Representation of the unitary shift operator

in factored form w.r.t. an orthogonal rational basis

Adhemar Bultheel

Dept. Computer Science, KU Leuven

Leipzig, November 2016

This is joint work with
Ruyman Cruz-Barroso, Andreas Lasarow

`http://nalag.cs.kuleuven.be/papers/ade/leipzig16`

Table of contents

- Krylov and ORF
- Orthogonal rational fuctions (ORF)
- Representation of the shift operator

Motivation

Classical Krylov

- $A \in \mathbb{C}^{n \times n}$ normal matrix: $A = U \Lambda U^*$, $U^* U = I_n$
- Krylov: $K_m = [v, Av, A^2 v, \ldots, A^m v]$
- Orthonormalize: $V_m = [v_0, v_1, \ldots, v_m]$, $V_m^* V_m = I_m$
- $v_k = \phi_k(A) v$, $\phi_k \in \mathcal{P}_k$ (polynomial)
- $Av_k \in \operatorname{span} K_{k+1}$
  i.e. $AV_m = V_m H_m + b_{m+1}[0, 0, \ldots, v_{m+1}]$, $H_m =$ upper Hessenberg
- $V_m^* A V_m = H_m$, $\sigma(A) \approx \sigma(H_m)$ (Ritz values)
- H_m is projection of A on a lower dimensional Krylov subspace

Motivation

Classical Krylov

- $A \in \mathbb{C}^{n \times n}$ normal matrix: $A = U \Lambda U^*$, $U^* U = I_n$
- Krylov: $K_m = [v, Av, A^2 v, \ldots, A^m v]$
- Orthonormalize: $V_m = [v_0, v_1, \ldots, v_m]$, $V_m^* V_m = I_m$
- $v_k = \phi_k(A) v$, $\phi_k \in \mathcal{P}_k$ (polynomial)
- $Av_k \in \operatorname{span} K_{k+1}$
  i.e. $AV_m = V_m H_m + b_{m+1}[0, 0, \ldots, v_{m+1}]$, $H_m =$ upper Hessenberg
- $V_m^* A V_m = H_m$, $\sigma(A) \approx \sigma(H_m)$ (Ritz values)
- H_m is projection of A on a lower dimensional Krylov subspace

Motivation

Classical Krylov

- $A \in \mathbb{C}^{n \times n}$ normal matrix: $A = U \Lambda U^*$, $U^* U = I_n$
- Krylov: $K_m = [v, Av, A^2 v, \ldots, A^m v]$
- Orthonormalize: $V_m = [v_0, v_1, \ldots, v_m]$, $V_m^* V_m = I_m$
- $v_k = \phi_k(A)v$, $\phi_k \in \mathcal{P}_k$ (polynomial)
- $Av_k \in \operatorname{span} K_{k+1}$
  i.e. $AV_m = V_m H_m + b_{m+1}[0, 0, \ldots, v_{m+1}]$, $H_m =$ upper Hessenberg
- $V_m^* A V_m = H_m$, $\sigma(A) \approx \sigma(H_m)$ (Ritz values)
- H_m is projection of A on a lower dimensional Krylov subspace

Motivation

Classical Krylov

- $A \in \mathbb{C}^{n \times n}$ normal matrix: $A = U \Lambda U^*$, $U^* U = I_n$
- Krylov: $K_m = [v, Av, A^2 v, \ldots, A^m v]$
- Orthonormalize: $V_m = [v_0, v_1, \ldots, v_m]$, $V_m^* V_m = I_m$
- $v_k = \phi_k(A)v$, $\phi_k \in \mathcal{P}_k$ (polynomial)
- $Av_k \in \operatorname{span} K_{k+1}$
  i.e. $AV_m = V_m H_m + b_{m+1}[0, 0, \ldots, v_{m+1}]$, $H_m =$ upper Hessenberg
- $V_m^* A V_m = H_m$, $\sigma(A) \approx \sigma(H_m)$ (Ritz values)
- H_m is projection of A on a lower dimensional Krylov subspace

Motivation

Classical Krylov

- $A \in \mathbb{C}^{n \times n}$ normal matrix: $A = U \Lambda U^*$, $U^* U = I_n$
- Krylov: $K_m = [v, Av, A^2 v, \ldots, A^m v]$
- Orthonormalize: $V_m = [v_0, v_1, \ldots, v_m]$, $V_m^* V_m = I_m$
- $v_k = \phi_k(A) v$, $\phi_k \in \mathcal{P}_k$ (polynomial)
- $Av_k \in \operatorname{span} K_{k+1}$
  i.e. $AV_m = V_m H_m + b_{m+1}[0, 0, \ldots, v_{m+1}]$, $H_m =$ upper Hessenberg
- $V_m^* A V_m = H_m$, $\sigma(A) \approx \sigma(H_m)$ (Ritz values)
- H_m is projection of A on a lower dimensional Krylov subspace

Motivation

Classical Krylov

- $A \in \mathbb{C}^{n \times n}$ normal matrix: $A = U \Lambda U^*$, $U^* U = I_n$
- Krylov: $K_m = [v, Av, A^2 v, \ldots, A^m v]$
- Orthonormalize: $V_m = [v_0, v_1, \ldots, v_m]$, $V_m^* V_m = I_m$
- $v_k = \phi_k(A) v$, $\phi_k \in \mathcal{P}_k$ (polynomial)
- $A v_k \in \operatorname{span} K_{k+1}$
  i.e. $AV_m = V_m H_m + b_{m+1}[0, 0, \ldots, v_{m+1}]$, $H_m =$ upper Hessenberg
- $V_m^* A V_m = H_m$, $\sigma(A) \approx \sigma(H_m)$ (Ritz values)
- H_m is projection of A on a lower dimensional Krylov subspace

Classical Krylov

- $A \in \mathbb{C}^{n \times n}$ normal matrix: $A = U \Lambda U^*$, $U^* U = I_n$
- Krylov: $K_m = [v, Av, A^2 v, \ldots, A^m v]$
- Orthonormalize: $V_m = [v_0, v_1, \ldots, v_m]$, $V_m^* V_m = I_m$
- $v_k = \phi_k(A) v$, $\phi_k \in \mathcal{P}_k$ (polynomial)
- $Av_k \in \operatorname{span} K_{k+1}$
  i.e. $AV_m = V_m H_m + b_{m+1}[0, 0, \ldots, v_{m+1}]$, $H_m =$ upper Hessenberg
- $V_m^* A V_m = H_m$, $\sigma(A) \approx \sigma(H_m)$ (Ritz values)
- H_m is projection of A on a lower dimensional Krylov subspace

Polynomials

- $V_m^* V_m = I_m$ i.e., $\delta_{k,\ell} = v_k^* v_\ell$

$$\begin{aligned}
\delta_{k,\ell} &= (\phi_k(A)v)^*(\phi_\ell(A)v) \\
&= v^*(\phi_k(A))^*(\phi_\ell(A))v, \qquad A = U^*\Lambda U \text{ or } \phi(A) = U^*\phi(\Lambda)U \\
&= v^*U^*(\phi_k(\Lambda))^* UU^*(\phi_\ell(\Lambda))Uv, \quad w = Uv \text{ and } UU^* = I_n \\
&= w^*(\phi_k(\Lambda))^*(\phi_\ell(\Lambda))w \\
&= \sum_{j=0}^{n} |w_j|^2 \overline{\phi_k(\lambda_j)}\phi_\ell(\lambda_j)
\end{aligned}$$

The $\{\phi_k\}$ are OP w.r.t. discrete inner product $d\mu_n = \sum_{j=0}^n |w_j|^2 \delta_{\lambda_j}$

$$\|1\|^2 = \sum_j |w_j|^2 = \|w\|^2 \Rightarrow \phi_0 = 1/\|w\|.$$

Polynomials

- $V_m^* V_m = I_m$ i.e., $\delta_{k,\ell} = v_k^* v_\ell$

$$\begin{aligned}
\delta_{k,\ell} &= (\phi_k(A)v)^*(\phi_\ell(A)v) \\
&= v^*(\phi_k(A))^*(\phi_\ell(A))v, \qquad A = U^*\Lambda U \text{ or } \phi(A) = U^*\phi(\Lambda)U \\
&= v^*U^*(\phi_k(\Lambda))^* UU^*(\phi_\ell(\Lambda))Uv, \quad w = Uv \text{ and } UU^* = I_n \\
&= w^*(\phi_k(\Lambda))^*(\phi_\ell(\Lambda))w \\
&= \sum_{j=0}^{n} |w_j|^2 \overline{\phi_k(\lambda_j)}\phi_\ell(\lambda_j)
\end{aligned}$$

The $\{\phi_k\}$ are OP w.r.t. discrete inner product $d\mu_n = \sum_{j=0}^n |w_j|^2 \delta_{\lambda_j}$

$$\|1\|^2 = \sum_j |w_j|^2 = \|w\|^2 \Rightarrow \phi_0 = 1/\|w\|.$$

Inverse eigenvalue problem

$$\delta_{k,\ell} = \sum_{j=0}^{n} |w_j|^2 \overline{\phi_k(\lambda_j)} \phi_\ell(\lambda_j) = \sum_{j=0}^{n} \underbrace{\overline{w_j \phi_k(\lambda_j)}}_{\overline{q_{jk}}} \underbrace{w_j \phi_\ell(\lambda_j)}_{q_{j\ell}}$$

with $D_w = \mathrm{diag}(w_0, \ldots, w_n)$ we define

$$\begin{aligned} q_k &= D_w \vec{\phi}_k = D_w [\phi_k(\lambda_0), \ldots, \phi_k(\lambda_n)]^T \\ &= [w_0 \phi_k(\lambda_0), \ldots, w_n \phi_k(\lambda_n)]^T \end{aligned}$$

then $q_k^* q_\ell = \delta_{k,\ell}$
hence

$$Q_m^* Q_m = I_m, \quad Q_m = [q_0, \ldots, q_m]$$

$Q_m^* q_0 = e_0$ but $\phi_0 = 1/\|w\|$ so $q_0 = w/\|w\|$ and thus $Q_m^* w = e_0 \|w\|$

