Top Quant
CHRD前海智库

金融科技
丛书

零起点
Python大数据与量化交易

何海群 著

Win Or Home

读完本书内容和配套的教学代码，
就能够编写简单的量化策略函数。

100%零基础，无须任何编程、交易经验，也不需要具备超强的数据分析能力，只要会使用Excel就可以轻松学会本书讲解的知识点。

中国工信出版集团
電子工業出版社

零起点 Python 大数据与量化交易

何海群　著

電子工業出版社
Publishing House of Electronics Industry
北京 • BEIJING

目　录

内容简介
前言
第1章 从故事开始学量化
1.1 亿万富翁的“神奇公式”
1.1.1 案例1-1：亿万富翁的“神奇公式”
1.1.2 案例分析：Python图表
1.1.3 matplotlib绘图模块库
1.1.4 案例分析：style绘图风格
1.1.5 案例分析：colormap颜色表
1.1.6 案例分析：颜色表关键词
1.1.7 深入浅出
1.2 股市“一月效应”
1.2.1 案例1-2：股市“一月效应”
1.2.2 案例分析：“一月效应”计算
1.2.3 案例分析：“一月效应”图表分析
1.2.4 案例分析：颜色表效果图
1.2.5 “一月效应”全文注解版Python源码
1.2.6 大数据·宏分析
1.3 量化交易流程与概念
1.3.1 数据分析I2O流程
1.3.2 量化交易不是高频交易、自动交易
1.3.3 小资、小白、韭菜
1.3.4 专业与业余
1.4 用户运行环境配置

1.4.1 程序目录结构
1.4.2 金融股票数据包
1.5 Python实战操作技巧
1.5.1 模块检测
1.5.2 Spyder编辑器界面设置
1.5.3 代码配色技巧
1.5.4 图像显示配置
1.5.5 Python2、Python 3双版本双开模式
1.5.6 单版本双开、多开模式
1.5.7 实战胜于一切
1.6 量化、中医与西医
第2章 常用量化技术指标与框架
2.1 案例2-1：SMA均线策略
2.1.1 案例要点与事件编程
2.1.2 量化程序结构
2.1.3 main程序主入口
2.1.4 KISS法则
2.2 Python量化系统框架
2.2.1 量化行业关键词
2.2.2 国外主流Python量化网站
2.2.3 我国主流Python量化网站
2.2.4 主流Python量化框架
2.3 常用量化软件包
2.3.1 常用量化软件包简介
2.3.2 案例2-2：模块库列表
2.4 常用量化技术指标
2.4.1 TA-Lib金融软件包

2.4.2 案例2-3：MA均线函数调用
2.4.3 TA-Lib函数调用
2.4.4 量化分析常用指标
2.5 经典量化策略
2.5.1 阿尔法（Alpha）策略
2.5.2 Beta策略
2.5.3 海龟交易法则
2.5.4 ETF套利策略
2.6 常用量化策略
2.6.1 动量交易策略
2.6.2 均值回归策略
2.6.3 其他常用量化策略
2.7 起点与终点
第3章 金融数据采集整理
3.1 常用数据源API与模块库
3.1.1 大数据综合API
3.1.2 专业财经数据API
3.1.3 专业数据模块库
3.2 案例3-1：zwDatX数据类
3.3 美股数据源模块库
3.4 开源文档库Read the Docs
3.5 案例3-2：下载美股数据
3.6 财经数据源模块库TuShare
3.6.1 沪深股票列表
3.6.2 案例3-3：下载股票代码数据
3.6.3 CSV文件处理
3.7 历史数据

3.7.1 历史行情
3.7.2 案例3-4：下载近期股票数据
3.7.3 历史复权数据
3.7.4 案例3-5：下载历史复权数据
3.8 其他交易数据
3.9 zwDat超大股票数据源与数据更新
3.9.1 案例3-6：A股基本概况数据下载
3.9.2 案例3-7：A股交易数据下载
3.9.3 案例3-8：A股指数行情数据下载
3.9.4 案例3-9：美股交易数据下载
3.10 数据归一化处理
3.10.1 中美股票数据格式差异
3.10.2 案例3-10：数据格式转化
3.10.3 案例3-11：A股策略PAT实盘分析
3.10.4 案例3-12：数据归一化
3.11 为有源头活水来
第4章 PAT案例汇编
4.1 投资组合与回报率
4.1.1 案例4-1：下载多组美股数据
4.1.2 案例4-2：投资组合收益计算
4.2 SMA均线策略
4.2.1 SMA简单移动平均线
4.2.2 案例4-3：原版SMA均线策略
4.2.3 案例4-4：增强版SMA均线策略
4.2.4 案例4-5：A股版SMA均线策略
4.3 均线交叉策略
4.3.1 案例4-6：均线交叉策略

4.3.2 案例4-7：A股版均线交叉策略
4.4 VWAP动量策略
4.4.1 案例4-8：VWAP动量策略
4.4.2 案例4-9：A股版VWAP动量策略
4.5 布林带策略
4.5.1 案例4-10：布林带策略
4.5.2 案例4-11：A股版布林带策略
4.6 RSI2策略
4.6.1 案例4-12：RSI2策略
4.6.2 案例4-13：A股版RSI2策略
4.7 案例与传承
第5章 zwQuant整体架构
5.1 发布前言
5.2 功能简介
5.2.1 目录结构
5.2.2 安装与更新
5.2.3 模块说明
5.2.4 zwSys模块：系统变量与类定义
5.2.5 zwTools模块：常用（非量化）工具函数
5.2.6 zwQTBox：常用“量化”工具函数集
5.2.7 zwQTDraw.py：量化绘图工具函数
5.2.8 zwBacktest：回溯测试工具函数
5.2.9 zwStrategy：策略工具函数
5.2.10 zw_TA-Lib：金融函数模块
5.3 示例程序
5.4 常用量化分析参数
5.5 回溯案例：对标测试

5.5.1 对标测试1：投资回报参数
5.5.2 对标测试2：VWAP策略
5.6 回报参数计算
5.7 主体框架
5.7.1 stkLib内存数据库
5.7.2 Bars数据包
5.7.3 案例：内存数据库&数据包
5.7.4 qxLib、xtrdLib
5.7.5 案例5-1：qxLib数据
5.7.6 量化系统的价格体系
5.7.7 数据预处理
5.7.8 绘图模板
5.8 新的起点
第6章 模块详解与实盘数据
6.1 回溯流程
6.1.1 案例6-1：投资回报率
6.1.2 代码构成
6.1.3 运行总流程
6.2 运行流程详解
6.2.1 设置股票数据源
6.2.2 设置策略参数
6.2.3 dataPre数据预处理
6.2.4 绑定策略函数
6.2.5 回溯测试：zwBackTest
6.2.6 输出回溯结果数据、图表
6.3 零点策略
6.3.1 mul多个时间点的交易&数据

6.3.2 案例6-2：多个时间点交易
6.4 不同数据源与格式修改
6.4.1 案例6-3：数据源修改
6.4.2 数据源格式修改
6.5 金融数据包与实盘数据更新
6.5.1 大盘指数文件升级
6.5.2 实盘数据更新
6.5.3 案例6-4：A股实盘数据更新
6.5.4 案例6-5：大盘指数更新
6.6 稳定第一
第7章 量化策略库
7.1 量化策略库简介
7.1.1 量化系统的三代目
7.1.2 通用数据预处理函数
7.2 SMA均线策略
7.2.1 案例7-1：SMA均线策略
7.2.2 实盘下单时机与推荐
7.2.3 案例7-2：实盘SMA均线策略
7.3 CMA均线交叉策略
7.3.1 案例7-3：均线交叉策略
7.3.2 对标测试误差分析
7.3.3 案例7-4：CMA均线交叉策略修改版
7.3.4 人工优化参数
7.4 VWAP策略
7.4.1 案例7-5：VWAP策略
7.4.2 案例7-6：实盘VWAP策略
7.5 BBands布林带策略

7.5.1 案例7-7：BBands布林带策略
7.5.2 案例7-8：实盘BBands布林带策略
7.6 大道至简1+1
第8章 海龟策略与自定义扩展
8.1 策略库
8.1.1 自定义策略
8.1.2 海龟投资策略
8.2 tur海龟策略v1：从零开始
8.3 案例8-1：海龟策略框架
8.4 tur海龟策略v2：策略初始化
8.5 案例8-2：策略初始化
8.6 tur海龟策略v3：数据预处理
8.7 案例8-3：数据预处理
8.8 tur海龟策略v4：策略分析
8.9 案例8-4：策略分析
8.10 tur海龟策略v5：数据图表输出
8.10.1 案例8-5：图表输出
8.10.2 参数优化
8.10.3 案例8-6：参数优化
8.11 tur海龟策略v9：加入策略库
8.12 案例8-7：入库
8.13 庖丁解牛
第9章 TA-Lib函数库与策略开发
9.1 TA-Lib技术指标
9.1.1 TA-Lib官网
9.1.2 矩阵版TA-Lib金融函数模块
9.2 MACD策略

9.2.1 MACD策略1
9.2.2 案例9-1：MACD_v1
9.2.3 MACD策略2
9.2.4 案例9-2：MACD_v2
9.3 KDJ策略
9.3.1 KDJ策略1
9.3.2 案例9-3：KDJ01
9.3.3 KDJ策略2
9.3.4 案例9-4：KDJ02
9.4 RSI策略
9.4.1 RSI取值的大小
9.4.2 RSI策略
9.4.3 预留参数优化接口
9.4.4 案例9-5：A股版RSI策略
9.5 基石、策略与灵感
第10章 扩展与未来
10.1 回顾案例2-1：SMA均线策略
案例10-1：SMA均线策略扩展
10.2 大盘指数资源
10.2.1 大盘指数文件升级
10.2.2 大盘指数内存数据库
10.2.3 扩展zwQuantX类变量
10.2.4 大盘指数读取函数
10.2.5 案例10-2：读取指数
10.2.6 大盘数据切割
10.2.7 案例10-3：inxCut数据切割
10.3 系统整合

10.3.1 案例10-4：整合设置
10.3.2 案例10-5：修改指数代码
10.3.3 修改sta_dataPre0xtim函数
10.3.4 案例10-6：整合数据切割
10.3.5 修改绘图函数
10.4 扩展完成
案例10-7：SMA均线扩展策略
10.5 其他扩展课题
10.5.1 复权数据冲突
10.5.2 波动率指标DVIX
10.5.3 修改回溯主函数zwBackTest
10.5.4 案例10-8：波动率
10.5.5 空头交易
10.5.6 虚拟空头交易
10.5.7 修改检查函数
10.5.8 案例10-9：空头数据
10.6 终点与起点
附录A zwPython开发平台用户手册
附录B Python量化学习路线图

图书在版编目（**CIP**）数据

零起点Python大数据与量化交易 / 何海群著. —北京：电子工业出版社，2017.2

（金融科技丛书）

ISBN 978-7-121-30659-4

I. ①零… II. ①何… III. ①软件工具－程序设计 IV. ①TP311.561

中国版本图书馆CIP数据核字（2016）第308385号

责任编辑：黄爱萍

印　　刷：三河市双峰印刷装订有限公司

装　　订：三河市双峰印刷装订有限公司

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：27.5

字　　数：528千字

版　　次：2017年2月第1版

印　　次：2017年2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，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至zlts@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。

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。

内容简介

本书是国内较早关于Python大数据与量化交易的原创图书，配合zwPython开发平台和zwQuant开源量化软件学习，是一套完整的大数据分析、量化交易的学习教材，可直接用于实盘交易。本书有三大特色：第一，以实盘个案分析为主，全程配有Python代码；第二，包含大量的图文案例和Python源码，无须专业编程基础，懂Excel即可开始学习；第三，配有专业的zwPython集成开发平台、zwQuant量化软件和zwDat数据包。

本书内容源自笔者的原版教学课件，虽然限于篇幅和载体，省略了视频和部分环节，但核心内容都有保留，配套的近百套Python教学程序没有进行任何删减。考虑到广大入门读者的需求，笔者在各个核心函数环节增添了函数流程图。

丛书编委会

主编：何海群

编委：

欧耘华，CHRD前海智库创始人，Python产业联盟发起人。科技金融、消费金融、艺术金融与产业政策专家。北京亚欧科技、深圳“中国科谷”等多家机构特聘专家。

刘志明，免费开源金融数据接口TuShare创始人，专业从事量化投资支持与服务。

杨适安，中信建投证券金融产品及创新业务部副总裁，对期权、期货等衍生品种有较深入的研究，擅长用中低频量化的方法配置大类资产和金融产品。

吴尚谦，金融交易实盘培训专家，纯实盘首席金融交易教练，15年股票、期货、外汇实盘交易经验，金融交易混元合一交易系统发明人，业内领先的现场实盘逻辑推演教学法资深讲师。北京正道阳光投资管理有限公司总经理。

曹嘉和，美国普林顿大学荣誉博士，北大当代企业文化研究所研究员。现任北大中国持续发展研究中心政府与社会资本合作研究所所长，未名湖智库秘书长。

王丁杰，QuantDigger量化软件开发者，擅长机器学习、NLP自然语言处理、人工智能。

王黎中，资深金融交易心理专家，北京正道阳光投资管理有限公司副总经理，幸福紫薇俱乐部董事长。

王家苍，经济学博士，李约瑟（《中国科学技术史》作者）隔代弟子，产业政策专家。长期致力于科技和文化的融合研究，参与科技金融、消费金融和艺术金融等课题的研究。

蔡磊，原高通公司工程师、项目经理，精通Python数据挖掘、数据库技术、机器学习、量化交易理论，熟悉4G无线技术、手机芯片架构及VOIP技术，极宽量化开源团队核心成员。

王硕，高级软件工程师，精通Python数据分析，擅长Java、JavaScript、HTML5和数据库技术，极宽量化开源团队核心成员。

吴娜，电信数据挖掘工程师，精通缠论和江恩波动理论，率先研发江恩波动法则量化模型，著有《游戏数据分析的艺术》，极宽量化开源团队发起成员。

余勤，AMD验证工程师，擅长数据分析，热爱Python量化分析，极宽量化开源团队发起成员。

孙洋洋，西南财经大学金融工程硕士，擅长机器学习、网络爬虫，有私募机构量化分析系统开发实盘经验，极宽量化开源团队发起成员。

李政隆，上海交大计算机硕士，分布式系统架构师，擅长金融数据抓取与分析，极宽量化开源团队发起成员。

前言

2014年，美国银行、美林证券的“石英”项目、摩根大通的“雅典娜”项目都不约而同地选择了Python作为金融行业的标准编程语言。

全世界的金融工程行业全部重新洗牌，这为中国的金融工程从业人员带来了前所未有的机遇。资本的力量是强大的，也是冷酷无情的。

2016年5月，《华尔街日报》报道，目前华尔街的三大编程语言是：C、Java与Python。其中，C与Java成为三大语言之一有两方面原因：一方面是由于历史积累，另一方面是系统架构设计的需要。而在应用领域Python更胜一筹，因为Python已经成为金融行业量化领域的标准编程语言。

本书是国内较早关于Python大数据与量化交易的原创图书。本书配合zwPython、zwQuant开源量化软件学习，是一套完整的大数据分析、量化交易的学习教材，可直接用于实盘交易。

本书内容包括：

- 近50万字的图文课件；
- 数十套结合课件的Python教学代码；
- 全套zwPython开源平台；
- 业内首套面向初学者的开源量化系统zwQuant；
- 国内较大的开源金融数据包zwDat，包括tick数据。

100%零基础，无须任何编程、交易经验，也不需要具备超强的数据分析能力，只要会使用Excel就可以轻松学会本书讲解的知识点。读完本书内容和配套的教学代码，就能够编写简单的量化策略函数。

本书的内容源自笔者的原版教学课件，虽然限于篇幅和载体，省略了视频和部分环节，但核心内容都有保留，配套的近百套Python教学程序没有进行任何删减。

考虑到广大入门读者的需求，笔者在各个核心函数环节增添了函数流程图。

量化新人学习指南

1．多参考笔者的字王量化网站（http://www.ziwang.com）和笔者博客（以原创为主，网址是http://blog.sina.com.cn/zbrow）。

2．本书配套程序可在百度网盘下载，网址是http://pan.baidu.com/s/ljIg944u。本书读者QQ群（zwPython量化总群）：124134140。网盘和群共享包含很多资源，读者可自行下载，也可以上传。建议先看“漫画学××系列”，该系列比较经典，而且简单；zwQuant开源量化软件在群共享和网盘都可下载。

3．强烈建议初学者先下载zwPython集成版，阅读zwPython中文手册，再开始学习Python，这样可以少走很多弯路。

资源下载

为配合本书出版，方便广大读者学习Python量化系统，笔者特意将与本书相关的教学资源打包成一个独立的教学版本压缩包，便于读者使用。

- 教学版为一个独立压缩包，解压即可，并配有说明文档。
- 教学版内置了zwPython2016M10版开发平台（Python 3.5版）、Python 2.7版开发平台、zwDat金融数据包、zw_down25金融数据下载更新程序包和zw_k10配套量化教学课件程序。

教学版软件和配套资源下载地址如下。

- 极宽公司网盘下载地址：http://pan.baidu.com/s/1kVO6T19，密码：v2ub。
- 出版社网站下载地址：http://www.broadview.com.cn/30659。

以上网址如果发生变动，请读者浏览字王网站（http://www.ziwang.com）或者极宽公司网站（http://www.zquant.cn）获取最新信息。

致谢

虽然很多网友在笔者博客留言，要求购买本书，但本书的创作和正式出版还是经历了许多波折。

如今本书终于出版，在此，要特别感谢电子工业出版社的黄爱萍和戴新编辑，感谢她们在选题策划和稿件整理方面做出的大量工作。

同时，在本书创作过程中，极宽开源量化团队和培训班的全体成

员，提出很多宝贵的意见，并对部分课件程序做了中文注解。

特别是吴娜、余勤两位同学，为极宽开源量化文库和zwQuant开源量化软件编写文档，以及在团队成员管理方面做了大量工作，为他们的付出表示感谢。

何海群（字王）
北京极宽科技有限公司CTO
2016年11月25日

第1章
从故事开始学量化

本书全程采用MBA个案教学模式，每章都有介绍案例，这些案例都配有纯Python源码，由浅到深，尽量覆盖量化交易、量化分析的各个环节，同时书中还配合有关案例，对Python编程技巧、金融量化领域的背景知识做了部分介绍。

即使非金融专业、没有经验的读者，也能够根据书中的案例举一反三，借鉴到自己的实盘操作当中。

至于案例背后的理论、算法、模型，有些非常复杂，属于专业课程，甚至学术范畴，初学者无须一一细究。

本章是入门课程，通过几个经典的案例，读者将学习以下知识：

- 量化算法、量化策略（投资策略）；
- 不同的量化算法、量化策略所带来的收益差别；
- 用Excel进行简单的量化分析；
- 用Python进行简单的量化分析。

很多人，特别是初学者，一提起量化投资就觉得很神秘、很复杂，特别是提到其中的交易算法、量化模型，更加觉得高不可攀。

本书一开始就从趣味投资学入手，通过几个简单的小故事揭开量化

投资的神秘面纱。

1.1　亿万富翁的“神奇公式”

1.1.1　案例1-1：亿万富翁的“神奇公式”

假设有一位年轻人，每年定期存款1.4万元，享受平均5%的利率，如此持续40年，他可以积累的财富为：

1.4×(1+5%)^40=169（万元）

注意，这个公式不是一般的数学计算公式，而是复利公式，其计算方法如下：

（1）本利和=本金×(1+年利率)^期限。

（2）Excel复利计算公式为“=FV(rate,nper,pv,pmt,type)”。

Excel操作为：“=FV(5%,40-1,-14000, -14000)”。

结果是：1 691 196.84（元）。

如果这位年轻人将每年应存的钱全部投资到股票或房地产市场，并假定能获得年均20%的投资回报率，则40年后，他能积累多少财富？

一般人猜测是200万元至800万元。然而，“神奇公式”给出的答案是1.0281亿元！

这个数据是依照财务学计算年金的公式得到的：1.4×[（1+20%）^（40-1）]=1.0281（亿元）

Excel操作为：“=FV(20%,40-1,-14000,-14000)”。

结果是：102 814 009.76（元）。

试着将这个天文数字与平均投资回报率仅为5%的定期存款结果相比，你会发现两者收益的差距达60多倍，多么令人惊讶！

这个“神奇公式”就是所有量化分析的基础公式：复利计算公式。

做量化分析、量化投资，任何算法、模型、短期收益的高低都无所谓，关键是稳定。只要稳定了，使用复利公式，哪怕是1%的增长，其增长速度都是几何裂变模式的。

量化投资所用的算法、模式、策略，其实只是数据分析的一种手段。

在“神奇公式”这个案例中，就涉及到两种不同的策略和算法公式。

- 保守投资策略：银行储蓄，5%的年收益。
- 激进投资策略：股市投资，20%的年收益。

保守的银行储蓄策略，5%的年收益基本可以实现。激进的股市投资策略，20%的年收益需要考虑风险因素，此次假设20%的年收益率已经扣除了各种风险因素。

下面用Excel分析一下这个“神奇公式”。

配套文件名是\zwpython\zw_k10\dat\k101.xls。

打开文件，如图1-1所示。

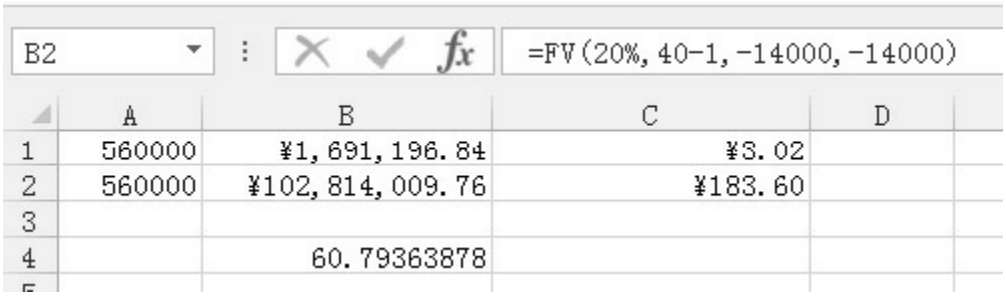


图1-1 “神奇公式”收益表

从“神奇公式”收益表我们可以看到：

- 用户累计投入56万元（1.4万元×40）；
- 保守模式，5%年收益率，40年总收益169万元，投资回报率3倍；
- 激进模式，20%年收益率，40年总收益1.02亿元，投资回报率184倍；
- 激进模式与保守模式两者的回报相差61倍。

下面再看看“神奇公式”Python版本的脚本程序。

脚本文件名为\zwpython\zw_k10\k101.py。

```
# -*- coding: utf-8 -*-

import numpy as np

# ======================
```

```
def sta001(k,nyear,xd):
     d2=np.fv(k,nyear,-xd,-xd);
     d2=round(d2)
     return d2

# ======================

d40=1.4*40
print("d40,40 x 1.4=",d40)
d=sta001(0.05,40-1,1.4);
print("01保守投资模式，",d,round(d/d40))

d2=sta001(0.20,40-1,1.4);
print("02激进投资模式，",d2,round(d2/d40))

dk=round(d2/d)
print("dk,两者差别（xx倍）：",dk)
```

以上程序运行结果如图1-2所示，和Excel的计算结果完全一致。

```
In [6]: runfile('F:/zwPython/demo/zw_k10/k101.py', wdir='F:/zwPython/demo/zw_k10')
d40,40 x 1.4= 56.0
01保守投资模式， 169.0 3.0
02激进投资模式， 10281.0 184.0
dk,两者差别（xx倍）： 61.0
```

图1-2 “神奇公式”Python程序运行结果

算法（Algorithm）是指解题方案准确而完整的描述，是一系列解决问题的清晰指令。算法代表着用系统的方法描述解决问题的策略机制。

也就是说，能够对一定规范的输入，在有限时间内获得所要求的输出。如果一个算法有缺陷，或不适合于某个问题，那么执行这个算法将不会解决这个问题。

案例1-1中的Python程序k101.py就是采用最简单的顺序算法，依照先后顺序，根据不同收益率，依次调用策略计算函数sta001计算相关的总收益。

读者仔细对比一下Excel与Python程序两个不同的分析过程，可以发现，其中的关键都是fv函数：

```
d2=np.fv(k,nyear,-xd,-xd);
```

fv为复利函数，看起来简单，其实具体编程非常复杂，这些现成函数读者知道如何调用就可以了。同样，量化分析的相关软件和算法原理都很专业，一般用户只需要知道如何使用相关的软件和工具就可以了。

有趣的是，Excel与Python程序的复利函数不仅名称一样，而且连调用的参数、顺序都是一致的。

所以笔者经常说，Pandas（潘达思）数据分析软件不过是Python版本的Excel报表而已，量化分析其实很简单，懂Excel就可以入门学习了。

事实上，微软公司的Excel电子表格软件是世界上最流行的数据分析、量化分析工具之一。Excel可以完成各种专业的数据分析，提供BI（商业智能）功能，甚至还能够提供服务器的功能。

不过，综合而言，专业的量化分析、数据分析使用Python语言和Pandas数据分析软件更加专业、简单和方便。

目前，大数据领域已经公认：十亿以下的大数据项目，Pandas数据分析软件是首选的工程一线的处理方案。

注意

本节的案例“神奇公式”Python版本的脚本程序文件名为\zwpython\zw_k10\k101.py。

在文件名“k101.py”中，字母k后的第一个字符1代表第1章的案例，即k101.py表示第1章的第一个Python案例程序。本书均采用这种命令方式，k2~k9表示第2~第9章的案例，kx表示第10章的案例。

1.1.2 案例分析：Python图表

下面在程序k101.py中加入图表绘制语句，再通过参数设置不同的年利率，看看不同的年利率最终对应的年收益是如何变化的。

我们采用四种不同的利率设置：

- 5%、10%、15%、20%。
- 对应参数：0.05、0.10、0.15、0.20。

脚本文件名为\zwpython\zw_k10\k101dr.py。

脚本代码如下：

```
# -*- coding: utf-8 -*-

import numpy as np
import scipy as sp
import pandas as pd
import matplotlib as mpl
import matplotlib.pyplot as plt
import seaborn as sns

# =======================
mpl.style.use('seaborn-whitegrid');

def sta001(k,nyear,xd):
     d2=np.fv(k,nyear,-xd,-xd);
     d2=round(d2)
     return d2

# =======================

d40=1.4*40
print("d40,40 x 1.4=",d40)
d=sta001(0.05,40-1,1.4);
print("01保守投资模式，",d,round(d/d40))
```

```
d2=sta001(0.20,40-1,1.4);
print("02激进投资模式，",d2,round(d2/d40))


dk=round(d2/d)
print("dk,两者差别（xx倍）：",dk)


dx05=[sta001(0.05,x,1.4) for x in range(0,40)]
dx10=[sta001(0.10,x,1.4) for x in range(0,40)]
dx15=[sta001(0.15,x,1.4) for x in range(0,40)]
dx20=[sta001(0.20,x,1.4) for x in range(0,40)]
#print(dx05)print(dx20)

df=pd.DataFrame(columns=['dx05','dx10','dx15','dx20']);
df['dx05']=dx05;df['dx10']=dx10;
df['dx15']=dx15;df['dx20']=dx20;
print("")
print(df.tail())
df.plot();
```

运行结果如图1-3所示。

```
In [14]: runfile('F:/zwPython/demo/zw_k10/k101dr.py', wdir='F:/zwPython/demo/zw_k10')
d40,40 x 1.4= 56.0
01保守投资模式， 169.0 3.0
02激进投资模式， 10281.0 184.0
dk,两者差别（xx倍）： 6..0

    dx05  dx10  dx15   dx20
35   134   419  1420   4955
36   142   462  1634   5947
37   151   510  1881   7138
38   160   562  2165   8567
39   169   620  2491  10281
```

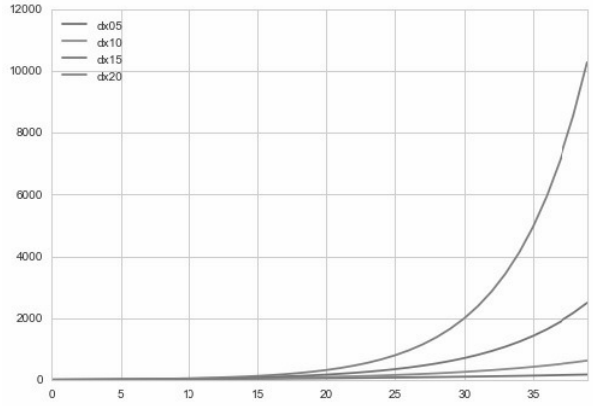


图1-3 “神奇公式”收益图

从图1-3可以看出，5%、10%、15%和20%的年利率对应的最终40年总收益分别是169万元、620万元、2491万元和10281万元。

注意在图1-3中，图形上面的最后一行字符：

```
39  169          620    2491    10281
```

这行数字表示不同的年利率对应的最后一年的收益结果，其中第一个数字是39，它是index（索引），之所以是39，是因为第一年只有本金，没有利息收入，所以40年的总收益必须扣除第一年的收益，在复利公式中按39年来计算时间。

1.1.3 matplotlib绘图模块库

单纯的数据分析非常枯燥，大量的数据令读者很头疼，所以需要把数据图形化显示。

可视化计算是量化分析、数据分析的一个重要组成部分，Python量化分析和数据分析最重要的绘图模块是matplotlib。

matplotlib绘图模块是Python编程语言最著名的绘图模块库，类似于MATLAB和R语言，它提供了一整套和MATLAB相似的命令API，非常适合交互式的制图。

matplotlib的官方网址是http://matplotlib.org/。

matplotlib绘图模块的文档相当完备，并且网站的Gallery页面中有上百幅缩略图，打开之后都有源程序。如果用户需要绘制某种类型的图，那么只需要在这个页面中浏览、复制、粘贴一下，基本上都能完成。

在Python语言中，比较著名的绘图模块还有gnuplot和seaborn，不过语法与Python差别太大，通常使用不多。

近年来，随着Python在大数据、金融量化方面的扩展，出现了许多新的绘图模块库，但matplotlib依然是最重要的Python绘图模块。

matplotlib 2.0版本正在开发中，同时也出现了许多优秀的、第三方matplotlib扩展库，如Basemap、seaborn、mplot3D、ggplot、prettyplotlib和iTerm2。

其中，seaborn和ggplot是统计绘图模块库和美化增强库：prettyplotlib模块库与数据分析、量化分析领域关联大，是用得较多的绘

图模块库。

这些绘图模块库在各自官方网站及Python网站的pyPI模块库下载中心都有下载。pyPI网址是https://pypi.python.org/pypi。

Python常用的模块库如图1-4~图1-6所示。

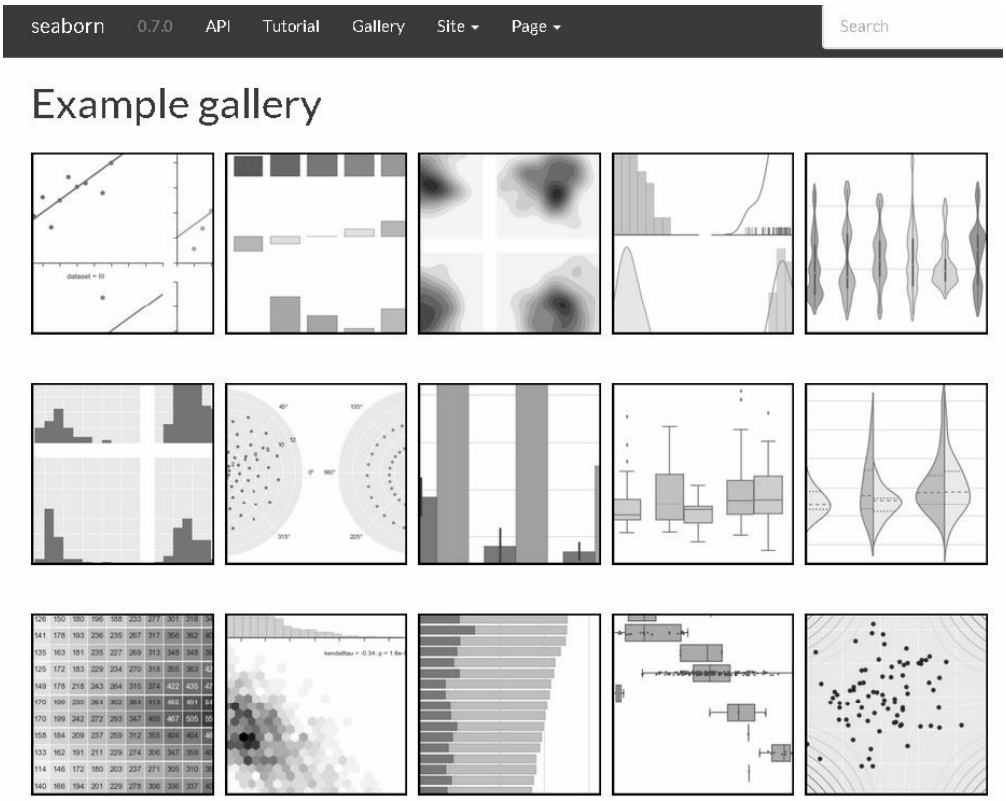


图1-4　seaborn绘图模块

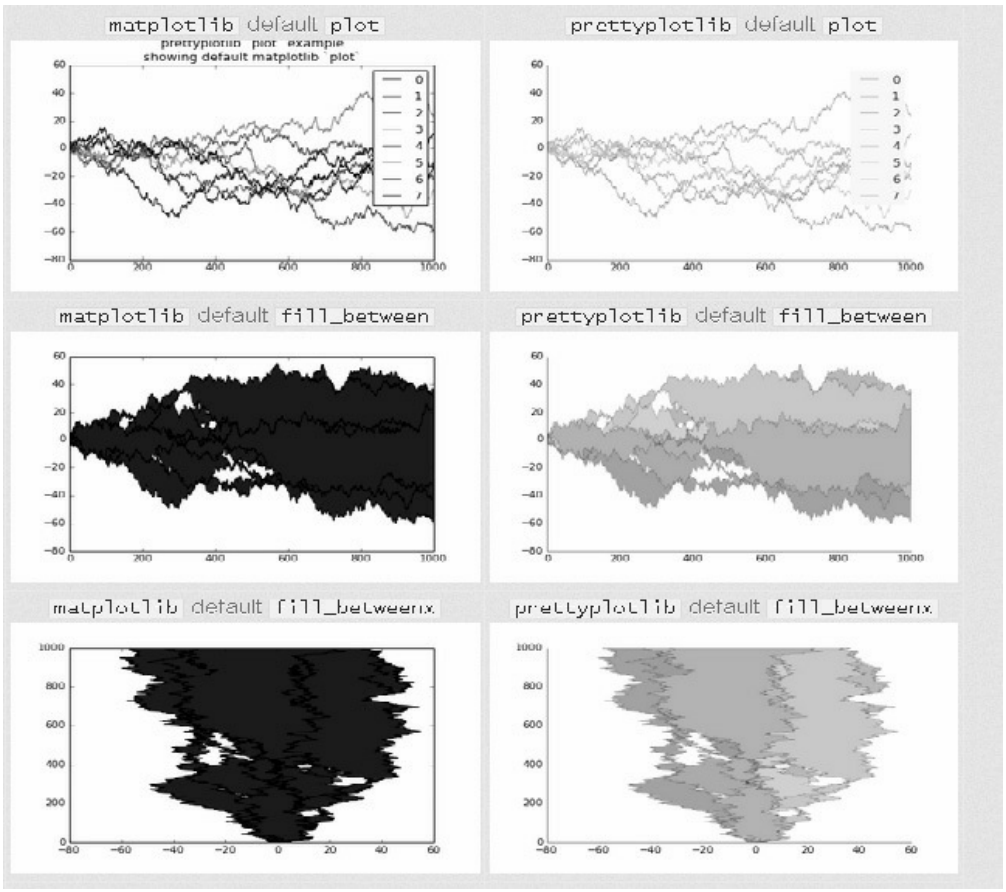


图1-5 prettyplotlib绘图模块

ggplot is powerful

```
ggplot(diamonds, aes(x='price', fill='cut')) +\
    geom_density(alpha=0.25) +\
    facet_wrap("clarity")
```

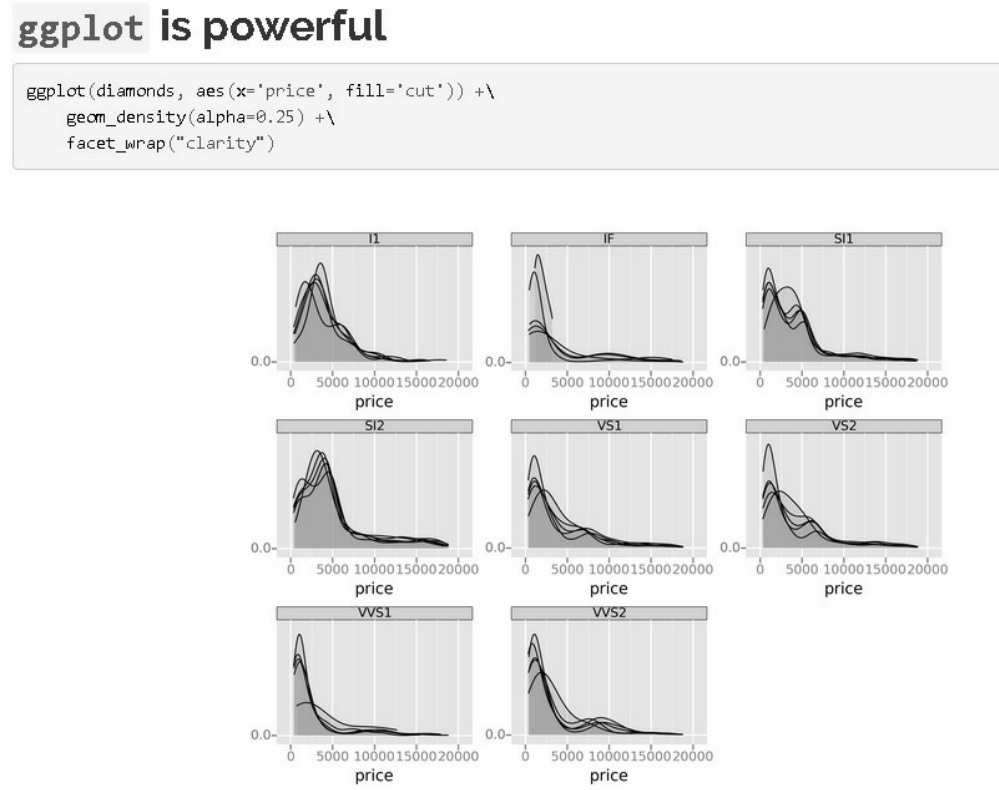


图1-6　ggplot绘图模块

matplotlib绘图模块库非常专业且庞大，有关的细节读者可慢慢研究，具体内容可浏览http://matplotlib.org/mpl_toolkits/index.html。

1.1.4　案例分析：style绘图风格

下面介绍matplotlib模块库中的style绘图风格函数，它是与绘图效果关联最大的一个函数。

脚本文件名为\zwpython\zw_k10\k101sty.py。

全部代码如下：

```
# -*- coding: utf-8 -*-

import numpy as np
import scipy as sp
import pandas as pd

import matplotlib as mpl
import matplotlib.pyplot as plt
import seaborn as sns
import prettyplotlib
import ggplot

# =======================
mpl.style.use('seaborn-whitegrid');

def sta001(k,nyear,xd):
     d2=np.fv(k,nyear,-xd,-xd);
     d2=round(d2)
     #print(nyear,d2)
     return d2

def dr_xtyp(_dat):
     #xtyp=['bmh','dark_background','fivethirtyeight','ggplot',
     i=0;
     for xss in plt.style.available:
         plt.figure()
         plt.style.use(xss);
```

```
        _dat.plot()
        fss="tmp\\k101_"+xss+".png";plt.savefig(fss);
        i+=1;
        print(i,xss,",",fss)
        plt.show()


# ======================

dx05=[sta001(0.05,x,1.4) for x in range(0,40)]
dx10=[sta001(0.10,x,1.4) for x in range(0,40)]
dx15=[sta001(0.15,x,1.4) for x in range(0,40)]
dx20=[sta001(0.20,x,1.4) for x in range(0,40)]

df=pd.DataFrame(columns=['dx05','dx10','dx15','dx20']);
df['dx05']=dx05;df['dx10']=dx10;
df['dx15']=dx15;df['dx20']=dx20;
print(df.tail())

dr_xtyp(df)
```

这段代码中最重要的语句是：

```
plt.style.use(xss);
```

另外，还可以设置绘图风格。在Python中，内置的绘图风格有21种，包括bmh、classic、dark_background、fivethirtyeight、ggplot、grayscale、seaborn-bright、seaborn-colorblind、seaborn-dark、seaborn-darkgrid、seaborn-dark-palette、seaborn-deep、seaborn-muted、seaborn-notebook、seaborn-paper、seaborn-pastel、seaborn-poster、seaborn-talk、seaborn-ticks、seaborn-white和seaborn-whitegrid。

其中，大部分是seaborn模块库引入的，运行后输出图保存在“tmp”目录下，具体效果如图1-7所示。

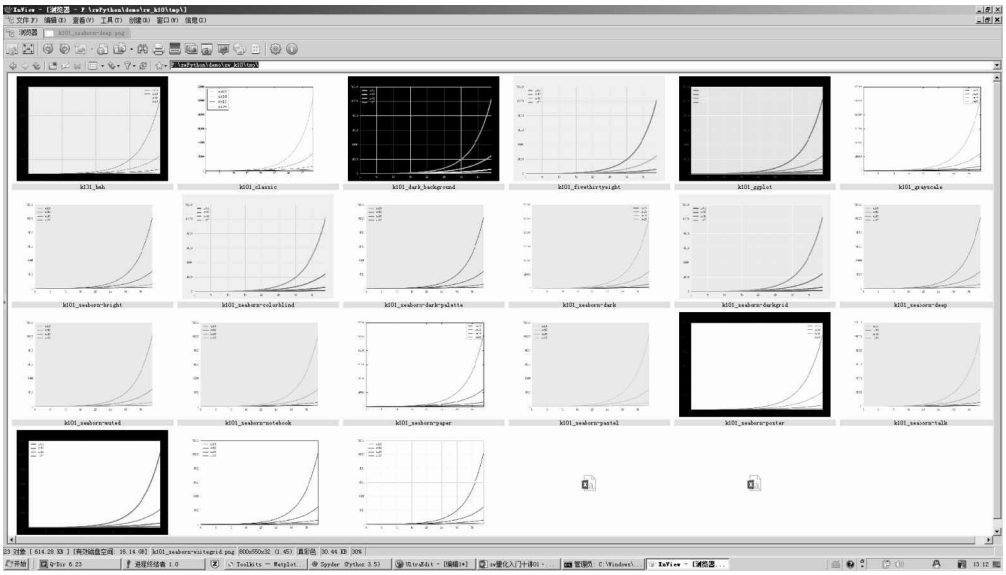


图1-7　matplotlib的不同style参数效果图

此外，在目录“x:\zwPython\zw_k10\pic\style\”中也有类似的效果图片。

通常，style绘图参数一般使用seaborn-whitegrid，即白色网格。在导入模块库后和其他程序语句前，用matplotlib设置为全局风格，代码如下：

```
mpl.style.use('seaborn-whitegrid');
```

1.1.5 案例分析：colormap颜色表

下面介绍颜色表colormap。

脚本文件名为\zwpython\zw_k10\k101cmap.py，全部代码如下：

```
# -*- coding: utf-8 -*-

import numpy as np
import scipy as sp
import pandas as pd

import matplotlib as mpl
import matplotlib.pyplot as plt
import seaborn as sns
import prettyplotlib
import ggplot

# =======================
mpl.style.use('seaborn-whitegrid');
```

```
def sta001(k,nyear,xd):
     d2=np.fv(k,nyear,-xd,-xd);
     d2=round(d2)

     return d2

def dr_cmap(_dat):
     fss='cor_maps.csv'
     cm8 = pd.read_csv(fss,encoding='gbk')
     i=0;
     for xss in cm8['name']:
         plt.figure()
         _dat.plot(colormap=xss)
         fss="tmp\\k101cor_"+xss+".png";plt.savefig(fss);
         i+=1;print(i,xss,",",fss)
         plt.show()

# ======================

dx05=[sta001(0.05,x,1.4) for x in range(0,40)]
dx10=[sta001(0.10,x,1.4) for x in range(0,40)]
dx15=[sta001(0.15,x,1.4) for x in range(0,40)]
dx20=[sta001(0.20,x,1.4) for x in range(0,40)]

df=pd.DataFrame(columns=['dx05','dx10','dx15','dx20']);
df['dx05']=dx05;df['dx10']=dx10;
df['dx15']=dx15;df['dx20']=dx20;
```

```
print(df.tail())


dr_cmap(df)
```

运行结果保存在“tmp”目录下。

colormap设置语句如下：

```
_dat.plot(colormap=xss)
```

在plot语句中，颜色表用关键词“colormap”指定，通常内置的颜色表有数十种，具体数量随着matplotlib版本的不同及扩充模块的不同而有所差异。在笔者测试的版本中，colormap颜色表有76种，具体参见文件cor_map.csv，它在“\zwpython\doc\”目录下。

此外，在目录“x:\zwPython\zw_k10\pic\cormap\”中也有相关的效果图片。

因为线条图色彩效果表现不佳，因此在后面的条形图中再进行介绍。

matplotlib不同的colormap颜色表参数效果图如图1-8所示。

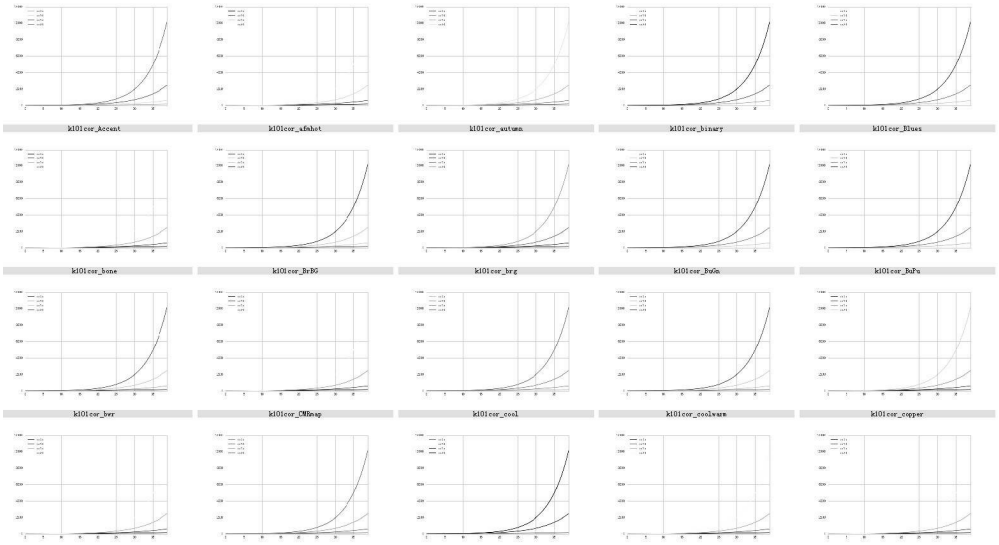


图1-8 matplotlib不同的colormap颜色表参数效果图1

1.1.6 案例分析：颜色表关键词

colormap颜色表可以使用预设的参数，也可以编写脚本来抓取系统内置的颜色表参数。

脚本文件名为\zwpython\zw_k10\k101cmap2.py，示例代码如下：

```
# -*- coding: utf-8 -*-


import numpy as np
import scipy as sp
import pandas as pd


import matplotlib as mpl
import matplotlib.pyplot as plt
```

```
import seaborn as sns
import prettyplotlib
import ggplot

# =======================
mpl.style.use('seaborn-whitegrid');

def sta001(k,nyear,xd):
     d2=np.fv(k,nyear,-xd,-xd);
     d2=round(d2)

     return d2

# =======================

dx05=[sta001(0.05,x,1.4) for x in range(0,40)]
dx10=[sta001(0.10,x,1.4) for x in range(0,40)]
dx15=[sta001(0.15,x,1.4) for x in range(0,40)]
dx20=[sta001(0.20,x,1.4) for x in range(0,40)]

df=pd.DataFrame(columns=['dx05','dx10','dx15','dx20']);
df['dx05']=dx05;df['dx10']=dx10;
df['dx15']=dx15;df['dx20']=dx20;
print(df.tail())

#df.plot(colormap='xss')
df.plot(colormap='hot')
```

运行结果如图1-9所示。

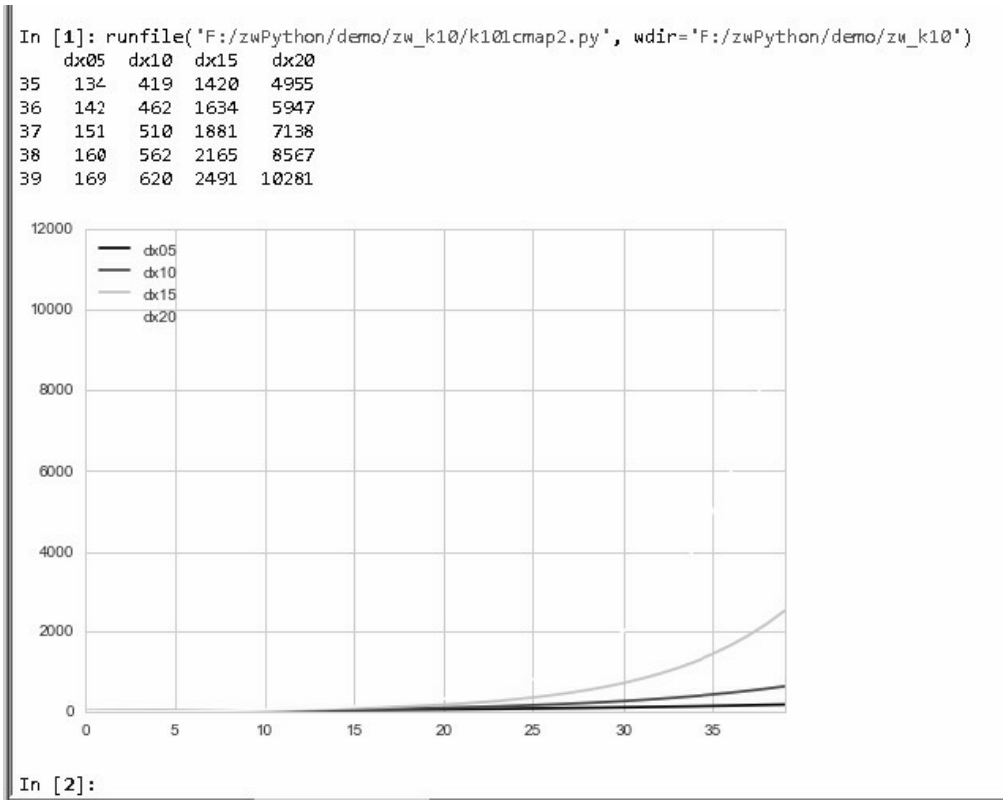


图1-9　matplotlib不同的colormap颜色表参数效果图2

下面修改一下代码：

- 取消第40行前的“#”注释符号；
- 在第41行前加上“#”注释符号。

修改后，相关代码如图1-10所示。

```
40 df.plot(colormap='xss')
41 #df.plot(colormap='hot')
42
```

图1-10　colormap颜色表参数代码修改

再次运行，因为“xss”是错误的颜色代码，因此运行出错，提示信息如图1-11所示。

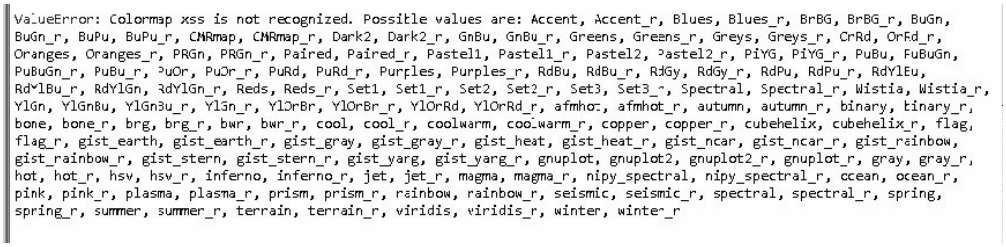


图1-11　错误colormap颜色表参数运行截图

复制“Possible values are:”后面的文字到记事本或UltraEdit编辑软件（推荐用UltraEdit软件编辑文本）。

把文字里的“,”（逗号）替换成“回车换行”符号，即可获得系统内置的可用colormap颜色列表。

需要注意的是，在文件cor_map.csv中只收录了正序的colormap颜色关键词。在关键词后面加上“_r”就可以组成同色系的、反序颜色表关键词，如Blues与Blues_r。

1.1.7　深入浅出

下面对1.1节的内容进行小结，1.1节主要讲述了以下内容：

- 趣味理财学之“神奇公式”；
- 神奇公式的Excel量化分析；
- 神奇公式的Python编程分析；
- Python图表绘制；
- Python绘图风格函数；

- Python颜色表colormap；
- Python颜色表关键词。

通常，讲到“神奇公式”的Python编程分析本节就可以结束了。但是，我们通过这个最简单的编程案例逐步深入，进一步学习了相关的Python绘图、style绘图风格、colormap颜色表等多种实战技巧。

授之以鱼，不如授之以渔。

做学问，其实很简单，把简单的问题复杂化，从多个不同的角度不断分析、深入，明白了原理、掌握了逻辑，就知道如何把复杂的问题简单化，从最简单、最直接的角度入手解决问题。笔者希望把这种学习的方法传授给读者。

学会了这种由浅入深、化繁为简的学习方法后，读者在面对各种新的领域、新的课题时就会知道如何把握全局、掌控方向。

1.2 股市“一月效应”

1.2.1 案例1-2：股市“一月效应”

“一月效应”是从统计学角度分析股市，发现一月份的回报率往往是“正数”，而且会比其他月份高，而在十二月，股市的回报率很多时候会呈现负值。

在传统股市，每逢进入新一年的一月，股市总是升的多、跌的少。最初出现“一月效应”的国家是美国，后来其他国家的学者也陆续发现“一月效应”存在于其他股市中。

对于“一月效应”的出现，学者认为这与美国“资本增值税”的税务安排、员工年终奖及美国年尾的重要假期（感恩节、圣诞节、元旦）有很大关系。

我国因为农历春节，通常“一月效应”会延后，大多在农历新年前后，即二月左右。

笔者挑选一些经典的传统投资策略进行具体的量化分析。“一月效应”从量化投资角度而言，就是一种投资策略。

不过，这种投资策略到底效果如何，我们可以通过Python编程进行具体的量化分析。

“一月效应”有两段示例代码：

- k102m1.py文件，用于计算美国股票、中国A股的全部股票，以及多种指数成分股的相关数据；
- k102m1dr.py文件，根据计算数据，绘制相关的结果图片。

由于是Python量化编程，所以我们不仅计算一月的升跌情况，还计算了其他各月的升跌情况。

本书采用的股票代码表源自zwDat股票数据包zwDat\inx。各文件含义如下。

（1）中国股票数据。

- inx_code.csv，中国A股大盘及各种指数代码。
- stk_base.csv，中国A股2810只股票代码公司概况。
- stk_code.csv，中国A股2810只股票代码。
- stk_hs300.csv，中国沪深300指数股票代码。
- stk_sz50.csv，中国上证50指数股票代码。
- stk_zz500.csv，中国中证500指数股票代码。

（2）美国股票数据。

- inxYahoo.csv，全部6688只美股代码。
- inxYahoo30sp.csv，道琼斯30指数美股代码。
- inxYahoo100ns.csv，纳斯达克100指数美股代码。
- inxYahoo100sp.csv，道琼斯100工业指数美股代码。
- inxYahoo500sp.csv，道琼斯30指数美股代码。
- inxYahoo600.csv，量化常用美股600股票代码。

1.2.2 案例分析：“一月效应”计算

相关脚本文件名为\zwpython\zw_k10\k102m1.py。

全部代码如下：

```
# -*- coding: utf-8 -*-

import sys
import pandas as pd
import numpy as np

import zwSys as zw  #::zwQT
import zwTools

#----------code

def zw_anz_m1sub(xcod,rss,kstr):
     fss=rss+xcod+".csv";print(fss)
     nSum=0;nAdd=0;nDec=0;
     knum=int(kstr);knum2=knum+1;
     try:
         df = pd.read_csv(fss,index_col=0,parse_dates=[0],encodi
         df =df.rename(columns={'Close':'close'});df =df.sort_in
         #
```

```
        _tim0=df.index[0];_ynum0=_tim0.year;
        _tim9=df.index[-1];_ynum9=_tim9.year+1;
        for ynum in range(_ynum0,_ynum9):
            ystr=str(ynum);_tim1x='-1';
            ystr2=ystr+"-"+kstr;  #print(ystr2,len(df),knum)
            if knum==12:
                ystr3=ystr+"-"+kstr+'-31';
                df2=df[(df.index>=ystr2)&(df.index<=ystr3)];
            else:
                kstr2=str(knum2);
                if knum2<10:kstr2='0'+kstr2;
                ystr3=ystr+"-"+kstr2+'-01';
                df2=df[(df.index>=ystr2)&(df.index<ystr3)];
            #print(ystr2,ystr3,len(df2))
            if (len(df2)>0):
                _tim1x=str(df2.index[0].month);
                if (len(_tim1x)<2):_tim1x='0'+_tim1x;
            if (_tim1x==kstr):
                df1=df2[ystr2];
                if (len(df1)>0):
                    xd1a=df1.ix[0];xd1z=df1.ix[-1];nSum+=1;
                    vd1a=xd1a['close'];vd1z=xd1z['close'];
                    if (vd1z>vd1a):nAdd+=1
                    else:nDec+=1;

    except IOError:
        pass;    #skip,error
```

```
    #print('nSum,nAdd,nDec,',nSum,nAdd,nDec);
    return nSum,nAdd,nDec

def zw_stk_anz_m01(qx,finx0,rss,ksgn):
    fss = qx.rdatInx+finx0+".csv";   #stk_code.csv,inxYahoo.csv
    dinx = pd.read_csv(fss,encoding='gbk')

    mx1={};mx1['finx']=finx0;mx1['ksgn']=ksgn;
    mx1['nSum']=0;mx1['nAdd']=0;mx1['nDec']=0;
    #nSum=0;nAdd=0;nDec=0;
    i=0;xn9=len(dinx['code']);mx1['nstk']=xn9;
    for xcod in dinx['code']:
        i+=1;
        if (not isinstance(xcod,str)):xcod="%06d" %xcod;

        dSum,dAdd,dDec=zw_anz_m1sub(xcod,rss,ksgn);

        mx1['nSum']=mx1['nSum']+dSum;
        mx1['nAdd']=mx1['nAdd']+dAdd;
        mx1['nDec']=mx1['nDec']+dDec;
        print(i,'/',xn9,mx1);

    mx1['kAdd']=np.round(mx1['nAdd']*100/ mx1['nSum']);
    mx1['kDec']=np.round(mx1['nDec']*100/ mx1['nSum']);

    return mx1
```

```
def zw_stk_anz_mx(qx,finx0,rss):
     c10=["finx","ksgn","nstk",'nSum','nAdd','nDec','kAdd','kDec
     df=pd.DataFrame(columns=c10);
     ftg="tmp\\mx_"+finx0+".csv";print(ftg)

     for i in range(12):
         ksgn=str(i+1);
         if i<9:ksgn='0'+ksgn;
         #print(ksgn)
         mx1=zw_stk_anz_m01(qx,finx0,rss,ksgn);

         ds1=pd.Series(mx1,index=c10);
         ds2=ds1.T;
         df=df.append(ds2,ignore_index=True);
         df.to_csv(ftg,index=False,encode='utf8');

def zw_stk_anz_mx_all(qx,xlst):
     for fx in xlst:
         if (fx.find('Yah')>0):
             rss=qx.rZWusDay
         else:
             if (fx=='inx_code'):rss=qx.rZWcnXDay
             else:rss=qx.rZWcnDay

         finx0=fx;
         zw_stk_anz_mx(qx,finx0,rss);
```

```
#============main

qx=zw.zwDatX(zw._rdatCN);

uslst=['inxYahoo30sp','inxYahoo100ns','inxYahoo100sp','inxYahoo6
zw_stk_anz_mx_all(qx,uslst)

cnlst=['inx_code','stk_sz50','stk_hs300','stk_zz500','stk_code',
zw_stk_anz_mx_all(qx,cnlst)
```

运行后，会在“zw_k10\tmp”目录下根据相应的股票代码文件名，生成对应的输出数据文件，全部是CSV格式。

为方便分析，我们缩短相关文件名，并复制到dat目录下“\zwpython\zw_k10\dat\”。

案例1-2，“一月效应”代码较长，为了便于初学者学习，我们将对关键代码，添加流程图，进行说明。

案例1-2程序代码，可分为三大模块：

- 文件头：import模块库库文件导入
- 函数定义：定义相关的函数
- 主控程序：注解main后面的语句

我们先看看案例1-2的主程序部分代码：

```
#============main
qx=zw.zwDatX(zw._rdatCN);

uslst=['inxYahoo30sp','inxYahoo100ns','inxYahoo100sp','inxYahoo6
zw_stk_anz_mx_all(qx,uslst)

cnlst=['inx_code','stk_sz50','stk_hs300','stk_zz500','stk_code',
zw_stk_anz_mx_all(qx,cnlst)
```

如图1-12所示为主程序流程图，主程序很简单：

- 初始化变量qx
- 先设置美股股票代码索引文件参数变量uslst，调用分析函数：zw_stk_anz_mx_all
- 设置中国A股股票代码索引文件参数变量cnlst，调用分析函数zw_stk_anz_mx_all

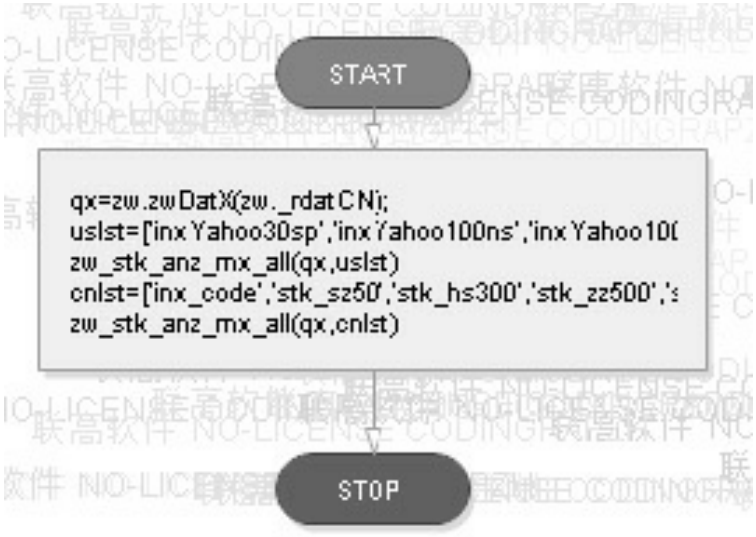


图1-12　流程图

案例1-2，“一月效应”的关键数据变量df，在函数zw_stk_anz_mx定义中：

```
def zw_stk_anz_mx(qx,finx0,rss):
     c10=["finx","ksgn","nstk",'nSum','nAdd','nDec','kAdd','kDec
     df=pd.DataFrame(columns=c10);
```

这里预定义变量df使用的是Pandas（潘达思）数据分析模块库中的DataFrame类型变量，而不是用传统的Python语言内置list（列表）类型变量。这样做是为了便于使用Pandas作为统一的数据格式，以及方便追加新的数据，代码如下：

```
ds1=pd.Series(mx1,index=c10);
ds2=ds1.T;
df=df.append(ds2,ignore_index=True);
```

同样，在下面的函数zw_stk_anz_m01中，技巧性地使用了“字典”这种数据格式作为数据变量，而不是直接使用传统的简单变量。因为Pandas数据分析模块库中的DataFrame、Series数据格式预设的就是字典模式。在量化程序中，变量使用字典模式，在与Pandas耦合方面会方便很多。其他有关细节，在后续相关章节将结合具体的图表进行讲解。

```
def zw_stk_anz_m01(qx,finx0,rss,ksgn):
     mx1={};mx1['finx']=finx0;mx1['ksgn']=ksgn;
     mx1['nSum']=0;mx1['nAdd']=0;mx1['nDec']=0;
```

1.2.3 案例分析：“一月效应”图表分析

有关代码与案例1-2类似，只是增加了图表输出代码，具体代码请大家参看脚本文件：\zwpython\zw_k10\k102m1dr.py。

在本节源码文件中，请注意下面的语句：

```
df2.plot(kind='bar',colormap='hot',rot=0,figsize=(20,5));
```

代码运行结果如图1-13所示。因为运行图较长，所以在plot绘图语句中特意使用了以下设置：

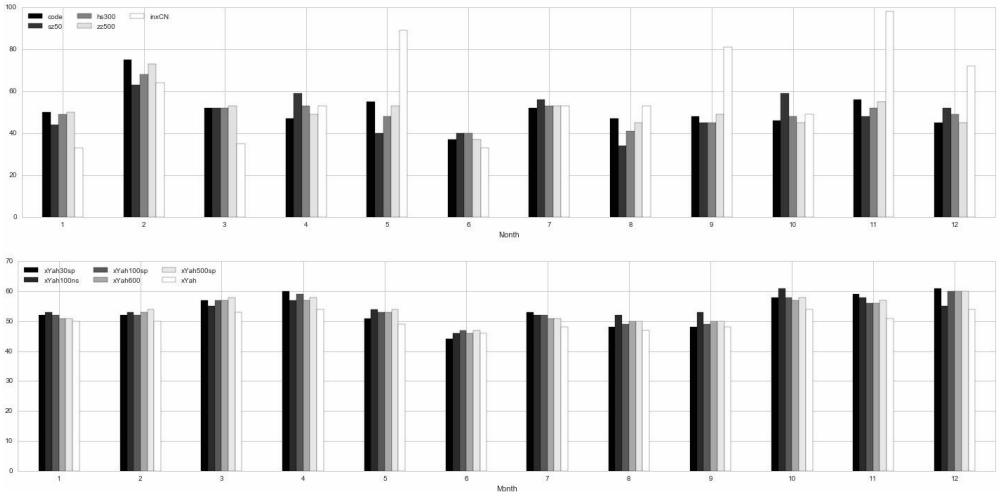


图1-13　股市“一月效应”运行结果

- figsize=(20,5)，把图像尺寸设置为长方形；
- 用rot=0把坐标显示出来，设置为无旋转；
- colormap='hot'，使色彩分隔明显。

此外，我们还使用了以下语句：

- plt.legend(ncol=3,loc=2)，图标采用3列模式，位置是左上角；
- plt.tight_layout()，紧缩四周空白，扩大绘图面积。

另外，请注意，为方便分析，我们缩短相关文件名，并复制到dat目录下“F:\zwPython\zw_k10\dat\”。

在图1-13中，上半部分为“一月效应”中国A股分析结果，下半部分为“一月效应”美股分析结果。

“一月效应”中国A股数据表如图1-14所示，对应的数据文件为\zwpython\ zw_k10\dat\mx_cn.csv。

	A	B	C	D	E	F
1	Month	code	sz50	hs300	zz500	inxCN
2	1	50	44	49	50	33
3	2	75	63	68	73	64
4	3	52	52	52	53	35
5	4	47	59	53	49	53
6	5	55	40	48	53	89
7	6	37	40	40	37	33
8	7	52	56	53	53	53
9	8	47	34	41	45	53
10	9	48	45	45	49	81
11	10	46	59	48	45	49
12	11	56	48	52	55	98
13	12	45	52	49	45	72

图1-14 “一月效应”中国A股数据表

由图1-13和图1-14可以看出，“一月效应”因为农历春节的影响具有滞后效果，到二月份，60%~70%的股票都是上涨的（二月）

另外，从图1-14中可以看出，inxCN混合指数的股票，在五月、九月、十一月三个月份也是上涨的，关联度远远高于传统的“一月效应”。可能是五一、国庆、元旦假期的黄金周效应，读者可以进一步研究分析。股票分析结果也是70%~80%为正收益，这是一个新策略，可以称为“A股黄金周策略”。

传统的算法、建模过程也和这个流程类似。对相关数据进行统计分析、数据挖掘，然后对高于K值的算法、框架进行验证整理和回溯测试，从而成为实战当中的操盘策略。

本节的案例分析因为是入门课程，所以只考虑了上涨和下跌两种情况，具体的上涨和下跌幅度没有考虑。在实际建模时，必须考虑这些因素，并且对相关指数对应的股票进行迭代分析，计算多个时间周期的盈利率。

若股票90%是上涨的，10%是下跌的，但90%的上涨幅度之和低于10%的下跌幅度之和，那么总收益还是负的。

“一月效应”美股数据表如图1-15所示，对应的数据文件为\zwpython\zw_k10\dat\mx_us.csv。

	A	B	C	D	E	F	G
1	Month	xYah30sp	xYah100ns	xYah100sp	xYah600	xYah500sp	xYah
2	1	52	53	52	51	51	50
3	2	52	53	52	53	54	50
4	3	57	55	57	57	58	53
5	4	60	57	59	57	58	54
6	5	51	54	53	53	54	49
7	6	44	46	47	46	47	46
8	7	53	52	52	51	51	48
9	8	48	52	49	50	50	47
10	9	48	53	49	50	50	48
11	10	58	61	58	57	58	54
12	11	59	58	56	56	57	51
13	12	61	55	60	60	60	54
14							

图1-15　股市“一月效应”美股数据表

从图1-15中可以看出，美股“一月效应”最强的是十二月份，可能和圣诞假期有关系。

从以上数据分析图表我们可以看出，“一月效应”的确存在，但并非指的就是“一月”，而是“二月”（中国A股）与“十二月”（美股）。

通过不同的指数组合可以发现，大企业等蓝筹股的“一月效应”更加明显，中小企业创业板股票相对较弱。

因此，对于各种投资策略绝对不能生搬硬套，要具体分析，根据实际分析的结果进行决策。

1.2.4　案例分析：颜色表效果图

本节代码请参看脚本文件\zwpython\zw_k10\k102m1cmap.py。

本节程序源码，在函数dr_cmap中，增加了一行绘图语句：

```
plt.axhline(50, color='r');
```

上面的绘图语句是在Y轴数值为50的地方绘制一条红色分割线表示中线。

本节案例使用colormap表现较弱，因此使用柱形图来展现相关效果，部分运行效果如图1-16所示。

在图1-16中，可以挑选自己喜欢的颜色风格。

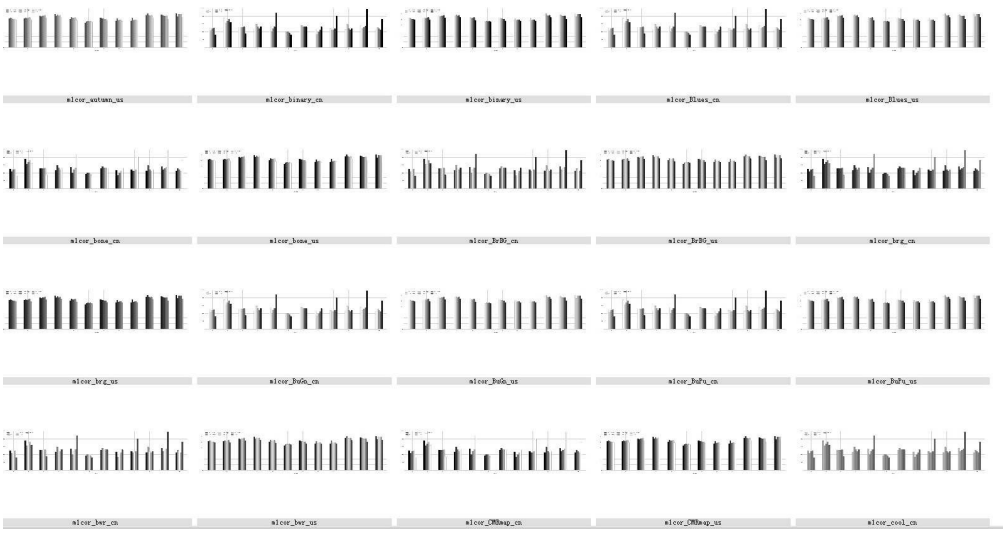


图1-16　股市“一月效应”柱形图效果

此外，在目录“x:\zwPython\zw_k10\pic\cormap”中也有相关的效果图片。

常用的颜色表有Accent、brg、coolwarm、Dark2、rainrow、gnuplot、hot、hsv、jet、prism、raibow和Set1。

1.2.5 “一月效应”全文注解版Python源码

“一月效应”全文注解版Python源码是笔者举办的“量化培训班”学员（昵称：“我怎能不戏猫”）在课件源码的基础上，根据自己的学习过程所做的修改。

这个全文注解版本几乎对每一行的代码都进行了详细的中文解释。

这种注解版本非常少见，不但对Python量化初学者，而且对普通Python入门读者都有很高的参考价值，故收录于此。

脚本文件名为\zwpython\zw_k10\k102m1_note.py。

代码如下：

```
# -*- coding: utf-8 -*-
#   zw量化开源团队
#   中文注释“我怎能不戏猫”（QQ： 316894075）
#
```

```
import sys
import pandas as pd
import numpy as np


import zwSys as zw  #::zwQT
import zwTools as zwt

#----------code


def zw_anz_m1sub(xcod,rss,monStr):#kstr表示月份
     fss=rss+xcod+".csv";print(fss) #文件名
     nSum,nAdd,nDec=0,0,0 #输入的月份数，其中上升的月份，其中下跌的月份
     kmon=int(monStr); #当前月     print('@m1sub',kstr,fss)
     try:
         df = pd.read_csv(fss,index_col=0,parse_dates=[0],encodi
         df =df.rename(columns={'Close':'close'});df =df.sort_in
         _tim0=df.index[0];_ynum0=_tim0.year; #解释时间模式，yy/mm/
         _tim9=df.index[-1];_ynum9=_tim9.year+1; #最后一年+1
         #print('@t',_tim0,_tim9)
         for ynum in range(_ynum0,_ynum9): #遍历所有年份
             ystr=str(ynum);last_day=zwt.lastDay(ynum,kmon);#年份
             dayStr='%02d'%last_day
             monStr1=''.join([ystr,'-',monStr,'-1'])        #当前月
             monStr9=''.join([ystr,'-',monStr,'-',dayStr])#当前月
```

```
            df2=df[(df.index>=monStr1)&(df.index<=monStr9)]; #选
            #print('@y',ystr1,ystr9,ystr,len(df2))
            if (len(df2)>0): #若存在交易日（处理月份用）
                _kmon5='%02d' %df2.index[0].month; #选取交易日期中
                if (_kmon5==monStr):  #若上述月份为函数输入的变量
                    xd1a=df2.ix[0];xd1z=df2.ix[-1];nSum+=1; #交
                    vd1a=xd1a['close'];vd1z=xd1z['close']; #选取
                    if (vd1z>vd1a):nAdd+=1 #比较收盘价位，判定升跌
                    else:nDec+=1;

    except IOError:
        pass;    #skip,error

    print('nSum,nAdd,nDec,',nSum,nAdd,nDec);
    return nSum,nAdd,nDec    #返回值为交易月份数量，上升，下跌

def zw_stk_anz_m01(qx,finx0,rss,ksgn): #对每只股票运算上一个函数
    fss = qx.rdatInx+finx0+".csv";   #stk_code.csv,inxYahoo.csv
    print('f',fss)
    dinx = pd.read_csv(fss,encoding='utf-8')  #读取CSV文件
    print(dinx.head())
    print('f2',fss)
    mx1={};mx1['finx']=finx0;mx1['ksgn']=ksgn;
    mx1['nSum']=0;mx1['nAdd']=0;mx1['nDec']=0;#字典，赋值
    #nSum=0;nAdd=0;nDec=0;
    xn9=len(dinx['code']);mx1['nstk']=xn9; #所读取的CSV文件的行数
    #遍历CSV文件中的code名，i 是计数器变量
```

```
    for i,xcod in enumerate(dinx['code']):
        if (not isinstance(xcod,str)):xcod="%06d" %xcod;

        dSum,dAdd,dDec=zw_anz_m1sub(xcod,rss,ksgn);

        mx1['nSum']=mx1['nSum']+dSum;
        mx1['nAdd']=mx1['nAdd']+dAdd;
        mx1['nDec']=mx1['nDec']+dDec;
        print(i,'/',xn9,xcod,mx1);
    #
    print('xn9',xn9)    ;#    print(len(mx1['nAdd']))
    mx1['kAdd']=np.round(mx1['nAdd']*100/ mx1['nSum']); #指数上升
    mx1['kDec']=np.round(mx1['nDec']*100/ mx1['nSum']); #指数下降


    return mx1

def zw_stk_anz_mx(qx,finx0,rss): #生成一个CSV文件
    c10=["finx","ksgn","nstk",'nSum','nAdd','nDec','kAdd','kDec
    df=pd.DataFrame(columns=c10); #定义dataframe
    ftg="tmp\\mx_"+finx0+".csv";print(ftg)  #打印CSV文件名

    for i in range(12):
        ksgn="%02d" %(i+1);
        #ksgn=str(i+1);#if i<9:ksgn='0'+ksgn; #1到12月

        #print(ksgn)
```

```
        mx1=zw_stk_anz_m01(qx,finx0,rss,ksgn); #利用上一个函数生成

        ds1=pd.Series(mx1,index=c10); #生成一个pandas中的series
        ds2=ds1.T; #.T=转置（矩阵转置）
        df=df.append(ds2,ignore_index=True);#在df中加上ds2
        df.to_csv(ftg,index=False,encode='utf8'); #保存为CSV文件,


def zw_stk_anz_mx_all(qx,xlst):    #遍历指定list中的股票
    for fx in xlst:
        if (fx.find('Yah')>0):
            #rss=qx.rZWusDay
            rss=qx.rdat+'\\us\\day\\'
        else:
            if (fx=='inx_code'):rss=qx.rdat+'\\cn\\xday\\' #rss
            else:rss=qx.rdat+'\\cn\\day\\' #rss=qx.rZWcnDay

        finx0=fx; #生成文件名
        zw_stk_anz_mx(qx,finx0,rss); #用上一个函数生成CSV文件


#============main
qx=zw.zwDatX(zw._rdat0);


uslst=['inxYahoo30sp','inxYahoo100ns','inxYahoo100sp','inxYahoo6
zw_stk_anz_mx_all(qx,uslst)
```

```
cnlst=['inx_code','stk_sz50','stk_hs300','stk_zz500','stk_code',
zw_stk_anz_mx_all(qx,cnlst)
```

案例“一月效应”主程序很简单，如图1-17所示为主程序流程图。

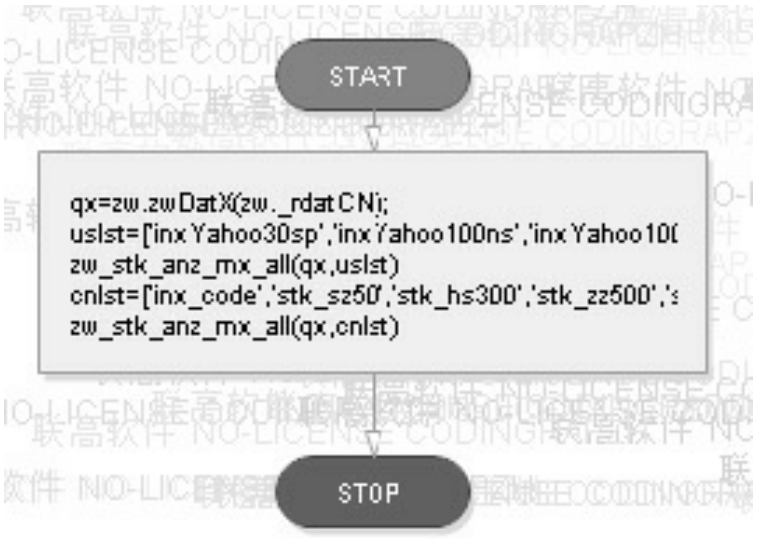


图1-17　主控程序流程图

- 初始化变量qx
- 先设置美股股票代码索引文件参数变量uslst，调用分析函数：zw_stk_anz_mx_all
- 设置中国A股股票代码索引文件参数变量cnlst，调用分析函数zw_stk_anz_mx_all

案例“一月效应”，重点在于几个自定义函数：

- zw_stk_anz_mx_all(qx,xlst)：遍历指定xlst中的股票
- zw_stk_anz_mx(qx,finx0,rss)：根据finx0股票指数文件，统计分析12个月各个月的具体数据

- zw_stk_anz_m01(qx,finx0,rss,ksgn)：根据finx0股票指数文件，分析单个月的数据
- zw_anz_m1sub(xcod,rss,kstr)：根据xcod具体股票代码，分析单只股票单个月的数据

为了便于初学者学习，我们分别绘出以上几个函数的流程图。

函数zw_stk_anz_mx_all(qx,xlst)，遍历指定xlst中的股票，流程图如图1-18所示。

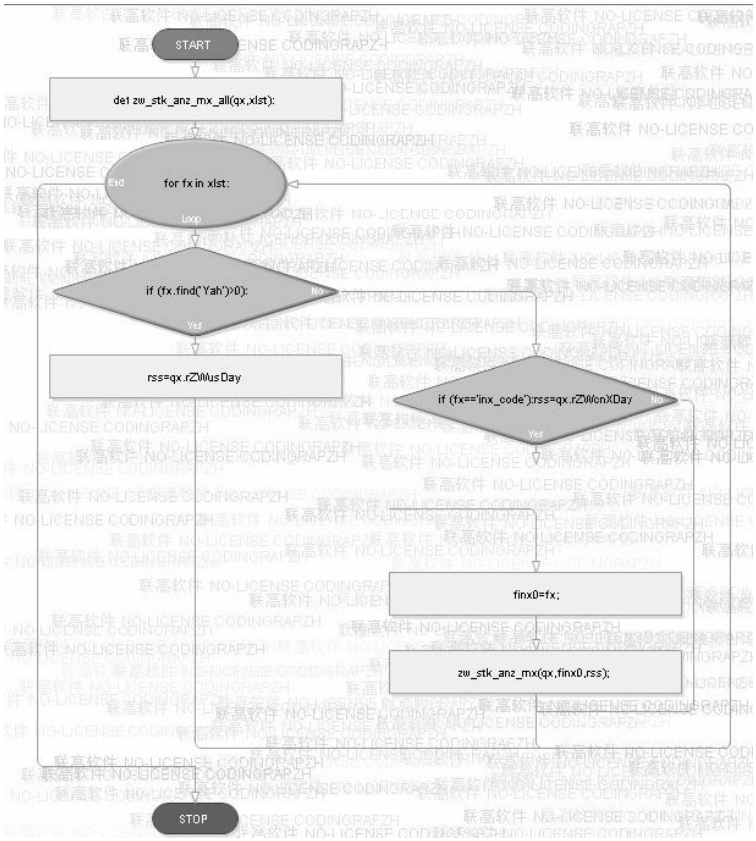


图1-18　函数zw_stk_anz_mx_all流程图

函数zw_stk_anz_mx(qx,finx0,rss)，根据finx0股票指数文件，统计分析12个月各个月的具体数据，流程图如图1-19所示。

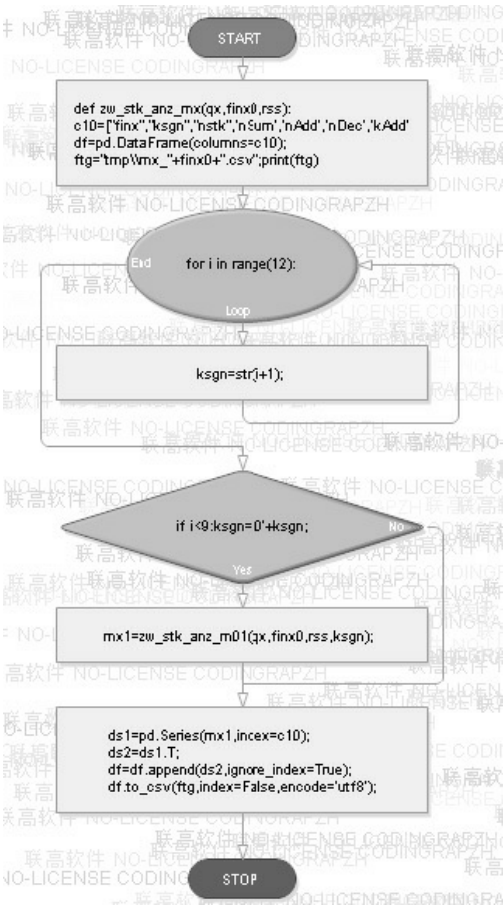


图1-19　函数zw_stk_anz_mx流程图

函数zw_stk_anz_m01(qx,finx0,rss,ksgn)，根据finx0股票指数文件，分析单个月的数据，流程图如图1-20所示。

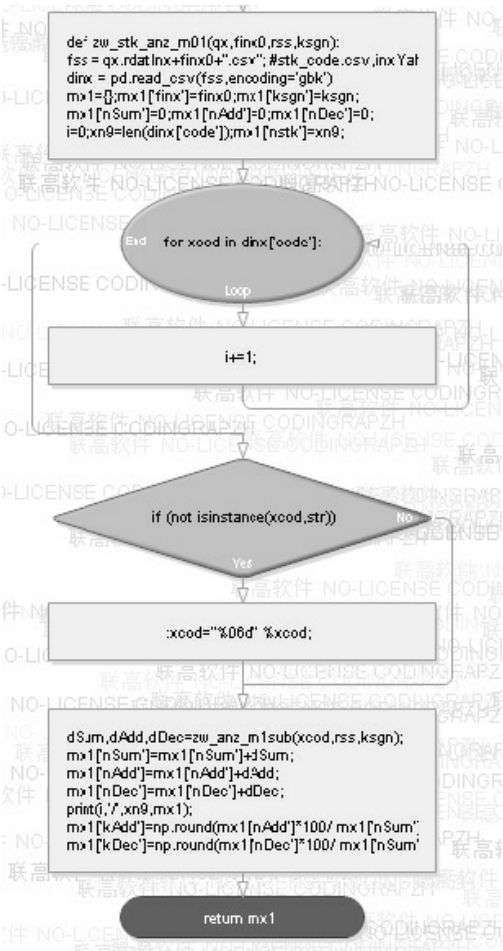


图1-20　函数zw_stk_anz_m01流程图

函数zw_anz_m1sub(xcod,rss,kstr)，根据xcod具体股票代码，分析单只股票单个月的数据，因为这个函数比较复杂，我们进行了简化，流程图如图1-21所示。

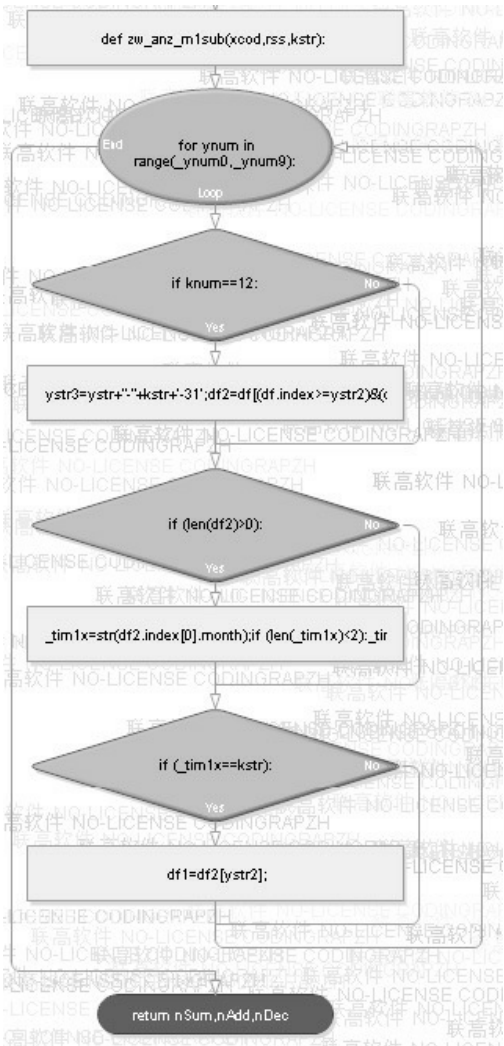


图1-21　函数zw_anz_m1sub流程图

1.2.6　大数据·宏分析

下面对1.2节的内容进行小结，本节主要讲解了以下内容：

- 股市“一月效应”；
- 股票指数代码表；
- “一月效应”量化分析；
- “一月效应”图表分析；

- 颜色表效果图；
- “一月效应”全文注解版。

从“一月效应”数据分析图表中我们可以看出，“一月效应”的确存在，但并非指的就是“一月”，而是“二月”（中国A股）与“十二月”（美股）。

因此，对于各种投资策略绝对不能生搬硬套，要具体分析，根据实际分析结果进行决策。

笔者在新浪博客《大数据·实战个例“宏”分析》中，有过专门阐述，博客截图如图1-22所示。



图1-22 《大数据·实战个例“宏”分析》博客截图

“尽信书，不如无书。”

对于Python量化，全世界都是零起点，都处于起步阶段。在这个阶

段，读者要充分发挥自己的主观能动性，任何权威、专家在最终的数据结果面前都是乏力的。

更何况，对于Python量化，目前整个行业没有任何权威。现在比的是大局观，看谁的方向更正确；比的是速度，看谁的动作更快，最终把传统的投资策略应用到Python量化框架中。

1.3　量化交易流程与概念

1.3.1　数据分析I2O流程

前面通过几个具体的案例讲述了量化分析的基本形态，下面介绍一些相关的基本概念。

对于量化，大部分行业教程和软件系统的名称都是量化交易。

IT行业和数据分析领域传统的流程划分模式如图1-23所示。



图1-23　I2O模式流程图

这个流程套用目前流行的网络名词，我们称之为I2O模式。

其中，“I”是输入，这个输入不仅是输入数据，也可以是用户的需求或者计算机的输入设备。

“O”是输出，同样，这个输出不仅是输出数据，也可以是输出相关的作品，或者是电脑的显示器、打印机。

“2”是“to”的谐音字符，在I2O模式中，最重要的环节就是“2”（to），即分析处理，类似计算机的CPU。

I2O模式不仅适用于IT行业和数据分析，实际上大部分行业和日常

工作都可以采用。

对于量化交易而言，其中的“I”就是数据采集，TuShare数据抓取模块和pandas数据分析模块的IO模块都是处理金融数据抓取的。

量化交易中的环节“O”就是trade（交易），即下单环节。PyAlgoTrade等量化软件都有专门的模块处理trade（交易）事务。

量化交易中的环节“2”（to）就是量化分析、Strategy（策略）分析，是其中最重要的环节。

Strategy（策略）是量化交易的核心，也是一切交易的核心。

量化的核心是策略分析，就是通过分析金融数据评估各种交易模型，建立各种交易策略。

策略分析相当于计算机的CPU、汽车的发动机。有了好的策略，就像美剧《疑犯追踪》中的计算机，哪怕只是提供一个社保号码，剩下的环节即使采用最古老的人工模式，也足以完成全部的交易过程。

对于大型金融机构而言，Strategy（策略）和Trade（交易）分离的好处就是保密，防止企业的核心商业机密和交易策略泄露。

1.3.2 量化交易不是高频交易、自动交易

很多人一谈到量化交易，认为就是程序化交易，计算机自动下单，这种观点其实是错误的。事实上，除了以秒、毫秒计算的高频交易，一

般的日内交易，甚至基于5分钟、15分钟等“分时数据”的中高频交易，都可以采用人工下单的模式。

量化交易与高频交易、程序化（自动）交易是完全不同的概念，但是初学者经常混淆。

量化交易的核心是策略分析，通过对历史数据、实时数据的分析，选择最佳的交易股票（期货、外汇等金融产品）的品种，以及交易（买进、卖出）的时间点，也就是传统的择股交易与择时交易。

高频交易是指从那些人们无法利用的、极为短暂的市场变化中寻求获利的计算机化交易。例如，某种证券买入价和卖出价差价的微小变化，或者某只股票在不同交易所之间的微小价差。

自动交易起源于美国1975年出现的“股票组合转让与交易”，随着技术的发展和计算机系统的应用，投资经理、经纪人可以实现股票组合的一次性买卖交易；20世纪80年代后被广泛应用于期货期权交易。

自动交易软件就是让计算机按照用户事先设定好的条件自动交易，当然赢亏结果取决于用户的交易计划设计得好坏。

这样，投资者就不用把主要精力与时间放在人工盯盘与手动操作上，可以有充分的时间不断完善自己的交易计划，然后让计算机自动执行。

金融市场的核心还是策略（Strategy），有了优秀的策略，至于使用什么方式下单，如计算机自动下单、人工下单，甚至电话委托、手机APP下单，只是流程问题。

不想每天人工下单，自己加一个自动下单的模块就可以。我国因为

政策限制，程序化交易API目前还没有开放，因此很多人采用外挂模式。

读者可以在运行回溯分析程序后，根据自己设定的交易策略，筛选出股票（期货）品种代码，以及买进和卖出的数量，编写“数据交换文件”，再运行交易模块，按照“数据交换文件”里面的股票（期货）代码和交易数量，由计算机替代人工完成下单的操作。

以zwQuant量化软件为例，自动交易模块可以加在回溯函数之后，也可以是独立的交易模块。

1.3.3 小资、小白、韭菜

在网络公开课中，笔者曾经说过：“小资、小白做股票，往往是韭菜的命”。

以“一月效应”为例，虽然没有计算具体的盈利率，不过一般情况下，可以假设盈利率与上涨率是正相关的。

依据假设，“一月效应”的确是一个正确的投资策略。小资、小白使用“一月效应”投资策略，往往只看到“一月效应”的表象，于是在一月初重仓“扫货”，到了月底“出货”。

但是，根据图1-14和图1-15两张数据表可以看出：

- 一月美股上涨的平均概率在51%左右；
- 一月中国A股上涨的概率在45%左右，混合指数股上涨的概率更是低至33%。

如果像小资、小白一样只看表面投资，那么这样投资操盘无疑是亏损的。

1.3.4 专业与业余

小资、小白做股票亏损居多，因为他们不专业。任何行业，只有专家才是赢家。

读者学习量化，也是希望尽快成为量化行业的赢家，这就需要具有职业化素质。

要做到职业化，通常需要做到以下三点：

- 了解行业的游戏规则；
- 熟悉本行业专业的软件工具和配套资源（例如zwQuant量化软件、zwDat股票数据包）；
- 了解行业背景知识。

职业化只是第一步，只是入门。常言道，“师傅领进门，修行在个人”，说的也是这个道理。

真正的赢家，必须是专家，超出职业化、专业化这两个阶段。要成为赢家，就必须采用比行业平均更高的水准来要求自己。

以zwPython开发平台为例，笔者曾经说过，zwPython是按工业级的标准制作的，可以直接应用到生产环境。这句话是什么意思呢？

做过IT企业CTO（技术总监）的人都知道，工业级的标准就是系统

可以承受每天24小时、每周7天不间断的连续工作。

在本书的创作过程中，几次批量下载A股数据，都近似24小时的压力测试，以及长达10天的中国A股历年的tick数据包下载，zwPython开发平台每次都顺利通过了，完全通过了7×24的实战测试。

zwDat股票数据包是目前行业较为专业的日线数据包，之所以这么说，原因如下。

- 地区覆盖广：同时包含了中国A股、美股纽约交易所、纳斯达克日线数据，未来还会增加中国香港、日本、欧洲等国家和地区的股票数据。
- 时间跨度大：中国A股是从1994年开市起收集的，美股是从1962年有数字化档案开始收集的。
- 品种全：基本上收录了全市场所有的活跃股票交易数据。

未来，如果找到合适的免费数据源API，zwDat股票数据包会扩展到期货、外汇领域，升级为zwDat金融数据包。

zwDat股票数据包比传统的股票数据包更加专业。

即使是现在，只有中国A股、美股份个历史数据的情况下，我们也能完成一些超出目前行业标准的工作。

以“一月效应”为例，“一月效应”其实是一个最简单的Backtest（回溯）测试。

在前面的章节中，把项目扩展到12个月的测试，绘出最终的结果图，省略了中间的回溯曲线图，如图1-24所示。

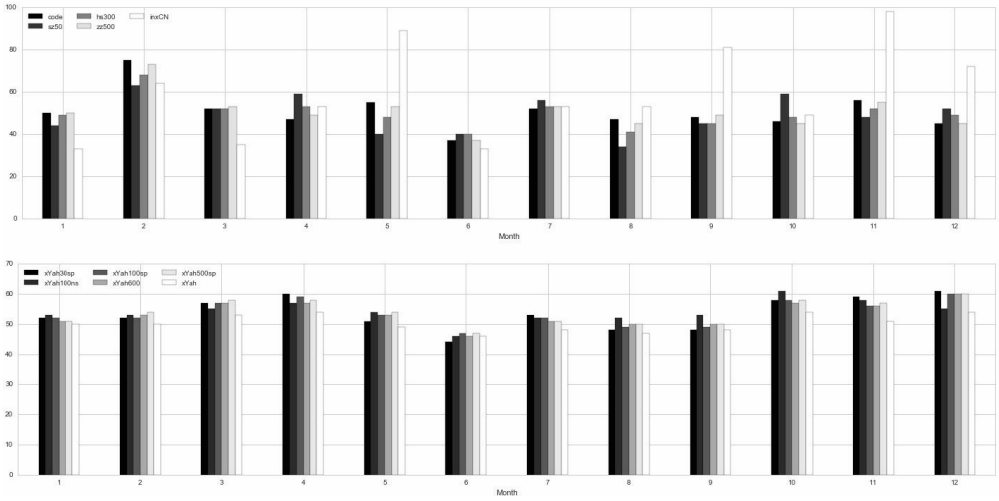


图1-24 “一月效应”统计图

本书第2章将会介绍一个完整的SMA（简单均线）量化策略，这个量化策略源自PyAlgoTrade（简称PAT，读者群、网盘共享有中文手册，读者可以自行下载）。

如图1-25所示是SMA策略收益回溯图。



图1-25 SMA策略收益回溯图

如图1-26所示是目前量化行业标准的SMA策略收益回溯图，很多在

线量化平台也是这种模式，其中的技术细节在后面的章节会具体介绍。

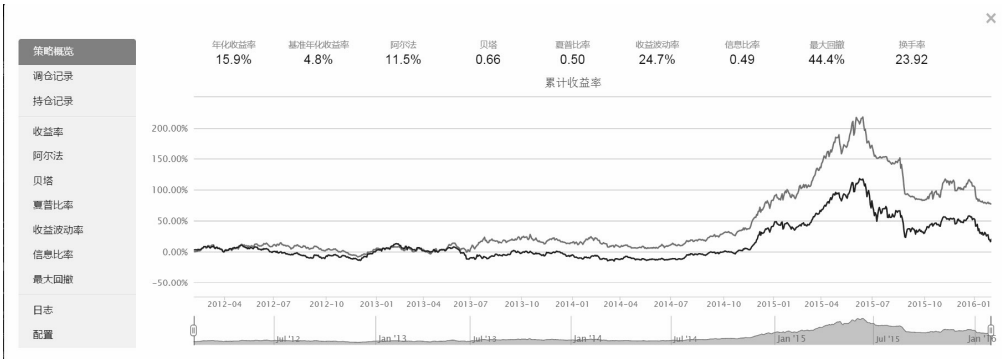


图1-26　量化行业标准的SMA策略收益回溯图

这种回溯测试只是单一周期、单一地区的回溯测试。zwQaunt量化分析软件因为只做策略模块，所以对于策略的回溯测试，一开始要求就比行业标准更高。

zwQuant量化分析软件采用的是MMM模式，是多周期、多地区、多频率的回溯测试，即Mul-Mul-Mul Backtest。

这种MMM模式部分源自笔者多年数据分析的实盘测试经验。

zwDat股票数据包有中国A股、美股两份数据，对跨区域的测试进行对比分析，已经有足够多的数据源支持。这种跨区域的测试，我们在“一月效应”里面已经使用过，增加其他地区的数据源后，这种跨区域的对比测试会更加强大。

相对滞后的是多频率回溯测试，就是对不同频率的分时数据进行回溯测试，包括1分钟线、5分钟线、15分钟线、30分钟线和60分钟线等。

目前，国内分时数据开放的API很少，而且分时数据的数据量太大，1分钟线的日数据量一般是10000组，是单只股票历年数据的3~5倍（2000~5000组）。

这部分的工作必须使用GPU工作站或者其他超算设备配合专业的实时tick数据源才能启动。

跨周期测试相对简单，对于前面两张回溯图：

- 如图1-25所示是PAT的演示截图，是2000年的单年回溯测试；
- 如图1-26所示是某个量化平台的截图，是2012~2016年的多年回溯测试。

不管是一年回溯测试，还是多年回溯测试，都只有一个时间测试周期。zwQuant量化分析软件采用的是多周期测试模式，每年一个测试轮回，以及n-1年一个测试轮回。

这种多周期测试，可以充分考察策略的稳定性，而稳定性是量化投资最重要的考量标准之一。

从图1-26所示的多年回溯测试截图中我们可以看出：同样的一个策略，2014年以前曲线是平坦的，收益也是不高的；到了2015年，变化猛然增大，7月前快速增长，7月后急速下跌。不管这个策略收益如何，在稳定性方面绝对是需要进一步调整和优化的。

