

法律声明

□ 本课件包括：演示文稿，示例，代码，题库，视频和声音等，量邦科技与小象学院拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利。

□ 课程详情请咨询

■ 微信公众号：小象

■ 新浪微博：ChinaHadoop



量化投资组合管理（一）

量化投资与机器学习系列课程

目录

- ☐ 背景知识
- ☐ 因子分类
- ☐ 单因子策略
- ☐ 策略讲解



量化投资组合管理 (一)

背景知识



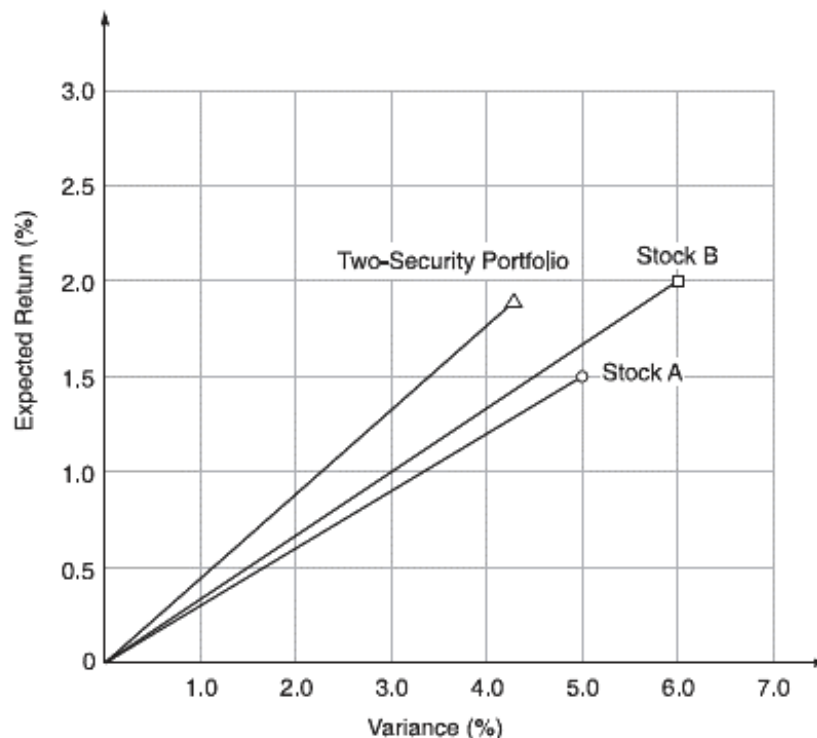
背景知识

□ 收益与风险：收益率与波动率（收益率标准差）

投资组合情形：

$$\begin{aligned} r_P &= \sum_{i=1}^n w_i r_i \\ &= w^T r \text{ (矩阵表达式)} \end{aligned}$$

$$\begin{aligned} \sigma_P^2 &= \sum_{i=1}^n \sum_{j=1}^n w_i w_j \rho_{ij} \sigma_i \sigma_j \\ &= w^T \Sigma w \text{ (矩阵表达式)} \end{aligned}$$

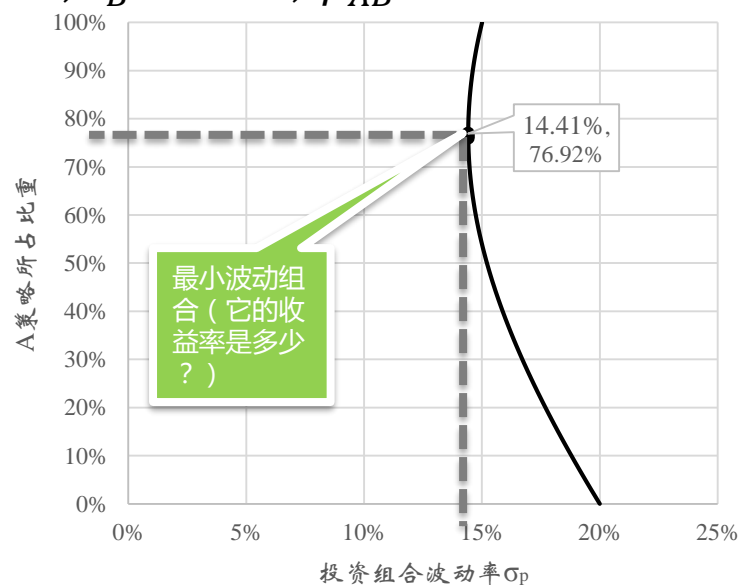
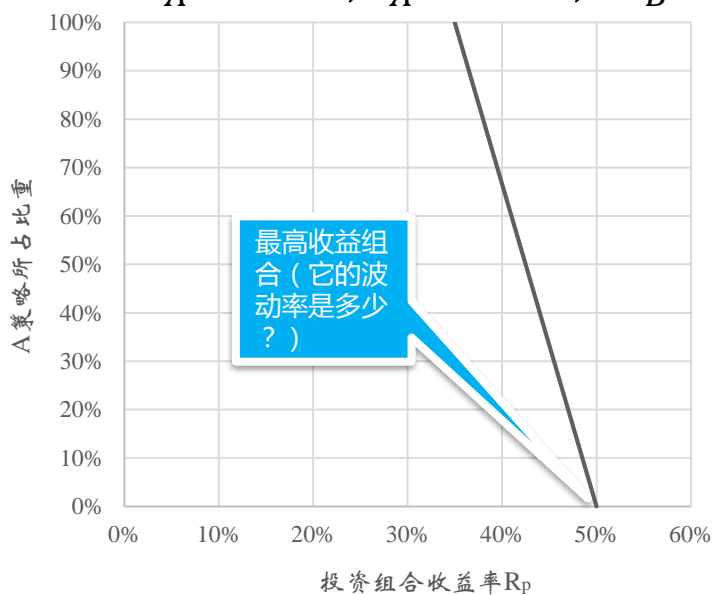


