# 데이터과학과 AI를 위한 파이썬

## 05강. 벡터와 공간

세종사이버대학교

김명배 교수

## 학습내용

- 스칼라, 벡터, 행렬 그리고 텐서
- 벡터의 특성과 특수한 벡터
- 선형 독립과 종속
- 부분공간의 기저와 랭크

## 학습목표

- 선형대수학에서의 스칼라와 벡터, 행렬, 텐서의 개념을 설명할 수 있다.
- 벡터의 특성과 특수한 벡터에 대해 설명할 수 있고, 파이썬으로 표현할 수 있다.
- 벡터 집합에서의 선형 독립과 선형 종속에 대해 설명할 수 있고, 식별할 수 있다.
- 벡터 공간에서 부분공간을 이해하고 기저와 랭크를 설명하고 파이썬으로 실행 할 수 있다.

# 1. 스칼라, 벡터, 행렬 그리고 텐서

**HR 데이터**

| 사원ID | 업무만족도 | 프로젝트 수 | 평균급여 | 업무시간 |
|---|---|---|---|---|
| 1 | 0.48 | 3 | 96 | 3 |
| 2 | 0.58 | 4 | 97 | 3 |
| 3 | 0.61 | 5 | 98 | 4 |
| 4 | 0.67 | 5 | 99 | 2 |
| 5 | 0.79 | 5 | 112 | 6 |
| 6 | 0.89 | 4 | 113 | 6 |

스칼라

텐서

벡터

행렬

# 1. 스칼라, 벡터, 행렬 그리고 텐서

## 1) 스칼라(scalar)

- 숫자 하나로 된 데이터
- 실수인 숫자 중 하나이므로 실수 집합 $R$의 원소로 표현

$$x \in R$$

## 2) 벡터(vector)

- 숫자 여러 개가 특정한 순서대로 모여 있는 것
- 방향이 있고 순서가 있고 차원의 공간에 존재

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad x \in R^n \qquad x_i: \text{원소(성분)}$$

# 1. 스칼라, 벡터, 행렬 그리고 텐서

### 3) 행렬(matrix)

- 벡터가 여러 개 있는 형태

행벡터

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \\ x_{51} & x_{52} & x_{53} & x_{54} \\ x_{61} & x_{62} & x_{63} & x_{64} \end{bmatrix}, \ X \in R^{6\times 4} \qquad , \ x_{mn}\text{: 성분}$$

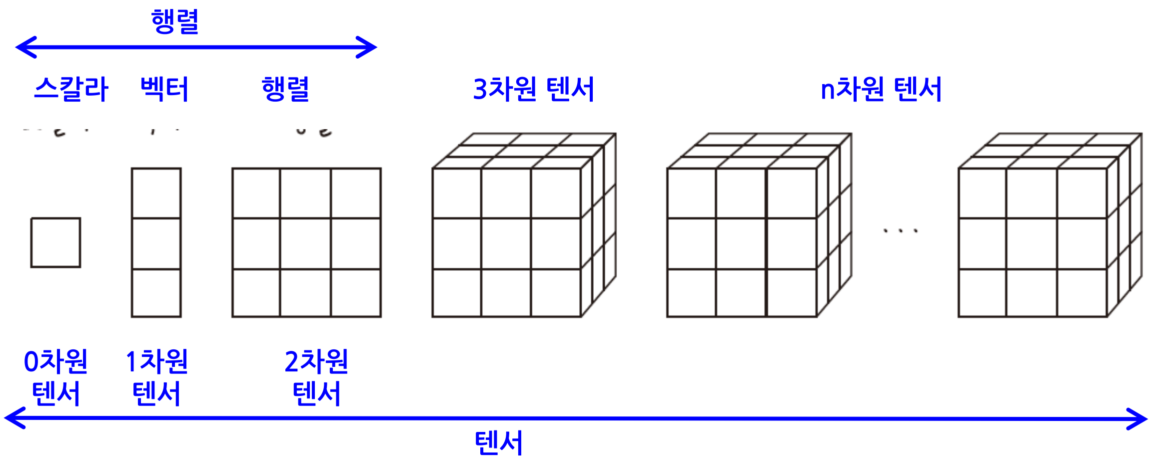
열벡터

- 스칼라 1×1 행렬, 벡터는 n×1 행렬

# 1. 스칼라, 벡터, 행렬 그리고 텐서

## 4) 텐서(tensor)

- 행렬이 여러 개 있는 형태



# 1. 스칼라, 벡터, 행렬 그리고 텐서

## 4) 텐서(tensor)

a) 그레이 스케일　　　　b) RGB 이미지

하나의 채널에 2차원 행렬(픽셀)

