

Python量化交易

张杨飞 / 编著

求人不如求己，用心打造自己的财富之井

本书将手把手地教会你用Python+vn.py流行组合

搭建属于自己的量化交易系统

且着重于量化交易策略的解析、应用与回测

Python量化交易

张杨飞 / 编著



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书本着能让新手快速上手量化交易的原则，循序渐进地讲解了量化交易入门所需要的知识，以及大量的开发技巧与交易技巧，具有很强的实用性。vn.py 是机构级别的量化交易软件，掌握 vn.py 框架原理并且熟练运用，有利于新手快速搭建属于自己的量化交易系统。Python 语言有非常强大的数据分析库，对于交易策略的研发具有天然优势，且其易学性也深受初学者喜爱。本书即以 Python+vn.py 这一流行组合写作，从量化交易的起源及其发展进程入手，在简单介绍 Python 量化编程基础，以及详细解析 vn.py 架构之后，全面地介绍了 CTA 策略、海龟策略，以及新策略的开发流程。

相对其他量化交易方面的书，本书不再讲述 Python 语言编程的大量细节，而将笔墨着重放在对量化交易策略的解析、应用与回测之上，这才是新手真正需要学习和实践的地方。本书适合所有对量化交易感兴趣的人员阅读，也适合相关院校和培训机构作为量化交易系统课程的教材。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目 (CIP) 数据

Python 量化交易 / 张杨飞编著. —北京: 电子工业出版社, 2019.5
ISBN 978-7-121-36140-1

I. ①P… II. ①张… III. ①股票交易—应用软件 IV. ①F830.91

中国版本图书馆 CIP 数据核字(2019)第 046455 号

责任编辑: 孙学瑛

印 刷: 北京京科印刷有限公司

装 订: 北京京科印刷有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编: 100036

开 本: 787×980 1/16 印张: 26 字数: 414 千字

版 次: 2019 年 5 月第 1 版

印 次: 2019 年 5 月第 1 次印刷

定 价: 99.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及订购电话：(010) 88254888, 88258888。

质量投诉请发邮件至 zltts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式: 010-51260888-819, faq@phei.com.cn。

前言

在证券交易领域中，量化交易脱胎于传统的主观交易：把投资者的交易理念、交易策略固化成计算机程序，通过算法快速自动下单，有效地防止投资者自身情绪的干扰，让其把精力放在研发交易策略上。量化交易的另一个优点是其成功的方法有迹可循，因为量化交易的过程是，运用现代统计学理论对历史数据进行数据分析，构建数学模型来预测市场未来价格的变化，然后通过计算机语言表达出来，从而实现自动交易。

vn.py 是机构级别的量化交易软件，掌握 vn.py 框架原理并且熟练使用，有利于新手快速入门量化交易，搭建属于自己的量化交易系统，也可以在机构中找到与量化岗位相关的工作。

为什么写作本书

本书使用的编程语言是 Python。尽管市面上关于 Python 量化交易方面的书籍不少，但是大部分着重讲述 Python 编程基础，而且主要是在股票交易中的应用，在期货市场的应用少之又少。就实践的层面来看，股票量化交易对于入门者几乎是不可能实现的，尽管 2019 年重新开放程序化交易 API，但是其资金门槛高达 5 亿元。

期货市场上程序化交易接口却无资金门槛，而且量化交易应用发展得比较成熟，这才是新手可以去学习和实践的地方。故本书的写作定位是交易策略在期货领域的应用和开发，力求填补这方面的空白。

本书特色

本书尽量以初学者的角度来讲述量化交易的内容，逐步填平量化交易入门要踩的“坑”，力求让读者快速熟悉这方面的知识，能够独立开发交易策略并且尝试进行仿真

交易。如果在 SinNow 仿真交易平台能够赢利，那么就可以上实盘去“跑”了。

本书的另一个特色是使用机构级别的开源交易软件：vn.py。机构级别软件对应的使用群体是做量化交易的机构投资者，如私募基金、证券自营，以及资管和期货资管等。这类软件虽然上手困难，但是熟练掌握后能更有效地深耕于量化交易领域，并且也有利于初学者入门量化交易。

本书主要内容

本书共包括 7 章，每章的主要内容如下。

第 1 章“量化交易速览”首先从狭义和广义两个方面介绍了“量化交易”的概念，然后介绍开拓出这个领域的先驱们的故事，接着讲述量化投资在美国与中国的历史发展进程，最后简单地介绍国内常用的量化交易策略及宽客这个专门从事量化交易的职业。

