



6장. 분류

되짚어보기



인공지능 모델

$\text{predict} = w_1 * x_1 + w_0$ (모델의 형태)

($W = [w_1, w_0]$ (매개변수(=가중치)))

모델 정의 -> 가중치에 따른 오차 확인 -> 가중치 최적화
(ex. 경사하강법)

회귀와 분류



회귀모델

모델(ex. $w_1 \cdot x_1 + w_0$) -> 평균제곱오차(MSE)(데이터와 모델의 예측값간의 거리차이)-> 가중치 최적화(ex. 경사하강법)

분류모델

로지스틱 회귀 모델

{모델(ex. $w_1 \cdot x_1 + w_0$) -> 확률적해석을 위한 시그모이드}

-> 교차엔트로피 오차(데이터의 정답 클래스에 대한 예측 확률값의 합)-> 가중치 최적화(ex. 경사하강법)

회귀와 분류



회귀

목표 데이터가 연속적인 수치값(선형)인것. 예측값

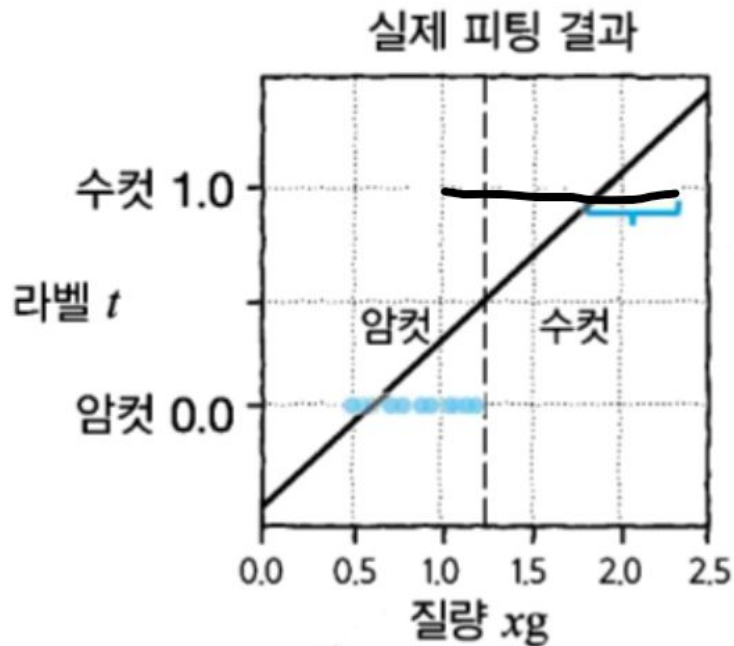
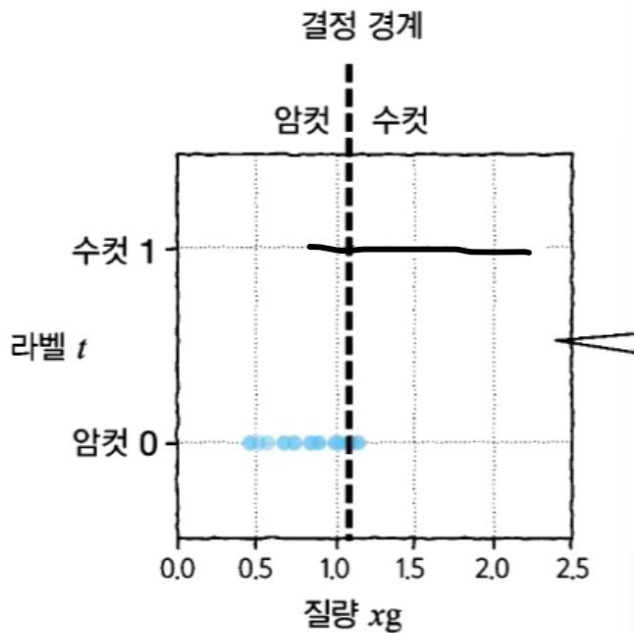
ex. 나이 몸무게 정보(input)을 이용해 키(output)를 예측

