Matrix Decomposition

LU $\quad \underline{A} = \underline{L}\,\underline{U}$

QR $\quad \underline{A} = \underline{Q}\,\underline{R}$

Spectral (square matrices) $\quad \underline{A} = \underline{Q}\,\underline{\Lambda}\,\underline{Q}^T$

SVD $\quad \underline{A} = \underline{U}\,\underline{\Sigma}\,\underline{V}^T$

$$\left.\begin{aligned} &\underline{A} = \underline{L}\,\underline{U} \\ &\underline{A} = \underline{Q}\,\underline{R} \\ &\underline{A} = \underline{Q}\,\underline{\Lambda}\,\underline{Q}^T \\ &\underline{A} = \underline{U}\,\underline{\Sigma}\,\underline{V}^T \end{aligned}\right\} \text{Next few weeks}$$

Motivate by examining row echelon form

Take a generic square matrix $\underline{A}$ and convert to an upper triangular matrix $\underline{U}$

Example:

$$\underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 6 & 10 \\ 3 & 14 & 28 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 3 & 14 & 28 \end{bmatrix}$$

Each step of this process, called Gaussian elimination, can be written as a matrix-matrix product. Here,

$$\underline{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\underline{E}_1 \underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 3 & 14 & 28 \end{bmatrix}$$

All of the elimination matrices are lower triangular

Characteristics of lower triangular matrices:

(1) Multiplication of lower triangular matrices results in a lower triangular matrix, i.e.,

$$\underline{L}_1 \underline{L}_2 = \underline{L}_3$$

(2) Inverse of a lower triangular matrix is lower triangular

$$\underline{L}_1^{-1} = \underline{L}_4$$

Gaussian elimination is nothing but repeated multiplication by elimination matrices, each of which is lower triangular

$$\underbrace{\underline{E}_n \underline{E}_{n-1} \cdots \underline{E}_3 \underline{E}_2 \underline{E}_1}_{\underline{L}^{-1}} \underline{A} = \underline{U}$$

row echelon form (upper triangular)

$\underline{L}^{-1}$ ~ a lower triangular

$\underline{L}^{-1} \sim$ a lower triangular matrix

$\underline{A}\,\underline{x} = \underline{b}$

$\underline{L}\,\underbrace{\underline{U}\,\underline{x}} = \underline{b}$

$\underline{L}^{-1}\underline{A} = \underline{U}$

$\underline{L}\,\underline{y} = \underline{b}$

$\underline{L}\,\underline{L}^{-1}\underline{A} = \underline{L}\,\underline{U}$

$\underline{U}\,\underline{x} = \underline{y}$

$\therefore\ \underline{A} = \underline{L}\,\underline{U} \Leftarrow$ LU decomposition!

(1) The diagonal of $\underline{L}$ must have all 1's

Consider

$$\underline{A} = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}$$

$$\left.\begin{aligned} l_{11}u_{11} &= 1 \\ l_{11}u_{12} &= 2 \\ l_{21}u_{11} &= 3 \\ l_{21}u_{12} + l_{22}u_{22} &= 5 \end{aligned}\right\}$$

6 unknowns, only 4 eqns

Set $l_{11} = 1$, $l_{22} = 1 \Rightarrow$ 6 eqns total

(2) Why do we care?

Solve $\underline{A}\,\underline{x} = \underline{b}$ by finding $\underline{A}^{-1}$

Then $\underline{x} = \underline{A}^{-1}\underline{b}$ ← Very expensive!

Look at $\underline{A}\,\underline{x} = \underline{b}$

$$\underline{A}\,\underline{x} = \underline{L}\,\underline{U}\,\underline{x} = \underline{b}$$

$$\underline{x} = (\underline{L}\,\underline{U})^{-1}\underline{b}$$

$$\underline{x} = \underbrace{\underline{U}^{-1}\underline{L}^{-1}}\underline{b}$$

$\underline{U}^{-1}$ + $\underline{L}^{-1}$ are cheap!
because both are triangular

$$\underline{L}\,\underline{y} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

(1) Compute $\underline{A} = \underline{L}\,\underline{U}$

(2) Solve $\underline{L}\,\underline{y} = \underline{b} \Rightarrow \underline{y} = \underline{L}^{-1}\underline{b}$

(2) Solve $\underline{L}\,\underline{y} = \underline{b} \Rightarrow \underline{y} = \underline{L}^{-1}\underline{b}$

