

基于主成分分析的 BP 神经网络在形变预测中的应用*

陈兴权¹⁾ 王解先^{1,2)} 谷 川¹⁾

(¹⁾同济大学测量与国土信息工程系,上海 200092
(²⁾现代工程测量国家测绘局重点实验室,上海 200092)

摘 要 为提高形变预测的精度,提出将主成分分析与改进的 BP 神经网络相结合用于形变监测数据处理。通过编程实现该算法,并用实测数据进行验证,结果表明:与其他方法相比,基于主成分分析的改进 BP 神经网络能取得更好的预测效果。

关键词 形变;主成分分析;神经网络;预测;形变监测

中图分类号:TD173

文献标识码:A

APPLICATION OF BP NEURAL NETWORK BASED ON PRINCIPAL COMPONENT ANALYSIS IN DEFORMATION FORECASTING

Chen Xingquan¹⁾, Wang Jiexian^{1, 2)} and Gu Chuan¹⁾

(¹⁾Department of Surveying and Geo-Informatics, Tongji University, Shanghai 200092
(²)Key Laboratory of Modern Engineering Surveying of SBSM, Shanghai 200092)

Abstract In order to improve deformation forecasting precision, application of principal component analysis and improved BP neural network in deformation monitoring is proposed. For the verification if the new method can enhance the precision and reliability of forecasted data, the proposed algorithm is programmed and verified by use of measured data. The results show that compared with other methods, the improved BP neural network based on principal component analysis can achieve better forecasting results.

Key words: deformation; principal component analysis; neural network; forecasting; deformation monitoring

1 引言

随着现代科技发展特别是计算机技术的进步,各种理论和方法为形变预测提供了广泛的研究途径,如回归分析方法,时间序列分析方法,灰色模型分析方法,人工神经网络方法等,预测的精度和可靠性不断提高。人工神经网络(Artificial Neural Networks,简称 ANN)是目前比较常用的预测方法,它是由大量简单的基本元件——神经元相互连接,模拟人的大脑神经处理信息的方式,进行信息并行处理

和非线性转换的复杂网络系统。人工神经网络有如下几个突出的优点:可以充分逼近任意复杂的非线性关系;所有定量或定性的信息都等势分布贮存于网络内的各神经元,有很强的鲁棒性和容错性;采用并行分布处理方法,使得快速进行大量运算成为可能;可学习和自适应不知道或不确定的系统;能够同时处理定量、定性知识。目前神经网络诸多算法中应用最广的是 BP(Back Propagation)网络,其结构比较简单,操作性强,据统计,80%~90%的神经网络模型采用了 BP 网络或者它的变化形式^[1-9]。

* 收稿日期:2007-10-25

作者简介:陈兴权,男,1982年生,硕士研究生,主要从事卫星大地测量研究。E-mail:chenxqagua@163.com

主成分分析 (Principal Component Analysis, 简称 PCA) 最早是由 Pearson 在 1901 年的生物学理论研究中引入, 后来在多元统计中广泛应用。PCA 的实质是 n 维空间的坐标旋转, 并不改变样本数据结构, 得到的主成分是原变量的线性组合且两两不相关, 能够最大程度地反映元变量所包含的信息, 在以一定标准选取前 K 个较重要的主成分之后, 原来的多维问题得以简化^[5]。

本文结合主成分分析法和 BP 网络的优点, 将这两种方法相结合用于形变预测, 并通过 Matlab 编程实现, 将预测结果与灰色模型和传统的 BP 神经网络预测得到的结果进行比较, 结论是基于主成分的 BP 神经网络预测结果较好。

2 主成分 BP 网络模型

2.1 主成分分析确定训练网络节点

假设某沉降监测网有 n 个沉降点, 观测了 m 期, 将各点每期高程表示 $n \times m$ 的矩阵 \mathbf{H} , 其元素 $h(i, j)$ ($i=1, 2, \dots, n$ $j=1, 2, \dots, m$) 为 i 点 j 期的高程, 合理假定 $m > n$ 。

