Welcome to Ma221! Lecture 27, Oct 25

Power Method $x_{i+1} = Ax_i / \| " \|_2$

Inverse Iteration $x_{i+1} = (A-\sigma I)^{-1} x_i / \| " \|_2$

Orthogonal Iteration

$Z_i^{n \times p}$ $Y_{i+1} = A Z_i$, $Z_{i+1} \cdot R_{i+1} = QR(Y_{i+1})$

Thm: Orthog Iter. with $Z_0 = I^{n \times n}$

$|\lambda_1| > |\lambda_2| > \cdots \Rightarrow$

$A_i = Z_i^T A Z_i \rightarrow$ Schur form

Next: QR iteration (let's us introduce inverse iter.)

$A_0 = A$

$i = 0$

repeat

$A_i = Q_i R_i$ ... QR decomp

$A_{i+1} = R_i Q_i$

$i = i+1$

until convergence

Thm: A_i from QR iteration same as $Z_i^T A Z_i$ from orthog. iter

proof (induction) assume $A_i = Z_i^T A Z_i$

One step of Orthog iter

$A Z_i = Z_{i+1} R_{i+1}$ ... QR decomp

$A_i = Z_i^T A Z_i$ ... induct. hypoth.

$$= \underbrace{Z_i^T Z_{i+1}}_{\text{orthog}} \underbrace{R_{i+1}}_{\triangle} \quad \ldots \text{ from orthog. iter.}$$

$= QR$ decomp of A_i, by uniqueness

$$Z_{i+1}^T A Z_{i+1} = (Z_{i+1}^T A Z_i)(Z_i^T Z_{i+1})$$
$$= (R_{i+1})(Z_i^T Z_{i+1})$$
$$= RQ$$
$$= A_{i+1}$$

Add inverse iteration to QR iteration

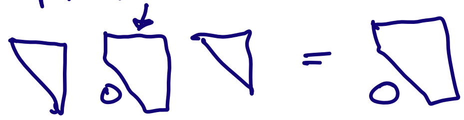
$A_0 = A$
$i = 0$
repeat
  factor $A_i - \sigma I = Q_i R_i$
  $A_{i+1} = R_i Q_i + \sigma I$
  $i = i+1$
until convergence

Lemma: A_i and A_{i+1} are orthog. similar

proof: $A_{i+1} = R_i Q_i + \sigma_i I$
$$= Q_i^T (Q_i R_i) Q_i + \sigma I$$
$$= Q_i^T (A_i - \sigma I) Q_i + \sigma I$$
$$= Q_i^T A_i Q_i \underbrace{- \sigma Q_i^T Q_i + \sigma I}_{\text{cancel}}$$
$$= Q_i^T A_i Q_i$$

Note: if R_i nonsingular, can also write

$$A_{i+1} = R_i Q_i + \sigma_i I$$
$$= R_i \underbrace{Q_i R_i} R_i^{-1} + \sigma_i I$$
$$= R_i (A_i - \sigma_i I) R_i^{-1} + \sigma_i I$$
$$= R_i A_i R_i^{-1} \underbrace{- \sigma_i R_i R_i^{-1} + \sigma_i I}_{\text{cancel}}$$
$$= R_i A_i R_i^{-1}$$



upper Hessenberg

If σ_i = exact eval of A

QR converges in one step

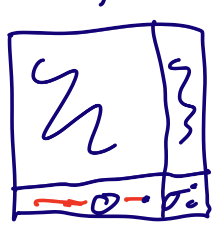
$A_i - \sigma_i I$ singular $\Rightarrow$ R_i singular

$\Rightarrow$ $R_i(k,k) = 0$, suppose $R_i(n,n) = 0$

$\Rightarrow$ last row of $R_i = 0$

$\Rightarrow$ last row of $R_i Q_i = 0$

$\Rightarrow$ last row of $A_{i+1} = R_i Q_i + \sigma_i I$ is zero, except for $A_{i+1}(n,n) = \sigma_i$



