CS 450: Numerical Anlaysis

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 $\boldsymbol{D} = \boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X}$, but find those of a slightly perturbed matrix $\boldsymbol{D} + \boldsymbol{\delta D} = \hat{\boldsymbol{X}}^{-1}(\boldsymbol{A} + \boldsymbol{\delta A})\hat{\boldsymbol{X}}$:

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

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

$$\operatorname{eig}(A + \delta A) = D + \delta D = \operatorname{eig}(D + \underline{X^{-1} \delta A X})$$

$$\| X^{-1} \delta A X \| \leq \kappa(X) \| \delta A \|$$

$$= \| X^{-1} \| \, \| X \| \, \| \delta A \|$$

$D = A$ (matrix is diagonal)

$X = I$

for some perturbation (non-diagonal) δA

$\mathrm{eig}(A + \delta A) = D + \delta D$ differ from D

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

$$\boxed{\begin{aligned} \hat{X} &= I + \delta X \\ \hat{X}^{-1} &\approx I - \delta X \end{aligned}}$$

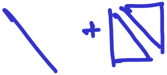
$$\delta D \approx \hat{X}^{-1}\,\delta A\,\hat{X} \underbrace{- \delta X \underset{D}{A} + \underset{D}{A}\,\delta X}_{\delta X\ \delta D\ \sim\ \delta A}$$

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

Gershgorin Theorem

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

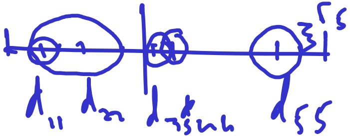
$$A = D + O$$

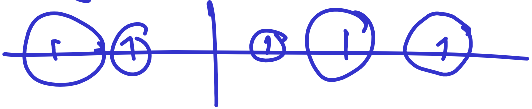


$$\lambda(A) \in \bigcup_i D(\underbrace{d_{ii}}_{\text{center}}, \underbrace{r_i}_{\text{radius}})$$

disk

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





Conditioning of Particular Eigenpairs

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

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

$\langle y, x\rangle, \quad \frac{\langle y, Ax\rangle}{\langle y, x\rangle} = \frac{\langle y, \lambda x\rangle}{\langle y, x\rangle}$

$\langle y, (A+\delta A)x\rangle - \lambda = \frac{\langle y, \delta A x\rangle}{\langle y, x\rangle} \leq \frac{\Delta}{\langle y, x\rangle}$ $\quad \|\delta A x\| \leq \Delta$

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

orthogonal $\quad T$

$A = [\quad][\quad][\quad]$

$Q \quad Q^T$

off-diagonal entries need be small

$\langle y, x\rangle$ is ≈ 1

Orthogonal Iteration via QR Iteration

hats ^ for orthogonal iteration

- In orthogonal iteration $\boxed{\hat{\boldsymbol{Q}}_{i+1}\hat{\boldsymbol{R}}_{i+1} = \boldsymbol{A}\hat{\boldsymbol{Q}}_i}$ QR iteration computes $\underline{A_i = Q_i R_i}$ $\boldsymbol{A}_{i+1} = \boldsymbol{R}_i\boldsymbol{Q}_i = \hat{\boldsymbol{Q}}_{i+1}^T\boldsymbol{A}\hat{\boldsymbol{Q}}_{i+1}$ at iteration i:

$U = \lim_{i\to\infty} A_i = \lim_{i\to\infty} \hat{Q}_{i+1}^T A \hat{Q}_{i+1}$

$[\hat{Q}_{i+1}, \hat{R}_{i+1}] = QR(A\hat{Q}_i)$

true → $\underline{A_i = \hat{Q}_i^T A \hat{Q}_i}$

induction

shows → $\underline{A_{i+1}} = \hat{Q}_{i+1}^T A \hat{Q}_{i+1}$

$A\hat{Q}_i = \hat{Q}_i A_i = \hat{Q}_i Q_i R_i$

$\underline{\hat{Q}_{i+1}\hat{R}_{i+1}} = \underline{\hat{Q}_i Q_i R_i}$

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

$= R_i Q_i$

QR Iteration with Shift

- Describe QR iteration with shifting

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

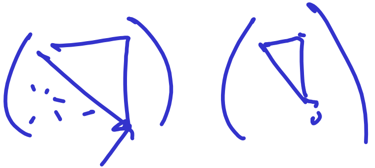
$$A_{i+1} = R_i Q_i + \sigma_i I$$

- Discuss how shift can be selected

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

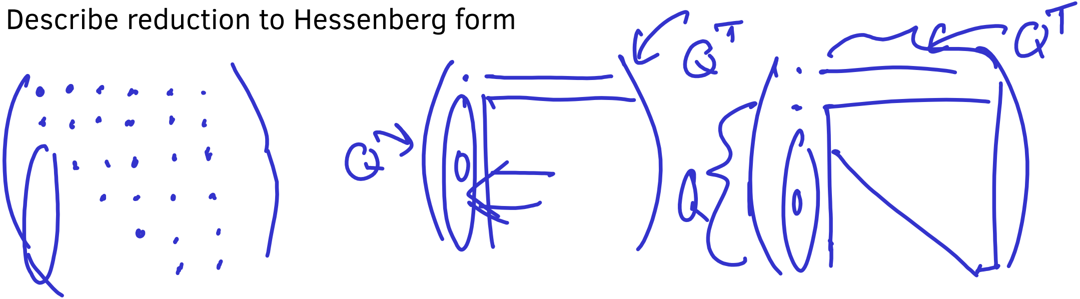
$$\sigma_i := (A_i)_{nn}$$

$A_i =$



Hessenberg and Tridiagonal Form

- Describe reduction to Hessenberg form



- Describe reduction to tridiagonal form in symmetric case

in symmetric case, similarity transformations introduces to rows and cols

QR Iteration Complexity

- Compare complexity of QR iteration for various matrices

with tridiagonal form QR iteration
costs $O(n)$ per iteration

Hessenberg form
$O(n^2)$ per iteration

General form
$O(n^3)$ per iteration