

doi: 10.3969/j.issn.0490-6756.2018.06.010

一种金融时间序列区域分割方法的研究

桑夏夏, 李旭伟

(四川大学计算机学院, 成都 610065)

摘要: 金融时间序列指描述不同金融产品诸如股票、汇率与基金等的时间序列. 它与金融市场中人类的各种经济活动密切相关, 呈现出复杂多变的状态. 为了从海量的金融数据中发现有价值的, 可用于投资的信息, 大量学者采用数据挖掘来对金融时间序列作数据提取和处理. 由于金融时间序列具有高噪声、非平稳性、潜在的周期性等特性, 如果直接在金融时间序列的原始数据的基础上进行数据挖掘, 会导致结果失败或是取得不理想的挖掘效果. 而在数据挖掘前能对原始数据进行数据清洗、数据集成等预处理, 数据挖掘质量将达到更好地效果. 作为金融时间序列的一个重要分支, 股票时间序列预测方法通常采用分段线性表示 PLR (Piecewise Linear Representation) 进行时间序列的预处理. 但是 PLR 算法存在采用单一的拟合误差作为阈值, 分段效果不太理想, 算法本身的通用性, 时间复杂度等性能都有待提高等缺点. 本文提出了金融时间序列区域分割方法, 该方法在定性和定量上都优于传统的分段线性方法.

关键词: 金融时间序列; 分段线性; 区域分割; 算法复杂度

中图分类号: TP391 **文献标识码:** A **文章编号:** 0490-6756(2018)06-1189-08

Research on the region-based segmentation method of financial time series

SANG Xia-Xia, LI Xu-Wei

(College of Computer Science, Sichuan University, Chengdu 610065, China)

Abstract: In financial time series prediction and data mining research, Piecewise Linear Representation (PLR) is generally used for time series data preprocessing, but the PLR algorithm uses a single fitting error as the threshold, and the segmentation effect is not ideal, and the versatility, time complexity and other shortcomings of PLR need to be improved. This paper proposes a financial time series region segmentation method, which is superior to the traditional piecewise linear method in both qualitative and quantitative.

Keywords: Financial time series; Piecewise linear; Region segmentation; Computation complexity

1 引言

面对海量数据, 直接去操作一个高维的、有噪声的数据空间是很困难的^[1]. 我们可以将数据映射到一个高维特征空间^[2], 进行维度约简, 在高效、方便的表示形式上进行有效的挖掘.

数据降维有助于分类、可视化、通讯和高维数据的存储^[3]. 为提高相似匹配效率, 学者提倡使用近似方法, 可采用有损耗的数据压缩模式, 如分段线性表示 (Piecewise Linear Representation, PLR) 方法、序列离散化、字符串匹配方法等^[4]. 其中, 分段线性表示方法在有效地提取序列中的趋势和压

收稿日期: 2018-04-13

基金项目: 国家自然科学基金(2018GZ0182)

作者简介: 桑夏夏(1994—), 女, 四川成都人, 硕士生, 研究方向为计算金融. E-mail: 1134011493@qq.com

通讯作者: 李旭伟. E-mail: lixuwei@scu.edu.cn

缩原始数据的同时,能够随着时间序列长度的增长对序列进行自动划分,具有高效、实现方法简便、效果直观的优点^[5],对于不同领域的数据适应性良好.传统的 PLR 算法对原始数据分段均采用单一的拟合误差作为阈值,分段效果不太理想,算法本身的通用性、时间复杂度等性能都有待提高.因此,提出一种新的方法来提高分段效果和降低算法复杂度很是迫切和必要的.

本文提出了金融时间序列区域分割方法,将股票时间序列划分为三种状态:上升(1),下降(-1),波动(0).并对本文提出的区域分割方法同分段线性方法进行定性、定量分析,验证本文区域分割方法的有效性.其中,定性分析主要是从区域划分直观图像方面进行对比;定量方面,一方面,以区域方差为评价指标进行评价,另一方面,引用文献^[13]中 RCS-Trader 模型,以回报率为评价指标,对比文中提出的区域分割法和文献中的人工区域标记法及分段线性单点标记法.

