

機器學習hw2

B1328009 AI 二 張祐甄



多項式基底函數

低階 ($1 \leq M \leq 3$) : Underfitting

- 訓練與驗證誤差皆維持在較高水準，模型缺乏足夠的表現力去擬合資料，也顯示模型無法捕捉資料的非線性結構

中階 ($4 \leq M \leq 7$) : 誤差快速下降

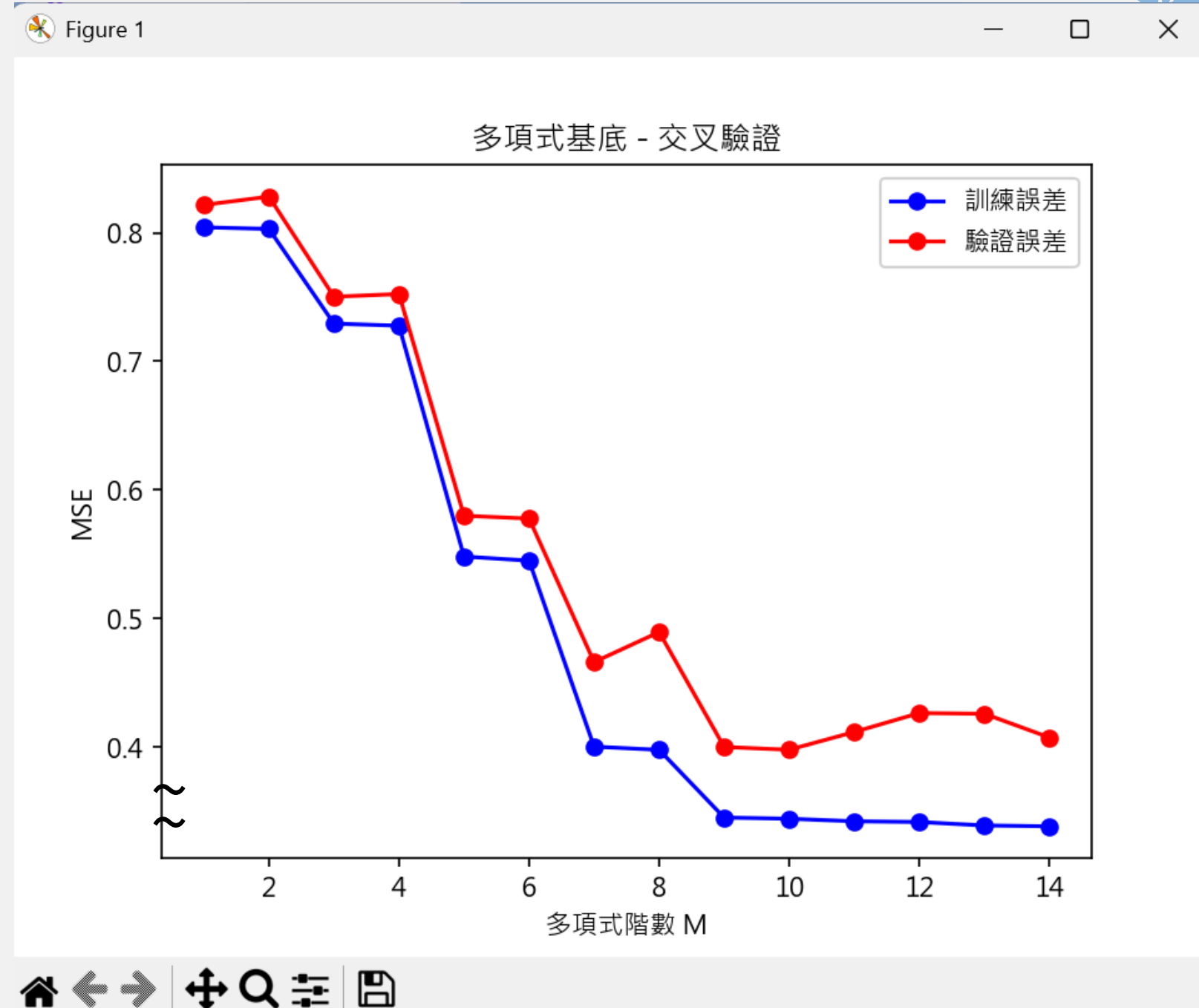
- 隨 M 增加，模型開始具備足夠的靈活度，能有效降低訓練與驗證誤差，代表模型擬合能力提升且尚未過度擬合

