

摘 要

随着经济发展的加快,中国的金融市场和金融体制不断完善。在国民人均可支配收入不断增长的趋势下,基金作为一种主要的大众投资方式,在近两年的时间内,呈现了迅猛发展的势头。基金逐渐成中国资本市场上的主要机构投资者,其投资组合直接作用于证券市场的价格波动。因此,建立有效的基金投资组合模型对中国金融市场的稳定发展具有十分重要的实践意义。

本论文共分为四个部分:第一部分,综述国内外有关文献研究现状,提出了研究的意义;第二部分,界定了基金和基金资产的概念,分析了基金资产的构成,剖析了现有投资组合模型的不足及相应的解决思路;第三部分对股票资产按照风险角度进行了分类,建立了基金资产分类的模糊数学模型。将股票和国债的单位风险所获得的超额收益作为遗传算法中的被操作对象,以遗传算法中的目标函数的形式来建立了投资组合模型,将投资组合模型作为遗传算法中的适应度函数,最终来求解不同风险偏好系数下的最优解。第四部分,选取上证 180 样本股数据为样本,应用遗传算法计算了两种不同风险偏好系数下的最优投资比例和投资绩效,与现有的三类基金的投资比例进行了比较分析,并从经济意义角度分析现有应用遗传算法求解投资组合模型与本文的模型的不足。

本论文的创新见解表现在两个方面:第一,在投资组合模型中加入风险偏好因素,解决了原有模型只适用于风险中性的不足,取代了投资者风险中性的假设,使投资者可以根据自身的风险偏好程度来选择投资组合。第二,以遗传算法中的目标函数的方式建立了基金投资组合模型,以风险调整的收益的最大值作为遗传算法中的目标函数,应用遗传算法进行了求解。避免了应用二次规划对原有模型求最优解的过程中,出现的非线性程度高,收敛缓慢并且收敛于局部解的问题。

关键词: 基金资产; 风险偏好系数; 遗传算法; 投资组合

A study on fund portfolio based on Genetic Algorithms

Abstract

With the development of economy, Chinese financial market and system makes rapid progress as well. At the situation of Chinese people's governable income' increase, as a kind of important investing method for Chinese people, Fund grew rapidly in recent two years. Fund is becoming the main institutional investor in Chinese capital market. Its investment combination directly effects the price movement on the stock market in China. This means that building an efficient model of Fund's investment combination will make great contribution in Chinese financial market' development

The content of chapter one analyzes the studies of domestic and foreign literatures, and points that the current investment combination theory is weak in calculating and applicability. The second part suggests that we can take use of GA to solve the problem that SQP cannot settle. Second, chapter two suggests bring risk taste coefficient into the Fund investment combination model, which improves the applicability of the model. The third part gives a new classification of Fund's assets, based on the excess risk coefficient and builds a model. It takes stock as chromosome on GA. It assumes risk taste coefficient and establish contact the excess risk coefficient with risk taste coefficient and build up a model based on the fit function of GA. It take the function that is built up above as target function in GA It uses the risk taste coefficient to ameliorate the model. Finally, in the chapter four ,the model is proved by the data in shanghai stock market from June to September in 2006. Compared with the normal fund data, the model is more efficiency.

The main characteristic and innovation of this paper is three. First, a new stock assets sort has been done from the stock risk. Second, taking risk taste coefficient into the model solve the problem of without risk taste in model. Third, it builds up a new model based on the thought of GA and mistiness mathematics as the fit function in GA. The biggest numerical value is the best result of the account. This method solves the problem that was brought by SQP which was used before.

Key Words: Fund assets; Risk averseness coefficient; Genetic Algorithm; Portfolio

独创性说明

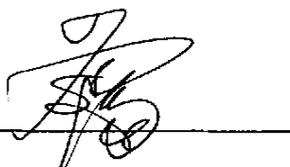
作者郑重声明：本硕士学位论文是我个人在导师指导下进行的研究工作及取得研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得大连理工大学或者其他单位的学位或证书所使用过的材料。与我一同工作的同志对本研究所做的贡献均已在论文中做了明确的说明并表示了谢意。

作者签名： 日期：2017.12.30

大连理工大学学位论文授权使用授权书

本学位论文作者及指导教师完全了解“大连理工大学硕士、博士学位论文版权使用规定”，同意大连理工大学保留并向国家有关部门或机构送交学位论文的复印件和电子版，允许论文被查阅和借阅。本人授权大连理工大学可以将本学位论文的全部或部分内
容编入有关数据库进行检索，也可采用影印、缩印或扫描等复制手段保存和汇编学位论文。

作者签名：



导师签名：



2007年12月30日

1 绪论

1.1 研究背景

1.1.1 问题的提出

中国已经成为全球最有潜力的投资市场。2006年,中国GDP突破20万亿元,增长了10.5%,人均可支配收入提高了10.4%,达到11759元;伴随着近年来金融市场与金融体制的完善,外部投资逐年增多;2007年中国加入世贸组织,完全兑现了当时加入世贸组织的承诺,减少了贸易壁垒,更是减少了货币流通壁垒,为外资流入中国提供了政策上的条件;人民币持续的升值,直接导致了外资在华投资类型的改变,由原有在华建厂投资方式改为更有利可图的投资于金融市场,尤其是证券市场。目前,中国的人民币存款利率较低,货币乘数效应比较明显,市场上热钱较多,热钱的最直接出路就是证券市场。资料表明:机构投资已经成为主流投资力量,机构包括一些基金,已经成为证券市场上的投资主体。2006年下半年开始,中国出现了强劲的牛市。伴随着证券市场的繁荣,基金市场也得到了跳跃式的发展,截至到2007年4月份,中国共有基金账户达到2000万户以上,日开户数量达到4761户,基金市场总净值突破1万亿元。基金市场异常火爆。

1952年,马科维茨的投资组合理论的诞生,奠定了现代投资理论的理论基础。各种有关资产组合的方法与模型纷纷问世^[1]。目前,中国大多数学者主要研究如何来优化现有的投资组合模型,并没有建立有效的投资组合模型来分析我国现有的金融市场,并且只是对原有的模型进行了简单的改进。本文研究将有机结合模糊数学与遗传算法两种数学方法,建立了一种有效的模型来确定基金投资组合的质量。本文研究将利用遗传算法的思想,建立基于遗传算法的基金市场投资组合模型,具有一定的理论意义。论文拟选择风险指标对基金资产加以分类,能够有效地帮助基金投资者认识基金业绩,以便投资者进行投资组合,从而降低非系统风险。因此,对投资者投资有具有一定的现实意义。

1.1.2 研究意义

2004年,中国的证券市场共有161只证券投资基金正式运作。2005年,共有223只证券投资基金,其中54只封闭式基金,169只开放式基金。截至2006年底,共有321只证券投资基金正式运作,其中封闭式基金53只,开放式基金268只。在三年之中,基金数量翻了一倍。随着新基金的不断设立和基金行业的超常规发展,恰当的监测、分析和评估基金的风险已越来越重要。证券投资基金作为证券市场的重要参与者,在稳定

市场秩序、倡导理性投资等方面发挥着不可替代的作用。基金作为一种建立在委托-代理关系基础上的金融产品出现在金融市场上。能否将广大投资者的资金保值、增值同时降低投资的风险,使得基金的风险评估问题显得尤为重要。对基金风险大小进行评估的前提是对基金有正确的认识,才能对基金的业绩评价更规范、更合理。

基金的主要投资对象是股票,债券和新推出的股指期货,如何认识基金资产、如何来进行投资组合便成为了首要问题,对投资者进行投资有一定的指导意义。目前,我国研究投资组合的学者较多,但是大多数是在原有的基础上进行改进,增减一些指标,来适应市场的发展,还有一些学者应用最新的数学方法,来使得原有的计算更精确,其中以遗传算法最为突出,但是这些研究主要还是在修正模型,而没有进行建立新模型,本研究运用遗传算法的思想和模糊数学的知识,建立了基于遗传算法的投资组合模型。使得本模型在优化问题和计算结果上兼顾了遗传算法的鲁棒性、并行性、自适应性和领域无关性,以及模糊数学的类别的合理性^[2]。

中国证券市场要得到长远的发展,成为一个成熟的市场,必须以基金作为机构投资者在证券市场上占据主导地位,只有这样,才能实现证券市场稳定的发展,最终通过证券市场的发展,来带动基金市场的发展与稳定。从基金经理的角度来分析,基金本身就是投资组合,将基金资产按照一定的比例进行分配和投资,建立准确的投资组合模型是十分必要的,基金的稳定发展又可以反向的促进证券市场的发展;从个人投资者的角度来分析,中国个人投资者的数量十分庞大,个人投资者在整个证券市场上所占比例也不容忽视,但由于信息不对称,选股能力以及个人能力的限制,导致个人投资者的风险比机构投资者的风险较大,这部分投资者若不进行合理的投资组合的话,势必增大整个证券市场的非系统风险,对整个市场的合理有序的发展不利,所以本文建立的投资组合模型无论从个人投资者和基金经理人的角度来分析,都具有现实意义。

原有的投资组合模型,对现实的假设条件要求太多,假设条件和现实不符,本论文通过将投资组合模型中加入风险偏好因素,解决了原有模型只适用于风险中性的不足,取代了投资者风险中性的假设,使投资者可以根据自身的风险偏好程度来选择投资组合,扩大了投资组合模型的应用范围。

1.2 国内外文献综述

1.2.1 投资组合模型方面

现代投资组合理论是研究有关对多种资产进行选择 and 组合的问题,即投资者在权衡风险与收益的基础上,如何使自身效用最大化以及由此对整个资本市场产生怎样的影

响。现代投资组合模型共经历了三个阶段的发展，分别由三个重要理论模型作为研究的依据。

第一阶段：1952年由马科维茨（Markowitz）提出来的均值-方差模型。该模型是一种局部均衡分析，即从单个投资者来考虑问题，试图确定在投资者个人效用最大化的情况下资产组合的构成，个人效用最大化的标准是：在风险一定的情况下实现收益的最大化，在收益一定的情况下实现风险的最小化^[3]。均值-方差模型的基本假设是：(1) 投资者遵循效用最大化原则；(2) 投资者是风险回避者，即在同等收益下，投资者选择最小风险的投资组合；(3) 每个投资者根据期望值、方差以及协方差来选择最佳投资组合；(4) 证券市场是完善的（Perfect）、无交易成本、且证券可以无限细分^[4]。

马克维茨的均值-方差模型为

$$\begin{aligned} \min \sigma_p^2 &= \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij} \\ \text{s.t.} \quad &\sum_{i=1}^n x_i E(R_i) = E(R_p) \\ &\sum_{i=1}^n x_i = 1 \\ &x_i \geq 0, \quad i = 1, 2, \dots, n \end{aligned} \quad (1.1)$$

