

基于轮动规则挖掘与风险控制的量化投资模型

孙如一

山东财经大学, 山东 济南 250022

DOI:10.61369/ER.2024060010

摘要：量化投资凭借其客观性、系统性。随着计算机技术的不断进步以及股票市场的完善和发展，越来越多的投资者们开始使用量化投资交易策略。研究表明，我国的股票市场中存在着明显的行业联动轮动现象，充分的挖掘、把握我国股票市场中的行业联动、轮动规律对于帮助投资者们获取超额投资效益具有重要意义。

关键词：关联规则；行业轮动；量化投资；风险控制

A Quantitative Investment Model Based on Rotation Rule Mining and Risk Control

Sun Ruyi

Shandong University of Finance and Economics, Jinan, Shandong 250022

Abstract：Quantitative investment, characterized by its objectivity and systematic approach, has gained increasing adoption among investors as computer technology advances and stock markets mature. Research indicates that China's stock market exhibits pronounced industry rotation patterns. Thoroughly analyzing these sector-specific dynamics.

Keywords：association rule; industry rotation; quantitative investment; risk control

一、轮动规则挖掘

行业轮动，是指股票市场上的投资热点随着股票市场行情的变化，也在不同行业之间发生轮转，从某些行业转为另一些行业。行业轮动规则相较于行业联动规则更为复杂，其是指一个或多个行业的表现是否能够引起未来某个行业上涨或者下跌的表现。因此把握市场中的行业轮动规则对于指导投资者构建量化投资策略具有极其重要的意义^[1]。

研究表明轮动规则挖掘的周期越长，可信度越高。我们从国泰安CSMAR数据库中筛选了申万2021修订版所划分的共31个一级行业指数的基本信息数据，以及各行业于2022年1月1日到2023年12月31日间的484条交易日历数据与15004条交易指数数据，对其进行了数据清洗以及预处理操作。为便于模型构建与数据回测，仅对A股市场行业间月周期维度下的一对一月度轮动规则进行了挖掘^[2]。计算各行业月度涨跌标记数据，规定轮动规则的最小支持度为0.20，最小置信度为0.70^[3]。调用Apriori挖掘算法得到部分月度轮动规则如下表所示。

表1月维度行业强轮动规则

轮动规则	支持度	置信度
电力设备--综合	0.26087	0.8571428571428571
计算机--通信	0.304348	0.7777777777777778
建筑装饰--通信	0.26087	0.75
机械设备--综合	0.347826	0.7272727272727273
电力设备--农林牧渔	0.217391	0.7142857142857143
电力设备--环保	0.217391	0.7142857142857143

由表1可知，支持度大于0.20，置信度大于0.70的轮动规则有12条。轮动规则“电力设备--综合”表示：若电力设备行业交易指数于第i月呈上涨趋势，则综合行业的交易指数于第i+1月有85.71%的概率也呈上涨趋势；轮动规则“计算机--通信”表示，若计算机行业交易指数于第j月呈上涨趋势，则通信行业的交易指数于第j+1月由77.78%的概率也呈上涨趋势。

二、量化投资模型的构建

考虑到长周期轮动规则所揭示的行业涨跌趋势较短周期更为的准确可信以及选择高置信度的轮动规则设计量化投资策略在一定程度上可以规避数据的噪声干扰以及股市中的各类风险因素，有助于投资策略获得超额收益。因此本文选择以月度为周期进行长线投资，并规定模型选取轮动规则设计量化投资策略的最小置信度为0.75，并使用2024年度交易数据对模型构建的投资策略的收益情况进行回测。

三、投资策略的设计

基于规定的最小置信度0.75，模型选取A股市场月度轮动规则中的“电力设备--综合”“计算机--通信”“建筑装饰--通信”三条轮动规则设计投资策略。以投资策略“电力设备--综合”为例，对其策略进行分析：若电力设备行业的交易指数于第i月呈上涨趋势时，则于第i+1月对综合行业进行投资。

