



量化投资课程实验报告

Experimental report of quantitative investment course

报告人：刘 烁 25120182202202

赵哲冕 17420182200850

二〇二一年五月二十九日

目录

第一章	绪论	3
第一节	研究背景与意义	3
一、	研究背景	3
二、	研究意义	3
第二节	文献综述	4
一、	国内外量化模型的发展状况	4
第二章	理论综述	6
第一节	金融市场相关理论综述	6
一、	对 A 股进行量化投资分析的优势	6
第二节	量化选股相关概念及理论	6
一、	Alpha 收益理论分析	6
二、	多因子选股相关概念	8
第三章	多因子模型在我国 A 股市场的实证分析	13
第一节	数据的选择和处理	13
一、	因子有效检验性	13
第二节	Alpha 策略下的多因子模型	14
第三节	回测结果	22
第四章	研究结论与建议	23
第一节	研究结论	23
第二节	相关建议	23
第三节	收获与体会	23
参考文献		25

第一章 绪论

第一节 研究背景与意义

一、研究背景

量化投资是指利用数学建模、计算机技术等方式进行数据分析与预测，找寻投资标的并实行投资策略的投资方法。量化投资是传统投资学的一次革命，在基本理念上二者是相似的，都是基于市场弱有用的理论，试图克服市场而获取超额收益。但与传统的投资方式相比，量化投资的分析方式有更多的数据作为支撑，故而相较于传统方式而言，量化投资更有可能使投资者获得更加连续的收益。

在 2020 年新冠疫情爆发之时从沪深两市来看，新冠肺炎爆发以来，沪深两市反应强烈。艾媒咨询发布的《2020 中国新冠肺炎疫情正/负影响行业及典型企业业绩影响分析报告》显示，28 个一级行业春节前后 20 个交易日内总体呈现涨少跌多态势，其中电子、计算机、医药生物涨幅靠前，休闲服务行业跌幅最大，下跌近 15%。疫情之下，餐饮、旅游、交通运输等行业受损严重，而生鲜配送、在线医疗、在线办公、在线教育、游戏娱乐等各类线上业务组成假期里的“宅经济”，在 2020 年初成为弯道超车的大热风口。

新型疫情影响，2020 年 2 月 3 日，上证指数收盘跌 7.72%报 2746.61 点；深证成指跌 8.45%报 9779.67 点；创业板指跌 6.85%报 1795.77 点。

A 股波动之外，更应看到当时国家的两大罕见操作。2.3-2.4 两天之内，央行为市场注入了极为富裕的流动性。在 2 月 3 日创纪录单日公开市场操作投放 1.2 万亿元(人民币，下同)后，4 日人民银行继续充足供应流动性，当日公开市场操作逆回购投放 5000 亿元。1.7 万亿元充分显示央行稳定市场预期、提振市场信心的决心。除政府驰援之外，商业资本也开始踊跃进入 A 股，包括北上资金、公募基金、险资及私募，由此 A 股迅速走出了修复行情。

故这次实验报告我们想探究疫情期间量化投资对疫情期间的股市是否有适用性、普遍性。

二、研究意义

建立一套科学的评估体系后，有利于发掘一些真正具有投资价值的企业，使资源分配到最需要的地方，降低资源的浪费的可能性，提高了市场的效率；可为企业在确定入市企业资格方面做给出参考。

同一企业，由于主体需要、偏好及判断能力的不同，也会呈现出不同的评估结果。本文主要从投资者与证券从业人员的角度出发，聚焦企业内在价值，探索分析量化内在价值的方法。理解并掌握内在价值的估算方法所在是投资管理和成功的关键，也是企业内部管理当局科学决策、理财人员进行有效的价值管理及寻求企业价值的不断提升实现价值最大化目标的关键。在现今 Python 等计算机技术的支持下，收集分析大数据之后的量化建模使得对公司内在价值计算更加便捷科学，进而更有助于投资者做出更优的投资决策。

就社会意义而言，相比于国外股票市场，我国股市发展时间较短，投资者个人情况参差不齐，投资理念尚不成熟，国内 A 股市场仍属于非有效市场，而投资者非理性行为也广泛存在，市场信心及政策信号的变化常常引起市场的过度反应或反应不足。因此急需一个可控标准来衡量 A 股投资价值，以给予广大股民一个投资趋势，来更好地进行理性投资。

第二节 文献综述

一、国内外量化模型的发展状况

马克维茨提出经典的均值-方差组合模型标志了量化投资理论的正式兴起，为投资者提供了最优的资产配置比例，使得在达到预期收益水平的情况下，尽量使风险最小化。但由于当时电脑并未普及，计算能力十分有限，导致该模型在实际应用过程中面临计算量大、难度和成本较高等问题。为了解决这些问题，在 1960 年代初，夏普、特里诺、莫辛通过深入挖掘并简化马克维茨模型，推导出了经典的现代投资学的奠基模型——资本资产定价模型（CAPM 模型）。该模型假设所有投资者对期望收益、方差和协方差等的估计完全相同，投资人可以自由借贷。基于这样的假设，资本资产定价模型得出了这样的公式：

$$E(R_q) - R_f = \beta_{qm} [E(R_m) - R_f]$$

其中， R_m 是指投资组合的收益率， β_{qm} 是指投资组合 q 的 Beta 值，反映了资产组合 q 的系统性风险， $E(R_q) - R_f$ 为风险投资组合 q 比无风险资产高出的期望收

益率，因此被称为风险溢价。这个公式研究的重点在于探求风险资产收益与风险的数量关系，即为了补偿某一特定程度的风险，投资者应该获得多少的报酬率。

1970 年代初，尤金·法玛提出了有效市场假说，把市场分成强有效市场、半强有效市场和弱有效市场。在之后几年，随着金融改革的进行，金融衍生产品的种类不断增多，其定价问题成为学术研究的重点之一。莫特于 1971 年提出期权合理定价理论，而后布莱克、肖尔斯几乎同时推导出期权定价模型（B-S-M 模型），成功解决了金融市场中衍生产品的合理定价问题。罗斯在随后提出 CAPM 模型的拓展——套利定价理论（APT 理论）。但与 CAPM 模型不同的是，APT 理论以多因子模型为研究基础，认为资产收益率与多个因素之间存在线性关系。许多学者以 APT 理论为研究基础，建立多因子选股模型，试图通过一些因子判断股票的投资价值，如法玛和阿斯内斯进行的相关研究。而 1980 年代以来，公司财务、商业银行和投资银行业务的迅猛变化导致了金融工程这一新兴学科的诞生。美国金融学教授芬纳蒂最早提出金融工程这一概念，创造性地使用金融市场中的产品和策略，解决现实当中金融和财务问题的一项应用性技术工程。这之后，在 1987 年美国股市大崩盘的教训下，金融学家们更加注重控制金融资产交易面临的市场风险。这其中，摩根的 VaR 模型是极重要的风险管理模型之一。与此同时，非线性研究方法在金融领域中也得到了广泛的应用，包括了遗传算法、ID3 算法，BP 神经网络等非线性回归方法。

上述有关多因子量化选股模型的文献从研究内容上其主要区别可以概括为以下三个方面：一是不同研究确定的影响股票价格波动的因子存在差异。股票的价格波动受很多因素的影响，且不同因子之间可能存在极强的相关性，因此在多因子量化选股模型的实证研究中，由于不同学者建立模型时选取的候选因子不同，得到的最终结果可能存在一定的差异；二是模型的判别方法不同。多因子选股模型有打分法和回归法两种判别方式；三是收益率选取的时间跨度不同。按照时间长度进行划分，股票收益率的选取在不同研究中主要分为三类，分别为长期收益率、中期收益率和短期收益率。其中，长期收益率以 5 年或 3 年累计收益率为代表，中期收益率主要是指年度收益率和半年度收益率，而短期收益率包括季度收益率、月收益率和周收益率等。

第二章 理论综述

第一节 金融市场相关理论综述

一、对 A 股进行量化投资分析的优势

数据优势。中国上市公司的财务数据披露完整，数据库更新迅速。通过常用的金融数据库，我们能够很容易地获取上市公司的财务数据。并且，数据库提供的数据也包含上市公司高管、股权结构、董事的数据，美国股票市场超过 80% 的交易量来自机构投资者。所以在美国股票市场，放眼望去，都是专业的机构投资者，具有极强的信息收集和、处理能力，但凡有一丁点超额收益的机会，机构投资者就会立刻发觉并利用。然而在中国股票市场上，根据上海证券交易所，量化投资的核心，在于提取与未来股票收益率相关的信息。这就需要市场满足两点：第一，市场是有效的，即与收益率、定价相关的信息终究会反映在价格当中。如果市场完全无效，价格不受信息的驱动，那么即使将一只股票方方面面的信息都掌握透彻，也无法对未来的收益率进行预测。第二，与股票定价相关的信息并不能够立刻、完全反映在价格当中，而是需要一定的时间。如果价格立刻反映信息，不给量化投资留任何分析、建仓的时间，那么量化投资也是没有用武之地的。也就是说，量化投资的适用，需要市场是有效的，但又不是那么有效，这样利用量化投资方法提取信息，才能对未来做出预测。

第二节 量化选股相关概念及理论

一、Alpha 收益理论分析

（一）Alpha 收益介绍

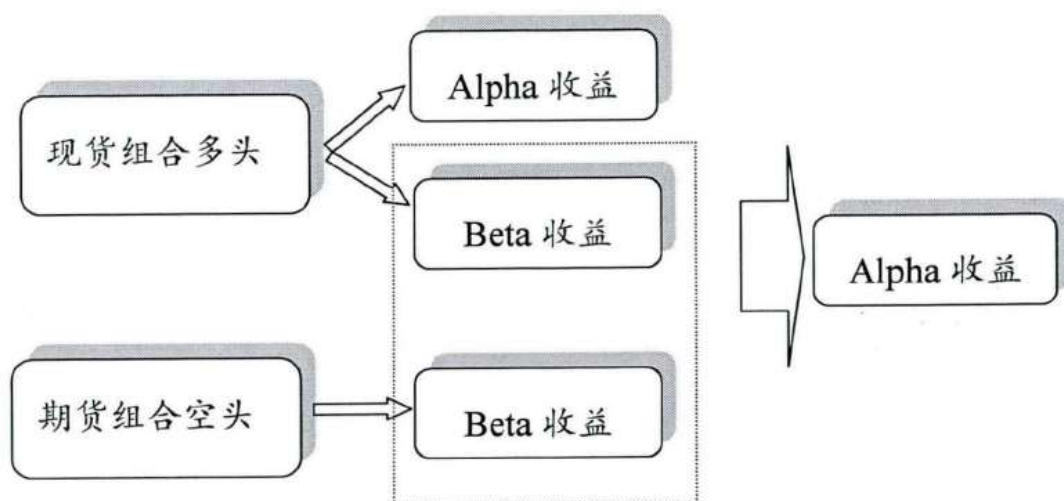
现在投资组合理论之父 Markowitz 在 1952 年首次系统深入提出投资组合理论，该理论提出用资产收益率的期望来代表收益，用其方差来代表风险，对现代投资组合研究具有划时代的里程碑意义。在投资组合理论的基础上 William Sharp 进一步提出资本资产定价模型（CAPM），该模型反映了资产应该依据其所承担的系统风险获取风险溢价补偿，也即资产的正常收益。利用 CAPM 模型我们可以得到资产的正常收益，而 Jensen's Alpha 收益，也即本文研究的策略所追求的收益，定义如下：

$$\alpha = R_i - E(R_i)$$

其中 R_i 表示资产的实际收益率， $E(R_i)$ 为用模型求得的资产预期收益率，即

$$E(R_i) = R_f + \beta(R_m - R_f)$$

R_f 为市场无风险收益， R_m 为有效市场组合的收益率。利用图例表示如下：



从而在实际操作中资产组合的实际收益可分为超额收益和按照其承担风险所获得的正常收益。

(二) Alpha 策略简介

传统的资产管理是基于收益风险平衡的管理哲学，同时追求平均市场收益（Beta 收益）和超额收益（Alpha 收益）。然而两者往往很难兼得。Beta 收益为按照承担风险所获取的期望补偿，与市场的波动性相关，具有很大的随机性。而 Alpha 收益由于对冲掉了系统风险，往往可以通过量化选股的办法使其相对稳定。因而国外很多大型对冲基金为获取稳定收益率常常采用此类办法将 Alpha 收益与 Beta 收益分离，进行相应的模型构建。

Alpha 策略具体是指在建立一定的资产组合后，通过统计手段求得其相应的 β 值，然后利用卖空衍生金融工具（比如股指期货）对冲 β 部分的风险，从而获取正的收益。

Alpha 策略并不依靠对股票组合或指数的趋势判断，而是研究其相对于大盘的投资价值。该策略的两个关键步骤分别是：1、利用选股模型选取能获得收益

的股票组合；2、将 Alpha 收益和 Beta 收益分离，利用衍生金融产品将 Beta 部分的系统风险对冲。

二、多因子选股相关概念

（一）基本概念

多因子量化选股模型相对单因子模型而言，即在选股过程中不单单只考虑一个因子对股价走势的影响，而是利用多个对股价走势有显著且有效的因子通过数量化的思想建立起来的选股模型。在整个选股过程中，满足多因子选股模型条件的股票则被买入，不满足的则卖出。多因子选股模型相对单因子选股模型来说相对稳定，克服了单因子模型可持续性差的缺点；在不同的市场条件下，总有一些因子能够发挥作用，避免单因子模型选股面窄、选股不科学等缺点。

（二）因子分析可行性检验

并不是所有的备选因子都对模型有效，选择有效的因子至关重要，因此有必要进行因子的有效性检验。

关于因子的有效性检测方面，前人通常采用排序法来检验备选因子的有效性。对于任一备选因子而言，在模型建立的第一个月初计算市场中每只股票该因子数值的大小，再按照一定（从小到大）次序对样本空间的所有股票进行排序，并将它们平均分成 i ($i=1, 2, \dots, n, n \in N$ 份)，计算每个月的收益率与排名，持有这些组合到期末。期末计算这 n 个组合的年化复合收益率、相对市场基准收益的超额收益率、各个组合跑赢市场基准的概率。为了检测选股因子的有效性，有如下标准：

(1) 如果某个因子能显著的影响组合的预期收益率，各个组合的排列序数应该跟组合的年化收益率有较大的相关关系。设序数为 i ($i=1, 2, \dots, n, n \in N$) 的组合年化复合收益率为 x_{ii} ($i=1, 2, \dots, n, n \in N$)，所有组合的排列序数就是 $I=[i_1, i_2, \dots, i_n]'$ ，与之对应的收益率就是 $X=[x_1, x_2, \dots, x_n]'$ ， X 与 I 的相关性绝对值 $Abs(Corr(X, I))$ 应该满足： $Abs(Corr(X, I)) \geq Min(Corr)$ 。

其中， $Min(Corr)$ 是模型所设定最小的相关性临界值。

(2) 设排在第一位的和最后一位的两个极端组合相对市场基准的超额收益分别是 AR_1 和 AR_n 。如果 $AR_1 > AR_n$ (因子越小, 收益越大), 两者须满足如下条件: $AR_1 > \text{Min } AR_{\text{top}} > 0$ 和 $AR_n < \text{Min } AR_{\text{bottom}} < 0$ 。反之, 如果 $AR_1 < AR_n$ (因子越小, 收益越小), 两者须满足: $AR_n > \text{Min } AR_{\text{top}} > 0$ 和 $AR_1 < \text{Min } AR_{\text{bottom}} < 0$ 。其中, $\text{Min } AR_{\text{top}}$ 和 $\text{Min } AR_{\text{bottom}}$ 分别是两个极端组合超出市场基准的最小临界值, 前者是赢市场基准的组合, 后者是跑输市场基准的组合。

(3) 在检验期间, 无论市场的走势是以何种形式上涨、下跌、振荡, 排在第一位的较高收益组合应该以较高的概率跑赢市场基准, 排在最后一位的较低收益组合应该以较高的概率跑输市场基准。如果某一因子同时满足上述三个条件, 至少能说明该因子在过去某一特定的时间段里可以表现出良好的选股能力, 这个因子也就可以作为被进一步筛选的有效因子。

(三) 候选因子的选取

候选因子的选择主要依赖于经济逻辑和市场经验, 但选择更多和更有效的因子无疑是增强模型信息捕获能力, 提高收益的关键因素。

例如: 在 2018 年 1 月 1 日, 选取流通市值最大的 50 支股票, 构建投资组合, 持有到 2018 年底, 则该组合可以获得 10% 的超额收益率。这就说明了在 2018 年这段时间, 流通市值与最终的收益率之间存在正相关关系。

从这个例子可以看出这个最简单的多因子模型说明了某个因子与未来一段时间收益率之间的关系。同样的, 可以选择其他的因子, 例如可能是一些基本面指标, 如 PB、PE、EPS 增长率等, 也可能是一些技术面指标, 如动量、换手率、波动等, 或者是其它指标, 如预期收益增长、分析师一致预期变化、宏观经济变量等。同样的持有时间段, 也是一个重要的参数指标, 到底是持有一个月, 还是两个月, 或者一年, 对最终的收益率影响很大。

因子类型可以概括为 9 类: 规模因子, 估值因子, 成长因子, 盈利因子, 动量反转因子, 交投因子, 波动率因子, 分析师预测因子。

1、规模类因子: 总市值, 流通市值, 自由流通市值

- 2、估值类因子：市盈率（TTM），市净率，市销率，市现率，企业价值倍数
- 3、成长类因子：营业收入同比增长率、营业利润同比增长率，归属于母公司的近利润同比增长率、经营活动产生的现金流金额同比增长率
- 4、盈利类因子：净资产收益率 ROE、总资产报酬率 ROA、销售毛利率、销售净利率
- 5、动量反转因子：前一个月涨跌幅，前 2 个月涨跌幅、前 3 个月涨跌幅、前 6 个月涨跌幅
- 6、交投因子：前一个月日均换手率
- 7、波动因子：前一个月的波动率，前一个月的振幅
- 8、股东因子：户均持股比例、户均持股比例变化、机构持股比例变化
- 9、分析师因子：预测当年净利润增长率、主营业务收入增长率、最近一个月预测净利润上调幅度、最近一个月越策主营业务收入上调幅度，最近一个月上调评级占比

（四）候选因子的评价方法

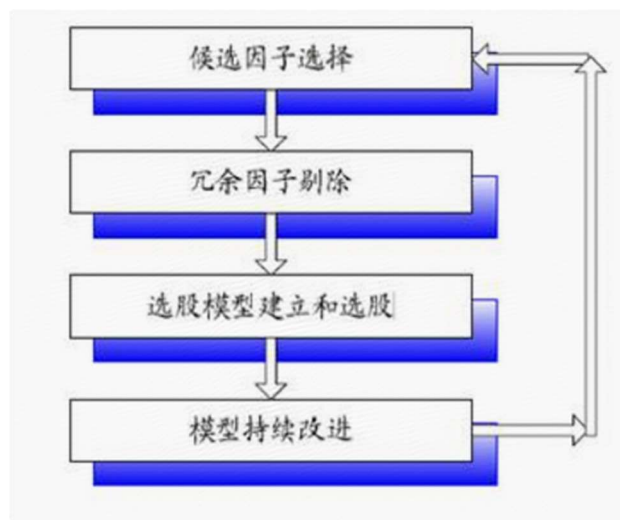
打分法和回归法

1、打分法的评价原理和流程

所谓打分法，就是根据各个因子的大小对股票进行打分，然后按照一定的权重加权得到一个总分，最后根据总分再对股票进行筛选。对于多因子模型的评价而言，实际通过评分法回测出的股票组合收益率，就能够对备选的选股模型做出优劣评价。

打分法的优点是相对比较稳健，不容易受到极端值的影响。但是打分法需要对各个因子的权重做一个相对比较主观的设定，这也是打分法在实际模型评价过程中，比较困难和需要模型求取的关键点所在。

进一步从打分法的流程来看，多因子选股模型的建立、评价和改进流程，大致可以分为 4 个步骤：



此外，对于量化选股打分法，专业人士还提示指出，一方面，多因子选股模型中有的因子会逐渐失效，而另一些新的因子可能被验证有效而加入到模型当中；另一方面，一些因子可能在过去的市场环境下比较有效，而随着市场风格的改变，这些因子可能短期内失效。在这种情况下，对综合评分选股模型的使用过程中，需要对选用的因子、模型本身做持续的再评价和不断的改进以适应变化的市场环境。除此之外，在计算综合评分的过程中，除了各因子得分的权重设计之外，交易成本和风险控制等因素，也同样需要予以综合考量。

2、多元线性回归简介

所谓回归法，就是用过去的股票的收益率对多因子进行回归，得到一个回归方程，然后再把最新的因子值代入回归方程得到一个对未来股票收益的预判，然后再以此为依据进行选股，并对选股模型的有效性和收益率进行评价。

回归法的优点是能够比较及时地调整股票对各因子的敏感性，而且不同的股票对不同的因子的敏感性也可以不同。回归法的缺点，则是容易受到极端值的影响，在股票对因子敏感度变化较大的市场情况下效果也比较差。

在线性回归分析中，如果有两个或两个以上的自变量，就称为多元线性回归。因此，通过多元线性回归对多因子选股模型进行评价，也能够得到一个直观的股票组合收益率结果，同时能够有效评价该选股模型的优劣。

从数学的角度来说，假设因变量 Y （预期收益率）是自变量 $X_1, X_2, X_3, \dots, X_k$ （候选因子）的线性函数，用方程来表示就是：

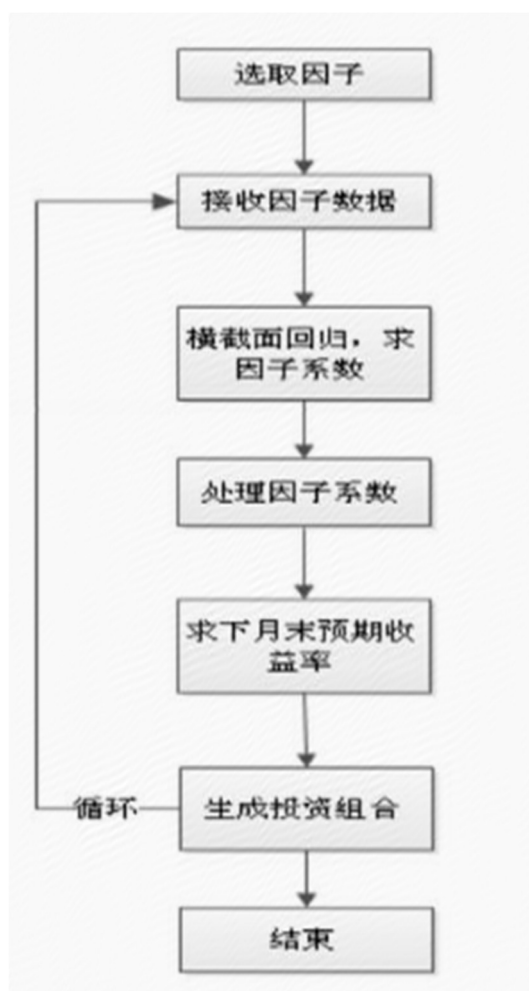
$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i$$

其中， Y_i 表示因变量（被解释变量）的第 i 个观测值，而 X_{ki} 则是第 k 个自变量（解释变量）的第 i 个观测值，是 X_k 的系数， ε_i 是第 i 组观测值的残差项。在金融领域， β_0 有时候会写成 α ，该方程来表示也可以写作：

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i$$

在此之中，多元线性回归通常采用普通最小二乘法（OLS）进行估计，普通最小二乘估计法的思路是改变 $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ ，使得残差的平方和最小。

从回归法的流程来看，多因子选股模型的建立、评价和改进流程，大致可以分为 6 个步骤：



第三章 多因子模型在我国 A 股市场的实证分析

第一节 数据的选择和处理

一、因子有效检验性

在多因子研究框架中，因子的有效性检验是不可避免的工作，其本质是衡量一个因子的选股能力。其检验方法主要包括以下几个内容。

（一）相关性检验

因子的相关性检验即检验单因子和收益率之间是否存在相关性。

（1）IC 值

计算同一时刻的个股的指标值和未来一段时间（通常为一个月）收益的相关性，也就是 IC 值（信息系数）。 $IC \in [-1, 1]$ ，绝对值越大，预测能力越好。IC 值的计算方法包括 Normal IC 和 Rank IC；分别对应 Pearson 相关系数和 Spearman 相关系数。

由 t 期因子载荷预测得到的 $t+1$ 期收益预测值与收益实际值的相关系数

$$ICA = \text{Pearson}(fA, r)$$

ICA : 因子 A 在该期的 IC 值

fA : 用 $t+1$ 期收益率的预测值

r : $t+1$ 期实际收益率

Normal IC 前提条件收益服从正态分布，也可以使用当期因子值与下期实际收益率的 Pearson 相关系数。

Rank IC 与 Normal IC 不同之处在于采用秩相关系数，适用于各种收益分布

$$IC_{rank} = \text{Pearson}(f_{index}, r_{index})$$

（2）IR 值的定义

IR 的是 Information Ratio 的缩写，即信息比率。IR 指的是超额收益的均值与标准差之比。需要多个调仓期，每一个调仓期计算出一个 IC 值。IR 代表因子获得稳定 Alpha 的能力。 $IR = (\text{IC 的多周期均值}) / (\text{IC 的标准方差})$ 。IR 有时指因子 IR，有时是策略 IR，对于策略 IR：策略 $IR = (\text{超出指数收益超出指数}) / (\text{收益的波动率策略})$ ，策略 IR 表示策略稳定战胜基准指数的能力。

（二）IC 与 IR

IR 等于多个调仓周期的 IC 均值除以这些 IC 的标准差，所以 IR 兼顾了因子的选股能力（由 IC 表征）和选股能力的稳定性（由 IC 的标准方差的倒数代表）。

IC 最大值为 1，表示该因子选股 100%准确，对应的排名分最高的股票，选出来的股票在下格调仓周期中涨幅最大；反之-1，则是反向指标。实际中对于多年的 IC 均值，当 $IC > 0.05$ 时，就可以视为有效因子，当 $IC > 0.1$ ，就可以认为是很好的阿尔法因子了。当 IC 均值接近 0，可视为无效因子。

IR 大于 0.5 时，因子稳定获取超额收益的能力较强。

（三）单调性检验

有时因子在统计上并不表现出对于未来收益率很好的预测能力，但是可能由于该因子的复杂逻辑，其在策略中仍能获得超额收益。于是可以用更直接的方法检验该因子的选股能力。

step1 每个调仓期，按照因子指标大小对股票池中所有股票分组，一般根据券池大小为 5 组或 10 组；step2 组内按等权重或者市值加权进行历史数据回测；step3 多次调仓期后观察回测结果，包括累计收益，最大回撤，IR 值，胜率等指标和净值曲线的层次划分。如果优势组各指标越好，净值曲线层次划分越明显，说明单调性越强，该因子越有效。

第二节 Alpha 策略下的多因子模型

Alpha 策略最初是由 William Sharpe 在 1964 年的著作《投资组合理论与资本市场》中首次提出的，并指出投资者在市场中交易面临系统性风险和非系统性风险，用公式表达如下：

$$E(R_p) = R_f + \beta * (R_m - R_f)$$

其中， $\beta = Cov(R_i, R_m) / Var(R_m)$, $E(R_p)$ 表示投资组合的期望收益率， R_f 表示无风险报酬率， R_m 表示市场组合收益率， β 为某一组合的系统风险系数。CAPM 模

型主要表示单个证券或投资组合系统风险收益率之间的关系，也就是说，单个投资组合的收益率等于无风险收益率与风险溢价之和。Alpha 策略的思想是通过衍生品来对冲投资组合的系统风险 β ，锁定超额收益 Alpha。因此首先需要寻找稳定的 Alpha，构建 Alpha 组合，进而计算组合的 β 来对冲风险。Alpha 策略成功的关键就是寻找到一个超越基准（具有股指期货等做空工具的基准）的策略。比如，可以构造指数增强组合+沪深 300 股指期货空头策略。这种策略隐含的投资逻辑是择时比较困难，不想承受市场风险。

市场中性 Alpha 策略是一套科学且成熟的投资研究方法，在实际应用中，我们通照用多因子模型整动、期货对冲来获取长期稳定的 Alpha。同时，使用多因子模型可有效地结合基本面和技术面，使策略更可靠，更稳健。另外，该策略的投资标的广市场容量大，基本没有产品规模限制。

通过上面的介绍，我们对市场中性 Alpha 策略有了较清晰的认识，这里利用多因子模型进一个股票组合，同时选取基准指数对应的期货进行对冲，该可以将理用照中性 Alpha 策略。

