Understanding the QR algorithm, Part X

David S. Watkins
watkins@math.wsu.edu

Department of Mathematics
Washington State University
1. Understanding the QR algorithm, SIAM Rev., 1982
1. Understanding the QR algorithm, SIAM Rev., 1982
1. Understanding the QR algorithm, SIAM Rev., 1982
3. Some perspectives on the eigenvalue problem, 1993
1. Understanding the QR algorithm, SIAM Rev., 1982
3. Some perspectives on the eigenvalue problem, 1993
4. QR-like algorithms—an overview of convergence theory and practice, AMS proceedings, 1996
1. Understanding the QR algorithm, SIAM Rev., 1982
3. Some perspectives on the eigenvalue problem, 1993
4. QR-like algorithms—an overview of convergence theory and practice, AMS proceedings, 1996
5. QR-like algorithms for eigenvalue problems, JCAM, 2000
1. Understanding the QR algorithm, SIAM Rev., 1982
3. Some perspectives on the eigenvalue problem, 1993
4. QR-like algorithms—an overview of convergence theory and practice, AMS proceedings, 1996
5. QR-like algorithms for eigenvalue problems, JCAM, 2000
1. Understanding the QR algorithm, SIAM Rev., 1982
3. Some perspectives on the eigenvalue problem, 1993
4. QR-like algorithms—an overview of convergence theory and practice, AMS proceedings, 1996
5. QR-like algorithms for eigenvalue problems, JCAM, 2000
1. Understanding the QR algorithm, SIAM Rev., 1982
3. Some perspectives on the eigenvalue problem, 1993
4. QR-like algorithms—an overview of convergence theory and practice, AMS proceedings, 1996
5. QR-like algorithms for eigenvalue problems, JCAM, 2000
Some names associated with the QR algorithm
Some names associated with the QR algorithm (short list)
Some names associated with the QR algorithm (short list)

- Rutishauser
Some names associated with the QR algorithm (short list)

- Rutishauser
- Kublanovskaya
Some names associated with the QR algorithm (short list)

- Rutishauser
- Kublanovskaya
- Francis
Some names associated with the QR algorithm (short list)

- Rutishauser
- Kublanovskaya
- Francis

Implicitly Shifted QR algorithm
Some names associated with the QR algorithm (short list)

- Rutishauser
- Kublanovskaya
- Francis

Implicitly Shifted QR algorithm

- How should we understand it?
Some names associated with the QR algorithm (short list)

- Rutishauser
- Kublanovskaya
- Francis

Implicitly Shifted QR algorithm

- How should we understand it? ... view it?
Some names associated with the QR algorithm (short list)

- Rutishauser
- Kublanovskaya
- Francis

Implicitly Shifted QR algorithm

- How should we understand it? ... view it? ... teach it to our students?
The Standard Approach …
The Standard Approach ...
... dating from the work of Francis
The Standard Approach . . .
...dating from the work of Francis

- Start with the basic algorithm . . .
The Standard Approach …
… dating from the work of Francis

- Start with the basic algorithm …

- $A = QR$
The Standard Approach …
…dating from the work of Francis

- Start with the basic algorithm …

- $A = QR \quad RQ = \hat{A}$
The Standard Approach . . .
...dating from the work of Francis

- Start with the basic algorithm . . .
- $A = QR \quad RQ = \hat{A}$ \quad repeat!
The Standard Approach . . .
...dating from the work of Francis

- Start with the basic algorithm . . .
- \( A = QR \quad RQ = \hat{A} \) repeat!
- This is simple,
The Standard Approach ... ... dating from the work of Francis

- Start with the basic algorithm ...
- $A = QR \quad RQ = \hat{A}$  repeat!
- This is simple, appealing,
The Standard Approach . . .
...dating from the work of Francis

- Start with the basic algorithm . . .
- $A = QR \quad RQ = \hat{A}$ repeat!
- This is simple, appealing, does not require much preparation,
The Standard Approach ... 
...dating from the work of Francis

- Start with the basic algorithm ...

- \( A = QR \quad RQ = \hat{A} \) repeat!

- This is simple, appealing, does not require much preparation, but ...
The Standard Approach ... 
... dating from the work of Francis

- Start with the basic algorithm ...

- \( A = QR \quad RQ = \hat{A} \) repeat!

- This is simple, appealing, does not require much preparation, but ...

