CS 450: Numerical Anlaysis

Lecture 11

Chapter 4 – Eigenvalue Problems

Direct Eigenvalue Solvers and the Symmetric Tridiagonal Eigenproblem

Edgar Solomonik

Department of Computer Science
University of Illinois at Urbana-Champaign

February 26, 2018

Eigenvalues and the Field of Values

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

$$W(\boldsymbol{A}) = \max_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\boldsymbol{x}^H \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^H \boldsymbol{x}}$$

- If and only if the matrix is normal, the field of values is the convex hull of the eigenvalues:
  For $\boldsymbol{A} = \boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{-1}$
  - *all eigenvalues are in the field of values,* $\forall i, d_{ii} \in W(\boldsymbol{A})$*.*
  - *if the matrix is normal,* $\boldsymbol{X}^{-1} = \boldsymbol{X}^T$*,*

$$W(\boldsymbol{A}) = \left\{ s : s = \sum_{i=1}^{n} x_i d_{ii}, ||\boldsymbol{x}||_1 \leq 1 \right\}$$

Canonical Forms

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

$$\boldsymbol{A} = \boldsymbol{X} \begin{bmatrix} \boldsymbol{J}_1 & & \\ & \ddots & \\ & & \boldsymbol{J}_k \end{bmatrix} \boldsymbol{X}^{-1}, \quad \forall i, \quad \boldsymbol{J}_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

  the Jordan form is unique modulo ordering of the diagonal Jordan blocks.

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

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^T$$

  where $\boldsymbol{T}$ is upper-triangular, so the eigenvalues of $\boldsymbol{A}$ is the diagonal of $\boldsymbol{T}$

Computing Eigenvectors of Matrices in Schur Form

- Given the eigenvectors of one matrix, we seek those of a similar matrix:
  Suppose that $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1}$ *and* $\boldsymbol{B} = \boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{-1}$ *where* $\boldsymbol{D}$ *is diagonal,*
  - *the eigenvalues of* $\boldsymbol{A}$ *are* $\boldsymbol{D}$
  - $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{B}\boldsymbol{S}^{-1} = \boldsymbol{S}\boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{-1}\boldsymbol{S}^{-1}$ *so* $\boldsymbol{S}\boldsymbol{X}$ *are the eigenvectors of* $\boldsymbol{A}$

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

  If $\boldsymbol{X}_1$ *are eigenvectors of* $\boldsymbol{T}_1$, $\begin{bmatrix} \boldsymbol{X_1} \\ \boldsymbol{0} \end{bmatrix}$ *are eigenvectors of* $\boldsymbol{T}$, *while if* $\boldsymbol{Y}_2$ *are eigenvectors of* $\boldsymbol{T}_2$, *then* $\begin{bmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \end{bmatrix}$ *are eigenvectors of* $\boldsymbol{T}$ *where* $\boldsymbol{Y}_1 = \boldsymbol{T}_1^{-1}\boldsymbol{T}_{12}\boldsymbol{T}_2$

Matrix Reductions

- Any matrix can be reduced by an orthogonal similarity transformation to upper-Hessenberg form $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{H}\boldsymbol{Q}^T$:

  We can reduce to upper-Hessenberg by successive Householder transformations

$$\boldsymbol{A} = \begin{bmatrix} h_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \\ \vdots & & \ddots \end{bmatrix} = \boldsymbol{Q}_1 \begin{bmatrix} h_{11} & a_{12} & \cdots \\ h_{21} & t_{22} & \cdots \\ \mathbf{0} & \vdots & \ddots \end{bmatrix} = \boldsymbol{Q}_1 \begin{bmatrix} h_{11} & h_{12} & \cdots \\ h_{21} & h_{22} & \cdots \\ \mathbf{0} & \vdots & \ddots \end{bmatrix} \boldsymbol{Q}_1^T = \cdots$$

  subsequent columns can be reduced by induction, so we always can and know how to reduce to upper-Hessenberg with roughly the same cost as QR.

- In the symmetric case, Hessenberg form implies tridiagonal:

  If $\boldsymbol{A} = \boldsymbol{A}^T$ then $\boldsymbol{H} = \boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T = \boldsymbol{H}^T$, and a symmetric upper-Hessenberg matrix must be tridiagonal

Solving Hessenberg Nonsymmetric Eigenproblems

- Eigenvalues of a Hessenberg matrix are usually computed by QR iteration:
  Using $\boldsymbol{A}_0 = \boldsymbol{H}$*, with a shift of* σ_i *at iteration* i *QR iteration is*

$$\boldsymbol{Q}_i \boldsymbol{R}_i = \boldsymbol{A}_i - \sigma_i \boldsymbol{I}$$
$$\boldsymbol{A}_{i+1} = \boldsymbol{R}_i \boldsymbol{Q}_i + \sigma_i \boldsymbol{I}$$

