

8-20-2020

## Polynomial Neural Network with Direct Solutions and Its Interpretation of Inputs

Shen Wei

*Department of Economics and Management of North China Electric Power University, Beijing 102206, China;*

Qiushi Li

*Department of Economics and Management of North China Electric Power University, Beijing 102206, China;*

Yukun Song

*Department of Economics and Management of North China Electric Power University, Beijing 102206, China;*

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research, Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

---

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

---

# Polynomial Neural Network with Direct Solutions and Its Interpretation of Inputs

## Abstract

**Abstract:** The generalized multivariate polynomial neural network and single-hidden-layer generalized multivariate polynomial neural network were designed and systematically studied, and the existence of the optimal weight vector was proved which could make the network the best approximation polynomial for an unknown function; The concepts of the natural upper and lower nodes in the hidden layer were created, and an indicator “value of importance” was creatively designed and the partial derivative analysis was introduced to solve the problem that the neural network was not able to interpret the relationship between variables. The weight vector directly was solved proving it optimal. In addition, Matlab-based graphical user interface was designed. Through this program, users could update stock data via the stock software, and predicted the stock index on specified date via different models.

## Keywords

generalized multivariate polynomial neural network, weights-direct-determination, importance value, stock index forecasting

## Recommended Citation

Shen Wei, Li Qiushi, Song Yukun. Polynomial Neural Network with Direct Solutions and Its Interpretation of Inputs[J]. Journal of System Simulation, 2015, 27(3): 559-570.

# 权值直解的多项式神经网络及其解释能力设计

沈巍, 李秋实, 宋玉坤

(华北电力大学经济与管理学院, 北京 102206)

**摘要:** 设计并系统研究了广义多元多项式神经网络, 单隐层广义多元多项式神经网络, 证明存在最优权值向量使该网络成为未知函数的最佳逼近多项式; 创造性地建立了隐层节点的自然次序上限和下限, 以及重要值等概念, 并引入了偏导数分析, 解决了神经网络不具备解释能力的弊病。设计了权值直接解法, 证明该解法所得的权值向量是迭代法逼近的最优权值向量。设计了基于 Matlab 的图形用户界面。通过该程序, 用户可通过炒股软件更新股票数据, 读取特定股票、特定容量的数据, 进行不同模型下指定日期的预测。

**关键词:** 广义多项式神经网络; 权值直接解法; 重要值; 股指预测

中图分类号: TP391.9

文献标识码: A

文章编号: 1004-731X (2015) 03-0559-12

## Polynomial Neural Network with Direct Solutions and Its Interpretation of Inputs

Shen Wei, Li Qiushi, Song Yukun

(Department of Economics and Management of North China Electric Power University, Beijing 102206, China)

**Abstract:** The generalized multivariate polynomial neural network and single-hidden-layer generalized multivariate polynomial neural network were designed and systematically studied, and the existence of the optimal weight vector was proved which could make the network the best approximation polynomial for an unknown function; The concepts of the natural upper and lower nodes in the hidden layer were created, and an indicator "value of importance" was creatively designed and the partial derivative analysis was introduced to solve the problem that the neural network was not able to interpret the relationship between variables. The weight vector directly was solved proving it optimal. In addition, Matlab-based graphical user interface was designed. Through this program, users could update stock data via the stock software, and predicted the stock index on specified date via different models.

**Keywords:** generalized multivariate polynomial neural network; weights-direct-determination; importance value; stock index forecasting

## 引言

1943 年, 神经生物学家 W·McCulloch 和数理逻辑学家 W·Pitts 进行合作, 用逻辑的数学工具研



收稿日期: 2014-02-03 修回日期: 2014-05-09;  
基金项目: 北京市自然科学基金(9132011);  
作者简介: 沈巍(1965-), 女, 辽宁朝阳, 博士, 副教授, 研究方向为预测理论与方法、人口预测、股指预测、智能预测; 李秋实(1991-), 男, 吉林, 研究方向为股指预测、智能预测; 宋玉坤(1989-), 男, 辽宁丹东, 硕士生, 研究方向为预测理论与方法、人口预测、智能预测。

究客观事件在形式神经网络中的表述, 开创了对神经网络的理论研究, 首次提出了神经元的数学模型, 即神经网络模型<sup>[1-2]</sup>。人工神经网络是一种模拟大脑神经突触联接的结构进行信息处理的数学模型, 这种模型的计算能力具有 2 个明显特点<sup>[3-4]</sup>: (1)大规模并行分布式结构; (2)神经网络学习能力以及由此而来的泛化能力。作为一种在外界信息的基础上改变内部结构而建立的一种自适应系统, 神经网络可以用来对输入和输出间复杂的关系进行建模, 或用来探索数据的模式, 因而被众多学者关

注并展开了广泛而深入的研究<sup>[5-6]</sup>。

然而, 一些学者对神经网络本身存在疑问态度, 因为作为一种自适应系统, 它自身是一个黑箱, 我们虽然可以得到较为准确的预测或识别结果, 但其中的关系不能得知<sup>[7-9]</sup>。如果能够将神经网络赋予解释能力, 那么神经网络的应用也就有了意义。不少学者对输入变量的重要性提出了自己的办法, 为改善神经网络黑箱特质做出了贡献。Shinying Huang, Ruahuan Tsaih, Wanying Lin(2012)利用非监督神经网络研究了欺诈性财务报告<sup>[10]</sup>。显然, 作者考虑到了输入变量可能不能够解释输出变量的变化的问题, 所以一开始用计量模型剔除了不显著的变量。但是, 如果作者打算运用非线性的神经网络, 用线性模型剔除不显著总是有问题的。Kelly E. Fish, Richard S. Segall(2004)移植了统计学的方法, 认为计量模型的 t 检验结果就是相应的偏系数是否为 0, 也就是各个偏导数是否为 0<sup>[11]</sup>。利用这种思想, 我们可以求出神经网络系统模型下, 各输入变量对输出变量的偏导数, 通过判断是否为 0 来说明输入变量是否重要, 并且如何影响结果。二人的研究为解释神经网络模型提供了一种新的方法, 并使这一模型免于不能够解释结果的弊病。然而, 虽然二人工作具有开创性, 并且运用了数学方法进行了推导, 但是最后的偏导数是否显著异于 0 是靠观察得出的, 因此缺乏一个标准。故下一步的研究方向可能就是如何判断这种偏导数显著异于 0。Shee Q. Wong, Nik R. Hassan, Ehsan Feroz(2007)通过研究资本增值从另一个角度提出了判断解释变量是否重要一种方法<sup>[12]</sup>。作者同时运用了传统统计法、神经网络法对资本增值进行了拟合, 结果表明, 神经网络在预测时获得了更小的平均绝对误差。在判断解释变量是否更重要时, 他们分别剔除了各个解释变量, 发现在删除股利收益率后平均绝对误差迅速增加, 最终得出股利收益率是影响资本增值的最重要的因素。实际上这个方法就是 Kelly E. Fish 等人的方法, 认为梯度下降越快的变量重要性越高。

