Northeastern University, Department of Mathematics

MATH 4570 - Matrix Methods in Data Analysis and Machine Learning

• Instructor: **He Wang** Email: **he.wang@northeastern.edu**

§2. Matrix Algebra and matrix factorizations.

2331 2341

CONTENTS

$\mathbb{F}$ any field $+: \mathbb{F}^{m\times n} \times \mathbb{F}^{m\times n} \longrightarrow \mathbb{F}^{m\times n}$

1. Sum and scalar product

Definition 1. • The ***sum*** $A + B$ of $m \times n$ matrices A and B is

$[a_{ij}]$

$$A+B = \left[a_{ij}+b_{ij}\right]$$

• The ***scalar product*** $r \cdot A$ of a scalar $r \in \mathbb{F}$ and A is

$$rA = \left[ra_{ij}\right]$$

$\mathbb{F}^{m\times n} := \{\text{all } m\times n \text{ matrices}\}$

$$\mathbb{F} \times \mathbb{F}^{m\times n} \longrightarrow \mathbb{F}^{m\times n}$$

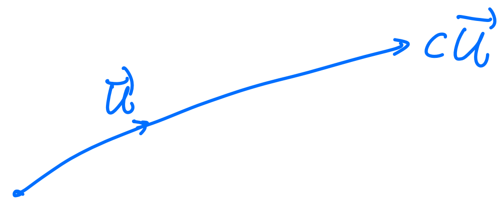
$$0 = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ 0 & 0 & \ddots & \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{n\times m}$$

Theorem 2. *For $n \times m$ matrices A, B, C and scalar r, s, the following hold.*

(1) $A + B = B + A$;
(2) $(A + B) + C = A + (B + C)$;
(3) $A + 0 = A$;
(4) $A + (-A) = 0$;
(5) $r(A + B) = rA + rB$;
(6) $(r + s)A = rA + sA$;
(7) $r(sA) = (rs)A$;
(8) $1A = A$.

$(\mathbb{F}^{n\times m}, +)$ abelian group

Geometric meanings of vectors:



Definition 3. A vector $\vec{b}$ in $\mathbb{F}^{m\times 1}$ is called ***linear combination*** of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ in $\mathbb{F}^m$ if

$$\vec{b} = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n$$

2. Matrix Product

- **Product of a matrix A and a vector $\vec{x} \in \mathbb{F}^n$.**

Definition 4. The ***product*** of A and $\vec{x}$ defined to be

$$A\vec{x} = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n.$$

The product of A and $\vec{x}$ can be computed as

$$A\vec{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}$$

Theorem 5 (**Algebraic Rules for $A\vec{x}$**)**.** *If A is an $m \times n$ matrix, $\vec{u}$ and $\vec{v}$ are vectors in $\mathbb{F}^n$ and c is a scalar, then*
(1.) $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$
(2.) $A(c\vec{u}) = c(A\vec{u})$.

Definition 6. The $\boxed{\textbf{dot product}}$ of two vectors

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \text{ and } \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Ex: $\mathbb{F} = \mathbb{R}$

is defined as

$$\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2 + \cdots u_nv_n$$

$$\mathbb{F}^{n\times 1} \times \mathbb{F}^{n\times 1} \longrightarrow \mathbb{F}$$

More generally, we can define the product of two matrices:

Definition 7. Let A be an $m \times n$ matrix and B be a $n \times p$ matrix. Define the ***product*** of A and B, to be the $m \times p$ matrix

$$AB := [A\vec{b}_1 \quad A\vec{b}_2 \quad \dots \quad A\vec{b}_p]$$

• The Row-Column Rule for Computing $A \cdot B$

The (i,j)-th entry of AB is

$$\sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj},$$

which equals the dot product of the i-th row of A with the j-th column of B

Theorem 8 (Properties of Matrix Multiplication)**.** *Let A be an $m \times n$ matrix, and let B and C be matrices for which the indicated operations are defined. Let I_n denote the $n \times n$ identity matrix.*

- $A(BC) = (AB)C.$
- $A(B + C) = AB + AC.$
- $(A + B)C = AC + BC.$
- $r(AB) = (rA)B$ *where r is any scalar.*
- $I_m A = A = AI_n.$

$(\mathbb{R}^{n\times n}, +, \text{product})$ ring

Proof.

