

分析师:

郑兆磊

zhengzhaolei@xyzq.com.cn

S0190520080006

花开股市，相似几何五：弱关联关系下的特异性 Alpha 因子挖掘

2022 年 6 月 22 日

报告关键点

新因子挖掘一般有三种方式：利用另类数据（比如专利等）、对已有因子的逻辑进行改进和修缮（比如估值因子各种改进）、利用老数据但采用新的加工方式（比如机器学习）。本文继续深耕相似动量思想，为投资者呈现相似的魅力。在相似动量体系下，因子的生成大体上包含三步：相似向量（看什么相似），基于向量得到相似矩阵、利用月度收益率进行降维生成新的选因子。我们对三个方向均进行了改进和探索。

相关报告

《猎金系列之三十二：财报季的财务效应研究和因子构建》
2021-10-20
《基于专利分类的科技动量因子研究》2019-06-25

投资要点

- 新因子挖掘一般有三种方式：利用另类数据（比如专利等）、对已有因子的逻辑进行改进和修缮（比如对估值因子的改进）、利用老数据但采用新的加工方式（比如机器学习）。本文继续深耕相似动量思想，为投资者呈现相似的魅力。

相似向量（看什么相似）、

基于向量得到相似矩阵、利用月度收益率进行降维生成新的选因子。我们对三个方向均进行了探索。

- **相似矩阵探索**：1）对相似矩阵尾部进行降噪，以提升因子的纯度，得到的因子效果相对于未经过降噪的提升非常显著。因子 IC 均值达到 0.029，ICIR 达到 0.48，T 值达到 5.78，且因子**特异性极强**；2）对矩阵信息进行另类利用，构建的因子（简称**相似信号**）IC 达到 0.042，ICIR 达到 0.68，T 值高达 8.18，且**特异性极强**。
- **降维方式探索**：除了利用月度收益率进行降维外，我们引入基金持仓。以财务作为向量为例，最终计算的基金持仓动量因子（简称为**持仓动量**）表示：以某考察标的为例，财务相似的公司基金持有的越多，接下来该标的将涨的越多。因子 IC 达到 0.032、ICIR 达到 0.54、T 值高达 6.3，分位数各组别严格单调，且因子**特异性极强**。
- **相似向量探索 1**：相似的基础实际上是相似向量。在往期报告中我们尝试了专利动量因子、财务动量因子。以**专利动量**因子为例，该因子报告发布于 2019 年 6 月份，样本外因子 IC 达到 0.023、ICIR 达到 0.5，T 值达到 2.8。分位数各组别相对严格单调，多头收益高达 29.3%。因子**样本外极佳、特异性极强**。
- **相似向量探索 2**：除了利用专利分布作为底层向量，我们选择其他因子作为底层向量构成。复合因子 IC 达到 0.035、ICIR 高达 0.57、T 值达到 6.78，因子分位数测试各组别严格单调，且因子**特异性极强**。
- **相似的极限**：考察以上维度生成的新因子（专利动量、财务动量 N、相似信号、持仓动量），四个因子相关性几乎接近 0。将四个因子等权合成：复合因子 IC0.053、ICIR 达到 0.79、T 值高达 9.5。因子分位数测试各组别严格单调。即便在中证 800 内，效果几乎没有衰弱。另外因子的**特异性强**。

风险提示: 本报告模型及结论全部基于对历史数据的分析,当市场环境变化时,存在模型失效风险。

请务必阅读正文之后的信息披露和重要声明

目录

1、引言	4 -
2、相似动量方法论回顾及改进维度介绍	5 -
2.1、相似动量方法论回顾	5 -
2.2、相似动量因子改进维度介绍	7 -
3、矩阵应用改进	8 -
3.1、相似矩阵应用改进一---删除尾部噪音	8 -
3.2、相似矩阵应用改进二---矩阵的另类利用	11 -
4、降维方法改进	12 -
5、相似底层向量改进	15 -
5.1、相似向量尝试改进一---专利动量因子	15 -
5.2、相似向量尝试改进二---持仓动量因子	19 -
6、相似方法论本质探讨和总结	21 -
6.1、百尺竿头更进一步---因子归结篇	21 -
6.2、相似的本质探索	25 -
6.3.....	26 -

图表目录

图表 1、兴证金工部分新因子展示	4 -
图表 2、不同维度的代表性因子	6 -
图表 3、F-Momentum 以及行业市值中性化后因子 IC 测试	6 -
图表 4、行业市值中性化后 F-Momentum 因子分位数测试	6 -
图表 5、F-Momentum 与各选股因子相关性分布图	7 -
图表 6、F_Momentum 因子生成示意图	8 -
图表 7、删除不同噪音水平下的 F-Momentum IC 测试	8 -
图表 8、删除不同噪音水平下的 F-Momentum 分位数测试对比	9 -
图表 9、F-Momentum Vs. F-Momentum-80Pec 因子多头净值	9 -
图表 10、F-Momentum Vs. F-Momentum-80Pec 因子多空净值	10 -
图表 11、删除不同噪音水平下的 F-Momentum 在宽基范围内 IC 测试	10 -
图表 12、删除不同噪音水平下的 F-Momentum 在沪深 300 内分位数测试对比	10 -
图表 13、删除不同噪音水平下的 F-Momentum 在中证 800 内分位数测试对比	11 -
ew 因子 IC 测试结果	11 -
图表 15、Link_New 因子分位数测试结果	11 -
图表 16、Link_New 因子分位数测试净值表现	12 -
图表 17、Link_New 与各选股因子相关性分布图	12 -
图表 18、Fund-Moment 因子 IC 测试结果	13 -
图表 19、Fund_Moment 因子分位数测试结果	13 -
图表 20、行业市值中性化后 Fund_Moment 因子分位数测试结果	14 -
图表 21、行业市值中性化后 Fund_Moment 因子净值	14 -
图表 22、行业市值中性化后 Fund_Moment 因子和底层所有因子相关性分析	15 -
图表 23、剥离相关性最高的因子后 Fund-Moment 因子 IC 测试结果	15 -
图表 24、IPC 一级分类示意图	16 -
图表 25、专利 IPC 分类原始数据示意	16 -
图表 26、IE-Momentum 因子分位数测试	17 -
图表 27、IE-Momentum 因子多头&多空净值曲线	18 -

图表 28、IE -Momentum 因子 IC 测试.....	- 18 -
图表 29、IE-Momentum 因子与各选股因子相关性分布图	- 18 -
图表 30、删除不同噪音水平下的 Sig-Momentum IC 测试	- 19 -
图表 31、行业市值中性化后 Sig-Momentum 因子分位数测试结果	- 20 -
图表 32、行业市值中性化后 Sig-Momentum 因子分位数净值	- 20 -
图表 33、行业市值中性化后 Sig_Moment 因子分位数净值.....	- 21 -
图表 34、各因子逻辑简介	- 21 -
图表 35、各因子对比表现-IC	- 22 -
图表 36、各因子相关性测试	- 22 -
图表 37、Link-Comb-Sig IC 测试	- 22 -
图表 38、Link-Comb-Sig 因子分位数测试结果—全市场	- 23 -
图表 39、Link-Comb-Sig 因子分位数测试结果—中证 800 内	- 23 -
图表 40、Link-Comb-Sig 因子分位数净值—全市场	- 23 -
图表 41、Link-Comb-Sig 因子分位数净值—中证 800 内	- 24 -
图表 42、Link-Comb-Sig 因子相关性测试	- 24 -
图表 43、Link-Comb-Sig 与 Barra 风格因子相关性测试	- 25 -

用户1671154545于2023-07-10日下载，仅供本人内部使用，不可传播与转载

报告正文

1、引言

量化选股的基础是因子，但目前新因子是一件可遇不可求的事情。在因子同质化比较严重的情况下，投资者更多的是在因子的选择和合成方式方面开始进行探索，这方面的收获亦颇为丰富：比如近年来通过机器学习进行因子选择和因子合成就是其中尤为关注的方向。

但我们不得不承认，基于逻辑演绎的新因子始终是最受欢迎最易接受的。而兴证金工在这方面一直不遗余力，孜孜不倦的展开新因子的探索。图表 1 展示了我们 2019 年以来撰写的部分新因子。纵观这些因子的逻辑，有些是从基本面量化的角度出发、有些是从另类数据出发、有些是机器学习的角度出发等等。这些新因子均是得到了市场的认可，同时样本外表现亦颇为优秀。

图表 1、兴证金工部分新因子展示

因子名称	发布日期	释义
专利相关因子	2018/12/20	我们对专利领域进行了深入的研究，构建了众多专利选股因子，样本外一直有跟踪，并且在诸多泛科技策略中进行应用
IE_Momentum	2019/06/25	计算任意两个公司专利分布相似、然后转化为专利动量因子，因子选股特异性极强
Adaboost_Sig	2019/06/21	利用 Adaboost 算法构建选股因子
EP_DR	2020/03/20	个股估值不仅仅可以从截面看，还可以从时序看，构建全新的时序估值因子（基于市盈率）
BP_DR	2020/03/20	逻辑同 EP_DR，基于市现率
SP_DR	2020/03/20	逻辑同 EP_DR，基于市销率
DP_Fwd12M	2020/03/29	以 Lintner 股利平稳评价模型为基础，对公司股利向目标股利调整速度进行估计，而后依据卖方分析师的一致预期净利润对股利预测，最终完成预期股息率因子（DP_Fwd12M）的构建
MPD	2020/04/30	基于 B-S-M 期权定价模型构造的公司违约概率因子，该因子最高的一组能够稳定跑输市场
OEP_TTM	2020/06/08	巴菲特在 1986 年提出的股东盈余概念，从净利润计算过程中所存在的问题出发，构建全新的估值因子
TLSUE	2020/10/23	基于公司实际净利润与历史净利润构造的 SUE 因子，反映公司业绩的超预期程度
F_Momentum	2021/10/20	启发于专利动量因子，我们从四大财务能力出发，选择代表因子，构建了全新的财务动量因子，特异性极强
Search_Sig	2021/12/06	得到个股网络搜索指数，并进一步构建网络热度因子，因子有一定的特异性，效果颇佳
SWF	2021/12/07	基于个股仓位测算结果，构建基本因子、横截面持仓分歧因子、时序均值因子、持仓相似度动量因子，最终将 4 个因子合成为个股仓位大类因子 SWF

