

$$J(\underline{w}) = ||X\underline{w} - t||^2$$

$$= \left\| \begin{bmatrix} x_0^2 & x_0 & 1 \\ x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{bmatrix} \begin{bmatrix} w_2 \\ w_1 \\ w_0 \end{bmatrix} - \begin{bmatrix} t_0 \\ t_1 \\ t_2 \\ t_3 \end{bmatrix} \right\|^2$$

$$J(w) = \sum_{n=1}^4 [w_2 x^2 + w_1 x + w_0 - t]^2$$

$$= [w_2 x_0^2 + w_1 x_0 + w_0 - t]^2 + \dots$$

$$= w_2^2 x_0^4 + w_1^2 x_0^2 + w_0^2 - t^2 + \dots$$

$$\nabla_{w_0} J(w) = \delta w_0$$

$$w_0(k+1) = w_0(k) + \eta \delta w_0$$

$$\nabla_{w_2} J(w) = 2w_2 x_0^4 + 2w_2 x_1^4 + 2w_2 x_2^4 + 2w_2 x_3^4$$

$$w_2(k+1) = w_2(k) + \eta (2w_2 x_0^4 + 2w_2 x_1^4 + 2w_2 x_2^4 + 2w_2 x_3^4)$$

2.) a.)

$x \backslash Y$	0	1
0	$\frac{1}{8}$	$\frac{3}{8}$
1	$\frac{3}{8}$	$\frac{1}{8}$

$$P(X=0 | Y=1) = \frac{3}{4}$$

b.)

$$P(X=0) = \frac{1}{2}$$

Not statistically independent because  $P(X=0) \neq P(X=0|Y=1)$

3.) The validation set is used for test error estimation. In the K-fold method the validation set is a piece of the training set you use to estimate test error while still training your model. Your test set you would not use during training at all.



