

基于滑动窗-RBF神经网络的现金流预测

刘洪久^{1,2,3}(博士) Robert Rieg² 马卫民³

(1.常熟理工学院管理学院 江苏常熟 215500 2.阿伦应用技术大学(Hochschule Aalen) 德国(Germany) 73430

3.同济大学经济与管理学院 上海 200092)

【摘要】 现金流预测是项目投资决策和评价企业未来价值的关键性因素。本文通过采用滑动窗技术确定RBF神经网络的训练样本和测试样本,然后通过变换不同的分布函数值对模型进行建模和仿真。实证研究结果表明,RBF神经网络模型训练和仿真结果稳定,预测效果良好。

【关键词】 滑动窗技术 RBF神经网络 现金流预测

现金流是企业绩效评价和投资项目价值评估的重要指标,合理的预测现金流量是决策者和投资者的重要工作。现金流预测的准确性依赖于取得详细的数据信息和采用合理的预测方法。理论上我们可以建立现金流与各影响因素之间的映射关系,但由于影响现金流量的内外因素众多并且十分复杂,有些因素难以量化,有的时候又难以收集到具体的数据,这使得多元回归分析方法的应用受到限制。而基于时间序列的方法由于建模简单、易于掌握,所以仍有很好的应用前景。

时间序列的传统方法主要有时间序列平滑预测法和曲线趋势预测法,如移动平均法、指数平滑法、可线性化的曲线趋势模型等。

近年来,以神经网络为代表的智能算法在预测方面的应用越来越广泛,智能算法的优点在于,即使我们不清楚模型输入和输出之间确切的函数关系,依然可以获得较好的预测效果。本文尝试使用小样本条件下的径向基(RBF)神经网络结合滑动窗技术探索预测现金流的新方法。

一、样本数据来源和滑动窗参数的确定

1. 样本数据来源。样本数据采用一汽轿车(股票代码:000800)从2002年12月31日到2011年3月15日34个现金流数据(每季度)。现金流量用 CF_t 表示($t=0,1,\dots,n$), n 表示时间序列的长度。

为了便于仿真,所有的数据都被归一化(0,1),如公式(1)所示:

$$CF'_t = \frac{CF_t - CF_{t \min}}{CF_{t \max} - CF_{t \min}} \quad (1)$$

其中: $CF_{t \max}$ 和 $CF_{t \min}$ 分别表示现金流时间序列的最大值和最小值。

2. 滑动窗参数确定。滑动窗口包括发送器和接收器,发送器和接收器的尺寸大小分别用 w_t 和 w_r 表示。为了研究在什

么样的尺寸下,模型具有良好的预测效果,需要确定合适的发送器的尺寸,为此我们设定 w_t 从1到8($w_t=1,2,\dots,8$)。接收器的大小为1($w_r=1$)(参见图1),因为模型的输出仅有一个结果。

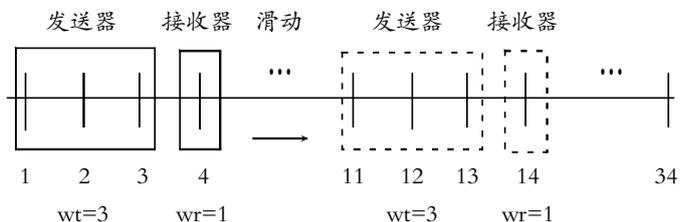


图1 滑动窗口的发送器和接收器($w_t=3$)

这样,当我们同时滑动发送器和接收器的时候,就会得到模型的输入矩阵 $P(i,j)$ ($i=1,2,\dots,w_t;j=1,2,\dots,n-w_t$)和输出矩阵 $T(j)$ ($j=1,2,\dots,n-w_t$),见公式(2)、(3):

$$P(i,j) = \begin{bmatrix} CF_1 & CF_2 & \dots & CF_{n-w_t} \\ CF_2 & CF_3 & \dots & CF_{n-w_t+1} \\ \dots & \dots & \dots & \dots \\ CF_{w_t} & CF_{w_t+1} & \dots & CF_{n-1} \end{bmatrix} \quad (2)$$

$$T(j) = (CF_{w_t+1}, CF_{w_t+2}, \dots, CF_n) \quad (3)$$

$P(i,j)$ 和 $T(j)$ 各自被划分为训练数据和测试数据两组,每组训练和测试数据都包括 $P(i,j)$ 和 $T(j)$ 。例如,数据组1由 $P(1,j)$ 和 $T(1)$ ($j=1,\dots,w_t$)。测试数据组中包含三对输入和输出向量用于检验模型的性能,训练数据组用来训练模型。

如果假定 t 时间的现金流量由 $t-1,t-2,\dots,t-w_t$ 时刻的现金流量决定,那么可以构建函数 $f: \mathbb{R}^{w_t} \rightarrow \mathbb{R}$ 的映射,从而得到 t 时刻的 CF 估计值:

$$CF(t) = f(CF(t-1), CF(t-2), \dots, CF(t-w_t)) \quad (4)$$

$$T(j) = f(P(i,j)) \quad (5)$$

$$(i=1,2,\dots,w_t;j=1,2,\dots,n-w_t)$$

二、RBF模型的建模和仿真

RBF神经网络嵌入一个两层前馈神经网络。网络由输入、输出和隐层构成，每层都使用径向基函数。RBF模型同样有强大的函数逼近能力。最常用的激活函数是高斯函数，如公式(6)所示：

$$\phi_j(X) = \exp[-(X-u_j)^T \Sigma_j^{-1}(X-u_j)] \quad (6)$$

对于 $j=1, \dots, L$ 来说， X 是输入特征向量， L 为隐层的数量， μ_j 和 Σ_j 是第 j 个高斯函数的均值和协方差矩阵。

模型的创建与运行均在Matlab神经网络工具箱中。RBF网络模型使用newrbf函数建立。在模型中，最重要的参数是分布函数SPREAD，它代表函数的平滑程度，其值越大，函数越平滑，但过大也会出现数值问题。为了保证模型的预测效果，SPREAD的取值范围为1~1 000。

