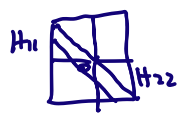
Welcome back to Ma221! Lecture 28, Oct 27

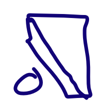
Making QR Iteration Practical

1) Reduce cost of 1 iteration from $O(n^3) \to O(n^2)$
2) How to converge to Real Schur form:
   shift by $\lambda, \bar{\lambda}$ in one step
   all imaginary parts cancel
3) Detect convergence:
   set any $|H(i+1,i)| \le O(\varepsilon) \cdot \|A\|$ to 0

   solve H_{11}, H_{22} independently
4) Reduce communication
   no known alg to reach lower bound
   $O\left(\frac{n^3}{\sqrt{\text{cache size}}}\right)$ that is deterministic,
   just randomized, Srivastava (2022)

More detail on Hessenberg QR



How to reduce $A = QHQ^T$, H Hessenberg

Analogous to QR decomposition,
use Householder transforms to
zero out columns of A

what if $A = A^T \rightarrow$ H tridiagonal

$$\begin{bmatrix} I_3 & \\ & H^{2\times2}\end{bmatrix}\begin{bmatrix} I_2 & \\ & H^{3\times3}\end{bmatrix}\begin{bmatrix} 1 & \\ & H^{4\times4}\end{bmatrix}\begin{bmatrix} x&x&x&x&x\\ x&x&x&x&x\\ x&x&x&x&x\\ x&x&x&x&x\\ x&x&x&x&x\end{bmatrix}\begin{bmatrix}1&0\\0&H^{4\times4}\end{bmatrix}^T\begin{bmatrix} I_2 & \\ & H^{3\times3}\end{bmatrix}^T\begin{bmatrix} I_3 & \\ & H^{2\times2}\end{bmatrix}^T$$

upper Hessenberg, can use BLAS3

Cost: $\frac{16}{3}n^3 + O(n^2)$ just for H

or $\frac{14}{3}n^3 + O(n^2)$ for Q too

much more than LU or QR, already

(next step more expensive)

SVD similar

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

B

$QAV = B$

bidiagonal

QR iteration on upper Hessenberg H in $O(n^2)$ flops

Lemma: upper Hessenberg preserved by QR iteration

pf: A upper Hess $\rightarrow$ $A - \sigma I$ upper Hess

$\Rightarrow A - \sigma I = QR$, Q upper Hess, because

col i of Q = linear comb of first i cols of $A - \sigma I$

$\Rightarrow$ 0's at bottom, rows $i+2:n$, stay 0

RQ = ◸ ◹ still upper Hess

How to do one step of QR iteration in $O(n^2)$ flops

Def: Assume H "unreduced" i.e. $H(i+1,i) \neq 0$ for all i

Implicit Q Theorem: suppose $Q^T A Q$ upper Hess and unreduced, then cols $2:n$ of Q uniquely determined by col 1 (proof later)

Step 1 of QR in $O(n^2)$ flops

$$\underset{\text{upper Hess}}{A - \sigma I} = QR \quad \text{first col of } A - \sigma I = \begin{bmatrix} A(1,1)-\sigma \\ A(2,1) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\Rightarrow$ know first col of Q

Let Q_1 be Givens rotation s.t

$$\underbrace{\begin{bmatrix} c & s & & & \\ -s & c & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}}_{Q_1^T} \begin{bmatrix} A(1,1)-\sigma \\ A(2,1) \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} x \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$Q_4^T \; Q_3^T \; Q_2^T \; Q_1^T \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ \oplus & x & x & x & x \\ & \oplus & x & x & x \\ & & \oplus & x & x \end{bmatrix} \overbrace{\underbrace{Q_1 \, Q_2 \, Q_3 \, Q_4}_{Q}}^{\text{first col of } Q \text{ same as } Q_1}$$

cost $= O(n^2)$

"bulge chasing"

proof of Implicit Q Theorem

let q_i be col i of Q

$$Q^T A Q = H \Rightarrow AQ = QH$$

Col 1: $(AQ)_1 = A q_1 = q_1 \cdot H(1,1) + q_2 \, H(2,1)$

$\Rightarrow$ determines q_2, $H(1,1)$ and $H(2,1)$ uniquely

via QR $[q_1, A q_1] = [q_1, q_2] \cdot \begin{bmatrix} 1 & H(1,1) \\ 0 & H(2,1) \end{bmatrix}$

Induction on column number i:

suppose we have $q_1, \ldots, q_i$

and columns $1, \ldots, i-1$ of H

Get next columns: $(AQ)_i = (QH)_i$

$$q_j^T \left(A q_i = \sum_{j=1}^{i+1} q_j \, H(j,i) \right)$$

$$q_j^T A q_i = H(j,i) \qquad \text{for } j = 1:i$$

$$A q_i - \sum_{j=1}^{i} q_j \, H(j,i) = q_{i+1} \cdot H(i+1,i)$$

gives us q_{i+1}, $H(i+1,i)$ QED

Used: in LAPACK xGEES (Schur)

xGEEV (evecs too)

eig(A) Matlab

Symmetric Eigenproblem + SVD

Goals: Perturbation Theory

Algorithms

need Perturbation Theory to understand Algs too

Real Symmetric $A = A^T$ ← do this case

Complex Hermitian $A = A^H$

$$A = Q \Lambda Q^T$$

$\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$

$Q = [q_1, \dots, q_n]$

Complex symmetric different

$\begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$ $i = \sqrt{-1}$ has 2 evals at 0, not diagonalizable

Most results apply to SVD (Thm 3.3 part 4), want svd(A)

$$B = \left[\begin{array}{c|c} 0 & A \\ \hline A^T & 0 \end{array}\right] = B^T$$

eigendecomp of B closely related to svd(A) $\Rightarrow$ algs, perturbation theory for sym eig extend to SVD some open problems extending fastest symeig algs to SVD

Def: Rayleigh Quotient $\rho(u,A) = \frac{u^T A u}{u^T u}$, $u \neq 0$

Properties: if $Au = \lambda u \Rightarrow \rho(u,A) = \lambda$

$u = \sum_{i=1}^{n} b_i q_i = Q b \qquad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$

$$\rho(u,A) = \frac{u^T A u}{u^T u} = \frac{(Qb)^T A (Qb)}{(Qb)^T (Qb)}$$

$$= \frac{b^T \overbrace{Q^T A Q}^{} b}{b^T \underbrace{Q^T Q}_{I} b} = \frac{b^T \Lambda b}{b^T b} \qquad A = Q \Lambda Q^T$$

$$= \frac{\sum_{i=1}^{n} \lambda_i b_i^2}{\sum_{i=1}^{n} b_i^2} = \text{convex combination of all evals}$$

$$\lambda_1 \geq \rho(u,A) \geq \lambda_n$$

all evals can be expressed using $\rho(u,A)$

Courant-Fischer Minimax Thm

R^j = j-dimensional subspace of $\mathbb{R}^n$

S^{n-j+1} = $n-j+1$ " " " "

$$\max_{R^j} \min_{\substack{r \in R^j \\ r \neq 0}} \rho(r,A) = \lambda_j = \min_{S^{n-j+1}} \max_{\substack{s \in S^{n-j+1} \\ s \neq 0}} \rho(s,A)$$

max over R^j attained by span$(q_1, \dots, q_j)$

min over $r \in R^j$ " " $r = q_j$

min over S^{n-j+1} " " span$(q_j, q_{j+1}, \dots q_n)$

max over $s \in S^{n-j+1}$ " " $s = q_j$