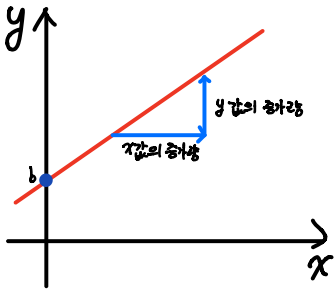


<2장. 답례임을 위한 기초 수학>

1. 일차 함수, 기울기과 y절편

$$y = ax + b$$

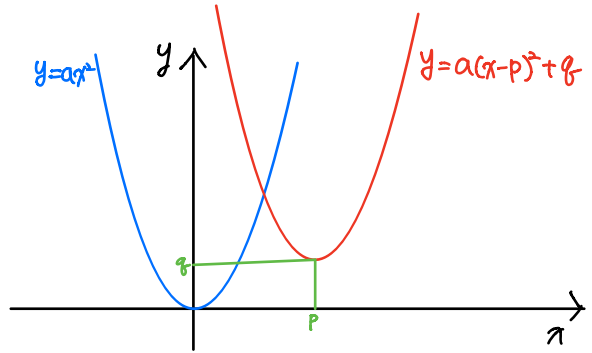


$$a(\text{기울기}) = \frac{\text{x값의 증가량}}{\text{y값의 증가량}}$$

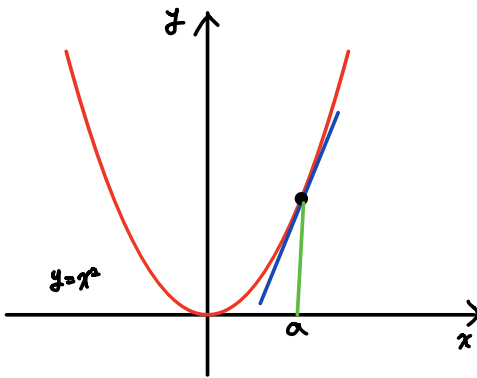
$$b = \text{y절편.}$$

2. 이차 함수와 최솟값

$$y = a(x-p)^2 + q$$



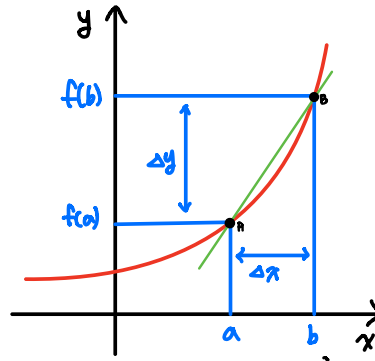
3. 미분, 순간 변화율과 기울기



a 에서의 순간 변화율 = 기울기.

미분 계수: 어느 순간에 어떤 변화가 일어나고 있는지를 숫자로 나타낸 것.

= 그래프에서의 기울기



$$\begin{aligned} \text{직선 AB의 기울기} &= \frac{\text{y값의 증가량}}{\text{x값의 증가량}} \\ &= \frac{f(b) - f(a)}{b - a} = \frac{f(a + \Delta x) - f(a)}{\Delta x} \end{aligned}$$

$b - a = \Delta x$
 $b = a + \Delta x$

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(a + \Delta x) - f(a)}{\Delta x}$$

(순간 변화율)

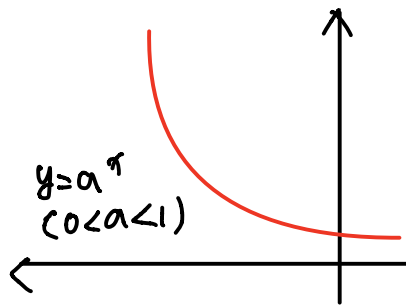
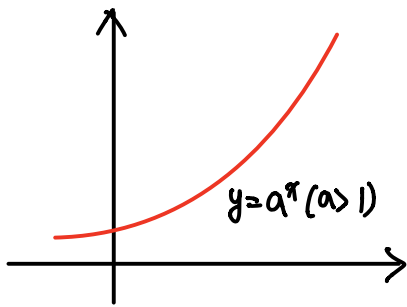
4. 편미분

