

学校代码: 10270

学号: 202502526

上海师范大学

硕士专业学位论文

基于 Informer 模型的期权交易策略 设计

学 院: 商学院

专业学位类别: 金融硕士

专 业 领 域: 金融

研 究 生 姓 名: 刘一凡

指 导 教 师: 傅毅

完 成 日 期: 2022. 4. 23

论文独创性声明

本论文是我个人在导师指导下进行的研究工作及取得的研究成果。论文中除了特别加以标注和致谢的地方外，不包含其他人或机构已经发表或撰写过的研究成果。其他同志对本研究的启发和所做的贡献均已在论文中做了明确的声明并表示了谢意。

作者签名：刘 - 凡 日期：2022 年 5 月 27 日

论文使用授权声明

本人完全了解上海师范大学有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其它手段保存论文。保密的论文在解密后遵守此规定。

作者签名：刘 - 凡 导师签名：傅敏 日期：2022 年 5 月 27 日

论文题目：基于 Informer 模型的期权交易策略设计

论文类型：交易策略设计

专业方向：金融数据分析

摘 要

2015 年 2 月 9 日, 随着上证 50ETF 期权的上市, 中国股市正式开启了 ETF 期权品种的交易市场。与早在 20 世纪 90 年代上市的股指期货相比, ETF 期权的标的资产 ETF 基金与中国 A 股股市整体之间的联系更为密切, 相比股票指数取得了更高的收益率, 并且 ETF 基金作为实物基金份额, 自身可以在二级市场进行上市交易, 定价效率高, 可以在一定程度上进行套利交易, ETF 期权还拥有相对较小的合约面值, 进一步提高了产品的流动性; 与个股期权相比, ETF 期权能够在相当大的程度上降低单个上市公司的非系统性风险, 在投资者分散化投资的基础上能够实现更优秀的对冲效果, 且 ETF 期权对证券交易的影响拥有更大的抗性, 市场价格的波动往往更小, 通过申购、赎回的方式间接实现 T+0 交易使得产品的组合更加灵活多变, 使得产品在套利方面拥有更大的优势。

ETF 期权产品拥有收益非对称的特点, 这使得它与波动率的交易之间建立了牢固的绑定关系, 在交易中, 波动率不再仅仅用来描述标的风险的大小, 而是更进一步被赋予了作为交易标的的性质。鉴于我国 ETF 期权市场发展至今仅仅 7 年, 相关市场规律与市场接受度仍在发展与提升的过程中, 与波动率相关的交易策略研究较少。本文通过对波动率进行预测, 提出了在相对较长的预测日期内的新的预测模型与因子选择, 并在此基础上制定了期权的交易策略, 进一步丰富了波动率预测与期权交易方面的研究。

本文先对采用了 Delta 中性策略的跨式期权的收益情况进行了理论层面的分析和阐述, 继而又通过建立隐含波动率预估模型、筛选合理的波动率预估因子等, 对波动率做出了相对精确的预测, 最后又在此基础上提出了以 Delta 中性策略为基础、以跨式 ETF 期权为标的的交易策略。文章最后的研究结果显示, 在波动率预测模型方面, 通过文章中构建的历史波动率因子、标的市场因子以及期权市场指标因子, 可以更加精确地对波动率的变动状况做出预估; 文章通过比较 OLS 模型、LSTM 模型和 Informer 模型这三类不同类型的时序预测模型的综合表现, 本文发现 Informer 模型在进行隐含波动率预测时拥有更高的准确率, 且随着预测日期的增加, 领先其他模型预测精度的程度也在增加。在期权交易策略构造方面, 本文综合考虑市场影响, 提出了三种不同维度下的交易信号, 并对多个参数进行优化, 总结出了最优的策略选择。从回测结果来看, 基于

Informer 模型预测的波动率结果构建出的期权交易策略以 11.65%的年化收益率、6.83%的最大回撤取得了最优的交易表现，证明了波动率预测模型的有效性与期权交易策略的稳健性。

关键词：隐含波动率；波动率预测；Delta 中性；期权定价

Abstract

On February 9, 2015, with the listing of Shanghai 50ETF options, the Chinese stock market officially opened the market for ETF options. Compared with the stock index options listed in the early 1990s, the underlying asset of the ETF options, the ETF Fund, has a closer relationship with the Chinese A-share stock market as a whole, and has achieved a higher yield rate than the stock index. As a real fund share, the ETF Fund itself can be listed and traded in the secondary market with high pricing efficiency and can carry out arbitrage transactions to a certain extent. The ETF options also have a relatively small contract value, further improving the liquidity of the products; Compared with individual stock options, ETF options can reduce the non-systematic risk of a single listed company to a considerable extent, and can achieve better hedging effect on the basis of diversified investment by investors. Moreover, ETF options have greater resistance to the impact of securities trading, and the fluctuation of market price is often smaller. The indirect realization of T+0 trading through subscription and redemption makes the product portfolio more flexible and changeable, and makes the product have greater advantages in arbitrage.

The ETF option product has the characteristic of asymmetric returns, which makes it establish a strong binding relationship with the transaction of volatility. In the transaction, volatility is no longer only used to describe the size of the underlying risk, but is further given the nature as the subject of the transaction. In view of the fact that China's ETF option market has only developed for 7 years, the relevant market rules and market acceptance are still in the process of development and improvement, and there is little research on the trading strategies related to volatility. Based on the forecast of volatility, this paper puts forward a new forecast model and factor selection in a relatively long forecast date, and formulates the trading strategy of options on this basis, which further enriches the research on volatility forecast and options trading.

In this paper, the returns of cross-type options with Delta-neutral strategy are analyzed and explained in theory. Then, the volatility is predicted relatively accurately by establishing implied volatility prediction model and selecting reasonable volatility prediction factors. Finally, a trading strategy based on Delta-neutral strategy and based on cross-type ETF options is proposed. The final research

results of the article show that, in terms of volatility forecast model, the historical volatility factor, underlying market factor and option market index factor constructed in the article can make more accurate forecast of volatility changes; By comparing the comprehensive performance of OLS model, LSTM model and Informer model which are three different types of time series forecasting models, this paper finds that the Informer model has higher accuracy when forecasting the implied volatility, and with the increase of forecast date, the degree of prediction accuracy ahead of other models is also increasing. In the aspect of option trading strategy construction, this paper comprehensively considers the market influence, puts forward the trading signals under three different dimensions, optimizes several parameters, and summarizes the optimal strategy selection. From the back test results, the option trading strategy based on the volatility results predicted by the Informer model achieves the optimal trading performance with an annualized return rate of 11.65% and a maximum retracement rate of 6.83%, which proves the effectiveness of the volatility prediction model and the robustness of the option trading strategy.

Keywords: Implied volatility; Volatility forecast; Delta neutral; Option pricing

目录

摘要	I
Abstract	III
第 1 章 绪论	1
1.1 研究的背景	1
1.2 研究的目的和意义	2
1.2.1 研究目的	2
1.2.2 研究意义	2
1.3 研究的内容、方法和技术路线	3
1.3.1 研究内容	3
1.3.2 研究方法	3
1.3.3 技术路线	4
1.4 本文的主要特点	5
第 2 章 相关理论回顾与文献综述	7
2.1 相关理论回顾	7
2.1.1 波动率相关理论	7
2.1.2 时间序列预测相关理论	8
2.1.3 期权交易相关理论	12
2.2 相关文献综述	14
2.2.1 有关波动率度量及预测模型的研究	14
2.2.2 有关序列预测的研究	15
2.2.3 有关期权及其交易的研究	15
2.2.4 相关文献述评	15
第 3 章 基于 Informer 模型的期权交易策略构思	20
3.1 波动率与期权价格	20
3.1.1 波动率预测方法的选择	20
3.1.2 期权交易策略的构思	20
3.2 波动率预测及交易策略的设计思路	21
3.2.1 波动率预测的设计思路	21
3.2.2 期权交易策略的设计思路	21
3.3 模型构建及策略设计的理论框架	23
3.3.1 SHAP 因子筛选原理	23

3.3.2 自注意力机制改进原理	24
3.3.3 期权 Delta 中性交易理论	24
3.4 模型的有效性评价	26
3.4.1 波动率预测评价方法	26
3.4.2 期权交易策略评价方法	26
第 4 章 基于波动率预测下的期权 Delta 中性交易策略设计方案	28
4.1 数据说明与变量说明	28
4.1.1 数据来源	28
4.1.2 预测因子选择及描述性分析	28
4.2 波动率预测模型构建	31
4.2.1 最小二乘法滚动预测	31
4.2.2 LSTM 模型预测	31
4.2.3 Informer 模型预测	31
4.3 策略实施	32
第 5 章 交易策略设计方案的有效性评价	34
5.1 交易策略方案的有效性评价	34
5.1.1 波动率预测效果评价	34
5.1.2 期权交易结果评价	36
5.2 交易策略方案的风险提示	38
第 6 章 结论	40
6.1 研究结论	40
6.2 文章的反思和改进方向	40
参考文献	42
致谢	46

第 1 章 绪论

1.1 研究的背景

作为金融市场的重要组成部分，期权市场及其产品能够更好地管理潜在风险，更合理地配置资产，且拥有更为合理的价格发现机制，不仅使得投资者在投资与资产管理方面有了更加精细、便于变动的选择，同时也丰富了市场的交易结构，提高了市场总体有效性。上证 50ETF 期权已于 2015 年 2 月 9 日挂牌交易至今，规模一直高速发展，2020 年的日均成交额达 13.12 亿，仅次于美国标普 500ETF 期权。2019 年底上市的沪深 300 股指期货期权两年来也处于快速扩张的势头，体现了我国期权市场正处于加速上升期，与期权交易相关的探讨与研究都有其必然性。

在“大资产管理”趋势的影响下，量化交易在资产管理中的运用方兴未艾，其中，波动率交易策略的壮大与期权市场的发展密不可分。波动率交易策略集中关注标的产品价格的波动与期权隐含波动率的变化情况，淡化方向性，强调变化程度，并以此为基础制定相应的交易策略。由产品的特质决定，对波动率的变化预测难度要小于对标的价格的预测，波动率交易策略往往也有相对更优的表现。目前对中国 ETF 期权市场波动率交易策略的研究还远没有系统化，没有进行充足的策略参数的比较分析，加深该方向的研究可以为期权市场波动率交易的研究提供广泛的指导。通过分析 50ETF 期权市场的特定情况，本文首先探索了理论上的波动率交易策略的收益来源，并比较如何在三种不同类型的时序预测模型中确定交易 50ETF 期权市场波动率的机会。最后，基于上证 50ETF 期权市场的波动率交易策略从理论与实践与市场实际情况的综合视角研究，比较了市场不同参数的决策结果，预测了波动率的变动，并据此研究了怎样把握好在国内期权市场进行交易的适当时机，具有重要的现实意义。

近年来，国内外学者针对波动率预测的建模方式开展了大量研究。一类方式为利用时间序列法，也就是通过对历史波动率变化从时间序列的层次上加以分析，从而找出历史波动率演变的特征变化规律，并由此对波动率做出预测，主要方式包括 GARCH 族模型法、指数加权移动平均法以及高频波动率模型法等；另一类方式为利用隐含波动率，由于在隐含波动率中蕴涵了期权的市场信号，因而能够用来反映期权价值的变动，主要涉及在 B-S 等各公式定价下的隐含波动率变化等。上述波动率预测方式利用整合了各种信号，来预测波动率的变动状况。本文通过预测波动率的变动状况及其幅度变化，在此基础上建立了

波动率的交易策略。本文在多因子预测模型的基础上，探索了预测准确度较高的模型，并引入了波动率锥与市场趋势等信号，从而形成了适应中国国内金融市场、符合市场交易规律的波动率交易策略，使中小投资者可以更好地运用期权交易实现资产的合理配置，对于不断加强我国多层次金融市场的稳定性有着重大意义，从而促进中国多层次期权市场结构的构建和中国金融市场的多维度发展。

1.2 研究的目的和意义

1.2.1 研究目的

本文试图利用我国证券市场中长期记忆性的存在以及波动率对期权定价的重要影响，通过建立一种基于 transformer 模型的经过改进的 Informer 模型来实现对波动率的预测，并与其他传统波动率预测模型的结果进行对比，最后输入不同参数进行不同条件下的期权交易策略回测，比较得出各预测模型下的最优收益，进而指导期权的交易方案。针对传统波动率预测模型在中长期预测方面能力的欠缺，本文通过 Informer 模型与传统波动率预测模型的预测效果对比，证明 Informer 模型在时间序列预测中的优越性。鉴于中国期权市场于 2015 年起步，时间较晚，本文旨在丰富我国在期权交易方面的研究。

1.2.2 研究意义

虽然现有文献研究已证实，我国证券市场具有长期记忆性，同时由于信息流动不畅通、市场交易机制尚不健全，从而确定了中国的期货交易市场有效性较弱，但受限于传统时间序列模型的性能，现有研究成果大都集中于定性指标分析方法，而量化分析方法则较少，因而有效的上证 50ETF 期权交易策略较少。期权交易策略的提出，将丰富中国股市现有的套期保值手段，减轻投资者的主观情绪对于交易风险的影响，帮助投资者对波动率有更清晰的认识，降低投资风险。

Informer 模型是一种类 Transformer 模型，相比 transformer 模型具有更好的长序列预测的能力，自 2017 年 transformer 模型问世以来，该模型实践相当广泛，特别是在变压器故障识别和文本翻译领域，但是在金融领域的应用实践则较为少见。之前有人在信贷贷款预测方面应用到该算法，应用效果较好，较为准确地实现了不良贷款的识别与分类工作。而在股票市场的应用研究还在开始阶段，暂时没有可供参考的相关文献。本文可能具有的理论意义如下：

第一，现有研究主要集中在期权行权日前的若干天进行短期预测，本文选择运用上证 ETF 期权价格进行 5 个交易日至 20 个交易日的波动率预测，丰富了我国期权交易中长序列预测的研究；

第二，Informer 模型自 2020 年 12 月提出以来，尚未得到广泛运用，本文将对 Informer 模型在长序列预测领域中的效果进行检验，并探究其与金融领域的适配性；

