

人工智能

用五个指数复制股票型基金指数

分析师: 包赞 S1230518090006
baozan@stocke.com.cn TEL: 021-80108127

◆研究背景

关于主动权益基金能不能战胜指数, 国内外学界与业界已有定论。当然学界严谨的计量经济学论证, 我们业界要利用好, 不能暴殄天物。如何用? 作者认为既然主动权益基金没有显著超额 alpha, 数学上讲就是可以被指数线性表示, 我们就可以用一些列的指数去复制。当然为了投资的方便, 用的指数个数越少越好, 所以, 本文只用了 50ETF、500ETF、300ETF、创业板、中小板这五个 ETF 来复制中证股票型基金指数 (H11021)。

◆算法介绍

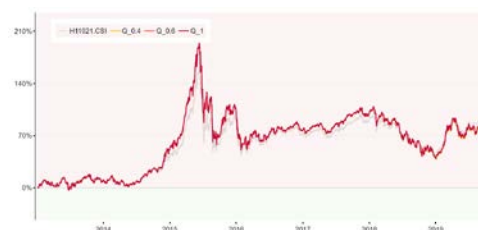
传统方法通过寻找定价因子, 计算目标组合 beta, 再通过优化技术让模仿组合 beta 和其相等, 以此达到复制的目的。流程较长, 每个流程都涉及不同的统计和优化技术。考虑复杂系统的不稳定性, 我们试图利用更直接的方法来复制, 设计一种优化算法, 直接利用基础资产的某种组合来跟踪目标组合的收益序列。事实证明, 在优化方法构造得当的情形下, 是可以利用较少股票组合来模仿目标组合的收益表现, 且大概率产生超额收益。需要特别说明的是, 文中介绍的方法, 计算效率高且无需日常维护。

◆回测展示

较长期的回测结果显示 (20130104-20190927), 在不同权重上限情景下, 季度调仓的经验跟踪误差目标函数组合 (etc) 都能较为稳定地跟踪目标指数。算法较为复杂, 为更好服务客户, 我们开发了网页端 APP:

<http://airione.org/shiny/HS300AI-BaoZan/>

20130104-20190927 以股票基金指数为目标的 etc 函数下模仿组合累计收益表现:



20190102-20190927 以股票基金指数为目标的 etc 函数下模仿组合累计收益表现:



正文目录

1. 引言	3
2. 回测	4
3. 算法	6

图表目录

图 1: 2013-2019 各指数累计收益对比	3
图 2: 20130104-20190927 以股票基金指数为目标的 ete 函数下模仿组合累计收益表现	4
图 3: 20190102-20190927 以股票基金指数为目标的 ete 函数下模仿组合累计收益表现	5
表 1: 2019 各指数收益情况对比	3
表 2: 2017-2019 各指数收益情况对比	3
表 3: 2013-2019 各指数收益情况对比	3
表 4: 用以跟踪股票基金指数的市场指数 ETF 概况	4
表 5: 20130104-20190927 以股票基金指数为目标的 ete 函数下模仿组合业绩表现	5
表 6: 20190102-20190927 以股票基金指数为目标的 ete 函数下模仿组合业绩表现	5

1. 引言

关于主动权益基金能不能战胜指数，国内外学界与业界已有定论。当然学界严谨的计量经济学论证，我们业界要利用好，不能暴殄天物。如何用？作者认为既然主动权益基金没有显著超额 alpha，数学上讲就是可以被指数线性表示，我们就可以用一些列的指数去复制。当然为了投资的方便，用的指数个数越少越好，所以，本文只用了 50ETF、500ETF、300ETF、创业板、中小板这五个 ETF 来复制中证股票型基金指数（H11021）。

下面简单展示下各指数的收益情况。

表 1：2019 各指数收益情况对比

	股票基金指数	上证 50	沪深 300	中证 500	中小板	创业板
期间收益率	34.07%	31.34%	30.38%	20.68%	30.07%	32.60%
年化夏普比	2.33	2.07	1.94	1.11	1.60	1.61
最大回撤	12.43%	11.72%	13.49%	21.65%	21.01%	20.32%

