

【Alpha 提纯】一种基于因子剥离的 FOF 择基逻辑

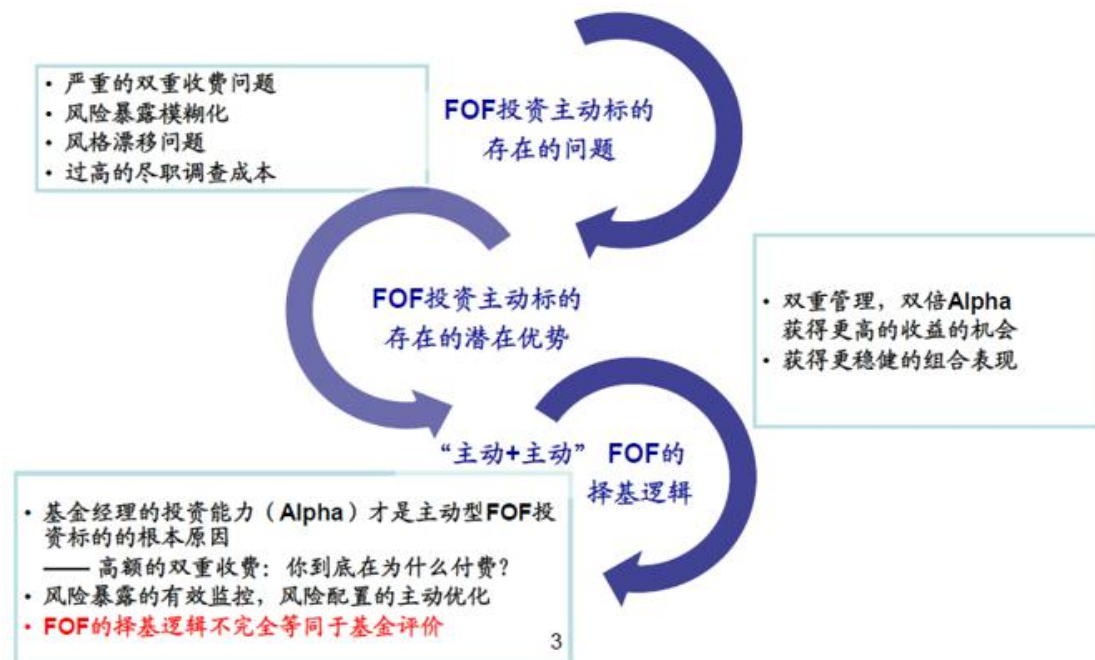
FOF 指的是将 80% 以上的基金资产投资于经证监会核准或注册的基金份额的基金，但不得持有具有复杂、衍生品性质的基金份额，也不得持有其他基金中基金。国内公募基金诞生 19 年以来，从零开始，穿越牛熊，发展已成为市场上举足轻重的角色。近期，随着监管层对公募 FOF 管理细则的不断出台，市场期盼已久的公募系 FOF 基金只待“临门一脚”，中国公募可谓进入 2.0 时代。那么，基金 2.0 版本的公募 FOF 该如何投资？今天私募云通就带大家来看一种基于因子剥离的 FOF 择基逻辑，也就是管理人常常挂在嘴边的 Alpha 提纯。



图片来源：海通证券金融工程

一．主动标的 vs. 被动标的

根据 2016 年私募 FOF 发展报告（点击名称进入链接），FOF 产品比照基础基金产品费率水平，存在重复收费问题，对投资业绩构成一定的拖累。另外，一个完整的 FOF 管理需要形成底层数据库和策略模型，而子基金管理人专业度参差不齐，因此存在过高的尽职调查成本。不论在投前筛选还是投中风控上，FOF 投资都存在风格暴露模糊、风格漂移等问题。