其中： $E(R_p)$ 为证券组合的期望收益率； σ_p^2 为证券组合的方差。 $x_i, i=1, 2, \dots, n$ 为投资者选择证券*i*的投资比例系数； σ_{ij} 为第*i*种与第*j*种证券收益率的协方差； $E(R_i), i=1, 2, \dots, n$ 为第*i*种证券的期望收益率。

从理论发展的角度来看，戴玉林（1990）对 Markowitz 模型产生的背景、模型的内容及其影响和模型的局限性做了较为全面的分析和评价；施东晖（1996）以 1993 年 4 月至 1996 年 5 月上海证交所的 50 种股票为样本，以双周收益率为指标，采用随机简单等权组合构成 50 个“*n* 种证券组合”（ $n=1, 2, \dots, 50$ ）来推断股市组合分散系统风险的能力，得出“中国股市未能通过采用模型有效分散风险”的结论；尹占华（2003）采用模糊概率方法对 Markowitz 模型进行了修正，优化原有模型；苏萍、肖东荣（2003）采用线性条件约束下的线性逼近法来求解 Markowitz 模型，线性逼近的方法是将二次型算法转化为线性方法解决，误差有所减少，优化了原有 Markowitz 模型；张树斌、白随平和姚立（2004）建立了含有交易成本的均值-方差-偏度资产组合优化模型，它是对传统均值-方差模型的直接扩展，通过加入交易成本及其它现实约束条件，使新模型与现有模型相比其实用性显著增强，并对模型进行了敏感性分析；杨明辉、张智光、任百琳、谢煜（2005）等人

通过引进投资者对待收益和风险的心理取向系数，将收益和风险两目标融为一体，从而将双目标优化模型转变成单目标优化模型，优化了 Markowitz 投资组合模型。

从实证角度来研究 Markowitz 投资组合模型的学者较多。张一弛（1997）分别选择了上海和深圳各 25 种股票，对 Markowitz 模型，单指数模型和 EGP 模型进行了比较实证研究；李善民、徐沛（2000）采用深沪股市 824 只股票自上市以来至 1998 年年底的每日价格变动数据，利用计算机建立数学模型协助研究，采用 Markowitz 模型、单指数模型、EGP 模型推算特定时间特定股票样本的有效组合，测算股票组合风险变动规律，随机等权股票组合的收益、方差，试图在全面评价 Markowitz 模型及相关简化算法基础上，结合我国股票市场的实际进行实证研究；苏敬勤，陈东晓（2002）在西部七省区的股票中筛选了 11 只股票，采用 Markowitz 模型测算股票组合的收益、方差，通过切点证券组合的办法，采用定量和定性的办法最终确定了最佳证券组合和个股；刘雪燕、张敬庭（2004）通过对台湾股票有效市场界面进行实证分析，来对马科维茨理论进行验证。

第二个阶段：由夏普等人在马科维茨的均值一方差模型的基础上提出来的资本资产定价模型（CAPM）^[5]。它是从整个资本市场的角度来考虑问题，即把所有投资者行为都包含在内。CAPM 基本理论假设是：（1）每一个投资者在关于证券的未来表现预测的基础上行事，预测以预期收益、收益的标准差和收益率的相关系数表示；（2）每一个投资者以投资组合理论建议他们应该行动的方式那样行动；（3）每个投资者在纯利率水平能够借入和贷出他所希望的或多或少的数额。无论投资者希望借入还是贷出，这一利率是相同的，且对每一个投资者都相同；（4）无交易成本；（5）投资者是风险厌恶者，追求期望效用的最大化^[6]。

CAPM 主要是研究证券市场中资产的预期收益率与其风险之间的关系，以及均衡价格是如何形成的。其主要特点是建立了测度公司证券的市场风险(系统风险)的 β 系数。

$$\beta = Cov(R_i, R_M) / \delta_M^2 \quad (1.2)$$

β 系数表示证券 i 的风险相对于市场风险的反应的灵敏度，一种资产的预期收益率可用这种资产的风险的相对测度 β 的值来衡量。

资本资产定价模型可以表示为：

$$R_i = R_F + \beta_i(R_M - R_F) \quad (1.3)$$

其中： β_i = 第 i 种证券的 β 系数； δ_M = 市场组合的收益率的标准差； R_i = 第 i 种证券的预期收益率； R_M = 市场证券组合的预期收益率； R_F = 无风险利率，如美国的 T-B 国债利率。

在一个均衡的市场中，每个投资者都面临最佳风险组合，这些投资组合集合就形成了投资组合的有效集，通过无风险证券点 R_f 作一条直线与有效集相切，得到的切点就为这个投资组合的最优解，通过无风险利率和有效组合点的直线被称为资本市场线 (CML)，它是同时包括无风险资产和投资组合的有效集，所有投资者最终都会位于资本市场线上的某一点。但并不是所有的证券或组合都在资本市场线上，事实上，从资本资产定价模型中推导有效边界的过程中可以知道，我们知道所有风险资产和无风险资产的组合，除了位于有效边界上，都在资本市场线的下方。资本市场线可以表示为通过无风险资产和一个风险组合的直线方程，所以资本市场线可以写成：

$$R_e = R_f + (R_M - R_f)\delta_e / \delta_M \quad (1.4)$$

其中： R_e = 有效投资组合的预期收益率； R_f = 无风险收益率； R_M = 市场组合的预期收益率； δ_e = 收益的标准差； δ_M = 市场组合的标准差

对所有有效组合来说，截距项可以被看成是风险的市场价格，它是增加一个单位有效组合的风险水平所得到的超额收益。无风险收益可以看成是时间的价格。因此，一个有效的组合的期望收益是：

期望收益 = 时间价格 + (风险价格) * (风险数量)

与资本市场线相对应的是证券市场线：

我们假设 $B = \delta_{im} / \delta_M^2$ ，那么我们将证券市场线写成：

$$R_i = R_f + \beta_i (R_M - R_f) \delta_{im} / \delta_M \times 1 / \delta_M \quad (1.5)$$

实际上，这是期望收益—— δ_{im} / δ_M 空间中的一条直线方程。因为 δ_{im} / δ_M 被定义为任何证券或组合的风险，我们可以看到证券市场线，就像资本市场线一样，表明任何证券的期望收益等于无风险利率加上风险的市场价格与该证券或组合的风险量的乘积。所以 CAPM 可以写成：

$$R_i = R_f + (R_M - R_f) \delta_{im} / \delta_M^2 \quad (1.6)$$

其中： $(R_M - R_f) \delta_{im} / \delta_M^2$ 定义为风险的市场价格； δ_{im} 为证券 i 的风险指标。 δ_{im} / δ_M 度量了一个证券的风险影响市场组合风险的程度。

国内有关 CAPM 模型的研究学者很多，但是对该理论研究集中在实证研究上的居多，景乃权 (2000) 结合夏普关于网络股的定价问题介绍了该模型在实际中的应用。陈浪南、屈文洲 (2000)，试图运用上海股票市场的数据，对 CAPM 模型进行实证检验，尤其注重在 CAPM 模型中举足轻重的 β 值的分析和测量上，并根据股市中的三种市场格局 (上升、下跌和横盘) 划分了若干的时间段加以分析，进而检验 β 的解释力^[7]。陈小

悦、孙爱军（2000）通过检验 CAPM 在中国股市的有效性，截面检验结果表明 β 对中国股市的平均收益不具有解释能力，从而否定了其在中国股市的有效性假设的结论。靳云汇、刘霖（2001）利用多种方法检验了 CAPM 在中国股票市场上的适用性，并得到了股票收益率不仅与贝塔之外的因子有关，而且与贝塔之间的关系也不是线性关系的结论。马静如（2001）利用深圳股票市场数据对该市场作了检验，得出 CAPM 模型不符合我国的深圳股票市场的结论。朱文辉、汪前明（2002）通过研究表明在上海股票市场上，非系统风险对投资收益率已经没有显著影响，而非系统风险的影响却是十分显著，投资收益率与系统风险的关系已经不符合 CAPM 模型，即投资收益率与贝塔并不呈现线性关系。黎金龙（2005）采用实证分析方法和经验验证手段，结合中国证券市场实践，对资本资产定价模型 CAPM 进行了有效性检验，得到 CAPM 在一定程度上可衡量上海股票市场风险与收益之间的关系，但有效性不强，非系统性风险对股票收益具有重大影响的结论。许涤龙、张钰（2005）通过研究也表明资本资产定价模型关于风险与收益之间关系的描述在上海股票市场中并不是完全成立的，并发现上海股票市场的投机性较强的结论^[8]。

在理论研究和模型修正的研究中，美国 New Orleans 大学的 Edward M·Miller 将投资者的无偏估计这一条件放宽，通过对投资者之间关于收益估计偏离的分析，得到一种关于投资收益与股票价格变化的新解释，并初步将这种偏离进行量化后，给出了一个对 Sharpe 理论的修正模型。周少甫，杜福林（2004）基于多元 GARCH 理论对资本资产定价模型进行了研究。李海涛、王建华和王永舵（2005）应用 CAPM 模型构造了 MGARCH(1.1)模型，并用其对上海股票市场的几支股票进行了条件 CAPM 实证研究。杨雪莱、梁四安、李善民一种新的 Downside Risk 度量方法—Shortfall 度量下导出了相应的 Sh-CAPM，得到了 Sh-CAPM 的 β 系数的表达式，利用我国股票市场的数据进行了实证研究，研究结果表明，原有的 CAPM 模型在我国股票市场的适用性已经降低，相反，Sh-CAPM 模型的解释能力更高的结论。余志红（2005）对 CAPM 在中国证券市场应用有效性的因素进行了分析，并提出了布莱克 CAPM 模型更适合实际情况的结论^[9]。

第三个阶段：罗斯提出来的套利定价理论（APT），APT 的基本思想是市场上一物一价，如果存在一物多价的情况就会产生无风险套利机会，而无风险套利将使一物多价消失恢复到一物一价的市场均衡状态^[10]。

国内有关套利定价理论的研究主要有：张妍（2000）利用套利理论对上海股票进行了经验检验，利用实际数据求解因子个数并进行多元线性回归的检验是基础，利用“自方差”和“证券规模”进行检验是在第一个检验的基础上对 APT 悖论的否定^[11]。曹红英、阳玉香（2005）利用公司规模、市值与账面价值比以及市盈率对股票收益率没有显

著影响，中国股票市场价格的变动是随机漫步的，即套利定价模型在我国证券市场是不适用的结论。董智勇、李长青、常秋萍、马宝霞（2005）根据上海股市的历史数据建立股市套利定价模型（APT），并考虑 ST 股票对建立 APT 模型的影响。得到了一方面存在其它因素影响 APT 模型，另一方面也说明 APT 模型本身也存在缺陷的结论。苏萍（2006）用探测性因子方法提取因子，并用二路径回归法对提取的因子进行验证。验证结果均表明：套利定价理论普遍适用于中国深市股市市场，确实存在一个以上的因素影响股票的收益率^[12]。

