

PROBLEM 4: EXAMPLE

We will assume the dataset for the examples for the problems 1 and 2 (page 1, 2). Let

$$x^{(1)} = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} \end{bmatrix}^T = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}^T, \text{ let } y^{(1)} = 3.$$

We calculated the gradient for a single training example as,

$$\nabla E = \begin{bmatrix} -(y - \hat{y}) & -(y - \hat{y})x_1 & -(y - \hat{y})x_2 & -(y - \hat{y})x_3 \end{bmatrix}^T.$$

We can write a function as,

```
def compGrad(x1, x2, x3, w0, w1, w2, w3):
    yp = w0 + w1*x1 + w2*x2 + w3*x3
    pdw0 = -(y - y_hat)*x1
    pdw1 = -(y - y_hat)*x2
    pdw2 = -(y - y_hat)*x2
    pdw3 = -(y - y_hat)*x3
    return (pdw0, pdw1, pdw2, pdw3)
```

$$x_1 = 1$$

$$x_2 = 2$$

$$x_3 = 1$$

$$w_0 = w_2 = 1$$

$$w_1 = w_3 = 2$$

$g = \text{compGrad}(x_1, x_2, x_3, w_0, w_1, w_2, w_3)$

print("the gradient is: ", g)