

基于两阶段优化算法的神经网络预测模型^①

邹昊飞¹, 夏国平¹, 杨方廷²

(1. 北京航空航天大学经济管理学院, 北京 100083; 2. 北京仿真中心, 北京 100854)

摘要:采用基于两阶段优化算法(multi-stage optimization approach, MSOA)的 GA 人工神经网络, 将训练集分为两部分, 在前一训练集训练后获得的网络基础上使用后一训练集进行进一步的训练获得更为优化的网络结构. 针对复杂系统建模输入节点难以确定的问题, 提出将其与自组织数据挖掘算法相结合, 利用 GMDH 算法获得神经网络的初始化节点, 使用训练好的神经网络模型进行预测. 将由此建立的预测模型应用于粮食价格的预测, 并进一步探讨了 MSOA 算法的收敛性问题. 结果表明基于 GMDH 和 MSOA 的神经网络组合预测模型能较大提高神经网络的全局收敛能力和收敛速度, 提高预测精度.

关键词: GMDH; MSOA; 遗传算法; 人工神经网络

中图分类号: TP183

文献标识码: A

文章编号: 1007-9807(2006)05-0028-08

0 引言

神经网络作为一种重要的人工智能技术, 近几年在财务风险评价、经济时间序列预测、组合证券优化、股票预测等经济领域得到应用研究, 取得了很好的效果^[1-3]. 在经济领域应用得最广的神经网络是 BP 网络, 但由于 BP 网络具有: ①学习过程收敛慢; ②容易陷入局部极小; ③鲁棒性不好, 网络性能差等缺点, 在实际应用过程中, 它总是和遗传算法(GA)相结合使用^[4-6]. 本文在此基础上用二阶段优化算法(the multi-stage optimization approach, MSOA)来优化三层 BP 神经网络的连接权和网络结构. 针对在复杂系统建模过程中 BP 网络输入变量难于确认的问题, 将其与自组织数据挖掘算法结合, 建立模型系统的优化算法(GMDH-GAANN). 这种新模型提高了神经网络的全局收敛能力和收敛速度, 增强了网络的适应能力, 使之能更有效地应用于复杂系统的建模, 将其应用到国家粮食安全预警系统模型中得到较好的预测效果.

1 自组织数据挖掘方法概述

乌克兰科学院 Ivakhnenko 院士提出复杂系统建模的自组织理论和方法^[4], Mueller 等人在 20 世纪 90 年代对该理论进一步发展, 形成了完备的自组织数据挖掘理论^[7], 很好地解决复杂动态系统的建模和预测问题. 自组织方法抛开了诸多主观假设, 主要从系统的输入输出关系上进行探讨, 较好地解决了复杂系统分析的客观性问题.

1.1 自组织数据挖掘原理

自组织数据挖掘的核心技术是 GMDH(Group Method of Data Handling), 它从参考函数构成的初始模型(函数)集合出发, 按一定的法则产生新的中间候选模型, 再经过选定的标准筛选(选择), 重复这样一个中间模型产生和选择过程, 使中间候选模型的复杂度不断增加, 直至得到最优复杂度模型. GMDH 可建立自变量与因变量之间存在的高阶多项式关系, 获得一个对因变量具有解释能力的多项式模型.

GMDH 网络结构图如图 1 所示^[7], 图中:

① 收稿日期: 2005-03-23; 修订日期: 2006-07-30.

基金项目: 国家自然科学基金资助项目(70371004); 教育部博士点基金资助项目(20040006023).

作者简介: 邹昊飞(1977—), 男, 湖北赤壁人, 博士生. Email: zouhaofei@sohu.com

\hat{Y} = GMDH 的预测值(输出变量).

x_{il} = 第 l 个样本中的第 i 个输入变量, $i = 1, 2, \dots, n, n =$ 自变量的个数.

\hat{y}_{jkl} = 第 l 个样本在第 j 层中第 k 个神经元时的预测值, $k = 1, 2, \dots, m$.

r_{jk}^2 = 第 j 层中第 k 个神经元的阈值集合的均方根.

R_j = 第 j 层的阈值.

