

AI+HI 系列（7）

## DecompGRU：基于趋势分解的时序+截面端到端模型

### ❖ 前言

近年基于 Transformer 的深度学习模型已有广泛运用，但我们仍能看到一些研究将注意力机制替换为 MLP、CNN 等模块，更简单的模型同样能取得有竞争力的表现，模型能力的贡献可能并不完全来自于特定的 token mixer，而是相当程度源于框架性的设计。受此启发，本篇报告我们设计一个模块简单，但在框架流程上不同于基线的新模型。

### ❖ 模型流程

我们延续并拓展先前报告“股票时序内的局部-整体交互+截面股票间交互”的整体思路，构建新的时序+截面端到端模型 DecompGRU。我们通过趋势分解模块得到时序趋势与残差分量解耦表征，随后使用单独的分支进行特征提取，其中股票的截面交互同样基于去均值处理实现；

### ❖ 因子测试

我们使用 150D 数据集，分别以 IC、MSE 两种损失函数训练模型，GRU 模型作为对比基线；DecompGRU 与基线模型最低相关性为 0.72，在风格倾向上与基线模型存在一定差异，在小市值选股域中优势更为明显；在中证全指 20190101~20250228 区间内，DecompGRU 的 10 日 RankIC 为 0.13、RankICIR 为 1.1；20 日 RankIC 为 0.141，RankICIR 为 1.23，均优于基线模型；周度分组测试中（不计交易成本），TOP 组年化收益为 52.82%，较基线提高 7.2%；在 300 / 500 选股域中，则整体表现弱于基线。

基于 DecompGRU-MSE 构建指增组合，在双边换手 30%的约束下，回测区间内 300/500/1000 指增组合年化超额收益分别为 8.48%、11.65%、17.11%，对应跟踪误差分别为 4.8、5.88、6.69。

### ❖ 风险提示：

策略基于历史数据回测，不保证未来数据的有效性。深度学习模型存在过拟合风险。深度学习模型受随机数影响。

### 相关研究报告

《AI+HI 系列（6）：对端到端模型泛化性的思考与改进——基于样本加权与风格约束》

2024-12-12

《AI+HI 系列（5）：CrossGRU-2：基于 Patch 与多尺度时序改进端到端模型》

2024-07-05

《AI+HI 系列（4）：CrossGRU：基于交叉注意力的时序+截面端到端模型》

2024-04-19

---

## 投资主题

### 报告亮点

时序深度学习模型有较多工作，我们注意到学术研究上有不少尝试使用CNN、MLP、RNN等模块来替换注意力机制的工作，这些模型同样能取得有竞争力的结果，因此本篇报告中，我们继续完善时序+截面的端到端框架，使用最为简单的模块来实现框架背后的思路，为广大投资者提供新的研究思路，助力量化投资的发展。

### 投资逻辑

时序与截面数据在量化投资领域处处可见，通过深度学习网络实现时序或截面信息的信息挖掘具有相当的潜力。本篇报告提出了基于深度学习的量价因子挖掘模型，进一步探索深度学习在量化领域上的运用。

---

# 目 录

|                  |    |
|------------------|----|
| 一、动机.....        | 5  |
| 二、模型介绍 .....     | 6  |
| (一) 时序趋势分解 ..... | 6  |
| (二) 双分支流程 .....  | 6  |
| 三、测试结果 .....     | 8  |
| (一) 数据集介绍 .....  | 8  |
| (二) 模型参数 .....   | 9  |
| (三) 因子测试结果 ..... | 9  |
| 1、IC 测试结果 .....  | 10 |
| 2、分组测试结果 .....   | 11 |
| (四) 指增测试 .....   | 15 |
| 四、总结 .....       | 17 |
| 五、风险提示 .....     | 17 |
| 六、参考文献 .....     | 17 |

