SANDIPAN HALDAR

18MF10050

PS-1 (CS60010)

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

Now $\vartheta = 1-\theta \rightarrow \theta = 1-\vartheta$

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

$$P(x|\vartheta) = \prod_{i=1}^{n} p(x_i|\vartheta) = \prod_{i=1}^{n} \vartheta^{1-x_i}(1-\vartheta)^{x_i}$$

$$\mathcal{L}(\vartheta) = \ln p(x|\vartheta) = \sum_{i}^{n} \ln\left(\vartheta^{1-x_i}(1-\vartheta)^{x_i}\right) = \sum^{n}(1-x_i)\ln\vartheta + \sum^{n} x_i \ln(1-\vartheta)$$

For MLE $\frac{\partial \mathcal{L}(\vartheta)}{\partial \vartheta} = 0$

$$\Rightarrow \sum^{n} \frac{(1-x_i)}{\vartheta} - \frac{\sum x_i}{1-\vartheta} = 0$$

let $\sum_{i=1}^{n} x_i = S$ $\quad \frac{n-S}{\vartheta} = \frac{S}{1-\vartheta} \Rightarrow \frac{n(1-\vartheta)-S}{(1-\vartheta)\vartheta} = 0$

$$\Rightarrow \vartheta = 1 - \frac{S}{n} = 1 - \frac{\sum x_i}{n}$$

Thus $g\left(\frac{\sum x_i}{n}\right)$ is MLE of $g(\theta)$ where $g(x) = 1-x$ (Proved)

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

$$p(x|\theta) = \prod_{i=1}^{2n} \frac{1}{2}e^{-|x_i-\theta|}$$

$$\ln(P(x|\theta)) = \sum_{i=1}^{2n}\left(\ln\frac{1}{2} + \ln e^{-|x_i-\theta|}\right) = -\sum_{i=1}^{2n}|x_i-\theta|$$

$$= -\sum^{2n} \operatorname{sgn}(x_i-\theta)(x_i-\theta) \quad (\operatorname{sgn}(x) = \text{signum func})$$

$$\frac{\partial \ln(P(x|\theta))}{\partial\theta} = \sum^{2n}\operatorname{sgn}(x_i-\theta) + 0$$

$\therefore$ for MLE $\frac{\partial}{\partial\theta}\ln(P(x|\theta)) = 0 \Rightarrow \sum^{2n}\operatorname{Sgn}(x_i-\theta) = 0$

$\therefore \theta$ is median of X

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

$$\ln P_i = \mathcal{L}_i(\theta) = \ln Z + d_1(y_i - \theta^T x_i) - \left(\frac{d_1+d_2}{2}\right) \ln\left(d_1 e^{2(y_i - \theta^T x_i)} + d_2\right)$$

$$\frac{\partial \mathcal{L}_i(\theta)}{\partial \theta} = 0 + (-d_1 x_i) - \frac{d_1+d_2}{2} \, \frac{2 \cdot d_1 x_i e^{2(y_i - \theta^T x_i)}}{d_1 e^{2(y_i - \theta^T x_i)} + d_2}$$

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

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

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

3.1) $P(y_i = 1 | \theta, x_i)$.

Now $y_i = \mathbb{1}[x\theta + \epsilon \geq 0]$

$$P(y_i = 1 | \theta, x_i) = P(\theta^T x_i + \epsilon_i \geq 0)$$

$$P(-\epsilon_i \leq \theta^T x_i) = F(\theta^T x_i) = \frac{1}{1 + e^{-\frac{(\theta^T x_i - \mu)}{s}}}$$

$\cdot\ \epsilon_i \sim \text{logistic}(0, \sigma_\epsilon) \Rightarrow \mu = 0 \quad \sigma_\epsilon = s$

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

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

$\therefore P(y_i = 0 | \theta, x_i) = 1 - \text{logistic}\left(\frac{\theta^T x_i}{\sigma_\epsilon}\right) \quad \sigma_\epsilon = 1$

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

3.3) $\log(P(y_i | \theta, x_i)) = y_i \log\left(\frac{1}{1+e^{-\theta^T x_i}}\right) + (1-y_i)\log\left(1 - \frac{1}{1+e^{-\theta^T x_i}}\right)$

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

let $e^{-\theta^T x_i} = z$

$$= -y_i \log(1+z) + (1-y_i)\log z - (1-y_i)\log(1+z)$$

$$= -y_i \log(1+z) + \log z - y_i \log z - \log(1+z) + y_i \log(1+z)$$

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

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

34) $\ell_{MLE}(\theta) = \sum_{i=1}^{n} \log P(y_i | x_i, \theta)$

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

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

$$\downarrow \quad - \cancel{\sum} \log\left(\prod_{i=1}^{n} \left(1 + e^{\theta^T x_i}\right)\right)$$

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

Now $\sum x_i$ = (where X is (n,1) shape)

$$= 1_{(1\times n)} \times X_{(n\times 1)}$$

Now $\sum \theta^T x_i = X\theta$ [θ is (d,1) x_i is (d,1)
$\theta^T x_i$ is (1,1) and $\sum \theta^T x_i$ is (n,1))

$X\theta = \sum_{i=1}^{n} X[i].\theta$

(X is (n,d) θ is (d,1) so $X\theta$ is (n,1))

x_i is $X[i]^T \Rightarrow$ ~~[struck out]~~

Now $AB = (B^T A^T)^T$

$\sum X[i]\theta = \sum \left(\theta^T X[i]^T\right)^T = \sum (\theta^T x_i)^T = \sum(\theta^T x_i)$

$\because \theta^T x_i$ is scalar.

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

$\sum \left(x_i^T \theta . y_i^T\right)^T - \sum \log\left(1_{n\times 1} + e^{X\theta}\right)$ [log of a matrix represent log of its values]

$y^T(X\theta) - 1_{1\times n} \log\left(1_{n\times 1} + e^{X\theta}\right)$ ($\because y_i$ is scalar)

35)

$$\frac{\partial L(\theta)}{\partial \theta} = (y^T X)^T - \frac{X^T e^{X\theta}}{1+e^{X\theta}} \quad \text{(dimension is } d \times 1\text{)}$$

$$= (y^T X)^T - \frac{X^T}{1+e^{-X\theta}}$$

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