Motivation

Inverse eigenvalue problem

- $AV_m = V_m H_m + b_{m+1}[0, 0, \ldots, v_{m+1}]$,

  $\Leftrightarrow$

  $z\Phi_m(z) = \Phi_m(z)H_m + b_{m+1}[0, \ldots, 0, \phi_{m+1}(z)]$
  where $\Phi_m(z) = [\phi_0(z), \ldots, \phi_m(z)]$.
- Evaluate for $z = \lambda_j$, $j = 0, \ldots, n$ to get
  $\Lambda\Phi_m = \Phi_m H_m + b_{m+1}\vec{\phi}_{m+1}e_m^T, \quad \Phi_m = [\vec{\phi}_0, \ldots, \vec{\phi}_m]$
- multiply with D_w and we get
  $\Lambda Q_m = Q_m H_m + [0, \ldots, 0, *]$
  multiply with Q_m^*: $Q_m^*\Lambda Q_m = H_m$

Motivation

Inverse eigenvalue problem

- $AV_m = V_m H_m + b_{m+1}[0, 0, \ldots, v_{m+1}]$,

  $\Leftrightarrow$

  $z\Phi_m(z) = \Phi_m(z)H_m + b_{m+1}[0, \ldots, 0, \phi_{m+1}(z)]$
  where $\Phi_m(z) = [\phi_0(z), \ldots, \phi_m(z)]$.
- Evaluate for $z = \lambda_j$, $j = 0, \ldots, n$ to get
  $\Lambda\Phi_m = \Phi_m H_m + b_{m+1}\vec{\phi}_{m+1}e_m^T, \quad \Phi_m = [\vec{\phi}_0, \ldots, \vec{\phi}_m]$
- multiply with D_w and we get
  $\Lambda Q_m = Q_m H_m + [0, \ldots, 0, *]$
  multiply with Q_m^*: $Q_m^*\Lambda Q_m = H_m$

Motivation

Inverse eigenvalue problem

- $AV_m = V_m H_m + b_{m+1}[0, 0, \ldots, v_{m+1}]$,
  $\Leftrightarrow$
  $z\Phi_m(z) = \Phi_m(z)H_m + b_{m+1}[0, \ldots, 0, \phi_{m+1}(z)]$
  where $\Phi_m(z) = [\phi_0(z), \ldots, \phi_m(z)]$.
- Evaluate for $z = \lambda_j$, $j = 0, \ldots, n$ to get
  $\Lambda\Phi_m = \Phi_m H_m + b_{m+1}\vec{\phi}_{m+1}e_m^T, \quad \Phi_m = [\vec{\phi}_0, \ldots, \vec{\phi}_m]$
- multiply with D_w and we get
  $\Lambda Q_m = Q_m H_m + [0, \ldots, 0, *]$
  multiply with Q_m^*: $Q_m^*\Lambda Q_m = H_m$

Motivation

Inverse eigenvalue problem

given $\Lambda = \mathrm{diag}(\lambda_0, \ldots, \lambda_n)$ and $w = [w_0, \ldots, w_n]^T$,
find upper Hessenberg H_m and Q_m such that $Q_m^* Q_m = I_m$
$Q_m^* \Lambda Q_m = H_m$ and $Q_m^* w = \|w\| e_0$

Recurrence orthogonal polynomials

given $\langle \cdot, \cdot \rangle$ on $\mathcal{P}_n$
find $\Phi_m(z) = [\phi_0(z), \ldots, \phi_m(z)]$ and H_m such that
$z\Phi_m(z) = \Phi_m(z) H_m + \phi_{m+1}(z) e_m^T$ and $\langle \Phi_m^*, \Phi_m \rangle = I_m$

Note: if $\lambda_0, \ldots, \lambda_m$ are zeros of $\phi_{m+1}(z) =$ eigenvalues of H_m
hence $\Lambda_m \Phi_m = \Phi_m H_m, \quad \Phi_m = [\phi_j(\lambda_i)]_{i,j=0}^m$,
$\tilde{\Phi}_m = D^{-1} \Phi_m$ (normalize rows=eigenvectors), set $w = \tilde{\Phi}_m e_0$
then $\langle p, q \rangle = \sum_{j=0}^m |w_j|^2 \overline{p(\lambda_j)} q(\lambda_j)$, $\forall p, q \in \mathcal{P}_m$.

Motivation

Inverse eigenvalue problem

given $\Lambda = \mathrm{diag}(\lambda_0, \ldots, \lambda_n)$ and $w = [w_0, \ldots, w_n]^T$,
find upper Hessenberg H_m and Q_m such that $Q_m^* Q_m = I_m$
$Q_m^* \Lambda Q_m = H_m$ and $Q_m^* w = \|w\| e_0$

Recurrence orthogonal polynomials

given $\langle \cdot, \cdot \rangle$ on $\mathcal{P}_n$
find $\Phi_m(z) = [\phi_0(z), \ldots, \phi_m(z)]$ and H_m such that
$z\Phi_m(z) = \Phi_m(z) H_m + \phi_{m+1}(z) e_m^T$ and $\langle \Phi_m^*, \Phi_m \rangle = I_m$

Note: if $\lambda_0, \ldots, \lambda_m$ are zeros of $\phi_{m+1}(z) =$ eigenvalues of H_m
hence $\Lambda_m \Phi_m = \Phi_m H_m, \quad \Phi_m = [\phi_j(\lambda_i)]_{i,j=0}^m$,
$\tilde{\Phi}_m = D^{-1} \Phi_m$ (normalize rows=eigenvectors), set $w = \tilde{\Phi}_m e_0$
then $\langle p, q \rangle = \sum_{j=0}^m |w_j|^2 \overline{p(\lambda_j)} q(\lambda_j)$, $\forall p, q \in \mathcal{P}_m$.

Motivation

Inverse eigenvalue problem

given $\Lambda = \mathrm{diag}(\lambda_0, \ldots, \lambda_n)$ and $w = [w_0, \ldots, w_n]^T$,
find upper Hessenberg H_m and Q_m such that $Q_m^* Q_m = I_m$
$Q_m^* \Lambda Q_m = H_m$ and $Q_m^* w = \|w\| e_0$

Recurrence orthogonal polynomials

given $\langle \cdot, \cdot \rangle$ on $\mathcal{P}_n$
find $\Phi_m(z) = [\phi_0(z), \ldots, \phi_m(z)]$ and H_m such that
$z\Phi_m(z) = \Phi_m(z) H_m + \phi_{m+1}(z) e_m^T$ and $\langle \Phi_m^*, \Phi_m \rangle = I_m$

Note: if $\lambda_0, \ldots, \lambda_m$ are zeros of $\phi_{m+1}(z) =$ eigenvalues of H_m
hence $\Lambda_m \Phi_m = \Phi_m H_m, \quad \Phi_m = [\phi_j(\lambda_i)]_{i,j=0}^m$,
$\tilde{\Phi}_m = D^{-1} \Phi_m$ (normalize rows=eigenvectors), set $w = \tilde{\Phi}_m e_0$
then $\langle p, q \rangle = \sum_{j=0}^m |w_j|^2 \overline{p(\lambda_j)} q(\lambda_j)$, $\forall p, q \in \mathcal{P}_m$.

Motivation

rational Krylov

- $K_m = [v, Av, A^2v, \ldots, A^mv] \rightarrow$ implicit = power method
  $\lambda \approx |\sigma(A)|_{\max}$
- $K_m = [v, A^{-1}v, A^{-2}v, \ldots] \rightarrow \lambda \approx |\sigma(A)|_{\min}$
- $K_m = [v, (A-\alpha I)^{-1}v, (A-\alpha I)^{-2}v, \ldots] \rightarrow \lambda \approx |\alpha|$
- $K_m = [\tilde{v}_0, \tilde{v}_1, \tilde{v}_2, \ldots]$ with $\tilde{v}_k = (A-\gamma_k I)^{-1}\tilde{v}_{k-1} = r_k(A)v$,
  $r_k(z)$ rational with poles $\{\gamma_j\}_{j=0}^{k}$
- need orthonormal basis of rational functions (ORF) $\{\phi_k\}$
- want to allow poles at any point in $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$,
  with or without repetition.
- study ORF and their recurrence $\Rightarrow H_m$ and compute its eigenvalues

Motivation

rational Krylov

- $K_m = [v, Av, A^2v, \ldots, A^mv] \rightarrow$ implicit = power method
  $\lambda \approx |\sigma(A)|_{\max}$
- $K_m = [v, A^{-1}v, A^{-2}v, \ldots] \rightarrow \lambda \approx |\sigma(A)|_{\min}$
- $K_m = [v, (A-\alpha I)^{-1}v, (A-\alpha I)^{-2}v, \ldots] \rightarrow \lambda \approx |\alpha|$
- $K_m = [\tilde{v}_0, \tilde{v}_1, \tilde{v}_2, \ldots]$ with $\tilde{v}_k = (A-\gamma_k I)^{-1}\tilde{v}_{k-1} = r_k(A)v$,
  $r_k(z)$ rational with poles $\{\gamma_j\}_{j=0}^{k}$
- need orthonormal basis of rational functions (ORF) $\{\phi_k\}$
- want to allow poles at any point in $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$,
  with or without repetition.
- study ORF and their recurrence $\Rightarrow H_m$ and compute its eigenvalues