此外, 神经网络模型还存在一些问题。例如,

神经网络训练速度慢, 有些模型如 BP 神经网络容易陷入最小值等<sup>[13]</sup>。大量学者对算法方面进行了研究, 但是绝大部分是软算法, 如遗传算法, 鱼群算法等<sup>[14]</sup>, 并没有从模型设计上提出改进的建议。近年来, 部分学者提出了以多项式为激励函数的神经网络, 这种模型能够从根本上解决训练速度慢的问题, 同时也可以以任意精度逼近未知函数。Rozaida Ghazali, Abir Jaafar Hussain, et al. (2009)分别将 Pi-Sigma 和脊多项式设定为激励函数, 设计了相应的神经网络模型<sup>[15]</sup>。他们发现脊多项式神经网络能够在绝大部分情况下捕获信号的混沌运动, 同时, 较其他激励函数而言, 这种多项式激励函数又能够快速收敛。随后, 作者又运用了随机游走模型对相同的金融信号进行了预测, 结果误差很大。M. Mehrabi, M. Sharifpur, J. P. Meyer(2012)分别运用了基于 FCM 的神经元模糊系统和用遗传算法的多项式神经网络, 对三氧化二铝的热传导比率进行了拟合和预测<sup>[16]</sup>。证明基于 FCM 的神经元模糊系统比多项式神经网络好一些, 但是不可否认, 多项式神经网络的误差也很小。Ivan Maric(2013)用基于 Levenberg-Marquardt 算法的自组织多项式神经网络对几种确定型非线性函数进行了拟合<sup>[17]</sup>。其结果显示, 自组织多项式神经网络能够获得令人满意的逼近精度。同时, 用 LM 算法进行优化的结果要优于粒子群算法, 但这两种算法都要优于迭代法。

但是, 大部分学者在求解神经网络的权值向量的时候都运用了迭代法的思想(遗传算法、粒子群法也是迭代法的一种), 如 E. Gomez-Ramirez, K. Najim, E. Ikonen(2007)<sup>[18]</sup>。而从某种意义上讲, 采用了迭代法进行求解就不可避免地出现局部最小点的问题, 表现为预测结果不准确, 预测结果不稳定等。Yunong Zhang, W. Li (2009)在求解 Genenbauer orthogonal basis neural network 时采用了伪逆矩阵的思想, 可以将神经网络的权值向量直接解出<sup>[19]</sup>。这时求解出的权值向量是在最小二乘意义下的最优解。张雨浓, 杨逸文, 李巍(2010)在文中设计了多种激励函数的多项式神经网络, 并指

出, 采用伪逆矩阵求解权值向量, 较迭代法而言, 其计算速度更为迅速, 且精度更高<sup>[20]</sup>。

基于以上学者的研究, 本文采用广义多项式作为激励函数的单隐层神经网络, 首先设计出权值向量的直接解法; 在此基础上, 设计出能够赋予模型解释能力的重要值等指标, 并通过该指标能够较为合理地筛选变量, 解释输入变量与输出变量之间的关系。最终, 我们将运用基于该模型的自己设计的软件对中国上证综指的 72 个交易日收盘价进行预测, 并与标准 BP 神经网络和基于遗传算法的 BP 神经网络预测结果进行对比。

## 1 广义多项式神经网络

### 1.1 广义多项式神经网络概述

一般幂函数的实线性组合称为广义多项式。广义多项式神经网络是指, 输入层到隐层采用线性恒等激励函数、隐层的整体输出采用 P 次广义多项式的单隐层前向神经网络。通过设定广义多项式的最高次数, 这种以多项式为激励函数的神经网络可以以任意精度逼近未知函数<sup>[21]</sup>。

### 1.2 广义多项式神经网络性质

为了方便描述广义多项式神经网络, 这里我们引入隐层神经元的自然次序的概念。隐层神经元的自然次序是指当我们将对广义多元多项式采取降序字典法排列的节点次序。计算自然次序的目的在于, 能够通过给出任意一组输入变量的次数, 判断出其排在第几位; 同时, 能够给定任意一个次数判断其中有多少个节点, 即求  $M_i$ 。

设输入变量  $x_1, x_2, \dots, x_k$ , 其次数分别为  $j_1, j_2, \dots, j_k$ 。

定理 1.1 设隐层多项式的最大次数为 P, 则在降序字典法的排序下, 给定总次数  $n$  ( $1 \leq n \leq p$ ), 其自然次序下限为  $C_{(n+k-1)}^k + 1$ 。

推论 1.1 在降序字典法的排序下, 给定总次数  $n$  ( $1 \leq n \leq p$ ), 其自然次序上限为  $C_{n+k}^k$ 。

推论 1.2 给定总次数  $0 \leq n \leq P$ , 总次数为  $n$

的节点共有  $C_{n+k-1}^{k-1}$  个。

推论 1.3 给定最高总次数 P, 网络一共有  $C_{k+P}^k$  个节点。

定理 1.2 设

$$A_n^m = \begin{cases} C_n^m & 0 \leq m \leq n \\ 0 & m > n \end{cases} \quad (1)$$

则给定次数  $j_1, j_2, \dots, j_k$ , 可以确定其自然次序的序号

$$O_{\{j_i\}_{i=1}^k}^P = \sum_{q=1}^k A_{\sum_{i=q}^k j_i + k - q}^{k-q+1} + 1 \quad (2)$$

### 1.3 广义多项式神经网络的一般形式

根据前 2 节, 我们可以给出广义多项式神经网络的一般形式。设  $w_{O_{\{j_i\}_{i=1}^k}^P}$  为第  $O_{\{j_i\}_{i=1}^k}^P$  个节点的权值, 则网络的输出为

$$\sum_{n=0}^P \left[ \sum_{j_1=0}^n \sum_{j_2=0}^{n-j_1} \dots \sum_{j_k=0}^{n-\sum_{i=1}^{k-1} j_i} \left( w_{O_{\{j_i\}_{i=1}^k}^P} \prod_{i=1}^k x_i^{j_i} \right) \right] \quad (3)$$

