Q1 eye(5)

Q2 xlabel...
ylabel

Q3. $\theta_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\theta_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

Gradient Desent

The size of theta is 2x1

$$h_\theta(x) = \sum_{j=0}^{n} \theta_j x_j$$

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix} \qquad x = \begin{bmatrix} 1 & x_0 \\ 1 & \vdots \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$$

2x1 $\qquad$ (n+1)x2 .

To calculate $\theta_j$, we know

$$\theta_j = \theta_j - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) x_j^{(i)},$$

In order to apply vectorization, as $\theta_j$ is a 2x1 matrix, $\alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) x_j^{(i)}$ has to be a 2X1 matrix as well.

In other words $\sum_{i=1}^{m}(h_\theta(x^{(i)})-y^{(i)})x_j^{(i)}$, has to be a $2 \times 1$ matrix.

As $x_j^{(i)}$ is a $(n+1) \times 2$ matrix, $(x_j^{(i)})^T$ will be $2 \times (n+1)$

Therefore, we need to mk sure $h_\theta(x^{(i)}-y^{(i)})$ to be a $(n+1) \times 1$ matrix.

$y^{(i)}$: $(n+1) \times 1$

$$\Rightarrow h_\theta x^{(i)} = \underbrace{\underset{(n+1)\times 2}{X} \cdot \underset{2 \times 1}{\theta}}_{(n+1)\times 1.}$$

Thus, put together:

$$X' * ((X * \theta) - y)$$

$$\Rightarrow \theta_j = \theta_j - \alpha \cdot \frac{1}{m} \cdot X' * ((X * \theta) - y)$$

Q4 ( Cost function)

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2.$$

Our goal is to get a 1x1 matrix:

$y^{(i)}$ : (n+1) x 1

$\theta$ : 2x1

$X$ : (n+1) x2

$(X\theta)^T$

1 x(n+1) * (n+1)x1

$$J(\theta) = \frac{1}{2m} * (X * \theta)' * (X * \theta)$$

# Optional:

- Feature Normalization:

  $\mu = mean(X);$

  As it's scaling based on standard deviation.

  $\therefore \sigma = std(X)$

  $$X\_Norm = \frac{(X-\mu)}{\sigma} \quad ((X-\mu)./\sigma)$$

- Gradient Descent : & Cost func :

  Check previous.

- learning rate.

  $\alpha = 0.01$

- Normal eq.

  $$\theta = (X^T X)^{-1} X^T y$$

  $$\theta = pinv(X' * X) * X' * y$$