资料来源：Wind、聚源、兴业证券经济与金融研究院整理

上述新因子客户复现度高，反馈好，同时我们兴证金工团队也在不同的策略中或多或少的使用了各个因子。比如基于时序估值偏离度构建的 EP_DR/BP_DR/SP_DR 在我们的沪深 300、中证 500 增强、科技医药基本面量化策略中均扮演重

请务必阅读正文之后的信息披露和重要声明

要角色。

在上述反馈较好的因子中,尤为需要注意的是 IE_Momentum(专利动量因子)以及 F_Momentum(财务动量因子)。两者均是从信号合成角度出发,利用余弦相似度衡量任意两个公司的信息相似程度,然后进一步转化为相应的动量因子。这种全新的因子构建方式得到了客户广泛的认可。我们兴证金工团队也在深入思考该方法论的逻辑,希望能进一步利用该体系。而这正是本报告的焦点。

报告框架如下:

第二章:相似动量方法论回顾和改进介绍;

第三章:相似矩阵改进;

第四章:相似动量降维方式改进;

第五章:相似向量尝试和改进;

第六章:相似方法论本质的探讨和总结;

2、相似动量方法论回顾及改进维度介绍

2.1、相似动量方法论回顾

新因子的挖掘一般有三种方式:利用另类数据(比如专利、产业链、供应链、网络文本数据等)、对已有因子的逻辑进行改进和修缮(比如前面提到的股东盈余等)、利用老数据但采用新的加工方式(比如目前非常流行的机器学习)。而相似动量实际上就是第三种方法的极致演绎。在《基于专利分类的科技动量因子研究》中,我们对相似进行了初步的尝试:通过余弦相似度(当然亦可以尝试其他方法)去捕捉任意两个公司的信号相似大小。如果结果越像,我们认为两个公司在信号上越接近,具体参见公式(1)。

$$\cos(\theta) = \frac{\sum_{i=1}^n A_i \cdot B_i}{\sqrt{\sum_{i=1}^n A_i^2} \cdot \sqrt{\sum_{i=1}^n B_i^2}} \quad (1)$$

公式(1)中 A_i 及 B_i 是向量空间 $\mathbf{A}(\mathbf{A} = (A_1, A_2, \dots, A_n))$ 及 $\mathbf{B}(\mathbf{B} = (B_1, B_2, \dots, B_n))$ 的细分维度。

在此基础上,我们构建了财务动量因子(《猎金系列之三十二:财报季的财务效应研究和因子构建》)。该因子从企业的盈利能力、偿债能力、营运能力和发展能力四个维度出发,对企业进行系统、全面、综合的分析,从而对企业的财务状况、经营成果和现金流量等作出整体的评价与判断。我们在每个维度选择代表性指标(部分指标参见图表2)形成一个一维向量,用该向量代表企业的财务属性,进而将专利动量的思路转化成财务动量因子(简称为 F_Momentum)。

图表 2、不同维度的代表性因子

因子名称	代表能力	释义
CurrentRatio_LR	偿债能力	流动资产合计_最新财报 / 流动负债合计_最新财报
Revenue_SQ_YoY	发展能力	单季度营业收入同比增长率
ROE_TTM_YoY	盈利能力	ROE_TTM / 去年同期的 ROE_TTM - 1
ReceivablesTurnover	营运能力	2*销售收入_TTM/(应收账款_最新财报+应收账款_去年同期)

资料来源：Wind、聚源、兴业证券经济与金融研究院整理

财务动量因子具有较强的选股能力:行业市值中性化后因子 IC 达到 0.026、T 值达到 4.9，分位数各组别严格单调，具体参见图表 3-4。

图表 3、F-Momentum 以及行业市值中性化后因子 IC 测试

	平均值	标准差	最小值	最大值	ICIR	T 统计量
F-Momentum	0.027	0.12	-0.24	0.32	0.23	2.7
中性化 F-Momentum	0.026	0.06	-0.10	0.21	0.41	4.9

资料来源：

备注：回测区间：2009 年 12 月 31 日-2021 年 12 月 31 日

图表 4、行业市值中性化后 F-Momentum 因子分位数测试

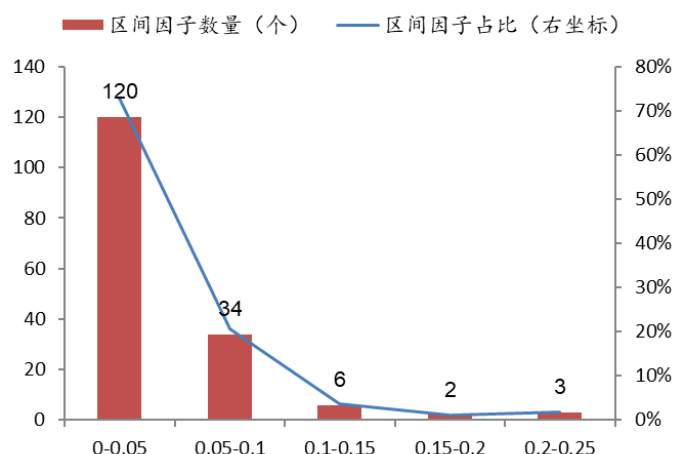
	年化收益率	波动率	Sharpe 比率	平均换手率	最大回撤率	年化超额收益率	跟踪误差
Top	11.0%	29.6%	0.37	153.4%	51.5%	0.8%	15.7%
1	12.1%	28.9%	0.42	171.7%	52.0%	1.7%	46.6%
2	10.2%	28.4%	0.36	175.0%	49.8%	-0.2%	-6.5%
3	10.4%	28.2%	0.37	175.2%	50.2%	-0.1%	-3.0%
4	10.8%	28.2%	0.38	174.0%	47.3%	0.3%	10.6%
5	9.0%	27.4%	0.33	174.0%	49.5%	-1.5%	-43.7%
6	6.8%	27.7%	0.25	174.7%	52.6%	-3.5%	-101.9%
7	6.2%	28.0%	0.22	174.6%	53.9%	-4.0%	-128.5%
8	4.9%	28.5%	0.17	170.4%	58.1%	-5.0%	-117.0%
Bottom	-0.4%	29.6%	-0.02	151.0%	62.0%	-9.7%	-188.2%
市场	10.4%	28.6%	0.36		44.9%		
L-S	11.1%	7.7%	1.45		12.3%		

资料来源：Wind、聚源、兴业证券经济与金融研究院整理

备注：回测区间：2009 年 12 月 31 日-2021 年 12 月 31 日

同时该因子具有极强的特异性。我们通过相关性分析（测试该因子和兴证金工底层所有因子相关性）、Fama-Mac 回归测试、合成方式探索（用余弦相似度合成与其他合成方式合成进行比对）、对兴证金工指数增强的改进作用四个维度对该因子特异性进行了全方位的分析。结果显示该因子的特异性极强（具体参见报告《猎金系列之三十二：财报季的财务效应研究和因子构建》）。

图表 5、F-Momentum 与各选股因子相关性分布图



资料来源: Wind、聚源、兴业证券经济与金融研究院整理

备注: 回测区间: 2009 年 12 月 31 日-2021 年 12 月 31 日

2.2、相似动量因子改进维度介绍

我们这里以财务动量因子为例来展示整个相似动量因子的流程图。相似动量因子的生成大体上包含三步（具体参见图表 6），其中经过第二步得到的是相似矩阵（图表 6 Step2），再进一步得到相似动量因子（图表 6 Step3）。那么针对于该因子的改进就可以从这三个步骤入手，大致如下：

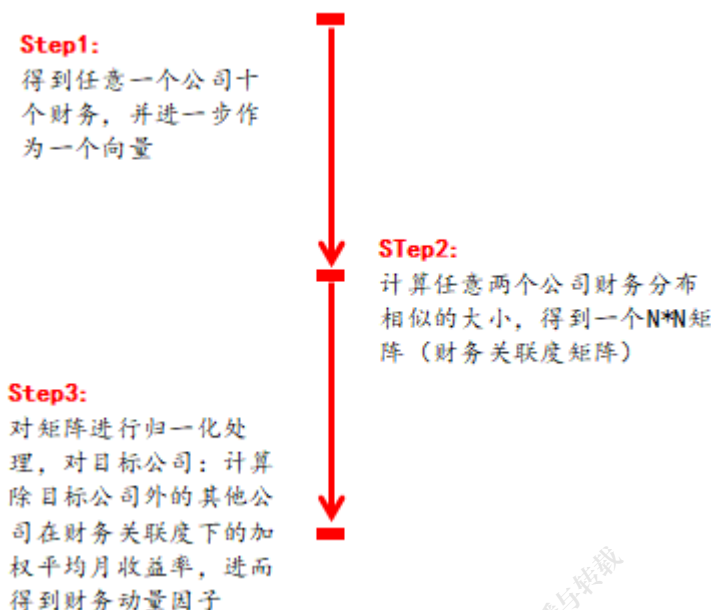
方式一（step2）：1）、以 2022 年 1 月底贵州茅台为例，在计算贵州茅台月底的财务动量时，我们将市场其余在市的 4630 只股票全部纳入（2022 年 1 月 28 日全市场在市的股票 4631 只），计算除贵州茅台外的其他公司在财务关联度下的加权平均月收益率。但实际上确实有很多公司和贵州茅台的财务分布极其不相似，这些股票的纳入反而会成为一种噪音，进一步影响最终的财务动量因子的稳健性。可否通过一些手段将这些股票删除，进而合成财务动量因子呢；2）、除了前面这种降噪之外，矩阵本身信息亦非常丰富，是否可以加以利用呢？

图表 6 中的第三步实际上是通过降维的方式得到了一个选股

因子。那么可否通过其他方式对图表 6 中第二步生成的矩阵进行降维进而形成一个选股因子呢？

方式三（step1）：在相似体系下，依赖什么信息去度量相似是整个研究的基础和核心。在财务动量中（参见团队 2021 年发布的《猎金系列之三十二：财报季的财务效应研究和因子构建》）我们输入的是财务信息作为向量；在专利动量中（参见团队 2019 年发布的《基于专利分类的科技动量因子研究》），我们输入的是专利分布，那么是否可以输入其他维度的信息呢？这里我们进行更广泛的探索。