最佳區間 ($8 \leq M \leq 10$)

- 驗證誤差達到相對最低值，約落在 0.4 左右
- 訓練誤差與驗證誤差的差距不大，代表此時的模型複雜度與泛化能力達到平衡

高階 ($M > 10$) : Overfitting 開始出現

- 訓練誤差持續下降，但驗證誤差反而略有回升，這代表模型已經開始記住訓練資料中的噪音，導致對驗證資料的表現下降



高斯基底函數

低階 ($1 \leq M \leq 5$) : Underfitting

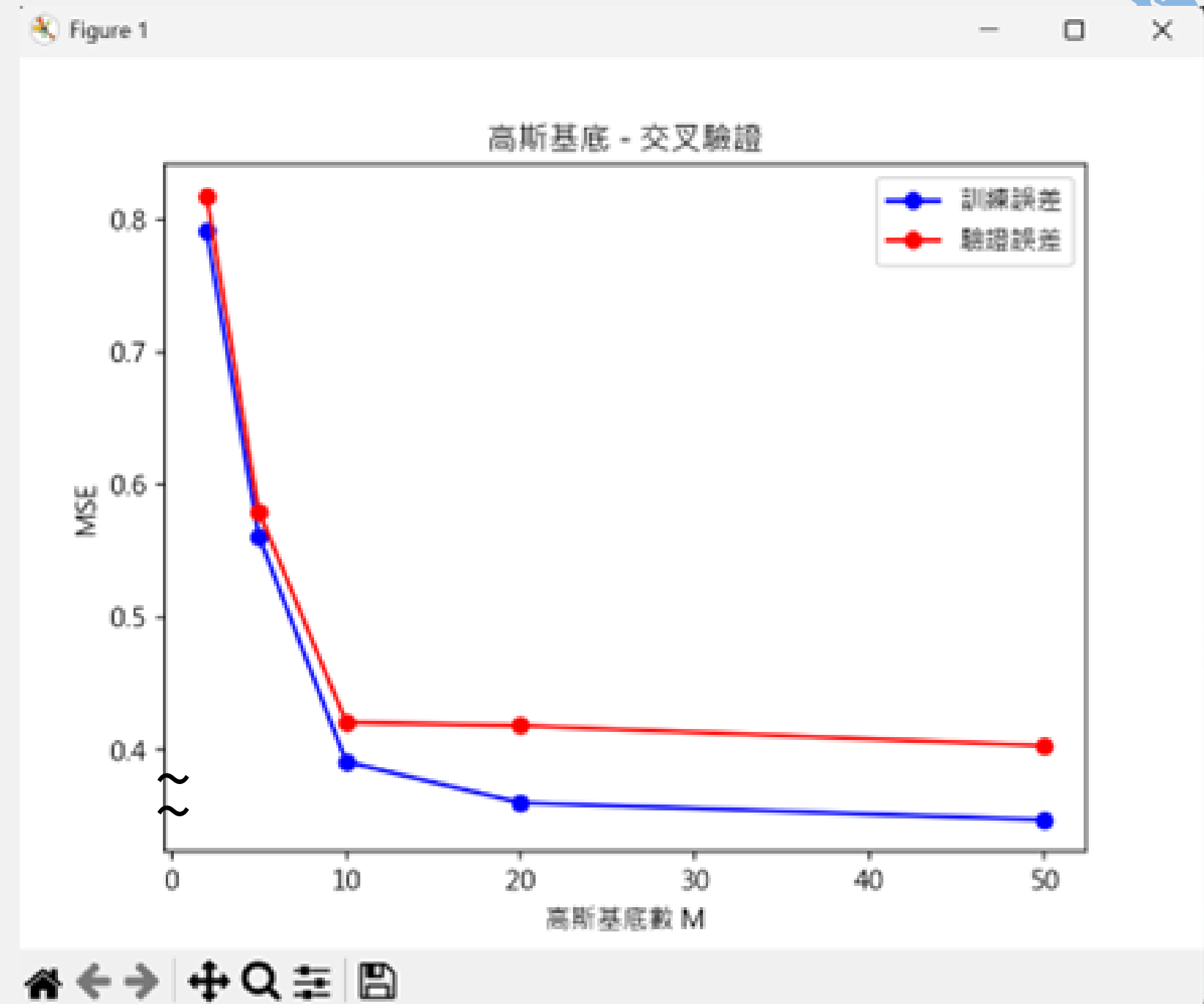
- 此時模型過於簡單，顯示高斯基底數量太少，捕捉不到非線性結構，無法有效描述資料中的變化。
- 訓練誤差與驗證誤差都偏高，約在 0.6 ~ 0.8 之間