$$[A(BC)]_{ij} = \sum_{k=1}^{n} a_{ik}(BC)_{kj} = \sum_{k=1}^{n} a_{ik}\left(\sum_{l=1}^{p} b_{kl}c_{lj}\right) = \sum_{k=1}^{n}\sum_{l=1}^{p} a_{ik}b_{kl}c_{lj}$$

$$[(AB)C]_{ij} = \sum_{l=1}^{p}(AB)_{il}c_{lj} = \sum_{l=1}^{p}\left(\sum_{k=1}^{n} a_{ik}b_{kl}\right)c_{lj} =$$

$$\sum_{l=1}^{p}\sum_{k=1}^{n} a_{ik}b_{kl}c_{lj} = \sum_{k=1}^{n}\sum_{l=1}^{p} a_{ik}b_{kl}c_{lj}$$

So, $A(BC) = (AB)C$. □

Example 9. $AB = \overset{A}{\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}} \overset{B}{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}} \neq BA$ $\Rightarrow \{\mathbb{R}^{n\times n}, \times\}$ not commutative

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 3 & 4 \end{bmatrix} = AC \quad \not\Rightarrow \quad B = C$$

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Definition 10. If A is an $n \times n$ matrix. We define the k-th power of A as

$$A^k = \underbrace{A \cdot A \cdot \cdots \cdot A}_{k \text{ factors}}.$$

Example 11. Calculate X^2, X^3, X^4, ... for the following matrices

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

A^k — transition matrix

B^k — Jorden block

C^k — nil potent matrix

D^k — diagonal

3. Gauss-Jordan Elimination

Go back to matrix $[A \mid \vec{b}]$.

The leftmost nonzero entry of a row is called **leading entry**(or **pivot**).

Definition 12. A matrix is in ***row-echelon form*** **(ref)** if
(1.) All entries in a column below a leading entry are zeros.
(2.) Each row above it contains a leading entry further to the left.
A matrix is in ***reduced row-echelon form*** **(rref)**, if it satisfies (1) (2) and
(3.) The leading entry in each nonzero row is 1.
(4.) All entries in a column above a leading entry are zeros.

Condition 2 implies that all zero rows are at the bottom of the matrix.

One example of **ref**, ($\blacksquare$: non-zero number, $*$ any number) and one example of **rref**

$$\textbf{ref} = \begin{bmatrix} \blacksquare & * & * & * & * & * \\ 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \to \cdots \to \qquad \textbf{rref} = \begin{bmatrix} 1 & * & 0 & 0 & 0 & * \\ 0 & 0 & 1 & 0 & 0 & * \\ 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Examples.

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 1 & 4 & 5 \\ 0 & 0 & 1 & 0 & 5 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 5 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 & 3 & 4 & 5 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 5 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 0 & 4 & 5 \\ 0 & 0 & 1 & 4 & 5 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & 0 & 4 & 0 \\ 0 & 0 & 1 & 7 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

***Elementary Row Operations*:**

(1.) **Scaling:** Multiply a row R_i by a nonzero scalar $k \neq 0$.

(2.) **Replacement:** Replace a row R_i by adding a multiple of another row kR_j.

(3.) **Interchange:** Interchange two rows.

$A \dashrightarrow \mathrm{rref}(A)$

Elementary row operations do not change solutions of the linear system.

> **Theorem 13.** *Using the elementary row operations, one can change a matrix to a reduced row-echelon form.*
>
> unique

Proof. Gauss-Jordan elimination:
1. Begin with the *leftmost* ***nonzero*** column.
2. Select a *nonzero* entry as a **pivot**, and interchange its row to the first row.
3. Use ERO to create zeros in all positions below the pivot.
4. Omit the first row and repeat this process.
5. Repeat the process until the last nonzero row.
6. Scale all pivots to 1's.
7. Beginning with the **rightmost** pivot and working upward and to the left. ◻

2331

$\mathbb{Z}_3 = \{[0],[1],[2]\}$

+	0	1	2
0	0	1	
1	1	2	
2			

$-[1]=[2]$

since $[1]+[2]=[0]$

$\boxed{-[a]} + [a] = [0]$

$A = \begin{bmatrix} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 0 & 2 \end{bmatrix} \xrightarrow{R_3 + (-1)R_2} \begin{bmatrix} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & [0] & 2 & [0] \end{bmatrix}$

$\boxed{\text{over } \mathbb{Z}_3}$

$[2]^{-1} = [2]$

$\xrightarrow{R_3 \cdot [2] \quad R_3/2} \begin{bmatrix} 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

$[2]\cdot[2]=[1]$

$\dashrightarrow \mathrm{rref}(A) = \begin{bmatrix} 1 & 2 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

$A_{\mathbb{Z}_2} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \xrightarrow[R_3 - R_2]{R_1 + R_2} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \dashrightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

over $\mathbb{Z}_2$

Theorem 14. *A matrix A has a unique reduced row echelon form* $\textbf{\textit{rref}}(A)$.

