

# 基于遗传算法优化股票投资组合研究

刘梓彤

青岛市崂山区培诺教育培训学校

DOI:10.32629/acair.v3i4.17938

**[摘要]** 本文聚焦股票投资组合优化问题展开相关研究,目的是在风险跟收益之间谋求平衡。本文采用智能优化手段——遗传算法,经由模拟自然选择和遗传进化的流程,对股票投资组合做优化,文章构建出基于历史数据的投资组合模型,采用遗传算法所涉及的编码、适应度函数、选择、交叉与变异操作,对投资的占比进行优化,实验部分选定若干股票数据开展实证研究。结果显示:遗传算法经多次迭代可收敛到较优的解,显著增强投资组合的收益-风险水平,本文把研究成果做了总结,点明了方法存在的不足以及未来的改进方向。

**[关键词]** 股票投资; 数字金融; 遗传算法

**中图分类号:** F830.59 **文献标识码:** A

## Research on Optimizing Stock Investment Portfolio Based on Genetic Algorithm

Zitong Liu

Painuo Education and Training School,Laoshan District,Qingdao City

**[Abstract]** This article focuses on the optimization problem of stock investment portfolios,with the aim of seeking a balance between risk and return.The intelligent optimization method – genetic algorithm – is used to optimize the stock investment portfolio by simulating the process of natural selection and genetic evolution.The article constructs an investment portfolio model based on historical data,and uses the encoding,fitness function,selection,crossover and mutation operations involved in genetic algorithm to optimize the proportion of investment.In the experimental part,several stock data are selected for empirical research.The results show that genetic algorithm can converge to a better solution after multiple iterations,significantly enhancing the return risk level of the investment portfolio.This article summarizes the research results,points out the shortcomings of the method and future improvement directions.

**[Key words]** Keywords; stock investment; digital finance; genetic algorithm

伴随着全球经济的不断发展以及人民生活水平上扬,投资理财逐渐成为现代社会里越来越重要的组成部分<sup>[1]</sup>,作为极具吸引力的投资领域之一,招来大量散户以及机构投资者参与,股票投资的核心诉求是实现资产的保值与增长,怎样从众多金融产品里选出合适的投资组合,成了投资者面对的重大考验。股票市场凸显的复杂情形,尤其是股票彼此的高关联性与波动特性,造成传统的投资分析方法在实际应用中显得捉襟见肘,对股票之间相关性开展研究能揭示市场潜在规律,并且为投资者提供关键的决策支持<sup>[2]</sup>,有关股票市场个股彼此相互关系的研究,虽说在一些领域已积累诸多成果,但大多数研究把焦点放在地域或行业层面的联动性分析,而对个股间的相似性与时间延迟效应关注不多,凭借对股票之间局部相似性的分析,可找出股票相互之间的潜在关联,进而为打造更科学合理的投资策略给予支持,探究股票市场里个股间的相互联系,探寻适合股票投资的

优化办法,具有重大的理论与实践意义。

在股票市场投资组合优化的研究中,近年来越来越多的研究者采用了不同的分析方法与算法来探索股票市场中的规律,提出了多种基于不同理论和方法的投资策略。

林星<sup>[3]</sup>以局部相似分析方法进行研究,研究了A股市场中各个股之间的相似性及时间延迟效应,他对2013至2016年深圳和上海交易所90支股票的收盘价进行了一番分析,发现个股彼此间存在显著的一同涨跌现象,且提出了基于局部相似情况分析的投资策略。实证分析的结果说明,2016年,深市样本的策略收益率攀升至12.78%,显著胜过同期深证成指12.46%的跌幅表现。

许文伟<sup>[4]</sup>把模糊数学和投资者行为联合起来,研发出一种基于模糊收益及波动现象的投资策略。此策略采用模糊推理机制,引入波动率、利率等要素,以应对股票市场的非线性、不确定性数据。还进一步搭建了基于BP神经网络的投资策略模型。实