中等 ($M \approx 10$) : 快速改善

- 當 M 增加到 10 左右時，訓練誤差與驗證誤差大幅下降
- 驗證誤差最低點約落在 0.42 左右，此時泛化效果最佳
- 模型能夠平衡擬合能力與泛化能力，是最佳超參數區間

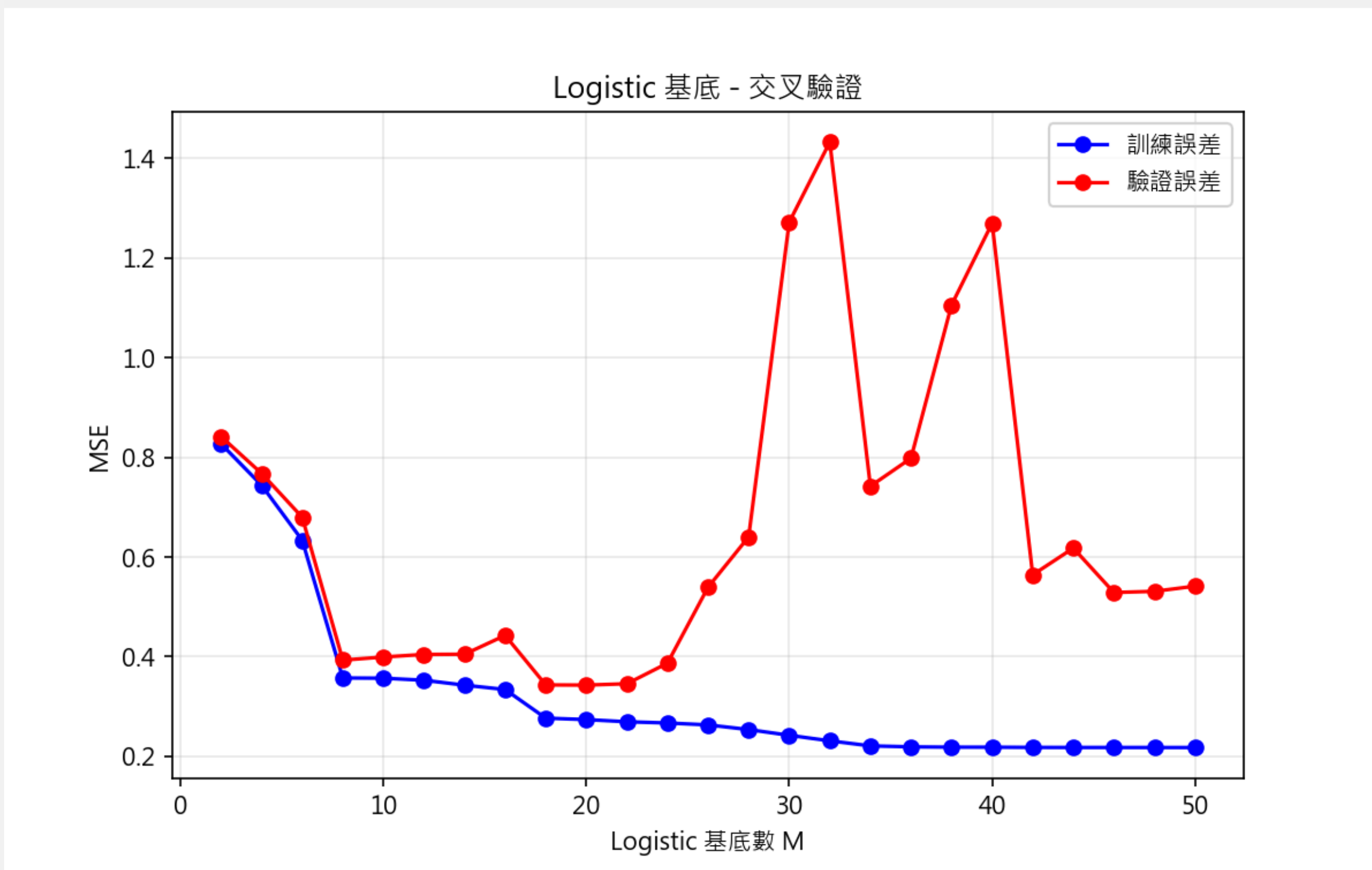
高階 ($M \geq 20$) : 出現 Overfitting 跡象

- 訓練誤差繼續下降，顯示模型幾乎記住了訓練資料
- 驗證誤差則沒有明顯改善，說明模型開始對噪音敏感，出現過度擬合現象。



logistic基底函數

Figure 1



logistic基底函數說明

低 ($1 \leq M \leq 6$) - Underfitting

- 基底太少、彈性不足，抓不到資料的非線性

中等 ($7 \leq M \leq 12$) - 最佳帶

- M 增加後 MSE 快速下降，驗證誤差在 $M \approx 10-12$ 達到最低，且訓練與驗證差距小 \rightarrow 泛化最好

過中等 ($13 \leq M \leq 20$) - 平臺/輕微過擬合

- 訓練誤差繼續降，但驗證誤差不再改善，甚至略升。

高 M (≥ 25) - 驟升與震盪 (明顯過擬合 + 數值條件不良)

- 驗證誤差在 $M \approx 28-35$ 、 $38-42$ 出現尖峰，但訓練誤差仍一路下降，這通常同時反映兩件事：

1. 容量過大：model 幾乎記住訓練資料，對看不到的驗證資料表現很差

2. 條件數爆炸 (Logistic 基底的多重共線性)：

- M 很大時，等距 centers 的 sigmoid 彼此形狀高度相關 \rightarrow 設計矩陣 Φ 的欄向量幾乎線性依賴。