图表 6、F_Momentum 因子生成示意图



资料来源：Wind、聚源、兴业证券经济与金融研究院整理

3、矩阵应用改进

3.1、相似矩阵应用改进一---删除尾部噪音

对于每只股票我们通过删除尾部财务分布极不相似股票，可以得到新的财务动量。这里可以设置极不相似的水平，我们共计设置了 20%、30%、50%、80%、90%的分档，得到的因子进一步进行动量、行业、市值中性化，之后得到的因子分别称之为 F-Momentum-20Pec、F-Momentum-30Pec、F-Momentum-50Pec、F-Momentum-80Pec、F-Momentum-90Pec。我们测试几个因子在 2009 年 12 月 31 日-2021 年 12 月 31 日期间的表现。

从结果看，这种方式稳健的提升因子表现：随着删除水平的提高，因子 IC、T 值均得到提升。以 F-Momentum-80Pec 为例：因子的 IC 均值达到 0.029，

T 值达到 5.78；因子分位数各组别严格单调，多空收益率达到 11.6%，夏普比达到 1.57，最大回撤 7.2%，因子整体表现显著优于之前的 F-Momentum 因子。

图表 7、删除不同噪音水平下的 F-Momentum IC 测试

	平均值	标准差	最小值	最大值	ICIR	T 统计量
F-Momentum	0.025	0.06	-0.10	0.21	0.39	4.62
F-Momentum-20Pec	0.025	0.06	-0.11	0.19	0.41	4.89
F-Momentum-30Pec	0.026	0.06	-0.10	0.20	0.40	4.86
F-Momentum-50Pec	0.027	0.06	-0.11	0.18	0.43	5.13
F-Momentum-80Pec	0.029	0.06	-0.11	0.18	0.48	5.78
F-Momentum-90Pec	0.031	0.06	-0.11	0.17	0.53	6.39

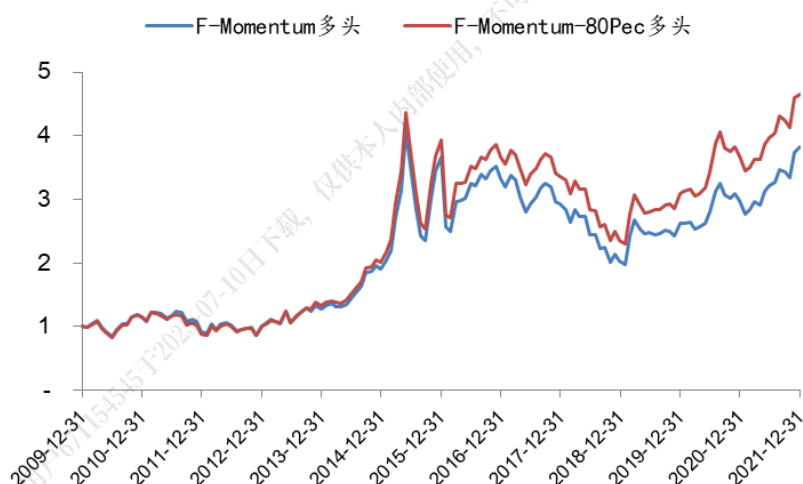
资料来源：Wind、聚源、兴业证券经济与金融研究院整理

图8、删除不同噪音水平下的 F-Momentum 分位数测试对比

分组	年化收益率	Sharpe比率	最大回撤率	分组	年化收益率	Sharpe比率	最大回撤率	分组	年化收益率	Sharpe比率	最大回撤率
F-Momentum				F-Momentum-20Pec				F-Momentum-30Pec			
Top	11.8%	0.40	51.4%	Top	12.2%	0.42	49.2%	Top	12.6%	0.43	48.1%
1	13.1%	0.46	51.3%	1	12.4%	0.44	47.9%	1	11.9%	0.42	49.5%
2	10.8%	0.38	50.2%	2	12.0%	0.42	50.3%	2	11.4%	0.40	51.1%
3	11.5%	0.41	48.9%	3	11.4%	0.41	49.0%	3	12.0%	0.43	47.3%
4	11.4%	0.41	47.7%	4	10.6%	0.39	49.7%	4	10.7%	0.39	51.4%
5	9.9%	0.36	49.3%	5	9.3%	0.34	52.2%	5	9.4%	0.34	52.3%
6	8.0%	0.29	51.8%	6	8.6%	0.31	50.7%	6	8.0%	0.29	52.0%
7	6.9%	0.25	54.7%	7	6.9%	0.25	56.8%	7	8.3%	0.30	54.6%
8	6.0%	0.21	58.0%	8	5.3%	0.19	57.6%	8	4.5%	0.16	58.8%
Bottom	0.9%	0.03	61.0%	Bottom	1.6%	0.06	60.1%	Bottom	1.7%	0.06	58.8%
L-S	10.5%	1.39	11.5%	L-S	11.1%	1.40	11.2%	L-S	10.3%	1.49	11.0%
F-Momentum-50Pec				F-Momentum-80Pec				F-Momentum-90Pec			
Top	13.3%	0.46	48.6%	Top	13.7%	0.47	47.3%	Top	13.1%	0.45	48.5%
1	11.9%	0.42	50.6%	1	12.2%	0.43	52.8%	1	14.6%	0.51	48.2%
2	12.6%	0.44	47.7%	2	12.8%	0.45	44.8%	2	11.8%	0.42	50.7%
3	12.0%	0.43	47.4%	3	11.5%	0.41	47.0%	3	11.4%	0.41	47.5%
4	10.0%	0.36	50.7%	4	11.1%	0.41	49.1%	4	10.3%	0.38	48.5%
5	8.9%	0.33	52.3%	5	8.7%	0.32	49.1%	5	10.0%	0.37	45.7%
6	8.3%	0.30	52.0%	6	7.5%	0.27	55.3%	6	7.0%	0.25	55.0%
7	7.1%	0.25	55.8%	7	6.9%	0.25	58.0%	7	6.5%	0.23	56.6%
8	5.2%	0.18	59.6%	8	4.6%	0.16	58.3%	8	3.7%	0.13	62.0%
Bottom	1.2%	0.04	59.4%	Bottom	1.5%	0.05	61.1%	Bottom	2.0%	0.07	60.1%
L-S	11.5%	1.60	7.6%	L-S	11.6%	1.57	7.2%	L-S	10.6%	1.48	6.8%

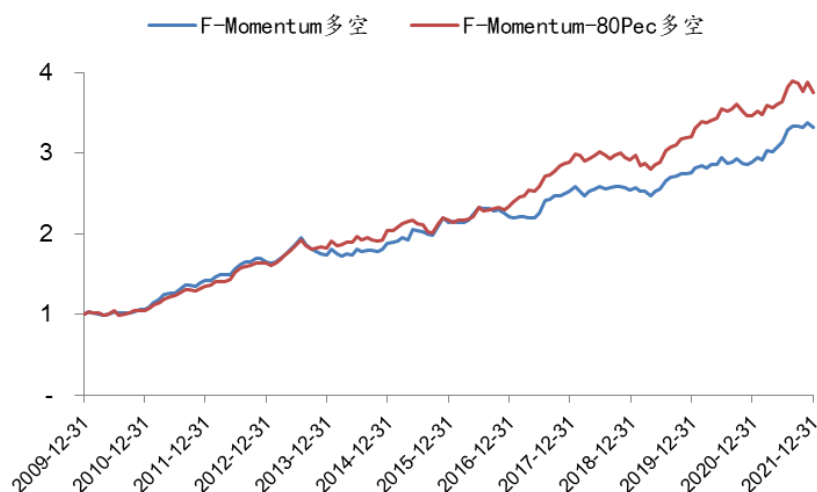
资料来源：Wind、聚源、兴业证券经济与金融研究院整理

图9、F-Momentum Vs. F-Momentum-80Pec 因子多头净值



资料来源：Wind、聚源、兴业证券经济与金融研究院整理

图 10、F-Momentum Vs. F-Momentum-80Pec 因子多空净值



资料来源：Wind、聚源、兴业证券经济与金融研究院整理

即便将测试范围缩小至宽基范围内，效果依然如此：随着删除水平的提升，在宽基范围内表现也逐步提升。我们以因子在中证 800 内的测试为例，随着删除水平的提升，因子的 IC 从 0.025 提升至 0.031，ICIR 从 0.32 提升至 0.44，T 值从 3.9 提升至 5.2，提升幅度显著。

图 11、删除不同噪音水平下的 F-Momentum 在宽基范围内 IC 测试

	沪深 300 内表现			中证 800 内表现		
	平均值	标准差	最小值	最大值	ICIR	T 统计量
F-Momentum	0.021	0.22	2.66	0.025	0.32	3.9
F-Momentum-20Pec	0.021	0.22	2.69	0.026	0.33	4.0
F-Momentum-30Pec	0.023	0.24	2.94	0.026	0.32	3.8
F-Momentum-50Pec	0.019	0.20	2.43	0.026	0.33	3.9
F-Momentum-80Pec	0.021	0.23	2.70	0.031	0.40	4.8
F-Momentum-90Pec	0.022	0.24	2.84	0.031	0.44	5.2

