MATH 312 LECTURE 3	: LU decomposition.

5] Matrix multiplication.

Q: Why do we multiply matrices the way we do?

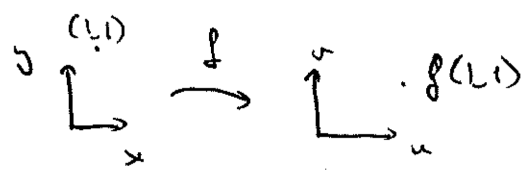
$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix} \to AB = \begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix}, \; BA = \begin{bmatrix} 3 & -1 \\ -1 & -3 \end{bmatrix}$$

But why?

Think of the following system $\left.\begin{array}{l} x - y = u \\ x + y = v \end{array}\right\}$

Here, given a point (x,y) we obtain a new one (u,v).

That is, we have a "formula" or function f:



Ex. $f(1,1) = (1-1, 1+1) = (0,2)$

In general, $(u,v) = f(x,y) = (x-y, x+y)$.

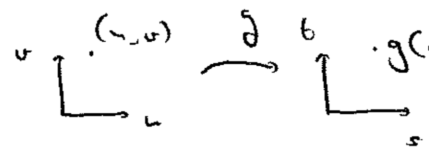
This function is LINEAR:

1) $f(ax, ay) = (ax - ay, ax + ay) = a(x-y, x+y) = a f(x,y)$.

2) $f(x+\tilde{x}, y+\tilde{y}) = f(x,y) + f(\tilde{x},\tilde{y})$

Consider another linear function that takes this (u,v) to a new point (s,t):

$$\left.\begin{aligned} s &= 2u + v \\ t &= u - 2v \end{aligned}\right\}$$



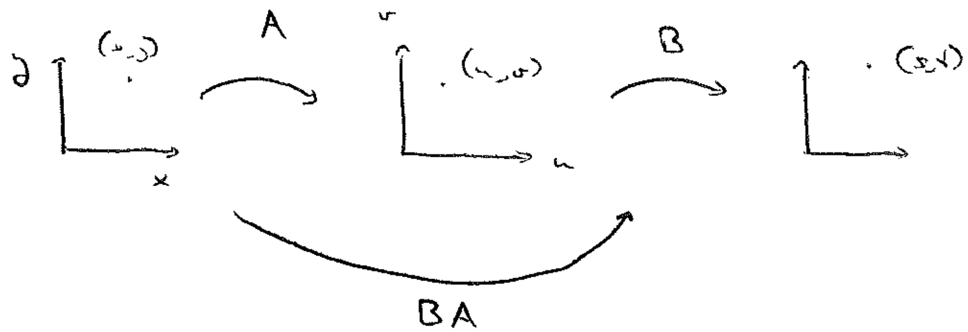
$\cdot\, g(u,v) = (s,t)$.

Question: Given (x,y), what is (s,t)? That is,

$$\left.\begin{aligned} s &= \ldots x + \ldots y \\ t &= \ldots x + \ldots y \end{aligned}\right\}$$

Well,
$$\begin{aligned} s &= 2u + v \xrightarrow{u = x - y} s = 2(x-y) + (x+y) \\ t &= u - 2v \xrightarrow{v = x + y} t = (x-y) - 2(x+y) \end{aligned} \longrightarrow$$

$$\longrightarrow \begin{aligned} s &= 3x - y \\ t &= -x - 3y \end{aligned} \quad \Big\| \quad \begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -1 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$



So we want BA to give this new function.

