

分 类 号 _____

学号 M201774226

学校代码 10487

密级 _____

华中科技大学

硕士学位论文

基于遗传算法优化 Elman 神经网络的 黑色金属期货跨品种套利策略

学 位 申 请 人： 唐瑞波

学 科 专 业： 金融硕士

指 导 教 师： 简志宏 教授

答 辩 日 期： 2019.5

**A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree for the Master of Finance**

**Cross-Variety Arbitrage Strategy of Ferrous
Metal Futures Based on Genetic Algorithm
Optimization of Elman Neural Network**

Candidate : Tang Ruibo
Major : Master of Finance
Supervisor : Prof. Jian Zhihong

Huazhong University of Science & Technology

Wuhan 430074, P.R.China

May, 2019

独创性声明

本人声明所呈交的学位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。尽我所知，除文中已经标明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对本文的研究做出贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：

日期： 年 月 日

学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，即：学校有权保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权华中科技大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本论文属于 ☐ 保密， 在_____年解密后适用本授权书。
☐ 不保密。

（请在以上方框内打“√”）

学位论文作者签名：

日期： 年 月 日

指导教师签名：

日期： 年 月 日

摘要

随着我国经济的高速发展和相关交易制度的不断完善,我国期货市场已日趋成熟,并开始朝着国际化的目标迈进。越来越丰富的期货品种既为不同的行业提供了套期保值的工具,也为套利者提供了更多的交易产品。

本文选取黑色金属产业链中具有代表性的螺纹钢、铁矿石和焦炭期货作为跨品种套利的研究对象。区别于统计套利方法,在根据协整检验的结果选择出合适的套利组合并利用 GARCH 模型确定了套利比例后,运用在动态系统拟合上比 BP 神经网络更优的 Elman 神经网络对价差进行预测,并使用遗传算法对其进行优化。基于神经网络的预测,本文构建了不同固定阈值下的趋势策略,相关的获利指标表明:随着开仓阈值的提高,神经网络策略的盈利能力不断降低,原因在于阈值的提高使得策略的交易次数明显减少;基于遗传算法优化 Elman 神经网络建立的策略比 BP 神经网络和 Elman 神经网络能够产生更高的年化收益率,且收益率在长期内保持了稳定增长,这得益于遗传算法加快了 Elman 神经网络的收敛速度,增加了网络的拟合精度并使策略产生更多的有效盈利次数,除此之外,遗传算法令网络在训练时获得全局最优解,从而在大部分时间段的预测都能维持较高的准确度。

本文在最后对神经网络和期货品种的后续研究进行了展望,并且为不同的市场参与主体提供了针对性的投资建议,这使得本文在具备一定理论价值的同时又具有较强的现实意义。

关键词: 跨品种套利; Elman 神经网络; 黑色金属期货; 遗传算法

Abstract

With the rapid development of China's economy and the continuous improvement of the relevant trading system, China's futures market has become increasingly mature and begins to move towards the goal of internationalization. More and more abundant futures varieties not only provide hedging tools for different industries, but also provide more trading products for arbitrageurs.

This paper chooses the representative rebar, iron ore and coke futures in the ferrous metal industry chain as the research object of cross-variety arbitrage. Different from statistical arbitrage method, after choosing the appropriate arbitrage combination according to the results of cointegration test and using GARCH model to determine the arbitrage ratio, Elman neural network, which is better than BP neural network in dynamic system fitting, is used to predict the spread, and genetic algorithm is used to optimize it. Based on the prediction of the neural network, this paper constructs the trend strategy under different fixed thresholds. The relevant profit indicators show that with the increase of the opening threshold, the profitability of the neural network strategy decreases continuously, because the increase of the threshold makes the number of trades of the strategy decrease significantly. The strategy of Elman neural network based on Genetic algorithm optimization is better than that of BP neural network and Elman neural network. The network can produce higher annual rate of return, and the rate of return keeps stable growth in the long run. This is due to the fact that genetic algorithm accelerates the convergence speed of Elman neural network, increases the fitting accuracy of the network and makes the strategy generate more effective profit times. In addition, genetic algorithm enables the network to obtain the global optimal solution during training, so that the prediction can maintain a high accuracy in most time periods.

Finally, this paper prospects the follow-up research of neural network and futures varieties, and provides specific investment suggestions for different market participants, which makes this paper not only have certain theoretical value, but also have strong practical significance.

Key words: Cross-Variety Arbitrage; Elman Neural Network; Ferrous Metal Futures; Genetic Algorithms

目 录

摘要.....	I
Abstract.....	II
目 录.....	III
1. 绪言.....	1
1.1 选题背景和意义.....	1
1.2 国内外研究文献综述.....	4
1.3 论文的主要内容和研究方法.....	8
1.4 创新点和不足之处.....	10
2.跨品种套利的理论基础.....	12
2.1 跨品种套利的概念、原理和分类.....	12
2.2 协整理论.....	15
2.3 ARCH 和 GARCH 模型.....	19
3.跨品种套利神经网络模型的构建.....	22
3.1 神经网络的预测功能.....	22
3.2 前馈式的 BP 神经网络.....	23
3.3 反馈式的 Elman 神经网络.....	25
3.4 基于遗传算法优化 Elman 神经网络.....	27
4.实证研究.....	31
4.1 数据选取和描述性统计.....	31
4.2 协整检验.....	33
4.3 套利比例的确定.....	38
4.4 基于神经网络拟合价差序列.....	40
4.5 套利的实现.....	48

5.结论.....	58
致谢.....	61
参考文献.....	62
附录.....	66

1. 绪言

1.1 选题背景和意义

1.1.1 选题背景和目的

在过去 30 年的时间里，我国期货市场经历了起步探索阶段、治理整治阶段、规范发展阶段，从无到有，从小到大，从无序变为有序，逐步发展并趋于成熟。在此前经历了两年的疲软态势之后，2018 年我国期货市场出现一定反弹，虽然累计成交量发生了极小幅度的下滑，但累计成交额却有了大幅度的攀升，在这之中，商品期货仍然占据主导地位，成交量和成交额分别占到总成交量和总成交额的 99.10% 和 87.61%。与此同时，我国期货市场开始逐步朝着国际化的方向迈进，3 月 26 日，我国首个引入境外交易者的期货品种—原油期货在上海国际能源中心挂牌上市，随后的 5 月 4 日和 11 月 30 日，铁矿石期货和 PTA 期货也成功引入境外交易者。可以说，整个市场在将来仍具有广阔的壮大空间。

近些年来，我国商品期货的品种日益丰富，基本覆盖了大部分的大宗商品，除了上面的原油期货外，2018 年我国还上市了纸浆期货、乙二醇期货两个商品期货新品种，目前已上市的商品期货品种已达 51 种。我国的经济体量正处于迅速扩张时期，无论是农产品产量，还是钢材加工量都十分巨大，这意味着许多行业都要对价格变动的风险进行管理，否则会影响到自身的效益。期货市场的主要功能在于价格发现和套期保值，这使得它与现货市场保持着极其紧密的联系，而当前我国的期货品种还不能完全满足实体经济发展的需要。不过相信随着我国期货市场国际化的推进，未来更多满足套期保值需求、市场条件具备的新的商品期货品种会被陆续推出，期货市场也将在风险管理方面发挥更为重要的作用。

作为商品期货市场的重要组成部分，目前我国黑色金属产业链期货上市品种已多达七个，基本覆盖了钢铁生产流程中所涉及的所有原料及主要初级产生品，虽然最早上市的螺纹钢期货距今才有 9 年，但主要品种的上市时间都在 3 年以上，市场

整体运行基本平稳，结构日趋完善。从近些年的发展趋势看，除了原有的钢铁厂外，如今中间贸易商乃至下游需求厂商也参与到黑色金属期货市场，使得该产业链期货的成熟度越来越高，同时受到市场的广泛认可。

在期货市场上，除了套期保值者外，套利者也是其中不可或缺的一大部分。套利者的加入能够加大期货的交易量，增强市场的流动性，使期货价格更具真实性和连续性。对于黑色金属产业链，跨品种套利是一种极为重要的套利形式，它是利用具有内在联系的商品期货之间短期内的价格差异所进行的套利活动，具体可分为相关商品套利和产业链套利，对于相关商品套利，螺纹钢和热卷都是钢铁厂重要的产成品，从原材料锻造成钢坯的过程均相同，只是从钢坯轧制成了不同形状特性的钢材，所以二者的价格走势具有较强的相关性，可以用于套利，对于产业链套利，铁矿石和焦炭都是冶炼螺纹钢的重要原料，三者的价格受到成本和利润的牵制，在形式上更为稳定，故而成为套利者的首选组合。

套利的方法是多样的，区别于传统基于均值回归思想的统计套利，人工智能领域的神经网络模型逐渐进入了相关研究者的视野。虽然人工智能的概念很早就被提出，但其发展历经曲折，不过进入 21 世纪的第二个十年后，人工智能在算法、计算能力方面都取得了较大的突破，从而成为当前最热门的科技焦点，例如 2016 年，AlphaGo 战胜人类围棋选手李世石，轰动一时，使得人们对 AI 的发展既满怀期待，又有所担忧，但大部分国家都意识到 AI 在未来国家科技实力竞争中的重要性，在这一方面，发达国家已走在前列，而我国也不甘落后，百度、腾讯、小米等企业纷纷推出各自的 AI 产品，以求占据未来科技革命制高点。

目前，人工智能已在多个领域发挥了重要的作用，如医疗卫生、金融、商业、安防等，在金融领域，智能投顾就是就是人工智能技术和金融理论融合的代表产物。作为人工智能的重要分支，神经网络也在各个行业得到了广泛的应用，并可用于金融时间序列的预测，基于此，那么将神经网络模型引入到跨品种套利的研究之中，不仅仅是促进了各个学科之间的交流，更是在如今人工智能大发展的时代下，对于以往金融学方法的一种变革与创新。

结合上述背景，本文的选题是出于以下几个方面的考虑：首先，在套利品种的

选择上，我们着重关注了黑色金属产业链的期货，这是因为经过大量的文献阅读，我们发现对于该产业链跨品种套利的研究相对偏少，可能是由于产业链中相关期货的上市时间不长所导致，通过本文的研究，我们希望能够发掘其中潜在的获利机会。其次，在套利形式的选择上，跨品种套利一般分为相关商品套利和产业链套利，对比相关商品套利，产业链套利的形式更为稳定，也更具实际意义，故本文选取了黑色金属期货中具有代表性的螺纹钢期货、铁矿石期货和焦炭期货展开实证研究。接着，在套利模型的选择上，为了对传统的套利策略和思想进行创新，本文并没有选用基于均值回归思想的统计套利模型，而是顺应了人工智能发展的潮流，引入了神经网络模型对价差进行预测，并在此基础上构建趋势策略。最后，在具体神经网络模型的选择上，虽然 BP 神经网络是最广泛使用的模型，但无法对动态系统作出有效拟合和预测，故本文采用了反馈式的 Elman 神经网络，并针对其收敛速度慢和易陷入局部最小值的缺点，利用遗传算法进行改进。

1.1.2 研究意义

从理论意义上看，在以往大部分的关于跨品种套利的理论研究中，都是以协整理论作为基础，本文则在此基础上运用了 ARCH 和 GARCH 模型，其目的在于消除异方差因素的干扰，确定不同种相关的期货之间的真实的套利比例，除此意外，本文所构建策略的模型基础不在以均值回复为中心思想，而是采用了跨学科研究的方法，引入人工智能领域的神经网络，实现对价差的有效预测，区别于传统的 BP 神经网络，本文选用的 Elman 神经网络具备极强的适应序列时变特性的能力，并用遗传算法改进其收敛速度慢和易陷入局部最小值的缺点，在此基础上建立的套利策略类似于趋势策略，是对传统套利策略的一种改进和创新。无论是在理论还是在模型方面，本文都极大丰富了跨品种套利领域的相关研究，对于后续从事该领域研究的学者提供了新的思路和改进的方向。

从实践意义上看，虽然我国期货市场仍在不断地发展，相关的交易制度也陆续得以完善，但期货市场很可能还只是处于弱式有效状态，需要提防市场的交易风险。尽管本文所选择的三种黑色系金属期货上市时间都不超过 10 年，可却是对市场交

易品种的一大丰富，为套期保值者提供了规避现货价格变动风险的交易机会，对于套利者而言，他们能够通过自身所掌握到的关于这些期货价格变动的信息来发掘其中潜在的获利机会，这是本文研究的一大出发点，我们希望在上市时间不长的期货品种上寻找新的赢利点，同时在方法上有所创新并成功应用于实践之中。本文建立的神经网络模型可能是对机构投资者在期货市场上进行跨品种套利策略的一个补充，适当忽略其中的一些假设条件，本文仍具有较强的实践意义。

1.2 国内外研究文献综述

1.2.1 国外文献综述

根据国内外研究跨品种套利的文献，整体来看，跨品种套利主要有统计套利和神经网络模型两种方法。按照套利的原理不同，跨品种套利可分为相关商品套利和产业链跨品种套利。

相关商品套利利用的是商品期货价格之间的相关性，相关性程度越高则不同期货价格的走势越一致。Wahab et al. (1994) 以及 Liu 和 Chou (2003) 都通过协整分析揭示了黄金和白银价格之间的比例关系，前者运用移动平均模型建立了跨品种套利的策略，后者则根据误差修正模型对二者的价差进行预测，并获得显著的无风险收益。

由于产业链跨品种套利在形式上更为稳定并且具有一定的现实意义，故以该类型的研究居多。在这一类型中，对大豆及其压榨品豆油和豆粕所组成的产业链研究最早也最为深入。Tzang 和 Leuthold (1990) 较早对大豆压榨过程进行了套利研究，发现对大豆和豆粕或大豆和豆油进行套期保值可以降低风险和持有成本；Johnson et al. (1991) 则是在研究统计套利的过程中，对大豆压榨价差进行了实证分析，但是其并没有给出一种适合日常交易的平仓方式，只是在持有一段时间后平仓；在此基础上，Rechner 和 Poitras (1993) 选取期货市场的日数据模拟大豆压榨过程，在忽略压榨成本的前提下，以 GPM 为标准，通过合理设置 GPM 的上下限并在规定的交易日进行平仓获得了较为理想的盈利；Simon (1999) 引入异方差模型研究了大豆压榨的利润，认为压榨价差的短期偏差最终会回到长期均衡，可以通过均衡状态的

偏差回归来获利；Mitchell 和 John（2010）对 Simon（1999）的工作进行了回顾，并通过研究发现，大豆压榨价差短期偏离长期均衡不一定会以震荡的方式进行回归，可以通过调整开平仓条件来增加价差的盈利比例。除了大豆压榨价差外，学者们还对原油及其裂化产成品的价差进行了研究，且在研究方法上有所创新。Girma 和 Paulson（1998）通过研究确定了原油、汽油和取暖油期货价格之间的协整关系后，利用移动平均线和标准差构建策略进行模拟交易，获利效果显著；与之类似，Haigh 和 Holt（2002）运用 MGARCH、BGARCH 等模型对上述期货进行套利效果的对比研究，结果发现 MGARCH 模型在交易中的获利能力最强；Cummins 和 Bucca（2011）则在前两者的基础上，增加了纽交所上市天然气期货作为研究对象，应用一种新的统计套利模型进行量化交易的实证研究，套利的结果显示盈利策略在大部分情况下的夏普比率都大于 2。

随着人工智能、机器学习研究的不断深入，国外学者逐渐开始重视利用神经网络模型对跨品种套利展开研究，这是因为神经网络具备良好的联想记忆和预测能力，能够实现对价差的有效预测。Dunis et al.（2006）将传统的回归分析作为更复杂模型的基准，运用多层感知器（MLP）、循环神经网络（RNN）和高阶神经网络（HONN）进行大豆压榨价差的交易，最终的套利结果表明，在样本内和样本外的交易费用发生之前，最优模型是 RNN，它产生比所有其它模型更高的风险调整收益，然而，如果考虑了价差交易的成本，对于大多数模型而言，任何利润潜力几乎都会被消除；Phoebe 和 David（2014）则利用径向基（RBF）神经网络和 LM 神经网络预测了大豆压榨价差在短期的变化，发现到期日较远的期货会影响到到期日较近的期货，从而导致价差出现波动，可依据此制定套利策略；Dunis et al.（2015）还运用神经网络研究了玉米/乙醇的压榨价差，将遗传规划算法（GPA）同多层感知器和高阶神经网络对比，在波动率较低的期间，使用具有杠杆作用的时变波动过滤器增强价差的交易表现，同时在高波动性时期避免交易，模拟了玉米/乙醇压榨价差交易，发现在没有杠杆作用时，GPA 实现了最高的风险调整收益，引入时变杠杆策略后，GPA 仍是最有利可图的模型。

综合上述文献，国外学者对跨品种套利两种方法的运用已经趋于成熟，但仍然

存在改进的空间。

1.2.2 国内文献综述

国内学者在跨品种套利的研究中所使用的方法仍主要是统计套利，且已十分深入，但效果各有不一。

在相关商品套利方面，汪媛和孙建明（2013）选用沪铜和沪铝连续作为研究对象，提出利用时间差和 MACD 指标进行套利操作的方法，获得比较稳定的盈利。殷晓梅等（2008）借助 Granger 因果检验考察了棕榈油和豆油期货价格的相互影响程度，认为二者之间的套利可行；基于此王丽（2017）构建了最优套利策略，获得极高的年化收益率；吴静雯（2018）则是对豆油、菜籽油和棕榈油的组合进行了研究，她利用价差波动建立了相应的指标，并考察了其均线联系，基于此进行套利。

在对大豆压榨价差套利的研究已经形成一套完整体系的基础上，学者们注重于对套利策略进行优化和引入新的方法展开研究，三种期货的价格资料都来源于我国尚未成熟的期货市场，但策略所取得的收益并不显著。武军伟和史永东（2008）利用压榨利润 5 日和 20 日的移动平均线设计了多头套利和空头套利策略，研究结果显示只有基于 20 日均线的多头套利和空头套利策略收益才都显著为正；顾全和雷星晖（2015）也有同样的发现，虽然豆类期货之间存在长期均衡关系，但利用三者进行统计套利的盈利能力并不明显；翁泽波（2014）则将研究的重心放在套利策略阈值的确定上，利用构建期望套利收益对于交易阈值的函数并对其求最大化的方法求出非对称的上下交易阈值，采用 VaR 方法求得风控阈值；张威波和胡艳英（2018）运用 VAR 模型进行了实证研究，他们着重考察了豆类期货之间的相互影响关系，结果发现，豆油对豆粕、豆粕对大豆存在明显的影响关系，而其他之间则存在微弱的影响关系。部分学者还尝试采用该产业链中两种期货进行套利。丁华玲和华仁海（2007）研究发现利用大豆与豆粕进行套利的获利能力并不强；在此基础上，陈承旭（2010）优化了原来的模型，使其具有较高的收益率和稳定性；殷晓梅（2009）则建立了趋势套利模型，她借鉴证券投资学的部分思想，设置了一系列的高低点，通过这些点来推测价差的变化完成套利。

近些年来,随着我国期货市场的不断发展和完善,越来越多新的商品期货品种被陆续推出,学者们逐渐开始关注对黑色金属产业链的研究。周亮(2017, a)根据 JB 统计量确定了螺纹钢、铁矿石和焦炭三者的套利组合,以此构建的策略在样本内外均取得了不错的收益;周亮(2017, b)还选取 I1705 和 RB1705 这两种期货合约的分时数据,分别以协整模型和 GARCH 模型为基础构建了均值回归策略,最终套利结果显示相比于一般的协整模型, GARCH 模型因能捕捉到价差的波动效应,所以具有更高的胜率;王镭和陈芳(2017)选取南华商品指数,用 HP 滤波方法将期货市场分为牛市和熊市,在两类市场中基于焦炭和螺纹钢分别构建了套利策略,发现牛市的套利效果好于熊市。总体而言,国内对于该产业链的研究仍然不够深入,其中还存在着诸多潜在的套利机会需要不断挖掘。

由于均值回复策略已经遇到了一定的瓶颈,我国学者逐步开始运用神经网络对跨品种套利进行研究,虽然此类研究起步的时间较晚,但仍取得了一定的成果。刘建和等(2016)在分析了豆类期货之间存在的长期均衡关系,并根据 Madhavan 和 Smidt(1991)提出的半周期公式对短期动态调整过程进行了估计后,分别利用均值回归策略和 Elman 神经网络策略进行套利,结果显示两种策略均能获得正向的套收益率,但是在不同的开平仓阈值下, Elman 神经网络的表现更优;靳朝翔等(2016)运用 BP 神经网络模型和 NAR 动态神经网络模型对焦炭、铁矿石和螺纹钢三者构建套利策略的效果进行了对比,他们也设置了一系列的阈值,结果显示 NAR 动态神经网络模型具备更精准的预测能力,故无论在何种阈值下,其收益率的表现都是更优的一方。