# 图表目录

图表 1 Transformer 抽象化通用框架..... 5

图表 2 DecompGRU 模型流程 ..... 6

图表 3 滚动训练流程 ..... 8

图表 4 基线 GRU 模型流程..... 9

图表 5 训练参数设定 ..... 9

图表 6 模型风格倾向 ..... 10

图表 7 模型相关性 ..... 10

图表 8 10 日 IC 统计结果 ..... 10

图表 9 20 日 IC 统计结果 ..... 11

图表 10 中证全指-20 分组年化收益对比 ..... 12

图表 11 中证全指-TOP 组超额走势对比..... 12

图表 12 中证全指-TOP 组逐年收益对比..... 12

图表 13 中证全指-TOP 组绩效统计..... 12

图表 14 沪深 300-TOP 组逐年收益对比..... 13

图表 15 沪深 300-TOP 组超额走势对比..... 13

图表 16 沪深 300-TOP 组绩效统计..... 13

图表 17 中证 500-TOP 组逐年收益对比..... 13

图表 18 中证 500-TOP 组超额走势对比..... 13

图表 19 中证 500-TOP 组绩效统计..... 14

图表 20 中证 1000-TOP 组逐年收益对比..... 14

图表 21 中证 1000-TOP 组超额走势对比..... 14

图表 22 中证 1000-TOP 组绩效统计..... 14

图表 23 沪深 300 指增超额走势 ..... 15

图表 24 中证 500 指增超额走势 ..... 16

图表 25 中证 1000 指增超额走势 ..... 16

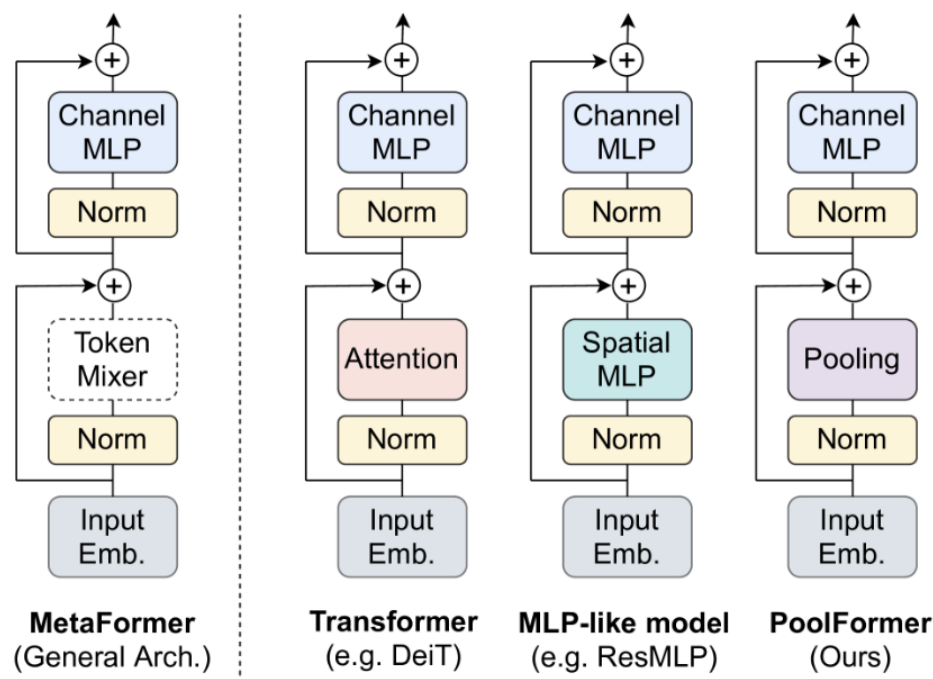
图表 26 指增组合统计 ..... 16

一、动机

在本篇报告，我们聚焦于端到端深度因子挖掘模型的设计优化。此前在《CrossGRU: 基于交叉注意力的时序+截面端到端模型》研究中，我们通过在基线 GRU 模型中引入交叉注意力机制，实现了股票间的截面信息交互；在后续《CrossGRU-2: 基于 Patch 与多尺度时序建模的端到端模型》报告中，采用双分支时序结构，通过交叉注意力机制融合不同时间尺度的“局部-整体”信息。

值得注意的是，在 CrossGRU-1 报告的消融实验部分表明，即便采用简单的均值截面模块，模型性能仍可超越基线。近年学术研究，例如 MetaFormer 等工作中，当用 MLP 或池化操作替代注意力机制时，模型仍可以取得有竞争力的性能，整体框架设计可能比具体模块实现更为重要：

图表 1 Transformer 抽象化通用框架



资料来源: Yu, Weihao, et al. "Metaformer is actually what you need for vision."

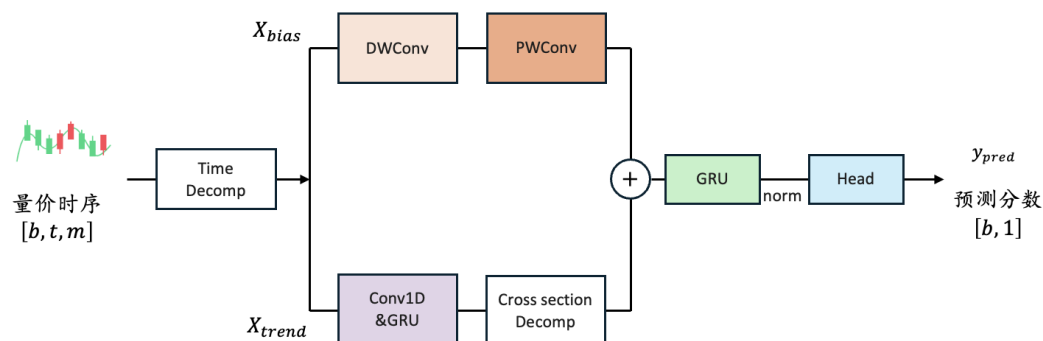
类似结论在时序领域同样：如时序 Patch、通道独立等与具体模块无关的设计对模型性能存在较大影响。因此，在本篇报告中，在继承先前报告的“股票时序内的局部-整体交互，以及截面股票间的交互”框架基础上，我们进行延续拓展，并将所有注意力机制替换为简单的去均值操作，构建一个更为简单的端到端模型——DecompGRU。

在时序模块部分，我们延续双分支设计，通过趋势分解生成差异化输入，维持多层次时序交互能力；截面模块部分，我们采用分组去均值策略，在保留每个时间步截面信息的同时大幅简化计算。

## 二、模型介绍

本篇报告的模型流程下图所示，模型每个 batch 输入为同截面上多只股票的量价时序，趋势分解模块将输入拆分为趋势与残差两项，分别进入一个分支进行建模，截面交互仅在趋势分支中被引入：

图表 2 DecompGRU 模型流程



资料来源：华创证券

DecompGRU 模型延续了先前报告的大致框架，使用更加简单的模块实现与拓展，下面我们具体介绍每个模块的实现细节。

### （一）时序趋势分解

时序分解是近年时序深度学习论文中常见的模块，例如 AutoFormer、Dlinear 等工作中使用最简单的移动均值方法来实现趋势分解，在进入骨干网络前，输入  $X$  拆分为长期趋势  $X_t$  与季节项  $X_s$ ：

$$X_t = \text{AvgPool}(\text{Padding}(X))$$

$$X_s = X - X_t$$

在后续流程，趋势分量和季节分量通常会使用单独的模块进行建模，最后进行合并两个分支得到最终预测。

我们使用上述相同的方法对模型输入进行处理，对股价场景而言，上述趋势分解模块的类似于技术指标 SMA 均线、均线乖离；padding 的作用为保证新序列长度与输入序列长度一致，我们将 padding 设置在时序的最远端，而非两端。

我们将分解模块输出的趋势分量和残差分量分别记为  $X_{trend}$  和  $X_{res}$ ，其中  $X_{trend}$  捕捉相对偏向整体的信息，而  $X_{res}$  则为更偏向局部的信息，两个分量被输入两个分支，进行不同处理。

### （二）双分支流程

#### （1）趋势分支

经过上一步，原始输入被拆分成了一项包含更多整体信息趋势分量、一项包含更多局部信息的残差分量。

首先，我们将趋势分量输入 1D 卷积 + GRU 来实现时序编码，每个股票在每个时间步被编码为一个  $d$  维的向量。

$$X_{trend} = GRU(Conv1D(X_{trend}))$$

其中，1D 卷积不使用 padding，卷积核大小等于步长，输出  $X_{trend} \in \mathbb{R}^{b,t',d}$

我们认为残差分量更多体现为个股的特质信息以及随机噪声，当我们需要考虑股票间的交互时，使用更偏向整体性的趋势分量可能更加合理，因此我们将简易截面模块只运用在趋势分支上。

我们直接在股票维度计算均值，将每个股票减去均值，得到趋势分量的最终表征  $X_{trendrepr} \in \mathbb{R}^{b,t',d}$ ：

$$X_{csrepr} = AvgPool(X_{trend}, dim = 0)$$

$$X_{trendrepr} = X_{trend} - X_{csrepr}$$

不同特征的股票可能存在截然不同的模式，直接在截面所有股票进行均值计算丢失信息较多。考虑到使用市值对截面股票进行划分是较为常见的选择，因此本篇报告中，我们引入分组进行去均值操作，根据截面股票的 SIZE 因子暴露大小进行 5 等分，将每个股票的表征减去同组股票的表征均值。

