Northeastern University, Department of Mathematics

MATH G5110: Applied Linear Algebra and Matrix Analysis. (Fall 2020)

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

Fix $\mathbb{F}$

e.g.: $\mathbb{F}=\mathbb{R}$ or $\mathbb{F}^{m\times n}$

§2. Matrix Algebra.

$a_{ij}\in\mathbb{F}$

Contents $\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & a_{ij} & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$

$M_{m\times n}(\mathbb{F})$

1. Sum and scalar product

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

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

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

$rA=[ra_{ij}]$

$$\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times n}\xrightarrow{+}\mathbb{R}^{m\times n}$$

$$\mathbb{R}\times\mathbb{R}^{m\times n}\xrightarrow[\text{prod.}]{\text{scalar}}\mathbb{R}^{m\times n}$$

$(\mathbb{R}^{m\times n},+)$ is a commutative group.

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) $1A=A$.

Geometric meanings of vectors: $\vec{u}, \vec{v} \in \mathbb{R}^n$

$\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$

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

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

Definition 4. The **dot product** of two vectors

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

is defined as

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

2. Matrix Product

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

Definition 5. The ***product*** of A and $\vec{x}$ defined to be

$$A\vec{x} = [\vec{a}_1 \;\; \vec{a}_2 \; \ldots \; \vec{a}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n.$$

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

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

Proposition 6. *Let A be an $m \times n$ matrix with columns $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$, and let $\vec{b}$ be a vector in $\mathbb{F}^m$. Then the matrix equation*

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

has the same solution set as the vector equation

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

which has the same solution set as the linear system with augmented matrix

$$\begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \ldots & \vec{a}_n & \vec{b} \end{bmatrix}.$$

$T_A : \mathbb{R}^n \to \mathbb{R}^m$

$\vec{u} \to A\vec{u}$

linear

Theorem 7 (**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.) $A(c\vec{u}) = c(A\vec{u})$.

$T_A(\vec{u}+\vec{v}) = T_A(\vec{u}) + T_A(\vec{v})$

$T_A(c\vec{u}) = cT_A(\vec{v})$

More generally,

Definition 8. 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

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

• **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

Example 9. Calculate AB for $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$.

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \left[A\begin{bmatrix} a \\ b \end{bmatrix} \quad A\begin{bmatrix} c \\ d \end{bmatrix} \right]$$

$$\mathbb{R}^{m\times n} \times \mathbb{R}^{n\times n} \xrightarrow{\times} \mathbb{R}^{m\times n}$$

$(\mathbb{D}^{n\times n}, +, \times, \text{scalar prod})$ $\qquad$ $(\mathbb{D}^{m\times n}, +, \text{scalar prod.})$ $\qquad$ $\begin{pmatrix} a_{11} & a_{12} & \cdots \\ - & - & \end{pmatrix}$ $\qquad$ $\mathbb{D}^{m\times n}$

abelian group

"Associative algebra" vector space

Theorem 10 (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.*
- $I_m A = A = A I_n$.

$(\mathbb{R}^{n\times n}, \times)$ group?

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

NOT commutative algebra!

$\mathbb{R}[x_1, x_2]$

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, $A(BC) = (AB)C$. $\square$

Example 11. $AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \neq BA$

"Lie Algebras"

$[A, B] := AB - BA$

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

$A^T A \vec{x} = A^T \vec{b}$

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \not\Rightarrow \quad A = 0 \text{ or } B = 0$$

Definition 12. If A is an $n \times n$ matrix. We define the k-th power of A as

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

$A = PDP^{-1}$

$A^k = PD^kP^{-1}$

Example 13. Calculate X^2, X^3, X^4, ... for the following matrices

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

Definition 14 (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.

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

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

$$\xrightarrow{kR_2} E_2(k) = \begin{bmatrix} 1 & & \\ & k & \\ & & 1 \end{bmatrix}$$

$$\xrightarrow{R_2 + kR_1} E_{21}(k) = \begin{bmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Proposition 15 (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.*

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

$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$

Product of block matrices.

If $A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$ and $B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$, then

$$AB = \begin{bmatrix} A_1B_1 + A_2B_3 & A_1B_2 + A_2B_4 \\ A_3B_1 + \cancel{A_4}B_3 & A_3B_2 + \cancel{A_4}B_4 \end{bmatrix}$$

3. Inverse of a matrix

Definition 16. 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 17. *If A is invertible, then it has only one inverse.*

$$\mathbb{R}^{n\times n} \xrightarrow{(\ )^{-1}} \mathbb{R}^{n\times n}$$

Theorem 18. *Let A and B be $n \times n$ invertible matrices.*

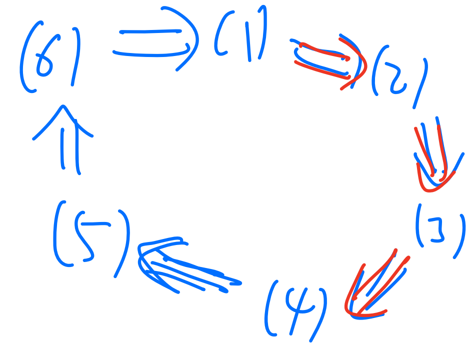
- A^{-1} is invertible and $(A^{-1})^{-1} = A$
- $k \neq 0$ (kA) is invertible and $(kA)^{-1} = \frac{1}{k}A^{-1}$
- (AB) is invertible and $(AB)^{-1} = B^{-1}A^{-1}$
- A^m is invertible, and $(A^m)^{-1} = (A^{-1})^m$
-
-

Example 19. The inverse of the elementary matrices.

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

$E_i(c)^{-1} =$

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



Theorem 20 (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. $\Leftrightarrow AB = BA = I$
(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. $A \xrightarrow{E_1} \cdots \xrightarrow{E_k} I$
(4) rank $A = n$*.*
(5) The reduced row echelon form of A is identity matrix, i.e. ***rref***$(A) = I_n$ $= E_k \cdots E_1 A$
(6) The matrix A is a product of elementary matrices. $A = E_1^{-1} \cdots E_k^{-1}$
(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 $\vec{b} \in \mathbb{F}^n$.

Proof. (1)⇒(2)

(2)⇒(3)

(3)⇒(4)

(4)⇒(5)

(5)⇒(6)

(6)⇒(1)

□

Theorem 21 (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 22. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible, find A^{-1}.

4. The transpose A^T

Definition 23. Given an $m \times n$ matrix A, we define the ***transpose matrix*** $A^T = [c_{ij}]$, as $c_{ij} = a_{ji}$.

Theorem 24 (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 25. *If AB is defined, then* $\operatorname{rank}(AB) \leq \operatorname{rank} A$.

Theorem 26. $\operatorname{rank}(A) = \operatorname{rank}(A^T)$.

Theorem 27. *If AB is defined, then* $\operatorname{rank}(AB) \leq \min\{\operatorname{rank} A, \operatorname{rank} B\}$.

5. LU factorizations and Gaussian elimination

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

Definition 28. 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 $i = 1, \ldots, m$.

Definition 29. 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:

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.

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.

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