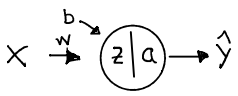

Propagación hacia adelante:

Para una muestra:

$z_i = x_i w + b$

$a_i = \sigma(z) = (1) z_i$ ← Lineal

$\hat{y}_i = a_i$ — Última capa

Para un lote:

$\underset{m\times 1}{Z} = \underset{m\times 1}{X}\ \underset{1\times 1}{W} + \underset{1\times 1}{B}$

$\underset{m\times 1}{A} = \sigma(Z) = IZ$ ← Lineal

$\hat{Y} = A$ — Última capa

Propagación hacia atrás:

Para una muestra:

MSE: $L_i = (y_i - \hat{y}_i)^2$

MAE: $L_i = |y_i - \hat{y}_i|$

$J = L_i$

Para un lote:

MSE: $L = (y - \hat{y})^2$

MAE: $L = |y - \hat{y}|$

$J = \frac{1}{m}\sum_{i=1}^{m} L_i$

Ejemplo:

	Year Center Norm (X)	Log N Transistors (Y)
0	0.2680	21.2539
1	0.9363	21.3651
2	1.0849	22.4280
3	-0.2518	16.4182

$\alpha = 0.1$

Epoca 1: $w = -0.25$, $b = 0$

$\hat{y} = 0.2680 \times -0.25 + 0 = -0.0670$

$J = L = [21.2539 - (-0.0670)]^2 = 454.58$ ← Costo

$\frac{\partial J}{\partial b} = -2[21.2539 - (-0.0670)] = -42.64$

$\frac{\partial J}{\partial w} = -2[21.2539 - (-0.0670)][0.2680] = -11.4280$

Actualizar θ:

$b := 0 - (0.1)(-42.64) = 4.264$ ↑ Incrementa

$w := -0.25 - (0.1)(-11.42) = 1.392$ ↑ Incrementa

Epoca 2: $w = 1.392$, $b = 4.264$

$\hat{y} = 0.9363 \times 1.392 + 4.264 = 5.5673$

$J = L = [21.3651 - (5.5673)]^2 = 249.57$ ← Costo

$\frac{\partial J}{\partial b} = -2[21.3651 - (5.5673)] = -31.5956$

$\frac{\partial J}{\partial w} = -2[21.3651 - (5.5673)][0.9363] = -29.583$

Actualizar θ:

$b := 4.264 - (0.1)\cdot(-31.5956) = 7.4236$ ↑ Incrementa

Gradiente:

$\theta = [b, w]$

$\theta := \theta - \alpha \nabla_\theta J(\theta)$

$\nabla_\theta = \left[\frac{\partial}{\partial b}, \frac{\partial}{\partial w}\right]$

$\nabla_\theta J(\theta) = \left[\frac{\partial J}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial b}, \frac{\partial J}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial w}\right]$ ← Para un lote

$\nabla_\theta J(\theta) = \left[\sum_{i=1}^{m} -2(y_i - \hat{y}_i), \sum_{i=1}^{m} -2(y_i - \hat{y}_i)x_i\right]$

$\nabla_\theta J(\theta) = \left[-2(y_i - \hat{y}_i), -2(y_i - \hat{y}_i)x_i\right]$

Actualizar θ:

$b := 4.264 - (0.1) \cdot (-31.5956) = 7.4236$ ↑ Incrementa

$w := 1.392 - (0.1) \cdot (-29.583) = 3.45$ ↑ Incrementa

Epoca 3: $\{w = 3.45,\ b = 7.4236\}$

$\hat{Y} = 1.0899 \times 3.45 + 7.4236 = 11.1665$

$J = L = [22.4280 - (11.1665)]^2 = 126.8213$

↖ Costo

$\frac{\partial J}{\partial b} = -2[22.4280 - (11.1665)] = -22.523$

$\frac{\partial J}{\partial w} = -2[22.4280 - (11.1665)][1.0899] = -24.93$

Actualizar θ:

$b := 7.4236 - (0.1) \cdot (-22.523) = 9.6759$ ↑ Incrementa

$w := 3.45 - (0.1) \cdot (-24.93) = 5.893$ ↑ Incrementa