由于投资者们无法就A股市场总的行业的交易指数股直接进

行买卖投资交易，但是可以考虑对行业内综合排名靠前、综合实力较强的上市公司的样本股进行投资^[4]。因此本小节基于“总体规模与投资效率指标的综合评价方法”，根据模型设计的投资策略，选取了综合行业、通信行业综合实力较强、质地较好的上市公司的样本股构成的投资组合来替代投资策略中对行业样本指数股的买卖投资。

“总体规模与投资效率指标的综合评价方法”，综合考虑了上市公司的整体竞争能力、市场抗风险能力、影响力以及投资效率等多方面因素。本文选定使用2024年1月1日-2024年12月31日作为模型投资策略收益率的回测周期，因此选取A股市场通信行业与综合行业的全部上市公司于2023年1月1日-2023年12月31日的指标字段数据来构建投资组合。借助pandas库对存在空值、与负值的指标数据字段进行清洗，仅保留了质地较好的上市公司；使用sklearn库中的StandardScaler函数模块对指标数据进行标准化处理，避免由于指标数据量纲差异造成的实验误差^[5]。

由于指标数据维度较大且数据字段间线性关系较强，本文采用sklearn库中的PCA模块对指标数据进行降维重构，再依据重构后的线性无关的主成分进行上市公司综合评价排名。设置PCA的参数n_components=0.95，根据公式上市公司综合排名F=，计算得到通信行业上市公司2023年度综合排名共93x3条数据；综合行业上市公司2023年度综合排名共11x3条数据。选取通信行业综合排名前10的上市公司（中国电信等）作为投资组合代表通信行业；选取综合行业仅有的得分为正的4家上市公司（浙农股份等）构建投资组合，代表综合行业。

四、投资策略收益率回测

根据模型设计的量化投资策略，我们选取电力设备、计算机、建筑装饰行业于2024年1月1日至2024年11月30日的收盘指数数据，以月度为周期分别计算三个行业的涨跌情况（上涨记为1，否则记为0）。选取通信行业投资组合、综合行业投资组合于2024年2月1日至2024年12月31日的收盘价数据。同时本文选取沪深300的同期收益率作为评价基准，若投资组合的收益率能够持续跑赢同期沪深300指数，说明模型设计的量化投资策略创造了超额收益^[6]；反之则表明我们仍需对模型设计的投资策略的合理性、市场的波动性以及风险管理等因素进行考量^[7]。

根据涨跌标记，电力设备行业的收盘指数于2月、3月、7月、9月、11月呈上涨趋势，则根据量化投资策略“电力设备--综合”，在3月、4月、8月、10月、12月对综合行业投资组合进行投资，计算得到投资策略以及沪深300的收益情况为：综合行业投资组合于3月亏损0.049%，4月亏损0.18%，8月亏损0.089%，10月亏损0.016%，12月亏损0.363%；而沪深300于同期3月亏损0.0009%，4月盈利0.002%，8月亏损0.029%，10月亏损0.086%，12月亏损0.003%。

计算机行业的收盘指数于2月、9月、10月、11月呈上涨趋势，则根据量化投资策略“计算机--通信”，在3月、10月、

11月、12月对通信行业投资组合进行投资，并计算该投资策略收益的收益情况以及基准大盘沪深300的同期收益情况。投资策略“计算机--通信”以及同期沪深300的收益情况为：通信行业投资组合于3月盈利0.046%，10月亏损0.81%，11月亏损0.117%，11月亏损0.117%，12月盈利0.665%；而沪深300于同期3月亏损0.0009%，10月亏损0.086%，11月盈利0.007%，12月亏损0.003%。