验结果表明, FSPVIS模型切实提升了投资决策的精准度, 尤其是在处理复杂多样的市场数据时, 体现出了不错的稳定性。

陈倩莹<sup>[5]</sup>就基于模糊逻辑和BP神经网络的投资策略展开研究, 在这项研究中, 提出了一种借助模糊推理机制的多目标组合投资策略。该策略把历史收益率、波动率等多因素进行了结合, 采用模糊多目标优化算法实现单目标线性规划。且借助仿真实验验证了该策略的有效性。该方法有效解决了股票市场里的不确定性难题, 进而提升了投资决策的稳定性及准确性。

谢旺旺<sup>[6]</sup>把目光投向了机器学习在量化投资中的应用, 尤其在疫情席卷的背景下, 探讨了噪声增强的股票市场环境对投资策略的相关影响。他采用随机森林与XGBoost模型对中证500成分股进行了预判, 形成了基于机器学习的投资组合策略。实证得出的结果表明, 机器学习模型在噪声环境里仍可取得超额收益, 尤其是在等权重和行业配置相关投资策略中, 表现明显胜过大盘基准。

陈家祥将强化学习应用于股票交易决策工作中, 推出了一套基于强化学习的自适应辅助决策系统。借助搭建股票交易模拟环境, 又引入分形股票理论, 采用DQN算法开展训练工作。实验结果说明, 该模型相较于传统策略而言, 大幅提高了年化收益率及夏普比率, 呈现出良好的自适应性及抗风险能力。

鉴于现有文献所存在的不足与研究漏洞, 本文构建了一个基于遗传算法的投资组合优化模型。该模型在目标函数中将股票的风险和收益纳入考量, 且利用遗传算法对投资比例进行优化处理, 旨在找到一个收益与风险相协调的最优投资组合。实验结果说明, 遗传算法可有效挖掘个股彼此间的潜在关联, 优化后的投资组合不仅能提高收益, 还能较好地把控风险。

## 1 遗传算法理论基础

遗传算法(GA)为一种模拟自然选择及遗传机制的优化算法, 它依靠选择、交叉和变异等操作, 模拟生物的进化阶段, 进而在复杂解空间里面搜索最佳解。GA的核心要义是保留优质基因块然后进行重组, 渐渐逼近最优解, 个体凭借适应度函数评估自身优劣, 把优质个体保留下来开展繁殖, 而交叉和变异操作引导算法探索新的解空间, 维持多样性以防过早陷入收敛。

GA适合处理高维复杂的优化难题, 其优势在于可处理非线性、无导数的目标函数和约束条件, 同时无需事先掌握问题空间的全貌, 依靠合适的编码样式、交叉及变异算子, GA能在较大搜索空间里以高效的方式找到最优解。为了躲开局部最优的困境, 设计算法时得平衡选择压力、探索与开发这三者的关系, 且在进化进程期间维持解的多样性。

GA的有效性跟问题的表示、选择机制以及算子设计紧密相关, 经过恰当设置参数, GA可适应各类优化任务, 但也会面临计算复杂程度大与收敛速度较慢的麻烦, 对于约束处理与多目标优化相关操作里, 一般采用罚函数、精英保留等策略以改善性能<sup>[7]</sup>。

## 2 实验方法

2.1问题建模。在进行股票投资组合优化事宜中, 投资者必须从给定的股票集合内, 明确每只股票的投资占比, 用以在风险

与收益之间实现恰当的平衡, 假设股票池里面一共存有N支股票, 分别记为 $S_1, S_2, \dots, S_N$ 。设第i支股票的预期收益率为 $r_i$ , 风险(通常以收益率方差或标准差表示)为 $\sigma_i$ 。投资者对各股票的投资比例记为 $w_i$ , 则投资组合可表示为向量

$$w = [w_1, w_2, \dots, w_N] \quad (2-1)$$

其中 $w_i \geq 0$ , 且满足约束条件

$$\sum_{i=1}^N w_i = 1 \quad (2-2)$$

表示全部资金均被分配完毕。

在投资组合理论中, 组合的整体预期收益率可表示为

$$R_p = \sum_{i=1}^N w_i r_i \quad (2-3)$$

其中 $\sigma_{i_3}$ 为股票 $S_i$ 与 $S_j$ 的协方差。

投资者一般都想在可承受的风险范围内获取更高的收益, 可把投资组合优化问题转化为一个多目标优化问题, 为了便于遗传算法顺利求解, 本文把这一情况简化为单目标形式, 即组建一个综合评价函数, 把收益与风险一同纳入考量, 现在定义目标函数为:

$$f(w) = R_p - \lambda \sigma_p^2 \quad (2-4)$$

其中, 参数 $\lambda$ 为风险厌恶系数, 用于平衡收益与风险的重要性。当 $\lambda$ 较大时, 算法倾向于选择低风险组合; 当 $\lambda$ 较小时, 算法则更注重收益的提升。

