \\ & & 0 & \times & \times \\ & & & ④ & \times \end{bmatrix}$$

$$K \quad \begin{bmatrix} \times & \times & \times & \times & \times \\ ❶ & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & ❸ & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

Algorithm & Initialization

- From extended pair (H, K) to new extended pair $(\tilde{H}, \tilde{K})$ implicitly.
- We initialize the iteration, for a chosen μ:

$$Q_0 e_1 \approx (H - \mu K)e_1, \quad \text{or} \quad e_1 \approx (Q_0^* H - \mu Q_0^* K)e_1,$$

  just as before.
- Clearly Q_0 is of the correct form: only operates on rows 1 and 2.
- We get as ratios of the subdiagonal elements:

$$\frac{(Q_0^* H)_{21}}{(Q_0^* K)_{21}} = \mu.$$

- Q_0 introduces the shift μ as a ratio on the subdiagonal.

Initialization

- Q_0 constructed to introduce the shift μ.
- Q_0 acts on rows 1 and 2.

$$Q_0^* H \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ \mu & \times & \times & \times & \times \\ & 2 & \times & \times & \times \\ & & 0 & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

$$Q_0^* K \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ \mu & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & 3 & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

Shift present in row 2.

Initialization

- Q_0 constructed to introduce the shift μ.
- Q_0 acts on rows 1 and 2.

$$Q_0^* H \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ \mu & \times & \times & \times & \times \\ & 2 & \times & \times & \times \\ & & 0 & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

$$Q_0^* K \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ \mu & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & 3 & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

Red block subjected to swapping!

Chase step 1

- The shift μ moved from row 2 to row 3.
- Q_1 and Z_1 execute swap.

$$Q_1^* Q_0^* H Z_1$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & \mu & \times & \times & \times \\ & & 0 & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

$$Q_1^* Q_0^* K Z_1$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & \mu & \times & \times & \times \\ & & 3 & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

Shift located in row 3, red block needs swapping!

Chase step 2

- The shift μ moved from row 3 to row 4.
- Q_2 and Z_2 execute swap.

$$Q_2^* Q_1^* Q_0^* H Z_1 Z_2$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & \color{red}{\mu} & \color{red}{\times} & \times \\ & & & \color{red}{4} & \times \end{bmatrix}$$

$$Q_2^* Q_1^* Q_0^* K Z_1 Z_2$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & \color{red}{\mu} & \color{red}{\times} & \times \\ & & & \color{red}{0} & \times \end{bmatrix}$$

Shift located in row 4, red block needs swapping!

Chase step 3

- The shift μ moved from row 4 to row 5.
- Q_3 and Z_3 execute swap.

$$Q_3^* Q_2^* Q_1^* Q_0^* H Z_1 Z_2 Z_3$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & 4 & \times & \times \\ & & & \mu & \times \end{bmatrix}$$

$$Q_3^* Q_2^* Q_1^* Q_0^* K Z_1 Z_2 Z_3$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & 0 & \times & \times \\ & & & \mu & \times \end{bmatrix}$$

Shift located in row 5!

Finalization Option 1

- Remove the shift and restore extended Hessenberg structure.
- Z_4 operates on columns 4 and 5, and removes the shift.

$$Q_3^* Q_2^* Q_1^* Q_0^* H Z_1 Z_2 Z_3 Z_4$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & 4 & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

$$Q_3^* Q_2^* Q_1^* Q_0^* K Z_1 Z_2 Z_3 Z_4$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & 0 & \times & \times \\ & & & 5 & \times \end{bmatrix}$$

Shift is removed! Extended structure restored!

Finalization Option 2

- Remove the shift and restore extended Hessenberg structure.
- Z_4 operates on columns 4 and 5, and removes the shift.

$$Q_3^* Q_2^* Q_1^* Q_0^* H Z_1 Z_2 Z_3 Z_4$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & 4 & \times & \times \\ & & & 5 & \times \end{bmatrix}$$

$$Q_3^* Q_2^* Q_1^* Q_0^* K Z_1 Z_2 Z_3 Z_4$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & 0 & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

Shift is removed! Extended structure restored!