下面是一个具体的市场中性策略的实例。

(1)策略配置如下。

回测区间：2020 年 1 月 1 日至 2029 年 8 月 26 日。

基准：沪深 300 指数。

股票池：沪深 300 动态成分股，在每个交易日都进行调仓。

(2)对因子的选取和处理如下。

对因子的选取：净利润增长率、净资产收益率、相对强弱指数（RSI）。

对因子的处理：用到了去极值（Winsorize）、标准化（Standardize）

中性化（Neutralize）处理。

组合构建：等权配置。

(3)空头期货对冲为：空头使用期货 IF 当月合约做空来进行对冲。如果距离合约到期日有 3 天，则移仓换月。IFL0 表示当月合约，IFL1 表示次月合约。

代码展示如下：

```

from CAL.PyCAL import *

import numpy as np

from pandas import DataFrame

start = '2020-01-01'                # 回测起始时间
end = '2020-08-26'                  # 回测结束时间
universe = DynamicUniverse('HS300') + ['IFL0','IFL1']    # 证券池，支持股票、基金、期货、指数四种资产
benchmark = 'HS300'                  # 策略参考标准
freq = 'd'
refresh_rate = 1
capital_base = 10000000
cal = Calendar('China.SSE')
period = Period('-1B')

#配置账户信息，支持多资产多账户
stock_commission = Commission(buycost=0.0, sellcost=0.0, unit='perValue')
futures_commission = Commission(buycost=0.0, sellcost=0.0, unit='perValue')
slippage = Slippage(value=0, unit='perValue')

accounts = {
    'stock_account':
AccountConfig(account_type='security',capital_base=capital_base,commission=stock
_commission,slippage=slippage),
    'futures_account':
AccountConfig(account_type='futures',capital_base=capital_base,commission=futures
_commission,slippage=slippage)
}

```



```

def initialize(context):
    context.signal_generator =
SignalGenerator(Signal('NetProfitGrowRate'),Signal('ROE'),Signal('RSI'))
    context.need_to_switch_position = False
    context.contract_holding = ''
    pass

def handle_data(context):
    universe = context.get_universe(exclude_halt=True)
    yesterday = context.previous_date
    signal_composite = DataFrame()

    # 净利润增长率
    NetProfitGrowRate = context.signal_result['NetProfitGrowRate']
    signal_NetProfitGrowRate =
standardize(neutralize(winsorize(NetProfitGrowRate),yesterday.strftime('%Y%m%d'))))
    signal_composite['NetProfitGrowRate'] = signal_NetProfitGrowRate

    # 权益收益率
    ROE = context.signal_result['ROE']
    signal_ROE =
standardize(neutralize(winsorize(ROE),yesterday.strftime('%Y%m%d'))))
    signal_composite['ROE'] = signal_ROE

    # RSI
    RSI = context.signal_result['RSI']
    signal_RSI =
standardize(neutralize(winsorize(RSI),yesterday.strftime('%Y%m%d'))))

```

```

signal_composite['RSI'] = signal_RSI

# 信号合成，各个因子权重
weight = np.array([0.6,0.3,0.1])
signal_composite['total_score'] = np.dot(signal_composite, weight)

# 订单委托
def handle_stock_orders(context, wts):
    account = context.get_account('stock_account')

def handle_futures_orders(context):
    stock_account = context.get_account('stock_account')
    future_account = context.get_account('futures_account')

# 组合构建
total_score = signal_composite['total_score'].to_dict()
wts = simple_long_only(total_score, yesterday.strftime('%Y%m%d'))
handle_stock_orders(context, wts)
handle_futures_orders(context)
# handle_futures_position_switch(context)

account = context.get_account('stock_account')
stock_account = context.get_account('stock_account')
future_account = context.get_account('futures_account')

log.info(u'时间: %s' % (context.current_date))

# 先卖出

```

```

sell_list = account.get_positions()

for stk in sell_list:
    account.order_to(stk,0)

# 再买入
buy_list = wts.keys()
total_money = account.portfolio_value
prices = account.reference_price
for stk in buy_list:
    if np.isnan(prices[stk]) or prices[stk] == 0: # 停牌或者还没有上市等
原因不能交易
        continue
    account.order(stk, int(total_money * wts[stk] / prices[stk] / 100) * 100)

# 将主力连续合约映射为实际合约
contract_current_month = context.get_symbol('IFL0')

# 判断是否需要移仓换月
contract_holding = context.contract_holding
if not contract_holding:
    contract_holding = contract_current_month

if contract_holding: # 这里还需要判断？什么情况没有？
    last_trade_date = get_asset(contract_holding).last_trade_date

# 当月合约离交割日只有 3 天
days_to_expire = (last_trade_date - context.current_date).days
if days_to_expire == 3:

```

```

        log.info(u'距离 %s 到期， 还有 %s 天' % (contract_holding,
days_to_expire))

        contract_next_month = context.get_symbol('IFL1')

        futures_position = future_account.get_position(contract_holding)

        if futures_position:

            current_holding = futures_position.short_amount

            log.info(u'移仓换月。【平仓旧合约: %s, 开仓新合约: %s,
手数: %s】' % (contract_holding, contract_next_month, int(current_holding)))

            if current_holding == 0: # 注意这里的判断

                return

        future_account.order(contract_holding,current_holding,'close')

        future_account.order(contract_next_month,-
1*current_holding,'open')

        context.contract_holding = contract_next_month

        return

    stock_position = stock_account.get_positions()

    # 有多头股票仓位，使用期货进行空头对冲

    if stock_position:

        stock_positions_value      =      stock_account.portfolio_value      -
stock_account.cash # 当前股票多头市值

        futures_position = future_account.get_position(contract_holding)

        if not futures_position: # 没有空头持仓， 则建仓进行对冲

            contract_current_month = context.get_symbol('IFL0')

            multiplier = get_asset(contract_current_month).multiplier

