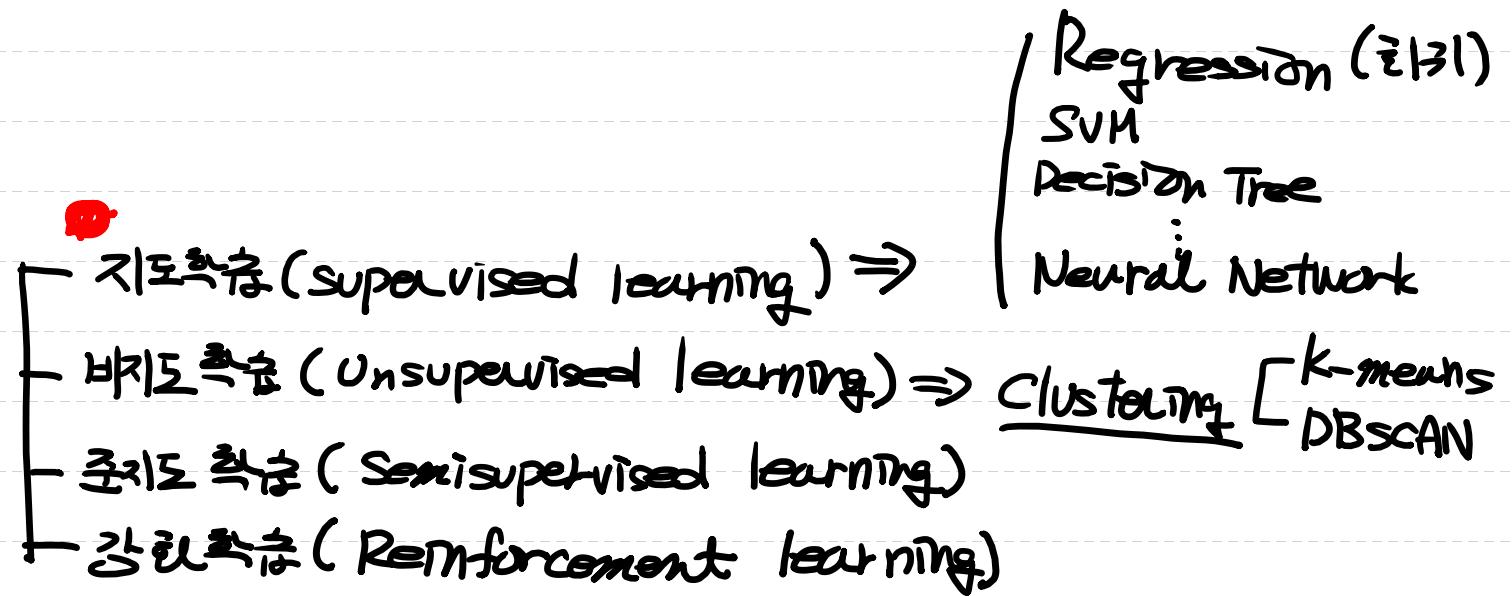


9/22. (화)