- ... it is far removed from versions of the \( QR \) algorithm that are actually used.
Refinements
Refinements

- shifts of origin
Refinements

- shifts of origin
- reduction to Hessenberg form
Refinements

- shifts of origin
- reduction to Hessenberg form
- implicit shift technique  (Francis)
Refinements

- shifts of origin
- reduction to Hessenberg form
- implicit shift technique (Francis)
- double shift $QR$
Refinements

- shifts of origin
- reduction to Hessenberg form
- implicit shift technique (Francis)
- double shift $QR$
- multiple shift $QR$
Refinements

- shifts of origin
- reduction to Hessenberg form
- implicit shift technique (Francis)
- double shift $QR$
- multiple shift $QR$
  - implicit-$Q$ theorem
Refinements

- shifts of origin
- reduction to Hessenberg form
- implicit shift technique (Francis)
- double shift $QR$
- multiple shift $QR$
  
  implicit-$Q$ theorem vs. Krylov subspaces
Refinements

- shifts of origin
- reduction to Hessenberg form
- implicit shift technique (Francis)
- double shift $QR$
- multiple shift $QR$

implicit-$Q$ theorem vs. Krylov subspaces

- Introducing Krylov subspaces improves understanding,
Refinements

- shifts of origin
- reduction to Hessenberg form
- implicit shift technique (Francis)
- double shift $QR$
- multiple shift $QR$
  - implicit-$Q$ theorem vs. Krylov subspaces
- Introducing Krylov subspaces improves understanding, allows more general results,
Refinements

- shifts of origin
- reduction to Hessenberg form
- implicit shift technique  (Francis)
- double shift $QR$
- multiple shift $QR$

  implicit-$Q$ theorem vs. Krylov subspaces

- Introducing Krylov subspaces improves understanding, allows more general results, and prepares students for Krylov subspace methods.
The Implicitly Shifted QR Iteration
The Implicitly Shifted QR Iteration

- matrix is in upper Hessenberg form
The Implicitly Shifted QR Iteration

- matrix is in upper Hessenberg form
- pick some shifts $\rho_1, \ldots, \rho_m$
The Implicitly Shifted QR Iteration

- matrix is in upper Hessenberg form
- pick some shifts \( \rho_1, \ldots, \rho_m \) \( (m = 1, 2, 4, 6) \)
The Implicitly Shifted QR Iteration

- matrix is in upper Hessenberg form
- pick some shifts $\rho_1, \ldots, \rho_m$ ($m = 1, 2, 4, 6$)
- $p(A) = (A - \rho_1 I) \cdots (A - \rho_m I)$
The Implicitly Shifted QR Iteration

- matrix is in upper Hessenberg form
- pick some shifts $\rho_1, \ldots, \rho_m$ ($m = 1, 2, 4, 6$)
- $p(A) = (A - \rho_1 I) \cdots (A - \rho_m I)$ expensive!
The Implicitly Shifted QR Iteration

- matrix is in upper Hessenberg form
- pick some shifts $\rho_1, \ldots, \rho_m$ \hspace{1cm} ($m = 1, 2, 4, 6$)
- $p(A) = (A - \rho_1 I) \cdots (A - \rho_m I)$ \hspace{1cm} expensive!
- compute $p(A)e_1$
The Implicitly Shifted QR Iteration

- matrix is in upper Hessenberg form
- pick some shifts $\rho_1, \ldots, \rho_m$ \hspace{1cm} ($m = 1, 2, 4, 6$)
- $p(A) = (A - \rho_1 I) \cdots (A - \rho_m I)$ \hspace{1cm} expensive!
- compute $p(A)e_1$ \hspace{1cm} cheap!
The Implicitly Shifted QR Iteration

- matrix is in upper Hessenberg form
- pick some shifts $\rho_1, \ldots, \rho_m$ ($m = 1, 2, 4, 6$)
- $p(A) = (A - \rho_1 I) \cdots (A - \rho_m I)$ expensive!
- compute $p(A)e_1$ cheap!
- Build unitary $Q_0$ with $q_1 = \alpha p(A)e_1$. 
The Implicitly Shifted QR Iteration

- The matrix is in upper Hessenberg form.
- Pick some shifts $\rho_1, \ldots, \rho_m$ ($m = 1, 2, 4, 6$).
- $p(A) = (A - \rho_1 I) \cdots (A - \rho_m I)$ expensive!
- Compute $p(A)e_1$ cheap!
- Build unitary $Q_0$ with $q_1 = \alpha p(A)e_1$.
- Perform similarity transform $A \rightarrow Q_0^*AQ_0$. 
The Implicitly Shifted QR Iteration

- matrix is in upper Hessenberg form
- pick some shifts $\rho_1, \ldots, \rho_m$ ($m = 1, 2, 4, 6$)
- $p(A) = (A - \rho_1 I) \cdots (A - \rho_m I)$ expensive!
- compute $p(A)e_1$ cheap!
- Build unitary $Q_0$ with $q_1 = \alpha p(A)e_1$.
- Perform similarity transform $A \rightarrow Q_0^*AQ_0$.
- Hessenberg form is disturbed.
An Upper Hessenberg Matrix
After the Transformation \((Q_0^*AQ_0)\)
After the Transformation \((Q_0^*AQ_0)\)

Now return the matrix to Hessenberg form.
Chasing the Bulge
Chasing the Bulge
Done
Done

The implicit $QR$ step is complete!
Summary of Implicit $QR$ Iteration
Summary of Implicit $QR$ Iteration

- Pick some shifts.
Summary of Implicit $QR$ Iteration

- Pick some shifts.
- Compute $p(A)e_1$. ($p$ determined by shifts)
Summary of Implicit $QR$ Iteration

- Pick some shifts.
- Compute $p(A)e_1$.  ($p$ determined by shifts)
- Build $Q_0$ with first column $q_1 = \alpha p(A)e_1$.  