构成协方差阵 \mathbf{B} :

$$\mathbf{B} = \mathbf{H}\mathbf{H}^T \quad (1)$$

显然 $\mathbf{B}(n \times n)$ 是实对称正定矩阵, 可分解为:

$$\mathbf{B} = \mathbf{D} \mathbf{\Lambda} \mathbf{D}^T \quad (2)$$

式中的 $\mathbf{\Lambda}$ 为 n 个特征根 ($\lambda_1, \lambda_2, \dots, \lambda_n$) 组成的对角阵, \mathbf{D} 的 n 个列为互相正交的特征向量 \mathbf{d}_k ($k=1, 2, \dots, n$), 它是 n 维向量的一组基, 对应于大特征根的特征向量称为主成分。

矩阵 \mathbf{H} 的 j 列 (即第 j 期高程) \mathbf{H}_j^T 可以表示为:

$$\mathbf{H}_j^T = \mathbf{a}\mathbf{D}^T \quad (3)$$

式中系数矩阵 \mathbf{a} 由 n 个值组成, 可由下式求得:

$$\mathbf{a} = \mathbf{H}_j^T \mathbf{D} \quad (4)$$

若提取主要形变信息, 利用主成分的贡献率求出需要选取的主成分个数, 将特征值从大到小排列, 贡献率公式为

$$p = \frac{\sum_{i=1}^k \lambda_i}{\sum_{j=1}^n \lambda_j} \quad (5)$$

如果前 k 个主成分的累积贡献率已到达 85%, 表明前 k 个主成分基本包含了几乎全部测量指标所具有的信息。将前 k 个特征值对应的特征向量保持不变, 其余的特征向量取为 0, 利用公式 (4) 算出 \mathbf{a}' , 再利用公式 (3) 算出 \mathbf{H}_j^T , 然后将变换后的 \mathbf{H} 作为神经网络的训练样本。

2.2 改进的 BP 神经网络

BP 神经网络通常由输入层、隐含层和输出层组成, 其信息处理过程由向前传播与向后学习两部分

组成, 网络学习的规则是误差从输出层到输入层向后传播并修正样本的过程, 学习的目的是使网络的实际输出逼近目标样本, BP 网络的典型结构如下:

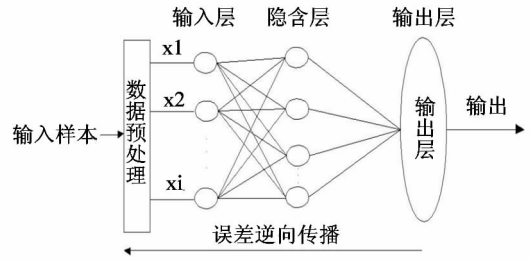


图1 BP神经网络结构图

Fig. 1 Structure of BP neural network

本文将 BP 神经网络进行了一些改进, 首先, 采用小波神经网络的思想将激励函数进行一些改进, 使小波变换良好的时频局域化特性和神经网络的自学习功能相结合, 充分发挥其函数逼近能力。小波基函数选用 Morlet 母函数, 因为 Morlet 小波是复值小波, 能够提取信号中的幅值和相位信息^[4], 公式如下:

$$h(x) = \cos(5x) e^{-\frac{x^2}{2}} \quad (6)$$

Morlet 母函数如图 2 所示:

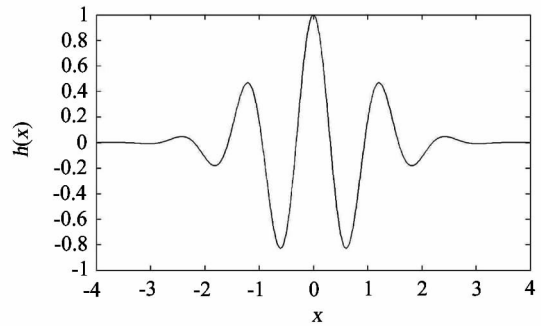


图2 Morlet母小波

Fig. 2 Morlet mother wavelet

设输入样本个数为 R , 输出节点个数为 N , 第 p 个样本第 n 个节点的输入可表示为:

$$f_{n,p} = \left[f \sum_{k=1}^K \omega_{n,k} \sum_{m=1}^M S^p(x_m) h\left(\frac{x_m - b_k}{a_k}\right) \right] \quad (7)$$

