

文章编号: 1672-8785(2006)11-0001-04

人工神经网络 BP 算法在近红外光谱分析中的应用

张 银, 周孟然

(安徽理工大学电气工程系, 安徽淮南 232001)

摘 要: 简要介绍了人工神经网络及其在近红外光谱分析中的应用。给出了三层 BP 网络的基本结构和对应的 BP 算法。通过测定润滑油的流动性、苯和甲苯的回收率及相对标准偏差的两个实例, 分析了人工神经网络 BP 算法的明显优势。最后指出了存在的部分问题。

关键词: 人工神经网络; BP 算法; 近红外光谱; 应用

中图分类号: O657 **文献标识码:** A

Application of Artificial Neural Network BP Algorithm in Near Infrared Spectroscopy

ZHANG Yin, ZHOU Meng-ran

(Department of Electric Engineering, Anhui University of Science and Technology, Huainan 232001, China)

Abstract: Artificial neural network and its application in near infrared spectroscopy are presented in brief. The basic structure of the BP network and its corresponding BP algorithm are given. The obvious superiority of the artificial neural network BP algorithm is shown through the determination of the fluidity of lubricating-oil and the recovery and RSD of benzene and methylbenzene. Finally, several existing problems are indicated.

Key words: artificial neural network (ANN); BP algorithm; near infrared spectroscopy; application

1 人工神经网络及其在近红外光谱分析中的应用

“人工神经网络”或“神经网络”是指用大量的简单计算单元(即神经元)构成的非线性系统,它是建立在现代神经科学研究成果基础上的一种抽象的数学模型,它在一定程度和层次上模仿了人脑神经系统的信息处理、存储及检索功能,因而具有学习、记忆和计算等智能处理功能。人工神经网络由于具有强大的信息处理能力,在电子、数学、物理、化学、生物工程、

新材料等领域都有着广泛的应用前景。

红外光谱分析中应用的人工神经网络是光谱分析中应用得最广泛的模式识别技术。Tanabe 等人利用神经网络系统对 1129 个红外谱图进行了识别,该系统由两部分组成,能在 0.1s 内鉴别未知谱。王志有等人采用 BP 人工神经网络光度法对复合维生素片中的 VB₁、VB₂、VB₆ 和 VPP 进行了同时测定,得到了较为准确的结果。张卓勇等人论述了人工神经网络在光谱分析重叠信号解析中的应用,他们将基于计算最大差异光谱的目标转换因子分析法用于解析混

收稿日期: 2006-06-02

基金项目: 国家自然科学基金资助项目(50574005)、安徽省教育厅自然科学基金资助项目(2005KJ081)

作者简介: 张银(1981—),男,安徽淮南市人,在读硕士研究生,研究方向为自动化控制与激光光谱技术。

合物的红外光谱和从混合物的红外光谱中解析纯组分光谱,得到了满意的结果。齐锋等人研究利用人工神经网络方法提高差分光学吸收光谱系统测量精度,对试验结果进行了比较,收到了良好的效果。该技术已被应用于广西的两个国家重点监测城市桂林和北海,并成功地向国家环境监测总站上报了数据。高建波等人提出的小波变换(WT)和人工神经网络相结合的方法,克服了光谱中由背景干扰所造成的数据处理困难,充分发挥了开放光路取样方法的优越性:

- ①通过 WT 数学方法扣除了样品光谱的背景干扰;
- ②利用人工神经网络技术消除了光谱中的非线性成分并获得了光谱透光度和组分浓度之间的定量关系;
- ③能够进行多组分气体的在线连续定量测量。在傅里叶变换红外光谱定量分析中,应用人工神经网络可以得到比经典最小二乘法更高的灵敏度^[1-5]。

