

电子科技大学
UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA

专业学位硕士学位论文

MASTER THESIS FOR PROFESSIONAL DEGREE



论文题目 商品期货跨品种套利策略设计
——以铁矿石产业链为例

专业学位类别 金融硕士

学 号 202022150409

作者姓名 敖艺璇

指导教师 李平 教授

学 院 经济与管理学院

分类号 _____ 密级 _____
UDC ^{注1} _____

学 位 论 文

商品期货跨品种套利策略设计 ——以铁矿石产业链为例

(题名和副题名)

敖艺璇

(作者姓名)

指导教师 李 平 教 授
电子科技大学 成 都
(姓名、职称、单位名称)

申请学位级别 硕士 专业学位类别 金融硕士
提交论文日期 2022 年 4 月 15 日 论文答辩日期 2022 年 5 月 20 日
学位授予单位和日期 电子科技大学 2022 年 6 月
答辩委员会主席 _____
评阅人 _____

注 1: 注明《国际十进分类法 UDC》的类号。

Commodity Futures Cross-variety Arbitrage Strategy Design: A Case Study of The Iron Ore Industry Chain

A Master Thesis Submitted to
University of Electronic Science and Technology of China

Discipline Master of Finance

Student ID 202022150409

Author Ao Yixuan

Supervisor Prof. Li Pin

School School of Management and Economics

摘 要

截止到 2021 年末,我国已有 64 个上市的商品期货品种,其数量不断地增加,使各期货品种的产业链也逐渐变得完善。期货市场过去以单一商品期货价格波动为主要特征,如今更多地体现出产业链整体波动的特点,这就为处于同一产业链上的商品期货提供了套利机会。在计算机技术快速发展的基础上,量化交易逐渐走进大众视野。同时商品期货市场拥有 T+0 交易机制、双向交易等特点,这为大众在期货市场进行量化投资提供了便利的条件,因此期货量化交易策略的构建也越来越受到企业和投资者的重视。

基于以上背景,本文选择从铁矿石产业链出发,寻找其产业链上存在的套利机会,在统计套利的基础上结合计算机技术构建期货跨品种套利策略。首先,本文对铁矿石产业链进行基本面分析,发现了铁矿石、螺纹钢、热轧卷板之间的基本面逻辑关系及套利机会,为后续设计跨品种套利策略奠定了基础。然后,本文基于协整理论对跨品种套利策略进行了实证分析,主要包括策略设计及回测两个阶段。

在策略设计部分,本文以 2019 年 1 月 1 日至 2021 年 12 月 31 日期间铁矿石、螺纹钢、热轧卷板期货活跃合约的日收盘价为样本数据做相关性分析及单位根检验,再将它们分别配对成三个组合进行协整检验,证明三个组合都具有协整性,可用以构建跨品种套利组合。在策略参数设计上,根据在移动时间窗口下计算的协整系数设计套利组合配比,并采用固定参数法将开平仓阈值和止损阈值分别设置为 ± 0.75 倍和 ± 2 倍的价差序列标准差。

在回测部分,本文以日频率对策略做了回测,再用遍历交易信号阈值的方式优化交易信号,发现三个组合分别在不同的交易信号阈值下获得了更好的策略绩效,其中铁矿石与螺纹钢组合在开平仓、止损阈值分别为 ± 0.75 倍、 ± 1.8 倍价差序列标准差时,策略收益率达到 47.11%,最大回撤为 13.78%;铁矿石与热轧卷板组合在开平仓、止损阈值分别为 ± 0.55 倍、 ± 1.8 倍价差序列标准差时,策略收益率达到 156.11%,最大回撤为 32.76%;螺纹钢与热轧卷板组合在开平仓、止损阈值分别为 ± 0.35 倍、 ± 1.5 倍价差序列标准差时,策略收益率达到 43.72%,最大回撤为 11.33%。回测结果表明,本文基于协整理论所设计的跨品种套利策略具有较好的适用性,但是对于不同品种组合还需对策略参数进行具体调整。

关键词: 商品期货, 跨品种套利, 协整与误差修正模型, 铁矿石

ABSTRACT

By the end of 2021, There have been 64 listed commodity futures varieties in China, and the number keeps increasing, which gradually improves the industrial chain of various futures varieties. In the past, the futures market was mainly characterized by the fluctuation of single commodity futures price, but now it is more characterized by the fluctuation of the whole industrial chain, which provides arbitrage opportunities for commodity futures in the same industrial chain. On the basis of the rapid development of computer technology, quantitative trading gradually entered the public view. At the same time, commodity futures market has the characteristics of T+0 trading mechanism and two-way trading, which provides convenient conditions for the public to make quantitative investment in the futures market. Therefore, the construction of quantitative futures trading strategy has been paid more and more attention by enterprises and investors.

Based on the above background, this thesis chooses to start from the iron ore industry chain, look for arbitrage opportunities in the industry chain, and build cross-variety futures arbitrage strategy based on statistical arbitrage combined with computer technology. This thesis firstly analyzes the fundamentals of the iron ore industry chain and discovers the fundamental logic relationship and arbitrage opportunities among iron ore, rebar and hot rolled coil, laying a foundation for the subsequent design of cross-variety arbitrage strategies. Then, the empirical analysis of cross-variety arbitrage strategy based on cointegration theory is made, including two stages of strategy design and backtesting.

In the part of strategy design, correlation analysis and unit root test are conducted on the daily closing prices of active futures contracts of iron ore, rebar and hot rolled coil from January 1, 2019 to December 31, 2021 as sample data, and then they are paired into three combinations for co-integration test to prove that all three combinations have co-integration. It can be used to construct cross-variety arbitrage portfolios. In the strategy parameter design, the arbitrage combination ratio is designed according to the cointegration coefficient calculated under the moving time window, and the fixed parameter method is adopted to set the open position threshold and stop loss threshold to ± 0.75 times and ± 2 times of the standard deviation of the spread series, respectively.

In the backtesting section, the strategy is backtested at a daily frequency, and then the trading signals are optimized by traversing the trading signal thresholds. Finally, each of the three portfolios achieved better strategy performance under different trading signal thresholds. The iron ore and rebar portfolio achieved a strategy return of 47.11% and a maximum retracement of 13.78% at an open and close threshold and a stop loss threshold of ± 0.75 times and ± 1.8 times the standard deviation of the spread series respectively. The iron ore and hot rolled coil portfolio achieved a strategy return of 156.11% and a maximum retracement of 32.76% at an open and close threshold and a stop loss threshold of ± 0.55 times and ± 1.8 times the standard deviation of the spread series respectively. The rebar and hot rolled coil portfolio achieved a strategy return of 43.72% and a maximum retracement of 11.33% at an open and close threshold and a stop loss threshold of ± 0.35 times and ± 1.5 times the standard deviation of the spread series respectively. The backtesting results show that the cross-variety arbitrage strategy designed in this paper based on cointegration theory has good applicability, but specific adjustments to the strategy parameters are needed for different variety combinations.

Keywords: Commodity Futures, Cross-variety Arbitrage, Cointegration and Error Correction Models, Iron Ore

目 录

第一章 绪论	1
1.1 研究背景及意义	1
1.1.1 研究背景	1
1.1.2 研究意义	2
1.2 国内外研究现状	3
1.2.1 量化投资研究现状	3
1.2.2 协整模型在统计套利中运用的研究现状	4
1.2.3 跨品种套利研究现状	6
1.2.4 文献评述	7
1.3 研究思路和内容安排	8
1.4 研究特色与创新	9
第二章 跨品种套利相关理论及方法	10
2.1 跨品种套利概述	10
2.2 跨品种套利相关方法及模型	10
2.2.1 基本面分析方法	11
2.2.2 统计套利方法	11
2.2.3 协整与误差修正模型	12
2.3 本章小结	13
第三章 铁矿石产业链分析	14
3.1 铁矿石基本信息	14
3.1.1 铁矿石简介	14
3.1.2 铁矿石产业链	14
3.1.3 铁矿石期货合约	15
3.2 铁矿石期货价格影响因素	16
3.2.1 供给	16
3.2.2 需求	18
3.2.3 库存	22
3.2.4 宏观因素	23
3.3 本章小结	23
第四章 跨品种套利策略设计	25

4.1 品种和数据选择	25
4.2 协整检验过程	26
4.2.1 相关性分析	26
4.2.2 单位根检验	27
4.2.3 协整关系检验	29
4.2.4 误差修正模型	30
4.3 策略设计	31
4.3.1 策略设计原理	31
4.3.2 套利组合的配比	32
4.3.3 交易信号的设计	33
4.3.4 具体策略设计	33
4.4 本章小结	35
第五章 策略回测分析	36
5.1 绩效评价指标	36
5.2 策略回测结果	37
5.3 交易信号阈值优化	39
5.4 本章小结	43
第六章 研究结论与展望	44
6.1 研究结论	44
6.2 研究展望	44
致 谢	46
参考文献	47
附 录	50

第一章 绪论

1.1 研究背景及意义

1.1.1 研究背景

我国期货市场最早起源于 1990 年的郑州粮食批发市场，发展至今，已经有了郑州商品期货交易所、上海期货交易所、大连商品期货交易所、中国金融期货交易所和上海能源交易所五家期货交易所。随着我国金融市场改革开放的不断推进，期货市场相关法律体系的逐步完善及其监管水平的不断提高，期货市场变得越来越规范化和法制化。我国期货市场运行机制的日趋健全，使之成为现代金融体系中不可缺少的一环，它的作用有套期保值、价格发现、投机套利等，其中期货套利主要有期现、跨期、跨市以及跨品种套利四种类型。

从我国期货市场成立到现在，市场上期货品种的交易量每年都在增加，其交易种类也从单一产品变得更加多样化，这都为进行期货跨品种套利创造了有利的条件。跨品种套利发生在相关性较高的期货品种之间，这种相关性通常是指各品种处于同一产业链上下游或是具有替代关系，这些品种具有相对稳定的价格关系，当价格关系发生变化时就存在套利的可能。这种套利行为能够提高期货市场的成交量与活跃程度，为套期保值操作提供帮助，也可以有效抑制期货市场上的过度投机行为，还可以帮助发现期货品种间的价格扭曲进而提供纠错的可能。另外对比其他套利行为，跨品种套利中存在的风险相对更加可控，其收益相对也更加稳定，因此它被广泛应用于国内外金融市场。

随着国内计算机技术的快速发展，量化交易也逐渐走进大众视野。量化投资是将数理模型与计算机程序相结合，用统计学方法组建金融资产组合，并从发现资产错误定价的过程中获得套利机会，再利用计算机程序设计买卖指令使其进行自动化交易，以期获取稳定投资回报的一种交易方式。与传统的投资方式相对比，量化投资具有更加客观、高效的特点，它只需要投资者在前期设计好交易机制，后期执行完全交给计算机，不需要随时关注资产价格变化。另外在量化交易下所设计的策略通常可以运用于多个标的资产或是多个市场，更具有全面性。相比国内股票市场，商品期货交易所拥有的 T+0 交易机制、双向交易等特点为投资者在期货市场中进行量化投资提供了更加便利的条件，因此期货量化交易策略的设计与实现也越来越受到企业和投资者的重视。

截止到 2021 年末,我国上市交易的商品期货品种已经达到 64 个,涉及农产品、金属、能源化工、黑色系等多个类别,商品期货品种数量的增加和规模的扩大,使各期货品种的产业链也逐渐变得完善。以往以单一商品期货价格波动为主要特征的期货市场,如今更多地表现出产业链整体波动的特点,这就为处于同一产业链上的商品期货增加了套利的可能性。在商品期货市场中,黑色系期货的热度一直都比较 高,是投资者重点关注的对象。以螺纹钢期货为例,自 2009 年最早推出后,它的市场热度就一直处于我国商品期货市场的前列,其成交金额、成交量也常年在我国甚至全球商品期货市场占据着极大的份额。钢铁产业是黑色系商品期货相对应的现货市场,是国家工业发展的基础,该行业的运营风险关系到国家经济能否平稳运行。因此,钢铁产业所涉及的商品期货的上市需求也日益增长。2013 年大连商品交易所推出铁矿石期货品种,这进一步为钢铁行业上的期货套利创造了有利条件。另外钢铁产业链上的套利具有操作灵活、风险相对可控、投资收益水平适中的特点,在期货市场上得到了众多企业以及投资者的青睐。

1.1.2 研究意义

在我国工业化和城市化进程中,钢铁产业涵盖了矿业开采、冶炼钢铁、轧制钢材等多个相关产业,其下游产品也被广泛应用于建筑业、制造业等多个领域,是一项影响我国国民经济的重要基础性产业。而我国正不断深入推进供给侧结构性改革,并出台了一系列与节能环保相关的政策,如“三去一降一补”、“碳中和、碳达峰”等,这使钢铁行业节能环保降耗的压力日益加大。另外我国钢铁行业集中度比较低,市场竞争加剧,从整体上看,该行业已经出现了市场价格波动幅度较大、利润较低的局面。从期货市场的价格波动中寻求钢铁产业链上的套利机会,有助于降低由钢材市场价格波动所带来的风险,从而加强我国钢铁产业对于风险的把控。这对促进我国钢铁产业的健康发展,保证我国经济的稳定运行有着重大的现实意义。

另外,我国铁矿石的消费量常年都位居世界前列,但是其供给高度依赖于国外进口,因此我国极度缺乏铁矿石的定价话语权。铁矿石价格的大幅变化会对整个钢铁行业产生影响,进而影响我国的国民经济。近几年来,铁矿石期货的价格波动幅度较大,对我国钢铁行业的相关企业产生了较大影响,同时也吸引了很多投资者的目光。开展铁矿石期货交易研究,可以帮助投资者更加全面地了解铁矿石期现货市场从而获取交易机会,还可以为钢铁企业套期保值提供辅助作用进而提高其风险管理意识,促进钢铁企业稳定发展。

基于此,本文以铁矿石产业链为切入点,探讨其产业链上各产品的基本面逻辑关系,从中选取铁矿石、热轧卷板、螺纹钢三大黑色系期货为研究对象,希望在协

整与误差修正模型下设计出一个有效的期货跨品种套利策略，这样一是能够丰富国内有关于钢铁产业套利方面的研究，让投资者了解跨品种套利的操作方法，并为其提供一个可选择的投资策略；二是将基本面分析方法和统计套利方法相结合来构建跨品种套利策略，可以为研究商品期货产业链上的套利提供一些思路；三是运用计算机程序化交易方法进行跨品种套利策略设计，能够让更多人了解量化交易的研究现状，从而参与到量化投资的行列中。

1.2 国内外研究现状

1.2.1 量化投资研究现状

量化投资起源于国外，它的发展建立在很多理论研究的基础上，最早是 Markowitz (1952)^[1]提出的均值一方差模型，在模型中他用投资组合的平均收益率和收益率的方差分别来衡量预期收益率和风险的大小，期望在一定风险下获取最大的投资回报。随后，Sharpe (1964)^[2]等人又在 Markowitz (1952)^[1]的基础上进行拓展，提出了资本资产定价模型，该模型主要是探讨资产的预期收益率与风险资产之间的均衡关系，也就是为了补偿一定风险，投资者应当获得的报酬水平。该模型是现在金融市场价格理论的支柱，在证券投资中得到广泛应用。由于证券市场会受到不同种类信息的影响，Fama (1970)^[3]基于此提出有效市场假说 (EMH)，该假说将市场的有效性分为强式有效、半强式有效以及弱式有效三个等级，这三类市场的信息传递效率也是不一样的，信息效率由强到弱递减，这也是实现基本面分析和统计套利的理论基础。而后 Ross (1976)^[4]在 Sharpe (1964)^[2]的基础上又提出套利定价模型，他认为套利行为是市场均衡价格的决定性影响因素之一，当市场没有达到均衡状态，如果不考虑交易成本，将会存在无风险的套利机会。

以上理论基本构成了现代金融学的主要理论基础，也是量化投资的重要理论依据，国外的量化投资在此基础上开始发展。国内的量化交易大多数是借鉴于国外的经典模型，学者丁鹏 (2012)^[5]所著的《量化投资》一书对量化投资有较为全面的介绍，他还对量化投资的类型和工具作了详细阐述，并总结了程序化交易下的八种量化交易策略，可应用于股票、期货、期权等市场，国内许多学者的相关研究都是以此为参考。

不论在国内外，量化交易在股票、期货市场都具有较强的适用性。但由于中国股市有 T+1 交易机制的限制，所以在股票市场，量化交易通常用于量化选股。相比之下，量化投资在期货市场的运用更加广泛，尤其是在 2010 年我国推出沪深 300 股指期货后，运用程序化交易来实现期货套利的投资者日益增加。黄适舜 (2011)^[6]将技术指标模型与计算机程序结合运用于期货投资，并对期货市场的三大产品

PTA 指数、郑棉指数、沪胶指数进行仿真交易，得到了一个收益稳定的程序化交易策略。李达捷（2013）^[7]基于机器学习构建了一个高频 CAT 策略，并在中国股指期货市场进行实证分析，策略获得了稳定的盈利。

利用计算机程序构建策略是量化投资在期货市场的主要应用，很多学者也开始引入多种数理模型到量化交易策略的设计中。陈艳，王宣承（2015）^[8]在对我国金融衍生品构建高频交易策略时，加入了对技术指标的选择并引入遗传网络规划方法，用黄金、铝和橡胶期货的高频交易数据进行回测以验证策略的效果，为量化投资领域提供了实践价值。李志鸿（2017）^[9]以用 Hurt 指数将分别基于海归法则和布林线的两个量化投资策略结合形成一个短期交易策略，并在实证结果中发现该策略可以获得一定的超额回报。李雁（2019）^[10]在 Alpha 模型和风险控制模型的基础上以 2011-2013 年的商品期货数据为样本构建商品期货量化交易策略，最终设计出高收益率和低风险的交易策略。周志中，俞祖卿（2021）^[11]在协整的基础上设计了一个商品信息融合的趋势跟踪策略，并在不同市场进行策略转换，实证结果表明信息融合及策略转换都可以提高策略绩效。