资料来源：Wind、聚源、兴业证券经济与金融研究院整理

图 12、删除不同噪音水平下的 F-Momentum 在沪深 300 内分位数测试对比

分组	年化收益率	Sharpe比率	最大回撤率	分组	年化收益率	Sharpe比率	最大回撤率	分组	年化收益率	Sharpe比率	最大回撤率
F-Momentum				F-Momentum-20Pec				F-Momentum-30Pec			
Top	5.8%	0.23	47.7%	Top	6.7%	0.28	46.4%	Top	7.1%	0.29	48.3%
1	3.6%	0.15	44.5%	1	4.0%	0.17	42.9%	1	3.6%	0.15	41.4%
2	5.8%	0.24	41.0%	2	2.9%	0.12	45.2%	2	4.0%	0.17	39.9%
3	4.4%	0.19	44.7%	3	3.2%	0.14	43.2%	3	2.5%	0.11	45.6%
Bottom	-1.3%	-0.05	57.4%	Bottom	1.4%	0.06	53.9%	Bottom	1.0%	0.04	53.1%
L-S	7.0%	0.80	14.6%	L-S	4%	0.57	21.7%	L-S	5.4%	0.58	22.3%
F-Momentum-50Pec				F-Momentum-80Pec				F-Momentum-90Pec			
Top	7.5%	0.31	45.4%	Top	7.7%	0.31	43.5%	Top	7.5%	0.29	45.8%
1	2.8%	0.12	43.5%	1	3.3%	0.14	44.6%	1	4.0%	0.17	44.1%
2	2.5%	0.10	44.3%	2	5.1%	0.22	38.9%	2	3.9%	0.17	40.5%
3	4.1%	0.17	43.9%	3	0.9%	0.04	52.0%	3	2.1%	0.09	51.9%
Bottom	1.4%	0.06	52.4%	Bottom	1.3%	0.05	52.1%	Bottom	0.9%	0.04	50.4%
L-S	5.6%	0.62	21.6%	L-S	6.0%	0.63	16.6%	L-S	6.3%	0.62	17.8%

资料来源：Wind、聚源、兴业证券经济与金融研究院整理

图表 13、删除不同噪音水平下的 F-Momentum 在中证 800 内分位数测试对比

分组	年化收益率	Sharpe比率	最大回撤率	分组	年化收益率	Sharpe比率	最大回撤率	分组	年化收益率	Sharpe比率	最大回撤率
F-Momentum				F-Momentum-20Pec				F-Momentum-30Pec			
Top	8.0%	0.31	50.9%	Top	8.3%	0.33	46.6%	Top	9.3%	0.36	47.7%
1	7.1%	0.29	45.9%	1	7.1%	0.29	47.9%	1	6.8%	0.27	45.5%
2	7.0%	0.29	43.1%	2	6.4%	0.26	42.5%	2	5.5%	0.22	46.0%
3	5.5%	0.22	47.3%	3	4.6%	0.19	50.8%	3	4.7%	0.19	49.9%
Bottom	-0.1%	-0.00	56.9%	Bottom	1.1%	0.04	55.1%	Bottom	1.3%	0.05	54.5%
L-S	7.8%	1.09	9.0%	L-S	5.8%	1.00	11.3%	L-S	7.5%	1.07	11.0%
F-Momentum-50Pec				F-Momentum-80Pec				F-Momentum-90Pec			
Top	8.6%	0.34	46.8%	Top	10.0%	0.39	44.4%	Top	9.9%	0.38	46.9%
1	7.7%	0.31	43.2%	1	7.4%	0.30	45.2%	1	6.9%	0.28	47.0%
2	4.9%	0.20	50.5%	2	5.6%	0.20	43.2%	2	6.5%	0.27	40.9%
3	5.3%	0.21	48.9%	3	3.7%	0.15	56.2%	3	3.6%	0.14	53.3%
Bottom	1.1%	0.04	54.0%	Bottom	0.9%	0.04	54.1%	Bottom	0.8%	0.03	54.9%
L-S	7.1%	0.97	10.8%	L-S	8.6%	1.14	10.9%	L-S	8.8%	1.20	11.6%

资料来源：Wind、聚源、兴业证券经济与金融研究院整理

3.2、相似矩阵应用改进二---矩阵的另类利用

在图表 6 因子流程图中，我们经过 Step2 得到了财务关联矩阵。矩阵捕捉了任意两个公司的财务相似大小。以贵州茅台为例，我们是否可以直接将其他所有公司的相似度进行直接操作，进而作为贵州茅台的因子值呢？通过这种方式我们得到全新的选股因子，称之为 Link_New。

从测试结果来看：Link_New 因子 IC 达到 0.042，T 值高达 8.18；因子分位数测试各组别严格单调，多空收益率达 14.7%，夏普比达 1.99，最大回撤 9.1%。另外值得注意的是近几年来因子有效性非常强。

图表 14、Link_New 因子 IC 测试结果

平均值	标准差	最小值	最大值	ICIR	T 统计量
0.042	0.06	-0.21	0.21	0.68	8.18

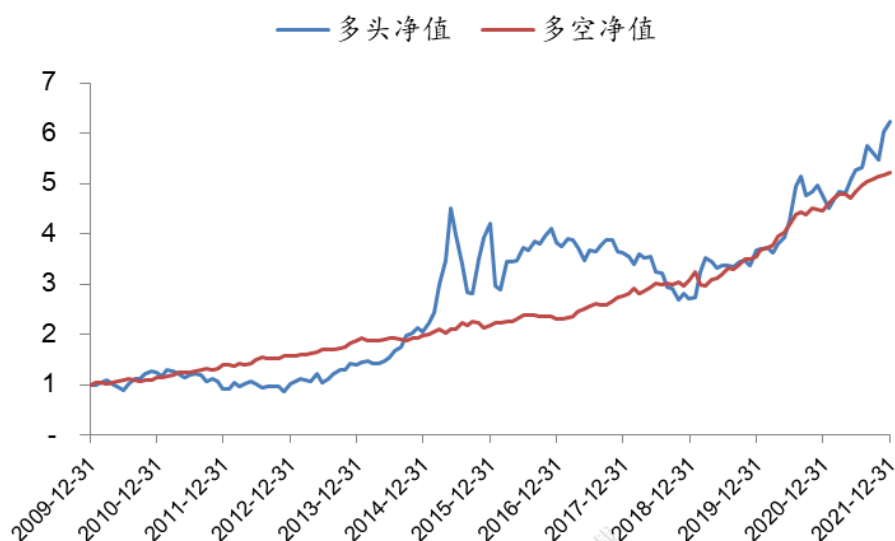
资料来源：Wind、聚源、兴业证券经济与金融研究院整理

图表 15、Link_New 因子分位数测试结果

	年化收益率	波动率	Sharpe比率	平均换手率	最大回撤率
Top	16.4%	27.8%	0.59	52.3%	40.4%
1	12.6%	27.4%	0.46	80.3%	47.4%
2	11.6%	27.6%	0.42	91.3%	48.5%
3	10.2%	28.0%	0.36	97.2%	50.0%
4	9.7%	27.6%	0.35	100.3%	52.6%
5	9.6%	27.4%	0.35	101.3%	50.0%
6	6.4%	28.4%	0.22	100.9%	59.4%
7	6.8%	28.6%	0.24	96.4%	59.1%
8	4.7%	29.4%	0.16	85.6%	61.3%
Bottom	0.4%	30.5%	0.01	56.6%	63.3%
L-S	14.7%	7.4%	1.99		9.1%

资料来源：Wind、聚源、兴业证券经济与金融研究院整理

图表 16、Link_New 因子分位数测试净值表现

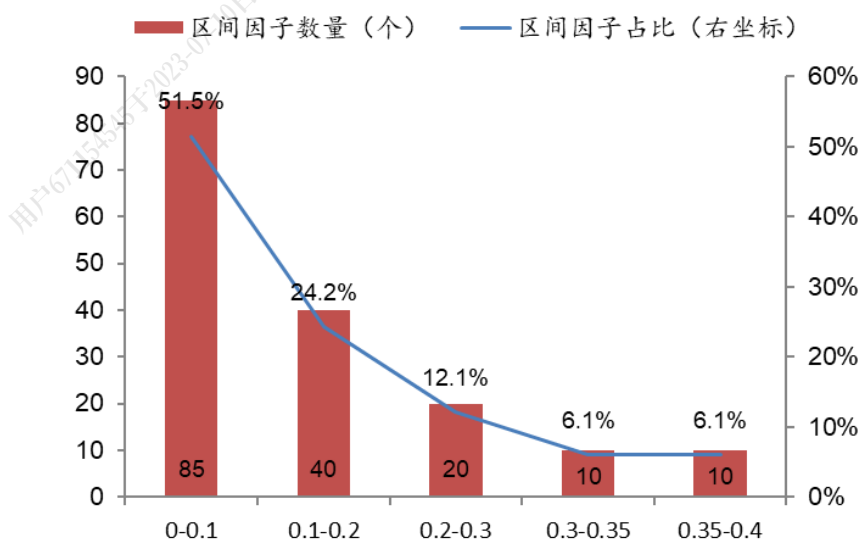


来源：聚源、兴业证券经济与金融研究院整理

资料来源：

针对于新构建的 Link_New 因子，我们进行了特异性分析：观察该因子和兴证金工底层 165 个选股因子的相关性，从因子相关性频率分布直方图来看，相关性高于 0.3 的占比为 12.2%，且相关性最高仅为 0.39。

图表 17、Link_New 与各选股因子相关性分布图



资料来源：Wind、聚源、兴业证券经济与金融研究院整理

4、降维方法改进

在财务动量因子构建的过程之中，我们通过月度收益降维最终构建了特异性极强的选股因子。但收益率降维只是其中一个方式，我们这里进行了全新的探索。

请务必阅读正文之后的信息披露和重要声明

这里保持步骤一和步骤二不变，而是将收益率改变成其他度量指标，比如基金重仓。 $FundHold_{jt}$ 表示基金持仓的公司 j 在 t 时刻的比例（当然这里需要注意的是仓位测算该如何进行？具体参见我们往期的仓位测算报告）。那么 $Fund-Momen_{it}$ 表示针对于公司 i ，利用财务关联度对基金持仓比例进行加权的结果。从这个因子构建逻辑来讲，因子值越大实际上表示和公司 i 财务关联度高的集群公司在考察月份（或者其他时间频次，我们后续如无特殊强调，均指月度频次）基金持仓越相似（配置幅度相同），我们认为这种情况下考察公司 i 接下来会有上涨的可能，简称为持仓动量因子（当然这里假设基金经理具有较强的选股能力）。

行业市值中性化的 Fund-Moment 因子 IC 达到 0.032、ICIR 达到 0.54、T 值高达 6.3，分位数各组别严格单调，有效性非常强。

图表 18、Fund-Moment 因子 IC 测试结果

					ICIR	T 值
Fund-Moment	0.031	0.12	-0.37	0.27	0.26	3.18
行业市值中性化后因子	0.032	0.06	-0.11	0.15	0.54	6.53