GMDH 网络建立输入变量和输出变量之间的一般函数关系. 自组织算法一般使用 Kolmogorov-Gabor 多项式参考函数:

$$y = a_0 + \sum_{i=1}^M a_i x_i + \sum_{i=1}^M \sum_{j=1}^M a_{ij} x_i x_j + \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^M a_{ijk} x_i x_j x_k + \dots$$

式中 $\mathbf{x}(x_1, x_2, \dots, x_M)$ 为输入变量矩阵, $\mathbf{a}(a_1, a_2, \dots, a_M)$ 为系数或权重矩阵, \mathbf{y} 为输出变量矩阵. GMDH 建模中的变量选择采用多层迭代算法进行, 借助计算机相对客观地选择变量之间的关系, 通过学习实现输入输出间的非线性映射, 用最小偏差准则选取最优模型.

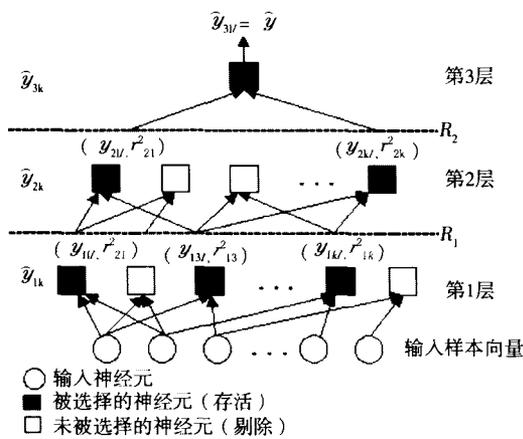


图 1 GMDH 网络结构图

Fig.1 Structure of GMDH network

1.2 自组织数据挖掘算法 GMDH 的基本步骤

GMDH 学习过程按如下步骤进行:

1) 数据准备. 建立预测模型, 将样本集分为学习集 A 、检测集 B 和校验集 C (checking set) ($N_w = N_a + N_b + N_c, w = a + b + c$). 并设定各层的阈值 R_j .

2) 选择一个外准则作为一个目标函数. 自组织算法允许众多选择准则, 为不同系统确定各自

的复杂性. 运用最小偏差准则^[5].

$$r_{jk}^2 = \frac{\sum_{l=N_a+1}^P (Y - \hat{y}_{jkl})^2}{Y^2}$$

其中, Y 因变量的输出值, $P = N_a + N_b$.

3) 产生初始的网络架构. 产生第 1 层 ($j = 1$) 的 m 个神经元, 建构成初始的网络.

4) 计算与检查 r_{jk}^2 . 可分为两个检查步骤:

① 当计算所得的 $r_{jk}^2 < R_j$, 则保留此神经元, 继续到步骤 ②; 反之到步骤 5).

② 找出第 j 层所保留重要神经元中最小的 r_{jk}^2 , 并与第 $j - 1$ 层最小的 $r_{j-1,k}^2$ 比较, 若 $r_{jk}^2 < r_{j-1,k}^2$, 到步骤 6); 反之到步骤 7).

5) 剔除不好的神经元. 根据步骤 4) 检查结果, 剔除不重要的神经元.

6) 产生下一层神经元, $j = j + 1$. 采用步骤 5) 保留下来的神经元, 产生下一层的神经元, 并回到步骤 4).

7) 完成 GMDH 学习. 当第 j 层最小的 r_{jk}^2 大于第 $j - 1$ 层最小的 $r_{j-1,k}^2$ 时, 则视为第 $j - 1$ 层第 k 各神经元找到最佳的参数, 完成第 j 层的学习.

2 遗传算法优化神经网络的 MSOA 算法

2.1 基于遗传算法的神经网络

遗传算法优化 BP 神经网络的基本思想: 利用遗传算法全局性搜索的特点, 为神经网络寻找最为合适的网络连接权和网络结构. 由于三层神经网络的输入层、输出层节点的个数由建模样本决定, 故而在优化 BP 网络结构时, 主要是优化它的隐节点的个数.

利用遗传算法求解如下的二次非线性优化问题, 将得到网络的结构和连接权, 代入 $\frac{1}{N - N_1} \sum_{k=1}^N \sum_{i=1}^n [y_k(t) - \hat{y}_i(t)]^2$ (记为式(1), 其中 N_1 为训练样本大小) 式计算误差 E_2 , 如果 E_2 小于设定的误差 ϵ_2 , 则可利用该模型进行实际应用. 本文采用两种不同的方法来优化三层 BP 神经网络, 其中基本的 GAANN 方法(称为方法 1) 的具体实现步骤如下:

1) 先采用三层 BP 网络来初步确定基本解空间, 确定网络连接权的取值范围. 采用 BP 算法对神经网络进行训练, 训练好的网络的连接权值中的最大值和最小值分别记为 u_{\max} 和 u_{\min} , 以该区间 $[u_{\min} - \delta_1, u_{\max} + \delta_2]$ (其中 δ_1, δ_2 为调节参数) 作为连接权的基本解空间.