从以上研究中可以发现，国内的量化交易起步虽然比较晚，但是在期货市场上的相关研究还是较为丰富，其中大多数是将数理模型和程序化交易相结合构建量化交易策略，这为本文后续商品期货量化策略的设计提供了夯实的基础。

1.2.2 协整模型在统计套利中运用的研究现状

统计套利的思想最早也是源自于国外，它是指用数理模型构建资产组合，再设计相关交易信号进行套利，以期获取超额收益的过程。国外学者在进行统计套利方面的相关研究时，也是在逐步完善统计套利的相关理论。统计套利涉及很多数理模型和方法，因此很多学者在研究过程中会不可避免地将这些方法进行对比，以选择最适用的方法。Carol（2005）^[12]将跟踪误差最小化和协整理论分别运用于统计套利中，对比发现将两种方法应用于简单指数跟踪时的效果相差不大，但是当跟踪变得困难时，基于协整理论的统计套利会更加稳定，套利效果也更好。Huck、Afawubo（2015）^[13]以标准普尔 500 指数成分股的日收盘价数据为研究对象，探讨基于各种配对方法所构建的交易策略的绩效，实证结果表明在控制交易成本和风险后，以协整方法构建的交易策略的盈利能力比距离法更强。Hossein Rad（2016）^[14]采用距离法、协整法、Copula 法分别设计三个不同的配对交易策略，并且利用 1962-2014 年美国股票市场的数据进行实证分析，最终发现三个策略在显著波动期间表现都很好，但是基于协整模型的策略在动荡的市场条件下有更好的表现。

从以上研究可以看出基于协整理论所构建的交易策略更加稳健,并且在这方面的研究也比较成熟,其中由 Vidyamurthy (2004)^[15]提出的以协整理论为基础构建配对交易策略的思想最为全面,他对构建该交易策略的步骤作了详细的阐述,包括套利组合的构建、协整关系的检验以及策略参数的设计等。这为协整理论在套利统计中的运用提供了主要参考,国内也有不少学者对此进行了相关研究,主要是在股票配对交易和期货跨期、跨品种套利等方面的应用。

在股票配对交易方面,王粹萃(2007)^[16]以协整理论为基础,从美国道琼斯工业平均指数的 30 家成份股中选取股票以构建不同的投资组合,并设计统计套利策略,各投资组合在实证阶段都获得了较高的夏普比率,结果表明策略是有效的。于玮婷(2011)^[17]将协整方法用于我国融资融券标的的股票配对中,所设计的策略在时间样本内外都获得了不错的收益。陈叠、王秋香(2016)^[18]等则是改进了传统的协整模型,采用滑动时间窗口更新数据,并且策略的交易信号阈值和仓位配比都随时间更新,实证结果表明新的策略可以获取更多的套利机会。吴员福(2018)^[19]也在协整理论的基础上以延后开仓、提前平仓以及引入 GARCH 模型三个措施来改进模型,并在 A 股市场上构建配对交易策略,实证结果表明传统策略稳定有效,而延后开仓以及引入 GARCH 模型的改进措施能够提高策略收益。

在跨期套利方面,仇中群、程希骏(2008)^[20]基于协整理论以我国沪深 300 股指期货为研究对象构建跨期套利策略,在进行仿真交易时有所获利。常宗琪(2008)^[21]以白糖期货合约作为研究对象,基于协整方法构建白糖期货合约 SR901 与 SR905 的跨期套利策略,在实证阶段策略获得了比较稳定的收益。何树红、张月秋(2013)^[22]等则在协整模型的基础上引入 GARCH 模型描述残差的异方差,并计算实际的价格差和方差,从而确定交易信号阈值,再以沪深 300 股指期货为研究对象进行策略效果检验,结果表明该模型有较强的灵活性,能够有效发现套利机会。张波(2017)^[23]在协整模型的基础上引入修正协整关系,以 EGARCH-M 模型构建股指期货跨期套利策略,在实证检验过程中发现 EGARCH-M 模型拟合效果比传统模型更好,并且拥有更好的套利效果。

在跨品种套利方面,丁秀玲和华仁海(2007)^[24]在验证大豆与豆粕期货价格间的协整关系后,分别构建多头套利与空头套利两种策略,并进行模拟交易检验策略的效果。扈文秀、牛静等(2013)^[25]利用协整方法对我国商品指数期货和大宗商品股票 ETF 进行研究,从中设计跨市场、跨品种的交易策略,两个策略在实证阶段都获取了较为稳定的收益,表明双跨套利的可行性。周亮(2017)^[26]依据协整理论在我国铁矿石、螺纹钢、焦炭期货中选取套利组合并设计套利交易策略,同时通过阈值测算获取最佳交易信号,策略在样本内外均获得了不错的收益。方军、李星野

(2019)^[27]在以传统协整理论构建统计套利模型的基础上引入离散 Fourier 变换的思想,构建了新的 DFT-GARCH 统计套利模型,并在同一交易规则下以铜期货为研究对象进行实证分析,结果表明在样本内传统模型更能捕捉套利机会,而在样本外新的模型对于价差波动的预测更好,策略收益也更好。

从上述研究中可以发现国内外有关于协整理论在统计套利中的应用已经比较成熟,无论是用以构建股票配对交易策略还是期货跨品种、跨期套利策略,都有较好的普适性,同时也可以从协整理论的基础上引入其他模型来提高策略绩效。

1.2.3 跨品种套利研究现状

跨品种套利一直以来都是国内外期货市场套利的研究热点, Liu(2005)^[28]基于生猪、玉米、豆粕期货价格之间的协整关系构建跨品种套利策略,策略在模拟交易过程中获得了稳定的收益。Dunis (2006)^[29]等分别使用协整、反向移动平均线、传统回归技术、神经网络回归模型对 WTI 原油期货和 BRENT 原油期货构建价差套利模型,并设计相关交易策略进行收益模拟分析,实证结果表明四个策略表现都较好,其中以反向平均线和传统回归技术所构建的策略表现更优。Mitchell (2010)^[30]基于协整和均值回复理论对大豆、豆粕、豆油期货进行实证研究,发现三者存在长期均衡关系并构建跨品种交易策略,再根据三者的价差变化调整开仓、平仓阈值,进而提高策略的盈利性。Kanamura (2010)^[31]等基于基本面的观点,对取暖油和天然气进行跨品种套利的研究,发现在构建跨品种套利策略时不仅需要考虑期货价格关系,还应当关注均值回复、季节性特征、波动率等影响因素。

随着国内期货市场的兴起以及期货品种数量的不断增加,国内有关于跨品种套利的研究也越来越广泛,主要包括相关产品套利以及产业链上下游产品套利两种类型。

相关产品套利主要是指具有可替代关系或是互补关系的期货品种之间产生的跨品种套利。张伟、景鸿息(2013)^[32]以传统协整模型构建黄金与白银期货间的跨品种套利策略,并以两者的 5 分钟收盘价数据为样本,在选定时间内分别做了样本内外回测,在样本内策略收益率为 135%,最大回撤仅为 21%,策略绩效很好,但在样本外策略表现不佳,后续可以通过改变交易信号阈值进行改进。黄腾(2019)^[33]研究的是豆油、菜籽油、棕榈油期货三者之间的套利,他以豆油期货与其他两个期货之间的价差为基础构建相关开仓、平仓信号,并以三者的日收盘价数据为样本进行实证分析,最终策略收益率为 20.40%,但是最大回撤有 48.123%,表明还需要进一步改进止损信号。张静文(2020)^[34]从焦炭和铁矿石的用途出发,结合协整

理论发现两者之间存在套利机会，以两者的日收盘价数据为样本构建价差套利模型，通过收益模拟分析发现两者之间能够进行跨品种套利。

产业链上下游产品套利一般是指在同一产业链上的原料、半成品和成品之间构成的跨品种套利。梁仁方(2016)^[35]在协整的基础上用 NAR 动态神经网络模型和 BP 神经网络模型对螺纹钢、铁矿石、焦炭期货设计跨品种套利策略，在模拟交易中发现以 NAR 动态神经网络模型构建的策略效果更好。孙建明,张伟楠(2017)^[36]等在分析玉米与淀粉之间的协整关系后，分别构建以价差、价比和技术指标为基础的套利策略，并以两者的日收盘价数据为样本进行实证分析，结果表明三种策略都可行且各有特色。刘一诺(2020)^[37]在验证焦煤、焦炭期货日收盘价的协整关系后，构建了以固定参数阈值为开仓、平仓、止损信号的跨品种套利策略，最终策略获得了 8.1% 的收益率。黄巍华、王伟(2021)^[38]对大豆产业链下的大豆、豆油、豆粕做了相关性分析和协整检验，确定三者之间的长期均衡关系，再将其两两组合设计跨品种套利策略，以三者的日收盘价数据为样本进行实证分析，结果表明该套利操作能够获得稳定收益。

从上述研究中可以发现国内外有关于跨品种套利的研究比较多，通常是运用数理模型探讨期货品种间的关联及其价格间所拥有的规律，再依据两者价格关系的变化进行套利。其中在数理模型的选择上，协整模型被运用地最为广泛。

1.2.4 文献评述

通过对国内外有关于量化投资、协整理论在统计套利中的运用、期货跨品种套利的文献梳理，可以发现国内外在期货跨品种套利方面已经拥有了一个比较全面的研究框架，这为本文接下来的研究提供了理论依据。基于以上研究，笔者主要发现了以下几点：

1. 量化投资在我国期货市场的运用较为广泛，主要表现为将数理模型与程序化交易相结合构建交易策略，而本文就是在这个基础上设计交易策略。

2. 相比统计套利所涉及的其他方法，协整模型在统计套利中的应用得到了普遍的认可，它在股票配对交易和期货跨期、跨品种套利中的运用也比较成熟，其中大多数研究都是基于传统协整模型，或者在传统模型基础上进行改进来设计相关统计套利策略。

3. 国内期货市场为期货跨品种套利提供了有利的条件，但从目前的研究来看更多地是运用定量的方法分析两个品种之间的关系，然后基于数理模型构建跨品种套利策略。

4.跨品种套利策略在数理模型上通常也是选择协整模型，一般都是选择固定时间内的样本做协整检验，利用分钟或者日收盘价数据进行收益模拟检验。

根据以上几点，本文将基于国内外众多相关的研究，结合使用定性与定量的方法，探究铁矿石产业链上的套利机会，在传统协整模型的基础上使用滚动时间窗口设计跨品种交易策略，并根据策略回测效果进行拓展，用遍历交易信号阈值的方式确定不同组合的最佳交易阈值以提高策略绩效。

1.3 研究思路和内容安排

本文根据理论基础和现实意义，分析铁矿石产业链上的相关商品期货之间是否存在套利机会，从而选取相关品种构建跨品种套利模型，再设计套利策略及相关参数，通过检验交易策略的效果进行策略拓展。

本文首先说明了本选题的研究背景、意义和现状，然后介绍了跨品种套利的相关理论及方法，再结合使用基本面分析方法和统计方法，验证铁矿石、螺纹钢、热轧卷板期货的相关性以及三者跨品种套利的可行性，然后基于协整模型设计黑色系期货跨品种套利策略，最后根据回测结果分析策略的有效性。

本文分为六章，具体安排如下：

第一章为绪论。这部分介绍了本文的研究背景及意义、国内外研究情况、本文的研究思路和内容安排、文章的创新点。

第二章为跨品种套利相关理论和方法。这部分介绍了商品期货跨品种套利的基本原理，以及基本面分析方法和统计套利方法中的协整与误差修正模型在跨品种套利中的运用。

第三章为铁矿石产业链分析。这部分首先介绍铁矿石产业链的情况，再从基本面角度分析铁矿石期货价格的影响因素，从基本面逻辑关系中发现铁矿石、螺纹钢、热轧卷板的紧密关联，进而探讨三个期货品种之间进行跨品种套利的可行性。

第四章为跨品种套利策略的设计。在上一章产业链分析的基础上运用数理方法对铁矿石、螺纹钢、热轧卷板期货收益率做相关性检验，再采用 ADF 检验、协整检验等方法分别检验铁矿石与螺纹钢组合、铁矿石与热轧卷板组合、热轧卷板与螺纹钢组合的协整关系，从中选择可以进行套利的组合。最后对策略的设计以及参数做了详细阐述，其中策略参数主要包括套利组合配比以及开平仓、止损阈值。

第五章为策略回测结果的分析。利用历史数据在聚宽平台上对策略进行回测，通过回测结果分析策略的有效性，并根据所存在的问题对策略进行拓展。

第六章为研究结论与展望。总结本文的研究工作，并对本文的不足之处提出改进意见，以确定后续的研究方向。

1.4 研究特色与创新

本文的创新点主要在以下两个方面：

一是在于研究对象，目前我国对铁矿石期货品种的研究相对较少，其中有关于铁矿石产业链基本面分析的研究更少。另外在期货跨品种套利方面，以往的研究更多地侧重在金融期货或是农产品期货之间，而在黑色系商品期货间的跨品种套利也大多数是只以某几个品种为研究对象。期货市场上对于以产业链为基础，将基本面分析方法与统计方法结合应用于跨品种套利的研究比较少。

二是在跨品种交易策略的设计上，本文与利用传统协整模型的价差套利方式不同。传统方法是在整个样本时间内计算出一个固定的回归方程，根据回归所得的固定参数值与价差值进行套利。而本文是采用移动时间窗口，在不同时间段分别计算出回归方程，使回归参数与价差能够及时得到调整，从而在一个变化的价差轨道中进行套利，提高套利的准确率。

第二章 跨品种套利相关理论及方法

2.1 跨品种套利概述

期货统计套利是把统计套利的理论和实践方法结合起来,根据期货市场的具体行情变化来设计相关的统计套利交易策略,目前,它已经被越来越多的投资者运用在期货市场中。期货套利不仅在市场上存在大量的机会,而且在经济意义上拥有基本面的支撑,对推动期货市场的价格发现功能有着举足轻重的作用,其中跨品种套利是期货套利中常见的一种类型。

期货跨品种套利是指两个或者多个在经济学或统计学意义上具有相关关系的期货品种之间的套利交易,当两个期货合约的价格关系发生变化,可以在其价差从非合理区间回归至均衡状态时获取利润,其具体操作为:当价差发生偏离时,做多被低估的期货,同时做空被高估的期货;当价差回归至合理水平时,反向平仓从中获利。期货跨品种套利是以多个品种之间的相关性为基础,这种相关性在经济学上通常分为两种,进而形成以下两种套利类型:

一是以产业链上下游的生产关系为基础的跨品种套利,比如大豆及豆制品产业链、黑色系期货产业链、原油产业链、能源化工产业链等。尽管处于同一产业链上的期货品种不一样,但是其市场供给需求状况却存在着同一性。当上游产品涨价时会推动下游产品价格的上升,因此上下游产品的价差应当保持在一个比较均衡的水平。当某个商品基本面发生变化影响其价格时,将会影响其期货所在产业链上其他商品期货的价格,导致产业链上商品期货之间的价差改变从而偏离合理水平。这时可以从期货品种的基本面逻辑关系中分析价差强弱,或是运用数理方法寻找价差变化的规律,从而进行跨品种套利操作。

二是具有替代关系的跨品种套利,比如豆粕与玉米、铜与铝、菜油与棕榈油等。品种之间存在替代关系是指两者能够实现相同的作用,从而在消费领域会出现重合,导致价格会出现此消彼长的现象。当某种商品的市价太高时,下游厂商会采用相对便宜的替代品,进而降低其购买量。如果这两种商品的供给情况不发生变化,那么高价的产品将会降价,其替代产品的价格会上涨,从而两种商品的价差又会回到原来的水平。从长期来看,这个价差在变动过程中还是会维持在一个比较稳定的水平,所以当价差发生偏离时,会产生套利的机会。

2.2 跨品种套利相关方法及模型

2.2.1 基本面分析方法

本文是以商品期货为研究对象，因此可以回归到商品期货的本质，从价格的影响因素入手，分析价格变动的一般规律，为进行基本面分析和套利奠定基础。大宗商品期货的基本面分析主要是探究基本面因素对于价格的影响，它建立在商品供需理论的基础上，也就是分析商品供需关系变化对于商品期货价格的影响，然后在这个过程中对商品期货未来的价格走势进行合理地预测。除了供需状况变化对价格变动有直接影响外，还会有一些其它的因素对期货价格产生一定的影响。总体而言，在探究大宗商品期货的价格影响因素时，主要需要考虑供给和需求状况、库存水平、宏观政策等基本面因素。

商品期货的基本面套利是在对商品进行基本面分析后，根据期货的基本面趋势变动进行套利的买卖，因此它也被称作基本面趋势套利。跨品种套利中的基本面套利是指相关期货品种未来基本面的强弱关系发生改变时，会导致价差也产生波动，从而存在套利的可能性。其具体思路是通过分析相关商品的基本面因素以及所在行业的发展状况，考虑市场行情的变化，根据相关期货品种的未来价格走势以及价差变化趋势来设计跨品种套利策略。

一般来说，在用基本面分析方法对商品期货的价格进行趋势跟踪时，会发现不论强势还是弱勢的期货合约都会将其基本面强弱延续一段时间，所以在构建跨品种套利策略时，通常是持有强势期货合约的多头以及弱勢期货合约的空头。但这种方法要求投资者拥有较强的市场洞察能力，以及较好的应变期货价格趋势变化的能力。如果投资者不能及时追踪趋势变化，可能会使套利策略中的多头和空头都产生损失。而且投资者如果只用基本面分析方法进行套利，将很难把控择时的效果。因此在进行跨品种套利时，通常是用基本面分析方法探讨相关品种间的基本面逻辑关系及其在经济学上的相关性，再结合统计学方法对套利的可行性做实证分析，进而设计相关跨品种套利策略。

