文章编号　1004-924X(2018)11-2805-09

基于 Pearson 关联度 BP 神经网络的时间序列预测

王　可[1,2*]，王慧琴[1,2]，殷　颖[1]，毛　力[1]，张　毅[1]
(1. 西安建筑科技大学 信息与控制工程学院，陕西 西安 710055；
2. 西安建筑科技大学 管理学院，陕西 西安 710055)

摘要：针对 BP 神经网络存在的过拟合问题，提出了基于 Pearson 关联度的神经网络预测模型。将传统的基于误差反向传播的 BP 神经网络中的误差函数替换为 Pearson 关联度函数，利用梯度上升法对训练过程中神经网络的连接权重和阈值的调整量进行了推导，并为调整量添加了动量项用于提高神经网络收敛速度，然后建立了关联度反向传播预测模型，并对其权重进行了阈值限制以及增加学习率来防止过拟合。对通用数据集进行时间序列预测实验，通过与改进的 RBF 和 BP 神经网络对比，表明对于多因素时间序列的预测 Pearson 关联度 BP 神经网络的预测误差精度 RMSE 降低了 4.02，收敛次数减少 1 690 代。实现了将关联分析与 BP 神经网络的结合，能够在保证效率的同时，解决过拟合问题，提高预测精度。

关　键　词：Pearson 关联度；BP 神经网络；时间序列预测
中图分类号：TP183　　**文献标识码**：A　　**doi**：10.3788/OPE.20182611.2805

Time series prediction method based on Pearson correlation BP neural network

WANG Ke[1,2*], WANG Hui-qin[1,2], YIN Ying[1], MAO Li[1], ZHANG Yi[1]

(1. *School of Information and Control Engineering*,
Xi'an University of Architecture and Technology, *Xi'an* 710055, *China*;
2. *School of Management*, *Xi'an University of Architecture and Technology*, *Xi'an* 710055, *China*)
* *Corresponding author*, *E-mail*: *wangke@xauat.edu.cn*

Abstract: In order to realize the over fitting problem existing in Back Propagation (BP) neural networks, a neural prediction model based on Pearson correlation was designed. It replaces the error function in a BP neural network based on error back propagation with the Pearson correlation function. By means of gradient ascent, the adjustment of connection weights and biases in training process is derived. Meanwhile, momentum is added to this adjustment to improve the convergence speed of the network. The Pearson correlation BP prediction model is built with weight threshold limiting and an increasing learning rate to prevent overfitting. Time series prediction experiments on a

收稿日期：2018-05-31；**修订日期**：2018-06-29.
基金项目：国家自然科学基金青年基金资助项目(No. 61701388)；住建部科学技术计划资助项目(No. 2017-K2-014)；陕西省科技厅国际合作资助项目(No. 2017KW-036)；陕西省自然科学基础研究计划资助项目(No. 2018JM5127，No. 2018JM6080)；西安市科技局软科学资助项目(No. 2016043SF/RK06(3))；西安市碑林区科技计划资助项目(No. GX1606)；西安建筑科技大学青年基金项目(No. QN1628)

standard dataset were performed. The results demonstrate that compared with improved the radial basis function and BP neural networks, the Pearson correlation BP neural network reduces root-mean-square error, and time to convergence in multi-factor time series prediction. Therefore, the Pearson correlation BP neural network realizes the integration of correlation analysis with neural networks, is able to ensure efficiency, and can solve fitting problems in the same time as other methods with higher accuracy.

Key words: Pearson correlation; BP neural network; time series prediction

1 引 言

作为机器学习理论中应用最为广泛的算法，BP(Back Propagation)神经网络[1]采用基于误差反向传播的梯度算法，以其良好的非线性函数和泛化能力以及其自组织、自适应性能和容错性等特点为非线性时间序列的建模、预测提供了一条行之有效的方法，有效地克服了时间序列预测方法[2]存在的平稳性不合理、非线性处理能力差等问题。

但BP网络在预测问题的应用中存在过拟合[3]问题，即网络拟合精度越高，预测性能越差，也就是泛化能力越差。为解决这一问题，相关研究如Yu[4]指出网络结构、样本质量、训练参数设置等对神经网络泛化能力有较为重要的影响。Srivastava N[5]证明通过向样本数据中加入随机噪声迫使神经网络不能精确地拟合数据来提高泛化能力；更有效的方法是在样本数据中添加正则项防止对网络学习过度。王林[6]针对BP神经网络学习算法随机初始化连接权值和阈值易使模型陷入局部极小点的缺点，设计了一种自适应差分进化算法优化BP神经网络的混合算法。部分地避免了过拟合现象但增加了算法的复杂度，降低了效率。

在多因素时间序列预测方面，与神经网络结合的方法主要是对可能影响待预测数据的因素与待预测数据进行关联分析[7-8]，为了降低网络的复杂度，筛选出关联系数较大的因素构造一个多输入单输出的神经网络，然后再传输到神经网络中。关联分析采用的方法主要有偏相关[9]、Pearson关联度[10]、灰色绝对关联度[11]等。

但以上研究成果在防止模型过拟合的参数设置上考虑较少，导致模型在其他问题上的应用存在阻碍。同时关联分析独立于神经网络之外，没有引入到神经网络的训练过程中，只是在预处理和衡量神经网络预测性能时使用，并没有与神经网络有机融合。

因而本文利用神经网络的实际输出序列与期望输出序列的关联度取代传统的代价函数，建立了基于关联度反向传播的神经网络(Correlation Back Propagation，Corr-BP)，Corr-BP从设计之初就没有以函数逼近作为直接目标，而是采用了实际输出和参考输出的关联度作为学习信号的来源，建立的是实际输出与参考输出的线性相关关系，为解决BP神经网络过拟合提供一种新的思路。从设计思想来说Corr-BP不适用于做函数逼近却具有直接的防止过拟合的优势，在训练过程中利用Pearson关联度与误差相结合的方式判断训练是否停止，用以提高泛化能力，最后利用线性回归建立关联度输出结果与期望输出的映射，该网络具有如下主要特点：

(1)将关联分析与预测集中到神经网络中计算，简化了预测步骤；

(2)利用关联度现有的典型相关评价指标，为防止过拟合参数的设定对样本有更好的适应性。

2 Corr-BP预测模型

2.1 Pearson关联度

Pearson关联度是用来衡量两个变量之间的线性相互关系的，取值范围在[−1，+1]之间，将关联度分级，分级越细越能够表示两个时间序列不同的相关程度，Pearson关联度与变量相关性的描述如表1所示[12]。

表 1　关联度取值与相关程度

Tab. 1　Correlation degree and relevance

关联度绝对值	意义
0.00～0.19	极低相关
0.20～0.39	低度相关
0.40～0.69	中度相关
0.70～0.89	高度相关
0.90～1.00	极高相关

从表 1 中看出，在神经网络的训练过程中，当实际输出与参考输出的关联度接近 1 时就越有可能出现过拟合现象，关联度过小在中度相关以下则不足以表示出相关关系，因而实际使用中只需要考虑高度相关和极高相关两种情况，当两个变量相关程度达到上述情况时，则可以利用最小二乘线性回归[13]建立彼此的线性函数关系。

2.2　设计思路

Pearson 关联度是一种线性关联度，表示的是两个时间序列的线性相关关系，当两个时间序列关联度较高时，就可以通过线性变换用其中一个时间序列表示另外一个，利用关联度代替神经网络的误差函数，建立实际输出与参考输出的线性关系，利用最小二乘法计算出二者的线性关系，能够完成对待预测数据的预测建模。关联度神经网络设计思路如下。

2.3　网络结构与训练方法

Corr-BP 神经网络结构有 4 层，分别为输入层、隐含层、关联度层、输出层，如图 1 所示。

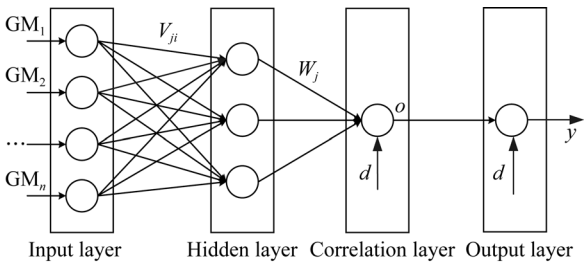


图 1　Corr-BP 神经网络模型

Fig. 1　Corr-BP neural network model

图 1 中样本数为 P，输入向量 $\mathbf{X}=(x_1, x_2, \cdots, x_i, \cdots, x_n)^{\mathrm{T}}$，隐含层输出向量 $\mathbf{Y}=(y_1, y_2, \cdots, y_j, \cdots, y_m)^{\mathrm{T}}$，关联度层输出 $\mathbf{O}=(o_1, o_2, \cdots, o_k, \cdots, o_l)^{\mathrm{T}}$，期望输出 $\boldsymbol{d}=(d_1, d_2, \cdots, d_k, \cdots, d_l)^{\mathrm{T}}$，输入层到隐含层之间的权重矩阵用 $\mathbf{V}$ 表示，隐含层到关联度层之间的权重矩阵用 $\mathbf{W}$ 表示。

隐含层的传输函数为单极性 Sigmoid 函数，如式(1)所示：

$$f(x)=\frac{1}{1+e^{-x}}. \tag{1}$$

$f(x)$ 具有连续、可导的特点，且有如式(2)所示性质：

$$f'(x)=f(x)[1-f(x)]. \tag{2}$$

相比于传统的 BP 神经网络，Corr-BP 神经网络在输出层前增加了一个关联度层，由于 Corr-BP 适用于做时间序列预测，其预测值的变化范围不一定为[－1，1]，故选择线性传输函数更为合适。选择线性函数作为关联度层传输函数如式(3)所示：

$$f(x)=x. \tag{3}$$

对于关联度层，其输入和输出如式(4)和式(5)：

$$o_k^p=f(net_k^p),\ p=1,2,\cdots,P, \tag{4}$$

$$net_k^p=\sum_{j=1}^{m} w_{kj} y_j^p-b_k,\ p=1,2,\cdots,P. \tag{5}$$

对于隐含层，其输入和输出如式(6)和式(7)：

$$y_j^p=f(net_j^p),\ p=1,2,\cdots,P, \tag{6}$$

$$net_j^p=\sum_{i=0}^{n} v_{ji} x_j^p-b_j,\ p=1,2,\cdots,P. \tag{7}$$

计算关联度层输出与实际输出的 Pearson 关联系数 $Corr(\mathbf{O},\boldsymbol{d})$ 如式(8)：

$$Corr(\mathbf{O},\boldsymbol{d})=\frac{\sum_{p=1}^{P}(o^p-\bar{o})(d^p-\bar{d})}{\sqrt{\sum_{p=1}^{P}(o^p-\bar{o})^2}\sqrt{\sum_{p=1}^{P}(d^p-\bar{d})^2}}=\frac{\sum_{p=1}^{P}o^p d^p-\frac{1}{P}\left(\sum_{p=1}^{P}o^p\right)\left(\sum_{p=1}^{P}d^p\right)}{\sqrt{\sum_{p=1}^{P}(o^p)^2-\frac{1}{P}\left(\sum_{p=1}^{P}o^p\right)^2}\sqrt{\sum_{p=1}^{P}(d^p)^2-\frac{1}{P}\left(\sum_{p=1}^{P}d^p\right)^2}}. \tag{8}$$

调整权重的原则就是使关联系数不断地增大，因此应使权重的调整量与关联度的梯度上升成正比，如式(9)所示：

$$\Delta w_{kj}=\eta\frac{\partial Corr}{\partial w_{kj}},\ j=1,2,\cdots,m;\ k=1,$$

$$\Delta v_{ji}=\eta\frac{\partial Corr}{\partial v_{ji}},\ i=1,2,\cdots,n;\ j=1,2,\cdots,m. \tag{9}$$

式中常数 $\eta\in(0,1)$ 表示学习率。进一步地，权重调整可表示为式(10)：

$$\Delta w_{kj}=\eta\frac{\partial Corr}{\partial w_{kj}}=\sum_{p=1}^{P}\frac{\partial Corr}{\partial o^p}\cdot\frac{\partial o^p}{\partial w_{kj}},$$

$$\Delta v_{ji}=\eta\frac{\partial Corr}{\partial v_{ji}}=\sum_{p=1}^{P}\frac{\partial Corr}{\partial o^p}\cdot\frac{\partial o^p}{\partial v_{ji}}. \quad (10)$$

由于关联度考虑了所有的输入，故其训练方式没有在线训练，只有批训练(Batch Training)，样本数较多时能够提高网络的训练速度，将式(10)中各项逐步展开，如式(11)所示：

$$\frac{\partial o^p}{\partial w_{kj}}=\frac{\partial o^p}{\partial net_k^p}\cdot\frac{\partial net_k^p}{\partial w_{kj}},$$

$$\frac{\partial o^p}{\partial v_{ji}}=\frac{\partial o^p}{\partial net_k^p}\cdot\frac{\partial net_k^p}{\partial y_j^p}\cdot\frac{\partial y_j^p}{\partial net_j^p}\cdot\frac{\partial net_j^p}{\partial v_{ji}}. \quad (11)$$

在给定的传输函数下计算式(10)和式(11)中各项偏导数如式(12)～式(17)所示：

$$\frac{\partial o^p}{net_k^p}=f'(net_k^p)=1, \quad (12)$$

$$\frac{\partial net_k^p}{\partial w_{kj}}=y_j^p, \quad (13)$$

$$\frac{\partial net_k^p}{\partial y_j^p}=w_{kj}, \quad (14)$$

$$\frac{\partial y_j^p}{\partial net_j^p}=f'(net_j^p)=y_j^p(1-y_j^p), \quad (15)$$

$$\frac{\partial net_j^p}{\partial v_{ji}}=x_i^p, \quad (16)$$

$$\frac{\partial Corr}{\partial o^p}=\frac{1}{\sqrt{\sum_{p=1}^{P}(o^p-\bar{o})^2\sum_{p=1}^{P}(d^p-\bar{d})^2}}\times\left[\left(d^p-\frac{1}{P}\sum_{p=1}^{P}d^p\right)\sum_{p=1}^{P}(o^p-\bar{o})^2-\left(\sum_{p=1}^{P}(o^p-\bar{o})(d^p-\bar{d})\right)\left(o^p-\sum_{p=1}^{P}o^p\right)\right]. \quad (17)$$

记 $\varepsilon^p=\frac{\partial Corr}{\partial o^p}$，完整的权重调整公式如式(18)所示：

$$\Delta w_{kj}=\eta\sum_{p=1}^{P}\varepsilon^p y_j^p,$$

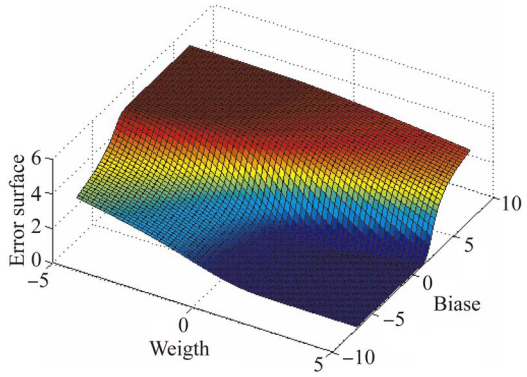
$$\Delta v_{ji}=\eta\sum_{p=1}^{P}\varepsilon^p w_{kj}y_j^p(1-y_j^p)x_i^p. \quad (18)$$

同理可推的神经元的偏置调整公式如式(19)所示：

$$\Delta b_k=\eta\sum_{p=1}^{P}\varepsilon^p,$$

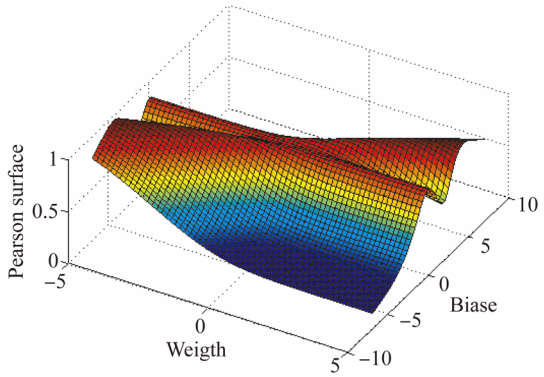
$$\Delta b_j=\eta\sum_{p=1}^{P}\varepsilon^p w_{kj}y_j^p(1-y_j^p). \quad (19)$$

Corr-BP 的关联度层相当于传统 BP 神经网络的输出层，所不同的是 Corr-BP 计算的是参考输出与实际输出的 Pearson 关联度，而传统 BP 神经网络计算的是它们的均方根误差，这两个学习信号的计算量相差不大，Corr-BP 完成的是一个线性最小二乘功能，其运算是非迭代的，实际中是训练完成之后的后处理，并不在迭代训练过程之列。如图 2 中(a)、(b)(彩图见期刊电子版)分别为相同的标准时间序列预测数据源下，随神经网络权重和偏置变化的误差曲面图和 Pearson 关联度曲面图，水平坐标分别为神经元权重 w 和神经元偏置 b，(c)、(d)分别为它们在水平面的投影，红色越深表示值越大，蓝色越深表示值越小。训练目标在误差曲面图表现为收敛至极小值点，在 Pearson 关联度曲面图表现为收敛至极大值点。



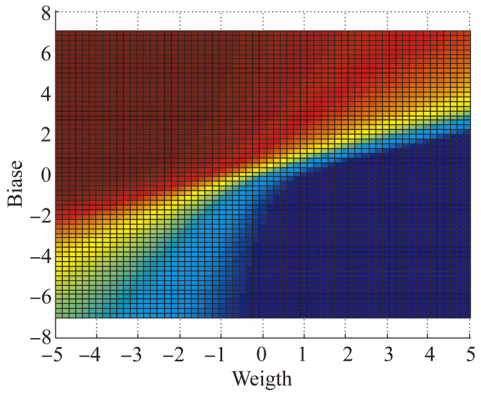
(a)误差曲面

(a)Error surface



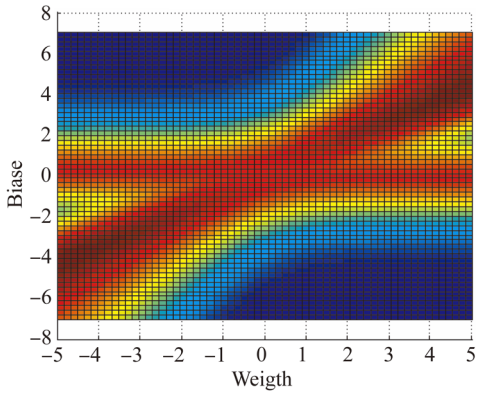
(b)Pearson 关联度曲面

(b)Pearson correlation degree surface



(c)误差曲面水平投影

(c)Error surface horizontal projection



(d)Pearson 关联度曲面水平投影

(d)Pearson correlation degree surface horizontal projection

图 2　误差曲面和 Pearson 关联度曲面

Fig. 2　Error surface and Pearson correlation degree surface

从图 2 中可以看出，Pearson 关联度的取值关于偏置 $b=0$ 对称，在训练过程中可以减少一半的迭代空间；误差曲面蓝色区域十分平坦不利于训练，Pearson 关联度的红色部分则相对起伏性强一些，说明 Corr-BP 能加快收敛，也就能够节省运算量。

在实践中不同数据源的 Pearson 关联度都具有类似的趋势，即中间都有一个类似于 X 形状的最大值区间，而误差曲面则各不相同，可以看出 Pearson 关联度的最优解的位置具有先验性，这一点对训练初始参数设置，预防过拟合具有很大的帮助。Corr-BP 出发点为趋势拟合，Pearson 关联度是一种无量纲的衡量标准，它与数据的尺度无关，因此在使用“提前终止法”结束训练设置网络训练参数更具有普适性；同时，相比于传统的 BP 误差，误差平面上不同的点在 Pearson 关联度平面上会反映为相同的值，这样给规避产生过拟合解带来了方便。

由于 Corr-BP 是对 BP 神经网络的改进，基本沿用了 BP 神经网络的结构，因而该模型依然存在数据量增加时收敛慢的问题，为了提高神经网络的收敛速度，在权重调整公式中增加动量项[14]，若用 代表某层权矩阵，则含有动量项的权重调整向量表达式如式(20)所示：

$$\boldsymbol{W}(t+1)=\boldsymbol{W}(t)+\eta\frac{\partial Corr}{\partial \boldsymbol{W}}+\alpha\Delta\boldsymbol{W}(t), \quad (20)$$

式中 $\alpha\in(0,1)$ 为动量系数，反映了以前积累的调整经验。动量项的取值为 0 时，权重修正只与当前梯度有关，当取值为 1 时则完全取决于上一次训练的结果，加入的动量项相当于阻尼项，能够减少学习过程的振荡趋势，从而改善了收敛性，因此应当尽量使其取较大的值，这一点与 BP 神经网络一致，有利于训练的稳定。

3　仿真实验和模型优化

为了验证 Corr-BP 对于时间序列预测算法的有效性，本文采用 5 个通用的非线性、非平稳随机时间序列数据集进行验证[15]。5 个用于实际的时间序列为：Sunspot database、Poland Electric Demand time series、ESTSP 2007_competition、Maunaloa、Forest Fires，其中前 3 个为单因素预测，后 2 个为多因素预测，实验所用的数据依时序按照 7∶2∶1 分为训练集、验证集和测试集，利用 Python3.5 对预测模型进行仿真，仿真平台为 Windows 7 操作系统，32 位，CPU 为 Pentium (R) E5300。为验证算法的有效性和适用性，对比实验选用文献[16]中改进的 RBF 神经网络(RBF)和改进的 BP 神经网络(BP)。

3.1　预测评价指标

预测模型对训练集进行预测而得出的准确率并不能很好地反映预测模型未来的性能，为了有效评价一个预测模型的性能，需要一组独立没有参加预测模型建立的数据集，并在该数据集上评价预测模型的准确率，这组独立的数据集叫做测试集。模型预测效果评价通常用相对/绝对误差(Relative / Absolute Error，R/AE)、平均绝对误差(Mean Absolute Percentage Error，MAPE)、均方误差(Root Mean Square Error，RMSE)等指标来衡量。

3.2 Corr-BP 模型拟合能力

首先验证模型对时间序列的拟合能力，仿真实验中的 Corr-BP 的参数设置参考经典 BP 神经网络，最大训练次数设置为 30 000，动量系数设为 0.9，学习率设为 0.01，初始权重设置为(0,1)之间的随机数，各神经元初始阈值统一设置为 0，训练终止条件为在测试集上的关联度与训练集上的差距在 0.05 以内，因为训练终止条件本身就要求要达到关联度 0.9 以上，故相差应设置小一些对其中两个数据集的拟合结果如图 3、图 4 所示。其中的 target 表示参考数据，test 表示经Corr-BP 建模后的测试数据结果，横坐标表示周为单位的时间，纵坐标表示序列值。从图中可以看出，Corr-BP 模型对时间序列的拟合能力较差，但已经能够表示出数据的变化趋势，这一问题的出现一是对于预测问题需要一定的泛化能力，对训练样本逼近程度越高，预测效果不一定越好，二是训练终止条件的设置仅凭经验，最终训练结果仅保证了关联度达到极高相关即停止训练，目的也是为了防止过拟合的出现。更加详细的对比见表 2，其中的 train_MAPE 表示训练阶段的平均相对误差，test_MAPE 表示预测阶段的平均相对误差。

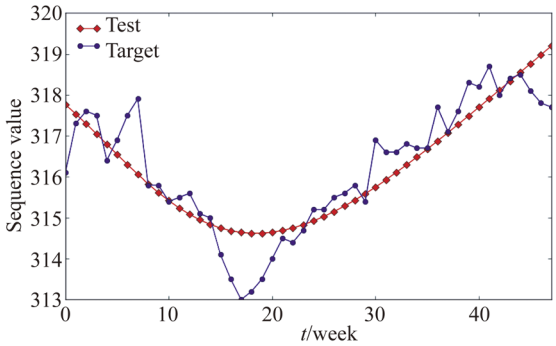


图 3　Maunaloa 拟合效果

Fig. 3　Maunaloa data fitting effect

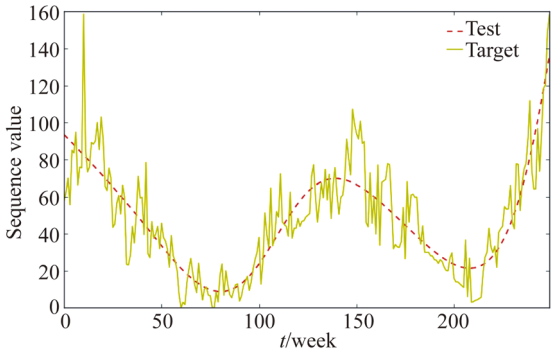


图 4　Sunspot 拟合效果

Fig. 4　Sunspot data fitting effect

表 2　各模型结果对比

Tab. 2　Comparison results of different models

数据集	算法	train_MAPE/%	test_MAPE/%	训练次数	运行时间/s
Sunspot	RBF	0.82	2.16	8 244	20.80
	BP	1.01	1.28	6 466	22.71
	Corr-BP	2.32	1.17	5 894	17.90
Electric	RBF	1.64	2.43	4 678	17.42
	BP	2.28	2.96	5 239	15.48
	Corr-BP	4.37	2.02	3 156	12.59
Maunaloa	RBF	0.31	2.03	7 456	26.45
	BP	0.49	2.45	6 802	29.43
	Corr-BP	0.62	0.49	6 489	18.98
ESTSP	RBF	5.48	6.32	16 854	65.62
	BP	5.38	6.58	18 752	72.54
	Corr-BP	6.84	6.57	30 000	110.23
Forest Fires	RBF	0.54	1.02	3 156	14.00
	BP	0.52	1.48	3 456	18.83
	Corr-BP	0.90	0.68	762	5.59

从表 2 中可以看出，相比于 RBF 神经网络和 BP 神经网络，Corr-BP 神经网络在函数逼近精度上存在明显的劣势，说明 Corr-BP 神经网络不适于用作函数拟合，同时可以看出，Corr-BP 神经网

络稳定性也劣于前两者，针对某些问题需要明显多的训练次数，性能波动较大。但 Corr-BP 的预测阶段误差则要优于前两者，和自身训练阶段的误差相比，具有相对较强的抗过拟合能力，同时也可以看出 Corr-BP 神经网络在 Maunaloa 和 Forest Fires 两个多因素的情况下，训练效率和逼近程度明显优于单因素，相对误差平均降低 0.94%，说明 Corr-BP 神经网络更适用于多因素预测问题。

3.3　Corr-BP 模型优化

从上一节的分析中可以看出，BP 神经网络的参数设置并不完全适用于 Corr-BP 神经网络，因此，本节将对 Corr-BP 神经网络中的关键步骤和参数设置进行讨论。以 ESTSP 的训练过程为例，训练过程中代价函数为参考输出与实际输出的关联度，其变化过程应当是逐渐增加直至超过 0.9 的，但实际训练过程中其关联度的变化却如图 5 所示。

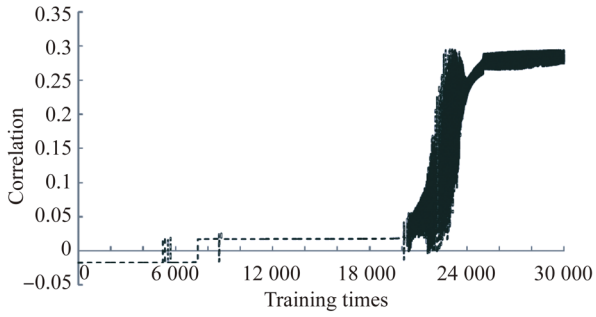


图 5　训练过程关联度变化

Fig. 5　Changes in relevance of training process

图中横坐标表示训练次数，纵坐标 Correlation 表示关联度，可以看出，训练过程中，关联度初期增长很慢，在 20 000 代左右才开始出现明显的提升，通过观察训练过程中权重的调整量 $\Delta \boldsymbol{W}$ 与权重 $\boldsymbol{W}$ 的数值可以发现，训练初期，$\boldsymbol{W}/\Delta \boldsymbol{W}$ 达到了 10^7，在这种情况下，对权重的调整几乎不受影响；直到 20 000 代左右，二者的比值才明显下降，训练也才开始初见成效，在训练次数达到 26 000 代左右，关联度远低于预期时，随着 $\Delta \boldsymbol{W}$ 对 $\boldsymbol{W}$ 调整的累加作用，虽然 $\Delta \boldsymbol{W}$ 已经达到 0.1 左右，但 $\boldsymbol{W}$ 的平均值已经增加到了 10^3 数量级，这样的情况下只能导致算法早熟，停止进化。

针对上述初始训练慢和后期部分连接权重过大的问题，将网络的初始权重设置为小权重，适当增加学习率同时添加正则项。加入正则项的网络代价函数如式(24)：

$$Cost = Corr(\boldsymbol{O}, \boldsymbol{d}) - \frac{\lambda}{2} \sum |\boldsymbol{W}|^2, \quad (24)$$

式中：后一项为 L2 正则项[18-19]，λ 为神经元的权重衰减比例系数，表示防止对训练样本学习过度需要对权重进行一定的衰减，同时防止训练初期某些神经元连接权重过早的固定，避免过拟合。在此基础上的权重变化量如式(25)：

$$\Delta w = \eta \frac{\partial Corr}{\partial w} - \eta \lambda w. \quad (25)$$

同时，针对训练后期连接权值过大的问题，对网络的连接权值按照式(20)进行门限设置，避免无限制增长。

图 6 为采取改进措施之后 Corr-BP 神经网络的进化曲线效果图，横坐标表示训练次数，纵坐标 Correlation 表示关联度。采取的方案是神经网络初始权重设置为(0，0.01)之间的随机数，同时神经元的权重衰减比例系数 λ 也设置为较大的 0.9，学习率调整为 0.2 并且对权值进行门限设置。可以看出训练能够达到要求，关联度不断进化至 0.9，总体的趋势是逐渐上升的，没有在训练早期陷入平坦。

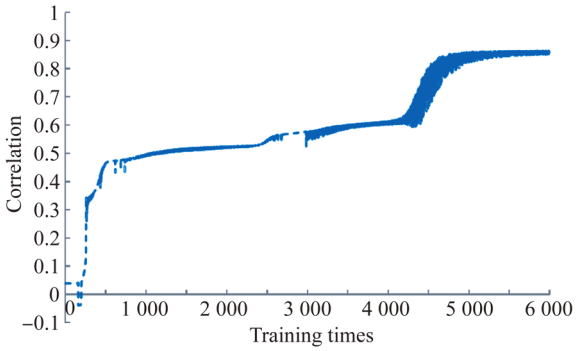


图 6　改进后的进化曲线

Fig. 6　Improved evolutionary curve

不同预测模型在预测阶段的差距较大，预测结果对比如图 7 所示。

图 7 中 WT-Corr-BP、ILR-Corr-BP、RBF 和 BP 分别代表权值门限 Corr-BP、增加学习率 Corr-BP、RBF 和 BP 各自的预测结果，横坐标为预测步长，纵坐标为序列值。从图中可以看出，在预测阶段使用 Corr-BP 比 RBF 和 BP 对参考信号的逼近程度都要好。在 RBF 和 BP 的曲线中都出现了比较明显的过拟合现象，具体的数据见表 3。

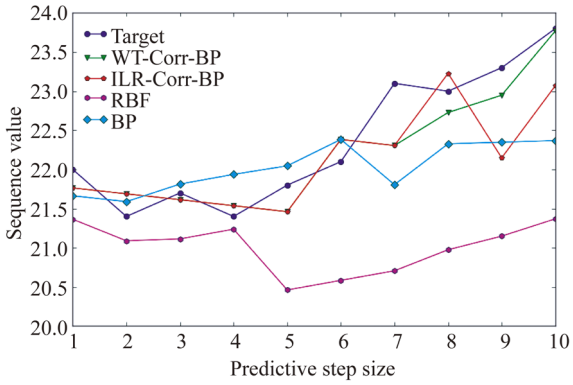


图 7　ESTSP 预测结果对比

Fig. 7　Comparison results of ESTSP prediction

表 3　性能数据对比

Tab. 3　Comparison of performance data

模型	训练 RMSE	测试 RMSE	训练次数
RBF	9.24	28.40	18 025
BP	9.35	18.84	18 752
ILR-Corr-BP	10.99	16.46	17 062
WT-Corr-BP	10.83	14.82	15 884

从表 3 中可以看出，各方法在训练阶段误差相差虽然并不明显，但 Corr-BP 仍然高于 RBF 和 BP 神经网络。测试阶段的误差 Corr-BP 要低于 RBF 和 BP，RMSE 值减少 4.02。在收敛速度方面对收敛速度有所提升，减少了 1 690 代，在预测方面对模型的误差控制要优于 RBF 和 BP 神经网络，其收敛速度则要明显优于 BP 算法。

4　结　论

本文提出了一种基于 Pearson 关联度的 Corr-BP 神经网络，利用关联度代替了 BP 神经网络的误差函数并采用梯度上升法对神经网络进行训练，对网络的参数设置进行了研究，在一系列标准预测数据集上进行的实验证明了 Corr-BP 神经网络在预防过拟合和训练效率上均优于传统的 BP 和 RBF 神经网络：测试阶段 Corr-BP 的 RMSE 值减少 4.02，训练次数减少 1 690 代；预测阶段的平均相对误差降低 0.94%，能够满足多因素时间序列预测问题的要求。

参考文献：

[1] LIU C X, SHU T, CHEN S. An improved grey neural network model for predicting transportation disruptions [J]. *Expert Systems with Applications*, 2016, 45: 331-340.

[2] TAN T, CHEN L, LIU F. Model of multiple seasonal autoregressive integrated moving average model and its application in prediction of the hand-foot-mouth disease incidence in Changsha [J]. *Journal of Central South University*, 2014, 39 (11): 1170-1176.

[3] 王民，赵渊，刘利，等. 基于量子粒子群优化广义回归神经网络的语音转换方法 [J]. 液晶与显示，2018，33(2):165-173.
WANG M, ZHAO Y, LIU L, *et al.*. Voice conversion based on quantum particle swarm optimization of generalized regression neural network [J]. *Chinese Journal of Liquid Crystals and Displays*, 2018, 33(2):165-173. (in Chinese)

[4] YU X L, GUO X L. Hourly photosynthetically active radiation estimation in midwestern united states from artificial neural networks and conventional regressions models [J]. *Int J Biometeorol*. 2016, 60 (8):1-13.

[5] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, *et al.*. Dropout: a simple way to prevent neural networks from overfitting [J]. *Journal of Machine Learning Research*, 2014, 15(1):1929-1958.

[6] 王林，彭璐，夏德，等. 自适应差分进化算法优化 BP 神经网络的时间序列预测 [J]. 计算机工程与科学，2015，37(12):2270-2275.
WANG L, PENG L, XIA D, *et al.*. BP neural network incorporating self-adaptive differential evolution algorithm for time series forecasting [J]. *Computer Engineering and Science*, 2015,37(12):2270-2275. (in Chinese)

[7] 雷雨，蔡宏兵，赵丹宁. 利用灰色关联极限学习机预报日长变化 [J]. 中国科学院大学学报，2015，32 (5): 588-593,604.
LEI Y, CAI H B, ZHAO D N. Prediction of length-of-day variation using grey relational analysis and extreme learning machine [J]. *Journal of University of Chinese Academy of Sciences*, 2015, 32 (5): 588-593,604. (in Chinese)

[8] 杜玉红，魏坤鹏，史屹君，等. 水质浊度红外光检测

及聚类灰色融合预测模型 [J]. 红外与激光工程, 2016, 45(10):264-270.
DU Y H, WEI K P, SHI Q J, *et al.*. Infrared detection and clustering grey fusion prediction model of water quality turbidity [J]. *Infrared and Laser Engineering*, 2016, 45(10):264-270. (in Chinese)

[9] BAI J, PERRON P. Estimating and testing linear models with multiple structural changes [J]. *Econometrics*, 2007, 75(2):459-502.

[10] SAAD Z, GLEN D G, BEAUCHAMP M, *et al.*. A new method for improving functional-to-structural MRI alignment using local Pearson correlation [J]. *Neuroimaging*, 2009, 44(3):839-848.

[11] NIU W, CHENG J, WANG G. Applications of extension grey prediction model for power system forecasting [J]. *Journal of Combinatorial Optimization*, 2013, 26(3):555-567.

[12] ARTHUR, MIRANDA, NETO. Pearson's correlation coefficient: a more realistic threshold for applications on autonomous Robotics [J]. *Computer Technology and Application*, 2014(2):69-72.

[13] GOLUB G H, LOAN C F V. An analysis of the total least squares problem [J]. *Siam Journal on Numerical Analysis*, 1980, 17(6):883-893.

[14] 丁思敏,吴军基. 改进模糊神经网络在负荷预测中的应用研究 [J]. 电力学报, 2009, 24(2):101-104.
DING S M, WU J J. Research on the use of improved fuzzy artificial neural network in load forecasting [J]. *Journal of Electric Power*, 2009, 24(2):101-104. (in Chinese)

[15] KARHUMEN J. Applications of Machine Learning Group[EB/OL]. (2015-10-21)[2018-01-15]. http://research.ics.aalto.fi/eiml/datasets.shtml.

[16] 张小红,王慧琴,于洪磊,等. 基于灰色相关分析的 GRFM 倾斜量预测模型 [J]. 西安建筑科技大学学报(自然科学版), 2016, 48(6):28-33.
ZHANG X H, WANG H Q, YU H L, *et al.*. GRFM forecasting model of inclination based on the grey relation analysis [J]. *Xi'an Univ. of Arch. & Tech (Natural Science Edition)*, 2016, 48(6):28-33. (in Chinese)

[17] 苗恩铭, 刘义, 董云飞, 等. 数控机床热误差时间序列模型预测稳健性的提升 [J]. 光学 精密工程, 2016, 34(10):2480-2489.
MIAO E M, LIU Y, DONG Y F, *et al.*. Improvement of forecasting robustness of time series model for thermal error on CNC machine tool [J]. *Opt. Precision Eng.*, 2016, 34(10):2480-2489. (in Chinese)

[18] 郭文月, 余岸竹, 刘海砚, 等. 正则化总体最小二乘用于光学线阵遥感影像定位 [J]. 光学 精密工程, 2017, 25(1):236-244.
GUO W Y, YU A ZH, LIU H Y, *et al.*. Regularized total least squares used in remote sensing image positioning of optical line array [J]. *Opt. Precision Eng.*, 2017, 25(1):236-244. (in Chinese)

[19] 王可, 王慧琴, 王展, 等. 基于改进 **R** 矩阵方法的光谱反射率重建研究 [J]. 液晶与显示, 2018, 33(4):78-86.
WANG K, WANG H Q, WANG ZH, *et al.*. Spectral reflectance reconstruction based on improved matrix **R** method [J]. *Chinese Journal of Liquid Crystals and Displays*, 2018, 33(4):78-86. (in Chinese).

作者简介:

王　可(1981－),男,山东荣成人,讲师,博士研究生,2004 年、2007 年于西南交通大学分别获得学士、硕士学位,主要从事数字图像处理及机器学习方面的研究。E-mail: wangke@xauat.edu.cn

王慧琴(1970－),女,山西长治人,教授,博士生导师,1992 年、1997 年于西安建筑科技大学分别获得学士、硕士学位,2002 年于西安交通大学获得博士学位,主要从事数字图像处理及信息管理系统方面的研究。E-mail: hqwang@xauat.edu.cn