AMS526: Numerical Analysis I (Numerical Linear Algebra)

Lecture 17: Eigenvalue Problems Cont'd

Xiangmin Jiao

SUNY Stony Brook

November 6, 2008

Review: Eigenvalue Decomposition

- *Eigenvalue problem* of $m \times m$ matrix $\boldsymbol{A}$ is $\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$
- *Eigenvalue decomposition* of $\boldsymbol{A}$ is $\boldsymbol{A} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}$
- *Geometric multiplicity* of λ is $\dim(\text{null}(\boldsymbol{A} - \lambda \boldsymbol{I}))$
- *Algebraic multiplicity* of λ is its multiplicity as a root of $p_{\boldsymbol{A}}$
- Algebraic multiplicity $\geq$ geometric multiplicity
- *Similar* matrices have the same eigenvalues, and algebraic and geometric multiplicities

Determinant and Trace

- Determinant of $\boldsymbol{A}$ is $\det(\boldsymbol{A}) = \prod_{j=1}^{m} \lambda_j$, because

$$\det(\boldsymbol{A}) = (-1)^m \det(-\boldsymbol{A}) = (-1)^m p_{\boldsymbol{A}}(0) = \prod_{j=1}^{m} \lambda_j$$

- Trace of $\boldsymbol{A}$ is $\mathrm{tr}(\boldsymbol{A}) = \sum_{j=1}^{m} \lambda_j$, since

$$p_{\boldsymbol{A}}(z) = \det(z\boldsymbol{I} - \boldsymbol{A}) = z^m - \sum_{j=1}^{m} a_{jj} z^{m-1} + O(z^{m-2})$$

$$p_{\boldsymbol{A}}(z) = \prod_{j=1}^{m} (z - \lambda_j) = z^m - \sum_{j=1}^{m} \lambda_j z^{m-1} + O(z^{m-2})$$

- Question: Are these results valid for defective or nondefective matrices?

Unitary Diagonalization and Schur Factorization

- A matrix $\boldsymbol{A}$ is unitarily diagonalizable if $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^*$ for a unitary matrix $\boldsymbol{Q}$
- A hermitian matrix is unitarily diagonalizable, with real eigenvalues
- A matrix $\boldsymbol{A}$ is *normal* if $\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^*$; unitarily diagonalizable $\Leftrightarrow$ normal
  - Examples of normal matrices include hermitian matrices, skew symmetric matrices
  - "$\Rightarrow$" is easy. Prove "$\Leftarrow$" by induction using Schur factorization below
- Every square matrix $\boldsymbol{A}$ has a Schur factorization $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^*$, where $\boldsymbol{Q}$ is unitary and $\boldsymbol{T}$ is upper triangular
- Eigenvalue-revealing factorization of square matrix $\boldsymbol{A}$
  - Diagonalization $\boldsymbol{A} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}$ (nondefective $\boldsymbol{A}$)
  - Unitary Diagonalization $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^*$ (normal $\boldsymbol{A}$)
  - Unitary triangularization (Schur factorization) $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^*$ (any $\boldsymbol{A}$)

Eigenvalue Algorithms

- The most obvious method is to find roots of characteristic polynomial $p_{\boldsymbol{A}}(\lambda)$, but it is very ill-conditioned
- Instead, compute Schur factorization $\boldsymbol{A} = \boldsymbol{QTQ}^*$ by introducing zeros
- However, this cannot be done in finite number of steps:

  Any eigenvalue solver must be iterative

- To see this, consider a general polynomial of degree m

$$p(z) = z^m + a_{m-1}z^{m-1} + \cdots + a_1 z + a_0$$

  There is no closed-form expression for the roots of p: (Abel, 1842)
  In general, the roots of polynomial equations higher than fourth degree cannot be written in terms of a finite number of operations

Eigenvalue Algorithms (Continued)

- However, the roots of $p_{\boldsymbol{A}}$ are the eigenvalues of the *companion matrix*

$$\boldsymbol{A} = \begin{bmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & \vdots \\ & & \ddots & 0 & -a_{m-2} \\ & & & 1 & -a_{m-1} \end{bmatrix}$$

- Therefore, in general, we cannot find the eigenvalues of a matrix in a finite number of steps
- In practice, however, there are algorithms that converge to desired accuracy in a few iterations

Schur Factorization and Diagonalization

- Most eigenvalue algorithms compute Schur factorization $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^*$ by transforming $\boldsymbol{A}$ with similarity transformations

$$\underbrace{\boldsymbol{Q}_j^* \cdots \boldsymbol{Q}_2^* \boldsymbol{Q}_1^*}_{\boldsymbol{Q}^*} \boldsymbol{A} \underbrace{\boldsymbol{Q}_1 \boldsymbol{Q}_2 \cdots \boldsymbol{Q}_j}_{\boldsymbol{Q}},$$

  where $\boldsymbol{Q}_i$ are unitary matrices, which converge to $\boldsymbol{T}$ as $j \to \infty$
- Note: Real matrices might need complex Schur forms and eigenvalues
- Question: For hermitian $\boldsymbol{A}$, what matrix will the sequence converge to?

