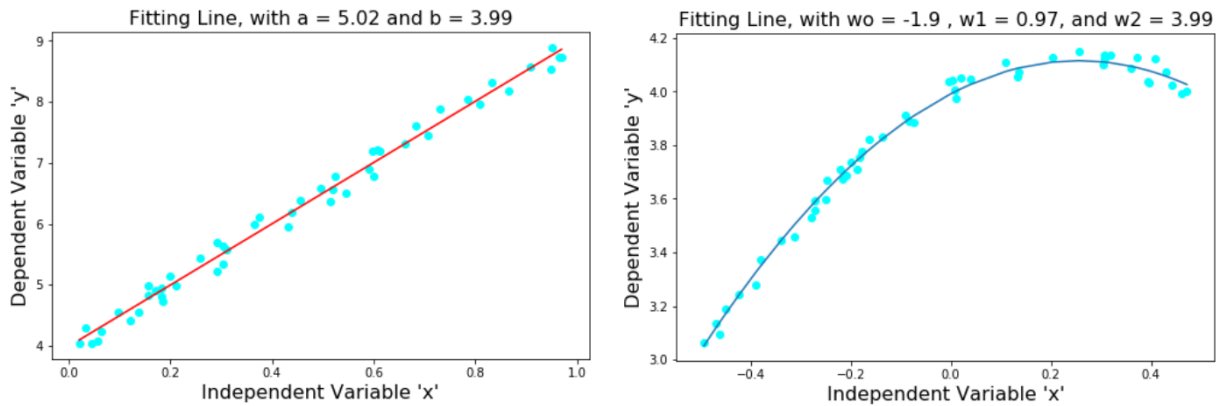


## Homework3 108003817 楊元福

### Problem A: Line Fitting

在 Problem A - Line Fitting 的過程中，跟 Problem B - License Plate Localization 比起來，比較沒有遇到困難與問題。參考助教所提供的 Pytorch 範例，就可以精準得跑出 fitting line 如下圖。



### Problem B: License Plate Localization

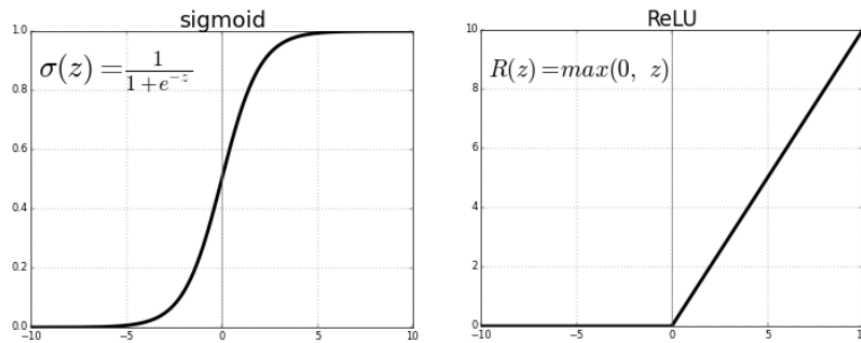
整體來說，在 Problem B 學習的過程中，首先遇到的問題是 underfitting，之後遇到的問題是 overfitting。整理這次深度學習的過程如下表：

No.	Net	Activation Function	Data Augmentation	Depth	Training Time (min)	RMSE- Training	RMSE- Testing	Result
1	6 層 Convolution/Max-pooling +1 Flatten + 3 dense layer	ReLU	N/A	16	10.1	69.08	-	underfitting
2	VGG16	ReLU	N/A	23	17.5	55.16	-	underfitting
3	ResNet50	ReLU	N/A	168	18.1	35.12	-	underfitting
4	ResNet50	Sigmoid	N/A	168	18.1	25.01	33.26	overfitting
5	ResNet50 (with Dropout at dense layer)	Sigmoid	N/A	168	18.2	33.57	-	underfitting
6	ResNet50	Sigmoid	隨機 Flipping & Rotation	168	50.5	26.05	28.14	generalized
7	InceptionResNetV2	Sigmoid	N/A	572	76.7	16.76	20.88	generalized

每一個過程遇到的問題與解決方法，詳細說明如下：

1. 首先我建立 6 層 Convolution/Max-pooling layer，搭配 1 層 Flatten 和 3 層 dense layer。在 30 次 Epoch 的訓練之後，Training RMSE 為 69.08。遠大於目標 35。此 model 嚴重 underfitting。我判斷可能的原因是模型深度不夠。

2. 承上, 我改用較深的 VGG16 模型 (如上表 No.2), 深度由 16 提升到 23 層. Training RMSE 降低到 55.16. 但是還是不理想, 因此我再改用更深的模型 ResNet50 (如上表 No.3).
3. 改用 ResNet50 之後, RMSE 再度降低到 35.12. 但是 ResNet50 的深度遠大於 VGG16 (168 vs 23), 改善成果不符合預期. 我判斷可能是最後一層的激活函數: ReLU 不適用. 根據 ReLU 的特性, 在  $x > 0$  之後, 呈現線性成長, 沒有限制. 這不符合我們位置座標的現象. 因此我改用 Sigmoid function, 限制 output 在一定的範圍內 (0~1)



4. 由於 sigmoid 限制 output 在 0 與 1 中間, 因此我必須要將 input 做轉換. 將座標位置(x, y), 改成該位置在整體影像的比例 ( $x/\text{width}$ ,  $y/\text{height}$ ), 以符合 sigmoid 輸出的特性. 訓練結果符合預期, training RMSE = 25.01. 但是 testing RMSE 為 33.26, 大於 training RMSE, 代表此模型由 underfitting 轉成 overfitting.
5. 承上, 在 No.5 中, 我添加了 Drop out function (0.1), 但是模型的 training RMSE 惡化 (25.01→33.57), 因此我不採用此方法.
6. 承上, 在 No.6 中, 我採用 Data Augmentation, 以避免過度擬合的問題. 我採用隨機 Flipping (水平/垂直) 與隨機 Rotation (正負 20 度以內). 由於增加了圖片處理, 訓練時間從 18 分鐘增加到 50 分鐘. 訓練的結果, training RMSE 降低到 26.05, testing RMS 也降低到 28.14. 符合預期.
7. 承上, 我改用更深層的網路架構 InceptionResNetV2 (深度: 572). Training RMSE 降低至 16. Testing RMSE 降低到 20. 而且是在沒有進行任何 Data Augmentation 所得到的結果. 最後我採用此模型作為最終提交結果.

這次的作業, 充分讓我了解與熟悉深度學習神經網路的建構與訓練. 心得主要有 3 點: (1) 模型的深度影響預測的結果. 深度不夠, 再多的訓練次數/資料增生也是沒有效果. (2) 模型深度越深, 或是進行 Data Augmentation, 都會嚴重增加訓練時間. 這對於 User 來說, 也是另外一個要考量的成本. (由於繳交作業時間問題, 我沒有再針對 InceptionResNetV2 進行資料增生). (3) 進行深度學習, 除了深度之外, 其他參數與函數的使用也會影響學習結果. 例如: 激活函數的選擇, Drop out function 的使用. 這些項目也是大家在建模時需要考量的重要因素.