머신러닝/딥러닝을 위한 수학

선형대수

고유값과 고유벡터
14강

딥러닝호형

고유값과 고유벡터

선형 변환(Linear Transformation)

$a \to L(a)$ 로 매핑한다고 표현

V, W가 벡터 공간이고 함수 $L : V \to W$이 다음을 만족하면 선형 변환이라고 한다.

($\boldsymbol{a}, \boldsymbol{b}$는 V에서의 벡터, r은 실수)

(1) $L(\boldsymbol{a} + \boldsymbol{b}) = L(\boldsymbol{a}) + L(\boldsymbol{b})$

(2) $L(r\boldsymbol{a}) = rL(\boldsymbol{a})$

$L : V \to W$이 선형 변환일 경우 다음을 만족한다. ($\boldsymbol{a}, \boldsymbol{b}$는 V에서의 벡터)

(1) $L(\mathbf{0}_v) = \mathbf{0}_w$

사용됨.

(2) $L(\boldsymbol{a} - \boldsymbol{b}) = L(\boldsymbol{a}) - L(\boldsymbol{b})$

$\therefore L(a-b) = L(a) + L(-b) = L(a) - L(b)$

$V = W$이면 $L : V \to V$를 V에서의 선형 연산(linear operator)라고 한다.

고유값과 고유벡터

x_1 x_2를 넣었을 때 $-x_2$ x_1으로 변환하는 함수

함수! 그냥 x_2를 위로 - 붙여서 올려주고 x_1을 하단으로 내려줌

예. $L:\mathbb{R}^2 \to \mathbb{R}^2$에 대해서 $L\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$라고 하자.

?

(1)

a, b 각 성분끼리 결합

-1 붙여서

$$\underbrace{L\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right.}_{a} + \underbrace{\left.\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right)}_{b} = L\left(\begin{pmatrix} x_1+y_1 \\ x_2+y_2 \end{pmatrix}\right) = \begin{pmatrix} -x_2-y_2 \\ x_1+y_1 \end{pmatrix}$$

이런형태로 come.

$$L\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) + L\left(\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} + \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix} = \begin{pmatrix} -x_2-y_2 \\ x_1+y_1 \end{pmatrix}$$

(2)

$$L\left(r\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} -rx_2 \\ rx_1 \end{pmatrix}$$

$$rL\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = r\begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} -rx_2 \\ rx_1 \end{pmatrix}$$

고유값과 고유벡터

고유값(eigenvalue)와 고유벡터(eigenvector)

$L : V \to V$이 n차원 벡터 공간 V에서의 선형 변환일 때, 영이 아닌 벡터 $\boldsymbol{x}$가 존재해서 $L(\boldsymbol{x}) = \lambda \boldsymbol{x}$를 만족한다면 λ를 고유값이라고 하며 이 때 방정식을 만족하는 영이 아닌 모든 벡터 $\boldsymbol{x}$에 대해서 고유벡터라고 한다.

↳ 있을수도 없을수도 있음.

예. 선형변환식 $L : \mathbb{R}^2 \to \mathbb{R}^2$에 대해서 $L\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$일 때,

$$L\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \left.\begin{aligned} -x_2 &= \lambda x_1 \\ x_1 &= \lambda x_2 \end{aligned}\right) \text{ 같은 성분끼리 비교}$$

고유값과 고유벡터

$$-x_2 = \lambda x_1, x_1 = \lambda x_2$$

대입

$$x_1 = -\lambda^2 x_1 \rightarrow x_1(\lambda^2 + 1) = 0 \rightarrow \lambda^2 + 1 = 0(if\, x_1 \neq 0)$$

$$\therefore \lambda = \pm i$$

$L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$이므로 위 식을 만족하는 벡터 $\boldsymbol{x}$는 없다.

선형변환은 실수의 2차원에서 일어나므로
만약 문제에서 $L: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ 처럼 복소수로 정의됐다면
i가 답이 될 수 있음.

고유값과 고유벡터

보통 $Ax=\lambda x$ 인 x를 구하라고함

예. 선형변환식 $L:\mathbb{R}^2 \to \mathbb{R}^2$에 대해서 $L\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$$\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \begin{matrix} x_1 & x_2 \\ 0 & -x_2 \end{matrix}$$

$$\to \begin{pmatrix} x_1 + x_2 \\ -x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \begin{matrix} \lambda x_1 \\ \lambda x_2 \end{matrix}$$

$$\to \begin{pmatrix} -\lambda + 1 & 1 \\ 0 & -\lambda - 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

↳ 고유값, 고유벡터를 구할때 1번째는 0이 아님

$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$가 0이 아님 근데 0이나옴 → 만약 그러면 → 이 행렬이

역행렬이 존재하면? $Px=0 \to P^{-1}Px = P^{-1}0 \to x=0$

단위행렬: 항등행렬 다 1인.

$$\to (\lambda - 1)(1 + \lambda) = 0$$

행렬식 만듦. → $f(\lambda)$임

$$\therefore \lambda = \pm 1$$

1) $\lambda = 1$인 경우

$$x_1 = 1, x_2 = 0$$

$$\therefore \boldsymbol{x} = \begin{pmatrix} s \\ 0 \end{pmatrix}$$

2) $\lambda = -1$인 경우

$$x_2 = -2x_1$$

$$\therefore \boldsymbol{x} = \begin{pmatrix} p \\ -2p \end{pmatrix}$$

(s, p는 모든 실수)

고유값과 고유벡터

A가 $n \times n$ 행렬일 때, $\lambda I_n - A$를 A의 characteristic polynomial이라고 한다.

