머신 러닝을 위한 수학

미적분, 선형대수

8기 교육부 임성연

목 차

미분

- 아주 짧은 순간의 함수에 대한 기울기를 구하는 것
- 도함수 : f’(x) = $\lim_{\Delta x \to \infty} \frac{f(a+\Delta x)-f(a)}{\Delta x}$
- x=a에 대하여 극한값이 실수로 존재하면 f는 x=a에서 미분 가능하며 f’(a)로 표시

미분

- $\frac{d}{dx}x^n = nx^{n-1}$
- $\frac{d}{dx}\ln x = \frac{1}{x}$
- $\frac{d}{dx}e^x = e^x$
- $\frac{d}{dx}(af + bg) = a\frac{df}{dx} + b\frac{dg}{dx}$
- $\frac{d}{dx}(fg) = f\frac{dg}{dx} + \frac{df}{dx}g$
- $\frac{d}{dx}\frac{1}{f} = -\frac{1}{f^2}\frac{df}{dx}$

머신 러닝에서의 활용

- 손실 함수의 값을 최소로 만들기 위해 다양한 기법을 사용
- 손실 함수를 미분하면 어떤 특정 지점에서 어느 정도의 기울기가 나오는지 알 수 있다.

상미분(Ordinary derivative)

- 변수가 하나만 있는 함수의 미분
- y= x^r일 때 $\frac{dy}{dx} = rx^{r-1}$
- $\frac{d}{dx}\{f(x) + g(x)\} = \frac{df(x)}{dx} + \frac{d(x)}{dx}$
- $\frac{d}{dx}\{kf(x)\} = k\frac{df(x)}{dx}$

편미분(Partial derivative)

- y를 상수라 생각하고 고정시킨 다음, x에 대해 편미분 : $\frac{\partial f(x,y)}{\partial x}$, f_x(x, y)
- X를 상수라 생각하고 고정시킨 다음, y에 대해 편미분 : $\frac{\partial f(x,y)}{\partial y}$, f_y(x, y)

f(x, y)=3x^2+5xy+3y^3

$\frac{\partial f(x,y)}{\partial x}$=6x+5y

$\frac{\partial f(x,y)}{\partial y}$=5x+9$y^2$

전미분(Total differentiation)

- **f(x, y)의 미분**
- $df = \frac{\partial f(\mathrm{x,y})}{\partial x}dx + \frac{\partial f(\mathrm{x,y})}{\partial x}dy$

f(x, y)=xy+x+y+1

$df(x, y) = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial \mathrm{y}}dy = (y+1)dx+(x+1)dy$

연쇄법칙(Chain Rule)

- **합성함수의 미분**
- $F(x) = (f \circ g)(x)$

  $F'(x) = f'(g(x))\, g'(x)$
- $y=f(u),\ u=g(x)$

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

연쇄법칙(Chain Rule)

- 합성함수의 미분

$$f(x, y) = (3x+1)^2 + (x+y+1)^3 \quad \rightarrow \quad f(x, y) = u^2 + v^3$$

$u = 3x+1$
$v = x+y+1$

$$\begin{aligned}\frac{\partial f(x,y)}{\partial x} &= \frac{\partial f(x,y)}{\partial u} \cdot \frac{\partial u}{\partial x} + \frac{\partial f(x,y)}{\partial v} \cdot \frac{\partial v}{\partial x} \\ &= \frac{\partial u^2}{\partial u} \cdot \frac{\partial u}{\partial x} + \frac{\partial v^3}{\partial v} \cdot \frac{\partial v}{\partial x} \\ &= 2u \cdot 3 + 3v^2 \cdot 1 \\ &= 6(3x+1) + 3(x+y+1)^2 \\ &= 3x^2 + (6y+24)x + 3y^2 + 6y + 9\end{aligned}$$

머신 러닝에서의 활용

- 신경망에서는 학습한 결과로 도출된 답이 정답 데이터에 가까워질 수 있도록 가중치(w)를 조정하는 과정을 반복한다.
- 이때, 실제 정답과 학습 결과 사이의 오차 값을 가중치로 편미분한 다음, 그 값을 가중치의 조정량으로 사용한다.
- 오차역전파법(Backproagation)

벡터(Vector)란?