其中: K 为隐含层单元数; M 为输入层单元数; $\omega_{n,k}$ 为隐含层第 k 个单元与输出层第 n 个单元之间的连接权值, a_k, b_k 分别为小波伸缩因子和平移因子, $S^p(x_m)$ 为输入信号。

其输出层函数如下:

$$f(z) = \frac{1 - e^{-z}}{1 + e^{-z}} \quad (8)$$

$f(z)$ 为对数 S 型 (Sigmoid) 激活函数, 可将输出值限制在 -1 和 1 之间。

其次, 利用附加动量法可降低网络对于误差曲

面局部细节的敏感性,有效抑制网络陷于局部极小。该方法是在反向传播法的基础上在每一个权值的变化上加上一项正比于前次权值变化量的值,并根据反向传播法来产生新的权值变化。传统 BP 算法权值修正的公式为:

$$\omega_{i,j}(t+1)=\omega_{i,j}(t)+\eta\Delta\omega_{i,j}(t+1)$$

(9)

带有附加动量因子的权值调节公式为:

$$\Delta\omega_{i,j}(t+1)=\eta\Delta\omega_{i,j}(t+1)+\alpha\Delta\omega_{i,j}(t)$$

(10)

其中, t 为训练次数, η 为网络的学习效率, α 为动量系数。

2.3 主成分神经网络算法的实现步骤

1) 进行数据预处理时,为了避免量纲不同而带来数据间无意义的比较,将数据进行标准化处理:

$$X'_i=\frac{1.8(X_i-X_{\min})}{\Delta X_i}-0.9$$

(11)

其中: X_i 为原始指标样本, X_{\min} 为最小值, ΔX_i 为的最大值与最小值之差。这样可使样本数据归划到(- 0.9,0.9) 区间;

2) 对标准化的数据进行主成分分析,确定神经网络输入节点。即公式(1) ~ (5);

3) 给 BP 网络进行初始化,并设置网络学习效率 η ,动量系数 α ,迭代容许误差 e ;

4) 为网络提供一组学习样本,将输入节点作为 BP 网络的输入向量,同时输入期望 $f_{n,p},n=1,2,\cdots,N;P=1,2,\cdots,R$;

5) 网络的自学习过程,利用当前网络参数计算网络的实际输出,如果输出误差小于容许误差 e 或到达限定的迭代次数,则训练结束;

6) 否则进行误差反向传播,使权值沿误差函数的负梯度方向改变,利用梯度下降法求网络参数的变化。利用带有附加动量因子的权值修正网络参数,见公式(10)。再转至第 4 步;

7) 采用训练好的改进的 BP 神经网络以及预测样本进行预测;

8) 将预测的数据进行反标准化处理,并与实测样本进行比较。

基于主成分分析的 BP 网络与传统网络相比,主要区别在于把主成分和神经网络相结合,使其预测效果更佳;以及使用非线性小波基 Morlet 函数取代了神经网络的非线性 Sigmoid 函数。主成分神经网络有如下优点:

1) 利用主成分分析提取形变数据的主要信息使多维问题得以简化,同时剔除了数据里的噪声误差,有利于神经网络平滑拟合,并且大大加快了神经网络的训练速度;

2) 一般的形变分析是对某一个点形变数据进行分析,但是实际工程中是对监测物的几个点进行分析,监测点间必定存在一定的联系,而主成分分析可以对几个样本同时处理,所以用主成分分析提取输入样本有一定的科学依据;