Q4.1

If σ_i not exact eval, but "close enough", i.e. $\|A_{i+1}(n,1:n-1)\|_2 = O(\varepsilon)$

set $A_{i+1}(n,1:n-1)=0$

backward stable

Previous analysis: expect $A_{i+1}(n,1:n-1)$ to shrink by factor

$$\frac{|\lambda_k-\sigma_i|}{\min_{j\neq k}|\lambda_j-\sigma_i|}$$

λ_k is closest eval to σ_i

$A_i-\sigma_i I = Q_i R_i$

$Q_i^T(A_i-\sigma_i I)=R_i$

if σ_i were $=\lambda_k \Rightarrow R_i(n,n)=0$

$\Rightarrow$ last row of $Q_i^T(A_i-\sigma_i I)=0$

$\Rightarrow$ last col of Q_i is left evec of A_i for σ_i

Now suppose σ_i just close to eval

$A_i-\sigma I = Q_i R_i$

$(A_i-\sigma_i I)^{-1} = R_i^{-1}Q_i^T$

$(A_i-\sigma_i I)^{-T} = Q_i R_i^{-T}$

$(A_i-\sigma_i I)^{-T}R_i^T = Q_i$

$\triangle$

$$(A_i - \sigma_i I)^{-T} e_n R_i(n,n) = Q_i e_n$$

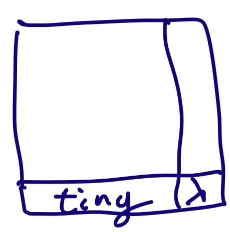
$\Rightarrow$ one step of inverse iteration starting with e_n, expect last col of Q_i to be approx evec

$\Rightarrow$ last col of Q_i closer to evec of A_i^T

$\Rightarrow$ last col of $A_i^T Q_i$ closer to $\lambda \cdot$ last col of Q_i

$\Rightarrow$ last col of $Q_i^T A_i^T Q_i$ closer to $\lambda \cdot e_n$

$\Rightarrow$ last row of $Q_i^T A_i Q_i$ closer to λe_n^T

$\Rightarrow$ 

where do we get σ_i?

use $\sigma_i = A_i(n,n)$

$\Rightarrow$ yields quadratic convergence

$$\|A_i(n,1:n-1)\|_2 = \varepsilon \ll 1$$

$$|A_i(n,n) - \lambda|_2 = O(\varepsilon) \leftarrow$$

expect $\|A_i(n,1:n-1)\|_2$ gets multiplied by $|A_i(n,n)-\lambda| = O(\varepsilon)$

$\Rightarrow$ becomes $O(\varepsilon^2)$

Matlab demo: see typed notes for code

Making QR iteration practical

1) Each iteration costs 1 QR + 1 matmul $= O(n^3)$, if we did only $O(1)$ iterations per eval $\Rightarrow O(n^4)$ cost total, want $O(n^3)$

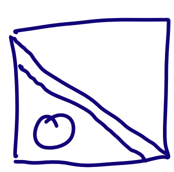
2) How to shift to converge to real Schur form?

3) How do decide on convergence?

4) How to minimize communication?

Answers:

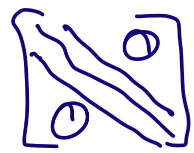
(1) preprocess $A = QHQ^T$ where Q orthog, H Upper Hessenberg



QR iteration preserves 0s, lowers cost to $O(n^2)$

$\Rightarrow$ cost $= O(n^3)$

$A = A^T \Rightarrow H = H^T \qquad H^T = (Q^T A Q)^T$

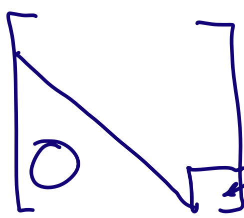
$\Rightarrow$ H tridiagonal 

$\Rightarrow$ cost $= O(n)$ (chap 5)

(2) Converge to Real Schur Form:

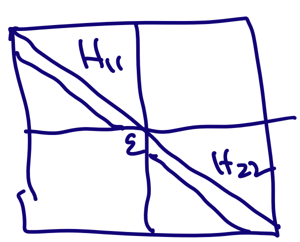
since evals appear in $\lambda, \bar{\lambda}$ pairs, turns out that 1 step of QR iteration with shift $= \lambda$, followed by $\bar{\lambda}$ as shift $\Rightarrow A_{i+2}$ real

$\Rightarrow$ don't compute any imaginary parts



$\lambda, \bar{\lambda}$ = eigenvalues of $H(n-1:n, n-1:n)$

(3) Detect convergence?



if $\varepsilon \le$ macheps $\cdot \|A\|$
set it to 0

work independently on H_{11}, H_{22}