

# 基于 SLSTM 的协方差预测及其在分层风险平价资产配置中的应用

盛积良<sup>1</sup>, 陈兰兮<sup>1</sup>, 曾燕<sup>2</sup>, 周骐<sup>3</sup>

(1. 江西财经大学统计与数据科学学院, 南昌 330013; 2. 中山大学岭南学院, 广州 510275; 3. 华南理工大学工商管理学院, 广州 510641)

**摘要** 本文将长短期记忆 (long short term memory, LSTM) 神经网络与选择性状态空间模型 (selective state space model, SSSM) 融合, 构建了选择性长短期记忆 (selective long short term memory, SLSTM) 神经网络模型, 提出了一种新的动态协方差矩阵预测方法, 并将其用于改进分层风险平价 (hierarchical risk parity, HRP) 资产配置策略。首先, 本文构建了具有长期记忆功能和动态调整机制的 SLSTM 模型, 并设计了一种复合损失函数用于动态协方差矩阵的预测; 然后, 将预测结果应用于 HRP 以优化金融资产的配置策略; 最后, 选取金融市场主流基金产品的数据, 通过实证分析和敏感性分析验证模型的有效性和合理性。结果表明: 1) 相比于 LSTM 模型, SLSTM 模型显著提高了动态协方差矩阵预测的有效性。2) 复合损失函数增强了模型对协方差矩阵关键属性的学习能力, 从而有效提高了模型对协方差矩阵预测的准确性。3) 相比于其他资产配置策略, 基于 SLSTM 模型预测协方差矩阵的 HRP 策略能够获得较高的风险调整后收益。

**关键词** 协方差预测; 分层风险平价; 资产配置

收稿日期: 2024-10-23

基金项目: 国家自然科学基金 (71973056, 72201106, 72371256); 国家社会科学基金项目 (24AZD019); 广东省自然科学基金卓越青年团队项目 (2023B1515040001); 江西省自然科学基金重点项目 (20232ACB201006)

Supported by National Natural Science Foundation of China (71973056, 72201106, 72371256); National Social Science Foundation of China (24AZD019); Natural Science Foundation of Guangdong Province of China (2023B1515040001); Natural Science Foundation of Jiangxi Province of China (20232ACB201006)

作者简介: 盛积良, 博士, 教授, 研究方向: 金融工程与风险管理, E-mail: shengjiliang@163.com; 陈兰兮, 博士研究生, 研究方向: 金融风险管理, E-mail: chenlanxi0112@163.com; 曾燕, 博士, 教授, 研究方向: 数字经济, 数字金融, 金融工程, 风险管理, 保险精算, E-mail: zengy36@mail.sysu.edu.cn; 通信作者: 周骐, 博士, 副教授, 研究方向: 金融风险管理, 资产配置, E-mail: zhouqi@scut.edu.cn.

# SLSTM-Based Covariance Prediction and Its Application in Hierarchical Risk Parity Asset Allocation

SHENG Jiliang<sup>1</sup>, CHEN Lanxi<sup>1</sup>, ZENG Yan<sup>2</sup>, ZHOU Qi<sup>3</sup>

(1. School of Statistics and Data Science, Jiangxi University of Finance and Economics, Nanchang 330013, China; 2. Lingnan College, Sun Yat-Sen University, Guangzhou 510275, China; 3. School of Business Administration, South China University of Technology, Guangzhou 510641, China)

**Abstract** In this paper, we integrate the selective state space model (SSSM) with the long short term memory (LSTM) neural network to construct a selective long short term memory (SLSTM) neural network model. We propose a novel dynamic covariance matrix prediction method and apply it to improve the hierarchical risk parity (HRP) asset allocation strategy. Firstly, we propose an SLSTM model with long-term memory function and dynamic adjustment mechanism and design a composite loss function for dynamic covariance matrix prediction. Secondly, we apply the prediction results to the HRP to optimize financial asset allocation strategies. Finally, we select data from mainstream fund products in the financial market and conduct empirical and sensitivity analyses to verify the effectiveness and rationality of the model. The results show that: 1) The SLSTM model significantly improves the effectiveness of dynamically predicted covariance matrix compared to the LSTM model. 2) The composite loss function enhances the ability of model to learn the key attributes of the covariance matrix, thereby effectively improving the accuracy of covariance matrix prediction. 3) Compared with other asset allocation strategies, the HRP asset allocation strategy based on covariance matrix prediction by using the SLSTM model can improve risk-adjusted returns.

**Keywords** covariance prediction; hierarchical risk parity; asset allocation

## 1 引言

均值-方差 (mean-variance, MV) 模型为现代投资组合理论奠定了基础 (Markowitz, 1952), 被全球投资者广泛采用. 相比只考虑行业配置数多元化的等权重 (equal weight, EW) 模型 (周骥等, 2022), MV 模型旨在通过优化资产配置实现风险和收益的最佳平衡. 然而, MV 模型对输入参数高度敏感, 尤其是预期收益率 (Britten-Jones, 1999), 任何微小的预期收益率变化都可能导致配置结果的剧烈变动, 即配置策略不稳健. 为缓解 MV 模型对收益率输入过于敏感的问题, 学者提出了风险平价 (risk parity, RP) 模型<sup>1</sup>. RP 模型通过均衡各资产对总体风险的贡献, 能够实现更为稳健的资产配置 (赵大萍和房勇, 2020). 尽管如此, 在资产收益率高度相关的环境下, RP 模型协方差矩阵逆的数值稳定性和精确度可能会降低, 从而削弱风险分散效果. 为解决该问题, 学者们进一步提出了分层风险平价 (hierarchical risk parity, HRP) 模型, 该模型通过分层聚类的方式, 能够更好地应对协方差矩阵的不稳定性和高

<sup>1</sup> 风险平价是对投资组合中不同资产分配相同的风险权重的一种资产配置理念, 其核心思想在于使每类资产在投资组合中对总风险的贡献保持相等, 通常以资产的波动率或标准差作为风险衡量指标.

相关性资产带来的潜在问题, 进而提高资产配置的稳健性和风险管理的有效性 (Lopez, 2016; Burggraf, 2021). 总体而言, 相较于传统的 MV 模型和 EW 模型, RP 类模型通常在风险调整后收益表现上有更为显著的提升, 尤其是在夏普比率和卡玛比率等指标上表现突出. MV 模型和 RP 类模型普遍依赖历史数据计算资产间的协方差, 以预测未来的市场变化 (DeMiguel et al., 2009). 但金融市场瞬息万变, 加上外部政治、经济黑天鹅事件频发, 仅依赖历史数据的简单外推方法无法准确捕捉市场的动态特征及潜在风险. 因此, 投资者需要采用更具灵活性和动态性的协方差矩阵估计方法, 以有效应对复杂多变的 market 环境和不确定性.

现有学者提出了多种提升协方差矩阵预测准确性的方法, 大致可以从模型类型的角度分为三类. 第一类, 基于统计模型的方法, 这类方法主要依赖历史数据的统计特性, 通过模型动态估计资产间的协方差矩阵. 代表性的模型包括 DCC-GARCH 模型 (Engle, 2002), 参数化 DCC 模型 (Buccheri et al., 2021; Hafner and Wang, 2023) 等. 与此同时, LASSO-CDRD 模型通过结合高频数据和正则化技术进一步提升了协方差矩阵在高维场景下的估计精度 (刘广应等, 2022). 第二类, 基于因子模型的方法, 这类方法通过分解协方差矩阵中的风险来源, 识别出不同资产间的共同因子, 从而提高预测的稳健性. Fan et al. (2008) 提出的多因子模型有效捕捉资产间的协方差结构, 并广泛应用于投资组合配置与风险管理. 随后, Fan et al. (2011) 进一步完善该模型, 通过系统分解风险来源, 降低估计误差并增强模型的稳健性. 在最新的研究中, Shi et al. (2023) 提出了一种多层次因子模型, 结合共同因子与群组特异性因子, 精准捕捉市场与行业动态, 从而提升大规模协方差矩阵的估计效率. 第三类, 基于传统机器学习和深度学习的方法, 这类方法不依赖于传统的分布假设, 能够从复杂的金融时间序列数据中自主学习和抽取特征 (Imajo et al., 2021). 典型预测协方差矩阵的模型包括支持向量机 (support vector machine, SVM) 模型 (Fiszeder and Orzeszko, 2021), 长短期记忆 (long short term memory, LSTM) 神经网络模型 (Jain and Jain, 2019) 等.

上述三类方法体现了从传统统计模型到现代机器学习模型的演进. 传统统计模型通过捕捉资产间的动态相关性和高频数据特征, 已在金融领域得到广泛应用, 但其在处理高维和非线性结构时存在局限性. 相比之下, 因子模型在降低估计误差和提高稳健性方面有明显优势, 尤其在投资组合配置和风险管理中表现突出. 然而, 其效果依赖于对因子结构的正确识别, 且在应对快速变化的市场环境时有所不足. 随着传统机器学习和深度学习的不断发展, 第三类方法通过引入更灵活的算法框架, 展现出优于传统模型的预测能力和适应性. 因此, 本文将主要考虑第三类方法对协方差矩阵进行预测, 以更有效地捕捉数据的潜在结构与特征, 获得更精准的协方差估计.

此外, 为了更好地捕捉金融时间序列中的关键特性, 本文讨论了现有损失函数方法的局限性并提出了改进思路. 在金融时间序列预测的神经网络模型中, 均方误差经常被选为目标损失函数, 以期通过最小化预测值与实际观测值之间的差异来优化模型参数. 尽管该方法在很多情形下都十分有效 (Karevan and Suykens, 2020), 但仅使用目标损失函数可能会导致模型对目标的关键属性 (如对称性和半正定性) 学习不足. 若协方差矩阵对称性缺失会引发资产间相关性估计偏差, 削弱组合的风险分散效果; 半正定性缺失, 则会产生负特征值, 使得优化问题不可解或产生不合理结果. 因此, 单纯依赖目标损失函数可能忽视这些关键特性, 降低模型在金融数据处理中的稳健性, 特别是在风险管理和资产配置中的应用效果. 因此, 本文考

虑对损失函数结构进行优化, 以进一步提升对协方差矩阵预测的准确性.

考虑到选择性状态空间模型 (selective state space model, SSSM) 能够根据输入特征动态调整其内部状态并精确提取信息 (Gu and Dao, 2023), 本文将 SSSM 作为一种注意力机制引入 LSTM 模型, 构建一种崭新的选择性长短期记忆 (selective long short term memory, SLSTM) 神经网络模型. 该模型不仅能够增强 LSTM 输出的特征表示能力, 还能有效地捕捉长期依赖关系和非线性动态特征. 此外, 本文在目标损失的基础上引入对称损失项, 以确保预测的协方差矩阵遵循对称性原则, 并引入半正定损失项, 用于惩罚违反半正定性条件的预测矩阵. 该做法使模型能够有效学习协方差矩阵的关键属性, 并进一步提升模型对协方差矩阵预测的准确性.

为验证本文提出的协方差矩阵预测模型的有效性, 本文将预测出的协方差矩阵应用于资产配置策略, 以进行实证检验. 具体地, 本文选取包括股票、债券、商品和其他资产类别在内的 11 只交易所交易基金 (ETFs) 作为样本, 时间跨度为 2010 年 2 月 12 日至 2024 年 2 月 29 日, 运用复合损失函数的 SLSTM 模型对动态协方差矩阵进行预测, 并优化 HRP 资产配置策略. 实证研究与敏感性分析结果表明: 在 HRP 资产配置策略中, 采用 SLSTM 模型对协方差矩阵进行预测, 不仅能在市场剧烈波动时表现出较强的稳健性, 还能在市场回暖时能迅速捕捉收益机会, 显著提升了风险调整后收益. 通过该结果, 本文验证了 SLSTM 模型在动态协方差矩阵预测中的实际应用价值.