背景知识

□ 一组股票可以构成多少种投资组合？哪个组合最好？

两只股票（A和B）的情形：

$$R_A = 35\%, \sigma_A = 15\%; R_B = 50\%, \sigma_B = 20\%; \rho_{AB} = 0.5$$



背景知识

□ 股票之间的相关性对投资组合有什么影响？

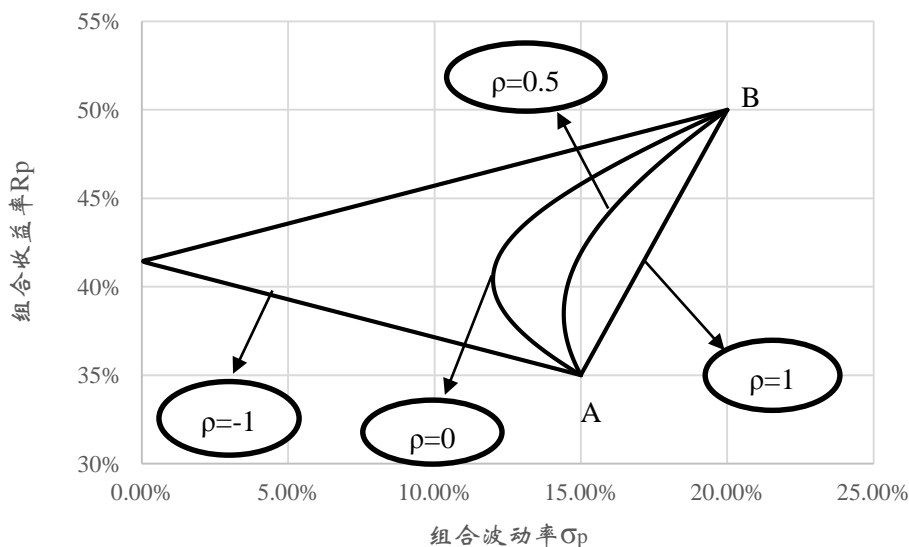
两只股票（A和B）的情形：

$$R_A = 35\%, \sigma_A = 15\%; R_B = 50\%, \sigma_B = 20\%;$$

不同相关系数下投资组合表现

图中哪个组合：

- 收益率最高？
- 波动率最低？
- 最好？



背景知识

□ 均值方差优化——“最优组合”选择方法

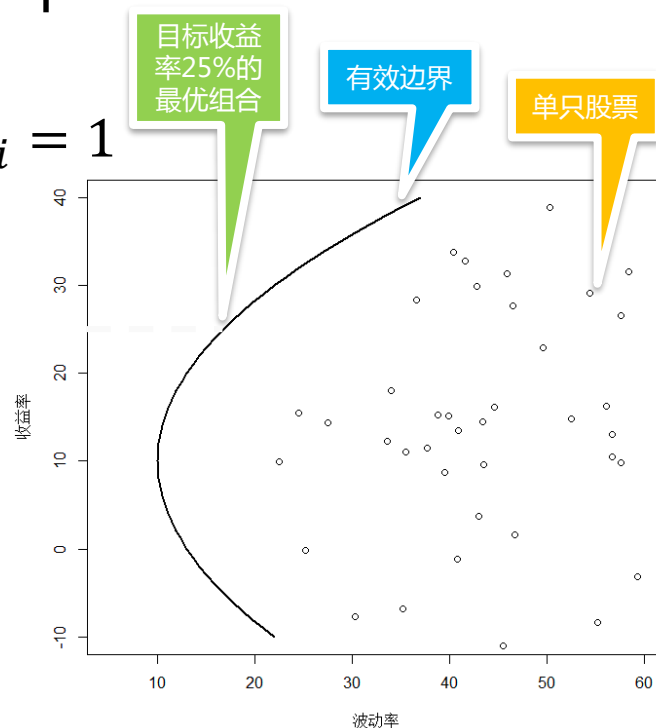
Markowitz Mean-Variance Optimization

优化目标： $\min\{w'\Sigma w\}$

约束条件： $w'r = r_{target}$ and $\sum_{i=1}^n w_i = 1$

优化结果：

