

CS 450: Numerical Analysis

Lecture 10

Chapter 4 – Eigenvalue Problems

Theory of Eigenvalue Solvers

Edgar Solomonik

Department of Computer Science
University of Illinois at Urbana-Champaign

February 16, 2018

Perturbation Analysis of Eigenvalue Problems

- Suppose we seek eigenvalues $D = X^{-1}AX$, but find those of a slightly perturbed matrix $D + \delta D = \hat{X}^{-1}(A + \underline{\delta A})\hat{X}$:

$$\begin{aligned}D + \delta D &= \hat{X}^{-1}A\hat{X} + \hat{X}^{-1}\delta A\hat{X} \\&= X^{-1}(A + \underline{\delta A})X = \underbrace{X^{-1}AX}_{D} + \underline{X^{-1}\delta AX}\end{aligned}$$

$$\text{eig}(A + \delta A) = D + \delta D = \text{eig}(D + \underbrace{X^{-1}\delta AX}_{\delta D})$$

$$\begin{aligned}\|\hat{X}^{-1}\delta A\hat{X}\| &\leq \kappa(\hat{X})\|\delta A\| \\&= \|X^{-1}\| \|X\| \|\delta A\|\end{aligned}$$

$$D = A \quad (\text{matrix is diagonal})$$

$$X = I$$

for some perturbation (non-diagonal) δA

$$\text{eig}(A + \delta A) = D + \delta D \quad \text{and after from } D$$

$$D + \delta D = \hat{X}^{-1} (A + \delta A) \hat{X}$$

$$\hat{X}(\delta D) = (\delta A)\hat{X}$$

$$\begin{cases} \hat{X} = I + \delta X \\ \hat{X} \approx I - \delta X \end{cases}$$

$$\delta D \approx \hat{X}^{-1} \delta A \hat{X} - S X A + A S X$$

$$\| \delta D \| = O(\delta A)$$

$$\underbrace{\delta X \delta D}_{\delta X \delta D \sim \delta A} \approx \delta A$$

Gershgorin Theorem

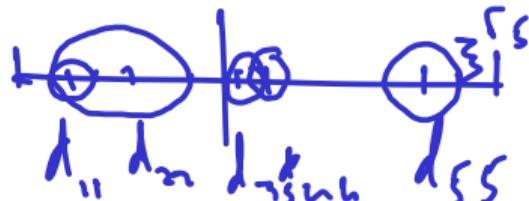
- ▶ Another way to show that the eigenvalues of a matrix are insensitive to perturbation is via Gershgorin theorem, which states that

$$A = D + O$$

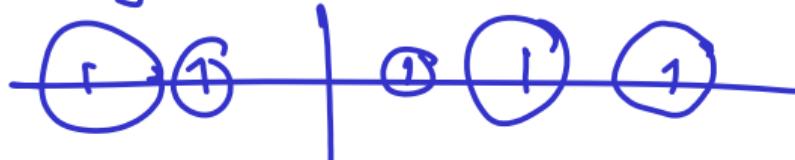
$$\downarrow + \begin{array}{c} \diagup \\ \diagdown \end{array}$$

$$\lambda(A) \in \bigcup_{i=1}^n p(\lambda_i, r_i)$$

↑ ↑ ↓
d.b. center r.adm



$$r_i = \sum_{j=1}^n |O_{ij}|$$



Conditioning of Particular Eigenpairs

- Consider the effect of a matrix perturbation on an eigenvalue λ associated with a right eigenvector x and a left eigenvector y^H , $\lambda = y^H A x / y^H x$

$$A = X^{-1} D X$$

$$\begin{bmatrix} 1 \\ x \end{bmatrix} \xrightarrow{\text{?}} \begin{bmatrix} y \\ \vdots \end{bmatrix}$$
$$\frac{\langle y, x \rangle}{\langle y, (A + \delta A)x \rangle} = \frac{\langle y, Ax \rangle}{\langle y, x \rangle} - \lambda = \frac{\langle y, \delta A x \rangle}{\langle y, x \rangle} \leq \frac{\|\delta A\|_F}{\langle y, x \rangle}$$