분류

목표 데이터가 비연속적인 class(비선형)인것. 카테고리분류

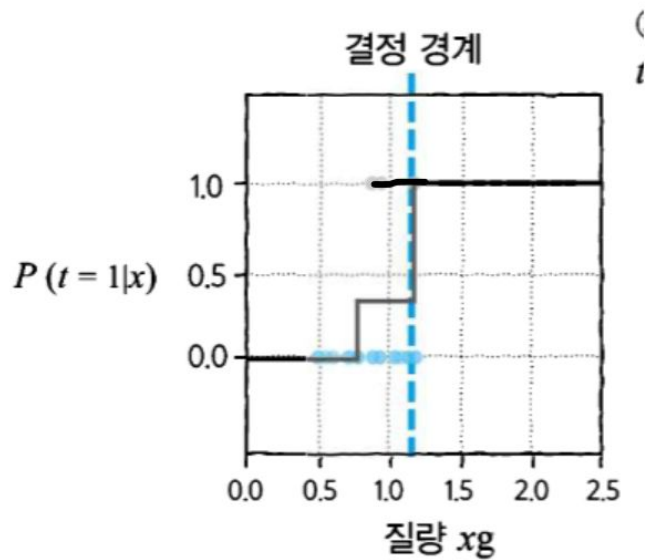
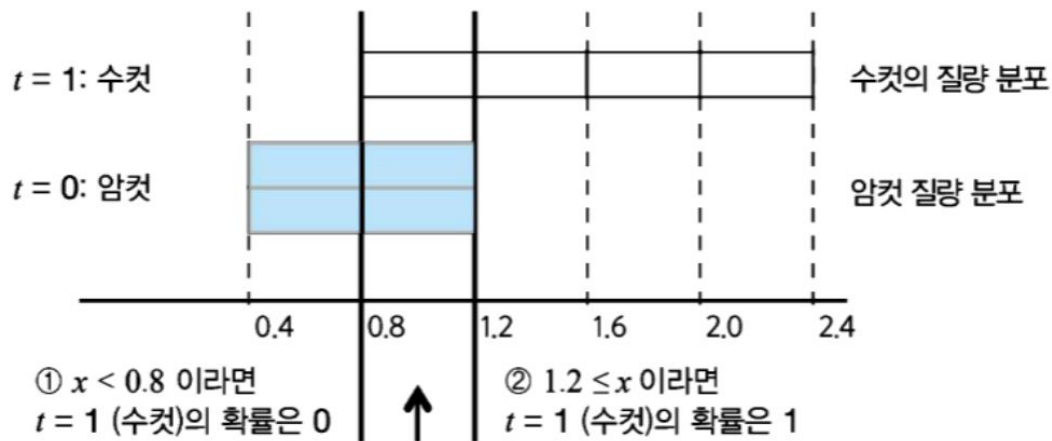
ex. 몸무게와 키 정보(input)을 이용하여 저체중과 정상체중, 과체중을 구분하는것

분류문제를 회귀모델로 접근했을때의 문제

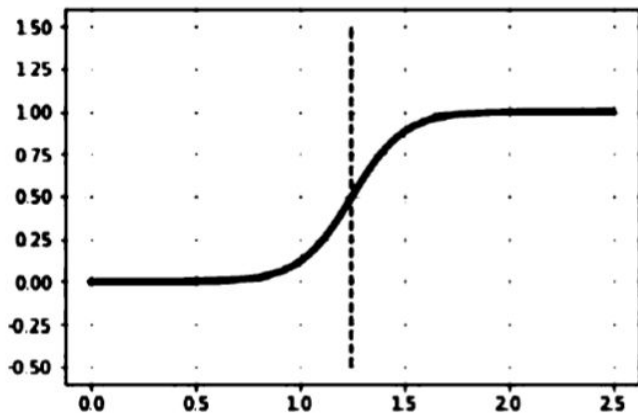


확률적 관점

질량 x 에 대한 '수컷일 확률' $P(t=1|x)$



로지스틱 회귀모델



시그모이드 함수

직선 모델

$$y = w_0x + w_1$$

직선 모델이 시그모이드
함수를 통과

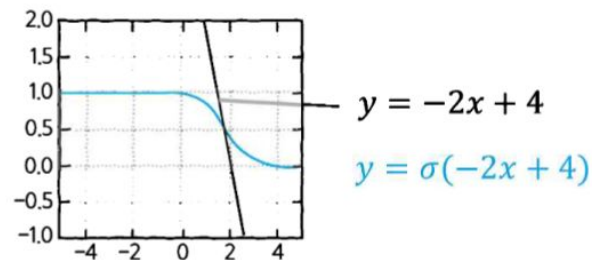
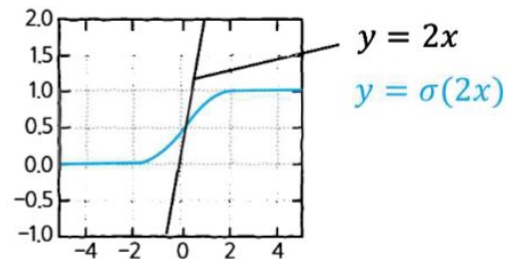
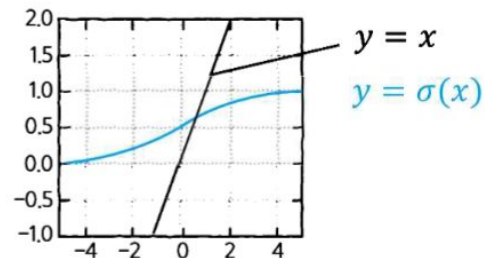
로지스틱 회귀 모델

$$y = \sigma(w_0x + w_1)$$

시그모이드 함수

$$\sigma(x) = \frac{1}{1 + \exp(-x)}$$

(4.7.5절 참조)



가능도(우도)



주어진 데이터 집합이 발생할 확률

$$T = [1, 1, 0, 1, 0] \quad P(t=1) = 0.3$$

T의 가능도는 $0.3 \times 0.3 \times 0.7 \times 0.3 \times 0.7$

$$= 0.3^3 (\text{T에서 } t=1 \text{인 횟수}) \times 0.7^2 (\text{T에서 } t=0 \text{인 횟수})$$

가능도(우도)

t_n (= 데이터 T에서 클래스 n의 출현횟수)

y_n (= 데이터 T에서 클래스 n의 발생확률) 이라면,

아까 예시는 $y_1^{t_1} * y_0^{t_0}$ 으로 볼 수 있다.