## (2) 残差分支

在残差分支我们使用更轻量的深度可分离卷积 (DSConv)，由深度卷积 (DWConv) 与点卷积 (PWConv) 构成，得到偏离分量的最终表征  $X_{resrepr} \in \mathbb{R}^{b,t',d}$ ：

$$X_{resrepr} = DSConv(X_{res})$$

其中，深度卷积无 padding，卷积核大小等于步长，数值设置同趋势分支中的 Conv1D 相同；点卷积将每个时间步特征维度映射至  $d$  维。

最后，我们将两个分支的结果相加合并，输入 GRU 编码器，取最后一个时间步的输出，通过输出维度为 1 的 Linear 得到每个股票的最终得分：

$$Y_{pred} = Linear\left(GRU(X_{trendrepr} + X_{resrepr})\right)$$

DecompGRU 包含 2 个不共享参数的 GRU 层，尽管增加了计算开销，但该设计对信息捕捉的是必要的：

- 在趋势分支，我们特意保留了 GRU 在每个时间步的输出，由于截面模块的简化，我们可以考虑每个时间步上的截面信息，而并非只使用最后一个时间步的表征进行截面交互；
- 残差分支专注局部特征增强，进一步扩充了每个时间步表征的信息。

换言之，趋势分支的 GRU 执行引入位置信息、基本时序编码任务。分支合并后，重构的股票表征输入第二个 GRU，以捕捉高阶时序演化规律。

三、测试结果

(一) 数据集介绍

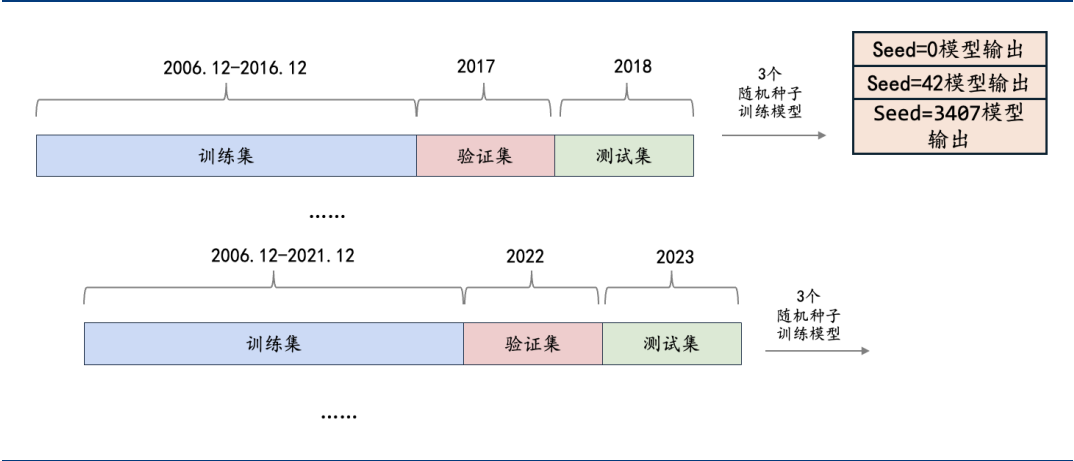
150D 数据集:

日频的高、开、低、收、均价、成交量 6 个特征;  
每个 batch 为同一截面多只股票的多变量时序, 序列长度为 150; batch 按日度采样;  
训练集的去极值、Zscore、填充缺失值预处理步骤与先前报告相同。

训练方式:

采用滚动训练, 每年年底进行模型重训练:

图表 3 滚动训练流程



资料来源: 华创证券

模型预测标签:

标签处理: 经过市值行业中性化的未来 10 日收益 (t+1 日~t+11 日以收盘价计算);  
标准化方法: 每个截面 rank / zscore 标准化, 后续分别对应 IC / MSE 损失函数。

损失函数:

我们分别测试 IC、MSE 作为损失函数时模型的表现。

其中 MSE 采用样本加权, 放大多头端权重:

$$s = \text{RELU}(y_{\text{true}} - \tau)$$

$$w = 1 + \alpha(\sigma(s) - 0.5)$$

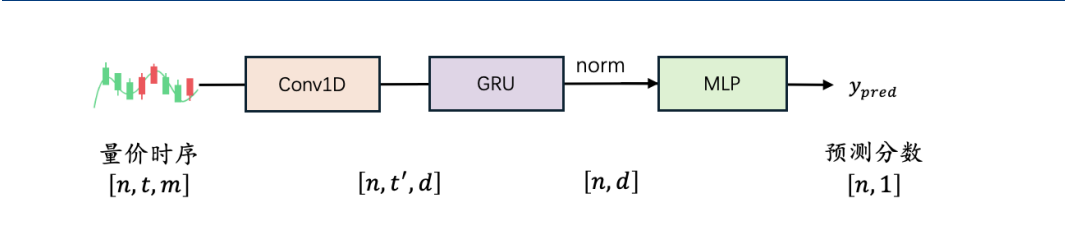
$$\mathcal{L} = w \cdot (y_{\text{pred}} - y_{\text{true}})^2$$

其中  $y_{\text{pred}}$  为模型预测值,  $y_{\text{true}}$  为经过标准化的 label, 阈值  $\tau$  设置为 1.0, 权重倍数  $\alpha$  设置为 2.0,  $\sigma(\cdot)$  表示 sigmoid 函数。

基线模型

基线模型为常用的 GRU 模型，流程如下：

图表 4 基线 GRU 模型流程



资料来源：华创证券

(二) 模型参数

基线 GRU 模型与 DecompGRU 共同的超参数设定如下：

- 实现 Patch 功能的 1D 卷积，卷积核大小等于步长，设置为 3；
- GRU 隐状态维度 64、层数为 1；
- 标准化层为 LayerNorm；

DecompGRU 模型涉及的额外超参数设定为：

- 时序趋势分解：计算 5 日均值；
- 趋势分支：根据 SIZE 因子暴露，将股票分为 5 组；
- 残差分支：深度卷积核大小等于步长，设置为 3；

所有训练流程涉及参数为：

图表 5 训练参数设定

| 变量名       | 参数设置       |
|-----------|------------|
| Patience  | 15         |
| Optimizer | Adam(1e-3) |
| Seed      | 0,42,3407  |

资料来源：华创证券

(三) 因子测试结果

回测区间：2019 年 1 月 1 日 ~ 2025 年 2 月 28 日；

不同模型因子与 barra 风格因子截面相关性均值如下图所示，DecompGRU 在与流动、残差波动风格上相关较高，相比基线 GRU，DecompGRU 模型更偏小市值、正 beta。