- 크기와 방향을 갖는 직선
- 벡터공간(Vector space)을 이루는 원소
- 행벡터(횡벡터) : 벡터의 원소를 가로로 표현한 경우
- 열벡터(종벡터) : 벡터의 원소를 세로로 표현한 경우

$$\mathbf{a} = [a_1, a_2, \dots, a_n] \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

덧셈과 뺄셈

- 서로 대응하는 성분끼리 덧셈 또는 뺄셈

$$\begin{bmatrix}1\\2\\3\end{bmatrix}+\begin{bmatrix}4\\5\\6\end{bmatrix}=\begin{bmatrix}1+4\\2+5\\3+6\end{bmatrix}=\begin{bmatrix}5\\7\\9\end{bmatrix}$$

$$\begin{bmatrix}1\\2\\3\end{bmatrix}-\begin{bmatrix}4\\5\\6\end{bmatrix}=\begin{bmatrix}1-4\\2-5\\3-6\end{bmatrix}=\begin{bmatrix}-3\\-3\\-3\end{bmatrix}$$

덧셈과 뺄셈

- 서로 다른 차원의 벡터끼리는 덧셈과 뺄셈을 할 수 없다.

$$\begin{bmatrix}1\\2\\3\end{bmatrix}+\begin{bmatrix}4\\5\\6\\7\end{bmatrix}=\text{계산불가}$$

스칼라배(Scalar multiple)

- 벡터의 모든 성분에 같은 수를 곱하는 것

$$2\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}2\times 1\\2\times 2\\2\times 3\end{bmatrix}=\begin{bmatrix}2\\4\\6\end{bmatrix}$$

머신 러닝에서의 활용

- 언어를 처리하기 위해 단어들을 벡터처럼 취급
- 일반적인 벡터 연산처럼 단어들도 덧셈과 뺄셈이 가능
- ‘왕‘-’남성‘+‘여성’=‘여왕‘
- ‘동경‘-’일본‘+’한국‘=‘서울‘

내적(Inner product)이란?

- 벡터에서 서로 대응하는 성분끼리 곱한 다음 그것들을 모두 더한 값
- 벡터와 벡터의 내적은 스칼라가 된다.
- 서로 다른 차원의 벡터끼리는 계산을 할 수 없다.

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

내적 표기

$$\langle \mathbf{a}, \mathbf{b} \rangle = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

내적(Inner product)이란?

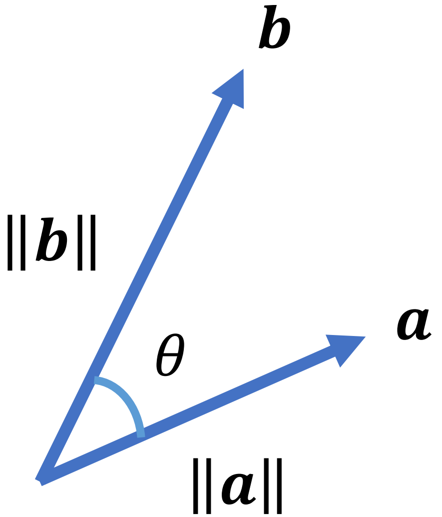
- 벡터에서 서로 대응하는 성분끼리 곱한 다음 그것들을 모두 더한 값
- 벡터와 벡터의 내적은 스칼라가 된다.
- 서로 다른 차원의 벡터끼리는 계산을 할 수 없다.

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \qquad \langle \mathbf{a}, \mathbf{b} \rangle = 1\times4+2\times5+3\times6 = 32$$

기하학적 정의

- 벡터 a와 b가 이루는 각이 θ일 때 <a, b>는 다음과 같이 정의

$$<\mathbf{a}, \mathbf{b}> \; = \|\boldsymbol{a}\|\|\boldsymbol{b}\| \cos \boldsymbol{\theta}$$



기하학적 정의

- 벡터 a와 b가 이루는 각이 θ일 때 <a, b>는 다음과 같이 정의

$$<\mathbf{a}, \mathbf{b}> = \|\boldsymbol{a}\|\|\boldsymbol{b}\|\cos\boldsymbol{\theta}$$

$$<\mathbf{a}, \mathbf{b}> = \|\boldsymbol{a}\|\|\boldsymbol{b}\|\cos\mathbf{45}° = \sqrt{5}\times\sqrt{10}\times\frac{\sqrt{2}}{2} = 5$$

