2.2 Matrices

행렬은 선형대수학에서 매우 중요한 역할을 합니다.

연립선형방정식을 나타내는 방법 또는 **linear mapping**을 표현하기도 합니다. (ch.2.7)

이번 장에서는 간단하게 행렬의 정의, 행렬과 관련된 연산과 특징 등을 다뤄볼 예정입니다.

(linear mapping에 대한 내용은 추후 다룰 예정입니다.)

행렬의 정의

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad a_{ij} \in \mathbb{R}.$$

($m, n \in \mathbb{N}$은 자연수) (m, n) 행렬 A 는 $m, n-$ 튜플로 표현되는 실수 a_{ij} $(i = 1, \cdots, m, j = 1, \cdots, n)$ 요소로 가지며 m행(rows)과 n열(columns)로 구성됩니다.

- $(1, n)-$행렬을 행(rows) 또는 행 벡터
- $(m, 1)-$행렬을 열(columns) 또는 열벡터

행렬의 덧셈

$$\boldsymbol{A} + \boldsymbol{B} := \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

- 행렬의 사이즈가 같은 것 끼리 더할 수 있습니다.
- **element-wise sum** (각 요소별 덧셈)

행렬의 곱셈

두 행렬, $A \in \mathbb{R}^{m\times n},\ B \in \mathbb{R}^{n\times \mathbb{k}}$ 이 있을 때, 두 행렬의 곱 $C = AB \in \mathbb{R}^{m\times \mathbb{k}}$의 요소 c_{ij}는 다음과 같이 계산됩니다.

$$c_{ij} = \sum_{l=1}^{n} a_{il}b_{lj}, \qquad i = 1,\ldots,m, \quad j = 1,\ldots,k.$$

- 자세히 살펴보면 c_{ij} 는 A의 i번째 행의 요소와 B의 j번째 열의 요소들을 곱해서 더한 것과 같습니다.
- 추후 다룰 예정이지만, 이를 **내적(dot product)**이라고 합니다.
- 내적이라는 것을 명시적으로 표시할 때는 $A \cdot B$로 표시합니다.

예시

$$\text{For } \boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 3}, \boldsymbol{B} = \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 2}, \text{ we obtain}$$

$$\boldsymbol{AB} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 2 & 5 \end{bmatrix} \in \mathbb{R}^{2\times 2},$$

$$\boldsymbol{BA} = \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 4 & 2 \\ -2 & 0 & 2 \\ 3 & 2 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 3}.$$

- $AB \neq BA$, 교환법칙이 성립하지 않습니다.

단위 행렬 (항등 행렬, Identity Matrix)

$$\boldsymbol{I}_n := \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}$$

- $n \times n$ 행렬 (정사각행렬)에서 대각 성분(digagonal)이 모두 1이고, 나머지 요소들은 0인 행렬을 말합니다.
- 행렬의 연산에서 항등원 역할을 합니다.

행렬의 성질인 결합법칙, 분배법칙, 단위 행렬과의 연산을 보여줍니다.

- *Associativity:*

$$\forall \boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{B} \in \mathbb{R}^{n \times p}, \boldsymbol{C} \in \mathbb{R}^{p \times q} : (\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$$

- *Distributivity:*

$$\begin{aligned} \forall \boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m \times n}, \boldsymbol{C}, \boldsymbol{D} \in \mathbb{R}^{n \times p} : (\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} &= \boldsymbol{AC} + \boldsymbol{BC} \\ \boldsymbol{A}(\boldsymbol{C} + \boldsymbol{D}) &= \boldsymbol{AC} + \boldsymbol{AD} \end{aligned}$$

- Multiplication with the identity matrix:

$$\forall \boldsymbol{A} \in \mathbb{R}^{m \times n} : \boldsymbol{I}_m \boldsymbol{A} = \boldsymbol{A} \boldsymbol{I}_n = \boldsymbol{A}$$

  Note that $\boldsymbol{I}_m \neq \boldsymbol{I}_n$ for $m \neq n$.

역행렬 (Inverse Matrix)

$$\boldsymbol{AB} = \boldsymbol{I}_n = \boldsymbol{BA}.$$

The matrices

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 1 \\ 4 & 4 & 5 \\ 6 & 7 & 7 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} -7 & -7 & 6 \\ 2 & 1 & -1 \\ 4 & 5 & -4 \end{bmatrix}$$

are inverse to each other since $\boldsymbol{AB} = \boldsymbol{I} = \boldsymbol{BA}$.

B를 A의 역행렬이라고 부르고, A^{-1}로 표기합니다.

모든 행렬이 역행렬이 존재하는 것은 아닙니다.

