

山东大学中泰证券金融研究院 2016 年项目结题报告

课题名称： 量化策略研究

山大课题负责人： 王 鑫

山大课题组成员： 郭禄禄、张志鑫、李艺迪
刘雅坤、孙佳欣、李 杨
李树阳、滕 斌

公司课题负责人： 王 伟

公司课题对接人： 王 伟

课题立项时间 ： 2016.4

中泰证券股份有限公司

山东大学中泰证券金融研究院课题结题报告

课题报告

2016 年度，课题组围绕“量化策略研究”这一课题，先后就股票配对交易策略和期权做市商策略两个问题进行了研究。

围绕股票配对交易策略这一主题，我们首先对我国 A 股市场进行了行业间、个股间、个股与大盘间的相关性分析。我们发现，专用设备、化学制品、通用机械三个行业在整个市场中重要的地位，个股与大盘的相关系数分布存在明显的尖峰厚尾现象等。

基于相关性研究的结果，在经典的配对交易原理的基础上，参考了这一领域近年发表在顶级量化金融期刊上的文献研究进展，我们采用比相关系数更稳健的欧式距离方法构建了配对交易策略。我们在沪深 300 指数成分股最近七年的历史数据上进行策略回测，结果表明，该策略在我国市场是有效的，策略具有稳定的盈利能力，回撤很小，在部分股票组合上，策略实现了 11% 的年化收益。配对交易收益较好的股票对有比较明显的聚集效应，在金融、工业、可选消费等行业中选取股票对有更大可能取得好的收益，此外，配对交易的收益取决于股票对价差的变化情况，与大盘走势相关性小。

围绕期权做市商策略这一主题，基于制度、需求和市场数据调研分析的结果，我们构建了理论上合理的报价逻辑，我们认为合理的报价逻辑和策略需要对市场行情的变化做出应对和调整，在同一类合约上的报价应具有一致性，而在不同类期权上的报价应该具有差异性，同时，做市商的报价应该根据希腊值的变化情况做出相应的调整，以实现合理的风险管控。（具体课题报告见附页）

课题承担单位公章

课题负责人（签字）_____

年 月 日

中泰证券课题立项单位评审意见

一、预期研究目标达成程度

对量化策略就中国 A 股市场进行了实证分析，得到了一些比较有趣的结论，预期研究目标达成，接下来对这些量化策略能否转化成稳定有效的实盘策略，还要再进行进一步的观察和分析。

二、课题报告中存在的问题

对一些局部特性的统计分析中偶尔会出现结论明显不正确的现象，希望今后研究中，编程的时候对一些数据的整理、是否用到未来数据及程序的编写逻辑上更细致一些。

三、后续是否有继续开展此项研究的必要（是否结题）

量化策略研究是一个很大的课题，策略范围广泛，需要持续进行研究。



公司课题负责人

2017年 7 月 3日

(纸面不敷，可另增独立页)

量化策略研究项目

总结报告

2016/12/23

项目负责人： 王 鑫
郭禄禄、张志鑫、李艺迪
项目组成员： 刘雅坤、孙佳欣、李 杨
李树阳、滕 斌
对接部门： 中泰证券衍生产品部

摘要.....	1
第一部分 配对交易策略研究.....	1
第 1 章 课题背景.....	1
第 2 章 问题描述.....	3
2.1 相关性研究的相关问题.....	3
2.2 统计套利策略的相关问题.....	4
第 3 章 研究思路.....	5
第 4 章 数据选取与处理.....	5
第 5 章 相关性分析与统计套利方法介绍.....	8
5.1 相关性的度量与分析.....	8
5.1.1 相关系数.....	8
5.1.2 距离.....	8
5.1.3 基于 Copula 函数的相关性分析.....	9
5.1.4 基于复杂网络的相关性分析.....	9
5.2 统计套利的研究方法.....	10
5.2.1 欧式距离的方法.....	10
5.2.2 协整检验的方法.....	10
5.2.3 时间序列的方法.....	11
5.2.4 随机控制的方法.....	11
5.2.5 其他方法.....	12
第 6 章 相关性分析.....	12
6.1 相关性度量指标比较与个股间相关性分析.....	12
6.2 行业间相关性分析.....	15
6.3 个股与大盘相关性分析.....	18
第 7 章 配对交易策略构建.....	20
第 8 章 实证结果与分析.....	22
8.1 策略收益与个股间距离的相关关系.....	22
8.2 优势股票对的分布情况.....	24
8.3 配对交易策略收益与大盘走势的相关性分析.....	25
8.4 结果总结与评价.....	26
第 9 章 总结.....	27
第二部分 期权做市策略研究.....	28
第 10 章 背景与问题.....	28
第 11 章 研究思路与方法.....	29

第 12 章 资料与数据分析.....	29
12.1 做市商基本义务.....	29
12.2 做市商基本需求.....	30
12.3 市场数据分析.....	30
第 13 章 报价逻辑构建.....	30
第 14 章 总结.....	32
第三部分 项目总结与展望.....	33

摘要

量化策略研究项目围绕股票配对交易策略和期权做市商策略两个问题进行研究。股票配对交易策略研究课题的目的是研究适合 A 股市场的量化交易策略。一方面，我们分析 A 股市场的数据特点，通过相关性分析等数学方法对交易数据进行深入了解，我们分析了常用的相关性分析指标的特点，详细分析了我国 A 股市场个股间相关性的特点、行业间相关性的规律以及个股与大盘的相关性情况，这些结果可以对交易策略的构建提供数据支撑。另一方面，我们基于相关性分析的结果，参考经典文献中配对交易的做法，用距离的方法构建了一个逻辑简单的配对交易策略，并在 A 股市场上进行策略有效性的检验。围绕期权做市商策略，我们进行了做市制度解读与做市商需求分析、文献资料学习与市场数据分析、报价逻辑与策略构建等研究环节，通过对做市政策制度、做市商做市需求、经典做市策略方法以及市场行情数据的全方位分析，我们给出了理论上合理的报价逻辑与策略。

第一部分 配对交易策略研究

第1章 课题背景

配对交易是一种统计套利策略。统计套利是通过对历史数据进行统计分析，寻找资产价格或收益率的某些变动规律，然后利用对冲手段构建投资组合捕获资产价格对其基本价值的暂时偏离而获取收益。统计套利技术基于定价理论、统计决策理论、博弈论、统计模式识别技术、时间序列分析方法、计量经济建模、现代计算方法等知识，是一种多学科的综合投资技术。统计套利是一种重要的市场中性策略。

Hogan 等在 2004 年放宽传统的无风险套利对风险的要求，给出了公式化的统计套利定义：

设初始投入为 0 的自融资交易策略，在时刻 t 经无风险利率折现后的价格为 $v(t)$ ，如果 $v(t)$ 满足：

$$v(0) = 0$$

$$\lim_{t \rightarrow \infty} E(v(t)) > 0$$

$$\lim_{t \rightarrow \infty} P(v(t) < 0) = 0$$

若 $\forall t < 0, P(v(t) < 0) > 0$, 则

$$\lim_{t \rightarrow \infty} \frac{Var(v(t))}{t} = 0$$

则称该策略为一个统计套利策略。

统计套利必需满足四个条件：一是自融资交易策略的初始投入为 0；二是经无风险利率折现后的价值的极限值为正；三是损失概率趋于 0，可以通过组合的重新调整或控制空头和多头头寸总额的比例来避免过度的净头寸暴露；四是如果损失概率在一定时间内不为 0，随着时间的推移，损失的概率趋于 0；第四个条件当且仅当损失概率大于 0 时成立，这说明统计套利是有风险的，即存在损失的可能，会因为投资者的判断失误或突发事件而造成投资者亏损。

统计套利主要有如下四种策略：配对交易策略、多因子套利策略、均值回归策略、协整套利策略。目前最成熟、应用最广泛的策略是配对交易策略和协整套利策略。

虽然统计套利的数学定义是 2004 年给出的，但统计套利的思想已经在实际交易中存在了很长时间。配对交易（Pairs Trading）的理念最早来源于上世纪 20 年代华尔街传奇交易员 Jesse Livermore 的姐妹股票对（sister stocks）交易策略。他首先在同一行业内选取业务相似，股价具备一定均衡关系的上市公司股票，然后做空近期的相对强势股，同时做多相对弱势股，等两者股价又回复均衡时，平掉所有仓位，了结交易。1985 年，Morgan Stanley 公司成立了一支由 Dr. Tartaglia 领导的量化团队，专门开展配对交易的研究，并于 1987 年投入实战，当年实现盈利 5000 万美元。不过该策略在之后两年连续亏损，研究团队被迫解散，小组成员散落到各家对冲基金，策略的思想也随之广为市场知晓。经过多年学术机构的研究和市场机构的实战，配对交易的理论框架和配套交易系统都日臻完善。

协整套利策略是指，利用协整分析寻找到价格或收益率序列长期存在的显著

协整关系的不同资产，即根据协整分析筛选配对股票进行套利。当短期内不同资产价格或收益率偏离长期均衡关系时，就可以通过买入低估值资产而卖空高估资产来进行统计套利。

据国外的研究，相对于其他量化投资方法，统计套利能够比较稳健地获得一定收益率，如果市场存在零初始成本、自融资的卖空机制的话，那么利用统计套利可以大大提升资产配置效率和市场有效性。与传统的无风险套利相比，统计套利是风险套利，因为基于历史数据得到的统计规律可能在未来不被遵从，不过研究发现，从长期来看，发生亏损的概率应该趋近于零。

在国内证券市场发展近二十年历程中，由于市场做空制度的缺乏，配对交易等统计套利方式并未为广大投资者熟知和使用。目前，该交易策略在国内证券市场实施的最后一块坚冰已被破除——融资融券业务已经在 A 股市场正式推出。尽管目前融资融券业务受到融券来源少、进入门槛较高、未对大部分机构投资者开放等瓶颈的制约，市场深度和广度均较难支撑配对交易的实现，但我们认为随着试点券商的逐批增加、融券标的证券范围扩大、转融通制度实施……，统计套利将在 A 股市场上被广泛应用。与之对应，对于该交易策略的研究也显得尤为重要。

