

基于双目标规划模型和蒙特卡洛模拟的基金组合投资研究

摘 要

本文针对设计基金组合投资策略这一问题，利用天天基金网旗下基金产品的相关数据，构建基金综合指标评价体系，通过 ARIMA 时间序列模型预测基金产品下一时期的预期收益，建立双目标规划模型，制定出不同风险偏好的投资者在不同时期的基金组合，使激进型和稳健型投资者都能取得风险和收益的动态平衡。

针对模型一，我们首先对基金产品数据进行预处理和分析，删除数据缺失比例超过 20% 的基金产品，使用同期基金产品的平均数据进行缺失值填补，从中提炼了基金评级、年化收益、最大回撤等 8 个指标，分别从基金历史、基金信誉、基金收益和基金风险四个方面对基金产品进行多维度评估，构建了基金综合指标评价体系（见图 1）。其次，使用以数据驱动为原理的熵权法给各指标赋权（见表 1），从而规避主观性对结果的影响。最后利用 TOPSIS 方法量化基金产品的综合指标，筛选出 20 家综合评分最高的基金产品（见表 2）。

针对模型二，我们建立了 ARIMA 时间序列模型，通过这 20 个基金产品的日度数据预测下一时期基金产品的预期收益，进而计算出基金产品的预期季度收益率（见图 5），为简化模型，我们假定投资者足够理性，买入基金产品后会持有一定时间，操作频率为每季度一次。

针对模型三，我们建立了双目标规划模型，针对不同风险偏好的投资者，我们建立以最小化最大风险和最大化预期收益的双目标规划模型，在模型求解时，由于需要进行多期组合投资决策，我们考虑使用迭代的方法，逐步迭代并更新约束条件，结合蒙特卡洛模拟对各个阶段的双目标规划问题进行求解。

通过以上三个模型，本文设计了针对激进型和稳健型投资者的基金组合投资策略，为检验该模型在基金产品市场中的实际执行效果，本文对两位具有 100 万初始资金的客户设计激进型和稳健型两种基金组合投资策略。模型结果显示，激进型组合投资最后总资产 2085426 元，稳健型组合投资最后总资产 1121002 元。本文展示了两种组合投资策略的第一季度基金购买策略（见表 7 和表 8）。相较于上证指数 2019 年 4 月至今的收益情况，本文设计的两种投资策略均能达到长期稳定盈利预期效果（见图 7 和图 8），模型可行度较高。在报告中计算出逐年收益率、最大回撤、夏普比率、波动率 4 个指标，分析两位客户的投资行为和模型的实际效果。

本文的亮点在于：首先，衡量基金综合指标评价体系的指标选取较为全面，赋权方式客观，使模型具有普适性；其次，对手续费率进行了稳定性检验、投资偏好系数进行了灵敏度检验；最后，模型的实际结果较为可观，增加了模型的说服力，便于模型的改进和推广。

目录

一、 问题重述	2
二、 问题分析	2
三、 模型假设	2
四、 符号说明	3
五、 模型的建立与求解	3
5.1 评价模型的建立与求解	3
5.2 基于 ARIMA 模型的参数预测	7
5.3 建立基金组合投资策略的多目标规划模型	12
5.4 投资组合报告	15
六、 模型的分析与检验	18
6.1 稳定性分析	18
6.2 灵敏度分析	18
七、 模型的评价、改进	19
7.1 模型的优点	19
7.2 模型的缺点	19
7.3 模型的推广	19
八、 参考文献	19
九、 附录	19

一、问题重述

1.1 问题背景

近年来,随着人们收入水平的不断提升以及金融消费观念的持续更新,人们的理财需求日益增长。基金作为一种适合不同人群的理财产品,深受投资者的喜爱。具有多元化组合投资性,专家理财性、流动性高,选择性强等优点。

虽然相对于股票来说,基金是一种间接的投资品种,但基金投资同样具有风险,因而如何选择基金产品十分关键。一般来说,同时选择多支基金来进行组合投资,可以分散风险。作为银行相关员工来说,一个重要的工作内容就是向客户推荐不同类型基金进行组合,来实现最佳的基金购买方案,从而稳定地获得长期增值。

1.2 问题重述

基于上述背景以及数据信息,我们需要建立数学模型,结合实际交易数据,分别设计出不同风险偏好的基金组合投资,并给出相应的基金组合投资报告,具体解决以下问题:

(1) 利用互联网公开的信息,收集各方面的数据,选择具有投资前景的基金产品。

(2) 针对消费者不同的风险偏好,设计出针对激进型和稳健型投资人的两种基金组合投资策略。

(2) 给出 2019 年 4 月 1 日至 2022 到 5 月 1 日的两种基金组合投资报告(包含对收益性、风险性、综合性指标的阐述)。

二、问题分析

文章的目标是针对激进型和稳健型投资人的两种基金组合投资策略,文章分为三个步骤进行。

第一步,需要缩小基金产品的范围,选出有投资前景的基金若干,有助于后期的投资决策。本文首先计划对基金产品数据进行预处理和分析,删除数据缺失比例超过 20% 的基金产品,使用同期基金产品的平均数据进行缺失值填补,并从基金的信息中确定基金投资前景的描述指标,并用考虑使用熵权法,客观确定权重,最后使用 TOPSIS 进行综合评价。

第二步,通过 ARIMA 模型,预测这 20 个基金产品的日度数据预测下一时期基金产品的预期收益率,进而计算出基金产品的预期季度收益率,由此求出变异系数。

第三步,建立双目标规划模型,考虑用迭代的方法,逐步迭代并更新约束结合蒙特卡洛模拟对各个阶段的双目标规划问题进行求解。

最后,文章利用模型建立求解的结果,给出 2019 年 4 月 1 日至 2022 到 5 月 1 日的两份基金组合投资报告。

三、模型假设

1. 不考虑负债投资;
2. 每个季度进行一次投资操作;
3. 在每个投资期 t 内, $r_i(t), r_0(t)$ 为定值,不受外界因素干扰;
4. 第 t 期的风险用投资基金 i 中最大的一个风险来度量;
5. 在每一期内,对每只基金最多只考虑一次买入或卖出操作;

6. 买入和卖出基金的手续费率相同，为定值。

四、符号说明

符号	说明	单位
n	操作期数	
t	季度数	
i	第 i 支基金	
$\hat{r}_i(t)$	第 i 支基金第 t 期的季度收益率预测值	
α	基金买入手续费率	
β	基金卖出手续费率	
q_i	第 i 支基金的风险度量系数（变异系数）	
$r_0(t)$	第 t 个季度的平均无风险收益率	
$x_i(t)$	第 t 期买第 i 种基金的花费资金	元
$M(t)$	第 t 个季度的资金量	元
s	投资偏好系数	
m	基金总数量	

五、模型的建立与求解

5.1 评价模型的建立与求解

5.1.1 特征变量选取与指标评价体系构建

● 基金历史度量

（1）成立年数

基金的成立年数在一定程度上可以反映基金的运营情况。设基金成立年数用 $year$ 表示，单位为年。实证研究表明，除了 2016 年外，成立时间久的基金均以微弱优势胜出，在 10 年中的 9 年时间里，成立时间久的基金赚钱概率都要比新基金平均高 1.8%。成立年数可以借助网络爬虫获取。

（2）基金规模

基金规模是该基金募集到的资金总额，设基金规模用 $scale$ 表示，单位为万元。投资者在进行投资时，基金规模是考虑因素之一。基金规模不能太小，因为基金规模太小，基金收取的各种费用可能无法覆盖各项成本（如果小于 5000 万元，该基金既可以被来保留，也能被清算和解散）。一般建议选择不低于 10 个亿的基金。基金规模可以借助网络爬虫获取。

● 基金风险度量

（3）最大回撤率

最大回撤率是指在选定周期内任一历史时点往后推，产品净值走到最低点时的收益率回撤幅度的最大值。设最大回撤率为 $drawdown$ 。最大回撤用来描述买入产品后可能出现的最糟糕的情况。最大回撤是一个重要的风险指标，对于对冲基金和数量化策略交易，该指标比波动率还重要。具体计算公式为

$$drawndown = \max \left(\frac{NV_t - NV_j}{NV_i} \right) (j > t, j, t \in N^+) \quad (1)$$

其中, NV_t 表示基金 i 在第 t 期的产品净值, NV_j 表示基金 i 在第 j 期的产品净值, 最大回撤值反应了基金在一段时间内收益率回撤的最大比率。

(4) 变异系数

变异系数又称“标准差率”, 是衡量资料中各观测值变异程度的另一个统计量。当进行两个或多个资料变异程度的比较时, 如果度量单位与平均数相同, 可以直接利用标准差来比较。如果单位和(或)平均数不同时, 比较其变异程度就不能采用标准差, 而需采用标准差与平均数的比值(相对值)来比较。设基金的变异系数为 q_i 。具体的计算公式为

$$q_i = \frac{\sum_{t=1}^n \frac{(r_i(t) - \hat{r}_i(t))^2}{n-1}}{\hat{r}_i(t)} \quad (2)$$

