

# 클린업 3주차

## 3팀 선형대수학

황정현  
고경현  
김지민  
반경림  
전효림

# INDEX

---

- 0. 지난주 REVIEW
- 1. 차원의 저주와 차원축소
  - 2. 고유값 분해 (EVD)
  - 3. 주성분 분석 (PCA)
  - 4. 특이값 분해 (SVD)
  - 5. 잠재요인분석 (LSA)
  - 6. 계층화분석법 (AHP)

0

지난주 REVIEW

## 행렬식과 선형변환

- 선형변환 전후의 넓이 관계 설명

지난주 이야기!



‘선형변환’의 관점에서 바라본 선형방정식  $Ax = b$

$$\begin{matrix} A & x & \\ \begin{pmatrix} 2 & -3 \\ 1 & 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \end{pmatrix} & = \begin{pmatrix} b \\ -1 \\ 2 \end{pmatrix} \end{matrix}$$

벡터  $x$ 에 행렬  $A$ 를 곱하는 선형변환을 통해 벡터  $b$  만들

input

벡터  $x$ 가 만드는  
공간의 넓이

Linear Transformation

$|\det(A)|$ 배 변화

output

벡터  $b$ 가 만드는  
공간의 넓이

## 선형부분공간

벡터공간

선형부분공간

*subspace*

span

column space

$\mathbb{R}^n$

선형부분공간

- 벡터공간 안의 몇 가지 벡터 집합으로 이루는 또다른 부분 벡터공간

☀ 선형 조건 ☀

**0**벡터(원점)가 존재한다

선형부분공간 내 벡터에 다른 벡터를 **합**한 것도 선형부분공간에 존재한다

선형부분공간 내 벡터에 **상수**를 곱한 것도 선형부분공간에 존재한다

## Null space

영공간  
*Null Space*

해가 없다



$Ax = 0$ 을 만족하는 해  $x$ 가 이루는 공간,  
즉 homogenous한 방정식의 해의 집합

$$Ax = 0$$

*homogenous*

일치한다

*Consistent* $x$ 가 모두 0

해가 유일하다

*trivial solution*

해가 무수히 많다

*non-trivial solution*

## 선형독립

- $c_1x_1 + c_2x_2 + \dots + c_nx_n = 0$ 에서  $c$ 가 모두 0인 경우를 제외한 어떠한  $c_i$ 의 조합으로도 식이 성립되지 않는 경우



$$A x = 0$$

- 선형방정식을 만족하는 벡터  $x$ 가 오직 0벡터뿐인 경우



- Null space에 0벡터만 있는 경우

## 투영벡터 (Projection)

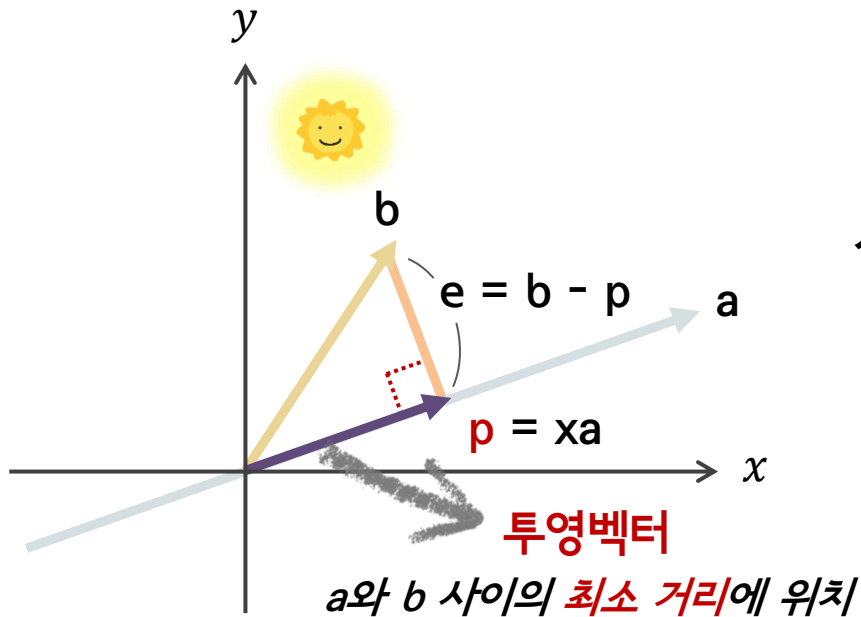
벡터  $b$ 를 직선  $a$  위의  
벡터로 매핑



벡터  $b$ 의 직선 공간으로의  
공간압축



공간압축  
선형변환



실제값( $b$ )과 예측값( $p$ )의 차이인  $e$ 와  
벡터  $a$ 가 직교함을 이용하여 계산

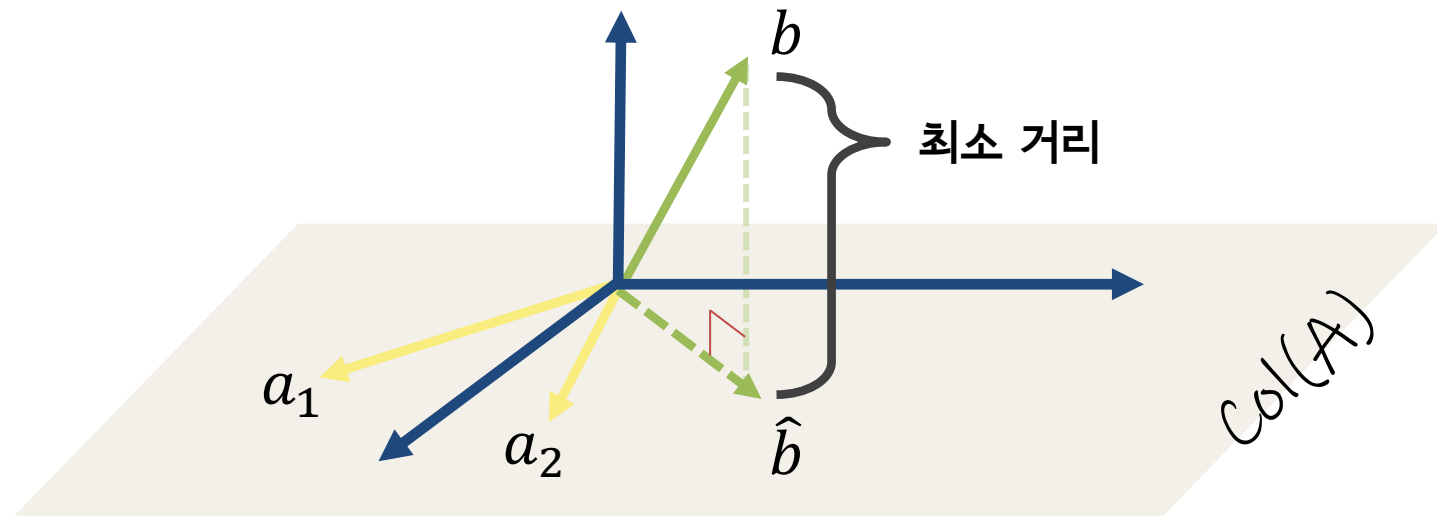


## 투영벡터와 Least Square Method

### 최소자승법

Least Square Method

$Ax = b$ 의 해가 없는 경우 **최소 거리**를 바탕으로  
가장 근접한  $A\hat{x} = \hat{b}$ 의 해  $\hat{x}$ 를 구하는 것



