$$\boxed{Ax = b}$$

LU Factorization - Algorithm

1. $A = LU$
2. $Ly = b$
3. $Ux = y$

2x2 LU Factorization (simple example)

$$\underbrace{\begin{pmatrix} A_{11} & A_{12} \\ \boxed{A_{21}} & \boxed{A_{22}} \end{pmatrix}} = \underbrace{\begin{pmatrix} 1 & 0 \\ \boxed{L_{21}} & 1 \end{pmatrix}} \underbrace{\begin{pmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{pmatrix}}$$

$$\begin{pmatrix} \boxed{A_{11}} & \boxed{A_{12}} \\ \boxed{A_{21}} & \boxed{A_{22}} \end{pmatrix} = \begin{pmatrix} \boxed{U_{11}} & \boxed{U_{12}} \\ \boxed{L_{21}U_{11}} & \boxed{L_{21}U_{12} + U_{22}} \end{pmatrix}$$

$$A_{21} = L_{21} U_{11} \rightarrow \boxed{L_{21} = \frac{A_{21}}{U_{11}} = \frac{A_{21}}{A_{11}}}$$

$$A_{22} = L_{21} U_{12} + U_{22} \rightarrow \boxed{U_{22} = A_{22} - L_{21} U_{12}}$$

LU Factorization

a_{11}: scalar
$\boldsymbol{a}_{12}$: row vector $(1\times(n-1))$
$\boldsymbol{a}_{21}$: column vector $((n-1)\times 1)$
$\boldsymbol{A}_{22}$: matrix $((n-1)\times(n-1))$

$A_{n\times n}$

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & \boldsymbol{a}_{12} \\ \boldsymbol{a}_{21} & \boldsymbol{A}_{22} \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{l}_{21} & \boldsymbol{L}_{22} \end{pmatrix} \begin{pmatrix} u_{11} & \boldsymbol{u}_{12} \\ \boldsymbol{0} & \boldsymbol{U}_{22} \end{pmatrix}$$

a_{11} $\boldsymbol{a}_{12}$ $\boldsymbol{a}_{21}$ $\boldsymbol{A}_{22}$

row

col

$u_{12} = a_{12}$

$u_{11} = a_{11}$

$$\begin{pmatrix} a_{11} & \boldsymbol{a}_{12} \\ \boldsymbol{a}_{21} & \boldsymbol{A}_{22} \end{pmatrix} = \begin{pmatrix} u_{11} & \boldsymbol{u}_{12} \\ u_{11}\boldsymbol{l}_{21} & \boldsymbol{l}_{21}\boldsymbol{u}_{12} + \boldsymbol{L}_{22}\boldsymbol{U}_{22} \end{pmatrix}$$

column vector

col

$l_{21} = \dfrac{a_{21}}{a_{11}}$ ← scalar

$A_{22} = l_{21} u_{12} + L_{22} U_{22}$

$L_{22} U_{22} = A_{22} - l_{21} u_{12}$

$L_{22} = ?$

$U_{22} = ?$

recursion

LU Factorization

a_{11}: scalar
$\boldsymbol{a}_{12}$: row vector $(1\times(n-1))$
$\boldsymbol{a}_{21}$: column vector $((n-1)\times 1)$
$\boldsymbol{A}_{22}$: matrix $((n-1)\times(n-1))$

a_{11} $\boldsymbol{a}_{12}$

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & \boldsymbol{a}_{12} \\ \boldsymbol{a}_{21} & \boldsymbol{A}_{22} \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \boldsymbol{l}_{21} & \boldsymbol{L}_{22} \end{pmatrix} \begin{pmatrix} u_{11} & \boldsymbol{u}_{12} \\ \mathbf{0} & \boldsymbol{U}_{22} \end{pmatrix}$$

$\boldsymbol{a}_{21}$ $\boldsymbol{A}_{22}$

1) First row of $\boldsymbol{U}$ is the first row of $\boldsymbol{A}$

$$\begin{pmatrix} a_{11} & \boldsymbol{a}_{12} \\ \boldsymbol{a}_{21} & \boldsymbol{A}_{22} \end{pmatrix} = \begin{pmatrix} u_{11} & \boldsymbol{u}_{12} \\ u_{11}\,\boldsymbol{l}_{21} & \boldsymbol{l}_{21}\boldsymbol{u}_{12} + \boldsymbol{L}_{22}\boldsymbol{U}_{22} \end{pmatrix}$$

2) $\boldsymbol{l}_{21} = \dfrac{1}{u_{11}}\boldsymbol{a}_{21}$

First column of $\boldsymbol{L}$ is the first column of $\boldsymbol{A}/\ u_{11}$

3) $\boldsymbol{M} = \boldsymbol{L}_{22}\boldsymbol{U}_{22} = \overbrace{\boldsymbol{A}_{22} - \boldsymbol{l}_{21}\boldsymbol{u}_{12}}^{\text{Known!}}$

Need another factorization!

