인공지능수학

1. 거리(distance)

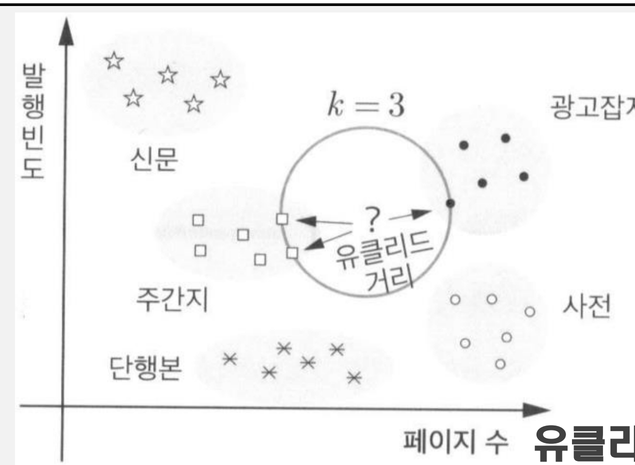
■ 인공지능에서는 이렇게 활용한다!

- 인공지능 분야에서는 과거의 데이터를 분석하여 최적의 모델을 만드는 학습 단계와, 그렇게 만들어진 모델을 사용해서 새로운 데이터에 대한 카테고리나 수치를 예측하는 추론 단계가 있습니다.
- 유클리드 거리는 인공지능 분야의 다양한 알고리즘에서 사용되는데, 그중 하나가 k-NN(k-nearest neighbors algorithm)이라는 분류 방법입니다. 이 방법은 소위 ‘교사가 있는 학습’, 즉 지도학습을 하기 때문에 미리 정답 데이터가 준비되어 있어야합니다.
- K-NN은 학습 단계에서 학습된 데이터를 벡터 공간상에 위치시킨 후, 추론 단계에서 새로운 데이터를 같은 공간에 배치합니다. 새 데이터가 어떤 카테고리에 속하는지 알기 위해서는 가까이에 있는 k개의 정답 데이터를 보고 추론하게 되는데, 이때 사용하는 것이 바로 유클리드 거리입니다.

1. 거리(distance)

■ 인공지능에서는 이렇게 활용한다!

➲ 이해를 돕기 위해 예를 하나 들어보겠습니다. 학습 단계에서 다음 그림처럼 ‘페이지 수’ 와 ‘발행빈도’ 의 축으로 만들어진 공간이 있습니다. 여기에 학습 데이터를 위치시키고 카테고리를 매칭합니다. 추론 단계에서는 새 데이터(그림에서 ‘? ‘로 표시)를 페이지 수와 발행빈도에 맞게 위치시킨 후, 추론을 시작합니다. 만약 k = 3이라고 할 때, 새 데이터에서 가까운 두 개의 데이터가 ‘주간지’ 카테고리에 속하고, 1개의 데이터가 ‘광고잡지’ 카테고리에 분류되어 있다면 새로운 데이터는 ‘주간지’ 카테고리에 속한다고 추론하게 되는 것입니다.



유클리드 거리

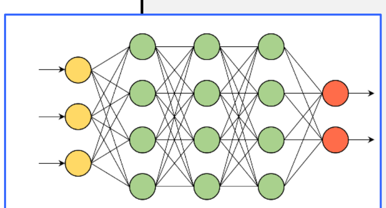
1. 함수 혹은 수열

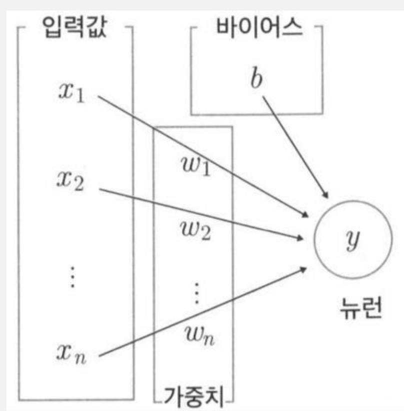
■ 인공지능에서는 이렇게 활용한다!

➲ 머신러닝 분야에서 주목받는 알고리즘의 하나인 ‘신경망’은 인간의 뇌 속에 있는 신경세포 ‘뉴런’과 그들의 연결 관계를 인공적으로 흉내 낸 것입니다.

➲ 뉴런에 입력되는 값은 여러 개의 ‘입력값’과 ‘가중치(ω)’의 곱을 모두 더한 다음, 여기에 상수를 더 추가한 것과 같습니다.

➲ 경우에 따라 다르겠지만 신경망에서는 이런 하나의 모델에서 이와 같은 덧셈이 수백만번 이루어지기도 합니다. 이런 계산식을 하나하나 쓰는 것은 사실상 불가능에 가깝기 때문에 $\sum$를 사용해서 표현합니다.





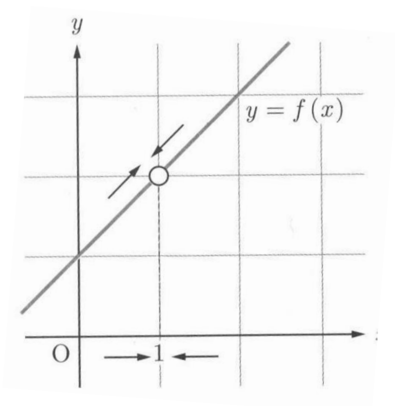
$$y = b + x_1 \cdot w_1 + x_2 \cdot w_2 + \cdots + x_n w_n$$
$$= \sum_{k=1}^{n} x_k w_k + b$$

신경망의 개요

2.1 미분 – 극한

■ 극한

➲ 수열이나 함숫값이 어떤 특정값에 한없이 가까워지는 것을 의미한다.



- $f(x) = \frac{x^2-1}{x-1}$
- lim 표현:

$$\lim_{x\to 1}\frac{x^2-1}{x-1} = \lim_{x\to 1}\frac{(x-1)(x+1)}{x-1} = \lim_{x\to 1}(x+1) = 2$$

➲ 용어

- 수렴(convergent) x의 값을 어떤 값 a에 최대한 가깝게 만들 때, 함수 $f(x)$의 값도 어떤 값 α에 최대한 가까워지는 모양
- 극한값(limit, limiting value) α 는 함수 $f(x)$에서 $x \to a$ 일 때의 극한값

➲ 수식 표현

- $\lim_{x\to a} f(x) = \alpha$
- $f(x) \to \alpha \quad (x \to a)$

2.2 미분의 기초

■ 평균 속도

➲ 단위 시간당 얼마나 이동

- $v = \dfrac{\Delta x}{\Delta t}$ 　　이동 거리를 이동시간으로 나눔

■ 순간 속도

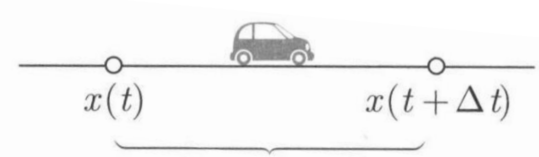
➲ 아주 짧은 구간의 속도

- $v = \lim\limits_{\Delta t \to 0} \dfrac{\Delta x}{\Delta t} = \lim\limits_{\Delta t \to 0} \dfrac{x(t+\Delta t) - x(t)}{\Delta t}$

