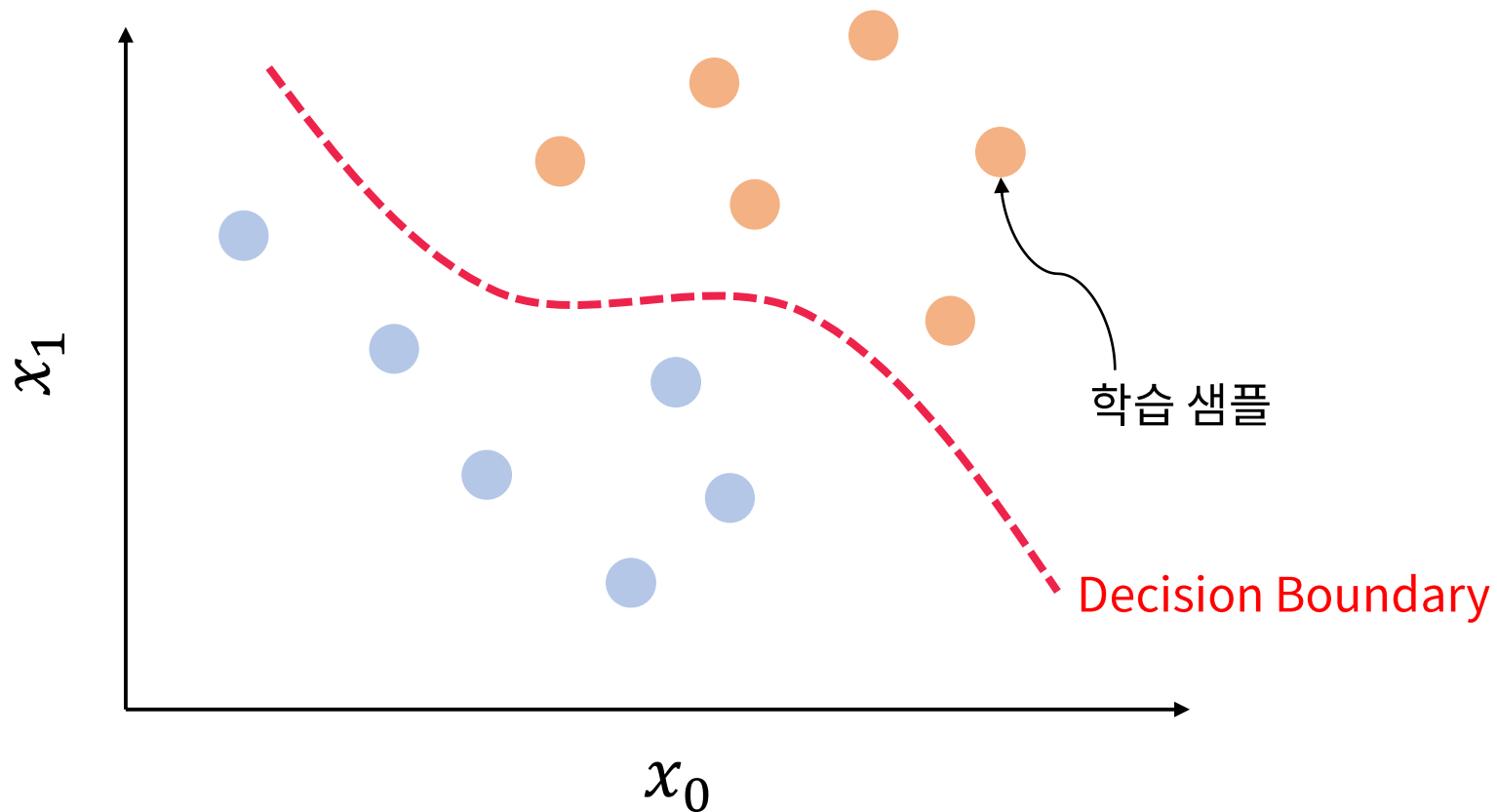


Chapter 02. 가장 단순한 신경망을 통해 작동 원리 이해하기

# STEP2. 이진 분류 문제

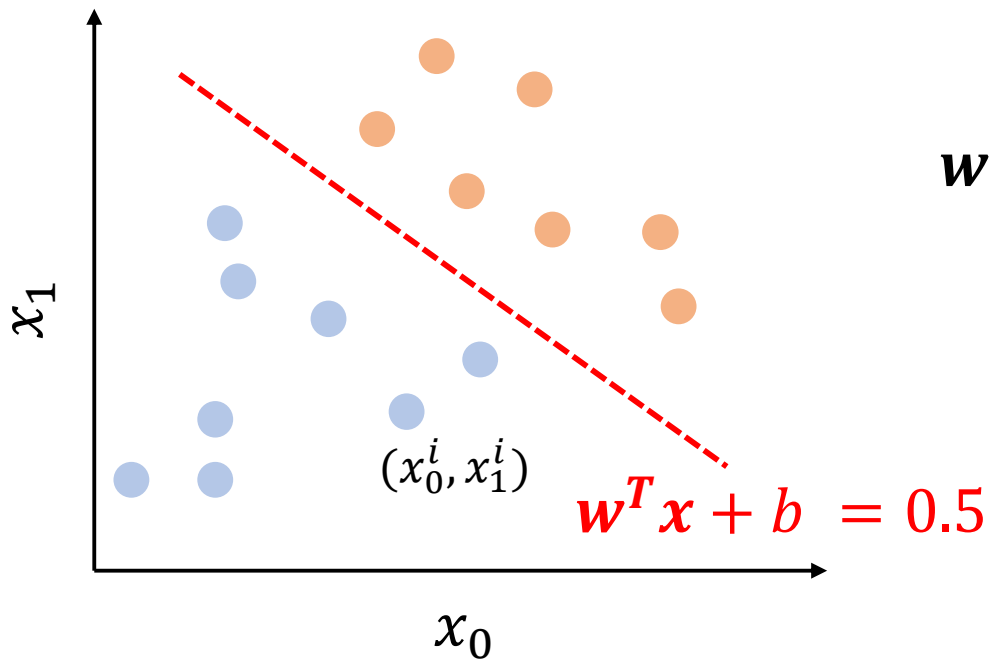
# 분류 (Classification)



입력의 범주를 추정하는 **분류(Classification)**에 대해 조금 더 자세히 알아보시다.

# 로지스틱 회귀

로지스틱 회귀(Logistic Regression): 범주형 데이터를 대상으로 하는 회귀. 분류 기법으로도 볼 수 있다.