其中, $r_i(t)$ 指第 t 个月的平均无风险收益率, $\hat{r}_i(t)$ 指第 i 支基金第 t 期的日均收益率预测值, n 指操作期数。

● 基金信誉度量

(5) 近三年排名

基金的排名情况, 是衡量基金发展水平的综合性指标, 具有较好的参考意义。选取基金 2016-2018 年的加权平均排名情况, 作为信誉度量标准之一。设基金近三年的加权平均排名情况为 $rank_{ave}$ 。具体计算公式为

$$rank_{ave} = rank_{2016} \times \omega_1 + rank_{2017} \times \omega_2 + rank_{2018} \times \omega_3 \quad (3)$$

其中基金各年的排名可通过网络爬虫获取, ω_1 、 ω_2 、 ω_3 分别取 0.3, 0.3, 0.4。

(6) 五星评级家数

基金的评级汇总分类是天天基金网的分类, 各机构分类详见机构评级。五星评级家数越多, 表示基金的信誉度量越好。设基金的五星评级家数为 $star$, 可借助网络爬虫获取。

● 基金收益度量

(7) 年化收益率

基金年化收益率是指通过购买基金产品可获得的预期收益率换算成年收益率来计算的, 设年化收益率用 R_p 表示。数据可以通过网络爬虫获取。

(8) 夏普比率

夏普比率就是一个可以同时收益与风险加以综合考虑的三大经典指标之一。目的是计算投资组合每承受一单位总风险, 会产生超额报酬。当投资组合内的资产皆为风险性资产时, 适用夏普比率。夏普指数代表投资人每多承担一分风险, 可以拿到几分超额报酬; 若大于 1, 代表基金报酬率高过波动风险; 若为小于 1, 代表基金操作风险大过于报酬率。这样一来, 每个投资组合都可以计算 Sharpe Ratio, 即投资回报与多冒风险的比例, 这个比例越高, 投资组合越佳。设基金的夏普比例为 $Sharperatio$ 。具体计算公式为:

$$Sharperatio = \frac{E(R_p) - R_f}{\sigma_p} \quad (4)$$

其中 $E(R_p)$ 表示投资组合预期年化报酬率, R_f 表示年化无风险利率, σ_p 表示投资组合年化报酬率的标准差。

综上所述, 本文总结了评估基金投资风险 4 个方面 8 个指标, 4 个方面分别为历史度量、风险度量、信誉度量以及收益度量, 8 个指标为成立年数、最大回撤率、基金规

模、变异系数、近三年排名、五星评级家数、年化收益率以及夏普比率。由此可以构建出基金投资风险评估指标体系，图 1 基金投资风险评估指标体系图所示。

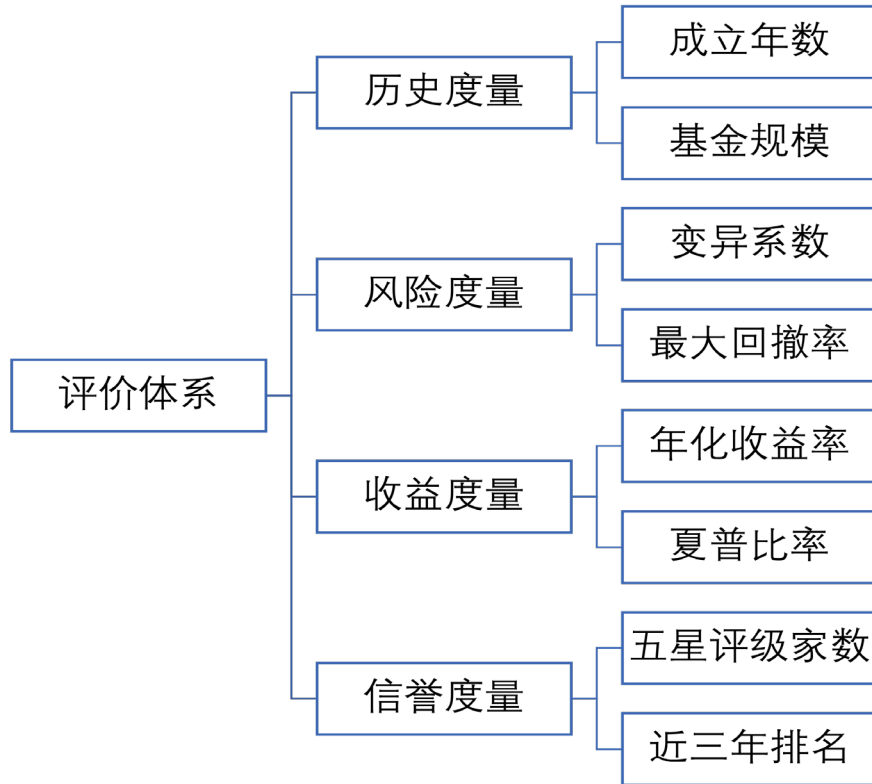


图 1 基金投资风险评估指标体系图

5.1.2 量化基金评价体系

依据上文所计算的各指标数值以及评价体系，文章可以量化出各基金的投资评级。首先需要确定各指标的权重。为了避免主观性对于量化结果的影响，本文采用依据数据驱动的权重计算方法-熵权法来进行赋权，在得到权重之后使用 TOPSIS 来量化每个基金的投资评级。

1) 熵权法确定指标权重

熵权法是一个客观的赋权方法,可以最大程度上避免主观性赋权对于基金评级量化结果的影响[2]。熵权法依据的原理是指标的变异程度，即变异程度越高则对应的权值也就越高。

方法步骤

Step1: 首先本文需要对企业指标数据进行正向化和归一化处理，保证数据的非负性。

$$z_{ij} = \frac{x_{ij} - x_{\min}}{x_{\max} - x_{\min}} \quad (5)$$

其中 z_{ij} 为归一化处理后的变量， x_{\max} 和 x_{\min} 分别为每个指标的最大值和最小值。

Step 2: 计算第 j 个基金评价指标下第 i 个基金所占的权重，将其看作计算信息熵时的概率 p 。

$$p_{ij} = \frac{z_{ij}}{\sum_{i=1}^n z_{ij}} \quad (6)$$

Step 3: 计算第 j 个基金评价指标的信息熵 e_j ，并计算对应信息效用值 d_j ，此处进行转换的原因是因为信息熵越大代表该基金评价指标的信息越少，引用信息效用值 d_j 就可以正向衡量信息量。

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n p_{ij} \ln p_{ij} \quad (7)$$

$$d_j = 1 - e_j \quad (8)$$

Step4: 最终归一化得到每个基金评价指标的熵权 w_j 。

$$w_j = \frac{d_j}{\sum_{j=1}^m d_j} \quad (9)$$

得到的指标权重分别为：

表 1 指标权重

指标	成立年数	基金规模	年化利率	变异系数
权重	0.0224	0.0555	0.1099	0.2535
指标	近三年排名	五星评级家数	夏普比率	最大回撤率
权重	0.0545	0.0752	0.2234	0.2056

2) TOPSIS 方法量化基金评级

TOPSIS 方法是基于数据对样本进行排序的一种方法，其基本思想是根据样本数据构造一个理想化的目标，比如在本例中就是构造一个各方面基金评价指标都达到最优的基金，然后测量实际基金和这个理想化基金的接近程度，越接近就代表其投资风险越低。

方法步骤

Step1: 找出每列也就是每个基金投资指标的最大值，记为 z_i^+ ($i=1,2,\dots,m$)，组成向量

$$Z^+ = \{z_1^+, z_2^+, \dots, z_m^+\} \quad (10)$$

Step 2: 该向量代表理想的基金，即每个正向化后的指标都达到了最小。同样的，找出每列也就是每个指标的最小值，记为 z_i^- ($i = 1, 2, \dots, m$)，组成向量

$$Z^- = \{z_1^-, z_2^-, \dots, z_m^-\} \quad (11)$$

该向量代表最不理想的企业，即每个正向化后的指标都达到了最小。

Step 3: 定义第 i 个企业与理想目标的距离为 D_i^+ ，计算公式为

$$D_i^+ = \sqrt{\sum_{j=1}^m (z_j^+ - z_{ij})^2} \quad (12)$$

Step 4: 定义第 i 个企业与不理想目标的距离为 D_i^- ，计算公式为

$$D_i^- = \sqrt{\sum_{j=1}^m (z_j^- - z_{ij})^2} \quad (13)$$

Step 4: 定义第 i 个企业的得分为 S_i ，计算公式为

$$S_i = \frac{D_i^-}{D_i^+ + D_i^-} \quad (14)$$

显然， S_i 位于 $[0,1]$ 之间。当 S_i 越接近于 1，说明基金 i 距离理想化目标越近，该基金的投资风险就越低。反之，当 S_i 越接近于 0，说明基金 i 距离理想化目标越远，该基金的投资风险就越高。

