

2022 第三届“大湾区杯”粤港澳 金融数学建模竞赛

题 目 基于宏观经济周期的大类资产预测模型

摘 要

对未来宏观经济的判断以及大类资产的预测是大类资产配置动态调整的核心，唯有准确预测出未来经济形势和资产表现的情况，才能有效地构建出较佳的大类资产配置策略，以获取满意的投资回报。本文主要研究以美林投资时钟理论为基础的宏观经济周期，判断了过去中国二十年的宏观经济环境，建立了基于层次分析、综合评价、时间序列分析、灰色预测以及资本资产定价的预测模型，通过 Python 软件对各参数求解，最终分析了大类资产投资组合在国内未来五年的收益情况。

针对问题一，首先，使用 Excel 软件对附件一进行数据预处理，为了提高数据的代表性与有效性，本文只选取 1995 年至 2022 年的数据。通过建立综合评价模型客观地选取了四个高频有效的宏观经济指标：CPI、GDP、国家财政收支、景气指数。接着，通过 Python 编程对这四项经济指标进行可视化分析。结合参考文献，最终将国内过去 20 年划分为七个经济状态：2001-2005 年为复苏期、2006-2008 年为过热期，2009-2013 年为滞涨期，2014-2015 年为衰退期，2016-2019 年为复苏期，2020 年为衰退期，2021 年为复苏期。

针对问题二，本文建立了时间序列分析二次预测模型，使用 Python 编程求解。除了使用 CPI、GDP、国家财政收支与景气指数这四个经济指标外，还增加了人民币存贷款利率这一项经济指标。首先，采用二次指数平滑法依次预测各经济指标下一年的数据，再采用非平稳时间序列分析做第二次预测，并对未来五年经济状态进行 Python 可视化分析。最终，本文的预测是未来五年国内经济状态将处在复苏期。

针对问题三，首先通过 Excel 软件对附件二进行数据预处理，挑选了过去十年大类资产指数的数据，接着进行标准归一化处理，再对这些数据进行方差计算。本文根据方差大小以及参考文献的结论，挑选出了沪深 300、南华商品指数、中债综合财富（7-10 年）指数以及货币基金。然后，计算出各指数的风险收益特征。最后，使用灰色关联度分析法分析了各大类资产指数之间的相关性，得知它们普遍具有较强的相关性。

针对问题四，基于问题二预测未来五年国内处于复苏期的结论，本文根据参考文献得知股票类资产占优，且配置策略的份额应是股票 > 债券 > 现金 > 大宗商品。本文建立灰色预测模型，并通过 Python 编程求解，得到未来五年国内各指数的年收益率和期望收益。然后，建立资本资产定价模型，计算各指数的期望收益率以及夏普比率。

综上所述，本文结合多种模型以及推导演算，全面分析了宏观经济周期的大类资产的收益情况。经分析论证，本文的模型与分析具有合理性。

关键词：综合评价 时间序列分析 灰色关联分析 灰色预测模型 资本资产定价模型

目 录

一、问题重述	4
1.1 问题背景	4
1.2 具体问题	4
二、问题分析	4
2.1 问题一的分析	4
2.2 问题二的分析	5
2.3 问题三的分析	5
2.4 问题四的分析	6
三、基本假设与符号说明	6
3.1 基本假设	6
3.2 符号说明	6
四、数据预处理	7
4.1 对附件一的数据预处理	7
4.2 对附件二的数据预处理	8
五、问题一的模型建立与求解	8
5.1 建模前的准备	8
5.2 模型的建立	9
5.3 选取高频有效的经济指标	11
5.4 高频有效的经济指标的时序图	11
5.5 综合分析与经济状况的划分	12
六、问题二的模型建立与求解	13
6.1 建模前的准备	13
6.2 模型的建立	14
6.3 中国未来五年经济状况分析	16
七、问题三的模型建立与求解	16

7.1	建模前的准备	16
7.2	模型的建立与求解结果的展示	16
7.3	风险收益特征	17
7.4	大类资产指数相关性	18
7.5	大类资产指数相关性的求解与展示	19
八、问题四的模型建立与求解		20
8.1	大类资产投资组合的确立	20
8.2	灰色预测模型	21
8.3	数字特征	23
8.4	资本资产定价模型 (CAPM)	23
8.5	夏普比率	24
8.6	模型的求解与展示	24
8.7	结论	25
九、模型的评估与改进		25
9.1	模型的优点	25
9.2	模型的缺点	25
9.3	模型的改进推广	25
参考文献		26

一、 问题重述

1.1 问题背景

相对于独立投资，大类资产配置通过分散投资为投资者平滑了投资组合风险，具有风险控制优势。通过将投资能力建立在大类资产配置和科技平台上的方式可以突破投资能力天花板，形成规模效应和网络效应。规模效应是指因规模增大带来的经济效益提高；网络效应是指一个产品或服务的用户越多，创造的价值就越大。目前全球资管公司竞争激烈但市场分散，未来的增长主要来自份额提高，领先的一站式大类资产配置平台预计将成为终极赢家。

由于投资组合中各类资产的未来表现是难以判断的，且各类资产的长期表现是否能够持续占优势也是难以明确的，因此，大类资产配置在实践过程中必然面临动态调整的问题。而常规操作是提高表现较佳的资产类别在整体组合的比例。

大类资产配置动态调整的核心在于对未来宏观经济的判断以及大类资产的预测。只有在准确预测出未来经济状况和资产表现的情况之下，我们才能有效辨别出表现较佳的资产类别。因此，针对中国的投资者与投资环境，我们需要有一套更适合的宏观经济与大类资产预测模型。

1.2 具体问题

题目要求我们综合附件所给数据建立合理的数学模型并解决以下问题：

问题一：找出两个或以上高频有效的宏观经济指标，并将 2001-2021 年国内的宏观经济运行状况划分为不同的经济状态。

问题二：通过建立宏观经济模型或其他数学模型模拟中国未来五年的经济增长、通胀、利率等宏观经济环境，并说明未来五年中国将面临的经济状态处于问题一划分出的经济状态中的什么经济状态。

问题三：选出能四类资产的四个指数，预测大类资产指数在问题一划分出的各类经济状态下的风险收益特征，以及大类资产指数之间的相关性。

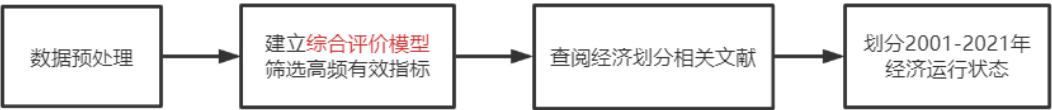
问题四：基于建立的模型预测国内未来五年的经济状态，选择合适的大类资产指数构建投资组合并预测该投资组合的风险收益特征。

二、 问题分析

2.1 问题一的分析

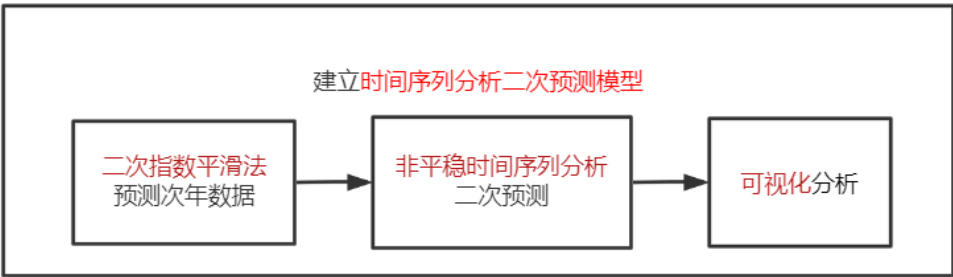
附件一提供了国民经济核算、工业、价格指数、银行与货币、利率与汇率、财政、就业与工资、景气指数以及人口共九大类宏观经济指标的数