3) 监测点间存在复杂的非线性关系,BP 神经网络的泛化功能和很强的容错能力为处理复杂非线性关系提供强有力的工具;

4) 输入函数利用 Morlet 小波母函数进行改进,由于 Morlet 函数具有紧支撑性,神经元之间的相互影响小,使得神经网络具有更快的收敛速度;

5) 相对于传统的权值修正公式,改进后的权值修正公式增加了动量项,这对于防止震荡和加速收敛起到了巨大作用。

3 算例分析

本文根据基于主成分的 BP 神经网络思想,采用 Matlab 语言编程实现上述算法,并用多组数据对其算法进行验证,神经网络学习后的输出结果与期望输出基本相符。下面利用某水库大坝 4 个监测点 1 ~ 70 期数据进行验证,对其第 71 ~ 75 期沉降变形量数据进行预测,并且将预测的结果与灰色模型和传统 BP 神经网络预测的结果进行比较,见表 1 ~ 4。

图 3 可以清楚地显示出基于主成分分析的 BP 网络与灰色模型和传统的 BP 网络预测数据能力的区别。

表 1 点 1 预测结果比较(单位:mm)
Tab.1 Comparison among forecast results with different models of point 1(unit:mm)

预测期数	实测值	灰色模型		传统 BP 神经网络		基于主成分的 BP 神经网络	
		预测值	绝对误差	预测值	绝对误差	预测值	绝对误差
71	-49.4	-47.71	0.31	-49.78	0.38	-49.79	0.39
72	-50.9	-48.24	-0.66	-50.54	-0.36	-50.63	-0.27
73	-51.2	-48.80	-0.80	-51.39	0.19	-51.57	-0.37
74	-53.0	-49.41	-0.79	-52.12	-0.88	-52.78	-0.22
75	-53.7	-50.02	-0.88	-52.93	-0.77	-54.18	0.48

表 2 点 2 预测结果比较(单位:mm)

Tab.2 Comparison among forecast results of point 2(unit:mm)

预测期数	实测值	灰色模型		传统 BP 神经网络		基于主成分的 BP 神经网络	
		预测值	绝对误差	预测值	绝对误差	预测值	绝对误差
71	-37.7	-37.78	0.08	-37.82	0.12	-37.78	0.08
72	-38.6	-38.30	-0.30	-38.33	-0.27	-38.62	0.02
73	-39.9	-38.97	-0.93	-39.02	-0.88	-39.38	-0.52
74	-40.3	-39.75	-0.55	-40.93	0.63	-40.46	0.16
75	-41.7	-40.59	-1.11	-42.71	1.01	-41.58	-0.12

表 3 点 3 预测结果比较(单位:mm)

Tab.3 Comparison among forecast results of point 3(unit:mm)

预测期数	实测值	灰色模型		传统 BP 神经网络		基于主成分的 BP 神经网络	
		预测值	绝对误差	预测值	绝对误差	预测值	绝对误差
71	-40.3	-40.74	0.44	-40.74	0.44	-40.57	0.27
72	-41.1	-40.69	0.59	-41.53	0.43	-40.98	-0.12
73	-41.9	-42.65	0.75	-42.21	0.31	-41.72	-0.18
74	-43.2	-43.64	0.44	-42.64	-0.56	-42.92	-0.28
75	-43.6	-44.64	1.04	-42.81	-0.79	-44.18	0.58

表 4 点 4 预测结果比较(单位:mm)

Tab.4 Comparison among forecast results of point 4(unit:mm)

预测期数	实测值	灰色模型		传统 BP 神经网络		基于主成分的 BP 神经网络	
		预测值	绝对误差	预测值	绝对误差	预测值	绝对误差
71	-45.4	-45.55	0.15	-45.52	0.12	-45.74	0.34
72	-46.9	-46.60	-0.30	-46.56	-0.33	-46.84	-0.06
73	-47.8	-47.72	-0.08	-47.43	-0.36	-47.59	-0.13
74	-48.2	-48.87	0.67	-48.91	0.71	-48.41	0.21
75	-48.6	-50.05	1.45	-49.58	-0.98	-49.48	0.88