第三，本文探究了在 5 个交易日至 20 个交易日下各市场参数的影响情况，其结果与短期预测的对比对市场风格变动相关研究具有一定的参考意义。

1.3 研究的内容、方法和技术路线

1.3.1 研究内容

在中国股市存在长期记忆性的前提下，更长时间序列数据的输入意味着更多的相关信息与更加准确的预测结果，同时也代表长序列预测具有其可行性与实际意义。但传统的 GARCH 类波动率预测模型与 HAR 模型只能在短期预测时保持一定准确度，在中长期时间序列预测上效果较差。为了更精准地刻画波动率在长期的趋势，更好地为制定期权交易方案服务，本文拟利用 Informer 模型在长序列预测方面的优势，进行期权长期波动率的预测。

为进一步完整地体现上证 50ETF 期权波动率的变化，同时兼顾数据有效性，本文选择日波动率数据作为输入变量，并对未来一个月的时间内波动率的变化情况进行预测，输出序列长度的增加限制了传统波动率预测模型的效果。因此本文拟利用 Informer 模型能够有效处理中长序列输出问题的优势，对波动率进行精准预测，并与传统波动率预测模型的效果进行比较。

(1) 中长期波动率的预测。在波动率预测问题中，传统波动率预测模型难以分析大规模数据且只能进行短序列输出，本文选择使用基于 transformer 模型改进的 Informer 模型进行中长期波动率的预测，旨在提高预测的准确度，得到更好的波动率预测结果。

(2) 基于波动率预测的期权交易策略。本文借助波动率的预测结果进行期权交易策略的制定，并在此基础上比较各窗口期策略收益的高低，进而制定更合理的期权交易方案。

1.3.2 研究方法

(1) 文献搜集法

本文对国内外现有的有关波动率预测理论与期权交易策略的有关研究成果进行了一定程度地总结和整理并做出了综合评价，归纳总结了已获得的重要成果以及还未能解决的有关问题，为文章今后的进一步深入研究奠定了理论基础。

（2）模型结合法

为进行中长期波动率预测，本文选择使用性能优异的 **Informer** 模型来实现。和其他传统波动率预测模型相比，**Informer** 模型不仅能并行化处理更大规模的数据，在模型记忆长度上也有大幅度的提高。**Informer** 模型的引入，能有效解决传统模型中长期预测失真的问题。

（3）对比分析法

为证明 **Informer** 模型在序列预测方向的优势，本文分别使用 **Informer** 模型和传统波动率预测模型进行 5 个交易日、10 个交易日和 20 天个交易日的波动率预测，时间跨度从一周到以一个月；同时，探究在不同窗口期下的不同参数值的表现，进而设计最优的期权交易策略。

（4）实证分析法

利用上证 50ETF 期权 500 个交易日的收益率数据，经过对数变换后代入 **Informer** 模型中进行实证分析，将其结果与其他模型相比较来体现本文的优越性，并进一步将预测结果用于期权交易策略收益的比较和期权交易方案的制定上。

1.3.3 技术路线

按照文章的主题的研究内涵、研究方式及其对变量的设计度量，可以形成文章的技术路线图，如图所示：

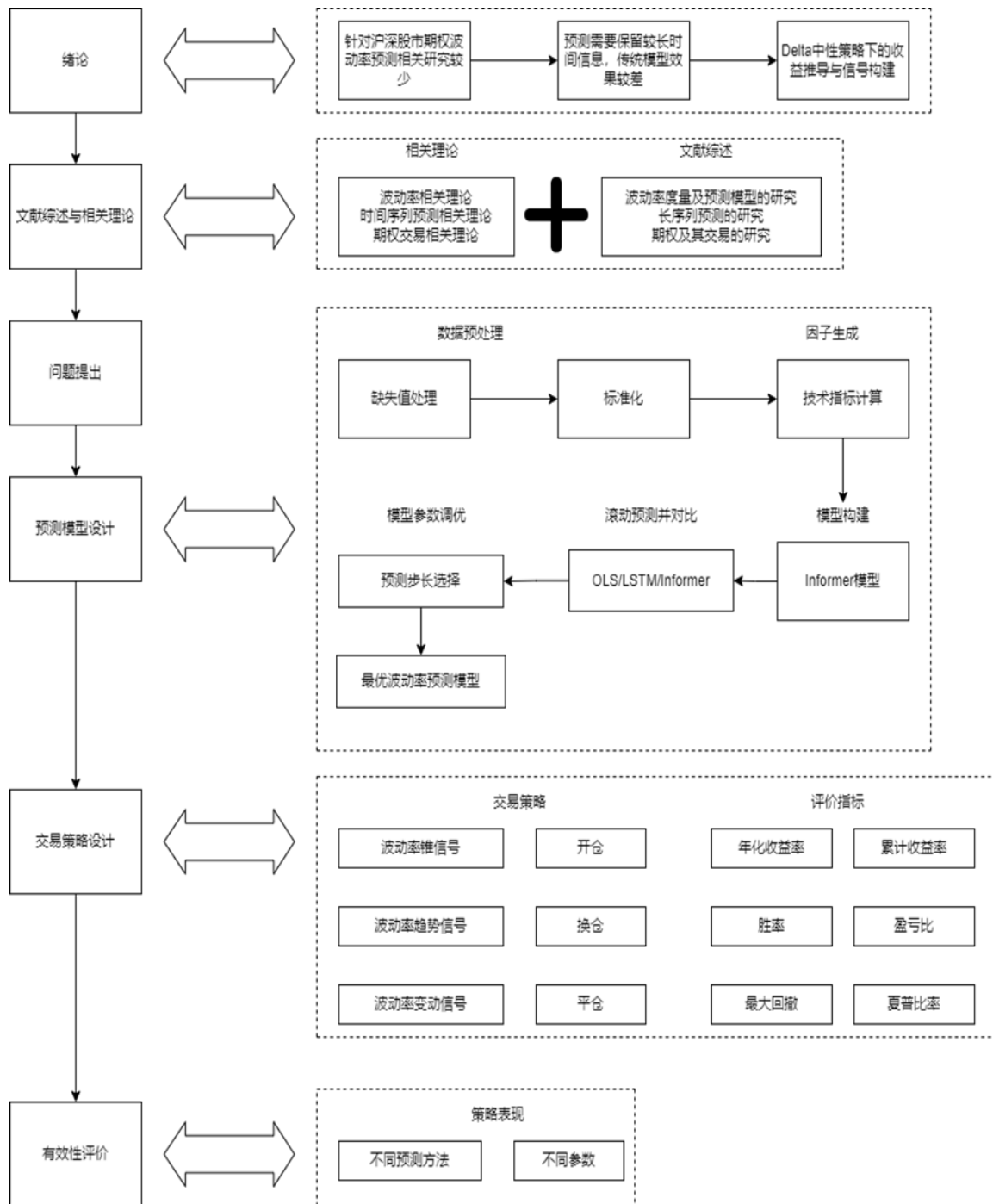


图 1-1 技术路线图

1.4 本文的主要特点

理论研究方面：

尽管现有研究证明中国股市因为市场不够成熟而存在明显的长期记忆性，但受限于长序列数据的处理问题，进行中长期数据的预测较为困难。运用

Informer 模型在中长序列输入和输出的预测结果，证明中长序列预测在中国股市的有效性。

未经简化的 **attention** 计算量高达 $O(L^2)$ ，使得模型每层时间复杂度和内存使用量较高，限制了输入数据的上限。本文拟通过保留 **self-attention** 中若干个较为显著的 **query** 值来达到简化模型的目的。

实证研究方面：

已有研究已经发现，中国期权市场往往具有不同于其他发达资本市场的运作规律，这可能是因为中国期权市场结构还不够健全导致的，在制定期权交易方案的同时，找到中国期权市场特有的规律也同样重要。通过历史数据预测未来波动率，并在此基础上进一步比较获得最优期权交易策略，为指导期权交易方案的制定提供思路。

第 2 章 相关理论回顾与文献综述

2.1 相关理论回顾

2.1.1 波动率相关理论

作为一种统计定义，波动率是指一定时期内资产收益率的年化标准差。波动率作为判别证券市场起伏和投资风险的关键性指标，是证券实证研究中不可忽视的变量。高波动率意味着证券资产价格波动频繁，证券产品的收益风险更大。同样，低波动率意味着证券资产的价格稳定波动，证券产品的收入更具确定性。波动率的预测是证券研究中的一项关键性任务，在过去的 20 年里，业内大量学者为了达到提高证券资产收益率、降低投资风险的投资目标，纷纷通过模型对波动率进行预测。

根据已有研究得出的结论，波动率往往具有如下 4 个特征：

（1）波动率聚集效应：波动率的变化通常具有一定程度的自相关性，即高波动率之间具有一定的连续性，反之亦然；

（2）波动率在黑天鹅事件等发生、对市场情绪造成重大冲击时，会出现大幅度地跳跃；

（3）均值回归效应：波动率在长期的变化过程中倾向于稳定在一个特定的值；

（4）非对称性效应：标的产品价格的上涨与下跌往往对波动率的变化会产生不同程度的影响。

根据不同的含义与对应的测量方式，波动率可分为以下 4 种类型：

1. 历史波动率

历史波动率是历史区间资产历史价格在统计学原理下测算的变动情况。历史波动率的数值一方面可以反映金融资产在统计时期的波动，另一方面也作为基石支撑其他波动率的预测。其测算步骤以文字概括为：首先获取固定时间间隔（例如每周或每月等）的市场资产价格；接下来计算该时间段结束时资产价格变化比率的自然对数；最后用这些对数值的标准乘以一年中包含的时间段数量的平方根。

其中： S_i 是第 i 时刻资产历史交易价格， u_i 是第 i 时刻资产相对前一时刻的对数收益率， m 是观察值数量， σ_n 代表在第 n 天观察过去 m 个对数收益的标准差。

$$u_i = \ln \frac{S_i}{S_{i-1}} \quad (2-1)$$

$$\sigma_n^2 = \frac{1}{m-1} \sum_{i=1}^m (u_i - \bar{u})^2 \quad (2-2)$$

$$\bar{u} = \frac{\sum_{i=1}^m u_{n-i}}{m} \quad (2-3)$$

2. 隐含波动率

隐含波动率是一种由期权定价中引申出的概念。按照期权定价原理，直接影响期权投资实际价值的主要因素有如下五个：标的资产实际价值，到期日期，波动率，无风险收益率以及执行价值。其中，波动率是一个不能直接观测到的变数，而期权投资实际价值又是能够观察到的，所以当期权投资的实际价值能够引入期权定价公式，我们就能够推导出波动率的数值，也便是隐含波动率的数值。它作为由期权的市场价值所决定的波动率，是金融市场内在价值的实际体现。而有效的金融市场内在价值既是供需平衡的成果，更是买卖双方博弈的成果。所以，隐含波动率也体现了金融市场上对标的资产市场价格未来波动情况的预测。

3. 未来实际波动率

未来实际波动率，也称为已实现波动率，即测量金融资产在未来一段时间内收益的波动程度，因为实际收益是一种随机过程，所以未来实际波动率始终是一种未知数。这就是说，实际波动值是无法事先准确测算的，人们可以采用各种方法求得对它的估计值。

4. 预期波动率

预期波动率，利用统计学方法和波动性率模型方法对市场未来的真实波动率做出预估，适用于期货计价和期权套利策略。典型的波动率模型还有 GARCH 族模型(广义自回归条件异方差模型)、EWMA 模型(指标加权偏移均值模型)、SV 模型(随即波动率模型)、高频统计波动率模型等。

2.1.2 时间序列预测相关理论

关于时间序列预测问题的研究一直延续了百年之久，因如此长期以来，不同的时间预测方式与预测模型都被提倡并运用到了人们日常的生产生活当中，并且这些方式都获得了数量巨大的研究成果。而传统的时间预测方式，一般都是根据关于时间序列的先验分布知识和统计的有关理论研究发展而来，但这种

方式更多应用于那些结构较为简单的一维的、分布也不太复杂的时间数据预测任务中。近年来，学习的发展及其它技术在自然语言处理、计算机视觉等应用领域的成熟运用，对时间序列预测理论产生了巨大的影响，时间序列预测应用领域的研究重心也逐步转为利用学习技术构建时间序列预测模型，而此类模型一般都是采用时间数据为驱动的，模型方式可以直接通过神经网络来掌握时间数据的分布特性，从而解决了传统时间预测方式中模型选取较难的问题，因而泛用性更强。近年来，各种基于神经网络的时间序列预测方式也陆续被提及，并获得了较为良好的预期效益。

根据每次预测未来序列的时间长度，将时间序列估计问题可分成单步式预测与多步预测，单步式预测利用时间序列的历史数据来预测未来每一个时间步骤的数量，而多步预测则能够预计更多的时间步骤。在实际运用中，多步预测是比较实际的。

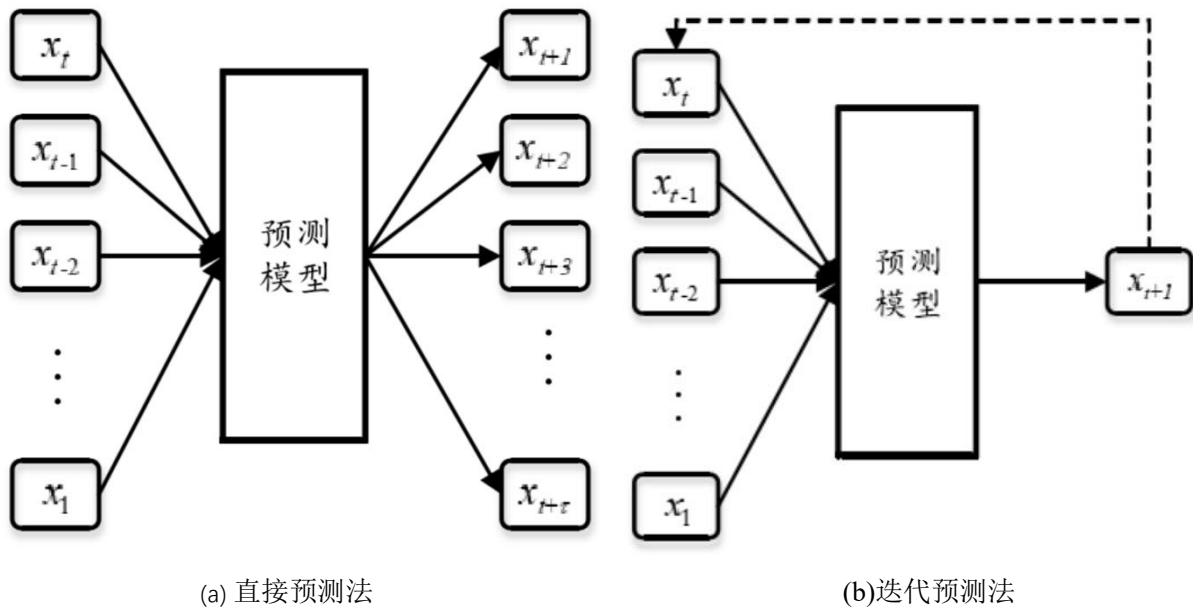


图 2-1 多步预测法的两种实现方式

图 2-1 中分别说明了多步预测法的 2 种方法，其中图 a 利用前 t 个连续时间的时序信息来预报未来总共 t 时间步长度的顺序信息，图 b 则利用前 t 个时点的时间信息来预测 $t+1$ 时点的时间信息。这二种方法都有利有弊。直接时间预测法通过多输入多计算的方法可以直接得出 t 的未来数值，但是这个方法由于要求掌握多对多的时间映射关系通常要求大量的数据处理，并且根据预计时间长短确定，但是如果要提高预计的时间长短，则需要重新训练网络。迭代预测法（在文中也被称之为自回归策略的预测模型）比预计的持续时间长短更灵活性，因为只需在一个数据集中训练好模型之后就可以胜任不同长度的时间序列预计任务，而且迭代预测法所需要的训练数量也比直接预测法要小许多，不过因为在预计时候就必须把网络的输出当做下次的输入，所以会产生误差累积现象。

一种减少误差积累的好方法就是提高网络系统的输入输出维度，但如果在估计流程中采用了网络系统的输入输出值，这个网络系统输入就会形成误差积累。图 2-1(b)是一种单步预测的迭代模型，受自然语言处理中序列建模的影响，迭代模型也可能是多输入输出的，所以这个建模也被叫做编码器-解码器序列建模。

(1) 最小二乘法

最小单位二乘法是一个多元分析模型，由 Wold^[1]在 1983 年首先提出，该方式综合了主成分解析，典型相关性分析和多元线性返回于一身，当其线性回归结果中出现多变数和多重共线性时仍具有着较好的表现力。假设在原始数据中的自变量 X 为 $N \times l$ 维矩阵，因变量 Y 为 $N \times m$ 维矩阵，也即自变量样本特征为 l 维，因变量样本特征为 m 维，样本量为 N 。偏最小二乘方法的建模原理为：

步骤 1：求出 X 与 Y 的第一对主成分 t_1 与 u_1 ，其中 $t_1 = X \times w_1$ ， $u_1 = Y \times c_1$ 。存在 t_1 和 c_1 使得 $COV(t_1, u_1)$ 最大，即：

$$\begin{aligned} & \text{Maximize: } COV(t_1, u_1) \\ & \text{subject to: } \|w_1\| = 1, \|c_1\| = 1 \end{aligned} \quad (2-4)$$

再建立下述回归方程，分别得到回归方程的系数和残差：

$$X = t_1 p_1 + X_1 \quad (2-5)$$

$$Y = u_1 q_1 + Y_1 \quad (2-6)$$

$$Y = t_1 r_1 + F \quad (2-7)$$

其中 X_1 为 X 对其第一主成分回归的残差， Y_1 为 Y 对其第一主成分回归的残差， F 为 Y 对 X 第一主成分回归的残差。

步骤 2：

将 X_1 与 Y_1 作为新的 X 与 Y ，步骤 1 循环进行回归，直到残差 F 达到精度要求，或者主成分数量达到上限。假设共有 k 个主成分，最后可将因变量 Y 表示为：

$$Y = t_1 r_1 + t_2 r_2 + \dots t_k r_k + F \quad (2-8)$$

(2) LSTM 模型

递归神经网络，是时序递归神经网络和结构递归神经网络的缩写，而 LSTM 神经网络也隶属于时序递归神经网络。通过对循环神经网络的优化，LSTM 神经网络可以克服传统循环神经网络的梯度问题，该算法的基本思想是摒弃原有的传统循环神经网络中的隐含层的神经元，而用较为特殊的 LSTM 的网络结构来代替。LSTM 神经网络包含存储器块(记忆单元 Cell)和三个自适应乘

法门，即输入门(Input Gate)、输出门(Output Gate)和遗忘门(Forget Gate)。一般来说，它并非直接把自己的输出值传给其他神经元，而只是负责在神经网络的某些部分或和记忆单元相连的边缘部设置权值，以作为选择性将记忆反馈的误差函数随梯度而减小的校正参考。LSTM 内部处理器完整架构的三重门具体构造如图 3-1 所示：

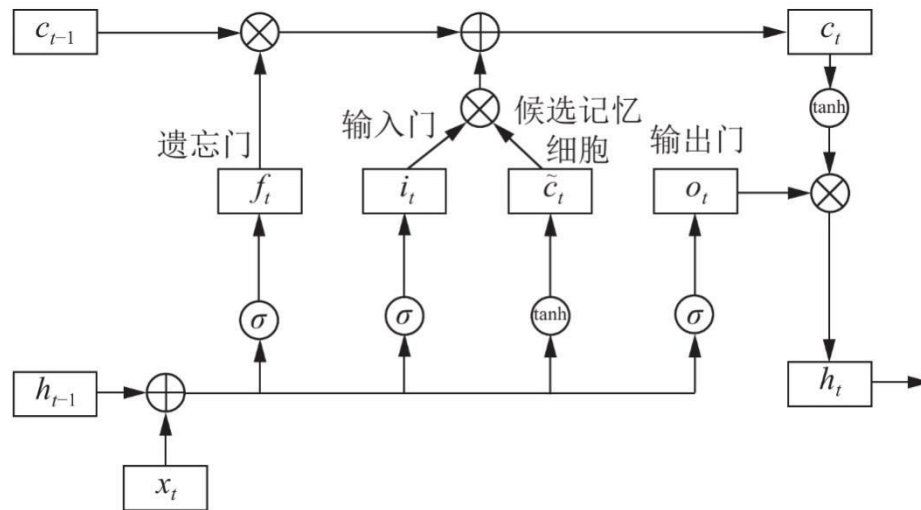


图 3-1 LSTM 三重门结构图

其中，Input Gate 即为输入门，Output Gate 即为输出门，在整个网络结构中用于接受信息以及输出、修正参数，分别为 I, ω 。Forget Gate 即遗忘门具有筛选作用，进行保留或者遗忘操作，即为 ϕ 。Cell 即存储器，代表神经元处于某种情境下的记忆，记 S_c^t ，可以通过这三个门以及存储器实现信息的存储与控制。

(3) Informer 模型

图 3-2 是 Informer 模型整体框架图。图中左边的结构是编码器 Encoder，其能接受长序列输入（绿色序列）。Encoder 采用了自注意力蒸馏机制（蓝色梯形），通过将层级输入减半，来提取主导的注意力，并大幅减少网络规模；同时我们在 Encoder 中用 Prob Sparse self-attention 替换掉了原始的 self-attention，提高长序列处理的性能。图中右边的结构是解码器 Decoder，其同样能接受长序列输入，并将需要预测的目标元素填充为零，通过将需要预测的该全零序列也作为特征图的加权注意力组成部分，用生成式的方式对预测序列（橙色序列）进行预测。

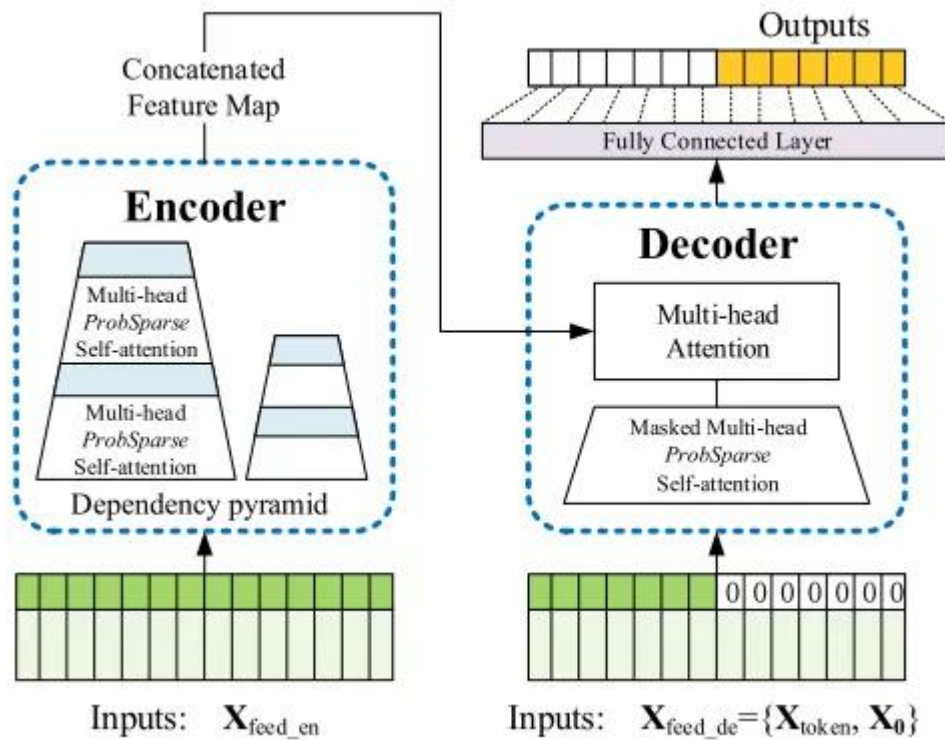


图 3-2 Informer 模型原理

来源: <https://arxiv.org/abs/2012.07436>

Informer 具有三个独特的特点: (1) Prob Sparse Self-attention 机制, 在时间复杂度和内存开销上实现了, 并且在序列依赖关系对齐上具有相当的性能; (2) 自注意力蒸馏机制, 通过将层级输入减半来突出主导注意力, 并高效处理极端的输入序列; (3) 生成式解码器, 虽然其在概念上简单, 但它能够以一次前向操作而不是分步操作的方式预测长时序序列, 极大地提高了长序列预测的推理速度。

2.1.3 期权交易相关理论

Bakshi 和 Kapadia^[2]两位学者, 在他们的研究成果中提供了持有期权和实施动态 Delta 对冲策略的收益分析方法。文中的期权策略主要是指同样拥有并出售看涨期权与看跌期权, 并对这种期货组合实施动态 Delta 对冲, 由于实质上两位学者的研究很相似, 所以在对 Delta 中性波动率决策的收益理论解析中, 本文主要参考了这两位学者的研究, 并在此基础上推论出了跨式化期货动态对冲风险的主要利润来源。首先假定了标的价值变动过程遵循几何布朗运动:

$$\frac{dS_t}{S_t} = u_t dt + \sigma_t dz_t \quad (2-9)$$

$$d\sigma_t = \theta_t dt + \eta_t dW_t \quad (2-10)$$

其中 z_t 和 W_t 为标准的维纳过程，两者之间的相关性系数为 ρ ，参数 u_t 为价格变化过程的瞬时漂移项，参数 σ_t 为瞬时波动率服从独立的随机过程， θ_t 为其漂移项， η_t 为其扩散项，两者均独立于 S_t 。假设跨式期权的价格为 f_t ，通过伊藤引理可得：

$$f_{t+\tau} = f_t + \int_t^{t+\tau} \frac{\partial f_u}{\partial S_u} dS_u + \int_t^{t+\tau} \frac{\partial f_u}{\partial \sigma_u} d\sigma_u + \int_t^{t+\tau} b_u du \quad (2-11)$$

其中

$$b_u \equiv \frac{\partial f_u}{\partial u} + \frac{1}{2} \sigma_u^2 S_u^2 \frac{\partial^2 f_u}{\partial S_u^2} + \frac{1}{2} \eta_u^2 \frac{\partial^2 f_u}{\partial \sigma_u^2} + \rho \eta_u \sigma_u S_u \frac{\partial^2 f_u}{\partial S_u \partial \sigma_u}$$

同时基于无套利定价，跨式期权的价格 f_t 必须满足下列公式：

$$\begin{aligned} & \frac{1}{2} \sigma_t^2 S_t^2 \frac{\partial^2 f_t}{\partial S_t^2} + \frac{1}{2} \eta_t^2 \frac{\partial^2 f_t}{\partial \sigma_t^2} + \rho \eta_t \sigma_t S_t \frac{\partial^2 f_t}{\partial S_t \partial \sigma_t} \\ & + r S_t \frac{\partial f_t}{\partial S_t} + (\theta_t - \lambda_t) \frac{\partial f_t}{\partial \sigma_t} + \frac{\partial f_t}{\partial t} - r f_t = 0 \end{aligned} \quad (2-12)$$

其中 r 为无风险利率， λ_t 表示波动率风险溢价。

整理公式（2-6）和公式（2-7）可得：

$$\begin{aligned} f_{t+\tau} = f_t & + \int_t^{t+\tau} \frac{\partial f_u}{\partial S_u} dS_u + \int_t^{t+\tau} \frac{\partial f_u}{\partial \sigma_u} d\sigma_u \\ & + \int_t^{t+\tau} \left(r \left(f_u - S_u \frac{\partial f_u}{\partial S_u} \right) - (\theta_u - \lambda_u) \frac{\partial f_u}{\partial \sigma_u} \right) du \end{aligned} \quad (2-13)$$

再将公式（2-5）代入公式（2-8）中再将 $d\sigma_u$ 进行化简可得：

$$\begin{aligned} f_{t+\tau} = f_t & + \int_t^{t+\tau} \frac{\partial f_u}{\partial S_u} dS_u + \int_t^{t+\tau} r \left(f_u - S_u \frac{\partial f_u}{\partial S_u} \right) du \\ & + \int_t^{t+\tau} \lambda_u \frac{\partial f_u}{\partial \sigma_u} d\sigma_u + \int_t^{t+\tau} \eta_u \frac{\partial f_u}{\partial \sigma_u} dW_u \end{aligned} \quad (2-9)$$

将公式（2-9）进行化简，由此可得 Delta 中性跨式期权的收益形式：

$$\pi_{t,t+\tau} = \int_t^{t+\tau} \lambda_u \frac{\partial f_u}{\partial \sigma_u} du + \int_t^{t+\tau} \eta_u \frac{\partial f_u}{\partial \sigma_u} dW_u \quad (2-14)$$

其中 λ_u 表示波动率风险溢价， η_u 为标的资产波动率服从随机分布的扩散项，将公式（2-10）进行简化和求导可以得到：

$$d\pi = \frac{\partial f_u}{\partial \sigma_u} (\lambda_u du + \eta_u dW_u) \quad (2-15)$$

$$d\pi = \frac{\partial f_u}{\partial \sigma_u} ((\lambda_u - \theta_u) du + d\sigma_u) \quad (2-16)$$