总体来看,在统计套利方法在占据主流的情况下,并未取得完全理想的套利收益,我国的期货市场不同于国外成熟的期货市场,可能还是处于弱势有效的状态,国外学者在研究期货的跨品种套利时,都是建立在期货市场达到相对有效的基础上,两种套利方法皆能取得比较好的效果,但对于我国的期货市场,当期货价格发生较大波动时,统计套利方法过分依赖于均值和方差等参数,以及价差迟迟未回复到均值,套利原理无法实现的缺点就暴露了出来,所以我们需要尝试采用神经网络模型去构建新的套利策略,不同于统计套利,神经网络方法并不依赖于外部参数和

价差回复这一特定条件，只需根据价差的涨跌就能完成套利。与已有的国内研究相比，本文并未选取大豆产业链作为研究对象，这是由于相关学者对该产业链的研究非常多，我们需要从新的期货品种中挖掘投资机会，于是本文选择了黑色金属产业链，部分学者已经用统计套利方法对产业链中的部分期货品种实现套利，获得了比较好的效果，与此不同，本文采用神经网络模型对黑色产业链进行跨品种套利的研究，同时并未选用最被广泛使用的 BP 神经网络，而是选择在动态建模方面更加具有优势的 Elman 神经网络，虽然该网络的结构也相对简单，可依旧能够达到良好的预测效果，我们还针对其缺点进行了进一步的优化和改进，这是本文与现有研究的联系以及不同之处。

1.3 论文的主要内容和研究方法

1.3.1 本文的主要内容和结构

本文的主要内容分为以下五个部分：

第一部分为绪论，包括研究目的和意义、国内外研究文献综述、论文的主要内容和研究方法以及研究的创新点和不足之处。

第二部分为跨品种套利的理论基础，介绍了跨品种套利的概念、原理和分类，引入了跨品种套利中需要用到的协整理论和 GARCH 模型。

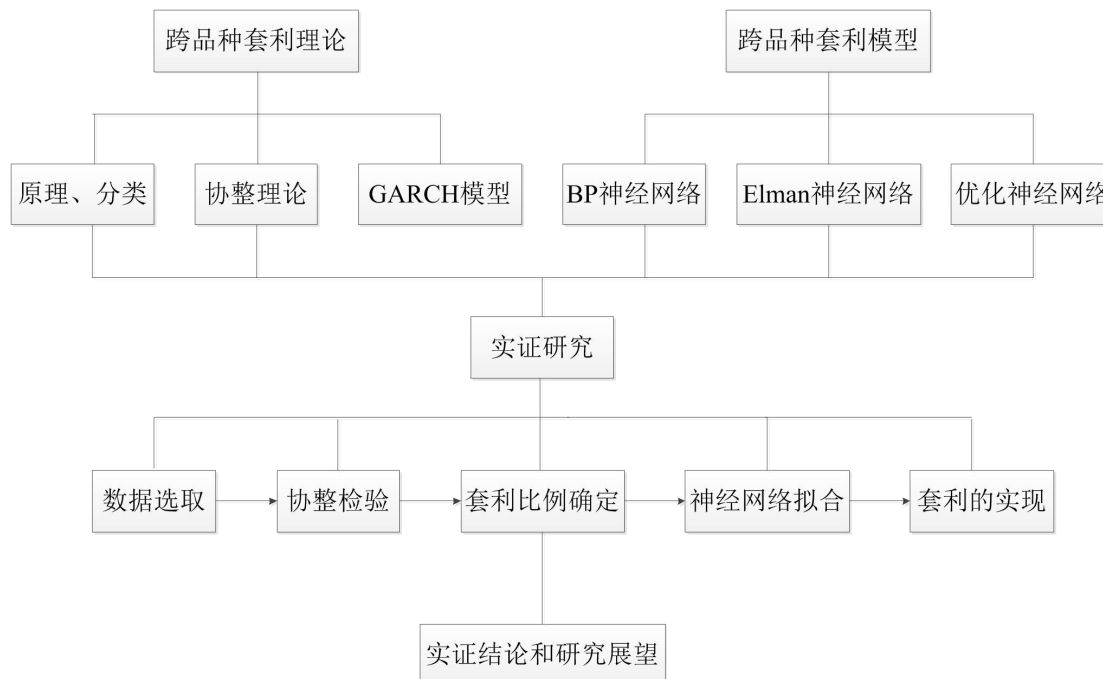
第三部分为跨品种套利神经网络模型的构建，介绍了本文构建跨品种套利策略的模型基础，即神经网络，该部分首先对神经网络的预测功能进行了介绍，然后依据神经网络的分类，着重概述了前馈式和反馈式网络中具有代表性的 BP 神经网络和 Elman 神经网络，并针对神经网络的固有缺点，利用遗传算法对性能更优 Elman 神经网络进行了优化。

第四部分为实证研究，需要将理论和模型应用于实际的套利过程中，该部分包括黑色金属期货高频数据的选取、套利组合的协整检验、套利比例的确定以及构建具体的策略实现套利。

第五部分为结论，对全文的内容进行总结并对下一步工作进行展望。

综上，本文研究的结构图大致如下：

图 1.1 本文研究思路图



1.3.2 研究方法

本文所采用的主要研究方法有：

一．文献研究法。通过阅读大量的文献，本文总结出国内外学者在研究跨品种套利时所采用的模型与方法，分析相关研究存在哪些缺陷和不足之处，根据这些总结和分析确定了本文的研究对象、方法和思路。

二．跨学科研究法。本文在研究跨品种套利时选择的模型为神经网络，神经网络的本质是一种深度的机器学习，机器学习是人工智能的核心，所以本文是将人工智能领域的方法应用到金融学的研究中，这使得学科之间产生了有益的交叉。

三．实证研究法。在建立了跨品种套利相应的理论基础和神经网络模型之后，本文以黑色金属产业链中三种代表性的期货作为研究对象，尽量模拟真实的市场环境构建了套利策略，使得本文具有在实践中应用的意义。

四．对比分析法。为了突出本文所建立的遗传算法优化 Elman 神经网络模型在构建套利策略时的优越性，本文将 BP 神经网络和 Elman 神经网络作为比较基准，

通过设立一些相关的指标来评价三种神经网络模型的获利能力。

1.4 创新点和不足之处

1.4.1 本文的创新点

本文有以下几个创新点：

一．套利方法。在跨品种套利的方法选择上，统计套利几乎占据了主流，国内的文献研究虽然在策略上不断改进，但仍以均值回复为基本套利思想。本文所采用的神经网络不仅仅构成了一种新的套利模型基础，更是在核心思想和产生交易信号的原理上与统计套利存在本质的区别。

二．模型改进。受制于学科交叉等多种因素的影响，学者们在考虑利用神经网络的预测功能时，往往会选择结构最为简单的 BP 神经网络，但该网络存在不少的缺陷，预测能力有限，特别是在我们要对动态的金融时间序列进行拟合和预测时，预测精度偏低，于是本文引入了性能更好的反馈式 Elman 神经网络进行价差的预测，同时运用遗传算法改进其收敛速度慢和易陷入局部最小值的缺点，取得了很好的效果。

1.4.2 本文的不足之处

整体来看，本文存在以下两点不足之处：

第一，在人工智能大发展的时代，神经网络更新速度也会不断加快，这主要在于其结构和算法上的创新，故可以预见的是神经网络的预测功能会逐渐地增强。本文所选用的 Elman 神经网络虽然在结构上比一般的 BP 神经网络更为复杂，预测精度也更高，但同现在主流的神经网络模型相比，仍然存在一定的差距，所以无论是在 Elman 神经网络的结构上有所改进，还是尝试用一种最新的神经网络来替代目前的模型，都将成为我们接下来主要的研究方向。

第二，本文在构建套利策略前，设置了诸多的假设条件，虽然这些条件能够帮助我们简化理论研究，可是从实际交易的角度出发，依然存在一些不妥之处。首先高频交易的一大特点就是在于交易次数比较多，这样会使手续费大量增加，极有可

能覆盖掉收益，本文在设定手续费率时考虑到了这一点，所以设置的手续费率偏小，不是特别符合真实情况，在今后的研究中，其实可以通过延长交易时间、提高开仓阈值以及寻找更优的模型等手段来减小手续费在获利中所占的比例。其次，本文利用 GARCH 模型确定了三种期货之间的套利比例，但是由于存在小数点，要求期货买卖不能按照完整的手数来进行，所以本文就假定期货合约是可分割的，显然这与实际的套利过程存在一定的出入，需要我们想办法改进。最后，实际交易远比理论研究来的复杂，比如我们在买卖期货的过程中，因为硬件或认为的因素导致下单价格与实际的成交价格有一点差异，这就是所谓的滑点，由于滑点的存在，我们需要尽量避开一些波动剧烈的行情。总而言之，减少假设条件、尽量贴近真实交易是我们在策略的研究中一直所追求的方向。

第三，本文只是将建立的神经网络模型运用到黑色金属期货的跨品种套利中，对于其它期货品种，该模型是否依然具有实用性，遗传算法优化的 Elman 神经网络是否仍然较 BP 神经网络和 Elman 神经网络更优，其稳定性如何，这都有待我们后续的进一步研究。

2.跨品种套利的理论基础

2.1 跨品种套利的概念、原理和分类

2.1.1 跨品种套利的概念和原理

在商品期货市场上，套利主要是指利用不同种期货合约价格差异的变化，同时进行方向相反的开仓交易，并在合适的时机将所有合约对冲平仓获利。按照研究对象的不同，期货套利可以分为跨期套利、跨品种套利和跨市场套利，它们分别是根据交割月份、商品种类和市场的不同开展操作，以此达到获利的目的。

跨品种套利，具体来讲，是指利用存在内在联系的商品期货合约实施套利，即在买进某种商品期货合约的同一时间点卖出与之价格走势类似的另外一种商品期货合约，当价格变动有利时将两种合约同时分别平仓获利。从本质上看，套利者的并非是在利用单个期货价格的变化来获利，他们更多地是在考察不同期货之间的价差波动，因此期货市场上的套利行为又被称为价差交易。

跨品种套利的原理在于，不同的商品因其内在的某种联系，它们的价格在长期内存在着稳定合理的关系，但这些有关联的商品的价格有时会偏离合理的区间，表现出一种商品期货价格被高估、另一种商品期货价格被低估的情况，从而产生了套利的机会。本文认为影响不同种商品期货之间的价格关系均衡的因素主要可以分为以下两类：

1.市场因素。期货市场的一大功能在于价格发现，也就是说，期货交易所聚集了众多的套期保值者和套利者，他们把自己所了解到的某种商品的行情信息集合到交易场内，在这样一个公开、透明的市场上，产生的期货价格具有真实性和预期性，一方面能够实际反映该商品的供求关系，另一方面可以体现人们对其未来价格走势的预期。如果市场是有效的，那么期货价格应该对新的信息及时作出反应，可从目前我国期货市场的成熟度看，价格对新信息的反应可能出现过度或延迟的情况，这就形成了期货价格之间的强弱对比关系，为投机者提供了机会。

2.外部因素，具体包括政策、季节等。从最为典型的季节因素看，在一年之中，对于农产品而言，从播种、成长到收获，都是有固定的时间节点，除此之外，每年的春季来临，钢铁行业总会迎来销售旺季，也就是说，无论是对于大宗商品，还是基建材料，它们的价格变化总会呈现出一定的规律性，考虑到现货与期货市场的关联性，故不同季节的期货合约价格的表现也会有强有弱，如 Simon（1999）通过实证研究发现，受季节性因素的影响，豆粕和豆油的价格相对于大豆价格会持续上升，但大豆压榨价差偏离其长期均衡只是暂时的，存在回复到最近 5 天平均水平的趋势，基于此制定交易策略是有利可图的。

假设商品期货合约 1、2 的价格分别表示为 m 和 n ，并且长时间内保持的线性关系为 $n = km + b$ （ k 、 b 为常数）。如果在 t_1 时刻有 $n_1 > km_1 + b$ ，那么套利者的策略有如下两种：

1.如果套利者认为短期内价差仍会扩大，选择买入 1 手商品期货 1，卖出 k 手商品期货 2，待到 t_2 时刻，以下条件可能满足时：

$$n_2 - km_2 > n_1 - km_1 \quad (2.1)$$

他会立刻卖出 1 手商品期货 1，买入 k 手商品期货 2，在不考虑手续费的情况下，他的盈利为 $n_2 - n_1 - k(m_2 - m_1)$ 。同样地，如果套利者认为短期价差会缩小，选择卖出 1 手商品期货 1，买入 k 手商品期货 2，等到以下条件可能满足时：

$$n_2 - km_2 < n_1 - km_1 \quad (2.2)$$

他会即刻买入 1 手商品期货 1，卖出 k 手商品期货 2，在这种情况下，他的盈利为 $n_1 - n_2 - k(m_1 - m_2)$ 。

2.即使是在半强有效或弱有效的市场，不同商品期货间的价格关系即使在短期内被打破平衡，可是从长期看总会逐渐趋于合理，此时价差会回到原来的均值水平（一般接近于 0），套利者就卖出 1 手商品期货 1，买入 k 手商品期货 2，并等到以下条件得到满足时：

$$n_2 \leq km_2 + b \quad (2.3)$$

他会立即买入 1 手商品期货 1，卖出 k 手商品期货 2，此时他的盈利将大于等于 $n_1 - km_1 - b$ 。

策略 2 是利用价差的回复特性来赚取收益，相较于策略 1，其获利更具有必然性和确定性，但无论是哪一种策略，它们都是利用期货价格关系脱离合理区间的机会获利。

2.1.2 跨品种套利的分类

根据套利原理的不同，跨品种套利可相应地分为相关商品套利和产业链跨品种套利。

相关商品套利利用的是商品期货价格之间的相关性，不同商品期货价格的相关性越强，其走势就越接近。例如在油脂类期货中，豆油和棕榈油的价格有很强的相关性，当某一时刻其中一种合约相对于另一种合约的价格发生了明显的偏离时，我们就可以买卖这一价差，待其恢复平均水平时再进行相反的操作。除此之外，农产品期货中的玉米和小麦，金属类期货中的黄金和白银、铜和铝，以及棉花与 PTA 等，都可以作为套利的对象。这些能够进行配对交易的商品往往具有类似的属性，大都属于替代品的关系，意味着当一种商品的价格上升（下降）时，它的需求随之减少（增加），而可用于替代的另一种商品需求则增加（减少），带来其价格的上升（下降），也就是说，两种商品的价格走势将趋于一致，使得套利存在可行性。一般来说，相关商品套利的效果取决于商品之间相关性的强弱程度，这就需要在选择套利组合时深入研究各品种的基本面，找到价差回归的必然因素，并在交易过程中控制好风险和仓位规模，严格止损。

与相关商品套利不同，产业链跨品种套利的原理为：在同一产业链上，不同的品种因存在上下游关系，受制于成本和利润的约束，故它们的价格变动基本保持一致。与相关商品套利相比，这一形式被赋予更多的实际意义且更加稳定。最典型的例子是大豆及其压榨产品豆粕和豆油之间的套利，又被称为大豆提油套利和反向提油套利。大豆压榨利润 C 可以表示为：

$$C = \text{豆油价格} \times \text{出油率} + \text{豆粕价格} \times \text{出粕率} - \text{大豆价格} - \text{加工费} \quad (2.4)$$

将这一公式运用于期货市场套利，在不考虑加工费的前提下，出油率和出粕率可直接由大豆压榨企业提供得到，由此得到三者的套利比例为：

$$\text{大豆：豆油：豆粕}=1:\text{出油率}:\text{出粕率} \quad (2.5)$$

此时大豆压榨利润 C 正好就是套利的盈利，期货市场上的交易就等同于现货市场的实际生产过程，从另一个角度看，这反映出三种期货价格之间的内在联系更具稳定性，套利的效果可能会更好。当收益 C 超过正常压榨利润时，套利者可以选择进行大豆提油套利，即买入大豆期货，卖出压榨品期货；当收益 C 低于正常压榨利润时，则可进行反向大豆提油套利，即卖出大豆期货，买入压榨品期货。同样，原油裂化套利利用的是原油及其主要加工制成品汽油和取暖油的价差关系，在期货市场上遵循 $X:Y:Z$ 的套利比例，这里的套利比例采用裂化比例，表示原油的产出结构，即 X 桶的原油对应 Y 桶的汽油和 Z 桶的取暖油，受限条件为 $X=Y+Z$ ，比较广泛使用的裂化比例有 3: 2: 1、5: 3: 2 和 2: 1: 1，其中 3: 2: 1 被使用最多，而一个汽油产出量稍低的炼油厂可能会采用 5: 3: 2 的组合。如果原油期货的价格相对于成品油期货的价格上升，也就是裂化价差（理论上的炼油收益）减小，那么套利者会买入裂化价差，即卖出原油合约的同时买入成品油合约，如果裂化价差增大，他会进行相反的操作。无论是大豆提油套利还是原油裂化套利，它们都是被大豆压榨企业和炼油厂用于对冲现货市场价格变动风险的重要手段，以此锁定压榨利润和炼油收益。正是由于这个特点，产业链跨品种套利被更多的用于实践和研究之中。

2.2 协整理论

2.2.1 协整的概念

跨品种套利的基础在于不同商品期货的价格存在长期均衡关系，因此我们需要引入协整的概念。

首先，我们要介绍单整的概念。考虑一个时间序列 $\{x_t\}$ ，如果它是平稳的，这时称序列 $\{x_t\}$ 为零阶单整序列，简记为 $x_t \sim I(0)$ ，假如序列 $\{x_t\}$ 经 1 阶差分后后平稳，

则称序列 $\{x_t\}$ 为 1 阶单整序列，也称为单位根过程，记为 $x_t \sim I(1)$ ，更一般地，假如序列 $\{x_t\}$ 至少需要进行 d 阶差分才能实现平稳，这时称序列 $\{x_t\}$ 为 d 阶单整序列。

在金融学领域的研究中，大多数时间序列是非平稳的，但它们可能同步，要考察不同变量间的内在联系，就必须建立回归模型，可这一过程中经常会出现伪回归的结果，即回归显示显著相关的变量实际上根本不相关，为了解决此类问题，我们引入协整理论，下面给出协整的定义：

设向量 $X_t = (x_{1t}, x_{2t}, \dots, x_{kt})$ 的各分量同为 d 阶单整，如果存在一个非零向量 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)$ ，使得线性组合 $\alpha X_t^T \sim I(d-b)(d \geq b \geq 0)$ ，则称 X_t 是 (d, b) 阶协整的，记为 $X_t \sim CI(d, b)$ ，向量 α 被称为协整向量。

2.2.2 协整检验的方法

在我们想要对多个变量建立回归模型时，首要步骤就是要对它们进行协整检验。协整检验的方法主要有 Engle-Granger 检验法和 Johansen 检验法，下面我们对这两种检验方法作出详述。

1. Engle-Granger 检验法

为了检验变量间是否存在协整关系，Engle 和 Granger (1987) 提出了 EG 检验法，又称 EG 两步法，实则是给出了协整的另一种定义。

设变量 y_t, x_t 都是 $I(1)$ 序列，并建立如下的回归模型（称为协整回归）：

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \quad (2.6)$$

这里的 ε_t 被称为非均衡误差，它描述的是对长期均衡的偏离程度。若长期均衡存在，意味着系统中的内在因素能够相互制约，从而达到某种平衡，即使变量在短期内偏离了均衡，那么系统的调节机制将会在以使其迅速回复到从前的平均水平， ε_t 会围绕 0 值附近波动。也就是说， ε_t 应当是一个平稳时间序列，满足 $\varepsilon_t \sim I(0)$ 且 $E(\varepsilon_t) = 0$ ，由此推得 $\varepsilon_t = y_t - \beta_0 - \beta_1 x_t \sim I(0)$ ，按照上述协整的定义，我们就可以说 y_t 和 x_t 是

(1,1)阶协整的, 即 $y_t, x_t \sim CI(1,1)$ 。

EG 检验法的关键就在于确定非均衡误差 ε_t 是否是平稳的, 具体的检验步骤为:

(1) 利用 OLS 法估计出模型参数 $\hat{\beta}_0, \hat{\beta}_1$, 计算回归方程的残差 e_t 作为非均衡误差 ε_t 的估计值:

$$e_t = y_t - \hat{y}_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t \quad (2.7)$$

(2) 检验残差序列 e_t 的平稳性, 常用的有 DF 检验或 ADF 检验。

2. Johansen 检验

如果模型中的变量在两个以上时, 如果仍采用 EG 检验法, 就必须以所有的变量各自为因变量, 其它变量为自变量, 建立多个回归方程来考察它们之间的协整关系, 因为此时可能存在多个协整向量。为了简化上述检验步骤, Johansen 和 Juselius (1990) 建立了一种针对多变量协整关系新的检验方法, 一般称为 Johansen 检验。

设 VAR 模型表示如下:

$$Y_t = c + \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \cdots + \Pi_p Y_{t-p} + u_t \quad (2.8)$$

