AMS526: Numerical Analysis I (Numerical Linear Algebra)

Lecture 14: Gaussian Elimination and LU Factorization

Xiangmin Jiao

SUNY Stony Brook

October 28, 2008

Gaussian Elimination and LU Factorization

- Gaussian elimination was covered intensively in AMS 505
- Gaussian elimination can be viewed as "triangular triangularization" of *nonsingular* $\boldsymbol{A} \in \mathbb{C}^{m \times m}$

$$\underbrace{\boldsymbol{L}_{m-1} \cdots \boldsymbol{L}_2 \boldsymbol{L}_1}_{\boldsymbol{L}^{-1}} \boldsymbol{A} = \boldsymbol{U}$$

analogous to Householder QR factorization of matrix $\boldsymbol{A} \in \mathbb{C}^{m \times n}$

$$\underbrace{\boldsymbol{Q}_n \cdots \boldsymbol{Q}_2 \boldsymbol{Q}_1}_{\boldsymbol{Q}^*} \boldsymbol{A} = \boldsymbol{R}$$

- Example of LU factorization of 4×4 matrix $\boldsymbol{A}$

$$\xrightarrow{\boldsymbol{L}_1} \underbrace{\begin{bmatrix} \times & \times & \times & \times \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \end{bmatrix}}_{\boldsymbol{L}_1 \boldsymbol{A}} \xrightarrow{\boldsymbol{L}_2} \underbrace{\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & 0 & \mathbf{x} & \mathbf{x} \\ & 0 & \mathbf{x} & \mathbf{x} \end{bmatrix}}_{\boldsymbol{L}_2 \boldsymbol{L}_1 \boldsymbol{A}} \xrightarrow{\boldsymbol{L}_3} \underbrace{\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & 0 & \mathbf{x} \end{bmatrix}}_{\boldsymbol{L}_3 \boldsymbol{L}_2 \boldsymbol{L}_1 \boldsymbol{A}}$$

What is Matrices $\boldsymbol{L}_k$

- At step k, eliminate entries below a_{kk}: let $\boldsymbol{x}_k$ be kth column of $\boldsymbol{L}_{k-1} \cdots \boldsymbol{L}_1 \boldsymbol{A}$,

$$\boldsymbol{x}_k = [x_{1,k}, x_{2,k}, \cdots x_{k,k}, x_{k+1,k}, \cdots x_{m,k}]^T$$
$$\boldsymbol{L}_k \boldsymbol{x}_k = [x_{1,k}, x_{2,k}, \cdots x_{k,k}, 0, \cdots 0]^T$$

- The multipliers $\ell_{jk} = x_{jk}/x_{kk}$ appear in $\boldsymbol{L}_k$

$$\boldsymbol{L}_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -\ell_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -\ell_{m,k} & & & 1 \end{bmatrix}$$

- Let $\boldsymbol{\ell}_k = [0, \cdots, 0, \ell_{k+1,k}, \cdots, \ell_{m,k}]^T$ and $\boldsymbol{e}_k = [\underbrace{0, \cdots, 0}_{k-1}, 1, \cdots, 0]^T$,

  then $\boldsymbol{L}_k = \boldsymbol{I} - \boldsymbol{\ell}_k \boldsymbol{e}_k^*$

Forming $\boldsymbol{L}$

- Luckily, the $\boldsymbol{L}$ matrix contains the *multipliers* $\ell_{jk} = x_{jk}/x_{kk}$

$$\boldsymbol{L} = \boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}\cdots\boldsymbol{L}_{m-1}^{-1} = \begin{bmatrix} 1 & & & & \\ \ell_{21} & 1 & & & \\ \ell_{31} & \ell_{32} & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ \ell_{m1} & \ell_{m2} & \dots & \ell_{m,m-1} & 1 \end{bmatrix}$$

  and is said to be a *unit lower triangular matrix*
- First, $\boldsymbol{L}_k^{-1} = \boldsymbol{I} + \boldsymbol{\ell}_k\boldsymbol{e}_k^*$, because $\boldsymbol{e}_k^*\boldsymbol{\ell}_k = 0$ and $(\boldsymbol{I} - \boldsymbol{\ell}_k\boldsymbol{e}_k^*)(\boldsymbol{I} + \boldsymbol{\ell}_k\boldsymbol{e}_k^*) = \boldsymbol{I} - \boldsymbol{\ell}_k\boldsymbol{e}_k^*\boldsymbol{\ell}_k\boldsymbol{e}_k^* = \boldsymbol{I}$
- Second, $\boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}\cdots\boldsymbol{L}_{k+1}^{-1} = \boldsymbol{I} + \sum_{j=1}^{k+1}\boldsymbol{\ell}_j\boldsymbol{e}_j^*$, since (prove by induction) $(\boldsymbol{I}+\sum_{j=1}^{k}\boldsymbol{\ell}_j\boldsymbol{e}_j^*)(\boldsymbol{I}+\boldsymbol{\ell}_{k+1}\boldsymbol{e}_{k+1}^*) = \boldsymbol{I}+\sum_{j=1}^{k+1}\boldsymbol{\ell}_j\boldsymbol{e}_j^*+\sum_{j=1}^{k}\boldsymbol{\ell}_j(\boldsymbol{e}_j^*\boldsymbol{\ell}_{k+1})\boldsymbol{e}_{k+1}^*$ where $\boldsymbol{e}_j^*\boldsymbol{\ell}_{k+1} = 0$ for $j < k+1$
- In other words, $\boldsymbol{L}$ is "union" of $\boldsymbol{L}_1^{-1}, \boldsymbol{L}_2^{-1}, \cdots, \boldsymbol{L}_{m-1}^{-1}$