资料来源：Wind、聚源、兴业证券经济与金融研究院整理

图表 19、Fund_Moment 因子分位数测试结果

	年化收益率	波动率	Sharpe比率	平均换手率	最大回撤率
Top	10.7%	29.2%	0.37	57.2%	50.7%
1	12.8%	29.3%	0.44	99.3%	51.6%
2	12.0%	28.3%	0.42	114.4%	52.6%
3	11.2%	28.0%	0.40	122.0%	54.4%
4	9.9%	27.4%	0.36	125.6%	49.8%
5	9.1%	27.8%	0.33	125.0%	51.3%
6	9.2%	27.9%	0.33	120.9%	51.2%
7	5.8%	28.1%	0.21	111.6%	57.5%
8	5.1%	28.9%	0.18	94.7%	59.6%
Bottom	1.3%	30.7%	0.04	59.9%	59.6%
L-S	7.6%	14.1%	0.54		19.1%

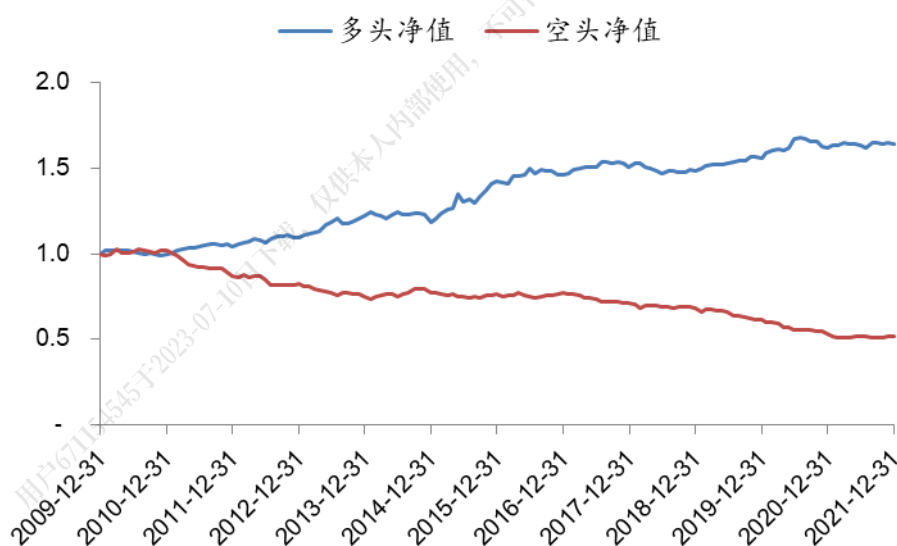
资料来源：Wind、聚源、兴业证券经济与金融研究院整理

图表 20、行业市值中性化后 Fund_Moment 因子分位数测试结果

	年化收益率	波动率	Sharpe比率	平均换手率	最大回撤率
Top	12.8%	30.2%	0.43	72.0%	49.9%
1	12.9%	28.3%	0.46	116.1%	44.7%
2	11.8%	27.9%	0.42	131.9%	48.8%
3	11.4%	27.5%	0.41	138.2%	52.1%
4	10.2%	27.2%	0.38	141.3%	48.5%
5	9.3%	26.9%	0.35	140.9%	49.7%
6	8.4%	27.3%	0.31	137.1%	52.2%
7	6.4%	27.6%	0.23	129.1%	56.4%
8	4.5%	28.6%	0.16	112.8%	61.3%
Bottom	2.5%	30.0%	0.08	70.2%	60.0%
L-S	9.7%	7.2%	1.35		10.1%

资料来源：Wind、聚源、兴业证券经济与金融研究院整理

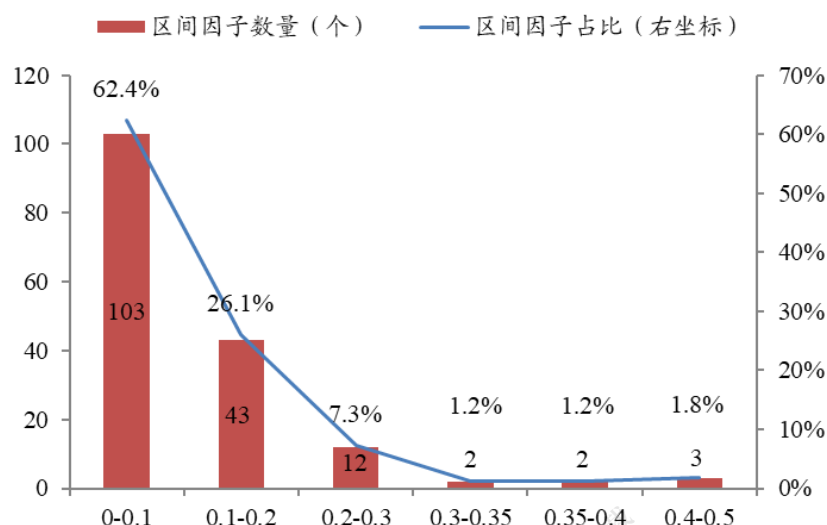
图表 21、行业市值中性化后 Fund_Moment 因子净值



资料来源：Wind、聚源、兴业证券经济与金融研究院整理

针对于新构建的 Fund-Moment 因子，我们分析其特异性分析：观察该因子和兴证金工底层 165 个选股因子的相关性。从因子相关性频率分布直方图来看，相关性高于 0.3 的共计 7 个因子，占比为 4.2%（相关性最高为 0.48）。即便剥离相关性最高的 5 个因子，其有效性依然很强，具体参见图表 23。

图 22、行业市值中性化后 Fund_Moment 因子和底层所有因子相关性分析



来源：聚源、兴业证券经济与金融研究院整理

资料来源：

图 23、剥离相关性最高的因子后 Fund-Moment 因子 IC 测试结果

平均值	标准差	最小值	最大值	ICIR	T 统计量
0.019	0.04	-0.10	0.13	0.42	4.23

资料来源：Wind、聚源、兴业证券经济与金融研究院整理

5、相似底层向量改进

接下来我们进入对图表 6 中 step1 的尝试，即改变输入向量的探索。在这里，我们首先对利用专利构建的因子样本外表现进行了回顾，同时尝试了输入其他因子进行建模。

5.1、相似向量尝试改进一---专利动量因子

➤ 1、专利动量因子数据&逻辑介绍

使用的专利数据由深圳德高行知识产权数据技术有限公司提供。

德高行是一家专业的知识产权解决方案供应商，其原始数据来自于国家专利局。专利数据的分析主要从**专利分类**（发明专利、实用新型专利、外观设计专利）、**专利所处状态**（专利申请、专利公开、实质审查、专利授权几个流程）以及**具体指标**（专利数量、权利要求项等）三个维度展开（具体详见团队于 2018 年 12 月 20 日发布的《从中国心到中国芯_由贸易战引发的专利选股有效研究》）。

除以上维度外，我们还能够从专利分类入手，通过定义专利关联度这一核心概念，进而聚焦于专利动量因子的研究。而专利分类依赖于《国际专利分类表》（IPC 分类）。该表是根据 1971 年签订的《国际专利分类斯特拉斯堡协定》编制的，是目前国际通用的专利文献分类和检索工具。国际专利分类系统按照技术主题设立类目，把整个技术领域分为 5 个不同等级：部、大类、小类、大组、小组，

分别对应着 8、145、670、3,000+、10,000+ 个类别。某种程度上 IPC 分类可以理解成为一种全新的专利分类（其中一级维度具体明细参见图表 24）。从数据来源角度讲，一个专利的 IPC 分类的原始数据是一个 JPG 格式的文件，从中进行解析即可得到 IPC 分类相应的数据，具体参见图表-25。

图表 24、IPC 一级分类示意图

类别代码	类别释义
A	电学
B	物理
C	机械工程、照明、加热、武器、爆破
D	固定建筑物
E	纺织、造纸
F	化学、冶金
G	作业、运输
H	人类生活需要（农、轻、医）

资料来源：德高行、兴业证券经济与金融研究院整理

图表 25、专利 IPC 分类原始数据示意

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.¹
 C02F 9/04
 C02F 9/02 A47J 31/00
 //(C02F9/04, 1: 2
 8, 1: 68)

[12] 实用新型专利说明书

[21] ZL 专利号 01246490.2

[45] 授权公告日 2002 年 5 月 22 日

[11] 授权公告号 CN 2492553Y

[22] 申请日 2001.7.11 [73] 专利权人 蔡忠春 地址 315411 浙江省余姚市丈亭镇宁波中联电器有限公司 [72] 设计人 蔡忠春	[21] 申请号 01246490.2 [74] 专利代理机构 浙江翔隆专利事务所 代理人 白拱长
---	--

权利要求书 1 页 说明书 4 页 附图页数 3 页

[54] 实用新型名称 饮水机净水矿化装置

资料来源：德高行、兴业证券经济与金融研究院整理

我们将所有上市公司过去 5 年的有效授权发明专利映射到 IPC 二级分类。计算任意两个公司的专利分布相似度，该指标本质上代表着两家公司的专利在类别上的相似程度：指标值越大，代表两个公司在申请的专利类别上布局越相似、创新研发方向也越趋于一致。这种相似性的衡量方式，已经跨域了传统财报以及行业分类所能提供的信息，在一定程度上反映了公司未来的发展战略方向。

正因为专利关联度高的公司在业务方向和创新研发上具有更高的相似性，所

以我们也有理由推测，其二级市场股价的联动性也将更为显著。有鉴于此，我们利用专利关联度和股票涨跌幅开发了专利动量因子，定义如下式：

$$tech_ret_{i,t} = \frac{\sum_{j \neq i} tech_{ij,t} \cdot ret_{j,t}}{\sum_{j \neq i} tech_{ij,t}}$$

$ret_{j,t}$ 代表公司 j 在 t 时刻月度收益率， $tech_{ij,t}$ 表示公司 i 和公司 j 在 t 时刻科技关联度， $tech_ret_{i,t}$ 代表公司 i 在 t 时刻的科技动量因子。简单来说，我们在每个月底，计算了除目标公司外的其他公司在专利关联度下的加权平均月收益率，进而形成专利动量因子。（针对于该因子的具体逻辑以及测试细节参见团队于 2019 年 6 月 25 日发布的《基于专利分类的专利动量因子研究》）。