图片来源：海通证券金融工程

当然，正是由于 FOF 投资于主动标的带来这一系列的问题，反过来给 FOF 的发展提供了很

大的成长空间，双重管理模式带来的双倍 Alpha 是投资者与管理者都喜闻乐见的，并且高收益与组合表现稳健的并存更促进了 FOF 从而呈现爆发式增长。要知道，FOF 投资的根本标的其实是基金经理的投资能力（Alpha），以及风险暴露的有效监控、风险配置的主动优化等等。可以说，FOF 的择基逻辑并不止于基金评价。

二 . FOF 的择基逻辑

传统的基金评价体系关注指标，如历史超额收益、历史波动率、历史最大回撤、夏普比率、卡玛比率、索提诺比率等等，虽然在一定程度上反映了单位不同风险下所获得的收益/超额收益，但并不能真实地反映这些基金都暴露于何种风险，不能反映有多大程度的收益是真正的 Alpha（基金经理的选股、择时、配对交易等能力）！

Manager	Historical excess return	Historical volatility	Maximum drawdown	Sharpe ratio	Calmar ratio
Event driven-A	4%	4%	5%	1.0	0.8
Multi-strategy-C	6%	7%	19%	0.9	0.3
Multi-strategy-B	7%	8%	15%	0.9	0.4
Long short equities-D	7%	10%	25%	0.7	0.3
Fixed income-E	6%	9%	22%	0.7	0.3
Long short equities-G	6%	9%	6%	0.6	1.0
Arbitrage-A	10%	18%	48%	0.6	0.2
Relative value-A	6%	12%	24%	0.5	0.3
Long short equities-E	2%	4%	5%	0.5	0.4
Macro-A	4%	10%	15%	0.4	0.3
CTA/managed futures-A	6%	17%	19%	0.4	0.3
Long short equities-F	8%	31%	71%	0.2	0.1
Long short equities-C	3%	14%	32%	0.2	0.1
Macro-B	3%	14%	30%	0.2	0.1
Long short equities-A	6%	29%	60%	0.2	0.1
Others-B	3%	19%	51%	0.2	0.1
Long short equities-B	5%	31%	76%	0.2	0.1
Others-A	1%	10%	26%	0.1	0.0
Long short equities-H	2%	28%	56%	0.1	0.0
Event driven-B	0%	12%	27%	0.0	0.0
Distressed debt-A	-2%	16%	50%	-0.1	0.0
Fixed income-F	-1%	3%	7%	-0.2	-0.1
Fixed income-D	-4%	14%	43%	-0.3	-0.1
Multi-strategy-A	-5%	16%	32%	-0.3	-0.2

图片来源：海通证券金融工程

因此我们要引入风险因子剥离模型，抽茧剥丝来对 Alpha 进行提纯。首先是基于夏普单因子模型，它将风险资产剥除一层共同因子（市场收益率）带来的影响。接着下一步，延伸至多因子模型，不管是 Fama-French 三因子模型还是 Carhart 四因子模型，都是继续剥离共同因子（大/小盘、成长/价值、动量）。

