

学校编码: 10384

分类号\_\_\_\_\_密级\_\_\_\_\_

学 号: 23220121153052

UDC\_\_\_\_\_

厦 门 大 学

硕 士 学 位 论 文

基于分段线性表示和支持向量机  
拐点预测的改进研究

Improvements for Inflection Point Prediction based on Piecewise  
Linear Representation and Weighted Support Vector Machine

徐 燕 茹

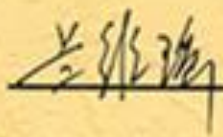
指导教师姓名: 罗 林 开 教 授

专 业 名 称: 模式识别与智能系统

论文提交日期: 2015 年 5 月

论文答辩时间: 2015 年 5 月

学位授予日期: 2015 年 月

答辩委员会主席: 

评 阅 人: \_\_\_\_\_

2015 年 月

厦门大学博硕士学位论文摘要库

## 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（  
（组）的研究成果，获得（  
实验室的资助，在（  
内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）： 徐燕茹

2015年 5月 20日

厦门大学博硕士学位论文摘要库

## 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

( ) 1. 经厦门大学保密委员会审查核定的保密学位论文，  
于 年 月 日解密，解密后适用上述授权。

(  ) 2. 不保密，适用上述授权。

(请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。)

声明人(签名): 徐燕茹

2015年5月20日

厦门大学博硕士学位论文摘要库

## 摘要

股市预测，作为经济预测的一个分支，一直是学术界的研究热点。近年来，一种结合分段线性表示（PLR）和加权支持向量机（WSVM）的方法（PLR-WSVM）被提出来预测证券的拐点，并取得了有益的成果。然而，该方法也存在一些不足的地方，例如，交易次数较少，投资周期长；采用 PLR 划分普通点和拐点的时候，它的参数设定由专家经验来决定，且对于不同趋势的股价预测都采用同一个参数。为了克服这些缺点，本文对 PLR-WSVM 预测股价拐点的方法进行了改进，主要表现为：

（1）将原带权重的四分类问题简化为带权重的两分类问题，既避免了样本类间权重系数的选取，又提高了模型的分类准确率和交易准确率。

（2）根据拐点当天股价与其前五天均价的比较判断给出交易信号，改善了完全由机器学习预测拐点作为交易信号的模式。

（3）提出了在拐点滞后一天进行实际交易操作的策略，进一步提高交易准确率，并提高了模型收益性能。

（4）提出一种 PLR 阈值参数选取的算法，根据股价的波动自适应选择合适的阈值，改进了原来模型依据经验统一设置的不足。

（5）设计了两种交易策略应用于模型，提高了模型的收益性能。

通过对来源于上海证券交易所 13 支股票的实证分析，在与 PLR-WSVM、买入并持有（BHS）策略方法的对比中，本文的改进方法 IPLR-WSVM 具有最好的稳定性和盈利能力，证实了 IPLR-WSVM 用于证券拐点预测的有效性与实用性。

**关键词：**拐点预测；分段线性表示；支持向量机

厦门大学博硕士学位论文摘要库



## Abstract

Stock market prediction, as a branch of economic forecasting, has been the research focus in academic, and the domestic and foreign researchers have established many outstanding models to forecast the trend of stock. Piecewise linear representation (PLR) and weighted support vector machine (WSVM) have been integrated, which has shown a great success for inflection point prediction recently. However, it provides few transactions for investors, which takes a long time for investors to complete a transaction. Also, the selection of the parameters in PLR-WSVM is another problem. In this paper, we made some improvements to PLR-WSVM. The main contributions of this paper are (1) the prediction of stock trading signal is formulated as a weighted two-class classification problem instead of a weighted four-class classification problem, in which it doesn't need to determine the weight coefficient between classes but also improves the accuracy of classification and of transaction; (2) the trading signals are determined by the comparison of the price of turning points with the average price on the previous five days of the turning point, which making use of historical data sufficiently and improves the accuracy of trading signals generated by machine learning absolutely; (3) Next-day trading instead of the turning point in trade strategy is designed, which furtherly improves the transaction accuracy and the profitability performance; (4) A threshold selection method is presented to choose a threshold for PLR according to the price trend in different segment of a financial time series; (5) two trading strategies are proposed based on the IPLR-WSVM framework, which improves the profitability performance under certain transaction cost. By the analysis of 13 shares from Shanghai Stock Exchange in China and the comparison with PLR-WSVM and buy-and-hold strategy (BHS) methods, this improved method IPLR-WSVM achieves the best stability and the best profitability, which confirmed the validity and practicality of IPLR-WSVM for inflection point prediction.

