




第4讲（B）：量化选股模型

运用基本面多因子模型预测
回报和风险



概览

- 投资流程：一般性框架
- 基于基本面多因子模型的量化选股策略：（Haugen和Baker，《金融经济学》，1996年）
 - 明确公司特定因素
 - 估计和预测因子报酬
 - 预测预期回报率
 - 构建多-空投资组合
 - 分析经风险调整后的投资组合回报



投资流程：三个主要步骤

- 初步筛选
- 选股
- 投资组合构建和风险控制



初步筛选

- 目标：剔除难以大量交易的股票
 - 美国证交所所有大约8,000只上市交易股票
 - 收缩到1,000-2,000只股票
- 标准：
 - 流动性
 - 能卖空
 - 市场影响力
 - 并购交易涉及的股票
 - 剔除某些板块和行业



初步筛选（续）

■ 流动性

- 高换手率要求股票易于借得和出售。
- 大多数基金经理将可投资股票的范围限定在流动性较好的大型公司。
 - 1,000-2,000只流动性最佳的国内股票
- 某些大公司一直受到华尔街分析师的密切关注。
 - 得到分析师关注提升了这些股票的价格效率。
 - 有些采用股市中性策略的基金经理会避免选择这类股票。
 - 还有一些基金经理会使用一种不同的因子模型来分析最大的几家公司。



初步筛选（续）

■ 能卖空

- 剔除那些被经纪商列入“难借”名单的股票。
- 报升规则进一步增加了卖空的难度。
 - 只有在股价出现上涨之后才能执行。
- 基金经理往往先对投资组合中的空头部分进行筛选，然后才对多头部分进行筛选。



初步筛选（续）

- 市场影响力

- 对某股票进行大宗交易可能对该股票的市价产生不利影响。
- 基金经理可以把这些股票从他们的可投资范围内剔除出去。
- 市场影响力衡量指标：
 - 占日均交易量百分比
 - 变现一定规模头寸所需的时间



初步筛选（续）

- 并购交易涉及的股票

- 剔除并购交易涉及的股票
- 一宗待定的并购交易，其结果的不确定性可能会拉低本来很不错的股票评级，因而削弱基金经理所采用的选股模型的预测力。

- 剔除某些板块和行业

- 剔除性态表现与模型预测相异的行业或部门
 - 传统的价值衡量指标对互联网股票几乎没有预测力。

选股

- 一般方法
 - 运用多因子选股模型对可投资股票创建相对评分
- 多因子选股模型

$$r = w_1 f_1 + w_2 f_2 + \dots + w_n f_n$$

- r = 预计Alpha值（超额回报）
- f = 基本面因子或技术面因子
- w = 因子权重

选股：主要步骤

- 第1步：
 - 针对每一只股票分析每一个因子，并对因子评分，如，1-10分（1分-最糟，10分=最好）
- 第2步：
 - 确定因子权重 w ，并且计算每一只股票的原始得分。
 - 原始得分 = $w_1 * (f_1 \text{的评分}) + w_2 * (f_2 \text{的评分}) + \dots + w_n * (f_n \text{的评分})$
- 第3步：
 - 依据原始得分对股票进行分类，进行十分位排名
 - 第1十分位 = 最没吸引力
 - 第10十分位 = 最有吸引力

选股：举例

- 一个假设的四因子模型
 - f_1 = 市盈率 (P/E), f_2 = 账面市值比 (B/M), f_3 = 股利收益率 (D/P), f_4 = 动量
 - $w_1 = 0.4, w_2 = 0.3, w_3 = 0.2, w_4 = 0.1$
- XYZ公司的因子排名如下：
 - P/E评分为1
 - B/M评分为6
 - D/P评分为2
 - 动量评分为10
- 计算原始得分：
 - $(0.4*1) + (0.3*6) + (0.2*2) + (0.1*10) = 3.6$

选股：因子权重

- 基金经理如何得出因子权重？
 - 平均加权：
 - 如果你没有线索
 - 主观判断
 - 以往数据的回归分析：
 - 对因子权重的估计值应该让预测回报与实际回报的相关性最大化。



