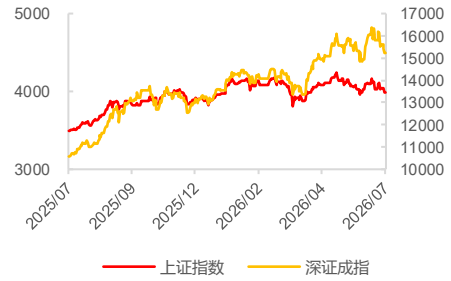
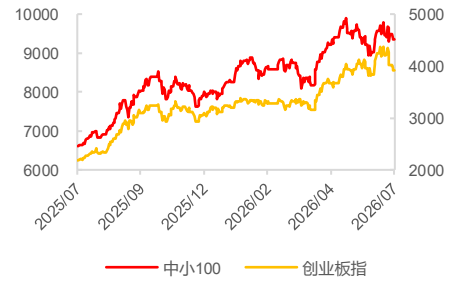
下载日志已记录，仅供内部参考

大盘指数





资料来源：聚源，中邮证券研究所

研究所

分析师:黄子鉴
SAC 登记编号:S1340523090002
Email:huangziyin@cnpsec.com

近期研究报告

《业绩之锚(7)：一季报定价中习惯与“甜区”策略》 - 2026.06.04

策略观点

市场脉搏(2):基于 LSTM-GHMM 混合方案的量化择时与动态仓位管理

- **投资要点**

基于系列首篇报告《市场脉搏（1)》的基础，本报告（2.0 版本）对择时框架进行了系统性迭代升级。

模型迭代：从离散到连续。系本篇将底层模型升级为高斯隐马尔可夫模型（GHMM)，观测变量由离散符号扩展为连续多维向量，发射概率由高斯分布刻画。特征维度不再受限，可将 25 维以上连续量价特征直接馈入模型；并且状态推断输出后验概率分布，软边界表征更贴近风格切换的渐进本质。

特征工程：LSTM 自编码器。与常规用法不同，本报告中的 LSTM 不以预测未来收益率为优化目标，而采用自编码器架构，以最小化输入序列的重建误差为训练目标。编码器将 90 个交易日×25 维的原始特征压缩为 10 维，解码器将其重建还原。训练集与验证集的重建损失同步收敛，且编码器输出的特征分布在训练集与测试集上保持高度一致，为 GHMM 的样本外推断提供了分布稳定的观测输入。

状态识别：过渡态与稳态。GHMM 在无监督条件下自发生成 5 个隐藏状态，其中 1 个为低自维持概率的过渡状态，其余 4 个为高自维持概率的稳态。这一结构表明，市场大部分时间运行于稳定状态中，状态切换通常经由过渡状态完成，而非不同稳态间的直接跳跃。从价格形态看，各状态与特定量价结构存在稳定的对应关系，状态识别在训练集与测试集上表现出较好的一致性。

回测表现与模型边界。2021 年以来多指数回测结果显示，该框架在大多数市场环境下能够有效控制回撤并积累超额收益，但对 2021 年（历史极低振幅下的高速轮动）与 2026 年（K 型极端分化）的结构性行情适应性偏弱。进一步归因发现，超额损失在万得全 A 指数上高度集中于某一特定状态，说明问题不在状态识别层面，而在仓位执行层面——基于历史频度统计的凯利公式在该状态下存在“均值回归”的保守倾向，对“假摔反包”类行情响应不足。据此引入了一个针对的轻量级干预规则，在不改变底层识别逻辑的前提下改善该状态下的适应性。

本模型的核心价值在于提供一套客观的、可重复验证的市场状态参照系，服务于主观择时决策的校准与风控环节。回测表明，该框架在 8 个主要指数上均能实现相对买入持有的正超额收益，具备较好的指数泛用性；同时，超额损失被精准锁定于特定状态，说明模型在识别端已趋于稳定，问题集中于执行端的策略适配。

- **风险提示：**

市场学习效益，统计误差等。

下载日志已记录，仅供内部参考

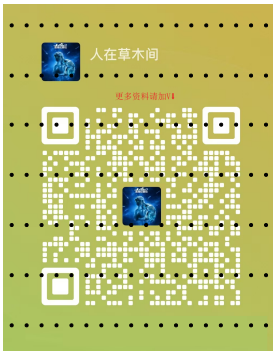
目录

1 从 1.0 到 2.0 的升级路线：让模型看到更多特征 5
1.1 纯 HMM 模型的优势和局限 5
1.2 LSTM-GHMM 混合方案的优势和升级路线 7
2 LSTM-HMM 混合模型架构：从时序特征到概率空间 9
2.1 LSTM-HMM 模型整体架构与流程 9
2.2 特征工程升级：更多的量价细节 10
2.3 LSTM 自编码器：从手工特征到深度特征 11
3 回测验证：2021 年以来多指数回测表现 13
3.1 LSTM-GHMM 的灰箱中学到了什么？ 13
3.2 多指数长期回测结果与模型局限性 15
3.3 基于状态特征的仓位优化思路和改进方向 20
4 风险提示 24

下载日志已记录，仅供内部参考

图表目录

图表 1： 隐马尔科夫链示意图 5
图表 2： 动态调节机制及模型完整择时流程示意图 6
图表 3： HMM 择时框架在万得全 A 上的应用 7
图表 4： HMM 择时框架在中证 1000 上的应用 7
图表 5： LSTM 训练集与验证集的重建损失 12
图表 6： LSTM 编码器输出的各特征维度分布 12
图表 7： 万得全 A 的 GHMM 状态转移矩阵 13
图表 8： 万得全 A 的 HMM 状态转移矩阵 13
图表 9： 万得全 A 训练集和测试集的状态分布 14
图表 10： 万得全 A 各状态的胜率(%)和期望收益 14
图表 11： 万得全 A 在长时间维度上的状态划分 15
图表 12： 上证指数 2021 年以来回测结果 15
图表 13： 沪深 300 指数 2021 年以来回测结果 15
图表 14： 万得全 A 指数 2021 年以来回测结果 16
图表 15： 万得微盘 2021 年以来回测结果 16
图表 16： 创业板指 2021 年以来回测结果 16
图表 17： 科创综指 2021 年以来回测结果 16
图表 18： 中证 1000 指数 2021 年以来回测结果 17
图表 19： 中证 2000 指数 2021 年以来回测结果 17
图表 20： 干预开度后的万得微盘指数回测结果 18
图表 21： 2026 年干预后仓位和回测结果 18
图表 22： 部分主要指数择时超额收益 18
图表 23： 2021 年以来回测结果 18
图表 24： 万得全 A 各状态择时超额走势 19
图表 25： 万得全 A 状态 2 择时超额走势 19
图表 26： 2021 年万得全 A 指数凯利仓位 20
图表 27： 2025 下半年以来万得全 A 指数凯利仓位 20
图表 28： 上证指数 2021 年以来补丁回测结果 21
图表 29： 沪深 300 指数 2021 年以来补丁回测结果 21
图表 30： 万得全 A 指数 2021 年以来补丁回测结果 21
图表 31： 万得微盘 2021 年以来补丁回测结果 21
图表 32： 创业板指 2021 年以来补丁回测结果 22
图表 33： 科创综指 2021 年以来补丁回测结果 22
图表 34： 中证 1000 指数 2021 年以来补丁回测结果 22
图表 35： 中证 2000 指数 2021 年以来补丁回测结果 22



下载日志已记录，仅供内部参考

图表 36： 上证指数补丁及原始相对收益 23
图表 37： 沪深 300 补丁及原始相对收益 23
图表 38： 万得全 A 补丁及原始相对收益 23
图表 39： 万得微盘补丁及原始相对收益 23
图表 40： 中证 1000 补丁及原始相对收益 24
图表 41： 中证 2000 补丁及原始相对收益 24
图表 42： 创业板指补丁及原始相对收益 24
图表 43： 科创综指补丁及原始相对收益 24

下载日志已记录，仅供内部参考

在《市场脉搏（1）：基于隐马尔科夫链与动态调制的量化择时方案》中曾提出过一种使用状态判断而非因子拟合的指数择时思路，在本篇报告中将在系列第一篇报告的基础上更近一步，通过 LSTM（长短期记忆网络）和 HMM（隐马尔可夫模型）的联动实现一个更具备精确性和普适性的动态仓位管理方案。

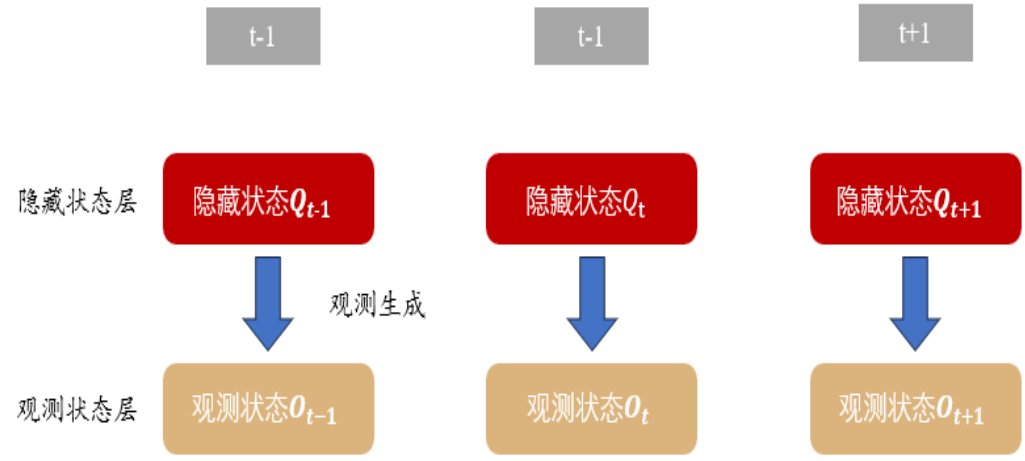
1 从 1.0 到 2.0 的升级路线：让模型看到更多特征

1.1 纯 HMM 模型的优势和局限

在本系列报告的第一篇《市场脉搏（1）：基于隐马尔科夫链与动态调制的量化择时方案》中，使用了 HMM 模型叠加动态调制矩阵的方式来实现一个简洁的宏微观联动择时机制，并且由于 HMM 模型划分市场状态的能力，也天然地近似解决了凯利公式在指数择时面临的“独立重复投注事件”假设问题。