(3) Solve $\underline{U}\,\underline{x} = \underline{y} \Rightarrow \underline{x} = \underline{U}^{-1}\underline{L}^{-1}\underline{b}$

Example: Compute LU of

$$\underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 6 & 10 \\ 3 & 14 & 28 \end{bmatrix}$$

1) Eliminate a_{21}

Multiply $\underline{A}$ by $\underline{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$\underline{E}_1\underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 3 & 14 & 28 \end{bmatrix}$$

2) Eliminate location (3,1)

$$\underline{E}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

$$\Rightarrow \underline{E}_2 \underline{E}_1 \underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 8 & 19 \end{bmatrix}$$

3) Eliminate location (3, 2)

$$\underline{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix}$$

$$\Rightarrow \underline{E}_3 \underline{E}_2 \underline{E}_1 \underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix} = \underline{U}$$

$$\underline{A} = \left(\underline{E}_3 \underline{E}_2 \underline{E}_1\right)^{-1} \underline{U}$$

$$\underline{A} = \underline{E}_1^{-1} \underline{E}_2^{-1} \underline{E}_3^{-1} \underline{U}$$

where

$$\underline{E}_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ +2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\underline{E}_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +3 & 0 & 1 \end{bmatrix}$$

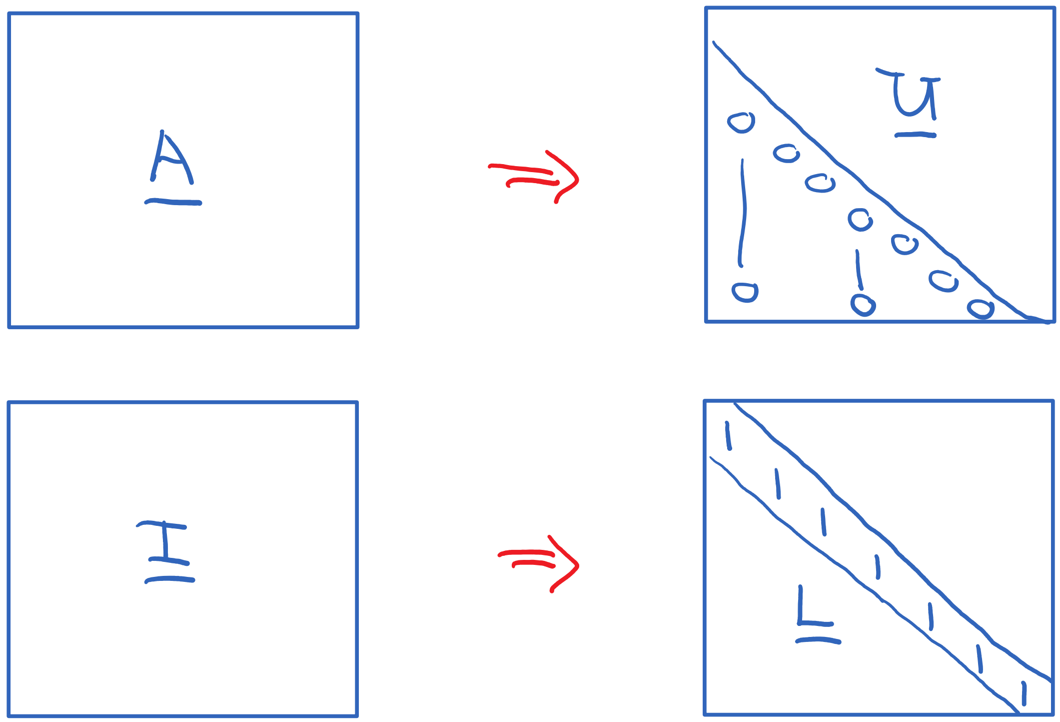
$$\underline{E}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & +4 & 1 \end{bmatrix}$$

$$\therefore \underline{E}_1^{-1}\underline{E}_2^{-1}\underline{E}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix}$$

The negative of the operation, or factor of the value divided by the pivot

$$\underline{A} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix}}_{\underline{L}} \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix}}_{\underline{U}}$$

LU Decomposition Algorithms



LU Pseudocode

Let $\underline{A} \in M_{mn}$ Note: Not necessarily a square matrix

$L = I_{mp}$ with $p = \min(m, n)$

1's on diagonals
0's elsewhere

$U = A$

for $i = 1 : \min(m-1, n)$ i^{th} column loop

for $j = i+1 : m$ pivot j^{th} row loop

$L(j,i) = U(j,i) / U(i,i)$ 1 op } Update

```
L(j,i) = U(j,i)/U(i,i)
U(j,i) = 0
for k = i+1:n
    U(j,k) = U(j,k) - L(j,i) * U(i,k)
end
```

