

文章编号: 0427-7104(2015)03-0313-07

基于动态剪枝算法的神经网络预测模型

张 瀛^{1,2}

(1. 华东师范大学 金融与统计学院, 上海 200241; 2. 复旦大学 经济学院, 上海 200433)

摘 要: 提出一种针对径向基函数网络动态剪枝算法, 该方法根据统计贡献度动态确定核函数最优数量, 在递归估计参数的同时根据核函数贡献度的大小动态消除冗余节点, 以达到最佳网络结构. 利用中国月度信贷数据进行实证分析表明, 新提出的模型与 SARIMA 和 SVR 等其他基准模型相比, 具有更好的预测稳健性和准确性.

关键词: 统计贡献度; 自回归单整移动平均季节模型; 支持向量回归

中图分类号: F 224.9; F 201

文献标志码: A

DOI:10.15943/j.cnki.fdxbs-jns.2015.03.008

随着中国对外开放程度的加深, 对政策部门决策水平提出了更高要求, 正确决策的前提是对客观情况做出科学准确的预测. 信贷政策是我国宏观经济政策的重要组成部分, 是中国人民银行根据国家宏观调控和产业政策要求对金融机构信贷总量和投向实施引导、调控和监督, 促使信贷投向不断优化, 实现信贷资金优化配置并促进经济结构调整的重要手段, 在支持和促进国民经济又好又快发展中发挥日益重要的作用. 2009 年受国际次贷危机的影响, 中国中外资金融机构人民币各项贷款呈现爆发性增长, 货币信贷规模高速增长是维持实体经济平稳发展的客观需要, 但也为未来可能出现的资产价格泡沫、通货膨胀以及不良贷款等问题埋下隐患. 由于中国信贷政策自身的特性和传导机制, 其实施效果与预期目标往往存在较大差异, 信贷政策的有效性受到各种因素制约, 因此, 研究中外资金融机构人民币贷款的发展变化规律, 进而对未来变化做出准确的预测, 这对提高目前我国信贷政策的前瞻性和有效性具有重要的意义.

目前, 实际应用的预测方法不下几百种, 包括传统线性时间序列模型、传统的非线性模型(如门限自回归和马尔科夫链)、灰色预测方法、计量经济学模型以及一些如人工神经网络(Artificial Neural Networks, ANN)、支持向量机(Support Vector Machine, SVM)等新的智能技术. 与其他国家不同, 信贷渠道是我国货币政策传导的重要路径, 我国中外资金融机构人民币贷款数据的主要特点, 一是信贷数据为月度数据, 属于低频数据, 数据时间跨度短, 样本数量较少; 二是由于国际和国内经济形势的多变, 尤其是国际金融危机对我国信贷政策产生冲击, 信贷数据呈现突变特征, 而传统时间序列模型和计量经济学模型仅适用于渐变的、线性的时间序列, 上述特点使得传统预测模型无法取得令人满意的结果. 传统预测方法大多是对因果关系回归模型和时间序列模型进行分析, 但由于系统内部关系十分复杂, 加之外界环境的多变性, 人们在预测时存在着正确选择模型的困难, 使得这些需要对系统运行机理有彻底了解的预测方法的应用受到很大限制, 从而带来预测失误, 因此, 人们逐渐转向了只利用现存历史数据去构造模型进而推测未来的非参数预测方法——基于神经网络的预测方法.

信贷规模是中国中央银行密切关注的经济变量, 但目前对我国信贷预测的研究较少且预测精度较低, 不少文献^[1,2]指出中央银行的预测存在较大误差和系统性偏差, 为此, 本文采用单变量时间序列和非参数计量方法建立中国的信贷规模预测模型, 利用单变量时间序列进行预测需要的数据信息较少, 预测结果更为稳定和精确^[3,4]. 神经网络和支持向量机是近年来发展较快的非参数计量模型, ANN 克服了经济计量模型常犯的模型设定错误^[5], 但 ANN 模型的预测结果并不稳定, 有时优于 SVM 和 ARMA(Autoregressive Moving Average)模型, 但也有研究表明 ANN 预测精度劣于 SVM 等模型^[6]. 预测结果

收稿日期: 2013-03-15

基金项目: 国家自然科学基金面上项目(71173043); 上海市教育委员会科研创新项目(11ZS09)

作者简介: 张 瀛(1971—), 男, 副教授, E-mail: zywelcomeu@sina.com.

