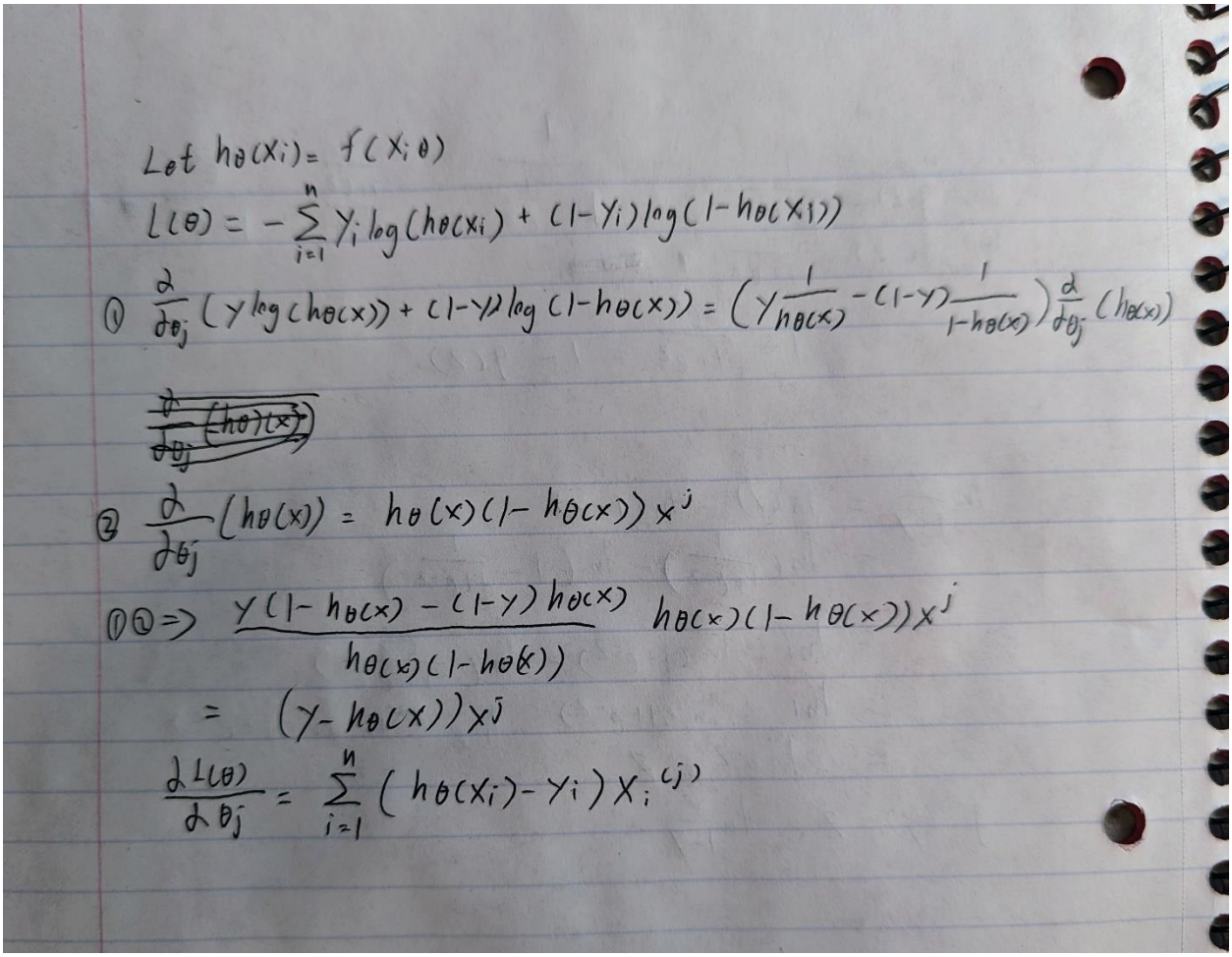
Q4

Derivation



Let $h_\theta(x_i) = f(x_i;\theta)$

$$L(\theta) = -\sum_{i=1}^{n} y_i \log(h_\theta(x_i)) + (1-y_i)\log(1-h_\theta(x_i))$$

① $$\frac{\partial}{\partial \theta_j}\left(y\log(h_\theta(x)) + (1-y)\log(1-h_\theta(x))\right) = \left(y\frac{1}{h_\theta(x)} - (1-y)\frac{1}{1-h_\theta(x)}\right)\frac{\partial}{\partial \theta_j}(h_\theta(x))$$