图表 6 模型风格倾向



资料来源: wind, 通联数据, 华创证券

下表展示不同模型因子的截面的秩相关性系数, 在时序上取均值的统计结果。使用不同损失函数时, DecompGRU 和 GRU 基线模型相关性最低, 为 72%。

图表 7 模型相关性

|               | GRU-IC | GRU-MSE | DecompGRU-IC | DecompGRU-MSE |
|---------------|--------|---------|--------------|---------------|
| GRU-IC        | 1      | 0.81    | 0.81         | <u>0.73</u>   |
| GRU-MSE       | 0.81   | 1       | <u>0.72</u>  | 0.77          |
| DecompGRU-IC  | 0.81   | 0.72    | 1            | 0.85          |
| DecompGRU-MSE | 0.73   | 0.77    | 0.85         | 1             |

资料来源: wind, 通联数据, 华创证券

1、IC 测试结果

下面我们统计不同模型的 10 日、20 日窗口的 IC 指标, 10 日 IC 收益计算区间为 T+1~T+11, 20 日以此类推。

图表 8 10 日 IC 统计结果

| 股票池    | 指标       | GRU          |       | DecompGRU   |       |
|--------|----------|--------------|-------|-------------|-------|
|        |          | IC           | MSE   | IC          | MSE   |
| 中证全指   | RankIC   | 0.125        | 0.107 | <u>0.13</u> | 0.11  |
|        | RankICIR | 0.99         | 1.1   | 1.1         | 1.09  |
|        | IC 胜率    | 0.86         | 0.87  | <u>0.88</u> | 0.87  |
| 沪深 300 | RankIC   | <u>0.094</u> | 0.079 | 0.09        | 0.07  |
|        | RankICIR | 0.55         | 0.52  | <u>0.58</u> | 0.45  |
|        | IC 胜率    | 0.72         | 0.7   | 0.72        | 0.66  |
| 中证 500 | RankIC   | <u>0.091</u> | 0.074 | 0.089       | 0.073 |
|        | RankICIR | 0.65         | 0.62  | <u>0.67</u> | 0.56  |
|        | IC 胜率    | 0.77         | 0.75  | 0.77        | 0.72  |

|         |          |       |       |              |             |
|---------|----------|-------|-------|--------------|-------------|
| 中证 1000 | RankIC   | 0.115 | 0.095 | <u>0.116</u> | 0.098       |
|         | RankICIR | 0.88  | 0.88  | <u>0.94</u>  | 0.85        |
|         | IC 胜率    | 0.83  | 0.82  | 0.83         | <u>0.84</u> |

资料来源: wind, 华创证券

图表 9 20 日 IC 统计结果

| 股票池     | 指标       | GRU          |       | DecompGRU    |             |
|---------|----------|--------------|-------|--------------|-------------|
|         |          | IC           | MSE   | IC           | MSE         |
| 中证全指    | RankIC   | 0.136        | 0.116 | <u>0.141</u> | 0.117       |
|         | RankICIR | 1.07         | 1.24  | 1.23         | <u>1.27</u> |
|         | IC 胜率    | 0.88         | 0.91  | 0.9          | <u>0.92</u> |
| 沪深 300  | RankIC   | <u>0.105</u> | 0.088 | 0.095        | 0.067       |
|         | RankICIR | 0.6          | 0.59  | 0.6          | 0.45        |
|         | IC 胜率    | 0.74         | 0.71  | <u>0.75</u>  | 0.66        |
| 中证 500  | RankIC   | <u>0.097</u> | 0.08  | 0.094        | 0.074       |
|         | RankICIR | 0.68         | 0.67  | <u>0.71</u>  | 0.58        |
|         | IC 胜率    | 0.79         | 0.8   | 0.78         | 0.74        |
| 中证 1000 | RankIC   | 0.127        | 0.107 | <u>0.129</u> | 0.107       |
|         | RankICIR | 0.99         | 1.05  | <u>1.09</u>  | 1           |
|         | IC 胜率    | 0.87         | 0.86  | 0.87         | 0.87        |

资料来源: wind, 华创证券

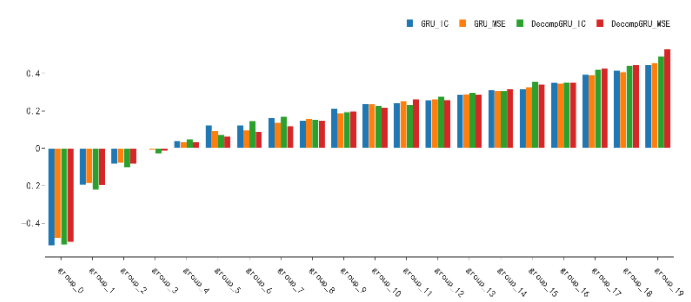
- 在中证全指股票池上，DecompGRU 综合表现强于基线模型，DecompGRU-IC 的 10 日 RankIC 为 0.13，RankIC 为 1.1，胜率为 88%，20 日 RankIC 为 0.141，RankICIR 为 1.23，胜率为 90%；
- 在 300、500 股票池内，DecompGRU 的 IC 表现弱于基线模型，在 1000 上略强于基线；
- 对比不同损失函数，使用 IC 损失函数训练的模型，在大多数 IC 相关指标的表现上优于 MSE 损失。

## 2、分组测试结果

本部分我们对因子进行分组测试。每周最后一个交易日根据当天因子值分组，以次周第一个交易日收盘价进行调仓操作，剔除涨跌停、停牌股票，不计交易成本；

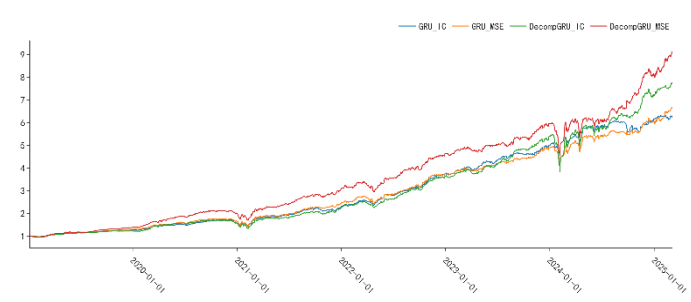
下图展示每一组的年化收益以及多头 TOP 组的超额收益时序走势对比，股票池为中证全指，分组数设置为 20。

