

Question 1

Ruotian Wang

$$y_i = \begin{cases} 1 & \text{if } y_i=1 \text{ is in the first class} \\ -1 & \end{cases}$$

$$\text{if } y_i=1 \quad P(y_i=1 | X, \beta) = \frac{1}{1+e^{-\beta^T x}} = \frac{1}{1+e^{-y_i \beta^T x}}$$

$$\text{if } y_i=-1 \quad P(y_i=-1 | X, \beta) = 1 - P(y_i=1 | X, \beta) = 1 - \frac{1}{1+e^{-\beta^T x}} = \frac{1}{1+e^{\beta^T x}} = \frac{1}{1+e^{-y_i \beta^T x}}$$

$$\therefore P(y | X, \beta) = G(y \beta^T x) = \frac{1}{1+e^{-y \beta^T x}}$$

$$L(\beta) = \prod_{i=1}^n P(y_i | X_i, \beta) = \prod_{i=1}^n \frac{1}{1+e^{-y_i \beta^T x_i}}$$

$$\text{we define } J(\beta) = \log(L(\beta))$$

$$\therefore J(\beta) = \log \left(\prod_{i=1}^n \frac{1}{1+e^{-y_i \beta^T x_i}} \right) = \sum_{i=1}^n \log \left(\frac{1}{1+e^{-y_i \beta^T x_i}} \right)$$

$$\therefore \hat{\beta}_{MLE} = \arg \min_{\beta} J(\beta)$$

$$\text{we assume } t_i = y_i \beta^T x_i$$

$$\begin{aligned} \therefore \frac{\partial J(\beta)}{\partial \beta} &= \sum_{i=1}^n \frac{\partial J(\beta)}{\partial t_i} \cdot \frac{\partial t_i}{\partial \beta} \\ &= \sum_{i=1}^n -\frac{1}{1+e^{-y_i \beta^T x_i}} \cdot e^{-y_i \beta^T x_i} \cdot -y_i x_i \\ &= \sum_{i=1}^n \frac{1}{1+e^{y_i \beta^T x_i}} \cdot y_i x_i \end{aligned}$$