

# 2023 年第四届“大湾区杯”粤港澳 金融数学建模竞赛

## 基于中国特色估值体系的股票模型分析和投资策略

### 摘要

本文针对中国特色估值体系下的现有估值模型，在政策背景和资本市场定位下，构建中国特色估值体系下股票的特征指标，并系统化描述其画像。从对其中股票分类入手，分析不同分类下的不同类型股票的投资特点。以经济环境的热点作为切入点，设计出一个基于热点事件改变特征指标权重的短期投资组合。然后在洞悉股票市场长期规律后通过改善后的中特估估值模型做长期投资组合。而在做长期投资组合的时候需要根据券商类型突出不同的特征指标然后完成投资组合。这样可以最大化地利用手中有限的投资股金。

针对**任务一（第五部分）**，我们先分析了现有的三个估值模型：市盈率（PE），市净率（PB）和现金流贴现（DCF）的优缺点。然后提出**中特估概念**，给出6个**中特估特征指标**：两个**市场性指标**（市净率（P/B）和**净资产收益率**（ROE））；两个**股票所有权指标**（**央企份额**和**企业性质**）；两个**流动性风险因素指标**（**流动性覆盖率**（LCR）和**资产负债率**（DB））。然后根据这些特征指标构建**中特估画像**。

针对**任务二（第六部分）**，我们随机挑选三个行业（科技，金融，制造业）里每个行业3个股票（大、中、小型），共9只。科技：小型-深科技（000021），中型-京东方A（000725），大型-海康威视（002415）。金融：小型-红塔证券（601236），中型-中国人保（601319），大型-工商银行（601398）。制造业：小型-弘元绿能（603185），中型-长城汽车（601633），大型-宁德时代（300750）。然后根据**特征指标数据**对这九个券商股票进行**聚类分析**。然后得出**蓝色的簇集**共有四个，为深科技（000021），红塔证券（601236），中国人保（601319）和工商银行（601398）。**绿色的簇集**共有两个，为京东方A（000725）和弘元绿能（603185）。**红色的簇集**共有三个，海康威视（002415），长城汽车（601633）和宁德时代（300750）。**蓝色簇集**：**金融类**的券商股票和**国有券商**的特征类型高度相似，可能被**中特估高估**。**红色簇集**：**制造业类**的券商股票拥有**高收益**的特征类型，可能被**中特估低估**。**绿色簇集**：**科技类**的券商股票拥有**高流动性和低风险**的特征类型，可能被**中特估高估**。然后是**投资特点**。**金融类**：选择**流动性覆盖率**（LCR）更高的以及**资产负债率**（DB）更低的券商股票。**制造业类**：选择**市净率**（P/B）和**净资产收益率**（ROE）更低的券商股票。**科技类**：选择**央企份额**更高的和**企业类型层级数**越低的。

针对**任务三（第七部分）**，我们先构建6个特征指标的市场性因素关于Y的多元线性回归模型，然后根据三种热点事件（战争，金融改革，基建）提出修正后的回归模型。

**战争类**：结果是海康威视（002415）>红塔证券（601236）>弘元绿能（603185）>深科技（000021）>京东方A（000725）>宁德时代（300750）>长城汽车（601633）>中国人保（601319）>工商银行（601398）。这个结果显然符合巴以冲突战争类事件（石油价

格飙升)使得科技类券商增值的情况。金融改革类: 结果是中国人保(601319) > 工商银行(601398) > 红塔证券(601236) > 海康威视(002415) > 深科技(000021) > 长城汽车(601633) > 京东方A(000725) > 宁德时代(300750) > 弘元绿能(603185) 这个结果显然符合金融改革类事件使得证券银行类股票券商增值的情况。基建类: 宁德时代(300750) > 海康威视(002415) > 长城汽车(601633) > 弘元绿能(603185) > 中国人保(601319) > 工商银行(601398) > 深科技(000021) > 红塔证券(601236) > 京东方A(000725) 这个结果显然符合基建类事件使得制造业类股票券商增值的情况。

针对任务四(第八部分), 我们根据 $Z$ 与特征指标的关系构建中特估体系参数函数

( $Z$ )。并根据 $Z$ 值的大小对每只股票的收益进行修正, 最后得出预期收益。结果如下: 提升估值的股票排序: 红塔证券(601236) > 工商银行(601398) > 京东方A(000725) > 中国人保(601319) > 深科技(000021) > 海康威视(002415)。需要降低估值的股票排序: 长城汽车(601633) > 弘元绿能(603185) > 宁德时代(300750)。最后是计算出的预计收益结果: 工商银行(601398): 0.431 > 海康威视(002415): 0.233 > 宁德时代(300750): 0.231 > 中国人保(601319): 0.205 > 弘元绿能(603185): 0.177 > 深科技(000021): 0.084 > 长城汽车(601633): 0.053 > 红塔证券(601236): 0.048 > 京东方A(000725): 0.014。

关键词: 特征指标与股票画像 kmeans聚类分析 多元线性回归模型 中特估体系参数



# 一、引言

## 1.1 背景介绍

证券投资的主要考虑是如何获得回报并降低风险，有效评估证券在市场交易中的价值是这一过程的核心。在股市中，为了评估公司的价值，人们通常使用一些常见的估值模型，例如市盈率估值模型、市净率估值模型以及现金流折现模型等。

然而这些模型的估值体系往往重点依赖于自由市场的资本环境，无法将一些政策导向或者国家控股的国有股份有限制公司的价值凸显出来。而且它们的评价体系不适用于我国很多重点关注的企业，忽略了诸如低估值股票，优质成长股，重点行业与领域以及国企改革所具有的超越原有估值的价值。所以根据2022年下半年证监会主席易会满的建议，我们需要着重思考不同类型上市公司的估值逻辑以及合理推测继而探索建立符合中国特色的估值体系。并且阐释中国特色估值体系的核心思想以及应用手段。

## 1.2 中特估概念及其在中国资本市场的重要性

主流财经媒体认为中国特色估值体系的核心要素是：1.政策导向，2.价值投资，3.资本配置，4.风险管理。政策导向是中特估的核心因素，往往获得政策倾斜的企业能获得更高的估值，因为它们符合战略方向，具有发展前景。国企改革就是政策导向的一个经典例子。价值投资强调成长性与可持续性，形成稳健长期的投资习惯。资本配置是指对资本市场的再分配，收到国家支持的企业往往具有更高的价值体现。风险管理则指国家会对市场进行监管，防止出现过多的金融泡沫，具有高风险的企业在中国特色估值体系中往往会有更低的估值，尽管它们拥有不低市盈率和市净率。

## 1.3 研究目的和意义

中国金融市场具有独特的特点和环境，传统的估值方法无法全面的预估所有种类的股票估值，而且根据资本市场的自由分配很多具有政策导向或重点布局的企业都无法拥有真正的估值。中国特色估值体系可以提供更准确更全面的估值体系，帮助投资者进行合适的投资，使市场更具有活力和秩序。

同时政府可以通过该体系更好地识别和评估市场风险和流动性风险，制定有效的政策以保证市场的稳定和健康发展。同时投资者可以根据该体系合理地制定投资决策，使得收益最大化。

# 二、问题分析

## 2.1 中特估股票特征指标与画像

基于中特估概念的政策背景，市场定位以及专家解析来分析什么是中特估，给出全面且准确的中特估的定义。通过题干中提到的中国股票特色估值体系的核心要素，着重

理解并实践政策导向，价值投资，资本配置和风险管理在中国股票市场中的影响和作用，得出符合中国特色的股票市场规律。根据这个找到的市场规律再分析各个类别的股票，计算出一个中特估的估值体系模型，构建中特估股票的特征指标，给出中特估股票的画像。

中国特色的股票估值体系在特定方面有一些不同之处，这些差异可能受到中国股市的独特性和监管环境的影响。政策上，中国股市受到政府政策的较大影响。政府可能会通过一系列政策来促进或限制某些行业或公司的发展，这可能需要被纳入估值考虑因素中。市场情绪上，股市可能更容易受到市场情绪和投资者情感的影响。因此，情绪指标和市场舆论可能在股票特征建立中占据更重要的位置。公司治理上，公司治理可能是一个重要的考虑因素。投资者可能更关注公司内部结构、董事会决策和公司的透明度等方面的信息。宏观经济因素里中国经济的宏观因素对股市有着直接的影响。因此，在股票特征建立中，考虑国家经济政策、GDP增长率等因素可能更为重要。市场流动性上，中国股市的流动性可能相对较低，这意味着在建立股票特征时需要更加关注交易量和流动性风险。中国特色的股票估值体系需要综合考虑国家政策、市场情绪、公司治理以及宏观经济等多个因素。这些因素可能与其他国家的股市存在一些差异，因此在建立股票特征时需要更具体地了解中国市场的特点。