Gaussian Elimination without Pivoting

- Factorize $\boldsymbol{A} \in \mathbb{C}^{m\times m}$ into $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$

Gaussian elimination without pivoting

$$
\begin{aligned}
&\boldsymbol{U} = \boldsymbol{A}, \boldsymbol{L} = \boldsymbol{I} \\
&\textbf{for } k = 1 \textbf{ to } m-1 \\
&\qquad \textbf{for } j = k+1 \textbf{ to } m \\
&\qquad\qquad \ell_{jk} = u_{jk}/u_{kk} \\
&\qquad\qquad u_{j,k:m} = u_{j,k:m} - \ell_{jk} u_{k,k:m}
\end{aligned}
$$

- Flop count $\sim \sum_{k=1}^{m} 2(m-k)(m-k) \sim 2\sum_{k=1}^{m} k^2 \sim 2m^3/3$
- In actually, $\boldsymbol{L}$ often overwrites lower-triangular part of $\boldsymbol{A}$ and $\boldsymbol{U}$ overwrites upper-triangular part of $\boldsymbol{A}$
- Question: What if u_{kk} is 0? Answer: The algorithm would break.

Partial Pivoting

- At step k, we divide by u_{kk}, which would break if u_{kk} is 0 (or close to 0), which can happen even if $\boldsymbol{A}$ is nonsingular

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ & x_{kk} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \times & \times & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} \times & \times & \times & \times & \times \\ & x_{kk} & \times & \times & \times \\ & 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \end{bmatrix}$$

- However, any nonzero entry in kth column below diagonal can also be used as *pivot*

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & x_{ik} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & \times & \times & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} \times & \times & \times & \times & \times \\ & 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & x_{ik} & \times & \times & \times \\ & 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \end{bmatrix}$$

  and we permute (interchange) row i with row k

- In general, we take nonzero entry with largest absolute value

More on Partial Pivoting

- kth step of Gaussian elimination of partial pivoting

$$\underset{\text{Pivot selection}}{\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & x_{ik} & \mathbf{x} & \mathbf{x} \\ & \times & \times & \times \end{bmatrix}} \xrightarrow{P_k} \underset{\text{Row interchange}}{\begin{bmatrix} \times & \times & \times & \times \\ & x_{kk} & \mathbf{x} & \mathbf{x} \\ & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & \times & \times & \times \end{bmatrix}} \xrightarrow{L_k} \underset{\text{Elimination}}{\begin{bmatrix} \times & \times & \times & \times \\ & x_{kk} & \times & \times \\ & 0 & \mathbf{x} & \mathbf{x} \\ & 0 & \mathbf{x} & \mathbf{x} \end{bmatrix}}$$

  and we interchange row i with row k
- In terms of matrices, it becomes $\underbrace{\boldsymbol{L}_{m-1}\boldsymbol{P}_{m-1}\cdots\boldsymbol{L}_2\boldsymbol{P}_2\boldsymbol{L}_1\boldsymbol{P}_1}_{\boldsymbol{L}^{-1}\boldsymbol{P}}\boldsymbol{A} = \boldsymbol{U}$
- $\boldsymbol{P} = \boldsymbol{P}_{m-1}\cdots\boldsymbol{P}_2\boldsymbol{P}_1$ and $\boldsymbol{L} = \left(\boldsymbol{L}'_{m-1}\cdots\boldsymbol{L}'_2\boldsymbol{L}'_1\right)^{-1}$, where $\boldsymbol{L}'_k = \boldsymbol{P}_{m-1}\cdots\boldsymbol{P}_{k+1}\boldsymbol{L}_k\boldsymbol{P}_{k+1}^{-1}\cdots\boldsymbol{P}_{m-1}^{-1}$
- It is easy to verify that $\boldsymbol{L}_{m-1}\boldsymbol{P}_{m-1}\cdots\boldsymbol{L}_2\boldsymbol{P}_2\boldsymbol{L}_1\boldsymbol{P}_1 = \left(\boldsymbol{L}'_{m-1}\cdots\boldsymbol{L}'_2\boldsymbol{L}'_1\right)\left(\boldsymbol{P}_{m-1}\cdots\boldsymbol{P}_2\boldsymbol{P}_1\right)$
- $\boldsymbol{L}'_k = \boldsymbol{I} - \boldsymbol{P}_{m-1}\cdots\boldsymbol{P}_{k+1}\boldsymbol{\ell}_k\boldsymbol{e}_k^*$ and $\boldsymbol{L}$ is "union" of $\left(\boldsymbol{L}'_k\right)^{-1} \equiv \boldsymbol{I} + \boldsymbol{P}_{m-1}\cdots\boldsymbol{P}_{k+1}\boldsymbol{\ell}_k\boldsymbol{e}_k^*$

