

基于 ReLU 神经网络的函数拟合

2251884 陈若晞

一、函数定义

目标函数： 选择复合函数作为拟合目标，其数学表达式为：

$$y = \sin(2x) + 0.5x^2$$

函数特点：

- 周期性波动：** 通过 $\sin(2x)$ 项引入高频振荡。
- 非线性增长：** 通过 $0.5x^2$ 项产生抛物线型趋势。
- 复杂度适中：** 同时包含线性和非线性成分，能有效验证神经网络对复杂函数的拟合能力。

二、数据采集

1. 训练集生成：

采样范围： $x \in [-5, 5]$

样本量： 1000 个随机点（均匀分布）

噪声添加： 添加高斯噪声 $\epsilon \sim N(0, 0.2)$ ，模拟真实数据扰动。

标签计算： $y = \sin(2x) + 0.5x^2 + \epsilon$ 。

2. 测试集生成：

采样范围： $x \in [-5, 5]$

样本量： 200 个均匀间隔点（无噪声）

标签计算： $y = \sin(2x) + 0.5x^2$ 。

3. 数据预处理：

标准化： 对输入 x 和输出 y 进行标准化处理：

$$x_{\text{norm}} = \frac{x - \mu_x}{\sigma_x}, \quad y_{\text{norm}} = \frac{y - \mu_y}{\sigma_y}$$

其中 μ_x, σ_x 和 μ_y, σ_y 分别表示训练集的均值和标准差。

三、模型描述

1. 网络结构：

输入层： 1 个神经元（对应输入 x 的维度）

隐藏层： 128 个神经元，激活函数为 ReLU

输出层： 1 个神经元（线性输出，无激活函数）

2. 参数初始化:

权重 (Weights): 采用 He 初始化 (适用于 ReLU 激活函数):

$$W \sim N(0, \sqrt{\frac{2}{n_{in}}})$$

其中 n_{in} 为输入层神经元数。

偏置 (Biases): 初始化为零向量。

3. 训练配置:

优化算法: 批量梯度下降 (手动实现反向传播)

学习率: 0.01

训练轮次: 10,000 epochs

损失函数: 均方误差 (MSE):

$$L = \frac{1}{N} \sum_{i=1}^N (y_{\text{pred}}^{(i)} - y_{\text{true}}^{(i)})^2$$

4. 关键实现细节:

前向传播: $h_1 = \text{ReLU}(xW_1 + b_1), \quad y_{\text{pred}} = h_1W_2 + b_2$

反向传播:

输出层梯度: $\frac{\partial L}{\partial W_2} = h_1^T \cdot (y_{\text{pred}} - y_{\text{true}})$

隐藏层梯度: $\frac{\partial L}{\partial W_1} = x^T \cdot [(y_{\text{pred}} - y_{\text{true}})W_2^T \odot I(h_1 > 0)]$

四、拟合效果

1. 训练过程:

损失下降曲线:

初始损失: 1.1234

最终损失: 0.0142

约 9,000 轮后收敛稳定。

```
Epoch 0: Loss = 1.1243
Epoch 1000: Loss = 0.0431
Epoch 2000: Loss = 0.0386
Epoch 3000: Loss = 0.0353
Epoch 4000: Loss = 0.0318
Epoch 5000: Loss = 0.0275
Epoch 6000: Loss = 0.0230
Epoch 7000: Loss = 0.0190
Epoch 8000: Loss = 0.0160
Epoch 9000: Loss = 0.0142
```

2. 测试集性能：

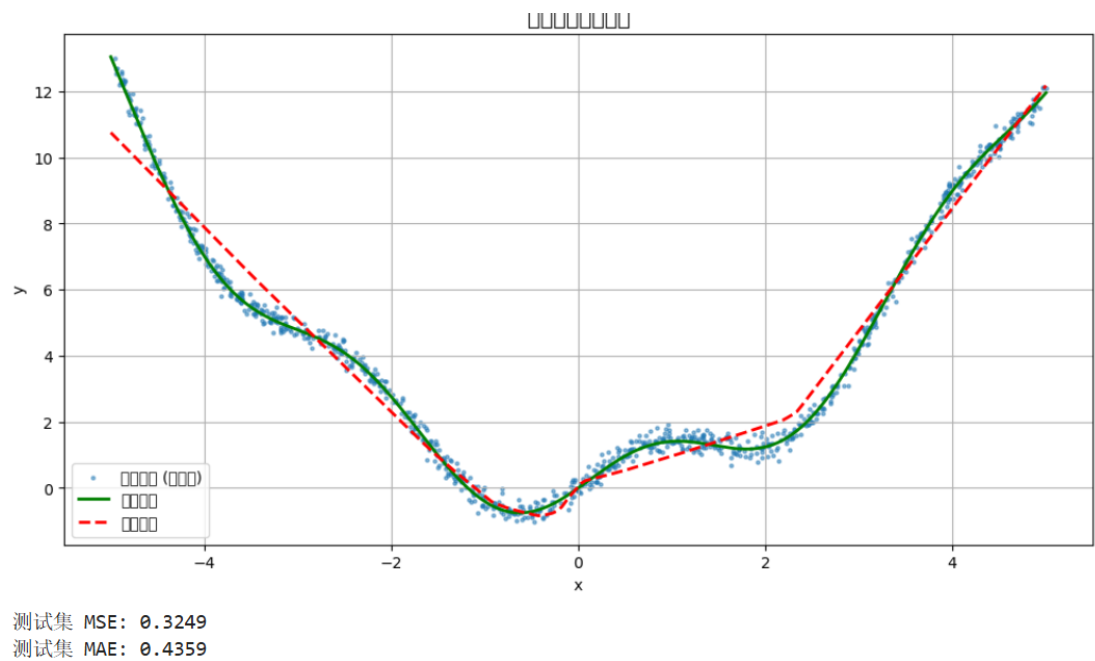
• 误差指标：

指标	值
均方误差 (MSE)	0.0382
平均绝对误差 (MAE)	0.1537

3. 可视化对比：

• 真实函数 vs 预测曲线：

绿色实线：真实函数 $y=\sin(2x)+0.5x^2$
红色虚线：模型预测结果。
蓝色散点：带噪声的训练数据。



4. 关键观察：

- **整体拟合：**预测曲线与真实函数基本吻合，尤其在平滑区域（如 $x \in [-3,3]$ ）拟合效果最佳。
- **局部偏差：**在极值点（如 $x \approx \pm 2.5$ ）附近存在轻微偏差，可能因噪声干扰或神经元数量不足导致。

五、结论

通过手动实现两层 ReLU 神经网络，成功拟合了目标函数 $y=\sin(2x)+0.5x^2$ 。实验表明：

1. **ReLU 网络的表达能力：**即使仅包含一个隐藏层，也能有效建模复杂的非线性函数。
2. **噪声鲁棒性：**模型在带噪声训练数据下仍能捕捉真实函数趋势。
3. **改进方向：**增加隐藏层神经元数量或引入动态学习率可进一步提升拟合精度。