## 2.2 中特估股票分类

根据建立的特征指标分类中特估中的股票可以方便分析投资组合，是非常重要的前提。

行业分类：将股票按照其所属行业进行分类。在中国，不同行业可能受到政府政策和宏观经济因素的不同影响。例如，科技、金融、制造业等行业可能有不同的特征。

按市值来看，根据公司的市值将股票分类为大型、中型和小型股。大型公司通常更稳定，而小型公司可能有更高的增长潜力和波动性。估值水平分类是根据估值指标（如市盈率、市净率）将股票分类为低估值、合理估值和高估值。这有助于识别潜在的投资机会或高风险股票。通过成长性分类将股票分类为价值股、成长股或收入股。这有助于了解公司的盈利模式和未来增长潜力。风险水平上根据股票的波动性和风险水平将其分类为低风险、中风险和高风险。这有助于投资者根据其风险偏好进行选择。财务健康上根据公司的财务状况将股票分类为强劲、稳健或弱势。这包括利润状况、负债水平等财务指标。

建立一个综合的分类系统可能需要考虑多个指标和因素，以更全面地了解股票的特征。后续通过统计工具或机器学习方法来自动化这一过程，以更有效地对大量股票进行分类。

## 2.3 短期股票投资组合

证券市场的行为非常依赖与市场环境，故近期市场热点（如巴以战争）会很大程度影响短期股票走势。我们将考虑近期三类热点事件：1. 战争类如巴以爆发新一轮大规模冲突；2. 金融改革类（证券类和银行类）如减半征收证券交易印花税，中央汇金公司增持四大行；3. 基建类如六部门近日联合印发的《算力基础设施高质量发展行动计划》。结合中特估股票行业、市值类别和自身特征构建短期组合投资模型并进行预测。



1. 首先我们考虑战争类事件巴以冲突。中东地区是全世界最重要的石油供给地区之一，巴以冲突将导致原油的供给的减少甚至中断，进而提高石油市场价格。所以在考虑这类短期事件发生的时候我们要着重考虑流动性覆盖率（LCR）的影响，在这些事件中石化能源类企业的优质流动资产会显著提高，故 $X_5$ 的值会显著提高，且需要更加重视 $X_5$ 的权重。
2. 其次我们考虑金融改革类中的银行类事件比如中央汇金公司增持四大行。此次汇金公司的增持传递出稳定资本市场的积极信号。四大行盈利稳定而且分红率高，但处于很低的估值水平。汇金公司增持彰显了控股股东对于低估值、高分红国有金融机构的信心。中央汇金公司增持四大行会增加四大行的分红率和央企份额，这里 $X_3$ 的值会显著提高，且需要更加重视 $X_3$ 的权重。
3. 《行动计划》的发布将吸引更多对刺激与基础算力相关的行业的投资，提高市场信心并刺激股票价值上涨。当政府政策投资过程中被投资的央企或国有企业的市净率（R/B）和净资产收益率（ROE）会提高，那么我们的中特估体系需要更加强调 $X_1$ 和 $X_2$ 的权重影响力。

## 2.4 长期股票投资组合

长期股票投资中，投资收益动态地受制于影响证券价格波动的各种因素。为了获得令人满意的投资收益，通过分析不同证券对影响因素的不同敏感度，相应地预测不同证券动态的收益率，并以此设立不同证券在投资组合中的动态投资比例，从而建立了基于影响因素分析的动态证券投资组合模型。利用这个模型，我们具体地讨论了在不同的投资环境中，不同的投资时期，如何根据具体的影响因素，动态地调节投资比例，使投资组合更好地适应市场变化，从而达到降低投资风险，提高收益的目的。

通过构建每个特征指标和中特估系数的长期关系，我们计算每个股票的 $Z$ 值， $Z$ 值越小的在中特估体系下越应该提升估值，然后根据原有估值加上 $Z$ 值的影响我们得出每个股票的长期收益。

## 三、符号说明

中特估体系数	$Z$
第三季度预期收益指数	$Y$
市净率（P/B）	$X_1$
净资产收益率（ROE）	$X_2$
央企份额	$X_3$

企业性质（层级数）	$X_4$
流动性覆盖率（LCR）	$X_5$
资产负债率（DB）	$X_6$

## 四、模型假设

1. 在考虑股票市场性指标时，现金流贴现模型（DCF）本身含有许多变量且计算十分复杂，同时权重的选择具有极高主观性。故我们在构建中特估特征指标时假设现金流贴现在一段时间内对考虑的股票都保持一致，故可以不在考虑。
2. 由于P/E模型同时受到盈利水平和风险因素的影响，其过分强调盈利水平而不考虑风险因素，故不适合作为中特估体系指标。其中我们假设其强调的盈利水平大致可以由ROE（净资产收益率）替代，而缺少的风险因素可以由资产负债率（DB）替代。
3. 流动性覆盖率（LCR）往往指代未来30天内优质流动性资产储备与资金净流出量之比，这里我们假设可以用30天内的LCR估计一年的流动性覆盖率，不考虑季节性差异。
4. 国企尤其是央企的分红率显著高于其他企业，故分红率与央企份额呈正相关的关系。我们假设央企份额高的企业分红率一定高，用更直观的央企份额作为特征指标。
5. 由于流动性覆盖率（LCR）只适用于个体银行，所以我们近似用公式：流值/（总值\*负债率）来近似假设每个券商股票的流动性覆盖率（LCR）。原理是LCR=优质流动性资产储备/资金净流出量。当公式求出来小于100%则用100%代替。

## 五、模型特征指标建立与初步应用

### 5.1 现有股票估值模型优缺点

在股市中，有基于公司状况和经济指标的三种常用股票估值模型，市盈率估值模型（P/E），市净率估值模型（PB）和现金流贴现模型（DCF）。

#### 1. P/E

P/E = 每股股价/每股盈利，以每股收益衡量盈利能力。此模型简单易行，运用了近期的盈利估计，而近期的盈利估计一般比较准，可以有广泛的参照比较。P/E适用于盈利稳定、成熟的公司。



同时 P/E 包含了几个明显的缺点：1) 市盈率估值模型主要关注当前的盈利水平，而忽略了公司未来的增长潜力。这可能导致低市盈率的公司在未来有高增长，但仍然被低估，缺少中特估强调的价值投资。2) 市盈率估值模型通常没有考虑到公司的风险因素，如市场风险、行业风险或公司特定风险，缺少中特估强调的风险管理。

## 2. P/B

P/B = 市场价值/账面价值，用于评估一个公司价值是否被低估或高估。P/B 易于计算，对于有大量有形资产的公司尤为重要。P/B 对于比较同一产业的公司十分有用。

P/B 的缺点：1) 对于拥有较多无形资产的公司估值会变得十分不准确。如今国企往往拥有很多的无形资产，比如政府信誉。2) P/B 同样的未考虑公司未来盈利和价值。很多刚起步的重点照顾企业往往拥有强劲的未来发展势头。3) 不适用于负债较多的情况，而往往很多国企持有大量国债。

## 3. DCF

现金流贴现模型 (DCF) 将公司未来特定时期内的现金流还原为现值，反映公司内在价值的本质。此模型有较多变量，更全面，考虑了公司时间价值和风险因素。

由于 DCF 模型含有较多变量，使得计算十分复杂。而对于不同变量的权重依赖于主观性。

# 5.2 中特估的概念

中国特色估值体系是一个由中国证监会主席易会满提出的概念，意思是透过发展符合中国国有企业特点的估值方法，以提高它们在金融市场的估值[3]。中国特色估值体系是结合中国经济的特征和发展阶段以及资本市场的体质来进行估值，通过深入研究成熟市场估值理论的适用场景，把握好不同类型上市公司的估值逻辑，探索建立具有中国特色的估值体系，促进市场资源配置功能更好发挥，这是契合中国式现代化特征的估值体系[2]。中国特色估值体系既要考虑企业的盈利能力、成长性、资产收益率、分红水平等传统因素，也要将盈利以外的国家战略、所处行业、产业安全、社会责任、ESG 等因素纳入估值体系。中国特色估值体系是中国特色现代资本市场的必然产物，应当是中国式现代化在资本市场的集中体现[1]。

# 5.3 中特估股票特征指标

我们构建一个总的中特估指标，记为  $Z$ ，表示为中特估体系常数。 $Z$  值越低，该股票在中特估体系里估值高的概率越大。因为中特估提出的概念就是在中特估体系下原本低估值的国企央企股票会获得更高的估值。所以设定的常数  $Z$  值与中特估估值体系里的高估值概率成反比。

首先我们考虑一些传统意义上的股票市场性指标，包括市净率估值模型(P/B)和净资产收益率模型(ROE)。根据{附件2}浙商证券发布的证券研究报告，证券板块里国有券商

板块估值极低。

1. 这体现在P/B均值里国有券商的平均估值较低（国有券商的PB平均估值为1.52倍，非国有券商P/B均值为2.57倍）。所以我们以市净率（P/B）指标为一个中特估指标，记为 $X_1$ ， $X_1$ 越低，Z越低。
2. 净资产收益率（ROE）为另一个中特估指标，往往国企央企的ROE较低，如军工企业。我们记ROE为 $X_2$ 。那么同样的 $X_2$ 越低，Z越低。