➔ <u>행렬</u>

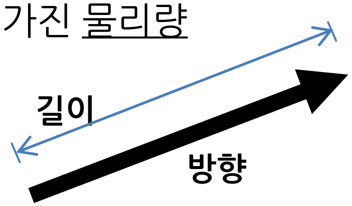
‘R’, ‘G’, ‘B’  세개 체널에
각각 2차원 행렬(픽셀)

➔ <u>텐서</u>

# 2. 벡터

## 1) 벡터란

- 기하학적 정의
  크기와 방향을 가진 물리량




- 대수적 정의
  실수 또는 함수를 성분으로 갖는 순서쌍 혹은 배열

- **파이썬** : numpy.array()

# 2. 벡터

**2) 벡터의 특징**

a) 방향성이 있기 때문에

b) 크기뿐만 아니라 방향도 중요

c) 크기와 방향이 같으면 같은 벡터

d) 같은 벡터라도 좌표평면상에 다르게 위치할 수 있음

3) 벡터의 표현

- 기하학적 표기

$$\overrightarrow{AB} \ or \ \vec{a}$$

- 대수적 표기

$$x = (x_1, x_2, \cdots, x_n) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

# 2. 벡터

### 3) 특수한 벡터

#### a) 단위벡터

- 길이가 1인 벡터
- 화살표가 아닌 헷으로 표현함 ($\hat{v}$)

※ 벡터의 크기 = 벡터의 길이 = 벡터의 norm(노름)

$\vec{v} = [x \;\; y]$일 때, $|\vec{v}| = \|\vec{v}\| = \sqrt{x^2 + y^2}$

단위벡터 $u = \left[\frac{x}{\sqrt{x^2+y^2}} \quad \frac{y}{\sqrt{x^2+y^2}}\right]$

[파이썬] scipy의 선형대수 서브패키지 linalg에서 norm() 함수 사용

# 2. 벡터

### 3) 특수한 벡터

#### a) 단위벡터

연습문제) $\vec{v} = [2\ 3\ 1]$일 때, 단위 벡터를 구하세요.

[풀이] $\|\vec{v}\| = \sqrt{2^2 + 3^2 + 1^2} = \sqrt{14}$

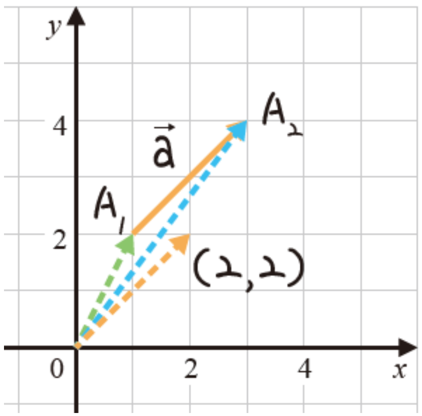
$$u = \begin{bmatrix} \frac{2}{\sqrt{14}} & \frac{3}{\sqrt{14}} & \frac{1}{\sqrt{14}} \end{bmatrix}$$

# 2. 벡터

### 3) 특수한 벡터

#### b) 위치 벡터

- 원점이 시작점인 벡터로 재표현한 벡터




**위치 벡터의 끝점**

$= \overrightarrow{OA_2} - \overrightarrow{OA_1}$

= (3, 4)-(1, 2)

= (2, 2)

# 2. 벡터

**3) 특수한 벡터**

**c) 영 벡터**

- 모든 성분이 **0**으로 구성된 벡터
- 크기가 0이고 시작점과 끝점이 같은 벡터 $\vec{0}\ or\ \begin{bmatrix}0\\0\end{bmatrix},\quad [0\ \ 0]$
- 모든 차원의 벡터 공간은 반드시 영벡터를 포함해야 함(선형결합을 위해)