# 1

## 차원의 저주와 차원 축소

## 차원의 저주



=

변수의  
개수

=

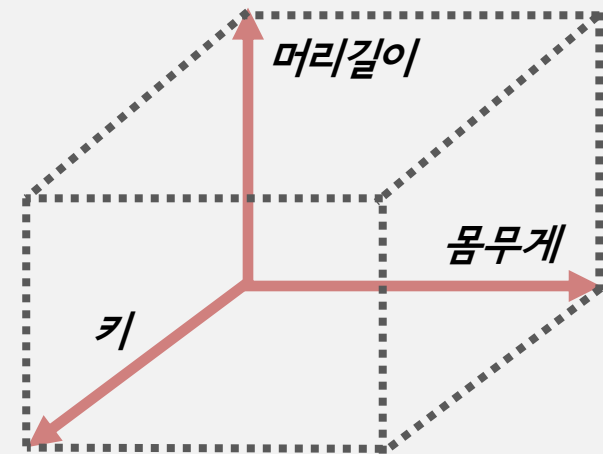
축의  
개수

변수 3개

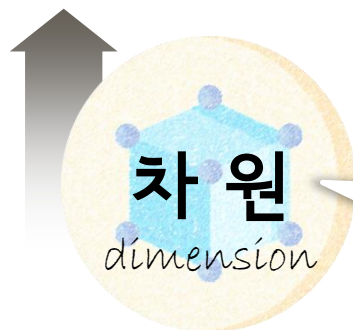
키	몸무게	머리 길이
168	59	20
150	48	16
175	70	20
181	77	30



3차원 공간



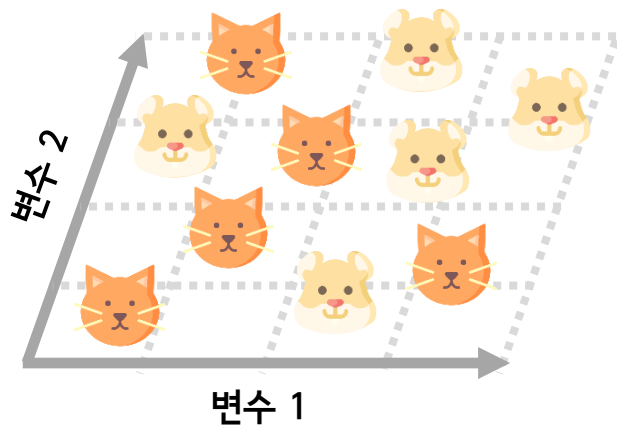
## 차원의 저주



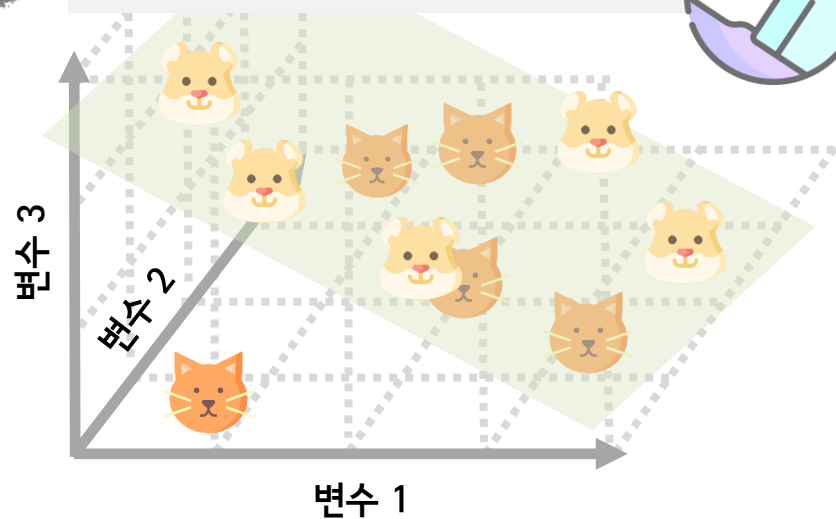
많은 정보를 가지고 있다

고차원일수록  
좋다?

2차원



3차원



## 차원의 저주



차원  
dimension

많은 정보를 가지고 있다

고차원일수록  
좋다?



하지만!

변수가 많아질수록

overfitting(과적합)의 문제 발생 가능성 농후

변수 간 관련성이 높은 경우 공간 낭비

변수 1

변수 1



## 차원의 저주

차원  
dimension

많은 정보를 가지고 있다

고차원일수록  
좋다?

하지만!

변수가 많아질수록

overfitting(과적합)의 문제 발생 가능성 농후

변수 간 관련성이 높은 경우 공간 낭비

차원의 저주

변수 1

변수 1


## 차원의 저주



차원  
dimension

많은 정보를 가지고 있다

고차원일수록  
좋다?



변수 추출

Feature Extraction

기존 변수를 조합해 새로운 변수 생성

변수 선택

Feature Selection

주요한 일부 변수만 사용

해결

차원의 저주

변수 1

변수 1



## 차원의 저주

차원  
dimension

많은 정보를 가지고 있다

고차원일수록  
좋다?

변수 추출

Feature Extraction

기존 변수를 조합해 새로운 변수 생성

변수 선택

Feature Selection

주요한 일부 변수만 사용

해결

차원의 저주

변수 1

변수 1





# 2

고유값 분해(EVD)

## 고유값과 고유벡터

## 1 개념

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad n \times n \text{ 정방행렬}$$

eigenvector  
고유벡터

$$A \neq 0 X = \lambda X$$

고유값  
eigenvalue

$$Ax = \lambda x$$

$$Ax - \lambda x = 0$$

$$(A - \lambda I)x = 0$$

## 고유값과 고유벡터

## 1 개념

Q

그렇다면 이 식은  
homogeneous 방정식인가요?



0이 아닌 해를 가진다는 조건을 근거로  
해가 무수히 많다고 해석할 수 있다

$$Ax = 0$$

homogeneous

x가 모두 0

해가 유일하다

trivial solution

해가 무수히 많다 ✓

non-trivial solution

eigenvector  
고유벡터

$$A \neq 0 X = \lambda X$$

고유값  
eigenvalue

$$Ax = \lambda x$$

$$Ax - \lambda x = 0$$

$$(A - \lambda I)x = 0$$

## 고유값과 고유벡터

1 개념

만약  $(A - \lambda I)$ 의 역행렬이 존재한다면?

$$(A - \lambda I)^{-1} (A - \lambda I)x = 0$$

 $x = 0$  (trivial solution)고유벡터  $\neq 0$  벡터 $(A - \lambda I)$ 의 역행렬이 존재하지 않는다,  $\det(A - \lambda I) = 0$ 이다!

## 고유값과 고유벡터

## 2 예시

$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ 일 때 고유값  $\lambda$ 와 고유벡터  $x$ 를 구해보자!

⊙ 고유값  $\lambda$  찾기  $\rightarrow \det(A - \lambda I) = 0$  이용

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = 0 \quad \rightarrow \begin{matrix} \lambda = 1 \\ \lambda = 3 \end{matrix}$$

⊙ 고유벡터  $x$  찾기  $\rightarrow (A - \lambda I)x = 0$  이용 ( $\lambda = 3$ 일 때)

$$\begin{pmatrix} 2-3 & 1 \\ 1 & 2-3 \end{pmatrix} = RREF \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \quad \rightarrow \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

## 고유값과 고유벡터

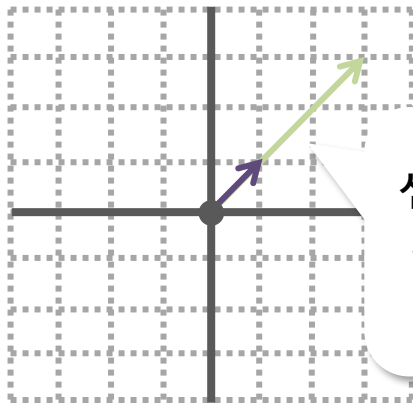
## 3 선형변환의 관점

$$\begin{matrix} A & x \\ \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{matrix} = \begin{matrix} b \\ \begin{pmatrix} 3 \\ 3 \end{pmatrix} \end{matrix} = \begin{matrix} A & x \\ \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{matrix} = \begin{matrix} \lambda & x \\ 3 & \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{matrix}$$

