

相关研究

《听海外高频交易专家讲解美国的高频交易》2019.06.11

《中证红利低波指数及华泰柏瑞中证红利低波ETF及其联接基金投资价值分析》2019.06.11

《选股因子系列研究（四十八）——探索A股的五因子模型》2019.05.28

金融科技（Fintech）和数据挖掘研究（三） ——量化因子的批量生产与集中管理

投资要点：

- **量化因子计算可以定义为一种基于初等计算函数与一阶谓词逻辑的递归过程。**量化因子的本质是股票的某种数量化特征，一般需要通过通过对股票某些数量信息进行计算得到。与计算机函数从“函数输入出发，输入函数计算，得到函数输出”相类似，量化计算过程也可以拆解为“从量化指标出发，输入量化计算，得到新量化指标”。由此可见，计算机函数构成的基本逻辑——递归过程也同样适用于因子计算架构。
- **定义量化指标的计算过程有计算机语言与 XML 文本两种方式。**由于 λ 算子与图灵机、递归函数等价，因此无论是基于 λ 算子的函数式编程语言，还是基于图灵机的过程、命令范式语言，都可以用来构建量化计算过程。在定义逻辑结构较为复杂的因子时，过程、命令范式语言较为合适。而当需要批量生产因子时，可以表达递归逻辑的 XML 文本也可以作为递归逻辑描述语言来使用。
- **利用自动因子生成器暴力挖掘因子是生产有效因子的重要方式。**利用自动因子生成器构建的 226 个突破型因子，其中有 20%-30% 在截面上有较好的选股效果。然而，经正交后，效果明显下降。这说明，暴力挖掘方式还是能获取一部分传统因子构建过程中没有挖掘到的信息，但毕竟两种方式具有相同的初始算根，有大量信息被传统因子所解释。
- **利用逐笔数据构建的高频因子有着非常突出的效果。**构建买入大单成交额占比因子，其截面表现非常优异。利用正交后的因子构建等权组合，可获得接近 20% 的多空年化收益与 9 左右的收益回撤比。高频数据的处理成本非常高，目前只能进行 15 年 8 月以来的短周期回测，所以，因子表现的稳定性需要进一步跟踪。
- **自动构建数据表是进行因子自动化管理的关键要素。**依据因子结构进行因子数据库管理是因子计算自动化的重要步骤，而自动构建 SQL 语句，对数据表进行增删查改是实现这种自动化管理的关键技术。
- **风险提示。**数据挖掘是从历史先验数据获取经验模型的方法，存在模型失效可能。

分析师:冯佳睿

Tel:(021)23219732

Email:fengjr@htsec.com

证书:S0850512080006

分析师:余浩淼

Tel:(021)23219883

Email:yhm9591@htsec.com

证书:S0850516050004

目 录

1. 计算机视角下的量化因子计算过程.....	5
2. 量化因子计算过程的实现	6
3. 基于日频数据的量化因子构建	6
3.1 日频初始量化指标计算过程	6
3.2 日频量化指标计算算子.....	7
3.3 利用计算机语言表达量化计算过程的递归逻辑	8
3.4 利用量化计算过程构建自动因子生成器	10
4. 基于高频数据的量化因子构建	12
4.1 高频初始量化指标计算过程	12
4.2 高频量化计算算子.....	13
4.3 基于高频初始计算过程的高频成交因子	14
5. 因子自动管理系统框架	16
6. 总结.....	17
7. 风险提示.....	17

图目录

图 1	量化因子计算的基本过程.....	5
图 2	量化因子计算递归流程简介	5
图 3	量化因子计算过程构建流程	6
图 4	量化指标计算语言示例	9
图 5	量化因子批量生产的逻辑流程图	10
图 6	暴力因子生成的 XML 脚本示意	11
图 7	组合一：多空净值 ($IC > 0.02$)	12
图 8	组合二：多空净值 (IC 分位数 $> 90\%$)	12
图 9	TotalBidAmount 等权组合超额收益	15
图 10	TotalBidAmount 等权组合多空收益	15
图 11	因子管理框架示意图	16
图 12	因子关系数据库中，用于建表的 SQL 语句示例	16

表目录

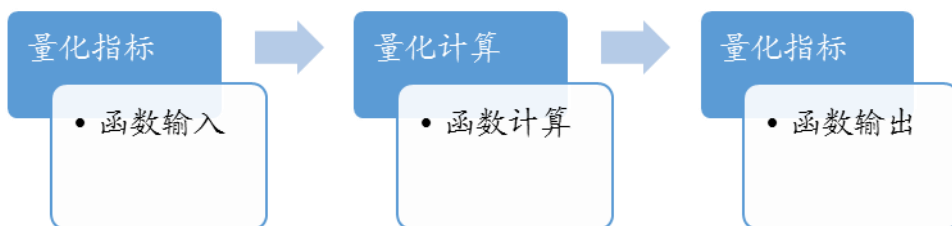
表 1 量化指标计算过程递归步骤	5
表 2 常用日频基础数据列表	7
表 3 常用量化指标计算算子	8
表 4 BARRA 风格因子定义	9
表 5 BARRA 因子 IC 与多空收益率（2010.01-2019.04）	10
表 6 均值突破因子截面表现（2010.01-2019.05）	11
表 7 正交后均值突破因子截面表现（2010.01-2019.05）	11
表 8 LEVEL2 高频数据列表	13
表 9 TotalBidAmount 因子 IC 与多空收益率（2015.08-2019.05）	14
表 10 正交 TotalBidAmount 因子 IC 与多空收益率（2015.08-2019.05）	15
表 11 正交 TotalBidAmount 因子的截面溢价（2015.08-2019.05）	15

量化因子是构建多因子组合的基本要素。随着越来越多不同种类的有效因子被发现,如何快速高效地计算因子,向前更新因子值,自动存储历史数据与新数据,方便快捷地访问因子值,对于量化组合管理来说,正变得越来越重要。本文希望从计算机的角度,构建一个完整的从计算到更新,再到管理的量化因子维护框架。

1. 计算机视角下的量化因子计算过程

量化因子的本质是股票的某种数量化特征，一般需要通过股票某些数量信息进行计算得到。转换成计算机语言为“以股票有关数据为输入，通过量化计算函数计算得到新的量化指标作为输出”。这一过程可以简单地通过下图来表达。

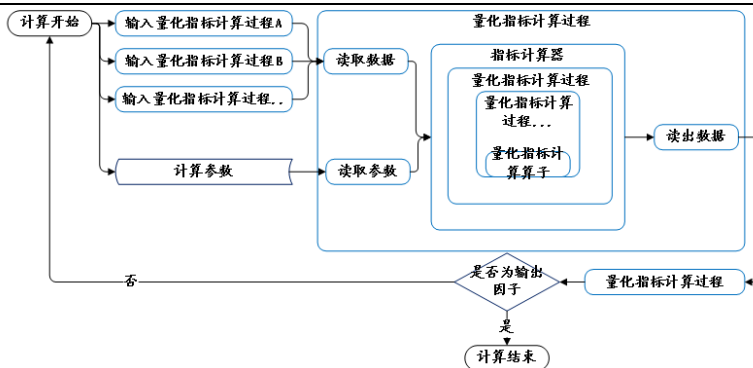
图1 量化因子计算的基本过程



资料来源：海通证券研究所整理

与计算机函数“从函数输入出发，输入函数计算，得到函数输出”相类似，量化计算过程也可以拆解为“从量化指标出发，输入量化计算，得到新量化指标”。由此可见，计算机函数构成的基本逻辑——递归过程也同样适用于因子计算架构。

图2 量化因子计算递归流程简介



资料来源：海通证券研究所整理

上图展示了因子计算的一般递归过程,其中所有实体都被定义为量化指标计算过程。这是以一个量化指标计算过程为输入,获得新量化指标计算过程的实体。量化指标计算过程的递归步骤如下表所示。

表 1 量化指标计算过程递归步骤

Step1:	从输入量化指标计算过程中读取其输出的量化指标或者参数
Step2:	将参数输入包含多个量化指标计算过程的指标计算器，计算获得输出量化指标计算过程
Step3:	判断输出量化指标计算过程是否为所求因子
Step4:	如果是所求因子，结束计算过程，获得所求输出。否则，将该量化指标计算过程作为输入参数，输入新的量化指标计算过程，从 Step1 重新开始计算。