不稳定的主要原因在于难以确定最优的模型结构,目前已有实验设计^[7]和预测随机复杂性^[5]等多种方法,但现有方法需要从为数众多的模型中进行选择,计算量非常大.为此,本文提出一种基于统计贡献度(statistical contribution)的 SC-RBF 模型,该方法在估计网络参数的同时,根据统计贡献度的大小动态双向调整核函数数目,得到一个最佳网络结构,提高了网络设计的效率和预测性能.

1 模型算法

1.1 参数估计

图 1 是结构为 $n-h-m$ 的 3 层 RBF(Radical Basis Function)前馈网络, $X=(x_1, x_2, \dots, x_n)^T \in \mathbf{R}^n$ 为网络输入, $y=(y_1, y_2, \dots, y_m)^T$ 为网络输出,将 m 取为 1 表示单输出. 中间层神经元数量为 h , 中间层核函数取为高斯核,图 1 实质是一个非线性回归函数^[5], X 和 y 分别为非线性回归中的解释变量和因变量. 给定样本 (X_j, y_j) , $j=1, 2, \dots, N$, 则第 j 个样本 X_j 的输出 y_j 为

$$y_j = f(X_j) = \sum_{i=1}^h w_i \varphi_i(\|X_j - c_i\|), \quad (1)$$

其中: $\|X - c_i\|$ 为 $X - c_i$ 的范数; $c_i \in \mathbf{R}^n$ 为隐层核函数中心参数, σ_i 为第 i 个核函数的宽度; $w \in \mathbf{R}^{h \times m}$ 为输出权值矩阵; φ_i 为第 i 个核传递函数. 一个 3 层前馈网络的中间层节点数 h 如果足够大,能够逼近任意连续函数,例如在 L^2 函数空间,给定解释变量 X 和因变量 y , 设未知的条件均值 $M(X) = E(y|X)$, 则根据 L^2 函数空间的逼近性质可以证明,如果 $M(X) \in L^2$, 则对任意 $\epsilon > 0$, 存在 h , 使得 $E|M(X) - f(X)|^2 < \epsilon$ 成立,前馈网络使用 $O(nh)$ 个参数,收敛速度可达到 $O(1/h)$, 而多项式、样条和三角级数展开等传统非参方法,需要 $O(h^n)$ 个参数才能达到相同收敛速度^[5], 上述特点使得前馈网络成为预测能力很高的非参数计量模型,其中 RBF 网络由于具有全局逼近性等优点在经济金融预测中得到广泛应用. 采用计量经济学中的非线性最小二乘法估计未知参数,参数初始值随机确定,步长 η_1, η_2 和 η_3 均取为 0.001, 令网络输出与真实输出的误差平方和 $Error$ 最小,得到递归估计如下:

$$w_i(n+1) = w_i(n) - \eta_1 (\partial Error / \partial w_i(n)), \quad (2)$$

$$c_i(n+1) = c_i(n) - \eta_2 (\partial Error / \partial c_i(n)), \quad (3)$$

$$\sigma_i(n+1) = \sigma_i(n) - \eta_3 (\partial Error / \partial \sigma_i(n)). \quad (4)$$

