Ans 1- Let a neural network contain a 1 hidden layer. Let $w_1$ be weights connecting input layer and 1st hidden layer and $w_2$ be connections b/w hidden layer and output layer. Let X be the input.

Output of this neural net =

$$= \tanh(X w_1) w_2$$

$$\text{Error} = \Sigma \left(Y - \tanh[X w_1) w_2\right)^2$$

If $w_1$ is zero and since tanh is zero at 0, the output at hidden layer will be zero. This will lead to zero vector at output layer. This will lead to

This will cause the evaluation of gradients equal to zero. Since gradient is zero the backpropagation will not update our weights and the model will not learn at all.

Therefore, origin in weight space is a stationary point.

2) (a) $E = \frac{1}{2} d_1 w_1^2 + \frac{1}{2} d_2 w_2^2$

$\frac{\partial E}{\partial w_1} = d_1 w_1$ ; $\frac{\partial E}{\partial w_2} = d_2 w_2$

$\frac{\partial E}{\partial w_1^2} = d_1$

$\frac{\partial E}{\partial w_2^2} = d^2$

$\frac{\partial E}{\partial w_1 \partial w_2} = 0$ ; $\frac{\partial E}{\partial w_2 \partial w_1} = 0$

$$H = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$$

$$\det (H - dI) = 0$$

$$\Rightarrow \begin{vmatrix} d_1 - d & 0 \\ 0 & d_2 - d \end{vmatrix} = 0$$

$$(d_1 - d)(d_2 - d) = 0$$

$d_1$ and $d_2$ are roots and hence eigen value of hessian matrix.

(b) We verify from our implementation that any learning less than 0.1 leads to convergence. The value of weights diverge for learning rate of 0.1.
We conclude the algorithm will converge if learning rate < 0.1.
0.1 is ratio of $d_1$ and $d_2$.