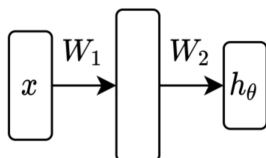


# 理论证明：一个两层的ReLU网络可以模拟任何函数

2150241 计科 赵彦平

在封闭集中采样密集的数据点，我们能够构造2层网络通过所有的这些数据点，并且网络是分段的线性函数，同时被拟合的函数是连续函数，通过采样尽量多的点实现无限逼近。



首先说明网络为什么是分段线性。

2层网络结构如图所示，其中 $x$ 是一维向量， $W_1, W_2$ 是网络权重，这里非线性函数选择ReLU，

$$h_{\theta}(x) = \text{ReLU}(XW_1)W_2 = \sum_{i,j} w_{2j} \max(0, w_{1i}x + b_i) b_j$$

可以简化为：

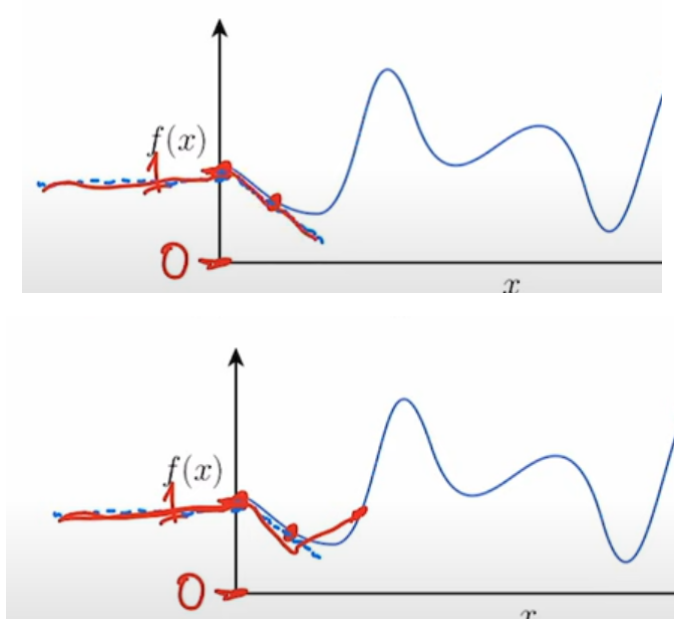
$$\hat{f}(x) = \sum_{i=1}^d \pm \max\{0, w_i x + b_i\}$$

函数可以由

$$\pm \max(0, w_i x + b_i)$$

线性组合叠加。

由于ReLU存在引入了非线性，当尝试用分段函数去表示的时候，就会发现下面的效果：



## 参考文献

[1] [为什么两层网络能够拟合任意函数 两层神经网络可以表达任意连续函数-CSDN博客](#)