在此先回顾一下 1.0 中的核心假设和设计机制。对于 A 股，如图 1 所示，可以假设存在一个不可直接观测的 “隐藏状态链”，并存在一个由隐藏状态驱动的可观测序列。在 1.0 中，这个假设为 A 股存在一个牛熊状态链，牛熊状态无法直接观测，但其会驱动股价、波动率、振幅等指标发生变化。

图表1：隐马尔科夫链示意图

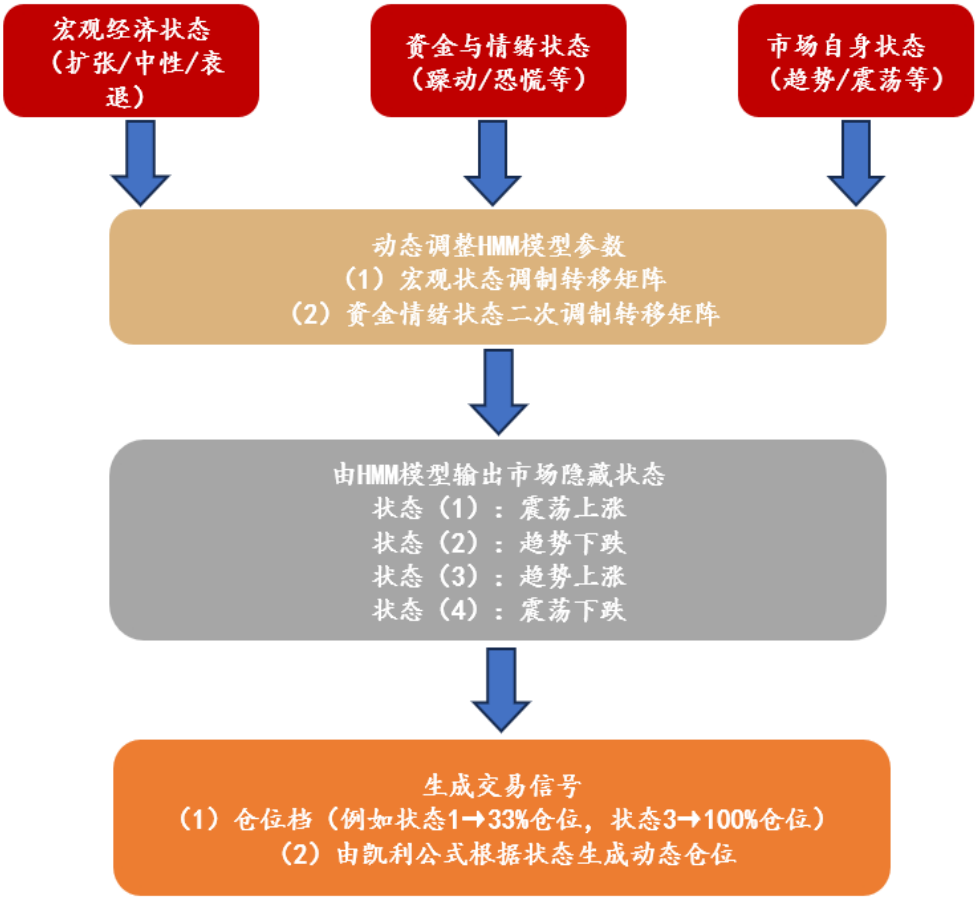


资料来源：中邮证券研究所

为了给 HMM 这个静态模型赋予灵活性，引入了两个基于宏观经济数据和资金情绪的调节矩阵，可以基于宏观环境和市场动态对 HMM 的输出结果进行动态调节。

下载日志已记录，仅供内部参考

图表2：动态调节机制及模型完整择时流程示意图



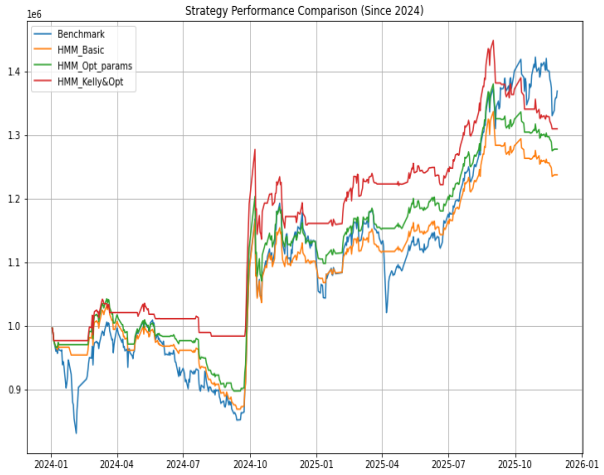
资料来源：中邮证券研究所

通过基于宏观经济数据和资金情绪状态的两个调制矩阵，可以在宏观经济走强、市场情绪亢奋和资金持续流入等状态时动态调高状态转移矩阵向牛市状态转移的概率，反之亦然

但是 1.0 版本的择时模型也存在着一定的局限性。首先，由于在状态划分上仅采用了 “趋势+波动”主观分类，这种划分方式在 2025 年末开始的高噪音市场环境中出现了一定的不适应性。如图 3-4 所示，在 2025 年末基于 1.0 版本的万得全 A 指数择时失效跑输，而中证 1000 在 1.0 版本的 HMM 择时框架下依然能够保持相对原指数的超额收益。这一方面是主观分类标准在不同指数之间的适配性不同，另一方面也是分类时所参考的特征类型过少在追求一般均衡的建模目标下对极端市场风格的适应性较弱。

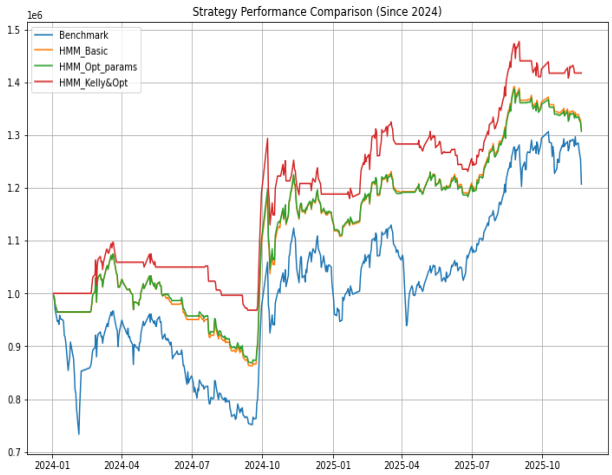
下载日志已记录，仅供内部参考

图表3： HMM 择时框架在万得全 A 上的应用



资料来源：Wind，中邮证券研究所。数据经标准化，基点为 1000

图表4： HMM 择时框架在中证 1000 上的应用



资料来源：Wind，中邮证券研究所。数据经标准化，基点为 1000

其次，由主观分类带来了状态边界切换的问题。由于在 1.0 版本中采取了主观分类的方法并且使用的是离散 HMM，因此在状态边缘的切换是突然且生硬的。简单来说就是市场会在一天完成从牛市到熊市的切换，而更符合金融上直觉的应当是“目前可能在从牛市向熊市切换”的软边界。

1.2 LSTM-GHMM 混合方案的优势和升级路线

由于 1.0 版本中最主要的问题是仅输入了主观划分的状态作为学习特征，所以最优先要升级的部分应当是增加整个模型所“看见”的特征类型。但增加特征类型和采取离散 HMM 基底的本意存在根本性的冲突，采取离散 HMM 本意是将一个填空题简化为选择题，而增加特征（在离散 HMM 中即增加状态）会大幅增加选项，这会导致 HMM 维度爆炸，一般认为 5-6 个状态已经是 HMM 的极限了，继续增加会导致严重的过拟合问题。

为了增加模型所“看见”的特征类型，在 2.0 版本中选择从离散 HMM 切换为高斯隐马尔可夫模型（Gaussian Hidden Markov Model，GHMM）。离散 HMM 与 GHMM 的核心区别在于观测变量的生成方式与数据类型假设。在离散 HMM 中，观测值必须是有限个离散符号（如“上涨/下跌/震荡”），观测概率矩阵描述的是在每个隐藏状态下出现某个离散符号的概率。这种设计天然要求事先对市场数据进行离散化编码，将连续的量价指标人为切割成几个有限的类别。这种“压缩”过程必然伴随信息损失，且离散类别的边界往往带有较强的主观性。更重要的是，当希望

下载日志已记录，仅供内部参考

引入更多维度的市场特征（如动量、波动率、价量关系、资金流向等）时，离散 HMM 的观测空间会呈指数级扩张，远超 HMM 有效推断的能力范围，这正是 1.0 版本不得不将特征限制在“趋势+波动”两个维度的根本原因。

相比之下，GHMM 将观测变量设定为连续的多维向量，其发射概率由高斯分布（正态分布）刻画：每个隐藏状态对应一个均值向量μ_s和一个协方差矩阵Σ_s，观测向量在该状态下出现的概率为：

$$P(o_t \mid S_t = s) = \frac{1}{(2\pi)^{d/2}|\Sigma_s|^{1/2}} \exp\left(-\frac{1}{2}(o_t - \mu_s)^\top \boldsymbol{\Sigma_s^{-1}} (o_t - \mu_s)\right)$$

这一设计带来的第一个关键改进是特征维度的解放：GHMM 直接接受连续数值向量作为输入，不再需要对每个特征进行离散化编码，可以自然地将数十维甚至上百维的连续特征直接馈入模型，而状态空间并不会因此扩张。状态数依然由设定的隐藏状态数量决定，与观测特征的维度无关。这恰好解决了 1.0 版本中“特征匮乏”的核心痛点。

第二个关键改进在于状态边界的软性化。在离散 HMM 中，观测值被硬性归入某一离散类别，状态推断也因此呈现"非此即彼"的硬切换特征。市场在模型判定下会在某一天从"趋势上涨"突然跳变为"趋势下跌"。而在 GHMM 中，每个时刻的状态推断是基于后验概率分布的：

$$P(S_t = s \mid o_1, \ldots, o_t) \propto P(o_t \mid S_t = s) \cdot \sum_i P(S_t = s \mid S_{t-1} = i) \cdot P(S_{t-1} = i \mid o_1, \ldots, o_{t-1})$$

这一概率分布天然给出了市场"当前处于每个状态的可能性有多大"的连续度量。当模型输出"趋势上涨概率 55%、震荡上涨概率 40%、趋势下跌概率 5%"时，其含义并非状态已经完成切换，而是市场正处于从某一状态向另一状态的过渡区间。这种软边界表征更贴近金融市场的真实运作方式，风格切换和牛熊转换从来都不是在一夜之间完成的，而是经历一个概率逐步累积和转移的过程。GHMM 的状态概率分布恰好提供了对这一渐变过程的数学描述，使得模型的择时信号更加平滑、更具前瞻性。

