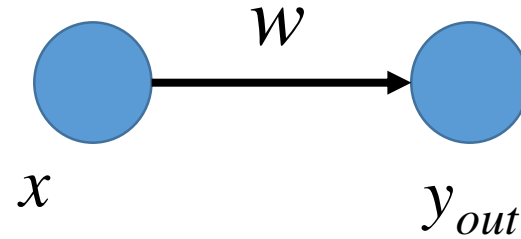# Train with 1 Data

학습데이타:　　　$x = 3, y_{true} = 6$

초기 weight:　　$w_0 = 3$

학습속도
(Learning rate):　$\alpha = 0.01$

$$Loss = (y_{true} - y_{out})^2$$

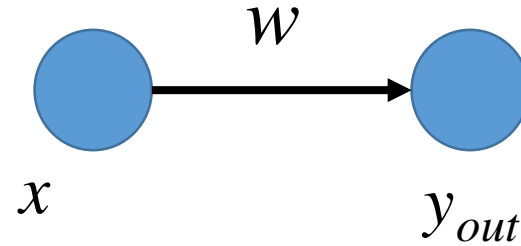
$$y_{out} = wx$$

뉴턴방법 (gradient Descent = 경사하강법)

$$w_k = w_{k-1} - \alpha \left. \frac{dLoss}{dw} \right|_{w=w_{k-1}}$$

학습데이타: $x = 3, y_{true} = 6$

초기 weight: $w_0 = 3$

학습속도
(Learning rate): $\alpha = 0.01$

$$y_{out} = wx$$

$$Loss = (y_{true} - y_{out})^2$$
$$= (y_{true} - wx)^2$$

$$w_1 = w_0 - \alpha \frac{dLoss}{dw}\bigg|_{w=w_0}$$

$$= 3 - 0.01 \cdot 2(6 - 3 \cdot 3)(-3) = 3 - 0.01 \cdot (-6)(-3)$$
$$= 3 - 0.18 = 2.82$$

$$\frac{dLoss}{dw} = 2(y_{true} - wx)(-x) = -2(y_{true} - wx)x$$

$$y_{out} = w_1 x = 2.82 \cdot 3 = 8.46$$

Train with 2 Data

학습데이타:
$$x_2 = 3, y_2^{true} = 6$$
$$x_1 = 4, y_1^{true} = 10$$
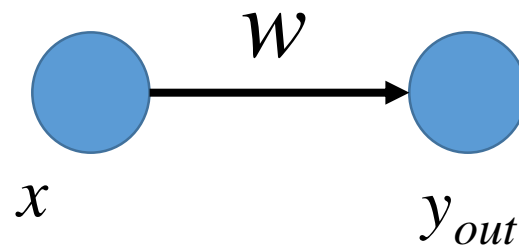
초기 weight: $w_0 = 3$

학습속도
(Learning rate): $\alpha = 0.01$

$$y_1^{out} = wx_1 = 12$$

$$y_2^{out} = wx_2 = 9$$

$$Loss = (y_1^{true} - y_1^{out})^2 + (y_2^{true} - y_2^{out})^2 = \sum_{n=1}^{2}(y_n^{true} - y_n^{out})^2$$

$$\frac{dLoss}{dw} = 2(y_1^{true} - y_1^{out})(-x_1) + 2(y_2^{true} - y_2^{out})(-x_2)$$

$$= 2(y_1^{true} - wx_1)(-x_1) + 2(y_2^{true} - wx_2)(-x_2)$$

학습데이타:
$$x_2 = 3, y_2^{true} = 6$$
$$x_1 = 4, y_1^{true} = 10$$

초기 weight: $w_0 = 3$

학습속도
(Learning rate): $\alpha = 0.01$

$$\frac{dLoss}{dw} = 2(y_1^{true} - y_1^{out})(-x_1) + 2(y_2^{true} - y_2^{out})(-x_2)$$