其中 Y_t 为 m 维向量, $\Pi_i (i=1, 2, \dots, p)$ 是 $m \times m$ 阶参数矩阵, $u_t \sim IID(0, \Omega)$ 。如果 $Y_t \sim I(1)$, 我们可以用它们的一阶差分建立一个平稳的模型, 但当它们之间存在协整关系时, 我们就无法得到蕴含在 u_t 中的重要信息, 因此对 VAR 模型作出变换:

$$\begin{aligned} \Delta Y_t = & c + (\Pi_1 + \Pi_2 + \cdots + \Pi_p - I)Y_{t-1} - (\Pi_2 + \Pi_3 + \cdots + \Pi_p)\Delta Y_{t-1} \\ & - (\Pi_3 + \Pi_4 + \cdots + \Pi_p)\Delta Y_{t-2} - \cdots - \Pi_p \Delta Y_{t-p+1} + u_t \end{aligned}$$

$$\text{令 } \Gamma_j = -\sum_{i=j+1}^p \Pi_i, \quad j=1, 2, \dots, p-1 \quad \Pi = \Pi_1 + \Pi_2 + \cdots + \Pi_p - I$$

则上式写为:

$$\Delta Y_t = c + \sum_{j=1}^{p-1} \Gamma_j \Delta Y_{t-j} + \Pi Y_{t-1} + u_t \quad (2.9)$$

这就是向量误差修正模型 (VECM)，矩阵 Γ_j 是对 ΔY_t 的短期调整，而 Π 是对 ΔY_t 的长期调整。除了 ΠY_{t-1} ，式 (9) 中的其它各项都是平稳的，故对 Y_t 各变量间协整关系的检验转化为讨论 Π 的性质问题。关于 Π ，存在以下三种情况：

- (1) $r(\Pi) = m$ ，则说明 $Y_t \sim I(0)$ ；
- (2) $r(\Pi) = 0$ ，此时 Y_t 中各变量不具有协整关系；
- (3) $0 < r(\Pi) \leq m-1$ ，表示存在 $r(\Pi)$ 个协整关系。此时， Π 可以拆开为两个 $m \times r(\Pi)$ 阶矩阵 α 和 β 的乘积，即 $\Pi = \alpha\beta'$ ($r(\alpha) = r(\beta) = r(\Pi)$)，其中 β 为协整向量矩阵， β' 的每一行都是一个协整向量， α 为调整系数矩阵， α 的每一行均是 r 个协整组合的一组权重。

根据以上原理，给出 Johansen 检验的具体过程：

(1) 检验前的准备

(i) 用 OLS 估计式 (10)、(11) 中的每一个方程，计算残差，得到 $m \times T$ (T 为样本的时间长度) 阶的残差矩阵 S_0 和 S_1 。

$$\Delta Y_t = c_1 + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + U_t \quad (2.10)$$

$$Y_{t-1} = c_2 + \sum_{j=1}^{p-1} \Gamma_j \Delta Y_{t-j} + V_t \quad (2.11)$$

(ii) 利用残差矩阵构造如下矩阵：

$$R_{00} = T^{-1} S_0 S_0' \quad R_{01} = T^{-1} S_0 S_1' \quad R_{10} = T^{-1} S_1 S_0' \quad R_{11} = T^{-1} S_1 S_1' \quad (2.12)$$

(iii) 计算矩阵 $R_{11}^{-1} R_{10} R_{00}^{-1} R_{01}$ 的特征值，该矩阵即为 Π 的估计矩阵。

(2) 迹检验和最大特征根检验

若矩阵的秩为 r ，则矩阵有 r 个非零特征根，从大到小排列为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$ ，

同时 $\lambda_{r+1} = \lambda_{r+2} = \dots = \lambda_m = 0$ ，但这并不意味着存在 r 个协整关系，我们还检验这些

特征根是否是显著的，对应有两种方法，分别是迹检验和最大特征根检验。

(i) 迹检验

设 $H_0: \lambda_{r+1} = \lambda_{r+2} = \dots = \lambda_m = 0$ ，即存在 r 个协整关系，检验统计量为：

$$\lambda_{\text{trace}}(r) = -T \sum_{i=r+1}^m \ln(1 - \lambda_i) \quad (2.13)$$

该统计量服从 Johansen 分布。首先从 $r=0$ 开始检验，如果接受 H_0 ，说明变量间不存在协整关系；如果拒绝 H_0 ，说明至少存在 1 种协整关系，必须往下检验 $r=1$ 时统计量值的显著性，如果接受 H_0 ，说明存在 1 种协整关系；如果拒绝 H_0 ，说明至少存在 2 种协整关系，必须往下检验 $r=2$ 时统计量值的显著性；重复上述步骤，直到 $r=h$ 时接受 H_0 为止，则表明系统中存在 h 种协整关系。

(ii) 最大特征根检验

设 H_0 ：存在 r 个协整关系，检验统计量为：

$$\lambda_{\max}(r, r+1) = -T \ln(1 - \lambda_{r+1}) \quad (2.14)$$

该统计量是基于最大特征值 λ_r 的，故检验过程是由下往上进行的，其它步骤均与迹检验相同。

2.3 ARCH 和 GARCH 模型

2.3.1 ARCH 模型和 ARCH 检验

大多数的金融时间序列会表现出如下特征：当我们用回归或者自回归模型获得它们的非平稳信息后，一般会认为平稳的残差序列不再留有任何信息，但通过实际的观察，可以发现有些残差序列会呈现出集群效应，即它们在某些时段的波动极其剧烈，我们将其称之为异方差性。异方差存在时，OLS 估计得到的参数估计量不再有效的，无法证明回归参数与真实值之间的关系，于是 Engle (1982) 提出了如下的 ARCH(q) 模型：

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \cdots + \beta_k x_{kt} + \varepsilon_t \quad \varepsilon_t \sim N(0, h_t) \quad (2.15)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 \quad (2.16)$$

式 (15) 是均值方程, 提取的是序列的非平稳信息, 这里给出的是一个简单的多元回归模型, 一般地也可以采用 AR 模型或 ARMA 模型的形式, 式 (16) 是条件方差方程, 包含残差项中蕴涵的波动相关信息, 它认为异方差性来源于某种自相关关系, 可以用自回归模型进行拟合。要想相对比较完整和精确的刻画时间序列, 就必须将上述两方面的信息结合起来。

要拟合 ARCH 模型, 首先需要判断 ε_t 是否具有异方差性, 为此必须进行 ARCH 检验。如果 ε_t 方差非齐, 那么式 (16) 应当显著成立, 相应的假设条件为:

$$H_0: \alpha_0 = \alpha_1 = \alpha_2 = \cdots = \alpha_q \leftrightarrow H_1: \alpha_0, \alpha_1, \alpha_2, \cdots, \alpha_q \text{ 不全为零}$$

该检验被称为 LM 检验。

记总误差平方和为 $SST = \sum_{t=q+1}^T \varepsilon_t^2$, 自由度为 $T - q - 1$, 自回归模型的残差平方和

为 $SSE = \sum_{t=q+1}^T e_t^2$, 自由度为 $T - 2q - 1$, 则 LM 检验统计量为:

$$LM(q) = \frac{(SST - SSE) / q}{SSE / (T - 2q - 1)} \quad (2.17)$$

当该检验统计量在一定的显著性水平下显著时, 就可以认为残差序列方差非齐。

2.3. 2GARCH 模型

尽管 ARCH 模型能够捕捉到序列的波动效应, 但是它仍然存在一定的缺点, 由于移动平均模型自相关系数具有 q 阶截尾性, 那么当 h_t 具有长期自相关性时, 为了更好地去拟合它, 我们就必须设置相当高的阶数, 增加模型中参数的个数, 从而导致估计的难度加大, 并不一定能够取得较好的拟合效果, 为了解决这个问题, Bollerslov (1985) 在 Engle (1982) 的基础上提出了 GARCH(p, q) 模型, 它的结构

如下：

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \gamma_j h_{t-j} \quad (2.18)$$

可以发现，GARCH 模型的实质就是在原有异方差函数的基础上增加了 h_t 的 p 阶滞后项，这使得它比 ARCH 模型能够更为广泛地拟合残差序列。

3.跨品种套利神经网络模型的构建

3.1 神经网络的预测功能

人工神经网络，也简称为神经网络或称作连接模型，它是一种以人脑为模型的机器学习，简单来说，就是从信息处理的角度对人脑神经元网络进行抽象，创造一个人工神经网络，这种网络依靠自身的复杂结构，通过调整内部大量神经元之间相互连接的关系以及权值和阈值，达到处理信息的目的。总之，神经网络可以看成是一个数学模型，在结合了一些算法之后，能够输出一些对我们有价值的信息。

作为人工智能领域刚刚兴起的热点，神经网络能够执行深度学习，它通过一种算法允许计算机合并新的数据来学习，进而使系统具有识别、分类、预测等多种功能。正是因其强大的学习能力，近些年来，神经网络被广泛应用于营销、金融、公司财务、医药等领域，成为大数据时代的象征。

在金融领域，投资者都希望借助于自身能得到的信息对金融产品未来价格的变动方向作出判断，提前进行买卖，并获得收益，因为从过去价格的历史信息中总能找到一些具有规律性的因素。运用神经网络，我们可以依据该产品的相关信息捕捉其价格模式，在对过去价格序列完成拟合的同时，对未来价格的变化作出有效预测。从实证研究的视角看，神经网络被运用于金融工具价格、汇率、利率等时间序列的预测中，都取得了不错的效果，例如 Kulkarni 和 Haider（2009）基于多层前馈神经网络对原油价格的短期变化进行了预测，Panella et al.（2011）运用神经网络对能源商品一年以内的价格进行了预测，汪劲松和石薇（2014）利用多种群遗传算法优化来优化 BP 神经网络，对我国股指进行预测，董皓舒（2010）结合指数回归模型和 BP 神经网络预测了上交所多个国债的到期收益率。对比传统的时间序列模型，在非线性系统的预测上，神经网络模型甚至更具优越性，比如 Donaldson 和 Kamstra（1997）利用神经网络预测伦敦、纽约、东京和多伦多股市收益率波动的情况，发现人工神经网络捕捉了 GARCH 和 EGARCH 模型所忽略的波动性效应，宋策（2017）

将 BP 神经网络模型和 ARIMA 模型预测我国黄金期货价格的效果进行了对比，结果显示 BP 神经网络在预测方面的精确性和稳定性都更高。

由于神经网络具有极强的预测能力，本文认为其同样可以运用到跨品种套利的策略构建中，因为价差是呈现出规律性的，我们就是要利用神经网络对非线性序列中的有效信息进行捕捉，从而精确预测样本外的价差。

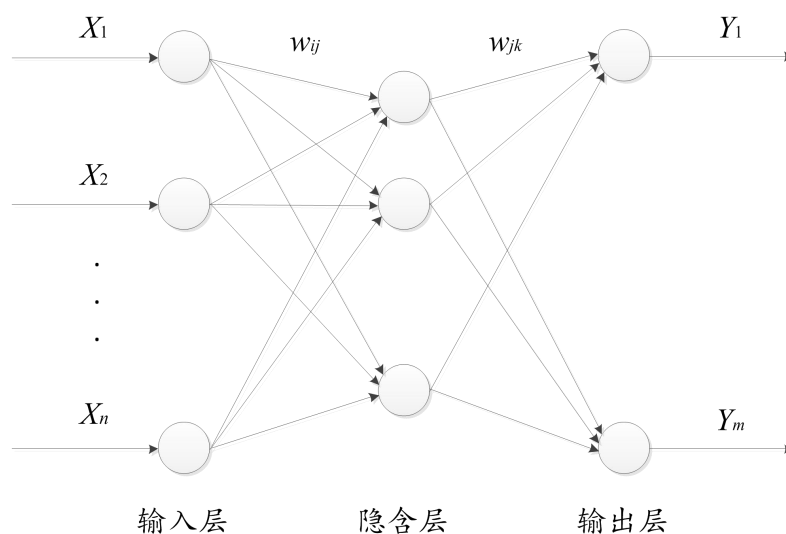
根据信息传递过程中是否存在反馈机制，神经网络可分为前馈式和反馈式两种类型，下面本文将选取这两种类型中最具有代表性的 BP 神经网络和 Elman 神经网络分别予以介绍。

3.2 前馈式的 BP 神经网络

前馈式网络的特征在于，在整个算法的计算流程中，网络的输出仅与当前的输入相关，而不会受到之前输入或输出的影响，也就是说，这种网络的结构相对比较简单。

BP 神经网络就是一种典型的前馈式网络，它是 1986 年由 Rumelhart 等学者所建立的模型，也是目前结构最为简单、使用最多的神经网络，BP 神经网络拓扑结构如图所示：

图 3.1 BP 神经网络拓扑结构图



图中展示了一个最普通的 BP 神经网络的结构，它包含输入层、隐含层和输出层，一般而言，隐含层可以有多个，为了方便解释说明，我们这里仅设置一个隐含层。输入层有 n 个神经元，隐含层有 l 个神经元，网络的输入值 X_1, X_2, \dots, X_n 通过激励函数 f 、权值 w_{ij} 和阈值 a_j 连接到隐含层，其中激励函数有多种表达形式：

$$\text{logsig}(x) = \frac{1}{1+e^{-x}} \quad \text{tansig}(x) = \frac{2}{1+e^{-2x}} - 1 \quad \text{purelin}(x) = x \quad (3.1)$$

最常用的为 $\text{logsig}(x)$ ，也称为 S 型函数，对应地，隐含层第 j 个神经元的输出为：

$$H_j = f\left(\sum_{i=1}^n w_{ij} X_i - a_j\right) \quad j=1, 2, \dots, l \quad (3.2)$$

同样，隐含层的输出通过激励函数 f 、权值 w_{jk} 和阈值 b_k 连接到输出层，输出层有 m 个神经元，第 k 个神经元的输出为：

$$O_k = f\left(\sum_{j=1}^l w_{jk} H_j - b_k\right) \quad k=1, 2, \dots, m \quad (3.3)$$

这里的 O_k 代表的是网络的预测输出，而并非实际输出 Y_1, Y_2, \dots, Y_m ，由此可以得出 BP 神经网络的特点之一在于信息是向前传递的，每一层神经元的输出仅受目前上一层输出的影响。为了使 O_k 尽量贴近其真实值 Y_k ，我们就必须对网络进行训练，这样的过程实际上是要对权值和阈值不断地进行修正，如果说神经网络表达了 n 个自变量到 m 个因变量的映射关系，那么我们就是在对其训练的过程中去拟合一个非线性函数。

限于篇幅的原因，我们不对 BP 神经网络的修正过程详细证明（证明过程中网络误差采用均方误差）。设每一次迭代后 w_{ij} 、 w_{jk} 、 a_j 、 b_k 的修正量分别为 Δw_{ij} 、 Δw_{jk} 、 Δa_j 、 Δb_k ，由于 BP 神经网络采用梯度下降法来逼近最优值，所以我们要以负梯度的方向对权值和阈值进行调整，在给定了学习速率 η ，并且隐含层和输出层的激励函数都是 S 型函数的前提下，令：

$$g_k = O_k(1 - O_k)(Y_k - O_k) \quad (3.4)$$

那么 Δw_{ij} 、 Δw_{jk} 、 Δa_j 、 Δb_k 对应的表达式为：

$$\Delta w_{ij} = \eta H_j(1 - H_j) X_i \sum_{k=1}^m w_{jk} g_k \quad (3.5)$$

$$\Delta w_{jk} = \eta g_k H_j \quad (3.6)$$

$$\Delta a_j = -\eta H_j(1 - H_j) \sum_{k=1}^m w_{jk} g_k \quad (3.7)$$

$$\Delta b_k = -\eta g_k \quad (3.8)$$

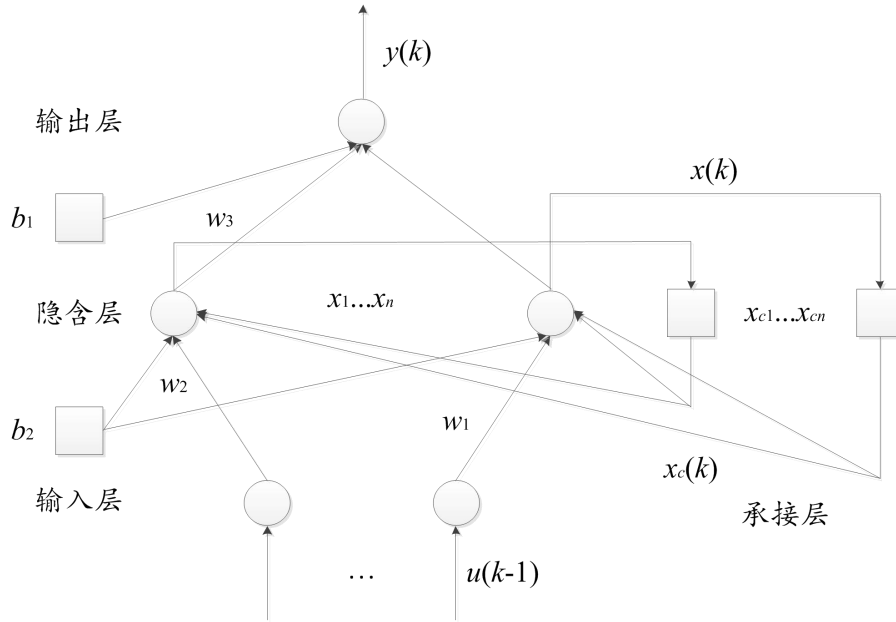
经过若干次迭代后， w_{ij} 、 w_{jk} 、 a_j 、 b_k 可以不断地逼近其最优值，BP 神经网络就拥有了较强的联想记忆的功能，将新的有效值输入后，就可以对样本外，即未来的信息作出精准地预测。

3.3 反馈式的 Elman 神经网络

反馈式神经网络也称递归网络或回归网络，它与前馈式网络最大的不同点在于：在前馈式网络中，当前的信息一旦被使用，就失去了价值，而反馈式则将这些有用的信息保留下来，并将其传递到神经元的输入中，使这些信息再次被利用。反馈式网络在本质上类似于时间序列模型，它认为过去的信息总会对未来产生一定的影响。如何使用之前的信息，取决于神经网络的反馈形式，一般来讲，反馈形式包括神经元自反馈、两层之间互相反馈、单层输出反馈等类型。反馈形式的不同，将决定神经网络学习特征的差异。几种比较重要的反馈式神经网络有 Hopfield 神经网络、Elman 神经网络以及 Boltzmann 机模型。根据具体问题的不同，我们需要选择对应的模型进行分析。

作为一种典型的反馈式网络，Elman 神经网络最早应用于语音处理领域，但后来学者们发现其在序列的预测方面具备非常好的性能，Elman 神经网络的拓扑结构如图所示：

图 3.2 Elman 神经网络拓扑结构图



图中展示了一个常见的 Elman 神经网络的结构，可以发现，它的结构与 BP 神经网络极其类似，只是在隐含层中增加了一个承接层，承接层又称上下文层或状态层，它在网络起着非常重要的作用，它可以将网络前几个时刻隐含层的输出值保留下来，这些输出值通过 w_3 连接到隐含层作为其下一个时刻的输入值，此时网络的输入在当前信息的基础上又考虑了历史信息的作用和影响，从而使整个网络能够根据系统状态的变化及时调整其输出。图 3.2 中，承接层保留的是上一时刻隐含层的输出值，依据此，从数学模型的角度看，Elman 神经网络的原理可写成如下形式：

$$y(k) = g(w_3 x(k)) \quad (3.9)$$

$$x(k) = f(w_1 x_c(k) + w_2(u(k-1))) \quad (3.10)$$

$$x_c(k) = x(k-1) \quad (3.11)$$

同 BP 神经网络一样，Elman 神经网络采用梯度下降法进行权值和阈值的修正，计算误差时仍使用均方误差公式。

Elman 神经网络最大的特点就在于状态层的加入，这使它具备很强拟合动态系统的能力。当我们想要对动态系统进行建模时，往往会考虑到动态系统几乎都带

有一些趋势性的特征,这种特征表现在如果过去的一段时间整个系统朝着正向的方向发生变化,那么当前时刻系统的变化方向很可能会保持不变,也就是说,系统具有一定历史敏感性,而 Elman 神经网络就是能够在忽略外部噪声的情况下,通过承接层对历史信息的记忆,完成对动态系统的精确拟合。

与前面的 BP 神经网络相比,Elman 神经网络是一种内部反馈型网络,能够存储和利用过去时刻的信息,既可以实现静态系统的建模,还可实现动态系统的映射并直接反映系统的动态特性,在计算能力及稳定性方面都更为优越。

3.4 基于遗传算法优化 Elman 神经网络

虽然 Elman 神经网络和 BP 神经网络对未来的数据都具有较强的预测能力,但它们存在共同的缺点:首先,由于初始权值和阈值的给定是随机的,那么在大多数情况下,采用梯度下降法进行对其进行修正时,网络的收敛速度可能会偏慢,其次,在大量样本存在的前提下,是无法保证所有的预测值都是完全接近样本值的,也就是说,网络会陷入局部最优。为了改善这几点不足之处,我们有必要引入遗传算法。

遗传算法,顾名思义,起源于对生物界进化规律的研究,是 1962 年由美国学者 Holland 提出的随机搜索最优化方法,主要借鉴的是达尔文的进化理论和孟德尔的遗传学说。众所周知,物竞天择、适者生存是自然界永恒不变的真理,在一个群体中,个体会通过繁衍产生新的个体,与此同时,基因变异会产生具有新的特征的个体,如此反复,不断循环,拥有优秀基因的个体被保留,其余的则被淘汰。遗传算法正是模拟了这一进化过程,它希望通过有限次数的迭代,来寻求问题的全局最优解。

本文利用遗传算法改进 Elman 神经网络,以期求出其最佳初始权值和阈值,在加快网络收敛速度的同时使其在训练过程中达到全局最优解,从而对价差序列作出更为精确和稳定地预测。遗传算法的流程如下:

(1) 种群初始化

根据我们所要解决的具体问题,需要确定种群的各项初始参数,这里包括了种群大小 M 、进化代数 G 、交叉概率 P_c 和变异概率 P_m 。

(2) 染色体编码

染色体编码是指对每个个体进行编码，从遗传学的角度看，就是建立表现型和基因型的映射关系。染色体编码的方法通常有二进制法和实数法，考虑到我们需要优化的具体对象，本文采用实数编码方法，在这里，每个个体均被编码为一个实数串，它是在根据输入输出序列确定 Elman 神经网络的结构后，由 w_1 、 w_2 、 w_3 、 b_1 、 b_2 五部分按照顺序排列转化而成的向量，实数串的长度即为染色体的长度。

(3) 确定适应度函数

在群体中，并非所有的个体都能够生存下来，因为它们对自然界环境的适应能力存在一定的差异，有较强适应能力的个体会生存下去，有较弱适应能力的个体则逐渐推出。为了确定哪些个体的适应能力更强，我们必须建立一个函数来衡量个体生存下来的能力，这就是适应度函数。

本文将每个个体（实数串）作为 Elman 神经网络的初始权值和阈值，用样本内数据对网络进行训练后，得到预测输出 o_i ，其与实际输出 y_i 之间的误差绝对值和即为个体适应度值 F ，计算公式为

$$F = k_1 \sum_{i=1}^m \text{abs}(y_i - o_i) \quad (3.12)$$

上式中， m 为网络输出层神经元的个数， k_1 为一个固定的常数。

(4) 遗传操作

(i) 选择操作

选择操作是从原有群体中选择具有较优适应度值的个体到新群体中，有轮盘赌法、锦标赛法等多种方法，本文选用轮盘赌法，每个个体 i 被选择的概率 p_i 为：

$$f_i = k_2 / F_i \quad (3.13)$$

$$p_i = \frac{f_i}{\sum_{j=1}^M f_j} \quad (3.14)$$

由于我们希望通过遗传算法找到能够使网络预测输出尽量贴近真实输出的权值和阈值,那么根据适应度值的计算方法, F_i 更小则为更优,根据 p_i 的定义,我们还需要对 F_i 求倒数。

(ii) 交叉操作

交叉操作模拟的是生物在进化过程中的繁衍行为,具体来讲是在种群中任意选择两个个体,以一定的交叉概率 P_c ,进行染色体的重组,从而产生新的个体。本文的每一个个体都代表了一个实数串,所以第 k 个个体 a_k 和第 l 个个体 a_l 在第 j 位的交叉方法如下:

$$a_{kj} = a_{kj}(1-b) + a_{lj}b \quad (3.15)$$

$$a_{lj} = a_{lj}(1-b) + a_{kj}b \quad (3.16)$$

上式中, b 为 $[0,1]$ 间的随机数。

(iii) 变异操作

变异操作模拟的是生物在进化时发生的基因突变状况,具体来说是在种群中任选一个个体,以一定的变异概率 P_m ,进行染色体的变异,以产生性状更优的个体。选取第 i 个个体的第 j 个基因 a_{ij} 进行变异方法如下:

$$a_{ij} = a_{ij} + (a_{ij} - a_{\max}) * f(g) \quad 0.5 < r < 1 \quad (3.17)$$

$$a_{ij} = a_{ij} + (a_{\min} - a_{ij}) * f(g) \quad 0 < r \leq 0.5 \quad (3.18)$$

这里

$$f(g) = (r(1 - \frac{g}{G}))^2 \quad (3.19)$$

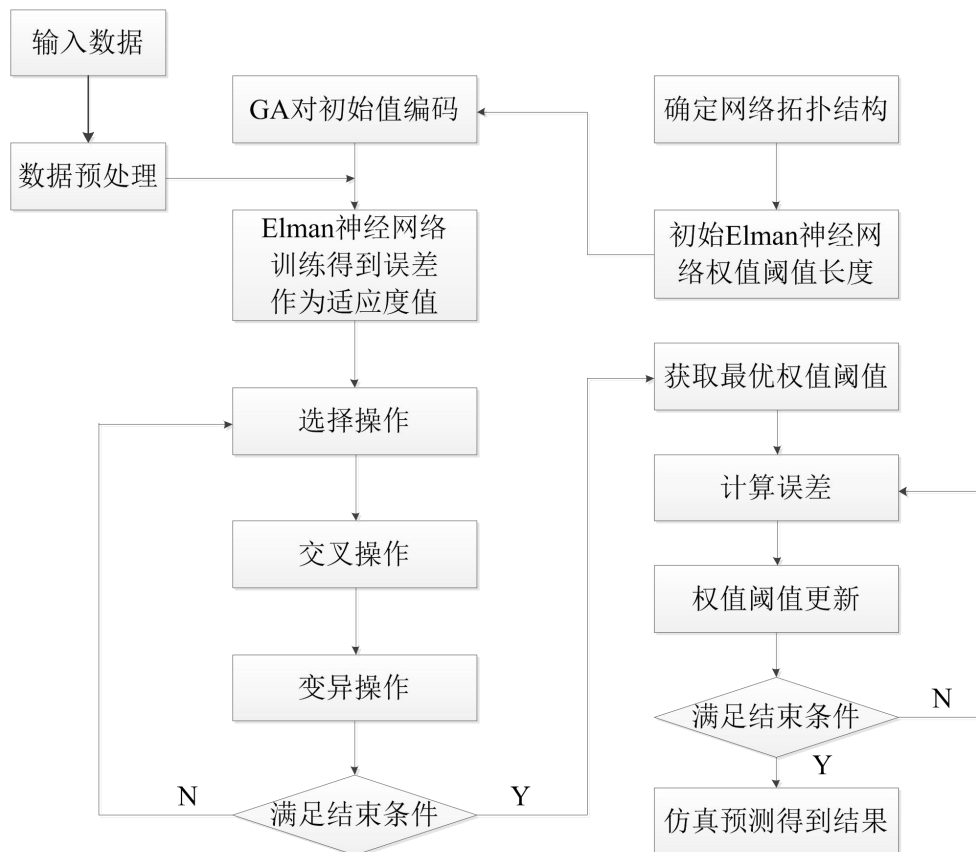
g 为此时的迭代次数。

经过遗传操作 G 次的迭代,也就是种群在进化了 G 代之后,我们就可以找到具有最优适应度值的个体,将其赋给 Elman 神经网络作为最佳权值和阈值,用于拟合

样本外的序列。

该部分的算法流程图如下：

图 3.3 算法流程图



4.实证研究

4.1 数据选取和描述性统计

黑色金属产业链主要包含 7 种商品期货：螺纹钢、热卷、线材、铁矿石、焦煤、焦炭、动力煤，其中多以螺纹钢、铁矿石和焦炭作为套利品种，因铁矿石和焦炭是生产螺纹钢的重要原料，矿厂和焦炭厂位于钢铁厂的上游，原料厂和冶炼厂都受到成本和利润的限制，故螺纹钢的价格同铁矿石和焦炭的价格可能存在较强的相关关系，除此之外，由于产业链跨品种套利比相关商品套利在形式上更具稳定性，所以本文选取上海期货交易所上市的螺纹钢期货、大连商品交易所上市的铁矿石期货和焦炭期货作为套利的研究对象。

在长期内，钢铁厂的冶炼技术会不断提升，从而影响到三种期货间的价格关系，那么研究的时间跨度就不宜选择过长，同时市场上一种商品对应着多个交割月份的期货合约，它们的交易活跃度有所不同，并且当我们进行换月交易时，若临近交割月份，套利条件仍然无法满足，也必须要平仓，为了简化交易、覆盖提前平仓的风险，我们采用三种期货成交量最大的合约（主力连续合约）的数据，最后，我们可以尝试增加套利的频数，一方面商品期货包含了日盘和夜盘交易，在螺纹钢、铁矿石和焦炭夜盘交易时间不同的情况下，需要取三者的交集，另一方面选用高频数据，但时间间隔不宜果断，否则手续费用的增加将影响获利。基于上述考虑，我们选取 2017 年 11 月 20 日至 2018 年 11 月 9 日螺纹钢、铁矿石和焦炭主力连续合约的日盘及夜盘的 5 分钟收盘价作为实证研究数据，所有数据均来自于通达讯金融终端。

表 4.1 螺纹钢、铁矿石和焦炭期货交易时间

商品期货	日 盘	夜 盘
螺纹钢	9:00-11:30 13:00-15:00	21:00-23:00
铁矿石		21:00-23:30
焦炭	10:15-10:30 休盘	21:00-23:30

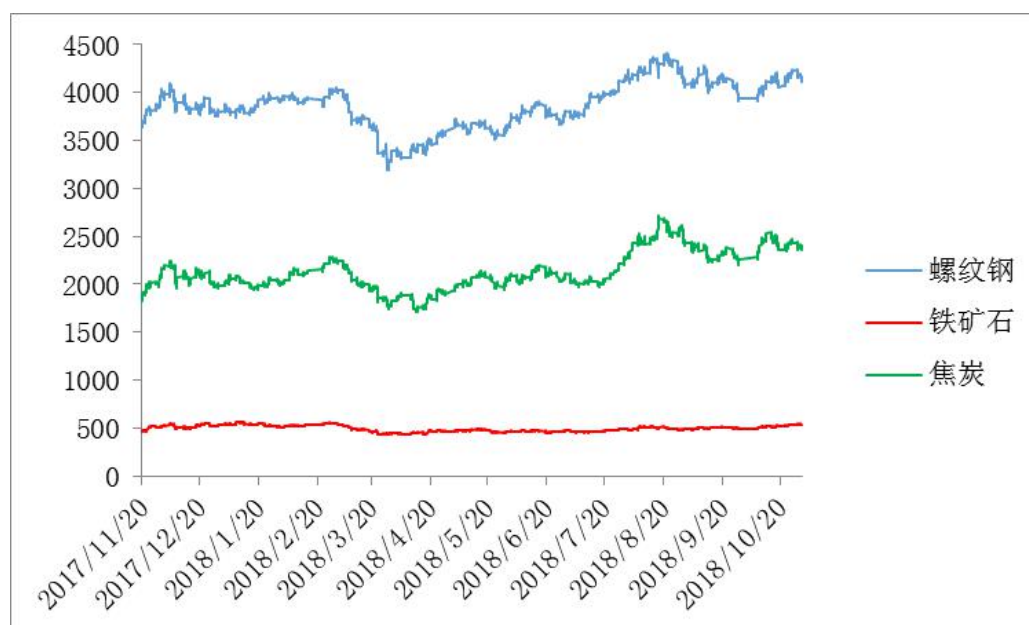
经过整理,我们的数据集内包含 3×15771 个期货 5 分钟收盘价的时间序列数据,其中样本内含 2017 年 11 月 20 日至 2018 年 9 月 28 日共 3×14553 个数据,用于协整检验和套利比例的确定,样本外含 2018 年 10 月 8 日至 10 月 31 日共 3×1218 个数据,用于神经网络预测和套利策略实现,分别用 rebar_t 、 iron_t 、 coke_t 表示螺纹钢期货、铁矿石期货、焦炭期货的价格序列,它们的描述性统计如下:

表 4.2 数据的描述性统计

	计数	最小值	最大值	均值	标准差	偏度	峰度
rebar_t	15771	3195.0	4408.0	3873.72	242.94	-0.2545	2.7252
iron_t	15771	428.0	562.5	493.89	30.57	0.1287	1.9661
coke_t	15771	1716.0	2711.5	2133.28	200.28	0.5861	2.7007

可以看到,三种期货的价格的波动性都比较大,焦炭期货在该时间段价格变动幅度接近 60%,螺纹钢期货和焦炭期货也有 30% 以上,不过三者的均值都基本位于最大值和最小值中间,且标准差均处于正常范围内,并没有出现价格在短期内发生极端异常波动的情况,说明我们收集的数据用于实际研究是满足要求的。

图 4.1 螺纹钢、铁矿石和焦炭期货收盘价的时间序列



通过观察螺纹钢、铁矿石和焦炭期货收盘价的时间序列，可以发现，螺纹钢期货和焦炭期货的价格走势几乎完全相同，从 2018 年的 4 月开始，价格都会大幅上涨，一直持续到 8 月左右，但铁矿石期货的价格变动趋势只在 2017 年 12 月和 2018 年上半年与二者保持一致，2018 年下半年并未随二者走出明显的上涨或下跌行情，也就是说，螺纹钢和焦炭期货的价格序列应该有比较强的相关性，而铁矿石期货的价格与二者可能存在一般的相关性。进一步地，我们通过计算三种期货的价格序列的相关系数来验证上述推论。

表 4.3 三种期货的价格序列的相关矩阵

	rebar_t	iron_t	coke_t
rebar_t	1.000000	0.894181	0.550217
iron_t	0.894181	1.000000	0.399032
coke_t	0.550217	0.399032	1.000000

相关系数矩阵显示， rebar_t 和 coke_t 的相关系数高达 0.894181，而 iron_t 与二者的相关系数分别为 0.550217 和 0.399032，表明三种期货之间可能存在长期稳定且合理的关系，必须利用协整检验的方法来考察它们之间的均衡关系。

4.2 协整检验

4.2.1 单位根检验

在对螺纹钢期货、铁矿石期货及焦炭期货价格序列之间的关系进行协整检验前，须对所有的序列及它们的差分序列进行单位根检验，以保证原序列皆为同阶单整，为此我们运用 ADF 法检验了三种期货的价格的原序列 rebar_t 、 iron_t 、 coke_t 及其一阶差分序列 Δrebar_t 、 Δiron_t 、 Δcoke_t 的平稳性。

表 4.4 螺纹钢、铁矿石和焦炭期货的价格序列及其一阶差分的 ADF 检验结果

	rebar_t	iron_t	coke_t	Δrebar_t	Δiron_t	Δcoke_t
t-Statistic	0.302169	0.125444	0.533091	-125.1948	-131.0981	-124.7131
Prob.	0.7732	0.7222	0.8311	0.0001	0.0001	0.0001

注：Test critical values: 1% level -2.565121 5% level -1.940846 10% level -1.616685
下面的 ADF 检验与之相同

根据检验结果，三种期货合约原价格序列的 ADF 检验统计量的值均高于 1%、5%、10%显著性水平上的临界值，故它们都是非平稳序列；价格序列经一阶差分后， Δrebar_t 、 Δiron_t 和 Δcoke_t 的 ADF 检验统计量的值均低于 1%、5%、10%显著性水平上的临界值，此时它们都是平稳序列，则 rebar_t 、 iron_t 和 coke_t 同为 $I(1)$ 序列，满足协整检验的条件。

4.2.2 套利组合的协整检验

由于 Johansen 检验需要确定 VAR 模型的滞后阶数以及检验其平稳性，过程比较复杂，因此本文采用 EG 两步法检验套利品种之间是否存在协整关系。在套利组合的选择上，我们既可以利用螺纹钢期货、铁矿石期货和焦炭期货中的两种进行套利，也可以进行三个品种之间的套利，不过前提条件是配对品种须通过协整检验。考虑两种商品间的套利，依据现货市场上商品上下游的关系，可构建螺纹钢期货和铁矿石期货、螺纹钢期货和焦炭期货这两对交易组合，并将其简记为 $\text{rebar}_t\text{-iron}_t$ 、 $\text{rebar}_t\text{-coke}_t$ 。我们首先用 OLS 法对它们的价格序列进行回归，得到如下估计结果：

表 4.5 $\text{rebar}_t\text{-iron}_t$ $\text{rebar}_t\text{-coke}_t$ 的 OLS 估计结果

	β_0	β_1	R^2
$\text{rebar}_t\text{-iron}_t$	1873.298*** (27.59968)	4.026511*** (0.056029)	0.261955
$\text{rebar}_t\text{-coke}_t$	1456.847*** (10.13128)	1.136040*** (0.004784)	0.794863

注：***、**和*分别表示系数在 1%、5%和 10%的显著性水平下显著，圆括号里面的数值为估

计标准误，最后一列为模型的拟合优度。（下同）

从回归结果看，系数都是显著的，螺纹钢期货的价格序列对焦炭期货的价格序列回归效果比较好，拟合度较高，这与二者价格走势相近密切相关，而同铁矿石期货的价格序列回归效果稍差，其原因在于 2018 年 6 月之前，铁矿石和螺纹钢期货的价格走势是相似的，但是此后铁矿石的价格却一直保持平稳，并未发生较大波动。接下来我们可以利用回归系数计算出残差 e_{1t} 和 e_{2t} ：

$$e_{1t} = \text{rebar}_t - 4.02651 \text{iron}_t - 1873.298 \quad (4.1)$$

$$e_{2t} = \text{rebar}_t - 1.136040 \text{coke}_t - 1456.847 \quad (4.2)$$

一般地，价差序列等同于残差 e_{1t} 和 e_{2t} 的计算公式中不考虑常数项，同时为了研究的方便，还需对其进行去中心化处理。本文则较为直接地采用残差序列代表已去中心化的价差序列，一方面能够简化计算的步骤，避免重复计算使数据丧失真实意义，另一方面残差的差值就是我们套利获得的实际收益，这样考虑概念上更为清晰，那么 $\text{rebar}_t - \text{iron}_t$ 和 $\text{rebar}_t - \text{coke}_t$ 的价差序列如图所示：

图 4.2 $\text{rebar}_t - \text{iron}_t$ 的价差序列

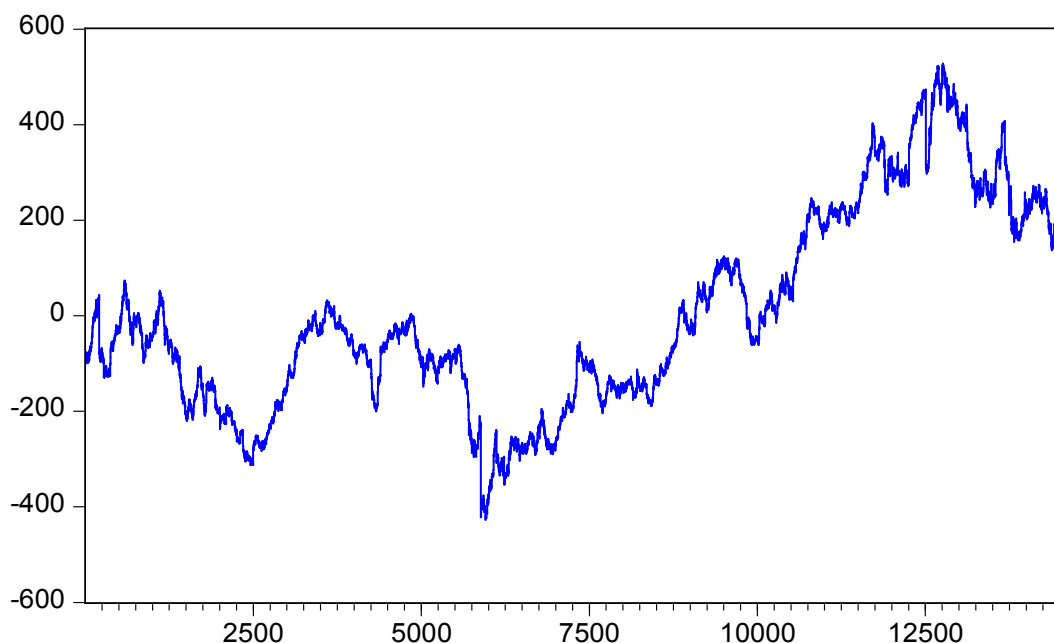
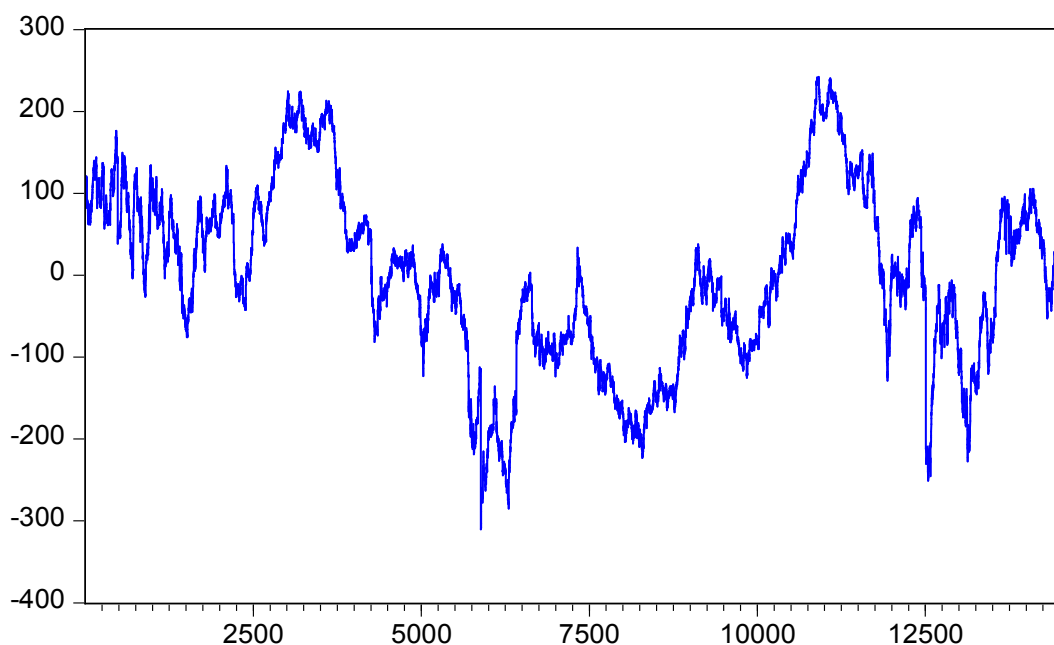