Δx - 이동 거리의 변화량
Δt - 이동 시간의 변화량

- 미분 표기

$$v = \frac{dx(t)}{dt}$$



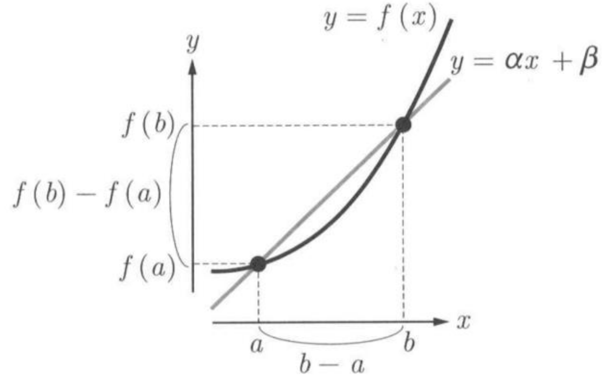
이 구간의 거리는 $\Delta x = x(t+\Delta t) - x(t)$ 라고 표현할 수 있습니다.

2.2 미분의 기초

■ 함수의 미분

➲ 두 점을 지나는 직선

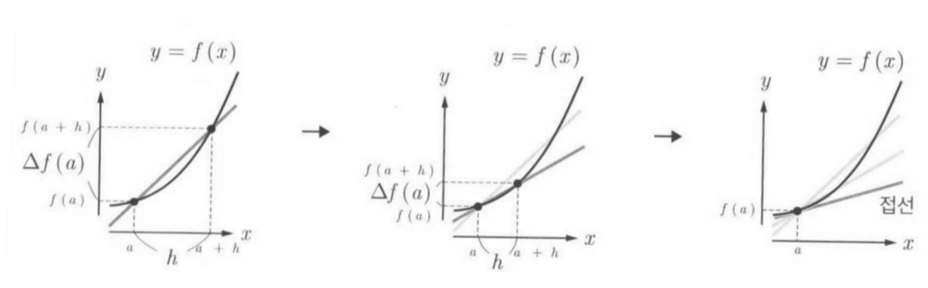
- $f(a) = \alpha a + \beta$
- $f(b) = \alpha b + \beta$
- 기울기 $\alpha = \frac{f(b)-f(a)}{b-a}$



➲ 극한과 접선

- 미분 : 특정한 지점에서의 기울기 구하는 것

$$\frac{df(a)}{dx} = \lim_{\Delta x \to 0} \frac{\Delta f(a)}{\Delta x} = \lim_{h \to 0} \frac{f(a+h)-f(a)}{(a+h)-a} = \lim_{h \to 0} \frac{f(a+h)-f(a)}{h}$$



- 접선 : 점$(a, f(a))$에서 $y = f(x)$접하는 직선
- 미분계수 : $x = a$일 때의 평균변화율의 극한값 α

2.2 미분의 기초

■ 함수의 미분

➲ 도함수

정의역의 모든 x에 대해 함수$f(x)$의 미분계수를 대응시키는 새로운 함수를 $f(x)$의 도함수라고 한다.

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{\Delta f(x)}{\Delta x} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

➲ 표기:

- 1계 미분 : y', $f'(x)$, $\frac{dy}{dx}$, $\frac{d}{dx}f(x)$
- 2계 미분 : y'', $f''(x)$, $\frac{d^2}{dx^2}f(x)$

2.2 미분의 기초

■ 인공지능에서는 이렇게 활용한다!

➲ 인공지능 분야에서는 함수의 값이 어느 지점에서 최소가 되는지를 알아내는 것이 중요합니다.

➲ 예를 들면, 손실 함수(loss function)는 정답과 예측값(측정값) 사이의 오차를 표현하는 함수인데, 인공지능 분야에서는 이 함수의 값을 최소로 만들기 위해 다양한 기법을 사용합니다.

➲ 손실 함수를 미분하면 어떤 특정 지점에서 어느 정도의 기울기가 나오는지 알 수 있는데, 이러한 기울기의 절댓값이 작아지는 방향으로 그 지점을 옮기다 보면 손실 함수의 최솟값을 구할 수 있습니다. 이 방법을 경사하강법(gradient descent) 이라 부릅니다.

2.3 상미분과 편미분

■ 상미분(ordinary derivative)

- ➲ 변수 하나만 있는 함수의 미분
- ➲ 주요 공식

① $y = x^r$ 일 때 $\frac{dy}{dx} = rx^{r-1}$ (r 은 임의의 실수)

② $\frac{\mathrm{d}}{\mathrm{dx}}\{\mathrm{f(x)} + \mathrm{g(x)}\} = \frac{df(x)}{dx} + \frac{dg(x)}{dx}$

③ $\frac{\mathrm{d}}{\mathrm{dx}}\{kf(x)\} = k\frac{df(x)}{dx}$

■ 전미분(total differentiation)

- ➲ 변수가 2개 이상인 다변수함수(multivariate function)의 미분

2.3 상미분과 편미분

■ 편미분(partial derivative)

➲ 다변수함수에 대하여, 그 중 하나의 변수에 주목하고 나머지 변수의 값을 고정시켜 놓고 그 변수로 미분하는 방법

- 예) 함수 $z = f(x, y) = 3x^2 + 2xy + 2y^2$
  - x에 대한 편미분 : y를 상수로 보고 x로 미분

$$\frac{\partial f(x,y)}{\partial x} = \lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x} = 6x + 2y$$

  - y에 대한 편미분 : x를 상수로 보고 y로 미분

$$\frac{\partial f(x,y)}{\partial y} = \lim_{\Delta y \to 0} \frac{\Delta z}{\Delta y} = 2x + 4y$$

➲ 표기법

- 라이프니츠 표기법: $\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y}$
- 라그랑주 표기법: $f_x(x, y), f_y(x, y)$