- Connect the notion of the angle between left and right eigenvectors to the magnitude of off-diagonal entries in the Schur form

$$A = \begin{bmatrix} \text{orthogonal} & \ddagger \\ & \ddagger \end{bmatrix} \begin{bmatrix} \text{upper} \\ \text{triangular} \end{bmatrix} \begin{bmatrix} & & & \\ & & & \\ & & & \\ & & & \end{bmatrix}$$

off-diagonal entries must be small
 $\langle y, x \rangle \approx 0$

Orthogonal Iteration via QR Iteration

hats¹ für orthogonal
Iteration

In orthogonal iteration $\hat{Q}_{i+1}\hat{R}_{i+1} = A\hat{Q}_i$ QR iteration computes

$$A := Q_i R_i, \quad A_{i+1} = R_i Q_i = \hat{Q}_{i+1}^T A \hat{Q}_{i+1} \text{ at iteration } i:$$

$$\begin{matrix} U = I_m & A_i = I_m \\ \downarrow \rightarrow \infty & \downarrow \rightarrow \infty \end{matrix}$$

$$\text{true} \rightarrow A_i = \hat{Q}_i^+ A \hat{Q}_i$$

induction

$$\text{show} \rightarrow \underline{A_{i+1}} = \hat{Q}_{i+1}^T A \hat{Q}_{i+1}$$

$$\boxed{[\hat{Q}_{:,1}, \dots, \hat{R}_{:,1}]} := \underline{\text{QR}(A\hat{Q})}$$

$$A\hat{Q}_i = \hat{Q}_i, \quad A_i = \hat{Q}_i \circledast \hat{Q}_i$$

$$\hat{Q}_{i+1} \hat{R}_{i+1} = \frac{\hat{Q}_i Q_i R_i}{\hat{Q}_{i+1}^T A \hat{Q}_{i+1}} = \hat{R}_{i+1} \hat{Q}_i^T \hat{Q}_{i+1}$$

$$= R_i \circledast Q_i$$

QR Iteration with Shift

- ▶ Describe QR iteration with shifting

$$[Q_i, R_i] = QR(A_i - \sigma_i I)$$

$$A_{i+1} = R_i Q_i + \sigma_i I$$

- ▶ Discuss how shift can be selected

$$A_{i+1} := Q_i^+ (A_i - \sigma_i I) Q_i + \sigma_i I$$

$$\sigma_i := (A_i)_{nn}$$

$$A_i = \begin{pmatrix} \text{triangular} \\ \text{dotted} \end{pmatrix} \quad \begin{pmatrix} \text{triangular} \\ \text{dotted} \end{pmatrix}$$

Hessenberg and Tridiagonal Form

- ▶ Describe reduction to Hessenberg form

A diagram showing the reduction of a general matrix to Hessenberg form. On the left, a general matrix with non-zero entries in the main and super-diagonals is shown. An arrow labeled $Q \rightarrow$ points to the right, where the matrix is transformed into a Hessenberg form (upper Hessenberg). This transformation is represented as $Q^{-1} \cdot \text{original matrix} \cdot Q$. The resulting Hessenberg matrix has zero entries below the super-diagonal. Brackets on the right side group the columns of Q and Q^T .

- ▶ Describe reduction to tridiagonal form in symmetric case

in symmetric case, similarly
transformer introduces to
rows and cols

QR Iteration Complexity

- ▶ Compare complexity of QR iteration for various matrices

with Tridiagonal form QR . iterates
 $\mathcal{O}(n)$ per . iteration

Hessenberg form

$\mathcal{O}(n^2)$ per . iteration

Circul form

$\mathcal{O}(n^3)$ per . iteration