```

```

        futures_price = context.current_price(contract_current_month)
        total_hedging_amount = int(stock_positions_value /
futures_price / multiplier)

        future_account.order(contract_current_month,-
1*total_hedging_amount,'open')

        context.contract_holding = contract_current_month
    else: # 已经有空头持仓， 则判断是否需要调仓

        contract_holding = context.contract_holding
        contract_current_month = context.get_symbol('IFL0')
        futures_price = context.current_price(contract_current_month)
        multiplier = get_asset(contract_holding).multiplier

        # 计算当前对冲需要的期货手数
        total_hedging_amount = int(stock_positions_value /
futures_price / multiplier)

        hedging_amount_diff = total_hedging_amount -
futures_position.short_amount

        # 调仓阈值，可以适当放大，防止反复调仓
        threshold = 2

        if hedging_amount_diff >= threshold:

            future_account.order(contract_holding,-
1*int(hedging_amount_diff),'open')

        elif hedging_amount_diff <= -threshold:

            log.info(u'hedging_amount_diff:%s abs:%s' %
(hedging_amount_diff, abs(hedging_amount_diff)))

        future_account.order(contract_holding,int(abs(hedging_amount_diff)),"close")

```

传统资本市场的套利主要是依靠市场一时的失灵或短暂的偏离，机理简单。

在这种情况下，大部分的投资者的操作模式会变得越来越趋同，由于不断增多的套利机构，套利的机会将会慢慢的消失，Alpha 套利策略是建立在投资者对时间和股票的选择，主动对资产进行配置的能力上，其能动性更加的强，获利的机会更多。

第三节 回测结果



通过对回测的结果进行分析可以发现，从整体上看，策略的收益曲线非常平稳。从2020年1月1日至2020年6月1日，受到新冠疫情的影响，策略的年化收益率仅有-5.3%，但基准的年化收益率为-7.6%，策略年化超额收益为2.3%。还可以发现，策略的贝塔值仅0.04，表明在发生新冠疫情这样的重大公共卫生事件时，该策略的对冲效果非常好，基本上对冲了市场风险，达到了防范重大风险的目的。

第四章 研究结论与建议

第一节 研究结论

本文从投资者实际操作的角度出发,首先介绍了我国金融市场的理论概述和量化投资选股模型的主要类别。经过分析后利用上证市场从2020年1月到2020年6月期间股票的财务数据及行情数据建立了多因子选股模型,并对模型的因子进行了选取。在因子选取方面,我们选择了对上市公司的财务因子从低到高的顺序进行赋权,对于单个财务因子的得分赋予不同的权重来求得总分。

第二节 相关建议

总的来说本文在上证A股市场所建立的多因子选股模型比较稳健,能稳定战胜突发风险。但是本文所建立的多因子选股模型并非一劳永逸,在因子选取、有效因子赋权以及选股方法上需要不断完善。随着时间的推移和市场变化,有些选股因子不再适合多因子选股模型,加入其他一些因子会使得模型表现更好,因此我们需要结合中国证券市场未来的新特点构建新的选股因子,增加对股票收益率解释力度,使得选股模型包含更多的信息,进一步提高多因子选股模型的选股能力。

多因子模型在中国股票市场表现较好,但是多因子模型也有它自己的局限性:首先模型收益并不稳定。其次在实务中,如果没有有效的对冲方法对风险进行控制,策略的生存性是值得怀疑的。另外,目前对于多因子的模型的研究对象还比较单一。大部分券商的研究主要集中在公司的财报数据上,对于市场中的行业分析以及宏观信息的研究并没有纳入模型的考虑范围中。如果能够考虑市场动态信息,将行业配置信息纳入模型中,那么多因子模型的有效性将得到更大的增强。

第三节 收获与体会

随着大数据时代的到来,以及数据挖掘算法的不断创新。相信在不久的将来,多因子模型将在中国量化领域占有更为重要的地位。研究多因子模型对于投资者了解市场,明确市场动态和方向,把握市场脉搏都具有十分重要的意义。从目前业界的研报中,我们也发现,业界在对于目前的量化投资重实践轻理论的风气还是比较普遍。由于缺乏有效的理论作为基础,部分量化投资策略在实盘中的收益难以保持长久。另外,提高模型的有效性和生存时间也是每一个交易策略开发人

员共同需要面对的问题，掌握的科学的思考方式和研究方法对于研究量化投资具有十分重要的意义。

参考文献

- [1]丁鹏. 量化投资亦能成为稳定市场的利器[J]. 中国经济周刊, 2015(31): 80-81.
- [2]范龙振, 王海涛. 上海股票市场股票收益率因素研究 [J]. 管理科学学报 2003 (1) 60-67.
- [3]黄吉平. 大数据时代量化投资: 功能、挑战与解决路径[J]. 金融纵横, 2015 (5): 10-15.
- [5]巨红岩, 李俊林, 董安强. 股票资金流强度的实证研究及评测 [J]. 科技和产业, 2015 (11): 133-136.
- [8]丁鹏. 量化投资策略与技术[M]. 北京电子工业出版社, 2012: 25-111.
- [9]陈德华, 孙成涛, 石建民. 证券市场多因素模型及其在沪深 A 股市场的应用 [J]. 生产力研究, 2009 (21): 112-115.