

## Gradiënt Descent en Partieel differentiëren

### Wat is gradient decent?

Gradient descent is een optimalisatie algoritme om de minimum van een functie te vinden. Het wordt vaak gebruikt bij deep learning modellen om gewichten en biases te veranderen en zo de model beter te trainen.

### Opdrachten

#### Partieel differentiëren

**Les 3**

1  $Z(x,y) = x^4 + x^3y^2 + y^5 + 8$

$$\frac{\partial Z}{\partial x} = 4x^3 + 3x^2y^2$$
$$\frac{\partial Z}{\partial y} = 2x^3y + 5y^4$$
$$\frac{\partial^2 Z}{\partial x \partial y} = 6x^2y \leftarrow \text{eerst } x \text{ dan } y$$

2  $Z(x,y) = x^3 + x^2y^4 - 2y^5$

$$\frac{\partial Z}{\partial x} = 3x^2 + 2xy^4$$
$$\frac{\partial Z}{\partial y} = 4x^2y^3 - 10y^4$$
$$\frac{\partial^2 Z}{\partial x \partial y} = 8xy^3$$

3  $Z(x,y) = -3x^3y^2 + 5$

$$\frac{\partial Z}{\partial x} = -9x^2y^2$$
$$\frac{\partial Z}{\partial y} = -6x^3y$$
$$\frac{\partial^2 Z}{\partial x \partial y} = -18x^2y$$

## Partieel afgeleide lineaire kostenfunctie

partieel afgeleide

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

$$\text{met } h_{\theta}(x^{(i)}) = \theta_0 + \theta_1 x^{(i)}$$

$\theta_0$

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

$\theta_1$

$$\begin{aligned} \frac{\partial}{\partial \theta_1} (J(\theta_0, \theta_1)) &= \frac{1}{n} \sum_{i=1}^n 2(h_{\theta}(x^{(i)}) - y^{(i)}) \cdot x^{(i)} \\ &= \frac{2}{n} \sum_{i=1}^n (h_{\theta}(x^{(i)}) - y^{(i)}) \cdot x^{(i)} \end{aligned}$$



## Gradiënt Descent

### Gradiënt Descent

kostenfunctie:  $3\theta^2$

1 random waarde  
 $\theta = 4$

2 afgeleide  
 $6\theta$

3 Helling  
 $6 \cdot 4 = 24$

4 Learning rate  
 $lr = 0.1$

5 Update  
 $\theta = 4 - (0.1 \cdot 24) = 1.6$

6 Repeat

- $\theta = 1.6$
- Helling =  $6 \cdot 1.6 = 9.6$
- update:  $\theta = 1.6 - (0.1 \cdot 9.6) = 0.64$



Kostenfunktion:  $\frac{1}{4}\theta_1^4 + \frac{1}{5}\theta_2^2$

1 Random waarden

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \end{pmatrix}$$

2 Afgeleide

$$\theta_1^3 + \frac{2}{5}\theta_2$$

3  $Lr = 0.1$

4 update

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} -4 - (1 \cdot -4^3 \cdot 0.1) \\ 2 - (\frac{2}{5} \cdot 2 \cdot 0.1) \end{pmatrix} = \begin{pmatrix} 2.4 \\ 1.92 \end{pmatrix}$$

5 repeat

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} 2.4 - (2.4^3 \cdot 0.1) \\ 1.92 - (\frac{2}{5} \cdot 1.92 \cdot 0.1) \end{pmatrix} = \begin{pmatrix} 1.02 \\ 1.84 \end{pmatrix}$$