LSTM 在 2.0 方案中扮演的核心角色是深度时序特征编码器，承担着将原始量价数据转化为 GHMM 所需的高质量观测向量的任务。具体而言，LSTM 接收过去的原始市场数据，利用其循环结构沿时间轴逐日扫描，最终将这段高维时间序列压缩提炼为一个固定长度的低维稠密向量，作为 GHMM 的观测输入。引入 LSTM 的必要性体现在两个层面：一是捕捉长程时序依赖，市场状态（如“趋势上涨”）并

下载日志已记录，仅供内部参考

非由某一日的价格决定，而是由过去数周甚至数月价格的演化路径所决定，LSTM 的门控机制天然擅长建模这种跨度长达 90 个交易日的依赖关系；二是自动特征提取与降维，LSTM 作为数据驱动的特征学习器，能够在训练过程中自动挖掘量价数据中隐含的高阶非线性交互模式，同时将高维的原始输入压缩，这一降维过程实质上是对数据施加了强正则化约束，有效规避了高维观测下 GHMM 协方差矩阵参数激增所引发的维度灾难与过拟合风险。

本方案中 LSTM 与常规用法的根本区别在于训练目标的不同。常规 LSTM 以预测未来为优化目标，需要依赖标注好的未来标签，在金融数据低信噪比的环境下极易过拟合，样本外表现往往大幅衰退。而本方案采用自编码器架构训练 LSTM，其优化目标是最小化输入序列的重建误差，即仅需当前的量价数据，无需任何未来标签。这一设计将 LSTM 的角色从“预测器”转变为“表征器”，不再试图猜测明天涨跌，而是专注学习如何将过去 90 天的市场信息精炼成一张信息富集、冗余剔除的“摘要卡片”。当市场风格发生切换时，这张卡片仍能忠实记录当下的市场微观结构特征，交由 GHMM 进行后续的概率推断和状态识别，从而实现了特征提取与状态推断的解耦，显著提升了模型对市场风格切换的适应性。

2 LSTM-HMM 混合模型架构：从时序特征到概率空间

2.1 LSTM-HMM 模型整体架构与流程

本报告提出的 LSTM-GHMM 混合择时模型，整体架构分为五个层次。如图 4 所示，模型沿“数据层 → 特征工程层 → LSTM 编码层 → GHMM 状态层 → 仓位决策层”逐级递进，形成从原始数据到最终交易信号的完整决策链条。

1.0 版本的本质是“人工定义特征 → 离散状态划分 → HMM 推断”，其中信息瓶颈在于人工定义的特征类型有限且状态硬边界不符合市场渐变规律。2.0 版本将这一流程改造为“原始数据 → LSTM 自动编码 → 连续特征 → GHMM 软推断”，LSTM 替代了人工特征工程，GHMM 替代了离散状态空间，两者结合使得模型能够从更丰富的信息中更平滑地感知市场状态变化，为后续的动态调制和仓位决策提供更高质量的输入。

下载日志已记录，仅供内部参考

2.2 特征工程升级：更多的量价细节

特征工程是整个模型的信息源头，其质量直接决定了后续 LSTM 编码与 GHMM 状态推断的上限。1.0 版本采用“趋势+波动”的二维离散划分，将 10 日收益率和 10 日 ATR 变化率相对于 90 天滚动窗口的均值±0.524σ阈值，将市场划分为趋势上涨、趋势下跌、温和上涨、温和下跌、宽幅震荡、窄幅震荡 6 种观测状态。这一划分方式虽然直观且具备经济学含义，但其局限性同样明显：仅使用两个指标描述市场的全部状态，相当于用两个坐标轴刻画高维金融市场的全貌，大量有用信息（如动量强度、价格位置、市场结构等）在离散化过程中被丢失。

2.0 版本将观测特征从 2 维扩展至 25+维连续特征集，全部特征采用 90 天滚动窗口进行 z-score 标准化，确保特征的时序可比性。扩展后的特征集涵盖六个类别：

①动量特征（4 维）：分别计算 10 日、20 日、60 日窗口的累计收益率，捕捉短、中、长三个时间维度的价格趋势强度。

②波动率特征（4 维）：分别计算 10 日、20 日、60 日窗口的已实现波动率（年化），以及基于滚动分位数的波动率状态指标（0-1 连续值），反映市场在不同时间尺度上的波动特征。

③价格位置特征（4 维）：包括价格相对于 20 日均线的偏离度、突破 20 日高点的幅度、跌破 20 日低点的幅度、以及价格在 20 日价格区间内的相对位置。这些特征刻画当前价格在历史价格结构中的相对位置，为状态判断提供锚点。

④市场结构特征（4 维）：包括收益率的一阶自相关系数（滞后 1 日）、20 日滚动偏度、20 日滚动峰度、以及价格趋势强度（方向一致性指标），从统计结构层面描述市场的微观行为模式。

⑤极端分位数特征（2 维）：20 日收益率分布的 95%分位数和 5%分位数，用于捕捉市场尾部风险的极端状态。

⑥趋势-波动 6 状态软概率（6 维）：这是 2.0 版本对 1.0 六状态划分理念的继承与升级。不同于 1.0 版本的离散硬分类（每个时刻唯一属于 6 种状态之一），2.0 版本采用非对称双 sigmoid 函数计算 6 种状态的软概率。

上述六个类别的特征中，前五类送入 LSTM 编码器进行时序压缩，第六类本身已是概率形式，直接作为 GHMM 的观测输入与 LSTM 输出拼接，形成最终的观测

下载日志已记录，仅供内部参考

向量。这一扩展特征集的设计背后，贯穿了模型设计的核心逻辑：在不引入未来信息的前提下，尽可能完整地呈现市场当前的多维状态。后续 LSTM 编码层的作用，正是从这 25 维特征中进一步提炼去除冗余后的核心市场状态。

2.3 LSTM 自编码器：从手工特征到深度特征

上一节所构建的 25 维扩展特征集，虽然在信息覆盖面上远优于 1.0 版本的二维离散状态，但直接将这些特征送入 GHMM 仍面临两个问题：其一，25 维观测向量意味着 GHMM 的协方差矩阵需要估计大量自由参数，在有限样本下极易过拟合；其二，这些特征在时间轴上存在显著的序列依赖关系，而 GHMM 本身不具备时序特征提取能力，它只能利用当期观测值进行状态推断。因此，需要在特征工程层与 GHMM 状态层之间引入一个时序特征编码器，将高维原始特征压缩为低维的、富含时序信息的稠密向量。

LSTM 自编码器（LSTM Autoencoder）正是为此设计。它由编码器（Encoder）和解码器（Decoder）两部分组成：编码器接收过去全部 25 维特征构成的时间窗口，利用 LSTM 沿时间轴逐日处理，将整个窗口的信息汇聚为最后一个时间步的隐藏状态，再通过一个线性映射层输出固定长度的瓶颈向量。解码器则接收瓶颈向量，通过另一个 LSTM 将其扩展回原始序列长度，并重建出与输入相同维度的时序数据。通过最小化输入序列与重建序列之间的均方误差（MSE），自编码器被强制学习如何在 10 维瓶颈向量中保留原始 90 天×25 维信息中最重要的部分，即那些与市场状态识别最相关的时序演化模式。

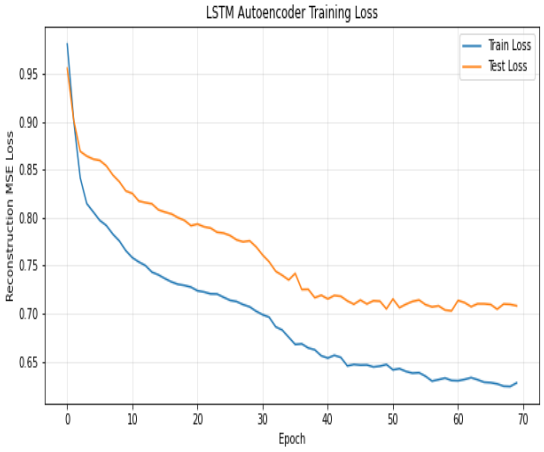
采用自编码器架构而非有监督预测架构，是出于对金融数据特性的审慎考量。常规 LSTM 以预测未来收益率为目标，本质上是在低信噪比的金融数据中寻找统计规律，极易捕捉到样本内噪声而非真实信号，导致样本外性能大幅衰退。自编码器的训练目标是重建输入，不涉及任何未来信息，仅学习刻画当前市场状态所需的时序特征压缩与还原，因此泛化能力更强，与 GHMM 的“状态识别”而非“未来预测”定位更为契合。

LSTM 自编码器的训练效果与特征质量可通过图 4-5 进行验证。如图 5 所示，训练集与验证集的重建损失在 100 个 epoch 内持续下降，两条曲线同步收敛且未出现显著分化，表明模型在有效学习时序特征压缩表征的同时，未发生明显的过拟合现象，L2 正则化与早停机制共同保障了训练的稳健性。如图 6 所示，LSTM

下载日志已记录，仅供内部参考

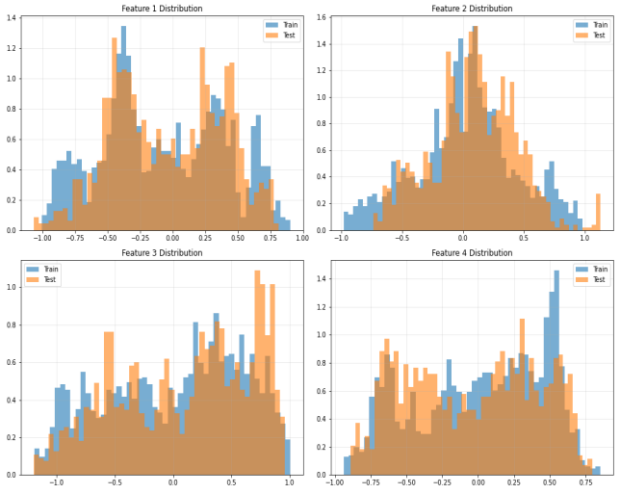
编码器输出的各特征维度在训练集与测试集上的分布形态高度一致，各特征的箱体位置、四分位区间及异常值范围在两组数据间基本重合，未见显著的分布漂移。这一性质对后续 GHMM 的样本外推断至关重要：GHMM 假设观测特征在各隐藏状态下服从固定的高斯分布，特征分布的跨期一致性是该假设成立的前提条件，也是模型在样本外保持稳健性的基础保障。综合训练损失的稳定收敛与特征分布的跨期一致性两方面的证据，可以认为 LSTM 自编码器成功将高维的原始特征（90 天×25 维）压缩为低维的稠密向量，在保留核心时序信息的同时消除了特征冗余，为 GHMM 的状态推断提供了信息富集且分布稳定的观测输入。