图表 10 中证全指-20 分组年化收益对比



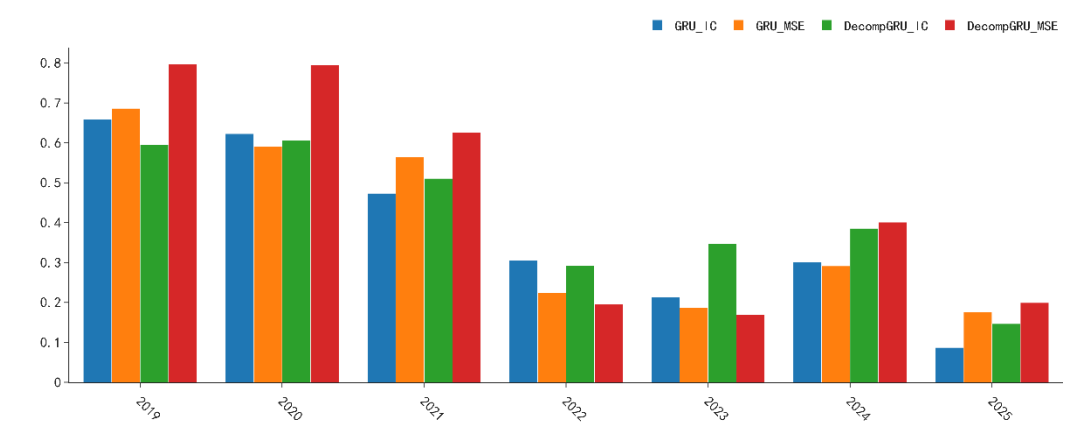
资料来源: wind, 华创证券

图表 11 中证全指-TOP 组超额走势对比



资料来源: wind, 华创证券

图表 12 中证全指-TOP 组逐年收益对比



资料来源: wind, 华创证券

图表 13 中证全指-TOP 组绩效统计

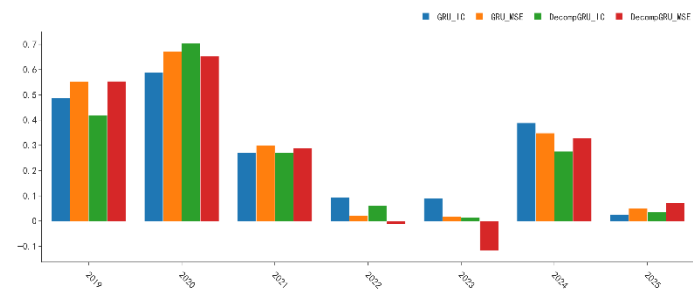
| 模型        | 损失函数 | 年化收益率(%) | 夏普比率 | 最大回撤(%) | 超额年化(%) | 平均单边换手 |
|-----------|------|----------|------|---------|---------|--------|
| GRU       | IC   | 44.61    | 1.61 | -33.31  | 38.53   | 0.53   |
|           | MSE  | 45.6     | 1.57 | -30.4   | 39.52   | 0.5    |
| DecompGRU | IC   | 49.18    | 1.58 | -36.73  | 43.1    | 0.52   |
|           | MSE  | 52.82    | 1.61 | -35.49  | 46.74   | 0.53   |

资料来源: wind, 华创证券

- 在中证全指股票池内，DecompGRU 表现优于基线模型，在回测区间内，DecompGRU-MSE 年化收益率为所有模型中最高，达到 52.82%，相比最优的 GRU-base 模型提高 7.2%；DecompGRU-IC 表现次之，为 49.18%；
- DecompGRU 在小市值风格的暴露更大，24 年年初回撤幅度明显大于基线模型；
- 从逐年角度，DecompGRU-IC 在近三年均跑赢基线，DecompGRU-MSE 在不同年份差异较大，在 19、20、21、24 年表现亮眼，显著强于基线模型，在 22、23 年则反之。

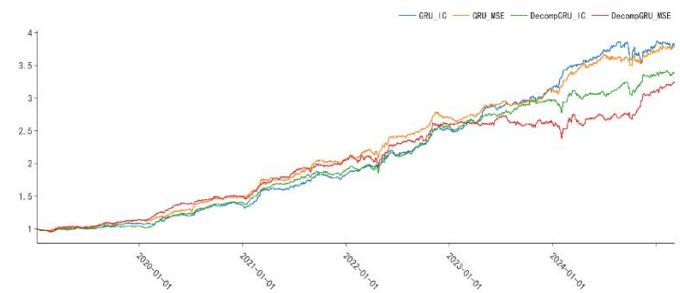
下面我们进一步展示 300、500、1000 宽基内分组测试的 TOP 组统计结果。中证 1000、中证 500 分组数设置为 10，沪深 300 分组数设置为 5，其余设置同上。