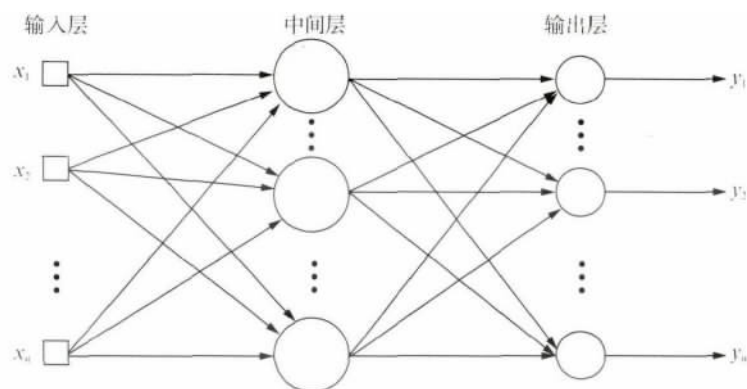


图 1 RBF 网络结构

Fig. 1 Structure of RBF

1.2 学习算法

Haykin(1998)^[8]指出具有最小结构的神经网络具有更好的预测能力,可由两种方式实现:一是网络增长法,即先设计一个小规模网络,然后根据网络性能自底向上逐步增加隐节点,如资源分配网络^[9];二是网络剪枝法,即先构造一个足够大的网络,然后自顶向下删除隐节点精简网络结构,如 Mozer 等^[10]提出的 Skeletonization 方法. 但上述方法均为单向设计,譬如资源分配网络^[9]中隐节点一旦添加就不能删除,

隐节点将越来越多,浪费系统资源(时间和空间)和减弱预测能力. Huang 等(2004)^[11]提出基于节点显著性的双向算法,可同时增加和删除隐节点,但该方法采用在线学习方法(sequential learning),学习结果受样本输入顺序的影响,并且在降低总体学习误差(即所有样本的误差平方和)方面不如批学习方法(batch learning),如果阈值有较小变化,在线学习方法会使隐节点数变化较多,严重影响网络的泛化能力^[12].

本文基于统计贡献度的概念,提出一种批学习的 RBF 网络隐节点动态调整的双向剪枝方法,该方法可以在估计未知参数的同时识别外移、萎缩和衰减等冗余核函数. 在参数递归估计中,某些核函数的中心参数 c 会太大、宽度 σ 太小或输出权重接近于 0,导致核函数输出为 0,分别称作核函数的外移、萎缩和衰减现象^[12],此时核函数对整个模型的贡献度很小,应当被删除. 如果核函数对整个网络的贡献度很大(大于阈值),就应该加进网络,因此,如何识别和去除多余的核函数是问题的关键.

(1) 修剪节点. 对于给定样本 $(X_j, y_j), j=1, 2, \dots, N$, 假设具有 h 个隐节点的 RBF 对输入向量 X_i 的单输出如式(1)所示,如果节点 q 被删除,剩下的 $h-1$ 个隐节点的输出为

$$y_q = \sum_{i=1}^{q-1} w_i \varphi(\|X - c_i\|) + \sum_{i=q+1}^h w_i \varphi(\|X - c_i\|), \quad (5)$$

对于输入 X_i , 删掉节点 q 产生的误差就是输出 y 和 y_q 之差的绝对值

$$\text{prunerr}(i, q) = |y - y_q| = |w_q| \varphi_q(\|X_i - c_q\|), \quad (6)$$

节点 q 的显著性定义为删除节点 q 的所有 N 个训练样本的平均误差:

$$\text{Sigprune}(q) = \frac{\sum_{i=1}^N \text{prunerr}(i, q)}{N} = \frac{|w_q|}{N} \sum_{i=1}^N \varphi_q(\|X_i - c_q\|), \quad (7)$$

(2) 增加节点. 增加新节点 $h+1$ 后 RBF 输出为

$$y_{h+1} = \sum_{i=1}^{h+1} w_i \varphi(\|X - c_i\|), \quad (8)$$

对于输入 X_i , 增加节点 $h+1$ 产生的输出误差为 $\text{growerr}(i, h+1) = |y - y_{h+1}| = |w_{h+1}| \varphi_{h+1}(\|X_i - c_{h+1}\|)$, 新增加的节点 $h+1$ 的显著性定义为增加节点 $h+1$ 的所有 N 个训练样本的平均误差:

$$\text{Siggrow}(h+1) = \frac{\sum_{i=1}^N \text{growerr}(i, h+1)}{N} = \frac{|w_{h+1}|}{N} \sum_{i=1}^N \varphi_{h+1}(\|X_i - c_{h+1}\|), \quad (9)$$