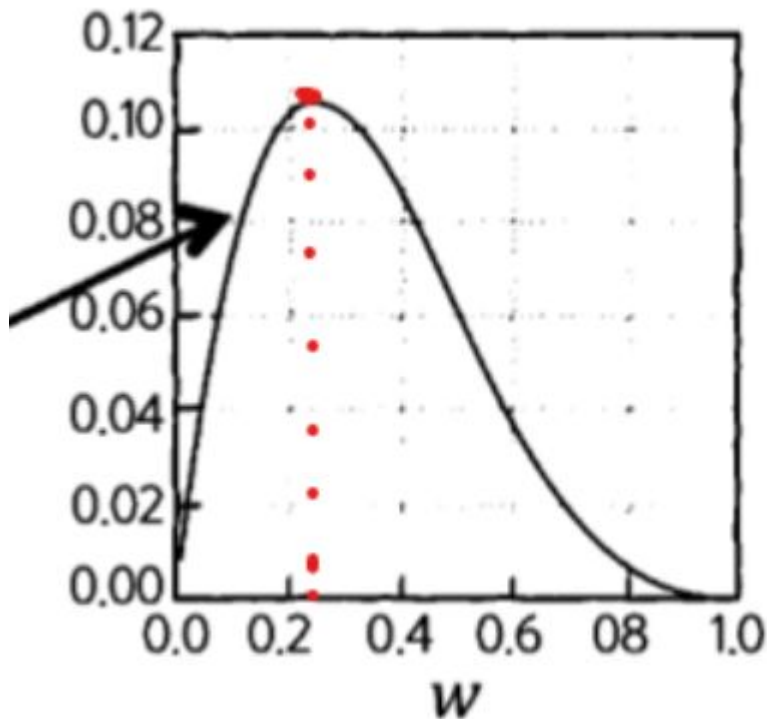
이를 n클래스가 있을때로 정리하면,

$$y_0^{t_0} * y_1^{t_1} * (...) * y_{(n-1)}^{t_{(n-1)}} = \prod_{n=0}^{N-1} y_n^{t_n} \Rightarrow \sum_{n=0}^{N-1} t_n \log y_n$$

최대가능도법

확률적으로 가장 발생할
확률(가능도)가 가장 높은
 $P(t=1)$ 가

주어진 데이터의 실제
발생확률일 것이다.



교차 엔트로피 오차

데이터의 T의 가능도를
모두 더하는 오차계산법

실제 X에 대한 데이터 T는 정답
클래스외엔 0으로 표시하므로

정답클래스에 대해 모델이 예측한
확률을 모으기 때문에 값이 최대가
될수록 좋다.

평균제곱오차처럼 최소를 구하기
위해 -1을 곱하여 사용한다.

$$-\prod_{n=0}^{N-1} \prod_{k=0}^{K-1} y_{nk}^{t_{nk}}$$

아래식은 계산의 편의성을 위하여 log를 취해준 식

$$-\sum_{n=0}^{N-1} \sum_{k=0}^{K-1} t_{nk} \log y_{nk}$$

평균 교차 엔트로피 오차



$$-\frac{1}{K} \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} t_{nk} \log y_{nk}$$

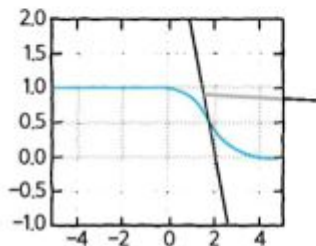
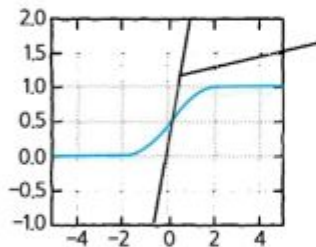
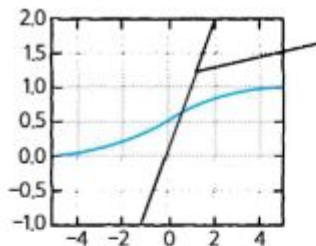
K는 데이터 수

정리

$w_{20}x_2 + w_{10}x_1 + w_0$
(클래스 0)

$w_{21}x_2 + w_{11}x_1 + w_0$
(클래스 1)

$w_{22}x_2 + w_{12}x_1 + w_0$
(클래스 2)



softmax
정규화

정답
클래스에
대한
확률만
수집
(교차엔트
로피오차)

→ 최적화
(경사
하강법)