也因此,拥有跨式期权和实施动态 Delta 对冲的收益可被划分为二个组成部分:波动率风险溢价和波动率变化。在式(2-12)中,从动态的角度来看, $(\lambda_u - \theta_u)$ 在较长时间中是一种符号不变化的常量值,在短期动态调整时,策略的收益会受到更多来自波动率变化的影响,在对预期波动率改变存在足够准确度的前提条件下,可以透过设定波动率的变化阈值,来抵消 $(\lambda_u - \theta_u)$ 的影响,能够获取正收益。

2.2 相关文献综述

2.2.1 有关波动率度量及预测模型的研究

自从 ARCH 和 GARCH 模型提出后,许多学者在预测金融时间序列波动方面展开了丰富的研究。郑振龙^[3]等在对香港恒生指数期货场所包含资讯的研究中首先发现,在估计未来波动率时,如果预期时限较短(一周),GARCH(1,1)的模型所包含资讯较多,预知力亦强,但如果预期时间较长期限(一个月),隐含波动率的所需信息较多,且预测能力较好。屈满学^[4]等的实验结果显示,在预计未来一周的证券市场风险情况方面,由上海市证券交易所编制的国内波动率指标(iVIX)的预期能力大于历史上已实现波动率和 GARCH 族波动率,但可能是因为国内期权交易市场并不是完全的自由交易市场,其预期能力并不如发达国家地区有效。因此马天平^[5]等以国内波指为研究对象,建立了基于机器学习算法的“SKRG 递进集成”新预警系统。该系统共选择了四十八个相关指数作为对中国波指预测的主要特征,结果显示基于机器学习的预测效果好于传统的 GARCH 模型。龚旭^[6]等人建立了 HAR-RV-J-SB、HAR-S-RV-J-SB 以及 PSlev-J-SB 模型,以美国 WTI 原油期货的五分钟高频成交数量为实验样品,展开了定量分析。实证结果表明:原油期货市场上出现了显著的跳跃风险和价格结构突变迹象。而相比于 HAR-RV、HAR-S-RV 和 PSlev 的模型,HAR-RV-J-SB、HAR-S-RV-J-SB 和 PSlev-J-SB 模型中对原油期货价格波动的除样本外预测准确度均明显较高且结果也比较稳定。Aparna^[7]考察了交易所交易的四种不同规格的美元-卢比期权隐含波动率微笑,发现历史波动性、汇率的动量和跳跃、到期时间、期权交易量和股票市场的波动性是导致波动率微笑的成因。王茹婷^[8]等人利用 HAR-RV 事件的扩散模型和日内跳跃 Logistic 模型进行的数据分析研究表明,中美国际贸易碰撞事故会对受到加增关税影响的行业指标、中国企业沪深三百指标等迅速形成短暂的冲击。此外,相比较于中国制裁公布,美国制裁公布后对国内金融市场的影响更大,引入国际贸易摩擦事件会提高市场波动度预测模型样本内外的预期效应。吴鑫育^[9]等以国际标准已完成

的 GARCH 模型为基准, 将条件方差贡献率乘性划分为长期方差、短时方差二部分, 并相对应的建立出了包含杠杆函数的长期方差方程与短时方差方程, 结果分析说明二指标的双因子准则已完成 GARCH 模型较国际标准已完成 GARCH 模型在样本内表达的似然函数估算价值更大。

除了传统的计量模型外, 近些年来, 随着文本分析能力的发展, 一些学者开始尝试将有关投资者情绪相关指标纳入到模型中去。李鸿翔^[10]等将中国波指 (iVIX) 用作影响投资者心态的代理变量, 研究结果显示, 上证五十指数和 iVIX 之间的关系长期呈负向, 并伴随着上证五十指数的回落, iVIX 会有明显的提高, 极端的股票价格暴跌和风险的急速增加有关, 投资人心态波动更大, 风险的非对称性特征较为明显。刘勇^[11]等将我国的上证五十 ETF 期权的隐含消息和投资者情绪导入了异质性自回归模型(HAR-RV), 实验结果表明, 期权隐含消息和投资者情绪将有助于改善 HAR-RV 模型对中国证券市场未来震荡率变化的预期效应。张同辉^[12]等在 ARMA 类和 HAR 类的模型中引入了投资者关系因子, 构建了全新的投资人关注度波动性率预估模型, 并发现了一方面投资人关注度低能够改善当前波动率预测模型的样本内拟合力量, 且当投资人关注度较高时, 波动率模型的样品外预期力量也都可以获得明显稳定的改善。Zhang W^[13]等利用文本分析技术构建了四个投资者情绪指数, 发现新兴市场的投资者情绪对股市波动的影响是非线性的, 同时改善了基于 LSTM 模型的股票波动预测。Lei B^[14]等通过前置关键词百度搜索指数构建投资者关注因子, 发现具有投资者关注的 TCN 模型的预测精度优于传统计量模型与具有投资者关注的 LSTM 模型。相较于传统计量模型相比, TCN 模型的多步预测结果依然稳健。

2.2.2 有关序列预测的研究

近年来, 人们研究长期记忆性问题的方法, 大致包括以下三种:

一是非参数方法, Mandelbrot^[15]首次把由 Hurst^[16]所提供的 r/S 分析方法, 运用于对经济历史时间顺序的长期记忆科学研究工作中。此后, 随着 R/S 分析技术在金融时间序列中获得了应用, 研究者们也着手深入探讨该方式检测长期记忆的稳定性。经典的 R/S 分析并不能区别短线关联性与长久依赖性, 为同样兼顾了短线关联性与长久关联性, Lo 给出了经过调整的 R/S 计算量, 并表明了美国股票的日收益率与年回报率之间都不存在长期记忆性。GiRaitis^[17]等人基于 KPSS 统计分析量和重标极差(R/S)方式, 给出一个新的检测序列长期记忆性的方法-重标方差和 V/S (ReScaled Variance)方式, 他既在理论上严格证实了 V/S 统计分析量存在着更强的长期稳定性, 也通过蒙特卡洛模拟实验比较了经过改进的 r/S 、KPSS 和 V/S 方式对于检测序列长时间记住的长期稳定性, 实验结果显

示: v/s 统计分析量比较平稳。他们也已经说明了, V/S 统计量在 q 的选择则上比修正的 R/S 、 $KPSS$ 等方法比较稳定, 但仍然比较敏感。

二是零点五参数方法, 由 Geweke 和 Porter-Hudak^[18]共同给出了对数周期性图(GPH)的方案, 对数周期性图是通过使用傅里叶变换把时刻顺序从时域映射变换到频率区间, 从而得出时刻顺序的周期性图, 之后再通过最小二乘计算对数周期性图回归微分方程的斜率。Robinson^[19-22]把对数时间图引入到了多元时间的序列模型中, 并放宽了假定: 谱密度矩阵可以趋向于无限和零, 对比 GPH 方式来说, 是更高效的一个计算方式。Robinson 首先给出了平均周期图法, 其思路是把平均周期图视为对谱分布函数的估估量, 从而可以根据谱分布函数的近似形状得出长期记忆参数的估测器。该方案进一步放松了 GPH 的假设前提, 并确认了所有建立在平均周期图上的估测器都是连续的。但是, 平均周期曲线对新引入的参数 q 的选择非常敏感, 因此 Lobato 和 Robinson 证实了当二分之一 Robinson 首先提出了局部 Whittle 分析方法, 该分析方法是对频域上高斯似然形式的一种最大化值近似, 可以通过求得局部 Whittle 目标函数的最小值来估计参数 H 或 $dd=H-1/2$ 。尽管该方式在运算上较前述方式更繁复, 但在运算时并没有引用任何主观变数如在平均时间图中引用的 q , 以及 Robinson 统计量中的 l 。它也证实了在更宽松的情况下其统计量将更具一致性, 并服务于渐进正态性。但 Shimotsu^[23]和 Phillips^[24]说明了在 $d=3/4$ 和 $d=1$ 时, 局部整体 Whittle(LW)方式的统计量满足的渐进分配不具备持续性, 当二分之一 1 时不具备一致统计, 因此当 $d>四分之三$ 时局部整体 Whittle 方式也并非最佳结果。因此, Shimotsu 和 Phillips 提供了精确的局部 Whittle(ELW)方式, 从而扩大了 d 的覆盖范围, 并证实了在 d 的时间间隔为等于二分之九时估计数量将遵循渐进的正态分配。Robinson 还依据 LR 检验结果提供了另一个检测长期记忆性问题的零点五参数估计方法-LM 方式。以上零点五参数方法都对带宽 m 的选取十分敏感, 因为不同的 m 会产生不同的计算结果。

三是参数方法, Sowell^[25]根据建立的 ARFIMA 模型, 推论出了无条件精准的极大似然函数方法来计算模型中全部的参量。极大似然函数法并不要求通过选取最佳的带宽 m 值, 就能够算出模型中全部的基本参数, 不过采用该方法需要事先选取适当的模型, 如果模型选取错误就会产生误导性的结果。

从理论上来说, 半参数方法比全参数方法的检验结果更为稳妥, 但 Baollie, Kongcharoen 和 Kapetanios^[26]对比了以 LW 方法为例的零点五参数方法和 ARFIMA 模型中的巨大似然函数估计的预期表现, 从二者的均方误差(MSE)比较看出, ARFIMA 模型的巨大似然函数估计方法预期效果较好, 同时还对比了最优估计带宽和带宽的样本空间的平方根下对 LW 的预期效果, 估计结果也显

示 LW 方法的有效性略差。在最近几十年里,改进 H 估计的尝试一直是一个热门话题,研究人员开发和测试了几种方法。可选的方法有多重分形去趋势波动分析、李雅普诺夫指数、Hudak 半参数方法、拟极大似然分析、中心移动平均法、广义赫斯特指数法、周期图法、小波方法学、基于几何方法的程序 (GM) 和压裂维数算法 (FD)。

Vaswani^[27]等在 2017 年发表的《Attention is all you need》介绍了以 self-attention 为基本单元的 transformer 模型使得注意力机制得到真正的成功运用。

Google 团队利用 transformer 代替了 Seq2Seq 的问题,再用自注意力代替了 LSTM,在翻译等任务中取得了更好的成绩。transformer 模型在自然语言处理的上下文语义等应用方面得到了广泛的应用。

Kitaev^[28]等将其中的一些结构进行优化,提出了新的模型 reformer。transformer 模型已经在生活中被广泛运用,如车牌、车票、身份信息识别^[29-36]。Zhou H^[37]等在 transformer 模型基础上提出了更适用于长序列输入和输出的 Informer 模型。

2.2.3 有关期权及其交易的研究

吴鸿超^[38]等以 GARCH 族, HAR 族和 ARFIMA 为基准,运用 50ETF 期权的高频价格数据建立了高频波动率预测模型。而将此模型运用于搭建 50ETF 期权中的波动率套利策略则发现,在波动率套利策略的各种模型中,在使用 EGARCH 模型的 Calmar 比率和胜率中最佳。而王西梅^[39]等人使用了参数法和非参数法两个方面对隐含波动率建立模型,并采用局部波动率模型对中国上证 50ETF 股指期货期权的价格进行实证调研。调查结果表明,非参数法对样品内数据模型化拟合的价格成果更高;而参数法对样品外数据分析的套期保值效果效果则更高。但相对于直接通过的市场统计模型,隐含波动率模型方法对期权定价与套期保值的有效性均较优。李雪飞^[40]等以历史波动率为条件变量,分析了自 2016 年 50ETF 期权上市不同期限上的标的商品未来实际波动率和期货价格隐含波动率,分析发现,2016 年起国内期货市场的价格有效性提高,但仍具有相当的价格套利空间。Li P^[41-42]发现在对期权进行估值时,就当月期权而言,使用平价隐含波动率在期权估值方面的表现优于其他波动率函数;对于下月期权,波动率函数的使用可以提高期权估值精度。杨昌辉^[43]等结合了嵌套长短时记忆神经网络模型 (NLSTM) 与 Heston 模型,基于华夏上证 50ETF、嘉实沪深 300ETF 和华泰柏瑞沪深 300ETF 的高频期权数据进行研究,发现 BS 价格模型和 Heston 价格模型对于 ETF 期权价值变化的描述都不精确,而将 NLSTM 神经网络模型和 Heston 模型融为一体的混合模型方法,则可以更有效的捕获 ETF 期

货的价格变化。姜晓晴^[44]以 B-S 模型和 CEV 模型为依据开展了实验研究,发现在值期权定价方面, B-S 模型与 CEV 模型在相同到期日下,对上证 50ETF 深度的市值期权价格体现最佳。由于 B-S 模型价格均精确于 CEV 模型,但这正好和中国的恰恰相反。孙友发^[45]等运用非仿射随机波动率模型处理了标的资产对数价值的特征函数问题,进而运用 Fourier Cosine 方式,推论出了欧式期权的拟闭型计价公式。研究人员发现推导的拟闭型定价公式相较于已有研究成果,在精度和效率方面都具有显著优势。郭精军^[46]等人提出了时间转换下的次分数布朗运动支付红利期权价格模型。使用次分数布朗运动扩散 B-S 模型确定带红利的欧式期权价格公式。然后再运用金融实际数据进行计算模拟,发现的新模型,能够反映金融资产的真实价值。吴鑫育^[47]等对时变风险厌恶条件下的期权价格问题进行深入研究后表明, TVRA-SV 期权价格模型的数值拟合效果明显高于传统的 CRA-SV 期权价格模型,从而可以更加全面地描摹标的上证五十 ETF 收益在客观和风险中性测度下的波动;而 TVRA-SV 期权价格模型的准确度则优于传统的 B-S 期权价格模型和 CRA-SV 期权价格模型。Recchioni MC^[48]等提出了多因素 Heston 模型的条件边际概率密度的两个新的表示公式,并通过非高斯函数在卷积中的渐近展开式,推导出欧式期权价格和隐含波动率的显式公式。Roul P^[49]等提出了一种求解欧式期权的时间分数阶 B-S (TFBS) 方程的数值方法,并将该方法应用于三个由 TFBS 方程控制的欧洲期权定价问题。

波动率策略的构建,主要由捕获买卖时机、对冲、管理风险等三步骤所构成。所谓捕获买卖时机,就是利用对买卖决策中的利润源泉加以分解,并利用实际经验或模拟找到有利可图的时点。对波动率买卖时机的捕获方面主要是从历史波动率的特征入手,利用历史波动率均值回复和聚集性这二个特征找到买卖时机。由 Burghardt 和 Lane^[50]建立的波动率锥,有助于很好的刻画波动性率在历史上的变化水平,被人们广泛的使用用来捕捉买卖时机。在中国国内的研究中,李璐^[54]利用十个日历史波动率的平均数与隐含波动率之差捕获买卖时机,从而形成了跨式期权实现套利。施智敏^[55]方面以隐含波动值在 IVIX 指数历史布局中的地位为信号时间节点,以实现跨式期权套利。另一方面也能够利用预测模型预估波动值的变动水平,来捕获买卖时机。谢佳珠^[56]则利用 GARCH、随机波动率模型, BS 模型等多种不同的波动值预估模型对波动值做出预估,以此实现跨式期权套利。

对冲是波动率交易中必不可少的环节,在对冲中必须在交易成本与风险抵减之间做出取舍。目前非系统化的对冲方式主要有三类:按照一定时刻间隔对冲、对冲于下一个 δ 带和按照标的资产价值变化对冲,这三类方式都主要依靠于交易者的主观判断。为使对冲投资风险的效果最大化,研究者们给出了基于效

用的对冲投资风险方式。Hodges 和 Neuberger^[57]认为在设定投资风险的厌恶系数时，在所有对冲投资风险中都必须存在一个临界点，使得保持未全部对冲头寸所产生的投资风险与完成全部对冲所产生的投资成本是无差异的，即空头与多头的对冲投资风险也有所不同。Whalley 和 Wilmott^[58]给出了 HN 理论上的一种近似最优预测解。而 Zakamouline^[59]所提供的双渐进最优预测解可以很好的贴合 HN 理论上的特点。与彭程，李爽，包莹，赵延龙^[60]等建立了外汇期货市场的摩擦系数，并通过对摩擦系数的长期监测，以不同的对冲方法增加了对冲效果。

风险管理原则是，通过止盈止损和对市场风险的重新评估限制高波动率策略的回撤，以减少重大损失的出现。在对中国的机构投资者的调研中，通常采用中国波动率指数作为风险识别的指数，当中国波动率指数大幅度上升时，由于市场的恐惧情绪上升，预期将无法生效，因此可选择平掉持仓，以控制损失。

2.2.4 相关文献述评

本文从波动率预测的相关研究、期权定价的相关研究以及中长序列预测与 transformer 模型的相关研究三个方面进行了已有文献的梳理与总结，对当前的研究现状进行了较为详细的描述。本文相比前人研究，主要存在以下改进之处：

第一，将中长期预测引入波动率预测模型中。传统波动率预测模型往往只侧重于短期个位数交易日，能向投资者提供的风险规避作用有限，中长期波动率的预测将带给资本市场更高的效率，帮助投资者在更大程度上识别机会与风险。

第二，基于不同的波动率信号制定期权交易策略。已有研究发现，随着预测日期的增加，期权定价公式的准确度也会随之发生变化。本文将在不同的预测窗口期内探究各参数信号的表现，以期归纳一个较为稳定的期权交易策略。

第3章 波动率预测问题分析与交易策略的构思

3.1 波动率与期权价格

3.1.1 波动率预测方法的选择

本文中为减少期权到期后对交易产生的价值波动性影响，而选择了在期权到期前1周进行换月处理，所以虽然本文已经决定预计期权在交割日前3周-前1周的隐含波动率变化情形，但现有研究成果已经证实，由于输入序列变化的时间长短与所隐含的价格信息成正向相关性，因此可以提高模型估计的准确性，为得到更良好的波动率估计结果，因此需要采用预言力更强大的长序列时间序列估计(Long sequence time-series forecasting, LSTF)模型。

长序列或时间序列估计都要求模型具备较大的估计力量，即可以更准确地捕捉产出和投入相互之间的持续依赖性关联。近年来，针对序列预测问题的研究主要都集中在短序列的预测上，输入序列越长，传统模型的计算复杂度越高，同时预测能力显著降低。Informer模型在transformer模型基础上增加了自注意力蒸馏机制，通过将低层级输入减零点五来突出主导注意力，从而有效增强了处理较长输入序列的能力。为验证Informer模型的表现，本文选择同时使用OLS和LSTM模型预测波动率与之对照。

3.1.2 期权交易策略的构思

期货作为一种不对称的衍生工具，通过对多个期货合约进行搭配，能够构造出多种不同的收益结构。主要有领口期货，跨式组合期权，蝶式组合期货和鹰式组合期权等多个期货合同所组成。但相对于蝶式组合和鹰式结合，期货的跨式结合则只应用在了二个期货。由于其结构简单，实用性较高，所以期货跨式结合也是目前国内外期货策略中，较为常见的一种期货交易策略。买入跨式结合由同一行权价和同一结束时段的一组认沽期权和一组认购人期货多头构成，当标的资产价值脱离运行价位较大区域时获益较多，但损失有限，且利润无限。所以，本文选用跨式期货结合为期货交易策略的主要投资合约。

卖出跨式结构由同一行权价和同一结束时段内的一组认沽期权和一组认购期货空头构成，当标的资产价位处在同一执行价格以下时收益最大，但利润有限，且损失无限。对期货跨式战略而言，由于看涨期权与看跌期权之间的Delta值恰恰相反，所以跨式期货合并对外披露的Delta值很小，这就是本文中选取

使用期货跨式合并为战略构建基准的另一种因素。

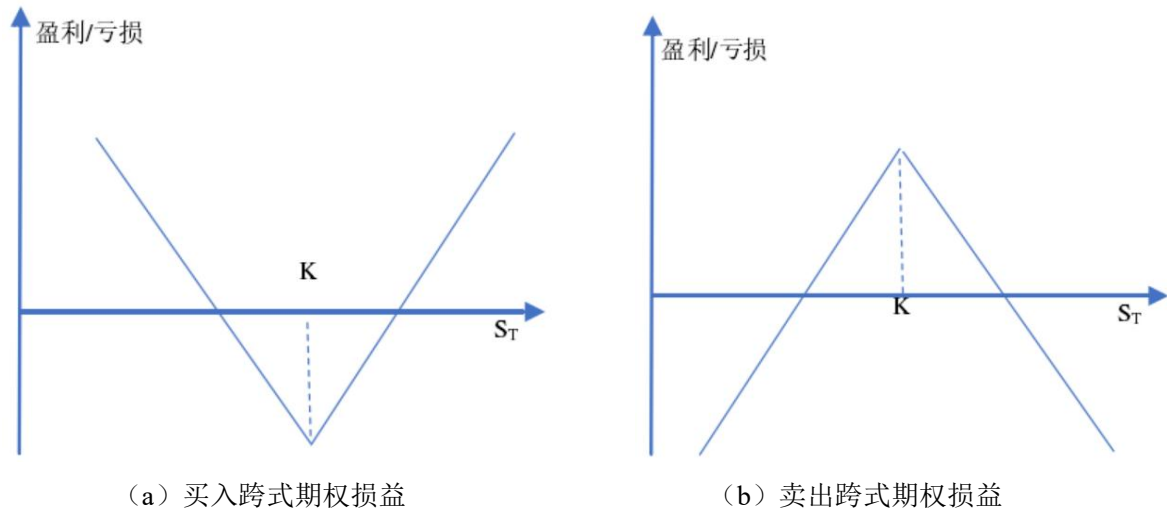


图 3-1 跨式期权组合损益图

3.2 波动率预测及交易策略的设计思路

3.2.1 波动率预测的设计思路

为获得准确的波动率预测结果并说明方法选择的契合性，本文分别通过最小二乘法、LSTM 模型和 Informer 模型进行波动率的预测，分别代表传统波动率预测模型、时间序列预测模型和引入自注意力机制的 transformer 类模型。

在因子筛选方面，现有研究已经证明了市场上存在大量会影响波动率变化的各种因子，且彼此之间并不是严格的线性相关关系，为了保证所选因子的有效性，本文选择使用 SHAP 进行因子筛选。SHAP 基于博弈论中的沙普利值法，通过计算各因子的贡献度为因子进行排序，综合 XGBoost、LightGBM 和随机森林三种树模型的筛选结果，最后选择 7 个因子作为模型的输入变量。

关于模型训练，最小二乘法和 LSTM 模型均采用滚动预测的模型，使用最新的信息进行一个时间步长的预测；Informer 模型具有一次性输出所有预测结果的功能，故设置为对未来一周、两周、1 个月进行预测。

3.2.2 期权交易策略的设计思路

(1) 策略构造原理

根据对上述波动率变化中关键信号的构建和对当前波动率变化趋势估计的

结果, 可进行 Delta 中性的波动率策略构建, 具体构成原理为:

1) 长期:

利用调整后的波动率锥作为信号构造基准, 以隐含波动率的高于百分之七十分位点和低于百分之三十分位点为主要突破位置, 并利用波动率均值逐渐回复的特征确定波动率在中长期变化的趋势。但同时由于考虑到波动率的高度集聚特征, 当发生高位波动率或低点波动率之后并没有立即恢复到均值, 往往还会一直维持在较高点或低点附近一段时间, 所以当持续发生了五天以上的相同突破信号时, 才进行开仓买卖。如当隐含波动率持续五日高于百分之七十分位点时考虑波动率即将下降, 抛出波动率; 当隐含波动率持续五日低于百分之三十分位点时考虑波动率即将提高, 买进波动率。选择连续五日的主要原因在于, 通过观察 IVIX 的走势可看出震荡率在一个较高点或低点的时间内通常不能超出两周, 而且通常在一个月过后又会有往相反方向移动的走势。

2) 中期:

当隐含震荡值并不是很明确的价格处在高位或低点状态时, 可以考虑利用中期的择时信号进一步顺应震荡值的变动方向, 以实现逆向买卖。择时信号选择利用 IVIX 的短时 MA 线与长久 MA 线加以构建, 当 IVIX 的短时 MA 线上行穿越长久 MA 线时, 波动率有上涨的趋势, 买进波动率; 当 IVIX 的短时 MA 线下行穿越中长期 MA 线时, 波动率有下降的趋势, 抛出波动率, 同时利用均线信息捕获波动率锥中无法捕获的信息。

3) 短期:

当证券市场隐含波动性率趋势并不明确时, 可以选择根据证券市场的交易心情、波动性率历史走势、标的证券市场走势等因数来预计次日波动性率趋势, 而由于所用到的因数大部分是短期内的证券市场影响因数, 所以该预测方法首要作用就是预测市场短期内的波动性率趋势, 从而可以掌握在短期内中特定证券市场影响作用下的波动性率趋势。首先选取了标的证券市场因子, 期权市场因数, 历史波动率因数, 预期波动性率因数以及海外市场因数等五种重要类型的因数, 采用了最小二乘模型, LSTM 模型和 Informer 模型进行估计, 并经过对比三个模型的预期精度及其在测试集的策略表现, 筛选出了最合适的模型。同时为提高获利的机率, 对买进开仓与卖出开仓各自设定门槛, 即当预期波动

性率改变超过买进门槛时买入波动性率，当预期波动性率改变低于卖出门槛时卖出波动性率。

(2) 期权组合选择

在期权组合的选取方面，一般选用跨式平值期权的期货组合。本文中对平值期权的界定范围比相对广泛，对 ATM 期货的界定范围一般是指等于 0.95 标的开盘价，或低于 1.05 标的开盘价的期货合约。对平值期货设定较宽的时间限制，也就可以掌握更多的波动率买卖机会，同时在期权离到期日不足七天时更换为下月合约持有。这种方法的好处在于：

1) 跨式期货同方向买卖的认沽与认购人期货，其 Delta 正负相抵，因此可以降低 Delta 的对冲交易生产成本；

2) 由于平值期权的市场流动性较好，因此选择较广范围的平值期权可以捕捉到更多的监管套利机会，从而提高了开仓次数，提高了利润；

3) 当期权离到期日比较接近时，由于 Theta 和 Gamma 的值都比较大，所以出售期权的 Gamma 风险很大，而买进期权的 Theta 的值风险也很大，策略组合的收益率很容易受这二种值的影响，会承受相当大风险，所以在离到期日七天更换合约就可以降低 Theta 和 Gamma 所暴露的风险。

3.3 模型构建及策略设计的理论框架

3.3.1 SHAP 因子筛选原理

Shapley value 可以通过组合来对整个模型做出全局的解释。我们对整个数据集中每一个样本计算他的 Shapley value，这样就会得到一个矩阵。这个矩阵的每一行表示一个数据，而每一列表示一个特征。于是我们就可以通过对这一个矩阵的分析得到整个模型的重要特征。首先我们会求出每一个特征的 Shapley value 的平均绝对值，计算过程如式 3-6 所示：

$$I_j = \sum_{i=1}^N \|\phi_j(x_i)\| \quad (3-1)$$

其中 I_j 表示第 j 个特征的平均 Shapley value。 $\phi_j(x_i)$ 表示第 i 个数据的第 j 个特征的 Shapley value。 N 表示数据集中共有多少数据。

在计算得到每个特征的平均 Shapley value $I = (I_1, I_2, \dots, I_N)$ 之后，我们对了

按 I 从大到小的顺序进行排序。于是我们得到的前 M 个特征就是模型对应的前 M 个重要特征。

3.3.2 自注意力机制改进原理

为提高 Transformer 类模型从长序列输入中提取鲁棒性的能力, Informer 模型在 Encoder 中使用了自注意力蒸馏机制, 利用蒸馏操作对具有主导地位的高级特征进行特权化, 在下一层生成一个集中的自注意力特征图, 这个操作能够大幅削减输入长度, Informer 采用的具体蒸馏方式如下:

$$X_{j+1}^t = \text{MaxPool} \left(\text{ELU} \left(\text{Conv1d} \left([X_j^t]_{AB} \right) \right) \right) \quad (3-2)$$

其中 $[\cdot]_{AB}$ 包含了多头 ProbSparse Self-attention 和注意力块中的基本运算, $\text{Conv1d}(\cdot)$ 表示 1-D 的卷积滤波器(kernel width=3), 并在时间维上使用 $\text{ELU}(\cdot)$ 激活函数。

