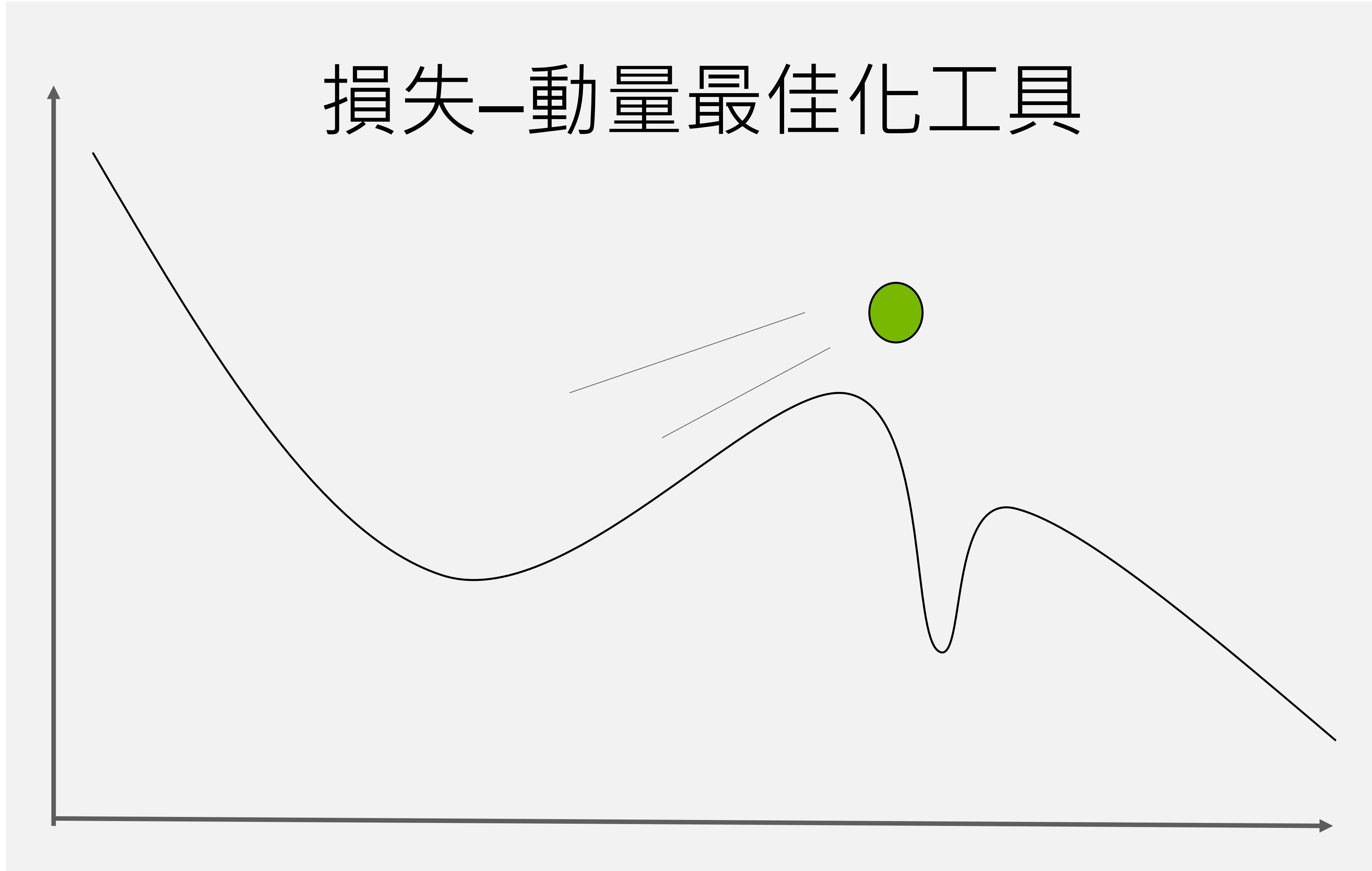
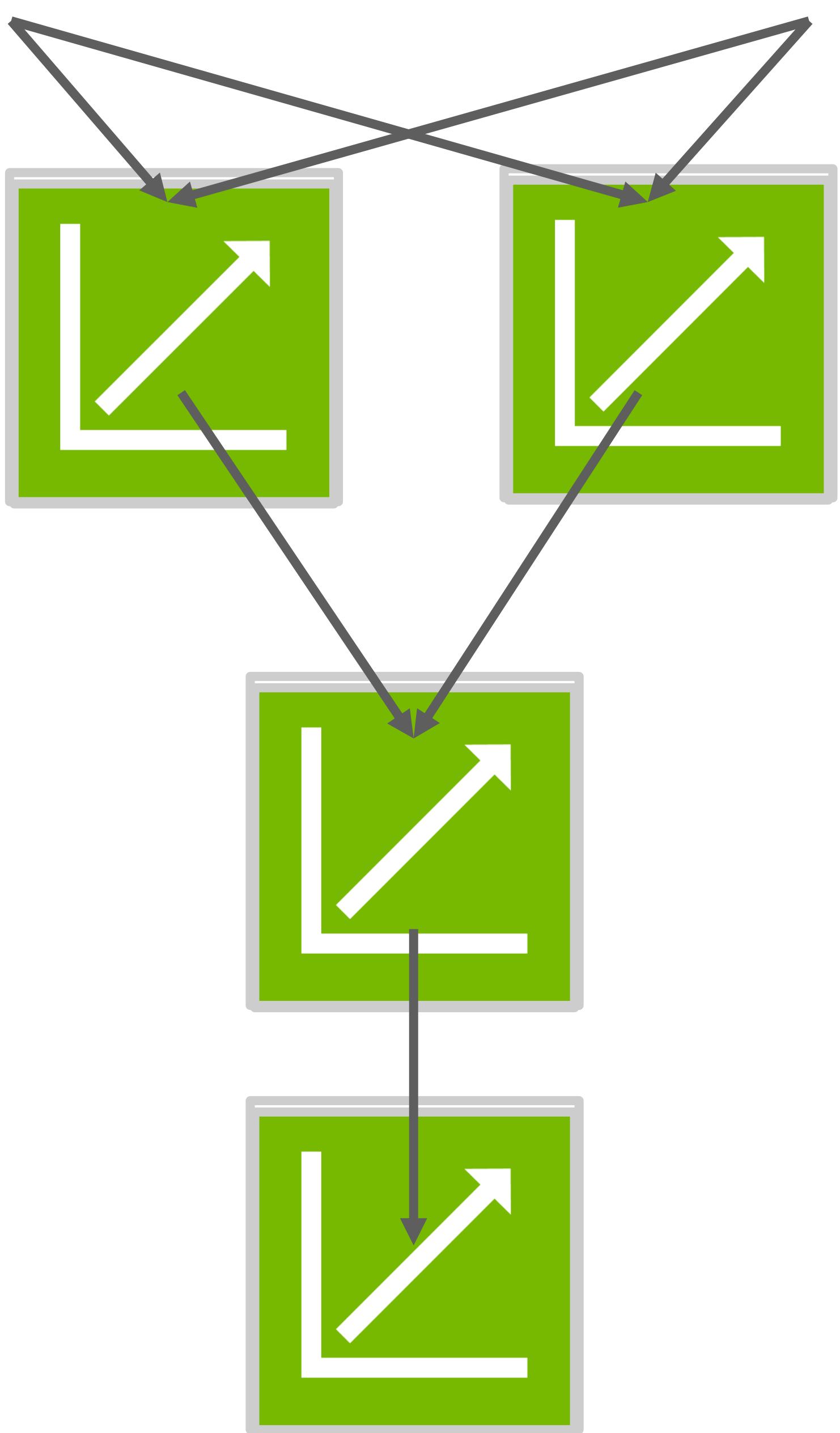


優化器 (Optimizer)



- . Adam
- . Adagrad
- . RMSprop
- . SGD

建立網路 (Network)

