3. $x_1 = [1, 1]$ $y_1 = 1$

$x_2 = [1, -1]$ $y_2 = -1$

a) what does the mean surface error look like?

The error function is the mean squared error which can be written as:

$$E = \frac{1}{2}\left[y - \sum_{i=1}^{n}(w_i^T x_i)\right]^2$$

n is the total co-ordinates of pt ⇒ error for a single point

$$E = \frac{1}{2}\left[y - w_1 x_1 - w_2 x_2\right]^2$$

Differentating wrt w_1, w_2, w_1^2, w_2^2, $w_1 w_2$ & $w_2 w_1$ for the entries in Hessian matrix, we get

$$\frac{\partial E}{\partial w_1} = [y - w_1 x_1 - w_2 x_2] \cdot (-x_1)$$

$$\frac{\partial E}{\partial w_1^2} = \frac{\partial}{\partial w_1}\left[-y x_1 + w_1 x_1^2 + w_2 x_1 x_2\right] = x_1^2$$

$$\frac{\partial E}{\partial w_1 \partial w_2} = x_1 x_2$$

$$\frac{\partial E}{\partial w_2} = [y - w_1 x_1 - w_2 x_2](-x_2)$$

$$\frac{\partial E}{\partial w_2^2} = \frac{\partial}{\partial w_2}\left(-y x_2 + w_1 x_1 x_2 + w_2 x_2^2\right) = x_2^2$$

$$\frac{\partial E}{\partial w_2 \partial w_1} = x_1 x_2$$

Hessian matrix for 2 variables $= \begin{bmatrix} \frac{\partial E}{\partial w_1^2} & \frac{\partial E}{\partial w_1 \partial w_2} \\ \frac{\partial E}{\partial w_2 \partial w_1} & \frac{\partial E}{\partial w_2^2} \end{bmatrix}$

$$= \begin{bmatrix} x_1^2 & x_1 x_2 \\ x_1 x_2 & x_2^2 \end{bmatrix}$$

Here each point is represented as (x_1, x_2)

Since we have 2 points, we get Hessian for the 2 points as

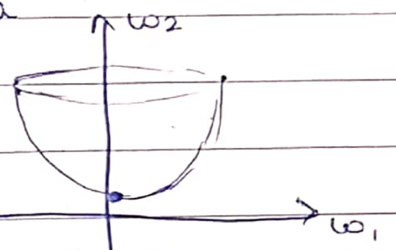
$\begin{array}{cc} x_1 & x_2 \end{array}$
$(1, 1)$ $\Downarrow$ $\quad\quad\quad\quad$ $(1, -1)$

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \qquad H_2 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

The total combined error is: $H = H_1 + H_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

$\boxed{\text{This } \cancel{\text{for}} \text{ error fn has curvature of convex}}$

A Hessian matrix with eigen values as positive (in this case 2, 2) gives the curvature as $\boxed{\text{concave - upwards}}$ [i.e. it has a ~~local~~ global minimum]. It takes the shape of a "convex function".



Note: In other case ~~wh~~ if the eigen values are negative, we get a 'concave - down' which gives ~~local~~ global maximum. If either one is 0, we get a saddle point

The minimum point is obtained by the eqn

$$\frac{\partial E}{\partial w_1} = 0 \implies -x_1 y + w_1 x_1^2 + w_2 x_1 x_2 = 0$$

For the 2 points we get $y = 1, -1$
$x_1 = 1 \quad x_2 = 1, -1$

(or) $\frac{\partial E}{\partial w_2} = 0$

we get

$$-1 + w_1 + w_2 = 0 \quad \&$$
$$+1 + w_1 - w_2 = 0$$

$$w_1 - w_2 = -1$$
$$w_1 + w_2 = 1$$

$\implies$ $\boxed{\begin{array}{l} w_1 = 0 \quad w_2 = 1 \\ \text{minimum is obtained at } (0, 1) \end{array}}$