图 4.3 $\text{rebar}_t - \text{coke}_t$ 的价差序列



观察 $\text{rebar}_t - \text{iron}_t$ 的价差序列，其并没有围绕 0 附近震荡，在相当长的一段时间内，螺纹钢期货的价格相对于铁矿石期货的价格持续上升，虽然之后出现了回落的趋势，但残差序列可能是非平稳的。 $\text{rebar}_t - \text{coke}_t$ 的价差序列则围绕 0 上下波动，其很可能是平稳的。为了论证上面的观点，我们对 $\text{rebar}_t - \text{iron}_t$ 、 $\text{rebar}_t - \text{coke}_t$ 的残差序列进行 ADF 检验，结果显示 e_{1t} 的 ADF 检验统计量的 t 值为 -1.423355，大于 10% 显著性水平上的临界值，表明它是非平稳序列，因此螺纹钢和铁矿石期货的价格序列之间不存在协整关系， e_{2t} 的 ADF 检验统计量的 t 值为 -2.774050，小于 1% 显著性水平上的临界值，则螺纹钢和焦炭期货的价格存在协整关系， $\text{rebar}_t - \text{coke}_t$ 可用于跨品种套利。

虽然在一定的显著性水平下，结果认为螺纹钢期货和铁矿石期货的价格不具有长期稳定的均衡关系，但它们的走势仍具有相似性，只是在某个时间段略微脱离了合理性，当然也可能与我们研究的区间较短有一定联系，不过这些并不影响我们考察三个期货品种之间的协整关系，同样地，我们将套利组合记为 $\text{rebar}_t - \text{iron}_t - \text{coke}_t$ ，

运用 OLS 法估计铁矿石和焦炭对螺纹钢的回归系数：

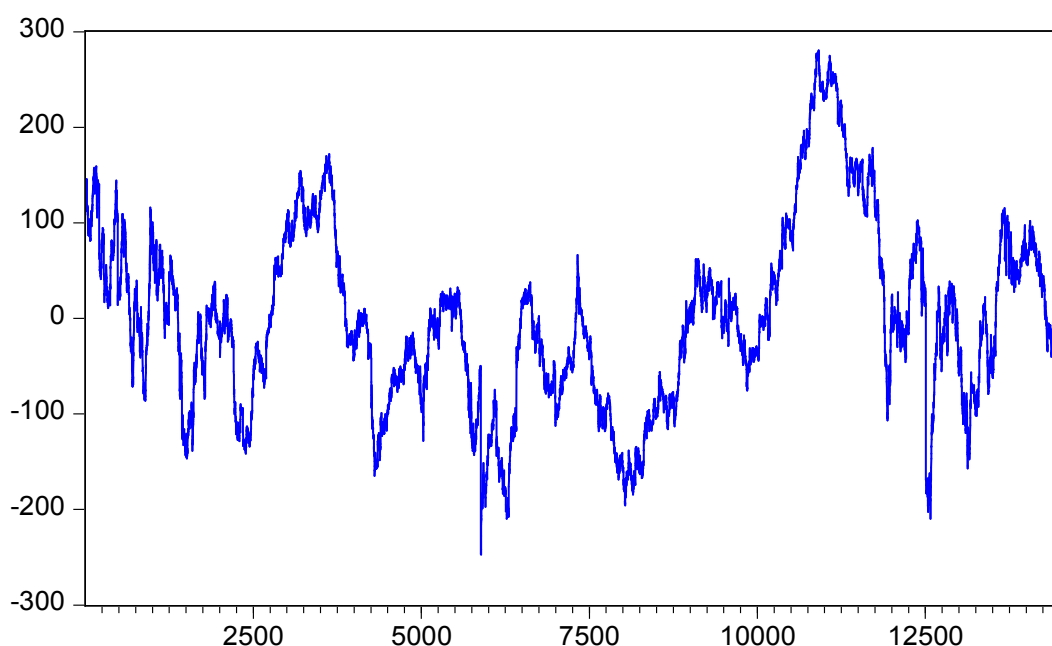
表 4.6 $\text{rebar}_t\text{-iron}_t\text{-coke}_t$ 的 OLS 估计结果

	β_0	β_1	β_2	R^2
$\text{rebar}_t\text{-iron}_t\text{-coke}_t$	748.5425***	1.880066***	1.033617***	0.845513
	(13.50838)	(0.027221)	(0.004409)	

现货市场上，生产 1 吨螺纹钢大致需要 1.6 吨铁矿石和 0.5 吨焦炭，回归系数在一定程度上支持了这一关系。可以看到，模型的拟合效果非常好，系数均呈显著，进一步地残差序列 e_{0t} 表示为：

$$e_{0t} = \text{rebar}_t - 1.880066\text{iron}_t - 1.033617\text{coke}_t - 748.8425 \quad (4.3)$$

图 4.4 $\text{rebar}_t\text{-iron}_t\text{-coke}_t$ 的价差序列



对其进行 ADF 检验， e_{0t} 的 ADF 检验统计量的 t 值为 -3.238771，在 1% 的显著性水平上是显著的，说明 e_{0t} 是平稳的，由此可知三种期货的价格存在协整关系，适合作为套利组合。

4.2.3 套利组合的选取

根据协整检验的结果, $\text{rebar}_t\text{-coke}_t$ 和 $\text{rebar}_t\text{-iron}_t\text{-coke}_t$ 都可作为跨品种套利的组合, 我们需要判断哪一对组合更优。首先, 虽然螺纹钢期货和铁矿石期货的价格不存在协整关系, 但三个品种间存在协整关系, 并且利用焦炭和铁矿石期货对螺纹钢期货做回归比单纯用焦炭期货的效果更好, 拟合优度更高; 其次, 从价差序列看, e_{2t} 的极差和标准差都大于 e_{0t} , 即 e_{2t} 的波动性高于 e_{0t} , 同时 e_{0t} 比 e_{2t} 更接近正态分布, 考虑价差的变化对套利的影响, $\text{rebar}_t\text{-iron}_t\text{-coke}_t$ 无疑是更适合作为套利组合的; 最后, 利用三种期货套利更具有现实意义, e_{0t} 代表的是钢铁厂的冶炼价值。综上, 本文将采用 $\text{rebar}_t\text{-iron}_t\text{-coke}_t$ 的组合进行套利研究。

表 4.7 e_{2t} 和 e_{0t} 的描述性统计

	计数	最小值	最大值	均值	标准差	偏度	峰度
e_{2t}	14553	-310.46	242.10	0.00	109.03	-0.0799	2.4874
e_{0t}	14553	-247.22	280.43	0.00	94.61	0.4288	3.0610

4.3 套利比例的确定

4.3.1 残差序列的 ARCH 检验

通过观察 e_{0t} , 可以发现其具有明显的集群效应, 那么残差序列就极有可能存在异方差性, 在利用 OLS 法估计参数时, 估计量不再是有效的, 系数的显著性检验结果不可依赖, 从而影响到我们对 $\text{rebar}_t\text{-iron}_t\text{-coke}_t$ 套利比例的确定, 因此有必要对 e_{0t} 进行 ARCH 检验。我们直接使用自回归模型拟合 e_{0t} , 这里令 $q=6$, 得到的回归结果如下:

表 4.8 e_{0t} 的 AR(6) 模型拟合结果

变量	系数	标准误差	t 值	P 值
C	23.19414	10.02274	2.314153	0.0207
ε_{t-1}^2	0.948634	0.008292	114.3979	0.0000
ε_{t-2}^2	0.039832	0.011430	3.484719	0.0005
ε_{t-3}^2	-0.017975	0.011435	-1.571924	0.1160
ε_{t-4}^2	0.010856	0.011434	0.949434	0.3424
ε_{t-5}^2	0.003116	0.011431	0.272614	0.7852
ε_{t-6}^2	0.012851	0.008292	1.549769	0.1212

模型的 $R^2 = 0.994059$ ，由 e_{0t} 和 AR(6) 模型的残差可计算出 LM 检验统计量的值为 405447.0， P 值为 0.0000，小于 0.05 的显著性水平，表明 e_{0t} 具有异方差性，需重新用 ARCH 模型或 GARCH 模型进行拟合。

4.3.2 GARCH 模型的拟合

根据上面的回归结果，残差序列的一阶和二阶滞后项的系数均是显著的，同时考虑到异方差函数可能存在长期自相关性，因此本文尝试利用 GARCH(2,2)模型重新拟合残差，模型的口径如下：

表 4.9 GARCH(2,2)模型的拟合结果

	β_0	β_1	β_2	α_0
e_{0t}	755.9692*** (7.538542)	1.843651*** (0.009895)	1.035053*** (0.002612)	2271.342*** (0.002162)
α_1	α_2	γ_1	γ_2	R^2
0.866015*** (0.056722)	1.029998*** (0.062314)	-1.190053*** (0.023278)	-0.304869*** (0.018786)	0.844536

模型的拟合效果良好，均值方程和条件方差方程中所有变量的系数均是显著的。根据回归系数，可以确定螺纹钢期货、铁矿石期货和焦炭期货三者真实的套利比例为：

$$\text{rebar}_t : \text{iron}_t : \text{coke}_t = 1 : 1.843651 : 1.035053 \quad (4.4)$$

此时，所有样本的实际价差为：

$$e_t = \text{rebar}_t - 1.843651 \text{iron}_t - 1.035053 \text{coke}_t - 755.9692 \quad (4.5)$$

4.4 基于神经网络拟合价差序列

4.4.1 BP 神经网络

神经网络经过训练后，具有良好的联想记忆和预测能力，利用这一点我们可以对价差序列进行拟合，从而基于此来构建套利策略。

在对 BP 神经网络进行训练前，首先要初始其结构，即各个层的神经元个数，其中输入层和输出层的节点数很大程度上取决于输入序列和输出序列的选择，这需要根据实际问题的特点来考察，我们的目的是预测未来的价差，由于事物过去的变化总能对未来的趋势产生一定的影响，那么根据目前所掌握到的信息来推测下一时间点的价差是可行的，相应的输入层节点数应设置为 1，代表下一时刻的价差，而我们现在所拥有的信息仅为前几个时间点的价差，据此本文将输入层节点数定为 10，即为前 10 个时间段的价差，也就是说，用前 10 个时间点的价差来预测下一时间点的价差，二者应呈非线性函数关系：

$$\bar{e}_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-10}) \quad (4.6)$$

这里的 \bar{e}_t 代表的是预测的价差，而并非真实的价差，神经网络的应用就体现在对非线性关系的拟合上。关于最佳隐含层节点数的选择，可参考如下公式：

$$l < n-1 \quad l < \sqrt{(m+n)} + a \quad l = \log_2 n \quad (4.7)$$

其中， n 为输入层节点数， l 为隐含层节点数， m 为输出层节点数， a 为 0~10 之间的常数。如果 $a=5$ ，那么隐含层节点数设置为 3 到 8 之间任意一个数都能满足要求，

从实际拟合的效果来看，我们令 $l=6$ 。

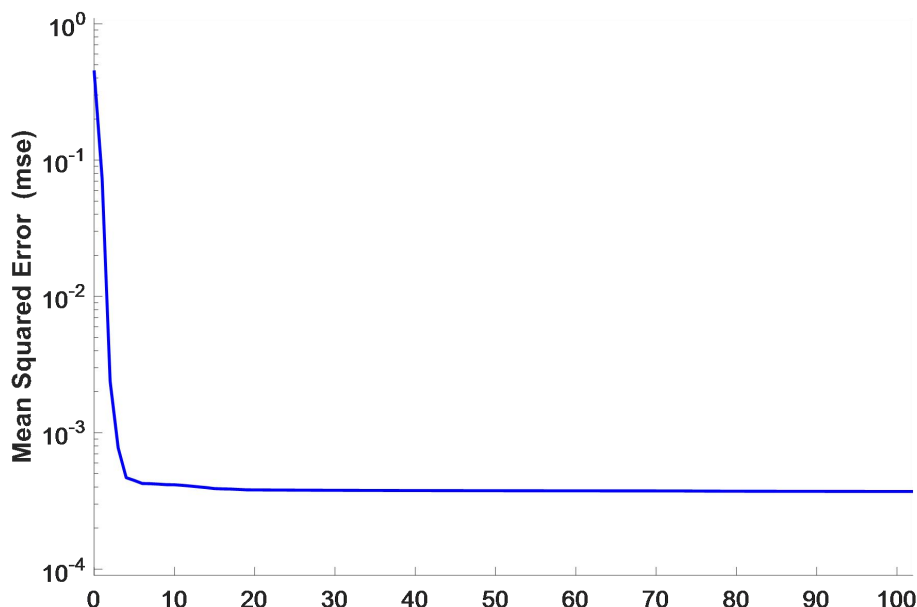
除此之外，我们必须对价差数据进行归一化处理，即将其转变成 $[0,1]$ 之间的数，这样做能够降低数据的数量级，从而减少网络在训练过程中的误差。本文采用最大最小法对数据进行归一化，其函数形式如下：

$$x_k = (x_k - x_{\min}) / (x_{\max} - x_{\min}) \quad (4.8)$$

最后还需要注意要对数据进行反归一化处理，否则就会丧失原有数据的实际意义。

由于包含现在时刻在内的前 10 个时段的价差是已知的，我们将其作为输入向量，而下一个时刻的价差作为输出向量。经过比较后，隐含层函数选用 $\text{logsig}(x)$ ，输出层函数选用 $\text{tansig}(x)$ ，利用 Matlab 自带的神经网络工具箱，对样本内的数据进行训练，迭代次数为 300 次，得到预测误差的收敛图如下：

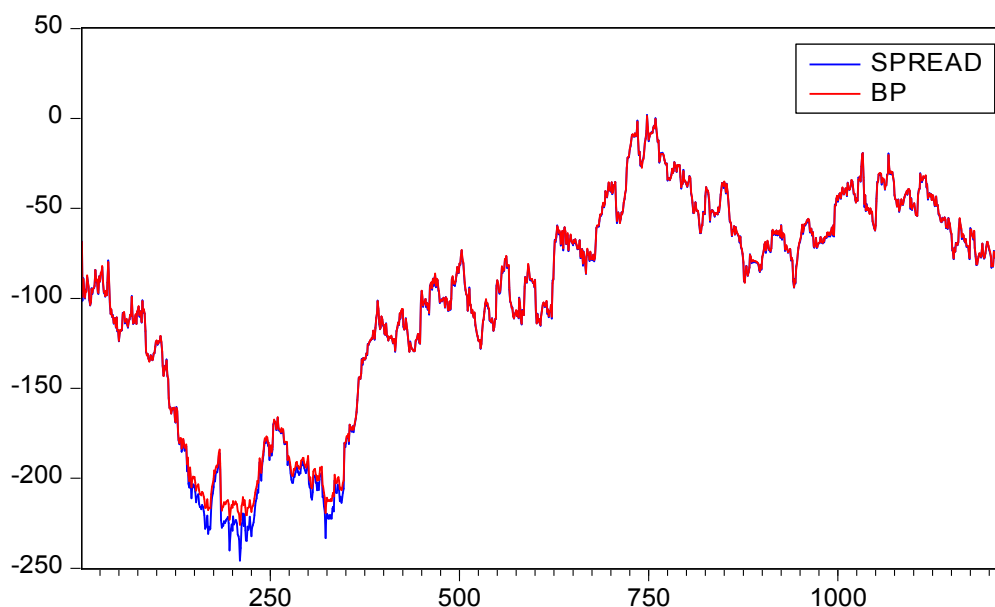
图 4.5 BP 神经网络预测误差收敛图



可以看到，因为 BP 神经网络的结构相对简单，所以收敛速度非常快，当迭代次数为 102 次时，系统就停止了运算，并且当迭代次数为 96 次时，网络的 MSE 达到最小值 0.00055087。运用训练好的 BP 神经网络，我们就可以对样本外的价差作

出预测：

图 4.6 BP 神经网络预测效果图



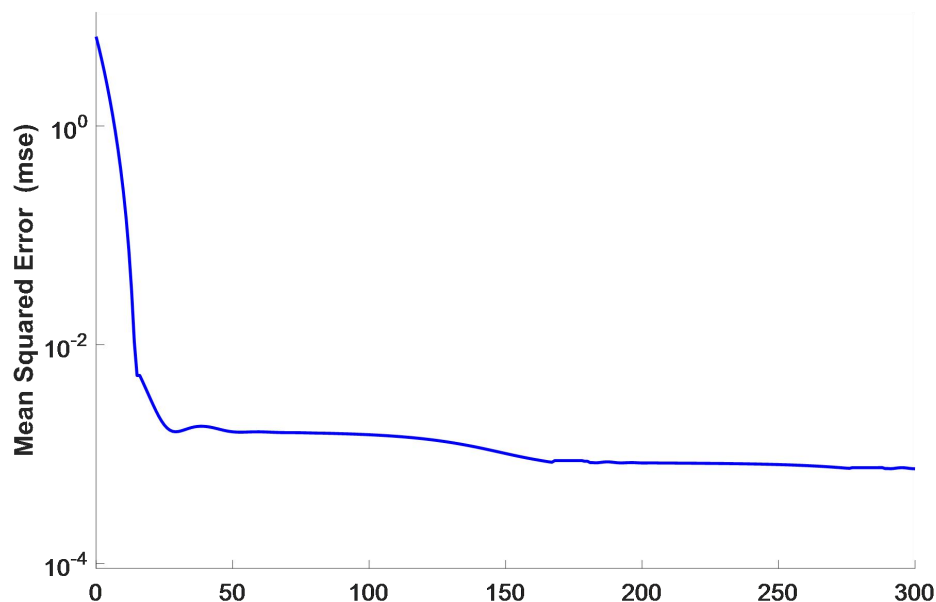
BP 神经网络的预测效果是比较好的，与真实的价差序列基本吻合，但是不难发现，当价差出现较大波动时，其预测效果较差，这可能是由于 BP 神经网络的信息是向前传播的，并不存在反馈的机制，故更适合非线性静态系统的预测，而价差的变化是一个动态的过程，所以在某些时间段内，BP 神经网络的预测结果精度较低，我们可以尝试用动态的 Elman 神经网络对价差序列重新进行预测。

4.4.2 Elman 神经网络

如前所述，Elman 神经网络的内部带有反馈的机制，增加的承接层能够存储和记忆之前输出的信息并传输给隐含层，增强了网络处理动态过程的能力，也就是说，此时隐含层的输入不仅包含前 10 个时刻的价差，同时包括上一次隐含层输出传递到承接层再以一定权重返回的输入值。

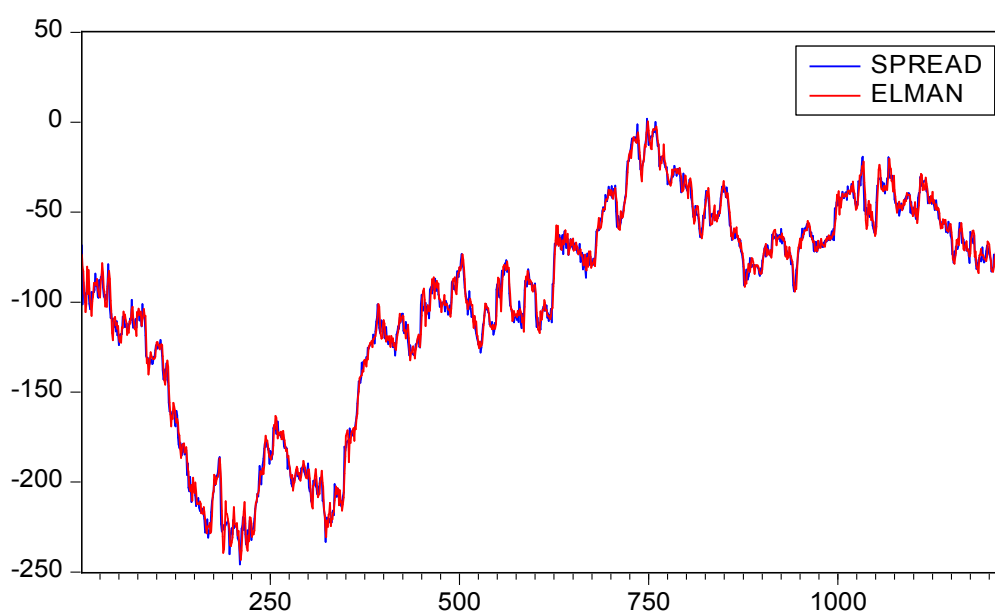
我们令隐含层节点数 $l=6$ ，承接层的节点数与之相同，通过对比，将隐含层函数和输出层函数都设置为 $\text{purelin}(x)$ ，利用 Elman 神经网络，对样本内的价差进行训练，出于一致性的考虑，迭代次数也为 300 次，收敛效果如下：