Motivation

rational Krylov

- $K_m = [v, Av, A^2v, \ldots, A^mv] \rightarrow$ implicit = power method
  $\lambda \approx |\sigma(A)|_{\max}$
- $K_m = [v, A^{-1}v, A^{-2}v, \ldots] \rightarrow \lambda \approx |\sigma(A)|_{\min}$
- $K_m = [v, (A-\alpha I)^{-1}v, (A-\alpha I)^{-2}v, \ldots] \rightarrow \lambda \approx |\alpha|$
- $K_m = [\tilde{v}_0, \tilde{v}_1, \tilde{v}_2, \ldots]$ with $\tilde{v}_k = (A-\gamma_k I)^{-1}\tilde{v}_{k-1} = r_k(A)v$,
  $r_k(z)$ rational with poles $\{\gamma_j\}_{j=0}^{k}$
- need orthonormal basis of rational functions (ORF) $\{\phi_k\}$
- want to allow poles at any point in $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$,
  with or without repetition.
- study ORF and their recurrence $\Rightarrow H_m$ and compute its eigenvalues

Motivation

rational Krylov

- $K_m = [v, Av, A^2v, \ldots, A^mv] \rightarrow$ implicit = power method
  $\lambda \approx |\sigma(A)|_{\max}$
- $K_m = [v, A^{-1}v, A^{-2}v, \ldots] \rightarrow \lambda \approx |\sigma(A)|_{\min}$
- $K_m = [v, (A-\alpha I)^{-1}v, (A-\alpha I)^{-2}v, \ldots] \rightarrow \lambda \approx |\alpha|$
- $K_m = [\tilde{v}_0, \tilde{v}_1, \tilde{v}_2, \ldots]$ with $\tilde{v}_k = (A-\gamma_k I)^{-1}\tilde{v}_{k-1} = r_k(A)v$,
  $r_k(z)$ rational with poles $\{\gamma_j\}_{j=0}^k$
- need orthonormal basis of rational functions (ORF) $\{\phi_k\}$
- want to allow poles at any point in $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$,
  with or without repetition.
- study ORF and their recurrence $\Rightarrow H_m$ and compute its eigenvalues

Motivation

rational Krylov

- $K_m = [v, Av, A^2v, \ldots, A^mv] \rightarrow$ implicit = power method
  $\lambda \approx |\sigma(A)|_{\max}$
- $K_m = [v, A^{-1}v, A^{-2}v, \ldots] \rightarrow \lambda \approx |\sigma(A)|_{\min}$
- $K_m = [v, (A-\alpha I)^{-1}v, (A-\alpha I)^{-2}v, \ldots] \rightarrow \lambda \approx |\alpha|$
- $K_m = [\tilde{v}_0, \tilde{v}_1, \tilde{v}_2, \ldots]$ with $\tilde{v}_k = (A-\gamma_k I)^{-1}\tilde{v}_{k-1} = r_k(A)v$,
  $r_k(z)$ rational with poles $\{\gamma_j\}_{j=0}^{k}$
- need orthonormal basis of rational functions (ORF) $\{\phi_k\}$
- want to allow poles at any point in $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$,
  with or without repetition.
- study ORF and their recurrence $\Rightarrow H_m$ and compute its eigenvalues

Motivation

rational Krylov

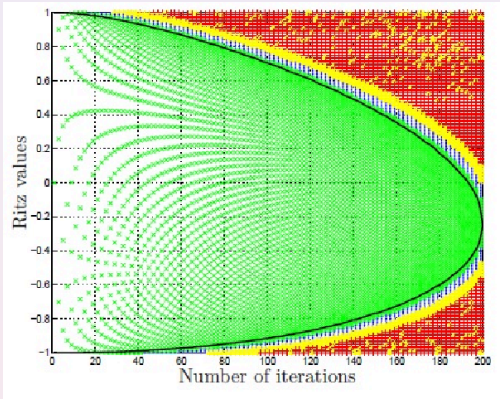
- $K_m = [v, Av, A^2v, \ldots, A^mv] \rightarrow$ implicit = power method
  $\lambda \approx |\sigma(A)|_{\max}$
- $K_m = [v, A^{-1}v, A^{-2}v, \ldots] \rightarrow \lambda \approx |\sigma(A)|_{\min}$
- $K_m = [v, (A-\alpha I)^{-1}v, (A-\alpha I)^{-2}v, \ldots] \rightarrow \lambda \approx |\alpha|$
- $K_m = [\tilde{v}_0, \tilde{v}_1, \tilde{v}_2, \ldots]$ with $\tilde{v}_k = (A-\gamma_k I)^{-1}\tilde{v}_{k-1} = r_k(A)v$,
  $r_k(z)$ rational with poles $\{\gamma_j\}_{j=0}^k$
- need orthonormal basis of rational functions (ORF) $\{\phi_k\}$
- want to allow poles at any point in $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$,
  with or without repetition.
- study ORF and their recurrence $\Rightarrow H_m$ and compute its eigenvalues

rational Krylov

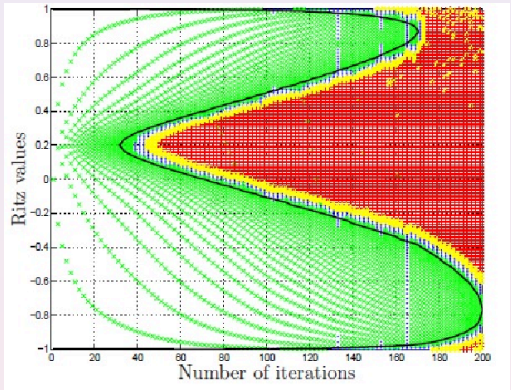
- $K_m = [v, Av, A^2v, \ldots, A^mv] \rightarrow$ implicit = power method
  $\lambda \approx |\sigma(A)|_{\max}$
- $K_m = [v, A^{-1}v, A^{-2}v, \ldots] \rightarrow \lambda \approx |\sigma(A)|_{\min}$
- $K_m = [v, (A-\alpha I)^{-1}v, (A-\alpha I)^{-2}v, \ldots] \rightarrow \lambda \approx |\alpha|$
- $K_m = [\tilde{v}_0, \tilde{v}_1, \tilde{v}_2, \ldots]$ with $\tilde{v}_k = (A-\gamma_k I)^{-1}\tilde{v}_{k-1} = r_k(A)v$,
  $r_k(z)$ rational with poles $\{\gamma_j\}_{j=0}^k$
- need orthonormal basis of rational functions (ORF) $\{\phi_k\}$
- want to allow poles at any point in $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$,
  with or without repetition.
- study ORF and their recurrence $\Rightarrow H_m$ and compute its eigenvalues

Motivation

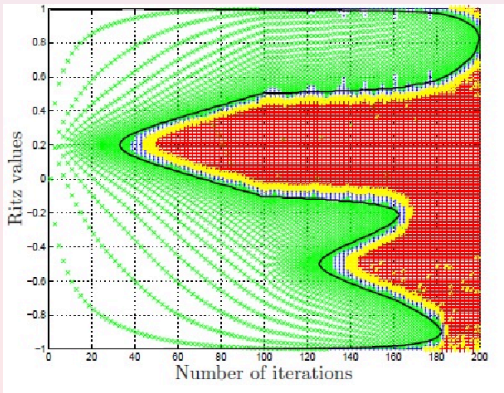
rational Krylov: equispaced eigenvalues in $[-1, 1]$



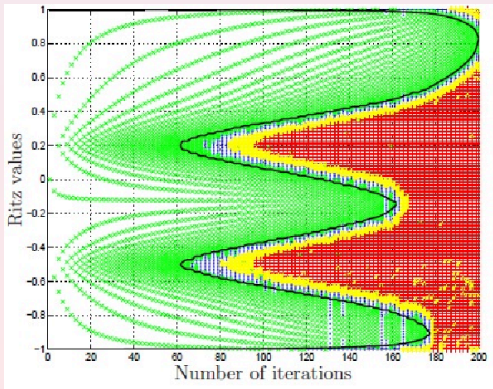
$200 \times \alpha = 2$



$200 \times \alpha = 0.2 + 0.1i$



$100 \times \alpha_1 = 0.2 + 0.1i$
$100 \times \alpha_2 = -0.5 + 0.1i$



$100 \times (\alpha_1 = 0.2 + 0.1i$
$\alpha_2 = -0.5 + 0.1i)$

Special cases

Polynomials

- $A = A^T$ and $Q^* A Q = H$ then $H = H^T$ and Hessenberg
  $\Rightarrow$ H tridiagonal (Jacobi) & $\lambda_j \in \mathbb{R}$: OPS on $\mathbb{R}$
- $A^* A = I$ and $Q^* A Q = H$ then $H^* H = I$ and Hessenberg
  $\Rightarrow$ H unitary Hessenberg & $\lambda_j \in \mathbb{T}$: OPS on $\mathbb{T}$

In this lecture assume unitary Hessenberg $\Rightarrow$ ORF on $\mathbb{T}$

Special cases

Polynomials

- $A = A^T$ and $Q^*AQ = H$ then $H = H^T$ and Hessenberg
  $\Rightarrow$ H tridiagonal (Jacobi) & $\lambda_j \in \mathbb{R}$: OPS on $\mathbb{R}$
- $A^*A = I$ and $Q^*AQ = H$ then $H^*H = I$ and Hessenberg
  $\Rightarrow$ H unitary Hessenberg & $\lambda_j \in \mathbb{T}$: OPS on $\mathbb{T}$

