

分 类 号 _____

学号 M201977969

学校代码 10487

密级 _____

华中科技大学

硕士学位论文

基于卡尔曼滤波及 BP 神经网络的商 品期货量化策略研究

学 位 申 请 人： 梁家林

学 科 专 业： 金融

指 导 教 师： 张学功

答 辩 日 期： 2021 年 5 月 14 日

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

**Study on Commodity Futures Quantitative
Strategy Based on Kalman Filter and BP Neural
Network**

Candidate : Jialin Liang

Major : Finance

Supervisor : Xuegong Zhang

Huazhong University of Science & Technology

Wuhan 430074, P.R.China

May, 2021

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：梁家林

日期：2021年5月20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：梁家林

日期：2021年5月20日

指导教师签名：张学功

日期：2021年5月20日

摘 要

商品期货出现以后,一直是生产企业、投资者和政府管理部门的关注重点。商品期货不仅有着价格发现的作用,而且还具有风险投资的功能,吸引着投资者进入市场。目前在商品期货市场的相关研究中,有一个重要的方向就是通过预测商品期货价格,进而构建出有效的量化策略,帮助投资者更好地规避风险,获得超额收益。

本文选取了商品期货中有代表性的豆粕、沥青、白银和 PVC 期货作为主要研究对象。在研究 BP 神经网络和卡尔曼滤波的原理及特点之后,基于卡尔曼滤波改进 BP 神经网络,并将改进模型应用到商品期货的价格预测中。通过商品期货的历史数据,使用卡尔曼滤波—BP 神经网络模型对其进行价格预测,并与其他预测模型比较。对比不同预测模型的预测结果和误差,发现商品期货的价格时间序列含有较多噪音, BP 神经网络模型对价格预测不够精确,预测价格的波动性大,而卡尔曼滤波—BP 神经网络模型可以有效降低输入数据中的噪音,训练效果更好,能够更准确地预测价格,预测价格曲线较平稳。最后基于卡尔曼滤波—BP 神经网络模型建立跨期套利策略,验证其在实际应用中的效果,并同传统基于协整的跨期套利策略相对比,发现基于神经网络的跨期套利策略机会更多,收益率提高。预测模型能在一定程度上指导投资者构建量化策略。

本文探讨了基于扩展卡尔曼滤波改进 BP 神经网络并将改进后的模型用于预测商品期货价格的可行性和有效性,并基于改进后神经网络模型构建量化投资策略,为投资者提供一定的参考意见。但是本文还存在一些缺陷,如商品期货样本的种类较少、时间跨度短。只有不断改进模型,才能更好地面对商品期货市场上复杂影响因素带来的挑战,更好地为投资者提供帮助。

关键词: 量化策略; BP 神经网络; 卡尔曼滤波; 价格预测; 商品期货

Abstract

Since commodity futures appeared, it has been the focus of attention of production enterprises, investors and government management departments. Commodity futures not only have the function of price discovery, but also have the function of venture capital, attracting investors to enter the market. At present, in the relevant research of commodity futures market, an important direction is to construct an effective quantitative strategy by predicting commodity futures prices, so as to help investors better avoid risks and obtain excess returns.

In this paper, soybean meal, asphalt, silver and PVC futures are selected as the main research objects. After studying the principle and characteristics of BP neural network and Kalman filter, the BP neural network is improved based on Kalman filter, and the improved model is applied to commodity futures price prediction. Based on the historical data of commodity futures, the Kalman filter-BP neural network model is used to forecast its price, and compared with other forecasting models. Comparing the prediction results of different forecasting model and error of commodity futures price time series were found to contain more noise, the BP neural network model to predict the price is not accurate, forecast the price volatility is big, and extended kalman filter, the BP neural network model can effectively reduce the noise in the input data, the better the training effect, can more accurately predict the price, forecast price curve is smooth. Finally, an intertemporal arbitrage strategy is established based on the Kalman Filter-BP neural network model to verify its effect in practical application. Compared with the traditional intertemporal arbitrage strategy based on co-integration, it is found that the intertemporal arbitrage strategy based on neural network has more opportunities and higher returns. Forecast model can guide investors to build quantitative strategy to some extent.

This paper discusses the feasibility and effectiveness of using the improved BP neural network based on extended Kalman filter to predict

commodity futures prices, and builds quantitative investment strategy based on the improved BP neural network model to provide some reference opinions for investors. However, there are still some defects in this paper, such as a small variety of commodity futures samples and a short time span. Only by constantly improving the model can we better face the challenges brought by the complex influencing factors in the commodity futures market and provide better help for investors.

Key words: Quantitative Strategy; BP Neural Network; Kalman Filter; Price Forecast; Commodity Futures

目 录

摘 要.....	I
Abstract	II
目 录.....	IV
1 绪论.....	1
1.1 研究背景与意义	1
1.2 国内外研究现状	3
1.3 本文的研究方法和结构安排	6
1.4 本文主要创新点	7
2 相关理论基础.....	8
2.1 人工神经网络	8
2.2 卡尔曼滤波理论基础	14
2.3 卡尔曼滤波—BP 神经网络组合模型	17
2.4 传统商品期货价格预测模型	19
3 期货价格预测实证研究.....	22
3.1 数据的选择与预处理	22
3.2 商品期货价格预测模型的构建与训练	25
3.3 预测结果分析	29
4 量化投资策略.....	33
4.1 跨期套利策略的理论基础	33
4.2 标准距离法跨期套利	34
4.3 卡尔曼滤波—BP 神经网络模型跨期套利	38
4.4 结论.....	39
5 总结与展望	41
5.1 全文研究总结	41

5.2	存在问题.....	41
5.3	未来工作展望	42
	致谢	43
	参考文献	44

1 绪论

1.1 研究背景与意义

随着经济全球化的进步,金融市场在不断发展,金融市场在发展中带来了新金融产品的出现。最早的金融衍生品是期货合约,由芝加哥交易所于 1865 年推出。期货合约是一种标准化协议,规定了在未来一段时间,买卖双方交易某种标的物时,以合约中双方约定的价格为交易价格。期货合约的出现,为交易者提供了套期保值、规避交易风险的方法。在我国的期货市场中,最早出现的期货合约是大豆、玉米和豆粕为标的物的商品期货,目前我国期货市场交易主要种类有农副产品、化工产品、金属产品和能源产品等。

商品期货合约出现的原因是防范商品价格风险,随后发现了商品期货也能用于投机套利。改革开放时,我国经济体制从计划经济转向了市场经济,国民经济发生重大变革,这为期货合约的出现奠定了基础。农村的经济体制改革,将原有落后的生产体制转变为以联产承包责任制为基础的生产体制,大大解放了农村生存力,调动了人民的生产积极性,全国粮食产量大幅度增加。虽然市场经济的出现放开了交易价格,降低了交易壁垒,但受到当时交易信息和商品流通的限制,买者找不到卖家,卖者找不到买家。市场达成交易的效率低,不同地区交易价格也不相同。为了规范商品市场,稳定价格,一部分学者提出建立粮食批发市场来解决问题。1990 年,中国郑州粮食批发市场经批准建立。郑州粮食批发市场作为全国第一个国家级的农产品交易所,为我国粮食市场引入了期货机制,规范化交易粮食,减少了买卖双方的沟通成本,降低了粮食市场的价格波动。随后,我国其他地方也开始筹备建立期货交易所,推动了国内期货市场的发展,为进一步降低信息壁垒,维持价格稳定做出贡献。

随着国内期货市场的数量和功能不断增加,商品期货可以更加有效地反映供求关系的变化,防范价格风险。由于期货市场本身所具有的风险投资功能,大批投资者被吸引进入期货市场。期货市场成为经营者、投资者和政府管理部门关注的焦点。如何

有效地预测期货价格,分析期货价格变动的规律,把握期货市场的动态,对于经营者防范市场风险、投资者制定正确的投资策略和管理部门宏观调控商品市场,有着非常重大的理论意义和现实意义。

由于市场交易会受外部环境因素的影响,如国家政策法规发生变动、不可预测的自然灾害等,商品期货价格序列具有高度的非平稳性和异方差性,简单统计和传统技术分析的价格预测方法存在滞后性且预测效果不佳。一些研究者表明期货价格是混沌序列,难以长期预测。研究者开始寻找更好的预测模型。其中,人工神经网络不需要依赖严格的假设和前提条件,可以通过对已有的数据进行处理,挖掘出潜在信息,能很好地处理复杂系统的非线性关系。将人工神经网络应用于期货价格研究成为一个研究方向。并且期货市场信息存在着高噪音、期货价格非线性特征等问题。卡尔曼滤波算法在工业和通信领域是常用的滤波算法,能够降低信息噪音。通过将卡尔曼滤波算法和神经网络结合,可以提高预测期货未来价格的准确度,具有一定的学术意义和可操作性。

投资者在期货市场中投资根据其交易策略可分为两种:一种投资者主要进行价值投资,依据基本面信息,通过价值上涨获利,一般是长期投资;另一种投资者基于计算机程序算法和数据的进行量化投资,从交易量、价格和技术指标等挖掘潜在信息,构建组合获利,一般是短期投资。尽管我国商品期货市场不断发展,但是很多学者表明我国商品期货市场仍是弱有效市场。通过量化投资能够捕抓潜在盈利机会从而获得超额收益。量化投资理论主要由国外学者提出并不断发展,近些年来我国一些基金公司也开始采用量化投资策略,且部分量化策略已经进行了实盘交易。

本文基于前人的研究,结合卡尔曼滤波改进BP神经网络并应用于商品期货市场,证明其在价格预测方面的有效性和可行性。再通过预测模型建立量化交易策略,制定量化交易组合,以获得更高的收益。以此说明价格预测模型可为投资者提供量化交易策略指导,帮助投资者更好地获利,具有一定的实际意义。

1.2 国内外研究现状

1.2.1 期货价格预测相关研究

针对商品期货市场,国内外研究者使用了不同方法建立理论模型,预测期货价格的未来走势和波动情况。最先出现的预测方法是基本分析法和技术分析法。前者根据商品期货市场的供求关系和相关因素来分析商品期货市场的价格未来走势,后者通过价格、涨跌幅和成交量等技术指标,借助技术图形分析商品期货市场的价格。基本分析法需要商品期货的全面信息,侧重于对商品期货的定性分析,适合预测长期价格变动。而技术分析因为人为因素较多,适合预测短期价格变动。王锐(2011)通过协整理论和向量自回归模型,研究了对我国农产品价格波动的影响因素。刘庆富和张金清(2010)通过贝叶斯 MCMC 模拟技术对我国大豆和小麦市场进行了实证分析,研究了我国商品期货隔夜信息对日间交易的预测能力。吴玉霞和温欣(2016)通过 ARIMA 模型对股票价格进行预测。Garcia(2007)通过同步估计方法,基于每日价格的限价波动对期货价格进行预测。Karali(2013)通过 GARCH 模型,解释美国农业和金属期货市场上的价格波动的影响因素。Clark 等(1996)做了早期的期货价格研究,发现期货价格同市场成交量之间,存在着正相关关系。王洪伟(2001)对商品期货同价格的关系进行了实证研究。华仁海(2002)研究国内期货市场上存在的日历效应。张方杰和胡燕京(2005)研究了影响期货价格预测的因素以及存在的问题。王习涛(2005)从期货价格序列的非平稳性着手,通过 ARMA 模型,成功让数据实现部分平稳化,更准确地预测期货价格的走势。刘璐等(2018)分析市场上实际需求和投机因素对价格的影响,结论是从短期来看是投机对价格影响大,从长期来看是实际需求对价格的影响更大。田利辉等(2015)通过分析 2002 年至 2012 年原油价格波动,发现金融市场中的投机行为会使价格剧烈波动。

1.2.2 神经网络价格预测相关研究

基于 ARMA 模型和 GARCH 模型的传统预测模型一般是线性预测模型,预测能力有限,且原始数据需要满足平稳性假设。与线性预测模型不同,非线性预测模型有更好的预测能力,能更好地挖掘变量之间的非线性关系。常用的非线性预测模型有支持向

量机、人工神经网络和随机森林等，其中以机器学习方法为代表。Robert（1997）提出期货市场的动力学特点，使得研究者开始使用神经网络预测期货价格。

人工神经网络在上个世纪 40 年代开始出现。人工神经网络模型拥有自组织、自适应、自学习的优势，在处理相互影响且有交叉效应和动态效应的变量间关系有独特的作用，于是人工神经网络模型在金融领域开始大放异彩。韩士专等（2012）利用 BP 神经网络模型，预测了玉米期货的价格走势，模型的预测精度高于其他模型。杨建辉（2011）年利用改进后的支持向量回归模型对期货价格进行预测，预测结果表明，支持向量回归模型适用于我国股票权证的价格预测。Avraam（2018）利用高频交易数据，通过卷积神经网络预测了股票价格。Pereira（2016）通过前馈神经网络，对原油期货合约的价格变化趋势进行预测。杨海民（2019）发现基于机器学习的预测方法虽然没有传统线性预测模型的问题，但是其对样本的要求较高。杨青（2019）基于 LSTM 神经网络模型，对全球股票指数进行预测，发现 LSTM 神经网络模型有很强的泛化能力，相比传统线性预测模型，预测精度在不同期限上均有提升。吴秋芳等（2013）将非线性预测模型和线性预测模型结合，通过 GARCH 类模型和 BP 神经网络模型对股市的量价关系进行了实证研究，发现非预期成交量与股市波动存在较明显的正相关关系。张炜（2003）运用属性聚类网络和属性识别理论解决模式识别问题，在期货价格趋势预测中获得成功的应用。Huchison 运用神经网络模型对欧式期权进行定价计算，通过对标准普尔 500 期货数据的实证，发现人工神经网络模型的定价结果优于 Black-Scholes 模型。张鸿彦和林辉等建立了混合小波神经网络模型并在香港市场的实证中发现混合小波神经网络模型要优于传统的 Black-Scholes 模型。孙红冰（2017）通过 BP 神经网络校正灰色 GARCH 模型预测残差实现股票价格的高精度预测。

