



投资全球更要投资自己



我的订阅

上一篇:

FOF与MOM, 傻傻  
分不清. 第20讲  
0

下一篇: ★

验血报告: 学会从风  
险的角度分析收益.  
第22讲



## 股票风险模型：从单因子到多因子. 第21讲



渔阳 2018-05-23 19:34

字数 4,798 阅读需 12分钟

“  
欢迎来到量化小学”

▲ 加入[“量化小学”校友圈儿](#)提问交流

来自特辑



量化小学

解放你的投资动手能力

最近更新

【学业总结】量化学习的脉络梳理，以及  
继续学习提高的路径

2019-04-12更新

进阶研究：集成学习和深度学习. 第31讲

2019-03-28更新

详细内容请在wifi环境下观看视频

<本期课程5008字，视频20分钟，请合理安排学习时间>



## 内容阅读

大家好，欢迎来到量化小学，今天是第21集，将开始新的一章《再探风险》。这一章我们谈一谈股票风险模型的两部分主要内容。一是从单因子到多因子模型。第二简单介绍一下商用的风险模型。

在本章的开始先给大家讲一讲，为什么风险模型这么重要？它其实是专业选手和业余选手的重要的分水岭！说到挑股票，大家都会有自己的观点，但风险模型玩的非常熟的、理解得非常透彻的，一定是专业选手。

因此，我建议大家多花一些时间弄清楚风险模型是怎么回事。学会解读风险报告，无论是做股票研究，甚至对找工作都会很有帮助，所以这章是量化小学重要的干货部分。

## 风险模型的直观理解

风险模型，先直观地理解它在讲什么，是前面讲过的一些内容的延续，先复习一下。第二章中讲过任何一只股票在未来一段时间的收益率都是一个概率分布。经典的理论认为是正态分布。

那么正态分布就有两个参数，一个是 $\mu$ ，一个是 $\sigma$ 。 $\mu$ 是它的预期收益率， $\sigma$ 是一种不确定性的度量，也就是所谓的风险。

一只股票它有 $\sigma$ 总体的不确定性，那么直观的想一想，影响到股票价格的涨跌会有很多因素，因此总体的不确定性一定是可以分解成为很多小的不确定性的，可以理解为一些小的 $\sigma$ 的叠加。这个数学公式如果有很多的不确定性，它们之间又有一个相关性矩阵，可以用一个线性代数把它们组合在一起，就能够算出股票的方差，是一个基本的公式。前面在讲相关性的时候也提过，大家可以复习一下。



数学并不是我们最关心的内容，我们最想知道的是这些 $\sigma$ 的金融含义是什么。这基本上是我们讲的那些风险因子了。这样的不确定性有两大类，第一类是所谓的共性因素，比如整个市场的涨跌，会影响到所有的股票，某一个行业的表现会影响到这个行业里所有的股票。比如医药股票大涨，那么这些行业中的医药股都会受到正面的影响，这是很显然的。

除此之外还有一些其它的因子，比如大家如果熟悉A股的话都知道有的时候大票表现好，有的时候小票表现好，这种相对表现也可以说是一个共性因素。这些共性因素就统称为风险因子。另外一些不确定性是来自股票自身的特殊的性质。比如这个公司产品怎么样，市场营销做的好不好，管理层的能力如何，等等。这些只影响到这只股票，因此称之为特异风险。

在这里要特别强调一下，在投资当中有一个盈亏同源的概念。在经典的投资理论中也是这样，风险是中性的，而不是一个负面的词汇。可以复习一下前面讲过的内容，有风险才有收益，这是最基本的金融理论。

## 基础模型：CAPM

接下来可以看到，怎么度量风险，到底是什么样的因子，从最基础的模型开始讲起。任何一本经典的金融教科书里面都会有，而且出现在比较靠前的章节。因为CAPM (Capital Market Pricing Model) 是1960年代就出来的经典的模型。

它跟前面讲到的有效市场假说,上一节讲到的现代组合理论都是有关系的。它的基本思想是什么,首先它只有一个风险因子，就是**市场系统性风险**，英文叫systemic risk。什么叫系统性风险，就是没有办法对冲，因为它是整个市场的涨跌，是没有办法消除掉的。

因此在金融的理论中，既然没有办法消除掉这个风险，那必然要有一定的预期收益，才会去承受这个风险，因此系统性风险也是预期收益的来源。除了系统性风险之外，剩下的就



是股票的特异风险，这是通过分散投资几乎可以完全消除的。前面讲过N个不相关的随机数，把它放在一起，它的风险是跟着根号N扩大的，所以股票足够多的时候，这风险就没有了。既然它能够被消除掉，也就不带来预期收益。

