

딥러닝 경사하강법 유도 / 뉴런 작성

정보컴퓨터공학과
권혁동

경사하강법 (Gradient Descent) 복습

경사하강법 유도

뉴런 작성

경사하강법 (Gradient Descent) 복습

- 선형회귀 기법 중 하나
 - **Weight(기울기)**와 **Bias(절편)**을 찾는 알고리즘
- 1. 무작위로 weight, bias 선택
- 2. 샘플 하나를 선택하여 예측 값 계산
- 3. 예측 값과 실제 값을 비교
- 4. 예측 값과 실제 값이 가까워지도록 weight, bias 조절
- 5. 모든 샘플에 대해서 처리가 될 때까지 2~4단계를 반복
- 변화율, 에러(오차)까지 고려해서 조절할 경우, 효과적인 진행 가능

경사하강법 유도

- 손실함수(loss function)

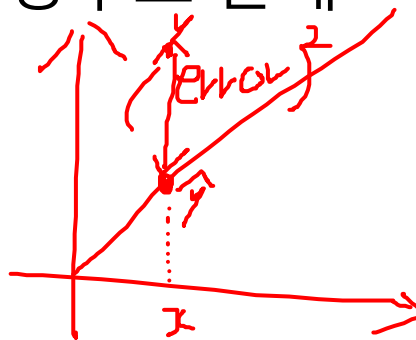
- 데이터를 토대로 산출한 **모델**의 예측 값과 실제 값의 **차이가 얼마나 나느냐**
- 모델 성능이 얼마나 나쁜지를 표현
- 여러 종류가 있으며, 직접 정의하는 경우도 존재

- 제곱 오차(Squared error)

- $$SE = (y - \hat{y})^2$$

- 선형회귀의 손실함수로 사용

- 오차가 클 수록 이 크기를 증가(벌칙) 시키기 위해 제곱을 사용



경사하강법 유도

- $SE = (y - \hat{y})^2$

- $\frac{dSE}{d\omega} = \frac{d}{d\omega} (y - \hat{y})^2$

$$y = f(g(x))$$
$$y' = F'(g(x)) \cdot g'(x)$$

$$g(x) = -\hat{y}$$

$$\hat{y} = w \cdot x + b$$

$$= 2(y - \hat{y}) \cdot \frac{d}{dw}(-\hat{y})$$

$$= 2(y - \hat{y}) \left(-\frac{d}{dw}(w \cdot x + b) \right)$$

$$= 2(y - \hat{y})(-x)$$

$$= -2(y - \hat{y})x$$

$$SE = \frac{1}{2}(y - \hat{y})^2$$

$$\rightarrow \boxed{\frac{dSE}{dw} = -(y - \hat{y})} \text{ (gradient)}$$

경사하강법 유도

- 유도한 경사하강법 식 (weight)

$$\frac{dSE}{d\omega} = -(y - \hat{y})x$$

- 이전 강의에서 **에러를 반영한 변화율 계산과 동일**

$$\begin{aligned}w_{\text{new}} &= w - \frac{dSE}{d\omega} \\&= w + (y - \hat{y})x\end{aligned}$$

경사하강법 유도

- 유도한 경사하강법 식 (bias)

$$\frac{dSE}{db} = -(y - \hat{y})$$

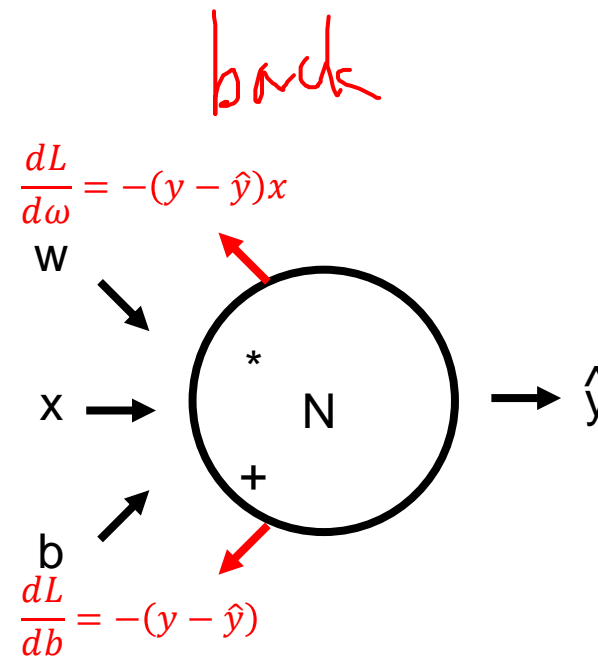
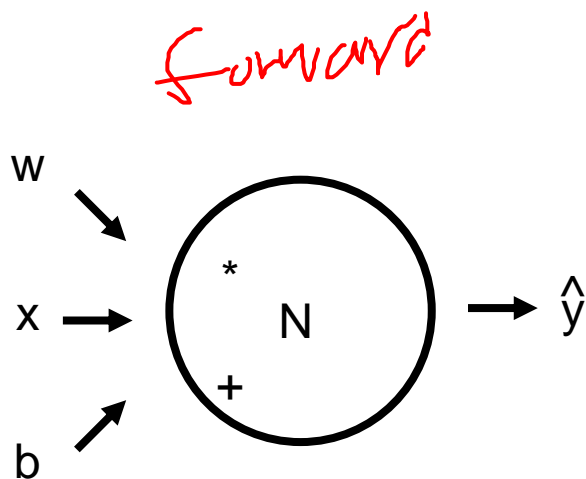
- 이전 강의에서 **에러를 반영한 변화율 계산과 동일**

$$b_{\text{new}} = b - \frac{dSE}{db} \quad \rightarrow w \cdot x + b$$
$$= b + (y - \hat{y})$$

$$\begin{aligned} \frac{dSE}{db} &= \frac{d}{db} \frac{1}{2} (y - \hat{y})^2 = (y - \hat{y}) \cdot \left(-\frac{d}{db} \hat{y} \right) \\ &= (y - \hat{y}) \cdot \left(-\frac{d}{db} (w \cdot x + b) \right) = -(y - \hat{y}) \end{aligned}$$

뉴런 작성

- 하나의 뉴런에는 ...
 - 정방향 계산 (Forward-pass)
 - 역방향 계산 (Backpropagation)
 - 훈련함수 (fit-function)



Q & A