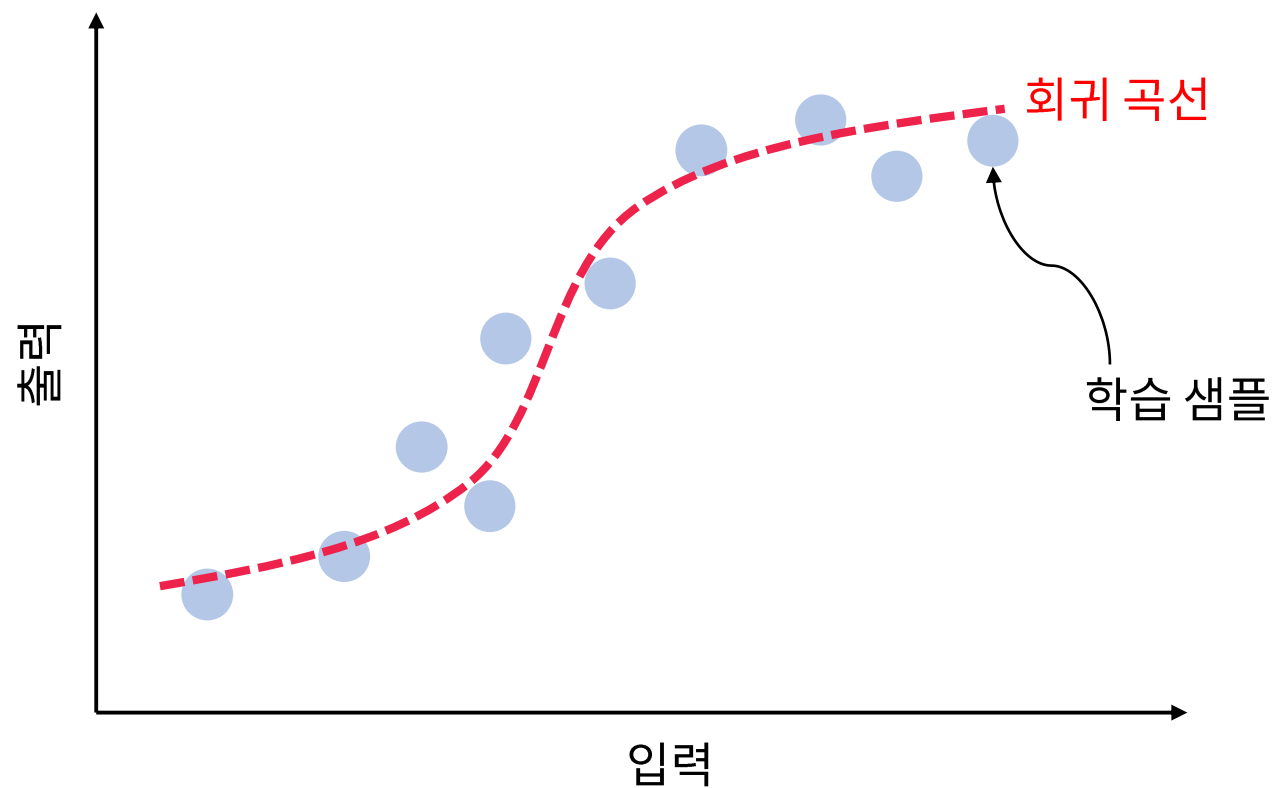


Chapter 02. 가장 단순한 신경망을 통해 작동 원리 이해하기

# STEP2. 회귀 문제의 이해

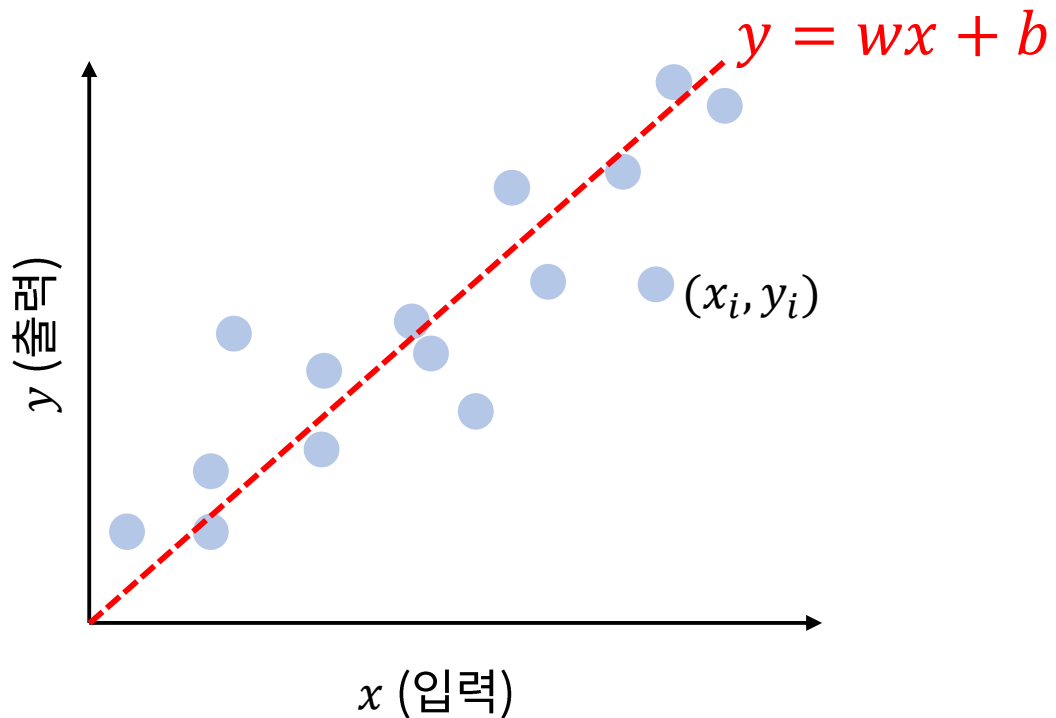
# 회귀 (Regression)



연속된 값을 추정하는 회귀(Regression)에 대해 조금 더 자세히 알아보시다.

# 단순 선형 회귀

선형 회귀(Linear Regression) : 데이터를 가장 잘 표현하는 선형식을 찾는 동작.