2) 定义遗传算法的适应度函数(目标函数最大值)为

$$F(w, v, \theta, r) = \frac{1}{\sqrt{\sum_{k=1}^{N_1} \sum_{t=1}^n [y_k(t) - \hat{y}_k(t)]^2}} \quad (2)$$

其中, $y_k(t)$ 为实际值, $\hat{y}_k(t)$ 为网络实际输出值, w, v 分别表示输入层至中间层和中间层至输出层的连接权, θ, r 分别表示中间层和输出层各单元的输出阈值.

3) 对基本解空间进行编码, 其中编码生成的码串由控制码和权重系数码两部分组成. 编码按一定的顺序级联成一个长串, 每个串对应一组网络结构和连接权.

4) 初始群体由 L 个体构成, 每个个体由两部分组成, 第 1 部分是串长为 l_1 的 0-1 串; 第 2 部分是区间 $[u_{\min} - \delta_1, u_{\max} + \delta_2]$ 上的 l_2 个均匀分布随机数.

5) 计算群体中每个个体的适应度, 由控制码得到网络的隐节点个数, 由权重系数码得到网络的连接权, 输入训练样本, 按照式(2) 计算个体的适应度.

6) 保留群体中适应度最高的个体. 对群体中的其它个体, 采用轮盘赌选择法进行选择.

7) 对于控制码的交叉和变异采用基本遗传算法中的方法, 在变异运算时, 当某个神经元被变异运算删除时, 相应的有关权重系数编码被置为 0, 而当变异运算增加某个神经元时, 则随机初始化有关权重系数编码.

8) 生成新一代群体.

9) 反复进行 5 ~ 8, 每进行一次, 群体就进化一代, 连续进化到 K 代(总的进化代数).

10) 从第 k 代中选适应度最高的个体解码得到相应的网络连接权和隐节点个数, 输入检测样本检验模型的泛化能力.

2.2 基于遗传算法的 MSOA 算法

利用训练样本, 求得满足式(2) 的极大值, 获得最后一代适应度最高的个体, 但其并不一定是

网络的全局最优解, 网络的全局最优点并不一定是极值点, 在网络权值的取值区间为紧致子集的情况下全局最优点就可能是边界点. 因此, 在上述的算法中, 以约束条件下个体适应度最大值作为目标函数, 求得的网络权系数和网络结构并不一定是网络的全局最优解. 另外遗传算法虽然能以较大的概率进化最优解的区域, 但是无法精确地确定最优解的位置, 由此也不能盲目认定目前适应度最高个体就是全局性的最优解.

建立时间序列的神经网络模型的一个需要重要考虑的问题是选择多长的样本空间才能更好的捕捉到时间序列中的动态本质和模式. 实验结果就表明并不是样本量越大神经网络预测模型的预测效果越好^[8]. 决定一个合适的样本区间来构建时间序列的神经网络预测模型是一项重要的任务. 从理论上讲, 当时间序列的潜在数据生成过程是平稳时, 样本量越大则能越能使模型剔除其中固有噪声的影响; 但若这个过程是不平稳的, 或者它的结构或参数随时间改变, 那么大的样本数据反倒可能影响模型的预测效果^[9]. 但是我们认为过去的样本数据对模型的构建是有益的, 不能轻易丢弃.

考虑到上述算法是用 N_1 个历史样本建模的基础上, 外推预测后面 $N - N_1$ 个时间周期的结果, 而没有考虑与其最近的样本对后面预测量的影响比较重要这个事实, 对离观测值越近的样本应赋予更重要的权重. 同时考虑到遗传算法搜索到最优解附近时, 却无法精确地确定最优解的位置, 为此在把上述基本方法进行修正, 得到改进的 GAANN 算法——MSOA 算法(称为方法 2), 进一步对网络进行优化, 以更加逼近全局最优解. 首先总样本 N 分成三个部分, 训练样本 ϕ_1 、训练样本 ϕ_2 和检测样本 ϕ_3

$$\phi_1 = \{(x_k, y_k) \mid x \in R^m, y \in R^n, \\ k = 1, 2, \dots, N_1; N_1 < N_2\}$$

$$\phi_2 = \{(x_k, y_k) \mid x \in R^m, y \in R^n, \\ k = N_1 + 1, N_1 + 2, \dots, N_2; N_2 < N\}$$

$$\phi_3 = \{(x_k, y_k) \mid x \in R^m, y \in R^n, \\ k = N_2 + 1, N_2 + 2, \dots, N\}$$

