

## 指数组合优化方法、模型与应用

作者：leson2006-09-02 17:58分类：默认分类标签：

指数组合优化方法、模型与应用

### 一、引言

随着证券价格指数的发展与演变,指数衍生品创新日益成为了当今全球金融市场上的一大亮点。指数基金、指数期货、指数期权、指数权证以及融合了众多基金产品优势的、被誉为第三代投资基金的交易所交易基金(ETFs)等在最近的几年中都得到了长足的发展。因此指数组合优化问题就变得日益重要,因为对指数基金、ETFs来讲,其产品本身实际上就是一个指数组合(只是其交易模式和产品架构有所不同),而对指数期货、指数期权等其它指数类金融产品来讲,通过构建指数组合来进行以对冲交易为主要目的的指数套利操作是规避金融风险的必不可少的操作手法。

指数组合优化问题从狭义上讲就是标的指数的优化复制(在本研究中指数组合优化与指数优化复制是一个概念)。指数复制的方法大体可以分为两类:即完全复制和不完全复制(即优化复制)。顾名思义,完全复制是购买标的指数中的所有成份证券,并且按照每种成份证券在标的指数中的权重确定购买的比例,以构建追踪组合从而达到复制指数的目的。而优化复制则是通过权重的优化再配置来寻找一个含有"部分"成份证券的最优的追踪组合,所谓的"最优"就是使得该组合相对标的指数的追踪误差最小或其它事先设定的标准最优。优化复制的方法又可以进一步细分为分层抽样复制(以下简称分层抽样)和优化抽样复制(以下简称优化抽样)两种。前者是两阶段优化法,即第一阶段是抽样,根据一定的标准选出样本证券;第二阶段则是权重的优化再配置,通过对最优化算法的应用,求出追踪组合内各样本证券的最优权重,使得组合的表现与标的指数一致,即追踪误差最小,同时保证较小的调整频率和追踪成本。与之不同,优化抽样复制则没有进行独立的抽样,而是让计算机来回答哪些成份证券该"购买",如果优化求解出来的某种成份证券的权重值为零,就剔除该成份证券;不为零,就买入该证券并按计算出来的权重值配置资产,因此优化抽样复制属于单阶段优化法。可见指数组合优化具体指的就是指数的优化抽样复制和分层抽样复制。(剖析主流资金真实目的,发现最佳获利机会!)

然而这两种方法具体如何应用,其在复制效果上又有何不同,包括与完全复制比有何不同,这是本研究所要探讨的问题。此外还要考虑个股以及组合整体的流动性问题,因为这是实务管理中一个非常重要问题,直接决定了指数复制的成败。因此本研究力图从国内证券市场的具体实际出发,并以上证180指数为标的对所有这些问题进行实证研究。

### 二、文献回顾

Roll(1992)在Markowitz建立的均值-方差模型的基础上较早研究了指数复制中的组合追踪误差最小化问题,其通过最小化组合收益率与标的指数收益率之差的平方和,即追踪误差波动率的最小化来改进组合对标的指数的复制效果。Y.Tabata和E. Takeda(1995)也在Markowitz的有效边界和均值-方差模型的基础上研究了指数基金的优化问题。其对追踪误差的定义与Roll的一致,不同的是Y.Tabata和E.Takeda在Markowitz的有效边界的基础上发展出了一套优化算法,并利用该算法进行了实证研究。与Roll、Y.Tabata和E.Takeda等人的研究不同,MarkusRudolf、Hans-Jurgen、Wolter和Heinz Zimmermann(1999)把追踪误差定义为组合收益率与标的指数收益率之间的差值,并因此提出了4个线性的追踪误差最小化模型,分别是均值绝对离差(MAD)模型、均值绝对向下偏差(MADD)模型、最小最大化(MinMax)模型、向下最小最大化(DMinMax)模型,并把以上4个线性模型应用于美国、日本、英国、德国、法国、瑞士六国的股票市场指数和摩根斯坦利资本国际指数(MSCI)的指数追踪实验,结果表明线性规划模型优于二次规划模型,把追踪误差定义为线性关系的组合收益率与标的指数收益率之间的差值也比传统的二次平方和的定义更接近现实。