资料来源：Wind，海通证券研究所

从上述过程中，我们可以得到计算量化指标最重要的 4 个要素。

量化指标计算过程：一个递归的计算实体，其输入与输出也均为量化指标计算过程。运算过程包括数据读取，参数读取，指标计算器，输出计算过程四步。

计算参数：一种常数量化指标计算过程。即，不需要任何参数或数据作为输入，便可输出返回一个固定的常数。

初始量化指标计算过程：仅以计算参数为输入，直接依据参数从数据源读取数据的过程。一般为从外部数据源获取股票成交价等量化数据的过程。

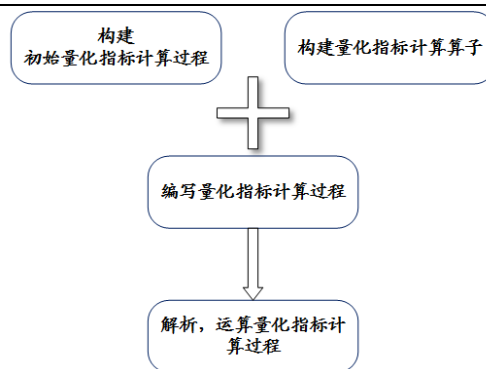
量化指标计算算子：一种指标计算器，仅包含初等基本函数计算或者简单一阶谓词逻辑计算的过程。由于常用的股票量化数据以时间序列数据或面板数据居多，因此这些基本函数计算一般为面板数据在截面上的计算函数或基础统计函数，或时间序列数据的一些基本运算。

2. 量化因子计算过程的实现

由上文量化指标计算过程的逻辑框架定义可以看出，实现该计算过程需要以下几个步骤：

1. 构建初始量化指标计算过程，也就是提取计算量化指标的原始数据。该过程是整个因子计算的起点。
2. 构建量化指标计算算子，即实现各种不同的基本函数计算或者一阶谓词逻辑的简单组合。
3. 采用某种递归逻辑表达方式，如程序语言，递归结构配置文件等，编写量化指标计算过程的递归过程。
4. 解析编写好的量化指标计算过程，运行并获取结果。

图3 量化因子计算过程构建流程



资料来源：海通证券研究所管理

由上述过程可知，在构建量化指标计算过程时，如何获取初始量化指标计算过程，定义哪些量化计算算子，用什么方式编写、记录量化计算过程的递归逻辑，是最重要的三个问题。后文我们将从这三点出发，介绍不同类型的量化因子是如何构建的。

3. 基于日频数据的量化因子构建

3.1 日频初始量化指标计算过程

从数据频率上看，日频数据是构建量化因子最常用的数据类型。以交易日为单位，市场每天会产生大量公开的价格及其衍生数据。此外，财报相关的低频率基本面数据，

也可以较为方便地加工成基于交易日的日频数据，用于因子计算。

表 2 常用日频基础数据列表

数据名称	数据周期类型	描述
价量数据		
开盘价 (OPEN)	交易日	股票每个交易日的开盘价
收盘价 (CLOSE)	交易日	股票每个交易日的收盘价
最高价 (HIGH)	交易日	股票每个交易日的最高价
最低价 (LOW)	交易日	股票每个交易日的最低价
前收盘价 (PRE_CLOSE)	交易日	股票前一交易日的收盘价按当日复权因子复权后价格
复权因子 (ADJFACTOR)	交易日	股票每个交易日的向后复权因子
成交量 (VOLUME)	交易日	股票每个交易日的成交量
成交额 (AMOUNT)	交易日	股票每个交易日的成交额
价格衍生数据		
总市值 (VAL_MV)	交易日	股票每个交易日的总市值
流通市值 (DQ_MV)	交易日	股票每个交易日的流通市值
市盈率 (PE)	交易日	股票每个交易日的市盈率
滚动市盈率 (PE_TTM)	交易日	股票每个交易日的滚动市盈率
市净率 (PB)	交易日	股票每个交易日的市净率
换手率 (TURN)	交易日	股票每个交易日的换手率
自由流通换手率 (FREETURNOVER)	交易日	股票每个交易日的自由流通换手率
一致预期净利润 (NET_REVENUE_FORCAST)	交易日	股票每个交易日的分析师一致预期净利润
非固定周期指标		
总股本 (TOT_SHARE)	公告变动日	每个股票不同日期公告的总股本
A 股流通股本 (FLOAT_A_SHARE)	公告变动日	每个股票不同日期公告的 A 股流通股本
自由流通股本 (FREE_SHARES)	公告变动日	每个股票不同日期公告的自由流通股本
净资产 (NET_ASSETS)	公告变动日	每个股票不同日期公告的净资产
不定期调入调出指标		
特别处理标志 (IS_ST)	调入调出日期	股票不定期被确认为 ST 或 ST 摘帽
所属申万一级行业代码 (SW_INDUS_I_CODE)	调入调出日期	股票被不定期调入或调出某申万一级行业
所属中信一级行业代码 (CITICS_INDUS_I_CODE)	调入调出日期	股票被不定期调入或调出某中信一级行业
会计期指标		
滚动营业收入 (OPER_REV_TTM)	财报期	股票滚动营业收入
滚动归属母公司净利润 (NET_PROFIT_PARENT_COMP_TTM)	财报期	股票滚动归属母公司净利润
资产净收益率 (ROE_TTM)	财报期	股票滚动资产净收益率

资料来源：Wind，海通证券研究所

上表列举了较为常用的几种不同周期的日频基础数据。一般情况下，日频因子的初始量化指标计算过程便是通过设置不同参数提取上述指标的过程。

量化因子计算的常见操作过程为，处理以交易日为时间序列、不同标的为截面维度的数据类型。上表中不同周期类型的数据，除交易日类型的价量数据、价格衍生数据外，其他基本面数据，如财报数据、不定期公告数据，均可以利用最近数值填充法，将其处理为以交易日为时间序列的面板数据。

3.2 日频量化指标计算算子

作为处理量化指标的原子运算单元，量化指标的面板数据结构决定了算子主要由时间序列算子与截面数据算子两种形式构成。常用的算子如下表所示。

表 3 常用量化指标计算算子

算子名称	算子类型	描述
最大值 (MAX)	时间序列	计算过去 N 个交易日指标最大值
最小值 (MIN)	时间序列	计算过去 N 个交易日指标最小值
均值 (MEAN)	时间序列	计算过去 N 个交易日指标均值
标准差 (SD)	时间序列	计算过去 N 个交易日指标标准差
偏度 (SKEWNESS)	时间序列	计算过去 N 个交易日指标偏度值
峰度 (KURTOSIS)	时间序列	计算过去 N 个交易日指标峰度值
半衰期加权求和 (DECAY_LINEAR)	时间序列	计算过去 N 个交易日某指标以距今交易日数为半衰期的加权求和值
延迟 (DELAY)	时间序列	计算某指标延迟 N 个交易日值
延迟做差 (DELTA)	时间序列	计算某指标延迟 N 个交易日值与当日的差值
累乘 (PROD)	时间序列	计算过去 N 个交易日指标累乘值
累加 (SUM)	时间序列	计算过去 N 个交易日指标累加值
分位数 (RANK)	时间序列/截面	计算过去 N 个交易日 同交易日所有标的指标分位数值
累计最小值 (ACC_MIN)	时间序列	计算过去 N 个交易日指标累计最小值
累计最大值 (ACC_MAX)	时间序列	计算过去 N 个交易日指标累计最大值
信息系数 (IC)	时间序列	计算过去 N 个交易日两指标间信息系数
秩信息系数 (RANK_IC)	时间序列	计算过去 N 个交易日两指标间秩信息系数
协方差 (COV)	时间序列	计算过去 N 个交易日两指标间协方差
BETA (BETA)	时间序列	计算过去 N 个交易日 A 指标与 B 指标线性回归系数
0,1 判别 (IF_ELSE)	截面	计算同交易日所有标的指标是否满足参数条件
初等函数运算	截面	利用类似 $x \Rightarrow x[0] + x[1]$ 初等函数公式, 对于同交易日所有标的的不同指标进行计算

资料来源：海通证券研究所整理

算子的输入和输出都必须为面板数据。上述算子包含了处理面板数据时，用到的绝大多数计算方法。通过上述算子的递归迭代，基本可以满足目前主要因子计算所需的因子计算过程。

3.3 利用计算机语言表达量化计算过程的递归逻辑

量化计算过程即为一种递归过程。通过某种可以描述递归逻辑的表达范式，可以编写范式脚本，描述预先定义好的递归逻辑过程，并通过专用程序解析范式脚本，进行数据计算。

计算机程序语言天然被设计用以描述这样的逻辑范式。由于 λ 算子与图灵机、递归函数等价，因此无论是基于 λ 算子的函数式编程语言，还是基于图灵机的过程、命令范式语言，都可以用来构建量化计算过程。