1.4 用户运行环境配置

本节主要讲解zwPython和zwDat的配置、应用流程方面的知识。

本书所有案例程序均采用纯Python语言开发，除特别指明外，默认使用Python 3语法，均经过zwPython平台测试。



图1-27 zwPython用户手册封面

zwPython是极宽公司推出的一个Python集成版本，功能强大，内置了数百种专业的Python模块库；无须安装，解压即用。有关zwPython的使用，可参考本书附录A。

本书所有案例程序可用于各种支持Python 3的设备平台，包括Linux操作系统、Mac苹果电脑，以及安卓系统，甚至树莓派。

其他非zwPython用户运行本书程序，如果出现问题，通常是缺少有关的Python模块库，可以根据调试信息安装相关的Python模块库，再运行相关程序。

限于篇幅，关于Python语言和pandas数据分析软件的基本操作，请读者查看有关图书。

1.4.1 程序目录结构

本书配套程序的工作目录是zwPython\zw_k10，这个目录也是默认的工作目录，凡是没有标注目录的脚本文件，一般都位于该目录。

有关的程序会定时在读者群发布更新，请读者及时下载。

zwPython目录结构如图1-28所示。

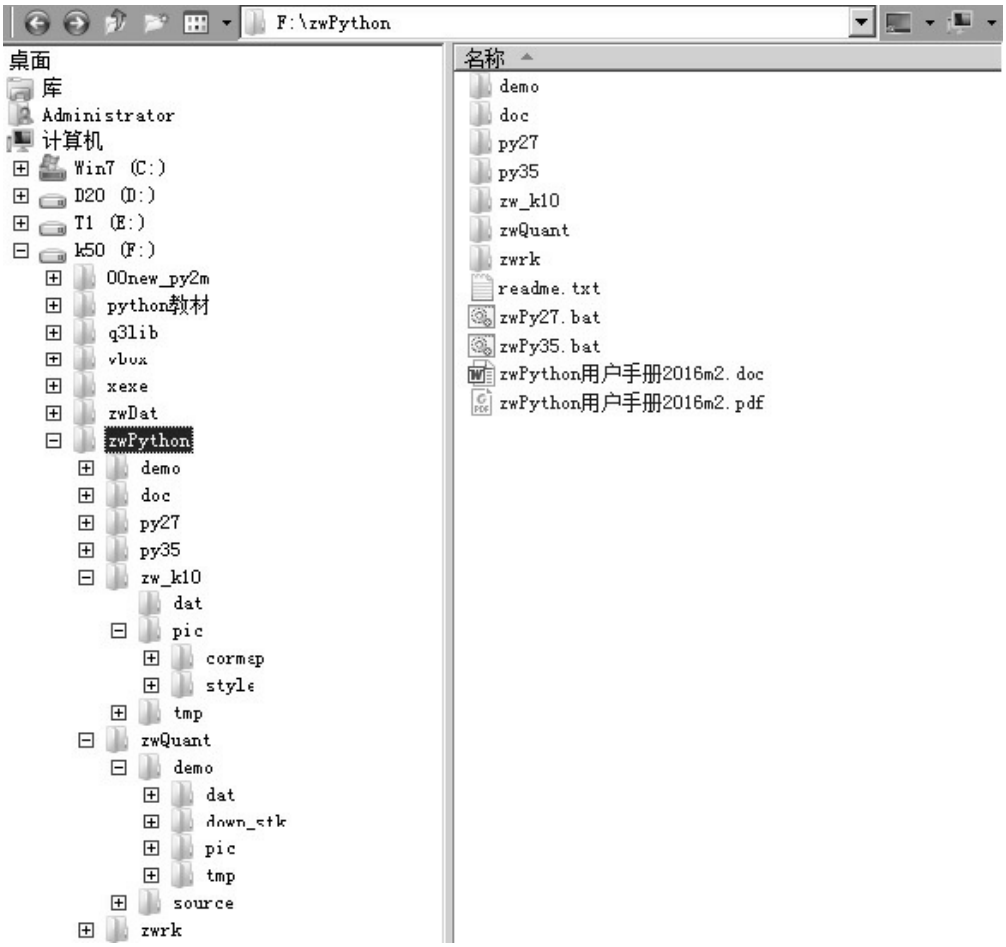


图1-28　zwPython目录结构

相比普通的Python版本，如图1-28所示的zwPython目录中多了一个zw_k10目录。

zw_k10目录收录的是相关培训课程的配套代码和所需数据。zw_k10目录也可以复制到其他目录，建议放到zwPython根目录下。

zwPython目录结构中的其他子目录如下。

- \zwPython\doc\：用户文档中心，包括用户手册和部分中文版的模块库资料，如fontforge、opencv和部分字王字库方面的文档。
- \zwPython\py27\：Python 2.7版本系统目录，除增加、删除模块库

外，一般不需要改动本目录下的文件，以免出错。

- \zwPython\py35\：Python 3.5版本系统目录，除增加、删除模块库外，一般不需要改动本目录下的文件，以免出错。
- \zwPython\demo\：示例脚本源码，包括opencv、pygame游戏设计、Pandas大数据分析、zw字王等相关的示例脚本源码。
- \zwPython\zwrk\：zw工作目录，用户编写的脚本代码文件建议放在本目录下。
- \zwPython\zwQuant\：zw开源量化工具箱和zwQuant量化分析开源软件。

zwQuant开源量化工具箱与zwPython进行了集成处理，可直接输入使用，支持Python 3.5。

移植时或使用其他Python环境时，可以把zwQuant目录下的脚本文件全部复制到自己的代码工作目录，注意zwSys.py代码里有关数据文件目录的设置。

1.4.2 金融股票数据包

1. 数据目录结构

zwDat股票数据包目录结构如图1-29所示。

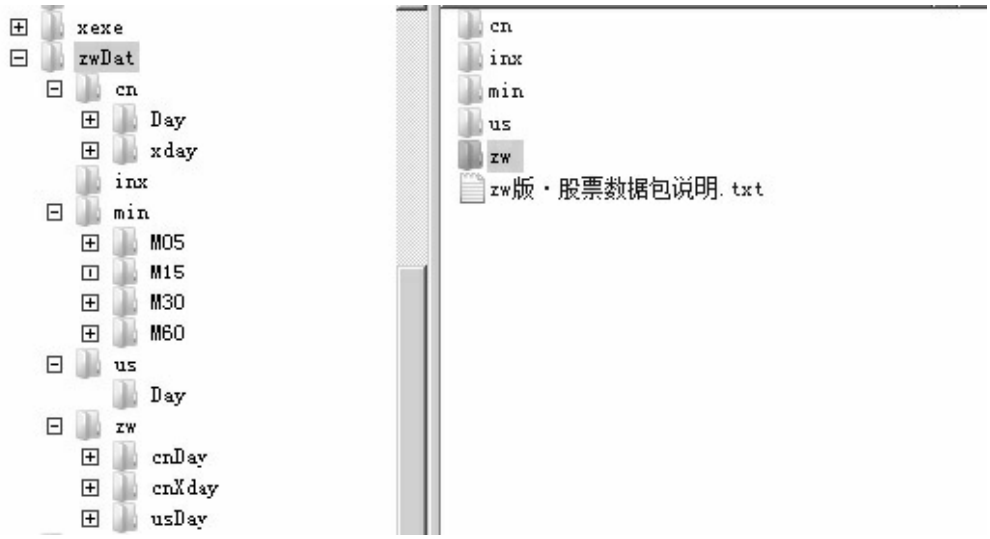


图1-29　zwDat股票数据包目录结构

其中，具体目录结构如下。

\zwDat\：zw数据源根目录。

cn\：中国股票数据。

Day\：A股日线数据。

xday\：大盘等指数日线数据。

us\：美股数据。

Day\：美股日线数据。

zw\：zw增强版数据包，包括中国A股、A股指数、美股数据。

cnDay\：zw增强版，A股日线数据。

cnXday\：zw增强版，大盘等指数日线数据。

usDay\：zw增强版，美股日线数据。

inx\：股票代码等参数数据。

min\：中国A股分时数据（仅提供部分演示数据）。

M05\：5分钟数据。

M15\：15分钟数据。

M30\：30分钟数据。

M60\：60分钟数据。

2. zwDat数据文件更新

zwDat的inx目录在首次发布后，有关股票、指数代码文件有大量更新，请读者自行下载。下载后直接覆盖原zwDat\inx目录下的文件即可。inx更新包文件名为inx.zip。

新的zwDat_cn数据包文件名为zwDat_cn_v2.7z，已经更新了inx目录下的数据，无须再下载inx更新文件包。

网盘地址为http://pan.baidu.com/s/ljIg944u。

- inx_code.csv：中国A股大盘及各种指数代码。
- stk_base.csv：中国A股2810只股票代码公司概况。
- stk_code.csv：中国A股2810只股票代码。
- stk_hs300.csv：中国沪深300指数股票代码。
- stk_sz50.csv：中国上证50指数股票代码。

- stk_zz500.csv：中国中证500指数股票代码。
- inxYahoo.csv：全部6688只美股代码。
- inxYahoo30sp.csv：道琼斯30指数美股代码。
- inxYahoo100ns.csv：纳斯达克100指数美股代码。
- inxYahoo100sp.csv：道琼斯100工业指数美股代码。
- inxYahoo500sp.csv：道琼斯500指数美股代码。
- inxYahoo600.csv：量化常用美股600股票代码。

1.5 Python实战操作技巧

1.5.1 模块检测

模块检测可运行相关脚本，具体代码请参见脚本文件：\zwpython\zw_k10\k103ver.py。

本节脚本程序用于检测量化相关的重点模块库是不是安装完毕。

本书常用的模块库与默认缩写字符串如下。

- import pandas：pd。
- import tushare：ts。
- import zipline：zp。
- import talib：ta。
- import statsmodels：sm。
- import numpy：np。
- import scipy：sp。
- import matplotlib：mpl。
- import matplotlib.pyplot：plt。
- import seaborn：sns。
- import ggplot：gpl。
- import pyalgotrade：pat。
- import zipline：zp。
- import zwSys：zw（#::zwQT）。

- import zwQTBox：zwBox。

1.5.2 Spyder编辑器界面设置

在设置界面前，把zwQuant\source\zwQTBox.py文件拖到Python编程语言编辑器Spyder的编辑框中，如图1-30所示。

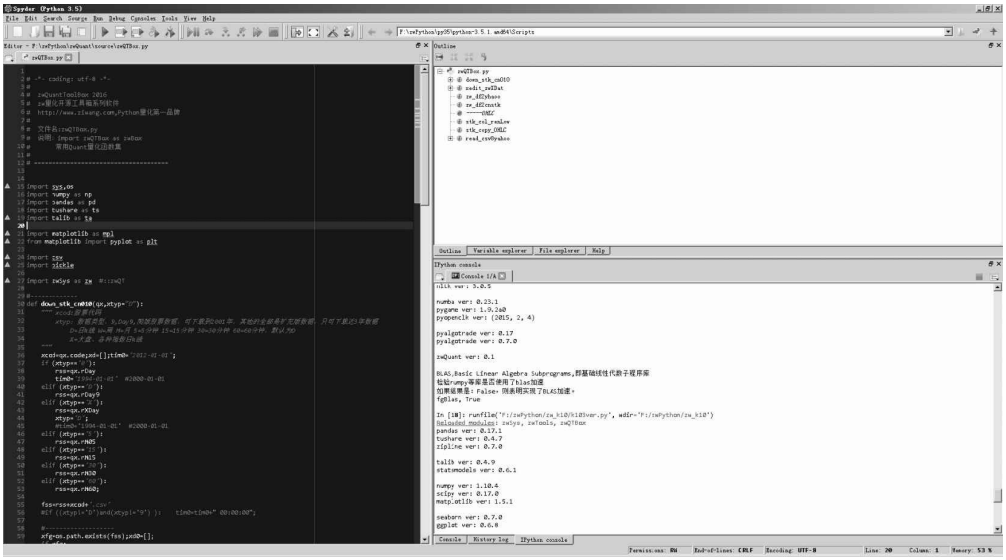


图1-30 Python编程语言编辑器Spyder编辑框界面

Spyder编辑器的界面设计非常合理，参考了MATLAB，特别适合量化分析。国际投行一般都选择这种布局作为标配。

通常需要优化的只有Outline（导航）面板，又称函数列表面板，类似Delphi语言的struct函数列表面板。

在Spyder编辑器默认配置中，Outline面板是不显示的，单击菜单View→Panes→Outlines，如图1-31所示，将显示Outline面板。

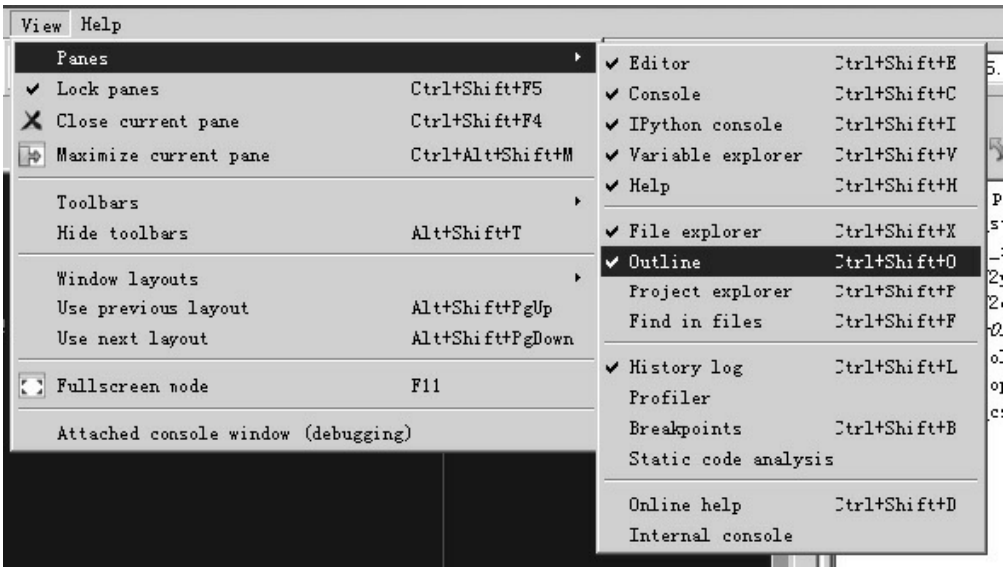


图1-31　显示Outline面板

Outline面板显示后，它的默认位置是在代码编辑器和右侧窗口的中间。

建议单击Outline面板左上角的“窗口缩放”按钮，拖动面板到右上方，将其与Var（变量）面板、File（文件）面板等合并。

Outline面板的作用是对代码中的函数、类、变量进行快速导航定位。单击Outline面板的函数、类、变量名称后，左侧代码编辑器就会自动移动到相关代码，如图1-32所示。对于大型项目而言，使用Outline面板可以提高效率。

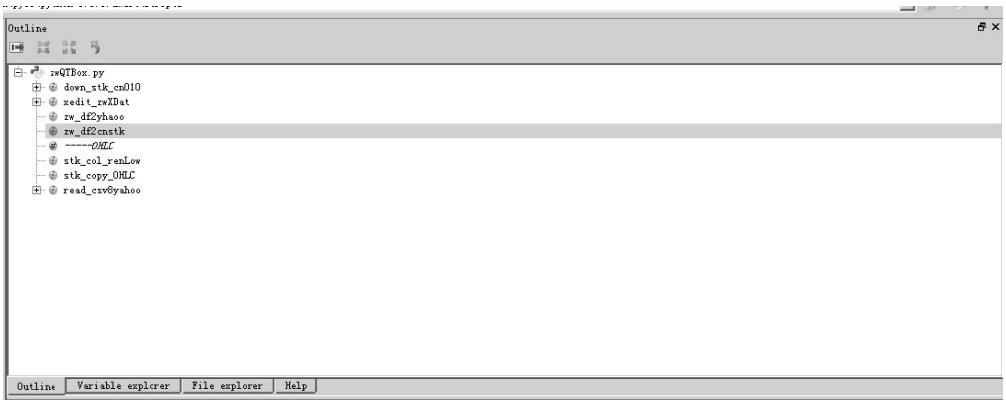


图1-32　Spyder编辑器Outline面板

1.5.3　代码配色技巧

zwPython的IDE代码编辑器是Spyder，默认配色是Spyder模式，采用白底黑字，与传统的IDE环境差别很大，如图1-33所示。

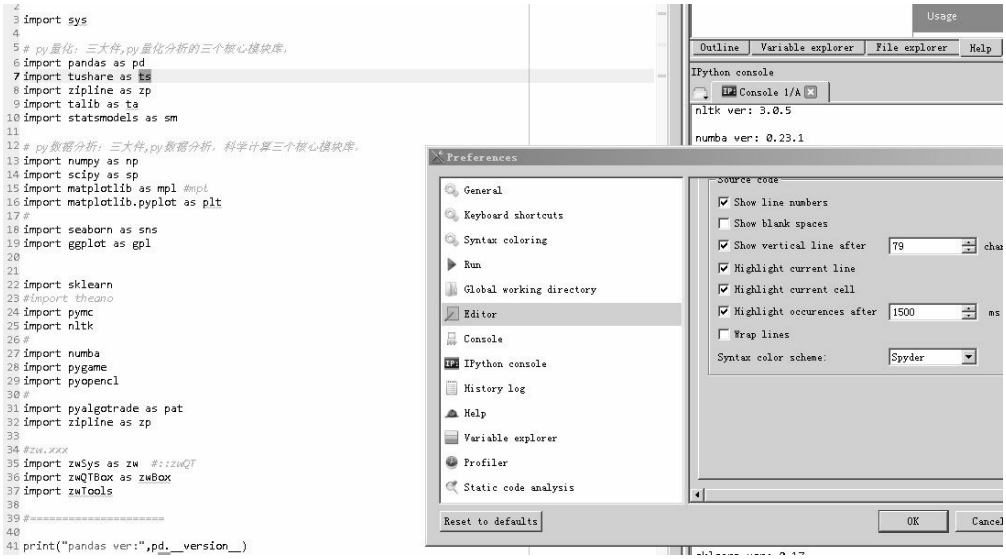


图1-33　Spyder编辑器配色模式

如图1-34所示是最新的Delphi-xe10的编辑器配色模式（Twilight模式）。

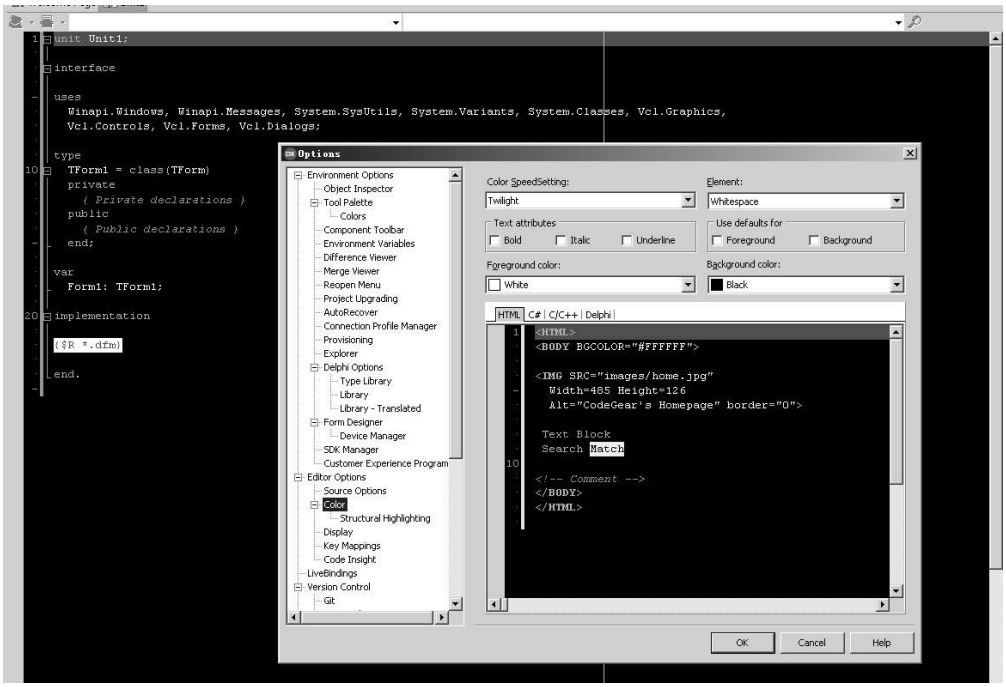


图1-34　Delphi-xe10编辑器配色模式

这种黑底模式也是微软等开发平台标准的代码编辑器配色模式。幸运的是，Spyder编辑器内置的配色模式里有类似的模式。

如图1-35所示，运行Spyder编辑器，单击菜单Tools→Preferences，打开Preferences对话框。在左侧的列表框中选择Editor（编辑器），在右侧的Syntax color scheme（语法的配色方案）下拉列表框中选择Spyder/Dark（暗调）模式即可。

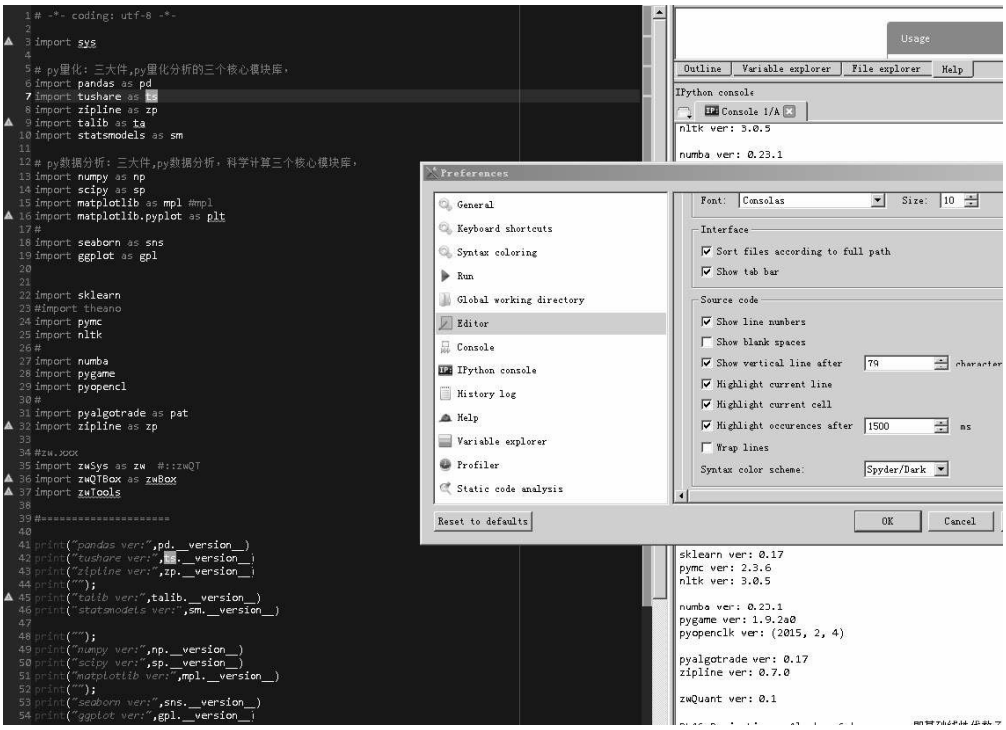


图1-35　调整Spyder编辑器配色模式

不同版本的Spyder编辑器调整细节会有所不同，请读者注意。

1.5.4　图像显示配置

Python语言的Spyder编辑器默认的图像显示尺寸对于高清显示器来说尺寸有些小，需要进行调整，具体步骤如下。

（1）单击菜单Tools→Preferences，打开Preferences对话框。

（2）单击左侧列表框中的IPython console（IPython控制台）。

（3）在对话框的右侧选择Graphics选项卡。

（4）在Graphics　backend选项区中，Backend选项默认为Inline，一

般不需要更改，如要进行交互分析，可以设置为Auto matic（自动）模式或者Qt模式。

（5）在Inline backend选项区中可以调整内置图像的大小，Width默认值为8，Height默认值为5，建议改成Width为10、Height为6。

此外，建议勾选对话框上部的Automatically load Pylab and NumPy modules复选框（会自动加载Pylab、NumPy模块），如图1-36所示。

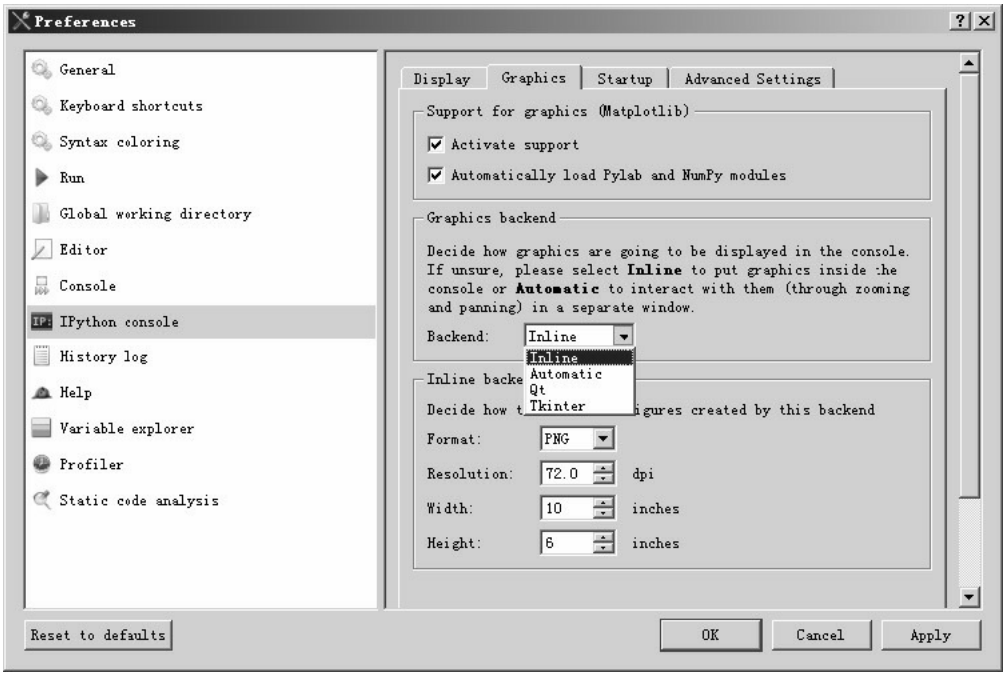


图1-36 调整Spyder编辑器图像显示尺寸

1.5.5 Python2、Python 3双版本双开模式

双开原本是网游工作室的术语，是指在一台计算机上同时打开两个客户端，通常会采用VBox等虚拟机技术。

zwPython开发平台在2016M10以前的版本中内置了Python 2.7和Python 3.x两种不同的版本。

通常很少出现两种环境同时并发工作的情况。双开模式最保险还是用VBox虚拟机，不过虚拟机安装麻烦，效率太低，偶然查询资料，用虚拟机有些大材小用。

由于很多Python量化库只支持Python 2.7，因此有时在运行Python 3.x版本的程序时，需要临时调用Python 2.7的资源，这时就需要用Python 2.7和Python 3.x两种不同版本的双开技术。

使用zwPython开发平台的双开模式很简单，分别运行py27.bat文件和py35.bat文件即可。为避免冲突，最好等一种环境完全加载完毕后，再启动另外一个版本。

如图1-37所示是双开模式运行截图。

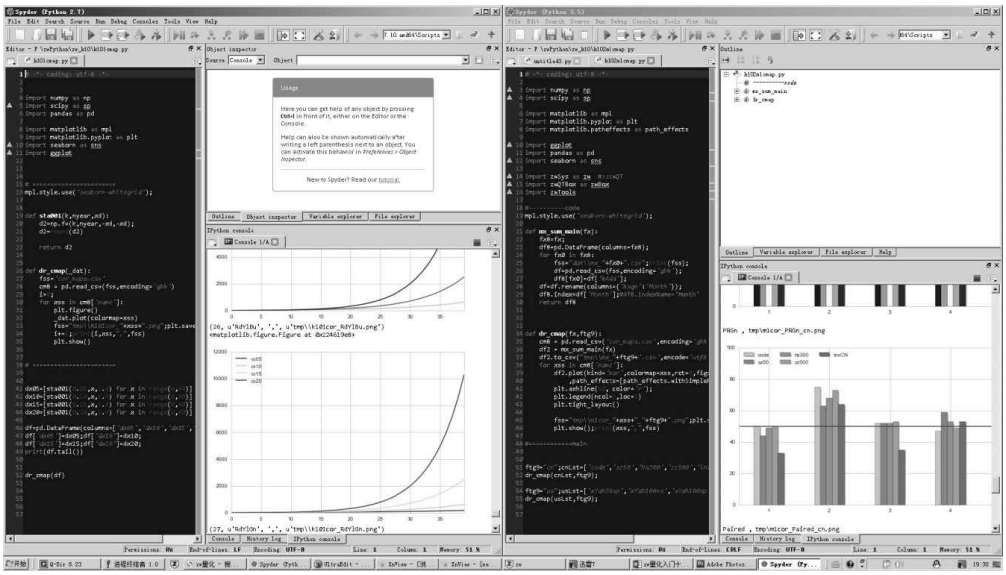


图1-37 双开模式运行截图

使用双开模式时，若显示器分辨率高，则可以左右分屏显示；若显

示器分辨率低，则只能单屏切换。在图1-37中：

- 左侧是Python 2.7版本，运行的脚本文件名是k101cmap.py；
- 右侧是Python 3.5版本，运行的脚本文件名是k102m1cmap.py。

1.5.6 单版本双开、多开模式

相比双版本双开模式，用得更多的是单版本的双开、多开模式。下面以Python 3.x为例，介绍单版本的双开、多开模式。

以colorcmap颜色表程序为例，操作步骤如下。

（1）加载k101cmap.py脚本文件，单击工具栏中的“运行”按钮或者按F5快捷键，运行第1组程序。

（2）加载k102m1cmap.py脚本文件，在代码编辑窗口选择第二个脚本文件k102m1cmap.py。

（3）用鼠标右键单击右下方iPython面板的Tab栏，在弹出的快捷菜单中选择Open a new console（建立一个新的iPython控制台），也可以按Ctrl+T快捷键，如图1-38所示。



图1-38 单版本双开模式运行设置

（4）等新的iPython控制台初始化完毕，单击工具栏中的“运行”按钮或者按F5快捷键，运行第2组程序。

单版本双开模式运行截图如图1-39所示。

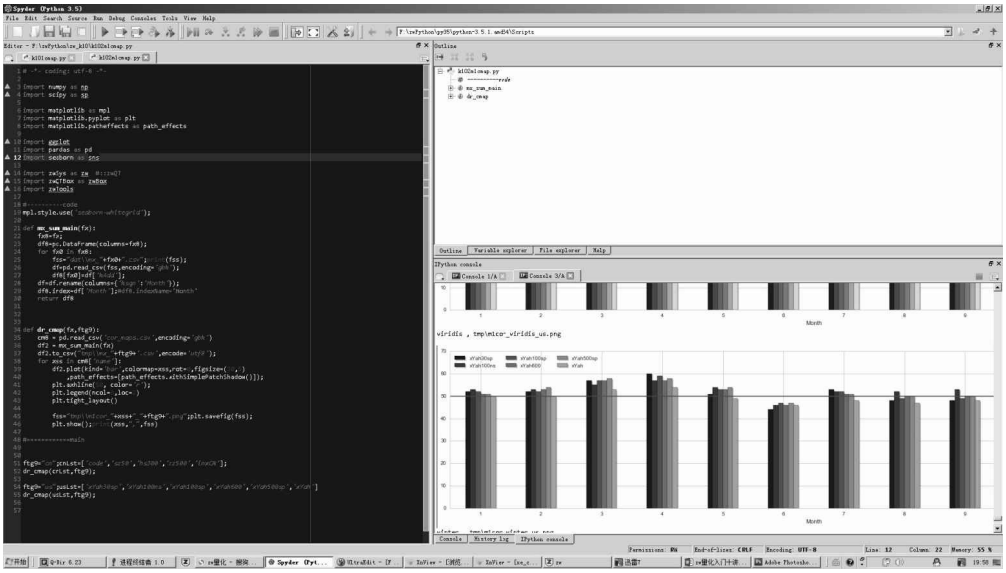


图1-39　单版本双开模式运行截图

如果有第3组、第4组程序，依次添加、运行即可。程序的具体上限没有测试，不过按经验，不要超过CPU的内核数，不然会影响效率。如果是4核的CPU，建议3开；如果是8核的CPU，建议6~7开，要留1~2个CPU内核给Windows系统。

为什么单版本双开、多开的用处比多版本用处大呢？

在数据分析时，有时需要下载大量的数据，例如下载中国A股历年的日线数据，就需要一天的时间。

这时，完全停工等待数据下载完毕，就会浪费时间。如果采用双开模式或多开模式，再开1~2个iPython控制台和代码编辑窗口编写、调试新代码，也占用不了多少CPU，却可以大大提高程序员的工作效率。这

样就使下载、编程双管齐下：前台编辑、调试新代码，后台下载数据包。

事实上，分析分时数据时，即使是5分钟数据线，几千只股票数据下载，也需要不少的时间。如果采用双开、多开模式，分析、下载同步进行，无疑效率会更高。

1.5.7 实战胜于一切

在1.5节中讲述了以下内容：

- zwPython操作技巧；
- 常用模块库与缩写；
- 界面设置；
- 代码配色技巧；
- 图像显示配置；
- Python 2.7、Python 3.x双版本双开模式；
- 单版本双开、多开模式。

本节介绍的大部分内容都是一线的实战技巧，这些技巧需要读者亲自模拟操作才能掌握，请读者学习完本节内容后多加练习。

1.6 量化、中医与西医

简单对比传统理财策略与量化投资，就像中医与西医。传统理财就像中医，靠的是经验、感觉，即使使用所谓的技术，看看K线图，也只能发现一些趋势的征兆，而无法给出具体的参数指标。量化分析就像西医，靠的是专业的工具、设备、软件，例如体温计、血压表、显微镜、CT、X光，把相关现象数字化，再根据病理进行治疗。

以简单的感冒发烧为例，中医通过传统手段能够知道病人身体发热，需要降温，可具体发烧到多少度，却无法提供准确的数据，只能凭经验，开一些降温药材。西医凭借简单的体温计就可以精确地知道病人发烧的程度，从而针对性治疗。

两相比较，优劣自现。

在本章“一月效应”案例中，传统理财者可以凭借经验知道在一月份投资盈利的可能性大，也知道相关的滞后现象，可如何操作却没有具体的数据支持。

我们通过Python程序，对历年的股票数据进行简单的统计分析，不仅知道美国的“一月效应”更多的是在十二月份圣诞期间、中国大部分在二月份春节期间，同时我们也计算出“一月效应”的盈利概率具体是多少。

通过不同的指数组合，还可以发现大企业等蓝筹股的“一月效应”更加明显、中小企业创业板股票相对较弱。至于A股市场，混合指数的“黄金周效应”只是意外收获。

不过，科学史和投资史上的偶然现象往往是机遇的开始：X光、微波炉、镭射……

第2章
常用量化技术指标与框架

第1章是入门引导，主要介绍Python操作技巧与几个趣味投资理财的案例，本章其实才是量化的第一课。本章将从宏观角度、从整体上对Python量化领域进行整体介绍。

- 通过SMA均线策略的实盘案例，介绍完整量化投资流程。
- 目前主流的Python量化框架、模块库、软件包。
- Ta-Lib金融模块库与量化分析常用的金融技术指标。
- 经典量化策略介绍：Alpha策略、Beta策略和海龟交易法则等。
- 量化策略的分类与要点。

2.1　案例2-1：SMA均线策略

下面介绍一个相对完整的量化分析案例。这是一个SMA均线策略案例，源自PyAlgoTrade开源量化软件文档，为方便教学并便于理解，笔者对相关代码进行了修改和汉化，并增加了绘图输出。SMA均线策略也是动量策略的一种，所以SMA均线策略也可称为SMA动量策略。

注意，下面的Python案例程序是与PyAlgoTrade相关的案例程序，只适用于Python 2.7。具体代码请参见脚本文件\zwpython\zw_k10\k201_sta_anz.py。

运行代码后，程序会输出数据，包括图像和文字信息，如图2-1所示为运行结果，即SMA均线策略收益曲线图。



图2-1　SMA均线策略收益曲线图

输出的文字信息如下：

```
最终资产价值 Final portfolio value: $13091.62
累计回报率 Cumulative returns: 30.92 %
夏普比率 Sharpe ratio: 0.72
最大回撤率 Max. drawdown: 23.69 %
最长回撤时间 Longest drawdown duration: 277 days, 0:00:00

总交易 Total trades: 13
平均利润 Avg. profit: $154
利润方差 Profits std. dev.: $1268
最大利润 Max. profit: $4197
最小利润 Min. profit: $-892
平均收益率 Avg. return:  2 %
收益率方差 Returns std. dev.: 13 %
最大收益率 Max. return: 46 %
最小收益率 Min. return: -7 %

盈利交易 Profitable trades: 3
平均利润 Avg. profit: $1964
利润方差 Profits std. dev.: $1586
最大利润 Max. profit: $4197
最小利润 Min. profit: $662
平均收益率 Avg. return: 21 %
收益率方差 Returns std. dev.: 18 %
最大收益率 Max. return: 46 %
最小收益率 Min. return:  6 %
```

```
亏损交易Unprofitable trades: 10
平均亏损 Avg. loss: $-389
亏损方差 Losses std. dev.: $238
最大亏损 Max. loss: $-892
最小亏损 Min. loss: $-44
平均收益率 Avg. return: -3 %
收益率方差 Returns std. dev.:  2 %
最大收益率 Max. return: -0 %
最小收益率 Min. return: -7 %
```

2.1.1 案例要点与事件编程

在案例2-1的SMA均线策略程序中，策略调用的是以下语句，即20日SMA平均线策略：

```
myStrategy = SMACrossOver(feed, "orcl", 20)
```

这是最简单的SMA均线策略，案例程序默认的参数是“20日平均线”，这个“20日”是优化后的结果，读者可以用5、15、29、21、25、30等不同的时间周期参数分别进行测试，这些参数的运行结果大部分都是负收益和低收益。

PyAlgoTrade（简称PAT）开源量化软件采用的是目前常用的事件

模式，而前面的案例程序采用的是最简单的SMA均线策略，核心代码是onBars事件函数，其他都是辅助功能，如数据源配置、绘图、计算回报率模块库等。

注意，函数开头字母为on，表示是事件函数，代码如下：

```
def onBars(self, bars):
    # If a position was not opened, check if we should enter a
    if self.__position is None:
        if cross.cross_above(self.__prices, self.__sma) > 0:
            shares = int(self.getBroker().getCash() * 0.9 / bar
            # Enter a buy market order. The order is good till
            self.__position = self.enterLong(self.__instrument,
    # Check if we have to exit the position.
    elif not self.__position.exitActive() and cross.cross_below
        self.__position.exitMarket()
```

SMA均线策略很简单：

- 当股票价格高于SMA平均线价格并且是向上趋势时，买入；
- 当股票价格低于SMA平均线价格并且是向下趋势时，卖出。

SMA均线策略和函数代码虽然很简单，但细节很啰嗦，而且是事件函数，初学者请参看流程图2-2。

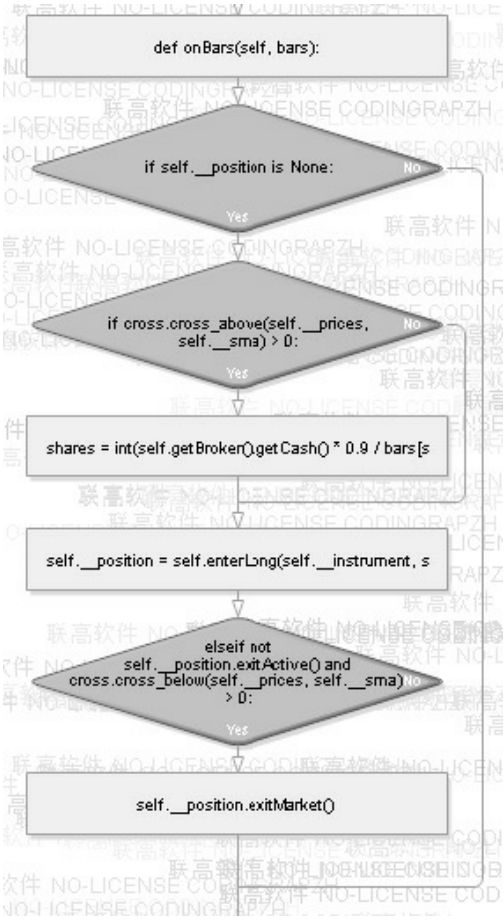


图2-2　SMA策略函数流程图

注意，事件模式对于数据采集、交易、分析一体化的量化系统来说，设计逻辑相对简单，是一体化的量化系统的主流模式。

github.org开源项目网站近来新增的量化项目大部分是独立策略分析模块，如backtrade、pyBacktest和zwQuant等，这些新一代量化软件专注策略和回溯的量化系统，更多采用传统的循环模式，以简化系统设计。

Python原本是脚本语言，不擅长做大工程、大项目，笔者曾经在博客和《zwPython开发平台用户手册》里提过：

“学习Python语言应该采用“面向过程”的方式，至少可以节省90%的精力。虽然Python语言是全对象设计，但Python的class类是弱类，不象

C++、Java那么复杂。对于Python语言的class类，笔者建议可以简单地把其视为一组预定义的函数集合或数据集合，类似C语言里的struc结构或Delphi语言里的带函数的记录。

对于Python程序，我们应尽量远离自省、attr开头或结尾的命令，同样，on开头的事件函数也尽量不要直接使用。因为on开头的事件函数，虽然有关的事件处理在函数内定义，但上级调用模块却可能放在其他的模块库，而Python语言对于大型软件的跨模块调试支持非常弱。”

本小节案例程序代码中的示例看起来很简单，是因为笔者把onBars这个函数单独分离出来了，否则一大堆on开头的函数没有先后顺序、没有入口函数，初学者很容易搞混，不知道如何下手，例如：

```
def onEnterCanceled(self, position):
def onExitOk(self, position):
def onExitCanceled(self, position):
def onBars(self, bars):
```

这个案例程序的代码是最简单的，只有一个class类定义；对于多个模块、多个class类定义，还要考虑各个class类之间的关系，代码的复杂度呈几何程度上升。

这里，我们再说几句题外话，笔者之所以选择PAT的案例程序作为教学参考案例，而不是直接采用quantopian等目前国内外流行的量化网站架构进行讲解，是因为以下几个原因：

- 目前，在线量化交易网站逐渐增多，每个网站应用环境都不同，

不可能一一讲到。

- 量化交易的核心是量化分析，PAT是离线版本，而且有源码，国内的量化网站大部分都是闭源的。
- 量化分析的核心就是onBars事件函数，各个量化网站差不多，只是函数名称、参数不同而已，其他的都是辅助功能。
- 掌握了onBars这个量化的“准”主函数，其他各种应用平台、操作都是大同小异的，只是细节不同，读者参考相关网站的说明资料基本就能够掌握了。
- PAT量化软件是开源的，能够对相关代码进行修改、汉化，强化教学功能，案例2-1的代码就是笔者根据初学者的水平进行了二次修改、汉化，并且增加了绘图输出，以方便读者学习。

2.1.2 量化程序结构

下面介绍代码的其他部分，以案例2-1为例，其程序相对比较简单，可以分为以下几个部分：

- Import为模块库导入，注意相关的缩写；
- class SMACrossOver(strategy.BacktestingStrategy)为类定义，注意是源自BacktestingStrategy；
- main主流程代码，数据源设置、调用等；
- 回报率计算，输出相关数据。

Import为模块库导入部分，是Python语言的常规流程，读者需注意相关模块库的缩写。

class SMACrossOver为类定义部分，比较复杂，现阶段读者只需把握好onBars函数就可以了。

回报率计算和数据输出已经汉化，很简单。虽然输出数据量很多，但都是行业惯例，而且已经分成了以下几组：

- 回报率概况；
- 总的交易情况；
- 盈利部分交易情况；
- 亏损部分交易情况。

2.1.3 main程序主入口

案例2-1程序代码的重点还是main程序主入口的流程代码，相关程序代码如下，已经增加了中文注释：

```
# Load the yahoo feed from the CSV file
# 初始化数据源变量feed
feed = yahoofeed.Feed()

# 加载数据源文件给feed
# 读者可以从zwDat的美股数据目录下复制其他数据文件进行测试
# 因为PAT使用的是美股数据，因此直接加载zw版和cn中国A股数据会出错，需要修改c
feed.addBarsFromCSV("orcl", "dat\\orcl-2000.csv")
```

```
# Evaluate the strategy with the feed's bars.
# 定义所用的策略myStrategy，源自class类定义SMACrossOver
# 定义时，已经导入了数据源feed，股票名称为orcl，SMA时长为20天
myStrategy = SMACrossOver(feed, "orcl", 20)

# Attach different analyzers to a strategy before executing it.
# 策略运行前，绑定相关分析模块

# 回报率分析
retAnalyzer = returns.Returns()
myStrategy.attachAnalyzer(retAnalyzer)

# 夏普指数分析
sharpeRatioAnalyzer = sharpe.SharpeRatio()
myStrategy.attachAnalyzer(sharpeRatioAnalyzer)

# 最大回撤率分析
drawDownAnalyzer = drawdown.DrawDown()
myStrategy.attachAnalyzer(drawDownAnalyzer)

# 交易分析
tradesAnalyzer = trades.Trades()
myStrategy.attachAnalyzer(tradesAnalyzer)

# ---set.plot
#设置绘图参数
plt = plotter.StrategyPlotter(myStrategy, True, False, True)
```

```
#设置绘图数据源，来自getSMA函数
plt.getInstrumentSubplot('orcl').addDataSeries("sma", myStrategy


# Run the strategy.
# 运行策略
myStrategy.run()


# 绘制策略回溯图
plt.plot()
```

这个案例程序对于初学者而言，即使汉化后难度也很大，毕竟这是一个完整的量化分析流程。

目前，读者不需要完全理解，我们在后面有关的章节中也会讲到这个案例。现阶段，读者通过这个案例对量化整体流程有一个全面的印象就可以了。

2.1.4 KISS法则

“KISS”是英语“Keep It Simple, Stupid”的首字母缩写。KISS法则是指在设计中应当注重简约的原则。KISS法则广泛应用在商业书信、电脑软件、动画和工程上。

“KISS”原文也有很多其他版本，包括“Keep It Sweet & Simple”、“Keep It Short & Simple”、“Keep it Simple, Sweetheart”和“Keep

it Simple, Sherlock”。

对于量化初学者而言，要掌握、理解案例2-1背后的整个量化流程，以及学会灵活利用其他策略取代案例程序中的SMA平均线策略。

对于量化专业人员而言，要掌握相关的代码构成，以及背后的基本逻辑，独立完成相关的Backtest（回溯）测试模块编程。

策略是量化交易的核心，无论是量化初学者，还是专业人员，都需要重点学习Backtest（回溯）测试模块部分。

至于量化程序前端的数据API接口与数据采集模块、后端的Trade（下单交易）模块，可以通过有关金融数据网站的API接口或者CSV等数据文件格式，甚至通过第三方软件实现；量化交易中的部分环节，特别是非高频的下单环节，也可以采用人工模式。这个也是KISS法则在量化方面的体现。

无独有偶，海龟交易法则的四大核心之一也是简单明了。

海龟交易法则的四大核心如下：

- 掌握优势，找到一个期望值为正的交易策略，因为从长期看，它能创造正的回报；
- 管理风险、控制风险、守住阵地，否则你可能等不到创造成果的那一天；
- 坚定不移，唯有坚定不移地执行你的策略，才能真正获得系统的成效；
- 简单明了，从长久看，简单的系统比复杂的系统更有生命力。

2.2 Python量化系统框架

量化系统的划分有很多标准，最简单的方式是根据软件架构分为以下两大类：

- 事件（消息）模式，如PyAlgoTrade等目前主流的交易一体化平台基本都是这种模式；
- 循环模式（类似暴力破解：穷举模式、词典模式），包括zwQuant、PyBacktest和Zipline等。

事件模式起步较早，不少大企业因为交易一体化的需要，目前依然很流行，但事件模式过于复杂、主次不分，对量化分析的核心即策略方面不够专业。

目前，GitHub等开源项目网站新增的量化项目，大多都只有独立的回溯Bakctest模块，例如BT、trading-backtest、backtrader等。

以上这种划分模式只是笔者的一家之言，读者也可以参考其他学者的划分方法。

2.2.1 量化行业关键词

在进一步介绍Python量化系统之前，先介绍几个常见的行业关键词。

- Quant：宽客，Quant的音译，英文是定量的意思，源自化学的定

量分析，许多量化交易软件的名称都包含Quant。

- strategy：投资策略，前面介绍的SMA策略、“一月效应”、A股黄金周效应都可以看做简单的量化投资策略。
- drawdown：最大回撤率，是指金融产品在历史上一段时间的最大跌幅。最大回撤率用来描述买入产品后可能出现的最糟糕的情况。最大回撤率是一个重要的风险指标，对于对冲基金和量化策略交易来说，该指标比波动率还重要。
- Backtest：回溯测试，简单来说，就是用历史数据测试相关的投资策略。
- Slippage：滑点，是指下单的价格点位，和最后成交的价格点位有差距。
- CAPM：资本资产定价模型，是Capital Asset Pricing Model的缩写。它是在资产组合理论的基础上发展起来的，是现代金融市场价格理论的支柱。CAPM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，CAPM模型研究的重点在于：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，投资者应该获得更多的报酬率。
- ETF：交易型开放式指数基金，通常又被称为交易所交易基金（Exchange Traded Funds，简称ETF），是一种在交易所上市交易的、基金份额可变的开放式基金。
- FOF：基金中的基金，与开放式基金最大的区别在于，基金中的基金以基金为投资标的，而基金是以股票、债券等有价证券为投资标的。它通过专业机构对基金进行筛选，帮助投资者优化基金投资效果。
- OHLC：外汇交易术语，O（Open）代表开盘价，H（High）代表最高价，L（Low）代表最低价，C（Close）代表收盘价。

2.2.2 国外主流Python量化网站

目前，全球量化投资很流行，发展很快，以下三家是全球知名的大型量化网站。

- quantopian，网址是https://www.quantopian.com/，网站首页如图2-3所示。目前为行业第一，Zipline就是该公司的产品，国内优矿等Web模式的量化平台基本上都是模仿Quantopian网站。

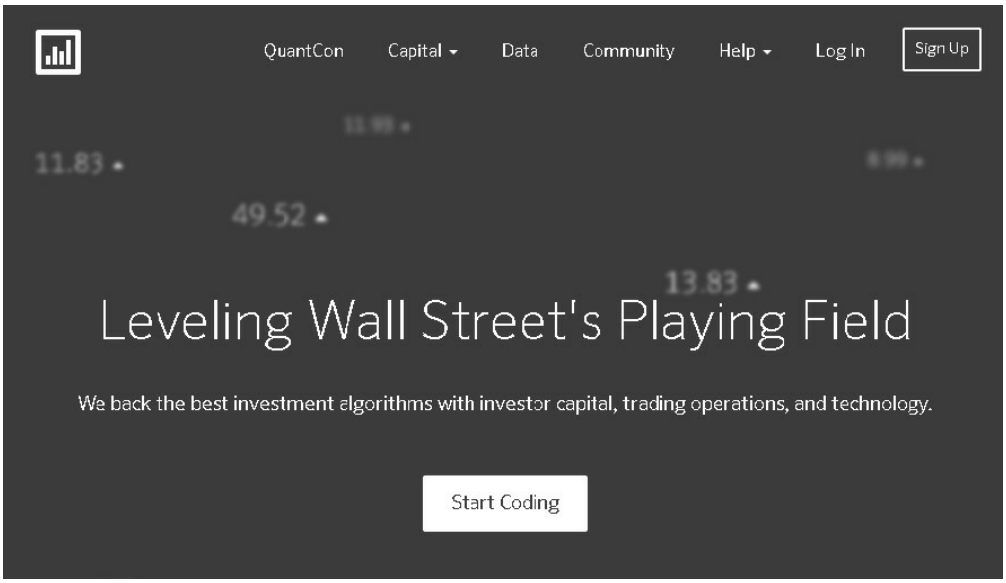


图2-3 quantopian量化网站首页

- Wilmott，网址是http://www.wilmott.com/，网站首页如图2-4所示，以策略研究为主。

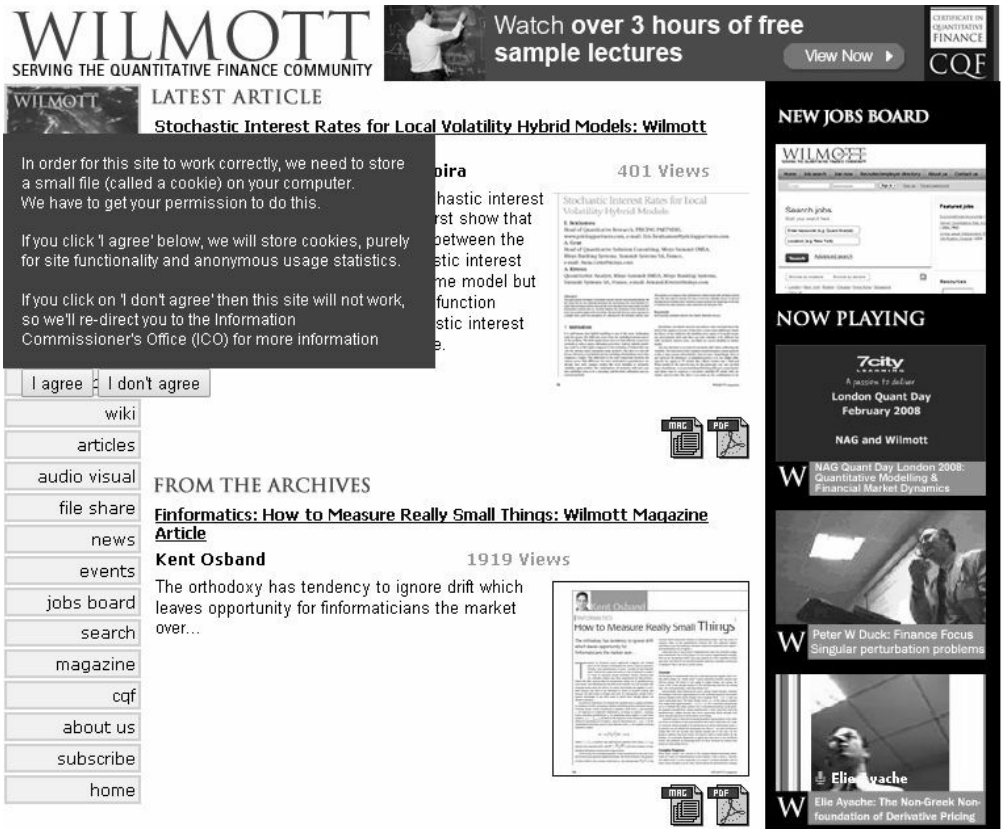


图2-4　Wilmott量化网站首页

- QuantPedia，网址是http://www.quantpedia.com/，网站首页如图2-5所示，以策略研究为主，免费用户只能看前几页资料。

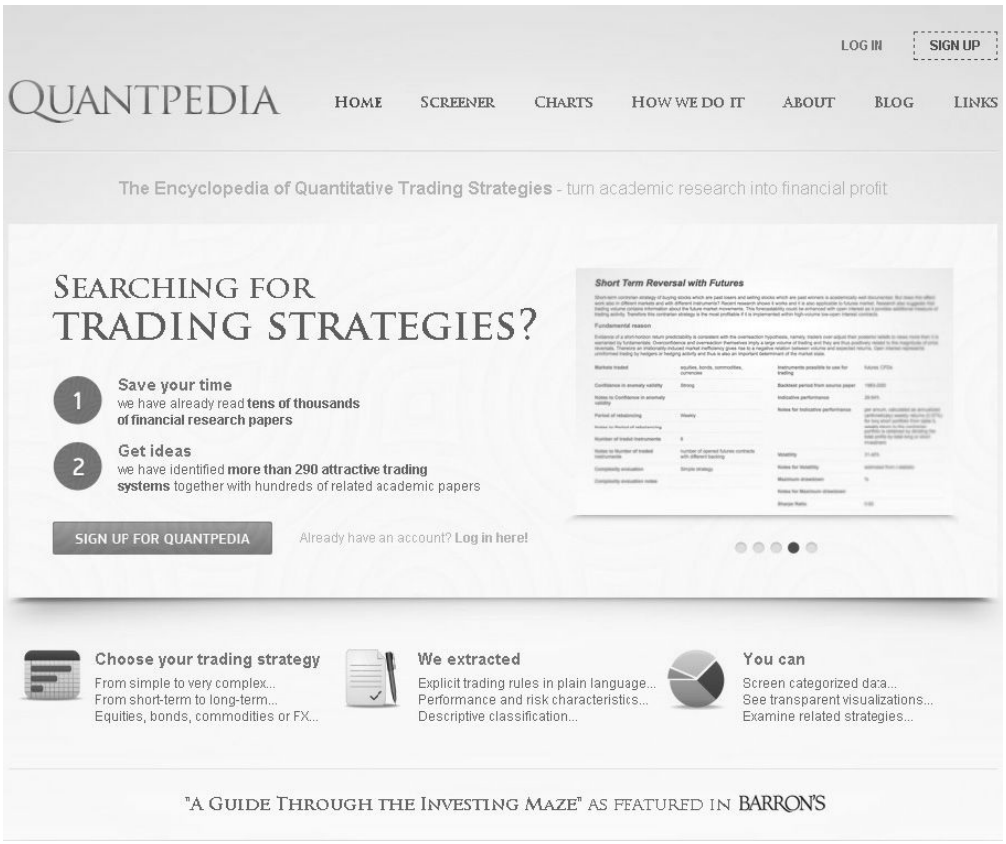


图2-5　QuantPedia量化网站首页

2.2.3　我国主流Python量化网站

下面是我国出现较早、业内知名的量化网站。

- zw量化，网址是www.ziwang.com，网站首页如图2-6所示。2016年免费发布了业内首套开源Python量化分析平台zwQuant，以及zwDat股票数据包和大量免费开源的量化课件，也是本书的技术支持网站。



图2-6　zw量化网站首页

- 极宽量化，网址是www.zQuant.cn，是zwQuant量化软件的升级版本，zQuant极宽量化软件的官方网站也是本书的技术支持网站，未来www.ziwang.com网站的量化模块会逐渐向极宽量化网站迁移。
- 掘金，网址是www.myquant.cn，网站首页如图2-7所示，老牌量化网站，有客户端软件，社区也不错。



图2-7　掘金网站首页

- 优矿，网址是www.uqer.io，网站首页如图2-8所示，老牌量化网站，是我国最大金融数据公司通联数据旗下的企业。



图2-8　优矿网站首页

- 聚宽，网址是www.joinquant.com，网站首页如图2-9所示，致力于打造最高效、易用的量化交易平台。基本上就是Quantopian网站的汉化版。



图2-9　聚宽网站首页

- RiceQuant，网址是www.ricequant.com，网站首页如图2-10所示，与聚宽类似，也是模拟Quantopian网站，不过增加了不少本地化的东西。

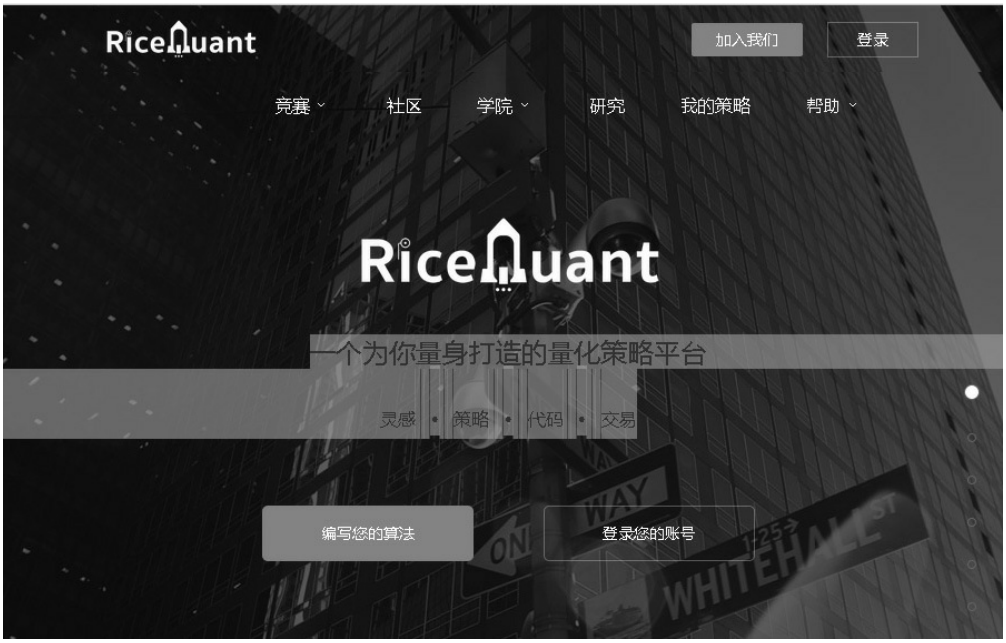


图2-10　RiceQuant网站首页

2.2.4 主流Python量化框架

目前，国内外比较成熟的Python量化框架可以分为三大类。

- Web在线网站模式，类似云计算网站，将数据采集、分析、交易一体化，如quantopian、优矿、聚宽、RichQuant等。
- 客户端模式，如PyAlgoTrade、掘金、vnpy、QuantDigger。
- 策略回溯模式，只做策略回溯部分，如zwQuant、pyQuanttest和Zipline。

1. Web在线网站模式量化系统

对于Web在线网站模式，目前全球quantopian量化网站一家独大，基本上可称之为quantopian模式。在全球最大的开源项目网站GitHub中使用关键词“quantopian”进行搜索，搜索结果中与quantopian相关的Python量化项目就有38个之多，如图2-11所示。

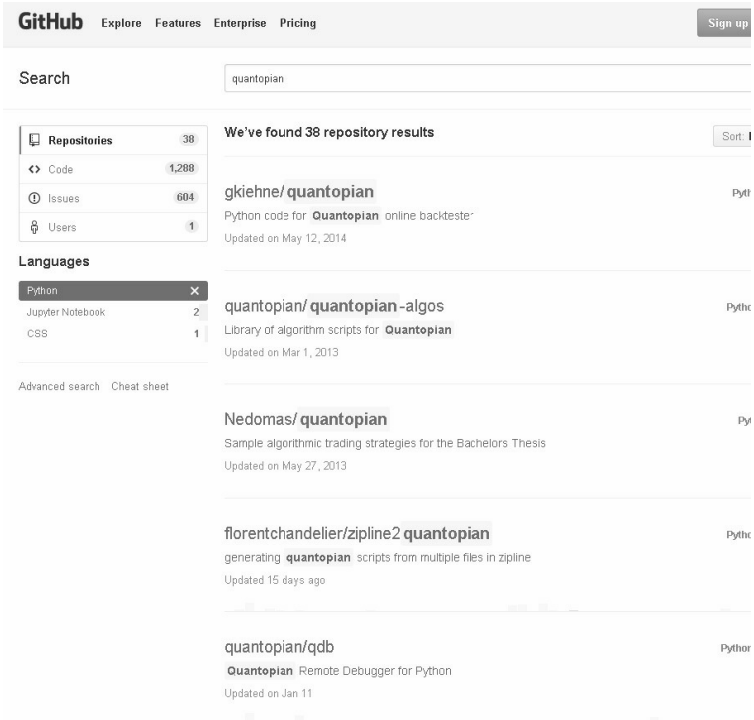


图2-11　GitHub开源网站与quantopian关联的项目

搜索结果的网址是https://github.com/quantopian。

单击第二个结果“quantopian-algos”左侧的“quantopian”链接，这个是项目开发商的链接（不一定是第二个结果，注意项目开发商的关键词“quantopian”），整理一下，有近56个项目。

笔者把目前的量化项目分为quantopian类别和非quantopian类别，虽然方法简单，但还是有一定道理的。

quantopian网站在GitHub上发布的开源项目大部分是Python版本，与数据分析、量化分析相关的项目虽然不少，但属于衍生项目，有名的项目也不少，量化领域最出名的就是Zipline和quantopian-algos。

2. quantopian项目清单

下面是quantopian在GitHub上发布的开源项目，如图2-12所示。

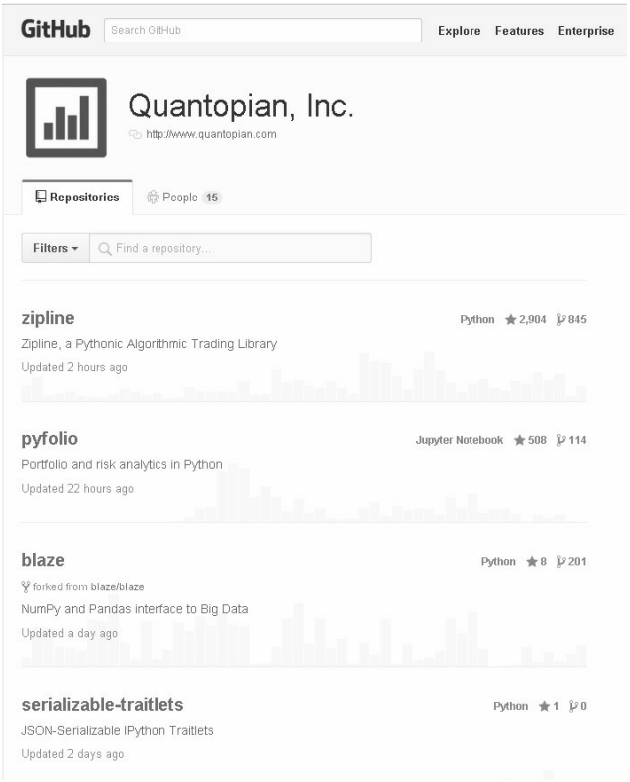


图2-12　quantopian开源项目

（1）DockORM，docker对象—关系映射。

（2）MongoDBProxy，MongoDB数据库独立工具。

（3）QuantSoftwareToolKit，QSTK量化工具包。

（4）Zipline，优秀的Python交易算法库。

（5）bcolz，数据列压缩工具。

（6）blaze，NumPy、Pandas与大数据的接口。

（7）cancan，ROR工具软件。

（8）cassandra_snapshotter，数据备份工具。

（9）chest，简单的字典数据格式，支持外溢至磁盘。

（10）coal-mine，监控任务的执行周期。

（11）conda-build，conda包管理工具箱。

（12）cytoolz，优化版本，高性能函数集。

（13）datashape，定义数据描述语言。

（14）dd-agent，数据代理。

（15）dockerspawner，docker工具。

（16）fakeredis，redis-rb开发和测试环境。

（17）gevent，gevent协同并发库工具。

（18）ipykernel，用于Jupyter的IPython 内核。

（19）ipython，专业的Python shell控制台。

（20）jsonrpc.py，JSON-RPC序列化、反序列化工具。

（21）jupyterhub，Jupyter notebooks多用户服务。

（22）letter_opener，浏览器邮件预览工具。

（23）logbook，log工具。

（24）metautils，metaclasses工具包。

（25）mobylette，ROR工具包。

（26）mongo-python-driver，mongo驱动。

（27）mongoid-encrypted-fields，mongo日期处理。

（28）mongoid-slug，mongo工具。

（29）mongoid_rails_migrations，mongo数据迁移工具。

（30）nose-parameterized，nose参数测试。

（31）nose_gevent_multiprocess，多进程gevent插件测试框架。

（32）nose_xunit_gevent，Xunit多任务测试工具。

（33）odo，数据处理，源自Blaze项目。

（34）pagedown-bootstrap，Markdown编辑工具。

（35）periodicbackup-plugin，备份工具。

（36）pgcontents，IPython内容管理工具。

（37）pyfolio，投资组合与风险分析。

（38）qdb，QDB型远程调试器。

（39）qgrid，IPython Notebook交互式的数据网格，支持数据排序、过滤。

（40）quantopian-algos，quantopian量化交易算法脚本集。

（41）quantopian-drafts，一些新的quantopian功能函数。

（42）rack-mini-profiler，Ruby应用微型分析器。

（43）research_public，quantopian公共研究项目。

（44）schematics，人性化的数据结构。

（45）sendgrid-python，SendGrid的Python模块库。

（46）serializable-traitlets，可序列化的JSON。

（47）sparklines，Ruby工具包。

（48）sqlalchemy_fdw，基于psycopg2的SQL数据库工具。

（49）stripe-ruby-mock，stripe-ruby-mock模拟库。

（50）TA-Lib，专业的金融财经函数集。

（51）tmpnb，Jupyter Notebook服务器，使用docker模式。

（52）toolz，Python标准函数库。

（53）tradingcalendar，交易所工具包。

（54）traitlets，轻量级Traits类模块。

（55）vanity，Ruby开发工具。

（56）xlrd，Excel工具。

3. 客户端模式量化系统

backtrader量化软件类似掘金量化软件，是客户端模式的，其网页简介中也介绍了不少目前主流的Python量化软件包。我们先到Python官网的模块库中搜索一下“backtrader”。

搜索结果的网址是https://pypi.python.org/pypi/backtrader/。

网页下方的Alternatives（替代品）栏中介绍了多个流行的Python量化软件包，如图2-13所示。

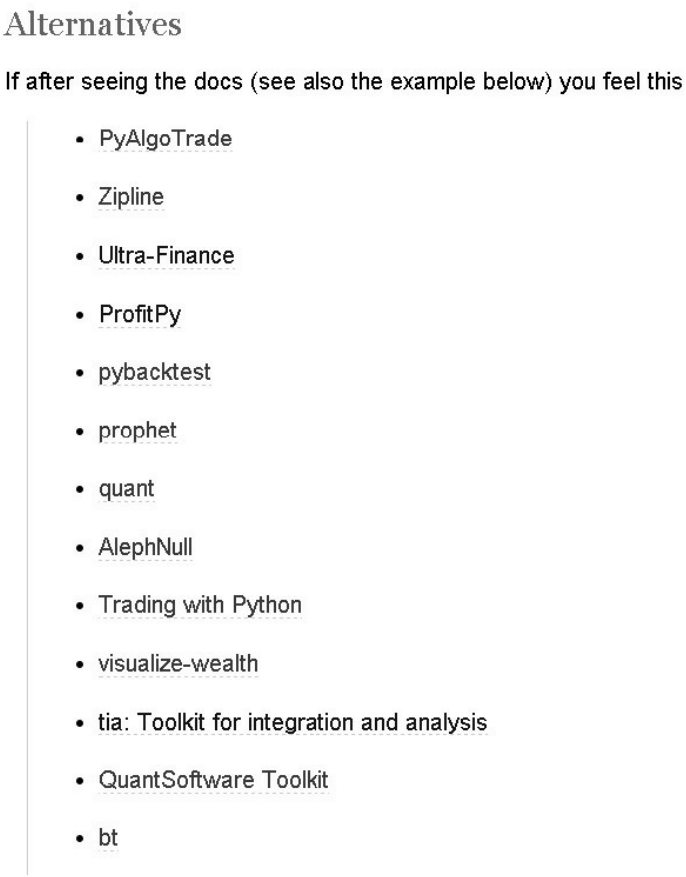


图2-13　Backtrade量化软件及关联项目

以上开源项目并不全是客户端量化系统，大部分如backtrader量化软件自身，着重于策略回溯模块。在目前Python量化的起步阶段，这种

举一反三的项目外延方式，读者可以多使用，尽量扩大自己的知识面。除了这些，国内的vnpy、QuantDigger和掘金都是客户端类的交易系统。

如图2-14和图2-15所示分别为vnpy量化项目和QuantDigger量化项目。

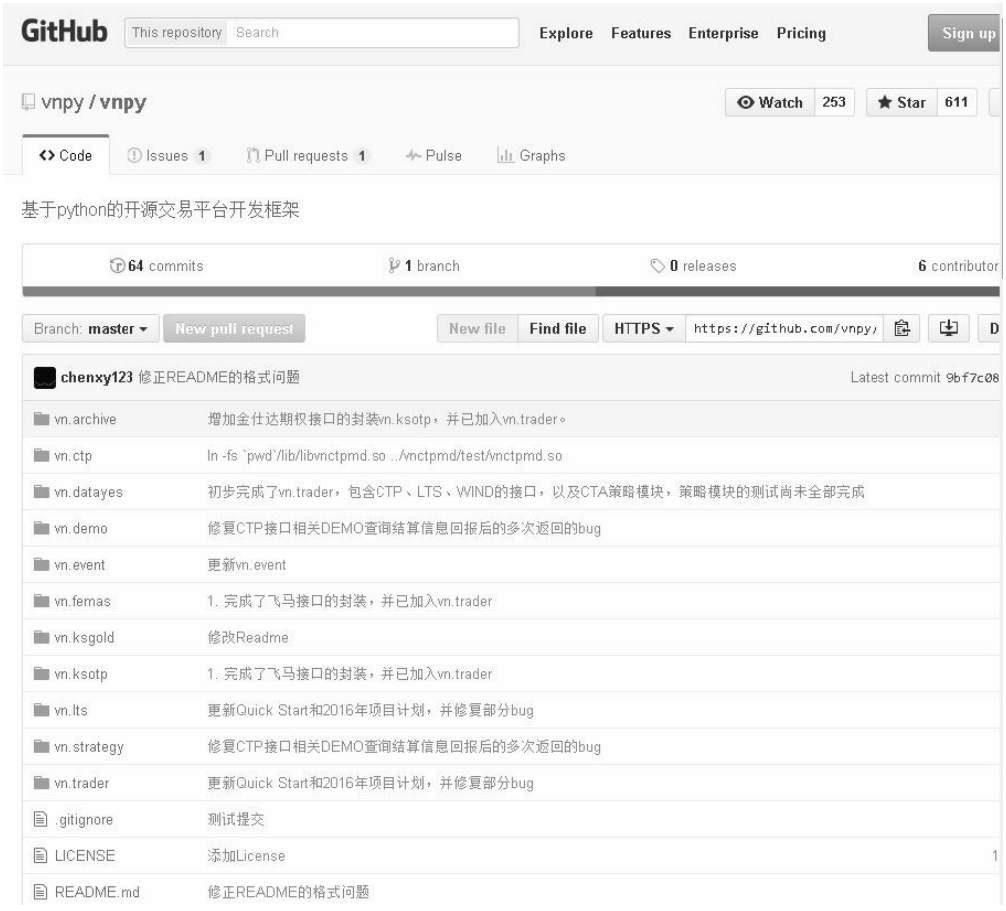


图2-14　vnpy量化项目

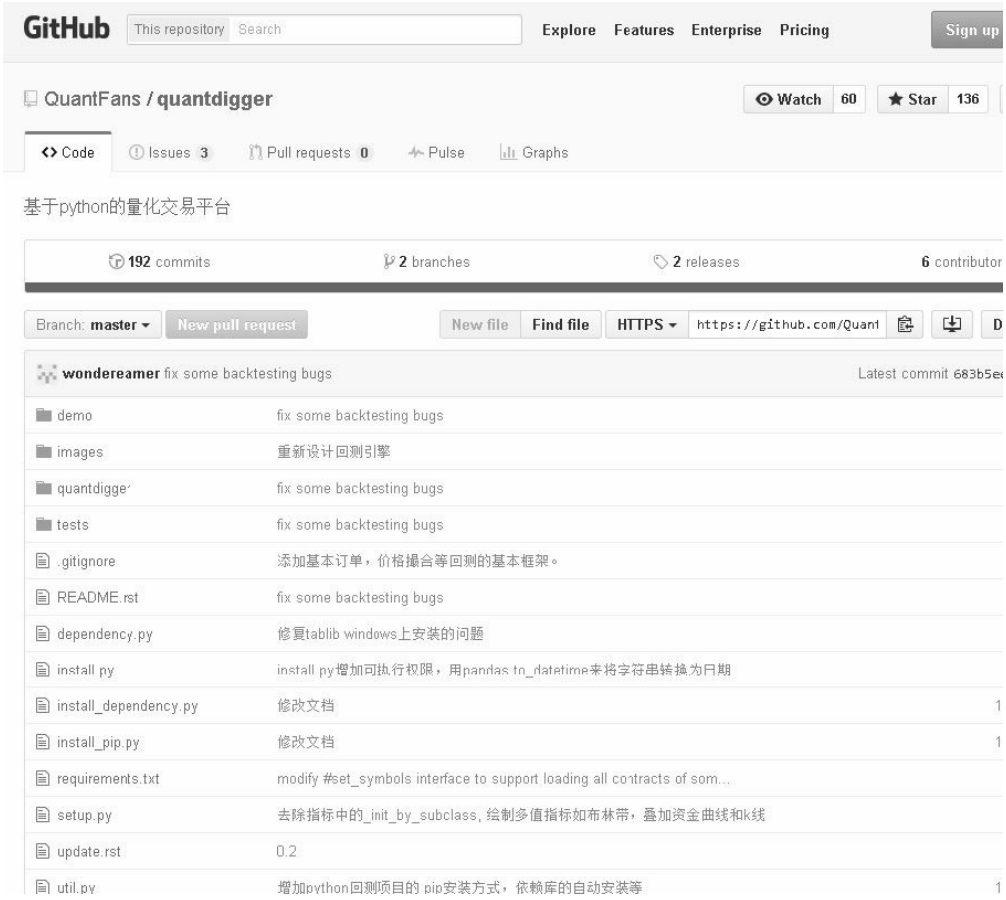


图2-15　QuantDigger量化项目

在GitHub开源项目网站上，中文量化项目发展迅速，大约占20%，如图2-16所示是关键词“量化”的搜索结果，读者可以用其他的金融类关键词再搜索看看。

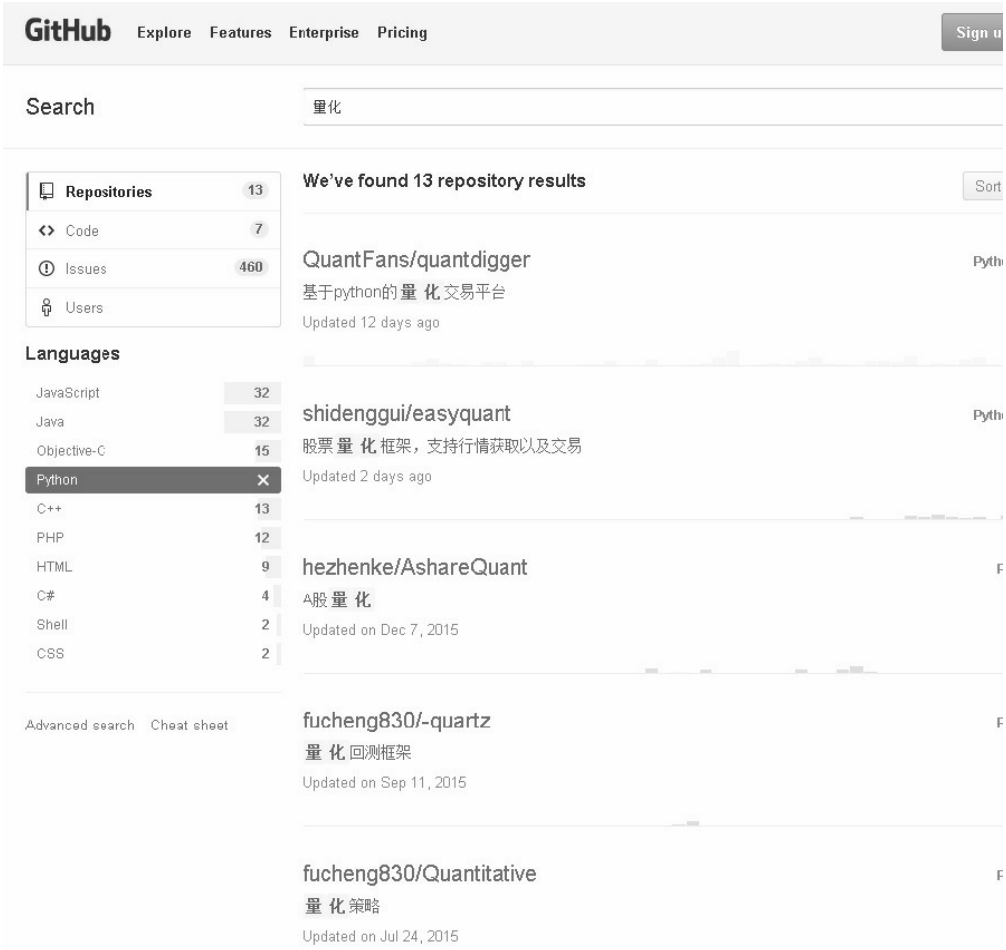


图2-16　GitHub开源网站中文量化项目

4. 策略回溯模式

策略回溯模式偏重于策略分析、回溯测试，虽然起步晚，但目前已经成为趋势。zwQuant、PyBacktest，backtrade、BT、Zipline等开源量化项目基本上都可以归为此类。

如图2-17和图2-18所示分别为PyBacktest开源量化项目和Zipline开源量化项目。

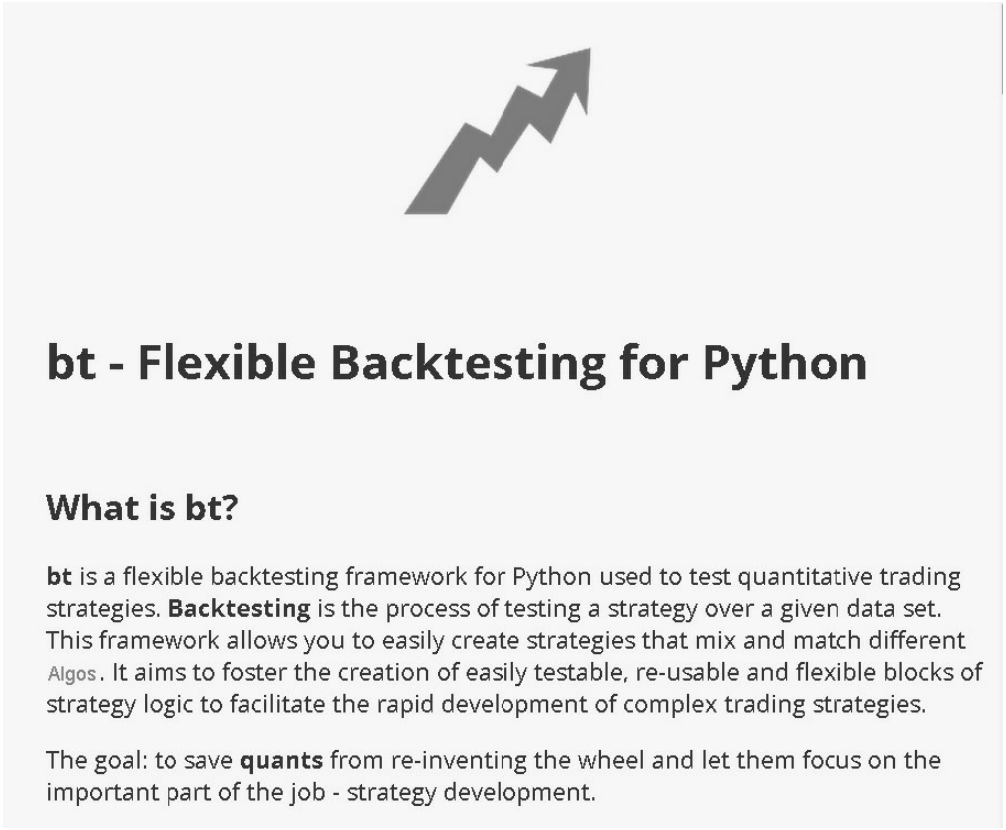


图2-17　PyBacktest开源量化项目



图2-18　Zipline开源量化项目

2.3 常用量化软件包

本书重点讲解策略分析和金融行业的数据分析，整体而言，与量化相关的软件包有如下几个。

- zwQuant，国内首个开源量化程序，纯Python源码，适用于Windows、Linux、Mac等多种操作系统，其网站（www.ziwang.com）还提供了大量的中文文档，是初学者首选的Python学习教材，也是国内金融机构、私募团队进行量化产品开发很好的参考模板。
- Pandas，是基于NumPy的数据分析工具，该工具是为了解决金融数据分析任务而创建的。Pandas纳入了大量库和一些标准的数据模型，提供了大量能使我们快速、便捷地处理数据的函数和方法。它使Python成为强大而高效的数据分析环境。
- Tushare，免费、开源的Python财经数据接口包。主要对股票等金融数据进行数据采集、清理加工和数据存储，能够为金融分析人员提供快速、整洁和多样的便于分析的数据，为他们在数据获取方面极大地减轻工作量，使他们更加专注于策略和模型的研究与实现。
- Ta-Lib，金融函数包，目前已经被金融行业广泛使用，用来对金融市场的数据进行技术分析。
- Zipline，用于量化分析，它是一个交易算法库，可以对历史数据进行投资算法的回溯检验。Zipline目前是quantopian的回溯检验引擎。它使用简单，包括常用的统计方法，如移动平均和线性回归；与现有Python生态圈能很好的融合；支持交易系统的开发、

数据分析和可视化。

- PyAlgoTrade，简称PAT，是开源的Python量化交易框架，支持策略回测和实盘交易，提供全面的技术分析接口。
- vnpy，中文版本的Python算法系统，是一个基于Python的开源量化交易框架项目。
- QuantDigger，中文版本的Python算法系统，策略语法更简单，类似MC、TB这些商业软件，但并不牺牲灵活性，保留了Python这门通用语言的所有功能。QuantDigger目前定位于研究工具，但是在设计上还是会从实盘交易的角度考虑，将来也会接入交易接口。
- QSTK，2012年的金融量化项目，框架有些旧。

以上量化行业相关的软件包中，最重要的还是Pandas数据分析模块、TuShare数据采集模块，以及TA-Lib、Zipline金融函数库。本书大量引用了PyAlgoTrade中的文档和代码，因此PyAlgoTrade读者也需重点关注。

PyAlgoTrade和TA-Lib的中文资料很少，笔者在本书网盘资源中提供了部分简单汉化版本，读者可以免费下载并参考。至于zwQuant、vnpy、TuShare和QuantDigger等量化软件，本身就是中文版的，文档也做得很好，读者可以自己下载，查看相关的源码和文档。

2.3.1 常用量化软件包简介

如图2-19~图2-22所示为常用的量化软件包。

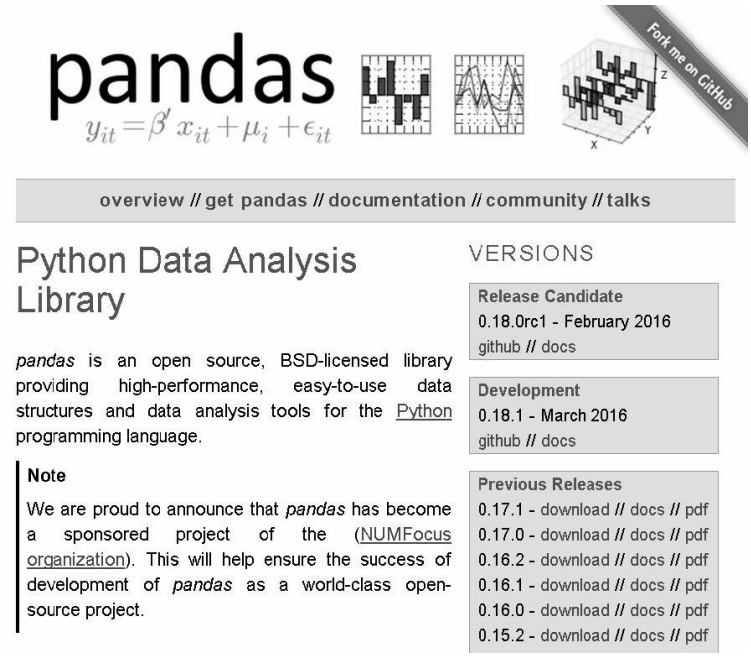


图2-19　Pandas开源数据分析软件



图2-20　TA-Lib金融软件包



图2-21 PyAlgoTrade开源量化软件



图2-22 TuShare开源金融数据抓取模块库

2.3.2 案例2-2：模块库列表

我们在使用Python时，经常需要安装大量的模块库，有时也需要整理Python平台现有的模块库。以下脚本很短，却可以很好地完成这些工作。

脚本文件名是\zwpython\zw_k10\k202_piplst.py。

代码如下：

```
# -*- coding: utf-8 -*-

import sys
import os
import pandas as pd
import pip

#=============
plst=pip.get_installed_distributions();
print(plst[10])

df=pd.DataFrame();
df['<name>']=plst;
print(df.tail())
fss="tmp\\m100.csv";print("\n"+fss)
```

```
df.to_csv(fss,index=False)
```

源码很简单，才十几行，支持Python 2.7和Python 3.x两种环境。在Python 2.7环境中运行，输出结果如图2-23所示。

```
In [11]: runfile('F:/zwPython/zw_k10/k202_piplst.py', wdir='F:/zwPython/zw_k10')
aggdraw 1.2a3
                  <name>
302     XlsxWriter 0.8.4
303           xray 0.7.0
304  yahoo-finance 1.2.1
305    yfinanceapi 1.0.2
306        zipline 0.7.0

tmp\m100.csv
```

图2-23　输出的内置模块库清单列表

在完整的输出结果中显示系统共有306个模块，输出CSV数据文件为tmp\m100.csv。注意，不同版本具体安装的模块库数目会有差异。

这段代码很简单，需要注意的是，我们采用Pandas数据分析模块进行数据的处理和保存等工作，而不是通常Python语言惯用的list（列表）模式。

Pandas数据分析模块库是Python量化的基础平台，已经整合了许多高端模块库。读者在量化编程时，尽量使用Pandas中的函数，一方面可以统一开发环境，另一方面也可简化配置。本例就直接调用Pandas数据分析模块的CSV函数，无需加载CSV模块库。

此外，借助Pandas数据分析模块，一条指令就可以查看最终的模块库清单，代码如下：

```
print(df.tail())
```

2.4 常用量化技术指标

股市中有很多经典的常识：

价格是最好的情绪，均线是最好的量化指标，成交量是最好的热点，强势不回调是最有利的利好，跌破支撑筹码密集区是最直接的利空。

这些常识，其实也是量化技术指标的一部分。

传统选股指标很多，例如成交量、均线、利润总额、每股收益、净资产收益率、威廉指标、布林带、DIF、RSI、DEA、MACD、KDJ、BBI等，这些指标已经形成了一个相对完整的技术指标体系，并从中衍生出许多技术流派。

因为本书篇幅有限，以上常用的技术指标和本书中其他技术指标、专业术语，在百度百科和各大金融网站都有介绍，不明之处，读者可自行查找。

在传统选股指标中，最真实的就是成交量，因为成交量都是“真金白银”，造不了假。

个股成交量的变化可以当做一个判断资金流向、市场情绪的标准。至于其他指标，如MACD、KDJ、BBI等，严格来说，都具有滞后性，是股票的量价走势决定了指标的走势，这个先后顺序是不能颠倒的。

量化技术指标可以分为以下两类：

- 常用金融股市技术指标；
- Python量化指标，是指目前已经有的、成熟的Python函数库的量化指标，如TA-Lib等。

之所以增加Python量化指标，是因为以下几个原因：

- 金融指数往往涉及到很多专业的理论知识，初学者缺乏足够的专业、理论背景，因此将有关指数代码化；
- 许多新生的、小众的指数缺乏足够的理论支持、实盘验证，初学者很容易被这些指数的短期效应迷惑；
- 常用的、基础的、成熟的技术指标大部分都有现成的函数库，足够支持相关的量化分析，而且可以大大简化程序结构，提高开发效率。

2.4.1 TA-Lib金融软件包

TA-Lib金融软件包是目前金融行业最常用的技术指标函数库，目前已经有Python版本接口。zwPython开发平台也已经集成了TA-Lib模块库。同时，本书配套的资源网盘还提供了《TA-Lib函数手册·汉化版》开源课件，请读者自行下载。

除个别技术指标和量化风险指数指标外，传统金融指标在TA-Lib函数库中基本都已经包括了。需要注意的是，因为TA-Lib是在国外开发的，因此个别术语和参数可能与国内的习惯用法不同，具体使用时请读者注意相关细节。

考虑到程序设计运行效率和与Pandas数据分析软件的集成，笔者与

极宽量化团队参考国外类似的开源软件，使用Pandas数据分析软件的矩阵模式，对TA-Lib重点函数进行了改写移植。目前，基本的移植版本已经发布，已经实现的函数有29个。

移植版本的文件名为zw_k10\pandas_talib.py。

有关函数名称如下。

ACCDIST(df,n)：积累/分配（Accumulation/Distribution）。

ADX(df,n,n_ADX)：定向运动平均指数（Average Directional Movement Index）。

ATR(df,n)：平均真实范围（Average True Range）。

BBANDS(df,n)：布林带（Bollinger Bands）。

CCI(df,n)：商品通道指数（Commodity Channel Index）。

COPP(df,n)：Coppock曲线（Coppock Curve）。

Chaikin(df)：蔡金振荡器（Chaikin Oscillator）。

DONCH(df,n)：奇安通道（Donchian Channel）。

EMA(df,n)：指数移动平均（Exponential Moving Average）。

EOM(df,n)：缓解运动（Ease of Movement）。

FORCE(df,n)：力指数（Force Index）。

KELCH(df,n)：Keltner通道（Keltner Channel）。

KST(df,r1,r2,r3,r4,n1,n2,n3,n4)：KST振荡器（KST Oscillator）。

MA(df,n)：移动平均（Moving Average）。

MACD(df,n_fast,n_slow)：MACD指标信号和MACD的区别（MACD Signal and MACD difference）。

MFI(df,n)：资金流量指标和比率（Money Flow Index and Ratio）。

MOM(df,n)：动量（Momentum）。

MassI(df)：质量指数（Mass Index）。

OBV(df,n)：平衡量（On-balance Volume）。

PPSR(df)：支点、支撑和阻力（Pivot Points,Supports and Resistances）。

ROC(df,n)：变化率（Rate of Change）。

RSI(df,n)：相对强弱指标（Relative Strength Index）。

STDDEV(df,n)：标准偏差（Standard Deviation）。

STO(df,n)：随机指标D（Stochastic oscillator %D）。

STOK(df)：随机指标K（Stochastic oscillator %K）。

TRIX(df,n)：矩阵（Trix）。

TSI(df,r,s)：真实强度指数（True Strength Index）。

ULTOSC(df)：最终振荡器（Ultimate Oscillator）。

Vortex(df,n)：涡指标（Vortex Indicator）。

2.4.2 案例2-3：MA均线函数调用

案例2-3调用Pandas版本的TA-Lib函数。

案例文件名为\zwpython\zw_k10\ k203ma_pdta.py。

全部代码如下：

```
# -*- coding: utf-8 -*-

import numpy as np
import scipy as sp

import matplotlib as mpl
import matplotlib.pyplot as plt
import pandas as pd
import pandas_talib as pdta

#----------code.init
mpl.style.use('seaborn-whitegrid');

#============main
```

```
fss="dat\\Appl2014.csv"
df=pd.read_csv(fss,encoding='gbk');
df=pdta.MA(df,5);
df=pdta.MA(df,10);
df=pdta.MA(df,30);
df=pdta.MA(df,50);
print(df.tail())

df['Close'].plot(figsize=(15,5));
df['MA_5'].plot();
df['MA_10'].plot();
df['MA_30'].plot();
df['MA_50'].plot();
plt.legend(ncol=5)
```

案例2-3程序很简单，只是调用了TA-Lib金融函数库里面最常用的MA均线函数，分别生成5日、10日、30日和50日的日均平均线数据。

调用MA均线函数的代码如下：

```
df=pdta.MA(df,5);
df=pdta.MA(df,10);
df=pdta.MA(df,30);
df=pdta.MA(df,50);
```

这里是首次介绍pandas_talib.py，下面看一下TA-Lib金融函数库里相关的MA均线函数源码。

pandas-talib模块库文件名为zwQuant\source\pandas_talib.py。

MA均线函数相关源码如下：

```
#Moving Average
def MA(df, n):
     MA = pd.Series(pd.rolling_mean(df['Close'], n), name = 'MA_
     df = df.join(MA)
     return df
```

MA均线函数定义很简单，需要注意以下几点。

- MA均线函数直接调用收盘价（Close）数据，生成日均线数据，而不是通过参数传递生成日均线数据，这样虽然保持了数据的一致性，但缺乏灵活性，无法生成开盘价、最高价、最低价及成交量等数据的日均平均数据。
- 数据格式采用的是雅虎网站的标准格式，首字母大写，与国内习惯全部小写不同，国内数据调用需要转换（zwQuant量化工具箱有转换函数）。
- 生成的日均数据作为一个新的数据列直接添加到原数据表格中。

案例2-3的运行结果如图2-24所示。

```
In [28]: runfile('F:/zwPython/zw_k10/k203ma_pd.py', wdir='F:/zwPython/zw_k10')
Reloaded modules: zwTools, pandas_talib, zwQTBox, zwSys
          Date       Open       High        Low      Close    Volume  \
247  2014-01-08  26.370001  27.680000  26.350000  27.629999  15733900
248  2014-01-07  27.200001  27.400000  26.670000  26.910000  11288700
249  2014-01-06  26.620001  27.200001  26.600000  27.030001  11271400
250  2014-01-03  25.750000  26.750000  25.510000  26.540001  13829600
251  2014-01-02  25.070000  25.820000  25.059999  25.360001   8997900

     Adj Close    MA_5   MA_10      MA_30    MA_50
247  27.256909  28.784  29.398  32.324667  34.3060
248  26.546632  28.392  28.969  32.033000  34.1130
249  26.665013  28.068  28.606  31.781000  33.9120
250  26.181629  27.506  28.258  31.522000  33.6958
251  25.017563  26.694  27.860  31.220333  33.4490
```

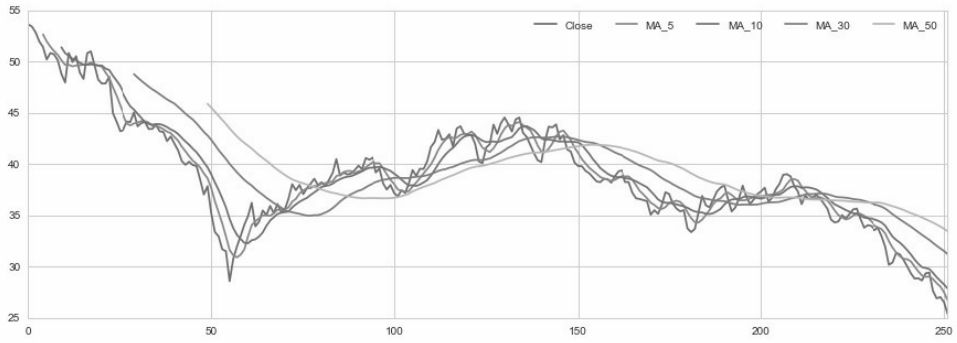


图2-24　MA均线策略运行结果

2.4.3　TA-Lib函数调用

TA-Lib金融函数库收录的函数已经超过100个技术指标，但部分复杂函数几乎很少用到。本书也只对涉及到的部分函数进行具体讲解，其他的函数请读者自行参考函数手册和网络资源进一步深入学习。

TA-Lib金融算法库原本是用C语言编写的，虽然有Python接口，但调用还是非常复杂。现阶段建议首先使用已经与Pandas数据分析软件整合的TA-Lib移植函数。

其他没有移植版本的函数，读者可以先看看PyAlgoTrade的二次封装版本，PyAlgoTrade已经在API层面考虑了与Pandas数据分析软件的整合，但目前仅支持Python 2.7版本，而且封装非常复杂。

例如，案例2-3中调用的MA均线函数很简单，PyAlgoTrade二次封

装版本的代码如下：

```
def MA(ds, count, timeperiod=-2**31, matype=0):
    """All Moving Average"""
    return call_talib_with_ds(ds, count, talib.MA, timeperiod,
```

系统不是直接调用均线函数，而是通过call_talib_with_ds间接调用函数。

call_talib_with_ds函数定义如下：

```
# Calls a talib function with the last values of a dataseries.
def call_talib_with_ds(ds, count, talibFunc, *args, **kwargs):
    data = value_ds_to_numpy(ds, count)
    if data is None:
        return None
    return talibFunc(data, *args, **kwargs)
```

由以上代码可以看出，TA-Lib金融函数库的直接调用相对比较复杂，因此尽量使用Pandas版本移植的TA-Lib金融函数，必须调用TA-Lib其他函数的，将在后面相关章节进行说明。