Example

$$\underline{\underline{M}} = \begin{pmatrix} 2 & 8 & 4 & 1 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 6 & 2 \\ 1 & 3 & 4 & 2 \end{pmatrix}$$

1) First row of $\boldsymbol{U}$ is the first row of $\boldsymbol{A}$

2) First column of $\boldsymbol{L}$ is the first column of $\boldsymbol{A}/\ u_{11}$

3) $\boldsymbol{L}_{22}\boldsymbol{U}_{22} = \boldsymbol{A}_{22} - \underline{\underline{\boldsymbol{l}_{21}\boldsymbol{u}_{12}}}$

u_{12}

$$\underline{\underline{L}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \end{pmatrix} \qquad \underline{\underline{U}} = \begin{pmatrix} 2 & 8 & 4 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

l_{21}

$$\boldsymbol{L}_{22}\boldsymbol{U}_{22} = \boldsymbol{A}_{22} - \boldsymbol{l}_{21}\boldsymbol{u}_{12} = \underline{\underline{\begin{pmatrix} 2 & 3 & 3 \\ 2 & 6 & 2 \\ 3 & 4 & 2 \end{pmatrix}}} - \underline{\underline{\begin{pmatrix} 4 & 2 & 0.5 \\ 4 & 2 & 0.5 \\ 4 & 2 & 0.5 \end{pmatrix}}}$$

$$\boxed{\boldsymbol{M} = \begin{pmatrix} 2 & 8 & 4 & 1 \\ 1 & -2 & 1 & 2.5 \\ 1 & -2 & 4 & 1.5 \\ 1 & -1 & 2 & 1.5 \end{pmatrix}}$$

$$\boldsymbol{M} = \begin{pmatrix} 2 & 8 & 4 & 1 \\ 1 & -2 & 1 & 2.5 \\ 1 & -2 & 4 & 1.5 \\ 1 & -1 & 2 & 1.5 \end{pmatrix}$$

$\frac{-2}{-2} = 1$ $\qquad$ $-\frac{1}{-2} = \frac{1}{2}$

$$\boldsymbol{L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 1 & 0 & 0 \\ 0.5 & 1 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \end{pmatrix} \qquad \boldsymbol{U} = \begin{pmatrix} 2 & 8 & 4 & 1 \\ 0 & -2 & 1 & 2.5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{L}_{22}\boldsymbol{U}_{22} = \boldsymbol{A}_{22} - \boldsymbol{l}_{21}\boldsymbol{u}_{12} = \begin{pmatrix} 4 & 1.5 \\ 2 & 1.5 \end{pmatrix} - \begin{pmatrix} 1 & 2.5 \\ 0.5 & 1.25 \end{pmatrix}$$

$$\boldsymbol{M} = \begin{pmatrix} 2 & 8 & 4 & 1 \\ 1 & -2 & 1 & 2.5 \\ 1 & -2 & 3 & -1 \\ 1 & -1 & 1.5 & 0.25 \end{pmatrix}$$

$$\boldsymbol{M} = \begin{pmatrix} 2 & 8 & 4 & 1 \\ 1 & -2 & 1 & 2.5 \\ 1 & -2 & 3 & -1 \\ 1 & -1 & 1.5 & \boxed{0.25} \end{pmatrix}$$

$$\boldsymbol{L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 1 & 0 & 0 \\ 0.5 & 1 & 1 & 0 \\ 0.5 & 0.5 & 0.5 & 0 \end{pmatrix} \quad \boldsymbol{U} = \begin{pmatrix} 2 & 8 & 4 & 1 \\ 0 & -2 & 1 & 2.5 \\ 0 & 0 & 3 & \boxed{-1} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{L}_{22}\boldsymbol{U}_{22} = \boldsymbol{A}_{22} - \boldsymbol{l}_{21}\boldsymbol{u}_{12} = \boxed{0.25} - (-0.5) = 0.75$$

$$\boldsymbol{L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 1 & 0 & 0 \\ 0.5 & 1 & 1 & 0 \\ 0.5 & 0.5 & 0.5 & 1 \end{pmatrix} \qquad \boldsymbol{U} = \begin{pmatrix} 2 & 8 & 4 & 1 \\ 0 & -2 & 1 & 2.5 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0.75 \end{pmatrix}$$

