

【应用经济学研究】

以广义回归神经网络预测共同基金报酬

潘文超

(兰阳技术学院 信息管理系, 台湾 台北 104)

摘要: 鉴于近年来许多相关文献成功地运用广义回归神经网络进行财经方面的预测, 以及国内共同基金净值之预测与报酬率评估。通过搜集国内基金资料, 以灰关联分析法进行各基金投资绩效分析, 挑选投资绩效良好的共同基金作为投资标的; 再以广义回归神经网络建立预测模型, 与灰预测模型、多元回归模型进行预测能力及报酬率的比较分析。5种预测绩效评价指数、5组数据交互验证散布图及报酬率分析表明: 广义回归神经网络在预测能力及预测报酬率上均有很好的表现。

关键词: 灰关联分析; 灰预测; 广义回归神经网络; 多元回归模型; 遗传算法

中图分类号: F7830.91

文献标志码: A

文章编号: 1671-6248(2007)04-0055-04

Forecast for mutual fund returns with general regression neural network

PAN Wen-chao

(Department of Information Management, Lanyang Institute of Technology, Taipei 104, Taiwan, China)

Abstract: In recent years, there are many relevant documents that are successfully in general regression neural network for the financial sector forecasts. This paper adopts the general regression neural network for the prediction of the net value of the domestic mutual fund and for evaluation of the return of the investment. The author picks out a lot of fund information at home, analyzes its investment performance with grey relational analysis, and select some good investment performances of mutual funds as investment targets. Through general regression neural network model, he sets up the prediction model and with grey prediction and multiple regression model, he conducts the comparative analysis on the accuracy of the prediction and the return rate. It is found that it is better to predict the return rate with general regression neural network than with grey prediction and multiple regression model. On the basis of the evaluation of the 5 indexes of the performance management, and 5 group interactive data validation map, the general regression neural network can perform well in prediction and the prediction of return rates.

Key words: grey relational analysis; grey prediction; general regression neural network; multiple regression; genetic algorithm

0 引言

共同基金是投资人最喜爱的投资工具之一, 其优点包括了专业的投资管理、流动性佳、分散投资风

险、合法节税管道与多元的投资标的等。然而以往学者或研究者使用人工智能方法进行基金净值之预测, 大多实行遗传算法^[1]或类神经网络^[2]建构预测模型。本文有别于以往, 采用类神经网络中较为新

收稿日期: 2007-08-13

作者简介: 潘文超(1966-), 男, 江苏南京人, 副教授。

型的广义回归神经网络进行共同基金投资报酬率之预测。本文首先自情报赢家数据库搜集国内 53 组开放平衡型基金数据,对报酬率数据进行灰关联分析,挑选出投资绩效前 2 名的基金作为投资标的,并搜集标的基金 2002 年至 2006 年周数据的净值市价,进行广义回归神经网络、灰预测及多元回归预测模型的建构与绩效评价。

广义回归神经网络在预测连续变量的问题上具有很好的广义化能力,目前已成功地应用于各种领域。Bowden 等人运用广义回归神经网络在水源分布系统中预测氯元素含量^[3],以有效控制水源分布区氯元素含量的多少,研究变量包括各区域的氯元素含量、水流量、温度及 PH 值等。Jeyamkondan 等人采用广义回归神经网络预测 *Aeromonas Hydrophila*、*Shigella Flexneri* 及 *Brochothrix Thermosphacta* 共 3 种微生物的成长^[4],使用的研究变量包括温度、氯化铀的百分比及 PH 值等。杜壮与陈勇雄应用广义回归神经网络建立快餐业的服务复原诊断系统^[5]。预测结果显示,广义回归神经网络优于倒传递类神经网络模型。

1 广义回归神经网络

Specht 提出机率神经网络架构^[6],它是一种监督式网络架构,其理论是建立在贝氏决策和无母数技巧上估计机率密度函数,此机率密度函数型态为高斯分布。叶怡成指出此一函数如式(1)所示^[7]。

$$f_k(X) = \left(\frac{1}{N_k}\right) \left[\frac{1}{(2\pi)^{m/2}} \right] \left(\frac{1}{\sigma^m}\right) \cdot \sum_{j=1}^{N_k} \exp\left(-\frac{\|X - X_{kj}\|^2}{2\sigma^2}\right) \quad (1)$$

因机率神经网络适用于一般分类的问题,假设特分类的特征向量必属于已知分类之一,则各分类的绝对机率值大小并不重要,只需考虑其相对大小值,因此式(1)中的 $\left[\frac{1}{(2\pi)^{m/2}}\right] \left(\frac{1}{\sigma^m}\right)$ 可忽略,则式(1)可简化为

$$f_k(X) = \left(\frac{1}{N_k}\right) \sum_{j=1}^{N_k} \exp\left(-\frac{\|X - X_{kj}\|^2}{2\sigma^2}\right) \quad (2)$$