图表5： LSTM 训练集与验证集的重建损失



资料来源：Wind，中邮证券研究所

图表6：LSTM 编码器输出的各特征维度分布



资料来源：Wind，中邮证券研究所

下载日志已记录，仅供内部参考

3 回测验证：2021 年以来多指数回测表现

3.1 LSTM-GHMM 的灰箱中学到了什么？

在以 2016-2020 年的 A 股数据为训练集进行训练后，对 2021 年以来的 A 股市场进行多指数回测。由于采取的模型底座是 GHMM，最核心的部分依然是对状态的划分。在此不一一列举各个指数的具体情况，仅以万得全 A 作为范式进行说明。如表 7 所示为万得全 A 指数在 5 个隐藏状态下的转移矩阵，作为对比的是表 8 所示 1.0 版本中万得全 A 指数在 4 个隐藏状态下的转移矩阵，两个状态转移矩阵均表现出了典型的金融时序数据常见的强自维持倾向。相较之下，2.0 版本的状态转移矩阵中状态 2-5 的自维持倾向要显著强于 1.0 版本，在 2.0 版本中的状态 1-4 很难发生状态的转移。2.0 版本最显著的差异在于出现了一个非常容易发生状态转移的状态 0，其自我维持的概率不足五成，是一个不稳定的过渡状态。

图表7： 万得全 A 的 GHMM 状态转移矩阵

	状态 0	状态 1	状态 2	状态 3	状态 4
状态 0	44.1%	11.5%	42.6%	0.0%	1.7%
状态 1	5.4%	87.1%	3.2%	1.8%	2.5%
状态 2	1.0%	2.7%	95.1%	0.7%	0.6%
状态 3	0.2%	2.1%	1.7%	93.8%	2.2%
状态 4	0.4%	1.5%	0.8%	1.1%	96.2%

资料来源：Wind，中邮证券研究所

图表8：万得全 A 的 HMM 状态转移矩阵

	震荡上涨	趋势下跌	趋势上涨	震荡下跌
震荡上涨	76.4%	5.6%	12.6%	5.4%
趋势下跌	10.7%	82.5%	0.4%	6.4%
趋势上涨	16.1%	0.0%	83.4%	0.4%
震荡下跌	13.1%	10.9%	0.5%	75.5%

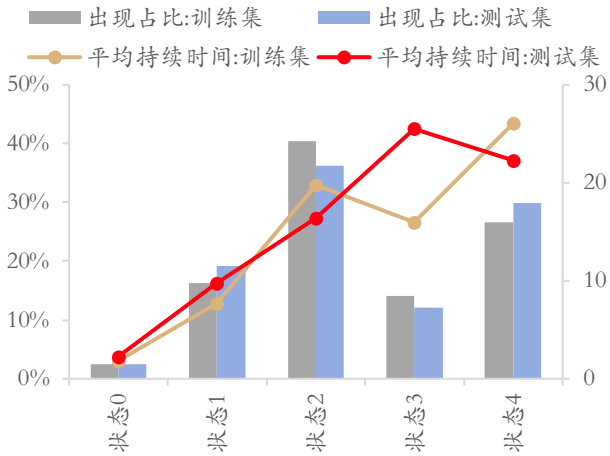
资料来源：Wind，中邮证券研究所

这个不稳定过渡状态的出现正是 2.0 版本在状态识别方面进步的体现。GHMM 从数据中自主发现了 5 个状态，其中 4 个是高度稳定的驻留状态，第 5 个则是低自维持的过渡状态。模型没有被预先告知“应该有几个状态”或“应该有什么类型的状态”，纯粹根据 LSTM 提取的连续特征向量和高斯似然推断，自主识别出了过渡状态的存在。这说明 GHMM 成功捕捉到了 1.0 版本中被人为压缩掉的、市场状态切换过程中的中间阶段。

下载日志已记录，仅供内部参考

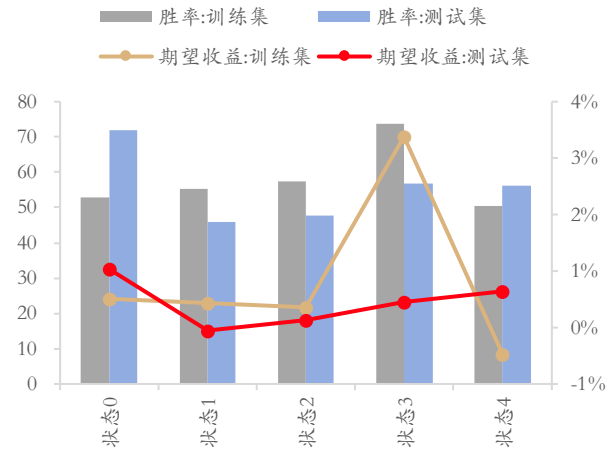
当然在正式进行回测之前还需要确定 GHMM 确实学习到了稳定的市场状态。如图 9 所示为针对万得全 A 训练的 GHMM 在训练集和测试集上的状态分布和其持续时间，其分布特征和平均时间基本相同，可以认为 GHMM 学习到的市场状态是稳定的。状态 3 的平均持续时间差异相对较大，但考虑到状态 3 较少的样本数量，这种差异是可以接受的。

图表9： 万得全 A 训练集和测试集的状态分布



资料来源：Wind，中邮证券研究所

图表10：万得全 A 各状态的胜率(%)和期望收益



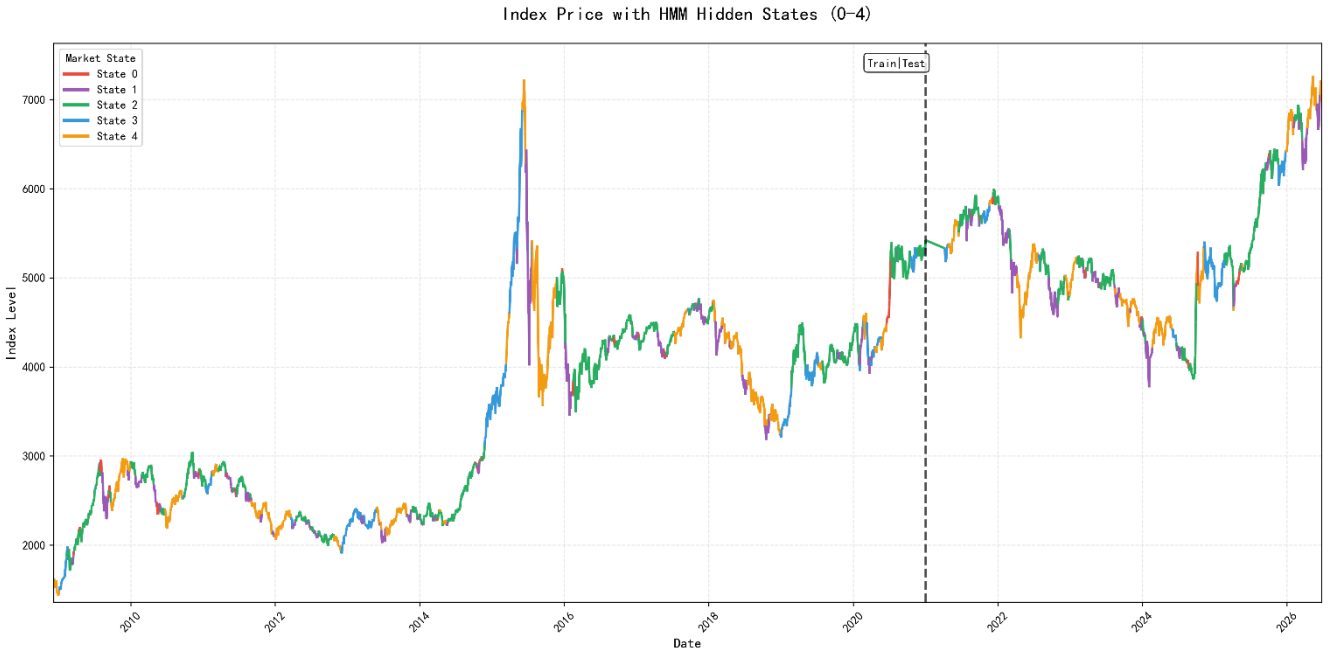
资料来源：Wind，中邮证券研究所

此时应该已经有读者注意到 2.0 版本中并没有对状态进行符合金融市场直觉的牛熊划分，主要原因在于引入了更多的特征类型后，GHMM 识别的结果显示牛熊并不一定是市场自身的状态，可能更多是人类的主观感受。如图 10 所示为 5 个状态在训练集和测试集中的每日胜率和期望收益率，即使考虑到市场环境不同带来的胜率和期望收益基准变动，4 个稳定状态代表的“牛熊状态”也并不稳固，此处就涉及到了关于 GHMM 状态识别结果的基础问题“在改为接受非主观标定的多类型量价特征后，学习到的状态底层逻辑究竟是什么？”

量价形态和波动大小是 GHMM 学习到的主要内容。如图 11 所示为在训练集和测试集上的各状态分段，可见相比显得缺乏规律的收益率统计结果，在图形上各状态在训练集和测试集上更具一致性，比如状态 1 常见于下跌底部的形态，而状态 4 则更多指向大幅的双向波动。显然，GHMM 基于 LSTM 特征向量的高斯似然聚类，确实在无监督条件下自发识别出了具有不同价格形态特征的时间段。尽管受限于深度特征的非线性纠缠特性，无法逐一还原模型具体依据了何种原始指标做出判定，但状态标签与价格形态之间的视觉一致性，从侧面验证了 LSTM-GHMM 在提取市场结构信息方面的有效性。