※ 영벡터의 특징

- 어떤 벡터와도 **평행**함
- 어떤 벡터와도 **같은 직선** 위에 있음
- 어떤 벡터와도 **수직**임
- 덧셈과 뺄셈의 **항등원**$(x + 0 = x,\ x - 0 = x)$

[파이썬] numpy의 zeros()함수를 사용

# 2. 벡터

## 4) 인공지능에서 벡터의 활용

### 범주형 변수의 숫자형 표현

| 연봉 | 연봉_숫자 | 연봉_low | 연봉_medium | 연봉_high |
|---|---|---|---|---|
| Low | 1 | 1 | 0 | 0 |
| Medium | 2 | 0 | 1 | 0 |
| High | 3 | 0 | 0 | 1 |

➔ 가변수(dummy variable)

# 3. 선형 결합과 선형 독립

**1) 선형 결합**

**벡터의 스칼라 곱과 덧셈을 조합하여 새로운 벡터를 얻는 연산**

두 벡터 $\vec{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$,

임의의 실수 $(C_1, C_2)$에 대해 선형결합:

$$C_1\vec{a} + C_2\vec{b} = \begin{bmatrix} C_1 a_1 & C_2 b_1 \\ C_1 a_2 & C_2 b_2 \end{bmatrix}$$

집합의 생성(span)벡터들의 선형결합으로 타나낼 수 있는 모든 벡터의 집합

$$span(\vec{a}, \vec{b})$$

# 3. 선형 결합과 선형 독립

**2) 선형 독립과 종속**

벡터의 집합 $A = \{a_1,\ a_2,\ a_3, \cdots,\ a_n\}$ 에 대해

$c_1 a_1 + c_2 a_2 + c_3 a_3 + \cdots + c_n a_n = 0$을 만족하는 $c_i$가 모두 0일 뿐일 때

➔ 선형 독립(linearly independent)

0이 아닌 $c_i$가 존재할 때

➔ 선형 종속(linearly dependent)

하나의 벡터를 나머지 벡터의 선형결합으로 표현할 수 있으면,
선형 종속이 됨

# 3. 선형 결합과 선형 독립

**2) 선형 독립과 종속**

[연습문제] 다음은 선형 독립인가? 선형 종속인가?

1) $\left\{\begin{bmatrix}4\\2\end{bmatrix}\begin{bmatrix}8\\4\end{bmatrix}\right\}$ ➔ $2\begin{bmatrix}4\\2\end{bmatrix}=\begin{bmatrix}8\\4\end{bmatrix}$

2) $\left\{\begin{bmatrix}5\\7\end{bmatrix}\begin{bmatrix}3\\2\end{bmatrix}\begin{bmatrix}2\\5\end{bmatrix}\right\}$ ➔ $\begin{bmatrix}3\\2\end{bmatrix}+\begin{bmatrix}2\\5\end{bmatrix}=\begin{bmatrix}5\\7\end{bmatrix}$

3) $\left\{\begin{bmatrix}3\\0\end{bmatrix}\begin{bmatrix}0\\-2\end{bmatrix}\right\}$ ➔ 직교관계

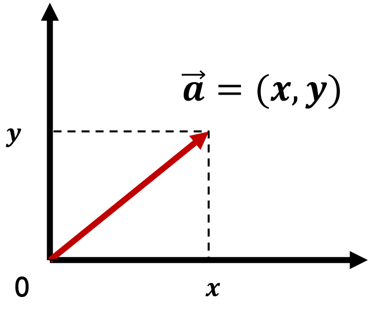
# 3. 선형 결합과 선형 독립

### 3) 벡터공간과 부분공간의 기저

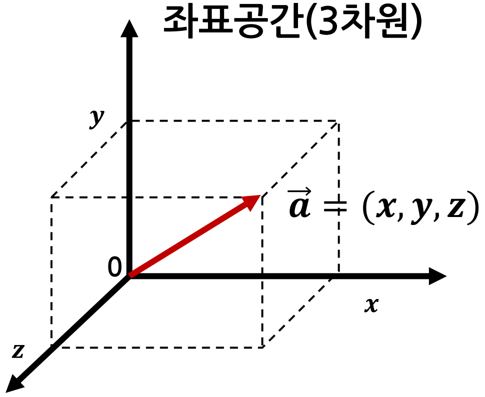
#### 가) 벡터공간

- 같은 공간에서 선형 결합 연산이 가능한 벡터가 모인 집합

좌표평면(2차원)




