

因子拥挤度专题 基于**BARRA**风格 因子构建**MSCI**因 子拥挤度构建

科大财经

大纲：

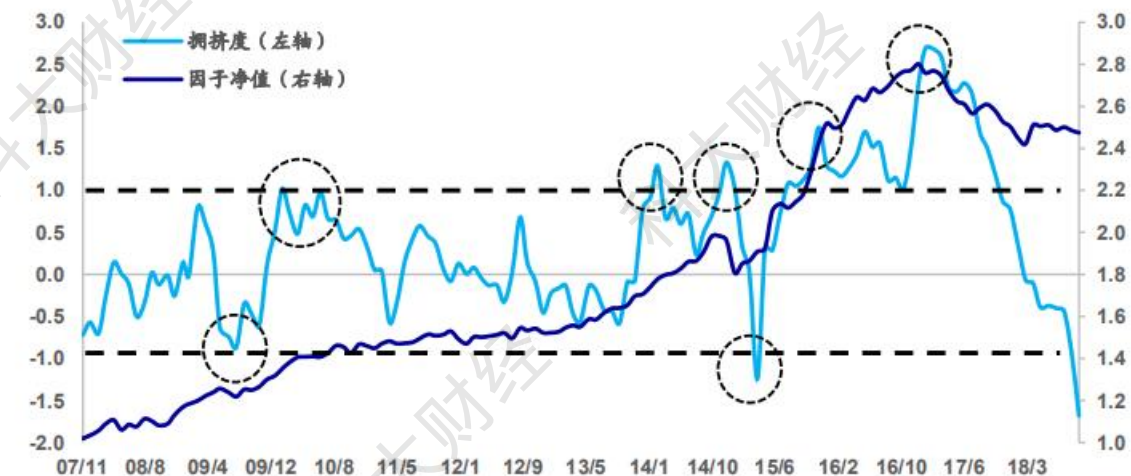
1. MSCI因子拥挤度概念介绍
 1. MSCI估值价差
 2. MSCI配对相关性
 3. MSCI因子波动率
 4. MSCI因子反转
2. 以BARRA动量因子为例构建拥挤度指标
3. 构建所有BARRA风格因子拥挤度直播
4. BARRA拥挤度指标预测（机器学习）

因子失效与因子拥挤度

因子失效也逐渐成为因子投资者重点研究的方向。在相关研究文献中，因子拥挤现象

(Factor Crowding) 常和因子的失效联系在一起。相关研究认为，有效的因子会受到资金的追捧，虽然资金的追捧会增强因子的收益，但是过多资金的追捧会导致因子的波动加大，并出现收益的减弱以及大幅回撤。因此，度量因子拥挤度就显得至关重要。通过因子拥挤度的监控，投资者可有效规避或者控制过度拥挤的因子给组合带来大幅回撤。

图25 小市值因子净值与复合拥挤度



资料来源: Wind, 海通证券研究所

(1) 估值价差

资金对于因子的追捧会进一步推升因子多头端的估值或者压低因子空头端的估值水平，由此加大因子多空组合的估值价差。估值价差指标越高，则说明该时点上因子拥挤度越高。

分别计算多空组合的相对估值指标的中位数，相对估值指标本文取通用的P/E(市盈率)与P/B(市净率)，为了解决估值指标在收益接近0时取值接近无穷等不良情况，本文取估值指标的倒数E/P和B/P，进行估值的刻画。

估值价差 = 空头组合的估值指标倒数中位数 - 多头组合的估值指标倒数中位数

总结：市净率表现会整体优于市盈率

(2) 配对相关性

配对相关性从股票同涨同跌的特征来度量因子的拥挤程度。资金对于因子的追捧会加剧这一类股票同涨同跌的特性，因此配对相关性指标越大则因子越拥挤。

在截面观测日回看3个月，将多空头组合内的成分股日收益等权，计算多空组合的日收益。分别计算组内成分股和多空头组合的日收益相关系数，并取均值，计算多空头配对相关性。

+ 多头配对相关性 = Mean(Corr(个股 i 3 个月, 多头组合均收益 3 个月))