2.4 그래프 그리기

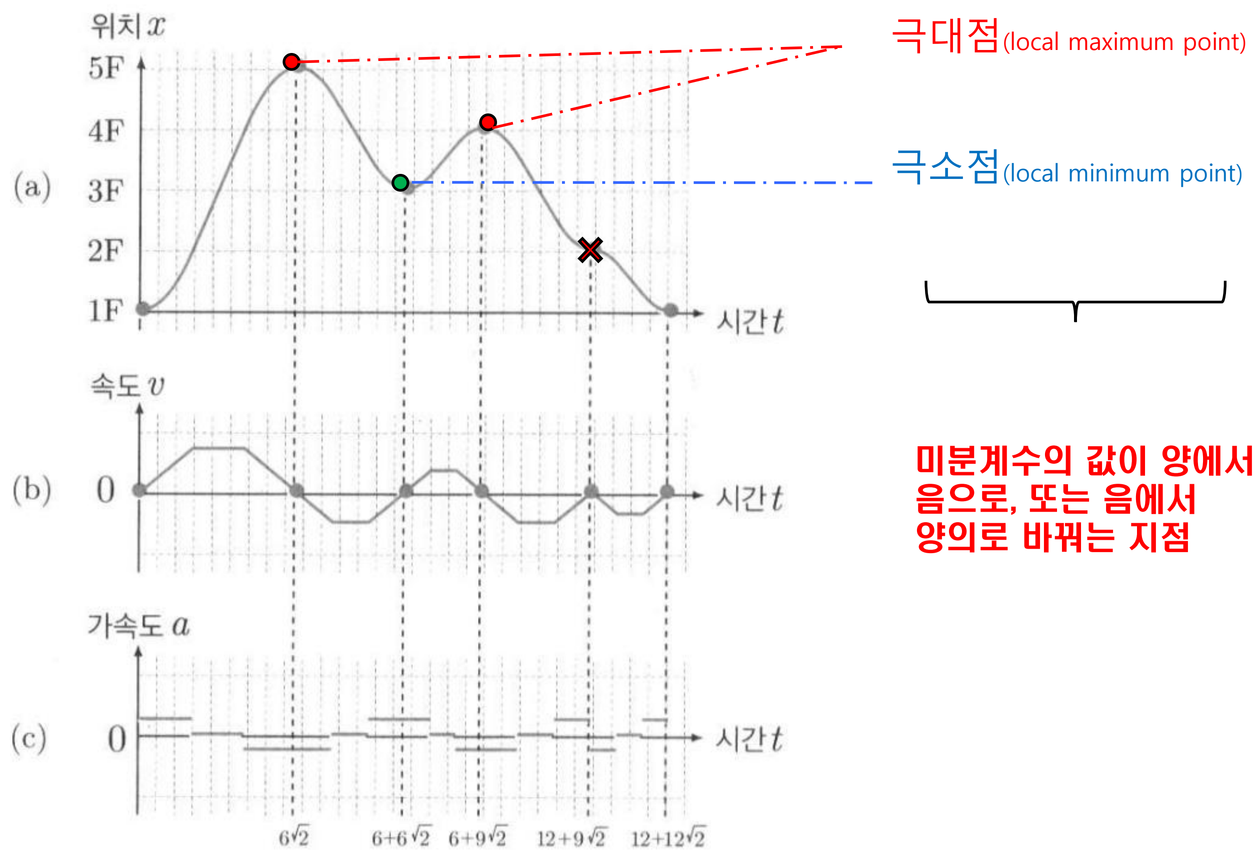
■ 그래프의 의미

➲ 예제) 엘레베이터가 1층→5층→3층→4층→2층→1층으로 이동, 시간 t에 대한 엘레베이터의 위치 x와 속도 v, 그리고 가속도 a를 그래프 표현

$x = f(t)$

1계 미분 $v = \frac{df(t)}{dt}$

2 계 미분 $a = \frac{d^2f(t)}{dt^2}$



미분계수의 값이 양에서 음으로, 또는 음에서 양의로 바뀌는 지점

2.4 그래프 그리기

■ 증감표

➲ 그래프의 특징을 표로 정리

극점 local extremum point
극값 local extremum value

시간 t	0	$\cdots$	$2\sqrt{2}$	$\cdots$	$4\sqrt{2}$	$\cdots$	$6\sqrt{2}$	$\cdots$	$2+6\sqrt{2}$	$\cdots$	$4+6\sqrt{2}$	$\cdots$	$6+6\sqrt{2}$	$\cdots$
속도 $v = \frac{dx}{dt}$	0			$+$			0			$-$			0	$+$
가속도 $a = \frac{d^2x}{dt^2}$		$+$		0			$-$			0			$+$	
위치 x	1F	⤴	2F	↗	4F	↱	5F (극대)	⤵	4.5F	↘	3.5F	⤷	3F (극소)	⤴

➲ 부호와 화살표 의미

$\frac{dx}{dt}$	$\frac{d^2x}{dt^2}$	화살표	의미
0		→	x 는 일정 $\left(\frac{dx}{dt} = 0\right)$
$+$	$+$	⤴	x 는 증가 $\left(\frac{dx}{dt} > 0\right)$ 하고, 증가율이 증가 $\left(\frac{d^2x}{dt^2} > 0\right)$
$+$	0	↗	x 는 증가 $\left(\frac{dx}{dt} > 0\right)$ 하고, 증가율이 일정 $\left(\frac{d^2x}{dt^2} = 0\right)$
$+$	$-$	↱	x 는 증가 $\left(\frac{dx}{dt} > 0\right)$ 하고, 증가율이 감소 $\left(\frac{d^2x}{dt^2} < 0\right)$
$-$	$+$	⤷	x 는 감소 $\left(\frac{dx}{dt} < 0\right)$ 하고, 감소율이 감소 $\left(-\frac{d^2x}{dt^2} < 0\right)$
$-$	0	↘	x 는 감소 $\left(\frac{dx}{dt} < 0\right)$ 하고, 감소율이 일정 $\left(-\frac{d^2x}{dt^2} = 0\right)$
$-$	$-$	⤵	x 는 감소 $\left(\frac{dx}{dt} < 0\right)$ 하고, 감소율이 증가 $\left(-\frac{d^2x}{dt^2} > 0\right)$

2.5 함수의 최대값과 최소값

■ 규칙

- ➲ 함수의 최대값, 최소값은 극점이나 구간의 양끝단에서 나온다.
- ➲ 1계 미분한 값이 0일 때 극값을 구할 수 있다.

■ 인공지능에서는 이렇게 활용한다!

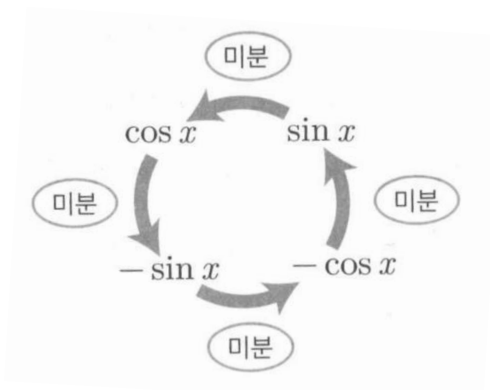
- ➲ 최소제곱법(least square method)은 여러 개의 오차들을 제곱하고, 그 제곱들의 합이 최소가 되는 관계식을 구한 것으로, 선형회귀(linear regression) 알고리즘에서 사용하는 최적화 기법입니다. 선형회귀 알고리즘은 인공지능 분야에서 자주 사용되는 기본 알고리즘 중의 하나입니다.
- ➲ $f(a, b)$ = (오차의 제곱들의 합)이라고 가정할 때, $\frac{\partial f(a,b)}{\partial a} = 0$이고 $\frac{\partial f(a,b)}{\partial b} = 0$이 되는 방정식을 풀면 오차의 제곱들을 합한 결과가 최소가 되는 관계식을 풀 수 있습니다.
- ➲ 이런 풀이는 2차함수를 1계 미분한 함수의 값이 0일 때, 극값(최소 제곱법에서는 최소값)을 구할 수 있다는 특징을 잘 활용한 예입니다.