最终根据投资风险进行排序，筛选出排名前 20 的基金产品，并进行基金组合投资策略的选择。

表 2 排名前二十的基金产品呈现

代码	基金名称	排名	代码	基金名称	排名
000015	华夏纯债债券 A	1	000066	诺安鸿鑫混合 A	11
001606	农银工业 4.0 混合	2	002245	泰康稳健增利债券 A	12
001900	诺安精选价值混合	3	002117	广发安享混合 C	13
002190	农银新能源主题	4	002220	南方瑞利灵活配置混合	14
002145	诺安景鑫灵活配置混合	5	002231	华夏新趋势混合 A	15
002071	长安产业精选混合 C	6	002259	鹏华健康环保混合	16
000431	鹏华品牌传承混合	7	002249	招商境远灵活配置混合	17
000566	华泰柏瑞创新升级混合 A	8	002158	汇添富安鑫智选混合 C	18
007469	中信建投精选混合 C	9	001898	易方达大健康混合	19
001884	中欧互通精选混合 E	10	001928	华夏消费升级灵活配置混合 C	20

5.2 基于 ARIMA 模型的参数预测

(1) ARIMA 模型的建立

ARIMA 模型是博克斯·詹金斯于 70 年代初提出的一种著名的时间序列预测方法。ARIMA (p, d, q) 又称为差分自回归移动平均模型[4]。模型通过 d 阶差分可由非平稳的时间序列得到稳定的时间序列，通过计算能够描述序列特征的统计量来确定 ARMA 模型的阶数。其中，AR 是自回归， p 为自回归项；MA 为移动平均， q 为移动平均项数， d 为时间序列成为平稳时所做的差分次数。

ARIMA 模型有三种基本类型：自回归模型(AR)、移动平均模型(MA)和自回归移动平均模型(ARMA)。下面对这三个类型进行介绍：

(一) 自回归模型(AR)

自回归模型描述当前值与历史值之间的关系，用变量自身的历史时间数据 y_{t-i} 对自身 y_t 进行预测。自回归模型的数据必须满足平稳性的要求，在模型建立时首先需要确定

一个阶数 p ，表示用几期的历史值来预测当前值。 p 阶自回归模型的公式定义为：

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \varepsilon_t \quad (15)$$

其中： y_t 是当前值， μ 是常数项， p 是阶数， γ_i 是自相关系数， ε_t 是误差。

（二）移动平均模型(MA)

移动平均过程作为自回归过程的补充，关注的是自回归模型中的误差项的累加，通常用于解决自回归方差中白噪声的求解问题， q 阶自回归过程的公式定义如下：

$$y_t = \mu + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (16)$$

（三）自回归移动平均模型(ARMA)

自回归模型 AR 和移动平均模型 MA 模型相结合，我们就得到了自回归移动平均模型 ARMA(p, q)， p 是自回归阶数， q 为移动平均阶数，回归方程表示为：

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (17)$$

ARIMA 模型由 ARMA(p, q) 的随机过程经过 d 次差分后可变换成为一个平稳过程，则称其为差分自回归移动平均过程，模型记为 ARIMA (p, d, q)。

由于 ARIMA 模型在经济预测过程中既考虑了经济现象在时间序列上的依存性，又考虑了随机波动的干扰性，对于经济运行短期趋势的预测准确率较高，是近年来应用比较广泛的方法之一，所以本文选用该模型对基金价格和收益率进行分析和预测。

（2）ARIMA 模型的求解

由于选取的基金总数较多，综合考虑天天基金网站的基金定投热销排行及 TOPSIS 的结果，本文选取其中一只基金（易方达蓝筹精选混合，基金代码 005827）为例，具体阐述 ARIMA 模型的求解过程，其余基金的模型求解过程类似。

选取易方达蓝筹精选混合（基金代码 005827）为例，具有一定的代表性和典型性，原因如下：

1.基金经理实力突出：2020 年底，易方达张坤管理的 5 只基金总规模已达到 1255.09 亿元，成为首个公募基金管理规模逾千亿的基金经理。此基金是市场上市场最大的基金，因此在一定程度上代表着中国股票市场的走势。

2.基金经理无变动：通常基金经理的变动意味着投资策略和风格的改变，在短期时间段内会造成基金净值和收益率的巨大变动，因此导致基金数据的不平稳，此时 ARIMA 模型的预测效果变差。

求解步骤

Step1: 数据采集

本文选取的数据的为易方达蓝筹精选混合(005827)自成立至今的日度基金净值数据，时间跨度 2018 年 10 月 8 日至 2022 年 4 月 30 日，将除去节假日的所有每日基金累计净值作为本次研究的样本数据，共有 855 条数据。数据来源于天天基金网。本文将依据这些样本数据对未来数据进行短期预测，以期为投资者提供投资指标，进而做出相应的投资行为。

Step2: 平稳性检验

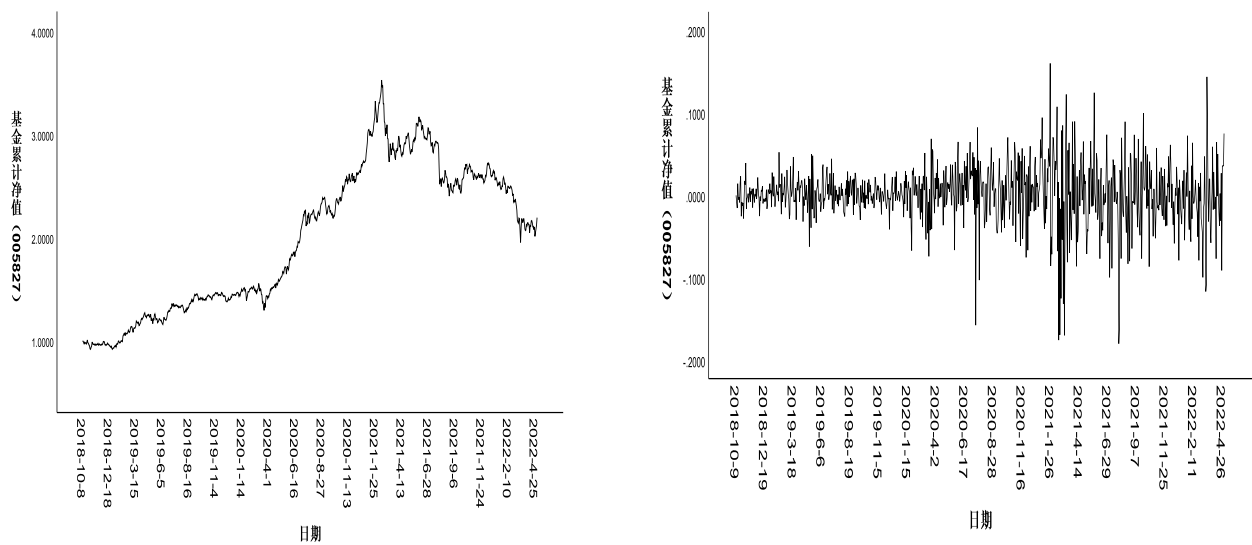
由于股市波动较大，基金是股票和其他投资的组合，因而基金的数据通常也具有较大的波动性，将原始序列（易方达蓝筹精选混合 005827 的日度基金净值）记为 X ，做出原始数据的序列图，可以看出该时间序列不平稳，需要进行平稳化处理。平稳化处理一

般有两种方法，一是取对数，二是进行差分，在本次研究中，选择通过差分进行平稳化处理。如表 3 所示,利用 R 软件对数据进行一阶差分处理(处理后的数据用 d_x 来表示)，并对差分过的数据进行单位根检验，发现 T 统计量的绝对值大于各个置信水平下单位根检验的临界值的绝对值，并且 P 值为 0.000，均小于相应的各个置信水平，因此一阶差分后的数据为平稳时间序列。

表 3 单位根检验结果

变量	检验形式	ADF 统计量	1%临界值	5%临界值	10%临界值	结论
X	(C,0,0)	0.8441	-3.4487	-2.8696	-2.5710	非平稳
	(C,T,0)	-2.7214	-3.9843	-3.4226	-3.1342	非平稳
d_x	(0,0,1)	-18.009*	-2.5716	-1.9417	-1.6160	平稳

一阶差分后，得到序列 DX ，即： $DX = X_t - X_{t-1}$ ， $t = 3,4,\dots,n$ ，做出易方达蓝筹精选混合(005827)的一阶差分序列图，由图 2 可知时间序列数据经过一阶差分后为平稳序列。