- ① 对于每一个目标收益率 (r_{target}) 都可以得到一个最优组合，这些最优组合构成一条“弧线”——称之为“有效边界”；
- ② 无法构造出“位于有效边界之外”（图中在有效边界左边）的投资组合；
- ③ 设定一个目标收益率，有效边界上满足这个收益率的组合即为风险最优组合。



背景知识

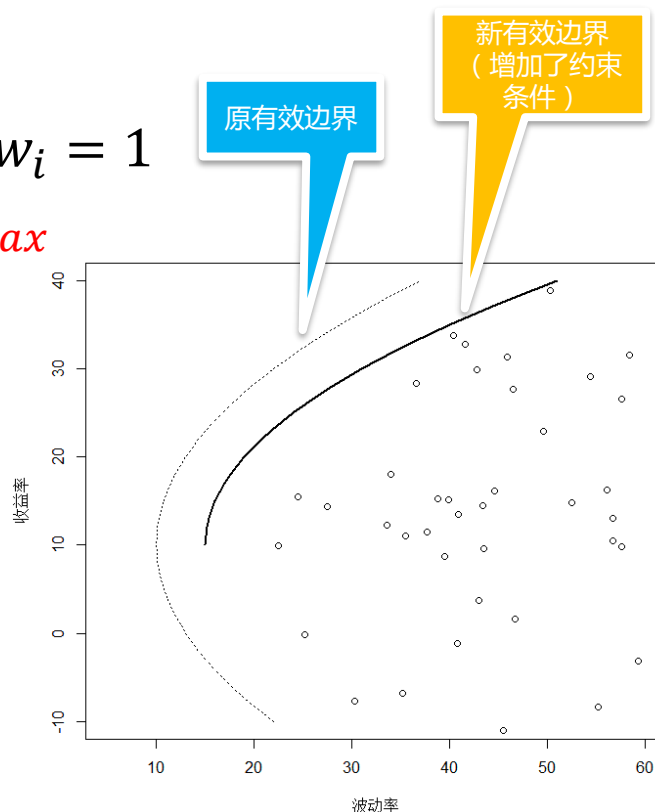
□ 均值方差优化—约束条件的影响

优化目标： $\min\{w'\Sigma w\}$

约束条件： $w'r = r_{target}$ and $\sum_{i=1}^n w_i = 1$
and $c_{min} \leq w_i \leq c_{max}$

结果：

- ① 新有效边界比原约束条件下的有效边界向右移动了，也就是说，能得到的最优组合有较小的收益率（或更大的波动率）。一般来说，约束条件越多，这种情况越明显。那么为什么还要约束？
 - a) 投资手段限制：无法进行卖空交易；
 - b) 标的仓位限制：必须在预定的范围内；
 - c) 参数（ r_i, Σ ）估计的不确定性：不增加约束代表完全信任估计参数，一旦估计不准，优化出来的组合表现有可能非常糟糕。