2.6.1 초등함수의 미분

■ 초등함수(elementary function)의 미분

원래 함수			원래 함수를 x로 미분한 것
멱함수		x^r	rx^{r-1}
지수함수		$e^x, \exp(x)$	$e^x, \exp(x)$
		a^x	$a^x \log_e a$
로그함수		$\log_e x \ (x > 0)$	$\frac{1}{x}$
삼각함수	사인함수	$\sin x$	$\cos x$
	코사인함수	$\cos x$	$-\sin x$
	탄젠트함수	$\tan x$	$\frac{1}{\cos^2 x}$

■ 삼각함수 미분의 관계



2.6.2 합성함수의 미분

■ 합성함수의 미분(연쇄법칙, 곱의 법칙)

➲ 합성함수(변수가 1개) $y = f(x)$의 미분법

- 연쇄법칙(chain rule)

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

임의의 식을 끼워 넣습니다.

➲ 합성함수(변수가 여러 개) $z = f(x, y)$의 미분법

- $\dfrac{\partial z}{\partial x} = \dfrac{\partial z}{\partial u} \cdot \dfrac{\partial u}{\partial x} + \dfrac{\partial z}{\partial v} \cdot \dfrac{\partial v}{\partial x}$

➲ 곱의 법칙

- $\dfrac{d}{dx}\{f(x)g(x)\} = \dfrac{df(x)}{dx} g(x) + f(x)\dfrac{dg(x)}{dx}$

2.6.2 합성함수의 미분

➲ 예제) 함수 $f(x) = (3x-4)^{50}$을 x에 대해 미분 $(u = 3x - 4)$

$$\frac{df(x)}{dx} = \frac{du^{50}}{du} \cdot \frac{d(3x-4)}{dx} = 50u^{49} \cdot 3 = 150(3x-4)^{49}$$

➲ 예제) 함수 $f(x,y) = (3x+1)^2 + (x+y+1)^3$을 x에 대해 미분

$u = 3x+1, v = x+y+1$ 가정하면 $f(x,y) = u^2 + v^3$

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

➲ 예제) 함수 $y = xe^x$를 x에 대해 미분$(f(x) = x, g(x) = e^x)$

$$\frac{dy}{dx} = \frac{df(x)}{dx} g(x) + f(x)\frac{dg(x)}{dx}$$
$$= 1 \cdot e^x + x \cdot e^x$$
$$= (1+x)e^x$$

c는 상수, n은 임의의 실수일 때 함수 $f(x)$와 $g(x)$의 도함수가 존재하면 다음 관계식이 성립한다.

$$\text{(1)}\quad \frac{d}{dx}[c] = 0$$

$$\text{(2)}\quad \frac{d}{dx}[cf(x)] = cf'(x)$$

$$\text{(3)}\quad \frac{d}{dx}[x^n] = nx^{n-1} \tag{1.5}$$

$$\text{(4)}\quad \frac{d}{dx}[f(x) \pm g(x)] = f'(x) \pm g'(x)$$

$$\text{(5)}\quad \frac{d}{dx}[f(x)g(x)] = f(x)g'(x) + f'(x)g(x)$$

$$\text{(6)}\quad \frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2} \qquad (g(x) \neq 0)$$

$$D_x u^r = ru^{r-1}u'(x)$$

$$D_x \ln u = \frac{1}{u}u'(x)$$

$$D_x e^u = e^u u'(x)$$

$$D_x \sin u = \cos u \cdot u'(x)$$

$$D_x \cos u = -\sin u \cdot u'(x)$$

$$D_x \tan u = \sec^2 u \cdot u'(x)$$

$$D_x \cot u = -\csc^2 u \cdot u'(x)$$

$$D_x \sec u = \sec u \tan u \cdot u'(x)$$

$$D_x \csc u = -\csc u \cot u \cdot u'(x)$$

$$D_x \sin^{-1} u = \frac{1}{\sqrt{1-u^2}}u'(x)$$

$$D_x \cos^{-1} u = -\frac{1}{\sqrt{1-u^2}}u'(x)$$

$$D_x \tan^{-1} u = \frac{1}{1+u^2}u'(x)$$

2.6.2 합성함수의 미분

■ 인공지능에서는 이렇게 활용한다!

- ➲ 신경망에서는 학습한 결과로 도출된 답이 정답 데이터에 가까워질 수 있도록 가중치(ω)를 조정하는 과정을 반복합니다.
- ➲ 이때, 실제 정답과 학습 결과 사이의 오차 값을 가중치로 편미분한 다음, 그 값을 가중치의 조정량으로 사용합니다.
- ➲ 이렇게 편미분을 하는 과정에서 앞서 살펴본 연쇄법칙이 사용되는데, 이런 일련의 기법을 오차역전파법(backpropagation)이라고 합니다.

3장 선형대수

개 요

벡터, 행렬

특징

노름, 거리, 차원,
내적, 행렬식

연산

벡터합, 스칼라곱,
행렬합, 행렬곱

특수형태

영행렬, 음행렬, 단위행렬,
역행렬, 대각행렬,
전치행렬 등

선형문제

- 선형결합
- 기본행변환
- 가우스-조르단 소거법

알고리즘

- LU분해
- QR분해
- 최소제곱법
- 그람-슈미트 방법

행렬의 특성값

- 고윳값&고유벡터
- 특이값
  - 특이값분해 (SVD)

"선형대수학"이란?

■ 선형대수학(Linear Algebra)이란?

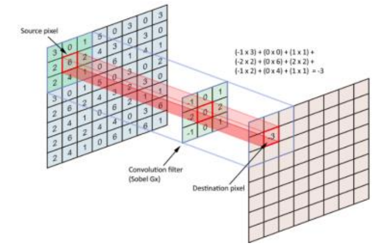
- ➲ 선형 방정식 문제를 풀기위해 사용되는 방법들을 연구하는 과목이다.
- ➲ 해의 존재성, 그리고 해를 "효율적"으로 찾는 것이 목표이다.
- ➲ 선형 시스템이 갖는 고유 특성을 이용하여 시스템을 분석한다.

