

基于4类神经网络的国债利率期限结构预测

刘 晓¹ 周荣喜^{2*} 李 杰¹

(1. 北京化工大学 经济管理学院, 北京 100029; 2. 对外经济贸易大学 金融学院, 北京 100029)

摘 要: 以2011年1月至2016年7月的国债月度数据为样本,研究了在利率期限结构预测中,反向传播神经网络(BPNN)、小波神经网络(WNN)、径向基神经网络(RBFNN)和广义回归神经网络(GRNN)4种网络中相关参数的设定对精度的影响,并对预测效果进行实证比较。研究表明:广义回归神经网络预测效果较好,反向传播神经网络预测结果波动性较小,小波神经网络和径向基神经网络预测结果波动性较大。

关键词: 利率期限结构; 预测; 神经网络; 参数优化

中图分类号: F830.9 **DOI:** 10.13543/j.bhxbzr.2017.03.019

引 言

利率期限结构是相同的风险水平下利率与到期期限之间的数量关系或理论上的零息票债券收益率曲线。利率期限结构在资产定价、金融风险管理和货币政策制定等方面具有重要作用,所以其获取方法是金融学研究的热点问题之一。

曲线拟合法是获取利率期限结构的传统方法之一,该方法通过假定收益率曲线的数学模型,并利用样本数据估计模型中的参数值,从而预测利率期限结构。常用的曲线拟合方法有:多项式样条函数利率期限结构模型^[1]、Nelson-Siegel(NS)模型^[2]、Svensson(SV)模型^[3-4]和双斜率因子动态利率期限结构模型^[5]等。但此类拟合模型均存在一个共性问题,即如果收益率曲线的真实情况不符合模型所设定的形式,那么必然存在模式设定误差,甚至模型设定错误,显然不利于利率期限结构预测。

神经网络作为一种新型的人工智能算法,具有自学习的特性和良好的非线性逼近能力,避免了传统计量模型主观性假定过多而造成的模型应用范围较小、对样本信息利用不充分的缺点。随着信息科技的发展,以神经网络为代表的人工智

能技术在经济预测问题中得到了广泛的应用,并取得了较好的效果^[6-9]。近年来,有学者尝试应用神经网络方法预测利率期限结构^[10-12],但是他们仅使用某一种神经网络,且只给出最终结果,并未针对网络参数的选取做进一步研究。本文将现有4类神经网络技术用于研究利率期限结构预测,利用部分国债到期收益率对网络进行训练,在此基础上通过对我国国债的到期收益率进行实证预测,比较了4个模型的优劣,并详细探讨网络参数选取对预测精度的影响,以期为准准确预测利率期限结构提供技术支持。

1 4类神经网络模型

1.1 反向传播神经网络

反向传播神经网络(BPNN)的拓扑结构如图1所示。

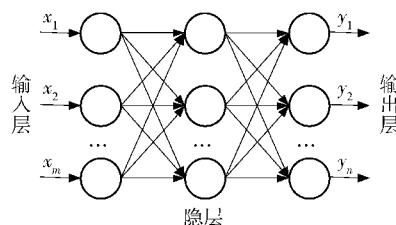


图1 BPNN的拓扑结构

Fig. 1 Topology of BPNN

输入序列 $\mathbf{X}_p = (x_1^p, x_2^p, \dots, x_m^p)^T$ ($p = 1, 2, \dots, W$), 其中 W 为样本总数。

第 p 个样本隐层第 j 个结点的输出值为:

$$h_j^p = f_j \left(a_{0j} + \sum_{i=1}^m \omega_{ij} x_i^p \right) \quad (j = 1, 2, \dots, L) \quad (1)$$

收稿日期: 2016-09-29

基金项目: 国家自然科学基金(71631005); 教育部人文社会科学规划基金(16YJA630078)