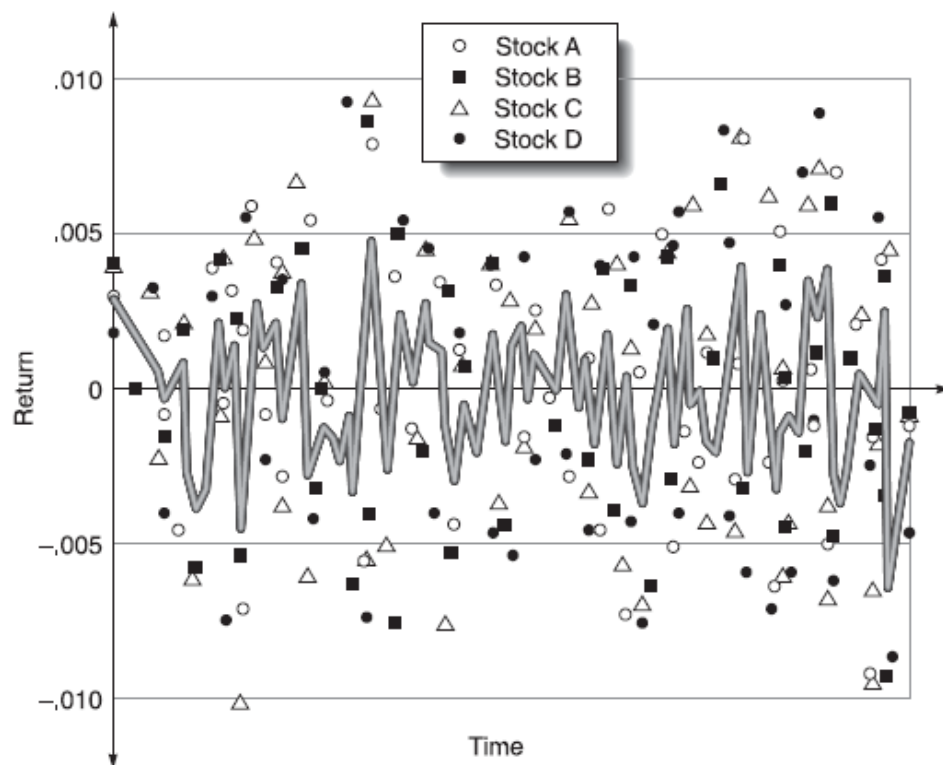
背景知识

□ 分散的效用

通过优化方法构建的有效股票组合可以在保持收益能力的情况下减小风险程度

实例：

右图中4只股票的波动率为16%，由它们构成的组合收益曲线明显更加稳定，波动率仅为5%！

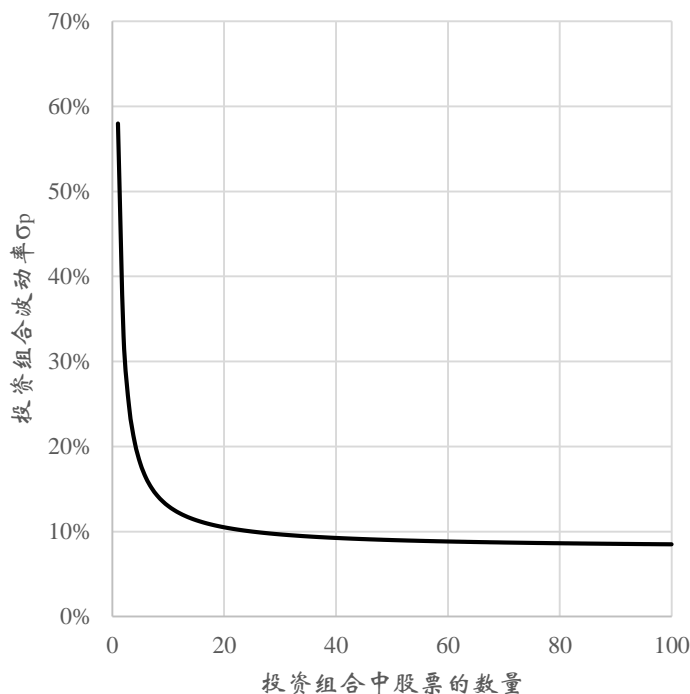


背景知识

□ 分散的局限

Evans-Archer的研究结果：

- ① 随着投资组合中股票数目的增加，组合波动率不断减小，不断的分散的确带来了好处；
- ② 分散的好处在组合中股票数目较少时较为明显，这时增加股票能够得到非常大的风险降低，但组合中股票越多的话这种风险降低效应会越小；
- ③ 当组合中股票数目达到一定数目时，进入“饱和”状态（在图中这种状态大致在20只股票左右），再增加股票所得到的风险降低效应微乎其微。



CAPM模型

□ Capital Asset Pricing Model (CAPM), Sharpe (1964)

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

相对应的股票定价模型为：

$$R_{it} = \alpha_i + \beta_i R_{mt} + e_{it}$$

从中可以看出，

Total Return = Alpha + Beta*Market Return

Total Risk = Specific Risk + Beta*Systematic Risk

结论：

- ① 股票风险有2部分组成：个股风险和市场风险；
- ② BETA反映股票跟市场的关联程度；
- ③ 构建有效组合可以分散风险，但是无法同时消除所有的个股风险和市场风险；