针对训练样本 ϕ_1 , 仍然采用 2.1 节中方法 1 (1 ~ 9) 的步骤, 直到进化 K 代后结束训练, 再把

最后一代的群体 L 个个体全部解码, 得到相应的 L 组神经网络的权系数和网络结构, 输入训练样本 ϕ_2 , 运用 GA 算法计算下式:

$$\begin{cases} \min E_2(w, v, \theta, r) = \\ \frac{1}{N_1 - N_2} \sum_{k=N_1}^{N_2} \sum_{t=1}^n [y_k(t) - \hat{y}_k(t)]^2 \\ \text{s.t. } w \in R^{m \times p}, v \in R^{p \times n}, \\ \theta \in R^p, r \in R^n \end{cases} \quad (3)$$

其中, E_2 为网络误差, $y_k(t)$, $\hat{y}_k(t)$, w, v, θ, r 含义同式(2). 为此可以得到新的一组网络的权系数和隐节点, 再用样本 ϕ_3 检验所得网络的泛化能力.

MSOA 算法(方法2)可以归纳为: 先通过训练样本 ϕ_1 , 进化得到一组网络权系数和网络结构, 再进一步用训练样本 ϕ_2 从中挑选优化, 最后检验网络的泛化能力. 在神经网络模型的建立过程中, 人们往往随机和武断地初始化神经网络的权重值矩阵, 因此也可将 MSOA 算法的第1步看作神经网络权重值矩阵的有效初始化过程.

3 基于自组织数据挖掘的神经网络组合预测模型

当研究复杂系统时, 为避免对因变量有重要影响的因素的漏选, 常用的方式是用定性分析方法先选出数量较多的对因变量较有影响的那些自变量, 再着手建立系统模型. 当这些因素很多时, 把它们都作为神经网络的输入, 显然会增加网络的复杂度, 降低网络性能, 大大增加计算运行的时间, 影响计算的精度; 另外, 这种定性选择的方法存在很大的主观性和随意性, 受人的主观因素和领域知识的影响较大, 很大程度上有赖于决策者的专家知识和经验. 自组织数据挖掘算法为解决这一难题提供了较好的方法. GMDH 是一种从数据中进行知识抽取的方法, 应用不完全归纳分类法, 可以对大量的输入变量进行处理, 对选取变量可尽量全面、广泛, 通过数据进行检验, 让数枯有“说话的权利”, 使因素的选择更为客观. 特别地, 当对所研究的系统知之甚少, 对其内部复杂的形成机制没有明确定义或具有很少的先验知识时, 这种方法更加地适用. 这样, 先用自组织算法

筛选出对因变量(网络输出)最有影响作用的变量(自变量)作为神经网络的输入节点, 再用改进的神经网络算法进行学习. 这样作的思想在于, 虽然自组织算法在筛选出对因变量最有影响的因素(自变量)方面有独到的优点, 然而由于自组织算法主要使用线性逼近得到结果, 其预测精度往往不如神经网络算法. 因而, 将两种方法有机地结合起来, 以期增强神经网络对复杂系统建模的能力.

自组织算法(GMDH)与改进型 GAANN 神经网络模型相结合的算法流程如图2所示.