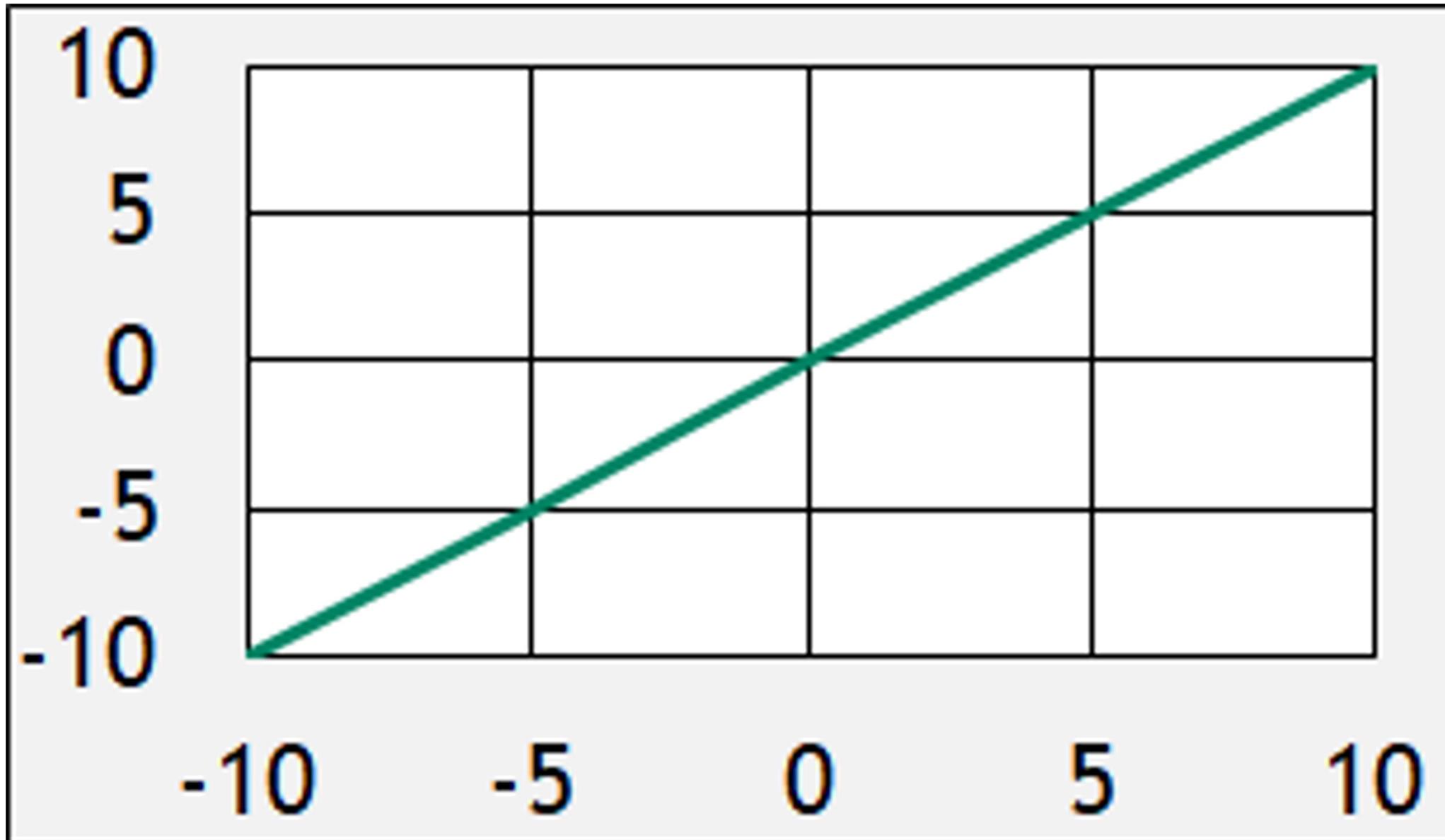


- . 擴充為更多輸入值
- . 可以連結神經元
- . 如果所有迴歸都是線性的，則輸出也會是線性迴歸

激活函數 (Activation Function)

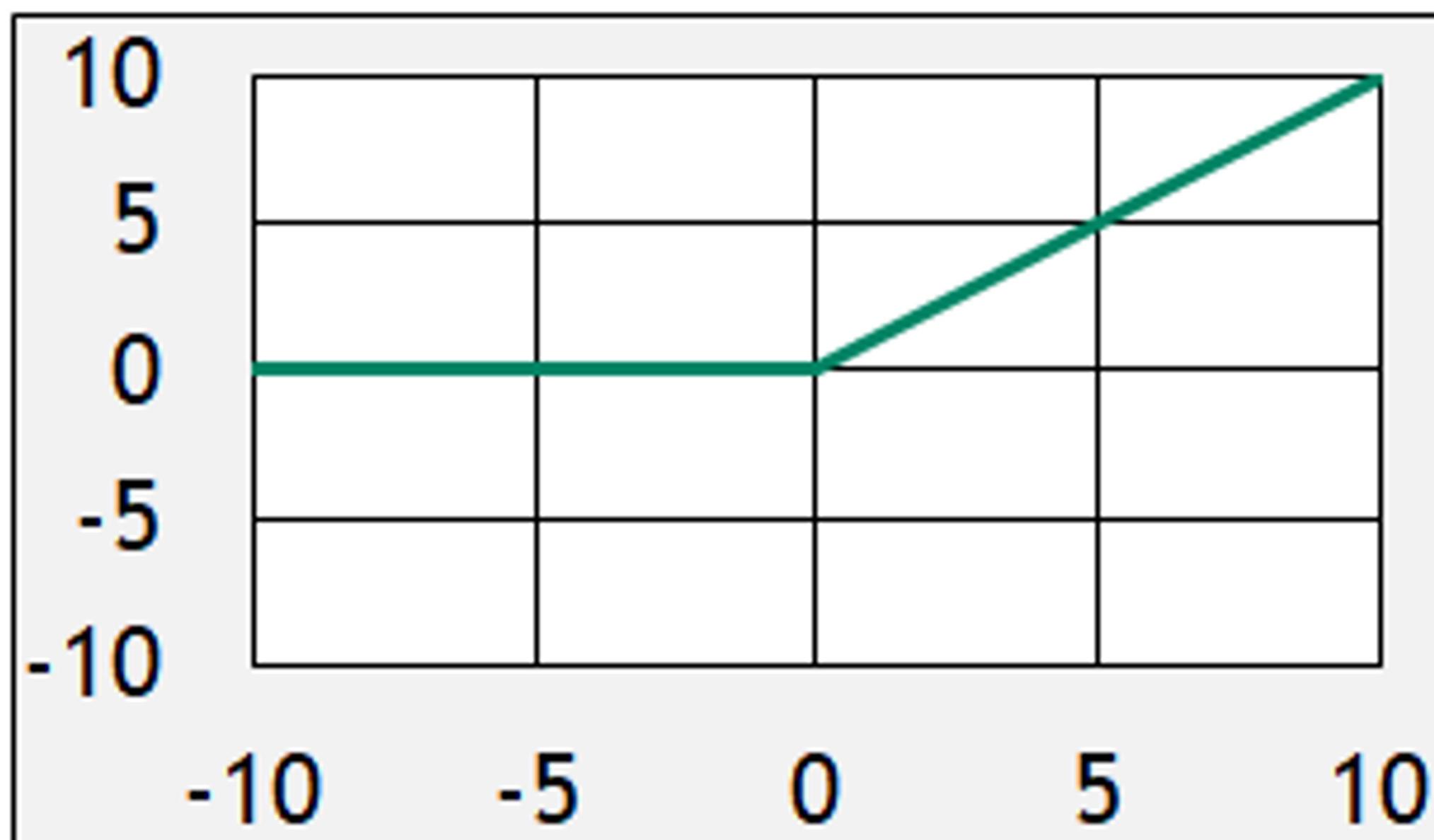
線性

```
1 # Multiply each input  
2 # with a weight (w) and  
3 # add intercept (b)  
4 y_hat = wx+b
```



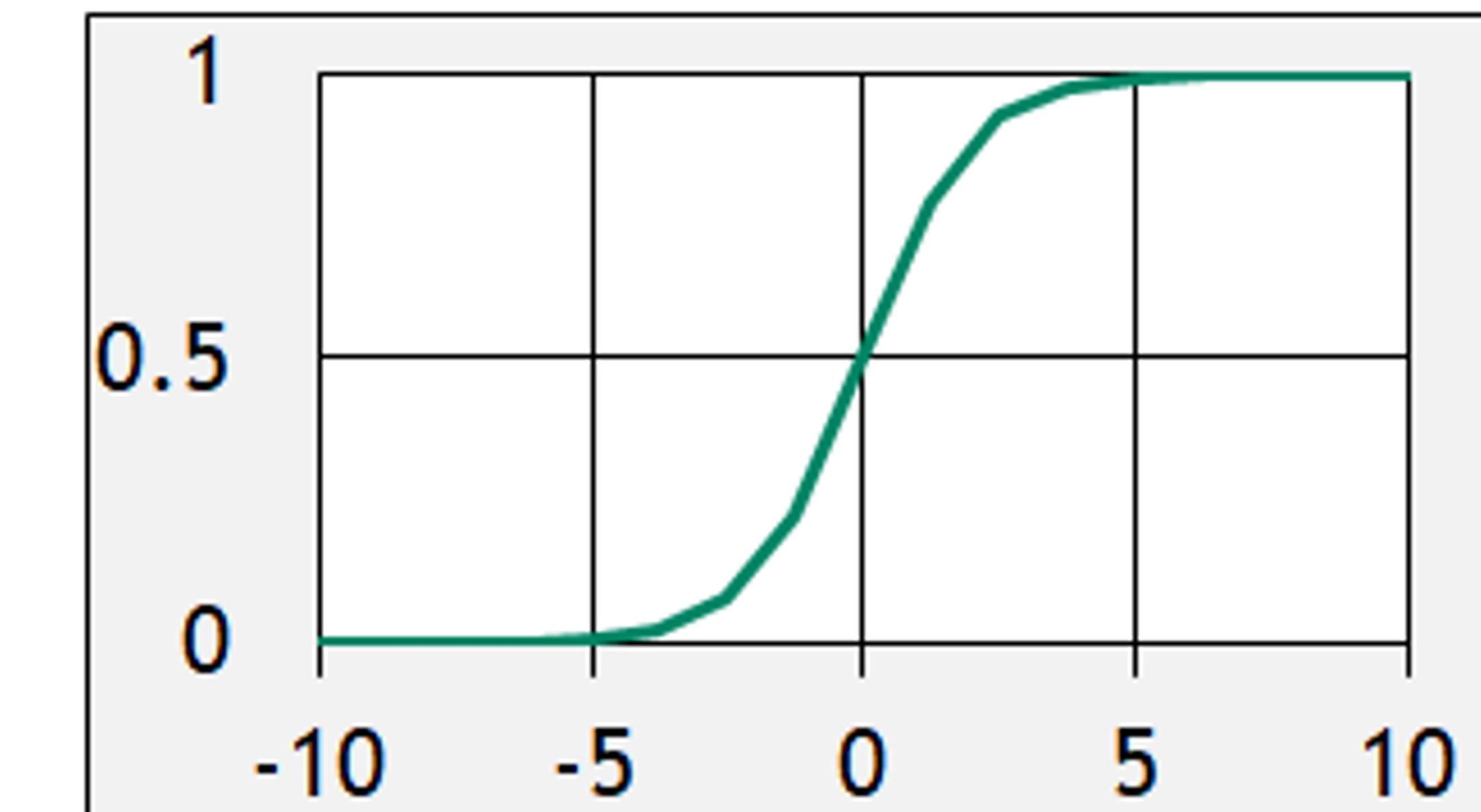
ReLU

```
1 # Only return result  
2 # if total is positive  
3 linear = wx+b  
4 y_hat = linear * (linear > 0)
```