Glasgow 2009 – p. 12
Summary of Implicit $QR$ Iteration

- Pick some shifts.
- Compute $p(A)e_1$. ($p$ determined by shifts)
- Build $Q_0$ with first column $q_1 = \alpha p(A)e_1$.
- Make a bulge. ($A \rightarrow Q_0^*AQ_0$)
Summary of Implicit $QR$ Iteration

- Pick some shifts.
- Compute $p(A)e_1$. ($p$ determined by shifts)
- Build $Q_0$ with first column $q_1 = \alpha p(A)e_1$.
- Make a bulge. ($A \rightarrow Q_0^*AQ_0$)
- Chase the bulge. (return to Hessenberg form)
Summary of Implicit $QR$ Iteration

- Pick some shifts.
- Compute $p(A)e_1$.  ($p$ determined by shifts)
- Build $Q_0$ with first column $q_1 = \alpha p(A)e_1$.
- Make a bulge.  ($A \rightarrow Q_0^*AQ_0$)
- Chase the bulge.  (return to Hessenberg form)
- $\hat{A} = Q^*AQ$
Question
Question

- This differs a lot from the basic $QR$ step.
Question

- This differs a lot from the basic $QR$ step.

\[ A = QR \quad RQ = \hat{A} \]
Question

- This differs a lot from the basic $QR$ step.

\[ A = QR \quad RQ = \hat{A} \]

- Can we carve a reasonable pedagogical path that leads directly to the implicitly-shifted $QR$ algorithm,
Question

- This differs a lot from the basic $QR$ step.

\[ A = QR \quad RQ = \hat{A} \]

- Can we carve a reasonable pedagogical path that leads directly to the implicitly-shifted $QR$ algorithm, bypassing the basic $QR$ algorithm entirely?
Question

- This differs a lot from the basic $QR$ step.

$$A = QR \quad RQ = \hat{A}$$

- Can we carve a reasonable pedagogical path that leads directly to the implicitly-shifted $QR$ algorithm, bypassing the basic $QR$ algorithm entirely?

- That’s what we are going to do today.
Ingredients
Ingredients

- subspace iteration (power method)
Ingredients

- subspace iteration (power method)
- Krylov subspaces
Ingredients

- subspace iteration (power method)
- Krylov subspaces and subspace iteration
Ingredients

- subspace iteration  (power method)
- Krylov subspaces  and subspace iteration
- (unitary) similarity transformation  
  (change of coordinate system)
Ingredients

- subspace iteration  (power method)
- Krylov subspaces  and subspace iteration
- (unitary) similarity transformation  
  (change of coordinate system)
- reduction to Hessenberg form
Ingredients

- subspace iteration  (power method)
- Krylov subspaces  and subspace iteration
- (unitary) similarity transformation  
  (change of coordinate system)
- reduction to Hessenberg form
- Hessenberg form and Krylov subspaces  
  (instead of implicit-$Q$ theorem)
Ingredients

- subspace iteration (power method)
- Krylov subspaces and subspace iteration
- (unitary) similarity transformation (change of coordinate system)
- reduction to Hessenberg form
- Hessenberg form and Krylov subspaces (instead of implicit-$Q$ theorem)

No Magic Shortcut!
Power Method, Subspace Iteration
Power Method, Subspace Iteration

\[ \nu, A\nu, A^2\nu, A^3\nu, \ldots \]
Power Method, Subspace Iteration

- $v, Av, A^2v, A^3v, \ldots$
- convergence rate $|\lambda_2/\lambda_1|$
Power Method, Subspace Iteration

- $v, Av, A^2v, A^3v, \ldots$
- convergence rate $|\lambda_2/\lambda_1|$
- $S, AS, A^2S, A^3S, \ldots$
Power Method, Subspace Iteration

- \( v, Av, A^2v, A^3v, \ldots \)
- convergence rate \( |\lambda_2/\lambda_1| \)
- \( S, AS, A^2S, A^3S, \ldots \)
- subspaces of dimension \( j \)
Power Method, Subspace Iteration

- $v, Av, A^2v, A^3v, \ldots$
- convergence rate $|\lambda_2/\lambda_1|$
- $S, AS, A^2S, A^3S, \ldots$
- subspaces of dimension $j$ ($|\lambda_{j+1}/\lambda_j|$)
Power Method, Subspace Iteration