Siggrow 度量了新增核函数 $h+1$ 对整个网络输出的贡献,如果显著性大于预设阈值,该核函数应该加入到网络. 算法流程如下: ① 为 RBF 的中心、权重和宽度等参数随机赋初值,按照梯度法进行训练; ② 训练一定次数后,判断是否修剪和增加节点; ③ 计算每一个隐节点显著性 Sigprune ,如果节点显著性小于阈值,则删除该节点. 如果没有隐节点被删除,则计算新节点的显著性 Siggrow ,如果 Siggrow 大于阈值,则增加该节点. 判断是否达到给定的训练次数,条件满足则终止学习,否则跳转到②.

2 实证分析

2.1 数据

选择 2001 年 1 月至 2011 年 6 月中外金融机构人民币贷款月度数据作为研究对象,取自然对数后一阶差分得到序列 $\Delta \ln y_t$. 由于 $\Delta \ln y_t$ 在滞后 12 阶处自相关和偏相关显著不为 0,说明存在季节性. 再进行季节差分得到 2002 年 2 月至 2011 年 6 月共 113 个观察值序列 $\Delta \Delta_{12} \ln y_t$ (图 2(见第 316 页)). $\Delta \Delta_{12} \ln y_t$ 在 1% 水平下通过单位根检验,说明经过差分已实现数据的平稳化. 实际应用中,模型预测效果不佳的一个主要原因是没有对经济变量时间序列的性质进行识别. 不同性质的序列应该用不同的预测模型,如果序列以随机为主,则应当采用 ARMA 等随机线性模型预测,如果序列存在非线性机制,只有采用非线性模

型才能得到更加准确的结果. 一个给定的时间序列如果是确定性的,则需要判断是周期的、准周期的还是混沌的. 一般来说,严格的周期和准周期比较容易鉴别,但对于貌似无规则变化复杂的时间序列,要识别是混沌还是随机是一个十分重要的问题.

我们采用 BDS 检验(Brock-Dechert-Scheinkman statistic)验证序列 $\Delta\Delta_{12}\ln y_t$ 是否是一个非线性混沌序列. 首先利用 $AR(p)$ 模型对序列 $\Delta\Delta_{12}\ln y_t$ 进行拟合,得到序列中的线性成分,然后原序列减去线性成分得到残差,对残差进行 BDS 检验. 利用 AIC(Akaike Information Criterion)准则确定 $AR(p)$ 模型为 $\Delta\Delta_{12}\ln y=0.372\Delta\Delta_{12}\ln y(-1)+0.271\Delta\Delta_{12}\ln y(-3)-0.273\Delta\Delta_{12}\ln y(-4)+0.162\Delta\Delta_{12}\ln y(-5)-0.323\Delta\Delta_{12}\ln y(-12)$, 其中, $\Delta\Delta_{12}\ln y(-1)$, $\Delta\Delta_{12}\ln y(-3)$, $\Delta\Delta_{12}\ln y(-4)$, $\Delta\Delta_{12}\ln y(-5)$ 和 $\Delta\Delta_{12}\ln y(-12)$ 分别为贷款月度数据平稳序列 $\Delta\Delta_{12}\ln y_t$ 的 1 阶、3 阶、4 阶、5 阶和 12 阶滞后值. $\Delta\Delta_{12}\ln y_t$ 减去线性成分得到的残差如图 3 所示,对残差进行 BDS 检验,嵌入维数最大设为 8(表 1),在 5% 的显著水平下检验结果接受零假设,即残差为独立同分布,因此, $\Delta\Delta_{12}\ln y_t$ 是一个线性序列. Maia 等^[13]指出,神经网络预测非线性时间序列优于线性模型,但在预测线性时间序列方面往往不如 ARIMA 模型这类传统线性方法,下面将检验本文模型能否在预测线性序列 $\Delta\Delta_{12}\ln y_t$ 上优于线性模型.

