

| 西南财经大学中国金融研究中心金融安全研究系列专著 |

量化投资四维逻辑

何诚颖 等◎著

DISCUSSION ON FOUR ASPECTS OF
QUANTITATIVE INVESTMENT

如何利用量化选股方法获取超额收益？
如何利用量化择时方法预测未来走势？
如何利用机器学习方法提高策略效果？
如何利用资产配置方法谋求稳步增长？

量化投资四维变革时代已经到来



中国财经出版传媒集团
中国财政经济出版社

| 西南财经大学中国金融研究中心金融安全研究系列专著 |



量化投资四维逻辑

何诚颖 等◎著

DISCUSSION ON FOUR ASPECTS OF
QUANTITATIVE INVESTMENT



中国财经出版传媒集团
中国财政经济出版社

图书在版编目 (CIP) 数据

量化投资四维逻辑 / 何诚颖等著. —北京: 中国财政经济出版社, 2018. 10
(西南财经大学中国金融研究中心金融安全研究系列专著)

ISBN 978 - 7 - 5095 - 8534 - 4

I. ①量… II. ①何… III. ①投资 - 研究 IV. ①F830.59

中国版本图书馆 CIP 数据核字 (2018) 第 216428 号

责任编辑: 郁东敏

责任校对: 胡永立

封面设计: 田 晗

中国财政经济出版社 出版

URL: <http://www.cfeph.cn>

E-mail: cfeph@cfeph.cn

(版权所有 翻印必究)

社址: 北京市海淀区阜成路甲 28 号 邮政编码: 100142

营销中心电话: 010 - 88191537 北京财经书店电话: 64033436 84041336

北京财经印刷厂印刷 各地新华书店经销

787 × 1092 毫米 16 开 31.5 印张 555 000 字

2018 年 10 月第 1 版 2018 年 10 月北京第 1 次印刷

定价: 88.00 元

ISBN 978 - 7 - 5095 - 8534 - 4

(图书出现印装问题, 本社负责调换)

本社质量投诉电话: 010 - 88190744

打击盗版举报电话: 010 - 88191661 QQ: 2242791300

序：量化投资四维逻辑

—

2011年，我在沃顿商学院访问的时候，量化投资在美国的发展已经到了相当成熟的阶段。量化投资给美国股市带来的变化是深刻的。量化交易利用公开数据，借助数学、统计学、信息技术的量化投资方法来管理投资组合，其组合构建注重的是对宏观数据、市场行为、企业财务数据、交易数据进行分析，利用数据挖掘技术、统计技术、计算方法等处理数据，得到最优的投资组合和投资机会。从方法论来看，量化投资大大降低了投资者对市场的人为干预。当时我就强烈感觉到，作为一种能够摆脱人性弱点并且相当稳定的投资方法，量化投资在A股也将会大有所为。回国后，我马上组织了一批具有丰富市场经验的博士后研究人员开展关于量化投资的研究，对量化投资这个舶来品的本土化率先做出了尝试，并出版了《解读量化投资之秘》一书。可以说，这是一本国内较早系统研究量化投资的专业书籍，一定程度上填补了本土量化投资策略开发研究的空白。

从国外量化投资整体发展来看，大概可以划分为三个阶段：第一阶段从1970年至1977年，1971年巴克莱国际投资管理公司发行了世界上第一只被动量化基金，标志着量化投资的开始，这一阶段量化投资处于崛起阶段。第二阶段从1977年到1995年，由于数据和计算机技术等限制，量化投资进入一个缓慢的发展阶段，但在行业调整过程中华尔街已开始对此模式极为重视。第三阶段从1995年至今，计算机技术的高速发展，特别是网络技术的突飞猛进，使得大量即时金融数据能够有效供给，量化投资进入高速发展时代。