第一作者: 女, 1993年生, 硕士生

* 通讯联系人

E-mail: zrx103@126.com

式中, $f_j(\cdot)$ 为激活函数, ω_{ij} 为输入层和隐层之间的连接权值, a_{0j} 为常数项, j 为隐层结点数。本文中激活函数使用如下形式

$$f_j(z) = \frac{\exp(z)}{1 + \exp(z)} \quad (2)$$

BPNN 第 p 个样本输出层第 k 个结点值为:

$$y_k^p = \sum_{j=1}^L \omega_{jk} h_j^p \quad (k=1, 2, \dots, n) \quad (3)$$

其中, ω_{jk} 为隐层到输出层的权值, h_j^p 为第 j 个隐层结点的输出值, k 为输出层结点数。

1.2 小波神经网络

小波神经网络(WNN)是在反向传播神经网络的基础上,用小波元代替神经元,将小波函数作为激活函数。小波函数通过尺度伸缩和平移变换,能够有效地提取信息,抗噪声能力强。

输入序列 $\mathbf{X}_p = (x_1^p, x_2^p, \dots, x_m^p)^T$ ($p=1, 2, \dots, W$), W 为样本总数。

第 p 个样本隐层第 j 个结点的输出值为:

$$h_j^p = f_j \left(\left(\sum_{i=1}^m \omega_{ij} x_i^p - b_j \right) / a_j \right) \quad (j=1, 2, \dots, L) \quad (4)$$

式中, $f_j(\cdot)$ 为小波基函数, ω_{ij} 为输入层和隐层之间的连接权值, b_j 为 $f_j(\cdot)$ 的平移因子, a_j 为 $f_j(\cdot)$ 的伸缩因子, j 为隐层结点数。 $f_j(\cdot)$ 一般选取 Morlet 小波函数,即 $f_j(z) = \cos(1.75z) e^{-z^2/2}$ 。

WNN 的第 p 个样本输出层第 k 个结点值为

$$y_k^p = \sum_{j=1}^L \omega_{jk} h_j^p \quad (k=1, 2, \dots, n)$$

1.3 径向基神经网络

径向基神经网络(RBFNN)是典型的局部逼近神经网络,与其他神经网络相比,RBFNN 仅需要训练隐层与输出层之间的激活函数,故学习速度更快;但 RBFNN 需要设置更多的隐层结点,以此保证对数据的响应,缺点是易出现过拟合现象。

输入序列 $\mathbf{X}_p = (x_1^p, x_2^p, \dots, x_m^p)^T$ ($p=1, 2, \dots, W$), 其中 W 为样本总数。

隐层采用径向基函数作为激活函数,用于实现非线性变换,隐层第 j 个结点的输出值为:

$$h_j^p = f_j(\|\mathbf{X}_p - \mathbf{c}_j\|) \quad (j=1, 2, \dots, L) \quad (5)$$

式中, $f_j(\cdot)$ 为径向基函数, \mathbf{c}_j 为第 j 个结点函数的中心参数, $\|\cdot\|$ 为欧式范数, j 为隐层结点数。 $f_j(\cdot)$ 选取高斯函数作为径向基函数:

$$f_j(z) = \exp \left(-\frac{\|z - \mathbf{c}_j\|^2}{2\sigma^2} \right) \quad (j=1, 2, \dots, L) \quad (6)$$

其中 σ 为高斯函数的宽度,也称为平滑参数。 σ 越小, RBFNN 的拟合效果越精确,但拟合结果不平滑; σ 越大则相反。

第 p 个样本输出层第 k 个结点的预测值为

$$y_k^p = \sum_{j=1}^L \omega_{jk} h_j^p \quad (k=1, 2, \dots, n)$$

1.4 广义回归神经网络

广义回归神经网络(GRNN)是 RBFNN 的一个分支,它的模式层仍使用径向基函数,求和层通过线性函数对模式层的结点进行求和。与其他神经网络相比,GRNN 的隐层结点数是确定的,只需给定一个平滑参数,受主观影响小。GRNN 需要进行训练的参数只有平滑参数,网络收敛速度更快,训练时间更短。GRNN 在小样本预测方面优势更大,其拓扑结构如图 2 所示。