2 BP 模型的基本结构与算法

神经网络的连接方式有很多种,使用较多的是 1986 年由 Rumelhart 等人提出的反向传输模型,即 BP(back-propagation)模型。该模型可以实现数据处理的定量预测,也可以用于模式识别。在近红外光谱的光谱数据处理中,无论是样品组成、质量指标的预测,还是模式识别分类,都可以用 BP 模型来完成。

2.1 BP 模型的基本结构

如图 1 所示, BP 模型由输入层、隐蔽层和输出层三部分组成。图中圆圈代表神经元。模型是将光谱数据的前四维主成分信息作为神经网络的 4 个输入层结点,每个主成分都是所有波长点的线性组合。它包括结点,上一层的每一个结点都同下层的每一个结点连接,同外界的联系通过输入、输出层的结点进行,隐含于内部的中间层决定系统的计算能力。在 ANN 的每条连接线上是权重因子,代表系统记忆能力的重要部分。输入层输入的变量经权重加和后进入隐含层进行非线性变换,隐含层的输出经权重加和后进入输出层,经线性变换得到输出结果。在网络训练过程中,权重的数值可随着正在网络

中流通的新信息而改变,输出的预测值(Y)与期望值(y)相比较,所得误差经反向传输,调整权重,使目标变量 $E = (y - Y)^T(y - Y)$ 为最小(T 为矩阵的转置),最终得到的权重用于人工神经网络的预测。

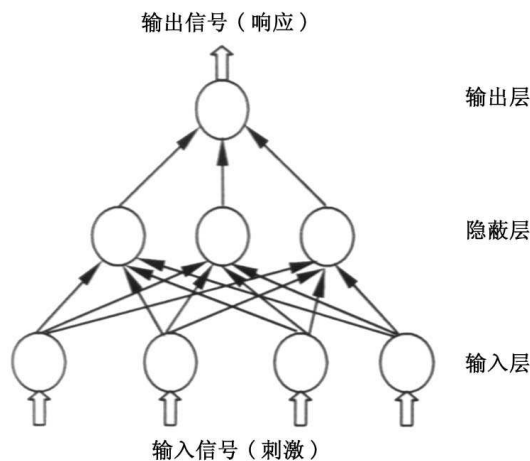


图 1 BP 模型的基本结构

2.2 BP 算法

假设 BP 网的输入矢量为 $x \in R^n$, 其中 $x = (x_0, x_1, \dots, x_{n-1})^T$; 第 1 隐层有 n_1 个神经元, 它们的输出为 $x' \in R^{n_1}$, $x' = (x'_0, x'_1, \dots, x'_{n_1-1})^T$; 第 2 隐层有 n_2 个神经元, 它们的输出为 $x'' \in R^{n_2}$, $x'' = (x''_0, x''_1, \dots, x''_{n_2-1})^T$; 输出层有 m 个神经元, 输出 $y \in R^m$, $y = (y_0, y_1, \dots, y_{m-1})^T$ 。设输入层到第 1 隐层的权为 w_{ij} , 阈值为 θ_j ; 第一隐层到第二隐层的权为 w'_{jk} , 阈值为 θ'_k ; 第二隐层到输出层的权为 w''_{kl} , 阈值为 θ''_l , 于是各层神经元的输出为

$$\begin{cases} y_l = f\left(\sum_{k=0}^{n_2-1} w''_{kl} x''_k - \theta''_l\right), & l = 0, 1, \dots, m-1 \\ x''_k = f\left(\sum_{j=0}^{n_1-1} w'_{jk} x'_j - \theta'_k\right), & k = 0, 1, \dots, n_2-1 \\ x'_j = f\left(\sum_{i=0}^{n-1} w_{ij} x_i - \theta_j\right), & j = 0, 1, \dots, n_1-1 \end{cases}$$