1.2.2 优化投资组合方法方面

Sharpe 的单指数模型。Sharpe 继 Markowitz 之后于 1963 年提出“单指数模型”，将“均值-方差模型”予以简化。他认为 Markowitz 的投资组合分析中，方差-协方差矩阵太过复杂不易计算，因此提出对角线模式来简化方差-协方差矩阵中的非对角线元素。此模型假设证券间彼此无关且各证券的收益率仅与市场因素有关，这一因素可能为股票市场的指数、国民生产总值、物价指数或任何对股票收益产生最大影响的因素。经由 Sharpe 的模型，任一股票收益率可由单一的外在指数来决定，大大简化了 Markowitz 模型的分析工作。夏普提出单因素模型的基本思想是：当市场股价指数上升时，市场中大量的股票价格走高；相反，当市场指数下滑时，大量股票价格趋于下跌。据此，可以用一种证券的收益率和股价指数的收益率的相关关系得出以下模型：

$$r_{it} = A_i + \beta_i r_{mt} + \varepsilon_{it} \quad (1.7)$$

该式揭示了证券收益与指数（一个因素）之间的相互关系。其中 r_{it} 为时期内 i 证券的收益率。 r_{mt} 为 t 时期内市场指数的收益率。 A_i 是截距，它反映市场收益率为 0 时，证券 i 的收益率大小。与上市公司本身基本面有关，与市场整体波动无关。因此 A_i 值是相对固定的。 β_i 为斜率，代表市场指数的波动对证券收益率的影响程度。 ε_{it} 为 t 时期内实际收益率与估算值之间的残差。

随后，Sharpe 有鉴于 Markowitz “均值-方差组合模型”，及其早期提出“单指数模型”中方差与投资比例不呈线性关系，必须用二次规划法求解，求解程序复杂。因而于 1967 年提出线性规划法，将 Markowitz 的组合模型以线性规划的方式求解。根据 Sharpe 进行的实证研究，当股票种类达 20 种以上时，投资组合的非系统风险逐渐趋于零，此时风险只生剩下系统风险，从而只与市场因素的方差有关，投资组合的标准差逐渐成为一个线性函数，因此可用“线性规划法”迅速找出有效边界。

林春艳、冯恩民（2002）根据各种资产的相关情况，分别建立了不同约束条件下的资产组合投资优化模型。詹正茂、陈刚、张伟（2003）采用预期收益率和收益率方差两

项指标,从风险控制的角度出发建立证券投资组合模型,来确定最优化的投资组合^[13]。黄向阳、陈学华、杨辉耀(2004)采用 RT Rockafellar 和 SUryasev 的一种优化算法,构造了一个以条件风险价值代替标准差度量风险的投资组合优化模型^[14]。何琳洁、文凤华、马超群(2005)提出了一种新的风险度量技术——一致性风险价值——来度量投资组合的信用风险,在此基础上建立了一致性风险价值的投资组合优化模型,并运用线性规划技术进行组合优化,并且得到了运用基于一致性风险价值的优化模型进行投资组合的结果,优于运用基于风险价值的优化模型的结论^[15]。李华、李兴斯(2005)基于熵以及差熵的概念,在研究其均值方差模型的基础上,提出用熵和差熵来作为风险的度量方法,从而建立了几种关于熵的证券投资组合优化模型^[16]。徐斌、方卫国、刘鲁(2006)在已有研究基础上讨论一种基于净现值和净现值成本的多项目多期投资组合优化的双目标模糊相关机会模型,并且提出了一种集模糊模拟、改进的神经网络与遗传算法于一体的人工智能算法,结合一种能够实时调整权数的评价函数法,从而对模型进行有效求解^[17]。刘晓星(2006)构建基于 CVaR 约束的投资组合优化模型,该模型虑及了投资组合资产的交易成本、交易限制、资金约束和投资者的风险承受度等条件建立模型^[18]。肖冬荣、黄静(2006)引入偏度水平,并用伸缩指标相应做出均值、方差和偏度三个模糊目标,形成一类新的非线性多目标投资组合模型。郭福华、邓其飞(2007)采用 Semi-A.D 的方法来度量投资组合的风险,建立了动态的 Semi-A.D 投资组合选择模型。鲁美娟(2007)基于 CVaR 和 RAROC 建立了投资组合优化模型。

随着遗传算法在求解最优投资组合比例的应用逐渐增多,有些学者应用遗传算法建立了投资组合模型,周群、孙德宝(2001)通过将遗传算法引入到证券投资分析领域,对最佳证券投资组合问题进行了优化计算^[19]。陈科燕、肖东荣(2003)通过研究遗传算法进行求解投资组合的有效集,即投资组合的有效边界,通过误差曲线与有效边界的交点来确定最优投资组合比例,得到了不同效用取向下的最优投资比例^[20]。金汉均、王洪峰(2004)分析了应用遗传算法求解马克维茨投资组合模型,提出了一种采用最优保存策略的遗传算法求解夏普模型的方法,并且实现了 N 种证券投资组合优化的模拟分析,得到了比二次规划更好的结论^[21]。江家宝、尤振燕、孙俊(2007)通过比较用微分进化算法和遗传算法(GA)优化同样的资产对象所得到的期望收益率均值与方差,建立了基于微分进化算法的多阶段投资组合优化模型^[22]。

1.3 论文研究框架与创新

1.3.1 研究思路和研究方法

论文研究采用了定性和定量研究相结合，理论与实证分析相结合的研究方法。本文明确了基金资产的分类，初步选取了基金资产分类的指标，研究了目前应用的投资组合模型不足。通过聚类分析的方法对基金资产分类的指标进行了筛选，运用模糊数学的知识对基金资产分类确定了其数学基础，将遗传算法的思想应用于投资组合模型的建立，选择指标对基金资产进行分类，将基金资产作为基因值，建立投资组合模型。然后，进行基因的“选择”、“交叉”和“变异”，运用主成分分析法确定了评价指标的权重，采用隶属度寒暑对指标进行单一指标评价，通过定量分析，测定在三种情况下的投资组合的最优解。在我国目前金融市场允许的情况下，测算最优投资比例，与现有基金十大重仓和其投资比例进行比较，最终验证模型的优越性。

论文的技术路线如图 1.1 所示：-

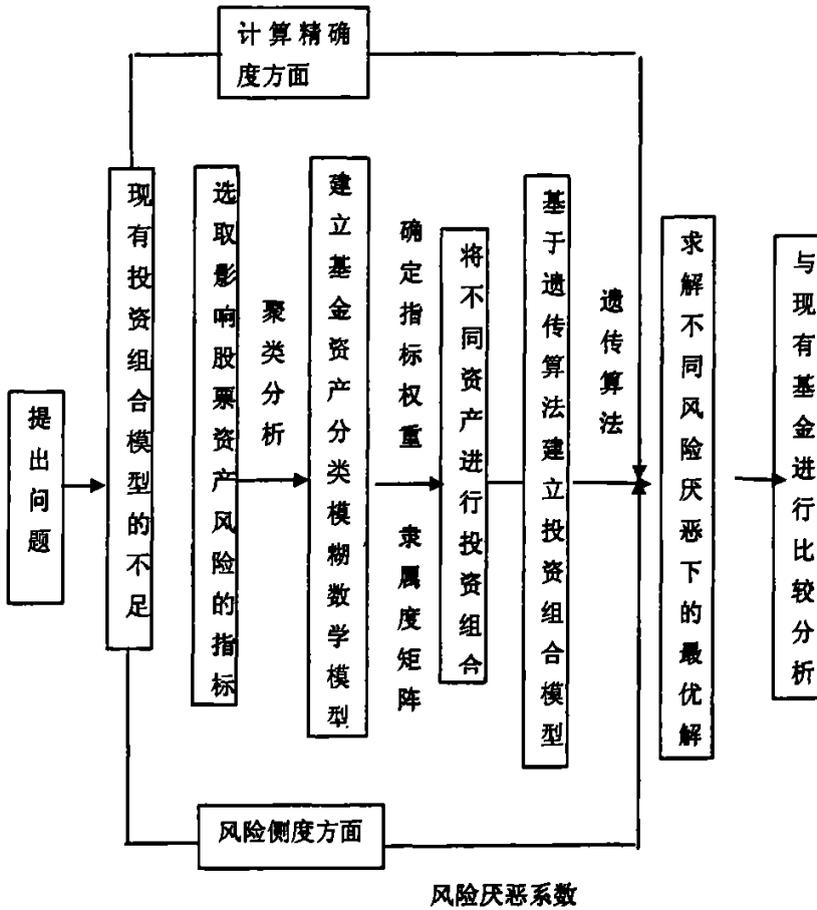


图 1.1 论文的技术路线
Fig. 1.1 Technology line of the paper

1.3.2 论文的创新点

本论文的创新见解表现在两个方面：第一，在投资组合模型中加入风险偏好因素，解决了原有模型只适用于风险中性的不足，取代了投资者风险中性的假设，使投资者可以根据自身的风险偏好程度来选择投资组合。第二，以遗传算法中的目标函数的方式建立了基金投资组合模型，以风险调整的收益的最大值作为遗传算法中的目标函数，应用遗传算法进行了求解。避免了应用二次规划对原有模型求最优解的过程中，出现的非线性程度高，收敛缓慢并且收敛于局部解的问题。

2 基金投资组合模型的理论分析

2.1 基金投资相关概念的界定

2.1.1 基金和基金资产

研究基金投资，有两个重要的概念必须界定清楚，即基金和基金资产。

从资金上讲，基金是用于特定目的并独立核算的资金；从组织上讲，基金是为特定目标而专门管理和运作资金的机构或组织，例如各种基金会。本文对基金的定义为：专门投资于资本市场上的并且独立核算的资金。

基金的资产是指基金拥有的所有资产的价值，包括现金、股票、债券、银行存款和其他有价证券。本文研究主要是针对中国的基金市场，因此将基金资产定义为，基金投资的所有金融产品，在中国主要包括债券和中国 A 股市场，所以基金资产包括国债资产和股票资产。

股票是一种有价证券，是股份公司为筹集资金发给出资人作为公司资本部分所有权的凭证，成为股东以此获得股息和红利。其中，股票还分为优先股和普通股，优先股是介于固定收益和可变收益证券之间的一种混合类型，其索取权仅局限于某一特定的金额。普通股则是公司在外发行的一般性股票，它是对公司的盈利具有剩余索取权的。

债券是具有固定收益的证券。债券的种类有很多种，最主要的还是政府债券和公司债券两种，由于政府可以印制钞票，所以，它发行的证券通常也认为是没有风险的，而公司债券可以通过很多种方式来进行发售，公司债券由于公司的经营状况的不安定因素导致了公司债券具有一定的风险，现在，公司债券的质量评级主要由两家权威的评价机构（标准·普尔公司和穆迪公司）来进行，其评价的依据主要是发行公司的财务状况和其他的相关因素，如发行公司所属行业的性质及其在本行业中的地位^[23]。