➤ 2、专利动量因子样本外表现

由于专利动量因子构建于 2019 年 6 月份，其样本外已经 2 年 6 个月有余。由于客观原因，我们期初并没有对该因子进行实时更新。近期将因子更新至最新发现，其样本外表现着实优异。

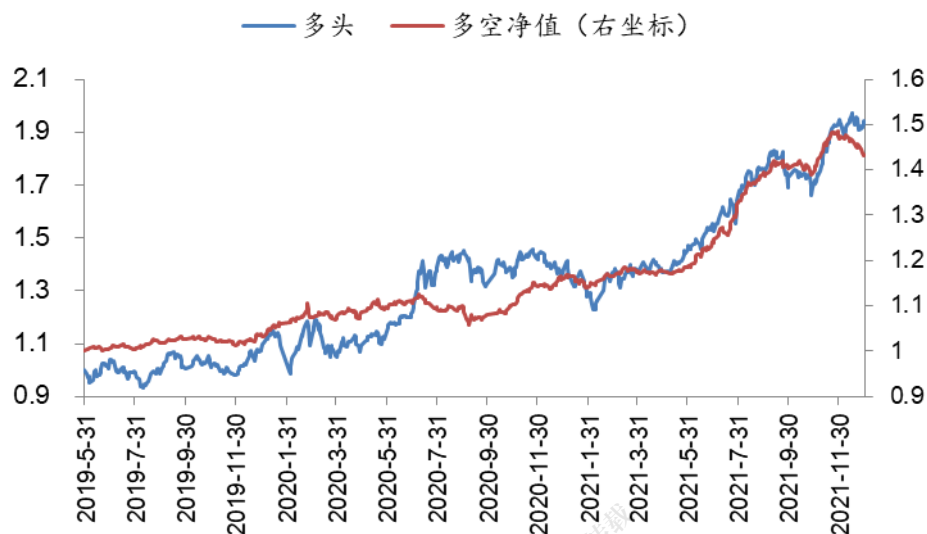
从因子样本外来看，其 IC 达到 0.023，T 值达到 2.8。分位数各组别相对严格单调，多头收益高达 29.3%，多空收益达到 14.9%，夏普比高达 2.7。无论是多头亦或是多空最大回撤均非常小（多头最大回撤 15.8%、多空最大回撤 5.8%）。

图表 26、IE-Momentum 因子分位数测试

	年化收益率	波动率	Sharpe比率	平均换手率	最大回撤率
Top	29.3%	21.3%	1.37	7.8%	15.6%
1	23.3%	20.9%	1.11	8.7%	15.5%
2	21.5%	20.8%	1.03	8.9%	15.9%
3	24.4%	20.6%	1.18	8.8%	16.5%
4	19.5%	20.5%	0.95	8.7%	16.6%
5	20.3%	20.2%	1.00	8.7%	15.6%
6	19.2%	20.6%	0.93	8.8%	18.3%
7	17.8%	20.7%	0.86	8.9%	20.2%
8	16.7%	20.6%	0.81	8.8%	20.2%
Bottom	13.1%	20.4%	0.64	7.8%	19.9%
L-S	14.9%	5.5%	2.71		5.8%

资料来源：德高行、Wind、聚源、兴业证券经济与金融研究院整理

图表 27、IE-Momentum 因子多头&多空净值曲线



资料来源：德高行、Wind、聚源、兴业证券经济与金融研究院整理

图表 28、IE -Momentum 因子 IC 测试

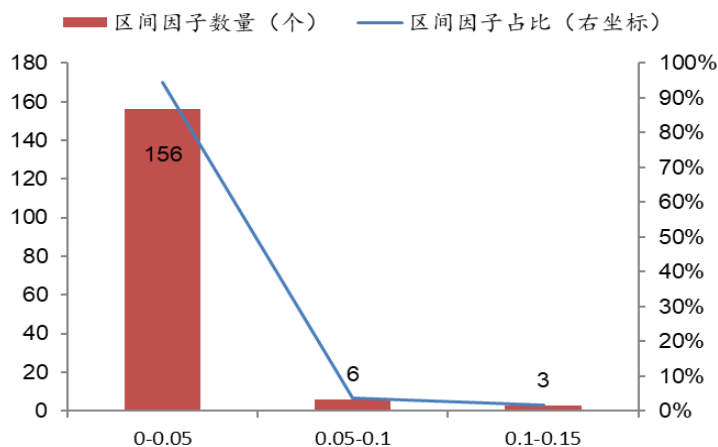
	平均值	标准差	最小值	最大值	ICIR	T 统计量
IE-Momentum	0.023	0.05	-0.06	0.13	0.50	2.79

资料来源：德高行、Wind、聚源、兴业证券经济与金融研究院整理

➤ 3、专利动量因子特异性分析

我们简单论证了该因子的特异性：通过测试该因子与兴证金工底层所有因子的相关性（兴证金融工程团队选股因子库共计 165 个因子，细分为价值、成长、质量、分析师情绪、动量反转、另类这六大类指标），我们发现相关性最高仅仅为 0.13，这充分论证了该因子的特异性。

图表 29、IE-Momentum 因子与各选股因子相关性分布图



资料来源：Wind、聚源、德高行、兴业证券经济与金融研究院整理

5.2、相似向量尝试改进二---持仓动量因子

除了前面的各种改进之外，我们甚至可以调整初始的向量构成，在初始的财物动量因子中，我们从企业盈利、偿债、营运、发展四个维度出发选择一些代表性指标，以此作为基准构建了财物动量因子并以此作为基本盘对其改进进行了全方位的探讨，效果颇佳。那么在本章中，我们将打破这个最底层的约束，而是选择其他因子作为输入，并将删除尾部不相似的 90% 样本，进一步构建相似度因子，算法如下：

$$Sig-link_{ijt} = \frac{\sum_{i=1}^n C_{itk} \cdot C_{jtk}}{\sqrt{\sum_{i=1}^n C_{itk}^2} \cdot \sqrt{\sum_{i=1}^n C_{jtk}^2}} \quad (4)$$

其中 k 代表新因子、i 代表公司、t 代表相应时点；

C_{itk} 代表公司 i 在 t 时刻时因子 k 的值大小；

$Sig-Link_{ijt}$ 代表公司 i 和 j 在 t 时刻的余弦相似度，简称为关联度指标。

关联度是任意两个公司维度相似性的衡量方式，在一个时点实际上有 3000*3000 对数值（假设市场截面共计 3000 个在上市公司）。在关联度的基础上，我们构建相似动量因子，具体参见公式（5）：

$$Sig-Momen_{it} = \frac{\sum_{i \neq j} F-link_{ijt} \cdot Ret_{jt}}{\sum_{i \neq j} F-link_{ijt}} \quad (5)$$

S.t:

$$F-link_{ijt} > 90\% \text{ 相似度}$$

我们测试该因子的有效性和特异性。从有效性结果来看：因子 IC 达到 0.035、ICIR 高达 0.57、T 值达到 6.78，因子分位数测试各组别严格单调，多头收益达到 15.0%，多空收益达到 13.6%。

图表 30、删除不同噪音水平下的 Sig-Momentum IC 测试

	平均值	标准差	最小值	最大值	ICIR	T 统计量
全市场测试结果	0.035	0.06	-0.12	0.21	0.57	6.78
中证 800 内测试结果	0.036	0.07	-0.16	0.23	0.51	6.10

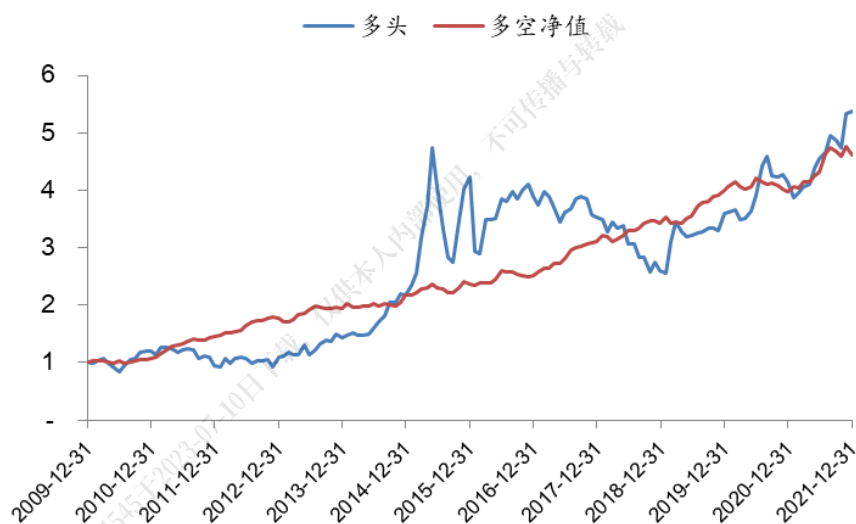
资料来源：Wind、聚源、兴业证券经济与金融研究院整理

图表 31、行业市值中性化后 Sig-Momentum 因子分位数测试结果

分组	年化收益率	Sharpe比率	最大回撤率	分组	年化收益率	Sharpe比率	最大回撤率
全市场				中证800内			
Top	15.0%	0.52	46.2%	Top	10.7%	0.41	42.0%
1	13.2%	0.47	48.6%	1	9.5%	0.37	45.0%
2	13.2%	0.46	46.7%	2	8.8%	0.36	47.6%
3	10.9%	0.39	48.6%	3	8.1%	0.33	41.3%
4	10.8%	0.39	48.4%	4	4.7%	0.19	47.9%
5	8.7%	0.32	50.3%	5	6.2%	0.25	48.7%
6	7.4%	0.26	55.3%	6	4.7%	0.19	51.2%
7	6.6%	0.23	55.5%	7	2.4%	0.10	53.3%
8	3.8%	0.14	60.1%	8	0.4%	0.02	54.4%
Bottom	1.0%	0.03	63.4%	Bottom	-0.5%	-0.02	57.1%
L-S	13.6%	1.79	6.6%	L-S	10.9%	1.15	17.8%