$a = (2, 1)$ $b = (1, 3)$ $\cos 45° = \frac{\sqrt{2}}{2}$

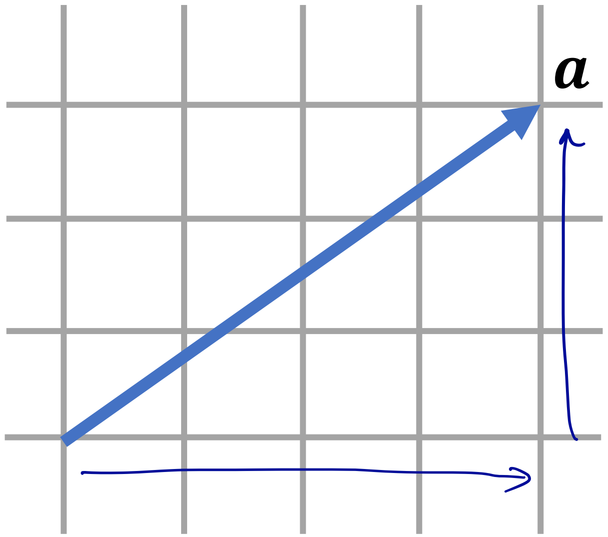
노름(Norm)이란?

- 벡터를 이루는 방향과 이동 거리 중 이동 거리를 의미
- $\|x\| = 0 \leftrightarrow x=0$
- $\|\alpha x\| = |\alpha| \, \|x\|$
- $\|x + y\| \leq \|x\| + \|y\|$

L1 노름

- 벡터 성분의 절대값을 모두 더한 값
- 맨해튼 거리

$$\|\boldsymbol{a}\|_1 = |a_1|+|a_2|+\ldots+|a_n| = \sum_{i=1}^{n} |a_i|$$

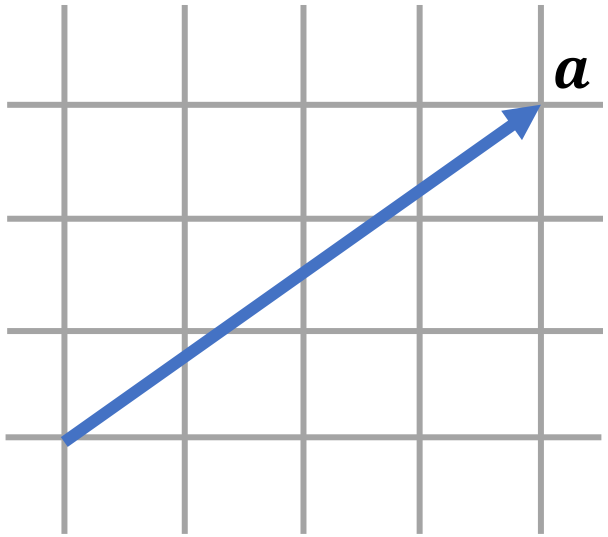


ex) $\|a\|_1 = 4+3=7$

L2 노름

- 시작점에서 목적지까지 직선으로 움직인 거리
- 유클리드 거리

$$\|\boldsymbol{a}\|_2 = \sqrt{\sum_{i=1}^{n} a_i^2} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$
$$= \sqrt{< \boldsymbol{a}, \boldsymbol{a} >}$$



ex) $\|a\|_2 = \sqrt{16+9} = 5$

머신 러닝에서의 활용

- 선형회귀 모델의 정규화 항
- 데이터셋을 학습 데이터와 테스트 데이터로 분류하는 과정에서 과적합
- 정규화 항을 덧붙여 계수의 절대값이나 제곱한 값이 너무 커지지 않도록 만들어줘야 한다.

