

學號：B03901144 系級：電機四 姓名：郭彥勝

請實做以下兩種不同 feature 的模型，回答第 (1) ~ (3) 題：

- (1) 抽全部 9 小時內的污染源 feature 的一次項(加 bias)
- (2) 抽全部 9 小時內 pm2.5 的一次項當作 feature(加 bias)

1. (2%)記錄誤差值 (RMSE)(根據 kaggle public+private 分數)，討論兩種 feature 的影響

(1) 使用 4 個 feature 以及 PM2.5

private	public
6.43557	8.61249

選擇過多的 feature 可能會造成準確度降低

(2) PM2.5 9hr

private	public
6.21691	6.92579

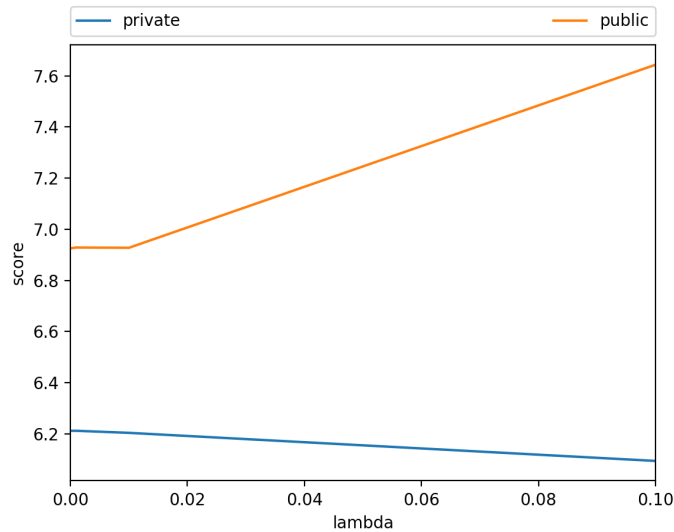
2. (1%)將 feature 從抽前 9 小時改成抽前 5 小時，討論其變化

PM2.5 5hr

private	public
6.41547	6.87568

Data 量過少也會降低準確度

3. (1%)Regularization on all the weight with $\lambda=0.1$ 、 0.01 、 0.001 、 0.0001 ，並作圖



4. Ans: $(X^T X)^{-1} X^T y$

$$\begin{aligned}
 & \min_{\theta} \sum_{j=1}^m [h_{\theta}(x^j) - y^j]^2 \\
 & \sum_{j=1}^m [h_{\theta}(x^j) - y^j]^2 = (X\theta - Y)^T (X\theta - Y) \\
 & J(\theta) = \frac{1}{2} \sum_{j=1}^m [h_{\theta}(x^j) - y^j]^2 = \frac{1}{2} (X\theta - Y)^T (X\theta - Y) \\
 & \frac{d}{d\theta} J(\theta) = \frac{1}{2} \frac{d}{d\theta} (X\theta - Y)^T (X\theta - Y) \\
 & \quad = \frac{1}{2} \frac{d}{d\theta} (\theta^T X^T X \theta - \theta^T X^T Y - Y^T X \theta + Y^T Y) \\
 & \quad = X^T X \theta - X^T Y \\
 & \frac{d}{d\theta} J(\theta) = 0 \quad (X^T X \theta - X^T Y) = 0 \\
 & \theta = (X^T X)^{-1} X^T Y
 \end{aligned}$$