接下来我们从企业的本体所有权入手，这包括企业央企份额和企业性质。

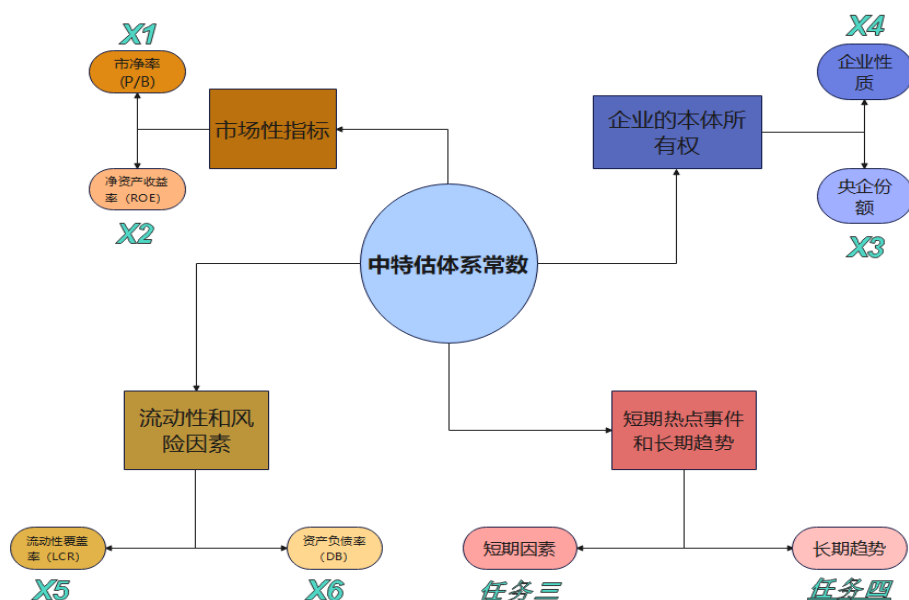
3. 央企份额占比越大的企业受国家政策导向与资源倾斜越多。我们记央企份额为 $X_3$ 。 $X_3$ 的值越高，Z值越低。
4. 企业性质，我们用层级数来代表企业性质，记为 $X_4$ 。其中当企业为中央国有企业时， $X_4 = 1$ ，当企业为地方国有企业， $X_4 = 2$ ，当企业为公众企业， $X_4 = 3$ ，当企业为民营企业， $X_4 = 4$ 。可得 $X_4$ 越低，Z值越低。因为国有企业往往会被估值体系低估。

然后我们考虑流动性和风险因素。

5. 考虑流动性时我们考虑以流动性覆盖率（LCR）为主，记为 $X_5$ 。 $X_5$ 越高，Z值越低。体现为流动性覆盖率高的情况下企业所拥有的优质流动资产储备越高。
6. 考虑风险时我们考虑资产负债率（DB）为主，记为 $X_6$ 。 $X_6$ 越高，Z值越高。体现为资产负债率高的情况下企业往往面对的风险越高，这不符合中特估体系的估值方法。

所以我们预设6个中特估特征指标，包括两个市场性指标（市净率（P/B）和净资产收益率（ROE））；两个股票所有权指标（央企份额和企业性质），其中央企份额是百分比精确指标而企业性质被设定为层级数指标；两个流动性风险因素指标（流动性覆盖率（LCR）和资产负债率（DB））。

## 5.4 中特估股票的画像





## 六、中特估股票的分类及投资特点

### 6.1 沪深A股的中特估股票分类

我们将从企业所属行业和市值两方面进行分类。对于企业行业，主要考虑科技、金融和制造业。对于市值将股票分为大型、中型和小型股。

我们首先考虑企业行业性质，分为科技、金融和制造业。我们将在沪深A股的所有上市券商里挑选9只股票，其中每一类行业的企业股票各找3只。这3只找到的股票将拥有不同大小的市值，即小型，中型，大型企业股票。然后对这9只股票进行聚类分析。其中我们对小型企业股票的定义为市值在200亿到500亿左右，中型企业股票的定义为市值在1500亿左右，大型企业股票的定义为市值在3000亿左右（含以上）。

同时这9只股票按照市值大小也可以分成三组，其中小型，中型，大型企业股票各有3只。这样也满足按市值大小分类的情况。

科技型小型企业我们选择深科技（000021），科技型中型企业我们选择京东方A（000725），科技型大型企业我们选择海康威视（002415）。

金融型小型企业我们选择红塔证券（601236），金融型中型企业我们选择中国人保（601319），金融型大型企业我们选择工商银行（601398）。

制造业小型企业我们选择弘元绿能（603185），制造业中型企业我们选择长城汽车（601633），制造业大型企业我们选择宁德时代（300750）。

我们将每个股票券商的6个中特估特征指标做成一张表格进行汇总。

注：每个股票对应的特征指标数据来源于2023年11月4日沪深A股实际数据

股票代码	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
000021	2.41	4.41%	38.43%	2	188.7%	52.96%
000725	1.14	0.71%	19.52%	4	183.5%	53.23%
002415	4.67	12.59%	42.61%	1	245.6%	39.75%
601236	1.58	0.90%	43.84%	2	235.6%	42.41%
601319	0.98	8.70%	94.27%	1	102.4%	78.38%
601398	0.51	7.92%	90.03%	1	100%	91.74%
603185	1.66	10.13%	7.23%	4	203%	48.71%

601633	3.89	7.54%	31.54%	3	110.6%	65.64%
300750	4.58	18.25%	10.98%	3	126.6%	69.92%

然后对这九个券商股票进行聚类分析。一下先是聚类分析的原理模型。

### 6.1.1 聚类分析的模型

K均值聚类是一种基于距离的聚类算法，其核心思想是将数据点划分为K个簇，使每个簇内的数据点彼此相似，而不同簇之间的数据点差异较大。以下是K均值聚类的数学公式：

#### 1. 初始化质心：

- 首先，选择K个初始质心（centroid）作为簇的中心点，通常是从数据集中随机选择的或者根据某种启发式方法选择的。

#### 2. 分配数据点：

- 对于每个数据点 $X_i$ ，计算它与K个质心之间的距离，通常使用欧氏距离或其他距离度量来衡量。
- 将数据点 $X_i$ 分配到距离最近的质心的簇中，可以用下面的数学公式表示：

$$\operatorname{argmin}_j \|X_i - \mu_j\|^2$$

其中 $\operatorname{argmin}_j$ 表示找到使得 $\|X_i - \mu_j\|^2$ 最小的 $j$ ， $\mu_j$ 表示第 $j$ 个质心。

#### 3. 更新质心：

- 重新计算每个簇的质心，将每个簇中所有数据点的均值作为新的质心，数学公式如下：

$$\mu_j = (1/|C_j|) \cdot \sum X_i$$

for ( $X_i$  in  $C_j$ )

其中 $\mu_j$ 表示第 $j$ 个质心， $C_j$ 表示第 $j$ 个簇， $|C_j|$ 表示簇 $C_j$ 中的数据点数量。

#### 4. 重复步骤2和步骤3，直到满足停止条件，通常是质心不再改变或达到预定的迭代次数。

- K均值聚类的目标是 minimize 每个簇内数据点与其所属质心之间的距离之和，即 minimize 以下目标函数：

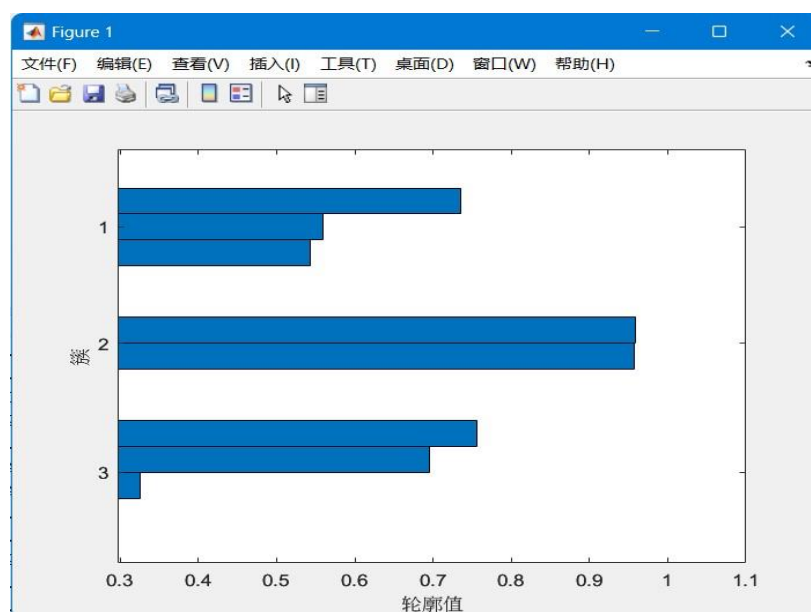
$$J = \sum \sum \|X_i - \mu_j\|^2$$

- 其中，第一个 $\sum$ 是对所有簇 $C_j$ 的求和，第二个 $\sum$ 是对每个簇中的数据点 $X_i$ 的求和。
- K均值聚类是一种迭代算法，通过不断分配数据点并更新质心，逐渐优化目标函数，直到收敛为止。最终，每个簇将包含与之最为相似的数据点，而不同簇之间的数据点将相对较远。