建筑装饰行业的收盘指数于2月、3月、9月、10月、11月呈上涨趋势，则根据量化投资策略“建筑装饰--通信”，在3月、4月、10月、11月、12月对通信行业投资组合进行投资，并计算该投资策略收益的收益情况以及基准大盘沪深300的同期收益情况。投资策略“建筑装饰--通信”以及同期沪深300的收益情况为：通信行业投资组合于3月盈利0.046%，4月盈利0.137%，10月亏损0.81%，11月亏损0.117%，11月亏损0.117%，12月盈利0.665%；而沪深300于同期3月亏损0.0009%，4月盈利0.002%，10月亏损0.086%，11月盈利0.007%，12月亏损0.003%。

虽然投资策略“计算机--通信”、“建筑装饰--通信”在大部分投资持有期中可以获得超额收益，但投资策略“电力设备--综合”的收益情况却几乎全线低于同期沪深300指数的收益情况：投资策略“计算机--通信”在持有期3、12月的收益情况跑赢同期沪深300指数；投资策略“建筑装饰--通信”在持有期3、4、12月的收益情况跑赢同期沪深300指数。

基于上述收益情况，我们推测模型受A股市场的风险因素影响导致其出现了投资策略效益不稳定的情况，即模型受噪声干扰出现了误差^[8]。

五、风险控制

在量化投资中，风险控制是至关重要的，它涉及到识别、评估和控制投资过程中可能遇到的风险^[9]。有效的风险控制可以帮助投资者保护资本，在避免重大损失的同时有效提高投资回报，所以投资者总是希望能够找到应对的方法来减少投资的风险而增加收益^[10]。

在上一小节的分析中我们认为，模型构建的投资策略收益情况可能受风险因素的影响存在较大波动性。因此本章将从分散投资的角度出发，根据投资组合中股票的不同表现，对其投资比例进行合理分配，将风险控制模型融入模型，对模型构建的投资策略进行改进，进而构建更加合理、可靠的量化投资策略与投资模型。

六、投资组合优化

本小节选取通信行业投资组合为例进行模型策略优化以及组合风险控制，具体实现为：通过市值加权、蒙特卡洛模拟以及夏普最优化率的方式对通信行业投资组合内的各股投资权重进行重新计算分配。并计算通信行业投资组合在三种风险优化下的累计

收益情况，与同期未进行风险控制优化的等权重投资组合累计收益情况以及基准大盘沪深300收益情况进行对比。本节中使用的实证回测数据均来自国泰安CSMAR数据库，时间维度均为2024年1月1日至2024年12月31日。

七、市值加权权重分配

个股总市值(简称市值)是股票价格与总股本的乘积，市值情况可以体现公司的规模与市场地位。在以上市公司股票的市值作为衡量标准时，我们认为市值高的股票可投资性更强、流动风险更小，因此分配给高市值股票大投资权重占比，当这些高市值占比的股票表现优异时，投资组合的整体收益情况也将极为可观。使用市值加权的方式对投资组合内各股投资权重进行重新分配，计算投资收益情况。获取市值加权投资组合权重为：0.1041827, 0.04473087, 0.1398353, 0.00488528, 0.43719653, 0.00977141, 0.12617345, 0.0399361, 0.04438967, 0.04889869。

(一) 蒙特卡罗最小风险权重分配

使用蒙特卡罗模拟生成10000次随机权重数据，计算每一次随机权重下投资组合的收益率，从所有的模拟结果中选择年化波动率最小的一组权重结果作为投资组合的最小风险权重，使用idxmin函数对有效边缘最小风险值进行筛选，最终得到的最小风险投资组合权重为(结果保留三位小数)：0.021, 0.041, 0.069, 0.013, 0.143, 0.236, 0.046, 0.211, 0.099, 0.121。

(二) 夏普比率权重分配

夏普比率(Sharpe Ratio)是衡量投资组合单位风险超额收益的经典指标其用以帮助投资者比较投资的回报和风险。其计算公式为： $\frac{R_p - R_f}{\sigma_p}$ 。其中是预期回报率，是无风险利率，是超额收益的标准差。