+ 空头配对相关性 = Mean(Corr(个股 i 3 个月, 空头组合均收益 3 个月))

+ 配对相关性指标 = 多头配对相关性 + 空头配对相关性

总结：空头配对相关性对于空头的剥离效果较好，多头配对相关性对于多头剥离效果较好

(3) 因子波动率

由于资金的流入会加剧因子收益率的波动，因此Bonne等(2018)从使用因子波动率衡量当前因子拥挤程度。本文中，在截面观测日回看3个月，将多空组合内的成分股日收益等权，计算多空组合的日收益，并计算多空组合日收益序列的标准差。将历史时点上的A股市场上符合要求的股票等权组合，构建市场组合，用同样的方法计算市场组合的收益并计算标准差，因此期望相对波动率度量与因子波动率正相关。

+ 多头组合因子波动 = $\text{STD}(\text{多头组合}) / \text{STD}(\text{市场组合})$

+ 多空组合因子波动 = $\text{STD}(\text{多空组合}) / \text{STD}(\text{市场组合})$

(4) 因子反转

De Bondt等(1985)表明美股在3到5年的中长期尺度上存在反转。由于投资者有追逐过去业绩的倾向，一个在一段时间内表现良好的因子可能已经积累了大量资本。这种性能追逐最初可能会为因子性能提供顺风，但最终可能会为视图或位置改变时的下降或均值回归奠定基础。异常强劲的表现通常不会永远持续下去。本文使用因子过去三年的累积收益率计算该指标。

+ 因子长期反转 = $\text{Ret}(\text{因子多空组合近三年})$

以BARRA动量因子为例构建拥挤度指标

BARRA因子研究的好处，我们可以拿到其纯因子收益率。因为其本身已经剔除了其他因素对于动量因子的影响，我们在计算因子拥挤度时，可以更具体的理解来源于本因子的收益的影响是多少。