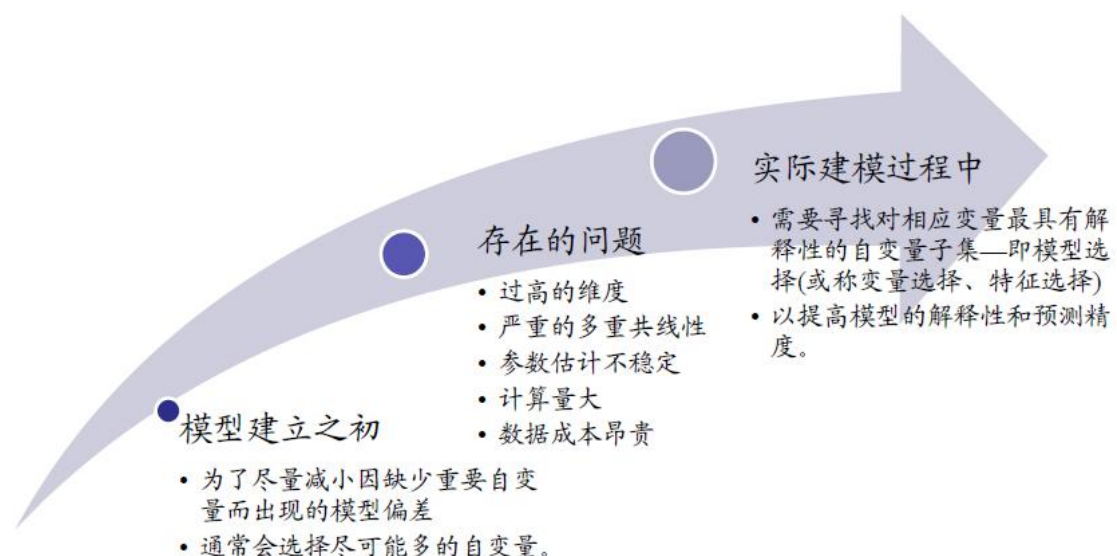


图片来源: 海通证券金融工程

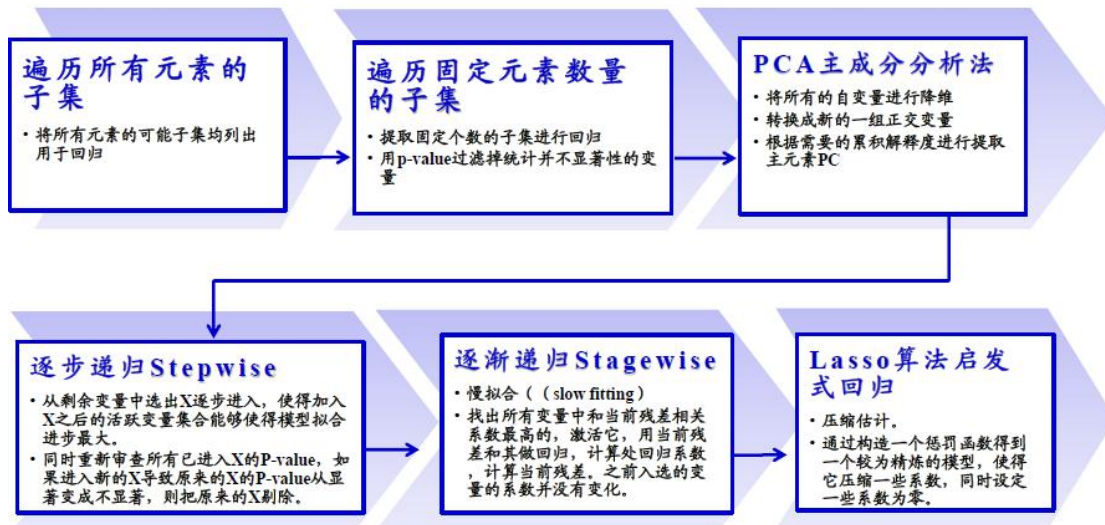
对于纯股票型基金基本已经能将 Alpha 提炼得比较纯净了, 然而带有其他类型资产的基金仍无法满足, 于是 PIMCO (太平洋投资管理公司) 在以上模型中又加入了久期、信用、商品等风险因子, 再过滤掉一层收益率。最后在海外投资中, 出现了基于不同资产的 SmartBeta 策略对于因子的进一步探索。

三 . 多因子拟合的算法优化

多因子模型在建立之初的目的是为了减小因缺少重要自变量而出现的模型偏差, 比如偏好大盘股的某只基金仅用资本资产定价模型 (CAPM), 是无法计算剔除大盘股指数之后这部分超额收益的。所以通常多因子模型会选择尽可能多的自变量, 如此便带来“维度诅咒”、多重共线性、参数不稳定等等问题, 在实际建模中需要多加考虑数据处理的细节、因子构造的逻辑, 以提高模型的解释力度与风险预测精度。



