

# 大作业 1 实验报告

---

舒意恒 计算机科学与技术系 MF20330067

## 设计方案

---

模型为一个输入神经元数量为 2，输出神经元数量为 1 的神经网络。其中包含 3 层隐藏层，隐藏层使用 LeakyReLU 激活函数，每层隐藏层神经元数量为 100。

由于拟合的函数值范围在  $[-2, 2]$ ，输出层未使用激活函数。模型使用 MSE 作为损失函数。

非线性拟合能力：LeakyReLU 是一个非线性激活函数，具有 ReLU 函数方便求梯度的特性，同时在  $x < 0$  的情况下也能得到非零梯度。在多个隐藏层中应用能展现出一定的非线性拟合能力。由于本题拟合函数值的范围，不应使用 Sigmoid、Tanh 等激活函数作为输出层的激活函数。实验发现，在使用这些激活函数作为输出层激活函数时，使用各种超参数都容易使模型 loss 陷入波动，难以收敛。

数据生成后无需 shuffle 即可开始训练。

## 编程实现

---

训练过程使用 AdamW 作为优化器，它可以自适应调整学习率。训练样本点为 10000，batch\_size 大小被设定为 25，即每个 epoch 中 batch 的数量是 400。

- 运行环境见 README.md
- 主函数位于 main.py 中，其中包含模型的定义，前向传播和反向传播的实现
- activate\_func.py 是激活函数的实现
- dataset.py 数据相关
- loss\_func.py 是损失函数的实现
- optim.py 是优化器的实现

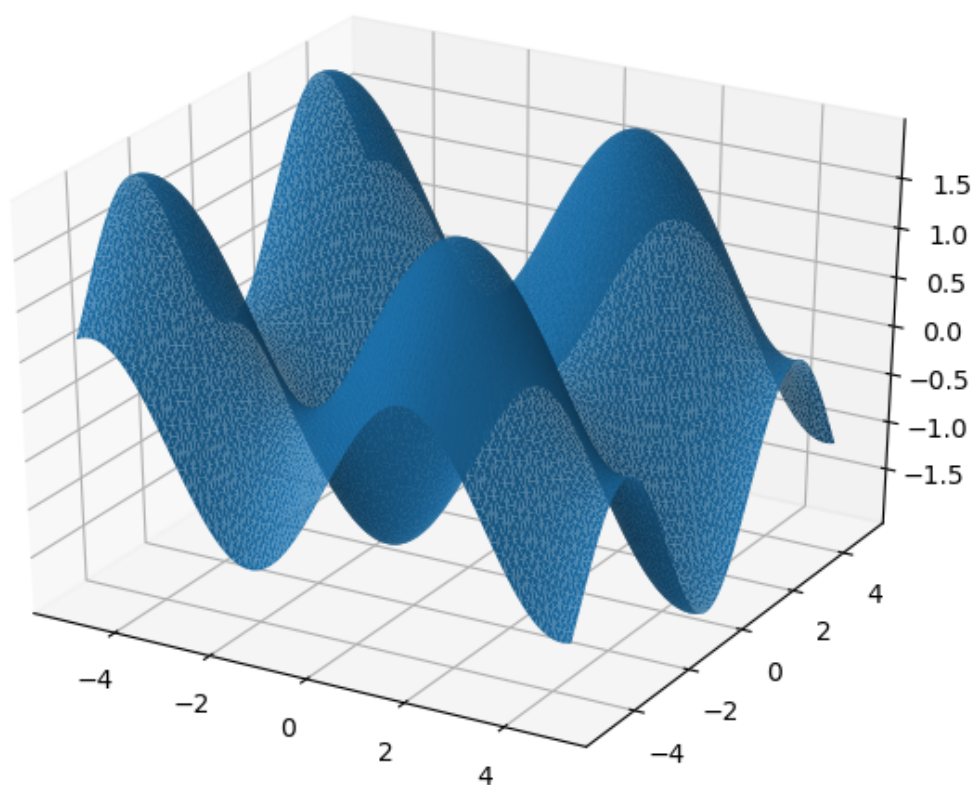
程序使用 numpy 执行矩阵运算，使用 matplotlib 绘制图像。

## 实验效果

---

训练日志在代码目录的 log.txt 文件中，包含训练过程中的 loss 变化。

希望拟合的函数图像如图（蓝色曲面）：



程序每训练 50 个 epoch 进行一次评估，并输出模型预测的函数图像，模型多次评估产生的图像（节选）如下所示（红色曲面），可观察到其逐步拟合的过程：

