CS 450: Numerical Anlaysis

Lecture 11

Chapter 4 – Eigenvalue Problems

Krylov Subspace Methods and the Generalized Eigenvalue Problem

Edgar Solomonik

Department of Computer Science
University of Illinois at Urbana-Champaign

February 21, 2018

Eigenvalues and the Field of Values

- The field of values is the set of possible Rayleigh quotients of matrix $\boldsymbol{A}$:

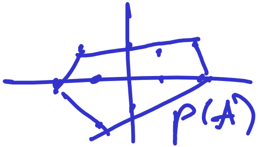
$$\rho(A) = \left\{ r : r = \frac{x^T A x}{x^T x},\ x \in \mathbb{C}^n \right\}$$

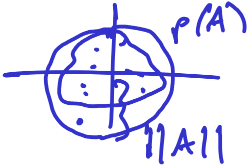
$$\forall r \in \rho(A),\ r \leq \sigma_{max} = \|A\|_2$$

- If and only if the matrix is normal, the field of values is the convex hull of the eigenvalues:

$$A^H A = A A^H$$

λ





Canonical Forms

- Any matrix is *similar* to a matrix in *Jordan* form:

$$A = X \underbrace{J}_{\begin{bmatrix} \lambda_1 & 1 & \\ & \lambda_1 & \\ & & \ddots \end{bmatrix}} X^{-1}$$

Jordan block

$$\begin{bmatrix} \lambda_1 & 1 & & & & \\ & \lambda_1 & & & & \\ & & \lambda_2 & & & \\ & & & \lambda_3 & 1 & \\ & & & & \lambda_3 & 1 \\ & & & & & \lambda_3 \end{bmatrix}$$

if A is diagonalizable

J is diagonal

- Any matrix is *orthogonally similar* to a matrix in *Schur* form:

$$A = Q T Q^H$$

T is upper-triangular

eigenvalues of A are $\mathrm{diag}(T)$

Computing Eigenvectors of Matrices in Schur Form

- Relate eigenvectors of matrices that are similar:

$$A = SBS^{-1}$$

$$A = \hat{X} D \hat{X}^{-1} = \underbrace{SX}_{\hat{X}} D X^{-1} S^{-1}$$

$$\text{if } B = XDX^{-1} \Rightarrow D = X^{-1} B X$$

- Its easy to obtain eigenvectors of triangular matrix $T = \begin{bmatrix} T_{11} & T_{12} \\ & T_{22} \end{bmatrix}$:

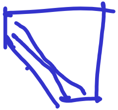
$$T_{11} X_1 = X_1 D_1$$

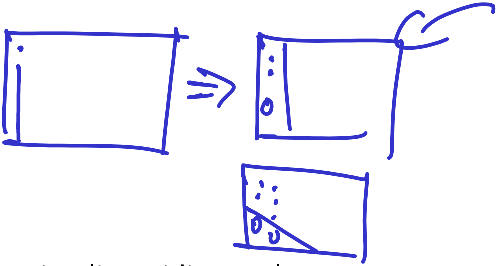
$$T \begin{bmatrix} X_1 \\ 0 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ & T_{22} \end{bmatrix} \begin{bmatrix} X_1 \\ 0 \end{bmatrix} = \begin{bmatrix} X_1 D_1 \\ 0 \end{bmatrix}$$

$$T_{22} Y_2 = Y_2 D_2 \Rightarrow T \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} D_2$$

Matrix Reductions

- Any matrix can be reduced by an orthogonal similarity transformation to Hessenberg form:

$Q^T A Q = H =$ 



- In the symmetric case, Hessenberg form implies tridiagonal:

Hessenberg + symmetric = tridiagonal

Solving Hessenberg Nonsymmetric Eigenproblems

- Eigenvalues of a Hessenberg matrix are usually computed by QR iteration:

$H_i = Q_i R_i$

$H_{i+1} = \underline{R_i Q_i}$

$\rightsquigarrow Q \cdot$

$\cdot\, Q =$

$n-1$ Givens rotations

simple shift $\sigma_i = (H_i)_{nn}$

- Good convergence guarantees given by Francis (Wilkinson) shift:

$H_i =$

$H_i - \sigma_i I = Q_i R_i$

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

eigenvalues of 2×2 block of H_i

$O(1)$ iterations per eigenvalue

$O(n^3)$

Solving Tridiagonal Symmetric Eigenproblems

$$\begin{bmatrix} 0 & u \\ h & 0 \end{bmatrix}$$

A rich variety of methods exists for the tridiagonal eigenproblem:

- QR iteration

  $n-1$ Givens rotations, each requires $O(1)$ work

  $$D = \dots Q_3 Q_2 Q_1 T \underbrace{Q_1 Q_2 Q_3 \dots Q_n}_{\text{can generally be dense}}$$

  can require $O(n^3)$ work

- Divide and conquer

  $$T = \left[\begin{array}{c|c} T_1 & u \\ \hline u & T_2 \end{array}\right] = \left[\begin{array}{c|c} T_1 & \\ \hline & T_2 \end{array}\right] + \underbrace{\begin{bmatrix} \\ z \\ h \\ \\ \end{bmatrix} \begin{bmatrix} & xy & \end{bmatrix}}_{\left[\begin{array}{c|c} & u \\ \hline u & \end{array}\right]}$$

Assume by induction that $T_1 = D_1 + u_1 v_1^T$

$T_2 = D_2 + u_2 v_2^T$

$$T = \begin{bmatrix} D_1 & \\ & D_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ 2 \\ u_2 \end{bmatrix} \begin{bmatrix} v_1 & 2 & v_2 \end{bmatrix}$$

tridiagonal $=$ $\underbrace{\text{diagonal} + \text{rank-1}}_{\text{determinant gives secular eqn.}}$

Solving Tridiagonal Symmetric Eigenproblems (II)

- Jacobi ~~iteration~~ rotations

$$A_{i+1} = Q_i^T A_i Q_i$$

Givens rotation defined to annihilate the entry of A with largest magnitude

- Bisection

partition spectrum to isolate eigenvalues

~~[illegible]~~ eigvals

inertia of $A - \sigma I$

$\hookrightarrow$ the number of positive / negative / zero eigenvalues

- Relatively robust representation (RRR and MRRR)

eigvals of T are sensitive to values in T but not to values in LDL^T

$B = XAX^T$

inertia is preserved

$A = \underline{LDL^T}$ (D diagonal)

inertia(D) = inertia(A)

Krylov Subspace Methods: Motivation

- Many important problems require computation of extremal eigenvalues of sparse matrices:

  e.g. dominant
  few largest
  few least

- QR iteration is too expensive in this case:

  sparse → QR iteration → dense matrix
  (Hessenberg reduction)

- Inverse iteration with deflation provides one approach:

  inverse iteration allows us to compute eigenvalue closest to σ or to 0

  $B \Leftarrow Q^{-1} A Q$

  $$Q A Q^{-1} = \begin{bmatrix} \lambda & b^T \\ 0 & B \end{bmatrix}$$

  → deflated subproblem may not be sparse

Krylov Subspace Methods

- Define k-dimensional Krylov subspace matrix $\boldsymbol{K}_k = \mathcal{K}(\boldsymbol{A}, \boldsymbol{x_0})$:

$$K_k = [x_0, Ax_0, A^2x_0, \dots, A^{k-1}x]$$

$S = \operatorname{span}(K_k)$

$p(A)x_0, \; p(x)$:= legal $k-1$ polynomial

- Show that $\boldsymbol{K}_k$ is a similarity transformation of $\boldsymbol{A}$ to a companion matrix $\boldsymbol{C}$

$$K_n^{-1}AK_n = K_n^{-1}[Ax_0, A^2x_0, \dots A^{n-1}x] = \begin{bmatrix} 0 & \\ I & c \end{bmatrix} = \begin{bmatrix} 0 & & & | \\ & 0 & & | \\ & & 0 & | \\ & & & 0 \; | \end{bmatrix} \rightarrow \text{companion matrix}$$

$$AK_k = K_k C_k$$

- QR factorization of $\boldsymbol{K}_k$ gives an orthogonal representation of the Krylov subspace:

$$\boxed{Q_k^T A Q_k = H_k}$$

$Q_k R_k = K_k$ then $K_n^{-1}AK = C_n$

upper-Hessenberg

$$Q_n^T A Q_n = R C_n R^{-1} = H$$

Rayleigh-Ritz Procedure

- The eigenvalues/eigenvectors of $\boldsymbol{H}_k$ are the *Ritz values/vectors*:

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

Ritz values of A

eigenvectors of H = Ritz vectors of A

- The Ritz vectors and values are the *best possible approximation* of the actual eigenvalues and eigenvectors given a k-dimensional Krylov subspace: