DOI:10. 19431/j. cnki. 1673-0062. 2020. 01. 014

基于 Softplus 函数的双隐含层 BP 神经网络的公路客运量预测

徐　森,崔舒为

(兰州交通大学 数理学院,甘肃 兰州 730070)

摘　要:为了提高预测模型的精度,提出一种基于 Softplus 激活函数的双隐含层 BP 神经网络的预测方法,提高了模型的非线性学习和泛化能力及预测精度,并改善了网络性能。将该方法应用于公路客运量实际预测中进行有效性验证,结果表明该方法对公路客运量有更好的非线性拟合能力和预测准确性。

关键词:Softplus 函数;双隐含层;BP 神经网络;公路客运量

中图分类号:U492. 413　　**文献标志码**:B　　**文章编号**:1673-0062(2020)01-0088-05

Highway Passenger Traffic Forecast Based on Double Implicit Layer BP Neural Network Based on Softplus Function

XU Sen, CUI Shuwei

(School of Mathematics and Physics, Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China)

Abstract: In order to improve the accuracy of the prediction model, a prediction method based on Softplus activation function for double hidden layer BP neural network is proposed, which improves the nonlinear learning and generalization ability and prediction accuracy of the model, and improves the network performance. The method is applied to the actual prediction of highway passenger traffic volume. The results show that the proposed method has better nonlinear fitting ability and prediction accuracy for highway passenger traffic.

key words: Softplus function; double hidden layer; BP neural network; highway passenger traffic

收稿日期:2019-09-06

作者简介:徐　森(1995-),男,硕士研究生,主要从事概率论与数理统计、数据预测方面的研究。E-mail:2048538446@qq. com

0　引　言

交通运输在国民经济和社会发展中起着重要的作用,对公路客运量的预测具有重要意义。公路客运量预测方法主要有指数平滑法、弹性系数法、回归分析法、灰色预测法、支持向量机、时间序列法、马尔可夫预测法等[1-2]。

国内外学者对公路客运量预测进行了大量研究,Garrido 等建立多项概率模型且用蒙特卡洛方法计算精度进行评价[3];陈鹏等运用灰色方法建立客运量预测模型并用马尔可夫方法修正预测值[4];孙崎峰等通过改进反向传播(back propagation,BP)神经网络进行公路旅游客流量预测[5]。

本文采用 1994—2009 年中国统计年鉴公路客运量数据,构造基于 Softplus 激活函数的双隐含层 BP 神经网络模型。并将预测结果与原始 BP 神经网络和双隐含层 BP 神经网络进行对比,旨在通过本次的数据对比,证明该方法对公路客运量具有更好的非线性拟合能力和预测准确性。

1　双隐含层 BP 神经网络

BP 神经网络是应用最广泛的神经网络模型之一。它是一种多层前向反馈神经网络,由输入层、隐含层和输出层以及每层神经元间的连接组成,其中隐含层可以有多层。理论上 3 层或 3 层以上的 BP 神经网络,只要有足够多的隐含层神经元,就能以任意精度逼近任意函数,因此在预测领域的研究中,学者大都采用 3 层的 BP 神经网络作为预测模型。而实际中,由于存在计算精度、四舍五入等因素,单隐藏层(3 层)神经网络难以准确逼近未知函数。因此,本文选择构造双隐含层 BP 神经网络模型[6-7]。

双隐含层 BP 神经网络模型如图 1 所示。从输入层进入,通过各隐含层节点,传递到输出层,每一层节点的输出结果只影响下一层节点的输出结果,相邻层每个节点通过适当的连接权值和阈值前向连接。

设输入层有 M 个输入信号,其中的任一输入信号用 m 表示;第 1 隐含层有 I 个神经元,其中任一神经元用 I 表示;第 2 隐含层有 J 个神经元,其中任一神经元 j 表示;输出层有 P 个神经元,其中任一神经元用 p 表示。输入层与第 1 隐含层的权用 ω_{mi} 表示,第 1 隐含层与第 2 隐含层的权用 ω_{ij} 表示,第 2 隐含层与输出层的权用 ω_{jp} 表示。用 u 表示神经元的输入,用 v 表示激活输出。U 和 v 的上标表示层,下标表示层中的神经元。训练样本集为 $X=[X_1,X_2,\cdots,X_k,\cdots,X_N]$,对应任一训练样本 $X_k=[x_{k1},x_{k2},\cdots,x_{kM}]'$,$k\in(1,2,\cdots,M)$。实际输出为 $Y_k=[y_{k1},y_{k2},\cdots,y_{kP}]'$,期望输出为 $D_k=[d_{k1},d_{k2},\cdots,d_{kp}]'$。

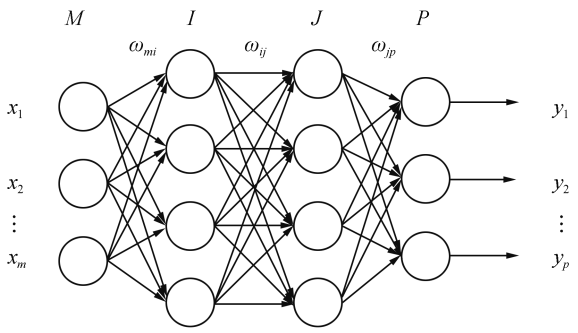

图 1　双隐含层 BP 神经网络

Fig. 1　Double hidden layer BP neural network

1)前向传播输出:从输入层到隐含层至输出层的逐层传播:

$$u_i^I=\sum_{m=1}^{M}\omega_{mi}x_{km},v_i^I=f(u_i^I),\quad i=1,2,\cdots,I \tag{1}$$

$$u_j^J=\sum_{i=1}^{I}\omega_{ij}v_i^I,v_j^J=\varphi(u_j^J),\quad j=1,2,\cdots,J \tag{2}$$

$$u_p^P=\sum_{j=1}^{J}\omega_{jp}v_j^J,v_p^P=\psi(u_p^P),\quad P=1,2,\cdots,P \tag{3}$$

$$y_{kp}=v_p^P=\psi\Big(\sum_{j=1}^{J}\omega_{jp}v_j^J\Big) \tag{4}$$

输出层第 p 个神经元的误差信号为

$$e_{kp}(n)=d_{kp}(n)-y_{kp}(n) \tag{5}$$

定义神经元的误差为 $e_{kp}^2(n)$,则输出层所有神经元的误差总和为 $E(n)$:

$$E(n)=\sum_{p=1}^{P}\frac{1}{2}e_{kp}^2 \tag{6}$$

2)反向误差逆传播:输出层经隐含层反向传播至输入层,逐层修正连接权。

输出层与第二隐含层、第二隐含层与第一隐含层、第一隐含层与输入层之间的权值修正量分别为 $\Delta\omega_{jp}(n)$、$\Delta\omega_{ij}(n)$、$\Delta\omega_{mi}(n)$。

$$\Delta\omega_{jp}(n)=-\eta\frac{\partial E(n)}{\partial\omega_{jp}(n)} \tag{7}$$

$$\Delta\omega_{ij}(n)=-\eta\frac{\partial E(n)}{\partial\omega_{ij}(n)} \tag{8}$$

$$\Delta\omega_{mi}(n) = -\eta\frac{\partial E(n)}{\partial\omega_{mi}(n)} \tag{9}$$

其中:η 表示学习步长。

2 激活函数

激活函数的性能表现,在神经网络中一直处于核心地位。Hornik 和 Leshno[8] 证明了,对于单隐含层的前馈神经网络,如果激活函数为连续的、有界的、非常数的非多项式,则它一定能够逼近任意连续的目标函数。因此,选择合适的激活函数对优化网络性能具有重要的作用。

Sigmoid 系函数主要包括 Sigmoid 函数和 Tanh 函数,是目前 BP 神经网络应用得最广泛的激活函数。但在深度神经网络中,Sigmoid 函数作为激活函数会导致梯度消失。

近年来,在深度学习领域,一种新的修正线性函数正逐渐取代 Sigmoid 函数成为主流,由该函数产生的单元被称为修正线性单元(rectified linear units,ReLU)[8],函数定义为

$$f(x) = \max(0,x) \tag{10}$$

它的特点是,如果计算值小于 0,则等于 0,否则原值保持不变。与 Sigmoid 函数相比,ReLU 函数更接近生物激活模型,且形式简单。由于不使用幂运算和除法运算,ReLU 函数具有更快的运算速度和更好的泛化能力。同时,该激活函数具有一定的稀疏能力,无需任何预训练即可直接规范网络。但 ReLU 函数的稀疏能力是通过一种简单的方式强制某些数据为 0 来实现的,可能会损害其预测能力,降低网络模型的平均性能。同时,ReLU 函数是分别在其大于和小于 0 的部分是线性函数,缺乏非线性映射。考虑到 ReLU 函数的这些缺点,文献[10]提出了 ReLU 的近似平滑的 Softplus 激活函数,其函数定义为

$$f(x) = \ln(1 + e^x) \tag{11}$$

Softplus 激活函数是一种非线性连续可微函数,其变化平缓,较 Sigmoid 函数也更接近生物激活模型,同时能避免 ReLU 函数强制性稀疏的问题,使整个网络的平均性能更好。因此,本文在双隐含层 BP 神经网络模型中引入了 Softplus 函数,代替 Sigmoid 函数作激活函数。

3 网络模型的构建

本文选择 1994—2004 年的全国公路客运量数据为神经网络的训练样本进行训练,以前 3 年的数据为输入向量,以第 4 年的数据为目标向量,可获取 10 组训练数据,2007—2010 年的数据作为网络的测试样本,用于验证预测的准确性。以 2007—2010 年的数据进行验证与测试,尝试建立预测模型;输入层和输出层中的节点数由研究对象决定,隐含层中的节点数需要主观确定。根据经验公式

$$m = \sqrt{n + l} + a \tag{12}$$

$$m = \log_2 n \tag{13}$$

$$m = \sqrt{nl} \tag{14}$$

其中 m 为隐含层节点数,n 为输入层节点数,l 为输出层节点数,a 为 1 到 10 的整数,通过试凑法,来确定隐含层节点数,本文网络模型结构确定为 3-5-2-1。

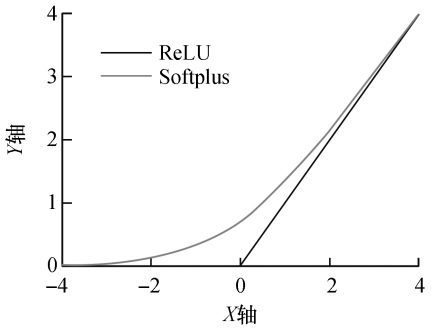

图 2 ReLU 和 Softplus 函数的图像

Fig. 2 Image of ReLU and Softplus functions

在网络学习过程中,为消除样本间的量纲差异以便更好地学习,需要对样本进行归一化处理:

采用公式

$$x'_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{15}$$

对数据进行归一化,其中,x'_i 表示第 i 个数据归一化后的数据;x_i 表示原始数据;$x_{\max}$、$x_{\min}$ 分别表示原始数据中最大值和最小值。

建模过程中,学习速率为 0.01,精度要求为 0.004,训练次数为 5 000。激活函数为 Softplus 函数。为定性比较模型的优劣,引入以下 2 种评价指标:

平均绝对百分比误差

$$E_{\text{MAPE}} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{16}$$

均方误差(MSE)

$$E_{\text{MSE}} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{17}$$

其中,y_i、$\hat{y}_i$ 分别表示实测值、预测值。

4　实例分析

4.1　数据

本文选取 1994—2010 年全国公路客运量数据,1997—2006 年公路客运量数据作为训练集,2007—2010 年数据作为测试集,如表 1 所示。

表 1　公路客运量数据

Table 1　Road passenger traffic data

年份	公路客运量/万人	年份	公路客运量/万人
1994	953 940	2003	1 464 335
1995	1 040 810	2004	1 624 526
1996	1 122 110	2005	1 697 381
1997	1 204 583	2006	1 860 487
1998	1 257 332	2007	2 050 680
1999	1 269 004	2008	2 682 114
2000	1 347 392	2009	2 779 081
2001	1 402 798	2010	3 052 738
2002	1 475 257		

4.2　结果与分析

从图 3 可以直观看出三种模型的预测结果,都能够很好地预测到结果,但单隐含层 BP 神经网络模型的预测值与实际值相差相对较大,双隐含层 BP 神经网络与 Softplus 函数改进的双隐含层 BP 神经网络则较逼近实际值,而后者对实际值的预测精确度明显优于前者。

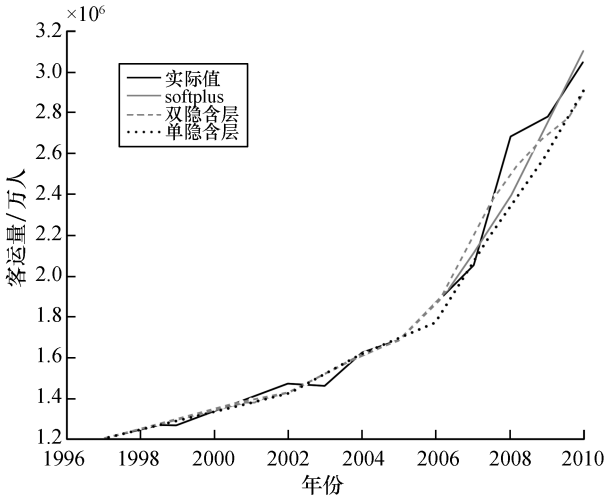

图 3　各模型预测结果比较

Fig. 3　Comparison of prediction results of each model

表 2 是三种模型的预测结果的统计,表 3 是三种模型的预测对比。由表 3 可见,Softplus 函数改进的双隐含层 BP 神经网络模型的预测效果最好,双隐含层 BP 神经网络模型次之。

表 2　各模型预测值对比

Table 2　Comparison of predicted values of each model

年份	实际值/10^{10} 人	单隐含层		双隐含层		Softplus 函数	
		预测值/10^{10} 人	相对误差/%	预测值/10^{10} 人	相对误差/%	预测值/10^{10} 人	相对误差/%
2007	2.05	2.09	1.95	2.21	7.97	2.11	2.93
2008	2.68	2.32	13.43	2.51	6.51	2.39	10.78
2009	2.78	2.59	6.83	2.69	3.03	2.75	1.08
2010	3.05	2.90	4.92	2.84	7.12	3.11	1.77

表 3　预测模型性能对比

Table 3　Comparison of predictive model performance

预测模型	网络结构	预测对比	
		E_{MAPE}	E_{MSE}
Softplus 函数	3-5-2-1	4.14%	2.28×10^{10}
双隐含层	3-5-2-1	6.16%	2.79×10^{10}
单隐含层	3-5-1	6.78%	4.74×10^{10}

5　结　论

本文提出了用基于 Softplus 函数的双隐含层 BP 神经网络预测公路客运量的方法。与单隐含层 BP 神经网络相比,双隐含层 BP 神经网络增强了网络输入输出之间复杂关系的映射能力,具有更强的逼近能力和容错能力。而使用 Softplus 函数作为激活函数,以保证网络在学习过程中不出现梯度消失造成无法收敛等问题。实验结果表明,该模型优于单隐含层 BP 神经网络和双隐含

层 BP 神经网络,有效降低了误差,具有更好的非线性拟合能力和预测准确性。

参考文献:

[1] 王生昌,白韶波,张慧. 公路客运量预测方法的比较[J]. 长安大学学报(自然科学版),2005,25(5):83-86.

[2] 张丽莉. 基于神经网络的公路货运量预测方法研究[J]. 华中师范大学学报(自科版),2015,49(2):186-189.

[3] GARRIDO R A,MAHMASSANI H S. Forecasting freight transportation demand with the space-time multinomial probit model[J]. Transportation research part B,2000,34(5):403-418.

[4] 陈鹏,孙全欣. 基于灰色马尔柯夫过程的铁路客运量预测方法研究[J]. 铁道运输与经济,2005,27(4):65-67.

[5] 孙崎峰,周栩,孙晓峰. 基于改进 BP 神经网络的公路旅游客流量预测[J]. 山东理工大学学报(自然科学版),2008,22(6):75-79.

[6] 吕砚山,赵正琦. BP 神经网络的优化及应用研究[J]. 北京化工大学学报(自然科学版),2001,28(1):67-69.

[7] 刘天舒. BP 神经网络的改进研究及应用[D]. 哈尔滨:东北农业大学,2011.

[8] LESHNO M,LIN V Y,PINKUS A,et al. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function[J]. Neural networks,1993,6(6):861-867.

[9] ZEILER M D,RANZATO M,MONGA R,et al. On rectified linear units for speech processing[C]//Acoustics,Speech and Signal Processing (ICASSP),2013 IEEE International Conference on. Vancouver:IEEE,2013:3517-3521.

[10] GLOROT X,BORDES A,BENGIO Y. Deep sparse rectifier networks[J]. Journal of machine learning research,2011,15:315-323.

(责任编辑:周泉)