Welcome to Ma221! Lec 16, Fall 2024

Thm (Schur Form)

Given an $n \times n$ A, $\exists$ unitary Q
$QQ^H = I$ s.t. $Q^H A Q = T =$ upper triangular

with evals $T(i,i)$, can appear in any order

proof: induction

choose any λ, corresponding x: $Ax = \lambda x$
$\|x\|_2 = 1$

Let $Q = [x, Q']$ be unitary

$$Q^H A Q = \begin{bmatrix} x^H \\ Q'^H \end{bmatrix} A \begin{bmatrix} x, & Q' \end{bmatrix}$$

$$= \begin{array}{c c} & \begin{array}{cc} 1 & n-1 \end{array} \\ \begin{array}{c} 1 \\ n-1 \end{array} & \left[\begin{array}{c|c} x^H A x & x^H A Q' \\ \hline Q'^H A x & Q'^H A Q' \end{array}\right] \end{array}$$

$$= \left[\begin{array}{c|c} x^H \lambda x & x^H A Q' \\ \hline \lambda Q'^H x & Q'^H A Q' \end{array}\right]$$

$$= \left[\begin{array}{c|c} \lambda & x^H A Q' \\ \hline 0 & Q'^H A Q' \end{array}\right]$$

apply induction to $Q'^H A Q' = U^H T U$

$$Q'^H A Q' = \underset{\text{unitary}}{U^H}\ \underset{\text{triangular}}{T}\ U$$

$$Q^H A Q = \left[\begin{array}{c|c} \lambda & x^H A Q' \\ \hline 0 & U^H T U \end{array}\right] = \left[\begin{array}{c|c} 1 & \\ \hline & U^H \end{array}\right] \left[\begin{array}{c|c} \lambda & x^H A Q' U^H \\ \hline 0 & T \end{array}\right] \cdot \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & U \end{array}\right]$$

$$\Rightarrow \underbrace{\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & U \end{array}\right] Q^H}_{\text{unitary}} A \underbrace{Q \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & U^H \end{array}\right]}_{\substack{\text{unitary} \\ \text{(inverse)}}} = \underbrace{\left[\begin{array}{c|c} \lambda & x^H A Q' U^H \\ \hline 0 & T \end{array}\right]}_{\text{upper triangular}}$$

Schur form! $\quad Q \in \mathbb{C}$

What about real matrices with complex evals?

Special case: $A = A^T$, all evals real
chap 5

Prefer real arithmetic for real A

- reduce # flops
- less memory
- make sure evals, evecs appear in complex conjugate pairs

$$\underset{\text{real}}{A}\ \underset{\text{complex}}{x} = \underset{\text{complex}}{\lambda}\ x \iff A\bar{x} = \bar{\lambda}\bar{x}$$

Instead of T = upper triangular
use block triangular

$$T = \begin{bmatrix} T_{11} & & T_{ij} \\ & T_{22} & \\ 0 & & \ddots \; T_{kk} \end{bmatrix}$$

evals of T are $\bigcup_i$ evals(T_{ii}) (HW 4.1)

each T_{ii} either 1×1 and = real eval
or 2×2 with evals $\lambda, \bar{\lambda}$

Thm (Real Schur Canonical Form)
Given any real A, $\exists$ real orthogonal Q
s.t. QAQ^T is block upper triangular
with 1×1 and 2×2 diagonal blocks

Generalize to "invariant subspaces"

Def: $V = \text{span}\{x_1, \ldots, x_m\} = \text{span}(X)$, $X = [x_1, \ldots, x_m]$
is a subspace of $\mathbb{R}^n$. V invariant
if $A \cdot V = \text{span}(AX) \subseteq V$

Ex: $V = \text{span}\{x\} = \{\alpha x, \text{ for all scalars } \alpha\}$
where $Ax = \lambda x$
$AV = \{A(\alpha x), \forall \alpha\} = \{\alpha \lambda x, \forall \alpha\}$
$\subseteq V$, $= V$ unless $\lambda = 0$

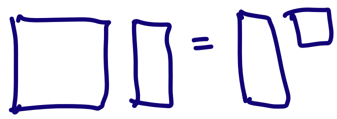
Ex: $V = \text{span}\{x_1, \ldots, x_k\} = \{\sum_{i=1}^{k} \alpha_i x_i, \ \forall \alpha_i\}$

where $Ax_i = \lambda_i x_i$

$$AV = \{ A(\sum_i \alpha_i x_i) \quad \forall \alpha_i \}$$
$$= \{ \sum_i (\alpha_i A x_i) \quad \forall \alpha_i \}$$
$$= \{ \sum_i (\alpha_i \lambda_i) x_i \quad \forall \alpha_i \} \subseteq V$$