2.1.2 基金投资组合

基金投资组合，是将债券和股票等金融产品同时进行投资，通过股票和债券的收益来得到收益，同时也承担了股票和债券的非系统风险。将相关性不同的资产放到一个投资空间内，利用不同资产之间的相关性不同，来降低投资的非系统风险，达到降低风险，增加收益的目的，这样的资产组成就是一个投资组合。投资组合的形式很多，但是主要还是利用资产之间的相关性差异，来达到降低风险，提高收益的作用，当然，单独投资一个金融产品的风险要远远的高于投资组合。由于投资者的理性假设的存在，投资者更

多的追求的是投资组合的安全性，所以投资组合才会得到投资者的认同，投资组合模型才会不断地发展。

基金投资组合与证券投资组合的不同之处在于，证券投资组合可以将所有的资产全部投资于一类，或者多类股票，可以持有也可以不持有债券（无风险资产），而基金投资组合必须持有一定的债券资产，来达到其基金业绩平稳增长的目的。所以，从收益上看，基金投资组合的收益不如证券投资组合，而风险远远比证券投资组合低。

2.2 现有的投资组合模型的不足

现有投资组合模型主要包括三类投资组合模型，即“均值一方差组合模型”、CAPM 资本资产定价模型、Sharpe 的单指数模型，这些模型虽然应用较广，但受现实条件的约束，还存在着两个方面的不足，本文将从这两个方面加以剖析。

2.2.1 计算方法精确度方面

现有的投资组合模型都是线形组合，都是将多元函数转化为一元函数，应用二次规划方法来求解，在求解过程中，容易出现一个方程的解与另一个方程的未知量相关，所以在求解线形方程的时候，解向量会出现多重共线性，这样最后的结果不准确，并且很容易得不到解，所以现在通过其他的方法来对模型进行求解就很有必要。

2.2.2 投资组合模型本身方面

除马科维茨理论不允许买空和卖空的假设与中国当前的金融证券市场的情况比较吻合外，投资组合理论与我国证券市场投资者组合投资实践尚存在众多的问题。

(1) 风险分散方式问题。现代证券投资组合理论的风险分散方式虽然也能够得到一个最优结果，但这种最优结果仅仅是由投资数量结构调整所产生，并非是由改进风险的收益和成本所决定，现有的模型中，都是应用收益率和方差来衡量投资组合的收益和风险大小，但是这样的判别标准并不是对投资组合绩效衡量的很好标准，如果一个资产的收益较大，同时风险也较大时，那么通过风险调整后的收益就不一定比收益水平一般，但是风险水平极低的资产好。

(2) 风险的测度问题。Markowitz 均值——方差模型是效用函数的特例，只有在证券收益率服从正态分布条件下，方差才是风险的有效测度。事实上，投资者对风险、收益的理解不对称，更谈不上均匀分布在均值左右，并不一定服从正态分布。

(3) 模型参数估计时效性问题。现实证券市场，证券收益具有非常强的时效性，这就要求证券投资决策方法也具有时变特性，而 Markowitz 的均值——方差模型中各参

数进行估计时,要求样本长度足够长,而样本长度过长会导致模型参数不能充分反映证券收益率的最新变化情况,因而它的时效性较差^[27]。

(4) 交易费用问题。Markowitz 模型没有考虑证券组合投资过程中的交易费用。在证券组合投资过程中,忽略交易费用的证券会导致非有效的证券组合投资。另外,该模型还假定投资者在作决策时仅持有有一定数量的资本金,而没有持有任何证券,在实际进行组合投资决策时,投资者往往已经持有有一定数量的证券,投资者进行投资决策,就是重新调整各风险证券的持有量。

(5) 市场有效性问题。自从 CAPM 模型建立起之后,无论从实证角度,还是从理论分析,市场的有效性都受到了很大的挑战和怀疑。只有当股票市场上股票价格能够及时且不偏不倚地充分反映市场上的所有信息时,市场才是有效的。由于市场本身可能存在失灵的现象,完全有效的股票市场是一种理想境界。但是在中国,上市公司信息披露存在着大量的虚假性、不充分性和不及时性,使得股票市场的有效性更低^[28]。

(6) 实际问题。现有的投资组合模型在计算上过程太过繁琐,出现误差的机会较大,不能真实反映现实情况;将收益率的期望值和标准差作为收益和风险的代表,真实情况显然会与这一假设有所不同;要求利用股票的历史数据求出其期望收益,标准差及相关系数,但未来与历史并不相同,用过去的数据来预测和判断未来显然是不够准确的;模型的假定情况是假定股票市场是均衡的,所有投资者对股票的预期都相同,这不符和实际;此外,该理论否定了专业人士识别证券的能力,排除了投资者比市场干的更好的可能性^[29]。并且,需要知道投资资产的全部信息,才能做出准确的判断,可是不可能知道投资资产的全部信息,所以,在实际问题中,投资组合模型并不能真正发挥降低投资组合风险,提高投资组合收益的效果。

2.3 应用遗传算法在基金投资组合中的优势

遗传算法(GA, genetic algorithm)是一种采用“适者生存”仿自然法则,基于种群(population)设计,并加入自然选择、交叉及变异操作的随机优化与搜索方法。它具有高度的鲁棒性、并行性及全局优化性等鲜明特征^[30]。其选择、交叉、变异运算均以一种概率方式进行,故增加了搜索过程的灵活性,且随着进化过程的进行,新种群中总会更多的产生出优良个体,如果在每代进化过程中保留最优解,则对于全局最优化问题,随着进化代数趋于无穷,遗传算法以概率 1 收敛于全局最优解^[31]。

2.3.1 用于求解投资组合最优问题

将遗传算法用于本文的投资组合问题是有效的，且由于遗传算法对函数有非常强的适应性，因此，如采用二次规划等较复杂的规划来研究证券投资基金模型，由于传统方法易陷入局部解，遗传算法擅长全局搜索寻优的优越性将更加突出^[32]。

(1) 遗传算法从问题解的中集开始搜索，而不是从单个解开始。这是遗传算法与传统优化算法的极大区别。传统优化算法是从单个初始值迭代求最优解的；容易误入局部最优解。遗传算法从初始染色体开始搜索，覆盖面大，利于全局择优。

(2) 遗传算法求解时使用特定问题的信息极少，容易形成通用算法程序。由于遗传算法使用适应值这一信息进行搜索，并不需要问题导数等与问题直接相关的信息。遗传算法只需适应值和编码等通用信息，故几乎可处理任何问题。

(3) 遗传算法有极强的容错能力。遗传算法的初始串集本身就带有大量与最优解甚远的信息；通过选择、交叉、变异操作能迅速排除与最优解相差极大的串；这是一个强烈的滤波过程；并且是一个并行滤波机制。故而，遗传算法有很高的容错能力。

(4) 遗传算法中的选择、交叉和变异都是随机操作，而不是确定的精确规则。这说明遗传算法是采用随机方法进行最优解搜索，选择体现了向最优解逼近，交叉体现了最优解的产生，变异体现了全局最优解的覆盖。

2.3.2 用于优化投资组合模型

通过研究遗传的特点及其特性，将遗传算法应用于优化投资组合模型也是可行的，主要体现在以下几点^[33]：

(1) 不必非常明确描述问题的全部特征，通用性和鲁棒性强，能很快适应问题和环境的变化；对领域知识依赖程度低，不受搜索空间限制性假设的约束，不必要求连续性、可导或单峰等；

(2) 从多点进行搜索，如同在搜索空间上覆盖的一张网，搜索的全局性强，不易陷入局部最优；

(3) 具有隐并行性，非常适合于并行计算。

综上所述，与传统优化方法相比，遗传算法具有很强的稳健性和很高的效率，以全局并行搜索方式来搜索优化群体中的最优个体，可求得满足要求的最优解。因此，本文采用遗传算法来建立基金市场投资组合模型。

3 基于遗传算法的基金投资组合模型建立

通过对基金资产配置原则的分析,我们可以看出,基金经理通过资产的重新配置来调整投资组合的 β 值以达到风险与收益的最优匹配^[34]。这一过程可以有两个具体措施,即资产调整和证券调整。前者是指调整组合中风险资产的比例(股票持仓比例)以达到调整整个组合 β 值的目的;后者是指通过不同 β 值的组合外证券与组合内证券之间的替换,从而以调整组合中单只股票 β 值大小的方式达到调整整个组合 β 值的目的。具体而言,当预测市场将下跌时,基金经理既可以减少组合中风险资产的持有比例,也可以从组合中调出系统性风险大的股票同时调入系统性风险小的股票;当预测下期市场将上涨时,基金经理则既可以加大风险资产的持有比例,又可以从组合中调出系统性风险小的股票而换入系统性风险大的股票。

3.1 基于遗传算法的基金投资组合模型框架

建立基于遗传算法的基金投资组合模型,首先要对投资的资产进行分析,分析资产产生风险的因素,通过因素分析来提取指标,利用聚类分析的方法对指标进行筛选,建立股票资产风险指标体系,建立指标的权重集,客观地分析各指标对资产产生风险的影响程度,建立隶属度矩阵,通过隶属度的值来判断资产的类别。完成了基金资产的分类,本论文将基金资产共分为3类,无风险资产、中风险资产和高风险资产。按照这三种资产来进行投资组合。将三种资产作为基因值,将三种资产中的国债和股票作为基因中的染色体,引入风险偏好系数,将风险偏好系数和超额风险系数联系,通过对投资资产的选择来判断投资者的风险偏好系数,建立基于遗传算法的基金投资组合模型,应用风险调整后的收益,作为投资资产的收益值,用投资组合的整体风险调整后的收益值作为衡量投资组合绩效的值,最大的值作为目标函数,应用遗传算法来对模型进行求解,通过选择,交叉和变异,最终求得最优投资比例。将风险偏好系数加入到投资组合模型中,求得具有风险偏好系数的最优投资比例。确定三种资产的最优投资比例。建立基于遗传算法的投资组合模型的框架如图3.1所示:

