

ProgrammingAssignment2_111652026劉馥瑞

1. 方法

利用一個多層感知器 (MLP) 類神經網路同時逼近 Runge 函數 ($f(x)=1+25x^{21}$) 和其一階導數 ($f'(x)=-(1+25x^2)250x$)。與僅使用函數值進行訓練的傳統方法不同，本方法採用一個創新的複合損失函數，其目標是同時最小化函數預測誤差和導數預測誤差。

神經網路架構為一個帶有兩個隱藏層的 MLP，每個隱藏層包含 50 個神經元，並使用 Tanh 為激活函數。這使得網路能夠有效地學習非線性關係。

損失函數定義為兩個均方誤差 (MSE) 損失之和：

1. 函數損失 (L_f): 預測函數值 $\hat{f}(x)$ 與真實函數值 $f(x)$ 之間的 MSE。
2. 導數損失 ($L_{f'}$): 預測導數值 $\hat{f}'(x)$ 與真實導數值 $f'(x)$ 之間的 MSE。

總損失為 $L_{total}=L_f+L_{f'}$ 。為了計算 $\hat{f}'(x)$ ，我們利用了 PyTorch 的自動微分功能 (`torch.autograd.grad`)，這使我們能夠精確且高效地計算模型輸出對其輸入的梯度。

模型使用 Adam 優化器進行訓練，並採用早停 (Early Stopping) 機制。如果驗證損失在連續 100 個訓練週期內沒有改善，訓練將提前停止，並載入最佳模型狀態以防止過度擬合。

2. 結果 (生成資料與圖檔都放在資料夾中)

該模型在訓練了 387 個週期後，由於驗證損失停滯而觸發了早停。

A. 訓練過程與最終誤差

訓練過程中，損失值從高點迅速下降，並最終收斂到非常小的值。

- **Epoch 1:** `train_loss=2.027369`, `val_loss=1.478172`
- **Epoch 200:** `train_loss=0.000006`, `val_loss=0.000007`
- 早停: 在 Epoch 387 發生

在獨立的測試集上，模型的最終性能評估如下：

- 函數 **MSE**: 1.088179×10^{-4}
- 函數最大絕對誤差: 1.655868×10^{-2}
- 導數 **MSE**: 3.489636×10^{-4}
- 導數最大絕對誤差: 7.878969×10^{-2}

B. 函數與導數預測圖

以下圖表展示了神經網路在 $[-1, 1]$ 區間內的預測結果。模型對 Runge 函數本身的預測與真實函數曲線完美吻合，並且對其一階導數的預測也非常精確。這證明了複合損失函數方法的有效性。

C. 訓練與驗證損失曲線

總體訓練損失和驗證損失曲線如圖所示。兩條曲線都穩定下降並最終收斂，表明模型正在有效地學習而沒有出現過度擬合的跡象。儘管在後期損失曲線出現小幅波動，但早停機制成功地在驗證損失達到平穩後停止了訓練。

註：本次coding與報告有使用chatgpt作為輔助工具。