1.2.3 卡尔曼滤波研究现状

卡尔曼滤波是一个最优化自回归数据处理方法。卡尔曼滤波可以通过状态方程和观测方程组合成一个线性的随机系统，进而构建出一个状态空间。卡尔曼滤波算法按照线性无偏最小均方误差估计准则，状态方程传递信息，新观测数据校准偏差，可以最优估计系统状态变量。卡尔曼滤波是目前使用最多的滤波算法之一，被广泛应用于

导航、控制和航天等领域。近年来卡尔曼滤波在经济和金融领域的应用开始增加。曹广福和路群（1999）利用卡尔曼滤波成功在股利贴现模型上将随机过程中的不确定性减弱，从而使现金流按目标规律变动。池丽旭等（2012）通过扩展卡尔曼滤波，结合了主成分分析方法，在中国股票市场成功获得情绪状态指标，发现投资者情绪的变动较投资者情绪能更好的预测股票价格的走势，获得更好的收益。彭继兵和唐春艳（2005）利用卡尔曼滤波对股票价格进行预测，发现基于变维分形理论的卡尔曼滤波能够有效预测股票价格。

1.2.4 量化投资研究现状

量化投资是建立在有效市场假说上的，被证明是弱有效市场中最好的分析理论。Olszewski（2001）通过多种趋势因子预测价格，成功将买卖信号量化，基于此构建了交易策略。国内的量化投资策略开展较晚，只进行了基本的理论研究。丁鹏（2012）认为量化投资是以计算机计算能力大幅提高为前提，通过大量的历史数据和理论模型，研究股票等金融产品的价格以图获得收益。李达捷（2013）利用较少人使用的高频信息，采用传统趋势理论，构建短线策略获利，说明我国金融市场仍是弱有效市场，量化策略适合我国金融市场。彭乐（2014）使用技术指标构建螺纹钢期货的量化交易策略。周星月（2014）基于布林线发现一个新的趋势指标，并在股指期货市场上证明该指标能应用于构建量化策略并获利。刘宗兴（2020）采用遗传算法对布林通道量化投资策略进行优化，发现优化改进后的策略具有更高的收益率和稳定性。黄宏远等（2017）利用 GA-SVM 模型解决了量化择时策略中存在的信号判断不准的问题。国内的量化交易不能快速发展的一大问题是国外已有的量化模型不一定适合国内市场，研究者和投资者需要对这些模型进行测试和适应调整。

1.2.5 简要评述

传统线性预测模型研究出许多影响商品期货价格的因素，如成交量和现货价格等。但由于传统线性预测模型对原始数据要求较高，限制了其预测能力的提高。研究者开始采用神经网络进行价格预测，并证明其预测能力优于传统线性预测模型。卡尔曼滤波是工业领域最常用的滤波算法之一，但在金融领域应用较少。量化投资在国外已经

广泛使用,而在中国使用较少。通过将卡尔曼滤波和神经网络结合能用于构建量化策略,给予投资者一定的投资建议。

1.3 本文的研究方法和结构安排

1.3.1 研究方法

本文采用比较分析法和实证研究法相结合,选取商品期货历史价格数据,通过比较改进后 BP 神经网络和其他预测模型对商品期货价格的预测能力,确定卡尔曼滤波能有效减少样本中的噪音,提高 BP 神经网络预测能力。然后基于改进后的 BP 神经网络模型构建量化投资策略,证明其能提高现有的量化策略的收益率,能指导投资者构建新的量化模型。

其次在定量分析中采用单位根检验和协整检验,证明豆粕合约 M2105.DCE 和 M2109.DCE 之间存在协整关系,能够构建跨期套利策略。

1.3.2 结构安排

BP 神经网络是一种常用的神经网络,被广泛应用于各个领域。本文基于卡尔曼滤波改进 BP 神经网络模型,选取豆粕、沥青、白银和 PVC 期货的主力合约作为研究对象,对期货合约的价格进行预测,证明卡尔曼滤波能改善 BP 神经网络的预测能力。然后根据价格预测模型提出量化投资策略,对原有的量化策略进行改良。文章的主要结构如下:

第一章绪论,主要介绍了本文的研究背景和研究意义、国内外关于期货价格预测、BP 神经网络、卡尔曼滤波和量化投资的研究现状。同时还叙述了本文的研究框架、方法和主要创新点。

第二章介绍基础理论,主要是人工神经网络、卡尔曼滤波和传统商品期货价格预测模型。同时构建卡尔曼滤波—BP 神经网络模型。

第三章选取四种商品期货作为样本,运用构建好的卡尔曼滤波—BP 神经网络模型,对商品期货价格预测进行实证研究,证明 BP 神经网络能有效预测商品期货的价格,且卡尔曼滤波能提高 BP 神经网络的预测能力,改进后的 BP 神经网络有更强的预测能力。

第四章基于卡尔曼滤波—BP 神经网络模型构建豆粕期货的跨期套利策略,研究其在实际应用的效果,并和传统基于协整的跨期套利模型做比较,证明价格预测模型能在一定程度上对量化投资策略提供建议。

第五章结论，总结全文的内容，提出文章存在的缺点以及对未来的研究方向做出展望。

1.4 本文主要创新点

在研究方法上，基于卡尔曼滤波改进 BP 神经网络算法主要应用于工程领域，在金融领域中应用较少。在研究对象上，国内对期货价格预测研究更多是选定股指期货作为研究对象，较少预测商品期货的价格，本文选取豆粕、沥青、白银和 PVC 期货主力合约作为研究对象，预测其合约价格。构建量化策略主要集中于黄金期货和股指期货等，本文采用豆粕期货作为研究对象构建量化策略。从实证模型上，本文将基于改进后的价格预测模型构建跨期套利策略，并和原有跨期套利比较，可以清晰地看出不同量化策略的优劣，帮助量化投资者更好地判断交易信号。

2 相关理论基础

2.1 人工神经网络

人工神经网络（Artificial Neural Networks）是一种模仿人类大脑学习和处理信息方式的模型。它通过学习生物神经细胞中的突触结构，可以实现分布式并行信息处理的功能。人工神经网络由一个个基础神经元组成，基础神经元的功能十分有限，但是当多个神经元共同构成神经网络时，神经网络就能依据自身的系统算法，调整不同神经元的连接关系，有效地处理复杂的信息。人工神经网络的结构决定了它的信息储存方式：不依靠神经网络的某个局部来存储信息，而是由所有节点及节点之间的连接来储存信息。信息的分散式储存使得神经网络拥有良好的容错能力，即使有一部分神经元发生不可逆地损坏，也不会因此导致人工神经网络的整体性能下降过多，并且在遇到问题信息时，如信息不完整，人工神经网络可以通过神经元之间的连接，将不完整信息修复，从而达到对不完整信息的正确处理。人工神经网络调整不同神经元的连接关系，可以逐渐构建出满足期望输出的结构，这就是人工神经网络的自学习自组织能力。

2.1.1 人工神经网络的拓扑结构及学习方式

人工神经网络学习人脑结构的，这造成了它的结构特点：由大量神经元相互连接构建成神经网络。对于神经网络来说，每一个神经元是一个计算节点，每一个计算节点一般用一种输出函数。这种输出函数称为激励函数。本文将使用 S 型函数作为激励函数，S 型函数的表达式为： $f(x) = \frac{1}{1+e^{-ax}}$ 。函数曲线如图 2-1 所示。

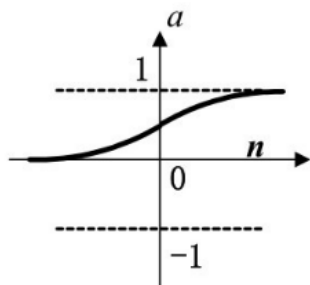


图 2-1 对数 S 型函数曲线

在神经网络中，每个神经元都同其他神经元连接。这种连接代表通过该连接信号的加权值，称为连接权重，人工神经网络的记忆就是通过这些权重实现。根据算法调整两个神经元的连接权重，这就是人工神经网络对输入的信息进行学习和储存方式。调整权重的过程就是神经网络的学习过程。

人工神经网络是对人脑的模仿，两者的区别在于人脑中有着数亿个生物神经元，而神经网络中神经元数量一般达不到这么大的量级，出于物理上的局限性和成本考虑，人工神经网络一般选择较少的神经元构成较简单的结构。目前已有近 40 种神经网络模型，如反传网络、感知器、向量机和自组织映射等。根据神经元连接的拓扑结构，可人工神经网络区分开，一种是前向网络，另一种是反馈网络。

在前向网络中，每级神经元接受上一级神经元的信息输入，然后输入到下一级神经元。神经网络中没有反馈，信息顺流而下，单向传播。前向网络可以实现输入信号从输入空间到输出空间的转化，它对输入信息的处理能力来源于简单的非线性函数多次复合。多层前向神经网络如图 2-2 所示，从图中可以看出，该神经网络由输入层、隐含层和输出层构成。多层前向神经网络中，神经元的传递函数有多种选择，当所有的神经元的传递函数都是符号函数时，此时的神经网络被称为多层离散感知器。

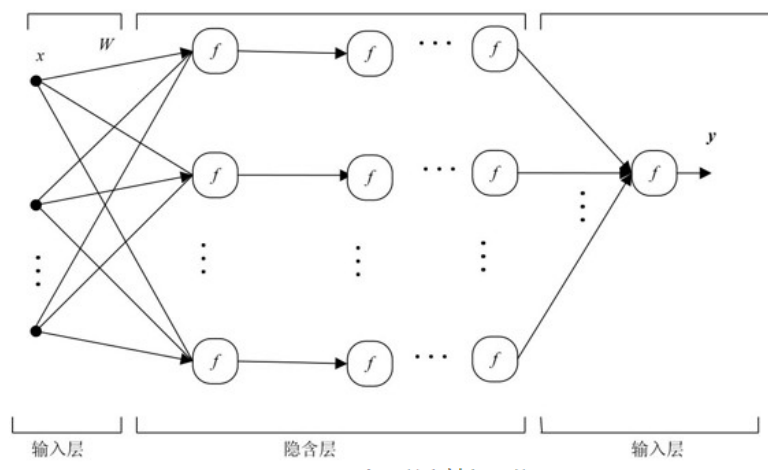


图 2-2 多层前向神经网络

反馈网络同前向网络不同，神经元的输出不仅和上一级神经元的传递信号有关，还和上一时刻的输出有关。在神经网络的训练过程中，从输出反馈到输入的信号发生改变，进而影响输出信号，这使得反馈神经网络拥有前向神经网络更好的训练能力，表现为震荡或者收敛的特性。这种信息处理方式可以用动力学系统理论解释。系统的稳定性和记忆功能非常有关系。反馈神经网络结构如图 2-3 所示，其中从 x 到 y 的信号为输入信号，从 y 到 x 的信号为输出反馈信号。

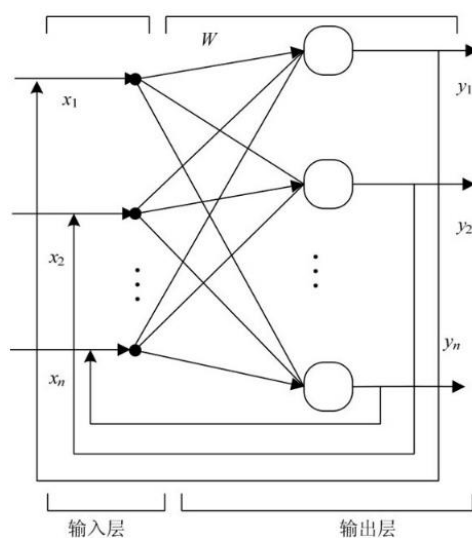


图 2-3 反馈神经网络的结构

人工神经网络的自适应性是通过学习或者说训练实现的，所以学习对于人工神经网络十分重要。人工神经网络的学习过程就是根据环境的不断变化，调整神经元之间的权重，改善输出信号的过程。根据学习的方式和算法的不同，学习方式分为有监督学习和非监督学习。

在有监督学习中，人工神经网络的学习过程会依据纠错原则，在将训练样本输入进人工神经网络后，将期望输出信号同神经网络的输出信号进行对比，得到一个误差信号，然后根据误差信号，按照一定的算法规则调整神经元之间的权重，改变神经元之间的连接强度，经过多次训练后，权重会收敛为一个确定的值。如果环境发生变化，训练样本不同，可以重新训练神经网络，再次修改神经元之间的权重，使得人工神经网络更好地适应新环境。使用监督学习的神经网络模型有 BP 神经网络、反传网络、感知器等。