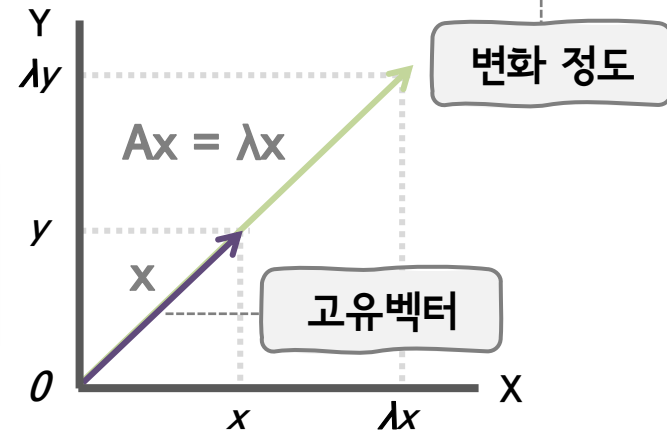
## 고유값과 고유벡터

## 3 선형변환의 관점

$$\begin{matrix} A \\ \left( \begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix} \right) \end{matrix} \begin{matrix} x \\ \left( \begin{smallmatrix} 1 \\ 1 \end{smallmatrix} \right) \end{matrix} = \begin{matrix} b \\ \left( \begin{smallmatrix} 3 \\ 3 \end{smallmatrix} \right) \end{matrix} = \begin{matrix} A \\ \left( \begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix} \right) \end{matrix} \begin{matrix} x \\ \left( \begin{smallmatrix} 1 \\ 1 \end{smallmatrix} \right) \end{matrix} = \begin{matrix} \lambda \\ 3 \end{matrix} \begin{matrix} x \\ \left( \begin{smallmatrix} 1 \\ 1 \end{smallmatrix} \right) \end{matrix}$$



선형변환을 취한 후에도  
같은 직선 위에 놓이는  
원점을 지나는 벡터



## 대각화와 고유값 분해

### 1 개념

$n \times n$  정방행렬  $A$

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

고유벡터 행렬  $P$

$$\begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$$

행렬  $A$ 의 선형독립인  
고유벡터  $n$ 개

고유값 대각행렬  $D$

$$\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

대각화

$$D = P^{-1} A P$$

고유값  
분해

$$A = P D P^{-1}$$



## 대각화와 고유값 분해

### 2 예시

$$A = \begin{pmatrix} 4 & 2 \\ 3 & 5 \end{pmatrix}, \lambda = 7 \text{일 때 } \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \lambda = 2 \text{일 때 } \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$P = \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}, D = \begin{pmatrix} 7 & 0 \\ 0 & 2 \end{pmatrix}, P^{-1} = \begin{pmatrix} 0.2 & -0.6 \\ 0.2 & 0.4 \end{pmatrix}$$

$$A = PDP^{-1} = \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 7 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0.2 & -0.6 \\ 0.2 & 0.4 \end{pmatrix}$$



## 대각화와 고유값 분해

## 3 활용

- $\det(A)$ ,  $A$ 의 거듭제곱, 역행렬 등의 계산 용이

*Example*

행렬  $A$ 의 거듭제곱  $A^k = PD^kP^{-1}$

$$\begin{aligned} A^2 &= (PDP^{-1})^2 = PDP^{-1}PDP^{-1} \\ &= PD^2P^{-1} \end{aligned}$$

# 3

## 주성분 분석(PCA)

## 공분산 행렬

## 공분산 행렬

Covariance Matrix

자료 간의 공분산을 행렬로 나타낸 것

~~~~~

수식으로 이해하기

 $X = (X_1, X_2, \dots, X_n)^T$ 일 때  $X$ 의 공분산 행렬은

$$\text{Cov}(X, X) = E[(X - E(X))(X - E(X))^T]$$

$$= \begin{pmatrix} \text{Cov}(X_1, X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_2, X_1) & \text{Cov}(X_2, X_2) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_n, X_1) & \text{Cov}(X_n, X_2) & \cdots & \text{Cov}(X_n, X_n) \end{pmatrix}$$

분산

## 공분산 행렬

## 공분산 행렬

Covariance Matrix

자료 간의 공분산을 행렬로 나타낸 것

~~~~~

선형변환으로 이해하기

$$\begin{bmatrix} \text{분산} & & \\ & 3 & 2 \\ & 2 & 4 \\ & & \text{공분산} \end{bmatrix}$$

분산

각각의 개별 변수가 퍼져 있는 정도

+

공분산

변수들의 공동적 움직임 설명

## 공분산 행렬

### 공분산 행렬

Covariance Matrix

자료 간의 공분산을 행렬로 나타낸 것

~~~~~

선형변환으로 이해하기



공분산  
행렬

- 분산과 공분산만큼 공간 변화
- 변수 간의 연관성 설명
- 변수들이 어떤 식으로 분포되어 있는지

## 공분산 행렬

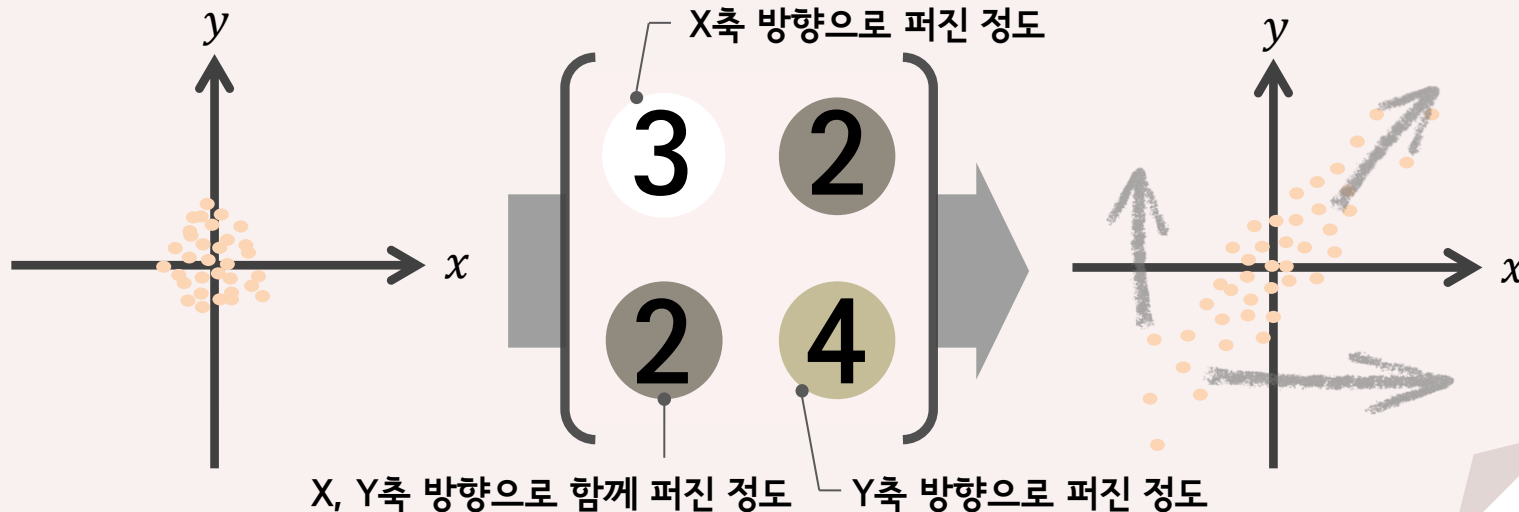
## 공분산 행렬

Covariance Matrix

자료 간의 공분산을 행렬로 나타낸 것

~~~~~

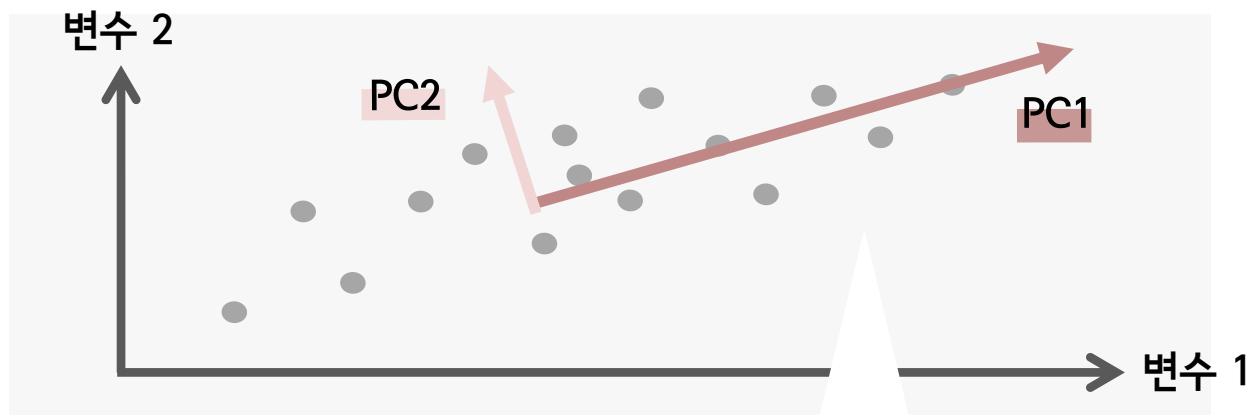
선형변환으로 이해하기



## 주성분 분석(PCA) 개념

*Principle Component Analysis*

데이터를 가장 잘 설명하는 주성분을 찾아내  
그 주성분이 이루는 공간으로 데이터를 **정사영**시켜 차원을 축소하는 방법



이 두 성분 중 데이터를 더 잘 설명하는 것은?