### 6.1.2 聚类分析

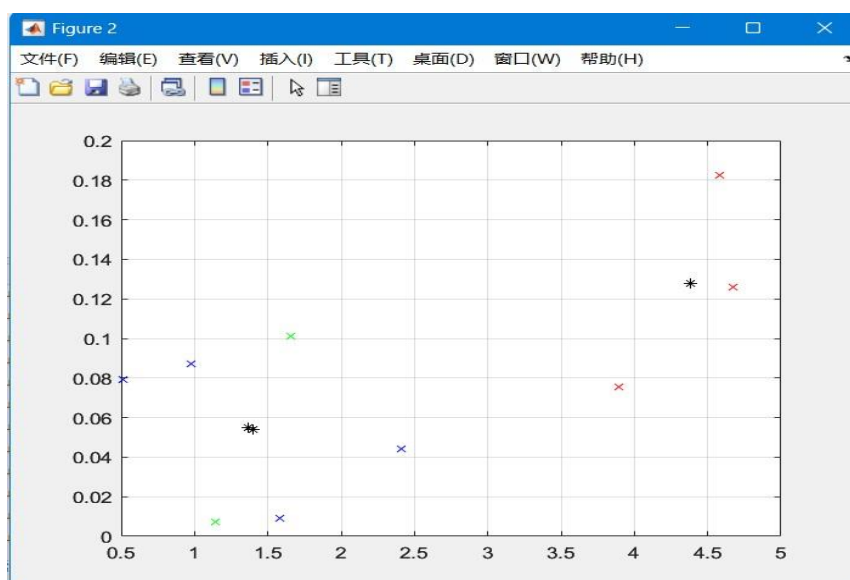
首先在进行聚类分析的时候将表格内的所有百分数都转换为小数，保留四位小数。自然数也保留四位小数。然后用matlab进行聚类分析，代码放在代码附录内（第六部分）。在用matlab进行聚类分析时我们重复100次“1次”或“2次”的迭代，将9个股票根据6个特征指标分为3组，用红绿蓝三种颜色标注且画上它们的质心坐标。结果如下图：

## 1. 轮廓系数



通过轮廓系数的图像可以分析得到第一个簇集的轮廓系数都大于0.55且有一个接近0.75；第二个簇集的轮廓系数都大于0.95且接近1；第三个簇集前两个的轮廓系数都大于0.75，只有一个的轮廓系数不高，大约0.33。在kmeans轮廓系数里系数越接近1说明聚类效果越好。根据这个图像分析我们得到尽管只有9个数据组的情况下聚类分析的效果还是比较不错的。

## 2. kmeans 聚类图



通过这张图我们可以看出蓝色的簇集共有四个，为深科技（000021），红塔证券（601236），中国人保（601319）和工商银行（601398）。绿色的簇集共有两个，为京东方

A（000725）和弘元绿能（603185）。红色的簇集共有三个，海康威视（002415），长城汽车（601633）和宁德时代（300750）。我们放一下6个特征指标的质心数据：

$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
1.370000000 00000	0.0548250000 000000	0.6664250000 00000	1.500000000 00000	1.566750000 00000	0.6637250000 00000
4.380000000 00000	0.1279333333 33333	0.2837666666 66667	2.333333333 33333	1.609333333 33333	0.5843666666 66667
1.400000000 00000	0.0542000000 000000	0.1337500000 00000	4.000000000 00000	1.932500000 00000	0.5097000000 00000

其中第一行蓝色的是蓝色簇集，以此类推。

蓝色簇集里包括了3个金融类券商股票和一个小型科技券商股票，故从行业来看蓝色簇集拥有金融类券商股票的特征指标的特点。通过质心数据我们发现蓝色簇类的市净率（P/B）和净资产收益率（ROE）基本是最低的或非常接近最低，同时企业央企份额最高，达到了66.6%，远超其他两类。且企业类型层级数又是最低（1.5），代表国企央企化最高。同时蓝色簇集拥有最低的流动性覆盖率（LCR，157%）和最高的债务率（DB，66.4%）。所以金融类的券商股票和国有券商的特征类型高度相似，相对来说是最符合中特估估值体系下应该提高估值的券商。

红色簇集包括一个大型科技券商股票和2个中、大型制造类券商股票，故从行业来看红色簇类拥有制造类券商股票的特征指标的特点。通过质心数据我们发现红色簇类的市净率（P/B，4.38）和净资产收益率（ROE，12.8%）远远高出其他两个，且央企份额，企业类型层级数，流动性覆盖率（LCR）和资产负债率（DB）均处于三个簇集的中位数。所以制造业类的券商股票拥有高收益的特征类型，相对来说会在中特估估值体系下因为高收益而被一定低估。

绿色簇集包括两个券商，京东方A（000725）和弘元绿能（603185），又因为弘元绿能本身就属于科技类能源制造商，所以我们近似认为从行业来看绿色簇类拥有科技类券商股票的特征指标的特点。通过质心数据我们发现绿色簇类的市净率（P/B）和净资产收益率（ROE）基本是最低的或非常接近最低，同时拥有最低的央企份额（13.4%）和企业类型层级数（4），代表两个都是民营企业，它们也具有最高的流动性覆盖率（LCR，193%）和最低的负债率（DB，51.0%）。故分析得知科技类的券商股票拥有高流动性和低风险的特征类型，相对来说会在中特估估值体系下因为高流动性低风险而被一定地提高估值。

## 6.2 分类股票的投资特点

根据以上的聚类分析，我们根据不同的行业进行投资特点的归类总结。

在考虑金融类券商股票时，市净率（P/B）和净资产收益率（ROE）已经比较小，符合中特估概念；央企份额最高且企业类型层级数最低，符合央企类型，符合中特估概念。而相对的流动性覆盖率（LCR）偏低以及资产负债率（DB）偏高。所以我们在进

行投资的时候需要重点关注流动性覆盖率（LCR）和资产负债率（DB），选择流动性覆盖率（LCR）更高的以及资产负债率（DB）更低的券商股票。

在考虑制造业类券商股票时，突出的特征指标是比较高的市净率（P/B）以及净资产收益率（ROE）。因为中特估体系往往更加看好低估值的券商股票，故我们在进行投资的时候需要重点关注市净率（P/B）和净资产收益率（ROE），选择市净率（P/B）和净资产收益率（ROE）更低的券商股票。

在考虑科技类券商股票时，市净率（P/B）和净资产收益率（ROE）已经比较小，符合中特估概念；流动性覆盖率（LCR）已经最高且资产负债率（DB）已经最低，符合中特估概念。而相对的央企份额占比偏低以及企业类型层级数偏高。所以我们在进行投资的时候需要重点关注央企份额和企业性质，选择央企份额更高的和企业类型层级数越低的。

## 七、中特估股票的短期投资组合

### 7.1 模型特征与市场热点的结合

根据2.3的问题分析我们已经将市场热点事件分成3类：1.战争类；2.金融改革类；3.基建类。且战争类事件发生的时候  $X_5$ ：流动性覆盖率（LCR）显著提高且权重变高。这是因为巴以冲突爆发的时候中东地区出口的石油量就会变少，相应的我国出口的石油量会大大增加。能源类企业现金流收入大大增加，这样会有更多的资金流进入股票市场，相对的流动性覆盖率（LCR）变得更高且更重要，我们需要根据具体情况调整流动性覆盖率（LCR）的权重。金融改革类事件发生的时候  $X_3$ ：央企份额显著提高且权重变高。这是因为金融改革时国有大型央企会增持四大行或其他央企的股票，这样的政策导向与资源倾斜会让央企份额在中特估体系里变得更加重要且变得更高，我们需要根据情况调整央企份额的权重。基建类事件发生的时候  $X_1$ ：市净率（P/B）和  $X_2$ ：净资产收益率（ROE）显著提高且权重变高。这是因为政府加大基建投资后可以拉动整个经济变得更好，创造更多就业机会和企业发展，这样的政府行为会让市净率（P/B）和净资产收益率（ROE）在中特估体系里变得更加重要且变得更高，我们需要根据情况调整P/B和ROE的权重。

### 7.2 短期投资组合的构建

#### 7.2.1 数据预处理

我们先探究  $X_1$  到  $X_6$  与  $Y$  的关系，其中  $Y$  指的是每个股票的第三季度预期收益指数，因为现在处在一年内的第三季度。 $Y$  与  $X_1$  到  $X_6$  的关系与  $Z$  和  $X_1$  到  $X_6$  的关系总体并不一致，这是因为  $Y$  与  $X_1$  到  $X_6$  的关系重点上是考虑股票的市场因素，而  $Z$  和  $X_1$  到  $X_6$  的关系是在中国股票特色估值体系下构建出来的，具有中特估特色与典型特征指标导向。且中特估的核心就是在政策导向下重点关注低估值的国企券商股票，提高它们的对应估值。

