

Alpha策略因子的选择与评价报告

2015-10-08 Alpha

请点击标题下方"**Alpha**"关注我们

作者：量化对冲实验室

摘要

本文主要概述了Alpha策略的整体思路，介绍了Alpha策略因子的分类方法、筛选和评价的一般方法和步骤，同时还较为详细地说明了因子的组合方法以及基于因子库的量化选股的一般方法和步骤。

目录

1.量化对冲Alpha策略简介

1.1 Alpha的含义

1.2 Alpha策略的基本思想

1.3 Alpha策略的分类

1.4 Alpha策略的优势

2.因子的分类

2.1 盈利性

2.2 估值

2.3 现金流

2.4 成长性

2.5 资产配置

2.6 价格动量

2.7 技术面因子

3.因子的筛选和评价

3.1 因子筛选的整体思路

3.2 因子评价的整体思路

3.3 基于因子胜率的因子评价

3.4 基于t检验的因子评价

3.5 基于IC和IR比率的因子评价

4.因子的组合

4.1 冗余因子的剔除

4.2 因子的降维

4.3 因子权重的确定

5. 基于因子库的选股

5.1 打分法

5.2 回归法

1.量化对冲Alpha策略简介

1.1 Alpha的含义

根据拓展的CAPM模型，我们知道，证券s的实际收益率满足：

其中， R_s 为现货组合的预期收益率， r_f 为无风险利率， R_m 为市场指数的预期收益， ϵ_s 为误差项， α 衡量了非系统性风险， β_s 衡量了系统性风险。投资者在市场交易中同时面临着系统性和非系统性风险，阿尔法策略通过对系统性风险进行度量并将其分离，从而获取超额绝对收益。

1.2 Alpha策略的基本思想

Alpha策略是典型的对冲策略，通过构建相对价值策略来超越指数，然后通过指数期货或期权等风险管理工具来对冲系统性风险。Alpha策略可以看成中性策略的一种，但是Alpha策略的约束更小，其Alpha来源可能是行业的、风格的或者其他的。Alpha策略注重选股，属于主动投资，相比之下，Beta策略注重对投资时机的选择，属于被动投资。股票的Alpha是它超过或低于通过CAPM模型预测的可能期望收益的部分，若股票定价公平，则其Alpha为0。

1.3 Alpha策略的分类

在实际中经常使用的Alpha策略主要有：多因子、风格轮动、行业轮动、资金流、动量反转等。

多因子是应用最为广泛的一种策略，该策略选择一系列因子搭建模型。通过这些因子筛选股票，满足则买入，不满足则卖出。多因子的最大优势在于，在不同的市场和行情下，因子库中总有一些因子能够发挥作用。

风格轮动是指利用市场的风格特征进行投资。市场有时会偏好小盘股，有时偏好大盘股。通过观察某些指标来判断市场的倾向性，在风格转换的初期介入，我们可以获得较大的超额收益。

行业轮动是指市场在经济周期的作用下对各个行业产生不同的偏好。在经济周期中，我们可以按照顺序依次对各个行业进行资产配置，从而获取相比于买入持有策略的超额收益。

资金流是根据资金的流向来进行选股。对于一只股票，资金流入时，股票的价格应该会上涨；资金流出时，股票的价格应该会下跌。通过观察资金流的情况，我们可以预测未来股价的变化。

动量反转是指股票的强弱变化情况，过去一段时间强的股票，在未来一段时间继续保持强势，过去一段时间弱的股票，在未来一段时间继续弱势，这叫做动量效应。过去一段时间强的股票在未来一段时间走弱，或者过去一段时间在弱的股票在未来一段时间走强，这叫做反转效应。通过判断动量和反转的有效性，筛选出应该购买的股票。