좌표공간(3차원)



# 3. 선형 결합과 선형 독립

### 3) 벡터공간과 부분공간의 기저

### 나) 벡터 공간을 만족하는 추가 공리

- 덧셈 관련 공리

벡터 공간 $V$는 덧셈에 대해 닫혀 있음(닫힘성),

즉 $a$ 및 $b$가 벡터 공간 $V$에 존재하면 $a + b$도 $V$에 존재해야 함

- 가환성(commutativity): $a + b = b + a$
- 결합성(associativity): $(a + b) + c = a + (b + c)$
- 영 벡터가 존재: $a + 0 = a$
- 덧셈 역원이 존재: $a + (-a) = 0$

# 3. 선형 결합과 선형 독립

### 3) 벡터공간과 부분공간의 기저

#### 나) 벡터 공간을 만족하는 추가 공리

- 스칼라 관련 공리

벡터 공간 $V$는 스칼라 곱셈에 대해 닫혀 있음(닫힘성),

즉 $a$가 $V$에 존재하고 $\alpha$가 스칼라이면, $\alpha a$도 $V$에 존재해야 함

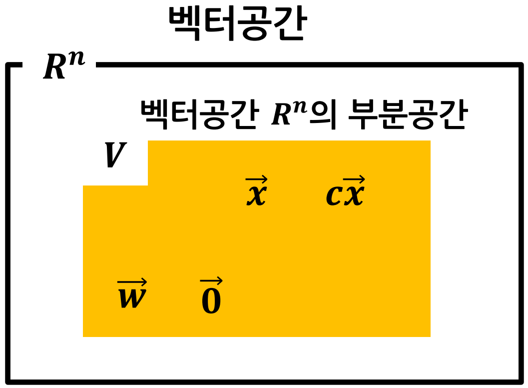
- 스칼라 결합성: $\alpha(\beta a) = (\alpha\beta)a$
- 단위원: $1\,a = a$
- 분배성(distributivity): $\alpha\,(a + b) = \alpha a + \alpha b$
- 분배성: $(\alpha + \beta)\,a = \alpha a + \beta a$

# 3. 선형 결합과 선형 독립

### 3) 벡터공간과 부분공간의 기저

#### 다) 부분공간

벡터 공간 $R^n$이 있을 때, $R^n$ 의 부분 공간($V$)으로 만든 공간을 전체 공간의 부분 공간이라고 함



# 3. 선형 결합과 선형 독립

### 3) 벡터공간과 부분공간의 기저

#### 다) 부분공간

부분공간의 충족 조건

- $R^n$의 부분 공간 $v$는 영 벡터($\vec{0}$)를 포함해야 함
- 벡터 $\vec{w}$가 벡터 공간 $v$에 포함되어 있다면, $\vec{w}$에 임의의 스칼라($c$)를 곱한 값 또한 $v$에 포함되어야 함(스칼라 곱셈에 대해 닫혀 있다)
- 벡터 $\vec{a}$와 $\vec{b}$가 $V$에 포함되어 있다면, $\vec{a} + \vec{b}$도 $v$에 포함되어야 함(덧셈에 대해 닫혀 있다)

# 3. 선형 결합과 선형 독립

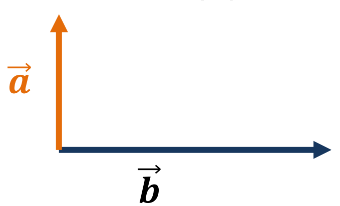
### 3) 벡터공간과 부분공간의 기저

#### 라) 부분공간의 기저(base)

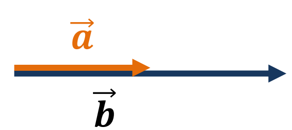
- 벡터 공간 $V$을 생성할 때 **최소한으로 필요한 선형 독립인 벡터**의 집합
- 기저 벡터의 선형결합으로 벡터 공간의 모든 벡터를 표현할 수 있음
- 차원에 따라 기저 벡터의 개수가 정해져 있음(**차원 = 기저**)

  2차원 : 기저 2개, 3차원 : 기저 3개

$R^2$의 기저(O)