■ 선형대수학의 적용분야

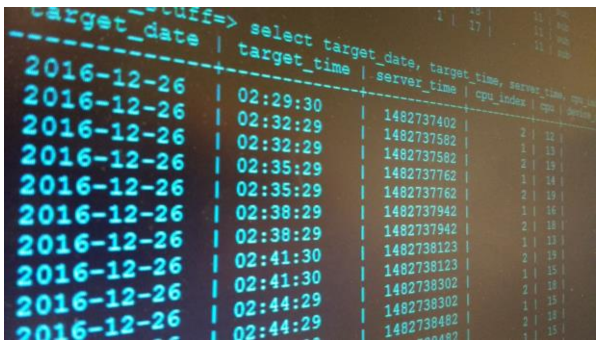
- ➲ 다변수 시스템 분야
  - ◆ 기계공학, 로봇공학, 위성통신, 기상학 등...
- ➲ 데이터를 취급하는 분야
  - ◆ 빅데이터 처리, 영상처리, 음향처리, 인공지능 등...

IT기술과 선형대수

데이터 처리



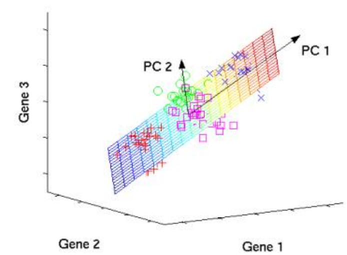
선형 데이터



행렬

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2n} \\ A_{31} & A_{32} & A_{33} & \cdots & A_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & A_{m3} & \cdots & A_{mn} \end{pmatrix}$$

데이터 분석

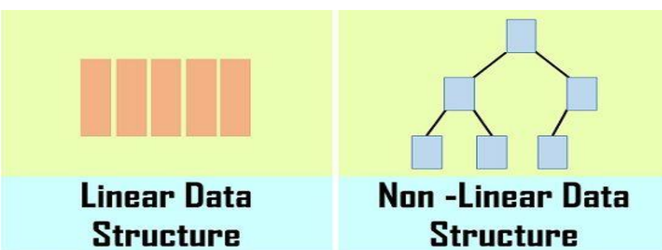


고속연산

컴퓨터공학과 선형대수

■ 컴퓨터공학에서 자주 사용되는 선형 데이터는 선형대수에서 벡터 또는 행렬로 다룰 수 있다.

■ 선형 데이터란 일렬로 나열되어 각 데이터의 순서(index)가 정의된 데이터를 의미한다.

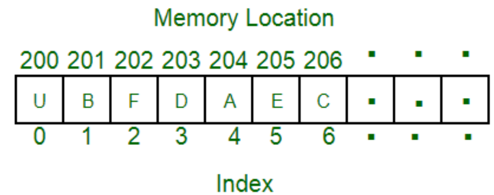


선형 데이터의 대표적 종류

테이블 데이터

Name	Stage	Applying for	Application Dat
Aaron Buxton	Pending	Event Support Agent	11/1/2019
Miguel Silva	Active	Event Support Agent	11/3/2019
Eric Ishida	On hold	Event Support Agent	10/16/2019
Kadji Bell	Active	Account Manager	10/13/2019
David Power	On hold	Event Support Agent	11/3/2019
Babak Shammas	Active	Support Team Lead	10/18/2019
Charlotte de Crum	Closed	Event Support Agent	11/5/2019
Joshua Vanburen	Active	Account Manager	10/11/2019
Keiko Tanaka	On hold	Event Support Agent	11/3/2019

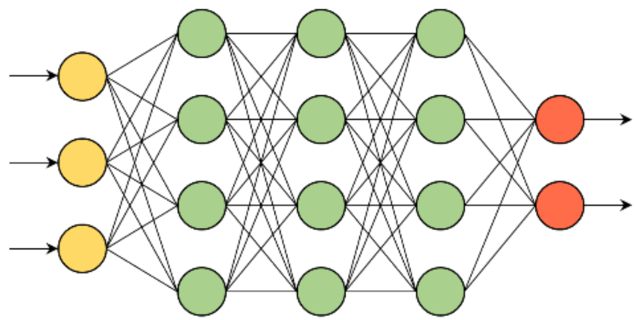
배열



이미지

인공지능과 선형대수

- 잘 알려진 인공지능 모형 중 하나인 딥러닝 모형은 선형대수 모형으로 표현이 가능하다.



||

$$\begin{pmatrix} w_{11}^{(4)} & w_{12}^{(4)} & w_{13}^{(4)} & w_{14}^{(4)} \\ w_{21}^{(4)} & w_{22}^{(4)} & w_{23}^{(4)} & w_{24}^{(4)} \end{pmatrix} \left(\cdots f \left(\begin{pmatrix} w_{11}^{(1)} & w_{12}^{(1)} & w_{13}^{(1)} \\ w_{21}^{(1)} & w_{22}^{(1)} & w_{23}^{(1)} \\ w_{31}^{(1)} & w_{32}^{(1)} & w_{33}^{(1)} \\ w_{41}^{(1)} & w_{42}^{(1)} & w_{43}^{(1)} \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) \right)$$

선형데이터의 표현 : 벡터&행렬

■ 선형데이터는 다음과 같이 벡터 또는 행렬로 표현할 수 있다.

선형 데이터

	A	B
1	Student	Score
2	A	90점
3	B	87점
4	C	79점
5	D	84점
6	E	96점

벡터

$$x = \begin{bmatrix} 90 \\ 87 \\ 79 \\ 84 \\ 96 \end{bmatrix}$$

선형 데이터

	A	B	C	D
1	Student	Score		
2		math	korean	english
3	A	90점	79점	80점
4	B	87점	61점	77점
5	C	79점	91점	86점
6	D	84점	88점	85점
7	E	96점	95점	100점

행렬

$$X = \begin{bmatrix} 90 & 79 & 80 \\ 87 & 61 & 77 \\ 79 & 91 & 86 \\ 84 & 88 & 85 \\ 96 & 95 & 100 \end{bmatrix}$$

3.1 벡터

■ 벡터는 다음과 같이 두가지 형태로 나타낼 수 있다.

선형 데이터

	A	B
1	Student	Score
2	A	90점
3	B	87점
4	C	79점
5	D	84점
6	E	96점

열벡터

$$x = \begin{bmatrix} 90 \\ 87 \\ 79 \\ 84 \\ 96 \end{bmatrix}$$

행벡터

$$x = \begin{bmatrix} 90 & 87 & 79 & 84 & 96 \end{bmatrix}$$

3.1 벡터

■ 벡터 성분

- 90, 87, 79, 84, 96는 벡터 x의 벡터 성분이다.

- 일반적으로 다음과 같이 표현한다.

$$x_1 = 90,\ \ x_2 = 87,\ \ \ldots\ \ ,\ \ x_5 = 96$$

■ n차원 벡터

- 벡터의 크기 = 벡터 성분의 개수

- 예시의 경우, x는 5차원 벡터이다.