第 2 章“Python 量化编程基础”介绍了将 Python 作为量化交易入门语言的理由，讲解了 Python 的基础概念，以及常用的数据分析库 NumPy 与 Pandas、机器学习库 scikit-learn，最后讲解绘图库 Matplotlib 的基本用法。

第 3 章“vn.py 入门”介绍了 vn.py 交易系统的概况、安装步骤、主交易界面的功能，具体讲述 vn.py 应用框架的结构，分别是底层接口、中层引擎及上层应用，最后对这 3 层结构的原理做一个具体说明。

第 4 章“在 vn.py 中实现 CTA 策略”介绍 vn.py 提供的数据解决方案，用于生成具体 CTA 策略的相关支出模块，如 K 线生成、K 线管理和策略模板，最后讲述回测和优化模块。

第 5 章“经典 CTA 策略”主要介绍 vn.py 官方提供的经典 CTA 策略，包括策略原理、代码解释、策略回测和参数优化。

第 6 章“海龟策略本地化实证”首先介绍海龟交易策略的起源、关键要素，然后解析 vn.py 下海龟策略的代码，通过交叉检验与筛选品种构成投资组合，最后基于构

建好的海龟组合对策略的各个关键要素进行研究。

第7章“新策略实战”首先介绍了开发新策略的流程，然后是搭建投资组合，并进行策略回测为实战做好准备，最后介绍在真实交易情况下接触的3套系统，并且分析策略回测与实战中结果不同的成因。

致谢

我首先要感谢“猴子聊人物”创始人，他的数据分析的课程让我快速上手 Python 语言；然后是“用 Python 的交易员”陈晓优先生，我也是受益于其知乎 Live 上对量化交易的推广才从传统的金融转到该领域的。

我还要真诚地感谢电子工业出版社优秀的 IT 编辑孙学瑛女士和电子工业出版社对本书的重视，以及他们为本书出版所做的一切。

读者服务

轻松注册成为博文视点社区用户 (www.broadview.com.cn)，扫码直达本书页面。

- **提交勘误**：您对书中内容的修改意见可在 [提交勘误](#) 处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。
- **交流互动**：在页面下方 [读者评论](#) 处留下您的疑问或观点，与我们和其他读者一同学习交流。

页面入口：<http://www.broadview.com.cn/36140>

目 录

第 1 章 量化交易速览.....	1
1.1 为何选择量化交易.....	1
1.1.1 量化交易的概念.....	1
1.1.2 主观交易与量化交易.....	2
1.2 量化交易的先驱们.....	5
1.2.1 朱尔斯·雷格纳特.....	5
1.2.2 爱德华·索普.....	6
1.2.3 托马斯·彼得菲.....	9
1.2.4 詹姆斯·西蒙斯.....	14
1.3 美国量化投资的发展历史.....	17
1.3.1 兴起阶段（1970—1990 年）.....	17
1.3.2 快速发展阶段（1990—2000 年）.....	18
1.3.3 稳步增长阶段（2000 年至今）.....	19
1.4 中国量化投资的发展历史.....	20
1.4.1 ETF 套利时代（2010 年以前）.....	20
1.4.2 多因子 Alpha 和高频交易称雄时代（2010—2015 年）.....	21
1.4.3 多元化投资时代（2016 年至今）.....	23
1.5 国内常用的量化交易策略.....	24
1.5.1 期货 CTA 策略.....	24
1.5.2 股票 Alpha 策略.....	32
1.5.3 期权波动率套利策略.....	41
1.5.4 高频交易策略.....	45
1.6 宽客.....	48
1.7 宽客的两大阵形：P 宗与 Q 宗.....	51