Conclusions

- Algorithm almost the same as the classical QR.
- Identical to the bulge hopping algorithm of Vandebril & Watkins.
- Introduce μ as ratio of subdiagonal elements.
- Chase μ by swapping subdiagonal elements.
- Finalization is different:
  - In the QR case: no options.
  - Here two options: position a zero left or right.
    This is equivalent to introducing a ratio ∞ or 0.
    Both are extended Hessenberg pairs, so that's fine.
- This choice has an influence on the convergence.
  Since forthcoming associated Krylov subspaces (next step) have changed!

Outline

Convergence Analysis

- Subspace iteration determined by the initialization:

$$Q_0 e_1 = (H - \mu K)e_1 \approx (HK^{-1})^{-1}(HK^{-1} - \mu I)e_1.$$

- Subspace iteration driven by $(HK^{-1})^{-1}(HK^{-1} - \mu I)$.
- Convergence:
  - lower right corner of $\tilde{H}\tilde{K}^{-1}$ gets pushed to μ (fast),
  - upper right corner gets pushed to 0 (slow).

Outline

What about rational Krylov

- Rational Krylov: 119.000 hits in Google.
- Introduced by Ruhe

- Beckermann, Güttel, Berljafa, Meerbergen, ...
- Applications
  - model reduction,
  - flexibility because of the pole selection.

Basis for the rational Krylov subspace

- Given matrix A and a (sufficiently random) starting vector v:

$$\begin{aligned} K_R(A,v) &= [v, (A-\sigma_1 I)^{-1}v, (A-\sigma_2 I)^{-1}(A-\sigma_1 I)^{-1}v, \ldots], \\ \mathcal{K}_R(A,v) &= \operatorname{span} K_R(A,v). \end{aligned}$$

- The σ_i's, named poles, can be chosen freely (can even be ∞, i.e., Av).
- We desire an orthonormal basis $Q = [q_1, \ldots, q_n]$ for $\mathcal{K}_R(A,v)$.
- Construct q_{i+1} out of the *continuation vector* $c \in \operatorname{span}\{q_1, \ldots, q_i\}$
  - For $\sigma_i \neq \infty$: q_{i+1} constructed out of $(A-\sigma_i I)^{-1}c$.
    For $\sigma_i = 0$: q_{i+1} constructed out of $A^{-1}c$. (Same as before.)
  - For $\sigma_i = \infty$: q_{i+1} constructed out of Ac. (Same as before.)

In Matrix Language

Generic relation between the vectors.

- $c \in \text{span}\{q_1, \ldots, q_i\}$.
  $(A - \sigma_i I)^{-1} c$ orthogonalized against $\text{span}\{q_1, \ldots, q_i\}$ to give rest r.
  Normalize r to get q_{i+1}.
- So $(A - \sigma_i I)^{-1} c \in \text{span}\{q_1, \ldots, q_i, q_{i+1}\}$.
- Or $c \in \text{span}\{(A - \sigma_i I) q_1, \ldots, (A - \sigma_i I) q_{i+1}\}$.
- This implies a relation between

$$[q_1, \ldots, q_{i+1}] \quad \text{and} \quad [Aq_1, \ldots, Aq_{i+1}].$$

Example: compute q_3 from Ac

- $c \in \text{span}\{q_1, q_2\}$.
- $(A - \sigma_2 I)^{-1} c \in \text{span}\{q_1, q_2, q_3\}$.
- $c \in \ (A - \sigma_2 I)\,\text{span}\{q_1, q_2, q_3\}$.
- We have

$$c = [q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \times \\ \times \\ 0 \\ 0 \\ 0 \end{bmatrix} = (A - \sigma_2 I)[q_1, q_2, q_3, q_4, q_5] \begin{bmatrix} \otimes \\ \otimes \\ \otimes \\ 0 \\ 0 \end{bmatrix}$$