Matrix multiplication is composition of linear functions!

$$\left.\begin{array}{l} \begin{bmatrix} u \\ v \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix} \\ \begin{bmatrix} s \\ t \end{bmatrix} = B \begin{bmatrix} u \\ v \end{bmatrix} \end{array}\right\} \quad \begin{bmatrix} s \\ t \end{bmatrix} = BA \begin{bmatrix} x \\ y \end{bmatrix} //.$$

• Ways of multiplying matrices:

$$
\begin{array}{c c}
 & \\
m & \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{A}
\end{array}
\begin{array}{c c}
 & p \\
n & \underbrace{\begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{bmatrix}}_{B}
\end{array}
=
\begin{array}{c c}
 & p \\
m & \underbrace{\begin{bmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & & \\ c_{m1} & \cdots & c_{mp} \end{bmatrix}}_{C}
\end{array}
$$

1) ~~Thetes~~, Classic: $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i*} \cdot b_{*j} \equiv$ dot product of row i with column j

2) By columns:

$$A \left[\vec{b}_1 \middle| \cdots \middle| \vec{b}_p \right] = \left[A\vec{b}_1 \middle| \cdots \middle| A\vec{b}_p \right]$$

3) By rows:

$$\begin{bmatrix} \vec{a}_1^T \\ \hline \vdots \\ \hline \vec{a}_m^T \end{bmatrix} B = \begin{bmatrix} \vec{a}_1^T B \\ \hline \vdots \\ \hline \vec{a}_m^T B \end{bmatrix}$$

• Example:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

(so $m=4$, $n=3$, $p=2$)

↳ $C = AB$ is 4×2

Classic way: $AB = \begin{bmatrix} \vec{a}_1^T \vec{b}_1 & \vec{a}_1 \vec{b}_2 \\ \vec{a}_2 \vec{b}_1 & \vec{a}_2 \vec{b}_2 \\ \vdots & \vdots \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 7 \\ 1 & 1 \end{bmatrix}$

By columns:

$$A\vec{b}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}, \quad A\vec{b}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 7 \\ 1 \end{bmatrix}$$

By rows:

$$\vec{a}_1^T B = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \end{bmatrix} \ldots$$

4) Block multiplication:

Example:

$$\left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right] \left[\begin{array}{cc} 1 & 2 \\ 3 & 4 \\ \hline 0 & 1 \\ 3 & 4 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix} =$$

$$= \begin{bmatrix} B_1 + B_2 \\ B_1 + B_2 \end{bmatrix} = \cancel{\begin{bmatrix} 2 & 4 \\ 6 & 8 \\ 0 & 2 \\ 6 & 8 \end{bmatrix}} \left[\begin{array}{cc} 1 & 3 \\ 6 & 8 \\ \hline 1 & 3 \\ 6 & 8 \end{array}\right]$$

in this particular problem $A_{ij} = I$

• Important particular case: Columns × Rows

Let $A = \left[\begin{array}{c|c|c} \vec{a}_1 & \cdots & \vec{a}_n \end{array}\right]$, $B = \left[\begin{array}{c} \vec{b}_1^T \\ \hline \vdots \\ \hline \vec{b}_n^T \end{array}\right]$

If we look at columns $\vec{a}_i$ and rows $\vec{b}_i^T$ as blocks, what is AB?

$$AB = \left[\begin{array}{c|c|c} A_1 & \cdots & A_n \end{array}\right] \left[\begin{array}{c} B_1 \\ \hline \vdots \\ \hline B_n \end{array}\right] = A_1 B_1 + \cdots + A_n B_n =$$

$$= \vec{a}_1 \vec{b}_1^T + \cdots + \vec{a}_n \vec{b}_n^T .$$

Examples:

1) $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}$

$$AB = \left[\begin{array}{c|c} 1 & 2 \\ 3 & 4 \end{array}\right] \left[\begin{array}{ccc} 1 & 0 & 1 \\ \hline 0 & 2 & 1 \end{array}\right] = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 1 \\ 3 & 0 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 4 & 2 \\ 0 & 8 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 3 \\ 3 & 8 & 7 \end{bmatrix}$$

2) $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$AB = \left[\begin{array}{c|c} 1 & 2 \\ 3 & 4 \end{array}\right] \left[\begin{array}{c} x_1 \\ \hline x_2 \end{array}\right] = \begin{bmatrix} 1 \\ 3 \end{bmatrix} x_1 + \begin{bmatrix} 2 \\ 4 \end{bmatrix} x_2$$ linear combination of the columns of A ‖

5) LU decomposition.

Goal: To find a lower triangular matrix L and an upper triangular one U, such that

$A = LU$.

