$$\boxed{w \leftarrow w - H(w)^{-1} \nabla L(w)}$$

Let $\quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}, \quad z = \begin{bmatrix} z_1 \\ \vdots \\ z_N \end{bmatrix}, \quad z_i = \sigma(w^T x_i)$

$$L(w) = \sum_i \left(y_i \log z_i + (1-y_i) \log(1-z_i) \right)$$

$$\frac{\partial}{\partial w} L(w) = \sum_i \left(\frac{y_i}{z_i} - \frac{1-y_i}{1-z_i} \right) \frac{\partial z_i}{\partial w} \quad (\because \text{Chain Rule})$$

$$= \sum_i \left(\frac{y_i(1-z_i) - z_i(1-y_i)}{z_i(1-z_i)} \right) \sigma(w^T x_i)(1-\sigma(w^T x_i)) \frac{\partial}{\partial w}(w^T x_i)$$

$(\because$ Chain Rule$)$

$$= \left(\sum_i (y_i - z_i) x_i \right)_{n \times 1}$$

$$H(w) = \nabla_w \left(\sum_i (y_i - z_i) x_i \right)$$

$$= \nabla_w \left(\sum_i (y_i - \sigma(w^T x_i)) x_i \right)$$

$$= \sum_i -x_i^T \sigma(w^T x_i) (1 - \sigma(w^T x_i)) \frac{\partial}{\partial w}(w^T x_i)$$

$$= -\sum_i x_i^T \left(\sigma(w^T x_i)(1-\sigma(w^T x_i)) \right) x_i$$

$$= -\begin{bmatrix} x_1 & x_2 & \cdots & x_N \end{bmatrix} \operatorname{diag}\left(z_1(1-z_1), z_2(1-z_2), \cdots, z_N(1-z_N) \right) \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

Let $S = \operatorname{diag}\left(z_1(1-z_1), z_2(1-z_2), \cdots, z_N(1-z_N) \right)$

$$H(w) = -X^T S X \quad \left(\text{Let } X = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \right)$$

$$\nabla L(w) = X^T(y - z)$$

$$w - H(w)^{-1}\nabla L(w) = w - (-X^T S X)^{-1} X^T (y - z)$$

$$\boxed{= w + (X^T S X)^{-1} X^T (y - z)} \quad \blacksquare$$