选择计算机语言作为量化计算过程的逻辑描述工具有两种方案可供选择，一是完全借鉴现有的 R、Matlab 等传统量化研究中比较常用的语言，构建其语法的解释器，并实现上述算子与初始量化指标计算过程。二是完全构建一套新的计算机语言，专门用于量化计算过程的逻辑范式描述，最大程度地保证对该语言的控制力。

出于快速实现的考虑，本文采用第二种方式，即构建了一套自己的语言，通过构建量化计算过程的方式，在服务器中计算并维护 BARRA 风格因子。

表 4 BARRA 风格因子定义

因子名称	算法描述
Size 市值	股票当日市值
Beta	过去 252 个交易日股票超额收益（股票收益-无风险收益）相对中证全指超额收益加权最小二乘法回归的 BETA 值
Momentum 动量	过去 252 个交易日 log 超额收益与过去 21 个交易日 log 超额收益的差值
DASTD	过去 252 个交易日股票相对于无风险收益率的超额收益的加权二阶中心距
CMRA	过去 252 个交易日, 最大累计月 log 超额收益与最小累计月 log 超额收益的差值
HSIGMA	过去 252 个交易日股票超额收益（股票收益-无风险收益）相对中证全指超额收益加权最小二乘法回归的残差值标准差
Residual/Volatility 残差波动率	$0.74 * DASTD + 0.16 * CMRA + 0.1 * HSIGMA$
NonLinearSize 非线性市值	市值的三次方
PB 市净率	市净率
PE 市盈率	市盈率
STOM	过去 21 个交易日 log 换手率之和
STOQ	过去 63 个交易日 log 换手率之和
STOA	过去 252 个交易日 log 换手率之和
Liquidity 流动性	$0.35 * STOM + 0.35 * STOQ + 0.3 * STOA$
EPFWD	未来一年盈利预测除以总资产
CETOP	滚动经营现金流除以总资产
ETOP	滚动归母净利润除以总资产
Earnings Yield 盈利收益率	$0.68 * EPFWD + 0.21 * CETOP + 0.11 * ETOP$
EGRSF	未来一年预测利润相对于目前利润增速
EGRO	过去四年利润回归斜率除以利润均值
SGRO	过去四年销售额回归斜率除以销售额均值
Growth 成长性	$0.17 * EGRSF + 0.3 * EGRO + 0.53 * SGRO$
MLEV	(市值+非流动性负债)/市值
DTOA	总资产/总负债
BLEV	(所有者权益+非流动性负债)/所有者权益
Leverage 杠杆	$0.38 * MLEV + 0.35 * DTOA + 0.27 * BLEV$

资料来源：The Barra US Equity Model (USE4)，海通证券研究所

BARRA 风格因子计算方式如上表所示。其中，除市值、市盈率等可以直接从初始量化指标计算过程中获得，其他指标均需要经过数次量化指标计算算子的递归迭代得到。其定义方式的迭代逻辑相对复杂，但逻辑过程固定，且需要计算的中间指标与最终指标的数量均有限。对于这种固定且数量较少，但相对需要多层次复杂迭代的递归逻辑，直接将逻辑转换为程序语言是一种较高效的方式。

图 4 量化指标计算语言示例

```

EndDate=20150919
RevenueStartDate=TradeDateCompute(StartDate,-251)
StockRevenueFactor=DoubleFactorExtract(("StockRevenue": "CASE WHEN AShareEODPrices.S_DQ_AMOUNT=0 THEN null ELSE AShareEODPrices.S_DQ_CLOSE/AShareEODPrices.S_DQ_AMOUNT * 1000000000000.0 : 1", "TradeCount"))
StockFilterFactor=DoubleFactorEntitySectionCompute(StockRevenueFactor, ["StockRevenue"], "x=>IsNaN(x[0]) ? 0 : 1", "TradeCount")
StockFilterFactor=DoubleFactorTimeDurationStatistic(StockFilterFactor, ["TradeCount"], [252], ["Sum"])
StockFilterFactor=DoubleFactorSelector(StockFilterFactor, ["TradeCount"])
LiquidityFactor=DoubleFactorSelector(StockRevenueFactor, ["Liquidity"])
LiquidityCodes=DoubleFactorCodesExtract(LiquidityFactor)
LiquidityDates=DoubleFactorDatesExtract(LiquidityFactor)
Factor=DoubleFactorEntityIn(LiquidityCodes, ["L"], LiquidityDates)
Factor=DoubleFactorCombine(Factor, LiquidityFactor)
DoubleFactorClear(LiquidityFactor)
Factor=DoubleFactorEntitySectionCompute(Factor, ["L", "Liquidity"], "x=>IsNaN(x[1]) ? 0 : x[1]", "Liquidity")
Factor=DoubleFactorSelector(Factor, ["Liquidity"])
Factor=DoubleFactorTimeDurationStatistic(Factor, ["Liquidity", "Liquidity"], [21, 63, 252], ["Sum"], ["STOM", "STOQ", "STOA"], [{"TRUE", "TRUE", "TRUE"}])
Factor=DoubleFactorEntitySectionCompute(Factor, ["STOM"], "x=>x[0]==0 ? double.NaN : Log(x[0])", "STOM")
Factor=DoubleFactorEntitySectionCompute(Factor, ["STOQ"], "x=>x[0]==0 ? double.NaN : Log(x[0]/3)", "STOQ")
Factor=DoubleFactorEntitySectionCompute(Factor, ["STOA"], "x=>x[0]==0 ? double.NaN : Log(x[0]/12)", "STOA")
Factor=DoubleFactorStandardize(Factor, ["STOM", "STOQ", "STOA"])
Factor=DoubleFactorEntitySectionCompute(Factor, ["STOM", "STOQ", "STOA"], "x=>(x[0]==0 || IsNaN(x[0]) || x[1] == 0 || IsNaN(x[1]) || x[2] == 0 || IsNaN(x[2])) ? 0 : (x[0]+x[1]+x[2])/3", "Liquidity")
IndexRevenueFactor=DoubleFactorExtract(("IndexRevenue": "AIndexEODPrices.S_DQ_CLOSE/AIndexEODPrices.S_DQ_PRECLOSE-1", NULL, NULL, RevenueStartDate, EndDate))
IndexRevenueFactor=DoubleFactorEntitySectionCompute(IndexRevenueFactor, "IndexRevenue", StockRevenueCodes, ("000985.SH": "IndexRevenue"))
StockRevenueCodes=DoubleFactorCodesExtract(StockRevenueFactor)
IndexRevenueFactor=DoubleFactorCombine(StockRevenueFactor, IndexRevenueFactor)
DoubleFactorClear(IndexRevenueFactor)
RiskFreeRevenueFactor=DoubleFactorExtract(("RiskFreeRevenue": "CBondBenchmark.B_INFO_RATE/100", NULL, NULL, RevenueStartDate, EndDate, ["CBondBenchmark.B_INFO_RATE/100"]))
RiskFreeRevenueFactor=TradeDateCount(RiskFreeRevenueFactor, "TradeCount")
RiskFreeRevenueFactor=DoubleFactorEntitySectionCompute(RiskFreeRevenueFactor, "RiskFreeRevenue", "TradeCount"), "x=>Pow(1+x[0], x[1]/365)-1", "RiskFreeRevenue")
RiskFreeRevenueFactor=DoubleFactorCombine(StockRevenueFactor, RiskFreeRevenueFactor)
StockRevenueFactor=DoubleFactorCombine(StockRevenueFactor, RiskFreeRevenueFactor)

```

资料来源：海通证券研究所整理

上图是利用我们自定义的量化指标计算语言，计算 BARRA 风格因子的部分示意。其中，以 $Value = Function(Value..., Param...)$ 的形式定义量化指标计算过程。Value 即为量化指标计算过程的输入或返回，Param 为过程参数。Function 可以代表不同类型的计算过程，既可以是预先定义的基本算子，也可以是在上下文中以 $function = Function(Value..., Param...)...end\ function$ 形式定义的量化指标计算过程计算器。从语法结构看，该语言是一种典型的基于图灵机的命令范式语言。

