변수가 여러개인 Linear Regression

⟨ Hypothesis ⟩

$$H(x) = Wx + b \quad \Rightarrow \quad H(x_1, x_2, x_3) = w_1x_1 + w_2x_2 + w_3x_3 + b$$

⟨ Cost function ⟩

$$cost(W,b) = \frac{1}{m}\sum_{i=1}^{m}\left(Hx^{(i)} - y^{(i)}\right)^2 \quad \Rightarrow \quad cost(W,b) = \frac{1}{m}\sum_{I=1}^{m}\left(\overbrace{\underbrace{H(x_1^{(i)}, x_2^{(i)}, x_3^{(i)})}_{\text{예측값}} - \underbrace{y^{(i)}}_{\text{실제값}}}^{\text{차이}}\right)^2$$

- 항이 많아질수록 식이 길어지는데? → **Matrix 를** 이용한다.

$$w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_nx_n$$ 너무길어진다!

$$(x_1 \;\; x_2 \;\; x_3) \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = (x_1w_1 + x_2w_2 + x_3w_3)$$

$$H(X) = XH$$

x와 w의 순서가 바뀌어도 사실상 같음

대문자로 쓰는건 Matrix라는 표현!

- Data Sample 의 값을 instance 라고 한다.

- instance 가 많아져도 Matrix 를 이용하면 일일이 계산안해도 됨!

$$\begin{array}{cc} & \text{Variable} \\ \text{instance} & \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{53} \end{pmatrix} \end{array} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} x_{11}w_1 + x_{12}w_2 + x_{13}w_3 \\ x_{21}w_1 + x_{22}w_2 + x_{23}w_3 \\ \vdots \\ + x_{53}w_3 \end{pmatrix}$$

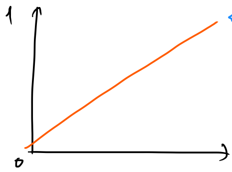
[5, 3] → Instance, Variable　　[3, 1]　　[5, 1] → Instance, Y

None이 들어갔다면? instance 수의 제한이 없다

Logistic Classification

- Classification = 분류
- 보통 **0과 1로** encoding 하여 학습시킨다. (binary classification)
- 문제점
- 0과 1로 linear Regression 을 수행하면 학습하는 모델에 따라 0과 1의 기준점이 계속 바뀌게 된다.
- linear Regression 에서 예측 함수가 0, 1보다 훨씬 큰 값을 뱉을 수도 있다.

$$H(x) = Wx + b$$

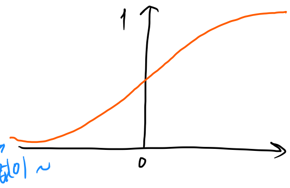


← 1을 뚫을 수 있다

Linear Regression

$$\downarrow$$

$$g(z) = \frac{1}{(1+e^{-z})}$$



← 1에 한없이 가까워지지만 1이 되지는 않음

0에 한없이 ~

Sigmoid function

⇒ cost 함수도 달라져야 한다!
gradient decent algolithm 사용못함!

{Hypothsis} $H(X) = XW \Rightarrow H(X) = \dfrac{1}{1+e^{-W^T X}}$

↳ log 함수

{Cost} $cost(W) = \frac{1}{m}\sum c(H(x), y) \Rightarrow c(H(x), y) = \begin{cases} -\log(H(x)) : y=1 \\ -\log(1-H(x)) : y=0 \end{cases}$

코드로 짜면 너무 복잡해서 수식 한 줄로 만듦

$$\Rightarrow c(H(x), y) = -y\log(H(x)) - (1-y)\log(1-H(x))$$

$$cost(W) = -\frac{1}{m}\sum y\log(H(x)) + (1-y)\log(1-H(x))$$