第2章 问题描述

2.1 相关性研究的相关问题

在现代金融市场中，不同市场、不同资产的价格/收益率之间往往存在着不同程度的相关性，相关性分析在投资决策、风险管理等各个方面都有非常重要的研究意义和实际作用。基于相关性分析构建投资策略是重要的投资手段，相关性分析是众多投资策略的基础。统计套利策略，尤其是配对交易策略就是一类基于相关性分析的投资策略。该类策略经过超过二十年的发展，理论日趋成熟，逻辑也愈发完整，虽然限于种种原因目前在在中国市场发展落后，但随着我国金融市场的发展，相关性分析和统计套利策略的研究越来越受到投资者的重视。进行相

相关性研究并基于相关性分析构建投资策略已经成为投资者尤其是机构投资者日益增长的需求，相关的研究有很重要的现实意义。

作为本项目的第一个研究内容，我们首先进行相关性分析的研究，我们将相关问题归结为以下几个方面。

首先，哪种（些）相关性的研究方法在我国金融市场上是更有效的？相关性的度量指标和分析方法有很多，常用的主要有：相关系数、距离（欧氏距离、马氏距离等）、基于 Copula 函数的相关性分析、基于复杂网络的相关性分析等等。不同的方法尤其不同的特点和优势，分别适应不同的研究需求。那么，具体到金融资产价格或者收益率序列的相关性研究这一问题上，哪些指标是更简单有效的呢？

其次，基于第一个问题分析出的有效指标，对我国股票市场的相关性进行分析。具体来讲，第一，不同个股之间价格变化的相关性如何？第二，一般而言，同行业的股票之间会有较高的相关性，那么不同行业之间的相关性如何？哪些行业之间存在明显的相关性？第三，个股与大盘指数的相关性如何？

2.2 统计套利策略的相关问题

统计套利已有超过二十年的历史，其理论日趋成熟，结构也逐渐完善，相关的研究方法包括欧式距离的方法、协整检验的方法、时间序列的方法，随机控制的方法等都非常完整，也已被国外的学者、投资者验证有效。但是由于制度的限制，统计套利在中国资本市场的规模非常有限。同时，统计套利和配对交易的研究在国内是相对落后的，策略的有效性没有得到合理的验证。

我们认为，尽管统计套利策略的具体实践目前仍受到一定的制度限制，但其有效性的验证、构建能用于实际交易的统计套利策略仍具有很高的实际价值。因此，围绕统计套利策略的研究，我们提出核心问题：基于经典理论和经典方法，构建配对交易策略，在我国金融市场上进行有效性的验证，并结合实证结果，对策略进行修正，得到完整有效的交易策略。

第3章 研究思路

整体而言，我们要研究的主要问题：相关性分析和统计套利策略，目前已经有非常完整、成熟的理论，因此，项目研究的第一阶段，我们主要进行文献资料的调研和学习。通过学习经典文献，熟悉相关性分析的主要方法，掌握统计套利策略的主流研究方法，清楚地把握各类方法的内容和特点，从理论上分析各种方法的优缺点和所适合的研究需求。通过学习研究报告，了解目前市场上相关性分析和配对交易策略构建的主流做法，给我们的项目研究提供参考和指导。

基于文献资料的学习整理，作为统计套利策略构建的基础，项目研究的第二阶段，我们进行相关性研究的实证分析。针对我国 A 股市场进行一系列的相关性研究，利用具体的相关性分析结果，比较各种相关性分析方法的优劣，找到更简单有效的适合我国 A 股市场的相关性度量指标和方法。利用找到的度量方法，对我国 A 股市场进行具体的相关性研究，包括：个股之间的相关性如何，各个行业（可以按照申万二级行业的分类）之间的相关性如何，个股与大盘之间的相关性如何；这些相关性有无规律，出现高相关性的原因是什么；理论上分析能否依据高相关性的结果构建交易策略等等。

项目第三阶段，我们进行配对交易策略的具体研究。依据理论学习和相关性研究的结果，我们基于最经典的配对交易方法，构建配对交易策略，在我国 A 股市场上进行策略回测，检验策略的有效性，根据回测结果，对交易策略进行适当的修正，最终给出一个理论合理，实证有效的配对交易策略。

第4章 数据选取与处理

在构架配对交易策略时，我们选取所有沪深 300 的成分股为样本空间。沪深 300 指数的成分股是 A 股市场中流动性最强和规模最大的代表性股票，成分股之间的走势具有较强的相关性，作为交易标的也有很强的可操作性。

我们选取 2010 年 1 月至 2016 年 12 月近 7 年的时间作为回测区间，进行配对交易的研究。整个回测区间包括各种市场行情走势，既包括 2010 年至 2014 年

初市场震荡下行的牛皮市阶段，同时包括 2014 年至 2015 年上半年的牛市和 2015 年下半年开始的熊市阶段，因此我们选择这段时间是合理的。由于每只股票上市时间不同，我们选取沪深 300 成分股中 2010 年 1 月 1 日之前上市的股票，这样在整个回测区间每只股票的数据都是完整的，经过筛选我们发现共有 232 只股票的数据是可用的。

由于我们的研究中需要对股票进行分行业研究，因此，我们按照中证指数公司对沪深 300 指数成分股的行业分类（具体分类标准如表 4.1 所示），将所研究的 232 只股票对应到其所属的行业，这 232 只股票的行业分布情况如图 4.1 所示。

表 4.1 中证指数公司对沪深 300 指数成分股行业分类标准

行业	行业小类
能源	能源设备与服务、石油、天然气与消费用燃料
原材料	化学制品、建筑材料、容器与包装、金属与采矿、纸类与林业产品
工业	航空航天与国防、建筑产品、建筑与工程、电气设备、机械制造、工业集团企业、贸易公司与经销商、商业服务与商业用品
	航空货运与物流、航空公司、海运、公路与铁路、交通基本设施
可选消费	汽车与汽车零部件、家庭耐用消费品、休闲设备与用品、纺织品、服装与奢侈品、酒店、餐馆与休闲、综合消费者服务、媒体、经销商、百货商店、专营零售
主要消费	食品与主要用品零售、饮料、食品、家常用品、个人用品
医药卫生	医疗保健、设备与用品、医疗保健提供商与服务、制药、生物科技
金融地产	银行、保险、资本市场、房地产管理和开发
信息技术	软件与服务、技术硬件与设备、半导体产品与设备
电信业务	综合电信业务、无线电信业务、通信设备
公用事业	电力公用事业、燃气公用事业、复合型公用事业、水公用事业

所研究股票的行业分布

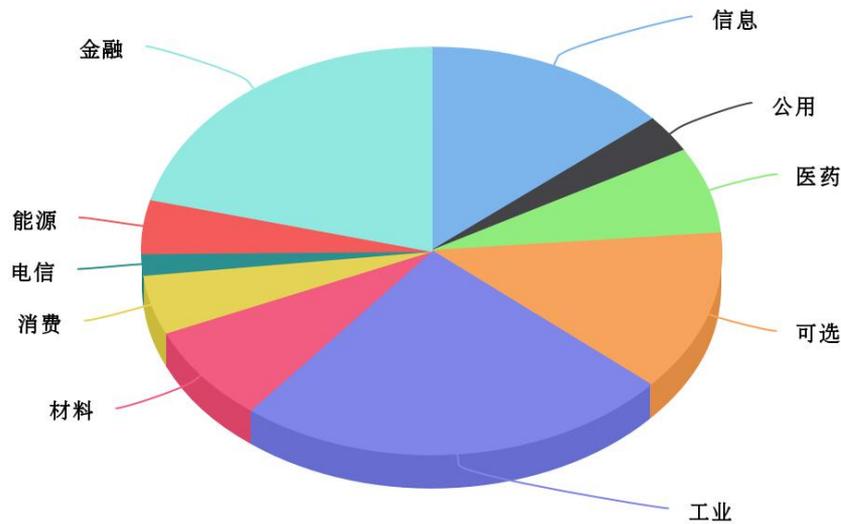


图 4.1 所研究 232 只股票行业分布情况

由于 A 股市场的 T+1 交易制度，无法实现日内交易，因此我们对个股日收盘价序列进行研究。为了使得所有股票的数据量相同，我们对个股停牌期间的股价进行插值处理。我们令股票停牌期间的股价为停牌前最后一个交易日的收盘价。

同时，为了使个股的股价是相对连续变动的，我们对所有股票的价格进行前复权处理。通过对股票原始数据的整理，我们得到了 232 个时间长度相同的股价时间序列。

不同个股之间股价差别较大，甚至有可能存在量级的差别，无法直接进行配对交易。我们对复权后的股价序列进行标准化处理，标准化的方法是对序列的标准差进行标准化，即将序列中的观察值减去该序列的平均数，然后除以该序列的标准差。经过标准差标准化后，股价序列的平均数为 0，标准差为 1。经标准化的数据是没单位的纯数量，消除了量纲（单位）影响和自身变异的影响。标准化的股价价差序列在 0 周围波动，是交易过程中计算交易参数和产生交易信号的数据。

在策略回测过程中，我们选取 12 个月作为样本内时间，即策略的生成区间，随后的 6 个月作为样本外区间进行交易，即策略的交易区间。同时，我们在整个回测区间采用滚动时间窗口的方式，高效利用所有的样本数据。滚动时间窗口的长度为 6 个月，使得回测区间样本外的时间是连续的。经过上述样本区间的划分，我们共有 12 个样本周期，由于 2010 年只能作为样本内区间，因此样本外时间从 2011 年 1 月开始，样本外总时间长度为 6 年。

第5章 相关性分析与统计套利方法介绍

5.1 相关性的度量与分析

常用的相关性的度量指标和分析方法主要有：相关系数、距离（欧氏距离、马氏距离等）、基于 Copula 函数的相关性分析、基于复杂网络的相关性分析等。

5.1.1 相关系数

相关系数是最为常见且常用的、也是最简单易实现的用以衡量相关性的指标。国内外研究者将股票之间的相关性定义为一段时间内股票市场中不同股票价格或是收益率的相关关系，一般用两者间的相关系数表示。相关系数的取值范围是 $[-1, 1]$ 。相关系数的绝对值越大，则表明所研究的对象之间相关性越高。根据相关系数数值大小，将不完全线性相关的密切程度分为四级： $0 \sim 0.3$ 为微弱相关； $0.3 \sim 0.5$ 为低度相关； $0.5 \sim 0.8$ 为中度相关； $0.8 \sim 1$ 为高度相关。在研究一组变量之间的两两相关性时，我们可以通过研究诸相关性的分布情况来研究相关性的相对大小。