在非监督学习时，不会将给定的标准样本输入到人工神经网络中，而是直接将人工神经网络置于环境中，相当于同时进行训练和工作。神经网络会一直从环境中接受

输入信息, 然后根据网络中神经元的连接方式和训练算法, 获取输入到神经网络的变量的特征, 从而发现隐含的规律, 再调整神经元之间的连接权重。在非监督学习时, 神经网络权重的调整并不取决于事前给定的样本和期望输出。Hebbian 学习算法是最简单的非监督学习算法。

2.1.2 BP 神经网络模型

反向传播 (Back Propagation) 网络简称为 BP 网络, 是一种按误差逆传播算法训练的多层前馈网络。BP 神经网络算法通过选定一组权重, 能够将事前给定的目标输出作为线性方程的代数和来构建出线性方程, 通过解出该方程, 得到所需要的权重。

BP 神经网络和基础的神经网络一样, 能够从训练中学习并储存信息。BP 神经网络常用学习算法是最速下降法, 通过误差的反向传播调整神经元之间的权重和阈值, 最终使得误差平方和最小。BP 神经网络模型由输入层、隐含层和输出层组成。图 2-4 为一个典型的三层 BP 神经网络的结构图, 层与层之间采用全互连方式, 同一层神经元之间不存在相互连接的关系, BP 神经网络的隐含层可以有一层或多层, 根据实际需求来决定。

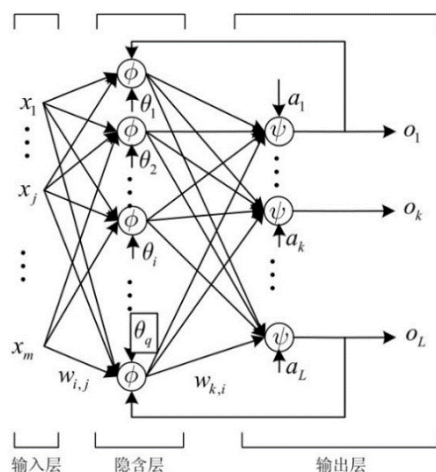


图 2-4 BP 神经网络结构图

其中, X_j 表示输入层第 j 个神经元的输入; $w_{i,j}$ 表示隐含层第 i 个神经元到输入层第 j 个神经元之间的权重; θ_i 表示隐含层第 i 个神经元的阈值; $w_{k,i}$ 表示输出层第 k 个神经元到隐含层第 i 个神经元之间的权重; a_k 表示输出层第 k 个神经元的阈值。

在 BP 神经网络中，有两种信号在流通，一种是函数信号，另一种是误差信号。函数信号正向传播，如图 2-4 所表示，从左向右正向流动，它是输入信号向前传播，并在输出端产生实际输出。误差信号逆向传播，在图 2-4 中从右向左流动，是网络实际输出和期望输出之间的差值。误差信号沿着原来的通路进行反向传播，根据 BP 网络的算法调整神经元之间的权重。

BP 网络的学习过程由前向训练过程和误差反向传播过程组成。在前向训练过程中，输入信号从输入层进入到网络，到达输出层，每层计算节点只会影响下一层计算节点。然后 BP 神经网络将输出信号和期望输出信号进行比较，得到误差，如果误差不能满足要求，则转入误差反向传播过程。误差信号沿原来的路线返回，不断调整 BP 网络各层计算节点之间的权重和阈值，最终到达输入层，再进行前向训练过程。这两个过程依次反复进行，不断调整权重和阈值，当误差最小或达到要求时，停止神经网络训练。

2.1.3 BP 神经网络的算法

BP 神经网络的算法由两个阶段组成。以一个简单的三层 BP 神经网络为例，如图 2-4 所示，设输入层神经元数为 m ，隐含层神经元数为 q ，输出层神经元数为 l ，输入层第 i 个神经元到隐含层第 j 个神经元之间的权值为 W_{ij} ，隐含层第 j 个神经元到输出层第 k 个神经元之间的权值为 w_{jk} ，隐含层第 i 个神经元的阈值为 θ_i ，输出层神经元的阈值为 a_k ，隐含层的激励函数为 ϕ ，输出层的激励函数为 ψ ，输入信号为 $X = (x_1, x_2, \dots, x_m)^T$ ，隐含层输出信号为 $Y = (y_1, y_2, \dots, y_q)^T$ ，输出信号为 $O = (o_1, o_2, \dots, o_l)^T$ ，则在前向传播中，各隐含层神经元的输出值为：

$$y_j = \phi(\sum_{i=1}^m W_{ij} x_i + \theta_i), j = 1, 2, \dots, q \quad (2.1)$$

各输出层神经元的输出值为：

$$o_k = \psi(\sum_{j=1}^q W_{kj} y_j + a_k), k = 1, 2, \dots, l \quad (2.2)$$

在误差的反向传播过程中，设神经元的输出误差函数为：

$$E = \frac{1}{2}(T - o)^T(T - o) = \frac{1}{2}e^T e \quad (2.3)$$

其中, $T = (t_1, t_2, \dots, t_L)^T$ 为神经网络的期望输出信号, $e = (e_1, e_2, \dots, e_L)^T$ 为误差向量。

当学习算法为最速下降法时, 神经网络权值和阈值的修改公式为:

输出层:

$$\Delta w_{j,k} = -\alpha \frac{\partial E}{\partial w_{j,k}} = \alpha (t_k - o_k) \psi' y_j \quad (2.4)$$

$$\Delta a_k = -\alpha \frac{\partial E}{\partial a_k} = \alpha (t_k - o_k) \psi' \quad (2.5)$$

隐含层:

$$\Delta W_{i,j} = -\beta \frac{\partial E}{\partial W_{i,j}} = \beta \left[\sum_{k=1}^l (t_k - o_k) \psi' w_{j,k} \right] \phi' x_i \quad (2.6)$$

$$\Delta \theta_j = -\beta \frac{\partial E}{\partial \theta_j} = \beta \left[\sum_{k=1}^l (t_k - o_k) \psi' w_{j,k} \right] \phi' \quad (2.7)$$

其中 α 和 β 为学习速率, 影响 BP 神经网络的收敛速度。

2.1.4 BP 神经网络的局限性和改善方式

BP 神经网络算法的最终目标就是使得误差函数最小化。BP 神经网络算法采用的是非线性规划中的最速下降方法, 按误差函数的负梯度方向修改权重, 因而通常存在一些问题。比如误差函数的 Hessian 矩阵非正定时, 有产生局部极小点的可能性。并且在实际运用神经网络中, 求解空间往往是多维曲面, 更加加大了陷入局部最小点的可能性。其次学习速率是固定的, 为了保证稳定性, 研究人员需要选择较小的学习速率, 这会导致 BP 神经网络学习效率低、收敛速度慢。还有因为 BP 神经网络采用最速下降法, 误差单调递减, 算法只有下坡能力而没有上坡能力, 容易陷入局部最小状态。针对以上问题, 一般常用三种方式对 BP 神经网络算法进行改进: 附加动量法、自适应学习速率和动量—自适应学习速率调整算法。当使用固定的动量因子改善 BP 神经网络的收敛时, 有可能引起神经网络随机波动, 很难达到收敛点的情况, 所以需要其他方法对 BP 神经网络进行改进。

2.2 卡尔曼滤波理论基础

2.2.1 卡尔曼滤波概述

卡尔曼滤波是一种从输入信号的观测量中,估计目标信号的滤波算法。卡尔曼滤波算法将状态空间概念引入了随机估计中,将目标信号作为被白噪音影响的线性系统的输出,通过状态方程来构建这种输入—输出的关系。卡尔曼滤波利用状态方程、观测方程、系统噪音矩阵和观测噪音矩阵来估计所需信号。卡尔曼滤波不仅仅可以对一维随机过程进行估计,也可以对非平稳且多维的随机过程进行估计,这是因为卡尔曼滤波用到的信号都是时域内的变量。卡尔曼滤波算法本质上是一种线性随机最小方差估计算法,通过状态空间来构建滤波器,利用系统噪音矩阵和观测噪音矩阵的特性,将系统信号的观测量作为滤波器的输入,以目标信号作为滤波器的输出,最后根据系统的状态方程和观测方程估计出所需要的信号。

卡尔曼滤波是最常用的滤波器之一,经常应用于控制、信号处理等领域,并在空间导航、航天技术等多领域投入实际使用。从上个世纪 90 年代开始,国内开始将卡尔曼滤波方法运用在经济领域。投资者可以利用金融高频数据,如期货五分钟交易数据等,就可以预测短期期货价格变化。投资者还可以利用卡尔曼滤波,将随机方程的不确定性减弱,更好地预测股利贴现模型中的现金流变动。

2.2.2 离散型卡尔曼滤波

离散形态的卡尔曼滤波公式主要由状态方程和观测方程两部分组成:

$$X_k = \phi_{k,k-1} X_{k-1} + \Gamma_{k-1} W_{k-1} + \omega_k \quad (2.8)$$

$$L_k = H_k X_k + V_k \quad (2.9)$$

X_k 为 t_k 时刻的 n 维状态参数向量; $\phi_{k,k-1}$ 为 $n \times n$ 阶状态转移矩阵; Γ_{k-1} 为 $m \times r$ 维矩阵,称为动态噪音矩阵; W_{k-1} 为 r 维高阶白噪音状态误差向量; ω_k 为残差项; L_k 为 t_k 时刻的 n 维观测参数向量; H_k 为 $m \times n$ 维矩阵,称为观测矩阵; V_k 为 t_k 时刻的观测噪音矩阵。

当 W 和 V 为高斯白噪音向量且相互独立时,即:

$$E(W_k) = E(V_k) = 0 \quad (2.10)$$

$$Cov(W_k, W_j) = Q_k \delta_{kj} \quad (2.11)$$

$$Cov(V_k, V_j) = R_k \delta_{kj} \quad (2.12)$$

$$Cov(W_k, V_j) = 0 \quad (2.13)$$

卡尔曼滤波的递推公式为：

预期状态

$$\hat{X}_{k,k-1} = \phi_{k,k-1} \hat{X}_{k-1} \quad (2.14)$$

预期状态协方差矩阵

$$P_{k,k-1} = \phi_{k,k-1} P_{k-1} + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T \quad (2.15)$$

状态估计

$$\hat{X}_k = \hat{X}_{k,k-1} + K_k (L_k - H_k \hat{X}_{k,k-1}) \quad (2.16)$$

状态协方差矩阵估计

$$P_k = (I - K_k H_k) P_{k,k-1} \quad (2.17)$$

其中， K_k 为卡尔曼滤波增益矩阵

$$K_k = P_{k,k-1} H_k^T (H_k P_{k,k-1} H_k^T + R_k)^{-1} \quad (2.18)$$

初始状态为

$$\hat{X}_0 = E(X_0) = \mu_0 \quad (2.19)$$

$$\hat{P}_0 = Var(X_0) \quad (2.20)$$

当已知 t_{k-1} 时刻系统的状态 X_{k-1} 时，就可以求出 t_k 时刻的预期状态 $\hat{X}_{k,k-1}$ ，然后根据 t_k 时刻的观测值，利用卡尔曼滤波模型对预测状态进行修正，从而得出 X_k ，通过

不断迭代，从而达到滤波的目的。

2.2.3 扩展卡尔曼滤波

在上述的卡尔曼滤波中，通常用于估计线性随机的离散过程，要求状态方程和观测方程都是线性的。但是有些情况下，系统的状态方程和观测方程无法满足线性的要求，不能直接使用卡尔曼滤波进行估计。这种情况考虑可以将非线性方程线性化，然后在利用线性系统卡尔曼滤波估计。

扩展卡尔曼滤波（EKF）是卡尔曼滤波在非线性的情形下的一种扩展形式，它将期望矩阵和协方差矩阵线性化来解决非线性滤波问题。扩展卡尔曼滤波与线性卡尔曼滤波相似，只是滤波公式中参数需要对非线性函数进行偏导计算，无法像线性卡尔曼滤波那样实现离线计算卡尔曼增益和协方差矩阵。对于离散时间非线性动态状态可以表示为下面的非线性差分方程：

$$X_k = f(X_{k-1}, W_{k-1}, k-1), k=1, 2, \dots \quad (2.21)$$

其中 f 为 n 维向量函数，对其自变量是非线性的， W_k 是 r 维动态噪声矩阵， X_k 是 n 维随机向量矩阵。考虑非线性观测时滤波问题，得出动态噪声的随机差分方程解

$$X_k = \phi(X_{k-1}, k-1) + \Gamma(X_{k-1}, k-1)W_{k-1}, k \geq 1 \quad (2.22)$$

其中 ϕ 为 n 维向量函数， Γ 为 $n \times r$ 阶矩阵函数， W_k 满足正态分布，初始状态 X_0 与 W_k 相互独立，且满足正态分布。则 X_k 是马尔科夫序列。设非线性测量方程为：

$$Z_k = H_k X_k + V_k \quad (2.23)$$

为了将非线性滤波线性化，设测量时刻 k 之前已得到滤波 \hat{X}_{k-1} ，将状态方程中的 ϕ 围绕 \hat{X}_{k-1} 进行泰勒展开，取其线性项，同时用 $\Gamma(\hat{X}_{k-1}, k-1)$ 替换 $\Gamma(X_{k-1}, k-1)$ ，得到以下表达式：

$$X_k = \phi(\hat{X}_{k-1}, k-1) + \frac{\partial \phi_{k-1}}{\partial \hat{X}_{k-1}} (X_{k-1} - \hat{X}_{k-1}) + \Gamma(\hat{X}_{k-1}, k-1) W_{k-1} \quad (2.24)$$