本小节将在蒙特卡罗模拟的结果上计算夏普比率，试图构建最小风险-最大收益的投资权重。设置无风险收益情况 free_risk_Returns=0, 使用 idxmax0 函数找到有效边缘，即最大夏普值的索引。最终获取的夏普最优单位风险-收益比投资权重为：0.031, 0.226, 0.004, 0.016, 0.059, 0.232, 0.032, 0.185, 0.002, 0.215。

(三) 优化结果分析

借助 pandas 库对三种投资权重分配方式下的投资组合累计收益

率以及同期沪深300累计收益率进行计算，并对其进行可视化输出。

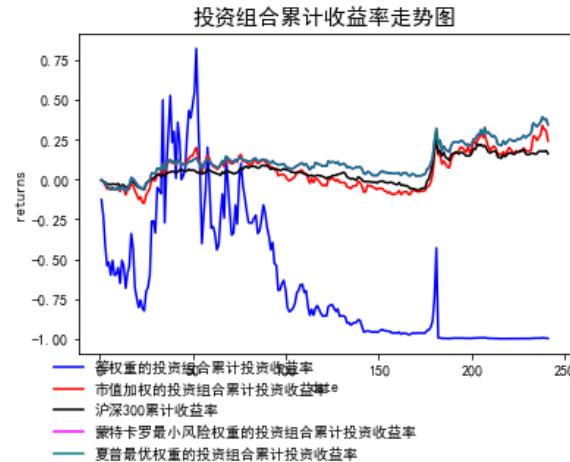


图1不同投资权重的投资组合累计收益率走势图

由图1可知：在高风险、高波动性的市场环境中，不进行任何风险控制的投资策略在短期内可能获取超额收益的(前提是构建了综合实力较强的投资组合)；但从长期来看该投资策略必然会受到市场高风险的反噬导致出现较大亏损；以市值加权、蒙特卡罗以及夏普比率的优化方式构建的投资组合累计收益情况无论是短线还是长线都取得了较好结果，且蒙特卡罗与夏普比率投资组合累计收益情况稳定跑赢同期沪深300，市值加权投资组合虽存在一定的波动但收益情况基本与沪深300持平或略优。

八、总结

本文构建了一个具有较强抗风险能力的量化投资模型，基于Apriori算法挖掘分析A股市场的行业轮动规则，选取置信度高于0.75的月度轮动规则设计量化投资策略，并采取市值加权、蒙特卡罗模拟以及夏普比率等方式对投资组合权重分配，抵御市场中的非系统风险。经回测实证表明，该量化投资风险控制模型设计的量化投资策略可以于长线投资中获得稳定的超额收益，为投资者们提供合理、有效的投资参考。

参考文献

- [1]周彩节.基于关联规则的我国股市行业轮动现象研究[J].中国市场,2020(26):32-33.
- [2]赖萱.基于经济周期视角的中国A股市场行业轮动投资策略研究[D].南宁:广西大学,2021.
- [3]殷丽凤,李梦琳.基于关联规则算法的A股股票联动性分析[J].云南民族大学学报(自然科学版).2022,31(4):486-492.
- [4]李艳玲,邓淋净.基于Apriori关联规则算法的辽宁省上市公司行业联动与轮动分析[J].商展经济.2023(17):082-085.
- [5]张牧云.经济周期框架下中国A股市场行业轮动与传导机制[D].安徽大学,2014.
- [6]赖美欣.基于挖掘算法的申银万国行业指数关联规则分析[J].中国管理信息化.2023,26(4):159-161.
- [7]燕吉吉.关联规则在中国股市行业轮动中的应用[D].青岛:青岛大学,2018.
- [8]刘欣雨.基于关联规则的中国A股市场行业轮动现象研究[D].北京邮电大学,2016.
- [9]李林.浅析资本市场量化投资策略和风控措施[J].当代经济,2018(19):122-123.
- [10]白一池.资本市场量化投资策略和风控措施探析[J].现代商业,2020(5):117-118.