classmate

Date

Page

Aditya Sawant 17CS10060

Q1.1

$$P(x|\theta) = \theta^{x}(1-\theta)^{1-x}$$

$$P(x|1-\theta) = (1-\theta)^{x}\theta^{1-x}$$

Joint probability

$$f(x_1, \ldots x_n; \theta) = \prod_i f(x_i;\theta) = \theta^{n-\sum_{i=1}^{n} x_i}(1-\theta)^{\sum_{i=1}^{n} x_i}$$

$$\log L(\theta) = \sum_i x_i \log(1-\theta) + n-\sum_i x_i \log(\theta)$$

$$\Rightarrow \frac{\partial \log L(\theta)}{\partial \theta} = -\frac{\sum_{i=1}^{n} x_i}{1-\theta} + \frac{n-\sum_{i=1}^{n} x_i}{\theta} = 0$$

$$\Rightarrow (n-\sum_i x_i)(1-\theta) = \theta \sum_i x_i$$

$$\Rightarrow n - \sum_i x_i - n\theta + \cancel{\theta \sum_i x_i} = \cancel{\theta \sum_i x_i}$$

$$\Rightarrow \theta = 1 - \frac{\sum_{i=1}^{n} x_i}{n}$$

MLE parameter of the distribution

$$1 - \frac{\sum_{i=1}^{n} x_i}{n}$$

Q1.2

$$P(x|\theta) = \frac{1}{2} e^{-|x-\theta|}$$

$$f(x_1, x_2, \dots x_{2n}; \theta) = \prod_i f(x_i; \theta)$$

$$= \left(\frac{1}{2}\right)^{2n} e^{-\sum_i |x_i - \theta|}$$

$$\log(L(\theta)) = 2n \log 1/2 - \sum_{i=1}^{2n} |x_i - \theta|$$

$$\frac{\partial \log L(\theta)}{\partial \theta} = \frac{\partial}{\partial \theta}\left[2n \log(1/2) - \sum_{i=1}^{2n} (x_i - \theta)\, sgn(x_i - \theta)\right]$$

$$\sum_{i=1}^{2n} sgn(x_i - \theta) = 0$$

no of $x_i > \theta$ = no of $x_i < \theta$

Let $m_1, m_2, \dots m_{2n}$ be the permutation such that $\{x_{m_i}\}$ is sorted

$\theta \in [x_{m_n}, x_{m_{n+1}}]$ where x_{m_n} and $x_{m_{n+1}}$ are two middle elements

for any term $|x_i - \theta|$

$$\frac{d|x_i - \theta|}{d\theta} = \begin{cases} 1 & \theta > x_i \\ -1 & \theta < x_i \end{cases}$$

Date
Page

Q2.1 $$P(y_i \mid x_i, \theta, d_1, d_2) = Z(d_1, d_2) \frac{e^{d_1(y_i - \theta^T x_i)}}{\left(d_1 e^{2(y_i - \theta^T x_i)} + d_2\right)^{\frac{d_1 + d_2}{2}}}$$

$$L_i(\theta) = L(\theta \mid x_i, y_i)$$
$$= \log(Z(d_1, d_2)) + d_1(y_i - \theta^T x_i) - \left(\frac{d_1 + d_2}{2}\right)\log\left(d_1 e^{2(y_i - \theta^T x_i)} + d_2\right)$$

$$\frac{\partial L_i(\theta)}{\partial \theta} = -d_1 x_i + \left(\frac{d_1 + d_2}{\cancel{2}}\right)\frac{d_1 e^{2(y_i - \theta^T x_i)} \cdot \cancel{2} x_i}{\left(d_1 e^{2(y_i - \theta^T x_i)} + d_2\right)}$$

$$\frac{d_1 d_2 x_i e^{2(y_i - \theta^T x_i)} - d_1^2 x_i e^{2(y_i - \theta^T x_i)} - \cancel{d_1 d_2 x_i e^{2(y_i - \theta^T x_i)}} + d_1^2 x_i e^{2(y_i - \theta^T x_i)}}{d_1 e^{2(y_i - \theta^T x_i)} + d_2}$$

$$\frac{d_1 d_2 x_i \left(e^{2(y_i - \theta^T x_i)} - 1\right)}{\left(d_1 e^{2(y_i - \theta^T x_i)} + d_2\right)}$$

$$\therefore \frac{\partial L_i(\theta)}{\partial \theta} = \frac{d_1 d_2 x_i \left(e^{2(y_i - \theta^T x_i)} - 1\right)}{d_1 e^{2(y_i - \theta^T x_i)} + d_2}$$

3.1)

Cummulative distribution function $F_{logistic}$ for logistic Distribution

$$F_{logistic} = \frac{1}{1+e^{-x/\sigma_\epsilon}}$$

Probability $P(y_i=1|\theta, x_i)$ is given as

$$\begin{aligned}
P(y_i=1, \theta, x_i) &= P(\theta^T x_i + \epsilon_i \geq 0) \\
&= 1 - P(\theta^T x_i + \epsilon_i < 0) \\
&= 1 - P(\epsilon_i < -\theta^T x_i) \\
&= 1 - F_{logistic}(-\theta^T x_i) \\
&= 1 - \frac{1}{1+e^{\frac{\theta^T x_i}{\sigma_\epsilon}}} \\
&= \frac{e^{\frac{\theta^T x_i}{\sigma_\epsilon}}}{1+e^{\frac{\theta^T x_i}{\sigma_\epsilon}}} \\
&= \frac{1}{1+e^{-\frac{\theta^T x_i}{\sigma_\epsilon}}}
\end{aligned}$$

$$\therefore P(y_i=1|\theta, x_i) = \text{logistic}\left(\frac{\theta^T x_i}{\sigma_\epsilon}\right)$$

