

## 证券研究报告—深度报告

## 金融工程

## 量化投资

## 数量化投资系列报告之四十三

2011年8月31日

相关研究报告:

专题报告

## 基于最优化方法的指数管理策略

证券分析师: 董艺婷

电话: 021-68866946

E-mail: dongyt@guosen.com.cn

证券投资咨询执业资格证书编码: S0980510120055

联系人: 毛甜

电话: 021-60933154

E-mail: maotian@guosen.com.cn

证券分析师: 葛新元

电话: 0755-82133332

E-mail: gexy@guosen.com.cn

证券投资咨询执业资格证书编码: S0980510120058

## 1. 数量投资在被动投资中的运用——最优化方法

伴随着日益增长的指数创新产品的需求, 指数投资的数量化管理技术变得尤为重要。本篇报告将介绍如何把最优化方法用于指数投资。在最优化方法下被动投资领域中的指数跟踪、成本管理、组合再平衡、业绩评价, 都转化成了一个约束条件下求解最优解的问题, 使得投资决策更为直接, 更为精细化。

## 2. 最优化抽样复制方法

从国内被动投资业务来看, 最优化方法能解决跨市场指数复制、大宽基指数复制、成份股替代这三种类型的问题。这些问题转化为优化模型, 即为求解目标函数为 Min (跟踪误差)、Max (信息比) 等, 约束条件为抽样股数、卖空、行业分布、冲击成本等的最优化问题。

## 3. 案例 1: 跨市场指数抽样复制

案例构想: A 股市场划分为上海和深圳两个市场, 这成为设置跨市场 ETF 的障碍, 那么是否能用上海市场的股票抽样复制跨市场的指数, 如中证 100, 中证 500? 结果分析及思考: 以信息比为优化目标的优化方法显然在收益上能战胜指数, 同时跟踪误差放大程度有限, 因此比较适用于样本量较大、抽样比例小的指数抽样, 例如, 富国上证综指 ETF。

## 4. 案例 2: 大宽基指数抽样复制

案例构想: 类似富国上证综指 ETF 的指数基金, 跟踪的基准指数的成份股数量庞大 (上证综合指数有 900 只以上的成份股), 这种类型的指数复制将非常适合用最优化的方法, 并以信息比作为优化目标, 既能保证跟上指数, 又能一定程度上获取 Alpha。

## 5. 案例 3: 托管行股票替代

利用最优化方法替代限购股, 从操作层面看, 投资逻辑明确, 可投资性比较强。首先, 在该成份股的行业内选择替代股, 风险也较小。其次, 替代权重累加到原有的基准权重中, 交易、换仓都非常方便, 不增加额外的交易成本。另外, 由于底仓的存在, 替代时可接受卖空权重。

## 6. 总结及后续研究

最优化方法属于数值计算方法, 能帮助基金经理更精细化、模型化的管理量化基金和被动基金。从国外经验和数量模型产品市场来看, 优化器软件是常用的商业软件, 如 (BARRA Optimizer, Axioma 优化器, Risk Matrixs), 它能广泛的解决投资决策模型中的大量求解问题。作为数量分析师, 利用通用的软件工具如 SAS、Matlab 等内置的优化算法, 结合数据处理, 并将之模块化, 能够完美的、灵活的实现优化器功能。后续研究方向如下:

将最优化方法引入指数基金管理和数量基金管理: 将优化技术与多因子模型, 收益率预测模型结合, 开发抽样复制技术、数量指数增强策略, 130/30 指数, 数量基金等。将最优化方法引入行业配置: 意在寻找有效的行业组合。定期跟踪成份股替代, 并发布替代结果。

## 独立性声明:

作者保证报告所采用的数据均来自合规渠道, 分析逻辑基于本人的职业理解, 通过合理判断并得出结论, 力求客观、公正, 结论不受任何第三方的授意、影响, 特此声明。

## 内容目录

引言 .....	3
最优化理论介绍 .....	3
数量技术在被动投资中的运用——最优化方法 .....	4
指数跟踪方法介绍 .....	4
基于最优化方法的指数复制 .....	5
目标函数 .....	5
约束条件 .....	6
指数跟踪优化模型 .....	7
案例分析 1——跨市场指数的抽样复制 .....	8
案例构想: .....	8
实证设置 .....	8
实证结果 .....	9
结果分析及思考 .....	11
案例分析 2——大宽基指数的抽样复制 .....	12
基于最优化方法的成分替代法 .....	12
案例分析 3——托管行股票替代 .....	13
实证设置 .....	13
实证结果 .....	14
结果分析及思考 .....	15
总结及后续研究 .....	15
附录: 共轭梯度法介绍 .....	17

## 图表目录

表 1: 中证 100 抽样复制结果 .....	9
图 1: 中证 100 最优化抽样_跟踪误差最小化_30 .....	10
图 2: 中证 100 最优化抽样_信息比最大化_30 .....	10
图 3: 中证 100 抽样_市值加权 .....	10
图 4: 中证 100 抽样_行业分层市值加权 .....	10
表 2: 中证 500 抽样复制结果 .....	10
图 5: 中证 500 最优化抽样_跟踪误差最小化_30 .....	11
图 6: 中证 500 最优化抽样_信息比最大化_20 .....	11
图 7: 中证 500 抽样_市值加权 .....	11
图 8: 中证 500 抽样_行业分层市值加权 .....	11
表 3: 工商银行 (601398.SH) 替代结果 .....	14
表 4: 建设银行 (601939.SH) 替代结果 .....	14
图 9: 工商银行 (601398.SH) 替代效果 .....	15
图 10: 建设银行(601939.SH)替代效果 .....	15

## 引言

伴随着日益增长的指数创新产品的需求，指数投资的数量化管理技术变得尤为重要——指数基金管理，包括指数编制、指数跟踪、成本管理、指数增强、组合再平衡、指数跟踪的业绩评价，也将走向精细化、结构化、数量化的趋势。

本篇报告将介绍如何把最优化方法用于指数投资，即基于最优化方法的指数投资管理策略。

## 最优化理论介绍

最优化方法是以数量分析为基础，以寻找具有确定的资源、技术约束的系统最大限度地满足特定活动目标要求的方案为目的，帮助决策者或决策计算机构对其所控制的活动进行实现优化决策的运筹学方法。最优化方法被人们广泛地应用到了工程技术、工农业生产、经济管理、国防等各个领域。