图表 14 沪深 300-TOP 组逐年收益对比



资料来源: wind, 华创证券

图表 15 沪深 300-TOP 组超额走势对比



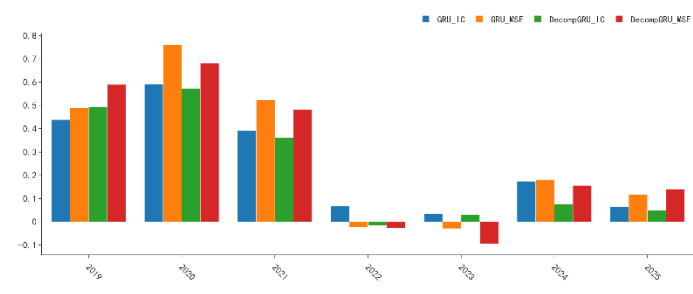
资料来源: wind, 华创证券

图表 16 沪深 300-TOP 组绩效统计

| 模型        | 损失函数 |           | 全区间          | 2019         | 2020         | 2021         | 2022         | 2023         | 2024         | 2025-02-28  |
|-----------|------|-----------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| GRU       | IC   | 超额收益(%)   | <u>25.34</u> | 8.95         | 24.67        | 34.9         | <u>38.05</u> | <u>24.44</u> | <u>21.95</u> | -0.7        |
|           |      | 超额夏普比     | <u>2.86</u>  | 1.58         | 3.09         | 3.4          | <u>4.27</u>  | <u>3.25</u>  | <u>2.15</u>  | -0.52       |
|           |      | 超额最大回撤(%) | -8.24        | -3.08        | -4.35        | -4.23        | -3.83        | -2.6         | -8.24        | -2.94       |
|           | MSE  | 超额收益(%)   | 25.25        | 14.06        | 32.02        | <u>38.04</u> | 29.7         | 16.49        | 18.79        | 1.74        |
|           |      | 超额夏普比     | 2.93         | <u>2.38</u>  | 3.64         | 3.78         | 3.09         | 2.99         | 2.07         | 1.43        |
|           |      | 超额最大回撤(%) | -6.11        | -2.51        | -3.32        | -2.8         | -6.11        | -2.53        | -4.54        | -1.65       |
| DecompGRU | IC   | 超额收益(%)   | 22.78        | 4.22         | <u>34.68</u> | 35.6         | 34.7         | 15.99        | 13.1         | 0.31        |
|           |      | 超额夏普比     | 2.71         | 0.73         | <u>4.25</u>  | 4.3          | 3.3          | 2.85         | 1.4          | 0.34        |
|           |      | 超额最大回撤(%) | -7.43        | -5.06        | -2.65        | -3.01        | -6.2         | -2.66        | -7.14        | -2.28       |
|           | MSE  | 超额收益(%)   | 21.93        | <u>14.36</u> | 30.78        | 37.9         | 25.72        | 1.53         | 18.32        | <u>3.9</u>  |
|           |      | 超额夏普比     | 2.29         | 2.2          | 3.71         | <u>4.33</u>  | 2.16         | 0.27         | 1.57         | <u>3.25</u> |
|           |      | 超额最大回撤(%) | -13.2        | -4.61        | -3.07        | -3.94        | -8.49        | -6.95        | -9.62        | -1.9        |

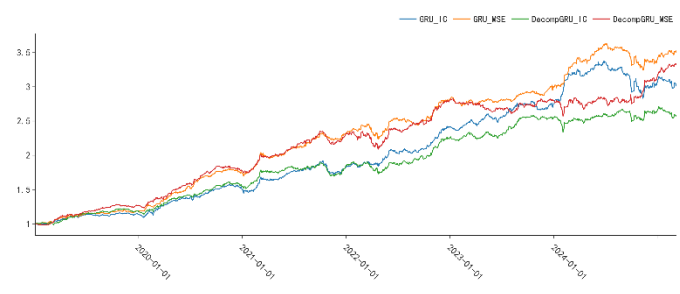
资料来源: wind, 华创证券

图表 17 中证 500-TOP 组逐年收益对比



资料来源: wind, 华创证券

图表 18 中证 500-TOP 组超额走势对比



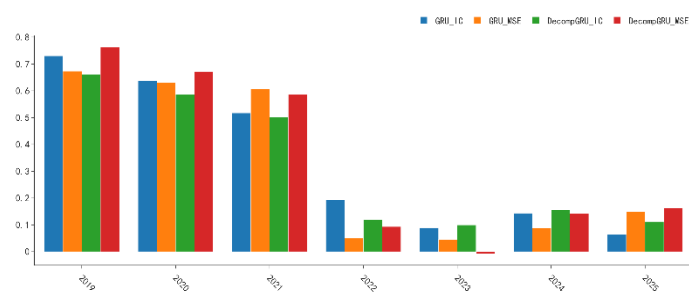
资料来源: wind, 华创证券

图表 19 中证 500-TOP 组绩效统计

| 模型         | 损失函数 |           | 全区间          | 2019         | 2020         | 2021         | 2022         | 2023         | 2024         | 2025-02-28  |
|------------|------|-----------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| GRU        | IC   | 超额收益(%)   | 20.52        | 14.53        | 30.8         | 22.03        | <u>31.98</u> | <u>13.43</u> | 12.85        | -2.18       |
|            |      | 超额夏普比     | 1.93         | 2.01         | 2.59         | 1.97         | <u>3.22</u>  | 1.69         | 1.09         | -1.14       |
|            |      | 超额最大回撤(%) | -16.34       | -3.4         | -5.26        | -9.86        | -3.81        | -4.71        | -16.34       | -5.74       |
|            | MSE  | 超额收益(%)   | <u>23.72</u> | 19.07        | <u>46.05</u> | <u>33.59</u> | 21.39        | 6.64         | <u>13.97</u> | 2.96        |
|            |      | 超额夏普比     | <u>2.37</u>  | 2.63         | <u>3.79</u>  | <u>3.03</u>  | 2.05         | 1.09         | <u>1.42</u>  | 2.27        |
|            |      | 超额最大回撤(%) | -10.72       | -3.06        | -4.28        | -5.05        | -9.16        | -4.52        | -10.72       | -2.01       |
| Decomp GRU | IC   | 超额收益(%)   | 17.28        | 18.8         | 30.09        | 19.37        | 22.08        | 13.02        | 4.62         | -3.46       |
|            |      | 超额夏普比     | 1.75         | 2.62         | 2.63         | 1.82         | 2.05         | <u>2.03</u>  | 0.51         | -2.3        |
|            |      | 超额最大回撤(%) | -10.16       | -4.33        | -4.06        | -10.16       | -9.55        | -4.54        | -8.81        | -6          |
|            | MSE  | 超额收益(%)   | 22.55        | <u>27.14</u> | 39.57        | 30.32        | 21.32        | -0.17        | 13.28        | <u>5.15</u> |
|            |      | 超额夏普比     | 2.11         | <u>3.51</u>  | 3.42         | 2.85         | 1.7          | 0.01         | 1.17         | <u>3.74</u> |
|            |      | 超额最大回撤(%) | -11.26       | -3.45        | -3.83        | -8.38        | -11.18       | -7.37        | -8.25        | -2.13       |