图 4.7 Elman 神经网络预测误差收敛图



Elman 神经网络的收敛速度要明显慢于 BP 神经网络, 迭代次数达到 150 次时才趋于平稳, 这是因为其网络结构比 BP 神经网络更为复杂, 在迭代了 300 次后, 虽然网络未达到系统设定的梯度, 但其 MSE 达到最小值 0.00073313。下面我们用训练好的 Elman 神经网络对样本外的数据进行预测:

图 4.8 Elman 神经网络预测效果图



通过观察,可以看到 Elman 神经网络的预测效果非常好,与实际的价差序列几乎完全吻合,即使是在价差波动率较大的时间段,也展现出良好的预测能力,这充分体现了该网络在处理动态信息时的优越性。

4.4.3 遗传算法的优化

根据吴仕勇(2006)等学者的研究,神经网络普遍存在收敛速度慢和易陷入局部最小值等缺点,故本文引入遗传算法对 Elman 神经网络进行优化,以获得网络最佳的初始权值和阈值。遗传算法的优化过程按照如下步骤进行:

(1) 确定运行参数

首先,关于种群大小 M ,较大会减慢算法的运行速度,较小会增大找到最优个体的难度,故本文将种群规模设定为 100,其次,遗传代数 G 尽量越多越好,这里令 $G=200$,最后,交叉概率 P_c 一般设置为 0.5,而变异概率则相对较低,于是我们令 $P_m=0.1$ 。

(2) 染色体编码

由于本文采用的染色体编码为实数编码,所以每个个体均为一个实数串,该实数串的长度是根据 Elman 神经网络的结构确定,Elman 神经网络的权值和阈值包括 5 个部分:隐含层和承接层的连接权值 w_1 有 36 个,输入层和隐含层的连接权值 w_2 有 60 个,隐含层和输出层的连接权值 w_3 有 6 个,隐含层阈值 b_1 有 6 个,输出层阈值 b_2 有 1 个,将这些权值和阈值按照顺序排列,就构成了一个染色体长度为 109 的个体。

(3) 个体适应度值的计算

本文个体适应度值 F 的计算方法为训练 Elman 神经网络后预测输出和期望输出之间的误差绝对值和,其中训练的迭代次数为 200。

(4) 遗传操作

通过 100 次进化,即 100 次选择、交叉和遗传操作后,找到适应度值最优的个体,并将实数串的值赋给 Elman 神经网络作为其初始权值和阈值。

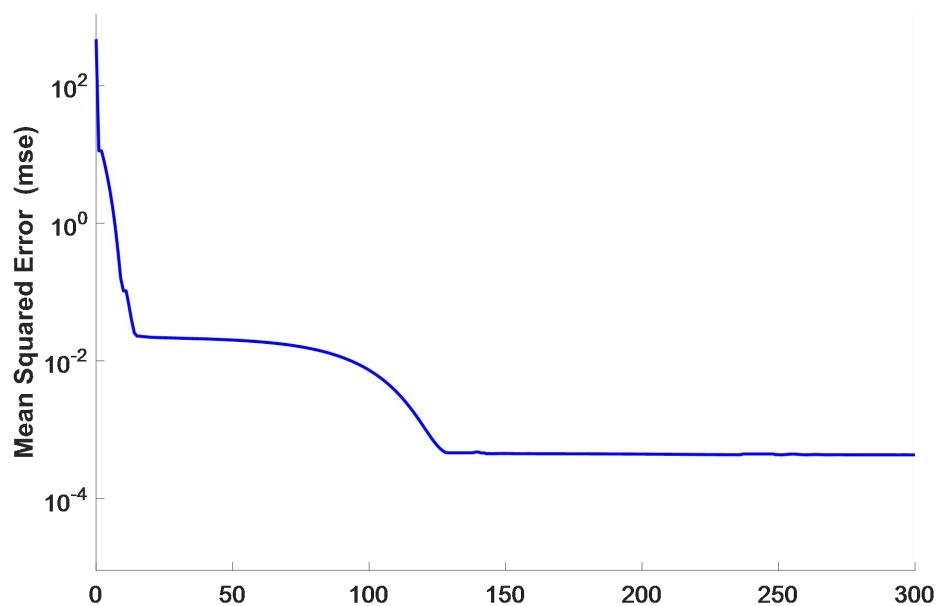
运用遗传算法,我们得到 Elman 神经网络的最优权值和阈值如下:

表 4.10 Elman 神经网络的最优权值阈值

w_1					
-1.9987	-2.222	0.0103	-2.5695	0.7551	0.2902
1.3424	0.6109	-0.2084	-2.6923	-0.8988	2.0141
-0.7657	-2.6156	-1.9246	-0.3942	1.5976	0.8947
2.8784	0.8205	0.611	-2.0766	0.0778	-0.484
2.0424	0.8759	-1.106	-2.3618	-0.0437	0.4131
1.2644	-2.2216	-2.9991	0.2971	-1.7548	0.8289
w_2					
0.8023	-1.3875	1.0535	-1.4925	0.6888	0.5897
-0.5314	1.6746	-1.3441	0.4297	2.1569	-1.1676
1.6675	-1.6855	2.716	-2.8532	-0.0847	2.5734
1.6577	-1.7394	-0.5179	1.0879	1.8438	0.9271
1.9989	-1.5093	-2.6736	2.5284	-0.584	-1.1546
0.6327	-0.239	-1.696	-0.8163	-1.2011	-2.0953
-2.2844	1.6311	2.5139	-2.5867	-1.0893	-2.4431
-2.2695	-0.2699	-1.8896	-0.5047	2.7992	-1.6858
-1.1701	0.8684	2.0897	-0.9047	1.5468	-0.4129
1.8075	1.6635	-0.074	-2.528	1.1639	0.1926
w_3					
1.8687	-1.0492	0.9697	-1.2045	2.7779	-1.796
b_1					
-1.9671	1.505	-2.374	-1.8804	-2.5621	1.8017
b_2					
1.8687					

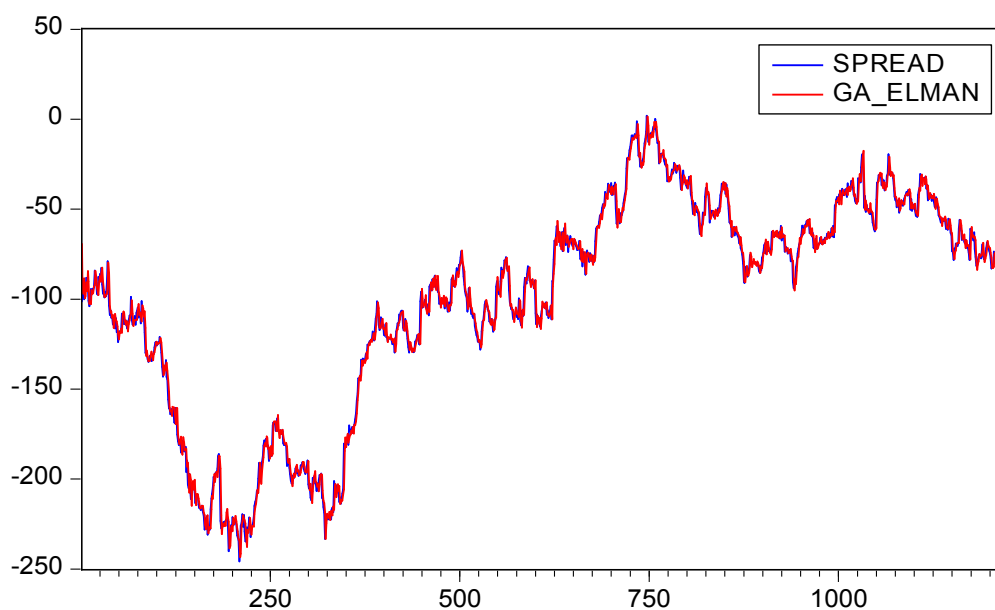
用优化好的 Elman 神经网络再次对样本内的价差进行训练，为了保持对比性，这里的迭代次数同样选择为 300 次。

图 4.9 遗传算法优化 Elman 神经网络预测误差收敛图



同时对样本外的价差进行预测：

图 4.10 遗传算法优化 Elman 神经网络预测效果图



从预测误差收敛图看,优化后的 Elman 神经网络在前 100 次迭代的过程中收敛速度不及原来的网络,但是在 100 次之后,其收敛速度急剧加快, MSE 快速逼近梯度允许范围内的极小值,当迭代次数达到 299 次时,优化的 Elman 神经网络预测误差的 MSE 达到最小值 0.00043057,远小于之前的模型,表明遗传算法能够提高网络的收敛速度。从预测效果图看,遗传算法优化后的 Elman 神经网络预测能力相当的强,与样本外的价差序列基本完全吻合。不过要比较优化前后哪一种模型更优,还需要我们建立具体的指标加以判断。

4.4.4 三种模型的比较

为了对三种模型进行比较,我们设置如下两个指标:

$$R^2 = \frac{(n \sum_{i=1}^n e_i \bar{e}_i - \sum_{j=1}^n e_j \sum_{k=1}^n \bar{e}_k)^2}{(n \sum_{j=1}^n e_j^2 - (\sum_{j=1}^n e_j)^2)(n \sum_{k=1}^n \bar{e}_k^2 - (\sum_{k=1}^n \bar{e}_k)^2)} \quad (4.9)$$

$$MSE = \frac{\sum_{i=1}^n (e_i - \bar{e}_i)^2}{n} \quad (4.10)$$

R^2 为平方相关系数,衡量的是两序列间相关关系的强弱,即预测序列对原序列的拟合程度, R^2 应当越大越好, MSE 为均方误,考察的是预测序列对原序列的平均偏离程度, MSE 应越小越好。我们分别计算三种神经网络这两项指标的值:

表 4.11 三种神经网络的 R^2 和 MSE

	R^2	MSE
BP 神经网络	0.9925	30.9117
Elman 神经网络	0.9878	42.9100
优化的 Elman 神经网络	0.9928	25.3764

可以看到,在两项指标上,遗传算法优化的 Elman 神经网络与另外两种神经网络相比都是更优的,说明遗传算法可以加快神经网络的收敛速度并使模拟的非线性

系统获得全局初始最优解,从而能够显著提高网络的预测能力,与此同时,Elman 神经网络的 R^2 小于 BP 神经网络,并且其 MSE 也大于 BP 神经网络,这是因为 Elman 神经网络的结构比较复杂,在相同的迭代次数下,并未达到理想的收敛效果,但是在对波动性较大的价差序列的估计上,Elman 神经网络显然更胜一筹,这得益于其结构中增加了一个承接层,使网络对历史信息的敏感性更强,增强了网络处理动态过程的能力。

4.5 套利的实现

4.4.1 套利的假设条件

理论研究不同于实际交易,一方面需要考虑实际交易中存在的诸多复杂因素,另一方面要使我们的理论研究变得简单易行,具有一定的说服力,因此在建立套利策略之前,我们需要作出一些适当的假设和简化部分条件,基于此,本文给出如下基本假设:

- (1) 三种期货的开平仓操作均以 5 分钟的收盘价进行交易的结算;
- (2) 交易所会根据市场情况不断调整期货的保证金率,根据最新规定,螺纹钢期货的保证金率为 9%,铁矿石期货和焦炭期货的保证金率同为 8%,为了尽量模拟实际交易,本文采用同样的保证金率标准,与此同时,现实情况下,螺纹钢期货的手续费率一般高于铁矿石期货和焦炭期货,出于高频交易的特点以及样本外时间区间长度的考虑,本文将螺纹钢期货的手续费率设置为万分之零点二,铁矿石期货和焦炭期货的手续费率设置为万分之零点一,手续费按照双边收取,真实的手续费率一般高于文章中的设定,并且依据开平仓的时间而有所不同,但回顾相关的研究,我们可以通过延长交易的时间段以及扩大阈值来覆盖掉多余的手续费,因此本文的规定是较为合理的。
- (3) 期货交易以 1 手为基本单位,螺纹钢期货 1 手为 10 吨,铁矿石期货和螺纹钢期货一手为 100 吨,而期货的价格单位是吨/元,所以在实际交易中,我们必须对三种期货的单位进行统一,除此之外,套利比例是存在小数点的,因此本文认为期货

合约是可分割的。综上，本文规定一次价差的多头套利开仓相当于买入 10 手螺纹钢期货，卖出 1.843651 手铁矿石期货和 1.035053 手焦炭期货，一次价差的空头套利开仓相当于卖出 10 手螺纹钢期货，买入 1.843651 手铁矿石期货和 1.035053 手焦炭期货。

(4) 期货交易具有杠杆效应，也就是说，保证金制度的存在，能够让投资者以数倍于本金的资金量参与到期货交易中，这些资金可以来源于市场上的借贷。本文将市场上的借贷利率设置为无风险利率，一般采用一年期贷款基准利率，目前为 4.35%。

(5) 根据期货交易的保证金率和手续费率、套利交易的买卖次数和手数以及无风险利率等因素综合考虑，本文假设保证金账户的资金为 100 万，其中包括初始自有资金 20 万和借入资金 80 万。

4.4.2 套利策略的构建

区别于传统的均值回归策略，本文将基于神经网络构建新的套利策略。首先利用神经网络，在已知包含目前时刻的前 10 个时间段的价差的情况下，我们可以对下一个 5 分钟的价差进行有效预测。这里，我们记下一时刻预测价差相对于前一刻（也就是目前时刻）实际价差的增量为 Δe_t ，表示如下：

$$\Delta e_t = \bar{e}_t - e_{t-1} \quad (4.11)$$

根据 Δe_t 的变化，我们就能快速进行三种期货的买卖，实现多头套利和空头套利，具体的操作方法为：（1）如果 $\Delta e_t > X$ （ X 是一个确定的正数），意味着价差在未来会上涨一定的幅度，我们需要采取多头套利，若当前持有的是多头头寸，就选择持仓不动，若当前持有的是空头头寸，就将该空头头寸平仓后建立多头头寸，若当前没有持有头寸，就重新建立多头头寸；（2）如果 $\Delta e_t < -X$ ，意味着价差在未来会下跌一定的幅度，此时我们需要采取空头套利，若当前持有的是多头头寸，就将多头头寸平仓后建立空头头寸，若当前持有的是空头头寸，就继续持仓不动，若目前没有持有头寸，就重新建立头寸；（3）如果 $-X < \Delta e_t < X$ ，不构成套利的条件，

就将当前可能持有的多头和空头头寸平仓。

我们可以初步考察神经网络是否能正确预测价差的变化，这里记下一时刻价差相对于前一时刻价差的实际增量为 $\Delta e'_t$ ，三种神经网络价差的预测增量和实际增量的对比图如下：

图 4.11 BP 神经网络价差的预测增量和实际增量

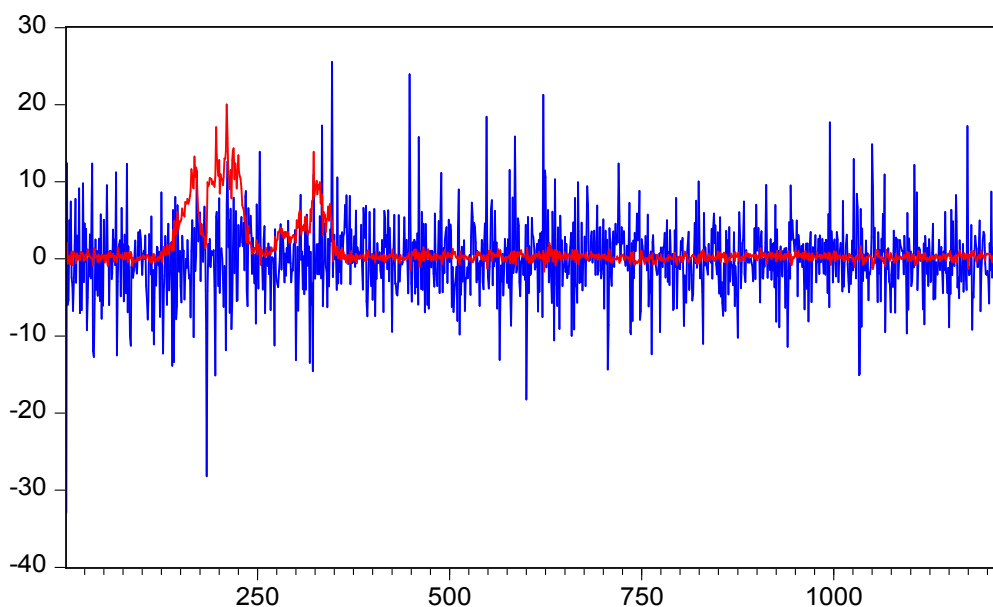


图 4.12 Elman 神经网络价差的预测增量和实际增量

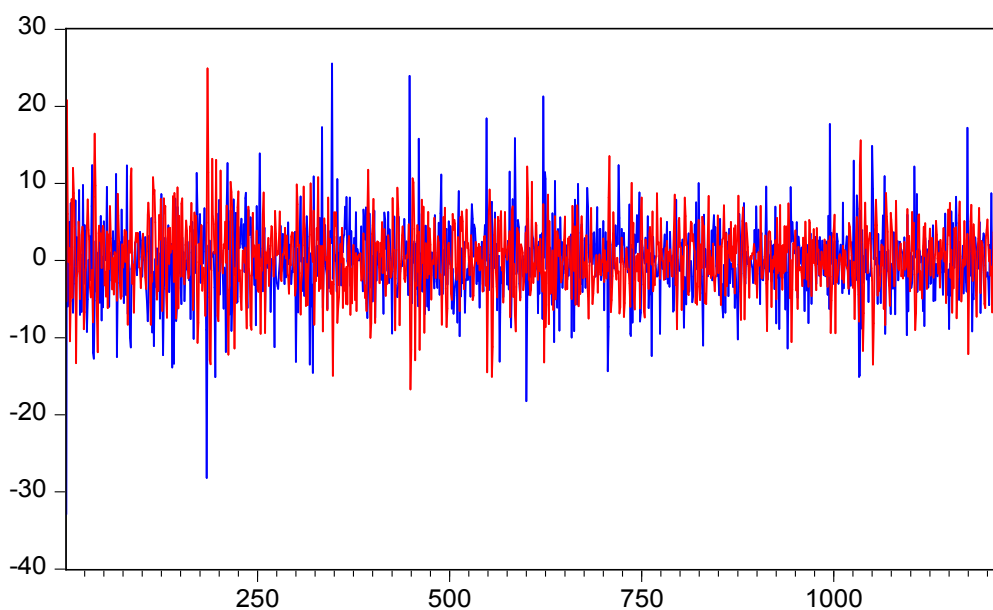
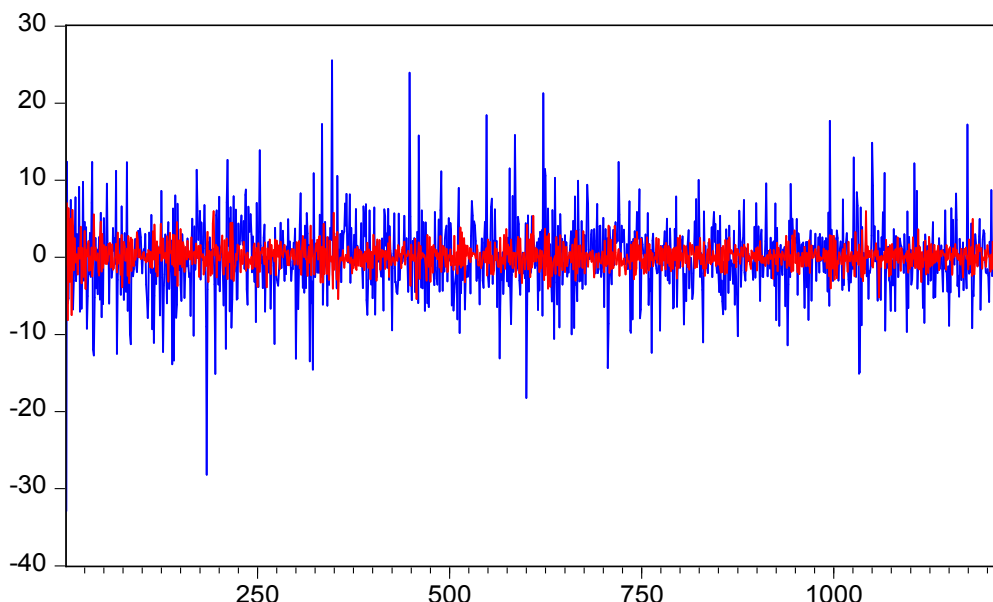


