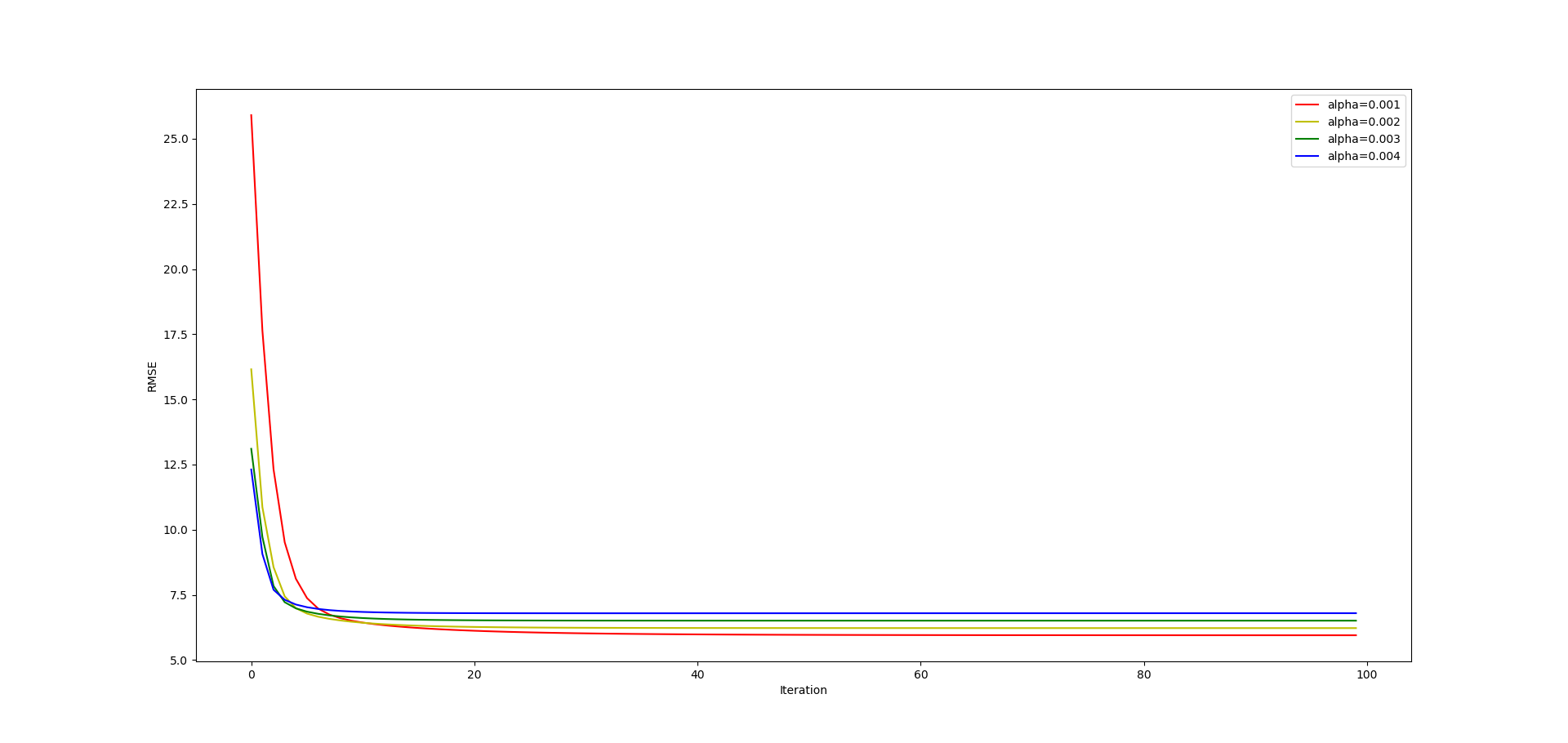
Homework 1 Report - PM2.5 Prediction

B05901022 電機三 許睿洋

1. (1%) 請分別使用至少**4**種不同數值的learning rate進行training（其他參數需一致），對其作圖，並且討論其收斂過程差異。

以下為示意圖：



在learning rate較小的時候，初始的loss相當大，但其收斂到的數值卻是比較小的；反之，在learnin rate較大的時候，雖然有相當好的初始loss，卻只能收斂到一個較大的數值。因此，慎選learning rate能夠確保訓練過程的loss能夠下降到較好的狀況。