其中  $O_{\{j_i\}_{i=1}^k}^P = \sum_{q=1}^k A_{\sum_{i=q}^k j_i + k - q}^{k-q+1} + 1$ , 网络共有  $C_{k+P}^k$  个节点。

## 2 权值解法

### 2.1 基于伪逆矩阵的直接解法

设如下线性方程组:

$$Ax = b \quad (4)$$

其中:  $A = (a_{ij})_{m \times n}$ ,  $x = (x_1, x_2, \dots, x_n)^T$ ,  $b = (b_1, b_2, \dots, b_m)^T$ 。

如果  $rank(A) = rank([A, b])$ , 则  $Ax = b$  有唯一解。然而, 在现实生活中, 我们很难做到上述的条件, 因此一般而言, (4) 不存在常规意义上的解。这时, 我们可以通过在最小二乘意义下利用伪逆矩阵进行求解。

定义 2.1 设实矩阵  $A \in R^{m \times n}$ , 若存在  $n \times m$  阶实矩阵 X 同时满足以下 4 个条件:

$$\begin{aligned} AXA &= A; & XAX &= X; \\ (AX)^T &= AX; & (AX)^T &= XA \end{aligned}$$

则称  $X$  为  $A$  的伪逆矩阵(或称广义逆矩阵、加号逆), 记为  $X = A^+ = (A^T A)^{-1} A^T$ 。

定理 2.1 矛盾线性方程组  $Ax = b$  的最小二乘解为  $x = A^+b$ ; 相容方程组  $Ax = b$  的极小范数解为  $x = A^+b$ 。

因此我们可以得到直接计算公式:

$$w^* = Q^+ \cdot \gamma \quad (5)$$

相比迭代法而言, 无需确定最优学习速率, 因此不存在计算时间过长、依赖经验选取学习速率等问题, 同时局部极小点问题也不复存在。

### 3 输入变量的重要性及筛选

神经网络的解释能力(即输入变量如何影响输出变量)一直是科研人员较为回避的话题, 因为目前为止, 尚未有较好的方法能够赋予网络以解释能力, 而一旦无法证明解释能力, 又从本质上否认神经网络的应用能力。现在研究有几种方法来筛选解释变量: 第 1 种是采取计量模型进行估计, 然后筛选不显著的变量, 之后再继续进行神经网络训练<sup>[10]</sup>。然而, 虽然剔除了线性关系不显著的变量, 但是不能保证这些变量与输出变量之间没有非线性的关系, 因此这种方法本质上是错误的; 第 2 种是通过神经网络进行改造, 从而赋予其一定的解释能力, 如将隐层修改为模糊化层和去模糊化层<sup>[22]</sup>, 由于模糊化层的规则由专家给出, 因此具备一定的解释能力。但是, 对于一个非线性的复杂问题(如股指的预测), 当我们选取了众多变量的时候, 它们在一起相互影响, 我们很难用专家经验进行分类, 因此这种方法有一定的局限性。第 3 种是逐步剔除输入变量, 然后观测平均绝对误差变化的幅度, 幅度越大, 说明该变量较为显著<sup>[5]</sup>。这种方法是正确的, 但是逐步删除变量十分笨拙麻烦。第 4 种方法在第 3 种方法上有所改进, 它计算各个输入变量的偏导数, 并通过观察偏导数的大小来判断解释变量的解释能力, 然后剔除偏导数一直接近 0 的解释变量。这种方法是最为科学的方法, 不过, “偏导数一直接近 0”又是一个不确定的经验性的判断标准。本

文创新性地设计出带权的重要值指标, 由偏导数入手, 能够完全解决网络的解释能力问题, 并且能为筛选变量提供依据, 解决了神经网络的本质问题。

定义 3.1 设单输出神经网络的输出为  $f(x_1, x_2, \dots, x_k)$ , 训练样本为  $d$  个, 定义  $x_i$  的重要值  $t_i$  为:

$$t_i = \sum_{s=1}^d \rho(s) \left| \frac{\partial f(x_1, x_2, \dots, x_k)}{\partial x_i} \right|_{x=x^{(s)}} \quad (6)$$

其中:  $\rho(s)$  为权函数, 且  $0 \leq \rho(s) \leq 1$ 。

#### 3.1 多项式神经网络重要值

多项式神经网络的优势在于, 不必重新计算偏导数矩阵, 可以直接在之前的结果上进行简单的运算, 便可求出各变量的重要值。

定理 3.1 广义多项式神经网络中,  $x_i$  的重要值  $t_i$  为

$$t_i = \sum_{s=1}^d \rho(s) \left| \sum_{l=1}^N \frac{j_l w_l g_l^{(s)}}{x_i} \right| \quad (7)$$

其中:  $l = O_{\{j_l\}_{i=1}^p}$ ,  $g_l^{(s)} = \prod_{i=1}^k (x_i^{(s)})^{j_i}$ ,  $N = C_{k+p}^k$ ,

$$O_{\{j_l\}_{i=1}^p} = \sum_{q=1}^k A_{\sum_{i=q}^k j_i+k-q}^{k-q+1} + 1。$$

定理 4.1 的好处在于, matlab 可实现矩阵的点乘, 即矩阵对应元素相乘, 因此我们可以在以前的计算结果上进行重要值的计算, 无疑节省了大量时间。

定理 3.2 广义多项式神经网络输入变量筛选定理

设输入输出变量已归一化, 规定第  $i$  个输入变量第  $s$  个样本的偏导数小于  $k$  即为不显著, 并称  $k$  为筛选因子, 则若一个输入变量重要值的绝对值低于下式即可认为该输入变量不显著, 可以剔除:

$$\sum_{s=1}^d k \rho(s) \quad (8)$$

## 4 GUI 设计及实证研究

### 4.1 GUI 设计

设计一个面向用户的用于预测股指的多项式神经网络应考虑以下几个问题:

(1) 应建立起与炒股软件的接口程序。用户在进行股价预测的时候, 如果能够从其常用的炒股软件中动态读取数据并进行计算将会非常方便, 同时也能够增强 GUI 的应用性。

(2) 应能根据用户输入的股票代码、预测日期、样本容量进行正确的识别。如, 在预测日期方面, 如果用户输入日期恰好是休市则应提醒用户重新选择预测日期, 如果用火输入日期早于或者晚于数据库, 则应提醒用户打开炒股软件进行数据更新。

