$$\frac{\partial J(\omega)}{\partial \omega_j} = \sum_i \frac{\partial}{\partial \omega_j} l(y_i; x_i; \omega)$$

$$\frac{\partial}{\partial w_j} l(y_i; x_i; w) = \frac{\partial}{\partial w_j}\Big[(1-y_i)\log(1-P(x_i;w)) + y_i \log P_1(x_i;w)\Big]$$

$$= (1-y_i)\frac{1}{1-P_1(x_i;w)}\left(-\frac{\partial P_1(x_i;w)}{\partial w_j}\right) + y_i \frac{1}{P_1(x_i;w)} \frac{\partial P_1(x_i;w)}{\partial w_j}$$

$$= \left[\frac{-(1-y_i)}{1-P_1(x_i;w)} + \frac{y_i}{P_1(x_i;w)}\right]\frac{\partial P_1(x_1;w)}{\partial w_j}$$

$$= \left[\frac{y_i - y_i P_1(x_i;w) - P_1(x_i;w) + y_i P_1(x_i;w)}{P_1(x_i;w)\,(1-P_1(x_i;w))}\right]\frac{\partial P_1(x_i;w)}{\partial w_j}$$

$$= \left[\frac{y_i - P_1(x_i;w)}{P_1(x_i;w)\,(1-P_1(x_i;w))}\right]\frac{\partial P_1(x_i;w)}{\partial w_j}$$

1

$$P_1(x_i;w) = \frac{1}{1+e^{-w\cdot x_i}}$$

$$\frac{\partial P_1(x_i;w)}{\partial w_j} = \frac{1}{(1+e^{-w\cdot x_i})^2} \frac{\partial}{\partial w_j}(1+e^{-w\cdot x_i}]$$

$$= \frac{1}{\{1+e^{(-w\cdot x_i)}\}^2} (e^{-w\cdot x_i}) \frac{\partial}{\partial w_j}(-w\cdot x_j)$$

$$= \frac{e^{(-w\cdot x_i)}}{(1+e^{-w\cdot x_i})} \frac{1}{(1+e^{-w\cdot x_i})} x_{ij}$$

$$= P_1(x_i;w) \; (1-P_1(x_i;w)) \; x_j.$$

(2)

---

(2) in (1)

$$\frac{\partial}{\partial w_j} l(y_i;x_i;w) = \left[\frac{(y - P_1(x_i;w))}{P_1(x_i;w)\,(1-P_1(x_i;w))}\right] P_1(x_i;w)(1-P_1(x_i;w)\, x_{ij}$$

$$= (y_i - P_1(x_i;w))\, x_{ij}$$

Over all gradiet

$$\frac{\partial J(w)}{\partial w_j} = \sum_i (y_i - P_1(x_i;w))\, x_{ij}$$