据，问题一要求我们在九大类宏观经济指标中找出两个或以上高频有效的宏观经济指标，并将 2001-2021 年国内的宏观经济运行状况划分为不同的经济状态。这里可以用综合评价模型找出高频有效的宏观经济指标，再查阅经济状态分类的相关文献，综合分析高频有效的宏观经济指标的表现，将 2001-2021 年间的经济运行状况划分为美林时钟模型的衰退期、复苏期、过热期和滞涨期。



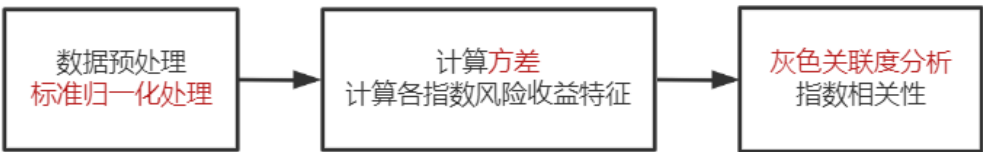
2.2 问题二的分析

问题二要求我们通过建模来模拟中国未来五年的经济增长、通胀、利率等宏观经济环境，并结合模型分析说明未来五年中国的经济状态。这里可以建立时间序列模型模拟过去十年的数据，逐年预测未来五年中国的经济运行状况的数据，再通过数据可视化分析说明未来五年中国的经济状态。



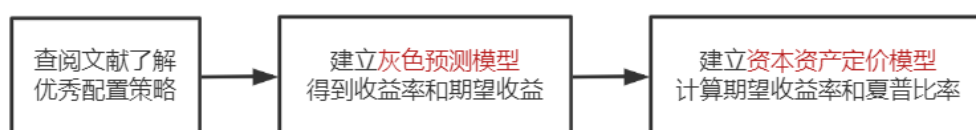
2.3 问题三的分析

附件二提供了股票、大宗商品、债券和现金及其等价物四类资产的指数行情数据。题目要求我们从附件二所给的指数行情数据中挑选出能够代表四类资产的四个指数，并预测大类资产指数在问题一划分出的各个经济状态下的风险收益特征以及大类资产指数之间的相关性。这里用标准归一化对数据进行处理，计算数据的方差，进而计算出各指数的风险收益特征，再运用灰色关联度分析各大类资产指数之间的相关性。



2.4 问题四的分析

问题四要求我们基于所建立的模型预测出的国内未来五年的经济状态，选出合适的大类资产指数构建投资组合，并预测投资组合的风险收益特征。我们沿用在问题三中已求出的四个大类资产指数的年收益率等数据，根据参考文献初步确定大类资产指数的投资组合，建立灰色预测模型预测收益率和期望收益等数据，并构建资本资产定价模型（CAPM）计算期望收益率和夏普比率。



三、基本假设与符号说明

3.1 基本假设

1. 假设附件数据没有修改经过修改，真实有效；
2. 假设未来五年，新冠疫情对全球经济的影响越来越小；
3. 假设世界未来五年不再发生重大战争、重大公共卫生事件以及重大金融危机。

3.2 符号说明

符 号	含 义	单 位
W_j	数据完整度	\
E_i	指标影响度	\
D_i	数据离散度	\
C_v	离散系数	\
CEM_i	评价总分	\
$E(X_{pj})$	某个时期某个指数的期望收益	\
$\zeta_i(k)$	灰色相关系数	\
ξ_j	灰色关联度	\
T_j	资产组合总收益率	\
$E(r_i)$	第 i 个资产的期望收益率	\

四、 数据预处理

本题提供了两个数据附件，整体数据量很大。附件一是宏观经济指标数据，共有九大类经济指标，分别为财政、工业、国民经济核算、价格指数、景气指数、就业与工资、利率汇率、人口以及银行与货币。附件二是大类资产指数行情数据。

4.1 对附件一的数据预处理

九大类经济指标里，一共有 31 个细致分化后的经济指标。本文对此附件一数据预处理的方法为去除数据量较大且复杂的指标和数据量缺失较多的指标。据统计，数据量较大且复杂的指标有工业企业主要经济指标：分行业（利润总额、营业利润、营业收入、主营业务收入）、中国宏观杠杆率（季）以及人民币存贷款利率。数据量缺失较多的指标有工业企业经济效益指标（月）、PMI、PPI、PPIM、RPI、CGPI、货币供应量、央行货币工具（日）、央行货币政策（日）、公开市场操作（月）。同时，数据既较大复杂，又缺失数据较多的指标有工业产品产量当月值和拆借回购利率。本文直接剔除以上 18 个经济指标，后文也不再考虑这些指标，所以初步得到 13 个具有一定适用度的经济指标，从而提高本文的推理分析与预测结果的有效性与准确度。

a. GDP

附件一给的是名义 GDP，有现价和支出法两种，本文用现价来继续进一步把名义 GDP 转化为实际 GDP，来减小通货膨胀的影响。

$$\text{实际 GDP} = \text{名义 GDP} \div (1 + \text{CPI})$$

b. 就业

附件一的数据分为就业人数和经济活动人口两类，本文将做进一步处理，得到就业率。

$$\text{就业率} = \text{就业人数} \div \text{经济活动人口} \times 100\%$$

c. 景气指数

附件一的数据分为预警指数、一致指数、先行指数和滞后指数，本文将它们做平均处理，得到平均景气指数。

$$\text{平均景气指数} = (\text{预警指数} + \text{一致指数} + \text{先行指数} + \text{滞后指数}) \div 4$$

d. CPI

本文使用附件一提供的当月同比的数据。

e. 工业增加值

本文使用附件一提供的全部工业增加值的数据。

f. 银行间同业拆借利率

本文使用附件一提供的银行间同业拆借: 加权平均利率 (当月值) 的数据。

g. 公开市场操作 (周)

本文使用附件一提供的公开市场操作 (货币净投放) 的数据。

h. 国家财政收支

本文使用附件一提供的财政收支差额 (当月值) 的数据。

i. 工资

本文使用附件一提供的平均工资的数据。

4.2 对附件二的数据预处理

附件二提供大量的大类资产指数行情的数据, 其中包括各类指数, 上证 50、沪深 300、中证 500、中证 1000、南华商品指数、标普高盛商品全收益指数、中债-综合财富指数 (1 年以下、3-5 年、7-10 年) 以及货币基金。提供统计, 发现除了标普高盛商品全收益指数, 其他指数在 1990 年至 2002 年都是缺失了数据。但基于这些指数本身就是权重平均的性质, 故缺失的数据对后续的计算影响不大, 故数据暂时保留。

五、 问题一的模型建立与求解

本题要求根据附件一的数据寻找到高频且有效的宏观经济指标, 然后将 2001 年-2021 年中国的宏观经济运行情况划分成不同的经济状态。本文至此的数据已经经过一轮的数据预处理, 但为了获得更加有效的数据, 在下文仍然会进行数据再处理。

5.1 建模前的准备

据有效统计发现, 可知附件一的数据仍然很多, 本题要求的是分析国内 2001 年-2021 年的宏观经济运行状态, 因为过早年份的数据对现在的影响较小, 因此本文的做法是只考虑 1995 年-2021 年的数据 (可能会出现有些经济指标不满足这个范围)。使用 Excel 软件做此数据再处理, 分别得到 GDP、工业增加值、CPI、中国大宗商品价格总指数、超储率、银行间同业拆借利率、公开市场操作、国家财政收支、就业 (就业率)、工资、景气指数、信心指数以及人口经济指标的 1995 年-2021 年的数据。

建立随机变量 i , 即 $i=1, 2, \dots, 13$ 依次代表 GDP、工业增加值、CPI、中国大宗商品价格总指数、超储率、银行间同业拆借利率、公开市场操作、国家财政收支、就业 (就业率)、工资、景气指数、信心指数以及人口。

5.2 模型的建立

为了寻找到高频有效的经济指标，本文将建立综合评价模型，然后再使用层次分析法给予各评价指标相应的权重，从而计算出各经济指标的分数，然后按分数由高到低依次进行排名，最后再综合考虑选择前几名的经济指标。

