《神经网络》大作业

姓名：陈越

学号：MG1937003

学院：人工智能学院

实验内容：

手写一个MLP对一个非线性函数近似。

设计方案：

首先构造能够产生数据的函数，函数很简单，直接调用numpy random函数生成即可，使用的是uniform采样。

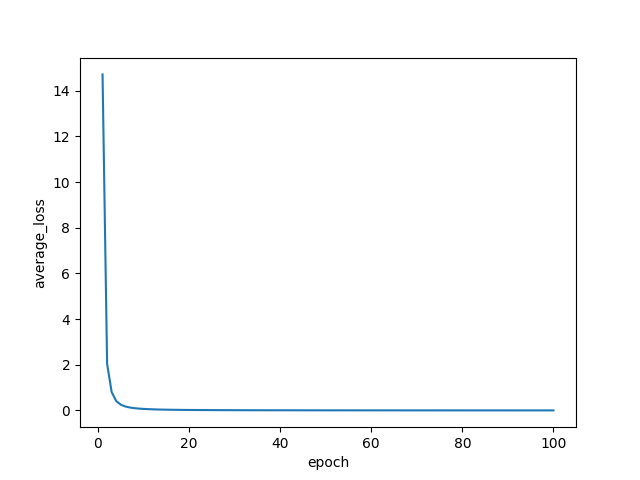
搭建模块化网络：

一个完整的神经网络应该由参数化的神经层以及相应的激活函数复合构成，为了方便整个网络的反向传播，我们对网络的神经层进行建模。

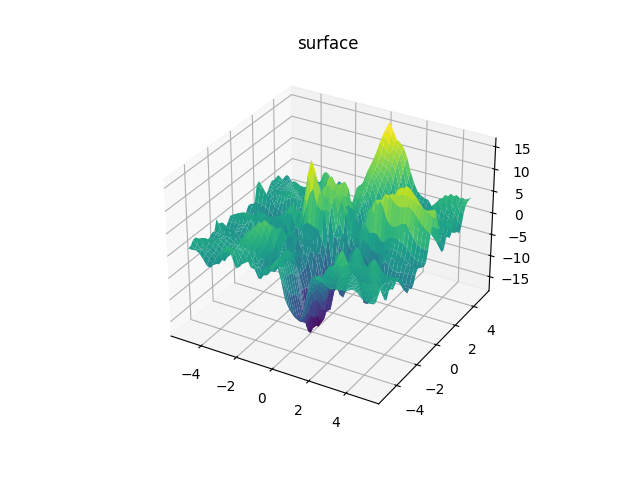
实现的激活函数有Sigmoid、ReLU、Identity。可以通过超参数对网络的输入大小、输出大小、激活函数类型进行设置。本次实验采用的梯度下降方法为最朴素的梯度下降方法。

实验结果：

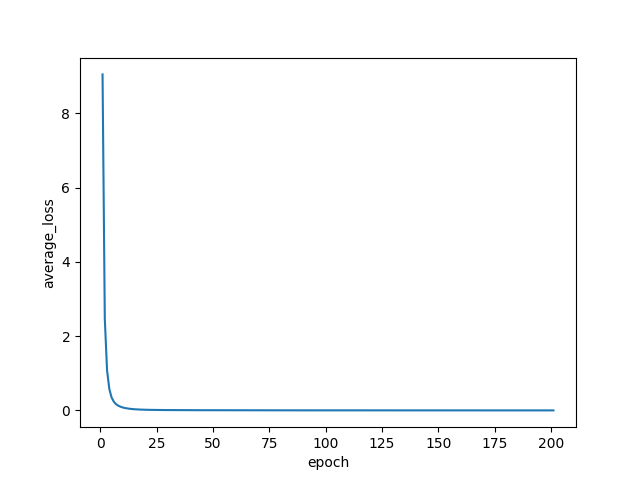
迭代100个Epoch，每个Epoch的loss曲线：

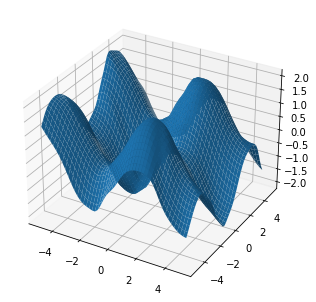


相应的重构图：



此时可以看到生成的数据不够好，此时算法只均匀采样了10000个样本。





总结：

在设计网络时，因为使用的数据远小于网络参数个数，所以可以成功拟合该函数。在这次实验中，发现随着训练轮数的增加，loss逐渐下降。模型越复杂，由于表示能力越强，也就可以越快地下降。应为ReLU函数的存在，如果不加batch norm的话很容易梯度爆炸，因此本次实验选用了sigmoid函数。但当模型深度达到一定程度时，使用sigmoid函数为激活函数会发生梯度消失现象，使得训练难以进行。