In this lecture assume unitary Hessenberg $\Rightarrow$ ORF on $\mathbb{T}$

Special cases

Polynomials

- $A = A^T$ and $Q^*AQ = H$ then $H = H^T$ and Hessenberg
  $\Rightarrow$ H tridiagonal (Jacobi) & $\lambda_j \in \mathbb{R}$: OPS on $\mathbb{R}$
- $A^*A = I$ and $Q^*AQ = H$ then $H^*H = I$ and Hessenberg
  $\Rightarrow$ H unitary Hessenberg & $\lambda_j \in \mathbb{T}$: OPS on $\mathbb{T}$

In this lecture assume unitary Hessenberg $\Rightarrow$ ORF on $\mathbb{T}$

Special cases

Computing eigenvalues of H by QR

$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

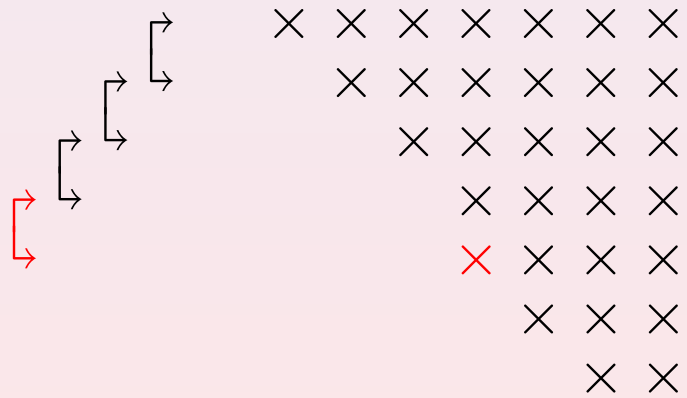
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

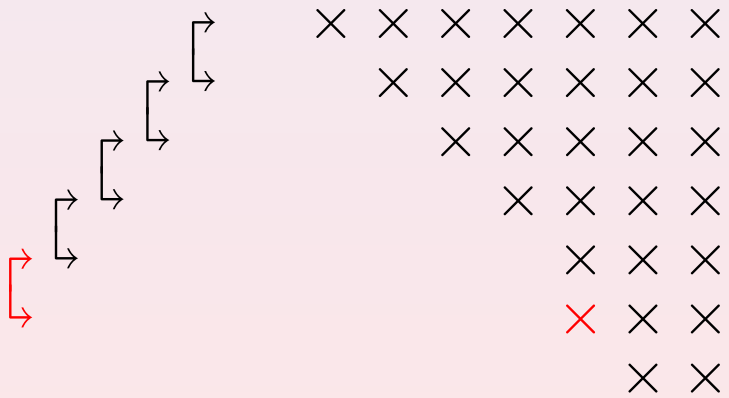
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

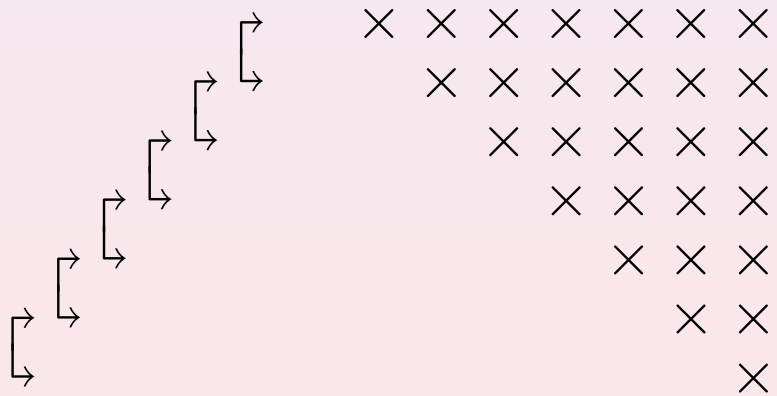
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

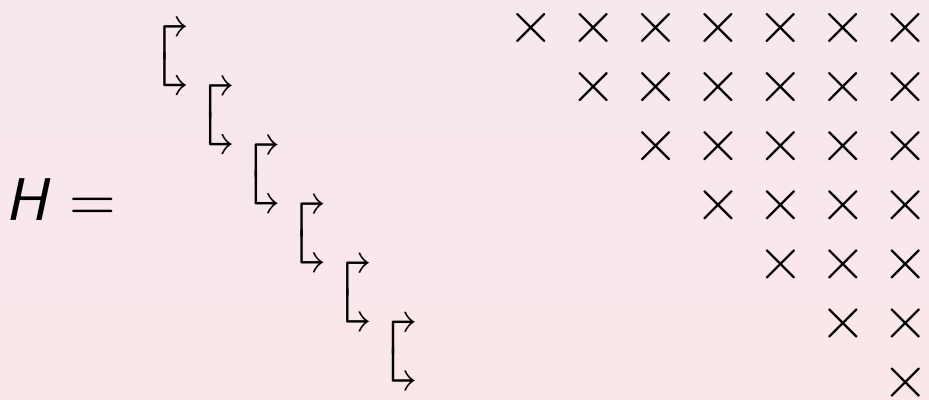
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

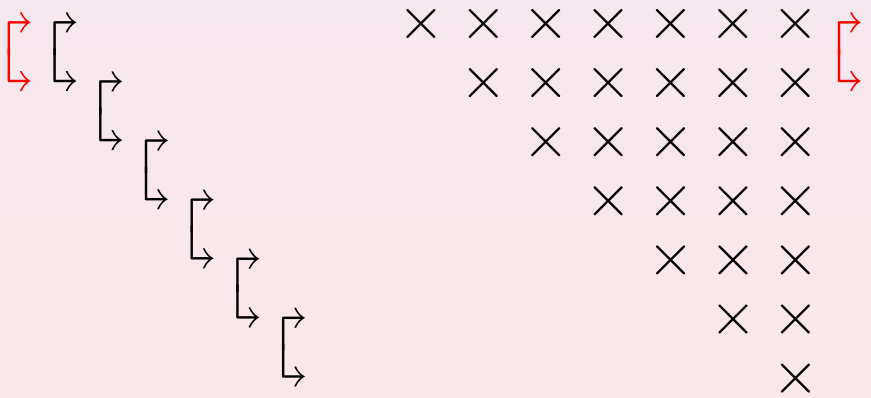
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$

$H =$ 

Special cases

Computing eigenvalues of H by QR

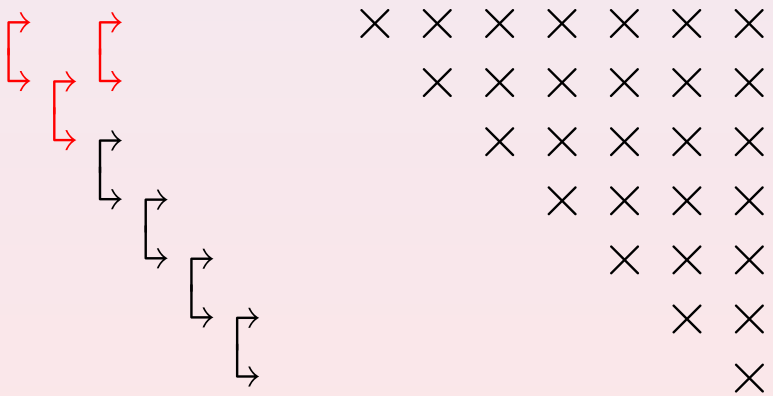
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

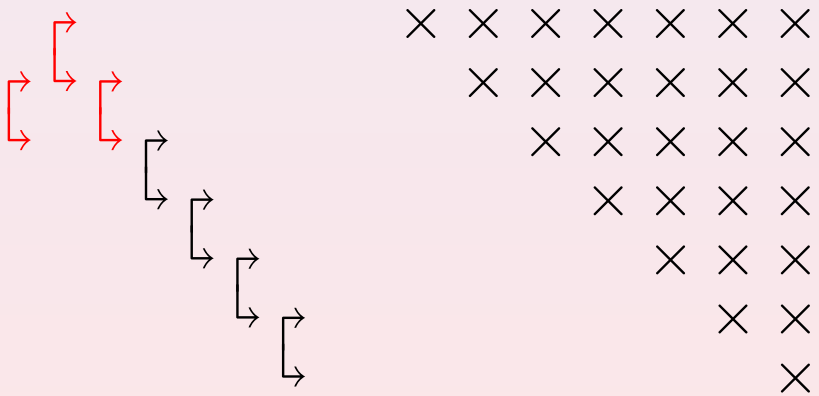
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

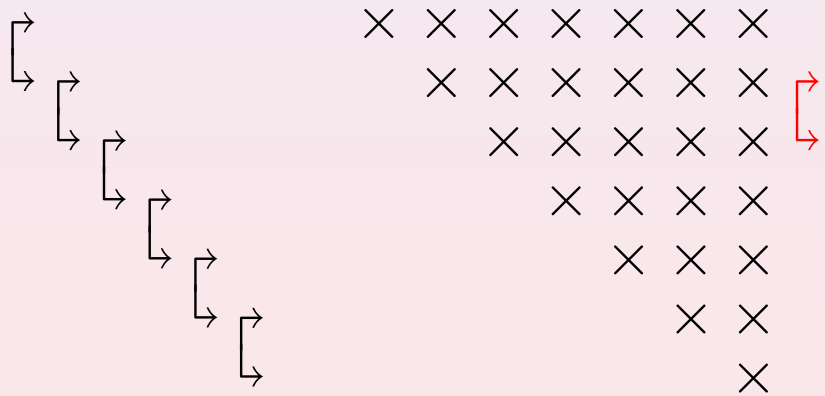
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