Sigmoid

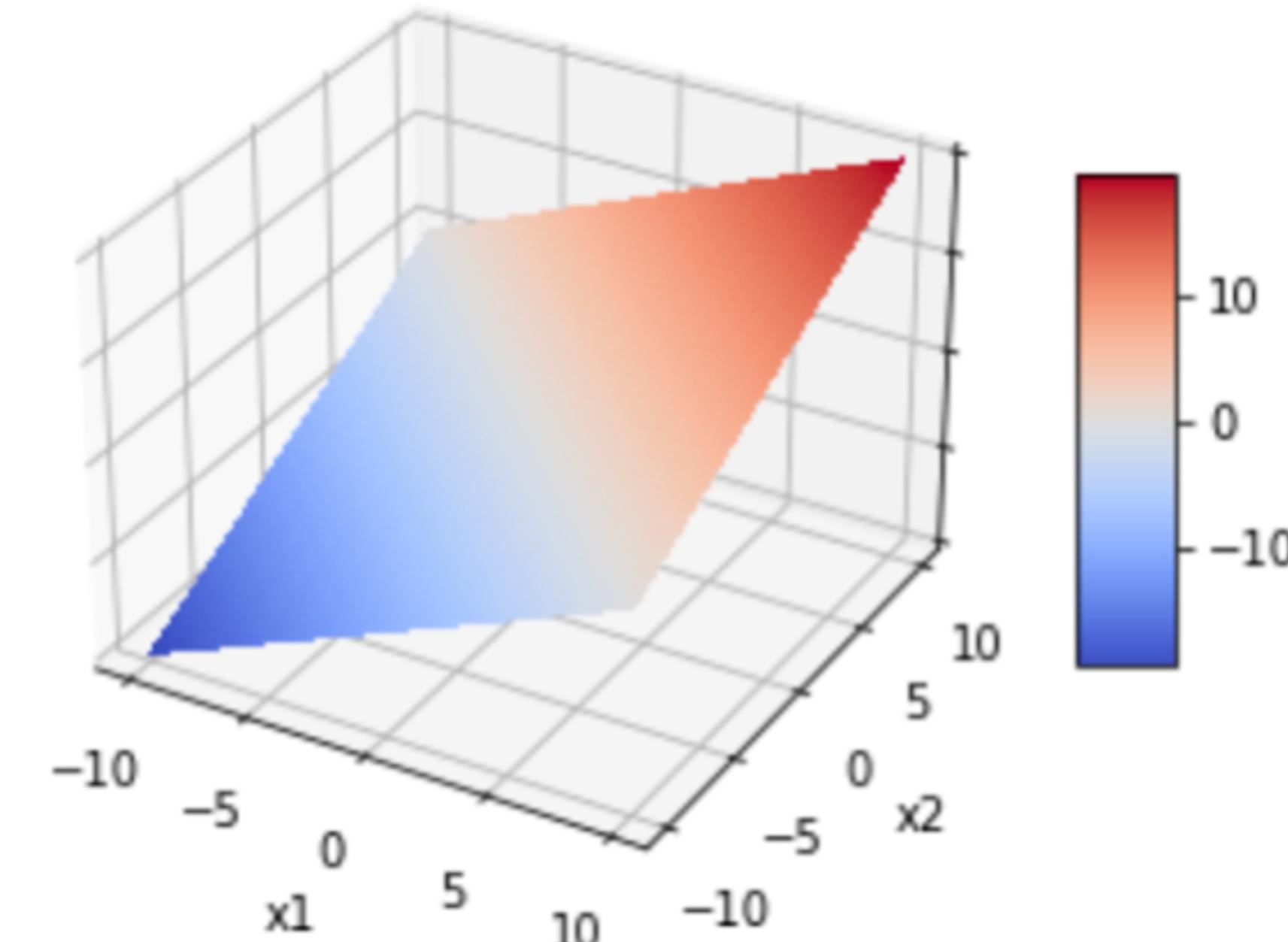
```
1 # Start with line  
2 linear = wx + b  
3 # Warp to -inf to 0  
4 inf_to_zero = np.exp(-1 * linear)  
5 # Squish to -1 to 1  
6 y_hat = 1 / (1 + inf_to_zero)
```



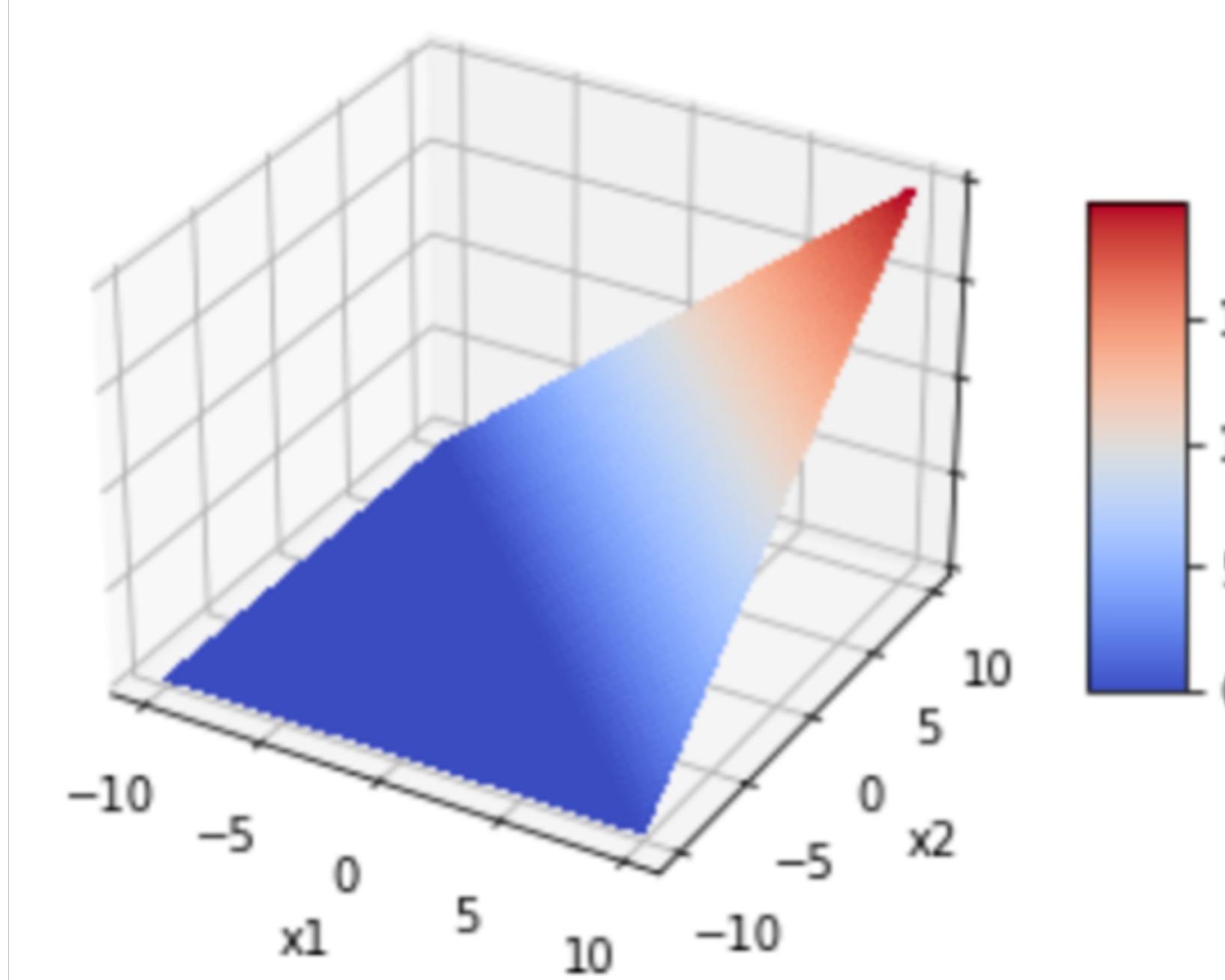
(Logistic Regression)

