Complexity In $\begin{cases}\text{Linear Regression}\\ \text{Logistic Regression}\end{cases}$

① linear Regression

Model: $y = w^T x + b = \theta^T x$

Loss function: $F(w,b) = F(\theta) = \sum_{i=1}^{n} (y_i - \theta^T x_i)^2$

$$= \sum_{i=1}^{n} F_i(\theta)$$

$$F_i(\theta) = (y_i - \theta^T x_i)^2$$

$$\nabla_\theta F_i(\theta) = \nabla_\theta z \quad \nabla_z F_i(z)$$

$$= x_i \cdot 2(y_i - \theta^T x_i)$$

$$= -x_i \cdot 2(y_i - \theta^T x_i)$$

$$\nabla_\theta^2 F_i(\theta) = \nabla_\theta [-2(y_i - \theta^T x_i) \cdot x_i)$$

$$= \nabla_\theta [2(\theta^T x_i - y_i) \cdot x_i]$$

$$= 2 x_i \cdot x_i^T \succeq 0 \quad \Rightarrow \quad F_i(\theta) \text{ convex on } \underline{\theta}$$

$$\Rightarrow \quad \sum F_i(\theta) = F(\theta) \text{ convex on } \underline{\theta}$$

② $\boxed{\text{Logistic Regression}}$

$$P(y=1 \mid x) = \sigma(\theta^T x)$$

$$\boxed{\sigma(z) = \frac{1}{1+\exp(-z)}}$$

$$= \frac{1}{1+\exp(-\theta^T x)}$$

$$1 - \sigma(\theta^T x) = \frac{\exp(-\theta^T x)}{1+\exp(-\theta^T x)}$$

Model (Cross-Entropy Loss) $\longleftrightarrow$ MLE

$$F(\theta) = -\left[\sum y_i \log \sigma(\theta^T x_i) + (1-y_i) \log (1 - \sigma(\theta^T x_i))\right]$$

$$= \sum_i y_i (-\log \sigma(\theta^T x_i)) + \sum_i (1-y_i)(-\log(1-\sigma(\theta^T x_i)))$$

Aim: $f(\cdot)$ & $g(\cdot)$ are convex

$\rightarrow f(\theta) = -\log \sigma(\theta^T x_i) \qquad g(\theta) = -\log(1 - \sigma(\theta^T x_i))$

① $f(\theta) = -\log\left(\frac{1}{1+\exp(-\theta^T X_i)}\right)$

$= \log(1+\exp(-\theta^T X_i))$

$\nabla f(\theta) = \nabla_\theta z \;\; \nabla_z \log(1+\exp(z))$

$= -X_i \frac{\exp(z)}{1+\exp(z)} \qquad z = -\theta^T X_i$

$\Rightarrow \boxed{f(\theta) \ \text{convex}}$

$= -X_i \frac{\exp(-\theta^T X_i)}{1+\exp(-\theta^T X_i)}$

$= X_i \, (\sigma(\theta^T X_i) - 1)$

$\nabla^2 f(\theta) = \nabla_\theta [(\sigma(\theta^T X_i) - 1) \cdot X_i]$

$= \sigma(\theta^T X_i) \cdot (1 - \sigma(\theta^T X_i)) \cdot X_i \cdot X_i^T \succeq 0$

② $g(\theta) = -\log\left(1 - \frac{1}{1+\exp(-\theta^T X_i)}\right)$

$= -\log\left(\frac{\exp(-\theta^T X_i)}{1+\exp(-\theta^T X_i)}\right)$

$= \theta^T X_i + \log(1+\exp(-\theta^T X_i))$

convex + affine
= convex

$\Rightarrow F(\theta) = \sum_{i=1}^{n} Y_i f_i(\theta) + \sum_{i=1}^{n} (1-Y_i) g_i(\theta) \rightsquigarrow \boxed{\text{convex}}$