图片来源：海通证券金融工程



图片来源：海通证券金融工程

我们画出多因子模型构建的优化流程，值得注意的是逐步递归与逐渐递归。

逐步递归：

1. 算法目标：最小化残差平方和
2. 初始系数为 0，找出一个变量，估计其回归系数使得这模型最能够减小当前的残差平方和，即最大化拟合优度
3. 从剩余变量中选出 1 个，使得加入这个变量后的活跃变量集合能够使得模型拟合进步最大
4. 用所有活跃集中的变量进行回归
5. 重复直到符合某种准则
6. 每一步的关键在于找出最能改进拟合情况的变量，然后对所有入选的变量，即被激活的变量集合重新回归，也即，每一步都做了多元回归

逐渐递归：

1. 找出和当前残差相关系数最高的变量，激活，用当前的残差和其做回归，估计出回归系数，计算当前残差
2. 再找出所有变量中与当前残差相关系数最高的，激活，用当前的残差和其做回归，估计出回归系数，计算当前残差，此次计算，之前入选的变量系数没有发生变化
3. 再找出所有变量中与当前残差相关系数最高的（有可能是已经入选的变量），激活，用当前的残差和其做回归，估计出回归系数，计算当前残差
4. 重复上述过程直至达到某种准则
5. 对于观测样本量大于变量数的情况，最终的结果就是普通的最小二乘估计

Lasso 算法的拟合思想简单来说：

$$\hat{\beta}(\lambda) = \arg \min_{\beta \in R^p} \{ \|Y - X\beta\|_2^2 + \lambda \|\beta\|_1 \},$$

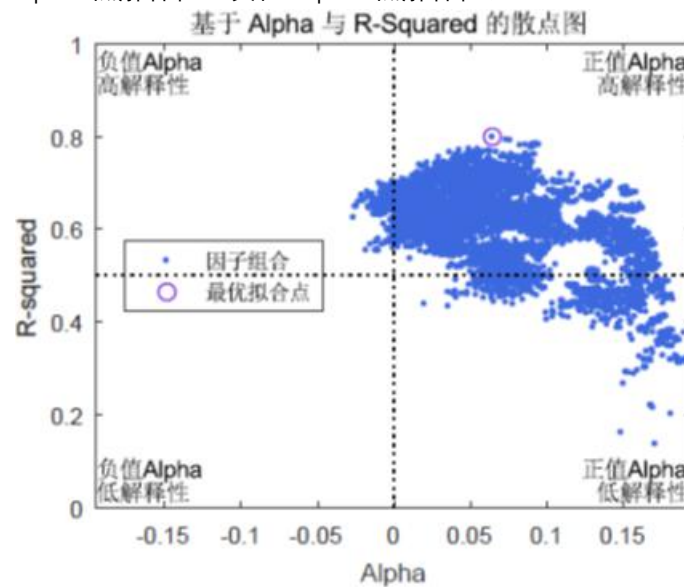
$$(\hat{\alpha}, \hat{\beta}) = \arg \min \left\{ \sum_{i=1}^N \left(y_i - \alpha - \sum_j \beta_j x_{ij} \right)^2 \right\} \quad \text{subject to } \sum_j |\beta_j| \leq t.$$