在数学上，所谓最优化问题，就是求解如下形式的最优解：

$$\begin{aligned} \min f(x) \\ \text{s.t. } x \in X \end{aligned}$$

其中  $f(x)$  称为目标函数， $x \in R^n$  是决策变量， $X$  为约束集，如果约束集  $X = R^n$ ，则最优化问题称为无约束最优化问题。约束最优化问题通常写为

$$\begin{aligned} \min f(x) \\ \text{s.t. } c_i(x) = 0, i \in E, \\ c_i(x) \geq 0, i \in I \end{aligned}$$

这里  $E$  和  $I$  分别是等式约束的指标集和不等式约束的指标集， $c_i(x)$  是约束函数。从数学意义上说，最优化方法是一种求极值的方法，即在一组约束为等式或不等式的条件下，使系统的目标函数达到极值，即最大值或最小值。

在优化问题中，根据决策变量、目标函数和约束函数的不同，最优化分成了线性规划、整数规划、二次规划、非线性规划、随机规划、动态规划、多目标规划等若干分支。

最优化方法通常采用迭代方法求解它的最优解，其基本思想是：给定一个初始点  $x_0$ ，按照某一迭代规则产生一个点列  $\{x_k\}$ ，使得当  $\{x_k\}$  是有穷点列时，其最后一个点是最优化模型问题的最优解  $x^*$ ，当是无穷点列是，它有极限点，且其极限点是最优化模型的最优解，一个好的算法应具备的典型特征为：迭代点能稳定地接近局部极小点的邻域，然后迅速收敛于  $x^*$ 。当给定的某种收敛准则满足时，迭代即终止。

## 数量技术在被动投资中的运用-----最优化方法

近年来，数量化投资的思想和方法越来越受到重视，已经发展成为一种新的投资趋势，逐渐深入到了主动投资和被动投资的各个环节。在各种数量化投资方法中，最优化方法非常重要，特别是在被动投资领域，是一项核心的技术。

在最优化方法下，被动投资领域中的指数的跟踪、指数跟踪过程中成本管理(包括交易成本、管理费用和现金成本等)、指数跟踪组合的再平衡、指数跟踪的业绩评价这些投资流程，都转化成了一个约束条件下求解最优解的问题，使得投资决策更为直接更为精细化。

### 指数跟踪方法介绍

指数管理技术的关键是指数的跟踪，即构造一个能够最大程度拟合指数走势的投资组合，也即指数复制。

常用的指数复制方法：完全复制法、抽样复制法（包含最优化复制法），我们下面逐一简要介绍。

**完全复制法。**完全复制法是在构建跟踪组合的时候，将目标指数中包含的所有成份股全部作为一个跟踪组合，并以指数编制中自然权重作为跟踪组合中每个股票的权重。全复制能够使得跟踪组合与指数具有最优的拟合，适用于成份股数量少且流动性高的指数。当成份股较多，特别是存在一些流动性较差的成份股时，就会产生零股、流动性不足、调整频率高、冲击成本大等等这些影响跟踪误差的问题。

**抽样复制法。**抽样复制法是指从指数的成份股中抽取若干股票构成用于复制指数跟踪组合的方法。常用的抽样复制法有：最大市值占比法、最大相关系数法、分层抽样法（如行业分层抽样、聚类分层抽样）。相对全复制，抽样在控制跟踪误差的基础上进一步减少了组合维护及再平衡的所需的费用，具有较强的可操作性。

抽样复制的一个缺点是无法保证由它抽样出来的样本股能够在所有影响股价波动特性的属性上精确的均匀覆盖，因而同样存在其它类型非系统性风险的可能。正是因为这一缺点，通常在采用抽样复制的时候会结合最优化的技术，将抽样复制分为两个阶段，第一阶段是抽样，第二阶段是优化组合权重，即优化抽样复制。

**最优化复制法。**最优化复制法是将构造组合时需要考虑的各种因素转化成完整的目标规划问题，通过各种数值方法求解出最优组合。其中，目标函数诸如跟踪误差最小化，或保证跟踪误差基础上的超额收益最大化，而约束条件则包括组合成份股数量，成份股是否可以卖空，成份股的行业分布，仓位调整成本和冲击成本等等。

由于优化抽样复制和最优化复制均采用的是最优化方法，我们这里把优化抽样复制和通常意义下的最优化复制，统称为基于最优化方法的指数复制法。相对于全复制、普通抽样复制，基于最优化方法的指数复制法更为复杂，也更为精细化。

从国外指数跟踪管理方法的研究状况来看，基于优化复制方法的研究占绝大部分，这主要是因为国外股票市场指数的成份股数量通常都非常大，因此在指数跟踪管理过程中采用最优化的方法，选用指数全部成份股的一个子集构造指数跟踪组合成为一种非常现实的选择。

## 基于最优化方法的指数复制

指数跟踪的过程，实际上是一个追求跟踪误差最小化、超额收益最大化以及二者的平衡的过程。

跟踪误差是跟踪组合与目标指数收益之间的差异，超额收益是一定时间内收益偏差的累积效应，这两项指标度量了指数管理者对跟踪组合的成份股的选择能力、跟踪组合的成分证券调整能力以及对各种约束的管理能力，如风险管理能力和成本管理能力等。

下面我们要构建的指数跟踪优化模型中，设定同时考虑跟踪误差和超额收益两个要素的指数跟踪目标函数，并考虑指数跟踪过程的约束条件，来求解最优跟踪组合。

### 目标函数

#### 1、跟踪误差最小化

##### 1) 跟踪误差（Tracking Error）的定义

跟踪误差通常用于刻画指数跟踪精度(Tracking Precision)，是指数跟踪业绩评价体系中最重要指标。一般来说，跟踪误差被定义为指数跟踪组合与目标指数收益之间差异的函数。

跟踪误差的度量方法有很多，通常分为线性跟踪误差和非线性跟踪误差。线性跟踪误差是跟踪组合与目标指数收益率之间的线性差异，即组合收益率与标的指数收益率之差。非线性跟踪误差将跟踪误差定义为跟踪组合与目标指数收益率之间的差异的方差或标准差。