**Key Words:** Inflection point prediction; PLR; SVM

厦门大学博硕士学位论文摘要库

## 目录

<b>第一章 绪论</b> .....	<b>1</b>
1.1 研究背景及意义 .....	1
1.2 国内外研究现状 .....	4
1.3 论文主要工作和结构安排 .....	7
<b>第二章 理论基础</b> .....	<b>10</b>
2.1 时间序列 .....	10
2.1.1 时间序列概述 .....	10
2.1.2 时间序列表示的主要方法 .....	12
2.2 统计学习理论 .....	17
2.2.1 机器学习 .....	17
2.2.2 VC 维 .....	18
2.2.3 推广性的界 .....	19
2.2.4 结构风险最小化 .....	19
2.3 支持向量机理论 .....	20
2.3.1 最优分类超平面 .....	20
2.3.2 SVM 基本理论 .....	21
2.3.3 SVM 常用核函数 .....	22
2.3.4 加权支持向量机 (WSVM) .....	23
2.4 本章小结 .....	24
<b>第三章 基于 PLR-WSVM 拐点预测的改进研究</b> .....	<b>25</b>
3.1 样本数据的分类 .....	25
3.2 滞后一天交易的设计 .....	29
3.3 PLR 阈值的选择 .....	34
3.4 本章小结 .....	36
<b>第四章 IPLR-WSVM 拐点预测的模型框架</b> .....	<b>37</b>
4.1 PLR 产生拐点 .....	37
4.2 输入变量的介绍 .....	39
4.3 WSVM 建立模型 .....	40
4.4 交易策略设计 .....	41
4.5 本章小结 .....	44
<b>第五章 IPLR-WSVM 有效性验证</b> .....	<b>45</b>
5.1 数据来源 .....	45
5.2 验证方案 .....	45
5.3 结果与分析 .....	45
5.4 本章小结 .....	49
<b>第六章 总结与展望</b> .....	<b>50</b>

6.1 全文总结 .....	50
6.2 存在的问题及进一步的研究工作 .....	50
参考文献 .....	52
附录 .....	59
致谢 .....	60

厦门大学博硕士论文摘要库

## CONTENTS

<b>Chapter 1 Introduction</b> .....	<b>1</b>
<b>1.1 Background and significance of the research</b> .....	1
<b>1.2 Research Status at home and abroad</b> .....	4
<b>1.3 Main Work and Structure of the Paper</b> .....	7
<b>Chapter 2 Theoretical Basis</b> .....	<b>10</b>
<b>2.1 Time Series</b> .....	10
2.1.1 Overview of the Time Series .....	10
2.1.2 The Main Method of Time Series Representation .....	12
<b>2.2 Statistical Learning Theory</b> .....	17
2.2.1 Machine Learning .....	17
2.2.2 Vapnik Chervonenkis Dimension .....	18
2.2.3 Generalization Bound .....	19
2.2.4 Structural Risk Minimization .....	19
<b>2.3 Support Vector Machines Theory</b> .....	20
2.3.1 Optimal Hyper-plane.....	20
2.3.2 Support Vector Machine Theory.....	21
2.3.3 Kernel Function of SVM.....	22
2.3.4 Weighted Support Vector Machine(WSVM) .....	23
<b>2.4 Brief Summary</b> .....	<b>24</b>
<b>Chapter 3 Improvement Research based on PLR-WSVM</b> .....	<b>25</b>
<b>3.1 Classification of the Sample Data</b> .....	25
<b>3.2 Next-day Trading Designing</b> .....	29
<b>3.3 Method of Choosing Threshold for PLR</b> .....	34
<b>3.4 Brief Summary</b> .....	36
<b>Chapter 4 Framework of IPLR-WSVM</b> .....	<b>37</b>
<b>4.1 Inflection Point Produced by the PLR Method</b> .....	37
<b>4.2 Input Variables</b> .....	39
<b>4.3 Model Establishment by WSVM</b> .....	40
<b>4.4 Trading Strategies Applied to IPLR-WSVM</b> .....	41
<b>4.5 Brief Summary</b> .....	44
<b>Chapter 5 Validation of IPLR-WSVM</b> .....	<b>45</b>
<b>5.1 Data Sources</b> .....	45
<b>5.2 Validation Scheme</b> .....	45
<b>5.3 Results and Analysis</b> .....	45
<b>5.4 Brief Summary</b> .....	49

<b>Chapter 6 Summary and Prospect.....</b>	<b>50</b>
<b>6.1 Summary .....</b>	<b>50</b>
<b>6.2 The Existing Problems and Further Research Work.....</b>	<b>50</b>
<b>References .....</b>	<b>52</b>
<b>Appendix .....</b>	<b>59</b>
<b>Acknowledgements .....</b>	<b>60</b>

厦门大学博硕士学位论文摘要库

## 第一章 绪论

### 1.1 研究背景及意义

在商品经济飞速发展的现代社会，人们已转变累积财富的观念，开始注重投资理财，而股票的高收益回报自然受到投资者的普遍关注。然而，股票投资是一种风险性较高的投资行为，股票市场收益和风险并存，且收益和风险往往是成正比，即投资者想要获得较高的投资收益，所承担的风险也较大。因此，如何正确分析和预测金融市场的走势，为投资者提供投资决策参考以降低投资风险，具有极为重要的现实意义。