- Good convergence guarantees given by Francis (Wilkinson) shift:
  To handle complex eigenvalues, diagonalize the bottom-right 2-by-2 block of $\boldsymbol{A}_i$ *and use the eigenvalues* $\sigma_i, \bar{\sigma}_i$ *as the next two shifts (also possible to reorganize and do a double-step with two shifts).*

Solving Tridiagonal Symmetric Eigenproblems

A rich variety of methods exists for the tridiagonal eigenproblem:

- QR iteration *requires* $O(1)$ *QR factorizations per eigenvalue,* $O(n^2)$ *cost to get eigenvalues,* $O(n^3)$ *for eigenvectors. The last cost leaves room for improvement.*
- Divide and conquer *reduces tridiagonal* $\boldsymbol{T}$ *by a similarity transformation to a rank-1 perturbation of identity, then computes its eigenvalues using roots of secular equation*

$$
\begin{aligned}
\boldsymbol{T} &= \begin{bmatrix} \boldsymbol{T}_1 & t_{n/2+1,n/2}\boldsymbol{e}_{n/2}\boldsymbol{e}_1^T \\ t_{n/2+1,n/2}\boldsymbol{e}_1\boldsymbol{e}_{n/2}^T & \boldsymbol{T}_2 \end{bmatrix} \\
&= \begin{bmatrix} \hat{\boldsymbol{T}}_1 & \\ & \hat{\boldsymbol{T}}_2 \end{bmatrix} + t_{n/2+1,n/2} \begin{bmatrix} \boldsymbol{e}_{n/2} \\ \boldsymbol{e}_1 \end{bmatrix} \begin{bmatrix} \boldsymbol{e}_{n/2}^T & \boldsymbol{e}_1^T \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}_1\boldsymbol{D}_1\boldsymbol{Q}_1^T & \\ & \boldsymbol{Q}_2\boldsymbol{D}_2\boldsymbol{Q}_2^T \end{bmatrix} + \cdots \\
&= \begin{bmatrix} \boldsymbol{Q}_1 & \\ & \boldsymbol{Q}_2 \end{bmatrix} \underbrace{\left(\begin{bmatrix} \boldsymbol{D}_1 & \\ & \boldsymbol{D}_2 \end{bmatrix} + t_{n/2+1,n/2} \begin{bmatrix} \boldsymbol{Q}_1^T\boldsymbol{e}_{n/2} \\ \boldsymbol{Q}_2^T\boldsymbol{e}_1 \end{bmatrix} \begin{bmatrix} \boldsymbol{e}_{n/2}^T\boldsymbol{Q}_1 & \boldsymbol{e}_1^T\boldsymbol{Q}_2 \end{bmatrix} \right)}_{\boldsymbol{D}+\alpha\boldsymbol{u}\boldsymbol{u}^T} \begin{bmatrix} \boldsymbol{Q}_1^T & \\ & \boldsymbol{Q}_2^T \end{bmatrix}
\end{aligned}
$$

Solving the Secular Equation

To solve the eigenproblem at each step, the divide and conquer method needs to diagonalize a rank-1 perturbation of a diagonal matrix

$$\boldsymbol{A} = \boldsymbol{D} + \alpha \boldsymbol{u}\boldsymbol{u}^T$$

- *the characteristic polynomial is*

$$f(\lambda) = 1 - \alpha \boldsymbol{u}^T(\lambda \boldsymbol{I} - \boldsymbol{D})^{-1}\boldsymbol{u} = 1 - \alpha \sum_{i=1}^{n} \frac{u_i^2}{\lambda - d_{ii}} = 0$$

- *this nonlinear equation can be solved efficiently by a variant of Newton's method, that uses hyperbolic rather than linear extrapolations at each step*

Solving Tridiagonal Symmetric Eigenproblems (II)

- Jacobi iteration *classically is performed to eliminate largest value in magnitude, requires* $O(1)$ *sweeps over all nonzeros,* $O(n^2)$ *cost to get eigenvalues,* $O(n^3)$ *to get eigenvectors*
- Bisection *finds a partition point using* $\boldsymbol{LDL}^T$ *factorization or Sturm sequence to compute inertia (#positive eigenvalues, #negatives eigenvalues #zero eigenvalues). Sylvester's inertia theorem shows that inertia is preserved that under any transformation* $\boldsymbol{A} = \boldsymbol{SBS}^T$ *where* $\boldsymbol{S}$ *is an invertible matrix. Consequently, the diagonal* $\boldsymbol{D}$ *matrix in the* $\boldsymbol{LDL}^T$ *factorization has the same inertia as* $\boldsymbol{A}$*. Computing this factorization with various shifts enables successive halving of the approximation interval.*
- Relatively robust representation (RRR and MRRR) *leverages stability of values in* $\boldsymbol{LDL}^T$ *and other techniques to compute all eigenvectors and eigenvalues in* $O(n^2)$ *cost. These factorized forms minimize sensitivity to round-off error.*