国内量化投资的起步比美国市场晚了30多年。由于中、美两国投资者结构和市场环境、监管政策的差异，简单对美国策略进行套用和模仿并不可行。几年来，我带领的团队一直在从事量化投资方面的研究，最大的感受是：在量化投资这个领域，做得好的人往往守口如瓶，做得不太好的人却到处高谈阔论。这给新入行的量化投资人带来极大的困扰。其实，中国的量化投资人不应亦步亦趋，而是应该以成熟的美国量化投资策略为基础，结合中国股票市场的投资者结构和市场特征进行本土化改造。于是，我希望率先作出尝试，通过择时、选股、机器学习和资产配置四大维度，构建本土化的量化投资策略，为广大量化投资人建立一个全新的分析框架。

二

近年来，量化投资在美国的发展可谓蒸蒸日上，市场规模和份额不断扩大。2013年，美国市场量化交易占美股总交易量的14%，2017年这一比率就上升到了27%。^①2017年第一季度美国市场量化策略对冲基金的管理规模为9320亿美元，占全部对冲基金管理规模的30%。反观中国市场，2017年第一季度股票量化基金1600亿元人民币，占基金总量的1.7%，占所有股票基金总量的5.7%。可见，中国市场上量化基金占比较低。这恰恰说明量化投资在中国的投资体系中还不是主流，也预示着中国量化投资未来发展潜力可期。

为什么量化投资在海外如此受欢迎呢？一方面，在过去的很长时间里，量化投资持续稳定的收益能力得到了投资者的认可；另一方面，海外投资者的投资理念更为成熟。为了分散风险和投资组合多元化，即使是在量化投资表现不好的年份，他们依然愿意使用。但是量化投资在中国市场的表现就远不如美国。在中国，量化基金还是比较边缘化的产品。私募基金的主流依然是主动管理的纯股票多头产品。2017年以来，量化

^① 《量化投资：华尔街爆款到中国为何不红了？》，《中国基金报》，2017年9月5日。

基金资金募集困难，赚钱的效应也在下降，一些小型的量化私募基金面临挑战。很多人疑惑，量化投资作为一种新兴的投资策略，在中国这个新兴的资本市场上是否会水土不服？

量化投资在国内市场不太受欢迎的主要原因有三点：一是国内市场历史有限，可以用来提供的数据不充足，行业数据的整理也不是很完善，质量也不够高，所以对于量化投资来说缺少足够的支持。二是国内投资者对量化投资的认识不到位，一些相关法律法规也不利于量化投资在我国的发展。三是美国市场的投资以机构为主，对于机构的博弈，量化、系统化的方法持续性更强，先进的 IT 技术、数据挖掘技术会不断应用于全球市场。

即使目前量化投资在国内表现不佳，但不可否认的是，量化投资的优势是十分明显的。量化投资把人的情绪排除于投资进程之外，使得整个投资进程完全按照预先设定的程序进行操作，这是量化投资极具吸引力的一个本质原因。同时，量化投资打破了传统投资在范围上的局限，信息爆炸式的接收和处理信息速度成为量化投资的优势；同时，其借助计算机技术提高效率，加快了策略更新。此外，由于策略无上限，投资基金就可以凭借量化投资管理更大规模的资金。

国内的量化投资人对量化投资的态度整体上也是乐观的。首先，国内资本市场一直推出新的衍生品；其次，市场上新的业态也逐渐增多，比如智能投顾公司、数据公司和金融科技公司等，这些都是量化投资的有机组成部分；最后，中国市场并不缺乏勇于承担风险和尝试量化策略的资金，只是大部分投资者对于量化投资的认识还有待加深。

三

进入新时代，随着供给侧改革的逐步推进，A股市场也正在发生深刻变化。并且，随着人工智能、金融科技、深度学习等新技术的出现和发展，量化投资又被赋予了新的内涵。从市场环境的变化来看，A股的市场化进程逐步加快，具体包括 IPO 常态化、投资者机构化、严格监管

持续化、内外投资一体化。A股的深刻变化要求量化投资也要随之改变。市场结构及供需关系的变化，改变了超额收益的来源，过去市值因子、动量反转、波动率、换手率等长期有效的因子未来不一定继续有效。我们认为，随着中国量化市场的成熟，传统基本面因子的有效性将越来越高，这一趋势是与A股的市场化进程一致的。可以说，量化投资的中国化，是量化与基本面紧密结合的一个过程。