表2-1　主要TA-Lib函数

函　数	名　称	函　数	名　称
AD	随机指数	MFI	货币流量指数
ADOSC	佳庆指标	MINUS_DI	负向指标
ADX	平均趋向指数	MINUS_DM	负向运动
ADXR	评估指数	MOM	动量
APO	绝对价格振荡指数	NATR	归一化平均值范围
AROON	阿隆指标	OBV	能量潮
AROONOSC	阿隆震荡线	PLUS_DI	更向指示器

续表

函　数	名　称	函　数	名　称
ATR	平均真实波幅	PLUS_DM	定向运动
AVGPRICE	平均价格	PPO	价格振荡百分比
BBANDS	布林带	ROC	变动率指标
BETA	贝塔指数	ROCP	价格变化差值比率
BOP	均势指标	ROCR	价格变化差值率，百分比模式
CCI	顺势指标	ROCR100	价格变化率，百分比模式
CMO	钱德动量摆动指标	RSI	相对强弱指标
CORREL	皮尔森相关系数	SAR	抛物线转向
DEMA	双指数移动平均线	SAREXT	增强型抛物线转向
DX	动向指数	SMA	简单移动平均
EMA	指数移动平均线	STOCH	随机指标，又名 KD 指标
HT_DCPERIOD	希尔伯特变换，主周期	STOCHF	快速 STOCH 指标
HT_DCPHASE	希尔伯特变换，主阶段	STOCHRSI	随机强弱指数
HT_PHASOR	希尔伯特变换，相成分	T3	T3 指标
HT_SINE	希尔伯特变换，正弦波	TEMA	三指数移动平均，特马指标
HT_TRENDLINE	希尔伯特变换，瞬时趋势	TRANGE	真实范围
HT_TRENDMODE	希尔伯特变换-趋势与周期模式	TRIMA	三指数移动平均
KAMA	适应性移动平均线	TRIX	三重指数平滑平均线
MA	移动平均线	TSF	时间序列预测
MACD	指数平滑移动平均线	TYPPRICE	典型价格
MACDEXT	MA 型可控 MACD	ULTOSC	极限振子
MACDFIX	移动平均收敛/发散修复	WCLPRICE	加权收盘价
MAMA	MESA 移动平均线	WILLR	威廉指标
MAVP	变周期均值	WMA	加权移动平均
MEDPRICE	中位数价格		

2.4.4　量化分析常用指标

量化交易属于新兴产业，TA-Lib金融软件包收录的函数虽然多，但大部分是传统技术指标，对于量化交易、量化分析领域有很大缺失。本

节主要介绍量化领域新出现的一些技术指标。

量化分析本质上是一种风险和收益的综合平衡分析，其经典三大指标是夏普比率、詹森指数和特雷诺指数。

- 夏普比率（Sharpe Ratio），也称夏普指数，缩写是SR。简单地说，它指的是投资回报与风险的比例。夏普比率代表投资人每多承担一分风险，就可以拿到几分报酬，若为正值，代表基金报酬率高过波动风险；若为负值，代表基金操作风险大于报酬率。通常，这个比例越高，投资组合越佳。
- 詹森指数（Jensen），又称为阿尔法值，是衡量基金超额收益大小的一种指标。这个指标综合考虑了基金收益与风险因素，比单纯的考虑基金收益大小要更科学。
- 特雷诺指数（Treynor Ratio），缩写是TR，是每单位风险获得的风险溢价。特雷诺指数越大，单位风险溢价越高，绩效越好；相反，特雷诺指数越小，单位风险溢价越低，绩效越差。

在量化领域，其他常用的指标如下。

- 策略收益（Total Returns），策略实际收益额。
- 基准收益（Benchmark Returns），策略最低收益额。
- 基准收益率（Benchmark Yield），也称基准折现率，投资项目最低标准的收益水平，即选择特定的投资机会或投资方案必须达到的预期收益率。
- 阿尔法（Alpha），是高于预期收益率的超额收益率。阿尔法策略基于CAPM模型。传统阿尔法策略是在基金经理建立了Beta部位的头寸后，通过衍生品对冲Beta部位的风险，从而获得正的阿尔法收益。

- 贝塔（Beta），按照CAPM模型的规定，Beta系数是用以度量一项资产系统风险的指数，是用来衡量一种证券或一个投资组合相对总体市场的波动性（volatility）的一种风险评估工具。
- 索提诺比率（Sortino Ratio），缩写是SR，与夏普比率类似，所不同的是它区分了波动的好坏，因此在计算波动率时它所采用的不是标准差，而是下行标准差。这其中的隐含条件是投资组合的上涨（正回报率）符合投资人的需求，不应计入风险调整。和夏普比率类似，这个比率越高，表明基金承担相同单位下行风险能获得更高的超额回报率。索提诺比率可以看做是夏普比率在衡量对冲基金/私募基金时的一种修正方式。
- 信息比率（Information Ratio），也称为绩效评估比率（Appraisal Ratio），缩写是IR。信息比率是从主动管理的角度描述风险调整后的收益，它不同于夏普比率从绝对收益和总风险角度来描述。信息比率越大，说明所获得的超额收益越高，策略收益持续优于大盘的程度越高。计算方式是：将基金报酬率减去同类基金或者大盘报酬率（剩下的值为超额报酬），再除以该超额报酬的标准差。
- 策略波动率（Algorithm Volatility），是指回报率在策略执行期间内所表现出的波动率。
- 基准波动率（Benchmark Volatility），基准指数的波动范围，用来测量基准的风险性，波动越大代表基准风险越高。
- 下行风险（Downside Risk），是指由于市场环境变化，未来价格走势有可能低于分析师或投资者所预期的目标价位。下行风险是投资可能出现的最坏的情况，也是投资者可能需要承担的损失。
- 最大回撤率（drawdown），是指该金融产品历史上一段时间的最大跌幅。最大回撤率用来描述买入产品后可能出现的最糟糕的情况。最大回撤是一个重要的风险指标，对于对冲基金和量化策

略交易来说，该指标比波动率还重要。

2.5 经典量化策略

2.5.1 阿尔法（Alpha）策略

使用Alpha策略，通俗地说就是少挣点没关系，别赔本就可以了。Alpha策略是指利用管理人选股和择时上的优势，寻找具有稳定Alpha收益的现货组合，通过衍生品（股指期货、期权、互换等）来分离Beta，获得与市场相关性较低的Alpha收益。

Alpha是高于经Beta调整后的预期收益率的超额收益率，也就是说超越基准指数的收益率。经统计，20世纪中叶约75%的股票型基金经理构建的投资组合，无法跑赢根据市值大小构建的简单组合或指数。

随着金融衍生品的诞生，不少基金取得了令人眩目的收益率，这说明通过积极的投资管理是可以获得超额收益率的。

Alpha策略的关键是：通过衍生品来对冲投资组合的系统风险Beta，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Beta，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和沪深300指数期货空头策略。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。

从实盘效果来看，成熟市场一般不存在或者很难寻找稳定的Alpha。因此，阿尔法策略一般运用在市场效率相对较弱的市场上，如

新兴股票市场、创业板市场等。我国的股票市场正是一个新兴的市场，效率相对较低，从过往的经验及研究来看，的确存在着超额收益Alpha。

Alpha策略成败的两个关键要素是：其一，现货组合的超额收益空间有多大；其二，交易成本的高低。两者相抵的结果才是Alpha策略可获得的利润空间。

Alpha策略的关键是获取稳定的Alpha超额收益率，稳定也是所有量化策略的核心。通常，Alpha机会通常存在于以下几个方面。

- 基金组合。国外很多关于Alpha策略的研究中，一般都是通过基金和衍生品的组合来构造Alpha组合的，这是因为基金较易获得超额收益。研究表明，基金的Alpha体现了基金经理的管理能力，过去的成功有很大的概率在未来时段得到延续，因此通过寻找具有稳定Alpha收益的基金来构建Alpha策略是可行的。以沪深300指数作为市场基准收益，计算基金的詹森（Jensen）指标、信息比率（IR）指标、夏普比率（SR）指标等，通过上述指标对基金进行排序，选取表现“好”的基金，然后按照等权重或等市值权重构建基金组合。
- 选股。市场上总是存在被低估的股票，如果能够准确地找出这些被低估的股票，买入这些股票构建投资组合，并通过衍生品对冲组合的系统风险，我们就可以获得稳定的Alpha收益。在选股的过程中，不仅可以采用基本面分析，还可以通过一系列量化方法来帮助选股。
- 传统的基本面分析。通过估值水平、盈利能力、盈利质量、成长能力、运营能力、负债水平等方面来综合评价上市公司，筛选出具有超额收益的股票。一套选股指标体系一般难以适应所有行

情，最有效的选股指标在不同行情下各有差异。因此，在不同的市场阶段挖掘超越指数的选股指标对基本面分析至关重要。

- 动量策略。在物理学上，运动的物体停止受力后，因为存在动量，仍会在原有的轨道上运行一段距离。金融市场同样存在类似的动量效应（momentum）。在一定时间内，如果某只股票或者某个股票组合在前一段时期表现较好，那么之后一段时期该股票或者股票投资组合仍将有良好表现，这就是金融市场上的动量效应。以此为基础，我们可以扩展Alpha动量策略、IR动量策略等，据此挑选股票。
- 波动捕获策略。在效率相对较低的市场，某些个股会有比市场指数更高的波动性。波动捕获策略就是寻找波动性大且相关性低的股票并构建组合，获取Alpha收益。
- 行业Alpha。由于个股的数量大、波动性大，在寻找Alpha时往往比较复杂。另一种思路是寻找行业Alpha，国内外很多基金的大部分收益都来源于正确的行业配置。类似的，在寻找行业Alpha方面，我们同样既可以从基本面角度来分析特定行业的运行特征，也可以将单个行业看成单只股票，利用与上文类似的量化方法进行选择。

此外，我们还可以通过股指期货套利、债券套利、ETF套利等套利策略来获得Alpha收益。

在我国证券市场中，沪深300股指期货给投资组合对冲系统风险提供了有力的工具。不过，目前我国期货市场还不完善，缺乏卖空机制，这方面的实盘操作需要非常谨慎。对于现货多头，可以通过卖空股指期货来对冲风险。

构建好Alpha策略组合后，需要密切跟踪组合的表现。即使历史时

期具有非常稳定和出色的Alpha收益，未来时段也可能发生改变。在期货头寸方面，除了需要关注系统风险Beta的变化之外，还应该留意期货行情的波动。当组合风险敞口达到一定阀值，就应该及时调整期货头寸，以匹配资产组合的系统风险。

值得注意的是，在动态调整之前，应该充分考虑交易的费用和成本。频繁地调整必然增加交易成本，降低整体收益。

2.5.2 Beta策略

简单来说，Beta策略就是当判断股票要上涨时增大Beta值，多买点股票，预测下跌时卖点股指期货或者股票。Beta策略是指被动跟踪指数的策略，从长期来讲，Beta策略是可能盈利的，但由于股票市场波动比较大，因此在某段特定时期内往往会出现亏损或被套住的情况。

Beta策略在上涨趋势和下跌趋势中都好于对冲策略。比如在上涨趋势中，要么只做多股票，要么只做多期货指数；在下跌趋势中，要么只做空期货指数，要么只融券卖出。当然，这要求投资者对行情中长期的趋势要有准确的判断。

Alpha策略与Beta策略是两类基于不同出发点来获取超过大盘表现的超额收益的投资策略。Alpha策略是依靠精选行业和个股来超越大盘；Beta策略是依靠准确地把握市场大势、准确择时来获得超越大盘的收益。不同的侧重点使得两者在投资理念、仓位控制、风险控制等方面都存在差异，进而在不同的市场行情中表现各异。

按照CAPM模型的规定，Beta系数是用以度量一项资产系统风险的

指针，是用来衡量证券或投资组合相对总体市场的波动性（volatility）的一种风险评估工具。

Beta系数表示的是相对于市场收益率变动、个别资产收益率同时发生变动的程度，是一个标准化的度量单项资产对市场组合方差贡献的指标。

若一只股票的价格和市场的价格波动性是一致的，那么这只股票的Beta值就是1。如果一只股票的Beta值是1.8，那么就意味着当市场价格上升10%时，该股票价格上升18%；在市场价格下降10%时，该股票的价格也会下降18%。Beta值是通过统计和分析同一时期市场每天的收益情况及单只股票每天的价格收益来计算的。

当Beta值处于较高位置时，投资者便会因为股票的风险高，而相应提升股票的预期回报率。

【示例】

如果一只股票的Beta值是2.5，无风险回报率是2%，市场回报率（Market Return）是10%，那么市场溢价（Equity Market Premium）就是8%（2%~10%），股票风险溢价（Risk Premium）为20%（2.5×8%，用Beta值乘以市场溢价），那么股票的预期回报率则为22%（20%+2%，即股票的风险溢价加上无风险回报率）。

上面的示例说明，一个风险投资者需要得到的溢价可以通过CAPM模型计算出来。换句话说，可通过CAPM模型知道股票的价格是否与其回报相吻合。

2.5.3 海龟交易法则

海龟交易法则源自于海龟实验的公开交易系统，在1982年由全球著名的商品投机家理查德·丹尼斯在一个交易员培训班上推广而闻名，它涵盖了交易系统的各个方面。海龟实验证明了交易可以被传授，证明了用一套简单的法则可以使仅有很少或根本没有交易经验的人成为优秀的交易员。

海龟（操盘手培训）实验是金融史上著名的实验，在实验之后的4年中，这些选手取得了年均复利80%的收益。简单来说，就是根据实盘经验或者统计分析，找出大概率的正盈利产品，设定好收益点、止损点，然后严格按照预设的程序进行操作。海龟交易法则非常适合计算机量化分析模式，所以在最近几年的量化投资热浪中再一次成为热门模式。

1. N系数

海龟交易法则用一个理查德·丹尼斯和比尔·埃克哈特称之为N系数的概念，来表示某个特定市场根本的波动性。

N系数就是TR（True Range，实际范围）的20日指数移动平均，现在更普遍地称之为ATR平均真实波幅参数。从概念上来看，N系数表示单个交易日某个特定市场所造成的价格波动的平均范围，它说明了开盘价的缺口，N系数同样用构成合约基础的点（points）进行度量。

每日实际范围的计算如下：

TR（实际范围）=max(H-L,H-PDC,PDC-L)

其中：

H，当日最高价；

L，当日最低价；

PDC，前一个交易日的收盘价。

可以用下面的公式计算N：

N=(19×PDN+TR)/20

其中：

PDN，前一个交易日的N值；

TR，当日的实际范围。

因为这个公式要用到前一个交易日的N值，所以必须从实际范围的20日简单平均开始计算初始值。

2. 价值量波动性的调整

海龟法则的核心是确定价值量波动性，也称市场价格波动性，用N值表示：

确定头寸规模的第一步是确定用市场价格波动性（用其N值定义）表示的价值量波动性。

价值量波动性=N×每点价值量

海龟交易法则按照我们所称的单位（Units）建立头寸。单位按大小

排列，如1N代表账户净值的1%。因此，特定市场或特定商品的单位可用下面的公式计算：

单位=账户的1%/市场价值量波动性

或

单位=账户的1%/(N×每点价值量)

3. 海龟法则四大核心

海龟交易法则的四大核心内容如下：

- 掌握优势。找到一个期望值为正的交易策略，因为从长期看，它能创造正的回报。
- 管理风险。控制风险、守住阵地，否则你可能等不到创造成果的那一天。
- 坚定不移。唯有坚定不移地执行你的策略，你才能真正获得系统的成效。
- 简单明了。从长久看，简单的系统比复杂的系统更有生命力。

2.5.4 ETF套利策略

交易型开放式指数基金通常又被称为交易所交易基金（Exchange Traded Funds，简称ETF），是在交易所上市交易的、基金份额可变的一种开放式基金。

ETF指数套利是指投资者同时交易股指期货合约和相对应的一篮子

股票的交易策略，以谋求从期货、现货市场同一组股票存在的价格差异中获利。

假设在某个时段中，某只ETF成分股暴跌，使得该ETF的净值迅速走低，但该ETF的市场价格未能及时跟上，两者短暂地出现了一个价差，这时就可以选择买入ETF一篮子股票组合申购成ETF（以净值计价），然后将ETF在二级市场上出售（以市场价格计价），从而实现低买高卖，获取价差。

ETF基金和自建的一篮子股票都可以理解为以指数为基础的衍生品，其定价都以指数为基础，而且可以同时交易，只要它们之间定价存在偏差，并且足以弥补交易费用和延时风险，便有可能产生一定的套利机会，这就是常说的ETF套利，也是国内证券市场近几年唯一被成功实践的指数套利模式。

2.6 常用量化策略

笔者曾经在博客中说过：“人人都大数据，就人人都无数据”。

市场本身是自适应的，任何公开的模型、策略用得人多了，市场就会产生预判效果，类似生物疫苗，产生免疫作用，从而无法长期稳定获利。在这方面，最典型的案例是MACD（指数平滑移动平均线）指标，读者可以看看MACD相关的金融史，就会发现，MACD这个指标从个别团体走向小众市场，再到大众领域，其正收益的关联性是如何逐步下降的。

目前，国际投行出于策略保密的需要，要求一线交易员直接编程，这个也是Python编程语言在量化领域崛起的重要原因。

Python语言是唯一适合非专业IT人员的通用编程语言，已经在天文学、化学、生物学、地理学领域得到证实。Python已经是这些领域的行业标准，属于编程工作语言。目前，随着数据分析、金融量化的发展，Python已经是无限接近“标准”级的编程工作语言。

量化策略一般分为三个类型：

- 动量交易策略；
- 均值回归策略；
- 其他策略。

2.6.1 动量交易策略

动量交易策略即预先对股票收益和交易量设定过滤准则，当股票收益或股票收益和交易量同时满足过滤准则就买入或卖出股票的投资策略。动量交易策略常见的操作手段有：VWAP动量策略、SMA简单交叉线策略、market_timing适时进出投资策略。

1. VWAP动量策略

VWAP（Volume Weighted Average Price）是成交量加权平均价的意思。VWAP策略是最常用的量化交易策略之一，具有简单、易操作等特点，让大部分成交价在指定的价格之下。策略的基本思想就是让自己的交易量提交比例与市场成交量比例尽可能地匹配，在减少对市场冲击的同时，获得市场成交均价的交易价格。

2. SMA简单交叉线策略

SMA简单交叉线策略就是利用简单移动平均线设定一个买卖系统。将两条平均线中较短的一条作为信号线。例如，当短期的平均线上穿较长期的平均线时，显示牛市在即；相反，当短期的平均线下穿较长期的平均线时，显示熊市在即。

简单移动平均线也可作为支持及阻力之用。当市况向跌破时沽出，升破时买进，简单移动平均线便会在部分情况下反映强烈地支持或阻止。

3. 适时进出投资策略

market_timing适时进出投资策略类似基金定投，指通过技术性指标的分析，将市场行情作为买卖基金和股票的依据，当预测行情即将下跌时，就卖出，减少手中的基金和股票份额；反之，就买进。适时进出投

资策略，试图用技术指标的各项分析来预测股市的走势，希望能在股价的相对底部买进，于相对顶点卖出，以赚取其中价差。据统计表明，要想运用适时进出投资策略获利，必须至少有70%以上的判断准确率，否则考虑到交易成本的存在，投资者至多打个平手，甚至可能得不偿失。

2.6.2 均值回归策略

均值回归是金融学的一个重要概念。均值回归是指股票价格无论高于或低于均值，都会以很高的概率向均值回归的趋势。根据这个理论，股票价格总是围绕其平均值上下波动。一种上涨或者下跌的趋势，不管其延续的时间多长都不能永远持续下去，最终均值回归的规律一定会出现：涨得太多了，就会向平均值移动（下跌）；跌得太多了，也会向平均值移动（上涨）。

均值回归投资策略的一个难点是，到目前为止，均值回归策略仍不能解决的或者不能预测的是回归的时间间隔，即回归的周期为“随机漫步”。不同的股票市场，回归的周期不一样，即使是同一个股票市场，每次回归的周期也不一样。

1. 布林带策略

布林带（Bollinger Bands）策略源自Boll指标（布林线指标）。布林线（Boll）由约翰·布林先生创造，利用统计原理，求出股价的标准差及其信赖区间，从而确定股价的波动范围及未来走势，利用波带显示股价的安全高低价位，因此也被称为布林带。

布林带上下限范围不固定，随股价的波动而变化。布林线指标和麦

克指标同属路径指标，股价波动在上限和下限的区间内，这条带状区的宽窄随着股价波动幅度的大小而变化。股价涨跌幅度加大时，带状区变宽；涨跌幅度狭小、盘整时，带状区则变窄。

在众多技术分析指标中，Boll指标属于比较特殊的一类指标。绝大多数技术分析指标都是通过数量的方法构造出来的，它们本身不依赖趋势分析和形态分析，而Boll指标却和股价的形态和趋势有着密不可分的联系。

Boll指标中的“股价通道”概念正是股价趋势理论的直观表现形式。Boll指标利用“股价通道”来显示股价的各种价位：

- 当股价波动很小、处于盘整时，股价通道就会变窄，这可能预示着股价的波动处于暂时的平静期；
- 当股价波动超出狭窄的股价通道的上轨时，预示着股价异常激烈的向上波动即将开始；
- 当股价波动超出狭窄的股价通道的下轨时，同样也预示着股价异常激烈的向下波动将开始。

2. RSI2策略

RSI（Relative Strength Index）为相对强弱指标，由尔斯·怀尔德（Welles Wilder）最早应用于期货买卖，后来人们发现，RSI极其适合于股票市场的短线投资，于是用于股票升跌的测量和分析中。RSI2策略是指同时对空头、多头的RSI指标进行对比分析，根据历史数据统计和分析，设置相关的参数，从而挑选出获利的股票产品和操作时机。

2.6.3 其他常用量化策略

量化策略通常包括量化选股策略和量化择时策略两个类型。

量化选股策略就是利用统计和分析的方法构建股票组合，期望该股票组合能够获得超越基准收益率的投资行为。量化选股策略总的来说可以分为两类：第一类是基本面选股，第二类是市场行为选股。

基本面选股包括多因子模型、风格轮动模型和行业轮动模型等。市场行为选股包括资金流模型、动量反转模型、一致预期模型、趋势追踪模型和筹码选股模型等。

量化择时策略就是利用量化分析的方法，通过对各种宏观和微观指标的量化分析，找到最佳的市场高位点、低位点，从而进行投资操作。常见的量化择时模式有趋势择时、市场情绪择时、有效资金模型、牛熊线、Hurst指数、SVM分类、SWARCH模型及异常指标模型等。

看到这么多专业的术语，读者可能有些眼花缭乱，这是正常的。做量化分析，包括进行策略函数编程，读者要有一个最基本的原则，就是：

用户是金融行业的，不是IT行业的程序员；

只需要做应用一级的金融数据分析，绝对不要涉及背后的理论研究、底层设计。

量化投资属于新兴产业，相关理论、名词发展变化很快，以上划分只是笔者的一家之言。相关的模型、框架、理论及具体的程序实现版本

很多，了解概况就可以了。真正实盘操作，还要参考具体的量化分析软件平台和平台的模块库。

1. 三连涨策略

下面举例说明三连涨策略。

假设通过分析历史数据得知，股票或期货如果连涨3天，则第4天也会涨（当然一般看期望而不是概率），那么就可以在3天连续牛市末建仓，在第4天末清仓。

三连涨策略其实也是一种动量策略。

2. MACD策略

MACD（Moving Average Convergence and Divergence）为指数平滑移动平均线。MACD是从双指数移动平均线发展而来的，由短期（常用为12日）的指数移动平均线（EMA）减去长期（常用为26日）的指数移动平均线，MACD的意义和双移动平均线基本相同，但阅读起来更方便。

当MACD从负数转向正数，是买的信号；当MACD从正数转向负数，是卖的信号。当MACD以大角度变化时，表示快的移动平均线和慢的移动平均线的差距迅速拉开，代表了一个市场大趋势的转变。

MACD策略的有关参数如下：

- 短期EMA：短期（例如12日）的收盘价指数移动平均值。
- 长期EMA：长期（例如26日）的收盘价指数移动平均值。
- DIF（Difference）线：短期EMA和长期EMA的离差值。

- DEA（Difference Exponential Average）线：DIF线的M日指数平滑移动平均线。
- MACD线：DIF线与DEA线的差。

MACD策略的参数有：SHORT（短期）、LONG（长期）、M天数（一般为12、26、9）。

指数加权平滑系数分为：

- 短期EMA平滑系数：2/(SHORT+1)。
- 长期EMA平滑系数：2/(LONG+1)。
- DEA线平滑系数：2/(M+1)。

MACD策略交易时机如下：

- DIF从下而上穿过DEA，买进；
- DIF从上往下穿过DEA，卖出。

2.7 起点与终点

案例2-1从某种程度上而言是量化学习的起点，也是终点。

对于量化初学者而言，整个学习过程就是掌握、理解案例2-1背后的整个量化流程，以及灵活运用其他策略，取代案例程序中的SMA平均线策略。

对于量化专业人员而言，学习相关的代码构成，以及背后的基本逻辑，独立完成相关的Backtest（回溯）测试模块编程。

策略是量化交易的核心，无论是量化初学者，还是专业人员，都需要重点学习Backtest（回溯）测试模块部分。

至于量化程序前端的数据API接口与数据采集模块、后端的Trade（下单交易）模块，可以通过有关金融数据网站的API接口或者CSV等数据文件格式，甚至通过第三方软件实现；量化交易中的部分环节，特别是非高频的下单环节，也可以采用人工模式。

本书以实战为主，理论方面尽量点到为止。本章从全局角度介绍Python量化的各个方面，把相关理论集中在一起。这样集中讲解，便于读者从整体上把握Python行业的整体状况，培养大局观，而且相关的理论资料汇集在一起，也便于资料查找。

第3章
金融数据采集整理

常言道：兵马未动，粮草先行。对于量化分析而言，这个“粮草”指的就是专业的量化分析软件与金融数据源。

Python编程语言是金融行业量化分析的标准编程语言；zwPython集成版是Python开发平台，也是目前行业内领先的量化分析平台，配备了大量的专业金融模块库；zwQuant开源量化分析软件，除了配套的中文用户手册外，还对各个模块的函数添加了中文注释，这种函数级的中文注释在开源软件中是非常罕见的，此外，www.ziwang.com网站也提供了大量的中文量化课件与文档，这些都为广大的量化初学者提供了极大的方便。

zwDat金融数据包是目前行业内领先的股票数据源。它提供了：

- A股数据，收集了近2700只中国股票的日线数据，从1994年—2016年2月7日（春节休市）；
- 美股数据，收集了近7000只美股的日线数据，从1960年—2016年1月；
- 独家提供归一化数据栏，为不同区域、行业的股票对比分析提供了基础数据源。

因为股票数据有时效性，需要每天实时更新数据，所以在介绍量化

分析前，先介绍相关的金融数据采集工作，以及简单的数据整理、归一化处理工作。

量化看起来很高深、很神秘，其实不过是一套高级版本的Excel电子表格软件。事实上，就软件工程而言，目前最复杂的Python量化系统也比Excel电子表格软件要简单得多。

注意

本书采用的案例脚本源自zwQuant股票下载案例程序，为了演示需要，增加了中间数据输出及部分绘图语句，因此运行效率方面会有所降低。为了避免对zwDat数据源的数据覆盖，书中对部分数据目录、文件名进行了修改。此外，为方便初学者学习，取消了批量自动下载功能，重点讲解单一股票的数据采集。

3.1 常用数据源API与模块库

近两年，随着大数据的兴起，开放的金融数据源API接口越来越多。此类数据源API从收费模式来看，分为收费API与免费API两种；从行业背景来看，分为大数据综合API和专业财经数据API。

3.1.1 大数据综合API

目前，随着互联网行业和大数据领域的综合，出现了许多综合性的大数据网站，这些网站可以提供各种各样的金融数据源，如图3-1所示。



图3-1　大数据综合网站

大数据综合网站有：

- 开发者数据，http://www.haoservice.com/。
- 聚合数据，https://www.juhe.cn/。
- 阿凡达数据，http://www.avatardata.cn/。
- 免费接口API汇集，http://www.apifree.net/。
- 91查，http://www.91cha.com/。
- 云聚数据，http://www.36wu.com/。
- JSON　　API免费接口，http://www.bejson.com/knownjson/webInterface/。
- 百度API集市，http://apistore.baidu.com/。

3.1.2　专业财经数据API

金融数据由于专业性和实时性，一般使用专业的金融财经数据公司提供的数据作为数据源。专业金融数据网站如图3-2所示。



图3-2　专业金融数据网站

专业金融数据网站有：

- TuShare中文财经API，http://tushare.org/。
- 通联数据商城，https://app.wmcloud.com/datamkt/dashboard?lang=zh。
- 百度股票大盘历史数据免费API，http://apistore.baidu.com/apiworks/servicedetail/1626.html。
- 大富翁数据中心，http://www.licai668.cn/index.asp。
- 新浪股票数据接口，以大秦铁路（股票代码：601006）为例，http://hq.sinajs.cn/list=sh601006。
- 百度财经数据，http://stock.baidu.com/。
- ChinaStock的WebService，http://www.webxml.com.cn/WebServices/ChinaStockWebService.asmx。

3.1.3 专业数据模块库

早期进行Python股票数据分析时，需要根据API进行网络编程并采集数据。数据采集难度很小，但很啰嗦，因为各个数据源分得很散，用户需要到不同的网站抓取信息。

现在，有了专业的财经数据模块库，用户就方便多了。国内财经数据采集主要使用TuShare模块库，国外财经数据可以直接用Pandas（潘达思）软件内置的数据采集模块。

3.2　案例3-1：zwDatX数据类

zwDatX数据类是zwSys.py模块定义的类，主要是保存初始化数据，如相关的目录设置。为便于讲解，笔者把相关代码移植到案例3-1。

本章开始会涉及部分zwQuant量化程序的案例，凡是与zwQuant量化软件相关的案例程序，笔者会用字母“zq”作为文件名开头，以便管理。

脚本文件名为\zwpython\zw_k10\ zq001.py。

全部代码如下：

```
# -*- coding: utf-8 -*-
import numpy as np
import pandas as pd
#import pandas.io.data as web
import pandas_datareader.data as web
#-----------
#----init.dir


_rdat0="\\zwDat\\"
_rdatCN=_rdat0+"cn\\"
_rdatUS=_rdat0+"us\\"
_rdatInx=_rdat0+"inx\\"
```

```
_rdatMin=_rdat0+"inx\\"
_rdatZW=_rdat0+"zw\\"

#----class
class zwDatX(object):
     def __init__(self,rs0=_rdat0):    #默认有此函数，实例化时首次运行
         self.rdat=rs0;
         self.rdatCN=_rdatCN;
         self.rdatUS=_rdatUS;
         self.rdatInx=_rdatInx;
         self.rdatMin=_rdatMin;

         self.rdatZW=_rdatZW;
         self.rZWcnXDay=_rdatZW+"cnXDay\\"
         self.rZWcnDay=_rdatZW+"cnDay\\"
         self.rZWusDay=_rdatZW+"usDay\\"

         self.rXDay=rs0+"xday\\"
         self.rDay9=rs0+"day9\\"
         self.rDay=rs0+"day\\"

         self.rM05=_rdatMin+"m05\\"
         self.rM15=_rdatMin+"m15\\"
         self.rM30=_rdatMin+"m30\\"
         self.rM60=_rdatMin+"m60\\"

         self.code=""
```

```
        self.cname=""

    def prDat(self):
        print('rdat',self.rdat)
        print('rdatCN',self.rdatCN)
        print('rdatUS',self.rdatUS)
        print('rdatInx',self.rdatInx)

        print('')
        print('rdatZW',self.rdatZW)
        print('rZWcnXDay',self.rZWcnXDay)
        print('rZWcnDay',self.rZWcnDay)
        print('rZWusDay',self.rZWusDay)

        print('')
        print('XDay',self.rXDay)
        print('Day9',self.rDay9)
        print('Day',self.rDay)

        print('')
        print('rdatMin',self.rdatMin)
        print('rM05',self.rM05)
        print('rM15',self.rM15)
        print('rM30',self.rM30)
        print('rM60',self.rM60)

        print('')
```

```
        print('code',self.code)
        print('name',self.cname)


#------------


rs0="\\zwDat\\us\\"
#rs0="\\zwDat\\cn\\"
qx=zwDatX(rs0);
qx.prDat();
```

上面的脚本看起来很长，是因为初始化需要设置很多变量参数，但其实就是针对两个函数：

- __init__内置参数设置初始化函数；
- prDat内置方法，打印变量数据。

整个代码分为以下几个部分：

- import模块库导入；
- 预定义变量设置，类似其他语言的const常数变量；
- class类定义，包括数据初始化、prDat方法定义；
- 测试脚本。

zwDatX数据类目前只有一个prDat方法，用于打印设置的数据。

初学者注意

Python的class类更多的和C语言的struc结构、Delph语言的record纪录类似，而不是Java。读者可以把其视为简单的数据变量集合和内置函数集合。除了__init__内置初始化函数外，尽量不要引用“__”结构的内置函数。

运行结果如下：

```
runfile('F:/zwPython/zw_k10/zq001.py', wdir='F:/zwPython/zw_k10'
rdat \zwDat\us\
rdatCN \zwDat\cn\
rdatUS \zwDat\us\
rdatInx \zwDat\inx\

rdatZW \zwDat\zw\
rZWcnXDay \zwDat\zw\cnXDay\
rZWcnDay \zwDat\zw\cnDay\
rZWusDay \zwDat\zw\usDay\

XDay \zwDat\us\xday\
Day9 \zwDat\us\day9\
Day \zwDat\us\day\

rdatMin \zwDat\min\
rM05 \zwDat\min\m05\
rM15 \zwDat\min\m15\
```

```
rM30 \zwDat\min\m30\
rM60 \zwDat\min\m60\


code
name
```

案例3-1是使用zwQuant量化分析软件的第一步。软件设计有多种模式，不过在项目起步阶段，定义一个class数据类用于保存系统参数变量是一种惯例，对于保持系统的统一性、调试都方便不少。

随着zwQuant量化分析软件的发展，未来实际工程中的zwDatX类定义可能与以上的案例版本有很大的不同，希望读者注意。

3.3 美股数据源模块库

采集美股数据源相对简单些，所以将其放在采集中国A股数据源的前面介绍。如图3-3所示是Pandas数据分析软件独立的datareader数据采集模块。

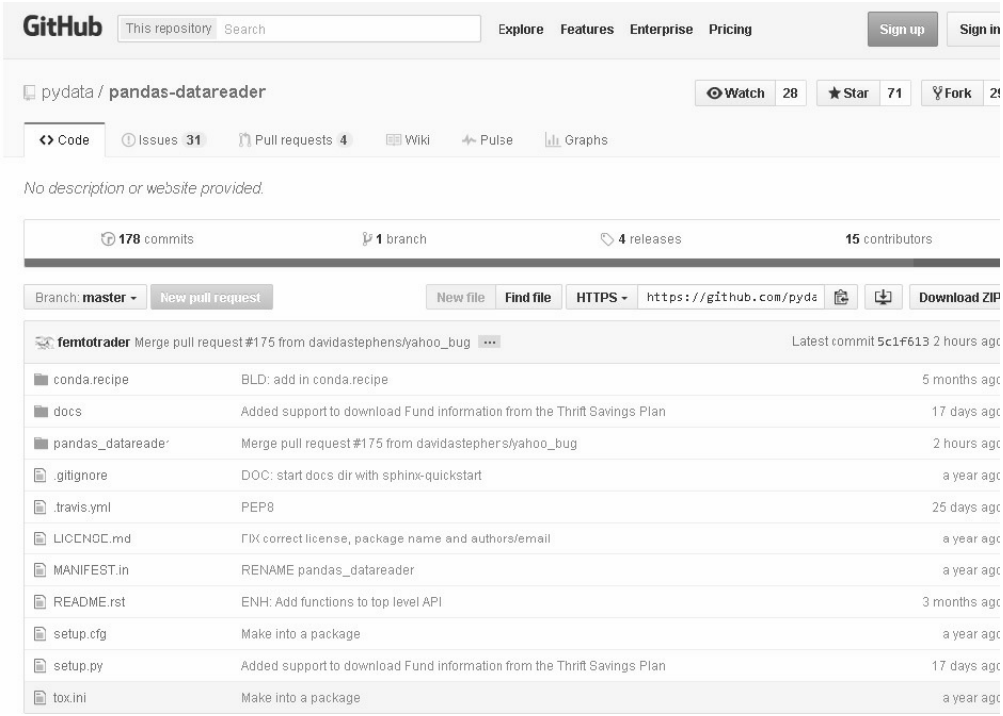


图3-3 Pandas数据分析软件独立的datareader数据采集模块

美股数据源主要使用Pandas数据分析软件内置的数据采集模块来采集。需要注意的是，Pandas 0.17版本后，相关的数据采集模块进一步独立，并且拆分成单独的项目，即pandas-datareader。

pandas-datareader的网址是https://github.com/pydata/pandas-datareader。其文档网址是http://pandas-datareader.readthedocs.org/en/latest/index.html。

原来数据采集模块的名称是pandas.io.data，现在需要改为pandas_datareader. data。

原语句：

```
from pandas.io import data, wb # becomes
```

新语句：

```
from pandas_datareader import data, wb
```

缩写规范：

```
import pandas_datareader as pdr
```

3.4　开源文档库Read the Docs

如图3-4所示是开源文档库Read the Docs的网站截图。



图3-4　开源文档库Read the Docs的网站截图

pandas-datareader文档在开源文档库中的网址是http://pandas-datareader. readthedocs.org/en/latest/index.html。

请读者注意，该网站的主域名为http://readthedocs.org。

Read the Docs网站是目前全球最大的开源文档中心，许多GitHub开源网站的项目直接把其作为项目的官方文档网站。网站有中文界面，无需注册就可以下载文档。读者若要寻找各种软件的说明文档，可以首先考虑这个网站。

Python量化常用的软件模块库，如Pandas（潘达思）数据分析软件、TA-Lib金融函数库、PyAlgoTrade量化软件、Zipline回溯模块等，

都可以在Read the Docs网站找到相关的用户文档，部分文档还有中文版本。

3.5　案例3-2：下载美股数据

下面通过具体案例来进一步介绍pandas-datareader。

案例文件名为\zwpython\zw_k10\k301_down_us.py。

全部代码如下：

```
# -*- coding: utf-8 -*-

import numpy as np
import pandas as pd
#import pandas.io.data as web
import pandas_datareader.data as web

#zw.Quant
import zwSys as zw
import zwQTBox as zwBox

#-----------

def zw_down_yahoo8code(qx):
     try:
         xcod=qx.stkCode;
         xdat= web.DataReader(xcod,"yahoo",start="1/1/1900");
```

```
        fss=qx.rDay+xcod+".csv";print(fss);
        xdat.to_csv(fss);
    except IOError:
        pass    #skip,error

#------------

qx=zw.zwDatX(zw._rdatUS);
qx.prDat();

#
code='USA';qx.stkCode=code;qx.rDay="tmp\\";
zw_down_yahoo8code(qx);
```

核心的代码就是下面这句，起始时间设为1900年，按指定的股票代码从Yahoo下载全部交易数据。

```
xdat= web.DataReader(xcod,"yahoo",start="1/1/1900");
```

运行结果如下：

```
runfile('F:/zwPython/zw_k10/k301_down_us.py', wdir='F:/zwPython/
Reloaded modules: zwQTBox, zwSys
rdat \zwDat\us\
```

```
rdatCN \zwDat\cn\
rdatUS \zwDat\us\
rdatInx \zwDat\inx\

rdatZW \zwDat\zw\
rZWcnXDay \zwDat\zw\cnXDay\
rZWcnDay \zwDat\zw\cnDay\
rZWusDay \zwDat\zw\usDay\

XDay \zwDat\us\xday\
Day9 \zwDat\us\day9\
Day \zwDat\us\day\

rdatMin \zwDat\inx\
rM05 \zwDat\inx\m05\
rM15 \zwDat\inx\m15\
rM30 \zwDat\inx\m30\
rM60 \zwDat\inx\m60\

code
name
tmp\USA.csv
```

最后一行的“tmp\USA.csv”是下载的美股数据文件名。如图3-5所示是美股数据文件。

	A	B	C	D	E	F	G
1	Date	Open	High	Low	Close	Volume	Adj Close
2	2015-1-14	5.73	5.76	5.7	5.76	236900	5.130823
3	2015-1-15	5.79	5.79	5.67	5.7	277100	5.077377
4	2015-1-16	5.68	5.77	5.67	5.77	387600	5.139731
5	2015-1-20	5.79	5.81	5.71	5.74	468100	5.113007
6	2015-1-21	5.64	5.68	5.62	5.65	217400	5.122074
7	2015-1-22	5.68	5.76	5.64	5.73	464400	5.194598
8	2015-1-23	5.72	5.75	5.72	5.74	324000	5.203664
9	2015-1-26	5.73	5.76	5.71	5.74	242300	5.203664
10	2015-1-27	5.68	5.74	5.65	5.72	335600	5.185533
11	2015-1-28	5.75	5.75	5.58	5.61	384500	5.085811
12	2015-1-29	5.63	5.64	5.54	5.62	381100	5.094876
13	2015-1-30	5.59	5.63	5.56	5.57	390500	5.049549
14	2015-2-2	5.6	5.61	5.52	5.6	1029300	5.076745
15	2015-2-3	5.66	5.7	5.65	5.68	544800	5.14927
16	2015-2-4	5.69	5.74	5.68	5.69	551200	5.158336
17	2015-2-5	5.7	5.78	5.7	5.75	554700	5.21273
18	2015-2-6	5.77	5.79	5.71	5.72	373900	5.185533

图3-5　美股数据文件

注意，表格的名称和排列顺序与我国的股票数据格式有所不同。如图3-6所示是我国股票数据文件。

	A	B	C	D	E	F	G
1	date	cpen	high	close	low	volume	amount
2	2016-2-4	7.23	7.45	7.32	7.16	32063180	234423328
3	2016-2-3	7.03	7.4	7.25	6.93	33325940	240750960
4	2016-2-2	6.98	7.25	7.13	6.83	29214640	207110576
5	2016-2-1	7.28	7.48	6.92	6.65	4009266	286401600
6	2016-1-29	6.29	6.96	6.96	6.21	1824298	123249080
7	2016-1-28	6.57	6.66	6.33	6.2	17300066	112924544
8	2016-1-27	6.86	6.96	6.69	6.18	26494432	172994000
9	2016-1-26	7.49	7.51	6.83	6.83	23455000	167900528
10	2016-1-25	7.65	7.72	7.59	7.51	14960245	114252776

图3-6　我国股票数据文件

此外，需注意的是，在案例代码中：

```
#import pandas.io.data as web
import pandas_datareader.data as web
```

被注释符号（#）屏蔽的语句是旧版本Pandas数据分析软件采用的数据采集模块库，下面的语句是新版本的调用格式。函数定义后，主流程的前面两行是zwQuant的数据初始化和打印语句。

```
qx=zw.zwDatX(zw._rdatUS);
qx.prDat();
```

pandas-datareader数据抓取模块支持以下多种不同的数据源。

- Yahoo!Finance（雅虎财经）。
- Google Finance（谷歌财经）。
- FRED（弗雷德）。
- Fama/French（法国Fama）。
- World Bank（世界银行）。
- OECD（经合组织）。
- Eurostat（欧盟）。
- EDGAR Index（埃德加，SEC美国证券交易委员会）。

因为各种原因，对于我国用户来说，目前股票数据方面真正有用的也就是雅虎财经。

3.6 财经数据源模块库TuShare

我国财经数据的采集目前主要使用TuShare模块库。TuShare项目网址是http://tushare.org/，这是一个纯中文的TuShare文档说明网址；TuShare项目本身放在GitHub开源项目网站，网址是https://github.com/waditu/tushare。

如图3-7所示是TuShare项目网站。



图3-7 TuShare项目网站

TuShare是一个免费、开源的Python财经数据接口包，主要实现对股票等金融数据的数据采集、清洗加工和数据存储，能够为金融分析人员提供快速、整洁和多样的数据，为分析人员在数据获取方面极大地减

轻工作量，使他们能够专注于策略和模型的研究与实现。

考虑到Pandas在金融量化分析中体现出的优势，TuShare返回的绝大部分的数据格式都是Pandas DataFrame类型，便于用Pandas/NumPy/matplotlib进行数据分析和可视化。

当然，如果用户习惯了用Excel或者关系型数据库做分析，也可以通过TuShare的数据存储功能将数据全部保存到本地后进行分析。应一些用户的请求，从0.2.5版本开始，TuShare同时兼容Python 2.x和Python 3.x，对部分代码进行了重构并优化了一些算法，确保数据获取的高效和稳定。

TuShare内置了通联数据开放平台数据接口，数据在丰富性和质量方面得到了质和量的全面提升，基本上满足了用户对全品类金融数据的需求。

通联数据是我国最大的专业金融数据供应商之一，是我国领先的优矿量化网站。目前，通联数据的部分接口是免费开放的，读者可以注册一个账号，特别适合对数据源要求较高的实盘用户。

此外，目前还有各种收费API，大部分费用在500元/月~2000元/月。本小节TuShare函数的介绍大部分源自TuShare说明文档，特此说明。

TuShare数据模块提供的数据源种类丰富，配套函数也很多，读者最好把相关文档通读一遍。对于量化初学者而言，一般常用的就是函数，除了基本面类数据外，还有交易数据。

基本面类数据提供所有股票的基本面情况，包括股本情况、业绩预

告和业绩报告等，主要包括以下类别：

- 沪深股票列表；
- 业绩预告；
- 业绩报告（主表）；
- 盈利能力数据；
- 营运能力数据；
- 成长能力数据；
- 偿债能力数据；
- 现金流量数据。

TuShare数据模块提供的基本面类数据全部来自新浪财经，常用的函数是get_stock_basics()（获取沪深股票列表）。

TuShare数据模块的交易数据提供股票的交易行情，通过简单的接口调用，可获取相应的DataFrame格式数据，主要包括以下类别：

- 历史行情数据，get_hist_data()；
- 复权历史数据，get_h_data()；
- 实时行情数据，get_today_all()；
- 历史分笔数据，get_tick_data()；
- 实时报价数据，get_realtime_quotes()；
- 当日历史分笔，get_today_ticks()；
- 大盘指数列表，get_index()；
- 大单交易数据，get_sina_dd()。

3.6.1 沪深股票列表

函数名称为get_stock_basics()。

调用参数：无。

使用get_stock_basics()函数可以获取沪深上市公司的基本情况，其属性如下：

- code（代码）。
- name（名称）。
- industry（所属行业）。
- area（地区）。
- pe（市盈率）。
- outstanding（流通股本）。
- totals（总股本（万元））。
- totalAssets（总资产（万元））。
- liquidAssets（流动资产）。
- fixedAssets（固定资产）。
- reserved（公积金）。
- reservedPerShare（每股公积金）。
- eps（每股收益）。
- bvps（每股净资）。
- pb（市净率）。
- timeToMarket（上市日期）。

调用代码如下：

```
import tushare as ts
```

```
ts.get_stock_basics()
```

运行结果如下：

```
code      name     industry    area       pe      outstanding      tot

600606   金丰投资     房产服务      上海       0.00      51832.01          51
002285   世联行       房产服务      深圳       71.04     76352.17          76
000861   海印股份     房产服务      广东       126.20     83775.50         11
000526   银润投资     房产服务      福建       2421.16    9619.50          96
000056   深国商       房产服务      深圳       0.00      14305.55          26
600895   张江高科     园区开发      上海       171.60    154868.95         15
600736   苏州高新     园区开发      江苏       48.68     105788.15         10
600663   陆家嘴       园区开发      上海       47.63     135808.41         18
600658   电子城       园区开发      北京       19.39     58009.73          58
600648   外高桥       园区开发      上海       65.36     81022.34          11
600639   浦东金桥     园区开发      上海       57.28     65664.88          28
600604   市北高新     园区开发      上海       692.87    33352.42          56
```

3.6.2 案例3-3：下载股票代码数据

案例3-3介绍通过TuShare数据模块下载我国A股的股票代码数据。

案例脚本文件名为\zwpython\zw_k10\k302_down_cnBase.py。

相关代码如下：

```
# -*- coding: utf-8 -*-

import sys,os
import tushare as ts

#----------
#ts.get_stock_basics()

def zw_stk_down_base():
     rss="tmp\\"
     fss=rss+'stk_base.csv';print(fss);
     dat = ts.get_stock_basics();
     dat.to_csv(fss,encoding='gbk');

     c20=['code','name','industry','area'];
     d20=dat.loc[:,c20]
     d20['code']=d20.index;

     fss=rss+'stk_code.csv';print(fss);
     d20.to_csv(fss,index=False,encoding='gbk');

     #上证50成份股，上证规模大、流动性好的50只股票，优质大盘企业
```

```
    #上证50，指数代码000016
    fss=rss+'stk_sz50.csv';print(fss);
    dat=ts.get_sz50s();
    dat.to_csv(fss,index=False,encoding='gbk');

    #沪深300
    fss=rss+'stk_hs300.csv';print(fss);
    dat=ts.get_hs300s();
    dat.to_csv(fss,index=False,encoding='gbk');

    #中证500(CSI 500)，上海代码000905，深圳代码399905
    fss=rss+'stk_zz500.csv';print(fss);
    dat=ts.get_zz500s();
    dat.to_csv(fss,index=False,encoding='gbk')

#----------
zw_stk_down_base()
```

运行结果如下：

```
runfile('F:/zwPython/zw_k10/k302_down_cnBase.py', wdir='F:/zwPyt
tmp\stk_base.csv
tmp\stk_code.csv
tmp\stk_sz50.csv
tmp\stk_hs300.csv
```

```
tmp\stk_zz500.csv
```

本案例程序是直接从zwDat数据包下载的脚本复制过来的，除了输出文件路径，未做其他修改。

在结果数据中，除了stk_base.csv（中国A股两千多只股票代码公司概况），还额外提供其他几个数据文件：

- stk_code.csv（中国A股2810只股票代码）。
- stk_hs300.csv（中国沪深300指数股票代码）。
- stk_sz50.csv（中国上证50指数股票代码）。
- stk_zz500.csv（中国中证500指数股票代码）。

stk_base.csv相关下载脚本如下：

```
dat = ts.get_stock_basics();
```

如图3-8所示是股票代码公司的数据。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	code	name	industry	area	pe	outstandi	totals	totalAssets	liquidAss	fixedAsse	reserved	reservedF	esp	bvps	pb	timeToMarket
2	1	平安银行	银行	深圳	6.1	1180406	1430868	259906000	0	440600	5932600	4.15	1.27	10.98	0.92	19910403
3	2	万　科A	全国地产	深圳	29.51	970832.8	1103913	57079348	52000132	404016.4	834236.9	0.76	0.621	8.05	3.04	19910129
4	4	国农科技	生物制药	深圳	846.18	8387.54	8397.66	35863.95	28597.28	3792.01	66.45	0.01	0.028	0.99	32.41	19910114
5	5	世纪星源	区域地产	深圳	0	91374.29	105853.7	133658.91	22092.81	2413.78	26976.98	0.25	-0.028	0.7	11.09	19901210
6	6	深振业A	区域地产	深圳	32.15	134148.7	134999.5	1238693.25	1015397	362.29	48432.15	0.36	0.205	3.15	2.79	19920427
7	7	全新好	酒店餐饮	深圳	0	20588.71	23096.53	52680.38	33271.51	7281.37	42324.12	1.83	-0.051	1.48	16.1	19920413
8	8	神州高铁	运输设备	北京	181.75	149759	266883.9	335976.16	146185	41418.96	31871.68	0.12	0.036	1.17	7.44	19920507
9	9	中国宝安	全国地产	深圳	22.76	148966.9	159210.7	1666130.63	1162050	168927.7	52228.1	0.33	0.427	2.81	4.62	19910625
10	10	深华新	建筑施工	北京	2156.58	40792.98	81985.46	190286.31	55589.3	4303.23	12939	0.16	0.003	0.93	10.39	19951027

图3-8　股票代码公司的数据

stk_code.csv股票代码文件是在stk_base.csv中提取的，相关脚本如下，请注意相关的数据代码：

```
c20=['code','name','industry','area'];
d20=dat.loc[:,c20]
d20['code']=d20.index;
```

如图3-9所示是股票代码数据，由上面的案例程序生成。

	A	B	C	D
1	code	name	industry	area
2	1	平安银行	银行	深圳
3	2	万　科A	全国地产	深圳
4	4	国农科技	生物制药	深圳
5	5	世纪星源	区域地产	深圳
6	6	深振业A	区域地产	深圳
7	7	全新好	酒店餐饮	深圳
8	8	神州高铁	运输设备	北京
9	9	中国宝安	全国地产	深圳
10	10	深华新	建筑施工	北京

图3-9　股票代码数据

沪深300指数、上证50指数、中证500指数的数据文件的相关下载脚本是：

```
dat=ts.get_sz50s();
dat=ts.get_hs300s();
dat=ts.get_zz500s();
```

3.6.3　CSV文件处理

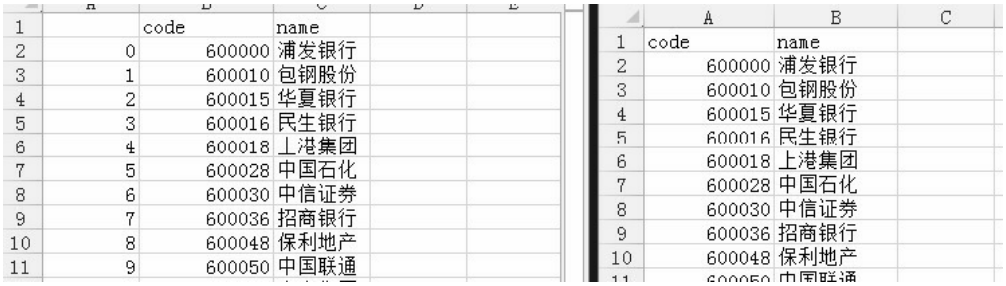
在案例脚本中，CSV数据文件的保存语句是：

```
dat.to_csv(fss,index=False,encoding='gbk')
```

语句中使用了以下两个小技巧。

- “index=False”：表示不保存索引列，否则数据文件会多一列排序数字。
- “encoding='gbk'”：因为在下载的A股数据中，公司名称有中文，而且是国标码，所以使用GBK编码保存。一般情况下使用utf-8编码保存。

如图3-10所示为不使用和使用“index=False”的差别。



		code	name
1			
2	0	600000	浦发银行
3	1	600010	包钢股份
4	2	600015	华夏银行
5	3	600016	民生银行
6	4	600018	上港集团
7	5	600028	中国石化
8	6	600030	中信证券
9	7	600036	招商银行
10	8	600048	保利地产
11	9	600050	中国联通

	A	B	C
1	code	name	
2	600000	浦发银行	
3	600010	包钢股份	
4	600015	华夏银行	
5	600016	民生银行	
6	600018	上港集团	
7	600028	中国石化	
8	600030	中信证券	
9	600036	招商银行	
10	600048	保利地产	

图3-10　不使用和使用“index=False”的差别

图3-10中的左图是没有使用“index=False”语句的数据文件效果，默认是不使用；右图是使用“index=False”语句的数据文件效果。Pandas数据软件最常用的数据文件接口就是CSV，其命令是：

- read_csv，读取csv数据文件；
- to_csv，保存csv数据文件。

read_csv与to_csv函数是Pandas数据软件中最常用的数据文件处理函数，其功能强大，参数很多，读者应用时多注意相关细节。

read_csv常用参数如下：

- filepath_or_buffer：默认的文件路径。
- dtype：用dict来表示读取某一行的文件类型，对于“0”开头的数字一定要用str来显示。
- encoding：写到CSV文件中的内容编码，本案例因为有公司名称，所以使用GBK编码。
- index_col：指定用某列作为index，第一列用0、第二列用1等。
- skiprows：指定先跳过多少行再读取，适用于大文件的分段读取。
- names：用list给读进来的列命名。
- nrows：指定读多少行，适用于大文件的分段读取。

to_csv常用参数如下：

- path_or_buf：默认的文件路径。
- encoding：写到CSV文件中的内容编码，这里因为有公司名称，所以使用GBK编码。
- header：是否在CSV文件里面写明每一行的抬头。
- index：bool值表示是否写入DataFrame里面的index，默认写入。
- chunksize：表示一口气写入多少行。
- mode：表示写入模式，默认是“W”。
- cols：可以用list来表明写入哪些行。
- date_format：用Format string来表示写入的时间类型。

3.7 历史数据

3.7.1 历史行情

历史行情数据抓取函数的名称为get_hist_data()。

获取个股历史交易数据（包括均线数据），可以通过参数设置获取日K线、周K线、月K线，以及5分钟、15分钟、30分钟和60分钟K线数据。该函数只能获取近3年的日线数据，适合搭配均线数据进行选股和分析，如果需要全部历史数据，则需调用函数get_h_data()。

（1）参数说明。

- code：股票代码（6位数字代码）或者指数代码（sh=上证指数，sz=深圳成指，hs300=沪深300指数，sz50=上证50，zxb=中小板，cyb=创业板）。
- start：开始日期，格式为YYYY-MM-DD。
- end：结束日期，格式为YYYY-MM-DD。
- ktype：数据类型，D=日K线，W=周K线，M=月K线，5=5分钟，15=15分钟，30=30分钟，60=60分钟，默认为D。
- retry_count：网络异常后的重试次数，默认为3。
- pause：重试时的停顿秒数，默认为0。

（2）返回值说明。

- date：日期。

- open：开盘价。
- high：最高价。
- close：收盘价。
- low：最低价。
- volume：成交量。
- price_change：价格变动。
- p_change：涨跌幅。
- ma5：5日均价。
- ma10：10日均价。
- ma20：20日均价。
- v_ma5：5日均量。
- v_ma10：10日均量。
- v_ma20：20日均量。
- turnover：换手率（指数无此项）。

调用方法如下：

```
import tushare as ts

ts.get_hist_data('600848') #一次性获取全部日K线数据
```

结果显示如下：

```
            open    high    close    low     volume    p_change
date
```

```
2012-01-11   6.880   7.380   7.060   6.880   14129.96     2.62
2012-01-12   7.050   7.100   6.980   6.900    7895.19    -1.13
2012-01-13   6.950   7.000   6.700   6.690    6611.87    -4.01
2012-01-16   6.680   6.750   6.510   6.480    2941.63    -2.84
2012-01-17   6.660   6.880   6.860   6.460    8642.57     5.38
2012-01-18   7.000   7.300   6.890   6.880   13075.40     0.44
2012-01-19   6.690   6.950   6.890   6.680    6117.32     0.00
2012-01-20   6.870   7.080   7.010   6.870    6813.09     1.74

              ma10      ma20      v_ma5      v_ma10     v_ma20
date
2012-01-11   7.060     7.060     14129.96   14129.96   14129.96
2012-01-12   7.020     7.020     11012.58   11012.58   11012.58
2012-01-13   6.913     6.913     9545.67    9545.67    9545.67
2012-01-16   6.813     6.813     7894.66    7894.66    7894.66
2012-01-17   6.822     6.822     8044.24    8044.24    8044.24
2012-01-18   6.833     6.833     7833.33    8882.77    8882.77
2012-01-19   6.841     6.841     7477.76    8487.71    8487.71
2012-01-20   6.863     6.863     7518.00    8278.38    8278.38
```

下面的代码设定历史数据的时间：

```
ts.get_hist_data('600848',start='2015-01-05',end='2015-01-09')
```

数据结果显示如下：

```
            open    high    close    low   volume    p_change
date
2015-01-05  11.160  11.390  11.260  10.890  46383.57  1.26  11.1
2015-01-06  11.130  11.660  11.610  11.030  59199.93  3.11  11.1
2015-01-07  11.580  11.990  11.920  11.480  86681.38  2.67  11.3
2015-01-08  11.700  11.920  11.670  11.640  56845.71  -2.10  11.
2015-01-09  11.680  11.710  11.230  11.190  44851.56  -3.77  11.
            ma20    v_ma5     v_ma10     v_ma20      turnover
date
2015-01-05  11.198  58648.75  68429.87   97141.81     1.59
2015-01-06  11.382  54854.38  63401.05   98686.98     2.03
2015-01-07  11.543  55049.74  61628.07  103010.58     2.97
2015-01-08  11.647  57268.99  61376.00  105823.50     1.95
2015-01-09  11.682  58792.43  60665.93  107924.27     1.54
```

其他获取数据的方法如下：

```
ts.get_hist_data('600848', ktype='W') #获取周K线数据
ts.get_hist_data('600848', ktype='M') #获取月K线数据
ts.get_hist_data('600848', ktype='5') #获取5分钟K线数据
ts.get_hist_data('600848', ktype='15') #获取15分钟K线数据
ts.get_hist_data('600848', ktype='30') #获取30分钟K线数据
```

```
ts.get_hist_data('600848', ktype='60') #获取60分钟K线数据
ts.get_hist_data('sh'）#获取上证指数K线数据，其他参数与个股一致，下同
ts.get_hist_data('sz'）#获取深圳成指K线数据
ts.get_hist_data('hs300'）#获取沪深300指数K线数据
ts.get_hist_data('sz50'）#获取上证50指数K线数据
ts.get_hist_data('zxb'）#获取中小板指数K线数据
ts.get_hist_data('cyb')#获取创业板指数K线数据
```

注意，虽然TuShare文档中说明使用get_hist_data函数可以抓取分时数据，但实际测试表明，只能抓取前350多项数据，不管是5分钟K线数据还是60分钟K线数据。

3.7.2 案例3-4：下载近期股票数据

案例3-4的脚本文件名为k303_down_hist_data.py。

相关代码如下：

```
# -*- coding: utf-8 -*-
import os
import tushare as ts
import pandas as pd
import matplotlib as mpl
```

```
#zw.xxx
import zwSys as zw  #::zwQT

#----------

def zw_down_stk_cn010(qx,xtyp="D"):
     xcod=qx.stkCode;xd=[];tim0='2012-01-01';
     rss="tmp\\";fss=rss+xcod+"_"+xtyp+'.csv';
     #-------------------
     xfg=os.path.exists(fss);xd0=[];
     if xfg:
         xd0=pd.read_csv(fss,index_col=0,parse_dates=[0],encodin
         xd0=xd0.sort_index(ascending=False);
         _xt=xd0.index[0];
         s2=str(_xt);tim0=s2.split(" ")[0]

     print('\n',xfg,fss);
     #-----------
     try:
         xd=ts.get_hist_data(xcod,start=tim0,end=None,retry_coun
         #-------------
         if xd is not None:
             if (len(xd0)>0):
                 xd2 =xd0.append(xd)
                 #  flt.dup
                 xd2["index"]=xd2.index
                 xd2.drop_duplicates(subset='index', keep='last'
```

```
                del(xd2["index"]);
                #xd2.index=pd.to_datetime(xd2.index)
                xd=xd2;

            xd.to_csv(fss,encoding='gbk')
    except IOError:
        pass    #skip,error

    return xd

  #----------
mpl.style.use('seaborn-whitegrid');
qx=zw.zwDatX(zw._rdatCN);
qx.stkCode="002739";#万达院线
#qx.prDat();

#
df=zw_down_stk_cn010(qx,"D")
#df=zw_down_stk_cn010(qx,"5")
df.index=pd.to_datetime(df.index)
df=df.sort_index(ascending=False);

df['close'].plot(figsize=(15,5),rot=20)
print(df.tail())
```

案例3-4的核心与难点是函数zw_down_stk_cn010，就是根据指定的

股票代码，参数，下载股票的交易数据，运行逻辑，请参看图3-11函数流程图：

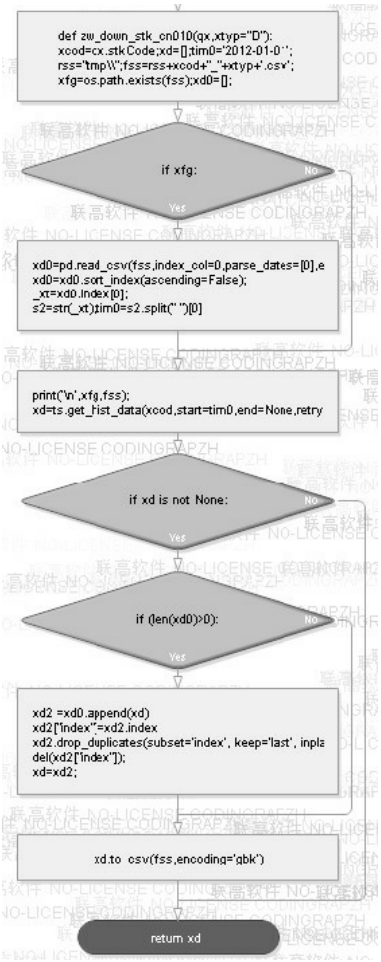


图3-11　函数流程图

运行结果如下：

```
runfile('F:/zwPython/zw_k10/k303_down_hist_data.py', wdir='F:/zw

  False tmp\002739_D.csv
            open   high  close   low    volume     price_change
date
```

```
2015-01-28  45.00  45.00  45.00  45.00  287.00          4.09
2015-01-27  40.91  40.91  40.91  40.91  124.00          3.72
2015-01-26  37.19  37.19  37.19  37.19  105.00          3.38
2015-01-23  33.81  33.81  33.81  33.81  113.29          3.07
2015-01-22  27.94  30.74  30.74  27.94  465.00          9.39

              ma5     ma10    ma20    v_ma5   v_ma10  v_ma20  turn
date
2015-01-28  37.530  37.530  37.530  218.86  218.86  218.86
2015-01-27  35.663  35.663  35.663  201.82  201.82  201.82
2015-01-26  33.913  33.913  33.913  227.76  227.76  227.76
2015-01-23  32.275  32.275  32.275  289.15  289.15  289.15
2015-01-22  30.740  30.740  30.740  465.00  465.00  465.00
```

如图3-12所示是股票数据走势图，即案例3-4的运行图。

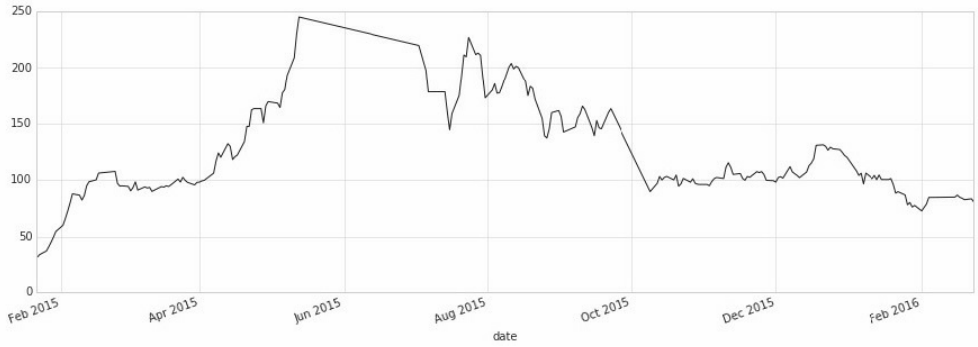


图3-12　股票数据走势图

运行后，首先输出文件名，检测数据文件是否存在。

```
False tmp\002739_D.csv
```

注意，在案例中，文件名的构成是股票代码和抓取频率。抓取频率为ktype，包括：

```
D=日K线，W=周，M=月，5=5分钟，15=15分钟，30=30分钟，60=60分钟，默认为D。
```

文件名等数据的输出由以下脚本执行：

```
print('\n',xfg,fss);
```

其中，fss是文件名；xfg用于判断数据文件是否存在，如果文件存在，就读入数据文件，并且提取最新的日期作为起始日期即tim0；由于get_hist_data只能抓取近3年的数据，所以默认为“tim0='2012-01-01'”。

```
if xfg:
     xd0=pd.read_csv(fss,index_col=0,parse_dates=[0],encoding='g
     xd0=xd0.sort_index(ascending=False);
     _xt=xd0.index[0];
     s2=str(_xt);tim0=s2.split(" ")[0]
```

案例3-4中的核心语句是以下数据抓取语句：

```
xd=ts.get_hist_data(xcod,start=tim0,end=None,retry_count=5,pause
```

注意，结束时间的设置为“end=None”，它表示抓取到最新一个交易日的数据。如果数据抓取成功，就运行以下脚本：

```
if xd is not None:
      if (len(xd0)>0):
          xd2 =xd0.append(xd)
          #  flt.dup
          xd2["index"]=xd2.index
          xd2.drop_duplicates(subset='index', keep='last', inpla
          del(xd2["index"]);
          #xd2.index=pd.to_datetime(xd2.index)
          xd=xd2;

xd.to_csv(fss,encoding='gbk')
```

上面这段脚本主要是合并新、旧数据，并去掉重复数据，然后保存数据文件。

最后介绍一下主控程序，相关脚本和注释如下：

```
mpl.style.use('seaborn-whitegrid');                    # 设置绘图环境
qx=zw.zwDatX(zw._rdatCN);                              # 设置数据目录等初始
qx.code="002739";                                      # 万达院线   # 设置

df=zw_down_stk_cn010(qx,"D")                           # 抓取日线数据
#df=zw_down_stk_cn010(qx,"5")
df.index=pd.to_datetime(df.index)                      # 反向排序数据
df=df.sort_index(ascending=False);

df['close'].plot(figsize=(15,5),rot=20)                # 绘制收盘价曲线图
print(df.tail())                                       # 输出尾部数据
```

使用get_hist_data()函数可以抓取分时数据，我们抓取5分钟数据测试一下，抓取代码如下：

```
df=zw_down_stk_cn010(qx,"5")
```

如图3-13所示为股票分时数据示意图，是以上脚本的运行结果。

```
True tmp\002739_5.csv
                     open   high  close    low  volume  price_change  \
date
2016-02-05 13:35:00  85.45  85.45  85.40  85.30     382         -0.05
2016-02-05 13:30:00  85.39  85.48  85.45  85.30    1592          0.06
2016-02-05 13:25:00  85.00  85.39  85.39  84.91    1296          0.39
2016-02-05 13:20:00  85.07  85.15  85.00  85.00    1093         -0.07
2016-02-05 13:15:00  85.30  85.30  85.11  85.08     993         -0.19

                     p_change     ma5    ma10     ma20    v_ma5   v_ma10  \
date
2016-02-05 13:35:00     -0.06  85.270  85.347  85.4580  1071.20   970.07
2016-02-05 13:30:00      0.07  85.242  85.356  85.5130  1122.60  1018.12
2016-02-05 13:25:00      0.46  85.252  85.364  85.5665   932.59   947.68
2016-02-05 13:20:00     -0.08  85.254  85.335  85.6255   850.59   896.78
2016-02-05 13:15:00     -0.22  85.350  85.390  85.6945   806.75   945.18

                      v_ma20  turnover
date
2016-02-05 13:35:00  1316.62      0.01
2016-02-05 13:30:00  1449.11      0.04
2016-02-05 13:25:00  1490.28      0.03
2016-02-05 13:20:00  1577.70      0.03
2016-02-05 13:15:00  1749.50      0.02
```

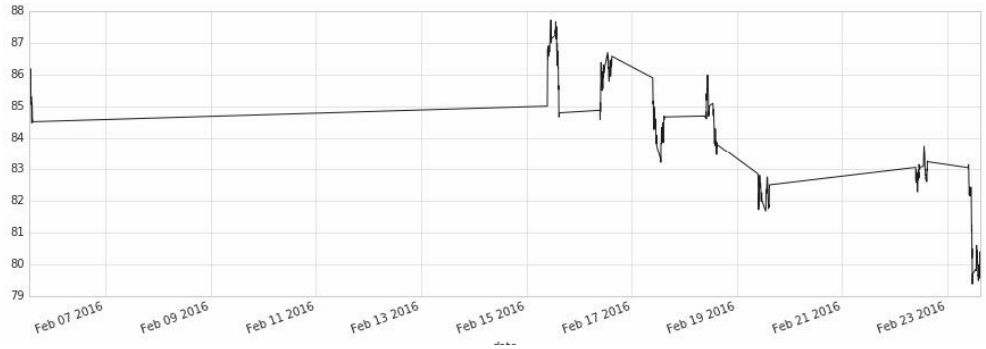


图3-13　股票分时数据示意图

3.7.3　历史复权数据

历史复权数据抓取函数的名称为get_h_data()。

历史复权数据分为前复权数据和后复权数据。接口提供股票上市以来所有的历史数据，默认为前复权。如果不设定开始日期和结束日期，则返回近一年的复权数据，从性能上考虑，推荐设定开始日期和结束日期，而且最好不要获取超过三年以上的全部历史数据，可分时段分步获取，获取到数据后及时在本地存储。

注意

虽然文档中说抓取的数据时间不要超过三年，但实际上可以抓取1994年我国股市开盘以来的所有数据。

get_h_data()函数的参数说明如下：

- code：string型，股票代码，例如600848。
- start：string型，开始日期，格式为YYYY-MM-DD，为空时取当前日期。
- end：string型，结束日期，格式为YYYY-MM-DD，为空时取去年今日。
- autype：string型，复权类型，qfq为前复权，hfq为后复权，none为不复权，默认为qfq。
- index：Boolean型，是否是大盘指数，默认为False。
- retry_count：int型，默认为3，指遇网络等问题重复执行的次数。
- pause：int型，默认为0，重复请求数据过程中暂停的秒数，防止请求间隔时间太短出现问题。

get_h_data()函数的返回值说明如下：

- date：交易日期（index）。
- open：开盘价。
- high：最高价。
- close：收盘价。
- low：最低价。
- volume：成交量。
- Amount：成交金额。

获取个股首个上市日期，可参考以下方法：

```
df = ts.get_stock_basics()
date = df.ix['600848']['timeToMarket'] #上市日期YYYYMMDD
```

本接口还提供大盘指数的全部历史数据，调用时务必设定index参数为True，由于大盘指数不存在复权的问题，故可以忽略autype参数。代码如下：

```
ts.get_h_data('002337') #前复权
ts.get_h_data('002337', autype='hfq') #后复权
ts.get_h_data('002337', autype=None) #不复权
ts.get_h_data('002337', start='2015-01-01', end='2015-03-16') #两

ts.get_h_data('399106', index=True) #深圳综合指数
```

运行结果如下：

```
date        open   high   close  low     volume      amount

2015-03-16  13.27  13.45  13.39  13.00   81212976  1073862784
2015-03-13  13.04  13.38  13.37  13.00   40548836   532739744
2015-03-12  13.29  13.95  13.28  12.96   71505720   962979904
```

```
2015-03-11  13.35  13.48  13.15  13.00   59110248   780300736
2015-03-10  13.16  13.67  13.59  12.72  105753088  1393819776
2015-03-09  13.77  14.73  14.13  13.70  139091552  1994454656
2015-03-06  12.17  13.39  13.39  12.17   89486704  1167752960
2015-03-05  12.79  12.80  12.17  12.08   26040832   966927360
2015-03-04  13.96  13.96  13.30  12.58   26636174  1060270720
2015-03-03  12.17  13.10  13.10  12.05   19290366   733336768
```

3.7.4 案例3-5：下载历史复权数据

本案例是下载历史复权数据，脚本文件名为\zwpython\zw_k10\k304_down_h_data.py。

相关代码如下：

```
# -*- coding: utf-8 -*-
import os
import tushare as ts
import pandas as pd
import matplotlib as mpl

import zwSys as zw  #::zwQT

#----------
```

```
def zw_down_stk_cn020(qx):
     xcod=qx.stkCode;xd=[];tim0='1994-01-01';
     rss="tmp\\";fss=rss+xcod+'.csv';
     #-------------------
     xfg=os.path.exists(fss);xd0=[];
     if xfg:
         xd0=pd.read_csv(fss,index_col=0,parse_dates=[0],encodin
         xd0=xd0.sort_index(ascending=False);
         _xt=xd0.index[0];
         s2=str(_xt);tim0=s2.split(" ")[0]

     print('\n',xfg,fss);
     #-----------
     try:
         xd=ts.get_h_data(xcod,start=tim0,end=None,retry_count=5
         #-------------
         if xd is not None:
             if (len(xd0)>0):
                 xd2 =xd0.append(xd)
                 #  flt.dup
                 xd2["index"]=xd2.index
                 xd2.drop_duplicates(subset='index', keep='last'
                 del(xd2["index"]);
                 #xd2.index=pd.to_datetime(xd2.index)
                 xd=xd2;
```

```
            xd.to_csv(fss,encoding='gbk')
    except IOError:
        pass    #skip,error

    return xd

  #----------
mpl.style.use('seaborn-whitegrid');
qx=zw.zwDatX(zw._rdatCN);
qx.stkCode="002739";#万达院线
#qx.prDat();

#
df=zw_down_stk_cn020(qx)

df.index=pd.to_datetime(df.index)
df=df.sort_index(ascending=False);

df['close'].plot(figsize=(15,5),rot=20)
print(df.tail())
```

本案例程序代码与案例3-4的代码类似，只是抓取数据的语句不同，采用的是以下语句：

```
xd=ts.get_h_data(xcod,start=tim0,end=None,retry_count=5,pause=1)
```

此外，使用get_h_data函数可以抓取A股开市以来所有的股票交易数据，所以默认起始时间可设为以下代码：

```
tim0='1994-01-01';
```

如图3-14所示是股票历史复权数据运行图。

```
In [66]: runfile('F:/zwPython/zw_k10/k304_down_h_data.py', wdir='F:/zwPython/zw_k10')

 True tmp\002739.csv
[Getting data:]            amount  close   high    low    ma10    ma20    ma5   open  \
date
2015-01-28     NaN  45.00  45.00  45.00  37.530  37.530  37.530  45.00
2015-01-27     NaN  40.91  40.91  40.91  35.663  35.663  35.663  40.91
2015-01-26     NaN  37.19  37.19  37.19  33.913  33.913  33.913  37.19
2015-01-23     NaN  33.81  33.81  33.81  32.275  32.275  32.275  33.81
2015-01-22     NaN  30.74  30.74  27.94  30.740  30.740  30.740  27.94

            p_change  price_change  turnover  v_ma10  v_ma20   v_ma5  volume
date
2015-01-28     10.00          4.09      0.05  218.86  218.86  218.86  287.00
2015-01-27     10.00          3.72      0.02  201.82  201.82  201.82  124.00
2015-01-26     10.00          3.38      0.02  227.76  227.76  227.76  105.00
2015-01-23      9.99          3.07      0.02  289.15  289.15  289.15  113.29
2015-01-22     43.98          9.39      0.08  465.00  465.00  465.00  465.00
F:\zwPython\py35\python-3.5.1.amd64\lib\site-packages\tushare\stock\trading.py:441: FutureWarning: sort(columns=....) is deprecated, use
sort_values(by=.....)
  df = df.sort('date', ascending=False)
```

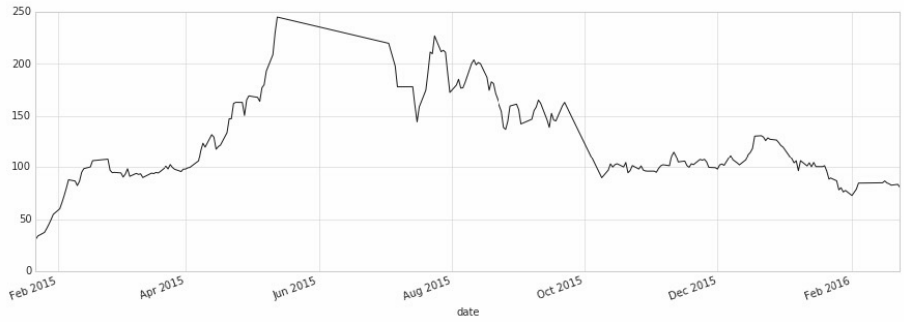


图3-14　股票历史复权数据运行图

3.8 其他交易数据

量化分析最主要的数据源是历史数据，此外还有实时交易数据等。除了历史行情、历史分时数据、历史复权数据外，常用的TuShare数据采集函数还有：

- 实时行情数据get_today_all()；
- 历史分笔数据get_tick_data()；
- 实时分笔数据get_realtime_quotes()；
- 当日历史分笔数据get_today_ticks()；
- 大盘指数行情列表get_index()；
- 大单交易数据get_sina_dd()。

1. 实时行情数据

实时行情数据抓取函数的名称是get_today_all()。

使用该函数可一次性获取当前交易的所有股票的行情数据（如果是节假日，即为上一个交易日，结果的显示速度取决于网速）。该函数无参数。

该函数返回值说明如下：

- code：代码。
- name：名称。
- changepercent：涨跌幅。
- trade：现价。

- open：开盘价。
- high：最高价。
- low：最低价。
- settlement：昨日收盘价。
- volume：成交量。
- turnoverratio：换手率。

该函数使用示例如下：

```
import tushare as ts

ts.get_today_all()
```

运行结果显示如下：

```
      code    name   changepercent  trade  open  high   low
0     002738  中矿资源       10.023  19.32  19.32  19.32  19.32
1     300410  正业科技       10.022  25.03  25.03  25.03  25.03
2     002736  国信证券       10.013  16.37  16.37  16.37  16.37
3     300412  迦南科技       10.010  31.54  31.54  31.54  31.54
4     300411  金盾股份       10.007  29.68  29.68  29.68  29.68
5     603636  南威软件       10.006  38.15  38.15  38.15  38.15
6     002664  信质电机       10.004  30.68  29.00  30.68  28.30
7     300367  东方网力       10.004  86.76  78.00  86.76  77.87
8     601299  中国北车       10.000  11.44  11.44  11.44  11.29
```

```
9     601880   大连港        10.000   5.72   5.34   5.72   5.22


         volume       turnoverratio
0        375100        1.25033
1         85800        0.57200
2       1058925        0.08824
3         69400        0.51791
4        252220        1.26110
5       1374630        5.49852
6       6448748        9.32700
7       2025030        6.88669
8     433453523        4.28056
9     323469835        9.61735
```

2. 历史分笔数据

历史分笔数据的函数名称为get_tick_data()。

使用该函数可以获取个股以往交易历史的分笔数据明细。通过分析分笔数据，可以大致判断资金的进出情况。在使用过程中，对于获取股票某个阶段的历史分笔数据，需要通过输入交易日参数并追加（append）到一个DataFrame变量，或者直接追加（append）到本地同一个文件里。

历史分笔接口只能获取当前交易日之前的数据，当日分笔历史数据应调用get_today_ticks()接口或者在当日18点后通过本接口获取。

get_tick_data()函数的参数说明如下：

- code：股票代码即6位数字代码。
- date：日期，格式为YYYY-MM-DD。
- retry_count：int型，默认为3，表示遇网络等问题重复执行的次数。
- pause：int型，默认为0，重复请求数据过程中暂停的秒数，防止请求间隔时间太短出现问题。

get_tick_data()函数的返回值说明如下：

- time：时间。
- price：成交价格。
- change：价格变动。
- volume：成交手。
- amount：成交金额（元）。
- type：买卖类型（买盘、卖盘、中性盘）。

历史分笔数据函数的调用方法如下：

```
import tushare as ts


df = ts.get_tick_data('600848',date='2014-01-09')
df.head(10)
```

运行结果显示如下：

```
     time       price change   volume  amount  type
0    15:00:00   6.05     --        8    4840     卖盘
1    14:59:55   6.05     --       50   30250     卖盘
2    14:59:35   6.05     --       20   12100     卖盘
3    14:59:30   6.05  -0.01      165   99825     卖盘
4    14:59:20   6.06   0.01        4    2424     买盘
5    14:59:05   6.05  -0.01        2    1210     卖盘
6    14:58:55   6.06     --        4    2424     买盘
7    14:58:45   6.06     --        2    1212     买盘
8    14:58:35   6.06   0.01        2    1212     买盘
9    14:58:25   6.05  -0.01       20   12100     卖盘
```

3. 实时分笔数据

实时分笔数据的函数名称为get_realtime_quotes()。

使用该函数可以实时取得股票的当前报价和成交信息，其中一种场景是，写一个Python定时程序来调用本接口（可两三秒执行一次，性能与行情软件基本一致），然后通过DataFrame矩阵进行计算，从而实现交易监控，可实时监测交易量和价格的变化。

该函数的参数为symbols，6位数字股票代码或者指数代码（sh=上证指数，sz=深圳成指，hs300=沪深300指数，sz50=上证50，zxb=中小板，cyb=创业板），可输入的类型有str、list、set和Pandas的Series对象。

该函数的返回值说明如下：

0：name，股票名字。

1：open，今日开盘价。

2：pre_close，昨日收盘价。

3：price，当前价格。

4：high，今日最高价。

5：low，今日最低价。

6：bid，竞买价，即“买一”报价。

7：ask，竞卖价，即“卖一”报价。

8：volume，成交量。

9：amount，成交金额（元）。

10：b1_v，委买一（笔数）。

11：b1_p，委买一（价格）。

12：b2_v，“买二”。

13：b2_p，“买二”。

14：b3_v，“买三”。

15：b3_p，“买三”。

16：b4_v，“买四”。

17：b4_p，“买四”。

18：b5_v，“买五”。

19：b5_p，“买五”。

20：a1_v，委卖一（笔数）。

21：a1_p，委卖一（价格）。

...

30：date，日期。

31：time，时间。

实时分笔函数的调用方法如下：

```
import tushare as ts

df = ts.get_realtime_quotes('000581') #Single stock symbol
df[['code','name','price','bid','ask','volume','amount','time']]
```

结果显示如下：

```
    code    name   price  bid    ask    volume   amount         t
0  000581  威孚高科  31.15  31.14  31.15  8183020  253494991.16  1
```

请求多个股票数据的方法（一次最好不要超过30个）如下：

```
#symbols from a list
ts.get_realtime_quotes(['600848','000980','000981'])
#from a Series
ts.get_realtime_quotes(df['code'].tail(10))  #一次获取10个股票的实时
```

获取实时指数的方法如下：

```
#上证指数
ts.get_realtime_quotes('sh')
#上证指数 深圳成指 沪深300指数 上证50 中小板 创业板
ts.get_realtime_quotes(['sh','sz','hs300','sz50','zxb','cyb'])
#或者混搭
ts.get_realtime_quotes(['sh','600848'])
```

4. 当日历史分笔数据

当日历史分笔数据的函数名称是get_today_ticks()。

该函数用于获取当前交易日（交易进行中使用）已经产生的分笔明细数据。

get_today_ticks()函数的参数说明如下：

- code：股票代码即6位数字代码。
- retry_count ： int型，默认为3，表示遇网络等问题重复执行的次数。
- pause ： int型，默认为0，重复请求数据过程中暂停的秒数，防止请求间隔时间太短出现的问题。

get_today_ticks()函数的返回值说明如下：

- time：时间。
- price：当前价格。
- pchange：涨跌幅。
- change：价格变动。
- volume：成交手。
- amount：成交金额（元）。
- type：买卖类型（买盘、卖盘、中性盘）。

当日历史分笔数据函数的调用方法如下：

```
import tushare as ts


df = ts.get_today_ticks('601333')
df.head(10)
```

结果显示如下：

```
      time       price  pchange  change  volume  amount type
0     11:30:07   5.77   -0.52    0.00     634   366372   买盘
1     11:29:57   5.77   -0.52    0.00     216   124632   买盘
2     11:29:52   5.77   -0.52    0.00     306   176562   买盘
3     11:29:42   5.77   -0.52    0.01     159    91766   买盘
4     11:29:37   5.76   -0.69    0.00     546   314496   卖盘
5     11:29:32   5.76   -0.69   -0.01     954   549504   卖盘
6     11:29:22   5.77   -0.52    0.00     374   215798   买盘
7     11:29:17   5.77   -0.52    0.00     762   439674   买盘
8     11:29:12   5.77   -0.52    0.00     164    95182   买盘
9     11:29:07   5.77   -0.52    0.00     303   174854   买盘
```

5. 大盘指数行情列表

大盘指数行情列表的函数名称是get_index()。

该函数用于获取大盘指数实时行情列表，以表格的形式展示大盘指数的实时行情。

get_index()函数的返回值说明如下：

- code：指数代码。
- name：指数名称。
- change：涨跌幅。

- open：开盘点位。
- preclose：昨日收盘点位。
- close：收盘点位。
- high：最高点位。
- low：最低点位。
- volume：成交量（手）。
- amount：成交金额（亿元）。

get_index()函数的调用方法如下：

```
import tushare as ts

df = ts.get_index()
```

结果显示如下：

```
    code       name   change   preclose    close      high
0   000001    上证指数  -1.13   4527.396   4476.215   4572.391
1   000002    A股指数   -1.13   4744.093   4690.628   4791.534
2   000003    B股指数   -2.15    403.694    395.018    405.795
3   000008    综合指数   0.79   3724.496   3753.906   3848.575
4   000009    上证380  -2.79   7689.128   7474.305   7695.329
5   000010    上证180  -1.13  10741.180  10619.610  10863.080
6   000011    基金指数  -1.02   7033.291   6961.659   7058.856
7   000012    国债指数   0.01    148.626    148.641    148.656
```

8	000016	上证50	-0.79	3308.454	3282.330	3370.025	3
9	000017	新综指	-1.13	3826.013	3782.936	3864.307	3
10	000300	沪深300	-1.37	4807.592	4741.861	4839.078	4
11	399001	深证成指	-0.69	14809.424	14707.245	14979.810	14
12	399002	深成指R	-0.69	17193.832	17075.202	17391.652	1
13	399003	成份B指	-1.93	9027.079	8853.081	9013.194	
14	399004	深证100R	-1.79	5994.881	5887.414	6036.322	
15	399005	中小板指	-3.34	8935.338	8637.195	8953.813	8
16	399006	创业板指	-2.17	2747.497	2687.974	2779.200	2
17	399100	新 指 数	-2.77	10091.194	9811.256	10111.664	
18	399101	中小板综	-3.31	12792.057	12368.868	12800.453	12
19	399106	深证综指	-2.76	2271.275	2208.561	2275.344	2
20	399107	深证A指	-2.77	2375.176	2309.466	2379.507	2
21	399108	深证B指	-1.77	1398.244	1373.512	1397.996	1
22	399333	中小板R	-3.34	9640.766	9319.085	9660.699	9
23	399606	创业板R	-2.16	2828.251	2767.127	2861.040	2

	volume	amount
0	767676416	10611.72
1	766188823	10599.65
2	1487592	12.07
3	263748855	3440.01
4	182628996	2531.04
5	464275133	6437.40
6	66280981	428.46
7	263420	2.74

```
8     266042859   3735.74
9     766077611  10596.65
10    608638545   8603.50
11  51106975785   6405.28
12   6357969430   1017.68
13     51206484      4.32
14  10418315890   1779.58
15   3071396395    830.54
16   1441659735    551.73
17  32943457787   6091.34
18  10450911278   2291.43
19  33395285515   6137.71
20  33274363870   6128.94
21    120921645      8.77
22   3071396395    830.54
23   1441659735    551.73
```

6. 大单交易数据

大单交易数据的函数名称是get_sina_dd()。

该函数用于获取大单交易数据，默认为大于等于400手，数据来源于新浪财经。

get_sina_dd()函数的参数说明如下：

- code：股票代码即6位数字代码。

- date：日期，格式为YYYY-MM-DD。
- vol：手数，默认为400手，要输入数值型参数。
- retry_count：int型，默认为3，表示遇网络等问题重复执行的次数。
- pause：int型，默认为0，表示重复请求数据过程中暂停的秒数，防止请求间隔时间太短出现问题。

get_sina_dd()函数的返回值说明如下：

- code：代码。
- name：名称。
- time：时间。
- price：当前价格。
- volume：成交手。
- preprice：上一笔价格。
- type：买卖类型（买盘、卖盘、中性盘）。

get_sina_dd()函数的调用方法如下：

```
import tushare as ts


df = ts.get_sina_dd('600848', date='2015-12-24') #默认400手
#df = ts.get_sina_dd('600848', date='2015-12-24', vol=500)  #指定
```

结果显示如下：

```
```

	code	name	time	price	volume	preprice	type
0	600848	上海临港	14:58:10	23.05	104309	23.05	卖盘
1	600848	上海临港	14:57:03	23.05	56500	23.07	卖盘
2	600848	上海临港	14:52:47	23.00	76750	23.04	卖盘
3	600848	上海临港	14:47:32	23.10	47000	23.09	买盘
4	600848	上海临港	14:16:03	23.00	60859	23.01	卖盘
5	600848	上海临港	14:15:38	23.01	68659	23.03	卖盘
6	600848	上海临港	14:00:34	23.10	66200	23.10	买盘
7	600848	上海临港	13:25:24	23.28	42000	23.09	买盘
8	600848	上海临港	13:23:54	23.28	79600	23.07	买盘
9	600848	上海临港	13:16:16	23.03	40000	23.08	卖盘

3.9 zwDat超大股票数据源与数据更新

数据源如此重要，所以在发布zwPython 2016的同时，又发布了zwDat，它是一个6GB的超大股票数据源，供读者分析。

zwDat股票数据包包括中、美两国不同的股票日交易数据。加入美股数据，是考虑到目前GitHub开源项目网站有大部分量化项目，采用的都是雅虎财经的美股数据。

- 美股数据，收集了近7000只美股和日线数据，从1960—2016年1月。
- A股数据，收集了近2700只中国股票和日线数据，从1994—2016年2月7日（春节休市）。

以上两个数据源虽然只下载到2016年初，但已经内置了自动更新功能，读者运行相关的脚本就可以自动更新、合并相关的交易数据到最新的交易日。

zwDat是内置的数据下载软件，作为极宽量化工具箱zwQuant示例程序的一部分，已经收录在zwPython 2016中，其目录是x:\zwPython\zwQuant\demo\down_stk\。

目录内有多个脚本源码文件，其中以下脚本就是用于数据下载更新的：

- zw_down_cnSTK.py，A股交易数据下载下载。

- zw_down_cnSTK_Base.py，A股基本概况数据下载。
- zw_down_cnSTK_inx.py，A股指数行情数据下载。
- zw_down_usSTK.py，美股交易数据下载。

下面我们逐一进行讲解。

3.9.1 案例3-6：A股基本概况数据下载

该案例的脚本文件名是zwQuant\demo\zw_down_cnSTK_Base.py。

相关代码如下：

```
#*- coding: utf-8 -*-


import sys,os
import tushare as ts


#----------
#ts.get_stock_basics()


def zw_stk_down_base():
     rss="..\\tmp\\"
     fss=rss+'stk_base.csv';print(fss);
     dat = ts.get_stock_basics();
```

```
    dat.to_csv(fss,encoding='gbk');

    c20=['code','name','industry','area'];
    d20=dat.loc[:,c20]
    d20['code']=d20.index;

    fss=rss+'stk_code.csv';print(fss);
    d20.to_csv(fss,index=False,encoding='gbk');

    #上证50成份股,上证规模大、流动性好的50只股票,优质大盘企业
    #上证50，指数代码000016
    fss=rss+'stk_sz50.csv';print(fss);
    dat=ts.get_sz50s();
    dat.to_csv(fss,index=False,encoding='gbk');

    #沪深300
    fss=rss+'stk_hs300.csv';print(fss);
    dat=ts.get_hs300s();
    dat.to_csv(fss,index=False,encoding='gbk');

    #中证500(CSI 500）,上海代码000905，深圳代码399905。
    fss=rss+'stk_zz500.csv';print(fss);
    dat=ts.get_zz500s();
    dat.to_csv(fss,index=False,encoding='gbk')

#----------
zw_stk_down_base()
```

程序很简单，就是一个zw_stk_down_base函数和一行调用代码。运行该函数的代码如下：

```
zw_stk_down_base()
```

无需任何参数，运行结果如下：

```
runfile('F:/zwPython/zwQuant/demo/down_stk/zw_down_cnSTK_Base.py
..\tmp\stk_base.csv
..\tmp\stk_code.csv
..\tmp\stk_sz50.csv
..\tmp\stk_hs300.csv
..\tmp\stk_zz500.csv
```

需注意的是，为防止覆盖原参数文件，输出的数据文件应保存在以下目录中：

```
zwQuant\demo\tmp
```

用户需要人工复制更新时，原数据文件就在zwDat\inx目录中。

在结果数据中，除了stk_base.csv（中国A股2810只股票代码公司概况），还额外提供其他几个数据文件：

- stk_code.csv，中国A股2810只股票代码。
- stk_hs300.csv，中国沪深300指数股票代码。
- stk_sz50.csv，中国上证50指数股票代码。
- stk_zz500.csv，中国中证500指数股票代码。

3.9.2 案例3-7：A股交易数据下载

本案例脚本文件名为zwQuant\demo\ zw_down_cnSTK.py。

相关代码如下：

```
# -*- coding: utf-8 -*-

import sys,os
import tushare as ts
import pandas as pd
from pandas.tseries.offsets import Day
import dateutil.parser as parse
import datetime as dt

#zwQuant
import zwSys as zw
```

```
import zwQTBox as zwBox

#----------

def zw_stk_down_all(qx,xtyp):
     fss=qx.rdatInx+'stk_code.csv';print(fss);
     dinx = pd.read_csv(fss,encoding='gbk')

     i=0;xn9=len(dinx['code']);
     for xc in dinx['code']:
         i+=1;
         code="%06d" %xc
         #code=zwTools.v2sk(xc,6);
         print("\n",i,"/",xn9,"code,",code)
         #---
         qx.code=code;
         zwBox.down_stk_cn010(qx,xtyp);

#----------

qdat=zw.zwDatX(zw._rdatCN);
#qdat.prDat();

#zw_stk_down_all(qdat,'T')
zw_stk_down_all(qdat,'0')
```

程序很简单，就是下面的一个函数：

```
def zw_stk_down_all(qx,xtyp):
```

核心脚本是以下语句：

```
zwBox.down_stk_cn010(qx,xtyp);
```

这段代码调用了极宽量化工具箱中的下载函数.down_stk_cn010

注意import模块导入部分的语句：

```
#zwQuant
import zwSys as zw
import zwQTBox as zwBox
```

这是zwQuant相关模块的约定缩写方式。

主控程序如下：

```
#----------
```

```
qdat=zw.zwDatX(zw._rdatCN);
#qdat.prDat();


#zw_stk_down_all(qdat,'T')
zw_stk_down_all(qdat,'0')
```

首先调用zw.zwDatX类定义，设置相关的初始化参数。

正常的下载、更新是运行以下语句：

```
zw_stk_down_all(qdat,'0')
```

因为是测试代码，为避免改变zwDat的原数据，我们把最后两行改一下：

```
zw_stk_down_all(qdat,'T')
#zw_stk_down_all(qdat,'0')
```

测试运行时，执行的是以下语句：

```
zw_stk_down_all(qdat,'T')
```

函数命令都是一样的，只是参数不同：

- 参数为0，表示从股票最早上市日期开始下载数据；
- 参数为T，表示是测试模式，结果数据保存在临时目录\zwPython\zwQuant\demo\tmp中，测试模式是下载全部股票数据。

参数“0”和“T”调用的都是TuShare中的函数get_hist_data()。

也可以为参数设置抓取频率ktype，包括：

```
D=日K线，W=周，M=月，5=5分钟，15=15分钟，30=30分钟，60=60分钟
```

还可以抓取分时数据等，调用的是TuShare中的函数get_h_data()。因为分时数据只能抓取前350项，所以对分析、实盘意义都不大。

案例代码运行结果如下：

```
 1 / 2813 code, 000001


 False \zwPython\zwQuant\demo\tmp\000001.csv , 2012-01-01
[Getting data:]################
```

用测试模式抓取2~3个演示数据后，应人工中断程序。不然会抓取所有A股2800只股票的历年数据大约需要10~12个小时，因网速、硬件

不同，会有所差异。

如果希望看到数据更新追加的效果，可以先从目录F:\zwDat\cn\Day中复制几个数据文件到临时目录\zwPython\zwQuant\demo\tmp中，一般是前几个数据文件：

```
000001.csv
000002.csv
```

为方便对比，把这几个数据文件再复制一份，文件名加一个后缀sr：

```
000001sr.csv
000002sr.csv
```

再以测试模式运行演示数据：

```
zw_stk_down_all(qdat,'T')
```

运行完两个测试数据包的更新，人工中断程序。

如图3-15所示是更新前的数据文件备份000001sr.csv。

	A	B	C	D	E	F	G
1	date	open	high	close	low	volume	amount
2	2016-2-4	9.89	10	9.95	9.88	37309948	370586176
3	2016-2-3	9.85	9.89	9.85	9.77	27457216	269997824
4	2016-2-2	9.8	10.03	9.95	9.78	36910416	367360512
5	2016-2-1	9.98	10.01	9.8	9.74	41773216	412635648
6	2016-1-29	9.74	10.08	10	9.69	54443576	540544448