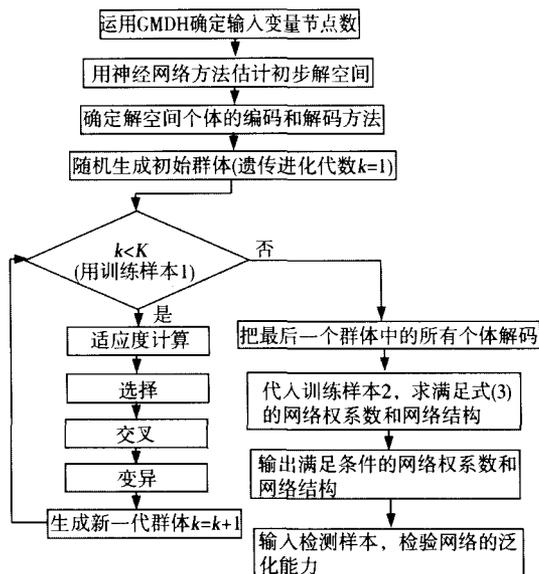


图2 基于 GMDH 的 GAANN 改进算法的建模流程图

Fig.2 Flow chart of the improved arithmetic of GAANN based GMDH

4 基于自组织数据挖掘的神经网络组合预测模型的应用

粮食安全事关国计民生和社会政治稳定的大事, 一直是各国经济发展的重要问题. 我国是一个粮食消费大国, 粮食市场价格对消费产生重要影响, 粮食市场价格频繁波动将对我国的粮食安全产生重要影响. 对国家粮食市场进行监控和预测是国家粮食安全预警系统中的一个重要课题, 本文选择一个粮食市场, 运用上述建模模型, 进行粮食的月度价格预测建模.

4.1 月度粮食价格自组织建模

本文选择的数据为中国郑州粮食交易市场的

三级白小麦月度价格序列,从 1996 年 1 月到 2005 年 7 月共 115 个观测值,具体数据可从中华粮网(www.cngrain.com)获得.以 1996 年 1 月—2004 年 7 月的当期月度小麦价格(X_t)为输出神经元,以滞后期为 1 到 6 期的月度小麦价格(X_{t-1} 、 X_{t-2} 、 X_{t-3} 、 X_{t-4} 、 X_{t-5} 、 X_{t-6})等 6 个因子为输入神经元.采用自组织算法进行模拟,得到模型方程为

$$X_t = 49.965 + 1.294X_{t-1} - 0.335X_{t-2}$$

$R^2 = 0.970$, 预测误差平方和:0.031 8, 平均绝对百分比误差:2.02%, 逼近误差方差:0.030. 从模型可以看出,GMDH 网络从 6 个因子中选择了 2 个(X_{t-1} 、 X_{t-2})对当期小麦价格产生重要影响的因素.

4.2 基于自组织数据挖掘的神经网络进行粮食价格预测

在建立基于遗传算法 BP 神经网络模型时,首先采用三层 BP 神经网络估计网络连接权的基本解空间,利用自组织算法筛选的变量作为神经网络

输入层节点(2 个),输出层节点个数 1 个.关于隐节点的选取问题,有学者^[10]建议隐含节点数目应介于输出节点和隐含节点之间,另有规则^[11]认为隐含结点数大约等于输入节点的 75%,作者采用试探方法确定网络隐节点 4 较为合适.输入训练样本和检测样本,按照 3.1 节的算法求得 $u_{\min} = -2.54$, $u_{\max} = 5.38$, 则连接权的解空间初步设为 $[-3, 6]$, 隐节点个数搜索范围为 2 ~ 6, 遗传算法进化过程中初始种群 $L = 100$, 总的进化代数 $K = 1\ 000$, 交叉概率 $p_c = 0.85$, 变异概率 $p_m = 0.008$. 在用方法 1 和方法 2 建模时训练样本都为 1996.01—2004.07 年 103 个样本,只是在方法 2 中把训练样本按约 70% : 30% 比例分成两部分,训练样本 ϕ_1 为 70 个(1996.01—2001.09)、训练样本 ϕ_2 为 33 个(2001.10—2004.07). 两个方法的检测样本 ϕ_3 均相同,为 2004.07—2005.07 年 12 个独立样本.建立的预报模型对独立样本的预报统计结果见表 1,图 3 显示了各点预测值的比较情况.

表 1 三种模型预测结果比较*

Table 1 Compare of forecasting results and error of three models

模型	样本内(In-sample)数据集			样本外(Out-sample)数据集		
	MAE	MSE	MAPE (%)	MAE	MSE	MAPE (%)
GMDH	26.227	1 737.636	2.032	16.559	358.106	0.931
方法 1(GAANN)	22.735	1 298.635	1.903	12.655	276.753	0.692
方法 2(MSOA)	17.528	1 023.564	1.245	10.475	205.125	0.509

* 样本内数据集:1996.01—2004.07;样本外数据集:2004.08—2005.07

从表 1 和图 3 结果比较可以看出,GMDH-GAANN 算法在预测精度上较 GMDH 方法提高了许多,而 GMDH-GAANN 改进算法(MSOA)又较 GMDH-GANN 在预测精度又有了进一步的提高,说明改进型的算法是切实可行并且能够更进一步的优化神经网络结构.