因此, 投资组合优化问题可表述为:

$$\begin{aligned} \max \quad & f(w) = \sum_{i=1}^N w_i r_i - \lambda \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij} \\ \text{s.t.} \quad & \sum_{i=1}^N w_i = 1, \quad w_i \geq 0, \quad i = 1, 2, \dots, N \end{aligned} \quad (2-5)$$

2.2遗传算法设计。针对之前所说的投资组合优化模型, 本文采用遗传算法作为求解手段, 遗传算法借助模拟自然界“优胜劣汰”的进化机制, 在复杂的搜索空间中求索最优解, 针对本研究聚焦的具体问题, 算法的设计与实现主要涉及以下几个方面。

在本研究中, 每个个体代表一个投资组合。采用实数编码, 将股票投资比例直接表示为一个向量 $w=[w_1, w_2, \dots, w_N]$ 。初始种群通过随机生成若干可行解得到, 并在生成与进化过程中对各个体进行归一化处理, 以确保投资比例非负且总和为1。这种编码方式直观、计算效率高, 且便于在交叉和变异后保持约束条件。

适应度函数能衡量个体的好坏水平, 作为遗传算法的核心构成, 本文把投资组合的综合目标值作为适应度考量, 综合权衡预期收益跟风险水平, 收益上扬、风险下降的组合, 其适配度就越大, 对不符合约束条件的个体, 适应度值按照一定比例实施惩罚, 以维持种群的可行特性与搜索稳定性。

选择操作判定哪些个体可遗传到下一代,本研究采用锦标赛选择法,即随机抽取若干个体进行比对,留下适应度高的个体进入下一代,该手段简单高效,能在维持种群多样性的同时加大优秀个体的传递概率。

交叉操作的用途是在个体间交换基因信息,产生别样的投资组合,本文借助算术交叉的途径,凭借随机系数在父代个体之间开展线性组合,进而在收益—风险空间里产生新的潜在解,交叉之后对结果向量再次归一化,保障投资比例依然契合约束要求。

变异操作采用随机微调部分投资比例手段,添加种群的多样性,避免算法落入局部最优解,具体实施办法是:随机挑选若干只股票的权重,在较小范畴内进行扰动,随即对整个组合开展归一化,对变异幅度进行小范围控制,避免大幅波动干扰解的可行性。

### 3 实验

3.1实验数据。为核实所构建的投资组合优化模型及遗传算法的有效性,本文拣选美国股市里若干具有代表性的股票作为研究对象,顾及行业分布与市场特质,选定科技、消费、能源等领域的7支股票,分别为:苹果公司(AAPL)、微软公司(MSFT)、特斯拉(TSLA)、英伟达(NVDA)、可口可乐公司(KO)、埃克森美孚公司(XOM)和强生公司(JNJ)。这些股票展现出较高的流动性与市值特征,能比较好地反映市场整体的面貌。

实验数据是从公开金融数据平台采集而来的,取样的时间范畴为2020年1月至2024年12月,数据采用月度收盘价作为周期,为降低短期波动形成的干扰,本文依据月收益率做计算,把收益率界定成:

$$r_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} \quad (3-1)$$

其中,  $P_{i,t}$  表示第  $i$  支股票在第  $t$  个月的收盘价,  $r_{i,t}$  为对应的月收益率。依照整个样本期的收益数据,计算各股票的平均收益率跟波动率,并进一步评定股票间的协方差矩阵,以反映不同资产相互间的风险关联程度,为保障数据的可比性与稳定性,全部价格数据均做过缺失值处理与对数收益率平滑化处理,最终算出的平均收益率和波动率统计结果见表3.1。