## 주성분 분석(PCA) 개념

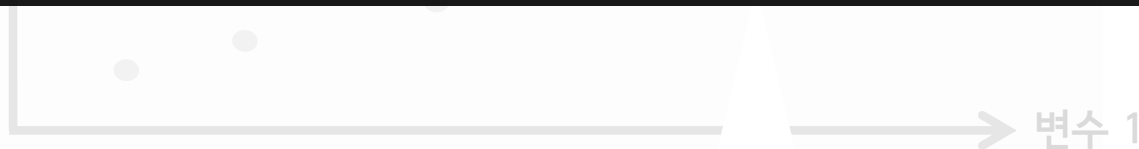
*Principle Component Analysis*

데이터를 가장 잘 설명하는 주성분을 찾아내



그 주성분이 이루는 공간으로 데이터를 정사영시켜 차원을 축소하는 방법

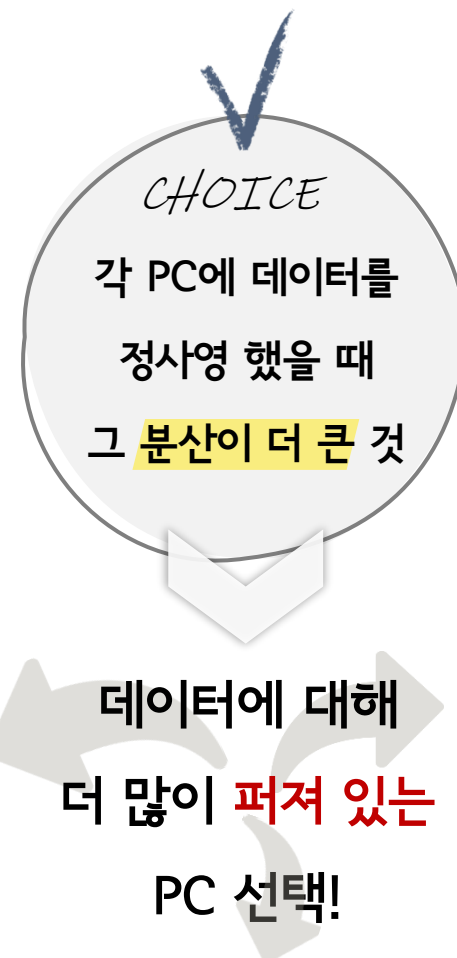
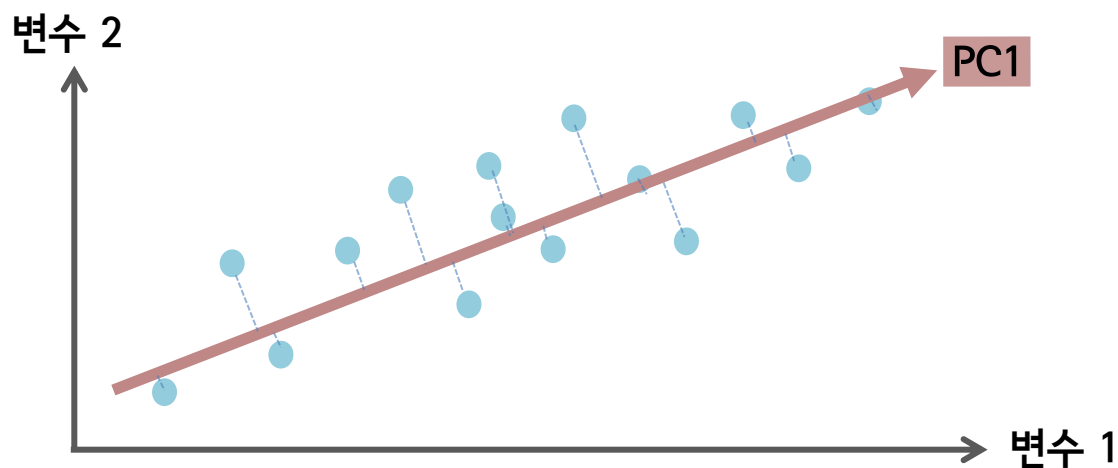
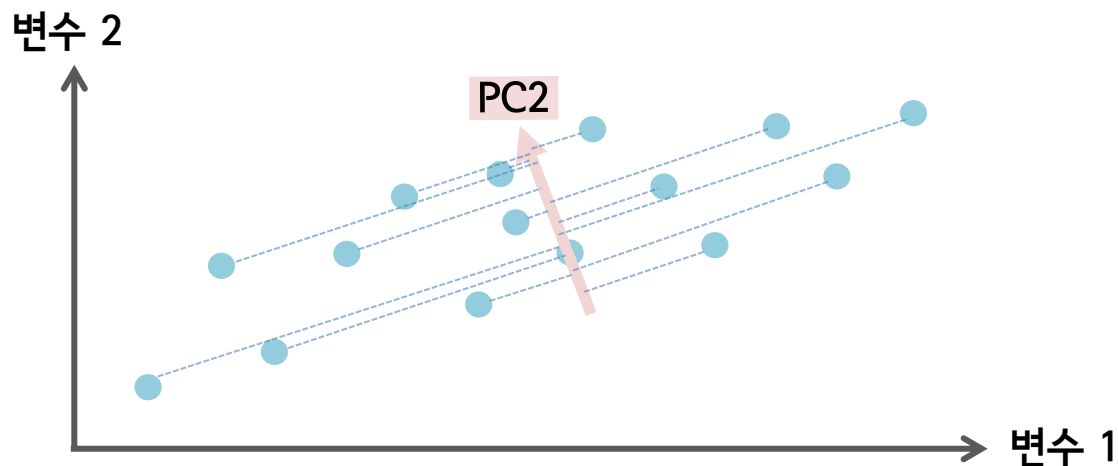
지난주에 살펴본 **정사영** 개념 등장!



이 두 성분 중 데이터를 더 잘 설명하는 것은?



## 주성분 분석(PCA) 개념



## 주성분 선택



PC 구하는 방법

공분산  
행렬

고유벡터 = PC

고유값 = 중요성

## 주성분 선택


 PC 구하는 방법


 공분산  
행렬

고유벡터 = PC

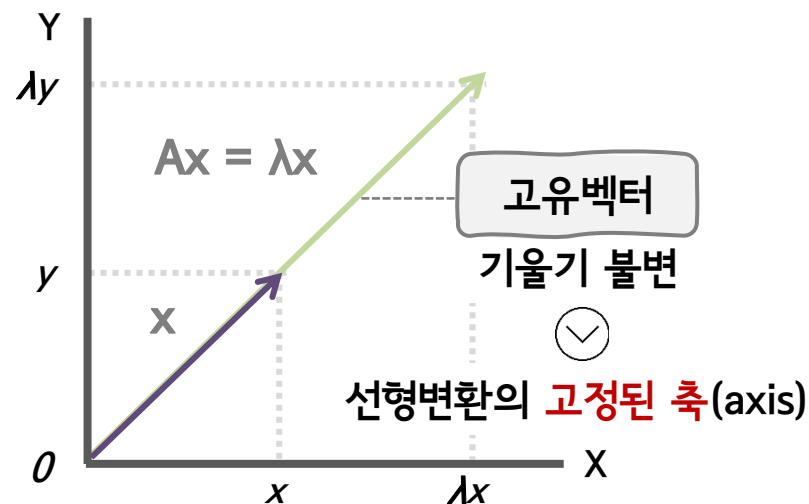
고유값 = 중요성

Q

고유벡터와 고유값을 사용하는 이유는 무엇인가요?