2.2.2 统计套利方法

统计套利是以均值回复理论为基础的套利方法，在面临多个金融资产时，通常用资产组合的价差或者价比来描述其长期均衡回复性，当资产价格在短期内产生波动时，价差或者价比所拥有的均值回复性能够带来套利机会。统计套利的具体操作就是以金融资产的历史价格数据为样本，使用数理统计学方法处理数据以构建金融资产投资组合，再结合统计分析所得的参数设计投资组合的多头和空头并进

行套利交易,从而获得稳定低风险的超额回报。这种套利通常不会受到市场趋势变化的影响,它主要是取决于资产组合之间的长期稳定关系以及短期内的价格波动。

在构建统计套利组合的过程中,涉及到的方法范围非常广,主要有距离法、协整法和时间序列分析方法等。在上一章的文献梳理中,笔者发现协整检验是最常用于统计套利的方法,主要用以衡量组合之间的长期均衡关系,从而选取跨品种套利组合,所以在构建跨品种套利策略时,本文将采用协整与误差修正模型。

2.2.3 协整与误差修正模型

协整与误差修正模型是一种常用的时间序列计量经济学模型,其中时间序列分析是在随机过程理论基础上,利用数理统计方法分析历史数据,从中发现样本数据之间存在的内在关系,根据这种关系推断数据整体所遵守的统计规律,再依据所归纳出的规律采用数量统计学的方法来预测未来发展。协整与误差修正模型主要包括协整理论与误差修正模型两个方面。

1. 协整理论

协整理论是由 Engle 和 Granger (1987)^[39]提出,他们表示某些不平稳的时间序列的线性组合可能具有平稳性。具体而言,当两个非平稳的时间序列经过 N 阶差分后能够得到两个平稳的序列,此时若有一个线性组合能将两个原始时间序列构成一个平稳的序列,则称两者之间存在协整关系。从以上可以发现,协整关系存在于两个不平稳的时间序列间,所以在进行协整关系检验之前还必须先检验原始数据的平稳性,再做 OLS 回归估计。这其中的理由为如果原始数据是平稳的时间序列,在进行 OLS 回归后将会产生一个伪回归模型,该模型的各项指标如 P 值、 T 统计量、 F 统计量都能通过检验,但这可能只是数据上的巧合,两者之间并没有经济意义上的关系。

在检验方法的选择上,时间序列平稳性检验最常用的是 ADF 单位根检验方法。如果在单位根检验后发现数据是非平稳的序列,则还需要对数据做差分处理再进行单位根检验,直到所处理的数据变成一个平稳的时间序列。检验数据之间的协整关系也有很多方法,其中最常使用的是 E-G 两步法,具体操作如下:

1) 完成平稳性检验后,对两个非平稳的原始时间序列 $\{X_t\}$ 、 $\{Y_t\}$ 做 OLS 回归估计,得到回归方程:

$$Y_t = a_0 + a_1 X_t + \varepsilon_t \quad (2-1)$$

根据式 (2-1) 可以计算出残差序列 μ_t :

$$\mu_t = Y_t - a_0 - a_1 X_t \quad (2-2)$$

2) 利用 ADF 检验来检验残差序列 μ_t 的平稳性, 如果检验结果发现该残差序列是平稳的, 则可以确定时间序列 $\{X_t\}$ 、 $\{Y_t\}$ 之间存在协整关系。

2. 误差修正模型

根据 Engle 与 Granger (1987)^[39] 所提出的格兰杰定理, 如果原始时间序列 $\{X_t\}$ 、 $\{Y_t\}$ 之间是协整的, 基于两者之间的长期均衡关系, 可以构建误差修正模型。误差修正模型是用于描述 $\{X_t\}$ 、 $\{Y_t\}$ 在短期内偏离均衡状态的程度, 也用于解释残差序列 μ_t 的自动修正作用。根据艾云辉、郭文娟 (2009)^[40] 对误差修正模型的研究, 一阶误差修正模型的推导过程具体如下:

假设时间序列 $\{X_t\}$ 、 $\{Y_t\}$ 之间存在如式 (2-1) 的均衡关系, 并且两者的 (1,1) 阶分布滞后形式如下:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t-1} + \mu Y_{t-1} + \varepsilon_t \quad (2-3)$$

将式 (2-3) 差分处理得到以下公式:

$$\Delta Y_t = \beta_0 + \beta_1 \Delta X_t + (\beta_1 + \beta_2) X_{t-1} - (1 - \mu) Y_{t-1} + \varepsilon_t \quad (2-4)$$

再将式 (2-4) 进行变换得到以下公式:

$$\Delta Y_t = \beta_1 \Delta X_t - \lambda (Y_{t-1} - \alpha_0 - \alpha_1 X_{t-1}) + \varepsilon_t \quad (2-5)$$

其中, $\lambda = 1 - \mu$; $\alpha_0 = \beta_0 / (1 - \mu)$; $\alpha_1 = (\beta_1 + \beta_2) / (1 - \mu)$ 。

式 (2-5) 就是一阶误差修正模型。

令 $ecm_{t-1} = Y_{t-1} - \alpha_0 - \alpha_1 X_{t-1}$, 式 (2-5) 可以简化为:

$$\Delta Y_t = \beta_1 \Delta X_t - \lambda ecm_{t-1} + \varepsilon_t \quad (2-6)$$

其中, ecm_{t-1} 表示的是 t-1 期的非均衡偏差项; λ 代表的是短期调整系数, 可以用来反映对短期偏离状态的修正程度。

2.3 本章小结

本章对跨品种套利所涉及的理论以及方法作了详细的阐述, 首先介绍了跨品种套利的原理, 包括跨品种套利的基本操作以及市场上主流的两种跨品种套利类型。其次介绍了跨品种套利中常用的两种方法, 包括基本面分析方法和统计套利方法, 随后本章也详细介绍了在统计套利中用以选择跨品种套利组合常用的计量经济学模型——协整与误差修正模型。另外考虑到基本面分析方法的局限性, 本文将采用基本面分析方法和统计套利方法相结合的方式构建跨品种套利组合以设计相关交易策略。

第三章 铁矿石产业链分析

3.1 铁矿石基本信息

3.1.1 铁矿石简介

铁矿石是一种含铁的矿石，具有很高的使用价值，它可以通过一系列的工艺过程选出铁，是冶炼钢铁必不可少的原材料。目前市场上已经出现了多种品位的铁矿石，品位是铁矿石质量的一个重要衡量指标，它代表铁矿石中铁元素的质量分数，而在当前的技术水平下，真正具有工业应用价值的也就是磁铁矿、赤铁矿、褐铁矿、菱铁矿。

全球铁矿可以分为主流矿与非主流矿，主流矿通常是指钢厂使用最多的 MNP（麦克粉、纽曼粉、PB 粉），其主要产地是国际四大矿山（澳大利亚必和必拓、力拓、福蒂斯丘以及巴西淡水河谷），而非主流矿就是除了主流矿以外的其他铁矿。在中国，铁矿石品位普遍较低，矿区资源一般是按区域划分，主要有 7 个大型矿区，其具体分布如表 3-1 所示。

表 3-1 中国主要矿区分布

矿区名称	所属区域	主要构成	含铁量
鞍山矿区	东北	磁铁矿、赤铁矿	平均 30%
迁滦矿区	华北	磁铁矿	/
邯邢矿区	华北	磁铁矿、赤铁矿	40-55%
大冶矿区	华中	磁铁矿、赤铁矿、黄铜矿、黄铁矿（铁铜共生矿）	40-50%
芜宁矿区	华东	磁铁矿、赤铁矿	富矿部分：50-60%
攀枝花矿区	西南	钒钛磁铁矿、伴生矿	27.8%
石碌矿区	海南	赤铁矿	平均 49.15%

3.1.2 铁矿石产业链

从广义上来说，铁矿石产业链除了涉及铁矿石挖掘采买以及冶炼加工的行业外，还应当包含其制造成钢材后的应用行业，其产业链具体如图 3-1 所示。

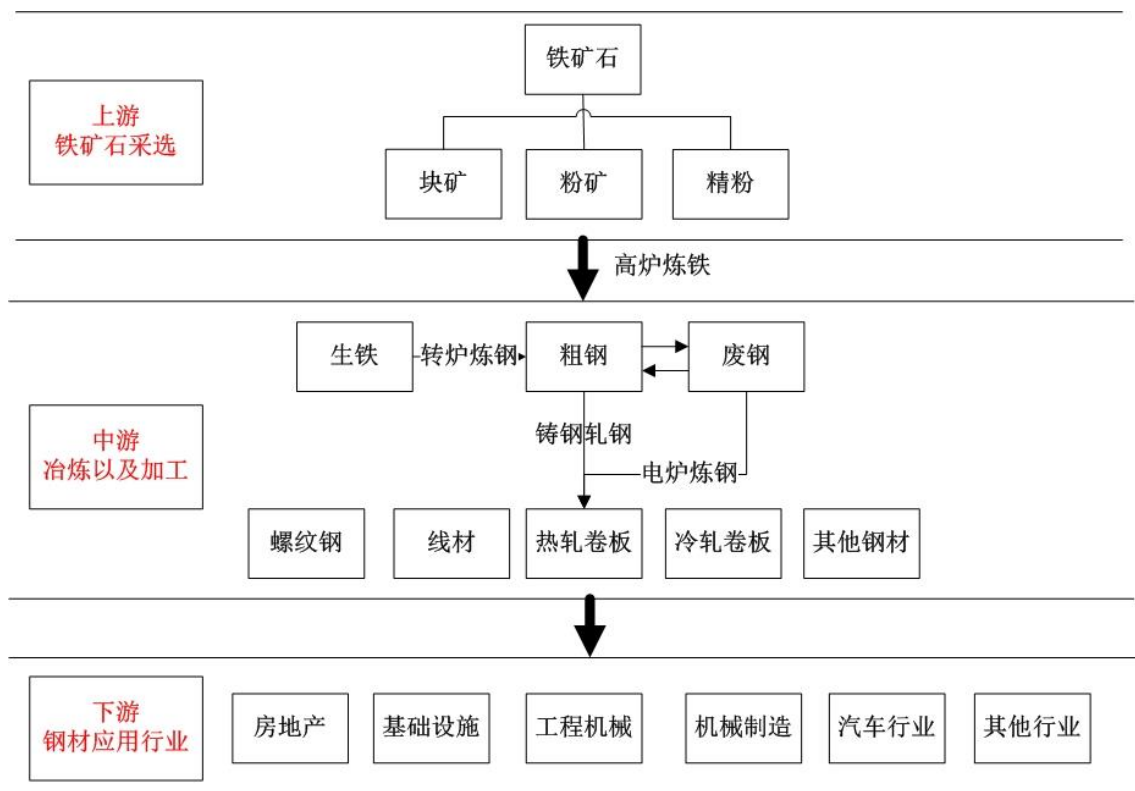


图 3-1 铁矿石产业链

从图 3-1 可以看到，作为钢铁冶炼的原材料，铁矿石处于其产业链的上游位置。

在铁矿石产业链的中游，主要是铁矿石的冶炼及加工，在经过一系列的炼铁炼钢过程后，最终生产出钢材产品，即铁矿石产业链的中游产物。钢材主要有型材、板材、管材、金属制品四种类型，细分的话则主要是螺纹钢、热轧卷板、线材等品种。

铁矿石产业链的下游是钢材产品的应用领域，主要是建筑业和制造业。其中建筑业主要是房地产行业、基建行业，制造业则主要集中在汽车、机械等行业。

3.1.3 铁矿石期货合约

中国大连商品交易所于 2013 年正式挂牌铁矿石期货品种，其具体合约形式如表 3-2 所示。大连商品交易所目前拥有 8 家交割仓库，其中天津 1 家，连云港 1 家，唐山 3 家，日照市 2 家，青岛 1 家，协议库存总和为 740 万吨。

表 3-2 大连商品交易所铁矿石期货合约

交易品种	铁矿石
交易单位	100 吨/手
报价单位	元（人民币）/吨
最小变动单位	0.5 元/吨
涨跌停板幅度	上一交易日结算价的 4%
合约月份	1、2、3、4、5、6、7、8、9、10、11、12 月
交易时间	每周一至周五上午 9:00~11:30，下午 13:30~15:00，以及交易所规定的其他时间
最后交易日	合约月份第 10 个交易日
最后交割日	最后交易日后第 3 个交易日
交割等级	大连商品交易所铁矿石交割质量标准（F/DCE I001-2019），可交割品牌和品牌升贴水由大连商品交易所另行规定。
交割地点	大连商品交易所铁矿石指定交割仓库及指定交割地点
最低保证金	合约价值的 5%
交割方式	实物交割
上市代码	I
上市交易所	大连商品交易所

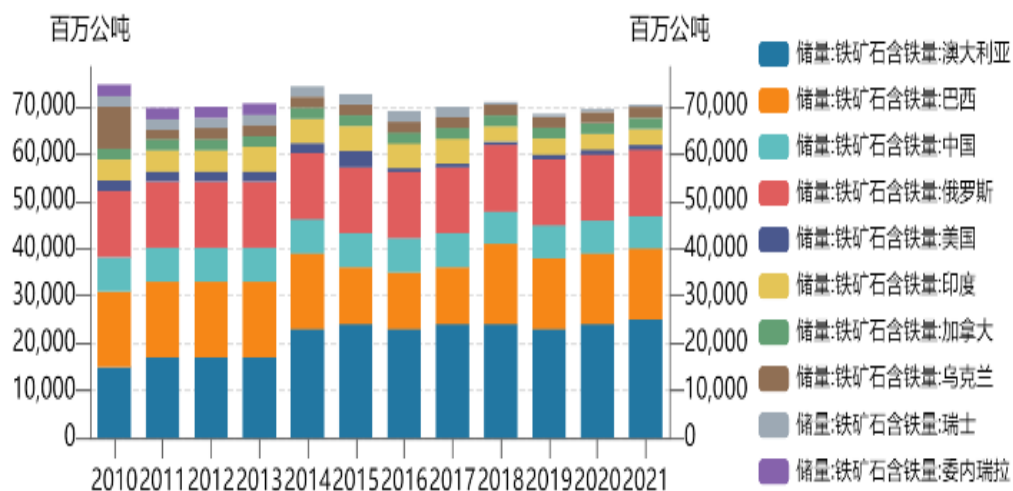
3.2 铁矿石期货价格影响因素

3.2.1 供给

与其他商品期货相似，铁矿石的供求关系对价格的影响最为直接。从产业链的角度来看，铁矿石作为上游产品，主要供给主要是受矿山供应量的影响。

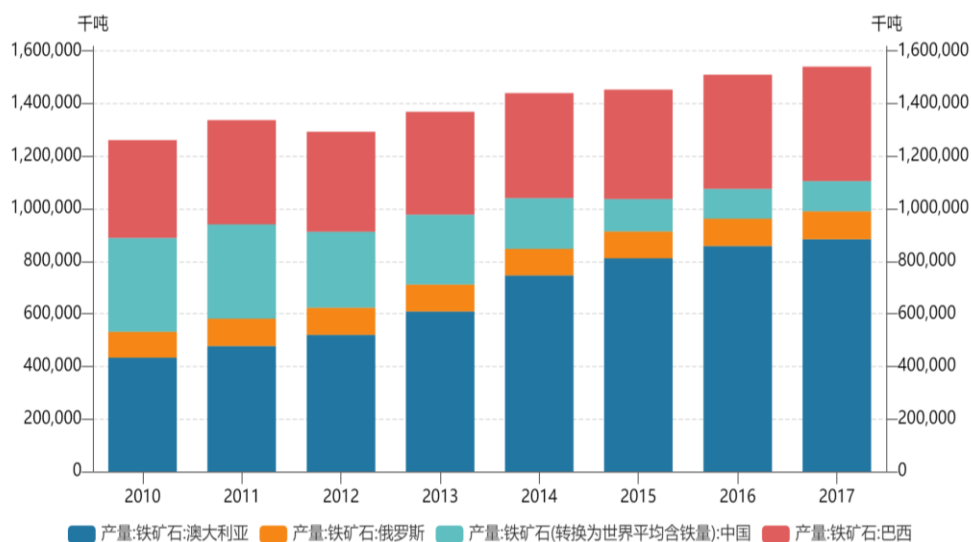
铁矿石是不可再生资源，它主要的供应首先要基于各个国家铁矿石的储量的大小，从图 3-2 可以看到近十年来全球铁矿石的储量也一直比较稳定，主要分布在澳大利亚、巴西、中国和俄罗斯，形成了铁矿石储量“四家比肩”的局面。

每个国家铁矿石储量的大小是决定其铁矿石产量大小的必要条件，但同时铁矿石的产量也和品位高低有关系，铁矿石平均含铁量高低会影响其需求，从而也影响其产出的多少，就上述铁矿储量最多的四个国家而言，其产量分布如图 3-3 所示。其中巴西和澳大利亚主要依靠国际“四大矿山”的铁矿石供给，其铁矿石的品位属于高位档次，同时这两个国家也是全球铁矿石的主要供应方；俄罗斯自从 1997 年整体平均含铁量提高之后就一直保持在 55%的水平上下，但是由于该国家政策的限制，其铁矿石产量在一定程度上只能满足本国的需求；中国的铁矿石平均含铁量则一直保持在较低的水平，近几年都维持在 35%的水平左右，另外由于开采铁矿石需要较高的成本，国内铁矿石的供给并不能满足本国的需求。