$X_1$  指的是市净率 (P/B) ,  $X_1$  越高,  $Y$  越高。所以  $Y$  与  $X_1$  正相关。

$X_2$  指的是净资产收益率 (ROE) ,  $X_2$  越高,  $Y$  越高。所以  $Y$  与  $X_2$  正相关。

$X_3$  指的是央企份额,  $X_3$  越高,  $Y$  越低。所以  $Y$  与  $X_3$  负相关。

$X_4$  指的是企业性质层级数,  $X_4$  越高,  $Y$  越高。所以  $Y$  与  $X_4$  正相关。

$X_5$  指的是流动性覆盖率 (LCR) ,  $X_5$  越高,  $Y$  越高。所以  $Y$  与  $X_5$  正相关。

$X_6$  指的是资产负债率 (DB) ,  $X_6$  越高,  $Y$  越低。所以  $Y$  与  $X_6$  负相关。

我们找到以上9只股票在2023年11月4号的第三季度收益指数作为  $Y$  值的数值。

代码	000021	000725	002415	601236	601319	601398	603185	601633	300750
$Y$	0.286	0.027	0.949	0.045	0.464	0.754	2.281	0.588	7.083

然后找到9只股票的特征指标矩阵。所有的百分数化简成小数, 加上  $Y$  值, 做成以下矩阵形式, 并将该矩阵标准化。标准化矩阵的matlab代码放在附录代码11.2部分。

股票代码	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$Y$
000021	0.285	0.155	0.249	0.256	0.360	0.282	0.038
000725	0.135	0.025	0.127	0.512	0.350	0.284	0.004
002415	0.552	0.442	0.276	0.128	0.468	0.212	0.125
601236	0.187	0.032	0.284	0.256	0.449	0.226	0.006
601319	0.116	0.306	0.612	0.128	0.195	0.418	0.061
601398	0.060	0.278	0.584	0.128	0.191	0.489	0.099
603185	0.196	0.356	0.047	0.512	0.387	0.260	0.301
601633	0.460	0.265	0.205	0.384	0.211	0.350	0.078
300750	0.542	0.641	0.071	0.384	0.241	0.373	0.934

### 7.2.2 多元线性回归的模型建立

我们根据不同的  $Y$  值变化来构建它与6个特征指标的多元线性回归模型, 并用R语言中的lm()的最佳fit方法来拟合该线性回归模型。

我们构建的线性回归模型如下:

$$Y = \theta_1 X_1 + \theta_2 X_2 + \theta_3 X_3 + \theta_4 X_4 + \theta_5 X_5 + \theta_6 X_6 + \varepsilon$$



其中  $\varepsilon$  指的是误差项,  $\theta_1$  到  $\theta_6$  是每个特征指标的权重系数,  $Y$  是我们要确定函数关系的第三季度预期收益指数。改模型没有设定常数项是因为在实际应用中任意股票的  $X_1$  到  $X_6$  值不会等于0, 否则将毫无意义。若存在常数项, 那么当  $X_1$  到  $X_6$  都等于0整个回归模型中的  $Y$  值还存在且不等于0, 这是显然不可能的。

以下是线性回归的模型解释:

拟合线性模型是一种用于建立线性关系的统计方法, 通常用于回归分析。线性模型的一般数学模型可以表示为:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon$$

其中:

- $y$  表示因变量 (要预测或解释的变量)。
- $\beta_0$  是截距 (常数项), 表示当所有自变量 ( $x_1, x_2, \dots, x_p$ ) 为零时的因变量的期望值。显然在该题分析中  $\beta_0 = 0$ 。
- $\beta_1, \beta_2, \dots, \beta_p$  是自变量的回归系数, 表示自变量与因变量之间的线性关系。这些系数表示自变量的影响力。
- $x_1, x_2, \dots, x_p$  是自变量 (特征) 的值。
- $\varepsilon$  是误差项, 表示模型无法解释的随机误差。

线性模型的目标是估计回归系数 ( $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ ), 以最小化因变量的预测值与观测值之间的残差平方和。最常用的方法是最小二乘法, 它通过最小化残差平方和来找到最优的系数估计值。

### 7.2.3 多元线性回归拟合的结果

根据R语言的结果 (代码放在附录代码的11.3部分), 我们有以下  $X_1$  到  $X_6$  系数值, 在结果图中表示为coefficients\_x1到coefficients\_x6。

```
call:
lm(formula = y ~ 0 + x1 + x2 + x3 + x4 + x5 + x6)

Residuals:
    1      2      3      4      5      6      7      8      9
-0.04781  0.01107 -0.07927  0.15240  0.08791 -0.09428 -0.08941 -0.07414  0.13741

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
x1    -0.4377      0.4954   -0.884   0.442
x2     0.9278      0.5254    1.766   0.176
x3    -1.9643      1.2016   -1.635   0.201
x4    -1.0850      1.0962   -0.990   0.395
x5     0.4567      0.7216    0.633   0.572
x6     2.3731      1.8341    1.294   0.286

Residual standard error: 0.1642 on 3 degrees of freedom
Multiple R-squared:  0.9191,    Adjusted R-squared:  0.7572
F-statistic: 5.678 on 6 and 3 DF,  p-value: 0.09123

> |
```

可以看到  $R^2$  数值达到了0.9191, 拟合的效果不错。

我们将多元线性回归的模型表示出来:

$$Y = -0.4377X_1 + 0.9278X_2 - 1.9643X_3 - 1.0850X_4 + 0.4567X_5 + 2.3731X_6 + \varepsilon$$

这仅仅是根据股票资本市场特征指标建立的模型，在短期考虑中特估体系时我们需要考虑短期热点事件，分析其对该回归模型的影响，然后调整模型参数，最后建立短期投资组合模型，然后得出结果。

### 7.3 热点事件的影响和结果

根据我们在2.3部分的分析我们将热点事件大致分成三类：1.战争类如巴以冲突，通过中国石油出口大大提高影响股票市场流动性覆盖率（LCR）。2.金融改革类如中央汇金公司增持四大行，通过增持国企股份使得央企份额大大提高。3.基建类如《行动计划》的发布使得国企央企的市净率（R/B）和净资产收益率（ROE）提高。

1. 当战争类事件发生的时候， $X_5$ 的权重会变得更高且更重要。随着战争风险溢价重返市场，美国WTI原油在亚洲时段交易中升到了每桶85美元上方。澳大利亚伍德赛德能源集团(Woodside Energy Group Ltd.)和澳洲天然气巨头Santos等亚洲生产商的股价均上涨了3%以上。[4]所以我们以该分析数据为准，因为流动性覆盖率（LCR）的原理是LCR=优质流动性资产储备/资金净流出量，在资金净流出量不变的情况下流动资产上涨3%，所以LCR也上涨3%。我们这里提高 $X_5$ 的系数 $\theta_5$ 至 $1.03\theta_5=0.4704$ ，有以下回归方程：

$$Y = -0.4377X_1 + 0.9278X_2 - 1.9643X_3 - 1.0850X_4 + 0.4704X_5 + 2.3731X_6 + \varepsilon$$

然后我们放上计算好的战争类收益变动图表，其中 $Y$ 指每个股票通过原始线性回归模型计算出来的收益指数，adjusted  $Y$ 指通过战争类线性回归模型计算出来的收益指数。它们两的差值就是短期增加的收益。

战争类回归模型数据								
$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$Y$	Adjusted $Y$	Difference
0.285	0.155	0.249	0.256	0.36	0.282	0.08582	0.09075236	0.0049324
0.135	0.025	0.127	0.512	0.35	0.284	-0.0070752	-0.00227985	0.0047953
0.552	0.442	0.276	0.128	0.468	0.212	0.2042832	0.210695268	0.0064121
0.187	0.032	0.284	0.256	0.449	0.226	-0.1464026	-0.140250851	0.0061517
0.116	0.306	0.612	0.128	0.195	0.418	-0.0268857	-0.024214005	0.0026717
0.06	0.278	0.584	0.128	0.191	0.489	0.1933108	0.195927691	0.0026169
0.196	0.356	0.047	0.512	0.387	0.26	0.3904144	0.395716687	0.0053023



0.46	0.265	0.205	0.384	0.211	0.35	0.1521522	0.155043111	0.0028909
0.542	0.641	0.071	0.384	0.241	0.373	0.7966121	0.799914041	0.0033019

根据图中Difference的大小我们得出短期投资组合在战争类事件中的结果：投资先后顺序：

海康威视（002415）>红塔证券（601236）>弘元绿能（603185）>深科技（000021）>京东方A（000725）>宁德时代（300750）>长城汽车（601633）>中国人保（601319）>工商银行（601398）。

这个结果显然符合巴以冲突战争类事件（石油价格飙升）使得科技类券商增值的情况。

2. 当金融改革类事件发生的时候， $X_3$ 的权重会变得更高且更重要。2023年10月12日，保险股开盘上涨，截至午盘，五大A股上市保险公司盘中涨幅均超过2%。[5]所以我们以该分析数据为准。因为 $\theta_3 < 0$ ，所以我们这里提高 $X_3$ 的系数 $\theta_3$ 至 $0.98\theta_3 = -1.925$ ，有以下回归方程：

$$Y = -0.4377X_1 + 0.9278X_2 - 1.9250X_3 - 1.0850X_4 + 0.4567X_5 + 2.3731X_6 + \varepsilon$$

