03 머신러닝의 기초를 다집니다

- 수치 예측

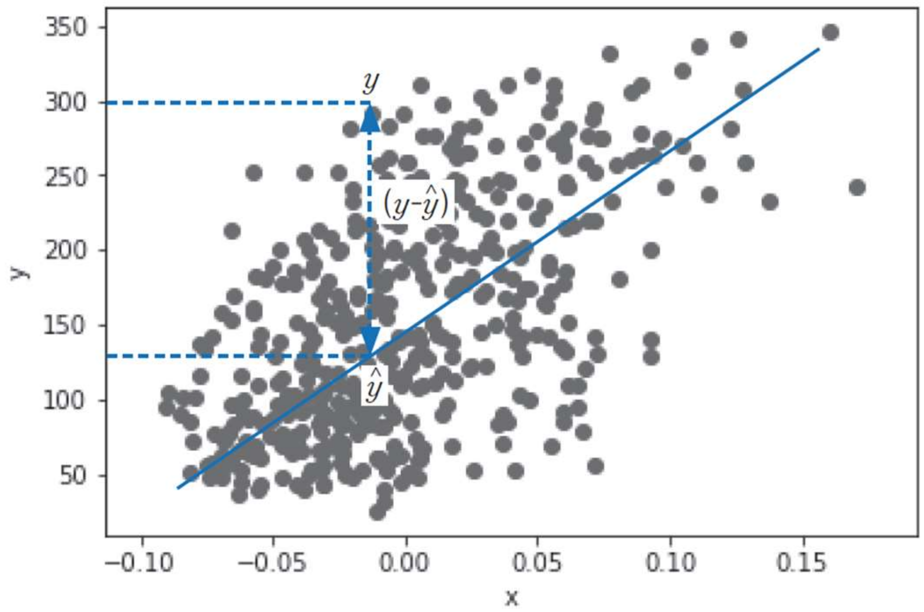
03-3 손실 함수와 경사 하강법의 관계를 알아봅니다

- 손실 함수는 예측한 값과 실제 타깃값의 차이를 측정합니다
- 손실 함수의 차이를 줄이는 방법으로 경사 하강법을 사용합니다
- 대표적인 회귀, 분류 등에는 널리 사용하는 손실 함수가 있습니다
- 복잡한 다른 문제에서는 자신만의 손실 함수를 정의하여 사용하기도 합니다

회귀의 손실 함수

제곱 오차(squared error)

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



손실 함수의 기울기를 찾기 위해 미분합니다

가중치에 대하여 제곱 오차 미분하기

$$\frac{\partial SE}{\partial w} = \frac{\partial}{\partial w}(y - \hat{y})^2 = 2(y - \hat{y})(-\frac{\partial}{\partial w}\hat{y}) = 2(y - \hat{y})(-x) = -2(y - \hat{y})x$$

$$\frac{\partial}{\partial w}(w \times x + b) = x$$

$SE = \frac{1}{2}(y - \hat{y})^2$ 라면 $\frac{\partial SE}{\partial w} = -(y - \hat{y})x$

미분 결과를 가중치에서 빼면 손실 함수의 낮은 쪽으로 이동

$$w = w - \frac{\partial SE}{\partial w} = w + (y - \hat{y})x$$

앞서 직관으로 계산한 오차 역전파가 제곱 오차를 미분한 것과 결과가 같군요!

```
y_hat = x_i * w + b
err = y_i - y_hat
w_rate = x_i
w = w + w_rate * err
```

절편에 대해 미분하고 업데이트하기

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

$$\frac{\partial}{\partial w}(w \times x + b) = x$$

그레이디언트(gradient)

$$b = b - \frac{\partial SE}{\partial b} = b + (y-\hat{y})$$

```
err = y_i - y_hat
b = b + 1 * err
```