~~$\frac{\partial}{\partial \theta_j}(h_\theta(x))$~~

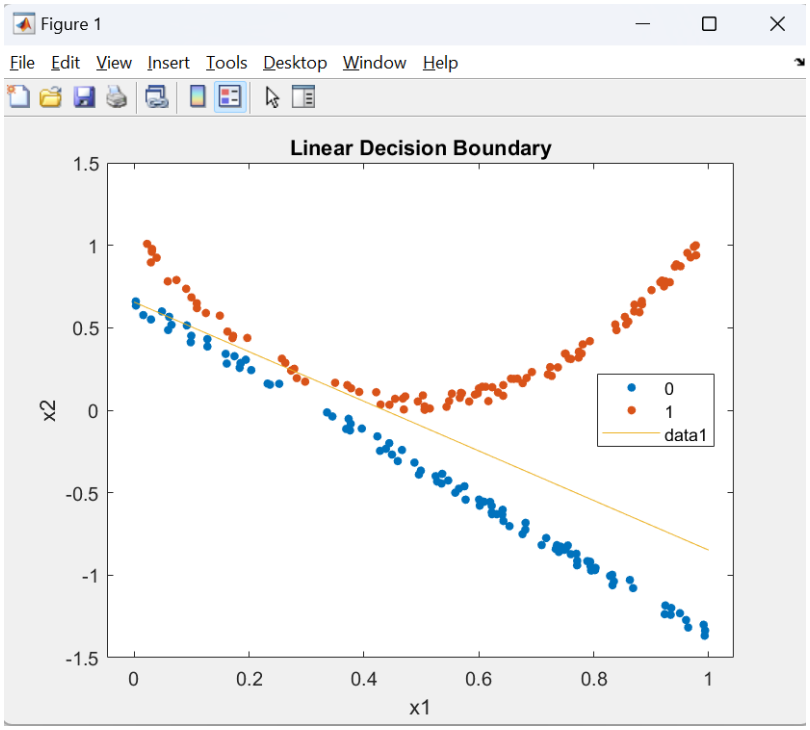
② $$\frac{\partial}{\partial \theta_j}(h_\theta(x)) = h_\theta(x)(1-h_\theta(x))x^j$$

①② $\Rightarrow$ $$\frac{y(1-h_\theta(x)) - (1-y)h_\theta(x)}{h_\theta(x)(1-h_\theta(x))} h_\theta(x)(1-h_\theta(x))x^j$$

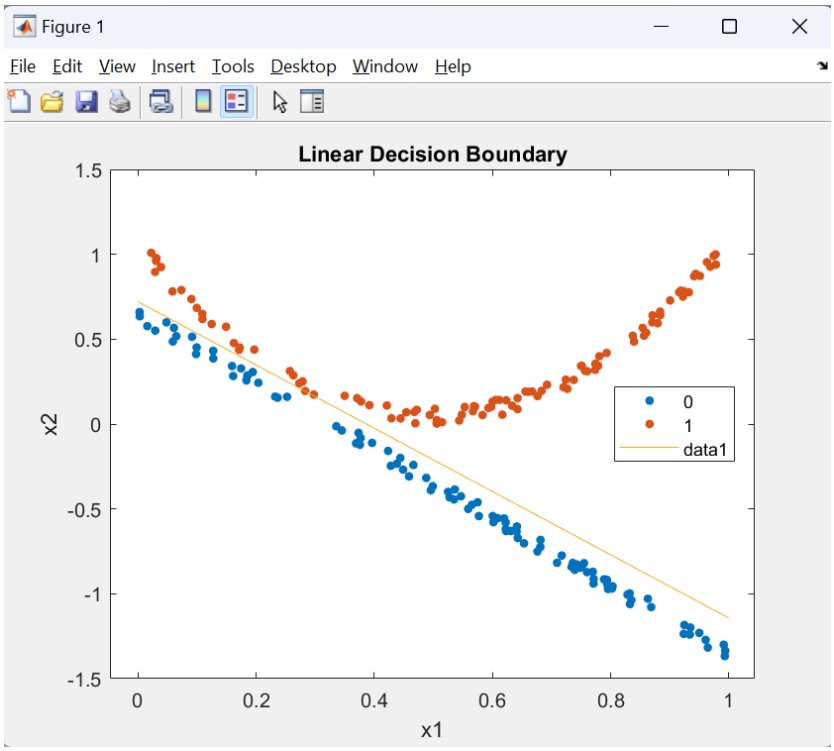
$$= (y - h_\theta(x))x^j$$

$$\frac{\partial L(\theta)}{\partial \theta_j} = \sum_{i=1}^{n}(h_\theta(x_i) - y_i)x_i^{(j)}$$

e = 0.01 and η = 0.01, total 796 iterations, errors 6, empirical risk 1.2565



e = 0.01 and η = 0.1, total 2231 iterations, errors 0, empirical risk 2.3638



e = 0.001 and η = 0.01, total 22480 iterations, errors 0, empirical risk 2.3637