图片来源：海通证券金融工程

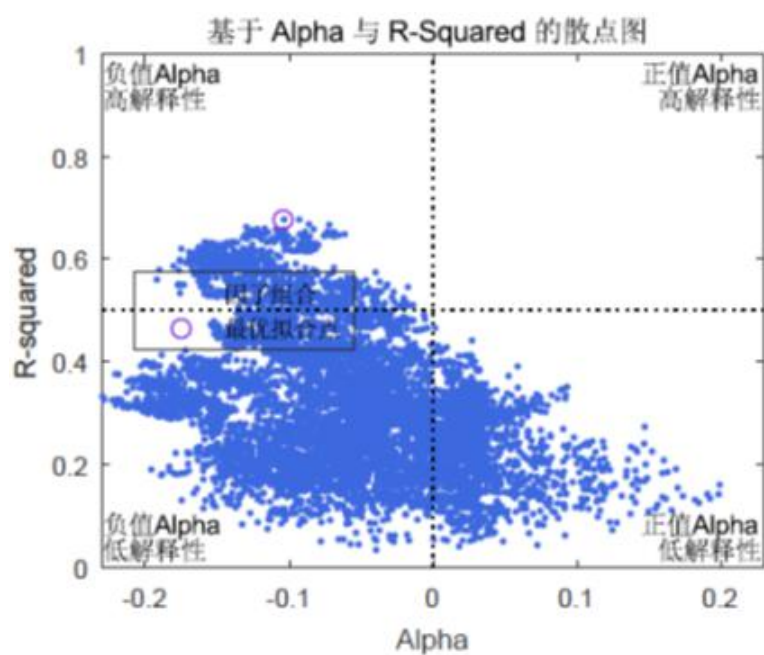
1. 核心：加入了惩罚项
2. 效果：对变量进行选择
3. 当 t 值小到一定程度时，该算法使得某些回归系数的估值为 0，起到变量选择作用
4. 当 t 值不断增大时，选入回归模型的变量会逐渐增多，当 t 值增大到某个值时，所有变量都入选了回归模型，此时得到的系数就是通常意义下的最小二乘估计

四．基于图像的分析逻辑

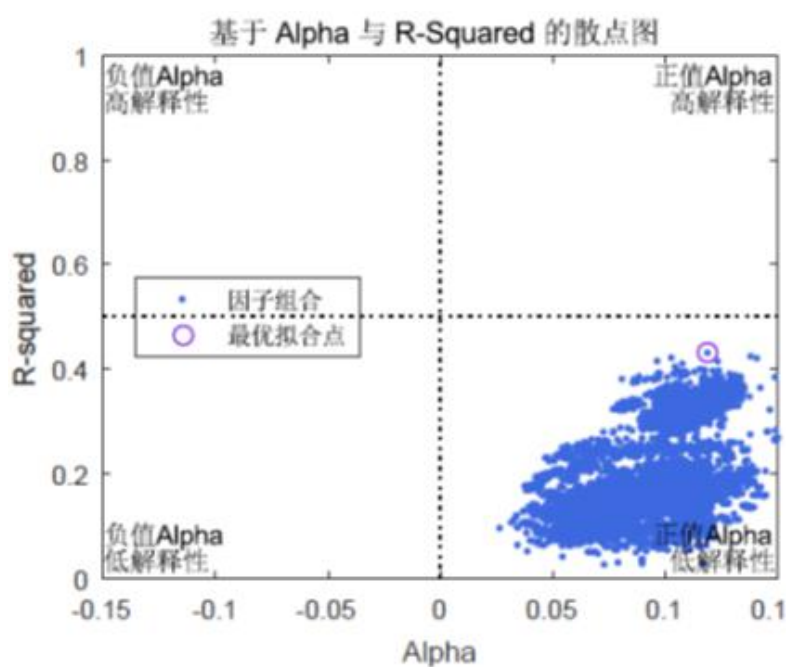
一种基于图像的分析逻辑是象限划分，将不同基金提纯而得的最终 Alpha 与 R-Squared（拟合优度）画成散点图，得到以下四个象限的最优拟合点：正值 Alpha/高解释性、负值 Alpha/高解释性、正值 Alpha/低解释性、负值 Alpha/低解释性。



