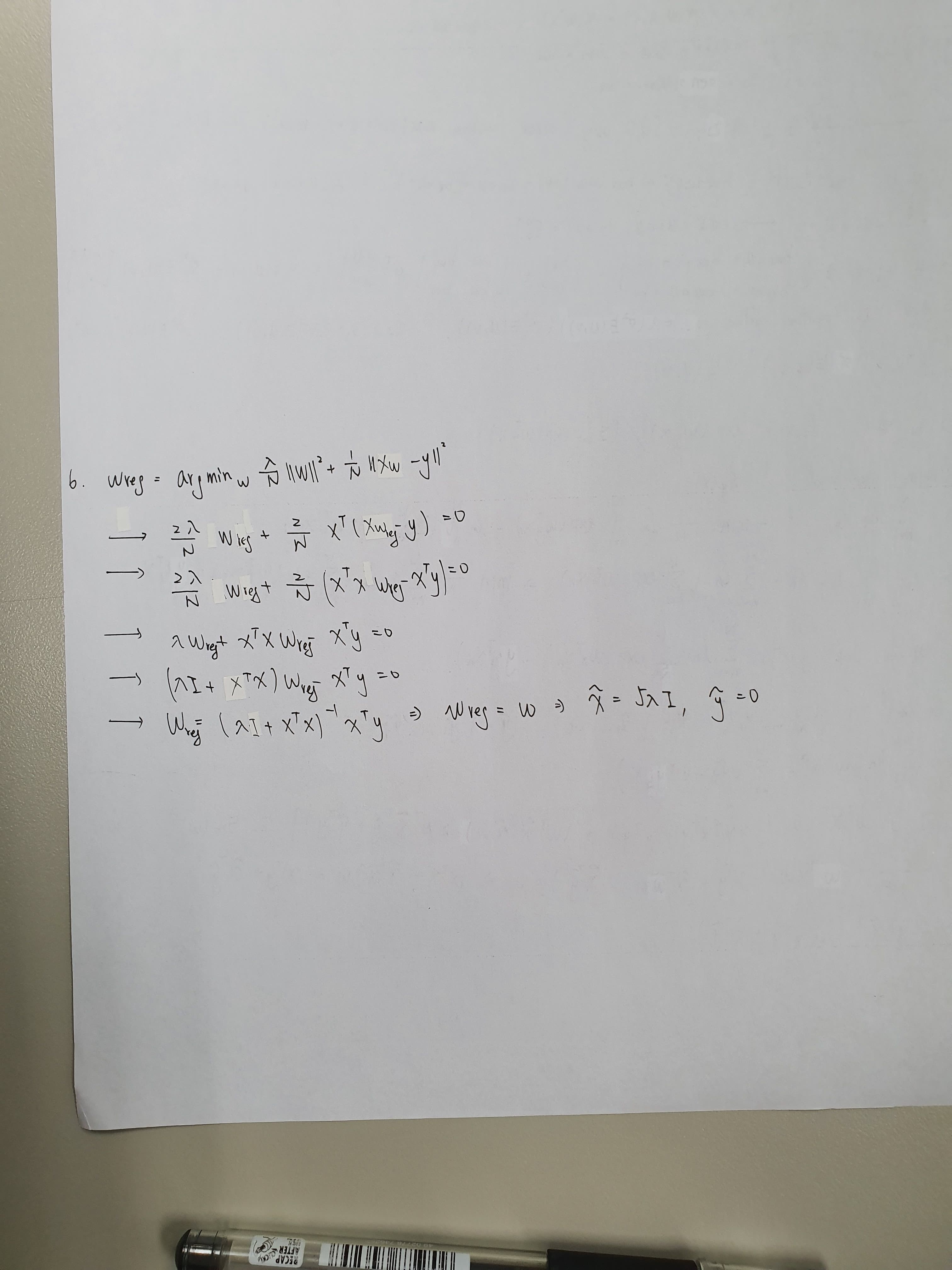
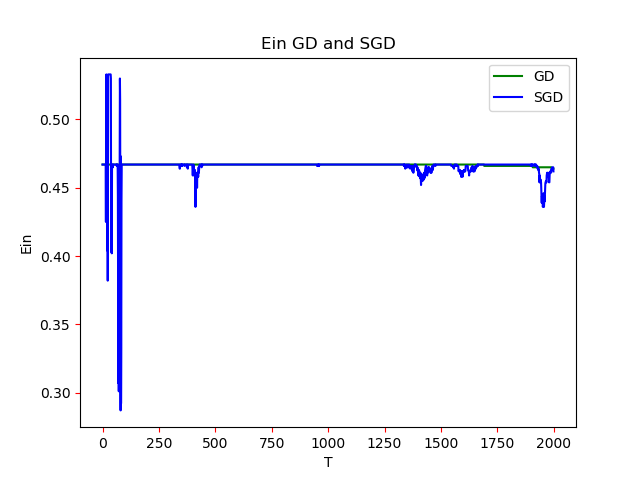
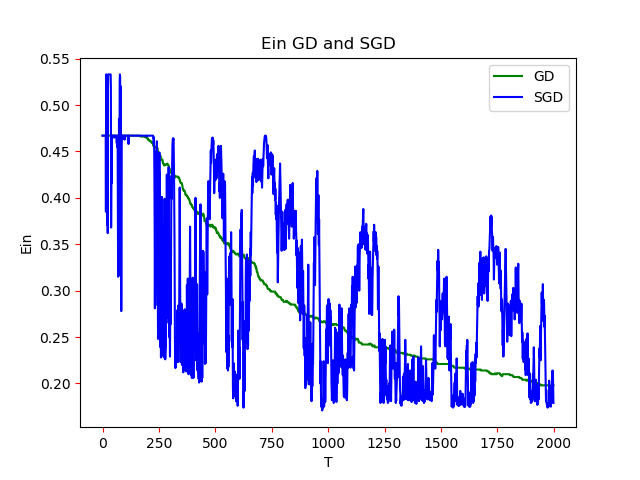


