

already know :  $\frac{\partial}{\partial w_k} g(w^T x^{(i)}) = g \cdot (1-g) x_k^{(i)}$ .

We have  $l(w) = \frac{1}{2} \sum (y^{(i)} - g)^2$

$$\frac{\partial}{\partial w_j} l(w) = \sum_i (y^{(i)} - g) \cdot g \cdot (1-g) \cdot x_j^{(i)}$$

$$H[k, j] = \frac{\partial}{\partial w_k} \frac{\partial}{\partial w_j} l(w) = \frac{\partial}{\partial w_k} \sum (y^{(i)} - g) \cdot g \cdot (1-g) \cdot x_j^{(i)}$$

$$= \sum \frac{\partial}{\partial w_k} y^{(i)} x_j^{(i)} g(1-g) - \frac{\partial}{\partial w_k} x_j^{(i)} g^2 (1-g)$$

$$= \sum \frac{\partial}{\partial w_k} y^{(i)} x_j^{(i)} g - \frac{\partial}{\partial w_k} y^{(i)} x_j^{(i)} g^2 - \frac{\partial}{\partial w_k} x_j^{(i)} g^2 + \frac{\partial}{\partial w_k} x_j^{(i)} g^3$$

$$= \sum y^{(i)} x_j^{(i)} g(1-g) x_k^{(i)} - 2 y^{(i)} x_j^{(i)} g \cdot g(1-g) x_k^{(i)} - 2 x_j^{(i)} g \cdot g(1-g) x_k^{(i)} + 3 g^2 \cdot g(1-g) x_k^{(i)}$$

$$= \sum g(1-g) x_k^{(i)} x_j^{(i)} y^{(i)} - 2 g^2 (1-g) x_j^{(i)} x_k^{(i)} y^{(i)} - 2 g^2 (1-g) x_j^{(i)} x_k^{(i)} + 3 g^3 (1-g) x_j^{(i)} x_k^{(i)}$$

$$= \sum g(1-g) x_j^{(i)} x_k^{(i)} [y^{(i)} - 2 g y^{(i)} - 2 g + 3 g^2]$$

$$H[l(w)] = X^T \text{DIAG}(g(1-g), X \cdot [\text{sum}(y^{(i)} - 2 g y^{(i)} - 2 g + 3 g^2)])$$

$$H[k, j] = \sum g(1-g) x_k^{(i)} x_j^{(i)} \cdot [y^{(i)} - 2 g y^{(i)} - 2 g + 3 g^2]$$

$$W := W + \eta [X^T \cdot \text{diag}(g(1-g)) \cdot X * \text{sum}(y^{(i)} - 2 g y^{(i)} - 2 g + 3 g^2)]^{-1} \cdot X * ydH \cdot g(1-g)$$

inverse Hessian gradient