图 4.13 遗传算法优化 Elman 神经网络价差的预测增量和实际增量



图中蓝线代表了价差的实际增量，红线代表了预测增量。通过观察可以发现，对比 BP 神经网络，Elman 神经网络和遗传算法优化的 Elman 神经网络能够捕捉到价差变动的波动特性，具有更好的预测能力。我们还统计了样本外 1218 个时间点神经网络对价差变化预测方向正确的个数，BP 神经网络共预测正确 639 个时间点，而 Elman 神经网络和遗传算法优化的 Elman 神经网络则分别预测正确 650 个和 643 个时间点，三种神经网络的预测正确率都超过了 50%，而一般趋势策略的胜率基本都在 20%-50% 之间，这说明神经网络对于价差未来变化的方向具有极强的判断能力，为我们的套利奠定了良好的基础。

4.4.3 套利策略的回测

根据上海期货交易所的规定，螺纹钢期货价格的最小变动单位为 1 元/吨，根据大连商品交易所的规定，铁矿石期货和焦炭期货价格的最小变动单位为 0.5 元/吨，由于本文采用的是高频策略，所以阈值 X 的取值不宜过大，否则会导致交易次数明显过少，为了对比三种模型的套利效果，同时考虑到期货价格的最小变动单位，本文考察了 X 的取值为 0、0.25 和 0.5 三种情况下套利策略的获利能力，运用的指标如下：

(1) 总盈利

指策略回测期间的收益减去手续费以及借贷的费用，手续费按照双边收取，借贷的费用包含休息日所产生的利息。

(2) 年化收益率

该指标的计算公式如下：

$$\text{年化收益率} = \frac{\text{总盈利}}{\text{初始自有保证金}} \times \frac{\text{回测天数}}{365} \times 100\% \quad (4.12)$$

其中本文的回测天数为 24 天。

(3) 交易次数

指超过阈值标准 X ，产生交易信号产生的次数，可以用开仓次数来衡量，开仓包括了多头开仓和空头开仓。

(4) 盈利次数

指在所有开仓交易中，平仓后获得收益的次数，其中收益不考虑套利过程中的利息支出。

(5) 胜率

指套利策略的成功率，计算的方法一般是用盈利次数除以交易次数。

(6) 最大单日收益

指在所有交易日中单个交易日所能获得的最大收益，这里同样不加入借贷费用。

(7) 单笔平均收益

指每笔套利交易获得的平均收益，计算方法是用总收益除以交易次数。

(8) 平均持有期

指每笔套利交易从开仓到最后平仓持有头寸的平均时间，由于本文所有数据的时间间隔为 5 分钟，所以这里采用持有头寸时间点的个数除以交易次数来计算该指标。

下面给出不同阈值下三种神经网络的年化收益率曲线：

图 4.14 $X = 0$ 时三种神经网络的年化收益率

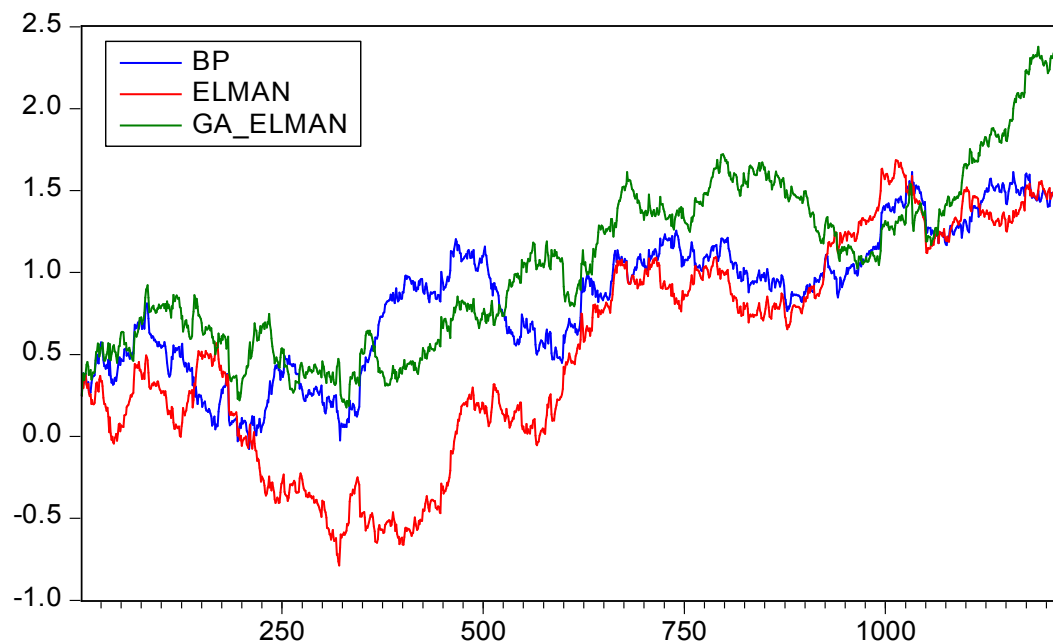


图 4.15 $X = 0.25$ 时三种神经网络的年化收益率

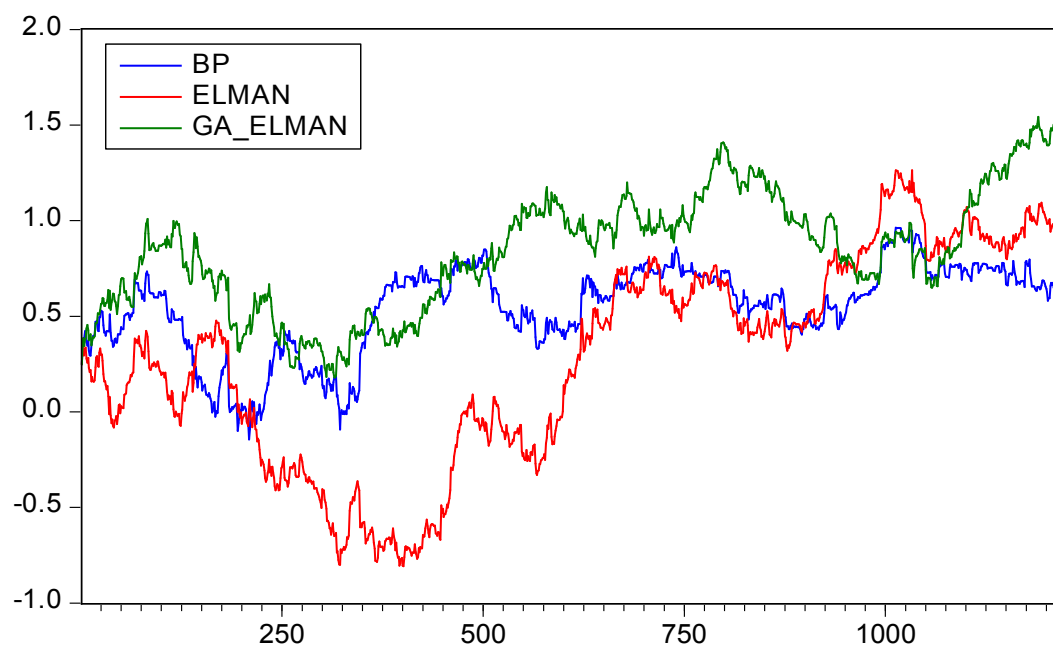
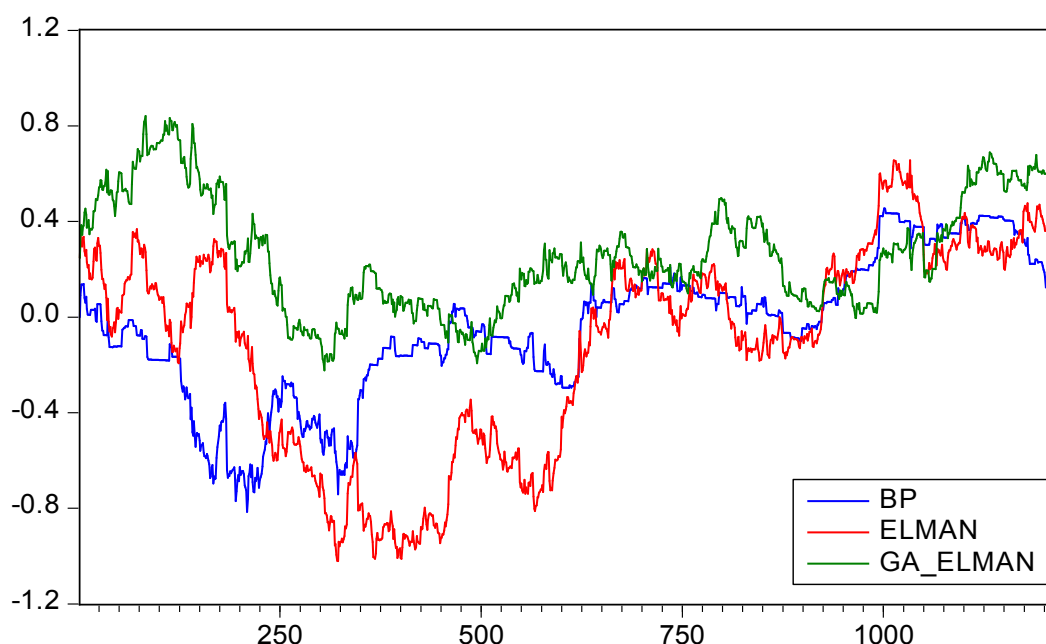


图 4.16 $X = 0.5$ 时三种神经网络的年化收益率



从年化收益率曲线上看，对于任何阈值，在大部分的时间段，遗传算法优化的Elman神经网络表现都是最优的，根据该网络预测的价差序列所建立的策略收益率保持了持续稳定的增长。对比而言，Elman神经网络的最终年化收益率尽管表现不错，但收益率的波动幅度过大，尤其是在前期，其收益率一度跌至-80%。与前两种神经网络相比，BP神经网络实现的收益率一般，虽然收益率基本维持增长的势头，但增长速度偏慢。下面我们将计算三种神经网络的获利能力指标：

表 4.12 $X = 0$ 时三种神经网络的套利指标评价

指标	BP	Elman	GA_Elman
总收益	19679.60	20464.35	28089.08
年化收益率	149.65%	155.61%	213.59%
交易次数	386	528	759
盈利次数	229	333	407
胜率	59.33%	63.07%	53.62%
单日最大收益	3268.18	3268.18	3268.18
单笔平均收益	50.98	38.76	37.01
平均持有期	3.16	2.31	1.60

表 4.13 $X = 0.25$ 时三种神经网络的套利指标评价

指标	BP	Elman	GA_Elman
总收益	8835.40	16079.34	17307.54
年化收益率	67.19%	104.87%	131.61%
交易次数	359	517	741
盈利次数	202	320	383
胜率	56.27%	61.89%	51.28%
单日最大收益	3268.18	3268.18	3268.18
单笔平均收益	24.14	30.22	23.36
平均持有期	3.38	2.29	1.64

表 4.14 $X = 0.5$ 时三种神经网络的套利指标评价

指标	BP	Elman	GA_Elman
总收益	2265.90	5599.54	6580.26
年化收益率	17.23%	42.58%	50.04%
交易次数	242	507	667
盈利次数	134	313	344
胜率	55.37%	61.73%	51.57%
单日最大收益	2541.57	3268.18	3268.18
单笔平均收益	9.36	11.04	9.86
平均持有期	5.03	2.40	1.82

为了确保策略是有效的，我们还需对上述指标进行检验，为此建立原假设：

H_0 ：交易策略能够获得盈利是出于偶然因素

构造的 Z 检验统计量如下：

$$Z = \frac{Y - np}{\sqrt{np(1-p)}} \quad (4.13)$$

其中， Y 为回测期策略盈利的次数， n 为回测期所有的交易次数， p 为策略取得收益的概率，一般被赋值为 0.5。三种模型不同阈值下的 Z 检验统计量的值如下：

表 4.15 三种神经网络不同阈值下 Z 检验统计量的值

	BP	Elman	GA_Elman
$X = 0$	3.66	6.01	2.00
$X = 0.25$	2.37	4.68	0.92
$X = 0.5$	1.67	4.13	0.81

注：Test critical values: 1% level 2.58 5% level 1.96 10% level 1.65

可以看到，在 10% 的显著性水平下，除了遗传算法优化的 Elman 神经网络在 $X = 0.25$ 和 0.5 时 Z 检验统计量的值无法拒绝原假设外（主要是由于交易次数较多，胜率一般所导致），其他模型在不同阈值下的统计值都是显著的，表明基于神经网络构建的策略基本有效。

从阈值 X 的变动对收益率的影响来看，随着 X 的增大，三种神经网络的总收益和年化收益率都出现了大幅度的下降，导致策略的单笔平均收益减少。本文认为，这主要由于阈值增大后，交易次数出现一定程度的减少所引起，当交易次数递减时，策略的平均持有期也会相应延长，弱化手续费率的影响，盈利当然会有所减少。对比三种神经网络， X 的变化对 BP 神经网络的影响是最大的，同时不难发现，该网络的交易次数在不同的阈值下也都是最少的，因为 BP 神经网络属于前馈式的网络，相较于反馈式的 Elman 神经网络，对历史信息的记忆能力明显不足，无法反映出价差动态变化过程的特性，这可以从价差的实际增量和预测增量图看出，BP 神经网络价差的实际增量大都分布在 0 到 1 的范围内，唯一出现较大波动的是在 128-254 和 256-348 两个时间段，上述时间段内，价差实际增量的表现是比较反常的，一般而言，神经网络对价差变化的预测不可能完全准确，所以预测增量会呈现出围绕 0 上下波动的特征，但根据 BP 神经网络的预测，两个时间段里，预测增量长期偏离了均衡，并停留在 0 的上方，与实际增量震荡的特征完全不符，当 $X = 0$ 时，就意味着我们在 128 和 256 的时间点执行多头开仓操作后，直到 254 和 348 两个时间点到来之时才能进行平仓，此时 BP 神经网络的预测完全失准，同样地，Elman 神经

网络也存在类似的问题，尽管该网络对动态系统具备良好的逼近能力，预测增量呈现出波动性的特点，但是受制于收敛速度一般的缺点，拟合精度有限，基于网络构建的策略产生交易信号的次数以及盈利的次数仍然明显少于遗传算法优化后的网络。依据我们之间的统计，三种网络在预测价差变化的方向时，正确的个数其实是差不多的，可遗传算法优化后的 Elman 神经网络之所以能够产生最多的收益，无外乎在于其交易次数和盈利次数是最多的，也就是说，该网络能够给出更多的有效预测，预测效果更加符合实际增量波动的特点，尽管胜率不及另外两种神经网络，可是考虑到网络的交易基数较大，胜率基本维持在 50% 以上，高于一般趋势策略 20%-50% 的胜率，这也是为我们所接受的。利用遗传算法获得 Elman 神经网络的最优初始权值和阈值后，除了 Elman 神经网络本身的优点外，我们加快了网络的收敛速度，使其具备更加良好的拟合精度和更强的预测能力，提高了策略的获利次数。遗传算法优化的 Elman 神经网络的优势还体现在能使我们持续稳定的盈利，这一点与 Elman 神经网络存在很大的不同，观察 Elman 神经网络的年化收益率曲线，可以发现其波动性太大，对于绝大部分市场投资者而言，是难以忍受超过 10% 以上的亏损的，但该网络的最大跌幅曾接近 80%。本文认为出现这种情况的原因在于 Elman 神经网络在收敛的过程中可能陷入了局部最优，这一点在预测的过程中体现了出来，Elman 神经网络的年化收益率曲线大致呈 U 字型，换句话说，该网络在策略前半时间段的预测几乎失准，但在策略后半时间段的预测却异常精确，运用遗传算法后，我们使神经网络获得了全局最优解，令网络在大部分时间的预测都趋于稳定，策略的风险得到有效改善。只不过，优化后的网络仍存在胜率一般、持有期过短导致手续费偏高的缺点，这也是今后值得我们不断改进的地方。

5.结论

本文摒弃了传统的统计套利方法，基于人工智能领域的研究热点—神经网络构建了跨品种套利的策略，取得了可观的收益。

在跨品种套利的策略选择上，与一直以来沿用的统计套利方法相比，神经网络具备诸多的优势。统计套利的核心思想是均值回归，认为价差偏离均值后一定会回复，某种程度上，这种回复具有不可测性，而神经网络运用的是其拟合和预测的功能，具体来讲，先是利用前几期的价差作为输入，下一期的价差作为输出来训练网络，再用网络对样本外的价差进行预测，所以神经网络方法具有一定的前瞻性，这也注定了二者产生交易信号的原理完全不同。对于统计套利，价差偏离达到一定幅度时会产生开仓信号，回复到均值附近时会提示提前进行平仓，而对于神经网络，只要预测下一期的价差上涨或下跌一定幅度就能产生开仓信号，如果价差并未发生趋势性的变化，就一直不会平仓。由上，神经网络模型的优点体现在：第一，神经网络拥有特殊的结构和误差修正机制，适合于非线性系统建模，反馈式的网络更是能捕捉到价差的波动性特征，换言之，神经网络方法可以依赖于价差本身的特点进行策略的构建，而统计套利方法则依靠于均值和标准差这样的外部参数，建立策略时会受到很多的限制。第二，均值回复策略的持仓时间取决于价差何时回复到均值附近，如果说价差迟迟无法回复，或者说在相当长的一段时间内都没有触及到平仓点，投资者就有可能选择提前平仓，不可避免地造成一定的损失，同时回复周期过长也会导致交易次数偏少，对比之下，神经网络策略的开平仓操作只取决于价差的上涨或下跌，这意味着它能够产生比统计套利更多的交易机会；第三，如同大多数金融时间序列一样，价差序列同样会在某些时间段呈现出集群效应，当价差发生巨大波动时，统计套利方法往往会设置一个止损点，也就是说，在这种情况下，损失的发生具有必然性，而神经网络方法能够抓住价差连续上涨或下跌的趋势进行套利，扩大盈利的空间。考虑到上述优点，本文选择神经网络方法对跨品种套利展开研究。

本文的研究对象是黑色金属产业链中三种具备代表性的期货：螺纹钢期货、铁矿石期货和焦炭期货，这三种期货的上市时间都不超过 10 年，且在国内的相关研究偏少，故具有一定的研究价值。我们运用 EG 两步法考察了三种期货两两配对以及全部结合的协整关系，计算出它们的价差序列，得到了螺纹钢期货对焦炭期货以及螺纹钢期货对铁矿石期货和焦炭期货两对可行的套利组合，根据回归的拟合程度、价差序列的特点与是否具有现实意义，确定三种期货组合为最终的套利对象。接着利用前 10 个交易时间点的价差作为输入向量，下一时间点的价差作为输出向量对神经网络进行了训练，在神经网络的选择上，前人大都使用的是静态的 BP 神经网络，该网络属于前馈式网络，信号只向前传递，无法完整捕捉到价差的波动特性，为此我们引入了反馈式的 Elman 神经网络，它在 BP 神经网络的基础上加入了一个承接层，用于记录历史信息，从而使网络具备更强的处理动态过程的能力，之后我们运用遗传算法，针对神经网络收敛速度慢和易陷入局部最小值的缺点，对 Elman 网络进行了优化。为了体现不同模型之间的区别，我们将三种训练好的神经网络同时用于价差的预测，都取得了不错的效果，其中尤以遗传算法优化的 Elman 神经网络对样本外价差序列拟合的效果最佳，预测的偏离程度最小。在此基础上，我们设定了诸多假设，以预测的价差增量是否超过一定的阈值作为开仓条件，构建了黑色金属期货跨品种套利的策略，发现在大部分时间段内，遗传算法优化的 Elman 神经网络能够获得最优的年化收益率，并且收益能够保持持续稳定的增长，这一方面得益于 Elman 神经网络自身在非线性动态系统建模方面的优势，在利用遗传算法优化后，网络收敛速度加快，预测正确率提高，产生了更多的交易次数和有效盈利次数，增加了收益，另一方面遗传算法赋予了神经网络能够达到全局最优的初始权值和阈值，使得网络在绝大部分时间都能达到一定的预测精度，增强了获利的稳定性。整体来看，在跨品种套利上，遗传算法优化的 Elman 神经网络是本文首要选择的最佳模型。

尽管如此，本文仍然存在改进的空间。首先，我们正处于人工智能大发展的时代，神经网络也会在解决实际问题的过程中不断被创新，本文采用的 Elman 神经网络虽然同一般的 BP 神经网络存在一定的区别，但是从结构上来讲依然是较为简单

的，能否改变 Elman 神经网络的结构或者是应用一种新的神经网络模型来提高对价差序列的拟合和预测精准度，这是值得我们今后所要关注的问题。除此之外，本文在实施跨品种套利的过程中设定了诸多假设条件，尤其是在手续费的设置上缺乏一定的合理性，实际交易过程要比本文预设的套利场景要复杂的多，例如滑点、期货合约不可分割的问题都会存在，这些同样是我们今后优化的重点。最后，本文运用神经网络模型构建了黑色金属产业链中具有代表性的三种期货的跨品种套利策略，那么对于其它相关性程度较高的期货品种，神经网络方法是否仍然适用，并且改进后的 Elman 神经网络是否仍在大部分的时间区间内保持最高的预测正确率和最优的年化收益率，这两点都有待后续的研究加以考量。