Why?

Say we want to solve $A\vec{x} = \vec{b}$ ~~[illegible]~~ for many different $\vec{b}$'s.

We can do:

1) Apply Gaussian elimination to $A\vec{x} = \vec{b} \rightsquigarrow [A \ \vec{b}]$ each time.

↳ The operations on A are repeated!

2) Find A^{-1}, save it and then do $\vec{x} = A^{-1}\vec{b}$ for each $\vec{b}$.

↳ Finding A^{-1} requires solving n systems (if A is $n \times n$).

↳ We have to multiply $A^{-1}\vec{b}$ each time.

↳ Important: A "sparse" $\not\Rightarrow$ A^{-1} sparse

3) Assume we are given L, U such that $A = LU$.

Then we can solve $A\vec{x} = \vec{b}$ as follows:

Step 1: Denote $\vec{y} = U\vec{x}$, and solve $L\vec{y} = \vec{b} \rightsquigarrow \vec{y}$

This easy and fast! L is lower triangular, so the system is solved by "forwards" substitution.

Step 2: Since now we know $\vec{y}$, solve

$U\vec{x} = \vec{y} \rightsquigarrow \underline{\underline{\vec{x}}}$ | easy too! → substitute backwards.

Conclussion: Once we have L and U, to solve $A\vec{x} = \vec{b}$ we only need to solve two triangular systems (i.e, direct substitution, no gaussian elimination).

Example: Given $A = LU$ with $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & -2 \\ 2 & -3 & 3 \end{bmatrix}$,

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 5/3 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & -3 \\ 0 & 0 & 6 \end{bmatrix},$$

solve $A\vec{x} = \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{\vec{b}}$ without perfoming gaussian elimination.

Sol:

$A\vec{x} = \vec{b} \rightarrow LU\vec{x} = \vec{b}$

- $\vec{y} = U\vec{x}$
- $L\vec{y} = \vec{b} \longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 5/3 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \rightarrow$ $y_1 = 1$ $\longrightarrow y_2 = 1 - y_1 = 0$ $\longrightarrow y_3 = 1 - 2y_1 - \frac{5}{3}y_2 =$

$= 1 - 2 = -1$

So,

$U\vec{x} = \vec{y} \rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & -3 \\ 0 & 0 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \rightarrow$ $\longrightarrow x_1 = 1 - x_2 - x_3 = 1$ $\longrightarrow -3x_2 = 3x_3 \rightarrow x_2 = \frac{1}{6}$ $x_3 = \frac{-1}{6}$

That is,

$$\vec{x} = \begin{bmatrix} 1 \\ 1/6 \\ -1/6 \end{bmatrix} //.$$

- So, how to find L and U?

~~Let's go back to the examp~~

Let's start with an example: $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & -2 \\ 2 & -3 & 3 \end{bmatrix}$

We can do (elementary) row operation until we get to an echelon form (upper triangular):

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & -2 \\ 2 & -3 & 3 \end{bmatrix} \xrightarrow[\cancel{R3' = R3 - 2R1}]{R2' = R2 - R1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & -3 \\ 2 & -3 & 3 \end{bmatrix} \xrightarrow{R3' = R3 - 2R1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & -3 \\ 0 & -5 & 1 \end{bmatrix} \xrightarrow{R3'' = R3' - \frac{-5}{-3}R2'}$$

$$\rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & -3 \\ 0 & 0 & 6 \end{bmatrix}.$$ Which were the corresponding elementary matrices?

$$E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \quad E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{5}{3} & 1 \end{bmatrix}$$

That is, $\| \; E_{32} E_{31} E_{21} A = U \; \|$ $\qquad A \xrightarrow{E_{21}} \xrightarrow{E_{31}} \xrightarrow{E_{32}} U$

Therefore,

$$A = (E_{32} E_{31} E_{21})^{-1} U \rightarrow L = (E_{32} E_{31} E_{21})^{-1} = E_{21}^{-1} E_{31}^{-1} E_{32}^{-1} =$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5/3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 5/3 & 1 \end{bmatrix}$$ Lower triangular

↑ these are the multipliers ‖.