下载日志已记录，仅供内部参考

图表11：万得全 A 在长时间维度上的状态划分

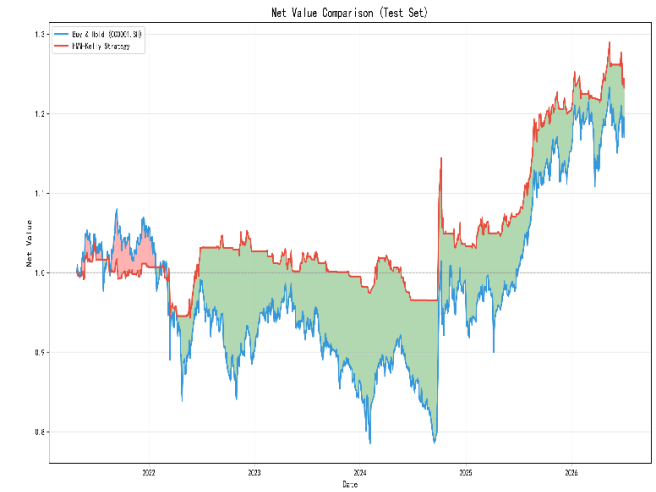


资料来源：Wind，中邮证券研究所

3.2 多指数长期回测结果与模型局限性

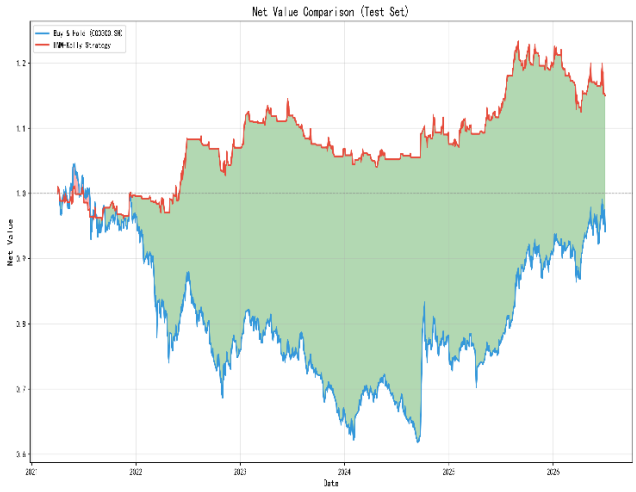
以 2008-2020 年为训练集（科创综指至 2022 年），2021 年至今为测试集，对 LSTM-GHMM 混合择时方案进行回测，如图 12-19 所示为主要指数的回测结果。在一个长回测周期中，选择的 8 个主要指数除了万得微盘均能获得相较买入持有策略的正相对收益，这说明 LSTM-GHMM 混合择时方案对于不同指数有着较好泛用性。

图表12：上证指数 2021 年以来回测结果



资料来源：Wind，中邮证券研究所

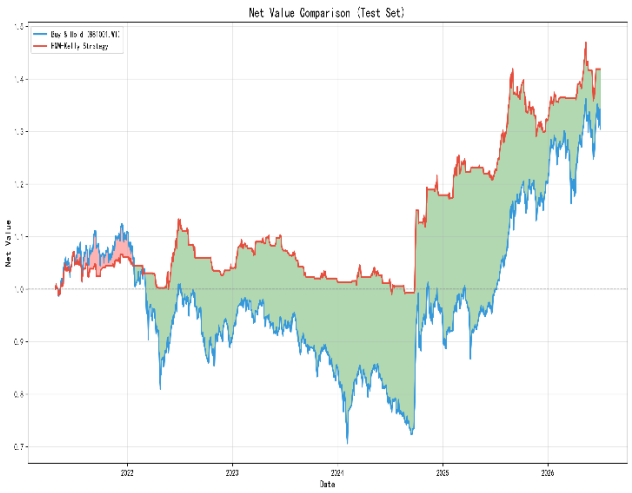
图表13：沪深 300 指数 2021 年以来回测结果



资料来源：Wind，中邮证券研究所

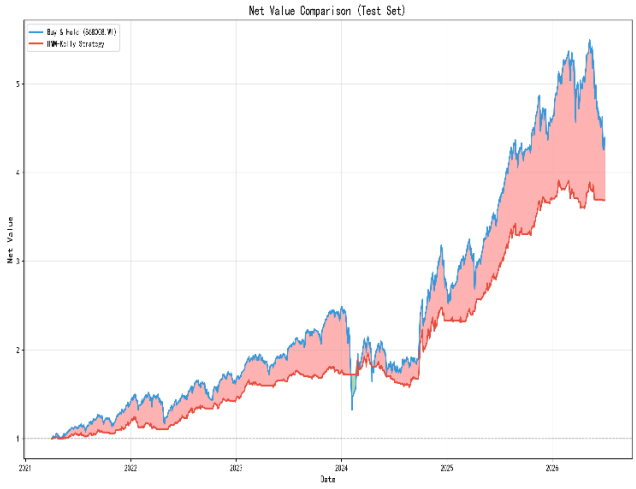
下载日志已记录，仅供内部参考

图表14： 万得全 A 指数 2021 年以来回测结果



资料来源：Wind，中邮证券研究所

图表15：万得微盘 2021 年以来回测结果



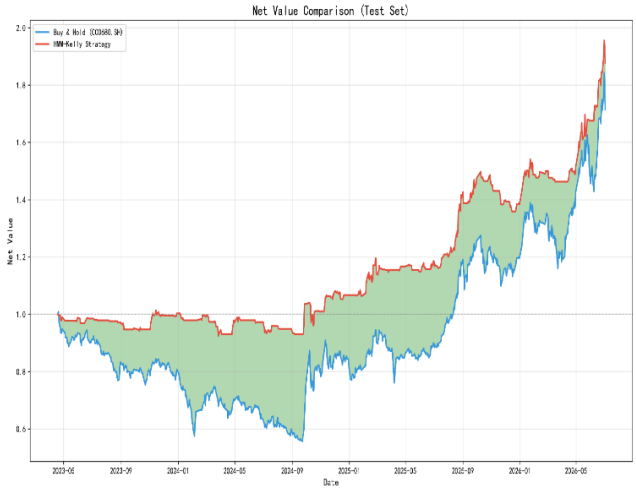
资料来源：Wind，中邮证券研究所

图表16：创业板指 2021 年以来回测结果



资料来源：Wind，中邮证券研究所

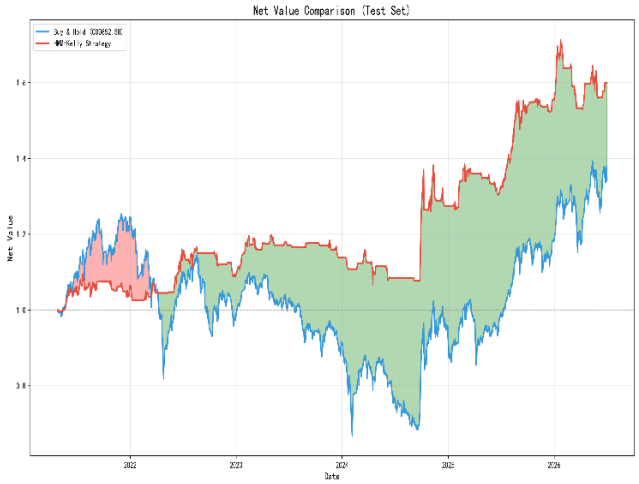
图表17：科创综指 2021 年以来回测结果



资料来源：Wind，中邮证券研究所

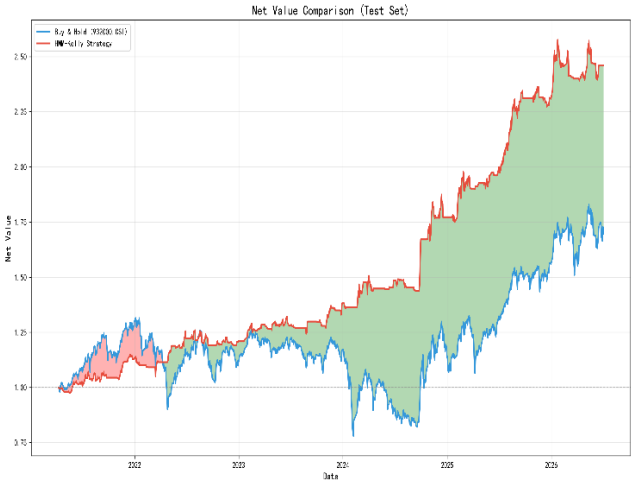
下载日志已记录，仅供内部参考

图表18：中证 1000 指数 2021 年以来回测结果



资料来源：Wind，中邮证券研究所

图表19：中证 2000 指数 2021 年以来回测结果



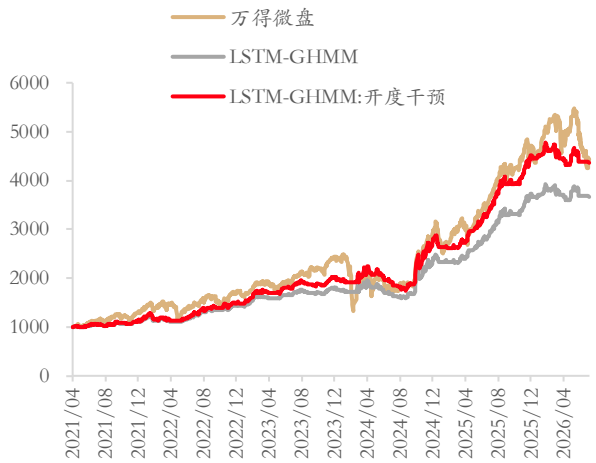
资料来源：Wind，中邮证券研究所

但在回测结果中也暴露出了两个问题：①即使 LSTM-GHMM 相较纯 HMM 方法在万得微盘的择时能力上有了很大提升，但依然在长期相较买入持有策略有着较大差距；②LSTM-GHMM 方法和纯 HMM 方法同样在处理 2021 年和 2026 年这样的单边上扬行情时会阶段性跑输导致超额收益收敛。

对于问题①，由于万得微盘特殊的编撰方式，其天然带有一个底部反转的趋势动量，这导致基本所有择时方法对于万得微盘在长时间区间都是无效的，LSTM-GHMM 能够实现对回测风险的识别和控制就已经满足设计目标了，想要获得相对超额可以通过干预模型仓位开度的方法解决。如图 20 所示，以 50%持仓以上即满仓的方式来干预，即可在大幅降低波动和回测的同时获得和买入持有基本相同的收益率。如图 21 所示为 2026 年以来的干预仓位和万得微盘走势，可见万得微盘趋势下跌中及时进行空仓控制回测。