(3) 应提供输入变量滞后期数的可能, 并对其进行正确的识别。如, 影响股票价格的因素很可能不是上 1 期, 而是滞后 2 期、3 期甚至 4 期, 这时我们就应该允许用户输入相应滞后期数进行变量输入; 同时对于滞后期数据超过数据库中的数据的时候, 应提醒用户重新输入或者更新数据库。

同时, 应提供设定预测变量类别的可能。如, 用户如欲预测收盘价, 则其在界面中输入“收盘价”

即可进行以收盘价为输出变量的预测; 用户如欲预测开盘价, 则其在界面中输入“开盘价”即可进行以开盘价为输出变量的预测。对于用户输入的数据库中不存在的预测变量, 应提醒用户重新修改。

此外, 应提供多项式神经网络、标准 BP 神经网络、基于遗传算法优化的 BP 神经网络(GA-BP 神经网络)的相应程序, 并输出相应的拟合误差、隐层节点个数、网络训练时间等。

另外, 需提供用户指定的输入变量的重要值及参考值, 以供用户筛选输入变量。与此同时, 应绘出各输入变量的偏导数变化的图像, 并能够让用户进行任意两个输入变量的偏导数分析。对于用户输入的非用户指定的输入, 应提醒其重新输入。

通过以上设计思路, 本文设计了如图 1 所示的 GUI 界面, 其中包含“读取数据”、“预测”、“对比” 3 个按钮。同时将其生成的 m 文件命名为“stock”。



图 1 多项式神经网络股指预测 GUI 设计

## 4.2 实证研究

### 4.2.1 归一化与权函数

就 BP 神经网络而言, 由于其采用的 sigmoid 函数值在 0, 1 附近的时候曲线变化平缓, 导数值

非常小, 容易导致计算时间过长。因此, 为了减少学习时间, 一般将输入和输出变换到[0.1, 0.9]之间, 这样 sigmoid 函数就可以在此区间内有较快的变化, 缩短训练时间。同时, 对于多项式神经网络

而言, 定理 3.1 中  $t_i = \sum_{s=1}^d \rho(s) \sum_{l=1}^N \frac{j_l w_l g_l^{(s)}}{x_i}$  的计算要求  $x_i$  不能为 0, 因此将数据变换到 [0.1, 0.9] 之中是合情合理的, 这也是本文所采用的方法。采用的公式为:

$$x^* = 0.1 + \frac{0.8(x - x_{\min})}{x_{\max} - x_{\min}} \quad (x^* \in [0.1, 0.9]) \quad (9)$$

在计算网络之后获得的价格数据应进行反归一化, 返回到原始数据的量纲, 采取的公式为:

$$x = \frac{(x^* - 0.1)(x_{\max} - x_{\min})}{0.8} + x_{\min} \quad (10)$$

由于预测股票价格的时候, 离预测日期越近的数据其影响越大, 而离预测日期越远的数据其影响越小, 因此引入权函数来区分不同时期的偏导数的不同重要程度, 即确定  $t_i = \sum_{s=1}^d \rho(s) \sum_{l=1}^N \frac{j_l w_l g_l^{(s)}}{x_i}$  中的  $\rho(\cdot)$ 。本文采用  $\rho(x) = x/n$ ,  $n$  为总项数, 这种权函数假定权重是以等差数列递增的。

#### 4.2.2 单次预测——多项式神经网络解释能力

为说明多项式神经网络具备解释能力, 现选择中国上证指数(000001)的收盘价作为预测对象, 开盘价、最高价、最低价、收盘价、成交金额和成交量的滞后一期、滞后二期作为模型的输入变量。设定预测日期为 2013-05-06, 筛选因子  $k=0.2$ , 目标预测精度 = 0.001, 样本容量为 10 (即训练样本是 2013-05-06 前 10 个交易日), 对收盘价进行预测。

初步预测, 多项式神经网络预测相对误差为 -0.88%。成交金额的滞后二期变量重要值为 0.81, 不重要, 因此可以剔除。剔除后, 相对误差上升为 -0.95%, 变化不大 (关于剔除变量是否有效的讨论详见下节)。从图 2 中可知, 滞后二期开盘价的最为重要, 滞后二期最高价和滞后二期收盘价对这一区间的收盘价影响也比较显著。相比之下滞后一期开盘价、滞后一期收盘价、滞后一期成交金额对这一区间的收盘价影响较少。



图 2 单次预测——剔除不重要变量后的相对误差、重要值及偏导数图像

从偏导数图像中可知, 在预测日期 2013-05-07 日之前的 10 个交易日样本区内, 大部分的变量对收盘价的影响由弱变强, 并有由强变弱的趋势。滞后一期开盘价、滞后一期最高价、滞后一期收盘价、滞后一期成交金额对收盘价的影响较为平缓、微弱, 而滞后二期开盘价、滞后二期最高价和滞后二期收盘价对收盘价影响较为剧烈, 波动幅度较大。

现为直观地说明各输入变量的重要性变化情况, 分别对比滞后二期开盘价与滞后二期最高价(高重要值对高重要值)、滞后一期收盘价与滞后一期最高价(低重要值对低重要值)、滞后二期开盘价与滞后一期收盘价(高重要值对低重要值)。

从图 3 的左图可知, 同样是对模型影响强的两个输入变量, 在本预测区间内, 滞后二期最高价对当期收盘价的影响倾向于负影响, 而滞后二期开盘价恰好相反。滞后二期的变量重要, 说明该类输入变量的趋势能够显著影响当期收盘价。随着时间的推移, 2 个变量对收盘价的影响都有由强变弱的趋势, 说明原本显著影响市场的变量开始更替, 或者市场承受来自外部的影响逐渐变强。外部因素包括: 货币政策、宏观经济景气度、外资、政治风险等。从图 3 的中图可知, 同样作为低重要值的输入变量, 其形成原理完全不同。滞后一期收盘价在开始阶段较为重要, 但其随后逐渐减弱, 以致其以时间加权后重要值很低; 滞后一期最高价则是因为其偏导数在整个预测区间内在零附近波动, 因此在整个区间内都表现出了低重要值。对于滞后一期收盘价来讲, 其可能在前几个预测日期内是重要性高的变量, 但是随后由于受到其他变量的替代或者外部因素逐渐增强的影响, 其对当期收盘价的影响逐渐衰弱。从图 3 的右图可知, 滞后二期开盘价与滞后一期收盘价对当期收盘价的影响都是正向的, 但是在整个预测区间内, 滞后二期开盘价比滞后一期收盘价重要得多。值得指出的是, 这两种变量的趋势比较相似, 说明这 2 种变量可能存在一定的相关性。