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = 3 \begin{pmatrix} x \\ 1 \end{pmatrix}$$



## 주성분 선택


 PC 구하는 방법


 공분산  
행렬

고유벡터 = PC

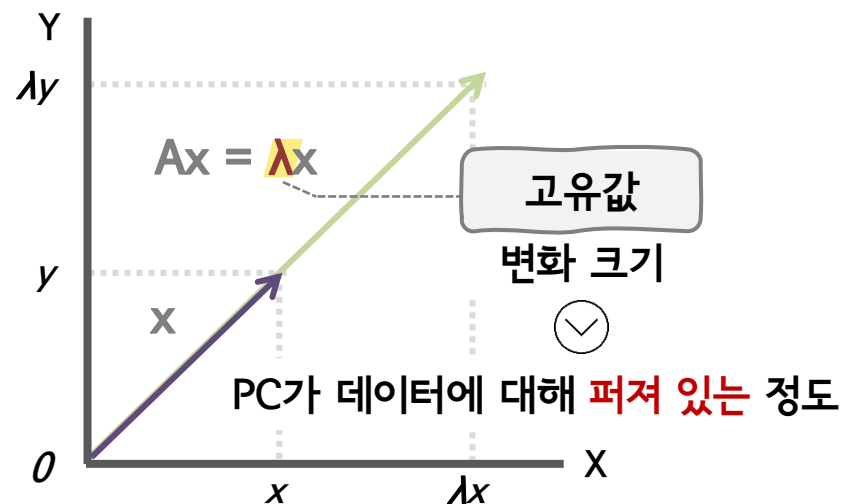
고유값 = 중요성

Q

고유벡터와 고유값을 사용하는 이유는 무엇인가요?

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

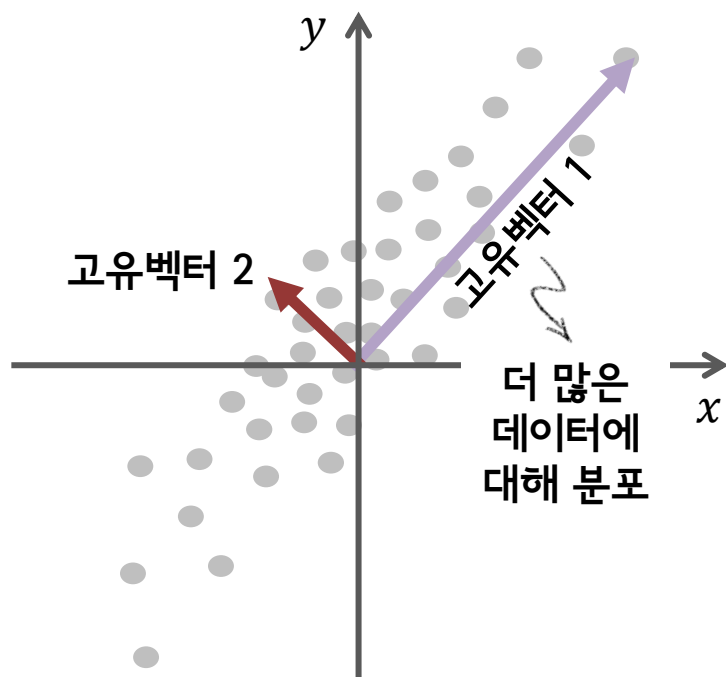


## 주성분 선택


 PC 구하는 방법
공분산  
행렬

고유벡터 = PC

고유값 = 중요성


 공분산 행렬
 
$$\begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix}$$

	중요도 高	
고유값	5.56	1.43

고유벡터	$\begin{bmatrix} 0.615 \\ 0.788 \end{bmatrix}$	$\begin{bmatrix} -0.788 \\ 0.615 \end{bmatrix}$
------	--	---

## 주성분 선택



### PC 개수 구하는 방법

PCA 결과의 summary

누적비율

Scree plot

엘보 포인트

## 주성분 선택

## PC 개수 구하는 방법

PCA 결과의 summary

누적비율

Scree plot

엘보 포인트

&gt; SUMMARY(PCA)

: 일반적으로 누적비율이 80% 정도에 이르는 주성분 개수 선택

	PC1	PC2	PC3	PC4
Standard deviation	1.7084	0.9560	0.38309	0.14393
Proportion of Variance	0.7296	0.2285	0.03669	0.00518
Cumulative Proportion	0.7296	0.9581	0.99482	1.00000

중요도 순

원 데이터의 99.48% 설명



변동의 99.5% 정도를  
설명할 수 있었으면 좋겠어요

4차원 데이터

차원 축소

3차원 데이터



## 주성분 선택

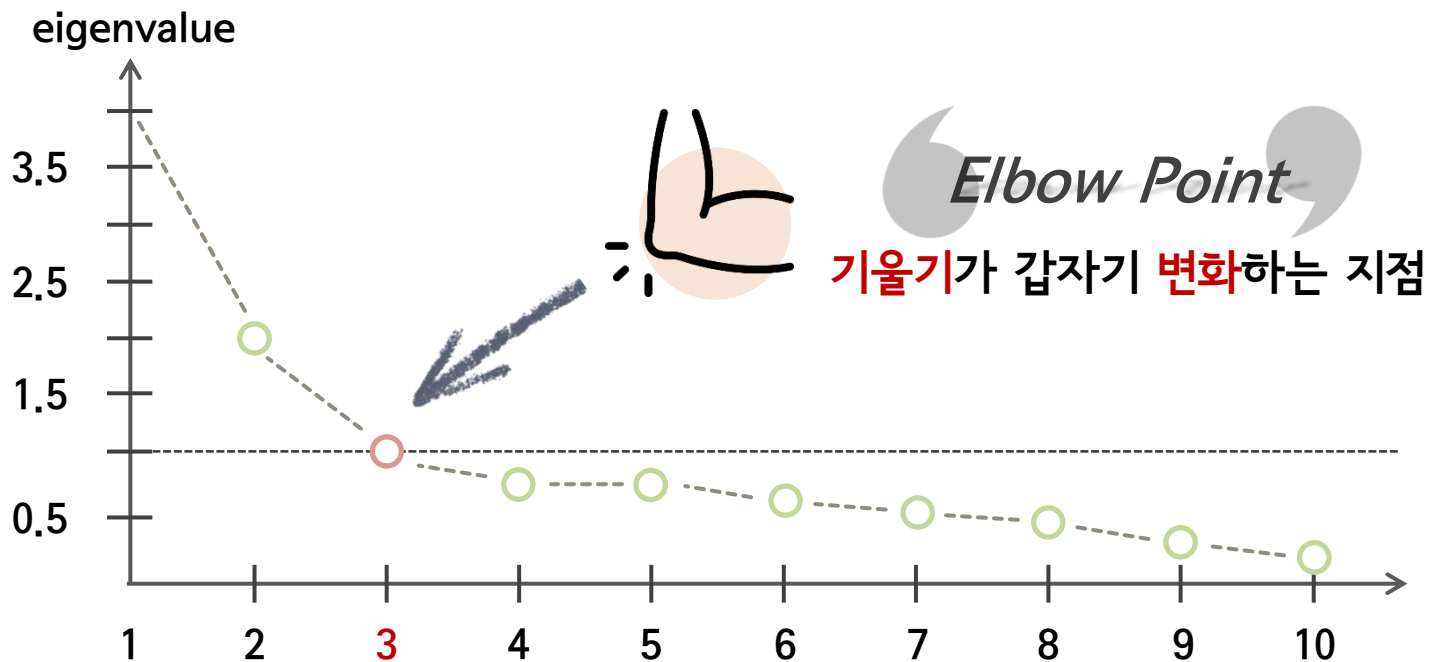
## PC 개수 구하는 방법

PCA 결과의 summary

누적비율

Scree plot

엘보 포인트

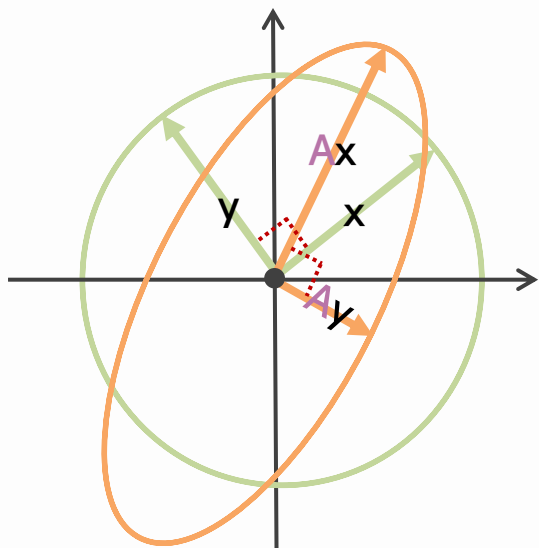


# 4

특이값 분해(SVD)