##### 2) 跟踪误差的计算方法

本报告采用非线性跟踪误差计算方式，如下：

$$TE = \sqrt{\frac{\sum_{t=1}^T (ER_t - \overline{ER})^2}{T-1}}$$

其中  $ER_t = R_{Pt} - R_{It}$ ， $R_{Pt}$  为跟踪组合在 t 时刻的收益率， $R_{It}$  为目标指数

在 t 时刻的收益率， $\overline{ER}$  为  $ER_t$  的均值。

##### 3) 跟踪误差最小化下的目标函数

根据跟踪误差的计算方法，跟踪误差最小化下的目标函数为：

$$\min \frac{ssq(ER - \overline{ER})}{T-1}$$

其中， $ER = R_p - R_I$ ， $R_p = \begin{bmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{T1} & \cdots & R_{Tn} \end{bmatrix}$  为跟踪组合收益率矩

阵,  $R_I = \begin{bmatrix} R_{I1} \\ \vdots \\ R_{IT} \end{bmatrix}$  为基准收益率矩阵,  $\begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix}$  为跟踪组合成份股权重矩阵。

## 2、信息比率最大化

### 1) 信息比率(Information Ratio)的定义

信息比率是以马可维茨均值方差理论为基础,表示超额风险带来的超额收益,用来衡量投资组合相对于业绩基准的风险收益情况,如果信息比率小于 0,说明基金经理跑输业绩基准,如果信息比率为正且比较高则说明基金经理能够以相对较小的超额风险换取较大的超额收益。

### 2) 信息比率的计算方法

具体的计算公式为:

$$IR = \overline{ER} / TE$$

### 3) 信息比率最大化下的目标函数

根据信息比率的计算方法,信息比率最大化下的目标函数为:

$$\max \overline{ER} / TE$$

其中  $TE = \sqrt{\frac{\sum (ER - \overline{ER})^2}{T - 1}}$ ,  $ER = R_p - R_I$ ,  $R_p = \begin{bmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{T1} & \cdots & R_{Tn} \end{bmatrix} \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix}$ , 为

跟踪组合收益率矩阵,  $R_I = \begin{bmatrix} R_{I1} \\ \vdots \\ R_{IT} \end{bmatrix}$  为基准收益率矩阵,  $\begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix}$  为跟踪组合成

份股权重矩阵。

## 约束条件

### 1. 个股权重限制

由于一些制度和法规的约束,跟踪组合内单个股票的权重应在一定的范围内,如不能卖空;基金持有 1 家上市公司的股票,其市值不超过基金资产净值的 10% (完全复制情形除外);基金与由基金管理人管理的其他基金持有 1 家公司发行的证券的总和,不超过该证券的 10%,完全复制情形除外)等等。这里,我们将个股权重的上限设置为 10%,另外,为了控制每一期优化权重的变化幅度,减少调仓成本,我们将个股权重的下限设置为个股在基准指数中的权重。

$$W_{original} \leq W \leq 0.1$$

$$\sum_{i=1}^n \omega_i = 1$$



其中,  $W = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix}$  为跟踪组合个股权重矩阵,  $W_{original} = \begin{bmatrix} \omega_{o1} \\ \vdots \\ \omega_{on} \end{bmatrix}$  为跟踪组合

个股在基准指数中的权重矩阵。

## 2. 行业权重限制

行业属性是上市公司最重要的属性之一, 行业板块的轮动效应表现的比其它类别的板块轮动效应都更加明显。为了抵消各行业轮动带来的收益偏差, 在较少跟踪样本数的前提下最大限度的获取和目标指数相同的收益水平, 需要让跟踪组合中包含的股票在行业这一属性上均匀覆盖, 并和目标指数中成份股在各行业的分布保持一致的比例。

这里, 我们将跟踪组合的各行业权重控制在该行业在指数中的权重上下浮动的一定范围内。

$$W_{sectororiginal} - \theta \leq W_{sector} \leq W_{sectororiginal} + \theta$$

其中,  $W_{s\_corc} = \begin{bmatrix} \omega_{s\_corc} \\ \vdots \\ \omega_{s\_corc} \end{bmatrix}$  为跟踪组合行业权重矩阵,

$W_{sectororiginal} = \begin{bmatrix} \omega_{sectoro1} \\ \vdots \\ \omega_{sectoroj} \end{bmatrix}$  为跟踪组合行业在基准指数中的权重矩阵。

## 指数跟踪优化模型

综上所述, 本报告所使用的优化模型为如下非线性规划模型:

$$\min \frac{ssq(ER - \overline{ER})}{T-1} \text{ 或 } \max \overline{ER} / TE$$

$$s.t. \quad W_{original} \leq W \leq 0.1$$

$$\sum_{i=1}^n \omega_i = 1$$

$$W_{sectororiginal} - \theta \leq W_{sector} \leq W_{sectororiginal} + \theta$$

其中,

$$a) TE = \sqrt{\frac{\sum (ER - \overline{ER})^2}{T-1}} \text{ 为跟踪误差,}$$

$$b) ER = R_p - R_f,$$

$$c) R_p = \begin{bmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{T1} & \cdots & R_{Tn} \end{bmatrix} \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix}, \text{ 为跟踪组合收益率矩阵,}$$

$$d) R_I = \begin{bmatrix} R_{I1} \\ \vdots \\ R_{IT} \end{bmatrix} \text{ 为基准收益率矩阵,}$$

$$e) W = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix} \text{ 为跟踪组合个股权重矩阵,}$$

$$f) W_{original} = \begin{bmatrix} \omega_{O1} \\ \vdots \\ \omega_{On} \end{bmatrix} \text{ 为跟踪组合个股在基准指数中的权重矩阵,}$$

$$g) W_{sector} = \begin{bmatrix} \omega_{sector1} \\ \vdots \\ \omega_{sectorj} \end{bmatrix} \text{ 为跟踪组合行业权重矩阵,}$$

$$h) W_{sectororiginal} = \begin{bmatrix} \omega_{sectoro1} \\ \vdots \\ \omega_{sectoroj} \end{bmatrix} \text{ 为跟踪组合行业在基准指数中的权重矩阵。}$$