其中,σ是机率神经网络的平滑参数,网络训练完成之后,其预测的精确度可由调整平滑参数σ而提高,其值愈大就会有愈平滑的逼近函数。如果平滑参数σ选择不当,会造成网络设计中神经元数目过多或过少,在函数逼近中就会造成过度配适和配适不当,使预测能力降低。令 $d_{ij}^2 = \|X - X_{kj}\|^2$ 为样本空间中X与 X_{kj} 二点的欧几里得距离平方,则式(2)可改

写成

$$f_k(X) = \left(\frac{1}{N_k}\right) \sum_{j=1}^{N_k} \exp\left[-\frac{1}{2}\left(\frac{d_{kj}}{\sigma}\right)^2\right] \quad (3)$$

机率神经网络广义化能力有赖于调整平滑参数σ。在式(1)中,当平滑参数σ趋近于零时, $f_k(X) = \frac{1}{N_k}$;如果 $X = X_{kj}$,则 $f_k(X) = 0$,此时机率神经网络完全依未分类样本最接近的已分类样本决定其分类。当平滑参数σ趋近于无限大时, $f_k(X) = 1$,此时机率神经网络接近盲目分类。然而,由于机率神经网络只能进行分类问题之研究,因此,Specht 又利用机率神经网络演变出广义回归神经网络^[8],以解决连续变量的问题。广义回归神经网络不仅可以用于分类问题的研究,而且对于建构预测模型或控制模型、线性或非线性的回归问题也有很好的预测能力。

广义回归神经网络类似机率神经网络为 4 层类神经网络(图 1),第一层为输入层,神经元个数为自变量的个数,并接受输入数据;第二层为样式层,存放每一笔训练数据;而样式层输出的数据会经过第三层(汇总层)的神经元而对应每一可能的类别,此层进行式(3)的计算;第四层为线性层,与机率神经网络不同,此层执行汇总层的输出加权计算,并且产生输出值。

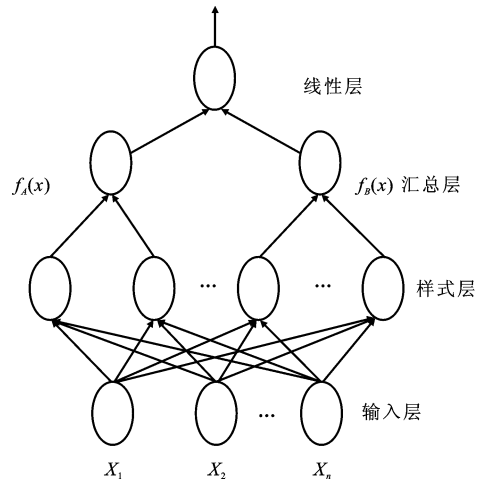


图 1 广义回归神经网络架构

2 样本数据与变量

本文由情报赢家数据库扣除数据不齐全者,共搜集 53 种开放平衡型基金之 8 项报酬率资料(包括 3 个月、6 个月、9 个月、1 年、2 年、3 年、今年以来以及自成立以来报酬率),并采用 Deng 所提出的局部灰关联分析进行投资绩效分析^[9],以挑选投资绩效

前 2 名的基金作为投资标的。通过分析, 投资绩效前 2 名的基金分别为元大双盈证券基金(Γ 值为 0.964 9)和保德信金平衡基金(Γ 值为 0.799)。因此, 本文以这 2 种基金作为研究对象, 搜集该基金自 2002 年至 2006 年共 254 笔净值周资料, 进行预测模型建构与绩效评估。表 1 为这 2 种基金净值资料之叙述统计值, 再由图 2 发现这 2 种基金净值都有逐渐上扬的趋势。

表 1 2 种基金净值之叙述统计值

企业	个数	最大值	最小值	平均值	标准差
元大	254	15.8	8.62	10.841	1.719
保德信	254	16.18	8.9	11.563	1.723

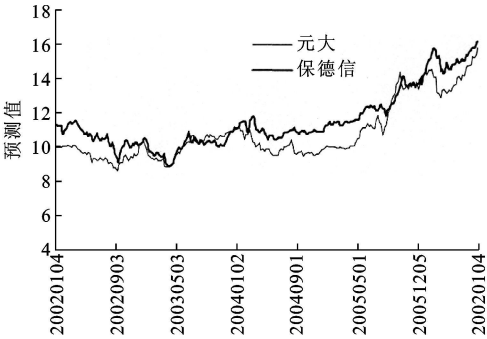


图 2 2 种基金净值趋势

为配合灰预测 4 点滚动建模, 本文不采用基金绩效指标, 而采用以前 4 期预测后 1 期的方式进行广义回归神经网络、灰预测及多元回归模型建构。因此, 将 254 笔数据分为 5 组, 前 4 组为训练数据, 后 1 组为测试数据, 然后进行数据的正规化, 将数据转换成 0 到 1 之间的数值。