- $v, Av, A^2v, A^3v, \ldots$
- convergence rate $|\lambda_2/\lambda_1|$
- $S, AS, A^2S, A^3S, \ldots$
- subspaces of dimension $j$ ($|\lambda_{j+1}/\lambda_j|$)
- Substitute $p(A)$ for $A$
Power Method, Subspace Iteration

- $v, Av, A^2v, A^3v, \ldots$
- convergence rate $|\lambda_2/\lambda_1|$
- $S, AS, A^2S, A^3S, \ldots$
- subspaces of dimension $j$ ($|\lambda_{j+1}/\lambda_j|$)
- Substitute $p(A)$ for $A$ (shifts, multiple steps)
Power Method, Subspace Iteration

- $v, Av, A^2v, A^3v, \ldots$
- convergence rate $|\lambda_2/\lambda_1|$
- $S, AS, A^2S, A^3S, \ldots$
- subspaces of dimension $j$ ($|\lambda_{j+1}/\lambda_j|$)
- Substitute $p(A)$ for $A$ (shifts, multiple steps)
- $S, p(A)S, p(A)^2S, p(A)^3S, \ldots$
Power Method, Subspace Iteration

- $v, Av, A^2v, A^3v, \ldots$
- convergence rate $|\lambda_2/\lambda_1|$
- $S, AS, A^2S, A^3S, \ldots$
- subspaces of dimension $j$ ($|\lambda_{j+1}/\lambda_j|$)
- Substitute $p(A)$ for $A$ (shifts, multiple steps)
- $S, p(A)S, p(A)^2S, p(A)^3S, \ldots$
- convergence rate $|p(\lambda_{j+1})/p(\lambda_j)|$
Krylov Subspaces ...
Krylov Subspaces …
…and Subspace Iteration
Krylov Subspaces . . .

...and Subspace Iteration

Def: $\mathcal{K}_j(A, q) = \text{span}\{q, Aq, A^2q, \ldots, A^{j-1}q\}$
Krylov Subspaces ...

... and Subspace Iteration

Def: $\mathcal{K}_j(A, q) = \text{span}\{q, Aq, A^2q, \ldots, A^{j-1}q\}$

$j = 1, 2, 3, \ldots$ (nested subspaces)
Krylov Subspaces . . .
...and Subspace Iteration

- Def: $\mathcal{K}_j(A, q) = \text{span}\{q, Aq, A^2q, \ldots, A^{j-1}q\}$
  
  $j = 1, 2, 3, \ldots$ (nested subspaces)

- $\mathcal{K}_j(A, q)$ are “determined by $q$”.

Glasgow 2009 – p. 16
Krylov Subspaces ...

... and Subspace Iteration

- Def: $\mathcal{K}_j(A, q) = \text{span}\{q, Aq, A^2q, \ldots, A^{j-1}q\}$
  
  $j = 1, 2, 3, \ldots$ (nested subspaces)

- $\mathcal{K}_j(A, q)$ are “determined by $q$”.

- $p(A)\mathcal{K}_j(A, q) = \mathcal{K}_j(A, p(A)q)$
Krylov Subspaces . . .

...and Subspace Iteration

- Def: $\mathcal{K}_j(A, q) = \text{span}\{q, Aq, A^2q, \ldots, A^{j-1}q\}$
  \[ j = 1, 2, 3, \ldots \text{ (nested subspaces)} \]

- $\mathcal{K}_j(A, q)$ are “determined by $q$”.

- $p(A)\mathcal{K}_j(A, q) = \mathcal{K}_j(A, p(A)q)$

- ...because $p(A)A = Ap(A)$
Krylov Subspaces ...
... and Subspace Iteration

- Def: \( \mathcal{K}_j(A, q) = \text{span}\{ q, Aq, A^2q, \ldots, A^{j-1}q \} \)
  \( j = 1, 2, 3, \ldots \) (nested subspaces)

- \( \mathcal{K}_j(A, q) \) are “determined by \( q \)”.

- \( p(A)\mathcal{K}_j(A, q) = \mathcal{K}_j(A, p(A)q) \)

- ... because \( p(A)A = Ap(A) \)

- Conclusion: Power method induces nested subspace iterations on Krylov subspaces.
power method: \[ p(A)^k q \]
- power method: $p(A)^k q$
- nested subspace iterations:
  
  $$p(A)^k \mathcal{K}_j(A, q) = \mathcal{K}_j(A, p(A)^k q) \quad j = 1, 2, 3, \ldots$$
- power method: \( p(A)^k q \)
- nested subspace iterations:
  \[
p(A)^k \mathcal{K}_j(A, q) = \mathcal{K}_j(A, p(A)^k q) \quad j = 1, 2, 3, \ldots
\]
- convergence rates:
  \[
  |p(\lambda_{j+1})/p(\lambda_j)|, \quad j = 1, 2, 3, \ldots
  \]
(Unitary) Similarity Transforms
(Unitary) Similarity Transforms