表 1 是 matlab 命令窗口返回的变量偏导数矩阵。由于该偏导数矩阵是用归一化后的数据计算的,

因此下面陈述中所涉及到的输入变量和输出变量全部是归一化后的变量。另外, 下面所述收盘价为当期收盘价简称。因此, 在预测日期 2013-05-07 之前的 10 个交易日样本区内:

滞后二期开盘价上升 1, 收盘价先是上升 2.65, 然后随时间其影响先由弱变强, 再由强变弱。至最后一日, 滞后二期开盘价上升 1, 收盘价上升 2.93;

滞后一期开盘价上升 1, 收盘价先是上升 0.47, 然后随时间其影响先由弱变强, 再由强变弱。至最后一日, 滞后一期开盘价上升 1, 收盘价下降 0.03, 影响力较小;

滞后二期最高价上升 1, 收盘价先是下降 1.36, 然后随时间其影响先由弱变强, 再由强变弱。至最后一日, 滞后二期最高价上升 1, 收盘价下降 0.89, 其影响力逐渐减弱;

滞后一期最高价上升 1, 收盘价先是上升 0.03, 随后波动, 至第 7 个样本日期, 其对收盘价的影响变为反向。至最后一日, 滞后一期最高价上升 1, 收盘价下降 0.33;

滞后二期最低价上升 1, 收盘价先下降 0.55。随后, 其对收盘价的影响先由弱变强, 再由强变弱。至最后一日, 滞后二期最低价上升 1, 收盘价先下降 0.45。

滞后一期最低价上升 1, 收盘价先上升 1.26, 随后, 其对收盘价的影响先由弱变强, 再由强变弱, 至最后一日, 滞后一期最低价上升 1, 收盘价上升 0.53;

滞后二期收盘价上升 1, 收盘价先上升 1.26。随后, 其对收盘价的影响先由弱变强, 再由强变弱。至最后一日, 滞后二期收盘价上升 1, 收盘价上升 0.56;

滞后一期收盘价上升 1, 收盘价先下降 0.09。在整个区间内, 其对收盘价的影响方向不断变化。至最后一日, 滞后一期收盘价上升 1, 收盘价上升 0.06;

滞后一期成交金额上升 1, 收盘价先下降 0.03。随后, 其对收盘价影响方向变为正向, 并先由弱变

强，再由强变弱，最后影响方向再次变为负向。至最后一期，滞后一期成交金额上升 1，收盘价下降 0.05；

滞后二期成交量上升 1，收盘价先下降 0.56，随后，其对收盘价的影响先由弱变强，再由强变弱。至最后一期，滞后二期成交量上升 1，收盘价下降 0.47；

滞后一期成交量上升 1，收盘价先下降 0.88。

随后，其对收盘价的影响不断震荡。至最后一期，滞后一期成交量上升 1，收盘价下降 0.64。

在整个区间内，模型的拟合误差为  $7.76 \times 10^{-30}$ ，说明模型所选取的变量仅不能解释收盘价  $7.76 \times 10^{-30}$  的变异。另外，大部分输入变量对收盘价的影响呈现出先增强后减弱的情况，说明市场中的外部因素逐渐变弱随后逐渐增强，甚至有可能影响股票市场，并导致模型预测的相对误差开始增加。

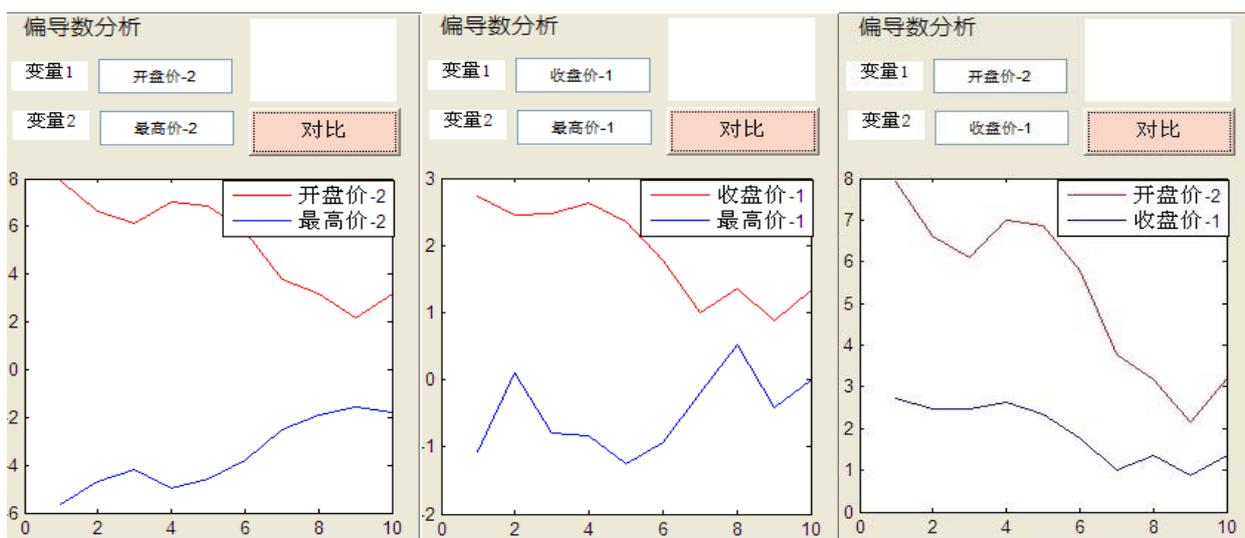


图 3 单次预测——偏导数图像分析

表 1 预测区间内各输入变量对收盘价的偏导数

日期序列	1	2	3	4	5	6	7	8	9	10
开盘价-2	2.65	1.27	2.67	1.75	4.44	5.75	5.62	2.40	3.96	2.93
开盘价-1	0.42	0.47	0.58	1.71	1.07	0.42	0.54	0.78	0.06	-0.03
最高价-2	-1.36	-1.52	-1.49	-3.44	-5.17	-4.33	-3.12	-3.33	-2.42	-0.89
最高价-1	0.03	0.50	0.56	0.95	0.50	0.07	-0.46	0.23	-0.23	-0.33
最低价-2	-0.55	-0.54	-0.45	-1.25	-2.28	-2.60	-1.30	-1.74	-1.37	-0.45
最低价-1	0.82	0.82	1.08	1.83	2.29	2.30	1.45	1.54	1.18	0.53
收盘价-2	1.26	1.29	1.54	3.47	3.74	3.01	2.22	2.58	1.45	0.56
收盘价-1	-0.09	0.30	0.41	-0.52	0.07	0.77	-0.50	-0.09	0.32	0.06
成交金额-1	-0.03	0.36	0.08	0.57	0.96	0.84	0.02	0.87	0.34	-0.05
成交量-2	-0.56	-0.80	-0.65	-1.22	-2.49	-2.05	-1.45	-1.48	-1.34	-0.47
成交量-1	-0.88	-0.24	-0.66	-1.31	-1.37	-1.42	-1.81	-0.87	-0.90	-0.64