5.1.2 距离

距离度量用于衡量个体在空间上存在的距离，标准化距离越大说明个体间的差异越大。常用的距离有欧氏距离、马氏距离等。

欧氏距离是最常见的距离度量，衡量的是多维空间中各个点之间的绝对距离，即每个对应分量值差的平方和再平方根，目的是计算其间的整体距离即不相似性。它将样品的不同属性（即各指标或各变量）之间的差别等同看待，这一点有时不能满足实际要求。因为计算是基于各维度特征的绝对数值，所以欧氏度量需要保证各维度指标在相同的刻度级别。

马氏距离表示数据的协方差距离。它是一种有效的计算两个未知样本集的相似度的方法。与欧式距离不同的是它考虑到各种特性之间的联系，并且是尺度

无关的，即独立于测量尺度。马氏距离不受量纲的影响，由标准化数据和中心化数据（即原始数据与均值之差）计算出的二点之间的马氏距离相同。

在计算马氏距离过程中，要求总体样本数大于样本的维数，否则得到的总体样本协方差矩阵逆矩阵不存在，这种情况下，用欧式距离来代替马氏距离，也可以理解为，如果样本数小于样本的维数，这种情况下求其中两个样本的距离，采用欧式距离计算即可。

5.1.3 基于 Copula 函数的相关性分析

根据 Sklar 理论，一个连续的多元分布函数，可以分成它的边缘分布与 Copula 函数，Copula 函数实际上是一种将联合分布与各自的边缘分布连接在一起的函数，因此也称为连接函数。基于 Copula 函数的相关性分析将问题转化为先确定边缘分布，做概率积分变换后，得到两者服从近似的均匀分布（上尾下尾有较强相关性），再用两步估计对 Copula 定参，之后从 Copula 可以得到相关性。

常用的相关系数是线性相关的度量指标，通常只在变量的线性变换下才不会发生改变，而由 Copula 函数导出的一致性和相关性测度，对于严格单调增的变换都不改变，因此应用范围和实用性更广。

5.1.4 基于复杂网络的相关性分析

具有自组织、自相似、吸引子、小世界、无标度中部分或全部性质的网络称为复杂网络。复杂网络有以下几个方面的特性。首先，复杂网络以简单的措辞描述了大多数网络尽管规模很大但是任意两个节（顶）点间有一条相当短的路径的事实。其次，集群即集聚程度。集聚程度的意义是网络集团化的程度；这是一种网络的内聚倾向。连通集团概念反映的是一个大网络中各集聚的小网络分布和相互联系的状况。最后，无标度特性。网络的度分布呈现指数分布的形式，更确切的说符合幂律分布。度指的是网络中某个节点（相当于一个个体）与其它节点关系（用网络中的边表达）的数量；度的相关性指节点之间关系的联系紧密性；介数是一个重要的全局几何量。

在股票相关性分析中，复杂网络有以下几个方面的应用。

1. 股票板块的划分。通过相关性构建的股票网络，能依据股票与股票间时间序列数据的相关性，从定量角度去划分股票板块。这样的量化处理使得板块内部的波动性更加一致。

2. 股票市场行业内部关键个股的节点作用。节点度服从幂律分布，幂函数曲线是一条下降相对缓慢的曲线，绝大部分节点度很小，只有极少数度极大的节点的存在，这些节点我们称之为 Hub 节点。Hub 节点通常数目很少，却对网络起着不可忽视的主导作用。其在股票市场运行过程中会对整个板块产生较大影

响,此类股票个体价格变化也会带动群体价格剧烈波动。节点中心性由节点介数、凝聚中心度、特征向量中心性刻画。节点的中心性可以体现节点在整个网络中的影响力。节点中心性较高的节点,其在整个网络中影响力极大。

5.2 统计套利的研究方法

目前统计套利有 5 种主流的方法:欧式距离的方法,协整检验的方法,时间序列的方法,随机控制的方法以及其他方法。

5.2.1 欧式距离的方法

距离的方法是在 Gatev 等人的开创性论文《Pairs Trading: Performance of a Relative-Value Arbitrage Rule》中最先提出的。该方法用欧式距离作为选取股票对的标准,实证结果也表明,这样的选取标准是有效的。

使用距离的方法优势很明显:没有使用经济模型,排除了模型使用中的错误估计、指标不明确等问题。欧式距离方法很容易实现,且数据探测稳健,在统计上有显著的风险调整后的超额收益。

然而,选择欧氏距离作为选择度量分析也有一定缺陷。我们假设一个理性投资者的目标是获取每对交易的超额收益。因此,配对交易者目标是找到价差的频繁剧烈的分离并且随后回归到均值的特性。换言之,利润最大化的理性投资者会寻找出具有较高方差的价差和较强均值回复的资产。这两个特性会增加配对交易的利润。

欧式距离为零的组合是不能产生利润的,那么我们为了实现利润最大化实际上是在选择一对次优的指标,限制最低的欧式距离等同于使得价差平方和最小与价差平方和均值最小,实证很难说明哪个主导了欧式距离最小化的过程,因此,欧式距离形成的方法间接造成了选出的股票配对组合有低价差波动的特性,会限制利润增长潜力。

5.2.2 协整检验的方法

Vidyamurthy (2004)提出了被引用最多的基于协整思想的配对交易。虽然当时缺乏实证检验,但是他们奠定了理论基础,后续有很多以协整关系为基础的交易策略研究。策略框架有三个关键步骤:

首先，预选出基于统计或基本结构相似如行业指标等的股票对。其次，可进行改进的 Engle - Granger 协整检验。最后，采用非参数方法优化入场/出场阈值。

5.2.3 时间序列的方法

在状态空间对价差进行建模，状态变量 x_k 表示价差，服从均值回复过程：

$$x_{k+1} - x_k = (a - bx_k)\tau + \sigma\sqrt{\tau}\varepsilon_{k+1}$$

$a, \sigma \in R^+, b \in R^+, \varepsilon_k \sim N(0,1), iid$ ，时间 $t_k = k\tau, k = 0, 1, 2, \dots$ 是离散的。均值为 a/b ，这里 b 为均值回复强度。方程也可以被写为：

$$x_{k+1} = a\tau + (1 - b\tau)x_k + \sigma\sqrt{\tau}\varepsilon_{k+1}$$

在连续时间下，可以用著名的 Ornstein-Uhlenbeck(OU)过程来描述这个过程：

$$dX_t = \rho(\mu - X_t)dt + \sigma dW_t$$

这里 dW_t 是一些概率空间下的标准布朗运动。参数 $\mu = a/b$ 表示 X_t 的均值， $\rho = b$ 代表均值回复的速度。状态空间模型的另一部分内容是度量方程，这里观测到的价差为状态变量 x_k 与高斯噪音 ω_k 的和：

$$y_k = x_k + D\omega_k, D > 0$$

在这个状态模型下，一个配对交易的入场条件为 $y_k \geq \mu + c(\sigma/\sqrt{2\rho})$ 或 $y_k \leq \mu - c(\sigma/\sqrt{2\rho})$ ，这里 c 为一个固定的常数，需要通过一定的方式进行确定。仓位的变化时间 T ，表示 OU 过程第一次穿过阈值的时间的解。

5.2.4 随机控制的方法

随机控制方法就是将长期资产分配在均值回复资产和无风险资产之间，在两者之间进行投资转换。它的重点在于确定最佳投资组合所持有的配对资产和其他资产的数量，在这一资产组合问题中，随机控制理论可以用来确定最优策略函数。根据描述资产价格模型的不同，这方面的研究主要有两种情形，一是资产价格用 OU 过程描述，另一种是资产价格用误差修正模型描述。

5.2.5 其他方法

其他的研究方法主要包括：机器学习与组合预测方法、Copula 方法、主成分分析方法等，这里不做具体的介绍。

第6章 相关性分析

6.1 相关性度量指标比较与个股间相关性分析

如前文所述，度量相关性的方法和指标有很多，我们首先研究这些指标刻画相关性的有效性。相关系数和欧氏距离是两个最常用的指标，因此，我们在这一部分通过研究个股数据，发现两种指标的关系，分析两种指标的特点，比较两种指标的优劣。本部分我们选取的数据的具体情况如下。

两个行业：证券行业与锂电池（概念）

时间区间：2015 年 8 月 27 日至 2016 年 8 月 17 日

数据频率：五分钟级别及日线级别

分析对象：股票价格及股票对数收益率

主要结果：相关系数矩阵及标准化后欧氏距离矩阵，具体如图 6.1 所示。

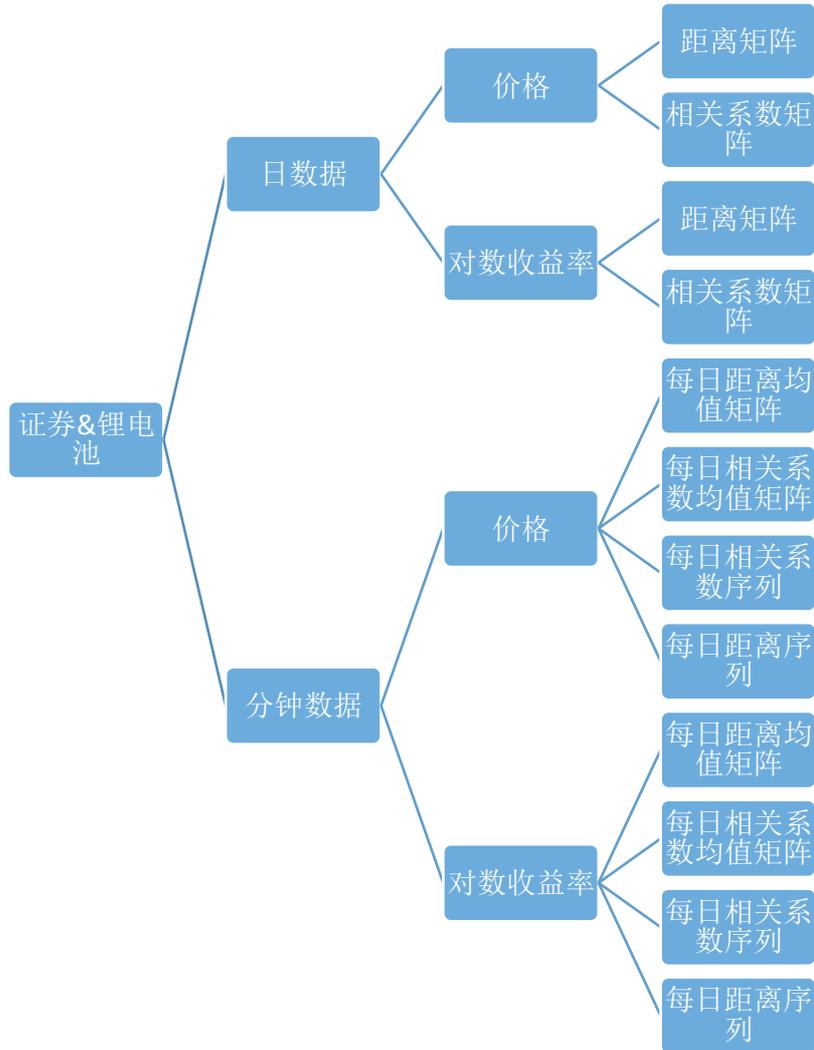


图 6.1 相关性度量指标比较分析的结果框架

通过分析具体的结果，我们得出来以下结论。

相关系数和距离具有较好的一致性。 同行业内股票两两之间的距离结果和相关系数结果显示，相关系数大的一对股票，其距离也较小，这也符合我们对于相关系数以及距离的理解，如图 6.2 所示。