코사인 유사도

$$\langle \mathbf{a}, \mathbf{b} \rangle = \|\boldsymbol{a}\|\|\boldsymbol{b}\|\cos\boldsymbol{\theta}$$

$$\cos\boldsymbol{\theta} = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|}$$

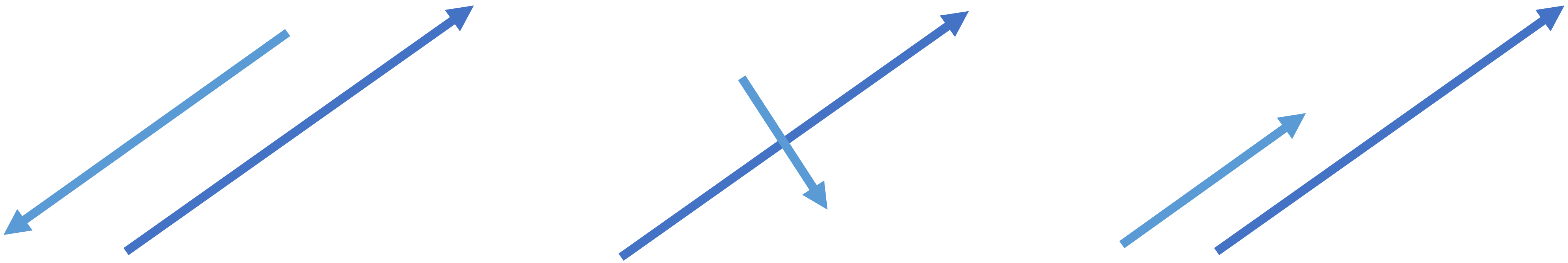
$$\langle \mathbf{a}, \mathbf{b} \rangle = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

$$\|\boldsymbol{a}\|_2 = \sqrt{\sum_{i=1}^{n} a_i^2} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

$$\cos\boldsymbol{\theta} = \cos(\boldsymbol{a}, \boldsymbol{b}) = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|} = \frac{\sum_{i=1}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} a_i^2}\sqrt{\sum_{i=1}^{n} b_i^2}}$$

코사인 유사도

- -1 이상 1 이하의 값을 가진다.
- 코사인 유사도가 높다는 말은 벡터가 더 비슷하다는 의미



머신 러닝에서의 활용

- 텍스트를 분석할 때 벡터로 만들어진 단어나 문장들의 관계성을 파악할때 활용
- 코사인 유사도가 높을수록 단어나 문장들은 더 가까운 관계라는 것을 의미

행렬이란?

- 같은 차원의 벡터를 모아 나열한 것
- 행렬의 크기 : m×n, m행 n열

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

행렬의 덧셈과 뺄셈

- 서로 대응하는 성분끼리 덧셈 또는 뺄셈

$$\begin{bmatrix} 0 & 7 & 2 & 2 \\ 1 & 2 & 6 & 1 \\ 5 & 3 & 3 & 4 \end{bmatrix} + \begin{bmatrix} 2 & 6 & 7 & -1 \\ 1 & 8 & 3 & 5 \\ 0 & -1 & 6 & 11 \end{bmatrix} = \begin{bmatrix} 0+2 & 7+6 & 2+7 & 2-1 \\ 1+1 & 2+8 & 6+3 & 1+5 \\ 5+0 & 3-1 & 3+6 & 4+11 \end{bmatrix} = \begin{bmatrix} 2 & 13 & 9 & 1 \\ 2 & 10 & 9 & 6 \\ 5 & 2 & 9 & 15 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 7 & 2 & 2 \\ 1 & 2 & 6 & 1 \\ 5 & 3 & 3 & 4 \end{bmatrix} - \begin{bmatrix} 2 & 6 & 7 & -1 \\ 1 & 8 & 3 & 5 \\ 0 & -1 & 6 & 11 \end{bmatrix} = \begin{bmatrix} 0-2 & 7-6 & 2-7 & 2+1 \\ 1-1 & 2-8 & 6-3 & 1-5 \\ 5-0 & 3+1 & 3-6 & 4-11 \end{bmatrix} = \begin{bmatrix} -2 & 1 & -5 & 3 \\ 0 & -6 & 3 & -4 \\ 5 & 4 & -3 & -7 \end{bmatrix}$$