## 案例分析 1——跨市场指数的抽样复制

### 案例构想：

A 股分为上海和深圳两个市场，这成为设置跨市场 ETF 的最大障碍。那么是否能用上海市场的股票抽样复制跨市场的指数，如中证 100、中证 500？

### 实证设置

我们对中证 100 和中证 500 进行了实证分析，实证的细节设置如下：

#### 1. 股票池的选取

- 1) 考虑到跨市场的问题，在实证中，我们将跟踪组合的成份股限定在上海市场，剔除掉深圳股票。
- 2) 由于优化时，需要采用历史收益率矩阵，我们在选取各期的成份股时，对上市时间和历史停牌天数进行了限制：上市时间超过三个月，过去 3 个月的停牌天数比例低于 15%。

#### 2. 成份股的数量

- 1) 对于中证 100，测试时所用到的成份股数量为：停牌和上市天数剔除后的



所有上海股票。

2) 对于中证 500，测试时所用到的成份股数量为：停牌和上市天数剔除后的所有上海股票。

### 3. 复制方法

最优化抽样复制、市值加权、行业分层市值加权。

### 4. 最优化方法的参数设置

目标函数：跟踪误差最小化；信息比最大化

历史收益率矩阵：20 日，30 日，40 日

行业限制： $\theta = 0.02$   $\theta = 0.02$

### 5. 最优化问题求解方法

共轭梯度法（见附录）

### 6. 再平衡周期设置

每月底进行一次再平衡。

### 7. 实证区间

中证 100：2007 年 1 月 4 日-2011 年 8 月 5 日

中证 500：2009 年 4 月 1 日-2011 年 8 月 5 日

## 实证结果

对于中证 100，市值加权法和行业分层市值加权法的年化跟踪误差分别为 0.102%，0.082%，信息比分别为-0.588%，-1.452%；跟踪误差最小化 20 日、30 日、40 日优化抽样的年化跟踪误差分别为 0.153%，0.150%，0.155%，信息比分别为 0.568%，-1.775%，0.642%；信息比率最大化 20 日、30 日、40 日优化抽样的年化跟踪误差分别为 0.155%，0.160%，0.161%，信息比分别为-0.616%，0.390%，-0.743%。

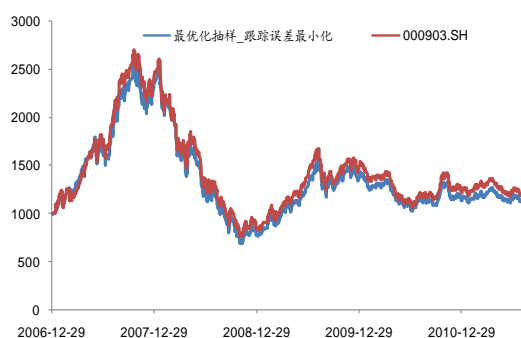
根据实证结果，相比之下，对于中证 100，历史收益率矩阵采用 30 天复制效果最好，20 天，40 天相差不大。

表 1：中证 100 抽样复制结果

跟踪误差及信息比	最优化抽样						市值加权法	行业分层市值加权法
	信息比最大化			跟踪误差最小化				
	20	30	40	20	30	40		
区间跟踪误差	0.329%	0.339%	0.340%	0.323%	0.318%	0.327%	0.216%	0.260%
年化跟踪误差	0.155%	0.160%	0.161%	0.153%	0.150%	0.155%	0.102%	0.082%
信息比率	-0.616%	0.390%	-0.743%	0.568%	-1.775%	0.642%	-0.588%	-1.452%
2007 年跟踪误差	0.450%	0.492%	0.512%	0.472%	0.458%	0.466%	0.284%	0.242%
2008 年跟踪误差	0.360%	0.342%	0.328%	0.348%	0.339%	0.344%	0.262%	0.210%
2009 年跟踪误差	0.271%	0.272%	0.277%	0.243%	0.259%	0.290%	0.167%	0.115%
2010 年跟踪误差	0.256%	0.265%	0.245%	0.247%	0.238%	0.238%	0.167%	0.128%
2011 年跟踪误差	0.212%	0.214%	0.208%	0.172%	0.171%	0.172%	0.131%	0.104%

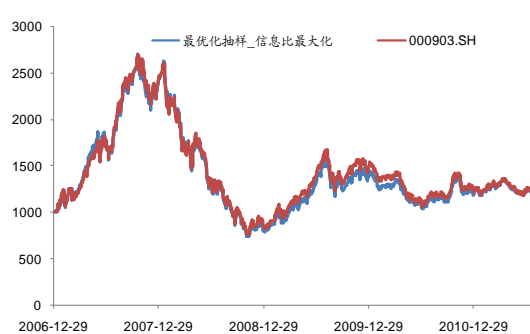
资料来源：国信证券经济研究所

图 1: 中证 100 最优化抽样\_跟踪误差最小化\_30



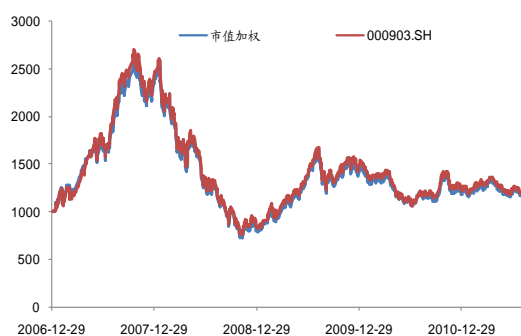
资料来源: 国信证券经济研究所, 中证指数公司, Wind 资讯

图 2: 中证 100 最优化抽样\_信息比最大化\_30



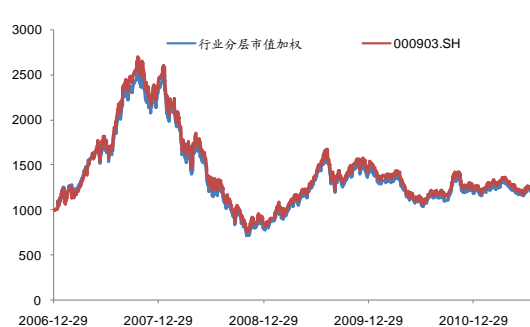
资料来源: 国信证券经济研究所, 中证指数公司, Wind 资讯

图 3: 中证 100 抽样\_市值加权



资料来源: 国信证券经济研究所, 中证指数公司, Wind 资讯

图 4: 中证 100 抽样\_行业分层市值加权



资料来源: 国信证券经济研究所, 中证指数公司, Wind 资讯

对于中证 500, 市值加权法和行业分层市值加权法的年化跟踪误差分别为 0.078%, 0.084%, 信息比分别为 2.508%, -0.117%; 跟踪误差最小化 20 日、30 日、40 日优化抽样的年化跟踪误差分别为 0.274%, 0.242%, 0.243%, 信息比分别为 0.449%, -7.055%, -4.808%; 信息比率最大化 20 日、30 日、40 日优化抽样的年化跟踪误差分别为 0.399%, 0.407%, 0.401%, 信息比分别为 7.926%, 2.547%, 4.823%。