2 时间序列区域分割方法与 RCS-Trader 模型

2.1 时间序列分段线性表示

Keogh 提出了一种时间序列分段线性表示方法,这种方法能够对时间序列进行有效的压缩,在一定程度上能反映时间序列的变化趋势,因而在时间序列数据挖掘中得到大量的应用.在 Keogh 的分段线性表示方法中,分段近似的目标是使原时间序列与其线性近似表示之间的残差平方和最小^[6],在这种目标函数的限制下,不能保证时间序列的每一分段内只具有单一的基本趋势,也就是只具有上升、下降中的一种,容易导致时间序列中某些点的基本趋势被错误提取.

此后,大量学者提出了很多新的时间序列分段线性表示方法,主要有基于特征点的分段线性表示方法、基于时态边缘算子的时间序列分段线性表示以及基于重要点的时间序列趋势特征提取的方法等.在这些研究方法中,对时间序列的处理都是保留原数据的一些重要数据点,没有考虑到时间序列具有局部特征(即局部上升,局部下降等区域).针对这些不足,本文提出区域分割方法,从区域上对时间序列进行分割.

2.2 时间序列区域分割方法

本节主要对股票时间序列区域分割进行介绍,如图 1 所示,主要包括股票数据的获取,股票数据

初始标签化,确定股票区域划分阈值,初始标签调节确定区域.

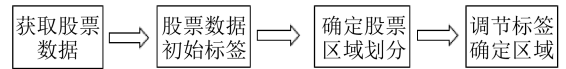


图 1 股票时间序列区域分割方法

Fig. 1 Stock data time series segmentation method

2.2.1 获取股票数据 本文实验使用的数据从免费、开源的 python 财经数据接口包 Tushare 获得,并截取其中的数据:时间、开盘价、最高价、最低价、收盘价、持仓量,以(date, open, high, low, close, volume)的格式保存到本地文件中.

2.2.2 股票数据初始标签 在传统的股指趋势预测中,移动平均线(Moving Average, MA)是一个重要的参考指标.而本文使用中心移动平均值(Centered Moving Average, CMA)代替移动平均值用于股票数据标签的标注.

MA 是一种传统的统计分析方法,它是由连续的移动平均值构成,在一段时间序列中,以一个时间节点为基准,按照移动平均周期依次计算每个交易周期收盘价的算数平均值,即可得到这段时间序列的移动平均值,将移动平均值用光滑的曲线相连即可得到移动平均线^[7],其计算方式如式(1)所示.

$$MA_t = \begin{cases} P_t^c, & t < m \\ \frac{1}{m} \sum_{t-m+1}^t P_t^c, & t \geq m \end{cases} \quad (1)$$

其中, MA_t 为时间节点 t 处的移动平均值; m 为移动平均周期; P_t^c 为 t 处的收盘点位.移动平均线反映的是一段时期的价格趋势.如图 2 所示,图中显示为农产品(股票代码:000061)的日 K 线和 5 日均线图像,从图中可以直观的看出股票的价格走势.

虽然在股票趋势预测中常采用均线作为参考指标,但移动平均线仅仅考虑了收盘价和历史时间的关系,没有区分不同时间段对当前价格的影响程度,一般在实验研究中,训练数据包含该时间点后面的数据.本文采用中心移动平均值来对这一现象做出改进.CMA 是统计学中用于中心化处理的一种方法,该方法广泛应用于动态数据的处理和采集上^[9].中心移动平均,是对移动平均值的改进,它以一个时间节点为中心,计算其左右两边相同时间跨度的平均值,再对第一次的平均值再次平均.也就是说,中心移动平均值是进行两次平均的结果^[8].对于时间节点 t 处的中心化移动平均值 CMA_t的计算方式如下所示.

