1주차(3/3)

행렬

파이썬으로 배우는 기계학습

한동대학교
김영섭 교수

함수와 뉴론

- 학습 목표
  - 행렬을 이해하고, 행렬 연산을 익힌다.

- 학습 내용
  - 왜 행렬인가?
  - 행렬에 관한 용어의 이해
  - 행렬의 사칙 연산
  - 행렬의 전치, 역행렬, 단위행렬

왜 행렬인가?

- 연립 방정식

$$
\begin{aligned}
4x_1 - 5x_2 &= -12 \\
-2x_1 + 3x_2 &= 8
\end{aligned}
$$

$$\Downarrow$$

$$Ax = b$$

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}, \; b = \begin{bmatrix} -12 \\ 8 \end{bmatrix}$$

왜 행렬인가?

- 연립 방정식
  - 행렬 표기
- 기계학습과 행렬
  - 간결한 표기
  - 실수 감소
  - 컴퓨터에 적합한 형식

용어의 이해

- 스칼라 – 크기
  - 질량, 부피, 거리, 속력, 실수, 정수
- 벡터 – 크기와 방향
  - 속도, 가속도, 운동량, 전기장
- 행렬
  - 2차원 배열
- 텐서
  - N 차원 배열

행렬의 정의

- **2 x 3 행렬 A**

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

행렬과 원소, 표기

- 원소 $a_{i,j}$
  - i 번째 행, j번째 열

$$A_{m,n} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

행렬과 원소, 표기

- 열 a_j
  - i 번째 행, j번째 열

$$A_{m,n} = \begin{pmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{pmatrix}$$

- 행 a_i^T

전치행렬

  - i 번째 행

$$A_{m,n} = \begin{pmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{pmatrix}$$

행 벡터와 열 벡터

- 행 벡터
  - 1 x m 행렬

$$\mathbf{w} = \begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix}$$

- 열 벡터

대개의 경우 열벡터로 !!! 기계학습에서는 특히

  - m x 1 행렬

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

행렬의 곱

- $A \in R^{m,n}, B \in R^{n,p}$행렬의 곱

$$C = AB \in \mathbb{R}^{m \times p} \qquad \text{단, } C_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

행렬 A와 B의 곱($AB = C$)

$$\boldsymbol{A} \qquad \boldsymbol{B} \qquad = \qquad \boldsymbol{C}$$

일치

$$(\boldsymbol{M} \times \boldsymbol{N}) \ (\boldsymbol{N} \times \boldsymbol{P}) \quad = \quad (\boldsymbol{M} \times \boldsymbol{P})$$

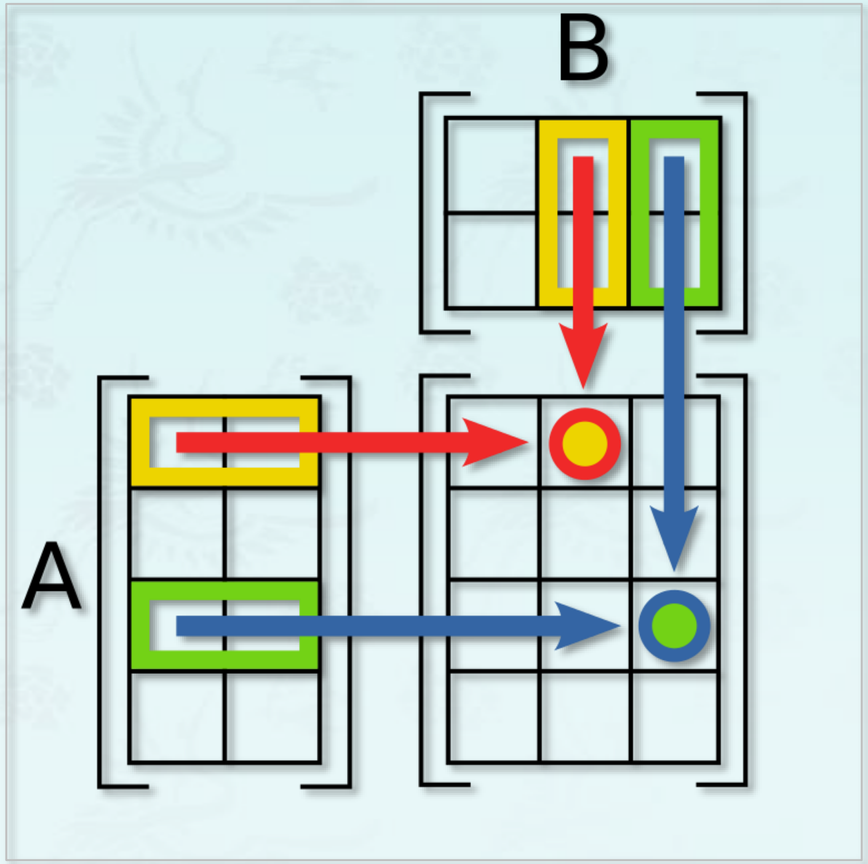
풀잎스쿨 찾아서 코드 비교!!