- $A \rightarrow Q^* AQ$ preserves eigenvalues
(Unitary) Similarity Transforms

- $A \rightarrow Q^*AQ$ preserves eigenvalues
- transforms eigenvectors in a simple way
  ($w \rightarrow Q^*w$)
(Unitary) Similarity Transforms

- $A \rightarrow Q^*AQ$ preserves eigenvalues
- transforms eigenvectors in a simple way $(w \rightarrow Q^*w)$
- is a change of coordinate system $(v \rightarrow Q^*v)$
(Unitary) Similarity Transforms

- $A \rightarrow Q^*AQ$ preserves eigenvalues
- transforms eigenvectors in a simple way ($w \rightarrow Q^*w$)
- is a change of coordinate system ($v \rightarrow Q^*v$)
- triangular form (eigenvalues)
(Unitary) Similarity Transforms

- $A \rightarrow Q^*AQ$ preserves eigenvalues
- transforms eigenvectors in a simple way ($w \rightarrow Q^*w$)
- is a change of coordinate system ($v \rightarrow Q^*v$)
- triangular form (eigenvalues)
- relationship of invariant subspaces to triangular form
Subspace Iteration
with change of coordinate system
Subspace Iteration with change of coordinate system

- take $S = \text{span}\{e_1, \ldots, e_j\}$
Subspace Iteration
with change of coordinate system

- take $S = \text{span}\{e_1, \ldots, e_j\}$

$$p(A)S = \text{span}\{p(A)e_1, \ldots, p(A)e_j\}$$

$$= \text{span}\{q_1, \ldots, q_j\} \text{ (orthonormal)}$$
Subspace Iteration
with change of coordinate system

- take $S = \text{span}\{e_1, \ldots, e_j\}$

  $p(A)S = \text{span}\{p(A)e_1, \ldots, p(A)e_j\}$

  $= \text{span}\{q_1, \ldots, q_j\}$ (orthonormal)

- build unitary $Q = [q_1 \cdots q_j \cdots]$
Subspace Iteration with change of coordinate system

- take $S = \text{span}\{e_1, \ldots, e_j\}$

\[ p(A)S = \text{span}\{p(A)e_1, \ldots, p(A)e_j\} = \text{span}\{q_1, \ldots, q_j\} \text{ (orthonormal)} \]

- build unitary $Q = [q_1 \cdots q_j \cdots]$

- change coordinate system: $\hat{A} = Q^*AQ$
Subspace Iteration with change of coordinate system

- take $S = \text{span}\{e_1, \ldots, e_j\}$

\[
p(A)S = \text{span}\{p(A)e_1, \ldots, p(A)e_j\} = \text{span}\{q_1, \ldots, q_j\} \text{ (orthonormal)}
\]

- build unitary $Q = [q_1 \cdots q_j \cdots]$

- change coordinate system: $\hat{A} = Q^*AQ$

- $q_k \rightarrow Q^*q_k = e_k$
Subspace Iteration with change of coordinate system

- take $S = \text{span}\{e_1, \ldots, e_j\}$
  
  $p(A)S = \text{span}\{p(A)e_1, \ldots, p(A)e_j\}$
  
  $= \text{span}\{q_1, \ldots, q_j\}$ (orthonormal)

- build unitary $Q = [q_1 \cdots q_j \cdots]$

- change coordinate system: $\hat{A} = Q^*AQ$

- $q_k \rightarrow Q^*q_k = e_k$

- $\text{span}\{q_1, \ldots, q_j\} \rightarrow \text{span}\{e_1, \ldots, e_j\}$
Subspace Iteration with change of coordinate system

- take $S = \text{span}\{e_1, \ldots, e_j\}$
  \[
p(A)S = \text{span}\{p(A)e_1, \ldots, p(A)e_j\}
  = \text{span}\{q_1, \ldots, q_j\} \quad \text{(orthonormal)}
\]

- build unitary $Q = [q_1 \cdots q_j \cdots]$
- change coordinate system: $\hat{A} = Q^*AQ$
- $q_k \rightarrow Q^*q_k = e_k$
- $\text{span}\{q_1, \ldots, q_j\} \rightarrow \text{span}\{e_1, \ldots, e_j\}$
- ready for next iteration
This version of subspace iteration ...
This version of subspace iteration . . .
  . . . holds the subspace fixed
This version of subspace iteration . . .

- . . . holds the subspace fixed
- while the matrix changes.
This version of subspace iteration . . .

- ... holds the subspace fixed
- while the matrix changes.
- ... moving toward a matrix under which

\[ \text{span}\{e_1, \ldots, e_j\} \]

is invariant.
This version of subspace iteration . . .