행렬의 곱셈

$$A = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} \mathbf{3} \\ \mathbf{2} \end{bmatrix}$$

$$A\mathbf{b} = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{3} \\ \mathbf{2} \end{bmatrix} = \begin{bmatrix} \mathbf{-1 \times 3 + 2 \times 2} \\ \mathbf{1 \times 3 + 1 \times 2} \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ \mathbf{5} \end{bmatrix}$$

$$A = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & 1 \\ 2 & 3 \end{bmatrix}$$

$$AB = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} -1 \times 3 + 2 \times 2 & -1 \times 1 + 2 \times 3 \\ 1 \times 3 + 1 \times 2 & 1 \times 1 + 1 \times 3 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ 5 & 4 \end{bmatrix}$$

행렬의 곱셈

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$A\mathbf{b} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_n \end{bmatrix} \mathbf{b} = \begin{bmatrix} < \boldsymbol{a}_1, \boldsymbol{b} > \\ < \boldsymbol{a}_2, \boldsymbol{b} > \\ \vdots \\ < \boldsymbol{a}_n, \boldsymbol{b} > \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_1 + a_{12}b_2 + \cdots + a_{1n}b_n \\ a_{21}b_1 + a_{22}b_2 + \cdots + a_{2n}b_n \\ \vdots \\ a_{m1}b_1 + a_{m2}b_2 + \cdots + a_{mn}b_n \end{bmatrix}$$

행렬의 곱셈

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_n \end{bmatrix} (\mathrm{b}_1 \quad \mathrm{b}_2 \quad \cdots \quad \mathrm{b}_l)$$

$$= \begin{bmatrix} <\boldsymbol{a}_1, \boldsymbol{b}_1> & <\boldsymbol{a}_1, \boldsymbol{b}_2> & \cdots & <\boldsymbol{a}_1, \boldsymbol{b}_l> \\ <\boldsymbol{a}_2, \boldsymbol{b}_1> & <\boldsymbol{a}_2, \boldsymbol{b}_2> & \cdots & <\boldsymbol{a}_2, \boldsymbol{b}_l> \\ \vdots & \vdots & \ddots & \vdots \\ <\boldsymbol{a}_m, \boldsymbol{b}_1> & <\boldsymbol{a}_m, \boldsymbol{b}_2> & \cdots & <\boldsymbol{a}_m, \boldsymbol{b}_l> \end{bmatrix}$$

행렬의 곱셈

- 첫번째 행렬의 행과 두번째 행렬의 열에 대하여 원소의 위치가 대응하는 값끼리 곱한 다음 더해준다.
- 행렬 간의 곱에서는 앞 행렬의 행과 뒤 행렬의 열의 수가 같아야 한다.

$$A\mathbf{b} = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} -1 \times 3 + 2 \times 2 & -1 \times 1 + 2 \times 3 \\ 1 \times 3 + 1 \times 2 & 1 \times 1 + 1 \times 3 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ 5 & 4 \end{bmatrix}$$

행렬의 곱셈

- A(BC) = (AB)C : 결합 법칙
- A(B+C) = AB+AC, (A+B)C = AC+BC : 분배 법칙
- AB $\neq$ BA : 교환 법칙은 성립하지 않는다.

$$\begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ 5 & 4 \end{bmatrix} \neq \begin{bmatrix} -2 & 7 \\ 1 & 7 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}$$

행렬의 스칼라배

- 행렬의 모든 성분에 같은 수를 곱하는 것

$$\frac{1}{2}\begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

행렬의 종류

- 영행렬(Zero matrix) : 행렬의 성분 전체가 0인 행렬

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

행렬의 종류

- 정방행렬(Square matrix) : 행과 열의 크기가 같은 행렬

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

행렬의 종류

- 대각행렬(Diagonal matrix) : 정방행렬에서 대각선에 해당되는 대각 성분만 0이 아니고, 나머지 비대각 성분의 원소 값이 0인 행렬

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

행렬의 종류

- 단위행렬(Identity matrix) : 대각행렬 중 대각 성분이 1로만 구성된 행렬
- 어떤 행렬이나 벡터에 단위행렬을 곱하면 결과가 달라지지 않는다
  : 항등사상

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

행렬의 종류

- 상삼각행렬(Upper triangular matrix) : 대각선을 기준으로 위쪽에만 원소 값이 있는 행렬

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$$

행렬의 종류

- 하삼각행렬(Lower triangular matrix) : 대각선을 기준으로 아래쪽에만 원소 값이 있는 행렬

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix}$$

행렬의 종류

- 대칭행렬(Symmetric matrix) : 대각 원소에 대하여 대칭인 원소가 같은 행렬
- A = A^T

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$$

행렬의 종류

- 직교행렬(Orthogonal matrix) : 열벡터들이 서로 직교하고 길이가 1인 행렬
- $PP^T = P^T P = I$

행렬식(Determinant)

