모두의연구소

수포자도 도전해 볼 만한

Mathematics in DeepLearning

Lecture1. Elementary Linear Algebra

Juhee Lee Ph.D.
Research Professor @ ewha womans university

Numbers and arrays

Scalar ： a

Vector ： $\boldsymbol{a}$

Matrix ： $\boldsymbol{A}$

Tensor ： $\boldsymbol{\mathsf{A}}$

Definition A vector $\boldsymbol{a}$ is an array of numbers

$$\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$$

a_1: the first element of $\boldsymbol{a}$

a_{i} : the i-th element of $\boldsymbol{a}$

Quiz) what is the A and B? $\mathrm{A} \leq i \leq B$

row vector $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$

column vector $\boldsymbol{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$

(row vector)^{T} = column vector

Operations of vectors : addition

$$\begin{aligned}\boldsymbol{a}+\boldsymbol{b} &= (a_1, a_2, \cdots, a_n) + (b_1, b_2, \cdots, b_n) \\ &= (a_1 + b_1,\ a_2 + b_2, \cdots, a_n + b_n)\end{aligned}$$

$$\begin{aligned}\boldsymbol{a}-\boldsymbol{b} &= (a_1, a_2, \cdots, a_n) - (b_1, b_2, \cdots, b_n) \\ &= (a_1 - b_1,\ a_2 - b_2, \cdots, a_n - b_n)\end{aligned}$$

Operations of vectors : multiplications

$$\mathrm{c}\boldsymbol{a} = (ca_1, ca_2, \cdots, ca_n)$$

$$\boldsymbol{a} \odot \boldsymbol{b} = (a_1, a_2, \cdots, a_n) \odot (b_1, b_2, \cdots, b_n)$$
$$= (a_1 b_1,\ a_2 b_2, \cdots, a_n b_n)$$

$$\boldsymbol{a} \cdot \boldsymbol{b} = (a_1, a_2, \cdots, a_n) \cdot (b_1, b_2, \cdots, b_n)$$
$$= (a_1 b_1 +\ a_2 b_2 + \cdots + a_n b_n)$$

Operations of vectors : multiplications

$$\boldsymbol{a} \times \boldsymbol{b} = (a_1, a_2, a_3) \times (b_1, b_2, b_3)$$
$$= (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1)$$

<u>Definition</u> A matrix **A** is a 2 D array of numbers

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{\mathrm{m}1} & \cdots & a_{\mathrm{m}n} \end{pmatrix} = \left(a_{\mathrm{i}j}\right)_{\mathrm{m}\times n}$$

$$= \begin{pmatrix} a_{11} & a_{1j} & a_{1n} \\ \vdots & a_{\mathrm{i}j} & \vdots \\ a_{\mathrm{m}1} & a_{\mathrm{m}j} & a_{\mathrm{m}n} \end{pmatrix} = \left(\boldsymbol{a_1}\, \boldsymbol{a_j}\, \boldsymbol{a_n}\right)$$

$$= \begin{pmatrix} \boldsymbol{a_1} \\ \boldsymbol{a_i} \\ \boldsymbol{a_m} \end{pmatrix}$$

Operations of matrices : addition, transpose

$$\mathbf{A}+\mathbf{B}=(a_{\mathrm{i}j})_{\mathrm{m}\times n}+(b_{\mathrm{i}j})_{\mathrm{m}\times n}$$
$$=(a_{\mathrm{i}j}+b_{\mathrm{i}j})_{\mathrm{m}\times n}$$

$$\mathbf{A}=(a_{\mathrm{i}j})_{\mathrm{m}\times n}\Rightarrow\ \mathbf{A}^{\mathrm{T}}=(a_{\mathrm{j}i})_{\mathrm{n}\times m}$$

Operations of matrices : multiplications

$$\mathrm{c}\mathbf{A} = \mathrm{c}\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{\mathrm{m}1} & \cdots & a_{\mathrm{m}n} \end{pmatrix} = \begin{pmatrix} ca_{11} & \cdots & ca_{1n} \\ \vdots & \ddots & \vdots \\ \boldsymbol{c}a_{\mathrm{m}1} & \cdots & ca_{\mathrm{mn}} \end{pmatrix}$$

$$\mathbf{A} \odot \mathbf{B} = \left(a_{\mathrm{i}j}\right)_{\mathrm{m}\times n} \odot \left(b_{\mathrm{i}j}\right)_{\mathrm{m}\times n}$$

$$= \begin{pmatrix} a_{11}\mathrm{b}_{11} & \cdots & a_{1n}b_{1n} \\ \vdots & \ddots & \vdots \\ a_{\mathrm{m}1}b_{\mathrm{m}1} & \cdots & a_{\mathrm{m}n}b_{\mathrm{mn}} \end{pmatrix}$$

Operations of matrices : multiplications

$$\mathbf{AB} = (a_{ij})_{m\times p} \quad (b_{ij})_{p\times n} = (c_{ij})_{m\times n}$$

$$= \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mp} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{p1} & \cdots & b_{pn} \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & c_{ij} & \vdots \\ c_{m1} & \cdots & c_{mn} \end{pmatrix}$$

i-th row

j-th column

$$= (c_{ij})_{m\times n} = (\boldsymbol{a}_i^T \cdot \boldsymbol{b}_j)_{m\times n}$$