对于个人投资者而言，在资金量有限且不具备交易背景的前提下，他们更不可能去研究本文的算法并应用于套利交易中，但是对于机构投资者，本文所提出的新的跨品种套利的思想和方法是符合目前研究趋势的，如今我们正处于人工智能大发展的时代，投资机构完全可以运用他们在信息和技术方面的优势掌握神经网络这一智能算法，并开发出更为出色的神经网络模型，将其应用到金融的各个领域，实现科技与金融的完美融合。另外，从本文研究的期货品种的角度出发，对于黑色金属产业链的下游企业，尤其是钢铁厂，他们也越来越广泛地加入到相关期货的交易中，神经网络模型可以对未来的价差，也就是其冶炼价值进行预测，从而更好地帮助这些企业达到套期保值的目的。总体来看，本文不仅具备了理论上的价值，还从实践的角度出发，为机构投资者提供了思路，为实体企业提供了宝贵的建议，这使得本文具有一定的现实意义。

致谢

本文是在导师简志宏教授的悉心指导下完成的，简老师在毕业设计的每一环节都给我提出了针对性的建议，使我受益良多，最终成功地完成了这篇论文。简老师丰富的学识和独特的人格魅力让我钦佩不已，并且在择业方面为我指明了方向，在此我想向简老师表达自己深深的感激之情。

作为金融专硕的一名学生，感谢经济学院两年以来对我的培养，感谢金融系曾经教过和帮助过我的老师，面对我的提问，你们总是不厌其烦地给予解答，在你们身上，我感受到的不仅仅是渊博的学识，更是严谨踏实的学风，我相信在未来的学习和工作过程中，我一定会牢记你们的教诲，继续努力学习、完善自我，争取成为一名高水平的金融人才，在此，请再让我向你们表达诚挚的感谢。

最后，感谢硕 1704 班的全体同学，在两年的时间里，学习上我们互相合作，共同提高，在生活中我们互相关心，共同成长，希望我们的友谊长存，还要感谢一直以来照顾和支持我的父母、家人，我一定会用实际行动来回报你们，一定会为了梦想继续不懈奋斗！

参考文献

- [1] Wahab M, Cohn R, Lashgari M. The Gold-silver Spread: Integration, Cointegration, Predictability and Ex-ante Arbitrage[J]. Journal of Futures Markets, 1994, 14(6): 709-756.
- [2] Liu S M, Chou C H. Parities and Spread Trading in Gold and Silver Markets: A Fractional Cointegration Analysis[J]. Applied Financial Economics, 2003, 13(12): 879-891.
- [3] Tzang D N, Leuthold R M. Hedge Ratios Under Inherent Risk Reduction in A Commodity Complex[J]. Journal of Futures Markets, 1990, 10(5): 497-504.
- [4] Johnson R L, Zulauf C R, Irwin S H, et al.. The Soybean Complex Spread: An Examination of Market Efficiency from the Viewpoint of a Production Process[J]. Journal of Futures Markets, 1991, 11(1): 25-37.
- [5] Rechner D, Poitras G. Putting on the crush: Day trading the soybean complex spread[J]. Journal of Futures Markets, 1993, 13(1): 61-75.
- [6] Simon D P. The Soybean Crush Spread: Empirical Evidence and Trading Strategies[J]. Journal of Futures Markets, 1999, 19(3): 271-289.
- [7] Mitchell, John B. Soybean Futures Crush Spread Arbitrage: Trading Strategies and Market Efficiency[J]. Journal of Risk & Financial Management, 2010, 3(1): 63-96.
- [8] Girma P B, Paulson A S. Seasonality in Petroleum Futures Spreads[J]. Journal of Futures Markets, 1998, 18(5): 581-598.
- [9] Haigh M S, Holt M T. Crack Spread Hedging: Accounting for Time-varying Volatility Spillovers in the Energy Futures Markets[J]. Journal of Applied Econometrics, 2002, 17(3): 269-289.
- [10] Cummins M, Bucca A. Quantitative Spread Trading on Crude Oil and Refined Products Markets[J]. Quantitative Finance, 2011, 12(12): 1857-1875.
- [11] Dunis C L, Laws J, Evans B. Modelling and trading the soybean-oil crush spread with

- recurrent and higher order networks:A comparative analysis[J].Neural Network World,2006,16(3):192-213.
- [12]Phoebe S W,David E.Nonlinear Modeling Using Neural Networks for Trading the Soybean Complex[J].Procedia Computer Science,2014,36:234-239.
- [13]Dunis C L,Laws J,Middleton P W,et al.Trading and hedging the corn/ethanol crush spread using time-varying leverage and nonlinear models[J].The European Journal of Finance,2015,21(4):352-375.
- [14]汪媛,孙建明.铜铝期货跨商品套利分析与研究[J].湖北经济学院学报(人文社会科学版)[J],2013,10(01):45-47.
- [15]殷晓梅,孙涛,徐正栋.基于高频数据的棕榈油与豆油期货跨商品套利可行性研究[J].农村经济与科技,2008,(08):84-86.
- [16]王丽.大连商品交易所棕榈油与豆油期货价格之间的套利研究[J].现代商业,2017,(05):84-85.
- [17]吴静雯.基于协整的期货跨品种套利策略分析[D].广州:暨南大学,2018.
- [18]武军伟,史永东.大连商品交易所大豆压榨利润套利的实证研究[J].金融经济,2008,(22):106-107.
- [19]顾全,雷星晖.基于协整的豆类期货统计套利实证研究[J].统计与决策,2015,(07):151-154.
- [20]翁泽波.中国豆类期货市场套利研究[D].成都:西南财经大学,2014.
- [21]张威波,胡艳英.基于高频数据的豆类期货套利实证研究[J].时代金融,2018,(06):155-156+158.
- [22]丁秀玲,华仁海.大连商品交易所大豆与豆粕期货价格之间的套利研究[J].统计研究,2007,(02):55-59.
- [23]陈承旭.国内期货市场跨商品套利模型及实证研究[D].成都:西南财经大学,2010.
- [24]殷晓梅.商品期货跨品种套利交易模型及其实证研究[D].南京:南京航空航天大学,2009.
- [25]周亮.基于协整的期货跨品种套利研究—以黑色系期货为例[J].价格理论与实

践,2017,(04):112-115.

[26]周亮.GARCH 模型与协整模型在跨商品套利中的比较研究—以铁矿石和螺纹钢期货为例[J].山东财经大学学报,2017,29(05):54-60+77.

[27]王镭,陈芳.焦炭与螺纹钢分市场跨品种套利策略[J].经贸实践,2017,(12):94.

[28]刘建和,梁仁方,王玉斌等.大豆期货合约均值回归套利策略和 Elman 神经网络套利策略对比研究[J].湖南财政经济学院学报,2016,32(03):13-20.

[29]Madhavan A,Smidt S.A Bayesian Model of Intraday Specialist Pricing[J].Journal of Financial Economics,1991,30(1): 99-134.

[30]靳朝翔,梁仁方,刘建和.基于神经网络模型的商品期货跨品种套利策略—以焦炭、铁矿石和螺纹钢为例[J].云南财经大学学报,2016,32,(04):150-160.

[31]Engle R F, Granger C W J. Cointegration and error correction: Representation, estimation, and testing[J]. Econometrical,1987,55(4):251-276.

[32]Johansen S,Juselius K. Maximum Likelihood Estimation and Inference on Cointegration: With Application to the Demand for Money[J]. Oxford Bulletin of Economics and Statistics, 1990,52(2):169-210.

[33]Engle R F.Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation[J]. Econometrica. 1982,50(4):987-1007.

[34]Bollerslev T.Generalized Autoregressive Conditional Heteroskedasticity.Journal of Econometrics[J]. 1986,31(3):307-327.

[35]Kulkarni S,Haider I.Forecasting Model for Crude Oil Price Using Artificial Neural Networks and Commodity Futures Prices[J].International Journal of Computer Science and Information Security,2009,2(1):81-88.

[36]Panella M,Barcellona F,Santucci V,et al..Neural Networks to Model Energy Commodity Price Dynamics[A].In Proceeding of the USAEE/IAEE North American Conference[C],2011.

[37]汪劲松,石薇.多种群遗传神经网络在股指预测中的应用[J].统计与决策,2014,(14) 76-79.

- [38]董皓舒.基于智能神经网络组合预测的国债利率期限结构建模与实证[J].中国市场,2010,(44):47-49.
- [39]Donaldson R G,Kamstra M.An artificial neural network-GARCH model for international stock return volatility[J].Journal of Empirical Finance,1997,4(1):17-46.
- [40]宋策.基于 BP 神经网络与灰色关联的黄金期货价格预测[J].上海工程技术大学学报,2017,31(01):90-94.
- [41]吴仕勇.基于数值计算方法的 BP 神经网络及遗传算法的优化研究[D].昆明:云南师范大学,2006.
- [42]Cheng-Few L,John C L.Handbook of Financial Econometrics and Statistics [M].Springer,2015:761.
- [43]王小川,史峰,郁磊等.MATLAB 神经网络 43 个案例分析[M].北京:北京航空航天大学出版社,2013.

附录

下面给出全文的部分代码:

数据导入及样本集和测试集的构建:

```
spreadt=xlsread('C:\Users\Administrator\Desktop\毕设\spread.xlsx');
Y_Test=spreadt(end-1700:end)';
Y_Train=spreadt(11:end-1701)';
X_Test=[spreadt(end-1710:end-10),spreadt(end-1709:end-9),spreadt(end-1708:end-8),spreadt(end-1707:end-7),...

spreadt(end-1706:end-6),spreadt(end-1705:end-5),spreadt(end-1704:end-4),
spreadt(end-1703:end-3),...
spreadt(end-1702:end-2),spreadt(end-1701:end-1)]';
X_Train=[spreadt(1:end-1711),spreadt(2:end-1710),spreadt(3:end-1709),spreadt(4:end-1708),...

spreadt(5:end-1707),spreadt(6:end-1706),spreadt(7:end-1705),spreadt(8:end-1704),...
spreadt(9:end-1703),spreadt(10:end-1702)]';
[inputn,inputps]=mapminmax(X_Train);
[outputn,outputps]=mapminmax(Y_Train);
```

BP 神经网络的训练和预测:

```
net=newff(inputn,outputn,6,{'logsig','tansig'});
net.trainParam.epochs=300;
net=train(net,inputn,outputn);
Xn_Test=mapminmax('apply',X_Test,inputps);
an=sim(net,Xn_Test);
y=mapminmax('reverse',an,outputps);
y1=y(1,:);
```

Elman 神经网络的训练和预测:

```
net=newelm(minmax(inputn),[6,1],{'purelin','purelin'});
net.trainParam.epochs=300;
net=train(net,inputn,outputn);
Xn_Test=mapminmax('apply',X_Test,inputps);
```

```
an=sim(net,Xn_Test);
y=mapminmax('reverse',an,outputps);
y1=y(1,:);
```

遗传算法优化:

关于这一部分,可参考《MATLAB 神经网络 43 个案例分析》(北京航空航天大学出版社),但需要根据本文研究的问题和 Elman 神经网络的结构做适当的修改,由于该部分的算法流程比较复杂,限于篇幅原因,本文不详细列出相关代码。

指标的计算:

```
n=length(Y_Test);
R2=(n*Y_Test*y1'-sum(Y_Test)*sum(y1))^2/...
    ((n*sum(Y_Test.^2)-sum(Y_Test)^2)*(n*sum(y1.^2)-sum(y1)^2));
MSE=sum((Y_Test-y1).^2)/n;
```

年化收益率的计算(这里给出 $X=0.5$ 的代码,另外的两个阈值与之类似):

```
tl=xlsread('C:\Users\Administrator\Desktop\毕设\szz.xlsx');
spread=tl(:,1);
bp=tl(:,2);
elman=tl(:,3);
ycelman=tl(:,4);
lwg=tl(:,5);
tks=tl(:,6);
jt=tl(:,7);
% 仓位 Pos=1 多头 1 手 Pos=-1 空头一手 假设期货是可分的
Pos=zeros(length(spread)-1,1);
% 初始资金
Initial=200000;
% 日收益记录
ReturnD=zeros(length(spread)-1,1);
Return=zeros(length(ReturnD),1);
Return1=zeros(length(ReturnD),1);
Return2=zeros(length(ReturnD),1);
delta=bp(1:1218,:)-spread(1:1218,:);
Pos(1)=0;
for t=2:length(spread)-1
    if delta(t)>0.5
        if Pos(t-1)==1
```

```
        Pos(t)=1;
        ReturnD(t-1)=(spread(t)-spread(t-1))*100;
        continue;
    end
    if Pos(t-1)==-1
        Pos(t)=1;

ReturnD(t-1)=lwg(t-1)*100-tks(t-1)*184.3651-jt(t-1)*103.5053-lwg(t)*100
*1.00004+tks(t)*184.3651*0.99998+jt(t)*103.5053*0.99998;
        continue;
    end
    if Pos(t-1)==0
        Pos(t)=1;

ReturnD(t-1)=-lwg(t)*100*0.00002-tks(t)*184.3651*0.00001-jt(t)*103.5053
*0.00001;
        continue;
    end
    if delta(t)<-0.5
        if Pos(t-1)==-1
            Pos(t)=-1;
            ReturnD(t-1)=(spread(t-1)-spread(t))*100;
            continue;
        end
        if Pos(t-1)==1
            Pos(t)=-1;

ReturnD(t-1)=lwg(t)*100*0.99996-tks(t)*184.3651*1.00002-jt(t)*103.5053*
1.00002-lwg(t-1)*100+tks(t-1)*184.3651+jt(t-1)*103.5053;
            continue;
        end
        if Pos(t-1)==0
            Pos(t)=-1;

ReturnD(t-1)=-lwg(t)*100*0.00002-tks(t)*184.3651*0.00001-jt(t)*103.5053
*0.00001;
            continue;
        end
    end
end
```

```
if delta(t)>-0.5&&delta(t)<0.5
    if Pos(t-1)==-1
        Pos(t)=0;

ReturnD(t-1)=lwg(t-1)*100-tks(t-1)*184.3651-jt(t-1)*103.5053-lwg(t)*100
*1.00002+tks(t)*184.3651*0.99999+jt(t)*103.5053*0.99999;
        continue;
    end
    if Pos(t-1)==1
        Pos(t)=0;

ReturnD(t-1)=lwg(t)*100*0.99998-tks(t)*184.3651*1.00001-jt(t)*103.5053*
1.00001-lwg(t-1)*100+tks(t-1)*184.3651+jt(t-1)*103.5053;
        continue;
    end
    if Pos(t-1)==0
        Pos(t)=0;
        ReturnD(t-1)=0;
        continue;
    end
end
end
for i=1:length(ReturnD)
    ReturnD(i)=ReturnD(i)-0.0435*24*800000/365/1218;
end
ReturnCum = cumsum(ReturnD);
for i=1:length(ReturnCum)
    Return(i)=ReturnCum(i)/Initial*365/24;
end
delta=elman(1:1218,:)-spread(1:1218,:);
ReturnD=zeros(length(spread)-1,1);
Return1=zeros(length(ReturnD),1);
Pos(1)=-1;
for t=2:length(spread)-1
    if delta(t)>0.5
        if Pos(t-1)==1
            Pos(t)=1;
            ReturnD(t-1)=(spread(t)-spread(t-1))*100;
            continue;
        end
```

```
if Pos(t-1)==-1
    Pos(t)=1;

ReturnD(t-1)=lwg(t-1)*100-tks(t-1)*184.3651-jt(t-1)*103.5053-lwg(t)*100
*1.00004+tks(t)*184.3651*0.99998+jt(t)*103.5053*0.99998;
    continue;
end
if Pos(t-1)==0
    Pos(t)=1;

ReturnD(t-1)=-lwg(t)*100*0.00002-tks(t)*184.3651*0.00001-jt(t)*103.5053
*0.00001;
    continue;
end
end
if delta(t)<-0.5
    if Pos(t-1)==-1
        Pos(t)=-1;
        ReturnD(t-1)=(spread(t-1)-spread(t))*100;
        continue;
    end
    if Pos(t-1)==1
        Pos(t)=-1;

ReturnD(t-1)=lwg(t)*100*0.99996-tks(t)*184.3651*1.00002-jt(t)*103.5053*
1.00002-lwg(t-1)*100+tks(t-1)*184.3651+jt(t-1)*103.5053;
        continue;
    end
    if Pos(t-1)==0
        Pos(t)=-1;

ReturnD(t-1)=-lwg(t)*100*0.00002-tks(t)*184.3651*0.00001-jt(t)*103.5053
*0.00001;
        continue;
    end
end
if delta(t)>-0.5&&delta(t)<0.5
    if Pos(t-1)==-1
        Pos(t)=0;
```

```
ReturnD(t-1)=lwg(t-1)*100-tks(t-1)*184.3651-jt(t-1)*103.5053-lwg(t)*100
*1.00002+tks(t)*184.3651*0.99999+jt(t)*103.5053*0.99999;
    continue;
end
if Pos(t-1)==1
    Pos(t)=0;

ReturnD(t-1)=lwg(t)*100*0.99998-tks(t)*184.3651*1.00001-jt(t)*103.5053*
1.00001-lwg(t-1)*100+tks(t-1)*184.3651+jt(t-1)*103.5053;
    continue;
end
if Pos(t-1)==0
    Pos(t)=0;
    ReturnD(t-1)=0;
    continue;
end
end
end
ReturnD(1)=ReturnD(1)-lwg(1)*0.00002-tks(1)*0.00001-jt(1)*0.00001;
ReturnD(1218)=lwg(1218)*100-tks(1218)*184.3651-jt(1218)*103.5053-lwg(12
19)*100*1.00002+tks(1219)*184.3651*0.99999+jt(1219)*103.5053*0.99999;
for i=1:length(ReturnD)
    ReturnD(i)=ReturnD(i)-0.0435*24*800000/365/1218;
end
ReturnCum = cumsum(ReturnD);
for i=1:length(ReturnCum)
    Return1(i)=ReturnCum(i)/Initial*365/24;
end
delta=ycelman(1:1218,:)-spread(1:1218,:);
ReturnD=zeros(length(spread)-1,1);
Return2=zeros(length(ReturnD),1);
Pos(1)=-1;
for t=2:length(spread)-1
    if delta(t)>0.5
        if Pos(t-1)==1
            Pos(t)=1;
            ReturnD(t-1)=(spread(t)-spread(t-1))*100;
            continue;
        end
        if Pos(t-1)==-1
```

```
Pos(t)=1;

ReturnD(t-1)=lwg(t-1)*100-tks(t-1)*184.3651-jt(t-1)*103.5053-lwg(t)*100
*1.00004+tks(t)*184.3651*0.99998+jt(t)*103.5053*0.99998;
    continue;
end
if Pos(t-1)==0
    Pos(t)=1;

ReturnD(t-1)=-lwg(t)*100*0.00002-tks(t)*184.3651*0.00001-jt(t)*103.5053
*0.00001;
    continue;
end
end
if delta(t)<-0.5
    if Pos(t-1)==-1
        Pos(t)=-1;
        ReturnD(t-1)=(spread(t-1)-spread(t))*100;
        continue;
    end
    if Pos(t-1)==1
        Pos(t)=-1;

ReturnD(t-1)=lwg(t)*100*0.99996-tks(t)*184.3651*1.00002-jt(t)*103.5053*
1.00002-lwg(t-1)*100+tks(t-1)*184.3651+jt(t-1)*103.5053;
    continue;
end
if Pos(t-1)==0
    Pos(t)=-1;

ReturnD(t-1)=-lwg(t)*100*0.00002-tks(t)*184.3651*0.00001-jt(t)*103.5053
*0.00001;
    continue;
end
end
if delta(t)>-0.5&&delta(t)<0.5
    if Pos(t-1)==-1
        Pos(t)=0;

ReturnD(t-1)=lwg(t-1)*100-tks(t-1)*184.3651-jt(t-1)*103.5053-lwg(t)*100
```

```
*1.00002+tks(t)*184.3651*0.99999+jt(t)*103.5053*0.99999;
    continue;
end
if Pos(t-1)==1
    Pos(t)=0;

ReturnD(t-1)=lwg(t)*100*0.99998-tks(t)*184.3651*1.00001-jt(t)*103.5053*
1.00001-lwg(t-1)*100+tks(t-1)*184.3651+jt(t-1)*103.5053;
    continue;
end
if Pos(t-1)==0
    Pos(t)=0;
    ReturnD(t-1)=0;
    continue;
end
end
end
ReturnD(1)=ReturnD(1)-lwg(1)*0.00002-tks(1)*0.00001-jt(1)*0.00001
ReturnD(1218)=lwg(1219)*100*0.99998-tks(1219)*184.3651*1.00001-jt(1219)
*103.5053*1.00001-lwg(1218)*100+tks(1218)*184.3651+jt(1218)*103.5053;
for i=1:length(ReturnD)
    ReturnD(i)=ReturnD(i)-0.0435*24*800000/365/1218;
end
ReturnCum = cumsum(ReturnD);
for i=1:length(ReturnCum)
    Return2(i)=ReturnCum(i)/Initial*365/24;
end
```