CAPM模型

□ Capital Asset Pricing Model (CAPM)

Alpha和Beta的估算—线性回归

最直观和常用的方法是对股票收益率和市场收益率进行回归，回归参数即为该股票的Alpha和Beta

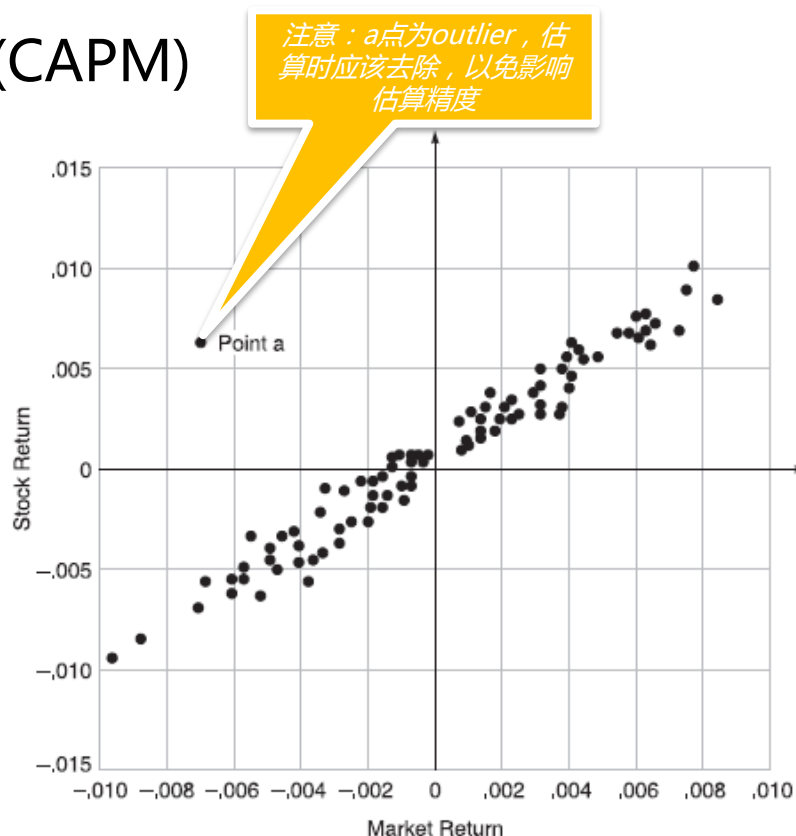
右图的估算结果为：

$$R_{it} = 0.001 + 1.060R_{mt} + e_{it}$$

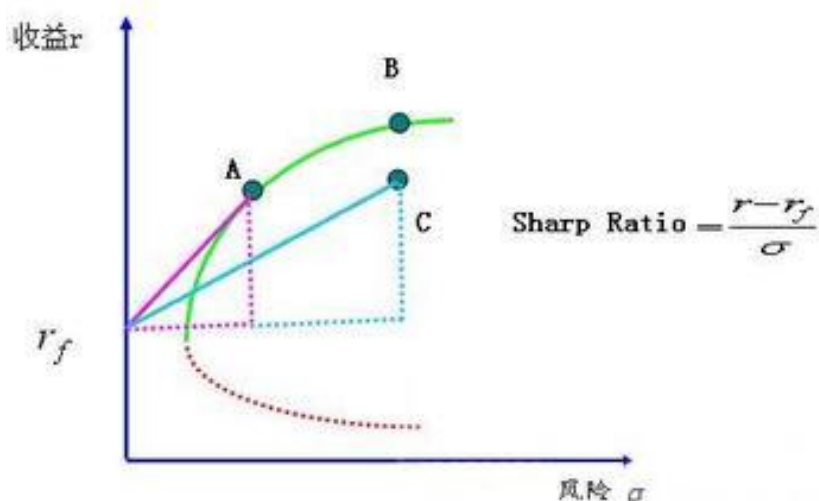
■ Alpha = 0.001

■ Beta = 1.060

市场收益变化1单位，股票收益变化1.06单位。



背景知识



- ❑ 风险预算
- ❑ 夏普比率：收益率除以标准差
- ❑ 无风险收益替换为投资基准收益，夏普比率变成信息比率
- ❑ 定性投资者觉得年化收益率最实在，量化投资者看夏普比率
- ❑ 奥妙在于，如果A策略的收益率低于B，但夏普比率高于B，A策略可以通过杠杆承担和B策略同样大的风险，但收益率会大大提高
- ❑ 夏普比率决定举债能力！