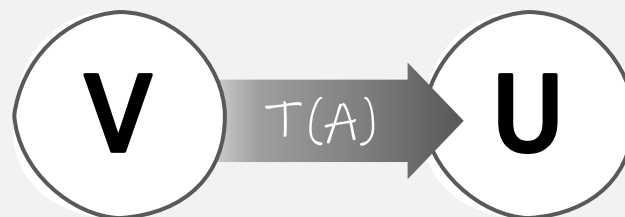
## 특이값 분해(SVD) 이해하기 : 정방행렬이 아닌 행렬에 고유값 분해(EVD)와 유사한 처리

기하학적 의미



직교하는 벡터 집합에 대하여  
선형변환 후에도  
여전히 직교하는 직교집합 찾기

수식적 의미



직교

직교 + 크기 변화( $\Sigma$ )

$$A (V_1 \ V_2) = (U_1 \ U_2) \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}$$

$$A V = U \Sigma$$

$$A V V^{-1} = U \Sigma V^{-1}$$

$$A = U \Sigma V^{-1}$$

## 특이값 분해(SVD) 이해하기

### 1 개념

- $m \times n$  크기의 일반적인 행렬에 대해 고유값 분해와 유사한 처리

$$\begin{array}{ccccccc}
 \begin{array}{c} n \\ m \end{array} \mathbf{A} & = & \begin{array}{c} m \\ m \end{array} \mathbf{U} & \times & \begin{array}{c} n \\ m \end{array} \mathbf{\Sigma} & \times & \begin{array}{c} n \\ n \end{array} \mathbf{V}^T \\
 m \times n \text{ 행렬} & & \text{선형변환 후의} & & \text{크기 변화} & & \text{선형변환 전의} \\
 & & m \times m \text{ 직교행렬} & & \text{나타내는} & & n \times n \text{ 직교행렬} \\
 & & & & m \times n \text{ 대각행렬} & &
 \end{array}$$

## 특이값 분해(SVD) 이해하기

## 1 개념



직교행렬(orthogonal matrix)이란?

$VV^T = V^TV = I$ 를 만족하는 행렬, 즉  $V^T = V^{-1}$ 인 행렬

$$\cancel{VV^T = I} \quad \begin{matrix} & V & \\ \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} & = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

## 특이값 분해(SVD) 이해하기

## 1 개념



직교행렬(orthogonal matrix)이란?

$VV^T = V^TV = I$ 를 만족하는 행렬, 즉  $V^T = V^{-1}$ 인 행렬



즉,  $A = U\Sigma V^T$ 은 곧  $A = U\Sigma V^{-1}$

$$V = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}, V^T = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

## 특이값 분해(SVD) 이해하기

## 1 개념

Q

대각행렬의 크기가 어떻게  $m \times n$  인가요? $m > n$ 

$$\begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & \sigma_n \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

 $m < n$ 

$$\begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & \sigma_m & 0 & \dots & 0 \end{bmatrix}$$

나 말하는 거야???!!

$$\begin{matrix} n \\ \bullet \quad \bullet \\ m \end{matrix} \Sigma$$

## 특이값 분해(SVD) 이해하기

## 2 활용

분해된 여러 개의 행렬을 다시 조합하는 과정에 주로 이용

The diagram shows the matrix equation  $U' \Sigma' V'^T = A'$  using colored blocks to represent dimensions and matrix types.

- $U'$** : An orange block with height  $m$  and width  $p$ . A yellow block of the same size is attached to its right side.
- $\Sigma'$** : A light blue block with height  $p$  and width  $p$ . A red block of size  $p \times p$  containing the symbol  $\Sigma'$  is on top. A lightning bolt symbol points from this red block to the text below.
- $V'^T$** : A block with height  $n$  and width  $p$ . The top part is orange and the bottom part is light green.
- $A'$** : An orange block with height  $m$  and width  $n$ .

White 'X' and '=' symbols indicate the matrix multiplication and equality. Below the  $\Sigma'$  block, the text reads: "대각원소(특이값) 中 상위  $p$ 개만 선택" (Select only the top  $p$  diagonal elements (singular values)).



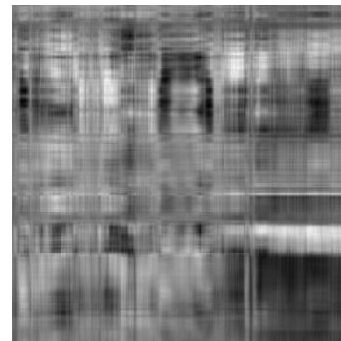
## 특이값 분해(SVD) 이해하기

## 3 이미지에 활용

정사각 사진의  
화질을 조정해보자!

$$\begin{array}{r} 2530 \\ \times \\ 2530 \end{array}$$

행렬 표현을 위해  
흑백 사진으로 변경



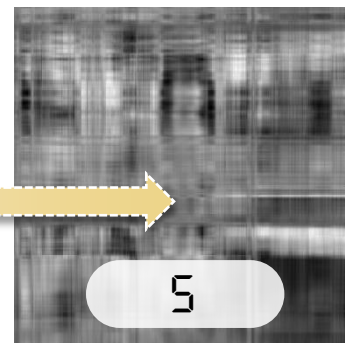
## 특이값 분해(SVD) 이해하기

## 3 이미지에 활용

정사각 사진의  
화질을 조정해보자!

$$\begin{array}{r} 2530 \\ \times \\ 2530 \end{array}$$

SVD를 통해  
적은 개수의  
**singular value**로  
사진 복원



## 특이값 분해(SVD) 이해하기

## 3 이미지에 활용

정사각 사진의  
화질을 조정해보자!

$$\begin{array}{r} 2530 \\ \times \\ 2530 \end{array}$$

SVD를 통해  
적은 개수의  
**singular value**로  
사진 복원

더 적은 정보로,  
더 적은 용량으로,  
더 빠른 성능으로  
분석 가능!

원본 500 1000

50 10

오이이

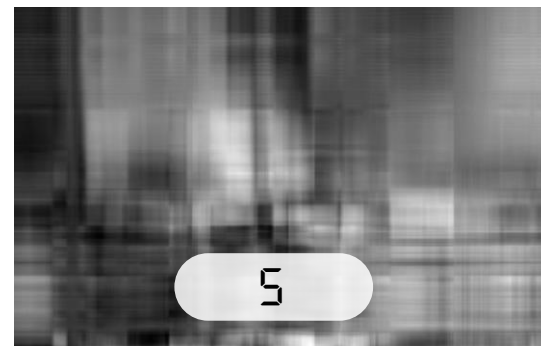
## 특이값 분해(SVD) 이해하기

### 3 이미지에 활용

직사각 사진의  
화질을 조정해보자!

$$\begin{array}{r} 952 \\ \times \\ 607 \\ \hline \end{array}$$

SVD를 통해  
적은 개수의  
**singular value**로  
사진 복원



# 5

## 잠재요인분석(LSA)

## 개념

## 잠재요인분석

*Latent Semantic Analysis*

자연어 처리에서 문서 집합의 추상적 주제를 발견하고자  
사용하는 통계적 모델 (토픽모델링)



## 예시



- /// Pizza
- /// Pizza Hamburger  
Cookie
- /// Hamburger
- /// Ramen
- /// Sushi
- /// Ramen sushi

