Last time: eigenvalue problems

$A \in \mathbb{C}^{n\times n}$, find $\lambda \in \mathbb{C}$, $v \in \mathbb{C}^n$ s.t.

$$Av = \lambda v \qquad (v \neq 0)$$

- Power iteration with Rayleigh quotient

Given $x_0 \in \mathbb{C}^n$

For $k = 1, 2, 3, \cdots$

$$\hat{x}_k = A x_{k-1}$$

$$x_k = \hat{x}_k / \|\hat{x}_k\|_2$$

$$m_k = R(x_k) := x_k^* A x_k$$

When $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots$ or $\lambda_1 = \lambda_2 = \cdots = \lambda_r$ but $|\lambda_1| > |\lambda_{r+1}|$

$$|m_k - \lambda_1| \approx \alpha \, |m_{k-1} - \lambda_1| \qquad \alpha = \frac{|\lambda_2|}{|\lambda_1|} \text{ or } \frac{|\lambda_{r+1}|}{|\lambda_1|}$$

No convergence when $|\lambda_1| = |\lambda_2|$

- Rayleigh Quotient iteration

Given $x_0 \in \mathbb{C}^n$, $\mu \in \mathbb{C}$

For $k = 1, 2, 3, \cdots$

$$m_k = R(x_k)$$

Solve $(A - m_k I)\hat{x}_k = x_{k-1}$

$$x_k = \frac{\hat{x}_k}{\|\hat{x}_k\|_2}$$

"linear" rate $\approx \max_{\ell \neq j} \dfrac{|\lambda_j - m_k|}{|\lambda_\ell - m_k|}$

Almost always globally converges for normal A

then $|m_k - \lambda_j| \approx |m_k - \lambda_j|^3$ locally cubically

- Simultaneous power iteration

  Given $Q_0 \in \mathbb{C}^{n\times n}$

  For $k = 1, 2, 3, \cdots$

  $X_k = A Q_{k-1}$

  Compute QR fact. $X_k = Q_k R_k$

  $T_k = Q_k^* A Q_k$

  when $Q_k \to Q$, as $k \to +\infty$,

  we know $T_k = Q_k^* A Q_k = Q_k^* A Q_{k-1} (Q_{k-1}^* Q_k) \approx R_k$

  thus $T_k \to T = Q^* A Q$ as $k \to +\infty$, the Schur form of A

  Convergence is guaranteed when $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$

  the i^{th} eig. val. of T_k $\to$ $|\lambda_i^k - \lambda_i| \approx \dfrac{|\lambda_{i+1}|}{|\lambda_i|} |\lambda_i^{k-1} - \lambda_i|$

  lower Δ entries $\to$ $|(T_k)_{ij}| \approx \dfrac{|\lambda_{j+1}|}{|\lambda_j|} |(T_{k-1})_{ij}|, \forall i > j$

$$
\begin{aligned}
T_{k-1} &= Q_{k-1}^* A Q_{k-1} \\
&= (Q_{k-1}^* Q_k) R_k \\
T_k &= Q_k^* A Q_k \\
&= Q_k^* A Q_{k-1} Q_{k-1}^* Q_k \\
&= Q_k^* Q_k R_k Q_{k-1}^* Q_k \\
&= R_k (Q_{k-1}^* Q_k)
\end{aligned}
$$

- QR iteration

  Given $Q_0 \in \mathbb{C}^{n\times n}$, $T_0 = Q_0^* A Q$

  For $k = 1, 2, 3, \cdots$

  Compute QR fact. $T_{k-1} = Q_k R_k$ $\leftarrow$ $\frac{4}{3} n^3$ (complex) flops

  $T_{k+1} = R_k Q_k$ $\leftarrow$ n^3 (complex) flops

  Cost? Convergence?

Today: Practical QR iteration

Idea: Reduce A to simpler form before applying QR

- Two phase algorithm:

<u>Step 1</u> : Transform A to an upper Hessenberg matrix.

i.e. $h_{ij} = 0$ for $i > j+1$

$$H = \begin{bmatrix} x & x & x & \cdots & x \\ x & x & x & \cdots & x \\ & x & x & \cdots & x \\ & & \ddots & \ddots & \vdots \\ & & & x & x \end{bmatrix} \in \mathbb{C}^{n\times n}$$

This is possible by applying unitary transform to A

i.e. find $U \in \mathbb{C}^{n\times n}$, unitary, $U^*AU = H$

unchanged

$$\underset{A}{\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix}} \xrightarrow[\begin{bmatrix} 1 & \\ & V_1 \end{bmatrix}]{U_1^* \cdot} \underset{U_1^*A}{\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{bmatrix}} \xrightarrow[\begin{bmatrix} 1 & \\ & V_1^* \end{bmatrix}]{\cdot U_1} \underset{U_1^*AU_1}{\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{bmatrix}}$$