传统的基本面投资者和量化投资者的分析和投资方法有着明显区别。基本面投资者是筛选公司后进行深入研究，少而精，偏艺术。量化投资者是挖掘市场规律、有广度，偏科学。但是在2008年全球金融危机之后，各国金融监管越来越多，纯量化策略仅仅在数理金融的抽象理论上深入是不够的，更需要结合现实业务和宏观环境。基于此量化基金把量化和基本面分析有效结合起来，创造出一个跨界的新流派——量化基本面分析法（Quantamental）。这种流派包含了基本面投资人自下而上的选股技巧，同时借助计算机技术和大数据系统完善模型。

与市场因子相比，基本面量化最大的差异是聚焦于企业的长期收入和利润。在基本面量化价值的判断里，首先要看清行业的发展趋势，建立一个宏观与行业的分析框架；其次是分析行业内公司的竞争格局。量化投资的竞争将由目前模型的竞争转变为对基本面的分析竞争。对此，建立基本面的研究框架，需要对公司盈利进行预测，进而对公司的内在价值进行判断，方能构造量化基本面策略。另外，对基本面的研究不能过于关注短期，而要构造基于基本面信息的“中长”持有期的选股策略，要善于在行业发展的高潮和低谷的时候做反向判断，做逆向投资。

大数据的发展使得价值投资的核心理念可以进行量化分析。近年来，在社会科学领域尤其是投资领域，人们产生、搜集和整理数据的类型和数量呈现出几何级数增长，研究人员建立数学模型发现规律的能力大大提升。过去，基金经理如果要调研一家公司的运营情况，常用做法是到这家公司门口观察每天卡车的进出量，通过计算卡车的进出量来评估公司的经营活跃度。而现在可使用低轨卫星监测工厂门口的卡车进出量的大数据推测出公司的营运状况。如果这个季度该公司的卡车量上升很多、

活跃度很高，我们就可以预测公司这个季度业绩比较好，从而预测了它的基本面。

格雷厄姆曾说：“市场短期是投票器，长期是称重机。”在中国价值投资理念已经开始起步、市场格局向机构主导的成熟转型过程中，经过市场检验、契合市场风格的量化价值产品正生逢其时。

四

随着金融科技（Fintech）时代的到来，中国金融业正经历着一场新的变革，并且这场变革不断升级。中国的金融科技行业由原来注重流量和模式的1.0时代，升级为以人工智能技术为主导、以数据为驱动力的2.0时代。量化投资借力人工智能技术，运用现代统计学和数学的方法，从大量的历史数据中寻找并搭建获得超额收益的投资策略，服务于个人投资者和机构，也成为金融科技新时代的领军者。

金融科技的迅猛发展赋予量化投资新的技术，量化投资中国化需要与金融科技深入结合，将大数据、云计算、人工智能等前沿技术运用到量化投资中来，开创量化投资研究新范式。我们已认识到这种发展趋势，尝试进行基于大数据的人工智能量化策略研究，希望从三个方面进行突破。

第一，利用人工智能，将非结构化大数据运用到量化投资策略上。非结构化式数据蕴含的信息非常丰富，比如文本的、图像的、声音的等等，以往受限于技术，很难或者需要花费巨大成本变成结构化数据，而人工智能的方法目前允许低成本非结构化数据处理成结构化数据来进行交易。

第二，借助人工智能的算法突破线性建模的限制。以往简单的线性关系建模已经不能准确刻画数据间的关系，但受制于算法，人们常常作线性假设。随着人工智能中非线性算法的不断完善，使用非线性人工智能算法构建量化策略已经成为现实，并且更有价值。

第三，加强对强化学习算法的研究。人工智能自主强化学习能力如

阿尔法元算法为量化策略提供了无限可能。人工智能算法的一项重要突破是在没有任何先验知识的前提下，通过完全的自学达到超越人类的境地。阿尔法元算法无须对数据标注，通过自动化地处理数据产生认知进行决策，既能突破人类自身能力的局限，避开策略局部收敛问题，又能够解决大数据中的信息噪声问题。

