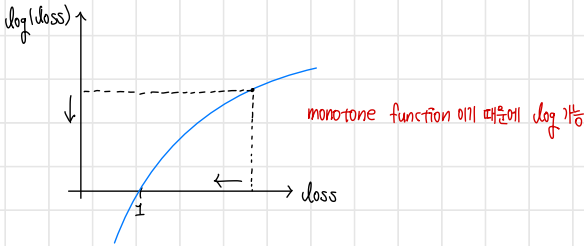


1. Binary Cross Loss Entropy

- ① Cross Entropy: 분류에 적합한 loss function
- ② Sigmoid를 이용한 이진 분류 → 출력값을 최대한 키우자. $\text{maximize } \log(-1 \pm p)^{\pm 1}$
- ③ 2장 이상일 경우 → loss function 이 계속해서 좁해지게 됨 → log를 취해준다.



2. Logistic regression

logit → sigmoid → probability

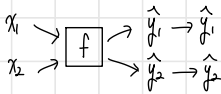
↳ sigmoid 통과 직전이 logit 이다. ⇔ logit을 linear regression 으로 봐도 된다.

3. 왜 MSE loss는 사용할 수 없을까

4. 다중 분류

① n 개의 class → n 개의 확률 필요 → 출력 노드를 이 확률이 control 하게 됨.

(cf. 이진 분류의 경우 2개의 class → 2개의 확률 분류)



② 머신의 출력으로 각 레이블의 확률이 출력되도록 하자

③ 새로운 Activation function 필요 → softmax function

→ input 으로 들어오는 확률값을 받아 output 으로 score 값을 뱉어냄.

④ MSE로 생각하자.

⑤ Multinomial logistic regression.