图3-15　更新前的数据文件备份

如图3-16所示是更新后的数据文件，自动追加4日以后的交易数据。

	A	B	C	D	E	F	G
1	date	open	high	close	low	volume	amount
2	2016-2-25	10.12	10.13	9.67	9.6	62207284	615004736
3	2016-2-24	10.13	10.22	10.22	10.08	30010360	302498016
4	2016-2-23	10.36	10.36	10.19	10.12	42587436	432315296
5	2016-2-22	10.2	10.38	10.36	10.13	61773944	630251520
6	2016-2-19	10.16	10.21	10.11	10.06	31889824	320939648
7	2016-2-18	10.25	10.29	10.16	10.16	40617824	412337568
8	2016-2-17	10.1	10.29	10.22	10.06	58516704	590538944
9	2016-2-16	9.91	10.1	10.08	9.9	42838640	427507776
10	2016-2-15	9.73	9.92	9.86	9.72	27849946	271173376
11	2016-2-5	10.03	10.04	9.99	9.98	27089334	269184384
12	2016-2-4	9.89	10	9.95	9.88	37309948	370586176
13	2016-2-3	9.85	9.89	9.85	9.77	27457216	269997824
14	2016-2-2	9.8	10.03	9.95	9.78	36910416	367360512
15	2016-2-1	9.98	10.01	9.8	9.74	41773216	412635648

图3-16　更新后的数据文件

追加是从最后更新的数据开始追加，如果每日15点交易截止后更新数据，那么只更新当天的数据，一般10~30分钟即可完成。

3.9.3　案例3-8：A股指数行情数据下载

脚本文件名为zwQuant\demo\ zw_down_cnSTK_inx.py。

相关代码如下：

```
# -*- coding: utf-8 -*-

import pandas as pd
from pandas.tseries.offsets import Day

#  zwQuant
import zwSys as zw
import zwQTBox as zwBox

#----------

def zw_stk_down_inx(qx):
     fss=qx.rdatInx+'inx_code.csv';print(fss);
     dinx = pd.read_csv(fss,encoding='gbk')

     i=0;xn9=len(dinx['code']);
     for xss in dinx['code']:
         i+=1;code=xss;#code=zwTools.v2sk(xc,6);
         print("\n",i,"/",xn9,"code,",code)
         #---
         qdat.code=code;
         zwBox.down_stk_cn010(qx,'X');

#----------
qdat=zw.zwDatX(zw._rdatCN);
#qdat.prDat();
```

```
zw_stk_down_inx(qdat);
```

程序很简单，就是以下函数：

```
def zw_stk_down_inx(qx):
```

核心脚本是以下语句：

```
zwBox.down_stk_cn010(qx,'X');
```

调用了zw量化工具箱中的下载函数down_stk_cn010。

指数行情数据下载与A股数据下载类似，只是股票代码范围不同。

索引文件名是F:\zwDat\inx\inx_code.csv，数据文件目录为F:\zwDat\cn\xday\。

如图3-17所示是inx_code.csv索引文件。

	A	B
1	code	name
2	sh	上证指数
3	sz	深圳成指
4	hs300	沪深300指数
5	sz50	上证50
6	zxb	中小板
7	cyb	创业板
8	399001	深证成指
9	399002	深成指R
10	399003	成份B指
11	399004	深证100R
12	399005	中小板指
13	399006	创业板指
14	399100	新指数
15	399101	中小板综
16	399106	深证综指
17	399107	深证A指
18	399108	深证B指
19	399333	中小板R
20	399606	创业板R

图3-17　inx_code.csv索引文件

3.9.4　案例3-9：美股交易数据下载

脚本文件名为zwQuant\demo\ zw_down_usSTK.py。

相关代码如下：

```
# -*- coding: utf-8 -*-


import pandas as pd
#import pandas.io.data as web
import pandas_datareader.data as web
```

```
#zw.xxx
import zwSys as zw  #::zwQT
#-----------

def zw_down_yahoo8code(qx):
     try:
         xcod=qx.code;
         xdat= web.DataReader(xcod,"yahoo",start="1/1/1900");
         fss=qx.rDay+xcod+".csv";print(fss);
         xdat.to_csv(fss);
     except IOError:
         pass    #skip,error

def zw_stk_down_yahoo(qx):
     fss = qx.rdatInx+'inxYahoo.csv';print(fss);
     dinx = pd.read_csv(fss,encoding='gbk')

     i=0;xn9=len(dinx['code']);
     for xcod in dinx['code']:
         i+=1;print("\n",i,"/",xn9,"code,",xcod)
         #
         qx.code=xcod;
         zw_down_yahoo8code(qx);

#------------
```

```
qdat=zw.zwDatX(zw._rdatUS);
#qdat.prDat();


#
zw_stk_down_yahoo(qdat)
```

美股数据下载相对复杂一点，其中有两个函数：

- zw_down_yahoo8code(qx)，按指定股票代码下载数据；
- zw_stk_down_yahoo(qx)，下载所有的美股数据，函数流程图如图3-18所示。

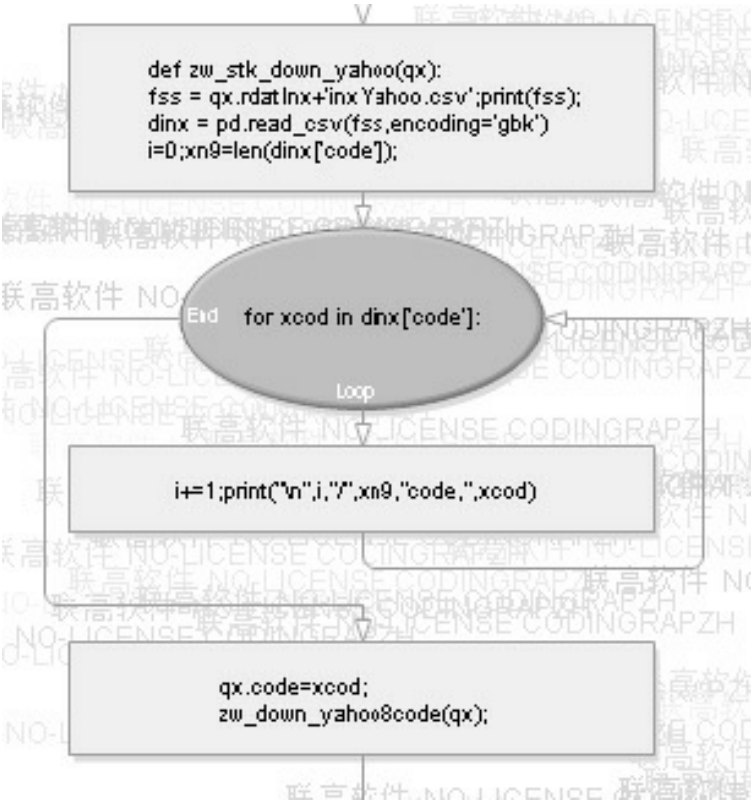


图3-18　函数流程图

美股数据下载使用的不是TuShare，而是pandas内置的数据抓取模块

pandas_datareader。

核心语句是：

```
xdat= web.DataReader(xcod,"yahoo",start="1/1/1900");
```

自0.17版本起，Pandas的数据抓取模块拆分成独立的项目pandas_datareader，原来的导入语句已经废弃：

```
#import pandas.io.data as web
```

新的导入语句是：

```
import pandas_datareader.data as web
```

此外，下载国外和美股数据时，需要注意时区的差异。

3.10　数据归一化处理

数据源很重要，所以在发布zwPython 2016的同时发布了zwDat，它是一个6GB的超大股票数据源，供读者分析。

3.10.1　中美股票数据格式差异

如图3-19所示是下载的美股数据文件。请注意，表格的名称和排列顺序与我国的股票数据格式有所不同。

	A	B	C	D	E	F	G
1	Date	Open	High	Low	Close	Volume	Adj Close
2	2015-1-14	5.73	5.76	5.7	5.76	236900	5.130823
3	2015-1-15	5.79	5.79	5.67	5.7	277100	5.077377
4	2015-1-16	5.68	5.77	5.67	5.77	387600	5.139731
5	2015-1-20	5.79	5.81	5.71	5.74	468100	5.113007
6	2015-1-21	5.64	5.68	5.62	5.65	217400	5.122074
7	2015-1-22	5.68	5.76	5.64	5.73	464400	5.194598
8	2015-1-23	5.72	5.75	5.72	5.74	324000	5.203664
9	2015-1-26	5.73	5.76	5.71	5.74	242300	5.203664
10	2015-1-27	5.68	5.74	5.65	5.72	335600	5.185533
11	2015-1-28	5.75	5.75	5.58	5.61	384500	5.085811
12	2015-1-29	5.63	5.64	5.54	5.62	381100	5.094876
13	2015-1-30	5.59	5.63	5.56	5.57	390500	5.049549
14	2015-2-2	5.6	5.61	5.52	5.6	1029300	5.076745
15	2015-2-3	5.66	5.7	5.65	5.68	544800	5.14927
16	2015-2-4	5.69	5.74	5.68	5.69	551200	5.158336
17	2015-2-5	5.7	5.78	5.7	5.75	554700	5.21273
18	2015-2-6	5.77	5.79	5.71	5.72	373900	5.185533

图3-19　美股数据文件

如图3-20所示是我国A股的数据。

	A	B	C	D	E	F	G
1	date	open	high	close	low	volume	amount
2	2016-2-4	7.23	7.45	7.32	7.16	32063180	234423328
3	2016-2-3	7.03	7.4	7.25	6.93	33325940	240750960
4	2016-2-2	6.98	7.25	7.13	6.83	29214640	207110576
5	2016-2-1	7.28	7.48	6.92	6.65	40092616	286401600
6	2016-1-29	6.29	6.96	6.96	6.21	18242978	123249080
7	2016-1-28	6.57	6.66	6.33	6.2	17300066	112924544
8	2016-1-27	6.86	6.96	6.69	6.18	26494432	172994000
9	2016-1-26	7.49	7.51	6.83	6.83	23455000	167900528
10	2016-1-25	7.65	7.72	7.59	7.51	14960245	114252776

图3-20　A股数据文件

对比图3-19和图3-20，可以发现以下不同之处：

- 美股数据有Adj　Close（调整收盘价）栏，我国A股没有，我国A股有amount（总成交金额）栏，美股数据没有；
- 美股的栏目名称首字母大写，我国A股的栏目名称全部是小写；
- 美股的Low（最低价）数据放在Close（收盘价）数据前面，与我国习惯放置位置不同。

zwDat股票数据包在对A股、美股数据进行归一化处理时，栏目名也进行了统一处理。

3.10.2　案例3-10：数据格式转化

案例脚本文件名为\zwpython\zw_k10\k305_pd_renam.py。

相关代码如下：

```
# -*- coding: utf-8 -*-
```

```
import pandas as pd
import matplotlib as mpl
import zwQTBox as zwBox

#============main
mpl.style.use('seaborn-whitegrid');

cod="002739";#万达院线
fss="dat\\"+cod+".csv";print(fss);
df=pd.read_csv(fss,encoding='gbk');
print("原数据")
print(df.head())

df2=zwBox.df2yhaoo(df);
print("\nYhaoo格式")
print(df2.head())

print('')
cod="ORCL-2000";
fss="dat\\"+cod+".csv";print(fss);
df=pd.read_csv(fss,encoding='gbk');
print("原数据")
print(df.head())

#df2=zwBox.zw_df2yhaoo(df);
df2=zwBox.df2cnstk(df);
print("\ncn中国A股格式")
```

```
print(df2.head())
```

运行结果如下：

```
runfile('F:/zwPython/zw_k10/k305_pd_renam.py', wdir='F:/zwPython
Reloaded modules: zwSys, zwQTBox
dat\002739.csv
原数据
        date   open   high  close   low    volume      amount
0  2016-02-04  80.15  86.11  84.55  80.11  14555358  1199210240
1  2016-02-03  74.97  79.09  78.65  74.65   9597928   739158144
2  2016-02-02  72.70  76.77  75.70  72.02   8403592   634427328
3  2016-02-01  76.00  76.50  72.59  71.30   9300339   691239616
4  2016-01-29  75.91  78.78  77.45  75.30  11651538   896822912

Yhaoo格式
            Open   High   Low   Close    Volume   Adj Close
Date
2016-02-04  80.15  86.11  80.11  84.55  14555358      84.55
2016-02-03  74.97  79.09  74.65  78.65   9597928      78.65
2016-02-02  72.70  76.77  72.02  75.70   8403592      75.70
2016-02-01  76.00  76.50  71.30  72.59   9300339      72.59
2016-01-29  75.91  78.78  75.30  77.45  11651538      77.45

dat\ORCL-2000.csv
```

```
原数据
         Date        Open        High         Low       Close
0  2000-01-03  124.625000  125.187523  111.625000  118.125000  9
1  2000-01-04  115.500000  118.625000  105.000000  107.687523 11
2  2000-01-05  101.625000  106.375000   96.000000  102.000000 16
3  2000-01-06  100.156242  105.000000   94.687523   96.000000 10
4  2000-01-07   95.000000  103.500000   93.562477  103.375000  9

   Adj Close
0  27.425189
1  25.001910
2  23.681433
3  22.288407
4  24.000668
```

查看以上结果，注意转换后栏目首字母大小写不同，以及Low（最低价）与Close（收盘价）的位置的不同。

此外，从A股数据转为Yahoo数据格式时，因为A股数据没有Adj Close栏目，就直接用Close数据代替。实盘时，如果有额外要求，注意用计算后的数据设置Adj Close栏目。

这种转换可以对不同区域的股票数据进行对比，在“一月效应”的案例中，我们已经利用了A股、美股两个不同地区的股票数据进行了对比分析。

此外，许多国外的量化程序采用的数据源是Yahoo财经的模式，无

法直接使用我国的A股数据。

3.10.3 案例3-11：A股策略PAT实盘分析

本脚本源自k201_sta_anz.py，是PyALgoTrade（简称PAT）的一个标准案例程序。为了简化程序，省略了结尾的回报率分析等文字输出，保留了基本的输出和数据绘图部分。

因为PyALgoTradedemo只支持Python 2.7，所以本案例程序只能在Python 2.7环境下调用，相关程序源码，请参见脚本文件：k306_sta_anz.py。

案例3-11程序代码与案例k201_sta_anz.py大同小异，只是在数据设置时增加了一段A股数据转换代码：

```
cod="002739";#万达院线
fss="dat\\"+cod+".csv";
df=pd.read_csv(fss,encoding='gbk');
#df2=zwBox.zw_df2yhaoo(df);
df2=zwBox.df2yhaoo(df);
fss="dat\\"+cod+"_yh.csv";print(fss);
df2.to_csv(fss,encoding='utf-8')
```

如图3-21所示是运行结果。

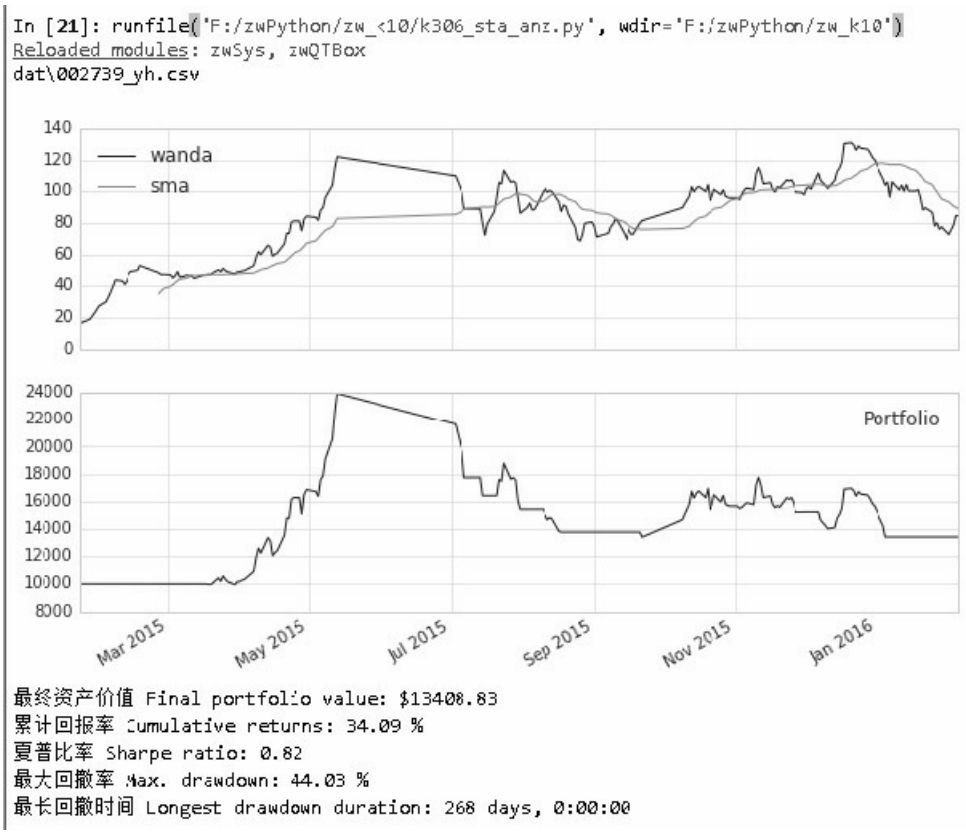


图3-21　运行结果

如果取消程序中第72行（具体请参考源码）的注释，直接使用A股数据作为数据源，会导致运行错误。

```
#fss="dat\\"+cod+".csv";
```

用户修改本书有关案例程序用于实盘或模拟盘时，只需在代码设置环节把变量cod对应的股票代码参数改为自己制定的股票代码，以及把后面几处“wanda”（万达拼音）字符串改为对应的股票代码拼音，就可以使用自己制定的股票数据进行SMA均线策略的回溯测试，代码如下：

```
cod="002739";#万达院线
fss="dat\\"+cod+".csv";
df=pd.read_csv(fss,encoding='gbk');
#df2=zwBox.zw_df2yhaoo(df);
df2=zwBox.df2yhaoo(df);
fss="dat\\"+cod+"_yh.csv";print(fss);
df2.to_csv(fss,encoding='utf-8')
```

3.10.4 案例3-12：数据归一化

案例脚本文件名为\zwpython\zw_k10\k309_zw_stk_xedit.py。

本脚本是zwDat数据包进行归一化处理的正式程序。相关代码如下：

```
# -*- coding: utf-8 -*-

import os
import numpy as np
import pandas as pd

#zwQuant
import zwSys as zw
```

```
import zwQTBox as zwBox

#-----------

def zw_stk_xedit_all(qx,finx,rsr,rtg):
     dinx = pd.read_csv(finx,index_col=False,encoding='gbk')

     i=0;xn9=len(dinx['code']);
     for xcod in dinx['code']:
         i+=1;print("\n",i,"/",xn9,"code,",xcod)
         if (not isinstance(xcod,str)):xcod="%06d" %xcod
         qx.code=xcod;
         fss=rsr+xcod+'.csv';
         if os.path.exists(fss):
             df= pd.read_csv(fss,index_col=0,parse_dates=True,en
             d30=zwBox.xedit_zwXDat(df)
             fss=rtg+xcod+'.csv';print(fss);
             d30.to_csv(fss,encoding='gbk');

#============main
#------init
qdat=zw.zwDatX(zw._rdatCN);
qdat.prDat();

#-----------
#   cnInx
finx=qdat.rdatInx+'inx_code.csv';
```

```
rsr=qdat.rdatCN+"XDay\\"
rtg=qdat.rZWcnXDay
zw_stk_xedit_all(qdat,finx,rsr,rtg);


#   cnSTK
finx=qdat.rdatInx+'stk_code.csv';
rsr=qdat.rdatCN+"Day\\"
rtg=qdat.rZWcnDay   #"\zwDat\zw\cnDay\"
zw_stk_xedit_all(qdat,finx,rsr,rtg);


#  usSTK
finx=qdat.rdatInx+'inxYahoo.csv';
rsr=qdat.rdatUS+"Day\\"
rtg=qdat.rZWusDay   #"\zwDat\zw\usDay\"
zw_stk_xedit_all(qdat,finx,rsr,rtg);
```

程序很简单，会自动批量化对下载的A股数据和美股数据进行归一化处理。处理后的数据保存在目录\zwDat\zw\中。

案例3-11的重点是函数zw_stk_xedit_all，运行逻辑请参看如图3-22所示的函数流程图。

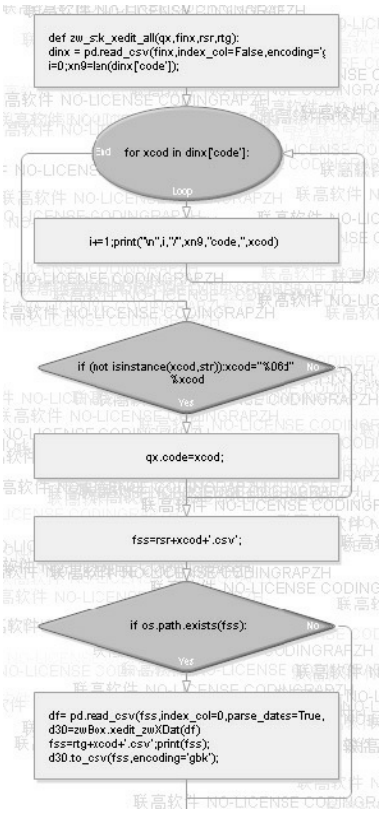

def zw_sk_xedit_all(qx,finx,rsr,rtg):
dinx = pd.read_csv(finx,index_col=False,encoding='
i=0;xn9=len(dinx['code']);
End
for xcod in dinx['code']:
Loop
i+=1;print("\n",i,"/",xn9,"code,",xcod)
if (not isinstance(xcod,str)):xcod="%06d"
%xcod
No
Yes
qx.code=xcod;
fss=rsr+xcod+'.csv';
if os.path.exists(fss):
No
Yes
df= pd.read_csv(fss,index_col=0,parse_dates=True,
d30=zwBox.xedit_zwXDat(df)
fss=rtg+xcod+'.csv';print(fss);
d30.to_csv(fss,encoding='gbk');

图3-22　函数流程图

3.11 为有源头活水来

笔者经常强调，无论是量化交易、高频交易，还是股票期货、外汇黄金，所有金融交易的核心都是策略。策略源自分析，这种分析就是我们常说的量化分析。量化分析的本质其实就是数据分析，如图3-23所示，数据源是最重要的环节之一，也是I2O模型中的三大主模块之一。



图3-23 I2O模型

本章介绍了金融数据的采集，重点是我国A股数据采集和美股数据采集两部分。GitHub开源网站上还有许多其他的金融数据采集项目，读者可以自己检索、学习。

本章在介绍我国A股数据采集模块时采用了部分TuShare文档说明，在此笔者向TuShare的作者及量化工作者表示感谢。

赠人玫瑰，手有余香，开源项目就是这样薪火相传的。

第4章
PAT案例汇编

我们常说，MBA最成功的地方就在于其个案教学体系。本章将集中介绍PyAlgoTrade（简称PAT）中的多个案例：

- 投资组合回报计算；
- SMA策略（简单平均线策略）；
- SMA_cross策略（均线交叉策略）；
- VWAP策略（交易量加权平均价格策略）；
- bbands策略（布林带策略）；
- RSI2策略（相对强弱指标策略）。

通过本章案例的介绍，希望读者能够更加全面地理解Python量化分析的各个方面。本章案例程序原版均来自PAT文档，数据源采用的是Yahoo财经数据，为方便读者理解，本章增加了以下两个修改版本：

- 增强版，对原版进行汉化，修改了数据源配置，文件名以“ed”字符结尾；
- A股版，在增强版的基础上改为我国A股数据，方便读者直接套用，文件名以“zw”字符结尾。

需要说明的是，因为PAT采用的事件编程模式，结构复杂，超出了初学者的学习范畴，因此读者在讲解这些案例时，重点讲解onbars事件

函数中对有关策略的定义和调用。

我们学习量化的目的是学习相关量化策略的操作逻辑，而不是程序编码方面的技术细节，这样在真正的实盘操作中，不管是用简化版本的zwQuant量化软件，还是用其他量化平台，都可以还原这些策略。

4.1　投资组合与回报率

这是PyAlgoTrade作者参加QSTK　2012“量化投资课程”的作业题。QSTK是2012年美国乔治理工大学开办的计算投资公开课，学习Python量化投资项目。当时Pandas数据分析、sci-learn（人工智能软件包）都刚起步，QSTK量化课程采用的算法也比较陈旧，与近两年新兴的Pandas数据分析软件采用的矩阵风格不匹配，已经被逐渐淘汰。

PAT软件作者在作业中使用了PyAlgoTrade软件（简称PAT）。

【作业】

选择4只股票，在2011年投资总计100万美元，计算以下数据：

- 最终组合价值；
- 年收益回报；
- 平均每日收益率；
- 日收益率方差；
- 夏普比率。

这个作业的实质，就是在年初购买4只股票，年内不进行任何交易，年底再计算相关的收益。

4.1.1　案例4-1：下载多组美股数据

在完成这个作业之前，需要下载相关的美股数据。PAT原教程使用

以下命令，通过Yahoo财经数据源下载所需数据：

```
python -c "from pyalgotrade.tools import yahoofinance; yahoofina
python -c "from pyalgotrade.tools import yahoofinance; yahoofina
python -c "from pyalgotrade.tools import yahoofinance; yahoofina
python -c "from pyalgotrade.tools import yahoofinance; yahoofina
```

实际下载数据使用的是极宽量化工具箱zwQuant里面的程序。

脚本文件名为\zwpython\zw_k10\k401_down_mul_us.py。

具体代码如下：

```
# -*- coding: utf-8 -*-

import numpy as np
import pandas as pd
#import pandas.io.data as web
import pandas_datareader.data as web

#zw.Quant
import zwSys as zw
import zwQTBox as zwBox

#--------
```

```
qx=zw.zwDatX(zw._rdatUS);


qx.code="AETI";fss="tmp\\"+qx.code+".csv";
zwBox.down_stk_yahoo010(qx,fss);


qx.code="EGAN";fss="tmp\\"+qx.code+".csv";
zwBox.down_stk_yahoo010(qx,fss)


qx.code="GLNG";fss="tmp\\"+qx.code+".csv";
zwBox.down_stk_yahoo010(qx,fss)


qx.code="SIMO";fss="tmp\\"+qx.code+".csv";
zwBox.down_stk_yahoo010(qx,fss)
```

运行后，在tmp目录下会生成4个抓取的数据文件，如图4-1所示。

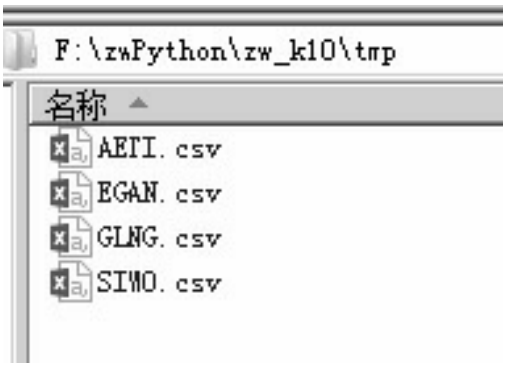
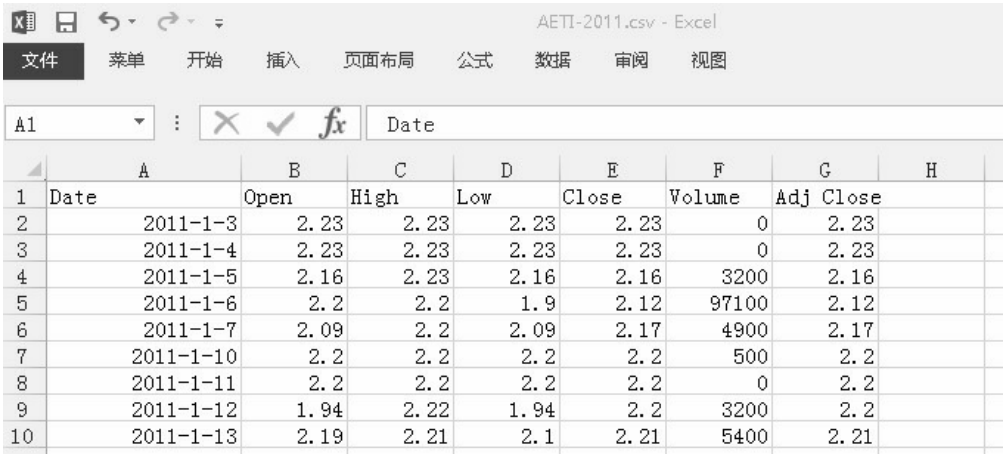


图4-1　下载的4个数据文件

在后面的案例程序中，使用的是2011年的数据，为此，笔者从中人工提取了2011年的数据，并在文件名后添加了“-2011”的字符，复制到dat目录下，供案例程序调用。

如图4-2所示是dat\ARTI-2011.csv数据文件的截图。



	A	B	C	D	E	F	G	H
1	Date	Open	High	Low	Close	Volume	Adj Close	
2	2011-1-3	2.23	2.23	2.23	2.23	0	2.23	
3	2011-1-4	2.23	2.23	2.23	2.23	0	2.23	
4	2011-1-5	2.16	2.23	2.16	2.16	3200	2.16	
5	2011-1-6	2.2	2.2	1.9	2.12	97100	2.12	
6	2011-1-7	2.09	2.2	2.09	2.17	4900	2.17	
7	2011-1-10	2.2	2.2	2.2	2.2	500	2.2	
8	2011-1-11	2.2	2.2	2.2	2.2	0	2.2	
9	2011-1-12	1.94	2.22	1.94	2.2	3200	2.2	
10	2011-1-13	2.19	2.21	2.1	2.21	5400	2.21	

图4-2　美股数据文件截图

4.1.2　案例4-2：投资组合收益计算

案例4-2的数据源是采用案例4-1的脚本下载的，因为案例4-2只使用了2011年的数据，所以笔者采用人工方式对2011年的数据进行了提取，保存在dat目录下。

案例4-2源自PyAlgoTrade的示例程序，原始程序文件名为compinv-1.py。

案例4-2对原始程序脚本的修改之处有：

- 修改了数据文件路径；
- 输出信息进行了部分汉化；
- 增加了自定义函数ret2csv，用于保存日回报率；
- 增加了日回报率的绘图输出。

具体代码请参见脚本文件\zwpython\zw_k10\k402_compinv01.py，仅适用于Python 2.7。

案例4-2的运行结果如图4-3所示。

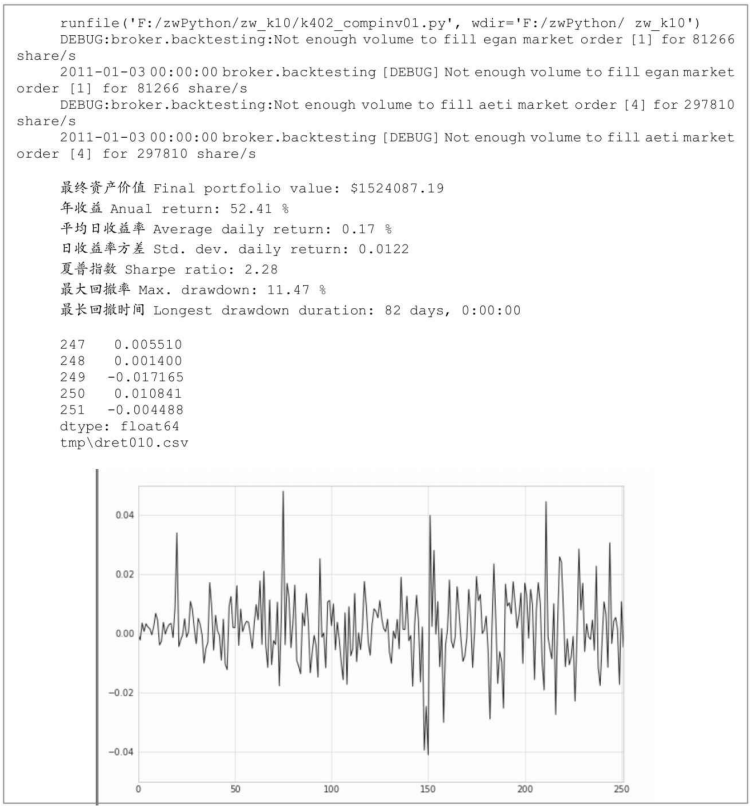


图4-3　运行结果

如图4-3所示是日回报率的绘图输出，横坐标是交易日，共252个交易日（0~251）。

案例4-2相对简单，不涉及策略交易，onBars函数是空函数：

```
def onBars(self, bars):
      pass
```

具体策略是：在年初下单购买指定份额的股票，持有一年，在年底结算。这个是在__init__初始化环节完成的，具体执行指令的是以下脚本：

```
for instrument, quantity in orders.items():
       self.marketOrder(instrument, quantity, onClose=True, allO
```

注意最终资产价值前面的几行DEBUG（调试）英文信息：

```
DEBUG:broker.backtesting:Not enough volume to fill egan market o
2011-01-03 00:00:00 broker.backtesting [DEBUG] Not enough volume
DEBUG:broker.backtesting:Not enough volume to fill aeti market o
2011-01-03 00:00:00 broker.backtesting [DEBUG] Not enough volume
```

以上代码表示市场上没有81266手的egan股票和297810手的aeti股票，因为预设的下单指令是allOrNone（全部成交或者取消）。

所以，股票代码为egan和aeti的这两只股票没有买到，这个也是QSTK作业中故意设置的市场场景。

下面就是最终资产价值、年收益、平均日收益率、日收益率方差、夏普指数、最大回撤率和最长回撤时间等常见的量化分析参数。

- 最终资产价值（Final portfolio value）: $1524087.19。
- 年收益（Anual return）: 52.41 %。

- 平均日收益率（Average daily return）: 0.17 %。
- 日收益率方差（Std. dev. daily return）: 0.0122。
- 夏普指数（Sharpe ratio）: 2.28。
- 最大回撤率（Max. drawdown）: 11.47 %。
- 最长回撤时间（Longest drawdown duration）: 82 days, 0:00:00。

这些参数是最基本也是最常用的量化投资评估参数，下一节会介绍自行编程分析投资策略来计算以上参数的方法。

最后几行输出数据如下：

```
247    0.005510
248    0.001400
249   -0.017165
250    0.010841
251   -0.004488
dtype: float64
tmp\dret010.csv
```

以上代码表示一年252个交易日，每天的日收益率；具体操作是由自定义函数ret2csv完成的：

```
def ret2csv(ftg):
     xd=retAnalyzer.getReturns()
     x8=[];
```

```
    for x1 in xd:
        x8=x8+[x1];
    xs1=pd.Series(x8);
    print(xs1.tail())
    xs1.to_csv(ftg)
    print(fss)

    return xs1
```

注意日收益率数据文件tmp\dret010.csv，在后面的案例中会用到里面的相关数据。为避免清理tmp目录时删除该数据文件，可以把该数据文件更名为k04-dret010.csv并复制到dat目录下，以便以后查阅。

4.2 SMA均线策略

按照PAT量化软件的说明文档，量化交易策略可以分为以下三类：

- 动量（Momentum），如VWAP动量策略、SMA均线交叉策略、market_timing交易时机策略等；
- 均值回归策略（Mean Reversion），如布林带策略、RSI2策略等；
- 其他策略。

4.2.1 SMA简单移动平均线

SMA（Simple Moving Average）是简单移动平均线，又称算术移动平均线，是指对特定期间的收盘价进行简单平均化的意思。一般提及移动平均线，即指简单移动平均线（SMA）。

简单移动平均线沿用最简单的统计学方式，将过去某特定时间内的价格取其平均值。简单移动平均线计算方法如同其名——简单。

以5天移动平均线为例，简单移动平均线的公式如下：

SMA=(C1+C2+C3+C4+C5)/5

一般公式：

SMA=(C1+C2+C3+C4+C5+.....+Cn)/n

Cn：第n日收盘价。

n：移动平均数周期。

大部分技术分析者利用简单移动平均线设定一个买卖系统。用两条平均线中较短期的一条线作为信号线。例如，当短期的平均线上穿较长期的平均线时，显示牛市在即；相反，当短期的平均线下穿较长期的平均线时，显示熊市在即。简单移动平均线也可作为支持及阻力之用。

4.2.2 案例4-3：原版SMA均线策略

下面用一个简单的SMA均线策略模拟一次实际的交易，策略很简单：

- 如果调整收盘价，高于SMA(15)：15日均线价格，输入多头位置（下单买进）。
- 如果调整收盘价，低于SMA(15)：15日均线价格，退出多头位置（空头，下单卖出）。

这个案例程序源于PyAlgoTrade量化程序，原始文件名是tutorial-4-trade.py。

案例4-3程序代码仅修改了文件路径，其他未做任何修改，有关脚本请参见案例文件\zwpython\zw_k10\k403_tur04sr.py。

案例4-3运行结果如下：

```
runfile('F:/zwPython/zw_k10/k403_tur04sr.py', wdir='F:/zwPython/
INFO:strategy:BUY at $26.35
2000-01-26 00:00:00 strategy [INFO] BUY at $26.35
……
……
2000-12-21 00:00:00 strategy [INFO] SELL at $25.83
INFO:strategy:BUY at $28.21
2000-12-22 00:00:00 strategy [INFO] BUY at $28.21
Final portfolio value: $974.53
```

案例4-3程序很简单，核心语句就一条：

```
myStrategy = MyStrategy(feed, "orcl", smaPeriod)
```

策略定义还是onBars函数：

```
def onBars(self, bars):
     # Wait for enough bars to be available to calculate a SMA.
     if self.__sma[-1] is None:
         return

     bar = bars[self.__instrument]
     # If a position was not opened, check if we should enter a
```

```
    if self.__position is None:
        if bar.getPrice() > self.__sma[-1]:
            # Enter a buy market order for 10 shares. The order
            self.__position = self.enterLong(self.__instrument,
    # Check if we have to exit the position.
    elif bar.getPrice() < self.__sma[-1] and not self.__positio
        self.__position.exitMarket()
```

策略很简单，采用的是15日SMA均线：

- 如果调整收盘价，高于前一天的日均线：SMA[-1]，-1表示使用前一个交易日的数据，下单买进10手。
- 如果调整收盘价，低于前一天的日均线：SMA[-1]，下单卖出10手。

如图4-4所示是策略函数的流程图。

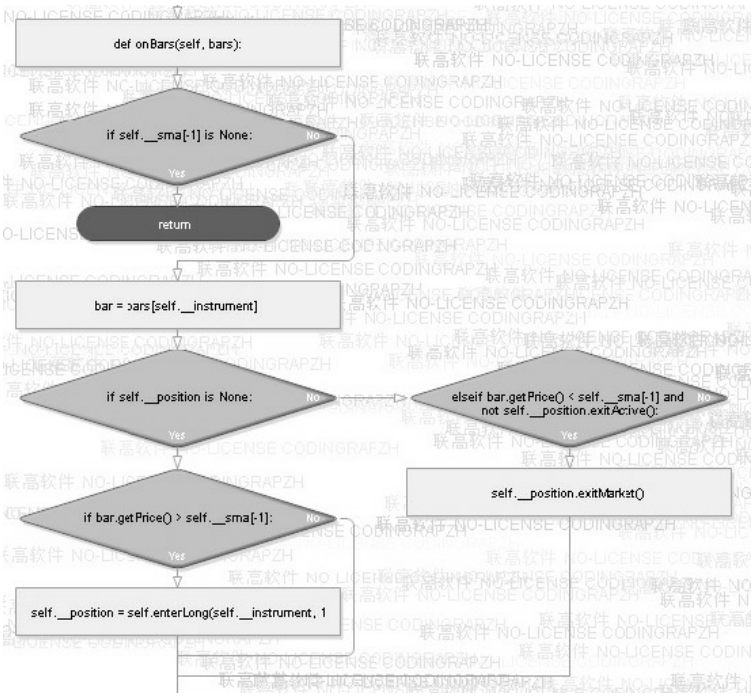


图4-4　策略函数流程图

我们对比案例4-3与案例2-1。

案例2-1的文件名为\zwpython\zw_k10\k201_sta_anz.py，该案例使用的是交叉平均线策略，比本案例更加复杂：

```
def onBars(self, bars):
    # If a position was not opened, check if we should enter a
    if self.__position is None:
        if cross.cross_above(self.__prices, self.__sma) > 0:
            shares = int(self.getBroker().getCash() * 0.9 / bar
            # Enter a buy market order. The order is good till
            self.__position = self.enterLong(self.__instrument,
    # Check if we have to exit the position.
    elif not self.__position.exitActive() and cross.cross_below
```

```
            self.__position.exitMarket()
```

案例2-1的策略函数流程图如图4-5所示。

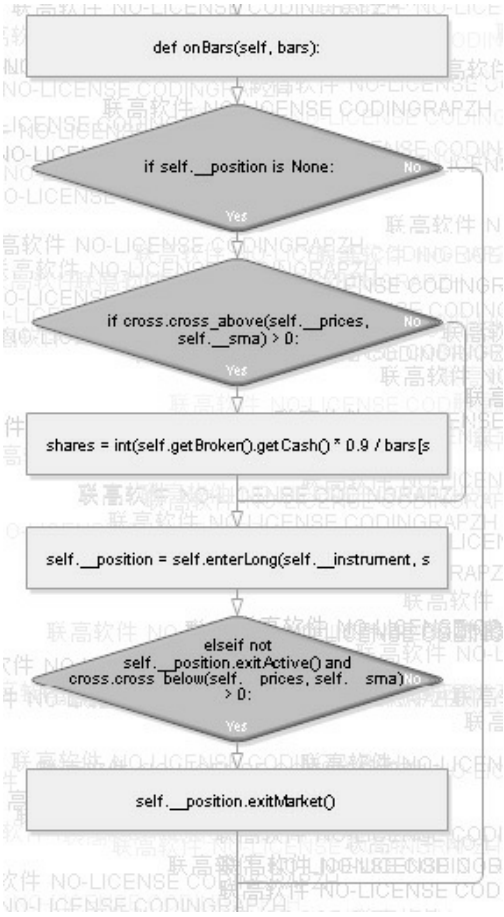


图4-5　案例2-1 SMA策略函数流程图

- 当股票价格高于SMA平均线价格，并且是向上趋势，买入；
- 当股票价格低于SMA平均线价格，并且是向下趋势，卖出。

案例2-1策略的函数流程图如图4-5所示。

对比案例4-3与案例2-1，读者可以体验一下，通过不同的onBars事件程序采用不同的分析策略。

4.2.3 案例4-4：增强版SMA均线策略

为了便于理解，根据案例2-1对案例4-3进行修改。

有关程序源码请参看脚本文件\zwpython\zw_k10\k404_tur04ed.py。

修改后的案例程序增加了绘图输出和回报率、夏普比率、最大回撤率等指标的计算。此外，因为绘图数据源自getSMA类函数：

```
plt.getInstrumentSubplot('orcl').addDataSeries("sma", myStrategy
```

所以，类定义增加了一个getSMA类函数：

```
def getSMA(self):
      return self.__sma
```

其他部分未做大的修改。如图4-6所示是运行结果。

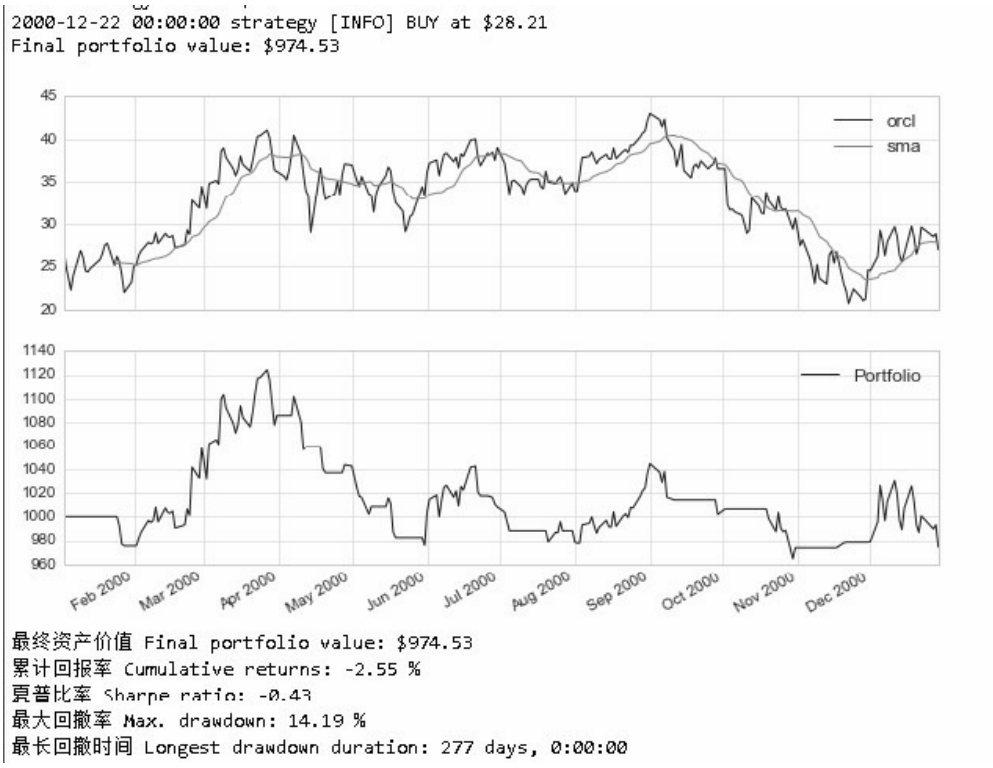


图4-6　运行结果

对照图4-6，再看看下面的类定义事件（on）定义事件（on）函数：

```
def onEnterOk(self, position):进入交易

def onExitOk(self, position):退出交易
```

以上事件（on）函数对应买卖交易流水清单，最终资产价值（Portfolio）的变化过程就更加清楚了。

（1）在2000年年初，有1000美元本金。

（2）2月底，略有亏损，资产总值大约是980美元。

（3）最高峰是4月下旬，资产总值大约1120美元。

（4）随后就是起起伏伏，到了10月，资产总值还处于盈利状态，大约为1030美元。

（5）12月略有亏损，最终资产总值为974.53美元，亏损2.55%。

4.2.4 案例4-5：A股版SMA均线策略

参照案例4-4，使用我国的A股数据对案例2-1进行修改。

案例4-5程序源码请参见脚本文件\zwpython\zw_k10\k405_tur04zw.py。

在案例4-5的代码中，修改部分在主流程开头，增加了一个数据格式转换部分，并且增加了一个cname变量，便于修改股票代码名称：

```
#--------数据格式转换，常用国内A股数据，转换为Yahoo财经格式
cod="002739";#万达院线
cname='wanda';
fss="dat\\"+cod+".csv";
df=pd.read_csv(fss,encoding='gbk');
#df2=zwBox.zw_df2yhaoo(df);
df2=zwBox.df2yhaoo(df);
```

```
cfn="dat\\"+cod+"_yh.csv";print(fss);
df2.to_csv(cfn,encoding='utf-8')
```

读者可以自行修改以上的股票代码变量“cod”和名称变量“cname”，并套用自己熟悉的股票数据。使用前要记得把指定的股票交易数据复制到下面的数据目录中：

```
x:\zwPython\zw_k10\dat\
```

案例4-5运行结果如图4-7所示。

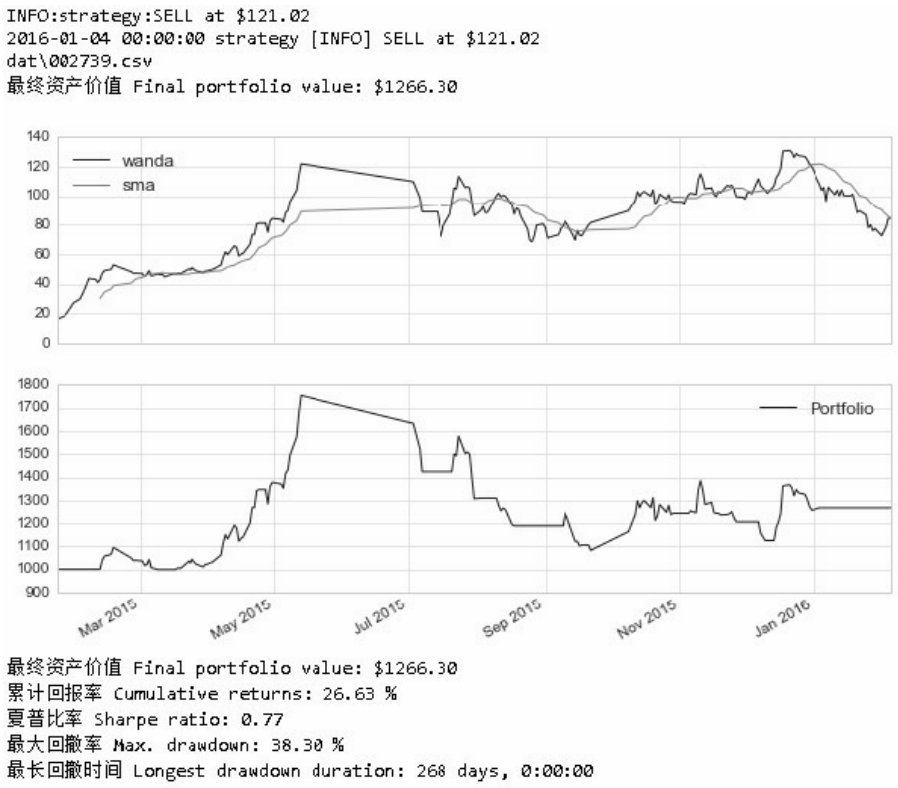


图4-7　案例4-5运行结果

4.3 均线交叉策略

均线是将每天的收盘价加权平均，从而得到一条带有趋势性的轨迹。均线系统是大多数分析者常用的技术工具，从技术角度看是影响技术分析者心理价位的决定因素，是技术分析者的良好参考工具，相比价格变化是滞后的。

均线指标实际上是移动平均线指标的简称。由于该指标是反映价格运行趋势的重要指标，其运行趋势一旦形成，将在一段时间内继续保持，趋势运行所形成的高点或低点又分别具有阻挡或支撑作用，因此均线指标所在的点位往往是十分重要的支撑或阻力位，这就为我们提供了买进或卖出的有利时机，均线系统的价值也正在于此。例如：

- 均线向上是均线多头；
- 均线向上产生的交叉是金叉，反之是死叉。

以每天的前9天和当天共10天的收盘价取算术平均值，再以若干天的这种算术平均值而连结的曲线就是10日均线。同样，有10分钟均线、10小时均线，还有以周、月、年等不同的时间单位做成的各种均线。

以上是常见的做法，还有人取每天的平均价和均权平均值等，做法不一。K线图中常标以MA5、MA10等。

由于均线对股价趋势有一定的比照作用，所以它对于技术分析相当重要。一般以日线MA5、MA10分析短期走势，以MA30、MA60分析中期走势，以M125和M250分析中长期走势。

由于从均线可以动态分析股价的走势，所以常有人以均线来设置止损点及止赚点（高抛点），其实就是起到一种通过技术分析而确定的活动标尺的作用，都有相对的参考价值。

4.3.1 案例4-6：均线交叉策略

案例4-6程序源码参见文件\zwpython\zw_k10\k406_sma_crossover_sample.py。

案例4-6运行结果代码如下，如图4-8所示是运行结果。

```
Sharpe ratio: 1.12
```

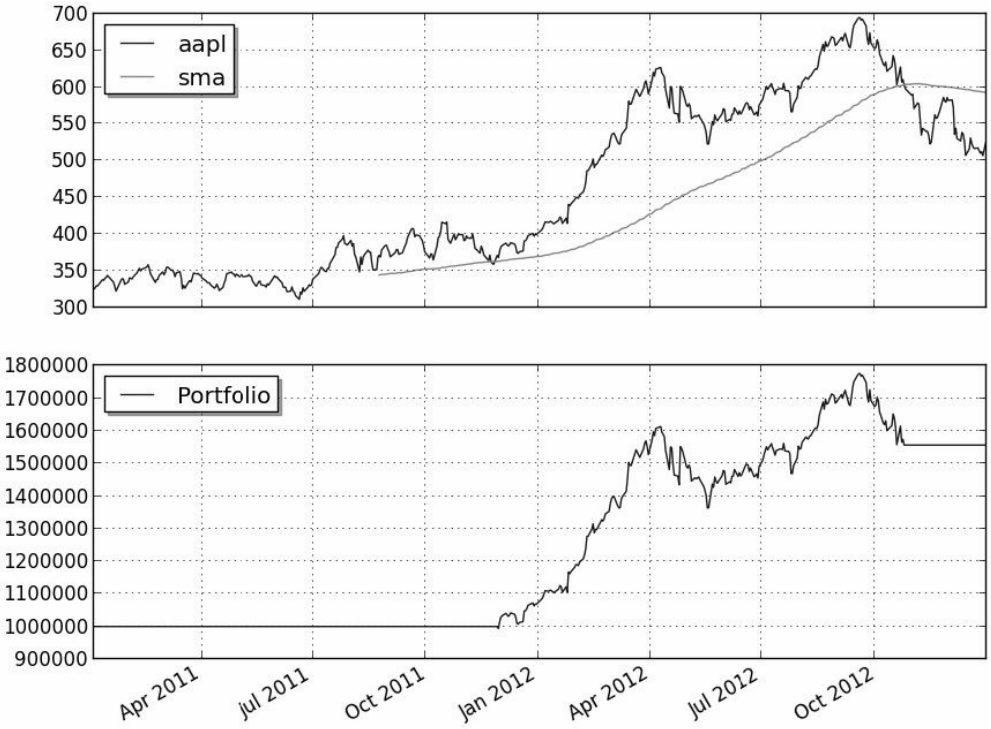


图4-8　运行结果

读者可以调整不同的参数重新运行测试。注意在案例4-6中定义了主入口函数：

```
def main(plot):
```

主入口函数名采用了C语言的习惯名称main。这个main不是必需的，可以更改，如main_SMA、SMA100等。

Python的主入口程序调用与其他编程语言的习惯不同，参见以下脚本：

```
if __name__ == "__main__":
     main(True)
```

有关细节读者可查看相关的Python用户文档，这里就不细说了。

此外，在案例4-6中，sma_crossover均线交叉策略采用了外部模块文件，用import导入：

```
import sma_crossover
```

sma_crossover.py文件源码如下：

```
from pyalgotrade import strategy
from pyalgotrade.technical import ma
from pyalgotrade.technical import cross

class SMACrossOver(strategy.BacktestingStrategy):
    def __init__(self, feed, instrument, smaPeriod):
        strategy.BacktestingStrategy.__init__(self, feed)
        self.__instrument = instrument
        self.__position = None
        # We'll use adjusted close values instead of regular cl
        self.setUseAdjustedValues(True)
        self.__prices = feed[instrument].getPriceDataSeries()
        self.__sma = ma.SMA(self.__prices, smaPeriod)

    def getSMA(self):
        return self.__sma

    def onEnterCanceled(self, position):
        self.__position = None

    def onExitOk(self, position):
        self.__position = None

    def onExitCanceled(self, position):
```

```
        # If the exit was canceled, re-submit it.
        self.__position.exitMarket()

    def onBars(self, bars):
        # If a position was not opened, check if we should ente
        if self.__position is None:
            if cross.cross_above(self.__prices, self.__sma) > 0
                shares = int(self.getBroker().getCash() * 0.9 /
                # Enter a buy market order. The order is good t
                self.__position = self.enterLong(self.__instrum
        # Check if we have to exit the position.
        elif not self.__position.exitActive() and cross.cross_b
            self.__position.exitMarket()
```

由源码可以看出，sma_crossover.py文件就是一个简单的类定义。这种主控文件（main）、策略等其他模块部分采用不同文件进行模块分离，也是Python项目的一个习惯。

4.3.2 案例4-7：A股版均线交叉策略

案例4-7是A股版的均线交叉策略，就是在案例4-6的基础上把数据源改为A股数据。

笔者常说，各种金融产品对于量化回溯而言，流程、架构都是相同的，只是数据源不同。如果读者采用每天最新的实时数据进行量化分

析，再按照分析的结果进行下单交易，就是实盘的量化交易；如果不下单，只是把回溯结果与最新的实盘数据进行交叉对比，就是量化交易模拟盘。当然，在进行实盘交易之前，笔者建议至少要进行300~500次的模拟盘训练，强化实战技巧。

在量化数据源方面，无论是A股还是美股，甚至是期货、外汇、黄金及其他各种金融产品，只要有稳定的数据源，都可以通过PAT或zwQuant等量化软件进行量化分析。因为对于各种量化软件而言，分析的都是数据。在使用不同的数据源之前，需要注意数据的格式，必须与量化软件的数据格式保持一致。

案例4-7程序源码请参看文件\zwpython\zw_k10\k407_sma_crossover_zw.py。

案例4-7是A股增强版本，除了增加数据转换部分外，还在第一行增加了以下代码：

```
# -*- coding: utf-8 -*-
```

因为在代码内使用了中文，所以加入以上语句，强制采用utf-8内码，否则可能会出现乱码，特别是Python　2.7。这里面的原因，有兴趣的读者可以自己检索相关资料。

如图4-9所示是本案例的运行结果。

```
In [11]: runfile('F:/zwPython/zw_k10/k407_sma_crossover_zw.py', wdir='F:/zwPython/zw_k10')
Reloaded modules: sma_crossover
dat\002739.csv
夏普指数 Sharpe ratio: 0.00
```

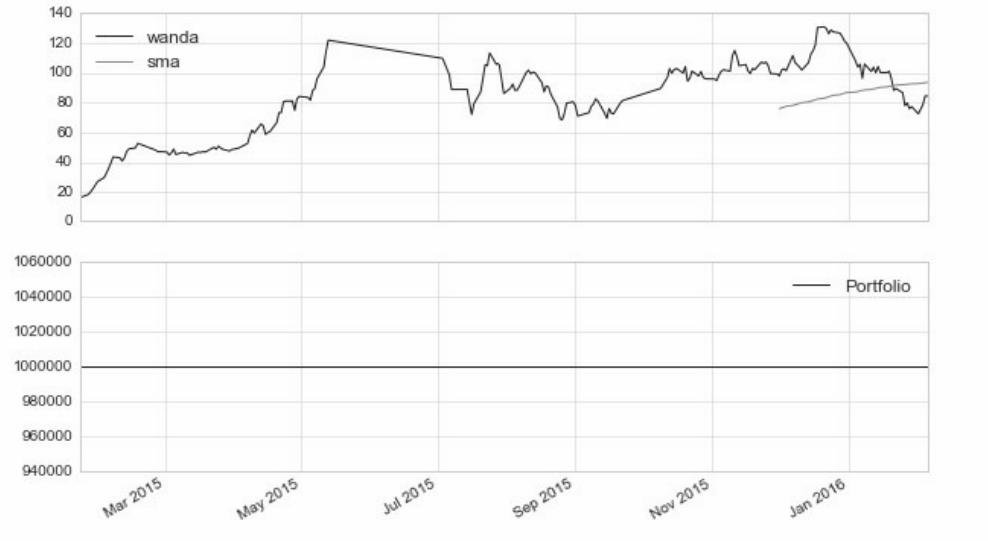


图4-9　运行结果

图4-9初看好像是正常的，可细看问题就大了：夏普指数居然为零；下面的图表中显示的Portfolio投资组合总值居然是一根直线；再看看SMA均线图形，居然只有短短的一截。

原因是：采用的A股数据源“万达”在2015年上市后才有数据，PAT量化软件原来的案例使用的是苹果2011—2012年两年的数据，而且原案例中采用的均线周期是163天，因此出现这个问题。

```
22    #smaPeriod = 163
23    smaPeriod = 20
```

我们把案例4-7代码中的第22行屏蔽，改用第23行的数据，均线周期设为20天，再运行看一下。

如图4-10所示是修正后的运行结果。

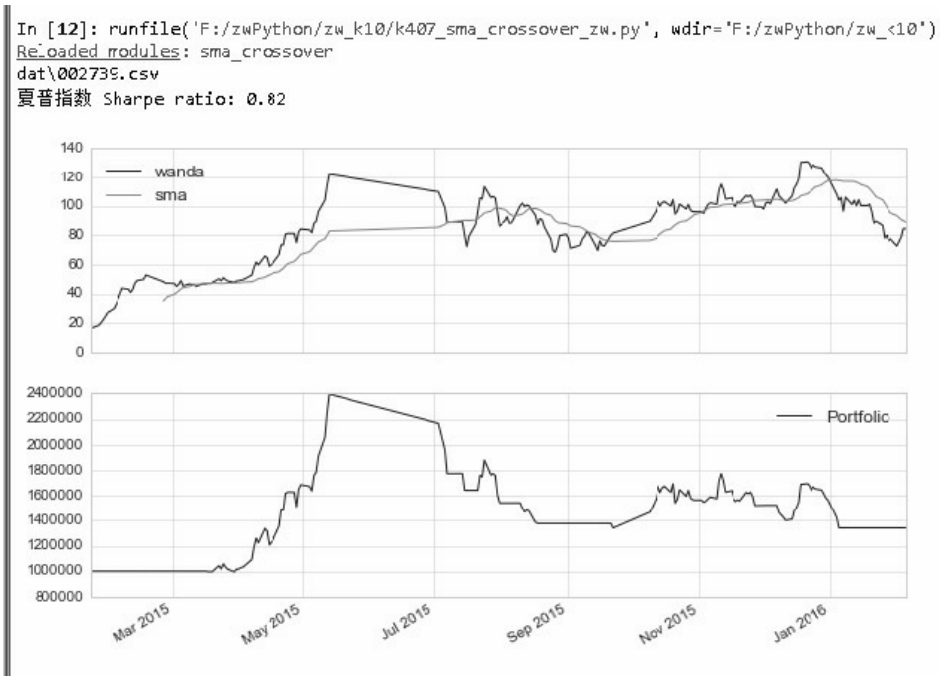


图4-10　修正后的运行结果

修正后的运行结果中，夏普指数、SMA均线图、Portfolio投资组合资产总值图表全部正常。

4.4 VWAP动量策略

VWAP（交易量加权平均价格）策略是比较常见的被动型交易策略。被动型交易策略主要有VWAP、TWAP、PEG算法等，下面进行简单介绍。

1. 交易量加权平均价格（VWAP）

交易量加权平均价格（VWAP）是使用最广泛的算法交易策略。据统计，国际证券市场使用算法交易完成的成交量中，超过一半的算法交易是使用VWAP及其衍生算法交易完成的。

VWAP算法交易的原则是每一段时间完成交易的总量占这段时间内市场总交易的比例恒定，理想的情况下，这个算法交易实现的成交价格等于一段时间内的市场成交均价。VWAP算法交易的目的是最小化冲击成本，并不寻求最小化所有成本。理论上，在没有额外的信息也没有针对股票价格趋势预测的情况下，VWAP是最优的算法交易策略。

2. 时间加权平均价格（TWAP）

时间加权平均价格（TWAP）策略与VWAP策略类似，不同的是，这个方法并不预测交易期内成交量的分布。TWAP算法交易把交易期划分为若干时间片以后，按每个时间片的长度权重分配该时间段内需完成的交易量，该策略其他交易程序与VWAP相同。

3. 盯住盘口策略（PEG）

盯住盘口策略（PEG）每时每刻都根据市场盘口的现状下达交易指令；当指令执行的需求比较迫切时，可以不在现有的盘口，而是选择市场的买卖中间价下达交易指令，以使我们的指令能够尽早执行。

4. IS策略

IS策略是按投资者的个人偏好权衡优化一笔交易的市场冲击与事件风险，尽量减小最终实际成交价格与目标价之间的差距。这里目标价可以是开盘价、收盘价，或者是到达价格，即交易指令下达时的市场价格。通常，该类策略都允许交易人员设置买入（出）时的最高（低）容忍价格，并按照交易速度的要求选择激进、中性和保守的策略风格。

5. SOR策略

SOR策略即下单路径优选策略，通过对不同渠道实时交易数据的分析，在保证成交量的前提下寻求最优价格。

以上是量化交易中使用最为频繁的5种被动型交易策略，此外还有一些机构为客户量身定制的策略，如隐身策略（Stealth）、游击策略（Guerrilla）、狙击策略（Sniper）和嗅探策略（Sniffer）等。

4.4.1 案例4-8：VWAP动量策略

本案例代码细节参见https://www.quantopian.com/posts/momentum-trade。

案例4-8源码参见文件

\zwpython\zw_k10\k408_vwap_momentum.py。

如图4-11所示是案例4-8 VWAP策略的运行结果。

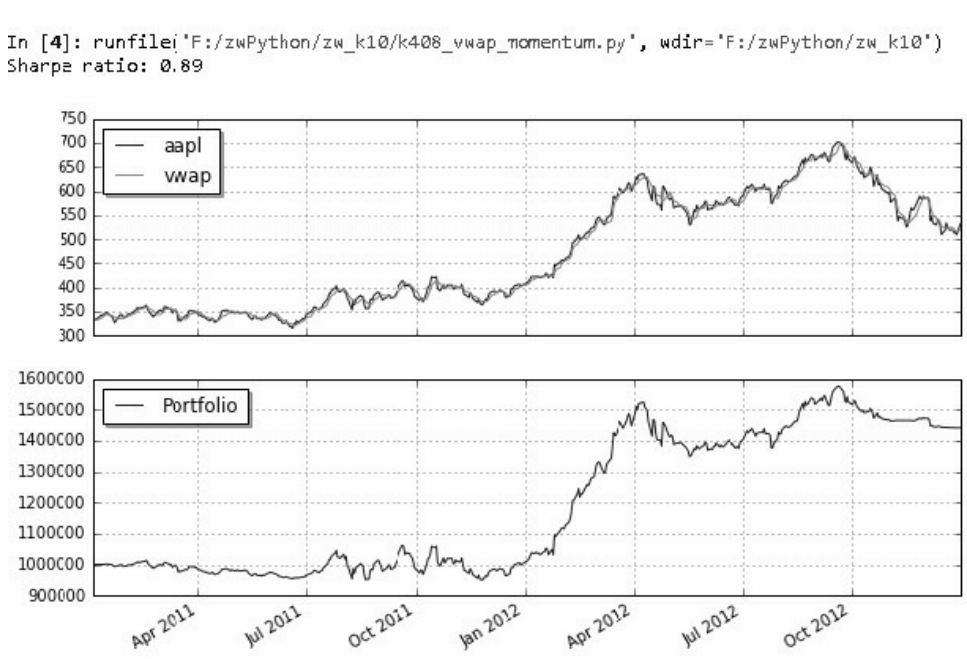


图4-11　VWAP策略运行结果

读者可以调整不同的参数，重新运行本案例进行测试。

4.4.2　案例4-9：A股版VWAP动量策略

本案例根据案例4-8进行了修改，把数据源改为A股数据。

案例4-9源码请参见文件\zwpython\zw_k10\k409_vwap_momentum_zw.py。

案例4-9是A股增强版本，除了增加数据转换部分外，还在第一行增加了utf-8编码指令。

需要说明的是，VWAP、TWAP和PEG被动型交易策略主要是大盘资金、庄家操作使用，个人用户一般用得很少。

对于规模较大的证券交易，如果一次性全部按市价下单，则该交易会造成巨大的市场冲击；如果分成几笔，在不同时间段内成交，投资者又面临市场价格和流动性发生不利变动的时间风险。个人用户资金量很少，通常无须考虑以上问题。

如图4-12所示是A股VWAP策略运行结果。

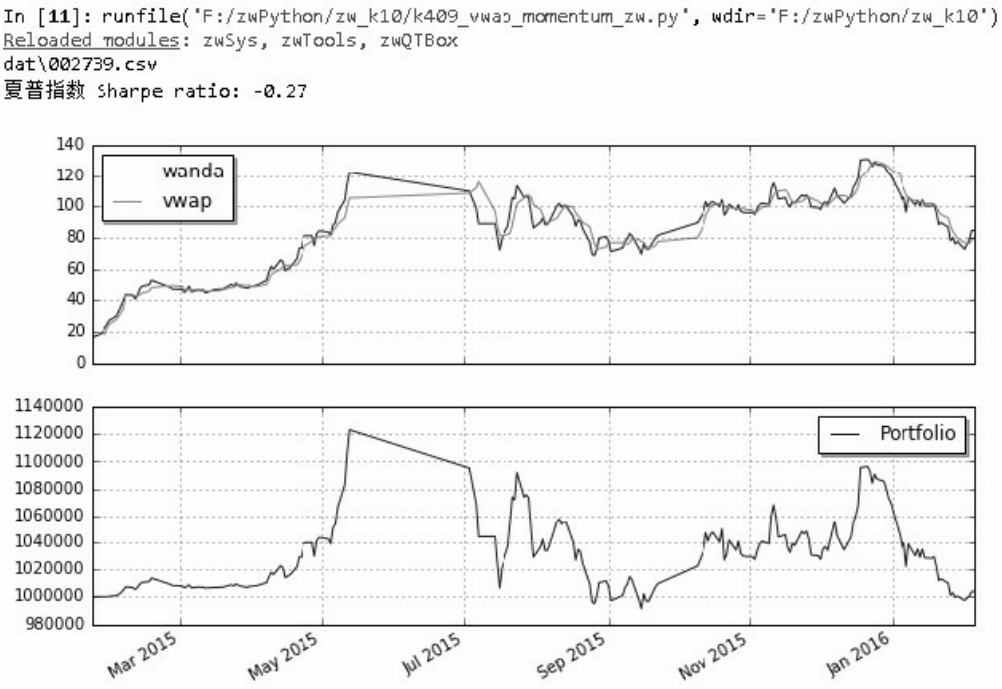


图4-12　A股VWAP策略运行结果

4.5 布林带策略

如图4-13所示是布林带（Bollinger Band）策略示意图。

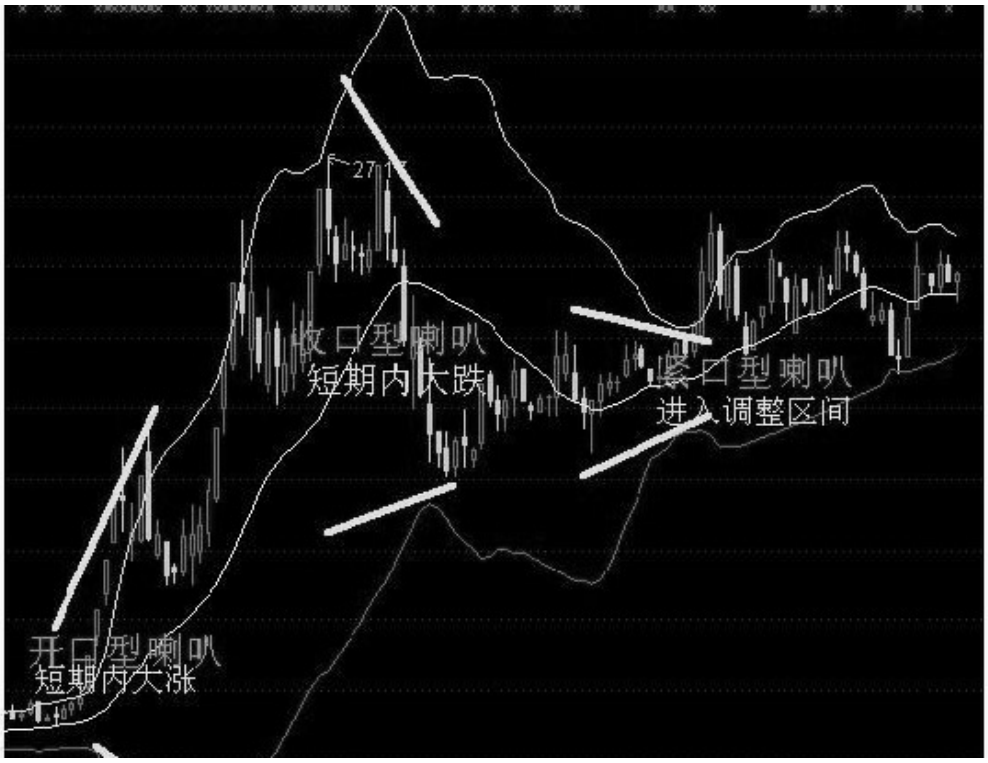


图4-13 布林带策略示意图

布林线（Boll）指标是通过计算股价的“标准差”，再求股价的“信赖区间”。该指标在图形上画出三条线，其中上下两条线可以分别看成是股价的压力线和支撑线，而在两条线之间还有一条股价平均线，布林线指标的参数最好设为20。一般来说，股价会运行在压力线和支撑线所形成的通道中。

在所有的计算指标中，Boll指标的计算方法是最复杂的指标之一，它引进了统计学中的标准差概念，涉及中轨线（MB）、上轨线（UP）

和下轨线（DN）的计算。

和其他指标的计算一样，由于选用的计算周期不同，Boll指标也包括日Boll指标、周Boll指标、月Boll指标、年Boll指标和分钟Boll指标等各种类型。经常用于股市研判的是日Boll指标和周Boll指标。虽然它们计算时的取值有所不同，但基本的计算方法一样。

以日Boll指标计算为例，其计算方法如下：

中轨线=N日的移动平均线

上轨线=中轨线+两倍的标准差

下轨线=中轨线－两倍的标准差

日Boll指标的计算过程如下。

（1）计算MA。

MA=N日内的收盘价之和÷N。

（2）计算标准差MD。

MD=平方根N日的（C－MA）的二次方之和÷N

（3）计算MB、UP、DN线。

MB=（N－1）日的MA

UP=MB+K×MD。

DN=MB－K×MD。

在股市分析软件中，Boll指标共由四条线组成，即上轨线UP、中轨线MB、下轨线DN和价格线。其中，上轨线UP是UP数值的连线，用黄色线表示；中轨线MB是MB数值的连线，用白色线表示；下轨线DN是DN数值的连线，用紫色线表示；价格线的颜色为浅蓝色。和其他技术指标一样，在实战中，投资者不需要进行Boll指标的计算，主要是了解Boll的计算方法和过程，以便更加深入地掌握Boll指标的实质，为运用指标打下基础。

4.5.1 案例4-10：布林带策略

本案例代码细节可参考http://www.investopedia.com/articles/trading/07/bollinger.asp。

案例4-10源码请参见文件\zwpython\zw_k10\k410_bbands.py。

运行结果如下：

```
DEBUG:broker.backtesting:Not enough cash to fill yhoo order [1]
2011-07-21 00:00:00 broker.backtesting [DEBUG] Not enough cash t
……
DEBUG:broker.backtesting:Not enough cash to fill yhoo order [9]
2012-02-21 00:00:00 broker.backtesting [DEBUG] Not enough cash t
Sharpe ratio: 0.71
```

如图4-14所示是布林带策略运行结果。

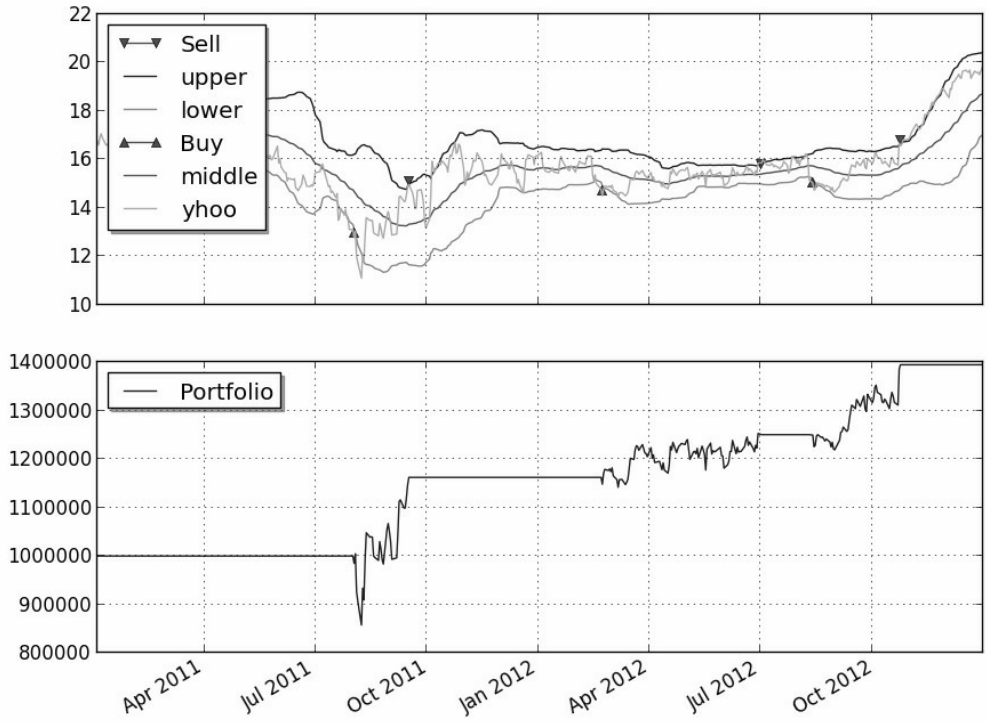


图4-14　布林带策略运行结果

注意输出信息中的代码：

```
2011-07-28 00:00:00 broker.backtesting [DEBUG] Not enough cash t
DEBUG:broker.backtesting:Not enough cash to fill yhoo order [5]
```

以上代码表示，在某些交易机会，因为没有足够的现金，而出现了“漏单”现象。

4.5.2　案例4-11：A股版布林带策略

案例4-10根据案例4-10进行了修改，把数据源改为A股数据。

有关源码请参见文件\zwpython\zw_k10\k411_bbands_zw.py。

案例4-11是A股增强版本，除了增加数据转换部分外，还在第一行增加了utf-8编码指令。

运行结果如下：

```
runfile('F:/zwPython/zw_k10/k411_bbands_zw.py', wdir='F:/zwPytho
DEBUG:broker.backtesting:Not enough cash to fill wanda order [1]
2015-08-26 00:00:00 broker.backtesting [DEBUG] Not enough cash t
DEBUG:broker.backtesting:Not enough cash to fill wanda order [2]
2015-08-27 00:00:00 broker.backtesting [DEBUG] Not enough cash t
dat\002739.csv
夏普指数 Sharpe ratio: 0.40
```

如图4-15所示是A股布林带策略运行结果。

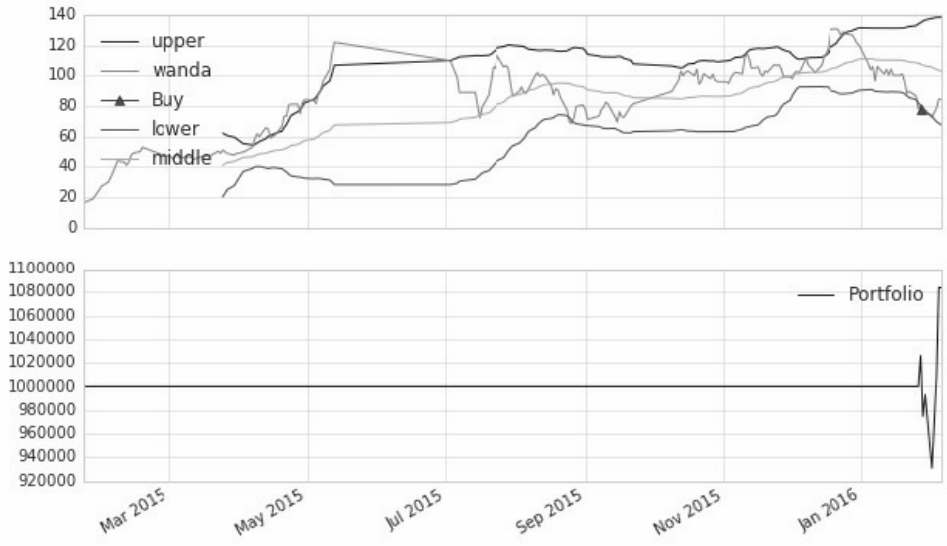


图4-15　A股布林带策略运行结果

有关参数的修改与前面A股版本案例类似，读者可自行查看。

从图4-15中，可以看到下方的Portfolio资产总值曲线大部分是一条直线。这是因为默认的时间周期是40天，将代码改为10天后再运行看一下。

参看程序源码文件，如图4-16所示，把第49行屏蔽，改用第50行的数据。

```
49    #bBandsPeriod = 40
50    bBandsPeriod = 10
```

图4-16　修改布林带策略参数

如图4-17所示是修改参数后的A股布林带策略运行结果。

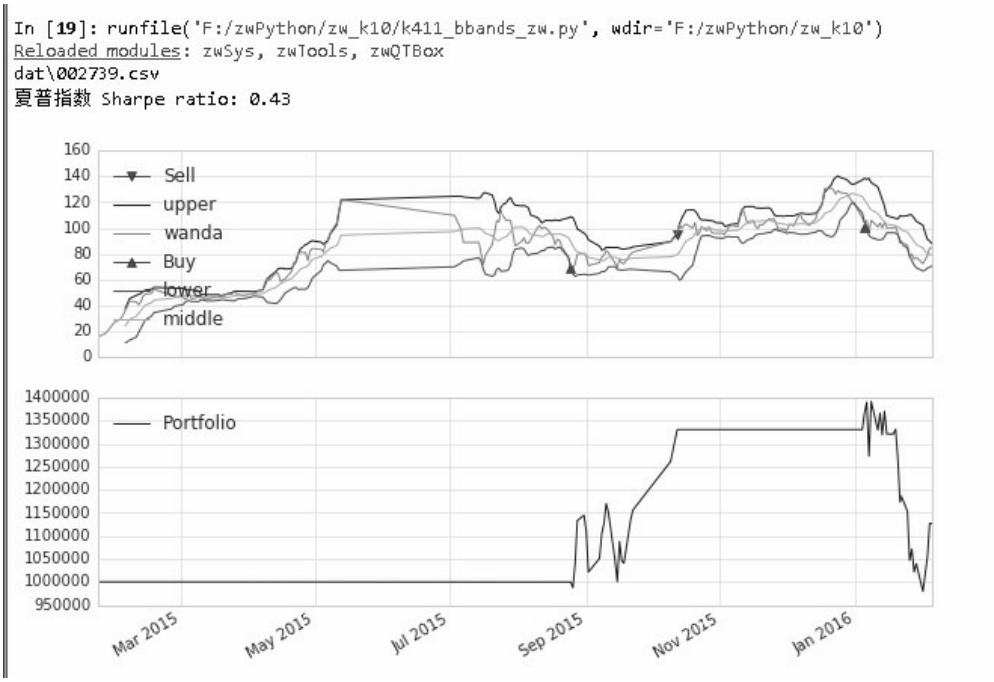


图4-17　修改参数后的A股布林带策略运行结果

修改参数后，没有漏单信息提示。

4.6　RSI2策略

相对强弱指标（RSI，Relative Strength Index）最早由威尔斯·威尔德（Welles Wilder）应用于期货买卖，后来人们发现在众多的图表技术分析中，强弱指标的理论和实践极其适合于股票市场的短线投资，于是被用于股票升跌的测量和分析中。如图4-18所示是RSI策略示意图。

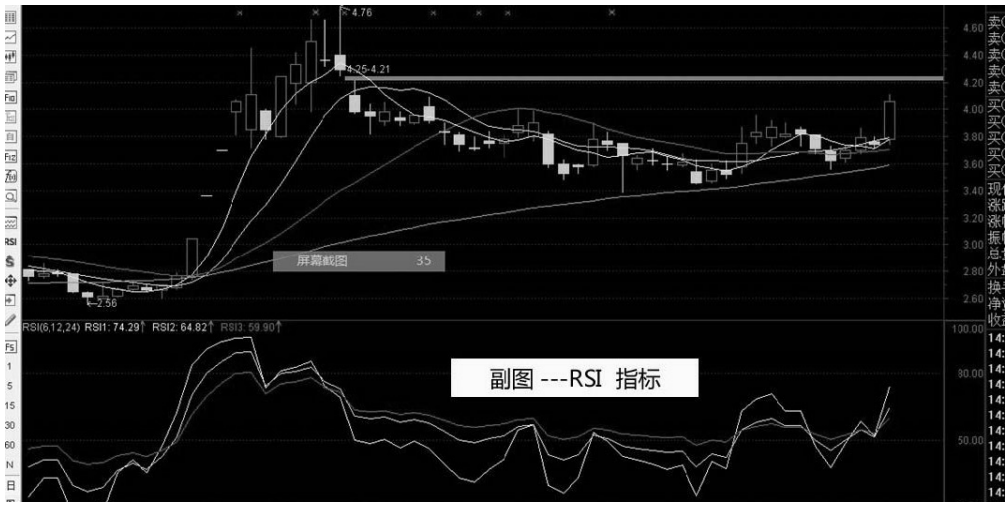


图4-18　RSI策略示意图

RSI指标是根据一定时期内上涨和下跌幅度之和的比率制作出的一种技术曲线，能够反映出市场在一定时期内的景气程度。由于RSI指标实用性很强，因此被多数投资者所喜爱。RSI指标非常适合做短线差价操作。该分析指标的设计是以三条线来反映价格走势的强弱，这种图形可以为投资者提供操作依据。RSI指数比较复杂，包括了交叉、数值、形态和背离等多方面的判断原则。

1. RSI指标

- RSI1一般是6日相对强弱指标。

- RSI2一般是12日相对强弱指标。
- RSI3一般是24日相对强弱指标。

RSI指标的计算公式如下：

RSI=100×RS/(1+RS)或RSI=100－100÷(1+RS)

其中，RS等于X天的平均上涨点数除以X天的平均下跌点数。

一般情况下，RS等于14天内收市价上涨数之和的平均值除以14天内收市价下跌数之和的平均值。

2. RSI指标变动范围

RSI的变动范围是0~100，我国单边做多的股市强弱指标值一般分布在20~80。

- 80~100，极强，卖出。
- 50~80，强，买入。
- 20~50，弱，观望。
- 0~20极弱，买入。

这里的“极强”、“强”、“弱”和“极弱”只是一个相对的分析概念，是一个相对的区域。有的投资者也把它们取值为30、70或15、85，不是绝对的。

3. RSI指标的五种用途

- 顶点及底点。指标值70及30通常为超买及超卖信号。
- 分歧。当市况创下新高（低）但RSI并不处于新高（低）时，通

常表明市场将出现反转。

- 支持及阻力。RSI能显示支持及阻力位，有时比价格图更能清晰反应支持及阻力。
- 价格趋势形态。与价格图相比，价格趋势形态如双顶及头肩在RSI上表现更清晰。
- 峰回路转。当RSI突破（超过前高或低点）时，可能表示价格将有突变，RSI需与其他指标配合使用，不能单独产生信号，价格的确认是决定入市价位的关键。

4.6.1 案例4-12：RSI2策略

本案例代码细节可参考http://stockcharts.com/school/doku.php?id=chart_school:trading_strategies:rsi2。

案例4-12建议使用如下参数：

- 识别SMA（简单移动平均线）趋势称为entrySMA，范围是150~250。
- 一小段SMA出口点称为exitSMA，范围是5~15。
- 一个RSI空头/多头位置称为rsiPeriod，范围是2~10。
- RSI多头超卖阈值称为overSoldThreshold，范围是5~25。
- RSI空头超买阈值称为overBoughtThreshold，范围是75~95。

案例代码参见文件\zwpython\zw_k10\k412_rsi2_sample.py与rsi2.py。

如图4-19所示是RSI策略运行结果。

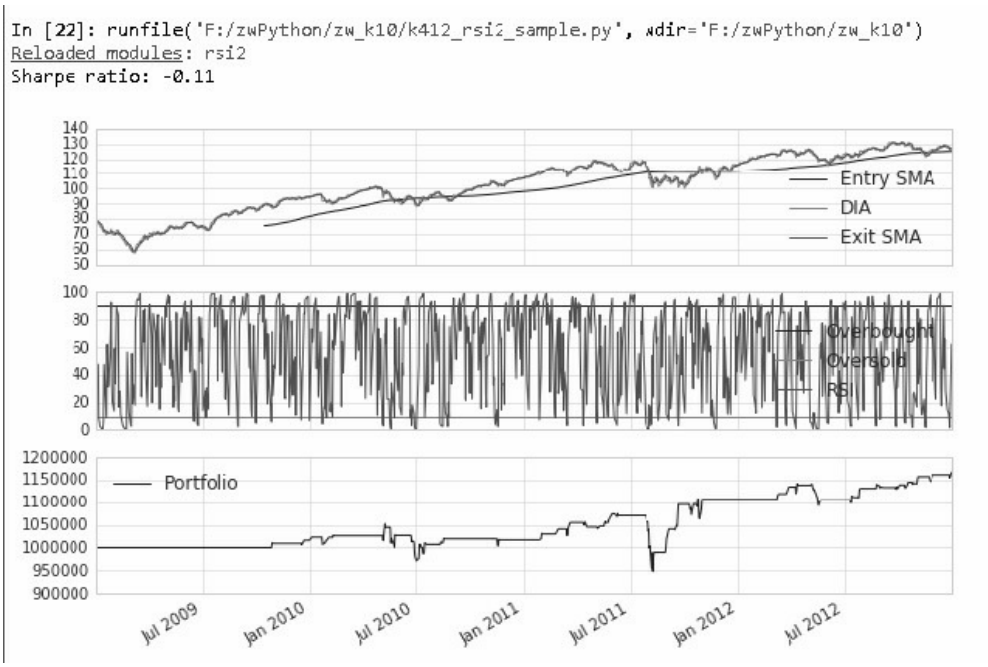


图4-19　RSI策略运行结果

读者可以调整不同的参数重新运行测试。

4.6.2　案例4-13：A股版RSI2策略

案例4-13是案例4-12的A股版本RSI2策略案例，程序源码请参见脚本文件\zwpython\zw_k10\k413_rsi2_zw.py。

如图4-20所示是A股版RSI策略运行结果。

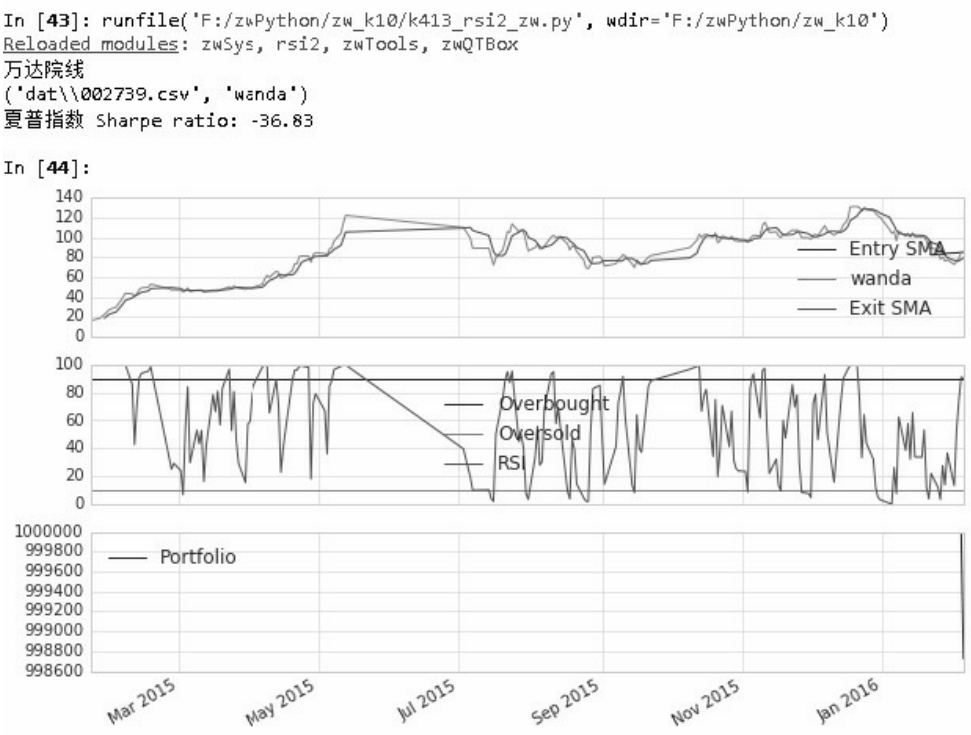


图4-20　A股版RSI策略运行结果

因为“万达院线”是2015年新上市的公司，数据比较少，故最下方的Portfolio资产总值几乎没有图形。

此外，作为本章的最后一个案例，笔者在输出信息方面做了一些优化，不仅可以输出股票的数据文件名，还可以输出股票的拼音：

```
万达院线
('dat\\002739.csv', 'wanda')
```

注意，本章使用了大量PAT量化软件的案例程序，运行环境是Python　2.7，对于中文支持比较弱。要成功输出中文信息，需要修改以下代码：

- 在文件首行增加“# -*- coding: utf-8 -*-”，强制使用utf-8编码；
- 在表示中文信息字符串的变量前加“u”，表示是unicode格式的字符串，如“csgn=u'万达院线'”；
- 输出时，如果采用Python 3函数格式，最好只使用一个中文信息变量print(csgn)，否则会出错。

例如，可以使用：

```
print(csgn,fss)
```

输出的结果中中文信息是乱码：

```
(u'\u4e07\u8fbe\u9662\u7ebf', 'dat\\002739.csv')
```

如果使用Python 2的格式，不用加括号，多个变量也可以成功输出中文信息：

```
print csgn,fss,cname
```

输出结果为：

```
万达院线 dat\002739.csv wanda
```

看来，虽然Python 2.7尽量兼容Python 3的语法，但在某些细节方面与Python 3还是有些差别。这也是笔者推荐读者尽量使用Python 3的原因。

为此，笔者在代码当中特意多准备了几组测试数据，读者可以逐一测试：

```
def main(plot):
     cod="002739";cname='wanda';csgn=u'万达院线'
     #cod="600663";cname='lujiazui';csgn=u'陆家嘴'
     #cod="600231";cname='lingang';csgn=u'凌钢股份'
     #cod='002046';cname='zouyan';csgn=u'轴研科技'
     #cod='300239';cname='donbao';csgn=u'东宝生物'
```

读者也可以按本章的案例程序添加自己喜欢的股票代码数据。运行前，要注意把相关的数据文件复制到dat目录下。

如图4-21~图4-24所示是以上几组数据的运行结果。

```
cod="600663";cname='lujiazui';csgn=u'陆家嘴'
```

```
n [46]: runfile('F:/zwPython/zw_k10/k413_rsi2_zw.py', wdir='F:/zwPython/zw_k10')
Reloaded modules: zwSys, rsi2, zwTools, zwQTBox
陆家嘴 lujiazui dat\600663.csv
夏普指数 Sharpe ratio: -0.20

In [47]:
```

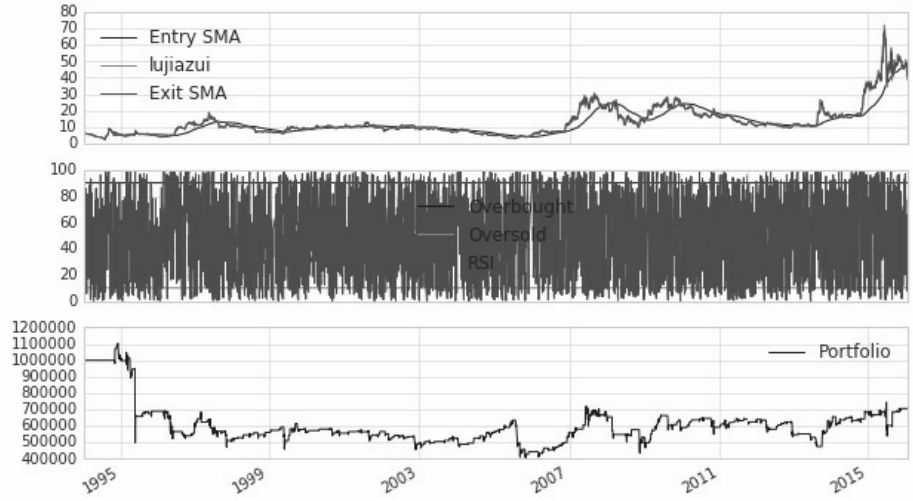


图4-21　A股版RSI策略运行结果：陆家嘴

```
cod="600231";cname='lingang';csgn=u'凌钢股份'
```

```
runfile('F:/zwPython/zw_k10/k413_rsi2_zw.py', wdir='F:/zwPython/zw_k10')
Reloaded modules: zwSys, rsi2, zwTools, zwQTBox
凌钢股份 lingang dat\600231.csv
夏普指数 Sharpe ratio: -0.27
```

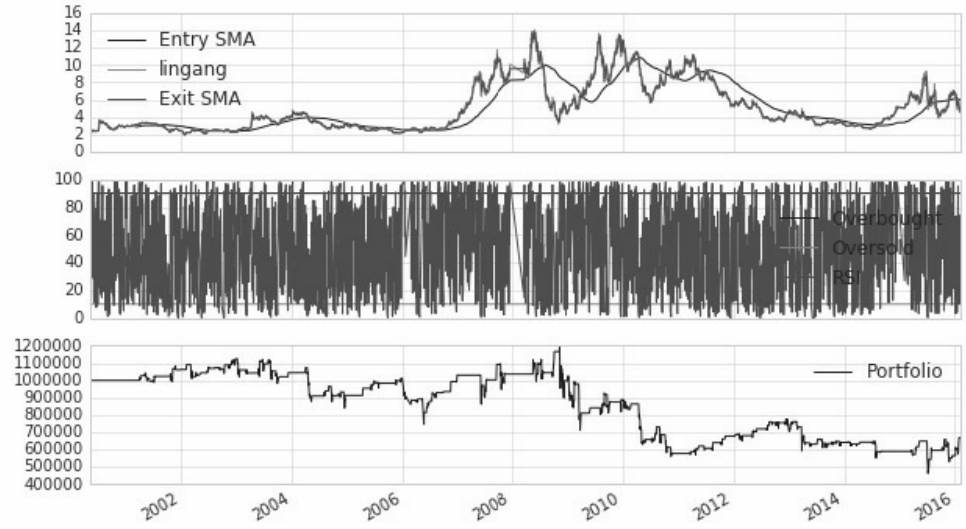


图4-22　A股版RSI策略运行结果：凌钢股份

```
cod='002046';cname='zouyan';csgn=u'轴研科技'
```

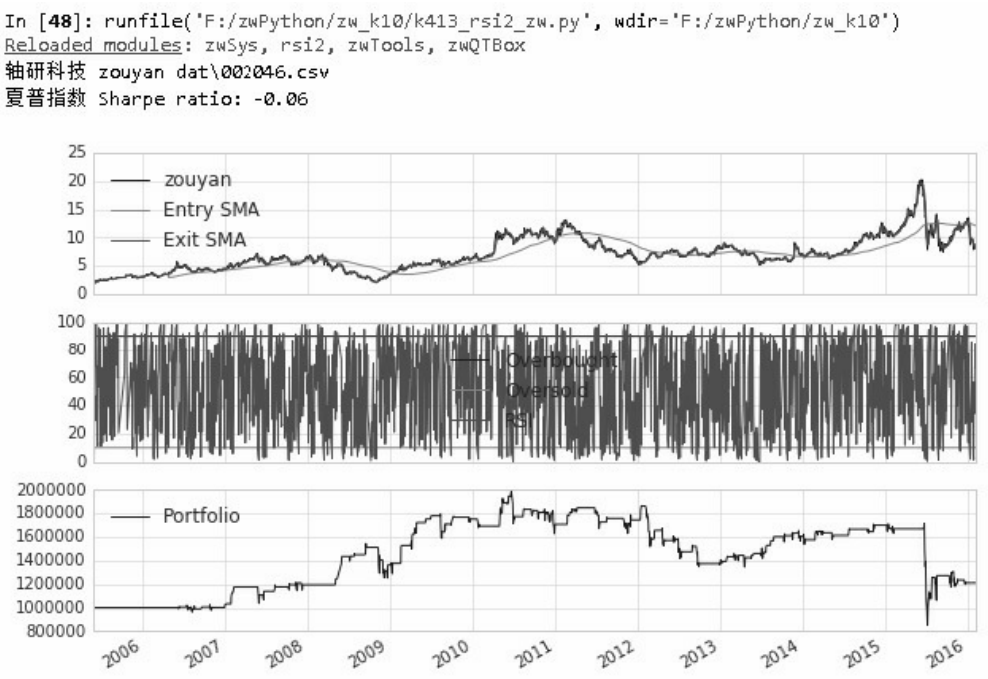


图4-23　A股版RSI策略运行结果：轴研科技

```
cod='300239';cname='donbao';csgn=u'东宝生物'
```

```
In [49]: runfile('F:/zwPython/zw_k10/k413_rsi2_zw.py', wdir='F:/zwPython/zw_k10')
Reloaded modules: zwSys, rsi2, zwTools, zwQTBcx
东宝生物 donbao dat\300239.csv
夏普指数 Sharpe ratio: -0.49

In [50]:
```