原始序列图一阶差分后的序列图
图 2 易方达蓝筹精选混合（005827）时间序列图

（三）模型的构建

综上所述，本文的单序列数据经过一阶差分后得到平稳的时间序列，所以我们选择 $ARIMA(p,d,q)$ 模型对数据进行建模分析。首先需要对模型进行定阶，通过上述的平稳化处理，经过一阶差分后由原始的不平稳的时间序列可以得到平稳时间序列，由此可以确定 $d = 1$ 。观察 ACF 图和 PACF 图（图 3 可知），偏自相关系数和自相关系数都是 4 阶截尾，可以看出初步估计 $p = 2$ ， $q = 2$ 。通过反复测试，根据信息最小化原则，选择 AIC 最小的对模型进行定阶，最终可以得出当 $p = 2$ ， $q = 2$ 时，AIC 的值最小，为 9.557，所以可以建立 $ARIMA(2,1,2)$ 模型，以此模型进行参数估计，得出结果（如表 4 所示）。基于时间序列 $ARIMA(2,1,2)$ 模型对数据进行拟合，通过调整参数，使其通过显著性检验，参数估计结果如表 4 所示各的参数估计 P 值小于 0.01，在 1% 的显著

水平下都是显著的，可以得到模型：

$$DX_t = -1.11471DX_{t-1} - 0.411636DX_{t-2} + \varepsilon_t + 1.206366\varepsilon_{t-1} + 0.588450\varepsilon_{t-2} \quad (18)$$

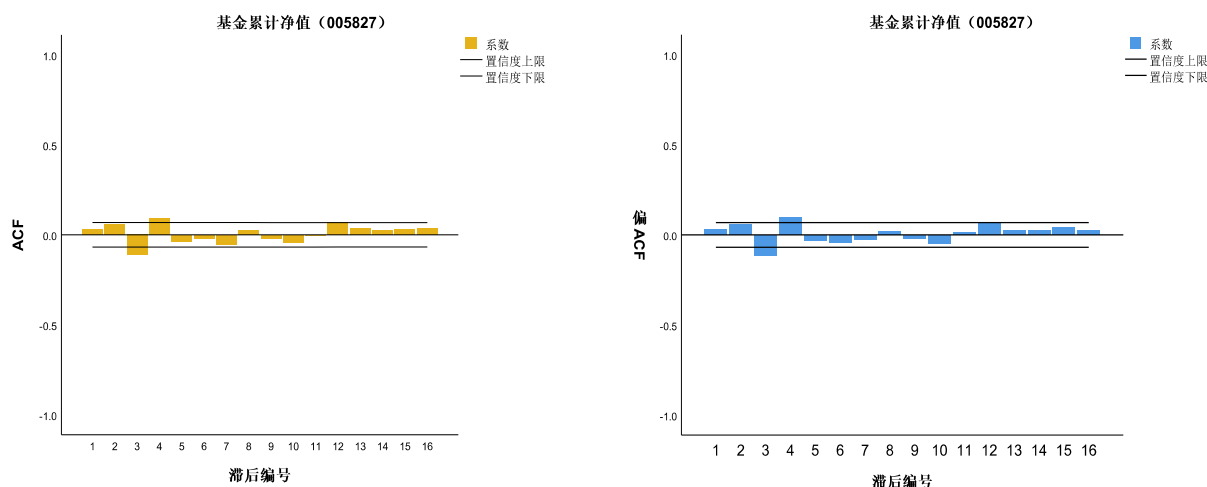


图 3 一阶差分序列相关图

表 4 ARIMA 模型回归结果

变量	系数	标准差	Z 统计量	P 值
AR(1)	-1.11471	0.155783	-7.323232	0.0000
AR(2)	-0.411636	0.152394	-3.349299	0.0000
MA(1)	1.206366	0.142894	7.596455	0.0000
MA(2)	0.588450	0.182233	4.694082	0.0000

(四) 模型的检验和预测

对最终确定的估计模型进行异方差检验，所得结果如表 5 所示。从表 5 中我们可以看出，该模型在置信水平下均通过了检验，即最终确定的模型不存在条件异方差。

表 5 对模型的异方差检验

F 统计量	0.642839	Prob. F(1,387)	0.4410
样本数可决系数	0.642318	Prob. Chi-Square(1)	0.4434

创立模型就是为了对未来某个时点的数据进行预测分析，所以我们采用已经拟合好的 ARIMA(2,1,1)模型来进行数据预测。由于我们所采集的是 2018 年 10 月 8 日至 2022 年 4 月 30 日的数据集，所以基于这些数据，本文将对 2020 年 4 月 25-29 日之间的交易进行模拟预测，得到下面的静态图和预测结果值（见图 4 和表 6）。从图 4 和表 6 所呈现的状况看，拟合的效果较好，与实际值高度吻合，具有较好的预测效果。并且为了进一步验证模型的实际意义，我们将预测值和实际值进行比较，得出误差率小于 10%。从预测结果来看，前两天的误差比不到 1%，精确度非常高，之后两天的误差比也在 5%以内。但随着时间的推移，预测 4 月 29 日的基金净值与真实值之间的误差比达到了 6.2%，说明该模型在短期内能够对基金净值有一个较好的预测，但不适用于长期预测。

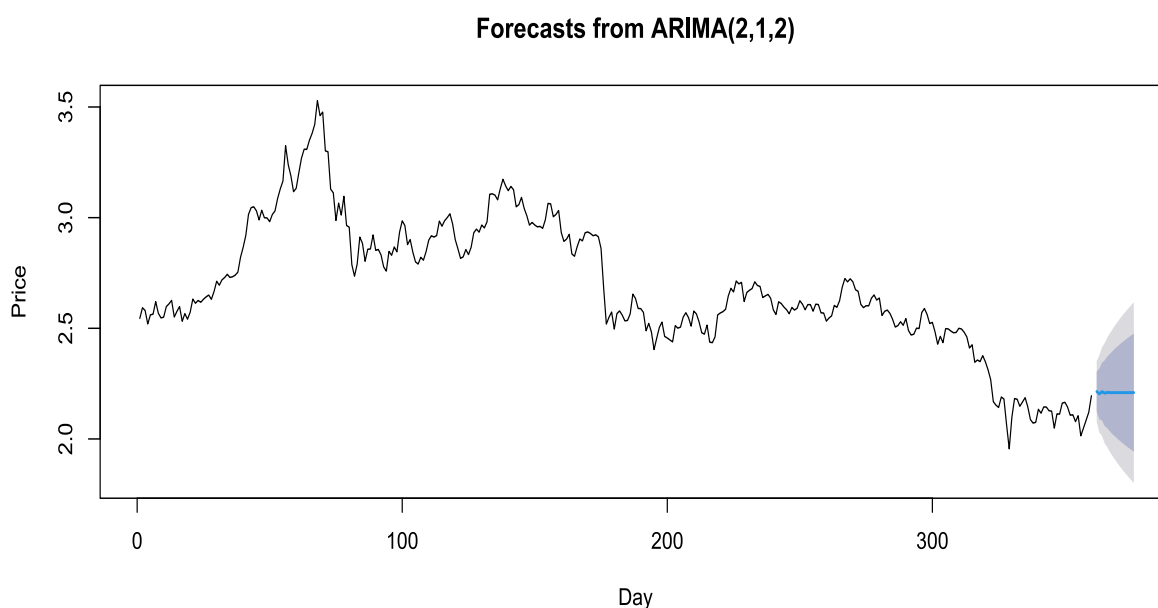


图 4 ARIMA 模型预测结果图

表 6 ARIMA 模型预测结果表

时间	真实值	预测值	误差	误差比
2022/4/25	2.0140	2.023438	0.009438	0.004686
2022/4/26	2.0490	2.034310	-0.01469	-0.071693
2022/4/27	2.0847	2.048383	-0.036317	-0.017420732
2022/4/28	2.1204	2.054283	-0.066117	-0.031181381
2022/4/29	2.1952	2.057921	-0.137279	-0.062535988

基于 ARIMA 模型，本文可以根据基金历史收益率预测下一期的收益率，并通过迭代的方式更新历史数据。由于基金在持有时间不满一个月时赎回费率较高，理性投资者的持有时长为 3-5 个月，因此本文假定投资者一个季度操作一次基金组合的更换和调整，基于 ARIMA 模型可以通过过去时间段的季度收益率，以上文选取的 20 支基金产品 2019 年第二季度之前的季度收益率数据，通过 ARIMA 模型预测出 2019 第二季度的预期季度收益率，如图 5 所示。通过比对 2019 年第二季度的实际收益情况，模型总体预测准确率为 87.3294%，其中预测误差最大为农银新能源主题（002190），预测误差百分比为 -13.496%，预测误差最小为华夏纯债债券 A（000015），预测误差百分比为 2.733%。