激活函數 (Activation Function)

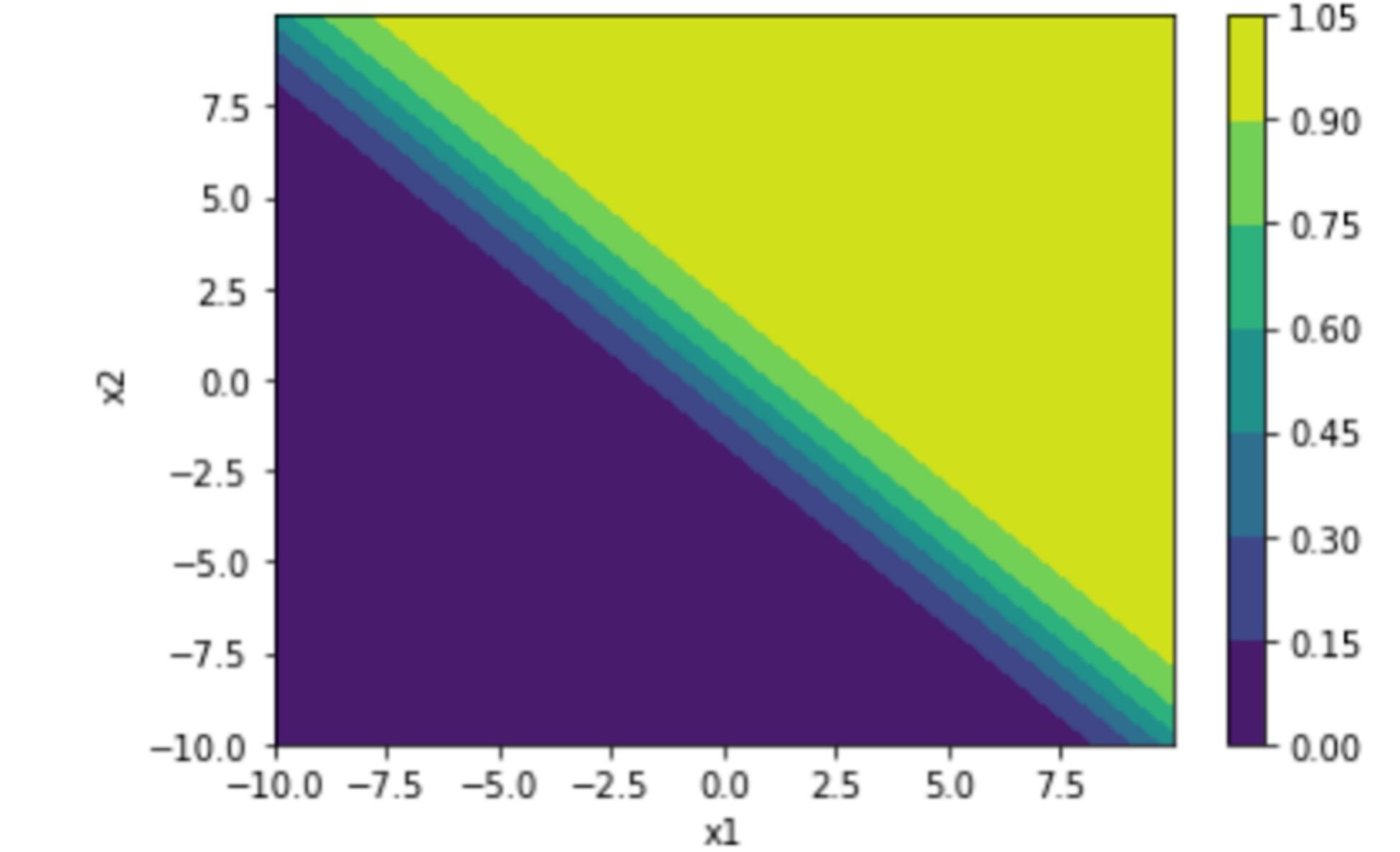
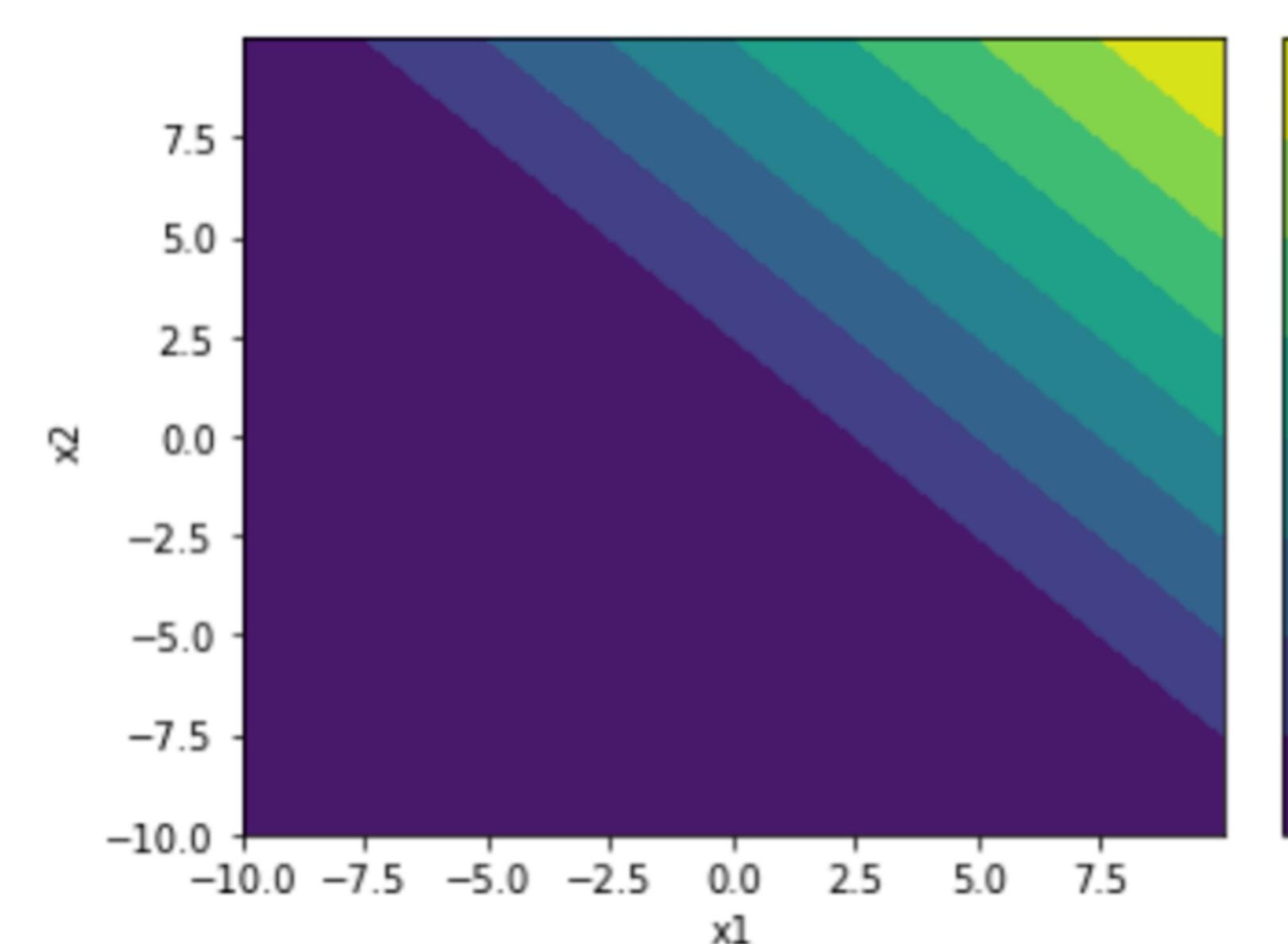