表 2: 中证 500 抽样复制结果

跟踪误差及信息比	最优化抽样						市值加权法	行业分层市值加权法
	信息比最大化			跟踪误差最小化				
	20	30	40	20	30	40		
区间跟踪误差	0.605%	0.617%	0.608%	0.415%	0.367%	0.368%	0.118%	0.127%
年化跟踪误差	0.399%	0.407%	0.401%	0.274%	0.242%	0.243%	0.078%	0.084%
信息比率	7.926%	2.547%	4.823%	0.449%	-7.055%	-4.808%	2.508%	-0.117%
2009 年跟踪误差	0.470%	0.463%	0.468%	0.359%	0.391%	0.399%	0.112%	0.114%
2010 年跟踪误差	0.616%	0.628%	0.640%	0.375%	0.331%	0.340%	0.117%	0.132%
2011 年跟踪误差	0.725%	0.747%	0.693%	0.521%	0.392%	0.373%	0.126%	0.131%

资料来源: 国信证券经济研究所, 中证指数公司, Wind 资讯

图 5: 中证 500 最优化抽样\_跟踪误差最小化\_30



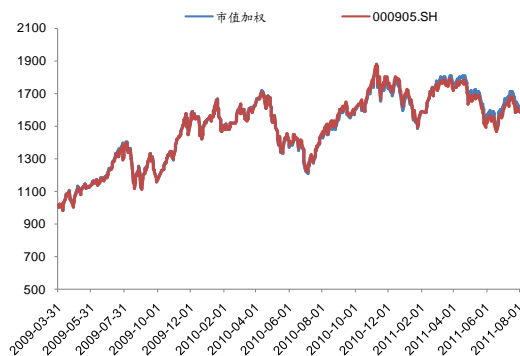
资料来源: 国信证券经济研究所, 中证指数公司, Wind 资讯

图 6: 中证 500 最优化抽样\_信息比最大化\_20



资料来源: 国信证券经济研究所, 中证指数公司, Wind 资讯

图 7: 中证 500 抽样\_市值加权



资料来源: 国信证券经济研究所, 中证指数公司, Wind 资讯

图 8: 中证 500 抽样\_行业分层市值加权



资料来源: 国信证券经济研究所, 中证指数公司, Wind 资讯

### 结果分析及思考

1. 以跟踪误差最小化为目标的优化抽样结果不理想, 相对市值加权, 跟踪误差稍高。  
分析原因为: 市值加权法和行业分层市值加权法的优点是抓住了股票最重要的两个特征, 行业和规模。在中证 100 和中证 500 中, 上海市场股票占大多数的市值和数量, 所以市值加权法的跟踪误差较小。举例来说, 市值加权的上证 180 指数的走势已经几乎可以代表整个上海市场。
2. 但是, 用市值加权的方式替代指数中的深圳部分权重, 本身是没有科学依据的。  
理论上最优化的方法可以改进跟踪误差 (Tracking Error), 没有改进的原因有两点。一是, 选择的成份股本身已经涵盖了指数的大部分市值; 二是, 最优化方法的换仓频率最高设为 1 个月, 不够频繁。
3. 以信息比 (Information Ratio) 为优化目标的优化方法, 显然在收益上能战胜指数, 同时跟踪误差放大程度有限, 因此比较适用于样本量较大, 抽样比例较小的指数抽样, 例如, 富国上证综指 ETF (510210)。

## 案例分析 2——大宽基指数的抽样复制

案例构想：

在案例 1 中，用来抽样的成份股仅限于上海市场（代码：6\*\*\*\*），股票数量占比较大，也不用主动的选股。

但是，类似富国上证综指 ETF（510210）的指数基金（上证综合指数（000001）有 900 只以上的成份股）如果抽样 100-200 只成份股作为基金持仓，那么抽样的比例将会比较小，而且抽样过程中包含了基金经理主动选股的过程，不论是主观选择还是数量模型选股，都会兼顾跟踪误差和相对收益这两个因素。

因此，这种类型指数复制，将非常适合用最优化的方法，并以信息比（Information Ratio）作为优化目标，既能保证跟上指数，又能一定程度上获取 Alpha，甚至在此基础上开发更激进的指数增强型基金。

## 基于最优化方法的成份替代法

在对指数进行全复制时，会遇到成份股限制持仓（基金资产不能投资于基金托管行发行的股票或与基金管理人、基金托管人有控股关系或其他重大利害关系的公司发行的股票）、成份股长期停牌、流动性不足这三种需要进行成份股替代的情形。

以上三种成份股替代情形中，长期停牌和流动性不足，长期来看，不会对组合的跟踪误差造成较大影响。

而对于持仓限制的情形，目前指数基金的托管行主要集中在工、农、中、建、交行，以及招商、民生、兴业等商业银行，在指数中占有较大的权重，如果不能买入相关的股票进行组合复制，将会造成较大的跟踪误差。

对于托管行股票的替代问题，大家通常都采用“近似”替代的方法：基于基本面的替代法：即寻找与标的股票基本面相似的股票作为替代，因为基本面相似的公司发行的股票价格趋势通常也很相近；基于相关关系的替代法：通过历史的数据的回归统计，寻找与标的股票价格走势长期具有较强且稳定的正相关性的股票作为替代。上述的两种方法均着重于“选股”，对权重没有做精确的处理，在跟踪误差的控制方面存在很大的改善空间。