数据来源：wind 数据库

图 3-2 全球铁矿石储量（2010-2021 年）



数据来源：wind 数据库

图 3-3 巴西、澳大利亚、中国、俄罗斯铁矿石产量（2010-2017 年）

由于我国铁矿石受开采成本及资源的限制，国内矿山的产量具有很高的刚性。从图 3-4 中可以看到，在近几年在各种环保政策不断推出的大背景下，我国铁矿石原矿产量从高位回落，产量增速不断收窄至 0，如果不出现较大变故，国产铁矿石原矿产量将不会出现较大突破，还是比较依赖于进口铁矿石。



数据来源：wind 数据库

图 3-4 我国铁矿石产量统计（2010-2021 年）

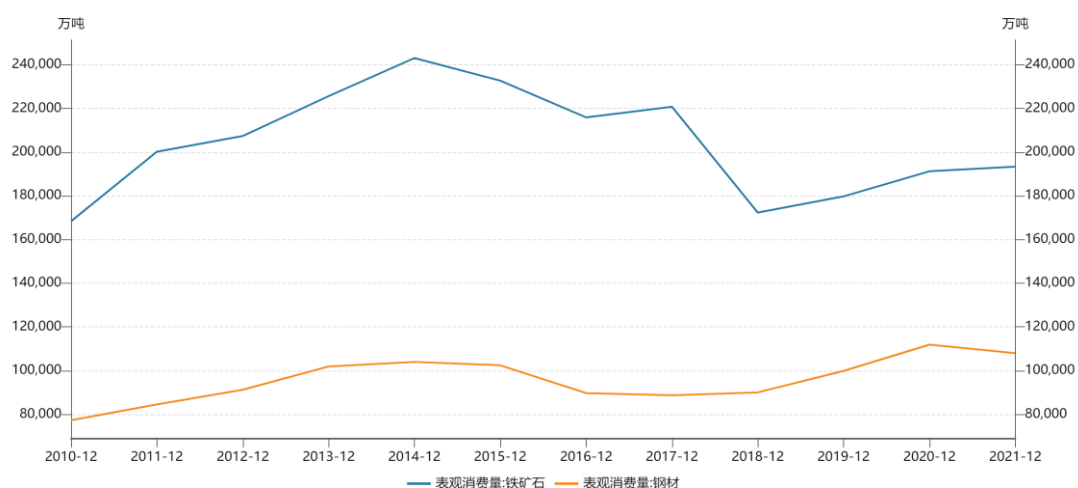
总而言之，我国铁矿石高度依赖于进口，巴西和澳大利亚的铁矿石发货量以及我国一些矿区的铁矿石产量决定了国内铁矿石的供给，这是影响铁矿石的价格走势的重要因素。

3.2.2 需求

与供给端相比，铁矿石的下游产业链相对较长，因此其需求方也就表现更加多元化。铁矿石的需求主要与中下游的产品有关，体现在两个方面：第一是铁矿石在加工生产方面的需求，进一步表现为钢材的需求，钢材的需求变化会对铁矿石需求产生比较直接的影响；第二是钢材产品下游应用行业的需求，主要以房地产、基建、机械、汽车等行业的需求为主。下游行业的用钢需求会直接影响钢材的需求，进而影响到铁矿石的需求。

1. 钢材

就钢铁的生产过程而言，目前最主要的炼钢工艺可以分为长流程和短流程炼钢，两者主要的区别为前者炼钢的原材料是铁矿石而后者是废钢。我国目前主要依赖于长流程炼钢，所以钢材的需求会对铁矿石的需求产生极大的影响，进而会影响铁矿石的价格。如图 3-5 所示，我国对于钢材的需求从长期来看基本属于稳步增长的状态，但在 2015 年有个拐点，其原因是 2015 年我国提出供给侧改革，推进“三去一降一补”政策，降低了我国对于钢材的需求，从而导致铁矿石的需求量也在减少，直至近几年来才有所好转，而铁矿石的需求也开始回升。从图 3-5 中钢材和铁矿石的需求曲线上来看，两者的相关性比较高，基本验证了铁矿石的需求会随钢材需求变化而变化。



数据来源：wind 数据库

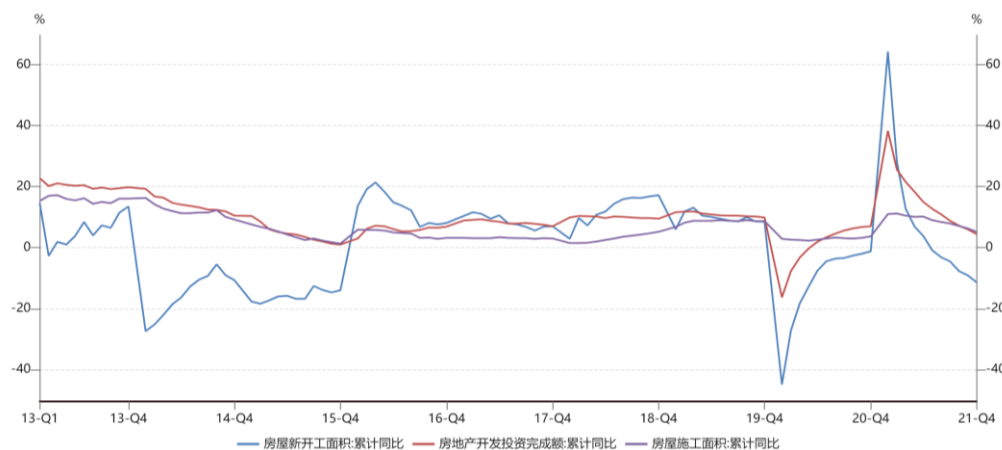
图 3-5 我国铁矿石与钢材的消费情况（2010-2021 年）

2. 下游应用领域

在钢铁产业下游的主要应用领域，我国建筑行业对于钢铁的需求量多年来处于第一位，其次是机械行业，再就是汽车行业，因此可以重点关注这几个行业的用钢需求。

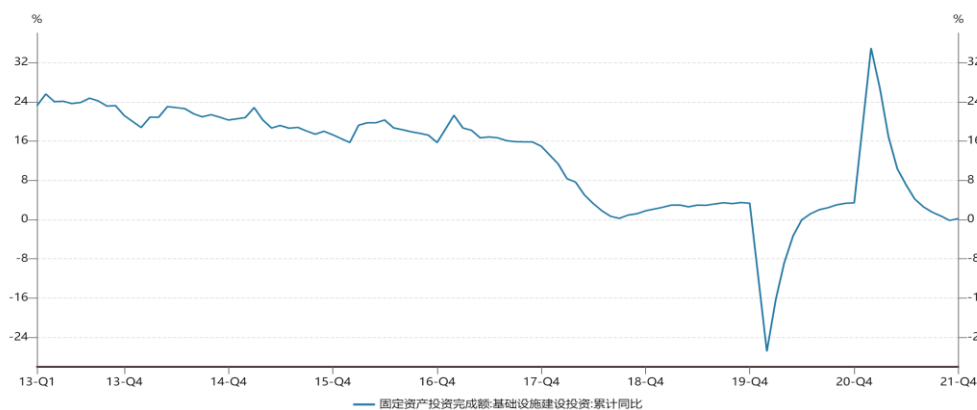
1) 建筑行业

我国建筑行业主要涉及的是基础设施建设（简称“基建”）以及房地产行业，考虑到螺纹钢是建筑行业中不可或缺的建材，它与房地产行业、基建行业的用钢需求有紧密的联系。我国房地产投资和开工情况、基建行业变化情况以及螺纹钢产量变化分别如图 3-6、图 3-7、图 3-8 所示。其中基础设施建设作为我国重点投资项目，其投资额在 2013-2017 年都是处于稳定增长的状态，直至 2018 年金融“去杠杆”政策的推出，基建投资出现断崖式下降。此时，我国在房地产方面的各项投资依旧是处于增长的状态，因此用钢需求的变化不是十分明显，对螺纹钢产量变化的影响也不大。2019 年末因为新冠疫情的影响，我国对于基建行业以及房地产行业的投资急骤下降，甚至出现负增长，用钢需求的减少致使螺纹钢的产量也急剧下降。直至经济回暖，全国各地开始复工，我国在房地产行业以及基建行业的投资才开始回增，螺纹钢的产量也出现增长的状态。2021 年全国推进由央行、住建部、银保监会对房地产企业提出的资金监测和融资管理规则（“三条红线”政策），这对房地产行业产生了一定的影响，房地产开发投资增速开始下降，其中房屋新开发面积甚至出现负增长。与此同时，我国宏观政策偏向结构化调整，基建投资的增速也下降至低位运行。房地产行业以及基建行业的用钢需求都在减少，此时螺纹钢产量的增速也呈现下降的趋势。



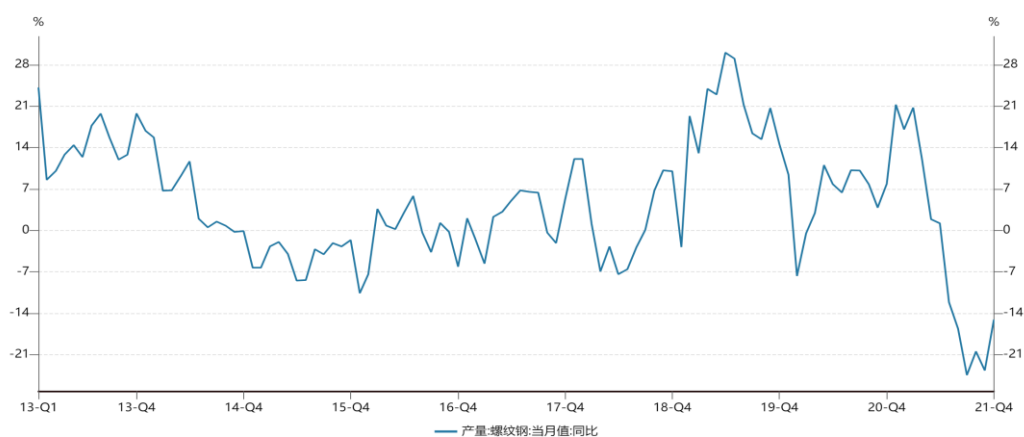
数据来源：wind 数据库

图 3-6 我国房地产市场投资、开工变化情况（2013-2021 年）



数据来源：wind 数据库

图 3-7 我国基础设施建设投资变化情况（2013-2021 年）



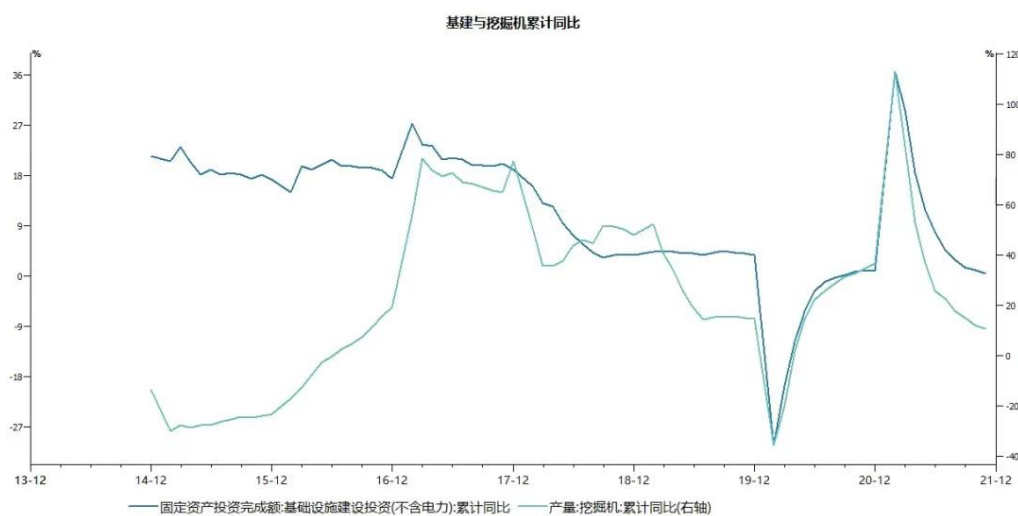
数据来源：wind 数据库

图 3-8 我国螺纹钢产量变化情况（2013-2021 年）

从房地产投资额、基建投资额的增速变化以及螺纹钢产量变化情况来看，螺纹钢需求的确会随房地产与基建行业用钢需求变化而变化，而螺纹钢需求量的变化会直接影响钢材的需求，进而影响铁矿石的需求。

2) 机械行业

对于工程机械行业而言，它的主要评价指标就是挖掘机的产量，而挖掘机的需求市场与基建行业也十分相关。从图 3-9 中可以看到挖掘机的产量变化走势与基建投资额变化基本一致，由此推断挖掘机的用钢需求和基建行业的用钢需求变化也基本一致，可以将两者结合共同考虑对于钢材需求的影响。

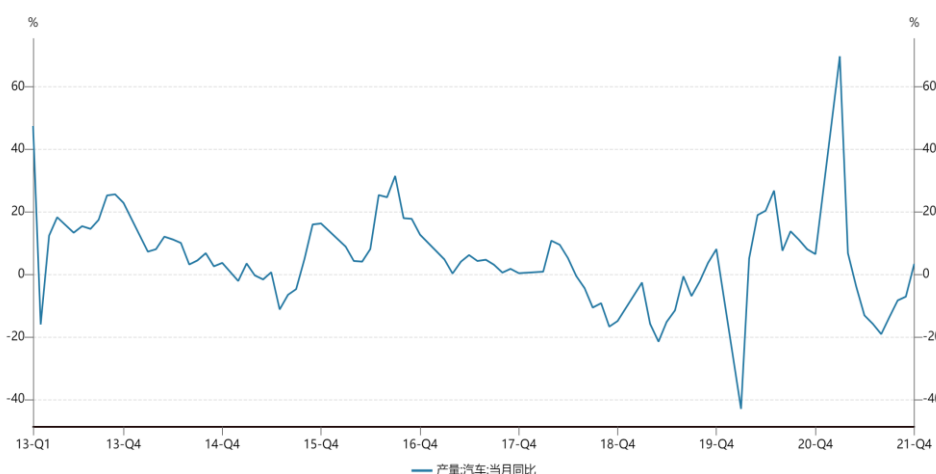


数据来源：wind 数据库

图 3-9 我国挖掘机产量与基础设施建设投资情况（2013-2021 年）

3) 汽车行业

对于我国的汽车行业，从图 3-10 可以看到在 2013-2019 年我国汽车生产总体上呈现增长的态势，但也存在一定的波动性。其中在 2016 年，受益于车辆购置税减半及新能源推广等政策，我国汽车行业增长明显。在 2018 年我国汽车行业不太景气，主要是因为国内宏观经济下行，消费力度整体下降，还有前几年的政策优惠所导致的汽车市场饱和。2019 年由于国内经济下行以及汽车行业出现缺乏芯片的问题，汽车行业需求延续 2018 年的水平，在 2019 年末产量甚至出现负增长趋势。直至 2021 年，随着缺芯问题得到边际缓解，汽车产量回升，因此也预计汽车行业的用钢需求会有所增加。



数据来源: wind 数据库

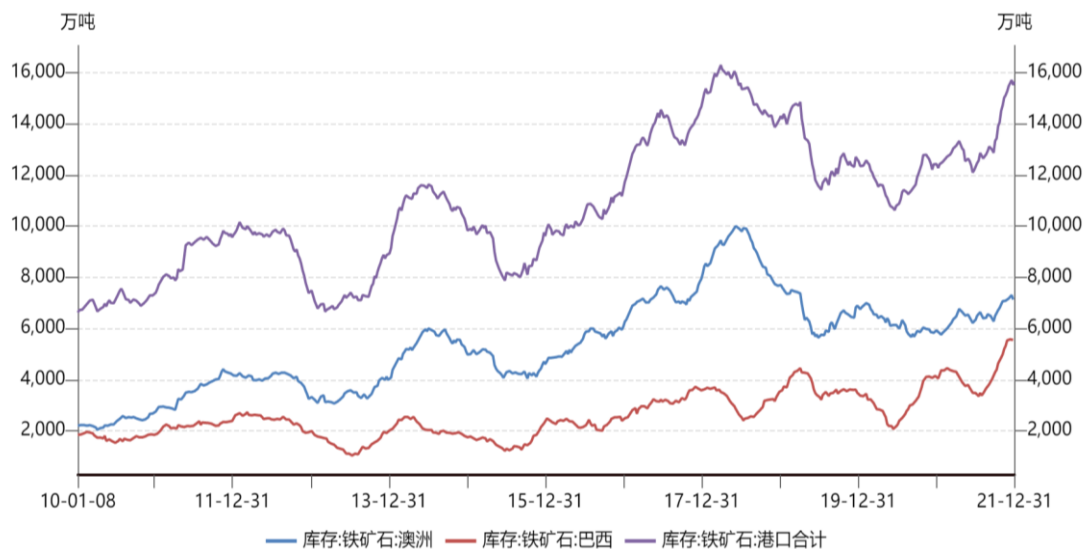
图 3-10 我国汽车产量情况（2013-2021 年）

考虑到热轧卷板主要应用于汽车以及工程机械行业，这两个行业的需求变化势必会影响到热轧卷板的需求，进而表现为钢材需求量的变化，而铁矿石的需求也将因此受到影响。

总体而言，铁矿石的需求来源比较多，不仅需要考虑下游钢铁产品的需求，而且需要考虑下游应用领域的用钢需求。在分析下游应用领域需求变化的过程中，本文发现下游主要应用行业对于钢材的需求会体现在对螺纹钢以及热轧卷板的需求中，因此可以推断螺纹钢、热轧卷板和铁矿石三者的需求存在较高的相关性。