将样本数据按公式(1)~(5)处理后代入模型，仿真结果显示当SPREAD在200~300之间时，测试数据的相对误差较小(见下表)。

RBF模型中不同wt条件下测试数据的相对误差

测试数据	wt=1	wt=2	wt=3	wt=4	wt=5	wt=6	wt=7
数据1	40.45	41.43	29.42	62.65	73.46	7.55	79.53
数据2	69.45	71.13	50.52	107.59	126.82	12.44	136.76
数据3	65.95	38.98	0.25	73.42	71.88	24.03	370.49
平均误差	58.61	50.51	26.73	81.22	90.72	14.67	195.59
均方差	0.025 1	0.023 8	0.021 9	0.018 8	0.009 2	0.005 2	0.004 3
SPREAD	200	200	200	200	300	300	300

由表1可以看出，滑动窗发送器的尺寸wt=6, SPREAD为300时模型的平均相对预测误差最小(14.67%)，预测效果最好；而滑动窗发送器的尺寸wt=7, SPREAD为300时模型的均方差最小(0.004 3)，训练效果最好。

滑动窗发送器的尺寸wt=6, SPREAD为300时的训练效果和预测效果见图2和图3。

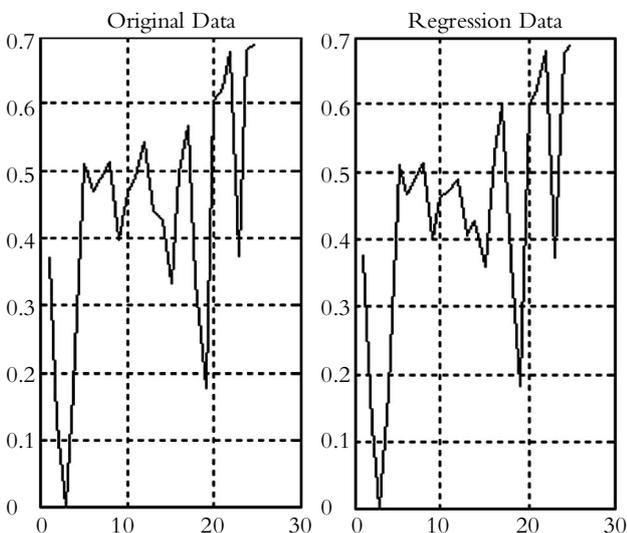


图2 RBF模型中wt=6的训练效果

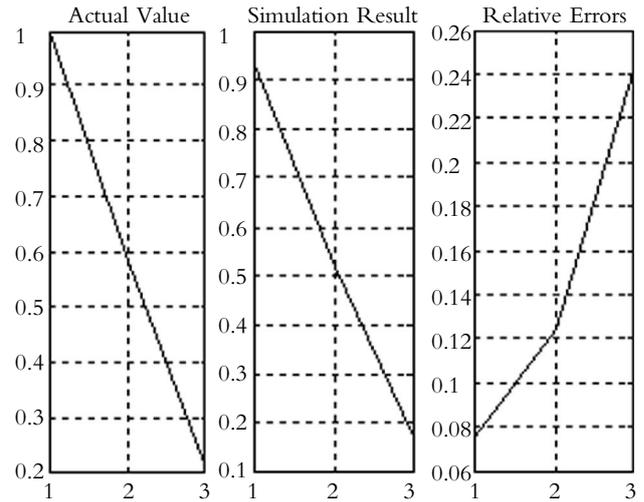


图3 RBF模型中wt=6的仿真效果

三、结论

本文通过对RBF神经网络的建模和仿真，发现该模型建模简单，训练效果良好；同时，与其他时间序列方法(指数平滑法、移动平均法、曲线趋势预测法和BP神经网络等)相比，该模型的预测功能较好。模型的训练效果和预测性能与滑动窗口的发送器尺寸wt和分布函数SPREAD的大小有关。实证研究表明，当滑动窗发送器尺寸wt等于6，分布函数SPREAD为300时的平均相对误差最小，预测效果最好；而滑动窗发送器的尺寸wt=7，权重函数为300时模型的均方差最小，训练效果最好。因此，应用RBF模型做预测时，要同时兼顾样本训练的结论和预测的效果。

【注】本文受国家自然科学基金(编号:71071113)、江苏省教育厅高校哲学社会科学研究项目(批准号:2010SJB790001)资助。

主要参考文献

1. 邹颖. 基于自由现金流假说和融资约束假说的投资行为研究. 财会月刊(理论), 2008; 11
2. Jafarizadeh B. Financial factor models for correlated inputs in the simulation of project cash flows. Information Sciences, 2010; 75
3. 何正文, 刘入境, 徐渝. 基于不同支付条件的现金流均衡项目调度优化. 管理科学学报, 2011; 8
4. 郁明, 丛爽. 基于径向基函数神经网络滑模控制的运动控制系统. 系统仿真学报, 2009; 3
5. Chu C. J. Time series segmentation: A sliding window approach. Information Sciences. Information Sciences, 1995; 85
6. Rojas I., Pomares H., Bernier J. L. Time series analysis using normalized PG-RBF network with regression weights, 2002; 42