• <u>Remark</u>: 1) E_{ij} is always lower triangular.

2) The inverse of a lower triangular is lower triangular.

3) The product of lower triag. matrices is also a lower triag. matrix.

Q: When does $A = LU$ exist?

Conclusion: If no row exchanges are needed, then any matrix A can be factorised as

$$A = \underbrace{L}_{\substack{\text{lower} \\ \text{triangular} \\ \text{with 1's on} \\ \text{diagonal}}} \; \underbrace{U}_{\substack{\text{upper} \\ \text{triangular}}} \quad , \text{ where}$$

- $U \equiv$ result of doing Gaussian elimination on A.
  (so if A is $m \times n$, so is U).

- $$L = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \\ \vdots & \vdots & \cdots & \cdots & 1 \end{bmatrix}$$ ("substract l_{21} times row1 from row2").

multipliers

$\rightarrow L \equiv$ product of elementary matrices

$\Rightarrow L$ is always invertible

~~Q: Is this factorization unique (when it exists)?~~

~~Answer: It is unique as A is invertible.~~

$\boxed{\text{Chapter 3}}$: Vector Spaces and Subspaces.

We are already used to the n-dimensional Euclidean vector space:

$\mathbb{R}^n = \left\{ \vec{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \text{ with } u_i \in \mathbb{R} \right\}$ with the two standard operations

- Addition of vectors: $\vec{u} + \vec{v} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix}$
- Scaling: $\lambda \vec{u} = \begin{bmatrix} \lambda u_1 \\ \vdots \\ \lambda u_n \end{bmatrix}$ for all $\lambda \in \mathbb{R}$

- <u>Def</u>: A linear combination of the vectors $\vec{u}_1, \dots, \vec{u}_n$ is ~~the~~ a new vector

  $\lambda_1 \vec{u}_1 + \dots + \lambda_n \vec{u}_n$ with $\lambda_1, \dots, \lambda_n$ scalars (for us, $\in \mathbb{R}$).

- <u>Def</u>: General vector space

  A vector space V is a collection of objects (called "vectors") ~~that satisfy~~ together with two operations (that we call "addition" and "scaling") satisfying that

$\rightarrow$ Any linear combination of objects in V also belong to V.

That is, if $\vec{u}, \vec{v} \in V$, then $\alpha \vec{u} + \beta \vec{v} \in V$ for any α, β scalars.

- **Def:** Subspace

A subspace V' of a vector space V is a subcollection of objects from V which itself forms a vector space.

→ In practice: check V is a subspace.

1) Closed under addition: $\left.\begin{array}{l}\vec{u} \in V \\ \vec{v} \in V\end{array}\right\} \Rightarrow \vec{u} + \vec{v} \in V$ (for any $\vec{u}, \vec{v} \in V$)

2) Closed under scaling: $\left.\begin{array}{l}\vec{u} \in V \\ \lambda \text{ scalar}\end{array}\right\} \Rightarrow \lambda \vec{u} \in V$ (for any $\vec{u} \in V$, λ scalar).

- **Example:**

1) Is $V = \left\{ \vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : x_1 x_2 = 0 \right\}$ a subspace of $\mathbb{R}^2$?

2) Is $V = \left\{ \vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : x_1 + x_2 = 1 \right\}$ a subspace of $\mathbb{R}^2$?

3) Is $V = \left\{ \vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : x_1 = 2x_2 \right\}$ " " " " ?

4) Is $V = \left\{ \vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : x_1 \geq 0 \right\}$ " " " " ?

5) Is $V = \mathbb{R}^2$ a subspace of $\mathbb{R}^2$?

6) Is $V = \left\{ \vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : x_1 = 0 = x_2 \right\} = \{ \vec{0} \}$ " " " " ?

7) Is the space of 2×2 matrices M_{22} with usual addition and scaling a vector space?

8) Is the space of 2×2 matrices of the form

$\begin{bmatrix} a & a \\ 0 & a \end{bmatrix}$ a subspace M_{22}?