- $\Phi^T \Phi$ 變得病態 (condition number 很大)，即使用 pinv，也會得到係數幅度超大的解；在不同 fold 上換一批資料，就可能暴衝。
- Logistic 還有飽和問題：遠離中心時輸出接近 0 或 1，許多欄變得近似常數，讓病態更嚴重。

比較

多項式基底

- 模型在 $M=8\sim 10$ 附近能達到最佳的泛化能力，驗證誤差最低。
- 當 M 過小時，模型表現出欠擬合；當 M 過大時，則逐漸出現過度擬合的現象。
- 這驗證了交叉驗證能有效幫助挑選合適的模型複雜度，以兼顧擬合能力與泛化能力。

高斯基底

- 模型在 $M \approx 10$ 時能達到最佳表現，訓練與驗證誤差皆顯著下降，且兩者差距最小，泛化能力良好。
- 當 M 過小，模型欠擬合，無法捕捉資料特徵；當 M 過大，則容易過度擬合，驗證誤差不再下降。
- 與多項式基底相比，高斯基底在中等複雜度下能提供更好的平滑性與泛化效果，避免了高階多項式容易出現的震盪問題。

Logistic基底

- 最佳泛化落在 $M \approx 10-12$ 。
- 當 M 再往上加，訓練誤差單調下降，但驗證誤差不再改善，且在 $M \geq 25$ 出現劇烈上升與震盪，這是過度擬合 + 數值不穩定的典型跡象。
- 與多項式/高斯相比，它需要較大的 M 才夠表現力，但也更需要良好的正則化與超參數設定來維持穩定。

(此圖為視覺上擬和最完整的結果)

三種基底函數最佳參數與過度擬合檢查

基底函數	最佳超參數 M	最佳驗證誤差 (MSE)	過度擬合檢查
多項式	10	0.398	在 M=11 之後出現過度擬合
羅吉斯	20	0.343	在 M=22 之後出現過度擬合
高斯(RBF)	50	0.34	無明顯過度擬合

