LA2 7

KYB

Thrn, it's a Fact

mathrnfact@gmail.com

October 30, 2020

Overview

The LU factorization

Observation

Consider a linear system $Ax = b$, where $A \in \mathbb{R}^{n\times n}$ is a nonsingular upper triangular matrix, i.e.

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ 0 & A_{22} & \cdots & A_{2n} \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

This system can be solved by

$$\begin{aligned} x_n &= b_n / A_{nn} \\ x_k &= \left(b_k - \sum_{i=k+1}^{n} A_{ki} x_i \right) / A_{kk}. \end{aligned}$$

Observation

If A is a nonsingular lower triangular matrix,

$$\begin{bmatrix} A_{11} & 0 & \cdots & 0 \\ A_{21} & A_{22} & \cdots & 0 \\ \vdots & & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

$$\begin{aligned} x_1 &= b_1 / A_{11} \\ x_k &= \left(b_k - \sum_{i=1}^{k-1} A_{ki} x_i \right) / A_{kk}. \end{aligned}$$

Observation

Hence if $A = LU$ where L is a nonsingular lower triangular matrix and U is a nonsingular upper triangular matrix, the system $Ax = b$ can be solved by the simple algorithm.
Moreover, if the diagonal entries of L are all 1, we can find the solution by more simple calculation.

The LU factorization

For given $A \in \mathbb{R}^{n \times n}$, if no row interchanges are required when Gaussian elimination is applied, then there is a nonsingular lower triangular matrix L such that $U = L^{-1}A$.

Theorem (378)

Let $L \in \mathbb{R}^{n\times n}$ be lower triangular and invertible. Then L^{-1} is also lower triangular. Similarly, if $U \in \mathbb{R}^{n\times n}$ is upper triangular and invertible, then U^{-1} is also upper triangular.

Theorem (379)

Let $L_1, L_2 \in \mathbb{R}^{n\times n}$ be two lower triangular matrices. Then $L_1 L_2$ is also lower triangular. Similarly, the product of two upper triangular matrices is upper triangular.

Theorem (381)

Let $A \in \mathbb{R}^{n\times n}$. For each $k = 1, 2, \cdots, n-1$, let $M^{(k)} \in \mathbb{R}^{k\times k}$ be the submatrix extracted from the upper left-hand corner of A. If each $M^{(k)}$ is nonsingular, then A has a unique LU factorization.

Proof

$$\left[\begin{array}{c|c} M^{(n-1)} & a \\ \hline b^T & A_{nn} \end{array}\right] = \left[\begin{array}{c|c} L_1 & 0 \\ \hline u^T & 1 \end{array}\right] \left[\begin{array}{c|c} U_1 & v \\ \hline 0 & \alpha \end{array}\right]$$

where $M^{(n-1)} = L_1 U_1$, $L_1 v = a$, $u^T U_1 = b^T$, and $u \cdot v + \alpha = A_{nn}$.

Ex 9.1.1

Find the LU factorization of

$$A = \begin{bmatrix} 1 & 3 & 2 & 4 \\ 3 & 10 & 5 & 15 \\ -1 & -7 & 3 & -17 \\ -2 & -6 & 1 & -12 \end{bmatrix}.$$

Proof

$$\begin{bmatrix} 1 & 3 & 2 & 4 \\ 3 & 10 & 5 & 15 \\ -1 & -7 & 3 & -17 \\ -2 & -6 & 1 & -12 \end{bmatrix} \xrightarrow{L_1} \begin{bmatrix} 1 & 3 & 2 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & -4 & 5 & -13 \\ 0 & 0 & 5 & -4 \end{bmatrix}$$

$$\xrightarrow{L_2} \begin{bmatrix} 1 & 3 & 2 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 5 & -4 \end{bmatrix} \xrightarrow{L_3} \begin{bmatrix} 1 & 3 & 2 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$L = L_1{}^{-1} L_2{}^{-1} L_3{}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -1 & -4 & 1 & 0 \\ -2 & 0 & 5 & 1 \end{bmatrix}$$

Ex 9.1.7

(a) Show that there do not exist matrices of the from

$$L = \begin{bmatrix} 1 & 0 \\ l & 1 \end{bmatrix}, U = \begin{bmatrix} u & v \\ 0 & w \end{bmatrix}$$

such that $LU = A$, where

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

(b) Let

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Show that A has infinitely many LU factorizations.