5.2.1 评价指标的建立

a. 数据完整度 W_i

由于经过了数据预处理之后的经济指标数据，理应数据缺失较少，但为了再次确保数据的完整性，所以需要再次做数据完整度的检查。本文用 W_i 代表数据完整度。

b. 指标影响度 E_i

指标影响度表示经济指标对宏观经济状况的影响力度。本文用 E_i 代表指标相关度。

c. 数据离散度 D_i

数据离散度表示的是各经济指标的离散系数。其中，离散系数又称变异系数，是统计学当中的常用统计指标。离散系数是测度数据离散程度的相对统计量，主要是用于比较不同样本数据的离散程度。离散系数大，说明数据的离散程度也大；离散系数小，说明数据的离散程度也小。本文用 D_i 代表数据离散度。本文给该评价指标打分时，会侧重离散系数大的给予更高的分数。选取离散系数大的经济指标，更加有利于后续划分经济状况。

5.2.2 离散系数的计算与结果展示

记 σ 为标准差， μ 为平均值， c_v 表示总离散系数和样本离散系数，离散系数的计算公式可表示为：

$$c_v = \frac{\sigma}{\mu} \quad (1)$$

通过 Excel 计算得到离散系数的计算结果如下：

指标	GDP	同业拆借利率	大宗商品价格总指数	公开市场操作	就业率
离散系数	0.81	0.07	0.174388426	182.12	0.012
指标	CPI	景气指数	工业增加值	国家财政收支	人口
离散系数	1.1	0.03	0.7242249	0.48	1.19
指标	工资	信心指数	超储率		
离散系数	0.82	0.030557	0.556		

注 1. 公开市场操作经济指标的离散系数与其他经济指标的差距过大，本文的求解将不采用这个经济指标。

5.2.3 评价指标的权重确立

为了更加客观的确立各评价指标的权重，本文采用层次分析法中的判断矩阵来进行权重的划分。层次分析法是指将与决策总是有关的元素分解成目标、准则、方案等层次，在此基础之上进行定性和定量分析的决策方法。为了高效地获得评价指标地权重信息，故只使用层次分析法中的判断矩阵，而不完整地使用层次分析法。

a. 构造判断 (成对比较) 矩阵

$$a_{xy} = \frac{1}{a_{yx}}, \quad \text{其中 } x, y = W, E, D \quad (2)$$

b. 判断矩阵的展示

$\begin{matrix} y \\ x \end{matrix}$	E	D	W
E	1	2	4
D	$\frac{1}{2}$	1	2
W	$\frac{1}{4}$	$\frac{1}{2}$	1

c. 权重确立

指标影响度 E 获得 0.57 权重比，数据离散度 D 获得 0.29 权重比，数据完整度 W 获得 0.14 权重比。

5.2.4 综合评价指标模型的建立

a. 综合评价计算方程

$$CEM_i = 0.57 \times E_i + 0.29 \times D_i + 0.14 \times W_i \quad (i = 1, 2, \dots, 13) \quad (3)$$

b. 数据完整度

指标	GDP	同业拆借利率	大宗商品价格总指数	公开市场操作	就业率
数据完整度	100	94	98	98	100
指标	CPI	景气指数	工业增加值	国家财政收支	人口
数据完整度	100	100	95	97	100
指标	工资	信心指数	超储率		
数据完整度	98	100	97		

c. 指标影响度

指标	GDP	同业拆借利率	大宗商品价格总指数	公开市场操作	就业率
指标影响度	100	82	83	85	90
指标	CPI	景气指数	工业增加值	国家财政收支	人口
指标影响度	99	96	78	93	70
指标	工资	信心指数	超储率		
指标影响度	78	90	89		

d. 数据离散度

指标	GDP	同业拆借利率	大宗商品价格总指数	公开市场操作	就业率
数据离散度	88	76	78	异常	70
指标	CPI	景气指数	工业增加值	国家财政收支	人口
数据离散度	96	72	84	80	100
指标	工资	信心指数	超储率		
数据离散度	92	74	82		

5.2.5 综合评价模型的求解与展示

经统计得出评价总分情况如下表所示：

指标	评价总分	指标	评价总分
CPI	98.27	就业率	85.6
GDP	96.52	工资	84.86
国家财政收支	91.5	大宗商品价格总指数	83.65
景气指数	89.6	人口	82.9
超储率	88.09	工业增加值	82.12
信心指数	86.76	同业拆借利率	81.94

5.3 选取高频有效的经济指标

本文选取 CPI、GDP、国家财政收支（差额）、（平均）景气指数作为高频有效的经济指标。

5.4 高频有效的经济指标的时序图

通过 Python 编程得到 CPI、GDP、（平均）景气指数和国家财政收支差额的时序图如下：

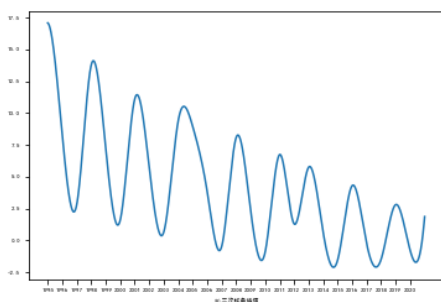


图 5-1: 2001-2021CPI

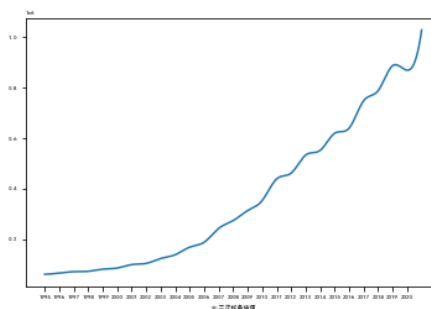


图 5-2: 2001-2021GDP

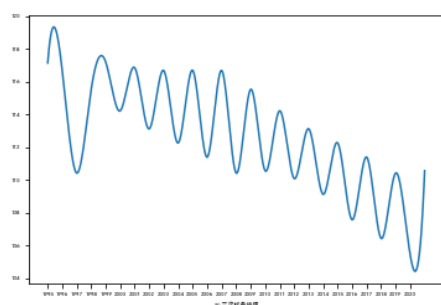


图 5-3: 2001-2021 景气指数

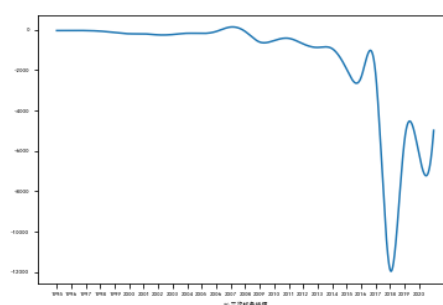


图 5-4: 2001-2021 财政收支差额

5.5 综合分析与经济状况的划分

将 CPI、GDP、宏观经济指数、财政收入差额这四组数据结合时序图进行综合分析，得到以下经济状况划分及具体分析：

● 2001-2005 年——复苏

CPI 波动较大，整体有下降趋势，居民消费指数波动大且消费潜力不断提高。GDP 保持稳定增长，财政赤字逐年减小，说明经济稳步提高，市场向好。宏观经济指数维持在 110-117 的区间在波动，宏观经济很景气，人们对未来的经济状况很有信心。因此，2001-2005 年期间划分为美林时钟模型的复苏阶段。

● 2006-2008 年——过热

CPI 波动较上一时期有所减小，甚至在 2007 年牛市中掉下零轴呈现负值。GDP 高速增长，股市呈现一片繁荣景象，财政收入呈正数，宏观经济指数维持在 110-116 之间，投资信心前所未有的乐观。

● 2009-2013 年——滞涨

经过 2006-2008 年经济大繁荣后经济过热滞涨，CPI 上涨，人口红利带来的 GDP 增长无法满足过热经济的需求。

● 2014-2015 年——衰退

GDP 增长速度放缓，财政赤字增大，经济景气指数下降到 110 位置，2015 年爆发股灾，CPI 降到历史最低位，市场中货币流通量减少，人们倾向于持有现金，投资信心大幅下降，该时期经济处于衰退阶段。