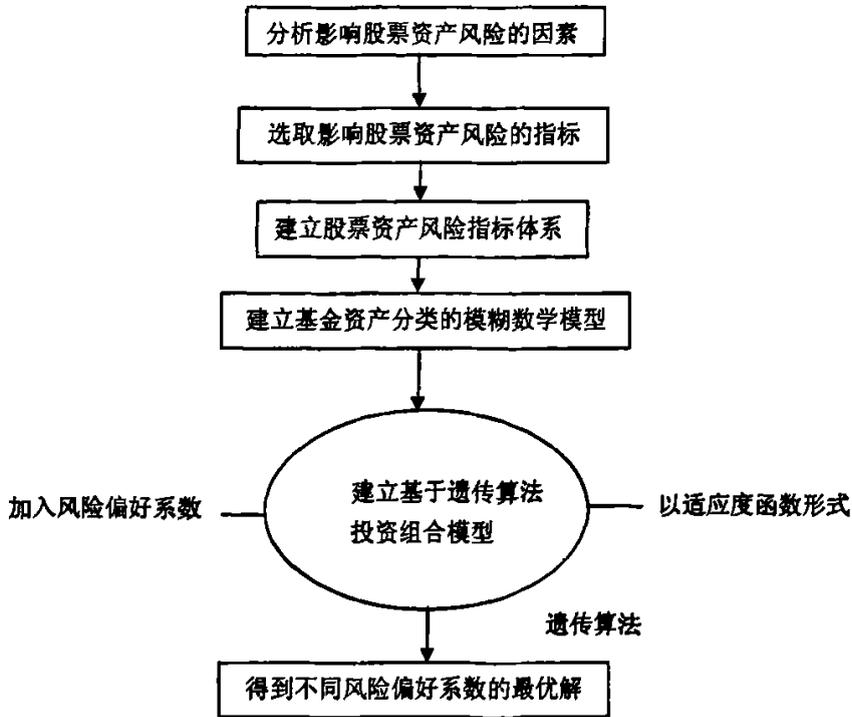


图 3.1 建立基于遗传算法的投资组合模型的框架

Tab 3.1 The frame of building the investment combination model based on GA

3.2 确立股票资产的分类型

股票资产是基金资产中的一部份，对基金资产的增长绝大部分来自于其股票资产的增长，当然，基金资产的风险也就大多数来自于股票资产收益的波动。所以，在进行投资组合之前，对基金资产中的股票资产从风险角度来分类，将会对投资产生的风险有了初步的判断。投资管理之所以成为高度专业化的职业或工作，其根本原因在于投资风险的存在，即投资收益的不确定性。它是投资者不想要的客观存在，也是投资管理的关键所在。投资风险分析是研究和掌握资产的风险特征，即在投资期内投资各资产可能面临的风险类型、风险大小以及产生风险的原因等。

3.2.1 股票资产风险因素分析

证券投资过程中由于宏观、产业、公司、市场等因素的变化，导致了不确定性，从而给证券价格和公司利润带来了波动，产生了证券投资风险^[34]。

第一，宏观因素。主要来源于社会、政治、经济、文化等方面。

(1) 宏观经济因素。即宏观经济环境状况及其变动导致证券投资风险，包括宏观经济运行的周期性波动等规律性因素和政府实施的经济政策等政策性因素。证券市场是整个市场体系的重要组成部分，上市公司是宏观经济运行的微观基础中的重要主体，因此，证券的价格和公司的利润当然会随宏观经济运行状况的变动而变动，会因宏观经济政策的调整而调整。

(2) 政治因素。一国的政局是否稳定直接影响着一国的证券市场的稳定，一般的，政局稳定则证券市场稳定；相反，则证券市场动荡。由于财务的杠杆效应以及全球证券市场的联动性，一个国家证券市场的剧烈动荡，很有可能带动别的国家金融市场的变动。

(3) 法律因素。一国的法律健全程度也可以影响一国的证券市场变动。一国的法律不健全，则证券市场更具有投机性，震荡剧烈，不正当的交易较多；相反地，法律法规体系比较完善，制度和监管机制比较健全的证券市场，证券从业人员营私舞弊的机会较少，证券市场受到人为操作的情况较少，因而表现的相对稳定和正常。总的来说，新兴的证券市场往往不够规范，而成熟的市场法律法规则比较完善。

(4) 军事因素。一般指军事冲突，军事冲突，小则，造成一个国家内部的或一个地区的社会经济生活动荡，大则，打破正常的国际秩序，它使证券市场的正常交易遭到破坏，因而必然导致相关的证券市场的剧烈动荡。

(5) 文化、自然因素。就文化而言，一个国家的文化传统在很大的程度上决定人们的边际储蓄倾向以及人们的风险喜好，喜欢冒险的国家，人们的风险喜好更倾向于风险喜好型，而保守的国家的人们更倾向于风险中性或者风险厌恶。如果一个国家的人口素质普遍偏低，投资更多的依靠感性认识，则一个的证券市场的波动性更强，证券市场会经常发生暴涨暴跌的现象；若投资者的抗风险的能力较差，则市场会经常地发生波动，不利用证券市场的发展；如果一国遭受自然灾害，则一个国家的正常经济生产生活都会受到影响，证券市场同样会受到一定的影响，比如 2003 年我国的“非典”，就对我国的证券市场影响较大。

第二，产业和区域因素。主要指产业发展的前景和区域经济的发展情况对上市公司的生产经营产生影响，从而影响证券价格的波动。产业方面，每个行业的发展都要经历由成长到衰退的过程，一个产业的生命周期通常分为四个阶段，即创业期、成长期、稳定期、衰退期，处于不同环节的产业在经营状况和发展前景方面有很大的差异，这必然会反映在公司利润和证券价格的波动上，从而给证券投资带来风险。在区域方面，由于区域经济的发展状况、信息沟通程度，区域内的投资活跃程度不同，分属各区域的企业生产经营产生差异，产生投资风险。

第三，公司因素。由于公司的外部经营环境和调价以及内部经营管理方面的问题造成公司的收入的变动而引起的证券投资者收益的不确定性。即上市公司的行业因素、业务因素和财务因素对证券投资产生风险。上市公司是股票的发行者，也是募集资金的使用者，上市公司的经营状况的好坏直接影响了股票的价格，上市公司的管理水平，研发水平，在行业中的竞争地位高低，财务状况，等一系列因素都可以影响上市公司的股票价格，这些因素都可以作为产生投资风险的因素。还有上市公司的信息不对称也会带来投资风险。经营风险主要来自公司本身的管理水平、技术与研发能力、经营方向、产品结构等内在因素。

第四，市场因素。影响证券市场的各种证券市场操作。不规范的市场操作很多。虚假重组风起云涌。近几年，股市里重组屡见不鲜，甚至还有什么“重组板块”。其中很多重组是虚假重组。被重组公司通过重组可以免于摘牌或获得配股资格，重组公司则可以借壳上市、扩大知名度、通过配股从市场上获取大量资金；庄家操纵投机盛行。庄家操纵股市一直以来就是广大投资者对于中国证券市场反映最多的问题之一，即使是现在仍有人把中国股市称为“消息股”，言下之意只要拥有及时的消息，没有炒不好的股。很多庄家可以提前获得内幕消息，并利用手中资金炒作，最终从中获利；对中小投资者权益的侵蚀和国有资产的流失。2000年中国股市情形一片大好，可是问问身边的朋友，却少有从中获利的。为什么会这样？显然，中小投资者的利益在某种意义上受到侵蚀。这种侵犯主要来自四个方面：上市公司董事和大股东的侵犯、信息披露不规范对投资者的损害、证券公司违规经营或破产的影响以及交易所系统出错和监管不严带来的损害等。

第五，投资者操作因素。在同一个市场上，有的投资者赚钱，有的投资者赔钱，可能不同的投资者投资同一个股票，结果也不同，产生这样结果的原因有很多，主要还是投资者自身操作产生的风险，不同的投资者由于本身的投资技巧、心理素质、投资喜好以及投资风格，对证券市场的判断标准都会产生投资风险。

3.2.2 股票资产风险指标的选取

基金资产按照风险分类可以分为无风险资产、中风险资产和高风险资产，但是光光从字面上来进行分类，其理论意义并不大，本节将从基金资产中的股票入手。股价的变动导致了股票风险的产生，从影响股价的因素着眼，分析产生股票风险的重要的影响因素，根据股票风险的影响因素，利用聚类分析的方法，对股票风险产生因素进行分类，建立了股票风险指标体系，通过权重集和隶属度矩阵分析，建立模糊隶属度矩阵，通过与标准隶属度矩阵的对比，对股票按照风险进行了分类，这种分类对后面的基于遗传算法的投资组合模型的构建，为基因值的选取建立基础。

建立一套科学、客观、实用的股票分类选择指标，对于识别股票风险，衡量股票风格，确定基金风险资产比例都有极大的作用。基于前文对股票风险影响因素的分析，我们得到建立股票风险分类指标选择的原则：

灵敏性原则。评价指标的灵敏性是决定股票分类选择分析准确性的重要前提。灵敏的评价指标可以反映企业经营活动的偏差信息，对不同企业经营活动的差异性做出灵活的反映，从而为正确分析、确定企业产融结合模式提供依据。

概括性原则。股票分类选择指标体系不可能囊括所有与股票风险有关的因素，但其应保证所反映的信息量与能够覆盖全部股票风险选择影响因素。在此基础上，采用较少的指标来得到等价的评价效果，一方面可以减少信息收集、加工、分析的工作量，另一方面也可以排除指标过多相互影响评价结果的现象。

代表性原则。指标间存在着一定的可替代性，即指标之间的关系不是相互孤立的，而是相互联系与制约的，且对于某类具体的影响因素，也常常表现出一个指标与几个指标或一组指标与另一组指标反映其特征几乎是等价的特性。因此，可以利用指标间的这种关系，选择具有较强代表性的指标，从而减少工作量、降低误差、提高工作效率。

可操作性原则。评价指标计算所依赖的数据必须是能够从企业的财务报表或其他统计途径中获得的，并保证数据能及时、真实、准确、完整。本文选取的评价指标基本都能通过企业的会计报表所得数据计算出来，且指标含义明确、计算过程简单，易于实际操作。

在股市中，板块是指具有某一共同特征的股票，被称为“XX 板块”或“XX 概念股”，它们因相同的特征而在价格走势中呈现较大的联动性。板块的划分是多样的，如按行业划分，有“钢铁板块”、“水泥板块”等；按地域划分，有“边疆概念股”、“首都概念股”等；按经营状况，有“绩优板块”、“亏损板块”。一家上市公司往往因为其不同题材时归属若干个不同的板块。如四川长虹，既属于“绩优板块”又属于“家电板块”。根据划分范围的大小，板块之间也存在交错关系，“生物制药板块”中的个股，理所当然包含在“高科技板块”之列。板块的人为划分，为市场的炒作提供了题材。板块中的股票在少数一两家个股（通常被称为板块的“领头羊”）率先突破的情况下齐涨齐跌，形成炒作效应。本研究将选取每个板块的领头股作为分析的对象，来衡量相关板块的整体风险水平。

按照不同的分析标准可以将股票分为不同的板块，本文的分析对象采用证监会的分类标准。证监会按照行业标准对基金资产中的股票进行分类，分别为：采掘业、初级产品工业、投资品工业、耐用消费品制造、经常性消费品工业、医药卫生、贸易和零售、社会服务业、交通运输和仓储业、金融业、房地产、信息技术业、公用事业和综合类，

共 13 个行业。根据此 13 个行业分类，选取相关的指标对基金资产中的股票进行分类，最终将这 13 类行业划分为中风险行业股票，高风险行业股票，根据这种分类进行投资组合。

影响股价的因素有很多，如前所述，分为宏观因素，产业和区域因素，市场因素，投资者操作因素，以及公司因素，总的来看，可以归纳为宏观因素和微观因素，所以指标的选取将从宏观和微观两个方向来选取。

