Welcome back to Ma 221! Lecture 25, Oct 20

Recall Invariant Subspaces

$V = \text{span}\{X\}$ $\quad X = [x_1, \dots, x_m]$

$AV \subseteq V$, then V invariant

If V invariant $\exists B$ $\quad AX = XB^{m \times m}$

all evals of B are evals of A

$X = QR$, let $[Q, Q']$ be orthogonal

$$[Q, Q']^T A [Q, Q'] = \begin{array}{cc} & m \\ m & \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] \end{array}$$

$A_{11} = RBR^{-1}$

recursively apply to A_{22}

Use this to get real Schur Form

one block per real λ, or complex conjugate pair $Ax = \lambda x$, $A\bar{x} = \bar{\lambda}\bar{x}$

$$X = [\text{re}(x), \text{im}(x)]$$

$$\text{Re}(Ax) = \text{Re}(\lambda x) \quad \text{and} \quad \text{Im}(Ax) = \text{Im}(\lambda x)$$

$$AX = X \cdot \begin{bmatrix} \text{re}(\lambda) & \text{im}(\lambda) \\ -\text{im}(\lambda) & \text{re}(\lambda) \end{bmatrix} = X \cdot B$$

$\text{evals}(B) = \{\lambda, \bar{\lambda}\}$

$\Rightarrow$ real Schur Form exists

Recall other eigenproblems

(1) ODE $x'(t) = K\,x(t)$

if $K\,x(0) = \lambda\,x(0) \Rightarrow x(t) = e^{\lambda t}x(0)$

similar if $x(0)$ = linear comb of evecs

(2) $M x''(t) + K\,x(t) = 0$

$\Rightarrow \lambda^2 M x(0) + K x(0) = 0$

"generalized eigenproblem"

$\det(\lambda' M + K) = 0$ where $\lambda' = \lambda^2$

(3) $M x''(t) + D\,x'(t) + K x(t) = 0$

$\Rightarrow$ "non linear eigenproblem"

$\lambda^2 M x(0) + \lambda D\,x(0) + K x(0) = 0$

reduce to linear eigenproblem

(2 matrices, 2x larger)

(4) $x'(t) = A \cdot x(t) + B \cdot u(t)$

"linear control system"

turns into "singular eigenproblem"

$n\left[\overset{n}{B}, \overset{m}{A}\right]$ and $n\left[\overset{n}{I}, \overset{m}{O}\right]$

All ideas of Chap 4 (Jordan Form, Schur form, perturbation theory, algorithms) extend to all these cases (Chap 4.5)

Concentrate on one nonsymmetric A

Perturbation Theory: Can I trust my answer?

Last time $A = I$ showed eigenvectors very ill conditioned, but eigenvalues "perfectly" conditioned

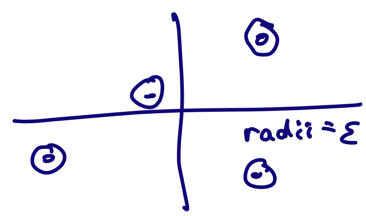
To describe how evals can be perturbed:

Def: Epsilon-pseudo-spectrum of A
= set of all evals of all matrices within distance ε of A

$$\Lambda_\varepsilon(A) = \{\lambda : (A+E)x = \lambda x \text{ for some } x \neq 0,\ \|E\|_2 \le \varepsilon\}$$

Smallest possible $\Lambda_\varepsilon(A)$: disks of radius ε around each eval of A

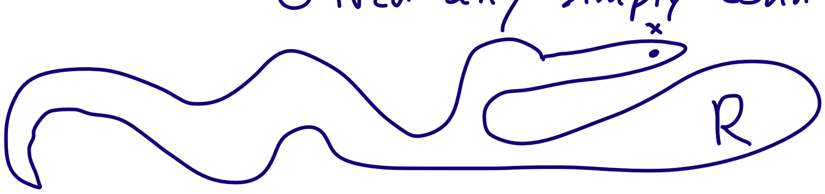
attained by $E = xI$
$|x| \le \varepsilon$



Worst case (most sensitive)

Thm (Trefethen + Reichel)

Given any simply connected subset of $\mathbb{C}$
$R \subseteq \mathbb{C}$
(no holes)



Given any $\varepsilon > 0$, given any $x \in \mathbb{R}$

$\exists A$: $\Lambda_0(A) = \{x\}$, $\Lambda_\varepsilon(A)$ fills out inside of $\mathbb{R}$

proof: Ma185 (Riemann Mapping Thm)

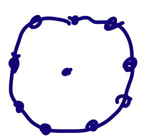
Ex: Perturb $n \times n$ Jordan Block, $\lambda = 0$ with $J(n,1) = \varepsilon$

$$\begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \varepsilon & & & 0 \end{bmatrix}$$