Gunter Bamberg 和Niklas Wagner(2000)等人首先对线性回归法应用于最优追踪组合的求解问题进行研究,结果发现无论是使用复权调整后的股价还是未复权调整的股价以及使用收益率还是累积收益率建立的线性回归模型,都将违反古典线性回归最小二乘法的一些经典假设。然而Gunter Bamberg和Niklas Wagner认为这不影响回归方法在求解最优指数组合问题中的应用,他们采用鲁棒回归法来避开最小二乘法的假设限制,并应用德国DAX指数进行了实证检验。结果表明,当样本区间的市场指数波动较大时,应用鲁棒回归法来计算最优追踪组合效果较好。Francesco Corielli和Massimiliano Marcellino(2002)则基于动态因素模型考虑了指数成份股动态变化的特征,也通过最小化样本期内损失函数的方法来解决追踪组合的最优化问题。他们以EuroSTOXX50指数为样本做了

实证研究,在具体的算法上使用了蒙特卡洛模拟来求解最优的成份股权重。Dirk Eddelbuttel和Marseilles则用到了混合遗传算法来解决寻找最优的追踪组合问题。他们对德国DAX指数进行实证研究的结论表明,通过选择少量的成份股并通过追踪误差的最小化来寻找最优权重方面涉及到的复杂的算法问题完全可以由混合遗传算法加以解决。

此外,研究者还尝试使用其它一些更复杂的方法来进行建模和求解,如J.E. Beasley、N.Meade和T.J.Chang(2001)以及M. Gilli和E.K llezi(2001)等人采用启发式算法(HA)、Paolo Dai Pra、Wolfgang J.Runggaldier、和Marco Tolotti(2002)则利用随机控制和顺向优化法、David D. Yao、Shuzhong Zhang、和Xun Yu Zhou(2003)等人采用随机线性二次控制法(SLQ)并利用半限量规划(SDP)等方法求解。尽管这些方法并不常用,但其运用无疑使得指数复制问题的研究更为深入。可见随着最优化理论的发展和计算机技术的革新,更多、更复杂的新理论与新方法可以被应用到指数基金以及更广义上的指数类金融产品的设计和管理上来,这也正是当前国际上有关指数复制问题研究的一个热点和趋势。

当然研究者还把视角投向了指数复制中的其它一些问题,如David M.Walsh、Kathleen D.Walsh和John P.Evans(1998)利用澳大利亚股市指数数据,研究了追踪误差的估计误差问题。M.Ammann和Jurg Tobler(2000)把追踪误差分解为几个部分,并研究了不同的部分对投资决策的影响问题。Soosung Hwang和Stephen E.Satchell(2001)还研究了事先和事后的追踪误差问题,M. Ammann和H.Zimmermann(2001)则研究了追踪误差与资产配置决策的关系问题。此外,G.Connor和H.Leland(1995)等人还研究了指数复制中的现金拖累问题。

然而从以上文献的简要回顾中可以看出,研究者更多地主要精力放在了追踪误差、优化模型以及求解算法上,忽略了对不同复制方法和其不同复制效果的对比研究,更关键的是很少考虑个股以及组合整体的流动性问题,然而这些都非常重要而又实际,实务操作中无论是指数产品设计、指数套利还是实施指数化投资策略都必须首先决定采取哪种复制方法,而且都要考虑个股和组合的流动性问题。因此本研究力图在这些方面有所改进,具体是本研究将同时考虑三种复制方法:优化抽样、分层抽样和完全复制,并且除了应用序贯二次规划法外,还将用到遗传算法进行优化求解,同时对结果加以比较。此外还将基于市场微观结构理论的一些研究成果并利用高频成交数据对个股的流动性问题进行综合评价,并在抽样复制中,以个股流动性指标抽样,力图提高指数组合的整体流动性。进而在此基础上,对原模型进行改进,加入基于市场冲击成本考虑的流动性惩罚约束,旨在没有恶化追踪误差的情况下,让模型优化计算的结果能把更多的资金配置在更有流动性的资产上。