选股：买卖规则

- 依据通过多因子选股模型得出的相对排名制定买卖决策。
- 达到一定排名的股票要纳入多头头寸。
- 低于一定排名的股票要卖空，等其上升到一定排名时再平仓。
- 买卖规则一般在风险最优化的框架内才发挥作用。
 - 构建最优投资组合（下一个话题）



投资组合构建和风险控制

- 目标：
 - 风险与回报的平衡
 - 构建由相等金额的高排名股票和低排名股票组成的最优组合，同时保持对选定系统性风险因子的净敞口为零。
- 除了市场风险以外，投资组合通过适当构建也能中和其他常见的风险因子。
- 优化工具：
 - 二次优化工具，如APT和BARRA
 - 专有优化工具



投资组合构建和风险控制：一般的风险因子

- **Beta值：市场风险**
- **板块或行业**
 - 在同一个板块或行业做空和做多，以避免板块风险。
- **市值**
 - 市值相近的股票受系统性、方向性价格变动的影响，而后者受以下因素的驱动
 - 流动性偏好
 - 投资者情绪变化
 - 在市值相近的股票群体中做多和做空，以消除市值风险。



投资组合构建和风险控制：一般的风险因子（续）

- **利率**
 - 有些公司对利率的变动比其他公司敏感：
 - 高杠杆公司
 - 大型放贷者
 - 在杠杆率相近的公司范围内做多和做空，以消除利率风险。
- **劳动用工**
 - 工会力量强大的公司容易面临劳资纠纷和罢工的风险。
 - 在工会力量相近的公司范围内做多和做空，以消除劳动用工风险。



投资组合构建和风险控制：一般的风险因子（续）

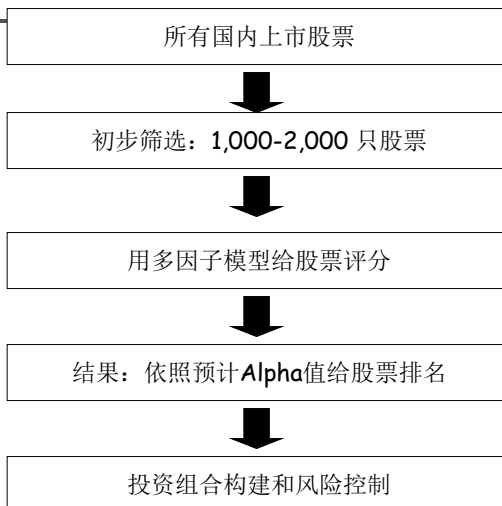
- 商品价格：油价
- 市盈率
- 市净率
-



投资组合构建和风险控制：执行和交易

- 基金经理必须设计有效的系统，以确保其交易得到正确执行，而且成本尽可能低。
 - 机会成本
 - 市场影响
 - 佣金

投资流程：概要



量化选股

- Haugen和Baker（《金融经济学》，1996年）
 - 导致股票回报率存在差异的基本面决定因素和技术面决定因素在不同时间和不同国家是稳定的。
 - 多因子预期回报模型的预测力很强。
 - 研究表明，依照预测的预期回报率确定的多-空投资组合具备盈利能力。



多因子模型中使用的公司特定属性：

- 风险因子
 - 风险越高 \Rightarrow 预期回报率就越高
- 流动性因子
 - 流动性越高 \Rightarrow 预期回报率就越低
- 表明相对于会计数字的价格水平的因子：
 - 比率越高 \Rightarrow 预期回报率就越低
- 表明增长潜力的因子
 - 增长潜力越高 \Rightarrow 预期回报率就越高
- 技术面因子
- 行业变量



风险因子

- 市场Beta值（过去60个月每月超额回报率的回归分析）
- APT Beta值（过去60个月的短期国债收益率、工业产值的变动百分比、通胀率、长短期政府债券收益率差，以及公司债和政府债的收益率差的回归分析）
- 总回报率的波动性
- 剩余方差
- 收益风险
- 债务权益比率
- 利息保障倍数（净经营收益/利息总额）
- 收益可变性（盈余、股利和现金流收益在过去5年里的波动性）