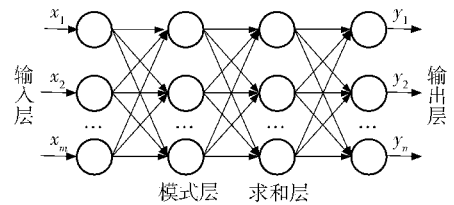


图2 GRNN 的拓扑结构

Fig. 2 Topology of GRNN

输入序列 $\mathbf{X}_p = (x_1^p, x_2^p, \dots, x_m^p)^T$, 期望目标值序列 $\mathbf{O}_p = (o_1^p, o_2^p, \dots, o_n^p)^T$ ($p=1, 2, \dots, W$), 其中 W 为样本总数。

模式层的结点数目等于样本总数,结点与样本一一对应。模式层第 j 个结点函数为:

$$h_j^p = \exp \left(-\frac{(\mathbf{X}_p - \mathbf{X}_j)^T (\mathbf{X}_p - \mathbf{X}_j)}{2\sigma^2} \right) \quad (7)$$

式中, $j=1, 2, \dots, W$, \mathbf{X}_j 为第 j 个结点对应的学习样本。

求和层中有两类神经元,一类神经元对模式层中所有结点进行算数求和,传递函数为

$$S^p = \sum_{j=1}^W h_j^p \quad (p=1, 2, \dots, W) \quad (8)$$

另一类神经元对模式层中所有结点进行加权求和,第 p 个样本的模式层中第 j 个结点的权值为第 j 个样本中的期望目标值序列中第 k 个元素。传递函数为

$$S_k^p = \sum_{j=1}^W o_j^p h_j^p \quad (k=1, 2, \dots, N) \quad (9)$$

第 p 个样本输出层第 k 个结点的预测值为

$$y_k^p = \frac{S_k^p}{S^p} \quad (k=1, 2, \dots, N) \quad (10)$$

2 神经网络拟合及预测模型

2.1 算法步骤

对于任意的神经网络,其一般算法步骤如下。

1)网络初始化。确定训练样本向量、目标向量和训练参数。

2)计算预测值序列。输入 (x_i^p, r_k^p) ($i=1,2,\cdots, m; k=1,2,\cdots, N$), x_i^p 为第 p 支国债的剩余年限, r_k^p 为第 p 支国债的到期收益率;设 y_k^p 为第 p 支国债的到期收益率的预测值,由式(11)计算总误差

$$E = \sum_{p=1}^W \sum_{k=1}^N (r_k^p - y_k^p)^2 \tag{11}$$

3)参数修正。修正网络权值和训练参数,使网络预测值逐步逼近期望值,如若达到最大迭代次数,则停止训练;否则,返回步骤 2)。

2.2 拟合及预测效果检验

为检验模型拟合及预测利率期限结构的效果,本文用训练神经网络后得到的利率函数作为贴现函数,对债券进行贴现并求得其理论价格。由于收益率计算出的偏差均值过小,难以比较出不同模型的区别,且债券定价是收益率预测的一个重要方面,本文采用理论价格和实际价格之间的绝对偏差均值来检验每日的拟合和预测效果,其计算公式为:

$$E_{MAD} = \frac{1}{W} \sum_{j=1}^W |\hat{P}_j - P_j|, j=1,2,\cdots, W \tag{12}$$

其中, \hat{P}_j 是理论价格, $\hat{P}_j = \sum_{i=1}^{n_j} C_j^i r(t_i)$ ($j=1,2,\cdots, W$); n_j 是第 j 只债券的现金流次数, C_j^i 是第 j 只债券第 i 次的现金流, $r(t_i)$ 是由神经网络得到的利率期限结构函数; P_j 是实际价格。

由样本期内误差的均值和标准差评价多日的拟合和预测效果,均值和标准差计算公式分别如式(13)、(14)所示:

$$E = \sum_{i=1}^M \frac{1}{M} E_{MAD(i)} \tag{13}$$

$$\sigma = \sqrt{\frac{1}{M-1} \sum_{i=1}^M (E_{MAD(i)} - E)^2} \tag{14}$$

其中, M 表示样本的天数。

3 实证比较与分析

为了避开 2008 年到 2010 年期间 4 万亿计划造成的债券市场短期的非正常数据,同时从利率期限结构和时效性的角度考虑,本文选取了 2011 年 1 月

到 2016 年 7 月国债数据(数据来自 wind 数据库, <http://www.wind.com.cn>)。每月月末交易日为观察日,若为非工作日即向前顺延至工作日,共计 67 组数据,每组数据中 80% 用于拟合,20% 用于预测。