- ... holds the subspace fixed
- while the matrix changes.
- ... moving toward a matrix under which

\[ \text{span}\{e_1, \ldots, e_j\} \]

is invariant.

- \( A \rightarrow \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \)
  \((A_{11} \text{ is } j \times j.\)
Reduction to Hessenberg form
Reduction to Hessenberg form

\[ Q \rightarrow Q^* AQ = H \quad (a \text{ similarity transformation}) \]
Reduction to Hessenberg form

- $Q \rightarrow Q^*AQ = H$ (a similarity transformation)
- can always be done (direct method, $O(n^3)$ flops)
Reduction to Hessenberg form

- $Q \rightarrow Q^*AQ = H$ (a similarity transformation)
- can always be done (direct method, $O(n^3)$ flops)
- brings us closer to triangular form
Reduction to Hessenberg form

- $Q \rightarrow Q^*AQ = H$ (a similarity transformation)
- can always be done (direct method, $O(n^3)$ flops)
- brings us closer to triangular form
- makes computations cheaper
Reduction to Hessenberg form

- $Q \rightarrow Q^*AQ = H$ (a similarity transformation)
- can always be done (direct method, $O(n^3)$ flops)
- brings us closer to triangular form
- makes computations cheaper
- First column $q_1$ can be chosen “arbitrarily”.
Reduction to Hessenberg form

- $Q \rightarrow Q^*AQ = H$ (a similarity transformation)
- can always be done (direct method, $O(n^3)$ flops)
- brings us closer to triangular form
- makes computations cheaper
- First column $q_1$ can be chosen “arbitrarily”.
- Example: $q_1 = \alpha p(A)e_1$
Krylov Subspaces ...
Krylov Subspaces ...

...and Hessenberg matrices ...
Krylov Subspaces . . .

...and Hessenberg matrices . . .

... go hand in hand.
Krylov Subspaces . . .
...and Hessenberg matrices . . .

- ... go hand in hand.
- $A$ properly upper Hessenberg $\implies$

$$\mathcal{K}_j(A, e_1) = \text{span}\{e_1, \ldots, e_j\}.$$
Krylov Subspaces ...  
... and Hessenberg matrices ...

- ... go hand in hand.

- A properly upper Hessenberg \( \implies \)

\[ K_j(A, e_1) = \text{span}\{e_1, \ldots, e_j\}. \]

- More generally ...
Krylov-Hessenberg Relationship
Krylov-Hessenberg Relationship

If $H = Q^* A Q$, and $H$ is properly upper Hessenberg, then for $j = 1, 2, 3, \ldots$,

$$\text{span}\{q_1, \ldots, q_j\} = \mathcal{K}_j(A, q_1).$$
Krylov-Hessenberg Relationship

If $H = Q^*AQ$, and $H$ is properly upper Hessenberg, then for $j = 1, 2, 3, \ldots$,

$$\text{span}\{q_1, \ldots, q_j\} = \mathcal{K}_j(A, q_1).$$

Proof (sketch):
Krylov-Hessenberg Relationship

If $H = Q^*AQ$, and $H$ is properly upper Hessenberg, then for $j = 1, 2, 3, \ldots$,

$$\text{span}\{q_1, \ldots, q_j\} = \mathcal{K}_j(A, q_1).$$

Proof (sketch): Induction on $j$. 

Glasgow 2009 – p. 23
Krylov-Hessenberg Relationship

If $H = Q^*AQ$, and $H$ is properly upper Hessenberg, then for $j = 1, 2, 3, \ldots$,

$$\text{span}\{q_1, \ldots, q_j\} = \mathcal{K}_j(A, q_1).$$

Proof (sketch): Induction on $j$. $AQ = QH$
Krylov-Hessenberg Relationship

If \( H = Q^* A Q \), and \( H \) is properly upper Hessenberg, then for \( j = 1, 2, 3, \ldots \),

\[
\text{span}\{q_1, \ldots, q_j\} = \mathcal{K}_j(A, q_1).
\]

Proof (sketch): Induction on \( j \). \( AQ = QH \)

\[
Aq_j = \sum_{i=1}^{n} q_i h_{ij} = \sum_{i=1}^{j} q_i h_{ij} + q_{j+1} h_{j+1,j}
\]
\[ Aq_j = \sum_{i=1}^{j} q_i h_{ij} + q_{j+1} h_{j+1,j} \]
\[ Aq_j = \sum_{i=1}^{j} q_i h_{ij} + q_{j+1} h_{j+1,j} \]

\[ q_{j+1} h_{j+1,j} = Aq_j - \sum_{i=1}^{j} q_i h_{ij} \]
\[ Aq_j = \sum_{i=1}^{j} q_i h_{ij} + q_{j+1} h_{j+1,j} \]

\[ q_{j+1} h_{j+1,j} = Aq_j - \sum_{i=1}^{j} q_i h_{ij} \]