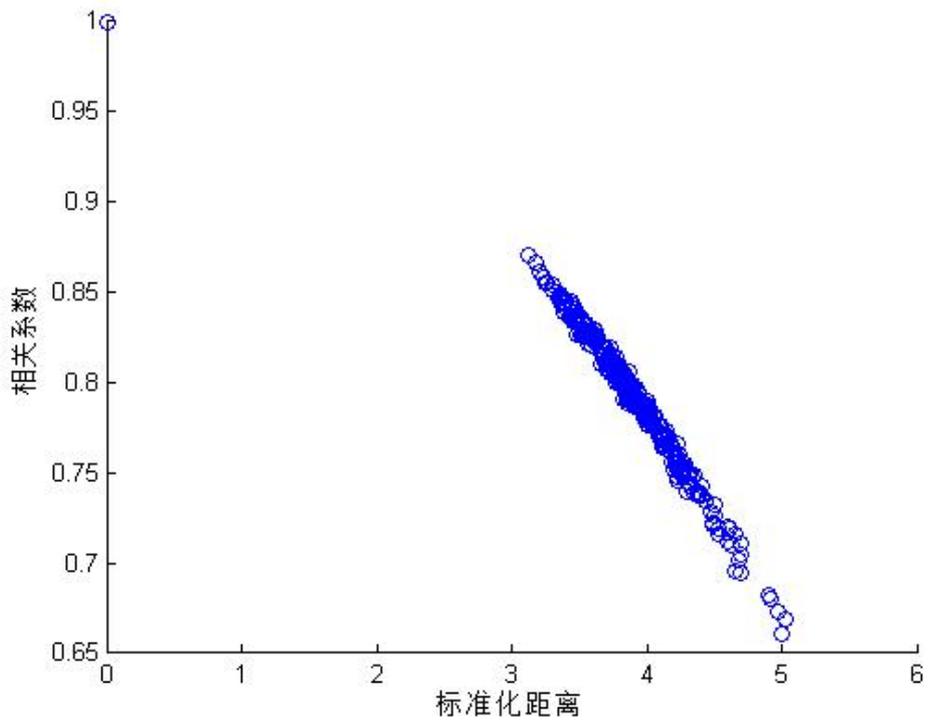


图 6.2-1 证券行业个股间相关系数与标准化距离的关系

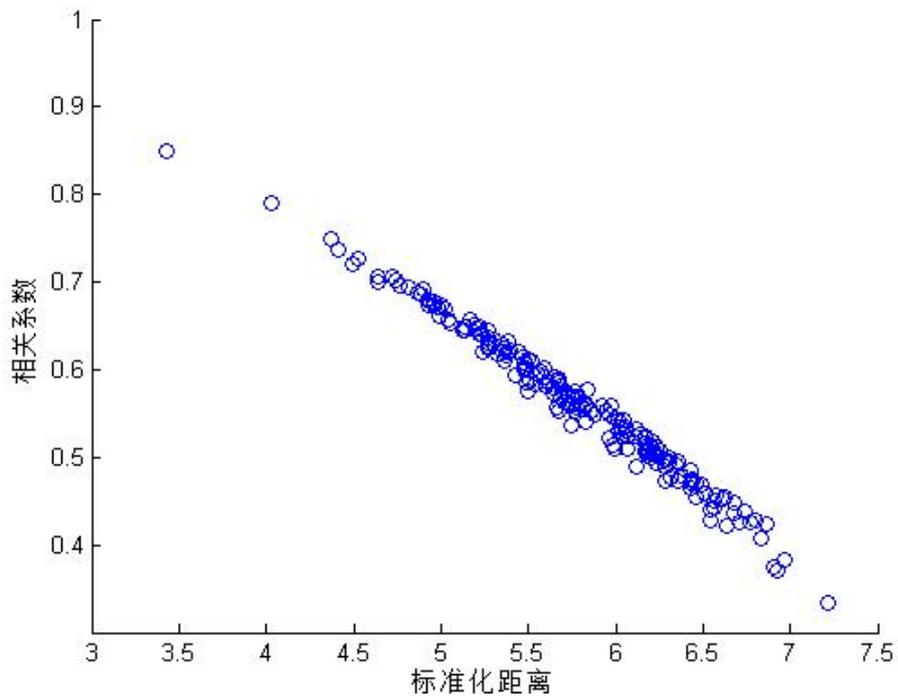


图 6.2-1 锂电池概念个股间相关系数与标准化距离的关系

这说明可以在一定程度上说明相关系数和距离间的一致性。即相关系数大的两只

股票之间的标准距离较小。对于锂电池概念，相关系数最大的股票组合是赣锋锂业和天齐锂业，赣锋锂业和多氟多。从距离矩阵上看，距离最小的股票组合也是赣锋锂业和天齐锂业，赣锋锂业和多氟多。对于证券板块，同样能体现相关性和距离的一致性。国金证券和西部证券的相关性最强，距离最小。

价格序列的相关性和收益率序列的相关性趋势高度一致。在证券和锂电池上用股价得到的相关系数阵与对数收益率得到的相关系数矩阵趋势是几乎相同的，只是数值上，应用对数收益率得到的略小。在距离矩阵上得到的结果与相关系数上的对比结果大体一致。

距离的波动性明显小于相关系数的波动性。我们推断距离比相关系数包含更多的信息，在一定程度上能更好的刻画相关性。

分析两个行业内个股间的相关性结果，我们得到了以下结论。

证券板块的相关系数和距离的表现明显优于锂电池概念。即证券板块个股之间的联动性更好。

相比在五分钟线上得到的结果，锂电池的五分钟数据上的相关系数与日线级别的相关系数差别还是很大的，总的来说，大部分日线上的相关系数要比于五分钟线上的相关系数大，而且不同股票对之间的相关系数变化也更为明显，具体来说，五分钟线上相关系数在 $0.4 \sim 0.6$ 之间波动，日线上没有这样明显的特征，而且在日线级别上相关系数出现了负值。类似的，证券上也有相似的结果出现，同样的股票对在日线上的相关性要明显强于五分钟线上的相关性。

这在一定程度上说明五分钟线上存在的噪音更多，那么相比在日线上的交易，它会存在更多的机会，当然也扩大了风险。

6.2 行业间相关性分析

在进行行业间相关性分析时，我们选取 2015 年 8 月 27 日至 2016 年 8 月 17 日 102 个申万二级行业的行业指数作为研究对象。我们着重用标准化的欧氏距离以及欧氏距离序列刻画行业指数之间的相关性。

专用设备、化学制品、通用机械三个行业的指数与其他大部分申万二级行业指数的相关性较高，银行、黄金、石油开采、保险四个行业与绝大多数申万二级行业的相关性较弱。

我们首先计算了 102 个申万二级行业的指数之间的欧式距离，共获得 5151 组数据结果。将这些距离从小到大排序，先选取前 5%（距离最小的 5%）和后 5%（距离最大的 5%）做初步的分析，各有 258 组数据。

距离最小的五组依次是：通用机械&专用设备、化学制品&专用设备、计算机设备&计算机应用、化学制品&通用机械、化学制品&汽车零部件。在前 5% 的距离中，出现次数最多的三个行业依次是专用设备、化学制品、通用机械。距离最大的五组依次是：银行&黄金、银行&农业综合、银行&餐饮、银行&地面兵装、银行&仪器仪表（均为银行与其他行业的距离）。在后 5% 的距离中，出现次数最多的四个行业依次是银行 II、黄金、石油开采、保险 II。

进一步，我们重点分析前 5% 的距离中占比较高的三个行业：专用设备、化学制品、通用机械与所有行业的距离以及后 5% 的距离中占比较高的四个行业：银行、黄金、石油开采、保险与所有行业的距离。