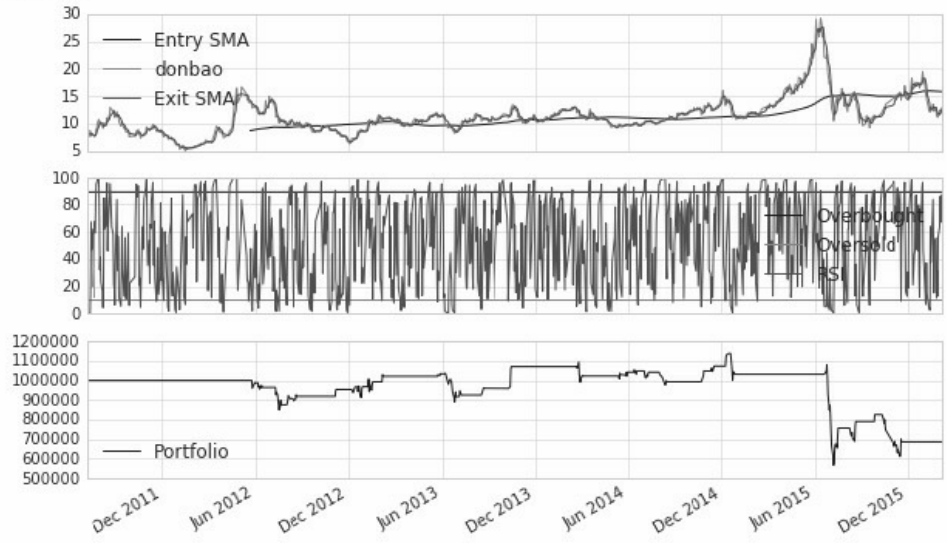


图4-24　A股版RSI策略运行结果：东宝生物

4.7 案例与传承

笔者撰写本书及创办Python量化培训班的原因有很多，不过，其中最重要的一条就是：传承。

Python培训班刚启动时，笔者当时和学员说，笔者原本学的是法律，却在IT、人工智能、中文字模智能设计、网络公关、Python量化分析等几个领域取得了一定的成绩，这说明笔者的学习方法还是有一定的可取之处的。

笔者希望，本书不仅传授Python量化方面的专业知识，更多的是把这种学习方法传授给读者。

这样，读者以后即使进入新的领域，也能够把握方向，在最短的时间内成为行业专家。

本章的案例源自PAT量化作者在2012年学习QSTK课程的作业。

QSTK是2012年美国乔治理工大学计算投资公开课。QSTK之花尚未绽放，便已凋零。QSTK项目因为架构陈旧，被业界淘汰，但是QSTK项目中的精髓却通过传承，通过PAT、quantopian等量化软件被继承下来，而且进一步发扬光大。

赠人玫瑰，手有余香。

开源项目就是这样薪火相传的。

第5章
zwQuant整体架构

5.1 发布前言

这个世界上聪明人太多了，金融行业更是如此。

既然大家都是聪明人，笔者就只能选择充当“愚人”了。于是，便有了zwQuant（简称ZWQT）“愚人版”的发布。

笔者的口号是：只谈原创。令人欣喜的是，zwQuant成功地将这个口号从字库行业“传承”到Python量化领域。

为保证zwQuant量化软件回溯测试数据的准确性，笔者特意与目前业界领先的PythonAlgoTrad量化软件（简称PAT）进行了多场合、多策略的对标测试。对标测试时，全部采用同一数据源和同一量化策略。

在绝大部分对标测试案例当中，核心的测试参数：回报率、夏普指数、回缩周期等方面均取得了一致的测试结果。个别数据略有差异是因为程序架构和数据定义等方面差别导致，但都在可以接受的范围之内。

zwQuant量化软件的特点如下。

- zwQuant率先基于Pandas数据分析软件矩阵模式进行开发，整体

架构简单清晰，全部代码不到1000行，全面涵盖了各个环节，如数据导入、策略分析、回溯分析、绘图统计等。

- zwQuant首创的dataPre数据预处理计算，借助Pandas和Numpy高度优化的运算模块，即使不采用其他优化策略，其整体回溯速度也比目前Python同类量化系统的回溯测试速度快5~10倍。
- 支持单只、多只股票回溯测试，类似于股票池，股票数目上不封顶，仅受硬件和内存限制。
- 量化策略基于传统过程函数，采用1+1模式，即策略分析函数+策略数据预处理函数，其编写难度比目前事件模式简化，并且，支持多策略联动分析。
- 策略函数、绘图输出，均采用准模板编程，用户可以自行扩充，不断增加策略函数库、绘图模板库。
- 支持tick和分时数据，可用于股票实盘、期货外货、贵重金属等项目，仅需修改数据源，不需要修改其他程序即可直接使用。

5.2 功能简介

zwQuant量化软件是一套自主开发的、纯Python中文量化策略分析系统。

zwQuant的正式名称是：极宽·量化策略分析及回溯测试系统开源版，简称：zwQT量化系统。

zwQuant是高度简化版的量化分析、回溯测试系统，只提供策略分析和回溯测试功能，也是一套全功能、全开源、纯Python量化策略分析系统，可以直接用于实盘操作。

数据源方面：配合tushare数据抓取模块、zwDat数据包和内置的开源数据下载程序down_stk可以完成单日A股、美股数据的日交易数据更新和去重。

策略方面：zw_talib开源金融函数库已经发布，大家借助TA-Lib和其他专业金融函数库，可以轻松编写策略，对zwQuant策略库进行扩充。

业务方面：zwQuant量化软件可根据用户需要和硬件运行环境，按天、小时或分钟生成推荐数据，辅助用户人工下单，或采用CSV数据文件导入第三方交易软件下单。

5.2.1 目录结构

zwQuant量化软件的文件目录随版本的变化可能会有所不同，请读者注意。目录结构如图5-1所示。



图5-1　zwQuant目录结构

各层级目录说明如下：

- demo，zwQuant演示程序目录。
- source，zwQuant源码程序。
- down_stk，zwDat股票数据下载程序。
- zwta，zw_talib金融函数库。

注意

zw_talib虽然是一个独立的项目，但在source目录下，也有zw_talib.py的复制文件，且内容相同，如果日后zw_talib升级，用户可以自己复制，覆盖source目录下的同名文件。

zwDat股票数据下载程序，因为源码较多，也是zwQuant的独立衍生项目，所以自成一个目录。

5.2.2 安装与更新

zwQuant极宽量化开源的全部源码，位于目录（x是硬盘盘符）：

```
x:\zwPython\zwQuant\source\
```

zwQuant开源量化软件与zwPython开发平台进行了集成处理，可直接导入使用，支持Python 3.x版本。升级时解压覆盖\zwPython\目录下的zwQuant子目录即可，注意，不要有双重zwQuant子目录。移植时或使用其他Python环境时，可以把zwQuant目录下的脚本文件全部复制到自己的代码工作目录中。

如果采用zwDat作为数据源，zwDat与zwPython必须位于同一个硬盘根目录下。强烈建议使用SSD固态硬盘，可以提高3~5倍运行效率。

5.2.3 模块说明

zwQuant量化软件的主模块和全部源码位于目录（x是硬盘盘符）：

```
x:\zwPython\zwQuant\source\
```

zwQuant量化软件主要有以下模块：

- zwBacktest.py，回溯测试工具函数。
- zwQTBox.py，zw常用“量化”工具函数集。
- zwQTDraw.py，绘图工具函数。
- zwStrategy.py，策略工具函数。
- zwSys.py，zw量化初始化和系统模块，主要定义基础数据类、全局路径、全局变量等。
- zwTools.py，zw常用（非量化）工具函数集。
- zw_talib.py，zw版的TA-Lib函数封装，主要基于Pandas进行移植zwBacktest.py。

为简化架构，除了zwSys.py系统模块以外，因全局数据变量需要定义了三个基本类class以及少量类函数，其他模块都是采用统一的函数定义，整体设计、逻辑清晰，便于用户修改和扩展。为统一起见，zwQuant量化软件的升级版本可能会将有关的类函数改为工具函数。

zwQuant量化软件最常用的三个模块如下。

zwSys.py：zw量化初始化和系统模块，主要定义基础数据类、全局路径、全局变量等。

zwTools.py：zw常用（非量化）工具函数集。

zwQTBox.py：zw常用量化工具函数集。

zwQuant量化软件内置模块默认缩写方式为：

```
import zwSys as zw
import zwTools as zwt
```

```
import zwQTBox as zwx
import zwBacktest as zwbt
import zwQTDraw as zwdr
import zwStrategy as zwsta
import zw_talib as zwta
```

以上为zwQuant量化软件的默认缩写方式，建议大家统一采用以上的缩写方式，以便于沟通学习。

5.2.4 zwSys模块：系统变量与类定义

zwSys是zwQuant量化软件的初始化和系统模块，主要定义基础数据类、全局路径和全局变量等，如图5-2所示。

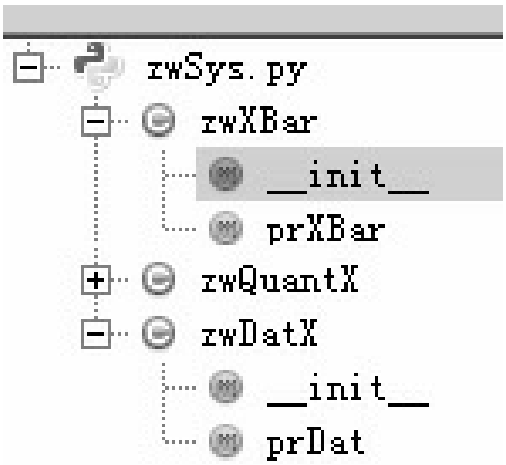


图5-2 zwSys结构图

zwSys模块定义了三个类class：zwXBar、zwDatX、zwQuantX。

其中，zwXBar和zwDatX类很简单，都只有一个类函数：

- zwXBar为数据包类，Bar是量化交易的传统数据单位，记录每笔交易；类函数prXBar输出当前Bar变量的数据。
- zwDatX设置各个数据目录；类函数prDat输出其他相关的目录信息。

zwSys模块相对复杂的是类zwQuantX，其定义了zwQuant量化交易所需要的各种变量参数以及相关的类函数，如图5-3所示。

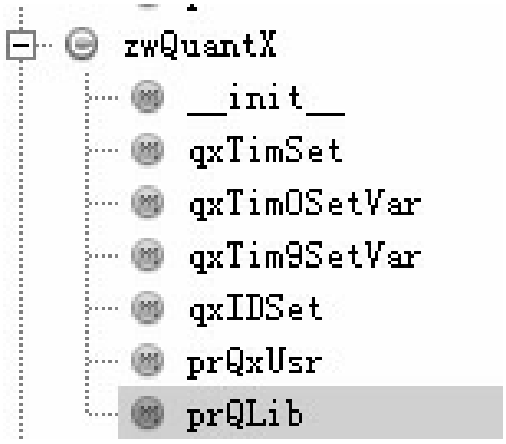


图5-3　zwQuantX结构图

类zwQuantX定义的函数包括：

- qxTimSet，设置时间参数。
- qxTim0SetVar，回溯测试时间点开始，初始化相关参数。
- qxTim9SetVar，回溯测试时间点结束，整理相关数据。
- qxIDSet，生成订单流水号编码ID。
- prQxUsr，输出用户变量保存的数据。
- prQLib，输出各种回溯交易测试数据，一般用于结束时。

5.2.5 zwTools模块：常用（非量化）工具函数

zwTools工具模块内置函数说明如下。

- iff2(kflag,x1,x0)：二选一函数，如果kflag为True，则返回值是x1；否则，返回值是x0。
- iff3(v,k,xn1,x0,x1)：三选一函数，如果v<k，则返回值是xn1；v=k，返回值是x0；v>k，返回值是x1。
- listPr(lst)：输出列表信息。
- listRd(fnam)：读取列表数据。
- listWr(fnam,lst)：保存列表数据。
- lst4dir(rss)：目录文件生成列表数据。
- wait(n,mstr='')：等待n秒。
- xdebug(xmod,mnam,fnam)：输出调试信息。
- xin(xk,k0sgn,k9sgn)：如果xk位于(x0sgn, x9sgn)间，不含等于，则返回True；否则，返回值是x0。
- xinEQ(d,k0,k9)：如果xk位于(x0sgn, x9sgn)间，包含等于，则返回True；否则，返回值是x0；
- xobj2str(xobj,xnamLst)：对象属性指根据属性列表，生成字符串。

5.2.6 zwQTBox：常用“量化”工具函数集

zwQTBox是量化函数模块，定义的函数如下。

- cross_Mod(qx,ksma)：均线交叉策略，判断均线向上、向下趋势。
- df2cnstk(df0)：股票数据格式转换，转换为中国A股格式。
- df2yhaoo(df0)：股票数据格式转换，转换为Yahoo格式。
- df2zw(df0)：股票数据格式转换，转换为zw格式。
- df2zwAdj(df0)：股票数据格式转换，转换为zw增强版格式，带adj close。
- downKMax(dlow,dhigh)：计算最大回缩率。
- down_stk_cn010(qx,xtyp="D")：中国A股数据下载子程序。
- down_stk_yahoo010(qx,ftg)：Yahoo美股数据下载子程序。
- qxObjSet(xtim,stkVal,dcash,dret)：设置xtrdLib单次交易节点数据。
- sharpe_rate(rets,rfRate,ntim=252)：计算夏普指数。
- stkGetPrice(qx,ksgn)：获取当前价格。
- stkGetPrice9x(qx,ksgn)：获取首个、末个交易日数据。
- stkGetVars(qx,ksgn)：获取股票代码，指定字段的数据。
- stkLibGetTimX(xcod)：返回指定股票代码首个、末个交易日时间数据。
- stkLibName8Code(xcod)：根据股票代码，返回股票中文、英文/拼音缩写名称。
- stkLibPr()：输出股票数据。
- stkLibRd(xlst,rdir)：读取指定的股票数据，可多只股票。
- stkLibSet8XTim(dtim0,dtim9)：根据时间段，切割股票数据。
- stkValCalc(qx,xdicts)：计算xdicts内所有的股票总价值。
- stk_col_renLow(dat)：股票数据格式转换，转换小写列名称。

- stk_copy_OHLC(dat0)：复制股票OHLC数据。
- xbarGet8Tim(xcod,xtim)：根据指定股票代码和时间获取数据包。
- xbarGet8TimExt(xcod,xtim)：根据指定股票代码和时间获取数据包及股票数据。
- xbarPr(bars)：输出数据包数据。
- xusrStkNum(qx,xcod)：返回用户持有的xcod股票数目。
- xusrUpdate(qx)：更新用户数据。
- xusr4xtrd(qx,b2)：根据交易数据，更新用户数据qxUsr。
- xedit_zwXDat(df)：编辑用户数据格式。
- xtrdChkFlag(qx)：检查是不是有效交易。
- xtrdLibAdd(qx)：添加交易到xtrdLib。
- xtrdObjSet(qx)：设置交易节点数据。
- zwRetPr(qx)：输出、计算回报率。
- zwRetTradeCalc(qx)：输出、计算交易数据。

5.2.7 zwQTDraw.py：量化绘图工具函数

zwQTDraw绘图模块库很简单，采用的是1+1模式，即1个函数设置绘图环境加1个函数生成绘图。目前只有一个三合一的绘图模板：Quant 3x，用户可以参考相关代码，自行添加。

- dr_quant3x_init(qx,w,h)：初始化绘图环境，w、h是图形大小尺寸。
- dr_quant3x(qx,ktop2,kbot,kmidlst,midSgn0='')：绘制图形。

- dr_quant3x(qx,ktop2,kbot,kmidlst,midSgn0='')函数，3x三合一，回溯测试绘图函数。

输入参数如下。

- qx：全局量化参数变量。
- ktop2：Top顶部成交量股票代码。
- kbot：Bot底部绘图列名称，一般是'val'，资产总价值；数据源为qx.qxLib。
- kmidlst：Mid中部绘图列名称、列表；子列表元素1为股票代码xcod，其他列名称格式为[[xcod1,nam1,nam2,...],[xcod2,nam1,nam2,...],[xcod3,nam1,nam2,...]]。注意，kmidlst数据源为stkLib[xcod]，包含预处理扩充的数据列。
- midSgn0：中部绘图区图标前缀，其中“<xcod>”为特殊符号，表示对应的股票代码。

5.2.8 zwBacktest：回溯测试工具函数

zwBacktest是回溯测试模块，其目前很简单，只有三个脚本函数，未来会根据需要增加不同时间周期、不同数据源的回溯测试函数。

函数说明如下。

- bt_init：设置回溯测试参数。
- zwBackTest100(qx)：回溯测试函数，是单一股票、单一时间点的

测试。

- zwBackTest(qx)：回溯测试主程序。

5.2.9 zwStrategy：策略工具函数

zwStrategy量化策略模块内置的分析函数采用1+1模式，即每个策略有两个函数：一个是策略数据初始化函数，另一个是策略分析函数。这种1+1策略函数组模式是zwQuant量化软件独创的，具有策略编写简单、逻辑清晰、运行效率极高的特点。用户可参考相关源码，自行编写、扩充策略函数库。

zwStrategy量化策略模块目前有以下策略函数。

- sta_dataPre0xtim(qx,xnam0)：策略参数设置子函数。
- BBANDS_dataPre(qx,xnam0,ksgn0)：布林带数据预处理函数。
- BBANDS_sta(qx)：布林带策略分析函数。
- CMA_dataPre(qx,xnam0,ksgn0)：均线交叉策略数据预处理函数。
- CMA_sta(qx)：均线交叉策略分析函数。
- SMA_dataPre(qx,xnam0,ksgn0)：简单均线策略数据预处理函数。
- SMA_sta(qx)：简单均线策略分析函数。
- VWAP_dataPre(qx,xnam0,ksgn0)：VWAP数据预处理函数。
- VWAP_sta(qx)：VWAP成交量加权平均价策略分析函数。

5.2.10 zw_TA-Lib：金融函数模块

zw_TA-Lib金融函数库是一个独立开源项目，基于Pandas数据分析软件的talib函数封装，属于zwQaunt量化项目的衍生项目。zw_TA-Lib金融函数库无须安装zwQaunt量化软件、模块和案例程序即可独立运行。

zwTA-Lib首批函数名称如下。

- ACCDIST(df, n)：集散指标（A/D），其英文全称是Accumulation/Distribution，是由价格和成交量的变化而决定的。
- ADX(df, n, n_ADX)：平均趋向指数，ADX指数是反映趋向变动的程度，而不是方向的本身；英文全称是Average Directional Index或者Average Directional Movement Index。
- ATR(df, n)：均幅指标（Average True Ranger），取一定时间周期内的股价波动幅度的移动平均值，主要用于研判买卖时机。
- BBANDS(df, n)：布林带，英文全称：Bollinger Bands。
- BBANDS_UpLow(df, n)：笔者改进版的布林带TA-Lib函数。
- CCI(df, n)：顺势指标（Commodity Channel Index），是由美国股市分析家唐纳德·蓝伯特（Donald Lambert）所创造的，是一种重点研判股价偏离度的股市分析工具。
- COPP(df, n)：估波指标（Coppock Curve），又称“估波曲线”，通过计算月度价格的变化速率的加权平均值来测量市场的动量，属于长线指标。估波指标主要用于判断牛市的到来，该指标用于研判大盘指数较为可靠，一般较少用于个股。
- Chaikin(df)：佳庆指标（Chaikin Oscillator），是由马可·蔡金（Marc Chaikin）提出来的，是集散指标（A/D）的改良版本。
- DONCH(df, n)：奇安通道指标（Donchian Channel），该指标是由Richard Donchian发明的，由3条不同颜色的曲线组成，该指标用周期内的最高价和最低价来显示市场的波动性；当其通道窄时

表示市场波动较小，通道宽则表示市场波动比较大。

- EMA(df, n)：指数平均数指标（Exponential Moving Average，EXPMA或EMA），指数平均数指标也叫EXPMA指标，也是一种趋向类指标，其构造原理是仍然对收盘价进行算术平均，并根据计算结果进行分析，用于判断价格未来走势的变动趋势。
- EOM(df, n)：简易波动指标（Ease of Movement Value），又称EMV指标，是根据等量图和压缩图的原理设计而成，目的是将价格与成交量的变化结合成一个波动指标来反映股价或指数的变动状况。由于股价的变化和成交量的变化都可以引发该指标数值的变动，因此，EMV实际上也是一个量价合成指标。
- FORCE(df, n)：劲道指数（Force Index）；劲道指数是衡量每个涨势中的多头劲道与每个跌势中的空头劲道。劲道指数结合三项主要的市场资讯，即价格变动的方向、幅度与成交量。它是从一个崭新而实用的角度，把成交量纳入交易决策中。
- KELCH(df, n)：肯特纳通道（Keltner Channel，KC），是一个移动平均通道，由三条线组合而成（上通道、中通道及下通道）。一般情况下，以上通道线和下通道线的分界作为买卖的最大可能性。若股价在边界出现不正常的波动，即表示出现买卖机会。
- KST(df, r1, r2, r3, r4, n1, n2, n3, n4)：确然指标（KST），又称为完定指标，该指标参考长、中、短期的变速率ROC，以了解不同时间循环对市场的影响。该指标将数个周期的价格变动率函数做加权，再平滑绘制长短曲线，其特色是通过修正价格变动组合来判断趋势，精准掌握转折买卖点。
- KST4(df, r1, r2, r3, r4, n1, n2, n3, n4)：笔者修订版KST确然指标。
- MA(df, n)：移动平均线（Moving Average），就是最常用的均线指标。
- MACD(df, n_fast, n_slow)：指数平滑移动平均线，是由一快、一

慢指数移动平均（EMA）之间的差计算出来。“快”指短时期的EMA，而“慢”则指长时期的EMA，最常用的是12日及26日EMA。

- MFI(df, n)：资金流量指标和比率（Money Flow Index and Ratio），资金流量指标又称为量相对强弱指标（Volume Relative Strength Index，VRSI）；根据成交量来计算市场供需关系和买卖力道。该指标是通过反映股价变动的四个元素（上涨的天数、下跌的天数、成交量增加幅度、成交量减少幅度）来研判量能的趋势，预测市场供求关系和买卖力道，属于量能反趋向指标。
- MOM(df, n)：动量线，英文全称是Momentum，简称MOM。动量可以视为一段期间内，股价涨跌变动的比率。
- MassI(df)：梅斯线（Mass Index），是累积股价波幅宽度之后所设计的震荡曲线。其最主要的作用在于寻找飙涨股或者极度弱势股的重要趋势反转点。MASS指标是所有区间震荡指标中风险系数最小的一个。
- OBV(df, n)：能量潮指标（On Balance Volume，OBV），是葛兰维（Joe Granville）于20世纪60年代提出的，并被广泛使用。股市技术分析的四大要素：价、量、时、空。OBV指标就是从“量”这个要素作为突破口，来发现热门股票、分析股价运动趋势的一种技术指标。它是将成交量与股价的关系数字化、直观化，以股市的成交量变化来衡量股市的推动力，从而研判股价的走势。关于成交量方面的研究，OBV能量潮指标是一种相当重要的分析指标之一。
- PPSR(df)：支点、支撑线和阻力线（Pivot Points，Supports and Resistances），PIVOT指标的观念很简单，不需要计算任何东西，纯粹只是一个分析反转点的方法而已。PIVOT是指“轴心”，轴心是用来确认反转的基准，所以PIVOT指标其实就是找轴心的

方法；PIVOT指标经常被与布林带数据一起分析。

- ROC(df, n)：变动率（Rate of change，ROC），是由当天的股价与一定的天数之前某一天的股价比较，其变动速度的大小反映股票市场变动的快慢程度。ROC也叫作变动速度指标、变动率指标或变化速率指标。
- RSI(df, n)：RSI，相对强弱指标（Relative Strength Index），也称相对强弱指数、相对力度指数；RSI是通过比较一段时期内的平均收盘涨数和平均收盘跌数来分析市场买沽盘的意向和实力，从而做出未来市场的走势。RSI通过特定时期内股价的变动情况计算市场买卖力量对比，来判断股票价格内部的本质强弱、推测价格未来的变动方向。
- RSI100(df, n)：笔者修改版的RSI相对强弱指数，取0~100之间的数值。
- STDDEV(df, n)：标准偏差（Standard Deviation）。
- STOD(df, n)：随机指标D值（Stochastic oscillator %D），随机指标又称为KD指标或KDJ指标；随机指标综合了动量观念、强弱指标及移动平均线的优点，用来度量股价脱离价格正常范围的变异程度。
- STOK(df)：随机指标K值（Stochastic oscillator %K）。
- TRIX(df, n)：TRIX指标又叫作三重指数平滑移动平均指标（Triple Exponentially Smoothed Average）。
- TSI(df, r, s)：真实强度指数（True Strength Index），是相对强弱指数（RSI）的变体。TSI使用价格动量的双重平滑指数移动平均线，剔除价格的震荡变化并发现趋势的变化。r一般取值25，s一般取值13。
- ULTOSC(df)：终极指标（Ultimate Oscillator）。先找出三个周期不同的振荡指标，再将这些周期参数按照反比例的方式制作成常

数因子。然后，依照加权的方式将三个周期不同的振荡指标分别乘以不同比例的常数，综合制作成UOS指标。经过一连串参数顺化的过程后，UOS指标比一般单一参数的振荡指标更能够顺应各种不同的市场状况。

- Vortex(df, n)：螺旋指标（Vortex Indicator）。

5.3 示例程序

为了方便读者学习zwQuant量化软件，特意增设了部分示例程序，这些案例大部分源自PAT的对标测试；所有内置的示例程序，都保存在目录：

```
x:\zwPython\zwQuant\demo
```

在内置的案例程序中，zwQuant量化软件有关的案例程序全部以字母zq开头；PAT量化软件内置的示例程序以字母k开头，其他的文件辅助子程序与数据文件。

笔者对部分源自PAT量化软件的案例进行了部分修改，主要是把数据源改为国内A股数据源，并且增加了绘图输出和结果统计，运行结果与PAT量化软件原版案例程序一样。

不是使用zwPython集成版开发平台的读者，或使用其他Python开发环境的读者，在运行本书案例程序之前，请把zwQuant所有程序：

```
x:\zwPython\zwQuant\source
```

复制到demo目录下，并自行安装所需要的模块库。

5.4 常用量化分析参数

在本书前面的章节当中，我们已经学过一些常见的量化分析参数，比如，最终资产价值、年收益、平均日收益率、日收益率方差、夏普指数、最大回撤率、最大回撤时间等。这些参数的计算和理论对于初学者而言有不少门槛，而且计算方式也不是简单的一两个函数，有些需要进行连续的迭代结算。zwQuantBox.py量化工具箱已经收录了以上常用量化参数等工具函数，大家运行回溯时，系统会自动调用相关的函数。

最终资产价值、年收益率、平均日收益率、日收益率方差几个参数相对比较简单，计算公式分别如下。

最终资产价值=持有现金+股票数目×结算日收盘价

年收益率=（年最终资产价值-初始投入）/初始投入

日收益率=（日最终资产价值-（前一日）最终资产价值）/（前一日）最终资产价值

平均日收益率=mean（日收益率列表）

日收益率方差=std_dev（日收益率列表）

以上参数是衡量一个投资策略好坏的最基本的参数，也是最重要的参数，特别是收益率。

需注意的是，以上参数的计算都非常简单，就是最基本的四则运算。平均日收益率和日收益率方差略微复杂一些，涉及方差计算，不过

有现成的公式函数，读者直接套用就可以了。

本章只是概述，介绍了部分对标测试的结果数据，对相关案例程序代码进行了省略，后面章节再具体讲解。

5.5 回溯案例：对标测试

为保证zwQuant量化软件回溯测试数据的准确性，zwQuant与目前业界领先的PythonAlgoTrad进行了多场合、多策略的对标测试。对标测试时，全部采用同一数据源、同一量化策略。

在绝大部分案例中，在回报率、夏普指数、回缩周期等方面均取得了一致的测试结果。个别数据略有差异是因为程序架构和数据定义等方面的差别所导致，但都在可以接受的范围之内。

5.5.1 对标测试1：投资回报参数

PAT量化软件版的“投资回报参数”案例，运行结果是：

```
DEBUG:broker.backtesting:Not enough volume to fill egan market o
2011-01-03 00:00:00 broker.backtesting [DEBUG] Not enough volume
DEBUG:broker.backtesting:Not enough volume to fill aeti market o
2011-01-03 00:00:00 broker.backtesting [DEBUG] Not enough volume

最终资产价值 Final portfolio value: $1524087.19
年收益 Anual return: 52.41 %
平均日收益率 Average daily return: 0.17 %
日收益率方差 Std. dev. daily return: 0.0122
```

```
夏普指数 Sharpe ratio: 2.28
最大回撤率 Max. drawdown: 11.47 %
最长回撤时间 Longest drawdown duration: 82 days, 0:00:00


2470.005510
2480.001400
249 -0.017165
2500.010841
251 -0.004488
dtype: float64
tmp\dret010.csv
```

A股版投资回报参数案例，运行结果如图5-4所示。

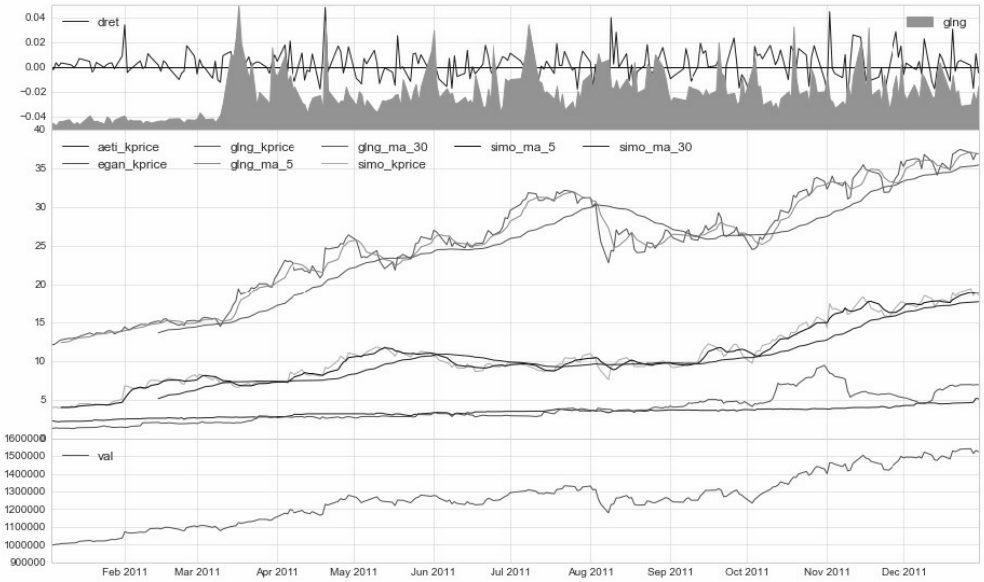


图5-4　A股版投资回报参数

A股版投资回报参数案例其他输出信息有如下：

交易总次数：2
交易总盈利：524087.19

盈利交易数：2
盈利交易金额：524087.19
亏损交易数：0
亏损交易金额：0.00

最终资产价值 Final portfolio value: $1524087.19
最终现金资产价值 Final cash portfolio value: $795739.26
最终证券资产价值 Final stock portfolio value: $728347.93
累计回报率 Cumulative returns: 52.41 %
平均日收益率 Average daily return: 0.175 %
日收益率方差 Std. dev. daily return:0.0121

夏普比率 Sharpe ratio: 2.020,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 2.279,（0利率）

最大回撤率 Max. drawdown: 11.4729 %
最长回撤时间 Longest drawdown duration: 82
回撤时间(最高点位) Time High. drawdown: 2011-12-27
回撤最高点位 High. drawdown: 1540988.216
回撤最低点位 Low. drawdown: 1514537.859

时间周期 Date lenght: 362 (Day)

```
时间周期（交易日） Date lenght(weekday): 252 (Day)
开始时间 Date begin: 2011-01-03
结束时间 Date lenght: 2011-12-30

项目名称 Project name: inv01
策略名称 Strategy name: tim0Trade
```

由以上的对比测试数据可以看出，zwQuant量化软件与PAT量化软件的回溯数据完全一致。

5.5.2 对标测试2：VWAP策略

下面，我们再看一组略为复杂的对标测试案例：VWAP策略，其又叫作成交量加权平均价策略，主要用于机构和庄家大资金进货、出货的操作。不过，目前也有人用于作为小资本的量化投资策略。

PAT版的VWAP示例源码，运行结果如图5-5所示。

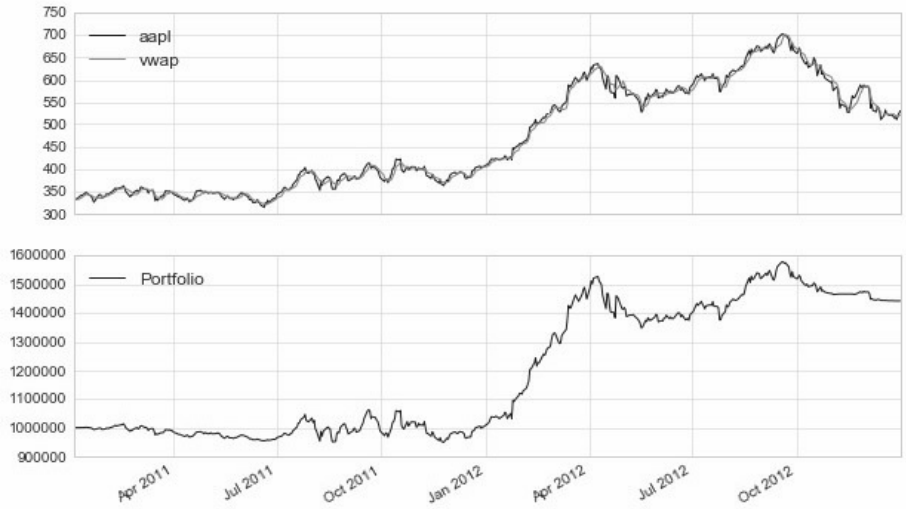


图5-5　PAT股版VWAP策略

PAT股版VWAP策略的其他输出信息如下：

```
最终资产价值 Final portfolio value: $1441776.00
累计回报率 Cumulative returns: 44.18 %
夏普比率 Sharpe ratio: 0.89
最大回撤率 Max. drawdown: 11.54 %
最长回撤时间 Longest drawdown duration: 153 days, 0:00:00
4970.000058
498 -0.000497
4990.000037
5000.000000
5010.000000
dtype: float64
tmp\k408vwap_dret.csv
```

A股版的VWAP案例程序，运行结果如5-6所示。

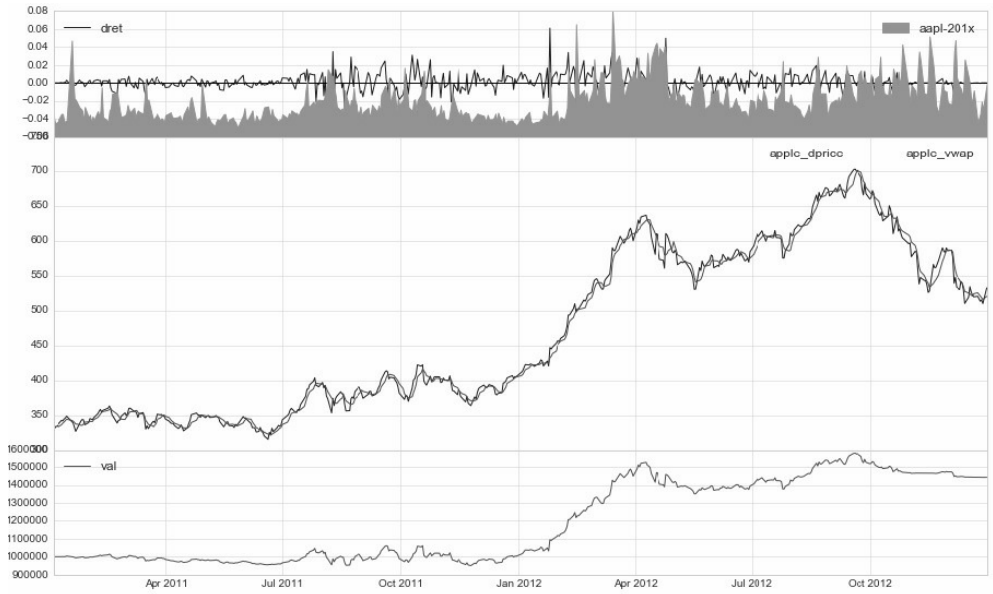


图5-6　A股版VWAP策略

A股版VWAP策略的其他输出信息如下：

```
交易总次数：234
交易总盈利：441776.00

盈利交易数：137
盈利交易金额：1706656.98
亏损交易数：97
亏损交易金额：-1264880.98

最终资产价值 Final portfolio value: $1441776.00
最终现金资产价值 Final cash portfolio value: $1441776.00
最终证券资产价值 Final stock portfolio value: $0.00
```

```
累计回报率 Cumulative returns: 44.18 %
平均日收益率 Average daily return: 0.078 %
日收益率方差 Std. dev. daily return:0.0103

夏普比率 Sharpe ratio: 0.895,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 1.199,（0利率）

最大回撤率 Max. drawdown: 11.5174 %
最长回撤时间 Longest drawdown duration: 153
回撤时间（最高点位）Time High. drawdown: 2012-09-19
回撤最高点位 High. drawdown: 1576163.027
回撤最低点位 Low. drawdown: 1441776.005

时间周期 Date lenght: 729 (Day)
时间周期（交易日） Date lenght(weekday): 502 (Day)
开始时间 Date begin: 2011-01-03
结束时间 Date lenght: 2012-12-31

项目名称 Project name: vwap
策略名称 Strategy name: vwap
```

对比图5-5和图5-6，再对比两个VWAP策略的输出信息，数据完全一致。

5.6　回报参数计算

下面介绍一下zwQuant量化软件中相关的回报参数计算函数。需要说明的是，这些函数的原理和编程对于初学者而言有一些难度，不要求大家全部掌握。本章介绍的相关函数，目的只是告诉大家函数的大概构成，知道相关参数的流程就可以了。对于其中具体的细节和原理，等大家日后实盘、编程水平提高之后，再深入学习。

以下函数如无特别注明，源码均位于模块文件：

```
zwQuant\source\ zwQTBox.py
```

1. 最大回撤率Max. drawdown

最大回撤率：在选定周期内任一历史时点往后推，产品净值走到最低点时的收益率回撤幅度的最大值。最大回撤用来描述买入产品后可能出现的最糟糕的情况，是一个重要的风险指标，对于对冲基金和数量化策略交易，该指标比波动率还重要。

相关代码如下：

```
def downKMax(dlow,dhigh):
    '''
    downKMax(dlow,dhigh):
```

```
    回缩率计算函数
    低位、高位指的是投资组合的市场总值
【输入】：
dlow，当前的低位，低水位，也称lowerWatermark
dhigh，当前的高位，高水位，也称highWatermark
【输出】
回缩率，百分比数值
'''
if dhigh>0:
    kmax=(dlow-dhigh)/float(dhigh)*100
else:
    kmax=0

return kmax
```

2. 夏普指数sharpe_rate

夏普指数（Sharpe Ratio）又称夏普比率，是经过风险调整后的绩效指标。夏普指数反映了单位风险基金净值增长率超过无风险收益率的程度。如果夏普比率为正值，说明在衡量期内基金的平均净值增长率超过了无风险利率，在以同期银行存款利率作为无风险利率的情况下，说明投资基金比银行存款要好。夏普比率越大，说明基金单位风险所获得的风险回报越高。反之，则说明在衡量期内基金的平均净值增长率低于无风险利率，在以同期银行存款利率作为无风险利率的情况下，说明投资基金比银行存款要差，基金的投资表现不如国债回购。而且当夏普比率为负时，按大小排序没有意义。（摘自《百度百科》）

夏普指数有几种计算方式，下面介绍其中的一种。

```
def sharpe_rate(rets,rfRate,ntim=252):
    """
    sharpe_rate(rets,rfRate,ntim=252):
        计算夏普指数
    【输入】
     ret，收益率数组（根据ntim，按日、小时保存）
     rfRate，无风险收益利润
     ntim，交易时间（按天、小时等计数）
         采用日线数据，ntim= 252.
         采用小时（60分钟）线数据，ntim= 252* 6.5 = 1638.
    【输出】
        夏普指数数值
    """
    rsharp= 0.0
    if len(rets):
        # print('rets',rets)
        rstd = np.array(rets).std(ddof=1)#np.stddev(rets, 1)#收
        #print('rstd',rstd,rets[-1]) #,returns[]
        if rstd != 0:
            rfPerRet = rfRate / float(ntim)
            rmean=np.array(rets).mean()
            avgExRet = rmean - rfPerRet
            dsharp = avgExRet / rstd
            rsharp = dsharp * np.sqrt(ntim)
```

```
            #print('rmean,avgExRet,dsharp',rmean,avgExRet)
            #print('dsharp,rshapr',dsharp,rsharp)
    return rsharp
```

夏普指数函数sharpe_rate流程图如图5-7所示。

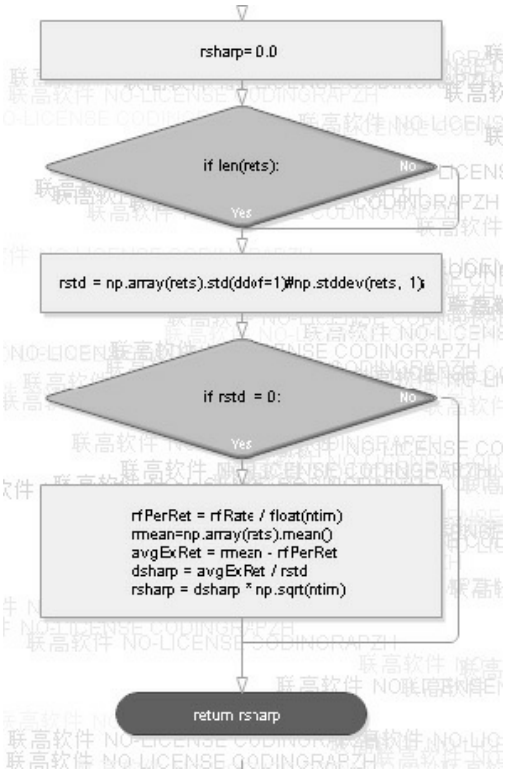


图5-7　夏普指数函数sharpe_rate流程图

3. 交易回报计算

这是回溯测试结束后总体的交易回报计算。不同的系统采用的参数和指标有所不同，但总体的回报率大体相同。

```
def zwRetTradeCalc(qx):
```

```
    trdNum=len(qx.xtrdLib);
    #print('trdNum',trdNum)
    #
    numAdd=0;numDec=0;
    sumAdd=0;sumDec=0;
    sumPut=0;sumGet=0;
    xbar=qx.qxLib.iloc[-1];
    xtim9=xbar['date']
    for xc in range(trdNum):
        xbar=qx.xtrdLib.iloc[xc];

        #print(xbar)
        #kprice=xbar['kprice']
        dnum=xbar['num']
        #print('dprice',dprice)
        qx.stkCode=xbar['code']
        #qx.xtim=xbar['date']
        price=xbar['kprice']
        #ksgn='dprice';

        #price=stkGetPrice(qx,ksgn);
        qx.xtim=xtim9;ksgn=qx.priceCalc;
        price0,price9=stkGetPrice9x(qx,ksgn);
        #print(qx.stkCode,dprice,'$',dprice0,dprice9)
        #---
        #if dnum>0:sumPut=sumPut+price*dnum;
        #if dnum<0:sumGet=sumGet+price9*dnum;:
```

```
        #
        dsum=dnum*(price9-price);
        #料敌从宽，平局（收益率为零时），考虑到交易成本，作为亏损处理，
        if dsum>0:
            numAdd+=1;sumAdd=sumAdd+dsum;
        else:
            numDec+=1;sumDec=sumDec+dsum;



        #print('trdNum',trdNum,numAdd,numDec,'$',dprice,dprice9
    #---------
    sum9=sumAdd+sumDec;#sumx=sumPut+sumGet;
    print('交易总次数：%d' %trdNum)
    print('交易总盈利：%.2f' %sum9)
    #print('交易总支出：%.2f' %sumPut)
    #print('交易总收入：%.2f' %sumGet)
    #print('交易收益差：%.2f' %sumx)
    print('')
    print('盈利交易数：%d' %numAdd)
    print('盈利交易金额：%.2f' %sumAdd)
    print('亏损交易数：%d' %numDec)
    print('亏损交易金额：%.2f' %sumDec)
    #print('@t',qx.xtim)
    qx.xtim=xtim9
```

回报率计算函数zwRetTradeCalc代码看起来很长，其实运行逻辑很

简单，如图5-8所示为函数流程图。

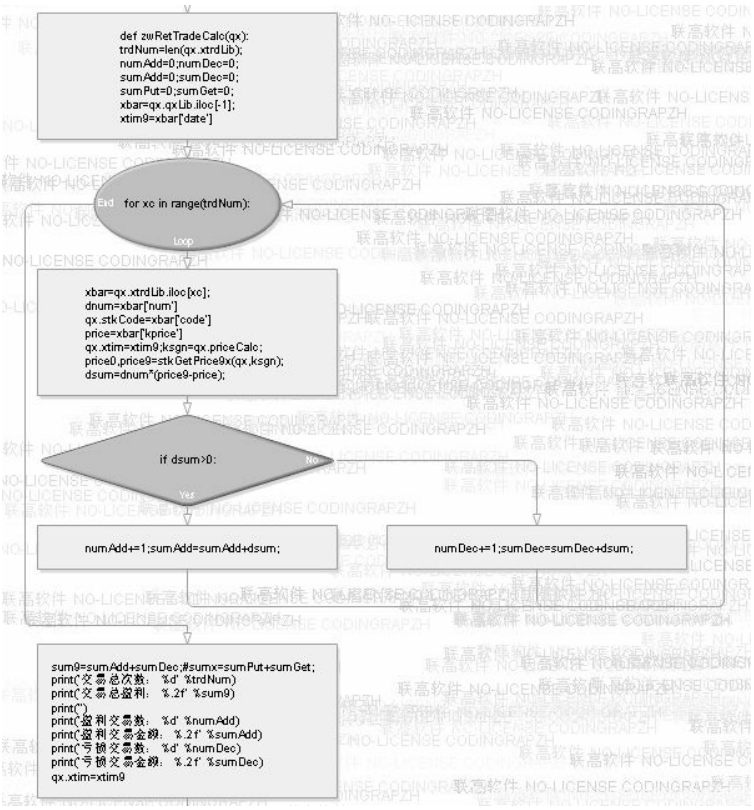


图5-8　回报率计算函数zwRetTradeCal流程图

4. 资产回报分析

这是回溯测试结束后总体的资产回报计算。不同的系统采用的参数指标有所不同，但总体的回报率大体相同。

```
def zwRetPr(qx):
    #---回报测试
    retAvg=np.mean(qx.qxLib['dret']);
    retStd=np.std(qx.qxLib['dret']);
    dsharp=sharpe_rate(qx.qxLib['dret'],qx.rfRate)
```

```
dsharp0=sharpe_rate(qx.qxLib['dret'],0)
dcash=qx.qxUsr['cash'];
dstk=stkValCalc(qx,qx.xbarUsr);
dval=dstk+dcash;
dret9=(dval-qx.mbase)/qx.mbase

print('')
print("最终资产价值 Final portfolio value: $%.2f" % dval)
print("最终现金资产价值 Final cash portfolio value: $%.2f" % c
print("最终证券资产价值 Final stock portfolio value: $%.2f" %
print("累计回报率 Cumulative returns: %.2f %%" % (dret9*100))
print("平均日收益率 Average daily return: %.3f %%" %(retAvg*1
print("日收益率方差 Std. dev. daily return:%.4f " %(retStd))
print('')
print("夏普比率 Sharpe ratio: %.3f,（%.2f利率）" % (dsharp,qx.
print("无风险利率 Risk Free Rate: %.2f" % (qx.rfRate))
print("夏普比率 Sharpe ratio: %.3f,（0利率）" % (dsharp0))
print('')
print("最大回撤率 Max. drawdown: %.4f %%" %(abs(qx.downKMax))
print("最长回撤时间 Longest drawdown duration:% d" %qx.downMa
print("回撤时间(最高点位) Time High. drawdown: " ,qx.downHighT
print("回撤最高点位 High. drawdown: %.3f" %qx.downHigh)
print("回撤最低点位 Low. drawdown: %.3f" %qx.downLow)
print('')
print("时间周期 Date lenght: %d (Day)" %qx.periodNDay)
print("时间周期（交易日） Date lenght(weekday): %d (Day)" %qx.
```

```
    print("开始时间 Date begin: %s" %qx.xtim0)
    print("结束时间 Date lenght: %s" %qx.xtim9)
    print('')
    print("项目名称 Project name: %s" % qx.prjName)
    print("策略名称 Strategy name: %s" % qx.staName)
    print('')
```

5.7 主体框架

5.7.1 stkLib内存数据库

在zwSys.py模块中，定义以下两个全局变量，用于保存股票数据源：

```
stkLib={} #全局变量，相关股票的交易数据，内存股票数据库
stkLibCode=[] #全局变量，相关股票的交易代码，内存股票数据库
```

其中，stkLib采用字典格式，保存了用户指定股票代码的股票数据，也就是当前使用的数据源。

stkLib字典，每一项key值就是股票代码，对应的赋值就是对应股票的历年交易数据，类似zwDat的csv交易数据文件，采用Pandas的DataFrame格式。

如统计沪深300，就是沪深300指数内的300只股票代码。stkLibCode是列表格式，是stkLib中股票的代码。

采用stkLib，而不采用mySQL数据库是为了简化程序结构，而且，stkLib数据赋值直接采用Pandas的DataFrame数据格式，中间无须进行格式转化，在效率上不会低于数据库，但未来实盘测试时，如果数据量

大，影响整体效率了再做优化处理。

注意

理论上，32位Python单个字典变量的数据上限是2GB。64位版本Python已经不受此限制，数据的上限几近于无限。

一线数据分析工程表明，十亿级别以下的数据量都可以用字典变量存储，配合Pandas直接分析，效率比Spark、Hadoop更高，而且极其方便。

stkLib的数据全部位于内存，可以看成一个Pandas版本的内存数据库。

如果采用zw的GPU工作站和全内存运算模式，即使不用GPU加速，32GB以上的用户，即使加载zwDat股票数据源的历年股票数据（约6GB），也可以全部导入内存，进行全内存计算。

内存的速度，比SSD要快50~100倍，比硬盘快近千倍，对于IO密集型的数据分析项目，整体效率可以提高10~20倍，一般的数据分析项目，也可以提高3~5倍。

【参见】

《zw·10倍速大数据与全内存计算》：http://ziwang.com/?p=186。

《zw开源量化GPU超算工作站草案》：http://ziwang.com/?

p=417。

5.7.2 Bars数据包

下面看一看zwXBars数据包的类定义：

```
class zwXBar(object):
     def __init__(self,xtim,mode='',code='',num=0,price=0):
         self.time=xtim;
         self.mode=mode;  #'',buy,sell
         self.code=code;
         self.num=num;
         self.price=price
         self.sum=price*num

     def prXBar(self):
         print('')
         print('date,',self.time);
         print('mode,',self.mode);
         print('code,',self.code);
         print('num,',self.num);
         print('price,',self.price);
         print('sum,',self.sum);
```

zwXBars数据包的类定义很简单，是最基本的交易数据记录单元，保存每一次的交易数据。其中的内部变量说明如下。

- time：交易时间。
- mode：交易模式，格式：字符串，可以为："（空字符）、'buy'（买进）、'sell'（卖出）。
- code：股票代码。
- num：交易数量。
- price：成交价格，即kprice。

此外，数据包zwXBars类还定义了一个方法：prXBar，用于输出类变量，保存的数据。

5.7.3 案例：内存数据库&数据包

变量stkLib的数据全部位于内存，可以看成是一个Pandas版本的内存数据库。

有些量化系统把与stkLib类似的变量也称为股票池，如stock、pool，类似传统IT编程的线程池。不管名称如何，目的都是把相关的股票数据集中处理，提高工作效率。传统的股票池、线程池一般采用静态镜像模式，pool池中的数据一般是不变的。

zwQuant的stkLib与传统的股票池、线程池技术略有不同。zwQuant的stkLib为提高效率，采用了动态处理的方式，在dataPre数据预处理阶段，会根据策略要求对股票池数据进行动态的衍生扩展。

目前stkLib是全局变量，在不影响效率的前提下，未来版本可能会因为zwQuant的升级，特别是多用户、实时版本的要求，整合到zwQuantX类当中，与qxLib、xtrdLib等一起作为zwQuantX类的子变量。

下面，我们以具体的案例来介绍stkLib流程数据库。

案例脚本文件名：\zwpython\zw_k10\zq501_stkLib.py，全部代码如下：

```
# -*- coding: utf-8 -*-
import numpy as np
import pandas as pd

import zwSys as zw          #zw.zwQuant
import zwTools as zwt
import zwQTBox as zwx

#======================

xlst=['aeti','egan','glng','simo']
zwx.stkLibRd(xlst,'dat\\');
print(zw.stkLibCode)

xcod='aeti';xtim='2011-01-03'
x1=zwx.xbarGet8Tim(xcod,xtim);
print('x1\n',x1)
```

```
xcod='aeti';xtim='2012-01-03'
x2=zwx.xbarGet8Tim(xcod,xtim);
print('x2\n',x2)

#def xbarGet8Tim(xcod,xtim):
#def xbarGet8TimExt(xcod,xtim):
x2,df2=zwx.xbarGet8TimExt(xcod,xtim);
print('x2\n',x2)

print('df')
print(df2.head())
print(df2.tail())
```

运行结果如下：

```
runfile('E:/zwPython/zw_k10x/zq501_stkLib.py', wdir='E:/zwPython
Reloaded modules: zwTools, zwQTBox, zwSys
['aeti', 'egan', 'glng', 'simo']
x1
              open  high   low  close  volume  adj close
date
2011-01-03  2.23  2.23  2.23   2.23       0       2.23
x2
  Empty DataFrame
Columns: [open, high, low, close, volume, adj close]
```

```
Index: []
x2
  Empty DataFrame
Columns: [open, high, low, close, volume, adj close]
Index: []
df
            open  high   low  close  volume  adj close
date
2011-01-03  2.23  2.23  2.23   2.23       0       2.23
2011-01-04  2.23  2.23  2.23   2.23       0       2.23
2011-01-05  2.16  2.23  2.16   2.16    3200       2.16
2011-01-06  2.20  2.20  1.90   2.12   97100       2.12
2011-01-07  2.09  2.20  2.09   2.17    4900       2.17
            open  high   low  close  volume  adj close
date
2011-12-23  4.57  4.58  4.40   4.55   23500       4.55
2011-12-27  4.59  4.60  4.41   4.60   33300       4.60
2011-12-28  4.47  4.65  4.47   4.64   32400       4.64
2011-12-29  4.65  5.34  4.48   5.15   73900       5.15
2011-12-30  5.00  5.16  4.85   5.08   51400       5.08
```

有关输出数据，请大家与脚本源码逐一对照，重点是以下三个函数，均位于模块：zwQTBox.py。

```
zwx.stkLibRd(xlst,'dat\\');
```

```
zwx.xbarGet8Tim(xcod,xtim);
zwx.xbarGet8TimExt(xcod,xtim);
```

其中，函数stkLibRd，用于读取数据文件，流程图如图5-9所示。

```
def stkLibRd(xlst,rdir):
     zw.stkLib={}    #全局变量，相关股票的交易数据
     zw.stkLibCode=[]  #全局变量，相关股票的交易代码
     zw.stkLibCode.extend(xlst);
     for xcod in xlst:
         #fss="dat\\"+xcod+"-2011.csv";
         fss=rdir+xcod+".csv";
         #df10=pd.read_csv(fss,index_col='date');
         df10=pd.read_csv(fss,index_col=0,parse_dates=[0]);
         df10=df2zwAdj(df10)
         zw.stkLib[xcod]=df10.sort_index();
```

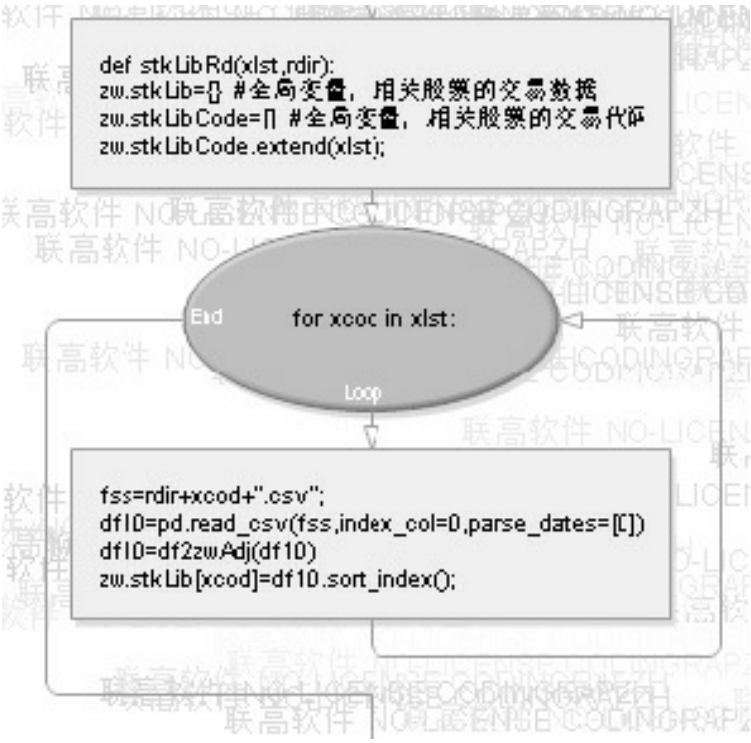


图5-9　函数stkLibRd流程图

函数stkLibRd可以一次性读取多个数据文件，股票代码在xlst列表案例程序中，脚本代码如下：

```
xlst=['aeti','egan','glng','simo']
zwx.stkLibRd(xlst,'dat\\');
print(zw.stkLibCode)
```

注意，stkLibRd函数，会自动对stkCdoe、stkLib进行清空处理，以免残留数据影响效率；另外，请注意stkLibRd函数中的以下脚本：

```
df10=df2zwAdj(df10)
```

可自动修改中国A股、Yahoo美股的数据格式，将其转换为为统一的zwQuant量化软件内部格式。其他格式的数据源，请看df2zwAdj函数源码，自行进行预处理，再进行导入。

用户做实盘时，修改stkLibRd函数的输入参数，rdir就可以设置不同的数据源，包括1分钟行情、5分钟行情等中、高频数据源。外汇、期货、原油、重金属期货等金融产品，根据数据入口stkLibRd函数和df2zwAdj函数源码对数据源进行预处理，即可直接使用zwQuant进行回溯测试、策略分析，以及实盘交易推荐。

函数xbarGet8Tim、xbarGet8TimExt根据指定的股票代码、时间，返回相关的OHLC等bar数据，如果没有，则返回空数据。

```
def xbarGet8Tim(xcod,xtim):
     d10=zw.stkLib[xcod]
     d02=d10[xtim:xtim];

     return d02

def xbarGet8TimExt(xcod,xtim):
     d10=zw.stkLib[xcod]
     d02=d10[xtim:xtim];

     return d02,d10
```

以下脚本可返回、输出数据信息：

```
xcod='aeti';xtim='2011-01-03'
x1=zwx.xbarGet8Tim(xcod,xtim);
print('x1\n',x1)

xcod='aeti';xtim='2012-01-03'
x2=zwx.xbarGet8Tim(xcod,xtim);
print('x2\n',x2)
```

对应的输出信息是：

```
x1
            open  high   low  close  volume  adj close
date
2011-01-03  2.23  2.23  2.23   2.23       0       2.23
x2
  Empty DataFrame
Columns: [open, high, low, close, volume, adj close]
Index: []
```

请注意，x2的输出信息是空数据，因为x2的对应时间点是'2012-01-03'，在案例中的数据源没有数据。xbar数据的来源是stkLib[xcod]对应的数据源。

5.7.4 qxLib、xtrdLib

在zwSys.py中定义以下两个全局变量，用于保存股票数据源：

```
stkLib={} #全局变量，相关股票的交易数据，内存股票数据库
stkLibCode=[] #全局变量，相关股票的交易代码，内存股票数据库
```

其中，变量stkLib采用字典格式保存了用户指定股票代码的股票数据，也就是当前使用的数据源。此外，我们在zwQuantX类还定义了一组类似的变量：

```
self.qxLib=pd.DataFrame(columns=qxLibName,index=['date']);     #所
self.xtrdLib=pd.DataFrame(columns=xtrdName,index=['date']);    #所
self.qxUsr=pd.Series(index=qxLibName)     #用户资产数据
self.qxUsrStk={};                         #用户持有的股票资产数据
```

以上交易记录和回溯分析数据都是按时间逐一保存的。

zwQuant量化分析系统是精简版的量化分析、回溯测试系统，面向的是单用户、单任务模式，因此可以采用这种简单的设计模式。

如果是多个用户，需要进行多种测试，例如，沪深300、上证50

等，可以设置不同的zwQuantX类变量和不同的参数，运行完一个任务，再运行第二个任务即可。

qxLib、xtrdLib直接使用Pandas的DataFrame格式，便于数据处理：

```
xtrdName=['date','ID','mode','code','dprice','num','kprice','sum
```

下面，我们介绍一下qxLib、xtrdLib的列命名，这类似于传统数据库的字段名称。

```
qxLibName=['date','stkVal','cash','dret','val','downLow','downHi
```

- date：交易时间，采用date是为了和股票数据源保持统一，此处的date表示时间列，不仅仅支持日期，也支持分钟、毫秒等分时数据。
- stkVal：当前时间点，用户手中持有的所有股票的市场价值。
- cach：当前时间点，用户手中持有的现金数额。
- downLow：到当前时间点为止，用户资产的最低价值水位。
- downHigh：到当前时间点为止，用户资产的最高价值水位。
- downDay：到当前时间点为止，用户回缩的最长时间。
- downKMax：到当前时间点为止，用户的最大回缩率。
- dret：单位回报率，如果time是日，则是“日回报率”。
- val：当前时间点，用户的全部资产价值。

需要注意的是，zwQuant类变量qxUsr采用的数据格式也是qxLibName。实际上，qxUsr是单个时间点用户的资产数据，qxLib是多个时间点用户数据的汇总表格。qxUsr还有一个衍生变量，即qxUsrStk，其定义如下：

```
self.qxUsrStk={};
```

变量qxUsrStk表示当前用户持有的所有股票数目，采用的是字典格式，每个股票对应一个数值。

下面我们看看xtrdLib中的列命名。

```
xtrdName=['date','ID','mode','code','dprice','num','kprice','sum
```

- date：交易时间，支持分钟、毫秒等分时数据。
- ID：用户交易订单ID编号。
- mode：订单模式。buy：买入；sell，卖出。
- code：股票代码。
- num：交易的股票数量。
- dprice：当前股票，策略分析采用的价格。
- kprice：当前股票，实际交易采用的价格。
- sum：当前股票，实际交易采用的总价格。
- cash：当前交易完毕，用户实际持有的现金数目。

xtrdLib是所有交易的记录列表，单个记录表示只股票交易订单的情况。

5.7.5 案例5-1：qxLib数据

下面，我们以具体的案例来介绍zwXBar数据包。

示例程序文件名：\zwpython\zw_k10\zq502_qxLib.py。

```
# -*- coding: utf-8 -*-
import numpy as np
import pandas as pd

import zwSys as zw          #zw.zwQuant
import zwQTBox as zwx

#======================
pd.set_option('display.width', 450)

qx=zw.zwQuantX('xbar',1000000); #100w
qx.priceCalc,qx.priceWrk,qx.priceBuy='close','close','close'
xlst=['aeti','egan','glng','simo']
zwx.stkLibRd(xlst,'dat\\');
#print(zw.stkCode)
#
```

```
xtim='2011-01-03';
qx.qxTim0SetVar(xtim);
qx.qxUsr=zwx.qxObjSet(qx.xtim0,0,qx.money,0);
#
qx.stkCode='simo';
qx.stkNum=100;
xfg,qx.xtrdChk=zwx.xtrdChkFlag(qx)
zwx.xtrdLibAdd(qx)
qx.qxTim9SetVar(xtim)

qx.prQLib()
```

运行结果如下：

```
runfile('E:/zwPython/zw_k10x/zq502_qxLib.py', wdir='E:/zwPython/

::qxUsr
date,2011-01-03; stkVal,441.0; cash,999559.0; dret,0.0; val,1000

::qxUsr.stk xbar
{'simo': 100}
::xtrdLib tmp\xbar_xtrdLib.csv
       date          ID mode  code  dprice    num  kprice    sum
0  2011-01-03  xbar_000001  buy  simo   4.41  100.0    4.41  441
```

```
::qxLib.head tmp\xbar_qxLib.csv
        date  stkVal      cash  dret       val    downLow   down
0  2011-01-03   441.0  999559.0   0.0  1000000.0  1000000.0  100

::qxLib.tail
        date  stkVal      cash  dret       val    downLow   dow
0  2011-01-03   441.0  999559.0   0.0  1000000.0  1000000.0  100
```

5.7.6 量化系统的价格体系

在zwQuant量化软件的类中，我们定义了多种价格变量：

```
#------设置各种环境的价格模式：
#    priceWrk，策略分析时使用的股票价格，一般是：dprice，复权开盘价
#    priceBuy，买入/卖出的股票价格，一般是：kprice，一般采用次日的复权开盘
#    priceCalc，最后结算使用的股票价格，一般是：adj close，复权收盘价
#    qxKPriceName=['open','high','low','close','adj close','dpri
self.priceWrk='dprice';
self.priceBuy='kdprice';
self.priceCalc='adj close';
```

其中，priceWrk、priceBuy、priceCalc分别表示用户进行策略分析、实际交易、以及最后结算时采用的价格模式。之所以采用多个变量

来表示价格，是由于实际业务的需要。

极宽Quant量化软件采用priceWrk、priceBuy、priceCalc作为策略分析的内部变量，用户只需在数据预处理时设置好相关参数即可，大大简化了策略分析内部流程的处理。

极宽Quant自带的案例程序对以上的各种价格模式都有涉及。对于初学者而言，起步阶段可以将priceWrk、priceBuy和priceCalc统一设置为“close”收盘价或“adj close”复权收盘价，在水平提高以后再进行优化。

在极宽Quant量化软件的价格体系中没有考虑佣金、税率等费用，这是因为相关费用不到交易总值的1%。

回溯测试和统计建模最终的结果都是选取大概率的策略，用户只需在最终盈利率参数上，根据保守、激进的操作需求，增加5%~10%的冗余进行修正即可。

这种源自工程一线的操作模式，对于策略评估、实盘操作的最终收益没有任何损失，却能大大简化量化系统的设计框架，提升回溯测试的速度与效率，简化量化策略的设计逻辑。

极宽Quant量化软件是一种优化模型的量化系统，在大数据和超算领域，这种模式已经是标准的优化策略。

Yahoo、谷歌的大数据、AI人工智能、机器学习和数据挖掘项目基本上都有这种模式。

最近火爆的alphaGO“人机围棋”大赛，其中最关键的算法，就是源自前两年，“蒙地卡罗”博弈算法的发布。

“蒙地卡罗”博弈算法的核心思想就是传统的“2：8定律”，既优先考虑：只需20%的计算量，胜率在80%的策略模型。

5.7.7 数据预处理

dataPre数据预处理模块是极宽Quant量化软件的一大创新。通常，金融产品的数据源只提供OHLC和成交量等基本数据，在任何量化回溯分析中都需要根据时间序列对这些数据进行衍生运算。传统的量化软件大部分是在策略分析时对每个数据进行衍生扩展。

在C、Java、Basic等传统编程语言中，这种模式类似lazy模式，对于整体项目的速度影响很小，有时，因为lazy是“按需计算”，还可提高整体运算项目。不过对于Python、Pandas、Numpy、MATLAB，以及GPU超算，这些基于矩阵的现代矢量数据处理框架显然不适合。

Pandas、Numpy都是经过高度优化的矢量矩阵模块库，以最简单的求和计算为例。

Pandas只需一个SUM函数，即可替代Python的数十行代码，而且速度快上百倍。

以下代码源自极宽Quant量化软件的策略函数库：zwStrategy.py，是最简单的均线平均策略，采用dataPre数据预处理函数，流程图如图5-10所示。

```
def SMA_dataPre(qx,xnam0,ksgn0):
     sta_dataPre0xtim(qx,xnam0);
     #----对各只股票数据进行预处理，提高后期运算速度
     ksgn,qx.priceCalc=ksgn0,ksgn0;  #'adj close';
     for xcod in zw.stkLibCode:
         d20=zw.stkLib[xcod];

         #  计算交易价格kprice和策略分析采用的价格dprice，kprice一般采用
         d20['dprice']=d20['open']*d20[ksgn]/d20['close']
         d20['kprice']=d20['dprice'].shift(-1)
         #d20['kprice']=d20['dprice']
         #
         d=qx.staVars[0];d20=zwta.MA(d20,d,ksgn);
         d=qx.staVars[1];d20=zwta.MA(d20,d,ksgn);
         #
         zw.stkLib[xcod]=d20;
         if qx.debugMod>0:
             print(d20.tail())
             #---
             fss='tmp\\'+qx.prjName+'_'+xcod+'.csv'
             d20.to_csv(fss)
```

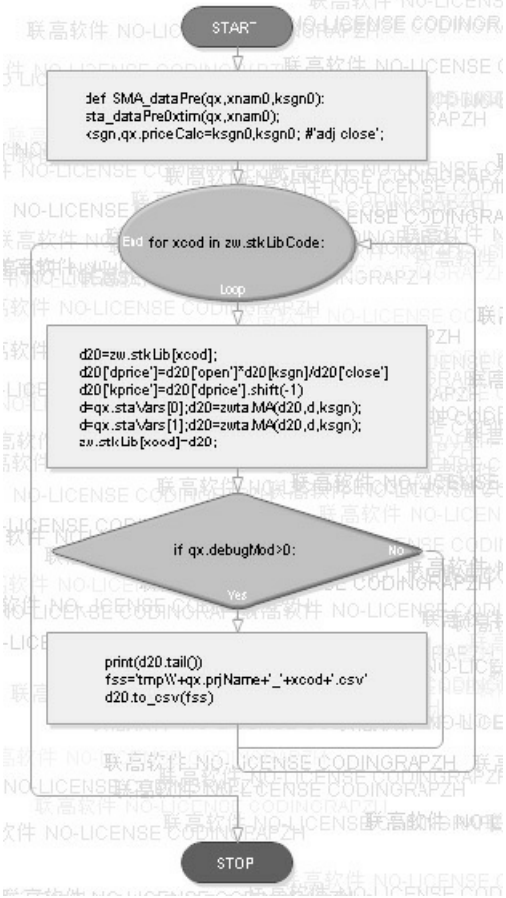


图5-10　数据预处理函数dataPre流程图

数据预处理函数dataPre程序代码当中，我们可以看到，函数自用了两行脚本：

```
d=qx.staVars[0];d20=zwta.MA(d20,d,ksgn);
d=qx.staVars[1];d20=zwta.MA(d20,d,ksgn);
```

就完成了5日、30日的均线数据计算。用户在进行策略分析时，无须计算，只需直接调用相关的数据列即可。

极宽Quant量化软件的dataPre数据预处理的具体细节，将在后面的

案例中进行讲解。极宽Quant量化软件的dataPre数据预处理函数，因为是基于Pandas、Numpy基础上的，所以具有以下优点：

- 简化量化系统、量化策略的整体代码量。
- 集中运算，大大提高回溯测试速度。
- 零成本，充分利用Pandas、Numpy等模块库底层的优化成果。
- 整体架构基于矩阵的矢量化运算，便于GPU优化。

5.7.8 绘图模板

极宽Quant的绘图模块文件名是：zwQTDraw.py。

极宽Quant绘图模块库很简单，采用的也是1+1模式：一个函数设置绘图环境，一个函数生成绘图。目前只有一个三合一的绘图模版：Quant 3x，用户可以参考相关代码自行添加。

- dr_quant3x_init(qx,w,h)：初始化绘图环境，w、h是图形大小尺寸。
- dr_quant3x(qx,ktop2,kbot,kmidlst,midSgn0='')：绘制图形。

3x三合一，回溯测试绘图函数。

【输入】

qx，全局量化参数变量

ktop2，Top顶部成交量股票代码

kbot，Bot底部绘图列名称，一般是'val'，资产总价值；数据源为qx.qxLib

kmidlst，Mid中部绘图列名称列表，为复合表格

子列表元素1，为股票代码xcod，其他列名称格式为：

[[xcod1,nam1,nam2,...],[xcod2,nam1,nam2,...],[xcod3,nam1,nam2,

注意，kmidlst数据源为：stkLib[xcod]，包含预处理扩充的数据列

midSgn0，中部绘图区图标前缀

其中，"<xcod>"为特殊符号，表示对应的股票代码

【输出】

无

如图5-11和图5-12所示是部分回溯案例截图。

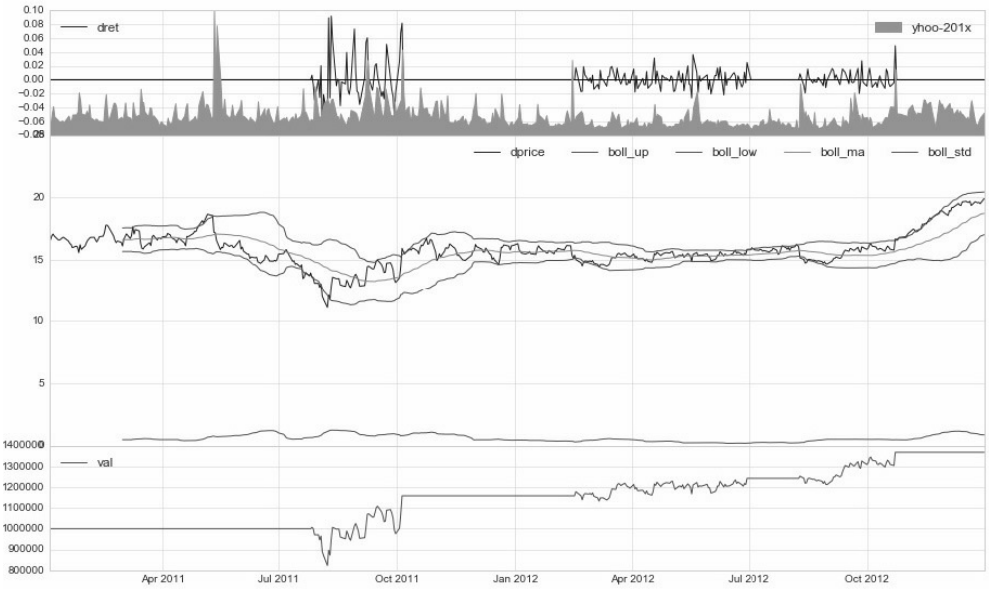


图5-11　案例脚本文件名：zq060_k410bbands.py

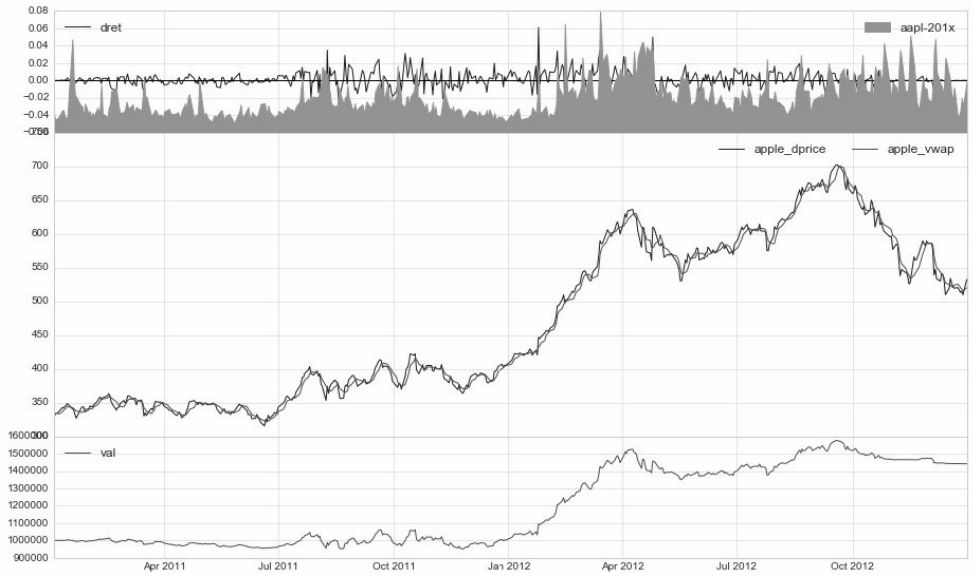


图5-12　案例脚本文件名：zq050_k408vwap.py

Quant　3x函数是个三合一的绘图模版，虽然输出的图形比较复杂，接近专业的金融图表，不过使用的模块库只是最基本的matplotlib绘图模块，没有使用seaborn新一代的Python图形模块库。

绘图报表属于比较专业的编程领域，对于初学者而言，初期掌握相关模版的调用方法即可。有经验的用户，可以参考Quant　3x函数和极宽Quant量化程序源码，自己设计图形模版。

5.8 新的起点

至此本书从宏观和微观角度介绍了目前全球量化领域，特别是Python量化领域的整体状况，并对Python量化的典型项目PAT量化软件进行了详细讲解。

在此基础上，本章借助zwQuant量化软件平台，进行正式的Python量化实盘策略分析。

本章的小结叫作新的起点，是因为对于各位读者而言：

- 本章是告别理论课程进行实盘操作和策略分析的新起点。
- 使用极宽Qaunt量化软件是大家告别传统K线图表，进入专业量化分析领域的新起点。

第6章
模块详解与实盘数据

本章我们将通过案例程序，讲解以下内容。

1. 回溯流程

- 案例：投资回报率
- 代码构成
- 运行总流程

2. 运行流程详解

- 设置股票数据源
- 设置策略参数
- dataPre数据预处理
- 绑定策略函数
- 回溯测试：zwBackTest
- 输出回溯结果数据、图表

3. 零点策略

- mul多个时间点的交易&数据
- 案例：多个时间点交易

4. 不同数据源与格式修改

- 案例：数据源修改
- 数据源格式修改

5. 金融数据包与实盘数据更新

- 大盘指数文件升级
- 实盘数据更新
- 案例：A股实盘数据更新
- 案例：大盘指数更新

6. 稳定第一

6.1　回溯流程

6.1.1　案例6-1：投资回报率

案例6-1，源码参见文件：\zwpython\zw_k10\zq601_k402inv.py。

运行结果，如图6-1所示。

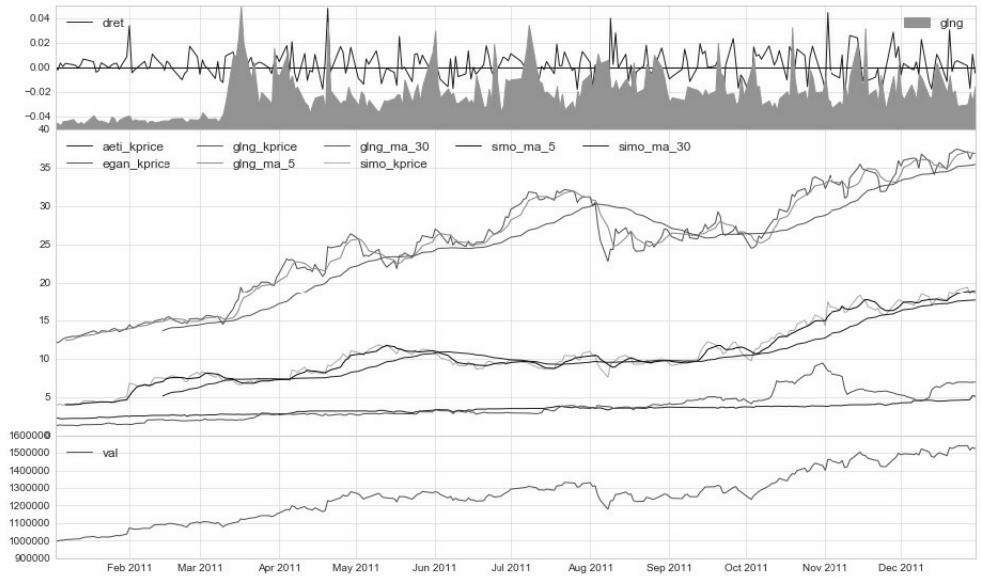


图6-1　运行结果

其他输出信息如下：

```
2011-01-03 2011-12-30
stkCode ['aeti', 'egan', 'glng', 'simo']
            open   high    low   close   volume   adj close   dprice
```

```
date
2011-12-23  4.57  4.58  4.40   4.55   23500       4.55    4.55
2011-12-27  4.59  4.60  4.41   4.60   33300       4.60    4.60
2011-12-28  4.47  4.65  4.47   4.64   32400       4.64    4.64
2011-12-29  4.65  5.34  4.48   5.15   73900       5.15    5.15
2011-12-30  5.00  5.16  4.85   5.08   51400       5.08    5.08
            open  high   low  close  volume  adj close  dprice
date
2011-12-23  6.83  7.00  6.83   6.91   16500       6.91    6.91
2011-12-27  6.97  7.12  6.79   6.92   23700       6.92    6.92
2011-12-28  6.85  6.95  6.76   6.87   13200       6.87    6.87
2011-12-29  6.89  7.00  6.83   6.91   12000       6.91    6.91
2011-12-30  6.96  7.01  6.85   6.91   17600       6.91    6.91
              open       high        low     close     volume  a
date
2011-12-23  45.380001  45.590000  44.910000  45.119999  278900
2011-12-27  45.110001  45.209999  44.540001  44.590000  299200
2011-12-28  44.849998  44.849998  43.500000  43.529999  456800
2011-12-29  43.669998  44.750000  43.669998  44.450001  340400
2011-12-30  44.779999  44.880001  43.820000  44.450001  526600
            open       high     low      close      volume  adj
date
2011-12-23  20.82  21.100000  20.719999  21.030001  396500  18.9
2011-12-27  21.09  21.600000  20.750000  21.480000  664100  19.3
2011-12-28  21.50  21.540001  20.350000  20.410000  515900  18.4
2011-12-29  20.40  20.980000  20.400000  20.920000  233300  18.8
2011-12-30  20.92  20.920000  20.320000  20.480000  280800  18.4
```

```
nday 362



::qxUsr
date,2011-12-30; stkVal,728347.930062; cash,795739.258782; dret,

::qxUsr.stk inv01
{'simo': 17293, 'glng': 11095}
::xtrdLib tmp\inv01_xtrdLib.csv
   date        ID mode       code  dprice  num     kprice     su
0  2011-01-03  inv01_000001  buy  glng  12.203890  11095.0  12.2
1  2011-01-03  inv01_000002  buy  simo   3.981876  17293.0   3.9

::qxLib.head tmp\inv01_qxLib.csv
  date         stkVal           cash         dret           val
0  2011-01-03  204260.741218  795739.258782  0.000000  1.000000e
1  2011-01-04  202111.937399  795739.258782 -0.002149  9.978512e
2  2011-01-05  205707.492065  795739.258782  0.003603  1.001447e
3  2011-01-06  206460.767876  795739.258782  0.000752  1.002200e
4  2011-01-07  209811.734705  795739.258782  0.003344  1.005551e

::qxLib.tail
     date         stkVal           cash       dret           val
247  2011-12-23  743094.145613  795739.258782  0.005510  1.53883
248  2011-12-27  745248.957007  795739.258782  0.001400  1.54098
249  2011-12-28  718798.600570  795739.258782 -0.017165  1.51453
250  2011-12-29  735218.179567  795739.258782  0.010841  1.53095
```

```
251  2011-12-30  728347.930062  795739.258782 -0.004488  1.52408
```

交易总次数：2

交易总盈利：524087.19

盈利交易数：2

盈利交易金额：524087.19

亏损交易数：0

亏损交易金额：0.00

最终资产价值 Final portfolio value: $1524087.19

最终现金资产价值 Final cash portfolio value: $795739.26

最终证券资产价值 Final stock portfolio value: $728347.93

累计回报率 Cumulative returns: 52.41 %

平均日收益率 Average daily return: 0.175 %

日收益率方差 Std. dev. daily return:0.0121

夏普比率 Sharpe ratio: 2.020,（0.05利率）

无风险利率 Risk Free Rate: 0.05

夏普比率 Sharpe ratio: 2.279,（0利率）

最大回撤率 Max. drawdown: 11.4729 %

最长回撤时间 Longest drawdown duration: 82

回撤时间(最高点位) Time High. drawdown: 2011-12-27

回撤最高点位 High. drawdown: 1540988.216

回撤最低点位 Low. drawdown: 1514537.859

```
时间周期 Date lenght: 362 (Day)
时间周期（交易日） Date lenght(weekday): 252 (Day)
开始时间 Date begin: 2011-01-03
结束时间 Date lenght: 2011-12-30

项目名称 Project name: inv01
策略名称 Strategy name: tim0Trade
```

案例6-1中的代码看起来有些长，其实，其中一半是注解，特别是代码尾部的大段注解是第4章案例4-2k402inv的运行结果，以方便进行对标测试。

本章通过本案例分析zwQuant量化软件的整个流程，同时学习其他量化软件的运作原理，因此，本案例虽然简单，但输出信息却繁多，这是因为我们在各个节点都添加了中间数据输出代码，便于初学者学习掌握。

这种在程序中间增加断点和输出数据的方式是一种标准的软件调试手段，特别是在浏览第三方程序源码时。

读者可能会觉得这段程序有些熟悉，的确，案例6-1正是第4章PAT量化软件案例4-2k402inv的移植版本，案例6-1虽然根据极宽Quant量化软件对有关代码重新进行了编写，但数据源和策略逻辑完全一致。

6.1.2 代码构成

下面，我们通过案例6-1简单讲解一下zwQuant量化软件的构成。

一套完整的极宽Quant回溯测试程序，通常包括以下部分：

- 文件头和库导入，即import语句等。
- 策略函数定义，如代码中的def tim0Trad(qx)函数。
- 策略数据预处理函数定义，如代码中的def tim0Trade_dataPre(qx,xnam0)函数。
- 策略回溯结果输出函数，如代码中的def bt_endRets(qx)函数。
- 策略回溯测试主流程也可以将主流程相关脚本编写成一个main函数。

6.1.3 运行总流程

zwQuant量化软件的运行流程主要是策略回溯的主流程。

案例6-1中的相关代码如下：

```
#--------设置参数
xlst=['aeti','egan','glng','simo']
qx=zwbt.bt_init(xlst,'dat\\','inv01',1000000);
#
#---设置策略参数
qx.staVars=[5,30,'','']
qx.debugMod=1
#---根据当前策略，对数据进行预处理
```

```
tim0Trade_dataPre(qx,'tim0Trade')
#
#---绑定策略函数&运行回溯主函数
qx.staFun=tim0Trad;
zwbt.zwBackTest(qx)
#
#   运行回溯结束，计算交易回报数据，绘制相关图表
bt_endRets(qx)
```

根据以上代码可以看到，zwQuant量化软件的回溯测试主流程是：

- 设置股票代码参数，初始化全局变量qx。
- 设置策略参数。
- 根据当前策略对数据进行预处理。
- 绑定策略函数。
- 运行回溯测试主函数zwBackTest。
- 输出回溯结果数据、图表。

6.2　运行流程详解

6.2.1　设置股票数据源

首先我们来看看流程的第一步：设置参数，初始化全局变量qx；有关程序代码如下：

```
xlst=['aeti','egan','glng','simo']
qx=zwbt.bt_init(xlst,'dat\\','inv01',1000000);
```

其中，变量xlst就是需要回溯测试的数据源：股票代码；变量xlst是Python标准的列表格式，可设置一次性支持多只股票，

基于指定的股票代码，把相关代码的交易数据读入stkLib内存数据库，即分析所用的内存数据库，或者股票池。

目前，变量xlst对应的都是股票代码，未来升级版本，可能会采用“<hs300>”等内置符号，表示股票集合的文件名，简化用户参数设置。极宽Quant量化软件在绘图模块中已经采用了这种内置符号技术。

qx是全局变量，保存了量化回溯测试和策略分析的各种参数。qx属于zwQuantX类变量，在函数bt_init里面申明，并作为返回数据。

函数bt_init是极宽Quant回溯测试的初始化参数设置函数，输入参

数，参见以下代码的注解：

```
def bt_init(xlst,rdat,prjNam,money0=1000000):
    '''
    qt_init(qx,xlst,rdat):
        初始化zw量化参数、stk股票内存数据库等
    【输入】:
        xlst，股票代码列表，例如：xlst=['aeti','egan','glng','simo'
        rdat，股票数据目录，可直接使用 zwDat的中美股票数据:\\zwdat\\cr
        prjNam，项目名称
        money0，启动资金，默认是：1000000(100w)

    【输出】:
        qx，程序化全局变量qx
    '''
    #qx=zw.zwQuantX('tur',1000000); #100w
    qx=zw.zwQuantX(prjNam,money0); #100w
    #------设置各种价格模式：
    #    priceWrk，策略分析时使用的股票价格，一般是：dprice，复权开盘价
    #    priceBuy，买入/卖出的股票价格，一般是：kprice，一般采用次日的复
    #    priceCalc，最后结算使用的股票价格，一般是：adj close，复权收盘
    qx.priceCalc='adj close';
    qx.priceWrk='dprice';
    qx.priceBuy='kprice';
    #----------设置绘图&数据输出格式
    mpl.style.use('seaborn-whitegrid');
```

```
    pd.set_option('display.width', 450)
    #-----设置数据源目录等场所，读取股票数据，stkLib
    qx.rdat=rdat;
    zwx.stkLibRd(xlst,rdat);

    #  读取股票数据，
    xtim0=parse('9999-01-01');xtim9=parse('1000-01-01');
    for xcod in xlst:
        xt0,xt9=zwx.stkLibGetTimX(xcod)
        if xtim0>xt0:xtim0=xt0
        if xtim9<xt9:xtim9=xt9
    #
    xtim0=xtim0.strftime('%Y-%m-%d');xtim9=xtim9.strftime('%Y-%
    qx.qxTimSet(xtim0,xtim9)
    print(xtim0,xtim9,'\nstkCode',zw.stkLibCode)    #zwx.stkLib

    #
    return qx
```

初始化函数bt_init流程图如图6-2所示。

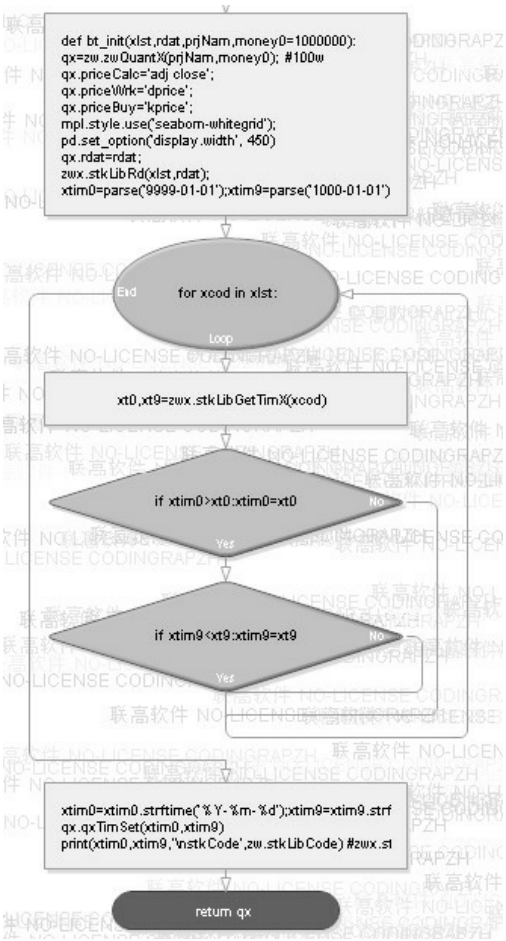


图6-2　初始化函数bt_init流程图

初始化函数bt_init主要有以下功能：

- 申明qx变量，并对部分qx内部属性变量进行设置。
- 设置数据源目录。
- xlst设定的股票代码，读取数据源，并修正数据源格式为zwQuant内部格式。
- 根据数据源设置回溯起始、结束时间。
- 这个只是目前版本初始化函数bt_init的作用，未来zwQuant量化软件升级后的相关功能、参数可能会有所变化，请大家以最新发布的源码为准。

这种函数和函数调用接口的变化，以及新旧版本接口的改变会对用户带来一些困惑，不过开源软件本身都提供源码，修改也很方便，这也是开源软件自身不断优化的一种特色。如Pandas数据分析软件，在v0.17版本以后，apply、roll等函数组都进行了彻底改变，而数据源模块IO，甚至成为一个独立的开源项目：pandas_datareader。

6.2.2 设置策略参数

策略参数的设置参见以下代码：

```
#---设置策略参数
qx.staVars=qx.staVars=[5,30,'','']
qx.debugMod=1
```

其中qx.debugMod=1，设置调试级别，用于控制调试和中间信息数据的输出。0为非调试状态；1为1级调试状态。目前只用到1级调试模式。

```
qx.staVars=qx.staVars=[5,30,'','']
```

qx的内置变量staVars是策略参数。因为各种策略是灵活多变的，所需的参数数目也是不同的，所以极宽Quant的zwQuantX类定义中采用列

表变量staVars传递策略参数。

staVars原名是staVarList，为精简结构，采用了短变量名称：staVars。

变量staVars是标准的Python列表格式，可传递数值、字符串、字典、列表等数据，不过，通常只传递数值和字符串参数。

默认状态staVars最后两位表示策略回溯测试的起始、结束时间：

- staVars[-2]，起始时间。
- staVars[-1]，结束时间。

空字符串表示使用数据源的起始、结束时间。

有关回溯时间的设置参见策略函数模块：zwStrategy.py。

```
函数名：def sta_dataPre0xtim(qx,xnam0):
```

其中，数据预处理函数def sta_dataPre0xtim流程图如图6-3所示。

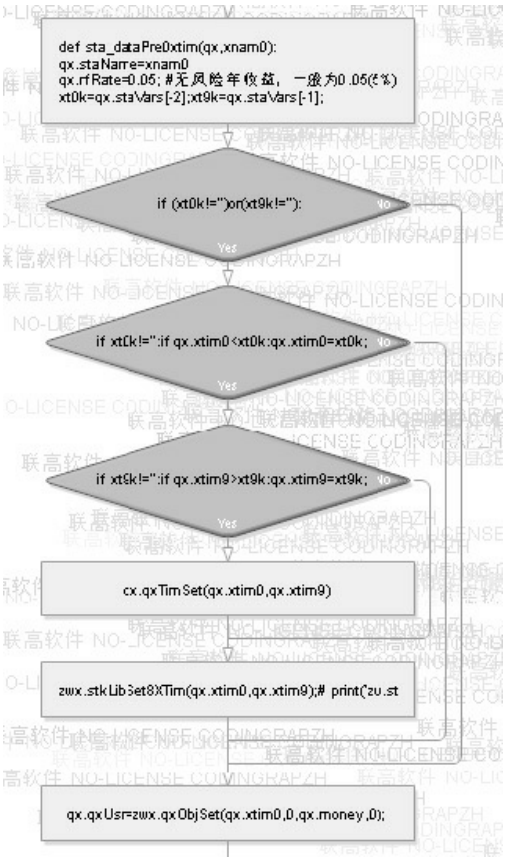


图6-3　数据预处理函数def sta_dataPre0xtim流程图

数据预处理函数def sta_dataPre0xtim，相关代码如下：

```
def sta_dataPre0xtim(qx,xnam0):
    '''
    设置策略参数，根据预设时间，裁剪数据源stkLib
    '''
    #设置当前策略的变量参数
    qx.staName=xnam0
    qx.rfRate=0.05;  #无风险年收益，一般为0.05(5%)，计算夏普指数等需要
    #qx.stkNum9=20000;    #每手交易，最多20000股
    #
```

```
    #按指定的时间周期，裁剪数据源
    xt0k=qx.staVars[-2];xt9k=qx.staVars[-1];
    if (xt0k!='')or(xt9k!=''):
        #xtim0=parse('9999-01-01');xtim9=parse('1000-01-01');
            #xtim0=xtim0.strftime('%Y-%m-%d');xtim9=xtim9.strft
        if xt0k!='':
            if qx.xtim0<xt0k:qx.xtim0=xt0k;
        if xt9k!='':
            if qx.xtim9>xt9k:qx.xtim9=xt9k;
        qx.qxTimSet(qx.xtim0,qx.xtim9)
        zwx.stkLibSet8XTim(qx.xtim0,qx.xtim9);#    print('zw.st
    #============
    #---设置qxUsr用户数据
    qx.qxUsr=zwx.qxObjSet(qx.xtim0,0,qx.money,0);
```

代码当中：

```
xt0k=qx.staVars[-2];xt9k=qx.staVars[-1];
```

表示staVars传递的起始、结束时间。

6.2.3 dataPre数据预处理

传统量化系统里也有数据预处理环节，但一般只是处理数据源格式的不同。与策略分析相关的数据，例如MA均线、MACD、布林带、RSI等全部在策略分析时进行实时运算。这种单点计算模式类似于lazy模式，对于传统编程语言而言，对效率影响不大。

但Python语言，特别是Pytrhon　3.5以后的版本，内置了矩阵运算，与MATLAB类似，配合Pandas、Numpy，在矩阵、矢量化运算方面，对于传统的单点计算模式具有百倍、甚至千倍的速度优势，数据越大，效果越明显。而且Python作为cuda三大官方编程语言，在未来版本中，GPU加速能带来更大的速度优势。

正因为如此，zwQuant量化软件在Python量化领域率先导入dataPre集中式数据预处理，针对各种不同策略，对相关参数统一进行预处理，在硬件方面，配合目前廉价的大内存，以内存空间换取运算时间效率的提升。测试表明，对于同样的策略，zwQuant量化软件即使不采用其他优化手段，策略回溯测试速度也是传统量化程序的3~5倍以上。

下面，我们看看案例zq601_k402inv.py当中的dataPre数据预处理函数，其流程图如图6-4所示。

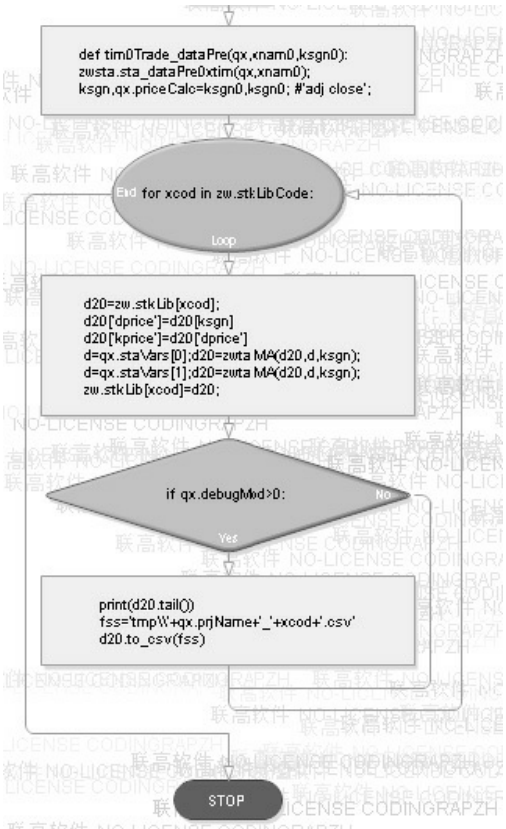


图6-4　案例zq601数据预处理函数流程图

案例zq601_k402inv中，数据预处理函数相关代码如下：

```
def tim0Trade_dataPre(qx,xnam0,ksgn0):
    zwsta.sta_dataPre0xtim(qx,xnam0);
    #-------
    ksgn,qx.priceCalc=ksgn0,ksgn0;  #'adj close';
    for xcod in zw.stkLibCode:
        d20=zw.stkLib[xcod];
        #  计算交易价格kprice和策略分析采用的价格dprice,kprice一般采用
        #d20['dprice']=d20['open']*d20[ksgn]/d20['close']
        d20['dprice']=d20[ksgn]
```

```
        #d20['kprice']=d20['dprice'].shift(-1)
        d20['kprice']=d20['dprice']
        #
        d=qx.staVars[0];d20=zwta.MA(d20,d,ksgn);
        d=qx.staVars[1];d20=zwta.MA(d20,d,ksgn);
        #
        zw.stkLib[xcod]=d20;
        if qx.debugMod>0:
            print(d20.tail())
            fss='tmp\\'+qx.prjName+'_'+xcod+'.csv'
            d20.to_csv(fss)
```