表3.1 数据展示

股票代码	平均月收益率(%)	波动率(%)
AAPL	1.85	6.42
MSFT	1.6	5.8
TSLA	2.95	10.5
NVDA	2.75	9.8
KO	0.95	3.1
XOM	1.3	5.2
JNJ	0.9	2.9

从表里面可以看出,科技类股票(例如TSLA、NVDA)虽说收益回报率较高,但相对而言波动性更大;而消费类与医药类的股票(如KO、JNJ)波动幅度不大,风险相对较易掌控,这种差异给投资组合优化搭建了基础条件,也为遗传算法对收益与风险进行平衡探索提供了大量搜索空间。

3.2实验过程。本文凭借前面提到的投资组合优化模型与遗传算法设计方案,做了实证实验,实验的主要目的是借助遗传算法在既定股票池中自动探寻投资比例,使投资组合在收益与风险的关系上达到较优平衡,实验于Python平台实现,主要凭借NumPy、Pandas和Matplotlib等科学计算库来进行数据处理与结果可视化,对股票的历史价格数据做预处理后,得出平均收益率、波动率与协方差矩阵,把这些指标当作模型输入,用以衡量不同投资组合的表现。

结合前期测试与相关文献的经验,遗传算法的参数如此设置:

表3.2 参数设置

参数名	参数值
种群规模	50
最大迭代次数	1000
交叉概率	0.8
变异概率	0.05
风险厌恶系数	3

初始种群采用随机手段生成,保障各个个体的权重达到非负与归一化约束,处于实验进行期间,每一代个体都要经过适应度评估、选择、交叉及变异操作,进而渐渐进化出表现更棒的投资组合。

为保障结果的稳定性,实验重复做10次,选用适应度最高的投资组合作为最终结果,各次运行结果的最优解相互间差异较小,说明算法呈现出较好的稳健性与收敛一致性。

3.3实验结果。为检验遗传算法在优化股票投资组合中的表现,本文对算法收敛过程、最优投资权重的分布以及组合收益表现进行了研究分析,实验得到的结果如下说明。

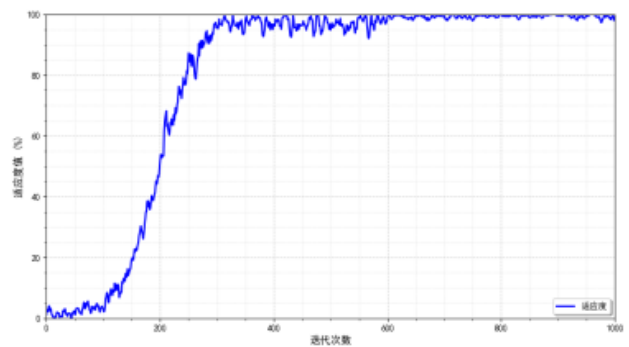


图3.1 适应度进化曲线

图3.1显示出遗传算法优化过程中适应度值随迭代次数的

变动情形,从图中不难看出,算法在差不多前200代时经历了明显的上升阶段,适应度的值迅速增高,暗示优质投资组合已逐步被发现且留存,曲线慢慢变得平稳,差不多到第300代过后时段,适应度值基本达到饱和,波动范围不大,说明算法已收敛至稳定的最优区域范围。

从结果可以明显看出,最优方案把资金大部分分配给稳健型股票,其中可口可乐(KO)和埃克森美孚(XOM)的权重比较高,各自占比25%和22%;苹果公司(AAPL)和微软公司(MSFT)所占权重分别为20%与18%;强生公司(JNJ)占据的比例为15%。特斯拉(TSLA)跟英伟达(NVDA)的投资比例经优化调整成0%,该结果体现出遗传算法在面临风险约束条件时的理性偏好:虽然科技股存在较高的收益潜力,但其波动幅度太大,对整体风险的影响程度高,于是在优化结果之后被彻底去除掉,算法更倾向于挑拣收益稳定、风险较低的消费与能源类股票,进而在风险与收益的权衡里实现总体最优。