我们看一看数学公式是什么样子的？

$$r = r_f + \beta(r_m - r_f) + \epsilon$$



一只股票的预期收益率等于第一部分是无风险收益率。在这个基础上，有一个 $r_m$ 是市场的收益率。实践中一般都用宽基的指数例如S&P500或者沪深300来代替。市场指数对于无风险利率的超额收益，就是系统性风险的因子。 $\beta$ 是线性回归的系数，就是这只股票对风险因子的暴露程度。 $\beta$ 越高，也就意味着市场涨的时候，它涨的越多，反之亦然。

最后一项是 $\epsilon$ ，也就是股票的特异风险，它也符合正态分布，也称之为残差收益率。可以注意到正态分布它的均值是零，为什么是零？刚才讲过股票的特异风险，可以被充分投资，在diversification分散掉了，买的足够多了，风险就没了。因此在经典的金融理论中，这样的东西是不会带来预期收益的。

因此CAPM也是有效市场派的逻辑链，就是股票的预期收益率仅尤其对市场的风险暴露有关，残差收益率没有带来预期收益。第二理性投资者只应该投资无风险资产和市场组合，因为投资个券，残差没有预期收益率，为什么要冒这个风险呢，因此只应该是部分的投资无风险资产，部分投资市场组合。前面其实是讲过的，大家可以复习第17讲现代组合理论的内容，应该是叫做Capital Market Line资本市场线。

## Fama—French三因子模型



说到会有一个问题，就是为什么只有市场风险这么一个因子了，是不是真的没有其它因子了？显然不是，金融学家就继续研究，到1990年代就有一个著名的Fama—French三因子模型就出来了。Fama应该是一个比较熟悉的名字，在讲有效市场假说的时候，提到这位芝加哥大学著名的教授，他提出的有效市场假说，得了诺贝尔奖，之后就找到了超额收益的因子。

Fama—French三因子模型它的基本思想是什么？就是除了市场风险之外，还有两个因子也能够解释股票的回报，他们是风险因子，因为有风险，所以也能够带来预期收益。

这两个风险因子是什么？第一个就是市值，一般的来说，小市值股票对大市值股票是有超额收益的。第二个是估值，估值低的股票相对于估值高的股票是有超额收益的，就是你应该去买比较便宜的股票。

跟刚才的模型一样，除了三个因子，剩下的是股票的特异风险，能够通过分散投资消除，也不带来预期收益。所以Fama—French三因子模型，数学公式中它的第一项跟刚才一样，还是对整个市场的超额收益有一个 $\beta$ 的暴露，剩下两项，第一个是SMB,这是度量小市值股票对大市值股票的超额收益的风险因子。

$$r = r_f + \beta(r_m - r_f) + b_s SMB + b_v HML + \epsilon$$

02

另外一个HML是度量价值股对成长股超额收益的风险因子。这两个都是从历史数据来计算的。具体的大家可以去看这个相关的资料，在量化小学里我们重点讲一讲怎么来理解这些模型。



$B_s$ 、 $B_v$ 就是某一只股票对这两个风险因子的暴露程度，它是通过线性回归来估计的参数。举一个直观的例子，如果是一只市值很大的股票的话，它对SMB的风险 $b$ 暴露应该是负的，因为它是大股票，最后 $E$ 一样还是股票的特异风险。

Fama—French三因子模型的解释力就更强了，多了两个因子，传统上来讲是可以带来超额收益的，但是必须指出这种超额收益都是很不稳定的。它们的夏普比率都是很低的。

## 多因子风险模型

那大家又问三个因子难道就只有这三个，没有其他因子了吗？显然不是。所以从三因子模型非常自然地就可以扩充到多因子模型。

$$r_i = r_f + \sum_{j=1}^n \beta_{i,j} F_j + \epsilon$$

03

多因子风险模型的数学公式看起来更复杂一点，其实跟刚才意思是一样的。把三个因子扩展到 $N$ 个，相当于每一只股票，第 $i$ 只股票都对这 $N$ 个因子有一个风险暴露，就是 $\beta_{i,j}$ 。

怎么来刻画股票的预期收益率，还是刚才那几部分，第一部分是无风险收益率，第二部分是它对这 $N$ 个因子，每个因子的风险暴露和这些因子本身的回报加起来，最后还有一个 $E$ ，就是这个风险模型不能解释的残差的部分。