' x 에 관하여 편미분하라' = $\frac{\partial f}{\partial x}$

$$f(x, y) = x^2 + xy + a$$

$$\frac{\partial f}{\partial x} = 2x + y, \quad \frac{\partial f}{\partial y} = 1$$

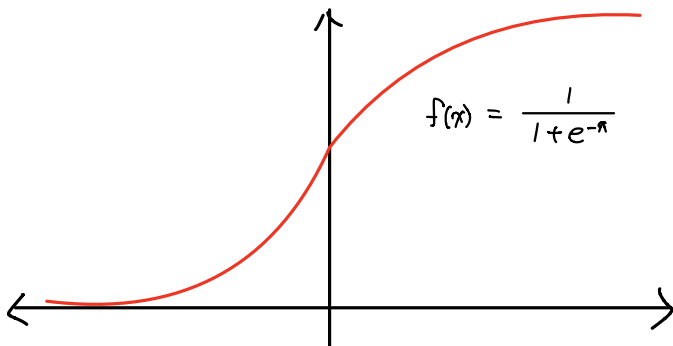
5. 지수과 지수함수



6. 시그모이드 함수.

e = 자연상수 / 자연로그의 밑 / 오일러의 수

시그모이드 함수 = e 가 지수함수에 포함되어 분자로 들어간 함수

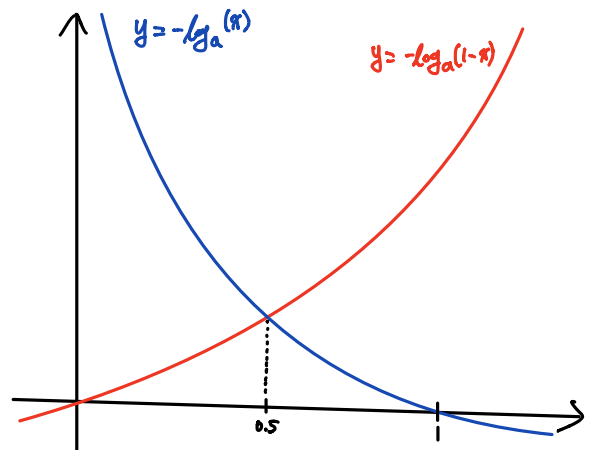
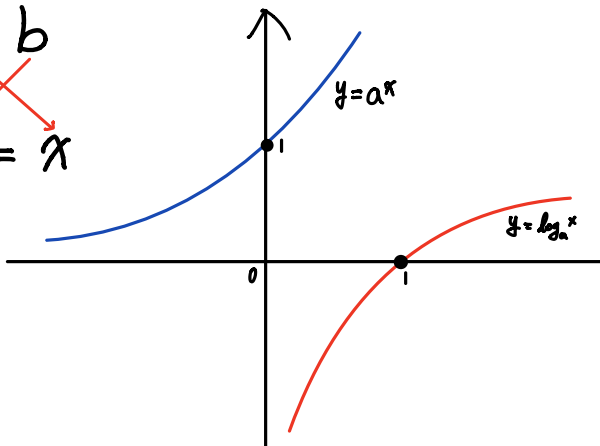


x 가 큰 값을 가지면 $f(x)$ 는 1에 가까워짐.
 x 가 작은 값을 가지면 $f(x)$ 는 0에 가까워짐.
 $\Rightarrow 0$ 또는 1, 두 개의 값 중 하나를 고를 때 유용.

7. 로그와 로그 함수

$$a^x = b$$

$$\log_a b = x$$



\Rightarrow 컴퓨터 기법 판단 장치에서 사용.