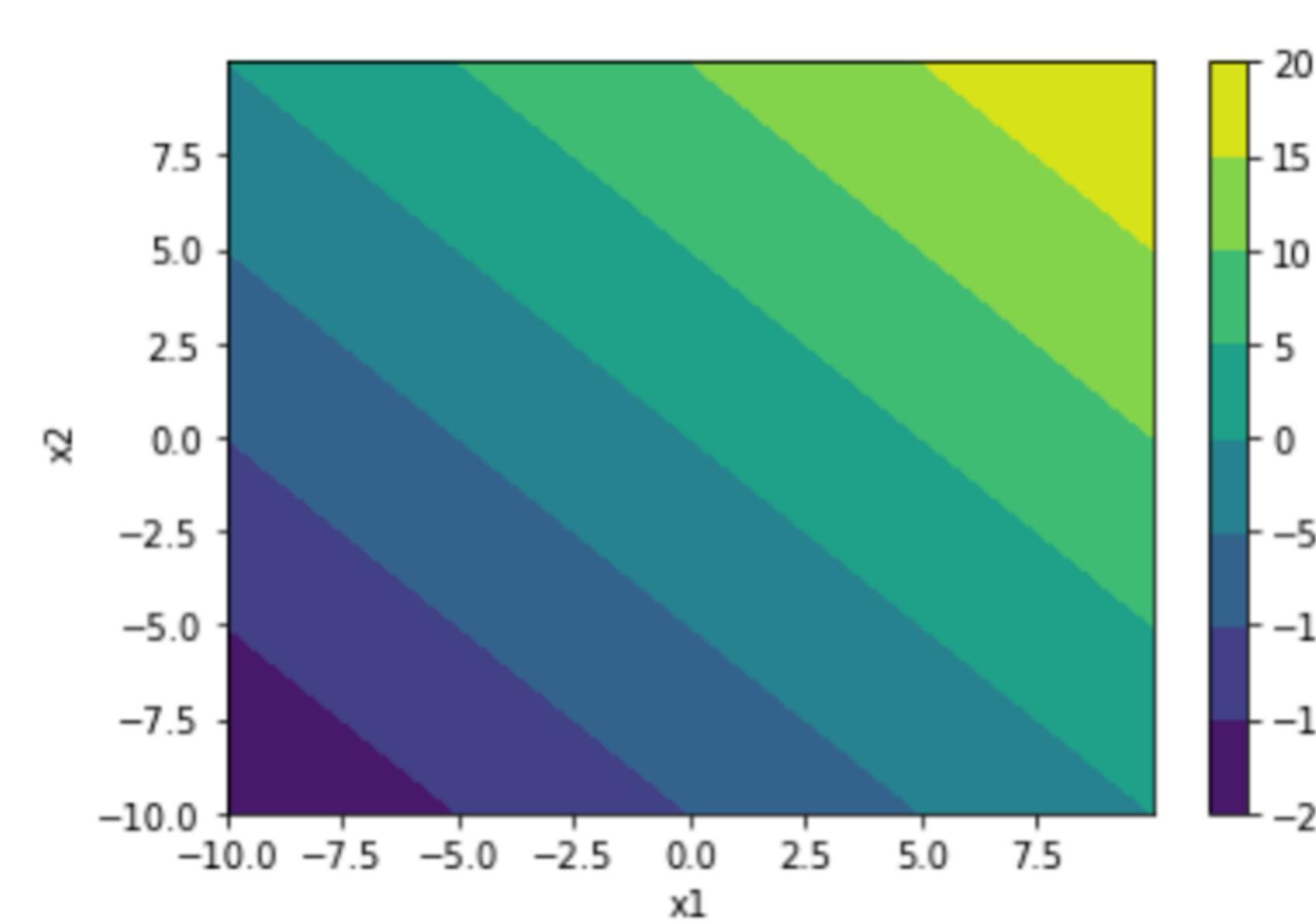
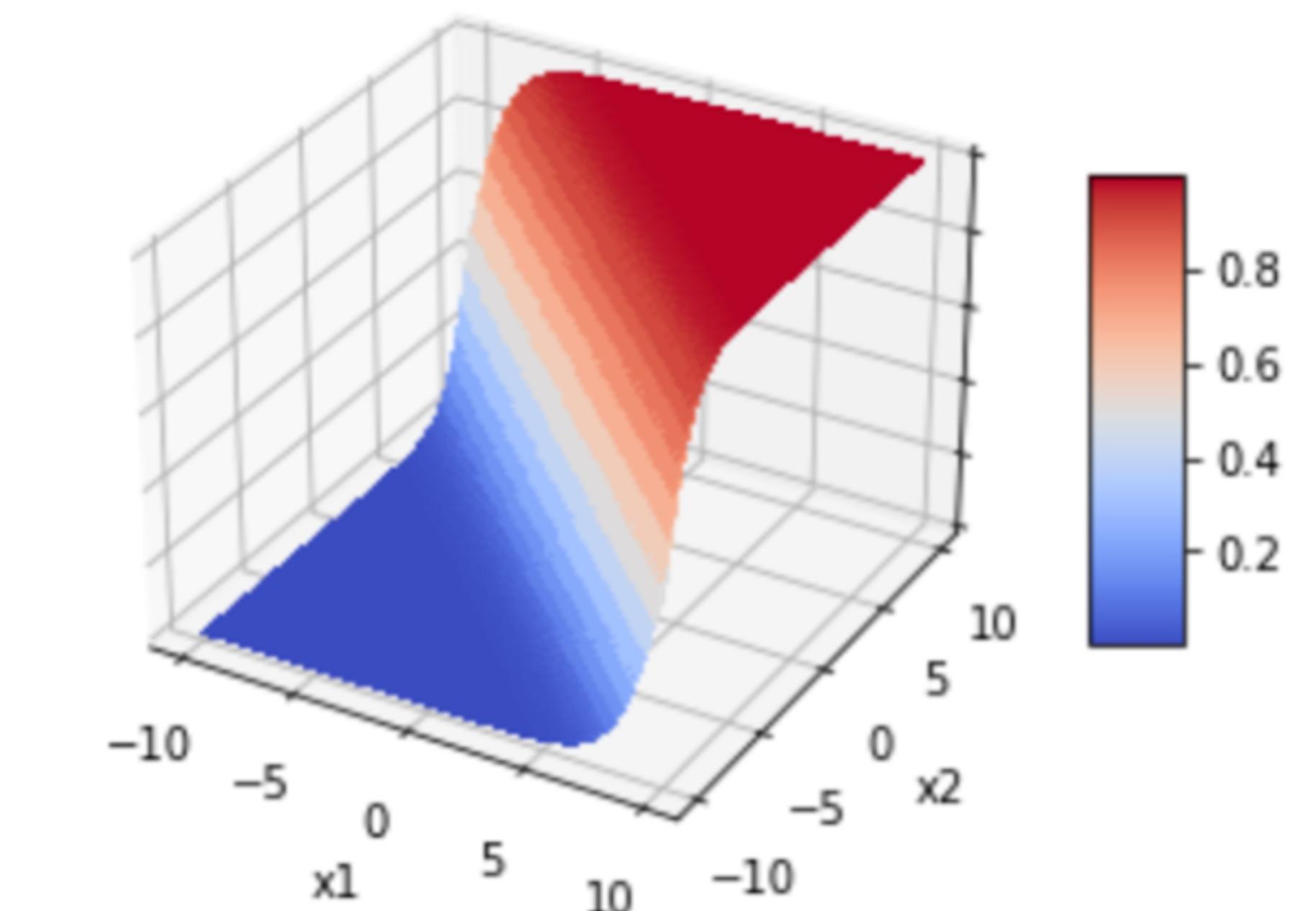
Linear



