學號：B03901144 系級： 電機四 姓名：郭彥滕

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

1. PM2.5 9hr

|  |  |
| --- | --- |
| private | public |
| 6.21691 | 6.92579 |

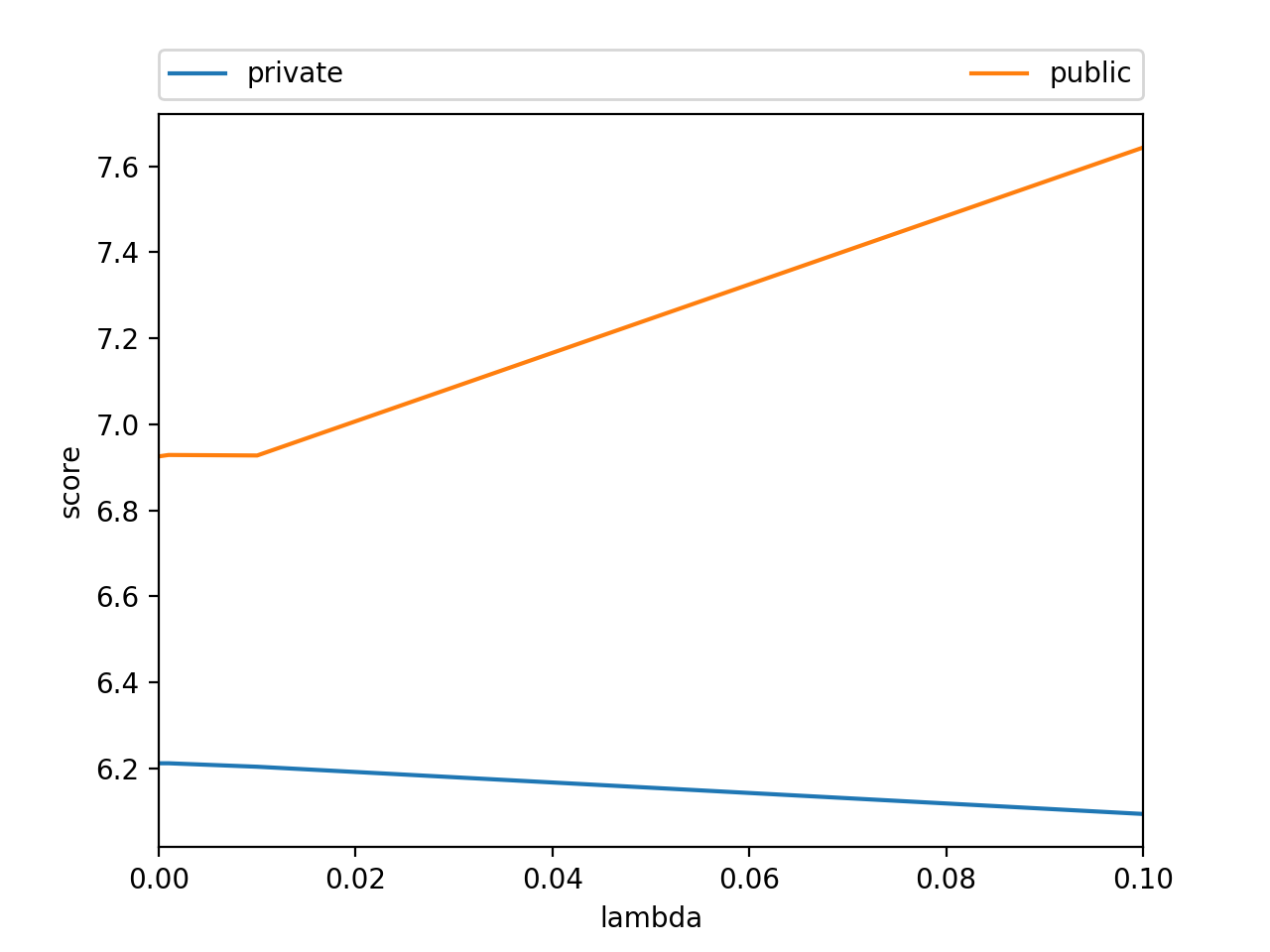
2. (1%)將feature從抽前9小時改成抽前5小時，討論其變化

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 λ=0.1、0.01、0.001、0.0001，並作圖



4. Ans: (XTX)-1XTy