LU Factorization

a_{11}: scalar
$\boldsymbol{a}_{12}$: row vector $(1\times(n-1))$
$\boldsymbol{a}_{21}$: column vector $((n-1)\times 1)$
$\boldsymbol{A}_{22}$: matrix $((n-1)\times(n-1))$

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & \boldsymbol{a}_{12} \\ \boldsymbol{a}_{21} & \boldsymbol{A}_{22} \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{l}_{21} & \boldsymbol{L}_{22} \end{pmatrix} \begin{pmatrix} u_{11} & \boldsymbol{u}_{12} \\ \boldsymbol{0} & \boldsymbol{U}_{22} \end{pmatrix}$$

a_{11} $\boldsymbol{a}_{12}$ $\boldsymbol{a}_{21}$ $\boldsymbol{A}_{22}$

$$\begin{pmatrix} a_{11} & \boldsymbol{a}_{12} \\ \boldsymbol{a}_{21} & \boldsymbol{A}_{22} \end{pmatrix} = \begin{pmatrix} u_{11} & \boldsymbol{u}_{12} \\ u_{11}\,\boldsymbol{l}_{21} & \boldsymbol{l}_{21}\boldsymbol{u}_{12} + \boldsymbol{L}_{22}\boldsymbol{U}_{22} \end{pmatrix}$$

1) First row of $\boldsymbol{U}$ is the first row of $\boldsymbol{A}$

2) $\boldsymbol{l}_{21} = \frac{1}{u_{11}}\boldsymbol{a}_{21}$

First column of $\boldsymbol{L}$ is the first column of $\boldsymbol{A}/\ u_{11}$

3) $\boldsymbol{M} = \boldsymbol{L}_{22}\boldsymbol{U}_{22} = \overbrace{\boldsymbol{A}_{22} - \boldsymbol{l}_{21}\boldsymbol{u}_{12}}^{\text{Known!}}$

Need another factorization!

Cost of solving linear system of equations

Cost of solving triangular systems

$$x_{n-1} = \frac{b_{n-1} - U_{n-1,n}\, x_n}{U_{n-1,n-1}}$$

$$x_n = b_n / U_{nn}$$

$$x_i = \frac{b_i - \sum_{j=i+1}^{n} U_{ij} x_j}{U_{ii}}, \qquad i = n-1, n-2, \dots, 1$$

	$i=n$	$i=n-1$	$i=n-2$	...	$i=n-(n-1)$ $=1$	TOTAL
#divisions	1	1	1	...	1	n
#multiplications	0	1	2	3 ...	$(n-1)$	$\frac{1}{2}n(n-1)$
#subtractions + additions	0	1	2	3 ...	$(n-1)$	$\frac{1}{2}n(n-1)$

$$\sum_{i=1}^{m} i = \frac{1}{2}m(m+1)$$

$+$

Total $= n + n(n-1)$

Comp. complexity $n \to \infty$ $O(n^2)$

Cost of solving triangular systems

$$x_n = b_n/U_{nn} \qquad x_i = \frac{b_i - \sum_{j=i+1}^{n} U_{ij}x_j}{U_{ii}}, \qquad i = n-1, n-2, \dots, 1$$

n divisions
$n(n-1)/2$ subtractions/additions
$n(n-1)/2$ multiplications

$\Rightarrow$ Computational complexity is $O(n^2)$

$n \to \infty$

$$x_1 = b_1/L_{11} \qquad x_i = \frac{b_i - \sum_{j=1}^{i-1} L_{ij}x_j}{L_{ii}}, \qquad i = 2,3, \dots, n$$

n divisions
$n(n-1)/2$ subtractions/additions
$n(n-1)/2$ multiplications

$\Rightarrow$ Computational complexity is $O(n^2)$

Cost of LU factorization

Side note:

$$\sum_{i=1}^{m} i = \frac{1}{2}m(m+1)$$

$$\sum_{i=1}^{m} i^2 = \frac{1}{6}m(m+1)(2m+1)$$

```
## Algorithm 1
## Factorization using the block-format,
## creating new matrices L and U
## and not modifying A
print("LU factorization using Algorithm 1")
L = np.zeros((n,n))
U = np.zeros((n,n))
M = A.copy()
for i in range(n):
    U[i,i:] = M[i,i:]
    L[i:,i] = M[i:,i]/U[i,i]
    M[i+1:,i+1:] -= np.outer(L[i+1:,i],U[i,i+1:])
```