案例中的数据预处理函数名是：tim0Trade_dataPre。不同的策略数据预处理函数都不同，习惯上使用“_dataPre”作为函数名尾字符（后缀）。

dataPre数据预处理函数，主要完成以下任务。

- 调用sta_dataPre0xtim函数，完成策略初始化参数设置，以及根据回溯时间切割数据源。
- 对股票数据源当中的所有股票：
  - 设置dpriec策略价格、kpriec交易价格、priecCalc结算价格。
  - 根据策略要求计算相关数据，本例当中是5日、30日MA均线数据。
  - 根据调试模式，按股票代码分别保存stkLib，用于调试、分析。

dataPre数据预处理函数生成的中间数据文件名格式为：项目名称+'_'+股票代码。以案例中的aeti股票代码为例，生成的中间数据如图6-5所示。

inv01_aeti.csv

	A	B	C	D	E	F	G	H	I	J	K
1	date	open	high	low	close	volume	adj close	dprice	kprice	ma_5	ma_30
2	2011-1-3	2.23	2.23	2.23	2.23	0	2.23	2.23	2.23		
3	2011-1-4	2.23	2.23	2.23	2.23	0	2.23	2.23	2.23		
4	2011-1-5	2.16	2.23	2.16	2.16	3200	2.16	2.16	2.16		
5	2011-1-6	2.2	2.2	1.9	2.12	97100	2.12	2.12	2.12		
6	2011-1-7	2.09	2.2	2.09	2.17	4900	2.17	2.17	2.17	2.182	
7	2011-1-10	2.2	2.2	2.2	2.2	500	2.2	2.2	2.2	2.176	
8	2011-1-11	2.2	2.2	2.2	2.2	0	2.2	2.2	2.2	2.17	
9	2011-1-12	1.94	2.22	1.94	2.2	3200	2.2	2.2	2.2	2.178	
10	2011-1-13	2.19	2.21	2.1	2.21	5400	2.21	2.21	2.21	2.196	
11	2011-1-14	2.2	2.2	2.19	2.19	2100	2.19	2.19	2.19	2.2	
12	2011-1-18	2.23	2.23	2.22	2.23	2200	2.23	2.23	2.23	2.206	
13	2011-1-19	2.23	2.25	2.23	2.25	3600	2.25	2.25	2.25	2.216	
14	2011-1-20	2.29	2.36	2.29	2.36	4900	2.36	2.36	2.36	2.248	
15	2011-1-21	2.24	2.36	2.24	2.35	2300	2.35	2.35	2.35	2.276	
16	2011-1-24	2.34	2.34	2.3	2.3	500	2.3	2.3	2.3	2.298	
17	2011-1-25	2.35	2.35	2.35	2.35	2400	2.35	2.35	2.35	2.322	
18	2011-1-26	2.38	2.46	2.35	2.35	3200	2.35	2.35	2.35	2.342	
19	2011-1-27	2.3	2.4	2.3	2.4	5500	2.4	2.4	2.4	2.35	
20	2011-1-28	2.41	2.45	2.3	2.43	3300	2.43	2.43	2.43	2.366	
21	2011-1-31	2.43	2.5	2.43	2.5	10200	2.5	2.5	2.5	2.406	
22	2011-2-1	2.49	2.5	2.45	2.45	1400	2.45	2.45	2.45	2.426	
23	2011-2-2	2.5	2.53	2.5	2.52	2600	2.52	2.52	2.52	2.46	
24	2011-2-3	2.66	2.76	2.3	2.5	13800	2.5	2.5	2.5	2.48	
25	2011-2-4	2.31	2.59	2.2	2.51	4100	2.51	2.51	2.51	2.496	
26	2011-2-7	2.51	2.51	2.51	2.51	0	2.51	2.51	2.51	2.498	
27	2011-2-8	2.51	2.55	2.51	2.55	400	2.55	2.55	2.55	2.518	
28	2011-2-9	2.56	2.56	2.56	2.56	800	2.56	2.56	2.56	2.526	
29	2011-2-10	2.57	2.59	2.52	2.59	900	2.59	2.59	2.59	2.544	
30	2011-2-11	2.44	2.59	2.4	2.59	4400	2.59	2.59	2.59	2.56	
31	2011-2-14	2.59	2.59	2.4	2.59	2700	2.59	2.59	2.59	2.576	2.36
32	2011-2-15	2.54	2.59	2.54	2.59	1400	2.59	2.59	2.59	2.584	2.372
33	2011-2-16	2.58	2.59	2.41	2.58	8600	2.58	2.58	2.58	2.588	2.383667
34	2011-2-17	2.57	2.59	2.46	2.59	5300	2.59	2.59	2.59	2.588	2.398

图6-5　aeti股票数据

从图6-5可以看出，dataPre数据预处理函数扩展后的数据除保留了数据源原本的OHLC、成交量等数据以外，还增加了dpriec、kpriec、ma_5、ma_10等数据列。

软件工程已经证明，在所有算法中，查表是最快的方式，无须计算，一个下标即可获得所需结果。不过，这种效率的提高是以牺牲内存为代价的。

大数据，大内存是必须的，个人电脑、笔记本电脑很容易扩充到32GB。机构用户，可能需要一次性导入所有股票数据到内存，可以采

用支持1T以上大内存的专业服务器。

tushare数据采集软件和Yahoo雅虎财经接口提供的日线数据，不管是不是包含复权，在遇到配股、复盘等特殊时间点时，两个交易日的价格差别很大，都会带来数据干扰。在数据预处理时，如果两个相邻的时间点数据差异过大，超过k值，就可视为特殊时间点，进行特殊处理。目前的dataPre数据预处理函数没有包含以上功能，未来升级版本会考虑增加相关功能。

6.2.4 绑定策略函数

绑定策略函数参见以下代码：

```
qx.staFun=tim0Trad;
```

请注意，在这里我们使用了“绑定”这个词，形象、充分地体现了Python语言的优雅、灵活与强大。只需一个赋值语句就把用户自定义的策略函数与staFun函数接口绑定了。

而且，调用也非常简单、优雅，代码如下：

```
qx.stkNum=qx.staFun(qx);
```

系统会根据用户预定义staFun函数接口“绑定”不同的函数，运行不同的策略。这种函数“绑定”模式，在传统的编程语言（如C语言）中是很难实现的。具体的策略在后面的章节再进行详细解说。

6.2.5 回溯测试：zwBackTest

zwQuant量化软件的回溯测试调用很简单，设置好相关参数后，直接调用回溯测试函数即可：

```
zwbt.zwBackTest(qx)
```

zwBackTest回溯测试函数位于模块：zwBacktest.py。

函数代码如下：

```
def zwBackTest(qx):
    '''
    zwQuant，回溯测试主程序
    #---debug,#上级主叫函数 sys._getframe().f_back.f_code.co_name
    #zwt.xdebug(qx.debugMod,__name__,sys._getframe().f_code.co_
    '''
    #计算回溯时间周期，也可以在此根据nday调整回溯周期长度
    #或者在 qt_init数据初始化话时，通过qx.qxTimSet(xtim0,xtim9)，设置
```

```
    nday=qx.periodNDay;print('nday',nday);   #nday=3;
    #按时间循环，进行回溯测试
    for tc in range(nday):
        tim5=qx.DTxtim0+dt.timedelta(days=tc)
        xtim=tim5.strftime('%Y-%m-%d')  #        print('tc',tc,
        #每个测试时间点开始时，清除qx相关参数
        qx.qxTim0SetVar(xtim);  #qx.prQxUsr();#qx.xtim=xtim;
        xpriceFlag=False;  #有效交易标志Flag
        #按设定的股票代码列表，循环进行回溯测试
        for xcod in zw.stkLibCode:
            qx.stkCode=xcod;    #print('xcod',xcod)
            #xdatWrk是当前xcod，=stkLib[xcod]
            #xbarWrk是当前时间点的stkLib[xcod]
            #注意,已经包括了qt_init里面的扩充数据列
            qx.xbarWrk,qx.xdatWrk=zwx.xbarGet8TimExt(xcod,qx.xt
            #print(xcod,'xbar\n',qx.xbarWrk)
            if not qx.xbarWrk[qx.priceWrk].empty:
                xpriceFlag=True
                # 调用回溯子程序，如果是有效交易，设置成功交易标志xtrdF
                zwBackTest100(qx)

        #如果所有股票代码列表循环完毕，成功交易标志为真
        #在当前测试时间点终止，设置有关交易参数
        if xpriceFlag:
            qx.wrkNDay+=1
            qx.qxTim9SetVar(qx.xtim);
```

zwBackTest回溯测试函数代码本身已经内置了大量的中文注解，流程图如图6-6所示。

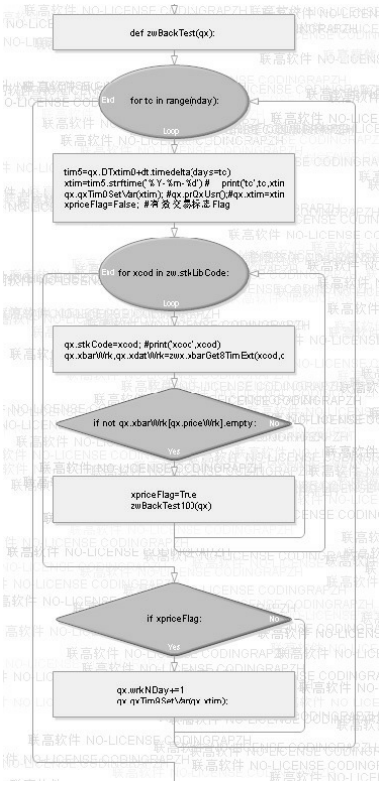


图6-6　zwBackTest回溯测试函数流程图

下面再根据函数代码，简单介绍一下zwBackTest回溯测试函数代码的运行流程。

- 获取nday：回溯测试周期，这个nday在建模时也可人工设定，如30、50、100，用于不同时间周期的回溯测试。
- 遍历nday时间周期：
  - 遍历股票池内的所有股票代码。
    - 设定回溯的时间点、股票代码。
    - 获取当前时间点、股票代码的xbar数据。

- 根据xbar的数据，跳过非交易日数据。
- 交易日，运行回溯测试子程序zwBackTest100。

- 当前时间点，遍历结束。

- 如果是交易日：设置交易数据qxTim9SetVar。

由以上流程可以看到，具体的回溯是在回溯测试子程序zwBackTest100。

相关代码如下：

```
def zwBackTest100(qx):
    '''
    zwBackTest100(qx):
    zwQT回溯测试子函数，测试一只股票xcod，在指定时间xtim的回溯表现数据
    会调用qx.staFun指定的策略分析函数，获取当前的股票交易数目 qx.stkNum
    并且根据股票交易数目 qx.stkNum，判定是不是有效的交易策略
    【输入】
            qx.stkCode，当前交易的股票代码
            qx.xtim，当前交易的时间
    【输出】
         无
    '''
    #----运行策略函数，进行策略分析
    qx.stkNum=qx.staFun(qx);
    #----
```

```
    if qx.stkNum!=0:
        #----检查，是不是有效交易
        xfg,qx.xtrdChk=zwx.xtrdChkFlag(qx)
        if xfg:
            #----如果是有效交易，加入交易列表
            zwx.xtrdLibAdd(qx)
            #qx.prQCap();
```

zwBackTest100回溯测试子程序，运行流程如下。

- 调用策略函数qx.staFun，返回成交的股票数目stkNum。
- 如果股票说明不等于0，则表示有交易。
  - 检查是不是是有效交易，函数xtrdChkFlag。
  - 如果是有效交易，则交易数据，加入到xtrdLib。

回溯测试子程序zwBackTest100，流程图如图6-7所示。

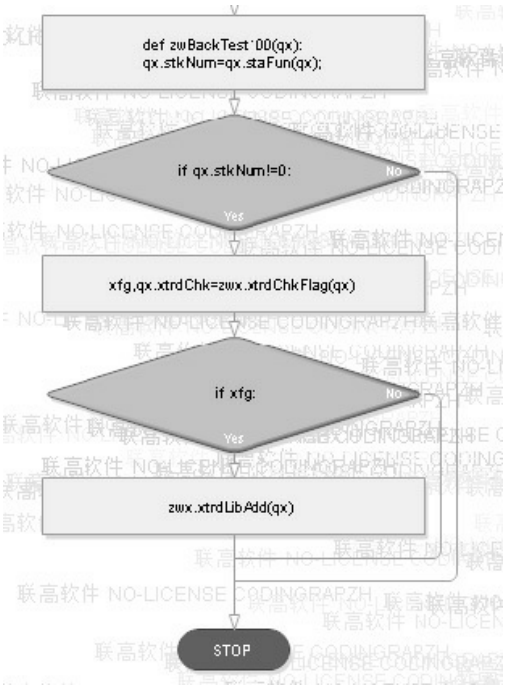


图6-7　回溯测试子程序zwBackTest100流程图

6.2.6　输出回溯结果数据、图表

极宽Quant量化软件的数据、图表输出，参见bt_endRets函数，代码如下：

```
def bt_endRets(qx):
    #---ok ，测试完毕
    # 保存测试数据，qxlib，每日收益等数据；xtrdLib，交易清单数据
    #qx.qxLib=qx.qxLib.round(4)
    qx.qxLib.to_csv(qx.fn_qxLib,index=False,encode='utf-8')
    qx.xtrdLib.to_csv(qx.fn_xtrdLib,index=False,encode='utf-8')
    qx.prQLib()
```

```
    #
    #-------计算交易回报数据
    zwx.zwRetTradeCalc(qx)
    zwx.zwRetPr(qx)

    #-------绘制相关图表，可采用不同的模板
    # 初始化绘图模板：dr_quant3x
    zwdr.dr_quant3x_init(qx,12,8);
    #  设置绘图相关参数
    xcod='glng';ksgn=qx.priceBuy;
    kmid8=[['aeti',ksgn],['egan',ksgn],['glng',ksgn,'ma_5','ma_
    # 绘图
    zwdr.dr_quant3x(qx,xcod,'val',kmid8,'<xcod>')
    # 可设置，中间图形窗口的标识
    #qx.pltMid.legend([]);
```

bt_endRets函数之所以和主函数放在一起，而没有收录到模块库中，是因为各种策略、输出的细节、绘图方面有所不同。输出图形如图6-8所示。

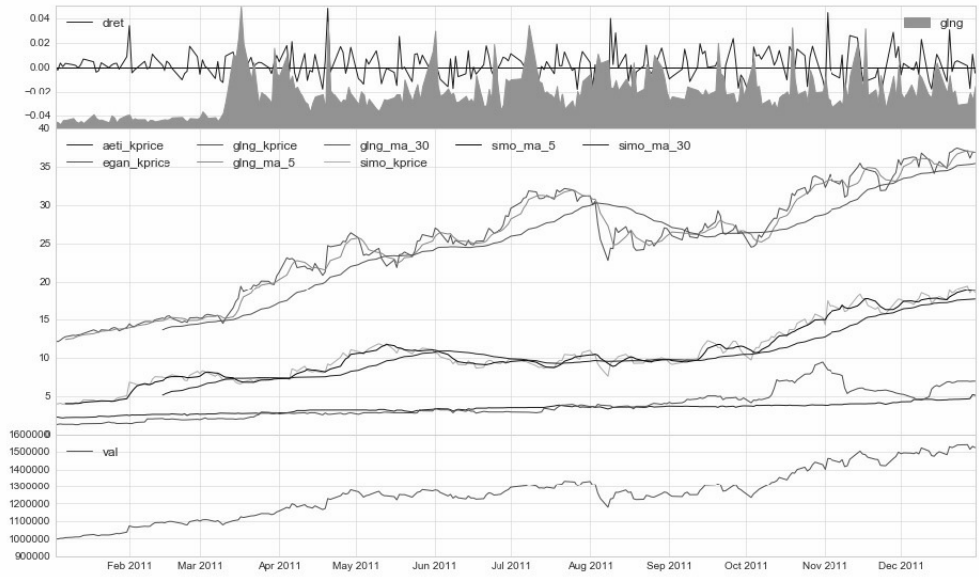


图6-8　bt_endRets函数图表输出

下面，我们根据运行流程，分别进行介绍。

- 保存回溯测试结果，并输出相关信息。
- 计算交易回报数据，并输出相关信息。
- 设置绘图参数。
- 输出图形。

bt_endRets函数输出图形以前的语句基本都是相同的，用户一般无须修改。

需要注意以下语句：

```
#   设置绘图相关参数
xcod='glng';ksgn=qx.priceBuy;
kmid8=[['aeti',ksgn],['egan',ksgn],['glng',ksgn,'ma_5','ma_30'],
# 绘图
```

```
zwdr.dr_quant3x(qx,xcod,'val',kmid8,'<xcod>')
```

其中变量kmid8用于设置中部绘图列名称列表为复合表格，子列表元素1，为股票代码xcod，其他列名称，格式为：[[xcod1,nam1,nam2,...], [xcod2,nam1,nam2,...], [xcod3,nam1,nam2,...]]。

注意，kmidlst数据源为stkLib[xcod]，包含预处理扩充的数据列。

这种复合表格定义有一些复杂，初学者可以参考实际调用案例代码，以便于理解学习。

绘图函数为dr_quant3x：

```
zwdr.dr_quant3x(qx,xcod,'val',kmid8,'<xcod>')
```

其中的“<xcod>”为特殊符号，表示对应的股票代码，如果为空，则采用以下语句：

```
zwdr.dr_quant3x(qx,xcod,'val',kmid8,'')
```

如图6-9所示，中部的图标数据因为没有股票代码，会出现重名，而引发混乱。

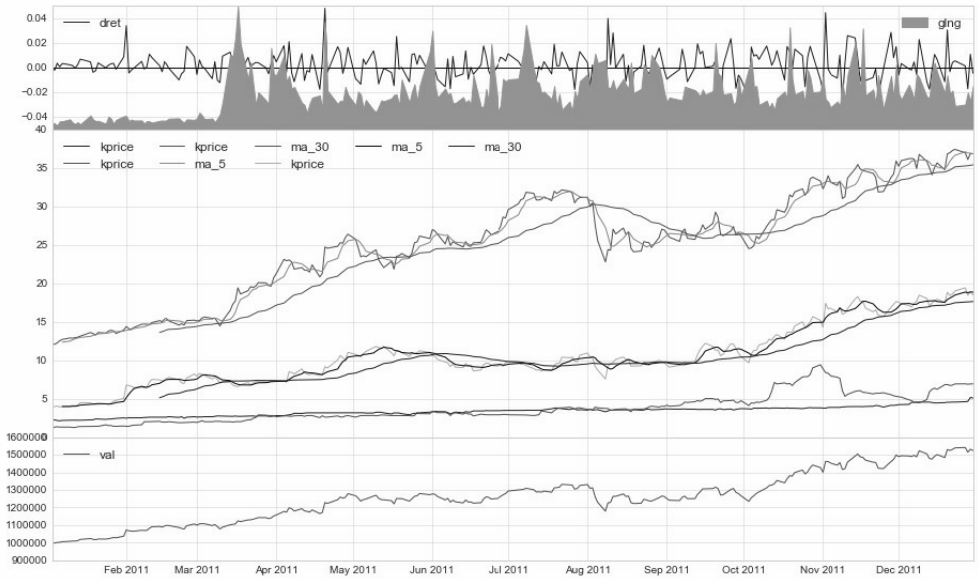


图6-9　图表输出错误：重名

6.3 零点策略

下面我们来分析讲解案例zq601_k402inv.py的操作策略。这个案例策略的实质是，在年初购买4只股票，年内不进行任何交易，年底再计算相关的收益。

许多私营企业主平常没有太多空闲时间，又想自己投资，往往采用这种模式。所有的交易都是在第一个交易日完成，所以也叫零点策略，相关代码如下：

```
def tim0Trad(qx):
    '''
        tim0Trad(qx):
        零点交易策略，可看作是基于时间的：事件模式·回溯测试
        【输入】
            qx.stkCode，当前交易的股票代码
            qx.xtim，当前交易的时间
        【输出】
            srkNum，当前股票代码xcod，交易的股票数目：>0，买入；<0，卖
    '''
    stknum=0;
    xtim=qx.xtim;
    xcod=qx.stkCode;
    xday=rrule.rrule(rrule.DAILY,dtstart=qx.DTxtim0,until=parse
    #第一天
```

```
    if xday==1:
        #print('xd',xday,dtim0,xtim)
        #按预设订单，下单购买股票
        if xcod=='aeti':stknum=297810;
        if xcod=='egan':stknum=81266;
        if xcod=='glng':stknum=11095;
        if xcod=='simo':stknum=17293;
        #print('tim',xtim,dtim0,xcod,stknum,dprice);

    if xday==-2:
        if xcod=='aeti':stknum=500;
        if xcod=='egan':stknum=500;
        if xcod=='glng':stknum=-1000;

    if xday==-3:
        if xcod=='aeti':stknum=-500;
        if xcod=='egan':stknum=-500;
        if xcod=='glng':stknum=1000;

    return stknum
```

零点策略函数tim0Trad流程图如图6-10所示。

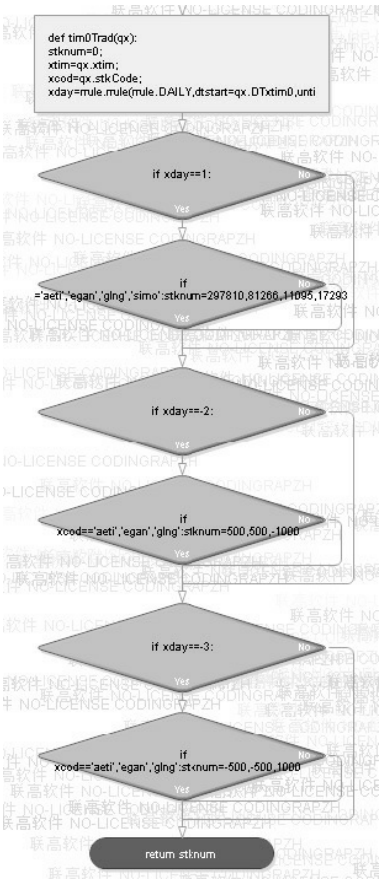


图6-10　零点策略函数tim0Trad流程图

零点策略函数tim0Trad的主要运作流程如下。

- 初始化stknum，从qx全局变量获取xcod股票代码和xtim策略执行时间点。
- 因为是零点策略，所有需要计算第几交易日：xday。
- 如果xday等于1，则表示是第一个交易日。
    - 按预设订单，下单购买股票。
- 返回成交的股票数目stknum。

变量xday是多个时间点交易预设代码，前面加负号“-”屏蔽，在此处不会自行运行。有关的交易订单保存在变量xtrLib中。

在运行结果中查看xtrdLib变量，会发现只有glng、simo的股票是有效交易。

```
::xtrdLib tmp\inv01_xtrdLib.csv
   date        ID mode      code  dprice   num     kprice    sum
0  2011-01-03  inv01_000001  buy  glng  12.203890  11095.0  12.2
1  2011-01-03  inv01_000002  buy  simo   3.981876  17293.0   3.9
```

在pat量化软件和QSTK量化最初案例，以及其他量化交易系统中，都对单次股票交易数目的上限有限制，zwQuant量化软件也对单次交易上限进行了设置。

在zwSys.py模块，类zwQuantX的init初始化函数中，每手交易最多20 000股。

```
self.stkNum9=20000;    #每手交易，最多20000股
```

当然，这个上限是可以调增的，在特殊情况下可以人工设置。修改语句必须在qx=zwbt.bt_init变量初始化代码之后。

6.3.1 mul多个时间点的交易&数据

下面，我们来看看扩充版的“零点策略”：多个时间点的交易。在介

绍具体代码之前，我们要对零点策略进行深度讲解。

在案例投资回报率中，有两段屏蔽的代码：

```
if xday==-2:
     if xcod=='aeti':stknum=500;
     if xcod=='egan':stknum=500;
     if xcod=='glng':stknum=-1000;

if xday==-3:
     if xcod=='aeti':stknum=-500;
     if xcod=='egan':stknum=-500;
     if xcod=='glng':stknum=1000;
```

大家知道，传统的量化程序大部分基于事件模式，事件模式的onXXX触发函数是Windows时代消息编程的产物，与传统的流程式编程完全不同，复杂很多。

但事件模式对于某些基于突发事件的交易策略而言存在巨大的商业机会。

量化投资本质上是通过统计模型套利这种套利必须基于动荡的市场，市场动荡越大，机会点越多，投资机会越多。

案例投资回报率中的零点策略函数可以看作是一个通用的事件策略入口函数模版。只是，案例中采用的是最简单的xday变量：第几个交易日。

如果换成日期变量xt：

- 如果xt等于911，则采用xx策略。
- 如果xt等于圣诞假期，则采用xx策略。

如果再复杂一点，大家可以试一下前面介绍的“一月效应”，重新编程，进行回溯测试。

这种扩展不仅是基于时间，而且可以基于任何可获取的数据。

例如：

- 如果是连续第n个跌停板/涨停板，则采用xx策略。
- 如果开盘价/收盘价/xx指数超过xx指标，则采用xx策略。

这些策略的具体编写专业性太强，对于初学者而言，可能有些复杂。但策略函数的入口在这里，水平提高后可自行深入学习。

6.3.2 案例6-2：多个时间点交易

案例6-2是多个时间点交易的代码，相关程序代码请参见程序文件名：zq602_mulTrade.py。

运行结果如下：

```
runfile('E:/zwPython/zw_k10x/zq602_mulTrade.py', wdir='E:/zwPyth
Reloaded modules: zwQTDraw, zwSys, zwQTBox, zwBacktest, zwTools,
```

```
2011-01-03 2011-12-30
stkCode ['aeti', 'egan', 'glng', 'simo']
            open  high   low  close  volume  adj close  dprice
date
2011-12-23  4.57  4.58  4.40   4.55   23500       4.55    4.57
2011-12-27  4.59  4.60  4.41   4.60   33300       4.60    4.59
2011-12-28  4.47  4.65  4.47   4.64   32400       4.64    4.47
2011-12-29  4.65  5.34  4.48   5.15   73900       5.15    4.65
2011-12-30  5.00  5.16  4.85   5.08   51400       5.08    5.00
            open  high   low  close  volume  adj close  dprice
date
2011-12-23  6.83  7.00  6.83   6.91   16500       6.91    6.83
2011-12-27  6.97  7.12  6.79   6.92   23700       6.92    6.97
2011-12-28  6.85  6.95  6.76   6.87   13200       6.87    6.85
2011-12-29  6.89  7.00  6.83   6.91   12000       6.91    6.89
2011-12-30  6.96  7.01  6.85   6.91   17600       6.91    6.96
            open       high        low       close      volume
date
2011-12-23  45.380001  45.590000  44.910000  45.119999  278900
2011-12-27  45.110001  45.209999  44.540001  44.590000  299200
2011-12-28  44.849998  44.849998  43.500000  43.529999  456800
2011-12-29  43.669998  44.750000  43.669998  44.450001  340400
2011-12-30  44.779999  44.880001  43.820000  44.450001  526600
            open      high       low      close    volume   adj
date
2011-12-23  20.82  21.100000  20.719999  21.030001  396500  18.9
2011-12-27  21.09  21.600000  20.750000  21.480000  664100  19.3
```

```
2011-12-28  21.50  21.540001  20.350000  20.410000  515900  18.4
2011-12-29  20.40  20.980000  20.400000  20.920000  233300  18.8
2011-12-30  20.92  20.920000  20.320000  20.480000  280800  18.4
nday 362



::qxUsr
date,2011-12-30; stkVal,408569.33642; cash,867669.253688; dret,0

::qxUsr.stk inv01
{'glng': 11095}
::xtrdLib tmp\inv01_xtrdLib.csv
         date           ID   mode  code    dprice      num     k
0  2011-01-03  inv01_000001   buy  glng  12.100265  11095.0  12.
1  2011-01-03  inv01_000002   buy  simo   3.918672  17293.0   3.
2  2011-01-04  inv01_000003   buy  aeti   2.230000    500.0   2.
3  2011-01-04  inv01_000004  sell  glng  12.227803  -1000.0  12.
4  2011-01-04  inv01_000005  sell  simo   4.017992 -17293.0   4.
5  2011-01-05  inv01_000006  sell  aeti   2.160000   -500.0   2.
6  2011-01-05  inv01_000007   buy  glng  11.988668   1000.0  11.

::qxLib.head tmp\inv01_qxLib.csv
         date         stkVal           cash      dret
0  2011-01-03  204260.741218  797981.976553  0.002243  1.002243e
1  2011-01-04  122784.351515  878577.921831 -0.000878  1.001362e
2  2011-01-05  135755.923625  867669.253688  0.002060  1.003425e
3  2011-01-06  138851.328770  867669.253688  0.003085  1.006521e
```

```
4  2011-01-07  140797.014540  867669.253688  0.001933  1.008466e

::qxLib.tail
          date         stkVal           cash      dret
247  2011-12-23  414727.727120  867669.253688  0.001220  1.28239
248  2011-12-27  409856.167805  867669.253688 -0.003799  1.27752
249  2011-12-28  400112.993700  867669.253688 -0.007627  1.26778
250  2011-12-29  408569.336420  867669.253688  0.006670  1.27623
251  2011-12-30  408569.336420  867669.253688  0.000000  1.27623

交易总次数：7
交易总盈利：276238.59

盈利交易数：4
盈利交易金额：552590.87
亏损交易数：3
亏损交易金额：-276352.28

最终资产价值 Final portfolio value: $1276238.59
最终现金资产价值 Final cash portfolio value: $867669.25
最终证券资产价值 Final stock portfolio value: $408569.34
累计回报率 Cumulative returns: 27.62 %
平均日收益率 Average daily return: 0.100 %
日收益率方差 Std. dev. daily return:0.0081

夏普比率 Sharpe ratio: 1.572,（0.05利率）
无风险利率 Risk Free Rate: 0.05
```

```
夏普比率 Sharpe ratio: 1.961,（0利率）

最大回撤率 Max. drawdown: 8.4762 %
最长回撤时间 Longest drawdown duration: 91
回撤时间(最高点位) Time High. drawdown:  2011-12-23
回撤最高点位 High. drawdown: 1282396.981
回撤最低点位 Low. drawdown: 1267782.247

时间周期 Date lenght: 362 (Day)
时间周期（交易日） Date lenght(weekday): 252 (Day)
开始时间 Date begin: 2011-01-03
结束时间 Date lenght: 2011-12-30

项目名称 Project name: inv01
策略名称 Strategy name: timXTrade
```

如图6-11所示是以上程序的输出图形。

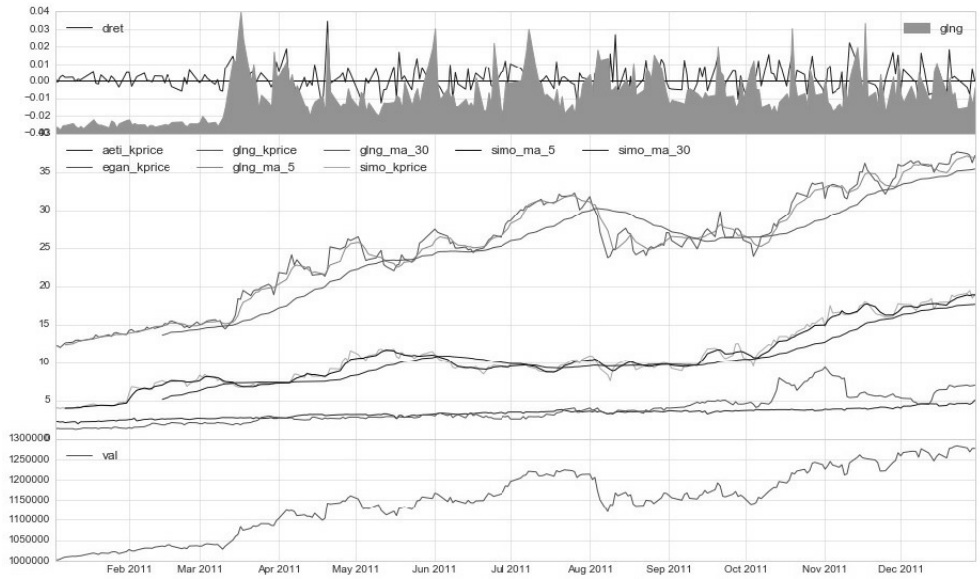


图6-11　案例6-2的输出图形

案例6-2多个时间点的交易的运行结果与案例6-1差不多，最大的差异在结果数据中xtrdLib的交易订单记录。

案例6-2中有效交易订单数是7个，有关数据如下：

```
::xtrdLib tmp\inv01_xtrdLib.csv
         date            ID  mode  code     dprice      num
0  2011-01-03  inv01_000001   buy  glng  12.100265  11095.0  12.
1  2011-01-03  inv01_000002   buy  simo   3.918672  17293.0   3.
2  2011-01-04  inv01_000003   buy  aeti   2.230000    500.0   2.
3  2011-01-04  inv01_000004  sell  glng  12.227803  -1000.0  12.
4  2011-01-04  inv01_000005  sell  simo   4.017992 -17293.0   4.
5  2011-01-05  inv01_000006  sell  aeti   2.160000   -500.0   2.
6  2011-01-05  inv01_000007   buy  glng  11.988668   1000.0  11.
```

案例6-1运行结果中的有效交易只有两个订单：

```
::xtrdLib tmp\inv01_xtrdLib.csv
         date            ID  mode  code     dprice     num     kp
0  2011-01-03  inv01_000001   buy  glng  12.203890  11095.0  12.2
1  2011-01-03  inv01_000002   buy  simo   3.981876  17293.0   3.9
```

案例6-1与案例6-2其他的细节差异，限于篇幅，这里就不展开了，请大家自己对照分析。

6.4　不同数据源与格式修改

zwQuant量化软件的数据源导入，主要由zwQTBox.py模块的stkLibRd函数完成。

前面，我们说过要对stkLibRd进行扩展，以便支持索引文件作为参数，以便一次性、成组导入大批量的股票数据源。

升级版的stkLibRd函数，相关代码如下：

```
def stkLibRd(xlst,rdir):
    zw.stkLib={}    #全局变量，相关股票的交易数据
    zw.stkLibCode=[]  #全局变量，相关股票的交易代码
    #
    x0=xlst[0];
    if x0.find('@')==0:
        fss=x0[1:];#print('fss',fss);  #fss=_rdatInx+fs0
        flst=pd.read_csv(fss, dtype=str,encoding='gbk')
        xlst=list(flst['code'])
        #print(xlst)
    for xcod in xlst:
        fss=rdir+xcod+".csv";
        xfg=os.path.exists(fss);
        if xfg:
            df10=pd.read_csv(fss,index_col=0,parse_dates=[0]);
```

```
            df10=df2zwAdj(df10)
            zw.stkLib[xcod]=df10.sort_index();
            zw.stkLibCode.append(xcod);
```

升级版的stkLibRd函数，流程图如图6-12所示。

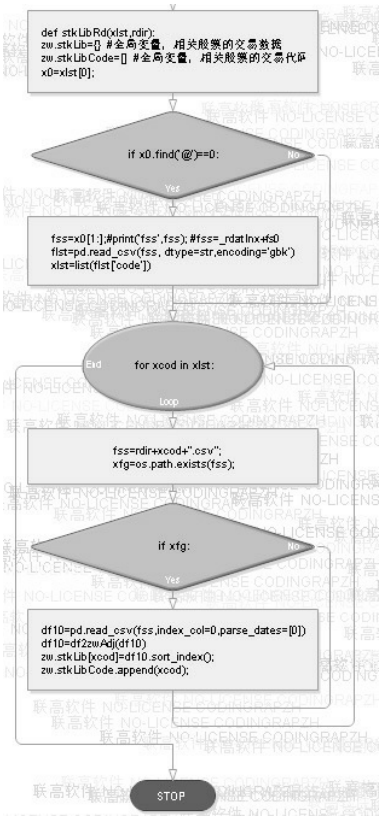


图6-12　升级版stkLibRd函数流程图

升级版的stkLibRd函数除支持使用索引文件作为参数，一次性导入大批量的股票数据源以外，还增加了容错支持，以及其他方面的细节修改。

读取索引文件，使用以下脚本：

```
flst=pd.read_csv(fss, dtype=str,encoding='gbk')
```

此处使用了很少用的“dtype=str”参数选项，因为A股股票代码是数值，例如，万科的股票代码是000002，如果不加此选项，默认读入的是数字2，则后面的文件名会出错。

此外，解码格式使用的是“encoding='gbk'”，而不是默认的utf-8。

这是因为Excel对于utf-8格式的csv文件支持不好，中文显示是乱码，而且目前没有有效的解决办法。

所以，zwQuant涉及中文内容的数据文件，只有zwDat的inx目录下的索引文件采用的是gbk格式，其他文件（如日交易数据），全部是英文、数字，采用的是utf-8格式。

没有全部采用gbk格式是因为gbk格式的英文、数字文件也是双字节保存，比utf-8格式大一倍。

6.4.1 案例6-3：数据源修改

案例6-3演示说明如何使用索引文件作为参数，一次性导入大批量的股票数据源。

程序文件名：\zwpython\zw_k10\zq603_stkLib_ext.py，全部代码如下：

```
# -*- coding: utf-8 -*-
import numpy as np
import pandas as pd

import zwSys as zw         #zw.zwQuant
import zwTools as zwt
import zwQTBox as zwx

#======================

xlst=['aeti','egan','glng','simo']
zwx.stkLibRd(xlst,'dat\\');
print(zw.stkLibCode);
xcod=zw.stkLibCode[0]
d2=zw.stkLib[xcod]
print('\nd2',xcod)
print(d2.head())

#
xlst=['@dat\\inx_code.csv']
zwx.stkLibRd(xlst,'\\zwDat\\cn\\xday\\');
print(zw.stkLibCode);
xcod=zw.stkLibCode[0]
d2=zw.stkLib[xcod]
print('\nd2',xcod)
print(d2.head())
```

运行结果如下：

```
runfile('E:/zwPython/zw_k10x/zq603_stkLib_ext.py', wdir='E:/zwPy
Reloaded modules: zwSys, zwTools, zwQTBox
['aeti', 'egan', 'glng', 'simo']

d2 aeti
            open  high   low  close  volume  adj close
date
2011-01-03  2.23  2.23  2.23   2.23       0       2.23
2011-01-04  2.23  2.23  2.23   2.23       0       2.23
2011-01-05  2.16  2.23  2.16   2.16    3200       2.16
2011-01-06  2.20  2.20  1.90   2.12   97100       2.12
2011-01-07  2.09  2.20  2.09   2.17    4900       2.17
['000001', '000002']

d2 000001
             open    high     low   close     volume  adj close
date
1994-01-03  837.70  840.65  831.66  833.90  101005600     833.90
1994-01-04  835.97  836.97  829.89  832.69   65274300     832.69
1994-01-05  829.30  847.05  823.10  846.98   89412100     846.98
1994-01-06  850.78  869.33  850.78  869.33  184511700     869.33
1994-01-07  875.18  883.99  873.01  879.64  168688400     879.64
```

案例6-3使用传统的股票代码，导入数据源，运行后，stkLibCode股票池代码变量保存的股票代码是：

```
['aeti', 'egan', 'glng', 'simo']
```

使用文件名作为变量值时，参数设置如下：

```
xlst=['@dat\\inx_code.csv']
```

需要注意，目前只支持单个索引文件，需要在索引文件名前加特殊字符“@”，表示是索引文件。

运行后，stkLibCode股票池代码变量保存的股票代码是：

```
['000001', '000002']
```

不同用户，因为'\zwDat\cn\xday\目录下的文件不同，内容可能会有所不同。需要说明的是，即使索引文件有股票代码，如果数据文件不存在，则相关股票代码的数据是无法导入的。

6.4.2 数据源格式修改

金融产品的各种数据尽管都是OHLC格式，但因为排列、名称不同，所以无法直接使用，为此，极宽Quant量化软件在导入数据源时，使用了一个格式转换函数：df2zwAdj()。

```
def df2zwAdj(df0):
     #date,open,high,close,low,volume,amount
     clst=["open","high","low","close","volume","adj close"];
     df2 =pd.DataFrame(columns=clst)
     df0 =df0.rename(columns={'Date':'date','Open':'open','High'
     #df0=df0.round(decimals=2)

     df0['date']=df0.index;
     df2['date']=df0['date'];
     df2['open']=df0['open'];df2['high']=df0['high'];
     df2['low']=df0['low'];df2['close']=df0['close'];
     df2['volume']=df0['volume'];
     #'adj close'
     ksgn='adj close'
     if ksgn in df0.columns:
         df2[ksgn]=df0[ksgn]
     else:
         df2[ksgn]=df0['close'];

     #----index
     df2=df2.set_index(['date'])
```

```
    return df2
```

df2zwAdj()格式转换函数可以直接把中国国内股市和Yahoo雅虎财经的美股数据，统一转换为极宽Quant量化软件的内部格式。其他量化软件内部也有自己的数据格式，请大家使用时注意。

df2zwAdj()格式转换函数代码的有关细节，对于初学者而言，可能有些难度，目前可以无须掌握。

使用tick、1分钟、5分钟分时数据的中高频的用户，或期货、外汇、石油、贵重金属等其他金融产品的用户可以参考以上代码，对数据源进行预处理，再导入极宽Quant量化软件，就可直接使用该软件对各种量化策略进行回溯分析。

6.5 金融数据包与实盘数据更新

zwDat金融数据包是极宽开源量化项目的重要组成部分，已经发布很久了，目前我们对zwDat进行了部分优化升级，可以称为zwDat 2.0，升级的部分包括：

- 修改了多个模块和down_stk数据下载程序相关代码。
- 中国A股数据源更新，对交易数据进行了升级。
- 国内大盘指数文件更新，主要是对inx目录下的文件重新进行了整理。
- 国内大盘指数全部统一采用六位的股票代码，相关数据增加到1994年开市起。
- 由于zwQuant量化软件的dataPre数据预处理模块可以快速对数据进行归一化处理，zwDat取消了极宽归一化数据格式，数据包所占空间减少50%。

注意

美股日线数据下载函数，因为没有去重、追加部分，每次需从头下载所有数据。所以，本次升级没有更新美股数据，有需要的用户可以自己下载美股数据。做美股实盘的用户，请参考A股下载代码，修改相关函数，增加数据去重、追加功能。

6.5.1 大盘指数文件升级

zwDat 2.0金融数据包变化最大的是大盘指数文件，原来的大盘指数文件采用字母、数字混用，只有19种，而且都只有近三年的数据。

此次对大盘指数数据下载进行了多处升级：

- 采用了全新的独立数据抓取函数。
- 大盘指数种类增加到24类。
- 指数名称统一采用6位数字模式。
- 大盘数据最早增加到1994年股市开市起。
- 索引文件增加了起始时间栏目tim0。
- 所有指数、起始时间，全部采用人工手动验证。

指数文件，保存在目录下：

```
x:\zwDat\cn\xday\
```

指数索引文件名是：E:\zwDat\inx\ inx_code.csv。

相关内容如表6-1所示。

表6-1 指数索引文件

code	name	tim0
000001	上证指数	1994-01-01
000002	A 股指数	1994-01-01
000003	B 股指数	1994-01-01
000008	综合指数	1994-01-01
399001	深证成指	1994-01-01
399002	深成指 R	1994-01-01
399003	成份 B 指	1994-01-01
000009	上证 380	2011-01-01
000010	上证 180	2000-01-01
000011	基金指数	2001-01-01
000012	国债指数	2003-01-01
000016	上证 50	2004-01-01
000017	新综指	2006-01-01
000300	沪深 300	2005-01-01

续表

code	name	tim0
399004	深证 100R	2003-01-01
399005	中小板指	2006-01-01
399006	创业板指	2010-01-01
399100	深证新指数	2006-01-01
399101	中小板综	2005-01-01
399106	深证综指	2000-01-01
399107	深证 A 指	2000-01-01
399108	深证 B 指	2000-01-01
399333	中小板 R	2006-01-01
399606	创业板 R	2010-01-01

6.5.2 实盘数据更新

下面我们介绍一下任何进行日线的实盘数据更新。

zwDat金融数据包内置的down_stk数据下载程序，本身就自带了数

据更新、追加去重等功能。最新版本的down_stk数据下载程序目录是：

```
x:\zwPython\zwQuant\down_stk\
```

和

```
x:\zwDat \down_stk\
```

因为zwDat金融数据包是一个独立的开源项目，所以用户下载极宽Quant量化软件、zwDat金融数据包，都可获得down_stk数据下载程序。

中国股票数据追加、更新，运行：zw_down_cnSTK.py。

中国股市大盘数据追加、更新，运行：zw_down_cnSTK_inx.py。

6.5.3 案例6-4：A股实盘数据更新

案例6-4以中国股票数据追加、更新为例，介绍实盘数据的更新。

日线数据更新一般在每天15点收市后进行：

- 运行zwPy35.bat，启动spyder。
- 打开zw_down_cnSTK.py文件。
- 运行zw_down_cnSTK.py即可。

案例6-4实盘数据的更新调用zwDat金融数据包内置的down_stk数据下载程序。

脚本文件名：zwDat\down_stk\zw_down_cnSTK.py，代码如下：

```
# -*- coding: utf-8 -*-
import os
import pandas as pd
import tushare as ts

#zwQuant
import zwSys as zw
import zwQTBox as zwx

#----------

#
def zw_stk_down_all(qx,xtyp):
    fss=qx.rdatInx+'stk_code.csv';print(fss);
    dinx = pd.read_csv(fss,encoding='gbk')

    i=0;xn9=len(dinx['code']);
    for xc in dinx['code']:
        i+=1;
        code="%06d" %xc
        #code=zwTools.v2sk(xc,6);
```

```
            print("\n",i,"/",xn9,"code,",code)
            #---
            qx.code=code;
            zwx.down_stk_cn010(qx,xtyp);

#----------

qdat=zw.zwDatX(zw._rdatCN);
#qdat.prDat();

#zw_stk_down_all(qdat,'T')
zw_stk_down_all(qdat,'0')
```

无须修改任何代码，直接运行即可。即使中间漏了数日，以后追加数据时也会自动补充。

案例6-4支持tick、1分钟数据、5分钟数据抓取，不过需要用户参考tushare金融数据模块修改相关的程序代码。操作期货、外汇等其他金融产品的用户可以参考以上代码，相关的数据源接口API自行修改即可。

需要说明的是tick、1分钟、5分钟等中高频数据属于实时数据，建议采用双平台运行模式，一台电脑采用循环模式不断更新数据源，数据源采用服务器共享目录方式。其他的电脑通过读取服务器的数据源进行实盘分析。

有了实时的数据就可以进行实盘数据分析，推荐根据策略分析的结果、股票代码、金融产品，以及买进、卖出操作方向等细节。

6.5.4 案例6-5：大盘指数更新

下面我们介绍一下，任何进行日线的大盘指数更新。

zwDat发布的down_stk本身就自带了大盘数据更新、追加去重等功能。最新版本的down_stk目录是：

```
x:\zwPython\zwQuant\down_stk\
```

中国股票数据追加、更新，运行：zw_down_cnSTK.py。

中国股市大盘数据追加、更新，运行：zw_down_cnSTK_inx.py。

大盘日线数据更新，一般在每天15点收市后进行：

- 运行zwPy35.bat，启动spyder。
- 打开zw_down_cnSTK_inx.py文件。
- 运行zw_down_cnSTK_inx.py即可。

案例6-5调用zwDat金融数据包内置的down_stk数据下载程序。

脚本文件名：zwDat\down_stk\zw_down_cnSTK_inx.py，相关代码如下：

```
# -*- coding: utf-8 -*-
```

```
import os
import numpy as np
import pandas as pd
import tushare as ts

#  zwQuant
import zwSys as zw
import zwQTBox as zwx
import zwTools as zwt

#----------

def zw_stk_down_inx(qx):
     fss=qx.rdatInx+'inx_code.csv';print(fss);
     dinx = pd.read_csv(fss,encoding='gbk')

     xn9=len(dinx['code']);
     for i in range(xn9):
     #for xc,xtim0 in dinx['code'],dinx['tim0']:
         d5=dinx.iloc[i]
         xc=d5['code'];xtim0=d5['tim0']
         i+=1;code="%06d" %xc
         print("\n",i,"/",xn9,"code,",code,xtim0)
         #---
         qx.code=code;
         zwx.down_stk_cn020inx(qx,xtim0)
```

```
#----------
qx=zw.zwDatX(zw._rdatCN);
qx.prDat();


#
zw_stk_down_inx(qx);
```

6.6 稳定第一

量化分析的参数有很多：最终资产价值，年收益、平均日收益率、日收益率方差、夏普指数、最大回撤率、最大回撤时间等。

以上参数从多个方面衡量投资策略的好坏。

其中最基本的参数，也是最重要的参数，即收益率。

如图6-13所示是案例4-2中的日收益率图，横坐标是交易日，一共252个交易日（0-251）。

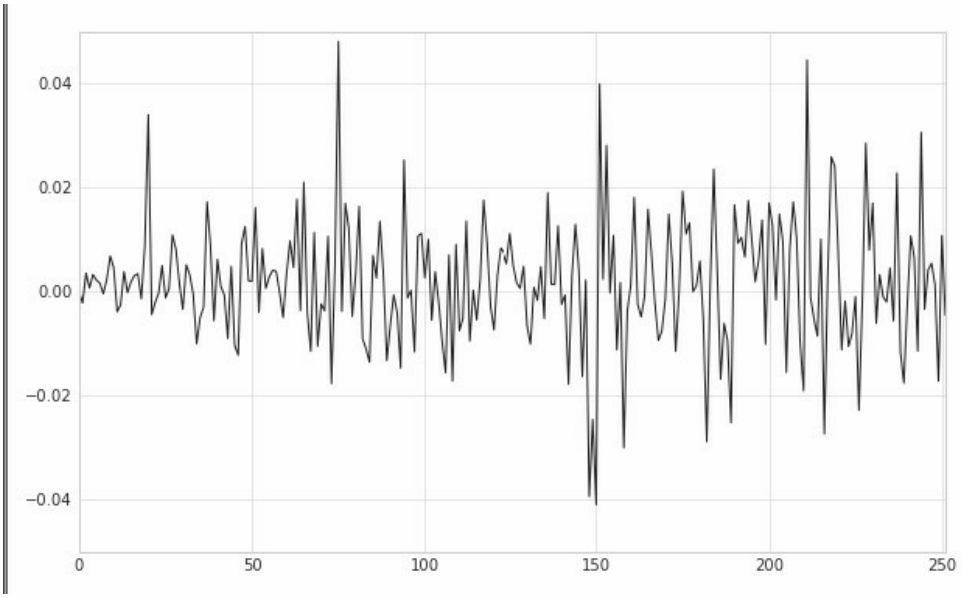


图6-13　日收益率截图

图6-13是最常见的日收益率波动曲线，上下波动的极限在正负5%之间，大部分波动幅度在正负2%之间。而我们学习的目的就是超越这种行业平均水平的波动，争取更加平稳的收益率。

所有的交易策略，无论是量化交易还是传统交易，目的都是盈利率。比盈利率更重要的是稳定的盈利率。盈利率高低无所谓，关键是稳定。

第7章
量化策略库

本章将详细介绍zwQuant（极宽Quant）量化软件内置的量化策略库，以及各种策略函数的使用，本章涉及的策略有SMA均线策略、CMA均线交叉策略、WAP策略和BBands布林带策略。

同时，本章还会结合实盘数据，对zwQuant量化软件和AT量化软件，在采用同样的数据源和量化策略的情况下进行对标测试，用实盘数据检验zwQuant量化软件的准确性。

7.1　量化策略库简介

zwQuant量化软件的策略模块库，文件名是zwStrategy.py，其与传统量化架构不同，zwQuant量化软件的策略库，采用的是1+1模式：策略数据预处理函数+策略分析函数。

zwQuant量化软件这种把数据预处理和策略分析分别采用独立函数模式，这在学术上和量化软件实践上是一种创新，所以zwQuant量化软件是基于矩阵运算的第三代量化系统。

这种矩阵模式可以简化设计逻辑，充分利用Numpy、Pandas数据分析等模块库，在底层、多核甚至GPU方面的优化可以大大提高回溯速度，同时，也便于未来GPU版本的移植。

7.1.1　量化系统的三代目

量化系统的三代目：

第一代，以西蒙斯为代表的古典量化派，采用传统数学、物理公式，从数学统计、理论物理、能量热学熵理论等角度，对金融产品进行量化分析，大部分采用人工计算，没有专业的金融量化分析软件。

第二代，在古典量化学派的基础上，结合现代电脑技术，采用专业的平台，对金融产品进行定量分析，如QSTK金融函数库、Q语言、R语言、MATLAB金融工具箱，以及部分早期的Python量化项目。

第三代，始于2014年摩根、高盛的雅典娜项目和黑石计划，全面向Python转型，基于Pandas等现代数据分析平台，以矢量化矩阵运算为基础，并结合GPU优化速度，全新一代的纯Python矩阵量化系统。极宽Quant量化系统属于第三代量化系统的拓荒者与尝鲜者。

zwQuant策略模块库程序源码较长，请大家参看模块文件zwStrategy.py。

模块文件zwStrategy.py已经内置了以下几组常用的策略函数：

- SMA均线策略
- CMA均线交叉策略
- VWAP策略
- BBAnds布林带策略

模块文件zwStrategy.py的代码看起来很长，但其架构很简单。

- 通用数据预处理函数：sta_dataPre0xtim(qx,xnam0)。
- 各种策略的数据预处理函数+策略分析函数。

7.1.2 通用数据预处理函数

下面我们先介绍一下通用数据预处理函数：sta_dataPre0xtim(qx,xnam0)，相关代码如下：

```
def sta_dataPre0xtim(qx,xnam0):
```

```
    ''' 策略参数设置子函数，根据预设时间，裁剪数据源stkLib

    Args:
        qx (zwQuantX): zwQuantX数据包
        xnam0 (str):  函数标签

    '''

    #设置当前策略的变量参数
    qx.staName=xnam0
    qx.rfRate=0.05;  #无风险年收益，一般为0.05(5%)，计算夏普指数等需要
    #qx.stkNum9=20000;    #每手交易，默认最多20000股
    #
    #按指定的时间周期，裁剪数据源
    xt0k=qx.staVars[-2];xt9k=qx.staVars[-1];
    if (xt0k!='')or(xt9k!=''):
        #xtim0=parse('9999-01-01');xtim9=parse('1000-01-01');
            #xtim0=xtim0.strftime('%Y-%m-%d');xtim9=xtim9.strft
        if xt0k!='':
            if qx.xtim0<xt0k:qx.xtim0=xt0k;
        if xt9k!='':
            if qx.xtim9>xt9k:qx.xtim9=xt9k;
        qx.qxTimSet(qx.xtim0,qx.xtim9)
        zwx.stkLibSet8XTim(qx.xtim0,qx.xtim9);#    print('zw. s
    #============
    #---设置qxUsr用户数据
    qx.qxUsr=zwx.qxObjSet(qx.xtim0,0,qx.money,0);
```

为便于理解，我们提供了数据预处理函数：sta_dataPre0xtim(qx,xnam0)的流程图，如图7-1所示。

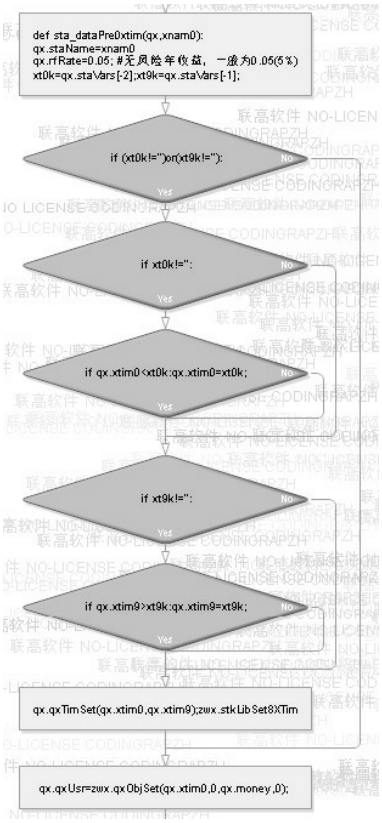


图7-1　数据预处理函数流程图

通用数据预处理函数sta_dataPre0xtim(qx,xnam0)内置的说明是：策略参数设置子函数，根据预设时间，裁剪数据源stkLib，运行流程如下：

- 根据输入参数xname设置策略名称。
- 设置默认的无风险年收益rfRate等参数。
- 按策略参数变量列表staVars指定的时间切割数据源。
- 设置qxUsr用户变量，初始数据。

注意

staVars策略变量列表中，最后两位是数据源：起始时间和结束时间参数，所以采用staVars[-2]、staVars[-1]反向提取模式。时间参数可以为空字符串，此时不切割数据源，采用数据源默认的起始、结束时间。

7.2　SMA均线策略

SMA均线策略很简单：

- 如果调整收盘价高于SMA(15)，则输入多头位置（下单买进）。
- 如果调整收盘价低于SMA(15)，则退出多头位置（空头，下单卖出）。

SMA简单均线策略源自PAT量化软件的内置案例，即本书案例4-4，文件名是：k404_tur04ed.py。

7.2.1　案例7-1：SMA均线策略

案例7-1源码较长，请大家参看程序文件：\zwpython\zw_k10\zq701_k404sma.py。

虽然案例7-1源码较长，但程序本身代码很短，文件尾部附加的注解比代码本身还长。因为案例7-1和本章其他案例一样，都是Quant量化软件和PAT量化软件的对标测试案例，所以，在代码尾部附有原PAT量化软件案例程序的输出结果，便于对比。

案例7-1运行结果，如图7-2所示。

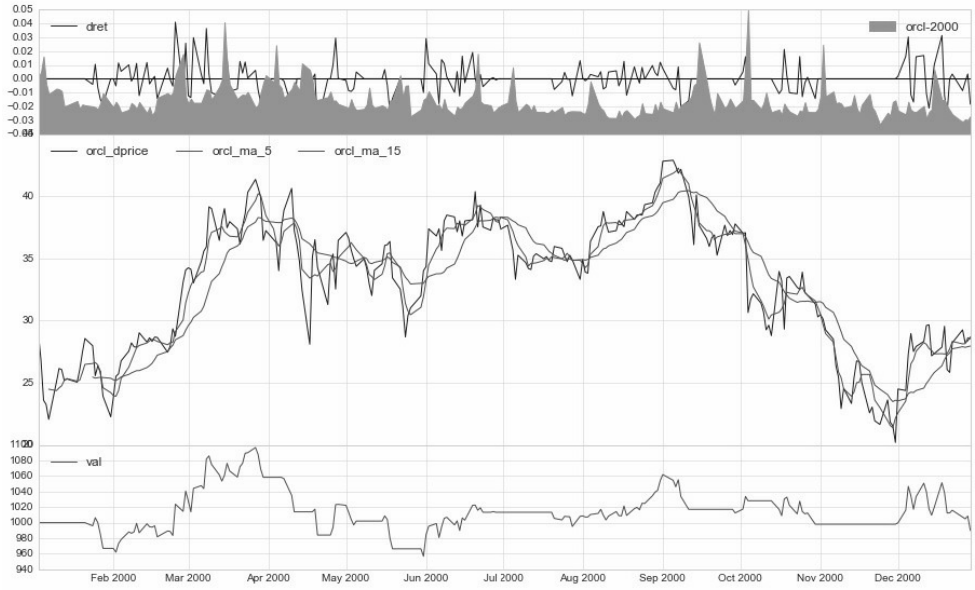


图7-2　SMA均线策略

案例7-1其他输出数据如下：

```
2000-01-03 2000-12-29

stkCode ['orcl-2000']
               open      high      low    close    volume  adj c
date
2000-12-22  30.3750  31.984381  30.0000  31.8750  35568200  29.6
2000-12-26  31.5000  32.187500  30.0000  30.9375  20589500  28.7
2000-12-27  30.3750  31.062500  29.3750  30.6875  26437500  28.4
2000-12-28  30.5625  31.625000  30.3750  31.0625  25053600  28.8
2000-12-29  30.8750  31.312500  28.6875  29.0625  31702200  26.9
nday 362
buy 2000-01-24 27.9765942422 25.4546377333 0
sell 2000-01-28 23.9136040317 25.4169096 10
```

```
……
……

sell 2000-12-20 26.0611849035 27.2065610667 10
buy 2000-12-22 28.2087655412 27.7289456 0

::qxUsr
date,2000-12-29; stkVal,269.89868; cash,719.943864547; dret,-0.0

::qxUsr.stk sma
{'orcl-2000': 10}
::xtrdLib tmp\sma_xtrdLib.csv
          date          ID  mode       code     dprice   num
0   2000-01-24  sma_000001   buy  orcl-2000  27.976594  10.0  25
1   2000-01-28  sma_000002  sell  orcl-2000  23.913604 -10.0  22
。。。。。。
。。。。。。

41  2000-12-20  sma_000042  sell  orcl-2000  26.061185 -10.0  25
42  2000-12-22  sma_000043   buy  orcl-2000  28.208766  10.0  29

::qxLib.head tmp\sma_qxLib.csv
         date  stkVal    cash  dret     val  downLow  downHigh
0  2000-01-03     0.0  1000.0   0.0  1000.0   1000.0    1000.0
1  2000-01-04     0.0  1000.0   0.0  1000.0   1000.0    1000.0
2  2000-01-05     0.0  1000.0   0.0  1000.0   1000.0    1000.0
3  2000-01-06     0.0  1000.0   0.0  1000.0   1000.0    1000.0
```

```
4  2000-01-07     0.0  1000.0   0.0  1000.0   1000.0    1000.0

::qxLib.tail
          date     stkVal        cash      dret          val
247  2000-12-22  296.01791  719.943865  0.003440  1015.961775  9
248  2000-12-26  287.31150  719.943865 -0.008570  1007.255365  9
249  2000-12-27  284.98979  719.943865 -0.002305  1004.933655  9
250  2000-12-28  288.47236  719.943865  0.003465  1008.416225  9
251  2000-12-29  269.89868  719.943865 -0.018419   989.842545  9

交易总次数：43
交易总盈利：-10.16

盈利交易数：22
盈利交易金额：1564.25
亏损交易数：21
亏损交易金额：-1574.41

最终资产价值 Final portfolio value: $989.84
最终现金资产价值 Final cash portfolio value: $719.94
最终证券资产价值 Final stock portfolio value: $269.90
累计回报率 Cumulative returns: -1.02 %
平均日收益率 Average daily return: 0.001 %
日收益率方差 Std. dev. daily return:0.0098

夏普比率 Sharpe ratio: -0.310,（0.05利率）
无风险利率 Risk Free Rate: 0.05
```

```
夏普比率 Sharpe ratio: 0.012,（0利率）


最大回撤率 Max. drawdown: 12.7556 %
最长回撤时间 Longest drawdown duration: 277
回撤时间(最高点位) Time High. drawdown:  2000-03-27
回撤最高点位 High. drawdown: 1096.641
回撤最低点位 Low. drawdown: 956.758


时间周期 Date lenght: 362 (Day)
时间周期（交易日） Date lenght(weekday): 252 (Day)
开始时间 Date begin: 2000-01-03
结束时间 Date lenght: 2000-12-29


项目名称 Project name: sma
策略名称 Strategy name: sma
```

7.2.2 实盘下单时机与推荐

我们学习量化的目的是为了实盘交易，通过zwQuant量化分析软件进行实盘操作，或者通过zwQuant量化软件提供实时的实盘股票代码，进行买入、卖出操作指示，有了具体的指示数据，后端无论是人工下单，还是导入第三方交易程序都是轻而易举的事情。而这个交易数据就在类变量qx内置的xtrdLib变量中。

在案例7-1中，xtrdLib变量的最后一行就是我们所需要的最新交易

指示数据。

```
::xtrdLib tmp\sma_xtrdLib.csv
          date           ID  mode       code     dprice   num

37  2000-10-18  sma_000038  sell  orcl-2000  29.311578 -10.0  33
38  2000-10-19  sma_000039   buy  orcl-2000  33.432611  10.0  33
39  2000-10-30  sma_000040  sell  orcl-2000  31.343073 -10.0  30
40  2000-12-01  sma_000041   buy  orcl-2000  24.494031  10.0  24
41  2000-12-20  sma_000042  sell  orcl-2000  26.061185 -10.0  25
42  2000-12-22  sma_000043   buy  orcl-2000  28.208766  10.0  29
```

在后面的案例中我们会采用最新的实盘数据，更新数据源。因为案例7-1的数据只有2000年的，所以最后的交易日期是2000-12-22。大家可以采用国内最新的股市数据进行真正的实盘案例策略分析。

因为本书读者大部分是做国内股票的，所以本书是基于A股实盘日线数据进行回溯测试。使用tick、1分钟、5分等分时数据的中高频交易用户，以及美股、外汇、期货等其他金融产品的用户，请参考zwQuant量化软件数据源部分相关的程序代码，自行修改数据源。

在实盘回溯测试前，请用户在股市收盘后定时更新日线数据，相关操作在前面的章节已经讲解过，在此不再赘述。

7.2.3 案例7-2：实盘SMA均线策略

案例7-2实盘SMA均线策略将对案例7-1进行一些修改：

- 采用国内最新的A股实时更新数据作为实盘数据源。
- 更新bt_endRets输出函数，增加每日交易推荐。

案例7-2的源码请参看程序文件：\zwpython\zw_k10\zq702_sma_xt.py。

由案例7-2程序代码可以看出，数据源设置部分代码如下（股票代码：603020，是爱普股份。）：

```
xlst=['603020']    #爱普股份
```

为此，特意对bt_endRets(qx)输出函数的中间绘图标识进行了修改：

```
zwdr.dr_quant3x(qx,xcod,'val',kmid8,'aipu')
```

改为爱普股份的拼音“aipu”。

数据源方面直接采用zwDat金融数据包作为数据源：

```
qx=zwbt.bt_init(xlst,'\\zwdat\\cn\\day\\','sma2',1000);
```

tick、1分钟等分时数据，以及外汇、期货等，请自行修改，配置数

据源。

此外，在结果数据输出函数bt_endRets(qx)尾部增加了以下代码：

```
print('')
print('每日交易推荐')
print('::xtrdLib',qx.fn_xtrdLib)
#print(self.xtrdLib.tail())
print(qx.xtrdLib.tail())
```

用于输出：每日交易推荐。

如图7-3所示是运行截图。

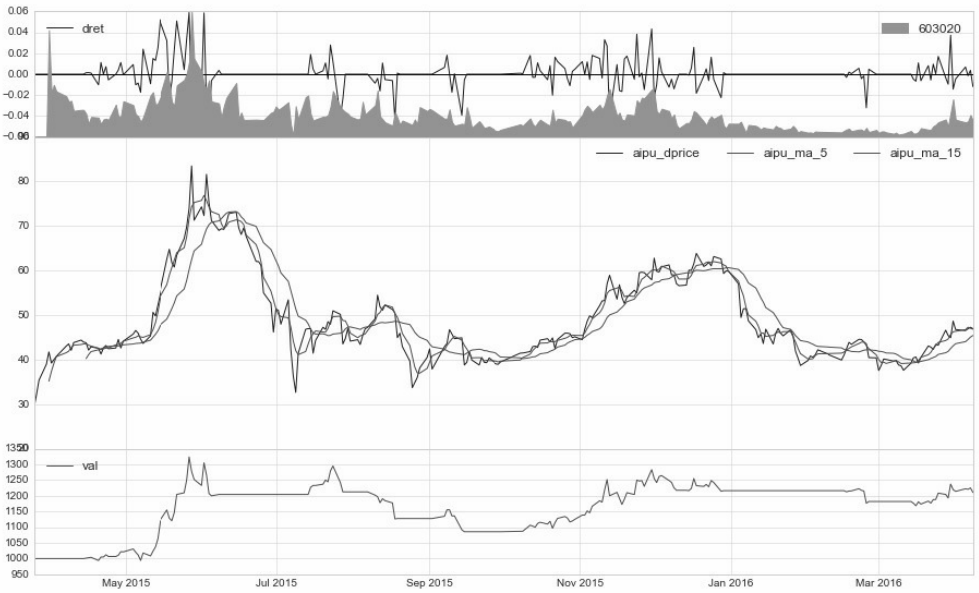


图7-3　实盘SMA均线策略运行图

案例7-2其他输出数据如下：

交易总次数：29
交易总盈利：210.50

盈利交易数：14
盈利交易金额：814.60
亏损交易数：15
亏损交易金额：-604.10

最终资产价值 Final portfolio value: $1210.50
最终现金资产价值 Final cash portfolio value: $749.40
最终证券资产价值 Final stock portfolio value: $461.10
累计回报率 Cumulative returns: 21.05 %
平均日收益率 Average daily return: 0.084 %
日收益率方差 Std. dev. daily return:0.0138

夏普比率 Sharpe ratio: 0.737,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 0.965,（0利率）

最大回撤率 Max. drawdown: 17.9950 %
最长回撤时间 Longest drawdown duration: 317
回撤时间（最高点位）Time High. drawdown: 2015-05-27
回撤最高点位 High. drawdown: 1323.700
回撤最低点位 Low. drawdown: 1085.500

时间周期 Date lenght: 381 (Day)

```
时间周期（交易日） Date lenght(weekday): 256 (Day)
开始时间 Date begin: 2015-03-25
结束时间 Date lenght: 2016-04-08

项目名称 Project name: sma2
策略名称 Strategy name: sma
```

案例7-2输出数据的最后几行是每日股票推荐信息：

```
每日交易推荐
::xtrdLib tmp\sma2_xtrdLib.csv
          date        ID  mode    code  dprice   num  kprice  su
24  2016-02-17 sma2_000025 buy   603020  42.50  10.0  43.95  439
25  2016-02-26 sma2_000026 sell  603020  40.55 -10.0  40.50 -405
26  2016-03-15 sma2_000027 buy   603020  40.50  10.0  40.69  406
27  2016-03-17 sma2_000028 sell  603020  39.35 -10.0  40.60 -406
28  2016-03-18 sma2_000029 buy   603020  40.60  10.0  43.15  431
```

案例7-2在bt_endRets(qx)函数尾部增加了部分代码，即每日交易推荐，其中各字母含义如下。

- date：交易日期，如果策略运行当日有推荐数据，则会出现在xtrdLib最后一行。
- mode：交易模式，buy为买进；sell为卖出。
- code：交易的股票代码。

- num：每次交易的股票数目，用户可以自己修改相关的策略函数，或在此集成上进行倍乘法交易。
- dprice：策略分析采用的价格。
- kprice：实际成交价，因为次日的价格尚未出来，所以默认采用当前的收盘价。

任何策略都不是每天都有推荐，当推荐数据的最后一行不是当前日期时，表明当天没有交易推荐，这个类似于足彩的每日推荐。

其他输出参数回溯周期的资金回报率是：

```
累计回报率 Cumulative returns: 21.05 %
```

这个回报率还是相当不错的，特别是在2015年80%的股民亏损的情况下。但该量化策略能不能作为通用的投资策略还需要进行多只股票的回溯。

如果硬件计算能力支持，可以类似于“一月效应”模式，对不同股票组合进行分组测试。例如全部股票、沪深300、上证50、创业版指数等。

同时，对于多个时间段进行组合测试。这些测试有些难度，而且需要专业的测试软件框架配合，属于专业领域。

目前阶段大家需要知道的是，任何量化策略必须结合多种股票组合和多个时间周期的测试才能用于实盘。

例如案例7-2的策略，回报率是21%，可同样的策略在案例7-1中是亏损的，回报率是-2.5%。

```
累计回报率 Cumulative returns: -2.55 %
```

7.3 CMA均线交叉策略

均线是将每天的收盘价加权平均，从而得到一条带有趋势性的轨迹。均线系统是大多分析者常用的技术工具，从技术角度看是影响技术分析者心理价位的决定因素，是技术分析者良好的参考工具，相比价格变化是滞后的。

均线指标实际上是移动平均线指标的简称。由于该指标是反映价格运行趋势的重要指标，其运行趋势一旦形成，将在一段时间内继续保持，趋势运行所形成的高点或低点又分别具有阻挡或支撑作用，因此均线指标所在的点位往往是十分重要的支撑或阻力位，这就为我们提供了买进或卖出的有利时机，均线系统的价值也正在于此。

7.3.1 案例7-3：均线交叉策略

案例7-3对标的PAT量化软件原程序是案例4-6：k406_sma_crossover_sample.py。

为了便于对标测试，我们对案例4-6进行了修改，在不影响功能的前提下，增加了部分中间数据输出，修改后的案例代码文件是：k406x_sma_crossover_sample.py。本节对标案例指的就是修改后的案例：k406x。

CMA均线交叉操作策略也很简单，当均线和日线数据交叉时：

- 均线趋势朝上，买进。
- 均线趋势朝下，卖出。

案例7-3程序源码，请参见代码文件：\zwpython\zw_k10\zq703_k406cma.py。

案例7-3对于zwQuant量化软件内置的策略库模块zwStrategy.py内的均线交叉策略、均线交叉数据预处理函数进行了升级。同样，zwQuant处于不断升级优化中，以最新发布的软件代码为准。

案例7-3运行结果如图7-4所示。



图7-4　CMA均线交叉策略运行结果

案例7-3其他输出数据如下：

```
交易总次数：2
交易总盈利：558441.42
```

盈利交易数：2

盈利交易金额：558441.42

亏损交易数：0

亏损交易金额：0.00

最终资产价值 Final portfolio value: $1558441.42

最终现金资产价值 Final cash portfolio value: $1558441.42

最终证券资产价值 Final stock portfolio value: $0.00

累计回报率 Cumulative returns: 55.84 %

平均日收益率 Average daily return: 0.094 %

日收益率方差 Std. dev. daily return:0.0106

夏普比率 Sharpe ratio: 1.109,（0.05利率）

无风险利率 Risk Free Rate: 0.05

夏普比率 Sharpe ratio: 1.405,（0利率）

最大回撤率 Max. drawdown: 15.5823 %

最长回撤时间 Longest drawdown duration: 326

回撤时间(最高点位) Time High. drawdown: 2012-09-19

回撤最高点位 High. drawdown: 1794192.432

回撤最低点位 Low. drawdown: 1558441.422

时间周期 Date lenght: 729 (Day)

时间周期（交易日） Date lenght(weekday): 502 (Day)

开始时间 Date begin: 2011-01-03

结束时间 Date lenght: 2012-12-31

```
项目名称 Project name: cma
策略名称 Strategy name: cma
```

7.3.2 对标测试误差分析

在案例7-3运行结果中：

```
最终资产价值 Final portfolio value: $1558441.42
累计回报率 Cumulative returns: 55.84 %
```

与此对标的案例4-6修改版本k406x，运行结果是：

```
最终资产价值 Final portfolio value: $1565183.40
累计回报率 Cumulative returns: 56.52 %
```

同样，案例7-1运行结果中：

```
最终资产价值 Final portfolio value: $989.84
```

PAT量化软件对标案例4-4：k404_tur04ed.py运行结果是：

```
最终资产价值 Final portfolio value: $974.53
```

以上数据说明，zwQuant量化软件与PAT量化软件的对标测试，在数值上略有差别。存在这种差别的原因有很多，例如价格体系细节方面的选择、浮点数的位数与精度、佣金的扣除方式、交易限额，以及交易时机的选择。

不过，对于案例7-3而言，最大的差异可能源自函数：cross_Mod(qx)，均线交叉趋势判断函数。

cross_Mod(qx)均线交叉趋势判断函数位于zwQTBox.py模块，相关代码如下：

```
def cross_Mod(qx):
     kma='ma_%d' %qx.staVars[0]
     xbar=qx.xbarWrk;
     dma,ma2n=xbar[kma][0],xbar['ma2n'][0]
     dp,dp2n=xbar['dprice'][0],xbar['dp2n'][0]
     #
     kmod=-9;
     if  (dp>dma)and(dp2n<ma2n)and(dp>dp2n):
         kmod=1;
         #print(kmod,'xbar',xbar)
```

```
    elif (dp<dma)and(dp2n>ma2n)and(dp<dp2n):
        kmod=-1;
        #print(kmod,'xbar',xbar)

    return kmod
```

为了便于理解，大家可以参考cross_Mod(qx)均线交叉趋势判断函数的流程图，如图7-5所示。

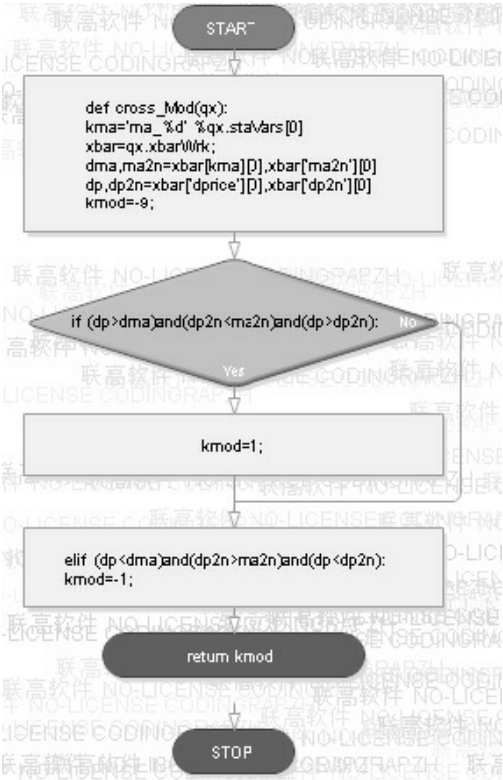


图7-5　cross_Mod(qx)均线交叉趋势判断函数流程图

对于初学者而言，理解cross_Mod均线交叉趋势判断函数可能有些难度，可以无须了解函数内部的细节，知道如何调用函数的功能即可。

案例7-3使用的cross_Mod均线交叉趋势判断函数与PAT量化软件对

应的函数差别很大，笔者对其进行了简化，但功能都是一样的。

在案例7-3的结果中，交易订单数据是：

```
::xtrdLib tmp\cma_xtrdLib.csv
        date          ID  mode       code  dprice      num  kpri
0  2011-11-28  cma_000001   buy  aapl-201x  49.760  18086.0  49.
1  2012-10-25  cma_000002  sell  aapl-201x  80.987 -18086.0  80.
```

对标案例4-6（k406）中也有一段对应的debug数据：

```
2011-11-28 00:00:00 strategy [INFO] above: $1000000.00,$49.76,48
2011-11-29 00:00:00 strategy [INFO] BUY at $49.72,$48.67,$100708
。。。。。。
2012-10-25 00:00:00 strategy [INFO] below: $100708.79,$80.99,81.
2012-10-26 00:00:00 strategy [INFO] Sell at $49.72,$81.06,$15651
```

对应的脚本是：

```
self.info("above: $%.2f,$%.2f,%.2f,%2d" % (dcash,dprice,self.__s
self.info("BUY at $%.2f,$%.2f,$%.2f" % (execInfo.getPrice(),self

self.info("below: $%.2f,$%.2f,%.2f" % (dcash,dprice,self.__sma[-
```

```
self.info("Sell at $%.2f,$%.2f,$%.2f" % (execInfo.getPrice(),sel
```

将两者对比可以发现，cross_mod均线交叉趋势判断函数与价格模式不同，案例7-3与对标的案例4-6（k406）存在以下异同：

- 买入、卖出交易时间是相同的，都是2011-11-29买入，2012-10-25卖出。
- 案例703的买入价kprice是：49.374，案例k406的above dprice买入价是：49.76。
- 案例703的卖出价kprice是：80.251，案例k406的bellow dprice卖出价是：80.99。
- 在案例k406x中，execInfo.getPrice返回值并非交易价格，所以采用策略分析的价格作为参考。

对于zwQuant量化软件和PAT量化软件的价格体系而言，最终采用的价格无论是当天的收盘价，或是次日的开盘价、平均价，作为实际成交价都是一种预设价格，这种预设价格与实盘价格是有差异的。因为实盘价格始终处于动态变化中，即使是基于tick、1分钟线的高频交易也有“滑点”漏单的情况。实盘中任何订单的买卖，特别是大额订单都会影响股票、期货等金融产品的实盘价格。回溯测试永远无法100%再现实盘交易。

根据大数原理，从长期来看，这些细微的差异是可以忽略的，回溯测试的模拟与实盘的吻合度超过99%，完全可以用于量化策略的分析。

所以，zwQuant量化软件与PAT等量化系统，由于架构的不同和价格模式方面的角度差异，所引发的回溯结果之间的细微差异属于正常情

况，如果数值差别不大，则可以接受。但是，如果出现较大的差异，就需要查明相关原因。

7.3.3 案例7-4：CMA均线交叉策略修改版

案例7-4是基于实盘的CMA均线策略回溯测试，是根据案例7-3修改的。因为案例7-3中CMA的均线默认周期是163日，案例7-2中的股票603020（爱普股份）数据不够，所以案例7-3改用股票601999（出版传媒）作为测试数据源。

在本书编写时，zwDat金融数据包数据源已经更新到2016-4-8，如果用户自行更新，部分输出数据可能有所差异，属于正常情况。

案例7-4是修改版的均线交叉策略，源码较长，请读者参看程序文件：\zwpython\zw_k10\zq704_cma_xt.py。

案例7-4主要运行结果如下：

```
最终资产价值 Final portfolio value: $6415.28
最终现金资产价值 Final cash portfolio value: $6415.28
最终证券资产价值 Final stock portfolio value: $0.00
累计回报率 Cumulative returns: -35.85 %
平均日收益率 Average daily return: -0.073 %
日收益率方差 Std. dev. daily return:0.0120
```

```
夏普比率 Sharpe ratio: -1.220,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: -0.958,（0利率）



每日交易推荐
::xtrdLib tmp\cma2_xtrdLib.csv
         date         ID  mode    code  dprice    num  kprice
0  2015-08-10 cma2_000001  buy   601999  13.81  651.0  13.68   8
1  2015-08-20 cma2_000002  sell  601999  12.26 -651.0  11.04  -7
2  2015-11-12 cma2_000003  buy   601999  13.18  565.0  13.71   7
3  2015-11-27 cma2_000004  sell  601999  12.30 -565.0  12.05  -6
4  2015-12-07 cma2_000005  buy   601999  13.04  506.0  12.71   6
5  2015-12-11 cma2_000006  sell  601999  12.19 -506.0  12.44  -6
6  2015-12-15 cma2_000007  buy   601999  12.60  514.0  12.79   6
7  2016-01-04 cma2_000008  sell  601999  11.44 -514.0  11.25  -5
```

7.3.4 人工优化参数

进行量化分析首先需要注意的就是回报率：

```
累计回报率 Cumulative returns: -35.85 %
```

由于2015年是股灾之年，上面代码中的-35.85%的收益率属于正常情况。对于2015年这种波动大的市场，采用163日作为均线，有些不合适。在这里，我们可以采用人工的方式逐一测试5~120日，作为均线的回报率。

修改代码第75行和第76行即可：

```
#qx.staVars=[163,'2014-01-01','']
qx.staVars=[30,'2014-01-01','']
```

把第75行原案例数据屏蔽，用第76行作为测试行。

测试结果如图7-6所示，保存在文件：dat\zq704xt.xls中，测试表明，20日均线回报率最高，达到74%。

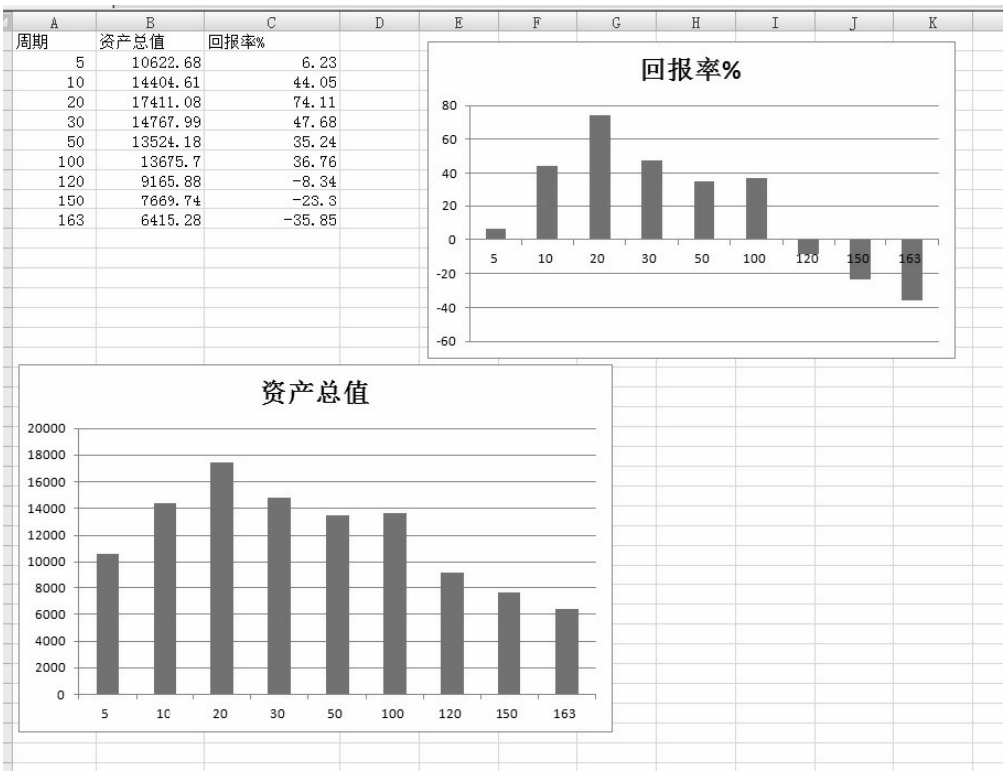


图7-6　参数优化

20日均线的CMA均线交叉策略高达74%的回报率，这已经是非常高的数字了，一般的量化交易有10%~20%的稳定回报率就已经不错了。

当然这个只是单一股票，2014—2016年4月的回溯，真正的回溯测试必须是多只股票、多个时间周期的测试。如果全靠手工进行测试显然是不现实的，必须采用专业的测试框架，并且对测试结果进行统计分析，然后对参数和操作策略进行优化。

这些已经涉及数据分析、统计建模、参数优化等领域，属于专业领域，了解一下就可以了。本章使用的这种手工测试方法是一种最基本的、应急测试手段，即在没有专业工具的情况下，通过人工记录测试数据对有关参数进行优化。

7.4 VWAP策略

VWAP策略又称作成交量加权平均价策略。传统上，VWAP、TWAP、PEG算法等被动型交易策略主要用于大盘资金、庄家的操作策略，个人用户一般用的很少。

对于规模较大的证券交易，如果一次性全部按市价下单，则该交易会造成巨大的市场冲击；如果分成几笔，在不同时间段内成交，投资者又面临市场价格和流动性发生不利变动的时间风险。不过，近年也有个人或机构将VWAP策略用于实盘投资中。

7.4.1 案例7-5：VWAP策略

案例7-5对标的PAT量化软件原案例程序文件名是：k408x_vwap_momentum.py。

为了便于对标测试，我们对案例4-8进行了修改，在不影响功能的前提下，增加了部分中间数据输出，修改后的案例代码文件是：k408x_vwap_momentum.py。

本节对标案例k408指的就是修改后的案例：k408x。

案例7-5源码较长，请大家自己参考程序文件：\zwpython\zw_k10\zq705_k408vwap.py。

案例7-5运行结果如图7-7所示。

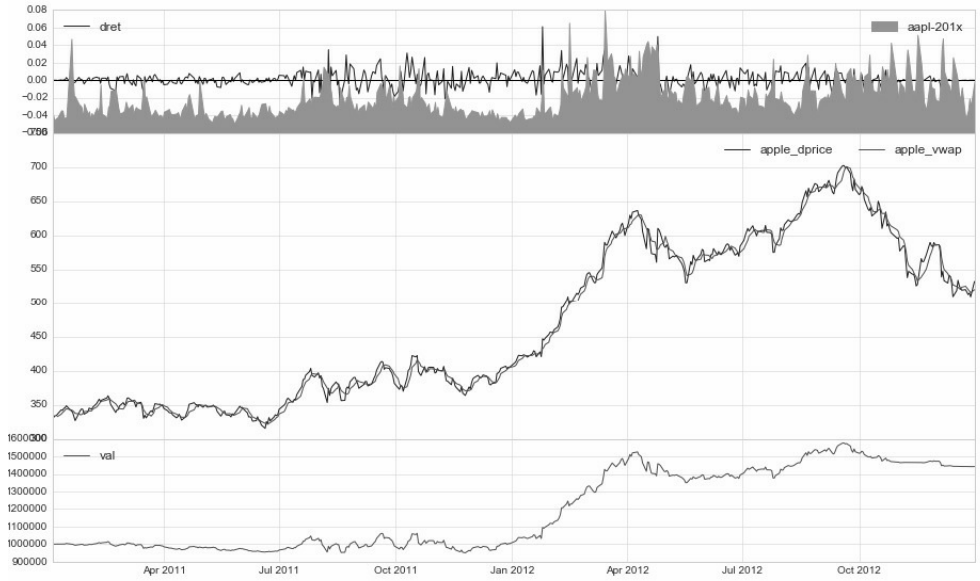


图7-7 VWAP策略

案例7-5其他输出数据有：

```
最终资产价值 Final portfolio value: $1441776.00
最终现金资产价值 Final cash portfolio value: $1441776.00
最终证券资产价值 Final stock portfolio value: $0.00
累计回报率 Cumulative returns: 44.18 %
平均日收益率 Average daily return: 0.078 %
日收益率方差 Std. dev. daily return:0.0103
夏普比率 Sharpe ratio: 0.895,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 1.199,（0利率）

最大回撤率 Max. drawdown: 11.5174 %
最长回撤时间 Longest drawdown duration: 153
回撤时间(最高点位) Time High. drawdown:  2012-09-19
```

```
回撤最高点位 High. drawdown: 1576163.027
回撤最低点位 Low. drawdown: 1441776.005

时间周期 Date lenght: 729 (Day)
时间周期（交易日） Date lenght(weekday): 502 (Day)
开始时间 Date begin: 2011-01-03
结束时间 Date lenght: 2012-12-31

项目名称 Project name: vwap
策略名称 Strategy name: vwap
```

对照以上数据可以发现，案例7-5与对标程序案例4-8的输出完全一致。

7.4.2 案例7-6：实盘VWAP策略

案例7-6是基于实盘的VWAP回溯测试案例。

程序源码参见文件：\zwpython\zw_k10\zq706_vwap_xt.py。

案例7-6主要运行结果如下：

```
最终资产价值 Final portfolio value: $12292.00
最终现金资产价值 Final cash portfolio value: $3634.00
最终证券资产价值 Final stock portfolio value: $8658.00
```

```
累计回报率 Cumulative returns: 22.92 %
平均日收益率 Average daily return: 0.057 %
日收益率方差 Std. dev. daily return:0.0199

夏普比率 Sharpe ratio: 0.298,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 0.456,（0利率）

最大回撤率 Max. drawdown: 27.3932 %
最长回撤时间 Longest drawdown duration: 235
回撤时间(最高点位) Time High. drawdown:  2015-08-17
回撤最高点位 High. drawdown: 15544.000
回撤最低点位 Low. drawdown: 11286.000

时间周期 Date lenght: 829 (Day)
时间周期（交易日） Date lenght(weekday): 553 (Day)
开始时间 Date begin: 2014-01-01
结束时间 Date lenght: 2016-04-08

项目名称 Project name: vwap2
策略名称 Strategy name: vwap
```

由运行结果发现，回报率为22.9%，也是个不错的策略。

7.5　BBands布林带策略

布林带又称布林线（Boll Bands），如图7-8所示，布林线指标是通过计算股价的“标准差”，再求股价的“信赖区间”。该指标在图形上画出三条线，其中上下两条线可以分别看成是股价的压力线和支撑线，而在两条线之间还有一条股价平均线，布林线指标的参数最好设为20。一般来说，股价会运行在压力线和支撑线所形成的通道中。

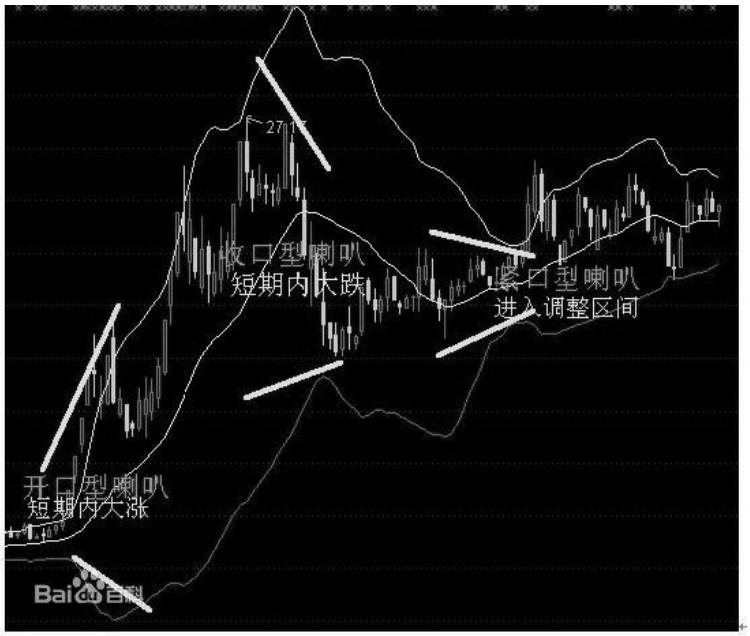


图7-8　BBands布林带策略

在所有的指标计算中，BOLL指标的计算方法是最复杂的指标之一，其中引进了统计学中的标准差概念，涉及中轨线（MB）、上轨线（UP）和下轨线（DN）的计算。

和其他指标的计算方法一样，由于选用的计算周期不同，BOLL指标也包括日BOLL指标、周BOLL指标、月BOLL指标、年BOLL指标以

及分钟BOLL指标等各种类型，经常用于股市研判的是日BOLL指标和周BOLL指标。虽然它们计算时的取值有所不同，但基本的计算方法一样。

7.5.1 案例7-7：BBands布林带策略

案例7-7对标的PAT原程序是：k410_bbands.py。

为了便于对标测试，我们对案例4-8进行了修改，在不影响功能的前提下，增加了部分中间数据输出，修改后的案例代码文件是：k410x_bbands.py，本节对标案例k410，指的就是修改后的案例：k410x。

案例7-7程序源码文件为：\zwpython\zw_k10\zq707_k410bbands.py。

案例7-7运行结果如图7-9所示。

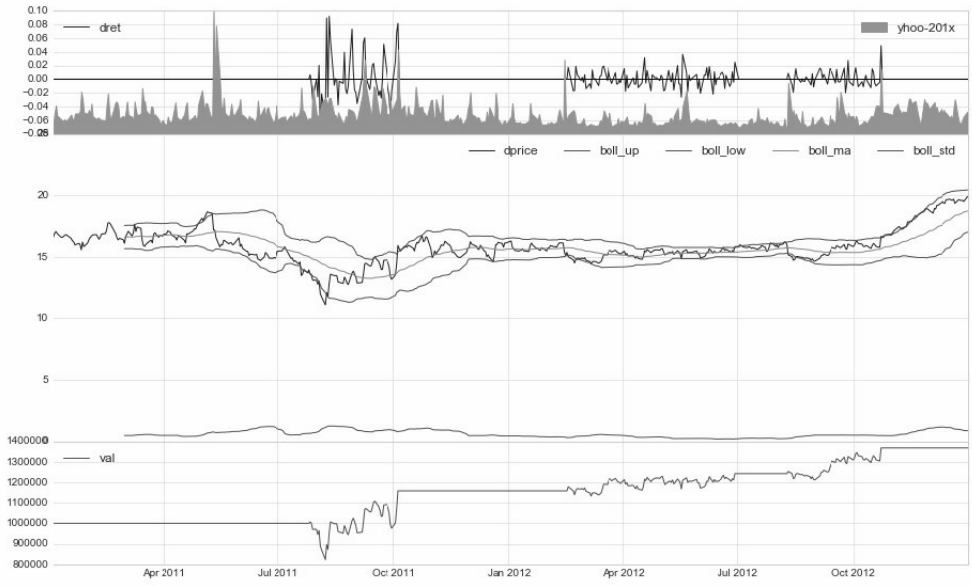


图7-9　BBands布林带策略案例运行结果

案例7-7其他输出信息如下：

```
最终资产价值 Final portfolio value: $1367770.03
最终现金资产价值 Final cash portfolio value: $1367770.03
最终证券资产价值 Final stock portfolio value: $0.00
累计回报率 Cumulative returns: 36.78 %
平均日收益率 Average daily return: 0.071 %
日收益率方差 Std. dev. daily return:0.0136

夏普比率 Sharpe ratio: 0.603,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 0.835,（0利率）

最大回撤率 Max. drawdown: 18.2747 %
最长回撤时间 Longest drawdown duration: 204
```

```
回撤时间（最高点位）Time High. drawdown:  2012-10-23
回撤最高点位 High. drawdown: 1367770.030
回撤最低点位 Low. drawdown: 1367770.030

时间周期 Date lenght: 729 (Day)
时间周期（交易日） Date lenght(weekday): 502 (Day)
开始时间 Date begin: 2011-01-03
结束时间 Date lenght: 2012-12-31

项目名称 Project name: bbands
策略名称 Strategy name: bbands
股票代码列表 Stock list:  ['yhoo-201x']
```

7.5.2 案例7-8：实盘BBands布林带策略

案例7-8是基于实盘的VWAP回溯测试案例。

案例7-8源码文件文件名是：\zwpython\zw_k10\zq708_bbands_xt.py。

案例7-8运行结果如图7-10所示。

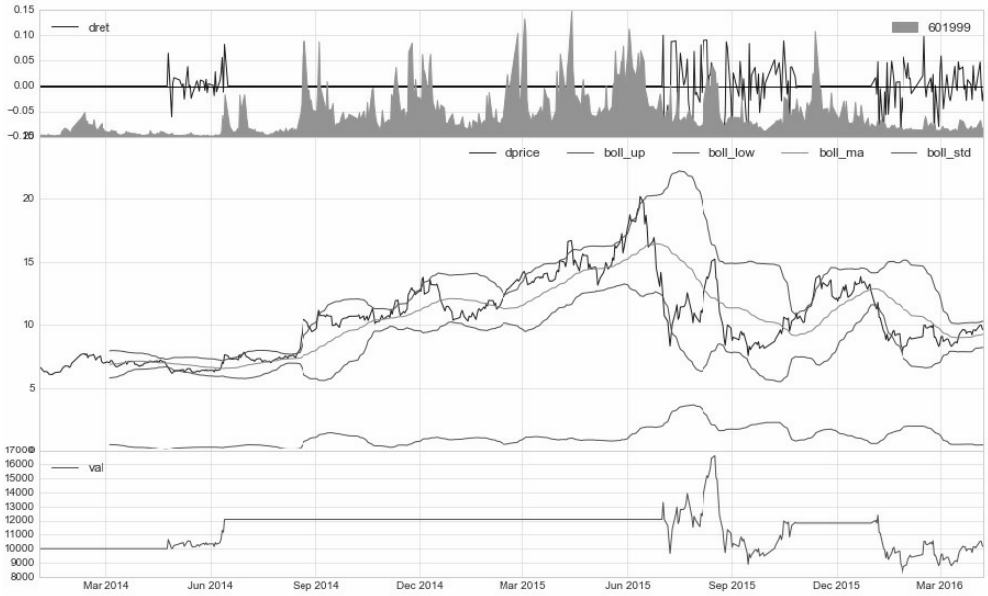


图7-10　BBands布林带策略实盘案例运行结果

案例7-8其他输出信息如下：

```
最终资产价值 Final portfolio value: $10136.89
最终现金资产价值 Final cash portfolio value: $199.43
最终证券资产价值 Final stock portfolio value: $9937.46
累计回报率 Cumulative returns: 1.37 %
平均日收益率 Average daily return: 0.037 %
日收益率方差 Std. dev. daily return:0.0262

夏普比率 Sharpe ratio: 0.104,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 0.225,（0利率）

最大回撤率 Max. drawdown: 50.3748 %
最长回撤时间 Longest drawdown duration: 383
```

```
回撤时间(最高点位) Time High. drawdown:  2015-08-17
回撤最高点位 High. drawdown: 16617.560
回撤最低点位 Low. drawdown: 8246.500

时间周期 Date lenght: 829 (Day)
时间周期（交易日） Date lenght(weekday): 553 (Day)
开始时间 Date begin: 2014-01-01
结束时间 Date lenght: 2016-04-08

项目名称 Project name: bbands2
策略名称 Strategy name: bbands
股票代码列表 Stock list:  ['601999']
```

案例7-8实盘回溯测试的收益率是1.37%，有些偏低，不过在2015—2016年4月出现过几次大的股灾，所以这个结果也算是跑赢大市，可以接受。

7.6　大道至简1+1

zwQuant量化软件没有采用通用的事件模式，走的是创新路线。笔者的口号是：只谈原创。

目前看来，这个创新的路子是对的；开发周期比预计的短，各个模块库代码也比预计的简短。特别是绘图模块，三合一函数的表现功能还是很强的，而且非常灵活，支持MA均线、RSI、布林带、MACD等各种技术指标。

zwQuant量化软件属于第三代矩阵式量化平台，基于GPU并行架构，其效率和逻辑都比传统的量化模型简单、清晰。这种简化对于初学者特别重要，可以节约大量的学习成本。

目前最完整的PAT汉化资料是极宽开源组2015年年底发布的《PyAlgoTrade用户手册·zw中文版.pdf版》，该版本类似“枪版”，可即便如此，依然是目前最完整的PAT汉化参考资料。

zwQuant量化软件看起来简单，可是整个量化流程需要的功能全部具备，而且使用纯函数式的策略库，扩充方便。

大家看完本章就可以套用现成的量化模型进行回溯测试了。保守的读者可以再做半年的模拟盘、实盘练习。激进的读者可以用以前的数据复盘，直接进入实盘操作。

通过本章的学习，大家已经可以直接将zwQuant量化软件用于实盘，所缺的不过是实盘策略，而策略是经验的积累。

大家可以参考传统的投资策略，还可以根据zwQuant、Ta-Lib、mt4等各种金融软件的参数进行创新。

至于，哪种策略好要问市场，要看实盘的盈利率，其他的都是空谈，一切靠数据说话。

第8章
海龟策略与自定义扩展

本章我们以海龟策略为例，说明如何采用自定义的方式，从零开始，一步一步自行扩展zwQuant量化软件的策略库。

8.1 策略库

zwQuant量化软件的策略模块库文件名是：zwStrategy.py，模块中的量化策略分析函数采用1+1模式，即每个策略一组两个函数。其中一个为策略数据初始化函数，另一个为策略分析函数。