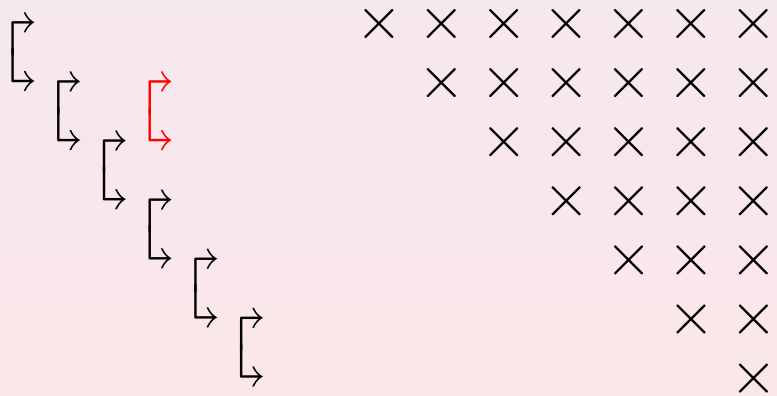
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

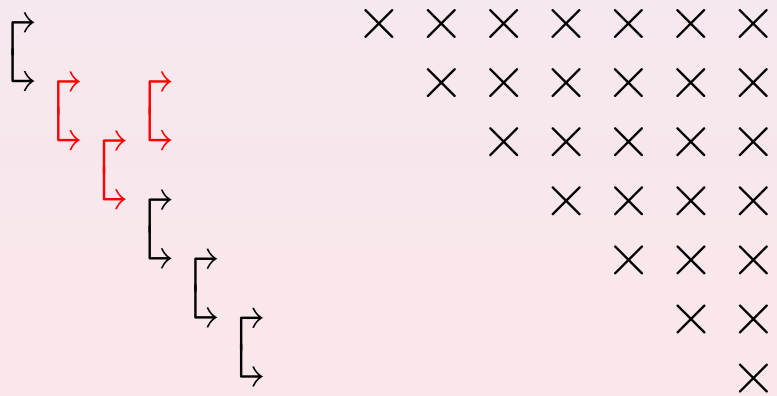
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

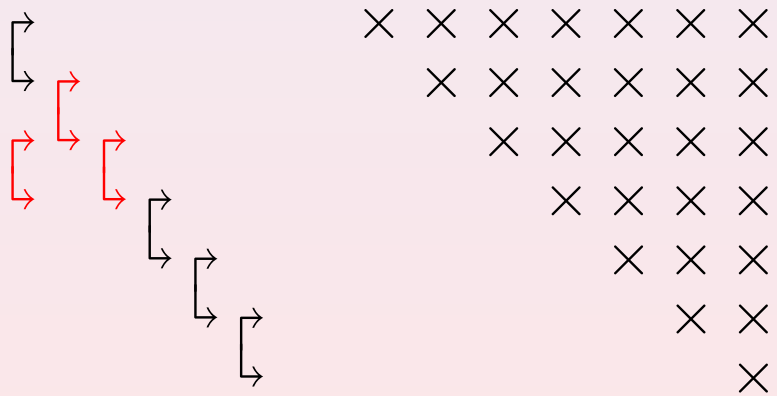
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

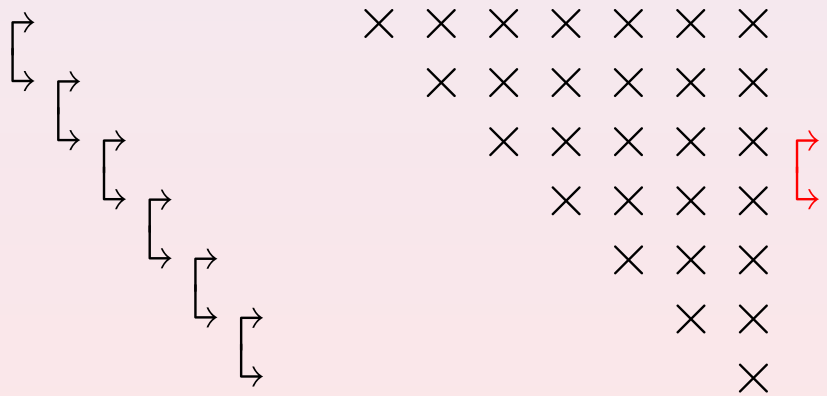
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

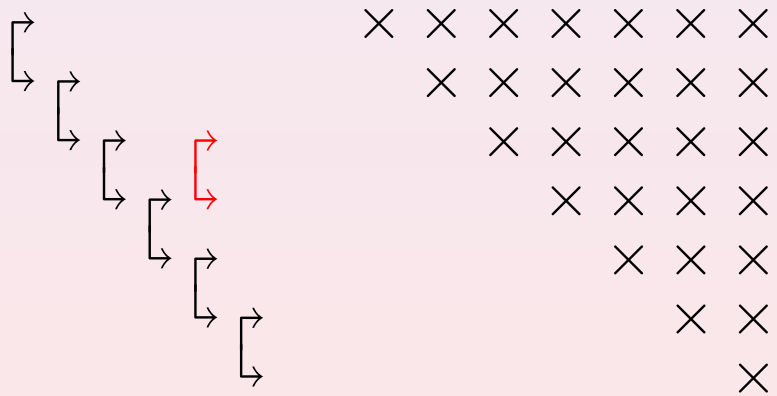
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

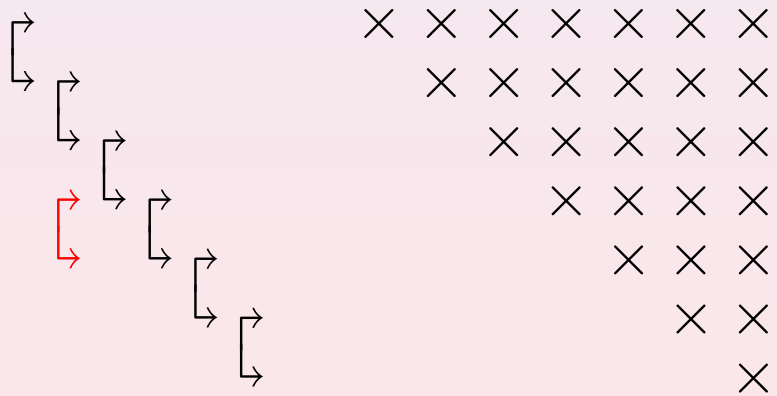
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

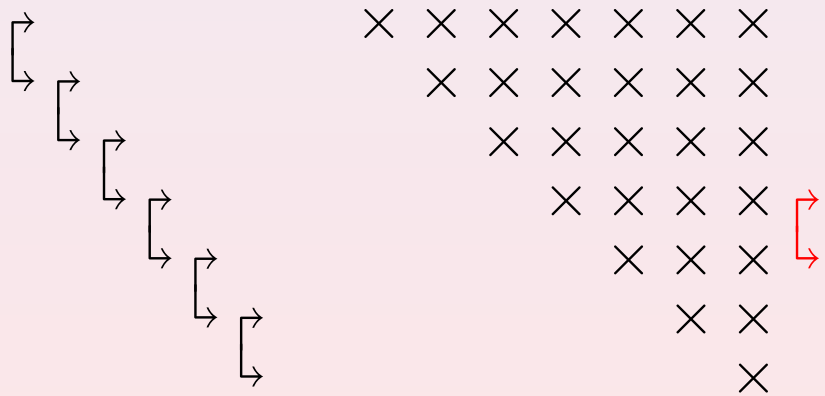
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

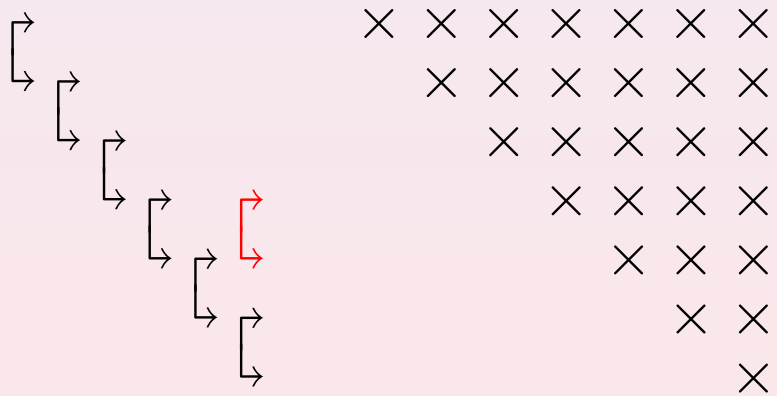
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

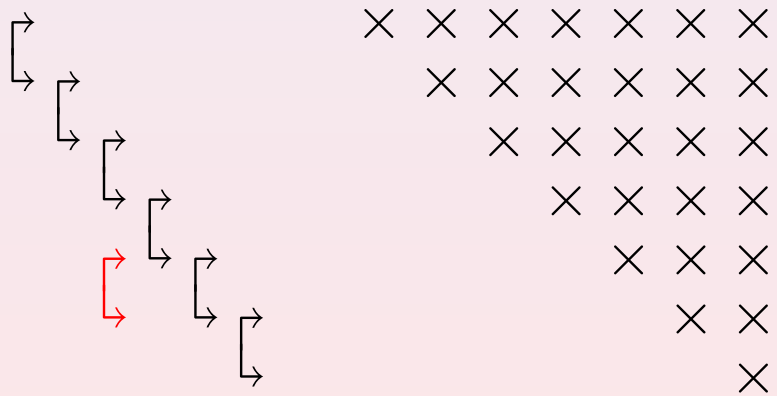
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