专用设备、化学制品、通用机械与所有行业的距离分布与所有行业距离的分布的对比如图 6.3 所示：

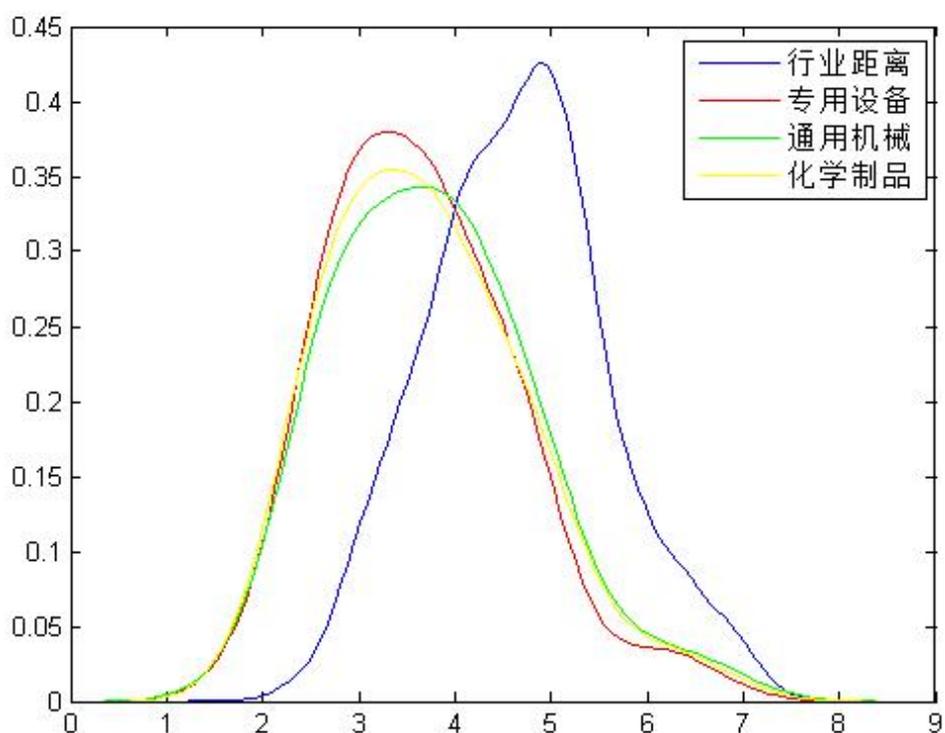


图 6.3 专用设备、化学制品、通用机械与所有行业的距离分布与所有行业距离的分布对比

专用设备、化学制品、通用机械三个行业的指数与其他大部分申万二级行业指数的距离都很小，三个行业之间的距离也非常小。专用设备、化学制品、通用机械包含的股票各有 123 只、159 只、105 只，我们认为，股票数量多，权重大以及支柱性的地位是以上三个行业与其他行业距离都较小的可能原因。专用设备与通用机械同属于一个申万一级行业，行业间距离较小符合直观认知，而专用设备与化学制品的高相关性可能具有较高的研究价值。

后 5% 的距离中占比较高的四个行业：银行、黄金、石油开采、保险与所有行业的距离分布与所有行业间距离的分布对比如图 6.4 所示。

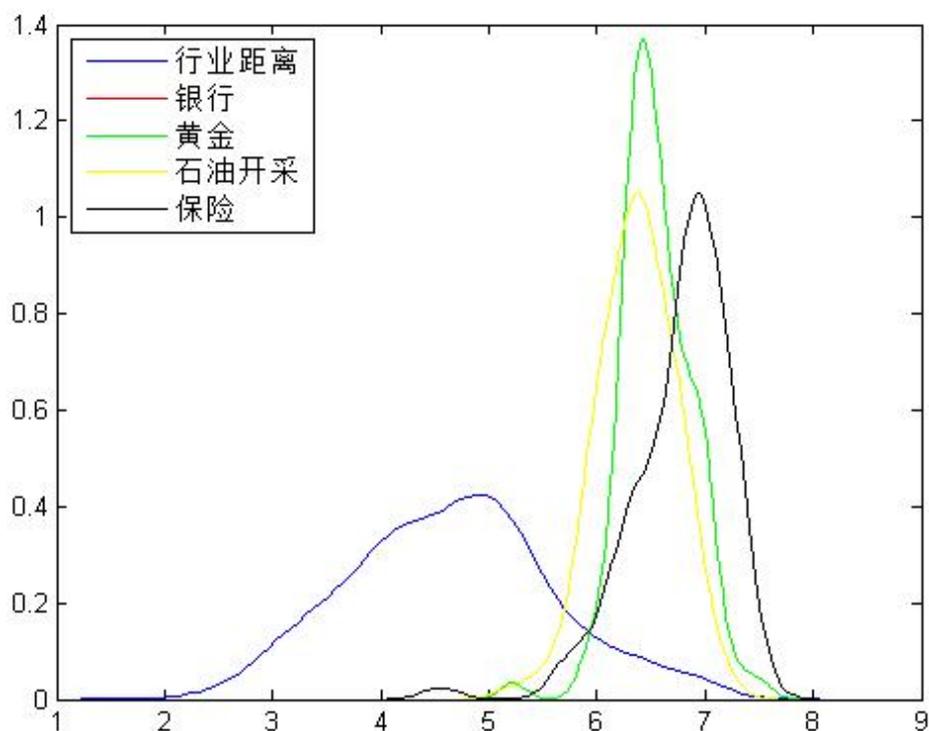


图 6.4 银行、黄金、石油开采、保险与所有行业的距离分布与所有行业间距离的分布对比

银行、黄金、石油开采、保险四个行业与绝大多数申万二级行业的距离都较大。过去一年中，银行、石油开采、保险的波动很小，而黄金的波动很大，这可能是以上四个行业与其他行业距离较大的原因之一。银行与黄金的距离在所有距离中是最大的，这两个行业的相关性我们认为也具有较高的研究价值。

6.3 个股与大盘相关性分析

我们用相关系数分析了所有个股与大盘的相关性，其分布如图 6.5 所示。

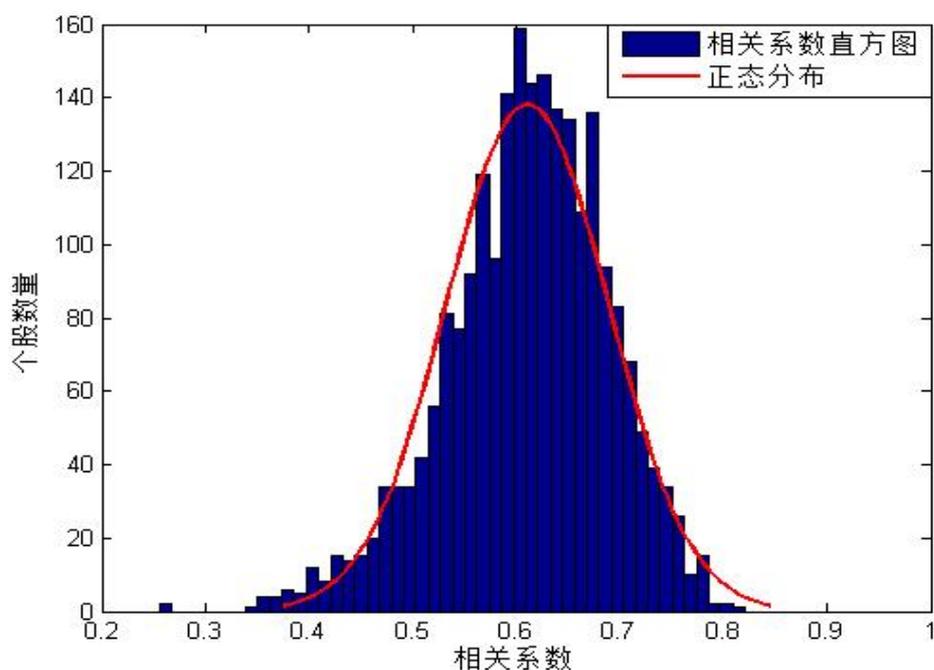


图 6.5 个股与大盘相关系数分布

从图 6.5 中我们可以看出，个股与大盘的相关系数分布接近正态分布，但也存在明显的尖峰厚尾现象。个股与大盘的相关系数集中在 0.5-0.7 之间。值得注意的是有少部分个股与大盘的相关系数小于 0.4，是较弱的相关关系。

通过简单的统计我们可以观察到，与大盘相关性最强的个股多来自证券或电力行业。这非常符合我们的直观。证券和电力属于权重板块，与大盘的相关性非常强。但是我们发现银行和石油等板块个股与大盘的相关性并不是太强，原因是银行和石油的波动性太弱。

与大盘相关性最弱的个股，部分来自黄金和有色行业。其他个股分布规律不明显。黄金和有色行业最近一年走势较强，短时间内波动也非常剧烈，因此与大盘的相关性较弱。这与我们在计算申万二级行业距离时得到的结果是一致的。另外，我们发现与大盘相关性较弱的个股往往在短时间内股价出现大幅波动，大涨或大跌的情况都有出现。

第7章 配对交易策略构建

我们构建配对交易策略的逻辑是从市场上找出历史股价走势相近的股票进行配对，即通过相关性筛选配对股票进行套利：当配对股票价差偏离历史均值时，则做空股价被高估的股票同时买进股价被低估的股票，等待价差回归到长期均衡关系，由此赚取价差收敛的收益。

我们以沪深 300 股票 2010 年 1 月至 2016 年 12 月的行情序列为样本空间，进行策略构建与回测，整个回测区间接近 7 年。经过简单筛选，我们发现共有 232 只股票的数据是完整可用的。在策略回测时，我们以 12 个月为样本内数据，优化开平仓参数，以随后的 6 个月为样本外数据，进行策略绩效的统计分析。并且在整个回测区间中，我们以 6 个月为周期时间窗口滚动，共产生 12 个样本区间。

策略构建的主要步骤如下：

1. 筛选股票组合

计算样本空间内所有股票在样本内的标准欧氏距离，筛选股票组合。根据我们相关性的研究，我们发现标准欧氏距离是刻画相关性的重要指标，因此我们利用标准欧氏距离筛选股票组合。

我们首先对原始股票日收盘价数据进行复权处理，使得股价时间序列是相对连续的。其次，我们对股价序列进行标准化处理，标准化的方法是对序列的标准差进行标准化，即将序列中的观察值减去该序列的平均数，然后除以该序列的标准差。经过标准差标准化后，股价序列的平均数为 0，标准差为 1。经标准化的数据是没单位的纯数量，消除了量纲（单位）影响和自身变异的影响。

2. 构建配对交易的回测系统

选定配对交易的开平仓参数，构建配对交易的回测系统。配对交易其实是一种多空投资组合，当配对股票的价差出现偏离时，做空股价暂时被高估的股票同时做多股票暂时被低估的股票，并在价差回归到一定水平时平仓获利。确定开平

仓参数是交易策略中的核心工作。

我们以配对股票标准化的价差序列作为标的的数据，计算开平仓参数并产生交易信号。通过经典文献的学习和我们的实际经验，我们选定标准化价差序列的两倍标准差作为开仓参数。在一个样本区间内，我们计算样本内标准化价差的标准差，作为样本外交易的开仓参数。由上述的标准化方法，可知标准化价差的均值为 0。当样本外标准化价差大于其两倍标准差时，做空价差；当标准化价差回复到 0 时，平仓获利了结头寸，即当价差大于两倍标准差时，做空第一只股票同时做多第二只股票；当价差回复到 0 时，同时平仓两只股票。反之，当样本外标准化价差小于负的两倍标准差时，做多价差。如果开仓后在样本区间结束前价差仍未回复到 0，以样本最后一天的收盘价计算盈亏。