### 三、模型及抽样方法

本部分首先在总结国外研究文献的基础上提出了符合我国国内证券市场实际的指数组合优化模型,进而出于对个股和组合流动性改进的目的,提出了基于市场冲击成本考虑的流动性惩罚约束模型作为新的指数组合优化模型,使得在没有恶化追踪误差的情况下,让模型优化计算的结果能把更多的资金配置在更有流动性的资产上。

#### 1、一般模型

这是当今国外研究文献中常见的一种模型,所以称之为"一般模型",是与改进后的流动性惩罚约束模型相对应的一个概念。首先是目标函数。优化复制的思想是通过权重的优化再配置来寻找一个含有"部分"成份证券的最优的追踪组合,并使得该组合相对标的指数的一些考核标准最优。这里的考核标准一般有两个,一是追踪误差,二是超额收益,由于本研究中设定的目标是追踪指数,而不是战胜指数,因此只考虑追踪误差,不考虑超额收益。在Roll(1992)、Buckly和Korn(1998)、Cornnor和Leland(1995)以及Larsen和Resnick(1998)等人的文献中把追踪误差(简称TE)定义为 $R_{pt}$ 和 $R_{it}$ 之差的方差,Beasley、Meade和Chang(2001)不同意这种定义,如果 $R_{pt}-R_{it}$ 为一个常数,则其方差为零,则基于这种定义的追踪误差为零,显然这种定义忽略了这种情况下的 $R_{pt}$ 与 $R_{it}$ 之间的偏差,因此他们采用了误差平方和的均值平方根的定义。本研究中对追踪误差的定义与Beasley等人一致,即目标函数如下:

令 $R_{it}=\ln(I_t/I_{t-1})$ 和 $R_{it}=\ln(P_{it}/P_{i,t-1})$ ,经代换得:

其中: $T$ 是基于历史时期即样本内时间区间 $0,1,2,\dots,T$ 来考察股票与指数的价值变化, $T+1$ 代表一个决策时点,在该时点把当前组合调整为一个新的追踪组合; $P_{it}$ 代表 $t$ 时刻( $t=0,\dots,T$ )个股 $i$ ( $i=1,2,\dots,N$ )的价格, $\ln(P_{it}/P_{i,t-1})$ 即个股对数收益率 $R_{it}$ ; $P_i$ 代表追踪组合调整(或构建)时刻的个股当日收盘价; $I_t$ 代表 $t$ 时刻指数的点位, $\ln(I_t/I_{t-1})$ 即指数对数收益率 $R_{it}$ ; $q_i$ 为决策变量(即待求变量),代表 $T+1$ 时点个股 $i$ 的持有数量。

需要考虑的约束条件有:

1)组合内个股集中度,即组合内单只股票的持有数量不得超过一定的比例,也不能小于一定的比例:

其 $\alpha_i$ 中为追踪组合中个股i的最小持有比例,实践中 $\alpha_i$ 代表个股i的"最小持有水平"; $\beta_i$ 为追踪组合中个股i的最大持有比例,则有 $0 \leq \alpha_i \leq \beta_i \leq 1$ , $\beta_i$ 限制了追踪组合中个股i的风险敞口;

2)资金规模,即持有的追踪组合市值加上仓位调整成本等于初始资金规模:

其中 $q_i$ 为组合调整之前的个股持有量,创设问题中 $q_i$ 为零, $\gamma$ 为交易中涉及到的佣金、印花税等费率的综合比例,而

为仓位调整所带来的交易成本; $V$ 为初始资金规模,其值等于每次调整前一日收盘后的追踪组合市值与因指数调整而被调出的个股悉数卖出而带来的交易成本的差额。

3)交易成本,即仓位调整所带来的交易成本不能超过初始资金规模的一定比例。因指数调整而被调出的个股需悉数抛出,因此而带来的交易成本是个给定值,无法优化。能够加以考虑的是因指数调整而被调入的个股以及保留在指数内的个股的仓位变动所带来的交易成本,其不能超过初始资金规模的一定比例,此中为 $c$ 。

追踪组合的初次建仓即创设问题中,由于 $q_i=0$ ,因此通过定义权重变量

而把上述模型变换为:

目标函数:

约束条件:

1) $\alpha_i \leq w_i \leq \beta_i$ ;

2)但要增加约束

3)由于决策变量为 $w_i$ ,所以不必考虑规模的约束;

4)由于创设时(初次建仓),交易成本是规模的一定比例的一个常数,所以对交易成本无需加以约束,原约束3取消;

## 2、流动性惩罚约束模型

除了进一步明确追踪组合中个股i的限制比例 $\alpha_i$ 和 $\beta_i$ ,并对交易成本约束赋予控制仓位调整数量的内涵外,关键的是在上述模型中加入一个基于市场冲击成本考虑的流动性惩罚约束,即:

$\Delta_i$ 就是个股的流动性惩罚因子, $\Delta_i | q_i - q_i^0 |$ 是个股i仓位调整所引起的惩罚值,总的惩罚值不能超过初始资金规模的一定比例,此中为 $d$ 。 $\Delta_i$ 的取值则由下式给定:

其中 $(\Delta A/A)$ 为市场成交金额的变化比例, $e$ 为市场冲击成本关于成交金额的平均弹性,因此 $(\Delta A/A) \times e$ 为市场冲击成本相对其自身的平均变化比例。把市场冲击成本占成交金额变化额的比例记为 $1$ (或为任意正小数,比如 $\eta$ ,但因为 $d$ 的取值与 $\eta$ 有关,在这个约束的不等式两边可以同时约去 $\eta$ ,因此这里可以记为 $1$ ),则 $1 \times (\Delta A/A) \times e$ 为市场冲击成本占成交金额变化额的平均比例,这个平均比例再根据个股不同的流动性综合评价价值进行线性分摊就得到个股不同的市场冲击成本占其成交金额变化额的比例,这就是 $\Delta_i$ 的经济含义。

在一般模型中加入流动性惩罚约束,则模型分别变为以下模型(1)和(2):

------(1)

创设问题的模型经变换为:

------(2)

优化抽样复制和分层抽样复制所不同的是前者让计算机自动抽样,后者则在优化计算之前先进行抽样,但不论怎样两者在以上优化模型上大致是一致的。限于篇幅,两者的区别就不赘述。

### 3、完全复制问题

为了比较,本文对完全复制问题也一并加以研究。而撇开实务中的动态调整不谈,完全复制中追踪组合的构建方法就是组合调整时,买入标的指数的所有成份证券,并按照每种成份证券该时刻在标的指数中的权重值来确定每种成份证券的购买比例。由于考虑到资金规模的约束,因此必须求解如下非线性方程组:

------(3)

因创设时交易成本是规模的一定比例,故直接计算得:

### 4、抽样方法

即在分层抽样复制中设计一套抽样方法来选出比较有流动性的股票,从而达到提高个股与组合整体流动性的目的。因此首先得衡量出个股的流动性,然后根据指数复制的其它要求设计抽样方法从而既要保证个股的流动性,又能很好地复制指数。本研究中将基于市场微观结构理论的一些研究成果并利用高频成交数据对个股的流动性问题进行综合评价,评价指标包括个股的市场宽度、市场深度以及市场弹性等三个方面,具体计算公式如下:

1)市场宽度:即交易价格偏离市场有效价格的程度,它是投资者支付的流动性升水。市场宽度的观测值是做市商的买卖差额,只有当这个差额为零时,才可认为这个市场完全达到流动性。由于我国内地的证券市场是指令驱动市场,没有做市商的买卖报价,因此对市场宽度的衡量需要基于内地证券市场的实际情况进行变通,本研究中对此界定如下:

$$W_{\alpha, \Delta} = 2 \times (H_{t, \Delta} - L_{t, \Delta}) / ((H_{t, \Delta} + L_{t, \Delta}))$$

即 $\Delta$ 时间间隔的市场宽度定义,其中H、L分别为达到换手率 $\alpha$ 的最高价与最低价,t和 $\Delta$ 分别指在 $\Delta$ 时间间隔内首次达到 $\alpha$ 换手率的时间;

2)市场深度:即在不影响当前价格下的成交量。从市场深度来看,流动性好意味着在当前价格下投资者可以大量买入或抛出股票。市场深度的观测值是某一时刻做市商在委托簿报出的委托数量,本研究中对此界定如下: $\Delta$ 时间间隔的深度 $D_{c, \Delta}$ ,是指自t时始到股票价格首次超越区间 $[P_t - c, P_t + c]$ 时的换手率,c为价格波动幅度;

3)市场弹性:即由交易引起的价格波动消失的速度,弹性通过时间来表示,它表示价格偏离均价后首次返回所需要的时间,本研究中对此界定如下:

$\Delta$ 时间间隔的弹性 $R_{c, \Delta}$ 是指自t起在 $\Delta$ 时间间隔内,股票价格偏离 $P_t$ 的幅度为c并首次返回 $P_t$ 的时间。

之后则利用距离评价法对这三个指标进行综合评价,得出个股的流动性综合指标,再在抽样复制中,以个股流动性综合指标和调整流通市值指标进行综合排序并综合考虑行业的分散性和代表性等因素加以抽样,从而力图提高指数组合的整体流动性。

#### 四、实证计算

本部分从内地证券市场的具体实际出发,并以上证180指数为标的进行实证研究,对如何设定指数追踪目标、如何限定约束条件、如何进行数学计算以及对三种复制方法的实际复制效果上的不同以及如何改进个股和组合的流动性水平等给出实证结论。

##### 1、样本选择和参数设定

本研究中以上证180指数为标的指数,根据180只成份股每次调整前半年的历史成交数据并同时考虑资金规模、交易成本、组合分散化和流动性惩罚约束等条件后来构建追踪组合,并且一直持有追踪组合直到上证180指数发生调整时,追踪组合也相应进行调整。优化抽样中直接把180只股票的数据代入模型进行优化计算,而分层抽样中则最终是抽样150只股票,再进行优化计算。然后进行资产配置和组合模拟,看看追踪组合在之后的5个交易日、10个交易日、30个交易日、60个交易日和截至下次指数调整日期间相对于上证180指数的表现。

此外模型中的一些参数设定如下,即最初的资金规模为10亿元,交易成本(包括佣金、印花税等的综合)为交易市值的0.003(即模型中的 $\gamma$ 值),优化复制中 $c$ 的取值为0.0005,抽样复制中 $c$ 的取值为0.0004,即仓位调整所带来的交易成本分别不超过初始资金规模的0.0005和0.0004,流动性惩罚因子 $\Delta i$ 以及 $d$ 的取值见上述模型的说明,权重上界 $\beta_i$ 为0.1,下界 $\alpha_i$ 在优化抽样复制中为零,在分层抽样复制中则限定其取值以保证每股最少的持有数量,限于篇幅 $\alpha_i$ 值的设定及模型推导就不赘述。