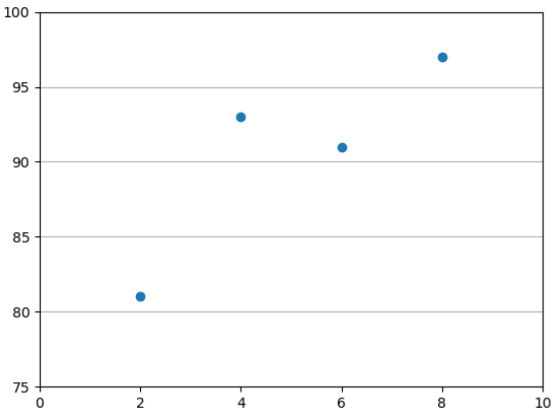
<3장. 가장 훌륭한 예측선 긋기 : 선형 회귀>

1. 선형 회귀의 정의

- 선형 회귀 : 독립 변수 x를 사용해 종속 변수 y의 움직임을 예측하고 설명하는 작업
- 단순 선형 회귀 : 하나의 x값만으로도 y값을 설명할 수 있을 때
 - 다중 선형 회귀 : x값이 여러 개 필요 할 때

2. 가장 훌륭한 예측선이란?

공부한 시간	2시간	4시간	6시간	8시간
성적	81점	93점	91점	97점



선형 회귀를 공부하는 과정
= 이 점들의 특징을 가장 잘 나타내는 정확한 선을 그리는 과정

여기에서 선은 직선이므로 곧 일차 함수 그래프이다.
 $y = ax + b$

정확하게 선을 그리기 위해서 a,b 값을 알아야함

3. 최소 제곱법

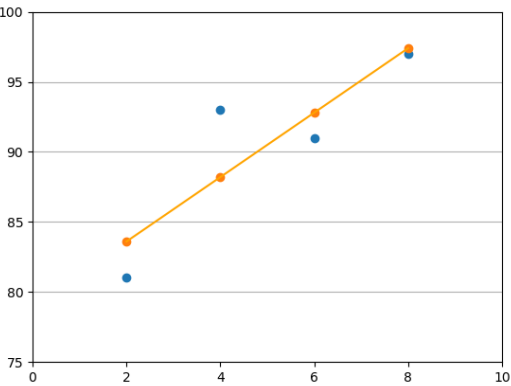
최소 제곱법 : 주어진 x의 값이 하나일 때 적용이 가능한 공식

$$a = \frac{(x - x_{\text{평균}})(y - y_{\text{평균}}) \text{의 합}}{(x - x_{\text{평균}})^2 \text{의 합}}$$
$$b = y \text{의 평균} - (x \text{의 평균} \times \text{기울기 } a)$$

$$\Rightarrow a = \frac{(2-5)(81-90.5) + (4-5)(93-90.5) + (6-5)(91-90.5) + (8-5)(97-90.5)}{(2-5)^2 + (4-5)^2 + (6-5)^2 + (8-5)^2}$$
$$= \frac{46}{20} = 2.3$$
$$b = 90.5 - (2.3 \times 5) = 79$$

※ $y = 2.3x + 79$

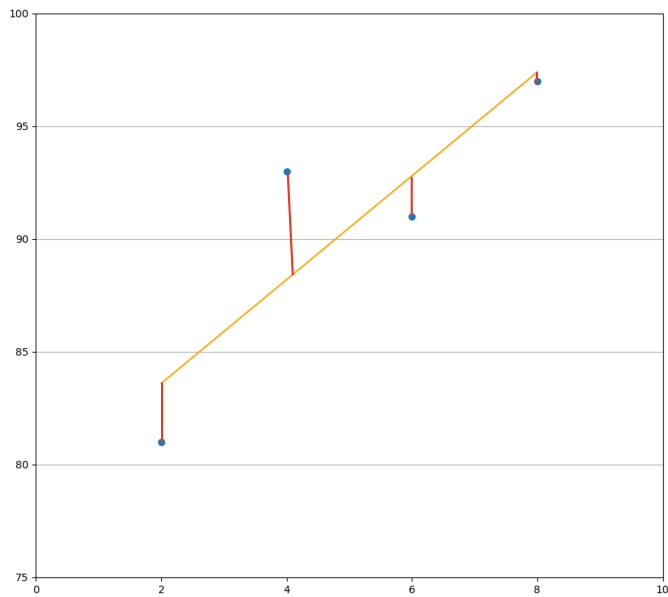
공부한 시간	2	4	6	8
성적	81	93	91	97
예측 값	83.6	88.2	92.8	97.4