3. 策略绩效统计分析

虽然每个样本周期样本外的时间长度为 6 个月，但是我们采用滚动窗口的方式，使得整个回测区间样本外时间是连续的。策略回测开始时，我们给每只股票配置相等的资金。这些资金在整个回测区间是可以再投资的，即每个样本区间内产生的收益在下一个样本区间内可以再投资，产生收益。

根据我们的回测方法，策略产生的收益是一系列在样本区间内随机分布的正的现金流，但是每个样本区间内的最后一笔盈亏是无法确定的。当然在样本区间内也可能没有现金流，因为在整个样本区间内价差的绝对值可能始终大于两倍的标准差。整个回测区间的前 12 个月，即 2010 年全年是无法交易的，因为这 12 个月是第一个样本区间的样本内数据。

我们计算每个股票组合在整个回测区间上的收益率以及收益率的统计性质，如收益率的均值、标准差、概率分布等。同时我们还会详细讨论配对股票的收益与其标准化欧氏距离之间的相关关系。最后，我们考虑配对交易策略的收益与大盘收益之间的关系。

第8章 实证结果与分析

8.1 策略收益与个股间距离的相关关系

我们将构建的配对交易策略在 232 只股票两两组合形成的 26796 组股票对上进行策略回测，统计分析 2011 年 1 月以来策略在各对股票上收益的相关情况。我们将股票对按照个股间标准化距离从小到达的顺序排列，各股票对中两股票间距离的分布如图 8.1 所示，相应的策略累计收益情况如图 8.2 所示。

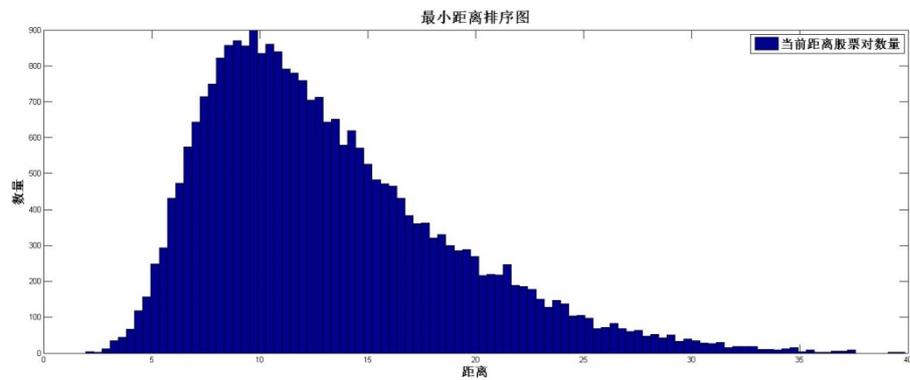


图 8.1 股票间距离分布直方图

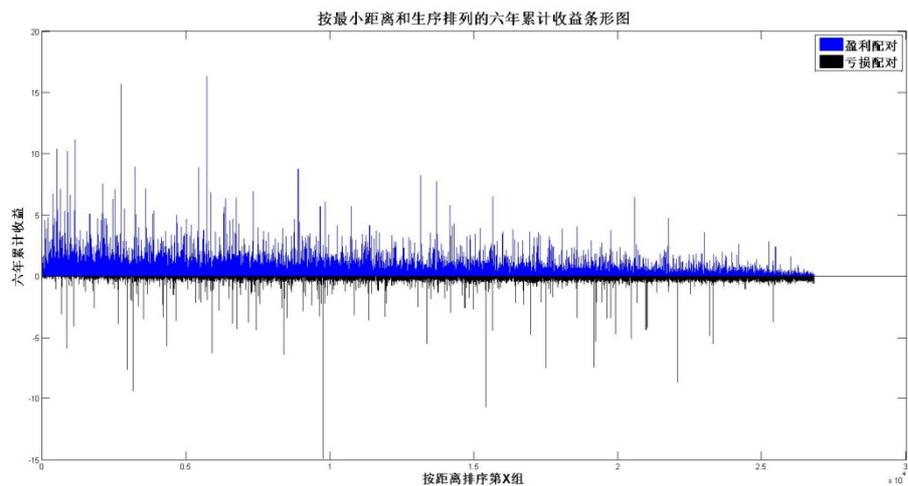


图 8.2 配对交易策略在每个股票对上的累计收益

图中横轴代表股票对编号，一共 26796 对股票，从左到右距离依次增大，纵轴表示相应股票对上策略自 2011 年 1 月以来的累计收益。从图 8.2 中不难发现，

在距离较小的股票对上，策略的整体收益、正收益比例明显高于在距离大的股票对上，随着距离的增大，策略的收益呈现明显的下降趋势。从这些宏观特点上分析，我们用距离的方法构建的配对交易策略是有效的。策略在每个股票对上的累计收益的分布直方图如图 8.3 所示。

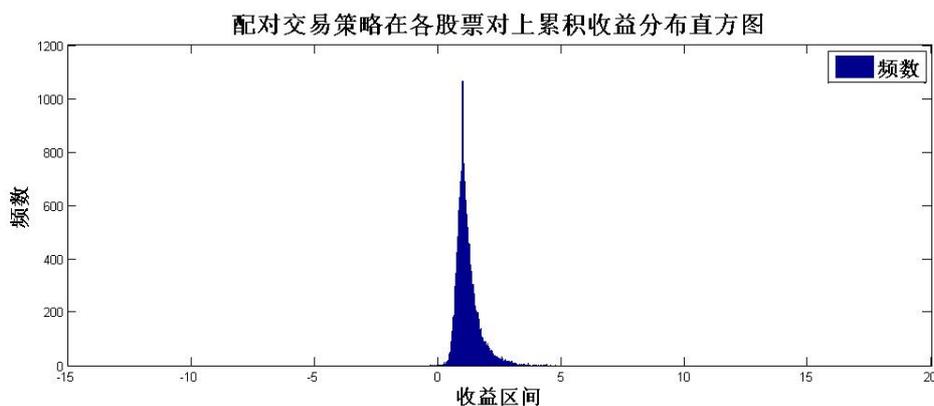


图 8.2 配对交易策略在各股票对上累积收益分布直方图

为分析个股间距离对策略收益的影响，我们在 26796 个股票对中重点分析四组股票对，分别是距离从小到大排列，前 100 对、前 1000 对、第 10001 对到 11000 对、第 25000 对到 26000 对，分析这四组股票对上策略收益的均值、正收益的比例、极值等指标，指标具体情况统计如表 8.1 所示。

表 8.1 配对交易策略在各组股票对上的收益情况统计

	前 100 对	前 1000 对	第 10001 对 至 11000 对	第 25001 对 至 26000 对	全部股票对
平均收益率	0.069	0.114	0.038	-0.008	0.036
正收益比例	0.940	0.913	0.632	0.351	0.609
中位数	0.274	0.432	0.110	-0.046	0.085
最大收益	2.223	10.359	5.670	2.829	16.302
最大亏损	-0.208	-5.916	-3.171	-3.742	-14.913

表中，所有的收益数据均为年化收益率，我们发现，前 1000 对股票上的平均年化收益率超过了 10%，这进一步证明了用距离方法构建的配对交易策略的有

效性。从图 8.1 的结果中，我们得出了“配对交易策略在距离小的股票对上的整体表现好于距离大的股票对”的结论，这一结论我们在上表中仍能得到验证，距离从小到大排序，前 100 对和前 1000 对上的策略收益明显优于第 10001 对至 11000 对和第 25001 对至 26000 对。但我们发现，距离最小的 100 对股票上，策略的收益情况并不是最好的，从策略在每组股票对上的平均收益、正收益比例、中位数等指标上来看，前 1000 对的结果全面优于前 100 对的结果。这些结论在各组股票对收益的概率密度分布对比图中也可以得到进一步验证，如图 8.4。

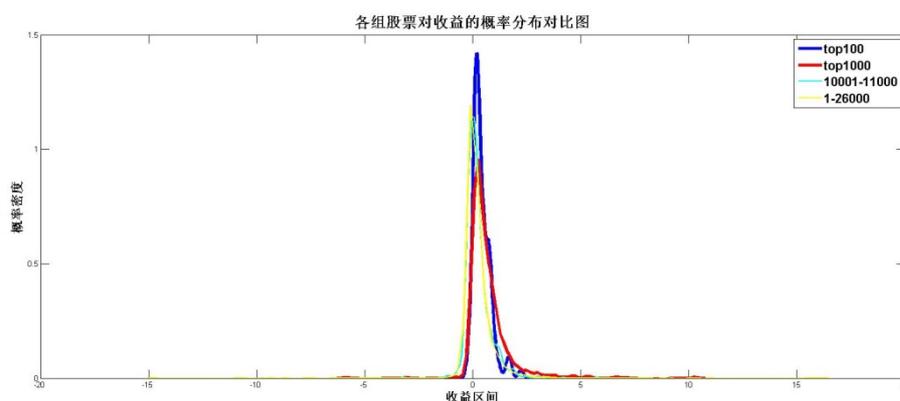


图 8.4 各组股票对收益的概率密度分布对比图

基于以上分析，我们可以得出如下几个结论。

1. 用距离的方法构建的配对交易策略是有效的。
2. 距离最小的股票对不一定是最适合做配对交易的。这也符合我们的直观认知，距离越小，价格差序列的波动也越小，因而出现的交易机会也少。

8.2 优势股票对的分布情况

本部分我们重点分析距离从小到大排序后的前 1000 对股票。基于以上的分析，配对交易策略在该组股票对上的整体收益是最好的，我们想分析这部分股票所属的行业或板块的情况，发现其中的聚集性规律，期望能找出这部分股票对集中分布的板块或行业，以给予实际交易选择备选股票池以相应的指导和建议。如果这部分股票对集中在某一个或几个行业，我们可以据此推断，这几个行业的股票更有可能通过配对交易获利。

仍然按照中证指数公司的行业分类，我们将 1000 对股票的所属的行业进行了统计，具体结果如表 8.2 所示。

表 8.2 距离最小的前 1000 对股票的行业分布

行业	信息	公用	医药	可选	工业	材料	消费	电信	能源	金融
信息	9	6	9	16	35	3	0	1	5	13
公用	6	3	5	18	29	2	0	1	4	42
医药	9	5	13	20	17	4	1	3	0	36
可选	16	18	20	19	59	5	3	7	7	103
工业	35	29	17	59	81	30	6	9	15	83
材料	3	2	4	5	30	18	3	8	8	14
消费	0	0	1	3	6	3	2	2	0	1
电信	1	1	3	7	9	8	2	1	1	9
能源	5	4	0	7	15	8	0	1	2	11
金融	13	42	36	103	83	14	1	9	11	198