五

中国 A 股市场相对于海外成熟市场还比较年轻，在量化投资所必需的一些基础设施上还有所欠缺。比如中国市场的金融衍生品不太发达，而量化投资需要有很多金融衍生品去对冲这些风险。此外，中国股市起步较晚，又经历了几次重大变革，可借鉴的数据非常有限。成熟量化交易模型往往需要在长期稳定的历史数据基础上进行归纳验证，但国内证券市场可供借鉴的历史数据十分有限。对中国量化投资而言，相比于纯粹的数据挖掘，建立模型更需要根据基本面和投资者行为搭建逻辑性框架，建立针对国内市场的多领域、多策略的投研体系；对每个策略的收益和风险系数进行把控，来权衡各类资产配置，然后根据市场变化产生的数据进行模型调整，从而开发出有中国特色的量化策略。

必须强调的是，中国量化投资的发展一定要加大力度培养相关人才。量化投研团队除了要有量化建模的能力之外，还要在经济金融、投资理论方面都有较深造诣，将社会科学和自然科学技术进行深刻理解，充分掌握两门学科并有机结合，并实现综合运用。国内量化投资行业从业人员相对年轻，对宏观经济、计算机中较复杂的技术、灵活运用量化产品上，还需要进一步提高。国外量化投资研究者多为数学、物理、计算机专业博士，他们缺乏金融专业知识。因此，我国量化投资本土化团队不仅要有研究经济金融和投资理论的人才，也要有自然科学方面的专业人才，从而实现团队本土化。

当务之急是将从美国引进来的量化投资中国化，找到中国化的量化投资策略，提高回报率，解决投资界面临的现实难题，才是我们应该探

索的。多年前我们就从股市资金流入手，希望结合大数据技术和方法改进量化投资策略，顺着择股、择时、机器学习和资产配置的投资逻辑，构建独特的四维投资之道，从而找到突破口，探寻一条符合中国特色的量化投资之路。

本书是对量化投资中国化的又一次尝试，希望可以起到抛砖引玉的作用，引领中国量化投资的发展进入一个新阶段。

何诚颖

2018年9月于深圳

目 录

第一篇 量化选股策略篇

1. 量化选股策略概述	(3)
1.1 量化选股的收益来源	(3)
1.2 量化选股策略的框架	(4)
2. 单指标策略	(5)
2.1 单指标/因子策略方法概述	(5)
2.2 资金流向策略	(15)
2.3 动量与反转策略	(74)
3. 多因子模型及策略	(128)
3.1 多因子策略方法概述	(128)
3.2 基于风险指标的多因子模型投资策略分析	(136)
3.3 基于指标被动筛选的多因子模型策略	(143)
3.4 多因子模型因子加权方法研究	(154)
4. 风险模型及其应用	(166)
4.1 风险模型概述	(166)
4.2 中国 A 股市场风险模型的行业因子优化	(172)
4.3 中国 A 股市场风险模型的应用	(186)
5. 投资组合构建	(195)
5.1 投资组合构建方法概述	(195)

5.2	基于规模暴露控制的投资组合构建	(199)
5.3	基于稳健优化的投资组合构建	(216)
6.	绩效归因与评价	(225)
6.1	绩效归因与评价概述	(225)
6.2	多因子策略的绩效评价	(228)

第二篇 量化择时策略篇

7.	量化择时策略概述	(243)
7.1	量化择时的理论基础——有效市场假说	(243)
7.2	量化择时策略的类型	(244)
8.	动量择时模型	(246)
8.1	均线择时系统：简单易行，最为常用的趋势交易	(246)
8.2	LLT 模型：高阶过滤，降低信号延迟	(263)
8.3	基于参数优化的 Dual - Thrust 交易策略：过滤震荡行情抓趋势	(273)
8.4	指数收益率奇数阶矩预测：放大价格波动抓趋势	(282)
8.5	海龟交易系统：完整的交易系统	(290)
9.	反转择时模型	(296)
9.1	DeMark Combo 策略：累计能量，逆市操作	(296)
9.2	抛物线拟合策略：寻找拐点，右侧交易	(305)
9.3	对数周期幂律模型及其基本应用：阶段性“顶”和“底”的预测	(312)
9.4	对数周期幂律模型的概率预测	(325)
9.5	对数周期幂律模型的相关检验	(334)
9.6	对数周期幂律模型的窗口初始点选择	(342)