- Or

$$(c =)\ Q \begin{bmatrix} \times \\ \times \\ 0 \\ 0 \\ 0 \end{bmatrix} = AQ \begin{bmatrix} \otimes \\ \otimes \\ \otimes \\ 0 \\ 0 \end{bmatrix} - \sigma_2 Q \begin{bmatrix} \otimes \\ \otimes \\ \otimes \\ 0 \\ 0 \end{bmatrix}$$

Example: compute q_3

- We had

$$Q\begin{bmatrix}\times\\\times\\0\\0\\0\end{bmatrix} = AQ\begin{bmatrix}\otimes\\\otimes\\\otimes\\0\\0\end{bmatrix} - \sigma_2 Q\begin{bmatrix}\otimes\\\otimes\\\otimes\\0\\0\end{bmatrix}$$

- Rewritten we get

$$AQ\begin{bmatrix}\otimes\\\otimes\\\otimes\\0\\0\end{bmatrix} = Q\begin{bmatrix}\times+\sigma_2\otimes\\\times+\sigma_2\otimes\\\sigma_2\otimes\\0\\0\end{bmatrix}.$$

- Note that we can extract the pole!

In matrix Language

- Matrix relation in detail

$$AQ \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix} = Q \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix},$$

- Matrix relation for K and H Hessenberg

$$AQK = QH.$$

- Poles as ratios of subdiagonal elements.
- We name (H, K) a rational Hessenberg pair.

Outline

Implicit Theorems

Fixed constraints for both Theorems

1. The matrix A;
2. unitary structure of Q;
3. starting vector $v = Qe_1$.

Different constraints and different outcomes

Implicit H-Theorem

- $\mathcal{K}_R(A, v) = \text{span}\{v, (A - \sigma_1 I)^{-1} v, \ldots\}$;
- $K_R(A, v) = QR$.

Implicit Q-Theorem

- Rational pair (H, K);
- $AQK = QH$.

Fixed outcomes for both Theorems

1. essentially unique elements of HK^{-1};
2. essentially unique elements of Q.

Summary

Krylov

- $\mathcal{K}(A, v) = \text{span}\{v, Av, A^2v, A^3v, \dots, \}$,
- Hessenberg pair (H, K).

Extended Krylov

- $\mathcal{K}_E(A, v) = \text{span}\{v, A^{-1}v, Av, A^{-2}v, A^2v, \dots, \}$,
- Extended Hessenberg pair (H, K).

Most general:

Rational Krylov

- $\mathcal{K}_R(A, v) = \text{span}\{v, (A - \sigma_1 I)^{-1}v, (A - \sigma_2 I)^{-1}(A - \sigma_1 I)^{-1}v, \dots, \}$,
- Rational Hessenberg pair (H, K).

Outline

Original

- We have poles as ratios of the subdiagonal elements.
- We use the following notation ①/❶ $= \sigma_1$, and so forth.

$$H$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ ① & \times & \times & \times & \times \\ & ② & \times & \times & \times \\ & & ③ & \times & \times \\ & & & ④ & \times \end{bmatrix}$$

$$K$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ ❶ & \times & \times & \times & \times \\ & ❷ & \times & \times & \times \\ & & ❸ & \times & \times \\ & & & ❹ & \times \end{bmatrix}$$

Algorithm & Initialization

- From rational pair (H, K) to new rational pair $(\tilde{H}, \tilde{K})$ implicit.
- We initialize the iteration for a chosen μ:

$$Q_0 e_1 \approx (H - \mu K) e_1,$$

  just as before.
- All relations still hold.
- Pole changing algorithm of Güttel and Berljafa.

Initialization

- Q_0 constructed to introduce the shift μ.
- Q_0 acts on rows 1 and 2.

$$Q_0^* H$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \mu & \times & \times & \times & \times \\ & 2 & \times & \times & \times \\ & & 3 & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

$$Q_0^* K$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \mu & \times & \times & \times & \times \\ & 2 & \times & \times & \times \\ & & 3 & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

Shift present in row 2.