图3.2表明了样本期内不同投资策略累积收益率的变化格局,囊括遗传算法优化而成的组合、全仓单只股票投资(TSLA与JNJ)以及随机投资组合作为对比样本,遗传算法组合于整个时间段内表现稳定,累计收益率不断稳步攀升,最终的收益明显比随机投资组合与防御型单股投资(JNJ)要高,且波动的程度明显低于激进型单股投资。

当收益增长的时候,遗传算法组合的波动幅度微弱,体现出不错的风险管控本事,遗传算法能有效地在收益跟风险之间寻求平衡,形成更具投资价值的组合计划。

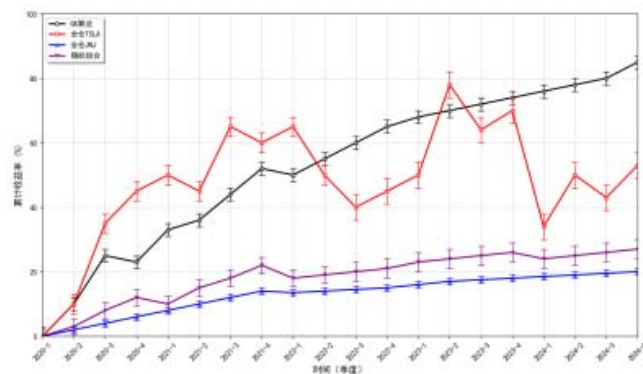


图3.2 累计收益计算

把三项结果综合考量,可以得出下面的结论:

- (1) 遗传算法能快速收敛,平稳输出较优的投资组合情况;
- (2) 要是风险厌恶系数设定得合乎理性,算法倾向去选择波动幅度小、收益稳定的股票;
- (3) 优化后的投资组合在长期收益及风险控制上都胜过随

机与单一投资策略,证实所提出方法具有实际的可行性。

实验结果验证出遗传算法对股票投资组合优化问题的有效性,其自适应搜索机制能在复杂的收益与风险所形成的空间里自动调整投资权重,呈现出较强的全局优化本领,为投资决策给出了可参考的量化凭据。

#### 4 结论

本论文以股票投资组合优化当作研究对象,依靠历史市场所记录的数据,创制了一个兼顾收益与风险的优化模型,而且采用遗传算法作为求解途径,借助对7支代表性的股票做实证分析,证实了遗传算法在投资组合优化中的有效性,跟随机投资和单一股票投资相比照,遗传算法优化后的投资组合在样本期内实现了更高且更平稳的累计收益,该方法可以在复杂的收益—风险空间之中有效开展寻优操作,兼顾投资回报与风险暴露,具备良好的稳健性与可解释性。

本研究依旧存在一些局限性,实验所用的样本数量偏少,只采用8支股票,结果的推广范围依旧有限;风险衡量采用历史波动率指标,未将市场结构变化以及极端风险因素纳入考虑;算法参数凭借经验完成设定,尚未采用自适应机制。本文验证了遗传算法于股票投资组合优化的可行性及有效性,为智能化投资决策给出了一种可以借鉴的技术路径,未来工作可在拓宽样本范围、改善风险度量方法以及提升算法自适应性能等方面进一步健全,以增强模型的实用效果和稳定性。

#### [参考文献]

- [1]吴敏.中国人保股票投资规模激增[N].华夏时报,2025-09-01(011).
- [2]胡永新.股票成险资下半年投资首选[N].北京商报,2025-08-25(007).
- [3]林星.基于局部相似性分析的股票市场投资策略研究[D].山东大学,2017.
- [4]许文伟.基于模糊逻辑的智能股票投资模型[J].中国新通信,2017,19(16):147-148.
- [5]陈倩莹.基于模糊理论的金融数据预测模型与交易策略算法研究[D].湖南大学,2018.
- [6]谢旺旺.基于机器学习的股票投资策略研究[D].哈尔滨工业大学,2022.
- [7]侯广林.股票投资组合中若干优化算法的比较研究[J].考试周刊,2013,(15):182-18.

#### 作者简介:

刘梓彤(2008--),女,汉族,江西省信丰县人,高中,研究方向:股票投资、数字金融、遗传算法。