3.2.3 库存

港口库存是供应紧张程度的一个体现，其变化会对铁矿石价格产生影响，比如当库存量增加时，会使市场担心供给过量，商家更加愿意出货，铁矿石价格会下降；如果库存量不足时，贸易商更加愿意囤货，这将导致铁矿石价格上涨。从图 3-11 可以看到，我国铁矿石进口库存在近几年波动较大，表示铁矿石的供求关系不稳定，当港口存量变化较大时，往往对于铁矿石现货和期货价格影响也会比较大，这就为套利创造了有利的条件。



数据来源：wind 数据库

图 3-11 我国铁矿石港口库存情况（2010-2021 年）

3.2.4 宏观因素

宏观经济因素是一个比较重要的影响期货价格的基本面因素，它能够对国家整体经济状况造成一定的影响，进而也会影响铁矿石的期货价格。具体而言，宏观经济上行会对钢材市场的下游应用领域产生正向影响，而下游市场的一系列变化则会带动铁矿石的需求，最终提振铁矿石的价格。

宏观因素中包含国家政策，国家针对钢铁行业制定的政策会对钢铁市场产生直接影响，从而影响钢铁产业中各商品现货以及期货的价格。自 2016 年以来，我国推出供给侧改革和钢铁行业“去产能”等举措，并不断出台各种环保政策以及推进一系列的环保措施，这对我国钢铁市场造成了极大影响。目前我国正处于“十四五”规划期间，深化供给侧改革依旧是主线任务，而在 2020 年所提出的“碳达峰、碳中和”的任务也还在持续推进，钢铁行业“去产能”的成果依旧是我国需要关注的重点。在这些背景下，持续关注国家政策对于把控铁矿石及钢材期货的价格变化具有重要意义。

3.3 本章小结

本章从铁矿石产业链出发，对铁矿石产业链做基本面分析之后，可以将其基本逻辑关系做以下总结：

从产业链的角度来看，铁矿石、螺纹钢、热轧卷板都是炼钢环节中不可或缺的元素。在铁矿石产业链的中游环节，铁矿石在经过一系列加工后可以冶炼成粗钢，

粗钢再经过塑形加工制成螺纹钢和热轧卷板，其中螺纹钢和热轧卷板除了塑型方式不同，其他的生产过程基本一致。基于上述产业链上的生产关系，可以发现铁矿石、螺纹钢、热轧卷板三者在本面上是紧密相关的。

从供需角度上来看，我国铁矿石主要依赖于国外进口，在供给方面受到国外铁矿石的供应量即四大矿山发货量的影响较大，而在其需求方面，重点考虑的是钢材的需求以及钢材下游应用领域的影响，供需关系的变化是影响铁矿石价格的重要因素。而钢材中的螺纹钢和热轧卷板主要被应用于建筑、机械、汽车等下游行业，它们的供需状况变化也受到这些下游应用行业的影响，因此其价格波动会使铁矿石期货的价格发生变化。

基于以上情况，无论从基本面还是价格关系联系上看，铁矿石、螺纹钢、热轧卷板三者之间都是有着紧密关联的，因此本章认为从铁矿石、螺纹钢、热轧卷板期货中选取跨品种套利组合以设计套利交易策略极具可行性。

第四章 跨品种套利策略设计

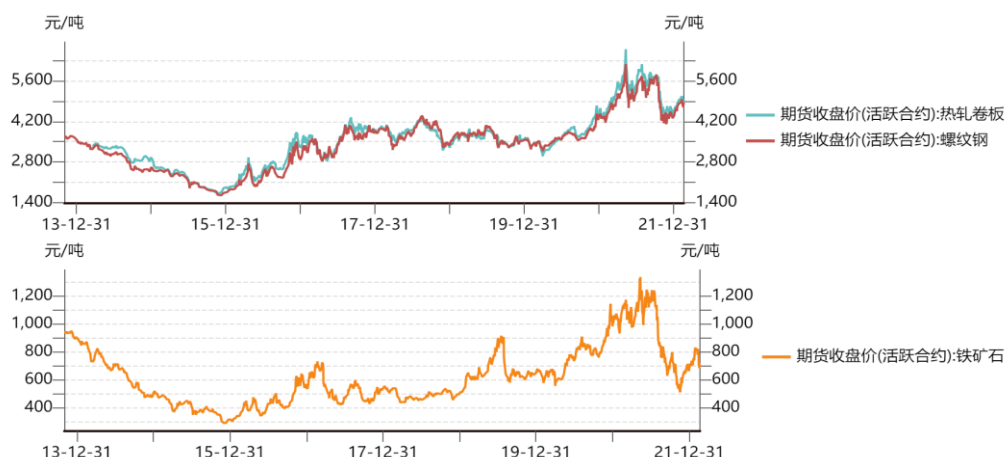
一个完整的跨品种套利策略通常包括以下两部分：

第一部分是策略设计阶段。资产组合具有稳定的价格关系是实现统计套利的首要条件，因此在设计策略之前要对跨品种套利组合的期货价格关系做稳定性检验。首先选取合理的样本数据，其次对所选取的数据进行稳定性分析，也就是判断两组价格序列是否处于一种均衡关系，只有处于长期均衡关系的资产组合才能够符合本文跨品种套利所需要的条件。具体操作就是用 ADF 检验判断所选取的样本数据是否平稳，再用 E-G 两步法、误差修正模型等计量方法检验数据的协整关系，当两组数据具备长期均衡关系时，构建相应的跨品种套利组合。最后设计交易策略，主要是对策略参数的设计，包括组合配比、交易阈值等。

第二部分是策略回测过程。本文运用聚宽平台对所设计的策略进行回测，根据回测结果的各项指标分析策略效果，并针对策略所存在的问题进行改进。

4.1 品种和数据选择

基于上一章的结论，从铁矿石产业链基本面的角度出发，可以考虑在铁矿石、螺纹钢、热轧卷板期货中选取跨品种套利组合。再根据价格关系分析，三种期货活跃合约的日收盘价如图 4-1 所示。



数据来源：wind 数据库

图 4-1 铁矿石、螺纹钢、热轧卷板期货的活跃合约日收盘价（2013-2021 年）

从图 4-1 的价格走势上看，三者的波动在长期具有相似性，应当具有较强的关联。另外从图 4-1 也可以发现，最近几年三个期货品种的价格波动都比较大，由此

推断三者之间存在的套利可能性也较大。再从活跃度上来看,铁矿石、螺纹钢和热轧卷板近三年的期货合约日平均成交量在两万手以上,市场成交量以及成交额都具有很好的流动性。所以无论是从经济联系还是活跃度上来看,铁矿石、螺纹钢和热轧卷板期货可以初步作为本文构建跨品种套利策略的商品期货对象。

基于此,本文选用铁矿石期货(I)、螺纹钢期货(RB)、热轧卷板期货(HC)自2019年1月2日至2021年12月31日的活跃合约日收盘价作为研究样本数据(数据来源于wind数据库),用以实证分析以及策略设计。在回测方面,本文以2019年1月1日至2021年12月31日为回测区间,在JionQuant聚宽量化交易平台上检验策略的适用性并分析其效果。

4.2 协整检验过程

4.2.1 相关性分析

在前文根据基本面分析初步获取了用以构建跨品种套利组合的商品期货品种后,还需要用定量分析法对其关系进行验证,即利用统计和计量的方法对前文认为可能相关的期货品种进行相关性检验。考虑到期货合约的价格序列是不平稳的,对其做相关性检验也不具参考性,所以在进行相关性分析时本文利用的是期货合约的对数收益率,计算公式如下:

$$R_t = \ln P_t - \ln P_{t-1} \quad (4-1)$$

其中, R_t 表示对数收益率; P_t 表示第 t 日的收盘价; P_{t-1} 表示第 $t-1$ 日的收盘价。

相关性分析则用以判断两个期货品种之间是否存在联系,是跨品种套利中选取配对合约最基本的判断方法,其中最重要的分析指标就是相关系数,通过计算相关系数可以确定各品种之间关系的紧密程度,相关系数的计算公式如下:

$$\rho_{XY} = r(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (4-2)$$

相关系数 ρ_{XY} 的取值在 $(-1, +1)$ 的范围内,当 $|\rho_{XY}| \leq 0.4$ 时, X 和 Y 被称作弱相关;当 $0.4 < |\rho_{XY}| \leq 0.6$ 时, X 和 Y 被称作中度相关;当 $0.6 < |\rho_{XY}| \leq 0.8$ 时, X 和 Y 被称作强相关;当 $0.8 < |\rho_{XY}| \leq 1$ 时, X 和 Y 被称作极强相关。

本文从 wind 数据库中分别获取铁矿石(I)、螺纹钢(RB)、热轧卷板(HC)3个期货在2019年1月2日至2021年12月31日的活跃合约日收盘价数据,根据

公式（4-1）计算出各期货品种的对数收益率，并按照公式（4-2）中相关系数的计算方法，分析各品种收益率的相关性，所得到的相关系数矩阵如表 4-1 所示。

表 4-1 铁矿石、螺纹钢、热轧卷板期货收益率之间的相关系数矩阵

相关系数	I	RB	HC
I	1	\	\
RB	0.5264	1	\
HC	0.4866	0.8245	1

从表 4-1 中可以看出，本文所选取的三种商品期货品种相关性都还算比较高，其中螺纹钢和热轧卷板的相关性最高，相关系数有 0.8245，达到了极强相关，而铁矿石和螺纹钢以及铁矿石和热轧卷板的相关性虽然较前者而言偏弱，但是相关系数也分别有 0.5264 和 0.4866，达到了中度相关。

无论是上章的产业链分析还是本节的相关性分析，都表明这三种商品期货品种之间具有较高的相关性，初步表明三者之间存在套利机会的假设成立。接下来本文将继续对该假设进行验证，判断这三种商品期货品种活跃合约日收盘价两两之间是否存在协整关系，从而为后续策略的设计提供配对组合。

4.2.2 单位根检验

为了更加直观的看出铁矿石期货（I）、螺纹钢期货（RB）、热轧卷板期货（HC）的日收盘价价格变化趋势，本文用 Eviews10 对铁矿石、螺纹钢、热轧卷板期货在 2019 年 1 月 2 日至 2021 年 12 月 31 日的样本内日收盘价数据绘制价格走势图，如图 4-2 所示。

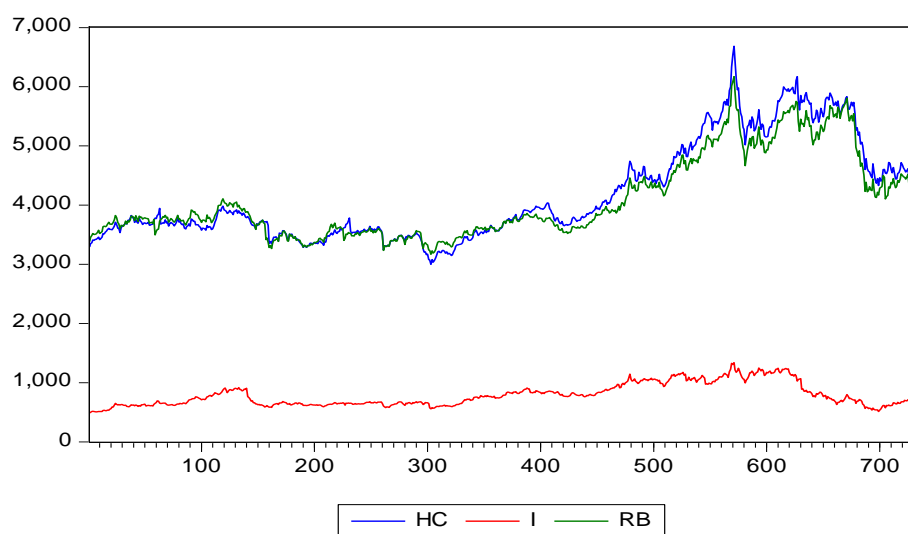


图 4-2 铁矿石、螺纹钢、热轧卷板期货的日收盘价（2019-2021 年）

从图 4-2 可以比较直观地看出三种商品期货的日收盘价的走势不存在规律性,初步判断三种商品期货的日收盘价是非平稳的,但是从价格长期走势来看,三者变动的趋势是比较一致的,符合本文之前对于三种商品期货的相关性分析,另外对于三者之间的协整关系还需要进一步检验。

协整关系的前提是时间序列必须为同阶单整的,所以在协整检验之前,首先要对数据做单位根检验,判断数据是否平稳。在单位根的检验结果中,有 1%、5%、10%的临界值反映不同的置信水平,通常是选择 5%的临界值作为参照水平,当 P 值 <0.05 时,拒绝原假设,说明时间序列是平稳的。

本文首先对 2019 年 1 月 2 日至 2021 年 12 月 31 日铁矿石期货、螺纹钢期货和热轧卷板期货的日收盘价取自然对数,获得 LGI, LGRB, LGHC 三个对数价格序列,再使用 Eviews10 对三个序列分别做单位根检验,得到的结果如表 4-2 所示。

表 4-2 LGI、LGRB、LGHC 的单位根检验结果

变量	P 值	t-Statistic	5% Critical Value
LGI	0.2173	-2.1710	-2.8652
LGRB	0.5062	-1.5529	-2.8652
LGHC	0.5870	-1.3927	-2.8652

从表 4-2 可以看到,三个序列在进行单位根检验后的 p 值都大于 0.05,且三者的 t 统计量均大于 5%显著水平下的临界值-2.8652,因此三个变量的单位根检验都不能拒绝原假设,即铁矿石、螺纹钢、热轧卷板期货的对数价格序列是非平稳的。

下一步对三个对数价格序列做一阶差分处理得到序列 $D(LGI)$ 、 $D(LGRB)$ 、 $D(LGHC)$,再对三者进行单位根检验,得到的结果如表 4-3 所示。

表 4-3 $D(LGI)$ 、 $D(LGRB)$ 、 $D(LGHC)$ 的单位根检验结果

变量	P 值	t-Statistic	5% Critical Value
$D(LGI)$	0.0000	-26.8135	-2.8652
$D(LGRB)$	0.0000	-28.0728	-2.8652
$D(LGHC)$	0.0000	-26.9247	-2.8652

从表 4-3 可以看到,三个新变量的 p 值均为 0.0000 小于给定的显著水平 0.05,且三者的 t 统计量也均远小于在 5%水平下的临界值,所以三个变量通过单位根检验,拒绝原假设,即三个商品期货的对数价格序列在一阶差分后是平稳的。

总而言之,在样本内铁矿石、螺纹钢、热轧卷板期货的对数价格序列具有一阶单整的性质,可以进行下一步协整关系的检验。

4.2.3 协整关系检验

对于两个一阶单整变量之间协整关系的检验，本文选用 E-G 两步检验法。

首先使用 Eviews10 分别对铁矿石、螺纹钢、热轧卷板期货的对数价格序列两两进行配对做 OLS 回归估计，构建相关回归方程：

1) 对铁矿石期货对数价格 (LGI) 和螺纹钢期货对数价格 (LGRB) 进行 OLS 回归估计，得到方程 (4-3)：

$$LGI_t = -1.2826 + 0.9546 * LGRB_t + \varepsilon_{1t} \quad (4-3)$$

$$(-3.9084) \quad (24.1497)$$

$$R^2 = 0.4447 \quad DW = 0.0190 \quad F = 583.2085$$

2) 对铁矿石期货对数价格 (LGI) 和热卷轧板期货对数价格 (LGHC) 进行 OLS 回归估计，得到方程 (4-4)：

$$LGI_t = -0.3941 + 0.8458 * LGHC_t + \varepsilon_{2t} \quad (4-4)$$

$$(-1.4227) \quad (25.4021)$$

$$R^2 = 0.4698 \quad DW = 0.0210 \quad F = 648.2691$$

3) 对螺纹钢期货对数价格 (LGRB) 和热轧卷板期货对数价格 (LGHC) 进行 OLS 回归估计，得到方程 (4-5)：

$$LGRB_t = 1.2143 + 0.8519 * LGHC_t + \varepsilon_{3t} \quad (4-5)$$

$$(29.9043) \quad (174.5375)$$

$$R^2 = 0.9766 \quad DW = 0.1423 \quad F = 30463.33$$

其次，计算上述三个回归方程的残差序列 μ_{1t} 、 μ_{2t} 、 μ_{3t} ，并分别进行单位根检验，分析各残差序列是否为平稳序列，由此判断各变量组合之间是否有协整关系。各残差序列的单位根检验的结果如表 4-4 所示。

表 4-4 残差序列 μ_{1t} 、 μ_{2t} 、 μ_{3t} 的单位根检验结果

序列	P 值	t-Statistic	1% Critical Value	5% Critical Value	10% Critical Value
μ_{1t}	0.0367	-2.0728	-2.5681	-1.9412	-1.6164
μ_{2t}	0.0321	-2.1289	-2.5681	-1.9412	-1.6164
μ_{3t}	0.0000	-5.1913	-2.5681	-1.9412	-1.6164

从表 4-4 可以看出，三个残差序列在经过单位根检验后的 P 值均小于显著水平 0.05，且 t 统计量也均小于在 5% 置信水平下的临界值，因此拒绝原假设，也就

是说上述三个残差序列均具有平稳性。

根据 E-G 两步检验法, 上述三种商品期货的对数价格序列是一阶单整的, 其两两组合进行 OLS 回归估计后得到的残差序列也都是平稳的, 由此判定铁矿石期货与螺纹钢期货、铁矿石期货与热轧卷板期货、螺纹钢期货与热轧卷板期货三个组合都具有协整关系。