将观测方程中的 H 围绕 \hat{X}_{k-1} 进行泰勒展开, 并取其线性项, 得到以下表达式:

$$Z_k = H(\hat{X}_{k,k-1}, k) + \frac{\partial H_k}{\partial X_{k,k-1}} (X_k - \hat{X}_{k,k-1}) + V_k \quad (2.25)$$

则非线性状态方程和非线性测量方程可分别线性化为以下模型:

$$X_k = \frac{\partial \phi_{k-1}}{\partial X_{k-1}} X_{k-1} + [\phi(\hat{X}_{k-1}, k-1) - \frac{\partial \phi_{k-1}}{\partial \hat{X}_{k-1}} \hat{X}_{k-1}] + \Gamma(\hat{X}_{k-1}, k-1) W_{k-1} \quad (2.26)$$

$$Z_k = \frac{\partial H_k}{\partial X_{k,k-1}} X_k + [H(\hat{X}_{k,k-1}, k) - \frac{\partial H_k}{\partial X_{k,k-1}} \hat{X}_{k,k-1}] + V_k \quad (2.27)$$

由上述式子得到扩展卡尔曼滤波方程:

$$\hat{X}_k = \hat{X}_{k,k-1} + K_k [Z_k - H(\hat{X}_{k,k-1}, k)] \quad (2.28)$$

$$\hat{X}_{k,k-1} = \phi(\hat{X}_{k-1}, k-1) \quad (2.29)$$

$$K_k = P_{k,k-1} \left(\frac{\partial H_k}{\partial \hat{X}_{k,k-1}} \right)^T \left[\frac{\partial H_k}{\partial \hat{X}_{k,k-1}} P_{k,k-1} \left(\frac{\partial H_k}{\partial \hat{X}_{k,k-1}} \right)^T + R_k \right]^{-1} \quad (2.30)$$

$$P_{k,k-1} = \frac{\partial \phi_{k-1}}{\partial X_{k-1}} P_{k-1} \left(\frac{\partial \phi_{k-1}}{\partial X_{k-1}} \right)^T + \Gamma(\hat{X}_{k-1}, k-1) Q_{k-1} \Gamma(\hat{X}_{k-1}, k-1)^T \quad (2.31)$$

$$P_k = (I - K_k \frac{\partial H_k}{\partial \hat{X}_{k,k-1}}) P_{k,k-1} \quad (2.32)$$

2.3 卡尔曼滤波—BP 神经网络组合模型

2.3.1 组合模型概述

由上文可知 BP 神经网络存在学习效率低, 算法收敛速度慢和容易收敛到局部最小值等问题, 并且对输入样本信号的要求较高。考虑将扩展卡尔曼滤波同 BP 神经网络结合, 组成扩展卡尔曼滤波—BP 神经网络组合模型, 利用扩展卡尔曼滤波算法,

改善神经网络自身的缺项。这不仅可以提高学习效率,还可以增强神经网络的抗干扰能力。BP 神经网络算法是在事前给定的输入信号和样本输出信号下,对神经网络的权重进行调节,所以可以将训练 BP 神经网络的过程看成是非线性状态估计问题,将 BP 神经网络的权重作为卡尔曼滤波的状态变量,将 BP 神经网络的权重调整过程转化为权重的最优估计问题,即滤波问题,从而利用扩展卡尔曼滤波来实现非线性网络的学习。

2.3.2 组合模型原理

假设 BP 神经网络有 l 层,每层的神经元个数用 $n_i (i=1,2,\dots,l)$ 表示。把 BP 神经网络中的所有权重和阈值构建状态向量 x , x 的具体表达式如下:

$$x = [w_{1,1}^1 w_{1,2}^1 \cdots w_{n_{l-1},n_l}^l]^T \quad (2.33)$$

状态向量 x 的维度 n_x 可用下式计算:

$$n_x = \sum_{k=1}^{l-1} (n_{k-1} + 1) \times n_k + n_{l-1} \times n_l \quad (2.34)$$

则可得非线性系统下扩展卡尔曼滤波的状态方程和观测方程:

$$x(k+1) = x(k) \quad (2.35)$$

$$y_d(k) = y^1(k) + \varepsilon(k) = h(x(k), y^0(k)) + \varepsilon(k) \quad (2.36)$$

其中, $y^0(k)$ 和 $y^1(k)$ 分别表示期望输出 $y_d(k)$ 的输入向量和输出向量,非线性函数 $h()$ 为 BP 神经网络输入向量、权重、输出向量之间的非线性函数关系,如果用 $NET(x(k), u(k))$ 来表示 BP 神经网络中的非线性映射关系,则有:

$$y^1(k) = h(x(k)) \quad (2.37)$$

$$y^0(k) = NET(x(k), u(k)) = f^l(w^l f^{l-1}(w^{l-1} \cdots f^1(w^1, y^0) \cdots)) \quad (2.38)$$

所以可将期望输出 $y_d(k)$ 泰勒展开:

$$y_d(k) = NET(\hat{x}(k|k), y^0(k)) + H(k)(x(k) - \hat{x}(k|k)) + \dots \quad (2.39)$$

$$H(k) = \left. \frac{\partial NET(x(k), y^0(k))}{\partial x(k)} \right|_{x(k)=\hat{x}(k|k)} \quad (2.40)$$

则上述扩展卡尔曼滤波可以简化为以下形式:

预测估计方程:

$$\hat{x}(k+1) = \hat{x}(k) + K(k+1)[y_d(k+1) - h(\hat{x}(k), y^0(k))] \quad (2.41)$$

预测卡尔曼增益方程:

$$K(k+1) = P(k)H^T(k+1)[H(k+1)P(k)H^T(k+1) + R(k+1)]^{-1} \quad (2.42)$$

预测协方差方程:

$$P(k+1) = [I - K(k+1)H(k+1)]P(k) \quad (2.43)$$

2.4 传统商品期货价格预测模型

2.4.1 ARIMA 模型

ARIMA 模型是指自回归移动平均模型, 是在 ARMA 模型的基础上发展得到, 由 Box 和 Jenkins 于上世纪 70 年代提出, 所以又可称为 B-J 模型。ARIMA 模型适用于预测平稳时间序列, 基本思想是将预测对象的时间序列视为一个随机序列, 通过数学模型来描述这个随机序列, 当模型确定后便可以通过该序列的过去值和现在值来预测未来值。ARIMA 是一种常用的计量经济模型。一个典型的 $ARIMA(p, d, q)$ 表达式如下:

$$y = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t - \omega_1 y_{t-1} - \omega_2 y_{t-2} - \dots - \omega_q y_{t-q} \quad (2.44)$$

其中 $\phi_1, \phi_2, \dots, \phi_p$ 是自回归系数, $\omega_1, \omega_2, \dots, \omega_q$ 是移动平均自回归系数, ε_t 是残差。

通过 ARIMA 模型预测商品期货的价格，一般可以通过以下步骤：

- (1) 先验证期货价格时间序列的平稳性，如果时间序列不平稳，需要进行平稳性处理，常用的平稳性处理方法有取自然对数和差分等。
- (2) 通过自相关 ACF 和偏自相关 PCAF 图来判断阶数和阶数。
- (3) 检验时间序列是否具有趋势性和季节性，然后利用最小二乘法来确定模型主要参数。
- (4) 检验模型中的残差项是否是白噪音，如果残差项是白噪音，说明该预测模型可以很好地预测，但大多数情况残差项非白噪音，需要进一步分解残差项。
- (5) 通过过去和现在的数据，利用建立好的模型进行预测商品期货未来价格。

2.4.2 GARCH 模型

GARCH 模型是扩展 ARCH 模型后得到，被称为广义 ARCH 模型，由 Bollerslev 在 1986 年提出的。由于大部分时间序列都存在异方差性，不适用于 ARIMA 模型，而 GARCH 模型相对普通回归模型对误差的方差进行了进一步建模，更适合用于时间序列的预测和分析。一个基本的 GARCH 模型表达式如下：

$$y_t = c + \beta x_t + \varepsilon_t \quad (2.45)$$

$$\varepsilon_t = \sqrt{h_t} \times v_t \quad (2.46)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \theta_j h_{t-j} \quad (2.47)$$

在预测期货价格中， y_t 代表所要预测的期货价格， x_t 为现货价格， β 为待估计误差参数向量， ε_t 为待估计误差项， h_t 为 ε_t 的条件方差， v_t 为标准正态分布， h_t 和 v_t 相互独立，此时的正态条件方差模型为 GARCH 模型。

GARCH 模型受到市场正负冲击对条件方差的影响是对称的，所以 GARCH 模型不能描述条件方差波动的非对称性。为了衡量市场冲击非对称性，在 GARCH 模型的基础上，提出了 EGARCH 模型。

2.4.3 基于卡尔曼滤波价格预测模型

建立期货价格的状态方程和观测方程，可以通过卡尔曼滤波估计期货的价格。卡尔曼滤波状态估计的目的就是在已知初始状态变量的统计特性（状态变量的初始期望和初始协方差等）时，从观测得到最优估计。商品期货价格是一维时间序列，可以

假定状态转移矩阵 ϕ 的初始值为 1，表示 k 时刻的价格和 $k+1$ 时刻的价格是一致的。

假定控制变量为 0，状态方程的残差项 ω_k 为白噪音。因为状态变量的初始值和扰动对系统影响不大，可假定状态变量的初始值为 0。对于商品期货来说，观测值和实际值是一一对应的，且观测值很大程度是直接代表实际值，可以假定观测矩阵为 1，观测噪音是随机白噪音。基于以上分析，我们得到下列方程：

期货价格状态方程：

$$X_{k+1} = \phi X_k + \omega_k \quad (2.48)$$

期货价格观测方程：

$$L_k = H_k X_k + V_k \quad (2.49)$$

接着通过卡尔曼递推算法，便可预测商品期货价格。上文已做推导，此处不在赘述。

3 期货价格预测实证研究

3.1 数据的选择与预处理

3.1.1 原始数据的选取

本文主要选取沥青、白银、PVC 和豆粕商品期货作为研究对象,选取原因是这四种商品期货长期交易活跃、成交量大,是商品期货市场的交易主力,比较有代表性,同时可以提高实证结果的可靠性。本文采用从 Choice 客户端中获取的沥青主力连续(BU0.SHF)、沪银主力连续(AG0.SHF)、PVC 主力连续(V0.DCE)和豆粕主力连续(M0.DCE)数据。主力连续是由主力合约,也就是成交量最大的合约的 K 线连接而成,主要目的是避免主力合约换月所造成的价格不连续。通过研究豆粕主力连续,可以知道长期豆粕期货价格的变化。本文选取的样本区间为 2016 年 1 月 4 日至 2021 年 3 月 17 日的日数据共 1267 期,选取前 80%作为训练样本,即 1013 期作为训练样本,后 254 期作为检验样本。

参考相关研究选取的变量,选取商品期货的开盘价 OP、最高价 HP、最低价 LP、收盘价 CP、成交量 V 以及持仓量 H 来描述交易情况,选取无风险利率 r 描述市场情况,总共 7 个变量。一般将合约的收盘价定为合约价格,通过上一日的变量来预测下一日的合约价格。无风险利率有很多种,如国债同期收益率、银行间市场质押式回购利率和上海银行间同业拆放利率(Shibor)等。Shibor 品种包括隔夜、1 周、2 周、1 个月、3 个月、6 个月、9 个月及 1 年。本文选择 Shibor 一年利率作为无风险利率。

3.1.2 数据的预处理

BP 神经网络的输出数据高度依赖输入数据,由于不同数据量纲和数量级不同,如果不进行处理,BP 神经网络的收敛速度会变慢或者预测结果出现较大偏差。表 3.1 为部分原始数据。豆粕期货的开盘价同最高价、最低价和收盘价在相同的数量级上,量纲相同,但远远小于成交量和持仓量的数量级,且量纲不同。如果不对数据进行预处理,会影响神经网络的收敛速度和预测结果。

表 3.1 部分原始数据

时间	开盘价	最高价	最低价	收盘价	成交量	持仓量	无风险利率
2016-01-04	2370	2377	2321	2326	1335646	1553566	3.350
2016-01-05	2328	2347	2325	2343	1139054	1504576	3.350
2016-01-06	2344	2352	2327	2347	1367216	1514752	3.347
2016-01-07	2344	2386	2339	2376	1975726	1457710	3.346
2016-01-08	2373	2407	2363	2404	1978374	1574678	3.345
2016-01-11	2406	2413	2377	2383	1490070	1470028	3.344
2016-01-12	2383	2397	2372	2378	1498406	1470060	3.340
2016-01-13	2380	2413	2376	2382	1877026	1533138	3.334
2016-01-14	2386	2415	2376	2406	1860008	1576008	3.318
2016-01-15	2406	2416	2395	2404	1355476	1569586	3.295
2016-01-18	2396	2462	2394	2458	2383820	1578096	3.271
2016-01-19	2458	2470	2450	2462	1453138	1558610	3.250
2016-01-20	2462	2471	2440	2453	1575738	1526076	3.253
2016-01-21	2448	2452	2419	2425	1485312	1477962	3.253
2016-01-22	2428	2459	2425	2452	1164004	1417696	3.256
2016-01-25	2452	2463	2428	2434	1147520	1384894	3.259
2016-01-26	2432	2453	2423	2443	839828	1384876	3.268
2016-01-27	2440	2452	2428	2448	1050482	1417870	3.265
2016-01-28	2446	2460	2435	2443	1263326	1381288	3.265
2016-01-29	2442	2444	2411	2419	1063362	1368258	3.265
2016-02-01	2422	2451	2416	2444	822598	1337758	3.265
2016-02-02	2440	2461	2437	2456	675808	1364016	3.265