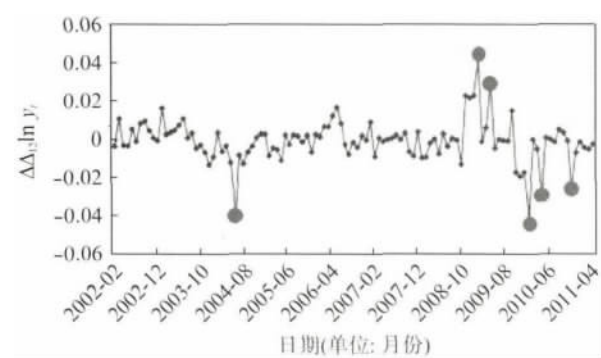


图 2 序列 $\Delta\Delta_{12}\ln y_t$ 的异常数据(异常数据以圆点标出)
Fig. 2 Outliers of time series $\Delta\Delta_{12}\ln y_t$ (Outlier marked in dots)

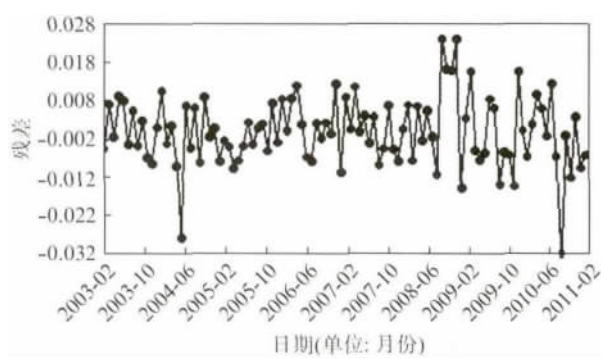


图 3 序列 $\Delta\Delta_{12}\ln y_t$ 的残差
Fig. 3 Residual of time series $\Delta\Delta_{12}\ln y_t$

表 1 残差的 BDS 检验
Tab. 1 BDS test for residual

嵌入维数	BDS 统计量	标准差	z-统计量	双侧概率	$C(m,n)$	$c(m,n)$	$C(1, n-(m-1))$	$c(1, n-(m-1))$	$c(1, n-(m-1))^k$
2	0.011 569	0.007 691	1.504 195	0.132 5	2 501.000	0.505 253	3 478.000	0.702 626	0.493 684
3	0.020 527	0.012 302	1.668 589	0.095 2	1 792.000	0.369 408	3 415.000	0.703 979	0.348 882
4	0.028 234	0.014 742	1.915 154	0.055 5	1 278.000	0.268 883	3 329.000	0.700 400	0.240 649
5	0.025 381	0.015 463	1.641 369	0.100 7	926.000 0	0.198 883	3 280.000	0.704 467	0.173 502
6	0.027 290	0.015 008	1.818 406	0.069 0	667.000 0	0.146 272	3 198.000	0.701 316	0.118 982
7	0.013 708	0.013 840	0.990 448	0.322 0	492.000 0	0.110 190	3 197.000	0.716 013	0.096 482
8	0.014 316	0.012 310	1.162 925	0.244 9	364.000 0	0.083 276	3 129.000	0.715 854	0.068 960

异常数据(outlier)对模型预测能力有重要影响,异常数据是指受到异常事件和干扰的影响,导致出现反常态势,会出现与时间序列大多数观测值不一样的数值^[14]. 解释变量中含有异常数据,会导致模型设定错误和有偏估计(biased estimates),因此,通常去掉或修正异常数据. 但研究表明,异常数据含有大量有用信息,包含异常数据会提高模型预测精度^[15],对异常数据进行修正并不能提高模型的预测精度^[16],因此,本文不对异常数据进行处理,比较存在异常数据情形下各种模型的预测精度. 本文采用切比雪夫不等式两阶段方法^[17]检验异常数据,得到的异常数据在图 2 中以圆点标出. 异常值包含中国央行的调控信息,例如 2009 年之后的信贷异常值,是由于 2008 年美国次贷金融危机对我国经济增长造成严重冲击,为此中国央行取消信贷配额限制,使信贷增量由 2008 年的 4.9 万亿元急剧升到 2009 年的 9.59 万亿元,超过市

