04.2. 회귀 문제 기울기

2020년 12월 29일 화요일 오후 8:55

- 손실 함수 : 오차제곱합
- 은닉층 활성화 함수 : 시그모이드
- 출력층 활성화 함수 : 항등 함수

- 출력층 기울기

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

$\frac{\partial \mathrm{E}}{\partial y_k}$ 는 손실 함수인 오차제곱합을 출력 y_k 로 편미분해서 구함

$$\frac{\partial \mathrm{E}}{\partial y_k} = \frac{\partial}{\partial y_k}\left(\frac{1}{2}\sum_{k=1}^{n}(y_k - t_k)^2\right)$$

$$= \frac{\partial}{\partial y_k}\left(\frac{1}{2}(y_0 - t_0)^2 + \frac{1}{2}(y_1 - t_1)^2 + \cdots + \frac{1}{2}(y_k - t_k)^2 + \cdots + \frac{1}{2}(y_n - t_n)^2\right)$$

$$= y_k - t_k$$

$\frac{\partial y_k}{\partial u_k}$ 는 출력층 활성화 함수를 미분해서 구함

출력층 활성화 함수는 항등함수 이므로 다음과 같음

$$\frac{\partial y_k}{\partial u_k} = \frac{\partial u_k}{\partial u_k} = 1$$

결국 $\delta_k = y_k - t_k$ 임

위 결과값을 이용하면 나머지 기울기값을 구할 수 있음

$$\delta_k = y_k - t_k$$

$$\partial \mathrm{w}_{jk} = y_j \delta_k$$

$$\partial \mathrm{b}_k = \delta_k$$

$$\partial \mathrm{y}_j = \sum_{r=1}^{n} \delta_r w_{jr}$$

● 은닉층 기울기

$$\delta_j = \partial \mathrm{y}_j \frac{\partial y_j}{\partial u_j}$$

위 식의 $\frac{\partial y_j}{\partial u_j}$ 은 활성화 함수를 편미분해서 구할 수 있음, 은닉층의 활성화 함수는 시그모이드 함수이고 시그모이드 함수 $f(x)$의 미분은 $f'(x) = (1 - f(x))f(x)$ 이므로 다음과 같다.

$$\frac{\partial y_j}{\partial u_j} = (1 - \mathrm{y}_j)\mathrm{y}_j$$

결국 δ_j 는 다음과 같다.

$$\delta_j = \partial \mathrm{y}_j (1 - \mathrm{y}_j)\mathrm{y}_j$$

위 결과값을 이용하면 나머지 기울기값을 구할 수 있음

$$\delta_j = \partial \mathrm{y}_j (1 - \mathrm{y}_j)\mathrm{y}_j$$

$$\partial \mathrm{w}_{ij} = y_i \delta_j$$

$$\partial \mathrm{b}_j = \delta_j$$

$$\partial \mathrm{y}_i = \sum_{q=1}^{m} \delta_q w_{iq}$$