这种1+1策略函数组模式是zwQuant量化软件独创的，具有策略编写简单、逻辑清晰、运行效率极高的特点。用户可参考相关源码，自行编写、扩充策略函数库。

极宽Quant量化软件的策略模块库目前的版本是v1.0，包括以下策略函数。

- sta_dataPre0xtim(qx,xnam0)：策略参数设置子函数。
- BBANDS_dataPre(qx,xnam0,ksgn0)：布林带数据预处理函数。
- BBANDS_sta(qx)：布林带策略分析函数。
- CMA_dataPre(qx,xnam0,ksgn0)：均线交叉策略数据预处理函数。
- CMA_sta(qx)：均线交叉策略分析函数。
- SMA_dataPre(qx,xnam0,ksgn0)：简单均线策略数据预处理函数。
- SMA_sta(qx)：简单均线策略分析函数。
- VWAP_dataPre(qx,xnam0,ksgn0)：VWAP数据预处理函数。
- VWAP_sta(qx)：VWAP成交量加权平均价策略分析函数。

8.1.1 自定义策略

下面我们将学习如何从零开始采用自定义的方式自行设计一个新的策略，并加入极宽Quant的策略模块库文件：zwStrategy.py。

量化策略的设计灵感、思路有很多：

- 传统的股票投资策略。
- 各种金融股票、量化网站上面的量化策略。
- TA-Lib等金融函数库，专业参数与组合。
- mt4、Tb、金字塔、同花顺等金融软件的类似选股工具。

8.1.2 海龟投资策略

海龟策略被称为最早的量化投资策略。海龟交易法则源自于海龟实验著名的公开交易系统。

1983年著名的商品投机家理查德·丹尼斯在一个交易员培训班上推广而使该策略闻名于世，它涵盖了交易系统的各个方面。

海龟（操盘手培训）实验是金融史上最著名的实验，在实验后的四年，这些选手取得了年均复利80%的收益。海龟实验证明了交易可以被传授，证明了用一套简单的法则可以使具有很少或根本没有交易经验的人成为优秀的交易员。

简单来说，就是根据实盘经验或者统计分析，找出大概率的正盈利产品，设定好收益点、止损点，然后，严格按照预设的程序，进行操作。海龟交易法则，非常适于计算机量化分析模式，所以在近年的量化投资热浪当中，再一次成为热门模式。

本节我们就以海龟策略为例，说明如何采用自定义的方式自行扩展zwQuant量化软件的策略函数库。

海龟策略经过多年发展，衍生出许多变种，具体模式和实现手段都略有不同。

我们在本书中采用的海龟策略是相对比较简单的一种：

- 当今天的收盘价大于过去N（默认是30）个交易日中的最高价时，以收盘价买入。
- 买入后，当收盘价小于过去N（默认是15）个交易日中的最低价时，以收盘价卖出。

8.2　tur海龟策略v1：从零开始

为了方便广大初学者快速建立策略函数，zwQuant量化软件内置了一个空白的策略函数模版，位于source目录下，文件名是：source\zw_sta00.py。

大家自定义策略函数的第一步，就是将模版文件复制到工作目录下并重新命名，这里我们把zw_sta00.py复制到zw_k10目录下，命名为：zw_k10\zq801_tur_v1.py。

需注意的是，模版文件使用的数据文件是：dat\600401（ST海润）。需要事先把该数据文件复制到zw_k10\dat目录下。

8.3　案例8-1：海龟策略框架

案例8-1程序文件名为：\zwpython\zw_k10\zq801_tur_v1.py。

案例8-1程序代码中“主流程”部分的代码暂时是屏蔽的，属于空策略模版，不过依然可以运行。运行结果如下：

```
runfile('E:/zwPython/zw_k10x/zq801_tur_v1.py', wdir='E:/zwPython
2003-09-24 2016-04-08
stkCode ['600401']
```

输出数据表示当前的数据源股票代码是：600401。数据源起始时间为：2003-09-24；数据源终止时间为：2016-04-08。

另外，大家可以看看相关策略模版的源码，熟悉zwQuant量化软件策略库函数的用户可能会发现，zw_sta00.py空策略函数中的几个关键函数与SMA均线策略非常相似：

- 策略分析sta00
- 数据预处理sta00_dataPre
- 数据输出bt_endRets

为了提供一个参考导引，zwQuant量化的空策略模版基本上就是直接套用最简单的SMA均线策略，作为策略模版，大家在此基础上修改即可。

8.4　tur海龟策略v2：策略初始化

下面，我们把zw_k10\zq801_tur_v1.py复制为一个新文件zw_k10\zq802_tur_v2.py，v1作为原始参考文件，v2作为我们动手修改的第一个文件。

专业开发系统如微软vs、delphi编程平台都内置有版本管理系统，每次存盘均会修改子版本的编号。目前Python语言需要人工进行换名的方式进行管理，但工作量不大。大家在自己编写策略或其他程序时，也要养成这种习惯，即在每次进行较大的修改前采用新的文件名。

需要修改的地方有：

- 文件头备注。
- 策略分析函数名称tur10。
- 数据预处理函数名称tur10_dataPre。
- 主流程参数设置、名称设置、策略绑定等相关部分。
- 取消主流程数据预处理函数的屏蔽。

8.5　案例8-2：策略初始化

案例8-2是根据案例8-1修改的，修改后的新文件名是：\zwpython\zw_k10\zq802_tur_v2.py。

读者查看案例8-2程序代码时需注意的是，目前策略函数tur10、数据预处理tur10_dataPre函数和测试主流程位于同一个文件，函数前面的策略模块库名称的缩写zwsta需要取消，不然会出错。

在正常情况下，修改后函数是可以运行的，案例8-2的运行结果如下：

```
runfile('E:/zwPython/zw_k10x/zq801_tur_v2.py', wdir='E:/zwPython
2003-09-24 2016-04-08
stkCode ['600401']
            open  high   low  close      volume   adj close  dpr
date
2016-04-01  2.54  2.62  2.53   2.60  111421383.0       2.60    2
2016-04-05  2.63  2.68  2.63   2.68  110368219.0       2.68    2
2016-04-06  2.68  2.80  2.66   2.73  171821194.0       2.73    2
2016-04-07  2.74  2.78  2.59   2.59  198460790.0       2.59    2
2016-04-08  2.56  2.65  2.55   2.63  108273174.0       2.63    2
```

以上信息是由数据预处理函数tur_dataPre运行的，以下代码的输

出：

```
print(d20.tail())
```

目的是用于检查数据预处理后的数据源，可以看到，数据源增加了dprice、kprice、ma_30、ma_15等数据列。

8.6　tur海龟策略v3：数据预处理

下面，我们把代码文件另存为一个新文件zw_k10\zq803_tur_v3.py。

在v2版中，我们已经对tur海龟策略对应的函数、参数等环节进行了修改，并且运行了数据预处理函数。

虽然代码能够正确运行，但处理后的数据并非我们所需要的结果。

下面，我们根据tur海龟策略的要求，对tur10_dataPre数据预处理函数进行修改。

我们采用的海龟策略是相对比较简单的一种。根据要求，数据预处理需要提供以下两组数据。

- 过去n日的最高价，数据列名称为：xhigh。
- 过去n日的最低价，数据列名称为：xlow。

8.7　案例8-3：数据预处理

因为v2与v3版本的代码只有tur10_dataPre数据预处理函数完全一致，所以我们只参看该函数的代码即可。

修改后的tur10_dataPre数据预处理函数代码如下：

```
def tur10_dataPre(qx,xnam0,ksgn0):
     zwx.sta_dataPre0xtim(qx,xnam0);
     #----对各只股票数据进行预处理，提高后期运算速度
     ksgn,qx.priceCalc=ksgn0,ksgn0;  #'adj close';
     for xcod in zw.stkLibCode:
         d20=zw.stkLib[xcod];

         #  计算交易价格kprice和策略分析采用的价格dprice，kprice一般采用
         #d20['dprice']=d20['open']*d20[ksgn]/d20['close']
         #d20['kprice']=d20['dprice'].shift(-1)
         d20['dprice']=d20['close']
         d20['kprice']=d20['dprice']
         #
         d=qx.staVars[0];ksgn='xhigh0';d20[ksgn]=pd.rolling_max(
         d=qx.staVars[1];ksgn='xlow0';d20[ksgn]=pd.rolling_min(d
         d20['xhigh']=d20['xhigh0'].shift(1)
         d20['xlow']=d20['xlow0'].shift(1)
         #
```

```
        zw.stkLib[xcod]=d20;
        if qx.debugMod>0:
            print(d20.tail())
            #---
            fss='tmp\\'+qx.prjName+'_'+xcod+'.csv'
            d20.to_csv(fss)
```

tur10_dataPre数据预处理函数，流程图如图8-1所示。

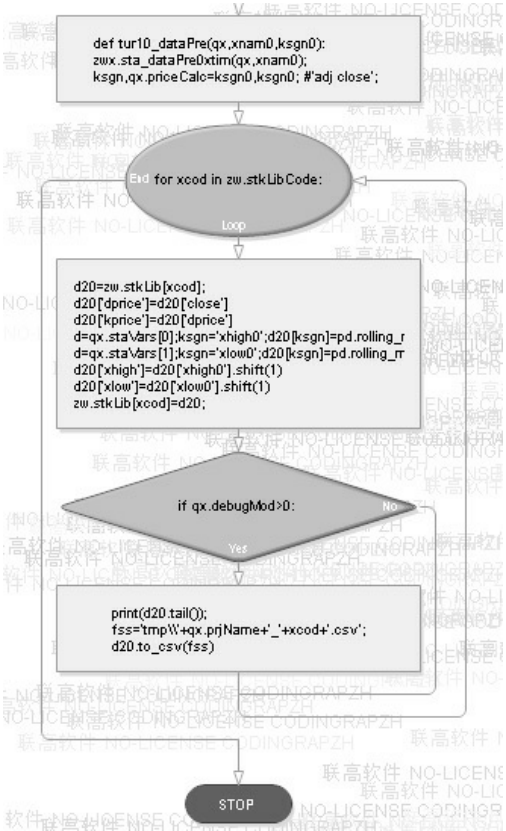


图8-1　tur10_dataPre数据预处理函数流程图

运行案例8-3（zq803_tur_v3.py），结果如下：

```
2003-09-24 2016-04-08
stkCode ['600401']
            open  high   low  close       volume  adj close  dpr
date
2016-04-01  2.54  2.62  2.53   2.60  111421383.0       2.60    2
2016-04-05  2.63  2.68  2.63   2.68  110368219.0       2.68    2
2016-04-06  2.68  2.80  2.66   2.73  171821194.0       2.73    2
2016-04-07  2.74  2.78  2.59   2.59  198460790.0       2.59    2
2016-04-08  2.56  2.65  2.55   2.63  108273174.0       2.63    2
```

从案例8-3的结果中可以看到，我们需要的两组数据xhigh和xlow均正常显示。此外，还有两个多余的数列：xhigh0、xlow0，这是因为xhigh、xlow数据是策略分析前一天的。看起来好像很复杂，但其实相关的脚本很简单，只有几行：

```
d=qx.staVars[0];ksgn='xhigh0';d20[ksgn]=pd.rolling_max(d20['high
d=qx.staVars[1];ksgn='xlow0';d20[ksgn]=pd.rolling_min(d20['low']
d20['xhigh']=d20['xhigh0'].shift(1)
d20['xlow']=d20['xlow0'].shift(1)
```

需要注意的是，xhigh对应的是high数据，xlow对应的是low数据，而不是dprice、close数据列。从上面的代码可以看到极宽Quant架构在策略逻辑、代码编写方面非常简洁。借助Pandas数据分析软件的rolling函

数和矩阵运算，一方面简化了代码编写，另一方面也大大提高了运行速度。

在数据预处理函数中，除了策略需要的特殊函数外，还有一段例程用于计算dprice、kpreice。

zwQuant量化软件的价格体系，参见如下说明：

```
#     priceWrk，策略分析时使用的股票价格，一般是：dprice，复权收盘价
#     priceBuy，买入/卖出的股票价格，一般是：kprice，一般采用次日的复权开盘
#     priceCalc，最后结算使用的股票价格，一般是：adj close，复权收盘价
```

不过，海龟策略对于价格方面有明确的指示：

```
当今天的收盘价大于过去n个交易日中的最高价时，以收盘价买入；
买入后，当收盘价小于过去n个交易日中的最低价时，以收盘价卖出。
```

全部采用收盘价的模式也是初学者常用的一种模式，简化了计算和编程。相关代码如下：

```
ksgn,qx.priceCalc=ksgn0,ksgn0;  #'adj close';
for xcod in zw.stkLibCode:
     d20=zw.stkLib[xcod];
```

```
    # 计算交易价格kprice和策略分析采用的价格dprice,kprice一般采用次日的
    #d20['dprice']=d20['open']*d20[ksgn]/d20['close']
    #d20['kprice']=d20['dprice'].shift(-1)
    d20['dprice']=d20['close']
    d20['kprice']=d20['dprice']
```

8.8　tur海龟策略v4：策略分析

下面，我们把代码文件另存为一个新文件：zw_k10\zq804_tur_v4.py。

在v3版中我们已经完成了数据预处理函数。

我们根据tur海龟策略的要求编写策略分析函数。

8.9 案例8-4：策略分析

因为v4版本的代码只有策略分析函数tur10与v3版完全一致，所以我们只参看该函数的代码即可。

修改后的tur10策略分析函数代码如下：

```
def tur10(qx):
    '''
    海龟策略:tur10
    当今天的收盘价大于过去n个交易日中的最高价时，以收盘价买入；
    买入后，当收盘价小于过去n个交易日中的最低价时，以收盘价卖出。

    '''
    stknum=0;
    xtim,xcod=qx.xtim,qx.stkCode
    dprice=qx.xbarWrk['dprice'][0];
    x9=qx.xbarWrk['xhigh'][0];
    x1=qx.xbarWrk['xlow'][0];
    dcash=qx.qxUsr['cash'];
    dnum0=zwx.xusrStkNum(qx,xcod)

    if dprice>x9:
        if dnum0==0:
            stknum = int(dcash*0.9 /dprice);#dsum=stknum*kprice
```

```
            #stknum = 500
            #print(xtim,stknum,dnum,'++b,%.2f,%.2f,%.2f,$,%.2f,
            #print(xtim,stknum,'++xd',xcod,dprice,x9,x1)
    elif (dprice<x1):
            #stknum = -500
            stknum = -1
            #stknum = -1;dsum=dnum*kprice

    if stknum!=0:
        #print(xtim,stknum,'xd',xcod,dprice,x9,x1)
        pass;

    return stknum
```

tur10策略分析函数流程图如图8-2所示。

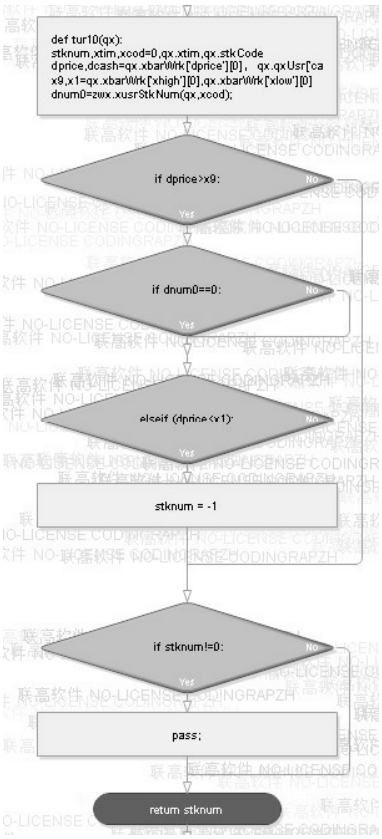


图8-2　tur10海龟策略流程图

量化策略函数设计的一般流程如下。

- 首先，将返回参数stknum成交的股票数目初始化设置为0。
- 通过qx全局变量提取xtim测试时间点、xcod股票代码。
- 提取dprice，策略分析时的股票价格。
- 根据策略，提取其他参数：
    - 在tur10海龟策略中，需要x9近期最高价、x1近期最低价、dcash用户持有的现金数、dnum0用户持有代码为xcod的股票数目。
- 根据策略判断分析，并计算相应的买进、卖出股票数目stknum。
- 返回交易的股票数目：大于0为买进；小于0为卖出；等于0为无

交易操作。

注意，函数尾部的代码：

```
if stknum!=0:
      #print(xtim,stknum,'xd',xcod,dprice,x9,x1)
      pass;
```

为临时测试代码，当发生交易时，根据策略输出相关的中间数据，便于调试。

在正式版本中，为了优化，可以删除以上代码片段。

编写好策略函数后，还需要修改zq804_tur_v4.py代码文件末尾的回溯语句，取消屏蔽。

```
zwbt.zwBackTest(qx)
```

v4版还没有修改输出函数，其输出和v3是一样的，所以这里不再重复。

8.10　tur海龟策略v5：数据图表输出

下面，我们把代码文件另存为一个新文件：zw_k10\zq805_tur_v5.py。在v4版中，我们已经完成了策略分析函数。下面，根据tur海龟策略的要求，修改bt_endRets(qx)数据图表输出函数。

8.10.1　案例8-5：图表输出

因为v5版本的代码，只有bt_endRets(qx)数据图表输出函数与v4版完全一致，所以我们只看该函数的代码即可。

修改后的bt_endRets(qx)数据图表输出函数，代码如下：

```
def bt_endRets(qx):
      #---ok ，测试完毕
      # 保存测试数据，qxlib，每日收益等数据；xtrdLib，交易清单数据
      #qx.qxLib=qx.qxLib.round(4)
      qx.qxLib.to_csv(qx.fn_qxLib,index=False,encode='utf-8')
      qx.xtrdLib.to_csv(qx.fn_xtrdLib,index=False,encode='utf-8')
      qx.prQLib()
      #
```

```
#-------计算交易回报数据
zwx.zwRetTradeCalc(qx)
zwx.zwRetPr(qx)

#-------绘制相关图表，可采用不同的模板
# 初始化绘图模板：dr_quant3x
zwdr.dr_quant3x_init(qx,12,8);
#  设置相关参数
xcod=zw.stkLibCode[0];ksgn=qx.priceBuy;
#xcod='glng';ksgn=qx.priceBuy;
#kmid8=[['aeti',ksgn],['egan',ksgn],['glng',ksgn,'ma_5','ma
kmid8=[[xcod,ksgn,'xhigh','xlow']]
# 绘图
zwdr.dr_quant3x(qx,xcod,'val',kmid8,'')
# 可设置中间图形窗口的标识
#qx.pltMid.legend([]);
#
print('')
print('每日交易推荐')
print('::xtrdLib',qx.fn_xtrdLib)
print(qx.xtrdLib.tail())
#print(qx.xtrdLib)
```

虽然bt_endRets(qx)数据图表输出函数很长，但一般修改的只是kmid中间图表，用于绘图的数据列。

相关的代码是：

```
xcod=zw.stkLibCode[0];ksgn=qx.priceBuy;
#xcod='glng';ksgn=qx.priceBuy;
#kmid8=[['aeti',ksgn],['egan',ksgn],['glng',ksgn,'ma_5','ma_30']
kmid8=[[xcod,ksgn,'xhigh','xlow']]
```

修改完输出函数后，还需要修改zq805_tur_v5.py代码文件末尾的输出语句，取消屏蔽。

```
bt_endRets(qx)
```

至此，案例8-5修改完成，现在可以运行了。案例文件名是：\zwpython\zw_k10\ zq805_tur_v5.py。运行结果如图8-3所示。

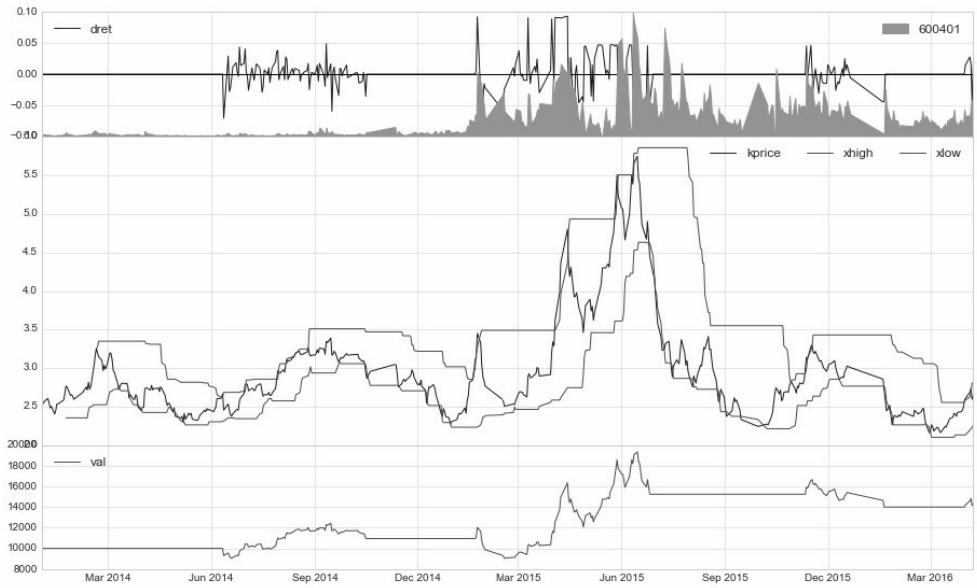


图8-3　海龟策略diy v5版

案例8-5其他输出数据如下：

```
交易总次数：7
交易总盈利：4339.67

盈利交易数：4
盈利交易金额：7545.17
亏损交易数：3
亏损交易金额：-3205.50

最终资产价值 Final portfolio value: $14339.67
最终现金资产价值 Final cash portfolio value: $1400.07
最终证券资产价值 Final stock portfolio value: $12939.60
累计回报率 Cumulative returns: 43.40 %
平均日收益率 Average daily return: 0.095 %
```

日收益率方差 Std. dev. daily return:0.0206

夏普比率 Sharpe ratio: 0.581,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 0.734,（0利率）

最大回撤率 Max. drawdown: 27.6250 %
最长回撤时间 Longest drawdown duration: 298
回撤时间(最高点位) Time High. drawdown: 2015-06-15
回撤最高点位 High. drawdown: 19337.160
回撤最低点位 Low. drawdown: 13995.270

时间周期 Date lenght: 829 (Day)
时间周期（交易日） Date lenght(weekday): 486 (Day)
开始时间 Date begin: 2014-01-01
结束时间 Date lenght: 2016-04-08

项目名称 Project name: tur10
策略名称 Strategy name: tur10
股票代码列表 Stock list: ['600401']

每日交易推荐
::xtrdLib tmp\tur10_xtrdLib.csv

	date	ID	mode	code	dprice	num	kpric
2	2015-01-21	tur10_000003	buy	600401	3.10	3178.0	3.1
3	2015-06-26	tur10_000004	sell	600401	4.45	-3178.0	4.4
4	2015-11-11	tur10_000005	buy	600401	2.98	4601.0	2.9

```
5  2016-01-20  tur10_000006  sell  600401    2.71 -4601.0    2.7
6  2016-03-31  tur10_000007   buy  600401    2.56  4920.0    2.5
```

案例8-5采用的数据源股票代码是：600401（ST海润），最后更新日期是：2016-04-08。

在输出结果中，每日交易推荐是根据最新数据变化的，如果是03-31后运行tur10海龟策略，则完全可以用于实盘。

8.10.2 参数优化

案例zq805_tur_v5.py运行结果中：

```
最终证券资产价值 Final stock portfolio value: $12939.60
累计回报率 Cumulative returns: 43.40 %
```

与此对应的输入参数是：

```
qx.staVars=[35,15,'2014-01-01',''] #  30,15,=14339.67,43.40 %
```

8.10.3 案例8-6：参数优化

下面把代码文件另存为一个新文件：zw_k10\zq806_tur_v6.py，来介绍相关的参数优化环节。因为目前没有专业的回溯参数测试框架，因此只能采用手动的方式修改输入参数，并记录输出结果。

在案例zq805_tur_v5.py中，回溯周期是从2014年开始的，大家知道在2015年6月，中国股市大涨过，如果从2014的数据开始回溯，那么大部分都是正收益，则不利于测试，所以，我们从2015年开始回溯。

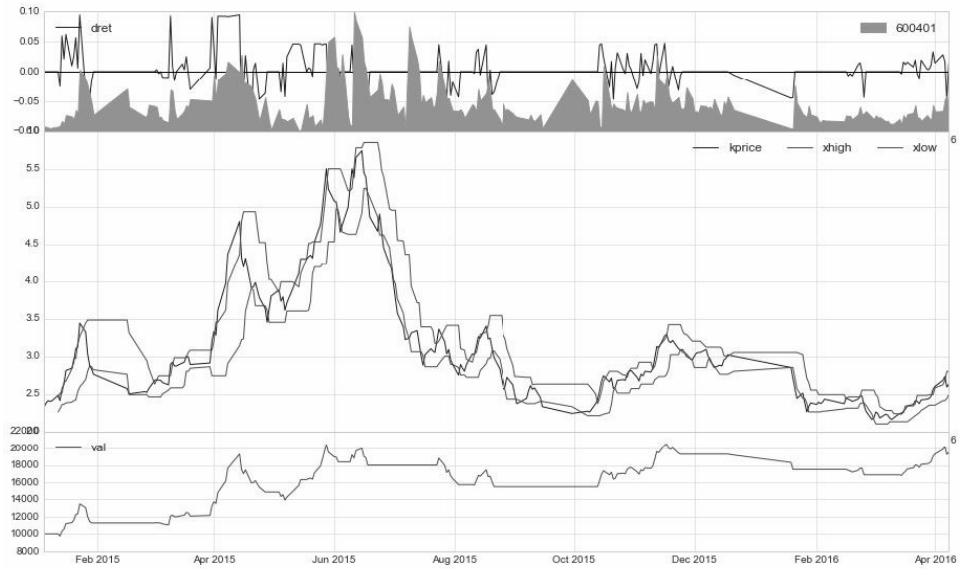
修改后的输入语句是：

```
qx.staVars=[30,10,'2015-01-01',''] #  30,15,= 9325.77, -6.74 %
```

案例8-6程序文件名是：\zwpython\zw_k10\zq806_tur_v6.py。

案例8-6参数的输出如图8-4所示。

```
qx.staVars=[5,5,'2015-01-01',''] #  30,15,= 9325.77, -6.74 %
```



每日交易推荐
::xtrdLib tmp\tur10_xtrdLib.csv

	date	ID	mode	code	dprice	num	kprice	sum	cash
14	2015-12-18	tur10_000015	buy	600401	3.02	5749.0	3.02	17361.98	1931.50
15	2016-01-20	tur10_000016	sell	600401	2.71	-5749.0	2.71	-15579.79	17511.29
16	2016-02-16	tur10_000017	buy	600401	2.45	6432.0	2.45	15758.40	1752.89
17	2016-02-25	tur10_000018	sell	600401	2.35	-6432.0	2.35	-15115.20	16868.09
18	2016-03-14	tur10_000019	buy	600401	2.24	6777.0	2.24	15180.48	1687.61

图8-4　海龟策略参数优化

手工记录的参数输入和对应的输出结果如下：

```
参数1,2,=最终资产价值 ,回报率
30,10,=$9325.77, -6.74 %
20,10,=$12407.49, 24.07 %
10,10,=$12544.90,25.45 %
5,10,=$15057.73, 50.58 %
5,5,=$19511.12, 95.11 %
```

这种参数优化手段完全靠人工，效率低而且不全面，所以我们需要专业的、全自动的测试框架。

zwQuant量化软件全自动的回溯测试框架暂定名称为：zwBacktestPro，功能如下：

- 全自动修改输入参数，并自动记录对应的回溯测试结果。
- 根据要求优化、整理、筛选相关的输出数据。
- 自动更换单一股票、多个股票分组数据源。
- 自动更换回溯测试起始、结束时间段等。

zwQuant量化软件的自动测试涉及许多专业领域，以及建模方面的知识，所以大家了解即可。

8.11　tur海龟策略v9：加入策略库

下面，我们把代码文件另存为一个新文件zw_k10\zq809_tur_v9.py，到v5、v6版本为止，我们已经完成了海龟策略tur10的全部编程工作，并且进行了参数优化。

下面，我们根据zwQuant的模式，把海龟策略tur10收录到zwQuant量化软件的策略库模块zwStrategy.py中去。

打开zwStrategy.py文件，一般新定义的策略添加在文件尾部，找到文件尾部，并加入注解行：

```
#---tur10海龟策略
```

然后，从zq809_tur_v9.py中，把海龟策略分析函数tur10、海龟数据与策略函数tur10_dataPre剪切到zwStrategy.py最后。

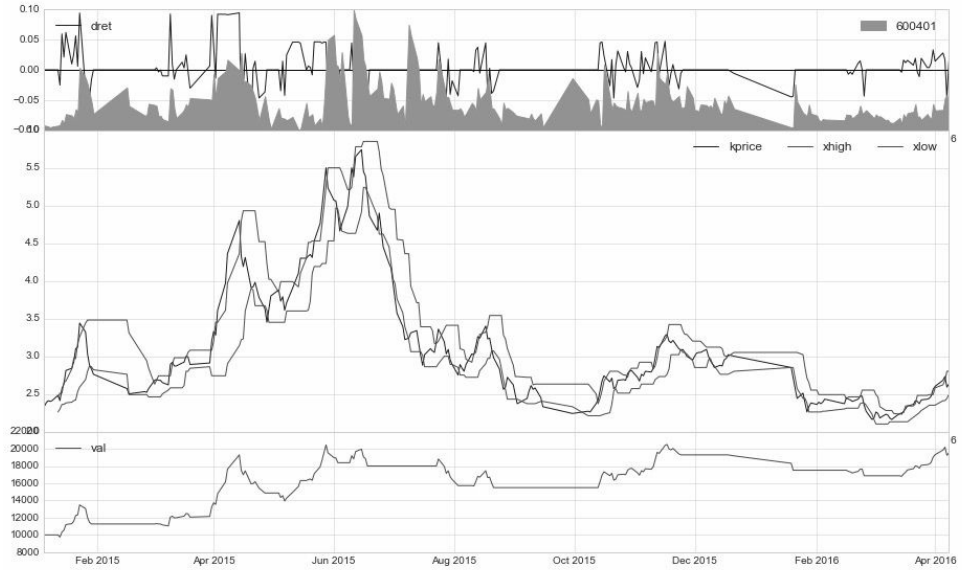
将tur10海龟策略增加到极宽Quant的策略库模块之后，其对应的主流程在策略分析、数据处理函数绑定方面也需要修改对应的代码：

```
#qx.staFun=tur10; #---绑定策略函数&运行回溯主函数
qx.staFun=zwsta.tur10; #---绑定策略函数&运行回溯主函数
#---根据当前策略，对数据进行预处理
#zwsta.tur10_dataPre(qx,'sta00','close')
```

```
zwsta.tur10_dataPre(qx,'tur10','close')
```

8.12　案例8-7：入库

案例8-9的脚本文件名是：\zwpython\zw_k10\zq809_tur_v9.py，其运行结果如图8-5所示。



每日交易推荐
::xtrdLib tmp\tur10_xtrdLib.csv

	date	ID	mode	code	dprice	num	kprice	sum	cash
14	2015-12-18	tur10_000015	buy	600401	3.02	5749.0	3.02	17361.98	1931.50
15	2016-01-20	tur10_000016	sell	600401	2.71	-5749.0	2.71	-15579.79	17511.29
16	2016-02-16	tur10_000017	buy	600401	2.45	6432.0	2.45	15758.40	1752.89
17	2016-02-25	tur10_000018	sell	600401	2.35	-6432.0	2.35	-15115.20	16868.09
18	2016-03-14	tur10_000019	buy	600401	2.24	6777.0	2.24	15180.48	1687.61

图8-5　海龟策略diy扩展v9版

案例8-9其他输出信息如下：

```
交易总次数：19
交易总盈利：9511.12

盈利交易数：12
盈利交易金额：31047.07
```

亏损交易数：7

亏损交易金额：-21535.95

最终资产价值 Final portfolio value: $19511.12

最终现金资产价值 Final cash portfolio value: $1687.61

最终证券资产价值 Final stock portfolio value: $17823.51

累计回报率 Cumulative returns: 95.11 %

平均日收益率 Average daily return: 0.297 %

日收益率方差 Std. dev. daily return:0.0274

夏普比率 Sharpe ratio: 1.608,（0.05利率）

无风险利率 Risk Free Rate: 0.05

夏普比率 Sharpe ratio: 1.723,（0利率）

最大回撤率 Max. drawdown: 27.7741 %

最长回撤时间 Longest drawdown duration: 172

回撤时间(最高点位) Time High. drawdown: 2015-11-17

回撤最高点位 High. drawdown: 20503.290

回撤最低点位 Low. drawdown: 16800.320

时间周期 Date lenght: 464 (Day)

时间周期（交易日） Date lenght(weekday): 257 (Day)

开始时间 Date begin: 2015-01-01

结束时间 Date lenght: 2016-04-08

```
项目名称 Project name: tur10
策略名称 Strategy name: tur10
股票代码列表 Stock list:  ['600401']
```

8.13　庖丁解牛

本章从零开始讲叙了如何完整地编写一个具体的量化策略函数，并添加到zwQuant量化软件的策略模块库中。

量化程序是一个专业的系统，虽然zwQuant量化软件对系统进行了高度简化，但依然是一套完整的量化系统。

只有通过庖丁解牛的方式，从整体到局部，从案例到策略，分分合合、多个层次、多个角度进行讲解，大家对于量化的本质：数据分析，才能有更深的了解，并更快进入实盘操作环节。

第9章
TA-Lib函数库与策略开发

本章，我们将通过多个案例介绍TA-Lib金融函数库，以及如何使用相关的TA-Lib金融函数，设计量化分析策略。

9.1 TA-Lib技术指标

TA-Lib金融函数库是目前金融行业最常用的技术指标函数库，目前已经有Python版本接口。zwPython集成版Python开发平台已经集成了TA-Lib。

传统金融指标除了个别技术指标和量化风险指数指标外，TA-Lib金融函数库基本都已经包括。

需注意的是，因为TA-Lib金融函数库是国外开发的，个别名称术语和参数与国内的习惯用法不同，具体使用时请大家注意相关细节。

9.1.1 TA-Lib官网

网站地址：http://ta-lib.org/，如图9-1所示。



图9-1 TA-Lib网站截图

虽然TA-Lib网站提供了源码和文档，但TA-Lib金融函数库源码是C

语言，直接下载安装Python有困难。一般都在LFD网站下载Python的二进制安装包。

LPD网址：http://www.lfd.uci.edu/~gohlke/pythonlibs/

不仅是TA-Lib金融函数库，基本上大型的非原生Python开发的Python模块库，都是在LFD网站下载二进制安装包，例如Opencv、Numpy等，LFD网站有专业团队维护，更新很快，而且提供多个版本支持。

TA-Lib金融函数库在LFD网站对应的下载说明是：

```
TA-Lib, a wrapper for the TA-LIB Technical Analysis Library.
o       TA_Lib-0.4.9-cp27-none-win32.whl
o       TA_Lib-0.4.9-cp27-none-win_amd64.whl
o       TA_Lib-0.4.9-cp34-none-win32.whl
o       TA_Lib-0.4.9-cp34-none-win_amd64.whl
o       TA_Lib-0.4.9-cp35-none-win32.whl
o       TA_Lib-0.4.9-cp35-none-win_amd64.whl
```

9.1.2 矩阵版TA-Lib金融函数模块

zw_TA-Lib是一个独立开源项目，是基于Pandas量化软件的TA-Lib函数封装，属于zwQaunt项目的衍生项目，是矩阵版的TA-Lib金融函数库。矩阵版zw_TA-Lib金融函数库无须安装zwQaunt量化软件即可独立运行。

zwQuant量化软件会逐渐采用Pandas矩阵模式对重点函数进行改写移植。zwQuant源码工具箱里面。已经收录了一个基本版本的pandas-talib源码。文件名：zwQuant\source\zw_talib.py。

目前，zw_TA-Lib作为zwQuant量化项目的衍生项目已经成为一个独立的开源项目，项目收录在zwQuant中，目录名称是：zwQuant\zwta\。

第一批以pandas_talib.py的29个函数为蓝本，默认数据格式采用zwDat标准，全部小写：

```
ohlcv:open,high,low,close,volumns
```

矩阵版zw_TA-Lib金融函数库v0.5版共有33个函数，均已测试通过运行平台：Python 3.x。

我们学习TA-Lib金融函数库的目的是为了使用TA-Lib金融函数库中的指标，用于量化策略的分析。本章，我们将通过多个案例介绍TA-Lib金融函数库，以及如何使用相关的TA-Lib金融函数，设计量化分析策略。

9.2 MACD策略

MACD称为指数平滑移动平均线，是从双指数移动平均线发展而来的，由快的指数移动平均线（EMA）减去慢的指数移动平均线，MACD的意义和双移动平均线基本相同，但阅读起来更方便。

当MACD从负数转向正数时，是买入的信号。当MACD从正数转向负数时，是卖出的信号。当MACD以大角度变化时，表示快的移动平均线和慢的移动平均线的差距是短时间内被迅速拉开的，代表了一个市场大趋势的转变。

9.2.1 MACD策略1

基于MACD的量化策略有很多，本章讲解的第一个MACD策略是最简单的：

- 当MACD>0，买进。
- 当MACD<0，卖出。

第一个MACD策略函数名为MACD10，相关的MACD策略函数和数据预处理函数的代码在文件：zwStrategy.py。

注意

为方便用户扩展策略，策略数据切割与设置函数sta_dataPre0xtim正式迁移到模块zwQTBox.py中，这样，大家在扩展策略时，无须修改函数sta_dataPre0xtim前面的模块名称。

此外，为简化策略设计，所有数据预处理函数的策略分析价格dprice和回溯成交价格kprice，都统一采用close收盘价，这个对于回溯测试结果影响不大。

MACD10与数据预处理函数相关代码如下：

```
def macd10(qx):
    '''
     MACD策略01
     MACD称为指数平滑移动平均线
    当 macd>0，买入；
    当 macd<0，卖出

    '''
    stknum=0;
    xtim,xcod=qx.xtim,qx.stkCode
    dprice=qx.xbarWrk['dprice'][0];
    xk=qx.xbarWrk['macd'][0];
    dcash=qx.qxUsr['cash'];
    dnum0=zwx.xusrStkNum(qx,xcod)
```

```
    if xk>0:
        if dnum0==0:
            stknum = int(dcash*0.9 /dprice);#dsum=stknum*kprice
            #stknum = 500
            #print(xtim,stknum,dnum,'++b,%.2f,%.2f,%.2f,$,%.2f,
            #print(xtim,stknum,'++xd',xcod,dprice,x9,x1)
    elif (xk<0):
            #stknum = -500
            stknum = -1
            #stknum = -1;dsum=dnum*kprice

    if stknum!=0:
        #print(xtim,stknum,'xd',xcod,dprice,x9,x1)
        pass;

    return stknum

def macd10_dataPre(qx,xnam0,ksgn0):
    '''
    MACD策略，数据预处理函数

    Args:
        qx (zwQuantX): zwQuantX数据包
        xnam0 (str): 函数标签
        ksgn0 (str): 价格列名称，一般是'adj close'
        '''
```

```
zwx.sta_dataPre0xtim(qx,xnam0);
#----对各只股票数据进行预处理，提高后期运算速度
ksgn,qx.priceCalc,qx.priceBuy=ksgn0,ksgn0,ksgn0  #'adj clos
for xcod in zw.stkLibCode:
    d20=zw.stkLib[xcod];

    #  计算交易价格kprice和策略分析采用的价格dprice，kprice一般采
    #d20['dprice']=d20['open']*d20[ksgn]/d20['close']
    #d20['kprice']=d20['dprice'].shift(-1)
    d20['dprice']=d20['close']
    d20['kprice']=d20['dprice']
    #
    d=qx.staVars[0];d2=qx.staVars[1];
    d20=zwta.MACD(d20,d,d2,'close');
    #d20['macd1n']=d20['macd'].shift(1)
    #d20['msign1n']=d20['msign'].shift(1)
    #
    zw.stkLib[xcod]=d20;

    if qx.debugMod>0:
        print(d20.tail())
        #---
        fss='tmp\\'+qx.prjName+'_'+xcod+'.csv'
        d20.to_csv(fss)
```

案例当中MACD策略函数MACD10流程图如图9-2所示。

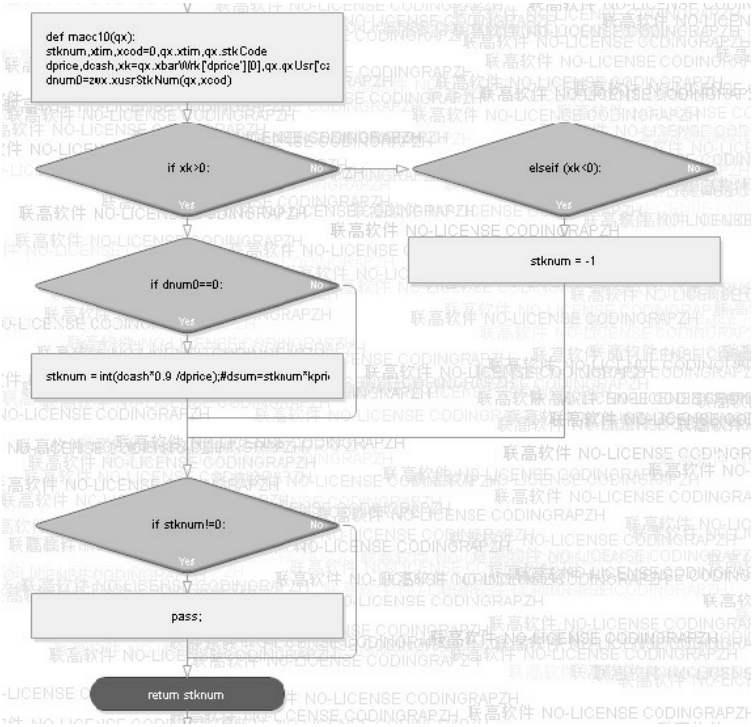


图9-2 MACD10策略函数流程图

案例当中数据预处理函数macd10_dataPre对应的输出数据是：

	open	high	low	close	volume	adj close	dpr
date							
2016-04-01	2.54	2.62	2.53	2.60	111421383.0	2.60	2
2016-04-05	2.63	2.68	2.63	2.68	110368219.0	2.68	2
2016-04-06	2.68	2.80	2.66	2.73	171821194.0	2.73	2
2016-04-07	2.74	2.78	2.59	2.59	198460790.0	2.59	2
2016-04-08	2.56	2.65	2.55	2.63	108273174.0	2.63	2

需要注意的是，在有些MACD策略中，实际调用的数据是

MACD1n、MSIGN1n数据列中的数据，即策略分析前一日的数据。zwQuant量化软件默认采用每天收市后更新当天的数据源，再进行策略分析，因此此处，直接使用的是MACD、MSIGN数据。

在数据预处理函数macd10_dataPre4中，数据列MACD1n、MSIGN1n对应的代码是屏蔽的：

```
#d20['macd1n']=d20['macd'].shift(1)
#d20['msign1n']=d20['msign'].shift(1)
```

9.2.2 案例9-1：MACD_v1

案例9-1程序文件名：\zwpython\zw_k10\zq901_macd_v1.py，运行结果如图9-3所示。

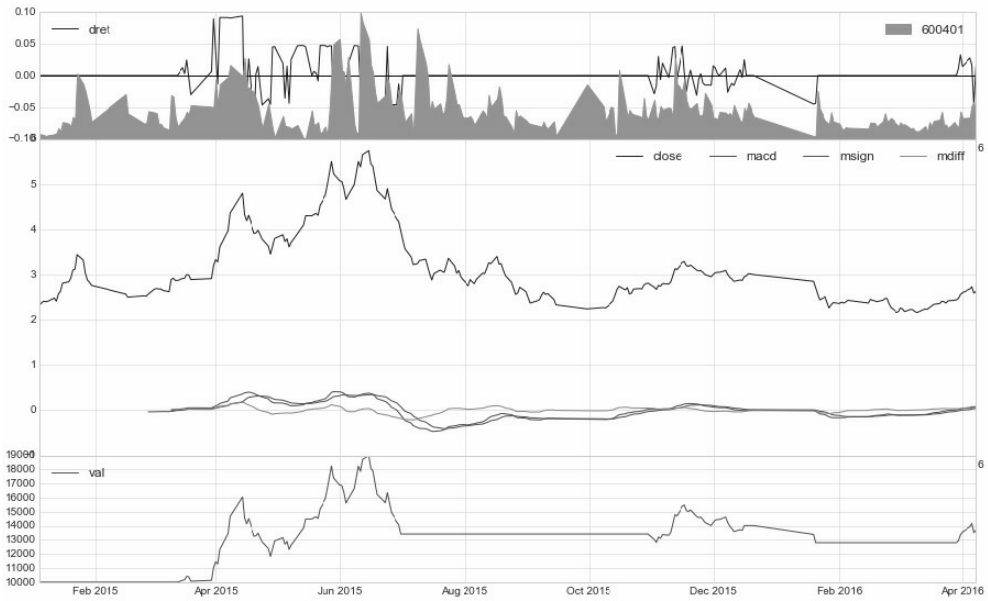


图9-3 MACD策略案例1

```
交易总次数：7
交易总盈利：4442.57

盈利交易数：4
盈利交易金额：7739.38
亏损交易数：3
亏损交易金额：-3296.81

最终资产价值 Final portfolio value: $14442.57
最终现金资产价值 Final cash portfolio value: $1345.17
最终证券资产价值 Final stock portfolio value: $13097.40
累计回报率 Cumulative returns: 44.43 %
平均日收益率 Average daily return: 0.171 %
日收益率方差 Std. dev. daily return:0.0239

夏普比率 Sharpe ratio: 1.005,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 1.136,（0利率）

最大回撤率 Max. drawdown: 28.9842 %
最长回撤时间 Longest drawdown duration: 298
回撤时间(最高点位) Time High. drawdown:  2015-06-15
回撤最高点位 High. drawdown: 18691.240
回撤最低点位 Low. drawdown: 13273.740
```

```
时间周期 Date lenght: 464 (Day)
时间周期（交易日） Date lenght(weekday): 257 (Day)
开始时间 Date begin: 2015-01-01
结束时间 Date lenght: 2016-04-08

项目名称 Project name: macd10
策略名称 Strategy name: macd10
股票代码列表 Stock list:  ['600401']
```

MACD10策略函数虽然简单，但2015年的回报率却是不错的，高达44%。

9.2.3 MACD策略2

第二个MACD策略略为复杂，策略函数：MACD20。策略函数代码文件：zwStrategy.py。

MACD策略相关的代码如下：

```
def macd20(qx):
     stknum=0;
     xtim,xcod=qx.xtim,qx.stkCode
     dprice=qx.xbarWrk['dprice'][0];
```

```
    xk=qx.xbarWrk['macd'][0];
    x2=qx.xbarWrk['msign'][0];
    dcash=qx.qxUsr['cash'];
    dnum0=zwx.xusrStkNum(qx,xcod)



    if xk>x2:
        if dnum0==0:
            stknum = int(dcash*0.9 /dprice);#dsum=stknum*kprice
            #stknum = 500
            #print(xtim,stknum,dnum,'++b,%.2f,%.2f,%.2f,$,%.2f,
            #print(xtim,stknum,'++xd',xcod,dprice,x9,x1)
    elif (xk<x2):
            #stknum = -500
            stknum = -1
            #stknum = -1;dsum=dnum*kprice


    if stknum!=0:
        #print(xtim,stknum,'xd',xcod,dprice,x9,x1)
        pass;


    return stknum
```

以上MACD策略函数的流程图如图9-4所示。

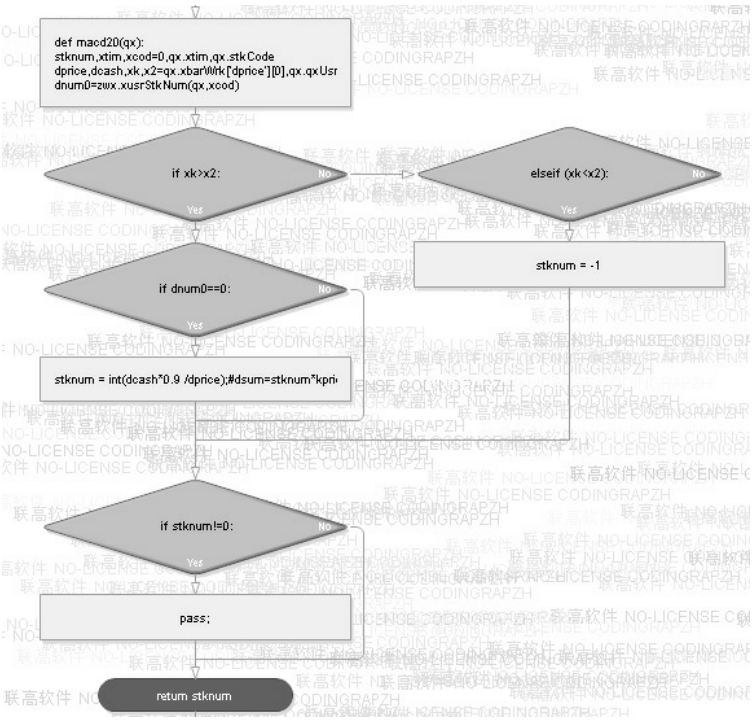


图9-4　MACD20策略流程图

由以上MACD策略函数代码和流程图可以看出，MACD20函数的策略是：

- 当macd>macd_sign，买入。
- 当macd<macd_sign0，卖出。

9.2.4　案例9-2：MACD_v2

案例9-2程序代码较长，请大家自己参考源码文件：\zwpython\zw_k10\ zq902_macd_v2.py。

案例9-2与案例9-1的代码基本相同，除了策略绑定语句以外：

```
qx.staFun=zwsta.macd20; #---绑定策略函数&运行回溯主函数
```

需要注意的是，案例9-2的数据预处理函数也是macd10_dataPre函数：

```
zwsta.macd10_dataPre(qx,'macd10','close')
```

因为MACD10和MACD20都是基于MACD指数的策略，而且数据源基本相同，为简化代码，所以采用了系统的数据预处理函数。

案例9-2运行结果如图9-5所示。

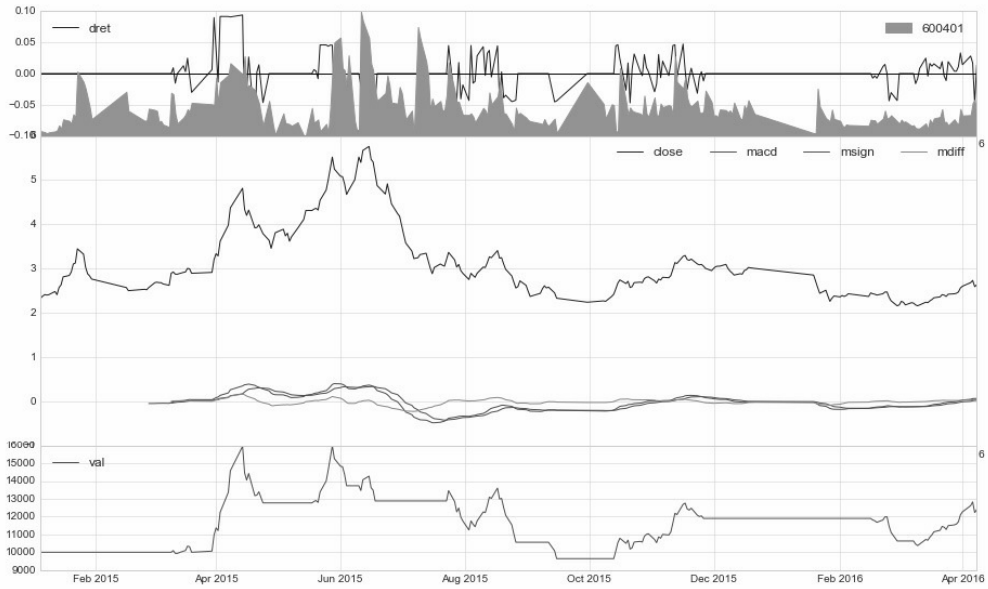


图9-5　MACD案例2

案例9-2，其他输出信息如下：

交易总次数：15
交易总盈利：2398.86

盈利交易数：8
盈利交易金额：19598.95
亏损交易数：7
亏损交易金额：-17200.09

最终资产价值 Final portfolio value: $12398.86
最终现金资产价值 Final cash portfolio value: $1063.56
最终证券资产价值 Final stock portfolio value: $11335.30
累计回报率 Cumulative returns: 23.99 %
平均日收益率 Average daily return: 0.114 %
日收益率方差 Std. dev. daily return:0.0247

夏普比率 Sharpe ratio: 0.604,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 0.732,（0利率）

最大回撤率 Max. drawdown: 39.6660 %
最长回撤时间 Longest drawdown duration: 316
回撤时间(最高点位) Time High. drawdown: 2015-05-28
回撤最高点位 High. drawdown: 15979.060
回撤最低点位 Low. drawdown: 9640.800

时间周期 Date lenght: 464 (Day)

```
时间周期（交易日） Date lenght(weekday): 257 (Day)
开始时间 Date begin: 2015-01-01
结束时间 Date lenght: 2016-04-08

项目名称 Project name: macd20
策略名称 Strategy name: macd20
股票代码列表 Stock list:  ['600401']
```

9.3 KDJ策略

KDJ指标又称随机指标，其综合了动量观念、强弱指标及移动平均线的优点，用来度量股价脱离价格正常范围的变异程度。KDJ指标考虑的不仅仅是收盘价，而且有近期的最高价和最低价，这避免了仅考虑收盘价而忽视真正波动幅度的弱点。

在TA-Lib金融函数库中，KDJ指标对应的函数是STOD函数，返回参数中的Stod、Stok分别对应KDJ指标中的K值、D值。因为J值很少用到，而且计算简单，没有直接返回，有需要的用户可以根据上面的公式另行计算。

9.3.1 KDJ策略1

基于KDJ随机指标的量化策略有很多，这第一个KDJ策略相对比较简单：

- 当K值stok>90，买进。
- 当K值stoks10，卖出。

KDJ10策略函数相关代码如下：

```
def kdj10(qx):
    '''
```

```
 KDJ策略10
 KDJ 指标，又称随机指标
当 stok>90，买入；
当 stok<10，卖出

'''
stknum=0;
xtim,xcod=qx.xtim,qx.stkCode
dprice=qx.xbarWrk['dprice'][0];
dcash=qx.qxUsr['cash'];
dnum0=zwx.xusrStkNum(qx,xcod)
#
ksgn1,ksgn2='stok','stod'
xk,xk2=qx.xbarWrk[ksgn1][0],qx.xbarWrk[ksgn2][0];



if xk>90:
    if dnum0==0:
        stknum = int(dcash*0.9 /dprice);#dsum=stknum*kprice
        #stknum = 500
        #print(xtim,stknum,dnum,'++b,%.2f,%.2f,%.2f,$,%.2f,
        #print(xtim,stknum,'++xd',xcod,dprice,x9,x1)
elif (xk<10):
        #stknum = -500
        stknum = -1
        #stknum = -1;dsum=dnum*kprice
```

```
    if stknum!=0:
        #print(xtim,stknum,'xd',xcod,dprice,x9,x1)
        pass;

    return stknum
```

为了便于理解，大家可以参考KDJ10策略函数的流程图，如图9-6所示。

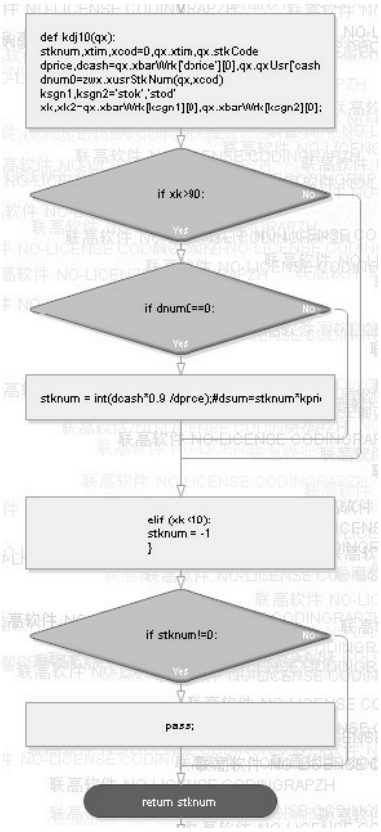


图9-6 KDJ10策略函数流程图

策略函数KDJ10对应的数据预处理函数是kdj10_dataPre，有关代码如下：

```
def kdj10_dataPre(qx,xnam0,ksgn0):
    zwx.sta_dataPre0xtim(qx,xnam0);
    #----对各只股票数据进行预处理，提高后期运算速度
    ksgn,qx.priceCalc,qx.priceBuy=ksgn0,ksgn0,ksgn0  #'adj clos
    for xcod in zw.stkLibCode:
        d20=zw.stkLib[xcod];

        #  计算交易价格kprice和策略分析采用的价格dprice，kprice一般采用
        #d20['dprice']=d20['open']*d20[ksgn]/d20['close']
        #d20['kprice']=d20['dprice'].shift(-1)
        d20['dprice']=d20['close']
        d20['kprice']=d20['dprice']
        #
        d=qx.staVars[0];#d2=qx.staVars[1];
        d20=zwta.STOD(d20,d,'close');
        d20['stod1n']=d20['stod'].shift(1)
        d20['stok1n']=d20['stok'].shift(1)
        #
        zw.stkLib[xcod]=d20;
        if qx.debugMod>0:
            print(d20.tail())
            #---
            fss='tmp\\'+qx.prjName+'_'+xcod+'.csv'
            d20.to_csv(fss)
```

为了便于理解，大家可以参考KDJ数据预处理策略函数的流程图，

如图9-7所示。

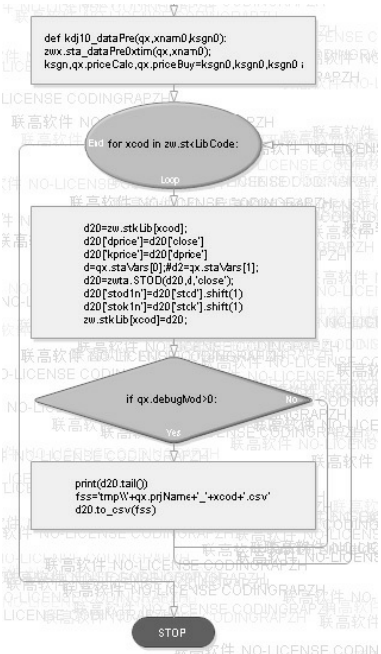


图9-7　KDJ10数据预处理函数流程图

9.3.2　案例9-3：KDJ01

案例9-3主要是介绍KDJ策略，案例程序文件名：\zwpython\zw_k10\zq903_kdj01.py。

案例9-3主要运行结果如图9-8所示。

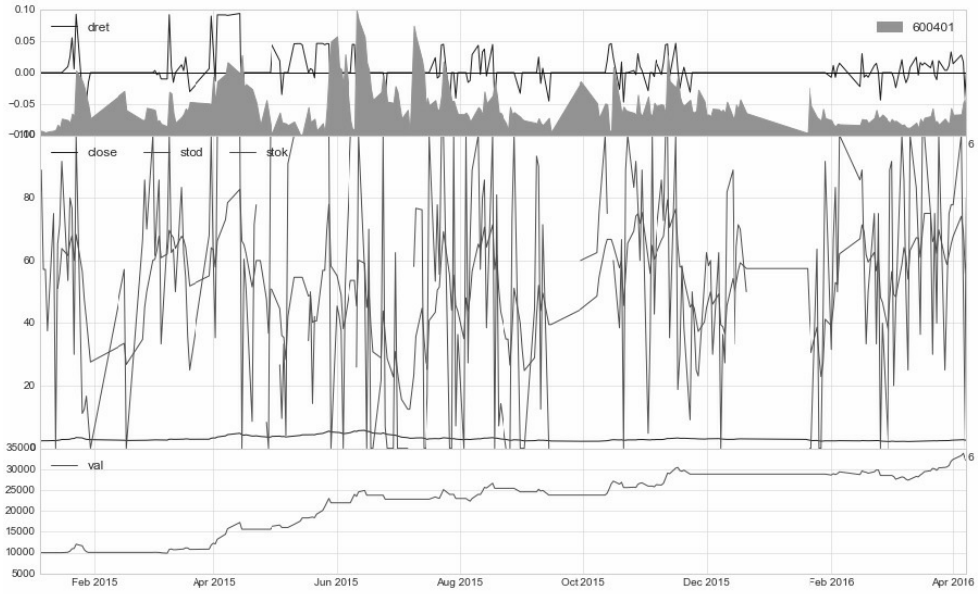


图9-8 KDJ案例1

案例9-3其他输出信息如下：

```
交易总次数：32
交易总盈利：22234.43

盈利交易数：15
盈利交易金额：65970.25
亏损交易数：17
亏损交易金额：-43735.82

最终资产价值 Final portfolio value: $32234.43
最终现金资产价值 Final cash portfolio value: $32234.43
最终证券资产价值 Final stock portfolio value: $0.00
累计回报率 Cumulative returns: 222.34 %
平均日收益率 Average daily return: 0.488 %
```

```
日收益率方差 Std. dev. daily return:0.0255

夏普比率 Sharpe ratio: 2.913,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 3.036,（0利率）

最大回撤率 Max. drawdown: 17.9405 %
最长回撤时间 Longest drawdown duration: 132
回撤时间(最高点位) Time High. drawdown:  2016-04-06
回撤最高点位 High. drawdown: 33822.730
回撤最低点位 Low. drawdown: 32234.430

时间周期 Date lenght: 464 (Day)
时间周期（交易日） Date lenght(weekday): 257 (Day)
开始时间 Date begin: 2015-01-01
结束时间 Date lenght: 2016-04-08

项目名称 Project name: kdj10
策略名称 Strategy name: kdj10
股票代码列表 Stock list:  ['600401']
```

输出的图形有些乱，因为KDJ的范围是基于0~100，需要对绘图模版进行修改，或者在数据预处理和策略分析时对KDJ参数除以10，以降低幅度。

总体来看，KDJ策略的回报率还是不错的，2015年回报率是222%。

9.3.3 KDJ策略2

第二个KDJ策略相对复杂一些：

- 当K值stok>D值stod，并且是向上趋势，买进。
- 当K值stok<D值stod，并且是向下趋势，卖出。

数据预处理函数还是一样的kdj10_dataPre。对应的策略函数KDJ20，相关代码如下：

```
def kdj20(qx):
     stknum=0;
     xtim,xcod=qx.xtim,qx.stkCode
     dprice=qx.xbarWrk['dprice'][0];
     dcash=qx.qxUsr['cash'];
     dnum0=zwx.xusrStkNum(qx,xcod)
     #
     ksgn1,ksgn2='stok','stod'
     xk,xk2=qx.xbarWrk[ksgn1][0],qx.xbarWrk[ksgn2][0];
     nksgn1,nksgn2='stok1n','stod1n'
     nxk,nxk2=qx.xbarWrk[nksgn1][0],qx.xbarWrk[nksgn2][0];

     if (xk>xk2)and(nxk<=nxk2):
         if dnum0==0:
             stknum = int(dcash*0.9 /dprice);#dsum=stknum*kprice
```

```
            #stknum = 500
            #print(xtim,stknum,dnum,'++b,%.2f,%.2f,%.2f,$,%.2f,
            #print(xtim,stknum,'++xd',xcod,dprice,x9,x1)
    elif (xk<xk2)and(nxk>=nxk2):
            #stknum = -500
            stknum = -1
            #stknum = -1;dsum=dnum*kprice

    if stknum!=0:
        #print(xtim,stknum,'xd',xcod,dprice,x9,x1)
        pass;

    return stknum
```

为了便于理解，大家可以参考KDJ策略函数2的流程图，如图9-9所示。

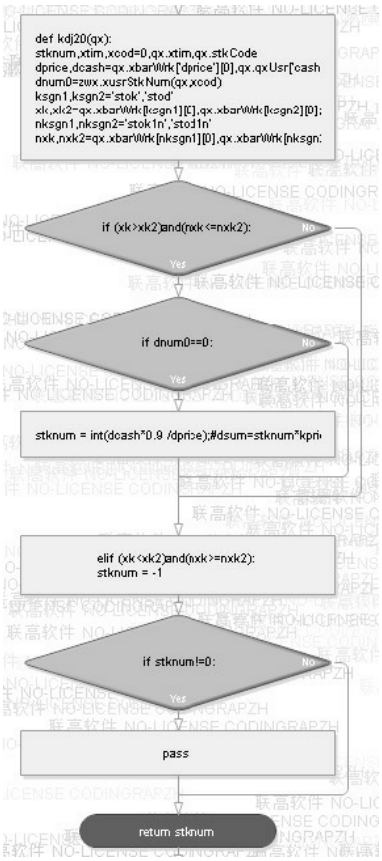


图9-9　KDJ20策略函数2流程图

9.3.4　案例9-4：KDJ02

案例9-4是介绍KDJ第2个策略，有关程序源码请参见程序文件：\zwpython\zw_k10\zq904_kdj02.py，运行结果如图9-10所示。

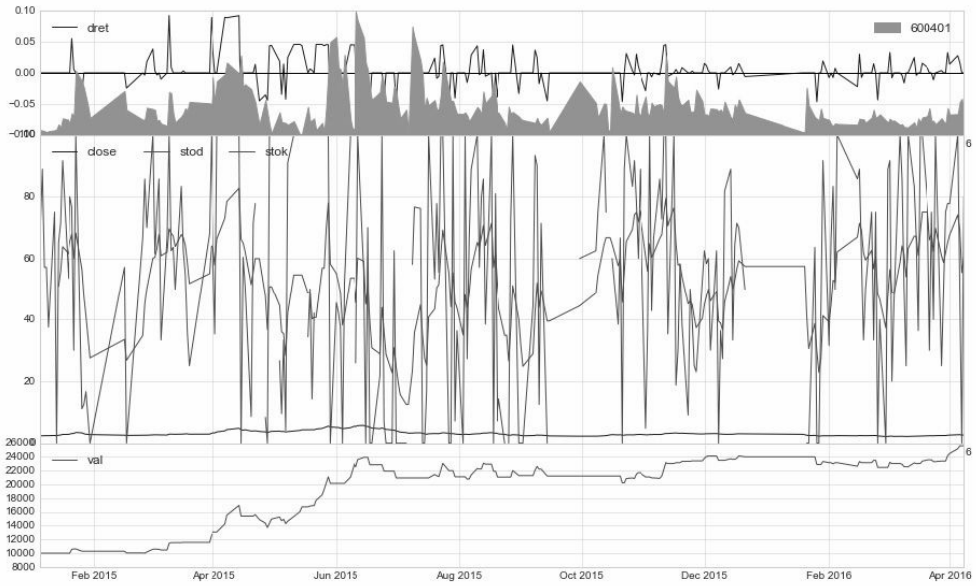


图9-10　KDJ策略案例2

案例9-4其他运行结果如下：

```
交易总次数：99
交易总盈利：15610.60

盈利交易数：52
盈利交易金额：147260.70
亏损交易数：47
亏损交易金额：-131650.10

最终资产价值 Final portfolio value: $25610.60
最终现金资产价值 Final cash portfolio value: $2561.28
最终证券资产价值 Final stock portfolio value: $23049.32
累计回报率 Cumulative returns: 156.11 %
平均日收益率 Average daily return: 0.395 %
```

```
日收益率方差 Std. dev. daily return:0.0239

夏普比率 Sharpe ratio: 2.488,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 2.619,（0利率）

最大回撤率 Max. drawdown: 19.0211 %
最长回撤时间 Longest drawdown duration: 169
回撤时间(最高点位) Time High. drawdown:  2016-04-06
回撤最高点位 High. drawdown: 25610.600
回撤最低点位 Low. drawdown: 25610.600

时间周期 Date lenght: 464 (Day)
时间周期（交易日） Date lenght(weekday): 257 (Day)
开始时间 Date begin: 2015-01-01
结束时间 Date lenght: 2016-04-08

项目名称 Project name: kdj20
策略名称 Strategy name: kdj20
股票代码列表 Stock list:  ['600401']
```

总体来看，KDJ20策略的回报率还是不错的，2015年的回报率高达156%。

9.4 RSI策略

RSI是相对强弱指标（Relative Strength Index，简称RSI），也称为相对强弱指数或相对力度指数。RSI相对强弱指标通过比较一段时期内的平均收盘涨数和平均收盘跌数来分析市场买沽盘的意向和实力，从而做出未来市场的走势。

RSI相对强弱指标最早应用于期货买卖，后来人们发现在众多的图表技术分析中，强弱指标的理论和实践极其适合股票市场的短线投资，于是被用于股票升跌的测量和分析中。

相比其他分析工具，RSI指标实用性很强，故一经推出便大受欢迎。

```
RSI计算公式和方法：
RSI＝[上升平均数÷（上升平均数＋下跌平均数）]×100
```

RSI相对强弱指标是根据一定时期内上涨和下跌幅度之和的比率制作出的一种技术曲线，能够反映出市场在一定时期内的景气程度。

由于RSI指标实用性很强，因而被多数投资者所喜爱。该技术指标非常适合做短线差价操作，该分析指标的设计是以三条线来反映价格走势的强弱，这种图形可以为投资者提供操作依据。

9.4.1 RSI取值的大小

```
RSI的变动范围在0~100之间，强弱指标值一般分布在20~80。
80—100极强卖出
50—80强买入
20—50弱观望
0—20极弱买入
```

这里的“极强”、“强”、“弱”和“极弱”只是一个相对的分析概念，即一个相对的区域，也可把它们取值为30、70或15、85，根据个人的喜好或习惯来分，没有绝对性。

9.4.2 RSI策略

基于RSI的量化策略有很多，下面所讲的RSI策略相对比较简单：

- 当rsi>kbuy（一般是70、80），买进。
- 当rsi<ksell（一般是30、20），卖出。

RSI策略函数名为RSI10，代码在文件：zwStrategy.py，RSI策略函数代码如下：

```
def rsi10(qx):
```

```
'''
RSI策略
RSI相对强弱指标
当 rsi>kbuy，一般是70、80买入
当 rsi<sell，一般是30、20卖出
'''
stknum=0;
xtim,xcod=qx.xtim,qx.stkCode
dprice=qx.xbarWrk['dprice'][0];
dcash=qx.qxUsr['cash'];
dnum0=zwx.xusrStkNum(qx,xcod)
#
d=qx.staVars[0];kstr1='rsi_{n}'.format(n=d)
xk=qx.xbarWrk[kstr1][0]
kbuy,ksell=qx.staVars[1],qx.staVars[2]

if xk>kbuy:
    if dnum0==0:
        stknum = int(dcash*0.9 /dprice);#dsum=stknum*kprice
        #stknum = 500
        #print(xtim,stknum,dnum,'++b,%.2f,%.2f,%.2f,$,%.2f,
        #print(xtim,stknum,'++xd',xcod,dprice,x9,x1)
elif xk<ksell:
        #stknum = -500
        stknum = -1
        #stknum = -1;dsum=dnum*kprice
```

```
    if stknum!=0:
        #print(xtim,stknum,'xd',xcod,dprice,x9,x1)
        pass;

    return stknum
```

为了便于理解，大家可以参考RSI策略函数的流程图，如图9-11所示。

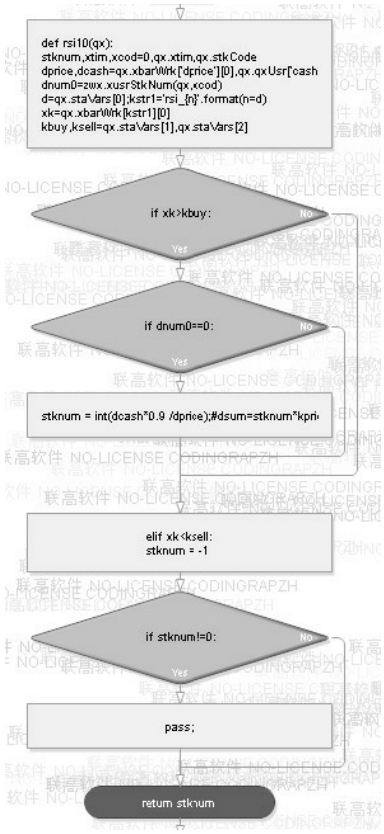


图9-11　RSI策略函数流程图

RSI策略函数对应的数据预处理函数是rsi10_dataPre，代码如下：

```
def rsi10_dataPre(qx,xnam0,ksgn0):
    '''
    RSI策略, 数据预处理函数

    Args:
        qx (zwQuantX): zwQuantX数据包
        xnam0 (str): 函数标签
        ksgn0 (str): 价格列名称，一般是'adj close'
        '''

    zwx.sta_dataPre0xtim(qx,xnam0);
    #----对各只股票数据进行预处理，提高后期运算速度
    ksgn,qx.priceCalc,qx.priceBuy=ksgn0,ksgn0,ksgn0  #'adj clos
    for xcod in zw.stkLibCode:
        d20=zw.stkLib[xcod];

        #  计算交易价格kprice和策略分析采用的价格dprice,kprice一般采用
        #d20['dprice']=d20['open']*d20[ksgn]/d20['close']
        #d20['kprice']=d20['dprice'].shift(-1)
        d20['dprice']=d20['close']
        d20['kprice']=d20['dprice']
        #
        d=qx.staVars[0];#d2=qx.staVars[1];
        d20=zwta.RSI(d20,d);
        #d20['macd1n']=d20['macd'].shift(1)
        #d20['msign1n']=d20['msign'].shift(1)
```

```
        #
        zw.stkLib[xcod]=d20;
        if qx.debugMod>0:
            print(d20.tail())
            #---
            fss='tmp\\'+qx.prjName+'_'+xcod+'.csv'
            d20.to_csv(fss)
```

为了便于理解，大家可以参考RSI策略预处理函数的流程图，如图9-12所示。

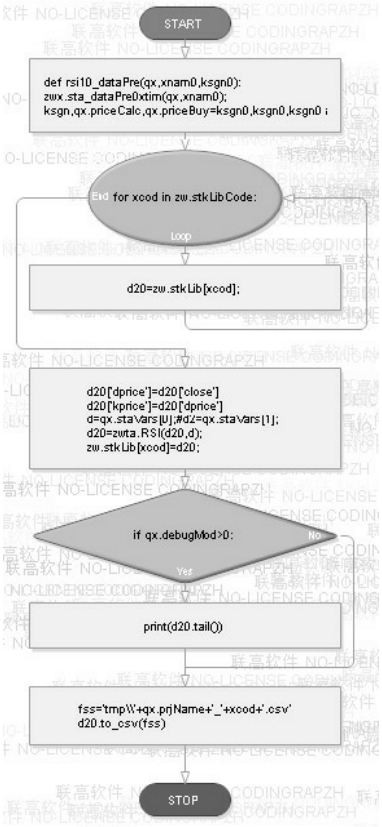


图9-12　RSI策略预处理函数流程图

以下是数据预处理函数，处理后的数据源：

```
            open  high   low  close      volume  adj close   dpr
date
2016-04-01  2.54  2.62  2.53   2.60  111421383.0       2.60    2
2016-04-05  2.63  2.68  2.63   2.68  110368219.0       2.68    2
2016-04-06  2.68  2.80  2.66   2.73  171821194.0       2.73    2
2016-04-07  2.74  2.78  2.59   2.59  198460790.0       2.59    2
2016-04-08  2.56  2.65  2.55   2.63  108273174.0       2.63    2
```

相比其他几个案例，RSI的数据列rsi_14，有一个数字14的后缀，这是因为我们输入的策略参数是：

```
qx.staVars=[14,70,30,'2015-01-01','']
```

其中第一个参数14是RSI的周期参数。

因此，在策略分析函数KDJ10中，数据列名称使用了以下代码：

```
d=qx.staVars[0];kstr1='rsi_{n}'.format(n=d)
xk=qx.xbarWrk[kstr1][0]
```

同样，为了绘制数据列rsi_14图形，我们在数据输出函数bt_endRets中也使用了以下代码：

```
d=qx.staVars[0];mstr1='rsi_{n}'.format(n=d)
kmid8=[[xcod,ksgn,mstr1]]
```

以上代码都是根据策略输入参数，生成对应的数据列名称。

9.4.3 预留参数优化接口

在策略分析函数RSI10中使用了一些技巧：

```
d=qx.staVars[0];kstr1='rsi_{n}'.format(n=d)
xk=qx.xbarWrk[kstr1][0]
kbuy,ksell=qx.staVars[1],qx.staVars[2]

if xk>kbuy:
     if dnum0==0:
          stknum = int(dcash*0.9 /dprice);#dsum=stknum*kprice
          #stknum = 500
          #print(xtim,stknum,dnum,'++b,%.2f,%.2f,%.2f,$,%.2f,%.2f
          #print(xtim,stknum,'++xd',xcod,dprice,x9,x1)
elif xk<ksell:
          #stknum = -500
          stknum = -1
          #stknum = -1;dsum=dnum*kprice
```

xk是数据列RSI参数比较的对象，不再是预设常数，而是kbuy、ksell两个变量。

```
kbuy,ksell=qx.staVars[1],qx.staVars[2]
```

由代码可见，kbuy、ksell两个变量源自staVars策略输入参数。

这样设计，虽然增加了一点代码，但为参数优化预留了测试接口。以后升级版本的zwBackTestPro量化程序中的全自动回溯测试模块，可以根据这些接口动态改变输入数值，进行最优化模型的测试。

9.4.4 案例9-5：A股版RSI策略

案例9-5是基于A股实盘数据的RSI策略。

程序代码文件是：\zwpython\zw_k10\ zq905_rsi_01.py，运行结果如图9-13所示。

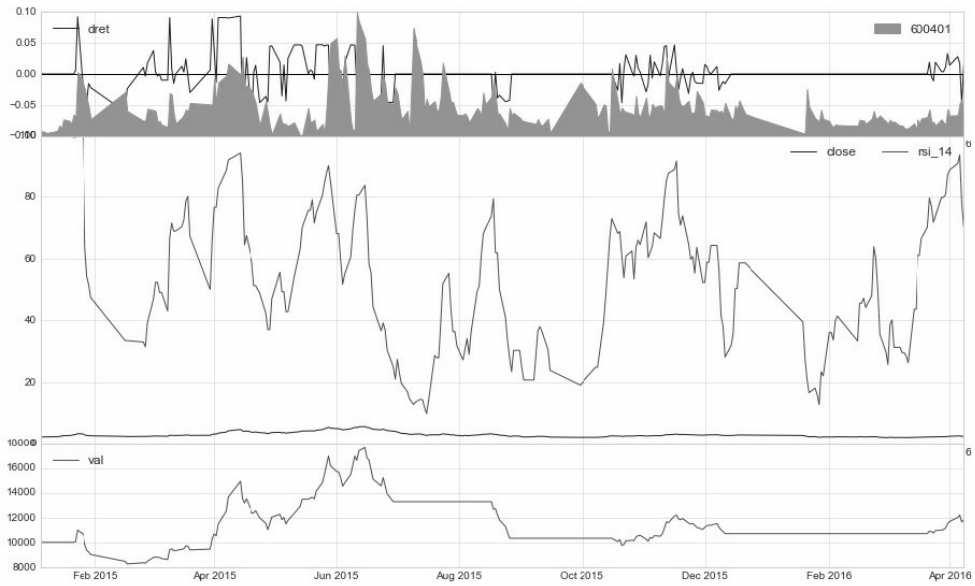


图9-13　RSI案例

案例9-5的其他输出信息如下：

```
交易总次数：7
交易总盈利：1802.51

盈利交易数：3
盈利交易金额：6492.64
亏损交易数：4
亏损交易金额：-4690.13

最终资产价值 Final portfolio value: $11802.51
最终现金资产价值 Final cash portfolio value: $1072.11
最终证券资产价值 Final stock portfolio value: $10730.40
累计回报率 Cumulative returns: 18.03 %
平均日收益率 Average daily return: 0.102 %
```

```
日收益率方差 Std. dev. daily return:0.0275

夏普比率 Sharpe ratio: 0.474,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: 0.588,（0利率）

最大回撤率 Max. drawdown: 44.7962 %
最长回撤时间 Longest drawdown duration: 298
回撤时间(最高点位) Time High. drawdown:  2015-06-15
回撤最高点位 High. drawdown: 17663.920
回撤最低点位 Low. drawdown: 9751.150

时间周期 Date lenght: 464 (Day)
时间周期（交易日） Date lenght(weekday): 257 (Day)
开始时间 Date begin: 2015-01-01
结束时间 Date lenght: 2016-04-08

项目名称 Project name: rsi10
策略名称 Strategy name: rsi10
股票代码列表 Stock list:  ['600401']
策略参数变量 staVars[]:  [14, 70, 30, '2015-01-01', '']
```

相比前面的价格案例，案例9-5的回报率有些低，才18%，不过，考虑到2015年的大盘行情，这已经是不错的策略。

9.5 基石、策略与灵感

李嘉诚曾经说过：“做生意，最重要的是地段、地段、还是地段。”

同理，做量化交易，最重要的就是：策略、策略、还是策略！

那么，这些策略从何而来？

量化策略需要收集、需要积累，同时更要善于借鉴其他人的成果。

TA-Lib金融函数库是久经考验的金融函数库，在金融量化领域基石般的存在。虽然，zw_TA-Lib目前只移植了30多个函数，但这30多个函数都是经典中的经典。举一反三，足以衍生出上百种不同的量化策略。如果这些经典指标形成组合策略，那么能够提供的策略数目成千上万。

这些全都是策略的灵感之源。所以，做量化其实并不缺策略，缺的只是稳定、经过测试的策略。

第10章
扩展与未来

在前面的章节我们已经介绍了zwQuant量化软件的整体架构、模块、函数、策略函数、数据预处理函数等内容。

第10章是本书正文的最后一章，我们将学习zwQuant量化软件自身的扩展。

zwQuant量化软件是“骨头版本”的开源量化系统，逻辑简单、清晰，全部代码（不含zw_talib金融函数库）去掉注解不到1000行，是学习Python量化程序理想的入门教程与系统。

同时，zwQuant量化软件只提供了最基本的Python量化功能，留下了大量的空白，供广大Python程序员和量化高手进行扩展。

本章，我们将通过多个案例详细讲解zwQuant量化软件的扩展方法。

10.1　回顾案例2-1：SMA均线策略

为了尽快深入到量化一线，我们在第2章就介绍过一个相对完整的量化分析案例2-1：SMA均线策略。

第2章案例2-1采用的是最简单的“SMA均线策略”案例，这个案例源自pyAlgoTrade量化软件文档，为便于理解，笔者对相关代码进行了修改和汉化，并增加了绘图输出。

案例2-1程序文件名：\zwpython\zw_k10\k201_sta_anz.py。

注意，本案例是与pyAlgoTrade相关的程序，只适用于Python 2.7。

有关代码前面已经写过，这里就此省略，请大家参看第2章案例2-1和程序文件。

案例2-1运行后，输出的数据包括图像和文字信息，输出图像如图10-1所示。

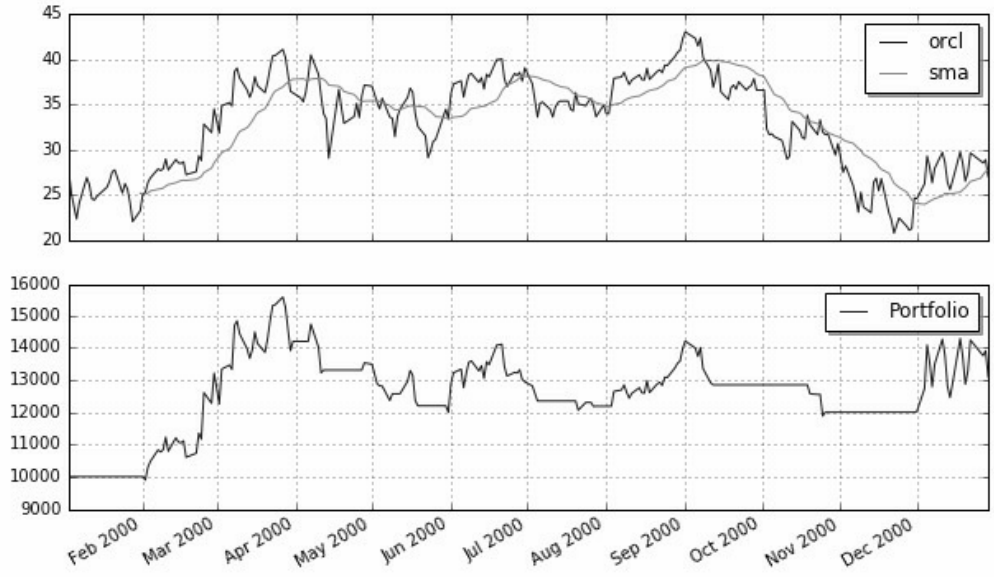


图10-1　SMA均线策略

案例2-1输出的其他信息如下：

```
最终资产价值 Final portfolio value: $13091.62
累计回报率 Cumulative returns: 30.92 %
夏普比率 Sharpe ratio: 0.72
最大回撤率 Max. drawdown: 23.69 %
最长回撤时间 Longest drawdown duration: 277 days, 0:00:00

总交易 Total trades: 13
平均利润 Avg. profit: $154
利润方差 Profits std. dev.: $1268
最大利润 Max. profit: $4197
最小利润 Min. profit: $-892
平均收益率 Avg. return:  2 %
收益率方差 Returns std. dev.: 13 %
```

```
最大收益率 Max. return: 46 %
最小收益率 Min. return: -7 %

赢利交易 Profitable trades: 3
平均利润 Avg. profit: $1964
利润方差 Profits std. dev.: $1586
最大利润 Max. profit: $4197
最小利润 Min. profit: $662
平均收益率 Avg. return: 21 %
收益率方差 Returns std. dev.: 18 %
最大收益率 Max. return: 46 %
最小收益率 Min. return:  6 %

亏损交易Unprofitable trades: 10
平均亏损 Avg. loss: $-389
亏损方差 Losses std. dev.: $238
最大亏损 Max. loss: $-892
最小亏损 Min. loss: $-44
平均收益率 Avg. return: -3 %
收益率方差 Returns std. dev.:  2 %
最大收益率 Max. return: -0 %
最小收益率 Min. return: -7 %
```

在案例2-1程序中，策略调用以下语句，20日SMA平均线策略：

```
myStrategy = SMACrossOver(feed, "orcl", 20)
```

这是最简单的SMA策略，程序默认的参数是20日平均线策略，这个20日是优化后的结果，大家可以用5、15、29、21、25、30等参数测试一下，大部分是负收益和低收益。

pyAlgoTrade量化程序采用的是目前常用的事件模式，案例2-1采用的是最简单的SMA均线策略，核心代码是onBars事件函数，其他的都是辅助功能：数据源配置、绘图、计算回报率等。

请注意函数定义的开头字母：on，表示是事件函数。

```
def onBars(self, bars):
     # If a position was not opened, check if we should enter a
     if self.__position is None:
         if cross.cross_above(self.__prices, self.__sma) > 0:
             shares = int(self.getBroker().getCash() * 0.9 / bar
             # Enter a buy market order. The order is good till
             self.__position = self.enterLong(self.__instrument,
     # Check if we have to exit the position.
     elif not self.__position.exitActive() and cross.cross_below
         self.__position.exitMarket()
```

SMA均线策略很简单：

- 当股票价格高于SMA平均线价格，并且是向上趋势时，买入。
- 当股票价格低于SMA平均线价格，并且是向下趋势时，卖出。

案例2-1从某种程度上而言，是大家学习量化交易的起点，也是学习量化交易的终点。

对于初学者而言，学习量化交易的整个过程就是让大家掌握、理解案例2-1背后的量化流程，以及如何灵活利用其他策略取代案例中的SMA平均线策略。

对于有更加专业要求的读者可以学习掌握相关的代码构成，以及背后的基本逻辑，独立完成相关的Backtest回溯测试模块编程。

策略是量化交易的核心，无论是初学者还是专业人员，只需要重点学习Backtest回溯测试部分就可以了。

前端的数据采集和后端Trade下单可以通过CSV等数据文件格式和第三方软件进行，部分环节，特别是非高频的下单环节也可以采用人工模式。

这个也是kiss法则在量化方面的体现。

案例10-1：SMA均线策略扩展

案例10-1是案例2-1的修改版本，“x”是罗马字母，是数字“10”的意思，所以第10章的案例程序文件以“zqx”作为文件名开头字母。

案例10-1脚本文件名是：\zwpython\zw_k10\zqx01_sma.py，具体关

代码请大家参看程序文件。

案例10-1运行结果如图10-2所示。

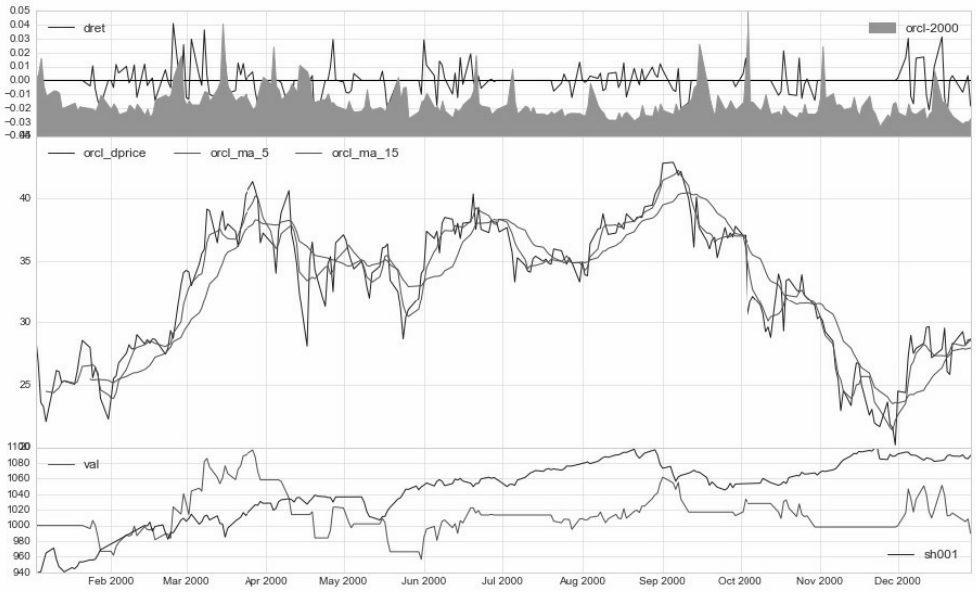


图10-2　SMA均线策略扩展

案例10-1其他输出信息如下：

```
交易总次数：43
交易总盈利：-10.16

盈利交易数：22
盈利交易金额：1564.25
亏损交易数：21
亏损交易金额：-1574.41

最终资产价值 Final portfolio value: $989.84
```

最终现金资产价值 Final cash portfolio value: $719.94

最终证券资产价值 Final stock portfolio value: $269.90

累计回报率 Cumulative returns: -1.02 %

平均日收益率 Average daily return: 0.001 %

日收益率方差 Std. dev. daily return:0.0098

夏普比率 Sharpe ratio: -0.310,（0.05利率）

无风险利率 Risk Free Rate: 0.05

夏普比率 Sharpe ratio: 0.012,（0利率）

最大回撤率 Max. drawdown: 12.7556 %

最长回撤时间 Longest drawdown duration: 277

回撤时间(最高点位) Time High. drawdown: 2000-03-27

回撤最高点位 High. drawdown: 1096.641

回撤最低点位 Low. drawdown: 956.758

时间周期 Date lenght: 362 (Day)

时间周期（交易日） Date lenght(weekday): 252 (Day)

开始时间 Date begin: 2000-01-03

结束时间 Date lenght: 2000-12-29

项目名称 Project name: sma

策略名称 Strategy name: sma

股票代码列表 Stock list: ['orcl-2000']

策略参数变量 staVars[]: [5, 15, '2000-01-01', '2000-12-31']

与前面几章的案例对比，该案例输出的文字信息略有增加，但变化不大。

最大的变化是图形输出，如图10-3所示。

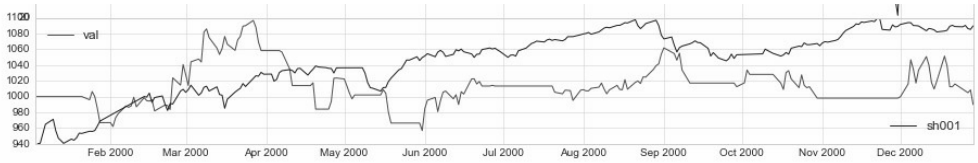


图10-3　大盘指数图形

以前的版本只有红色线条：val资产总值，新的版本增加了蓝色线条：大盘指数。在案例10-1中蓝色图标是sh001，表示上证指数。这个名称用户可以自己修改。

这个新增的大盘指数及其绘图输出是我们新的扩展，也是本章的主要内容。

10.2　大盘指数资源

大盘指数数据其实就是各种大盘的日线数据，对于做高频的用户而言指的是对应的分时数据。大盘指数数据是zwDat金融数据包项目的一部分。

zwDat金融数据包是极宽开源量化项目的重要组成部分，随着本书的发布，我们对zwDAT金融数据包进行了部分优化升级，并改名为zwDat2.0，升级的部分包括：

- 修改了多个模块和down_stk数据下载代码。
- 中国A股数据源更新，升级到最新时间。
- 国内大盘指数文件更新，主要是inx目录下的文件重新进行了整理。
- 国内大盘指数全部统一采用六位的股票代码，相关数据增加到1994年开市起。
- 由于zwQuant量化软件的dataPre数据预处理模块可以快速对数据进行归一化处理，zwDat金融数据包取消了归一化数据格式，数据包所占空间减少50%。

注意

美股日线数据下载函数没有去重、追加部分，每次需要从头下载所有数据，所以，本次升级没有更新美股数据，有需要的用户可以自己下载美股数据。做美股实盘的用户请参考A股下载代码修改相关函

数，增加数据去重、追加功能。

10.2.1 大盘指数文件升级

zwDat2.0金融数据包变化最大的是大盘指数文件，原来的大盘指数文件采用字母、数字混用，只有19种，而且只有近三年的数据。此次升级，对大盘指数数据下载，进行了多处升级：

- 采用了全新的独立数据抓取函数。
- 大盘指数种类增加到24类。
- 指数名称统一采用6位数字模式。
- 大盘数据最早增加到1994年股市开市起。
- 索引文件增加了起始时间栏目：tim0。
- 所有指数、起始时间全部采用人工手动验证。

指数文件，保存在目录下：

```
x:\zwDat\cn\xday\
```

指数索引文件名是：E:\zwDat\inx\ inx_code.csv。

相关内容如表10-1所示。

表10-1 大盘指数

code	name	tim0
000001	上证指数	1994-01-01
000002	A股指数	1994-01-01
000003	B股指数	1994-01-01
000008	综合指数	1994-01-01
399001	深证成指	1994-01-01
399002	深成指R	1994-01-01
399003	成份B指	1994-01-01
000009	上证380	2011-01-01
000010	上证180	2000-01-01
000011	基金指数	2001-01-01
000012	国债指数	2003-01-01
000016	上证50	2004-01-01
000017	新综指	2006-01-01
000300	沪深300	2005-01-01
399004	深证100R	2003-01-01
399005	中小板指	2006-01-01
399006	创业板指	2010-01-01
399100	深证新指数	2006-01-01
399101	中小板综	2005-01-01
399106	深证综指	2000-01-01
399107	深证A指	2000-01-01
399108	深证B指	2000-01-01
399333	中小板R	2006-01-01
399606	创业板R	2010-01-01

10.2.2 大盘指数内存数据库

zwQuant量化软件的股票数据全部保存在全局变量stkLib中。stkLib的数据全部位于内存，可以看成一个Pandas数据分析软件版本的内存数据库。

zwQuant量化软件的stkLib与传统的股票池、线程池技术略有不同。zwQuant量化软件的stkLib为提高效率采用了动态处理的方式，在dataPre数据预处理阶段根据策略要求对股票池数据进行动态的衍生扩展。

同样，为保存大盘数据，我们在zwSys模块定义了一个新的全局变

量：

```
stkInxLib=None   #全局变量，大盘指数，内存股票数据库
```

变量stkInxLIb是全局变量，用于保存大盘数据，也是一种内存数据库的模式。

stkInxLIb只保存一组大盘数据，不像stkLib，采用list模式，保存多个不同股票的数据。

stkInxLIb与stkLib不同，大盘指数往往只是作为背景数据，本身并不参与到回溯测试当中。

如果回溯测试需要大盘数据，可以将大盘指数作为股票添加到stkLib中去。

10.2.3　扩展zwQuantX类变量

为了成功加载数据源，我们还修改了zwSys模块中的zwQuantX类。为zwQuantX类增加了一组stkInx大盘指数相关的内部类变量：

```
# stkInx,大盘指数
self.stkInxCode=''     #大盘指数代码
self.stkInxName=''     #大盘指数名称，拼音
```

```
self.stkInxCName=''     #大盘指数中文名称，拼音
self.stkInxPriceName='close'   #大盘指数数据列名称，默认是:close
self.stkInxRDat=''     #大盘指数数据源路径
```

10.2.4 大盘指数读取函数

同时，为了读取大盘指数文件，在zwQTBox模块还增加了一个大盘指数数据读取函数：

```
def stkInxLibRd(qx):
     '''
        读取指定的大盘数据到zw.stkInxLib

        Args:

     :
     qx.stkInxRDat='\\zwdat\\cn\\xday\\''     #大盘指数数据源路径
     qx.stkInxCode='000001'     #大盘指数代码
     qx.stkInxName='sz001'     #大盘指数名称，拼音
     qx.stkInxCName='上证指数'     #大盘指数中文名称，拼音
     #
     zw.stkInxLib=None   #全局变量，大盘指数，内存股票数据库

     '''
```

```
    if qx.stkInxCode!='':
        fss=qx.stkInxRDat+qx.stkInxCode+".csv";
        xfg=os.path.exists(fss);
        if xfg:
            df10=pd.read_csv(fss,index_col=0,parse_dates=[0]);
            df10=df2zwAdj(df10)
            zw.stkInxLib=df10.sort_index();
```

由此可见，对于zwQuant量化软件本身的修改，与策略函数扩展、绘图模版扩展不同，需要全局可虑，牵一发而动全身。

10.2.5 案例10-2：读取指数

完成以上各种准备工作后，开始测试如何读取大盘指数文件。

案例10-2程序文件名：\zwpython\zw_k10\zqx02_inxRd.py，相关代码如下：

```
# -*- coding: utf-8 -*-


#zwQuant
import zwSys as zw
import zwQTBox as zwx
```

```
#======================


qx=zw.zwQuantX('rdInx',1000);
qx.stkInxRDat='\\zwdat\\cn\\xday\\'     #大盘指数数据源路径
#大盘指数代码，名称拼音，中文名称
qx.stkInxCode,qx.stkInxName,qx.stkInxCName='000001','sh001','上证

#读取大盘指数
zwx.stkInxLibRd(qx)

#输出大盘指数
print(zw.stkInxLib)
```

案例10-2输出结果如下：

```
runfile('E:/zwPython/zw_k10x/zqx02_inxRd.py', wdir='E:/zwPython/
 date          open      high       low     close       volume  a
1994-01-03   837.700   840.650   831.660   833.900    101005600
1994-01-04   835.970   836.970   829.890   832.690     65274300
1994-01-05   829.300   847.050   823.100   846.980     89412100
1994-01-06   850.780   869.330   850.780   869.330    184511700
1994-01-07   875.180   883.990   873.010   879.640    168688400
1994-01-10   891.990   900.300   889.730   900.300    187595100
            ...       ...       ...       ...         ...
2016-03-30 2941.218 3001.106 2941.218 3000.645 21037145100  3000
```

```
2016-03-31 3009.367 3023.409 2992.916 3003.915 22041344500  3003
2016-04-01 2997.088 3009.673 2956.247 3009.530 20654586600  3009
2016-04-05 3000.938 3057.330 2993.148 3053.065 25673034400  3053
2016-04-06 3039.745 3059.794 3029.005 3050.592 23226205100  3050
2016-04-07 3058.339 3062.358 3007.060 3008.420 22091950600  3008
2016-04-08 2988.200 2996.171 2960.460 2984.958 18790463300  2984

[5413 rows x 6 columns]
```