对应的我们放上金融改革类收益图表：

金融改革类回归模型数据								
x1	x2	x3	x4	x5	x6	$\hat{Y}$	adjusted $\hat{Y}$	difference
0.285	0.155	0.249	0.256	0.36	0.282	0.08582	0.0956057	0.0097857
0.135	0.025	0.127	0.512	0.35	0.284	-0.0070752	-0.0020841	0.0049911
0.552	0.442	0.276	0.128	0.468	0.212	0.2042832	0.21513	0.0108468
0.187	0.032	0.284	0.256	0.449	0.226	-0.1464026	-0.1352414	0.0111612
0.116	0.306	0.612	0.128	0.195	0.418	-0.0268857	-0.0028341	0.0240516
0.06	0.278	0.584	0.128	0.191	0.489	0.1933108	0.216262	0.0229512
0.196	0.356	0.047	0.512	0.387	0.26	0.3904144	0.3922615	0.0018471
0.46	0.265	0.205	0.384	0.211	0.35	0.1521522	0.1602087	0.0080565
0.542	0.641	0.071	0.384	0.241	0.373	0.7966121	0.7994024	0.0027903

根据图中Difference的大小我们得出短期投资组合在金融改革类事件中的结果：投资先后顺序：

中国人保（601319）>工商银行（601398）>红塔证券（601236）>海康威视（002415）>深科技（000021）>长城汽车（601633）>京东方A（000725）>宁德时代（300750）>弘元绿能（603185）

这个结果显然符合金融改革类事件使得证券银行类股票券商增值的情况。

3. 当基建类事件发生的时候， $X_1$ 和 $X_2$ 的权重会变得更高且更重要。中信证券更是给出预计基建投资一季度同比增长7%的积极观点。[6]我们近似以该数据作为 $X_1$ 和 $X_2$ 提升的百分比。因为 $\theta_1 < 0$ ，所以我们这里提高 $X_1$ 的系数 $\theta_1$ 至 $0.93\theta_1 = -0.4071$ 。我们这里提高 $X_2$ 的系数 $\theta_2$ 至 $1.07\theta_2 = 0.9927$ ，有以下回归方程：

$$Y = -0.4071X_1 + 0.9927X_2 - 1.9643X_3 - 1.0850X_4 + 0.4567X_5 + 2.3731X_6 + \varepsilon$$

对应的我们放上基建类收益图表：

基建类回归模型数据								
x1	x2	x3	x4	x5	x6	$\hat{Y}$	adjusted $\hat{Y}$	difference
0.285	0.155	0.249	0.256	0.36	0.282	0.08582	0.1046005	0.0187805
0.135	0.025	0.127	0.512	0.35	0.284	-0.0070752	-0.0013217	0.0057535
0.552	0.442	0.276	0.128	0.468	0.212	0.2042832	0.2498602	0.045577
0.187	0.032	0.284	0.256	0.449	0.226	-0.1464026	-0.1386036	0.007799
0.116	0.306	0.612	0.128	0.195	0.418	-0.0268857	-0.0034767	0.023409
0.06	0.278	0.584	0.128	0.191	0.489	0.1933108	0.213189	0.0198782
0.196	0.356	0.047	0.512	0.387	0.26	0.3904144	0.4195164	0.029102
0.46	0.265	0.205	0.384	0.211	0.35	0.1521522	0.1834267	0.0312745
0.542	0.641	0.071	0.384	0.241	0.373	0.7966121	0.8547982	0.0581861

根据图中Difference的大小我们得出短期投资组合在基建类事件中的结果：投资先后顺序：

宁德时代（300750）>海康威视（002415）>长城汽车（601633）>弘元绿能（603185）>中国人保（601319）>工商银行（601398）>深科技（000021）>红塔证券（601236）>京东方A（000725）

这个结果显然符合基建类事件使得制造业类股票券商增值的情况。

## 八、中特估股票的长期投资组合模型

### 8.1 模型特征与长期投资的关联性

模型中的特征应与长期投资的目标相一致，选择具有长期增长潜力的股票。中特估的考虑因素应更注重企业的长期盈利能力和未来增长前景。共同关注公司的基本面。通过深入的财务分析，确保选取的股票具备稳健的财务状况和业务模式。考虑政策和宏观经济因素的模型特征，在长期内预测企业的经营环境和市场动态。长期投资和模型特征中的风险管理需关联，确保投资组合在不同市场环境下都能保持稳健。在设计模型特征时，确保这些特征与长期投资的理念相契合，能够在长期内为投资组合带来稳定的、可持续的增长。

故我们为了找到中特估体系下符合长期投资目标的特征指标关联性，我们开始构建中特估体系参数（Z值）。上文提到Y与 $X_1$ 到 $X_6$ 的关系与Z和 $X_1$ 到 $X_6$ 的关系不一致。Y与 $X_1$ 到 $X_6$ 的关系更多的是根据股票市场的资本市场性因素决定的，而Z和 $X_1$ 到 $X_6$ 的关系就需要构建中特估体系，找到 $X_1$ 到 $X_6$ 不同的侧重点。

根据前面5.3的特征指标分析，我们直接写出下列关系：

- $X_1$ 越低，Z越低。
- $X_2$ 越低，Z越低。
- $X_3$ 的值越高，Z值越低。
- $X_4$ 越小，Z值越低。
- $X_5$ 越高，Z值越低。
- $X_6$ 越高，Z值越高。

显然我们可以根据 $X_1$ 到 $X_6$ 与Z的关联度构建一个Z的函数指标，其中函数指标的参

数设定与前面关于 $Y$ 的回归模型的系数绝对值有关。然后根据 $Z$ 的大小来修改 $Y$ 的收益值，最后比较调整后的收益值。

## 8.2 长期投资组合的构建

根据上述分析我们可以总结： $X_3$ 和 $X_5$ 与 $Z$ 成反比，其他与 $Z$ 成正比。其中 $\theta_1$ 到 $\theta_6$ 是 $X_1$ 到 $X_6$ 的回归系数，我们再一次展现一下多元回归模型：

$$Y = \theta_1 X_1 + \theta_2 X_2 + \theta_3 X_3 + \theta_4 X_4 + \theta_5 X_5 + \theta_6 X_6 + \varepsilon$$

$$Y = -0.4377X_1 + 0.9278X_2 - 1.9643X_3 - 1.0850X_4 + 0.4567X_5 + 2.3731X_6 + \varepsilon$$

所以我们构建：

$$Z = \frac{X_1' X_2' X_4' X_6'}{X_3' X_5'} (1 + \beta_1)(1 + \beta_2)(1 + \beta_3)(1 + \beta_4)(1 + \beta_5)(1 + \beta_6)$$

其中：

- $X_i'$ 指的是标准化后的矩阵中 $X_i$ 的数值， $i=1,2,\dots,6$ 。

- $\beta_i = \frac{|\theta_i|}{\sqrt{\sum_{i=1}^6 (\theta_i^2)}}$ ， $i=1,2,\dots,6$ 。所以我们有：

$$\beta_1 = 0.1268, \beta_2 = 0.2687, \beta_3 = 0.5689, \beta_4 = 0.3142, \beta_5 = 0.1323, \beta_6 = 0.6873$$

所以最后我们有：

$$Z = \frac{5.6314 X_1' X_2' X_4' X_6'}{X_3' X_5'}$$

这就是我们根据上述 $Y$ 的多元回归模型中每个变量前的系数得出的中特估股票估值体

系常数，其中求 $\beta_i = \frac{|\theta_i|}{\sqrt{\sum_{i=1}^6 (\theta_i^2)}}$ 的过程就是在求平方差后的标准权重。

首先这个模型是根据“ $X_3$ 和 $X_5$ 与 $Z$ 成反比，其他与 $Z$ 成正比”这一个关系构建的。

$X_3'$ 和 $X_5'$ 放在分母，其他的放在分子。又因为我们已经将上述矩阵标准化，所以带入的变量是标准化矩阵下 $X_1$ 到 $X_6$ 的量。然后我们需要考虑上述 $Y$ 的回归模型中系数的影响力，作为 $X_1'$ 到 $X_6'$ 在中特估体系里的权重。具体方法是用每个变量的参数绝对值除以6个变量的参数平方和的开方。

下一部分我们对9只股票进行 $Z$ 值分析且根据 $Z$ 值得出长期投资组合的收益结果。

## 8.3 预测长期投资组合的收益

我们将 $Z$ 值计算结果展示如下：

x1	x2	x3	x4	x5	x6	Y	$\bar{Y}$	$\bar{Z}$
0.285	0.155	0.249	0.256	0.360	0.282	0.038	0.086077	0.200516
0.135	0.025	0.127	0.512	0.350	0.284	0.004	-0.00673	0.062147
0.552	0.442	0.276	0.128	0.468	0.212	0.125	0.203698	0.288468
0.187	0.032	0.284	0.256	0.449	0.226	0.006	-0.14702	0.015084
0.116	0.306	0.612	0.128	0.195	0.418	0.061	-0.02618	0.08942
0.060	0.278	0.584	0.128	0.191	0.489	0.099	0.193783	0.053165
0.196	0.356	0.047	0.512	0.387	0.260	0.301	0.389793	2.883497
0.460	0.265	0.205	0.384	0.211	0.350	0.078	0.152877	2.138711
0.542	0.641	0.071	0.384	0.241	0.373	0.934	0.796394	16.29195