所以在机器学习中，将样本数据输入 BP 神经网络训练前，首先需要对样本数据进行标准化处理。通过标准化处理的数据将转化为无量纲、无数量级差异的数值，可以消除不同变量之间因属性不同给 BP 神经网络的训练结果造成的负面影响，从而使 BP 神经网络的预测数据更具有可比性。常用的标准化处理方法为极差标准化法、Z-score 标准化法、线性比例标准化法和归一法等，本文采用归一法进行标准化处理。本文通过使用 MATLAB2017 的 mapminmax 函数，该函数将样本数据大小转换为 $[-1, 1]$ 之间的值，函数表达式为：

$$x = \frac{2(y - \min y)}{\max y - \min y} - 1 \quad (3.1)$$

其中, x 为处理后的数据, y 为原数据, $\max y$ 、 $\min y$ 分别为该变量中的最大值和最小值。豆粕主力连续部分归一化后的数据如表 3.2 所示:

表 3.2 部分归一化后的数据

时间	开盘价	最高价	最低价	收盘价	成交量	持仓量	无风险利率
2016-01-04	-0.8728	-0.8957	-0.9267	-0.9491	-0.7172	-0.5300	-0.3739
2016-01-05	-0.9422	-0.9431	-0.9201	-0.9212	-0.7706	-0.5709	-0.3739
2016-01-06	-0.9158	-0.9352	-0.9167	-0.9146	-0.7086	-0.5624	-0.3768
2016-01-07	-0.9158	-0.8815	-0.8968	-0.8670	-0.5432	-0.6100	-0.3778
2016-01-08	-0.8679	-0.8483	-0.8568	-0.8210	-0.5425	-0.5123	-0.3787
2016-01-11	-0.8134	-0.8389	-0.8335	-0.8555	-0.6752	-0.5997	-0.3797
2016-01-12	-0.8514	-0.8641	-0.8418	-0.8637	-0.6729	-0.5997	-0.3836
2016-01-13	-0.8563	-0.8389	-0.8351	-0.8571	-0.5700	-0.5470	-0.3890
2016-01-14	-0.8464	-0.8357	-0.8351	-0.8177	-0.5746	-0.5112	-0.4050
2016-01-15	-0.8134	-0.8341	-0.8035	-0.8210	-0.7118	-0.5166	-0.4274
2016-01-18	-0.8299	-0.7615	-0.8052	-0.7324	-0.4322	-0.5095	-0.4512
2016-01-19	-0.7275	-0.7488	-0.7119	-0.7258	-0.6853	-0.5257	-0.4711
2016-01-20	-0.7209	-0.7472	-0.7286	-0.7406	-0.6519	-0.5529	-0.4682
2016-01-21	-0.7440	-0.7773	-0.7635	-0.7865	-0.6765	-0.5931	-0.4682
2016-01-22	-0.7770	-0.7662	-0.7535	-0.7422	-0.7639	-0.6434	-0.4653
2016-01-25	-0.7374	-0.7599	-0.7485	-0.7718	-0.7683	-0.6708	-0.4624
2016-01-26	-0.7704	-0.7757	-0.7569	-0.7570	-0.8520	-0.6708	-0.4534
2016-01-27	-0.7572	-0.7773	-0.7485	-0.7488	-0.7947	-0.6433	-0.4565
2016-01-28	-0.7473	-0.7646	-0.7369	-0.7570	-0.7369	-0.6738	-0.4565
2016-01-29	-0.7539	-0.7899	-0.7769	-0.7964	-0.7912	-0.6847	-0.4565
2016-02-01	-0.7870	-0.7788	-0.7685	-0.7553	-0.8567	-0.7102	-0.4565
2016-02-02	-0.7572	-0.7630	-0.7336	-0.7356	-0.8966	-0.6882	-0.4565

在获得 BP 神经网络输出数据后, 为了将实验结果真同实收盘数据相比较, 需要对输出数据进行反归一化处理。表达式为:

$$z = 0.5(x + 1)(\max y - \min y) + \min y \quad (3.2)$$

其中, x 为输出数据, z 为反归一化后的数据。

3.2 商品期货价格预测模型的构建与训练

3.2.1 BP 神经网络模型

本文采用 MATLAB 构建 BP 神经网络,采用 MATLAB 中函数 newff 构建 BP 神经网络。newff 函数在 2010 年进行更新,使用中需要注意新版本 newff 函数和旧版本 newff 函数的区别。最主要的区别是新版本的函数中,不需要规定隐含层的神经元数,函数可以从样本输出向量中判断出输出层的神经元数。

(1) BP 神经网络的层数

随着隐含层数及其神经元数量的增加,BP 神经网络的拟合程度会有所提升,但是过多的层数和神经元数量会使 BP 神经网络更加复杂,导致训练时间增长,或者出现训练中拟合程度好,但是泛化能力差的情况。根据 Kolmogorov 定理可知:当 BP 神经网络架构合理且神经元之间的权重合适时,BP 神经网络仅需拥有三层的结构便可以无限逼近一个连续函数。基于以上考虑,本文选择只有一个隐含层的 BP 神经网络。

(2) 激励函数

BP 神经网络的激励函数通常使用 S 型函数,最常使用的函数有 tansig、logsig 和 purelin 函数,本文将采用 tansig 函数作为输入层到隐含层的激励函数,数学表达式为 $f(x) = \frac{2}{1 + e^{-2x}} - 1$,由于 BP 神经网络的输出范围大于 $[-1,1]$,所以采用 purlin 函数为隐含层到输出层的激励函数。

(3) 神经网络的算法

BP 神经网络常用的学习算法有很多种:如梯度下降 BP 神经网络训练算法 traingd、动量反传的 BP 神经网络算法 traingdm 等。本文的 BP 神经网络学习算法采用动量反传和动态自适应学习率的梯度下降 BP 神经网络算法 traingdx,这是一种带有动量项的自适应学习算法,优点是学习速度要比 traingdm 快。

(4) 输入层的节点数

输入层的节点数根据样本输入向量可得,本文共有 7 个输入变量,所以 BP 神经网络的输入层的节点数为 7,每一个神经节点对应一个变量或者说特征值。

(5) 输出层的节点数

在最新版本的 newff 函数中,函数可以根据样本输出向量自行计算出所需的输出变量数,即输出层节点数。所以现在的 BP 神经网络不需要提前设定好输出层节点数。本文的输出变量只有商品期货的收盘价格,故输出变量为 1,输出层节点数为 1。

(6) 隐含层的节点数

根据相关研究表明,BP 神经网络的隐含层节点数一般可以根据经验公式 $r = \sqrt{n+m} + a$ 进行确定范围,其中 r 、 n 、 m 分别为隐含层、输入层和输出层的节点数, a 为[1,10]内的所有自然常数。本文的输入层有 7 个节点,输出层有 1 个节点,所以依据上述经验公式,本文的神经网络隐含层的节点数在[4, 13]之间。同时,还要遵循在客观反映输入和输出变量关系为目标时,选择节点较少的神经网络的原则,在神经网络的拟合效果和复杂程度之间进行衡量考虑。

BP 神经网络衡量训练效果,即分析神经网络输出向量和期望输出向量之间的误差时,通常用性能函数来进行衡量。本文采用平均绝对误差 MAE 来衡量。

为确定 BP 神经网络最佳的隐含层节点数,本文依次对有着不同隐含层节点数的 BP 神经网络进行训练,通过均方误差和平均绝对误差的大小,衡量隐含层节点数和训练效果的最佳组合,从而确定好最佳的隐含层节点数。BP 神经网络的最大训练次数设定为 5000,最大运行误差设定为 0.001,学习效率设定为 0.01。

以豆粕期货为例,将神经网络训练结果进行比较,结果如表 3.3 所示。

表 3.3 不同隐含层节点的 BP 神经网络训练误差

隐含层节点数	MAE
4	0.0504
5	0.0460
6	0.0450
7	0.0525
8	0.0612
9	0.0636
10	0.0789
11	0.1125
12	0.1032
13	0.7905

从表 3.3 中可以看出,不同的隐含层节点数的 BP 神经网络训练效果不同,神经网络的泛化能力先随着隐含层节点数的增加而增加,但增加到一定程度后开始降低。一般来说,出现这种情况的原因是随着隐含层节点数增加,神经网络能更好地对样本拟合,但是隐含层节点数过多会导致神经网络过拟合,预测结果变差。当隐含层节点数为 6 时,平均绝对误差最小,泛化能力最好,同时隐含层节点数也较少,所以在本文选择采用隐含层节点数为 6 的 BP 神经网络进行商品期货价格预测。对于预测其他三种商品期货,使用同样的方法来确定最佳隐含层节点数。

3.2.2 卡尔曼滤波-BP 神经网络模型

卡尔曼滤波有良好的降噪能力,而 BP 神经网络训练依赖输入样本,所以将两者结合,再对商品期货价格进行预测。扩展卡尔曼滤波-BP 神经网络模型基本步骤如下:

- (1) 将样本向量输入 BP 神经网络中,得到初始神经网络权重、阈值、误差矩阵和神经网络输出。
- (2) 将神经网络的权重和阈值作为状态向量,神经网络输出作为卡尔曼滤波的观测向量,进行卡尔曼滤波的更新,得到新的状态向量,即 BP 神经网络的权重。
- (3) 将更新后的权重再次输入 BP 神经网络,再进行预测。
- (4) 重新进行下一轮卡尔曼滤波的更新。
- (5) 重复以上步骤,直至 BP 神经网络满足要求最小误差,得到预测值。

卡尔曼滤波-BP 神经网络的层数、激励函数和算法等和原来的 BP 神经网络相同,但是卡尔曼滤波在降噪的同时会减少样本中的信息,需要重新估计卡尔曼滤波-BP 神经网络佳的隐含层节点数。依次对不同隐含层节点数的神经网络进行训练,并用平均绝对误差衡量训练结果,并与 BP 神经网络训练结果进行比较。结果如表 3.4 所示。

表 3.4 不同隐含层节点的卡尔曼滤波-BP 神经网络训练误差

隐含层节点数	改进后 MAE	MAE
4	0.0557	0.0504
5	0.0509	0.0460
6	0.0547	0.0450
7	0.0508	0.0525
8	0.0505	0.0612
9	0.0332	0.0636
10	0.0561	0.0789
11	0.0525	0.1125
12	0.0587	0.1032
13	0.0572	0.7905

从表 3.4 可以看出,改进后 BP 神经网络的平均绝对误差在总体上有所减少,表明 BP 神经网络泛化能力提高,对豆粕期货价格的预测精确度增加。隐含层节点数为 9 时,均方误差最小,所以考虑采用隐含层节点数为 9 的神经网络进行预测研究。对于预测其他三种商品期货,使用同样的方法来确定最佳隐含层节点数。

3.2.3 卡尔曼滤波模型

卡尔曼滤波预测流程如下:

- (1) 输入样本。
- (2) 设定状态噪声和观测噪声。
- (3) 根据 $k-1$ 时刻的状态预测 k 时刻的状态 $\hat{X}_{k,k-1}$ 。
- (4) 估计 $k-1$ 时刻的预测误差估计 k 时刻预测误差 $P_{k,k-1}$ 。
- (5) 计算卡尔曼滤波增益。
- (6) 计算系统最优估计 \hat{X}_k 。
- (7) 计算系统当前预测误差 P_k 。
- (8) 重复上诉流程。