市场上量化投资者使用的Alpha策略还有很多，在此不做赘述。

1.4 Alpha策略的优势

Alpha策略有三大优势：一是回避了择时这一难题，仅需专注于选股；二是波动较单边买入持有策略要低；三是在单边下跌的市场下也能盈利，Alpha与市场的相关性理论上为0。在熊市或者盘整期，可以采用“现货多头+期货空头”的方法，一方面建立能够获取超额收益的投资组合的多头头寸，另一方面建立股指期货的空头头寸以对冲现货组合的系统风险，从而获取正的绝对收益。

2.因子的分类

对于因子的分类方法很多，整体而言，因子可以被分为基本面因子和技术面因子。基于对一只股票的不同特征的刻画，我们可以将因子更加细致地分为：盈利性、估值、现金流、成长性、资产配置、价格动量和技术面因子。表格 1是对因子分类方法的总结。

表格 1

类型	常见因子
盈利性	投入资本回报率（ROIC）、已利用资本回报率（ROCE）、净资产收益率（ROE）、总资产收益率（ROA）、边际利润、人居收入、经济利润、投资超额收益
估值	自由现金流价格比、外部融资总资产比、企业价值与EBITDA比（EV/EBITDA）、市盈率、股息率
现金流	自由现金流（FCF）和营业收入之比、投入资本现金回报率（ROIC）

成长性	每股自由现金流、每股盈余
资产配置	净回购股份与投入资本比、一年流通股减少量、一年长期债务减少量、外部融资和总资产比、三年平均资本支出和投入资本比、收购与投入资产比
价格动量	相对强弱、价格范围、相对强弱指数
技术面因子	强弱指标(RSI)、随机指标(KD)、趋向指标(DMI)、平滑异同平均线(MACD)、能量潮（OBV）

2.1 盈利性

与盈利性相关的因子主要反应了公司利用现有资源实现收益的能力。公司的盈利能力可以通过许多方法来衡量，例如投入资本回报率（ROIC）、已利用资本回报率（ROCE）、净资产收益率（ROE）、总资产收益率（ROA）、边际利润、人居收入、经济利润、投资增额收益率。整体而言，盈利性因子是一类效果较好的因子，即盈利性好的公司股票具有显著的正超额收益，而盈利性差的公司股票具有显著的负超额收益。

2.2 估值

估值因子主要反映了股票作为一种资产的价值，这种资产价值和股票的价格相关，但其决定性因素是该公司为其客户创造价值的多少。估值可以通过许多方法得到，但都包括了一定的定性分析和对未来的预测。常见的估值因子有：自由现金流价格比、外部融资总资产比、企业价值与EBITDA比（EV/EBITDA）、市盈率、股息率等。

2.3 现金流

现金流可以分为经营性、投资性和融资活动三类。其中，我们主要关注经营性现金流。经营性现金流，包括从商品销售和服务得到的现金减去生产这些产品和提供这些服务需要支付的现金流出，其中包括为利润支付的现金税和为债务融资支付的利息。一个公司产生的经营性现金流水平是衡量未来股市回报的一个重要指标。常见的现金流因子有：自由现金流（FCF）和营业收入之比、投入资本现金回报率（ROIC）等。

2.4 成长性

成长性因子在市场中通常获得的超额收益较为微弱。这主要是因为成长性投资更多地是定性而非定量的，更加依赖投资者独到且有前瞻性的眼光而非精确的数量分析，更加偏向于”艺术“而非”量化“。尽管如此，成长性仍然是我们因子库中重要的一部分。正如成长性投资者们所说的：”我所知道的投资中的最大的一个错误，就是对那些最伟大公司和其他普通投资一视同仁。“在实际使用成长型因子的过程中，我们常常和其他因子结合使用，以弥补其预测性不足的劣势。

常见的成长性因子有：每股自由现金流、每股盈余等。

2.5 资产配置

资产配置主要涉及一家公司资本资源的使用情况，我们主要考虑现金来源和现金使用两方面的内容。现金来源主要包括业务经营、资产和投销售收入、发行股票和发行债券等。现金使用主要包括经营费用、业务投