在这里，我们提出基于最优化方法的成份股替代方法，即结合基本面、相关性选定替代股票，利用优化技术，以跟踪误差最小或信息比最大为目标，得到替代组合的优化权重。

基于最优化方法的成份股替代方法和前文所述的指数复制方法类似，只需把跟踪标的替换为所需替代的成份股，约束条件稍加修改即可，模型如下：

$$\min \frac{ssq(ER - \overline{ER})}{T-1} \text{ 或 } \max \overline{ER} / TE$$

$$s.t. \quad 0 \leq W \leq 1$$

$$\sum_{i=1}^n \omega_i = 1$$

其中,

$$a) \quad TE = \sqrt{\frac{\sum (ER - \overline{ER})^2}{T - I}}, \text{ 为跟踪误差,}$$

$$b) \quad ER = R_p - R_{sub},$$

$$c) \quad R_p = \begin{bmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{T1} & \cdots & R_{Tn} \end{bmatrix} \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix}, \text{ 为替代组合收益率矩阵,}$$

$$d) \quad R_{sub} = \begin{bmatrix} R_{sub1} \\ \vdots \\ R_{subT} \end{bmatrix} \text{ 为需替代的成份股收益率矩阵,}$$

$$e) \quad W = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix} \text{ 为替代组合个股权重矩阵.}$$

## 案例分析 3——托管行股票替代

### 实证设置

我们对中证 100 中的工商银行 (601398.SH) 和建设银行 (601939.SH) 的替代进行了实证分析, 实证的细节设置如下:

#### 1. 替代组合的股票选取

1) 考虑到跨市场的问题, 在实证中, 我们将替代组合的成份股限定在上海市场的银行股。

2) 由于优化时, 需要采用历史收益率矩阵, 我们在选取各期的成份股时, 对上市时间和历史停牌天数进行了限制: 上市时间超过三个月, 过去 3 个月的停牌天数比例低于 15%。

#### 2. 替代方法

最优化方法、市值加权法、最大相关系数法。

#### 3. 最优化方法的参数设置

目标函数: 跟踪误差最小化

历史收益率矩阵: 20 日, 30 日, 40 日

#### 4. 再平衡周期设置

每月底进行一次再平衡

#### 5. 最优化问题求解方法

共轭梯度法（见附录）

#### 6. 实证区间

1) 工商银行（601398.SH）：2007 年 1 月 4 日-2011 年 8 月 5 日

2) 建设银行（601939.SH）：2008 年 7 月 1 日-2011 年 8 月 5 日

### 实证结果

对于工商银行（601398.SH）采用优化替代的年化跟踪误差分别为 0.0051（20 日），0.0051（30 日），0.0050（40 日），采用市值加权法和最大相关系数法的年化跟踪误差分别为 0.0074，0.0063。

**表 3: 工商银行（601398.SH）替代结果**

跟踪误差	优化替代（30 日）	优化替代（30 日）	优化替代（40 日）	市值加权法	最大相关系数法（20 日）
区间跟踪误差	0.0108	0.0107	0.0106	0.0156	0.0134
年化跟踪误差	0.0051	0.0051	0.0050	0.0074	0.0063
2007 年跟踪误差	0.0116	0.0115	0.0113	0.0199	0.0135
2008 年跟踪误差	0.0127	0.0127	0.0129	0.0187	0.0180
2009 年跟踪误差	0.0093	0.0085	0.0085	0.0123	0.0098
2010 年跟踪误差	0.0114	0.0114	0.0112	0.0132	0.0135

资料来源：国信证券经济研究所，中证指数公司，Wind 资讯

对于建设银行（601939.SH）采用优化替代的年化跟踪误差分别为 0.0091（20 日），0.0089（30 日），0.0088（40 日），采用市值加权法和最大相关系数法的年化跟踪误差分别为 0.0132，0.0111。

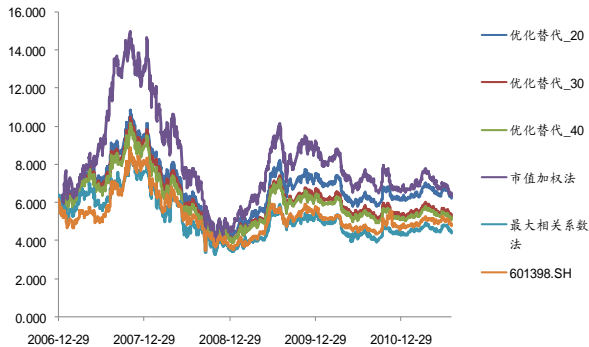
**表 4: 建设银行（601939.SH）替代结果**

跟踪误差	优化替代（20 日）	优化替代（30 日）	优化替代（40 日）	市值加权法	最大相关系数法（20 日）
区间跟踪误差	0.0091	0.0089	0.0088	0.0132	0.0111
年化跟踪误差	0.0051	0.0050	0.0049	0.0074	0.0063
2007 年跟踪误差	0.0103	0.0101	0.0100	0.0154	0.0142
2008 年跟踪误差	0.0077	0.0075	0.0073	0.0140	0.0089
2009 年跟踪误差	0.0094	0.0092	0.0091	0.0097	0.0100
2010 年跟踪误差	0.0049	0.0049	0.0049	0.0116	0.0076