Informer 用一个步长为 2 的最大池化层, 在每一层之后将 X^t 下采样到其一半的长度, 这样能够将整体内存使用量减少为 $O((2 - \varepsilon)L \log L)$, 其中是一个小数。而为了增强蒸馏操作的稳健性, Informer 对 Encoder 的主结构进行堆叠组合, 通过使用不同输入范围和输入长度得到多个关注不同时间区间的 Encoder 结构, 再将它们的输出进行拼接得到完整 Encoder 的输出。

3.3.3 期权 Delta 中性交易理论

(1) 波动率锥信号的构建

波动率锥是由 Burghardt 和 Lane 最先发明, 于是有了更多的期货工作者和相关研究专家学者使用波动率锥来发掘波动率交易机遇。由于波动率锥对标的资产历史波动率在各种时期宽度下的散布情形做出了刻画, 又由于短期内波动率的散布范围较广, 从短时到长时间波动率的散布都呈锥形, 于是称为波动率锥。隐含波动率和历史波动率之间有很大的相互关联, 据此可利用波动率锥确定隐含波动率在历史发展中所处的水平, 对未来隐含波动率的走向做出研判, 利用波动率锥也有助于找到长期买卖时机。当利用历史波动率锥对隐含波动率做出预报时, 大致具有如下几个优点:

- 1) 历史波动率的时间, 与隐含波动率对应的距结束日时间长相匹配;
- 2) 立体的显示当前隐含波动值所处的水平, 从而确定了波动率交易时机;
- 3) 把当前的波动率信息放在历史信息中, 更加全面。

本文中在建立波动率锥时使用高频的已实现波动率对其波动率水平进行测量，第 t 天的已实现波动率 RV_t 的计算方法为：

$$R_{ti} = X_{ti} - X_{ti-1}, (i = 1, 2, \dots, N) \quad (3-3)$$

$$RV_t = \sum_{i=1}^N R_{ti}^2 \quad (3-4)$$

这里， R_{ti} 为第 t 天第 i 个取样区间的自然对数收益率，而 X_{ti} 则为第 t 天第 i 个取样点数据的自然对数。设定参数 N 和 T ，其中 N 为建立波动率锥信号所需要的样本区域宽度， T 为期权的剩余到期长度。对每一天滚动选择前 N 天的标的资产高频交易数据作为估计波动率的样本数据，在样本数据中多次随机选择连续的 T 日，计算年化已实现波动率。从中选择 90 分位点（per90）、70 分位点（per70）、50 分位点（per50）、30 分位点（per30）和 10 分位点（per10）构造剩余到期时间为 T 日的波动率分布情况图，每一日中的每个剩余到期时间对应着同一个波动率分布。以 2020-2021 年上证 50ETF 的收益率数据为基准，测算出的与期权到期年限相符合的波动率锥信号，如图 3-2 所示，5 条波动率锥信号曲线分别代表当以 360 天为波动率锥计算周期时，在不同到期时间下 90 分位点（per90）点、70 分位点（per70）、50 分位点（per50）、30 分位点（per30）和 10 分位点（per10）的波动率，横轴为距到期时间，纵轴为波动率，从图 3-2 可以看出波动率锥的左侧陡峭部分反映了短期波动率的分布，而右侧平缓部分则体现了中长期波动率的分布。

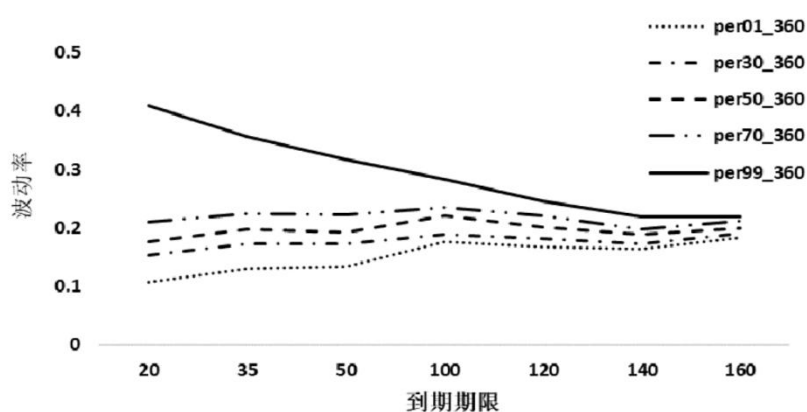


图 3-2 2019-2020 年波动率锥

在本文的计算方法中，波动率锥由已实现波动率得到，但是隐含波动率对已实现波动率有一定的溢价。所以本文在使用波动率锥时对波动率锥进行简单的调整，从而使得波动率锥对隐含波动率分布的刻画更加准确。本文选择的调整方法为使用前 M 天作为波动率溢价计算的时间周期，通过计算在前 M 天里标的资产波动率与 IVIX（中国波动率指数）之差的均值对波动率锥进行调整。

计算公式为:

$$mean_{IVIX_t} = \frac{1}{M} \sum_{i=t-M}^{t-1} IVIX_i \quad (3-5)$$

$$mean_{RV_t} = \frac{1}{M} \sum_{i=t-M}^{t-1} RV_i \quad (3-6)$$

$$adjvalue_t = mean_{IVIX_t} - mean_{RV_t} \quad (3-7)$$

(2) 波动率趋势信号的构建

为比较精确的捕捉到波动性率趋势变化的走向,我们选择了运用波动性率均线配合来对波动性率走势变化规律加以深入刻画,而顺应波动性率走势变化规律通常就可以把握波动性率大幅度范围变动的机会获得较多的收益。因此我们采用了 IVIX 指数作为波动性率均线构建基准,依次构建出了 IVIX 五日均线, IVIX10 日均线, IVIX15 日均线和 IVIX20 日均线,并利用长度均线的配合确定了波动性率走势。因此,当短均线上穿长均线时,可以说明波动性率在近一个月内的上涨势头明显而强势,从而发出了波动性率上升的信号;当短期内均线下跌越过中长期均线时,说明了震荡率在近一个月内下滑趋向强烈,给出了震荡量逐渐减少的讯号。

3.4 模型的有效性评价

3.4.1 波动率预测评价方法

对于评估所构建的三种模型的最大波动度预测准确性,本文将采用最大损失函数 MAE、MSE 进行对比,计算方法如下表所示:

表 3-1 损失函数说明

名称	计算公式
平均绝对误差	$MAE = \frac{1}{T} \sum_{i=1}^T IV_t - \sigma_t^2 $
均方误差	$MSE = \frac{1}{T} \sum_{i=1}^T (IV_t - \sigma_t^2)^2$

这里, IV_t 代表已实现波动率, σ_t^2 代表模型中预测得到的波动率, T 代表预测的样本数量。损失函数的值越小,就说明模型的预测性能越好。

3.4.2 期权交易策略评价方法

本文在回测中采用的评级指数,主要包括:年化收益率、累计收益率、胜率、

最大赢利与亏本的比率、最大回撤和夏普比率。

N_c 表示持有看涨期权手数, N_p 表示持有看跌期权手数, $ClosePrice_c$ 表示看涨期权收盘价格, $ClosePrice_p$ 表示看跌期权收盘价格, $OpenPrice_c$ 表示看涨期权开盘价格, $OpenPrice_p$ 表示看跌期权开盘价格, $yearlyreturn$ 表示年化均值收益率, $yearlystd$ 表示年化波动率, 合约乘数为 10000 份, 无风险利率设为 3%, 则以当日开盘开仓为例, 其余开仓平仓收益情况使用相应的开盘价格、收盘价格, 收益率为:

$$return_t = \frac{(ClosePrice_{c,t} - OpenPrice_{c,t}) \times N_{c,t} + (ClosePrice_{p,t} - OpenPrice_{p,t}) \times N_{p,t}}{ClosePrice_{c,t} \times N_{c,t} + ClosePrice_{p,t} \times N_{p,t}} \quad (3-8)$$

累计收益率为:

$$accretreturn_t = \prod_0^t (1 + return_t) \quad (3-9)$$

每日盈亏为:

$$pnl_t = \left(\begin{array}{l} (ClosePrice_{c,t} - OpenPrice_{c,t}) \times N_{c,t} \\ + (ClosePrice_{p,t} - OpenPrice_{p,t}) \times N_{p,t} \end{array} \right) \times 10000 \quad (3-10)$$

夏普比例为:

$$sharp = \frac{yearlyreturn - 0.03}{yearlystd} \quad (3-11)$$

第 4 章 基于波动率预测下的期权 Delta 中性交易策略设计方案

4.1 数据说明与变量说明

4.1.1 数据来源

本文将选择 2015.02.09-2021.12.31 的数据分析当作总样本集，2015.02.09-2019.12.31 的数据当作训练集，2020.01.01-2021.12.31 的数据当作验证集，以进行建模参数优化和预测。文中使用了华夏上证 50ETF，以及上证 50ETF 期货的有关财务数据，而上述财务数据均来自于万得数据库。

不同于静态预测的方式，训练需要将所有数据集合按照 8:2 的比率进行分配，只完成了一个模拟训练和一次预测，而滚动预测方法则要求按照实际模型更新频率的需要对训练集和测试集中分配。本文在设置训练集长度时考虑到，上证 50ETF 期权平均一年约有 250 个交易日，而每个月份一般只有 20 个交易日，所以将模型输入长度设定为 500 交易日，即两年的交易数据；而测试集中则设定为 20 个交易日，即每月的交易数据。当完成了一次练习训练和预测的工作后，滚动预测方法要求将训练集和测试集中数据整体向后移动 20 个交易日的长度，然后通过添加新数据并向前移动练习训练集中进行下一轮的模拟更新。在采用了滚动预估的方式之后，本文将对上证 50ETF 期权自 2017 年 3 月 2 日至 2021 年 12 月 31 日期间的当月隐含波动值作出预估。

4.1.2 预测因子选择及描述性分析

在因子的选取中，本文考察了如下几大类因子：

（1）历史波动率因子：

波动率之间具有较强的自相似性，使得本文选取了历史波动率相对因子作为预测因子。历史波动率类因子选择法采用了二个以不同方式统计后得到的历史波动率，它们是低收盘-高收盘的标准差方法和高频已完成历史波动率方法。

（2）标的市场因子：

标的交易市场的变动状况和期货交易市场密切相关，故文章中选取了利用标的交易市场的相关信息估计指标。标的行业因子选择了与标的股票定价和成交量变化状况密切相关的因子，包含了定价极差、振幅、真实波幅、动量，

以及相应强弱指标等所有与波动和情绪变化密切相关的因子。

(3) 期权市场因子:

期权市场中的价格信息和隐含波动值的变动之间关联得较为密切。在期权市场类因子中,选取了隐含波动率、期权成交量、期权持仓量等相关领域方面的重要因子。

因子的具体情况如下:

表 4-1 预测因子计算方法说明

因子类别	因子名称	因子缩写	计算说明
历史波动率因子	前 30 日历史波动率	HV30	最近 30 个交易日的平均对数收益率
	前 60 日历史波动率	HV60	最近 60/90 个交易日的平均对数收益率
	前 90 日历史波动率	HV90	最近 90 个交易日的平均对数收益率
标的市场因子	收盘价	close	当日收盘价
	成交量	vol	当日成交量
	成交金额	amount	当日成交金额
	涨跌幅	%Chg	当日涨跌幅
	换手率	turnover	当日换手率
	振幅	swing	(最高价-最低价)/昨日收盘价
	前 5 日真实波幅	ATR_5	Max(前 5 日振幅, 今日最高-昨收, 今日最低-昨收)
	前 10 日真实波幅	ATR_10	Max(前 5 日振幅, 今日最高-昨收, 今日最低-昨收)
	相对强弱指数	RSI	$[上升平均数 \div (上升平均数 + 下跌平均数)] \times 100$
	下月隐含波动率	IV_nm	下月到期期权价格隐含波动率
期权市场指标因子	当季隐含波动率	IV_ts	该季月到期期权隐含波动率
	下季隐含波动率	IV_ns	下季月到期期权隐含波动率
	P/C ratio	PCratio	看跌期权成交量/看涨期权成交量
	机构净买入量	np_i	金融机构净买入量

表 4-2 各项变量的描述性统计

变量	名称	均值	标准差	最小值	最大值
close	收盘价	2.4997	0.4022	1.7005	3.5878
vol	成交量	6.1489	6.7394	0.7785	91.4683
amount	成交金额	16.9425	19.2284	1.7197	249.0883
chg	涨跌幅	0.0499	1.5131	-9.9824	8.8484
turnover	换手率	4.4495	4.8975	0.6360	69.6547
iv_nm	下月隐含波动率	0.2181	0.0983	0.05	0.87
iv_ts	当季隐含波动率	0.2189	0.0869	0.06	0.58
iv_ns	下季隐含波动率	0.2213	0.0798	0.05	0.57
hv30	前 30 日历史波动率	0.2134	0.1188	0.06	0.65
hv60	前 60 日历史波动率	0.2211	0.1131	0.07	0.61
hv90	前 90 日历史波动率	0.2260	0.1083	0.08	0.55
swing	振幅	1.7250	1.3355	0.3091	14.5545
atr_5	前 5 日真实波幅	0.0456	0.0286	0.0115	0.2577
atr_10	前 10 日真实波幅	0.0456	0.0268	0.0139	0.2348
rsi_9	相对强弱指数	53.5372	14.4131	11.7175	93.9027
pcratio	P/C ratio	0.8028	0.1343	0.4370	2.0643
np_i	机构净买入量	0.1881	1.2355	-6.0222	8.5866
iv	当月隐含波动率	0.2035	0.1039	0.03	1.25

机器学习方法通常用于对许多种属性描述的对象进行分类，在这种类型的应用中，很大一部分属性可能与分类问题完全无关，并且通常无法先验地确定哪些属性是相关的。SHAP 将输出值归因到每一个特征的 shapely 值上，去除由于变量、随机树模型、参数等带来的误导性影响。将上文提取的 17 个变量放入 SHAP 模型中，基分类器选择随机森林，并将其结果和 XGBoost、LightGBM 两种树分类器的排序进行综合评估。筛选结果前 10 名如下所示：

表 4-3 变量筛选

综合排名	变量
1	IV_nm
2	ATR_10
3	ATR_5
4	RSI_9
5	HV30
6	turnover
7	np_i
8	close
9	%Chg
10	vol