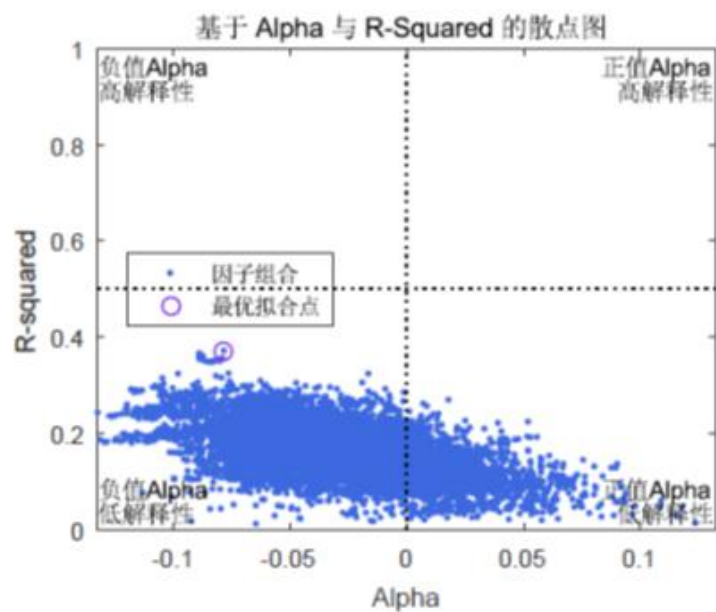
图片来源：海通证券金融工程



图片来源：海通证券金融工程

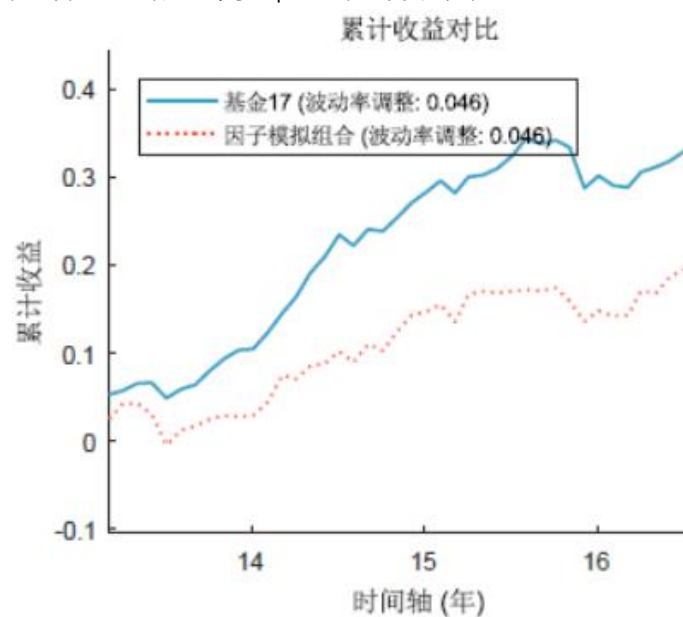


图片来源：海通证券金融工程

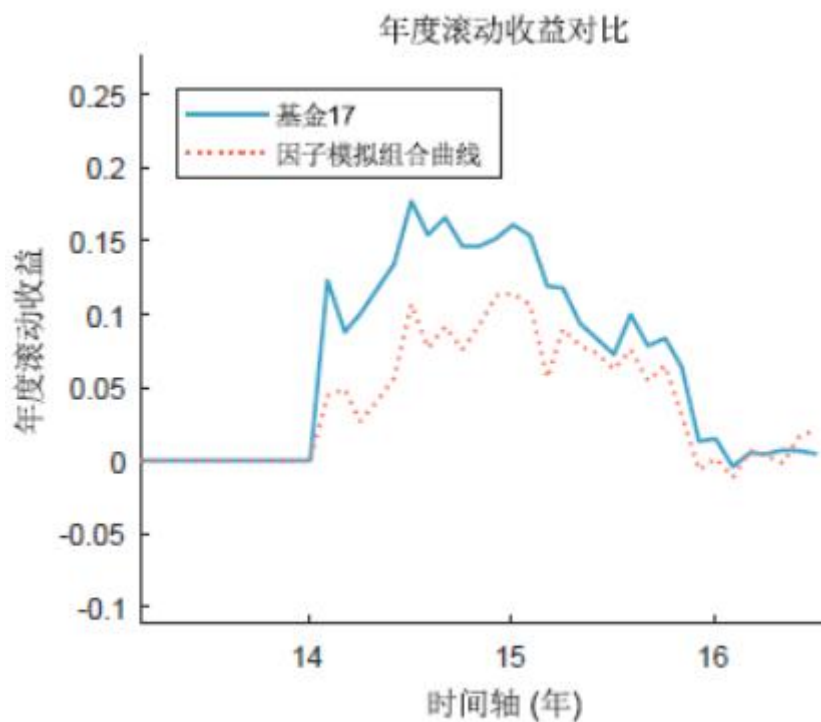


图片来源：海通证券金融工程

另一种基于图像的分析逻辑则是探究 Alpha 的可持续性，从而对净值曲线进行补充。

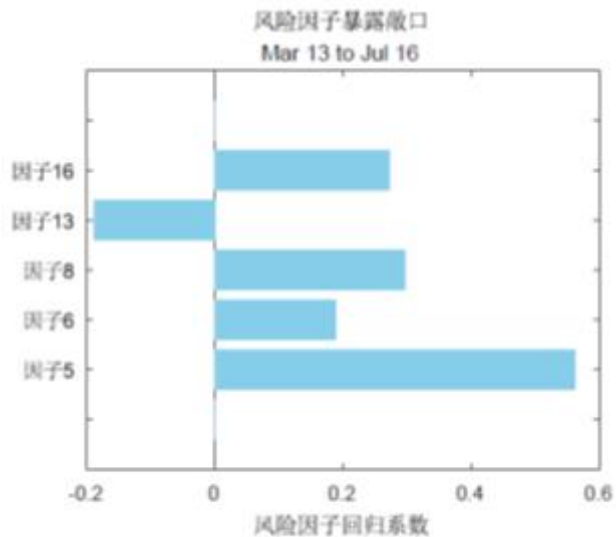


图片来源：海通证券金融工程

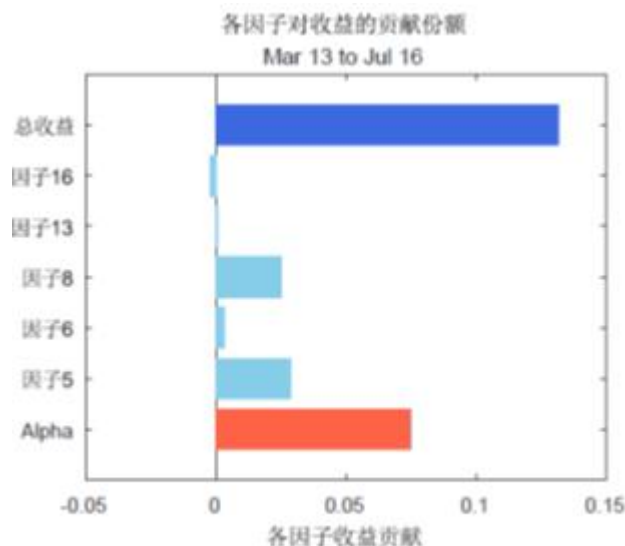


图片来源：海通证券金融工程

两条曲线之间的差就是提纯的 Alpha，从而可见其持续情况。
最后也是最常见的一种图像分析逻辑，将因子贡献程度分解。



图片来源：海通证券金融工程

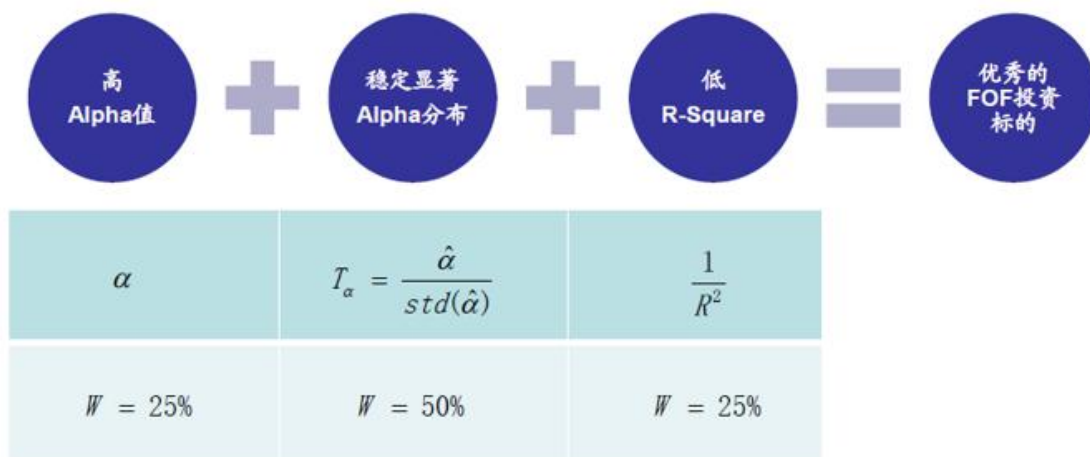


图片来源：海通证券金融工程

对于中国市场的一些基金来说，投资者必须明确自己在为什么风险买单。根据基于因子剥离的 FOF 择基逻辑，可以使用一些 ETF 来替代获取高透明度与稳定性，在监管和政策可行的范围内，使用期货衍生品实现做空，同时规避风险漂移，实现资产配置的高效化，为风险暴露敞口的控制提供参考。

五．总结

