14-1. 고유값과 고유벡터

고유값, 고유벡터

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \lambda\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

$$\begin{cases} 1a_1 + 2a_2 = \lambda a_1 \\ 3a_1 + 4a_2 = \lambda a_2 \end{cases}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \underbrace{\lambda}_{\text{고유값}}\underbrace{\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}}_{\text{고유벡터}}$$

주어지는 행렬은 항상 NxN 행렬
행 또는 열 개수에 맞는 고유값과 고유벡터가 계산됨
상수값은 고유값이며, 고유값에 대해 만족하는 벡터는 고유벡터라고 함

고유값, 고유벡터

1

$$\begin{pmatrix} -10 & 6 \\ -18 & 11 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \lambda \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

2

$$\begin{cases} -10a_1 + 6a_2 = \lambda a_1 \\ -18a_1 + 11a_2 = \lambda a_2 \end{cases}$$

3

$$\begin{cases} (-10-\lambda)a_1 + 6a_2 = 0 \\ -18a_1 + (11-\lambda)a_2 = 0 \end{cases}$$

*6

4

$$(-10-\lambda)(11-\lambda)a_1 + 6(11-\lambda)a_2 = 0$$

$$-\quad (-18)(6)a_1 + 6(11-\lambda)a_2 = 0$$

$$\{(-10-\lambda)(11-\lambda) - (-18)(6)\}a_1 = 0$$

고유값, 고유벡터

$$\lambda^2 - \lambda - 110 + 108 = 0$$

$$\lambda^2 - \lambda - 2 = 0 \Rightarrow (\lambda - 2)(\lambda + 1) = 0$$

3

$$\begin{cases} (-10-\lambda)a_1 + 6a_2 = 0 \\ -18a_1 + (11-\lambda)a_2 = 0 \end{cases}$$

3식에 람다=2 대입

3식에 람다=1 대입

$\lambda = 2$를 입력, $\begin{cases} -12a_1 + 6a_2 = 0 \\ -18a_1 + 9a_2 = 0 \end{cases}$

$$\downarrow$$

$$\begin{cases} -2a_1 + 1a_2 = 0 \\ -2a_1 + 1a_2 = 0 \end{cases}$$

$\lambda = -1$ 입력, $\begin{cases} -9a_1 + 6a_2 = 0 \\ -18a_1 + 12a_2 = 0 \end{cases}$