1.8	宽客的 3 种职能分类.....	52
1.8.1	量化 IT 工程师.....	52
1.8.2	量化研究员.....	53
1.8.3	量化交易员.....	54
1.9	宽客的四大派系.....	55
1.9.1	券商资管.....	55
1.9.2	公募基金.....	56
1.9.3	私募基金.....	57
1.9.4	期货市场.....	57
第 2 章	Python 量化编程基础.....	59
2.1	Python 运行环境搭建.....	60
2.1.1	安装 Anaconda2-5.0.0 (32 位).....	61
2.1.2	设置 Anancoda 环境.....	62
2.1.3	创建共享环境.....	64
2.1.4	列出共享环境.....	64
2.1.5	安装 Jupyter Notebook.....	65
2.2	数据.....	66
2.2.1	字符串.....	66
2.2.2	数字.....	68
2.2.3	容器.....	68
2.2.4	布尔值.....	73
2.2.5	空值.....	74
2.3	函数.....	74
2.3.1	自定义函数.....	74
2.3.2	第三方库的函数.....	75
2.4	条件判断.....	75
2.5	循环.....	77
2.6	类和实例.....	79
2.6.1	定义学生父类.....	79
2.6.2	定义父类实例.....	81

2.6.3	定义团体子类.....	82
2.6.4	定义子类实例.....	83
2.7	NumPy 与 Pandas	84
2.7.1	一维数组	85
2.7.2	二维数组	88
2.8	scikit-learn 机器学习库.....	93
2.8.1	机器学习的步骤.....	93
2.8.2	线性回归	94
2.8.3	逻辑回归	101
2.9	Matplotlib 绘图库.....	104
2.9.1	用列表绘制线条.....	105
2.9.2	用数组绘图.....	106
2.9.3	多个图的绘制.....	109
第 3 章	vn.py 入门.....	111
3.1	vn.py 介绍	111
3.2	搭建 vn.py 运行环境.....	115
3.2.1	安装 Visual Studio 2013 社区版（特定版本）	115
3.2.2	安装代码编辑器工具：Sublime Text.....	116
3.2.3	安装 Wing IDE	117
3.2.4	安装 MongoDB 数据库	117
3.2.5	安装 Robo 3T.....	120
3.2.6	安装 vn.py.....	121
3.2.7	更新 vn.py.....	123
3.3	VnTrader 界面功能介绍.....	124
3.3.1	连接 CTP	124
3.3.2	界面说明	125
3.4	vn.py 架构	126
3.4.1	底层接口	127
3.4.2	中层引擎	128
3.4.3	上层应用	129

3.5	底层接口	130
3.5.1	CTP API 的工作原理	130
3.5.2	CTP API 的 Python 封装设计	135
3.5.3	CTP API 对接中层引擎原理	137
3.6	事件引擎	140
3.6.1	时间驱动	140
3.6.2	事件驱动	141
3.6.3	事件引擎工作流程	142
3.6.4	事件引擎结构	143
3.7	上层应用	145
3.7.1	PyQt 介绍	145
3.7.2	GUI 组件构成	146
第 4 章	在 vn.py 中实现 CTA 策略	149
4.1	数据解决方案	149
4.1.1	CSV 加载模块	149
4.1.2	开发新的 CSV 导入模块	154
4.1.3	数据下载模块	157
4.2	K 线生成模块	159
4.2.1	1 分钟 K 线合成	160
4.2.2	X 分钟 K 线合成	163
4.3	K 线管理模块	164
4.3.1	初始化参数	164
4.3.2	生成时间序列	165
4.3.3	定义属性函数	166
4.3.4	生成计算指标	167
4.4	CTA 策略模块	169
4.4.1	定义成员变量	170
4.4.2	构造函数	171
4.4.3	回调函数	172
4.4.4	主动函数	173

4.5	策略回测模块.....	176
4.5.1	CTA 回测引擎.....	176
4.5.2	参数优化设置.....	180
4.5.3	调用回测和优化模块.....	180
第 5 章	经典 CTA 策略.....	187
5.1	双均线策略.....	187
5.1.1	策略原理.....	187
5.1.2	向量回测.....	188
5.1.3	vn.py 回测.....	193
5.2	Dual Thrust 策略.....	202
5.2.1	策略原理.....	202
5.2.2	策略代码解析.....	203
5.2.3	策略回测.....	208
5.2.4	策略优化.....	210
5.2.5	滚动回测.....	213
5.3	AtrRsi 策略.....	214
5.3.1	ATR 指标.....	215
5.3.2	RSI 指标.....	217
5.3.3	策略原理.....	218
5.3.4	策略代码解析.....	219
5.3.5	策略回测.....	222
5.3.6	滚动回测.....	223
5.4	金肯特纳通道策略.....	225
5.4.1	策略原理.....	225
5.4.2	策略代码解析.....	226
5.4.3	策略回测.....	231
5.4.4	滚动回测.....	231
5.5	布林带通道策略.....	233
5.5.1	策略原理.....	233
5.5.2	CCI 指标.....	234