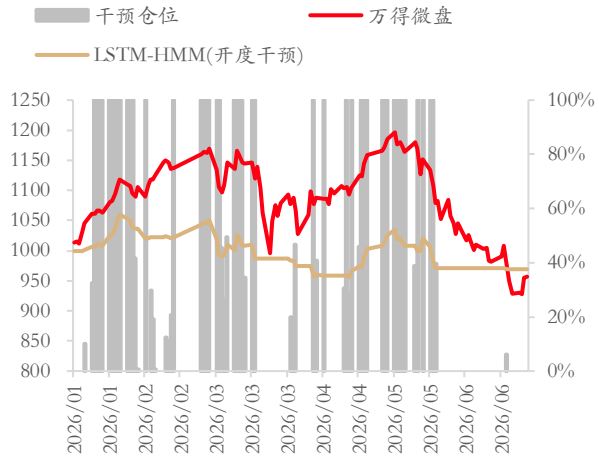
下载日志已记录，仅供内部参考

图表20：干预开度后的万得微盘指数回测结果



资料来源：Wind，中邮证券研究所

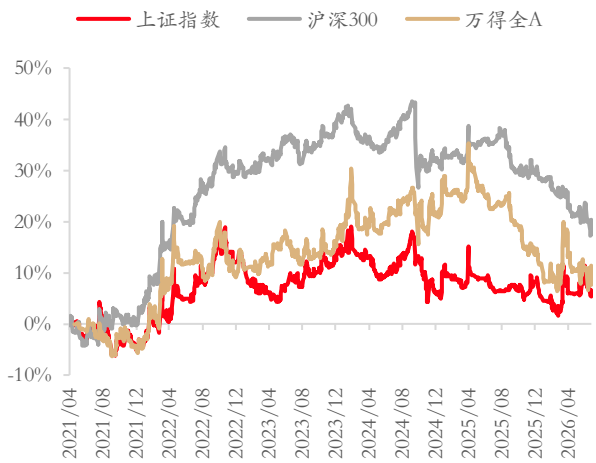
图表21：2026 年干预后仓位和回测结果



资料来源：Wind，中邮证券研究所

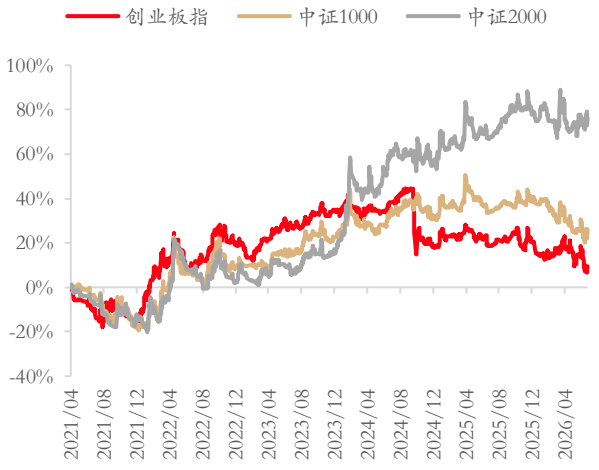
对于问题②，这并非模型参数的偶然失效，而是以 HMM 为底座的择时方案在底层设计逻辑上固有的缺陷。HMM 本质上是一个基于历史频度统计的泛用模型，其状态转移概率受训练集中“均值回归”这一强先验支配。如图 22-23 所示，2021 年与 2025 下半年出现的极端单边逼空行情，累计超额收益在这两个阶段中均出现了明显的收敛。这种行情在 A 股历史长周期中属于典型的小样本事件，其持续的正反馈动量特征与 HMM 训练集所蕴含的震荡回归统计规律存在结构性冲突。当市场进入模型从未见过的“非典型”状态时，HMM 出于对历史大概率事件的尊重，被迫选择降低仓位，从而导致阶段性超额收益收敛。

图表22：部分主要指数择时超额收益



资料来源：Wind，中邮证券研究所

图表23：2021 年以来回测结果



资料来源：Wind，中邮证券研究所

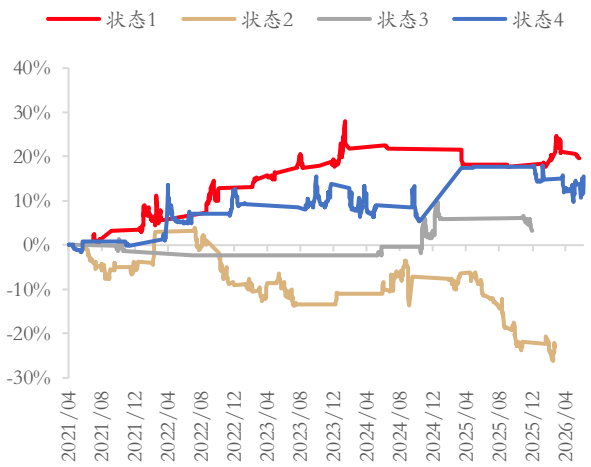
在解决思路上可以考虑 2 种方法：一是和微盘股的处理方式一样，在更宏观地识别到单边上扬行情后干预仓位开度，但这种方法不够优雅；二是考虑到如图

下载日志已记录，仅供内部参考

24 所示损失相对超额的状态较为固定，既然超额收益的损失和特定状态有强对应关系，因此可以针对特定状态外置拐点判断开关，以其他更灵敏的策略完成拐点识别进而干预反弹时的仓位开度。

这种固定状态中的超额损失对于构建择时策略来说是值得庆幸的，在 LSTM-GHMM 的框架下，只要是能被状态机正确识别被归类的类重复独立事件都可以通过更换仓位控制算法的方式进行优化。在传统的端到端预测模型中，超额损失散布于整个时间轴，难以追溯归因；而在状态机框架下，每个状态对应一类可被识别、可被回测、可被独立优化的市场微观结构。既然状态 2 的超额损失具有系统性和可重复性，就说明问题出在仓位控制算法对该状态的适应性不足，而非状态识别本身失效。比如在此展示的 LSTM-GHMM 框架在万得全 A 择时习惯在状态 2 中跑输，那么要考虑的就是使用采样生成凯利仓位的算法对于这个状态不是那么合适而已。如图 25 所示状态 2 的识别基本对应着趋势波动行情，由于确定仓位的凯利公式依赖过去 15 个同状态样本的胜率和赔率进行决策，状态 2 就天然会对应着较低的仓位选择。

图表24：万得全 A 各状态择时超额走势



资料来源：Wind，中邮证券研究所

图表25：万得全 A 状态 2 择时超额走势



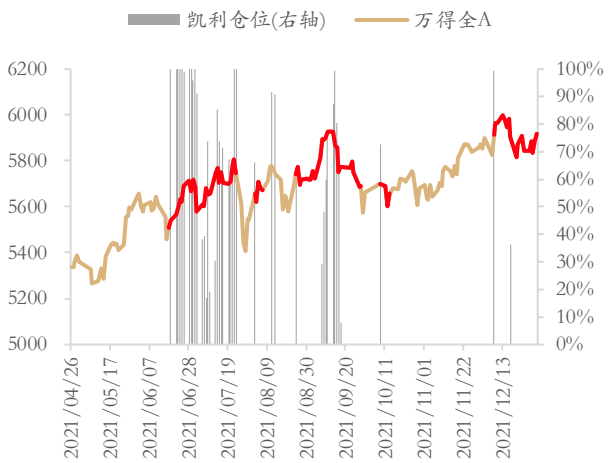
资料来源：Wind，中邮证券研究所

如图 26-27 所示为 2021 年和 2025 下半年两次 LSTM-GHMM 在状态 2 中损失相对收益时的凯利仓位，可见模型损失相对收益的主要原因是对趋势上涨行情中的回撤“过分重视”。在一个显著的回撤后，凯利公式的赔率会大幅下降，这就导致模型会降低仓位，而这样就错过了回撤后的反包。因此对于状态 2 的优化思路可以从识别“假摔反包”出发，以触发特定信号后对凯利仓位进行覆盖的方式来

下载日志已记录，仅供内部参考

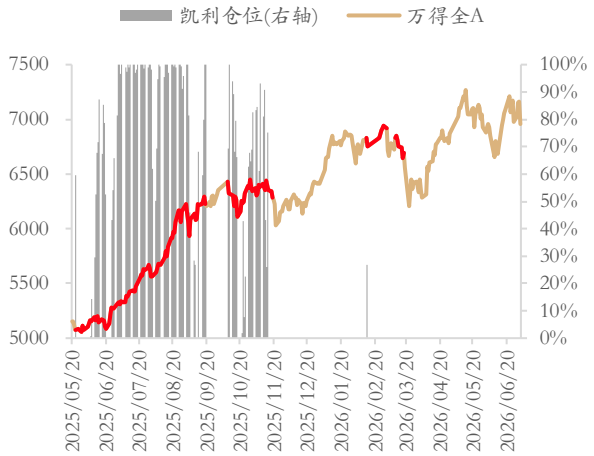
进行干预。考虑到不管在测试集还是训练集中，状态 2 都不是一个有着显著大幅回撤风险的稳定状态，因此这种进行信号干预的“冒险”是可以接受的。

图表26：2021 年万得全 A 指数凯利仓位



资料来源：Wind，中邮证券研究所

图表27：2025 下半年以来万得全 A 指数凯利仓位



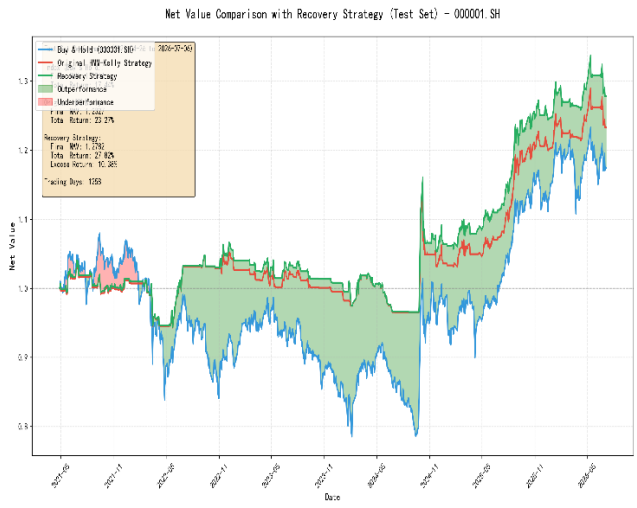
资料来源：Wind，中邮证券研究所

3.3 基于状态特征的仓位优化思路和改进方向

依照上一节所提到的思路，一个简单的优化路线也就浮现出来：如果在一个至少指向横向震荡的状态中识别到底部快速修复的形态，那么强制介入仓位管理，使仓位固定为 50%+50%凯利。如图 28-35 所示采取底部修复补丁之后的策略回测结果，红色净值曲线为原始的 LSTM-HMM+凯利仓位管理，绿色净值曲线为带底部修复的策略。可见在采取此种优化路线后主要指数的净值回测表现都能出现明显提升。