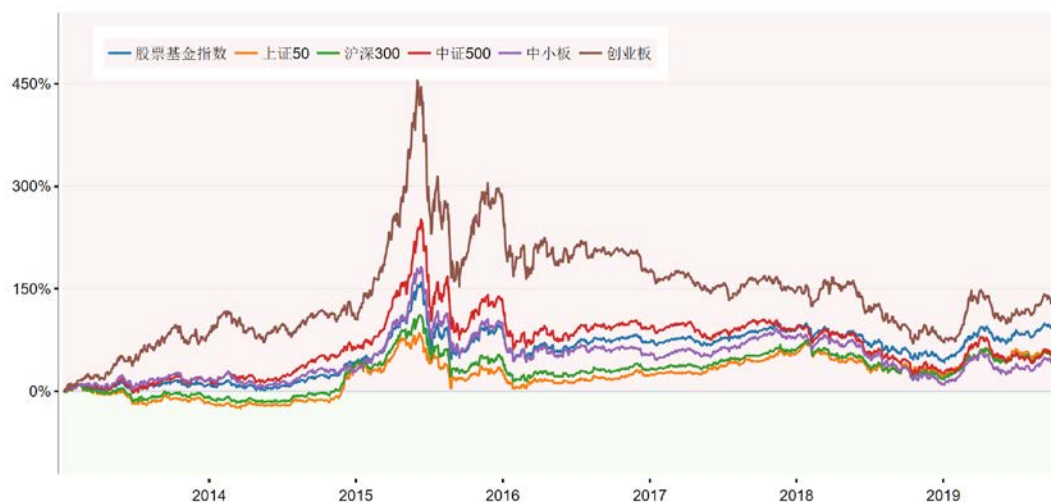
表 2：2017-2019 各指数收益情况对比

	股票基金指数	上证 50	沪深 300	中证 500	中小板	创业板
期间收益率	13.96%	31.70%	18.58%	-19.70%	-5.48%	-15.48%
年化收益率	4.98%	10.78%	6.54%	-7.83%	-2.07%	-6.06%
年化夏普比	0.28	0.58	0.36	-0.36	-0.09	-0.25
最大回撤	28.11%	28.87%	32.46%	40.11%	43.73%	39.49%

表 3：2013-2019 各指数收益情况对比

	股票基金指数	上证 50	沪深 300	中证 500	中小板	创业板
期间收益率	95.22%	62.12%	55.58%	53.54%	44.39%	132.29%
年化收益率	10.76%	7.66%	6.99%	6.77%	5.77%	13.75%
年化夏普比	0.46	0.31	0.29	0.25	0.21	0.43
最大回撤	45.11%	44.70%	46.70%	65.20%	61.75%	69.74%

图 1：2013-2019 各指数累计收益对比



资料来源：浙商证券研究所

2. 回测

用到的工具指数：

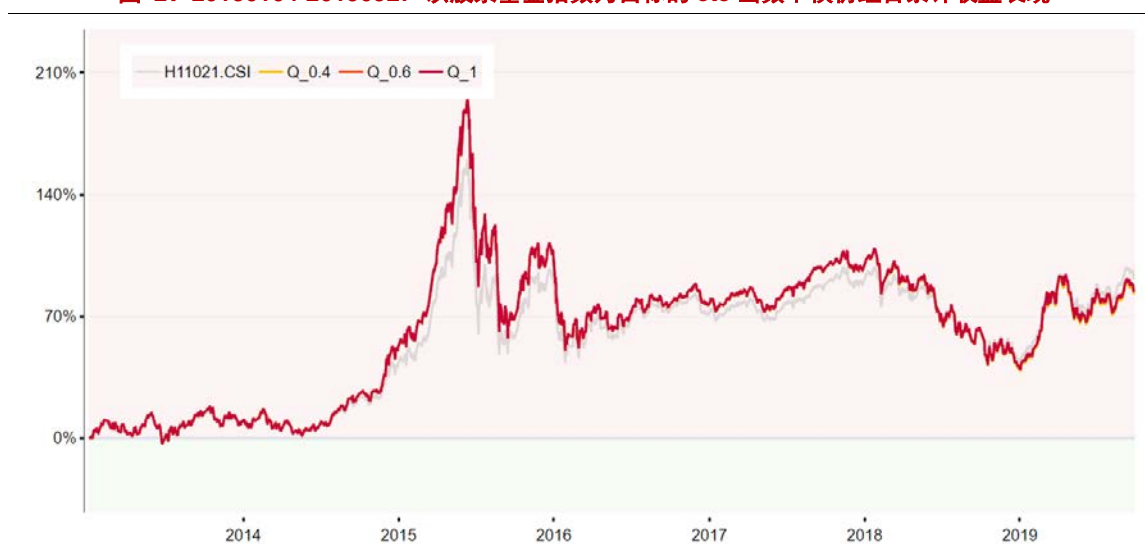
表 4：用以跟踪股票基金指数的市场指数 ETF 概况

etf 代码	etf 简称	跟踪指数代码	跟踪指数简称	规模（亿）	上市日期
510050.SH	50ETF	000016.SH	上证 50	492.59	2005-02-23
510500.SH	500ETF	000905.SH	中证 500	481.42	2013-03-15
510300.SH	300ETF	000300.SH	沪深 300	364.68	2012-05-28
159915.SZ	创业板	399006.SZ	创业板指	229.86	2011-12-09
159902.SZ	中小板	399005.SZ	中小板指	23.10	2006-09-05