案例10-2使用的是上证指数，从输出信息可以看出，自1994年中国股票开市以来的上证大盘数据全部都有。其他大盘指数起始时间可能有所不同。

10.2.6 大盘数据切割

为了和回溯数据的时间同步，还需要对大盘数据进行切割。为此，在zwQTBox增加了大盘数据切割函数：stkInxLibSet8XTim。

```
def stkInxLibSet8XTim(qx,dtim0,dtim9):
    ''' 根据时间段，切割大盘指数数据 zw.stkInxLib

    Args:
        dtim0（str）: 起始时间
        dtim9（str）:结束时间
```

```
:ivar
zw.stkInxLib，大盘指数数据
'''
df10=zw.stkInxLib
if dtim0=='':
    df20=df10;
else:
    df20=df10[(df10.index>=dtim0)&(df10.index<=dtim9)]

zw.stkInxLib=df20.sort_index();
```

为了便于理解，我们根据大盘数据切割函数stkInxLibSet8XTim绘制了流程图，如图10-4所示。

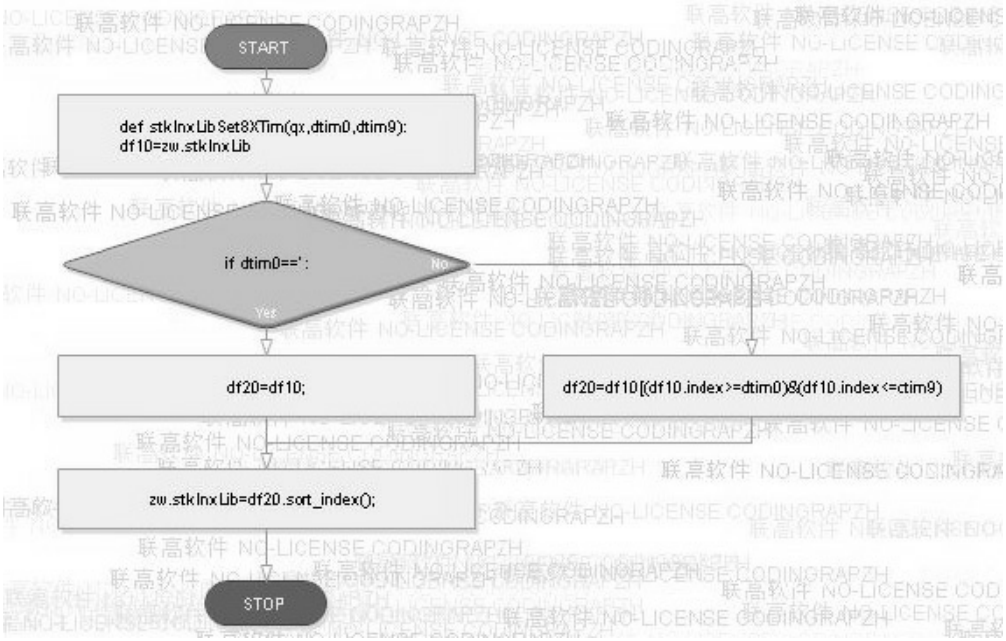


图10-4　大盘数据切割函数：stkInxLibSet8XTim流程图

10.2.7　案例10-3：inxCut数据切割

案例10-3进一步说明如何对大盘数据进行切割。

案例10-3文件名：\zwpython\zw_k10\zqx03_inxCut.py，相关代码如下：

```
# -*- coding: utf-8 -*-

#zwQuant
import zwSys as zw
import zwQTBox as zwx

#=======================

qx=zw.zwQuantX('rdInx',1000);
qx.stkInxRDat='\\zwdat\\cn\\xday\\'     #大盘指数数据源路径
#大盘指数代码，名称拼音，中文名称
qx.stkInxCode,qx.stkInxName,qx.stkInxCName='000001','sh001','上证

#读取大盘指数
zwx.stkInxLibRd(qx)

#输出大盘指数
print(zw.stkInxLib)
```

```
#切割大盘指数
zwx.stkInxLibSet8XTim(qx,'2015-01-01','2016-04-05')

#输出大盘指数
print(zw.stkInxLib)
```

运行案例10-3，得到的切割后的输出数据是：

```
                open      high       low     close       volume
date
2015-01-05  3258.627  3369.281  3253.883  3350.519  53135238400
2015-01-06  3330.799  3394.224  3303.184  3351.446  50166169600
2015-01-07  3326.649  3374.896  3312.211  3373.954  39191888000
2015-01-08  3371.957  3381.566  3285.095  3293.456  37113116800
2015-01-09  3276.965  3404.834  3267.509  3285.412  41024086400
              ...       ...       ...       ...          ...
2016-03-25  2956.200  2981.408  2952.106  2979.434  17507434400
2016-03-28  2988.010  3008.168  2948.526  2957.820  20202124100
2016-03-29  2956.705  2962.203  2905.251  2919.832  18296534500
2016-03-30  2941.218  3001.106  2941.218  3000.645  21037145100
2016-03-31  3009.367  3023.409  2992.916  3003.915  22041344500
2016-04-01  2997.088  3009.673  2956.247  3009.530  20654586600
2016-04-05  3000.938  3057.330  2993.148  3053.065  25673034400
```

```
[305 rows x 6 columns]
```

请注意，切割后的数据包含了起始时间和结束时间。上面的输出数据起始日期是：2015-01-05，因为前面几天是元旦假期，不是交易日。

最后一条数据：2016-04-05，与切割函数的终止日期是一样的。

```
#切割大盘指数
zwx.stkInxLibSet8XTim(qx,'2015-01-01','2016-04-05')
```

10.3　系统整合

通过前面的案例，我们已经知道如何读取大盘数据，并且根据时间进行数据切割。为了提高效率，我们把相关的stkInx大盘指数读取、切割等工作，整合到zwQuant量化软件现有的函数中，这样，大家无须修改策略函数就可以直接使用增强的功能。

这种兼容式的扩展方式在案例10-1中已经可以看到，案例10-1与案例7-1的程序代码完全一样，但案例10-1的输出图形中已经包含了大盘指数的图形曲线。

与zwBacktest.py模块的策略初始化bt_init函数进行整合，把相关的stkInx参数设置放到函数bt_init中。

bt_init函数程序比较长，具体源码请看程序文件，如图10-5所示是bt_init函数的流程图。

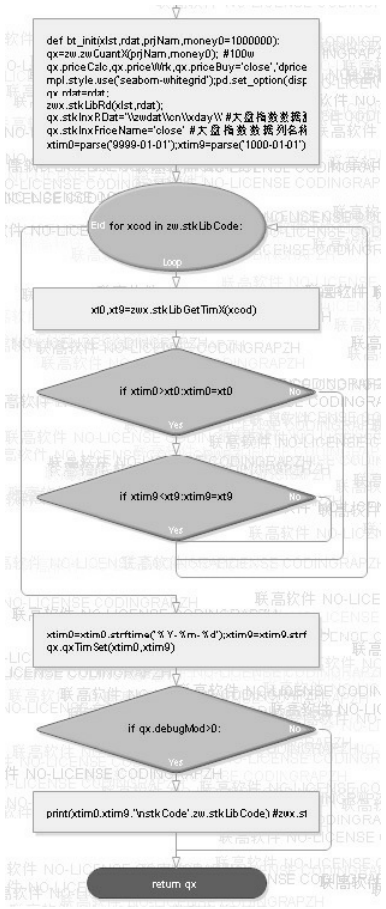


图10-5　bt_init初始化函数流程图

bt_init初始化函数主要是增加了以下几条代码，用于stkInx大盘指数的变量设置语句：

```
#大盘指数.设置
#zw.stkInxLib=None   #全局变量，大盘指数，内存股票数据库
qx.stkInxRDat='\\zwdat\\cn\\xday\\'     #大盘指数数据源路径
qx.stkInxPriceName='close'   #大盘指数数据列名称，默认是:close
#大盘指数代码,名称拼音,中文名称
qx.stkInxCode,qx.stkInxName,qx.stkInxCName='000001','sh001','上证
```

bt_init函数默认使用的大盘指数是上证指数，这个也是用的最多的。

如果使用其他大盘指数，请在主流程调用bt_init后，修改stkInxCode等变量：

```
#大盘指数代码，名称拼音，中文名称
qx.stkInxCode,qx.stkInxName,qx.stkInxCName='000001','sh001','上证
```

10.3.1 案例10-4：整合设置

案例10-4介绍如何调用bt_init初始化函数，设置stkInx相关参数，并且根据读取的股票数据起始时间、终止时间，切割大盘指数数据。

案例10-4程序文件名：\zwpython\zw_k10\zqx04_inxCut2，相关代码如下：

```
# -*- coding: utf-8 -*-

#zwQuant
import zwSys as zw
import zwQTBox as zwx
import zwBacktest as zwbt
```

```
#======================
rss='dat\\'  #rss='\\zwdat\\cn\\day\\'
xlst=['600401']   #600401,*ST海润
qx=zwbt.bt_init(xlst,rss,'inx',10000);



#读取大盘指数
zwx.stkInxLibRd(qx)

#切割大盘指数
zwx.stkInxLibSet8XTim(qx,qx.xtim0,qx.xtim9)

#输出大盘指数
print(zw.stkInxLib)
```

案例10-4运行结果是：

```
runfile('E:/zwPython/zw_k10x/zqx04_inxCut2.py', wdir='E:/zwPytho
2003-09-24 2016-04-08
stkCode ['600401']
               open      high       low     close       volume
date
2003-09-24  1391.404  1396.945  1386.788  1394.229   591952200
2003-09-25  1394.494  1394.737  1374.791  1376.291   478096900
2003-09-26  1374.602  1377.871  1367.746  1370.843   454288900
```

```
2003-09-29  1369.437  1369.477  1353.245  1355.335   487950800
2003-09-30  1354.648  1371.073  1348.222  1367.161   605960700
              ...       ...       ...       ...         ...
2016-04-01  2997.088  3009.673  2956.247  3009.530  20654586600
2016-04-05  3000.938  3057.330  2993.148  3053.065  25673034400
2016-04-06  3039.745  3059.794  3029.005  3050.592  23226205100
2016-04-07  3058.339  3062.358  3007.060  3008.420  22091950600
2016-04-08  2988.200  2996.171  2960.460  2984.958  18790463300

[3044 rows x 6 columns]
```

由以上结果可以看出，切割后，大盘指数数据与股票600401的起始时间、终止时间完全一样。

10.3.2 案例10-5：修改指数代码

案例10-5演示如何修改大盘指数代码，使用新的大盘指数数据。

案例10-5文件名：\zwpython\zw_k10\zqx05_inxNam，全部代码如下：

```
zqx05_inxNam.py代码如下：
# -*- coding: utf-8 -*-
```

```
#zwQuant
import zwSys as zw
import zwQTBox as zwx
import zwBacktest as zwbt

#======================
rss='dat\\'  #rss='\\zwdat\\cn\\day\\'
xlst=['600401']   #600401,*ST海润
qx=zwbt.bt_init(xlst,rss,'inx',10000);
#000300,沪深300,2005-01-01
qx.stkInxCode,qx.stkInxName,qx.stkInxCName='000300','hs300','沪深

#读取大盘指数
zwx.stkInxLibRd(qx)



#切割大盘指数
zwx.stkInxLibSet8XTim(qx,qx.xtim0,qx.xtim9)

#输出大盘指数
print(zw.stkInxLib)
```

案例10-5输出结果如下：

```
runfile('E:/zwPython/zw_k10x/zqx05_inxNam.py', wdir='E:/zwPython
```

```
2003-09-24 2016-04-08
stkCode ['600401']
               open      high       low     close       volume
date
2005-01-04   994.769   994.769   980.658   982.794    741286894
2005-01-05   981.577   997.323   979.877   992.564    711910898
2005-01-06   993.331   993.788   980.330   983.174    628802905
2005-01-07   983.045   995.711   979.812   983.958    729869409
2005-01-10   983.760   993.959   979.789   993.879    579169799
              ...       ...       ...       ...          ...
2016-03-31  3229.200  3241.927  3208.665  3218.088  12155207400
2016-04-01  3213.674  3222.612  3165.856  3221.895  11812213900
2016-04-05  3211.305  3271.930  3205.207  3264.486  14604777200
2016-04-06  3250.522  3267.640  3236.196  3257.528  12019630500
2016-04-07  3266.288  3270.820  3208.644  3209.290  11240946700
2016-04-08  3189.846  3197.774  3163.305  3185.726  10100562500

[2735 rows x 6 columns]
```

从案例10-5输出数据可以看出，股票600401的起始时间是2003-09-24。但沪深300大盘指数最早只搜集到2005年，所以大盘指数数据起始时间是：2005-01-04。

10.3.3 修改sta_dataPre0xtim函数

在zwQuant量化软件中，数据切割部分放在模块zwQTBox的sta_dataPre0xtim函数中。

下面，我们进一步整合，把大盘指数的数据切割整合到sta_dataPre0xtim函数。

修改后的sta_dataPre0xtim数据切割函数代码如下：

```
def sta_dataPre0xtim(qx,xnam0):
    ''' 策略参数设置子函数，根据预设时间，裁剪数据源stkLib

    Args:
        qx (zwQuantX): zwQuantX数据包
        xnam0 (str):  函数标签

    '''

    #设置当前策略的变量参数
    qx.staName=xnam0
    qx.rfRate=0.05;  #无风险年收益，一般为0.05(5%)，计算夏普指数等需要
    #qx.stkNum9=20000;    #每手交易，默认最多20000股
    #
    #按指定的时间周期，裁剪数据源
    xt0k=qx.staVars[-2];xt9k=qx.staVars[-1];
    if (xt0k!='')or(xt9k!=''):
        #xtim0=parse('9999-01-01');xtim9=parse('1000-01-01');
            #xtim0=xtim0.strftime('%Y-%m-%d');xtim9=xtim9.strft
```

```
        if xt0k!='':
            if qx.xtim0<xt0k:qx.xtim0=xt0k;
        if xt9k!='':
            if qx.xtim9>xt9k:qx.xtim9=xt9k;
        qx.qxTimSet(qx.xtim0,qx.xtim9)
        stkLibSet8XTim(qx.xtim0,qx.xtim9);#    print('zw.stkLib
    #---stkInx 读取大盘指数数据，并裁剪数据源
    if qx.stkInxCode!='':
        stkInxLibRd(qx)
        stkInxLibSet8XTim(qx,qx.xtim0,qx.xtim9)

    #============
    #---设置qxUsr用户变量，起始数据
    qx.qxUsr=qxObjSet(qx.xtim0,0,qx.money,0);
```

sta_dataPre0xtim数据切割函数流程图如图10-6所示。

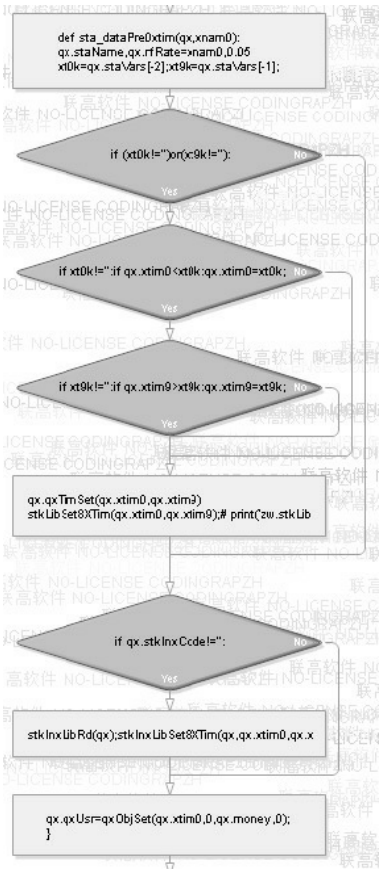


图10-6　sta_dataPre0xtim数据切割函数流程图

sta_dataPre0xtim数据切割函数其他部分没有改动，只是增加了以下几条stkInx大盘指数切割的代码：

```
#---stkInx 读取大盘指数数据，并裁剪数据源
if qx.stkInxCode!='':
     stkInxLibRd(qx)
     stkInxLibSet8XTim(qx,qx.xtim0,qx.xtim9)
```

如果大家不希望使用大盘指数数据，把qx.stkInxCode变量设为空字符串即可。

10.3.4 案例10-6：整合数据切割

案例10-6演示如何使用整合后的sta_dataPre0xtim函数切割大盘指数数据。

案例文件名：\zwpython\zw_k10\ zqx06_cut3.py，代码如下：

```
# -*- coding: utf-8 -*-

#zwQuant
import zwSys as zw
import zwQTBox as zwx
import zwBacktest as zwbt

#=======================
rss='dat\\'  #rss='\\zwdat\\cn\\day\\'
xlst=['600401']   #600401,*ST海润
qx=zwbt.bt_init(xlst,rss,'inx',10000);

#读取大盘指数
zwx.stkInxLibRd(qx)

#切割大盘指数
#zwx.stkInxLibSet8XTim(qx,qx.xtim0,qx.xtim9)
qx.staVars=[0,'','']
```

```
zwx.sta_dataPre0xtim(qx,'inx')


#输出大盘指数
print(zw.stkInxLib)
```

案例10-6输出如下：

```
2003-09-24 2016-04-08
stkCode ['600401']
               open      high       low     close       volume
date
2003-09-24  1391.404  1396.945  1386.788  1394.229    591952200
2003-09-25  1394.494  1394.737  1374.791  1376.291    478096900
2003-09-26  1374.602  1377.871  1367.746  1370.843    454288900
2003-09-29  1369.437  1369.477  1353.245  1355.335    487950800
2003-09-30  1354.648  1371.073  1348.222  1367.161    605960700
2003-10-08  1366.201  1373.081  1359.497  1371.685    432480500
              ...       ...       ...       ...          ...
2016-03-29  2956.705  2962.203  2905.251  2919.832  18296534500
2016-03-30  2941.218  3001.106  2941.218  3000.645  21037145100
2016-03-31  3009.367  3023.409  2992.916  3003.915  22041344500
2016-04-01  2997.088  3009.673  2956.247  3009.530  20654586600
2016-04-05  3000.938  3057.330  2993.148  3053.065  25673034400
2016-04-06  3039.745  3059.794  3029.005  3050.592  23226205100
2016-04-07  3058.339  3062.358  3007.060  3008.420  22091950600
```

```
2016-04-08  2988.200  2996.171  2960.460  2984.958  18790463300

[3044 rows x 6 columns]
```

需要注意的是，在调用sta_dataPre0xtim函数、切割大盘指数数据之前，需要设置好qx.staVars策略变量，不然可能会出错。

```
qx.staVars=[0,'','']
zwx.sta_dataPre0xtim(qx,'inx')
```

此外，也可以通过设置qx.staVars参数自行设置起始时间和终止时间。

10.3.5 修改绘图函数

自此，zwQuant量化软件大盘指数扩展功能接近完工，只差最后一个环节：绘图输出。

zwQTDraw.py绘图模块目前只有以下两个函数。

- dr_quant3x_init：zwQuant三合一模版，绘图模块初始化。
- dr_quant3x：zwQuant三合一模版，绘图输出。

大盘指数图形的位置，对于zwQuant量化软件选择的三合一模版有

上、中、下三种选择。

- 上部，已经是dret日均回报率+成交量模式，没有位置了，除非把其他一块移走。
- 中部，原本是重点考虑的位置，不过大盘指数和个股价格数据差别太大，影响曲线幅度。
- 下部，只有val资产总值和空位。

最终，大盘指数的位置定在下部区域，虽然不是很理想，但这种布局也在可以接受的范围之内。其他的方案（例如采用全新的模版，四和一矩阵等模式）代码工作量太大，兼容性也有问题，就不在考虑范围之内了。

dr_quant3x_init图形数据初始化函数改动不大，只是把原来屏蔽的几条代码解除屏蔽即可。

主要修改的是dr_quant3x绘图函数，为了避免破坏原函数，把原来的dr_quant3x函数改名为dr_quant3x00，大家可以对比参考。

dr_quant3x绘图函数源码比较长，请大家自己查看程序文件，如图10-7所示是函数的流程图。

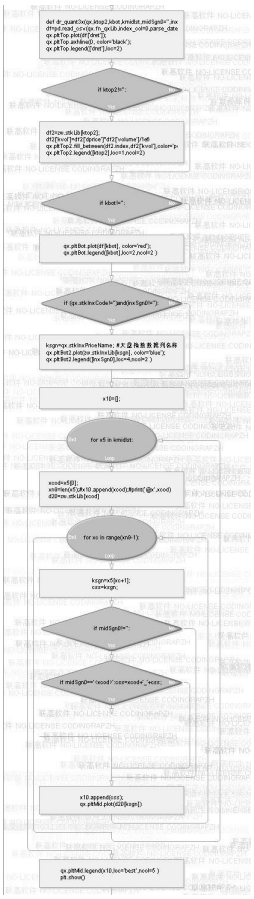


图10-7　dr_quant3x绘图函数流程图

dr_quant3x绘图函数其他部分改动不大，只是增加了几条stkInx大盘指数相关的代码：

```
#---stkInx，大盘指数数据
if  (qx.stkInxCode!='')and(inxSgn0!=''):
     ksgn=qx.stkInxPriceName;  #大盘指数数据列名称，默认是:close
     qx.pltBot2.plot(zw.stkInxLib[ksgn], color='blue');
     qx.pltBot2.legend([inxSgn0],loc=4,ncol=2 )
```

另外，为保持兼容，在函数变量方面增加了一个大盘指数名称变

量：inxSgn0。

```
def dr_quant3x(qx,ktop2,kbot,kmidlst,midSgn0='',inxSgn0=''):
```

大盘指数名称变量inxSgn0默认是空字符串，不绘制大盘指数图形。

dr_quant3x绘图函数只有在大盘指数有数据，而且inxSgn0有赋值时，才绘制大盘图形。同时，inxSgn0的赋值作为图标名称显示在下部图形框中。

10.4　扩展完成

自此，zwQuant量化软件大盘指数扩展功能便全部完工，我们通过案例来看看扩展后的结果。

案例10-7：SMA均线扩展策略

案例10-7是案例10-1的A股修改版，本质上，案例10-7也是案例2-1的衍生版本，改动不多：

- 数据源改为国内股票数据。
- 使用沪深300指数作为大盘指数。

案例10-7有关代码请浏览程序文件：\zwpython\zw_k10\zqx07_sma300.py。

案例10-7与其他案例类似，修改绘图函数dr_quant3x的调用参数：

```
zwdr.dr_quant3x(qx,xcod,'val',kmid8,'','hs300')
```

注意，与其他案例不同的是，函数最后增加了一个字符串赋值：hs300，表示大盘指数是沪深300。

运行结果如图10-8所示。

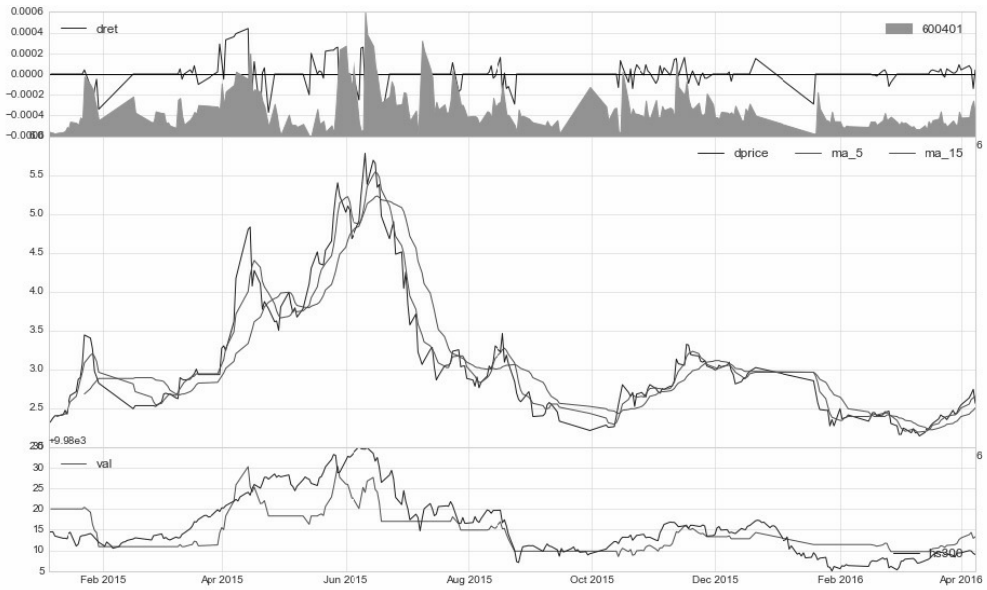


图10-8　运行结果

案例10-8其他输出数据如下：

```
交易总次数：25
交易总盈利：-6.70

盈利交易数：12
盈利交易金额：75.00
亏损交易数：13
亏损交易金额：-81.70

最终资产价值 Final portfolio value: $9993.30
最终现金资产价值 Final cash portfolio value: $9967.00
最终证券资产价值 Final stock portfolio value: $26.30
累计回报率 Cumulative returns: -0.07 %
平均日收益率 Average daily return: -0.000 %
```

```
日收益率方差 Std. dev. daily return:0.0001

夏普比率 Sharpe ratio: -28.075,（0.05利率）
无风险利率 Risk Free Rate: 0.05
夏普比率 Sharpe ratio: -0.363,（0利率）

最大回撤率 Max. drawdown: 0.2178 %
最长回撤时间 Longest drawdown duration: 316
回撤时间(最高点位) Time High. drawdown:  2015-05-28
回撤最高点位 High. drawdown: 10010.300
回撤最低点位 Low. drawdown: 9988.500

时间周期 Date lenght: 464 (Day)
时间周期（交易日） Date lenght(weekday): 257 (Day)
开始时间 Date begin: 2015-01-01
结束时间 Date lenght: 2016-04-08

项目名称 Project name: sma
策略名称 Strategy name: sma
股票代码列表 Stock list:  ['600401']
策略参数变量 staVars[]:  [5, 15, '2015-01-01', '']
```

所有运行结果、数据和图形完全正常。

10.5　其他扩展课题

虽然zwQuant量化程序大盘指数扩展功能已经完成，但zwQuant量化程序的扩展工作只是刚刚开始。下面再介绍两个实用的扩展功能：

- 复权数据冲突
- 支持空头交易

10.5.1　复权数据冲突

复权数据冲突是因为有些股票在拆分、配股、停盘前后数据差异太大，出现异常波动。特别是雅虎财经API、tusahre数据抓取模块等这些公众版本的免费数据源。

对于复权数据冲突，一种简单的解决方案就是：从数据源开始，全部采用复权数据，类似雅虎的adj close复权收盘价模式。

zwQuant量化软件采用tushare模块库的get_h_data复权抓取函数，默认数据模式是：前复权。

不过，即使是采用adj　close复权收盘价模式，或者专业的收费财经数据源，依然会碰到股票在拆分、配股、停盘时的这些问题。这种方式还有一个缺陷，即对于均线数据无法采用这种简单的方式进行处理。

因此，还是需要对这些出现数据异常的节点进行标明，对于异常节点时间出现的交易视为无效交易，尽量减少异常数据对于回溯测试的干

扰。

10.5.2 波动率指标DVIX

我们需要在原始的数据源基础上增加一个相邻日期的波动率指标。波动率指标的英文缩写是VIX。根据国内涨停板、跌停板政策，凡是变动率上下超过10%都属于无效数据，当天交易属于无效交易，这个可以在zwQTBox模块的有效交易检查函数xtrdChkFlag中查看。

波动率上限默认是10，但是可以人工修改，美股数据也可以修改波动率上限。

扩展zwQuant量化软件时，增加VIX波动率指标最简单的方法是在读取数据源的时候增加。zwQuant量化软件量化系统的价格体系有多种模式，需要在策略数据预处理函数中指定策略采用的价格模式。因此，VIX波动率指标只能在策略数据预处理函数之后进行扩展，而且必须支持单个、多个股票数据源模式。

在实际扩展中是基于dprice策略分析的价格计算VIX波动率参数，所以，数据列名称是DVIX。DVIX数据列扩展代码可以增加在策略数据预处理函数中，不过，这样需要修改每个策略对应的数据预处理函数，工作量较大。

因此，zwQuant量化软件采用独立的函数：stkLibSetDVix，代码函数在zwQTBox模块，相关代码如下：

```
def stkLibSetDVix():
    ''' 根据时间段，切割股票数据

    Args:
        dtim0（str）: 起始时间
        dtim9（str）:结束时间

    :ivar xcod (int): 股票代码
    '''
    for xcod in zw.stkLibCode:
        df10=zw.stkLib[xcod]
        df10['dvix']=df10['dprice']/df10['dprice'].shift(1)*100
        #
        zw.stkLib[xcod]=df10;
```

stkLibSetDVix函数很简单，只是增加一个DVIX每日波动率的数据列。

10.5.3 修改回溯主函数zwBackTest

函数的调用接口在zwBacktest.py回溯模块的回溯主函数：zwBackTest。

zwBackTest回溯函数代码较长，请大家参看模块程序文件

zwBacktest.py，如图10-9所示是zwBackTest回溯函数的流程图。

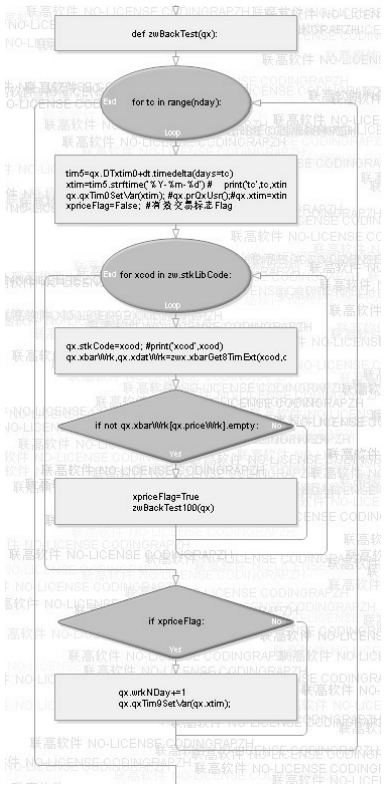


图10-9　zwBackTest回溯函数流程图

修改后的zwBackTest回溯函数，在函数开始部分增加了以下几条语句：

```
# 增加数据源波动率参数
    zwx.stkLibSetDVix()


    if qx.debugMod>0:
        xcod=zw.stkLibCode[0];
        print(zw.stkLib[xcod].tail())
```

增加的代码，放在zwBackTest函数开头的好处是：这个位置刚好是策略数据预处理运行完、策略分析还没开始的环节。在这里增加stkLibSetDVix函数代码可以为日后扩展其他功能、留下接口和参考。下面的代码是输出增加DIVX后的数据源，供大家参考。

此外，还需要在zwQuantX类定义里面增加DVIX波动率的上限值与下限值。

```
self.dvix_k0=80        #dvix波动率下限
self.dvix_k9=120       #dvix波动率上限
```

用户可以自行修改这两个数值，也可以用于美股或其他领域。

虽然国内政策的涨停板、跌停板的幅度是10%，但在实际回溯测试时，因为价格模式等数据的问题，如果采用上限90、下限110作为默认参数，会经常出现误报，所以将数值放宽到80%~120%。

波动率的检查代码也在zwBackTest函数内：

```
if not qx.xbarWrk[qx.priceWrk].empty:
          #-----dvix 波动率检查
          dvix=zwx.stkGetVars(qx,'dvix');#dvixFlag=False;
          dvixFlag=zwt.xinEQ(dvix,qx.dvix_k0,qx.dvix_k9)or(dvix==
          if dvixFlag:
              xpriceFlag=True
```

```
            # 调用回溯子程序，如果是有效交易，设置成功交易标志xtrdFlag
            zwBackTest100(qx)
        else:
            print('@dvix',xcod,xtim,dvix);
            pass;
```

如果DVIX波动率在正常范围内，再调用回溯子程序zwBackTest100。如果变动率异常，则会输出提示信息，包括股票代码、异常时间对应的DVIX波动率。

将DVIX波动率检查放在调用zwBackTest100子程序之前，而不是放在zwQTBox模块中的xtrdChkFlag有效交易检查函数中是为了提高效率。

10.5.4 案例10-8：波动率

案例10-8是介绍波动率，相关代码请看程序文件：\zwpython\zw_k10\zqx08_ dvix.py。

案例10-8需要注意的是，为了检查DVIX变动率的效果，在程序中策略回溯时间跨度很长，是从2000年开始，到2016年4月结束，运行时间较长。

```
qx.staVars=[5,15,'2000-01-01','']
```

案例10-8运行结果如图10-10所示。

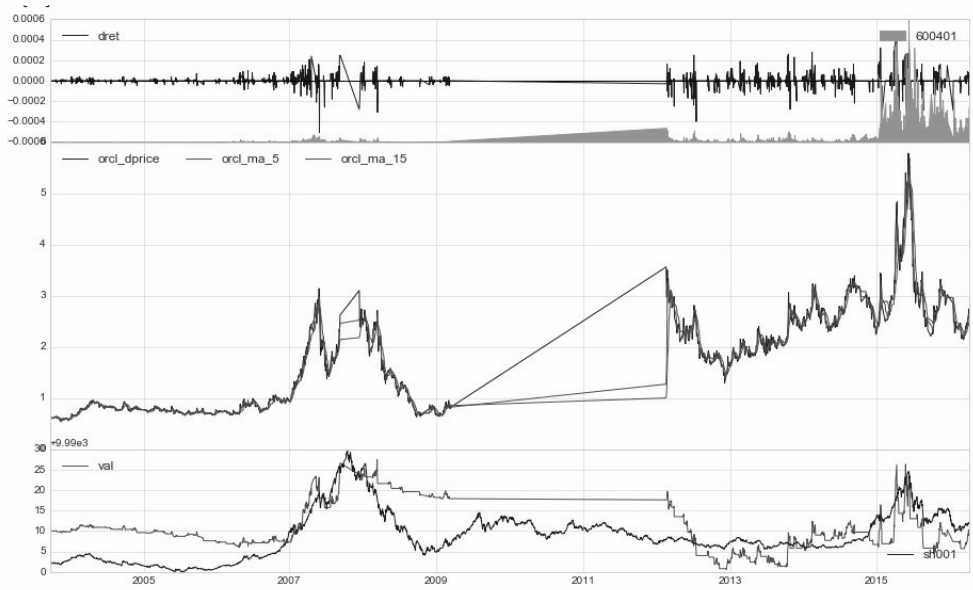


图10-10 DIVX波动率

案例10-8其他运行数据如下：

```
最终资产价值 Final portfolio value: $9999.30
最终现金资产价值 Final cash portfolio value: $9973.00
最终证券资产价值 Final stock portfolio value: $26.30
累计回报率 Cumulative returns: -0.01 %
平均日收益率 Average daily return: -0.000 %
日收益率方差 Std. dev. daily return:0.0001

             open  high   low  close       volume  adj close  dpr
date
2016-04-01  2.54  2.62  2.53   2.60  111421383.0       2.60    2
2016-04-05  2.63  2.68  2.63   2.68  110368219.0       2.68    2
```

```
2016-04-06  2.68  2.80  2.66   2.73  171821194.0       2.73    2
2016-04-07  2.74  2.78  2.59   2.59  198460790.0       2.59    2
2016-04-08  2.56  2.65  2.55   2.63  108273174.0       2.63    2
nday 4581
@dvix 600401 2003-09-24 nan
@dvix 600401 2007-05-08 122.17
@dvix 600401 2007-12-17 77.42
@dvix 600401 2012-02-17 434.15
```

注意，输出信息中的DVIX数据列，在运行结果中的回报率是-0.01%，这个明显是不正常的。此外，还有@divx异常提示信息，我们注意最后一条：

```
@dvix 600401 2012-02-17 434.15
```

变动率是434%，明显有异常情况发生，通过百度检索发现原因是：600401这只股票在2012年2月17日不再是ST股了。从ST申龙重组改名为海润光伏。前期重组完成后，经过长期停牌恢复上市，第一个交易日可以随意涨，也可以随意跌。这就叫作重组复牌第一个交易日不设涨跌幅（度）。

对于这种异常数据，zwQuant量化软件的解决方法是作为无效交易处理，类似Pandas数据分析软件的NaN（空）数据，这个也是数据清洗的一种常用手段。

10.5.5 空头交易

原本zwQuant量化软件不支持空头交易，不过买空、卖空是期货行业的基本要求。下面，我们通过案例介绍如何为zwQuant量化软件增加空头交易支持。

通过多个方案的测试发现zwQuant量化软件的空头交易模块类似伪命题，因为zwQuant量化软件只涉及策略分析、策略回溯，并不进行真正的下单交易；具体的交易程序和资金管理其实是由第三方交易平台、软件完成的。而回溯测试无所谓多头、空头，你只需根据策略进行买进、卖出操作即可。

实盘测试盈利的策略通常在空头操作模式下，类似配资的杠杆作用，也是盈利的，当然，谨慎的用户可以把对策略的预期盈利率放宽10%~20%。

“空头”的“空”是相对于未来的，对于面向过去的回溯测试没有所谓的空头，全部是实盘已经产生的数据。对于实盘操作而言，期货操盘手根据回溯测试的策略、系统推荐的产品及操作指令买进、卖出即可。

对于空头交易，zwQuant量化软件原本有两条限制：

- 买入时，交易金额不能超过用户持有的现金数额。
- 卖出时，用户手中必须持有足够数目的对应股票。

这两条限制都在zwQTBox模块的xtrdChkFlag交易检查函数中。我们解除xtrdChkFlag交易检查函数中的以上两条限制，即可为zwQuant量化软件增加空头交易功能。为了回避空头交易对最后回溯统计的影响，

减少代码修改的环节，我们采用虚拟空头交易的模式。

10.5.6 虚拟空头交易

所谓“虚拟空头交易”模式就是空头订单，即只发出交易提示信息，并不统计到最后的回溯测试结果中。

增加虚拟空头交易功能后，在回溯或实盘时，再遇到这种情况系统会提示有符合策略要求的交易，并提供股票名称和交易时间。为了保存这些虚拟交易记录，我们在zwQuantX模拟xtrdLib交易记录清单增加了一个xtrdNilLib变量，用于保存虚拟交易记录：

```
self.xtrdLib=pd.DataFrame(columns=xtrdName,index=['date']);    #所
self.xtrdLib.dropna(inplace=True);
self.xtrdNilLib=pd.DataFrame(columns=xtrdName,index=['date']);
self.xtrdNilLib.dropna(inplace=True);
```

同时，zwQuantX类，还增加了一个布尔变量，用于表示空头交易标志：

```
self.trdNilFlag=False;  #空头交易标志
```

为了方便输出交易记录，我们为zwQuantX增加了内部方法函数：

prTrdLib。

```
def prTrdLib(self):
    ''' 输出xtrdLib、xtrdNilLib 各种实盘、空头交易数据，一般用于结束时
    '''
    print('\n::xtrdNilLib 空头交易')
    print(self.xtrdNilLib)
    print('\n::xtrdLib 实盘交易',self.fn_xtrdLib)
    print(self.xtrdLib)
```

10.5.7 修改检查函数

zwQuant量化软件需要修改的地方还有zwQTBox模块的xtrdChkFlag交易检查函数，以及zwBackTest回溯模块的回溯测试子函数：zwBackTest100。

xtrdChkFlag交易检查函数代码较长，请大家自己浏览zwQTBox模块代码文件，如图10-11所示是xtrdChkFlag交易检查函数流程图。

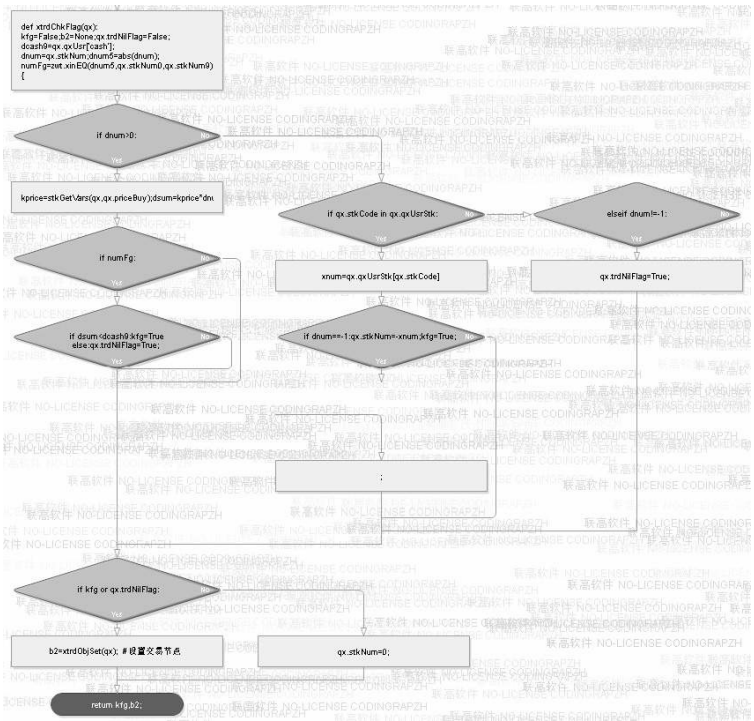


图10-11　xtrdChkFlag交易检查函数流程图

原来的xtrdChkFlag交易检查函数，模块文件里面已经更名为xtrdChkFlag00，大家可以对比一下。

zwBackTest回溯模块的回溯测试子函数zwBackTest100变动不大，修改后的代码如下：

```
def zwBackTest100(qx):
    '''
    zwBackTest100(qx):
    zwQT回溯测试子函数，测试一只股票xcod，在指定时间xtim的回溯表现数据
    会调用qx.staFun指定的策略分析函数，获取当前的股票交易数目 qx.stkNum
    并且根据股票交易数目 qx.stkNum，判定是不是有效的交易策略
    【输入】
```

```
        qx (zwQuantx): zwQuantx数据包
        #qx.stkCode，当前交易的股票代码
        #qx.xtim，当前交易的时间
    【输出】
         无
         '''

    #----运行策略函数，进行策略分析
    qx.stkNum=qx.staFun(qx);
    #----
    if qx.stkNum!=0:
        #----检查是不是有效交易
        xfg,qx.xtrdChk=zwx.xtrdChkFlag(qx)
        if xfg:
            #----如果是有效交易，加入交易列表
            zwx.xtrdLibAdd(qx)
            #qx.prQCap();
        elif qx.trdNilFlag:
            zwx.xtrdLibNilAdd(qx)
```

和原来的版本相比只是增加了一个trdNilFlag空头标志检查，

```
elif qx.trdNilFlag:
       zwx.xtrdLibNilAdd(qx)
```

符合trdNilFlag要求的调用xtrdLibNilAdd空头贸易记录函数，保存相关的空头贸易。

xtrdLibNilAdd空头贸易记录函数是新增加的函数，位于zwQTBox模块，代码如下：

```
def xtrdLibNilAdd(qx):
    ''' 添加交易到空头记录 xtrdNilLib

    Args:
        qx (zwQuantX): zwQuantX数据包

    '''
    qx.xtrdChk['ID']='nil';
    qx.xtrdNilLib=qx.xtrdNilLib.append(qx.xtrdChk.T,ignore_ ind
```

10.5.8 案例10-9：空头数据

案例10-9介绍如何使用空头模式。

空头案例10-9源自案例6-1，文件名：zq601_k402inv.py，案例6-1是PAT量化程序中的QSTK作业之一，因为支持人工下单，便于说明空头模式。

案例10-9程序代码较长，请大家参看程序文件：

\zwpython\zw_k10\zqx09_nil.py。

案例10-9与其他案例基本类似，只是在数据输出函数bt_endRets中，最后新增的qx.prTrdLib函数增加了空头、实盘的交易记录输出。同时，为了强化空头交易，我们把起点资金改为10000元，原来是100万元。

在运行结果中，我们只看最后的qx.prTrdLib函数，输出数据即可：

```
::xtrdNilLib 空头交易
         date   ID  mode  code  dprice     num  kprice      sum
0  2011-01-03  nil  sell  aeti    2.23 -2000.0    2.23  -4460.0
1  2011-01-03  nil  sell  egan    1.20 -1000.0    1.20  -1200.0
2  2011-01-03  nil   buy  glng   12.20  1000.0   12.20  12200.0
3  2011-01-04  nil  sell  glng   12.05 -2000.0   12.05 -24100.0
4  2011-01-05  nil   buy  glng   12.24  1000.0   12.24  12240.0

::xtrdLib 实盘交易 tmp\inv01_xtrdLib.csv
         date            ID  mode  code  dprice     num  kprice
0  2011-01-03  inv01_000001   buy  simo    3.98  1000.0    3.98
1  2011-01-04  inv01_000002   buy  aeti    2.23   500.0    2.23
2  2011-01-04  inv01_000003   buy  egan    1.30   500.0    1.30
3  2011-01-05  inv01_000004  sell  aeti    2.16  -500.0    2.16
4  2011-01-05  inv01_000005  sell  egan    1.34  -500.0    1.34
```

从以上结果可以看出，当用户手中资金不足，或没有对应的现货

时，系统会提示相关的股票代码，以及交易模式（mode）是买进还是卖出，至于具体数额虽然也有提示，但只供参考，并不计入回溯结果。用户可以根据材料要求，按比例（例如90%）或定额（例如100手），进行买进、卖出的操作。对于实盘业绩，目前的交易软件都有资金管理模块，可以自动计算用户最终的业绩、盈利率等指标。

zwQuant量化软件的扩展不仅DVIX和空头交易这两点，未来版本会不断对zwQuant量化软件进行扩展与优化。

对于初学者而言，zwQuant量化软件的扩展有些复杂。zwQuant量化软件的扩展有些难度，将此作为本书的最后一章，也是给大家一些启发，发挥承前启后的作用。

10.6 终点与起点

自此，本书内容就全部结束了，相信大家通过对本书的学习，以及对zwQuant量化软件和配套代码的练习，基本上都能够掌握基本的Python量化策略的应用与编写。同时，对于Python量化领域的整体状况、运作流程，从宏观到微观，都有了一定的认识。

进阶课程

对于初学者而言，下一步，需要学习和补充的课程如下。

- 实盘、实盘练习、大量的实盘练习。
- 各种策略的回溯测试。
- 至少掌握一种全自动的量化回溯测试系统。
- 常用数据挖掘、分析统计的算法。
- 常用的量化分析、统计建模手段与编程。
- zwQuant量化软件架构与模块库的扩展与维护。
- 常用Python性能优化模块库和手段。
- 常用人工智能、机器学习、统计分析模块库。
- 常用Python绘图模块库。
- 常用Python语义分析模块库。

……

以上内容看起来不多，其实，其中的每一项都可以作为一个博士课题进行深入研究。关键是学会、学懂。

附录A
zwPython开发平台用户手册

Talk is cheap, show me the code!

——Linus（linux发明人）

zwPython，字王集成式Python开发平台，比PythonXY更强大、更方便，具体如下：

- 内置Opencv、CUDA/Opencl、NLTK自然语言、Pygame游戏设计等多个重量级模块库。
- 绿色设计，无须安装，解压即可直接使用，可作为U盘便携式、随身开发平台。

zwPython主要应用领域：

- Opencv视觉/人脸识别。
- 机器学习，支持scikit-learn、Theano、Pattern等。
- 大数据处理，内置Pandas等模块。
- 自然语言处理，NLTK、SpaCy等模块。
- CUDA/Opencl，高性能技术GPU编程，内置Python Cuda、Python Opencl模块。
- 图像处理，内置PIL、MATPLOTLIB等模块。
- 字体设计，Fonttools模块。

- 游戏设计，支持Pygame。

……

系统内置多个模块库示例源码作为配套，保存在demo目录下。

简而言之：

zwPython=图像处理+游戏开发+数据分析+机器学习+人工智能+字体设计+GPU并行计算+……

为什么选择**Python**

“Talk is cheap，show me the code！”这是IT界第一牛人，Linux发明人林纳斯的一句名言。就像华尔街流行的商业名言，谈任何项目、投资，请先：Show me the money。

之所以选择Python，是因为Python是目前IT行业唯一的入门简单、功能强大的工业级开发平台。事实上，Python已经超越普通编程语言，成为IT行业的万能开发平台。

入门简单

任何熟悉JavaScript脚本、Visual Basic、C、Delphi的用户，通常一天即可学会Python。

即使是不会编程的设计师、打字员，一周内也能熟练掌握Python，其学习难度绝对不会比学习五笔高。至少笔者现在还不会使用五笔。

功能强大

海量级的Python模块库提供了IT行业最前沿的开发功能。

- 大数据：Pandas已经逐步碾压R语言。
- CUDA：高性能计算，Python、C（C++）、FORTRAN是NV官方认可的三种编程语言，也是目前唯一适合PC平台的CUDA编程工具。
- 机器学习：scikit-learn、Theano、Pattern是国际上最热门的机器学习平台。
- 自然语言：NLTK，全球首选的自然语言处理平台；spaCy，工业级NLP平台。
- 人脸识别：OpenCV，光流算法、图像匹配、人脸算法，简单、优雅。
- 游戏开发：Pygame提供图像、音频、视频、手柄、AI等全套游戏开发模块库。
- 字体设计：Fontforge，是唯一商业级的字体设计开源软件，内置脚本和底层核心的Fonttools都是用Python开发的。
- 电脑设计：Blend、GIMP、Inkscape、MAYA、3DS都内置或扩展了Python语言支持。

目前热门的iOS、安卓、WP等手机APP的应用开发也可以使用Python，但基本都是商业收费模块，因此未集成到zwPython软件包，大家可以自行搜索。

既然Python如此美好，而且是100%免费的开源软件，学习Python的人越来越多，为什么Python始终还只是一个小众语言呢（相对于Visual Basic、JavaScript、Visual Basic、C、C#来说）？

笔者认为，Python的“大众化”之路存在以下两个瓶颈。

- 配置：软件行业有句俗话，“搞懂了软件配置，就学会了一半”。对于Python和Linux以及许多开源项目而言，80%的问题都出现在配置方面、尤其是模块库的配置。
- OOP（面向对象程序设计）：大部分人都认为Python是一种“面向对象”的编程语言，而OOP是业界公认的无比繁杂的编程风格。

如果能够解决好以上两个问题，学习Python的难度可以降低90%，而在应用领域和开发效能上，可以瞬间提升十倍，而且这种提升是零成本的。

zwPython难度降低**90%**，性能提高十倍

为此，笔者在WinPython软件包的基础上推出了“zwPython”：字王集成式Python开发平台。

- 业界首次提出“零配置、零对象”研发理念，绿色软件封装模式，类似MAC开箱即用风格，无须安装，解压即可直接使用，更可放入U盘，支持MobileAPP移动式开发编程。
- “外挂”式“核弹”级开发功能，内置N多IT前沿具有强大功能的开发模块库，例如OpenCV视觉/人脸识别、CUDA高性能GPU并行计算（OpenCL）、Pandas大数据分析、机器学习、NLTK自然语言处理等。
- 便于扩展：用户可以轻松增删相关模块库，全程智能配置，无须用户干预，就像拷贝文件一样简单，而且支持U盘移动便携模

式，真正实现了“一次安装，随处可用”。

- 针对中文开发文档缺乏、零散放入的缺点，内置多部中文版OpenCV、Fontforge和Python入门教材。
- 内置多款中文开源truetype字库。
- 大量示例脚本源码，涵盖：OpenCV、CUDA/OpenCL、Pygame等。

……

推出zwPython就是为了便于IT行业外的用户学习Python时能够“零起步”快速入门，并且短时间内应用到生产环节中去。

zwPython的前身是zw2015sdk（字王智能字模设计平台），原设计目标是向广大设计师提供一款统一的、可编程的字体设计平台，以便于大家交流。经过多次扩充，发现zwPython已经基本覆盖了目前IT编程领域90%的应用，因此又增加了一些模块，将zwPython扩展成一个通用的、集成式Python开发平台。

“零对象”编程模式

虽然，很多人认为Python是一种“面向对象”的编程语言。不过，对于初学者而言，把Python视为一种Basic风格的、过程式入门语言，学习难度可以降低90%，甚至学习一个小时，即可动手编写代码。

关于“零对象”补充以下几点。

- 不写“面向对象”风格的代码，不等于不能使用，对于各种采用对

象模式开发的模块库，仍然可以直接调用。

- 将Python视为非“面向对象”语言，并非大逆不道，事实上，许多人认为Python也是一种“函数”编程语言。
- 笔者从事编程工作十几年，从未用过“面向对象”模式编写过一行“Class”（类对象）代码，依然可以应对各种编程工作。
- “面向对象”过于复杂，与“人生苦短，我用python”的优雅风格天生不合。

zwQuant量化家族成员

经过极宽开源团队成员的不断努力，目前，zwQuant量化包括以下项目。

- zwPython：业界领先Python数据分析、量化回溯开发平台，绿色版本，开箱即用。
- zwQuant：业界首套第三代、基于矩阵运算的量化回溯测试系统。
- zwDat：国内首个大型的免费开源金融数据包，数据量超过300GB，提供tick级分笔数据。
- zw-down：金融数据下载程序，提供源码，支持中国A股、美股数据下载、更新、去重、追加，并提供tick数据下载，以及tick与分时数据转换程序。
- zw-TA-Lib：zw版TA-Lib金融函数包，与Pandas无缝集成，直接调用。

目前，极宽开源团队已有上百位活跃成员，已完成的项目包括：

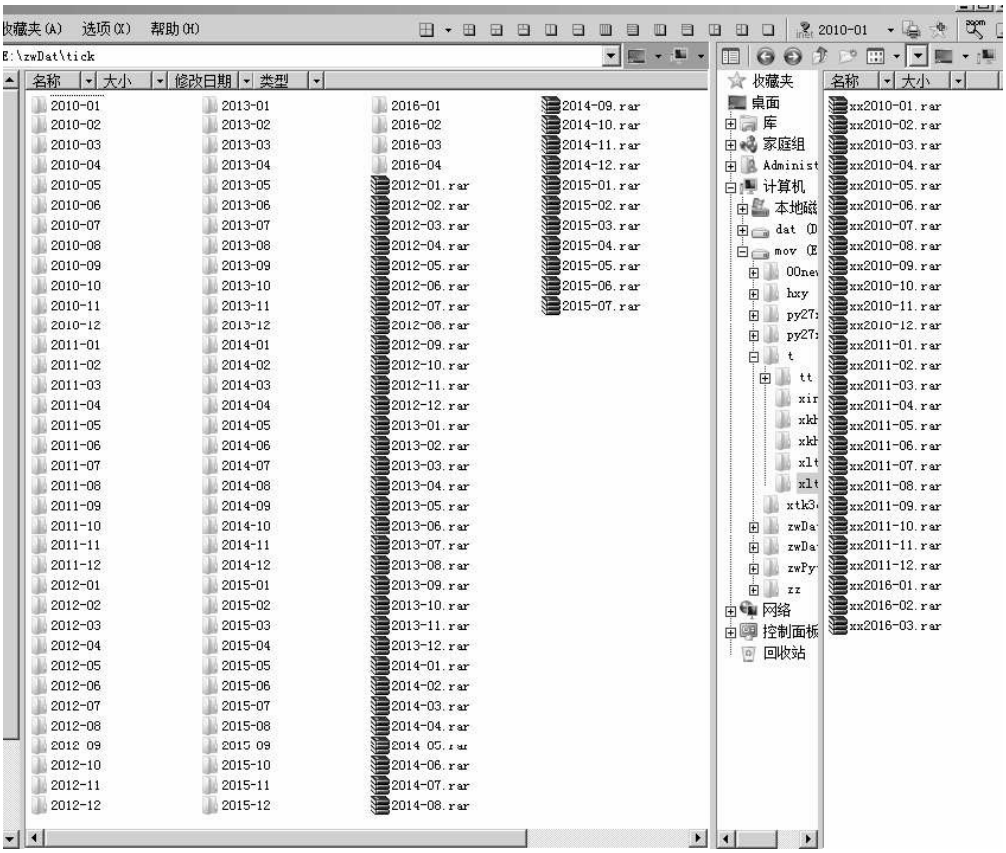
《Python量化项目总览》、《Seabborn中文指南》、zwQuant系列程序的函数级中文注解等。极宽系列开源课件是业内首个专业户的量化开源课件系列，目前已经推出十多部作品。《Python量化·入门十讲》是国内首个专业的Python量化入门教学课件系列，包括PDF图文课件、高清视频和配套的教学代码。

欢迎大家加入Python量化QQ总群：124134140，极宽开源量化团队QQ群：533233771（面向技术人员）。Python量化网站http://www.ziwang.com，如图A-1所示。



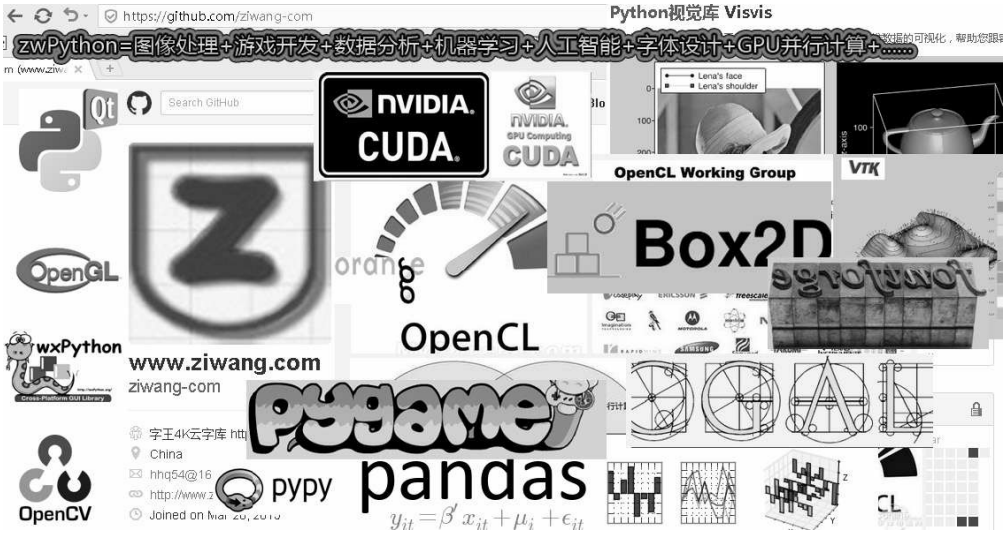
图A-1 Python量化网站http://www.ziwang.com

如图A-2所示是zwDat金融数据包文件截图，zwDat大型开源金融数据包总数据超过300GB，提供tick级别的分笔数据。



图A-2　zwDat金融数据包

zwPython开发平台是工业级Python数据分析、量化回溯开发平台，如图A-3所示。



图A-3　zwPython开发平台

zwPython与winPython、pythonXY

- zwPython是对WinPython进行的二次封装和扩展，增加了大量的模块，无须用户下载安装即可直接使用，类似于Ubuntu和Debian的关系。
- WinPython是PythonXY作者的新作品，对PythonXY进行了精简，并提供了Python3版本的支持。
- zwPython对于WinPython的扩展，从很多功能方面已经超越了PythonXY，可以视为PythonXY++版本，而WinPython类似于PythonXY的精简版）。

相对于WinPython软件，zwPython开发平台在以下方面做了增强。

- 增加多个重量级模块库：OpenCV、CUDA、OpenCL、Pygame、Pandas，强化了系统在大数据、机器学习、人工智能、游戏开发方面的功能。
- 内置多套开源中文Truetype字库。
- 作为首款源自中国的Python集成式开发平台，增加了大量的中文资源和文档支持Python入门教材。
- 大量的demo示例源码，包括：OpenCV、Pygame、OpenCL、Pandas等。
- 二次封装大大简化了用户接口；内置CURL程序。

zwPython升级要点

相比WinPython，zwPython增加的模块约30个，其中大部分为重量

级模块，如OpenCV、Pycuda、Pygame等。

部分重点模块说明如下。

- PyOpenGL：三维图像OpenGL接口。
- SciTools：Python工具函数库。
- ScientificPython：一组经过挑选的Python程序模块，用于科学计算。
- CGAL：计算几何算法库（Computational Geometry Algorithms Library，简称CGAL），提供计算几何相关的数据结构和算法。
- Chardet：字符编码探测器，可以自动检测文本、网页、XML的编码。
- Cytoolz：Toolz的Cython实现，高性能的函数编程工具。
- Epydoc：从源码注释中生成各种格式文档的工具。
- Eventlet：开销很少的多线程模块，使用的是 green threads 概念，例如，pool = eventlet.GreenPool(10000)这样一条语句便创建了一个可以处理10000个客户端连接的线程池。类似Gevent线程库。
- Fonttools：字库设计工具包。
- Gensim：一个相当专业的主题模型Python工具包，无论是代码还是文档，都可用于计算两个文档的相似度。
- Gevent：多线程模块。
- Libsvm：Libsvm是林智仁（Lin Chih-Jen）教授等开发设计的一个简单、易于使用和快速有效的SVM模式识别与回归的软件包，该软件对SVM所涉及的参数调节相对比较少，提供了很多的默认参数，利用这些默认参数可以解决很多问题；并提供了交互检验（Cross Validation）的功能。该软件可以解决C-SVM、V-SVM等问题，包括基于一对一算法的多类模式识别问题。

- MILK：机器学习工具箱，其重点是提供监督分类法与几种有效的分类分析，比如SVMs、K-NN、随机森林经济和决策树，还可以进行特征选择。这些分类可以在许多方面相结合，形成不同的分类系统。对于无监督学习提供K-means和affinity propagation聚类算法。
- OpenCV：目前最好的开源图像/视觉库，包括图像处理和计算机视觉和机器学习的很多通用算法。
- PyHook：基于Python的“钩子库”，主要用于监听当前计算机上鼠标和键盘的事件。这个库依赖于另一个Python库PyWin32，如同其名，PyWin32只能运行在Windows平台，所以PyHook也只能运行在Windows平台。
- Pycairo：开源矢量绘图库Cairo（开罗）的Python接口，Cairo提供在多个背景下做2-D的绘图，高级的可以使用硬件加速功能。
- Pycuda：CUDA高性能GPU并行计算。
- Pycurl：网络模块库。
- Pygame：著名的游戏开发套件。
- Pymvpa2：为大数据集提供统计学习分析的Python工具包，它提供了一个灵活可扩展的框架，提供的功能有分类、回归、特征选择、数据导入导出、可视化等。
- Pyopencl：通用高性能GPU并行计算。
- Pytools：著名的Python通用函数、工具包。
- boost_python-1.59：借助Boost.Python库可以将C/C++代码方便、快捷地移植到Python模块中，实现对Python模块的扩充。
- fann2-1.0.7：快速人工神经网络库。
- GDAL-1.11.3：地理空间数据模块库。
- jieba-0.37：结巴中文分词。
- mahotas-1.4.0：图像处理函数库。

- numba-0.22.1：NumPy动态编译器，Python加速神器。
- pgmagick-0.6.0：GraphicsMagick图像库接口API。
- python_Levenshtein-0.12.0：计算字符串的距离和相似之处。
- pythonnet-2.1.0：ms-dot-net4.0接口，运行时（CLR）和.NET应用程序的脚本工具。
- quantLib_Python-1.6.1：金融业量化交易分析模块库。
- regex-2015.11.22：正则表达式。
- reikna-0.6.6：纯Python开发的GPU并行运算库，基于Pyopencl或Pycuda。
- statsmodels-0.6.1：提供统计函数库，需要Pandas支持。
- Theano-0.7.0：Python深度学习库。
- Willow-0.2.1：图像库，Pillow的替代品。

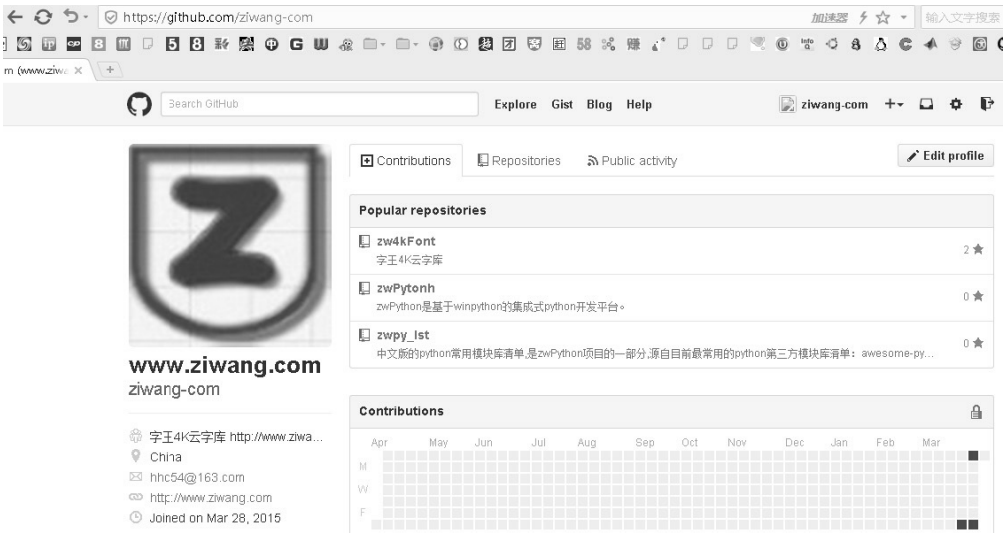
下载与安装

zwPython开发平台是基于WinPython的集成式Python开发平台，比PythonXY更加强大、更加方便，绿色设计，无须安装，解压即可直接使用，可解压到U盘，作为便携式随身开发平台。如果用U盘安装，建议采用8GB以上的内存。

字王项目所有下载集中在百度网盘，网盘地址：http://pan.baidu.com/s/1jIg944u。

还可以访问以下网址，查看项目的最新的进展。

- 网站：http://www.ziwang.com。
- 字王Github项目总览：https://github.com/ziwang-com/，如图A-4所示。



图A-4　字王Github项目总览

zwPython集成式Python开发平台，使用环境：

- 操作系统是Windows 7/Windows 10等64位平台。
- 编程语言是Python 3.x（64位，推荐）、Python 2.7（64位）。
- 内存建议在8GB以上。
- 硬盘，zwPython+zwDat数据包，建议有10GB以上空余空间，如安装到U盘，建议空间在8GB以上。

zwPython开发原本面向电脑美工、设计师，零门槛，无须安装，开箱即用。

为统一开发环境，zwPython需解压到安装盘根目录下（建议为D盘），目录为：d:\zwPython，安装到SSD固态硬盘可加速5~10倍。最新的zwDat、tick数据解压后约150GB，建议使用240GB以上的SSD（压缩模式）。

运行系统

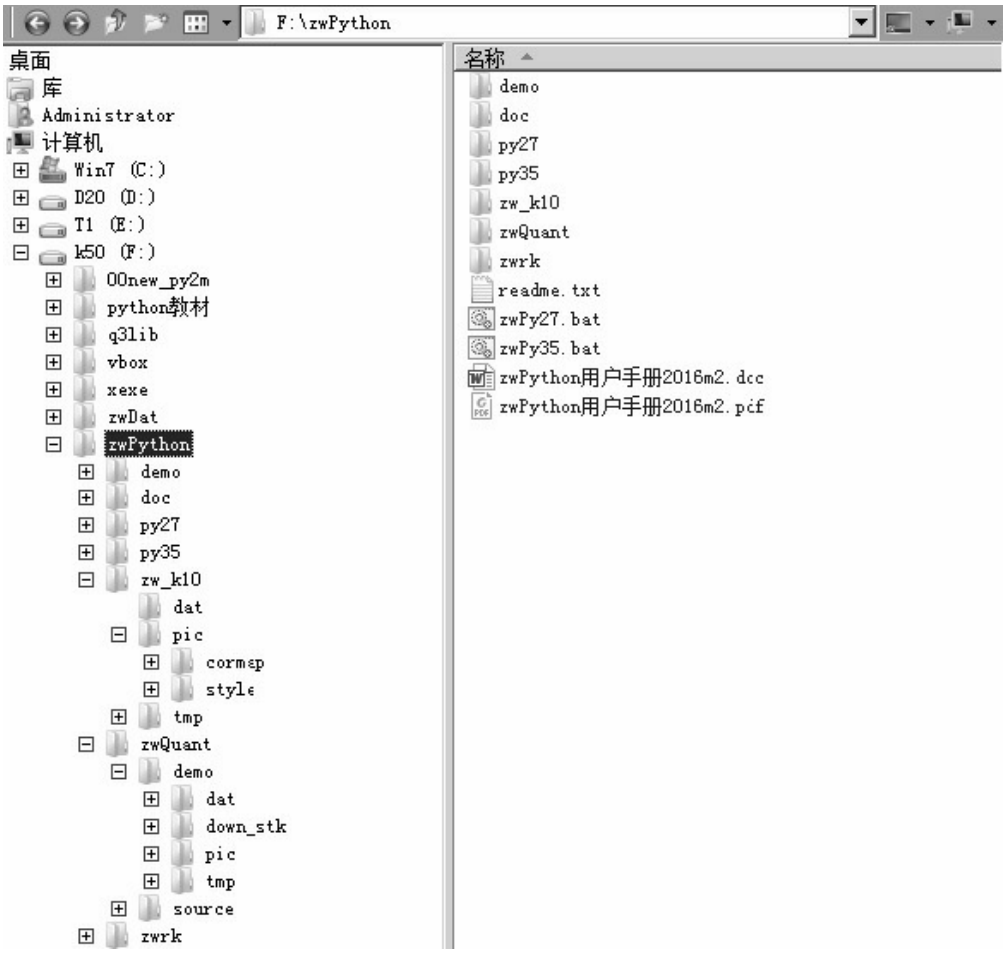
解压后，打开zwPython目录，可以发现文件结构非常简单，只有两个文件和几个子目录。

- zwPy27.bat：Python 2.7版本，运行文件，单击即可进入Python 2.7开发平台。
- zwPy35.bat：Python 3.5版本，运行文件，单击即可进入Python 3.5开发平台。
- Readme.txt：帮助文件。

如果运行zwPy27.bat或zwPy35.bat无法进入Python系统，可能是因为Windows系统缺少VC运行库，请自行到zw网盘下载安装，文件名：vc2015_redist.x64.exe。

目录结构

建议解压zwPython到根目录下，目录结构如图A-5所示。



图A-5　目录结构

zwPython子目录的介绍分别如下。

- \zwPython\doc\：用户文档中心，包括用户手册和部分中文版的模块库资料，如Fontforge、Opencv，以及部分字王字库方面的文档。
- \zwPython\Py27\：Python　2.7版本系统目录，除了增加或删除模块库以外，一般不要改动本目录下的文件，以免出错。
- \zwPython\Py35\：Python　3.5版本系统目录，除了增加或删除模块库以外，一般不要改动本目录下的文件，以免出错。
- \zwPython\demo\：示例脚本源码，包括OpenCV、Pygame游戏设计、Pandas大数据分析、zw字王等相关的示例脚本源码。

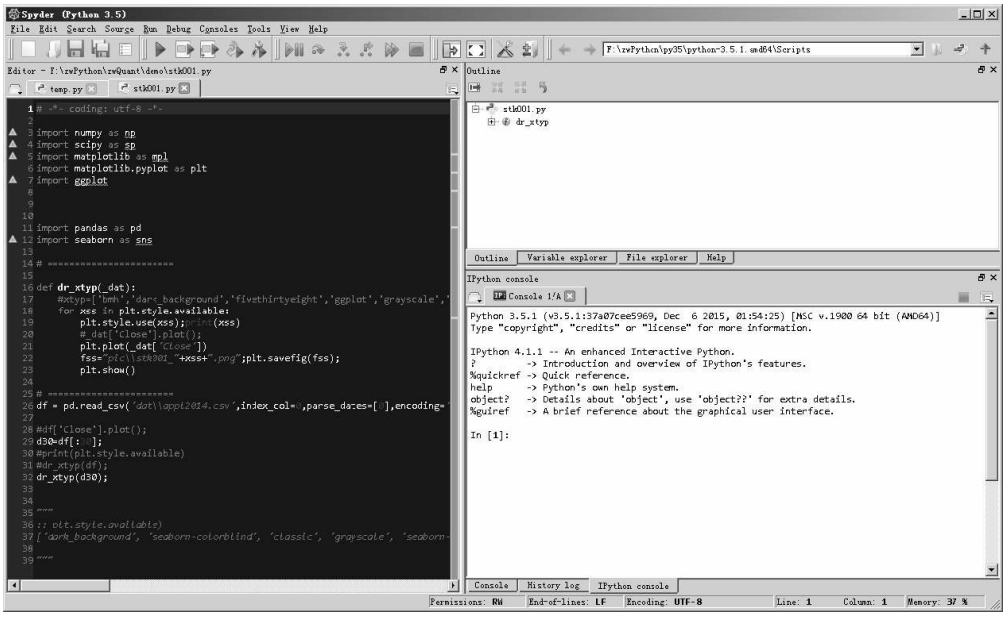
- \zwPython\zwrk\：zw工作目录，用户编写的脚本代码文件，建议放在本目录下。
- \zwPython\zwQuant\：zw开源量化工具箱及zwQuant量化分析开源软件。

zwQuant开源量化软件与zwPython进行了集成处理，可直接使用，支持Python 3.x。移植或使用其他Python环境时，可以把zwQuant目录下的脚本文件全部拷贝到自己的代码工作目录。移植时注意zwSys.py代码中有关的数据文件目录设置。

界面说明

运行zwPy27.bat或zwPy35.bat脚本文件，即可进入Python 2.7、或Python 3.5开发平台，建议使用Python 3.5开发平台，主要是考虑Python 2.7与部分现有项目的兼容问题，所以参考查证后再使用。

zwPython采用的IDE是Spyder，界面类似轻量级MATLAB，进入开发平台后，窗口如图A-6所示。



图A-6　Spyder编辑界面

菜单栏下面是工具栏，常用的工具栏按钮如下。

- “新建”脚本按钮。
- 脚本文件读取、保存按钮，在“新建”按钮右边。
- “运行”按钮，绿色的“▶”按钮。

Spyder界面下方分为以下四部分。

- 左部：代码编辑窗口。
- 右上：参数和工具面板，以及文件和函数导航面板。
- 右下：输出调试窗口。

通常，只要注意编辑窗口和输出窗口即可。关于Spyder程序的更多资料，请参看Spyder、WinPython相关资料和手册。

第一次编程

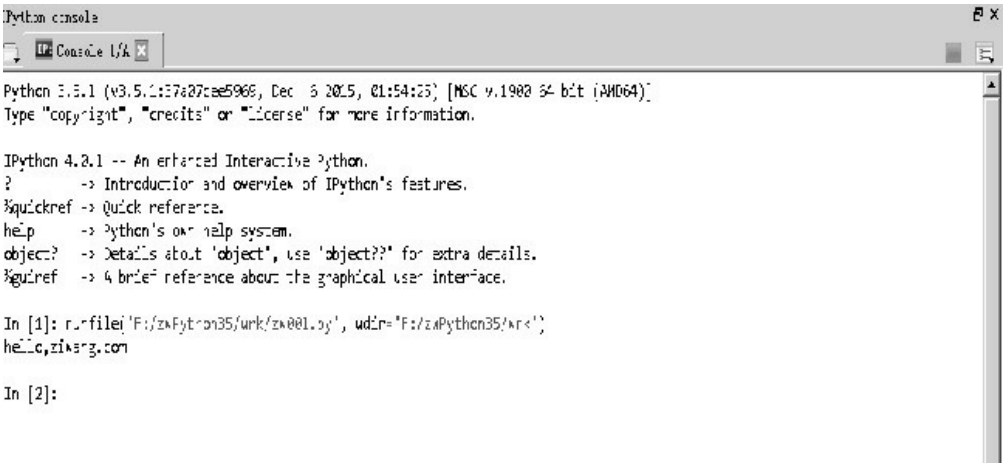
至此软件安装完毕，我们可以开始编写、运行Python脚本程序了。

- 单击工具栏的“读取”按钮，打开≒”zwrk\”目录下的zw001.py脚本文件。
- 单击工具栏中部的绿色“▶”运行按钮。

```
print("hello,ziwang.com")
```

zw001.py很简单，只有一行代码。如图A-7所示，运行后，在右下

角的输出窗口可以看到“hello,ziwang.com”的字样，表示运行成功（注意，输出面板是IPython console）。

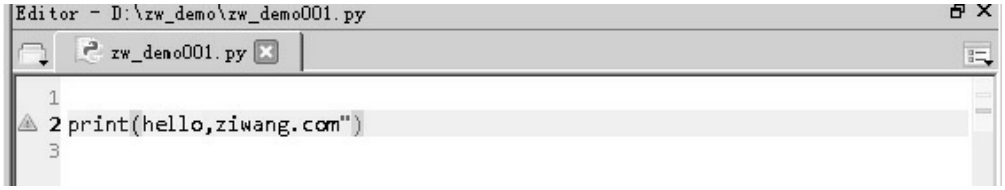


图A-7　程序运行输出窗口

大家可以自己修改引号里面的文字，看看输出效果，注意，此处必须是英文字符和标点。

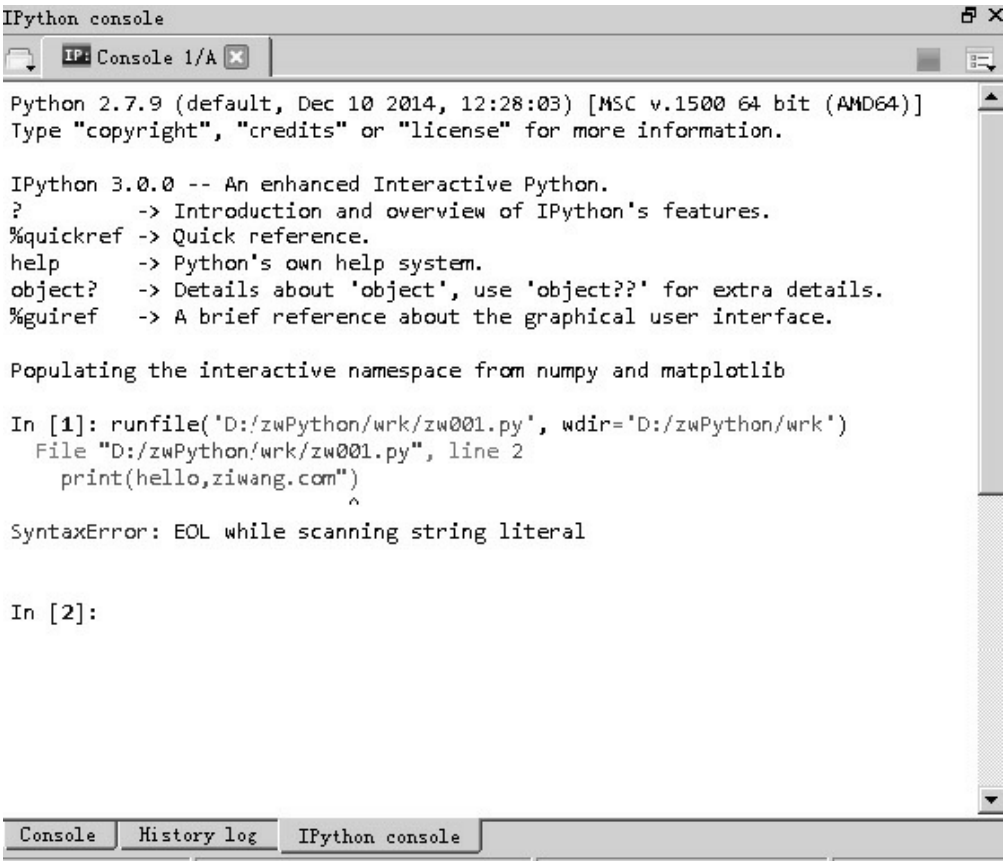
简单调试

下面学习最简单的调试，如图A-8所示。



图A-8　代码编辑窗口

去掉代码左边的引号，再按工具栏上面的“▶”运行按钮，右下角的输出窗口显示如图A-9所示。



图A-9　程序运行输出窗口

图A-9的输出信息表示有错误，注意这行代码：

```
File "D:/zwPython/wrk/zw001.py", line 2
```

其中的“line 2”表示出错的代码，位于第二行。

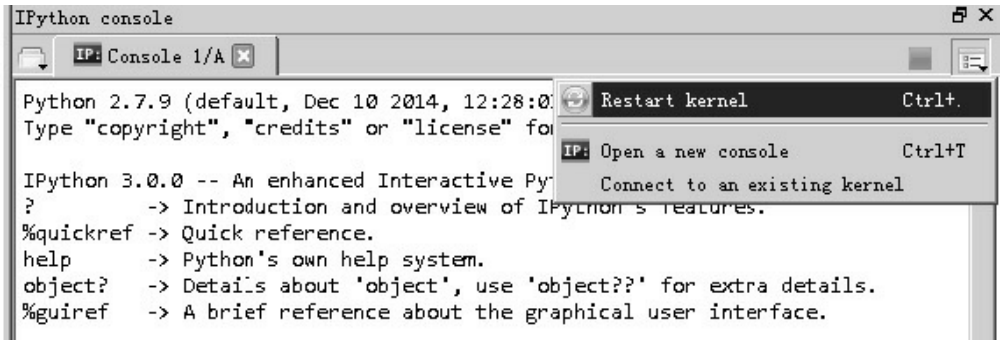
出错信息：

```
SyntaxError: EOL while scanning string literal
```

表示是字符串应用错误，我们加上引号即可。

系统复位

有时，由于脚本代码、或者其他原因，可能引发严重错误，系统运行时出现死循环、崩溃问题，这时，可单击IDE右侧中部的“Restart”下拉菜单和按钮，重新复位即可，如图A-10所示。



图A-10　控制台复位

增强版“hello，zwiang”

下面我们运行一个增强版的“hello，ziwang”：

- 单击工具栏的“读取”按钮，打开“wrk\”目录下的zw002.py脚本文件。
- 单击工具栏中部的绿色“▶”运行按钮。

zw002.py脚本文件也很简单，但功能非常强大，除输出文字“hello，ziwang”以外，还可提供中文输出，以及检测系统多个重量级模块，如OpenCV、Pycuda、OpenCL、Pygame、Pandas等。zw002.py脚本表明，即使是初学者采用Basic的过程模式编写简单代码，也能完成很复杂的功能。

```
# -*- coding: utf-8 -*-
import sys
import cv2
import pandas
import pygame
import tushare as ts

print("hello,zwPython 3.0")
print("字王·4K云字库·系列demo脚本, ziwang.com")
print("python ver:",sys.version)
print("")

print("opencv ver:",cv2.__version__)
print("pandas ver:",pandas.__version__)
print("pygame ver:",pygame.ver)
print("tushare ver:",ts.__version__)
```

以上是zw002.py脚本代码，如果运行无误，输出面板，结果如图A-11所示（不同版本，细节略有差别）。



图A-11　增强版hello，zwiang

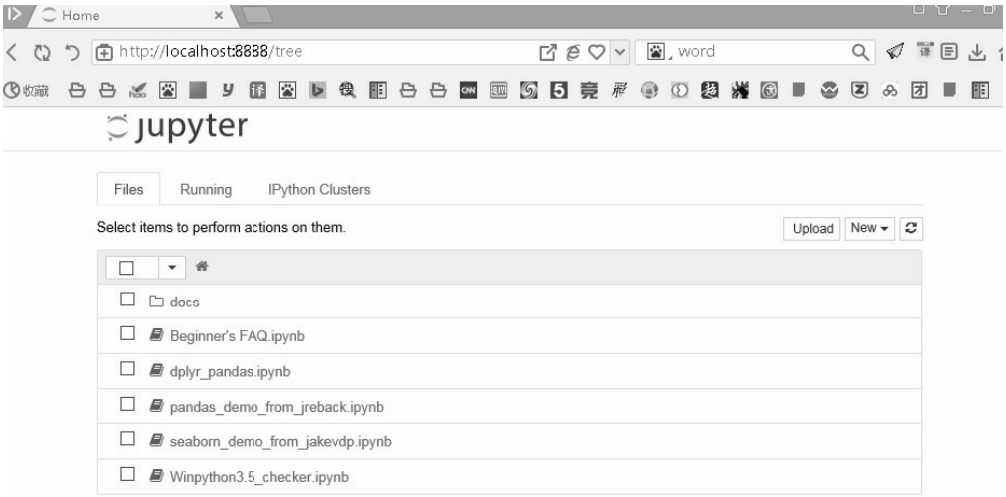
Notebook模式

zwPython内置的Notebook支持模式，事实上目前已经是Web标准模式。Notebook模式文件后缀名是.ipynb，类似IE的MHT网页打包格式，支持文字格式、排版和图像。

运行方法如下：

- 进入Python27或Python35目录。
- 单击运行Jupyter Notebook.exe程序。

Jupyter Notebook.exe程序类似单机的本地Web服务器软件。如图A-12所示，运行后会自动调用默认浏览器，并访问默认网址：http://localhost:8888/tree。

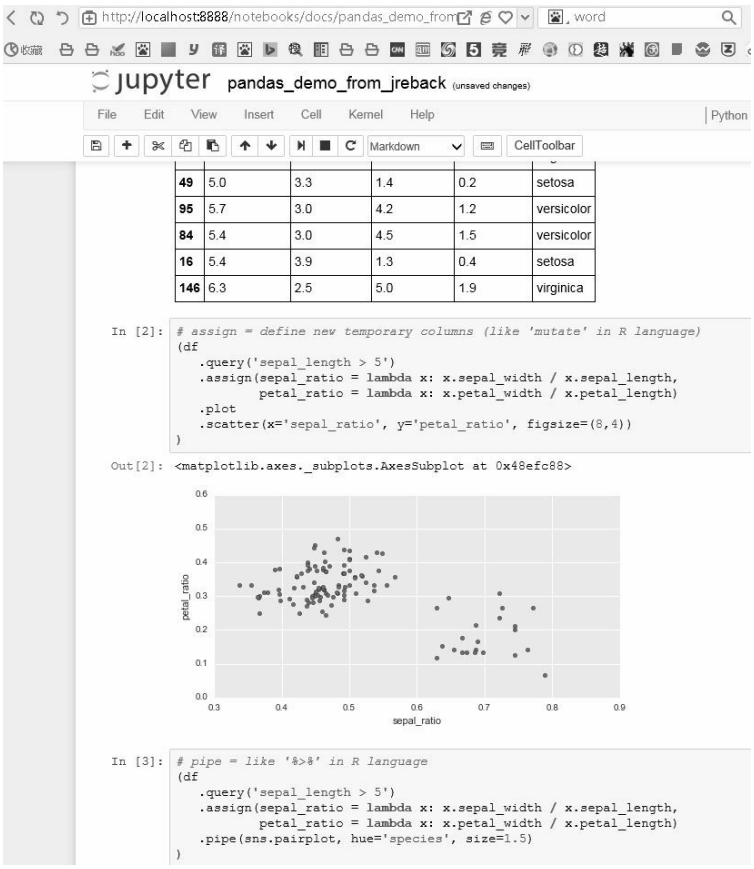


图A-12　Notebook模式

“.ipynb”格式文件使用方法如下。

- 运行Jupyter Notebook.exe程序，进入Notebook模式。
- 单击右上角的“upload”按钮，或者用鼠标直接拖放“.ipynb”格式的文件到浏览器窗口。
- 再单击文件名右侧“upload”按钮，即可上传文件。
- 上传文件后，单击相应的文件名即可看到相应的脚本内容、运行结果和图片。

具体效果如图A-13所示，根据文件内容不同会有所不同。



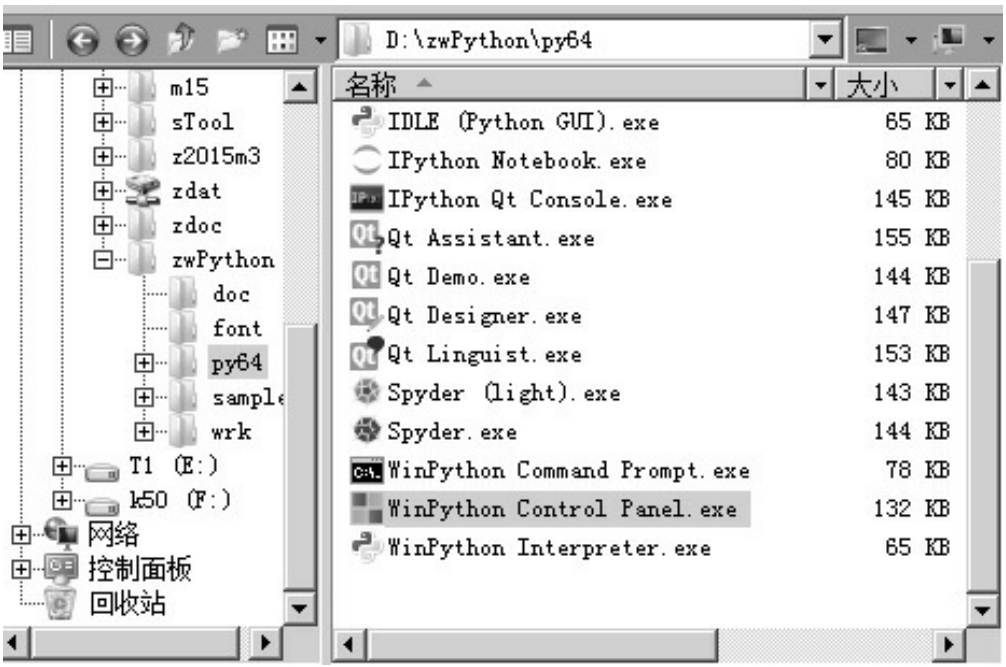
图A-13　Notebook模式运行效果

模块库控制面板

一些简单的模块或Python函数可以直接复制到目录：py64\python-2.7.9.amd64\Lib，zw_lib.py字王函数库就是复制到此目录。Python语言的强大和方便，除了体现在海量的内置模块以外，还体现在绿色、灵活的模块库管理功能。

模块库更新与增删

zwPython的模块库管理，直接使用的是WinPython的模块库控制面板程序：WinPython Control Panel.exe，如图A-14所示。



图A-14　模块库控制面板

模块库控制面板程序：WinPython Control Panel.exe，位于Python 64目录下，运行后的界面如图A-15所示。



图A-15　模块库控制面板启动界面

zwPython模块库的安装流程

zwPython模块库的安装流程如下。

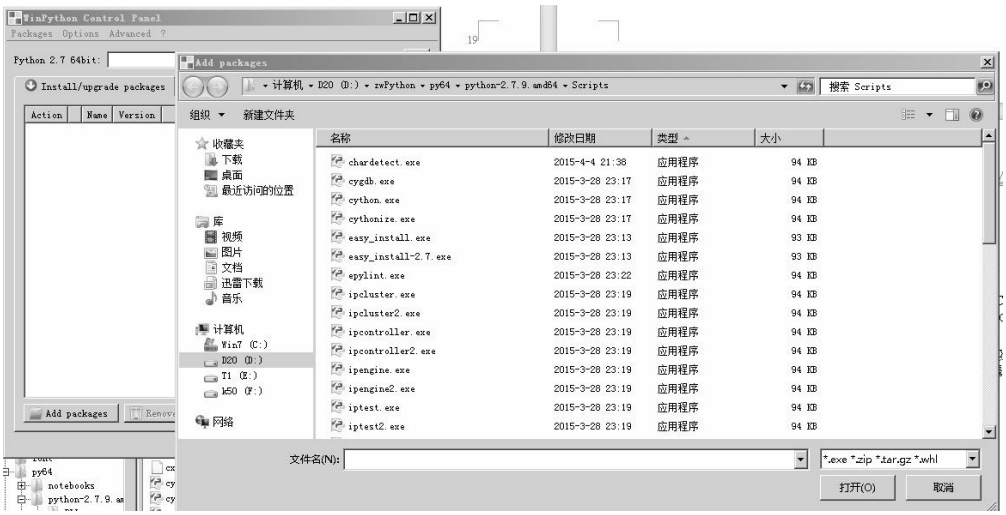
- 把下载的Python模块库复制到任意目录，系统会自动从指定的目录把模块程序复制到系统模块库目录，安装后可以删除模块库源软件。
- 单击右下角的“Add　packages”按钮，即可完成对模块库的添加。一次可选择添加多个模块库文件，如果模块库版本不对，会出现提示对话框，显示相关的出错模块名称；zwPython系统是基于64位Python 2.7版本，因此下载模块请选择对应的版本。
- 添加完毕后，单击右下角的“Install　packages”即可完成模块库的安装。

模块库资源

zwPython模块库资源主要来自以下三个方面。

- 各大网络Python社区：主要是“.zip”、“.gz”格式。
- PyPI（Python　Package　Index）：Python官方模块库，主要是“.zip"、”.gz"格式。
- LFD：欧文加州大学的非官方Python集成模块库，主要是“.exe”、“.whl”格式。

运行控制面板程序py64\WinPython　Control　Panel.exe后，单击右下角的“Add　packages”按钮，如图A-16所示，可以发现，系统支持多种格式的模块库安装：“.zip”、“.gz”、“.exe”、“.whl”。



图A-16　模块库安装

zwPython集成式开发平台在模块库安装方面的强大体现以下几个方面。

- 支持多种格式：除官方的“.zip”、“.gz”格式以外，还支持LFD的“.exe”、“.whl”格式。
- 绿色安装，一次安装，随处运行。

模块库维护更新

控制面板程序py64\WinPython Control Panel.exe还提供了模块库的维护和升级功能，单击菜单Option--Repair packages，如图A-17所示。



图A-17　模块库维护

系统关联

控制面板程序py64\WinPython Control Panel.exe还提供系统关联功能，如图A-18所示。



图A-18　系统关联设置

- 单击菜单Advanced--Register...，即可将zwPython关联到Windows系统，关联后，可以直接在资源浏览器运行“.py”脚本文件，另外增加鼠标右键的“.py”脚本文件与“spyder”IDE程序的关联编辑功能。
- 单击菜单Advanced--Unregister...，即可解除关联。

通常，用户无须采用关联模式。

通常，用户无须采用关联模式。

使用pip更新模块库

有时会出于各种原因，使用控制台安装模块库，并出现失败现象，或者需要批量更新模块库，这时，可以使用pip模块管理程序。

pip常用命令

列出已安装的包：pip freeze or pip list。

导出requirements.txt：pip freeze ><目录>/requirements.txt。

在线安装<安装包>（模块库）：pip install <包名>或pip install -r requirements.txt。

指定版本：通过使用== >= <= ><来指定版本，不写则安装最新版。

requirements.txt内容格式：

```
APScheduler==2.1.2
Django==1.5.4
MySQL-Connector-Python==2.0.1
MySQL-python==1.2.3
PIL==1.1.7
South==1.0.2
django-grappelli==2.6.3
django-pagination==1.0.7
```

安装本地安装包：

```
pip install <目录>/<文件名>或 pip install --use-wheel --no-index -
```

<包名>前有空格，可简写为：

```
pip install --no-index -f=<目录>/ <包名>
```

卸载包：pip uninstall <包名>或 pip uninstall -r requirements.txt。

升级包：pip install -U <包名>。

升级pip：pip install -U pip。

显示包所在的目录：pip show -f <包名>。

搜索包：pip search <搜索关键字>。

查询可升级的包：pip list –o。

下载包而不安装：pip install <包名> -d <目录>或 pip install -d <目录> -r requirements.txt。

打包：pip wheel <包名>。

其他更多的有关pip使用细节，大家可以自行百度。

进入Python命令行模式

Python命令行模式与普通的命令行模式不同，因为集成了Python的运行环境参数，所以，许多新用户使用Windows或其他软件自带的Dos

命令，并运行时会出错。正确的方法是，运行Python目录下的WinPython Command Prompt.exe程序，如图A-19所示。



图A-19　Python命令模式

运行后自动进入Python对应的子目录。

- Python 2.7版本目录是:\zwPython\py27\python-2.7.10.amd64\。
- Python 3.5版本进行了优化，目录是E:\zwPython\py35\scripts\。

pip安装模版

为了方便大家使用pip安装新的模块库，zwPython集成了一个pip01.bat批命令模版，位于不同版本对应的目录下。

pip01.bat批命令，内容是：

```
pip install --upgrade pattern  -i https://pypi.tuna.tsinghua.edu
```

其中，pattern是示例的模块库名称，请大家自行改为需要安装更新的模块库名称。

这个pip01.bat批命令会自动更新指定的模块库，如果找不到对应的模块，会重新安装。

因为Python官网速度很慢，所以，我们在pip01.bat批命令中使用了国内的镜像源，如果出现网络问题，大家可以自行搜索、更换对应的镜

像网站即可。

pip参数解释

pip参数解释如表A-1所示。

表A-1　pip参数解释

install	安装包
uninstall	卸载包
freeze	按着一定格式输出已安装包列表
list	列出已安装包
show	显示包详细信息
search	搜索包，类似 yum 里的 search
wheel	按 requirements.要去创建模块包
unzip	不推荐. Unzip individual packages
bundle	不推荐. Create pybundles
zip	不推荐. Zip individual packages
download	下载模块
hash	计算模块包的 hash 哈希数值
help	当前帮助
General Options	常规选项
-h, --help	显示帮助
-v, --verbose	更多的输出，最多可以使用 3 次

续表

-V, --version	现实版本信息然后退出
-q, --quiet	最少的输出
--log-file <path>	覆盖的方式记录 verbose 错误日志，默认文件：/root/.pip/pip.log
--log <path>	不覆盖记录 verbose 输出的日志
--proxy <proxy>	Specify a proxy in the form [user:passwd@]proxy.server:port
--retries <retries>	重试次数（默认 5 次）
--trusted-host <hostname>	可信任站点，无须 https 站点
--timeout <sec>	连接超时时间（默认 15 秒）
--exists-action <action>	Default action when a path already exists: (s)witch, (i)gnore, (w)ipe, (b)ackup
--cert <path>	证书
--cache-dir <dir>	cache 目录
--isolated	绝对模式，无视 Python 环境和用户设置
--upgrade	如果已安装就升级到最新版

pip-install参数选项

install是最常用的pip参数，install参数选项如表A-2所示。

表A-2 install参数选项

-c, --constraint <file>	约束，使用给定的约束限制版本文件。此选项可多次使用
-e, --editable <path/url>	可编辑的<路径/网址>在“开发模式”下安装一个项目
-r, --requirement <file>	要求<文件>，按给定要求文件，安装模块
-b, --build <dir>	建造模块包
-t, --target <dir>	目标<目录>
-d, --download <dir>	下载到<目录>
--src <dir> SRC	目录查看编辑项目
-U, --upgrade	所有的包升级到最新
--force-reinstall	-升级时强制重新安装，即使他们已经是最新的了
-I, --ignore-installed	忽略已经安装的模块包（与 reinstalling 相反）
--no-deps	不安装依赖包
--install-option <options>	安装选项<选项>，使用 setup.py 的额外参数
--egg	采用 eggs 模式安装，不用默认的“flat”模式
--root <dir>	根目录，安装使用的根目录
--global-option <options>	全局选项
--prefix <dir>	安装前缀目录

续表

--compile	编译 py 文件为 pyc 代码
--no-compile	不编译 py 文件为 pyc 代码
--no-use-wheel	不使用 wheel 模块包
--no-binary <format_control>	不使用二进制模块包
--only-binary <format_control>	不使用源码模块包，只用二进制模块包
--pre	包括预处理和开发版本
--no-clean	不清除 build 创建目录
--require-hashes	使用 hash 验证

附录B
Python量化学习路线图

- 花1周时间学习Python基础知识，先学习zwPython用户手册，可以少走很多弯路。
- 学好Python基础后，建议先通读本书的内容一两遍。通读时碰到问题没关系，记录一下先跳过去；然后进行精读，正式学习，每章的代码，一定要认真学习，并自己运行一两遍，培养编程感觉。
- 多学习zwQuant量化软件的源码，zwQuant源码的中文注解已经到了函数一级，网站也有完善的中文文档。
- 要根据代码学习画流程图，有流程图就可以把握程序逻辑，重点是策略逻辑。

极宽开源量化团队·成立纪念

笔者做项目，喜欢先做整体扫描，在宏观上有个全面的把握。

通过对Python官网和github两个网站进行检索，全球Python Quant量化方面活跃的软件包，基本上都可以找到。

因为很多都是原版的英文资料，第二天，笔者在zwPython量化QQ群（群号：124134140）和大家互动了一下，希望找几个网友一起翻

译。没想到，大家的热情很高，短短两个小时就有20人报名，于是，原本的“极宽开源量化团队”讨论组，就直接升级为了“极宽开源量化团队”（QQ群：533233771）。

同时，也宣布了：“极宽开源量化团队”正式成立。

这一天，是中国量化历史的一个里程碑。因为从这一天开始，一群有着共同理想的年轻人，从全国各地，通过网络聚集到这个小小的“极宽开源量化团队”。

从零开始，逐步向国人翻译介绍世界上最前先进的金融量化技术，同时，开始逐步编写中国人自己的开源量化交易软件。

相信，未来的历史，特别是中国的量化金融历史，会记住这一天：2016年1月18日（腊月初九，乙未年，羊年，己丑月，己亥日）。

Top Quant极宽公司简介

北京极宽科技有限公司，英文名称：Top　Quant，是极宽量化项目的母公司，负责zwQuant量化软件、zwDat金融数据包、zRoboto（机器人）等开源项目的运营和发展。

- zwQuant量化软件新版本名为zQuant极宽量化配套，也是免费的开源软件。
- 商务版，名为zQuantPro，在免费版基础上增加了参数自动寻优，MTrd并发运算等优化模块。
- 新版本zQuant量化软件、zwDat数据包，统一在zQuant.cn网站发

布和维护。

- zRoboto.com网站负责Python机器人和Python芯片项目的维护与发展。