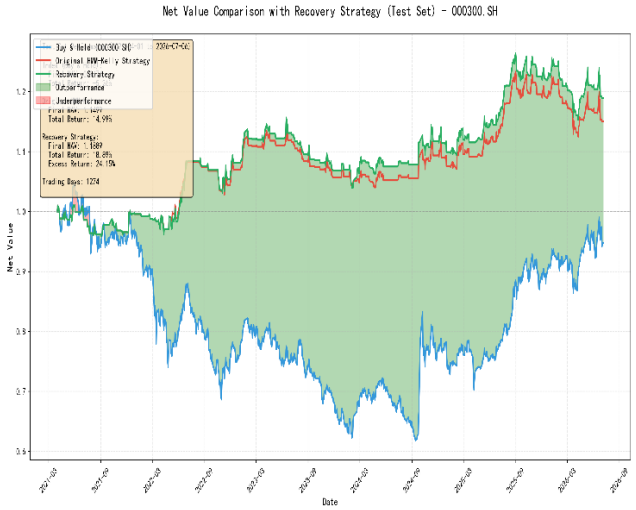
下载日志已记录，仅供内部参考

图表28：上证指数 2021 年以来补丁回测结果



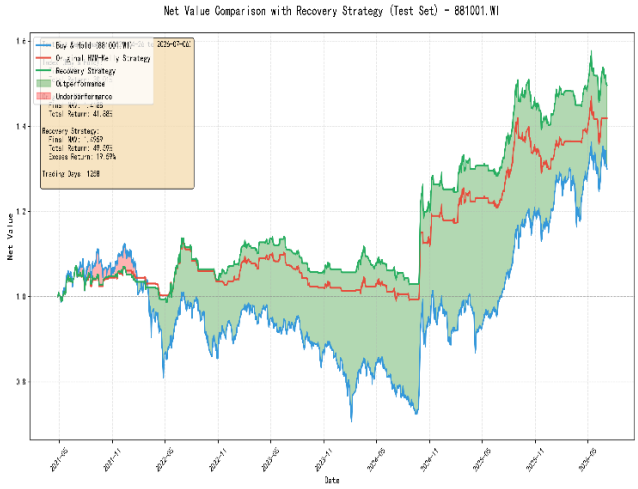
资料来源：Wind，中邮证券研究所

图表29：沪深 300 指数 2021 年以来补丁回测结果



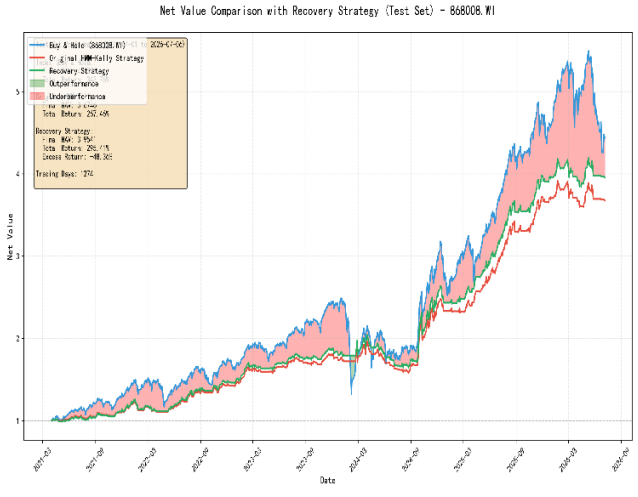
资料来源：Wind，中邮证券研究所

图表30：万得全 A 指数 2021 年以来补丁回测结果



资料来源：Wind，中邮证券研究所

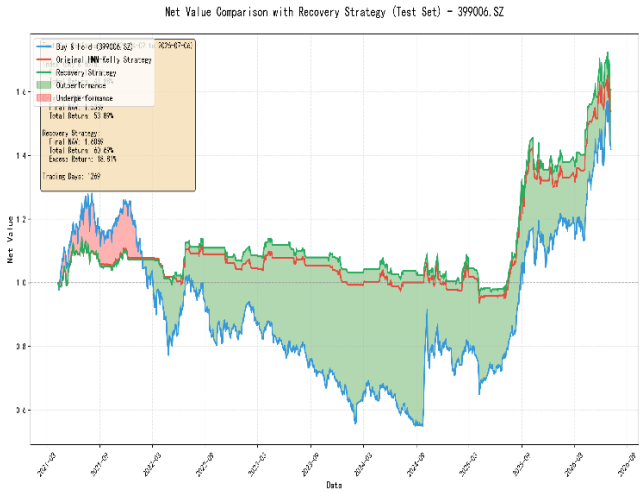
图表31：万得微盘 2021 年以来补丁回测结果



资料来源：Wind，中邮证券研究所

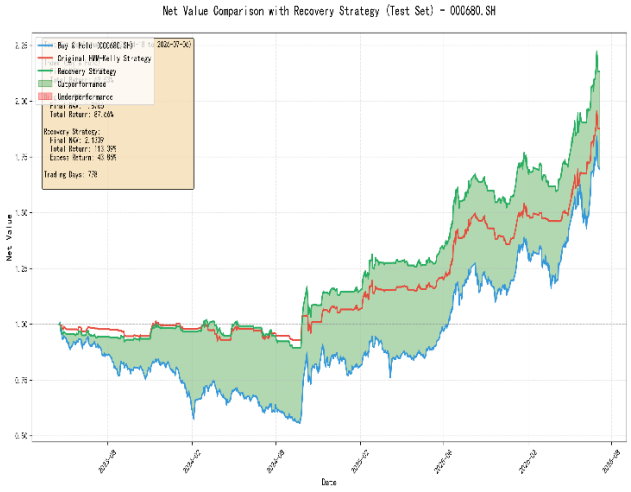
下载日志已记录，仅供内部参考

图表32：创业板指 2021 年以来补丁回测结果



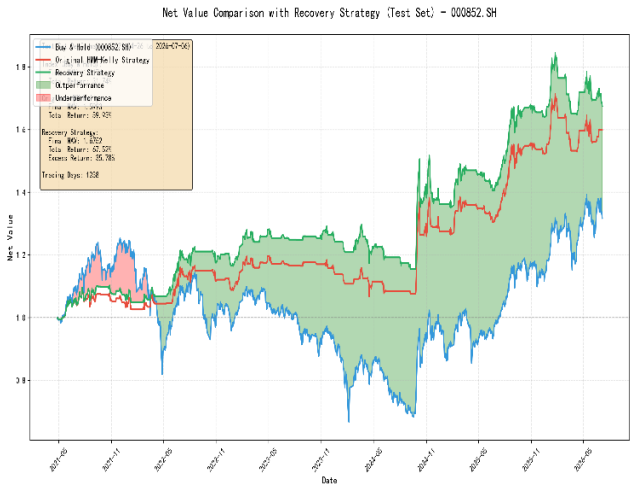
资料来源：Wind，中邮证券研究所

图表33：科创综指 2021 年以来补丁回测结果



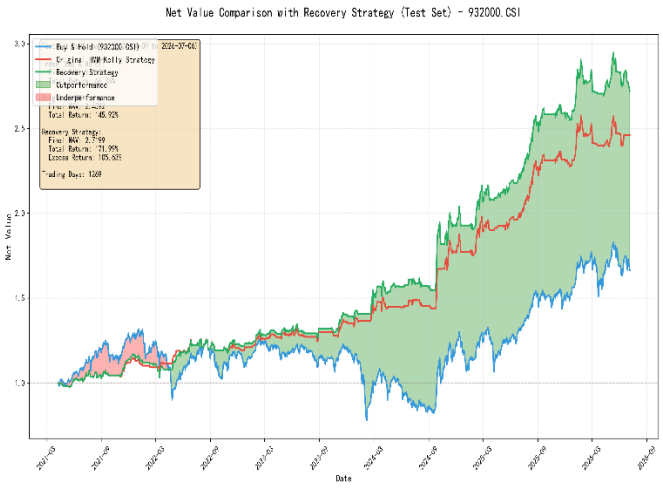
资料来源：Wind，中邮证券研究所

图表34：中证 1000 指数 2021 年以来补丁回测结果



资料来源：Wind，中邮证券研究所

图表35：中证 2000 指数 2021 年以来补丁回测结果

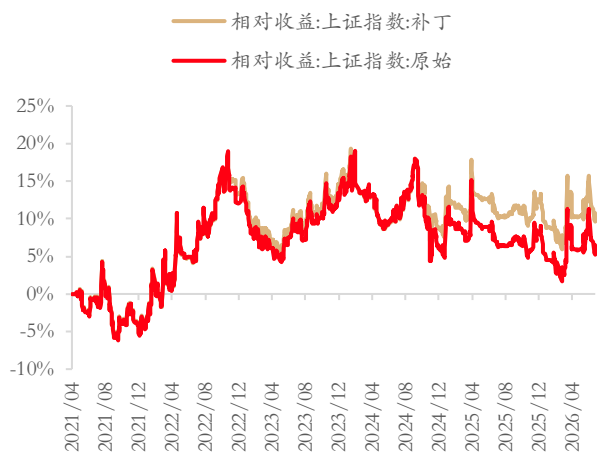


资料来源：Wind，中邮证券研究所

在使用底部修复补丁之后的相对收益改善效果具体如图 36-43 所示。虽然这个针对性的底部反转补丁在各种指数上表现出了极好的泛用性，但除了科创综指外都没有能够从根本上扭转相对超额收敛的趋势。并且也可以注意到这种底部修复补丁的做法虽然不至于在某些市场环境带来负相对收益，但也确实对具体指数和市场环境敏感。