表 5 BARRA 因子 IC 与多空收益率（2010.01-2019.04）

因子名称	IC	IC-IR	IC 胜率	RIC	RIC-IR	RIC 胜率	多空收益	多空波动	多空 IR
Size 市值	-0.054	-1.654	31.5%	-0.052	-1.304	33.3%	26.8%	17.0%	1.581
Beta	0.008	0.291	48.6%	-0.001	-0.016	45.0%	5.3%	9.9%	0.536
Momentum 动量	-0.009	-0.338	49.5%	-0.015	-0.495	46.8%	3.6%	9.6%	0.377
ResidualVolatility 残差波动率	-0.015	-0.499	40.5%	-0.043	-1.253	33.3%	4.2%	11.8%	0.357
NonLinearSize 非线性市值	-0.053	-1.616	34.2%	-0.052	-1.307	36.0%	28.0%	18.5%	1.510
PB 市净率	-0.003	-0.237	49.5%	-0.015	-0.706	39.6%	4.7%	9.4%	0.499
PE 市盈率	-0.004	-0.542	44.1%	-0.013	-0.742	39.6%	4.3%	6.3%	0.679
Liquidity 流动性	-0.099	-2.073	29.7%	-0.054	-1.405	32.4%	21.0%	14.5%	1.445
EarningsYield 盈利预期	0.010	0.736	61.3%	0.009	0.451	57.7%	7.6%	9.2%	0.831
Growth 成长性	-0.002	-0.118	41.4%	-0.001	-0.031	45.9%	0.7%	6.0%	0.109
Leverage 杠杆	-0.002	-0.127	45.0%	0.002	0.082	51.4%	0.7%	10.7%	0.063

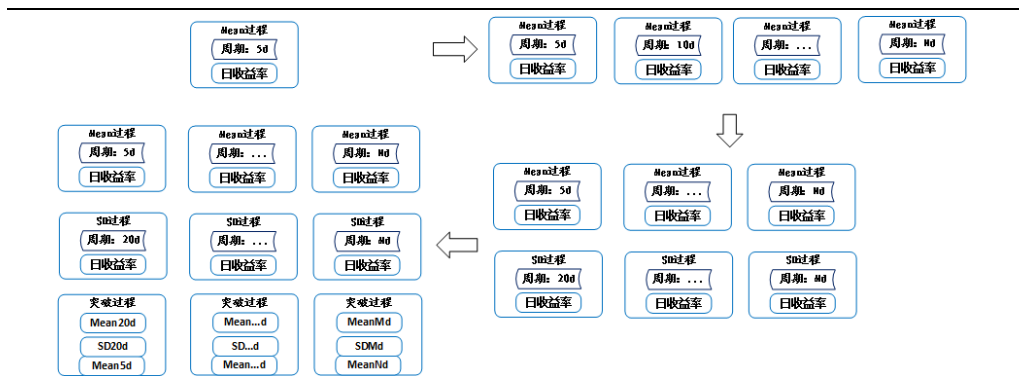
资料来源：Wind，海通证券研究所

上表罗列了我们计算的 BARRA 因子的截面表现。除市值，流动性外，绝大部分因子的收益预测能力已严重下降。因此，它们主要作为风险因子，被用来控制组合风险暴露。而寻找有效的 ALPHA 因子，则可以通过自动因子生成器的方式进行探索。

3.4 利用量化计算过程构建自动因子生成器

将因子的计算数据与逻辑单元原子化后，便可以利用这些原子单元的自动递归迭代，批量生产新的因子。具体过程如下图所示。

图5 量化因子批量生产的逻辑流程图



资料来源：海通证券研究所整理

由单一的收益率这一初始量化计算过程，叠加不同周期的均值（MEAN）算子，便可以构成 N 个短周期均值因子和 M 个长周期均值因子。利用标准差（SD）算子，构建 M 个和长周期均值因子一一对应的标准差因子。最后，将 N 个短周期均值或最新值与不同的 M 个长周期均值做差，再除以对应的标准差，便可以构建 $N*M$ 个突破因子。

假设使用的初始量化过程为 P 个，则突破因子可以有 $P*N*M$ 个。突破因子同时又可以通过别的算子与其他量化计算过程迭代，每多一个初始量化过程数据，每多一种周期，每多一个算子，均可以让生产的因子成倍增加。

由于初始量化计算过程、量化计算因子以及可计算周期为有穷，则通过这些元素生

成的因子也将是有穷的。理论上，遍历所有递归情况可以穷举所有因子，但目前并不推荐这种做法，原因主要有三。

首先，某些数据之间进行初等函数计算是没有意义的。如，成交量加减成交价格。

其次，某些算子多层次叠加并不会增加蕴含的信息量。如，计算数据截面分位数与将数据求三次方后再计算其截面分位数。

最后，穷尽所有可能性对运算资源的消耗非常严重。而且穷举得到的最终结果往往包含大量噪音，降噪过程也会极大地占用计算资源，使得因子构建得不偿失。

自动因子生成器仅通过遍历周期与算子对量化计算过程进行迭代计算，其逻辑路径判断比较简单。因此，除通过计算机语言构建递归逻辑外，还可以通过 XML 文本对递归过程进行描述。

图6 暴力因子生成的 XML 脚本示意

```
<?xml version="1.0" encoding="utf-8"?>
<Attribute>
  <MaxDate_PeriodType='YEAR' PeriodCount='1' />
  <BasePeriod>
    <PeerNode Name='CLOSE_BREAK' Type='Multi' FuncStr="x->x[2] == 0 ? 0 : (x[0]-x[1])/x[2]">
      <Nodes Seq="0">
        <BaseData Table='AShareEODPrices' SQL='[S_DQ_ADJCLOSE]' Name='ADJCLOSE' />
      </Nodes>
      <Nodes Seq="1">
        <PeerNode Name='ADJCLOSE' Type='SamePeriodSingleTimeSeriesArray'>
          <BaseData Table='AShareEODPrices' SQL='[S_DQ_ADJCLOSE]' Name='ADJCLOSE' />
          <Operators Type='Select'>
            <Operator Name='MEAN' />
            <Operator Name='SD' />
          </Operators>
          <Periods Type='Base' />
        </PeerNode>
      </Nodes>
    </PeerNode>
    <PeerNode Name='CLOSE_SHORT_BREAK' Type='Multi' FuncStr="x->x[2] == 0 ? 0 : (x[0]-x[1])/x[2]">
```

资料来源：海通证券研究所整理

XML 是计算机领域一种主流的标记计算机可理解信息的语言。以 XML 节点为单元，每一个节点又递归地包含 N 个子节点。我们以 BaseData 节点表示初始量化计算过程，以包含 FuncStr 属性的 PeerNode 节点表示包含初等函数算子的量化计算过程，以包含 Operators 与 Periods 节点的 PeerNode 节点，表示 X 个时间序列计算算子的量化计算过程。其中，X 等于时间序列算子数 N 乘以周期数 M。

我们以日收益率、换手率、自由流通换手率、PB、PE 等为初始量化计算过程，通过叠加均值，标准差算子，构建了 226 个依据以上基础数据的不同短周期相对不同长周期的突破因子。当因子 IC 或 RIC 均值的绝对值大于 0.02，且 IR 大于 1.5 时，我们认定其有效。下表给出了这一标准下，全部有效因子的平均表现。

表 6 均值突破因子截面表现 (2010.01-2019.05)

	有效因子数	有效因子 IC 均值	有效因子 IC 胜率	有效因子多空收益	有效因子多空波动	有效因子多空 IR
IC 有效因子	47/226	0.042	76.3%	13.0%	7.89%	1.634
RIC 有效因子	81/226	0.048	73.8%	11.9%	8.29%	1.441

资料来源：Wind，海通证券研究所

2010 年 1 月到 2019 年 4 月，我们构建的 226 个因子中，约有 20%到 30%的因子在截面上长期有效。考虑到这些因子均是由构建传统因子的基础数据经初等基本函数计算得出，我们将其与传统九因子正交，剥离后的效果如下表所示。

表 7 正交后均值突破因子截面表现 (2010.01-2019.05)

	有效因子数	有效因子 IC 均值	有效因子 IC 胜率	有效因子多空收益	有效因子多空波动	有效因子多空 IR
IC 有效因子	66/226	0.024	71.1%	7.6%	4.93%	1.568
RIC 有效因子	95/226	0.026	73.2%	7.1%	4.68%	1.527