2. (1%) 請分別使用每筆data9小時內所有feature的一次項（含bias項）以及每筆data9小時內PM2.5的一次項（含bias項）進行training，比較並討論這兩種模型的root mean-square error（根據kaggle上的public/private score）。

|  |  |  |
| --- | --- | --- |
|  | Public Score | Private Score |
| With all features | 7.65725 | 7.44538 |
| Only with PM2.5 | 8.54986 | 8.37499 |

由於PM2.5的濃度還可能受到其他因素影響(例如:下雨)，只針對PM2.5進行訓練可能會有較差的結果。

3. (1%)請分別使用至少四種不同數值的regulization parameter λ進行training（其他參數需一至），討論及討論其RMSE(traning, testing)（testing根據kaggle上的public/private

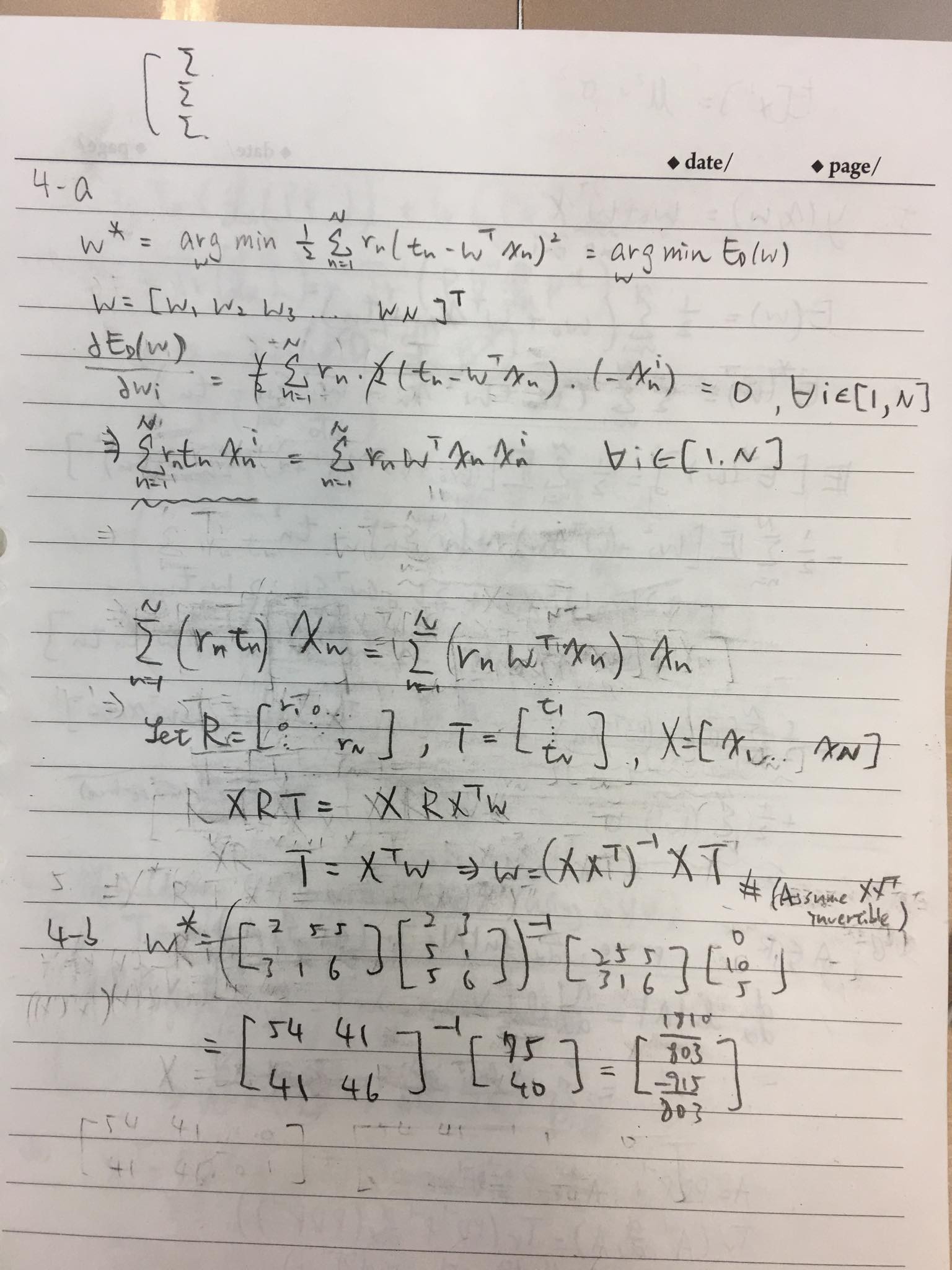
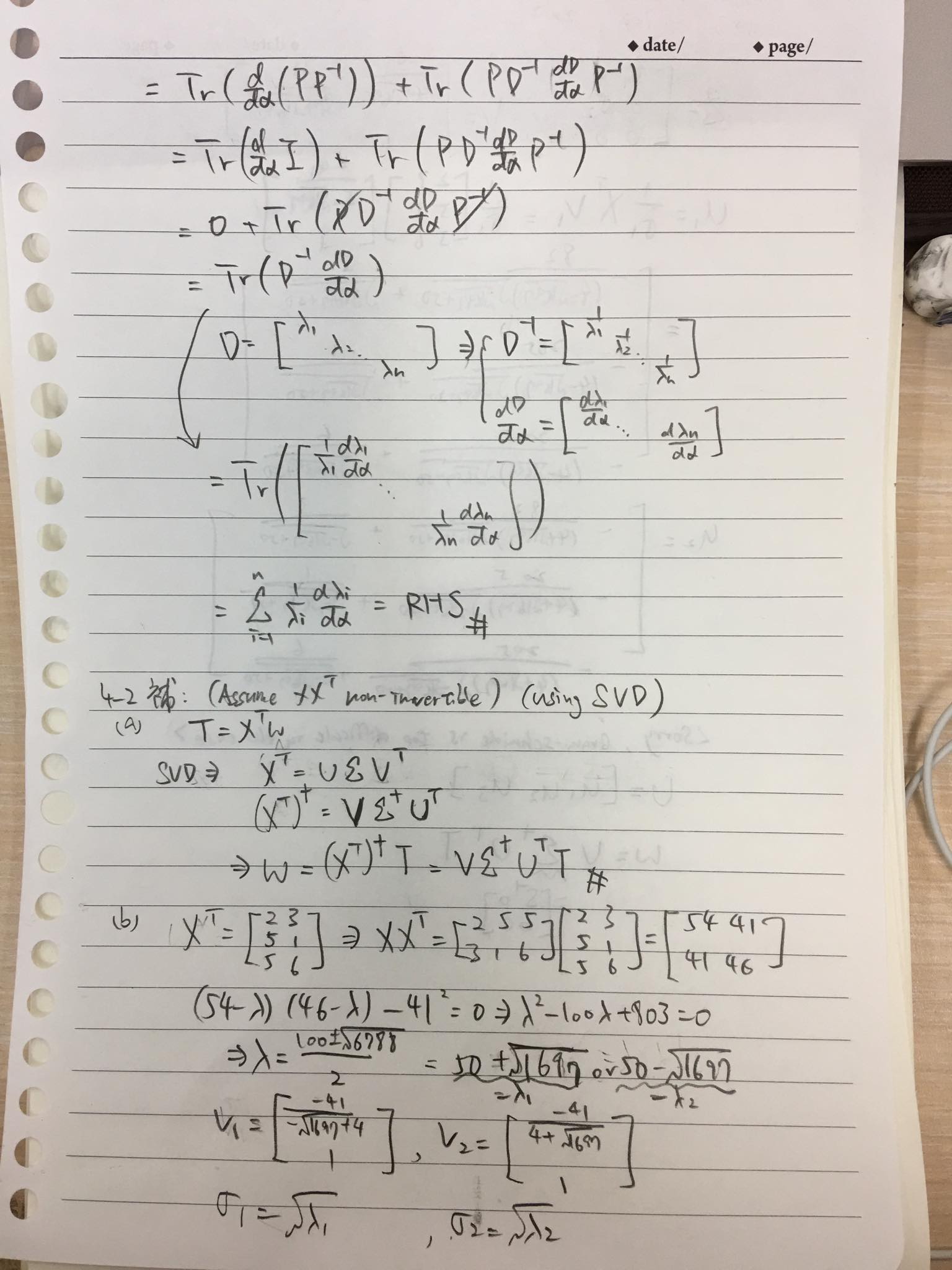
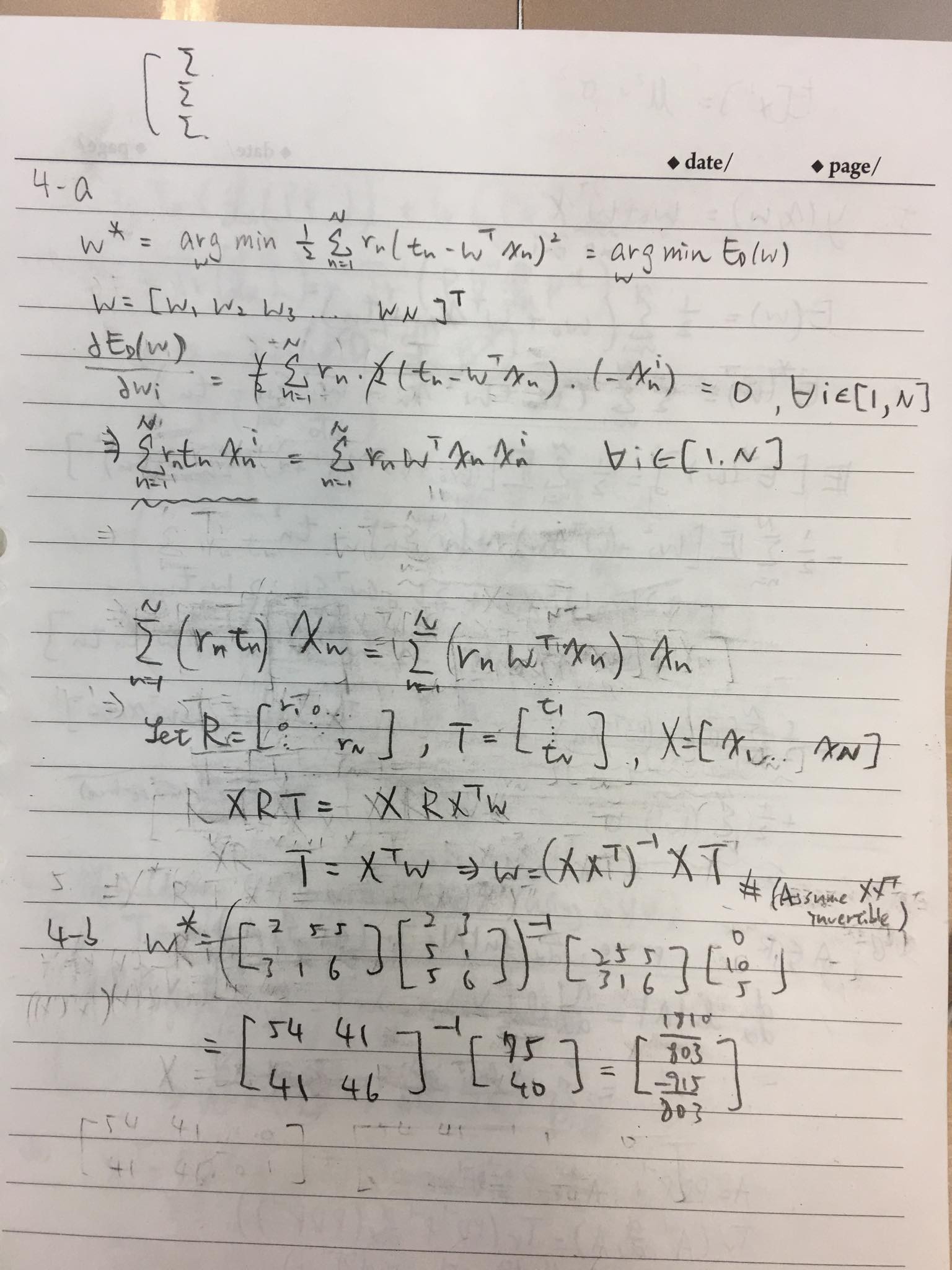
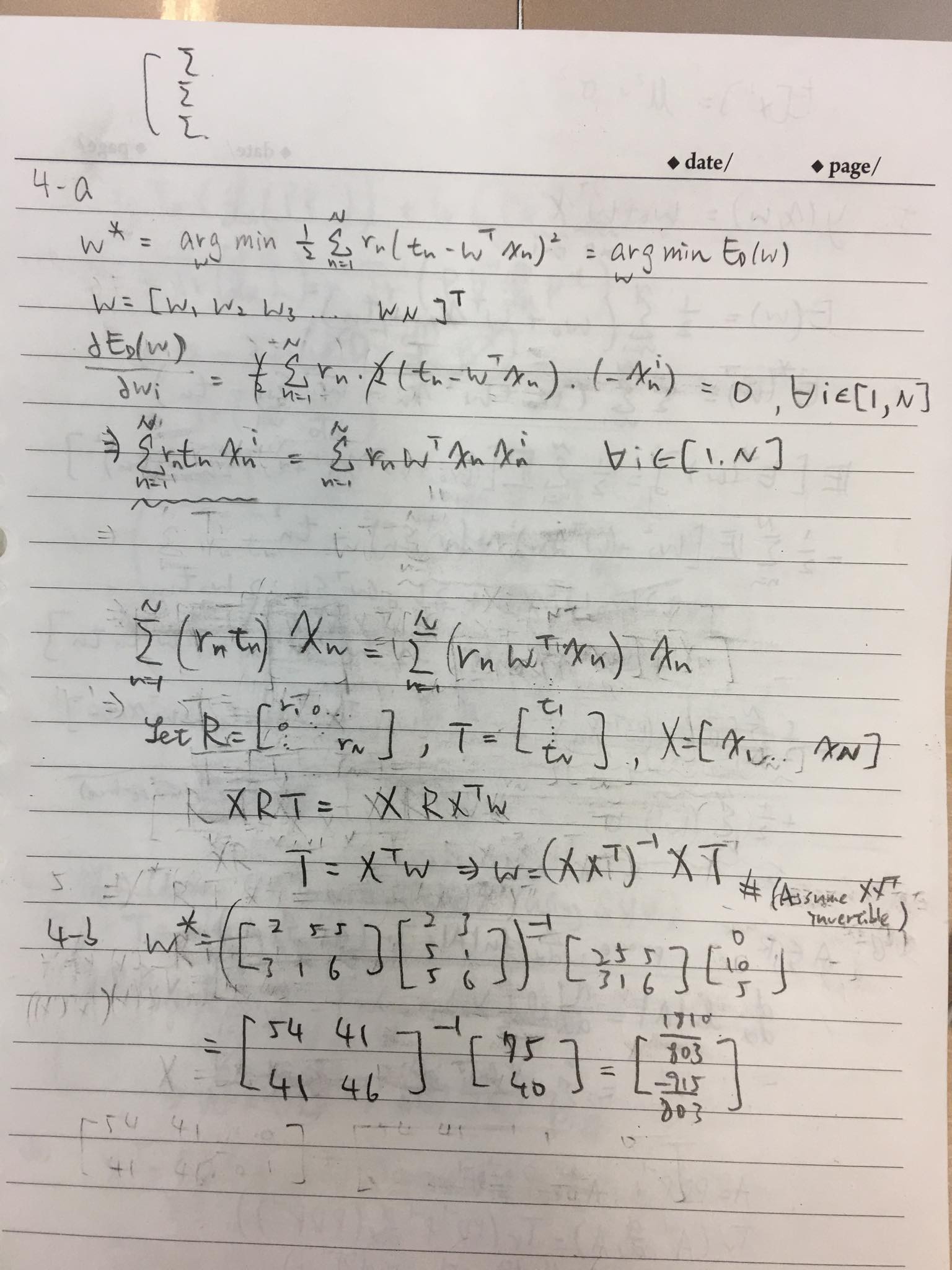
score）以及參數weight的L2 norm。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Training | Public | Private | L2Norm |
|  | 6.426675 | 7.66664 | 7.45865 | 15.669713 |
|  | 6.537951 | 7.71292 | 7.36896 | 15.019553 |
|  | 7.280320 | 7.97906 | 7.56481 | 12.144796 |
|  | 11.062287 | 9.96809 | 8.06805 | 8.504792 |