예시)

$$x = \begin{bmatrix} 90 \\ 87 \\ 79 \\ 84 \\ 96 \end{bmatrix}$$

3.2 벡터의 연산

■ 벡터합

$$x + y = \begin{bmatrix}1\\2\\3\\4\end{bmatrix} + \begin{bmatrix}1\\2\\1\\2\end{bmatrix} = \begin{bmatrix}1+1\\2+2\\3+1\\4+2\end{bmatrix} = \begin{bmatrix}2\\4\\4\\6\end{bmatrix}$$

■ 스칼라곱

$$\alpha x = 3 \times \begin{bmatrix}1\\2\\3\\4\end{bmatrix} = \begin{bmatrix}3\times 1\\3\times 2\\3\times 3\\3\times 4\end{bmatrix} = \begin{bmatrix}3\\6\\9\\12\end{bmatrix}$$

예시)

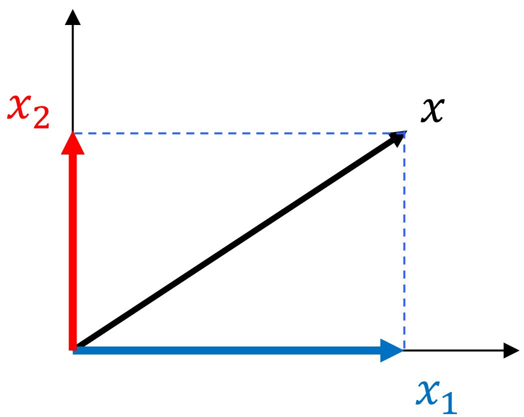
$$x = \begin{bmatrix}1\\2\\3\\4\end{bmatrix},\ y = \begin{bmatrix}1\\2\\1\\2\end{bmatrix},$$

$$\alpha = 3$$

3.2 벡터의 연산

■ 그래프 위의 표현(3.3 유향성분)

벡터 & 벡터성분

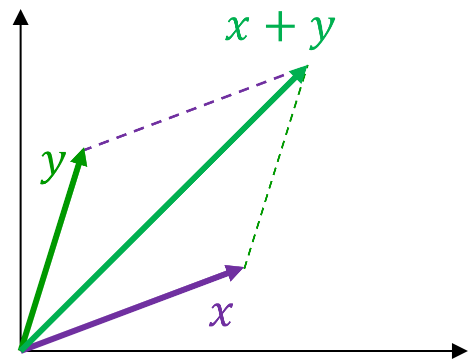


x_1와 x_2 는

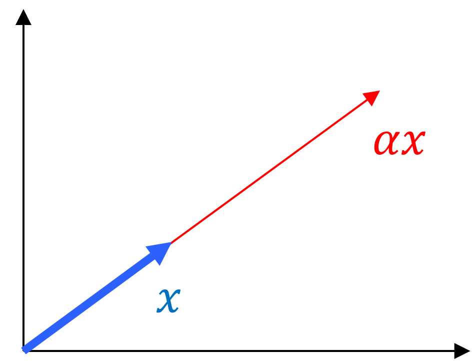
x의 벡터성분

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

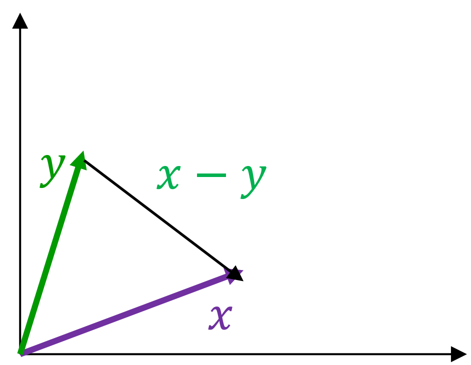
벡터의 덧셈



스칼라곱



벡터의 뺄셈



영벡터와 음벡터

■ 영벡터

- 모든 벡터성분이 0인 벡터

  예를들면, $a = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, $b = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

- 일반적으로 영벡터는 0으로 표현한다.

■ 음벡터

- 주어진 벡터의 모든 벡터성분에 -1을 곱한 벡터
- 간단하게, $-x$을 x의 음벡터로 나타낸다.
- 주어진 예시에 대해 다음 벡터가 음벡터이다.

$$-x = \begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix} \qquad (※\ x + (-x) = 0)$$

예시)

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

3.4 벡터의 내적

■ 벡터의 내적

- $x \cdot y = \sum_{i=1}^{n} x_i y_i$
- 예를 들면,

$$x \cdot y = (1 \times 2) + (2 \times 3) + (1 \times 1) = 9$$

예시)

$$x = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, y = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

- **Cauchy–Schwarz 부등식(Theorem)**

For all vectors x and y,

$$|x \cdot y| \leq ||x|| \times ||y||$$

- **내적을 기하학적인 특징으로 정의 (두 벡터의 직교성은?)**

벡터 x 와 y 가 이루는 각이 θ일 때 두 벡터의 내적은

$$x \cdot y = ||x|| ||y|| \cos \theta$$

3.7 벡터의 노름

■ 노름

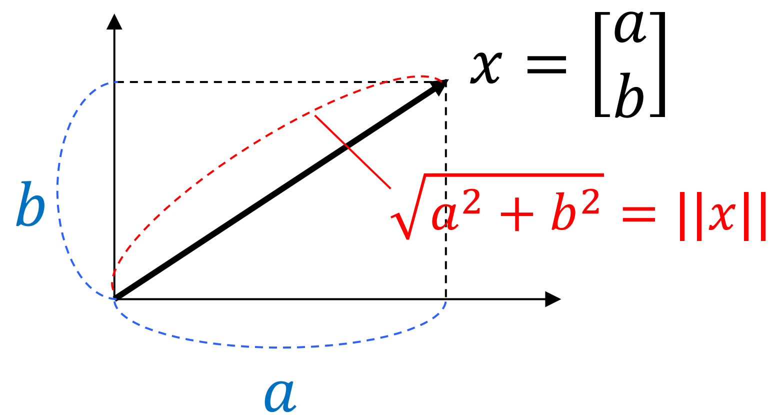
- 벡터의 크기 또는 원점으로부터의 거리
- $||x|| = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$
- 예를 들면, $||x|| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$

예시)