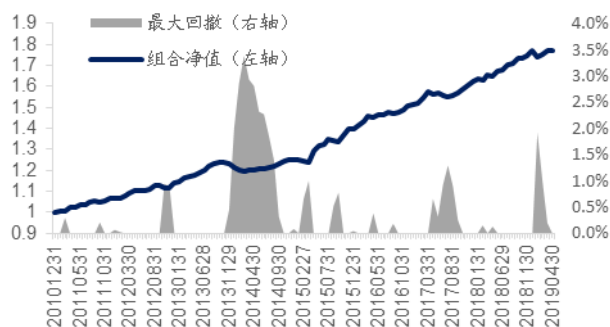
资料来源：Wind，海通证券研究所

由上表可见，由于计算所用的原始数据高度重合，剥离传统因子后，因子 IC 稳定性得到了提升，但有效性也显著下降。

我们利用两种因子筛选标准，以一年为滚动窗口构建两个因子和个股加权方式均为等权的组合。组合一选取过去一年 IC 均值大于 0.02，且 IC-IR 大于 1.5 的因子；组合二选取过去一年 IC 均值分位数大于 90%，且 IC-IR 大于 1.5 的因子。

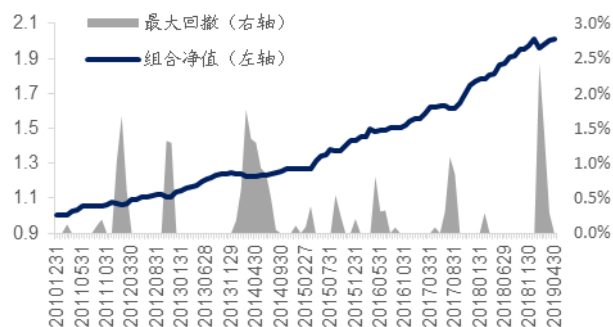
由下图可见，将筛选条件从一变为二后，多空组合的业绩表现有了明显的提高。年化收益率从 7.04% 上升到 8.65%，最大回撤从 3.39% 下降到 2.42%。

图7 组合一：多空净值 (IC>0.02)



资料来源：Wind，海通证券研究所

图8 组合二：多空净值 (IC 分位数>90%)



资料来源：Wind，海通证券研究所

由此可见，通过改变筛选因子的条件，可以有效提升组合的收益表现。但由于上述计算过程与九因子同源，正交后单因子的有效性较弱，故如何将这类由暴力生成方法得到的因子与传统因子结合使用，是发挥因子效应的关键。

以上分析表明，利用量化计算过程构建的自动因子生成器可以得到有效的因子簇，提供传统因子未包含的信息。目前为止，我们只是简单尝试了均值突破这一最直接的因子构建逻辑，就得到了令人兴奋的结果。可以期待，扩展更复杂的因子构建逻辑、增加更多有效初始量化计算过程与量化计算因子，或许会对现有组合带来更多有益的补充。

此外，由于通过这种方式构建的因子的数目相较于以往会有几何级的增长，因此优化因子筛选机制也会使组合的表现获得显著的提升。当因子维度成千上万时，简单的基于 OLS 的因子处理方法将会失效。所以，如何处理高维度的因子数据，高效地从其中提炼有效的因子信息，也是利用好自动因子生成器面临的重要课题。

4. 基于高频数据的量化因子构建

随着传统量化选股方式的广泛应用，经典日频因子的效果逐渐降低，并逐渐转化为风险因子。因此，除利用自动因子生成方式，尽可能挖掘传统基础数据的有效信息外，引入新的量化因子数据源是提供组合 ALPHA 的重要方式。自 2010 年交易所推出 LEVEL2 行情源以来，透过描述市场成交细节的日内高频行情信息，分析、寻找价格形成的细节脉络成为构建新的 ALPHA 因子的重要方式。

相较于日频数据，高频数据虽然拥有更高的数据密度与不同的数据结构，但就单个标的来看，依然可以将其定义为一种时间序列数据进行处理。因此，对应的因子计算过程的大体思路依然可以沿用上一节所述的量化因子计算过程这一框架。

4.1 高频初始量化指标计算过程

下表列示了交易所的 LEVEL2 行情产品，基于此可以构建不同种类的高频初始量化指标计算过程。

表 8 LEVEL2 高频数据列表

数据名称	最小数据间隔	描述
分钟 K 线		
开盘价 (OPEN)	分钟	分钟 K 线开盘价
收盘价 (CLOSE)	分钟	分钟 K 线收盘价
最高价 (HIGH)	分钟	分钟 K 线最高价
最低价 (LOW)	分钟	分钟 K 线最低价
成交量 (VOLUME)	分钟	分钟成交量
成交额 (AMOUNT)	分钟	分钟成交额
成交笔数 (TRADE_NUM)	分钟	分钟成交笔数
Tick 盘口快照		
成交价 (MATCH)	3 秒	每三秒最后成交价
成交量 (VOLUME)	3 秒	每三秒最后成交量
成交笔数 (TRADE_NUM)	3 秒	每三秒内成交笔数
BS 标志 (BS_FLAG)	3 秒	每三秒最后成交方向
累计成交量 (ACC_VOLUME)	3 秒	开盘到最新 Tick 累计成交量
累计成交额 (ACC_AMOUNT)	3 秒	开盘到最新 Tick 累计成交额
前收盘价 (PRE_CLOSE)	3 秒	股票前收盘价
开盘价 (OPEN)	3 秒	股票开盘价
最高价 (HIGH)	3 秒	开盘到最新 Tick 最高价
最低价 (LOW)	3 秒	开盘到最新 Tick 最低价
买卖十档行情	3 秒	每三秒盘口截面买卖十档挂单价格
平均委卖价格 (AVG_ASK_PRICE)	3 秒	每三秒盘口截面平均委卖价格
叫卖总量 (ASK_TOT_VOL)	3 秒	每三秒盘口截面总委卖量
平均委买价格 (AVG_BID_PRICE)	3 秒	每三秒盘口截面平均委买价格
叫买总量 (BID_TOT_VOL)	3 秒	每三秒盘口截面总委买量
逐笔成交		
成交编号 (INDEX)	毫秒	每笔成交编号
成交代码 (FUNC_CODE)	毫秒	每笔成交类型 (深市有意义)
BS 标志 (BS_FLAG)	毫秒	每笔成交 BS 标志
成交价格 (TRADE_PRICE)	毫秒	每笔成交价格
成交量 (TRADE_VOLUME)	毫秒	每笔成交量
叫卖序号 (ASK_ORDER)	毫秒	成交叫卖单编号
叫买序号 (BID_ORDER)	毫秒	成交叫买单编号
逐笔委托 (仅深市)		
委托编号 (INDEX)	毫秒	每笔委托编号
委托号 (ORDER)	毫秒	委托单唯一编号
委托代码 (FUNCTION_ID)	毫秒	委托代码: B、S、C
委托价格 (PRICE)	毫秒	委托单价格
委托数量 (VOLUME)	毫秒	委托单数量

资料来源: Wind, 海通证券研究所

高频与日频数据间的最大差异在于不同的时间周期与数据结构。与传统的日频数据相比,除单日数据量大增外,时间单位的不整齐也是高频数据的另一重要特征。除分钟数据严格的在每个交易日的 240 分钟均有数据外, Tick 数据、逐笔数据均会因标的成交热度不同而有较大差异,其时间序列有明显的时间间隔不均匀特性。

4.2 高频量化计算算子

从算子类型角度看,与日频数据相同,高频数据算子依然由初等函数、一阶谓词逻辑以及基本的时间序列函数和截面统计函数构成。由于高频数据具有数据规模大、时间间隔不均匀的特性,在处理过程中有以下几点需要注意。

分段的数据处理过程。由于高频数据规模很大，因此单次数据处理往往需要按标的或者按时间窗口对数据规模进行划分。一次性计算所有指标对应的所有时段数据，很有可能因为计算资源有限而无法实现。

高频数据的低频化。间隔不均匀的时间序列数据往往需要先变为均匀时间间隔的数据后，才能进一步处理。将数据处理为最小时间间隔单位可以保证不损失信息，然而却会因为空值填充给本来就需要降噪的高频数据增加更多噪音。因此，将高频数据利用统计或者抽样方式处理为低频数据是更常用的方式。

计算的时间窗口应当避免跨交易日。日内数据包含当天的市场交易观点和信息，很多时候依赖随机过程模型进行统计抽样。集合竞价的交易制度与隔夜信息会扭曲数据随机过程的特性，对于抽样后的再处理会有一定程度的干扰。因此在降低数据频率，选择计算的时间窗口时应尽可能规避跨交易日。而且，对集合竞价数据也需特别处理。