资、业务收购、项目或证券投资、支付现金股利、偿还债务及回购股份等。常见的资产配置因子有：净回购股份与投入资本比、一年流通股减少量、一年长期债务减少量、外部融资和总资产比、三年平均资本支出和投入资本比、收购与投入资产比等。

2.6 价格动量

价格动量因子能够在一定意义上反映市场周期和投资者情绪，并依此对未来进行预测。价格动量一般的衡量指标是价格变化的速度，或一段时间内价格的变化率。正的价格动量意味着某只股票的买家数量正在不断增加，需求大于供给；负的价格动量则意味着供需平衡向卖家倾斜，供给大于需求。常见的价格动量因子有：相对强弱、价格范围、相对强弱指数等。

2.7 技术面因子

技术面因子相比于基本面因子，数据更新时间快，更加注重市场的微观结构，而非股票的价值。常见的技术面因子有：强弱指标(RSI)、随机指标(KD)、趋向指标(DMI)、平滑异同平均线(MACD)、能量潮 (OBV)等。由于技术面因子的Alpha往往不稳定，所以在实际应用中较为少见。

3.因子的筛选和评价

对于一家上市公司而言，我们能够获取到大量的数据。这些数据中有的能够很好地描述上市公司的特性，我们通过对这些数据进行分析，就有可能从中找出具有显著超额收益的因子。

3.1 因子筛选的整体思路

上市公司的数据多而繁，在对这些数据进行处理之前，我们需要对数据进行选择，即对因子进行初步的筛选。我们筛选因子的主要原则有：**a.** 数据的准确性和真实性；**b.** 数据的完整性；**c.** 数据来源的稳定性。一般情况下，我们可以通过选择一个专业的、满足要求的金融数据库来完成因子的初步筛选。

3.2 因子评价的整体思路

在完成了因子的初步筛选之后，我们可以对因子进行进一步的评价。因子评价的整体思路是研究各个股票的超额收益和因子参数之间的关系。我们试图找出这样的因子：对于绝大多数股票而言，当该因子参数越大/越小时，超额收益越大/越小，或者关系恰好相反。总之，我们试图找出那些能够有效预测一只股票未来超额收益的因子，无论两者是正相关还是负相关。

研究股票超额收益和因子参数之间关系的方法主要有两种。第一种是根据因子参数的大小对股票进行分组，计算每组的平均超额收益，并依次进行因子胜率、t检验。第二类是在每一个时间点上，计算全体股票截面上的超额收益率和因子参数大小的相关系数以及信息比率。

3.3 基于因子胜率因子评价

对于每一个时间点，我们都可以得到每只股票的因子参数，根据这些因子参数，我们将股票平均分成五组，计算这五组股票的超额收益。我们可以根据经验设定一些阈值，当参数超过阈值时我们认为该因子有效。

对于每一个时间点，我们可以计算每组股票的超额收益序列和组数序列的相关系数，当时，我们认为该因子

在这个时间点有效。其中，MinCorr为一个阈值， x_i 为第*i*组的平均超额收益。

对于每一个时间点上第一组和最后一组的平均超额收益，我们还可以进行判定，当或者时，我们认为该因子在这个时间点上明显跑赢或者跑输市场，即该因子有效。其中，为两个阈值。

对每个时间点都进行上述判定，我们可以得出某个因子在多少个时间点上有效。用有效的时间点的个数除以时间点总的个数，我们可以得到因子的胜率。胜率越大，我们认为因子的有效性持续时间越长，即有效性越好。

3.4 基于t检验的因子评价

和3.3中的思路相同，我们仍然计算每个时间点上各个分组的收益率均值。我们可以对这个均值序列进行假设检验。原假设为 $H_0:\mu_1=\mu_2=\mu_3=\mu_4=\mu_5$ ，根据t检验，我们可以得出P值。当P值小于20%的时候，我们认为可以拒绝原假设，即我们认为五个组的收益率均值不等。P值越小，该因子越有效。