Definition 15. If $A \xrightarrow{ERO} \cdots \xrightarrow{ERO} B$, then A is called **row-equivalent** to B.

Proposition 16. *Row-equivalent is an equivalent relation.*

Proof. 1. (reflexive)

2. (symmetric)

3. (transitive) □

Theorem 17. *A linear system* $[A|\vec{b}]$ *is inconsistent (no solution)* if and only if $\textbf{\textit{rref}}([A|\vec{b}])$ *has a row*

$$[\ 0\ 0\ 0\ \ldots\ 0\ |\ 1\ ].$$

If a linear system is consistent, it has either
- *a unique solution (no free variables), or*
- *infinitely many solutions (at least one free variable).*

Definition 18. The **rank** of a matrix A is

$$\operatorname{rank}(A) = \text{the number of pivots in } \textbf{rref}(A).$$

Proposition 19. *Row-equivalent matrices have the same rank.*

Example 20. Suppose the coefficient matrix A is of size $m \times n$. Then,

1. $\text{rank}(A) \leq m$ and $\text{rank}(A) \leq n$.
2. If the system is inconsistent, then $\text{rank}(A) < m$.
3. If the system has exactly one solution, then $\text{rank}(A) = n$.
4. If the system has infinitely many solutions, then $\text{rank}(A) < n$.

Definition 21. An $m \times n$ matrix A has **full rank**, if $\text{rank}(A) = \min(m, n)$.

Proposition 22. *A linear system with an $n \times n$ coefficient matrix A has exactly one solution if and only if* $\text{rank}(A) = n$ *if and only if* ***rref***$(A) = I_n$.

$\mathbb{Z}_q$ ring

Remark:

1. We can apply Gaussian elimination over any field (including $\mathbb{Z}_p$). p prime
2. We can apply Gaussian elimination over integers $\mathbb{Z}$. However, we can not achieve **rref**.
3. Buchberger's algorithm is a generalization of Gaussian elimination to polynomials to obtain a Grobnear basis in commutative algebra.

$$\begin{pmatrix} x^2+x & 3x+1 & x^4 \\ 3x^4 & \cdot & \cdot \\ - & - & \cdot \end{pmatrix}$$

4. **Inverse of a matrix**

Definition 23. An $n \times n$ matrix A is called **invertible** if there exists an $n \times n$ matrix B such that

$$AB = BA = I_n.$$

Proposition 24. *If A is invertible, then it has only one inverse.*

Theorem 25. *Let A and B be $n \times n$ invertible matrices.*

(1) The matrix A is invertible.
(2) There is a square matrix B such that $BA = I$.
(3) The linear system $A\vec{x} = \vec{0}$ has only the trivial solution.
(4) rank $A = n$.
(5) The reduced row echelon form of A is identity matrix, i.e. ***rref***$(A) = I_n$.
(6) The matrix A is a product of elementary matrices.
(7) There is a square matrix C such that $AC = I_n$.
(8) The linear system $A\vec{x} = \vec{b}$ has a unique solution for every $\vec{b} \in \mathbb{F}^n$.

then

- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^{-1})^{-1} = A$
- $(A^T)^{-1} = (A^{-1})^T$

$$A \xrightarrow{R_1 \leftrightarrow R_2} B$$

Definition 26 (Elementary matrices)**.**

- E_{ij} denotes the matrix obtained by switching the i-th and j-th rows of I_n.

$$I \xrightarrow{R_i \leftrightarrow R_j} E_{ij}$$

- $E_i(c)$ denotes the matrix obtained by multiplying the i-th row by a nonzero c.

$$I \xrightarrow{cR_i} E_i(c)$$

- $E_{ij}(d)$ denotes the matrix adding d times the j-th row to the i-th row.