4.2.3 滚动预测——多项式神经网络预测能力

一个模型的预测能力应该包括：模型的预测准确度、网络拟合误差、网络计算时间和网络隐层节点数等。在衡量模型的准确度时，可以选用绝对误差、相对误差和绝对相对误差。为尽可能地科学衡

量预测准确度，选择绝对相对误差作为衡量指标。

选用上证指数(000001)的收盘价作为预测对象，开盘价、最高价、最低价、收盘价、成交金额和成交量的滞后 1 期~滞后 5 期作为模型的输入变量。选取这些输入变量的原因在于，收盘价的滞后

变量包含收盘价的趋势信息, 而开盘价、最高价、最低价、成交金额和成交量包含投资者的心理因素。之所以没有选取大部分文献选取的大量技术指标, 是因为技术指标是由上述六种变量通过线性或非线性关系构造、以期预测股票走势的关系式, 而多项式神经网络本身就具备非线性逼近的能力, 因此选用上述 6 种变量的同时再选用技术指标会导致网络重复和冗杂, 并且也没有必要。对于宏观因素, 如利率、货币政策等, 由于其对股市的影响是逐渐渗透、循序渐进的, 并且炒股软件不能实时提供这些数据, 因此模型暂不考虑将这些因素作为输入。股市的交易周期一般为 5 日, 因此选取五期的滞后变量作为输入变量。

现设定筛选因子  $k = 0.2$ , 目标预测精度 = 0.001, 样本容量为 10, 对 2013-01-01~2013-05-01 之间共 72 个交易日的上证指数收盘价进行滚动预测。

#### (1) 预测准确度

图 4 显示了 2013-01-01~2013-05-01 多项式神经网络、标准 BP 神经网络、基于遗传算法的 BP 神经网络和筛选变量后的多项式神经网络预测精确度。从图中可以看到, 多项式神经网络的整体预测精确度较高, 绝对相对误差较为稳定; GA+BP 模型的预测精确度较差, 并且波动性大。采用了筛选变量法

后, 模型的预测精度变化不大, 说明我们剔除的变量对于模型没有解释能力, 筛选变量法是有效的。

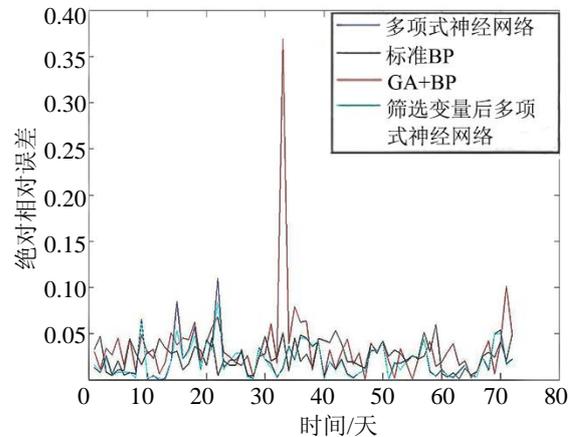


图 4 2013-01-01~2013-05-01 各模型预测绝对相对误差

表 2 列示了各模型的平均绝对相对误差。其中, 遗传算法 BP 神经网络的预测精度最差, 而筛选变量后的多项式网络预测精度最高。此外, 标准 BP 神经网络的预测精度要优于基于遗传算法的 BP 神经网络。当筛选变量后, 多项式神经网络的预测精度有所提高, 平均绝对相对误差由 2.19% 下降到 2.03%。值得指出的是, BP 类神经网络由于采用迭代算法, 因此其权值矩阵每次都不相同, 导致反复预测其结果不一, 这为其应用带来了困难。

表 2 各模型在预测区间内的平均绝对相对误差

	多项式网络	标准 BP 神经网络	遗传算法 BP 神经网络	筛选变量后的多项式神经网络
平均绝对相对误差	2.19%	2.68%	3.59%	2.03%

综合图 4 和表 2, 我们大致可知, 各模型预测精度的关系为, 基于遗传算法的 BP 神经网络次于标准 BP 神经网络, 标准 BP 神经网络次于多项式神经网络, 多项式神经网络又次于筛选变量后的多项式神经网络。为验证这一关系, 现做假设检验。

设基于遗传算法的 BP 神经网络、标准 BP 神经网络、多项式神经网络、筛选变量后的多项式神经网络的绝对相对误差总体均值分别为  $\mu_1, \mu_2, \mu_3, \mu_4$ 。由于样本数为 72, 根据大数定律, 可采用 Z 统计量进行检验。取  $\alpha = 0.05$ 。

从表 3 中可以发现, 在置信度为 0.05 的条件

下, 基于遗传算法的 BP 神经网络绝对相对误差显著高于标准 BP 神经网络, 标准 BP 神经网络绝对相对误差显著高于多项式神经网络。基于遗传算法的 BP 神经网络采用遗传算法计算最优初始权值向量, 因此计算的权值向量应比标准 BP 神经网络更接近真实值, 而现在基于遗传算法的 BP 神经网络预测精度反而不能比得上标准 BP 神经网络, 说明 BP 神经网络的逼近效果不如多项式神经网络。此外, 不能说明多项式神经网络绝对相对误差显著高于筛选变量后的多项式神经网络。因此, 进行下一步假设检验, 结果见表 4。

由表 4 可知, 在置信度为 0.05 的条件下, 多项式神经网络绝对相对误差等于筛选变量后的多项式神经网络绝对相对误差。虽然这一结果说明, 筛选变量并不能显著提高模型的预测能力, 但是可以验证, 通过本文的方法剔除的变量是不重要的变量, 剔除后不会对模型预测能力产生显著影响, 因此本文筛选变量的方法是有效果的、能够剔除不重要变量的。