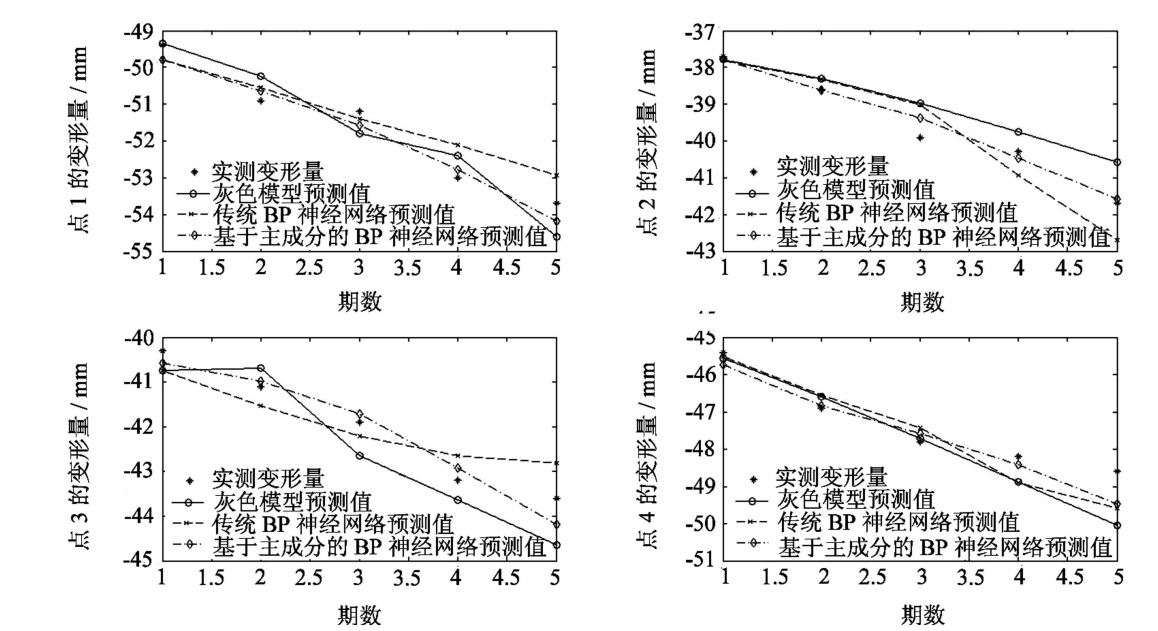


图 3 各点预测结果比较

Fig.3 Comparison among forecasting results all points

由图 3 可见,基于主成分的 BP 神经网络的预测的绝对误差基本都在 $\pm 0.4\text{ mm}$ 内,能够满足形变监测的需要,与其他两种方法相比预测的绝对误差较小,特别是超过 4 期时灰色模型和传统的 BP 神经网络预测的绝对误差会较大。表 5 列出各方法预测结果的中误差以及平均预测中误差。

表 5 各方法预测结果中误差(单位:mm)
Tab. 5 MS of forecasting results with different methods(unit:mm)

预测方法	点 1 预 测 MS	点 2 预 测 MS	点 3 预 测 MS	点 4 预 测 MS	平均预 测 MS
灰色模型	0.72	0.71	0.69	0.73	0.71
传统 BP 神经网络	0.58	0.68	0.53	0.59	0.60
PCA-BP 神经网络	0.36	0.25	0.33	0.44	0.35

基于主成分的 BP 神经网络与传统的 BP 网络相比,基于主成分的 BP 网络的训练时间要比传统的 BP 网络短,而且迭代的次数较少。在进行预测时,传统的 BP 神经网络偶尔陷入极小,而基于主成分的 BP 网络几乎不会陷入极小。