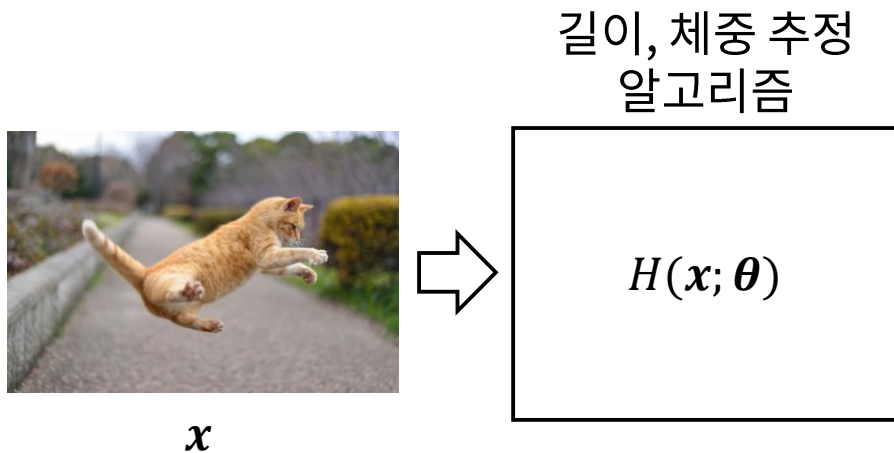
$$w^* = \arg \min_w \frac{1}{N} \left\{ \sum_i \frac{(y - y_i)^2}{2} \right\}$$

평균제곱에러 (Mean Squared Error; MSE)

“MSE를 최소로 하는  $w$ 를 찾아라.”

단순 선형 회귀의 예. 독립 변수(입력)이 하나이므로, 추정해야 할 변수도 하나.  
단, 편향을 포함하면 2개가 된다.