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

- L1노름은 절대값의 합
- L2 노름은 유클리드 거리

■ 그래프 표현



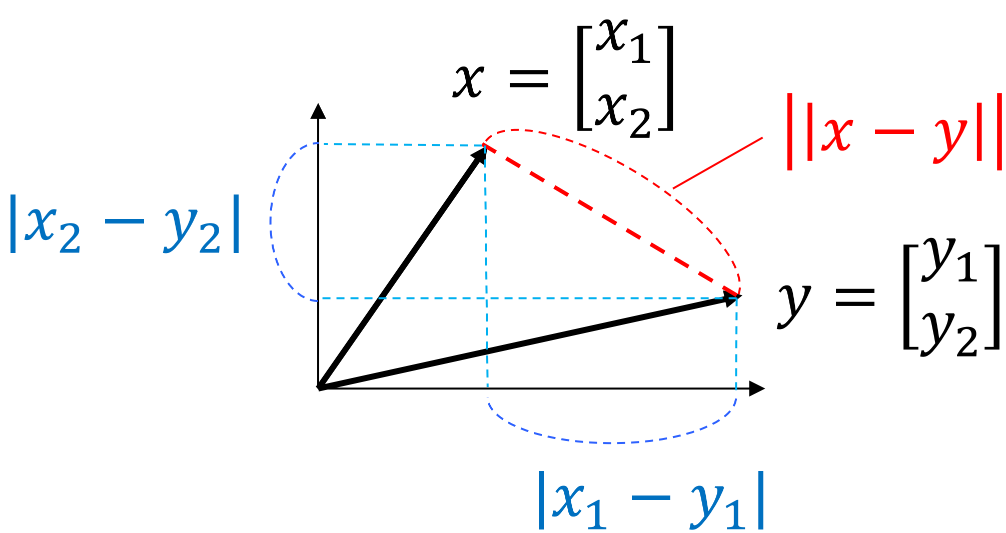
벡터의 거리

■ 거리

- $||x - y|| = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$
- 예를 들면,

$||x - y|| = \sqrt{(1-2)^2 + (2-3)^2 + (1-1)^2} = \sqrt{2}$

■ 그래프 표현



예시)

$x = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, y = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$

3.8 코사인 유사도

■ 벡터의 내적으로부터

$$\cos\theta = x \cdot y / ||x||||y||$$

$$||x|| = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

$$||y|| = \sqrt{\sum_{i=1}^{n} y_i^2} = \sqrt{y_1^2 + y_2^2 + \cdots + y_n^2}$$

- **코사인 유사도(Cosine Similarity) $\cos(\boldsymbol{x}, \boldsymbol{y})$**

$$\cos(\boldsymbol{x}, \boldsymbol{y}) = \frac{x \cdot y}{||x||||\mathbf{y}||} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}}$$

$--\rightarrow \quad -1 \leqq \cos(x, y) \leqq 1$ **(두 벡터의 관계성)**

3.9 행렬

선형 데이터

	A	B	C	D
1	Student	Score		
2		math	korean	english
3	A	90점	79점	80점
4	B	87점	61점	77점
5	C	79점	91점	86점
6	D	84점	88점	85점
7	E	96점	95점	100점

행렬

$$X = \begin{bmatrix} 90 & 79 & 80 \\ 87 & 61 & 77 \\ 79 & 91 & 86 \\ 84 & 88 & 85 \\ 96 & 95 & 100 \end{bmatrix}$$

5 행

3 열

벡터처럼 행렬도 행렬의 성분을 갖는다.

***행렬은 "n 행, m 열"* 의 크기를 갖는다.**

예를들면, X 는 "5 행, 3 열인 행렬이다".

3.9 행렬

■ 다양한 형태의 행렬이 존재한다.

정방행렬

$$X = \begin{bmatrix} 3 & 9 & 0 \\ 4 & 1 & 2 \\ 5 & 1 & 5 \end{bmatrix}$$

대각행렬

$$X = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

영행렬

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

대칭행렬

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

상삼각행렬

$$X = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{bmatrix}$$

열벡터와 닮은 행렬

$$X = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

3.9.1 행렬의 벡터표현

선형 데이터

	A	B	C	D
1	Student	Score		
2		math	korean	english
3	A	90점	79점	80점
4	B	87점	61점	77점
5	C	79점	91점	86점
6	D	84점	88점	85점
7	E	96점	95점	100점

행렬

$$X = \begin{array}{cc} \begin{bmatrix} 90 & 79 & 80 \\ 87 & 61 & 77 \\ 79 & 91 & 86 \\ 84 & 88 & 85 \\ 96 & 95 & 100 \end{bmatrix} & \begin{array}{c} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{array} \\ \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} & \end{array}$$

$$= \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{11} \end{bmatrix}$$

행렬의 벡터표현

$$X = [v_1 \ \ v_2 \ \ v_3]$$

$$X = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix}$$

3.9.2 행렬의 연산

■ 행렬의 덧셈

$$A + B = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 2 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1+1 & 1+3 \\ 2+2 & 0+2 \\ 3+1 & 1+1 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 2 \\ 4 & 2 \end{bmatrix}$$

■ 행렬의 스칼라곱

$$\alpha A = 2 \times \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 2\times 1 & 2\times 1 \\ 2\times 2 & 2\times 0 \\ 2\times 3 & 2\times 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 4 & 0 \\ 6 & 2 \end{bmatrix}$$

예시)

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 3 & 1 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & 3 \\ 2 & 2 \\ 1 & 1 \end{bmatrix},$$

$$\alpha = 2$$

3.9.3 영행렬과 음행렬

■ 영행렬

- 행렬의 모든 성분이 0인 행렬

  예를 들면, $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
- 일반적으로 영행렬을 0으로 표현한다.

예시)

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 3 \\ 3 & 0 & 4 \end{bmatrix}$$

■ 음행렬

- 주어진 행렬의 모든 성분에 -1을 곱하여 만든 행렬 $-A$.
- 예를들면,

$$-A = \begin{bmatrix} -1 & -2 & -2 \\ -2 & -1 & -3 \\ -3 & 0 & -4 \end{bmatrix} \qquad * \ A + (-A) = 0$$

3.9.4 전치행렬과 대칭행렬

■ 전치행렬

- **전치행렬**이란 대각성분을 기준으로 모든 성분의 행과 열의 위치를 바꾼 행렬이다.
- 예를들면,

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 1 \\ 3 & 0 \end{bmatrix} \quad \Rightarrow \quad A^T = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 1 & 0 \end{bmatrix}$$

■ 대칭행렬

- 만약 주어진 행렬이 "$A = A^T$"라면,
행렬 A를 ***대칭행렬*** 이라한다.

예를들면,

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

3.10 행렬의 곱셈

■ 행렬곱(행렬-행렬곱)