这是数学公式，重在理解数学公式在讲些什么，这些因子都有什么。它主要分为两大类，第一类是行业因子，这跟刚才说的市场因子是非常类似的。它涉及到A股，各家的划分是有所不同的，但差不多都是30个左右，比如金融、医药、钢铁、电子等等都是行业。



很显然每一个行业里的股票更倾向于同涨同跌，都是做类似生意的。第二点绝大部分股票只有一个行业有暴露，也就是它在这30个左右的行业，只有一个 $\beta$ 等于1，剩下的都是0，因为绝大部分股票都只属于一个行业。美国有个别的股票，属于两个甚至三个行业，这是一个个例。A股中基本上应该都是一个行业，这是第一块市场因子。

第二块刚才讲过，小票是有超额收益的，低价股是有超额收益的，这类的因子统称为风格因子。除了市直和价值之外，还有一些其他的東西，比如成长、动量、流动性等等。这些风格因子不是每一个都一定能带来超额收益，但是它们自身是会有波动性的，这种波动一定程度上能解释股票收益率的波动，所以它们就是风格因子，是风险因子的一部分。

## 商用多因子模型

一般来说有十几个风格因子，在A股市场影响力最大的是市值因子，在接下来看风险报告的时候也会有更加直观的体会。多因子模型里边的参数很多，显然不是每个人都有能力或者时间自己计算这些参数。因此就有一些商用的第三方的数据提供商来提供这些数据。

大家的想法都比较类似，都是多因子的框架，基于基本面的，当然具体因子的定义会有所不同。比较大的数据厂商，最大的显然是Barra，几十年了一直在做风险模型这块，像Axioma是Barra的竞争者。它卖的数据分为几个部分，第一它会告诉你每一个因子在每一个周期的回报率，第二会告诉你因子之间的协方差矩阵，也就是这些因子之间是什么关系。大家可以复习一下上一章中度量相关性的部分。

第三块会告诉你每一只个股的因子暴露，就是前面讲到 $\beta$ 是多少。最后它会对个股的特异风险有一个数据上度量，所有的这些数据都是通过回归的方法生成的，道理上不难，但实践中有很多坑，这也是人家为什么能够卖钱的原因。因为它花了很多时间，数据质量高，把工作都做了，最后就可以卖钱，这数据几十万一年。



我们关心的还是它到底在讲什么？这里就简单的列了几个最重要的风格因子，看一下是怎么定义的。

风格因子定义举例		
	Barra CNE5	Axioma AXCN
市值 ( Size )	市值自然对数	市值自然对数
价值 ( Value , Book-to-Price )	市净率的倒数	市净率的倒数
成长 ( Growth )	$0.47*SGRO+0.24*EGRO+0.18*EGIBS+0.11*EGIBS_s$	Plowback * ROE
动量 ( Momentum )	504个交易日数据计算的RSI	过去12个月除掉过去1个月的收益率

首先是市值因子，也是刚才Fama—French三因子模型中的一个因子， Barra和Axioma两家都是用市值的自然对数来定义的，为什么要取一个自然对数，是因为股票的市值彼此之间相差很大。

以A股为例，有几十亿市值的，也有几千亿市值的，相差了上百倍，数学上不是正态分布所以比较难处理。那取一个自然对数之后，就更接近于正态分布，更便于线性模型进行处理，两家的定义是一样的。第二是价值，通常是用Book-to-Price表示，就是一个公司的Book Value除以它的市场股价，中文就是市净率的倒数。

很显然越便宜的公司，它的净值越大，市值价格是越低的，也就是Book-to-Price是越大的，这两家的定义也是一样的。第三个涉及到成长，这个就出现不一样了，这个公式挺复杂的，初学的时候不用特别的想这些细节。比如这个定义里SGRO、EGRO无非是它销售额的增长率、净利润的增长率等等。

它按照一定的模式给组合出来，造了一个它觉得有用的因子，可以度量一公司的成长性。Axioma类似，它把几个不同的参数组合在一起来度量成长。动量也一样， Barra用500多天





的数据计算的一个相对强弱指标，Axioma比较直观，就是拿过去12个月的收益，扣掉过去这个月的收益，这是因子的定义。

其实做为用户首要的关注点并不是到底里面的参数是怎么样的，而是如何解读这些数据生成的风险报告。下一次就具体的看一看，怎么样解读风险报告。

## 课后思考题

这次的内容，可能有的同学觉得有些抽象，为了便于大家理解，出一个课后思考题。刚才讲了一些风格因子，具体到两只股票，中国平安和招商银行都是著名的白马股，大家可以思考一下，他们对下面这几个重要的风格因子，暴露应该是正的还是负的，为什么？