资料来源：Wind、聚源、兴业证券经济与金融研究院整理

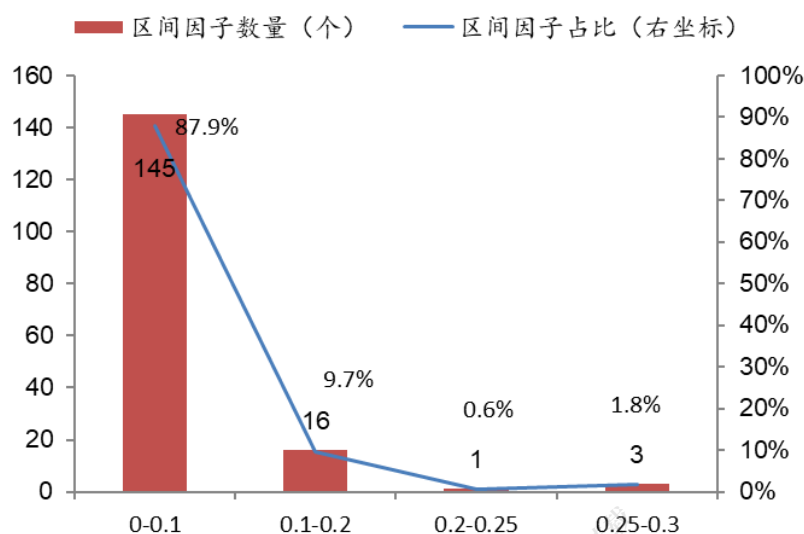
图表 32、行业市值中性化后 Sig-Momentum 因子分位数净值



资料来源：Wind、聚源、兴业证券经济与金融研究院整理

从特异性结果来看：该因子和兴证金工底层所有因子相关性最高的仅仅 0.29（净利润同比增速因子）。

图表 33、行业市值中性化后 Sig_Moment 因子分位数净值



来源、兴业证券经济与金融研究院整理

资料来源：

6、相似方法论本质探讨和总结

6.1、百尺竿头更进一步---因子归结篇

➤ 相似动量改进维度归纳

至此，我们完成了相似动量目前所有的改进。百尺竿头更进一步，我们对上述各个改进的步骤和结果进行简单的归纳总结，为进一步构建复合因子打下基础。

图表 34、各因子逻辑简介

改进维度	简介
矩阵操作--删除尾部噪音	由于市场一些股票和考察标的确实不相似，所以通过删除能提纯信息，进而提升表现（具体参见章节 3.1）→ F-Momentum-80Pec
矩阵操作—信	由于相似度矩阵不太适合选股，所以期初我们通过收益率进行降维，（章节 3.2） → Link_New
降维方式--基金持仓	基于相似度矩阵，我们利用月度收益率进行降维构建财务动量因子。将月度收益率替换成基金持仓，全新的因子，效果颇佳（具体参见第四章）→ Fund-Moment
改变输入向量—专利	前面的所有的分析都是基于财务进行分析。我们多次强调，相似是一种信息提取的方式。这里我们将财务动量换成专利的分布，进而构建全新的专利动量，效果颇佳（具体参见 5.1）→ IE-Momentum
改变输入向量—其他	除了利用专利之外，我们亦可以利用其它维度信息进行研究以形成向量基础，并进一步构建相似度动量因子，同时在构建该因子时，我们借鉴了删除尾部噪音（具体参见 5.2）→ Sig-Momentum。 这里需要注意的是，由于 Sig-Momentum 构建的基础向量是在

请务必阅读正文之后的信息披露和重要声明

F-Momentum-80Pec 之上，同时也利用了删除尾部噪音。所以后续复合因子时，我们不在考虑 F-Momentum-80Pec

资料来源：德高行、Wind、聚源、兴业证券经济与金融研究院整理

图表 35、各因子对比表现-IC

	平均值	标准差	最小值	最大值	ICIR	T 统计量
Link_New	0.042	0.06	-0.21	0.21	0.68	8.2
Fund_Momentum	0.032	0.06	-0.11	0.15	0.54	6.5
IE_Momentum	0.023	0.05	-0.06	0.13	0.5	2.8
Sig_Momentum	0.036	0.07	-0.16	0.23	0.51	6.1

资料来源：德高行、Wind、聚源、兴业证券经济与金融研究院整理

➤ 复合相似动量表现分析

总结完毕后，我们简单测试了各个因子的相关性，从结果来看，各因子相关性非常低。

图表 36、各因子相关性测试

	Sig-Momentum	Fund-Moment	Link_New	IE-Momentum
Sig-Momentum		0.11	0.09	0.05
Fund-Moment			0.32	0.01
Link_New				-0.00
IE-Momentum				

资料来源：德高行、Wind、聚源、兴业证券经济与金融研究院整理

进一步我们将 Link_New、Fund-Moment、IE-Momentum、Sig-Momentum 等权合成，构建复合因子，简称为 Link-Comb-Sig。从表现来看，复合因子在全市场 IC 高达 0.053、ICIR 达到 0.79、T 值高达 9.5。因子分位数测试各组别严格单调，以在全市场测试为例：多头收益高达 21.4%、多空收益高达 21.9%、多空夏普比达到 2.64、最大回撤 4.5%。因子整体有效性非常强。

进一步我们测试了复合因子在宽基范围内表现并没有出现衰减（IC 维度），

其在中证 800 内表现与全市场几乎持平。

图表 37、Link-Comb-Sig IC 测试

	平均值	标准差	最小值	最大值	ICIR	T 统计量
全市场测试结果	0.053	0.07	-0.11	0.22	0.79	9.50
中证 800 内测试结果	0.049	0.07	-0.11	0.25	0.67	8.08

资料来源：德高行、Wind、聚源、兴业证券经济与金融研究院整理

图表 38、Link-Comb-Sig 因子分位数测试结果—全市场

	年化收益率	波动率	Sharpe比率	平均换手率	最大回撤率
Top	21.4%	29.3%	0.73	147.5%	37.8%
1	15.1%	28.3%	0.53	166.2%	43.7%
2	11.9%	27.6%	0.43	171.9%	48.1%
3	11.3%	27.7%	0.41	174.4%	48.4%
4	10.2%	28.3%	0.36	175.3%	50.5%
5	7.7%	27.8%	0.28	175.2%	58.2%
6	4.6%	27.6%	0.17	174.9%	60.2%
7	5.3%	28.5%	0.19	171.4%	60.5%
8	2.4%	28.4%	0.08	164.4%	61.3%
Bottom	-0.9%	29.4%	-0.03	134.4%	64.1%
市场	8.8%	28.1%	0.31		53.5%
L-S	21.9%	8.3%	2.64		4.5%

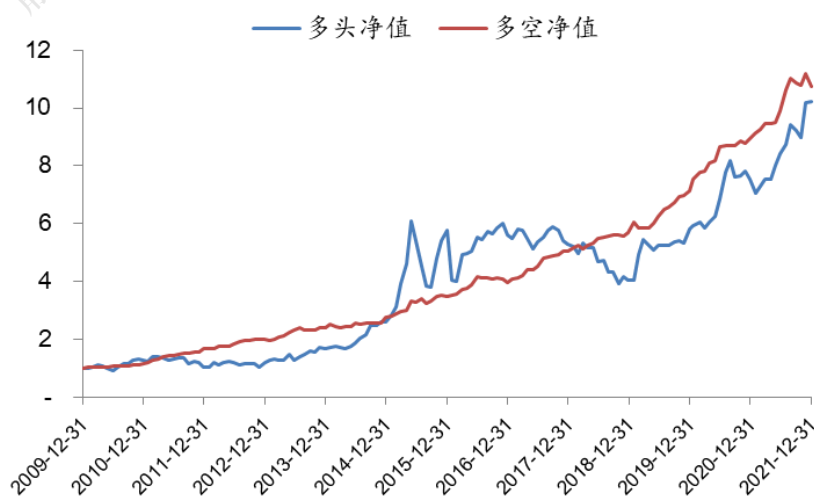
资料来源：德高行、Wind、聚源、兴业证券经济与金融研究院整理

图表 39、Link-Comb-Sig 因子分位数测试结果—中证 800 内

	年化收益率	波动率	Sharpe比率	平均换手率	最大回撤率
Top	13.9%	25.6%	0.54	146.3%	39.9%
1	10.6%	24.7%	0.43	166.6%	39.0%
2	7.5%	24.7%	0.31	171.5%	41.5%
3	6.9%	24.9%	0.28	173.5%	50.1%
4	7.3%	24.6%	0.30	173.3%	42.9%
5	5.4%	24.8%	0.22	175.0%	55.2%
6	2.9%	25.0%	0.12	173.8%	51.7%
7	1.9%	25.5%	0.07	171.1%	58.8%
8	0.5%	25.6%	0.02	163.7%	55.5%
Bottom	-4.1%	26.9%	-0.15	133.0%	66.0%
市场	5.3%	24.9%	0.21		50.0%
L-S	17.8%	9.4%	1.89		10.7%