Algorithm of Gaussian Elimination with Partial Pivoting

- Factorize $\boldsymbol{A} \in \mathbb{C}^{m \times m}$ into $\boldsymbol{PA} = \boldsymbol{LU}$

Gaussian elimination with partial pivoting

$\boldsymbol{U} = \boldsymbol{A}, \boldsymbol{L} = \boldsymbol{I}, \boldsymbol{P} = \boldsymbol{I}$
for $k = 1$ **to** $m - 1$
  $i \leftarrow \arg\max_{i \geq k} |u_{ik}|$
  $\boldsymbol{u}_{k,k:m} \leftrightarrow \boldsymbol{u}_{i,k:m}$
  $\boldsymbol{\ell}_{k,1:k-1} \leftrightarrow \boldsymbol{\ell}_{i,1:k-1}$
  $p_{k,:} \leftrightarrow p_{i,:}$
  for $j = k + 1$ **to** m
    $\ell_{jk} = u_{jk} / u_{kk}$
    $u_{j,k:m} = u_{j,k:m} - \ell_{jk} u_{k,k:m}$

- Question: What if u_{kk} is 0?
- Flop count $\sim \sum_{k=1}^{m} 2(m-k)(m-k) \sim 2 \sum_{k=1}^{m} k^2 \sim 2m^3/3$, same as without pivoting

An Alternative Implementation

- In practice, $\boldsymbol{L}$ and $\boldsymbol{U}$ overwrite $\boldsymbol{A}$ and $\boldsymbol{P}$ is represented by a vector

Gaussian elimination with partial pivoting (alternative)

$$
\begin{aligned}
&\boldsymbol{p} = [1, 2, \cdots, m]; \\
&\textbf{for } k = 1 \textbf{ to } m - 1 \\
&\quad i \leftarrow \arg\max_{i \geq k} |a_{ik}| \\
&\quad \boldsymbol{a}_{k,1:m} \leftrightarrow \boldsymbol{a}_{i,1:m} \\
&\quad p_k \leftrightarrow p_i \\
&\quad \boldsymbol{a}_{k+1:m,k} \leftarrow \boldsymbol{a}_{k+1:m,k} / a_{k,k} \\
&\quad \boldsymbol{A}_{k+1:m,k+1:m} \leftarrow \boldsymbol{A}_{k+1:m,k+1:m} - \boldsymbol{a}_{k+1:m,k} * \boldsymbol{a}_{k,k+1:m}
\end{aligned}
$$

- Using LU factorization to solve $\boldsymbol{Ax} = \boldsymbol{b}$:
  1. $\boldsymbol{PA} = \boldsymbol{LU}$; (LU factorization with partial pivoting)
  2. $\boldsymbol{Ly} = \boldsymbol{P}^T\boldsymbol{b}$; (Forward substitution)
  3. $\boldsymbol{Ux} = \boldsymbol{y}$; (Back substitution)

Complete Pivoting

- More generally, we can use any nonzero entry
- In theory, any nonzero entry (i, j), $i \geq k$, $j \geq k$

$$\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & \mathbf{x} & x_{ij} & \mathbf{x} \\ & \times & \times & \times \end{bmatrix} \rightarrow \begin{bmatrix} \times & \times & \times & \times \\ & \mathbf{x} & 0 & \mathbf{x} \\ & \times & x_{ij} & \times \\ & \mathbf{x} & 0 & \mathbf{x} \end{bmatrix}$$

  and we then permute row i with row k, column j with column k
- In matrix operations, it can be expressed as

$$\underbrace{\boldsymbol{L}_{m-1}\boldsymbol{P}_{m-1}\cdots\boldsymbol{L}_2\boldsymbol{P}_2\boldsymbol{L}_1\boldsymbol{P}_1}_{\boldsymbol{L}^{-1}\boldsymbol{P}}\boldsymbol{A}\underbrace{\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_{m-1}}_{\boldsymbol{Q}} = \boldsymbol{U}$$

- Therefore, $\boldsymbol{PAQ} = \boldsymbol{LU}$ where $\boldsymbol{P} = \boldsymbol{P}_{m-1}\cdots\boldsymbol{P}_2\boldsymbol{P}_1$ and $\boldsymbol{L} = \left(\boldsymbol{L}'_{m-1}\cdots\boldsymbol{L}'_2\boldsymbol{L}'_1\right)^{-1}$
- However, complete pivoting is typically not used in practice because it increases cost in search of pivot and complexity of implementation