Q.4

b) We have seen in Q.3(c) that solution for w is given by

$$(X^T R X)^{-1} X^T R Y,$$

R is a diagonal matrix of weights

Y is the vector of target values.

X is the design matrix.

We also know that, Newton-Raphson Update Equation for Logistic Regression is given by

$$W_{new} = W_{old} - (X^T W X)^{-1} X^T (Y - P)$$

Now we apply the Newton-Raphson update to the cross-entropy error function for the logistic regression model.

We see that the gradient and Hessian of this error function are given by

$$\nabla E(W) = \sum_{n=1}^{N} (y_n - t_n) \phi_n$$

$$= \phi^T (Y - t)$$

$$H = \nabla\nabla E(w) = \sum_{n=1}^{N} y_n(1-y_n)\phi_n \phi_n^T$$

$$= \phi^T R \phi$$

where we know that $\frac{d\sigma}{da} = \sigma(1-\sigma)$

R is $N \times N$ diagonal matrix

The Newton-Raphson update formula for the logistic regression model then

$$w^{(new)} = w^{(old)} - (\phi^T R \phi)^{-1} \phi^T (y-t)$$

$$= (\phi^T R \phi)^{-1} \{\phi^T R \phi w^{(old)} - \phi^T (y-t)\}$$

$$= (\phi^T R \phi)^{-1} \phi R z. \rightarrow (a)$$

where z is an N-dimesional vector with elements

$$z = \phi w^{(old)} - R^{-1}(y-t)$$

$\Rightarrow$ we see that the update formala (a) ~~is~~ is in WLS wher R depends on the w.

at each iteration we use new weights. vector w to compute a revised weighting matrix for this reason, the algorithm is known as iterative reweighted least square.