本文主要有两方面的贡献: 一是将 LSTM 模型与 SSSM 融合, 构建了一种崭新的 SLSTM 模型, 并将其应用于动态协方差矩阵预测. SLSTM 模型结合了 LSTM 的长期记忆功能和 SSSM 的动态调整机制, 能够有效捕捉资产间的长期依赖和非线性动态特征, 提高动态协方差矩阵预测的精度和模型适应市场变化的能力. 二是为进一步提高对协方差矩阵预测的准确性和模型的拟合效果, 构建了一种复合损失函数, 该函数结合了目标损失、对称损失和半正定损失. 这种多维度损失函数的设计使模型在参数优化时能更全面地考虑协方差矩阵的关键属性, 提升了模型在动态金融市场中的适用性和可靠性.

本文余下部分安排如下: 第二部分构建基于 SLSTM 的协方差矩阵预测模型, 并将该模型应用于 HRP 资产配置策略; 第三部分通过将不同的协方差矩阵预测神经网络模型与 HRP 等资产配置策略相结合, 对比优化前后的投资表现, 评估协方差矩阵预测模型的经济效益; 第四部分对 SLSTM 模型进行敏感性分析; 第五部分总结本文的结论.

## 2 模型构建

协方差矩阵的估计是资产配置和风险管理中的基本步骤. 传统的历史数据估计方法通过计算各资产在不同时刻收益率与其样本平均收益率之差的乘积, 并对所有时刻求和取均值得到资产之间的协方差, 进而构造出协方差矩阵, 具体为:

$$\text{Cov}(R_i, R_j) = \frac{1}{T-1} \sum_{t=1}^T (R_i(t) - \mu_i)(R_j(t) - \mu_j), \quad (1)$$

其中  $\text{Cov}(R_i, R_j)$  表示第  $i$  个和第  $j$  个资产之间的协方差,  $R_i(t)$  表示第  $t$  期第  $i$  个资产的收益率,  $T$  为样本期数,  $\mu_i$  表示第  $i$  个资产的平均收益率. 由此得到的协方差矩阵  $\Sigma$  的第  $(i, j)$

元素即为  $Cov(R_i, R_j)$ . 然而, 历史数据估计法计算协方差矩阵主要依赖过去的市场数据, 在市场结构快速变化时常显不足, 难以捕捉数据的非线性和复杂动态特征. 为解决上述问题, 下文将首先构建一种崭新的 SLSTM 模型; 然后, 详细探讨如何利用 SLSTM 模型来预测协方差矩阵; 最后, 本文将 SLSTM 模型预测的协方差矩阵应用于 HRP 策略, 构建基于 SLSTM 预测协方差矩阵的分层风险平价模型.

### 2.1 基于 SLSTM 的协方差矩阵预测模型

为了有效捕捉市场的时间序列特性及其复杂性, 本文构建了融合 LSTM 模型与 SSSM 的 SLSTM 模型. 该模型利用 LSTM 独特的门控机制, 有效捕捉时间序列数据中的长期依赖关系; 通过 SSSM 动态调整状态转移矩阵和输入矩阵, 使模型能够根据输入数据的不同特征, 更好地捕捉时间序列中的非线性和非平稳特性. 本文将 SSSM 的动态调整功能作为一种注意力机制, 增强了模型处理复杂时间序列数据的能力. SLSTM 模型具体框架结构如图 1 所示.

SLSTM 模型的构建过程如下:

首先, 将数据输入至 SLSTM 模型的第一部分即 LSTM 网络. 该网络由两层 LSTM 组成, LSTM 结构如图 1(a) (崔炎炎和刘立新, 2022). 该部分的主要功能是从输入数据中提取时间特征, 捕捉数据的短期和长期依赖关系. LSTM 通过其门控机制有效地处理了时间序列数据的长短期依赖性问题, 使得模型能够学习复杂的时间序列模式. 每个 LSTM 网络包含三个主要部分: 遗忘门 ( $f_t$ )、输入门 ( $i_t$ )、输出门 ( $o_t$ ) 以及一个单元状态 ( $c_t$ ). LSTM 层将处理整个序列, 并返回序列的完整输出并供下一层使用, 其更新公式为:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f),$$

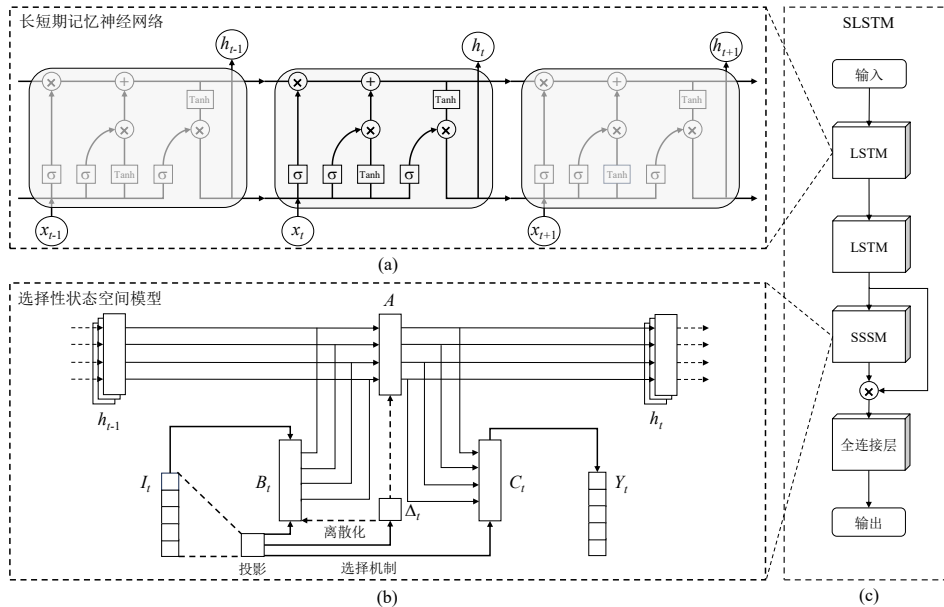


图 1 SLSTM 网络结构模型框架图

$$\begin{aligned}
i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\
o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \\
\tilde{c}_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \\
c_t &= f_t * c_{t-1} + i_t * \tilde{c}_t, \\
h_t &= o_t * \tanh(c_t),
\end{aligned} \tag{2}$$

其中  $x_t$  为  $t$  时刻的输入;  $h_t$  为  $t$  时刻的隐藏状态, 它是当前单元状态和输出门输出的组合;  $W_k$  表示各个单元的输入权重矩阵,  $b_k$  为偏置矩阵,  $k = f, i, o, c$ ;  $\tilde{c}_t$  为  $t$  时刻候选单元状态, 提供一个更新当前单元状态的可能选项;  $\tanh(x)$  和  $\sigma(x)$  为激活函数<sup>2</sup>.

在第一层中, LSTM 主要接收处理后的时间序列数据, 通过输入门、遗忘门和输出门动态管理记忆状态. 在每个时间步, LSTM 结合当前输入与前一时间步的隐藏状态, 通过非线性变换筛选需保留的信息, 从而捕捉时间序列中的长期依赖性. 在第二层中, LSTM 进一步处理第一层的输出特征, 以优化特征表示, 从而增强模型提取全局动态特征的能力. 与第一层 LSTM 捕捉输入数据的直接时序关系不同, 第二层 LSTM 通过处理第一层的输出特征, 能够学习更加抽象的时序模式, 为后续的 SSSM 提供更高层次的时序信息, 输出特征记为  $I$ .

然后, LSTM 网络的输出特征  $I$  被传递到 SLSTM 模型的第二部分即 SSSM 中, SSSM 结构如图 1(b) 所示 (Gu and Dao, 2023). SSSM 通过一个可学习的隐藏状态  $h(t)$  将输入  $I(t)$  转换成输出  $y(t)$ , 其参数为  $A$ ,  $B$  和  $C$ , 具体如下:

$$h'(t) = Ah(t) + BI(t), \tag{3}$$

$$y(t) = Ch(t), \tag{4}$$

其中  $h'(t)$  和  $y(t)$  分别表示连续时间模型中的状态和输出. 为提高计算效率, 现将连续参数  $A, B, C$  转换为离散参数. 具体而言, 假设时间间隔为  $\Delta$ , 应用零阶保持法<sup>3</sup>以获得离散参数  $\bar{A}$ ,  $\bar{B}$  和  $\bar{C}$ , 具体如下:

$$\begin{aligned}
\bar{A} &= \exp(\Delta A), \\
\bar{B} &= (\Delta A)^{-1}(\exp(\Delta A) - E) \cdot (\Delta B), \\
\bar{C} &= C,
\end{aligned} \tag{5}$$

其中  $E$  是单位矩阵. 通过离散化处理, 状态更新被转化为递归关系, 最终用离散参数表示式 (3) 和式 (4):

$$h(t) = \bar{A}h(t-1) + \bar{B}I(t), \tag{6}$$

$$y(t) = \bar{C}h(t), \tag{7}$$

<sup>2</sup>激活函数  $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ ,  $\sigma(x) = \text{Sigmoid}(x) = \frac{1}{1 + e^{-x}}$ .

<sup>3</sup>零阶保持法是一种信号采样方法, 即在每个采样周期内保持最后一次采样值不变, 以此来实现由连续系统到离散系统的近似转换.

其中  $l$  为离散化后采用的时间步长 (即状态更新时的延迟长度). SSSM 通过状态方程和观测方程进一步处理和提取特征, 并通过有选择地将数据压缩到状态中, 从而作为一种注意力机制输入  $I$  中每个特征的重要性, 对  $I$  中有效度高的特征和有效度低的特征分别赋予高权重和低权重, 再将 SSSM 输出的特征权重与  $I$  相乘得到输出  $Y$ , 如图 1(c) 所示. 该过程如式 (8) 所示:

$$Y = I \times \text{SSSM}(I), \quad (8)$$

其中  $\text{SSSM}(I)$  表示 SSSM 提取输入  $I$  的特征. SSSM 不仅提高了模型的灵活性和适应性, 还增强了模型处理高维数据和捕捉潜在状态动态变化的能力.

最后, 将  $Y$  输入到全连接层 (FC), 使 SSSM 提取到的特征得到进一步处理, 即将当前输出映射为最终生成的预测协方差矩阵结果  $\hat{\Sigma}$ . 该过程如式 (9) 所示:

$$\hat{\Sigma} = \text{FC}(Y). \quad (9)$$

为进一步优化模型的训练效果并提高预测精度, 本文对损失函数结构进行改进以有效引导模型学习动态市场条件下的复杂关联结构, 并增强 SLSTM 模型对协方差矩阵预测的准确性. 考虑到协方差矩阵具有对称性和半正定性, 本文提出在目标损失的基础上引入对称损失和半正定损失, 使神经网络模型在捕捉资产收益的长期依赖和非线性动态时, 进一步增强模型对协方差矩阵对称性和半正定性的学习能力. 损失函数的优化模型如下:

$$\begin{aligned} \min_{\alpha, \beta} L &= L_{\text{Target}} + \alpha L_{\text{Symmetric}} + \beta L_{\text{PSD}}, \\ &= \frac{1}{T} \sum_{t=1}^T \left( \|\Sigma_t - \hat{\Sigma}_t\|_F^2 + \alpha \|\hat{\Sigma}_t - \hat{\Sigma}_t'\|_F^2 + \beta \|\hat{\Sigma}_t - \hat{\Sigma}_t \cdot \hat{\Sigma}_t'\|_F^2 \right), \\ \text{s.t. } \Lambda_i &= \max(\text{RE}(\Lambda_i), 0), \end{aligned} \quad (10)$$

其中  $L$  表示模型的损失函数, 它包括协方差矩阵的均方误差、对称损失和半正定损失;  $L_{\text{Target}}$  为目标损失, 衡量预测协方差矩阵与实际协方差矩阵之间的差异, 用于最小化预测误差;  $L_{\text{Symmetric}}$  为对称损失, 用以确保矩阵的对称性;  $L_{\text{PSD}}$  为半正定损失, 用于确保矩阵半正定, 即所有特征值非负, 从而保证预测矩阵的数学合理性;  $\Sigma_t$  是在时间  $t$  的真实协方差矩阵;  $\hat{\Sigma}_t$  是模型在时间  $t$  的预测协方差矩阵;  $\hat{\Sigma}_t'$  表示  $\hat{\Sigma}_t$  的转置;  $T$  是总时间长度;  $\|\cdot\|_F$  表示 Frobenius 范数;  $\alpha$ 、 $\beta$  是权衡不同损失项重要性的调节参数; 分解预测矩阵  $\hat{\Sigma}$  得到的特征值及特征向量分别表示为  $\Lambda_i, V_i, i \in [1, n]$ . 因非实对称矩阵特征值分解时可能出现复数特征值, 为简化分析, 本文在特征值约束部分只考虑特征值的实数部分是否大于 0, 即  $\Lambda_i = \max(\text{RE}(\Lambda_i), 0)$ , 其中  $\text{RE}$  表示取实数部分.