本文中的样本即四种商品期货的价格时间序列,初始价格设定为 1,噪音为随机生成的白噪音。通过 MATLAB 上代码运算,便可得到商品期货预测价格。

3.3 预测结果分析

使用 BP 神经网络模型，卡尔曼滤波-BP 神经网络模型和基于卡尔曼滤波预测不同商品期货的价格，结果如下所示：

(1) PVC 期货价格预测

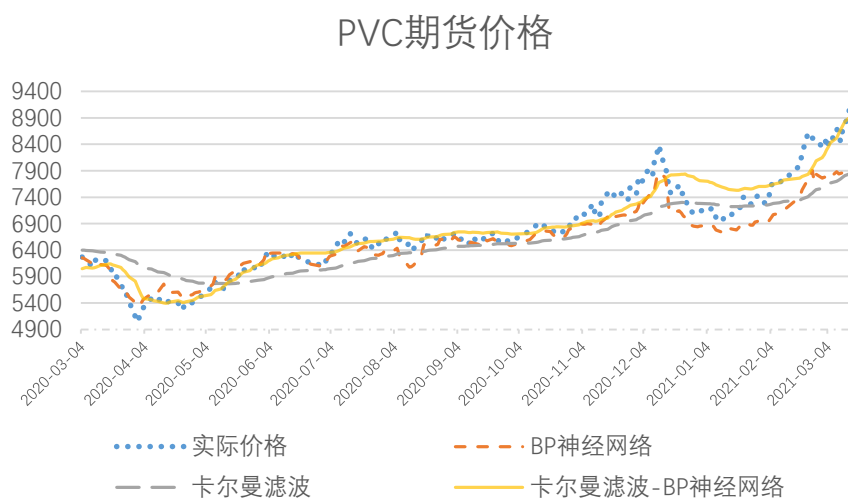


图 3-1 PVC 期货价格预测结果

(2) 豆粕期货价格预测

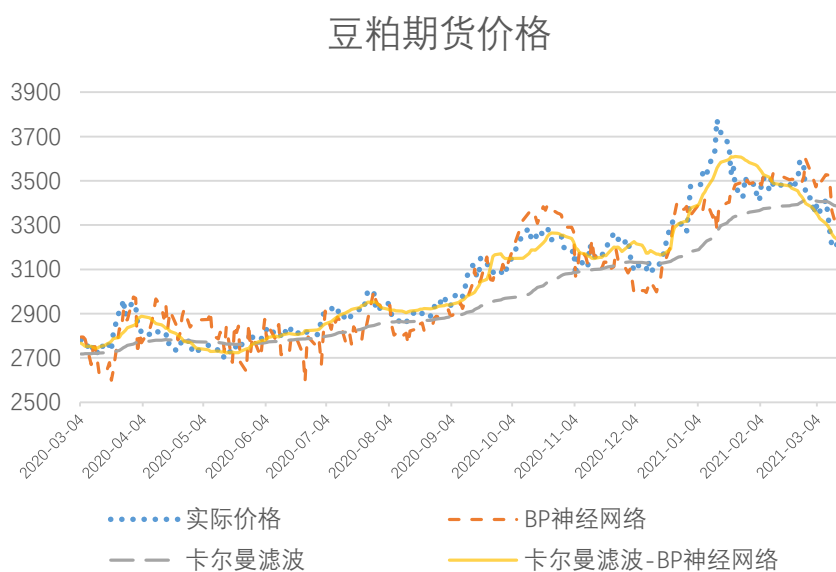


图 3-2 豆粕期货价格预测结果

(3) 沪银期货价格预测

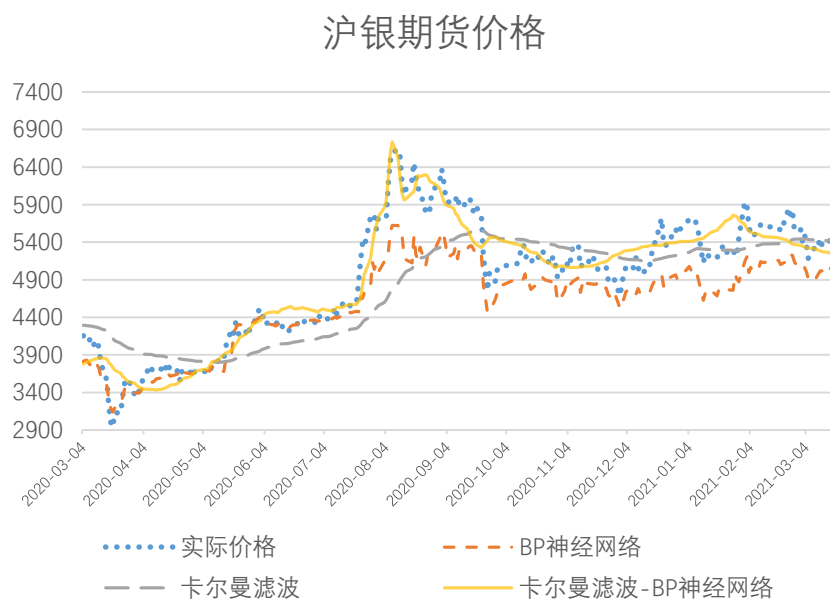


图 3-3 沪银期货价格预测结果

(4) 沥青期货价格预测

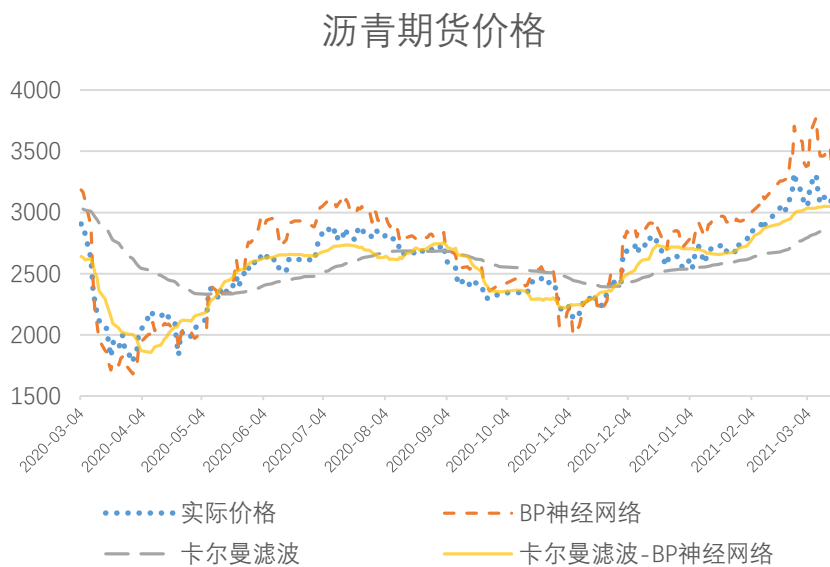


图 3-4 沥青期货价格预测结果

评价不同模型的预测效果可以计算预测的精确度来衡量, 价格预测模型的精确度是指预测值和实际值之间的误差。主要通过均方误差、绝对误差和相对误差来判断误差大小, 这三种指标没有绝对的优劣, 都能从一个方面反映误差。本文采用均方误差指标 (MSE)、平均绝对误差 (MAE) 和平均相对误差 (MRE)。计算不同种类期货的预测误差, 结果如下

(1) PVC 期货

表 3.4 PVC 期货价格预测误差

预测模型	MSE	MAE	MRE
BP 神经网络	0.0427	0.0455	0.0333
卡尔曼滤波	0.0830	0.0366	0.0536
卡尔曼滤波-BP 神经网络	0.0258	0.01816	0.0262

(2) 豆粕期货

表 3.5 豆粕期货价格预测误差

预测模型	MSE	MAE	MRE
BP 神经网络	0.0409	0.0450	0.0257
卡尔曼滤波	0.0691	0.1155	0.0304
卡尔曼滤波-BP 神经网络	0.0150	0.0332	0.0148

(3) 沪银期货价格预测

表 3.6 沪银期货价格预测误差

预测模型	MSE	MAE	MRE
BP 神经网络	0.0753	0.0351	0.0654
卡尔曼滤波	0.0844	0.0374	0.0759
卡尔曼滤波-BP 神经网络	0.0251	0.0197	0.0413

(4) 沥青期货价格预测

表 3.7 沥青期货价格预测误差

预测模型	MSE	MAE	MRE
BP 神经网络	0.0132	0.0156	0.0594
卡尔曼滤波	0.0348	0.0228	0.0956
卡尔曼滤波-BP 神经网络	0.0061	0.0190	0.0401

由上述图表得出以下结论:

(1) 基于卡尔曼滤波的预测模型对期货价格的预测能力最差，对价格的预测更多的是对价格走势的预判。分析主要原因是实际中观测噪音矩阵和状态噪音并非白噪音，导致模型误差过大。

(2) BP 神经网络能在一定程度上预测期货价格，但是预测能力一般，预测价格波动率大。基于图 3-2 和图 3-3 可以看出 BP 神经网络面对价格突然大幅度波动时，并不能很好的预测出来，对样本区间外数据的泛化能力较弱。且 BP 神经网络的预测价格在后期会逐渐偏离真实价格。分析原因，在训练阶段，BP 神经网络通过误差的逆传播可以更好地拟合样本，但是在工作阶段，BP 神经网络没有误差的逆传播的过程，就失去了降噪能力。随着误差不断积累，预测效果不断下降，预测价格逐渐偏离实际价格。

(3) 卡尔曼滤波-BP 神经网络的价格预测能力最好，预测误差远远小于其他神经网络。且将卡尔曼滤波-BP 神经网络的预测价格曲线同 BP 神经网络的预测价格曲线相比较，发现前者更加平稳，后者波动剧烈。分析原因是因为卡尔曼滤波过滤掉一部分噪音，BP 神经网络受到输入数据中的噪音影响变小。改进后的神经网络泛化能力提高，改进后的神经网络对期货价格预测不会逐渐偏离实际价格，或者说偏离程度远远小于 BP 神经网络。

4 量化投资策略

为了验证价格预测模型在实际中的作用,基于预测模型构建跨期套利策略,并和传统的标准距离法套利相比较。

4.1 跨期套利策略的理论基础

跨期套利策略是以交易金额或交易数量相同的方式,同时买入并卖出同种商品不同交割月份的期货合约。同一种商品期货在不同月份进行交割,合约距离交割日期的时间不同,所以不同合约受到的市场因素影响程度不同,导致合约价格变动程度不同。通过两个期货合约的价差扩大或减小,达到套利的目的。在一般情况下,商品期货市场是稳定的,投资者对未来预期不会发生太大波动,所以不同交割日期的商品期货合约的价差也应当是稳定的。在这种平稳状态下,不存在套利的可能性。但是如果市场发生突然变化,出现利多或者利空的消息,商品期货市场会发生波动,交割日期较近的期货合约波动较大,交割日期远的期货合约波动较小,这就使得两者价差发生变化,当价差的变动幅度大于无套利区间时,就可以进行跨期套利。对于跨期套利来说,收益的主要来源是合约价差同预期方向变动,风险的主要来源是合约价差变动方向和预期的相反。

进行跨期套利的投资者需要对不同交割日期的期货合约的价格进行预测,这是一种投机性行为。跨期套利的主要方式是通过预测未来不同交割日期的期货合约之间的价差是怎么样变动的,并不需要预测出期货合约价格整体走势情况。

根据期货合约的交割月份以及买卖远近期货合约的不同,商品期货跨期套利可以分为一下三种方式:

(1) 牛市套利。当期货市场总体情况较好时,合约价格总体上呈上涨趋势。在牛市套利中,投资者预期交割日期较近的商品期货的价格波动会大于交割日期较远的商品期货价格波动,比如说交割日期较近的价格上涨幅度大于交割日期较远的,投资者通过买入交割日期较近的期货合约,同时卖出交割日期较远的期货合约,便可获

得收益。

(2) 熊市套利。当期货市场总体情况较差时, 合约价格总体上呈下降趋势。在熊市套利中, 刚好和牛市套利相反, 投资者预期交割日期较近的价格下跌幅度大于交割日期较远的, 此时通过买入交割日期较远的期货合约, 同时卖出交割日期较近的期货合约, 便可获得收益。

(3) 蝶式套利。蝶式套利可以看成是以上两种套利的组合, 当期货市场情况不稳定, 合约价格出现上升, 下降, 再上升的情况时, 由于不同交割日期的合约的价格强弱不同, 投资者利用三种不同交割月份的商品期货合约进行构建套利模型。比如投资者买入交割日期较近和交割日期较远的合约, 卖出中间月份的合约。

4.2 标准距离法跨期套利

采用标准距离法构建套利策略, 即价差的均值回复。计算出合约价差的均值, 并设置均值上下 a 个标准差作为无套利区间。当合约价差大于价差均值加上 a 个标准差时, 处于无套利区间的上方, 做空价差套利, 即卖出距离交割日期较近的合约, 买入距离交割日期较远的合约开仓; 当合约价差小于价差均值减去 a 个标准差时, 处于无套利区间的下方, 则做多价差套利, 卖出距离交割日期较远的合约, 买入距离交割日期较近的合约开仓。当合约价差回复到价差均值时进行平仓。因为合约价差并不总是同预期变化, 所以当做空价差开仓时, 如果合约价差没有按预期回复均值, 而是穿出无套利区间, 需要及时平仓止损, 对于做多价差开仓也是如此。

4.2.1 合约选择

选择豆粕期货合约作为研究对象。由于供求关系和季节性特征等因素的影响, 同时市场上进行交易的不同交割月份的豆粕期货合约一般只有八个, 目前在交易的豆粕期货合约有 M2105.DCE、M2107.DCE、M2108.DCE、M2109.DCE、M2111.DCE、M2112.DCE、M2201.DCE 和 M2203.DCE。由于后四个合约上市日期较晚, 数据量较少, 用于构建策略的机会有限, 所以舍弃掉后四个合约, 主要研究前四个豆粕合约。

不同交割日期的豆粕合约交易情况不同, 有些豆粕合约交易市场十分活跃, 有些豆粕合约交易市场则十分冷淡。交易活跃的豆粕合约一般是主力合约, 有大量机构投资者进入, 资金流动量大, 合约的成交量大。而成交量大的合约自然也会导致合约的持仓量大。所以持仓量大且成交量大的合约, 可以为确定主力合约。且跨期套利策略

是通过价差变化来获利，价差变化越频繁，构建套期策略的机会也越多，获得收益的可能性更高。需要选择交易活跃的主力合约，才能更好地进行跨期套利策略。

通过计算不同合约在上市以后的日均成交量和日均持仓量，结果如图 4-1 和图 4-2 所示，确定 M2105.DCE 和 M2109.DCE 是主力合约。选择这两个豆粕期货合约作为研究对象。选取豆粕期货合约的每日收盘价作为期货合约的价格，M2109.DCE 上市时间为 2020 年 9 月 15 日，所以样本区间为 2020 年 9 月 15 日至 2021 年 3 月 17 日。因为选取两个豆粕期货合约作为研究对象，无法采用蝶式套利，所以主要采用牛市套利和熊市套利策略。