Style: **Momentum**

Definition: 1.00 RSTR

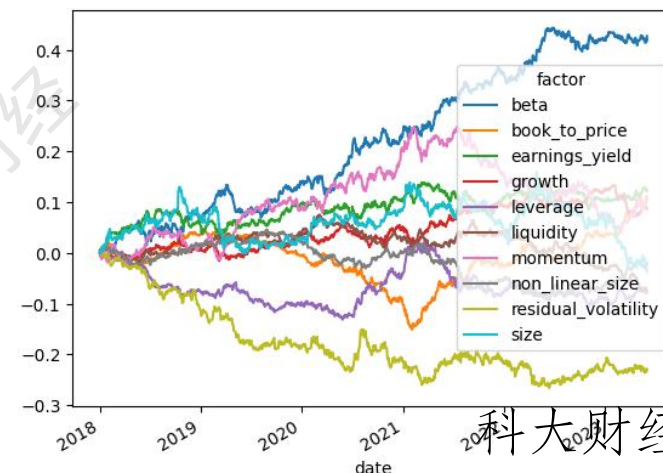
Components: RSTR

Relative strength

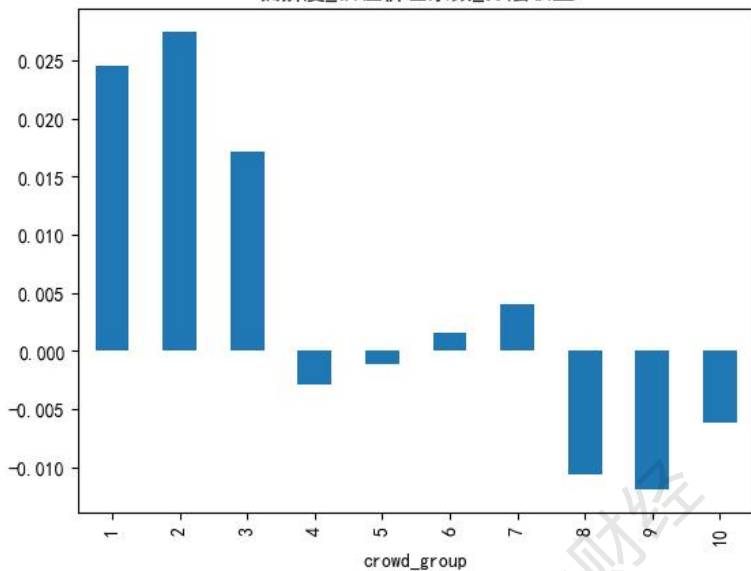
Computed as the sum of excess log returns over the trailing $T = 504$ trading days with a lag of $L = 21$ trading days,

$$RSTR = \sum_{t=L}^{T+L} w_t [\ln(1 + r_t) - \ln(1 + r_{ft})] , \quad (2)$$

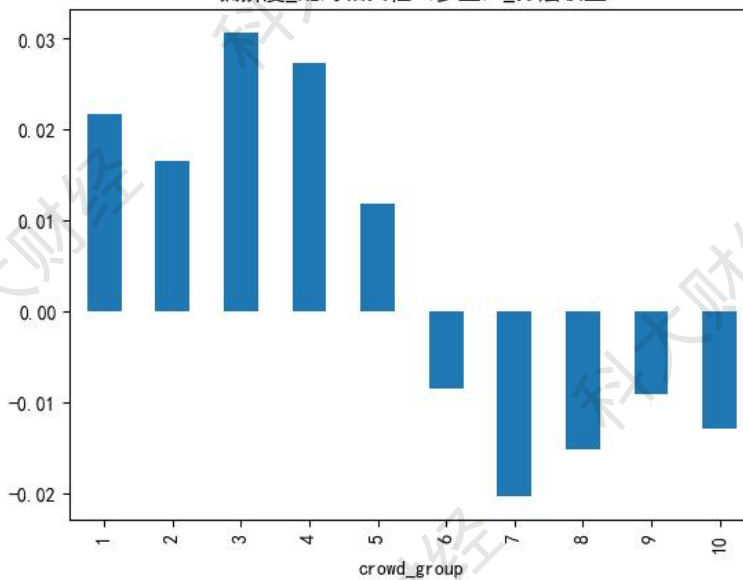
where, r_t is the stock return on day t , r_{ft} is the risk-free return, and w_t is an exponential weight with a half-life of 126 trading days.