优化问题 (10) 的目标是通过调整模型参数, 实现最小化总损失函数  $L$ . 常见的优化方法之一是 Adam (adaptive moment estimation) 算法, 该算法结合了动量和自适应学习率, 能够有效处理稀疏梯度和非平稳目标 (杨青和王晨蔚, 2019). 因此, 本文使用 Adam 算法对神经网络模型进行反馈, 并通过迭代更新优化模型参数. Adam 算法具体步骤如下: 首先, 计算当前梯度的动量估计, 即梯度的加权平均; 其次, 计算梯度的平方的动量估计; 然后, 对这两个

估计进行偏差校正,使其在初始阶段更加准确;最后,更新模型参数,使用动量估计的均值和均方根值来调整每个参数的学习率,从而实现自适应学习,提高训练的效率和稳定性.通过上述优化过程,SLSTM模型能够更有效地处理和预测金融时间序列数据中的协方差矩阵,为HRP策略提供稳定可靠的数据支持.

## 2.2 基于 SLSTM 的分层风险平价模型

本文基于 SLSTM 模型对动态协方差矩阵进行预测,并将预测矩阵应用于 HRP 模型,该框架结构如图 2 所示.该结构包括数据处理、协方差矩阵预测、分层聚类、准对角化和递归二分五个部分.其中分层聚类、准对角化和递归二分三个部分为 HRP 资产配置过程 (Raffinot, 2018).

在数据预处理阶段,首先对原始数据的收益率序列进行检查,以识别数据集中缺失值的位置<sup>4</sup>.对于检测到的缺失值,采用前后一个交易日数据的均值作为填充值,确保数据的完整性和准确性;随后对数据进行标准化处理,以保证输入数据的质量和适应性;最后将数据集划分为训练集,验证集和测试集,其中训练集和验证集为样本内数据,测试集为样本外数据.该划分方法旨在确保后续模型训练和测试的有效性和合理性,进而提升模型的预测性能和稳定性.

数据经过预处理后,作为协方差矩阵预测阶段的输入.在该阶段中,首先采用 SLSTM 模型对样本内数据进行训练.训练完成后,输出预测矩阵并计算对称损失与半正定损失,同时对预测矩阵进行对称化与半正定化,以保证最终得到的目标矩阵  $\hat{\Sigma}$  具有协方差矩阵的数学属性.其中,对称化与半正定化的实现依赖于分解预测矩阵得到矩阵的特征值  $\Lambda_i, i \in [1, n]$  及其对应的特征向量  $V_i, i \in [1, n]$ ,将负特征值置 0 对预测矩阵进行重构,重构过程如式 (11) 所示:

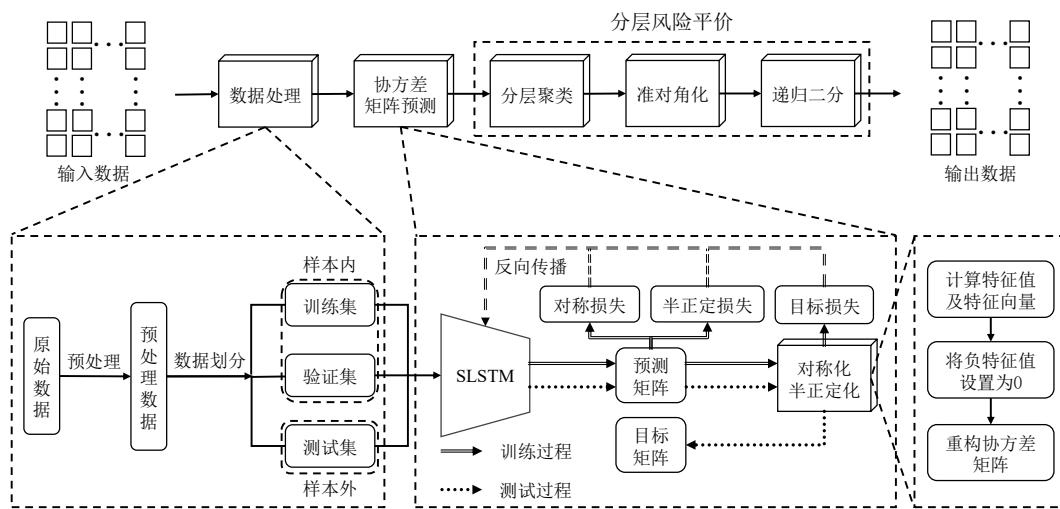


图 2 基于 SLSTM 的分层风险平价模型框架结构图

<sup>4</sup>由于停牌、数据传输延迟或交易日历不一致等原因,故部分交易日的收益率数据可能出现缺失.

$$\hat{\Sigma} = V\Lambda V', \quad (11)$$

其中  $V$  和  $V'$  分别表示特征向量构成的矩阵及其转置矩阵;  $\Lambda$  表示由特征值构成的对角矩阵;  $\hat{\Sigma}$  表示目标矩阵。

根据得到的目标矩阵计算出目标损失, 进而构建一个复合损失函数。该复合损失函数结合了目标损失、对称损失和半正定损失, 有效提高了神经网络的拟合效果和泛化性。通过该复合损失函数进行反向传播, 神经网络在训练过程中通过梯度下降法降低损失, 进而得到有效的神经网络预训练模型。在模型训练完成后, 将样本外数据输入至预训练模型中, 得到输出的动态协方差矩阵可直接应用于 HRP 资产配置策略。

这种通过利用神经网络预测动态协方差矩阵, 从而优化 HRP 资产配置策略的方法, 不仅能够提高资产配置的精度和效率, 还能强化投资组合在面对市场波动时的稳健性。该方法可以使得投资决策不再单纯依赖历史数据, 而是能够通过预测的信息动态反映市场的最新变化, 从而为投资者带来更可靠的信息和更稳定的回报。

### 3 实证研究

ETFs 作为现代金融市场的重要创新之一, 是机构投资者进行实时指数化投资的主要工具 (姜富伟等, 2022)。本研究根据大类资产分类标准, 选择股票、债券、商品及其他资产类别的 ETFs 作为研究样本, 数据时间跨度为 2010 年 2 月 12 日至 2024 年 2 月 29 日, 数据来源为 WIND 金融研究数据库<sup>5</sup>。样本时间跨度包含诸如 2011 年美国信用评级下调、2015 年中国股市暴跌、2020 年新冠疫情暴发等多个影响全球资本市场的重大事件, 这些事件对各类资产的收益率波动产生了显著影响。因此, 选取该时间跨度能够突出考察不同资产类别在高波动环境下的收益稳定性与风险管理效果, 为评估模型的鲁棒性与适应性提供实证依据。选择的投资标的的具体如表 1 所示。

表 1 资产类别描述

ETFs 类别	ETFs 代码	描述
股票类 ETFs	SPY	美国股市
	EEM	新兴市场股市
	VGK	欧洲市场股市
	TLT	美国国债 (但通常与股票一起分析, 因其在资产配置中的重要性)
债券类 ETFs	BND	美国总债券市场 (包括政府债券和公司债券)
	LQD	美国高信用等级公司债
	HYG	信用评级较低的公司债, 即高风险债券市场
商品类 ETFs	GLD	黄金价格
	USO	原油价格
其他资产类别	VNQ	房地产市场
	ACWI	全球股市

<sup>5</sup>本文所使用的数据和代码请参见科学数据银行 (ScienceDB) 期刊社区, DOI: 10.57760/sciencedb.21254 和 CSTR: 31253.11.sciencedb.21254。若使用文中数据信息, 请注明引文和数据出处。

上述 ETFs 能够全面覆盖不同的市场区域, 如美国、欧洲和新兴市场, 以及不同的资产类别, 如股票和债券, 从而在分析中更好地反映全球资产配置的多样性和复杂性. 由于金融时间序列数据具有时效性, 本文采取滑动窗口法划分数据集, 并进行滚动投资. 为尽可能地保证模型的训练效果及其泛化能力, 时间窗口长度设置为 30 个交易日. 在训练阶段, 将样本内数据按 8:2 的比例进行划分, 对神经网络模型进行训练与验证. 其中 2010 年 2 月至 2019 年 1 月期间数据为训练集, 2019 年 2 月至 2022 年 12 月期间数据为验证集. 在测试阶段, 将整个样本外数据作为测试集, 其时间段为 2023 年 1 月到 2024 年 2 月, 共计 14 个月. 该阶段旨在评估模型在样本外环境下的预测能力, 从而验证其有效性 (Drobtz et al., 2025).

### 3.1 绩效评价指标及实验参数设置

#### 1) 绩效评价指标

为检验神经网络模型预测的有效性以及评估不同投资组合策略的样本外表现, 本文采用均方误差和均方根误差等误差指标, 以及收益、风险和风险调整后收益等多维度指标进行分析, 具体如下表 2 所示.