4.2.4 误差修正模型

在 4.2.3 节已经验证了本文所选的三种期货品种两两之间存在协整关系, 也就是说铁矿石期货与螺纹钢期货、铁矿石期货与热轧卷板期货、螺纹钢期货与热轧卷板期货三个组合的价差都是长期稳定的。在这种长期均衡状态下, 可能会有短期的价格异常变化打破这种状态, 但因为协整关系的存在, 短期内的偏离最终还是会被修正, 价差会恢复到原来的均衡状态, 而这种短期内的偏离则带来了实现跨品种套利的机会。误差修正模型就是用来解释价差在短期内偏离均衡后, 如何不断进行调整直至恢复均衡状态。

在上一节协整关系检验的基础上, 本文分别以铁矿石期货、螺纹钢期货、热轧卷板期货的对数价格序列为变量, 构建以下三组一阶误差修正模型, 其具体表达式分别如下:

$$\Delta LGI_t^1 = \beta_1 \Delta LGRB_t^1 - \lambda_1 ecm_{1(t-1)} + \varepsilon_{1t} \quad (4-6)$$

$$\Delta LGI_t^2 = \beta_2 \Delta LGHC_t^2 - \lambda_2 ecm_{2(t-1)} + \varepsilon_{2t} \quad (4-7)$$

$$\Delta LGRB_t^3 = \beta_3 \Delta LGHC_t^3 - \lambda_3 ecm_{3(t-1)} + \varepsilon_{3t} \quad (4-8)$$

上述公式中, ΔLGI_t 、 $\Delta LGRB_t$ 、 $\Delta LGHC_t$ 分别表示的是铁矿石期货、螺纹钢期货、热轧卷板期货的对数价格在 t 时刻的一阶差分项; λ_1 、 λ_2 、 λ_3 是误差修正系数; $ecm_{1(t-1)}$ 、 $ecm_{2(t-1)}$ 、 $ecm_{3(t-1)}$ 分别代表的是回归方程式 (4-3)、式 (4-4)、式 (4-5) 所得的残差序列的一阶滞后项; ε_{1t} 、 ε_{2t} 、 ε_{3t} 代表的是随机误差项。

下一步本文用 Eviews10 对所构建的误差修正模型进行 OLS 回归估计, 回归得到如表 4-5 所示的结果。

从表 4-5 中可以看到三个组合的误差修正项 ECM 系数都在 $(-1, 0)$ 的区间内, 这表明上述回归方程的残差序列符合负反馈的修正机制。对于铁矿石与螺纹钢期货组合而言, 当前一期两个期货合约的对数价格序列偏离均衡时, 在这一期误差修正项将会以 0.0102 的力度进行反向修正, 直到恢复均衡状态。相似地, 对于其他两组期货组合的误差修正分析也是这样的, 这也充分表明了三个组合的残差序列的确拥有均值回复的规律, 这是跨品种组合获取套利机会的重要前提。

表 4-5 误差修正模型的估计结果

解释变量	估计系数	t-Statistic	P 值
C_1	0.0001	0.1734	0.8623
$\Delta LGRB_t^1$	0.9006	16.8185	0.0000
$ecm_1(-1)$	-0.0102	-2.0078	0.0450
C_2	0.0001	0.1129	0.9101
$\Delta LGHC_t^2$	0.8782	15.1808	0.0000
$ecm_2(-1)$	-0.0116	-2.1643	0.0308
C_3	-1.42E-05	-0.0417	0.9667
$\Delta LGHC_t^3$	0.8682	40.0755	0.0000
$ecm_3(-1)$	-0.0714	-5.1994	0.0000

4.3 策略设计

4.3.1 策略设计原理

在前文对铁矿石、螺纹钢、热轧卷板期货品种进行了定性以及定量的分析之后，可以发现这三种期货之间存在协整关系且存在极大的套利可能性，因此在本节将分别构建铁矿石——螺纹钢、铁矿石——热轧卷板、螺纹钢——热轧卷板三个商品期货跨品种套利组合，并对此设计相应的交易策略。

商品期货的跨品种套利是以均值回复为基础，利用两者的价差来构造套利区间。具体而言就是，假设两个商品期货的对数价格序列分别为 LGX 和 LGY，当其为同阶单整的序列时，进行 OLS 回归估计后，可计算出残差序列如下：

$$resid = LGY - \alpha - \beta * LGX \quad (4-9)$$

在式 (4-9) 中，残差序列 resid 近似于看作商品期货 X 和期货 Y 的价差序列。

然后进行协整检验，也就是对价差序列 resid 做单位根检验。当价差序列 resid 通过单位根检验时，认为价差序列 resid 拥有均值回复的规律，因此当 resid 偏离均值时，可以进行交易。根据均值回复理论，如果 resid 向上偏离均衡值，应当做空价差，也就是做空被高估的商品期货品种 Y，同时做多被低估的商品期品种货 X，反之，如果 resid 向下偏离均衡值，可以做多价差。另外当 resid 回复至均值范围时，就做平仓操作。但并不是所有的套利组合的价差在偏离均值后都能够按照期望回复至均值附近，也有可能会一直偏离，甚至到组合间不存在协整关系。面对以上情况，就还需要设置相应的止损条件以控制风险。

另外在以往的研究中，策略思想通常是获取固定时间内的样本数据，然后对样本数据做 OLS 回归构建回归方程，通过回归所获得残差序列计算得到固定的均值，

并根据这个方程的回归估计参数以及商品期货的价格计算出新的价差, 通过比较新价差和固定均值的偏离程度获取套利机会。但是这个方法实际上不是很客观, 因为在长期均衡关系中, 商品期货基本面信息的改变可能会导致两个商品期货的均衡关系发生改变, 这并不仅仅是指两者的均衡关系不成立, 还有可能是两者之间的价差以及回归方程的参数发生改变, 从而导致前后均值差异较大。如果交易策略还是按照初始回归方程的系数设计相关的策略参数以及计算新的价差, 在这种情况下可能会有错误的套利信号产生, 致使策略出现错误套利。

针对以上情况, 本文将采用变动的均值和系数来设定策略参数, 也就是在移动时间窗口内构建回归方程以更新参数。具体操作是策略在进行交易判断之前先获取前一段时间 t 的期货价格为样本数据, 再对获取的数据做 OLS 回归估计构建回归方程, 根据该方程实时计算价差序列和新价差。在这个过程中, 样本数据会根据交易的时间一直进行更新, 从而使得策略参数能够得到及时地调整, 策略参数的具体设计将在下文详细介绍。

4.3.2 套利组合的配比

跨品种套利主要构建的是风险中性的策略, 确定合理的商品期货头寸是满足风险中性的一个重要条件。本文根据式 (4-9) 中所得到的两个商品期货组合的协整向量 $(1, \beta)$ 来建立交易的头寸, 从组合的意义上来说, 应当是在买进 1 单位的期货 Y 同时卖出 β 单位的期货 X。考虑到期货交易是存在杠杆的, 所以需要进行合理地资金管理来控制风险, 在策略设计时, 本文采用五分之一的账户资金对期货 Y 建立仓位, 以降低穿仓风险。另外程序的交易指令需要具体到数量, 期货在实际交易中是按照合约的整手数下单, 当买入 1 手商品期货 Y 时需要的资金为: 期货 Y 的价格 * 期货 Y 的合约乘数 * 保证金比例, 因此可以通过账户资金来计算出期货组合的具体下单数量。根据上述的组合头寸比例和资金管理, 可以按照以下公式分别计算组合中两个期货的下单数量:

$$\text{期货Y的下单数量} M = \text{celing} \left(\frac{20\% * S}{F * H * R} \right) \quad (4-10)$$

$$\text{期货X的下单数量} N = \text{celing} (M * \beta) \quad (4-11)$$

式 (4-10) 中, S 表示账户资金; F 表示期货 Y 的主力合约收盘价; R 表示期货 Y 的保证金比例; H 表示期货 Y 的合约乘数 (铁矿石期货为 100 吨/手, 螺纹钢期货和热轧卷板期货为 10 吨/手); *celing* 表示向上取整。

式 (4-11) 中, M 是期货 Y 的下单数量; β 是式 (4-9) 中所得的协整系数; *celing* 是向上取整。

4.3.3 交易信号的设计

除了套利组合的配比,策略设计中最主要的参数就是开仓、平仓以及止损信号,只有交易信号设计地比较完善,策略才能尽可能地在保证收益的同时降低风险。根据式(4-9),可以得到在时间窗口 t 时段的价差序列 resid 及其标准差 $\text{std}(\text{resid})$,再计算出在 $t+1$ 时刻的新价差 resid' ,而交易信号就是依据新价差 resid' 与价差序列均值的偏离程度来设计。

考虑到价差序列 resid 是符合标准正态分布的,所以在设计期货跨品种套利策略的交易信号时,可以以价差序列 resid 的 N 倍标准差来设计相关交易信号阈值。开仓阈值的选择需要权衡,过高的阈值会使开仓次数减少,从而可能错过获利的机会;过小的阈值则会导致价差在偏离均值不够大的时候下开仓,这种情况下交易盈利的可能性或许不太高。综合考虑下,本文参考 Vidyamurthy (2004)^[15] 的思想以固定参数设计交易阈值,他表示假设去中心化的价差波动是一个白噪声序列,其最大收益的交易边界阈值是 ± 0.75 倍的标准差。因此,本文将开平仓阈值的 N 设为 ± 0.75 ,将止损边界阈值的 N 设置为 ± 2 。

另外对于交易信号的设计,本文还需要特别说明以下三点:

1.开仓时不仅仅是新价差需要通过所设定的阈值,而且期货 X 和期货 Y 也要通过协整检验。所以在进行开仓条件判断之前,首先要对价差序列 resid 做单位根检验,如果该价差序列没有通过单位根检验,便认为期货 X 和期货 Y 之间不存在协整关系,没有套利的机会。当期货 X 和期货 Y 通过协整检验后,再进行开仓条件的判断;

2.只在空仓的时候才进行开仓,不出现重复开仓的操作,另外只有当仓位被平仓后才开始下一次开仓;

3.考虑到期货主力合约变更会导致手中合约无法及时平仓的情况,所以当主力合约发生变更时,本文所设计的策略会进行强制换仓操作,移仓至次主力合约。

4.3.4 具体策略设计

根据前文所设置的组合配比以及交易信号,接下来将对策略设计进行具体地阐述:

首先,在每天开盘前对移动时间窗口 60 日内的期货 X 和期货 Y 的活跃合约日收盘价数据做 OLS 回归,得到价差序列 resid 及其标准差 $\text{std}(\text{resid})$;

其次,对价差序列 resid 做 ADF 检验,如果它通过单位根检验,获取前一日的期货 X 和期货 Y 的活跃合约日收盘价,再将其代入 OLS 回归方程中计算出新的价差 resid' ;

最后, 根据仓位情况以及新价差 resid' 的大小做开仓判断, 具体操作如下:

当处于空仓的状况时:

如果价差 resid' 上穿开仓点, 即 resid' 大于 $0.75 \cdot \text{std}(\text{resid})$ 且小于 $2 \cdot \text{std}(\text{resid})$ 时, 本文认为在此刻期货 Y 被高估期货 X 被低估, 所以应当做多期货合约 X 做空期货合约 Y, 也就是买入 $M \cdot \beta$ 手的期货合约 X, 卖出 M 手的期货合约 Y, 其中, M 的计算方法如式 (4-10)。

如果价差 resid' 下穿开仓点, 即 resid' 小于 $-0.75 \cdot \text{std}(\text{resid})$ 且大于 $-2 \cdot \text{std}(\text{resid})$ 时, 本文认为在此刻期货 X 被高估期货 Y 被低估, 所以应当做多期货合约 Y 做空期货合约 X, 也就是买入 M 手的期货合约 Y, 卖出 $M \cdot \beta$ 手的期货合约 X。

当处于非空仓的状况时:

如果价差 resid' 的绝对值处于 $(0, 0.75 \cdot \text{std}(\text{resid}))$ 的区间内, 本文认为此时价差 resid' 处于均衡状态, 不存在套利机会, 所以需要対目前所持仓位做反向平仓的操作。

如果价差 resid' 的绝对值大于 $2 \cdot \text{std}(\text{resid})$, 本文认为此时组合的协整关系可能不成立, 均值回复的特征大概率不会出现, 所以需要対当前的仓位做止损平仓操作。

该策略的具体思路构架如图 4-3 所示:

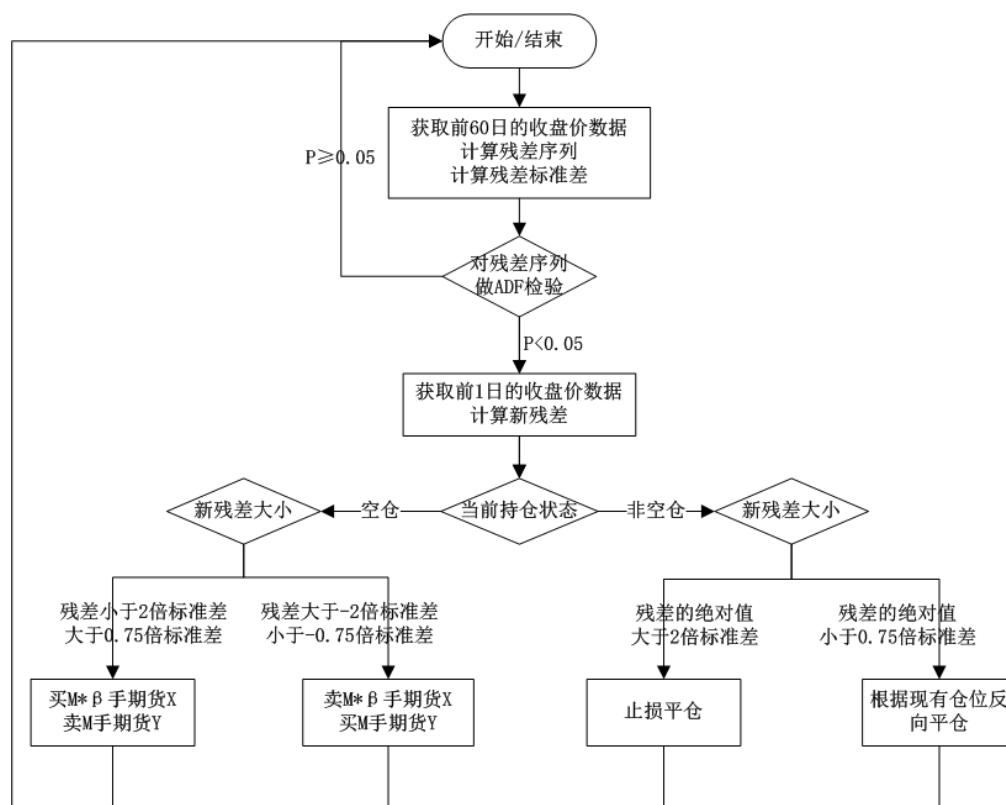


图 4-3 策略思路构架

4.4 本章小结

本章对跨品种套利策略的设计做了详细的介绍，其中包括品种的选取、跨品种套利组合的构建以及策略的具体设计。

在上一章对铁矿石产业链分析的基础上，本章结合期货品种之间的经济联系和活跃程度，确定了铁矿石、螺纹钢、热轧卷板期货为构建跨品种套利组合的主要对象，再用协整模型分析三个品种构建套利组合的可行性。本章以三个期货活跃合约在 2019 年 1 月 2 日至 2021 年 12 月 31 日时间段的日收盘价为样本数据，通过相关性分析发现三个期货之间有较强的相关性。其次用 ADF 检验验证了三个期货对数价格序列是一阶单整的时间序列。再采用 E-G 两步法检验三个期货对数价格序列之间的协整关系，发现三个对数价格序列两两之间存在长期均衡的关系，并基于误差修正模型，确定了其价差序列拥有均值回复的规律，三个期货两两之间可以构建跨品种套利组合。

最后是基于价差均值回复的规律设计策略，利用价差变化构建套利区间，将开平仓、止损信号阈值分别设置为 ± 0.75 倍、 ± 2 倍价差序列的标准差，并以期货价格序列在滚动时间窗口内回归所得的协整系数设计套利组合配比。

第五章 策略回测分析

5.1 绩效评价指标

在投资的过程中有很多评价指标可以用于绩效评估，其中最主要的关注点就是组合的风险以及收益。本文用以进行策略回测的平台是聚宽量化交易平台，在回测结束后，该平台会提供相应的评价指标来帮助判别交易策略的效果，其中包括收益类指标、风险类指标以及综合类指标等。本文主要选取了以下一些指标来评价所构建的期货跨品种交易策略的效果：

1. 策略收益率

它是指策略在回测时间段所获得的总收益率，其具体表达式如下：

$$R_{total} = \frac{(P_{end} - P_{start})}{P_{start}} * 100\% \quad (5-1)$$

式中， R_{total} 代表策略收益率； P_{start} 和 P_{end} 分别代表的是策略所持有的期货头寸以及现金最开始的价值以及最后的价值。

2. 策略年化收益率

它是指将策略总收益率换成相应的年收益率，可以用于不同策略的收益率的对比，其具体表达式如下：

$$R_{annual} = \left((R_{total} + 1)^{\frac{250}{n}} - 1 \right) * 100\% \quad (5-2)$$

式中， R_{annual} 代表策略年化收益率； R_{total} 代表策略总收益率； n 代表策略回测所运行的天数。

3. 策略波动率

它是指策略收益的波动变化，用以衡量策略的风险大小，其具体表达式如下：

$$\sigma_p = \sqrt{\frac{250}{n-1} \sum_{i=1}^n (R_{day} - \overline{R_{day}})^2} \quad (5-3)$$

式中， σ_p 代表策略波动率； R_{day} 代表的是策略日收益率； $\overline{R_{day}}$ 代表的是它的平均值； n 代表策略回测所运行的天数。