$$\downarrow$$

$$\begin{cases} 3a_1 - 2a_2 = 0 \\ 3a_1 - 2a_2 = 0 \end{cases}$$

고유벡터

- 상수값만을 곱하여 설명할 수 있도록 하는 새로운 벡터를 찾는 것
- 복잡한 기하변환 연산을 간단하게 바꿔줄 수 있는 새로운 좌표계를 정의

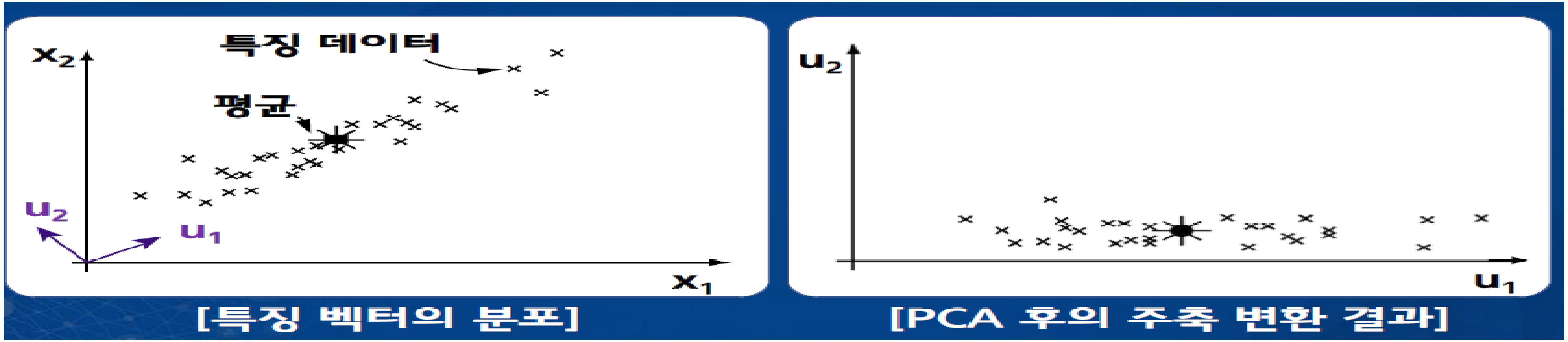
14-2. PCA를 통한 차원축소

차원축소

- 주성분 분석법(PCA) 선형판별 분석법(LDA) 등 존재
- 벡터의 차원이 높아짐에 따라 생길 수 있는 문제 – 차원의 저주
- 연산 시간 증가 및 Visualization하지 못함

주성분 분석(PCA)

- 데이터를 구성하는 특징 벡터의 각 성분들 중에 주성분을 찾아 특징 벡터를 재배치 하는 것
- 데이터의 분포들에 대한 공분산 행렬과 그에 대한 고유값과 고유 벡터를 구함
- 2개의 고유벡터 중 하나는 데이터를 굉장히 잘 표현 하는 고유벡터임
- 프로젝션 시켰을 때 데이터의 분포가 넓게 흩어져 있음(u1축을 기준으로 프로젝션)
- 2차원에서 1차원으로 차원 축소



주성분 분석(PCA)

Q) 2차원 평면상의 5개의 점 (0, 5), (1, 4), (2, 3), (3, 2), (4, 1)를 PCA 방법으로 차원축소 하시오.

① 공분산 행렬 ∑를 계산

	X	Y	X편차	Y편차	X분산	Y분산	공분산
	0	5	-2	2	4	4	-4
	1	4	-1	1	1	1	-1
	2	3	0	0	0	0	0
	3	2	1	-1	1	1	-1
	4	1	2	-2	4	4	-4
평균	*2*	*3*	*0*	*0*	*2*	*2*	*-2*

$$\Sigma = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

공분산 = X편차 * Y편차
공분산 행렬 = [X분산 평균 공분산 평균]
[공분산 평균 Y분산 평균]
공분산 행렬의 고유벡터를 구하면
그 고유벡터들은 무조건 직교함

주성분 분석(PCA)

② 공분산 행렬의 고유값 계산

$$\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}$$

$(2-\lambda)x - 2y = 0$ ...... ①

$-2x + (2-\lambda)y = 0$ ...... ②

$2(2-\lambda)x - 4y = 0$ ...... ①×2

$-2(2-\lambda)x + (2-\lambda)(2-\lambda)y = 0$ ...... ②×(2-λ)

② 공분산 행렬의 고유값 계산

$2(2-\lambda)x - 4y = 0$ ...... ①×2

$-2(2-\lambda)x + (2-\lambda)(2-\lambda)y = 0$ ...... ②×(2-λ)