选取经验跟踪误差最小化目标函数,对不同权重参数下组合的历史表现进行回测分析。模型计算为季度调仓,计算用到的历史时间序列区间长度一律设为 125 个交易日。图表中的 w 表示优化过程中组合权重的上限。

较长期的回测结果显示 (20130104-20190927), 在不同权重上限情景下, 季度调仓的经验跟踪误差目标函数组合 (ete) 都能较为稳定地跟踪目标指数, 见图 1。具体算法见下文。算法较为复杂, 为更好服务客户, 我们开发了网页端 APP: <http://airione.org/shiny/HS300AI-BaoZan/>

图 2：20130104-20190927 以股票基金指数为目标的 ete 函数下模仿组合累计收益表现



注：业绩基准为股票基金指数 (H11021)；Q_0.4 代表季度调仓、优化中的权重上限为 40%

表 5：20130104-20190927 以股票基金指数为目标的 ete 函数下模仿组合业绩表现

	H11021.CSI	w=40%	w=60%	w=100%
累计收益率	92.50%	83.74%	85.27%	85.27%
年化收益率	10.59%	9.80%	9.94%	9.94%
年化夏普比	0.45	0.40	0.40	0.40
标准差	1.47%	1.56%	1.56%	1.56%
最大回撤	45.11%	53.18%	52.91%	52.91%
跟踪误差	-	3.34%	3.30%	3.30%
月度胜率	-	53.09%	53.09%	53.09%

图 3：20190102-20190927 以股票基金指数为目标的 ete 函数下模仿组合累计收益表现



注：业绩基准为股票基金指数（H11021）；Q_0.4 代表季度调仓、优化中的权重上限为 40%

表 6：20190102-20190927 以股票基金指数为目标的 ete 函数下模仿组合业绩表现

	H11021.CSI	w=40%	w=60%	w=100%
累计收益率	32.21%	30.41%	30.56%	30.57%
年化夏普比	2.28	1.96	1.97	1.97
标准差	1.30%	1.42%	1.43%	1.43%
最大回撤	12.43%	14.28%	14.28%	14.28%
跟踪误差	-	3.42%	3.43%	3.44%
月度胜率	-	33.33%	33.33%	33.33%

3. 算法

具体见研报《“指数增强”新思维——人工智能+传统金融》，下文摘录主要部分。

传统方法通过寻找定价因子，计算目标组合 beta，再通过优化技术让模仿组合 beta 和其相等，以此达到复制的目的。流程较长，每个流程都涉及不同的统计和优化技术。考虑复杂系统的不稳定性，我们试图利用更直接的方法来复制，设计一种优化算法，直接利用基础资产的某种组合来跟踪目标组合的收益序列。事实证明，在优化方法构造得当的情形下，是可以利用较少股票组合来模仿目标组合的收益表现，且大概率产生超额收益。需要特别说明的是，下文介绍的方法，计算效率高且无需日常维护。

直接用基础资产去模仿目标资产，其实就是跟踪误差最小化，当然了，我们的目标不是样本内的最小化，而是模仿组合在未来的一个时期能够持续模仿目标资产并且能够产生超额收益，这就对算法技术提出了较高的要求，也是本文的创新点。

假设一个目标组合是一个指数，由 N 项资产组成。记 $\mathbf{r}^b = [r_1^b, \dots, r_T^b]^\top \in R^T$ ， $\mathbf{X} = [r_1, \dots, r_T]^\top \in R^T \times N$ 分别为该指数及 N 项资产过去 T 天的（算术）净收益率，其中 $\mathbf{r}_t \in R^N$ 为 N 项资产在第 t 天的净收益。

我们的目标是设计稀疏投资组合 $\mathbf{w} \in R_+^N$ ，满足 $\mathbf{w}^\top \mathbf{1} = 1$ ，以追踪指数，使得 $\mathbf{X}\mathbf{w} \approx \mathbf{r}^b$ 。相当于要解决的优化问题为：

$$\begin{aligned} & \underset{\mathbf{w}}{\text{minimize}} && \text{TE}(\mathbf{w}) + \lambda \|\mathbf{w}\|_0 \\ & \text{subject to} && \mathbf{w}^\top \mathbf{1} = 1 \\ & && 0 \leq \mathbf{w} \leq u\mathbf{1} \end{aligned}$$