Lemma: Let $V = \text{span}\{x_1, \dots, x_m\} = \text{span}(X)$ be invariant subspace of A

Then $\exists B: \ AX = XB$



and $\text{evals}(B) \subseteq \text{evals}(A)$

proof: existence of B follows from def:

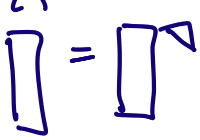
$Ax_i \in V \Rightarrow \exists$ scalars $B(1,i), B(2,i) \cdots B(m,i)$

s.t. $Ax_i = \sum_{j=1}^{m} x_j B(j,i)$ i.e. $AX = XB$

$$By = \lambda y \Rightarrow A(Xy) = XBy$$
$$\Rightarrow A(Xy) = X(\lambda y)$$
$$\Rightarrow A(Xy) = \lambda(Xy)$$

Lemma: $V = \text{span}(X)$ be m-dimensional invariant subspace of A, so $AX = XB$

$X = QR$ Let $\begin{array}{cc} m & n-m \\ \end{array}$ $[Q, Q']$ be square, orthogonal



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

$A_{11} = RBR^{-1}$ has same evals as B

proof: $$[Q, Q']^T A [Q, Q'] = \left[\begin{array}{c|c} Q^T A Q & Q^T A Q' \\ \hline Q'^T A Q & Q'^T A Q' \end{array}\right]$$

$$= \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{c|c} RBR^{-1} & A_{12} \\ \hline 0 & A_{22} \end{array}\right]$$

$$AQ = AXR^{-1} = XBR^{-1} = QRBR^{-1}$$

$$A_{11} = Q^T A Q = (\underbrace{Q^T Q}_{I}) RBR^{-1} = RBR^{-1}$$

$$A_{21} = \underbrace{Q'^T (Q}_{0} RBR^{-1}) = 0 \qquad \text{QED}$$

Proof of Real Schur Form

if $Ax = \lambda x$, x, λ real, reduces to

$n-1 \times n-1$ problem as before

if x, λ complex, take real, imag parts of $Ax = \lambda x$

$$X = [re(x), im(x)], \quad B = \begin{bmatrix} re(\lambda) & im(\lambda) \\ -im(\lambda) & re(\lambda) \end{bmatrix}$$

$AX = XB$ first col: $real(Ax) = real(\lambda x)$

2nd col: $imag(Ax) = imag(\lambda x)$

X invariant subspace, real

evals(B) are $\lambda, \bar{\lambda}$

use Lemma to do real orthogonal similarity to get
$$= \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$
QED

Recall other eigenproblems:

(1) ODE $x'(t) = Kx(t)$ $\quad x(0)$ given

if $Kx(0) = \lambda x(0)$ then $x(t) = e^{\lambda t}x(0)$

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

(2) $Mx''(t) + Kx(t) = 0$

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

"generalized eval problem"

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

(3) $Mx''(t) + Dx'(t) + Kx(t) = 0$

$\Rightarrow$ nonlinear eval problems

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

reduce this to linear matrices

(2 matrices) of 2x size

(4) $x'(t) = Ax(t) + Bu(t)$

"linear control system"

turns into eval problem: "singular"

with $[B, A]$ and $[0, I]$ rectangular

All ideas of Chap 4:

Jordan form, Schur form, algs,

perturbation theory, generalize to these

We will concentrate on one square nonsymmetric matrix

other cases: 4.5

symmetric A: Chap 5 (including SVD)

Perturbation Theory: can I trust my answers?

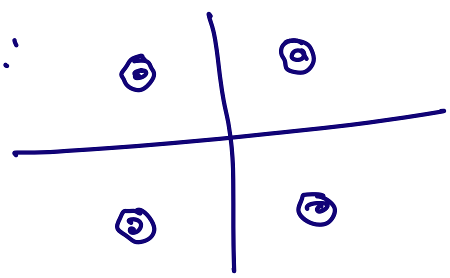
Last time: $A = I^{2\times 2}$: showed evecs very sensitive: close evals $\Rightarrow$ sensitive evecs

Describe perturbation theory for evals:

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

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

Best case: smallest possible $\Lambda_\varepsilon(A)$: each disk around eval of A, radius ε, attained by $E = \varepsilon' I$ $\quad |\varepsilon'| \le \varepsilon$

