2.) So the given function is: –

$$P\left(y=1/x,w\right) = 1/2\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)$$

and,

$$P(y=0|x,w) = 1 - 1/2\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)$$

So the bernouli form is:

$$P\left(y|x,w\right) = \begin{cases} \left(1/2\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right)^{y_i} & y_i=1 \\ \left(1-1/2\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right)^{1-y_i} & y_i=0 \end{cases}$$

So, for one independent event: –

$$P_i\left(y_i|x,w\right) = \left(1/2\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right)^{y_i} \left(1-1/2\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right)^{1-y_i}$$

So for n data points (independent & identical): –

$$P\left(y|x,w\right) = \prod_{i=1}^{N}\left(\frac{1}{2}\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right)^{y_i} \times \left(1-1/2\left(1+\frac{w^Tx}{\ldots}\right)\right)^{1-y_i}$$

So let's take the loglikelihood :-

$$ll(w) = \log\left[\prod_{i=1}^{N}\left[\tfrac{1}{2}\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right]^{y_i}\left[1-\tfrac{1}{2}\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right]^{1-y_i}\right]$$

So,

$$ll(w) = \sum_{i=1}^{N} \log\left[\tfrac{1}{2}\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right]^{y_i} + \log\left[1-\tfrac{1}{2}\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right]^{1-y_i}$$

$$= \sum_{i=1}^{N} y_i \log\left[\tfrac{1}{2}\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right] + (1-y_i)\log\left[1-\tfrac{1}{2}\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right]$$

$$= \sum_{i=1}^{N} y_i \log\left[\tfrac{1}{2}\left(1+\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right] + (1-y_i)\left[1-\tfrac{1}{2}-\tfrac{1}{2}\left(\frac{w^Tx}{\sqrt{1+(w^Tx)^2}}\right)\right]$$

$$= n\log(\tfrac{1}{2}) + \sum_{i=1}^{N} y_i \log\left(1+\frac{w^Tx_i}{\sqrt{1+(w^Tx)^2}}\right) + (1-y_i)\log\left[1-\frac{w^Tx_i}{\sqrt{1+(w^Tx)^2}}\right]$$

$$\frac{\partial\, ll(w)}{\partial w} = \sum_{i=1}^{N} y \left(\frac{1}{1 + \frac{w^T x_i}{\sqrt{1+(w^T x_i)^2}}} \right)$$

$$\left((w^T x_i) \times \cancel{2}^{-1} \times \left(1+(w^T x)^2\right)^{-3/2} \times \cancel{2} \times (w^T x_i)\, x_i \right)$$

$$+ (1-y_i) \frac{1}{\left(1 - \frac{w^T x_i}{\sqrt{1+(w^T x_i)^2}}\right)} \left((-w^T x_i)(\cancel{2})^{-1} (1+(w^T x_i) \times \cancel{2} \times (w^T x_i) + (1+(w^T x_i)^2 \cdot x_i \right)$$

$$= \sum_{i=1}^{N} y_i \left(\frac{\sqrt{1+(w^T x_i)^2}}{\sqrt{1+(w^T x_i)^2} + w^T x_i} \right) \times$$

$$\left(\frac{(-w^T x_i)^2 x_i}{(1+(w^T x_i)^2)^{3/2}} + \frac{x_i}{(1+(w^T x)^2)^{1/2}} \right)$$

$$- (1-y_i) \left(\frac{\sqrt{1+(w^T x_i)^2}}{\sqrt{1+(w^T x_i)^2} - w^T x_i} \right)$$

$$\left(\frac{-(w^T x_i)^2 x_i}{(1+(w^T x_i)^2)^{3/2}} + \frac{x_i}{(1+(w^T x_i)^2)^{1/2}} \right)$$

$$= \sum_{i=1}^{N} x_i y \left(\sqrt{1+(w^T x_i)^2} - w^T x_{ij} \right) \left(\frac{-(w^T x_{ij})^2}{1+(w^T x_i)^2} + 1 \right)$$

$$- x_i (1-y_i) \left(\sqrt{1+(w^T x_i)^2} + w^T x_i \right) \left(\frac{-(w^T x_i)^2}{1+(w^T x_i)} + 1 \right)$$

So,

$$= \sum_{i=1}^{N} \frac{y_i \left(\sqrt{1+(w^T x_i)^2} - w^T x_i\right) x_i - (1-y_i)\left(\sqrt{1+(w^T x_i)^2} + w^T x_i\right) x_i}{1+(w^T x_i)^2}$$

$$= \sum_{i=1}^{N} x_i \frac{\sqrt{1+(w^T x_i)^2} \cdot y - y_i \cancel{w^T x_i} - \sqrt{1+(w^T x_i)^2} - w^T x_i + \sqrt{1+(w^T x_i)^2}\, y + \cancel{w^T x_i y}}{1+(w^T x_i)^2}$$

$$= \sum_{i=1}^{N} x_i \frac{2y\sqrt{1+(w^T x_i)^2} - \sqrt{1+(w^T x_i)^2} - w^T x_i}{1+(w^T x_i)^2}$$

$$= \sum_{i=1}^{N} x_i \frac{(2y-1)\left(\sqrt{1+(w^T x_i)^2}\right) - w^T x_i}{1+(w^T x_i)^2}$$

So, the negitive log-liklihood is,

$$= -\sum_{i=1}^{N} x_i \frac{(2y-1)\left(\sqrt{1+(w^T x_i)^2}\right)}{1+(w^T x_i)^2} - \frac{w^T x_i}{1+(w^T x)^2}$$

So,

$$= \cancel{X}_i \left[\frac{1-2Y}{\sqrt{1+(w^T x)^2}} + \frac{w^T x}{1+(w^T x)^2} \right.$$

Since we cannot compute a solution as it is not in closed form.

we will use ~~step function~~ & step size iteratively to update weight using the above ~~function~~ gradient.

— x —

Scanned by CamScanner