=====================================

$(\lambda^2 - 4\lambda)y = 0$

$\lambda_1 = 0,\ \lambda_2 = 4$

② 공분산 행렬의 고유벡터 계산

$\lambda_1 = 0$일 때,

$(2-\lambda)x - 2y = 0$

$x = y$ 이므로,

$EV_1 = (1, 1)$

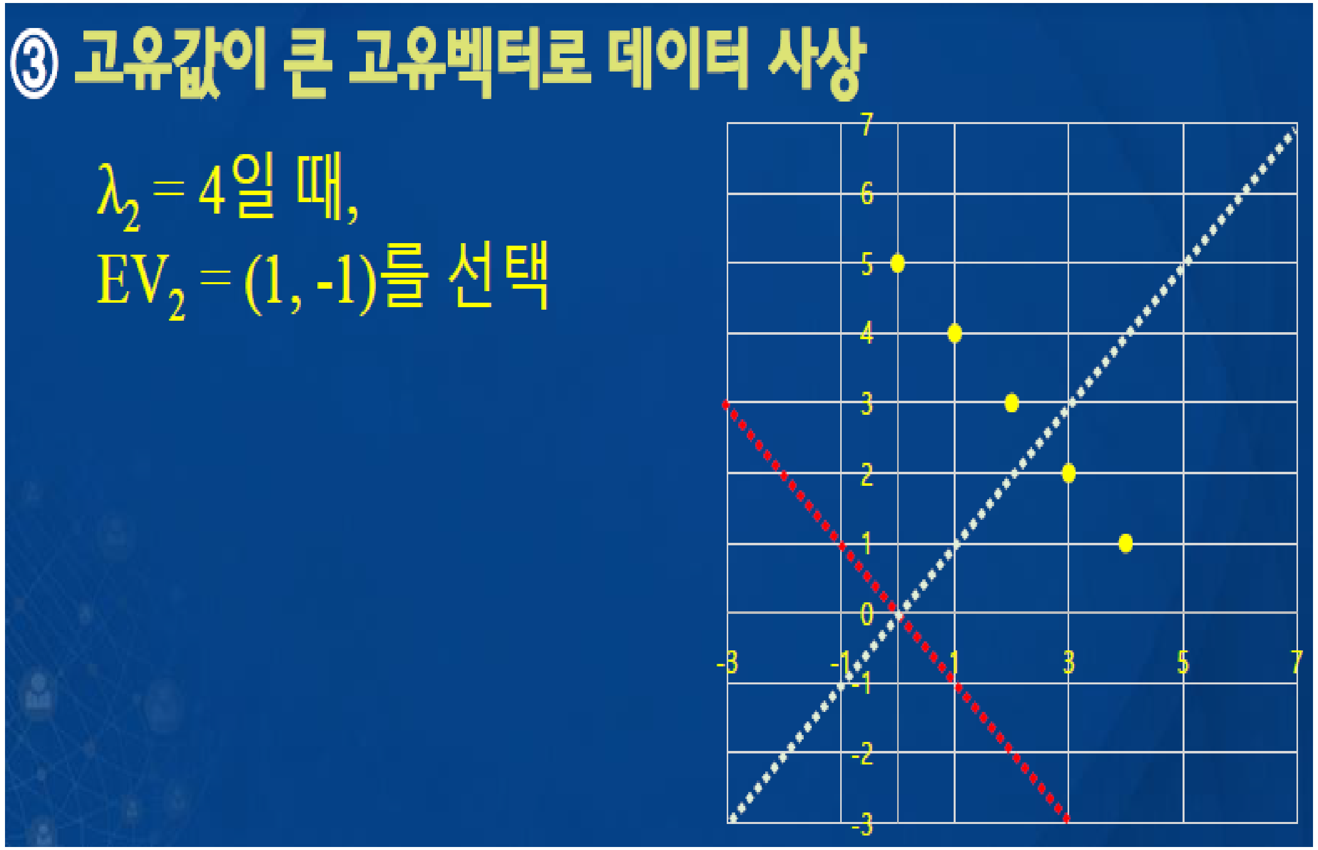
$\lambda_2 = 4$일 때,

$(2-\lambda)x - 2y = 0$

$x = -y$ 이므로,

$EV_2 = (1, -1)$

주성분 분석(PCA)



첫번째 고유벡터 : 흰색 line (고유값 0)
두번째 고유벡터 : 빨간색 line (고유값 4)

데이터를 두번째 고유벡터(고유값이 큰 고유벡터)에 프로젝션 시키면 원래 데이터의 거리가 그대로 유지된다.

공분산 행렬의 고유값의 의미

- 주어진 데이터를 고유벡터 상으로 프로젝션 시켰을 때 데이터들의 분산값
- 한 점으로 다 모임(편차, 분산 0)

주성분 분석(PCA) 한계

주성분 분석 효과적인 데이터 분포

주성분 분석 효과적이지 못한 데이터 분포

활용 분야 : 얼굴인식
30x30 픽셀 = 900, 픽셀마다 0~255개의 값을 가짐(900차원)
900쌍의 eigenvalue를 크기 순으로 sorting하여 eigenvalue 값이 큰 순서로 20~23개 선택(차원축소)
900차원과 성능 차이가 크게 나지 않음

PCA : 주어진 데이터를 한 덩어리로 놓고, 데이터의 분포를 잘 표현하는 축을 찾음
LDA : 주어진 데이터의 클래스를 고려하여 클래스가 잘 구분되는 축을 찾음