量化投资组合管理 (一)

因子分类

多因子模型

□ 多因子模型 (Multi-Factor Model)

以时间序列回归模型为例：

$$R_{it} = \alpha_i + \beta_{i1}f_{1t} + \cdots + \beta_{iK}f_{Kt} + e_{it}$$

β_{ik} ：因子暴露 (factor exposure)

f_{kt} ：因子溢价 (factor premium)

可以看出，多因子模型在市场因子是唯一因子时等价于CAPM

因子分类

□ 一种常见的因子分类方式：

因子类型	输入	参数估计方法	输出
宏观经济因子	证券收益与宏观经济变量	时间序列回归	证券的因子beta
统计学因子	证券收益	迭代的时间序列或横截面回归	统计学因子和证券的因子beta
基本面因子	证券收益与证券特征	横截面回归	基本面因子

- ① 宏观经济因子示例：GDP增长率、失业率、通货膨胀率、消费指数、商品价格指数等；
- ② 基本面因子示例：公司规模、估值、运营效率、金融风险、流通能力、分红、行业属性等；
- ③ 统计学因子示例：运用主成份分析、因子分析等方法从证券收益的协方差矩阵中提炼因子

因子分类

□ 其他因子分类方式：

① 技术因子：

根据股票行情构造的技术分析类因子（主要包括趋势类、反转类、量能类、情绪类）；

② 分析师因子：

根据分析师对股票的预测结果构造的因子；

③ 事件因子：

根据新闻报道、突发事件构造的因子；

④ 舆情因子：

根据舆情讨论热度监控等构造的因子；

⑤

量化投资组合管理 (一)

单因子策略



单因子策略

□ 单因子策略

之所以能获取Alpha超额收益，是因为市场上各股票的收益存在分化，表现出不同的分布，从而可以通过某种恰当的因子将股票进行划分，赚取股票与股票之间的收益差、或股票与大盘之间的收益差。

其主要步骤如下：

- ① 选取合适的因子将不同的股票加以区分
- ② 研判因子值处于不同区间的股票的收益情况，持有因子值处于某一恰当区间的股票，以获得超越市场平均收益的超额收益
- ③ 在此基础上还可通过做空股指期货、或融券做空因子值处于另一恰当区间的股票，来获取相对更为稳定的alpha收益

单因子策略

□ 以价值因子PB为例

- 众多研究表明，高账面市值比公司的股票倾向于具有较高的收益，而低账面市值比公司的股票则通常具有相对较低的收益；上市公司的账面市值比反映了某种风险特征，部分投资者愿意支付一定的风险溢价来回避相应的风险，由此，价值因子从长期来看，可以获得统计显著的超额收益
- 市净率PB是账面市值比的倒数，这里通过选取低PB的股票来获取超额收益，其主要步骤如下：
 - ① 获取市场上各股票的市值和股东权益合计（即账面价值）；
 - ② 计算各股票的市净率PB；
 - ③ 按PB大小对股票进行排序；
 - ④ 选取PB较小的若干股票构建投资组合；
 - ⑤ （可选：做空股指期货或PB较大的若干股票进行对冲）

单因子策略

□ 常见因子示例

- ① 盈利类因子：销售净利率、净资产收益率、总资产收益率、毛利率、营业费用比例、财务费用比例、息税前利润与营业收入比
- ② 成长类因子：每股净资产增长率、主营业务收入增长率、净利润增长率、总资产增长率、股东权益增长率、经营活动产生的现金流量净额增长率
- ③ 杠杆类因子：负债权益比、资产负债率、长期负债比率、流动负债比
- ④ 流动性因子：1个月平均换手率、3个月平均换手率
- ⑤ 动量因子：1个月股价动量、3个月股价动量、6个月股价动量、12个月股价动量
- ⑥ 规模因子：流通与总市值比、市值自然对数、流通市值自然对数、总资产自然对数、股权集中度
- ⑦ 红利因子：股息率、股息支付率
- ⑧ 现金流因子：经营活动产生的现金流量净额比总市值、经营活动产生的现金流量净额比营业收入、经营活动产生的现金流量净额比营业收入净收益
- ⑨ 相对估值因子：相对PE、相对PB、相对PS、相对PCF
- ⑩ 波动性因子：6个月振幅、12个月振幅、3个月日收益标准差、6个月日收益标准差
- ⑪ 一致预期因子：分析师预测净利润增长、分析师预测主营业务增长