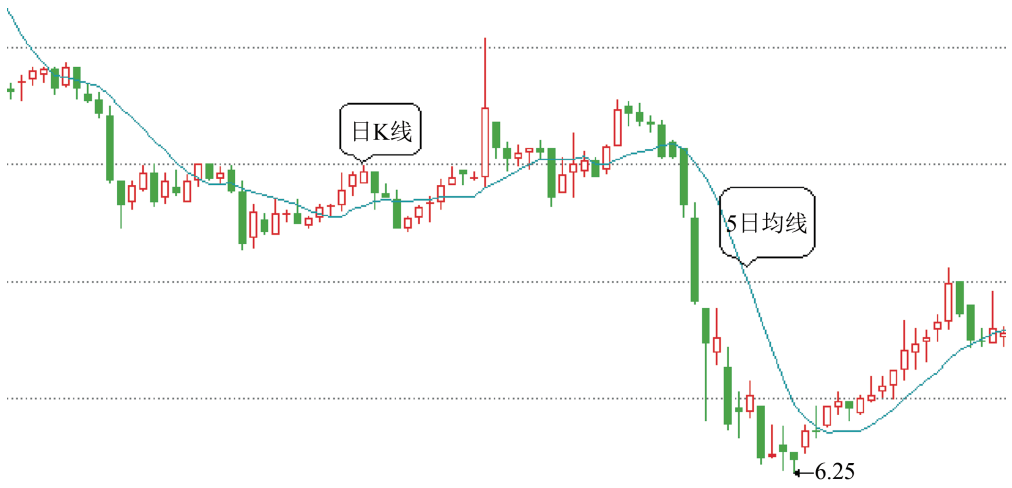


图 2 农产品(000061)日 K 线及 5 日均线图
Fig. 2 Agricultural products(000061) candlestick and MA5 chart

$$X_{t-0.5}^c = \frac{X_{t-2} + X_{t-1} + X_t + X_{t+1}}{4} \quad (2)$$

$$X_{t+0.5}^c = \frac{X_{t-1} + X_t + X_{t+1} + X_{t+2}}{4} \quad (3)$$

$$CMA_t = \frac{X_{t-0.5}^c + X_{t+0.5}^c}{2} = \frac{X_{t-2} + 2X_{t-1} + 2X_t + 2X_{t+1} + X_{t+2}}{8} \quad (4)$$

如式(2)~式(4)所示,其中 CMA_t 为时间节点 t 处的中心化移动平均值; X_t^c 为 t 处的收盘点位. 根据 CMA 计算公式,本文以一个时间节点 t 为中心,前后各取 2 个交易周期的数据,即一个 CMA 的计算周期为 $[t-2, t-1, t, t+1, t+2]$, 共计 5 个交易周期. 如图 3 所示,图中 CMA 与 K 线的走势也非常吻合,表明 CMA 可以用来标记股票时间序列的标签.

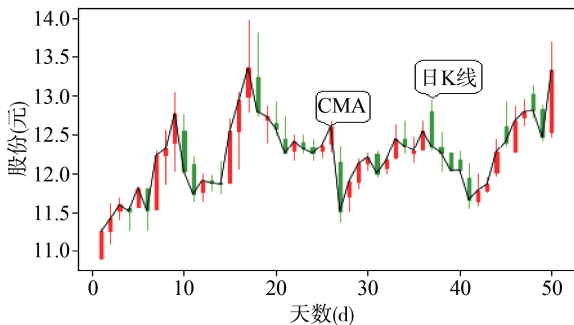


图 3 产品(000061)日 K 线及 CMA 图
Fig. 3 Agricultural products (000061) candlestick and CMA chart

本文采用基于 CMA 的趋势标注方法,文中初

始化类别标签选用 $y(t) \in \{+1; -1\}$ 分别标记价格趋势的上涨和下跌. 其中, $+1$ 表示为上升趋势, -1 表示为下降趋势. 具体判断依据如式(5)所示.

$$y(t) = \begin{cases} +1, & CMA(t) > \text{close}(t) \text{ and } CMA(t+3) > CMA(t+1) \\ -1, & CMA(t) < \text{close}(t) \text{ and } CMA(t+3) < CMA(t+1) \end{cases} \quad (5)$$

若不满足上述条件, t 时刻的标签与 $t-1$ 时刻的标签保持一致, 即 $y_t = y_{t-1}$; $y(0)$ 初始化定义为 $+1$.

2.2.3 确定股票区域划分阈值 通过上一节中确定的股票数据标签,通过式(6)的加权平均的方法来确定初始判定区域为上升集和下降集的阈值. 以上升阈值为例, $x_1 \cdots x_k$ 表示连续上升(即原始标签连续为 $+1$ 的个数), $f_1 \cdots f_k$ 中 f_i 表示对应连续 i 个 $+1$ 标签的天数.