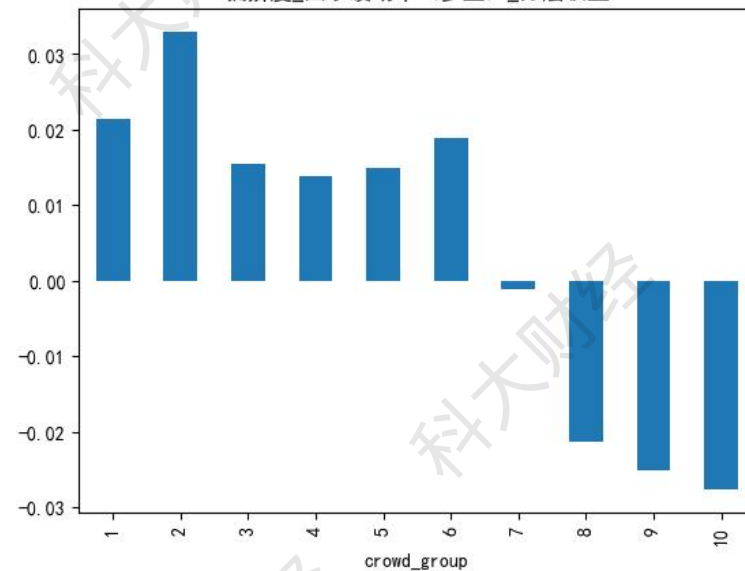
拥挤度_估值价差系数_分层收益



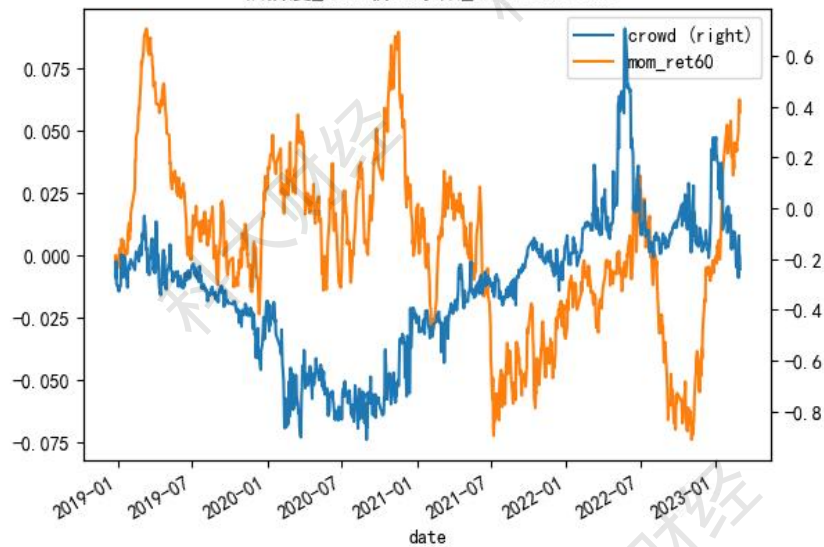
拥挤度_配对相关性（多空）_分层收益



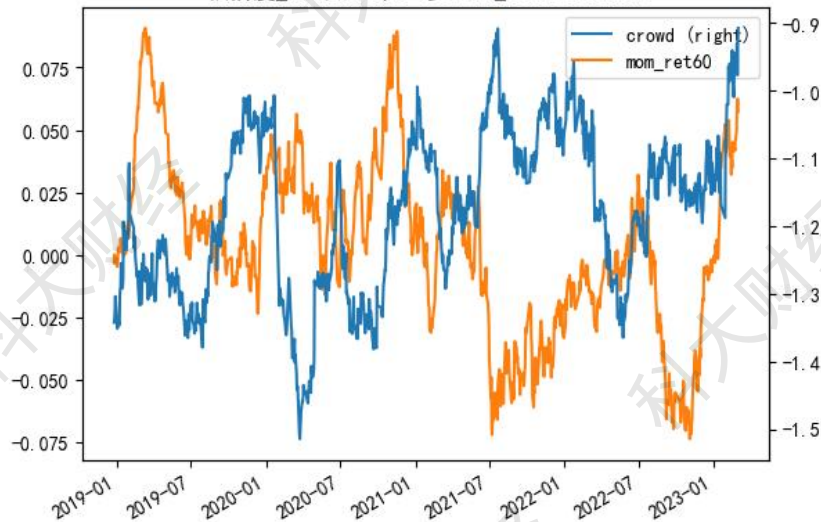