其中 $\text{TE}(\mathbf{w})$ 代表一般跟踪误差， λ 为控制投资组合稀疏性的正则化参数，以及 u 为组合权重上限。

ℓ_0 范数由连续和可微（对于 $w \geq 0$ ）函数近似：

$$\rho_{p,u}(w) = \frac{\log(1+w/p)}{\log(1+u/p)}$$

其中 $p > 0$ 是控制估计的参数。于是转变成以下近似问题：

$$\begin{aligned} & \underset{\mathbf{w}}{\text{minimize}} && \text{TE}(\mathbf{w}) + \lambda \mathbf{1}^\top \boldsymbol{\rho}_{p,u}(\mathbf{w}) \\ & \text{subject to} && \mathbf{w}^\top \mathbf{1} = 1 \\ & && 0 \leq \mathbf{w} \leq u\mathbf{1} \end{aligned} \tag{1}$$

其中 $\boldsymbol{\rho}_{p,u}(\mathbf{w}) = [\rho_{p,u}(w_1), \dots, \rho_{p,u}(w_N)]^\top$ 。

有多种类型的目标函数可供选择，例如经验跟踪误差（ETE）、下行风险（DR）、Huber 经验跟踪误差（HETE）、Huber 下行风险（HDR）等。此处，我们选取经验跟踪误差（ETE）为目标函数：

$$\text{ETE}(\mathbf{w}) = \frac{1}{T} \|\mathbf{r}^b - \mathbf{X}\mathbf{w}\|_2^2$$

因为经验跟踪误差就是我们常规的指数复制技术，选取该目标函数主要是为了验证优化算法，如果这个目标函数下，模仿组合大概率产生超额收益，那么说明这套优化算法是合适的。

无论选择何种跟踪误差类型，问题（1）都可以通过一种迭代的闭合形式更新算法——优化最小化来求解（迭代次数用 k 表示）。可以看出，上述所有变化归结为以下凸问题的迭代优化：

$$\begin{aligned} \underset{\mathbf{w}}{\text{minimize}} \quad & \mathbf{w}^\top \mathbf{w} + \mathbf{q}^{(k)\top} \mathbf{w} \\ \text{subject to} \quad & \mathbf{w} \in \mathcal{W}_u \end{aligned}$$

其中 $\mathcal{W}_u = \{\mathbf{w} \mid \mathbf{w}^\top \mathbf{1} = 1, 0 \leq \mathbf{w} \leq u\mathbf{1}\}$, $\mathbf{q}^{(k)} \in R^N$

求解上述优化问题，用的是优化最小化方法。

股票投资评级说明

以报告日后的 6 个月内，证券相对于沪深 300 指数的涨跌幅为标准，定义如下：

- 1、买入：相对于沪深 300 指数表现 +20% 以上；
- 2、增持：相对于沪深 300 指数表现 +10% ~ +20%；
- 3、中性：相对于沪深 300 指数表现 -10% ~ +10% 之间波动；
- 4、减持：相对于沪深 300 指数表现 -10% 以下。

行业的投资评级：

以报告日后的 6 个月内，行业指数相对于沪深 300 指数的涨跌幅为标准，定义如下：

- 1、看好：行业指数相对于沪深 300 指数表现 +10% 以上；
- 2、中性：行业指数相对于沪深 300 指数表现 -10% ~ +10% 以上；
- 3、看淡：行业指数相对于沪深 300 指数表现 -10% 以下。

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重。

建议：投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者不应仅仅依靠投资评级来推断结论

法律声明及风险提示

本报告由浙商证券股份有限公司（已具备中国证监会批复的证券投资咨询业务资格，经营许可证编号为：Z39833000）制作。本报告中的信息均来源于我们认为可靠的已公开资料，但浙商证券股份有限公司及其关联机构（以下统称“本公司”）对这些信息的真实性、准确性及完整性不作任何保证，也不保证所包含的信息和建议不发生任何变更。本公司没有将变更的信息和建议向报告所有接收者进行更新的义务。

本报告仅供本公司的客户作参考之用。本公司不会因接收人收到本报告而视其为本公司的当然客户。

本报告仅反映报告作者的出具日的观点和判断，在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议，投资者应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求。对依据或者使用本报告所造成的一切后果，本公司及/或其关联人员均不承担任何法律责任。

本公司的交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。本公司没有将此意见及建议向报告所有接收者进行更新的义务。本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

本报告版权均归本公司所有，未经本公司事先书面授权，任何机构或个人不得以任何形式复制、发布、传播本报告的全部或部分内容。经授权刊载、转发本报告或者摘要的，应当注明本报告发布人和发布日期，并提示使用本报告的风险。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

浙商证券研究所

上海市杨高南路 729 号陆家嘴世纪金融广场 1 号楼 29 层

邮政编码：200120

电话：(8621)80108518

传真：(8621)80106010

浙商证券研究所：<http://research.stocke.com.cn>