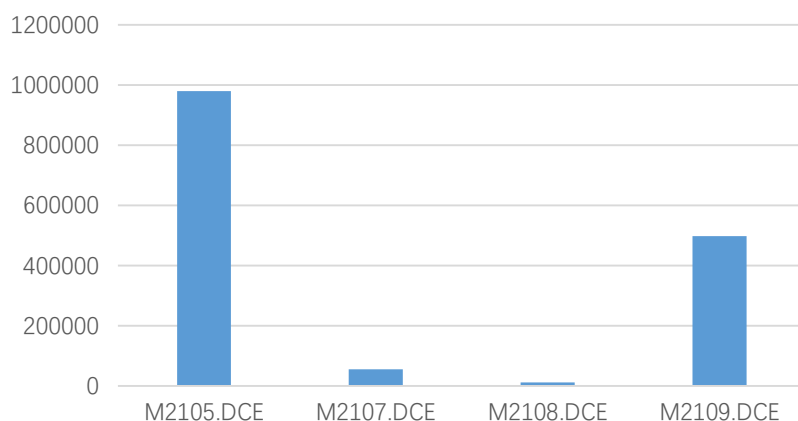


图 4-1 日均持仓量

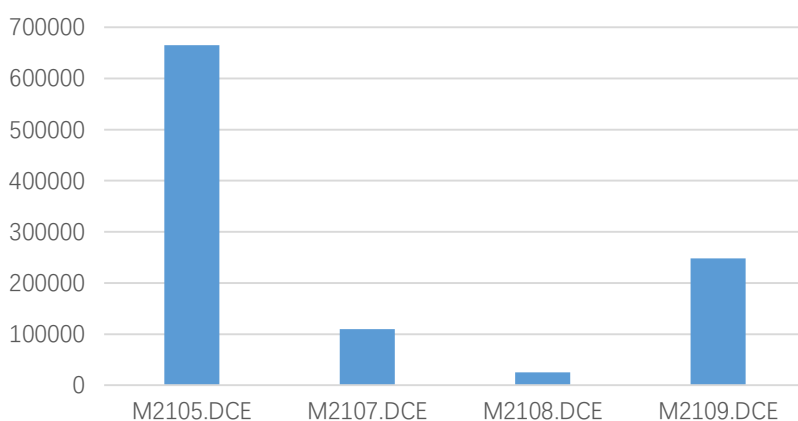


图 4-2 日均成交量

4.2.2 协整检验

对豆粕期货进行套利分析前，首先要进行协整检验，如果用于套利的价格序列存在协整关系，则说明进行跨期套利是合适的，反之则可能因为两者的相关关系发生改变而导致价差发生重大变化，进而导致跨期套利失败。

首先可以从两个商品期货的价格走势来初步判断。图 4-3 为两个合约价格随时间的变化图，可以看出两个合约价格的走势十分一致。在 2020 年 10 月，M2105.DCE 的价格曲线从下往上穿过了 M2109.DCE 的价格曲线，查询历史消息，主要原因是受到当时豆粕现货价格大幅上涨的影响，所以近月合约波动比远月合约的波动剧烈，存在着跨期套利的机会。

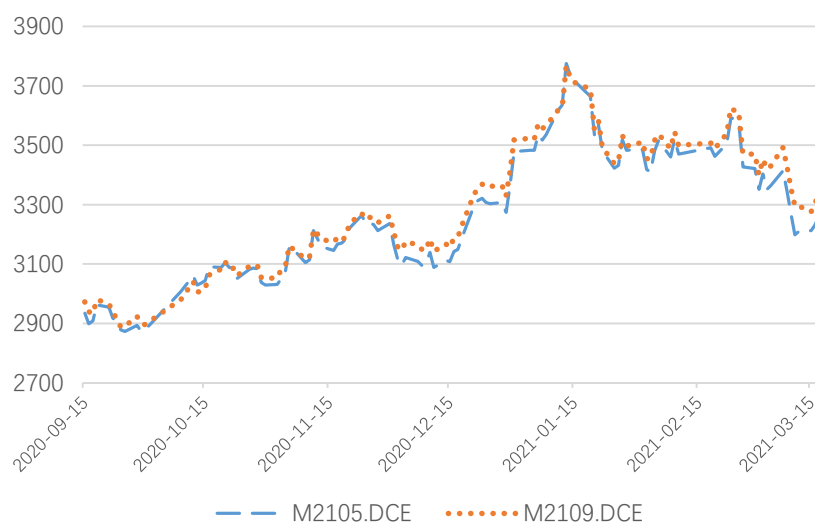


图 4-3 价格变化图

首先对两个时间序列和差分序列进行单位根检验，结果如表 4.1：

表 4.1 单位根检验结果

	T 统计量	1%	5%	10%	结果
price2105	-1.804	-3.504	-2.889	-2.579	不平稳
Δ price2105	-7.151	-3.504	-2.889	-2.579	平稳
price2109	-1.651	-3.504	-2.889	-2.579	不平稳
Δ price2109	-7.281	-3.504	-2.889	-2.579	平稳

其中，price2105 和 price2109 分别是 M2105.DCE 和 M2109.DCE 的时间序列，可

以看出 price2105 和 price2109 两个时间序列都是一阶单整的，满足协整要求两个序列是同阶单整的前提，继续进行协整分析。

采用 E—G 两步法检验 price2105 和 price2109 两个时间序列的协整关系，表 4.2 是回归方程残差的单位根检验结果，可以看出残差在 1%的显著性水平下是平稳的，因此两价格序列间存在着稳定的协整关系，回归模型下面所示，其中所有系数在 1%的水平下显著。

表 4.2 残差项单位根检验结果

	T 统计量	1%	5%	10%	结果
残差	-21.976	-3.990	-3.387	-3.080	平稳

$$price_{2109} = -19.49 + 1.008 * price_{2105} \quad (4.1)$$

4.2.3 构建套利组合及收益

式子中，price2109 的系数为 1.008，由于豆粕期货合约不可拆分，且系数与 1 相近，所以直接采用卖出一手近月合约同时买入一手远月合约或者买入一手近月合约同时买入卖出一手远月合约构建套利组合。目前豆粕期货市场的保证金比例 5%，交易成本为 1.5 元/手，且豆粕期货合约的价格最高不超过 3900 元，根据稳健性原则，保证金按照 400 元/次计算，交易成本按照 3 元/次计算，以 5 倍保证金作为本金，即 2000 元，保证顺利构建跨期套利组合。根据不同 a 值来套利，套利结果如下：

表 4.3 跨期套利结果

a	1	1.5	2	2.5
交易次数	12	10	10	6
胜率	50%	60%	50%	66.70%
年化收益率	25.66%	24.84%	32.18%	20.18%
最大回测比率	0.0021	0.0019	0.0017	0.0019
夏普比率	1.2473	1.2285	1.4083	1.2617

4.3 卡尔曼滤波—BP 神经网络模型跨期套利

卡尔曼滤波—BP 神经网络模型相对其他两个预测模型，可以更加准确的预测出商品期货价格，所以可以应用卡尔曼滤波—BP 神经网络模型构建跨期套利模型。

套利策略如下：先通过卡尔曼滤波—BP 神经网络模型预测期货合约价格，然后通过价格计算得到预测价差。设定一个常数 b ， b 根据价差的标准差来判断，价差的标准差由历史数据得到。当预测价差扩大，且价差变动值大于 b 时，做熊市套利，买入远月合约，卖出近月合约；当预测价差减小，且价差变动值小于 $-b$ 时，做牛市套利，卖出远月合约，买入近月合约；当持有头寸且预测价差的绝对值小于 b 时，进行平仓。选取豆粕主力合约 M2105.DCE 和 M2109.DCE 作为样本，时间区间为 2020 年 9 月 15 日至 2021 年 3 月 17 日。通过卡尔曼滤波—BP 神经网络模型预测 M2105.DCE 和 M2109.DCE 的价格，结果如图 4-4 和图 4-5 所示，其中纵轴坐标为归一化后的价格。

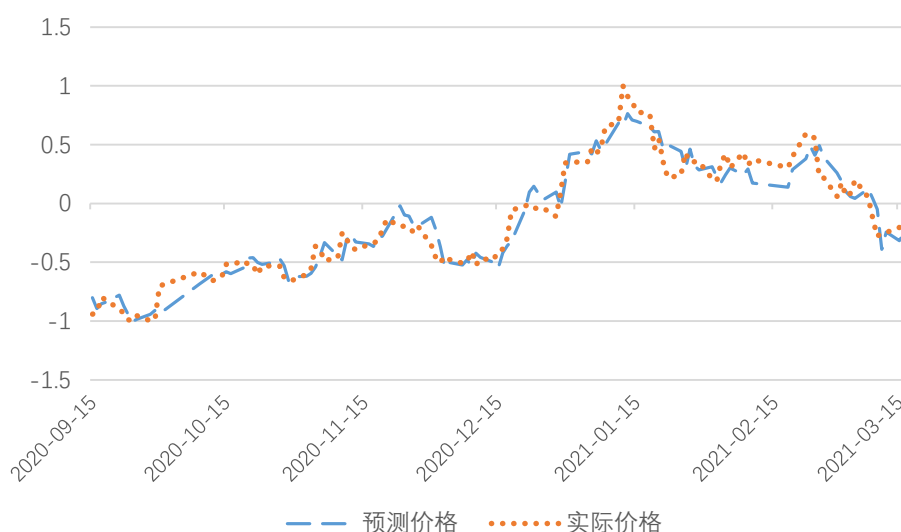


图 4-4 M2105.DCE 的价格

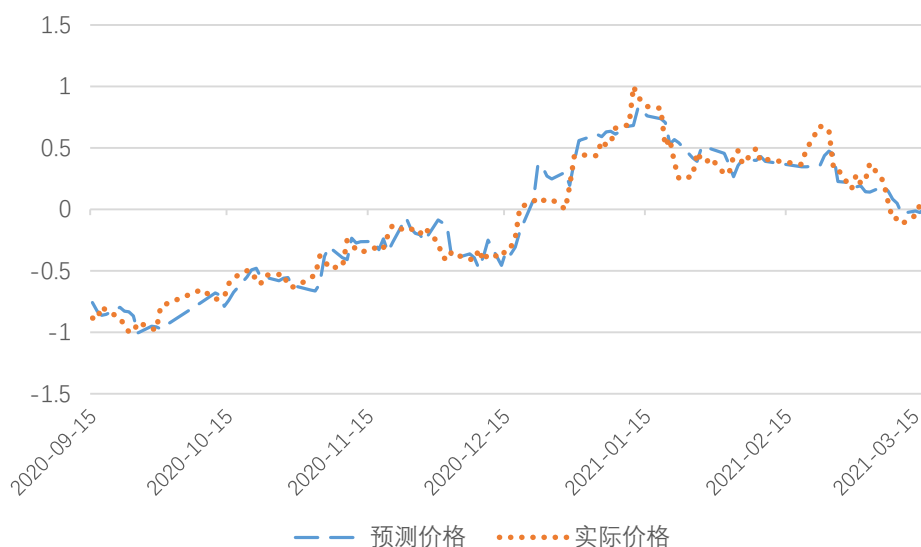


图 4-5M2109 的价格

然后通过预测价格计算得到预测价差，表 4.4 为预测价差同实际价差的误差。

表 4.4 预测价差的误差

预测模型	MSE	MAE	MRE
卡尔曼滤波-BP 神经网络	0.0476	0.0181	0.0228

从表中可以看出基于卡尔曼滤波—BP 神经网络模型能够较好地预测出价差，随后利用预测价差构建跨期套利策略组合。表 4.5 为基于卡尔曼滤波—BP 神经网络模型构建跨期套利的结果，其中 σ 为根据过去 30 个交易日计算得到的价差标准差。

表 4.5 跨期套利结果

b	0.25σ	0.5σ	0.75σ	1σ
交易次数	30	26	17	9
胜率	50%	46.15%	58.82%	77.78%
年化收益率	18.65%	27.94%	46.35%	35.44%
最大回测比率	0.0131	0.0053	0.0038	0.0127
夏普比率	1.8172	1.6617	1.5108	1.4950

4.4 结论

根据表 4.5 可以看出，基于卡尔曼滤波—BP 神经网络模型构建的跨期套利策略发现套利机会多于标准距离法跨期套利。且随着 b 值增大，交易机会逐渐减少。从胜

率来看,卡尔曼滤波—BP神经网络模型套利胜率略优于标准距离法套利,从盈利能力来看,当 b 值大于 0.25σ 后,卡尔曼滤波—BP神经网络模型套利的盈利能力要强于标准距离法套利。综上所述,基于卡尔曼滤波—BP神经网络模型构跨期套利策略是可行的,且表现优于传统套利方法。

通过价格预测模型可以指导传统量化策略,提高传统量化策略的成功率和减少损失。如投资者进行一个牛市套利,持有一个近月合约的多头和一个远月合约的空头。当基于价格预测模型发现未来近月合约将会下跌而远月合约价格将会上涨时,投资者可以及时平仓,避免受到较大的损失。相对传统的跨期套利模型中,价差由2倍标准差扩大到3倍标准差才进行触动平仓条件,基于价格预测模型的平仓条件更加提前,且损失少。不过也有因为价格预测模型不准确导致提前平仓降低收益的可能性。