$R^2$의 기저(X)



# 3. 선형 결합과 선형 독립

### 3) 벡터공간과 부분공간의 기저

#### 라) 부분공간의 기저(base)

- 기저 벡터 중에서도 여러 원소 중 하나만 값이 1이고,

다른 값은 0인 경우를 **표준기저 벡터(standard basis vector)**라고 함

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, v_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad [v_1, v_2, v_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

# 3. 선형 결합과 선형 독립

### 3) 벡터공간과 부분공간의 기저

#### 라) 부분공간의 기저(base)

- 행렬 A에서 선형 독립인 행 **혹**은 **열**의 개수를 **랭크(rank)**라고 함

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

$$rk(A)\ or\ rank(A) = 2$$
$$rank(A) = rank(A^T)$$

**[파이썬]** numpy.linalg.matrix_rank() 함수 사용

# 3. 선형 결합과 선형 독립

### 3) 벡터공간과 부분공간의 기저

#### 라) 부분공간의 기저(base)

[연습문제]

① $\vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ 일 때, $\vec{a} + \vec{b}$의 랭크를 구하세요.

② $\vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$ 일 때, $\vec{a} + \vec{b}$의 랭크를 구하세요.

# 실습 동영상

# 정리하기

## 1. 스칼라, 벡터, 행렬 그리고 텐서

- 스칼라 : 숫자로 된 데이터
- 벡터 : 숫자 여러 개가 특정한 순서대로 모여 있는 형태
- 행렬 : 벡터 여러 개 있는 형태로 행과 열로 구성된 형태
- 텐서 : 3차원 이상의 차원으로 구성된 형태

## 2. 벡터의 특성과 특수한 벡터

- 벡터 : 크기와 방향이 있는 물리량, 실수 또는 함수를 성분으로 갖는 순서쌍 혹은 배열
- 크기와 방향이 같으면 같은 벡터이기 때문에 좌표평면 상에 다르게 위치할 수 있음
- 벡터의 표현방법

$$\overrightarrow{AB} \ or \ \vec{a} \qquad x = (x_1, x_2, \cdots, x_n) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- 단위벡터 :길이가 1인 벡터($\hat{v}$)
- 노름 : 벡터의 크기($\|\vec{v}\|, |\vec{v}|$)
- 위치벡터 : 원점이 시작점인 벡터로 재표현한 벡터
- 영벡터 : 모든 성분이 **0**으로 구성된 벡터 $\vec{0} \ or \ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, [0 \ \ 0]$

# 정리하기

## 3. 선형 독립과 종속

- 선형결합 : 벡터의 스칼라 곱과 덧셈을 조합하여 새로운 벡터를 얻는 연산($span(\vec{a},\vec{b})$)
- 벡터의 집합 $A$={$a\_1$, $a\_2$, $a\_3$,⋯, $a\_n$ } 에 대해
  $c_1\, a_1 + c_2\, a_2 + c_3\, a_3 + \cdots + c_n\, a_n = 0$을 만족하는 0이아닌 $c_i$가 존재할 때,
  선형 종속, 그렇지 않으면 선형 독립

## 4. 부분공간의 기저와 랭크

- 벡터 공간 $R^n$이 있을 때, $R^n$ 의 부분 공간($V$)으로 만든 공간을 전체 공간의 부분 공간이라고 함
- 부분공간의 기저 : 벡터 공간 $V$을 생성할 때 최소한으로 필요한 선형 독립인 벡터의 집합
- 표준기저 : 기저 벡터 중에서도 여러 원소 중 하나만 값이 1이고, 다른 값은 0인 경우

$$v_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, v_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}, v_1 = \begin{bmatrix}0\\0\\1\end{bmatrix} \qquad [v_1, v_2, v_3] = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$$

- 행렬 A에서 선형 독립인 행 **혹**은 열의 개수를 랭크(rank)라고 함

$$rk(A)\ or\ rank(A)$$
$$rank(A) = rank(A^T)$$

# 정리하기

**5. 파이썬 활용**

- 벡터 : numpy.array() 함수
- 노름 : scipy의 선형대수 서브패키지 linalg에서 norm() 함수
- 영벡터 : numpy의 zeros() 함수
- 랭크 : numpy의 linalg.matrix_rank() 함수