第三篇 机器学习量化策略

10.	机器学习算法简介	(357)
-----	----------	-------

10.1	机器学习的定义和学科定位	(358)
10.2	机器学习方法的核心概念	(359)
10.3	机器学习方法的分类	(361)
11.	机器学习策略开发流程	(363)
11.1	机器学习算法的应用	(363)
11.2	机器学习策略的流程：以 GB 决策树为例	(365)
11.3	机器学习策略的预测能力	(366)
12.	GB 决策树模型	(368)
12.1	决策树模型介绍	(368)
12.2	决策树模型选股策略的参数选择与拟合效果	(371)
12.3	基于 GB 决策树的市值策略及其比较	(378)
12.4	结论	(383)
13.	长短期记忆神经网络模型	(385)
13.1	LSTM 的机理与优势	(385)
13.2	长短期记忆模型量化策略	(392)
13.3	结论	(396)
14.	随机森林模型	(398)
14.1	模型介绍	(398)
14.2	随机森林量化投资策略	(403)
14.3	结论	(408)
15.	简单决策树模型	(410)
15.1	决策树的概念及其实例	(410)
15.2	基于历史涨跌来预测未来涨跌	(411)
15.3	基于技术指标预测股指未来涨跌	(415)
15.4	多空策略及表现	(418)
15.5	结论	(422)
16.	临近取样模型	(424)
16.1	临近取样算法简介	(424)

16.2	策略描述	(425)
16.3	策略表现	(426)
16.4	结论	(431)

第四篇 资产配置篇

17.	资产配置概述	(435)
17.1	投资组合管理过程	(435)
17.2	资产配置决策	(436)
17.3	资产配置策略	(436)
17.4	结论与展望	(439)
18.	基于风险平价方法的资产配置策略	(440)
18.1	风险平价策略的数学模型	(443)
18.2	基于风险平价策略的股债配置	(445)
18.3	结论与展望	(449)
19.	基于风险预算方法的资产配置策略	(451)
19.1	风险预算配置方法	(452)
19.2	基于风险预算方法的股债配置实证分析	(454)
19.3	结论与展望	(461)
20.	基于风险平价方法的行业配置策略	(462)
20.1	行业组合配置方法	(462)
20.2	基于风险平价方法的 A 股行业配置实证分析	(465)
20.3	结论与展望	(474)
21.	基于 Black - Litterman 模型的行业配置策略	(476)
21.1	Black - Litterman 模型	(477)
21.2	Black - Litterman 模型在我国股市行业配置中的应用	(482)
21.3	结论与展望	(487)
后	记	(489)

第一篇

量化选股策略篇

1.

量化选股策略概述

1.1 量化选股的收益来源

市场并不是充分有效的，但却是激烈竞争的。市场及其投资者在数据和自身经验中吸取知识，使得过去的市场异象变成当前的随机噪音。利益驱动的市场机制使得投资者不断寻找下一个信号和模型来战胜市场。一旦新的信号和模型被证明有用，其收益最终都将会被套走，市场将变得越来越有效。因此，对于量化管理者而言，永远站在研究的第一线是唯一能持续获得超额收益率的方法。

真正的 alpha 收益很稀缺，持续战胜市场也很困难。市场可能并不有效，但绝对是充分竞争的。传统的因子不再能增加持续的 alpha 收益。如果一个因子确实能产生持续的超额收益，那么越来越多的投资者会投资这一信号直到其失去预测能力。这一点同样适用于模型技术与数据源，可以通过寻找新的指标和因子的构建技巧、收益风险模型与投资组合技术来获取超额收益。

比如预测股票收益的多因子模型，提高其效果的途径并不多，最好的方法是选择一些新的、使用较少的数据源，但毕竟这样的数据源有限。因此，在使用传统数据方法上有更深入的研究。传统的数据与因子不大可能产生持续的 alpha 收益。但如果从不同的角度考察现有的因子，则有可能产生更好、相关性更小的因子。此外，模型构建技术也是 alpha 收益的一个重要来源，但并没有受到同等的重视。等权