# 평균 제곱 오차 Mean Squared Error



$$y = [10.5, 3.6]$$

$$\Rightarrow \tilde{y} = [8.3, 4.6] \rightarrow E = (10.5 - 8.3)^2 + (3.6 - 4.6)^2 = 5.8$$

$$\tilde{y} = [9.4, 3.8] \rightarrow E = (10.5 - 9.4)^2 + (3.6 - 3.8)^2 = 1.2$$

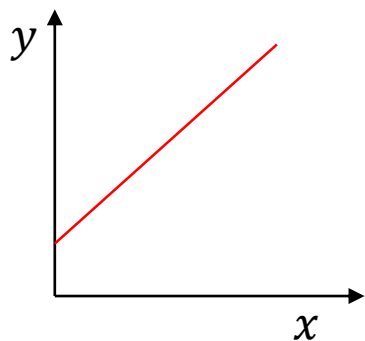
MSE를 이용해, 고양이의 길이와 체중의 오차를 종합적으로 판단할 수 있다.

# 다중 선형 회귀

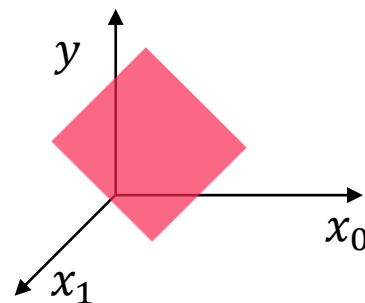
$$\begin{aligned} y = wx + b & \quad \longrightarrow \quad y = w_0x_0 + w_1x_1 + w_2x_2 + \cdots + w_{N-1}x_{N-1} + b \\ & \quad \quad \quad = \sum_{i=0}^{N-1} w_i x_i + b \\ & \quad \quad \quad = \mathbf{w}^T \mathbf{x} + b \end{aligned}$$

단일 입력이 아닌, 다중 입력을 받을 경우에는 변수가 확장되어 **벡터의 내적**이 된다.