3 3 种预测模型之建构与预测能力综合比较

本文在广义回归神经网络方面采用 MATLAB 自行撰写程序进行模型建构, 在测试期间采用遗传算法调整广义回归神经网络的平滑参数 σ。在 MATLAB 程序中, 广义回归神经网络可调整其参数 Spread 数值(也就是平滑参数 σ)。遗传算法的参数设定值包括: 最大遗传代数 为 200、染色体数目 为 4、基因交配率为 0.7、突变率为 0.01 及平滑参数初始值为 1。然而调整参数的做法是将训练数据带入广义回归神经网络进行网络训练, 再将广义回归神经网络的平滑参数 σ 进行编码换算, 并且将数值带入广义回归神经网络, 同时将测试数据带入广义回归神经网络进行一世代的网络测试, 得到一输出值, 并与因变量 Y 计算出误差均方根, 其公式为

$$\sqrt{\frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N}}$$
。此误差均方根就是遗传算法中的

目标函数, 再将此目标函数转换成适应函数, 然后进行遗传算法的 3 种基本操作(复制、交配及突变)。

本文在复制过程中, 所采用的方式是轮盘选择法, 就是按照适应函数值大小分布面积在轮盘上, 以供随机选取。在交配的过程中, 采用单点交配; 也就是, 在一对染色体中, 以单点的方式随机选取互换基因。在突变的过程中, 采取单点突变, 在单一染色体中, 以单点的方式随机产生基因突变。经过 200 个世代的演化之后, 会使得预测值逐渐逼近目标值。图 3 显示了测试数据, 经过了 200 个世代的演化后误差均方根值逐渐下降, 预测值逐渐逼近目标值的情况。

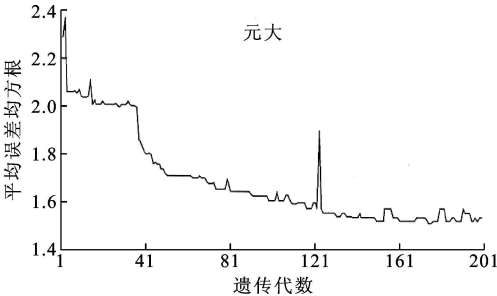


图 3 广义回归神经网络测试结果之误差均方根趋势

在灰预测方面, 本文采用 4 点滚动建模方式进行预测模型的建构及测试, 而多元回归模型的配适方式则采用全部选取法。3 种模型预测能力的比较, 采用 5 种评价指标进行, 分别为: 误差均方根公

式为 $\sqrt{\frac{\sum_{t=1}^n x_t - \hat{x}_t}{N}}$; 修正 Theil 不等系数公式为

$$\frac{\sum_{t=1}^N (X_t - \hat{X}_t)^2}{\sum_{t=1}^N X_t^2}$$
; 平均绝对误差公式为 $\frac{1}{M} \sum_{t=1}^M |Z_{t+1} -$

$Z_t(I)|$ (其中 M 为预测值数目; Z_{t+1} 为 1 时刻之观测值; $Z_t(I)$ 为 1 时刻之估计值); 平均绝对误差百分

比公式为 $\left\{ \frac{1}{M} \sum_{t=1}^M \left| \frac{Z_{t+1} - Z_t(I)}{Z_{t+1}} \right| \right\} \times 100\%$; 效率

系数公式为 $1 - \frac{\sum (x_t - \hat{x}_t)^2}{\sum (x_t - \bar{x}_t)^2}$ 。

在这 5 种评价指标中, 第一到第四种指标愈接近 0, 模型准确度愈高。而第五种指标愈接近 1, 模型准确度也愈高。分析测试资料结果如表 2 所示。

表 2 5 种评价指标分析结果

基金	模型	误差均方根	修正 Theil 不等系数	平均绝对误差	平均绝对误差百分比	效率系数
元大	灰预测	0.697	0.051	0.515	0.039	0.921
	广义回归神经网络	0.363	0.026	0.280	0.021	0.979
	多元回归预测	0.598	0.042	0.501	0.068	0.942
保德信	灰预测	0.503	0.250	0.667	0.029	0.975
	广义回归神经网络	0.274	0.019	0.346	0.015	0.922
	多元回归预测	0.522	0.035	0.761	0.033	0.973