$$I \xrightarrow{R_i + dR_j} E_{ij}(d)$$

Ex

$$I_3 = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$I_3 \xrightarrow{k \cdot R_2} E_i(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (k \neq 0)$$

$$I_3 \xrightarrow{R_2 + dR_1} E_{21}(d) = \begin{bmatrix} 1 & 0 & 0 \\ d & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Proposition 27 (Elementary matrices multiplications)**.** *Multiply a matrix A with an elementary on the **left** side is equivalent to an elementary row operation is performed on the matrix A.*

(1) $A \xrightarrow{R_i \leftrightarrow R_j} B = E_{ij} A$

(2) $A \xrightarrow{kR_i} B = E_i(k) A$

(3) $A \xrightarrow{R_i + dR_j} B = E_{ij}(d) A$

Example 28. The inverse of the elementary matrices.

$E_{ij}^{-1} = E_{ij}$

$E_i(c)^{-1} = E_i\left(\frac{1}{c}\right)$

$E_{ij}(d)^{-1} = E_{ij}(-d)$

Theorem 29 (The inverse matrix theorem). *Let A be an $n \times n$ matrix. Then the next statements are all equivalent (that is, they are either all true or all false).*
(1) The matrix A is invertible.
(2) There is a square matrix B such that $BA = I$.
(3) The linear system $A\vec{x} = \vec{0}$ has only the trivial solution.
(4) $\operatorname{rank} A = n$.
(5) The reduced row echelon form of A is identity matrix, i.e. ***rref***$(A) = I_n$.
(6) The matrix A is a product of elementary matrices.
(7) There is a square matrix C such that $AC = I_n$.
(8) The linear system $A\vec{x} = \vec{b}$ has a unique solution for every $\vec{b} \in \mathbb{F}^n$.

$A \to \cdots\cdots \to I_n$

$E_? E_? E_? E_? A = I_n$

Theorem 30 (Algorithm for Computing A^{-1})**.** *Given an $n \times n$ matrix A.*

1. Define an $n \times 2n$ "augmented matrix "

$$[A \mid I_n]$$

2. Find ***rref***$[A \mid I_n]$ *using elementary row operations to*

Example. Find the inverse of matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 8 & 2 \end{bmatrix}$

Example 31. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible, find A^{-1}.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

$$\begin{tikzcd} \mathbb{R}^{m\times n} \arrow[d] \\ \boxed{(\ )^T} \arrow[d] \\ \mathbb{R}^{n\times m} \end{tikzcd}$$

5. The transpose A^T

Definition 32. Given an $m \times n$ matrix A, we define the ***transpose matrix***
$A^T = [c_{ij}]$, as $c_{ij} = a_{ji}$.

$$(\ )^T : \mathbb{R}^{m\times n} \longrightarrow \mathbb{R}^{n\times m}$$

Theorem 33 (Properties of Matrix Transposition)**.** *Let A and B be matrices such that the indicated operations are well defined.*

- $(A^T)^T = A.$
- $(A + B)^T = A^T + B^T.$
- $(rA)^T = rA^T$ *for any scalar r.*
- $(AB)^T = B^T A^T.$

Proof. Compare the (i, j)-entry of the matrix.

$$[(AB)^T]_{ij} = [AB]_{ji} = \sum_k a_{jk} b_{ki}$$

$$[B^T A^T]_{ij} = \sum_k [B^T]_{ik} [A^T]_{kj} = \sum_k b_{ki} a_{jk} = \sum_k a_{jk} b_{ki}.$$

□

Theorem 34. *If AB is defined, then* $\operatorname{rank}(AB) \leq \operatorname{rank} A$*.*

Theorem 35. $\boxed{\operatorname{rank}(A) = \operatorname{rank}(A^T).}$

Theorem 36. *If AB is defined, then* $\boxed{\operatorname{rank}(AB)} \leq \min\{\operatorname{rank} A, \operatorname{rank} B\}$*.*

$\boxed{A = PDP^{-1}}$

$A = QR$

$A = U\Sigma V^T$

6. LU factorizations and Gaussian elimination

LU-decomposition is a matrix product version of Gaussian elimination.