이 때 $p(\lambda) = det(\lambda I_n - A) = 0$을 A의 characteristic equation이라고 한다.

예.

$$p(\lambda) = \lambda^2 - 1 = 0$$

$$p(\lambda) = \lambda^3 - 4\lambda^2 + \lambda + 2 = 0$$

우리의 목표: $AX = \lambda x$를 찾는거

$Ax - \lambda x = 0 \rightarrow (\lambda I_n - A)x = 0$

I_n을 끄집어내는 이유 → λ가 실수라서

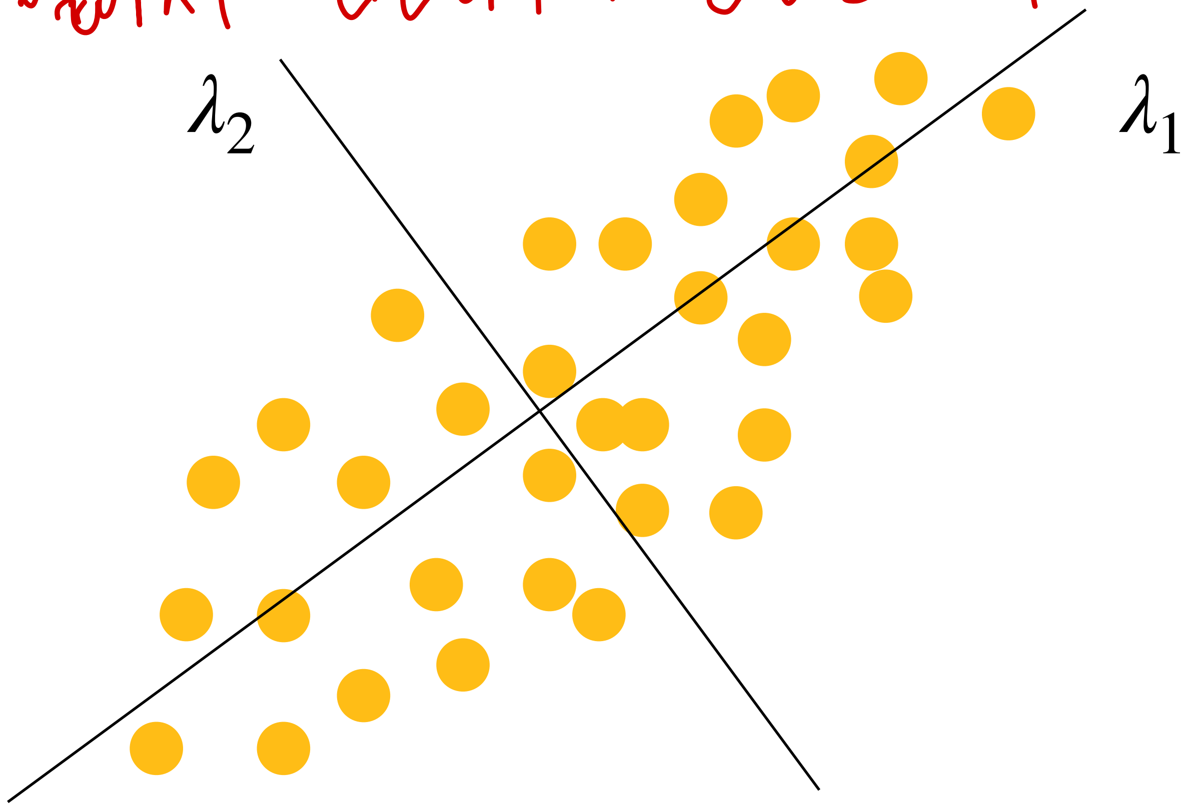
x는 0이 아님 $\therefore \lambda I_n - A = 0$.

고유값과 고유벡터

머신러닝/딥러닝: 어디서 사용되나요? 뒤에 나옴

고유값과 고유벡터는 차원 축소법인 PCA, 특이값 분해인 SVD 등 매우 다양한 방법에서 활용되며 축소법 및 분해법은 머신러닝에서 빼놓을 수 없는 기술이다.

어떤 행렬의 를 뽑아서 간다. 계산되었을하게



- 주성분 분석(PCA) -