- |A| = det A = ad-bc

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

- |A| = det A = $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$ = a$\begin{vmatrix} e & f \\ h & i \end{vmatrix}$-b $\begin{vmatrix} d & f \\ g & i \end{vmatrix}$+c $\begin{vmatrix} d & e \\ g & h \end{vmatrix}$

역행렬(Inverse matrix)

- $AA^{-1} = A^{-1}A = I$를 만족하는 행렬 A

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$\frac{1}{|A|}$

역행렬(Inverse matrix)

- 정방행렬이어야한다.
- 행렬식이 0인 경우 역행렬이 존재하지 않는다. (nonsingular matrix)
- 2×2 행렬보다 큰 정방행렬의 역행렬은 가우스 소거법이나 여인자 전개와 같은 방법으로 계산

역행렬(Inverse matrix)의 특징

- 항등행렬 : $AA^{-1} = A^{-1}A = I$
- 행렬 곱셈 전체에 대한 역행렬이 존재한다면, 곱한 행렬도 각각 역행렬이 존재한다. : $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$
- A의 역행렬이 존재한다면, A의 역행렬도 역행렬이 존재한다.
  : A^{-1}=B, $B^{-1} = A$
- A의 전치행렬(Transpose matrix)의 역행렬은 A의 역행렬의 전치행렬과 같다. : $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

★ 다시보기

고윳값(Eigenvalue)와 고유벡터(Eigenvector)

- 다음의 식을 만족하는 열벡터 x(x≠0)가 존재할 때, λ를 행렬 A의 고윳값이라 하고, x를 고유벡터라고 한다.

$$A\mathbf{x} = \lambda I\mathbf{x}$$

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

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

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

$\det(A - \lambda I) = 0$ → 고유방정식(Eigen Equation)

고윳값과 고유벡터 계산

$$A = \begin{bmatrix} 2 & 4 \\ -1 & -3 \end{bmatrix}$$

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

$$\det\begin{pmatrix} 2-\lambda & 4 \\ -1 & -3-\lambda \end{pmatrix} = 0$$

$$(2-\lambda)(-3-\lambda)-4(-1) = 0$$

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

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

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

고윳값과 고유벡터 계산

λ = -2 일때,

(A-(-2)I)**x** = $\begin{pmatrix} 2-(-2) & 4 \\ -1 & -3-(-2) \end{pmatrix}$**x** = $\begin{pmatrix} 4 & 4 \\ -1 & -1 \end{pmatrix}$**x** = **0**

x = $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$이라 가정

$\begin{pmatrix} 4 & 4 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\alpha + \beta = 0$

x = t$\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

고윳값과 고유벡터 계산

λ = 1 일때,

(A-I)$\mathbf{x}$ = $\begin{pmatrix} 2-1 & 4 \\ -1 & -3-1 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 1 & 4 \\ -1 & -4 \end{pmatrix}\mathbf{x} = \mathbf{0}$

x = $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$이라 가정

$\begin{pmatrix} 1 & 4 \\ -1 & -4 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\alpha + 4\beta = 0$

$\mathbf{x} = t\begin{pmatrix} 1 \\ -\frac{1}{4} \end{pmatrix} = t\frac{1}{4}\begin{pmatrix} 4 \\ -1 \end{pmatrix}$

머신 러닝에서의 활용

- 데이터가 많이 흩어져 있는 분포 상황에서는 주어진 문제를 해결하기 위해 고윳값과 고유벡터의 도움을 받는다.
- 고윳값은 어떤 데이터가 가진 특징을 얼마나 잘 설명할 수 있는지 가늠할 때 사용된다.
- 기여율은 각 고유벡터에 대응하는 고윳값을 전체 고윳값들의 총합으로 나눈 것으로, 우리가 가진 데이터를 얼마나 잘 설명할 수 있는지 평가하는 척도로 사용된다.

Q & A

감사합니다 :)