宏观方面影响股价的因素也有很多，比如一国的利率水平，一国的利率水平会影响到上市公司的资金的成本，影响公司的盈利能力和偿债能力，间接的影响了上市公司的业绩，影响了股票的价格，对股价的波动产生了一定的影响。

微观指标主要选取主要从公司因素和操纵者因素两方面考虑。通过财务指标评价方法来进行评估股票的非系统风险。财务指标评估法的基本原则，是利用企业财务报表和市场价格提供的信息，编制或挑选一套指标体系，反映企业的经营实绩和企业发行的股票的质量，在此基础之上，通过比较和动态的分析，加入评估者的主观判断，确定股票的风险度。财务指标可以是绝对量指标，如反映股票收益量的指标；也可以是相对量指标，如反映股票收益变动水平，偿债能力大小的各类比率等。常用的财务指标一般有三类：第一类是股票盈利能力指标，它包括每股收益、派息率、每股股息、市盈率和投资收益率等；第二类是企业偿债能力指标，包括流动比率、速动比率、现金比率、应收帐款周转率、存货周转率、自有资金比率、负债比率、股东权益对总资产比率、负债对股东权益比率、周转资金比率等；第三类是市场价格趋向指标，其中主要的指标是股票价格指数。

表 3.1 股票风险指标分类

Tab 3.1 the distribution sort of stock risk guide line

股票盈利能力指标	企业偿债能力指标	市场操纵指标	市场价格趋向指标
X1 每股收益	X8 负债比率	X13 日成交额	X16 股票价格指数
X2 每股净资产	X9 净利润	X14 日成交手数	
X3 净资产收益率	X10 现金比率	X15 换手率	
X4 利润增长率	X11 速动比率		
X5 总资产	X12 流动比率		
X6 总股本			
X7 流通股本			

3.2.3 股票资产风险指标的聚类分析

聚类分析法是一种通过判断指标之间的相似程度来筛选指标的统计分析方法。通过聚类分析,可以删除具有较大相关性的指标。聚类分析的过程为,设有 N 个指标,将每个指标看成一类,根据指标间的相似程度,通过比较类间距离进行分类。把距离最小的两类加以合并,此时余下 $N-1$ 类;选择类间距离最小的加以合并。反复进行这一过程,最终形成由小到大的分类系统。最后,根据指标间的相似关系确定指标体系中所包含的指标个数。本文选取编网聚类法对初选指标进行分析,具体步骤如下:

首先,度量指标间的相似程度,运用相关系数度量指标间的相似程度,根据 N 个指标的历史资料,分别计算每两个指标间的相关系数,形成相关系数矩阵 R ,以相关系数矩阵表示指标间的相关关系。

然后,根据编网法对初选指标进行聚类分析,分类中选出最具有代表性的指标。即从每一类中选取指标之间的相关系数平均值较大的评价指标,作为该类指标的代表性指标。回归分析各指标的相关性,做出相关性检验,筛选指标,留取有用指标。对指标进行筛选的聚类分析属于变量的聚类,即对所选样本的变量进行归类,使得每一类指标能代表样本某一方面的特征。对变量进行的聚类分析属于 R 型聚类分析。

在 R 型聚类分析中,使用何种聚类标准以及用那种方法计算类间的距离是两大重要的因素。本文选用适于进行 R 型聚类的皮尔逊相关性(Pearson correlation)度量指标间的相似性,选用组间连接距离法(Between-groups linkage)测量类间距离,组间连接距离法使得合并两类后,不同类间的平均距离达到最小。皮尔逊相关系数计算见式(3.1):

$$r_{ij} = \frac{\sum_{k=1}^n (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)}{n-1} \quad (3.1)$$

其中, $i=1, 2, \dots, n$; $j=1, 2, \dots, n$; n 为样本指标的个数。

根据从上证 180 指数中挑选的 42 只股票的数据,运用 Matlab7.0 进行聚类分析,可以得到各指标的相关系数矩阵,再根据相关系数矩阵进行编网法聚类分析,取 $\alpha = 0.6$ 。

从前所诉得知,16 个指标是按照 4 种因素进行选取的,可以分为以下几类:
 $\{X1, X2\}, \{X3, X4, X11\}, \{X5, X6, X7\}, \{X8\}, \{X9, X10\}, \{X11, X12\}, \{X13, X14\}, \{X15\}, \{X16\}$
 由于聚类分析中聚为一类的指标具有较大的相关性,因此选择同类中的任何一个指标都可以作为该类的代表指标。本文为了提高筛选指标所反映的信息度,根据相关系数均值较大的指标可以代表其他指标较多的信息,每类中选择相关系数最大的指标作为该类的典型代表指标,删除其余指标。

表 3.2 股票风险指标相关系数矩阵
Tab.3.2 Correlation Matrix of stock risk indexes

	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16
X1	0.483	1															
X2	0.361	0.501	1														
X3	0.354	0.831	0.141	1													
X4	0.238	0.25	-0.075	0.182	1												
X5	-0.028	0.03	0.312	0.001	-0.1	1											
X6	-0.168	0.04	0.08	-0.81	-0.103	0.825	1										
X7	-0.135	0.25	0.209	0.088	-0.861	0.75	0.65	1									
X8	0.001	-0.05	-0.136	0.002	0.04	0.07	-0.126	-0.134	1								
X9	0.143	0.502	0.38	0.245	-0.023	0.76	0.42	0.746	-0.076	1							
X10	0.06	0.361	0.382	0.105	-0.041	0.868	0.61	0.67	0.178	0.720	1						
X11	0.12	0.475	0.42	0.153	-0.046	0.887	0.59	0.811	0.026	0.917	0.900	1					
X12	0.163	0.116	-0.1	0.065	0.69	0.103	0.81	0.059	-0.070	0.050	-0.051	-0.052	1				
X13	0.142	0.208	0.12	0.127	-0.016	0.374	0.378	0.671	-0.140	0.525	0.300	0.400	-0.001	1			
X14	0.021	0.161	0.066	0.1	-0.034	0.369	0.41	0.68	-0.123	0.478	0.278	0.470	-0.010	0.99	1		
X15	0.267	0.323	0.124	0.16	0.104	0.064	0.08	0.156	-0.016	0.124	0.090	0.088	0.080	0.59	0.58	1	
X16	0.324	0.412	0.102	0.002	0.001	0.365	0.411	0.32	0.001	0.007	0.314	0.114	0.213	0.781	0.756	-0.62	1

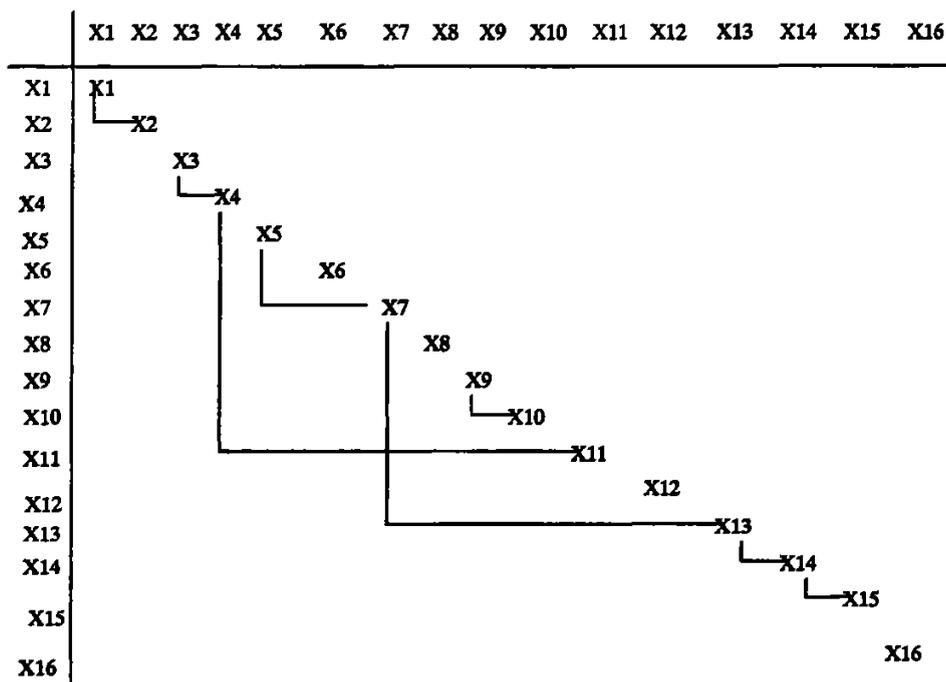


图 3.2 编网法聚类分析
Fig. 3.2 Netting clustering analysis

X1、X2 为同一类指标，两个指标具有较大的相似性，相关系数较大，从相关系数矩阵中可以看到，X1 的相关系数均值较大，根据前面所诉的原则选择保留相关系数均值较大的 X1，同理，X3、X4、X11 三个指标中，X3 的相关系数均值较大，保留 X3，X5、X6、X7 中，选取 X6 作为衡量指标，X9、X10 中 X9 的相关系数均值更大，保留 X9 作为衡量指标，X11、X12 中选取 X12 作为保留指标，X13 和 X14 具有明显相关性，根据原则，选取 X13 作为衡量指标。

3.2.4 股票资产分类指标体系的确定

根据聚类分析的结果得知，从原来的 16 个指标中剔出了 7 个指标，剩下的 9 个指标是：X1 每股收益，X3 净资产收益率，X6 总股本，X8 负债比率，X9 净利润，X12 流动比率，X13 日成交额，X15 换手率，X16 股票价格指数。从指标体系的图形中可以

看到，指标体系共分两层，第一层为准则层，包括股票盈利能力，企业偿债能力，市场操纵，市场价格趋向四个因素，第二层为指标层，由筛选出的 9 个指标构成。

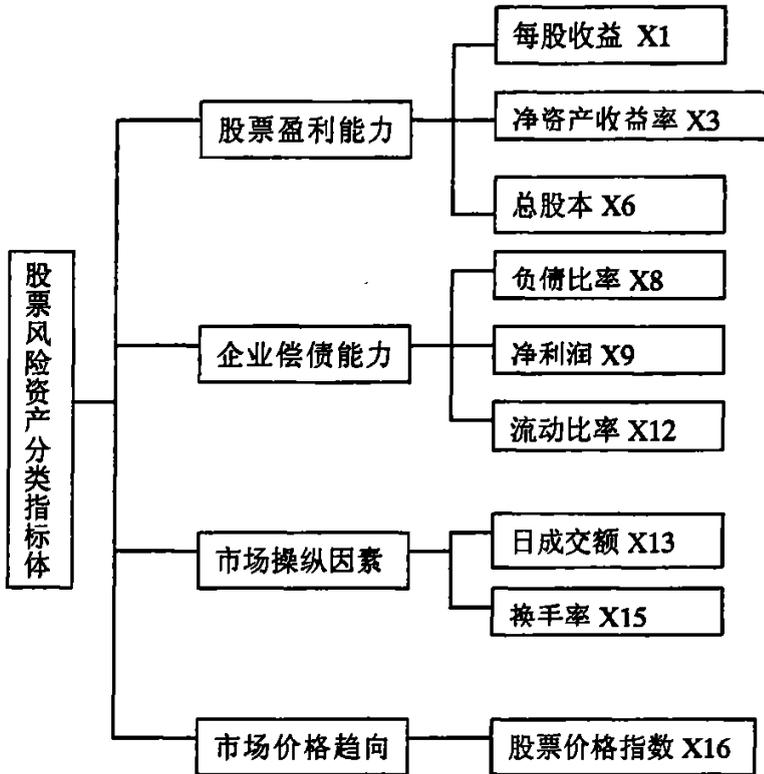


图 3.3 股票风险资产分类指标体系

Fig. 3.3 The index system of stock risk assets

3.3 建立基金资产分类的模糊数学模型

3.3.1 权重集的确定

每一个评价指标对股票的风险影响程度不同，因此，需要根据各指标的需要程度或对综合评价结果的贡献度给予各指标赋予相应的权重，以区分不同指标对整体评价的不同影响。

$$A = (A_1, A_2, A_3, A_4) \text{ 且 } \sum_{i=1}^4 A_i = 1, A_i \geq 0 \quad (3.2)$$