4. 夏普比率

它是指风险与收益之间的关系，也就是投资人每承担一单位的风险所能收获的超额收益的大小，其具体表达式如下：

$$Sharpe\ Ratio = (R_{annual} - R_f) / \sigma_p \quad (5-4)$$

式中， $Sharpe\ Ratio$ 代表夏普比率； R_{annual} 代表策略的年化收益率； R_f 代表无风险利率。

5. 最大回撤

它是指策略在回测过程中收益率最大的回撤幅度，可以用于衡量策略最严重的亏损状况，其具体表达式如下：

$$Max\ Drawdown = \frac{Max(P_m - P_n)}{P_m} \quad (5-5)$$

式中， $Max\ Drawdown$ 代表最大回撤； P_m 和 P_n 代表的是策略在 m 日和 n 日 ($m > n$) 分别所持有的期货头寸以及现金的总价值。

6. 胜率

它是指策略在回测期间盈利次数在总交易次数中的占比。

5.2 策略回测结果

根据上章制定交易机制以及相关交易信号，利用聚宽量化交易平台对所设计的跨品种策略进行回测，回测区间为 2019 年 1 月 1 日至 2021 年 12 月 31 日，移动时间窗口为 60 日，买入、卖出手续费为万分之一，保证金比例为 12%，回测频率为每日，初始资金为 100 万。三个期货组合的回测结果分别如下：

1. 螺纹钢-铁矿石 (RB-I) 跨品种组合的回测结果

策略收益	策略年化收益	超额收益	基准收益	阿尔法	贝塔	夏普比率	胜率	盈亏比	最大回撤 ②	索提诺比率
31.18%	9.74%	-4.34%	37.13%	0.060	-0.039	0.333	0.534	1.469	14.63%	0.415
日均超额收益	超额收益最大回撤	超额收益夏普比率	日胜率	盈利次数	亏损次数	信息比率	策略波动率	基准波动率	最大回撤区间	
0.03%	55.61%	-0.128	0.460	31	27	-0.040	0.172	0.366	2020/08/10, 2021/03/22	



图 5-1 螺纹钢与铁矿石期货跨品种交易策略回测结果

从图 5-1 可知,在以日频率回测的情况下,本文所构建的螺纹钢与铁矿石期货跨品种交易策略收益率为 31.18%,年化收益率为 9.74%,单从收益率上来看该策略表现一般。最大回撤为 14.63%,策略波动率为 17.2%,说明该策略也还是存在一定程度的风险。另外该策略的夏普比率仅为 0.333,胜率为 0.534 也只有一半,说明该策略还有改进的空间。

从上述评价指标可以发现,螺纹钢与铁矿石期货组合套利策略在日频交易中表现比较稳定,但策略绩效一般,后续可以通过改变交易信号阈值来优化策略绩效。

2.热轧卷板-铁矿石(HC-I)跨品种组合的回测结果

策略收益	策略年化收益	超额收益	基准收益	阿尔法	贝塔	夏普比率	胜率	盈亏比	最大回撤②	索提诺比率
39.01%	11.94%	1.37%	37.13%	0.069	0.146	0.242	0.619	1.451	53.22%	0.299
日均超额收益	超额收益最大回撤	超额收益夏普比率	日胜率	盈利次数	亏损次数	信息比率	策略波动率	基准波动率	最大回撤区间	
0.04%	63.60%	-0.077	0.492	39	24	0.012	0.328	0.366	2020/02/10,2021/09/09	



图 5-2 热轧卷板与铁矿石期货跨品种交易策略回测结果

从图 5-2 可知,在以日频率回测的情况下,热轧卷板与铁矿石期货组合的策略收益率为 39.01%,其年化收益率为 11.94%,夏普比率为 0.242,胜率为 0.619。但该策略的最大回撤达到了 53.22%,策略波动率为 32.8%,可见热轧卷板与铁矿石期货组合收益波动幅度较大,存在较高的风险,这也说明该策略在止损方面还有待进一步改进。

从上述评价指标可以发现,热轧卷板与铁矿石期货组合套利策略在日频交易的最大回撤较大,风险较高,后续可以通过调整交易信号阈值对策略进行改进。

3.螺纹钢-热轧卷板(RB-HC)跨品种组合的回测结果

从图 5-3 可以看到,在以日频率回测的情况下,以螺纹钢与热轧卷板期货组合构建的交易策略的收益率为 23.06%,年化收益率为 7.37%,策略收益并不是特别高。但该策略的最大回撤为 7.53%,策略波动率为 8.5%,可见该策略的风险还是比较低的。另外该策略的夏普比为 0.394,胜率为 0.538 也只有一半,可见策略还有优化的空间。

通过上述评价指标可以发现，螺纹钢与热轧卷板期货组合套利策略在以日频率的交易中表现比较稳定，相对而言策略效果还不错。但从收益走势图上看，该组合在回测期间的收益走势都十分平缓，原因是热轧卷板期货与螺纹钢期货在这三年的价差波动不大，长期处于均衡状态，由此推断可能是开仓阈值的大小设置不当影响了策略绩效，后续可以通过调整交易阈值来改善这一情况。

策略收益	策略年化收益	超额收益	基准收益	阿尔法	贝塔	夏普比率	胜率	盈亏比	最大回撤①	索罗斯比率
23.06%	7.37%	-3.03%	26.91%	0.033	0.019	0.394	0.538	1.096	7.53%	0.552
日均超额收益	超额收益最大回撤	超额收益夏普比率	日胜率	盈利次数	亏损次数	信息比率	策略波动率	基准波动率	最大回撤区间	
0.01%	46.02%	-0.212	0.478	49	42	-0.049	0.085	0.222	2021/10/25,2021/11/19	



图 5-3 螺纹钢与热轧卷板期货跨品种交易策略回测结果

5.3 交易信号阈值优化

根据本文所设计的策略的开仓条件，两个期货品种在具备协整关系的同时其价差也需要超过开仓信号阈值，这时才能进行开仓操作。因此，开仓信号阈值的大小对于策略绩效也很有影响。根据 5.2 节的回测结果可以发现，在有些时间段策略的收益走势基本不动，除了组合协整关系不存在的情况外，也有可能是价差波动过小或者过大导致没有满足开仓条件。

因此参照三个组合的价差序列，在本节将开仓信号阈值设置在 0.35 倍价差序列标准差到 1.45 倍价差序列标准差的区间，并以 0.1 倍价差序列标准差为步长进行遍历。综合考虑策略收益以及风险后，为各组合选取最佳的开仓信号阈值。参照 5.2 节所设置的回测条件，三个组合的回测结果如下。

1. 螺纹钢-铁矿石 (RB-I) 跨品种组合开仓阈值遍历回测结果

从表 5-1 的回测结果可以看到，当开仓阈值为 0.75 倍价差标准差时，此时构建的螺纹钢与铁矿石期货组合交易策略收益最大，收益率为 31.18%，年化收益率为 9.74%，夏普比最大为 0.333，最大回撤为 14.63%。对比其他回测结果，本文认为 0.75 倍价差标准差是螺纹钢与铁矿石组合的最佳开仓信号阈值。

表 5-1 螺纹钢与铁矿石组合开仓阈值遍历回测结果

开仓阈值	策略收益	策略年化收益	夏普比率	胜率	策略波动率	最大回撤
0.35	1.20%	0.41%	-0.126	0.531	28.5%	52.52%
0.45	-3.42%	-1.18%	-0.194	0.500	26.8%	48.45%
0.55	-27.61%	-10.48%	-0.652	0.484	22.2%	47.31%
0.65	-9.51%	-3.37%	-0.350	0.536	21.1%	37.73%
0.75	31.18%	9.74%	0.333	0.534	17.2%	14.63%
0.85	18.21%	5.90%	0.111	0.459	17.1%	16.46%
0.95	22.67%	7.25%	0.200	0.529	16.2%	16.56%
1.05	14.49%	4.74%	0.051	0.522	14.5%	14.30%
1.15	16.48%	5.36%	0.099	0.528	13.7%	14.74%
1.25	18.44%	5.97%	0.168	0.500	11.7%	15.29%
1.35	15.84%	5.16%	0.099	0.467	11.7%	15.04%
1.45	21.87%	7.01%	0.299	0.429	10.0%	10.23%

2. 热轧卷板-铁矿石（HC-I）跨品种组合开仓阈值遍历回测结果

表 5-2 热轧卷板与铁矿石组合开仓阈值遍历回测结果

开仓阈值	策略收益	策略年化收益	夏普比率	胜率	策略波动率	最大回撤
0.35	86.88%	23.88%	0.501	0.631	39.7%	57.01%
0.45	60.31%	17.54%	0.349	0.594	38.8%	56.75%
0.55	90.12%	24.61%	0.540	0.638	38.2%	54.82%
0.65	24.27%	7.73%	0.112	0.615	33.4%	56.41%
0.75	39.01%	11.94%	0.242	0.619	32.8%	53.22%
0.85	25.63%	8.13%	0.142	0.623	29.2%	48.64%
0.95	8.30%	2.77%	-0.044	0.563	28.1%	49.51%
1.05	-7.49%	-2.63%	-0.247	0.607	26.8%	51.66%
1.15	-22.21%	-8.24%	-0.461	0.588	26.5%	52.33%
1.25	-38.71%	-15.44%	-0.763	0.550	25.5%	53.96%
1.35	-37.89%	-15.05%	-0.759	0.559	25.1%	53.19%
1.45	-56.27%	13.88%	0.529	0.429	18.7%	17.35%

根据表 5-2 的回测结果可以发现，热轧卷板与铁矿石组合在开仓阈值为 0.55 倍价差标准差时策略收益达到最大，策略收益为 90.12%，年化收益率为 24.61%，此时最大回撤为 54.82%，策略波动率为 38.2%，夏普比率为 0.540。对比其他回测结果，此时是风险相对较小而收益较大的情况，因此可以得到热轧卷板与铁矿石组合的最佳开仓信号阈值为 0.55 倍价差标准差。

3.螺纹钢-热轧卷板（RB-HC）跨品种组合开仓阈值遍历回测结果

表 5-3 螺纹钢与热轧卷板组合开仓阈值遍历回测结果

开仓阈值	策略收益	策略年化收益	夏普比率	胜率	策略波动率	最大回撤
0.35	38.16%	11.71%	0.667	0.517	11.6%	11.91%
0.45	33.33%	10.35%	0.586	0.537	10.8%	11.94%
0.55	34.42%	10.66%	0.658	0.547	10.1%	11.95%
0.65	28.68%	9.02%	0.548	0.548	9.2%	10.96%
0.75	23.06%	7.37%	0.394	0.538	8.5%	7.53%
0.85	23.25%	7.42%	0.403	0.517	8.5%	7.25%
0.95	23.42%	7.47%	0.358	0.507	9.7%	9.45%
1.05	24.18%	7.70%	0.476	0.541	7.8%	7.47%
1.15	24.25%	7.72%	0.485	0.529	7.7%	7.47%
1.25	17.02%	5.53%	0.217	0.533	7.1%	7.54%
1.35	11.24%	3.71%	-0.045	0.533	6.3%	7.72%
1.45	13.21%	4.34%	0.056	0.542	6.1%	7.60%

从表 5-3 的回测结果可以看到，当开仓阈值为 0.35 倍价差标准差时，此时螺纹钢与热轧卷板组合的跨品种交易策略收益率达到最大值为 38.16%，年化收益率为 11.71%，另外此时夏普比率最大为 0.667，而策略最大回撤也仅仅只有 11.91%。无论是从收益还是控制风险的角度来看，应当选择 0.35 倍价差标准差作为螺纹钢与热轧卷板组合跨品种交易策略的最佳开仓信号阈值。

从以上回测结果中可以发现，在进行开仓信号阈值优化之后，三个组合的策略绩效都得到了提升，但其中热轧卷板与铁矿石组合的风险依旧很高，可见在止损方面还有待改进。通过对各策略绩效的分析，发现可能是止损信号设置过高导致策略无法及时止损，因此接下来在开仓信号阈值为各组合最佳开仓信号阈值的基础上，将止损信号阈值设置在 1.5 倍价差序列标准差到 2 倍价差序列标准差的区间，并以 0.1 倍价差序列标准差为步长进行遍历，为各组合选择更优的止损信号阈值，以提高策略绩效。

参照 5.2 节所设置的回测条件，三个组合的回测结果如下：

1.螺纹钢-铁矿石（RB-I）跨品种组合止损阈值遍历回测结果

从表 5-4 的回测结果可以看到，当开仓阈值为 0.75 倍价差标准差、止损阈值为 1.8 倍价差标准差时，此时构建的螺纹钢与铁矿石期货组合交易策略收益最大，收益率为 47.11%，年化收益率为 14.13%，另外最大回撤为 13.78%，策略波动率为 16.0%，夏普比为 0.632。无论是从收益还是风险上来看，策略绩效都得到了优化。

表 5-4 螺纹钢与铁矿石组合止损阈值遍历回测结果

止损阈值	策略收益	策略年化收益	夏普比率	胜率	策略波动率	最大回撤
1.5	35.64%	11.00%	0.478	0.577	14.6%	14.39%
1.6	35.39%	10.93%	0.469	0.615	14.8%	13.87%
1.7	27.39%	8.64%	0.303	0.589	15.3%	13.91%
1.8	47.11%	14.13%	0.632	0.589	16.0%	13.78%
1.9	39.05%	11.95%	0.473	0.571	16.8%	14.90%
2	31.18%	9.74%	0.333	0.534	17.2%	14.63%

2. 热轧卷板-铁矿石（HC-I）跨品种组合止损阈值遍历回测结果

表 5-5 热轧卷板与铁矿石组合止损阈值遍历回测结果

止损阈值	策略收益	策略年化收益	夏普比率	胜率	策略波动率	最大回撤
1.5	129.53%	32.92%	0.969	0.531	29.8%	33.76%
1.6	124.83%	31.98%	0.914	0.580	30.6%	34.66%
1.7	125.00%	32.01%	0.913	0.595	30.7%	36.31%
1.8	156.11%	38.00%	1.008	0.613	33.7%	32.76%
1.9	90.12%	24.61%	0.540	0.638	38.2%	54.82%
2	90.12%	24.61%	0.540	0.638	38.2%	54.82%

根据表 5-5 的回测结果可以发现，当开仓阈值为 0.55 倍价差标准差、止损阈值为 1.8 倍价差标准差时，热轧卷板与铁矿石组合的策略收益率达到最大为 156.11%，年化收益率为 38.00%，此时最大回撤也显著降低至 32.76%，策略波动率为 33.7%，夏普比率也提高至 1.008。总体而言，策略绩效得到有效提升。

3. 螺纹钢-热轧卷板（RB-HC）跨品种组合止损阈值遍历回测结果

表 5-6 螺纹钢与热轧卷板组合止损阈值遍历回测结果

止损阈值	策略收益	策略年化收益	夏普比率	胜率	策略波动率	最大回撤
1.5	43.72%	13.23%	0.903	0.556	10.2%	11.33%
1.6	34.16%	10.59%	0.584	0.526	11.3%	11.98%
1.7	40.89%	12.46%	0.735	0.527	11.5%	11.74%
1.8	42.04%	12.77%	0.763	0.516	11.5%	11.75%
1.9	42.04%	12.77%	0.763	0.516	11.5%	11.75%
2	38.16%	11.71%	0.667	0.517	11.6%	11.91%

根据表 5-6 的回测结果可以发现，螺纹钢与热轧卷板组合在开仓阈值为 0.35 倍价差标准差、止损阈值为 1.5 倍价差标准差时，策略收益率达到最大为 43.72%，

年化收益率为 13.23%，夏普比率提高至 0.903，最大回撤为 11.33%，策略波动率为 10.2%。总体而言，该策略的绩效得到了提升。

以上三个组合在多个交易信号阈值下回测结果表明，本文所构建的跨品种交易策略能适用于具有协整关系的商品期货组合，但是同一固定参数信号阈值并不能给每个组合都带来最大的盈利机会，策略参数的不同使策略绩效水平也不一样，如果想要获得更好的绩效水平，需要根据期货品种具体情况进行设计。

5.4 本章小结

本章对上一章所设计的三个跨品种套利交易策略进行了回测，再根据回测的结果对策略进行了拓展。

首先是以日频率进行了回测，其中铁矿石与螺纹钢组合的策略收益率为 31.18%，年化收益率为 9.74%，最大回撤为 14.63%；铁矿石与热轧卷板组合的策略收益率为 39.01%，年化收益率为 11.94%，最大回撤为 53.22%；螺纹钢与热轧卷板组合的策略收益率为 23.06%，年化收益率为 7.37%，最大回撤为 7.53%。单从策略收益上看，三个组合表现都相对较好，但在止损方面该策略还有待改进。另外三个组合的收益走势都比较平稳，交易次数偏少，可以通过调整交易阈值进行策略优化。

接着本章在日频率回测的基础上做了阈值优化，将开仓信号阈值以 0.1 倍价差标准差为步长，从 0.35 倍价差标准差到 1.45 倍价差标准差进行遍历，寻找各组合的最佳开仓阈值。再根据回测所得的结果，将止损信号阈值以 0.1 倍价差标准差为步长，从 1.5 倍价差标准差到 2 倍价差标准差进行遍历，以提升各组合的策略绩效。最终发现铁矿石与螺纹钢组合在开仓阈值为 ± 0.75 倍价差标准差、止损阈值为 ± 1.8 倍价差标准差时，策略收益率达到 47.11%，最大回撤为 13.78%；铁矿石与热轧卷板组合在开仓阈值为 ± 0.55 倍价差标准差、止损阈值为 ± 1.8 倍价差标准差时，策略收益率达到 156.11%，最大回撤为 32.76%；螺纹钢与热轧卷板组合在开仓阈值为 ± 0.35 倍价差标准差、止损阈值为 ± 1.5 倍价差标准差时，策略收益率达到 43.72%，最大回撤为 11.33%。总体而言，三个组合的策略绩效均得以有效提升。