市场预期。

2.2 SC-RBF 模型预测

采用扩展递归预测方法,训练集的起始点不变,随着新信息的到达,训练集样本数逐渐扩大,不断包含最新的信息,进行1步向前预测。把序列 $\Delta\Delta_{12}\ln y_t$ 中从2002年2月到2009年12月共95个观察值作为训练集,将2010年1月到2011年6月的18个数据作为预测集。预测前对序列 $\Delta\Delta_{12}\ln y_t$ 进行了归一化处理,得到的预测结果进行反归一化处理得到实际预测值。

采用交叉核实方法(cross validation)确定本文 SC-RBF 模型的阈值和网络结构,将训练集进一步分为训练集和测试集两部分,以预测2010年4月的信贷值为例划分方法如下:训练样本为2002年2月~2010年3月,选择该集中距离2010年4月最近的2009年3月~2010年3月的数据作为测试集 Test1,剩余数据作为训练集。选择2008年2月~2009年2月的数据作为测试集 Test2,剩余数据作为训练集。选择距离2010年4月较远的2005年6月~2006年6月的数据作为测试集 Test3,剩余数据作为训练集,并将同月份的数据作为测试集 Test4。评价指标采用测试误差和预测方向变化的正确率(Sign Correctness Percentage, SCP)。由于信贷数据具有12个月的季节周期性,所以将滞后期上限 N 设为15,测试集样本数设为13,可以预测12个方向变化值。

2.3 预测比较基准

本文以自回归单整移动平均季节线性模型 SARIMA(Seasonal Autoregressive Integrated Moving Average)和 ϵ -SVR(ϵ -Support Vector Regression)作为与本文模型比较的基准,采用 Box^[14] 提出的 SARIMA(p, d, q)(P, D, Q)^{*} 验证线性模型预测线性序列的精度优于神经网络的结论^[13]对本文模型是否成立。模型 ϵ -SVR 是 SVM 中的一种,主要用来估计和预测不同领域的时间序列, ϵ -SVR 求解下列二次规划问题:

$$\min_{\omega, b, \xi_i^-, \xi_i^+} \frac{1}{2} \omega' \omega + C \sum_{i=1}^n (\xi_i^- + \xi_i^+) \quad (10)$$

约束条件为:

$$\begin{aligned} y_i - (\omega' \phi(x_i) + b) &\leq \epsilon + \xi_i^- \quad \forall i, \\ (\omega' \phi(x_i) + b) - y_i &\leq \epsilon + \xi_i^+, \\ \xi_i^- &\geq 0, \xi_i^+ \geq 0, \end{aligned}$$

C 是惩罚参数, b 是常数项。上述问题的最优解为

$$f(x) = \sum_{i=1}^m (\lambda_i - \lambda_i^*) K(x_i, x) + b, \quad (11)$$

λ_i 和 λ_i^* 是拉格朗日乘子, $K(x_i, x_j) = \phi(x_i)' \phi(x_j)$ 是向量 x_i 和 x_j 的内积,内积函数的具体形式是一个核函数,本文选择了4种核函数(线性核、多项式核、高斯核和 Sigmoid 核)全面比较 SVR 的预测能力。

3 预测结果分析

首先采用预测误差绝对值平均值(Mean Absolute Error, MAE)、正规化的预测误差均方(Normalized Mean Squared Error, NMSE)和预测方向精度(SCP)等3个标准比较预测精度^[4],然后采用包含检验(encompassing test)对预测结果进行统计显著性检验。如果一种预测方法相对于另一种方法没有带来更多信息,则称前者被后一种方法包含,也即后一预测方法优于前一种方法,本文采用 ES(Ericsson Statistic)和 HLN(Harvey-Leybourne-Newbold statistic)^[18] 统计量检验。

