

基于神经网络的曲线拟合

何宜晖

西安交通大学

heyihui@stu.xjtu.edu.cn

2011年3月

摘要

人工神经网络是近年来发展起来的模拟人脑生物过程的人工智能技术，具有自学习、自组织、自适应和很强的非线性映射能力。在人工神经网络的实际应用中，常采用BP神经网络或它的变化形式，BP神经网络是一种多层神经网络，因采用BP算法而得名，主要应用于模式识别和分类、函数逼近、数据压缩等领域 [3]。BP网络是一种多层前馈神经网络，由输入层、隐层和输出层组成。层与层之间采用全互连方式，同一层之间不存在相互连接，隐层可以有一个或多个。BP算法的学习过程由前向计算过程和误差反向传播过程组成，在前向计算过程中，输入信息从输入层经隐层逐层计算，并传向输出层，每层神经元的状态只影响下一层神经元的状态。如输出层不能得到期望的输出，则转入误差反向传播过程，误差信号沿原来的连接通路返回，通过修改各层的神经元的权值，使得网络系统误差最小，最终实现网络的实际输出与各自所对应的期望输出逼近。本文利用BP网络修改权值对 $y = \sin(x)$ 曲线实现拟合，并比较了各种实现细节。

1 实现

要对BP网络进行训练，必须准备训练样本。对样本数据的获取，可以通过用元素列表直接输入、创建数据文件，从数据文件中读取等方式，具体采用哪种方法，取决于数据的多少，数据文件的格式等。本使用均匀分布在线生成100个训练样本，以及100个测试样本。同时采用归一化处理，使得 $x \in [-1, 1]$, $y \in [-1, 1]$ 。归一化后的训练样本如下图：

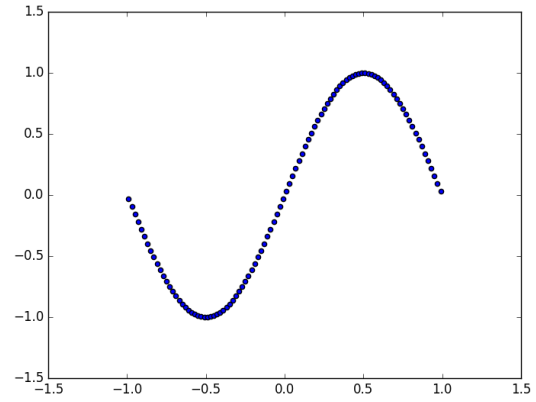


图 1: 训练集

根据系统输入输出序列，确定网络输入层节点数为1，隐含层节点数H为100，输出层节点数为1。初始化输入层、隐含层和输出层神经元之间的连接权值[1]

$$w_{l+1} = N(0, \sqrt{2/\hat{n}_l}) \quad (1)$$

其中 n_l 为前一层神经元个数， l 为层数，

给定学习速率 10^{-2} ，给定算法迭代次数100000，同时给定神经元激励函数 \tanh 。

激励函数，可设置为多种形式，本实验中所选函数为 \tanh ，从直觉讲，它的曲线较为平滑，输出在-1到1，更利于正弦函数拟合，如图2：

$$\tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \quad (2)$$

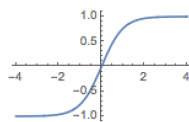


图 2: tanh

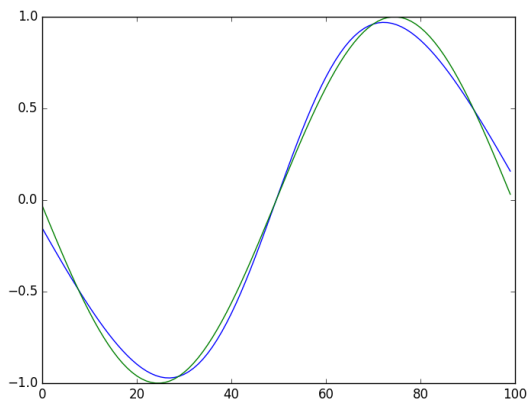


图 3: 初步拟合

我们简单设置了参数xx, 初步拟合sin函数, 如图3. 我们使用MSE作为我们的误差评判标准:

$$MSE = \frac{1}{n} \sum_{k=0}^n (y - f(x_k))^2 \quad (3)$$

初步拟合的误差为 10^{-3} , 可以看出, 拟合的效果并不好。

为了进一步提高拟合的准确率, 我们作出了两点改进。学习率的改进, 不再采用固定学习率, 而采用阶梯学习率

$$lr = lr_0 \gamma^{\lfloor \frac{iter}{step} \rfloor} \quad (4)$$

其中 lr_0 为初始学习率, $iter$ 为目前迭代次数, γ 为衰减系数, $step$ 为一个阶梯的迭代次数。我们一共设置3个阶梯。

另外, 我们将神经网络开始到结束每次迭代的结果, 用指数衰减结合起来[2]。

$$W_{ij}^k = (1 - \gamma)W_{ij} + \gamma W_{ij}^{k-1} \quad (5)$$

其中 γ 为指数衰变系数, W^k 为我们的当前最佳模型。

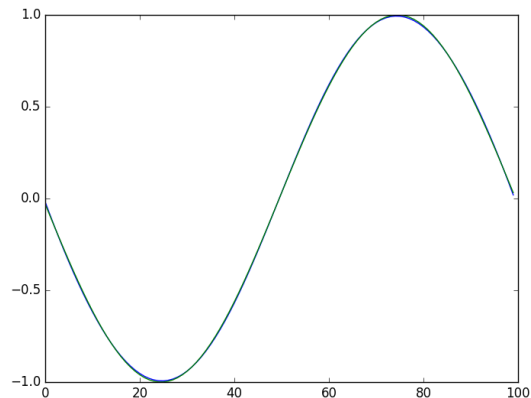


图 4: 最佳拟合

改良后, 我们的训练结果如图4, 误差为 10^{-6}

由图5可以看出, 惩罚基本一直处于下降阶段, 当学习达到第一个阶梯(1/3处), 学习率下降时, 惩罚也同样下降很多

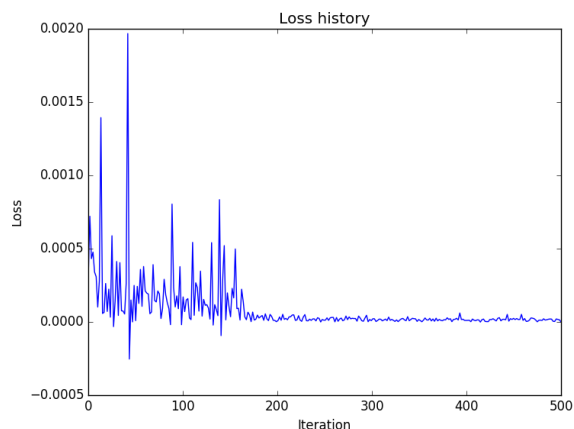


图 5: 最佳拟合

2 结论

通过上面分析可以看出, 神经网络基本上能将函数 $y = \sin(x)$ 拟合出来。虽然BP神经网络具有较

高的拟合能力，但是预测结果仍然存在一定的误差，基本的BP神经网络对于一些复杂系统的预测能力会比较差，其拟合能力存在局限性。要想达到精确拟合，需要针对问题设计训练策略。

参考文献

- [1] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1026–1034, 2015.
- [2] A. Karpathy, F. Li, and J. Johnson. Cs231n convolutional neural network for visual recognition,”. *Online Course*, 2016.
- [3] 鲍军鹏，张选平. 人工智能导论, 2010.