5 总结与展望

5.1 全文研究总结

本文总结国内外对商品期货价格预测方面的研究,发现传统线性预测模型对期货价格预测存在滞后性且预测效果不佳,目前主流方法是应用机器学习算法预测或者将机器学习算法和传统预测模型相结合预测。基于 BP 神经网络和卡尔曼滤波,本文的主要工作和结论如下:

(1) 基于卡尔曼滤波算法,改进 BP 神经网络,并将改进后的 BP 神经网络算法应用到商品期货市场中,选取豆粕、沥青、白银和 PVC 期货的主力合约作为研究对象,预测合约未来价格,并同其他预测模型相比较。证明卡尔曼滤波消除噪音的能力可以提高 BP 神经网络对输入样本的抗干扰能力和泛化能力。改进后的 BP 神经网络优于卡尔曼滤波和改进前的 BP 神经网络。

(2) 基于卡尔曼滤波-BP 神经网络构建跨期套利策略,并同基于协整的跨期套利投资策略相比较,基于神经网络的跨期套利策略能更好地发现潜在的套利机会,并规避潜在的风险。卡尔曼滤波-BP 神经网络模型对构建量化投资策略有一定的指导作用。且将预测模型和传统跨期套利策略相结合,能够更好地判断交易信号,规避潜在风险。

(3) 套利策略通过风险对冲,减少了遭受重大损失的可能性,且收益更加稳定。但是当价差结构发生变化时,投资者仍然可能遭受损失。在实际应用中,需要不断改善套利模型,增强对风险的预测能力。BP 神经网络模型依赖于神经网络的层数、学习算法和隐含层节点数,通过不断调整参数,可以增加预测模型的准确度,更好地帮助投资者构建量化策略。

5.2 存在问题

(1) 本文只是通过卡尔曼滤波算法改进 BP 神经网络的学习算法,将 BP 神经网络的权重调整过程转化为权重的最优估计问题,难以推广卡尔曼滤波同其他神经网络结合。如何将卡尔曼滤波消除噪音的性能应用到更多神经网络中,需要进一步学习和研究。

(2) 横向比较少,本文只将改进后的 BP 神经网络模型同卡尔曼滤波预测模型和原 BP 神经网络预测模型进行比较,没有同其他预测模型进行对比分析。如果 BP

神经网络模型本身较其他预测模型的预测能力相差较大时，即使卡尔曼滤波能改善 BP 神经网络的预测能力，最后得到的改进预测模型也不一定由于其他预测模型。

（2）在构建跨期套利策略时，大部分研究者都是通过历史分钟数据来构建的，但基于数据来源的限制，本文采用了日数据，所以样本量较少，构建策略套利的机会也更少。且跨期套利策略中的交易成本简化为手续费，在实际中还存在更多的成本，比如说保证金带来的利息损失、难以按照预定的价格买入或者卖出合约等等。所以本文只能在一定程度上证明卡尔曼滤波-BP 神经网络在实践中是可行的，能够构建量化策略，为投资者提供参考意见。

5.3 未来工作展望

BP 神经网络是一个常用的神经网络，有较多关于预测方向的研究。通过卡尔曼滤波改进 BP 神经网络，实证发现改进后的预测模型优于原预测模型，说明 BP 神经网络仍然有继续提升的空间。关于卡尔曼滤波及 BP 神经网络对商品期货价格预测研究，未来可以从以下几个方面进行：

（1）输入变量。输入变量作为直接影响预测结果的因素，研究人员挖掘了多种选取指标的方式。本文仅根据七个变量进行预测，。在后续研究中可以考虑增加其他变量如技术指标等，找到更好的变量选择标准来选取变量，提升预测模型的预测准确度和泛化能力。

（2）样本。本文仅选取四种商品期货主力合约来进行实证研究，研究样本种类较少。未来研究可以针对不同种类的商品期货价格预测来开展。

（3）神经网络。通过将卡尔曼滤波算法同其他神经网络算法结合，增强其他神经网络算法对样本的抗干扰能力，以及对样本外数据的拟合能力。

致谢

不知不觉中，我的研究生生涯即将结束，两年时间竟是如此短暂。我在离校一年后再次考回华科，尽管熟悉的同学都已毕业，但是熟悉的环境还在。如今又要面临分别，再次回来不知是何时，但是母校的人和事都刻在我心，永不忘怀。感谢华中科技大学，给我提供一个良好的平台，让我不断学习、不断升华自己。无论是专业知识，还是做人道理，都让我受益匪浅，这是我人生道路上最宝贵的财富。感谢那些引导我、激励我和帮助我的人，我的内心充满无限感激！

首先，非常感谢我的导师张学功老师，是他指引我的学术方向，并拨开我论文研究过程中的迷雾。论文从开题、定题以及定稿，张老师都倾注了大量心血。尽管导师事情繁重，但总能抽出时间回答我的问题，帮助我快速学到新知识。张老师温和的为人处世和严谨的治学态度，润物细无声，不断浸润着我的学习和生活。再一次感谢张老师，感谢两年的照顾和关怀！

同时也非常感谢我的师姐们，谢谢你们传授学习经验，让我避开了很多弯路。同时感谢我的同窗好友，有你们的陪伴是我一直走下去的动力。特别感谢丘同学，本科时已是同窗，研究生时再次一起毕业，无论是学习还是生活，你都给我很大的助力。兄弟友谊，不会忘怀！还要感谢女排，感谢你们让我重新爱上排球，让原本枯燥无味的时光变得丰富多彩。在华科，遇到太多好人好事，感谢你们一路陪伴！

最后，感谢我最可爱的父母，你们的默默付出，是我前进道路上最坚实的支柱。无论我做出什么选择，你们的无理由的支持，让我不会畏惧路上的重重困难，你们永远是我前进的动力和温暖的港湾！

参考文献

- [1] Hassan R, Nath B, Kirley M. A fusion model of HMM, ANN and GA for stock market forecasting[J]. Expert Systems with Applications, 2007, 33(1): 171-180.
- [2] Selvin S, Vinayakumar R, Gopalakrishnan E A, et al. Stock price prediction using LSTM, RNN and CNN-sliding window model[C]. International Conference on Advances in Computing, 2017: 1643-1647.
- [3] Alsayed H, McGroarty F. Optimal Portfolio Selection in Nonlinear Arbitrage Spreads[J]. European Journal of Finance, 2010, 19(3) : 206-227.
- [4] Faragher R. Understanding the Basis of the Kalman Filter Via a Simple and Intuitive Derivation [Lecture Notes][J]. IEEE Signal Processing Magazine, 2012, 29(5): 128-132.
- [5] Zhuge Q, Xu L, Zhang G. LSTM neural network with emotional analysis for prediction of stock price[J]. Engineering Letters, 2017, 25(2): 167-175.
- [6] Galanis, Louka, Katsafados, et al. Applications of Kalman filters based on non-linear functions to numerical weather predictions[J]. Annales Geophysicae, 2006, 24(10): 2451-2460.
- [7] Yu S, Zhu K, Diao F. A dynamic all parameters adaptive BP neural networks model and its application on oil reservoir prediction[J]. Applied Mathematics & Computation, 2008, 195(1): 66-75.
- [8] Cheng Y, Liu Z. Optimized selection of sigma points in the unscented Kalman filter[J]. IEEE, 2011: 3073-3075.
- [9] Mackinley A C. Index-futures arbitrage and the behavior of stock index futures prices[J]. Review of Financial Studies, 1988, 1(2): 137-158.
- [10] Kalman R E. A New Approach To Linear Filtering and Prediction Problems[J]. Journal of Basic Engineering, 1960, 82D: 35-45.
- [11] Baur D G, Beckmann J, Czudaj R. A melting pot — Gold price forecasts under model and parameter uncertainty[J]. International Review of Financial Analysis, 2016, 48(DEC.): 282-291.
- [12] Haiyan, Mo, Jun, et al. Exponent back propagation neural network forecasting for financial cross-correlation relationship[J]. Expert Systems with Applications, 2016, 53(1): 106-116.
- [13] Hassan M R, Nath B. Stock market forecasting using hidden Markov model: a new approach[C]. International Conference on Intelligent Systems Design & Applications, 2005: 192-196.
- [14] Xiao H, Prokhorov D V, Donald I I J N. Time series prediction with a weighted bidirectional multi-stream extended Kalman filter[J], 2007, 70(13-15): 2392-2399.
- [15] 丁鹏. 量化投资:策略与技术[M]. 北京: 电子工业出版社, 2014: 2-4.
- [16] 朱丽蓉, 苏辛, 周勇. 基于我国期货市场的跨期套利研究[J]. 运筹与管理, 2015, 000(003): 179-188.
- [17] 孙冰洁, 唐瑞, 左毅, 等. 小波分析下的神经网络股票预测研究[J]. 计算机与数字工程, 2016, 44(6): 1031-1034.
- [18] 向小东. 基于小波神经网络的非线性组合预测方法及其应用研究[J]. 现代管理科学, 2007(06): 32-33.
- [19] 李东, 苏小红, 马双玉. 基于新维灰色马尔科夫模型的股价预测算法[J]. 哈尔滨工业大学学报, 2003(02): 244-248.
- [20] 骆桦, 陈艳飞. 基于小波与动态 GM(1,1)-ARIMA 模型的股价预测研究[J]. 浙江理工大学学报,

- 2017(4): 575-579.
- [21] 廉小虎, 姜铁兵. 基于改进 GM(1,1)模型的电力市场期货价格的预测[J]. 水电能源科学, 2006(1): 19-21.
- [22] 彭继兵, 唐春艳. 基于变维分形理论的卡尔曼滤波实时跟踪预测模型在股票价格预测中的应用[J]. 计算机工程与应用, 2005, 041(013): 218-220,223.
- [23] 潘迪夫, 刘辉, 李燕飞. 基于时间序列分析和卡尔曼滤波算法的风电场风速预测优化模型[J]. 电网技术, 2008, 32(7): 82-86.
- [24] 舒服华, 马勇军. 卡尔曼滤波算法在我国钢产量预测中的运用[J]. 徐州工程学院学报(自然科学版), 2018, 33(002): 8-12.
- [25] 谢合亮, 张砣. 卡尔曼滤波在高频金融时间序列模型预测中的应用[J]. 统计与决策, 2017, 000(013): 82-84.
- [26] 张智勇, 张丹丹, 贾建林, 等. 基于改进卡尔曼滤波的轨道交通站台短时客流预测[J]. 武汉理工大学学报(交通科学与工程版), 2017, 06(v.41): 84-87.
- [27] 曹广福, 路群. Kalman 滤波在股票交易模型中的应用[J]. 松辽学刊(自然科学版), 1999(4): 9-12.
- [28] 王法胜, 郭权. 基于扩展卡尔曼粒子滤波算法的神经网络训练[J]. 计算机工程与科学, 2010, 32(5): 48-50.
- [29] 池丽旭, 张广胜, 庄新田, 等. 投资者情绪指标与股票市场——基于扩展卡尔曼滤波方法的研究[J]. 管理工程学报, 2012(03): 126-132+169.
- [30] 刘明堂, 田壮壮, 齐慧勤, 等. 基于 Kalman-BP 协同融合模型的含沙量测量[J], 2016(5): 970-977.
- [31] 李明明, 王旭皓, 孙舒曼, 等. 基于 ARIMA 与季节指数组合模型的居民储蓄存款预测[J]. 经济研究导刊, 2015, 000(005): 192-193.
- [32] 沈腾飞. 中国市场条件下股指期货,ETF 期权与 ETF 的套利策略分析[J]. 天津财经大学, 2015.
- [33] 宋家骥. 我国 ETF 套利策略研究[D]. 复旦大学, 2014.
- [34] 梁媚. 中国股指期货期现套利研究及策略设计[D]. 浙江大学, 2017.
- [35] 巩兰杰, 张龙斌. 一种考虑基差非对称影响的期货波动性预测模型研究——基于上海铜期货市场的实证分析[J]. 北京理工大学学报(社会科学版), 2008(04): 27-30.
- [36] 王一多, 谢达. 基于 BP 神经网络的黄金期货价格预测[J]. 内蒙古煤炭经济, 2013, 000(008): 26-27.
- [37] 肖国荣. BP 神经网络在基金价格预测中的应用研究[J]. 计算机仿真, 2011, 28(003): 373-376.
- [38] 王伟国, 赵新民. 基于 EMD-BP 神经网络的我国棉花期货价格预测方法研究[J]. 石河子大学学报(哲学社会科学版), 2013(01): 84-86.
- [39] 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(12): 26-29.
- [40] 李留洋. 沪深 300 股指期货跨期套利策略及其有效性研究[D]. 西南大学, 2019.
- [41] 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(03): 70-85.
- [42] 尹力博, 杨清元, 韩立岩. 技术指标能够预测商品期货价格吗?来自中国的证据[J]. 管理科学学报, 2018, 21(006): 99-109.
- [43] 王健, 黄祖辉. 我国大豆期货市场有效性的实证研究[J]. 商业研究, 2007, 000(007): 190-194.
- [44] 周亮. 投资者情绪对商品期货价格及波动率的影响研究——以螺纹钢期货为例[J]. 武汉金融, 2019(6): 27-32.