资料来源：国信证券经济研究所，中证指数公司，Wind 资讯

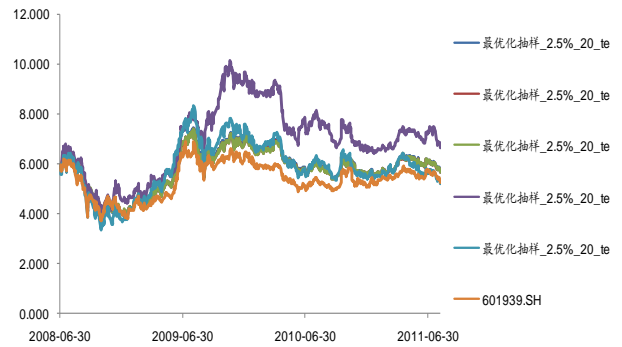


图 9: 工商银行 (601398.SH) 替代效果



资料来源: 国信证券经济研究所, 中证指数公司, Wind 资讯

图 10: 建设银行(601939.SH)替代效果



资料来源: 国信证券经济研究所, 中证指数公司, Wind 资讯

### 结果分析及思考

1. 从成份股替代的结果看, 都显著的降低了跟踪误差, 效果较好。
2. 从此方法的操作层面看, 投资逻辑明确, 可投资性比较强。首先, 在该成份股的行业内选择替代股, 符合股票的最本质特征, 风险也较小。其次, 计算出来的替代权重, 累加到原有的基准权重中, 交易、换仓都非常方便, 不增加额外的交易成本。
3. 此方法, 可以接受卖空权重。如: 计算得到替代 1 手 A 股票的权重是: 1.5 手 B 和 -0.5 手 C, 只要原来基准仓位中, 有超过 0.5 手的 C。就可以实现替代策略。
4. 此类基于最优化的替代方法能用于: 1、有较大比例限购的指数基金; 2、有长期停牌成份股的指数基金; 3、某时间段内看空某些成份股, 希望替代得到 Beta 和 Alpha。

### 总结及后续研究

最优化方法属于数值计算方法, 能有效的把握股票走势的关系, 虽然对选股、收益率预测没有太大帮助, 但是能帮助基金经理更精细化、模型化的管理量化基金和被动基金。

从国外经验和数量模型产品市场来看, 优化器软件是常用的商业软件, 如 (BARRA Optimizer, Axioma 优化器, Risk Matrixs), 它能广泛的解决投资决策模型中的大量求解问题。但是直接购买优化器软件最大的弊端就是: 1、黑箱操作, 对机理不了解; 2、不能灵活运用优化技术; 3、如果直接用于投资, 第三方风险较大。因此, 作为数量分析师, 利用通用的软件工具如 SAS、Matlab 等内置的优化算法, 结合数据处理, 并将之模块化, 能够完美的、灵活的实现优化器的功能。

未来研究计划

1. 将最优化方法引入指数基金管理和数量基金管理。将优化技术与多因子模型, 收益率预测模型结合, 开发抽样复制技术、数量指数增强策略, 130/30 指数, 数量基金等。

2. 将最优化的方法引入数量基金管理。将多因子选股、收益率预测等模型和优化器模型结合，开发数量基金和数量增强基金。
3. 将最优化方法引入行业配置，意在寻找有效的行业组合。
4. 定期跟踪成份股替代，并发布替代结果。

## 附录：共轭梯度法介绍

共轭梯度法最早是由 Magnus R. Hestenes 和 Eduard Stiefel 于 1952 年提出来的,用于解正定系数矩阵的线性方程组,在这个基础上,Fletcher 和 Reeves 于 1964 年首先提出了解非线性最优化问题的共轭梯度法。由于共轭梯度法且有较快的收敛速度和二次终止性,不用求矩阵的逆,在使用计算机求解时,所需存储量较小,适合于求解变量较多的大规模问题,现在共轭梯度法已经广泛地应用与实际问题中。共轭梯度法的具体算法步骤如下:

第一步:选取初始数据。

选取初始点  $x^{(0)}$ , 给出终止误差  $\varepsilon > 0$ 。

第二步:求初始点梯度。

计算  $\nabla f(x^{(0)})$ 。若  $\|\nabla f(x^{(0)})\| \leq \varepsilon$ , 停止迭代,输出  $x^{(0)}$ , 否则转第三步。

第三步:构造初始搜索方向。

令  $d^{(0)} = -\nabla f(x^{(0)})$ 。

令  $k:=0$ , 进行第四步。

第四步:进行一维搜索,求  $\lambda_k$ , 使

$$f(x^{(k)} + \lambda_k d^{(k)}) = \min_{\lambda \geq 0} (x^{(k)} + \lambda_k d^{(k)})$$

令  $x^{(k+1)} = x^{(k)} + \lambda_k d^{(k)}$ , 转第五步。

第五步,求梯度向量。

计算  $\nabla f(x^{(k+1)})$ 。若  $\|\nabla f(x^{(k+1)})\| \leq \varepsilon$ , 停止迭代,输出  $x^*$  的近似值,  $x^* \approx x^{(k+1)}$ 。

否则进行第六步。

第六步,检查迭代次数。

若  $k+1=n$ , 令  $x^{(0)} := x^{(n)}$ , 转第三步, 否则进行第七步。

第七步:构造搜索方向,令  $k := k+1$ , 转四步。

## 国信证券投资评级

类别	级别	定义
股票 投资评级	推荐	预计 6 个月内，股价表现优于市场指数 20%以上
	谨慎推荐	预计 6 个月内，股价表现优于市场指数 10%-20%之间
	中性	预计 6 个月内，股价表现介于市场指数 $\pm 10\%$ 之间
	回避	预计 6 个月内，股价表现弱于市场指数 10%以上
行业 投资评级	推荐	预计 6 个月内，行业指数表现优于市场指数 10%以上
	谨慎推荐	预计 6 个月内，行业指数表现优于市场指数 5%-10% 之间
	中性	预计 6 个月内，行业指数表现介于市场指数 $\pm 5\%$ 之间
	回避	预计 6 个月内，行业指数表现弱于市场指数 5%以上

## 风险提示

本报告信息均来源于公开资料，我公司对这些信息的准确性和完整性不作任何保证。报告中的内容和意见仅供参考，并不构成对所述证券买卖的出价或询价。我公司及其雇员对使用本报告及其内容所引发的任何直接或间接损失概不负责。我公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行业务服务。本报告版权归国信证券所有，未经书面许可任何机构和个人不得以任何形式翻版、复制、刊登。

## 证券投资咨询业务的说明

证券投资咨询业务是指取得监管部门颁发的相关资格的机构及其咨询人员为证券投资者或客户提供证券投资的相关信息、分析、预测或建议，并直接或间接收取服务费用的活动。

证券研究报告是证券投资咨询业务的一种基本形式，指证券公司、证券投资咨询机构对证券及证券相关产品的价值、市场走势或者相关影响因素进行分析，形成证券估值、投资评级等投资分析意见，制作证券研究报告，并向客户发布的行为。

