Unit 5 The LU and Cholesky Factorizations

Gaussian Elimination (LU right-looking)

$A = GE(A)$

$A \rightarrow \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right)$ A_{TL} is 0×0

while $n(A_{TL}) < n(A)$

$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \rightarrow \left(\begin{array}{c|cc} A_{00} & a_{01} & A_{02} \\ \hline a_{10}^T & \alpha_{11} & a_{12}^T \\ A_{20} & a_{21} & A_{22} \end{array}\right)$

$a_{21} := a_{21} / \alpha_{11}$

$A_{22} := A_{22} - a_{21} a_{12}^T$

$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \leftarrow \left(\begin{array}{cc|c} A_{00} & a_{01} & A_{02} \\ a_{10}^T & \alpha_{11} & a_{12}^T \\ \hline A_{20} & a_{21} & A_{22} \end{array}\right)$

endwhile

Thm 5.2.2.1

Cost of algorithm when applied to $n \times n$ matrix is $\frac{2}{3} n^3$

Thm 5.2.2.2

Cost of algorithm when applied to $m \times n$ matrix is $mn^2 - \frac{1}{3} n^3$

Definition 5.2.3.1

Given matrix $A \in \mathbb{C}^{m \times n}$ $(m \geq n)$, its LU ~~factoriz~~ factorization is given by $A = LU$ where $L \in \mathbb{C}^{m \times n}$ is a unit lower trapezoidal and $U \in \mathbb{C}^{n \times n}$ is upper triangular with nonzeros on its diagonal

Definition 5.2.3.2 Principal leading submatrix

For $k \leq n$, the $k \times k$ principal leading submatrix of a matrix A is defined to be a square matrix $A_{TL} \in \mathbb{C}^{k \times k}$ such that $A = \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right)$

Lemma 5.2.3.3

Let $L \in \mathbb{C}^{n \times n}$ be ~~lower~~ unit lower triangular matrix and $U \in \mathbb{C}^{n \times n}$ be an upper triangular matrix. Then $A = LU$ is nonsingular if only if U has no zeros on its diagonal.

Therem 5.2.3.4 Existence of the LU factorization

Let $A \in \mathbb{C}^{m \times n}$ and $m \geq n$ have linearly independent cols. Then A has a (unique) LU factorization if only if all its principal leading submatrices are nonsingular

LU left-looking

$A \rightarrow \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right)$ A_{TL} is 0×0

while $n(A_{TL}) < n(A)$

$$\begin{array}{cc} A_{TL} & A_{TR} \\ A_{BL} & A_{BR} \end{array} \rightarrow \left(\begin{array}{c|c|c} A_{00} & a_{01} & A_{02} \\ a_{10}^T & \alpha_{11} & a_{12}^T \\ A_{20} & a_{21} & A_{22} \end{array}\right)$$

Solve $L_{00} u_{01} = a_{01}$ overwriting a_{01} with u_{01}

$\alpha_{11} := v_{11} = \alpha_{11} - a_{10}^T a_{01}$

$a_{21} := a_{21} - A_{20} a_{01}$

$a_{21} := l_{21} = a_{21} / \alpha_{11}$

endwhile

Hw 5.2.3.2

Cost of left-looking LU factorization is $mn^2 - \frac{1}{3} n^3$

Gaussian Elimination via Gauss transforms

$L_{n-1} \ldots L_1 L_0 A = U$

$\Rightarrow A = L_{n-1}^{-1} \ldots L_1^{-1} L_0^{-1}$

$\Rightarrow A = \underbrace{L_0^{-1} L_1^{-1} \ldots L_{n-1}^{-1}}_{L} U$

where $L_k^{-1} = \begin{pmatrix} I_k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & l_{21} & I \end{pmatrix}$