Ex 9.1.10

Show that, if L is $n \times n$ and unit lower triangular, then the cost of solving $Lc = b$ is $n^2 - n$ arithmetic operations.

Proof

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ldots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Then for $k = 2, \cdots, n$, $x_k = b_n - \sum_{i=1}^{k-1} l_{ki} x_i$. So to compute x_k, need $k-1$ muliplications and $k-1$ subtractions. Thus $\sum_{k=2}^{n} 2(k-1) = \sum_{k=1}^{n-1} 2k = 2\frac{(n-1)n}{2} = n^2 - n$ operations are needed.

Partial pivoting

Floating number in Python

```
import sys
print(sys.float_info)

Result : sys.float_info(max=1.7976931348623157e+308,
max_exp=1024, max_10_exp=308, min=2.2250738585072014e-308,
min_exp=-1021, min_10_exp=-307, dig=15, mant_dig=53,
epsilon=2.220446049250313e-16, radix=2, rounds=1)
```

A round-off error

If $\epsilon > 0$ is sufficiently small, a computer system treats ϵ as 0. For example,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 10^{-10} \end{bmatrix}$$

is theorically nonsingular. In practice, however, $\det(A) = 10^{-10} \approx 0$. So a computer may think A is singular.

Partial pivoting

Consider

$$\begin{bmatrix} 10^{-5} & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

This system has the exact unique root $(1-10^{-5}, 1-\frac{1}{10^5-1}) \approx (1,1)$.
If we take a row reduction algorithm,

$$\left[\begin{array}{cc|c} 10^{-5} & 1 & 1 \\ 1 & 1 & 2 \end{array}\right] \to \left[\begin{array}{cc|c} 10^{-5} & 1 & 1 \\ 0 & 1-10^5 & 2-10^5 \end{array}\right]$$

$$\xrightarrow{\text{round off}} \left[\begin{array}{cc|c} 10^{-5} & 1 & 1 \\ 0 & -10^5 & -10^5 \end{array}\right] \to \left[\begin{array}{cc|c} 10^{-5} & 1 & 1 \\ 0 & 1 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{cc|c} 10^{-5} & 0 & 0 \\ 0 & 1 & 1 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 1 \end{array}\right]$$

So a practical solution is $(0,1)$.

Partial pivoting

This situation happens because we divide 1 by 10^{-5}. To avoid this situation, for each row reduction step, pivot rows so that the absolute value of A_{11} is maximum over A_1.

$$\left[\begin{array}{cc|c} 10^{-5} & 1 & 1 \\ 1 & 1 & 2 \end{array}\right] \xrightarrow{\text{pivot}} \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 10^{-5} & 1 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1-10^{-5} & 1-2\cdot 10^{-5} \end{array}\right] \xrightarrow{\text{round off}} \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right]$$

Then we get another practical solution $(1, 1)$.

Ex 9.2.1

$$A = \begin{bmatrix} -2.30 & 14.40 & 8.00 \\ 1.20 & -3.50 & 9.40 \\ 3.10 & 6.20 & -9.90 \end{bmatrix}, b = \begin{bmatrix} 1.80 \\ -10.50 \\ 22.30 \end{bmatrix}.$$