- Proof by induction follows immediately.
\[ Aq_j = \sum_{i=1}^{j} q_i h_{ij} + q_{j+1} h_{j+1,j} \]

\[ q_{j+1} h_{j+1,j} = Aq_j - \sum_{i=1}^{j} q_i h_{ij} \]

- Proof by induction follows immediately.
- This also gives the student a preview of the Arnoldi process,
Aq_j = \sum_{i=1}^{j} q_i h_{ij} + q_{j+1} h_{j+1,j}

q_{j+1} h_{j+1,j} = Aq_j - \sum_{i=1}^{j} q_i h_{ij}

- Proof by induction follows immediately.
- This also gives the student a preview of the Arnoldi process, the most important Krylov subspace method.
and now,
and now, the Implicit QR Iteration
and now, the Implicit QR Iteration

Work with Hessenberg form to get …
and now, the Implicit QR Iteration

- Work with Hessenberg form to get . . .
  - . . . efficiency.
and now, the Implicit QR Iteration

- Work with Hessenberg form to get ...
  - ...efficiency.
  - ...automatic nested subspace iterations.
and now, the Implicit QR Iteration

- Work with Hessenberg form to get ...
  - ... efficiency.
  - ... automatic nested subspace iterations.
- Get some shifts $\rho_1, \ldots, \rho_m$ to define $p$. 
and now, the Implicit QR Iteration

- Work with Hessenberg form to get . . .
  - . . . efficiency.
  - . . . automatic nested subspace iterations.
- Get some shifts $\rho_1, \ldots, \rho_m$ to define $p$.
- Compute $p(A)e_1$. (power method)
and now, the Implicit QR Iteration

- Work with Hessenberg form to get ...
  - ...efficiency.
  - ...automatic nested subspace iterations.
- Get some shifts \( \rho_1, \ldots, \rho_m \) to define \( p \).
- Compute \( p(A)e_1 \). (power method)
- Transform \( A \) to upper Hessenberg form:

\[
\hat{A} = Q^*AQ
\]

by a matrix \( Q \) that has \( q_1 = \alpha p(A)e_1 \).
\[ \hat{A} = Q^*AQ \quad \text{where} \quad q_1 = \alpha p(A)e_1. \]
\[ \hat{A} = Q^* A Q \quad \text{where} \quad q_1 = \alpha p(A) e_1. \]

\[ q_1 \rightarrow Q^* q_1 = e_1 \]
\[ \hat{A} = Q^*AQ \quad \text{where} \quad q_1 = \alpha p(A)e_1. \]

- \( q_1 \rightarrow Q^*q_1 = e_1 \)

- power method with a change of coordinate system. Moreover \ldots
\[ \hat{A} = Q^*AQ \quad \text{where} \quad q_1 = \alpha p(A)e_1. \]

- \( q_1 \to Q^*q_1 = e_1 \)
- power method with a change of coordinate system. Moreover . . .
- \( p(A)\mathcal{K}_j(A, e_1) = \mathcal{K}_j(A, p(A)e_1) \)
\[ \hat{A} = Q^*AQ \quad \text{where} \quad q_1 = \alpha p(A)e_1. \]

\begin{itemize}
  \item $q_1 \rightarrow Q^*q_1 = e_1$
  \item power method with a change of coordinate system. Moreover \ldots
  \item $p(A)K_j(A, e_1) = K_j(A, p(A)e_1)$
  \item i.e. $p(A)\text{span}\{e_1, \ldots, e_j\} = \text{span}\{q_1, \ldots, q_j\}$
\end{itemize}
\[ \hat{A} = Q^*AQ \quad \text{where} \quad q_1 = \alpha p(A)e_1. \]

- \( q_1 \rightarrow Q^*q_1 = e_1 \)

- power method with a change of coordinate system. Moreover ...

- \( p(A)\mathcal{K}_j(A, e_1) = \mathcal{K}_j(A, p(A)e_1) \)

- i.e. \( p(A)\text{span}\{e_1, \ldots, e_j\} = \text{span}\{q_1, \ldots, q_j\} \)

- subspace iteration with a change of coordinate system
\[ \hat{A} = Q^* A Q \quad \text{where} \quad q_1 = \alpha p(A) e_1. \]

- \( q_1 \to Q^* q_1 = e_1 \)
- power method with a change of coordinate system. Moreover ...
- \( p(A) \mathcal{K}_j(A, e_1) = \mathcal{K}_j(A, p(A) e_1) \)
- i.e. \( p(A) \text{span}\{e_1, \ldots, e_j\} = \text{span}\{q_1, \ldots, q_j\} \)
- subspace iteration with a change of coordinate system
- \( j = 1, 2, 3, \ldots, n - 1 \)
\[ \hat{A} = Q^* A Q \quad \text{where} \quad q_1 = \alpha p(A)e_1. \]

- \( q_1 \to Q^* q_1 = e_1 \)
- power method with a change of coordinate system. Moreover ...