4.3 基于高频初始计算过程的高频成交因子

高频数据从一个新的维度揭示了市场的交易信息，利用前述量化计算因子，我们尝试利用逐笔成交数据构建大额买单成交金额占比因子。

首先，我们定义大单如下：

$$IsBigOrder_{i,t} = \begin{cases} 1, & OrderVol_{i,t} - Mean(OrderVol) \geq l * SD(OrderVol) \\ 0, & OrderVol_{i,t} - Mean(OrderVol) < l * SD(OrderVol) \end{cases}$$

当订单*i*的成交量减去当日所有订单成交量均值的差大于*l*倍成交量标准差时，将该订单定义为大单。其中，*l*一般取 1、2、3。

其次，定义如下的大额买单成交金额指标。

$$BidAmount_i = \left(\sum_{t=1}^T OrderPrice_{i,t} * OrderVol_{i,t} \right) * IsBidOrder_i$$

假如订单*i*全天有*T*笔成交，将每笔成交价乘以对应成交量得到单笔成交额。把该订单所有单笔成交额加总，并乘以该订单买卖标志，即可得到该买单的总成交额。

第三，将所有大额买单的成交额加总，即可得到大额买单成交金额。

$$TotalBidAmount_i = \sum_{i=1}^n BidAmount_i * IsBigOrder_{i,t}$$

最后，将上式以当日总成交额，得到大额买单成交金额占比。在使用该因子选股时，我们进一步将每天的因子值按月求平均，考察其截面表现（见下表）。

表 9 TotalBidAmount 因子 IC 与多空收益率（2015.08-2019.05）

因子名称	IC	IC-IR	IC 胜率	RIC	RIC-IR	RIC 胜率	多空收益	多空波动	多空 IR
1_SD	0.062	2.576	75.6%	0.072	2.804	73.3%	22.0%	9.0%	2.435
2_SD	0.055	2.815	75.6%	0.066	3.245	82.2%	19.1%	6.3%	3.035
3_SD	0.043	2.534	75.6%	0.053	2.793	77.8%	14.1%	5.0%	2.827

资料来源：Wind，海通证券研究所

由上表可见，原始因子表现良好。与常用 9 因子正交后的结果如下表所示。

表 10 正交 TotalBidAmount 因子 IC 与多空收益率 (2015.08-2019.05)

因子名称	IC	IC-IR	IC 胜率	RIC	RIC-IR	RIC 胜率	多空收益	多空波动	多空 IR
1_SD	0.040	3.158	80.0%	0.048	3.320	75.6%	14.3%	4.4%	3.225
2_SD	0.041	3.171	80.0%	0.049	3.543	82.2%	14.0%	4.5%	3.141
3_SD	0.036	2.850	80.0%	0.043	3.240	77.8%	12.5%	4.0%	3.135

资料来源: Wind, 海通证券研究所

因子正交后, IC 有所下滑, 但稳定性显著提升。这对我们用历史的 IC 预测未来, 更有价值。

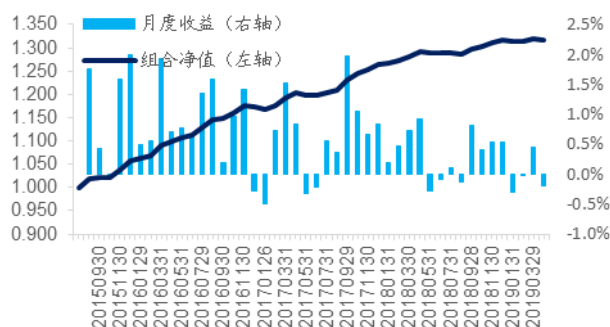
表 11 正交 TotalBidAmount 因子的截面溢价 (2015.08-2019.05)

	市值	非线性 市值	换手	非流动性	反转	波动	PB	盈利	增长	1_SD	2_SD	3_SD
月均溢价	-0.34%	0.23%	-0.70%	0.31%	-0.44%	0.49%	-0.17%	0.16%	0.12%	0.40%		
T 统计量	-1.251	3.419	-3.946	1.900	-3.096	5.231	-1.847	2.227	2.100	5.892		
为正比率	48.9%	66.7%	26.7%	62.2%	35.6%	82.2%	33.3%	55.6%	73.3%	80.0%		
月均溢价	-0.33%	0.22%	-0.70%	0.31%	-0.44%	0.49%	-0.18%	0.18%	0.12%		0.40%	
T 统计量	-1.237	3.390	-3.961	1.887	-3.091	5.172	-1.907	2.236	2.106		6.093	
为正比率	48.9%	66.7%	26.7%	62.2%	35.6%	82.2%	33.3%	55.6%	73.3%		80.0%	
月均溢价	-0.33%	0.22%	-0.70%	0.31%	-0.44%	0.49%	-0.17%	0.17%	0.12%			0.36%
T 统计量	-1.240	3.384	-3.973	1.894	-3.088	5.195	-1.864	2.225	2.092			5.620
为正比率	48.9%	66.7%	26.7%	62.2%	35.6%	82.2%	33.3%	55.6%	73.3%			80.0%

资料来源: Wind, 海通证券研究所

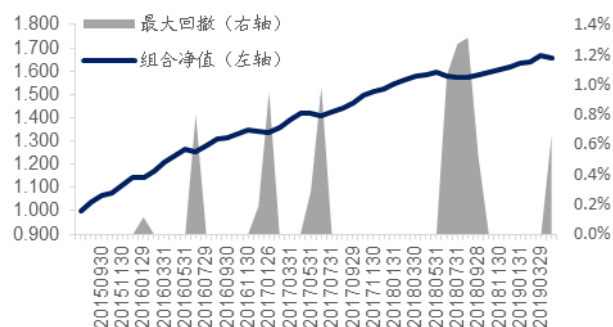
上表展示了大额买单成交金额占比因子在常见因子外, 所能贡献的收益。总的来看, 三个因子均存在稳定、显著的正向截面溢价。

图9 TotalBidAmount 等权组合超额收益



资料来源: Wind, 海通证券研究所

图10 TotalBidAmount 等权组合多空收益



资料来源: Wind, 海通证券研究所

根据该因子构建等权组合（见上图），2015年8月至今，年化收益20.1%，最大回撤1.31%。相对全市场等权组合的年化超额收益为10.53%，最大回撤0.78%，表现十分出色。

高频因子的计算复杂度非常高，而且需要对于存录的高频数据进行清洗、整理等操作，是一种工作量较大的因子计算方式。因此，在现有的基础设施条件下，我们尚未采取自动挖掘的方式对高频数据构建因子。

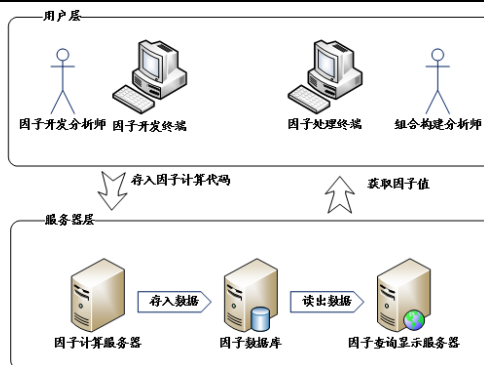
此外，由于数据整理进度的限制，基于逐笔数据的高频因子仅向前回溯至2015年8月。因此，目前只能对最近3年多的效果进行回测，因子稳定性有待进一步考察和检验。

高频数据中所隐藏的有效信息非常丰富，且与传统日频因子相关性较低。所以，无论是分钟、Tick 盘口还是逐笔成交数据，都可以让我们从中获取大量有用的信息。有关高频领域因子的深入探索，我们会在未来选股因子系列研究专题中进行更新与跟踪。

5. 因子自动管理系统框架

上文详细介绍了如何从基础数据出发，最终计算出量化因子的全部过程，在实际应用中，保证因子的“开发、计算、验证、维护、使用”这一流程，能够在一个高效的自动化系统中完成，是因子计算过程落地实践，并最终帮助我们构建组合的重要课题。

图11 因子管理框架示意图



资料来源：海通证券研究所整理

如上图所示，因子管理框架由服务器层与用户层两部分构成。用户层即分析师定义因子，编写量化指标计算过程的开发终端。服务器层主要包括因子计算服务器、因子数据库与因子查询显示服务器三类。

因子自动计算维护的过程有以下几点需要注意。

因子须有固定数据结构。通常来讲，一个带有 code，dateTime 字段的数据序列即可以标识某标的在某一时刻的一组因子值。关系数据库目前依然是存储因子值最常用的方法，因而保证因子关系数据库结构的稳定是服务器层可以进行自动管理的核心要素。