表2(见第318页)显示 SC-RBF 模型预测误差均小于 SARIMA 方法和 SVR 方法,SC-RBF 预测变化方向的准确度 SCP 也高于其他方法。表3(见第318页)给出 ES 和 HLN 统计量检验结果,ES 统计量和 HLN 统计量的渐进分布均为标准正态分布 $N(0, 1)$,单边假设下 10%, 5% 和 1% 显著水平对应分位点分别为 1.281 6, 1.645 和 2.326 3,检验均显示 SC-RBF 优于 SVR 方法和 SARIMA 方法,例如 SC-RBF 对

SARIMA 的 ES 统计量为 0.642 7, 小于 5% 显著水平的分位点 1.645, 因此零假设被拒绝, SC-RBF 在 5% 显著水平下优于 SARIMA. 根据上述评价标准可以看到, 在预测包含异常数据的信贷数据时, SC-RBF 方法统计上显著优于 SARIMA 方法和 SVR 方法. SARIMA 模型对正常数据预测精度非常高, 如 2010 年 8 月和 11 月的预测值明显优于其他非参数方法, 说明 SARIMA 模型在预测线性时间序列方面具有优势, 这与 Maia 等^[13]的结论一致. 但本文时间序列中包含异常数据, 使得 SARIMA 模型的

预测精度下降, 例如 SARIMA 模型在预测 2010 年 5 月、2010 年 9 月和 2011 年 4 月的数据时产生较大误差, 预测精度下降的原因在于用于估计 SARIMA 的样本中含有异常数据, 导致 SARIMA 模型出现设定错误和有偏估计, 因此, 线性模型的建模稳健性较差, 相比之下, SC-RBF 方法和 SVR 方法稳健性更好, 并且 SC-RBF 方法优于 SVR 方法.

表 2 预测精度评估

Tab. 2 Evaluation of forecast accuracy

预测方法	检验标准					
	NMSE(%)	排序	MAE	排序	SCP(%)	排序
SARIMA	176.284 8	6	0.104 6	6	58.823	3
SC-RBF	6.268	1	0.021	1	82.353	1
线性核	47.544 2	5	0.051 8	5	76.470	2
多项式核	34.493 1	3	0.044 7	3	76.470	2
高斯核	29.621 4	2	0.041 4	2	76.470	2
Sigmoid 核	36.106 2	4	0.047 1	4	76.470	2

表 3 包含检验

Tab. 3 Encompassing test

预测方法	ES 统计量					HLN 统计量				
	SARIMA	线性核	多项式核	高斯核	Sigmoid 核	SARIMA	线性核	多项式核	高斯核	Sigmoid 核
SC-RBF	0.642 7	-1.887 5	-0.638 0	-1.542 1	-0.743 3	0.427 2	-0.993 2	-0.372 2	-0.860 7	-0.437 2
SARIMA		11.068 5	14.523 0	12.061 0	13.387 2		1.985 2	2.042 0	2.009 9	2.003 4
线性核			2.905 8	3.592 1	2.433 4			1.240 4	2.024 5	1.455 9
多项式核				1.954 9	0.273 3				1.199 4	0.395 3
高斯核					0.415 9					0.217 0

4 结 论

中国是一个处于社会转型和经济转轨的发展中国家, 经济行为即遵循基本的经济规律, 又受到中国独特的具有不可重复性事件的影响, 如制度因素、国际事件的冲击等, 影响经济系统的因素较多, 使得包括信贷规模在内的许多经济变量的变化更加复杂, 这种环境条件下要提高经济预测的准确性, 必须加强包括神经网络在内的非参数智能预测模型的研究和运用, 这对提升中国的经济预测能力和决策的前瞻性具有重要意义. 本文提出一种简单有效的 SC-RBF 预测模型, 该模型根据统计贡献度动态消除冗余节点, 确定最优的核函数数量和网络结构. 通过对中国信贷序列进行预测, 并与 SARIMA 和 SVM 等基准方法进行比较, 结果表明, 本文模型无论根据预测精度评估标准还是嵌套检验都优于其他方法, 而线性模型 SARIMA 在预测线性时间序列方面具有优势, 但对异常数据的预测精度下降明显, 且易受异常数据的干扰, 预测稳健性最差. 因此, SC-RBF 模型在实时预报方面具有显著的可行性, 对于政府决策部门预测信贷规模提高政策有效性和前瞻性具有很好的应用价值.