$(n-1)$

$(n-1)$

$(n-1)^2$

$(n-2)^2$

divi $(n-1)+(n-2)+(n-3)+\cdots+1 = n(n-1)/2$

mult. $(n-1)^2+(n-2)^2+\cdots+1 = \boxed{\frac{n^3}{3}} - \frac{n^2}{2}+\frac{n}{6}$

sub/ad: $(n-1)^2+(n-2)^2+\cdots+1 = \boxed{\frac{n^3}{3}} - \frac{n}{2}+\frac{n}{6}$

TOTAL $n\to\infty$ $\quad \frac{2n^3}{3} \longrightarrow O(n^3)$

Solving linear systems

In general, we can solve a linear system of equations following the steps:

1) Factorize the matrix $\boldsymbol{A}$: $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$ (complexity $O(n^3)$)
2) Solve $\boldsymbol{L}\,\boldsymbol{y} = \boldsymbol{b}$ (complexity $O(n^2)$)
3) Solve $\boldsymbol{U}\,\boldsymbol{x} = \boldsymbol{y}$ (complexity $O(n^2)$)

But why should we decouple the factorization from the actual solve? (Remember from Linear Algebra, Gaussian Elimination does not decouple these two steps…)

① LU factorization : $\sim \frac{2n^3}{3}$ $O(n^3)$ $n \to \infty$

② Cholesky : Factorization $\sim \frac{n^3}{3} \to O(n^3)$

③ Matrix-matrix multiplication $\sim 2n^3 \to O(n^3)$

Example

Let's assume that when solving the system of equations $\boldsymbol{K}\,\boldsymbol{U} = \boldsymbol{F}$, we observe the following:

- When the matrix $\boldsymbol{K}$ has dimensions (100,100), computing the LU factorization takes about 1 second and each solve (forward + backward substitution) takes about 0.01 seconds.

Estimate the total time it will take to find the response $\boldsymbol{U}$ corresponding to 10 different vectors $\boldsymbol{F}$ when the matrix $\boldsymbol{K}$ has dimensions (1000,1000)?

$A)\ \sim 10\ seconds$
$B)\ \sim 10^2\ seconds$
$C)\ \sim 10^3\ seconds$
$D)\ \sim 10^4\ seconds$
$E)\ \sim 10^5\ seconds$

LU Factorization with pivoting

$$\underbrace{\begin{bmatrix} c & 1 \\ -1 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}}_{b}$$

1) We want to solve for x

2) But first we will "construct the problem".

- start with the true solution

$$x_t = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

- set $c = 10^{-1}$ (for example)
- compute matrix-vector multiplication to find $b = Ax_t$

3) Now we can perform the solve $Ax = b$ to find x

4) If "all goes well", $x = x_t$. Is it?

What can go wrong with the previous algorithm for LU factorization?

$$\boldsymbol{M} = \begin{pmatrix} 2 & 8 & 4 & 1 \\ 1 & 4 & 3 & 3 \\ 1 & 2 & 6 & 2 \\ 1 & 3 & 4 & 2 \end{pmatrix} \qquad \boldsymbol{L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \end{pmatrix} \qquad \boldsymbol{U} = \begin{pmatrix} 2 & 8 & 4 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{l}_{21}\boldsymbol{u}_{12} = \begin{pmatrix} 4 & 2 & 0.5 \\ 4 & 2 & 0.5 \\ 4 & 2 & 0.5 \end{pmatrix} \qquad \boldsymbol{M} - \boldsymbol{l}_{21}\boldsymbol{u}_{12} = \begin{pmatrix} 2 & 8 & 4 & 1 \\ 1 & 0 & 1 & 2.5 \\ 1 & -2 & 4 & 1.5 \\ 1 & -1 & 2 & 1.5 \end{pmatrix}$$

The next update for the lower triangular matrix will result in a division by zero! LU factorization fails.

What can we do to get something like an LU factorization?

Pivoting

$A = LU$

$A = PLU$

Approach:

1. Swap rows if there is a zero entry in the diagonal
2. Even better idea: Find the largest entry (by absolute value) and swap it to the top row.

The entry we divide by is called the pivot.

Swapping rows to get a bigger pivot is called (partial) pivoting.

$$\begin{pmatrix} a_{11} & \boldsymbol{a}_{12} \\ \boxed{\boldsymbol{a}_{21}} & \boldsymbol{A}_{22} \end{pmatrix} = \begin{pmatrix} u_{11} & \boldsymbol{u}_{12} \\ u_{11}\,\boldsymbol{l}_{21} & \boldsymbol{l}_{21}\boldsymbol{u}_{12} + \boldsymbol{L}_{22}\boldsymbol{U}_{22} \end{pmatrix}$$

Find the largest entry (in magnitude)