Definition 37. An $m \times m$ matrix L with entries l_{ij} is called
- **lower triangular** if $l_{ij} = 0$ whenever $j > i$.
- **unit lower triangular** if it is lower triangular, and $l_{ii} = 1$ for each $i = 1, \ldots, m$.

$$\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}$$

(replacement) elementary row operation

$m \times n$ $\qquad\qquad$ $m \times n$

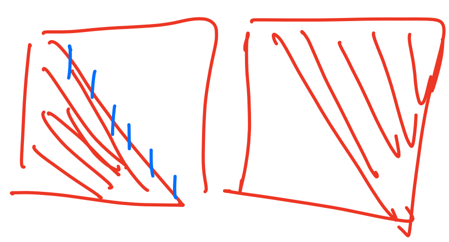
$$A \xrightarrow{O_1} \xrightarrow{O_2} \xrightarrow{O_3} \cdots \xrightarrow{O_s} \mathrm{ref}(A)$$

elementary matrix corresponding to O_1

$$E_s \cdots E_3 E_2 E_1 A = \mathrm{ref}(A)$$

$$A = E_1^{-1} E_2^{-1} \cdots E_s^{-1}\, \mathrm{ref}(A)$$

$$= L\,U =$$



unit lower triangular

- $U = \mathrm{ref}(A)$

- $L = E_1^{-1} E_2^{-1} \cdots E_s^{-1} \cdot I$

$$I \xrightarrow{O_s^{-1}} \cdots \xrightarrow{O_2^{-1}} \xrightarrow{O_1^{-1}} L$$

$$\boxed{A = LU}$$

$$A = \begin{bmatrix} 2 & -1 & 2 \\ -6 & 0 & -2 \\ 8 & -1 & 5 \end{bmatrix} \xrightarrow[R_3 - 4R_1]{R_2 + 3R_1} \begin{bmatrix} 2 & -1 & 2 \\ 0 & -3 & 4 \\ 0 & 3 & -3 \end{bmatrix} \xrightarrow{R_3 + R_2} \begin{bmatrix} 2 & -1 & 2 \\ 0 & -3 & 4 \\ 0 & 0 & 1 \end{bmatrix} = U$$

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_3 - R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \xrightarrow[R_2 - 3R_1]{R_3 + 4R_1} \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -1 & 1 \end{bmatrix} = L$$

Definition 38. Let A be an $m \times n$ matrix. An **LU factorization** for A is given by writing A as the product

$$A = L \cdot U$$

with L a unit lower triangular $m \times m$ matrix, and U an $m \times n$ matrix in **ref**.

Use of LU factorizations:

Solve Q: $A\vec{x} = \vec{b}$ ← Same A different $\vec{b}$

$$L \underbrace{\boxed{U\vec{x}}}_{=\vec{y}} = \vec{b}$$

$\Longleftrightarrow$ Solve $\boxed{\begin{array}{l} L\vec{y} = \vec{b} \\ U\vec{x} = \vec{y} \end{array}}$ ← unique solu.

Algorithm for Finding an LU Factorization:

Suppose A is an $m \times n$ matrix that can be transformed into a matrix in echelon form by using only Row-Replacement operations.

Then an LU factorization of A can be obtained as follows.

1. Reduce A to echelon form U using only Row-Replacement operations.

2. Let L be the matrix obtained from I_m by applying the inverse Row-Replacement operations from Step 1, in reverse order.

$$A = \begin{bmatrix} 2 & -1 & 2 \\ -6 & 0 & -2 \\ 8 & -1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 2 \\ 0 & -3 & 4 \\ 0 & 0 & 1 \end{bmatrix} = LU$$

Remark: There are several variations of LU-factorization: e.g.,

1. LDU-decomposition. $A = LDU$. Here D means a diagonal matrix and U is an unit upper triangular matrix.

$$A = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} = \begin{bmatrix} 1 & & \\ -3 & 1 & \\ 4 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & & \\ & -3 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 1 \\ & 1 & \frac{4}{-3} \\ & & 1 \end{bmatrix}$$

2. LU-factorization (with pivoting) $PA = LU$. Here P is a permutation matrix, obtained by multiplication of elementary matrices E_{ij}.

$$I \xrightarrow{R_i \leftrightarrow R_j} \rightarrow \cdots \xrightarrow{R_k \leftrightarrow R_l} P$$

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

Ex $A = \begin{bmatrix} 10^{-20} & 1 \\ 1 & 2 \end{bmatrix} = LU$