Lecture 15
Reduction to Hessenberg Form

Songting Luo

Department of Mathematics
Iowa State University

MATH 562 Numerical Analysis II

Outline

1. Reduction to Hessenberg Form

Outline

1. Reduction to Hessenberg Form

Eigenvalue Revealing Factorization

- Eigenvalue-revealing factorization of square matrix $\mathbf{A}$
  - Diagonalization $\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1}$ (nondefective $\mathbf{A}$)
  - Unitary Diagonalization $\mathbf{A} = \mathbf{Q}\Lambda\mathbf{Q}^*$ (normal $\mathbf{A}$)
  - Unitary triangularization (Schur factorization) $\mathbf{A} = \mathbf{QTQ}^*$ (any $\mathbf{A}$)
  - Jordan normal form $\mathbf{A} = \mathbf{XJX}^{-1}$ where $\mathbf{J}$ block diagonal with

$$\mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

- **In general, Schur factorization is used**, because
  - Unitary matrices are involved, so algorithm tends to be more stable
  - If $\mathbf{A}$ is normal, then Schur form is diagonal

"Obvious" Algorithms

- Most obvious method is to find roots of characteristic polynomial $p_{\mathbf{A}}(\lambda)$, but it is very ill-conditioned.
- Another idea is power iteration, using fact that

$$\frac{\mathbf{x}}{\|\mathbf{x}\|}, \frac{\mathbf{A}\mathbf{x}}{\|\mathbf{A}\mathbf{x}\|}, \frac{\mathbf{A}^2\mathbf{x}}{\|\mathbf{A}^2\mathbf{x}\|}, \frac{\mathbf{A}^3\mathbf{x}}{\|\mathbf{A}^3\mathbf{x}\|}, \dots$$

converge to an eigenvector corresponding to the largest eigenvalue of **A** in absolute value, but it may converge very slowly

"Obvious" Algorithms

- Most obvious method is to find roots of characteristic polynomial $p_{\mathbf{A}}(\lambda)$, but it is very ill-conditioned.
- Another idea is power iteration, using fact that

$$\frac{\mathbf{x}}{\|\mathbf{x}\|}, \frac{\mathbf{A}\mathbf{x}}{\|\mathbf{A}\mathbf{x}\|}, \frac{\mathbf{A}^2\mathbf{x}}{\|\mathbf{A}^2\mathbf{x}\|}, \frac{\mathbf{A}^3\mathbf{x}}{\|\mathbf{A}^3\mathbf{x}\|}, \dots$$

  converge to an eigenvector corresponding to the largest eigenvalue of $\mathbf{A}$ in absolute value, but it may converge very slowly
- Instead, compute a eigenvalue-revealing factorization, such as Schur factorization

$$\mathbf{A} = \mathbf{Q}\mathbf{T}\mathbf{Q}^*$$

  by introducing zeros, using algorithms similar to QR factorization

A Fundamental Difficulty

- However, eigenvalue-revealing factorization 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

A Fundamental Difficulty Cont'd

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

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

- 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 precision in a few iterations

Schur Factorization and Diagonalization

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

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

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

Two Phases of Eigenvalue Computations

- General **A**: First convert to upper-Hessenberg form, then to upper triangular

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

- Hermitian **A**: First convert to tridiagonal form, then to diagonal

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

- In general, phase 1 is direct and requires $O(m^3)$ flops, and phase 2 is iterative and requires $O(m)$ iterations, and $O(m^3)$ flops for non-Hermitian matrices and $O(m^2)$ flops for Hermitian matrices

Introducing Zeros by Similarity Transformations

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

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

- Unfortunately, the right multiplication destroys the zeros introduced by $\mathbf{Q}_1^*$.
- This would not work because of Abel?s theorem
- However, the subdiagonal entries typically decrease in magnitude (Later with QR)

The Hessenberg Form

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

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

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

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

The Hessenberg Form

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

$$\mathbf{Q}_{m-2}^* \cdots \mathbf{Q}_2^* \mathbf{Q}_1^* \mathbf{A} \mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_{m-2} = \mathbf{H} = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}$$

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

$$\mathbf{Q}_{m-2}^* \cdots \mathbf{Q}_2^* \mathbf{Q}_1^* \mathbf{A} \mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_{m-2} = \mathbf{H} = \begin{bmatrix} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & \times & \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}$$

Householder Reduction to Hessenberg

Householder Reduction to Hessenberg Form

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

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

- Note: **Q** is never formed explicitly.
- Operation count

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

Reduction to Tridiagonal Form

- If **A** is hermitian, then

$$\mathbf{Q}_{m-2}^* \cdots \mathbf{Q}_2^* \mathbf{Q}_1^* \mathbf{A} \mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_{m-2} = \mathbf{H} = \begin{bmatrix} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & \times & \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}$$

- For Hermitian **A**, operation count would be same as Householder QR: $4m^3/3$
  - First, taking advantage of sparsity, cost of applying right reflectors is also $4(m-k)^2$ instead of $4m(m-k)$, so cost is

$$\sim \sum_{k=1}^{m-2} 8(m-k)^2 \sim 8m^3/3$$

  - Second, taking advantage of symmetry, cost is reduced by 50% to $4m^3/3$

Stability of Hessenberg Reduction

Theorem

Householder reduction to Hessenberg form is backward stable, in that

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

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

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