然后我们得出在中特估体系下需要提升估值的股票排序（由左至右为提升估值大到小）：

红塔证券（601236）>工商银行（601398）>京东方A（000725）>中国人保（601319）>深科技（000021）>海康威视（002415）

然后我们得出在中特估体系下需要降低估值的股票排序（由左至右为降低估值小到）：

长城汽车（601633）>弘元绿能（603185）>宁德时代（300750）

接下来我们根据 $Z$ 值进行收益指数修正。因为 $Y$ 值已经是标准化后的数值，我们也需要先对 $Z$ 值进行标准化。这里的标准化是开方，因为上述模型有一次平方和求值。然后根据标准化后 $Z$ 值的修正 $Y$ 值，结果如下（收益从大到小排列）：

股票代码	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$Y$
000021	0.285	0.155	0.249	0.256	0.360	0.282	0.038
000725	0.135	0.025	0.127	0.512	0.350	0.284	0.004
002415	0.552	0.442	0.276	0.128	0.468	0.212	0.125
601236	0.187	0.032	0.284	0.256	0.449	0.226	0.006
601319	0.116	0.306	0.612	0.128	0.195	0.418	0.061
601398	0.060	0.278	0.584	0.128	0.191	0.489	0.099
603185	0.196	0.356	0.047	0.512	0.387	0.260	0.301
601633	0.460	0.265	0.205	0.384	0.211	0.350	0.078
300750	0.542	0.641	0.071	0.384	0.241	0.373	0.934

x1	x2	x3	x4	x5	x6	Y	$\bar{Y}$	$\bar{Z}$	stand $\bar{Z}$	adjusted Y
0.060	0.278	0.584	0.128	0.191	0.489	0.099	0.193783	0.053165	0.230575157	0.431296251
0.552	0.442	0.276	0.128	0.468	0.212	0.125	0.203698	0.288468	0.53709241	0.233041919
0.542	0.641	0.071	0.384	0.241	0.373	0.934	0.796394	16.29195	4.036329194	0.231444846
0.116	0.306	0.612	0.128	0.195	0.418	0.061	-0.02618	0.08942	0.299031992	0.204652558
0.196	0.356	0.047	0.512	0.387	0.260	0.301	0.389793	2.883497	1.698086272	0.177166818
0.285	0.155	0.249	0.256	0.360	0.282	0.038	0.086077	0.200516	0.447790372	0.084238017
0.460	0.265	0.205	0.384	0.211	0.350	0.078	0.152877	2.138711	1.46243331	0.05302957
0.187	0.032	0.284	0.256	0.449	0.226	0.006	-0.14702	0.015084	0.122816045	0.048325266
0.135	0.025	0.127	0.512	0.350	0.284	0.004	-0.00673	0.062147	0.249293514	0.014284651

工商银行(601398): 0.431>海康威视(002415): 0.233>宁德时代(300750): 0.231>中国人保(601319): 0.205>弘元绿能(603185): 0.177>深科技(000021): 0.084>长城汽车(601633): 0.053>红塔证券(601236): 0.048>京东方A(000725): 0.014

这显然符合央企提升估值比率多(如工商银行,中国人保), 且原本估值极低的股票在中特估下仍然不会太高。(如红塔证券, 京东方A)

## 九、结论

### 9.1 模型评价

模型优点:

- 1.考虑的中特估特征指标比较全面, 既包括市场性指标, 企业性质类指标同时也包括流动性风险类指标。这基本涵盖了中国股票特色估值体系的主要核心思想。
- 2.九只股票的选择并不是基于6个挑选出的特征指标, 且是随机抽取科技类, 金融类和制造业类各三只, 且符合大中小型规模。既可以规避可能存在的共线性且覆盖面基本全面, 不管是基于行业还是基于体量。
- 3.构建中特估体系的过程分成两个阶段, 先在第七部分做关于Y的多元线性回归, 充分考虑现有估值体系, 再构建中特估参数Z, 根据Z值调整Y值, 这样显得更加严谨。
- 4.在构建关于Z的函数设计的时候, 我们充分考虑了基于Y的多元线性回归的每个特征指标的权重, 而不是仅仅将它们视为同等重要。

模型缺点:

- 1.关于Z的函数设计并没有考虑权重相加的模型可能, 并没有做到互相对比, 找出相对更合适的模型。所以这一构建的函数体系在思维上存在一定主观性。
- 2.模型中挑选的数据点并不是很多, 仍具有一定偶然因素。如果可以找到更多的数据集会对模型的建立起到更好的作用。

### 9.2 模型的推广

该模型充分考虑了中国股票市场的资本性质和政策导向性质，找出基本股票预期收益指数的一般表达式，再根据中特估特征指标体系在一般表达式上进行修正，得出最符合中特估体系的收益指数。该模型对原始数据有充分的分析，首先是分析不同行业的中特估特征指标特点，再根据该特点进行投资特点的分析。

我们先得出一般股票的收益指数多元线性回归模型，再根据三类热点事件分别构建基于热点事件的多元线性回归模型，进行短期投资组合。然后构建中特估参数体系，根据长期特征指标建立函数关系。通过Z值分析各个股票的收益指数提升或下降，最后根据Z值完成收益指数修正，得出最终收益。

总体上来说，该模型可以应用与不同行业不同体量的股票券商，又可以结合短期热点事件，长期特征指标特点，进行收益指数预估。

我们同时发掘了每种热点事件的内在特点，其如何影响市场特征指标，从而影响整个收益指数。故该模型还可以为以后研究热点事件对股票资本市场影响的研究者提供一些总体思路的参考。

## 十、参考文献

1. 探索中国特色估值体系 做中国特色价值投资 - 腾讯网. <https://new.qq.com/rain/a/20230217A0516L00>.
2. 中国特色估值体系的内涵及其投资逻辑 - 腾讯网. <https://new.qq.com/rain/a/20230321A08NIW00>.
3. 重磅新名词！什么是中国特色估值体系？中字头概念拉升，券商：银行、国企估值有提升空间\_腾讯新闻. <https://new.qq.com/rain/a/20221122A055QP00>.
4. 从芯片巨头到造车新贵！一文读懂：巴以冲突会影响全球哪些股票？ - 东方财富网. <https://finance.eastmoney.com/a/202310092863763357.html>
5. 汇金增持四大行后 下一步怎么走？历史规律显示未来有增持券商股、保险股可能. <https://finance.eastmoney.com/a/202310122867987298.html>
6. 大盘强势走高，建筑板块功不可没 <http://www.chinavalue.net/Investment/Blog/2022-1-18/1937642.aspx#:~:text=%E4%B8%A%E5%8D%88%EF%BC%8C%E4%B8%89%E5%A4%A7%E6%8C%87%E6%95%B0%E7%94%B1%E5%BC%B1%E8%BD%AC%E5%BC%BA%EF%BC%8C%E9%80%90%E6%AD%A5%E5%B8%82%E5%9C%BA%E5%92%8C%E5%88%9B%E4%B8%9A%E6%9D%BF%E5%BD%A2%E6%88%90%E6%98%8E%E6%98%BE%E7%9A%84%E8%B7%B7%E8%B7%B7%E6%9D%BF%E6%95%88%E5%BA%94%E3%80%82,%E5%88%86%E6%97%B6%E5%9B%BE%E4%B8%8A%EF%BC%8C%E9%BB%84%E7%99%BD%E7%BA%BF%E5%88%86%E5%8C%96%E6%AF%94%E8%BE%83%E6%98%8E%E6%98%BE%EF%BC%8C%E7%99%BD%E4%B8%8A%E3%80%81%E9%BB%84%E4%B8%8B%E7%9A%84%E6%A0%BC%E5%B1%80%EF%BC%8C%E8%AF%B4%E6%98%8E%E5%B8%82%E5%9C%BA%E4%B8%BB%E8%A6%81%E4%BE%9D%E9%>



9D%A0%E6%9D%83%E9%87%8D%E8%82%A1%E7%9A%84%E5%B8%A6%E  
5%8A%A8%E3%80%82%20%E6%AF%94%E5%A6%82%E6%88%91%E4%BB%A  
C%E7%9C%8B%E5%88%B0%EF%BC%8C%E5%BB%BA%E7%AD%91%E3%8  
0%81%E6%88%BF%E5%9C%B0%E4%BA%A7%E3%80%81%E9%92%A2%E9%9  
3%81%E3%80%81%E5%BB%BA%E6%9D%90%E7%AD%89%E4%BC%A0%E7%  
BB%9F%E5%9F%BA%E5%BB%BA%E8%81%94%E5%8A%A8%E6%8B%89%E  
5%8D%87%EF%BC%8C%E5%85%B6%E4%B8%AD%E5%BB%BA%E7%AD%9  
1%E6%9D%BF%E5%9D%97%E4%BD%8D%E5%B1%85%E6%B6%A8%E5%B9%  
85%E6%A6%9C%E5%89%8D%E5%88%97%E3%80%82