拥挤度_因子波动率（多空）_分层收益



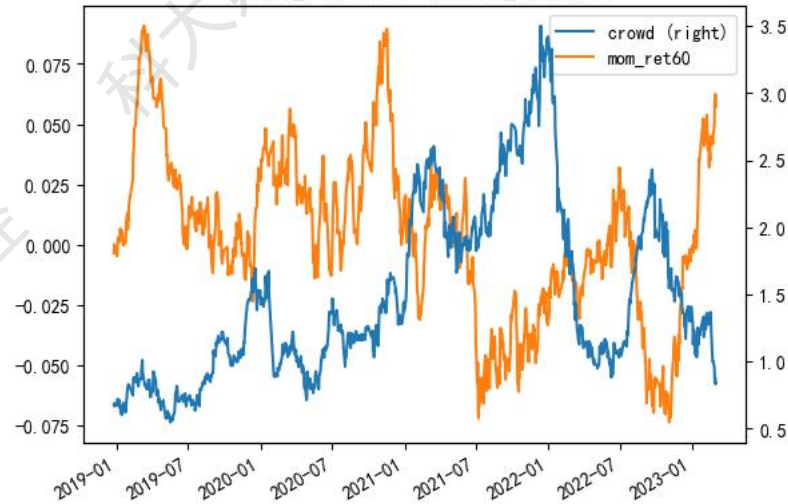
拥挤度_估值价差系数_未来60日收益



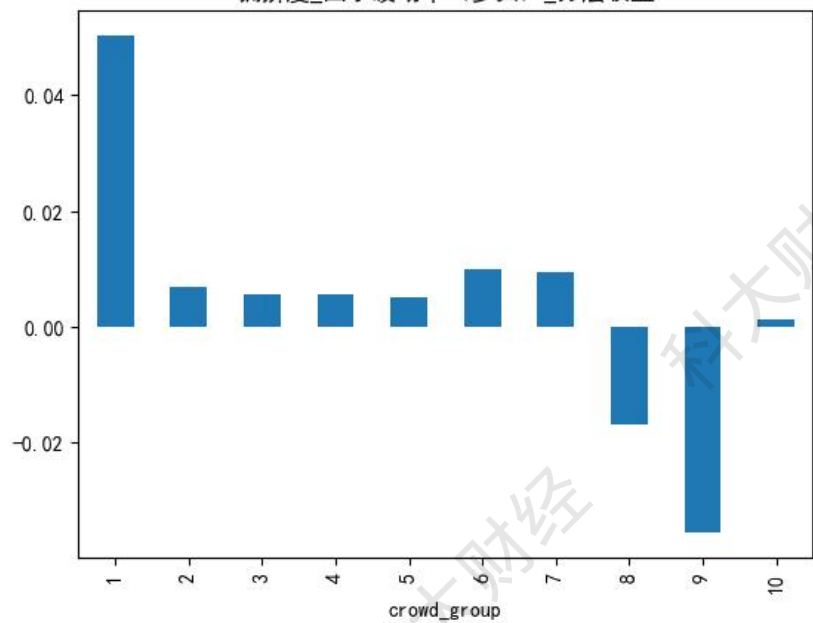
拥挤度_配对相关性（多空）_未来60日收益