最终选择下月隐含波动率、10 交易日真实波幅、5 交易日真实波幅、9 交易

日 RSI、前 30 日历史波动率、换手率、机构净买入量、收盘价、涨跌幅和成交量 10 个因子作为后续章节中输入模型的变量。

4.2 波动率预测模型构建

4.2.1 普通最小二乘法滚动预测

OLS 滚动回归的波动率方法的基本原理，是把 $T+1$ 期间的波动率变化值视为被解释变量， T 期间的因数值则视为解释变量。首先利用滚动方法得出每窗时的因数 IC 值，即因数和被解释变量间的相互关系，在每窗时选取 IC 值最大的变数为解析变数。然后再利用下窗口期的样本数对求解变量与被解释变量进行拟合。然后再利用模型拟合得出的系数估计下一期的波动率变化值。具体的过程为：

过程 1：每窗口时间内有 500 个样本，在各个窗口期分别统计各种因子的因子 IC，并选取 IC 中绝对值超过零点一的因子。

过程 2：在窗口期内用次日的隐含波动率作为因变量，用当日的所选因子作为自变量，然后通过 OLS 回归拟合得到回归系数。

过程 3：利用得到的标准化回归系数，估计窗口期下一日的隐含波动率，然后滚动窗口期，得出全部的预测值。

4.2.2 LSTM 模型预测

模型中基本的参数设定如下：学习率 lr 默认设定为 0.0006；隐含层单元 rnn_unit 设定为 10；因为要预测的是本月隐含波动率，所以模型输出维度 output_size 设定为 5；输入维度 input_size 为 18；经过多次试验，本文设定 batch size 为 60，即每一批次训练 60 个样例，约为 1 个季度的数据；迭代次数设定为 500；时间步长 time step 默认为 20。

4.2.3 Informer 模型预测

Informer 模型的参数如下所示：Encoder 输入序列的长度 seq_len 设定为 500；Decoder 中的 start token 的长度 label_len 设定为 10；预测序列长度 pred_len 设定为 5；模型的维数 d_model 设定为 512；模型中全卷积神经网络的维数 d_ff 设定为 2048；激活函数设定为 'gelu'；encoder 和 decoder 的输入宽度均为 11；输出宽度为 1；encoder 层数为 2，decoder 层数为 1；batch_size 设定为 60，learning rate 设定为 0.0001。

4.3 策略实施

4.3.1 波动率策略构造步骤

本文的波动率策略构造步骤为：

(1) 开仓：

本文的主要策略包括三个开仓的时间节点，当波动率锥信号为 1 或-1 时，即隐含波动率持续五日突破上下同分位置时予以开仓；当波动率锥的信号为零，即 IVIX 均线信号为 1 或-1 时，即长线无走势，中期波动性率有顺均线方向移动的走势时予以开仓；当波浪率锥的信号为零，均线组合的信号为零，或预期的波动率变化绝对值大于 0.1 时，予以开仓。具体开仓信息根据当日或之前的历史数据确定，一般采用次日开盘价予以开仓。

(2) 换仓：

在据期权离结束日还剩七个自然日时，由于充分考虑到接近结束的期权经营风险很大，所以必须在据结束日还剩七日时变更为次月合同，并按照最新的合同内容进行调整确定头寸。换仓信号以当日的开市数量确定，同时以开市数量对旧合约进行平仓，对新合约进行开仓。

(3) 平仓：

当波浪率锥、与均线交叉的信号值为 0 并且预测的波浪率变化绝对值也小于零点一时，可以平掉全部头寸。平仓信号根据当日或之前的实际情况确定，一般采用次日的开盘价位进行平仓。

(4) 策略对冲方法：

考虑到上证 50ETF 现货做空成本较高，且上证 50 指数期货的对冲效率也不高，因此文章中考虑了不采用对标的有关资产实现 Delta 对冲，而只是采用调整看涨期权与看跌期权的持股比率，使每日的平均 Delta 值维持为零。Joshua D 和 Tyler^[61]在他们的研究中给出了通过权重配比使得跨式组合期权的 Beta 保持中性的方法。文章参考了两位学者的思路，利用权重配比使跨式组合期权的平均 Delta 维持中性。假设看涨期权的 Delta 为 δ_c ，看跌期权的 Delta 为 δ_p ，看涨期权的持有比例为 w ，则可以解下列方程求得看涨期权的持有比例：

$$w\delta_c + (1 - w)\delta_p = 0 \quad (4-1)$$

可以解得上述公式中看跌期权的持有比例 $w = \frac{-\delta_p}{\delta_c - \delta_p}$ ，则看跌期权的持有比例例为 $\frac{\delta_c}{\delta_c - \delta_p}$ 。

利用以上公式，在每天开盘时都可以依次求出每天的看涨期权和看跌期权

的总仓位保持比率，并通过调节开盘价格对看涨期权和看跌期权头寸加以调节，从而保证了每日 Delta 中性。

（5）策略模拟交易设置：

初始资金设为 200 万，无风险利率设置为 3%。

我们把所有的数据集合分成训练集和测试集，在训练集对策略参数加以调优，在检测集种对策略的调优结果加以检测，并且检查策略的稳定性。本章的训练集样本区间为 2015.12.7-2019.12.31，验证集的样本区间为 2020.1.1-2021.12.31。训练集并未在 2015.2.9 开始的主要原因是，在进行波动率变化估计时，采用 200 个样本的滚动窗口，因此在 2015.12.7 进行了第一个预测。

下面将详尽说明在训练集中波动率策略下的结果和在测试集参数的调优方式和结论。

第5章 交易策略设计方案的有效性评价

5.1 交易策略方案的有效性评价

5.1.1 波动率预测效果评价

在 OLS 模型中, 在全样本下, 首先解析自变数和因变数间的关联性, 以过滤相关更多的因数, 在 17 个因子中有 8 个因子和因变数的关联很大, 如表 5-1 所示:

表 5-1 核心变量的相关系数矩阵

变量	iv	iv_nm	iv_nm	iv_nm	pcratio	vol	chg	rsi	np_i
iv	1								
iv_n	0.189***	1							
m									
iv_ts	-0.411*	0.009***	1						
iv_n	-0.403***	0.032***	0.411***	1					
s									
pcrat	-0.197**	0.004***	0.889***	0.033***	1				
io									
vol	0.121***	0.008***	0.828**	-0.004**	0.984***	1			
chg	-0.128***	0.004***	0.815*	-0.026***	0.981**	0.978***	1		
rsi	-0.113***	-0.032**	0.178***	0.124***	0.115*	0.101***	0.086**	1	
							*		
np_i	0.103**	0.026***	0.106***	0.008***	0.109***	0.097**	0.107**	-	1
							*	0.206***	

注: **、*、*分别表示在 1%、5%、10%的水平(双侧)上显著。

在表 5-1 中, 可发现隐含波动率变化趋势与上月隐含波动率变化趋、机构的净买量成比率关系, 表示了当远期隐含波动举办率高时隐含波动率有上升的态势, 当机构在市场上净买较多时隐含波动率也有上升的态势; 当季、次季隐含波动率变化趋势均为隐含波动率的历史数据, 隐含波动率变化趋势的历史数据与 P/C 值呈反比例关联, 表示了隐含波动率偏离平均数较远或出现较大幅度变动时, 隐含波动率变化趋势有反向回复的趋向。然后, 再通过以上因子对

隐含波动率变化趋势进行回归，结果如表 5-2 所示：

表 5-2 全样本下 OLS 多元回归结果

变量	系数	标准差	T-Statistic	P 值
iv_nm	-0.1562	0.0707	-2.2081	0.0274
iv_ts	-1.4562	0.5632	-2.5791	0.0100
iv_ns	1.3967	0.2441	5.7215	0.0000
pcratio	0.3616	0.2162	1.6719	0.0948
vol	0.0593	0.0145	4.0681	0.0001
chg	0.5383	0.1033	5.2093	0.0000
rsi	0.00009	0.00001	8.151782	0.0000
np_i	-0.3888	0.1805	-2.1532	0.0315
R-squared		0.3773		
Adjusted R-squared		0.3736		

由表 5-2 中可以得到，在整个样本中，这八个因子对隐含波动率变化趋势在 90%的置信水平下均具有显著的影响，并且总体的模型拟合优度超过了 0.37，模型调整后的拟合优度也较高，证实了在全样本下的因子池中选择的因子，可以对隐含波动率的变动做出合理的预期，于是在本文中通过上述因子池可以对各个时期的隐含波动率变化趋势做出合理滚动预期。

综合各个模型在上证 50ETF 期权隐含波动率验证集上的评价结果如表 5-3 所示：

表 5-3 三种预测模型下的波动率预测情况

模型名称	评价指标	预测长度		
		5	10	20
OLS	MSE	0.108	0.175	0.396
	MAE	0.284	0.424	0.504
LSTM	MSE	0.114	0.193	0.236
	MAE	0.272	0.358	0.392
Informer	MSE	0.098	0.158	0.183
	MAE	0.247	0.319	0.346

由表 5-3 可知，在未来 5 个交易日（1 周）、10 个交易日（近半个月）和 20 个交易日（近 1 个月）3 种预测长度下，Informer 整体表现均优于 OLS 模型和 LSTM 模型。其中，当预测长度为 5 个交易日时，3 个模型的 MSE 和 MAE 差距均不大，且 OLS 模型的 MSE 要略小于 LSTM；但随着预测长度的上升，与 Informer 模型相比，OLS 模型和 LSTM 模型在预测能力上皆表现出了较大幅度的退化，其中 OLS 模型尤为明显，这也体现了线性回归模型在多步预测类问题上的弱势。

5.1.2 期权交易结果评价

（1）不同预测方法下的策略表现

为研究前三预测模型中单一信号的策略表现，本章将波动率交易策略的开仓信号设定为第三章中三个模型所预测出的波动率变化情况，即当模型预测波动率上升时买进波动率，当模型预测波动率下降时卖出波动率，并同时设定买进阈值与卖出阈值设定为 0。在此基础上，研究波动率变化预测模型的单信号表达，以及比较各种波动率变化预测模型在策略实证中的表现，如表 5-4 所示。

表 5-4 三种预测模型下的策略具体表现

	OLS	LSTM	Informer
年均收益	3.27%	5.29%	7.11%
胜率	46.23%	51.33%	56.67%
盈亏比	1.12	1.27	1.41
最大回撤	10.34%	7.92%	6.31%
夏普比率	0.66	0.94	1.15
累计收益率	9.81%	14.67%	24.85%

从三个预测模型的累计收益率数据分析，三个模型在训练集均实现了正收益率，Informer 模型对波动率下跌时的估计准确度最大，而且在现实中的 2020 年为期权交易卖方最大获益市场，所以在波动率策略的实证分析中体现得最佳。

（2）不同参数下的策略表现

按照本文波动率策略的构成逻辑，文中包括了三种信号，依次是波动率锥信号、均线信号、预测模型信号，在这三种信号的形成中，文章还设定了多种参数和阈值，试图找出最能适合于中国市场规模的决策模型。本文中共有六组参数可以进行优化设定，分别是：

- 1) 波动率锥信号的时间跨度 N : 即表示以历史 N 日的数据构成波动率锥, 以不同的历史周期反应不同时间长度中的波动度分布状况。
- 2) 波动率溢价的时间跨度 M : 即以历史 M 日的数据计算溢价, 不同历史周期反应了不同时期长度内的波动率溢价状况。
- 3) 波动率均线组合: 即为短中期波动性率的变动而捕捉信息, 相对长期移动均线配合而言, 各种类型的长短期移动均线配合可以反映各种波动性率改变时间。
- 4) 波动率预测模型: 不同的波动率预测模型将具备不同的预期能力, 它与长期信息以及中期信号的结合, 也将产生不同的效应。
- 5) 买入阈值: 即在市场预期的价格波动率上涨或达到某个阈值时会买进, 以保证在一定的情形下获得真正的收益率。
- 6) 卖出阈值: 即在市场预期的价格波动率下跌或达到某个阈值时会出售, 而由于卖出期权的人天然地能够得到一定时间上的收益, 从而可以对买进与抛出设定不同的阈值。

根据以上策略参考, 本章一共使用了如下表格中的参数组合:

表 5-5 波动率策略参数组合

波 动 率 锥 周期 N	波 动 率 溢 价周期 M	均线组合	预测方法	买入阈值	卖出阈值
180	30	(5, 10)	OLS	0	0
360	60	(5, 15)	LSTM	0.005	-0.005
540	90	(5, 20)	Informer	0.01	-0.01
	120	(10, 15)		0.015	-0.015
		(10, 20)		0.02	-0.02
				0.025	-0.025
				0.03	-0.03

对上述参数结合在训练集进行了参数测试, 参数优化结果表明, 在波动率锥信号时间跨度为 540 天, 波动率溢价时间跨度为 60 天, 波动率均线组合为 MA5 和 MA20, 波动性率预测方法为 Informer 模型, 买进阈值为零, 卖出阈值为 0.005 时, 波动性率决策的综合表现结果比较良好。在进行了参数优选和信号优化之后, 策略的平均收益率达到了 11.65%, 夏普比率为 1.32, 累计收益率为 33.28%, 相比于优化以前的所有决策都有了显著的提高。虽然平均胜率有所降低, 但赢利与亏本的比率却增加了显著, 这主要是得益于加入了趋势性信息

后在顺趋势而行时更容易取得较大的盈利。在整套策略中最高回撤为 6.83%，由于回撤的波动幅度相对而言较小，因此策略的稳定能力较好，收益率也较高，因此波动率策略在测试集中有较好的表现。

表 5-6 最优参数下的策略具体表现

模型	Informer ⁺
年均收益	11.65%
胜率	54.3%
盈亏比	1.7
最大回撤	6.83%
夏普比率	1.32
累计收益率	33.28%

5.2 期权交易策略的风险提示

由于中国期货交易市场相较于外国起步晚滞，其金融市场体系还没有完善，所以投机性比较强烈，因此对于期货的投机风险也较高。尽管如此，本文设计的投资策略还是在回测过程中获得了相当可观的利润，不过采用机器学习的投资策略并不是百分百可行。Informer 模型最早运用在解决了天气研究预测的问题中，因为在天气问题中信息量大且规律性较强，所以这一模型可以有很好的体现，不过由于对于金融市场产品的相关数据，数量较小、规律性不强使得估计结果存在着一定程度上的不确定性，从而缺乏相对稳定的预期结论。另外，由于期货市场风险很大，和股市比较要求更大的资本体量，所以投资人也需要具备相当的风险承受能力。金融市场上没有稳赚不赔的策略，投资人切莫盲目投资，理性管理好个人财产。