在实际操作中，为了判断一个因子是否在某个时间点有效，我们可以先进行t检验，再进行相关性检验，最后观察第一组和第五组的超额收益情况。若因子能通过上述三个测试，则认为该因子在该时间点有效。

3.5 基于IC和IR比率的因子评价

对于某个时间点A，我们可以将所有股票的因子参数看成是一个序列，将下一个时间点对应的所有股票的超额收益看成是一个序列，计算着两个序列的相关系数，我们就可以得出IC在时间点A上的取值。越高，说明因子对下一个时间点超额收益的预测越好。时间点间隔的选择根据调仓周期变化。

在计算出所有的之后，我们得到了IC的时间序列，求这个时间序列的均值和标准差并求出两个之比，我们便得到了IR，即

IR的绝对值越大，说明该因子的预测效果越好，即越有效。

4. 因子的组合

4.1 冗余因子的剔除

由于很多因子内在的驱动原因大致相同，所以通过这些因子选出的股票往往很相似。实际运用中，我们往往会考虑剔除掉一些有效，但是信息冗余的因子，即在同类的因子中，我们只需要保留收益最好、区分度最高的那一个。

剔除冗余因子的一般方法如下：**a.** 我们取出各个有效因子在各个时间点上关于分组的序列；**b.** 计算这些序列的相关性矩阵；**c.** 得到相关性矩阵关于时间的序列，求这个序列的均值；**d.** 通过这个均值矩阵挑出相关性较大的因子组；**e.** 对于每个因子组，挑选其中有效性最好或者收益最好的一个因子作为最终的因子。

4.2 因子的降维

在多因子模型的实际应用中，我们希望将有效的因子加以组合和处理，减少模型中变量的个数，这种减少自变量的过程叫做降维。降维的主要方法有：因子简单平均降维法、因子历史平均收益率加权平均降维法、逐

步回归、主成分分析等，下面依次简述这些方法。

因子简单平均降维法：因子简单平均降维法就是对同类的因子进行简单的等权平均处理，我们对因子参数求平均，作为新的复合因子。

因子历史平均收益率：加权平均降维法就是对同类的因子按照历史平均收益求加权平均，因子的历史平均收益取各个时间点分组的第一组的收益。

逐步回归：在实际的多元回归问题中，我们总试图找到所谓“最优”回归方程，主要是指希望在回归方程中包含所有对因变量y影响显著的自变量而不包含对y影响不显著的自变量的回归方程。逐步回归分析正是根据这种原则提出来的一种回归分析方法。它的主要思路是在考虑的全部自变量中按其对于y的作用大小,显著程度大小或者说贡献大小,由大到小地逐个引入回归方程,而对那些对y作用不显著的变量可能始终不被引入回归方程。另外，已被引入回归方程的变量在引入新变量后也可能失去重要性,而需要从回归方程中剔除出去。引入一个变量或者从回归方程中剔除一个变量都称为逐步回归的一步,每一步都要进行F检验,以保证在引入新变量前回归方程中只含有对y影响显著的变量,而不显著的变量已被剔除。逐步回归分析的实施过程是每一步都要对已引入回归方程的变量计算其偏回归平方和(即贡献),然后选一个偏回归平方和最小的变量,在预先给定的F水平下进行显著性检验,如果显著则该变量不必从回归方程中剔除,这时方程中其它的几个变量也都不需要剔除(因为其它的几个变量的偏回归平方和都大于最小的一个更不需要剔除)。相反,如果不显著,则该变量要剔除,然后按偏回归平方和由小到大地依次对方程中其它变量进行F检验。将对y影响不显著的变量全部剔除,保留的都是显著的。接着再对未引入回归方程中的变量分别计算其偏回归平方和,并选其中偏回归平方和最大的一个变量,同样在给定F水平下作显著性检验,如果显著则将该变量引入回归方程,这一过程一直继续下去,直到在回归方程中的变量都不能剔除而又无新变量可以引入时为止,这时逐步回归过程结束，从而得到最优的回归方程。在逐步回归之后留下的因子即为降维后剩下的因子。