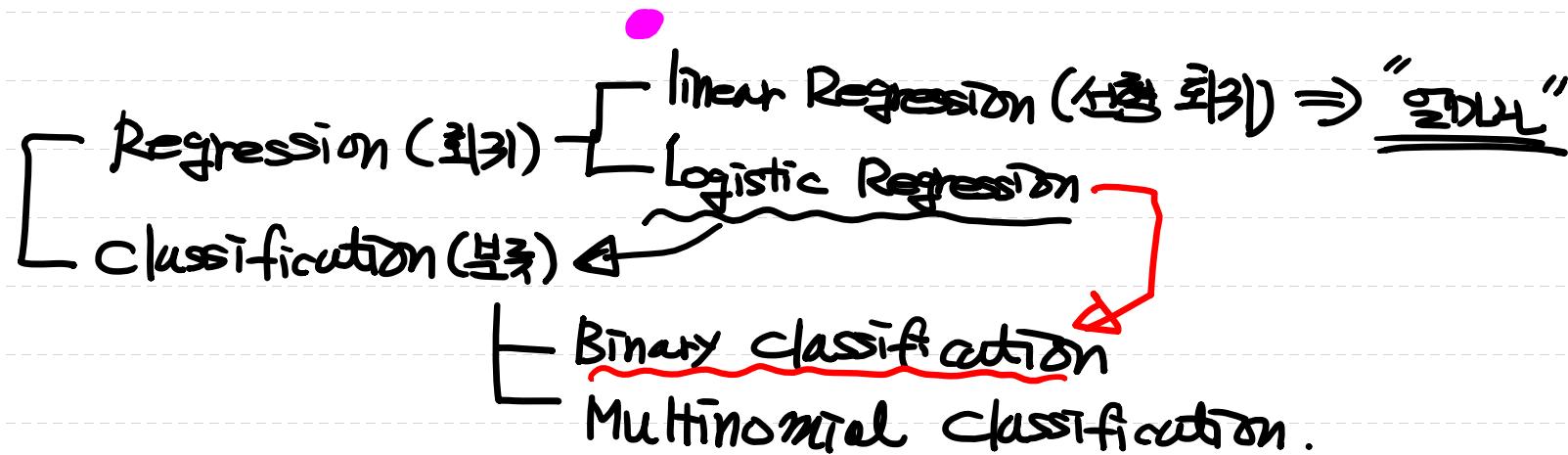
• AI

• Machine learning



• Deep learning \rightarrow NN (Neural Network), CNN, RNN, LSTM, GAN

• 지도학습



- Numerical derivative
differentiation

- 수치미분을 구현할 때

\Rightarrow 중앙차분

- "미분 기본공식"

- 편미분(partial derivative)

$$* \frac{\partial f(x,y)}{\partial x}$$

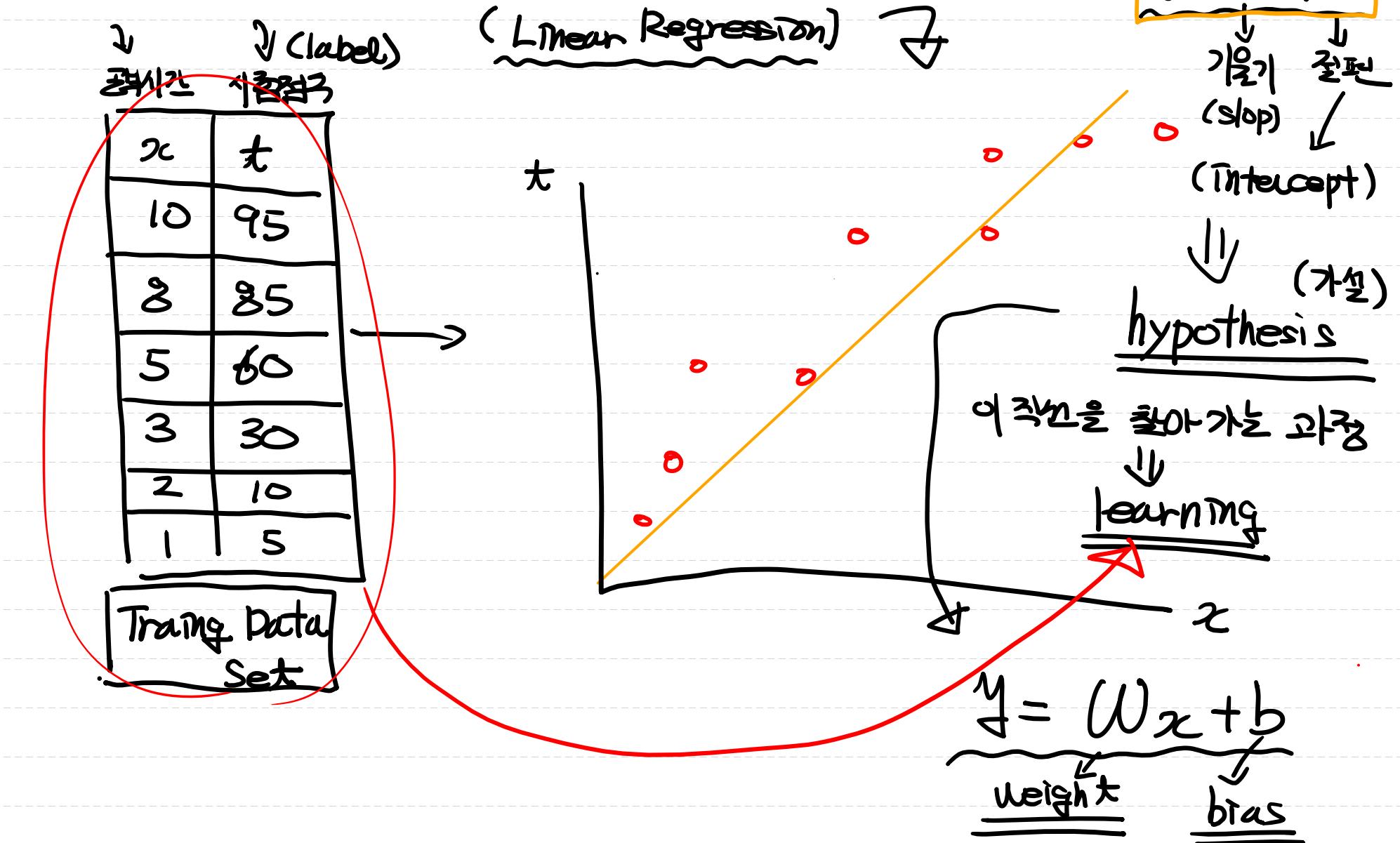
- * 연쇄법칙(chain Rule)

Python으로 구현

- 1. 일변수 함수에 대한 수치미분 코드
- 2. 다변수 함수에 대한 수치미분 코드.