在高斯基底的交叉驗證結果中，最佳超參數 M 明顯大於多項式與羅吉斯模型。這是因為高斯屬於局部基底函數，為了覆蓋整個輸入空間並捕捉細節變化，需要較多的基底中心。然而，透過固定的正則化係數 $\lambda=0.01$ ，模型在高維度下還是保持穩定，驗證誤差沒有明顯上升，顯示出高斯模型具良好的泛化能力。這也符合理論上的偏差－變異權衡：增加基底數降低偏差，而正則化項則控制變異。

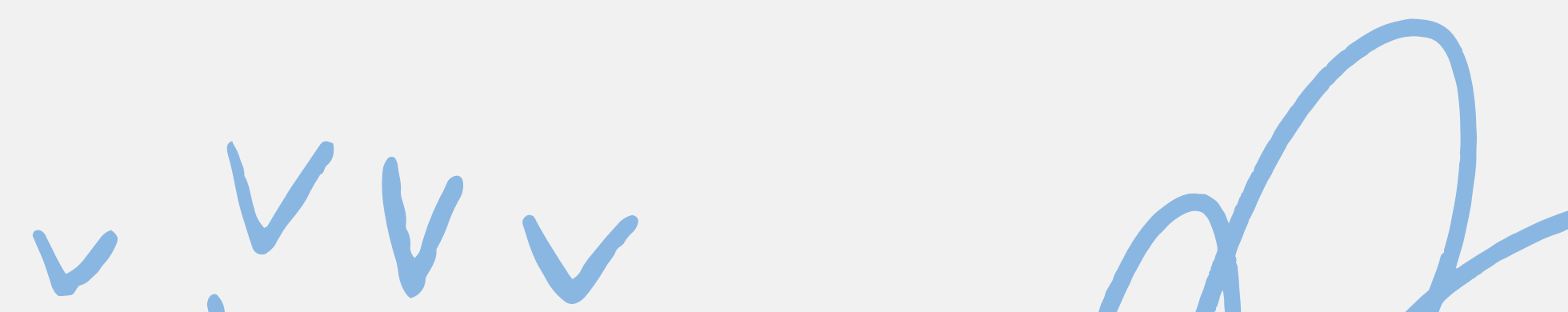


分析

在三種基底模型的交叉驗證結果中，可以觀察到三者的最佳驗證誤差（MSE）非常接近，顯示這組資料的主要趨勢能被不同基底函數有效捕捉。由於資料分佈平滑、噪音量有限，三種模型都具有足夠的表達能力以近似真實函數，因此在最小化平方誤差上達成類似的收斂結果。

這也呼應通用近似理論（Universal Approximation Theorem），即多項式展開、Sigmoid 函數組合以及高斯基底在理論上皆可於足夠基底數下逼近任意連續函數。