$p(\lambda) = \lambda^n - \varepsilon$ = characteristic polynomial

$\Rightarrow \lambda = \sqrt[n]{\varepsilon}$

uniformly spaced evals on circle of radius $\sqrt[n]{\varepsilon}$



$\varepsilon = 10^{-16}$ $\quad n = 16$

$\sqrt[n]{\varepsilon} = .1$

(1) evals are continuous functions of A
not necessarily differentiable
(slope of $\sqrt[n]{\varepsilon} = \infty$ at $\varepsilon = 0$)

(2) expect sensitive evals when evals nearly multiple

Condition number of simple (nonmultiple) evals (else ∞)

Thm: λ simple eigenvalue of A
$Ax = \lambda x$, $y^H A = y^H \lambda$, $\|x\|_2 = \|y\|_2 = 1$

If we perturb A to $A+E$

λ perturbed to $\lambda + \delta\lambda$

$$\delta\lambda = \frac{y^H E x}{y^H x} + O(\|E\|^2)$$

$$|\delta\lambda| \le \frac{\|E\|_2}{|y^H x|} + O(\|E\|^2)$$

$$= \sec\theta \cdot \|E\|_2 + O(\|E\|^2)$$

$$\theta = \angle(x,y)$$

$\sec\theta$ = condition number

proof: $(A+E)(x+\delta x) = (\lambda+\delta\lambda)(x+\delta x)$

$\underline{A\cdot x} + A\delta x + E x + \underline{E\delta x} = \underline{\lambda\cdot x} + \lambda\delta x + \delta\lambda\cdot x + \underline{\delta\lambda\cdot\delta x}$

cancel

second order, ignore

$y^H(A\cdot\delta x + E x = \lambda\cdot\delta x + \delta\lambda\cdot x)$

$\underline{y^H A}\delta x + y^H E x = \underline{y^H \lambda}\delta x + y^H\cdot\delta\lambda\, x$

cancel

$$y^H E x = \delta\lambda\cdot y^H x$$

$$\frac{y^H E x}{y^H x} = \delta\lambda \qquad \text{QED}$$

Special case 1: $A = A^H$ or "normal": $A\cdot A^H = A^H\cdot A$

$\Rightarrow$ A has orthonormal evecs (HW Q4.2)

Cor: If A normal, perturbing A to $A+E$

$\Rightarrow$ $|\delta\lambda| \le \|E\|_2 + O(\|E\|_2^2)$ i.e. cond# $= 1$

proof: $A = Q \Lambda Q^T$ cols of Q = evecs are orthonormal

$\Rightarrow AQ = Q\Lambda$

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

$\Rightarrow$ right evecs = cols of Q = rows of Q^T

= left evecs

$\Rightarrow x = y \quad y^H x = \|x\|^2 = 1$

Later: (Chap 5) if $A = A^H$ and $E = E^H$

then $|\delta\lambda| \le \|E\|_2$ i.e $O(\|E\|_2^2)$ disappears

Special Case 2: A = Jordan Block

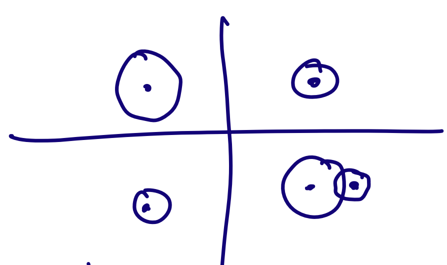
$$A = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} \qquad x = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad y = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$\Rightarrow y^H x = 0 \Rightarrow$ cond # $= \infty$

Thm: (Bauer-Fike Thm) If A has all simple λ_i with right and left evecs x_i and y_i, $\|x_i\|_2 = \|y_i\|_2 = 1$

Then for any E: the evals of $A+E$ lie in disks in $\mathbb{C}$ centered at λ_i with radii $\dfrac{n \cdot \|E\|_2}{|y_i^H x_i|}$

$n = \dim(A)$



if k disks overlap, k evals must lie in their union   (proof in book)

Algorithms for Nonsymmetric Eigenproblem

Ultimate Algorithm:

Hessenberg QR (HQR)

takes any $n \times n$ (nonsymmetric) dense A

computes Schur form $A = Q^* T Q$

in $O(n^3)$ flops

Build it via a sequence of simpler algs, also used in practice, eg. to find just a few evals and evecs see Chap 7

Plan: Power Method: Just repeated multiplication of x by A

converges to evec for largest eval in absolute value

Inverse Iteration: Apply power method to $B = (A - \sigma I)^{-1}$ which has same evecs as A and largest eval of B corresponds to eval of A closest to σ

$\Rightarrow$ by choosing "shift" σ carefully, converge to any (evec, eval)

Orthogonal Iteration: Extends power method from one evec to an invariant subspace

QR iteration: combine orthog iteration
and inverse iteration

Other techniques
to get to $O(n^3)$
real Schur form
use BLAS3 / minimize comm.
(discuss some of these)