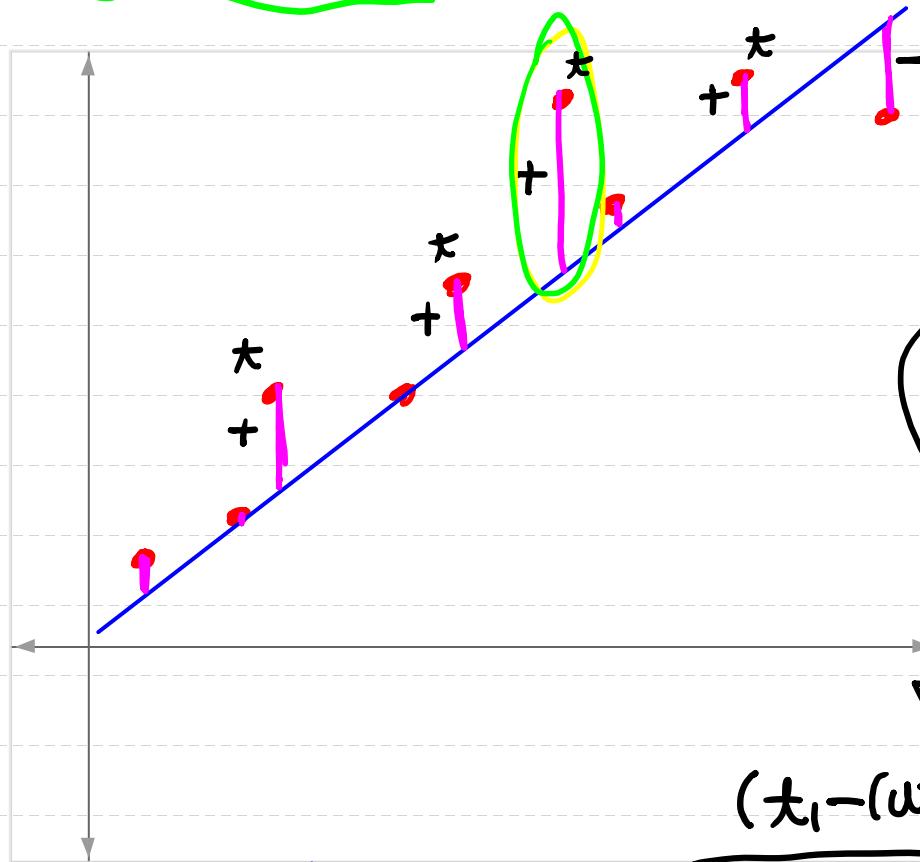
코드가 어려워요!

Machine Learning의 Regression!



$$y = \underline{w}^T x + b$$

(Hypothesis) $\begin{bmatrix} \text{초기 } w \\ b \end{bmatrix}$ random



오차(error)를 이용해서 학습 과정이
데이터를 잘 표현하고 있는지를 구하는

오차의 총이 작으면 좋은 예측입니다

최소제곱법 (least squared Method)

⇒ 오차들을 각각 제곱해서 평균을 구해요!

$$\frac{(t_1 - y_1)^2 + (t_2 - y_2)^2 + \dots + (t_n - y_n)^2}{n}$$

$$\frac{(t_1 - (w_1 x_1 + b))^2 + (t_2 - (w_1 x_2 + b))^2 + \dots}{n}$$

(
수학적
비용함수)

$\begin{bmatrix} \text{Loss function} \\ \text{Cost function} \end{bmatrix} \Rightarrow$

$$\frac{1}{n} \sum_{i=1}^n (t_i - (w_i x_i + b))^2$$

지금까지의 과정을 그림으로 표현.

* $y = w_2x + b$

(Linear Regression Model)
Hypothesis



loss function

$$E(w, b) = \frac{1}{n} \sum_{i=1}^n (t_i - (w_2x_i + b))^2$$

(시즈) (값)

x	t
10	95
8	80
5	60
2	10
1	5

Loss function
값 계산

계산하기

계산된 loss
function이 값이
최대인가??