5. 평균 제곱 오차

최소 제곱법만을 이용하여 여러 개의 입력을 처리하기 어렵기 때문에 오차를 평가하는 방법이 필요
=> 평균 제곱 오차를 사용

6. 잘못 그은 선 바로잡기



공부한 시간	2	4	6	8
성적	81	93	91	97
예측 값	83.6	88.2	92.8	97.4
오차	1	-5	3	3

오차의 합 = $\sum_i^n (y_i - \hat{y}_i)^2$

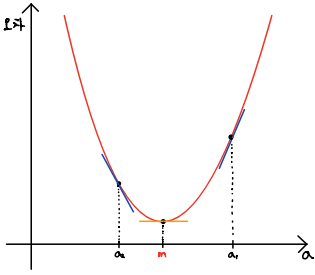
평균 제곱 오차(MSE) = $\frac{1}{n} \sum_i^n (y_i - \hat{y}_i)^2$

=> 평균 제곱 오차를 구하면
 $\frac{1}{4} \times (1 + 25 + 9 + 9) = 11$

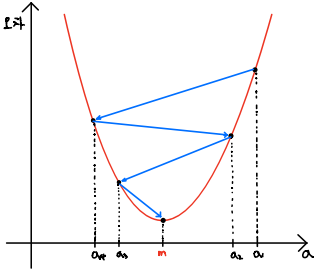
- 선형 회귀란
 - 임의의 직선을 그어 이에 대한 평균 제곱 오차를 구하고, 이 값을 가장 작게 만들어 주는 a와 b 값을 찾아가는 작업

<4장. 오차 수정하기 : 경사 하강법>

1. 경사 하강법의 개요



최솟값 m 에서의 순간 기울기 = 0
 \Rightarrow 우리는 미분 값이 0인 지점을 찾아야 한다!



- 1) a_1 에서 미분을 구한다.
- 2) 구해진 기울기의 반대 방향(기울기가 +면 음의 방향, -면 양의 방향)으로 이동시킨 a_2 에서 미분을 구한다.
- 3) 위에서 구한 미분 값이 0이 아니면 위 과정을 반복

\Rightarrow 경사 하강법
 : 이렇게 반복적으로 기울기 a 를 변화시켜 m 의 값을 찾아내는 방법

- 기울기의 부호를 바꿔 이동시킬 때 적절한 거리를 찾지 못해 너무 멀리 이동 시키면 a 값이 한 점으로 모이지 않고 위로 치솟아 버림. 따라서 어느 만큼 이동시킬지를 정하는 것이 중요. 이 때 이동거리를 정해주는 것이 학습률
- 학습률 : 기울기의 부호를 바꿔 이동 시킬 적절한 거리

<경사 하강법 계산식>

$$\frac{1}{n} \sum_i^n (y_i - \hat{y}_i)^2 \Rightarrow \frac{1}{n} \sum_i^n (y_i - (ax_i + b))^2$$