# 다중 선형 회귀의 기하학적 해석



$$y = wx + b$$



$$y = w_0x_0 + w_1x_1 + b$$

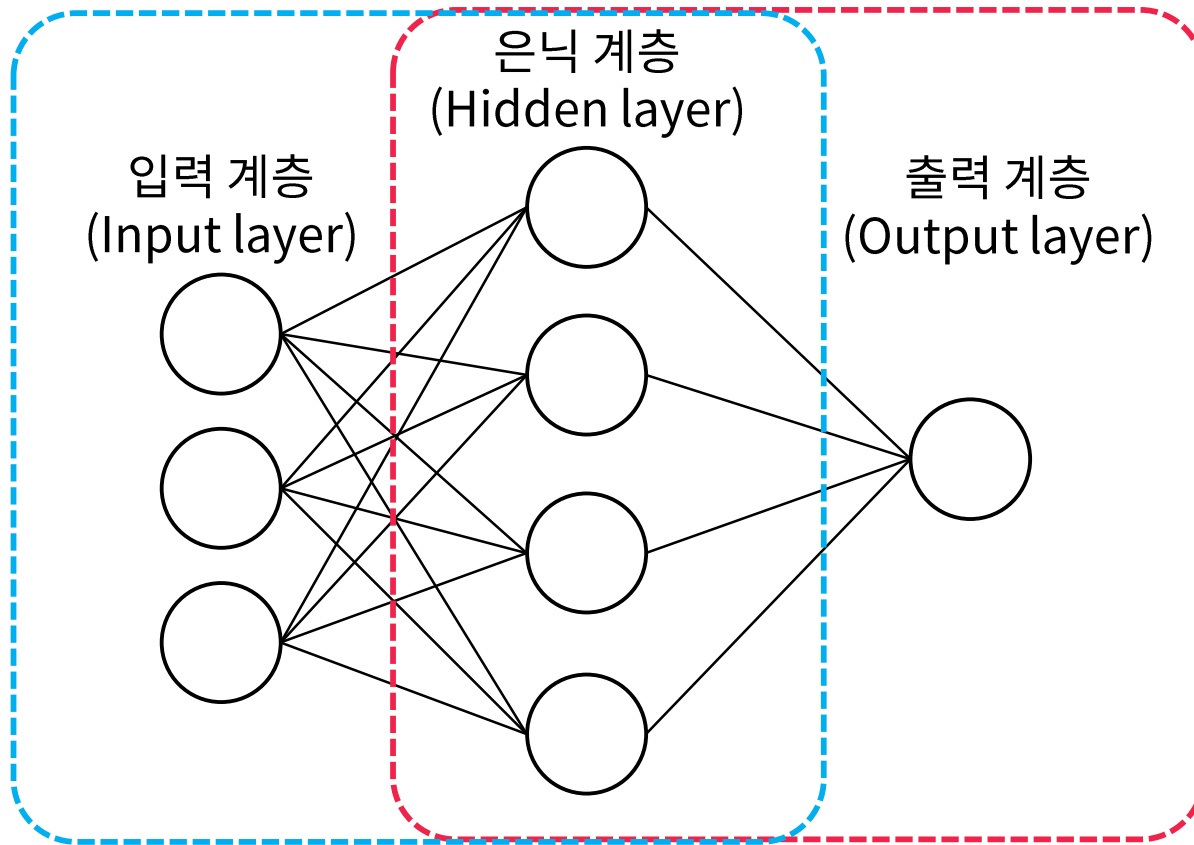


Hyper-Plane!!

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

변수가 하나 추가될 때 마다 차원이 하나씩 추가된다. 직선 → 평면 → 초평면

# 얇은 신경망과 회귀 알고리즘



Identity

$$y = a_o(\mathbf{w}_o^T \mathbf{h} + b_o)$$

$$= \mathbf{w}_o^T \mathbf{h} + b_o$$

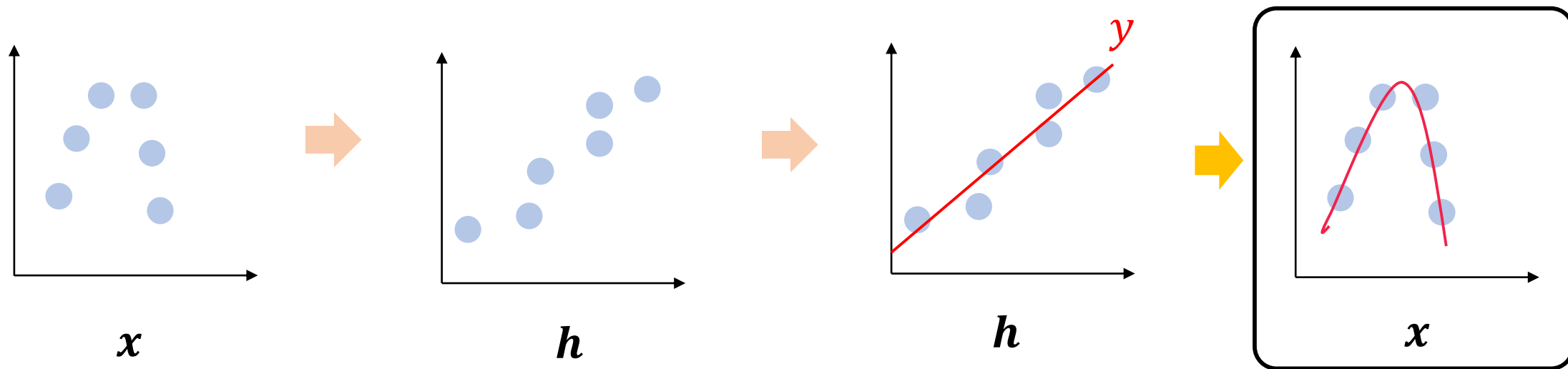
= 선형 회귀

$$\mathbf{h} = a_h(\mathbf{W}_h \mathbf{x} + b_h)$$

얇은 신경망으로 회귀를 수행할 경우, **출력 계층은 선형 회귀와 동일**하다.

입력 계층에서 **은닉 계층으로 추가적인 변환**이 있다는 것이 다른 점!

# 은닉 계층과 회귀



선형적으로 분포되지 않는 입력  $\rightarrow$  선형적으로 분포하는 은닉 계층(특징)  $\rightarrow$  선형 회귀  
입력 Space를 기준으로 보면 회귀 곡선이 된다!