2021학년도 2학기 언어와 컴퓨터

제14강 벡터, 통계, 데이터 시각화 (2)

박수지

서울대학교 인문대학 언어학과

2021년 11월 1일 월요일

오늘의 목표

1. 주어진 행렬을 더하거나 곱할 수 있다; 할 수 없다면 그 이유를 말할 수 있다.
2. 확률값의 두 가지 특징을 말할 수 있다.
3. 조건부확률의 정의를 말할 수 있다.

복습

벡터

(주로 수치형) 데이터를 표현하는 방식

성분 표시 $\vec{v} = (v_1, v_2, \ldots, v_n) \leftarrow$ 차원의 크기가 n인 벡터

연산
- 덧셈 $(5, 0, 1) + (7, 5, 0) = (12, 5, 1)$
- 상수배 $2 \times (5, 0, 1) = (10, 0, 2)$

속성
- 내적 $(5, 0, 1) \cdot (7, 5, 0) = (5 \times 7) + (0 \times 5) + (1 \times 0) = 35$
- 길이 $\|(5, 0, 1)\| = \sqrt{(5, 0, 1) \cdot (5, 0, 1)} = \sqrt{26}$
- 거리 $\|(5, 0, 1) - (7, 5, 0)\| = \|(-2, -5, 1)\| = \sqrt{30}$

행렬

(벡터를 쌓아서) 직사각형으로 배열한 것

성분 표시 $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ddots & a_{mn} \end{bmatrix} = (a_{ij}) \leftarrow (m \times n)$ 행렬

연산

- 덧셈(addition) $\begin{bmatrix} 1 & 3 & 1 \\ 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 5 \\ 7 & 5 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 6 \\ 8 & 5 & 0 \end{bmatrix}$
- 상수배(scalar multiplication) $2 \cdot \begin{bmatrix} 1 & 8 & -3 \\ 4 & -2 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 16 & -3 \\ 8 & -4 & 10 \end{bmatrix}$
- 전치(transposition) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 6 & 7 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ 2 & 6 \\ 3 & 7 \end{bmatrix}$

행렬

- 영행렬(zero matrix): 모든 성분의 값이 0인 행렬
  - 예: $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
- 항등행렬(identitiy matrix): 대각선 성분의 값은 1이고 나머지 성분은 모두 0인 정사각행렬
  - 예: $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

벡터와 행렬

$(1, 2, 3)$ 길이 3인 벡터

$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ (1×3) 행렬

$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ (2×3) 행렬

$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$ (2×4) 행렬

행렬의 곱셈

예시

$$A = \begin{array}{|c|c|c|c|}\hline & \text{라면} & \text{김밥} & \text{두유} \\ \hline \text{학관} & 1000 & 700 & 1300 \\ \hline \text{중도} & 900 & 850 & 1000 \\ \hline \end{array}$$

$$B = \begin{array}{|c|r|r|r|}\hline & \text{라면} & \text{김밥} & \text{두유} \\ \hline \text{아침} & 0 & 1 & 1 \\ \hline \text{점심} & 1 & 2 & 0 \\ \hline \end{array} \Rightarrow B^T = \begin{array}{|c|r|r|}\hline & \text{아침} & \text{점심} \\ \hline \text{라면} & 0 & 1 \\ \hline \text{김밥} & 1 & 2 \\ \hline \text{두유} & 1 & 0 \\ \hline \end{array}$$

$$AB^T = \begin{bmatrix} \text{학관} \cdot \text{아침} & \text{학관} \cdot \text{점심} \\ \text{중도} \cdot \text{아침} & \text{중도} \cdot \text{점심} \end{bmatrix} = \begin{bmatrix} 700 + 1300 & 1000 + 700 \cdot 2 \\ 850 + 1000 & 900 + 850 \cdot 2 \end{bmatrix}$$

행렬의 곱셈

다른 예시

$$A = \begin{bmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \vec{r}_3 \end{bmatrix}, B = \begin{bmatrix} \vec{c}_1 & \vec{c}_2 \end{bmatrix} \Rightarrow A \cdot B = \begin{bmatrix} \vec{r}_1 \cdot \vec{c}_1 & \vec{r}_1 \cdot \vec{c}_2 \\ \vec{r}_2 \cdot \vec{c}_1 & \vec{r}_2 \cdot \vec{c}_2 \\ \vec{r}_3 \cdot \vec{c}_1 & \vec{r}_3 \cdot \vec{c}_2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 3 & 4 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1000 \\ 1 & 100 \\ 0 & 10 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 1000 \\ 1 & 100 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} 2 & 3 & 4 \\ 1 & 0 & 0 \end{bmatrix} =$$

행렬의 곱셈

성질

- 결합법칙(associativity) 성립함 $(AB)C = A(BC)$
- 분배법칙(distributivity) 성립함 $A(B + C) = AB + AC, (A + B)C = AC + BC$
- 교환법칙(commutativity) 성립 안 함 $AB \neq BA$

확률

표본공간 확률 실험의 가능한 모든 결과의 집합
- 예: 6면체 주사위의 표본공간 {⚀, ⚁, ⚂, ⚃, ⚄, ⚅}

확률변수 표본공간의 각 원소에 실수 값을 대응시키는 함수
- 예: 주사위를 던져서 나오는 눈의 값 $X \in \{1, 2, 3, 4, 5, 6\}$

확률질량함수 확률변수의 각 원소에 확률 값을 대응시키는 함수
- 예: $f(x) = P(X = x) = \frac{1}{6}$

확률의 속성

1. 모든 확률 값은 0보다 크거나 0과 같다.
2. 모든 확률 값을 더하면 1이 된다.

조건부확률의 정의

사건 A가 일어났다는 가정 하에 사건 B가 일어날 확률

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{(\text{사건 } A\text{와 } B\text{가 모두 일어날 확률})}{(\text{사건 } A\text{가 일어날 확률})}$$

조건부확률의 활용

- 단어 w_1 가 나타났을 때 단어 w_2 가 나올 확률 (SLP3 3장 n-그램 언어 모형)
- 문서 d가 긍정적인 내용일 때 단어 w가 나올 확률 (SLP3 4장 단순 베이즈 분류기)

다음 시간에 할 일

SLP 3장 읽어 오기