3.1 最优参数选取

在神经网络中,一些重要参数直接影响神经网络的预测效果,如 BPNN 和 WNN 的隐层结点数和迭代次数、RBFNN 的隐层结点数和平滑参数及 GRNN 的平滑参数。神经网络参数选择的标准和范围目前尚无通行的法则,本文依据预测结果的表现来确定相关参数的取值。以 2016 年 7 月 29 日的数据为样本,设定 BPNN 和 WNN 的隐层节点数的变化区间为[1,15],递增量为 1,网络的迭代次数的变化区间为[1000,10000],递增量为 1000。先针对 GRNN 中的平滑参数进行测试,变化区间为[0.0001,0.1],递增量为 0.0001;再根据 GRNN 最优值出现的区间,设定 RBFNN 中平滑参数的变化区间为[0.01,0.1],递增量为 0.01,隐层节点数的变化区间为[1,15],递增量为 1。由公式(12)计算出的 4 类网络的拟合与预测误差如图 3~6 所示。

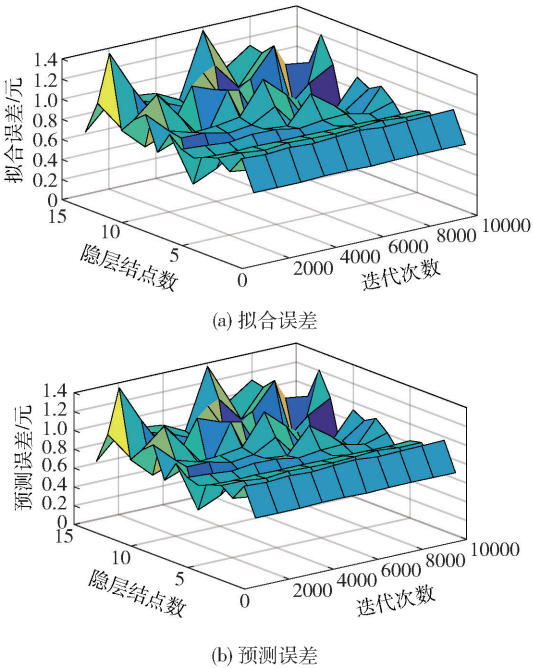


图 3 BPNN 隐层结点数和网络迭代次数对精度的影响
Fig. 3 Effect of hidden layer number and network iterations number on precision in BPNN

由图 3~5 看出,隐层结点数超过一定值时,网络的精度会下降,说明隐层结点数的增加并不会一直提高网络的精度。如图 3 中隐层结点数为 1、图 4

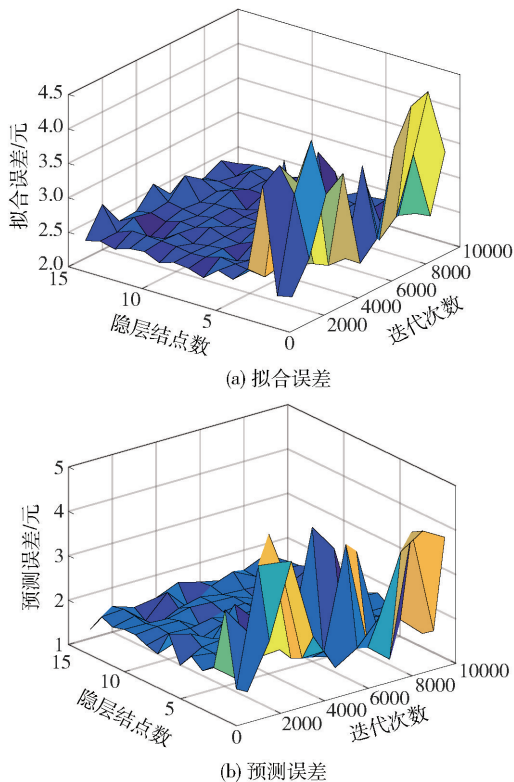


图4 WNN 隐层结点数和网络迭代次数对精度的影响

Fig.4 Effect of hidden layer number and network iterations number on precision in WNN