6.

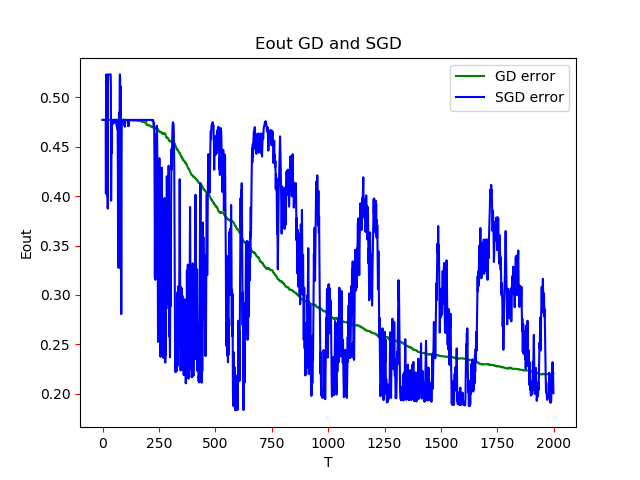


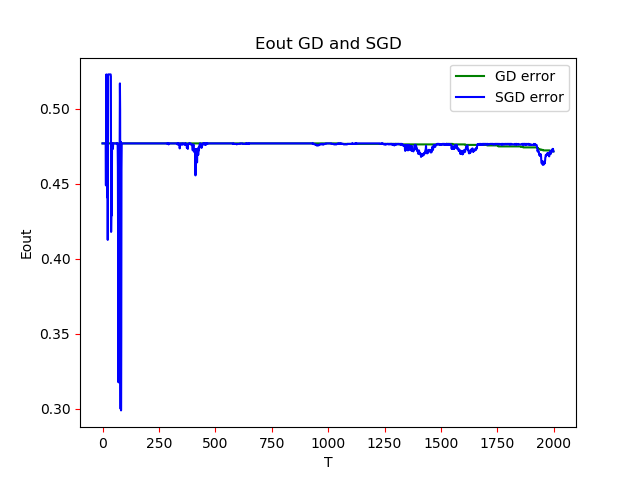
7.





Learning rate 0.01時，GD 與SGD的error rate 有明顯的下降，但Learning rate 0.001時，error rate變化不明顯，但SGD還是有機會達到error較低的時候。

8.   




Learning rate 0.01時，SGD有較明顯的震盪，Learning rate 0.001時，SGD前幾次更新有較劇烈的震盪，但最後error rate趨近GD。