##### 2、数据来源与数据处理

本研究中的数据来源于上海证券交易所提供的日成交数据和1分钟成交数据。由于上证180指数对成份股的现金分红不做调整,即让价格自然回落,对指数不做修正,因此本研究中在做复权处理时也对个股的现金分红不做调整,即计算未考虑现金分红的复权日交易数据。在此基础上,再计算指数和个股的日收益率数据。1分钟成交数据则用于抽样复制中计算个股的流动性指标。在具体计算上,本研究所有涉及到的优化计算是基于Matlab6.5优化工具箱的编程技术进行的,完全复制中则基于Matlab6.5编程来求解非线性方程组,抽样复制中的抽样工作则是在SAS8.2环境下进行的。优化计算中用到的算法的基本思想是用于解决非线性约束下的非线性优化问题的序贯二次规划法(SQP),可以选择的参数包括标准算法与大规模算法、Levenberg-Marquardt搜索法与Gauss-Newton搜索法等;SQP的思想以及相关参数的含义参见Matlab6.5的帮助文件和有关优化计算的文献资料。限于篇幅,数据处理过程以及优化计算编程在此也不赘述。

#### 五、算法改进

本研究中除了用到序贯二次规划法(SQP)外,还将用到遗传算法(GA)进行优化求解,一是GA算法通常情况下比SQP算法较有可能找到全局最优解;二是通过不同算法的应用,比较其在优化求解,尤其是本课题的优化求解中的具体应用和优化效果上有何不同,探讨应用GA来改进指数组合优化效果的可行性。

遗传算法最早由John Holland于1975年提出来,其模拟达尔文的遗传选择和自然淘汰的生物进化过程,以一种群体中的所有个体为对象,并利用随机化技术指导对一个被编码的参数空间进行高效的搜索,因而遗传算法是一种基于生物自然选择与遗传机理的随机搜索算法。与传统搜索算法不同,遗传算法从一组随机产生的称为"种群"的初始解开始搜索过程。种群中的每一个个体是问题的一个解,称为"染色体"。染色体是一串符号,比如一个二进制字符串。这些染色体在后续迭代中不断进化,称为"遗传"。在每一代中用"适应度函数"来测量染色体的好坏,生成的下一代染色体称为"后代"。后代是由前一代染色体通过"交叉"或"变异"运算形成的。在新的一代形成过程中,根据适应度的大小选择部分后代,淘汰部分后代,从而保持种群大小是个常数。适应度高的染色体被选中的概率较高。这样经过若干代进化之后,算法收敛于最好的染色体上,这个最好的染色体就是待求问题的最优解或次优解。

应用MATLAB语言编程而成的遗传算法工具箱比较多,比较常用的有Sheffield教授于1994年编写的1.2版本的遗传算法工具箱、还有C.R.Houck、J.A.Joines以及M.G. Kay等人于1998年编写的基于MATLAB5.x的遗传算法工具箱等。最新版本的MATLAB7.0(Release 14)已经加载有遗传算法与直接搜索工具箱模块。由于MATLAB7.0中

的遗传算法与直接搜索工具箱模块比较适用于解决无约束优化问题,对非线性约束的非线性规划问题的求解效率很低。本研究在综合比较了各工具箱的优劣之后,选择了C.R.Houck、J.A.Joines以及M.G.Kay等人的遗传算法工具箱来解决本课题中的优化求解问题。

注:上海证券交易所副总经理刘啸东博士对本文亦有贡献,特此致谢!