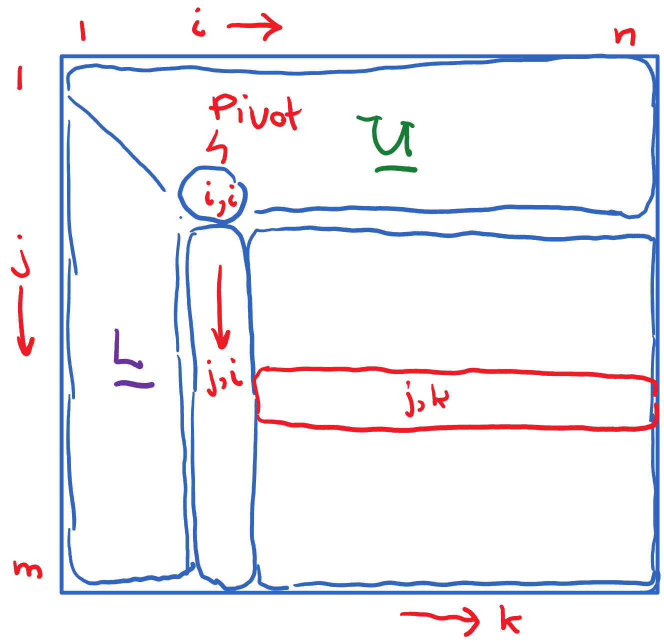
1 op } Update j,i elements

k^{th} column loop

2 ops } Update remainder of j^{th} row

```
  end
end

if m > n
    U = U(1:n,1:n)
end
```



Matlab: $[L,U] = lu(A)$

Next, consider total operation counts in terms of floating point operations (FLOPS)

Let $m = n$ (square matrix)

T: Total operation count

Focus on loops in pseudocode

$$\Rightarrow T = \sum_{i=1}^{n-1}\left(\sum_{j=i+1}^{n}\left(1 + \sum_{k=i+1}^{n} 2\right)\right)$$

$$T = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 1 + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\sum_{k=i+1}^{n} 2$$

$$T = \sum_{i=1}^{n-1}(n-i) + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 2(n-i)$$

$$T = \sum_{i=1}^{n-1}(n-i) + \sum_{i=1}^{n-1} 2(n-i)(n-i)$$

$$T = \sum_{i=1}^{n-1}\left[(n-i) + 2(n-i)(n-i)\right]$$

$$T = \sum_{i=1}^{n-1}\left[2n^2 - 4ni + n + 2i^2 - i\right]$$

However,

$$\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}$$

$$\sum_{i=1}^{n-1} i^2 = \frac{n(n-1)(2n-1)}{6}$$

$$\therefore T = 2n^2(n-1) - 4\frac{n(n-1)}{2}n + n(n-1) + \frac{2n(n-1)(2n-1)}{6} - \frac{(n-1)n}{2}$$

$$T = \frac{2n^3}{3} - \frac{n^2}{2} - \frac{n}{6}$$

Introduce "Big O" notation: A function $f(x)$ is $O(g(x))$, if as $x \to a$ there exists δ and M, such that

$$|f(x)| \leq M|g(x)| \text{ for } |x-a| < \delta$$

For the operation count, this means that as n becomes large, the operation count T becomes dominated by the n^3 term

$$\Rightarrow T = \frac{2n^3}{3} + O(n^2) \quad \text{or } T \text{ is } O(n^3)$$

Thus, in performing LU decomposition,

the computational cost scales as n^3

However, once you have $\underline{A} = \underline{L}\,\underline{U}$,
solving $\underline{A}\,\underline{x} = \underline{b} \Rightarrow \underline{L}\,\underline{U}\,\underline{x} = \underline{b}$ is only $\mathcal{O}(n^2)$

$\therefore$ Factorization is the expensive part of LU

Can anything go wrong?

Failure of Gaussian Elimination

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

Full rank $\rightarrow$ $\underline{A}^{-1}$ exists

$\kappa(\underline{A}) \cong 2.62$

Problem 1: How to eliminate the (2,1) location?