$$= 2(y_1^{true} - wx_1)(-x_1) + 2(y_2^{true} - wx_2)(-x_2)$$

$$= 2(10 - 3 \cdot 4)(-4) + 2(6 - 3 \cdot 3)(-3)$$

$$= 16 + 18 = 34$$

초기 weight: $w_0 = 3$

학습속도
(Learning rate): $\alpha = 0.01$

$$w_1 = w_0 - \alpha \left. \frac{dLoss}{dw} \right|_{w=w_0}$$

$$= 3 - 0.01 \cdot 34 = 3 - 0.34 = 2.66$$

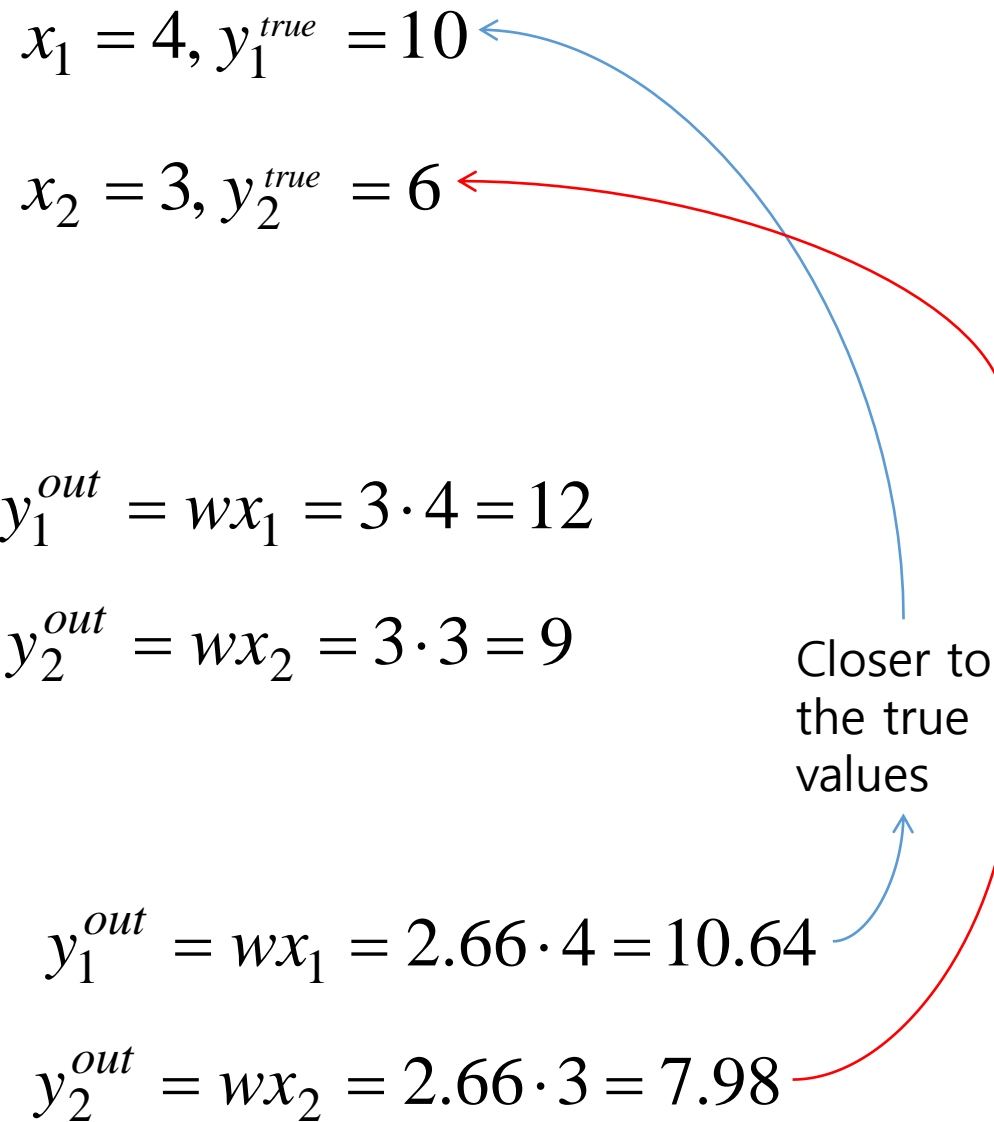$$x_1 = 4, \ y_1^{true} = 10$$

$$x_2 = 3, \ y_2^{true} = 6$$

이전 출력: $y_1^{out} = wx_1 = 3 \cdot 4 = 12$

$$y_2^{out} = wx_2 = 3 \cdot 3 = 9$$
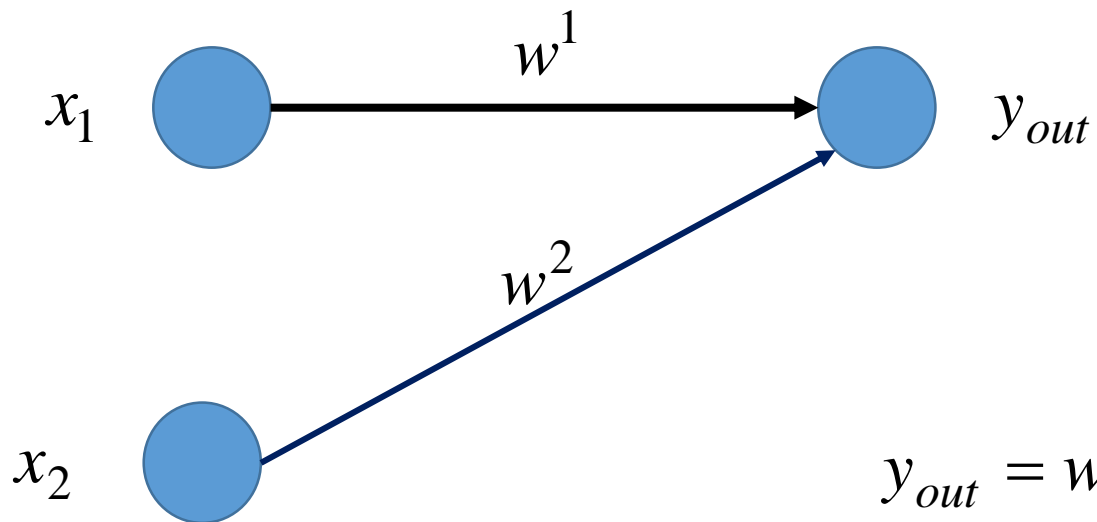
Closer to the true values

지금 출력: $y_1^{out} = wx_1 = 2.66 \cdot 4 = 10.64$

$$y_2^{out} = wx_2 = 2.66 \cdot 3 = 7.98$$

Train with 1 Data

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

x_1 w^1

y_{out}

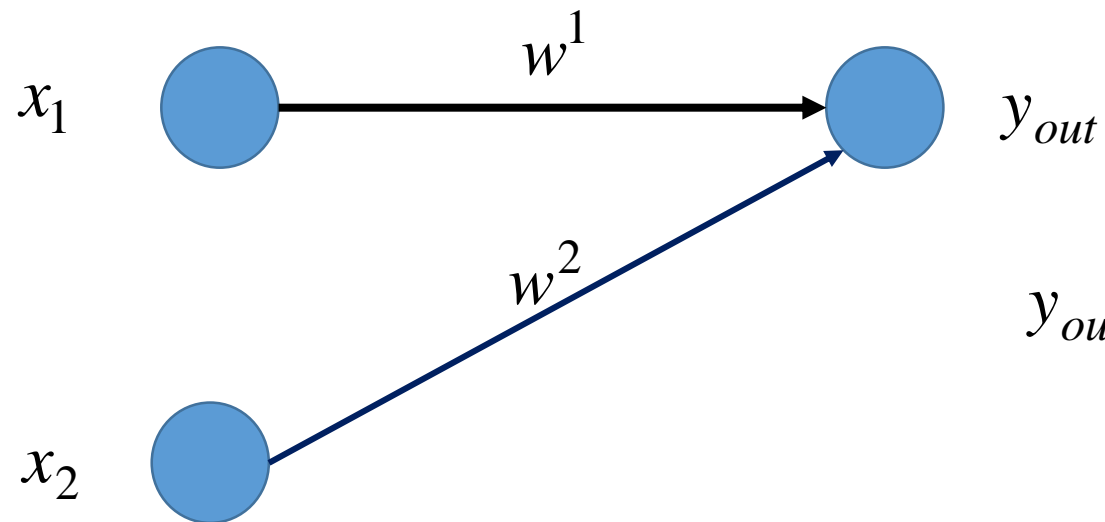x_2 w^2

$$y_{out} = w^1 x_1 + w^2 x_2$$

학습데이타: $\quad \mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, y_{true} = 6$

$$Loss = (y_{true} - y_{out})^2$$

초기 weight: $\quad w^1 = 1, w^2 = 1$