存储因子计算过程而不只存储因子值。利用计算机语言或者 XML 进行因子计算过程的逻辑定义既保证了因子计算过程可以用文本的形式进行存储，又实现了计算机的自动解析运行。存储因子计算过程可以有效保证因子计算的一致性，避免分析师在使用、分析、修改现有因子值时，产生歧义。

因子计算服务器需要有自构建 SQL，并进行因子数据库管理的能力。增删查改是关系数据库基本的操作类型，无论哪种功能都是以 SQL 语句为媒介进行的。下图展示了建表的 SQL 语句示例。

图12 因子关系数据库中，用于建表的 SQL 语句示例

```
CREATE TABLE [dbo].[sec_kline_rt_201104] (
    [sec_code] [nvarchar](50) NULL,
    [date] [nvarchar](50) NULL,
    [time] [nvarchar](50) NULL,
    [close_price] [decimal](18, 4) NULL,
    [trade_amount] [decimal](18, 4) NULL,
    [trade_value] [decimal](18, 4) NULL,
    [trade_num] [decimal](18, 4) NULL,
    [no_valid] [decimal](18, 4) NULL,
    [high_price] [decimal](18, 4) NULL,
    [low_price] [decimal](18, 4) NULL,
    [open_price] [decimal](18, 4) NULL
) ON [PRIMARY]
GO
```

资料来源：海通证券研究所

SQL 语句是一种逻辑高度一致化的语言，可由程序根据需要自动构建和运行。目前，已有非常多的工具可以依据用户定义，自动构建 SQL 语句进行数据库管理。出于功能的可扩展性，自己编写代码实现 SQL 语句生成器也是一种可选的解决方案。

因子计算服务器在识别出预计算因子结构后，可以直接将所计算因子的数据结构与因子数据库中现有的数据进行对比，检查是否为新的数据或数据结构已发生更改。由于关系数据库建表与删表的速度远高于数据表修改，因此当发现结构不匹配时，删除老数据表，与新数据一样重新构建数据表是一种更为高效的办法。

6. 总结

作为构建量化多因子组合的重要工具，因子可以通过计算机进行全自动的生产与管理，从而让因子的使用者摆脱管理、更新因子数据的繁重工作，将时间与精力投入到有效因子的挖掘与构建中。

因子计算只是构建多因子组合的第一步。在获得因子值后，后续还有如因子有效性评价，因子的过滤和筛选，因子权重配比，组合权重配比等大量工作。以量化的方法处理其中任何一步，都会有大量的数据计算过程。而这些过程的绝大多数，都可以用一阶谓词逻辑与初等函数计算的递归过程进行定义。即，可以将上述步骤中的绝大部分工作拆解为计算机可识别、可处理的模块进行自动化的计算与管理。如何利用计算机工具，进一步在量化研究和投资过程中提升自动化水平，将是本系列未来所关注的重点方向。

7. 风险提示

数据挖掘是从历史先验数据获取经验模型的方法，存在模型失效可能。

信息披露

分析师声明

冯佳睿 金融工程研究团队
余浩淼 金融工程研究团队

本人具有中国证券业协会授予的证券投资咨询执业资格，以勤勉的职业态度，独立、客观地出具本报告。本报告所采用的数据和信息均来自市场公开信息，本人不保证该等信息的准确性或完整性。分析逻辑基于作者的职业理解，清晰准确地反映了作者的研究观点，结论不受任何第三方的授意或影响，特此声明。

法律声明

本报告仅供海通证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议。在任何情况下，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。

本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。

市场有风险，投资需谨慎。本报告所载的信息、材料及结论只提供特定客户作参考，不构成投资建议，也没有考虑到个别客户特殊的投资目标、财务状况或需要。客户应考虑本报告中的任何意见或建议是否符合其特定状况。在法律许可的情况下，海通证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

本报告仅向特定客户传送，未经海通证券研究所书面授权，本研究报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。如欲引用或转载本文内容，务必联络海通证券研究所并获得许可，并需注明出处为海通证券研究所，且不得对本文进行有悖原意的引用和删改。

根据中国证监会核发的经营证券业务许可，海通证券股份有限公司的经营经营范围包括证券投资咨询业务。

海通证券股份有限公司研究所

路颖 所长
(021)23219403 luying@htsec.com

高道德 副所长
(021)63411586 gaodd@htsec.com

姜超 副所长
(021)23212042 jc9001@htsec.com

邓勇 副所长
(021)23219404 dengyong@htsec.com

荀玉根 副所长
(021)23219658 xyg6052@htsec.com

涂力磊 所长助理
(021)23219747 tll5535@htsec.com

宏观经济研究团队

姜超(021)23212042 jc9001@htsec.com
于博(021)23219820 yb9744@htsec.com
李金柳(021)23219885 lj11087@htsec.com
联系人
宋潇(021)23154483 sx11788@htsec.com
陈兴(021)23154504 cx12025@htsec.com

金融工程研究团队

高道德(021)63411586 gaodd@htsec.com
冯佳睿(021)23219732 fengjr@htsec.com
郑雅斌(021)23219395 zhengyb@htsec.com
罗蕾(021)23219984 ll9773@htsec.com
沈泽承(021)23212067 szc9633@htsec.com
余浩淼(021)23219883 yhm9591@htsec.com
袁林青(021)23212230 ylq9619@htsec.com
姚石(021)23219443 ys10481@htsec.com
吕丽颖(021)23219745 lly10892@htsec.com
周一洋(021)23219774 zyy10866@htsec.com
张振岗(021)23154386 zzg11641@htsec.com
联系人
颜伟(021)23219914 yw10384@htsec.com
梁镇(021)23219449 lz11936@htsec.com

金融产品研究团队

高道德(021)63411586 gaodd@htsec.com
倪韵婷(021)23219419 niyt@htsec.com
陈瑶(021)23219645 chenyaoyao@htsec.com
唐洋运(021)23219004 tangyy@htsec.com
宋家骥(021)23212231 sjj9710@htsec.com
皮灵(021)23154168 pl10382@htsec.com
徐燕红(021)23219326 xyh10763@htsec.com
谈鑫(021)23219686 tx10771@htsec.com
王毅(021)23219819 wy10876@htsec.com
蔡思圆(021)23219433 csy11033@htsec.com
庄梓恺(021)23219370 zzk11560@htsec.com
联系人
谭实宏(021)23219445 tsh12355@htsec.com
吴其右(021)23154167 wqy12576@htsec.com

固定收益研究团队

姜超(021)23212042 jc9001@htsec.com
朱征星(021)23219981 zzx9770@htsec.com
周霞(021)23219807 zx6701@htsec.com
姜珊珊(021)23154121 jps10296@htsec.com
杜佳(021)23154149 dj11195@htsec.com
联系人
李波(021)23154484 lb11789@htsec.com

策略研究团队

荀玉根(021)23219658 xyg6052@htsec.com
钟青(010)56760096 zq10540@htsec.com
高上(021)23154132 gs10373@htsec.com
李影(021)23154117 ly11082@htsec.com
姚佩(021)23154184 yp11059@htsec.com
周旭辉 zxx12382@htsec.com
张向伟(021)23154141 zxw10402@htsec.com
李姝醒(021)23219401 lsx11330@htsec.com
曾知(021)23219810 zz9612@htsec.com
联系人
唐一杰(021)23219406 tyj11545@htsec.com
郑子勋(021)23219733 zzx12149@htsec.com
王一潇(021)23219400 wyx12372@htsec.com

中小市值团队

张宇(021)23219583 zy9957@htsec.com
钮宇鸣(021)23219420 ymniu@htsec.com
孔维娜(021)23219223 kongwn@htsec.com
潘莹练(021)23154122 pyl10297@htsec.com
联系人
程碧升(021)23154171 cbs10969@htsec.com
相姜(021)23219945 xj11211@htsec.com

政策研究团队

李明亮(021)23219434 lml@htsec.com
陈久红(021)23219393 chenjiuhong@htsec.com
吴一萍(021)23219387 wuyiping@htsec.com
朱蕾(021)23219946 zl8316@htsec.com
周洪荣(021)23219953 zhr8381@htsec.com
王旭(021)23219396 wx5937@htsec.com

石油化工行业

邓勇(021)23219404 dengyong@htsec.com
朱军军(021)23154143 zjj10419@htsec.com
联系人
胡歆(021)23154505 hx11853@htsec.com
张璇(021)23219411 zx12361@htsec.com