A_1, A_2, A_3, A_4 分别表示股票盈利能力因素, 企业偿债能力因素, 市场操纵因素, 市场价格趋向因素。

$$A_1 = (a_1, a_3, a_6) \text{ 且 } a_1 + a_3 + a_6 = 1, a_i \geq 0 \quad (3.3)$$

$$A_2 = (a_8, a_9, a_{12}) \text{ 且 } a_8 + a_9 + a_{12} = 1, a_i \geq 0 \quad (3.4)$$

$$A_3 = (a_{13}, a_{15}) \text{ 且 } \sum_{i=2}^2 (a_{13} + a_{15}) = 1, a_i \geq 0 \quad (3.5)$$

$$A_4 = (a_{16}), a_i \geq 0 \quad (3.6)$$

3.3.2 建立隶属度矩阵及评语集

在自然科学、社会科学和工程技术等领域, 存在着很多模糊或不确定性, 它们概念的内涵和外延是不能用“是”或“不是”能够判断的。而模糊数学以精确数学为基础, 提出了隶属函数理论, 确定了某一事物在多大程度上属于所讲的概念, 或者不属于所讲的概念。在给出的范围内, 某一事物与所讲的概念不一定只有“是”或“不是”的关系, 而且还存在着中间状态, 用介于“0”和“1”之间的实数来表示其隶属的程度。这样描述模糊性问题比精确数学更为合理。同样, 对股票按照风险分类也属于一种模糊现象, 投资者到底应选择什么样子的股票呢, 用精确数学“是”或“不是”的概念很难做出判断。应用模糊数学对股票分类做出综合评判, 属于模糊决策的范畴。模糊综合评判问题又称为模糊综合决策问题, 它解决的问题是在考虑多种因素的影响下对某种事物做出综合决策。

选取相关指标值, 作为指标矩阵 $C_{m \times n}$

$$C_{m \times n} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix} = (c_{ij}),$$

$$i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n \quad (3.7)$$

其中： c_{ij} ----各项指标值

根据股票风险质量评价的分级标准， m 项参数的 t 级分级标准可以组成 $m \times t$ 阶风险分级标准矩阵。

$$S_{m \times t} = \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1t} \\ s_{21} & s_{22} & \dots & s_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ s_{m1} & s_{m2} & \dots & s_{mt} \end{bmatrix} = (s_{ij}),$$

$$i = 1, 2, \dots, m, \quad l = 1, 2, \dots, t \quad (3.8)$$

根据矩阵(2)中的分级标准值，按如下公式求出矩阵(1)中各参数指标对股票风险的相对隶属度，得到各参数的相对隶属度矩阵(5)。

对于基本面较好，对股票风险影响较大的参数，隶属度采用：

$$f_{ij} = \begin{cases} 1 & , c_{ij} > s_{it} \\ \frac{c_{ij} - s_{i1}}{s_{it} - s_{i1}} & , s_{i1} < c_{ij} < s_{it} \\ 0 & , c_{ij} < s_{i1} \end{cases} \quad (3.9)$$

对于基本面较好，对股票风险影响较大的参数，隶属度公式为：

$$f_{ij} = \begin{cases} 0 & , c_{ij} > s_{i1} \\ \frac{s_{i1} - c_{ij}}{s_{i1} - s_{it}} & , s_{it} \leq c_{ij} \leq s_{i1} \\ 1 & , c_{ij} < s_{it} \end{cases} \quad (3.10)$$

于是有各样本风险值隶属度矩阵

$$F_{m \times n} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ & & \cdots & \\ f_{m1} & f_{m2} & \cdots & f_{mn} \end{bmatrix} \quad (3.11)$$

另一方面,还应考虑到 m 项参数指标对股票风险大小的影响程度不完全相同,应根据对风险影响程度的大小赋予相应的权重, m 项指标的权向量为:

$$\vec{W} = (w_1', w_2', \cdots, w_m'), \quad \sum_{i=1}^m w_i' = 1.0.$$

综合考虑两类权重,即超标权重与指标权重,则 n 个样本 m 项参数指标的综合权重矩阵为:

$$A = \begin{bmatrix} w_1' & & & \\ & w_2' & & \\ & & \ddots & \\ & & & w_m' \end{bmatrix} \cdot \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ & & \cdots & \\ f_{m1} & f_{m2} & \cdots & f_{mn} \end{bmatrix} = \begin{bmatrix} w_1' f_{11} & w_1' f_{12} & \cdots & w_1' f_{1n} \\ w_2' f_{21} & w_2' f_{22} & \cdots & w_2' f_{2n} \\ & & \cdots & \\ w_m' f_{m1} & w_m' f_{m2} & \cdots & w_m' f_{mn} \end{bmatrix} \quad (3.12)$$

将矩阵(3.12)作归一化处理,将其压缩在 $[0, 1]$ 区间,得归一化的指标综合权重矩阵(3.13)

$$W_{m \times n} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ & & \cdots & \\ w_{m1} & w_{m2} & \cdots & w_{mn} \end{bmatrix} = (w_{ij}),$$

$$i = 1, 2, \cdots, m, \quad j = 1, 2, \cdots, n. \quad (3.13)$$

w_{ij} 为样本 j 的指标 i 的综合权重,表达式为:

$$w_{ij} = \frac{w_i' f_{ij}}{\sum_{i=1}^m w_i' f_{ij}}, \quad \sum_{i=1}^m w_{ij} = 1.0. \quad (3.14)$$

同样也可根据公式(3.15)把分级标准矩阵转化为分级标准模糊矩阵(3.16)

$$p_{il} = \frac{s_{il} - s_{i1}}{s_{il} - s_{i1}}, \quad p_{il} \in [0, 1], \quad (3.15)$$

$$P_{m \times t} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1t} \\ p_{21} & p_{22} & \cdots & p_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mt} \end{bmatrix} = (p_{il}),$$

$$i = 1, 2, \dots, n, \quad l = 1, 2, \dots, t \quad (3.16)$$

矩阵(3.16)的意义也可表达为分级标准的标准权重矩阵,第j个采样样本的超标权重

向量表示为: $\vec{f}_j = (f_{1j}, f_{2j}, \dots, f_{mj})^T.$ (3.17)

第l级股票分级标准的标准权重向量表示为:

$$\vec{p}_l = (p_{1l}, p_{2l}, \dots, p_{ml})^T. \quad (3.18)$$

设第j个样本隶属于l级标准的隶属度为 u_{jl} , 则n个样本对t个级别风险评价标准的隶属度以模糊矩阵(3.19)来表示:

$$U_{n \times t} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1t} \\ u_{21} & u_{22} & \cdots & u_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nt} \end{bmatrix} = (u_{jl}),$$

$$j = 1, 2, \dots, n, \quad l = 1, 2, \dots, t. \quad (3.19)$$

约束条件为: $\sum_{l=1}^t u_{jl} = 1.0$, 且 $u_{jl} \geq 0$.

用线性规划法求得第 j 个样本隶属于 l 级风险的隶属度的计算公式为:

$$u_{jl} = \frac{1}{\sum_{i=1}^m (w_{ij} |f_{ij} - p_{il}|)} \cdot \frac{1}{\sum_{k=1}^t \frac{1}{\sum_{i=1}^m (w_{ij} |f_{ij} - p_{ik}|)}} \quad (3.20)$$

本文称公式(3.20)为股票风险评定综合模糊优化模型。根据它可以计算确定最优分级矩阵的每一个元素值。优选出隶属度矩阵 u_{nl} 后便可以从每行中挑选出一个最大值, 其对应级别便为该行所对应样本的风险级别。最后按各测样本点所代表的股本大小来加权平均, 便得出股票风险综合评价级别。

评语集是对被评价对象各种评价的集合。设股票资产选择评语集为 $V, V=\{V1, V2\}$, 其中 m 为评语的个数。根据前文中介绍股票资产的划分标准, 将股票资产划分为中风险资产和高风险资产两种, 分别用 $V1$ 和 $V2$ 表示, 即 $V=\{\text{中风险资产}, \text{高风险资产}\}$ 。

(1) 中风险资产: 主要指基金投资对象中的股票资产, 这类资产的收益调整后的风险值在整体样本加权平均值之下。此时, $u_{jl} < 0.5$

(2) 高风险资产: 主要指基金投资对象中的股票资产, 这类资产的收益调整后的风险值在整体样本加权平均值之上。此时, $u_{jl} > 0.5$

在模糊集合中, 指标体系中的各指标与股票风险之间的关系用隶属度 U 来表示。隶属度 $U(x)$ 越接近于 1, 表示 x 属于 $V2$ 的程度越高, $U(x)$ 越接近于 0 表示 x 属于 $V1$ 的程度越高。

3.3.3 基金资产的分類

根据前文所述，将基金资产分为了国债资产和股票资产，通过 3.3.2 对股票资产的进一步分类，最终根据收益调整的风险大小将基金资产分为以下三类：

(1) 无风险/低风险资产，只要指那些有固定的或基本的投资收益的资产。例如，可自由流通或交易的国库券、级别高的政府或企业的债券、银行存款等由于这类资产没有违约风险，利率固定且事先已知，流动性风险不能说没有，但是非常的低。对这类资产值得注意的是利率风险，当市场或银行利率不稳定时，投资这类资产仍然有一定的风险，特别是短期的风险比较大。但是从长期来看，风险较小，收益固定。

(2) 中风险资产，指主要指基金投资对象中的股票资产，这类资产的收益调整后的风险值在整体样本加权平均值之下。比如那些内需型的股票，股票整体暴涨暴跌现象较少，波动幅度不大，这类股票风险较科技类股票小，但是要比一般的债券高很多。

