Part 5 - Newton's method

Now use these results to update the θ vector using Newton's method. We have a 2D training set composed of the data matrix X and the vector y.

The matrix X is:

$$\begin{bmatrix} 0 & 3 \\ 1 & 3 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

The vector y is:

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Prepend a 1 to each $x^{(i)}$ in the training set so that we can model the intercept or bias term in θ.

- State the θ update equation for an iteration of Newton's method for this problem
- Assume a starting $\theta = [0, -1, 1]^T$ and a regularization parameter $\lambda = 0.07$. Compute and provide the values of θ after the first and second iteration of Newton's method, using a Python script.

Solution

By Newton's method:

$$\theta^{(t+1)} = \theta^{(t)} - \left(H^{-1}\nabla J(\theta^{(t)})\right)$$

where

- $\theta^{(t)}$ is the parameter vector at iteration t.
- H is the Hessian matrix of the cost function $J(\theta)$.
- $\nabla J(\theta)$ is the gradient vector of the cost function.

Gradient vector:

$$\nabla J(\theta) = \frac{1}{m}X^T(h_\theta(X) - y) + \frac{\lambda}{m}\theta$$

Hessian matrix:

$$H = \frac{1}{m}X^T SX + \frac{\lambda}{m}I$$

Final theta: $[-5.84223232e + 006.36552915e - 172.92111616e + 00]$

See **newtons_method.py**