단어-문서 행렬 A

문장	Pizza	Pizza Hamburger Cookie	Hamburger	Ramen	Sushi	Ramen Sushi
Pizza	1	1	0	0	0	0
Hamburger	0	1	1	0	0	0
Cookie	0	1	0	0	0	0
Ramen	0	0	0	1	0	1
Sushi	0	0	0	0	1	1

단어

$$A = U \Sigma V^T$$

특이값 분해(SVD) 진행

## 예시

$$A = U \Sigma V^T$$

U

	t1	t2	t3	t4	t5
w1	0.6	0	0	0.7	-0.3
w2	0.6	0	0	-0.7	-0.3
w3	0.5	0	0	0	0.9
w4	0	0.7	-0.7	0	0
w5	0	0.7	0.7	0	0

토픽을 위한 단어 행렬

V<sup>T</sup>

	d1	d2	d3	d4	d5	D6
t1	0.3	0.9	0.3	0	0	0
t2	0	0	0	0.4	0.4	0.8
t3	0	0	0	-0.7	0.7	0
t4	0.7	0	-0.7	0	0	0
t5	-0.6	0.5	-0.6	0	0	0
t6	0	0	0	-0.6	-0.6	0.6

토픽을 위한 문장 행렬



예시

$$A = U \Sigma V^T$$

 $\Sigma$ 

	t1	t2	t3	t4	t5	t6
t1	1.9	0	0	0	0	0
t2	0	1.7	0	0	0	0
t3	0	0	1	0	0	0
t4	0	0	0	1	0	0
t5	0	0	0	0	0.5	0

토픽의 강도

## 예시

$$A = U \Sigma V^T$$

U

	t1	t2	t3	t4	t5
w1	0.6	0	0	0.7	-0.3
w2	0.6	0	0	-0.7	-0.3
w3	0.5	0	0	0	0.9
w4	0	0.7	-0.7	0	0
w5	0	0.7	0.7	0	0

 $\Sigma$ 

	t1	t2	t3	t4	t5	t6
t1	1.9	0	0	0	0	0
t2	0	1.7	0	0	0	0
t3	0	0	1	0	0	0
t4	0	0	0	1	0	0
t5	0	0	0	0	0.5	0

상위 2개의  
특이값만 남김

 $V^T$ 

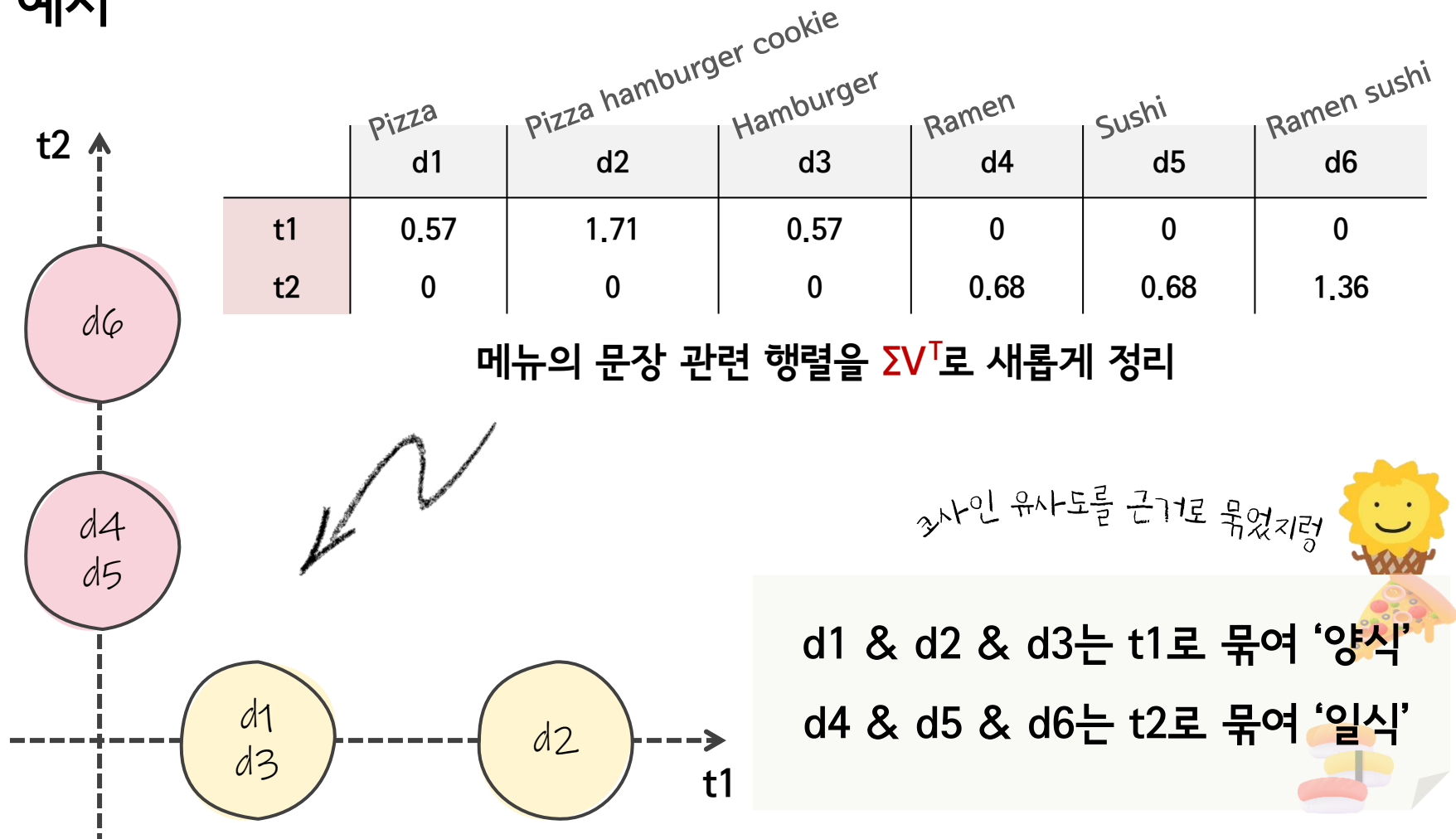
	d1	d2	d3	d4	d5	D6
t1	0.3	0.9	0.3	0	0	0
t2	0	0	0	0.4	0.4	0.8
t3	0	0	0	-0.7	0.7	0
t4	0.7	0	-0.7	0	0	0
t5	-0.6	0.5	-0.6	0	0	0
t6	0	0	0	-0.6	-0.6	0.6