Problem 2: Consider $\underline{A}$ with a slight perturbation

Let $\underline{A} = \begin{bmatrix} 10^{-20} & 1 \\ 1 & 1 \end{bmatrix}$

Exact LU decomposition is

$$\underline{L} = \begin{bmatrix} 1 & 0 \\ 10^{20} & 1 \end{bmatrix}, \quad \underline{U} = \begin{bmatrix} 10^{-20} & 1 \\ 0 & 1-10^{20} \end{bmatrix}$$

However, $1-10^{20}$ cannot be represented exactly in finite (double) precision, i.e.

$$1-10^{20} \sim -10^{20}$$

Then, approximate LU

$$\underline{\tilde{L}} = \underline{L}, \quad \underline{\tilde{U}} = \begin{bmatrix} 10^{-20} & 1 \\ 0 & -10^{20} \end{bmatrix}$$

$$\underline{\tilde{L}}\,\underline{\tilde{U}} = \begin{bmatrix} 10^{-20} & 1 \\ 1 & 0 \end{bmatrix} \neq \underline{A}$$

Now what?

Pivots to the rescue!

A pivot matrix $\underline{P}$ is one that simply swaps rows. (Technically, this is partial pivoting)

$$\underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{\underline{P}} \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}}_{\underline{A}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \underline{U}$$

When partial pivoting is used with LU, then one actually has

$$\underline{P}\,\underline{A} = \underline{L}\,\underline{U}$$

Solving $\underline{A}\,\underline{x} = \underline{b}$

$$\underline{P}\,\underline{A}\,\underline{x} = \underline{P}\,\underline{b}$$

$$\underline{L}\,\underline{U}\,\underline{x} = \underline{P}\,\underline{b}$$

$$\Rightarrow \underline{x} = \underline{U}^{-1}\underline{L}^{-1}(\underline{P}\,\underline{b})$$

Check perturbed case:

$$\underset{\sim}{A} = \begin{bmatrix} 10^{-20} & 1 \\ 1 & 1 \end{bmatrix}, \quad \underset{\sim}{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\underset{\sim}{P}\underset{\sim}{A} = \begin{bmatrix} 1 & 1 \\ 10^{-20} & 1 \end{bmatrix}$$

Let $\underset{\sim}{E}_1 = \begin{bmatrix} 1 & 0 \\ -10^{-20} & 1 \end{bmatrix}$

$$\rightarrow \underset{\sim}{E}_1^{-1} = \begin{bmatrix} 1 & 0 \\ 10^{-20} & 1 \end{bmatrix} = \underset{\sim}{L}$$

$$\underset{\sim}{E}_1 \underset{\sim}{P} \underset{\sim}{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1-10^{-20} \end{bmatrix} = \underset{\sim}{U}$$

$$\tilde{\underset{\sim}{U}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\underset{\sim}{P}\underset{\sim}{A} \simeq \underset{\sim}{L}\tilde{\underset{\sim}{U}} = \begin{bmatrix} 1 & 0 \\ 10^{-20} & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 10^{-20} & 1+10^{-20} \end{bmatrix} \cong \begin{bmatrix} 1 & 1 \\ 10^{-20} & 1 \end{bmatrix}$$

Then $\underline{P}^{-1}\underline{P}\tilde{\underline{A}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 10^{-20} & 1 \end{bmatrix}$

$$= \begin{bmatrix} 10^{-20} & 1 \\ 1 & 1 \end{bmatrix} = \underline{A} \quad \checkmark$$

Perturbed matrix!

Generic LU Pseudocode w/ Partial Pivoting

Assume $\underline{A} \in M_{nn}$ (square matrix for simplicity)

$\underline{L} = \underline{I}$, $\underline{U} = \underline{A}$, $\underline{P} = \underline{I}$, $tol_p = 1e{-8}$

for $k = 1:n-1$

select $i \geq k$ that maximizes $|U_{ik}|$

if $|U_{ik}| > tol_p$ % check for pivot zero

$U(k,k:n) \leftrightarrow U(i,k:n)$ % swap row

```
        L(k,1:k-1) <-> L(i,1:k-1)   % swap row
        P(k,:) <-> P(i,:)           % swap row
        for j = k+1:n
            L(j,k) = U(j,k)/U(k,k)
            U(j,k:n) = U(j,k:n)
                       - L(j,k) U(k,k:n)
        end
    end
end
```

What is operation count? $O(n^3)$

Matlab: `A([i,j],:) = A([j,i],:)` swaps rows i & j in-place

Matlab: `[L,U,P] = lu(A)`

`[L,U] = lu(A)`