Householder transform

unchanged

$$\xrightarrow[\begin{bmatrix} 1 & & \\ & 1 & \\ & & V_2 \end{bmatrix}]{U_2^* \cdot} \underset{U_2^*U_1^*AU_1}{\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & x & x & x \end{bmatrix}} \xrightarrow[\begin{bmatrix} 1 & & \\ & 1 & \\ & & V_2^* \end{bmatrix}]{\cdot U_2} \underset{U_2^*U_1^*AU_1U_2}{\begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & x & x & x \end{bmatrix}} \Rightarrow \cdots$$

unchanged

$$H = \underbrace{U_{n-2}^* \cdots U_1^*}_{U^*} A \underbrace{U_1 \cdots U_{n-2}}_{U}$$

— If A is non-Hermitian, A is reduced to upper Hessenberg

A is Hermitian, A is reduced to tridiagonal

— Total cost : $\approx 2 \times \frac{4}{3} n^3$ (complex) flop for general A

$\approx \frac{4}{3} n^3$ (complex) flop for Hermitian A
(it suffices to work on lower triangular part)

Step 2: Hessenberg QR iteration

QR iteration with H is significanty faster than A!

- Compute $H = QR$

  We only need to zero out one subdiagonal entry each time

$$
H = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}
\xrightarrow[\begin{bmatrix} F_1 & \\ & I_{n-2} \end{bmatrix}]{Q_1^T\cdot}
\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}}_{Q_1^T H}
\xrightarrow[\begin{bmatrix} 1 & & \\ & F_2 & \\ & & I_{n-3} \end{bmatrix}]{Q_2^T\cdot}
\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ & 0 & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}}_{Q_2^T Q_1^T H}
\Rightarrow \cdots \quad R = Q_{n-1}^T \cdots Q_1^T H
$$

$F_1, F_2, \cdots, F_{n-1}$ are 2×2 Householder transform

Note that Q_k^T only changes k^{th} and $k+1^{th}$ row

For general A (H is upper Hessenberg), (cost of $H=QR$) $\approx 3n^2$ flops (complex)

For Hermitian A (H is tri diagonal), (cost of $H=QR$) $\approx 6n$ flops (complex)

- Compute RQ

  Apply Q in the last step to R get RQ, but are the nice properties (upper Hessenberg / tridiagonal) preserved?

$$
R = \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}
\xrightarrow[\begin{bmatrix} F_1^T & \\ & I_{n-1} \end{bmatrix}]{\cdot Q_1}
\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}
\xrightarrow[\begin{bmatrix} 1 & & \\ & F_2^T & \\ & & I_{n-3} \end{bmatrix}]{\cdot Q_2}
\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}
\Rightarrow \cdots \quad RQ_1 \cdots Q_{n-1} = H
$$

Note that Q_k only changes k^{th} and $k+1^{th}$ columns.

When A is general (Hermitian), RQ is still upper Hessenberg (tridiagonal)!

Cost of RQ step: for general A, $\approx 3n^2$ (complex) flops

Hermitian A, $\approx 6n$ (complex) flops

Total cost of step 2 $\approx$ # iterations $\cdot\, 6n^2$ flops ($\begin{smallmatrix}\text{general}\\ A\end{smallmatrix}$)

$\cdots\cdots\cdots\cdot\, 12n$ flops ($\begin{smallmatrix}\text{Hermitian}\\ A\end{smallmatrix}$)

Note: By choosing suitable shifting, convergence of QR iteration is usually very fast (in a few iterations), so the dominate cost comes from step 1 $\approx O(n^3)$

Note: Apply QR iter. to a Hessenberg matrix, the p^{th} subdiagonal entry in H converges to zero with linear rate $\frac{|\lambda_{p+1}|}{|\lambda_p|}$

- Deflation

In practice, when a subdiagonal entry in H is sufficiently small,

for example, $h_{p+1,p} \leq C\, \varepsilon_{mach} \left(|h_{pp}| + |h_{p+1,p+1}| \right)$

we can justifiably set it to be zero, (b.c. comparable to rounding error)

In this case, we can decouple the problem into two small problems:

suppose we have at some step,

$$H = \begin{array}{c} \left[\begin{array}{cc} H_{11} & H_{12} \\ 0 & H_{22} \end{array} \right] \begin{array}{l} p \\ n-p \end{array} \\ \begin{array}{cc} p \quad & \quad n-p \end{array} \end{array}, \qquad 1 \leq p < n$$

if we are able to find Schur form: $H_{11} = Q_1 T_{11} Q_1^*$

$H_{22} = Q_2 T_{22} Q_2^*$

then

$$H = \begin{bmatrix} Q_1 & \\ & Q_2 \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ & T_{22} \end{bmatrix} \begin{bmatrix} Q_1^* & \\ & Q_2^* \end{bmatrix} \quad \text{with } T_{12} := Q_1^* H_{12} Q_2$$

is the Schur fact. of H

If this happens when $p = n-1$ or $n-2$ we call it <u>deflation.</u>

- How to incorperate shifting in QR ?