a 로 편미분 한 경과 = $\frac{2}{n} \sum (ax_i + b - y_i)x_i$

b 로 편미분 한 경과 = $\frac{2}{n} \sum (ax_i + b - y_i)$

↗ 학습률

$$a = a - lr \times \frac{2}{n} \sum (a_1x_{1i} + b - y_i)x_{1i}$$

$$b = b - lr \times \frac{2}{n} \sum (ax_i + b - y_i)$$

↘ a, b 값 업데이트

4. 다중 선형 회귀란

- 더 정확한 예측을 위해 추가 정보(독립 변수)를 추가해 새로운 예측 값을 구하는 것

공부한 시간(x_1)	2	4	6	8
과외 수업 횟수(x_2)	0	4	2	3
성적	81	93	91	97

$$a_1 = a_1 - lr \times \frac{2}{n} \sum (a_1x_{1i} + b - y_i)x_{1i}$$

$$a_2 = a_2 - lr \times \frac{2}{n} \sum (a_2x_{2i} + b - y_i)x_{2i}$$

$$b = b - lr \times \frac{2}{n} \sum (ax_{1i} + b - y_i)$$

$$y = a_1x_1 + a_2x_2 + b$$

a_1, a_2, b 값 업데이트

< 5장. 참 거짓 판단 장치 : 로지스틱 회귀 >

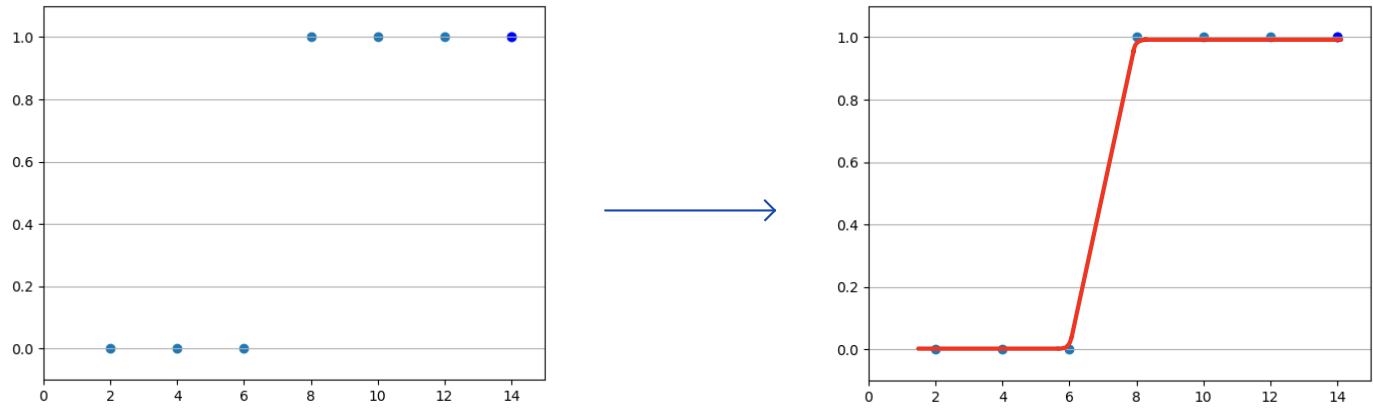
- 딥러닝을 수행 한다는 것은 복잡한 연산을 해낸 끝에 최적의 예측 값을 내놓는 작업이라고 할 수 있음.
=> 참과 거짓 중에 하나를 내놓는 과정은 로지스틱 회귀의 원리를 거쳐 이루어짐.

• 딥러닝 동작 원리

1. 로지스틱 회귀의 원리를 이용해 주어진 입력 값의 특징을 추출하여 모델을 만든다.
2. 비슷한 질문을 하면 지금까지 만들어 놓은 모델을 꺼내어 답을 한다.