流动性因子

- 市值
- 每股市场价格
- 交易量/市值（过去12个月的月交易量/市值的均值）
- 交易量趋势（月交易量在过去5年里的趋势）



表明相对于会计数字的价格水平的因子

- 收益价格比（ E/P ）（可获得的最近四个季度的数据，收益/当前市价）
- 账面市值比（ B/P ）（可获得的最近数据，账面价值/当前市价）
- 股利价格比（ D/P ）（可获得的最近四个季度的数据，股利/当前市价）
- 现金流价格比（ CF/P ）（可获得的最近数据，(每股收益+每股折旧)/当前市价）
- 销售价格比（ S/P ）（可获得的最近四个季度的数据，每股总销售额/当前市价）



表明相对于会计数字的价格水平的因子（续）

- 最近有研究表明，市价/当前现金流比率较低的股票在最近数十年赚取的回报率较高。
- 较高回报的来源：理性 vs. 行为
 - 价值股是“降临凡间的天使”，更具风险性。
 - 更高的回报率补偿更高的系统性风险。
 - Chan和Chen (1991)；Fama和French (1992)
 - 价值股的溢价回报是不可预料的。
 - 如果价值股的投资者预见到财务困境将长期持续，就可能把股价压得过低（相对于公允价值）。
 - Chopra, Lakonishok和Ritter (1992), Lakonishok, Shleifer和Vishney (1994)以及Haugen (1995)



增长潜力因子

- 利润率（净经营收益/总销售额）
- 资本周转率（总销售额/总资产）
- 资产回报率（净经营收益/总资产）
- 权益回报率（净收益/总权益资本的账面价值）
- 收益增长率（过去5年每股收益的季度时间趋势除以过去5年的每股收益均值）



增长潜力因子（续）

- 增长潜力因子表明股票的收益和股利在未来实现高（低）于平均增长水平的概率。
- 运用当前和以往获利能力的衡量指标作为未来增长率的预测因子。
 - 基础假设：当前盈利的公司在未来有更大的增长潜力。



技术面因子

- 前1个月的超额回报
- 前2个月的超额回报
- 前3个月的超额回报
- 前6个月的超额回报
- 前12个月的超额回报
- 前24个月的超额回报
- 前60个月的超额回报

衡量相对于标普500指数的超额回报。



技术面因子（续）

- 技术面因子描述了一只股票的历史价格。
- 最近有研究表明了历史回报与未来预期回报之间的三种关系。
 - 短期逆转
 - 中期惰性
 - 长期逆转



短期逆转

- 回报产生非常短期（1-2个月）的逆转模式：
 - 如果股票在上个月价格显著上涨，这可能标志着下个月会出现逆转。
- 解释：
 - 投资者试图快速买入或抛售大量特定股票，由此产生了价格压力。
 - Jegadeesh (1990)
 - 买卖报价反弹：
 - Ball, Kothari和Wasley (1995)
- 预测：
 - 对一只股票在过去1-2个月内的表现的负向报酬。



中期惰性（动量）

- 过去**6-12**个月内表现不俗（糟糕）的股票未来会有不俗（糟糕）的前景。
- 解释：Jegadeesh和Titman (1993)
 - 对每一份收益报告的滞后反应
 - 对公司超常高/低利润率的初始报告反应不足
- 预测：
 - 对一只股票在过去**6-12**个月内的表现的正向报酬。



长期逆转

- 过去**3-5**年内表现不俗（糟糕）的股票，其未来的回报将呈现逆转模式。
- 解释：
 - 市场对一系列有关良好（糟糕）收益的正面（负面）报道的过度反应。
 - 投资者相信这种趋势将会继续延长到未来的一段时间，就会把股价抬得过高（压得过低）。
- 预测：
 - 对一只股票在过去**2-5**年里的表现的负向报酬。



行业变量

- 反映公司主要业务线的0/1虚拟变量
 - 耐用品
 - 非耐用品
 - 公用事业
 - 能源
 - 建筑
 - 商业设备
 - 制造
 - 运输
 - 金融和商业服务



估计因子报酬

- 对于每个月份，运用最小二乘法横截面回归分析同时估计每一种风险因子的月度报酬。

$$r_{j,t} = \sum_i P_{i,t} * F_{j,i,t-1} + u_{j,t}$$

其中，j表示股票，i表示因子。

上述回归分析是对若干连续月份进行估算，以获取各种因子的历史报酬情况。



估计因子报酬（续）

- $r_{j,t}$ = 股票j在月份t的回报率
- $P_{i,t}$ = 因子i在月份 t的回归系数或报酬
- $F_{j,i,t-1}$ = 股票j在月份（t-1）末对因子i的敞口（公司特点，比如APT Beta值、规模、获利能力等）
 - $F_{j,i,t-1} = (E_{j,i,t-1} - \text{均值}) / (\text{标准差})$
- $u_{j,t}$ = 股票j在月份t的回报中的未解释部分