显然, 它将完成 n 维空间矢量到 m 维空间的映射。其网络训练过程如图 2 所示。

3 BP 算法在近红外光谱数据处理中的应用

在近红外光谱定量分析中,一部分分析采用线性数据处理工具能够得到满足要求的预测结果,例如对于燃料油的光谱数据分析,使用多元线性回归(MLR)、主成分回归(PCR)、偏最小二乘法(PLS)等能够满足要求,但对于像润滑油这样的组成结构复杂、相对分子质量较大的分析对象,它们的性质与其近红外光谱信息之间往往是非线性的关系,所以用线性数学处理工具产生的误差很大,而采用人工神经网络来处理数据则可以得到很好的预测结果。表 1 和表 2 是分别使用偏二乘法和人工神经网络 BP 算法分析 1000nm~1600nm 波段数据后获得的不同润

滑油的流动性能的结果。从表中可以看出,用人工神经网络法获得的结果较好^[6]。

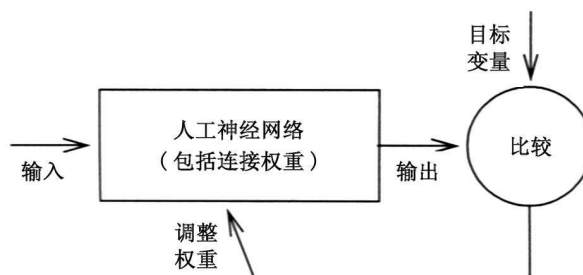


图 2 网络训练过程

再如:由于苯和甲苯化学结构相近,两者的

表 1 偏最小二乘法预测润滑油流动性能

序号	粘度 (40 °C)/(mm ² /s)			粘度 (100 °C)/(mm ² /s)			粘度指数		
	实测值	预测值	误差	实测值	预测值	误差	实测值	预测值	误差
1	139.0	159.0	-11.9	14.40	12.98	1.42	100.0	93.5	6.5
2	141.0	118.8	22.2	14.50	16.34	-1.84	100.0	119.9	-19.9
3	110.0	108.6	1.4	14.47	15.24	-0.77	135.0	138.9	-3.9
4	140.0	105.0	35.0	14.50	16.77	-2.27	107.0	108.6	-1.6
5	106.9	84.0	22.9	16.50	18.58	-2.08	168.0	150.3	17.7

表 2 人工神经网络 BP 算法预测润滑油流动性能

序号	粘度 (40 °C)/(mm ² /s)			粘度 (100 °C)/(mm ² /s)			粘度指数		
	实测值	预测值	误差	实测值	预测值	误差	实测值	预测值	误差
1	139.0	139.25	0.25	14.40	14.45	0.052	100.0	99.68	-0.32
2	141.0	141.18	0.18	14.50	14.48	-0.024	100.0	100.30	0.30
3	110.0	110.21	0.21	14.47	14.48	0.012	135.0	135.49	0.49
4	140.0	139.96	-0.04	14.50	14.47	-0.035	107.0	107.02	0.02
5	106.9	106.92	0.02	16.50	16.55	0.052	168.0	167.86	-0.14

表 3 苯和甲苯的回收率和相对标准偏差

序号	样品值 (μg/L)		加入量 (μg/L)		输出值 (μg/L)		回收率 (%)	
	苯	甲苯	苯	甲苯	苯	甲苯	苯	甲苯
1	2.82	12.01	3.52	3.16	6.21	15.08	96.5	97.3
2	6.34	27.81	3.52	1.58	9.91	29.42	101.5	102.2
3	8.80	7.90	7.04	6.32	15.79	14.29	99.4	101.2
4	13.38	24.65	7.04	6.32	20.22	30.69	97.1	95.7
5	18.30	5.69	10.56	9.48	28.73	15.21	98.8	100.4
平均	—	—	—	—	—	—	98.7	99.4
RSD(%)	—	—	—	—	—	—	2.0	2.8

光谱严重重叠, 实验配制 5 份与训练样本不同的人工合成样品溶液, 按实验方法测定各自不同波长下的紫外吸光度。将测得的数据输入人工神经网络模型, 用 BP 算法算出输出值后, 计算回收率及相对标准偏差, 数据及结果见表 3。测得的苯和甲苯的回收率分别为 98.7% 和 99.4%, 测定结果的相对标准偏差分别为 2.0% 和 2.8%, 取得了满意的效果^[7]。