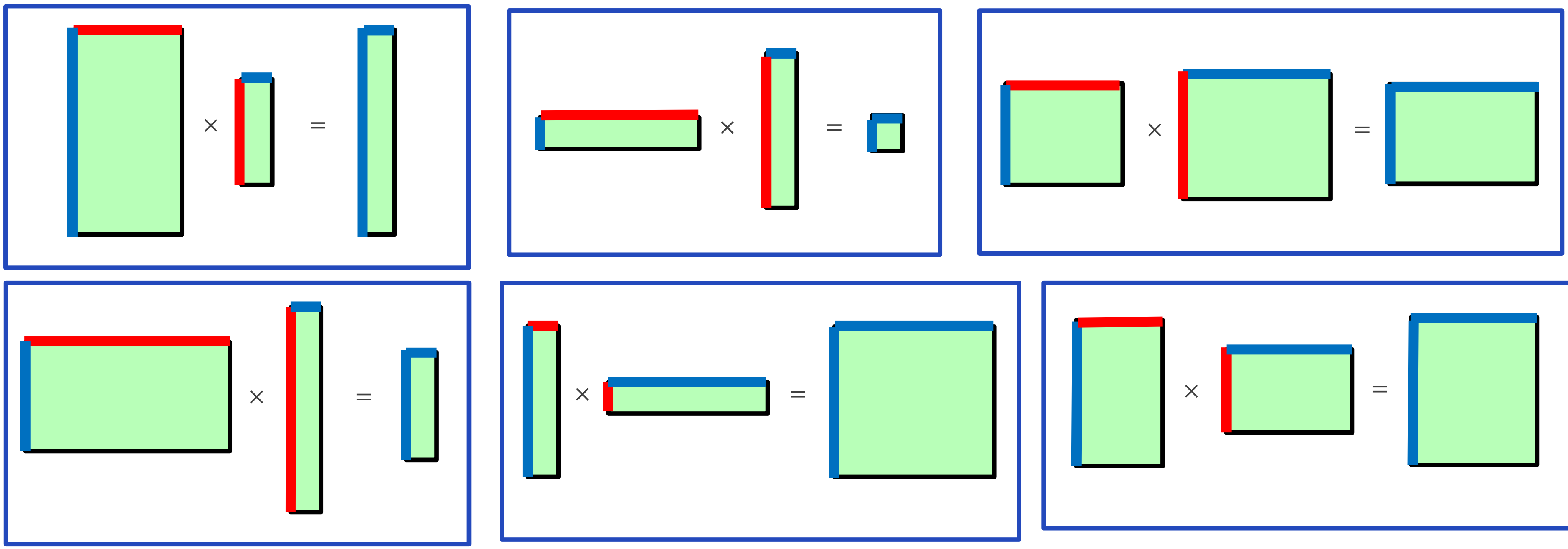
$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} = \begin{bmatrix} x_{11}a_1 + x_{12}b_1 + x_{13}c_1 & x_{11}a_2 + x_{12}b_2 + x_{13}c_2 \\ x_{21}a_1 + x_{22}b_1 + x_{23}c_1 & x_{21}a_2 + x_{22}b_2 + x_{23}c_2 \end{bmatrix}$$

예를 들면,

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 4 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1\times1+2\times0+3\times2 & 1\times2+2\times4+3\times1 \\ 2\times1+0\times0+1\times2 & 2\times2+0\times4+1\times1 \\ 3\times1+1\times0+1\times2 & 3\times2+1\times4+1\times1 \end{bmatrix} = \begin{bmatrix} 7 & 13 \\ 4 & 5 \\ 5 & 11 \end{bmatrix}$$

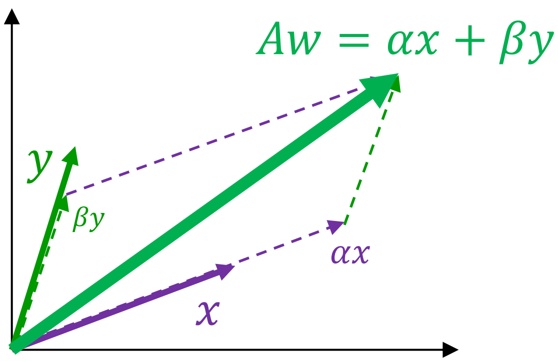
3.10 행렬의 곱셈

■ 행렬연산의 기하학적 이해



행렬곱의 그래프 표현

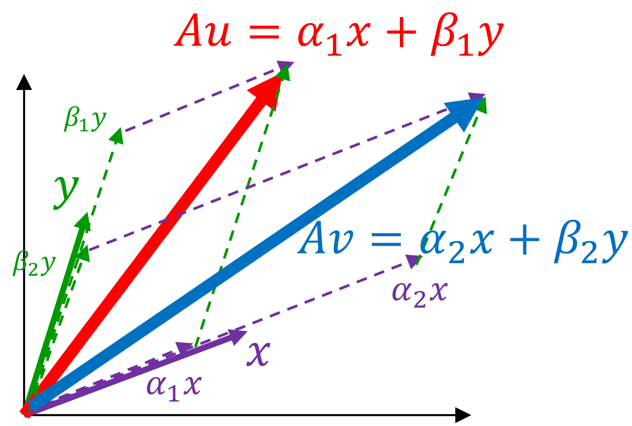
행렬-벡터곱



* 행렬을 열벡터들의 모임으로 생각한다면, 행렬-벡터곱은 열벡터들로 새로운 벡터를 만들어낸다.

행렬-행렬곱

$$AB = [Au \;\; Av]$$



* 행렬-벡터곱으로 만든 여러개의 열벡터를 모아 새로운 행렬을 만들어낸다.

예시)

$$A = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} = [x \quad y]$$

$$w = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

$$B = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{bmatrix} = [u \quad v]$$

3.10.1 행렬-벡터곱

■ 행렬-벡터곱

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x_{11}a + x_{12}b + x_{13}c \\ x_{21}a + x_{22}b + x_{23}c \\ x_{31}a + x_{32}b + x_{33}c \end{bmatrix}$$

예시)

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \end{bmatrix},$$

$$x = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix},$$

예를 들면,

$$Ax = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \times 4 + 2 \times 5 + 3 \times 6 \\ 2 \times 4 + 0 \times 5 + 1 \times 6 \\ 3 \times 4 + 1 \times 5 + 1 \times 6 \end{bmatrix} = \begin{bmatrix} 32 \\ 14 \\ 23 \end{bmatrix}$$

■ 단위행렬

- 모든 대각성분이 1로 이루어진 대각행렬.

예를 들면, $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

- 모든 정방행렬 A에 대해,

$$AI = IA = A$$

예를 들면,

$$AI = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \end{bmatrix} = A$$

3.11 역행렬

■ 역행렬

- 주어진 정방행렬 A에 대해,
만약 $AA^{-1} = A^{-1}A = I$를 만족하는 A^{-1}가 있다면,
A^{-1}를 A의 **역행렬**이라 한다.

예를 들면,

1) $AB = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} -2 & 3 \\ 3 & -4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow B = A^{-1}$

2) $CD = \begin{bmatrix} 2 & 3 & -1 \\ 4 & 4 & -3 \\ -2 & 3 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 & -\frac{1}{4} \\ \frac{1}{2} & -\frac{1}{5} & \frac{1}{10} \\ 1 & -\frac{3}{5} & -\frac{1}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow D = C^{-1}$

정 리

벡터, 행렬		
특징	**연산**	**특수형태**
노름, 거리, 차원, 내적, 행렬식	벡터합, 스칼라곱, 행렬합, 행렬곱	영행렬, 음행렬, 단위행렬, 역행렬, 대각행렬, 전치행렬 등

- 선형문제
  - 선형결합
  - 기본행변환
  - 가우스-조르단 소거법
- 알고리즘
  - LU분해
  - QR분해
  - 최소제곱법
  - 그람-슈미트 방법
- 행렬의 특성값
  - 고윳값&고유벡터
  - 특잇값
    - 특잇값분해 (SVD)