

INITIAL PARAMETERS

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 \quad \theta_0 = 1 \quad \theta_1 = -0,5 \quad \theta_2 = 0,8$$

DATASET

$$X = \begin{bmatrix} 3 & -2 \\ 1 & 5 \\ -4 & 3 \end{bmatrix} \quad y = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \quad \lambda = 0,05$$

ADD BIAS AND CALCULATE PREDICTIONS

$$\hat{y} = x\theta$$

$$\hat{y} = \begin{bmatrix} 1 & 3 & -2 \\ 1 & 1 & 5 \\ 1 & -4 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -0,5 \\ 0,8 \end{bmatrix} = \begin{bmatrix} -2,1 \\ 4,5 \\ 5,4 \end{bmatrix}$$

RESIDUALS (VECTORIZED)

$$e = \hat{y} - y = x\theta - y$$

$$e = \begin{bmatrix} -2,1 \\ 4,5 \\ 5,4 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -3,1 \\ 2,5 \\ 5,4 \end{bmatrix}$$

LOSS

$$\frac{1}{2m} e^T e$$

$$\text{Loss} = \frac{1}{2 \cdot 3} (-3,1)^2 + (2,5)^2 + (5,4)^2 = \frac{45,02}{6} \approx 7,5$$

FULL BATCH GD

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{m} \sum_i e^{\hat{y}_i} \quad \frac{\partial J}{\partial \theta_1} = \frac{1}{m} \sum_i e^{\hat{y}_i} x_i^1$$

$$\frac{\partial J}{\partial \theta_2} = \frac{1}{m} \sum_i e^{\hat{y}_i} x_i^2$$

$$\nabla_{\theta} = \frac{1}{m} X^T e$$

$$\nabla_{\theta} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 3 & 1 & -4 \\ -2 & 5 & 3 \end{bmatrix} \begin{bmatrix} -3,1 \\ 2,5 \\ 5,4 \end{bmatrix} = \begin{bmatrix} 1,6 \\ -9,40 \\ 11,63 \end{bmatrix}$$

PARAMETERS UPDATE

$$\lambda = 0,05$$

$$\theta_{\text{NEW}} = \theta - \lambda \nabla_{\theta}$$

$$\theta_{\text{NEW}} = \begin{bmatrix} 1 \\ -0,5 \\ 0,8 \end{bmatrix} - 0,05 \begin{bmatrix} 1,6 \\ -9,40 \\ 11,63 \end{bmatrix} = \begin{bmatrix} 0,92 \\ -0,027 \\ 0,218 \end{bmatrix}$$

SGD

$$X = \begin{bmatrix} 3 & -2 \end{bmatrix} \quad Y = \begin{bmatrix} 1 \end{bmatrix}$$

$$\hat{y} = x_{\theta} = \begin{bmatrix} 1 & 3 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ -0,5 \\ 0,8 \end{bmatrix} = -2,1$$

$$\ell = x_{\theta} - y = -2,1 - 1 = [-3,1]$$

$$X^T e = \begin{bmatrix} 1 \\ -3 \\ -2 \end{bmatrix} \begin{bmatrix} -3, 1 \\ -9, 3 \\ 6, 2 \end{bmatrix} = \begin{bmatrix} -3, 1 \\ -9, 3 \\ 6, 2 \end{bmatrix}$$

$$\Theta_{NEW} = \begin{bmatrix} 1 \\ -0,15 \\ 0,8 \end{bmatrix} - 0,05 \begin{bmatrix} -3, 1 \\ -9, 3 \\ 6, 2 \end{bmatrix} = \begin{bmatrix} 1, 155 \\ -0,035 \\ 0,49 \end{bmatrix}$$