Proof

$$\left[\begin{array}{ccc|c} -2.30 & 14.40 & 8.00 & 1.80 \\ 1.20 & -3.50 & 9.40 & -10.50 \\ 3.10 & 6.20 & -9.90 & 22.30 \end{array}\right] \xrightarrow{\text{pivot}} \left[\begin{array}{ccc|c} 3.10 & 6.20 & -9.90 & 22.30 \\ -2.30 & 14.40 & 8.00 & 1.80 \\ 1.20 & -3.50 & 9.40 & -10.50 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|c} 3.10 & 6.20 & -9.90 & 22.30 \\ 0 & 19.00 & 0.66 & 18.25 \\ 0 & -5.90 & 13.23 & -19.13 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 3.10 & 6.20 & -9.90 & 22.30 \\ 0 & 19.00 & 0.66 & 18.25 \\ 0 & 0 & 13.44 & -13.44 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|c} 3.10 & 6.20 & -9.90 & 22.30 \\ 0 & 19.00 & 0.66 & 18.25 \\ 0 & 0 & 1.00 & -1.00 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 3.10 & 6.20 & 0 & 12.40 \\ 0 & 19.00 & 0 & 19.0 \\ 0 & 0 & 1.00 & -1.00 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 3.10 & 6.20 & 0 & 12.40 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1.00 & -1.00 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 3.10 & 0 & 0 & 6.20 \\ 0 & 1.00 & 0 & 1.00 \\ 0 & 0 & 1.00 & -1.00 \end{array}\right]$$

Hence the solution is $(2.00, 1.00, -1.00)$

Ex 9.2.5

Suppose $A \in \mathbb{R}^{n \times n}$ has an LU decomposition. Prove that the product of the diagonal entries of U equals the product of the eigenvalues of A.

Ex 9.2.6

Suppose $\{e_1, \cdots, e_n\}$ is the standard basis for $\mathbb{R}^n$ and $(i_1, \cdots, i_n)$ is a permutation of $(1, \cdots, n)$. Let P be the permutation matrix with rows $e_{i_1}, \cdots, e_{i_n}$.

(a) Let A be an $n \times m$ matrix. Prove that if the rows of A are $r_1, \cdots, r_n$, then the rows of PA are $r_{i_1}, \cdots, r_{i_n}$.

(b) Let A be an $m \times n$ matrix. Prove that if the columns of A are $A_1, \cdots, A_n$, then the rows of AP^T are $A_{i_1}, \cdots, A_{i_n}$.

The Cholesky factorization

Lemma (386)

Let $A \in \mathbb{R}^{n\times n}$ be symmetric and positive definite. Then all the diagonal entries of A are positive. Moreover, the largest entry in magnitude of A lies on the diagonal.

Lemma (387)

If a step of Gaussian elimination is applied to an SPD matrix $A \in \mathbb{R}^{n\times n}$, the result has the from

$$A^{(2)} = \left[\begin{array}{c|c} A_{11} & a^T \\ \hline 0 & B \end{array}\right],$$

where $B \in \mathbb{R}^{(n-1)\times(n-1)}$ is also SPD and $a \in \mathbb{R}^{n-1}$. Moreover,

$$\max\{|B_{ij}|\} \leq \max\{|A_{ij}|\}.$$

Proof

For $i, j > 1$,

$$A_{ij}^{(2)} = A_{ij} - \frac{A_{i1}}{A_{11}} A_{1j}.$$

Therefore,

$$A_{ij}^{(2)} = A_{ji}^{(2)},$$

or B is symmetric.

Proof

Let $x \in \mathbb{R}^{n-1}$. Write $y = \left[\begin{array}{c} y_1 \\ \hline x \end{array}\right]$ for some $y_1 \in \mathbb{R}$.