上表中的数字是股票对数的统计，每一个位置的数字表示股票对的数量，这些股票对满足，一只股票属于它所在行的行业，另一只股票属于它所在列的行业，例如，表中左上角第一个数字 9 表示的含义是，在 1000 对股票中，有 9 个股票对完全属于信息行业。从表中我们不难看出，前 1000 对股票的聚集效果比较明显，大部分股票对（543 对）属于金融行业、材料行业和可选消费行业或他们的交叉。据此，我们推断，在金融、材料、可选消费行业按照距离选取股票对构建配对交易策略，有更大可能性获利。

8.3 配对交易策略收益与大盘走势的相关性分析

本部分，我们分析以上选取的每一组股票对的平均收益情况、所有股票对的平均收益情况与同时段内沪深 300 指数的行情走势作比较，分析配对交易收益与大盘走势的关系。

如图 8.2 所示，我们计算配对交易策略在前 100 对股票、前 1000 对股票、第 10001 对到 11000 对股票、第 25001 对到 26000 对股票以及全部 26796 对股票上的动态权益变化，与同时段沪深 300 指数行情作比较。

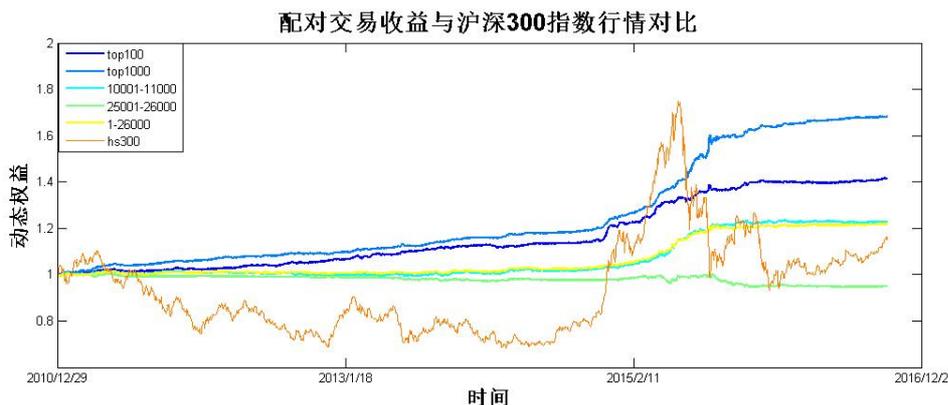


图 8.2 配对交易收益与沪深 300 指数行情对比

分析上图，我们可以得出以下结论。

上图更加直观的展示了我们基于距离构建的配对交易策略在四组股票对以及所有股票对上的表现情况，结果与我们统计的收益数据高度一致。距离较小的股票对的收益相较于距离大的股票对的收益优势明显，距离从小到大排序后的前 1000 对股票的平均收益情况要好于前 100 对股票的平均收益。

分析前 1000 对股票的平均收益的具体走势我们发现，配对交易的收益表现整体平稳。在大盘行情平稳时，配对交易能取得稳定收益；在大盘持续上涨时，配对交易仍能取得不错收益，但涨势弱于大盘走势；在大盘走势急剧下跌时，配对交易的收益有较大的涨势。基于此结果和配对交易的逻辑，我们认为，配对交易的收益取决于股票对价差的变化情况，与大盘走势关系不大。

8.4 结果总结与评价

基于以上结果的分析，我们得出以下三方面的结论。

1. 用距离的方法构建配对交易策略在我国股票市场是有效的，能够获得稳定收益。距离小的股票对上进行配对交易，能获得较好的收益，但距离最小的股票对上收益不见得最好。
2. 配对交易收益较好的股票对有比较明显的聚集效应，在金融、工业、可选消费等行业中选取股票对有更大可能取得好的收益。
3. 配对交易的收益取决于股票对价差的变化情况，与大盘走势关系不大。

基于相关性分析的结果，参考经典文献中配对交易策略的做法，我们构建了一个逻辑相对简单的配对交易策略。策略收益走势平稳，部分股票组合下，策略

年化收益率超过 10%，分析策略表现得出的以上结论也能为实际交易提供很多有价值的参考和支持。

第9章 总结

围绕相关性分析和配对交易策略，我们进行了理论学习、数据统计分析、模型构建、实证分析等一系列研究工作。

通过相关性分析，我们对 A 股市场的数据特点有了更深刻的了解。我们发现相关系数和标准欧式距离在刻画相关性时有高度一致性。即相关系数大的两只股票之间的标准距离较小。同时距离的波动性明显小于相关系数的波动性，我们认为距离在一定程度上更能刻画相关性。

在行业相关性分析中，我们利用标准欧氏距离构建复杂网络，可以发现专用设备、化学制品、通用机械是行业网络中的关键节点，即上述三个行业在整个市场中有重要的地位。

在个股与大盘的相关性分析中，数据结果显示个股与大盘的相关系数分布接近正态分布，但也存在明显的尖峰厚尾现象。与大盘相关性最强的个股多来自证券或电力行业。与大盘相关性最弱的个股，部分来自黄金和有色行业。其他个股分布规律不明显。这些相关性的结果大都符合我们的认知，也能为策略层面的研究提供数据支持。

基于相关性研究的诸多结果，参考经典的配对交易的做法，我们用距离的方法构建了配对交易策略，并在我国 A 股市场进行有效性检验，得出了非常有价值的结论。

首先，用距离的方法构建配对交易策略在我国股票市场是有效的，能够获得稳定收益。距离小的股票对上进行配对交易，能获得较好的收益，整体来看，距离小的股票对上配对交易策略的收益要明显好于距离大的股票对，但距离最小的股票对上收益不见得最好，这是因为配对交易要求股票的价差有一定的波动性，而距离很小的股票之间价差的波动相应的也很小。

其次，配对交易收益较好的股票对有比较明显的聚集效应，在金融、工业、可选消费等行业中选取股票对有更大可能取得好的收益。配对交易的收益取决于

股票对价差的变化情况，与大盘走势关系不大。

在本项目的研究中，我们对我国股票市场进行了大量的相关性研究，给出了一个逻辑相对简单的配对交易策略，策略的有效性也得到了验证。这些结果可以给实际的交易提供很多有价值的参考和支撑。基于已有的结果，我们计划在接下来的研究中充分利用相关性分析的结果，进一步完善策略的结构和逻辑，完善策略回测机制，对策略进行更加准确的回测，根据回测结果对策略进行修正和完善。

第二部分 期权做市策略研究

围绕期权做市商策略这一问题，我们先后进行了做市制度解读与做市商需求分析、文献调研学习与市场数据分析、做市商报价逻辑的建立等研究环节。具体情况如下。

第10章 背景与问题

2015年2月9日，中国第一个场内期权产品——上证50ETF期权合约品种在上海证券交易所上市交易，中国正式进入期权时代，随着我国金融市场的不断发展和相关制度的日益完善，50ETF期权上市以来，市场运行平稳有序，期权经济功能开始逐步得到发挥。我国的期权市场将有很大的发展空间和潜力。

50ETF期权自上市之初就引入了做市商制度。做市商与交易所签订做市协议，按照交易所的规定，履行报价义务，交易所根据做市商义务的履行情况对做市商给予手续费上的优惠。做市商在期权交易中的角色包括流动性的提供者、期权交易中的风险承担者和转移者、期权价格的稳定者等。中泰证券股份有限公司作为50ETF期权首批8家做市商之一，做市商策略的研究具有非常重要的现实意义。

50ETF期权交易采用在竞价交易制度下引入竞争性做市商的混合交易制度。期权做市商需按照交易所规定和要求，为期权合约提供流动性服务，包括向投资者提供双边持续报价，对投资者询价提供双边回应报价等。做市商在履行做市义务的过程中，容易被动地吃到合约订单，累积一定的期权头寸，承受市场风险，

也可能因为被动吃单导致亏损。因此，从做市策略的角度考虑，维持做市义务的前提下控制累积头寸数量、尽可能减小被动交易的亏损甚至从中获利是做市商的基本需求，如何围绕这一基本需求构建做市策略是做市商关心的问题。完整的做市策略是一个复杂庞大的体系，本课题选取这一体系中的基础环节——构建完整的报价逻辑作为研究重点，从做市商做市的义务和需求出发，通过分析市场数据，构建较为完整的做市商报价逻辑，并在此基础上构建报价策略。

第11章 研究思路与方法

我们按照如下研究思路展开研究。

第一、做市制度解读与做市商需求分析。我们先对交易所现行的做市商制度进行完整解读，明确做市商的基本义务，做市策略的构建需要符合做市商的基本需求，包括履行基本义务的需求、管理风险的需求等等。基于以上所述的基本义务和基本需求，建立合理可行的报价逻辑。

第二、文献资料学习与市场数据分析。通过文献资料的调研，研究经典的做市策略，并将其中有价值的研究成果与我国期权市场进行有机结合。另外，做市商的报价逻辑要根据市场行情做出相应的调整，反过来，做市商的报价又会影响市场价格走势。做市商的风险管理主要是结合希腊值进行管控，因此，我们需要对我国场内期权市场的期权合约数据进行宏观特点和微观结构的统计分析。

第三、建立报价逻辑与报价策略。基于做市商做市的基本需求，结合经典的做市策略，利用数据分析的结果，尝试构建合理可行的报价逻辑，兼顾履行做市义务、控制持仓头寸、管理账户风险、降低交易亏损等多种需求，并在报价逻辑的基础上完善报价策略。

第12章 资料与数据分析

12.1 做市商基本义务

做市商开展做市业务，应当符合交易所规定的做市报价指标要求，主要的指

标包括：最大买卖价差、最小报价数量、持续报价参与率、合约覆盖率、回应报价最长时限、回应报价保留最短时间、回应报价参与率等。

12.2 做市商基本需求

做市商的基本需求包括报价管理、风险管理等。做市商在履行做市义务的过程中，容易被动吃单，累积期权头寸，承受市场风险，也可能因为被动吃单导致亏损。因此，从做市策略的角度考虑，维持做市义务的前提下控制单边累积头寸数量、尽可能减小被动交易的亏损甚至从中获利是做市商的基本需求。