估计因子报酬（续）

因子	1979/01 – 1986/06		1986/07 – 1993/12	
	均值	T统计值	均值	T统计值
1个月超额回报率	-0.97%	-17.04	-0.72%	-11.04
12个月超额回报率	0.52%	7.09	0.52%	7.09
交易量/市值	-0.35%	-5.28	-0.20%	-2.33
2个月超额回报率	-0.20%	-4.97	-0.11%	-2.37
收益价格比	0.27%	4.56	0.26%	4.42
权益回报率	0.24%	4.34	0.13%	2.06
账面实值比	0.35%	3.90	0.39%	6.72
交易量趋势	-0.10%	-3.17	-0.09	-2.58
6个月超额回报率	0.24%	3.01	0.19%	2.55
现金流价格比	0.13%	2.64	0.26%	4.42
现金流价格比可变性	-0.11%	-2.55	-0.15%	-3.38

估计预期回报率

- 运用估计的历史报酬情况对后几期的未来回报率进行样本外预测
- 针对需要估计预期回报率的那个月，使用该月之前**12**个月中观察到的报酬均值

$$E(r_{j,t}) = \sum_i E(P_{i,t}) * F_{j,i,t-1}$$

因子*i*在月份*t*中的预期报酬（过去**12**个月的估计报酬的算术平均值）

形成投资组合

- 从**1979年1月**到 **1993年12月**：
 - 依据观察到的因子敞口和预测的因子报酬计算每只股票的预期回报率。
 - 依据相对的预期回报率对股票进行排名
 - 把这些股票归入十个等权重的十分位：
 - 第**1**十分位：预期回报率最低的股票
 - 第**10**十分位：预期回报率最高的股票
 - 每个月都重复上述流程
- 多-空投资组合：
 - 对第**10**十分位的股票做多，对第**1**十分位的股票做空
 - 每个月对投资组合进行再调整



实现的年化回报率

年	第1十分位	第10十分位
1981年	-15.6%	16.2%
1983年	11.8%	54.5%
1985年	4.3%	45.7%
1987年	-23.8%	-5.1%
1989年	-3.0%	28.7%
1991年	23.9%	57.4%
1993年	6.4%	22.2%
均值: 1979-1993	-4.5%	30.9%
标准差: 1979-1993	22.62%	18.50%



实现的年化回报率：综述

- 在每一年，当我们从第1十分位观察到第10十分位时，我们发现，实现的回报率往往变得越来越大，差距大得令人感到意外。
- 在整个时期，第10十分位与第1十分位之间的差距大约达到**35%**。
- 当我们从第1十分位向第10十分位移动，波动性就逐渐减小。
 - 投资者可能没有把第10十分位的股票视为具有高度风险。



各个十分位内股票的一般特点

- 我们从第**1**十分位移到第**10**十分位时，股票表现出：
 - 财务杠杆度越来越低
 - 利息保障倍数越来越高
 - 市场Beta值越来越低
 - 总回报率的波动性越来越小
 - 收益增长率越来越高
 - 盈利率越来越高
 - 市值越来越大
 - 往年绩效更好



如何解释回报率差异？

- 不太可能的情况：
 - 第**10**十分位的股票是困境公司，被视为比第**1**十分位中的股票风险大。
- 可能的情况：
 - 第**10**十分位的股票是“价廉物美”的“高质量股票”。
 - 表明市场可能失灵。



经风险调整后的实现回报率

- 以**Fama-French**模型中的三种因子为自变量，以月回报率的十分位数为因变量，进行回归分析：
 - 较低（高）的回报率十分位数的截距显著为负（正）。
 - 较高的回报率十分位数对以下因子的依赖性较小：
 - 规模因子：常见于大盘股；
 - 价值因子：包含具有强劲增长特征的股票。
 - 市场因子：反映出较低水平的市场相关风险。