換句話說，誤差相近並不代表模型表現無差別，而是表示在此資料情境下，三者皆已達到充分擬合的效果，模型差異主要體現在對樣本規模與正則化的敏感度上，而非最終誤差值。



分析

MSE：在做超參數搜尋時，在驗證資料上得到的最小平均誤差，它用來決定哪個模型複雜度最會泛化
算MSE的方法：

1.將資料切成五個folds

2.對於每個候選 M ：

- 依序做 5 次：用 4 份訓練、1 份驗證，算一次驗證 MSE。
- 把 5 次的驗證 MSE 取平均 → 得到該 M 的「平均驗證誤差」

3.把所有 M 的平均驗證誤差相比較，最小的那一個就是最佳驗證誤差，對應的 M 就是「最佳超參數」。

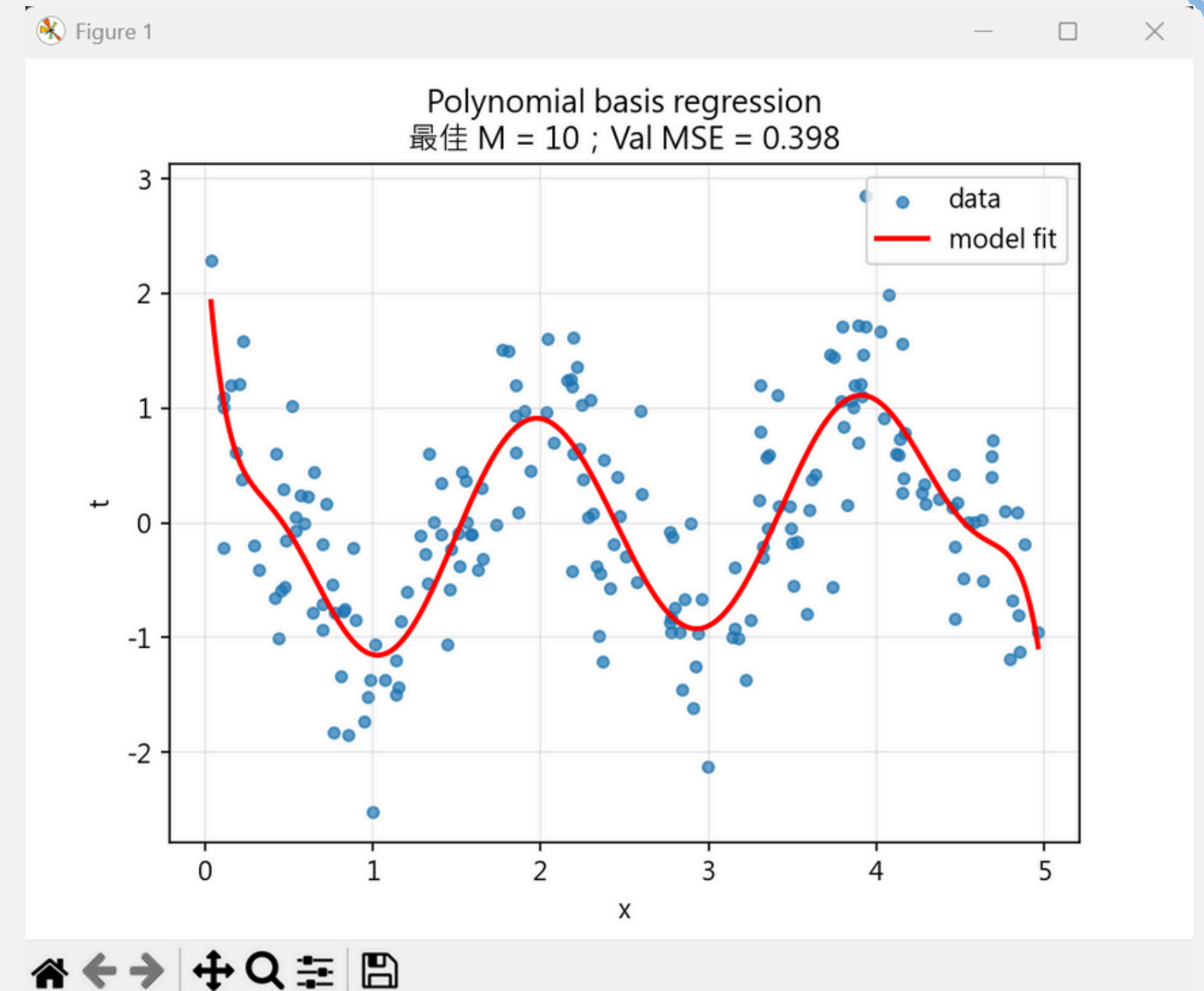
結果圖-多項式

本次實驗中，以 $M=10$ 作為最佳超參數，對資料集進行回歸建模，並得到驗證誤差 $\text{Val MSE}=0.398$ 。

紅色曲線為模型的預測結果，藍色散點為實際資料分佈。

觀察：

當多項式階數提升至 $M=10$ 時，曲線能有效捕捉三個主要波峰與波谷的變化趨勢，且整體平滑，未出現劇烈振盪。這表示模型在此階數下成功平衡了偏差與變異，達到良好的泛化能力。