我在这列了四个，第一个是市值因子（Size），第二个是价值因子（Value），然后是Growth和Momentum。

### 思考题

中国平安和招商银行两只股票，对下列风格因子的暴露是正的、还是负的，为什么？

- Size
- Value
- Growth
- Momentum



提示：

- 多因子框架下，对某个风格因子的暴露需要“全局看问题”，考虑到其他因子特别是行业因子的影响
- 可以用万得、东方财富、同花顺等金融终端查阅市净率（PB）、市盈率（PE）等相关信息
- 也可以用quant<OS>写程序直接查，还可以做同行业对比。量化研究的标准方法



给大家一些提示，首先在多因子框架当中，某个风格因子的暴露要全局的看问题。因为对这么多因子的风险暴露是用多元线性回归算出来的。很多时候可以理解为我是控制了其它因子影响的前提下，对这个风险因子的暴露是什么样的？比如要考虑在行业内部究竟处于什么样的位置，很多时候就决定了你对因子的风险暴露。

第二从什么地方得到这个信息。其实就是两个主要的途径。可以直接看，从万得终端、东方财富、同花顺等等都可以直接查PB、PE等信息。根据这样的PB、PE你觉得它的成长性、价值性怎么样等等。当然做量化通常是会写程序直接取数据的，quant<OS>可以做这个事情,这是标准的方法。

其实更加的好用可以简单的就做同行业对比，看中国平安、招商银行在他们行业内部处于一个什么样的位置。刚才讲过有些风险暴露得看相对的。给了不少提示了，希望大家先思考一下。

这次数学公式讲得比较多，下一次就讲一些非常具体的风险报告，通过股票具体的组合，让大家对风险报告有一个直观的认识。所以下一期的题目就叫验血报告。就好像当一个医生，看了有几十个指标的验血报告，用这个报告来判断投资组合它的风险是什么样子的。

## 学习资料

最后是一些学习资料，首先还是quant<OS>金融终端的教学频道有样例代码，主要教大家怎样去取PB、PE进行行业内的比较。（[戳链接进入](#)）

第二和第三建议大家去看一看风险模型的手册（[见下方](#)），包括一些其他的相关的材料。前面讲过风险模型是区分专业和业余的重要标志，把这些风险模型手册看明白了，专业化程度会大幅度提升。



- Barra CNE5 Model Handbook （请自行下载）
- Axioma AXCN Model Handbook （请自行下载）

涉及到版权的因素，不能直接把这些提供给大家，但在网上搜一搜都是可以找到的并且可以下载。希望系统掌握现代金融学理论的同学以及在金融界找到正规工作的同学一定要仔细的看一看。

今天的课程就到这里，谢谢大家。

-END-

加入“量化小学”的见识圈，关注动态

感谢您订阅本特辑，扫描下方二维码或[点击圈子链接](#)，即可加入专属见识圈子提问交流





## 量化小学



渔生

小学而大不遗，量化师生联谊会

感谢大家订阅《量化小学》，这里是学校见识社群，你可以随时提问、随时互动，我们一起投资，一起分享！



风险提示及免责条款

市场有风险，投资需谨慎。本文不构成个人投资建议，也未考虑到个别用户特殊的投资目标、财务状况或需要。用户应考虑本文中的任何意见、观点或结论是否符合其特定状况。据此投资，责任自负。

写评论

请发表您的评论



表情

图片

发表评论

华尔街见闻

- 关于我们
- 广告投放
- 版权与商务合作
- 联系方式
- 意见反馈

法律信息

- 版权声明
- 用户协议
- 付费内容订阅协议
- 隐私政策



华尔街见闻APP



华尔街见闻公众号



微博@华尔街见闻



## 声明

未经许可，任何人不得复制、转载、或以其他方式使用本网站的内容。

评论前请阅读网站 [“跟帖评论自律管理承诺书”](#)

## 违法和不良信息举报

举报电话: 021-60675200 (周一到周五9:30-11:30, 13:00-18:30)

举报邮箱: [contact@wallstreetcn.com](mailto:contact@wallstreetcn.com)

网站举报: [点击这里](#)

[违法和不良信息举报受理和处置管理办法](#)

[清朗·财经违规内容专项整治公告](#)



中央网信办  
违法和不良信息举报中心



上海市互联网  
违法和不良信息举报中心

## 友情链接

[腾讯财经](#) | [财经网](#) | [澎湃新闻](#) | [界面新闻](#) | [全景财经](#) | [陆家嘴金融网](#) | [富途牛牛](#) | [网易财经](#) | [凤凰网财经](#) | [虎嗅](#)