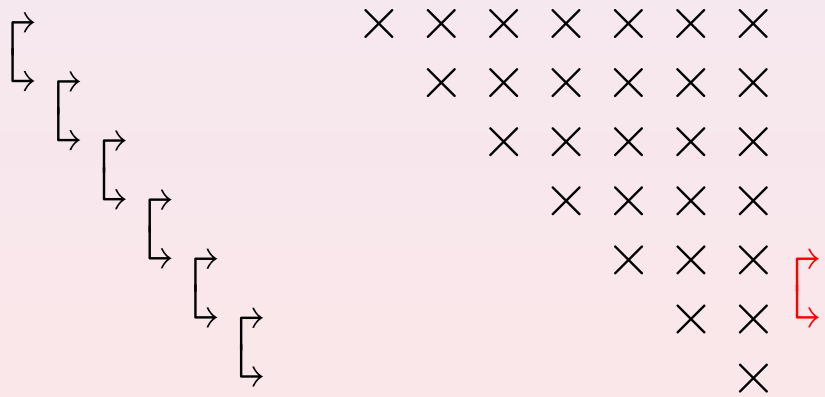
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

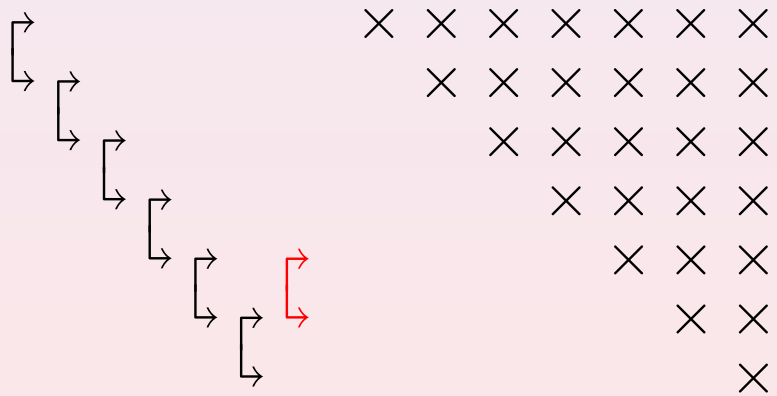
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

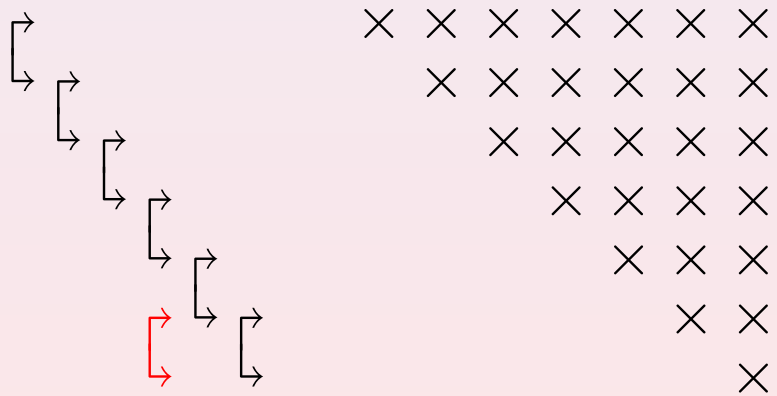
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

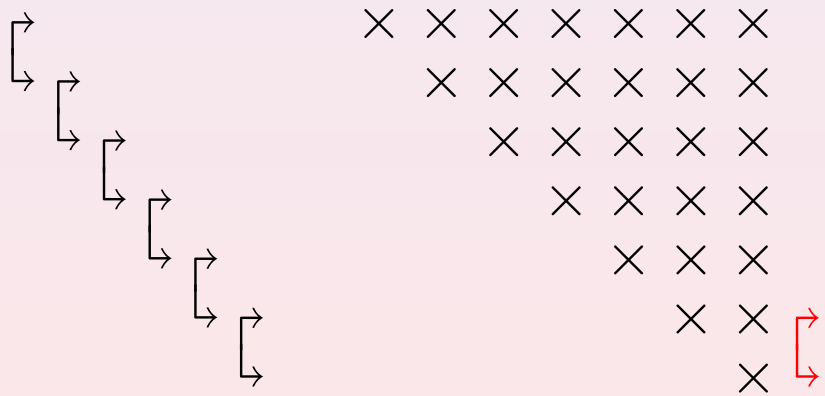
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

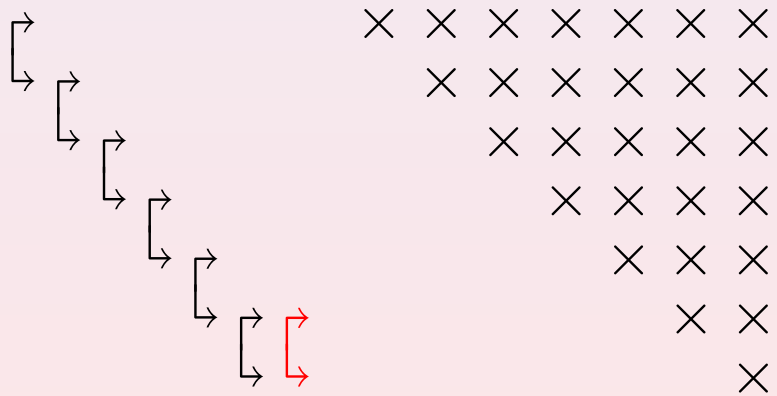
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

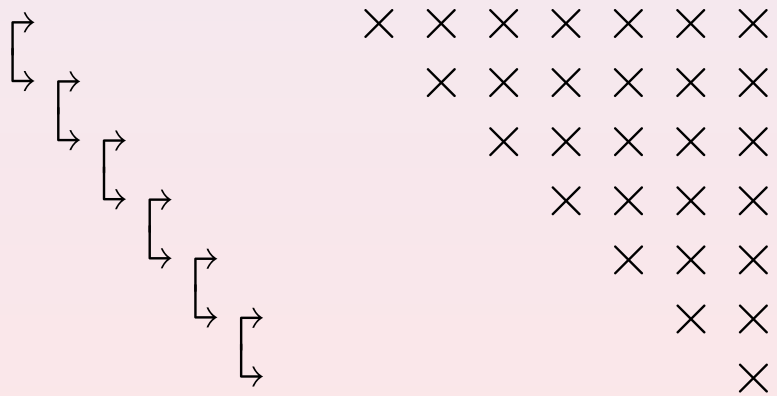
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Computing eigenvalues of H by QR

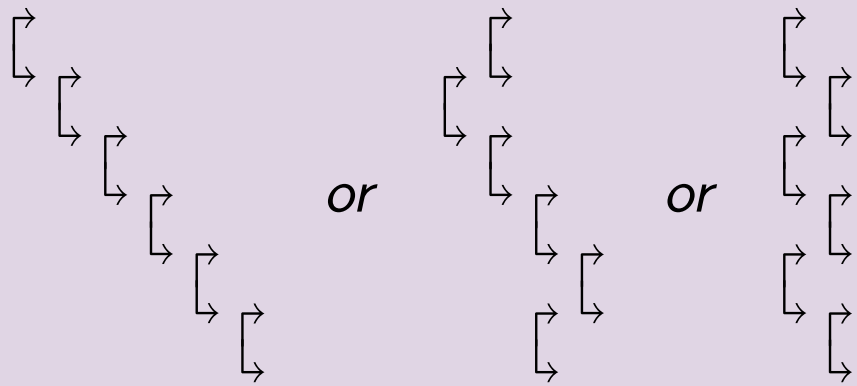
$$
\begin{aligned}
H = Q_0 R_0 \quad &\rightarrow \quad H_1 = R_0 Q_0 = Q_0^* H Q_0 = Q_1 R_1 \\
&\rightarrow \quad H_2 = R_1 Q_1 = Q_1^* Q_0^* H Q_0 Q_1 = \cdots
\end{aligned}
$$



Special cases

Theorem

If $H = G_1 \cdots G_n$ is a unitary Hessenberg matrix factored as a product of core transforms then its spectrum does not depend on the order in which the core transforms are mutiplied.



all have the same spectrum.

Note: The latter is the CMV form (Cantero-Moral-Velázquez)
$(G_1 G_3 G_5 \cdots)(G_2 G_4 G_6 \cdots)$ is not Hessenberg but 5-diagonal.

Cantero-Moral-Velázquez



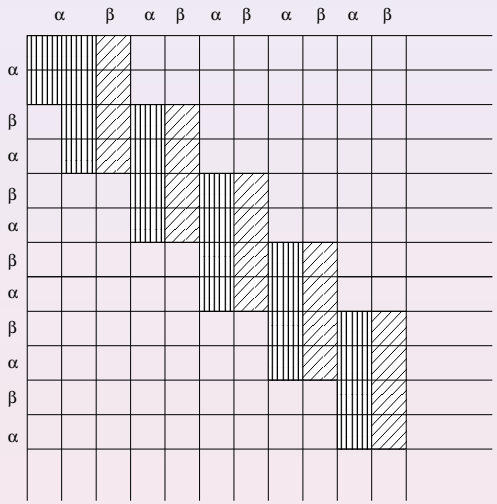
Shape of the matrix

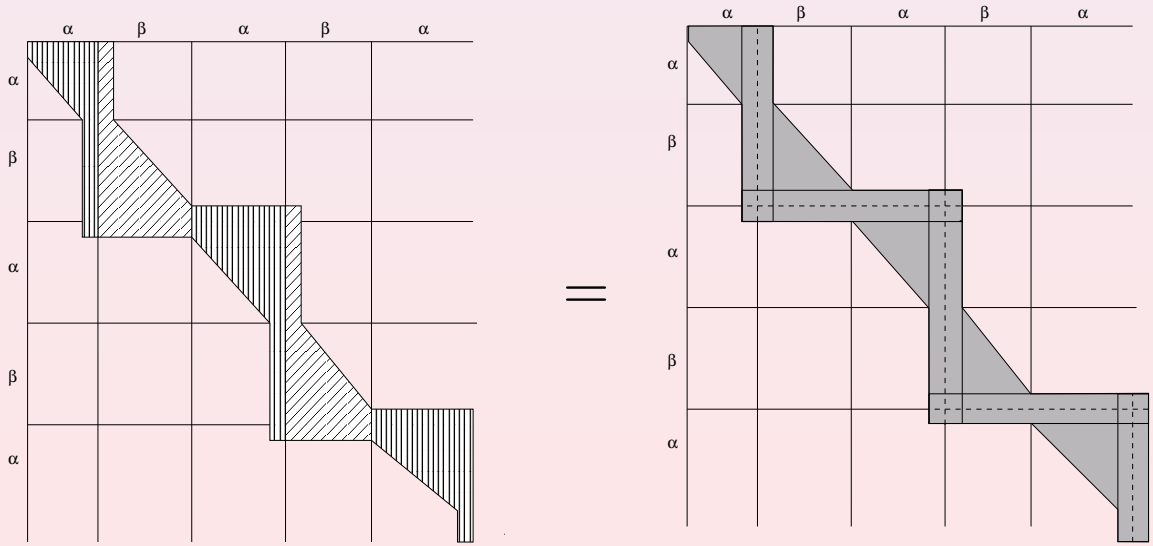
The order in which one multiplies defines the shape of the product.
For example CMV (5-diagonal)