학습속도
(Learning rate): $\quad \alpha = 0.01$

$$Loss = (y_{true} - y_{out})^2$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

x_1 w^1 y_{out}

x_2 w^2

$$y_{out} = w^1 x_1 + w^2 x_2$$

학습데이타: $\mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, y_{true} = 6$

$$w^1_{k+1} = w^1_k - \alpha \left.\frac{dLoss}{dw^1}\right|_{w^1 = w^1_k}$$

초기 weight: $w^1 = 1, w^2 = 1$

$$w^2_{k+1} = w^2_k - \alpha \left.\frac{dLoss}{dw^2}\right|_{w^2 = w^2_k}$$

학습속도
(Learning rate): $\alpha = 0.01$

$$w^1_{k+1} = w^1_k - \alpha \frac{dLoss}{dw^1}\bigg|_{w^1 = w^1_k}$$

$$y_{out} = w^1 x_1 + w^2 x_2$$

첫번째 업데이트

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, y_{true} = 6$$

$$w^1_1 = w^1_0 - \alpha \frac{\partial Loss}{\partial w^1}\bigg|_{w^1 = w^1_k} = w^1_0 - \alpha \frac{\partial Loss}{\partial M}\frac{\partial M}{\partial w^1}\bigg|_{w^1 = w^1_k}$$

$$= w^1_0 - \alpha 2(y_{true} - y_{output})\frac{\partial(y_{true} - y_{output})}{\partial w^1}\bigg|_{w^1 = w^1_k}$$