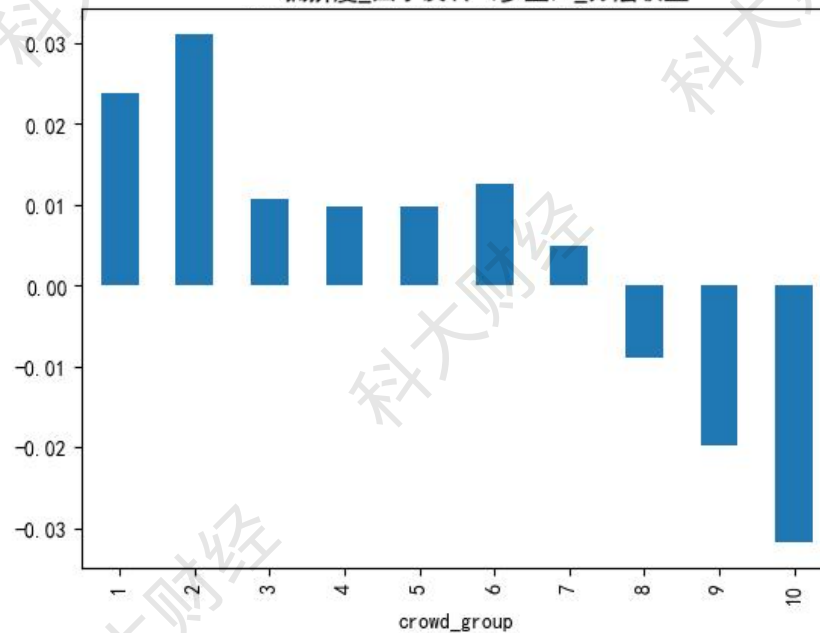
拥挤度_因子波动率（多空）_未来60日收益



拥挤度_因子波动率（多头）_分层收益



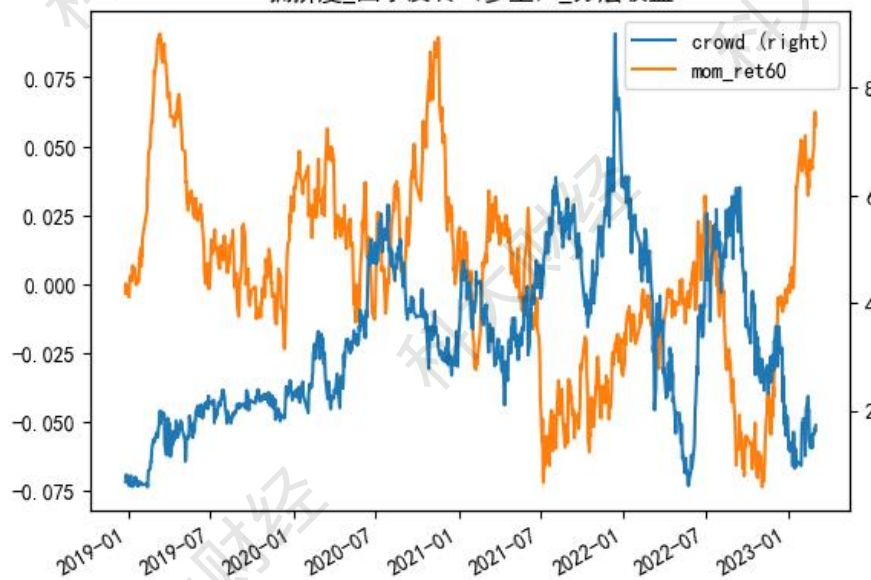
拥挤度_因子反转（多空）_分层收益