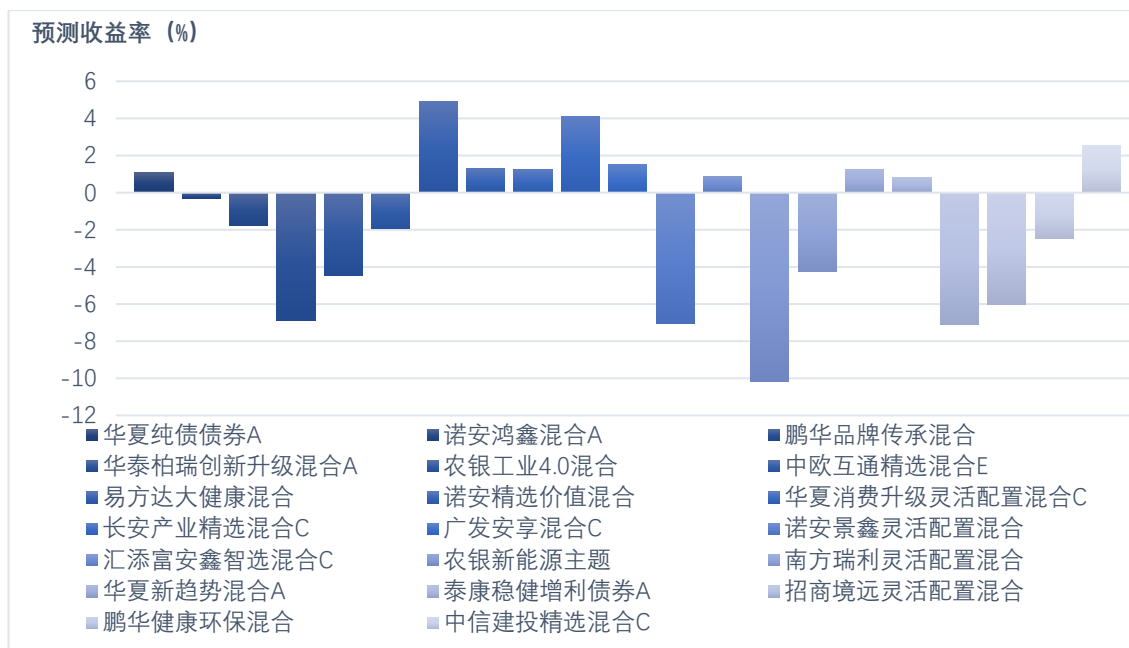


图 5 ARIMA 预测 2019 年第二季度收益率

5.3 建立基金组合投资策略的多目标规划模型

(1) 问题分析:

问题一针对消费者不同的风险偏好，设计出针对激进型和稳健型投资人的两种基金组合投资策略。我们以一系列题目要求为约束条件，找到合适的决策变量，以达到最大化当期预测收益和最小化投资风险的目标，建立多目标规划模型，求得最优解[1]。

(2) 模型简化:

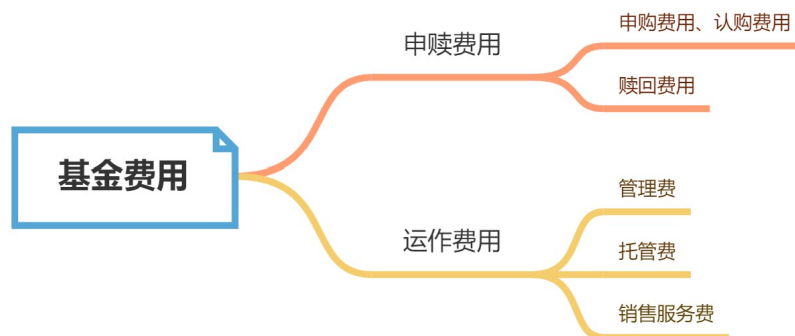


图 6 基金费率明细图

- 基金手续费是指在买或卖基金时支付给销售机构的费用，用来补偿基金销售机构因为办理基金买卖手续而产生的成本。假设用户不会在购买的 7 天内频繁操作，且平台的手续费因为折扣等于原因在 0.15%左右。因此买入和卖出的手续费较小，对于基金的选择无明显干扰，我们假设 $\alpha = \beta = 0.15\%$ ，即买入卖出手续费相同。
- 在实际过程中，考虑到用户不会频繁操作，所以我们以季度为周期进行基金投资操作。

(3) 决策变量:

决策变量是 $m \times n$ 矩阵，表明在第 t 期内买入、买入第 i 支基金所花的钱。表达方式如

下

$$X = (x(1) \quad \cdots \quad x(n)) = \begin{pmatrix} x_1(1) & \cdots & x_1(n) \\ \vdots & \ddots & \vdots \\ x_m(1) & \cdots & x_m(n) \end{pmatrix} \quad (19)$$

其中, m 表示可供选择的基金支数, n 表示总共进行基金买入、卖出操作的总期数, $x_i(t) > 0$ 时表示买入, $x_i(t) < 0$ 时表示卖出, $x_i(t) = 0$ 时表示第 t 期不对第 i 支基金做任何买入、卖出操作。

(4) 目标函数:

要使净收益尽可能大, 总体风险尽可能小, 这是一个多目标规划模型: 最大化当期预测收益和最小化投资风险目标。

- 最大化当期预测收益: $\max\{x_i(t)r_i(t) + \alpha \sum_{i=1}^m |x_i(t)|\}$
- 最小化投资风险: $\min\{\max\{q_i x_i(t)\}\}$

而在实际投资中, 投资者在权衡资产风险和预期收益两方面时, 希望选择一个令自己满意的投资组合。因此对风险、收益分别赋予权重 $s(0 \leq s \leq 1)$ 和 $(1-s)$, s 称为投资偏好系数[3]。对于激进投资者, s 的数值趋向 0; 对于稳健投资者, s 的数值趋向 1。因此改进后的目标函数为:

$$\min s \{ \max\{q_i x_i(t)\} \} - (1-s) \left(\sum_{i=0}^m x_i(t)r_i(t) + \alpha \sum_{i=1}^m |x_i(t)| \right) \quad (20)$$

(5) 约束条件:

$$\text{s.t.} \begin{cases} \sum_{i=0}^m x_i(t) + \alpha \sum_{i=1}^m |x_i(t)| = 0 \\ \sum_{i=1}^m (1+\alpha) I_{\{x_i(t)>0\}}(x_i(t))(x_i(t)) - M(t-1) \leq 0 \\ I_{\{x_i(t)>0\}}(x_i(t)) = \begin{cases} 1, & x_i(t) > 0 \\ 0, & \text{其他} \end{cases} \\ -\sum_{i=1}^k x_i(t) \leq 0 (\forall i \in [1, m], \forall k \in [1, n]) \\ M(t) = M(t-1) + x_0(t-1) \\ M(0) = 1000000 \end{cases} \quad (21)$$

约束条件说明:

- 即当期资金的买卖费用与现金持有变化量保持平衡;
- 由假设 1 可知, 不考虑负债投资, 买入的基金不能大于当前持有的资金量;
- $I_{\{x_i(t)>0\}}(x_i(t))$ 作为分段函数, 控制对买入的基金的筛选;
- 基金不能持负;
- 资金呈现递归形式, t 期的资金为 $(t-1)$ 期的初始资金和资金变化量的加和;
- 初始资金为 100 万元, 全部用来投资基金组合。

(6) 算法及模型的求解

算法的核心思想

由于需要进行多期投资组合决策, 我们考虑是用迭代的方法, 逐步迭代并更新约束结合蒙特卡洛模拟对各个阶段的双目标规划问题进行求解。由于生活中有许多事件发生的结果都有无限种可能, 因此我们不可能枚举出所有的结果, 这时候就只能通过蒙

特卡罗模拟，将一个不确定性的问题转化成很多个确定性问题，并得到一个近似解，因此蒙特卡罗算法也可以看成是枚举法的一种变异。

方法步骤

STEP 1: 设置包括预估收益率、实际收益率、当前持有现金等各变量初始值。

STEP 2: 每次从约束条件中求出每个变量受限的大致范围。

STEP 3: 在上述范围中用随机数生成 20000000 组试验点，并验证它们是否满足所有的约束条件，若满足，则将其划分到可行组，再从可行组中找到函数的最值。作为当期投资组合。

STEP 4: 迭代更新当前持有现金及各基金价格，重复进行 **STEP1-STEP3** 直至做完所有期数。

求解结果:

数据太多，选取部分进行展示，对于激进型和稳健型，分别展示进行投资至今，每期的预期净收益、实际净收益和当前总资产，如图 7，8 所示。



图 7 激进型投资预测

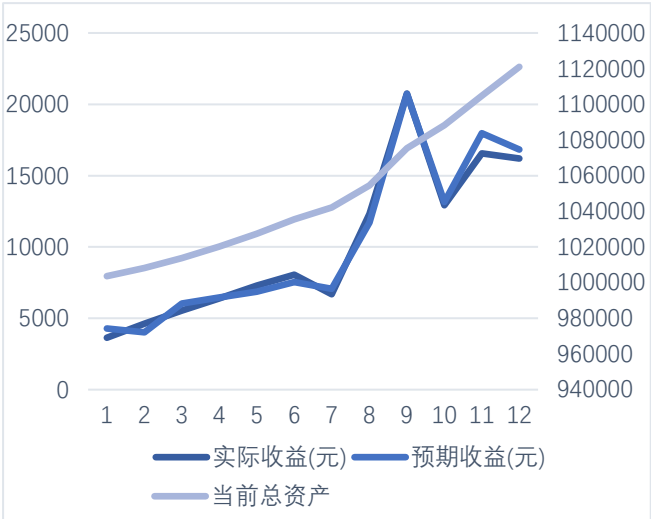


图 8 稳健型投资预测

从图中可以看出，实际收益与预期收益较为接近，说明 ARIMA 平均收益率预测较为准确，同时净收益呈现波动式上升。经过 12 期基金组合投资至今，激进型投资策略可达到总资产 **2085426 元**，稳健型投资策略可达到总资产 **1121002.866 元**。下面我们展示第 1 期激进型投资组合策略进行的不同基金买卖操作，如表 7 所示。