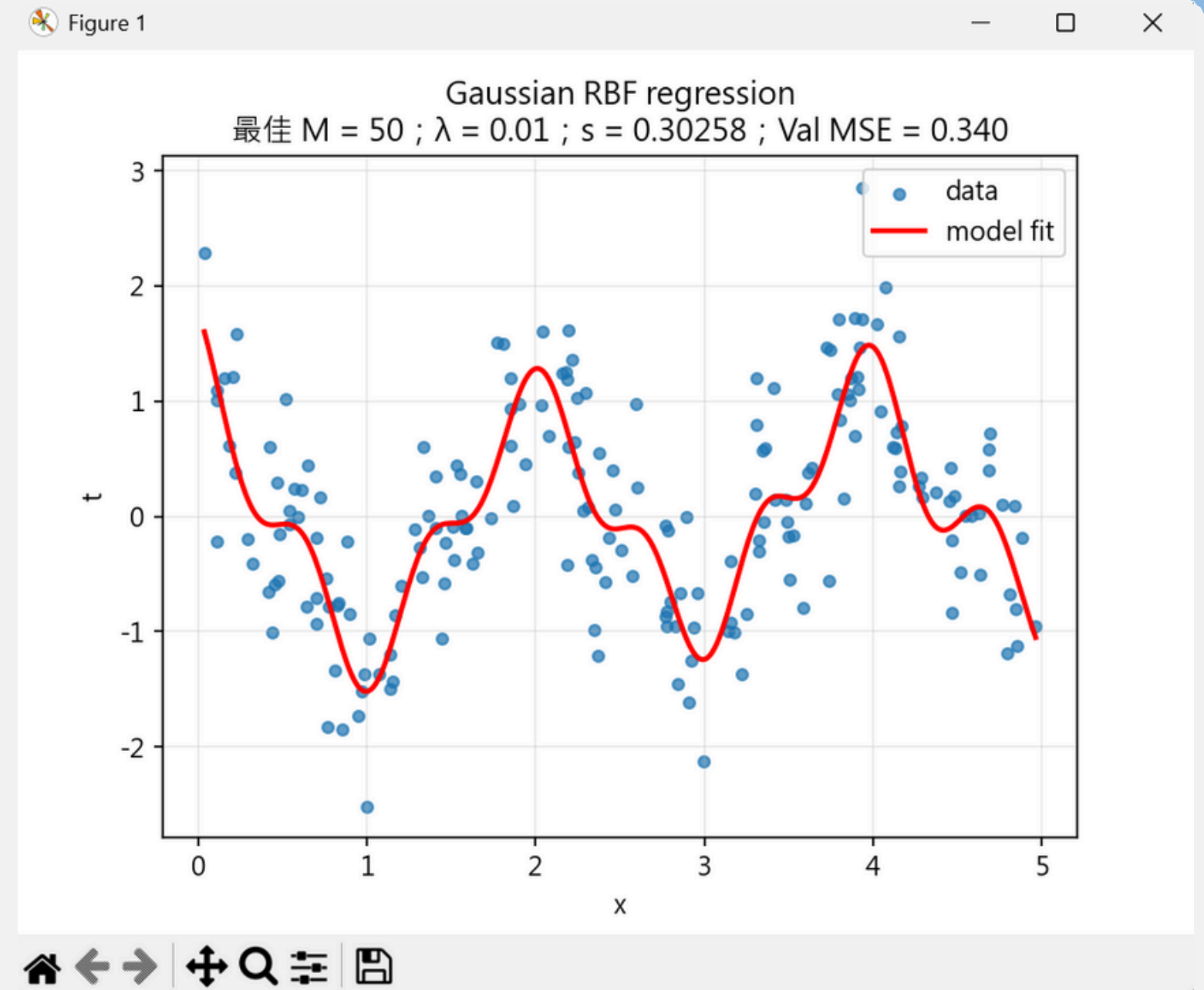
結果圖-高斯

本實驗使用高斯基底作為特徵轉換函數，並透過 5 折交叉驗證決定最佳超參數組合，得到 $M=50$, $\lambda=0.01$, $s=0.30258$, Val MSE=0.340，模型以 $M=50$ 個高斯基底進行特徵展開，並加入正則化項控制過擬合。