表 2 绩效评价指标

维度指标	指标名称	描述	计算公式
误差指标	均方误差	预测值与真实值之间误差平方的平均值	均方误差 = $\frac{1}{\text{样本数量}} \sum_{i=1}^{\text{样本数量}} (\text{真实值}_i - \text{预测值}_i)^2$
	均方根误差	均方误差的平方根	均方根误差 = $\sqrt{\frac{1}{\text{样本数量}} \sum_{i=1}^{\text{样本数量}} (\text{真实值}_i - \text{预测值}_i)^2}$
	平均绝对误差	预测值与真实值之间绝对误差的平均值	平均绝对误差 = $\frac{1}{\text{样本数量}} \sum_{i=1}^{\text{样本数量}}  \text{真实值}_i - \text{预测值}_i $
收益指标	累计收益率	显示整个分析周期内资产的总增值百分比	累计收益率 = $\prod_{t=1}^{\text{总期数}} (1 + \text{第 } t \text{ 期收益率}) - 1$
	年化收益率	将累计收益率年化, 比较不同投资期间的收益表现	年化收益率 = $(1 + \text{累计收益率})^{\frac{365}{\text{天数}}} - 1$
风险指标	年化波动率	度量收益率的标准差, 反映投资的波动性	年化波动率 = 标准差 $\times \sqrt{252}$
	最大回撤率	从峰值到谷底的最大损失百分比	最大回撤率 = $(\text{峰值} - \text{谷底}) / \text{峰值} \times 100\%$
风险调整后收益指标	夏普比率	衡量每承担一单位总风险所获得的超额回报	夏普比率 = $(\text{年化收益率} - \text{无风险利率}) / \text{年化波动率}$
	确定等价回报	根据投资者的风险偏好调整的回回报率	确定等价回报 = $\text{年化收益率} - (\text{投资者风险规避值} / 2) \times \text{预期收益率的方差}$
	卡玛比率	最大回撤率与年化收益率比值	卡玛比率 = $\text{最大回撤率} / \text{年化收益率}$

#### 2) 实验参数设置

基于投资者在实际市场中普遍表现出的风险态度, 本文将投资者的风险规避值设置为 3, 无风险利率设置为 3.5%. 本文的 SLSTM 模型使用了两层 LSTM, 其网络参数如表 3 所示. 每层 LSTM 包含 64 个隐藏单元, 该配置可以有效捕捉时间序列数据中的长期依赖关系. 同

时, 引入的 SSSM 同样设定状态维度为 64, 以增强模型对金融市场动态的捕捉能力. 为了高效地训练神经网络, 本文采用批量大小为 64 的批处理策略, 该策略有助于平衡训练速度与内存使用效率. 此外, 本文针对 SLSTM 模型的损失函数结构进行网格搜索, 将参数  $\alpha$ 、 $\beta$  的取值范围设置为  $\{0.5, 1, 1.5\}$ , 结果表明  $\alpha$ 、 $\beta$  取值均为 1 时效果最佳. 选择 Adam 算法作为优化算法是因其具有自适应学习率优势, 适合处理大规模数据训练. 初始学习率设定为 0.001, 可以在训练初期快速收敛, 随后通过内部机制逐渐调整学习率, 以细致优化模型参数. 模型中激活函数选用 ReLU<sup>6</sup>, 因其具有非线性特性和计算简便性, 有助于增强网络的表达能力并防止出现梯度消失问题. 同时, 为防止过拟合, 本文在网络的两层 LSTM 中均引入了 L2 正则化, 促使网络学习到更加有效的特征.

表 3 SLSTM 网络参数表

网络层	输入大小	层数	隐藏层单元数	状态维度	输出大小
LSTM	(64, 11, 30)	1	64	-	(64, 11, 64)
LSTM	(64, 11, 64)	1	64	-	(64, 11, 64)
SSSM	(64, 11, 64)	1	-	64	(64, 11, 64)
FC	(64, 11 × 64)	1	-	-	(64, 11 × 11)

### 3.2 数据描述性分析

通过对 2010 年 2 月 12 日至 2024 年 2 月 29 日期间 11 种 ETFs 的日收益率进行分析, 本文评估了各 ETFs 的收益表现及其风险特征, 为优化投资决策提供了依据, 具体结果见表 4.

由表 4 可知, 不同资产类别的日均收益率表现存在明显的差异. 股票类 ETFs (SPY, EEM, VIG, ACWI) 中, SPY 日收益率高达 0.043%, 反映了其显著的投资回报表现. 债券类资产如 BND 和 LQD 显示出了较低的低波动率, 表明其适合作为投资组合中的低风险资产. 在商品类资产中, USO 日收益率最大值为 213.45%, 主要源于该 ETF 采用期货合约滚动机制, 当合约切换日遇到较大价差时可能出现极端收益率表现, 而 GLD 表现出较为稳健的收益, 波动性相对较小.

### 3.3 样本外表现

本文以 EW 资产配置策略为基础进行模型对比, 并将反向传播 (back propagation, BP) 神经网络、循环神经网络 (RNN)、LSTM 和 SLSTM 四种神经网络模型<sup>7</sup>用于优化全局最小方差 (GMV)<sup>8</sup>、MV、RP 和 HRP 四种资产配置策略, 以探索这些模型在金融时间序列分析中的应用效果. 五种资产配置策略见附录.

<sup>6</sup>ReLU 是一种常用的神经网络激活函数, 其数学表达式为  $f(x) = \max(0, x)$ .

<sup>7</sup>BP、RNN、LSTM 和 SLSTM 模型的主要区别在于其时间依赖处理能力. 以 BP 为基础的前馈神经网络, 无法处理时间序列数据; RNN 通过反馈机制处理时间序列, 但容易面临梯度消失问题; LSTM 引入记忆门机制, 改善了对长时间依赖的捕捉能力; 而 SLSTM 则在 LSTM 基础上结合了 SSSM, 进一步提升了对复杂动态协方差矩阵的预测能力. 本文选择这四个模型是为了对比传统神经网络模型和改进模型在金融时间序列中的表现差异.

<sup>8</sup>GMV 策略是一种投资组合优化方法, 旨在通过最小化组合的整体波动率来实现风险的最优分散.

表 4 描述性统计

资产类别	资产	日收益率			
		均值	最小值	最大值	标准差
股票	SPY	0.043%	-11.589%	8.673%	0.012
	EEM	-0.005%	-13.329%	8.841%	0.015
	VGK	0.002%	-3.375%	2.371%	0.004
	TLT	0.002%	-6.901%	7.250%	0.010
债券	BND	-0.004%	-5.592%	4.134%	0.003
	LQD	-0.002%	-5.133%	7.131%	0.005
	HYG	-0.006%	-5.653%	6.341%	0.006
商品	GLD	0.013%	-9.191%	6.957%	0.010
	USO	0.032%	-29.189%	213.425%	0.046
其他	VNQ	0.011%	-19.514%	8.713%	0.014
	ACWI	0.026%	-11.896%	7.822%	0.012

### 3.3.1 损失可视化

为评估四种神经网络模型在协方差矩阵预测中的效果, 现对训练损失 (Train\_Loss) 和验证损失 (Val\_Loss) 进行可视化分析, 如图 3 所示。

由图 3 可知, 每个模型都随着训练周期 (Epoch) 的增加, 损失逐渐下降, 其中相比于 BP、RNN 和 LSTM 模型, SLSTM 模型表现出更低的损失和更快的收敛, 表明其在协方差矩阵预测中的优越性。具体而言, SLSTM 模型通过结合长短期记忆单元和选择机制, 能够更有效地捕捉时间序列数据中的复杂非线性关系和长期依赖性。这种改进不仅能够提高预测精度、减少训练和验证损失, 还能加快模型的收敛速度, 从而显著提升计算效率和模型稳定性。将四种神经网络模型预测出的协方差矩阵作为预测矩阵, 经过对称化和半正定化处理后的矩阵作为目标矩阵。四种神经网络模型在预测矩阵和目标矩阵下误差的描述性统计情况如表 5 所示。

从表 5 可得, 四种神经网络模型在协方差矩阵预测中的效果存在显著差异。具体而言, RNN 模型的误差值相对较高, 这可能是由于其在处理长时间序列数据时容易出现梯度消失问题所致。而 LSTM 模型通过引入记忆单元, 有效解决了 RNN 的梯度消失问题, 其误差显著降低, 表现出较好的预测效果。SLSTM 模型在 LSTM 的基础上进行了改进, 进一步增强了模型表达能力, 其误差值最小。对目标矩阵进行对称化与半正定化处理后, SLSTM 模型的误差指标均有显著改善, 这一结果表明了该处理方法的合理性和必要性。

表 5 协方差矩阵预测误差

	均方误差		均方根误差		平均绝对误差	
	预测矩阵	目标矩阵	预测矩阵	目标矩阵	预测矩阵	目标矩阵
BP	0.00202	0.00029	0.04410	0.01510	0.03440	0.01190
RNN	0.00950	0.00060	0.09750	0.02380	0.04580	0.01970
LSTM	0.00044	0.00020	0.02100	0.01230	0.02330	0.01020
SLSTM	0.00014	0.00004	0.01160	0.00617	0.00983	0.00476

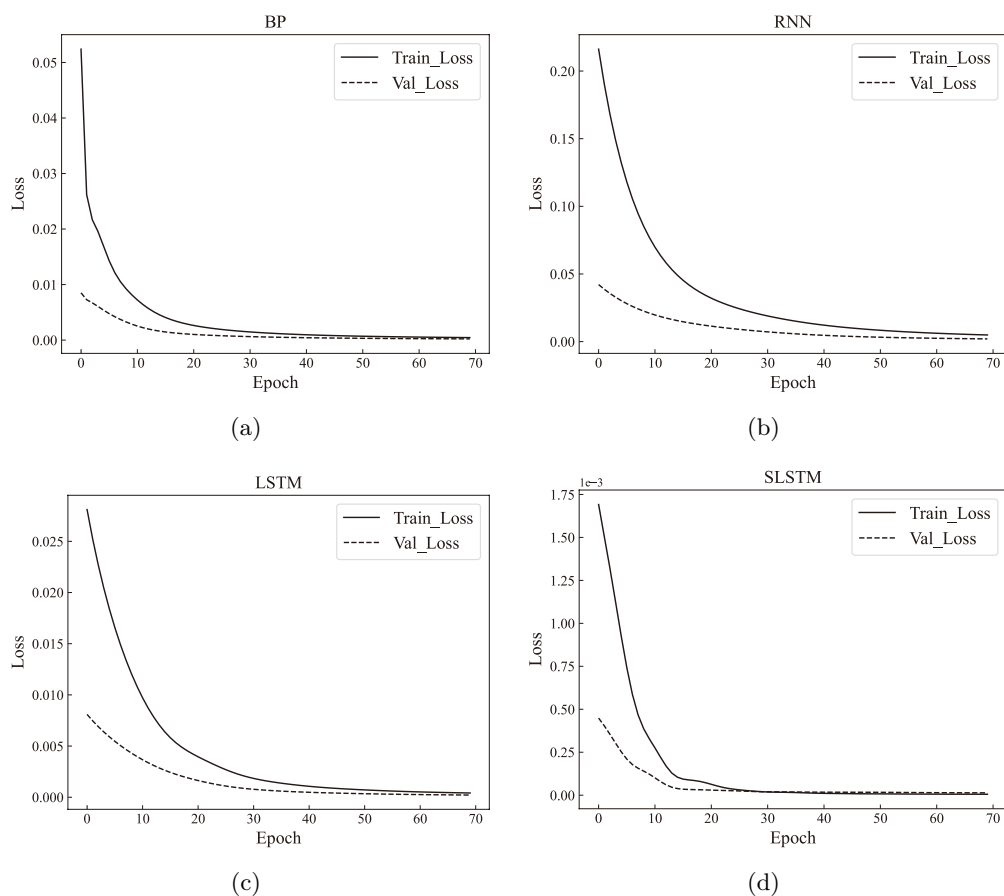


图 3 损失可视化

### 3.3.2 样本外收益表现

本文使用四种神经网络模型预测协方差矩阵以优化四种资产配置策略, 其累计收益率表现如图 4 所示. 通过图 4 可知, 在基于 GMV 的优化策略中, 使用 GMV-SLSTM 模型预测协方差矩阵的表现最为优越, 累计收益率显著高于 BP、RNN 和 LSTM 模型; 在基于 MV 的优化策略中, MV-SLSTM 同样表现出较高的累计收益率, 明显优于其他三种神经网络模型, 进一步验证了 SLSTM 在传统 MV 框架下的有效性; 在基于 RP 的优化策略中, RP-SLSTM 在收益率提升和波动性控制方面表现出色, 展示了较好的风险管理能力; 在基于 HRP 的优化策略中, HRP-SLSTM 实现了较高的累计收益率, 并保持了较低的波动, 表明 SLSTM 在分层风险平价策略中具有更优的资产配置效果.