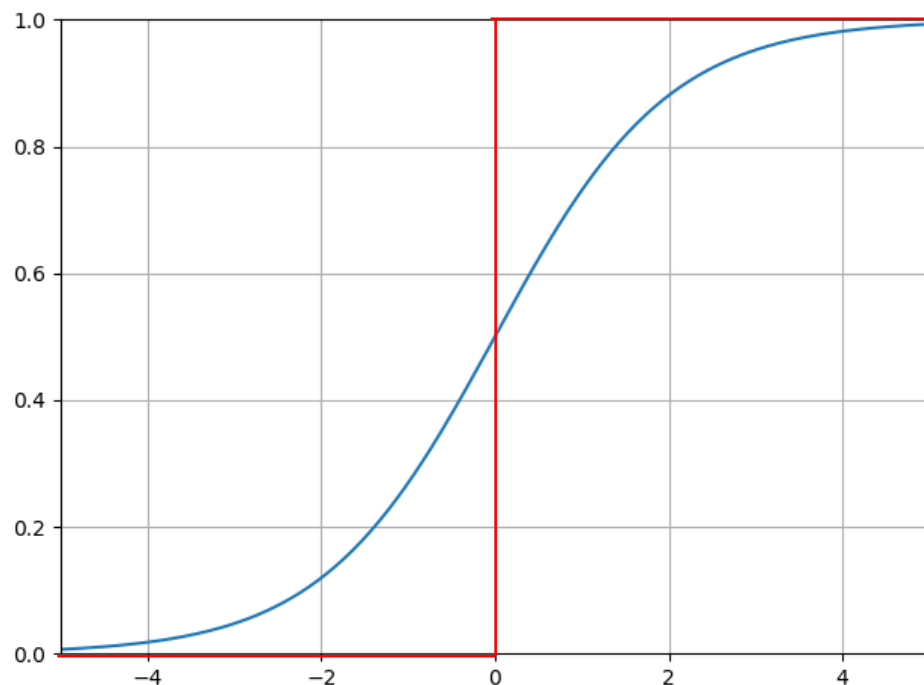


$$\mathbf{w}^* = \underset{\mathbf{w}}{\operatorname{argmin}} \{ -\sum_i y \log \tilde{y} + (1 - y) \log(1 - \tilde{y}) \}$$

“이진 교차 엔트로피를 최소화한다”

로지스틱 회귀는 선형 회귀와 비슷하나, 범주형 데이터를 분류하는 방향으로 선을 긋는다.