Hw 5.2.4.2

$$\begin{pmatrix} I_k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -l_{21} & I \end{pmatrix}^{-1} = \begin{pmatrix} I_k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & l_{21} & I \end{pmatrix}$$

where I_k denotes the $k \times k$ identity matrix

Hw 5.2.4.3

$\tilde{L}_{k-1} = L_0^{-1} L_1^{-1} \ldots L_{k-1}^{-1} = \begin{pmatrix} L_{00} & 0 & 0 \\ l_{10}^T & 1 & 0 \\ L_{20} & 0 & I \end{pmatrix}$ $\quad L_k^{-1} = \begin{pmatrix} I_k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & l_{21} & I \end{pmatrix}$

$\tilde{L}_k = \tilde{L}_{k-1} L_k^{-1} = \begin{pmatrix} L_{00} & 0 & 0 \\ l_{10}^T & 1 & 0 \\ L_{20} & l_{21} & I \end{pmatrix}$

Gaussian Elimination with row exchanges

Permutation Matrices Revision

Example:

$P\left(\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} e_1^T \\ e_2^T \\ e_0^T \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$

Definition Permutation Matrix

Given $p = \begin{pmatrix} \pi_0 \\ \vdots \\ \pi_{n-1} \end{pmatrix}$ where $\{\pi_0, \ldots \pi_{n-1}\}$ is the permutation of the integers $\{0, 1, \ldots n-1\}$. Then $P(p) = \begin{pmatrix} e_{\pi_0}^T \\ \vdots \\ e_{\pi_{n-1}}^T \end{pmatrix}$

HW 5.3.2.1

$$p = \begin{pmatrix} \pi_0 \\ \vdots \\ \pi_{n-1} \end{pmatrix} \text{ and } x = \begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix}$$

$$P(p)\,x = \begin{pmatrix} e_{\pi_0}^T \\ \vdots \\ e_{\pi_{n-1}}^T \end{pmatrix} \begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} x_{\pi_0} \\ \vdots \\ x_{\pi_{n-1}} \end{pmatrix}$$

HW 5.3.2.2

$$p = \begin{pmatrix} \pi_0 \\ \vdots \\ \pi_{n-1} \end{pmatrix} \text{ and } A = \begin{pmatrix} \tilde{a}_0^T \\ \vdots \\ \tilde{a}_{n-1}^T \end{pmatrix}$$

$$P(p)\,A = \begin{pmatrix} e_{\pi_0}^T \\ \vdots \\ e_{\pi_{n-1}}^T \end{pmatrix} \begin{pmatrix} \tilde{a}_0^T \\ \vdots \\ \tilde{a}_{n-1}^T \end{pmatrix} = \begin{pmatrix} \tilde{a}_{\pi_0}^T \\ \vdots \\ \tilde{a}_{\pi_{n-1}}^T \end{pmatrix}$$

$\Rightarrow$ Apply $P(p)$ to the left permutes rows

HW 5.3.2.3

$$p = \begin{pmatrix} \pi_0 \\ \vdots \\ \pi_{n-1} \end{pmatrix} \text{ and } A = (a_0 \mid \dots \mid a_{n-1})$$

$$\begin{aligned} A\,P(p)^T &= (a_0 \dots a_{n-1})\,(e_{\pi_0} \dots e_{\pi_{n-1}}) \\ &= (A e_{\pi_0} \dots A e_{\pi_{n-1}}) \\ &= (a_{\pi_0} \dots a_{\pi_{n-1}}) \end{aligned}$$

$\Rightarrow$ Apply $P(p)^T$ to the right permutes columns

HW 5.3.2.4

$$\begin{aligned} P(p)\,P(p)^T &= \begin{pmatrix} e_{\pi_0}^T \\ \vdots \\ e_{\pi_{n-1}}^T \end{pmatrix} (e_{\pi_0} \dots e_{\pi_{n-1}}) \\ &= I. \end{aligned}$$