表 6 传统 BP 神经网络与 PCA-BP 神经网络比较
Tab. 6 Comparison between traditional BP ANN and PCA-BP ANN

比较项	平均训练次数	平均训练时间(s)
传统的 BP 神经网络	23	2.796 9
PCA-BP 神经网络	7	1.515 6

4 结论

由本文的模型和算例可以得出,基于主成分分析的 BP 神经网络的预测精度优于其他两种方法。利用主成分分析提取形变的主要信息是去噪的过程,将去噪后的样本数据进行神经网络训练,其学习速度比没有经过主成分分析的样本要快。同时,将 BP 网络进行了一些改进,利用小波 Morlet 函数作为神经网络激励函数,以及利用附加动量法降低网络对于误差曲面局部细节的敏感性,有效的抑制网络陷于局部极小,其收敛性比灰色模型和传统的 BP 网络要好。由此可以得出结论,基于主成分分析的 BP 神经网络能较好地用于形变预测。

参 考 文 献

1 贺清碧,周建丽. BP 神经网络收敛性问题的改进措施[J]. 重庆交通学院学报,2005,2:143-145.
1 He Qingbi and Zhou Jianli. The convergence and improvements of BP neural network[J]. Journal of Chongqing Jiao-

tong University, 2005,2:143-145. (in Chinese)
2 潘国荣,谷川. 变形监测数据的小波神经网络预测方法[J]. 大地测量与地球动力学,2007,27(4):47-50.
2 Pan Guorong and Gu Chuan. Wavelet neural network prediction method of deformation monitoring data[J]. Journal of Geodesy and Geodynamics,2007,27(4):47-45. (in Chinese)
3 飞思科技产品研发中心. MATLAB6.5 辅助小波分析与应用[M]. 北京:电子工业出版社,2005.
3 Study and Development Center of Fecit Science and Technology Product. Matlab6.5 assistance wavelet analyses and application[M]. Beijing: Publishing House of Electronice Industry,2005. (in Chinese)
4 郑治真,等. 小波变换及其 MATLAB 工具的应用[M]. 北京:地震出版社,2001.
4 Zheng Zhizhen,et al. Application of wavelet transform method and MATLAB Tool[M]. Beijing: Publishing House of Earthquake,2001. (in Chinese)
5 杨超,李志强. 基于主成分-神经网络相集成的多策略评价赋权方法[J]. 当代财经,2005,7:124-128.
5 Yang Chao and Li Zhiqiang. The more strategy estimate method based on the principal component-neutral networks[J]. Contemporary Finance & Economics, 2005,7:124-128. (in Chinese)
6 飞思科技产品研发中心. 神经网络理论与 MATLAB7 实现[M]. 北京:电子工业出版社,2005.
6 Study and Development Center of Fecit Science and Technology Product. The implement of neutral networks theory and Matlab7[M]. Beijing: Publishing House of Electronice Industry,2005. (in Chinese)
7 林国玺,宣慧玉. 遗传算法和 BP 人工神经网络在税收预测中的应用[J]. 系统工程理论方法应用,2005,4:145-148.
7 Lin Guoxi and Xuan Huiyu. Application research on tax forecasting based on genetic algorithms and BP artificial neural networks[J]. Systems Engineering - Theory Methodology Applications, 2005,4:145-148. (in Chinese)
8 陈桦,程云艳. BP 神经网络算法的改进及在 MATLAB 中的实现[J]. 陕西科技大学学报,2004,4:45-47.
8 Chen Hua and Cheng Yunyan. Design and realize the software of machining techniques manual[J]. Journal of Shanxi University of Science & Technology,2004,4:45-47. (in Chinese)
9 李杰,等. 利用小波变换分析跨断层形变异常[J]. 大地测量与地球动力学,2004,24(3):73-78.
9 Li Jie, et al. Analysis of anomaly of cross-fault deformation after wavelet transform method[J]. Journal of Geodesy and Geodynamics, 2004,24(3):73-78. (in Chinese)