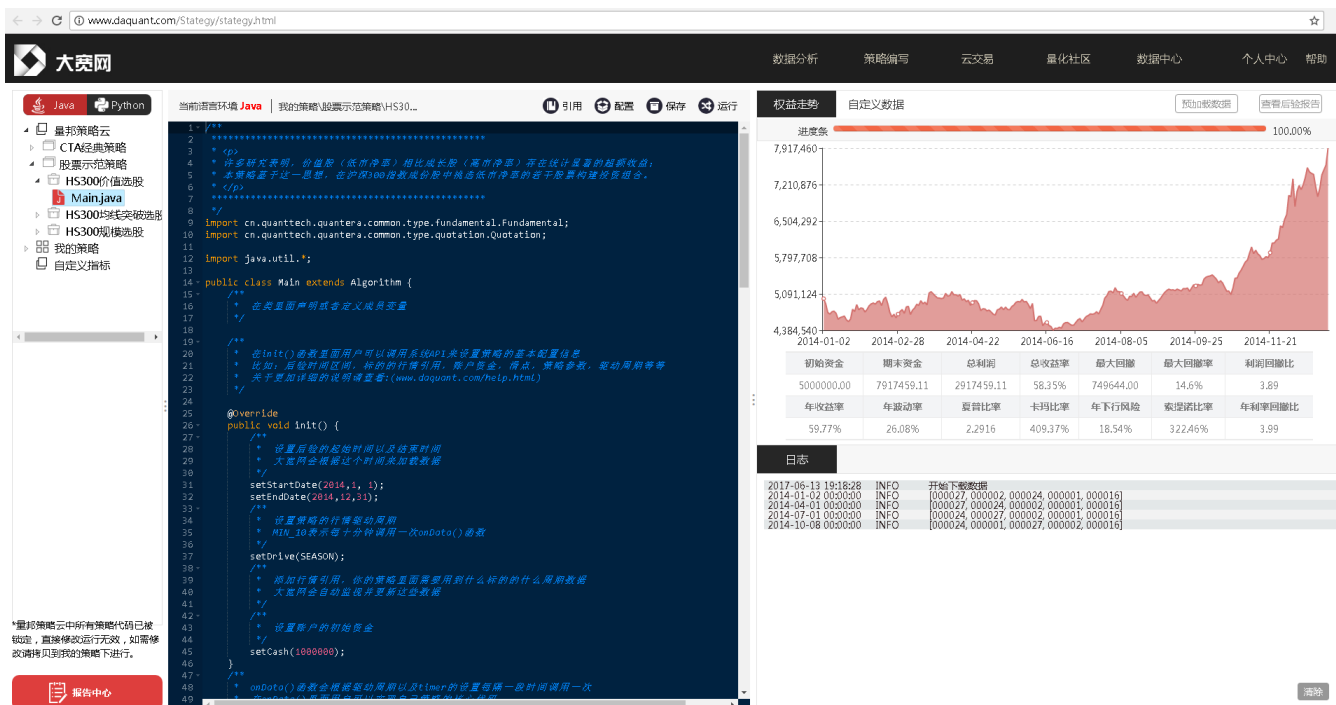
量化投资组合管理 (一)

策略讲解



策略实现

□ 在大宽网实现因子选股策略 支持 Java 与 Python 两种语言

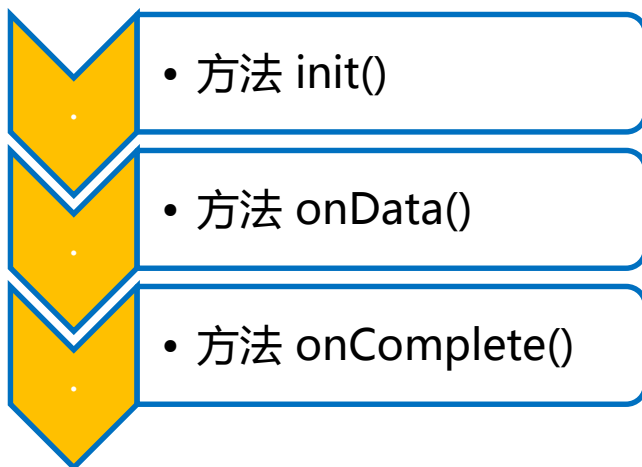


策略实现

□ 在大宽网实现因子选股策略

对大宽网而言，策略就是一个代码结构完整的程序。程序需实现三个方法（函数）：init(), onData() 及 onComplete()。具体的策略思想、交易条件在 onData() 中实现

