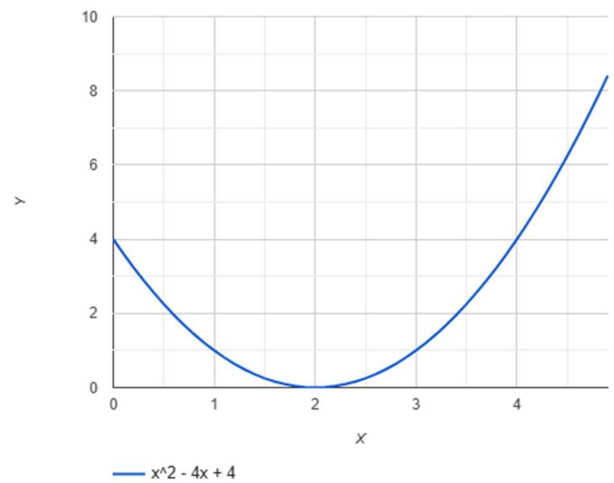
Gradient descent case study:

Task:

- find the global minimum of the non-linear quadratic function

$$f(x) = x^2 - 4x + 4 \qquad (1)$$



Gradient descent:

$$x_{n+1} = x_n - \alpha.\nabla\mathrm{f}(x_n) = x_n - \alpha\frac{\partial f(x_n)}{\partial x} \qquad (2)$$

- where **α** is a learning rate

Derivative of gradient descent:

- for the global minimum the derivative is zero:

$$\nabla\mathrm{f}(x_n) = \frac{\partial f(x_n)}{\partial x} = \frac{\partial(x_n^2 - 4x_n + 4)}{\partial x_n} = 2x_n - 4 \qquad (3)$$

- expected result = global minimum of the function:

$$\nabla\mathrm{f}(x) = \frac{\partial f(x)}{\partial x} = 0$$

$$2x - 4 = 0 \qquad (3)$$

$$\boldsymbol{x = 2}$$

Example of the iterations:

- starting values:
  - $\alpha = 0.1$
  - $x_0 = 0$
- iteration 0:

$$x_1 = x_0 - \alpha.\nabla\mathrm{f}(x_0) = x_0 - \alpha(2x_0 - 4) = 0 - 0.1(2 * 0 - 4) = 0.4 \qquad (4)$$

- iteration 1:

$$x_2 = x_1 - \alpha.\nabla \mathrm{f}(x_1) = x_1 - \alpha(2x_1 - 4) = 0.4 - 0.1(2 * 0.4 - 4) = 0.68 \tag{5}$$

- iteration 2:

$$x_3 = x_2 - \alpha.\nabla \mathrm{f}(x_2) = x_2 - \alpha(2x_2 - 4) = 0.68 - 0.1(2 * 0.68 - 4) = 0.88 \tag{6}$$

- iteration 3:

$$x_4 = x_3 - \alpha.\nabla \mathrm{f}(x_3) = x_3 - \alpha(2x_3 - 4) = 0.88 - 0.1(2 * 0.88 - 4) = 1.08 \tag{7}$$

- iteration 3:

$$x_5 = x_4 - \alpha.\nabla \mathrm{f}(x_4) = x_4 - \alpha(2x_4 - 4) = 1.08 - 0.1(2 * 1.08 - 4) = 1.28 \tag{8}$$

- from the iterations we can see that we are getting closer to the expected global minimum which is **2**.