表 7 激进型投资组合策略

基金	操作	基金	操作
华夏纯债债券 A	17295.3206803167	广发安享混合 C	61669.8409005016
诺安鸿鑫混合 A	19456.4215396427	诺安景鑫灵活配置混合	69417.9982060002
鹏华品牌传承混合	20177.2347895326	汇添富安鑫智选混合 C	82428.0649894002

华泰柏瑞创新升级混合 A	80818.513212007	农银新能源主题	66638.9015005683
农银工业 4.0 混合	43684.0282105515	南方瑞利灵活配置混合	77822.9512120864
中欧互通精选混合 E	28941.3927723508	华夏新趋势混合 A	71788.4616279914
易方达大健康混合	52045.6992484042	泰康稳健增利债券 A	47699.3705314351
诺安精选价值混合	83264.6409098615	招商境远灵活配置混合	82393.6151985893
华夏消费升级灵活配置混合 C	7693.07570993312	鹏华健康环保混合	21651.0935362666
长安产业精选混合 C	3713.05448788129	中信建投精选混合 C	59571.0357949168

通过第 1 期对该 20 支进行买入操作，剩余现金进行存款操作。第 1 季度结束后，实际可以达到 **29945.122** 元的净收益。第 1 期稳健型投资组合策略进行的不同基金买卖操作，如表 8 所示。

表 8 稳健型投资组合策略

基金	操作	基金	操作
华夏纯债债券 A	16805.0868697699	广发安享混合 C	79923.3518375279
诺安鸿鑫混合 A	53351.8261229034	诺安景鑫灵活配置混合	31799.6610230164
鹏华品牌传承混合	21048.4574283686	汇添富安鑫智选混合 C	9213.3194400631
华泰柏瑞创新升级混合 A	5934.5476497742	农银新能源主题	28327.1910112938
农银工业 4.0 混合	17909.1595296958	南方瑞利灵活配置混合	5522.57594835178
中欧互通精选混合 E	11800.426761429	华夏新趋势混合 A	28624.3553376162
易方达大健康混合	15048.9187769295	泰康稳健增利债券 A	66734.4669760189
诺安精选价值混合	14179.6288419487	招商境远灵活配置混合	2624.94523905466
华夏消费升级灵活配置混合 C	3274.47023492655	鹏华健康环保混合	9268.26210800002
长安产业精选混合 C	16199.6657610426	中信建投精选混合 C	47562.951072409

通过第 1 期对该 20 支进行买入操作，剩余现金进行存款操作。第 1 季度结束后，实际可以达到 **3638.6495** 元的净收益。

对比激进型投资组合与稳健型投资组合，可以看出，激进型投资组合购买基金的金
额明显大于稳健型投资，结果符合实际。

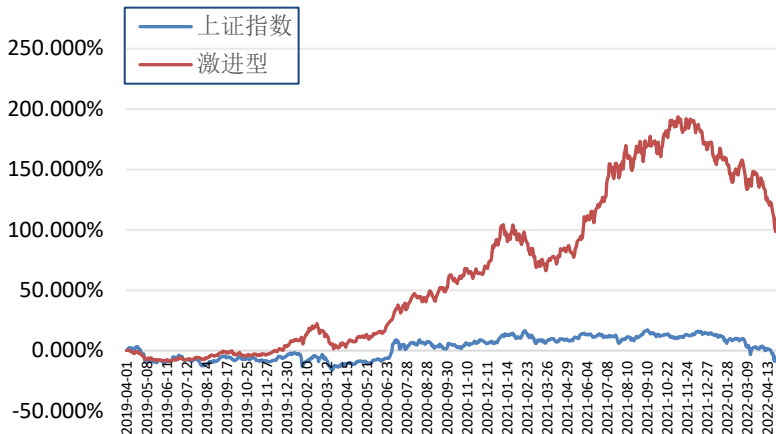
5.4 投资组合报告



“激进投资，收益超预期”—— 针对激进型投资者的投资报告

投资概况分析

2019年4月至2022年4月激进型投资组合收益率



数据来源：模型结果

投资业绩分析

27.85%

近三年年化收益率

收益较高

38.91%

近三年最大回撤率

回撤较高

34.28%

近三年波动率

波动较高

0.61

近三年夏普比率

性价比一般

投资改进建议



市场上行时
考虑保持策略



市场下行时
考虑更换策略

■ 总体收益极高，远超上证指数

该基金组合投资策略适合激进型投资者持有，该策略执行于2019年4月至2022年4月，获得总收益108.54%，平均年化收益为27.85%，期间大盘指数总收益为-3.89%，可见基金的选取贴合市场热点，能够为投资者获取超额收益。

■ 资产净值波动率大，风险较高

该基金组合投资策略风险较高，在2021年第四季度至今，市场风格切换，该组合投资策略回撤较大。说明市场趋势的变化会导致风险的剧烈增加。

对投资策略的业绩评判可从收益、风险、风险调整后收益三个维度进行评估：

■ 该组合投资策略收益很高

偏股型基金和混合型基金在该组合投资策略中占比较大，增加市场上行时的获利机会。

■ 该组合投资策略风险很高

该组合投资策略的近三年最大回撤率为38.91%，资金波动幅度很大，较为激进。

■ 该基金组合投资策略性价比一般

由夏普比率可知，该策略性价比一般。

■ 市场上行时考虑保持策略

激进型的投资者往往能在市场上行时获取超额收益。因此，在市场上行时，激进型投资者可积极追随市场趋势，保持高风险高收益的投资策略。

■ 市场下行时考虑更换策略

在市场下行时，该投资策略通常会加大资金净值的回撤，因此市场下行时激进型投资者需要更换风险系数更小的组合投资策略。



“稳健投资，收益如期”—— 针对稳健型投资者的投资报告

投资概况分析

2019年4月至2022年4月稳健型投资组合收益率



投资业绩分析

3.90%

近三年年化收益率

收益一般

17.47%

近三年最大回撤率

回撤较小

26.75%

近三年波动率

波动尚可

0.78

性价比高

近三年夏普比率

投资改进建议



市场上行时
考虑更换策略



市场下行时
考虑保持策略

■ 总体收益良好，持续跑赢指数

该基金组合投资策略适合稳健型投资者持有，该策略执行于2019年4月至2022年4月，获得总收益12.10%，平均年化收益为3.90%，期间大盘指数总收益为-3.89%，可见基金选取较为合理，能够抵抗通胀，降低投资风险，实现长期稳定盈利。

■ 资产净值波动率小，风险较低

该基金组合投资策略能够较大程度地降低市场风险，2020年第二季度全球遭遇新冠危机，经济普遍下行，此投资组合仍能获得持续的正向收益，说明该策略风险抵抗性较强。

对投资策略的业绩评判可从收益、风险、风险调整后收益三个维度进行评估：

■ 该组合投资策略收益一般

债券型基金和货币型基金在该组合投资策略中占比较大，减少市场上行时的获利机会。

■ 该组合投资策略风险较小

该组合投资策略的近三年最大回撤率为17.47%，资金波动幅度较小，较为稳健。

■ 该基金组合投资策略性价比高

由夏普比率可知，该策略性价比较高。

■ 市场上行时考虑更换策略

在市场上行过程中，一些基金能够迎合市场趋势，获取超额收益，稳健型投资者在市场上行时应积极关注偏股型和混合型基金的业绩情况并调整投资策略。

■ 市场下行时考虑保持策略

在市场下行时，该投资策略能较好的抵御市场风险并获得预期收益，因此稳健型投资者在市场下行时可保持原策略。

六、模型的分析与检验

6.1 稳定性分析

双目标规划模型的稳定性分析

在实际问题中，手续费率 α 通常不是一成不变的，受此影响，最后的总收益与各策略会存在偏差。本小节将在数据中加入均值为 0，方差为 0.0001 的正态噪声，然后用双目标规划模型，重复多次求解，对比求得的最总收益是否存在明显偏差。这里我们对激进型、稳健型基金组合投资双目标规划模型重复求解 10 次，绘制求解结果：