5.5.3	ATR 指标	236
5.5.4	策略回测	237
5.5.5	滚动回测	238
5.6	跨时间周期策略	240
5.6.1	策略原理	241
5.6.2	策略代码解析	241
5.6.3	策略回测	245
5.6.4	滚动回测	246
5.7	多信号组合策略	247
5.7.1	策略原理	248
5.7.2	信号生成部分	248
5.7.3	交易管理部分	253
5.7.4	多信号策略的重构	258
第 6 章	海龟策略本地化实证	261
6.1	海龟策略速览	261
6.1.1	海龟策略的故事	261
6.1.2	海龟策略的局限性	262
6.1.3	原版海龟策略	263
6.1.4	策略回测效果	268
6.2	本地化实现困境与解决方案	270
6.2.1	本地化实现困境	270
6.2.2	理想解决方案	272
6.3	vn.py 实现的海龟策略	273
6.3.1	工具准备	273
6.3.2	数据准备	274
6.3.3	海龟策略代码结构	277
6.3.4	海龟策略 6 大要素代码解析	280
6.3.5	海龟策略的回测	285
6.4	品种选择验证	287
6.4.1	原版投资组合测试	287

6.4.2	筛选品种的传统方法	288
6.4.3	构建海龟组合的难点	297
6.4.4	海龟组合筛选的解决方案	298
6.4.5	重新构建投资组合	302
6.5	长短周期信号检验	322
6.6	上一笔赢利过滤检验	324
6.7	手续费、滑点测试	326
6.8	单位头寸限制检验	326
6.9	关于海龟策略的其他研究方向	330
第 7 章	新策略实战	331
7.1	开发新的策略	331
7.1.1	策略思路	331
7.1.2	增加 AROON 函数	333
7.1.3	策略代码解析	334
7.1.4	策略回测	336
7.2	多策略的组合回测	338
7.2.1	历史表现	339
7.2.2	预测表现	342
7.2.3	回测的注意事项	342
7.3	模拟测试	349
7.3.1	策略文件目录	349
7.3.2	实盘/模拟盘配置文件	350
7.4	真实交易环境	353
7.4.1	交易环境的 3 套系统	353
7.4.2	交易环境的数据流	354
7.5	实际操作注意事项	355
7.5.1	计算错误	355
7.5.2	数据使用误差	356
7.5.3	过拟合	357
7.5.4	幸存者偏差	358

7.5.5	策略周期	359
7.5.6	动态变化的现实环境	360
7.5.7	人为干预	361
附录 A	主流交易品种	362
A.1	股票	362
A.1.1	股票的定义	362
A.1.2	股票交易所	363
A.1.3	股票竞价规则	364
A.1.4	T+1 制度	368
A.1.5	股票交易策略	370
A.2	期货	372
A.2.1	期货的定义	372
A.2.2	期货交易所	372
A.2.3	期货交易策略	375
A.3	期权	377
A.3.1	期权的定义	377
A.3.2	期权的分类	380
A.3.3	期权的影响因素	382
A.3.4	期权投资组合	384
A.3.5	期权波动率套利策略	387
A.4	外汇	388
A.4.1	外汇的定义	388
A.4.2	外汇市场的结构	390
A.4.3	外汇市场的组织形式	393
A.4.4	主要外汇交易中心	394
A.4.5	外汇交易策略	396
	参考文献	399

第 1 章

量化交易速览

本章将带领读者迅速了解量化交易。首先，从“量化交易”概念入手，通过行业先驱故事来阐述量化交易的重要性，然后分别讲述量化交易在美国与中国的历史发展进程，最后简单地介绍国内常用的量化交易策略，以及“宽客”这个职业。

1.1 为何选择量化交易

1.1.1 量化交易的概念

在回答“为何选择量化交易”这个问题前，先理解量化交易的概念。

严格地说，量化交易是运用复杂的统计学方法和数学模型，从庞大的历史数据中海选出能带来超额收益的多种“大概率”事件以找出规律、制定策略，并且能用数据模型验证、固化这些规律和策略，然后用计算机来严格、高效地执行之。

这一定义涉及统计学、金融学和计算机科学等多门学科，看起来高不可攀，但是，通俗点说，量化交易是指利用统计学、数学、计算机技术和现代的金融理论来辅助投资者更好地赢利。这些量化的方法可用于分析海量历史数据，也可用