Initialization

- Q_0 constructed to introduce the shift μ.
- Q_0 acts on rows 1 and 2.

$$Q_0^* H \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ \mu & \times & \times & \times & \times \\ & 2 & \times & \times & \times \\ & & 3 & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

$$Q_0^* K \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ \mu & \times & \times & \times & \times \\ & 2 & \times & \times & \times \\ & & 3 & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

Red block subjected to swapping!

Chase step 1

- The shift μ moved from row 2 to row 3.
- Q_1 and Z_1 execute swap.

$$Q_1^* Q_0^* H Z_1 \qquad\qquad Q_1^* Q_0^* K Z_1$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & \mu & \times & \times & \times \\ & & 3 & \times & \times \\ & & & 4 & \times \end{bmatrix} \qquad \begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & \mu & \times & \times & \times \\ & & 3 & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

Shift located in row 3, red block needs swapping!

Chase step 2

- The shift μ moved from row 3 to row 4.
- Q_2 and Z_2 execute swap.

$$Q_2^* Q_1^* Q_0^* H Z_1 Z_2$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & \mu & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

$$Q_2^* Q_1^* Q_0^* K Z_1 Z_2$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & \mu & \times & \times \\ & & & 4 & \times \end{bmatrix}$$

Shift located in row 4, red block needs swapping!

Chase step 3

- The shift μ moved from row 4 to row 5.
- Q_3 and Z_3 execute swap.

$$Q_3^* Q_2^* Q_1^* Q_0^* H Z_1 Z_2 Z_3$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & 4 & \times & \times \\ & & & \mu & \times \end{bmatrix}$$

$$Q_3^* Q_2^* Q_1^* Q_0^* K Z_1 Z_2 Z_3$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & 4 & \times & \times \\ & & & \mu & \times \end{bmatrix}$$

Shift located in row 5!

Finalization Option 1

- Remove the shift and restore rational Hessenberg structure.
- Z_4 operates on columns 4 and 5, and removes the shift.

$$Q_3^* Q_2^* Q_1^* Q_0^* H Z_1 Z_2 Z_3 Z_4$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & 4 & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

$$Q_3^* Q_2^* Q_1^* Q_0^* K Z_1 Z_2 Z_3 Z_4$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & 4 & \times & \times \\ & & & 5 & \times \end{bmatrix}$$

Shift removed! Structure Restored! Pole 0 introduced!

Finalization Option 2

- Remove the shift and restore rational Hessenberg structure.
- Z_4 operates on columns 4 and 5, and removes the shift.

$$Q_3^* Q_2^* Q_1^* Q_0^* H Z_1 Z_2 Z_3 Z_4$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & 4 & \times & \times \\ & & & 5 & \times \end{bmatrix}$$

$$Q_3^* Q_2^* Q_1^* Q_0^* K Z_1 Z_2 Z_3 Z_4$$

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 2 & \times & \times & \times & \times \\ & 3 & \times & \times & \times \\ & & 4 & \times & \times \\ & & & 0 & \times \end{bmatrix}$$

Shift removed! Structure Restored! Pole ∞ introduced!

Finalization Option 3

- Remove the shift and restore rational Hessenberg structure.
- Z_4 operates on columns 4 and 5, and removes the shift.

$$Q_3^* Q_2^* Q_1^* Q_0^* H Z_1 Z_2 Z_3 Z_4 \quad \begin{bmatrix} \times & \times & \times & \times & \times \\ ② & \times & \times & \times & \times \\ & ③ & \times & \times & \times \\ & & ④ & \times & \times \\ & & & ⑤ & \times \end{bmatrix}$$

$$Q_3^* Q_2^* Q_1^* Q_0^* K Z_1 Z_2 Z_3 Z_4 \quad \begin{bmatrix} \times & \times & \times & \times & \times \\ ❷ & \times & \times & \times & \times \\ & ❸ & \times & \times & \times \\ & & ❹ & \times & \times \\ & & & ❺ & \times \end{bmatrix}$$

Shift removed! Structure Restored! Pole σ_5 introduced!