Tensor

An array with more than two axes

$\mathbf{A} = (a_{ijk})_{m \times n \times p}$

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{\mathrm{m}1} & \cdots & a_{\mathrm{m}n} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & \cdots & b_{1q} \\ \vdots & \ddots & \vdots \\ b_{\mathrm{p}1} & \cdots & b_{\mathrm{pq}} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{\mathrm{m}1}\mathbf{B} & \cdots & a_{\mathrm{m}n}\mathbf{B} \end{pmatrix}$$

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{\mathrm{m}1}\mathbf{B} & \cdots & a_{\mathrm{m}n}\mathbf{B} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} \\ a_{\mathrm{m}1} \\ a_{1\,n} \\ a_{\mathrm{m}n} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & \cdots & b_{1q} \\ \vdots & \ddots & \vdots \\ b_{\mathrm{p}1} & \cdots & b_{\mathrm{pq}} \end{pmatrix}$$

$$= \left(a_{\mathrm{i}jk}\right)_{\mathrm{mn}\times\, p \times \mathrm{q}}$$

Note.

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$$
$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$$

$$\boldsymbol{a}^{\boldsymbol{T}}\boldsymbol{b} = \boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{a}$$

$$\left(\mathbf{A}^{\boldsymbol{T}}\right)^{\boldsymbol{T}} = \mathbf{A}$$
$$(\mathbf{AB})^{\boldsymbol{T}} = \mathbf{B}^{\boldsymbol{T}}\mathbf{A}^{\boldsymbol{T}}$$
$$(\mathbf{A}+\mathbf{B})^{\boldsymbol{T}} = \mathbf{A}^{\mathbf{T}}+\mathbf{B}^{\mathbf{T}}$$

Identity & Inverse Matrices

Definition $\mathbf{I_n} \in R^{n\times n}$ is the identity matrix

$$\mathbf{I_n} = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}$$ identity matrix of n

The matrix inverse of $\mathbf{A}$ is denoted as $\mathbf{A^{-1}}$ and is defined as the matrix such as

$\mathbf{AB} = \mathbf{I}_{\boldsymbol{n}} = \mathbf{B}\boldsymbol{A}$. Then $\mathbf{B} = \mathbf{A^{-1}}$

A linear system $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ is defined as

$$\sum_{\mathrm{j}=1}^{n} a_{ij} x_j = b_i \,,\ \ i \in \{1, \cdots, n\}$$

If there exist $\mathbf{A}^{-\mathbf{1}}$, $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ can be solved

$$\boldsymbol{x} = \mathbf{A}^{-\mathbf{1}}\boldsymbol{b}$$

For there exist $\mathbf{A}^{-1}$,

we can check the determinant of the matrix.

$$\det(\mathbf{A}) \neq 0 \Rightarrow \exists \mathbf{A}^{-1}$$
$$\det(\mathbf{A}) = 0 \Rightarrow \nexists \mathbf{A}^{-1}$$

Find $\mathbf{C}^{-\mathbf{1}}$ using $\det(\mathbf{C})$

$$\mathbf{C} = \begin{pmatrix} -1 & 3 & 2 \\ 0 & -2 & 1 \\ 1 & 0 & -2 \end{pmatrix}$$

Norm

Definition Norm is a function. We usually measure the size of vectors using the function

$$\|a\|_p = \left(\sum_i |x_i|^p\right)^{1/p} \quad p \geq 1$$

A norm is any function f such that

- $f(x) = 0 \Longrightarrow x = 0$
- $f(x + y) \leq f(x) + f(y)$
- $\forall c \in \mathbb{R}, f(ax) = |a| f(x)$

<u>Definition</u> Norm is a function. We usually measure the size of matrix using the function

$$\|\mathbf{A}\|_p = \Big(\sum_i |x_i|^p\Big)^{1/p} \qquad p \geq 1$$

A norm is any function f such that

- $f(x) = 0 \Longrightarrow x = 0$
- $f(x + y) \leq f(x) + f(y)$
- $\forall c \in \mathbb{R}, f(ax) = |a| f(x)$

Note.

$\mathbf{D} = (\mathbf{d_{ii}})_{\mathbf{n}\times\boldsymbol{n}}$: diagonal matrix

$\mathbf{A^T} = \mathbf{A}$: symmetric matrix

$(\mathbf{A}\mathbf{A}^{\boldsymbol{T}}) \quad = \mathbf{I}_{\boldsymbol{n}}$ $\boldsymbol{or}$ $\mathbf{A^T} = \mathbf{A^{-1}}$: Orthogonal matrix

$\mathbf{trA} = \sum_{\boldsymbol{i}=\mathbf{1}} \boldsymbol{a}_{\boldsymbol{ii}}$

Matrix decomposion

LU, QR, …, decomposition

Spectral decomposition

Single Value Decomposition (SVD)

Q & A