5 MSOA 算法收敛性的进一步探讨

本节利用第 4 节提供的样本数据采用试验方法,进一步探讨 MSOA 算法在提高神经网络的收敛能力和收敛速度上的有效性.训练神经网络的收敛标准为平均绝对误差小于 0.001 或对普通 GAANN 算法(方法 1)训练次数最大为 1 000 次,对于 MSOA 算法(方法 2)训练次数最大为 500 次. ANN 模型以 MataLab 6.5 构建,神经网络的输入层

节点限定为 6 个(即 1~6 个滞后期),初始网络的隐含层节点数限定为 1~8 个,输出层节点为 1 个,对共计 $6 \times 8 = 48$ 种初始神经网络结构进行试验.每次试验进行 10 次取平均值作为最后结果进行分析计算.训练样本的划分同 4.1 节.

从试验结果可以看出,以神经网络结构为 $3 \times 5 \times 1$ 为例,使用普通 GAANN 算法(方法 1)大部分神经网络都遇到收敛速度慢的问题,在设定的 1 000 训练次数内不能达到预定的训练目标(0.001)(如图 4-a);而使用改进算法——MSOA 算法(方法 2),在第 1 步训练过程中同样遇到收敛速度慢的问题(如图 4-b),然而在第 2 步训练过程中,大部分不同结构的神经网络都能够迅速收敛到目标值(如图 4-c).从试验结果可以清楚观察到 MSOA 算法良好的收敛特性.图 5 显示了结构为 $3 \times 5 \times 1$ 的神经网络,相同条件下采用两种算法进行 10 次训练所用训练时间的比较,使用 MSOA 算

法所需的总的神经网络训练时间较之普通算法显著减少;MSOA 方法的平均训练时间仅为 4.25 s 而普通算法为 7.83 s. 结果表明使用 MSOA 方法能较大提高神经网络的全局收敛能力和收敛速度,实验证明本算法不需要付出更多的计算代价.

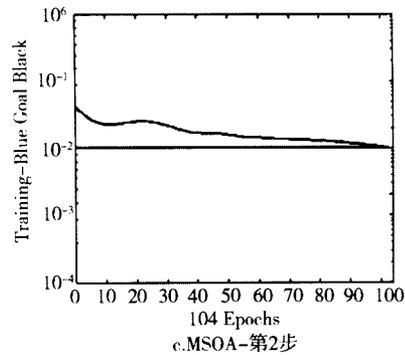
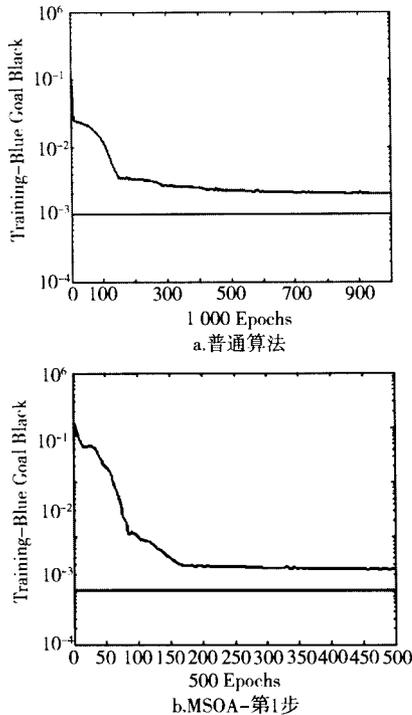


图 4 使用普通算法和 MSOA 的训练过程及 MSE 值 (以神经网络结构为 $3 \times 5 \times 1$ 为例)

Fig.4 MSE in the training process using two different algorithms (the neural network structure is $3 \times 5 \times 1$, for example)

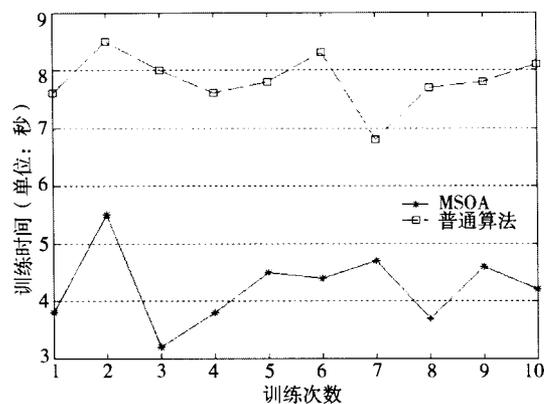


图 5 两种训练方法每轮的训练时间比较 (以神经网络结构为 $3 \times 5 \times 1$ 为例)