中为9、图5中为8时神经网络的精度达到最优。

由图3、4可以看出,隐层结点数过多时,随着迭代次数的增加,网络的预测效果时好时坏,具有一定的随机性,说明过多的隐层结点数导致过拟合出现时,迭代次数的增加将使误差进一步增大。所以选择迭代次数最小值1000为最优值。

由图5、6看出,随着平滑参数的增大,网络拟合和预测误差呈先降后升的趋势。当RBFNN的平滑参数为0.08,GRNN的平滑参数为0.0212时误差最小;平滑参数小于最优值时,网络的局部拟合能力较强,但对距离中心较远的数据无法响应,造成误差过大,此时平滑参数的增大有助于改善网络的精度;平滑参数超过最优值时,其局部拟合能力降低,整体误差变大。

由以上对图3~6分析结果,得出4类神经网络的最优相关参数如表1所示。

3.2 预测效果比较

根据所确定的最优参数对67个样本进行拟合,得出4类神经网络下的拟合及预测误差的时间序列变化趋势如图7所示,拟合及预测误差的均值和标准差如表2所示。

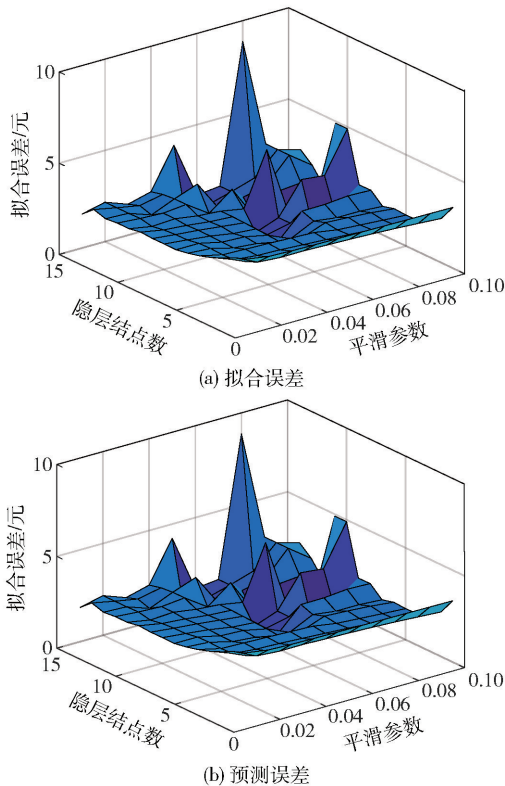


图5 RBFNN 隐层结点数和平滑参数对精度的影响

Fig.5 Effect of hidden layer number and smoothing parameters on precision in RBFNN