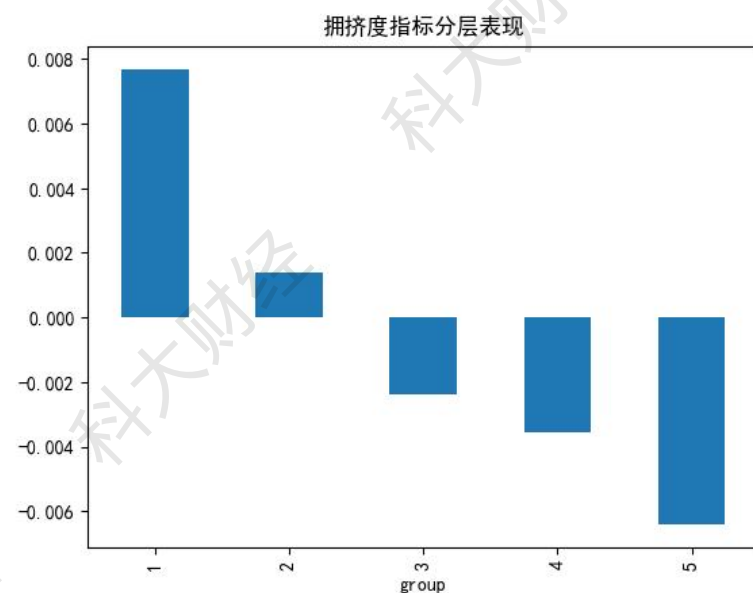
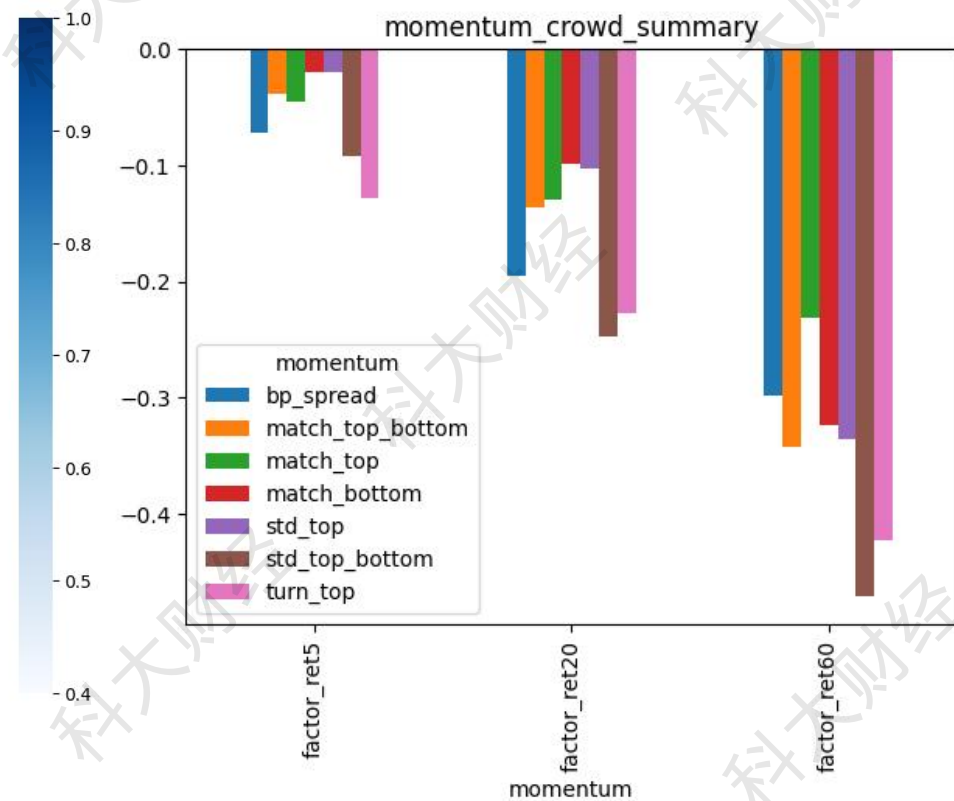
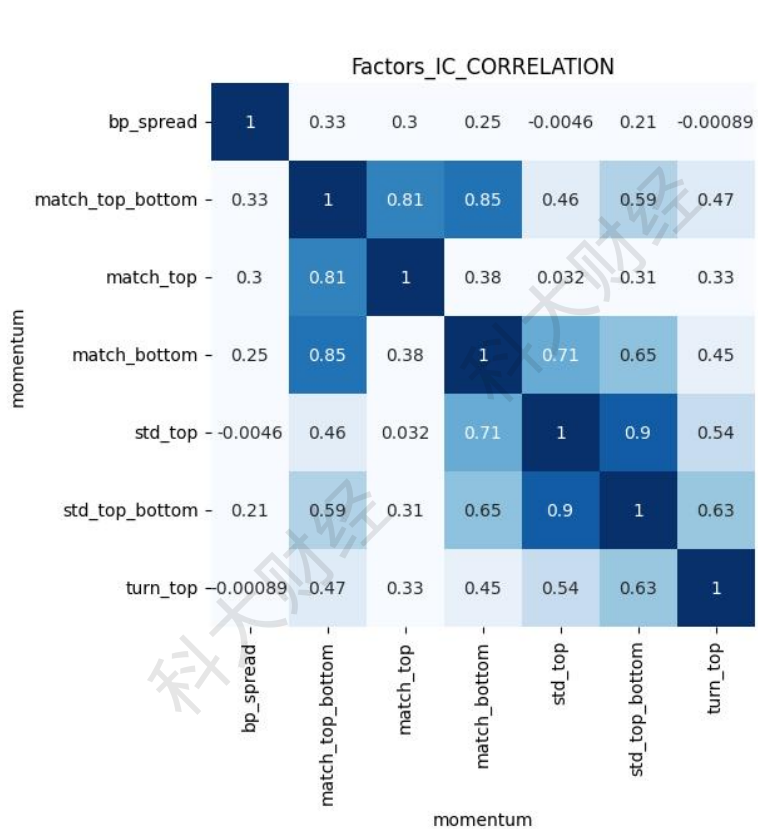
拥挤度_因子波动率（多空）_分层收益