综上所述, SLSTM 模型在预测协方差矩阵后, 于四种资产配置策略中均展现出优异的表现, 其捕捉市场动态和优化资产配置方面的优势尤为明显. 这些结果表明, 通过将 SLSTM 模型应用于协方差矩阵的预测, 并结合不同的资产配置策略, 可以显著提高投资组合的收益表现, 同时有效控制风险. 随后, 下文选用收益指标、风险指标和风险调整后收益指标等能够体

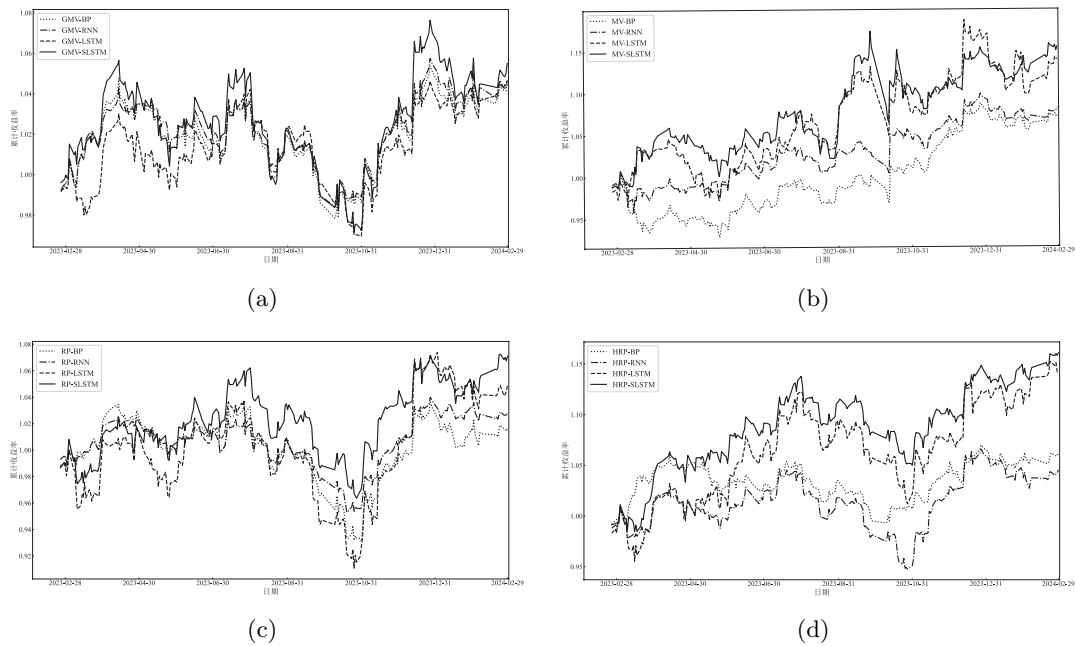


图 4 不同神经网络优化资产配置模型的累计收益率对比图

现投资组合特征的指标, 将使用不同神经网络模型预测的协方差矩阵, 分别应用于不同的资产配置策略, 并与简单的 EW 模型综合对比分析, 结果如表 6 所示。

从表 6 中收益指标和风险指标的对比可得, 不同资产配置模型在收益表现和风险管理方面存在显著差异。从 GMV 策略来看, GMV-SLSTM 模型的年化收益率达到 6.13%, 显著优于其他神经网络模型; 在 MV 策略下, MV-SLSTM 模型的年化收益率进一步提升至 17.24%。尽管此时波动率略微上升至 14.77%, 但其在夏普比率和确定等价回报上仍然表现出色。这表明 SLSTM 模型在提高收益的同时, 能够有效管理风险; 在 RP 策略中, RP-SLSTM 模型的年化收益率从基准模型的 3.10% 显著提升至 7.92%, 收益能力显著增强; 收益表现最优的是 HRP-SLSTM 模型, 其年化收益率最高, 达到了 18.04%。同时, 该模型的夏普比率高达 145.24%, 卡玛比率达到 257.22%, 这表明将 SLSTM 模型应用于优化 HRP 策略能够提升资产配置的收益能力和稳定性。

综上所述, 结合 LSTM 和 SSSM 模型的 SLSTM 模型通过预测动态协方差矩阵, 并将其应用于不同资产配置策略中, 增强了对市场变化的适应性和风险管理能力。相较于其他使用 SLSTM 优化的资产配置模型, HRP-SLSTM 模型的综合表现更为优越。HRP-SLSTM 模型不仅维持了 HRP 策略的风险均等化特点, 还在面对市场波动和结构性变化时展现出快速适应和调整的能力。该结合利用 SLSTM 在处理复杂数据和捕捉市场动态中非线性关系的优势, 提高了预测的准确性和有效性。

此外, 本文设置 60 个交易日的时间窗口, 对比不同神经网络模型预测的协方差矩阵在 HRP 资产配置中的表现, 以验证 SLSTM 模型的稳健性, 具体结果如表 7 所示。具体而言,

表 6 不同神经网络优化资产配置模型收益表现对比

投资策略	神经网络模型	收益指标		风险指标		风险调整后收益指标		
		累计收益率	年化收益率	年化波动率	最大回撤率	夏普比率	确定等价回报	卡玛比率
EW	-	4.31%	4.80%	8.33%	7.34%	15.59%	3.76%	65.35%
	-	0.32%	0.36%	3.52%	4.39%	-89.27%	0.17%	8.09%
GMV	BP	4.47%	4.97%	7.52%	6.43%	19.58%	4.12%	77.38%
	RNN	4.65%	5.18%	7.23%	5.49%	23.19%	4.39%	94.23%
	LSTM	4.66%	5.19%	8.24%	6.87%	20.44%	4.17%	75.45%
	SLSTM	5.50%	6.13%	9.65%	7.76%	27.23%	4.73%	78.96%
MV	-	13.79%	15.42%	14.58%	8.18%	81.81%	12.24%	188.59%
	BP	7.22%	8.04%	10.21%	7.61%	44.52%	6.48%	105.68%
	RNN	8.11%	9.04%	9.92%	4.93%	55.89%	7.57%	183.42%
	LSTM	13.95%	15.60%	15.63%	6.58%	77.45%	11.94%	237.03%
	SLSTM	15.41%	17.24%	14.77%	8.19%	93.07%	13.97%	210.65%
RP	-	2.79%	3.10%	6.68%	5.97%	-6.04%	2.43%	51.87%
	BP	1.71%	1.90%	9.62%	9.95%	-16.64%	0.51%	19.07%
	RNN	2.84%	3.16%	6.69%	7.02%	-5.15%	2.48%	44.92%
	LSTM	4.95%	5.51%	12.45%	12.42%	16.18%	3.19%	44.41%
	SLSTM	7.11%	7.92%	10.43%	9.07%	42.42%	6.29%	87.34%
HRP	-	2.98%	3.32%	5.13%	4.99%	-3.53%	2.92%	66.51%
	BP	6.32%	7.04%	7.56%	6.18%	46.80%	6.18%	113.87%
	RNN	4.57%	5.08%	9.30%	9.27%	17.01%	3.78%	54.80%
	LSTM	14.35%	16.05%	13.25%	9.17%	94.69%	13.41%	175.10%
	SLSTM	16.11%	18.04%	10.01%	7.01%	145.24%	16.53%	257.22%

表 7 不同神经网络优化 HRP 策略收益表现对比

投资策略	收益指标		风险指标		风险调整后收益指标		
	累计收益	年化收益率	年化波动率	最大回撤率	夏普比率	确定等价回报	卡玛比率
HRP	0.50%	0.64%	5.00%	5.16%	-57.31%	0.26%	12.33%
HRP-BP	1.19%	1.52%	8.48%	7.87%	-23.32%	0.44%	19.33%
HRP-RNN	1.76%	2.26%	8.10%	7.45%	-15.28%	1.28%	30.38%
HRP-LSTM	2.58%	3.32%	13.91%	12.01%	-1.32%	0.41%	27.61%
HRP-SLSTM	4.74%	6.11%	8.37%	6.71%	31.11%	5.05%	90.99%

HRP-SLSTM 模型的累计收益率为 4.74%，年化收益率为 6.11%，显著高于其他模型。同时，风险调整后收益指标如夏普比率为 31.11%，卡玛比率为 90.99%，均优于其他策略。该结果表明，SLSTM 模型在优化 HRP 策略时展现出明显的稳健性和显著的性能优势，特别是在应对复杂市场环境和动态风险管理需求方面表现突出。这表明在时间窗口为 60 个交易日时，

HRP-SLSTM 模型依然保持了较强的稳健性, 确保了协方差矩阵预测的准确性, 同时优化了资产配置的风险调整后收益.

## 4 敏感性分析

基于上述分析, 本节将从以下三个方面对 HRP-SLSTM 模型在时间窗口为 30 个交易日情形下进行敏感性分析: 其一, 探究注意力机制结构对模型的影响; 其二, 剖析损失函数结构对模型的影响; 其三, 分析不同资产组合对模型的影响. 通过检验模型在不同参数变化下的表现, 本节旨在进一步验证 HRP-SLSTM 模型在应对各类因素变化时的稳健性, 从而为资产配置和风险管理策略提供更全面的依据.

### 4.1 不同注意力机制的影响

Transformer 的自注意力机制 (self-attention mechanism, SAM) 能够有效捕捉长距离依赖关系, 在时间序列预测中得到广泛应用 (Du et al., 2023). Xu et al. (2022) 研究表明, 引入 SAM 结构能够使模型更有效地学习数据的时间和空间相关性, 从而更好地预测协方差矩阵. 因此, 本节在 HRP 策略中分别应用融合 SAM 结构的 LSTM 模型 (ALSTM) 和融合 SSSM 的 LSTM 模型 (SLSTM), 并进行对比分析. 通过比较这两种混合模型在协方差矩阵预测中的表现, 验证 SSSM 作为注意力机制的优越性, 其应用于 HRP 资产配置策略在累计收益率方面的对比走势图如图 5 所示.

从图 5 可得, HRP-SLSTM 模型的表现明显优于 HRP-ALSTM 模型. 具体而言, HRP-SLSTM 模型在整个样本外期间的累计收益率较高, 波动较小. 此结果表明, 将 SSSM 与 LSTM 模型相结合的模型能更准确地捕捉市场动态并提升组合表现. 相比之下, 尽管 HRP-ALSTM 模型通过 SAM 增强了特征捕捉能力, 但其在市场动态变化中的表现不佳, 收益率波动较大. 总体而言, 结合 SSSM 的 LSTM 模型在实际应用中展示了更强优势, 在风险管理和收益优化方面尤其明显. HRP-ALSTM 模型与 HRP-SLSTM 模型收益表现对比如表 8 所示.

由表 8 可得, HRP-SLSTM 模型的累计收益率为 16.11%, 高于 HRP-ALSTM 模型的 12.35%. 在年化收益率方面, HRP-SLSTM 模型也以 18.04% 超过 HRP-ALSTM 模型的

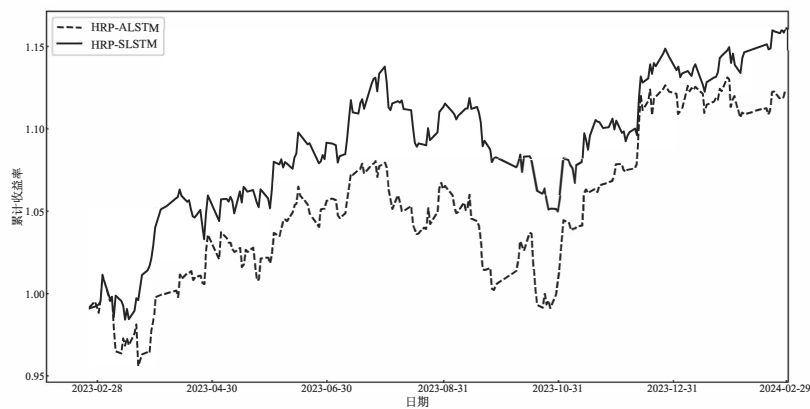


图 5 HRP-ALSTM 与 HRP-SLSTM 模型累计收益率对比图

13.80%, 显示了其更高的收益能力. 在风险调整后的回报率方面, HRP-SLSTM 模型同样展示了其优越性, 卡玛比率达到了 257.22%, 明显高于 HRP-ALSTM 模型的 176.39%. 这一数据表明, 在考虑最大回撤的情况下, HRP-SLSTM 模型提供了更高的风险调整回报. 此外, HRP-SLSTM 模型在夏普比率上也达到了 145.24%, 远超 HRP-ALSTM 模型的 89.23%, 进一步证明了其在承担单位风险时所带来收益的高效性.