초기 weight: $w^1 = 1, w^2 = 1$

$$= w^1_0 - \alpha 2(y_{true} - y_{output})(-x_1)\big|_{w^1 = w^1_k}$$

학습속도
(Learning rate): $\alpha = 0.01$

$$= 1 - 0.01 \cdot 2(6 - 1 \cdot 2 - 1 \cdot 3)(-2) = 1 - 0.01 \cdot (-4) = 1.04$$

$$w_{k+1}^2 = w_k^2 - \alpha \left. \frac{\partial Loss}{\partial w^2} \right|_{w^2 = w_k^2}$$

$$y_{out} = w^1 x_1 + w^2 x_2$$

첫번째 업데이트

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, y_{true} = 6$$

$$w_1^2 = w_0^2 - \alpha \left. \frac{\partial Loss}{\partial w^2} \right|_{w^2 = w_k^2} = w_0^2 - \alpha \left. \frac{\partial Loss}{\partial M} \frac{\partial M}{\partial w^2} \right|_{w^2 = w_k^2}$$

$$= w_0^2 - \alpha 2(y_{true} - y_{output}) \left. \frac{\partial (y_{true} - y_{output})}{\partial w^2} \right|_{w^2 = w_k^2}$$

초기 weight: $w^1 = 1, w^2 = 1$

학습속도
(Learning rate): $\alpha = 0.01$

$$= w_0^2 - \alpha 2(y_{true} - y_{output})(-x_2) \Big|_{w^2 = w_k^2}$$

$$= 1 - 0.01 \cdot 2(6 - 1 \cdot 2 - 1 \cdot 3)(-3) = 1 - 0.01 \cdot (-6) = 1.06$$

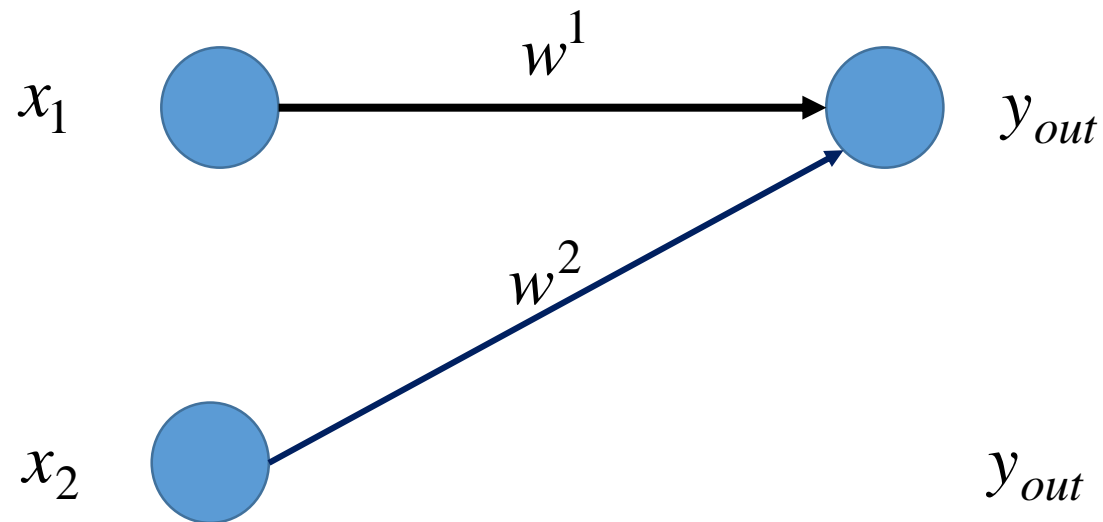이전 출력: $y_{out} = w^1 x_1 + w^2 x_2 = 1 \cdot 2 + 1 \cdot 3 = 5$

지금 출력: $y_{out} = w^1 x_1 + w^2 x_2 = 1.04 \cdot 2 + 1.06 \cdot 3 = 5.26$

참값에 더 가까와졌다.

$y_{true} = 6$

Train with 2 Data

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$y_{out} = w^1 x_1 + w^2 x_2$$

$$Loss = (y_{true} - y_{out})^2$$

학습데이타: $\mathbf{x}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, y_1^{true} = 6 \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, y_2^{true} = 4$

초기 weight: $w^1 = 1, w^2 = 1$

학습속도
(Learning rate): $\alpha = 0.01$