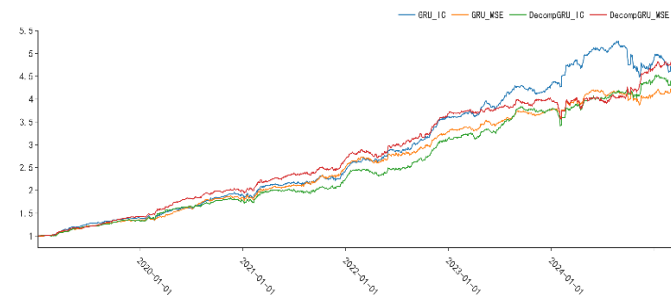
资料来源: wind, 华创证券

图表 20 中证 1000-TOP 组逐年收益对比



资料来源: wind, 华创证券

图表 21 中证 1000-TOP 组超额走势对比



资料来源: wind, 华创证券

图表 22 中证 1000-TOP 组绩效统计

| 模型         | 损失函数 |           | 全区间          | 2019         | 2020         | 2021         | 2022         | 2023         | 2024         | 2025-02-28  |
|------------|------|-----------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------------|
| GRU        | IC   | 超额收益(%)   | 29.44        | 39.69        | 35.62        | 28.35        | <u>48.44</u> | 18.88        | 13.91        | -5.61       |
|            |      | 超额夏普比     | 2.5          | 5.27         | 2.93         | 2.1          | <u>4.75</u>  | 2.69         | 1.03         | -2.37       |
|            |      | 超额最大回撤(%) | -15.11       | -1.82        | -7.96        | -5.16        | -2.74        | -4.57        | -15.11       | -8.33       |
|            | MSE  | 超额收益(%)   | 27.49        | 35.19        | 36.14        | 36.05        | 31.27        | 14.43        | 9.48         | 2.38        |
|            |      | 超额夏普比     | 2.67         | 5.09         | <u>3.51</u>  | 2.7          | 3.09         | 2.54         | 0.89         | 1.87        |
|            |      | 超额最大回撤(%) | -8.07        | -1.62        | -4.97        | -5.33        | -5.06        | -3.29        | -8.07        | -2.51       |
| Decomp GRU | IC   | 超额收益(%)   | 28.24        | 34.1         | 31.93        | 27.03        | 39.53        | <u>20.25</u> | <u>17.33</u> | -1.33       |
|            |      | 超额夏普比     | 2.51         | 5.21         | 2.65         | 2.03         | 3.68         | <u>2.96</u>  | <u>1.39</u>  | -0.7        |
|            |      | 超额最大回撤(%) | -10.91       | -2.74        | -4.52        | -4.81        | -6.78        | -3.99        | -10.12       | -5.39       |
|            | MSE  | 超额收益(%)   | <u>30.56</u> | <u>42.86</u> | <u>39.44</u> | <u>34.71</u> | 36.91        | 9.06         | 16.87        | <u>3.7</u>  |
|            |      | 超额夏普比     | <u>2.81</u>  | <u>6.09</u>  | 3.4          | <u>2.85</u>  | 3.54         | 1.47         | 1.31         | <u>2.68</u> |
|            |      | 超额最大回撤(%) | -11.67       | -1.86        | -3.42        | -3.35        | -6.27        | -3.67        | -11.11       | -2.84       |

资料来源: wind, 华创证券

综合以上测试结果：

- 从全区间看，DecompGRU 在小市值股票池的表现相对更好；
- 在 300、500 股票池内，DecompGRU 的整体表现跑输基线模型，其中 22-24 年跑输基线较多；24 年末基线 GRU 超额回撤区间，DecompGRU 相对更强；
- 在 1000 股票池内，DecompGRU 整体表现略强于基线模型，DecompGRU-MSE 在全区间超额年化收益为 30.56%，除 23 年大幅跑输基线外，其余年份均强于其他模型；DecompGRU-IC 则与基线模型相近。

总体而言，DecompGRU 在小市值股票池中相比基线 GRU 提高较为明显；而在 300、500 股票池中，DecompGRU 在整体统计上弱于基线模型，仅在局部区间能跑赢基线。

DecompGRU 在 GRU 的基础上引入时序的趋势分解以及截面交互，在框架流程上区别于基线模型。模型虽然仍有明显低波、低流偏好，但在市值、贝塔等风格的暴露不同于基线，能在近年风格多变的市场环境下提供更多选择。此外，我们还注意到加权 MSE 损失同样对模型的风格倾向存在较大影响，在保持输入端为 OHLCV 等基本字段时，不同模型间局部差异仍可以是显著的，为后续模型融合研究提供更多样的选择。

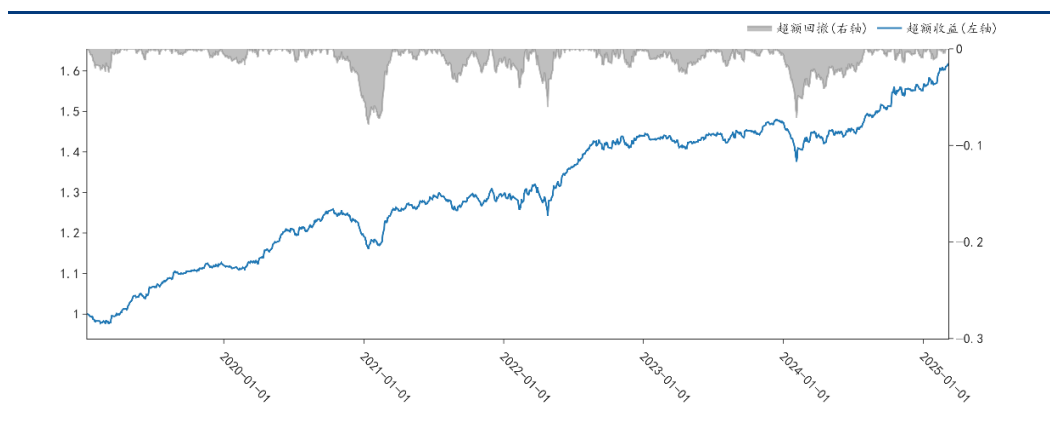
#### （四）指增测试

在本节，我们使用基于 MSE 损失的 DecompGRU-MSE，在沪深 300、中证 500 以及中证 1000 上进行指增测试；测试区间为 20190101~20250228。

- 约束条件为（1）指数成分股占比大于 80%（2）单只股票权重相对偏离上限为 0.8%（3）Barra 风格暴露约束 0.3，行业暴露约束 0.02（5）换手约束双边 30%；
- 调仓频率为周频，根据每周最后一个交易日的因子值在次周首个交易日进行调仓，剔除 ST、涨跌停、停牌股票；
- 回测设定成本为双边千 2.5。

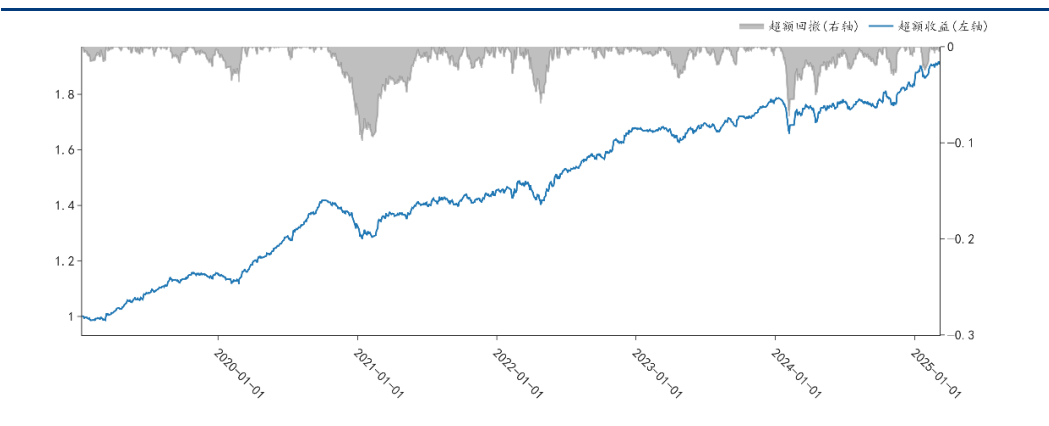
DecompGRU-MSE 在 300 / 500 / 1000 指增组合的年化超额收益分别为 8.48%、11.65%、17.11%，跟踪误差分别为 4.8、5.88、6.69，平均双边换手为 0.32、0.34、0.34，其他回测图表如下：