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \cdots + x_k f_k}{k} \quad (6)$$

2.2.4 调节标签确定区域 在股票数据中,上升趋势中可能夹杂细微的震荡,下降趋势中也可能存在小幅反弹,但这些波动并不会影响股票的趋势. 如图 4 所示,区域 1 和区域 3 为依据阈值确定的上升和下降区域(即连续 $+1$ 大于上升阈值),但区域 2 和区域 4 存在明显的上升和下降趋势,需要调整使得区域 1 扩展为区域 2,区域 3 扩展到区域 4.

本文的区域调整方法如下,在以连续 n (划分阈值)个相同标签为一个趋势标准,若存在小幅波

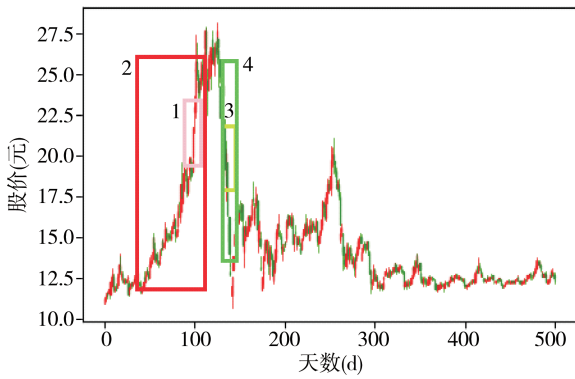


图 4 股票时间序列区域调整
Fig. 4 Regional adjustment of stock time series

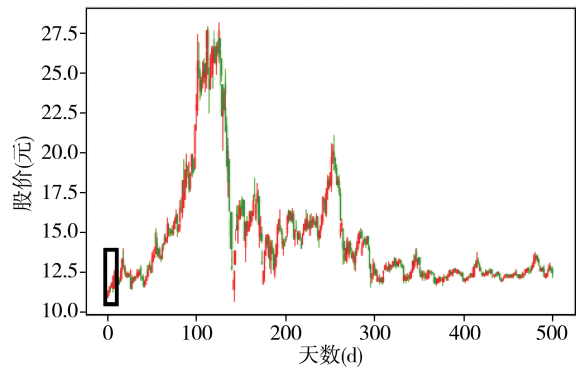


图 5 股票时间序列幅度变化
Fig. 5 Stock time series amplitude changes

动的标签个数小于趋势周期 n 的 $1/3$ 时,则将这些小幅波动调整为与该趋势相同的标签;反之,若波动幅度大于等于趋势周期 n 的 $1/3$,则把该段数据归为波动数据,划分到波动区域.但初始标签只是表明 $+1$ 标签为某一时段的上升趋势, -1 为某一时段的下降趋势,没有考虑上升、下降的幅度.即可能存在如图 5 所示的情况,黑色区域 $+1$ 连续个数达到上升阈值,因此应该划分为上升区域,但是这段区域上升幅度很小应该划分为波动区域.因此,在调整股票区域时,需要加上对区域起始点变化幅度的判断,本文采用区域终止点股票价格相对起始点的 $1/3$ 为判断标准.

2.3 RCS-Trader 模型

文献[13]中作者提出以区域标记法和代价敏感支持向量机为基础的 RCS-Trader 股票交易决策预测模型,文献中在预处理阶段采用了区域标记的方法,但其区域标记方法为人工经验标注.存在操作复杂,且区域划分结果存在主观因素,标记不准确等问题.

文献[13]中作者对 RCS-Trader 模型的构建如图 6 所示.文中主要对图 6 中第三个步骤,样本区域标记步骤进行改进和对比.下面介绍模型剩余部分.

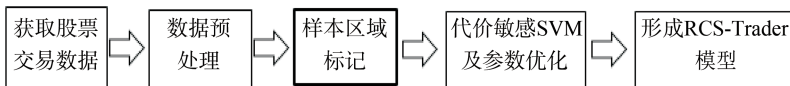


图 6 RCS-Trader 模型
Fig. 6 RCS-Trader model

2.3.1 数据预处理 RCS-Trader 采用 17 个特征作为输入,特征可由本地 csv 文件中原始数据 (date, open, close, high, low, volume) 计算得到.再对 17 个训练集特征进行归一化处理和 PCA 降维处理,获得一个 6 维数据.

2.3.2 代价敏感 SVM 及参数优化 代价敏感 SVM(支持向量机)是在 SVM 基础上引入松弛变量而构成的.

SVM(支持向量机)是在训练集上先训练一个分割超平面^[10],然后该平面就是分类的决策边界,目的是找到到分割超平面最近的点,确保他们离分割面的距离尽可能远,以达到最好分类的效果.

由于样本点存在类别不平衡问题,为了增强支持向量机的容错性,在支持向量机的基础上引入松弛变量(即允许一些样本点到超平面的距离不满足支持向量机的条件),模型中将损失作为惩罚因子

C 加入到目标函数中^[11],最终,代价敏感支持向量机优化问题可表示为式(7).

$$\begin{cases} \min_{w, b, \epsilon_i} \frac{1}{2} \|w\|^2 + C \sum \epsilon_i \\ \text{s. t } y_i (w^T x_i + b) \geq 1 - \epsilon_i \\ \epsilon_i \geq 0, i = 1, 2, \dots, m \end{cases} \quad (7)$$

式(7)表示对给定训练样本集 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, $y_i \in \{-1, 1\}$ 划分超平面,其中, w 是超平面的法向量; b 是偏差值; C 是惩罚因子(表示分类器对离群点的重视程度).

模型使用 F_s 最大化作为模型的优化目标,如式(8)所示.

$$F_s = (2 \times F_{\beta 1} F_{\beta 2}) / (F_{\beta 1} + F_{\beta 2}) \quad (8)$$

其中, $F_{\beta 1}$ 和 $F_{\beta 2}$ 分别代表买入类和卖出类的 F_{β} , F_{β}

为准确率(precision)和召回率(recall)的加权调和平均值,计算公式如式(9)所示.

$$F_s = ((1 + \beta^2) \times precision \times recall) / (\beta^2 \times precision + recall) \tag{9}$$

$F_{\beta 1}$ 和 $F_{\beta 2}$ 采用相同的权重, β 默认使用模型中的默认值 0.3.

3 实验与分析

3.1 实验方案

实验方案包括数据选取,实验参数选取和区域调整策略.

3.2 实验数据选取

为验证区域划分模型的有效性,本文随机选用了万科 A(股票代码:000002)、农产品(股票代码:000061)进行区域划分,选择数据的时间跨度从 2014 年 12 月 10 日到 2017 年 12 月 8 日.

3.3 实验参数的选择

由于实验中用到了区域划分阈值,实验应当先确定区域阈值,图 7 为上升(+1)标签和下降(-1)标签的分布图.图中标签分布呈正偏态分布,以上升标签为例,文中分别采用众数 $Ru_1=4$,加权平均值 $Ru_2=7$,随机数 $Ru_3=10$ 作为区域划分阈值,通过对比,以加权平均值为阈值,区域划分效果最好.所以,阈值以股票时间序列的加权平均值为准.

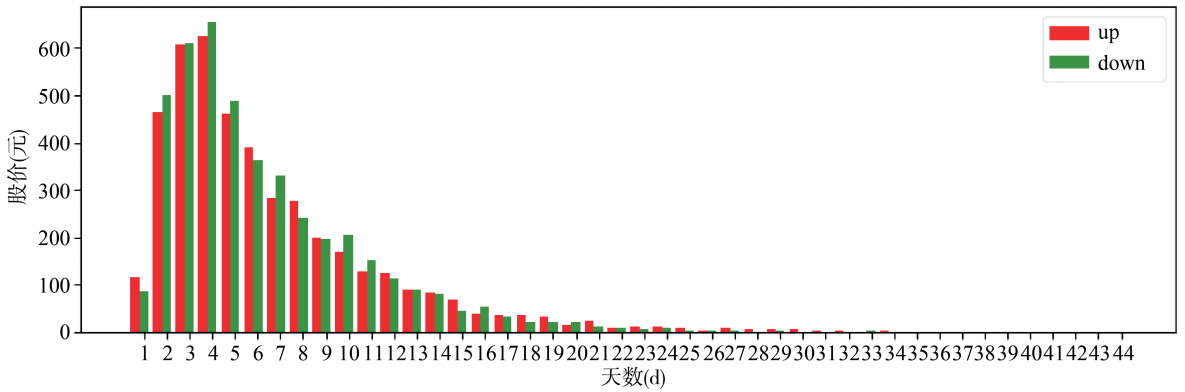


图 7 上升、下降标签分布图
Fig. 7 Ascending and descending label distribution map

3.4 区域分割

通过本文提出的区域迭代调整方法,确定原始股票的上升、下降和平稳区域,并在股票原数据的 K 线图上进行绘制,从而进行直观比较.如图 8 为万科 A(股票代码:000002)的区域表示.图 9 为农产品(股票代码:000061)的区域表示.

