

## MSE & Cross Entropy

### Mean Squared Error

De MSE verteld hoe dicht bij een regressielijn is tot de een aantal waardes. Des te kleiner de MSE, des te beter de fit is. Eerst worden de afstanden van de punten tot de regressielijn bij elkaar opgeteld. De afstanden zijn de errors. Het wordt gedeeld door de aantal punten en uiteindelijk gekwadeerd. Door te kwadrateren zorg je ervoor dat het getal altijd wel positief is en dat grotere waardes sterker afgestraft worden dan veel kleinere waardes.

$$\text{MSE} = (1/n) * \sum (\text{actual} - \text{forecast})^2$$

### Cross Entropy

De cross entropy is ook een loss function net als de MSE, maar deze wordt gebruikt voor classificatie problemen. Ook hier is het doel om de CE-waarde te minimaliseren

## Opdrachten

### Mean Squared Error

$$J(\theta_0, \theta_1) = \frac{1}{n} \sum_{i=1}^n (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

Hypothese:

Stel  $\theta_0 = 0$ , dan geldt  $\rightarrow h_{\theta}(x^{(i)}) = \theta_1 x^{(i)}$

- resultaat voor  $J(\theta_1)$  als  $\theta_1 = 1.5$

$$\begin{aligned} J(0; 1.5) &= \frac{1}{3} \sum_{i=1}^3 (1.5x^{(i)} - y^{(i)})^2 \\ &= \frac{1}{3} ((1.5 \cdot 1 - 1)^2 + (1.5 \cdot 2 - 2)^2 + (1.5 \cdot 3 - 3)^2) \\ &= \frac{1}{3} (0.5^2 + 1^2 + 1.5^2) \\ &= \frac{1}{3} (0.25 + 1 + 2.25) \\ &= 1.167 \end{aligned}$$

- resultaat voor  $J(\theta_1)$  als  $\theta_1 = 1.75$

$$\begin{aligned} J(0; 1.75) &= \frac{1}{3} \sum_{i=1}^3 (1.75x^{(i)} - y^{(i)})^2 \\ &= \frac{1}{3} ((1.75 - 1)^2 + (3.5 - 2)^2 + (5.25 - 3)^2) \\ &= \frac{1}{3} (0.5625 + 2.25 + 5.0625) \\ &= 2.625 \end{aligned}$$



softmax

$$x_{1,2,3} = \begin{pmatrix} 1,4 \\ 2,5 \\ 0,7 \end{pmatrix}$$

$$\delta(x)_1 = \frac{e^{1,4}}{(e^{1,4} + e^{2,5} + e^{0,7})} = 0,22$$

$$\delta(x)_2 = \frac{e^{2,5}}{(e^{1,4} + e^{2,5} + e^{0,7})} = 0,67$$

$$\delta(x)_3 = \frac{e^{0,7}}{(e^{1,4} + e^{2,5} + e^{0,7})} = 0,11$$

$$\delta(x)_i = \begin{pmatrix} 0,22 \\ 0,67 \\ 0,11 \end{pmatrix}$$