- \( p(A)K_j(A, e_1) = K_j(A, p(A)e_1) \)
- i.e. \( p(A)\text{span}\{e_1, \ldots, e_j\} = \text{span}\{q_1, \ldots, q_j\} \)
- subspace iteration with a change of coordinate system

- \( j = 1, 2, 3, \ldots, n - 1 \)
- \( \left| p(\lambda_{j+1}) / p(\lambda_j) \right| \quad j = 1, 2, 3, \ldots, n - 1 \)
Details

- choice of shifts
Details

- choice of shifts
- We change the shifts at each step.
Details

- choice of shifts
- We change the shifts at each step.
- $\Rightarrow$ quadratic or cubic convergence
Details

- choice of shifts
- We change the shifts at each step.
- $\Rightarrow$ quadratic or cubic convergence

Other Questions
Details

- choice of shifts
- We change the shifts at each step.
- \( \Rightarrow \) quadratic or cubic convergence

Other Questions

- \ldots how to get BLAS 3 speed?
- \ldots how to parallelize?
In Conclusion
In Conclusion

A careful study of
In Conclusion
A careful study of the power method and its extensions,
In Conclusion
A careful study of the power method and its extensions, similarity transformations,
In Conclusion
A careful study of the power method and its extensions, similarity transformations, Hessenberg form,
In Conclusion
A careful study of the power method and its extensions, similarity transformations, Hessenberg form, and Krylov subspaces
In Conclusion
A careful study of the power method and its extensions, similarity transformations, Hessenberg form, and Krylov subspaces leads directly to the implicitly shifted $QR$ algorithm.
In Conclusion
A careful study of the power method and its extensions, similarity transformations, Hessenberg form, and Krylov subspaces leads directly to the implicitly shifted $QR$ algorithm.

- The basic, explicit $QR$ algorithm is skipped.
In Conclusion
A careful study of the power method and its extensions, similarity transformations, Hessenberg form, and Krylov subspaces leads directly to the implicitly shifted $QR$ algorithm.

- The basic, explicit $QR$ algorithm is skipped.
- The implicit-$Q$ theorem is avoided.
In Conclusion

A careful study of the power method and its extensions, similarity transformations, Hessenberg form, and Krylov subspaces leads directly to the implicitly shifted $QR$ algorithm.

- The basic, explicit $QR$ algorithm is skipped.
- The implicit-$Q$ theorem is avoided.
- Krylov subspaces are emphasized.
In Conclusion
A careful study of the power method and its extensions, similarity transformations, Hessenberg form, and Krylov subspaces leads directly to the implicitly shifted $QR$ algorithm.

- The basic, explicit $QR$ algorithm is skipped.
- The implicit-$Q$ theorem is avoided.
- Krylov subspaces are emphasized.
- Krylov subspace methods are foreshadowed.
One Last Question
One Last Question

- In the implicitly shifted $QR$ algorithm
One Last Question

- In the implicitly shifted $QR$ algorithm, the $QR$ decomposition is nowhere to be seen.
One Last Question

- In the implicitly shifted $QR$ algorithm the $QR$ decomposition is nowhere to be seen.
- Should the implicitly-shifted $QR$ algorithm be given a different name?
One Last Question

- In the implicitly shifted $QR$ algorithm the $QR$ decomposition is nowhere to be seen.
- Should the implicitly-shifted $QR$ algorithm be given a different name? Some possibilities: ...
One Last Question

- In the implicitly shifted $QR$ algorithm, the $QR$ decomposition is nowhere to be seen.
- Should the implicitly-shifted $QR$ algorithm be given a different name? Some possibilities: ...
- ...unitary bulge-chasing algorithm
One Last Question

- In the implicitly shifted $QR$ algorithm the $QR$ decomposition is nowhere to be seen.
- Should the implicitly-shifted $QR$ algorithm be given a different name? Some possibilities: ...
- ... unitary bulge-chasing algorithm
- ... Hessenberg-Krylov nonstationary progressive nested subspace iteration
One Last Question

- In the implicitly shifted $QR$ algorithm the $QR$ decomposition is nowhere to be seen.
- Should the implicitly-shifted $QR$ algorithm be given a different name? Some possibilities: …
  - …unitary bulge-chasing algorithm
  - …Hessenberg-Krylov nonstationary progressive nested subspace iteration
  - …Francis’s algorithm
One Last Question

- In the implicitly shifted $QR$ algorithm the $QR$ decomposition is nowhere to be seen.
- Should the implicitly-shifted $QR$ algorithm be given a different name? Some possibilities: ...
  - ...unitary bulge-chasing algorithm
  - ...Hessenberg-Krylov nonstationary progressive nested subspace iteration
  - ...Francis’s algorithm
- Thank you for your attention.