由表 2 可知, 广义回归神经网络在 5 种评价指标中, 准确度都相对高于灰预测及多元回归模型的。本文进一步将 5 组样本数据进行交互验证, 得到预测结果与实际值资料共 250 种画成散布图(图 4)。图 4 中散布的各个观测点愈集中于中央对角线, 模型之预测能力也愈高。这显示广义回归神经网络在元大双盈证券基金各组资料交互验证预测结果中, 有很好的预测能力。本文紧接着进行各模型之报酬率的计算, 在 3 种模型的累积投资报酬率计算中, 报酬率的公式为 $\frac{P_t-P_0}{P_0} \times 100\%$, 其中 P_t 为下一期基金净值, P_0 为本期基金净值。

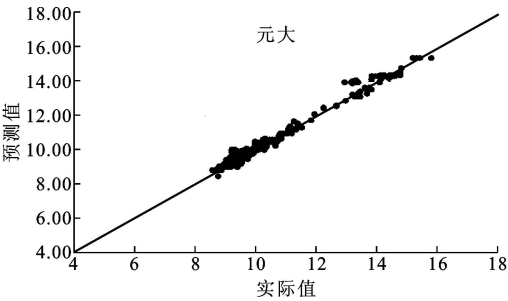


图 4 广义回归神经网络交互验证之测试值与实际值散布

投资决策为各模型测试数据的预测结果, 若下一期预测值高于本期预测值, 则实施买进策略。若下一期基金净值的投资报酬率为正, 则该模型可得正的投资报酬率; 反之, 若下一期基金净值投资报酬率为负, 则该模型的投资报酬率为负。将所有测试数据的投资报酬率加以累计, 结果如表 3 所示。

由表 3 可知, 广义回归神经网络所预测的报酬率最高, 元大双盈证券基金的报酬率为 41.922, 保德信金平衡基金的报酬率为 23.719, 多元回归模型次之, 灰预测模型最差。再观察这 2 种基金的预测平均投资报酬率, 以元大双盈证券基金较高, 平均报酬为 20.883。

表 3 3 种预测模型的累积投资报酬率 %

基金名	灰预测	广义回归神经网络	多元回归预测	平均
元大	2.014	41.922	18.713	20.883
保德信	15.812	23.719	20.843	20.124

4 结 语

本文搜集国内 53 组开放平衡型基金资料进行局部灰关联分析, 选取投资报酬绩效较佳的基金作为投资标的, 并搜集标的基金的净值市价, 进行广义回归神经网络、灰预测及多元回归预测模型的建构与报酬率评估。研究发现, 广义回归神经网络在模型预测绩效或累积报酬率上都优于灰预测及多元回归预测模型。而在 2 档基金的预测投资报酬率上, 以元大双盈证券基金为最高。此外, 本文发现属于资料勘探技术的灰预测模型, 其预测报酬率不如传统多元回归模型。然而, 本文是以基金周数据进行模型建构与预测, 建议未来可采用月数据进行模型预测分析。另外, 本文是以前 4 期预测后 1 期的方式进行模型建置, 未来可考率使用基金绩效指标(如 Sharpe、Jensen)作为研究变量进行建模。

参考文献:

[1] 李沃墙. 共同基金的绩效评估及净值预测-遗传算法的运用[J]. 台湾经济金融月刊, 1997, 33(12): 45-62.

[2] 余尚武, 许意铃. 共同基金净值之预测——灰色理论、类神经网络及适应性类神经模糊推论系统之应用[J]. 中华管理评论, 2005, 8(3): 1-33.

[3] Bowden G J, Nixon J B, Dandy G C, et al. Forecasting chlorine residuals in a water distribution system using a general regression neural network[J]. Mathematical and Computer Modelling, 2006, 44(5/6): 469-484.

[4] Jeyamkondan S, Jayas D S, Holley R A. Microbial growth modelling with artificial neural networks[J]. International Journal of Food Microbiology, 2004, 64(3): 343-354.

[5] 杜 壮, 陈勇雄. 智能型服务复原诊断系统——以快餐业为例[M]. 台北: 儒林书局, 2001.

[6] Specht D F. Probabilistic neural networks and the polynomial adaline as complementary techniques for classification[J]. IEEE Trans on Neural Networks, 1990, 1(1): 111-121.

[7] 叶怡成. 应用类神经网络[M]. 台北: 儒林书局, 1998.

[8] Specht D F. A general regression neural network[J]. IEEE Tras Neural Networks, 1991, 2(6): 568-576.

[9] DENG Ju-long. The control problems of grey system [J]. System & Control Letters, 1982, 3(5): 288-294.