Outline

Convergence Analysis

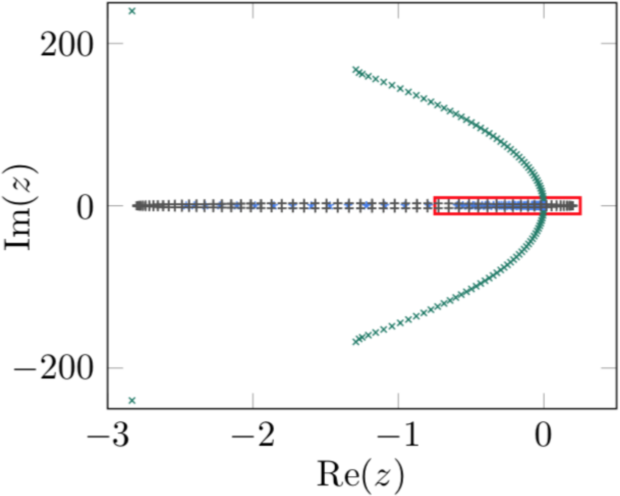
- Subspace iteration determined by the initialization:

$$Q_0 e_1 = (H - \mu K) e_1 \approx (HK^{-1} - \sigma_1)^{-1} \left(HK^{-1} - \mu I\right) e_1.$$

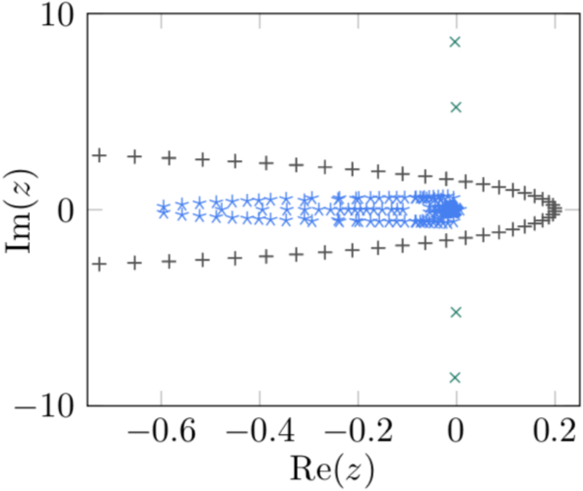
- Subspace iteration driven by $(HK^{-1} - \sigma_1)^{-1} \left(HK^{-1} - \mu I\right)$.
- Note, σ_1 is pushed off, so next time, next driving function.
- Convergence:
  - lower right corner of $\tilde{H}\tilde{K}^{-1}$ gets pushed to μ (fast),
  - upper right corner gets pushed to σ_i (slow).

Numerical Experiment

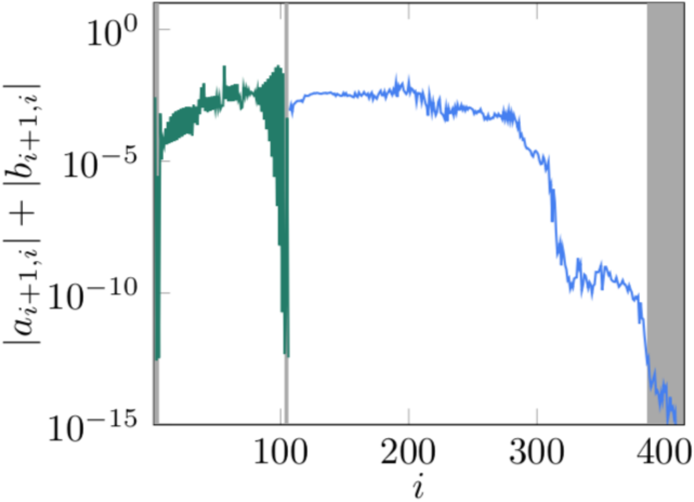
- 5 % – 10 % faster than the classical QZ, for standard pole choice.
- More interesting is the link with contour integration.



Numerical Experiments



Numerical Experiments




"That's all Folks!"
Isberg©