행렬 곱의 예시

- $A \in R^{m,n}, B \in R^{n,p}$ 행렬의 곱

$$AB = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ x_{41} & x_{42} \end{bmatrix} \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \end{bmatrix}$$

$$= \begin{bmatrix} x_{11}y_{11} + x_{12}y_{21} & x_{11}y_{12} + x_{12}y_{22} & x_{11}y_{13} + x_{12}y_{23} \\ x_{21}y_{11} + x_{22}y_{21} & x_{21}y_{12} + x_{22}y_{22} & x_{21}y_{13} + x_{22}y_{23} \\ x_{31}y_{11} + x_{32}y_{21} & x_{31}y_{12} + x_{32}y_{22} & x_{31}y_{13} + x_{32}y_{23} \\ x_{41}y_{11} + x_{42}y_{21} & x_{41}y_{12} + x_{42}y_{22} & x_{41}y_{13} + x_{42}y_{23} \end{bmatrix}$$



출처: 위키백과, 2018

두 벡터의 내적

- $x, y \in R^n$ 두 벡터의 곱(내적, $x^T y$)

$$\mathbf{x}^T\mathbf{y} \in \mathbb{R} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

$$= \sum_{i=1}^{n} x_i y_i$$

닷 프로덕트 Dot product

X를 전치후 계산

전제조건은 두 벡터의 크기가 같다

특별한 경우에 해당한다

둘중 아무거나 순서 상관없고

결과값이 스칼라이다

두 벡터의 외적

- $A \in R^m, B \in R^n$ 두 벡터의 곱
  (외적, xy^T)

$$\mathbf{x}^T\mathbf{y} \in \mathbb{R}^{m\times n} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix} = \begin{pmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ \ldots & & \cdots & \cdots \\ x_my_1 & x_my_2 & \cdots & x_my_n \end{pmatrix}$$

행렬 곱의 특징

- 결합법칙
  - $A(BC) = (AB)C$

연산이 가능하다는 전제로

- 분배법칙
  - $A(B + C) = AB + AC$
  - $(B + C)A = BA + CA$
  - $k(AB) = (kA)B = A(kB)$
- 교환법칙

주의 교환법칙은 성립하지 않는다

  - $A \cdot B \neq B \cdot A$

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

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

행렬의 덧셈과 뺄셈

- 대응하는 원소끼리 연산

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} + \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} = \begin{pmatrix} x_{11} + y_{11} & x_{12} + y_{12} \\ x_{21} + y_{21} & x_{22} + y_{22} \end{pmatrix}$$

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} - \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} = \begin{pmatrix} x_{11} - y_{11} & x_{12} - y_{12} \\ x_{21} - y_{21} & x_{22} - y_{22} \end{pmatrix}$$

행렬의 전치

트렌스포즈

- 전치: 행과 열의 인덱스를 바꾼 것.
  - 대각선을 기준으로 서로 위치 전환함

- 전치의 특성
- $(A^T)^T = A$
- $(AB)^T = B^T A^T$
- $(A+B)^T = A^T + B^T$

A

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

단위 행렬

- 주 대각선이 모두 1이고 다른 성분은 모 두 0인 행렬이며 I 로 나타냅니다.

곱셈의 항등원이 된다.

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdots$$

행렬

- 학습 정리
  - 행렬의 기본 이해
  - 스칼라, 벡터, 텐서의 정의
  - 행렬의 사칙 연산방법
  - 행렬의 전치, 역행렬, 단위행렬 개념