第六章 研究结论与展望

6.1 研究结论

本文从铁矿石产业链出发,分析了铁矿石及其下游产品的基本面逻辑关系,并选择铁矿石、螺纹钢、热轧卷板三大黑色系期货为构建跨品种套利组合的对象。然后通过相关性分析、协整检验、误差修正模型证明了三个期货品种两两构建跨品种套利组合的可行性。再基于协整模型和跨品种套利原理,设计了一个商品期货跨品种套利交易策略。最后对所构建的策略进行了日频率的回测,并且通过遍历交易信号阈值的方式优化策略绩效。根据策略的回测结果,本文主要得到以下结论:

第一、在开仓阈值为 ± 0.75 倍价差标准差、止损阈值为 ± 2 倍价差标准差时,铁矿石与螺纹钢组合获得了 31.18%的策略收益率,最大回撤为 14.63%;铁矿石与热轧卷板组合获得了 39.01%的策略收益率,最大回撤为 53.22%;螺纹钢与热轧卷板组合获得了 23.06%的策略收益率,最大回撤为 7.53%。对比来看,三个组合均获得了一定的收益,但其中铁矿石与热轧卷板组合在收益较高的同时也存在较大的风险,铁矿石与螺纹钢组合风险和收益适中,螺纹钢与热轧卷板组合的收益最低,但风险相对也更低。总体而言,本文基于协整模型所构建的跨品种套利策略有一定的适用性。

第二、在对交易信号阈值进行优化后,发现三个组合分别在不同的交易信号阈值下获得了更佳策略绩效,其中铁矿石与螺纹钢组合在开仓阈值为 ± 0.75 倍价差标准差、止损阈值为 ± 1.8 倍价差标准差时,策略收益率达到 47.11%,最大回撤为 13.78%;铁矿石与热轧卷板组合在开仓阈值为 ± 0.55 倍价差标准差、止损阈值为 ± 1.8 倍价差标准差时,策略收益率达到 156.11%,最大回撤为 32.76%;螺纹钢与热轧卷板组合在开仓阈值为 ± 0.35 倍价差标准差、止损阈值为 ± 1.5 倍价差标准差时,策略收益率达到 43.72%,最大回撤为 11.33%。由此可见,同一固定参数信号阈值并不能给每个组合都带来最大的盈利机会。在策略设计的过程中,需要结合市场行情和交易品种的特点,调整策略所用模型和交易信号,以优化交易策略提高绩效水平。

6.2 研究展望

虽然本文所设计的策略在回测阶段获得了比较好的效果,但是由于笔者能力有限,仍有许多地方还需要进一步完善:

第一、本文在选择标的时虽然采用了基本面分析方法，但仅仅是定性地分析品种间的基本面关系逻辑以选择相关性较高的期货品种，没有将其运用到策略设计中，后续可以考虑在策略设计中加入基本面指标进行策略优化。

第二、在策略设计上本文选择以固定参数阈值为交易信号，虽然在策略回测时也进行了参数阈值优化，但仅仅是采用以固定步长对交易信号进行遍历的方式选择最优值。虽然最优阈值会因为品种不同存在差异，但是这种遍历方式存在很大的局限性。另外固定参数法认为价差序列是符合正态分布的，但通常来说价差序列也可能不是方差齐性的，后续可以引入 GRACH 等模型，对回归方程的参数估计做进一步研究以选择交易阈值。

第三、本文在回测时虽然发现以日频率回测的策略具有较好的适用性，但也存在不足之处，比如策略的交易次数偏少会使资金长期处于闲置状态，这容易使交易者丧失信心而收回资金。后续可以选择以期货合约的日内 1 分钟、5 分钟、15 分钟等高频数据作为研究样本，对高频数据下的期货组合协整关系进行实证分析，并设计相关高频交易策略。

致 谢

行文至此，终于到了提笔写致谢的时刻，这一刻我的内心百感交集，时光总是过的很快，研究生时期马上就要结束，而上一次的毕业季仿佛还停留在我眼前。来到电子科技大学学习是我求学生涯中很重要的决定，在这里我获得了更好的平台，也认识了更多优秀的人。回首过往两年时光，我收获的帮助有很多，有太多的人需要感谢。

感谢我的导师——李平老师，李老师在刚成为我的指导老师时，就给我推荐了许多文献以及学习资料并且为我的论文选题提供建议，正是他认真负责的态度，让我在写论文之前有了充足的前期准备工作，从而能够顺利地完成本论文的研究与写作。同时我也很感谢李老师在研究生期间对我的提点以及教诲，他所给予我的建议将使我受益一生。

感谢我的室友和朋友们，在我写论文遇到问题时和我一起讨论，帮助我调整逻辑走出困境；在我心绪烦乱时给我建议以及安慰，让我能够及时调整心态。来到成电求学是我二十三年以来离家最远的一次，两年的时光里对这个城市总会有陌生的时刻，幸而有你们一直以来的陪伴，能在成电能与你们相识，我何其幸运。

感谢我的家人，在我做任何决定时都无条件的支持我，一直以来都给我提供最好的环境以及最大的帮助，让我在无数次迷茫时能够找到方向。长路漫漫，毕业也是个开始，在未来我会更加努力工作与生活，不负你们的期待。

太多的感谢用言语来表述依旧显得很苍白，最后就还是希望所有人都能平安喜乐、所愿皆所得！

参考文献

- [1] Markowitz H M. Portfolio selection[J]. Journal of Finance, 1952, 7(1): 77-91.
- [2] Sharpe W F. Capital asset prices: A theory of market equilibrium under conditions of risk[J]. Journal of Finance, 1964, 19(3): 425-442.
- [3] Malkiel B G, Fama E F. Efficient capital markets: A review of theory and empirical work[J]. Journal of Finance, 1970, 25(2): 383-417.
- [4] Ross S A. The arbitrage theory of capital asset pricing[J]. Journal of Economic Theory, 1976, 13(3): 341-360.
- [5] 丁鹏. 量化投资[M]. 北京: 电子工业出版社, 2012, 2-4.
- [6] 黄道舜. 程序化交易模型在中国期货市场的应用[J]. 中国外资(下半月), 2011, (10): 218+220.
- [7] 李达捷. 股指期货高频交易的实证研究[J]. 经济视角(下), 2013, (1): 53-55.
- [8] 陈艳, 王宣承. 基于变量选择和遗传网络规划的期货高频交易策略研究[J]. 中国管理科学, 2015, 23(10): 47-56.
- [9] 李志鸿. 国内商品期货“短周期”量化投资策略研究[D]. 浙江大学, 2017.
- [10] 李雁, 刘金山, 杨镇纲, 等. 商品期货量化交易策略[J]. 中国商论, 2019, (01): 54-57.
- [11] 周志中, 俞祖卿. 基于信息融合和策略转换的商品期货量化投资策略[J]. 系统管理学报, 2021, 30(2): 253-263.
- [12] Carol A, Anca D. Indexing and statistical arbitrage[J]. Journal of Portfolio Management, 2005, 31(2): 50-63.
- [13] Huck N, Afawubo K. Pairs trading and selection methods: Is cointegration superior?[J]. Applied Economics, 2015, 47(6): 599-613.
- [14] Rad H, Low R K Y, Faff R. The profitability of pairs trading strategies: Distance, cointegration and copula methods[J]. Quantitative Finance, 2016, 16(10): 1541-1558.
- [15] Vidyamurthy G. Pairs Trading: Quantitative Methods and Analysis[M]. Hoboken, New Jersey: John Wiley & Sons, Inc, 2004.
- [16] 王粹萃. 基于协整方法的统计套利策略实证检验[D]. 吉林大学, 2007.
- [17] 于玮婷. 基于协整方法的统计套利策略的实证分析[J]. 科学决策, 2011, (3): 70-85.
- [18] 陈叠, 王秋香, 华夏, 等. 基于协整的统计套利策略优化与实证研究[J]. 时代金融, 2016, (15): 335-336+342.

- [19] 吴员福. 融资融券标的证券配对交易策略研究——基于协整理论与 GARCH 模型[J]. 经贸实践, 2018, (11): 150.
- [20] 仇中群, 程希骏. 基于协整的股指期货跨期套利策略模型[J]. 系统工程, 2008, 26(12): 26-29.
- [21] 常宗琪. 白糖统计套利理论模式研究及实例分析[J]. 经济师, 2008, (11): 30-31.
- [22] 何树红, 张月秋, 张文. 基于 GARCH 模型的股指期货协整跨期套利实证研究[J]. 数学的实践与认识, 2013, 43(20): 274-279.
- [23] 张波, 刘晓倩. 基于 EGARCH-M 模型的沪深 300 股指期货跨期套利研究——一种修正的协整关系[J]. 统计与信息论坛, 2017, 32(04): 34-40.
- [24] 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, (2): 55-59.
- [25] 扈文秀, 牛静, 李芳, 等. 基于统计套利模型的商品指数期货双跨套利方案研究[J]. 管理评论, 2013, 25(9): 100-107.
- [26] 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, (4): 112-115.
- [27] 方军, 李星野. 基于协整理论的 DFT-GARCH 模型的统计套利研究[J]. 经济数学, 2019, 36(2): 57-62.
- [28] Liu Q F. Price relations among hog, corn, and soybean meal futures[J]. Journal of Futures Markets, 2005, 25(5): 491-514.
- [29] Dunis C L, Laws J, Evans B. Trading futures spreads: An application of correlation and threshold filters[J]. Applied Financial Economics, 2006, 16(12): 903-914.
- [30] Mitchell J B. Soybean futures crush spread arbitrage: Trading strategies and market efficiency[J]. Social Science Electronic Publishing, 2010, 3(1): 63-96.
- [31] Kanamura T, Rachev S T, Fabozzi F J. A profit model for spread trading with an application to energy futures[J]. Journal of Trading, 2010, 5(1): 48-62.
- [32] 张伟, 景鸿. 基于协整关系的统计套利研究——黄金白银期货的跨品种套利研究[J]. 现代经济信息, 2013, (20): 248-250.
- [33] 黄腾. 期货跨品种套利策略[J]. 现代经济信息, 2019, (21): 271-272.
- [34] 张静文. 商品期货跨品种套利策略的实证研究——以铁矿石、焦炭为例[J]. 现代商业, 2020, (28): 90-92.
- [35] 梁仁方. 基于 BP 神经网络模型和 NAR 动态神经网络模型的期货跨品种套利策略对比研究[D]. 浙江财经大学, 2016.

- [36] 孙建明, 张伟楠, 张华. 玉米与淀粉跨产品套利研究——基于价格协整关系的讨论[J]. 价格理论与实践, 2017, (7): 137-140.
- [37] 刘一诺. 基于协整的焦炭与焦煤期货跨品种套利实证研究[J]. 现代商业, 2020, (24): 85-86.
- [38] 黄巍华, 王伟. 农产品期货波动溢出效应及其跨品种套利研究——基于大豆、豆油和豆粕期货价格相关性的分析[J]. 价格理论与实践, 2021, (8): 1-4.
- [39] Engle R F, Granger C. Cointegration and error-correction: Representation, estimation and testing[J]. Econometrica, 1987, 55(2): 251-276.
- [40] 艾云辉, 郭文娟. 关于协整与误差修正模型的建立[J]. 中国经济与管理科学, 2009, (1): 7-8.

附 录

跨品种套利策略核心代码

```
# 导入函数库
import statsmodels.api as sm
from statsmodels import regression
import numpy as np
import pandas as pd
import time
from datetime import date
from jqdata import *
import datetime
from dateutil.relativedelta import relativedelta
from statsmodels.tsa.stattools import adfuller
# 初始化函数，设定基准等等
def initialize(context):
    # 设定跨品种组合中的某个作为基准
    set_benchmark('I9999.XDCE')
    # 开启动态复权模式(真实价格)
    set_option('use_real_price', True)
    # 设定账户为金融账户
    set_subportfolios([SubPortfolioConfig(cash=context.portfolio.starting_cash,
type='index_futures')])
    #交易时的手续费是：买入时万分之 1,卖出时万分之 1
    set_order_cost(OrderCost(open_commission=0.0001,
close_commission=0.0001,close_today_commission=0.0001), type='index_futures')
    # 设定保证金比例
    set_option('futures_margin_rate',0.12)
    #设置滑点
    set_slippage(FixedSlippage(0))
    g.bsFlag_A,g.bsFlag_B = False,False
    g.bsFlag = True
```

```

g.cash = context.portfolio.total_value
run_daily(market_open,time='open', reference_security='I9999.XDCE')

#开盘时运行
def market_open(context):
    #设置时间窗口
    g.gap = 60
    #获取合约
    f1= get_dominant_future('RB', date=context.current_dt)
    f2 = get_dominant_future('I', date=context.current_dt)
    priceX = history(g.gap, '1d', 'close', f1)
    priceY = history(g.gap, '1d', 'close', f2)
    priceX_lg = np.log(priceX [f1])
    priceY_lg = np.log(priceY [f2])
    x1=priceX_lg.values
    y1=priceY_lg.values
    x=np.array(x1)
    X=sm.add_constant(x)
    y=np.array(y1)
    results = sm.OLS(y, X).fit()
    resid=y-(results.params[0]+results.params[1]*x)
    alpha=results.params[0]
    beta=results.params[1]
    #记录残差项
    resid1=pd.Series(resid)
    resid1.dropna(inplace=True)
    sigma=resid.std()
    b=0.75*sigma
    c=2*sigma
    if len(resid1)>0:
        #对残差项进行 ADF 检验，并记录检验 p 值
        p_value=testStationarity(resid1)['p-value']
        if(p_value<0.05):

```

```

f_01=attribute_history(f1,1,'1d',fields='close')
f_02=attribute_history(f2,1,'1d',fields='close')
m=f_01.close.values
n=f_02.close.values
lot_Y=math.ceil(0.2*context.portfolio.available_cash/(n*100*0.12))
lot_X=math.ceil(lot_Y*beta)
f_1=np.log(f_01)
f_2=np.log(f_02)
#代入之前回归得到的方程计算新残差 resid_2
mean=np.mean(resid)
resid_2=np.array(f_2)-results.params[0]-results.params[1]*np.array(f_1)-mean
if g.bsFlag:
    if resid_2>b and resid_2<c:
        #记录开仓 买 X 卖 Y
        g.bsFlag = False
        #记录开仓方向
        g.bsFlag_A = True
        #记录交易合约以防止主力合约变动
        g.tradecode_A,g.tradecode_B = f1,f2
        #下单
        order(f1, lot_X, side = 'long', pindex = 0)
        print('多单:%s'%f1)
        order(f2, lot_Y, side = 'short', pindex = 0)
        print('空单:%s'%f2)
        g.f2 = f2
        g.f1 = f1
    elif (resid_2<-b and resid_2>-c):
        #记录开仓 卖 Y 卖 X
        g.bsFlag_B = True
        #记录开仓方向
        g.bsFlag = False
        #记录交易合约以防止主力合约变动
        g.tradecode_A,g.tradecode_B = f1,f2

```

```

#下单
order(f1, lot_X, side = 'short', pindex = 0)
order(f2, lot_Y, side = 'long', pindex = 0)
print('空单:%s'%f1)
print('多单:%s'%f2)
g.f2 = f2
g.f1 = f1

else:
    if(abs(resid_2)<b or abs(resid_2)>c):
        if g.bsFlag_A:
            order_target_value(g.tradecode_A, 0,side = 'long', pindex = 0)
            order_target_value(g.tradecode_B, 0, side = 'short', pindex = 0)
            print('止损平仓')
            g.bsFlag_A = False
            g.bsFlag = True
        if g.bsFlag_B:
            order_target_value(g.tradecode_B, 0, side = 'long', pindex = 0)
            order_target_value(g.tradecode_A, 0, side = 'short', pindex = 0)
            print('止损平仓')
            g.bsFlag_B = False
            g.bsFlag = True

#ADF 检验
def testStationarity(ts):
    dftest = adfuller(ts)
    dfoutput = pd.Series(dftest[0:4], index=['Test Statistic','p-value','#Lags Used','Number of
Observations Used'])
    for key,value in dftest[4].items():
        dfoutput['Critical Value (%s)%key'] = value
    return dfoutput

#强制移仓
def move_month(context,pindex=0):

```

```

long_p = context.subportfolios[pindex].long_positions
short_p = context.subportfolios[pindex].short_positions
need_change={'long': {}, 'short': {}}
for future in long_p:
    is_last_day = get_security_info(future).end_date==context.current_dt.date()
    today_dominant = get_dominant_future(future[:-9])
    if get_security_info(today_dominant).end_date==context.current_dt.date():
        today_dominant = get_future_contracts(today_dominant[:-9:])[1]
        log.error( '%s 遇到结算日,提前切换主力'%future)
    if future!=today_dominant:
        try:
            order_ = order_target(future,0,side='long',pindex=pindex)
            order(today_dominant,order_.amount,side='long',pindex=pindex)
        except:
            log.error('标的 %s 切换 %s 失败'%(future,today_dominant))
            need_change['long'][future] = today_dominant
for future in short_p:
    is_last_day = get_security_info(future).end_date==context.current_dt.date()
    today_dominant = get_dominant_future(future[:-9])
    if get_security_info(today_dominant).end_date==context.current_dt.date():
        today_dominant = get_future_contracts(today_dominant[:-9:])[1]
        log.error( '%s 遇到结算日,提前切换主力'%future)
    if future!=today_dominant:
        try:
            order_ = order_target(future,0,side='short',pindex=pindex)
            order(today_dominant,order_.amount,side='short',pindex=pindex)
        except:
            log.error('标的 %s 切换 %s 失败'%(future,today_dominant))
            need_change['short'][future] = today_dominant
return need_change

```