4 结语

从上面的分析结果中可以看出, 人工神经网络在近红外光谱分析中有着不可替代的作用, 特别是在解决非线性数据处理方面优势明显。但仍然存在的许多问题值得探讨, 如网络训练速度、过拟合、全局寻优、网络结构与网络能力等问题都存在着未知因素。相信在不久的将来, 关于人工神经网络的方法、参数、训练样本、学习过程以及预测结果等都会得到更深入的发展, 人工神经网络在红外光谱分析中的应用也必将得到更广泛的推广。

参考文献

[1] 许禄, 胡昌玉. 化学中的人工神经网络法[J]. 化学进展, 2000, 12(1): 18-31.

- [2] 王志有, 等. BP 人工神经网络 - 光度法同时测定 VB₁、VB₂、VB₆ 和 VPP [J]. 鞍山师范学院学报, 2005, 7(4): 36-38.
- [3] 张卓勇, 刘思东. 反向传播神经网络在分析化学重叠信号解析中的应 [J]. 抚顺石油学院学报, 1996, 16(3): 17-2.
- [4] 齐锋, 等. 利用人工神经网络方法提高差分光学吸收光谱系统测量精度研究 [J]. 光学学报, 2002, 22(11):
- [5] 高建波, 胡东成. 小波变换和神经网络用于红外光谱定量分析 [J]. 清华大学学报 (自然科学版), 2001, 41(3): 121-124.
- [6] 冯斯沛, 史永刚, 等. 近红外光谱及其在石油产品分析中的应用 [M]. 北京: 中国石化出版社, 2002.
- [7] 吴军, 等. 神经网络辅助紫外光谱测定苯和甲苯的含量 [J]. 新疆医科大学学报, 2005, 28(6):
- [8] 相燕. 神经网络在分析化学中的应用效果研究 [J]. 江汉石油学院学报, 2004, 26(6):
- [9] Norgaard L, Ssudland A, Wagner J, et al. Interval Partial Least-squares Regression (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy []. 2000, 54(3):
- [10] Saarinen P, Kauppinen J. Multicomponent Analysis of FT-IR Spectra [J]. Applied Spectroscopy, 1991, 45:

国外专利介绍

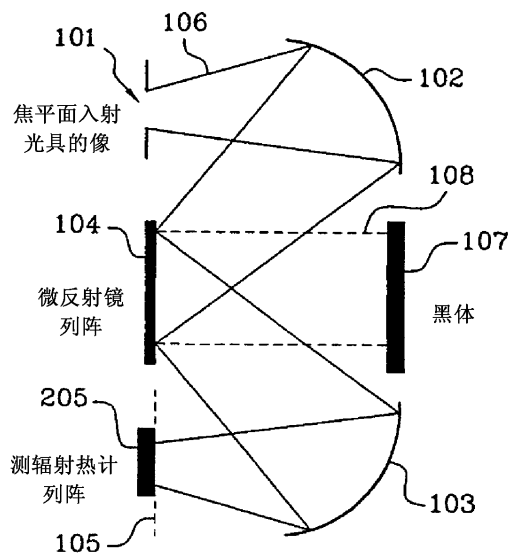
非致冷光学成像器件

美国专利 US7091488

(2006 年 8 月 15 日授权)

为了获得一种无需致冷但具有适当特性的光学成像器件特别是在红外波段工作的光学成像器件, 本发明提出了一种无需采用低温装置的红外成像器件。该红外成像器件通过用一系列 MEM 微反射镜对入射在传感器阵列上的光束进行调制, 使其热噪声的影响减小到了最低限度。它特别适合在摄像机和望远镜中用于红外成像。

本专利说明书共 5 页, 其中有 1 张插图。



高编译