### 国信证券经济研究所团队成员

<b>宏观</b>		<b>固定收益</b>		<b>策略</b>	
周炳林	0755-82130638	李怀定	021-60933152	黄学军	021-60933142
林松立	010-66026312	侯慧梯	021-60875161	林丽梅	021-60933157
崔 嵘	021-60933159	张 旭	010-66026340	<b>技术分析</b>	
				闫 莉	010-88005316
<b>交通运输</b>		<b>银行</b>		<b>房地产</b>	
郑 武	0755- 82130422	邱志承	021- 60875167	方 焱	0755-82130648
陈建生	0755- 82133766	黄 飙	0755-82133476	区瑞明	0755-82130678
岳 鑫	0755- 82130432	谈 煜	010- 66025229	黄道立	0755- 82133397
周 俊	0755-82130833-6215				
糜怀清	021-60933167				
<b>商业贸易</b>		<b>汽车及零配件</b>		<b>钢铁及新材料</b>	
孙菲菲	0755-82130722	左 涛	021-60933164	郑 东	010- 66025270
祝 彬	021-60933156			秦 波	010-66026317
常 伟				郭 莹	010-88005303
<b>机械</b>		<b>基础化工</b>		<b>医药</b>	
郑 武	0755- 82130422	刘旭明	010-66025272	贺平鸽	0755-82133396
陈 玲	0755-82130646	张栋梁	0755-82130532	丁 丹	0755- 82139908
杨 森	0755-82133343	罗 洋	0755-82150633	杜佐远	0755-82130473
后立尧	010-88005327	吴琳琳	0755-82130833-1867	胡博新	0755-82133263
		梁 丹	0755- 82134323	刘 勍	0755-82130833-1845
<b>电力设备与新能源</b>		<b>传媒</b>		<b>有色金属</b>	
杨敬梅	021-60933160	陈财茂	010-88005322	彭 波	0755-82133909
张 弢	010-88005311	刘 明	010-88005319	龙 飞	
<b>电力与公用事业</b>		<b>非银行金融</b>		<b>通信</b>	
谢达成	021-60933161	邵子钦	0755- 82130468	严 平	021-60875165
		田 良	0755-82130513	唐俊杰	021-60875160
		童成敦	0755-82130513		
<b>造纸</b>		<b>家电</b>		<b>计算机</b>	
李世新	0755-82130565	王念春	0755-82130407	段迎晟	0755- 82130761
邵 达	0755-82130706			欧阳仕华	0755-82151833
<b>电子元器件</b>		<b>纺织服装</b>		<b>农业</b>	
段迎晟	0755- 82130761	方军平	021-60933158	张 如	021-60933151
高耀华	0755-82130771				
熊 丹	0755-82133528				
<b>建材</b>		<b>旅游</b>		<b>食品饮料</b>	
郑 东	010- 66025270	曾 光	0755-82150809	黄 茂	0755-82138922
马 彦	010-88005304				
<b>建筑</b>		<b>新兴产业</b>		<b>研究支持</b>	
邱 波	0755-82133390	陈 健	010-66022025	沈 瑞	0755-82132998
刘 萍	0755-82130678	李筱筠	010-66026326	雷 达	0755-82132098
		孙 伟	010- 66026320	余 辉	0755-82130741
				王越明	0755-82130478
<b>量化投资产品</b>		<b>基金评价与研究</b>		<b>量化投资策略</b>	
焦 健	0755-82133928	杨 涛	0755-82133339	葛新元	0755-82133332
阳 瑾	0755-82133538	康 亢	010-66026337	董艺婷	021-60933155
周 琦	0755-82133568	刘舒宇	0755-82133568	郑 云	021-60875163
邓 岳	0755- 82150533	李 腾	0755-82130833-6223	毛 甜	021-60933154
		刘 洋	0755-82150566	李荣兴	021-60933165
		潘小果	0755-82130843	郑亚斌	
		蔡乐祥	0755-82130833-1368		
		钱 晶	0755-82130833-1367		
<b>量化交易策略与技术</b>		<b>数据与系统支持</b>			
戴 军	0755-82133129	赵斯尘	021-60875174		
黄志文	0755-82133928	徐左乾	0755-82133090		

彭甘霖	0755-82133259	李扬之	0755-82136165
秦国文	0755-82133528	陈爱华	0755-82133397
韦 敏	0755-82130833-3772	袁 剑	0755-82139918
张璐楠	0755-82130833- 1379		

### 国信证券机构销售团队

华北区（机构销售一部）		华东区（机构销售二部）		华南区（机构销售三部）	
王立法	010-66026352 13910524551 wanglf@guosen.com.cn	盛建平	021-60875169 15821778133 shengjp@guosen.com.cn	魏 宁	0755-82133492 13823515980 weining@guosen.com.cn
王晓建	010-66026342 13701099132 wangxj@guosen.com.cn	马小丹	021-60875172 13801832154 maxd@guosen.com.cn	邵燕芳	0755-82133148 13480668226 shaoyf@guosen.com.cn
焦 戡	010-66026343 13601094018 jiaojian@guosen.com.cn	郑 毅	021-60875171 13795229060 zhengyi@guosen.com.cn	林 莉	0755-82133197 13824397011 linli2@guosen.com.cn
李文英	010-88005334 13910793700 liwying@guosen.com.cn	黄胜蓝	021-60875166 13761873797 huangsl@guosen.com.cn	王昊文	0755-82130818 18925287888 wanghaow@guosen.com.cn
赵海英	010-66025249 13810917275 zhaohy@guosen.com.cn	刘 塑	021-60875177 13817906789 liusu@guosen.com.cn	甘 墨	0755-82133456 15013851021 ganmo@guosen.com.cn
原 祎	010-88005332 15910551936 yuanyi@guosen.com.cn	叶琳菲	021-60875178 13817758288 yelf@guosen.com.cn	段莉娟	0755-82130509 18675575010 duanlj@guosen.com.cn
		孔华强	021-60875170 13681669123 konghq@guosen.com.cn	徐 冉	0755-82130655 13632580795 xuran1@guosen.com.cn
				颜小燕	0755-82133147 13590436977 yanxy@guosen.com.cn
				赵晓曦	0755-82134356 15999667170 zhaoxxi@guosen.com.cn
				郑 灿	0755-82133043 13421837630 zhengcan@guosen.com.cn