$$\begin{bmatrix} c_{00} & c_{01} & c_{02} & 0 & 0 & 0 & 0 & \cdots \\ c_{10} & c_{11} & c_{12} & 0 & 0 & 0 & 0 & \cdots \\ 0 & c_{21} & c_{22} & c_{23} & c_{24} & 0 & 0 & \cdots \\ 0 & c_{31} & c_{32} & c_{33} & c_{34} & 0 & \cdots & \\ 0 & 0 & 0 & c_{43} & c_{44} & c_{45} & 0 & \cdots \\ 0 & 0 & 0 & c_{53} & c_{54} & c_{55} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \mathcal{C}_o(\delta)\mathcal{C}_e(\delta)$$

$$\mathcal{C}_o(\delta) = \begin{bmatrix} -\delta_1 & \epsilon_1 & 0 & 0 & \cdots \\ \epsilon_1 & \delta_1 & 0 & 0 & \cdots \\ 0 & 0 & -\delta_3 & \epsilon_3 & \cdots \\ 0 & 0 & \epsilon_3 & \delta_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}; \quad \mathcal{C}_e(\delta) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & -\delta_2 & \epsilon_2 & 0 & 0 & \cdots \\ 0 & \epsilon_2 & \delta_2 & 0 & 0 & \cdots \\ 0 & 0 & 0 & -\delta_4 & \epsilon_4 & \cdots \\ 0 & 0 & 0 & \epsilon_4 & \delta_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Shape of the matrix





What to do?

How to generalize to the rational case

- Is there a Hessenberg matrix associated with ORF?
- Can it be written as a product of elementary unitary factors?
- Does the order of these factors matter?
- If not what is the most efficient ordering?

Generalization of the polynomial case - Definitions

$$\mathbb{P}_n = \text{span}\{z^k : k = 0, \ldots, n\} \longrightarrow \mathcal{L}_n = \frac{\mathbb{P}_n}{\pi_n(z)}$$

$$\pi_n(z) = \prod_{k=1}^{n} (1 - \overline{\alpha}_k z)$$

Define

$$\zeta_k(z) = \frac{z - \alpha_k}{1 - \overline{\alpha}_k z}, \quad k = 1, 2, \ldots$$

$$B_0 = 1, \quad B_n(z) = B_{n-1}(z)\zeta_n(z) = \prod_{k=1}^{n} \zeta_k(z), \quad n = 1, 2, \ldots$$

Then

$$\Rightarrow \mathcal{L}_n = \text{span}\{B_0, \ldots, B_n\}$$

Note, if all $\alpha_k = 0$ then $B_n(z) = z^n$ and $\mathcal{L}_n = \mathbb{P}_n$.

Generalization of the polynomial case - Recurrence

As before: $f_*(z) = \overline{f(1/\overline{z})}, \quad \langle f, g \rangle = \int_{\mathbb{T}} f(z) g_*(z) d\mu(z)$
ORF: $\varphi_n \in \mathcal{L}_n \setminus \mathcal{L}_{n-1}$ and $\varphi_n \perp \mathcal{L}_{n-1}$
Recurrence: initialize with $[\varphi_0, \varphi_0^*] = [1,\ 1]$ and

$$\eta_n \begin{bmatrix} \varphi_n(z) \\ \varphi_n^*(z) \end{bmatrix} = \frac{1 - \overline{\alpha}_{n-1} z}{1 - \overline{\alpha}_n z} \begin{bmatrix} 1 & \delta_n \\ \overline{\delta_n} & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n-1}(z) \varphi_{n-1}(z) \\ \varphi_{n-1}^*(z) \end{bmatrix}, \quad n \geq 1$$

Note however

$$OP : \varphi_n^*(z) = z^n \varphi_{n*}(z) \quad \longrightarrow \quad ORF : \varphi_n^*(z) = B_n(z) \varphi_{n*}(z).$$

$$\begin{array}{ccc} p_n(z) = a_0 + a_1 z + \cdots + a_n z^n & & \phi_n(z) = p_n(z)/\pi_n(z) \\ \Downarrow & \longrightarrow & \Downarrow \\ p_n^*(z) = \overline{a}_n + \cdots + \overline{a}_1 z^{n-1} + \overline{a}_0 z^n & & \phi_n^*(z) = p_n^*(z)/\pi_n(z) \end{array}$$

Shift operator

$$z\mathcal{I} : \mathcal{L} \to \mathcal{L} : f(z) \mapsto zf(z), \quad \mathcal{L} = \text{clos.span}\{\phi_0, \phi_1, \ldots\}$$

This means that

$$f(z) = \sum_{k=0}^{\infty} a_k \varphi_k(z) \quad \& \quad zf(z) = \sum_{k=0}^{\infty} b_k \varphi_k(z)$$

or with $\Phi(z) = [\phi_0, \phi_1(z), \phi_2(z), \ldots]$

$$f(z) = \Phi(z)\mathbf{a} \quad \text{and} \quad zf(z) = \Phi(z)\mathbf{b}$$

then by $z\Phi(z) = \Phi(z)\mathcal{H}$:

$$zf(z) = z\Phi(z)\mathbf{a} = \Phi(z)\mathcal{H}\mathbf{a} = \Phi(z)\mathbf{b} \quad \text{or} \quad \mathbf{b} = \mathcal{H}\mathbf{a}$$

Shift operator

$z\mathcal{I} : \mathcal{L} \to \mathcal{L} : f(z) \mapsto zf(z)$ was Hessenberg $\mathcal{H}$ in OP case

Intuition: (how did we move from KRY to OP to ORF?)

$$KRY : \text{span}\{v, Av, A^2v, \ldots, A^nv\}, \quad A^k = A \cdot A \cdots A$$
$$OP : \text{span}\{1, z, \ldots, z^n\}, \quad z^k = z \cdot z \cdots z$$
$$ORF : \text{span}\{B_0, B_1(z), \ldots, B_n(z)\}, \quad B_k(z) = \zeta_1(z) \cdot \zeta_2(z) \cdots \zeta_k(z)$$

Hence if $\Phi(z) = [\phi_0, \phi_1(z), \ldots]$ and $\Phi(A)v = [\phi_0(A)v, \phi_1(A)v, \ldots]$ then

$$\text{KRY} : \quad \Phi(A)Av = \Phi(A^*)v\mathcal{H}$$
$$\text{OP} : \quad \Phi(z)z\mathcal{I} = \Phi(z)\mathcal{H}$$
$$\text{ORF} : \quad \Phi(z)(z\mathcal{I} - \mathcal{A})\eta_{\mathcal{A}}^{-1} = \Phi(z)(\mathcal{I} - \mathcal{A}^*z)\eta_{\mathcal{A}}^{-1}\mathcal{H},$$
$$\mathcal{A} = \text{diag}(\alpha_0, \alpha_1, \ldots), \quad \eta_{\mathcal{A}} = (\mathcal{I} - \mathcal{A}^*\mathcal{A})^{1/2}$$

Shift operator

$$z\mathcal{I} : \mathcal{L} \to \mathcal{L} : f(z) \mapsto zf(z)$$

$$\Phi(z)(z\mathcal{I} - \mathcal{A})\eta_{\mathcal{A}}^{-1} = \Phi(z)(\mathcal{I} - \mathcal{A}^* z)\eta_{\mathcal{A}}^{-1}\mathcal{H}$$

If $\Phi(z)$ complete in $\mathcal{L}$ then

$$(z\mathcal{I} - \mathcal{A})\eta_{\mathcal{A}}^{-1} = (\mathcal{I} - \mathcal{A}^* z)\eta_{\mathcal{A}}^{-1}\mathcal{H}$$

or

$$\mathcal{H} = \zeta_{\mathcal{A}}(z\mathcal{I}) = \eta_{\mathcal{A}}^{-1}(\mathcal{I} - \mathcal{A}^* z)^{-1}(z\mathcal{I} - \mathcal{A})\eta_{\mathcal{A}}$$

or

$$z\mathcal{I} = \zeta_{\mathcal{A}}^{-1}(\mathcal{H}) = \eta_{\mathcal{A}}^{-1}(\mathcal{H} + \mathcal{A})(\mathcal{I} + \mathcal{A}^*\mathcal{H})^{-1}\eta_{\mathcal{A}}$$

Hessenberg factorization

From recursion:

$$[(z-\alpha_{n-1})\phi_{n-1}(z)|(1-\overline{\alpha}_n z)\phi_n^*(z)]\eta_n^{-1}$$
$$= [(1-\overline{\alpha}_{n-1}z)\phi_{n-1}^*(z)|(1-\overline{\alpha}_n z)\phi_n(z)]\eta_n^{-1} G_n$$

$$\eta_n = \begin{bmatrix} \sqrt{1-|\alpha_{n-1}|^2} & 0 \\ 0 & \sqrt{1-|\alpha_n|^2} \end{bmatrix}$$

$$G_n = \begin{bmatrix} \delta_n & \epsilon_n \\ \epsilon_n & \overline{\delta}_n \end{bmatrix}, \quad \epsilon_n = \sqrt{1-|\delta_n|^2}$$

with appropriate normalization of the $\{\phi_k\}$.