Definition 5.3.2.2 Elementary pivot matrix

Given $\pi \in \{0, \dots, n-1\}$

$$\cancel{\tilde{P}(\pi) = \begin{pmatrix} e_\pi^T \\ \hline e_1^T \\ \vdots \\ e_{\pi-1}^T \\ \hline e_0^T \\ \hline e \end{pmatrix}} \qquad \tilde{P}(\pi) = \begin{pmatrix} e_\pi^T \\ \hline e_1^T \\ \vdots \\ e_{\pi-1}^T \\ \hline e_0^T \\ \hline e_{\pi+1}^T \\ \vdots \\ e_{n-1}^T \end{pmatrix}$$

or equivalently,

$$\tilde{P}(\pi) = \begin{cases} I_n & \text{if } \pi = 0 \\ \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & I_{\pi-1} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{n-\pi-1} \end{pmatrix} & \text{otherwise} \end{cases}$$

Giấy độ trắng tự nhiên
không hại mắt
fahasa

In LU factorization with pivoting, we use the elementary pivot matrices in a very specific way

$$p = \begin{pmatrix} \pi_0 \\ \vdots \\ \pi_{k-1} \end{pmatrix}, \text{ where } 1 \leq k \leq n,\ 0 \leq \pi_i < n-i$$

$$\widetilde{P}(p) = \begin{pmatrix} I_{k-1} & 0 \\ 0 & \widetilde{P}(\pi_{k-1}) \end{pmatrix} \begin{pmatrix} I_{k-2} & 0 \\ 0 & \widetilde{P}(\pi_{k-2}) \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 \\ 0 & \widetilde{P}(\pi_1) \end{pmatrix} \widetilde{P}(\pi_0)$$

There ~~the~~ the π_i is relative to the current row index

LU right looking with pivot

while $n(A_{TL}) < n(A)$

$$\begin{pmatrix} A_{TL} & A_{TR} \\ A_{BL} & A_{BR} \end{pmatrix} \to \left(\begin{array}{c|cc} A_{00} & a_{01} & A_{02} \\ \hline a_{10}^T & \alpha_{11} & a_{12}^T \\ A_{20} & a_{21} & A_{22} \end{array}\right) \qquad \left(\begin{array}{c} p_T \\ \hline p_B \end{array}\right) \to \left(\begin{array}{c} p_0 \\ \hline \pi_1 \\ p_2 \end{array}\right)$$

$$\pi_1 := \max i \left(\begin{array}{c} \alpha_{11} \\ \hline a_{21} \end{array}\right)$$

$$\left(\begin{array}{c|cc} A_{00} & a_{01} & A_{02} \\ \hline a_{10}^T & \alpha_{11} & a_{12}^T \\ A_{20} & a_{21} & A_{22} \end{array}\right) := \left(\begin{array}{c|c} I & 0 \\ \hline 0 & P(\pi_1) \end{array}\right) \left(\begin{array}{c|cc} A_{00} & a_{01} & A_{02} \\ \hline a_{10}^T & \alpha_{11} & a_{12}^T \\ A_{20} & a_{21} & A_{22} \end{array}\right)$$

$a_{21} := a_{21} / \alpha_{11}$

$A_{22} := A_{22} - a_{21} a_{12}^T$

endwhile

We can look at this algorithm as

$$L_{n-1} P_{n-1} \ldots L_1 P_1 L_0 P_0 A = U$$

$$\Leftrightarrow A = P_0^T L_0^{-1} P_1^T L_1^{-1} \ldots P_{n-1}^T L_{n-1}^{-1} U$$

$\langle$and $P_0^T = P_0 \rangle$

Also, we can prove that $A = P_0 P_1 \ldots P_{n-1} \underbrace{L_0^* L_1^* \ldots L_{n-1}^*}_{L} U$

$$\Leftrightarrow \boxed{P_{n-1} \ldots P_1 P_0 A = LU}$$

Solving $Ax = y$ via LU factorization with pivoting

Since $\widetilde{P}(p) A = LU$

$\Rightarrow \widetilde{P}(p) Ax = \underbrace{\widetilde{P}(p) b}_{y}$

$\Leftrightarrow L(\underbrace{Ux}_{z}) = y$

Now we solve $Lz = y$

And then $Ux = z$

Discussion

Computation tends to be more efficient when matrices are accessed by column. Hence Variant 1 of solving $Ax = y$ is better than Variant 2

Personal notes.