医药行业

余文心(0755)82780398 ywx9461@htsec.com
郑琴(021)23219808 zq6670@htsec.com
贺文斌(010)68067998 hwb10850@htsec.com
联系人
范国钦(021)232154384 fgq12116@htsec.com
梁广楷(010)56760096 lgk12371@htsec.com
吴佳桂 0755-82900465 wjs11852@htsec.com

汽车行业

王猛(021)23154017 wm10860@htsec.com
杜威(0755)82900463 dw11213@htsec.com
联系人
曹雅倩(021)23154145 cyq12265@htsec.com

公用事业

吴杰(021)23154113 wj10521@htsec.com
张磊(021)23212001 zl10996@htsec.com
戴元灿(021)23154146 dyc10422@htsec.com
联系人
傅逸帆(021)23154398 fty11758@htsec.com

批发和零售贸易行业

汪立亭(021)23219399 wanglt@htsec.com
李宏科(021)23154125 lkh11523@htsec.com
联系人
史岳 sy11542@htsec.com
高瑜(021)23219415 gy12362@htsec.com
谢茂莹 xmx12344@htsec.com

互联网及传媒

郝艳辉(010)58067906 hyh11052@htsec.com
孙小雯(021)23154120 sxw10268@htsec.com
毛云聪(010)58067907 myc11153@htsec.com
联系人
陈星光(021)23219104 cxg11774@htsec.com

有色金属行业

施毅(021)23219480 sy8486@htsec.com
联系人
陈晓航(021)23154392 cxh11840@htsec.com
甘嘉尧(021)23154394 gjy11909@htsec.com

房地产行业

涂力磊(021)23219747 tll5535@htsec.com
谢盐(021)23219436 xiey@htsec.com
金晶(021)23154128 jj10777@htsec.com
杨凡(021)23219812 yf11127@htsec.com

电子行业

陈 平(021)23219646 cp9808@htsec.com
尹 岑(021)23154119 yl11569@htsec.com
谢 磊(021)23212214 xl10881@htsec.com
联系人
石 坚(010)58067942 sj11855@htsec.com

煤炭行业

李 淼(010)58067998 lm10779@htsec.com
戴元灿(021)23154146 dyc10422@htsec.com
吴 杰(021)23154113 wj10521@htsec.com
联系人
王 涛(021)23219760 wt12363@htsec.com

电力设备及新能源行业

张一弛(021)23219402 zyc9637@htsec.com
房 青(021)23219692 fangq@htsec.com
曾 彪(021)23154148 zb10242@htsec.com
徐柏乔(021)23219171 x bq6583@htsec.com
联系人
陈佳彬(021)23154513 cjb11782@htsec.com

基础化工行业

刘 威(0755)82764281 lw10053@htsec.com
刘海荣(021)23154130 lhr10342@htsec.com
张翠翠(021)23214397 zcc11726@htsec.com
孙维容(021)23219431 swr12178@htsec.com
联系人
李 智(021)23219392 lz11785@htsec.com

计算机行业

郑宏达(021)23219392 zhd10834@htsec.com
杨 林(021)23154174 yl11036@htsec.com
鲁 立(021)23154138 ll11383@htsec.com
于成龙 ycl12224@htsec.com
黄竞晶(021)23154131 hjj10361@htsec.com
洪 琳(021)23154137 hl11570@htsec.com

通信行业

朱劲松(010)50949926 zjs10213@htsec.com
余伟民(010)50949926 ywm11574@htsec.com
张 弋 01050949962 zy12258@htsec.com
张峥青(021)23219383 zzq11650@htsec.com

非银行金融行业

孙 婷(010)50949926 st9998@htsec.com
何 婷(021)23219634 ht10515@htsec.com
联系人
李芳洲(021)23154127 l fz11585@htsec.com

交通运输行业

虞 楠(021)23219382 yun@htsec.com
罗月江 (010) 56760091 lyj12399@htsec.com
联系人
李 丹(021)23154401 ld11766@htsec.com

纺织服装行业

梁 希(021)23219407 lx11040@htsec.com
联系人
盛 开(021)23154510 sk11787@htsec.com
刘 溢(021)23219748 ly12337@htsec.com

建筑建材行业

冯晨阳(021)23212081 fcy10886@htsec.com
潘莹练(021)23154122 pyl10297@htsec.com
联系人
申 浩(021)23154114 sh12219@htsec.com

机械行业

余炜超(021)23219816 swc11480@htsec.com
耿 耘(021)23219814 gy10234@htsec.com
杨 震(021)23154124 yz10334@htsec.com
沈伟杰(021)23219963 swj11496@htsec.com
周 丹 zd12213@htsec.com

钢铁行业

刘彦奇(021)23219391 liuyq@htsec.com
刘 璇(0755)82900465 lx11212@htsec.com
联系人
周慧琳(021)23154399 zhl11756@htsec.com

建筑工程行业

杜市伟(0755)82945368 dsw11227@htsec.com
张欣劼 z xj12156@htsec.com
李富华(021)23154134 lf h12225@htsec.com

农林牧渔行业

丁 频(021)23219405 dingpin@htsec.com
陈雪丽(021)23219164 cxl9730@htsec.com
陈 阳(021)23212041 cy10867@htsec.com
联系人
孟亚琦 myq12354@htsec.com

食品饮料行业

闻宏伟(010)58067941 whw9587@htsec.com
成 珊(021)23212207 cs9703@htsec.com
唐 宇(021)23219389 ty11049@htsec.com

军工行业

蒋 俊(021)23154170 jj11200@htsec.com
刘 磊(010)50949922 ll11322@htsec.com
张恒昶 zhx10170@htsec.com
联系人
张宇轩(021)23154172 zyx11631@htsec.com

银行行业

孙 婷(010)50949926 st9998@htsec.com
解巍巍 xww12276@htsec.com
林加力(021)23214395 lj12245@htsec.com
谭敏沂(0755)82900489 tmy10908@htsec.com

社会服务行业

汪立亭(021)23219399 wanglt@htsec.com
陈扬扬(021)23219671 cyy10636@htsec.com
许樱之 xyz11630@htsec.com

家电行业

陈子仪(021)23219244 chenzy@htsec.com
李 阳(021)23154382 ly11194@htsec.com
朱默辰(021)23154383 zmc11316@htsec.com
联系人
刘 璐(021)23214390 ll11838@htsec.com

造纸轻工行业

衣楦永(021)23212208 yzy12003@htsec.com
赵 洋(021)23154126 zy10340@htsec.com

研究所销售团队

深广地区销售团队

蔡铁清(0755)82775962 ctq5979@htsec.com
伏财勇(0755)23607963 fcy7498@htsec.com
辜丽娟(0755)83253022 gulj@htsec.com
刘晶晶(0755)83255933 liujj4900@htsec.com
王雅清(0755)83254133 wyq10541@htsec.com
饶 伟(0755)82775282 rw10588@htsec.com
欧阳梦楚(0755)23617160
oymc11039@htsec.com
巩柏舍 gbh11537@htsec.com

上海地区销售团队

胡雪梅(021)23219385 huxm@htsec.com
朱 健(021)23219592 zhuj@htsec.com
李唯佳(021)23219384 lijw@htsec.com
黄 毓(021)23219410 huangyu@htsec.com
漆冠男(021)23219281 qgn10768@htsec.com
胡宇欣(021)23154192 hyx10493@htsec.com
黄 诚(021)23219397 hc10482@htsec.com
毛文英(021)23219373 mwy10474@htsec.com
马晓男 mxn11376@htsec.com
杨祎昕(021)23212268 yyx10310@htsec.com
张思宇 zsy11797@htsec.com
慈晓聪(021)23219989 cxc11643@htsec.com
王朝领 wcl11854@htsec.com
邵亚杰 23214650 syj12493@htsec.com
李 寅 021-23219691 ly12488@htsec.com

北京地区销售团队

殷怡琦(010)58067988 yyq9989@htsec.com
郭 楠 010-5806 7936 gn12384@htsec.com
张丽莹(010)58067931 zlx11191@htsec.com
杨羽莎(010)58067977 yys10962@htsec.com
杜 飞 df12021@htsec.com
张 杨(021)23219442 zy9937@htsec.com
何 嘉(010)58067929 hj12311@htsec.com
李 婕 lj12330@htsec.com
欧阳亚群 oyyq12331@htsec.com

海通证券股份有限公司研究所
地址：上海市黄浦区广东路 689 号海通证券大厦 9 楼
电话：(021) 23219000
传真：(021) 23219392
网址：www.htsec.com