于具体信号生成，或者控制持仓大小、进行风险控制等。

因此，我们可以惊喜地发现量化交易与传统的主观交易不再是二元对立，量化交易也包含主观交易，如比较常见的期货跨市套利策略、期权波动率套利等就属于半自动交易。它们需要交易员综合历史均值回归，以及对宏观政策的主观解读，开盘前调整好参数让计算机严格执行策略。

综上所述，量化交易是对主观交易的升华，剔除部分人为不稳定的因素，让用户专注于寻求超额收益（即 Alpha）。所以量化交易必然会成为历史发展的趋势。

1.1.2 主观交易与量化交易

量化交易总是被人们拿来与主观交易讨论，孰高孰低、众说纷纭。为了更好地分析它们之间的辩证关系，下面分别解读之。

1. 主观交易

运用主观交易的投资者需要关注国际环境及财经新闻、券商的研究报告、公司的财务报告、K 线趋势、个股新闻、价量经验、大人物的演讲、朋友圈情绪、甚至一些小道消息，并对其进行定性或者定量分析。定性分析可以是体验公司产品或者直接实地考察，定量分析则可根据自己的选股原则，用各种指标进行打分，加权汇总后，买入分数高的股票，卖出分数低的股票，以形成交易信号。在手动下单方面，注意不要记错代码并且要避免“胖手指失误”¹（fat-finger error），若单子太大还需要拆分成小单来降低成本。开仓之后要有风控意识并且严格遵循自己的止赢止损原则。

主观交易具有较强的主观能动性，即在同样的选股原则和止赢止损策略下，100 个人操盘会有 100 种不同的结果，基于大数定律和类正态分布，可以发现有一小撮人的资金曲线非常平滑和漂亮，但是绝大部分人都会低于均线水平，这里的“均线”指的是计算机严格执行交易策略所得到的资金曲线。原因往往是，理

¹ 即交易员在交易时因键盘操作失误而造成的输入错误。

性总是被情绪打败的。

20 世纪 60 年代，保罗·D·麦克莱恩（Paul D. MacLean）从生理角度提出了“三位一体的大脑”理论。此理论根据在进化史上出现的先后顺序，将人类大脑分成“爬行动物脑”“古哺乳动物脑”和“新哺乳动物脑”三大部分。每部分“脑”通过神经纤维与其他两者相连，但各自作为相对独立的系统分别运行，各司其职。“新哺乳动物脑”又称为边缘系统，它参与调解本能和情感行为，其主要作用是维持自身生存和物种延续。其中，杏仁核负责创造情绪并产生与之相关的记忆，海马结构能将短期记忆转化为长期记忆。杏仁核作为大脑的“恐惧中心”，它触发的保命技能可是顶尖的：远远看到狮子，恐惧会瞬间攫取你对身体的自主权并快速逃跑。杏仁核在漫长的人类演进过程中很好地提高了人类的生存概率，但是现在却阻碍了止赢止损策略的执行，让赢利变得困难。交易中，K 线的每一次跳动都会刺激着杏仁核，K 线越接近止损线，我们的恐惧越深，同时消极情绪的慢慢积累也会改变交易心态，为了解脱就选择提前离场，自我安慰着少亏就是赢，至少落袋为安。但是最后没想到行情随后一路上涨，一直接近原先设好的止赢线。蓦然回首，发现交易策略其实是对的，因情绪影响而没执行好策略。

在盯盘过程中，能够慢慢地读懂市场、看懂盘面，基于自己交易原则的基础上随机应变的人少之又少。天赋与运气缺一不可。这也就为什么在 A 股市场、期货市场和外汇市场这些偏短线投资的领域中，手动交易员如同大浪淘沙一样，换了一批又一批，能够稳定赢利的少之又少。“股神”巴菲特最厉害之处，未必是他的选股眼光，而是他控制“新哺乳动物脑”的功力。

2. 量化交易

量化交易不仅关注历史行情数据、基本面指标数据，而且会把一些非传统的数据，如市场情绪、财经新闻的关键字转化成机器可以理解的指标。数据越原始越好，例如可以直接购买交易所最原始的、未经清洗的行情数据。第三方数据提供商尽管物美价廉，但是在数据清洗过程中可能会把看似无用、实则隐含赢利机会的数据去掉。