基于 Informer 模型的价格波动率预测和期权交易策略的主要风险有如下四个方面：

1、模型和策略中的收益测算采用历史数据，尽管增加了无风险收益等条件可以更加模拟实际的交易情形，但其具体的测算值仍然是采用了理论上的可能性值，即在实盘交易的过程中，由于特定事件所导致的市场价值波动并不在本文模拟的测算范围之中，且预测值也并没有充分体现实际状况下的期权价值；

2、由于本策略选择的合约乘数为 10000 份，且初始资金量约为 200 万元，因此在投资初期由于持仓比率过高，一旦市场的真实情况与策略交易方向背离时，将可能会产生爆仓，而在后期一旦产生了可观利润之后，则会产生持仓量

相对较少，甚至资金闲置的状况，进而危害到投资人的利益；

3、期货市场目前投机气息浓重，市场变动难以预知，投资人必须具备充足的期货投资经验，并准确研判市场趋势，策略中的利润仅是理论利润，无法代表实际的市场交易环境。

第 6 章 结论

6.1 研究结论

在中国期权交易市场规模逐步扩大的时机，波动率交易也逐步走进了人们的视线，但是因为中国期权交易市场规模相对较小，所以目前关于期权波动率交易的有关研究成果相对较少。论文以实务的视角对期权波动率决策的关键理论，构成机理，构建过程等各个方面展开了深入细致的理论分析与实验研究，探究出在当前金融市场中不同因子之间对波动率变化趋势的预期能力，各种波动率变化趋势预测模型的决策与实践体现，各种波动率策略参数的选择方式与决策实践体现以及波动率策略的收益分析等，为投资人参与期权市场具有重要借鉴意义，为金融市场管理人员了解波动率变化趋势也具有重要参考意义。

通过理论研究的结果能够看出，Delta 中性期权交易波动率策略的利润源泉，大部分是来自隐含波动率变化与波动率风险溢价，但在较短时间的日频交易过程中，隐含波动率的变动对策略利润影响更大。本文的交易策略建设主要基础，便是通过对隐含波动率的改变走向作出预测并掌握隐含波动率趋势走向，确立波动率交易策略。经过本文的研究可以看出在目前金融市场中，主要有隐含波动率的历史走势，以及期权期限构成，而期货市场成交数量等这几种因子，对隐含波动率的变动具有很大的预知能力，这也就意味着在对短期波动率改变水准的预期中，对于短期内的成交冲击，或者短期内的波动率变化水准突变会对预期结果更加可靠。在普通的最小二乘回归、LSTM 模型和 Informer 模型这三个预测模型中，Informer 模型在对有隐含波动率变化的方向预测中具有最高的准确性，而 LSTM 模型在对较短序列隐含波动率变化的方向预测中具有较大的准确率，但局限于数据量较小，模型在对隐含波动率上升的方向预测中有效性较差。随着期权市场的深入发展，机器学习模型将在波动率交易中发挥更大的作用。

6.2 不足和改进方向

经过对文章的相关研究，波动率策略的总体构建原理和方式已相当清晰。文章的主要缺陷在于没有对对冲方法展开更广泛的对比分析，主要是采用现货对冲，设定对冲阈值和频率等，文章的下一个方面可能需要加入对对冲方法的研究，让波动率策略更加完备和细致。其次，文章中对于若干市场因子未展开研究，如期权市场流动性指数、投资人情绪指数等，文章在下一个研究方面还可以尽量多的加入多种因子。最后，文章的预测模型的估计准确度仍有进一步的提高空间，可以通过改进模型构建方法，以及研究更多的机器学习算法来进一步提高估计模型的准确度。

参考文献

- [1] Wold S, Martens H, Wold H. The multivariate calibration problem in chemistry solved by the PLS method[J]. Springer Berlin Heidelberg, 1983.
- [2] Kapadia B N . Delta-Hedged Gains and the Negative Market Volatility Risk Premium[J]. Review of Financial Studies, 2003, 16(2):527-566.
- [3] 郑振龙, 黄薏舟. 波动率预测: GARCH 模型与隐含波动率[J]. 数量经济技术经济研究, 2010, 27(01):140-150.
- [4] 屈满学, 王鹏飞. 我国波动率指数预测能力研究——基于隐含波动率的信息比较[J]. 经济问题, 2017(01):60-66.
- [5] 马天平, 吴卫星. 基于机器学习算法的金融期权波动率预测[J]. 学海, 2018(05):201-209.
- [6] 龚旭, 林伯强. 跳跃风险、结构突变与原油期货价格波动预测[J]. 中国管理科学, 2018, 26(11):11-21.
- [7] Aparna, Prasad, Bhat. The economic determinants of the implied volatility function for currency options: Evidence from India[J]. International Journal of Emerging Markets, 2018, 13(6):1798-1819.
- [8] 王茹婷, 李文奇, 黄诒蓉. 贸易摩擦、日内跳跃与股市波动——基于中国高频数据的经验证据[J]. 国际金融研究, 2019(12):63-73.
- [9] 吴鑫育, 侯信盟. 基于双因子已实现 GARCH 模型的波动率预测研究[J]. 运筹与管理, 2020, 29(12):207-214.
- [10] 李鸿翔. 投资者情绪与股票收益率的关系研究——基于上证 50ETF 波动率指数的实证分析[J]. 中国物价, 2020(08):74-77.
- [11] 刘勇, 白小滢. 投资者情绪、期权隐含信息与股市波动率预测——基于上证 50ETF 期权的经验研究[J]. 证券市场导报, 2020(01):54-61.
- [12] 张同辉, 苑莹, 曾文. 投资者关注能提高市场波动率预测精度吗?——基于中国股票市场高频数据的实证研究[J]. 中国管理科学, 2020, 28(11):192-205.
- [13] Zhang W, X Gong, Wang C, et al. Predicting stock market volatility based on textual sentiment: a nonlinear analysis[J]. Journal of Forecasting, 2021;1-22.
- [14] Lei B, Zhang B, Song Y. Volatility Forecasting for High-Frequency Financial Data Based on Web Search Index and Deep Learning Model[J]. 2021, 9(4):320-336.
- [15] Mandelbrot B. When can price be arbitrated efficiently? A limit to validity of the random walk and martingale models[J]. Review of Economics and Statistics, 1971, 53(2):225-236.

-
- [16]Hurst H .Long-term storage capacity of reservoirs[J].Transactions of the American Society of Civil Engineer,1951,116:770-799.
- [17]Giraitis L,Kokoszka P,Leipus R,Teyssiere G.Rescaled variance and related tests for long memory in volatility and levels[J].Journal of Econometrics,2003, 112:265-294.
- [18]Geweke J,Porter-Hudak S.The estimation and application of long memory time series models[J].Time Series Analysis,1983,4:221-238.
- [19]Robinson PM. Gaussian semiparametric estimation of long range dependence[J].The Annals of Statistics,1995b,23(5):1630-1661.
- [20]Robinson PM.Real and spurious long-memory properties of stock-market data: Comment[J].Journal of Business and Economic Statistics,1998,16(3),276-279.
- [21]Robinson PM.Log-periodogram regression of time series with long range dependence[J].The Annals of Statistics,1995a,23(3):1048-1072.
- [22]Robinson PM.Semiparametric analysis of long-memory time series[J].The Annals of Statistics,1994,22(1):515-539.
- [23]Shimotsu K.Simple (but effective) tests of long memory versus structural breaks[R]. Queens University,2007.
- [24]Shimotsu K,Phillips PCB.Exact local whittle estimation of fractional integration[J].The Annals of Statistics,2005,33(4):1890-1933.
- [25]Sowell F. Maximum likelihood estimation of stationary univariate fractionally integrated time series models[J].Journal of Economics,1992,53: 165-188.
- [26]Baollie RT,Kongcharoen C,Kapetanios G.Prediction from ARFIMA models: Comparisons between MLE and semiparametric estimation procedures[J]. International Journal of Forecasting,2012,28:46-53.
- [27]Vaswani A , Shazeer N , Parmar N , et al. Attention Is All You Need[J]. arXiv, 2017.
- [28]KITAEV N, KAISER Ł, LEVSKAYA A. Reformer:The efficient transformer [EB/OL]. [2020-3-11]. <https://arxiv.org/pdf/2001.04451.pdf>.
- [29]Guo Q , Qiu X , Xue X , et al. Low-Rank and Locality Constrained Self-Attention for Sequence Modeling[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2019, 27(12):2213-2222.
- [30]杨文瑞,沈韬,朱艳,曾凯,刘英莉.融合 ELMo 词嵌入的多模态 Transformer 的图像描述算法[J/OL].计算机工程与应用:1-15[2021-06-10].<http://kns.cnki.net/kcms/detail/11.2127.TP.20210527.0938.004.html>.
- [31]李铁飞,生龙,吴迪.BERT-TECNN 模型的文本分类方法研究[J/OL].计算机工程与应用:1-10[2021-06-10].<http://kns.cnki.net/kcms/detail/11.2127.TP.20210329.1358.011.html>.

-
- [32]Lo AW.Long-term memory in stock market prices[J].Econometrics,1991,59(5):1279-1313.
- [33]MNIH V, HEES N, GRAVES A. Recurrent models of visual attention[J]. Advances in neural information processing systems. 2014, 3(6): 2204-2212.
- [34]KATHAROPOULOS A, VYAS A, PAPPAS N, et al. Transformers are rnns: Fast autoregressive transformers with linear attention [C]//Proceedings of the 2020 International Conference on Machine Learning.Cham: Springer, 2020: 1-4.
- [35]王凯诚, 鲁华祥, 龚国良, 等. 基于注意力机制的显著性目标检测方法[J].智能系统学报, 2020,15(4): 1-9.
- [36]王亚飞. 带注意力机制的车辆目标检测与识别[D]. 上海: 华东师范大学, 2020: 25-31.
- [37]Zhou H , Zhang S , Peng J , et al. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting[J]. arXiv, 2020.07
- [38]吴鸿超,刘美尧,包悦妍.高频波动率预测模型在期权波动率套利中的比较分析——基于 50ETF 金融高频数据[J].产业科技创新,2020,2(01):54-56.
- [39]王西梅,赵延龙,史若诗,包莹.基于局部波动率模型的上证 50ETF 期权定价研究[J].系统工程理论与实践,2019,39(10):2487-2501.
- [40]李雪飞,赵冰,严高剑.波动率期限结构分析及在期权定价交易中的应用[J].证券市场导报,2019(02):19-25.
- [41]Li P, Lin Y, Zhong Y. Patterns of 50 ETF Options Implied Volatility in China: On Implied Volatility Functions. 2021,1(24):135-145
- [42]Li P, Xian A, Lin Y. What determines volatility smile in China? [J]. Economic Modelling, 2021,(96):326-335
- [43]杨昌辉,邵臻,刘辰,付超.一种混合建模方法及其在 ETF 期权定价中的应用[J].中国管理科学,2020,28(12):44-53.
- [44]姜晓晴.上证 50ETF 期权定价研究——基于 B-S 模型与 CEV 模型[J].吉林工商学院学报,2020,36(03):64-70+75.
- [45]孙有发,吴碧云,郭婷,刘彩燕.非仿射随机波动率模型下的 50ETF 期权定价:基于 Fourier-Cosine 方法[J].系统工程理论与实践,2020,40(04):888-904.
- [46]郭精军,宋彦玲.基于时间变换和分数型过程下的期权定价及模拟分析[J].应用概率统计,2020,36(01):59-70.
- [47]吴鑫育,赵凯,李心丹,马超群.时变风险厌恶下的期权定价——基于上证 50ETF 期权的实证研究[J].中国管理科学,2019,27(11):11-22.
- [48]Recchioni M C, Iori G, Tedeschi G, et al. The complete Gaussian kernel in the multi-factor Heston model: Option pricing and implied volatility applications[J]. European Journal of Operational Research, 2020,293(1):336-360

-
- [49] Roul P , Goura V . A compact finite difference scheme for fractional Black-Scholes option pricing model[J]. Applied Numerical Mathematics, 2021,166:40-60
- [50] Burghardt et al. How to tell if options are cheap[J]. Journal of Portfolio Management, 1990, 135(4):72-78.
- [51] 马天平,吴卫星.基于机器学习算法的金融期权波动率预测[J].学海,2018(05):201-209.
- [52] 刘晓倩,王健,吴广.基于高频数据 HAR-CVX 模型的沪深 300 指数的预测研究[J].中国管理科学,2017,25(06):1-10.
- [53] 王西梅,赵延龙,史若诗,包莹.基于局部波动率模型的上证 50ETF 期权定价研究[J].系统工程理论与实践,2019,39(10):2487-2501.
- [54] 李璐. 期权波动率策略在我国市场上的实证研究[D].山东大学,2018.
- [55] 施智敏. 上证 50ETF 波动率指数及期权交易策略的实证研究[D].上海师范大学,2018.
- [56] 谢佳珠. 基于不同波动率模型的期权跨式组合策略研究[D].上海师范大学,2018.
- [57] Britten-Jones M, Neuberger A. Option prices, implied price processes, and stochastic volatility[J]. Journal of Finance, 2000, 55(2): 839-866.
- [58] Whalley A E , Wilmott P . An Asymptotic Analysis of an Optimal Hedging Model for Option Pricing with Transaction Costs[J]. Mathematical Finance, 1997.
- [59] Valeri, Zakamouline, Steen, et al. Portfolio performance evaluation with generalized Sharpe ratios: Beyond the mean and variance[J]. Journal of Banking & Finance, 2009.
- [60] 彭程,李爽,包莹,赵延龙.外汇欧式期权在市场不完备下的对冲误差分析[J].系统工程理论与实践,2019,39(11):2739-2749.
- [61] Joshua D. Coval, Tyler Shumway. Expected Option Returns[J].The Journal of Finance,2001,56(3).

致谢

首先，我要感谢我的导师傅毅老师在我两年的研究生学习期间对我的教导，傅老师素来对学术抱有很高的热情，在对我们的教学中谆谆善诱，在论文的撰写过程中为我答疑解惑，指明方向，不仅为我提供了很多关于论文的研究思路与方法，同时也对论文的质量严格把关。不管是学习工作上还是生活中，傅毅老师都教会了我很多。

其次，我要感谢金融专硕的全体老师，是他们的悉心教导让我在研究生阶段丰富扎实了自己的专业知识水平，掌握了很多与专业相关的知识和技能，切实提高了自身能力。

最后，我还要感谢我的家人、亲友和同学，是他们赠予我的关怀、帮助与陪伴，支撑着我不断前进，顺利完成我的学业。