역행렬이 존재하는 행렬을 **regular / invertible / nonsingular** 라고 부르며, 역행렬이 존재하지 않는 행렬은 **singular / noninvertible** 이라고 부릅니다.

전치행렬 (Transpose Matrix)

두 행렬 $A \in \mathbb{R}^{m\times n},\ B \in \mathbb{R}^{n\times m}$ 에서 $b_{ij} = a_{ji}$를 만족하면 B행렬을 A의 전치행렬 이라고 합니다. 표기는 $B = A^T$로 표기합니다.

역행렬과 전치행렬의 중요한 속성

$$
\begin{aligned}
\boldsymbol{A}\boldsymbol{A}^{-1} &= \boldsymbol{I} = \boldsymbol{A}^{-1}\boldsymbol{A} \\
(\boldsymbol{A}\boldsymbol{B})^{-1} &= \boldsymbol{B}^{-1}\boldsymbol{A}^{-1} \\
(\boldsymbol{A}+\boldsymbol{B})^{-1} &\neq \boldsymbol{A}^{-1} + \boldsymbol{B}^{-1} \\
(\boldsymbol{A}^{\top})^{\top} &= \boldsymbol{A} \\
(\boldsymbol{A}\boldsymbol{B})^{\top} &= \boldsymbol{B}^{\top}\boldsymbol{A}^{\top} \\
(\boldsymbol{A}+\boldsymbol{B})^{\top} &= \boldsymbol{A}^{\top} + \boldsymbol{B}^{\top}
\end{aligned}
$$

대칭 행렬 (symmetric matrix)

- $A = A^T$을 만족할 때, 행렬 A를 symmetric 하다고 표현합니다.

스칼라 곱(배), Multiplication by a Scalar

스칼라 값 $\lambda \in \mathbb{R}$ 와 행렬 $A \in \mathbb{R}^{m\times n}$ 를 곱하면 그 결과는 다음과 같습니다.

$$\lambda \boldsymbol{A} = \boldsymbol{K},\ K_{ij} = \lambda\, a_{ij}.$$

두 스칼라 값 $\lambda, \psi \in \mathbb{R}$ 에 대하여 다음과 같은 성질이 성립합니다.

- *Associativity:*
  $(\lambda\psi)\boldsymbol{C} = \lambda(\psi\boldsymbol{C}), \quad \boldsymbol{C} \in \mathbb{R}^{m\times n}$
- $\lambda(\boldsymbol{BC}) = (\lambda\boldsymbol{B})\boldsymbol{C} = \boldsymbol{B}(\lambda\boldsymbol{C}) = (\boldsymbol{BC})\lambda, \quad \boldsymbol{B} \in \mathbb{R}^{m\times n}, \boldsymbol{C} \in \mathbb{R}^{n\times k}$.
  Note that this allows us to move scalar values around.
- $(\lambda\boldsymbol{C})^\top = \boldsymbol{C}^\top\lambda^\top = \boldsymbol{C}^\top\lambda = \lambda\boldsymbol{C}^\top$ since $\lambda = \lambda^\top$ for all $\lambda \in \mathbb{R}$.
- *Distributivity:*
  $(\lambda + \psi)\boldsymbol{C} = \lambda\boldsymbol{C} + \psi\boldsymbol{C}, \quad \boldsymbol{C} \in \mathbb{R}^{m\times n}$
  $\lambda(\boldsymbol{B} + \boldsymbol{C}) = \lambda\boldsymbol{B} + \lambda\boldsymbol{C}, \quad \boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{m\times n}$

연립선형방정식 표현

$$\begin{aligned} 2x_1 + 3x_2 + 5x_3 &= 1 \\ 4x_1 - 2x_2 - 7x_3 &= 8 \\ 9x_1 + 5x_2 - 3x_3 &= 2 \end{aligned}$$

다음과 같이 연립선형방정식이 있을 때, 행렬의 곱셈 법칙을 활용하여 간략한 형태로 작성할 수 있었습니다.

$$\begin{bmatrix} 2 & 3 & 5 \\ 4 & -2 & -7 \\ 9 & 5 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \\ 2 \end{bmatrix}$$

일반적으로는 모든 연립선형방정식은 $Ax = b$ 형태로 표현할 수 있습니다.

출처

- Mathmatics for Machine Learning (https://github.com/mml-book/mml-book.github.io)
- https://junstar92.github.io/mml-study-note/2022/06/30/ch2-2.html
- https://www.youtube.com/watch?v=XqOvyfMUAwA&list=PL_iJu012NOxdZDxoGsYidMf2_bERIQaP0&index=10
- https://www.youtube.com/watch?v=bvB5uQXX7WY&list=PLkoaXOTFHiqhVDo0nWybNmihCP_4BjOFR&index=8