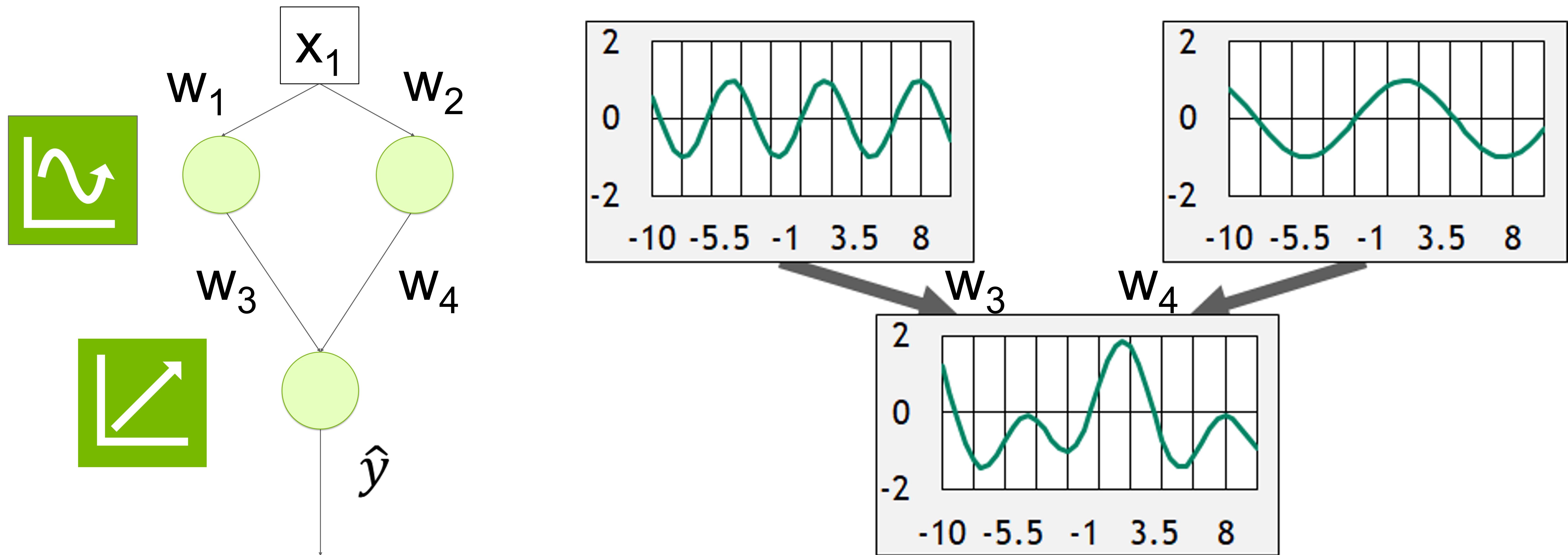
ReLU



Sigmoid

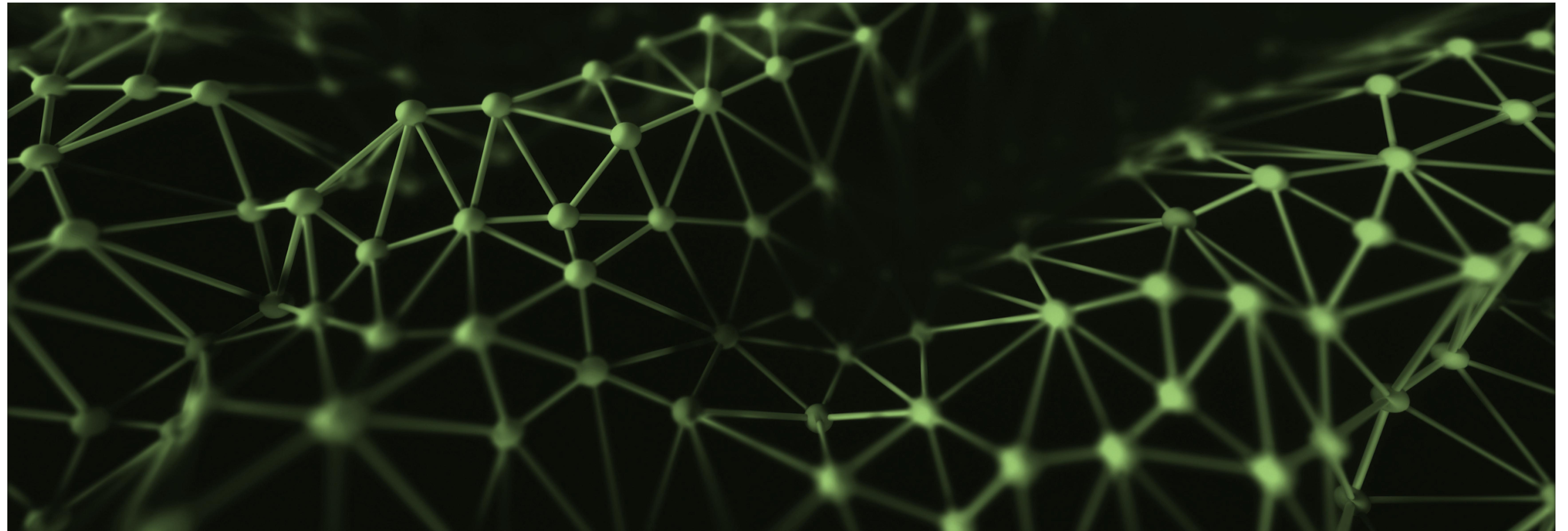


激活函數 (Activation Function)



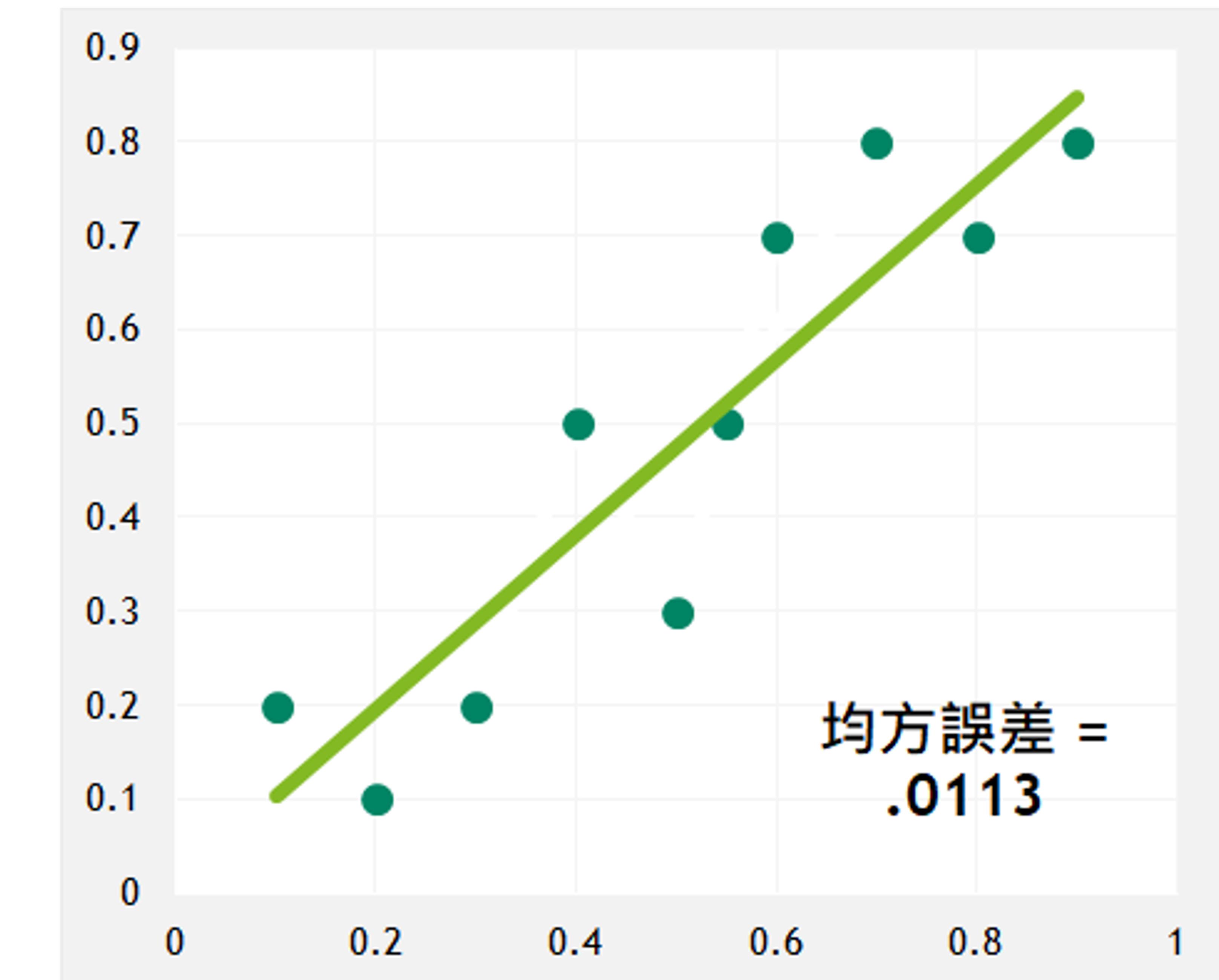
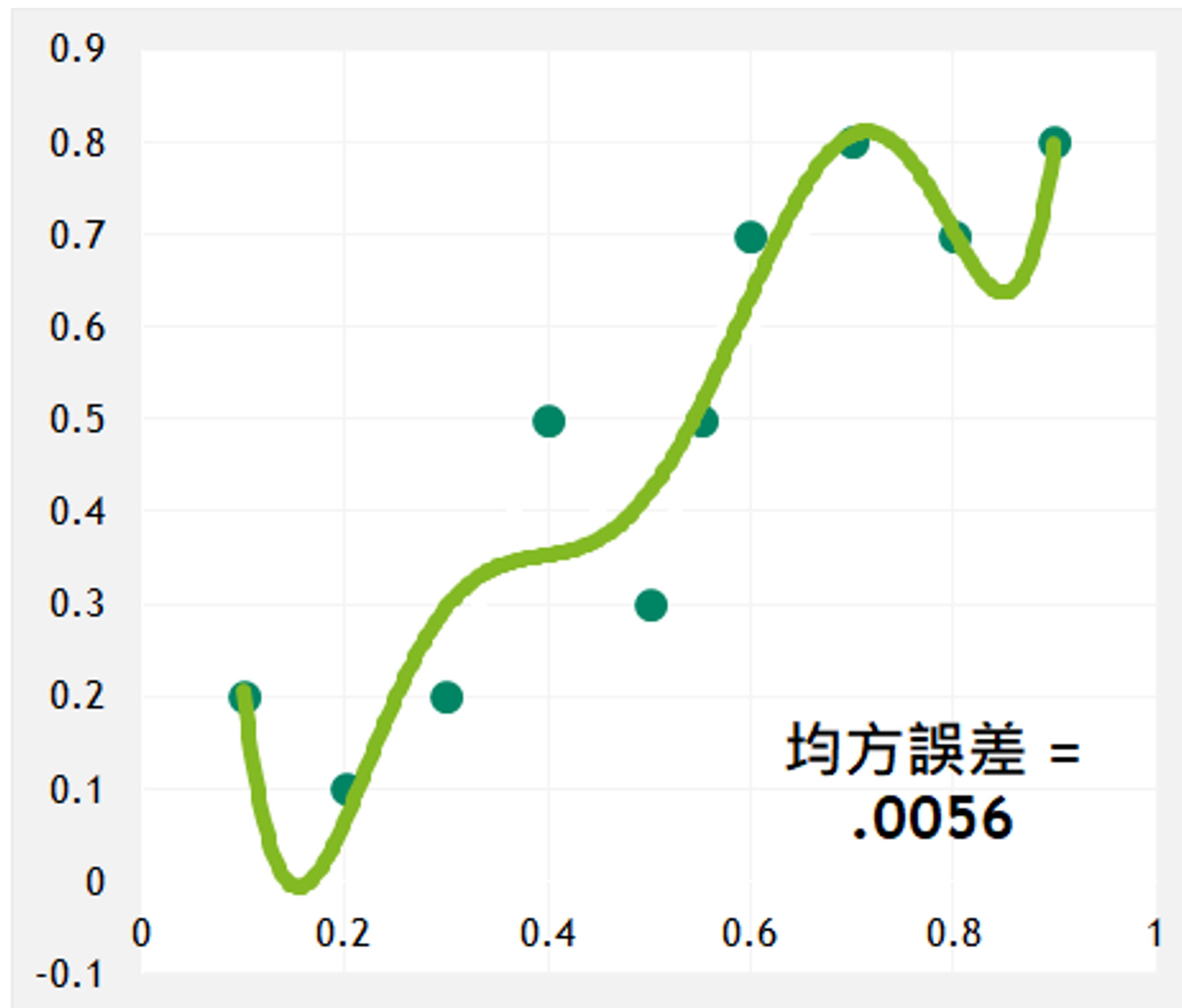
過度擬合 (Overfitting)