Two Phases of Eigenvalue Computations

- General $\boldsymbol{A}$: First convert to *upper-Hessenberg* form, then to upper triangular

$$\underset{A \neq A^*}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}} \xrightarrow{\text{Phase 1}} \underset{\text{upper-Hessenberg}}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}} \xrightarrow{\text{Phase 2}} \underset{\text{triangular}}{\begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}}$$

- Hermitian $\boldsymbol{A}$: First convert to *tridiagonal* form, then to diagonal

$$\underset{A = A^*}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}} \xrightarrow{\text{Phase 1}} \underset{\text{tridiagonal}}{\begin{bmatrix} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & \times & \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}} \xrightarrow{\text{Phase 2}} \underset{\text{diagonal}}{\begin{bmatrix} \times & & & & \\ & \times & & & \\ & & \times & & \\ & & & \times & \\ & & & & \times \end{bmatrix}}$$

Introducing Zeros by Similarity Transformations

- First attempt: Compute Schur factorization $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^*$ by applying Householder reflectors from both left and right

$$\underset{\boldsymbol{A}}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}} \xrightarrow{\boldsymbol{Q}_1^* \cdot} \underset{\boldsymbol{Q}_1^*\boldsymbol{A}}{\begin{bmatrix} \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} \\ 0 & \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} \\ 0 & \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} \\ 0 & \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} \\ 0 & \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} \end{bmatrix}} \xrightarrow{\cdot \boldsymbol{Q}_1} \underset{\boldsymbol{Q}_1^*\boldsymbol{A}\boldsymbol{Q}_1}{\begin{bmatrix} \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} \\ \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} \\ \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} \\ \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} \\ \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} & \mathrm{x} \end{bmatrix}}$$

- Unfortunately, the right multiplication destroys the zeros introduced by $\boldsymbol{Q}_1^*$
- This would not work because of Abel's theorem
- However, the subdiagonal entries typically decrease in magnitude

The Hessenberg Form

- Second attempt: try to compute upper Hessenberg matrix $\boldsymbol{H}$ similar to $\boldsymbol{A}$:

$$\underset{\boldsymbol{A}}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}} \xrightarrow{\boldsymbol{Q}_1^* \cdot} \underset{\boldsymbol{Q}_1^* \boldsymbol{A}}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \end{bmatrix}} \xrightarrow{\cdot \boldsymbol{Q}_1} \underset{\boldsymbol{Q}_1^* \boldsymbol{A} \boldsymbol{Q}_1}{\begin{bmatrix} \times & \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ \times & \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \end{bmatrix}}$$

- The zeros introduced by $\boldsymbol{Q}_1^* \boldsymbol{A}$ were not destroyed this time!
- Continue with remaining columns would result in Hessenberg form:

$$\xrightarrow{\boldsymbol{Q}_2^* \cdot} \underset{\boldsymbol{Q}_2^* \boldsymbol{Q}_1^* \boldsymbol{A} \boldsymbol{Q}_1}{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \end{bmatrix}} \xrightarrow{\cdot \boldsymbol{Q}_2} \underset{\boldsymbol{Q}_2^* \boldsymbol{Q}_1^* \boldsymbol{A} \boldsymbol{Q}_1 \boldsymbol{Q}_2}{\begin{bmatrix} \times & \times & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ \times & \times & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & \times & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & & \mathbf{x} & \mathbf{x} & \mathbf{x} \end{bmatrix}} \cdots$$

The Hessenberg Form

- After $m - 2$ steps, we obtain the Hessenberg form:

$$\underbrace{Q_{m-2}^* \cdots Q_2^* Q_1^*}_{Q^*} A \underbrace{Q_1 Q_2 \cdots Q_{m-2}}_{Q} = H = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}$$

- For hermitian matrix A, H is symmetric and hence is tridiagonal

Householder Reduction to Hessenberg

Householder Reduction to Hessenberg Form

for $k = 1$ **to** $m-2$

$$\boldsymbol{x} = \boldsymbol{A}_{k+1:m,k}$$
$$\boldsymbol{v}_k = \text{sign}(x_1)\|\boldsymbol{x}\|_2\boldsymbol{e}_1 + \boldsymbol{x}$$
$$\boldsymbol{v}_k = \boldsymbol{v}_k/\|\boldsymbol{v}_k\|_2$$
$$\boldsymbol{A}_{k+1:m,k:m} = \boldsymbol{A}_{k+1:m,k:m} - 2\boldsymbol{v}_k(\boldsymbol{v}_k^*\boldsymbol{A}_{k+1:m,k:m})$$
$$\boldsymbol{A}_{1:m,k+1:m} = \boldsymbol{A}_{1:m,k+1:m} - 2(\boldsymbol{A}_{1:m,k+1:m}\boldsymbol{v}_k)\boldsymbol{v}_k^*$$

- Note: $\boldsymbol{Q}$ is never formed explicitly.
- Operation count

$$\sim \sum_{k=1}^{m-2} 4(m-k)^2 + 4m(m-k) \sim 4m^3/3 + 4m^3 - 4m^3/2 = 10m^3/3$$

- For Hermitian $\boldsymbol{A}$, operation count would be same as Householder QR: $4m^3/3$

Stability of Hessenberg Reduction

Theorem

Householder reduction to Hessenberg form is backward stable, in that

$$\tilde{Q}\tilde{H}\tilde{Q}^* = A + \delta A, \qquad \frac{\|\delta A\|}{\|A\|} = O(\epsilon_{machine})$$

for some $\delta A \in \mathbb{C}^{m \times m}$

Note: Similar to Householder QR, $\tilde{Q}$ is exactly unitary based on some $\tilde{v}_k$