结果表明, 相较于 SAM 结构与 LSTM 模型结合, 将 SSSM 与 LSTM 模型结合可以使 LSTM 模型更加灵活地适应金融市场波动. 通过敏感性分析评估了这两种机制在提高预测准确性和稳定性方面的具体表现, 验证了 SSSM 的有效性, 为投资决策提供更加科学和精确的支持.

表 8 HRP-ALSTM 与 HRP-SLSTM 模型收益表现对比

投资策略	收益指标		风险指标		风险调整后收益指标		
	累计收益率	年化收益率	年化波动率	最大回撤率	夏普比率	确定等价回报	卡玛比率
HRP-ALSTM	12.35%	13.80%	11.54%	7.82%	89.23%	11.80%	176.39%
HRP-SLSTM	16.11%	18.04%	10.01%	7.01%	145.24%	16.53%	257.22%

#### 4.2 不同损失函数结构的影响

本节在四种神经网络模型预测协方差矩阵并应用于 HRP 策略的基础上, 通过对目标损失、对称损失和半正定损失这三种损失函数的不同组合结构<sup>9</sup> 进行敏感性分析, 旨在验证复合损失函数对模型训练拟合的有效性和测试的泛化性. 不同神经网络模型针对样本内训练集和验证集数据, 其不同损失函数结构的变化趋势如图 6 所示.

由图 6 可知, 每种神经网络模型均尝试了四种不同的损失函数配置: 单一目标损失 (Loss\_T)、目标损失加上对称损失 (Loss\_T\_S)、目标损失加上半正定损失 (Loss\_T\_PSD) 以及复合损失函数 (Loss\_ALL). 这些配置的对比显示了复合损失函数在训练和验证过程中对神经网络模型性能的积极影响. 具体而言, 采用复合损失函数的模型在验证损失上展示出更平稳和更低的下降趋势, 这一特征在 SLSTM 模型中尤为明显. 这表明, 通过结合对称损失和半正定损失可以有效提高模型对动态协方差矩阵预测的准确性. 此外, SLSTM 模型采用复合损失函数时显示出最低的验证损失, 这突出了复合损失函数在处理复杂时间序列数据时的优势. 使用四种不同损失函数结构的神经网络模型预测动态协方差矩阵后, 在 HRP 资产配置下的收益表现对比如表 9 所示.

从表 9 可知, 复合损失函数能够显著提高神经网络模型在动态金融环境中的表现. 以 BP 模型为例, 当仅使用目标损失时, 策略的年化收益率为 1.64%, 在此基础上分别添加对称损失和半正定损失, 年化收益率分别提升至 5.25% 和 2.01%, 而在综合使用复合损失函数后,

<sup>9</sup>T 代表目标损失函数 (Target Loss), 通常使用均方误差作为目标函数; T\_S 则为目标损失函数加上对称性损失函数 (Symmetry Loss), 保证预测矩阵的对称性; T\_PSD 表示目标损失函数加上半正定损失函数 (Positive Semi-Definite Loss), 用于确保预测协方差矩阵的半正定性; ALL 代表复合损失函数, 结合了目标损失、半正定损失和对称损失.

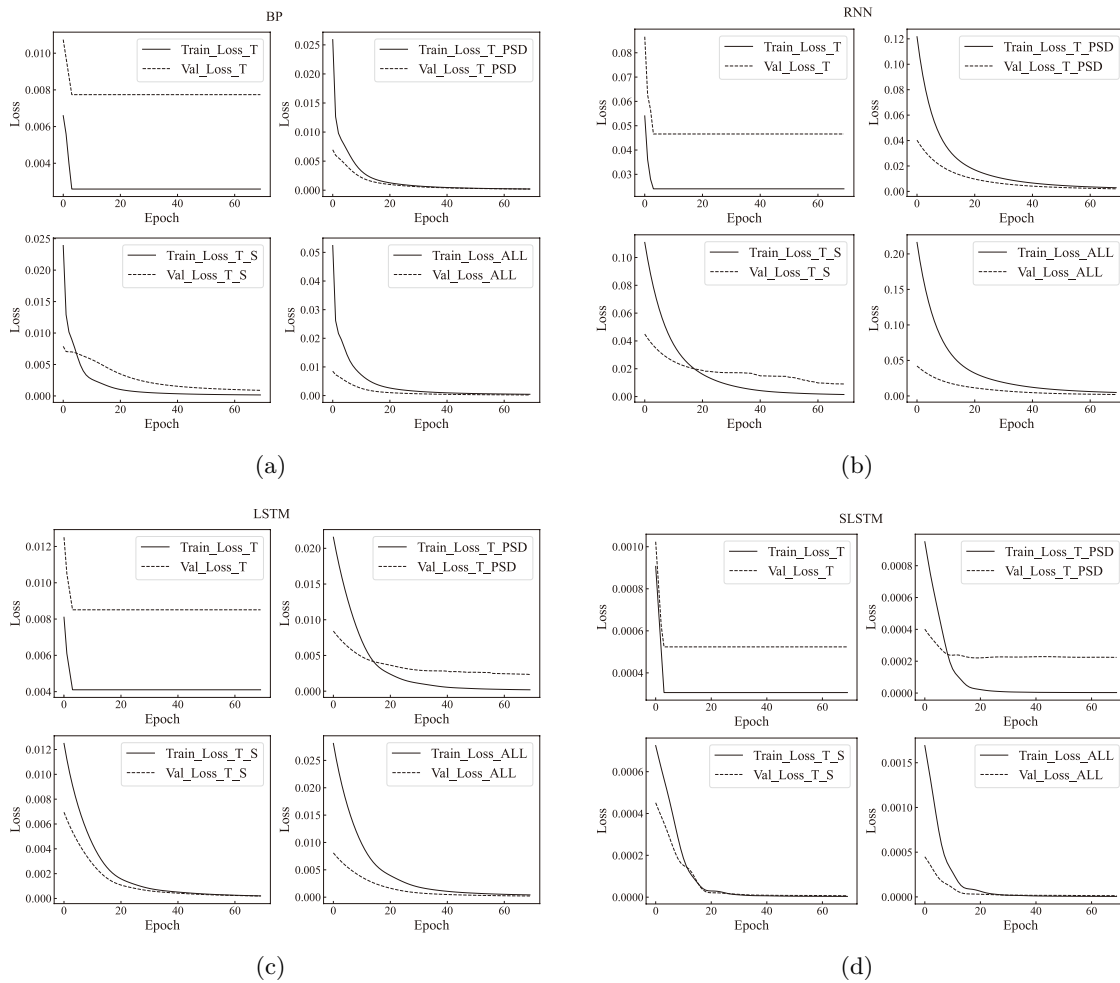


图 6 不同损失函数结构可视化

策略的年化收益率提升至 7.04%；对于 RNN 模型，当仅使用目标损失时，策略的年化收益率为 3.44%，在此基础上分别添加对称损失和半正定损失，年化收益率分别提升至 3.96% 和 3.76%，最后使用复合损失函数，年化收益率达到 5.08%，夏普比率为 17.01%；LSTM 模型的表现也类似，策略从单一目标损失的 6.42% 的年化收益率，在此基础上分别添加对称损失和半正定损失，年化收益率分别提升至 13.80% 和 7.49%，最后在复合损失函数应用后，年化收益率提高至 16.05%；SLSTM 模型通过引入复合损失函数，其年化收益率能够达到 18.04%，夏普比率高达 145.24%，并且最大回撤率降至 7.01%，显示出最优的风险调整后收益和稳健性。这些结果显示，复合损失函数显著提升了模型的收益能力和风险调整后的表现。

综上所述，添加对称损失和半正定损失能够有效提高模型的收益和风险管理能力，有助于全面地捕捉协方差矩阵的关键特性（对称性和正定性）。这不仅提高了模型对时间序列数据结构变化的适应性，而且增强了预测的稳定性和准确性。

表 9 不同损失函数结构的神经网络优化 HRP 模型收益表现对比

神经网络	损失函数	收益指标		风险指标		风险调整后收益指标		
		累计收益率	年化收益率	年化波动率	最大回撤率	夏普比率	确定等价回报	卡玛比率
BP	Loss_T	1.47%	1.64%	9.45%	9.93%	-19.69%	0.30%	16.50%
	Loss_T_PSD	1.81%	2.01%	8.38%	7.50%	-17.79%	0.96%	26.81%
	Loss_T_S	4.72%	5.25%	10.59%	8.90%	16.51%	3.57%	58.97%
	Loss_ALL	6.32%	7.04%	7.56%	6.18%	46.80%	6.18%	113.87%
RNN	Loss_T	3.09%	3.44%	10.06%	11.48%	-0.63%	1.92%	29.93%
	Loss_T_PSD	3.38%	3.76%	8.46%	8.83%	3.10%	2.69%	42.60%
	Loss_T_S	3.56%	3.96%	5.72%	5.18%	8.06%	3.47%	76.43%
	Loss_ALL	4.57%	5.08%	9.30%	9.27%	17.01%	3.78%	54.80%
LSTM	Loss_T	5.77%	6.42%	7.40%	6.77%	39.53%	5.60%	94.87%
	Loss_T_PSD	6.73%	7.49%	6.68%	4.91%	59.80%	6.82%	152.56%
	Loss_T_S	12.35%	13.80%	11.54%	7.82%	89.23%	11.80%	176.39%
	Loss_ALL	14.35%	16.05%	13.25%	9.17%	94.69%	13.41%	175.10%
SLSTM	Loss_T	6.93%	7.72%	8.93%	8.39%	47.24%	6.52%	92.01%
	Loss_T_PSD	8.82%	9.84%	9.21%	8.48%	68.81%	8.57%	116.08%
	Loss_T_S	9.79%	10.93%	11.75%	8.73%	63.23%	8.86%	125.16%
	Loss_ALL	16.11%	18.04%	10.01%	7.01%	145.24%	16.53%	257.22%

### 4.3 不同资产组合的影响

为分析 ETFs 数据对 HRP-SLSTM 模型的敏感性, 4.3.1 节考察了 ETFs 追踪的大类资产指数收益对模型表现的影响, 并与 ETFs 收益表现进行对比. 此外, 考虑到 ETFs 底层资产的重叠问题, 4.3.2 节在剔除重叠 ETFs 后重新评估模型表现. 通过使用不同资产组合对 HRP-SLSTM 模型进行敏感性分析, 并与最优基准模型<sup>10</sup>结果进行对比, 验证了 HRP-SLSTM 模型在不同资产组合下的适应性与稳健性.

#### 4.3.1 ETFs 追踪指数的问题分析

本节基于 HRP-SLSTM 模型, 分别在 ETFs 及其追踪的指数数据下进行资产配置分析, 旨在探讨不同资产组合对 HRP-SLSTM 模型收益表现的敏感性. 为了便于理解不同 ETFs 及其追踪指数的构成, 表 10 展示了所选 ETFs 及其追踪的大类资产指数<sup>11</sup>.

本节对表 10 所选 ETFs 及其追踪的大类资产指数在 HRP-SLSTM 模型中的表现进行

<sup>10</sup>最优基准模型指选取 3.3 节所有模型中整体表现最优的模型, 即使用 11 个 ETFs 的 HRP-SLSTM 模型, 由于篇幅限制, 本文后续仅比较最优基准模型结果与其他模型结果.