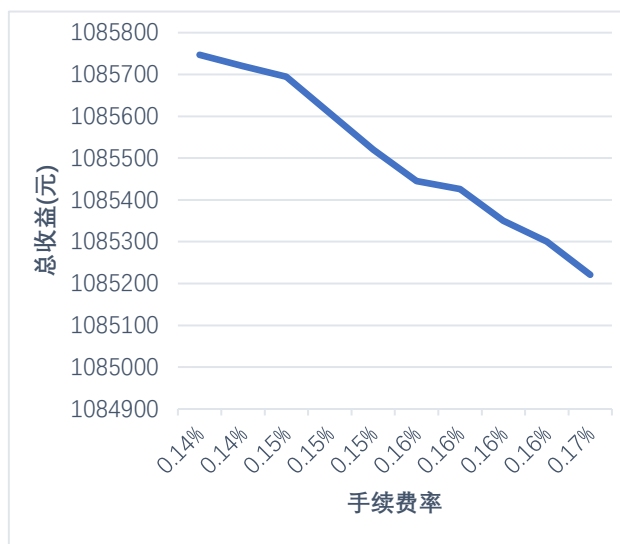


图 9 激进型投资策略稳定性分析

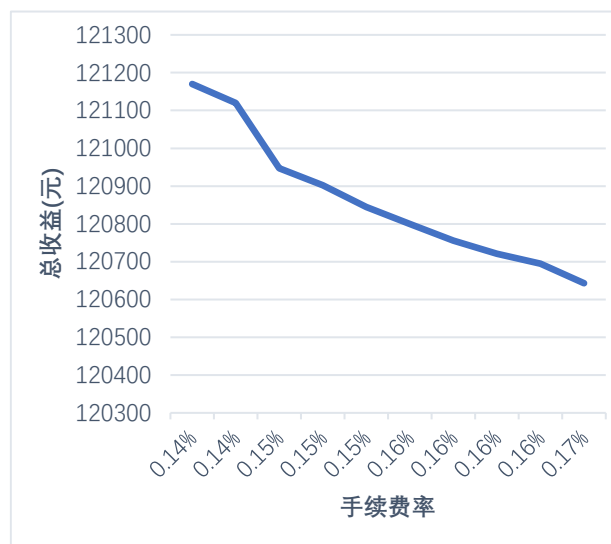


图 10 稳健型投资策略稳定性分析

从图 9 中可以看出，激进型手续费率在 0.14%—0.17%波动时，总收益在 108500 元—1085200 元间波动，波动范围相比总收益基数不大；图 10 中，稳健型手续费率在 0.14%—0.17%波动时，总收益在 121200—120600 间波动，波动范围相比总收益基数不大。这些波动在实际中是可接受的，说明双目标规划模型对手续费率波动稳定性较高。

6.2 灵敏度分析

关于投资偏好系数 s 的灵敏度

在最小化风险最大化预期收益的问题中存在可供选择的系数变量 s 。对于激进型和稳健型，分别取 $s = 0, s = 1$ 。实际中，可根据个人风险偏好，对 s 进行选择， s 越大，越偏好风险； s 越小，越趋于稳健投资，令 s 在 0 到 1 内变化，取 10 个不同的 s 系数，绘制总收益关于 s 系数的之间变化的示意图：

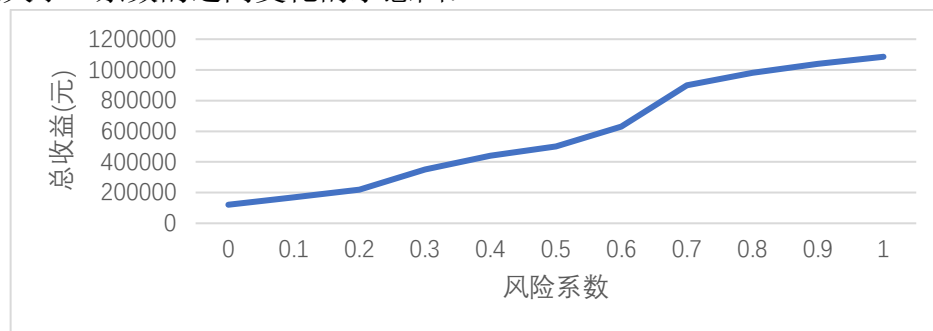


图 11 灵敏度分析

图 11 中，随着风险系数的增大，总收益逐渐增大，风险增大到 0.7 左右总收益增长趋于平缓。从中可以看出，总收益关于风险系数的波动较为敏感，实际进行基金组合投资时，应特别注意这一点。

七、模型的评价、改进

7.1 模型的优点

1. 在建立基金评价模型时，通过充分的论证，选取合理、有效、可获得的指标，衡量的指标多，设计方面广，可以全面衡量基金的优劣。此外，熵权法赋值比较客观，可以规避主观性对结果的影响。该模型的应用有效缩小了问题的规模，便于后续模型方法的实际推广与应用；

2. 在对基金收益率、风险系数的预测时，我们采用了迭代的 ARIMA 模型，用历史数据估计当期收益率，进而结合双目标规划可以做出买入、卖出、持平操作的决策，决策过程符合客观规律；

3. 在进行蒙特卡洛求解时，我们合理设置了投点空间，优化了算法，缩小了求解空间，同时分散了基金购买组合，有效规避了风险，算法效率与精确度显著提升。

7.2 模型的缺点

1. 为便于模型的建立与求解，我们假设了买入、卖出手续费率一致，并且不考虑手续费率随时间发生变化。而在实际购买基金时，卖出手续费会随着持有基金时间及持有基金的多少而发生变化，将该过程过于理想化可能会对结果带来一定的误差；

2. 在对复杂的双目标规划模型进行求解时，采用了蒙特卡洛算法，其精度依赖于随机投点的次数，以及约束解空间的限制。

7.3 模型的推广

本文使用双目标非线性规划等模型，从收益与风险角度对基金组合投资问题进行规划与决策，其中对单支基金收益率预测结合多阶段双目标规划求解问题的思路可以推广到更多的投资问题中，有深远的实用价值，为考虑问题提供了一个新的思路。

八、参考文献

- [1]孙卫军,金芝.证券投资风险的分析及防范[J].中国集体经济,2022(01):99-100.
- [2]司守奎.数学建模算法与应用[M].国防工业出版社,2011.8.
- [3]童文兵.风险投资组合的线性规划模型的建立及求解[J].统计与决策,2016(09):89-91.
- [4]杨宇堦,张梅.基于 ARIMA 模型的股票价格实证分析[J].科技资讯,2021,19(29):121-123+127.

九、附录

#数据爬取

```
1. import requests
2. import re
```

```

3. import csv
4. headers = {
5. 'Cookie': 'AUTH_FUND.EASTMONEY.COM_GSJZ=AUTH*TTJJ*TOKEN; em_hq_fls
    =js; em-quote-
    version=topspeed; intellpositionL=1010.67px; intellpositionT=1056.
    33px; qgqp_b_id=9bb5e67a98ecd30f313ca5710463a0f1; HAList=a-sz-
    000615-%u5965%u56ED%u7F8E%u8C37%2Ca-sh-
    600121-%u90D1%u5DDE%u7164%u7535%2Ca-sh-
    603085-%u5929%u6210%u81EA%u63A7%2Ca-sh-
    600585-%u6D77%u87BA%u6C34%u6CE5%2Ca-sz-
    300450-%u5148%u5BFC%u667A%u80FD%2Ca-sz-
    002340-%u683C%u6797%u7F8E%2Ca-sz-
    300014-%u4EBF%u7EAC%u9502%u80FD%2Ca-sh-
    600030-%u4E2D%u4FE1%u8BC1%u5238%2Ca-sh-
    600436-%u7247%u4ED4%u7640%2Ca-sz-
    300750-%u5B81%u5FB7%u65F6%u4EE3%2Ca-sz-
    300059-%u4E1C%u65B9%u8D22%u5BCC; st_si=65095510686414; st_asi=dele
    te; ASP.NET_SessionId=kihuw0mtd0qjlytvzd2vyj4k; EMFUND1=null; EMFU
    ND2=null; EMFUND3=null; EMFUND4=null; EMFUND5=null; EMFUND6=null;
    EMFUND7=null; EMFUND8=null; EMFUND0=null; EMFUND9=04-
    30 13:55:06@#%$%u534E%u590F%u6052%u751F%u4E92%u8054%u7F51%u79D1%u62
    80%u4E1AETF%u8054%u63A5%28QDII%29A@%23%24013171; st_pvi=5049306067
    6704; st_sp=2020-12-
    31%2022%3A09%3A53; st_inirUrl=https%3A%2F%2Fwww.baidu.com%2Fs; st_
    sn=23; st_psi=20220430171920947-112200312936-9337642130',
6. 'Host': 'fund.eastmoney.com',
7. 'Referer': 'http://fund.eastmoney.com/data/fundranking.html',
8. 'User-
    Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537
    .36 (KHTML, like Gecko) Chrome/97.0.4692.99 Safari/537.36 Edg/97.0
    .1072.76',
9. }
10. with open('股票.csv',mode='a',encoding='utf-8',newline='') as f:
11. f.write('基金代码,基金简称,基金英文,日期,单位净值,累计净值,日增长率,近
    一周,近一个月,近三个月,近六个月,近一年,近两年,近三年来,近年来,成立来,上
    市日期,null,自定义,null,手续费,null,null')
12. f.write('\n')
13. for page in range(1,216):
14. url='http://fund.eastmoney.com/data/rankhandler.aspx?op=ph&dt=kf&f
    t=all&rs=&gs=0&sc=6yzf&st=desc&sd=2021-01-01&ed=2021-03-
    31&qdii=&tabSubtype=,,,,,&pi=1&pn=50&dx=1&v=0.13460597403073082'
15. response=requests.get(url=url,headers=headers)
16. data=response.text