If apply QR directly to $A - \mu I$. i.e.

Given $Q_0 \in \mathbb{C}^{n \times n}$ unitary, Let $T_0 := Q_0^*(A - \mu I) Q_0$

For $k = 1, 2, 3, \dots$

Compute QR fact. $T_{k-1} = Q_k R_k$

$T_k = R_k Q_k$

Problem: How to change μ during iter. ?

Instead, let $H_k := T_k + \mu I$.

<u>Shifted QR Iteration</u>:

Given $Q_0 \in \mathbb{C}^{n \times n}$ unitary. Let $H_0 := Q_0^* A Q_0$

For $k = 1, 2, 3, \dots$

Determine a scalar $\mu_k \in \mathbb{C}$

Compute QR fact. $H_{k-1} - \mu_k I = Q_k R_k$

$H_k = R_k Q_k + \mu_k I$

If we order eig. val. of A so that

$$|\lambda_1 - \mu_k| \geq \cdots \geq |\lambda_n - \mu_k|$$

then the p^{th} subdiagonal entry in H converges to zero with rate $\dfrac{|\lambda_{p+1} - \mu_k|}{|\lambda_p - \mu_k|}$

In the extreme case, when μ is an eig. val of A, we get the exact eig. val. in a single step.

<u>Thm</u> Let μ be an eig. val. of a Hessenberg matrix H with all $h_{i+1,i} \neq 0$, $i = 1, 2, \cdots, n-1$ ← called "unreduced" when this is not true, we can decouple the problem

Then after a single shifted QR step, we have $h_{n,n-1} = 0$ and $h_{nn} = \mu$

Pf: Since $H - \mu I = QR$ and $H - \mu I$ is singular, we know $r_{11} \cdots r_{nn} = 0$.

Since H is unreduced, the first $n-1$ columns of H is linearly independent. thus $r_{ii} \neq 0$ $\forall i = 1, \cdots, n-1$

$\Rightarrow r_{nn} = 0$

$$\Rightarrow \mu I + RQ = \begin{bmatrix} * & * \\ 0 & \mu \end{bmatrix} \quad \blacksquare$$

- How to choose μ_k?

For simplicity, we assume A is Hermitian.

<u>Option 1</u>: Rayleigh quotient shift

Recall $Q_k^* H Q_k = T_k$

$$Q_k = [\, q_k^{(1)} \cdots q_k^{(n)} \,]$$

$$\mu_k := R(q_k^{(n)}) = q_k^{(n)*} A q_k^{(n)} = (T_k)_{n,n}$$

Rayleigh quotient shift gives (local) cubic convergence in generic case (global convergence and cubic local convergence can fail in some corner cases)

ex. $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$\mu = 0$, QR fact. $A = QR = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$$RQ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = A$$

<u>Option 2</u>: Wilkinson's shift

If
$$H = \begin{bmatrix} a_1 & b_1 & \cdots & 0 \\ b_1 & a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & b_{n-1} \\ 0 & \cdots & b_{n-1} & a_n \end{bmatrix}$$

Wilkinson's shift choose μ to be eigenvalue of

$$T(n-1:n,\ n-1:n) = \begin{bmatrix} a_{n-1} & b_{n-1} \\ b_{n-1} & a_n \end{bmatrix}$$

and set μ to be the eig. val. of $T(n-1:n, n-1:n)$ that is closer to a_n.

$$\mu = a_n + d - \operatorname{sign}(d)\sqrt{d^2 + b_{n-1}^2}$$

$$= a_n - \frac{\operatorname{sign}(d)\, b_{n-1}^2}{|d| + \sqrt{d^2 + b_{n-1}^2}} \quad \text{(numericall stable)}$$

with $d = (a_{n-1} - a_n)/2$

It is gauranteed that Wilkinson's shift always converges (globally) and at least locally quadrotically, and almost always cubically.

<u>Option 3</u>: (bulge chasing) choose μ based on largest off-diagonal entry of H ......

Others : If stalled, perturb the shift to break the cycle ...

Other problems:

1) Perform QR iter. in real number ?

Schur fact. is not true in real number

Complex eig. val. how to choose real shifts ?

Double - Implicit - Shift Strategy (Francis QR)

two sucessive shifts

combine QR fact. and RQ in a single step (good when $\mu >> a_{ii}$ for some i)

2) Compute selected eig. vectors ?

3) A lot others ...