12.3 市场数据分析

我们详细分析了 2016 年 8 月、9 月市场上的大部分期权合约的相关数据，包括行情走势、行情相关性、合约希腊值等。我们分析了每只期权合约日间行情的宏观特点，同时基于期权价格的 tick 数据分析了期权合约价格日内变化的微观结构；我们将所有的期权合约按照期权的主要要素（看涨看跌、执行价格、到期日等）进行不同的分类，研究了每一类期权合约在宏观特点和微观结构上变化的一致性、各类期权之间的差异性；我们研究了期权价格走势与标的价格走势的相关性，并分析了不同类别期权价格走势与标的行情相关性的强弱差异；另外，我们分析了所有期权合约的希腊值数据，以给报价逻辑中的风险管理提供参考。

第13章 报价逻辑构建

基于以上分析，我们从做市商做市的基本义务和基本需求出发，结合经典理论经典方法，尝试构建了理论上合理的报价逻辑。我们认为，一个合理的报价逻辑，需要兼顾以下几个要素。

1. 报价逻辑能对市场行情的变化做出应对和调整。例如，对于单边趋势行情，通过报价调整进行一定的趋势追踪，避免过多被动吃单导致亏损。

2. 在同一类合约（如所有的接近平值的期权、所有的深度实值或虚值的期权等）上报价的一致性。数据分析结果表明，同一类期权在行情变化上具有很好的一致性，因此，做市商的报价逻辑在同一类期权上也应相同或接近。

3. 在不同类期权上逻辑的差异性。不同类型的期权合约，比如平值期权与

实值期权之间，行情变化会有较大差异性，那么相应的，做市商在不同类别的期权合约上的报价也应有差异性，这种差异性可以表现为逻辑本身能对期权行情做出准确识别并自动调整报价，也可以是做市商分别针对不同类别制定报价逻辑。

4. 希腊值的管理。风险管理是做市商的基本需求之一，通过希腊值管理风险是常用的手段之一，做市商的报价应根据希腊值的变化情况做出相应的调整。

具体的策略如下。

策略一：考虑报价位置和报价价差的策略

考虑报价位置和报价价差的策略直接利用市场订单数据和市场价格进行报价，具体可以有以下四个具体策略。

1. 直接用当前市场上的卖一、买一价格作为报单的 ask、bid 价进行报价；
2. 当前市场上的卖一、买一价格加/减一跳作为报单的 ask、bid 价进行报价；
3. 以最新成交价格为基准价，基准价加减一跳作为报单的 ask、bid 价进行报价：

$$\text{Ask Price: } P_t + t_s, \quad \text{Bid Price: } P_t - t_s$$

其中， P_t 表示最新成交价格， t_s 表示一个价格最小变动；

4. 以最新成交价格为基准价，基准价加减两跳作为报单的 ask、bid 价进行报价。

策略二：考虑市场波动和报价不平衡性的策略

只利用订单数据和市场价格进行报价的报价策略在诸如行情持续趋势等情况下会有问题，比如，价格持续上涨时，做市商的买单很难被成交，因此容易积累大量的空头头寸导致亏损。所以考虑市场波动性和订单簿的不平衡性是必要的。基于这样的考虑，我们把以最新成交价格为基准价，基准价加减一跳作为报单的 ask、bid 价进行报价的策略作为基础报价策略，提出三个具体的策略。

1. 考虑市场波动性的报价策略：

当市场波动性高时，做市商会在基础报价的基础上加或减一部分进行报价：

$$\text{Ask Price: } P_t + (|P_t - P_{t-1}| + 1)t_s, \quad \text{Bid Price: } P_t - (|P_t - P_{t-1}| + 1)t_s$$

P_t 和 t_s 的含义同上，用 $|P_t - P_{t-1}|$ 刻画市场价格的波动；

2. 考虑订单簿报单数量不平衡性的报价策略：

当做市商发现订单簿的不平衡报价时，做市商会调整报价，当买单数量与卖单数量之差明显偏大时，做市商认为价格有上涨趋势，因此会提高卖单报价；同样的，当卖单数量明显多于买单数量时，做市商会降低买单报价，具体报价方法为：

Ask Price	Bid Price
$P_t + t_s * 2, \text{ if } (q_b - q_s) / (q_b + q_s) > 0.5$	$P_t, \text{ if } (q_b - q_s) / (q_b + q_s) > 0.5$
$P_t, \text{ if } (q_s - q_b) / (q_b + q_s) > 0.5$	$P_t - t_s * 2, \text{ if } (q_s - q_b) / (q_b + q_s) > 0.5$
$P_t + t_s, \text{ others}$	$P_t - t_s, \text{ others}$

其中， q_b, q_s 分别表示订单簿中买单和卖单的数量， $(q_b - q_s) / (q_b + q_s) > 0.5$ 表示买单数量明显多于卖单数量，反之， $(q_s - q_b) / (q_b + q_s) > 0.5$ 表示卖单数量明显多于买单数量；

3. 综合考虑市场波动性和订单簿报单数量不平衡性的报价策略：

Ask Price	Bid Price	Condition
$P_t + (P_t - P_{t-1} + 2) * t_s$	$P_t - P_t - P_{t-1} * t_s$	$\text{if } (q_b - q_s) / (q_b + q_s) > 0.5$
$P_t + P_t - P_{t-1} * t_s$	$P_t - (P_t - P_{t-1} + 2) * t_s$	$\text{if } (q_s - q_b) / (q_b + q_s) > 0.5$
$P_t + (P_t - P_{t-1} + 1) * t_s$	$P_t - (P_t - P_{t-1} + 1) * t_s$	<i>others</i>

第14章 总结

本课题的研究，我们围绕做市策略的报价逻辑这一核心问题，从做市商做市的基本义务和基本需求出发，参考经典理论和经典方法，分析市场行情数据，并尝试构建了做市商做市的报价逻辑和报价策略。由于做市商在市场中的地位和交易规模，做市商的报价和交易会对市场价格产生显著的影响，用历史数据进行回测有较大难度，因此我们并未实现报价逻辑的精确回测，策略的实际有效性需进一步验证。出于策略构建的完整性和有效性的考虑，我们计划对本课题进行持续的后续研究，完善报价逻辑与报价策略，同时，借助我们团队在上一期研究项目中研发的期权工具箱，尽可能实现报价逻辑的准确回测，通过回测结果验证报价策略的有效性，并对策略进行修正与完善。

第三部分 项目总结与展望

围绕相关性分析和配对交易策略，我们进行了理论学习、数据统计分析、模型构建、实证分析等一系列研究工作。

通过相关性分析，我们对 A 股市场的数据特点有了更深刻的了解。我们发现相关系数和标准欧式距离在刻画相关性时有高度一致性。即相关系数大的两只股票之间的标准距离较小。同时距离的波动性明显小于相关系数的波动性，我们认为距离在一定程度上更能刻画相关性。

在行业相关性分析中，我们利用标准欧氏距离构建复杂网络，可以发现专用设备、化学制品、通用机械是行业网络中的关键节点，即上述三个行业在整个市场中有重要的地位。

在个股与大盘的相关性分析中，数据结果显示个股与大盘的相关系数分布接近正态分布，但也存在明显的尖峰厚尾现象。与大盘相关性最强的个股多来自证券或电力行业。与大盘相关性最弱的个股，部分来自黄金和有色行业。其他个股分布规律不明显。这些相关性的结果大都符合我们的认知，也能为策略层面的研究提供数据支持。

基于相关性研究的诸多结果，参考经典的配对交易的做法，我们用距离的方法构建了配对交易策略，并在我国 A 股市场进行有效性检验，得出了非常有价值的结论。

首先，用距离的方法构建配对交易策略在我国股票市场是有效的，能够获得稳定收益。距离小的股票对上进行配对交易，能获得较好的收益，整体来看，距离小的股票对上配对交易策略的收益要明显好于距离大的股票对，但距离最小的股票对上收益不见得最好，这是因为配对交易要求股票的价差有一定的波动性，而距离很小的股票之间价差的波动相应的也很小。

其次，配对交易收益较好的股票对有比较明显的聚集效应，在金融、工业、可选消费等行业中选取股票对有更大可能取得好的收益。配对交易的收益取决于股票对价差的变化情况，与大盘走势关系不大。

在本项目的研究中，我们对我国股票市场进行了大量的相关性研究，给出了

一个逻辑相对简单的配对交易策略，策略的有效性也得到了验证。这些结果可以给实际的交易提供很多有价值的参考和支撑。基于已有的结果，我们计划在接下来的研究中充分利用相关性分析的结果，进一步完善策略的结构和逻辑，完善策略回测机制，对策略进行更加准确的回测，根据回测结果对策略进行修正和完善。

围绕期权做市商策略这一主题，我们详细解读了我国期权市场现行的做市制度，分析了做市商做市的基本需求，这一需求包括维持做市义务，控制单边累积头寸数量，合理管控风险、尽可能减小被动交易的亏损等。作为项目研究的基础，我们详细分析了我国期权市场的行情数据，包括期权合约日间行情的宏观特点和日内变化的微观结构，每一类期权合约在宏观特点和微观结构上变化的一致性、各类期权之间的差异性，期权价格走势与标的价格走势的相关性，不同类别期权价格走势与标的行情相关性的强弱差异以及期权合约的希腊值数据。制度需求以及市场数据的分析为报价逻辑与策略的构建提供了有效的理论与数据支撑。

基于制度、需求和市场数据调研分析的结果，我们构建了理论上合理的报价逻辑，我们认为合理的报价逻辑和策略需要对市场行情的变化做出应对和调整，在同一类合约（如所有的接近平值的期权、所有的深度实值或虚值的期权等）上的报价应具有 consistency，而在不同类期权上的报价应该具有差异性，同时，做市商的报价应该根据希腊值的变化情况做出相应的调整，以实现合理的风险管控。

由于做市商在市场中的地位和交易规模，做市商的报价和交易会对市场价格产生显著的影响，用历史数据进行回测有较大难度，因此我们并未实现对报价逻辑的精确回测，策略的实际有效性需进一步验证。出于策略构建的完整性和有效性的考虑，我们计划对本课题进行持续的后续研究，完善报价逻辑与报价策略，同时，借助我们团队在上一期研究项目中研发的期权工具箱，尽可能实现报价逻辑的准确回测，通过回测结果验证报价策略的有效性，并对策略进行修正与完善。