토픽을 위한 단어 행렬

토픽의 강도

토픽을 위한 문장 행렬

## 예시



# 6

## 계층화분석법(AHP)

## 개념



인간의 의사결정은 계층적이고 상대적인 원칙을 따른다



## 계층화분석법

Analytic Hierarchy Process

의사결정문제가 다수의 평가 기준으로 이루어져 있을 때,  
평가 기준을 계층화한 뒤 이에 따라 중요도를 정해가는  
다기준 의사결정기법

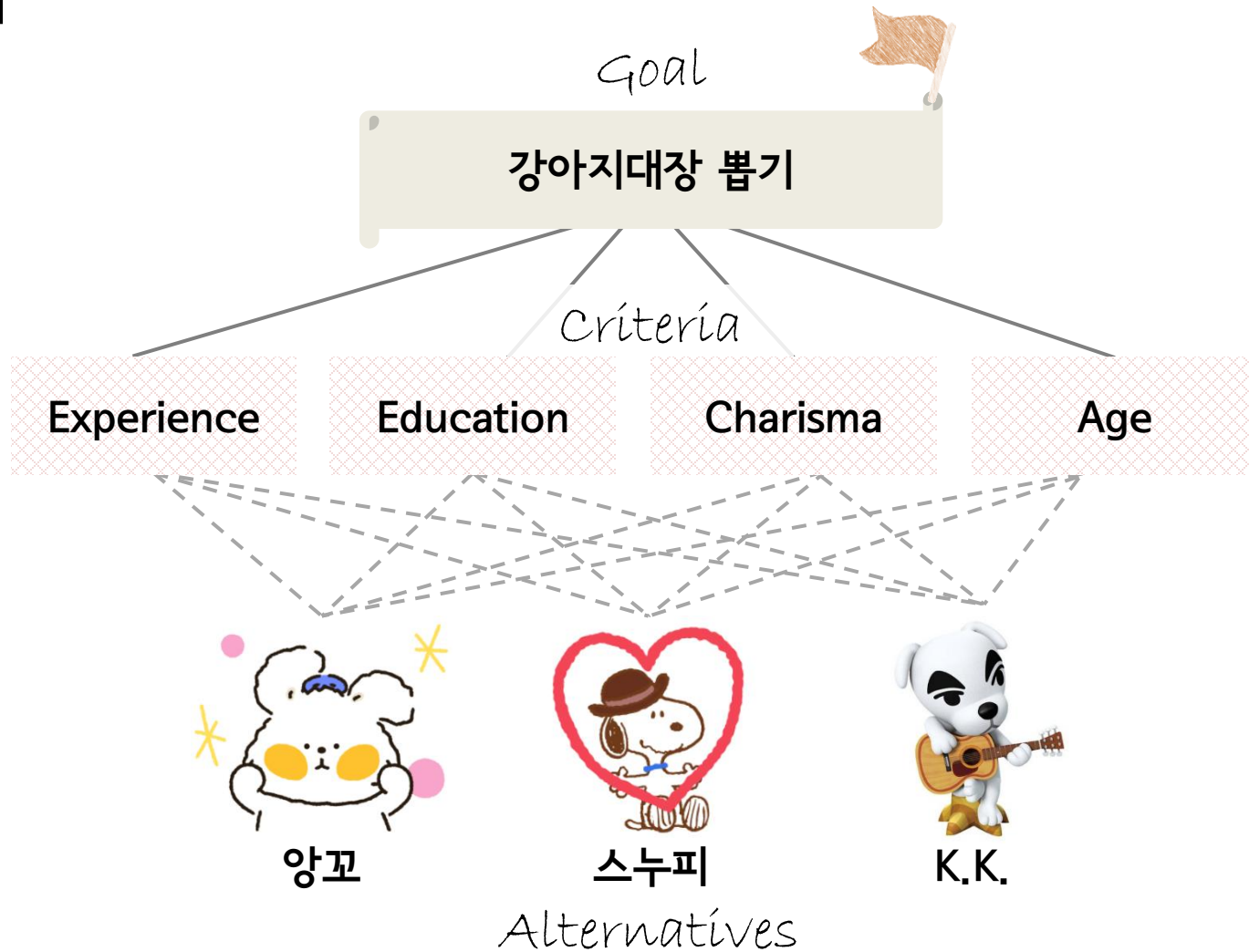
쌍대비교의 반복

행렬을 이용한  
단계적 가중치 산정법

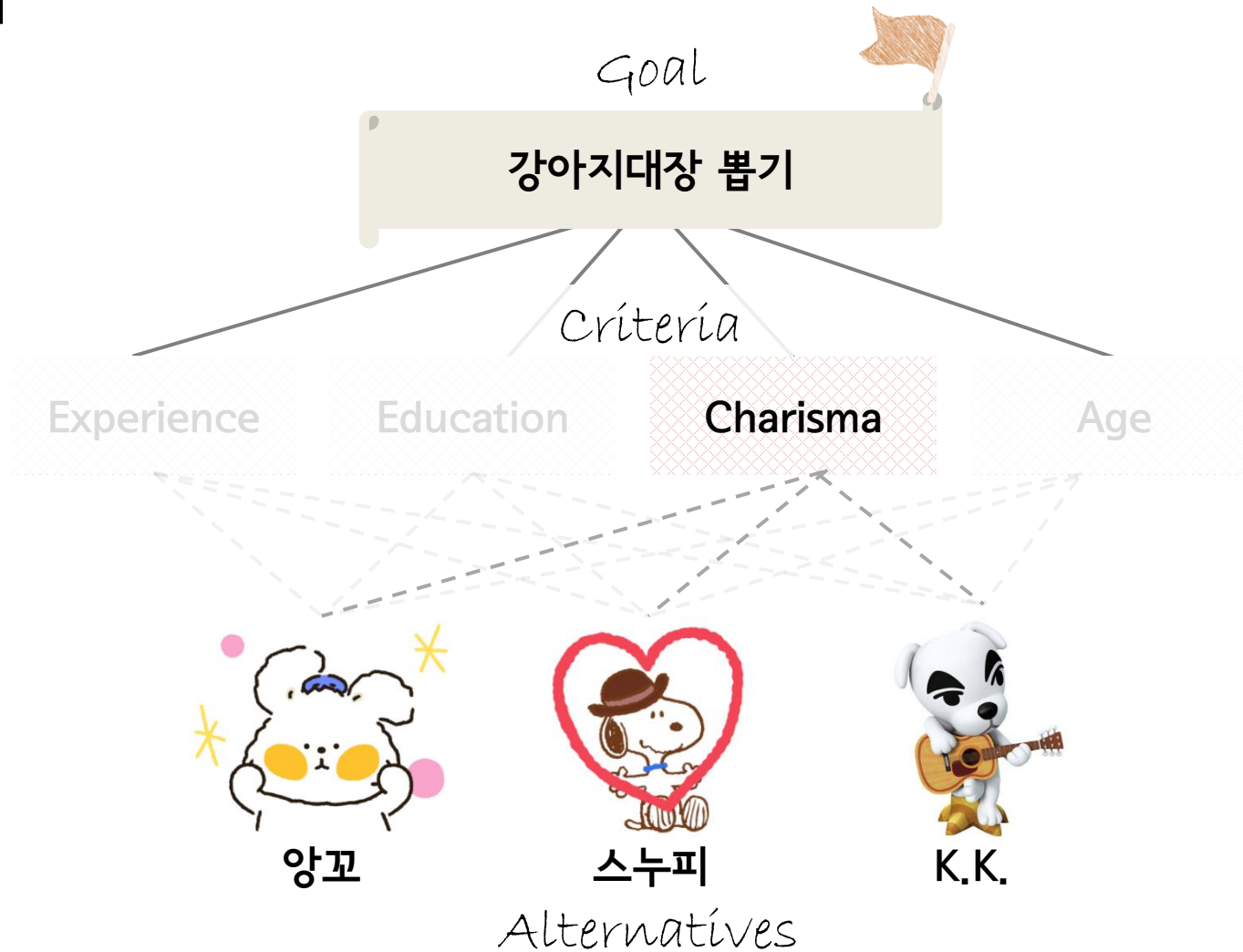
포인트



## 예시



## 예시



## 예시

‘**상대 비교**’두 개의 factor를 **상호 비교**

A와 B의 비교

동시에

B와 A의 비교



앙꼬



스누피



K.K.

Alternatives



## 예시

Charisma	앙꼬	스누피	K.K.
앙꼬	1	5	9
스누피	1/5	1	4
K.K.	1/9	1/4	1

앙꼬가 K.K.보다  
9만큼 더  
카리스마 있다!

3번의 비교

Charisma



앙꼬



스누피



K.K.

Alternatives

## 예시

두 행렬의 곱						행간의 합	각 행의 비율
1	5	9	1	5	9	53.24	0.75
1/5	1	4	×	1/5	1	13.64	0.19
1/9	1/4	1		1/9	1/4	4.31	0.06

## 예시

두 행렬의 곱						행간의 합	각 행의 비율			
1	5	9	1	5	9	53.24	0.75			
1/5	1	4	×	1/5	1	4	=	13.64	>	0.19
1/9	1/4	1		1/9	1/4	1		4.31		0.06

## 기준에 대한 가중치

Experience	0.547
Education	0.127
Charisma	0.270
Age	0.056

∴ 최종 점수가 가장 높은 스누피가 바로 강아지대장!

	Experience	Education	Charisma	Age	Goal
양꼬	0.119	0.024	0.201	0.015	0.358
스누피	0.392	0.010	0.052	0.038	0.492
K.K.	0.036	0.093	0.017	0.004	0.149
Totals	0.547	0.127	0.270	0.056	1.000

감사함멍





**THANK YOU**