● 2016-2019 年——复苏

财政赤字开始大幅增加，政府运用宏观调控手段调整经济萧条景象，开始追求经济转型升级。CPI 下降到 5% 甚至 3%，价格指数稳定在 $-0.5\% \sim 3\%$ 之间。GDP 实现翻倍的年限较以往周期大幅缩短，仅用四年时间实现了从 500k 亿元到 1000k 亿元的增长。宏观经济景气指数维持在 105-115 之间，仍处于经济景气区间，下降速度减缓，经济处于复苏状态。

● 2020 年——衰退

遭受疫情影响，全国停工停业，GDP 急剧下跌，宏观经济景气指数饱受打击，CPI 上涨，医疗方面财政支出增大，给予的经济方面支出相对减少，人们对市场信心下降，宏观经济出现短暂但严重的衰退。

● 2021 年——复苏

疫情得到控制，全国陆续复工复产，财政支出增大以促进消费带动生产，GDP 快速上升恢复到正常水平，同时 CPI 略降，人们投资和消费信心提高，宏观经济景气指数高涨。

六、 问题二的模型建立与求解

在问题一的求解过程中，本文通过建立综合评价模型选出了四个高频有效的宏观经济指标，分别是 CPI、GDP、国家财政收支（差额）、（平均）景气指数。问题二是在问题一的基础上，通过建立宏观经济模型模拟中国未来五年的经济增长、通胀、利率（反映货币政策松紧程度）等宏观经济环境，并解释未来五年中国将面临问题一中划分的经济状况里的什么经济状态。未来五年，即 2023 年、2024 年、2025 年、2026 年以及 2027 年。本文将使用时间序列分析对其进行预测分析。

6.1 建模前的准备

由于在问题一处理时，数据年份大部分是到 2021 年的，且大部分经济指标也缺乏完整的 2022 年数据，因此，本文将 2022 年也作为预测年。此外，本文除了继续采用 CPI、GDP、国家财政收支（差额）、（平均）景气指数这四个经济指标外，还将加入附件一数据中的人民币存贷款利率。同时，也需要对其进行数据处理。通过观察人民币存贷款利率的数据发现，其数值在过去保持在 $4 \sim 6$ 之间，趋于稳定。因此，对其的数据处理将与上文的方式

不同。本文对其只采用过去 11 年的数据，即 2011 年 7 月至 2022 年 8 月。

6.2 模型的建立

6.2.1 指数平滑法模型的建立

指数平滑法是一种特殊的加权移动平均法。它是在移动平均法基础上发展起来的一种时间序列分析预测法，是通过计算指数平滑值，配合一定的时间序列预测模型对现象的未来进行预测。其原理是任一期的指数平滑值都是本期实际观察值与前期指数平滑值的加权平均。根据平滑的次数不同，可分为一次指数平滑法和二次指数平滑法，或者更高次。当时间序列的变动出现直线趋势时，用一次指数平滑法进行预测，仍存在明显的滞后偏差。因此，也必须对其加以修正，再做二次指数平滑，利用滞后偏差的规律建立直线趋势模型。故本文使用的是二次指数平滑法。

计算及推导过程如下：

$$\begin{cases} S_t^{(1)} = \alpha y_t + (1 - \alpha)S_{t-1}^{(1)}, \\ S_t^{(2)} = \alpha S_t^{(1)} + (1 - \alpha)S_{t-1}^{(2)} \end{cases} \quad (4)$$

式中 $S_t^{(1)}$ 为一次指数的平滑值； $S_t^{(2)}$ 为二次指数的平滑值。当时间序列 $\{y_t\}$ 从某时期开始具有直线趋势时，可用直线趋势模型

$$\hat{y}_{t+m} = a_t + b_t m, \quad m = 1, 2, \dots \quad (5)$$

$$\begin{cases} a_t = 2S_t^{(1)} - S_t^{(2)} \\ b_t = \frac{\alpha}{1-\alpha} (S_t^{(1)} - S_t^{(2)}) \end{cases} \quad (6)$$

令 $m=1$ ，可得

$$\hat{y}_{t+1} = 2S_t^{(1)} - S_t^{(2)} + \frac{\alpha}{1-\alpha} (S_t^{(1)} - S_t^{(2)}) \quad (7)$$

6.2.2 非平稳时间序列模型的建立

经过二次指数平滑法的预测，已经初步得到了 2022 年-2027 年预测值。但由于二次指数平滑法的缺陷，造成了不可避免的误差出现。因此，为了减少误差，本文将继续使用非平稳时间序列模型继续优化误差，并做最后一步的预测。

a. 差分运算

差分运算是一种非常方便且有效的确定性信息提取方法。

$$\nabla^d X_t = (1 - B)^d X_t = \sum_{i=0}^d (-1)^i C_d^i X_{t-i} \quad (8)$$

b. ARIMA 模型

对差分运算后得到的平稳序列可用 ARIMA 模型进行拟合。

$$\begin{cases} \phi(B)\nabla^d X_t = \theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, \quad \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, \quad E(\varepsilon_t \varepsilon_s) = 0, \quad s \neq t \\ E(X_s \varepsilon_t) = 0, \quad \forall s < t \end{cases} \quad (9)$$

6.2.3 模型的求解与展示

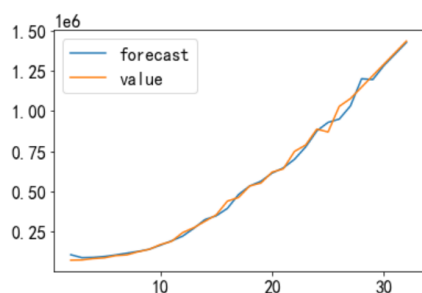


图 6-5: GDP

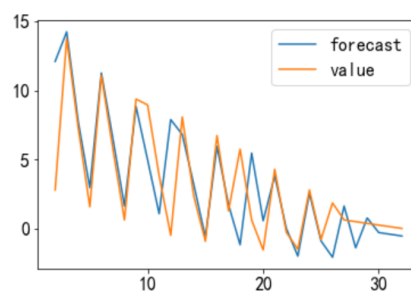


图 6-6: CPI

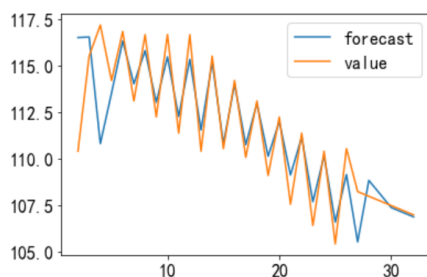


图 6-7: 景气指数

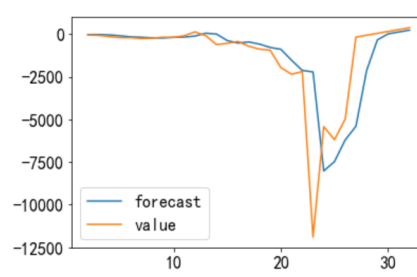


图 6-8: 财政收支差额

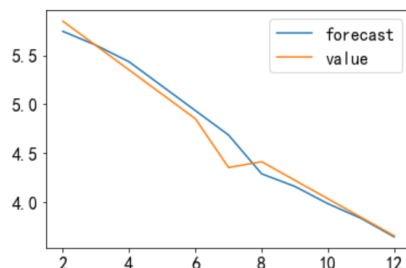


图 6-9: 人民币存贷款利率

注 2. *value*: 原始数据; *forecast*: 预测数据.