表 2 利用两种算法对神经网络训练 10 次试验结果比较

Table 2 The comparison of both algorithms after 10 training runs

试验次数		1	2	3	4	5	6	7	8	9	10
MSOA	训练时间/s	3.8	5.5	3.2	3.8	4.5	4.4	4.7	3.7	4.6	4.2
	性能值/ 10^{-3}	0.43	0.97	0.48	0.45	0.57	0.98	1	1	0.99	1
	收敛性	0.97	Y	Y	Y	Y	Y	Y	Y	Y	Y
普通算法	训练时间/s	7.6	8.5	8	7.6	7.8	8.3	6.8	7.7	7.8	8.1
	性能值/ 10^{-3}	1.75	2.19	2.1	1.75	2.11	2.16	1	2.12	1.82	1.89
	收敛性	N	N	N	Y	N	N	N	N	N	N

* Y: 神经网络训练后收敛到目标值(≤ 0.001); N: 神经网络训练后没有收敛到目标值(> 0.001)

6 结束语

本文利用自组织数据挖掘(GMDH)与神经网络模型结合,为解决复杂系统建立神经网络模型过程中难于确定网路输入节点问题提供了较好的方法.同时考虑到遗传算法在进化过程中能以较大概率搜

索到全局最优解存在的区域,在遗传算法搜索到最优解附近时,却无法精确地确定最优解的位置,为此本文进一步利用进化后的结果,采用两阶段优化方法(MSOA),挑选最佳网络权系数和网络结构,建立基于遗传算法的神经网络组合预测模型的改进算法,充分发挥了遗传算法和神经网络的长处,综合了神经网络泛化映射能力和遗传算法全局收敛能力.

将该神经网络组合预测模型应用于粮食价格预测,取得良好的效果.同时文章进一步探讨了MSOA算法用于训练神经网络的收敛性问题,试验研究表明

该算法能够提高神经网络的全局收敛能力和收敛速度,进一步提高预测能力,同时改进算法在计算成本上并没有提高,甚至有一定程度的减少.

参 考 文 献:

- [1]周 敏,王新宇.基于模糊优选和神经网络的企业财务危机预警[J].管理科学学报,2002,5(3):86—90.
Zhou Min, Wang Xin-yu. Pre-warning systems of enterprise financial crisis based on fuzzy selection and artificial neural networks [J]. Journal of Management Sciences in China, 2002, 5(3): 86—90. (in Chinese)
- [2]于立勇.商业银行信用风险评估预测模型研究[J].管理科学学报,2003,6(5):46—52.
Yu Li-yong. Study on credit risk assessment and forecasting model in commercial bank[J]. Journal of Management Sciences in China, 2003, 6(5): 46—52. (in Chinese)
- [3]樊 智,张世英.非线性协整建模研究及沪深股市实证分析[J].管理科学学报,2005,8(1):73—77.
Fan Zhi, Zhang Shi-ying. Study on modeling of nonlinear cointegration and empirical analysis of China's stock markets[J]. Journal of Management Sciences in China, 2005, 8(1): 73—77. (in Chinese)
- [4]Bi T S, Ni Y X, Shen C M, et al. A Novel ANN Fault Diagnosis Systems Using Dual GA Loops in ANN Training[C]. 2000 IEEE Power Engineering Society Summer Meeting, 2000, 1: 425—430.
- [5]Gupta J N D, Sexton R S. Comparing backpropagation with a genetic algorithm for neural network training[J]. Omega, The International Journal of Management Science, 1999, 27(6): 679—684.
- [6]Srinivasan D. Evolving artificial neural networks for short term load forecasting[J]. Neurocomputing, 1998, 23: 265—276.
- [7]Mueller J A, Lemke F. Self-Organizing Data Mining[M]. Berlin: Dresden, 1999.
- [8]Walczak S. An empirical analysis of data requirements for financial forecasting with neural networks[J]. Journal of Management Information Systems, 2001, 17: 203—222.
- [9]Morantz B H, Whalen T, Zhang G P. A weighted window approach to neural network time series forecasting[A]. In: Zhang, G. P., (Ed.), Neural Networks in Business Forecasting, 2004. 251—265.
- [10]Subramanian V, Ming S H, Hu M Y. Multilayer feedforward networks are universal approximators[J]. Computers and Operations Research, 1993, 20(7): 769—782.
- [11]Salchenberger L M, Clnar E M, Lash N A. Neural Networks: A new tool for predicting thrift failures[J]. Decision Sciences, 1992, 23(4): 899—916.
- [12]Farlow J S. Self-Organizing Methods in Modeling[M]. New York: Marcel Dekker, 1984.
- [13]Sexton R S, Dorsey R E, Johnson J D. Toward global optimization for artificial neural networks: A comparison of the genetic algorithm and back propagation[J]. Decision Support Systems, 1998, 22(2): 171—185.
- [14]Honik K. Approximation capabilities of multilayer feedforward networks neural[J]. Neural Network, 1991, 4: 551—557.
- [15]郑立群,吴育华,周伯康等.人工神经网络方法在投资风险评价中的应用[J].管理科学学报,1999,2(4):93—95.
Zheng Li-qun, Wu Yu-hua, Zhou Bo-kang, et al. The application of artificial neural network(ANN) on investment risk evaluation [J]. Journal of Management Sciences in China, 1999, 2(4): 93—95. (in Chinese)
- [16]荣莉莉,王众托.基于知识和模糊神经网络的学习型评价系统[J].管理科学学报,2003,6(3):1—7.
Rong Li-li, Wang Zhong-tuo. Learning evaluation system based on knowledge and fuzzy neural networks[J]. Journal of Management Sciences in China, 2003, 6(3): 1—7. (in Chinese)
- [17]何 跃,鲍爱根,贺昌政.自组织建模方法和GDP增长模型研究[J].中国管理科学,2004,12(2):134—142.
He Yue, Bao Ai-gen, He Chang-zheng. Self-organizing modeling methods and a model study on the growth of GDP[J]. Chinese Journal of Management Science, 2004, 12(2): 134—142. (in Chinese)
- [18]袁曾仁.人工神经元网络及其应用[M].北京:清华大学出版社,1999.
Yuan Zengren. Artificial Neural Networks and Applications[M]. Beijing: Tsinghua University Press, 1999. (in Chinese)
- [19]贺昌政,刘建平.自组织数据挖掘理论与经济系统的复杂性研究[J].系统工程理论与实践,2001,12(12):1—5.
He Chang-zheng, Liu Jian-ping. Study of self-organizing data mining theory and the complexity of economics systems[J]. Systems

Engineering-Theory & Practice, 2001, 12(12): 1--5(in Chinese)

[20] 李晓峰, 徐玖平, 王荫清, 等. BP 神经网络自适应学习算法的建立与应用[J]. 系统工程理论与实践, 2004, 5: 1—8.

Li Xiao-feng, Xu Jiu-ping, Wang Yin-qing, *et al.* The establishment of self-adapting algorithm of BP neural network and its application[J]. Systems Engineering-Theory & Practice, 2004, 5: 1—8. (in Chinese)

[21] 董 聪. 前向网络全局最优化问题研究[J]. 中国科学基金, 1997, 1: 23—29.

Dong Cong. Advances and prospect of the global optimization for feedforward networks[J]. Bulletin of National Natural Science Foundation of China, 1997, 1: 23—29. (in Chinese)

Neural network forecasting model using multi-stage optimization approach based on GMDH and genetic algorithm

ZOU Hao-fei¹, XIA Guo-ping¹, YANG Fang-ting²

1. School of Economics & Management, Beihang University, Beijing 100083, China;

2. Beijing Simulation Center, Beijing 100854, China

Abstract: This paper introduces a multi-stage optimization approach (MSOA) used in genetic algorithm (GA) for training neural networks to forecast the Chinese food grain price. We divide the training sample of neural networks into two parts considering the truth that the recent observations should be more important than the older ones. Firstly, we use the first training sample to train the neural network and achieve the network structure; Secondly, we continue to use the second training sample to further optimize the structure of neural network based on the previous step. Aiming at the characteristics of neural network structure, a model using a hybrid GMDH and artificial neural network is established. It can make the selection of input-lay units easy and improve the ability of rate of studying and the adaptability of neural network. Empirical results show that the neural networks based on MSOA can improve greatly the global convergence ability and convergence speed of most networks. Furthermore the result indicates that the combined model can be an effective way to improve forecasting accuracy.

Key words: GMDH; multi-stage optimization approach; genetic algorithm; artificial neural network