## (2) 网络拟合误差和网络计算时间

表 5 列示了各模型的拟合误差。虽然我们事先规定了目标精度为 0.001, 但是由于各模型采用算法不同, 其结果也不同。标准 BP 神经网络由于采用迭代法并且达到目标精度迭代终止, 因此其拟合误差和目标精度基本一致。相反, 多项式神经网络采用权值直接确定法确定权值矩阵, 因此其拟合误差就是该隐层神经元数目的最小误差。遗传算法的 BP 神经网络采用遗传算法计算最优初始权值, 因此拟合误差也不局限于用户设定的精度。通过比较四种模型的拟合误差, 我们发现, 标准 BP 神经网络的拟合误差最大, 逼近效果最差, 基于遗传算法的 BP 神经网络能够显著提高 BP 神经网络的逼近程度, 但是与多项式神经网络相比, 其拟合误差显得很大。至于筛选变量前后的多项式神经网络, 由于

删减了不重要的变量, 因此拟合误差变化不大。

下面检验筛选变量后多项式神经网络的拟合误差是否显著低于筛选前的多项式神经网络。设筛选前和筛选后多项式神经网络的拟合误差总体均值分别为  $\mu_5$ ,  $\mu_6$ 。取  $\alpha = 0.05$ , 结果见表 6。从中可知, 筛选变量前后, 多项式神经网络的拟合误差没有显著变化。

表 7 列示了各模型的计算时间。通过比较, 基于遗传算法 BP 的神经网络计算快于标准 BP 网络。实际上, 当将目标精度设定为更高时, 标准 BP 神经网络计算时间大幅增加, 通常超过 1 min, 而基于遗传算法 BP 神经网络的计算时间变化幅度较小。当然, 这些都无法与多项式神经网络的计算时间相比。由于采用直接确定法而非迭代法, 网络的权值计算相当快, 能达到低于 0.1 s。这一特性为日后将多项式神经网络用于实时网络预报提供了可能。

现检验剔除不重要变量后, 多项式网络计算时间是否显著降低。设筛选前和筛选后多项式神经网络的计算时间总体均值分别为  $\mu_7$ ,  $\mu_8$ 。取  $\alpha=0.05$ , 结果见表 8。通过计算, 我们发现, 筛选变量对神经网络的计算时间没有显著降低。

表 3 各模型预测精度比较的假设检验,  $\alpha = 0.05$

序号	原假设 $H_0$	备择假设 $H_1$	Z 统计量	临界值	决策
1	$\mu_1 - \mu_2 \leq 0$	$\mu_1 - \mu_2 > 0$	1.667 6	1.644 9	拒绝 $H_0$
2	$\mu_2 - \mu_3 \leq 0$	$\mu_2 - \mu_3 > 0$	1.705 8	1.644 9	拒绝 $H_0$
3	$\mu_3 - \mu_4 \leq 0$	$\mu_3 - \mu_4 > 0$	0.508 3	1.644 9	不能拒绝 $H_0$

表 4 多项式神经网络与筛选变量后多项式神经网络的进一步比较,  $\alpha=0.05$

序号	原假设 $H_0$	备择假设 $H_1$	Z 统计量	临界值	决策
4	$\mu_3 - \mu_4 = 0$	$\mu_3 - \mu_4 \neq 0$	1.667 6	1.960 0	不能拒绝 $H_0$

表 5 各模型的网络拟合误差

	多项式网络	标准 BP 神经网络	遗传算法 BP 神经网络	筛选变量后的多项式神经网络
网络拟合误差	$2.216 1 \times 10^{-29}$	$9.941 0 \times 10^{-4}$	$3.267 4 \times 10^{-13}$	$1.959 8 \times 10^{-29}$

表 6 筛选前、后多项式神经网络拟合误差检验,  $\alpha = 0.05$

序号	原假设 $H_0$	备择假设 $H_1$	Z 统计量	临界值	决策
5	$\mu_5 - \mu_6 \leq 0$	$\mu_5 - \mu_6 > 0$	0.441 9	1.644 9	不能拒绝 $H_0$
6	$\mu_5 - \mu_6 = 0$	$\mu_5 - \mu_6 \neq 0$	0.441 9	1.960 0	不能拒绝 $H_0$

表 7 各模型的网络计算时间/s

	多项式网络	标准 BP 神经网络	遗传算法 BP 神经网络	筛选变量后的多项式神经网络
网络计算时间	0.066 8	4.564 5	3.624 9	0.061 4

表 8 筛选前、后多项式神经网络计算时间检验,  $\alpha=0.05$ 

序号	原假设 $H_0$	备择假设 $H_1$	Z 统计量	临界值	决策
7	$\mu_8 - \mu_7 \geq 0$	$\mu_8 - \mu_7 < 0$	-1.177 4	-1.644 9	不能拒绝 $H_0$
8	$\mu_8 - \mu_7 = 0$	$\mu_8 - \mu_7 \neq 0$	-1.177 4	-1.960 0	不能拒绝 $H_0$

## 5 结论

主要研究了国内外都尚无系统研究的多项式神经网络, 并运用多项式神经网络对中国上证指数进行了滚动预测。

多项式神经网络具有计算速度快、解释能力强、拟合精度好等优点, 为系统研究多项式神经网络, 本文首先创造性地设计了自然次序上限、自然次序下限, 并证明了其表达式, 从而得出网络隐层节点数目的相关公式。其次, 证明了在降序字典法下, 通过给定各解释变量的次数, 可以确定该节点的次序。然后, 给出了多项式神经网络的数学表达通式。此外, 提出并证明了多项式神经网络权值的迭代解法和权值直接确定法。其中, 权值直接确定法首次被运用于多元多项式神经网络的求解之中。最后, 创造性地设计了重要值及筛选定理。通过计算重要值, 神经网络可以利用自身进行输入变量的筛选。同时, 利用偏导数可以解释各输入变量对输出变量的贡献, 从而赋予模型解释能力, 这在以前的研究中几无涉及。

通过设计的基于 Matlab 的图形用户界面, 用户可通过炒股软件更新股票数据, 并读取特定股票、特定容量的数据, 并进行指定日期的预测。通过实证研究, 首先用单次预测说明了多项式神经网络在实际运用中筛选变量和解释模型的方法, 并验证了多项式神经网络具备筛选变量与解释的能力。其次, 通过对 2013-01-01~2013-05-01 共 72 个交易日的收盘价滚动预测, 发现在置信度为 0.05 的条件下, 多项式神经网络预测精度显著优于标准 BP 神经网络与基于遗传算法的 BP 神经网络, 同时其计算时间往往  $< 0.01$  s, 远远小于上述 2 类 BP 神经