策略主程序 Main Class



- init() 方法用于处理策略的前置条件：
定义策略后验的开始结束时间;
设定策略标的、驱动周期、初始资金、滑点;
- onData() 方法是整个策略中的重点，
onData 方法在行情数据到来时调用，并处理内部逻辑。在 onData 方法里需要实现交易思想，开平仓规则。
- onComplete() 方法在程序运行完成时调用，不作数据处理。

策略实现

□ 在大宽网实现PB价值因子选股策略

```
11 class Main(Algorithm):
12     # 在类里面声明或者定义成员变量
13
14     def init(self):
15         # 设置后验的起始时间以及结束时间
16         # 大宽网会根据这个时间来加载数据
17
18         self.setStartDate(2014, 1, 1)
19         self.setEndDate(2014, 12, 31)
20
21         # 设置策略的行情驱动周期
22         # MIN_10表示每十分钟调用一次onData()函数
23
24         self.setDrive(TimeFrame.SEASON)
25
26         # 设置账户的初始资金
27         # self.setCash(10000)
28
```

init() 方法

设置策略后验起始与结束时间

设置策略的驱动周期，即隔多长的周期驱动一次 onData() 方法，例如这里是每季度驱动一次

设置策略的初始资金，股票组合策略默认初始资金500万，无需设置



策略实现

```
29 # onData() 函数会根据驱动周期以及timer的设置每隔一段时间调用一次
30 # 在onData()里面用户可以实现自己策略的核心代码
31 def onData(self):
32
33     # access HS300 components stock list
34     HS300 = self.getIndexStocks("000300")[0:9]
35
36     # access quotation data and financial report data of HS300 components
37     # calculate corresponding book-to-market ratios and store them into a dictionary
38     bookToMarket = {}
39     for iterCode in HS300:
40         # self.info(iterCode)
41         iterClosePrice = self.PickQuoteData(iterCode, TimeFrame.DAY, self.time().toLocalDate()).price()
42         # self.info(iterClosePrice)
43         iterFundamentals = self.getFundamentals( iterCode, Category.BALANCE, self.time().toLocalDate().plusDays(-180), self.time().toLocalDate() )
44         if iterFundamentals:
45             iterTotalStock = (iterFundamentals[0]).totalStock
46             # self.info(iterTotalStock)
47             iterTotalOwnerEquity = (iterFundamentals[0]).totalOwnerEquity
48             # self.info(iterTotalOwnerEquity)
49             iterBM = iterTotalOwnerEquity / iterClosePrice / iterTotalStock
50             # self.info(iterBM)
51             bookToMarket.update({iterCode:iterBM})
52         else:
53             self.info("Fundamental data of " + str(iterCode) + " is empty.")
54     # self.info(bookToMarket)
55
56     # sort book-to-market ratio in descending order
57     sortedBM = sorted(bookToMarket.items(), key = lambda d:d[1], reverse = True)
58     self.info(sortedBM)
59
60     # select the 5 highest book-to-market ratio stocks and stock them into a list
61     selectedStockPool = []
62     for iterSortedResult in sortedBM:
63         selectedStockPool.append(iterSortedResult[0])
64         if len(selectedStockPool) >= 5:
65             break
66     # self.info(selectedStockPool)
67
68     # set equal portfolio weight for selectedStockPool
69     self.lazyRecordWeightBook(self.createEqualWeight(selectedStockPool),self.time().toLocalDate())
```

onData() 方法

获取沪深300成份股

获取沪深300成份股的收盘价、股本、股东权益合计等数据，并计算相应的账面市值比

按账面市值比对股票进行排序，筛选账面市值比最高的若干股票

给筛选出的股票设置权重，构建股票组合，这里设置了等权重

策略实现

□ 在大宽网实现因子选股策略

```
77
78 # onComplete() 在策略运行结束的时候调用
79 def onComplete(self):
80
81     self.lazyEvaluate()
82     self.info('策略运行完毕')
83
```

onComplete() 方法

lazyEvaluate() 方法根据 onData() 方法中的 lazyRecordWeightBook() 方法所生成的股票权重进行相应的股票买入卖出操作

策略实现

□ 在大宽网实现因子选股策略

策略后验及报告



疑问

□ 问题答疑：<http://www.xxwenda.com/>

■ 可邀请老师或者其他人回答问题

□ 量化投资策略研发平台：<http://www.daquant.com>

相关微信号推荐



量化投资华山之巅



量化投资大家学



联系我们

小象学院：互联网新技术在线教育领航者

- 微信公众号：大数据分析挖掘
- 新浪微博：ChinaHadoop