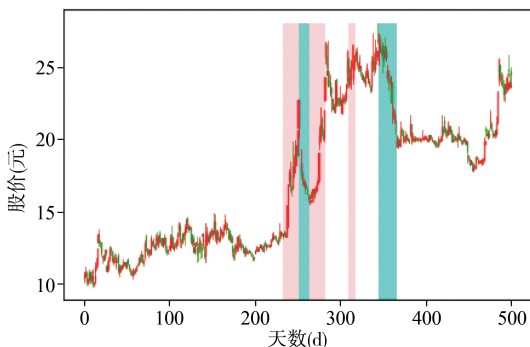


图 8 万科 A(股票代码:000002)的区域表示
Fig. 8 The area of vanke A (stock code:000002)

图中横轴表示时间节点,纵轴表示股票价格,红色、绿色曲线为股票原始序列的 K 线图,红色区

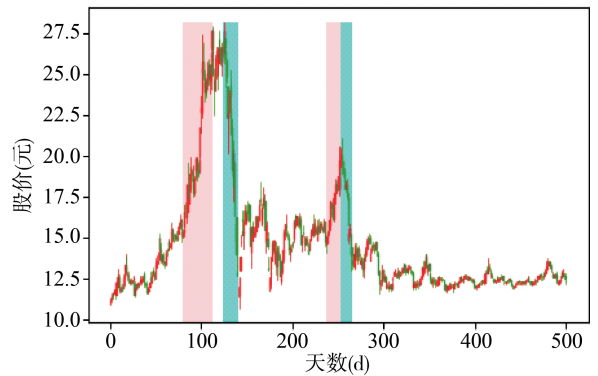


图 9 农产品(股票代码:000061)的区域表示
Fig. 9 Agricultural products (stock code:000061) regional representation

域为上升、蓝色区域为下降区域,未标注颜色区域为波动区.

3.5 RCS-Trader 模型人工区域标记法同区域分割法对比

本节通过对比 RCS-Trader 模型人工区域标记和本文提出的区域分割法,来显示实验效果,证明本文区域分割方法的可行性.

文献[13]中对比人工区域标记法和分段线性单点标记,因此,此部分实验不仅同 RCS-Trader 模型的人工区域标记进行了对比,也同分段线性单点标记进行了对比。

实验中除了样本标记方法不同,其他条件均保持一致。实验仍采用文献[13]中作者采用的交易规

$$Rate = \left\{ \sum_{i=1}^k \frac{[(1 - tax - charge) \times sell_i - (1 + tax) \times buy_i]}{[(1 + tax) \times buy_i]} \right\} \times 100\%,$$

$$i = 1, 2, 3, \dots, k \tag{10}$$

其中, tax 是税率; $charge$ 是股票交易中的手续费; k 是交易次数; $sell_i$ 是股票在第 i 次交易中卖出价格; buy_i 是股票在第 i 次交易的买入价格。

本节选用万科 A(股票代码:000002)、农产品(股票代码:000061)作为实验数据。

3.5.1 万科 A 区域分割法、人工区域标记法、分段线性单点标记法的预测比较 在本次实验中,获取了万科 A(股票代码:000002)2014 年 12 月 10 日至 2017 年 12 月 8 日数据作为历史交易数据,使用 2015 年 12 月 8 日至 2017 年 12 月 8 日数据作为测试集数据。

实验通过参数寻优得到代价敏感支持向量机的最佳参数是选择 rbf 核函数, $C=0.84$, $gamma=1.02$ 。通过对测试集进行预测,实验投资回报率(Rate)如表 1 所示。

表 1 万科 A(股票代码:000002)区域分割法、人工区域标记法、分段线性单点标记法预测对比

Tab.1 Vanke A (stock code:000002) regional segmentation method, artificial regional labeling method, PLR for prediction and comparison

	区域分割法	人工区域标记法	分段线性单点标记法
买入类 预测结果	precision	0.183	0.3333
	recall	0.3333	0.3333
	$F_{\beta 1}$	0.1902	0.3333
卖出类 预测结果	precision	0.3966	0.2857
	recall	0.511	0.6667
	$F_{\beta 2}$	0.404	0.2999
F_s	0.3267	0.2984	0.3056
回报率	33.87%	32.54%	29.66%