图表 23 沪深 300 指增超额走势



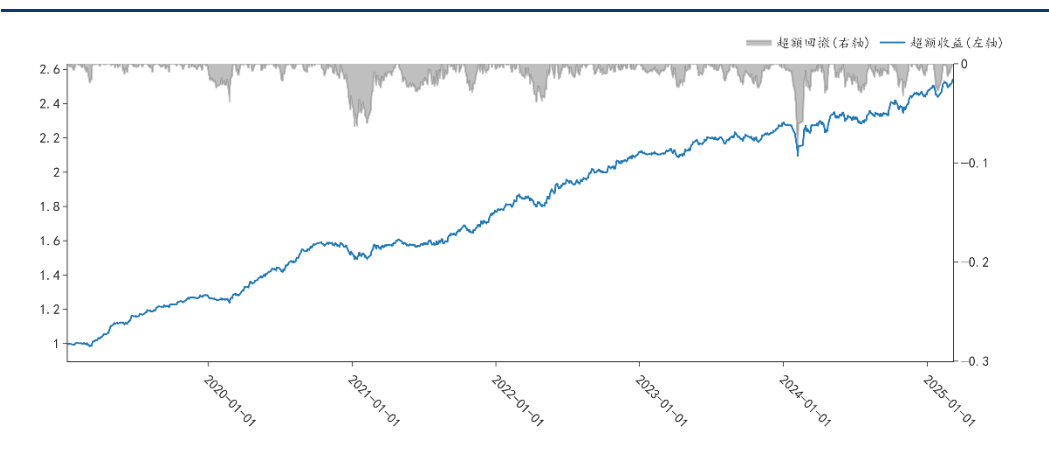
资料来源：wind，华创证券

图表 24 中证 500 指增超额走势



资料来源: wind, 华创证券

图表 25 中证 1000 指增超额走势



资料来源: wind, 华创证券

图表 26 指增组合统计

|         | 基准        | 全区间   | 2019  | 2020  | 2021  | 2022  | 2023  | 2024  | 2025-02-28 |
|---------|-----------|-------|-------|-------|-------|-------|-------|-------|------------|
| 沪深 300  | 超额收益(%)   | 8.48  | 12.03 | 6.44  | 8.53  | 11.21 | 2.46  | 5.85  | 3.69       |
|         | 超额夏普      | 1.72  | 3.32  | 1.56  | 1.72  | 1.73  | 0.75  | 1.16  | 4.75       |
|         | 超额最大回撤(%) | -7.84 | -2.47 | -5.31 | -3.5  | -6.04 | -2.62 | -6.71 | -1.1       |
|         | Calmar 比率 | 1.08  | 5.18  | 1.26  | 2.53  | 1.94  | 0.98  | 0.91  | 3.35       |
| 中证 500  | 超额收益(%)   | 11.65 | 15.58 | 14.94 | 9.55  | 15.35 | 5.66  | 4.38  | 3.67       |
|         | 超额夏普      | 1.91  | 3.46  | 2.53  | 1.48  | 2.2   | 1.37  | 0.7   | 3.37       |
|         | 超额最大回撤(%) | -9.79 | -2.05 | -7.15 | -3.72 | -5.89 | -3.28 | -7.21 | -2.44      |
|         | Calmar 比率 | 1.19  | 8.06  | 2.17  | 2.67  | 2.72  | 1.8   | 0.63  | 1.50       |
| 中证 1000 | 超额收益(%)   | 17.11 | 28.09 | 19.44 | 15.56 | 19.8  | 7.89  | 7.85  | 3.29       |
|         | 超额夏普      | 2.4   | 5.27  | 2.94  | 1.93  | 2.89  | 1.64  | 0.97  | 2.99       |
|         | 超额最大回撤(%) | -8.6  | -1.94 | -4.69 | -2.91 | -3.9  | -2.47 | -8.6  | -2.75      |
|         | Calmar 比率 | 1.99  | 15.48 | 4.31  | 5.56  | 5.31  | 3.34  | 0.95  | 1.19       |

资料来源: wind, 华创证券

#### 四、总结

- 本篇报告我们提出了 DecompGRU 模型，在先前报告的模型框架的基础上进行拓展；我们使用趋势分解模块将初始输入拆分为趋势与残差分量，并用双分支结构进行分别建模；在趋势分支，我们加入一个基于去均值操作的简单截面模块；
- 我们在 150D 数据集，分别以 IC、MSE 两种损失函数训练模型；DecompGRU 与基线模型的相关性最低为 72%，在风格倾向上与基线模型存在差异，在小市值选股域中提高更为明显；
- 在中证全指 20190101~20250228 区间内，以 GRU 模型为基线，DecompGRU 的 10 日 RankIC 为 0.13、RankICIR 为 1.1；20 日 RankIC 为 0.141，RankICIR 为 1.23，均优于基线模型；分组测试中，DecompGRU 的 TOP 组年化收益最高达 52.82%，相比基线实现了 7.2% 的提高；在 300、500 选股域中则整体表现弱于基线；
- 基于 DecompGRU 构建指增组合，在双边换手 30% 的约束下，回测区间内 300、500、1000 指增组合的年化超额收益分别为 8.48%、11.65%、17.11%，跟踪误差为分别 4.8、5.88、6.69。

#### 五、风险提示

策略基于历史数据回测，不保证未来数据的有效性。深度学习模型存在过拟合风险。深度学习模型受随机数影响。本文的模型实现和相关参考文献不完全相同。

#### 六、参考文献

Zeng, Ailing, et al. "Are transformers effective for time series forecasting?" Proceedings of the AAAI conference on artificial intelligence. Vol. 37. No. 9. 2023.

Wu, Haixu, et al. "Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting." Advances in neural information processing systems 34 (2021): 22419-22430.

Yu, Weihao, et al. "Metaformer is actually what you need for vision." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.