6.3 中国未来五年经济状况分析

由预测数据可知，未来五年 GDP 将持续上升，CPI 将持续下降，景气指数波动较大但维持在 110~120 较景气区间中，政府的财政收支维持在安全稳定的赤字区间，人民币存贷款利率下降，增加了市场货币流通量，促进了经济发展。

综上，未来五年中国经济状况处于复苏期。

七、 问题三的模型建立与求解

本问要求挑选出能够代表四类大类资产的四个指数，然后预测大类资产指数在问题一中划分出的各种经济状态下的风险收益特征，以及分析大类资产指数之间的相关性。基于这类指数本身的属性，本文决定采用计算它们的方差来进行选择哪些指数，然后接着计算其风险收益特征，最后用灰色关联分析来分析各类大类资产指数的相关性。

7.1 建模前的准备

因为年代越久远的指数数据对当下的影响呈递减趋势，故本问综合考虑决定只采用过去十年的指数数据，即 2012 年 6 月 29 日到 2022 年 6 月 30 日之间的数据。因为各指数数据的大小相差巨大，不能直接将其进行方差计算然后对比，所以本文用 Python 软件对其进行最大-最小标准归一化处理。

最大-最小标准化是对原始数据进行线性变换，设 $\min A$ 和 $\max A$ 分别是属性 A 的最小值和最大值，将 A 的一个原始值 x 通过最大-最小标准化映射到区间 $[0,1]$ 的值 x' ，那么公式如下：

$$x' = \frac{x - \min A}{\max A - \min A} \quad (10)$$

7.2 模型的建立与求解结果的展示

方差是在概率论和统计衡量随机变量或一组数据时离散程度的度量。概率论中方差用来度量随机变量和其数学期望（即均值）之间的偏离程度。统计中的方差（样本方差）是每个样本值与全体样本值的平均数之差的平方值的平均数。方差越大，数据的波动也就越大；方差越小，数据的波动也就越小。

7.2.1 方差的计算公式

$$D(x) = \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x}_j)^2 \quad (11)$$

其中 $j=1,2,\dots,10$ （依次代表上证 50，沪深 300， \dots ，货币基金）， $D(x)$ 表示方差， x_j 依次表示各个指数的各个值， \bar{x}_j 依次表示各个指数的平均值。

7.2.2 方差的计算结果

通过 Excel 函数求出方差的计算结果如下：

指数名称	方差
上证 50	0.056
沪深 300	0.058
中证 500	0.245
中证 1000	0.243
南华商品指数	0.038
标普高盛商品全收益指数	0.074
中债-综合财富(一年以下) 指数	0.089
中债-综合财富(3-5 年) 指数	0.091
中债-综合财富(7-10 年) 指数	0.08
货币基金	0.09

7.2.3 指数的选取

本题的目标是从股票指数中（上证 50、沪深 300、中证 500 以及中证 1000）四选一，从大宗商品指数中（南华商品指数与标普高盛商品全收益指数）二选一，从债券指数中（中债-综合财富(1 年以下、3-5 年以及 7-10 年) 指数中三选一。因现金及其等价物指数里附件二数据只提供了货币基金，故本文将直接选择它。

根据参考文献可知，相关研究人员表示上证 50 指数在代表性、可投资性、强抗操纵性以及参考价值方面都不如沪深 300 指数。又在股指期货定价中,特别是对上证 50 的定价,完全可以使用不完全市场条件下隐含收益率定价模型,而中证 500 股指期货的定价还需要结合国内的金融政策进行分析。同时本文对于股票指数的方差考虑选择方差大的指数，波动大说明可能存在高风险、高回报的时机。基于平稳性，其他指数的选择考虑方差小的指数。

最终本文选择**沪深 300、南华商品指数、中债-综合财富(7-10 年) 指数以及货币基金**。

7.3 风险收益特征

对于风险收益特征，本文选择期望收益。

7.3.1 期望收益

$$E(X_{pj}) = \frac{1}{n} \sum_{t=1}^n X_{pjt} \quad (12)$$

其中, $E(X_{pj})$ 表示某个时期某个指数的期望收益, $p = 1, 2, \dots, 7$ (依次代表 2001-2005 年, 2006-2008 年, 2009-2013 年, 2014-2015 年, 2016-2019 年, 2020 年, 2021 年), t 表示日期。

7.3.2 风险收益特征模型的求解与展示

	沪深 300	南华商品指数	中债-综合财指	货币基金
2001-2005(复苏)	1143	4532	104.6	978
2006-2008(过热)	2840	6815	116.5	1038
2009-2013(滞涨)	2770	4652	138.6	1157
2014-2015(衰退)	3150	3743	158.6	1318
2016-2019(复苏)	3565	2413	182.5	1457
2020(衰退)	4383	1805	206.5	1567
2021(复苏)	5088	2515	213.3	1601

7.4 大类资产指数相关性

1. 数据标准化

前文已经将数据标准化, 设标准化后的数据矩阵元素为 R_{ij} , 由上可得指标正向化后数据矩阵元素为 $(x'_{ij})'$, 则有:

$$R_{ij} = \frac{x'_{ij} - \min(x'_j)}{\max(x'_j) - \min(x'_j)} \quad (13)$$

2. 计算灰色相关系数

常见的灰色相关系数表达式如下:

$$\zeta_i(k) = \frac{\min_i \min_k |s_0(k) - s_i(k)| + \rho \cdot \max_i \max_k |s_0(k) - s_i(k)|}{|s_0(k) - s_i(k)| + \rho \cdot \max_i \max_k |s_0(k) - s_i(k)|} \quad (14)$$

设 $s_0(k)$ 是参考列, ρ 是分辨系数。它的范围为 0~1, 它的作用是为控制区分度, 它的值越小, 区分度越大; 它的值越大, 区分度越小。常常取 0.5。

3. 计算灰色关联度系数

a. 参考向量的选择

假设我们是研究 s_1 指标与 s_2 指标之间的灰色关联度。所以将 s_2 列作为参考向量, 即要研究与谁的关系, 就将谁作为参考。设参考向量为 $Y_1 = s_2$, 生成新的数据矩阵 $S_2 = s_1$ 。

b. 生成绝对值矩阵

设 A 即为生成的绝对值矩阵 $A = [S_1 - Y_1]$, 设 d_{\max} 为绝对值矩阵 A 的最大值, d_{\min} 为绝对值矩阵 A 的最小值。

