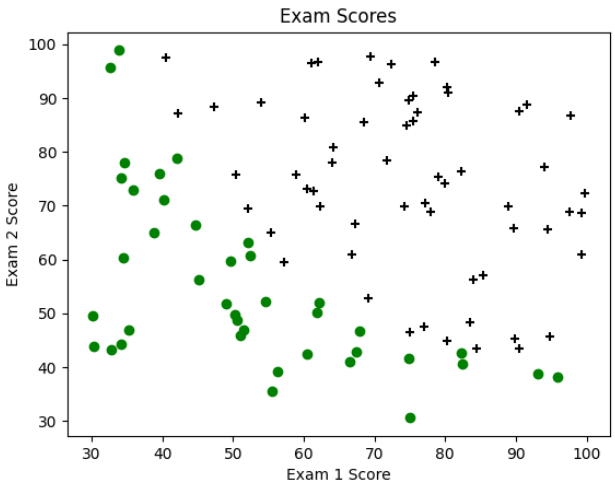
Keshav Shankar
ECE 1395
Homework 3

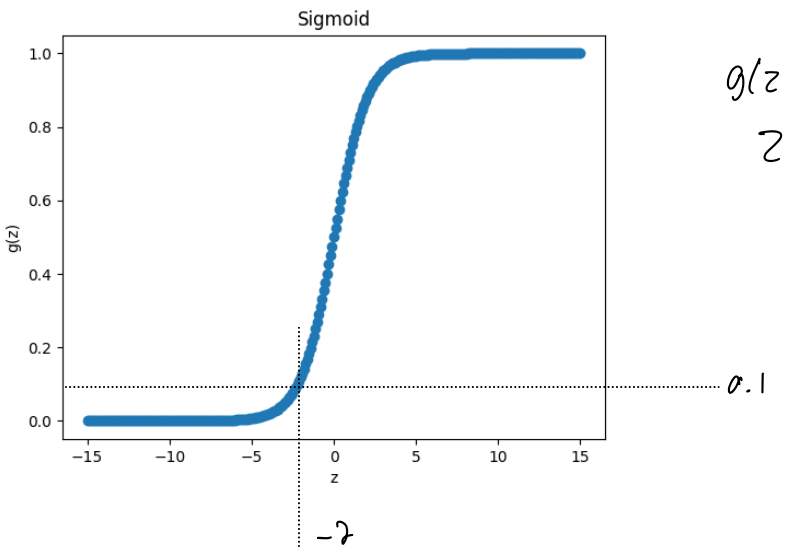
1a)

```
X has 100 data points and 3 features
Y has 100 data points and 1 labels
```

1b)



1d)



g(z) reaches 0.1 at z = -2

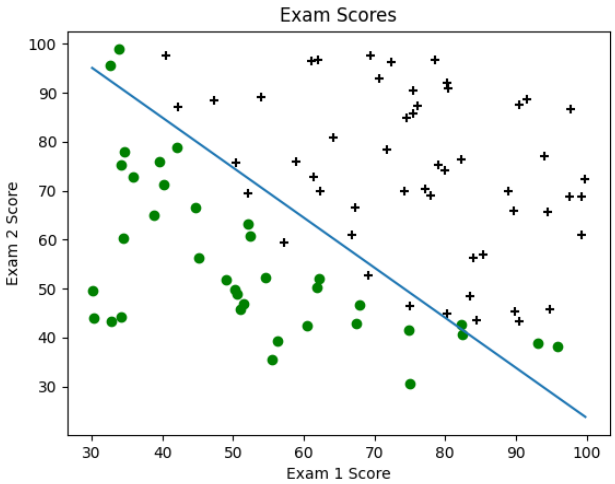
1e)

```
Cost is: [1.12692801]
Gradient of cost is:
[0.38079708 0.76159416 0.01159416]
```

1f)

```
Optimal theta is:
[-25.14997327   0.20429777   0.19992576]
Cost with optimal theta is: 0.20379428943086322
```

1g)



1i)

```
The accuracy of the prediction is: 0.6%
The admission probability is: 0.53%
The admission decision is: admitted
```

The decision should be admitted.

1j)

* BONUS *

Sigmoid : $h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$ (hypothesis function)

$\sigma(x) = \frac{1}{1+e^{-x}}$

$\frac{d(\sigma(x))}{dx} = \frac{(0)(1+e^{-x}) - (1)(e^{-x}(-1))}{(1+e^{-x})^2} = \frac{e^{-x}}{(1+e^{-x})^2}$

$= \frac{1}{1+e^{-x}}\left(1 - \frac{1}{1+e^{-x}}\right) = \sigma(x)(1-\sigma(x))$ (deriv.)

$$\frac{\partial(J(\theta))}{\partial(\theta_j)} = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)} \left(\frac{1}{h_\theta(x^{(i)})} \right) \left(\frac{\partial(h_\theta(x^{(i)}))}{\partial(\theta_j)} \right) \right]$$

$$+ \sum_{i=1}^{m}\left[(1-y^{(i)}) \left(\frac{1}{1-h_\theta(x^{(i)})} \right) \left(\frac{\partial(1-h_\theta(x^{(i)}))}{\partial(\theta_j)} \right) \right] \quad \text{(Chain rule)}$$

* Used online evaluvater to simplify this *

$$= -\frac{1}{m}\left[\sum_{i=1}^{m}\left[\left(y^{(i)} \frac{1}{h_\theta(x^{(i)})} h_\theta(x^{(i)}) \right) \left(1-h_\theta(x^{(i)}) \right) \left(x_j^i \right) \right] \right.$$

$$\left. + \sum_{i=1}^{m}\left[(1-y^{(i)}) \left(\frac{1}{1-h_\theta(x^{(i)})} \right) \left(-h_\theta(x^{(i)}) \right) \left(1-h_\theta(x^{(i)}) \right) \left(x_j^i \right) \right] \right]$$

$$= -\frac{1}{m}\left[\sum_{i=1}^{m}\left[y^{(i)} \left(1-h_\theta(x^{(i)}) \right) \left(x_j^i \right) - (1-y^{(i)}) \left(h_\theta(x^{(i)}) \right) \right] \left(x_j^i \right) \right]$$

$$-\frac{1}{m}\left[\sum_{i=1}^{m}\left[y^{(i)} - y^{(i)}\left(h_\theta(x^{(i)}) \right) - h_\theta(x^{(i)}) + y^{(i)}\left(h_\theta(x^{(i)}) \right) \right] \left(x_j^i \right) \right]$$

$$-\frac{1}{m}\left[\sum_{i=1}^{m}\left[y^{(i)} - h_\theta(x^{(i)}) \right] x_j^i \right]$$

This in matrix form now if you remove the summations

$$\therefore = \boxed{\frac{1}{m}\left(h_\theta(x) - y \right) x^T}$$

2a)

```
Learned model parameters:
[[ 2.19256369e+05]
 [-7.75885454e+02]
 [ 1.06170504e+01]]
```

2b)