為何不建構一個超大型神經網路？



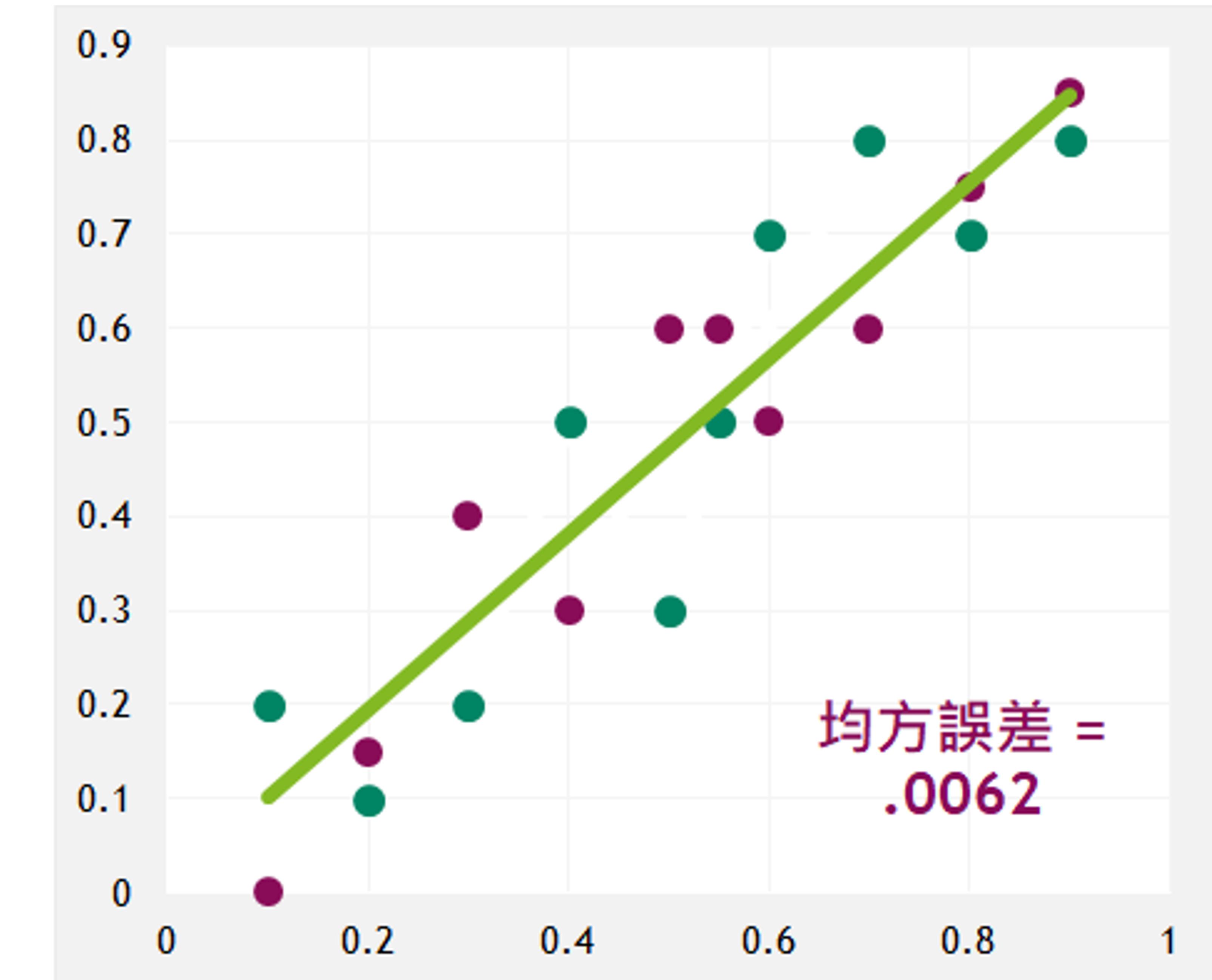
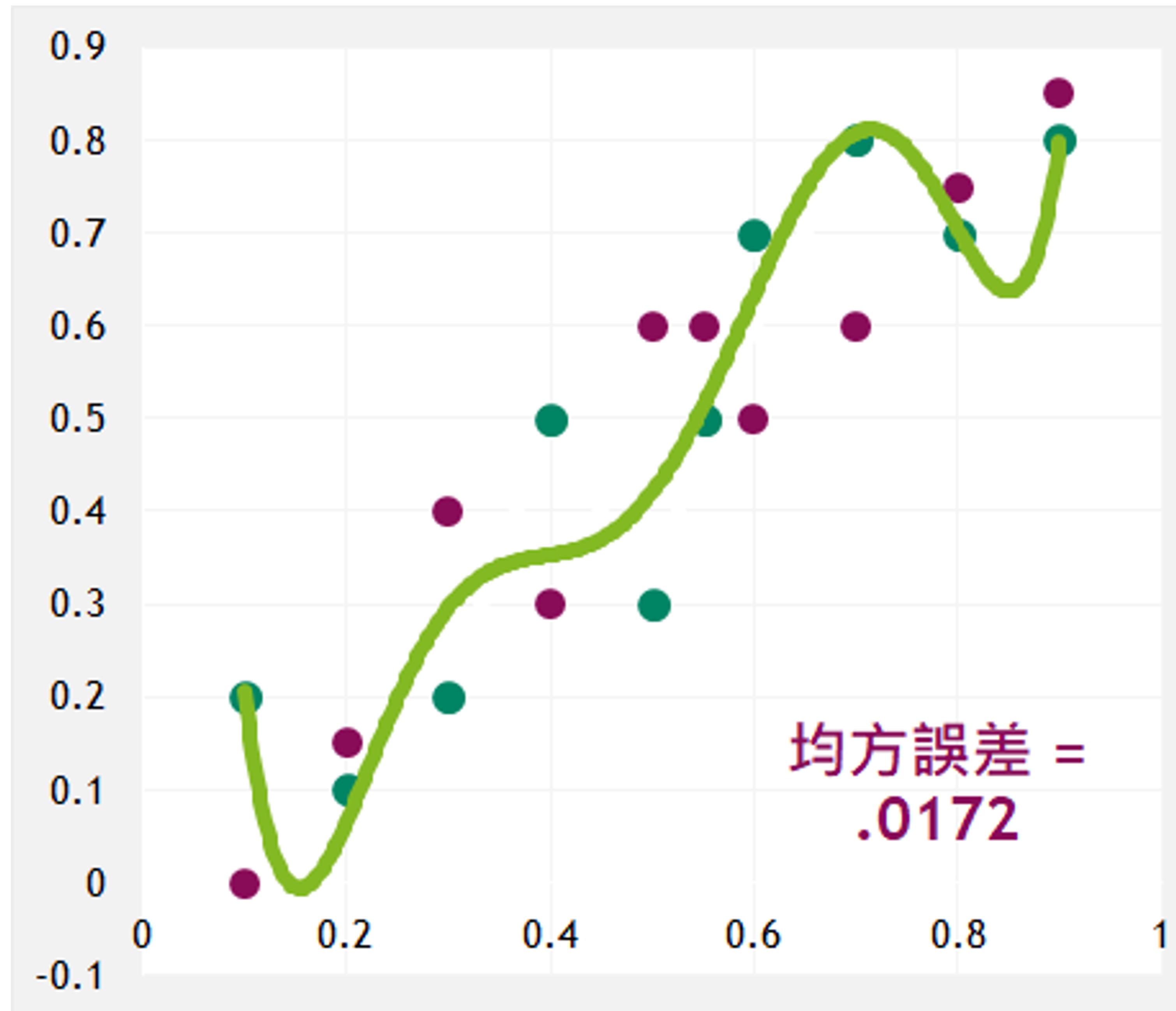
過度擬合 (Overfitting)