1. 로지스틱 회귀의 정의

공부한 시간	2	4	6	8	10	12	14
합격 여부	불합격	불합격	불합격	합격	합격	합격	합격

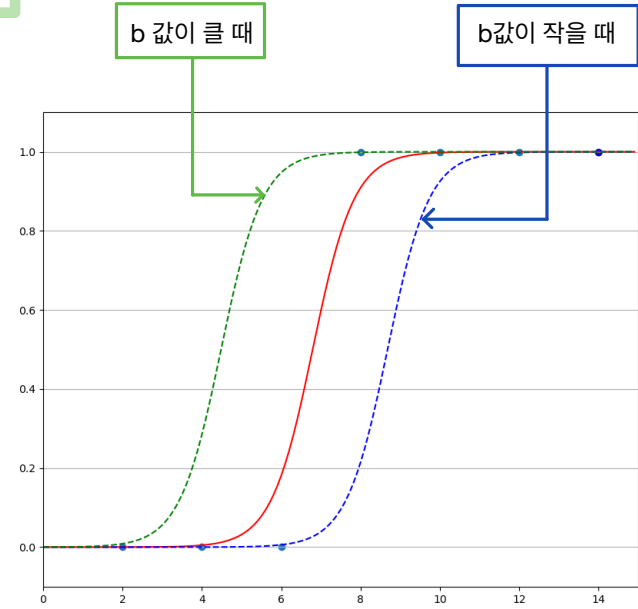
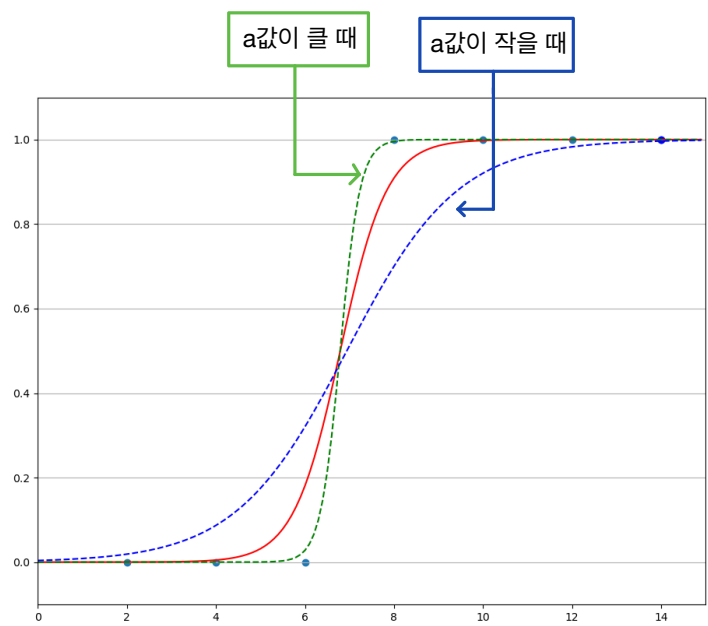


- 로지스틱 회귀 : 참(1)과 거짓(0) 사이를 구분하는 S자 형태의 선을 그어 주는 작업

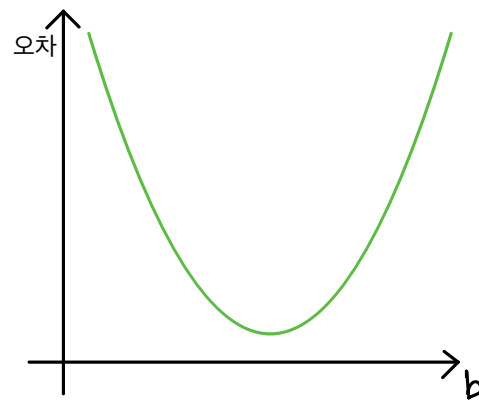
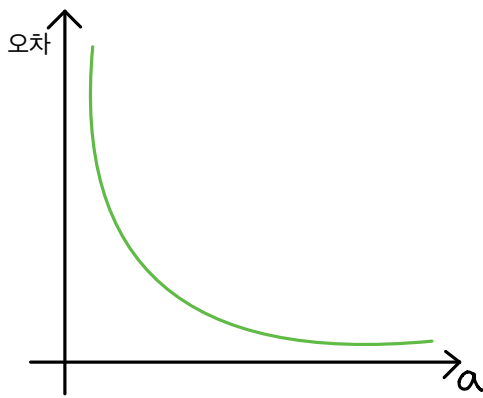
2. 시그모이드 함수

- 시그모이드 함수(sigmoid function) : S자 형태로 그래프를 그려주는 함수

$$y = \frac{1}{1 + e^{-(ax+b)}}$$
 → a와 b 값에 따라 s형태의 그래프가 달라짐



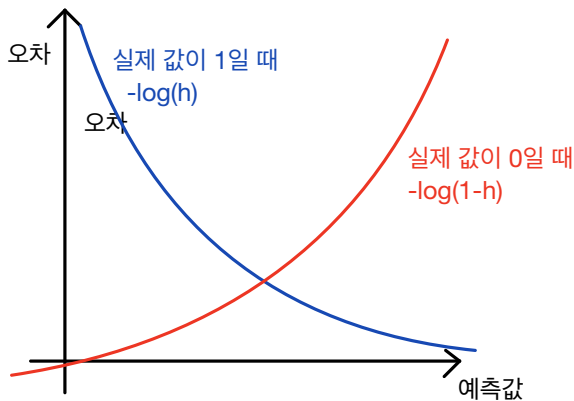
- a와 b 값에 따라 오차가 변함



3. 오차 공식

- 시그모이드 함수의 특징 : y값이 0과 1 사이
 - 실제 값이 1일 때 예측 값이 0에 가까워지면 오차가 커짐
 - 실제 값이 0일 때 예측 값이 1에 가까워지면 오차가 커짐
- => 공식으로 만들 수 있게 해주는 함수 = 로그함수