```



```

17.data_str=re.findall('\[(.*?)\]',data)[0]
18.tuple_data=eval(data_str)
19.
20.for t in tuple_data:
21.t_list=t.split(",")
22.with open('股票.csv',mode='a',encoding='utf-8',newline='') as f:
23.csv_write=csv.writer(f)
24.csv_write.writerow(t_list)

```

#TOPSIS 模型

```

1. %% 第一步：把数据复制到工作区，并将这个矩阵命名为 X
2. clear;clc
3. load TOPSIS2.mat
4. %% 第二步：判断是否需要正向化
5. [n,m] = size(X);
6. disp(['共有' num2str(n) '个评价对象，' num2str(m) '个评价指标'])
7. Judge = input(['这' num2str(m) '个指标是否需要经过正向化处理，需要请输入 1，不需要输入 0: ']);
8.
9. if Judge == 1
10.     Position = input('请输入需要正向化处理的指标所在的列，例如第 2、3、6 三列需要处理，那么你需要输入[2,3,6]: '); %[2,3,4]
11.     disp('请输入需要处理的这些列的指标类型（1：极小型， 2：中间型， 3：区间型） ')
12.     Type = input('例如：第 2 列是极小型，第 3 列是区间型，第 6 列是中间型，就输入[1,3,2]: '); %[2,1,3]
13.     % 注意，Position 和 Type 是两个同维度的行向量
14.     for i = 1 : size(Position,2) %这里需要对这些列分别处理，因此我们需要知道一共要处理的次数，即循环的次数
15.         X(:,Position(i)) = Positivization(X(:,Position(i)),Type(i),Position(i));
16.         % Positivization 是我们自己定义的函数，其作用是进行正向化，其一共接收三个参数
17.         % 第一个参数是要正向化处理的那一列向量 X(:,Position(i)) 回顾上一讲的知识，X(:,n)表示取第 n 列的全部元素
18.         % 第二个参数是对应的这一列的指标类型（1：极小型， 2：中间型， 3：区间型）
19.         % 第三个参数是告诉函数我们正在处理的是原始矩阵中的哪一列
20.         % 该函数有一个返回值，它返回正向化之后的指标，我们可以将其直接赋值给我们原始要处理的那一列向量
21.     end
22.     disp('正向化后的矩阵 X = ')
23.     disp(X)

```

```

24.end
25.%% 第三步：对正向化后的矩阵进行标准化
26.Z = X ./ repmat(sum(X.*X) .^ 0.5, n, 1);
27.disp('标准化矩阵 Z = ')
28.disp(Z)
29.%% 让用户判断是否需要增加权重
30.disp("请输入是否需要增加权重向量，需要输入 1，不需要输入 0")
31.Judge = input('请输入是否需要增加权重： ');
32.if Judge == 1
33.    Judge = input('使用熵权法确定权重请输入 1，否则输入 0： ');
34.    if Judge == 1
35.        if sum(sum(Z<0)) >0 % 如果之前标准化后的 Z 矩阵中存在负数，
            则重新对 X 进行标准化
36.            disp('原来标准化得到的 Z 矩阵中存在负数，所以需要对 X 重新标
                准化')
37.            for i = 1:n
38.                for j = 1:m
39.                    Z(i,j) = [X(i,j) - min(X(:,j))]/ [max(X(:,j))
                        - min(X(:,j))];
40.                end
41.            end
42.            disp('X 重新进行标准化得到的标准化矩阵 Z 为： ')
43.            disp(Z)
44.        end
45.        weight = Entropy_Method(Z);
46.        disp('熵权法确定的权重为： ')
47.        disp(weight)
48.    else
49.        disp(['如果你有 3 个指标，你就需要输入 3 个权重，例如它们分别为
            0.25,0.25,0.5，则你需要输入[0.25,0.25,0.5]']);
50.        weight = input(['你需要输入' num2str(m) '个权数。' '请以行向
            量的形式输入这' num2str(m) '个权重： ']);
51.        OK = 0; % 用来判断用户的输入格式是否正确
52.        while OK == 0
53.            if abs(sum(weight) -
                1)<0.000001 && size(weight,1) == 1 && size(weight,2) == m % 注意，
                Matlab 中浮点数的比较要小心
54.                OK =1;
55.            else
56.                weight = input('你输入的有误，请重新输入权重行向
                    量： ');
57.            end
58.        end

```

```

59.     end
60. else
61.     weight = ones(1,m) ./ m ; %如果不需要加权重就默认权重都相同，即都
        为 1/m
62. end
63.%% 第四步：计算与最大值的距离和最小值的距离，并算出得分
64.D_P = sum([(Z - repmat(max(Z),n,1)) .^ 2] .* repmat(weight,n,1) ,
        2) .^ 0.5;    % D+ 与最大值的距离向量
65.D_N = sum([(Z - repmat(min(Z),n,1)) .^ 2] .* repmat(weight,n,1) ,
        2) .^ 0.5;    % D- 与最小值的距离向量
66.S = D_N ./ (D_P+D_N);    % 未归一化的得分
67.disp('最后的得分为: ')
68.stand_S = S / sum(S)
69.[sorted_S,index] = sort(stand_S , 'descend')

```

#ARIMA 预测模型

```

1. library(ggplot)
2. (tseries)
3. library(report)
4. getwd()
5. setwd(dir="C:/Users/USER/Desktop/招银杯")
6. data<-read.csv(file="1234.csv",header=TRUE,encoding = "UTF-8")
7. str(data)
8. d1=as.character(data[,1])
9. d2=as.Date(d1,format="%Y%m%d")
10. library(zoo)
11. library(forecast)
12. dt.zoo<-zoo(data$X.U.00BC..U.06F8..f1.,d2)
13. dt.zoo1<-dt.zoo[496:856,]
14. dt.auto<-auto.arima(dt.zoo1)
15. dt.fore<-forecast(dt.auto,15)
16. plot(dt.fore,xlab = "Day",ylab="Price")
17. accuracy(dt.auto)
18. report(dt.auto)

```

#双目标规划模型

```

1. %风险偏好系数
2. s=0;
3. %基金数 m 支
4. m=20;
5. %手续费率
6. alpha=0.0015;
7. %初始现金

```



```

8. M(1)=1000000;
9. n=20000000;
10. Q=unifrnd(0,1000000/12,n,m);
11. sumX=zeros(1,m);
12. factR=factR./100;
13. R=factR+(unifrnd(-0.00001,0.0005,m,12));
14. %银行利率
15. R0=R0./100;
16. %循环变量 t, 表示第 t 季度, 12 (期数)
17. for t=1:12
18. r=R(1:m,t);
19. factr=factR(1:m,t);
20. r0=R0(t);
21. % 初始化函数 f 的最大值为负无穷
22. fmax=-inf;
23. %开始对 n 组随机数组进行试验, 比较大小
24. for i=1:n
25.     x = Q(i,1:m); %构造 x 向量,取第 i 行为本次模拟的值
26.     x0=-sum(x)-alpha*sum(abs(x));%做了投资之后有的现金
27.     I=zeros(1,m);
28.     for j=1:m
29.         if(x(j)>0)
30.             I(j)=1;
31.         else
32.             I(j)=0;
33.         end
34.     end
35.     if ((1+alpha)*sum(I.*x)<=M(t)) % 判断是否满足条件
36.         result = s*((-
            sum(x)+alpha*sum(abs(x)))*r0+x*r+sum(abs(x))*alpha)-(1-
            s)*max(q.*x);
37.         factresult=(-
            sum(x)+alpha*sum(abs(x)))*r0+x*factr+sum(abs(x))*alpha;
38.         if (result > fmax) % 如果这个函数值大于我们之前计算出来的
            最大值
39.             fmax = result; % 那么就更新这个函数值为新的最大值
40.             X = x;
41.             factfmax=factresult;
42.             yuqishouyi=(-
                sum(x)+alpha*sum(abs(x)))*r0+x*r+sum(abs(x))*alpha;
43.             M(t+1)=M(t)+x0;
44.         end
45.     end

```

```

46. end
47. sumX=sumX+X;
48. for j=1:m
49. Q(1:n,j)=unifrnd(-sumX(j),M(t+1)/12,n,1);
50. end
51. disp(strcat('第',num2str(t),'次实际收益为',num2str(factfmax)))
52. disp(strcat('第',num2str(t),'次各基金投资变化量为'))
53. disp(X)
54. disp(strcat('第',num2str(t),'次预期最大函数值为',num2str(fmax)))
55. disp(strcat('第',num2str(t),'次预期收益
    为',num2str(yuqishouyi)))
56. end
57. %最后各基金持有数量
58. disp(sumX)

```