Yes

정답
종료!

(w) (b)
Weight, bias update

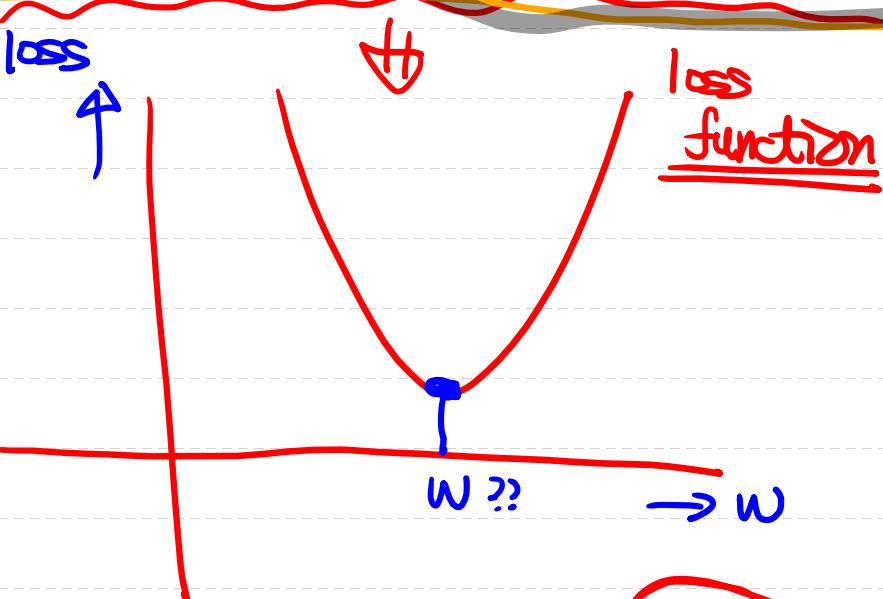
No

$$w = w - \alpha \frac{\partial E(w, b)}{\partial w}, \quad b = b - \alpha \frac{\partial E(w, b)}{\partial b}$$

Training Data Set

* loss function

$$\bullet E(w, b) = \frac{1}{n} \sum_{i=1}^n (t_i - (w x_i + b))^2$$



loss function의 값을 당연히

작을게 좋아요! !

loss function 을 최소로 만드는

w, b 를 찾을때까지 엎다변 ..

아마도 w, b 가 Hypothesis 를 만드는
값이 되요!!

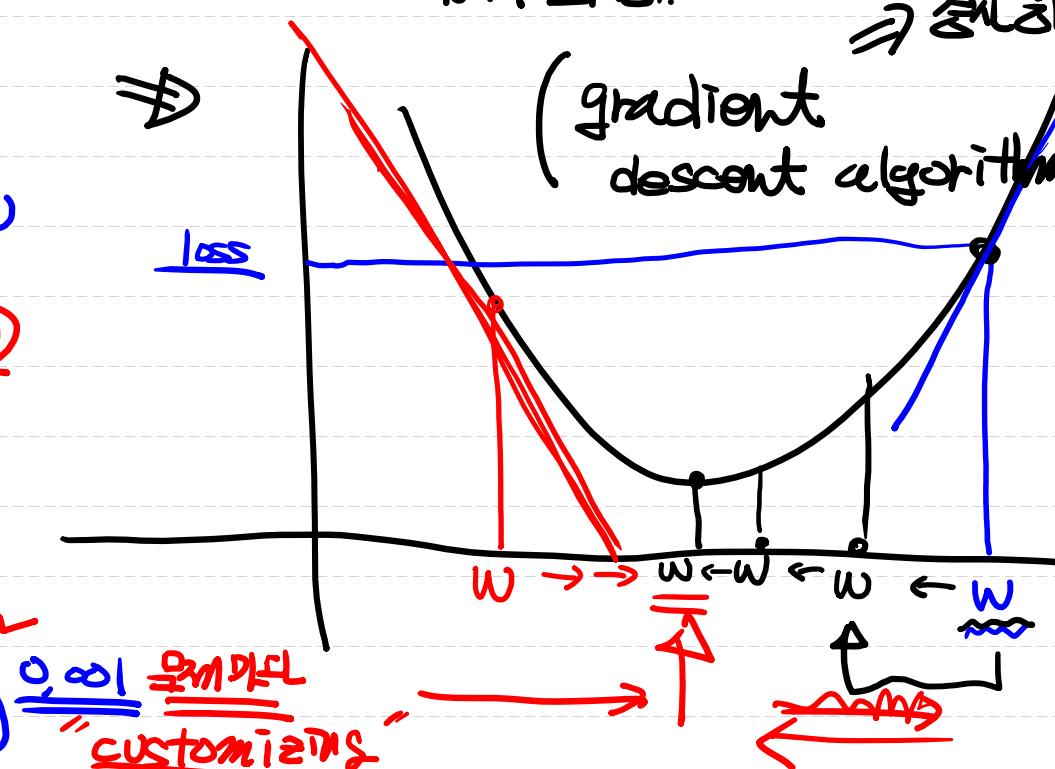
⇒ 경사하강법.