4. 로그함수



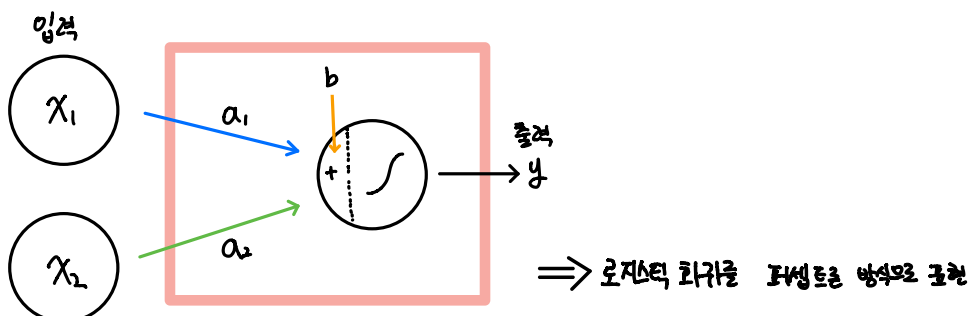
$$-\underbrace{\{y_{data} \log h\}}_A + \underbrace{(1 - y_{data}) \log(1 - h)}_B$$

y_data가 1이면 B부분이 없어지고
y_data가 0이면 A부분이 없어져 오차를 알 수 있다.

