머신러닝/딥러닝을 위한 수학

선형대수

대각화

15강

딥러닝호형

대각화

유사성(Similarity)

역행렬을 갖는 행렬

A, B를 $n \times n$ 행렬이라고 하자. 이 때 nonsingular 행렬 P가 존재해서 $B = P^{-1}AP$이면 "B는 A와 유사(similar)하다"라고 한다.

$$A = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \qquad P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$B = P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

대각화

대각화(Diagonalization)

n차원 벡터 공간 V에서의 선형 연산 $L : V \to V$이 정의($L(\boldsymbol{x}) = A\boldsymbol{x}$) 될 때, V의 basis S가 존재해서 L이 S에 대한 대각 행렬 B로 표현이 된다면 "L을 대각화 할 수 있다.(diagonalizable)"라고 한다. (L: 대각화 가능 선형 변환)

가장 기본이 되는? 행렬

각 열벡터들이

$$P = \begin{pmatrix} \boxed{\begin{matrix}1\\0\\1\end{matrix}} & \boxed{\begin{matrix}0\\1\\0\end{matrix}} & \begin{matrix}1\\1\\0\end{matrix} \end{pmatrix}$$ 열벡터의 모임이 V의 basis S

$$A = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \qquad B = P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

B가 대각행렬이 나옴

→ 대각화 할 수 있다

대각화

유사 행렬 A, B는 동일한 고유값을 갖는다.

$$\text{즉, } p_B(\lambda) = p_A(\lambda)$$

왜?, $P^{-1}I_n = P^{-1}$, $P^{-1}P = I_n$

I_n에 $P^{-1}P$를 곱함 결국 I_n 나옴

$$p_B(\lambda) = det(\lambda I_n - B) = det(\lambda I_n - P^{-1}AP) = det(\lambda P^{-1}I_nP - P^{-1}AP)$$

$$= det(P^{-1}(\lambda I_n - A)P) = \underbrace{det(P^{-1})}_{\text{숫자}}det(\lambda I_n - A)\underbrace{det(P)}_{\text{숫자}}$$

↳ P^{-1}를 앞쪽으로 옮기고 P를 뒤쪽으로 옮김

$$= det(P^{-1})det(P)det(\lambda I_n - A) = det(\lambda I_n - A) = p_A(\lambda)$$

$det(AB) = det(A)det(B)$

$det(P^{-1}) = \frac{1}{det(P)}$

대각화

1. n차원 벡터 공간 V에서의 선형 연산 $L : V \to V$이 정의 될 때 다음은 동치다.

(1) L을 대각화가 가능하다.

(2) V는 L의 고유벡터의 basis S를 갖는다. $L(x) = Ax \therefore D = P^{-1}AP$

→ P가 존재한다 = basis S를 갖는다.

2. 다음은 동치다.

(1) $n \times n$ 행렬 A가 대각 행렬 D와 유사하다.

(2) A는 일차 독립 n개의 고유벡터를 갖는다.

특히, D의 대각 원소들은 A의 고유값이다. $(\lambda_1, \lambda_2 \cdots$

3. $p_A(\lambda)$의 근이 모두 다르면 A는 대각화가 가능하다.

대각화

예. 14강에서 다룬 $A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$의 고유값은 $\lambda = \pm 1$이고 각각의 고유벡터는 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \end{pmatrix}$이다. 즉, $P = \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}, D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$이므로 A는 대각화가 가능하다.

$\therefore D = P^{-1}AP$

$$D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0.5 \\ 0 & -0.5 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix} = P^{-1}AP$$

이걸로 대각행렬

1) $\lambda = 1$인 경우	2) $\lambda = -1$인 경우
$x_1 = 1, x_2 = 0$	$x_2 = -2x_1$
$\therefore \mathbf{x} = \begin{pmatrix} s \\ 0 \end{pmatrix}$	$\therefore \mathbf{x} = \begin{pmatrix} p \\ -2p \end{pmatrix}$

이걸로 P 만듦

대각화

대칭행렬과 대각화

$n \times n$ 행렬 A이 대칭행렬(symmetric matrix)일 때

(1) $p_A(\lambda)$의 근은 모두 실수이다.

(2) A의 서로 다른 고유 벡터는 orthogonal하다.

예. $A = \begin{pmatrix} 0 & 2 \\ 2 & 3 \end{pmatrix}$이면

$$det(\lambda I_n - A) = det\left(\begin{pmatrix} \lambda & -2 \\ -2 & \lambda - 3 \end{pmatrix}\right) = 0 \rightarrow \lambda(\lambda - 3) - 4 = 0 \rightarrow (\lambda - 4)(\lambda + 1) = 0$$

$\lambda = -1$일 때 고유벡터는$\begin{pmatrix} -2 \\ 1 \end{pmatrix}$이고 $\lambda = 4$일 때 고유벡터는$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$이다. (orthogonal)

따라서 $P = \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}, D = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}$이다.

대각화

대칭행렬과 대각화

이 식을 얻어낼 수 있음

$n \times n$ 행렬 A이 $A^T = A^{-1}$일 경우 A를 orthogonal이라고 한다.(즉, $A^TA = I$)

$A^{-1}A = I$

다음은 동치다.

(1) A는 orthogonal이다. (| | |): 어떤 행렬의 행 1열이 만족하면 이 전체행렬은 직교행렬이라는 함 $A^{-1} = A^T$ ∴ $A^TA = I$ 길이가 1인 벡터

(2) A의 행(열)벡터가 orthonormal set을 형성한다.

(orthonormal: orthogonal하고 각각의 크기가 1인 단위벡터이면) ?

두개의 내적하면 0이 나옴

예. $A = \begin{pmatrix} 0 & 2 \\ 2 & 3 \end{pmatrix}$이면 $P = \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}$, $D = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}$이다. $P = \begin{bmatrix} \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$라 내적? 위 아래 + - 하면

↳ 아까 λ값이 P를 만든다는 함 $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$

하면 $D = P^{-1}AP = P^TAP$이다. → $Ax = \lambda x$를 만족하는 것 中 하나 동일 스칼라 곱해도 可

$\begin{pmatrix} -2 \\ 1 \end{pmatrix}$의 크기 = $\sqrt{(-2)^2 + (1)^2}$ ∴ $\sqrt{5}$로 divide

대각화

대각화 순서 - Spectral Decomposition

1. 고유값을 구한다. 즉, $p_A(\lambda) = det(\lambda I_n - A) = 0$
2. n개의 일차독립 고유벡터를 구한다. (없으면 A는 대각화 할 수 없다.)
3. orthogonal 고유벡터를 구한다. (symmetric 행렬인지 or Gram-schmidt)
4. orthonormal 고유벡터를 구한다. (단위벡터 계산) 즉, $P^T = P^{-1}$
5. $D = P^{-1}AP = P^TAP$

이거면 바로 직교성 나오고

이거면 직교화한 애 구하기 가능