$A = \begin{pmatrix} \alpha_{11} & a_{12}^T \\ a_{21} & A_{22} \end{pmatrix}$

When trying to apply Gaussian to A, through pivoting we try to pick α_{11} as large as possible in ~~neg~~ magnitude. This is because multiplier $m = \frac{a_{21}}{\alpha_{11}}$ and ~~m~~ multiplier will be apply in all the elements of matrix A so we want the multiplier to be as ~~sm~~ small as possible to avoid amplifying any error in matrix A.

Definition Hermitian positive definite matrix

Matrix $A \in \mathbb{C}^{n \times n}$ is Hermitian positive definite (HPD) i.o.i it is Hermitian ($A^H = A$) ~~and~~ and $\boxed{x \neq 0 \Rightarrow x^H A x > 0}$ for all $x \in \mathbb{C}^n$.

If $A \in \mathbb{R}^{n \times n}$, then it is called symmetric positive definite (SPD)

Thm 5.4.1.1

Let $B \in \mathbb{C}^{m \times n}$. Then matrix $A = B^H B$ is HPD

Thm 5.4.1.2

Let $A \in \mathbb{C}^{m \times m}$ be HPD

Then the diagonal elements of A are real and positive

Thm 5.4.1.3

Let $A \in \mathbb{C}^{m \times m}$ be HPD and $A = \begin{pmatrix} \alpha_{11} & a_{21}^H \\ a_{21} & A_{22} \end{pmatrix}$

A_{22} is HPD

Theorem 5.4.2.1 Cholesky Factorization

Given an HPD matrix A, there exists a lower triangular matrix L such that:

$A = L L^H$

(If diagonal elements of L are restricted to be positive, L is unique)

Obviously, there also exists matrix U (upper triangular matrix) such that:

$A = U^H U$

Cholesky Factorization Algorithm

while $n(A_{TL}) < n(A)$

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \rightarrow \left(\begin{array}{c|cc} A_{00} & a_{01} & A_{02} \\ \hline a_{10}^T & \alpha_{11} & a_{12}^T \\ A_{20} & a_{21} & A_{22} \end{array}\right)$$

$\alpha_{11} := \lambda_{11} = \sqrt{\alpha_{11}}$

$a_{21} := l_{21} = a_{21}/\alpha_{11}$

$A_{22} := A_{22} - a_{21} a_{21}^H$

...

endwhile

Thm 5.4.3.1

The cost of the algorithm when apply to an $n \times n$ matrix is $\frac{1}{3}n^3$

The cost of Cholesky Factorization is half of that of LU Factorization by taking advantage of symmetry.

Lemma 5.4.4.1

Let $A \in \mathbb{C}^{n \times n}$ be HPD. Then α_{11} is real and positive

Lemma 5.4.4.2

Let $A \in \mathbb{C}^{n \times n}$ be HPD, $l_{21} = a_{21}/\sqrt{\alpha_{11}}$.

Then $A_{22} - l_{21} l_{21}^H$ is HPD

Solving LLS with Cholesky factorization

We try to solve $\| b - A\hat{x} \|_2 = \min_{x \in \mathbb{C}^n} \| b - Ax \|_2$ using the normal equation:

$$\underbrace{A^H A}_{B} \hat{x} = \underbrace{A^H b}_{y}$$

- Form $B = A^H A$ — mn^2 flops
- Factor $B = LL^H$ — $\frac{1}{3}n^3$ flops
- Compute $y = A^H b$ — $2mn$ flops
- Solve $Lz = y$ — n^2 flops
- Solve $L^H \hat{x} = z$ — n^2 flops

Total cost $mn^2 + \frac{1}{3}n^3$ flops

Ponder 5.4.5.1

The Cholesky factorization of $A^H A$ ($A = QR$)

$A^H A = R^H R$ is directly related to the QR factorize of A