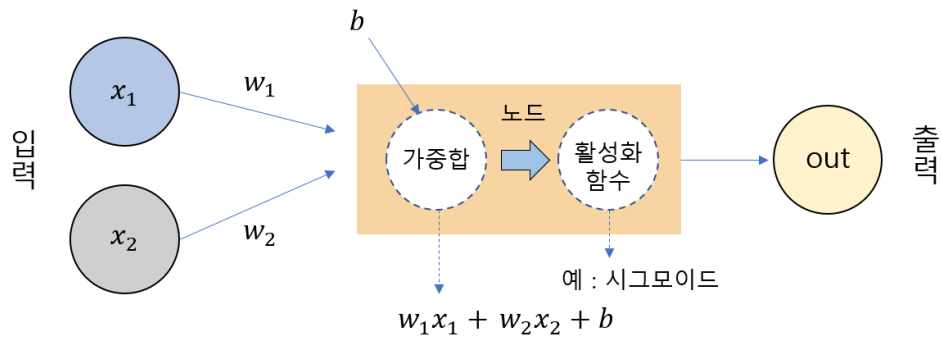
- 로지스틱 회귀의 오차 함수

$$-\frac{1}{n} \sum \{y_{data} \log h + (1 - y_{data}) \log(1 - h)\}$$

6. 로지스틱 회귀에서 퍼셉트론으로



<6장. 퍼셉트론>



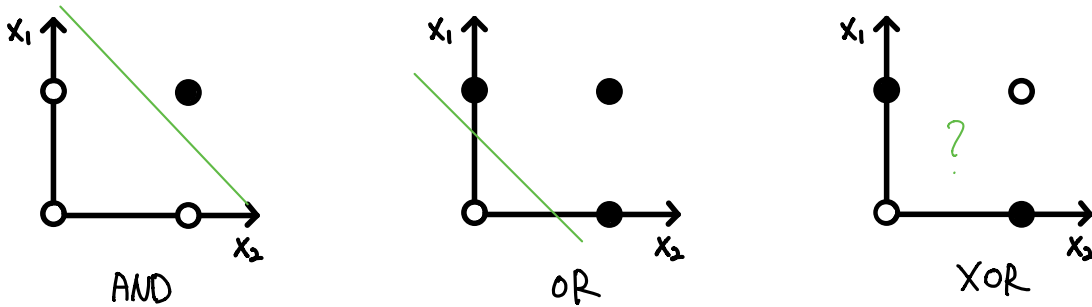
1. 가중치, 가중합, 바이어스, 활성화 함수

$$y = wx + b \quad (w = \text{가중치}, b = \text{바이어스})$$

가중합 : 입력값(x)과 가중치(w)의 곱을 모두 더한 다음 바이어스(b)를 더한 값
가중합의 결과를 놓고 1 또는 0을 출력해서 다음으로 보냄

활성화 함수 : 0과 1을 판단하는 함수 ex) 시그모이드 함수

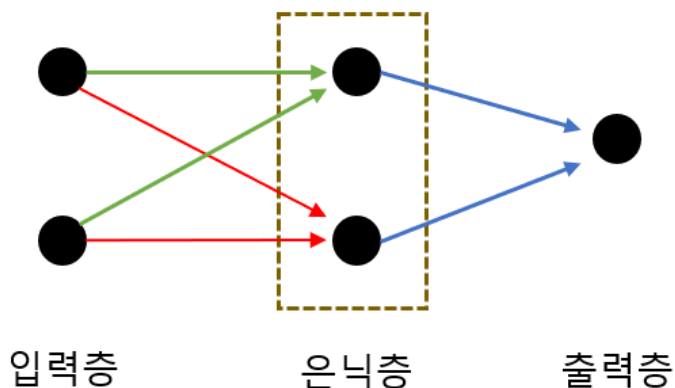
3. XOR 문제



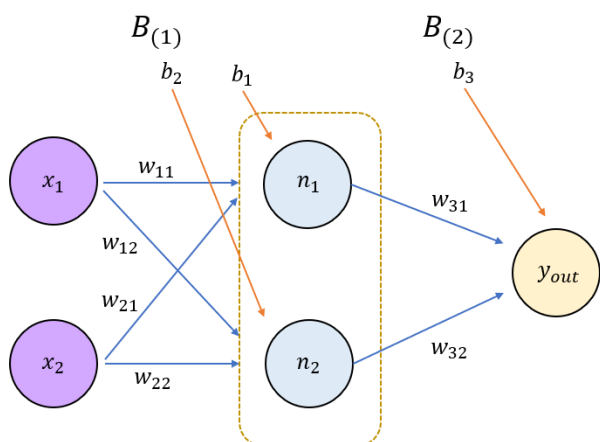
XOR문제를 어떻게 해결할 것인가? => 다층 퍼셉트론

다층 퍼셉트론

두 개의 퍼셉트론을 한 번에 계산할 수 있도록 숨어있는 층, 즉 은닉층을 만들어야함



1. 다층 퍼셉트론의 설계



1. 은닉층으로 퍼셉트론이 각각 자신의 가중치와 바이어스값을 보낸다.
2. 은닉층에서 모인 값이 활성화 함수를 이용해 최종 값으로 결과를 보낸다.

- 노드(n1, n2) : 은닉층에 모이는 중간 정거장
- n1과 n2의 값은 각각 단일 퍼셉트론의 값과 같음
- 활성화 함수 => 시그모이드 함수(σ)

$$\begin{aligned} n_1 &= \sigma(x_1w_{11} + x_2w_{21} + b_1) \\ n_2 &= \sigma(x_1w_{12} + x_2w_{22} + b_2) \\ y_{out} &= \sigma(n_1w_{31} + n_2w_{32} + b_3) \end{aligned}$$

두 식의 결과값이 출력층으로 보내짐
출력값

- 가중치와 바이어스의 값 정하기
- 은닉층을 포함해 가중치 6개와 바이어스 3개가 필요

$$W(1) = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \quad B(1) = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$W(2) = \begin{bmatrix} w_{31} \\ w_{32} \end{bmatrix} \quad B(2) = [b_3]$$

⇒ 이 부분
보완하기
(4/11)