主成分分析：主成分分析的基本思路是将原来具有相关性的一些指标组合成一组新的互相无关的综合指数来代替原来的指标。一般情况下，我们去原来指标的线性组合作为新的综合指标。我们认为一个综合指标的方差越大，其包含的信息也就越多。因此，在所有线性组合中，我们去方差最大的那一个作为第一主成分。如果认为第一主成分不能有效地反映原来的信息，我们就取另一个和第一主成分相关系数为0的线性组合作为第二主成分，以此类推。具体操作中，我们求出各个参数的协方差矩阵，并求出矩阵的特征值和特征向量，将特征值按照从大到小的顺序排列，各个特征值对应的特征向量即为各个主成分线性组合对应的系数。在完成了主成分分析之后，我们可以选择能够较为完成反映信息的前几个主成分来代替原来的因子，达到降维的目的。

4.3 因子权重的确定

在完成了因子的筛选和降维之后，我们就需要确定因子权重。对因子赋权的方法有很多，在此简要介绍三种：等权赋值、回归赋值、IC-IR因子赋值。

等权赋值：等权赋值是指在组合各个因子时对各个因子赋以相等的权重。

回归赋值：回归赋值是指在组合各个因子时，我们对某个时间区间上的收益率和参数因子进行最小二乘法回归，回归所得的系数向量即为各个因子的权重向量。具体而言，对于某个时间区间[t1,t2]，我们考虑每个时间点上的截面回归： $Y_t=X_t\beta+\epsilon_t, t \in [t1,t2]$ 。其中，Yt为时间点t截面个股的收益率向量，Xt为为时间点t个股的

因子载荷参数，我们通过最小二乘的思路求解 β ，使得下面的式子取得最小值： $\sum(Y_t-X_t\beta)'(Y_t-X_t\beta)$ 。解得的 β_0 即为各个因子的权重。

IC-IR因子赋值：IC-IR因子赋值是指在组合各个因子时，考虑因子的IC序列，优化因子组合的IR值，取使得IR值最大的组合权重为最终的权重。我们假设有n个因子，其IC值的均值向量为

IC的协方差矩阵为 Σ 。设权重向量为，我们能够得出：

为了得到IR的最大值，对IR求的偏导数，得：

令，解出令IR最大的为。我们按照o对各个因子赋以权重即可。

5.基于因子库的选股

在完成了因子的筛选和组合之后，我们就建立起了自己的Alpha因子库。基于这个因子库，我们可以筛选出这些因子较为突出的股票，并通过这些股票实现因子的超额收益。常见的选股方法有两种，分别是打分法和回归法。

5.1 打分法

打分法就是根据各个因子的大小对股票进行打分，然后按照一定的权重相加得到一个总分，通过分数的高低进行股票的筛选。打分法的特点是比较稳健，不易受到特殊值得影响。

5.2 回归法

回归法就是用过去的股票的收益率对多因子模型进行回归，得到回归方程，然后把最新的因子值代入到回归方程中得到一个对于未来股票值得预测，根据这个预测来进行股票的筛选。回归法的优点是能够比较及时地调整股票对各个因子的敏感性，但是回归法比较容易受到极端值得影响，导致选股失败。

对冲基金人才协会（Hedge fund Talents Association-HTA）是国内首个面向全国对冲基金高级管理人才的专业类协会。协会宗旨是建立由本土资产管理行业精英为主的行业平台，打造对冲基金界的高端人才交流平台，提供帮助海内外对冲基金落户，融资，招聘等一系列促进行业发展的服务。目前协会机构会员单位包括中金公司，中信证券，平安证券，海通证券，永安期货等，专家会员来自重阳，朱雀，景林等国内大型对冲基金董事总经理级别。个人会员所在机构管理资产国内超过3千亿人民币。对冲基金人才协会帮助优秀对冲基金投顾对接组合基金，孵化器资金。合作请发邮件至info@cnhta.org。

[阅读原文](#)