3.5.2 农产品区域分割法、人工区域标记法、分段线性单点标记法预测比较 本次实验选取农产品(股票代码:000061)2014 年 12 月 10 日至 2017 年

12 月 8 日数据作为历史交易数据,2015 年 12 月 8 日至 2017 年 12 月 8 日数据作为测试集数据。

实验通过参数寻优得到代价敏感支持向量机的最佳参数是选择 rbf 核函数, $C=0.91$, $gamma=1.02$ 。通过对测试集进行预测,实验投资回报率(Rate)如表 2 所示。

12 月 8 日数据作为历史交易数据,2015 年 12 月 8 日至 2017 年 12 月 8 日数据作为测试集数据。

实验通过参数寻优得到代价敏感支持向量机的最佳参数是选择 rbf 核函数, $C=0.91$, $gamma=1.02$ 。通过对测试集进行预测,实验投资回报率(Rate)如表 2 所示。

表 2 农产品(股票代码:000061)区域分割法、人工区域标记法、分段线性单点标记法预测对比

Tab.2 Agricultural products (stock code:000061) regional segmentation method, artificial regional labeling method, PLR for prediction and comparison

	区域分割法	人工区域标记法	分段线性单点标记法
买入类 预测结果	precision	0.183	0.3333
	recall	0.3333	0.3333
	$F_{\beta 1}$	0.1902	0.3333
卖出类 预测结果	precision	0.3966	0.2857
	recall	0.511	0.6667
	$F_{\beta 2}$	0.404	0.2999
F_s	0.3308	0.3096	0.2537
回报率	18.03%	15.47%	12.96%

3.6 实验结果分析和评估

实验结果从定性和定量两个方面对比文中提出的区域分割方法和分段线性方法。

定性方面,如图所示,图 10 为分段线性表示方法拟合的股票数据图,图 11 为区域分割方法标记的股票时间序列趋势图。对比图 10 和图 11,图 10 中,绿色曲线为原股票时间序列收盘价的变化曲线,红色折线为分段线性表示的拟合曲线,图中,蓝色区域相对整个时间序列价格波动较小,红色区域有明显上升趋势,而分段线性方法将蓝色区域和红色区域划分为一个区域,即只考虑到了时间序列趋势变化转折点中的极值点,而没有考虑到非极值点

的变化,不但序列的拟合误差较大,而且对于序列压缩程度的调整也不够灵活^[5].图 11 中,绿色曲线为股票时间序列的收盘价变化曲线,粉色区域为区域分割方法确定的上升区域,绿色区域为区域分割方法确定的下降区域,未标注区域为波动区域,图 11 可以很好的拟合出上升和下降区域.

定量方面,分别以 3.4 节中图 8(万科 A)和图 9(农产品)进行区域分割的结果为例,计算两支股票不

同区域的方差,结果如表 3 所示(表中 1,2.. 分别代表对应图 8 和图 9 中第 1,2 个上升、下降或波动集),表中俩支股票上升集和下降集方差较大,波动集方差较小;另一方面,对 3.5.1 节和 3.5.2 节中实验结果进行分析,表 1 和表 2 中区域分割法的回报率均大于人工区域标记法和分段线性单点标记法,表明本文提出的区域分割法在预测模型中的可行性.

