CS 450: Numerical Analysis
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
Eigenvalues and the Field of Values

- The field of values is the set of possible Rayleigh quotients of matrix $A$:

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

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

For $A = XDX^{-1}$

- all eigenvalues are in the field of values, $\forall i, d_{ii} \in W(A)$.
- if the matrix is normal, $X^{-1} = X^T$,

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

- Any matrix is \textit{similar} to a matrix in \textit{Jordan form}:

\[
A = X \begin{bmatrix} J_1 & \cdots & \cdots \\ \end{bmatrix} X^{-1}, \quad \forall i, \quad J_i = \begin{bmatrix} \lambda_i & 1 \\ \vdots & \ddots & \ddots \\ \vdots & \ddots & 1 \\ 1 & \ddots & \ddots & \ddots \\ \end{bmatrix}
\]

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

- Any diagonalizable matrix is \textit{orthogonally similar} to a matrix in \textit{Schur form}:

\[
A = QTQ^T
\]

where \( T \) is upper-triangular, so the eigenvalues of \( A \) is the diagonal of \( T \).
Given the eigenvectors of one matrix, we seek those of a similar matrix:

Suppose that $A = SBS^{-1}$ and $B = XDX^{-1}$ where $D$ is diagonal,

- the eigenvalues of $A$ are $D$
- $A = SBS^{-1} = SXD X^{-1} S^{-1}$ so $SX$ are the eigenvectors of $A$

It's easy to obtain eigenvectors of triangular matrix $T = \begin{bmatrix} T_{11} & T_{12} \\ T_{22} \end{bmatrix}$:

If $X_1$ are eigenvectors of $T_1$, $\begin{bmatrix} X_1 \\ 0 \end{bmatrix}$ are eigenvectors of $T$, while if $Y_2$ are eigenvectors of $T_2$, then $\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ are eigenvectors of $T$ where $Y_1 = T_1^{-1} T_{12} T_2$
Any matrix can be reduced by an orthogonal similarity transformation to upper-Hessenberg form $A = QHQ^T$:

We can reduce to upper-Hessenberg by successive Householder transformations

$$A = \begin{bmatrix} h_{11} & a_{12} & \cdots \\ a_{21} & a_{22} \\ \vdots & \ddots & \ddots \end{bmatrix} = Q_1 \begin{bmatrix} h_{11} & a_{12} & \cdots \\ h_{21} & t_{22} & \cdots \\ 0 & \ddots & \ddots \end{bmatrix} = Q_1 \begin{bmatrix} h_{11} & h_{12} & \cdots \\ h_{21} & h_{22} & \cdots \\ 0 & \ddots & \ddots \end{bmatrix} Q_1^T = \cdots$$

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 $A = A^T$ then $H = QAQ^T = H^T$, 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 \( A_0 = H \), with a shift of \( \sigma_i \) at iteration \( i \) QR iteration is

  \[
  Q_i R_i = A_i - \sigma_i I \\
  A_{i+1} = R_i Q_i + \sigma_i I
  \]

- Good convergence guarantees given by Francis (Wilkinson) shift:
  
  To handle complex eigenvalues, diagonalize the bottom-right 2-by-2 block of \( A_i \) 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 $T$ by a similarity transformation to a rank-1 perturbation of identity, then computes its eigenvalues using roots of secular equation

\[
T = \begin{bmatrix}
T_1 & t_{n/2+1,n/2}e_{n/2}e_1^T \\
t_{n/2+1,n/2}e_{n/2}^Te_1 & T_2
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\hat{T}_1 & 0 \\
0 & \hat{T}_2
\end{bmatrix} + t_{n/2+1,n/2} \begin{bmatrix}
e_{n/2} \\
e_1
\end{bmatrix} \begin{bmatrix}
e_{n/2}^T \\
e_1^T
\end{bmatrix} = Q_1D_1Q_1^T Q_2D_2Q_2^T + \cdots
\]

\[
= \begin{bmatrix}
Q_1 & Q_2
\end{bmatrix} \left( \begin{bmatrix}D_1 & 0 \\0 & D_2\end{bmatrix} + t_{n/2+1,n/2} \begin{bmatrix}Q_1^Te_{n/2} \\Q_2^Te_1\end{bmatrix} \begin{bmatrix}e_{n/2}^TQ_1 \\e_1^TQ_2\end{bmatrix} \right) \begin{bmatrix}
Q_1^T \\
Q_2^T
\end{bmatrix}
\]
To solve the eigenproblem at each step, the divide and conquer method needs to diagonalize a rank-1 perturbation of a diagonal matrix

$$A = D + \alpha uu^T$$

-the characteristic polynomial is

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

-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 $LDL^T$ 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 $A = SBS^T$ where $S$ is an invertible matrix. Consequently, the diagonal $D$ matrix in the $LDL^T$ factorization has the same inertia as $A$. Computing this factorization with various shifts enables successive halving of the approximation interval.

- Relatively robust representation (RRR and MRRR) leverages stability of values in $LDL^T$ and other techniques to compute all eigenvectors and eigenvalues in $O(n^2)$ cost. These factorized forms minimize sensitivity to round-off error.