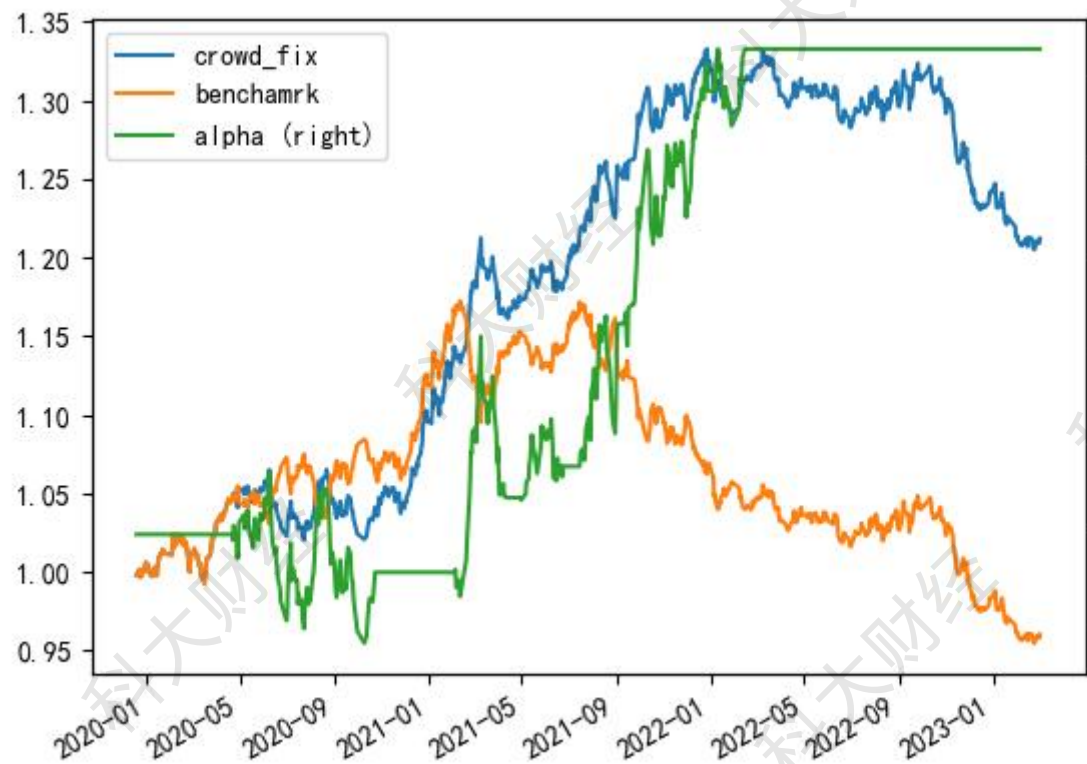
拥挤度_因子反转（多空）_分层收益



拥挤度指标汇总



可做空



不可做空



BARRA拥挤度指标预测（机器学习）

基于随机森林算法将拥挤度数据对于每个BARRA因子收益的方向进行预测

原始表现



预测表现

