

# Gradient descent

---

In the linear case:

$$\frac{\partial R}{\partial \beta_0} = \frac{1}{n} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) \times (-1) = 0$$

$$\frac{\partial R}{\partial \beta_1} = \frac{1}{n} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) \times (-x_i)$$

Repeat until convergence:

Update **simultaneously** all  $\beta_j$  for ( $j = 0$  and  $j = 1$ )

$$\beta_0 := \beta_0 - \alpha \frac{1}{n} \sum_{i=1}^n (\beta_0 + \beta_1 x_i - y_i)$$

$$\beta_1 := \beta_1 - \alpha \frac{1}{n} \sum_{i=1}^n (\beta_0 + \beta_1 x_i - y_i)(x_i)$$