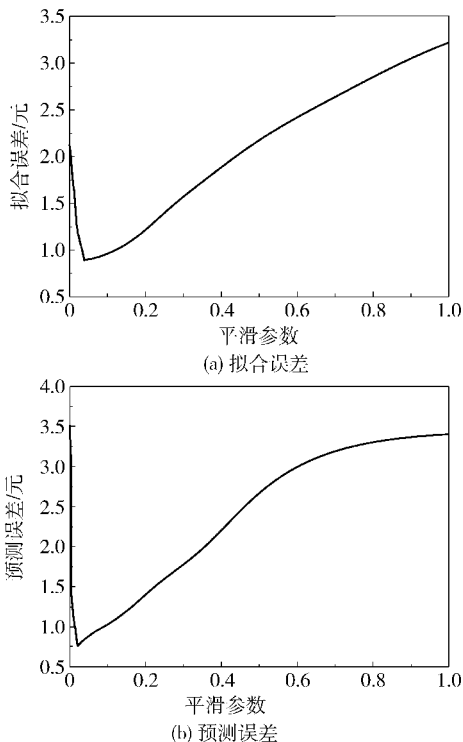


图6 GRNN 平滑参数对精度的影响

Fig.6 Effect of smoothing parameters on precision in GRNN

表 1 4 类神经网络相关最优参数

Table 1 Related parameters of four kinds of neural networks

模型	隐层结点数	迭代次数	平滑参数
BPNN	1	1000	—
WNN	9	1000	—
RBFNN	8	—	0.08
GRNN	—	—	0.0212

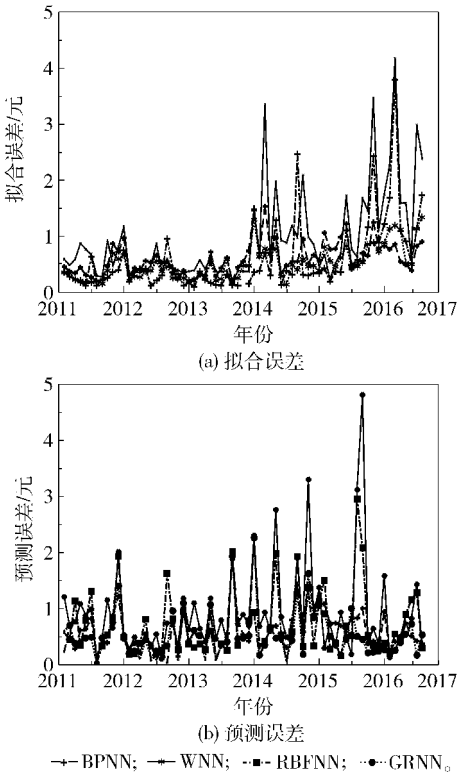


图 7 4 类神经网络误差比较

Fig. 7 Error comparison of four kinds of neural networks

表 2 4 类神经网络拟合及预测误差均值和标准差

Table 2 The mean and standard deviation of fitting and prediction error of four kinds of neural networks

模型	均值		标准差	
	拟合误差	预测误差	拟合误差	预测误差
BPNN	0.5497	0.6205	0.2431	0.4165
WNN	0.6962	0.9159	0.3152	0.8172
RBFNN	0.3688	0.6936	0.3126	0.5975
GRNN	0.4168	0.5376	0.1917	0.4261

由图 7 和表 2 可以看出:4 类神经网络中,RBFNN 虽然拟合效果较好,但是预测效果较差,且预测误差的波动性过高;GRNN 的预测效果最优,同时预测稳定性较好。说明 RBFNN 和 GRNN 更易出

现过拟合现象。这两种神经网络均为局部响应型神经网络,只能通过选取适当的参数来降低过拟合出现的可能性,但不能完全避免。在 4 类神经网络中,BPNN 的拟合误差和预测误差之间的波动最小,且预测结果也较为稳定;WNN 的拟合及预测效果最差,且误差的波动性较大。原因是一方面与 BPNN 相比,WNN 需要训练的参数更多,而在国债利率期限结构中,可获得的样本数量较少,使得网络训练不充分;另一方面,WNN 的优势在于对复杂信号的分解^[13],而国债利率期限结构中,数据结构相对简单,故 WNN 的优势难以体现。

4 结束语

本文在结合以往文献的基础上,筛选出适用于经济预测问题的 4 类神经网络,分别建立预测模型,分析了不同神经网络中参数选取对于网络拟合及预测精度的影响,确定了 4 类神经网络中相关参数的最优取值;并以 67 个月的月末数据为样本,对比了 4 类神经网络的拟合及预测误差。实证结果表明:广义回归神经网络预测效果较好,反向传播神经网络预测结果波动性较小。当网络参数选取恰当时,神经网络可以作为一种我国国债利率期限结构的预测技术。

参考文献:

[1] 周荣喜,邱莞华. 基于多项式样条函数的利率期限结构模型实证比较[J]. 系统工程, 2004, 22(6): 39-43.
Zhou R X, Qiu W H. Empirical comparison of term structure of interest rate based on polynomial spline functions [J]. Systems Engineering, 2004, 22(6): 39-43. (in Chinese)

[2] Diebold F X, Canlin L. Forecasting the term structure of government bond yields [J]. Journal of Econometrics, 2006, 130(2): 337-364.