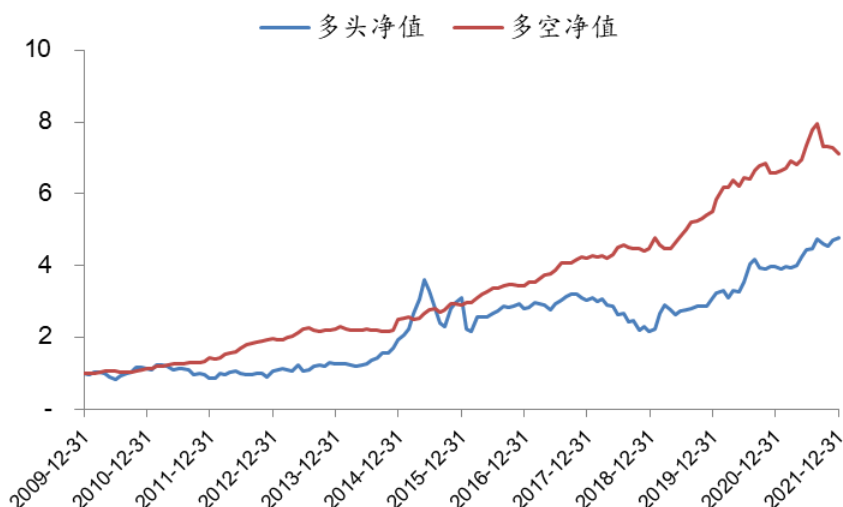
资料来源：德高行、Wind、聚源、兴业证券经济与金融研究院整理

图表 40、Link-Comb-Sig 因子分位数净值—全市场



资料来源：德高行、Wind、聚源、兴业证券经济与金融研究院整理

图表 41、Link-Comb-Sig 因子分位数净值—中证 800 内



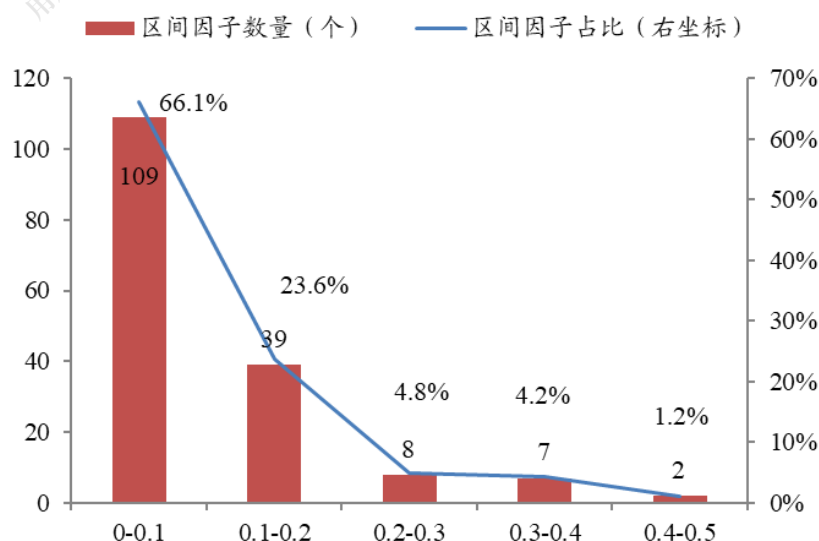
资料来源：德高行、Wind、聚源、兴业证券经济与金融研究院整理

➤ 复合相似动量特异性分析

进一步观察因子特异性，我们通过三种方式来观察因子的特异性：直接测试该因子和底层所有因子相关性（这种方式之前相同）；观察因子和 Barra 风格因子的相关性；将该因子放入到我们的复合模型中去看增量效应。

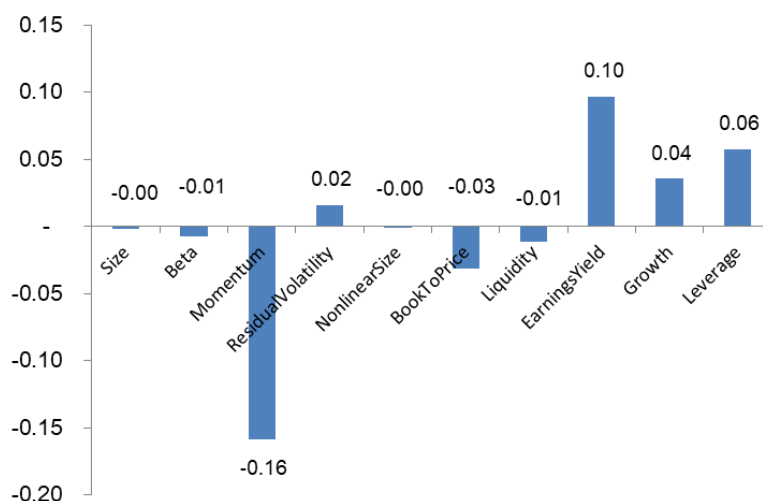
我们仍然通过极为严格的和底层所有因子相关性测试为基准进行分析，结果发现：底层所有因子和复合因子相关性极低，相关性高于 0.4 的仅仅只有 2 个。如果观察该因子和 Barra 的十个风格因子，我们发现其相关性也非常低，具体参见图表 43。

图表 42、Link-Comb-Sig 因子相关性测试



资料来源：Wind、聚源、兴业证券经济与金融研究院整理

图表 43、Link-Comb-Sig 与 Barra 风格因子相关性测试



资料来源：Wind、聚源、兴业证券经济与金融研究院整理

6.2、相似的本质探索

截至目前，我们通过《基于专利分类的科技动量因子研究》、《猎金系列之三十二：财报季的财务效应研究和因子构建》和本篇报告全方位探索了基于相似思想探索新因子的研究成果。尤其是在本篇报告中，我们给出了 4 个全新的选股因子，且这些因子的相关性非常低。那么追根溯源，这种方法到底在捕捉什么呢？在这里我们试图进行一定程度的解释。

从归宿上来讲，投资的方法论目前包括归纳法和演绎法。而无论归纳和演绎均是逻辑和统计规律的体现：要么是满足逻辑进行外推，要么是满足大数定律进而外推。而无论是归纳法亦或是演绎法均可归结为寻找“相似”：

- 1) 从权益的角度来讲：以新能源为例，当新能源的风口来临时，市场开始投资宁德时代等引领者，随着这些标的的估值拔高、投资标的的性价比变低，很多投资者开始关注产业链上的其他投资标的。这里的产业链实际上给出了一种投资逻辑的外延，通俗一点讲就是“一人得道鸡犬升天”，这是一种逻辑的外延，也是相似度量方式（过去怎么样，未来就会怎么样或者不会怎么样）。但关键是多数情况下大家关注的是**显性传导标的**，而隐性的和新能源相关的标的可能会被大家忽略；
- 2) 从量化投资的角度来讲，我们基于样本内得到某种规律，然后默认其可以外推，一般称之为风格动量或者因子动量等。然后样本内的统计规律，我们认为可以外延，这也是一种相似的应用体现。

其实上面两种形式均是在捕捉契合当下/过去投资逻辑的潜在机会，只不过殊途同归。相思动量正是这两种思想的应用和尝试，尤其是**隐性的、弱关联的、经**

常被大家忽略 Alpha，亦能被这种方法捕捉到，因此其特异性极强、稳定性极强（体现在目前构建的所有相似动量因子的 ICIR 均非常高）。我们认为这是该因子起效的本质和原因。

6.3、总结

本篇幅是前面两篇相似的延伸，但并不意味着我们在此方向止步不前。相反，我们将继续前行，一方面从深度上对相似体系的研究进行切入（比如尝试更多的场景、更多的改进），另一方面从广度上对相似的应用进行探索（比如尝试在选基、资产配置上）。我们真诚希望能通过扎实的研究为各位投资者带来启发和价值。让我们期待与下一篇相似研究的相遇。

风险提示：本报告模型及结论全部基于对历史数据的分析，当市场环境变化时，存在模型失效风险。

分析师声明

本人具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立、客观地出具本报告。本报告清晰准确地反映了本人的研究观点。本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

投资评级说明

投资建议的评级标准	类别	评级	说明
报告中投资建议所涉及的评级分为股票评级和行业评级(另有说明的除外)。评级标准为报告发布日后的12个月内公司股价(或行业指数)相对同期相关证券市场代表性指数的涨跌幅。其中：A股市场以上证综指或深圳成指为基准，香港市场以恒生指数为基准；美国市场以标普500或纳斯达克综合指数为基准。	股票评级	买入	相对同期相关证券市场代表性指数涨幅大于15%
		审慎增持	相对同期相关证券市场代表性指数涨幅在5%~15%之间
		中性	相对同期相关证券市场代表性指数涨幅在-5%~5%之间
		减持	相对同期相关证券市场代表性指数涨幅小于-5%
		无评级	由于我们无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使我们无法给出明确的投资评级
	行业评级	推荐	相对表现优于同期相关证券市场代表性指数
		中性	相对表现与同期相关证券市场代表性指数持平
		回避	相对表现弱于同期相关证券市场代表性指数

信息披露

本公司在知晓的范围内履行信息披露义务。客户可登录 www.xyzq.com.cn 内幕交易防控栏内查询静默期安排和关联公司持股情况。

使用本研究报告的风险提示及法律声明

兴业证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本报告仅供兴业证券股份有限公司(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。本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约，投资者自主作出投资决策并自行承担投资风险，任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效，任何有关本报告的摘要或节选都不代表本报告正式完整的观点，一切须以本公司向客户发布的本报告完整版本为准。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户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，必要时就法律、商业、财务、税收等方面咨询专家的意见。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。

本报告所载资料的来源被认为是可靠的，但本公司不保证其准确性或完整性，也不保证所包含的信息和建议不会发生任何变更。本公司并不对使用本报告所包含的材料产生的任何直接或间接损失或与此相关的其他任何损失承担任何责任。

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、价值及投资收入可升可跌，过往表现不应作为日后的表现依据；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；本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

除非另行说明，本报告中所引用的关于业绩的数据代表过往表现。过往的业绩表现亦不应作为日后回报的预示。我们不承诺也不保证，任何所预示的回报会得以实现。分析中所做的回报预测可能是基于相应的假设。任何假设的变化可能会显著地影响所预测的回报。

本公司的销售人员、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。

本报告并非针对或意图发送予或为任何就发送、发布、可得到或使用此报告而使兴业证券股份有限公司及其关联子公司

限公司受制于相关法律或法规的任何地区、国家或其他管辖区域的公民或

居民，包括但不限于美国及美国公民(1934年美国《证券交易所》第15a-6条例定义为本「主要美国机构投资者」除外)。

本报告的版权归本公司所有。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本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。未经授权的转载，本公司不承担任何转载责任。

特别声明

在法律许可的情况下，兴业证券股份有限公司可能会持有本报告中提及公司所发行的证券头寸并进行交易，也可能为这些公司提供或争取提供投资银行业务服务。因此，投资者应当考虑到兴业证券股份有限公司及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。

兴业证券研究

上海	北京	深圳
地址：上海浦东新区长柳路36号兴业证券大厦15层	地址：北京市朝阳区建国门大街甲6号SK大厦32层01-08单元	地址：深圳市福田区皇岗路5001号深业上城T2座52楼
邮编：200135	邮编：100020	邮编：518035
邮箱： research@xyzq.com.cn	邮箱： research@xyzq.com.cn	邮箱： research@xyzq.com.cn