网络, 并且, 多项式神经网络的拟合误差能够达到  $10^{-28}$  以下, 远远优于基于遗传算法的 BP 神经网络的  $10^{-13}$  以下及标准 BP 神经网络的用户指定的低精度。此外, 当采用了根据重要值筛选变量的方法后, 网络预测精度、拟合误差、计算时间没有显著变化, 但是隐层节点数却显著下降, 说明通过剔除不重要变量, 网络变得更为简洁, 并且不再有与输出变量无关的输入变量影响网络的解释能力。

通过本文讨论的多项式神经网络及相应的 Matlab 程序, 投资者可提前掌握指定日期指定股票的价格信息, 从而对其投资策略提供帮助; 通过重要值和偏导数图像, 投资者可进一步分析各输入变量的重要性增减程度, 从而判断其他因素(货币政策、宏观经济景气度、外资、政治风险等)是否存在于股票市场。同时, 政府相关人员可以用该模型得到指定日期的指数预测值, 并通过检验预测相对误差来判断市场是否存在其他因素、是否按预期的目标进行调整, 并据此进行监管和调控。此外, 多项式神经网络不像 BP 类神经网络那样每次预测结果都不同, 并且计算缓慢, 其良好的计算能力、预测能力、筛选能力和解释能力将支持未来股票市场实时预测。

## 参考文献:

- [1] 袁曾任. 人工神经网络及其应用 [M]. 北京: 清华大学出版社, 1999.
- [2] 刘永红. 神经网络理论的发展与前沿问题 [J]. 信息与控制, 1999, 28(1): 31-46.
- [3] Simon Haykin. 神经网络原理 [M]. 叶世伟, 史忠植译. 北京: 机械工业出版社, 2004.
- [4] 魏海坤, 徐嗣鑫, 宋文忠. 神经网络的泛化理论和泛化方法 [J]. 自动化学报, 2001, 27(6): 806-815.

- [5] 杨旭华. 神经网络及其在控制中的应用研究 [D]. 杭州: 浙江大学, 2004.
- [6] 孙富春, 李莉, 孙增圻. 非线性系统神经网络自适应控制的发展现状及展望 [J]. 控制理论与应用, 2005, 22(2): 254-260.
- [7] 董聪. 人工神经网络: 当前的进展与问题 [J]. 科技导报, 1999(7): 26-29.
- [8] 沈政. 人工神经网络研究的局限性 [J]. 心理学动态, 1991(2): 5-8.
- [9] 蒲小平, 陈克兴, 戴桂康. 神经网络的哲学意义 [J]. 自然辩证法研究, 1993, 9(1): 18-23.
- [10] Shinying Huang, RuahuanTsaih, Wanying Lin. Unsupervised neural networks approach for understanding fraudulent financial reporting [J]. Industrial Management & Data Systems (S0263-5577), 2012, 112(2): 224-244.
- [11] Kelly E Fish, Richard S Segall. A visual analysis of learning rule effects and variable importance for neural networks data mining operations [J]. Kybernetes (S0368-492X), 2004, 33(7): 1127-1142.
- [12] Shee Q Wong, Nik R Hassan, Ehsan Feroz. The equity premium puzzle: an artificial neural network approach [J]. Review of Accounting and Finance (S1475-7702), 2007, 6(2): 150-161.
- [13] 贺昌政, 李晓峰, 俞海. BP 神经网络模型的新改进及其应用 [J]. 数学的实践与认识, 2002, 32(4): 554-561.
- [14] 沈巍. 股指波动预测模型的方法研究及应用 [M]. 北京: 知识产权出版社, 2011.
- [15] Rozaida Ghazali, AbirJaafar Hussain, Nazri Mohd Nawi, et al. Non-stationary and stationary prediction of financial time series using dynamic ridge polynomial neural network [J]. Neurocomputing (S0925-2312), 2009 (72): 2359-2367.
- [16] M Mehrabi, M Sharifpur, J P Meyer. Application of the FCM-based neuro-fuzzy inference system and genetic algorithm-polynomial neural network approaches to modeling the thermal conductivity of alumina-water nanofluids [J]. International Communication in Heat and Mass Transfer (S0735-1933), 2012 (39): 971-977.
- [17] Ivan Maric. Optimization of Self-organizing poly neural networks [J]. Expert Systems with Applications, (S0957-4174), 2013, 40(11): 4528-4538.
- [18] E Gomez-Ramirez, K Najim, E Ikonen. Forecasting time series with a new architecture for polynomial artificial neural network [J]. ScienceDirect (S1568-4946), 2007, 7(4): 1209-1216.
- [19] Yunong Zhang, W Li. Genenbauer orthogonal basis neural network and its weights-direct-determination method [J]. Electronics Letters (S0013-5194), 2009, 45(23): 1184-1185.
- [20] 张雨浓, 杨逸文, 李巍. 神经网络权值直接确定法 [M]. 广州: 中山大学出版社, 2010.
- [21] 王华成, 尹美群. BP 神经网络模型对公司价值评估中自由现金流量的时序预测 [J]. 统计与决策, 2005(18): 15-17.
- [22] 汤凌冰, 廖福源, 罗键. 模糊神经网络在股价预测中的应用 [J]. 系统工程, 2004, 22(2): 107-109.

## 《系统仿真学报》荣获“2014 中国最具国际影响力学术期刊”证书

由中国学术期刊（光盘版）电子杂志社与清华大学图书馆联合成立的中国学术文献国际评价研究中心，以美国汤森路透 Web of Science 收录的 1.2 万余种期刊为引文统计源，首次研制发布了 2012《中国学术期刊国际引证年报》（CAJ-IJCR 年报）。第一次给出了我国 5 600 余种中外文学术期刊总被引频次、影响因子、半衰期等各项国际引证指标，并采用了新的国际影响力综合评价指标 CI 对期刊排序，发布了“中国最具国际影响力学术期刊”（排序 TOP5%）和“中国国际影响力优秀学术期刊”（排序 TOP5-10%），在国内外学术界产生了较大反响。之后，2013 年版年报，将引文统计源期刊扩展到 1.44 万多种。目前，2014 版国际、国内年报与 TOP5% 和 TOP5-10% 期刊的遴选业已完成，《系统仿真学报》被列入“2014 中国最具国际影响力学术期刊”行列。

我学报连续 2 年被列入 TOP5% 国内一流的中国最具国际影响力学术期刊，走向世界，进入国际一流，指日可待！

<http://www.china-simulation.com>