c. 计算灰色关联矩阵

设灰色关联矩阵为 B，则有：

$$B_{ij} = \frac{d_{\min} + \rho d_{\max}}{A_{ij} + \rho d_{\max}} \quad (15)$$

d. 计算灰色关联度

$$\xi_j = \frac{\sum_{i=1}^m B_{ij}}{m} \quad (16)$$

7.5 大类资产指数相关性的求解与展示

使用 Python 软件求解结果如下：

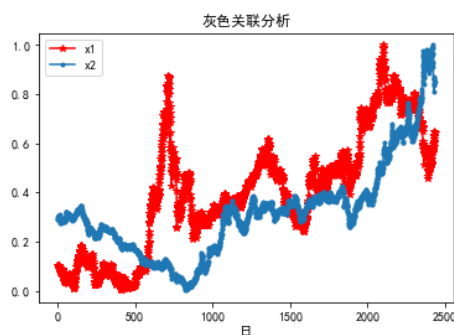


图 7-10: 沪深 300 与南华商指

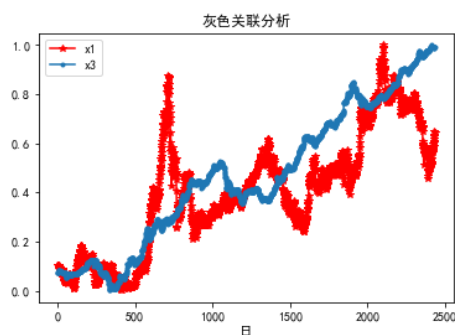


图 7-11: 沪深 300 与中债-综合财指

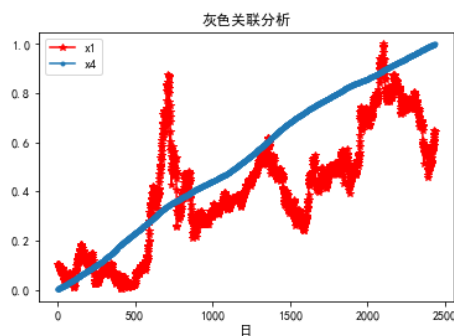


图 7-12: 沪深 300 与货币基金

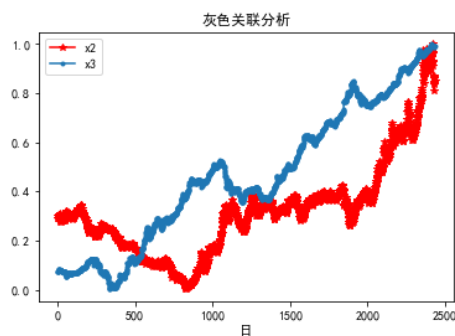


图 7-13: 南华商指与中债-综合财指

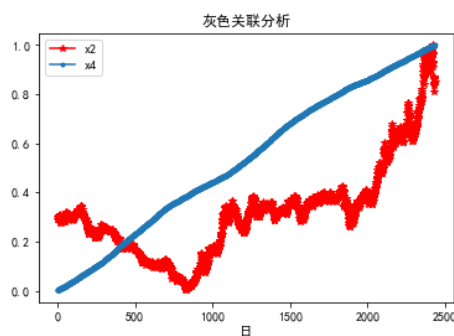


图 7-14: 南华商指与货币基金

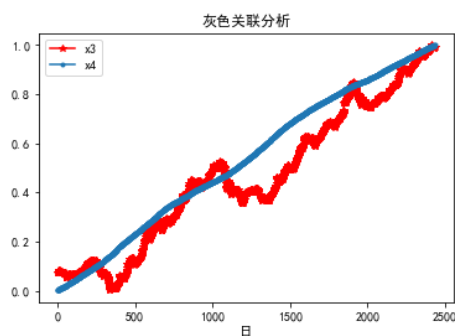


图 7-15: 中债-综合财指与货币基金

- a. 沪深 300 与南华商品指数的灰色关联度为：0.749.
- b. 沪深 300 与中债-综合财富 (7-10 年) 指数的灰色关联度为：0.815.
- c. 沪深 300 与货币基金的灰色关联度为：0.766.
- d. 南华商品指数与中债-综合财富 (7-10 年) 指数的灰色关联度为：0.718.
- e. 南华商品指数与货币基金的灰色关联度为：0.667.
- f. 中债-综合财富 (7-10 年) 指数与货币基金的灰色关联度为：0.865.

八、 问题四的模型建立与求解

本问的要求是基于本问的模型预测出的国内未来五年的经济状态，然后选择出合适的大类资产指数构建投资组合，并预测投资组合的风险收益特征。本问接着使用在问题三中已求出来的四个大类资产指数（沪深 300、南华商品指数、中债-综合财富 (7-10 年) 指数以及货币基金）的年收益率等数据。然后，本文根据参考文献初步确定大类资产指数的投资组合。对于未来五年的各个数据的预测，将使用灰色预测模型进行预测。并且构建资本资产定价模型（CAPM）计算期望收益率等。

8.1 大类资产投资组合的确立

根据上文的预测，基于美林时钟投资理论，国内未来五年的经济状态为复苏期。

8.1.1 数据处理

为了求四个指标的年收益率，使用 Python 软件求得沪深 300、南华商品指数、中债、货币基金四个指标的 2012 年 6 月 29 到 2022 的 6 月 30 日的日收益率，再通过 Excel 将四个指标的日收益率整合在同一个 Excel 表中，然后通过四个指标的日收益率进行取平均值得到每一年的年收益率。

8.1.2 投资组合的选取

根据参考文献的结论得，美林投资时钟的复苏期，经济特征为经济开始上行；通货膨胀率继续下降；利率也继续下降；物价也下行，需求逐渐增加。资产特征为此阶段股票相对债券和现金更具有可能超额收益。故股票类资产应该占优。同时，大类资产份额配置应是：股票 > 债券 > 现金 > 大宗商品。所以本文对于投资组合的权重份额 W_i 分配，也是如此，股票 > 债券 > 现金 > 大宗商品。

W_i 为权重份额， $i=1, 2, 3, 4$ ，分别代表沪深 300，南华商品指数、中债-综合财富 (7-10 年) 指数以及货币基金。其中的任意一个元素 W_i 是财富

分配在第 i 种资产上的比例，分配在所有资产上的财富比例之和应该为 1。

$$\sum_{i=1}^n w_i = 1 \quad (17)$$

8.1.3 权重份额的讨论

对于资本资产定价模型的计算，本文将会设置两种权重进行讨论。第一种权重为， $W_1 = 0.4$ ， $W_2 = 0.3$ ， $W_3 = 0.2$ ， $W_4 = 0.1$ ；第二种权重为 $W_1 = 0.6$ ， $W_2 = 0.2$ ， $W_3 = 0.15$ ， $W_4 = 0.05$ 。

8.1.4 资产组合总收益率

$$T_j = \sum_{i=1}^n W_i \cdot E(r_i) \quad (18)$$

T_j 表示资产组合总收益率， $j=1,2$ ； $E(r_i)$ 表示第 i 个资产的期望收益率。

8.2 灰色预测模型

灰色预测模型是通过少量的、不完全的信息，建立数学模型并做出预测的一种预测方法。

8.2.1 灰色预测模型的建立

A. 数据的检验

为了保证模型的有效性，需要对数据作必要的检验处理。设置参考序列的级比：

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, \quad k = 2, 3, \dots, n \quad (19)$$

B. 模型的建立

a. 建立微分方程模型（对应的白化方程）

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b \quad (20)$$

b. 为了识别模型参数 a, b ，在区间 $k-1 < t \leq k$ 上，令

$$\begin{aligned} x^{(1)}(t) &= z^{(1)}(k) = \frac{1}{2} [x^{(1)}(k-1) + x^{(1)}(k)] \\ \frac{dx^{(1)}(t)}{dt} &= x^{(1)}(k) - x^{(1)}(k-1) = x^{(0)}(k) \end{aligned} \quad (21)$$

c. 灰色微分方程

$$x^{(0)}(k) + az^{(1)}(k) = b, \quad k = 2, 3, \dots, n \quad (22)$$

d. 记 $\mathbf{u} = [a, b]^T, \mathbf{Y} = [x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)]^T$,

$$\mathbf{B} = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix},$$

则由最小二乘法求得使 $J(\mathbf{u}) = (\mathbf{Y} - \mathbf{Bu})^T(\mathbf{Y} - \mathbf{Bu})$ 达到最小值的 \mathbf{u} 的估计值 $\hat{\mathbf{u}} = [\hat{a}, \hat{b}]^T = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{Y}$.

接着求解 a 的微分方程得

$$\hat{x}^{(1)}(t) = \left(x^{(0)}(1) - \frac{\hat{b}}{\hat{a}} \right) e^{-\hat{a}t} + \frac{\hat{b}}{\hat{a}}, \quad (23)$$

e. 即得到预测值

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{\hat{b}}{\hat{a}} \right) e^{-\hat{a}k} + \frac{\hat{b}}{\hat{a}}, \quad k = 0, 1, 2, \dots, \quad (24)$$

而且 $\hat{x}^{(0)}(1) = \hat{x}^{(1)}(1), \hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), k = 1, 2, \dots$.

C. 误差检验（相对误差检验）

计算相对误差：

$$\delta(k) = \frac{|x^{(0)}(k) - \hat{x}^{(0)}(k)|}{x^{(0)}(k)}, \quad k = 1, 2, \dots, n \quad (25)$$

这里 $\hat{x}^{(0)}(1) = x^{(0)}(1)$. 如果 $\delta(k) < 0.2$, 则可认为达到一般要求; 如果 $\delta(k) < 0.1$, 则认为达到较高的要求.