$$
\begin{aligned}
x \cdot Bx &= \sum_{i=2}^{n} \sum_{j=2}^{n} A_{ij}^{(2)} y_i y_j \\
&= \sum_{i=2}^{n} \sum_{j=2}^{n} \left(A_{ij} - \frac{A_{i1}}{A_{11}} A_{1j} \right) y_i y_j \\
&= \sum_{i=2}^{n} \sum_{j=2}^{n} A_{ij} y_i y_j - \frac{1}{A_{11}} \sum_{i=2}^{n} \sum_{j=2}^{n} A_{i1} A_{1j} y_i y_j \\
&= \sum_{i=2}^{n} \sum_{j=2}^{n} A_{ij} y_i y_j - \frac{1}{A_{11}} (a \cdot x)^2. \\
y \cdot Ay &= \sum_{i=2}^{n} \sum_{j=2}^{n} A_{ij} y_i y_j + A_{11} y_1^2 + 2(a \cdot x) y_1
\end{aligned}
$$

Proof

If we can choose y_1 so that

$$A_{11}y_1^2 + 2(a \cdot x)y_1 = -\frac{1}{A_{11}}(a \cdot x)^2$$

we get $x \cdot Bx = y \cdot Ay$. Thus B is positive definite.

Finally,

$$B_{ii} = A_{ii}^{(2)} = A_{ii} - \frac{A_{i1}}{A_{11}}A_{1i} = A_{ii} - \frac{A_{i1}^2}{A_{11}} \le A_{ii}.$$

Theorem (388)

Let $A \in \mathbb{R}^{n\times n}$ be SPD. Then Gaussian elimination can be performed without partial pivoting. Moreover, the largest entry in any of the intermediate matrices during Gaussian elimination iles in the original matrix A. Finally, in the factorization $A = LU$, all the diagonal entries of U are positive.

Corollary (389)

- *$A = LU$, where L is unit lower triangular and U is upper triangular.*
- *$A = LDL^T$, where L is unit lower triangular and D is diagonal with positive diagonal entries.*
- *$A = R^T R$(the Cholesky factorization), where R is upper triangular with positive diagonal entries.*

The Cholesky factorization

Suppose $A \in \mathbb{R}^{n\times n}$ has the Cholesky factorization, $A = R^T R$. By calculating $R^T R$ directly, we get

$$R_{ii} = \sqrt{A_{ii} - \sum_{k=1}^{i-1} R_{ki}^2}$$

$$R_{ij} = \frac{A_{ij} - \sum_{k=1}^{i-1} R_{ki}R_{kj}}{R_{ii}} \text{ for } j = i+1, \cdot, n.$$

$$\begin{bmatrix} & \vdots & \\ \cdots & A_{ij} & \cdots \\ & \vdots & \end{bmatrix} = \begin{bmatrix} R_{11} & 0 & \cdots & 0 \\ R_{12} & R_{22} & \cdots & 0 \\ \vdots & & & \vdots \\ R_{1n} & R_{2n} & \cdots & R_{nn} \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1n} \\ 0 & R_{22} & \cdots & R_{2n} \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & R_{nn} \end{bmatrix}$$

Ex 9.3.1

Let

$$\begin{bmatrix} 1 & -3 & 2 \\ -3 & 13 & -10 \\ 2 & -10 & 12 \end{bmatrix}.$$

(a) Find the Cholesky factorization, $A = R^T R$.

(b) Using R, find the factorization $A = LU$ and $A = LDL^T$.

Proof

$$
\begin{aligned}
\begin{bmatrix} 1 & -3 & 2 \\ -3 & 13 & -10 \\ 2 & -10 & 12 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ -3 & 2 & 0 \\ 2 & -2 & 2 \end{bmatrix} \begin{bmatrix} 1 & -3 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -3 & 2 \\ 0 & 4 & -4 \\ 0 & 0 & 4 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & -3 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}
$$

Ex 9.3.5

Let $A \in \mathbb{R}^{n\times n}$ be SPD. Prove that the Cholesky factorization of A is unique.

Ex 9.3.6

Suppose $A \in \mathbb{R}^{n\times n}$ is symmetric and there exists an upper triangular matrix $R \in \mathbb{R}^{n\times n}$ with positive diagonal entries such that $A = R^T R$. Prove that A is positive definite.

The End