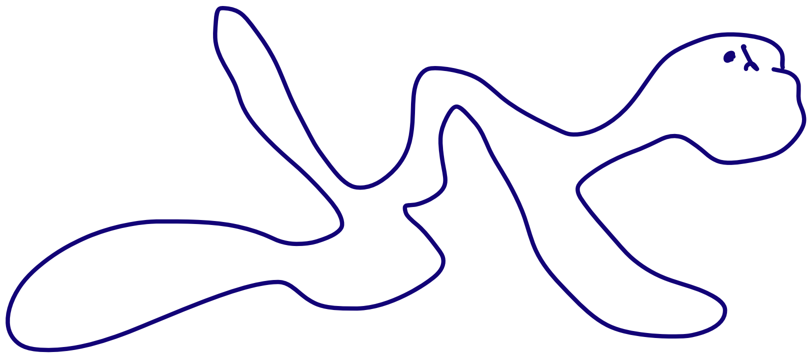


attained when $A = A^H$ (Chap 5)

Worst case (Trefethen + Reichel)

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

any λ inside R, any $\varepsilon > 0$



$\exists A$ s.t. A has one eval at λ, $\Lambda_\varepsilon(A)$ fills out R

proof: Ma185 (Riemann Mapping Thm)

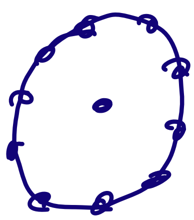
Ex: Perturb $n \times n$ Jordan Block $\lambda = 0$ with $J(n,1) = \varepsilon$

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

$p(\lambda) = \lambda^n \pm \varepsilon$

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

uniformly spaced numbers on circle around origin, radius $\sqrt[n]{|\varepsilon|}$



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

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

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

(2) expect sensitive evals when evals nearly multiple

Condition number of simple (non-multiple) evals (else ∞)

Thm: Let λ be simple eval of A

$$Ax = \lambda x, \quad y^H A = y^H \lambda, \quad \|y\|_2 = \|x\|_2 = 1$$

If we perturb A to $A+E$

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

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

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

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

θ = angle between x and y

$\sec(\theta)$ = condition number

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

$$Ax + A\delta x + Ex + E\delta x = \lambda x + \lambda\delta x + \delta\lambda x + \delta\lambda \cdot \delta x$$

cancel

second order: ignore

$$y^H \left(A\delta x + Ex = \lambda\delta x + \delta\lambda \cdot x \right)$$

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

cancel

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

Special Case 1: $A = A^H$ or "normal"

$AA^H = A^H A$ (HWQ 4.2)

$\Rightarrow A$ has orthonormal evecs

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

$\Rightarrow |\delta\lambda| \le \|E\|_2 + O(\|E\|_2^2)$

i.e. cond # = 1

proof: $A = Q\Lambda Q^H \Rightarrow$

$AQ = Q\Lambda$ and $Q^H A = \Lambda Q^H$

right evecs = cols of Q

left evecs = rows of Q^H = cols of Q

$\Rightarrow x = y \Rightarrow y^H x = 1$

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

then $|\delta\lambda| \le \|E\|_2$ (no $\|E\|_2^2$ term)

for all E

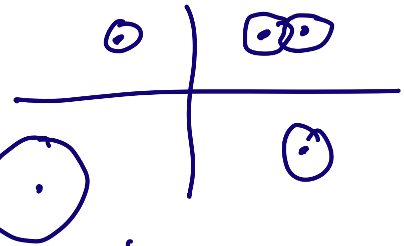
Special case 2: A = Jordan Block

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

$y^H x = 0 \Rightarrow$ cond # $= \sec(\theta) = \infty$

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

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



If k disks overlap, k evals lie in their union (proof in Book)

Algorithms for NonSymmetric Eigenproblem

Ultimate Algorithm: Hessenberg QR (HQR)

Takes any nonsymmetric dense A

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

in $O(n^3)$ flops

Build up to it via sequence of simpler algs, also used in practice to find just a few evals/evecs of large/sparse matrices; HQR also used used for large/sparse matrices, by "approximating" big matrix by smaller projection (dense), solve projection with HQR (Chap 7)

Plan:

Power Method: Just repeated multiplication of x by A, converges to evec for eval of largest magnitude

Inverse Iteration: Apply power method to $B = (A - \sigma I)^{-1}$ which has same evecs as A

largest eval of B corresponds to eval of A closest to σ.

By choosing σ carefully, can converge to any eval/evec pair I want.

Orthogonal Iteration: Extend power method from 1 evec to whole invariant subspace

QR Iteration: Combine Orthogonal Iteration and Inverse Iteration

Other techniques

how to get down to $O(n^3)$

real Schur form

minimize communication

guarantee convergence

(discuss some of these later)