8.2.2 模型的求解与展示

a. 年收益率

指数 年份	沪深 300	南华商品指数	中债-综合财富 (7-10 年) 指数	货币基金
2023	0.000084	0.0116	0.000177	0.000083
2024	0.000006	-0.4823	0.000123	0.000074
2025	\	-0.0000002	0.000145	0.000071
2026	\	\	0.00017	0.000066
2027	\	\	0.00013	0.000052

平均处理，最终得到国内未来五年各指数的预测年收益率：

指数名称	沪深 300	南华商品指数	中债-综合财富 (7-10 年) 指数	货币基金
平均收益率	0.000045	-0.0000002	0.000149	0.000069

注 3. 因沪深 300 与南华商品指数的数值带有负数，所以此算法预测的效果不佳，沪深 300 的 2025-2027 年无法预测，南华商品指数的整体误差较大，所以使用 2025 年的数值来代表年收益率。

b. 期望收益

	沪深 300	南华商品指数	中债-综合财富 (7-10 年) 指数	货币基金
2023	5145	2219	233.83	1706.79
2024	5415	2467	244.52	1752.99
2025	5591	2743	254.65	1800.95
2026	6053	3042	266.9	1851.67
2027	6402	3354	281.73	1902.81

8.3 数字特征

a. 方差 $\text{Var}(\text{rm})$

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N} \quad (26)$$

b. 标准差

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \quad (27)$$

c. 协方差

$$\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y] \quad (28)$$

8.4 资本资产定价模型 (CAPM)

资本资产定价模型是在资产组合理论和资本市场理论的基础上发展起来的，主要研究证券市场中资产的预期收益率与风险资产之间的关系，以及均衡价格是如何形成的，是现代金融市场价格理论的支柱，广泛应用于投资决策和公司理财领域。

资本资产定价方程：

$$E(r_i) = r_f + \beta_{im}(E(r_m) - r_f) \quad (29)$$

其中, $E(r_i)$ 是资产 i 的预期回报率 (或普通股的资本成本率), r_f 是无风险利率 (通常以短期国债的利率来近似代替), β_{im} 是 Beta 系数, 即资产 i 的系统性风险。

$$\beta = \frac{Cov(r_i, r_m)}{Var(r_m)} \quad (30)$$

$E(r_m)$ 是市场 m 的预期市场回报率 (通常用股票价格指数收益率的平均值或者所有股票的平均收益率来代替), $E(r_m) - r_f$ 是市场风险溢价, 即预期市场回报率与无风险回报率之差。

8.5 夏普比率

夏普比 (SR) 是一个在证券投资中被普遍使用的指标。夏普比衡量了通过承担更多的风险 (更大的波动率) 来获得更高期望回报率的效率。夏普比越高的资产, 承担同样风险则能获得更多期望回报率的上升。

计算公式为:

$$SR_i = \frac{E(r_i) - r_f}{\sigma(r_i)} \quad (31)$$

其中, SR_i 代表第 i 项资产的夏普比率。

8.6 模型的求解与展示

8.6.1 前提说明

(1) 无风险利率 r_f 的数据来源中国货币网 2015 年-2021 年的 3 个月期国债发行利率, 然后做均值处理, 获得平均短期国债利率, 即 $r_f = 0.0248$ 。

(2) $E(r_m)$ 市场投资组合 m 的期望收益率用各指数的平均收益率来代替, 即 $E(r_m) = 0.000206$ 。

(3) 各指数年收益率方差为

	沪深 300	南华商品指数	中债-综合财富 (7-10 年) 指数	货币基金
年收益率方差	0.00000007	0.00000006	0.000000003	0.000000001

(4) 各指数年收益率协方差由于经过实际计算后的数值实在太小了, 不易进行后续的计算, 所以决定用 0.000000001 来代替。

(5) 各指数的标准差为 0.014

8.6.2 资本资产定价模型的求解结果

a. 各个指数的资产定价结果

	沪深 300	南华商品指数	中债-综合财富 (7-10 年) 指数	货币基金
$E(r_i)$	0.024765	0.024759	0.008198	0.000206

b. 资产组合总收益率的求解结果

	组合 1	组合 2
总收益率	0.019	0.021

8.6.3 夏普比率的求解结果

	沪深 300	南华商品指数	中债-综合财富 (7-10 年) 指数	货币基金
夏普比	-1.758	-1.751	-1.756	-1.762

8.7 结论

最终经过资本资产定价模型计算，第二种权重的投资组合比第一种的投资组合占优。

九、模型的评估与改进

9.1 模型的优点

- (1) 本文的各种模型全面，逻辑思路完整，文字表述准确；
- (2) 本文的代码简洁且高效，预测的结果准确；
- (3) 本文所选择的四个大类资产指数覆盖了大部分国内股票市场，具有一定的代表性。

9.2 模型的缺点

- (1) 本文的综合评价模型打分，部分存在主观性，不具全面客观性；
- (2) 由于缺乏相对应的收益率数据，所以本文部分收益率数据由附件数据推算出来，误差难以避免；
- (3) 由于部分数据集的数值过小，算法预测出来的数值可能存在较大误差。

9.3 模型的改进推广

本文建立的预测模型，由于缺乏历年短期国债利率和历年股票收益率等数据，所以本文的部分求解结果不具普遍性。如果获取到这些数据，本文的预测模型求解结果的预测结果将大大提升。本文还应该对美林投资时钟理论进行发展推广，继续划分更加详细的周期，从而提高预测的准确性。此外，还可以以美林投资时钟理论为基础，建立真正符合中国大类资产市场的周期模型。

参考文献

- [1] 甘倩倩. 中证 500 与上证 50 股指期货定价模型对比研究 [D]. 兰州大学,2018.
- [2] 姜雪伟. 上证综指该让位给沪深 300 了吗?[J]. 金融经济,2010,No.338(16):90.
- [3] 王思韵. 主流大类资产配置量化策略的适用性分析及实证研究 [D]. 西南财经大学,2021.DOI:10.27412/d.cnki.gxncu.2021.002955.
- [4] 孟德峰. 经济周期、行业轮动与 A 股市场投资策略 [D]. 中南财经政法大学,2019.
- [5] 周亮. 基于美林投资时钟的我国大类资产配置探讨 [J]. 上海经济,2018,No.280(01):105-117.
- [6] 刘冰, 林子赫. 投资时钟视角下我国大类资产轮动配置策略 [J]. 山东工商学院学报,2019,33(05):1-20.
- [7] 姚琼. 美林时钟对于中国大类资产配置的有效性检验、改进及预测 [D]. 对外经济贸易大学,2019.DOI:10.27015/d.cnki.gdwju.2019.000116.
- [8] 肖曼君, 刘梦雨. 我国通货膨胀预测模型研究 [J]. 武汉金融,2015,No.188(08):12-15.
- [9] 隋占林. 中国经济增长与通货膨胀预测及影响因素研究 [D]. 东北财经大学,2016.
- [10] 林思亮. 美林时钟视角下中国金融市场资产配置策略研究 [D]. 华侨大学,2019.
- [11] 司守奎, 孙玺菁著,Python 数学实验与建模, 科学出版社,2020 年 4 月第一版.
- [12] 孔威. 时间序列分析在 CPI 中的应用研究 [D]. 延边大学,2014.
- [13] 汤岩. 时间序列分析的研究与应用 [D]. 东北农业大学,2007.
- [14] 李懿洋. 甘肃省产业结构与经济增长的灰色关联分析 [J]. 企业经济,2011,30(05):20-23.DOI:10.13529/j.cnki.enterprise.economy.2011.05.025.
- [15] 王珏. 基于数据特征的灰色预测模型构建及应用研究 [D]. 江南大学,2021.DOI:10.27169/d.cnki.gwqgu.2021.001604.

- [16] 丁圆苹. 灰色预测模型的优化研究 [D]. 河南农业大学,2022.DOI:10.27117/d.cnki.ghenu.2022.000096.
- [17] 余谦. 灰色预测模型对中国经济预测分析 [J]. 现代商贸工业,2011,23(19):12-13.DOI:10.19311/j.cnki.1672-3198.2011.19.006.
- [18] 隋少堂. 风险平价策略在大类资产配置中的应用 [D]. 首都经济贸易大学,2019.DOI:10.27338/d.cnki.gsjmu.2019.000065.
- [19] 张鸿捷. 大类资产配置的主流量化模型实证研究及部分改进 [D]. 山东大学,2018.
- [20] 臧金娟, 黄一黎, 赵学军. 大类资产配置中各类资产最优代表选择的研究 [J]. 金融理论与实践,2017(09):11-18.