<sup>11</sup>除部分 ETFs (SPY, EEM, VGK, ACWI) 明确公布了其追踪的指数外, 对于未明确公布具体追踪指数的 ETFs, 本文采用其他相关数据作为其大类资产替代, 具体包括: TLT 参考 30 年期美国国债; BND 参考美国总债券市场指数; LQD 参考 IBOXX 美元债总回报; HYG 参考亚洲美元企业高收益债券指数; GLD 参考黄金现货市场; USO 参考 WTI 原油; VNQ 参考 MSCI 美国房地产投资信托指数, 以确保数据分析的完整性和一致性, 数据均来源于 WIND 金融数据库.

比较, 结果如表 11 所示. 研究结果表明, 无论是使用 ETFs 还是其所追踪的大类资产指数, HRP-SLSTM 模型在不同资产配置下均展现出较强的稳健性, 都能够有效实现风险优化与收益平衡. 具体而言, 尽管 ETFs 的风险水平略高于大类资产指数, 但其收益表现明显优越, 尤其在风险调整后的收益指标上, ETFs 的夏普比率和卡玛比率均高于大类资产指数. 这表明 HRP-SLSTM 模型能够有效地平衡风险与收益, 验证了其在不同资产组合下的适应性和稳健性.

表 10 ETFs 及其追踪指数

ETFs 类别	ETFs 代码	追踪指数
股票类 ETFs	SPY	标准普尔 500 指数
	EEM	MSCI 新兴市场指数
	VGK	FTSE 欧洲股票指数
	TLT	美国长期国债
债券类 ETFs	BND	美国总债券市场
	LQD	iBoxx 投资级公司债指数
	HYG	iBoxx 高收益公司债指数
商品类 ETFs	GLD	黄金现货市场
	USO	原油价格 WTI 原油期货合约
其他资产类别	VNQ	MSCI 美国房地产指数
	ACWI	MSCI 全球指数

表 11 ETFs 及其追踪的大类资产指数的 HRP-SLSTM 模型收益表现对比

资产组合	收益指标		风险指标		风险调整后收益指标		
	累计收益率	年化收益率	年化波动率	最大回撤率	夏普比率	确定等价回报	卡玛比率
ETFs	16.11%	18.04%	10.01%	7.01%	145.24%	16.53%	2.57
大类资产	11.00%	11.63%	7.74%	5.67%	105.01%	10.73%	2.05

#### 4.3.2 ETFs 底层资产重叠的问题分析

ETFs 所追踪的资产类别及其构成成分对投资组合的风险与收益具有重要影响. 因此, 分析 ETFs 间底层资产的重叠现象对优化投资策略至关重要. 故本节在剔除存在底层资产重叠的 ETFs 资产后, 重新评估 HRP-SLSTM 模型表现, 以进一步验证不同资产配置对 HRP-SLSTM 模型的敏感性. 参考表 10 所列的 ETFs 及其追踪指数, 本节对部分存在底层资产重叠的 ETFs (TLT, LQD, ACWI) 进行了剔除<sup>12</sup>. 剔除部分 ETFs 后, 本节重新评估 HRP-SLSTM 模型表现, 并将其与最优基准模型表现进行对比, 结果如表 12 所示.

<sup>12</sup>TLT 与 BND 在长期美国国债部分存在重叠, 由于 BND 已涵盖长期国债成分, 保留 BND 可以有效覆盖 TLT 的相关投资, 因此剔除 TLT. LQD 与 BND 在投资级公司债部分也存在重叠, 保留 BND 即可有效覆盖 LQD 的相关成分, 因此剔除 LQD. SPY 与 ACWI 在美国股票成分上完全重合, 鉴于 SPY 已能全面代表美国股市表现, 保留 SPY 能够有效避免对美国股票市场的重复风险暴露, 因此剔除 ACWI.

表 12 包含重叠 ETFs 和剔除重叠 ETFs 后的 HRP-SLSTM 模型收益表现对比

资产组合	收益指标		风险指标		风险调整后收益指标		
	累计 收益率	年化 收益率	年化 波动率	最大 回撤率	夏普 比率	确定等 价回报	卡玛 比率
包含重叠 ETFs	16.11%	18.04%	10.01%	7.01%	145.24%	16.53%	2.57
剔除重叠 ETFs	10.88%	12.15%	7.77%	5.52%	111.28%	11.24%	2.20

结果显示, 最优基准模型表现在收益方面表现较为优越, 年化收益率为 18.04%, 夏普比率为 145.24%。而剔除重叠 ETFs 后, 虽然风险指标有所改善, 年化波动率从 10.01% 降低至 7.77%, 最大回撤率从 7.01% 降低至 5.52%, 但收益表现有所下降, 年化收益率降至 12.15%, 夏普比率降至 111.28%。综上所述, 剔除重叠 ETFs 后, HRP-SLSTM 模型的风险控制能力得到提升, 但收益表现有所下降。与包含重叠 ETFs 的结果对比, 剔除重叠资产操作并未显著改善模型的整体表现, 进一步验证了 HRP-SLSTM 模型在不同资产配置下的稳健性与适应性。

## 5 结论

资产配置是投资管理的核心环节, 通过优化不同资产类别之间的权重分配, 能够实现风险和收益的最佳平衡。随着全球经济和市场环境的不断变化, 有效的资产配置策略对于增强投资组合表现和抵御市场波动尤为重要。因此, 本文在现有研究基础上构建资产配置理论模型, 选取多种类型的交易所交易基金作为研究样本, 采用结合 LSTM 模型和 SSSM 的 SLSTM 模型以及复合损失函数来预测动态协方差矩阵, 并将其应用于 HRP 资产配置策略中。本文主要结论如下:

第一, 通过对比传统神经网络模型与构建的 SLSTM 模型的性能, 本文发现 SLSTM 模型在预测动态协方差矩阵方面具有显著优势。具体而言, SLSTM 模型结合了 LSTM 的长期记忆能力和 SSSM 的动态调整功能, 能够更加精确地捕捉和预测金融资产间复杂的波动性关联, 从而显著提高模型在实际金融决策中的应用效率和有效性。

第二, 通过融合目标损失、对称损失和半正定损失的复合损失函数, 能够有效增强模型对数据特征的学习能力。具体而言, 传统的单一损失函数在处理复杂金融数据时可能面临模型对目标关键属性学习不足的问题, 进而导致预测性能不佳。本文通过引入复合损失函数, 不仅增强了模型保留协方差矩阵关键属性的能力, 还显著提高了其在复杂数据环境中的稳定性与泛化能力。

第三, 与其他传统资产配置模型相比, 基于 SLSTM 的 HRP 模型展现出更为优越的风险调整后收益表现。该优势源自 SLSTM 模型对动态协方差的精确预测以及更为合理的资产权重分配, 使 HRP 策略能够更有效地降低潜在的投资风险, 同时提升投资组合的整体收益率。该结果不仅验证了 SLSTM 模型在金融资产管理中的实用性, 还强调了在现代金融市场条件下, 采用先进的神经网络模型对传统资产配置方法进行改进的重要性和必要性。

以上结论表明, 通过结合 SLSTM 模型与 HRP 策略, 不仅提高了协方差矩阵预测的准确性和稳定性, 还优化了资产配置的风险调整后收益, 为现代金融市场中的资产管理提供了新的思路和方法。为进一步推进该成果的实践应用, 建议金融机构在资产配置中加强智能化

算法的应用, 定期更新并动态调整风险模型. 同时, 监管机构可通过推广该类技术方法, 鼓励投资者在高波动市场环境下采用更加灵活的风险管理工具, 以提高市场整体稳定性.

## 参 考 文 献

- 崔炎炎, 刘立新, (2022). 网络舆情赋能金融科技股票收盘价预测研究 [J]. 统计研究, 39(6): 148–160.  
Cui Y Y, Liu L X, (2022). Research on the Impact of Internet Public Opinion on Fintech Stock Closing Price Forecast[J]. Statistical Research, 39(6): 148–160.
- 姜宏伟, 宁炜, 薛浩, (2022). 机构投资与金融稳定: 基于 A 股 ETF 套利交易的视角 [J]. 管理世界, 38(4): 29–49.  
Jiang F W, Ning W, Xue H, (2022). Institutional Investment and Financial Stability: Evidence from A-Share ETF Arbitrage Trading[J]. Journal of Management World, 38(4): 29–49.
- 刘广应, 包悦妍, 林金官, (2022). 基于金融高频数据的 LASSO-CDRD 协方差矩阵预测模型 [J]. 统计研究, 39(9): 145–160.  
Liu G Y, Bao Y Y, Lin J G, (2022). LASSO-CDRD Covariance Matrix Prediction Model Based on High Frequency Financial Data[J]. Statistical Research, 39(9): 145–160.
- 杨青, 王晨蔚, (2019). 基于深度学习 LSTM 神经网络的全球股票指数预测研究 [J]. 统计研究, 36(3): 65–77.  
Yang Q, Wang C W, (2019). A Study on Forecast of Global Stock Indices Based on Deep LSTM Neural Network[J]. Statistical Research, 36(3): 65–77.
- 赵大萍, 房勇, (2020). 基于 VaR 的风险平价投资策略及应用 [J]. 系统工程学报, 35(5): 623–641.  
Zhao D P, Fang Y, (2020). Var-based Risk Parity Investment Strategy and Its Application[J]. Journal of Systems Engineering, 35(5): 623–641.
- 周骥, 李仲飞, 曾燕, (2022). 复杂网络视角下行业风险传染与银行信贷配置 [J]. 管理科学学报, 25(2): 24–46.  
Zhou Q, Li Z F, Zeng Y, (2022). Industry Risk Contagion and Bank Credit Allocation from the Perspective of Complex Network[J]. Journal of Management Sciences in China, 25(2): 24–46.
- Britten-Jones M, (1999). The Sampling Error in Estimates of Mean-variance Efficient Portfolio Weights[J]. Journal of Finance, 54(2): 655–671.
- Buccheri G, Borgetti G, Corsi F, Lillo F, (2021). A Score-driven Conditional Correlation Model for Noisy and Asynchronous Data: An Application to High-frequency Covariance Dynamics[J]. Journal of Business & Economic Statistics, 39(4): 920–936.
- Burggraf T, (2021). Beyond Risk Parity — A Machine Learning-based Hierarchical Risk Parity Approach on Cryptocurrencies[J]. Finance Research Letters, 38: 101523.
- DeMiguel V, Garlappi L, Uppal R, (2009). Optimal Versus Naive Diversification: How Inefficient is the  $1/N$  Portfolio Strategy?[J]. Review of Financial Studies, 22(5): 1915–1953.
- Drobotz W, Hollstein F, Otto T, Prokopczuk M, (2025). Estimating Stock Market Betas via Machine Learning[J]. Journal of Financial and Quantitative Analysis, 60(3): 1074–1110.
- Du W, Côté D, Liu Y, (2023). Saits: Self-attention-based Imputation for Time Series[J]. Expert Systems with Applications, 219: 119619.
- Engle R, (2002). Dynamic Conditional Correlation[J]. Journal of Business & Economic Statistics, 20(3): 339–350.