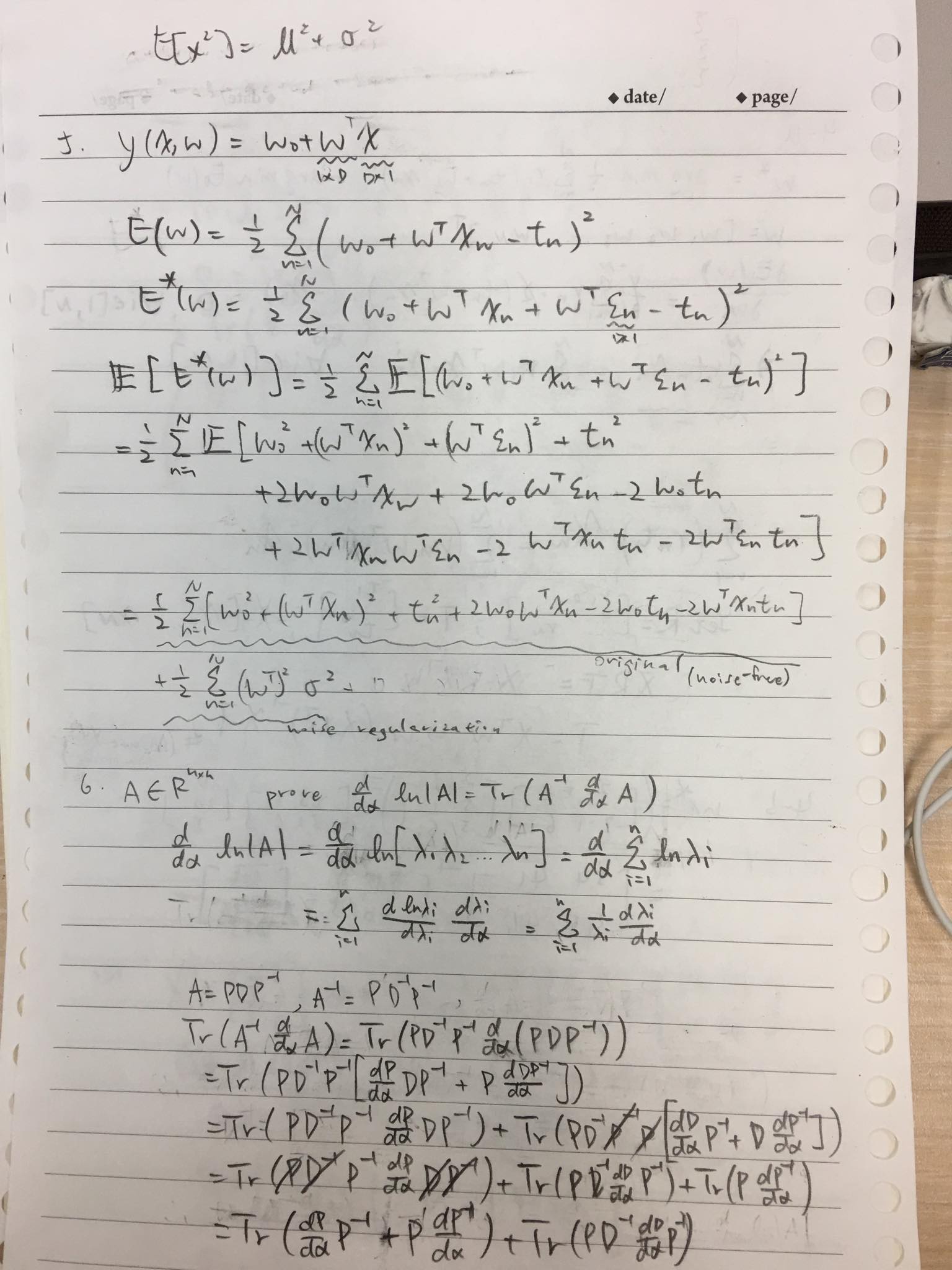
加入regularization可以使loss在訓練過程中能夠有相對穩定的變化，但是在這次的case中，加入過大的可能會使訓練結果變差，且可以從結果看出來可能是L2Norm被大幅壓縮所導致的惡化現象，因此如若要考慮加入regularization，慎選參數是必要的。

4~6數學題:

4.Collaborator: B05901009高瑋聰、B05901034劉奎元



5.Collaborator:(Self only)



6.Collaborator:

B05901009高瑋聰、B05901092歐瀚墨、B05901011許秉倫、

B05901034劉奎元、B05901082楊晟甫、B05901101陳泓廷