附录

附录只包含了涉及核心思想的 Python 源代码，其余可见于支撑材料。

支撑材料清单

问题一 CPI.py

GDP.py

财政支出.py

经济景气.py

第一题数据.xlsx

数据预处理文件夹中 12 个宏观经济指标 (.xlsx) 文件

问题二 指数平滑法.py

非平稳时间序列.py

问题二数据文件夹中 5 个预测值 (.csv) 文件

预测值数据可视化文件夹中 5 个预测值的可视化图 (.png) 文件

问题三 期望收益.xlsx

大类归一化.xlsx

灰色关联度文件夹中四个指数两两之间的指数相关性 (.py) 文件共 6 个

四个指数的收益率文件夹中指数的.py、.xls、.xlsx 文件各 4 个

问题四 ERI.xlsx

各指数年收益率.xlsx

灰色预测模型.py

年收益率初始值.xlsx

四期望收益.xlsx

3 个月期国债发行利率历史数据 (2015-2020 年 6 个.xlsx) 文件

代码

代码一：问题一中 Python 数据可视化处理：

```
# -*- coding: utf-8 -*-
"""
Created on Sun Nov 6 11:04:06 2022

@author: ww
"""

#CPI
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib
from scipy.interpolate import interp1d

x = np.arange(1995,2022,1)
y = (17.07,8.33,2.81,13.76,7.28,1.59,11.09,5.75,0.64,9.38,
      8.95,3.84,-0.47,8.08,2.34,-0.90,6.74,1.30,5.75,0.64,
      -1.53,4.30,-0.25,-1.47,2.81,-0.78,1.87)

xnew = np.linspace(1995,2021,5000) #插值点

f1 = interp1d(x,y,'cubic')
y1= f1(xnew)
plt.rc('font',size=5)
plt.rc('font',family='SimHei')
plt.xticks(range(1995,2021,1))
plt.plot(xnew,y1);plt.xlabel("(B) 三次样条插值")

plt.show()
```

代码二：问题二中 Python 对数据进行指数平滑处理：

```
# 二次指数平滑法
import numpy as np

y = np.array([1,2,3,4,5,6,7,8,9,10]) #在这里修改数据集
n = len(y)
alpha = 0.35
yh = np.zeros(n)

s1 = np.zeros(n)
```

```

s2 = np.zeros(n)
s1[0] = y[0]
s2[0] = y[0]
for i in range(1,n):
    s1[i] = alpha*y[i]+(1-alpha)*s1[i-1]
    s2[i] = alpha*s1[i]+(1-alpha)*s2[i-1]
    yh[i] = 2*s1[i-1]-s2[i-1]+alpha/(1-alpha)*(s1[-1]-s2[-1])

at = 2*s1[-1]-s2[-1]
bt = alpha/(1-alpha)*(s1[-1]-s2[-1])
m = np.array([1,2])
yh2 = at+bt*m
print("预测值为：",yh2)

```

代码三：问题二中 Python 对数据进行非平稳序列处理：

```

import pandas as pd
from statsmodels.graphics.tsaplots import plot_acf
import pylab as plt
from statsmodels.tsa.arima_model import ARIMA

plt.rc('axes',unicode_minus=False)
plt.rc('font',size=16)
plt.rc('font',family='SimHei')
df = pd.read_csv('二利率预测.csv') #在这里修改所需要的数据
plt.subplot(122)
plt.plot(df.value.diff())
plt.title('一次差分')
ax2 = plt.subplot(122)
plot_acf(df.value.diff().dropna(),ax=ax2,title='自相关')

md = ARIMA(df.value,order=(2,1,0))
mdf = md.fit(dis=0)
print(mdf.summary())

residuals = pd.DataFrame(mdf.resid)
fig,ax = plt.subplots(1,2)
residuals.plot(title="残差",ax = ax[0])
residuals.plot(kind='kde',title='密度',ax=ax[1])
plt.legend('')
plt.ylabel('')

mdf.plot_predict() #原始数据与预测值对比图
plt.show()

```

代码四：问题三中 Python 灰色关联度计算两两指数之间的相关性：

```
# -*- coding: utf-8 -*-
"""
Created on Sat Nov 5 16:46:27 2022

@author: ww
"""

#x1到x4

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import matplotlib as mpl
%matplotlib
mpl.rcParams['font.sans-serif'] = ['SimHei']
mpl.rcParams['axes.unicode_minus'] = False
data=pd.read_excel('data数据.xlsx')
print(data)

label_need=data.keys()[1:]
print(label_need)

data1=data[label_need].values
print(data1)

[m,n]=data1.shape
data2=data1.astype('float')
data3=data2
ymin=0.002
ymax=1
for j in range(0,n):
    d_max=max(data2[:,j])
    d_min=min(data2[:,j])
    data3[:,j]=(ymax-ymin)*(data2[:,j]-d_min)/(d_max-d_min)+ymin
print(data3)

t=range(1,2433)
plt.plot(t,data3[:,0], '*-',c='red')

plt.plot(t,data3[:,3], '-.')
plt.xlabel('日')
plt.legend(['x1','x4'])
plt.title('灰色关联分析')
```

```

for i in range(3,4):
    data3[:,i]=np.abs(data3[:,i]-data3[:,0])
data4=data3[:,3:4]
d_max=np.max(data2)
d_min=np.min(data4)
a=0.5 #定义分辨系数
data4=(d_min+a*d_max)/(data4+a*d_max)
xishu=np.mean(data4, axis=0)
print(' x4与 x1之间的灰色关联度为： ')
print(xishu)

```

代码五：问题三中 Python 计算指数的收益率：

```

# -*- coding: utf-8 -*-
"""
Created on Sat Nov 5 19:31:49 2022

@author: ww
"""

import numpy as np
import pandas as pd
from pandas import DataFrame, Series
data = pd.read_excel('货币基金.xlsx')
                #index_col=[0],parse_dates=[0])

#对数收益率
data['货币基金每日收益率']=np.log(data['close']/data['close'].shift(1))[1:]
data1=data['货币基金每日收益率']
x=data1
x.to_excel('货币基金每日收益率(1).xls')

```

代码六：问题四中 Python 建立灰色预测模型：

```

import numpy as np
import sympy as sp
from matplotlib.pyplot import plot,show,rc,legend,xticks
rc('font',size=16); rc('font',family='SimHei')
x0=np.array([1,2,3,4,5,6,7,8,9,10]) #在这里面修改数据
n=len(x0); jibi=x0[:-1]/x0[1:] #求级比
bd1=[jibi.min(),jibi.max()] #求级比范围
bd2=[np.exp(-2/(n+1)),np.exp(2/(n+1))] #q求级比的容许范围
x1=np.cumsum(x0) #求累加序列
z=(x1[:-1]+x1[1:])/2.0
B=np.vstack([-z,np.ones(n-1)]).T
u=np.linalg.pinv(B)@x0[1:] #最小二乘法拟合参数

```



```

sp.var('t'); sp.var('x',cls=sp.Function) #定义符号变量和函数
eq=x(t).diff(t)+u[0]*x(t)-u[1] #定义符号微分方程
xt=sp.dsolve(eq,ics={x(0):x0[0]}) #求解符号微分方程
xt=xt.args[1] #提取方程中的符号解
xt=sp.lambdify(t,xt,'numpy') #转换为匿名函数
t=np.arange(n+1)
xt1=xt(t) #求模型的预测值
x0_pred=np.hstack([x0[0],np.diff(xt1)]) #还原数据
x2023=x0_pred[-1] #提取2002年的预测值
cha=x0-x0_pred[:-1]; delta=np.abs(cha/x0)*100
print('2012~2023的预测值： ',x0_pred)
print('\n-----\n','相对误差',delta)
t0=np.arange(2012,2023); plot(t0,x0,'*--')
plot(t0,x0_pred[:-1],'s-'); legend(('实际值','预测值'));
xticks(np.arange(2012,2023));
show()

```