参考文献:

- [1] Clements M P, Fred J, Herman O S. An evaluation of the forecasts of the federal reserve: A pooled approach [J]. *Journal of Applied Econometrics*, 2007, **22**(1): 121-136.
- [2] Edge R M, Michael T K, Jean P L. A comparison of forecast performance between federal reserve staff forecasts, simple reduced-form models, and a DSGE model [J]. *Journal of Applied Econometrics*, 2010, **25**(2): 720-754.

- [3] Cassino V, Misich P, Barry J. Forecasting the demand of currency [J]. *Reserve Bank Bulletin*, 1997, **60**(1): 27-33.
- [4] Chatfield C. Time-series forecasting [M]. Boca, Raton, London, New York, Washington, D. C: Chapman and Hall/CRC Press, 2000.
- [5] Kuan C M, Liu T. Forecasting exchange rates using feedforward and recurrent neural networks [J]. *Journal of Applied Econometrics*, 1995, **10**(4): 347-364.
- [6] Kim K. Financial time series forecasting using support vector machines [J]. *Neurocomputing*, 2003, **55**(1-2): 307-319.
- [7] Balestrassi P P, Popova E, Paiva A P, et al. Design of experiments on neural network's training for nonlinear time series forecasting [J]. *Neurocomputing*, 2009, **72**(4): 1160-1178.
- [8] Haykin S. Neural networks: A comprehensive foundation [M]. Englewood Cliffs, NJ: Prentice Hall, 1998.
- [9] Platt J. A resource-allocating network for function interpolation [J]. *Neurocomputing*, 1991, **43**(3): 213-225.
- [10] Mozer M C, Smolensky P. Skeletonization: A technique for trimming the fat from a network via relevance assessment, in advances in neural inform [J]. *Processing Syst*, 1988(2): 107-115.
- [11] Huang G B, Saratchandran P, Sundararajan N. An efficient sequential learning algorithm for growing and pruning RBF networks [J]. *IEEE Transactions on Systems Man and Cybernetics*, 2004, **34**(6): 2284-2292.
- [12] 魏海坤. 神经网络结构设计的理论与方法 [M]. 北京: 国防工业出版社, 2005: 33-35.
- [13] Maia A L S, Carvalho F A, Ludermit T B. Forecasting models for interval-valued time series [J]. *Neurocomputing*, 2008, **71**(16-18): 3344-3352.
- [14] Box G J, Reinsel G. Time series analysis: Forecasting & control [M]. Englewood Cliffs, NJ: Prentice Hall, 1994.
- [15] Tolvi J. Outliers and predictability in monthly stock market index returns [J]. *The Finnish Journal of Business Economics*, 2002, **4**: 369-380.
- [16] Burman J P, Mark C O. Outliers in time series [R]. Bureau of the Census Statistical Research Division Report Series, SRD Research Report, 1988, CENSUS/SRD/RR-88114.
- [17] Amidan B G, Ferryman T A, Cooley S K. Data outlier detection using the Chebyshev theorem [C]//2005 IEEE Aerospace Conference Proceedings. New York: IEEE, 2005: 3814-3819.
- [18] Harvey D I, Stephen J L, Paul N. Tests for forecast encompassing [J]. *Journal of Business and Economic Statistics*, 1998, **16**(2): 254-259.

ANN Forecast Model Based on Dynamic Pruning Algorithm

ZHANG Ying^{1,2}

(1. Management School of Finance and Statistics, East China Normal University, Shanghai 200241, China;
2. School of Economics, Fudan University, Shanghai 200433, China)

Abstract: A new method for pruning RBF network is introduced. Networks are selected based on the Kernel function's statistical contribution to the overall performance of the network, and redundant nodes can be deleted dynamically according to the statistical contribution. The forecasting abilities of this method are compared with SARIMA and SVR approaches based on empirical RMB monthly China loan data series. The results show that the proposed method has the strongest forecast ability among all methods.

Keywords: statistical contribution; SARIMA; support vector regression