04.3. 분류 문제 기울기

- 손실 함수 : 교차 엔트로피 오차
- 은닉층 활성화 함수 : 시그모이드
- 출력층 활성화 함수 : 소프트맥스 함수
- 출력층 가중치 기울기

$$\delta_k = \frac{\partial \mathrm{E}}{\partial y_k}\frac{\partial y_k}{\partial u_k}$$

교차 엔트로피 오차는 다음과 같이 표현

$$\mathrm{E} = -\sum_{k=1}^{n} t_k \ \log(y_k)$$

소프트맥스 함수는 다음과 같음

$$y_k = \frac{\exp(u_k)}{\sum_{k=1}^{n} \exp(u_k)}$$

교차 엔트로피 함수의 y_k 부분에 소프트맥스 함수를 대입하면 아래와 같음

$$\mathrm{E} = -\sum_{k=1}^{n} t_k \ \log(\frac{\exp(u_k)}{\sum_{k=1}^{n} \exp(u_k)})$$

$\log\frac{p}{q} = \log p - \log q$ 이므로 위의 식은 다름과 같이 전개 가능

$$\mathrm{E} = -\sum_{k=1}^{n}(t_k \ \log(\exp(u_k)) \ - \ t_k \sum_{k=1}^{n} \exp(u_k))$$

$$= -\ \textstyle\sum_{k=1}^{n}(t_k \ \log(\exp(u_k))) \ - \ \sum_{k=1}^{n}(t_k log \sum_{k=1}^{n} \exp(u_k))$$

$$= -\ \textstyle\sum_{k=1}^{n}(t_k \ \log(\exp(u_k))) \ - (\sum_{k=1}^{n} t_k)(log \sum_{k=1}^{n} \exp(u_k))$$

$\log(\exp(x)) = x$ 이고, 분류 문제에서 단 하나의 정답만이 1이고 나머지는 모두 0이므로,
$\sum_{k=1}^{n} t_k = 1$ 임
따라서 위 식은 다음과 같이 정리됨

$$E = -\sum_{k=1}^{n} t_k u_k + \log \sum_{k=1}^{n} \exp(u_k)$$

이 식을 δ_k 식에 대입하면

$$\delta_k = \frac{\partial E}{\partial u_k}$$

$$= \frac{\partial}{\partial u_k} \left(-\sum_{k=1}^{n} t_k u_k + \log \sum_{k=1}^{n} \exp(u_k)\right)$$

$$= -t_k + \frac{\exp(u_k)}{\sum_{k=1}^{n} \exp(u_k)}$$

$$= -t_k + y_k$$

$$= y_k - t_k$$

회귀 문제의 경우와 동일한 결과 산출, 이를 이용하여 나머지 기울기도 계산 가능

<출력층 기울기>

$$\delta_k = y_k - t_k$$

$$\partial w_{jk} = y_j \delta_k$$

$$\partial b_k = \delta_k$$

$$\partial y_j = \sum_{r=1}^{n} \delta_r w_{jr}$$

● 은닉층 기울기

$$\delta_j = \partial y_j \frac{\partial y_j}{\partial u_j}$$

은닉층의 활성화 함수는 시그모이드이므로 회귀 문제와 같은 방법으로 다음을 유도

<은닉층 기울기>

$$\delta_j = \partial y_j (1 - y_j) y_j$$

$$\partial w_{ij} = y_i \delta_j$$

$$\partial b_j = \delta_j$$

$$\partial y_i = \sum_{q=1}^{m} \delta_q w_{iq}$$

결국 회귀 문제와 분류 문제는 코드 재사용 가능