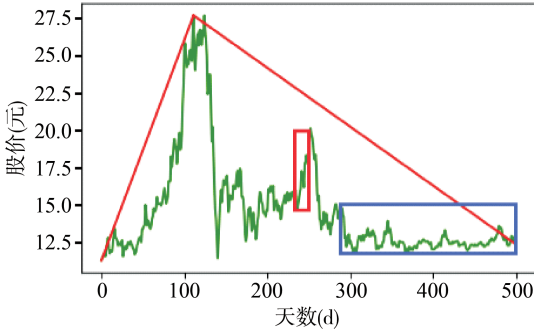


图 10 分段线性表示

Fig. 10 Piecewise Linear Representation

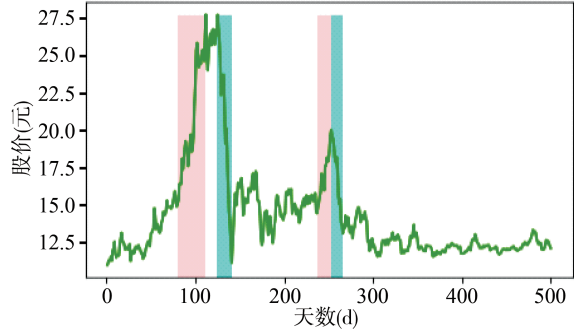


图 11 区域分割表示

Fig. 11 Regional segmentation

表 3 万科 A(股票代码:000002)和农产品(股票代码:000061)区域表示方差对比

Tab. 3 The variance of the area of Vanke A (stock code:000002) and agricultural products(stock code:000061)

股票名称	上升集		下降集		波动集		
	1	2	1	2	1	2	3
万科 A	6.29	9.11	1.43	4.89	0.89	0.96	0.85
农产品	12.44	2.05	24.13	4.64	0.62	0.55	0.87

4 结 论

大数据时代,数据信息尤为重要,股票市场系列一般是动态的、非参数的、混沌的^[12],股票领域中要对股票交易进行评估,离不开对股票数据的分析处理.面对海量数据,直接操作一个高维的数据空间是很困难的.因此,需要研究合适的数据表示形式,进行维度约简,在高效、方便的表示形式上进行有效的挖掘.目前常用的 PLR 算法对原始数据分段均采用单一的拟合误差作为阈值,分段效果不太理想,算法本身的通用性、时间复杂度等性能都有待提高.本文提出区域划分的方法,在区域阈值的基础上进行调整确定新的上升、下降区域,避免了参数的单一性,对算法复杂度也有所提高.实验结果证明本文的有效性.

参考文献:

[1] Cunningham J P, Ghahramani Z. Linear dimension-

ality reduction: survey, insights, and generalizations [J]. J Mach Learn Res, 2015, 2859: 2900.

[2] Yang Y J, Yang Y, Li J P. Research on financial time series forecasting based on SVM [J]. IEEE, 2016, 5090: 6126.

[3] Waqar M, Dawood H, Shahnawaz M B, et al. Prediction of stock market by principal component analysis [J]. IEEE, 2017, 7695: 6341.

[4] 郭永东. 移动平均法在数据采集和分析中的应用 [J]. 电子科技, 2014, 27: 118.

[5] Chang P C, Wu J L, Lin J J. A takagi-sugeno fuzzy model combined with a support vector regression for stock trading forecasting [J]. Appl Soft Comput, 2016, 38: 831.

[6] Hu Y P, Jiang Z F, Li X Q. A novel multi-resolution representation for streaming time series [J]. Pro Comput, 2018, 178: 184.

[7] 向为民, 牛玉龙, 王霜. 公司基本面、市场行为与股票投资收益 [J]. 重庆理工大学学报: 社会科学, 2017, 67: 75.

- [8] 王威娜, 王严冬. 基于自动聚类 and 自回归模型的时间序列预测[J]. 吉林化工学院学报, 2017, 86: 89.
- [9] 吕涛, 郝泳涛. 基于 K 线序列相似性搜索的股票价格预测[J]. 计算机应用, 2017, 229: 235.
- [10] Li Q, Chen Y, Jiang L L. A tensor-based information framework for predicting the stock market[J]. ACM Trans Inform Syst, 2016, 34: 11.
- [11] Luo L, You S, Xu Y, *et al.* Improving the integration of piece wise linear representation and weighted support vector machine for stock trading signal prediction[J]. Appl Soft Comput, 2017, 56: 199.
- [12] Khaidem L, Saha S, Dey S R. Prediction the direction of stock market prices using random forest[J]. App Math Fin, 2016, 1: 20.
- [13] 秦璐, 李旭伟. 基于区域标记法的代价敏感支持向量机在股票预测中的研究[J]. 四川大学学报: 自然科学版, 2018, 57: 277.

引用本文格式:

- 中文: 桑夏夏, 李旭伟. 一种金融时间序列区域分割方法的研究[J]. 四川大学学报: 自然科学版, 2018, 55: 1189.
- 英文: Sang X X, Li X W. Research on the region-based segmentation method of financial time series [J]. J Sichuan Univ: Nat Sci Ed, 2018, 55: 1189.