3.2) $P(y_i \mid \theta, x_i)$

$$y_i = [\theta^T x_i + \epsilon_i \geq 0]$$

$$P(y_i \mid x_i ; \theta) = P(y_i = 1 \mid x_i, \theta)^{y_i} \cdot P(y_i = 0 \mid x_i, \theta)^{1-y_i}$$

$$P(y_i = 1 \mid x_i, \theta) = \text{logistic}(\theta^T x_i)$$

$$P(y_i = 0 \mid x_i, \theta) = 1 - \text{logistic}(\theta^T x_i)$$

$$P(y_i \mid \theta, x_i) = (\text{logistic}(\theta^T x_i))^{y_i} (1 - \text{logistic}(\theta^T x_i))^{1-y_i}$$

3.3

$$\log(P(y_i \mid \theta^T x_i))$$

$$= \log\left[\left(\frac{1}{1+e^{-\theta^T x_i}}\right)^{y_i} \left(\frac{e^{-\theta^T x_i}}{1+e^{-\theta^T x_i}}\right)^{1-y_i}\right]$$

$$= \log\left[\left(\frac{e^{\theta^T x_i}}{1+e^{\theta^T x_i}}\right)^{y_i} \left(\frac{1}{1+e^{\theta^T x_i}}\right)^{1-y_i}\right]$$

$$= y_i \theta^T x_i - \log(1+e^{\theta^T x_i})$$

3.4)

$$L_{MLE}(\theta) = \sum_{i}^{n} \log P(y_i \mid \theta, x_i)$$

$$= \sum_{i=1}^{n} \left[y_i \theta^T x_i - \log\left(1 + e^{\theta^T x_i}\right) \right] \quad \text{from 3.3}$$

$$= \sum_{i=1}^{n} y_i \left(\theta_1 x_{i1} + \theta_2 x_{i2} + \dots + \theta_d x_{id} \right) - \sum_{i=1}^{n} \log\left(1 + e^{\theta^T x_i}\right)$$

$$= \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix} \begin{bmatrix} \theta_1 x_{11} + \theta_2 x_{12} + \dots + \theta_d x_{1d} \\ \vdots \\ \theta_1 x_{n1} + \dots + \theta_d x_{nd} \end{bmatrix}_{n \times d} - 1^T_{n \times 1} \begin{bmatrix} \log\left(1 + e^{\theta^T x_1}\right) \\ \log\left(1 + e^{\theta^T x_2}\right) \\ \vdots \\ \log\left(1 + e^{\theta^T x_n}\right) \end{bmatrix}$$

$$= y^T X \theta - 1^T_{n \times 1} \begin{bmatrix} \log\left(1 + e^{\theta^T x_1}\right) \\ \log\left(1 + e^{\theta^T x_2}\right) \\ \vdots \\ \log\left(1 + e^{\theta^T x_n}\right) \end{bmatrix}$$

$$= y^T X \theta - 1^T_{n \times 1} \log \begin{bmatrix} 1 + e^{\theta^T x_1} \\ 1 + e^{\theta^T x_2} \\ \vdots \\ 1 + e^{\theta^T x_n} \end{bmatrix}$$

$$= y^T X \theta - 1^T_{n \times 1} \log \left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}_{n \times 1} + \begin{bmatrix} e^{\theta^T x_1} \\ e^{\theta^T x_2} \\ e^{\theta^T x_n} \end{bmatrix} \right)_{n \times 1}$$

$$= y^T X \theta - 1^T_{n \times 1} \log\left(1_{n \times 1} + e^{X\theta} \right)$$

3.5 $$\frac{\partial L_{MLE}(\theta)}{\partial \theta} = \frac{\partial}{\partial \theta}\left(y^T X\theta - 1^T_{n\times 1}\log\left(1_{n\times 1} + e^{X\theta}\right)\right)$$ from 3.4

$$= \frac{\partial}{\partial \theta}\left(\sum_{i=1}^{n}\left(y_i\theta^T x_i - \log\left(1+e^{\theta^T x_i}\right)\right)\right)$$ from 3.3

$$= \sum_{i=1}^{n}\left(\frac{\partial}{\partial \theta}\left(\underbrace{y_i}_{1\times 1}\underbrace{\theta^T}_{d\times 1}\underbrace{x_i}_{d\times 1} - \log\underbrace{\left(1+e^{\theta^T x_i}\right)}_{1\times 1}\right)\right)$$

$$\frac{\partial}{\partial x} x^T b = b$$

$$= \sum_{i=1}^{n}\left(y_i x_i - \frac{e^{\theta^T x_i}}{1+e^{\theta^T x_i}} x_i\right)$$

$$= \sum_{i=1}^{n} x_i\left(y_i - \frac{e^{\theta^T x_i}}{1+e^{\theta^T x_i}}\right)$$

$$= \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & & x_{n2} \\ \vdots & & & \\ x_{1d} & & & x_{nd} \end{bmatrix}_{d\times n}\left(\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} e^{\theta^T x_1}/(1+e^{\theta^T x_1}) \\ \vdots \\ e^{\theta^T x_n}/(1+e^{\theta^T x_n}) \end{bmatrix}\right)$$

$$= X^T\left(y - \begin{bmatrix} \text{logistic}(\theta^T x_1) \\ \vdots \\ \text{logistic}(\theta^T x_n) \end{bmatrix}_{n\times 1}\right)$$

$$= X^T\left(y - \text{logistic}(X\theta)\right)$$