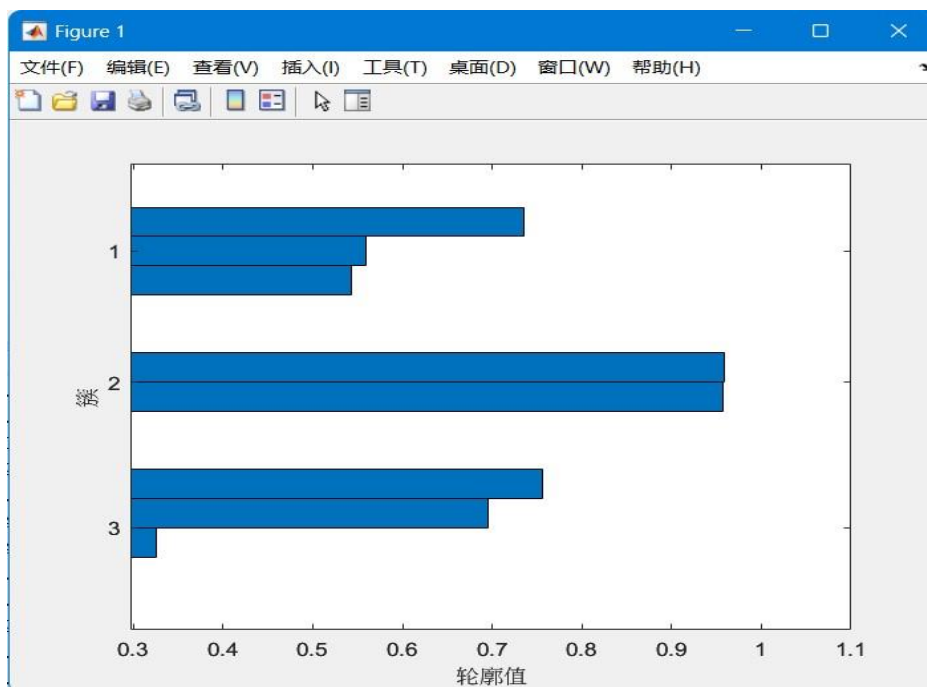
## 十一、附录代码

### 11.1 第六部分6.1.1 kmeans聚类分析的matlab代码

```
clear;clc;
x=[2.4100    0.0441  0.3843  2.0000  1.8870  0.5296
1.1400  0.0071  0.1952  4.0000  1.8350  0.5323
4.6700  0.1259  0.4261  1.0000  2.4560  0.3975
1.5800  0.0090  0.4384  2.0000  2.3560  0.4241
0.9800  0.0870  0.9427  1.0000  1.0240  0.7838
0.5100  0.0792  0.9003  1.0000  1.0000  0.9174
1.6600  0.1013  0.0723  4.0000  2.0300  0.4871
3.8900  0.0754  0.3154  3.0000  1.1060  0.6564
4.5800  0.1825  0.1098  3.0000  1.2660  0.6992
];
[idx,cen]=kmeans(x,3,'replicates',100,'display','final');
%轮廓值、轮廓系数
silhouette(x,idx)
color=['r','g','b'];
figure,
for i=1:3
plot(x(idx==i,1),x(idx==i,2),'color',color(i),'linestyle','none','marker','x')
hold on
end
%质心坐标
a=cen(:,1);
b=cen(:,2);
plot(a,b,'color','k','linestyle','none','marker','*')
hold off
grid on
```

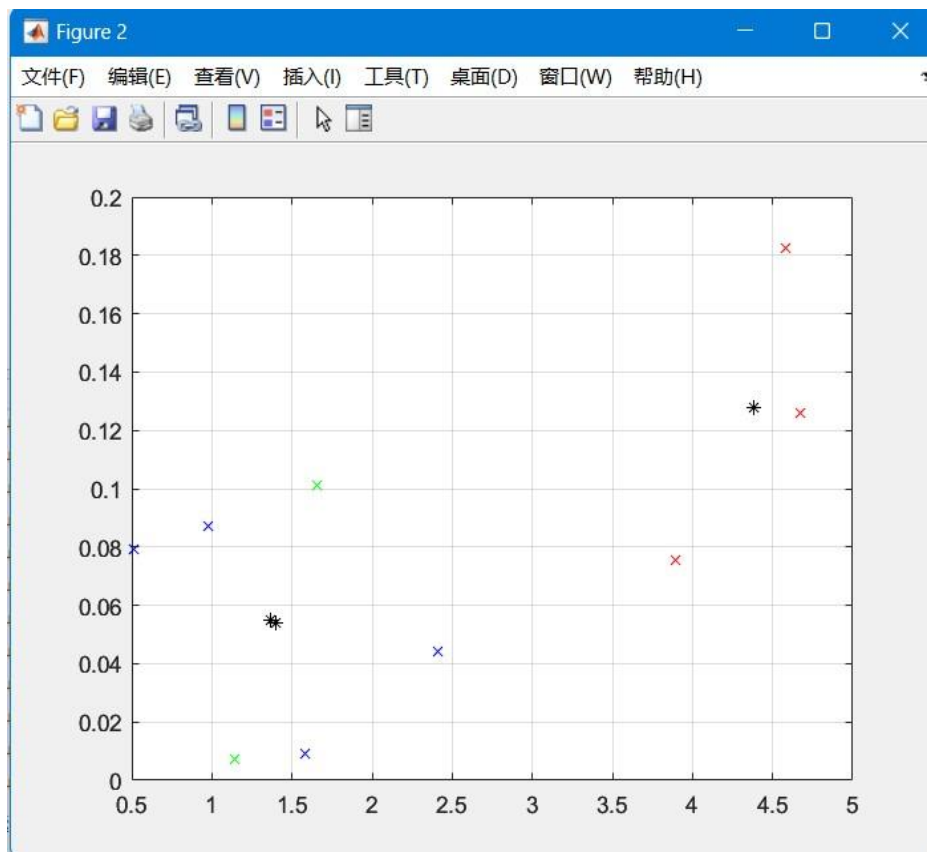
两张结果图：

#### 1. 轮廓系数





## 2. kmeans 聚类图



### 11.2 第七部分7.2.1 矩阵标准化的matlab代码

%%矩阵标准化

```
clear;clc;
X=[2.410.0441 0.3843 2      1.887  0.5296 0.286
1.14  0.0071 0.1952 4      1.835  0.5323 0.027
4.67  0.1259 0.4261 1      2.456  0.3975 0.949
1.58  0.0090 0.4384 2      2.356  0.4241 0.045
0.98  0.0870 0.9427 1      1.024  0.7838 0.464
0.51  0.0792 0.9003 1      1.000  0.9174 0.754
1.66  0.1013 0.0723 4      2.030  0.4871 2.281
3.89  0.0754 0.3154 3      1.106  0.6564 0.588
4.58  0.1825 0.1098 3      1.266  0.6992 7.083
];
[a,b]=size(X);
for i=1:b
    J(1,i)=norm(X(:,i));
end
J= repmat(J,a,1);
Y=X./J;
Y
```

### 11.3 第七部分7.2.2 多元线性回归的R语言代码

```
x1=c(0.285,0.135,0.552,0.187,0.116,0.060,0.196,0.460,0.542)
```

```
x2=c(0.155,0.025,0.442,0.032,0.306,0.278,0.356,0.265,0.641)
```

```
x3=c(0.249,0.127,0.276,0.284,0.612,0.584,0.047,0.205,0.071)
```

```
x4=c(0.256,0.512,0.128,0.256,0.128,0.128,0.512,0.384,0.384)
```

```
x5=c(0.360,0.350,0.468,0.449,0.195,0.191,0.387,0.211,0.241)
```

```
x6=c(0.282,0.284,0.212,0.226,0.418,0.489,0.260,0.350,0.373)
```

```
y=c(0.038,0.004,0.125,0.006,0.061,0.099,0.301,0.078,0.934)
```

```
fit=lm(y~0+x1+x2+x3+x4+x5+x6)
```

```
summary(fit)
```

输出结论：

```
Call:
```

```
lm(formula = y ~ 0 + x1 + x2 + x3 + x4 + x5 + x6)
```

```
Residuals:
```

```
      1      2      3      4      5      6      7      8      9  
-0.04781  0.01107 -0.07927  0.15240  0.08791 -0.09428 -0.08941 -0.07414  0.13741
```

```
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t )
x1	-0.4377	0.4954	-0.884	0.442
x2	0.9278	0.5254	1.766	0.176
x3	-1.9643	1.2016	-1.635	0.201
x4	-1.0850	1.0962	-0.990	0.395
x5	0.4567	0.7216	0.633	0.572
x6	2.3731	1.8341	1.294	0.286

```
Residual standard error: 0.1642 on 3 degrees of freedom
```

```
Multiple R-squared:  0.9191,    Adjusted R-squared:  0.7572
```

```
F-statistic: 5.678 on 6 and 3 DF,  p-value: 0.09123
```

```
> |
```