[3] 赵晶,张洋,尹浩明. 中国国债利率期限结构的动态特征: 基于 NS 和 SV 模型的实证判别[J]. 管理世界, 2014(11): 164-165.
Zhao J, Zhang Y, Yin H M. Dynamic characteristics of interest rate term structure of China's national debt: evidence based on NS and SV [J]. Management World, 2014(11): 164-165. (in Chinese)

[4] 周荣喜,刘雯宇,牛伟宁. 基于 SV 利率期限结构模型的国债价差研究[J]. 北京化工大学学报(自然科学版), 2011, 38(3): 129-133.
Zhou R X, Liu W Y, Niu W N. A study of bond spreads

- based on the stochastic volatility model of the term structure of interest rates[J]. Journal of Beijing University of Chemical Technology: Natural Science, 2011, 38(3): 129-133. (in Chinese)
- [5] 沈根祥, 陈映洲. 双斜率因子动态 Nelson-Siegel 利率期限结构模型及其应用[J]. 中国管理科学, 2015, 23(10): 1-10.
- Shen G X, Chen Y Z. Nelson-Siegel dynamic term structure model with double slope factors and its applications [J]. Chinese Journal of Management Science, 2015, 23(10): 1-10. (in Chinese)
- [6] 马力. 基于神经网络的旅游收入预测模型建立研究——以山西为例[J]. 科学决策, 2010(9): 66-72.
- Ma L. The study of tourism demand forecasting model based on ANN—it takes Shanxi as example[J]. Scientific Decision Making, 2010(9): 66-72. (in Chinese)
- [7] Hejazi S A, Jackson K R. A neural network approach to efficient valuation of large portfolios of variable annuities [J]. Insurance Mathematics & Economics, 2016, 70: 169-181.
- [8] Jammazi R, Aloui C. Crude oil price forecasting: experimental evidence from wavelet decomposition and neural network modeling [J]. Energy Economics, 2012, 34(3): 828-841.
- [9] 熊志斌. ARIMA 融合神经网络的人民币汇率预测模型研究[J]. 数量经济技术经济研究, 2011(6): 64-76.
- Xiong Z B. Research on RMB exchange rate forecasting model based on combining ARIMA with neural networks [J]. The Journal of Quantitative & Technical Economics, 2011(6): 64-76. (in Chinese)
- [10] 姜德峰. 基于神经网络的利率期限结构研究[J]. 青海金融, 2007(5): 21-23.
- Jiang D F. Study on interest rate term structure based on neural network [J]. Qinghai Finance, 2007(5): 21-23. (in Chinese)
- [11] 周万隆, 胡雅丽. 神经网络在国债利率期限结构中的建模与实证[J]. 商业研究, 2003(4): 110-112.
- Zhou W L, Hu Y L. Modeling and demonstration of neural network in term structure of national debt interest rate [J]. Commercial Research, 2003(4): 110-112. (in Chinese)
- [12] 董皓舒. 基于智能神经网络组合预测的国债利率期限结构建模与实证[J]. 中国市场, 2010(44): 47-49.
- Dong H S. Modeling and demonstration of interest rate term structure of national debt based on combination forecasting of intelligent Neural Network[J]. China Market, 2010(44): 47-49. (in Chinese)
- [13] 郑承利, 黄冬冬. 小波神经网络对沪深 A300 的分析和预测[J]. 上海管理科学, 2011, 33(3): 34-37.
- Zheng C L, Huang D D. The analysis and prediction of a 300 index based on wavelet neural network[J]. Shanghai Management Science, 2011, 33(3): 34-37. (in Chinese)

Forecasting the term structure of interest rates with four kinds of neural networks

LIU Xiao¹ ZHOU RongXi^{2*} LI Jie¹

(1. School of Economics and Management, Beijing University of Chemical Technology, Beijing 100029;

2. School of Finance and Banking, University of International Business and Economics, Beijing 100029, China)

Abstract: The study sample selected in this paper is the monthly government bonds data from January 2011 to July 2016. We study the effect of related parameters which are selected for back propagation neural network (BPNN), wavelet neural network (WNN), radial basis function neural network (RBFNN) and generalized regression neural network (GRNN) on the prediction accuracy, and compare the prediction results of the four types of neural networks empirically. The results demonstrate that the GRNN performs best and the prediction results of the BPNN have smaller volatility, whilst the WNN and RBFNN prediction results are more volatile.

Key words: term structure of interest; prediction; neural networks; parameters optimization

(责任编辑:汪 琴)