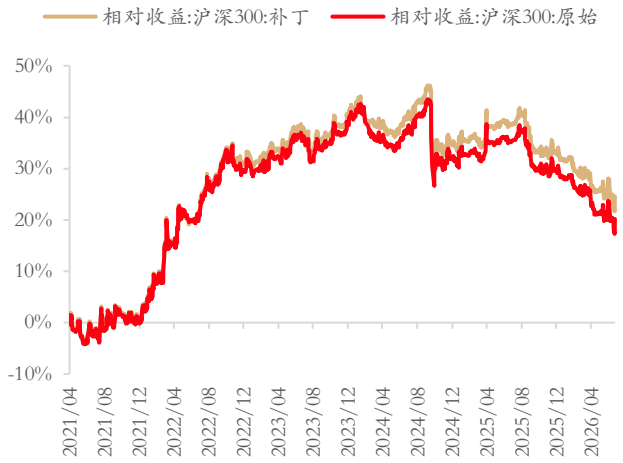
下载日志已记录，仅供内部参考

图表36：上证指数补丁及原始相对收益



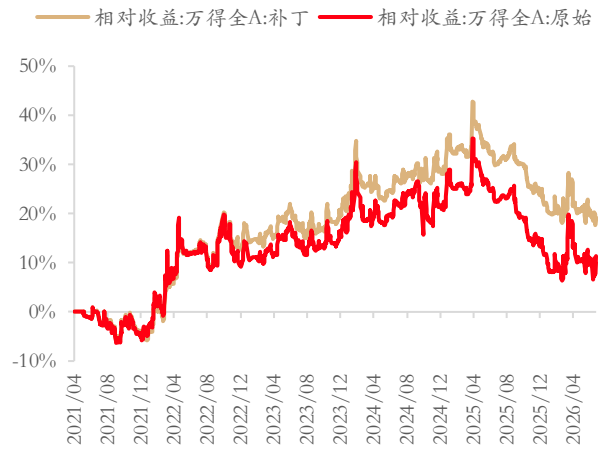
资料来源：Wind，中邮证券研究所

图表37：沪深 300 补丁及原始相对收益



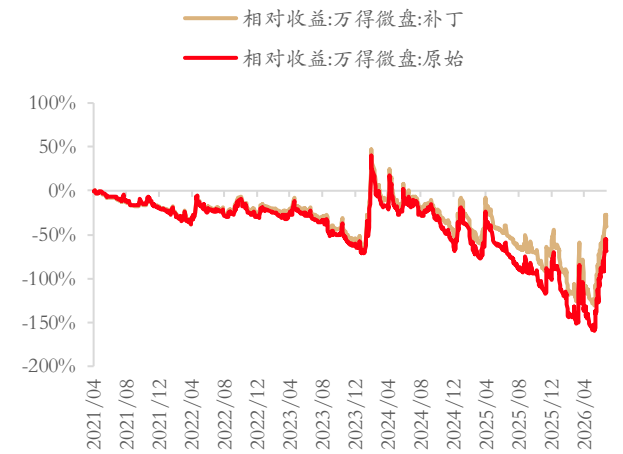
资料来源：Wind，中邮证券研究所

图表38：万得全 A 补丁及原始相对收益



资料来源：Wind，中邮证券研究所

图表39：万得微盘补丁及原始相对收益



资料来源：Wind，中邮证券研究所

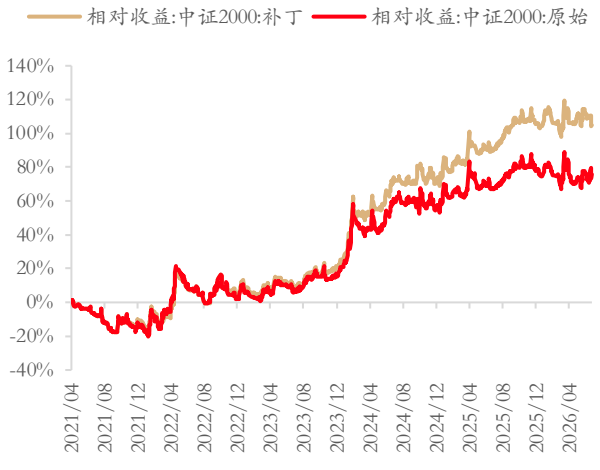
下载日志已记录，仅供内部参考

图表40：中证 1000 补丁及原始相对收益



资料来源：Wind，中邮证券研究所

图表41：中证 2000 补丁及原始相对收益



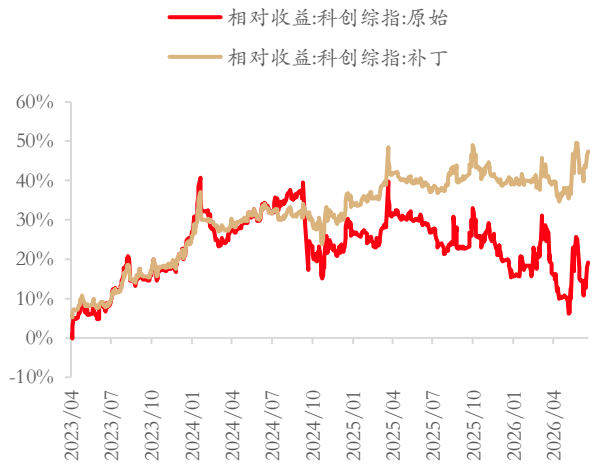
资料来源：Wind，中邮证券研究所

图表42：创业板指补丁及原始相对收益



资料来源：Wind，中邮证券研究所

图表43：科创综指补丁及原始相对收益



资料来源：Wind，中邮证券研究所

但这也提供了一个在 LSTM-GHMM 既有框架下进一步改良的思路：既然当前的框架可以精确完成市场状态识别的工作，那么就可以针对每一个状态去开发更适配的择时策略，而不是把所有仓位决定工作都交给一个带有均值回归思想的凯利公式，完全可以在带有趋势和反转特征的状态上做更多细化策略。

4 风险提示

市场学习效益，统计误差等。

下载日志已记录，仅供内部参考

中邮证券投资评级说明

投资评级标准	类型	评级	说明
报告中投资建议的评级标准：报告发布日后的6个月内的相对市场表现，即报告发布日后的6个月内的公司股价（或行业指数、可转债价格）的涨跌幅相对同期相关证券市场基准指数的涨跌幅。 市场基准指数的选取：A股市场以沪深300指数为基准；新三板市场以三板成指为基准；可转债市场以中信标普可转债指数为基准；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	预期个股相对同期基准指数涨幅在20%以上
		增持	预期个股相对同期基准指数涨幅在10%与20%之间
		中性	预期个股相对同期基准指数涨幅在-10%与10%之间
		回避	预期个股相对同期基准指数涨幅在-10%以下
	行业评级	强于大市	预期行业相对同期基准指数涨幅在10%以上
		中性	预期行业相对同期基准指数涨幅在-10%与10%之间
		弱于大市	预期行业相对同期基准指数涨幅在-10%以下
	可转债评级	推荐	预期可转债相对同期基准指数涨幅在10%以上
		谨慎推荐	预期可转债相对同期基准指数涨幅在5%与10%之间
		中性	预期可转债相对同期基准指数涨幅在-5%与5%之间
		回避	预期可转债相对同期基准指数涨幅在-5%以下

分析师声明

撰写此报告的分析师（一人或多人）承诺本机构、本人以及财产利害关系人与所评价或推荐的证券无利害关系。

本报告所采用的数据均来自我们认为可靠的目前已公开的信息，并通过独立判断并得出结论，力求独立、客观、公平，报告结论不受本公司其他部门和人员以及证券发行人、上市公司、基金公司、证券资产管理公司、特定客户等利益相关方的干涉和影响，特此声明。

免责声明

中邮证券有限责任公司（以下简称“中邮证券”）具备经中国证监会批准的开展证券投资咨询业务的资格。

本报告信息均来源于公开资料或者我们认为可靠的资料，我们力求但不保证这些信息的准确性和完整性。报告内容仅供参考，报告中的信息或所表达观点不构成所涉证券买卖的出价或询价，中邮证券不对因使用本报告的内容而导致的损失承担任何责任。客户不应以本报告取代其独立判断或仅根据本报告做出决策。

本报告所载的意见、评估及预测仅为本报告出具日的观点和判断。该等意见、评估及预测无需通知即可随时更改。过往的表现亦不应作为日后表现的预示和担保。在不同时期，中邮证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。

中邮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头寸并进行交易，也可能为这些公司提供或者计划提供投资银行、财务顾问或者其他金融产品等相关服务。

《证券期货投资者适当性管理办法》于2017年7月1日起正式实施，本报告仅供中邮证券签约客户使用，若您非中邮证券签约客户，为控制投资风险，请取消接收、订阅或使用本报告中的任何信息。本公司不会因接收人收到、阅读或关注本报告中的内容而视其为签约客户。

本报告版权归中邮证券所有，未经书面许可，任何机构或个人不得存在对本报告以任何形式进行翻版、修改、节选、复制、发布，或对本报告进行改编、汇编等侵犯知识产权的行为，亦不得存在其他有损中邮证券商业性权益的任何情形。如经中邮证券授权后引用发布，需注明出处为中邮证券研究所，且不得对本报告进行有悖原意的引用、删节或修改。

中邮证券对于本申明具有最终解释权。

下载日志已记录，仅供内部参考

公司简介

中邮证券有限责任公司于2002年9月经中国证券监督管理委员会批准设立,公司注册资本61.68亿元人民币，是中国邮政集团有限公司绝对控股的证券类金融子公司，公司是中邮创业基金管理股份有限公司的第二大股东。

公司经营范围包括:证券经纪，证券自营，证券投资咨询，证券资产管理，融资融券，证券投资基金销售，证券承销与保荐，代理销售金融产品，与证券交易、证券投资活动有关的财务顾问，具备展业的各项资格。截至2025年10月底，公司在全国设有58家分支机构(含29家分公司、29家营业部)，1家资产管理分公司和1家另类投资子公司。

中邮证券紧密依托中国邮政集团有限公司的雄厚实力，通过强化“自营+协同”发展模式，实现快速发展，当前服务的经纪客户已超过260万人。公司始终坚持诚信经营、践行金融为民，为社会大众提供全方位专业化的证券投融资服务，努力成为员工自豪、股东放心、客户信赖、社会尊重的优秀企业，打造契合中国邮政资源禀赋和市场地位的特色精品券商。

中邮证券研究所

北京
邮箱：yanjiusuo@cnpsec.com
地址：北京市丰台区北甲地路2号院6甲1号，玺萌大厦南塔
邮编：100050

上海
邮箱：yanjiusuo@cnpsec.com
地址：上海市虹口区东大名路1080号邮储银行大厦3楼
邮编：200000

深圳
邮箱：yanjiusuo@cnpsec.com
地址：深圳市福田区滨河大道9023号国通大厦二楼
邮编：518048