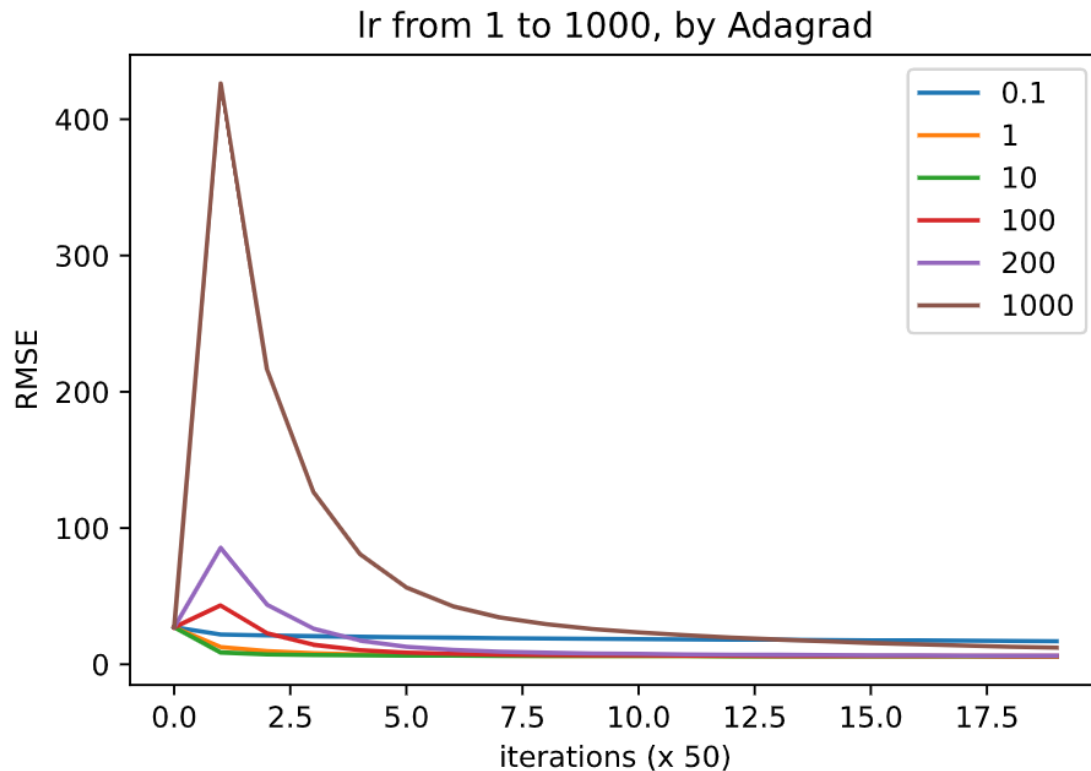


1. (2%) 使用四種不同的 learning rate 進行 training (其他參數需一致)，作圖並討論其收斂過程（橫軸為 iteration 次數，縱軸為 loss 的大小，四種 learning rate 的收斂線請以不同顏色呈現在一張圖裡做比較）。



可以發現當 adagrad 中 learning rate 若是越高，由於對一次 iter 的更新值相對更低，故在起始搜尋的效率不佳，容易使得 rmse 飆高，而開始抵達收斂區域後提供一個相對低 learning rate 更平滑的曲線。

而低 learning rate 帶給來更快速的起始搜索，但在抵達收斂區域周遭時，其更新速率會較為緩慢。

2. (1%) 比較取前 5 hrs 和前 9 hrs 的資料 ( $5 \times 18 + 1$  v.s  $9 \times 18 + 1$ ) 在 validation set 上預測的結果，並說明造成的可能原因（1. 因為 testing set 預測結果要上傳 Kaggle 後才能得知，所以在報告中並不要求同學們呈現 testing set 的結果，至於什麼是 validation set 請參考：

[https://youtu.be/D\\_S6y0Jm6dQ?t=1949](https://youtu.be/D_S6y0Jm6dQ?t=1949) 2. 9hr:取前 9 小時預測第 10 小時的 PM2.5；5hr:在前面的那些 features 中，以 5~9hr 預測第 10 小時的 PM2.5。這樣兩者在相同的 validation set 比例下，會有一樣筆數的資料）。

此處我自己手刻了 Adam，參數如下：

learning\_rate = 0.001

iter\_time = 50000

beta\_1 = 0.9

beta\_2 = 0.999

eps = 1e-07

而得到的結果為：

window	Training loss	Validation loss	size
5	5.8608062430859365	5.676321916384947	( 4560, 91 )
9	5.716747152704477	5.668570291510258	( 4521, 163 )

Training loss 上可以看到 9 比 5 還要小，可見在 Training 上 9 由於一次給予的資料量以及配適的模型參數數量較多，故可以達到更好的效果，而在 Validation 上 9 的表現也比 5 好，並沒有發生 over fitting 的問題，故可見此 case 上設定 window 為 9 比 5 來的佳。

**3. (1%) 比較只取前 9 hrs 的 PM2.5 和取所有前 9 hrs 的 features (9\*1 + 1 vs. 9\*18 + 1) 在 validation set 上預測的結果，並說明造成的可能原因。**

**optimizer** 參數設定同上

Features	Training loss	Validation loss	size
1	6.194031358522437	5.861093589266463	( 4521, 10 )
18	5.716747152704477	5.668570291510258	( 4521, 163 )

可以看到使用單一 feature 的效果較使用所有 feature 的效果來的差，不論是在 train 與 valid，使用所有 feature 皆有比較高的表現。

因為 Linear Regression 在參數越多時，越可以表達資料的變異，但也會更傾向學習 train set 資料所代表的參數空間，故需要注意的便是在使用多個變量時是否會導致 over fitting 的問題。

4. (2%) 請說明你超越 baseline 的 model(最後選擇在 Kaggle 上提交的) 是如何實作的（例如：怎麼進行 feature selection, 有沒有做 pre-processing、learning rate 的調整、advanced gradient descent 技術、不同的 model 等等）。

將 optimizer 自 tutorial 的 Adagrad 調整為 Adam，參數設定如下：

```
learning_rate = 0.001
```

```
iter_time = 50000
```

```
beta_1 = 0.9
```

```
beta_2 = 0.999
```

```
eps = 1e-07
```

其餘設定包含 data preprocessing 皆如 tutorial  
最後的 Kaggle public score 為 : 5.46443