

# BP 神经网络预测的 MATLAB 实现

焦淑华<sup>1</sup>, 夏 冰<sup>1</sup>, 徐海静<sup>2</sup>, 刘 莹<sup>3</sup>

(1. 哈尔滨金融高等专科学校 基础部, 黑龙江 哈尔滨 150030; 2. 黑龙江大学 信息管理学院, 黑龙江 哈尔滨 150080; 3. 黑龙江科技学院 数力系, 黑龙江 哈尔滨 150027)

**摘 要:**在人工神经网络的实际应用中,大多数的人工神经网络模型是采用前馈反向传播网络(Back - Propagation - Network,简称 BP 网络)或它的变化形式。它是前向网络的核心,体现了人工神经网络最精华的部分。近年来 MATLAB 因其编程效率高,易学易懂,被广泛应用。以旅游需求预测为例,说明 MATLAB 可以实现 BP 神经网络的预测。

**关键词:**神经网络;MATLAB;预测

## 一、问题的提出

在旅游经济的发展过程中,的确有某种规律性的东西,或者称为模式。这种规律性的东西常常蕴涵于看似杂乱无章的大量历史数据中,因此,要想获得有效的预测,关键在于根据历史数据发现规律。从数学的角度看,就是建立某种映射关系(函数),并进行函数的拟合。神经网络是解决这一问题的最佳工具之一。

经过几十年的发展,已经形成了上百种人工神经网络。1974 年, P. Werbos 在其博士论文中提出了第一个适合多层网络的学习算法,但该算法并未受到足够重视和广泛的应用。直到 20 世纪 80 年代中期, David Rumelhart, Geoffrey Hinton 和 Ronald Williams, David Parkr, 以及 Yann Le Cun 分别独立发现了 BP 算法。1986 年,美国加利福尼亚的 PDP(parallel distributed procession)小组发表了《Parallel Distributed Processing》一书,将该算法应用于神经网络的研究,才使之称为迄今为止最著名的多层网络学习算法——BP 算法。由此算法训练的神经网络,称之为 BP 神经网络。标准的 BP 网络是根据 W - H 学习规则,采用梯度下降算法,对非线性可微分函数进行权值训练的多层网络。由 Kolmogorov 定理和 BP 定理可知,对于一个三层 BP 神经网络,只要隐层节点数足够多,就具有逼近任意复杂的非线性映射的能力。

目前,在人工神经网络的实际应用中,大多数的人工神经网络模型是采用前馈反向传播网络(Back - Propagation - Network,简称 BP 网络)或它的变化形式。它是前向网络的核心,体现了人工神经网络最精华的部分。

BP 神经网络,即多层前馈式误差反传神经网络,通常

由输入层、输出层和若干隐含层构成:每一层都由若干个节点组成,每一个节点表示一个神经元,上层节点与下层节点之间通过权连接,层与层之间的节点采用全互联的连接方式,每层内节点之间没有联系。以一个三层结构的 BP 神经网络为例,即含有一个输入层、一个输出层和一个隐含层,其结构如图 1:

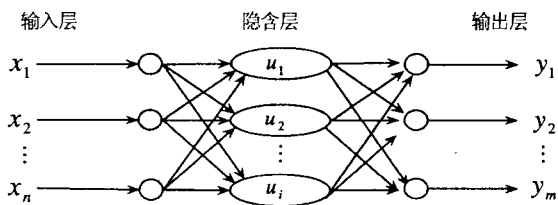


图 1

## 二、建立 BP 神经网络预测模型

在进行 BP 网络预测模型设计时,主要考虑网络的层数和每层中神经元的个数。

### (一) 网络层数

BP 网络是通过输入层到输出层的计算来完成的。多于一层的隐含层虽然能在速度上提高网络的训练,但是在实际应用中需要较多的训练时间,而训练速度可以用增加隐含层节点个数来实现,因此在应用 BP 神经网络进行预测时,选取只有一个隐含层的三层 BP 神经网络就足够了。

### (二) 网络各层中神经元的个数

输入、输出节点是与样本紧密相关的,与其应用的领域有关。根据旅游人数的历史数据,确定输入层神经元数为

5,即输入变量为连续5年的旅游人数;输出层神经元数为1,即输出变量为第六年的旅游人数。

如果隐含层神经元数目过少,网络很难识别样本,难以完成训练,并且网络的容错性也会降低;如果数目过多,则会增加网络的迭代次数,从而延长网络的训练时间,同时也会降低网络的泛化能力,导致预测能力下降。在具体设计时,首先根据经验公式初步确定隐含层神经元个数,然后通过对不同神经元数的网络进行训练对比,再最终确定神经元数。通用的隐含层神经元数的确定经验公式有:

$$i = \sqrt{n + m} + a$$

其中*i*为隐含层神经元的个数,*n*为输入层神经元的个数,*m*为输出层神经元的个数,*a*为常数且1 < *a* < 10(参见图1)。

由此,可以设定隐含层的神经元数为12。

### 三、BP神经网络预测的MATLAB实现

#### (一)数据样本的预处理

本文的数据样本来源于《中国旅游统计年鉴》。为了保证数据为同一数量级,首先需要对神经网络的输入和输出数据进行一定的预处理:将原数据乘以 $10^{-5}$ 。同时,将样本分为训练集和测试集,1993-1998年外国人入境旅游人数的处理结果作为训练集,即1993-1997年外国人入境旅游人数的处理结果作为训练输入;1998年外国人入境旅游人数的处理结果作为训练输出;1994-1999年外国人入境旅游人数的处理结果作为测试集,即1994-1998年外国人入境旅游人数的处理结果作为测试输入,1999年外国人入境旅游人数的处理结果作为测试输出。

#### (二)确定激活函数

根据处理后的数据范围,本文选取tansig和purelin作为激活函数。

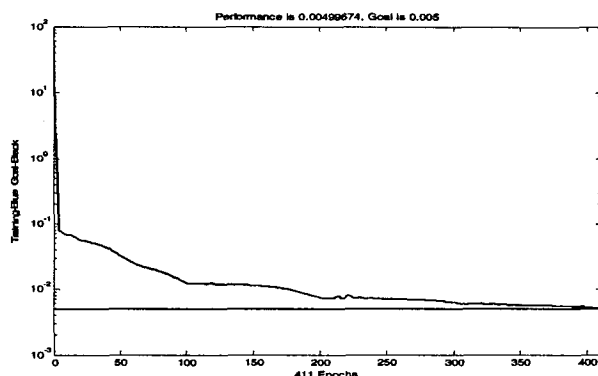
(三)设定网络的最大学习迭代次数为6000次。

(四)设定网络的学习精度为0.005

(五)创建和训练BP神经网络的MATLAB程序

```
% 旅游需求预测lyyc
>> clear all;
>> P = [2.16552 2.32935 3.13709 3.20267 3.36514
3.44702 3.35891 3.86590 4.15380 4.34453 3.33477
3.35957
2.63099 3.11346 4.01663 3.85012 3.94548 3.80859
3.94455 4.31889 4.35980 4.92129 4.04607 3.60270
3.18748 2.95487 4.24434 4.51493 4.57562 4.43475
4.27831 4.58795 4.96079 5.47786 4.79511 3.80859
3.44154 3.48779 4.53406 4.77124 4.80918 4.84979
5.19510 5.79808 5.45335 6.18219 5.37913 4.96571
4.46305 3.84784 5.62327 6.11243 5.65543 5.48965
5.69776 6.23772 5.92457 6.96998 5.93636 5.48528];
>> T = [4.99714 4.70636 6.28258 6.2659 6.33821
5.89043 6.08856 6.94419 6.96108 7.56545 6.48063
5.59884];
>> net = newff(minmax(P),[12 1],{'tansig' 'purelin'
},traingdx',learnrgdm');
>> net.trainParam.epochs = 10000;
>> net.trainParam.goal = 0.005;
```

```
>> net.trainParam.show = 500;
>> net = train(net,P,T)
```



显示的数据与所设计的网络模型相符,且如图显示网络学习迭代到411次时,就达到了学习精度0.00499674,其学习速度较快。

#### (六)测试BP神经网络

将测试的输出数据还原,与实际人数比较(见表1),说明BP神经网络预测的MATLAB实现是可行的。

表1 1999年1—12月份外国人入境旅游人数

1999年	实际人数	BP神经网络预测人数
1月	529 323	528 500
2月	494 216	486 010
3月	690 393	661 960
4月	716 292	715 160
5月	724 188	722 710
6月	693 599	710 690
7月	718 341	713 530
8月	769 209	764 460
9月	769 967	770 810
10月	887 492	805 620
11月	776 649	775 250
12月	662 627	613 980

### 四、小结

神经网络通过计算机程序实现对非线性映射逼近,在众多语言中,MATLAB语言允许数学形式的语言编写程序,比其他语言更接近我们书写计算公式的思维方式。因此编程效率高,易学易懂。

#### 参考文献:

- [1]叶世伟.神经网络原理[M].北京:机械工业出版社,2004.
- [2]周开利,康耀红.神经网络模型及其MATLAB仿真程序设计[M].北京:清华大学出版社,2005.
- [3]高隼.人工神经网络原理及仿真实例[M].北京:机械工业出版社,2003.
- [4]飞思科技产品研发中心.MATLAB辅助神经网络分析与设计[M].北京:电子工业出版社,2003.

责任编辑:王丽华