哪條趨勢線更理想？



過度擬合 (Overfitting)

哪條趨勢線更理想？



訓練 VS 驗證資料

避免背誦

訓練資料 (Training Dataset)

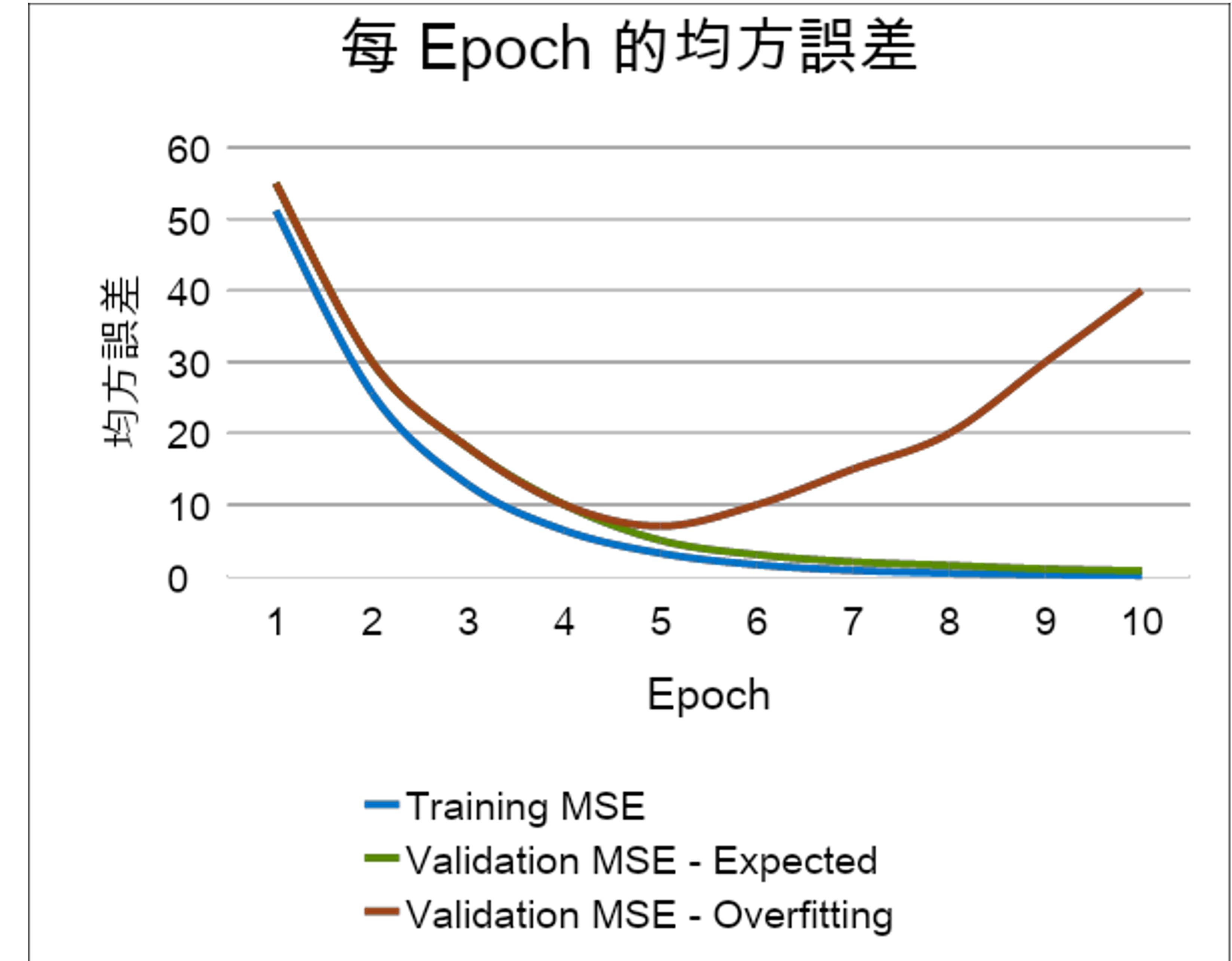
- 讓模型進行學習的核心資料集

驗證資料 (Validation Dataset)

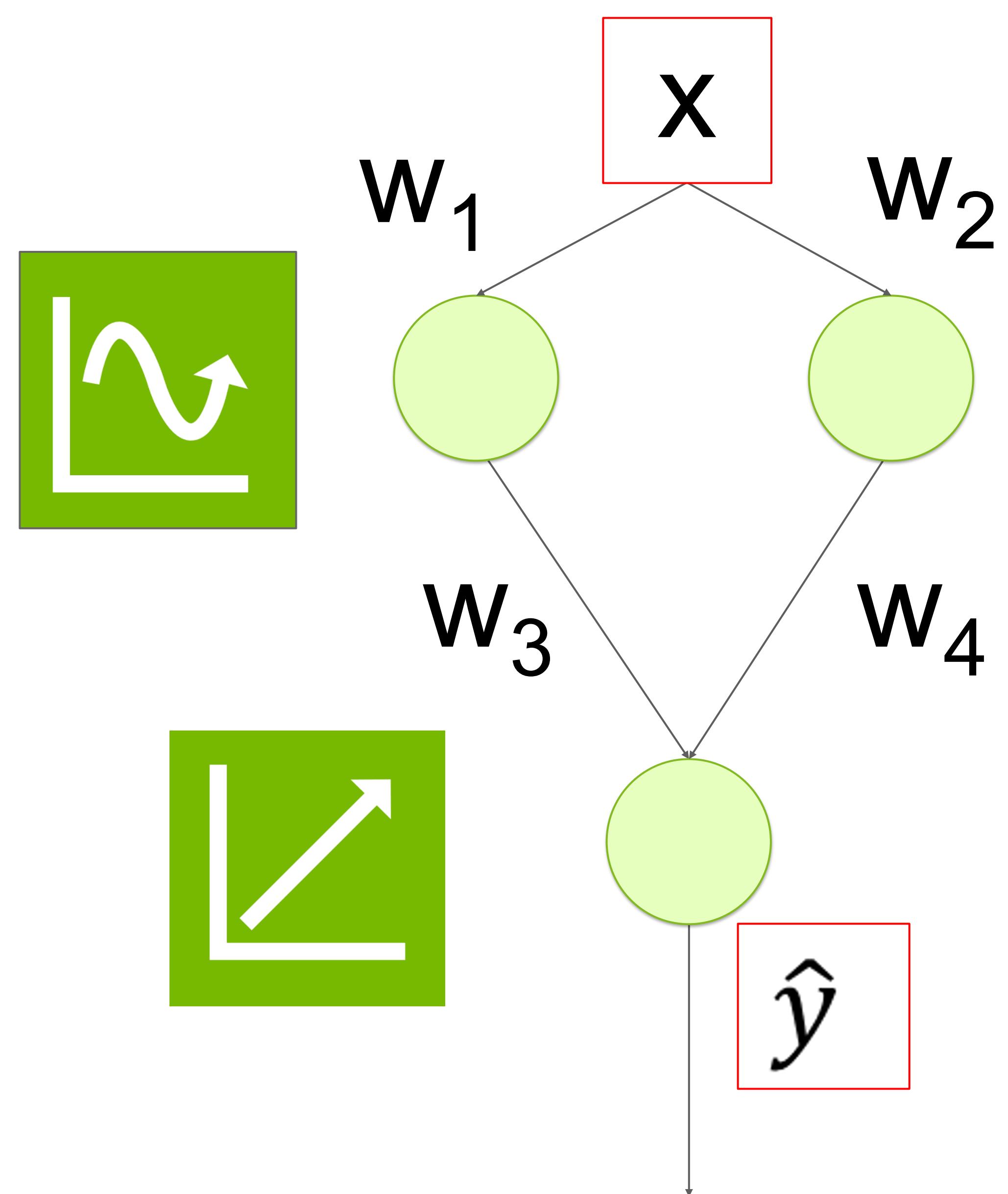
- 用來測試模型是否真的瞭解規則的新資料 (可歸納)

過度擬合 (Overfitting)

- 模型在訓練資料上效果很好，在驗證資料上卻效果不佳 (背誦的證據)
- 理想情況下，兩個資料集的精確度和損失率應該趨於接近



Recipe of Deep Learning



$$MSE = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2$$

