一、目標函數數學特性

擬合對象為**Runge函數**：

$$f(x) = \frac{1}{1 + 25x^2}$$

- **解析特性**：在
  $$x \in [-1, 1]$$
  區間內呈現中央平緩、邊緣陡峭的形態
- **逼近難點**：
  - 高曲率區域：當
    $$|x| > 0.6$$
    時，二階導數
    $$f''(x)$$
    急劇增大
  - 邊界震盪：傳統多項式插值在端點處會出現Runge現象

二、神經網路逼近原理

1. 通用近似定理 (Universal Approximation Theorem)

- **核心主張**：單隱藏層前饋網路，在隱藏層寬度足夠時，能以任意精度逼近緊緻集上的連續函數
- **數學表達**：

  $$f(x) \approx W_2 \cdot \sigma(W_1 x + b_1) + b_2$$

  其中：
  - $$W_1, W_2$$
    為權重矩陣
  - $$\sigma$$
    為非線性激活函數（此處使用ReLU）

2. 深度堆疊效應

程式碼採用**3層架構**：

$$\hat{y} = W_3 \cdot \mathrm{ReLU}(W_2 \cdot \mathrm{ReLU}(W_1 x + b_1) + b_2) + b_3$$

- **層級分解**：
  - 第1層：線性映射
    $$\mathbb{R}^1 \to \mathbb{R}^{20}$$
  - 第2層：特徵空間的非線性變換
  - 第3層：回歸到輸出空間

三、誤差最小化機制

1. 損失函數

採用**均方誤差(MSE)**：

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

- **特性**：對大誤差敏感，適合平滑函數逼近
- **梯度計算**：通過反向傳播自動微分

2. 優化過程

使用**Adam優化器**更新參數：

$$\theta_{t+1} = \theta_t - \eta \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

其中：

- $\hat{m}_t$：偏差修正後的一階動量
- $\hat{v}_t$：偏差修正後的二階動量
- $\eta = 0.01$ 為學習率

四、數值穩定性分析

1. 激活函數選擇

- **ReLU函數**：

$$\mathrm{ReLU}(z) = \max(0, z)$$

  - **優點**：緩解梯度消失問題
  - **缺點**：在負區間產生"死神經元"，影響邊緣區域擬合

2. 誤差分布特性

實驗結果顯示：

- **MSE**：2.56×10^{-4}
- **最大誤差**：4.53×10^{-2}（出現在 $x \approx \pm 0.9$ 處）

*本文件使用AI工具輔助生成