(gradient
descent algorithm.)

* $w = w - \alpha \frac{\partial E(w, b)}{\partial w}$

가장 가까운

α = learning rate (학습률) 0 ~ 1 무제한
"customizing"



Hypothesis (가설)

Linear Regress Model

$$y = \underline{w}x + b \quad \text{1개} *$$

2차원
(시즈) (점3) 2차원 행렬

H \downarrow $wx + b$

" $y = X \cdot w + b$ "

x	t
10	95
8	78
5	40
2	10
1	3

$$(1 \times 1) \cdot (5 \times 1) + b = (5, 1)$$

$$\underline{w}$$

$$\underline{x}$$

$$\underline{y}$$

$$\begin{aligned} x_1 w + b &= y_1 \\ x_2 w + b &= y_2 \\ x_3 w + b &= y_3 \\ x_4 w + b &= y_4 \\ x_5 w + b &= y_5 \end{aligned}$$

\Rightarrow

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$$

$$\bullet w^T = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix}$$

$$\begin{bmatrix} 5, 1 \end{bmatrix} \begin{bmatrix} 1, 1 \end{bmatrix} = [5, 1]$$

회귀모델이란?

(어떠한 데이터에 대해
그 값에 영향을 주는 조건을 고려해서
데이터의 흐름을 구하기 위한 흐름)

회귀(Regression) \Rightarrow 흐름으로의 회귀

데이터가 회귀모델 제작하는 추세 \rightarrow 회귀(돌아간다)
(regression coefficient)
회귀계수

$$\tilde{y} = h(x_1, \tilde{x}_2, \tilde{x}_3, \dots, \tilde{x}_k; \beta_1, \beta_2, \beta_3, \dots, \beta_k) + \epsilon$$

↑
평균을 구하는 흐름 (예측)
↓

실제치
점수

최적목표 \rightarrow $h(\cdot)$ 를 최대한 실제의 가깝게 만드는게 목표.

우리가 모델을 왜 만드는 걸까?? 우리가 해결해야 하는 현실을 좀 더 드러낸 형태로 표현하기 위해서. 만약 현실 세계의 문제를 드러내기 위해 Regression model을 만들면 독립변수와 종속변수의 의존 관계를 규칙화(추적) \rightarrow 미래의 예측치

우리가 해줄해야 하는 문제를
언제 시기적으론 어떻게 해야 학나요?? \Rightarrow 불필요한 정보는 버려요..

(1) 설정 \rightarrow 회귀모델의 assumption
(가정)

\rightarrow 회귀모델을 만들 때,

\rightarrow 실제 데이터는 일정한 특성을 가지고 있어요.

이런 가정이 많을수록 \rightarrow 모델은 유효해집니다.

"다양한 회귀모델이 존재"

[Linear Regression Model]
[Non-linear " Model]

회귀모델에서 (선형)은 구불한 기준이
비선형)은 직선이나 원뿔 등의 경우
 $(+, -)$

Linear Regression Model은 회귀계수(β)가 선형으로 결정되는 모델

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

$$\begin{aligned} y &= \beta_0 x^{\beta_1} \log(\beta_0 x^{\beta_1}) \\ \log(y) &= \log(\beta_0) + \log(x^{\beta_1}) \\ &= \log(\beta_0) + \beta_1 \log x \end{aligned}$$

$$\text{비선형 회귀} \Rightarrow \hat{y} = \frac{\beta_1 x}{\beta_2 + x}$$



Linear Regression (정렬 → 회귀계수를 추정하기 쉽고 해석하기 쉬워요.)

$$\hat{y} = \underline{w} \cdot \underline{x} + b$$

서로운 특성 (feature)을 추가해 추가로 학습하세요.

중복변수의 개수

단변량 회귀모델

다변량 //

classical Linear Regression Model



$$\hat{y} = \beta_0 + \sum_{i=1}^p \beta_i x_i$$

독립변수의 개수 → 한 Simple Linear

Regression M.

$$\hat{y} = \beta_0 + \beta_1 x_1$$

$$\rightarrow \hat{y} = \cancel{\alpha} x + b \rightarrow \hat{y} = \underline{w} x + b$$

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$