- Fan J, Fan Y, Lv J, (2008). High Dimensional Covariance Matrix Estimation Using a Factor Model[J]. *Journal of Econometrics*, 147(1): 186–197.
- Fan J Q, Liao Y, Mincheva M, (2011). High Dimensional Covariance Matrix Estimation in Approximate Factor Models[J]. *Annals of Statistics*, 39(6): 3320.
- Fiszeder P, Orzeszko W, (2021). Covariance Matrix Forecasting Using Support Vector Regression[J]. *Applied Intelligence*, 51(10): 7029–7042.
- Gu A, Dao T, (2023). Mamba: Linear-time Sequence Modeling with Selective State Spaces[J]. *arXiv Preprint arXiv: 2312.00752*.
- Hafner C M, Wang L, (2023). A Dynamic Conditional Score Model for the Log Correlation Matrix[J]. *Journal of Econometrics*, 237(2): 1–18.
- Imajo K, Minami K, Ito K, Nakagawa K, (2021). Deep Portfolio Optimization via Distributional Prediction of Residual Factors[C]// *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(1): 213–222.
- Jain P, Jain S, (2019). Can Machine Learning-Based Portfolios Outperform Traditional Risk-Based Portfolios? The Need to Account for Covariance Misspecification[J]. *Risks*, 7(3): 74.
- Karevan Z, Suykens J A K, (2020). Transductive LSTM for Time-Series Prediction: An Application to Weather Forecasting[J]. *Neural Networks*, 125: 1–9.
- Lopez de Prado M, (2016). Building Diversified Portfolios That Outperform Out of Sample[J]. *Journal of Portfolio Management*, 42(4): 59–69.
- Markowitz H, (1952). Portfolio Selection[J]. *Journal of Finance*, 7(1): 77–91.
- Raffinot T, (2018). Hierarchical Clustering-Based Asset Allocation[J]. *Journal of Portfolio Management*, 44(2): 89–99.
- Shi Y, Ai C, Shi Y, Ying T, Xu Q, (2023). Large Covariance Estimation Using a Factor Model with Common and Group-Specific Factors[J]. *Journal of Forecasting*, 42(8): 2217–2248.
- Xu H, Luo H, Wu Z, Wu F, Bao L, (2022). Towards Predicting the Measurement Noise Covariance with a Transformer and Residual Denoising Autoencoder for GNSS/INS Tightly-Coupled Integrated Navigation[J]. *Remote Sensing*, 14(7): 1691.

## 附录

### 等权重 (EW) 的资产配置过程

等权重模型是资产配置诸多理论中最为直观的风险配置模型, 也可称之为恒定比投资法, 即投资的每个标的都拥有相同的持仓比例, 具体而言:

$$x_i = 1/N, \quad (\text{A1})$$

其中  $x_i$  表示分配到每个资产  $i$  的权重, 并且资产数量为  $N$ .

### 全局最小方差 (GMV) 的资产配置过程

在市场不存在卖空的情况下, 传统的全局最小方差投资组合模型的最优权重  $x^*$  为如下问题的最优解:

$$\begin{aligned} x^* &= \arg \min x^T \Sigma x, \\ \text{s.t. } \sum_{i=1}^n x_i &= 1, \quad 0 \leq x_i \leq 1, \end{aligned} \quad (\text{A2})$$

其中  $\Sigma$  表示资产收益的协方差矩阵,  $x_i$  表示分配到每个资产  $i$  的权重.

#### 均值-方差 (MV) 的资产配置过程

MV 模型是现代投资组合理论的基础, 旨在优化资产权重以在给定的风险水平下最大化预期收益或在给定的预期收益水平下最小化风险. 基于 MV 模型的假设, 若以投资者效用最大为目标, MV 模型的最优权重  $x^*$  为如下问题的最优解:

$$\begin{aligned} x^* &= \arg \max \left( x^T R - \frac{\lambda}{2} x^T \Sigma x \right), \\ \text{s.t. } \sum_{i=1}^N x_i &= 1, \quad 0 \leq x_i \leq 1, \end{aligned} \quad (\text{A3})$$

其中  $R = (r_1, r_2, \dots, r_n)^T$  为各资产的期望收益率,  $x_i$  为资产  $i$  的配置权重,  $\Sigma$  表示资产收益的协方差矩阵,  $\lambda$  为风险厌恶水平.

#### 风险平价 (RP) 的资产配置过程

风险平价投资组合 (risk parity portfolios) 也被称为等风险贡献投资组合 (equal risk contribution portfolios), 其构造投资组合模型的核心思想是使不同资产对投资组合的总风险的贡献具有一致性, 该做法能够有效降低投资组合的波动, 从根本上分散风险. 风险平价模型使用的是等风险的方法计算最佳资产配置权重, 即依据各资产对总投资组合风险的等贡献原则进行权重分配, 区别于传统的均等波动率分配方式, 其目的在于实现资产之间风险贡献的均衡, 而非资产本身波动率的简单等量.

若在一个包含  $N$  种金融资产的组合中, 资产  $i$  的收益率记为  $r_i$ , 对应的配置权重记为  $x_i$ , 那么该投资组合的收益率  $r_p$  及其标准差  $\sigma_p$  可分别表示为:

$$r_p = \sum_{i=1}^n x_i r_i, \quad (\text{A4})$$

$$\sigma_p = \sqrt{\sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij}}, \quad (\text{A5})$$

其中  $\sigma_{ij}$  为资产  $i$  和  $j$  的协方差.

首先, 将投资组合的标准差  $\sigma_p$  关于资产  $i$  的配置权重  $x_i$  求偏导数, 得到单个资产  $i$  的边际风险贡献  $\text{MRC}_i$  (即每增加一单位资产  $i$  所增加的风险).

$$\text{MRC}_i(x) = \frac{\partial \sigma_p}{\partial x_i} = \frac{\sum_{j=1}^n x_j \sigma_{ij}}{\sigma_p} = \frac{\text{cov}(r_i, r_p)}{\sigma_p}, \quad (\text{A6})$$

其中  $\text{cov}(r_i, r_p)$  为资产  $i$  的收益率  $r_i$  与投资组合的收益率  $r_p$  的协方差.

其次, 由于边际风险贡献并不能反映资产的增加对投资组合整体风险的影响, 故引入总风险贡献值来衡量每个资产对整个投资组合风险的贡献. 资产  $i$  对于投资组合的总风险贡献  $\text{TRC}_i(x)$  等于资产  $i$  的边际风险贡献与其资产配置权重  $x_i$  相乘, 即:

$$\text{TRC}_i(x) = x_i \text{MRC}_i(x) = x_i \frac{\sum_{j=1}^n x_j \sigma_{ij}}{\sigma_p} = x_i \frac{\text{cov}(r_i, r_p)}{\sigma_p}. \quad (\text{A7})$$

最后, 根据风险平价模型定义 (对各资产均等分配风险), 可得:

$$x_i \frac{\partial \sigma_p}{\partial x_i} = x_j \frac{\partial \sigma_p}{\partial x_j}, \quad \forall i, j, \quad (\text{A8})$$

即资产  $i$  对于投资组合的总风险贡献  $\text{TRC}_i(x)$  等于资产  $j$  对于投资组合的总风险贡献  $\text{TRC}_j(x)$ .

在市场不存在卖空的情况下, 风险平价投资组合模型的最优权重  $x^*$  为如下问题的最优解:

$$\begin{aligned} x^* = \arg \min & \sum_{i=1}^n \sum_{j=1}^n \left( x_i \frac{\partial \sigma_p}{\partial x_i} - x_j \frac{\partial \sigma_p}{\partial x_j} \right)^2, \\ \text{s.t.} & \sum_{i=1}^n x_i = 1, \quad 0 \leq x_i \leq 1. \end{aligned} \quad (\text{A9})$$

### 分层风险平价 (HRP) 的资产配置过程

HRP 模型通过分析资产之间的相关性, 并重新配置资产权重以达到风险的均衡分配. HRP 的实施包括三个主要步骤: 分层聚类、准对角化处理和递归二分法权重分配, 具体步骤如下所示.

#### 1) 层次聚类

对于两个资产  $i$  和  $j$ , 将估计出的协方差矩阵  $\Sigma$  转换为相关性矩阵  $R$ , 其中  $R_{i,j} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii} \cdot \Sigma_{jj}}}$ . 接着, 根据相关性矩阵  $R$  计算相关距离矩阵  $D$ , 其中距离被定义为  $D(i, j) = \sqrt{\frac{1}{2}(1 - R_{ij})}$ , 该定义能够捕捉到资产间的非相似性. 然后, 计算距离矩阵  $D$  中任意两列之间的欧式距离, 得到增广的距离矩阵  $\bar{D}$ :

$$\bar{D}(i, j) = \sqrt{\sum_{k=1}^N (D(k, i) - D(k, j))^2}, \quad (\text{A10})$$

其中  $N$  为资产数量.

根据式 (A10), 使用递归来构造聚类树. 将簇的集合定义为  $U$ , 其中第一个簇  $U[1]$  计算方式为:

$$U[1] = \arg \min_{i,j} \bar{D}(i, j). \quad (\text{A11})$$

通过使用单连接聚类计算簇  $U[1]$  中其他资产的距离来更新增广的距离矩阵  $\bar{D}$ . 因此, 对于簇外的任意资产  $i$ , 计算到新形成簇的距离:

$$\bar{D}(i, U[1]) = \min(\bar{D}(i, j^*), \bar{D}(i, j^*)). \quad (\text{A12})$$

根据式 (A12) 递归地将投资组合中的资产组合成簇, 并更新距离矩阵  $\bar{D}$ , 直到只剩下一个单一的簇.

#### 2) 准对角化

从聚类树中按照聚类的紧密程度遍历资产,生成一个新的资产排序索引,用一个向量  $\sigma$  来表示,其中  $\sigma[i]$  表示原始索引为  $i$  的资产在新顺序中的位置.

根据排序索引  $\sigma$ , 构建置换矩阵  $S$ . 置换矩阵是一个  $N \times N$  的矩阵,其中  $N$  是资产的总数,  $\sigma[i]$  表示原始索引为  $i$  的资产在新顺序中的位置. 如果资产  $i$  被移到了位置  $j$ , 则矩阵  $S$  的第  $i$  行和第  $j$  列的元素为 1, 其余元素为 0. 置换矩阵  $S$  可以表示为:

$$S_{ij} = \begin{cases} 1, & j = \sigma(i), \\ 0, & j \neq \sigma(i). \end{cases} \quad (\text{A13})$$

在根据排序索引  $\sigma$  得到置换矩阵  $S$  后, HRP 模型采用准对角化的方法优化协方差矩阵的结构. 准对角化的目的是将相似的资产放置在协方差矩阵的邻近位置,从而简化资产之间复杂的依赖关系,该过程通过以下变换实现:

$$\tilde{\Sigma} = S\Sigma S^T, \quad (\text{A14})$$

其中  $\Sigma$  是基于历史数据估计出的协方差矩阵,  $\tilde{\Sigma}$  为新协方差矩阵. 在矩阵  $\tilde{\Sigma}$  中,相关性高的资产会被放置在一起. 准对角化后的协方差矩阵  $\tilde{\Sigma}$  更容易处理,该矩阵使得相关性强的资产在协方差矩阵中位置相邻,对角线附近的元素较大,远离对角线的元素较小.

### 3) 递归二分法分配资产权重

对准对角化后的协方差矩阵  $\tilde{\Sigma}$ , 计算每个子簇的方差. 假设在递归过程中得到了两个子簇  $C_1$  和  $C_2$ , 则方差为:

$$\text{Var}(C_1) = w_{C_1}^T \tilde{\Sigma}_{C_1} w_{C_1}, \quad (\text{A15})$$

$$\text{Var}(C_2) = w_{C_2}^T \tilde{\Sigma}_{C_2} w_{C_2}. \quad (\text{A16})$$

利用递归二分法,将总风险预算在子簇之间进行分配. 权重分配的公式为:

$$w_{C_1} = \frac{\text{Var}(C_2)}{\text{Var}(C_1) + \text{Var}(C_2)}, \quad (\text{A17})$$

$$w_{C_2} = \frac{\text{Var}(C_1)}{\text{Var}(C_1) + \text{Var}(C_2)}. \quad (\text{A18})$$

按照上述步骤将权重逐层分配给更小的子簇,直到所有资产的权重都被分配. 上述过程使得 HRP 模型不仅优化了资产间的风险分散,还通过结构化的层次方法增加了投资组合的整体稳定性和抗风险能力.