股票市场充满着不确定性，股票价格受诸多因素影响，如国内外政治变化、宏观经济形势、市场前景预测、投资者心理等，其波动的内部规律也是非常复杂的，因此，股市的可预测性就是一个充满争议的问题。有效市场假说认为，股票价格总是能够“完全反映”市场的所有信息，股价变化是随机游走的，在统计上没有“记忆性”，因此投资者不能根据历史数据来预测未来股价的走势。而现实中，有效市场是很难达成的，投资者总是寻求风险最小化下的利润最大化。研究表明，股票市场既不符合有效市场假说也不符合随机游走模型，股票价格在历史价格模式以及特定“基础性”价值度量基础上存在一定程度上可预测性，即股票市场是可预测的。当然，股票市场并不是能够 100% 的准确预见，股票可预测是指可以利用一些预测方法建立预测模型，该预测模型能够在误差范围内预测股价的变动<sup>[1]</sup>。

随着金融市场的不断深入发展，相关理论知识及实践经验不断的积累，各种股市预测方法相继涌现。在上世纪 30 年代，Benjamin Graham 最先将基本分析引用到证券分析中，试图通过对证券价格的基本因素进行研究，确定证券的真实价值，从而找出那些被市场低估的证券。基本分析能够对证券价格长期走势做出比较精确的预测，但它对市场的反映比较迟钝，预测结果往往较长时间落后于真实走势，在实际应用中被认可的程度不高。随后，又出现了各种技术分析法，包括：K 线理论、波浪理论、技术指标、循环周期、切线理论、形态理论等等。技术分析方法是对市场所表现的数据进行分析，通过图表、技术指标等来发现股票

价格变化的规律,找到最好的买点和卖点。这些分析方法是建立在图表的基础上,而图表的分析与指标的选择受个人的主观判断影响很大。由于股票交易的模式和相关信息是非常复杂的,对于这种严重依靠经验来分析的方法,其可靠性是会受到质疑的<sup>[2]</sup>。

到了 20 世纪 60 年代,时间序列分析方法开始被人们尝试着用来预测股市。时间序列分析的主要思想是:通过分析历史数据来发现时间序列内在规律,并对时间序列未来的发展趋势进行预测,为投资决策提供指导。常用的时间序列分析法主要是建立自回归模型(AR)、移动平均模型(MA)、自回归移动平均模型(ARMA)和齐次非平稳模型(ARIMA),通过选择模型的参数和辨识模型的系数来拟合时间序列,并用拟合好的模型预测未来<sup>[3-7]</sup>。利用时间序列分析法的前提是股票间价格是线性关系,因此局限性非常明显。股票市场瞬息万变,信号是非常复杂的,很少有完美的线性关系,因此这些时间序列分析方法的预测效果并不理想<sup>[8]</sup>。

随着人工智能理论的不断发展与完善,其在模式识别和回归分析等方面的效果良好,并开始被广泛应用于经济、金融等领域<sup>[9-16]</sup>。正是意识到证券市场的复杂性、时间序列分析方法的局限性,人们开始利用人工神经网络技术对股票价格进行分析和预测。与传统方法相比,神经网络在股市预测中显现出一定的优势<sup>[17-19]</sup>:传统方法通常是依靠推导数学模型、寻找最优参数来建立预测模型,往往因参数寻优过于精确而使得模型具有一定的局限性;而神经网络可以逼近任意复杂连续函数关系,可以通过学习来掌握数据间的依存关系,具有广泛的适应能力、学习能力和映射能力<sup>[20]</sup>。股价走势通常呈现高度非线性,而股价变动的内在规律和特点隐藏于成交价、成交量等数据中,神经网络可以通过对历史交易数据进行分析与学习,从而找出各种影响因子之间的关系,并得到股价变动的规律和特点<sup>[11]</sup>。目前,用于证券价格预测的模式主要有:误差反向传播网络(Back-Propagation Network)<sup>[21-25]</sup>、径向基神经网络(Radial Basis Function Network)<sup>[21]</sup>、概率神经网络(Probabilistic Neural Networks)<sup>[26]</sup>和 Elman 神经网络<sup>[27]</sup>等。但是,随着人们对神经网络的深入研究和广泛应用,也发现该方法存在一些缺陷:比如,计算量大,以迭代方式更新连结权重,费时而且耗费计算机资源;训练过程中隐藏层参数、学习率等参数难以确定,需要依赖实验者丰富的经验;



Degree papers are in the “[Xiamen University Electronic Theses and Dissertations Database](#)”.

Fulltexts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to [etd@xmu.edu.cn](mailto:etd@xmu.edu.cn) for delivery details.