紅色曲線代表模型擬合結果，藍色散點為原始資料分佈。

觀察：

與多項式模型相比，高斯模型的擬合曲線在整體上更貼近資料分佈，在區域變化（ex. 局部極值點）表現更靈活，同時保持曲線的整體連續性與穩定性。



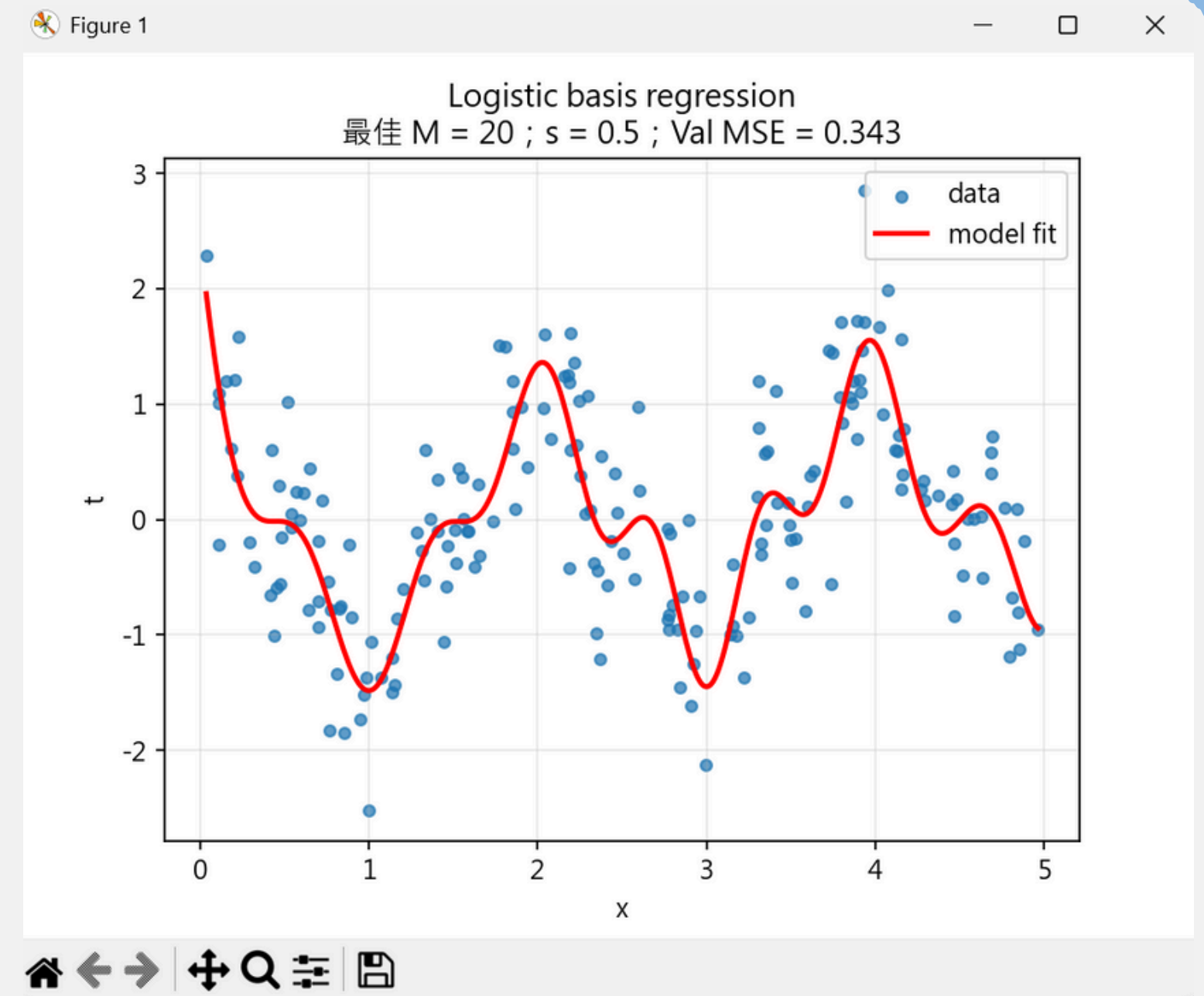
結果圖-羅吉斯

本實驗採用羅吉斯函數作為基底，以非線性轉換方式將輸入資料映射至較高維的特徵空間。使用 5 折交叉驗證進行模型評估，得到 $M=20, s=0.5, \text{Val MSE}=0.343$ 。

紅色曲線代表模型擬合結果，藍色散點為原始資料分佈。

觀察：

與多項式模型相比，羅吉斯具有半局部性的特質，每個基底在其中心區域內有明顯作用，但在遠端則逐漸趨於飽和，因此能避免多項式在邊界區域出現震盪。



The background is a light gray color, decorated with various hand-drawn blue doodles. These include several loops and swirls at the top, a star-like shape on the right, a wavy line at the bottom center, and several checkmarks at the bottom right. The central text is in a bold, black, sans-serif font with a white drop shadow.

**Thank you
very much!**