## 主要研究结论

本研究从2002年7月1日上证180指数首次发布之日起进行跟踪,分别考查了四个组合(初次建仓以及随后的三次调整)的追踪效果,并把流动性惩罚约束模型下的计算结果与一般模型下的计算结果进行比较,此外还应用遗传算法对流动性惩罚约束模型进行求解,主要结论如下:

1、优化抽样和分层抽样无论在样本内还是在样本外的追踪效果都相当不错。样本内时间区间为半年且按照最小交易手数取整的情况下,优化抽样和分层抽样的样本内日均追踪误差在 $6.6128e-5$ 至 $1.9053e-4$ 之间,样本内半年累积追踪误差在0.0069至0.0205之间,指数与追踪组合收益率相关系数在0.9888至0.9995之间;样本外5日、10日、30日、60日以及截至下次调整日的日均追踪误差在 $1.0828e-4$ 至 $7.5426e-4$ 之间,累积追踪误差在0.0022至0.0196之间,相关系数在0.9874至0.9999之间。

2、从复制精度看,完全复制最高,其次是优化抽样,最后是分层抽样。完全复制的追踪误差最小,优化抽样其次,分层抽样最大,而且样本外时间越长,完全复制在追踪误差上的优势越明显。而分层抽样由于剔除了30只"问题证券"从而提高了追踪组合的质量,但却以损害复制精度为代价,其相对于优化抽样和完全复制来讲,追踪误差加大。可见样本股数目与追踪误差之间存在着一种此涨彼伏、相互抵消的关系。

3、从交易成本看,分层抽样最小,其次是优化抽样,最后是完全复制。同样10亿元规模的最初资金,整个追踪下来,分层抽样的交易总成本为4,461,460元,优化抽样为4,501,820元,而完全复制为4,943,540元,完全复制分别比分层抽样和优化抽样高出482,080元和441,720元。可见完全复制的高精度是以较大的交易成本为代价的。而分层抽样剔除了30只"问题证券",损害了复制精度但却减少了追踪成本。可见追踪误差和追踪成本之间也存在着相互抵消的关系。

4、其它条件不变的情况下,样本内的时间窗越长,追踪效果越好。第三次调整中由于上证180指数纳入2003年11月18日才上市的长江电力(600900)为成分股,因此样本内的时间统一从2003年11月18日起算至2004年1月2日,只有32个交易日,时间短,样本内的追踪误差就大,并进而影响到样本外的追踪效果。可见,其它条件不变的情况下,样本内的时间窗越长,追踪效果越好,而且样本内追踪效果越好,样本外的表现也越好。

5、资产配置时由于最小交易手数的限制而进行取整,从现金差额看,其影响几乎可以忽略不计。平均现金差额在两三千元左右,占总资金规模的百万分之三左右。而取整给样本内追踪误差带来的影响尽管因时间段的不同而有所差异,但多数情况下取整加大了样本内的追踪误差,加大的幅度在创设问题中是10%左右,在调整问题中是3%左右。

6、从追踪组合中个股的集中度看,分层抽样较高,优化抽样其次,而完全复制最分散。不过无论哪种方法,个股占组合总市值的比例都在可以接受的范围之内。分层抽样中由于考虑了行业的分散性和代表性,因此尽管个股集中度略高一点,也不影响组合的分散化。

7、从仓位变动对市场的冲击成本看,尽管没有直接度量,但从个股仓位调整占该股流通股比例的描述性统计中可以大致了解一点。创设时,这个比例在千分之二左右,调整时则在万分之四左右,而且由于允许分阶段建仓,分层抽样中还考虑了个股的流动性,因此冲击成本应该比较小。

8、尽管加入流动性惩罚约束之后,组合的追踪误差有了一定的变化,但组合的总体流动性明显提高了。测算结果显示,相对于一般模型的结果,流动性惩罚约束模型情况下资金更多地配置于更有流动性的资产上。可见,流动性惩罚约束模型达到了预期的效果,即在没有严重恶化组合追踪误差的情况下,使得模型优化计算的结果能把更多的资金配置在比较有流动性的资产上。

9、尽管遗传算法(GA)在序贯二次规划法(SQP)计算结果的基础上都有所改进,但改进的幅度非常小。而且并没有比单纯应用SQP求解的结果好。可见SQP算法已经比较好地找到了模型的最优解,GA算法能够改进的余地很小。