ECE 565 Project

Luc Bouchard

November 19, 2020

$$\begin{aligned}&(\vec{x}_i, y_i) \quad i \in 1, 2, 3...n\\&X \in \ N(0, \sigma^2 I_d)\\&Y \in \ 0, 1\end{aligned} \tag{1}$$

1 Model

$$P(y = 1|\vec{x}; \vec{w}) = \frac{e^{\vec{w}^T \vec{x}}}{1 + e^{\vec{w}^T \vec{x}}} \tag{2}$$

2 Parameters

$$\begin{aligned}&\vec{\theta} = \vec{w}\\&\vec{y} = [y_1, y_2, y_3, ...y_n]^T\end{aligned} \tag{3}$$

3 Liklihood Function

$$f(y_1...y_n; \vec{w}) = \prod_{i=1}^{n} \Big(\frac{e^{\vec{w}^T \vec{x}_i}}{1 + e^{\vec{w}^T \vec{x}_i}}\Big)^{y_i} \Big(1 - \frac{e^{\vec{w}^T \vec{x}_i}}{1 + e^{\vec{w}^T \vec{x}_i}}\Big)^{1-y_i} \tag{4}$$

4 CRLB

$$CRLB = I(\theta)^{-1} \tag{5}$$

$$I_{ij}(\vec{w}) = -E_Y \Big[\frac{\partial^2 \log f(y_1...y_n; \vec{w})}{\partial w_i \partial w_j}\Big] \tag{6}$$

Note

$$\frac{\partial^2 \log f(y_1...y_n; \vec{w})}{\partial w_i \partial w_j} = \sum_{z=1}^{n} \frac{\partial \log f(y_z|\vec{x}_z; \vec{w})}{\partial w_i \partial w_j} \tag{7}$$

Consider $f(y|\vec{x}; \vec{w})$,

$$f(y|\vec{x};\vec{w}) = \Big(\frac{e^{\vec{w}^T\vec{x}}}{1+e^{\vec{w}^T\vec{x}}}\Big)^y \Big(1 - \frac{e^{\vec{w}^T\vec{x}}}{1+e^{\vec{w}^T\vec{x}}}\Big)^{1-y} \tag{8}$$

$$f(y|\vec{x};\vec{w}) = \Big(\frac{e^{\vec{w}^T\vec{x}}}{1+e^{\vec{w}^T\vec{x}}}\Big)^y \Big(\frac{1}{1+e^{\vec{w}^T\vec{x}}}\Big)^{1-y} \tag{9}$$

$$\log f = y\vec{w}^T\vec{x} - \log(1+e^{\vec{w}^T\vec{x}}) \tag{10}$$

$$\frac{\partial \log f}{\partial w_i} = x_i\Big(y - \frac{e^{\vec{w}^T\vec{x}}}{1+e^{\vec{w}^T\vec{x}}}\Big) \tag{11}$$

$$\frac{\partial^2 \log f}{\partial w_i \partial w_j} = -P(y=1|\vec{x};\vec{w})\ P(y=0|\vec{x};\vec{w})\ x_i x_j \tag{12}$$

Note that equation (12) has no y dependence, so the expectation is trivially computed:

$$E_Y\Big[\frac{\partial^2 \log f}{\partial w_i \partial w_j}\Big] = -P(y=1|\vec{x};\vec{w})\ P(y=0|\vec{x};\vec{w})\ x_i x_j \tag{13}$$

$$E_Y\Big[\frac{\partial^2 \log f}{\partial \vec{w}^2}\Big] = -P(y=1|\vec{x};\vec{w})\ P(y=0|\vec{x};\vec{w})\ \vec{x}\vec{x}^T \tag{14}$$

Finally,

$$I(\vec{w}) = \sum_{i=0}^{n} P(y=1|\vec{x}_i;\vec{w})\ P(y=0|\vec{x}_i;\vec{w})\ \vec{x}_i\vec{x}_i^T \tag{15}$$

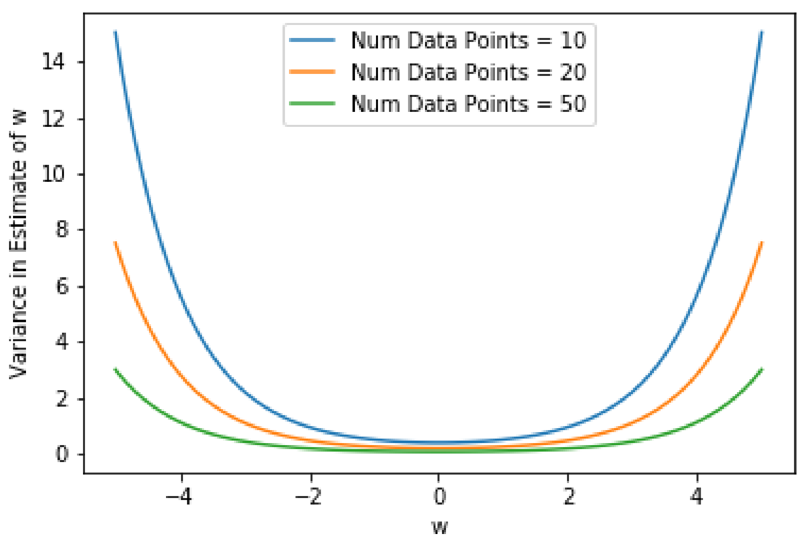
5 Consider 1D Case ($d = 1$, $x \in R, w \in R$)

$$var(\hat{w}) \geq \Big(\sum_{i=1}^{n} x_i^2\ P(y=1|x_i;w)\ P(y=0|x_i;w)\Big)^{-1} \tag{16}$$

Note that w in the equation above is the true w, and $\hat{w}$ is the estimate for w.

Let's assume all $x_i = 1$. This reduces equation (16) to the following:

$$var(\hat{w}) \geq \frac{1}{n\Big(\frac{e^w}{1+e^w}\Big)\Big(\frac{1}{1+e^w}\Big)} = \frac{(1+e^w)^2}{ne^w} \tag{17}$$



Num Data Points = 10
Num Data Points = 20
Num Data Points = 50
Variance in Estimate of w
w