Problem: $|\alpha_k| > 1$ and $|\delta_n| > 1$

Hessenberg factorization

From recursion:

$$[(z-\alpha_{n-1})\phi_{n-1}(z)|(1-\overline{\alpha}_n z)\phi_n^*(z)]\eta_n^{-1} = [(1-\overline{\alpha}_{n-1}z)\phi_{n-1}^*(z)|(1-\overline{\alpha}_n z)\phi_n(z)]\eta_n^{-1}G_n$$

$$\eta_n = \begin{bmatrix} \sqrt{1-|\alpha_{n-1}|^2} & 0 \\ 0 & \sqrt{1-|\alpha_n|^2} \end{bmatrix}$$

$$G_n = \begin{bmatrix} \delta_n & \epsilon_n \\ \epsilon_n & \overline{\delta}_n \end{bmatrix}, \quad \epsilon_n = \sqrt{1-|\delta_n|^2}$$

with appropriate normalization of the $\{\phi_k\}$.
Problem: $|\alpha_k| > 1$ and $|\delta_n| > 1$

Hessenberg factorization

Problem: $|\alpha_k| > 1$ and $|\delta_n| > 1$

Solutions:

Classical: take all $\alpha_k \in \mathbb{D} \Rightarrow \mathcal{H} = G_1 G_2 G_3 \cdots$ is full Hessenberg

CMV: Use previous ϕ_n^α if $\alpha_n \in \mathbb{D}$ and ϕ_n^β, $\beta_n = 1/\overline{\alpha}_n$ if $\alpha_n \in \mathbb{E}$

$$\phi_n^\alpha(z) = \frac{p_n^\alpha(z)}{\prod_k (1 - \overline{\alpha}_k z)}, \quad \phi_n^\beta(z) = \frac{p_n^\beta(z)}{\prod_k (1 - \overline{\beta}_k z)}$$

The G_j for a β-block have to be multiplied in reverse order

$$\mathcal{H} = \underbrace{G_1 \cdots G_{n-1}}_{\alpha's} \underbrace{G_k G_{k-1} \cdots G_n}_{\beta's} \underbrace{G_{k+1} G_{k+2} \cdots}_{\alpha's} \cdots$$

Classical CMV: alternate $\alpha_0, \beta_1, \alpha_2, \beta_3, \ldots$

$$\mathcal{H} = G_2 G_1 \; G_4 G_3 \; G_6 G_5 \; \cdots = (G_2 G_4 G_6 \cdots)(G_1 G_3 G_5 \cdots) = G_e G_o$$

Hessenberg factorization

$$[(z-\alpha_{n-1})\phi_{n-1}(z)|(1-\overline{\alpha}_n z)\phi_n^*(z)]\eta_n^{-1}$$
$$= [(1-\overline{\alpha}_{n-1}z)\phi_{n-1}^*(z)|(1-\overline{\alpha}_n z)\phi_n(z)]\eta_n^{-1} G_n$$

$$\eta_n = \begin{bmatrix} \sqrt{1-|\alpha_{n-1}|^2} & 0 \\ 0 & \sqrt{1-|\alpha_n|^2} \end{bmatrix}$$

$$G_n = \begin{bmatrix} \delta_n & \epsilon_n \\ \epsilon_n & \overline{\delta}_n \end{bmatrix}, \quad \epsilon_n = \sqrt{1-|\delta_n|^2}$$

with appropriate normalization of the $\{\phi_k\}$.

Can still be used but not for $\alpha_k = \infty$.

Solution: Switch to $\beta_n = 1/\overline{\alpha}_n$ only if $\alpha_n = \infty$

Then use reverse order for the product of the G_k only for the $\alpha_k = \infty$.

Hessenberg factorization

$$[(z-\alpha_{n-1})\phi_{n-1}(z)|(1-\overline{\alpha}_n z)\phi_n^*(z)]\eta_n^{-1}$$
$$= [(1-\overline{\alpha}_{n-1}z)\phi_{n-1}^*(z)|(1-\overline{\alpha}_n z)\phi_n(z)]\eta_n^{-1} G_n$$

$$\eta_n = \begin{bmatrix} \sqrt{1-|\alpha_{n-1}|^2} & 0 \\ 0 & \sqrt{1-|\alpha_n|^2} \end{bmatrix}$$

$$G_n = \begin{bmatrix} \delta_n & \epsilon_n \\ \epsilon_n & \overline{\delta}_n \end{bmatrix}, \quad \epsilon_n = \sqrt{1-|\delta_n|^2}$$

with appropriate normalization of the $\{\phi_k\}$.
Can still be used but not for $\alpha_k = \infty$.
Solution: Switch to $\beta_n = 1/\overline{\alpha}_n$ only if $\alpha_n = \infty$
Then use reverse order for the product of the G_k only for the $\alpha_k = \infty$.

Alternatives from literature

$$\Phi(z)(z\mathcal{I} - \mathcal{A})\eta_{\mathcal{A}}^{-1} = \Phi(z)(\mathcal{I} - \mathcal{A}^* z)\eta_{\mathcal{A}}^{-1}\mathcal{H}$$

says that

$$(z - \alpha_n)\phi_n(z) \in \operatorname{span}\{\Phi_{n+1}(z)(\mathcal{I}_{n+1} - \mathcal{A}_{n+1}^* z)\} = \frac{\mathbb{P}_{n+1}}{\pi_n(z)}$$

But of course then also

$$\phi_n(z) = \frac{p_n(z)}{\pi_n(z)} \in \operatorname{span}\{\Phi_{n+1}(z)(\mathcal{I}_{n+1} - \mathcal{A}_{n+1}^* z)\} = \frac{\mathbb{P}_{n+1}}{\pi_n(z)}$$

leading to

$$\Phi(z) = \Phi(z)(\mathcal{I} - \mathcal{A}^* z)\mathcal{H}', \quad \mathcal{H}' \text{ upper Hessenberg}$$

Alternatives from literature

M. Van Barel, D. Fasino, L. Gemignani, N. Mastronardi (2005) use denomitator $\pi_n(z) = \prod_k (z - \alpha_k)$ instead of our $\prod_k (1 - \overline{\alpha}_k z)$. So

$$\Phi(z) = \Phi(z)(z\mathcal{I} - \mathcal{A})\mathcal{H}', \quad \mathcal{H}' \text{ upper Hessenberg}$$

With $\mathcal{S} = (\mathcal{H}')^{-1}$, which is a semiseparable matrix

$$z\Phi(z) = \Phi(z)(\mathcal{S} + \mathcal{A})$$

with all α_k finite in the context of inverse eigenvalue problem.

Given $(\lambda_k, w_k, \alpha_k)_{k=0}^n$

Find Q_n unitary and $\mathcal{S}_n$ semiseparable such that

$Q_n^* w = e_0 \|w\|$ and $Q_n^* \Lambda Q_n = \mathcal{S}_n + \mathcal{A}_n$

Alternatives from literature

T. Mach, M. Van Barel, R. Vandebril (2014)
extend this to include poles at ∞ (i.e. polynomials $\equiv$ classical Krylov)
using the β-trick for $\alpha_k = \infty$.

In several other papers this is used in a Krylov context to solve PDE of compute matrix functions: $f(A)$ for large networks.

References



M.J. Cantero, L. Moral, L. Velázquez.

Minimal representation of unitary operators
Lin. Alg. Appl., 362:29-56, 2003.



M.J. Cantero, R. Cruz-Barroso, P. González-Vera

A matrix approach to the computation of quadrature formulas on the unit circle
Appl. Numerical Math. 58:296–318, 2008.



L. Velázquez.

Spectral methods for orthogonal rational functions
J. Functional Anal., 254(2):954-986, 2008.



A. Bultheel, M.J. Cantero.

A matricial computation of rational quadrature formulas on the unit circle
Numer. Algorithms, 52(1):47-68, 2009



A. Bultheel, M.J. Cantero, R. Cruz-Barroso.

Matrix methods for quadrature formulas on the unit circle. A survey
J. Comput. Appl. Math., 284:78-100, 2015

References



J.L. Aurentz, T. Mach, R. Vandebril, D.S. Watkins.

Fast and stable unitary QR algorithm

Electr. Trans. Numer. Anal., 44:327-341, 2015



A. Chesnokov, K. Deckers and M. Van Barel.

A numerical solution of the constrained weighted energy problem,

In *J. Comput. Appl. Math*, 235(4):950-965, 2010.



T. Mach, R. Vandebril, M. Van Barel.

Inverse eigenvalue problems for extended Hessenberg and extended tridiagonal matrices,

J. Comput. Appl. Math., 272: 377-398, 2014.



C. Mertens, R. Vandebril.

Short recurrences for computing extended Krylov bases for Hermitian and unitary matrices

Numer. Math., 131(2):303-328, 2015.



M. Van Barel, D. Fasino, L. Gemignani, N. Mastronardi.

Orthogonal rational functions and structured matrices,

SIAM J. Matrix Anal. Appl. 26(3):810-829, 2005.



R. Vandebril, M. Van Barel, and N. Mastronardi.

A quasiseparable approach to solve the definite generalized eigenvalue problem,

SIAM J. Matrix Anal. Appl., 31(1):154-174, 2009.