# Sigmoid Function



Sigmoid 함수

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$$

- 값이 작아질 수록 0, 커질 수록 1에 수렴
- 모든 실수 입력 값에 대해 출력이 정의됨
- 출력이 0~1 사이로, '확률'을 표현할 수 있음
- 입력 값이 0에 가까울 수록 출력이 빠르게 변함
- 모든 점에서 미분 가능

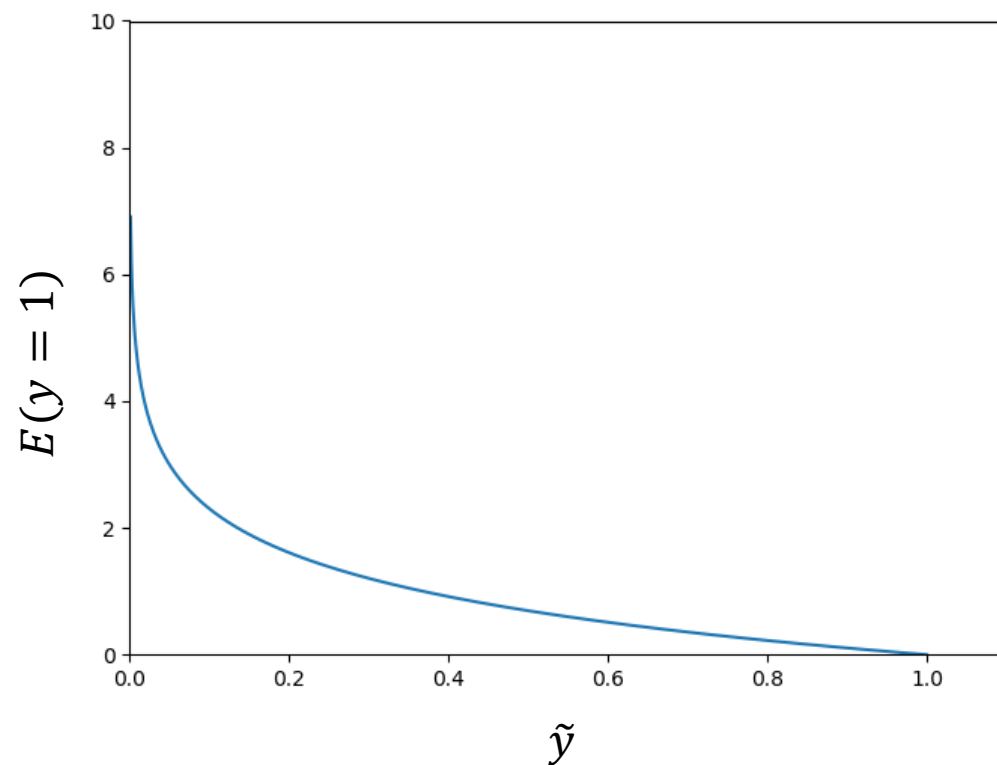
# 교차 엔트로피 오차

교차 엔트로피 오차 (Cross entropy error; CEE)

$$E = - \sum_i y \log \tilde{y} + (1 - y) \log(1 - \tilde{y})$$

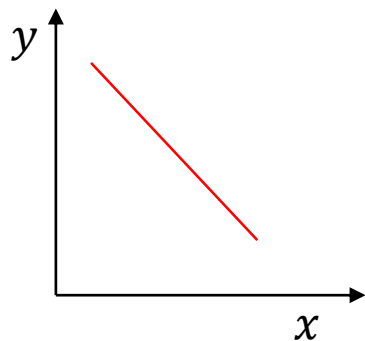
$y$  : 학습 데이터 정답 (0 or 1)

$\tilde{y}$  : 학습 데이터 입력으로 추정한 출력 (0 ~ 1)

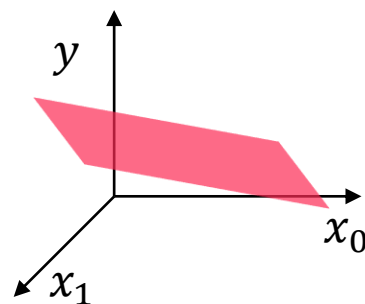


정확히 맞추면 오차가 0, 틀릴수록 오차가 무한히 증가하는 특징이 있다.