总的来说，基于因子剥离的 FOF 择基是一种数值统计方法，其实就是选出了高 Alpha、且 Alpha 分布稳定、低 R-Square 的基金。



图片来源：海通证券金融工程

具体来说是根据 Alpha 的显著性指标 T-Alpha，通过 Alpha 的标准差将其标准化，得到的下一期预测估计会比 Alpha 本身更稳健。通过以上三个指标基于横截面 z-score 标准化后进行加权打分。

Manager	Alpha	T-alpha	R-square	Rank
Long short equities-D	7%	2.6	17%	1
Multi-strategy-B	7%	3.1	57%	2
Multi-strategy-C	6%	3.0	51%	3
Long short equities-G	5%	1.8	10%	4
Arbitrage-A	10%	2.0	48%	5
CTA/managed futures-A	6%	1.0	4%	6
Long short equities-E	2%	1.5	4%	7
Event driven-A	4%	2.9	44%	8
Macro-A	4%	1.1	14%	9
Fixed income-E	5%	3.0	76%	10
Long short equities-A	7%	1.0	58%	11
Others-A	1%	0.5	2%	12
Long short equities-F	5%	1.0	71%	13
Event driven-B	-1%	0.0	8%	14
Long short equities-C	2%	0.5	58%	14
Macro-B	2%	0.7	75%	16
Long short equities-H	0%	0.0	53%	17
Fixed income-F	-1%	-0.2	14%	18
Relative value-A	-1%	-0.2	35%	19
Distressed debt-A	-1%	-0.2	18%	20
Long short equities-B	0%	0.2	81%	20
Others-B	0%	0.1	71%	22
Fixed income-D	-5%	-0.7	46%	23
Multi-strategy-A	-11%	-1.1	21%	24

图片来源：海通证券金融工程

特别声明

以上乃本公众号于 2017/6/13 发布的【Alpha 提纯】一种基于因子剥离的 FOF 择基逻辑一文，关于其中图片、数据等，均出自于海通证券金融工程的研究：抽丝剥茧与 Alpha 提纯——一种基于因子剥离的 FOF 择基逻辑，作者高道德、吕丽颖等，查看原文请关注海通证券金工微信公众号。

免责声明

本文中所提及的基金及投顾公司、所用图片数据等仅为展示说明功能，不构成任何投资参考意见，亦不构成财务、法律、税务、投资咨询意见或其他意见，对任何因直接或间接使用本文所涉及的信息和内容或者据此进行投资所造成的一切后果或损失，私募云通不承担任何法律责任。