1 February 9th, 2021

1.1 Pivoting LU Decomposition

Example 1.1

Consider $Ax = b$:

$$\begin{bmatrix} 2 & 1 & -1 \\ 4 & 5 & -3 \\ -2 & 5 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ -8 \end{bmatrix}$$

At each step, we have:

1. $$A = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 3 & -1 \\ -1 & 6 & -3 \end{bmatrix}$$

2. $$A = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 3 & -1 \\ -1 & 2 & -1 \end{bmatrix}$$

Thus, we have:

$$L = \begin{bmatrix} 1 & & \\ 2 & 1 & \\ -1 & 2 & 1 \end{bmatrix} \quad U = \begin{bmatrix} 2 & 1 & -1 \\ & 3 & -1 \\ & & -1 \end{bmatrix}$$

Now we solve $Ly = b$ by forward substitution:

$$\begin{bmatrix} 1 & & \\ 2 & 1 & \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ -8 \end{bmatrix} \implies \begin{cases} y_1 = 1 \\ y_2 = -5 \\ y_3 = 3 \end{cases}$$

Then we solve $Ux = y$ by back substitution:

$$\begin{bmatrix} 2 & 1 & -1 \\ & 3 & -1 \\ & & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -5 \\ 3 \end{bmatrix} \implies \begin{cases} y_1 = 1/3 \\ y_2 = -8/3 \\ y_3 = -3 \end{cases}$$

In LU Decomposition, $A(k, k)$ at step k is in the denominator, which means:

1. The procedure cannot continue if $A(k, k) = 0$

   Example 1.2

   $$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

2. If $A(k, k)$ is small, then the computation becomes inaccurate.

   As such, LU Decomposition is unstable unless $A(k, k)$ is large.

Definition 1.3. We call $A(k, k)$ at step k a **pivot entry**.

In order to avoid small pivots, we use "pivoting", where we interchange the row or columns of the matrix. There are many pivoting schemes, for example Row Pivoting, where at each step k, we choose the largest entry in the k column under $A(k, k)$.

Algorithm 1 Row Pivoting

```
for k = 1 : n − 1 do
    Find the max entry in abs. value in A(k : n, k) denoted by i_k
    A(i_k, :) ⟺ A(k, :)                                   ▷ Swap row i_k and k
    A(k + 1 : n, k) = A(k + 1 : n, k)/A(k, k)
    A(k + 1 : n, k + 1 : n) = A(k + 1 : n, k + 1 : n) − A(k + 1 : n, k)A(k, k + 1 : n)
end for
```

Row interchanging (row $i_k \iff$ row k) is equivalent to Left multiplying by a permutation matrix:

$$P_k = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

Thus when we perform the LU decomposition, we have:

$$L_{n-1}P_{n-1}\dots L_2P_2L_1P_1A = U$$

However, this is too complicated, so we will do some simplification. It is easy to check that $P_kP_k = I$, since we are swapping the two rows again, and that $P_k^T = P_k$.

Lemma 1.4

$$P_{n-1}P_{n-2}\dots P_{k+1}L_kP_{k+1}\dots P_{n-2}P_{n-1} = I - (P_{n-1}P_{n-2}\dots P_{k+1}\ell_k)e_k^T$$

Proof. We have:

$$\begin{aligned} L_k &= I - \ell e_k^T \\ P_{k+1}L_kP_{k+1} &= P_{k+1}(I - \ell_k e_k^T)P_{k_1} = I - P_{k+1}\ell_k e_k^T P_{k+1} \\ &= I - (P_{k+1}\ell_k)e_k^T \\ &\vdots \end{aligned}$$

□

If we denote $P_{n-1}P_{n-2}\dots P_{k+1}\ell_k = \tilde{\ell}_k$, we have:

$$\begin{aligned} &L_{n-1}P_{n-1}\dots L_2P_2L_1P_1A \\ &= L_{n-1}(P_{n-1}L_{n-2}P_{n-1})(P_{n-1}P_{n-2}\dots L_2P_2L_1P_1A \\ &= \tilde{L}_{n-1}\tilde{L}_{n-2}\dots\tilde{L}_1P_{n-1}P_{n-2}\dots P_1A. \end{aligned}$$

Denoting $P_{n-1}P_{n-2}\dots P_1 = P$, we have:

$$PA = \tilde{L}_1^{-1}\tilde{L}_2^{-1}\dots\tilde{L}_{n-1}^{-1}U$$

Giving us

$$PA = LU$$

Example 1.5

Using the same example as before, $A = \begin{bmatrix} 2 & 1 & -1 \\ 4 & 5 & -3 \\ -2 & 5 & -2 \end{bmatrix}$ we have:

1. $$\begin{bmatrix} 2 & 1 & -1 \\ 4 & 5 & -3 \\ -2 & 5 & -2 \end{bmatrix} \xrightarrow{\text{row } 1 \leftrightarrow \text{row } 2} \begin{bmatrix} 4 & 5 & -3 \\ 2 & 1 & -1 \\ -2 & 5 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 5 & -3 \\ \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{15}{2} & -\frac{7}{2} \end{bmatrix}$$

2. $$\begin{bmatrix} 4 & 5 & -3 \\ \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{15}{2} & -\frac{7}{2} \end{bmatrix} \xrightarrow{\text{row } 2 \leftrightarrow \text{row } 3} \begin{bmatrix} 4 & 5 & -3 \\ \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{15}{2} & -\frac{7}{2} \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 5 & -3 \\ -\frac{1}{2} & \frac{15}{2} & -\frac{7}{2} \\ \frac{1}{2} & -\frac{1}{5} & -\frac{1}{5} \end{bmatrix}$$

Giving us an output:

$$PA = \begin{bmatrix} 4 & 5 & -3 \\ -\frac{1}{2} & \frac{15}{2} & -\frac{7}{2} \\ \frac{1}{2} & -\frac{1}{5} & -\frac{1}{5} \end{bmatrix}$$

With this, we have:

$$L = \begin{bmatrix} 1 & & \\ -\frac{1}{2} & 1 & \\ \frac{1}{2} & -\frac{1}{5} & 1 \end{bmatrix} \quad U = \begin{bmatrix} 4 & 5 & -3 \\ & \frac{15}{2} & -\frac{7}{2} \\ & & -\frac{1}{5} \end{bmatrix} \quad P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Remark 1.6 — All entries in L have an abs. value smaller than 1.

There are also other pivoting strategy, such as **full pivoting** where you swap the columns as well go get the maximum pivot in the lower submatrix. This would give us:

$$PAQ = LU$$

where Q is also a permutation matrix.

Remark 1.7 — Full pivoting not only rearranges the equation, but it also rearranges the unknowns.

Remark 1.8 — Full pivoting is more stable than row pivoting but it is computationally more expensive.

For any non-singular matrix, we can solve it using pivoting LU decomposition. However, we can consider the structure of the matrix to improve the algorithm.

1.2 LU Decomposition on SPD Matrices

Definition 1.9. $A \in \mathbb{R}^{n\times n}$ is **symmetric positive definite (SPD)** if it is:

- Symmetric: $A = A^T$
- Positive Definite: $x^T Ax > 0$ for all $x \in \mathbb{R}^n$ and $x \neq 0$

Example 1.10

Examples of SPD matrices include:

- Discrete Laplacian is SPD
- Normal equation of Least Squares
- Hessian of strictly convex functions

For general matrices, the LU decomposition is $\frac{2}{3}n^3 + O(n^2)$. For SPD matrices, we can reduce this to $\frac{1}{3}n^3 + O(n^2)$, meaning we can reduce the computation cost by half. In addition, no pivoting is necessary for SPD matrices.

There is a slight modification of LU for SPD matrices called the **Cholesky Decomposition**.

1.3 Cholesky Decomposition

Assume $A \in \mathbb{R}^{n\times n}$ is SPD. Then there exists a decomposition such that:

$$A = LL^T$$

where L is a lower triangular matrix. With Cholesky decomposition, we can solve the linear system:

$$Ax = b \iff LL^Tx = b \iff \begin{cases} Ly = b \\ L^Tx = y \end{cases}$$

Since $A = LL^T$ with:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix} \quad L = L = \begin{bmatrix} \ell_{11} & & & \\ \ell_{21} & \ell_{22} & & \\ \vdots & \vdots & \ddots & \\ \ell_{n1} & \ell_{n2} & & \ell_{nn} \end{bmatrix} \quad L^T = \begin{bmatrix} \ell_{11} & \ell_{21} & \dots & \ell_{n_1} \\ & \ell_{22} & \dots & \ell_{n_2} \\ & & \ddots & \vdots \\ & & & \ell_{nn} \end{bmatrix}$$

We have:

- $a_{11} = \ell_{11}^2 \implies \ell_{11} = \sqrt{a_{11}}$
- $a_{21} = \ell_{21}\ell_{11} \implies \ell_{21} = a_{21}/\ell_{11}$
- $a_{n1} = \ell_{n1}\ell_{11} \implies \ell_{n1} = a_{n1}/\ell_{11}$

To be continued in the next lecture.