# 다중 로지스틱 회귀의 기하학적 해석



$$y = wx + b$$



$$y = w_0x_0 + w_1x_1 + b$$

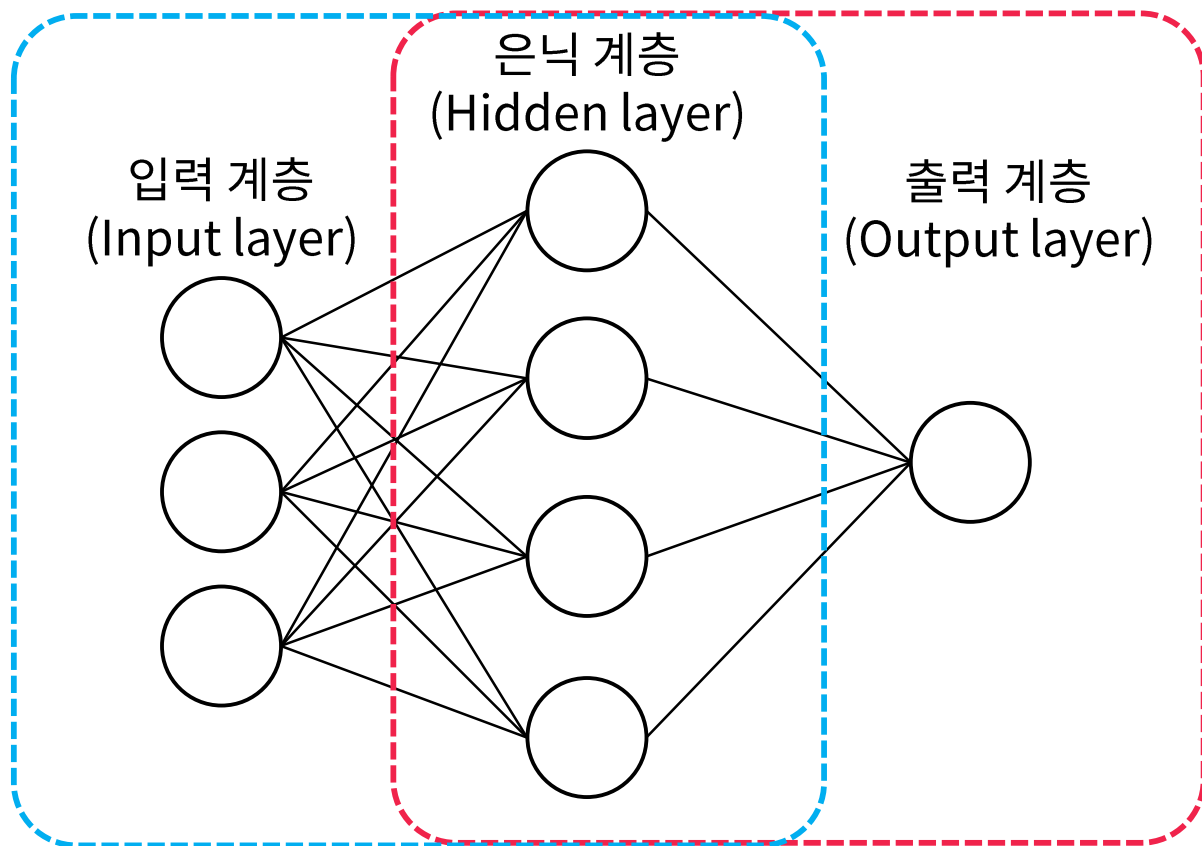


Hyper-Plane!!

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

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

# 얇은 신경망과 분류 알고리즘



Sigmoid

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

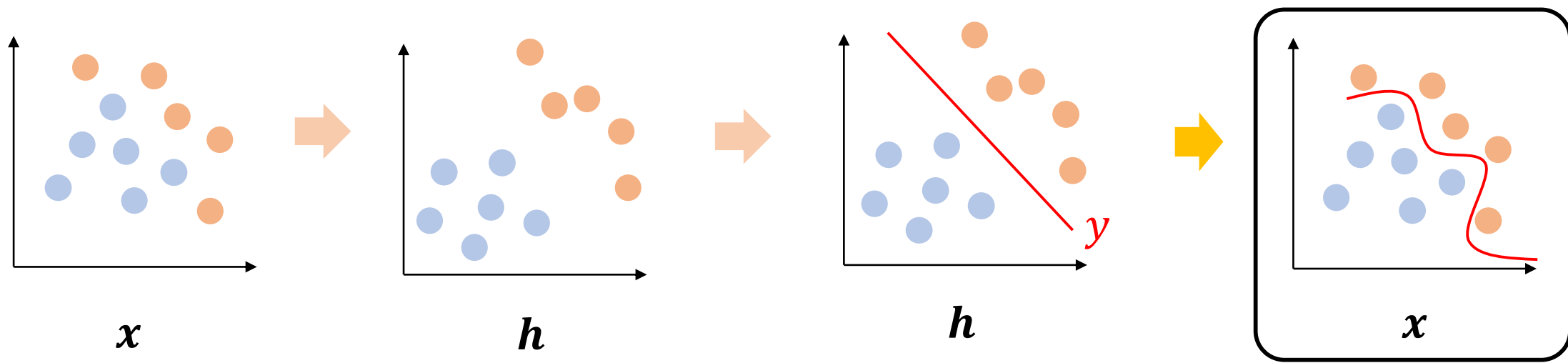
= 로지스틱 회귀

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

얇은 신경망으로 Classification을 수행할 경우, **출력 계층은 로지스틱 회귀와 동일**하다.

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

# 은닉 계층과 분류



선형적으로 분리되지 않는 Class → 선형적으로 분리되는 은닉 계층(특징) → 로지스틱 회귀  
입력 Space를 기준으로 보면 Decision Boundary가 곡선이 된다!