(3) 高风险资产，主要指基金投资对象中的股票资产，这类资产的收益调整后的风险值在整体样本加权平均值之上。主要是小型高科技公司的(Venture Capital)、垃圾债券、ST 类股票以及衍生证券。这类资产的高风险来自于高科技项目的高失败率、高违约率、面临退市的风险以及较高的财务杠杆效应。而这些资产之所以也在投资组合的范围之内，是因为他们通常具有较高的潜在收益或当前价格很有吸引力。

本论文将基金资产分为无风险资产 (R_f)、中风险资产 (R_L) 和高风险资产 (R_H) 三类，划分这三类资产的依据是，根据每类资产的在一段时期内的标准差大小，计算其收益率，将标准差按照收益率来进行调整，通过收益调整后的标准差大小来对资产进行分类。其中，将收益调整之后的标准差定义为超额风险系数：单位收益下承担超额风险。用 δ 表示， σ 表示资产的风险， R 为资产的收益率。 $\bar{\delta}$ 为 δ 的加权平均值。

$$\delta = \frac{\sigma}{R}, \quad \delta > 0 \quad (3.21)$$

在本文中，无风险资产主要是指基金投资的国债，以及其他在一定时期内，具有固定收益，风险小的投资项目及金融产品。中风险资产特指在基金资产中的股票资产，并且其收益调整后的风险值在整体样本空间的加权平均值以下。高风险资产特指在基金资产中的股票资产，并且其收益调整后的风险值在整体样本空间的加权平均值以上。

3.4 基于遗传算法的投资组合模型的确定

现代投资组合理论的精髓之一，就是利用不同资产收益之间的相关性来降低投资组合的非系统风险，将无风险资产，中风险资产、高风险资产进行投资组合，遵循投资收

益最大化、风险最小化的最优条件，应用遗传算法建立基金的投资组合模型。所以首先要验证资产间的相关性，只有资产的相关性较小，才能将其组合在一起，作为一个投资组合。

3.4.1 基金资产的相关性检验

资产收益的相关性实质上是不同资产在运动变化过程中表现出的协调一致性。如果资产的投资收益是同向变化，那么相关度就高，反之，相关度就低。资产投资收益的相关性取决于两个方面：影响资产收益的因素和这些因素对资产收益的影响方向。如果不同资产的收益各自取决于互不相关的因素，则相关性较低，例如，欧洲的债券市场与美国的股票市场的相关性就较低；如果资产收益受共同因素的影响且影响方向相同，则通常会有较高的相关性，比如，一国的证券市场上的股票，同时受国家政策影响程度较高。

通过前述对基金资产的划分，计算各资产在样本期间的收益率，从而计算三种资产之间的相关系数矩阵。

表 3.3 基金资产相关系数矩阵

Tab. 3.3 The correlation matrix of fund capital

	Rf	RL	RH
Rf	1	0.001	0.003
RL	0.001	1	0.214
RH	0.003	0.214	1

得到从相关系数矩阵可以看出，无风险资产和中风险资产、高风险资产的相关性不显著，中风险资产和高风险资产具有一定的相关性，相关系数为 0.214，从相关系数可以得知，有些因素共同影响了两种资产，比如说，股票价格指数（X16），在一定程度上对两种资产都有影响，有的对一种资产的影响要比另一种的影响更大，比如说，负债比率（X8）对中风险资产的影响就要高于高风险资产。

$$\text{建立回归方程: } F(r) = c_1 + c_2 R_f + c_3 E_{rl} + c_4 E_{rh} + \varepsilon \quad (3.22)$$

其中 $F(r)$ 为被解释变量，其经济学意义为总收益， R_f, E_{rl}, E_{rh} 为解释变量，其经济学意义为每种资产的收益。利用历史数据，对回归方程进行回归分析，得到结果为：

$$Y = 36.39640408 + 0.6459562148 * X1 + 0.4738835939 * X2 - 0.2695909231 * X3 \quad (3.23)$$

DW 值为 1.89554，F 值为 13.45608，R-squared 为 0.5286，分析结果，DW 值 F 值通过检验，拟合值 R 的平方值远小于 1，拟合效果不好，分析原因为数据选取得范围不够广泛。进行调整后，得到新的方程为：

$$Y = 39.458408752 + 0.26569884231 * X1 + 0.4125584744 * X2 - 0.3185475213 * X3 \quad (3.24)$$

调整后的拟合值 R 的平方值为 0.7561, 接近于 1, 通过检验。通过检验发现, 三种资产的相关性不显著, 可以进行分类。

3.4.2 基因值的选取和组合

不同的动物有不同类型的基因, 基因主要构成为染色体, 以人的基因和染色体为例, 人的基因构成为双螺旋结构的基因链, 染色体有 23 对。通过不同的染色体的选择, 交叉, 变异, 最终得到不同的生物种类。中风险资产和高风险资产就好比人的两条基因链, 而蕴含在两种资产中的 13 板块的股票, 就好比是构成基因链的染色体, 通过“选择”, “交叉”, “变异”, 最终得到不同的投资组合, 在这些组合当中只有一个是满足风险最小、收益最大的条件的, 最终是这个投资组合被选择作为最佳的投资组合。

在遗传算法中, 要求解的问题的可行解称为染色体或个体。采用遗传算法求解整数规划问题时, 必须考虑对个体如何编码, 以及在这些编码上进行的遗传, 借鉴对连续变量采用浮点数编码的思想, 我们采用了对整数变量最自然的整数形式编码, 因此, 在进化群体中的每个染色体直接用整数向量 $k=(k1, k2, \dots, kn)$ 来表示。这样可以减少变换编码带来的麻烦, 并且可以很直观的看到结果, 有利求解最优解。

基因值的选取主要是三条“基因链”作为基因(投资组合)的主要构成, 这三条“基因链”分别为无风险资产、中风险资产和高风险资产三部分。无风险资产单独作为一类资产而存在, 选取上证 180 中的 13 个板块的股票作为染色体的基因值选取对象, 选取每只股票的收益风险比作为染色体的基因值, 之后按照整数形式进行编码。初始种群规模为 180 只股票和 3 个月国债, 确定初始种群规模=181。

本文将 $\eta_1, \eta_2, \eta_3, f(r)$ 作为所要寻优的参数, η_1, η_2, η_3 作为投资三种资产的投资比例, 所以 η_1, η_2, η_3 的选取范围为从 0 到 1, 选取 $f(r)$ 作为适应度函数, $\max f(r)$ 作为操作过程中的评价函数。计算每只股票的收益率和收益率的均值的方差, 并且求其比值, 作为遗传算法的操作对象。

3.4.3 建立基金投资组合模型

不同的人对待风险的态度不同, 经济学认定有些人是风险厌恶的, 有些人是风险喜好的, 有些人是风险中性的。如图, 三种不同的风险偏好程度对应着三种不同效用类型的曲线。

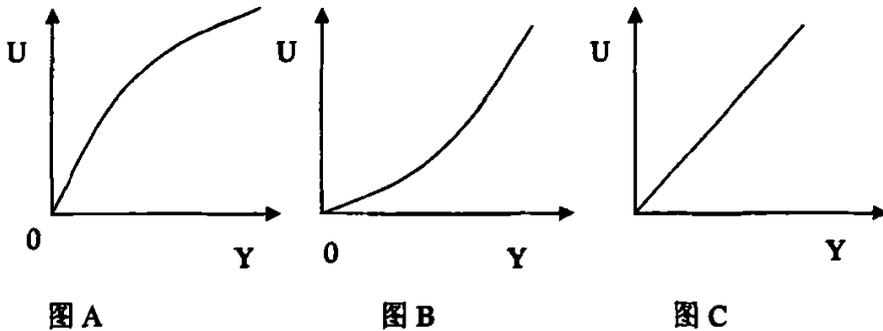


图 3.4 不同风险偏好下的效用曲线

Tab 3.4 The different utility curve in different risk taste

图 3.4 A 表示风险厌恶者的效用曲线， B 表示风险喜好者的效用曲线， C 表示风险中性者的效用曲线。

由图 3.4 A 知，图中曲线的二阶导数为负值，一阶导数递减，即随着收入的增加，效用递减，在预期收入相同的情况下，将选择无风险的投资项目。这类投资者为风险厌恶型。图 B，图中曲线的二阶导数为正，一阶导数递增，效用曲线为这种的投资者，随着收入的增加，效用也同时增加，但是效用的增加幅度要更大，这类投资者是风险喜好型。图 C，图中曲线为一条直线，即随着收入的增加，效用同比例的增加，这类投资者的风险类型为风险中性。

每个投资者的效用曲线是不同的，但是每个投资者的目的是相同的，就是整个投资的效用最大化，每个投资者可以根据本身的效用函数来进行最大化决策，由于每个投资者对待风险的态度不同，在此引入风险偏好系数来衡量每个投资者的风险偏好程度。图 3.5 说明了在不同风险偏好程度下的效用函数。

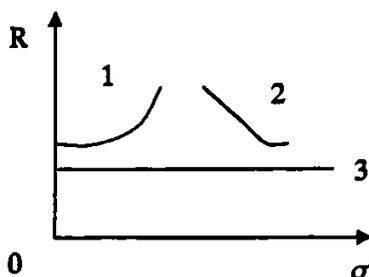


图 3.5 不同风险偏好系数的效用函数性质

Tab 3.5 The character of utility function with different risk taste coefficient

风险偏好系数描述了投资者对不确定性的厌恶程度。设风险偏好系数为 λ ， $\lambda \in [0,1]$ 。

$$\text{设函数 } f(R) = \lambda R(\sigma) - (1 - \lambda)R_f \quad (3.25)$$

$f(R)$ 是无风险收益 R_f 和风险收益 $R(\sigma)$ 的函数，风险偏好系数 λ 和 $1 - \lambda$ 就表示投资于风险资产和无风险资产的比例，当 $f(R, \sigma) = 0$ 时，公式 3.24 就改写成

$$f(R) = \lambda R(\sigma) - (1 - \lambda)R_f = 0 \quad (3.26)$$

求出投资于两种资产的比例。若 λ 大于 $1 - \lambda$ ，则 λ 大于 0.5，此时表明投资于风险资产的比例大于投资于无风险资产的比例，投资更看重高风险带来的高收益；若 λ 小于 $1 - \lambda$ ，则 λ 小于 0.5，此时表明投资于风险资产的比例小于投资于无风险资产的比例，投资更看重投资的风险，风险越小越好；若 λ 等于 $1 - \lambda$ ，则 λ 等于 0.5，此时表明投资于风险资产的比例等于投资于无风险资产的比例。所以，当 $\lambda \in [0, 0.5)$ 时，表示投资者是风险厌恶者，当 $\lambda = 0.5$ 时表示投资者是风险中性者，当 $\lambda \in (0.5, 1)$ 时表示投资者是风险喜好者。文中的风险偏好，风险中性，及风险厌恶的判别标准既是按照此标准进行划分的。

对公式 3.26 变形得：

$$\frac{R}{\sigma} = \frac{1}{1 - \lambda} - 1 \quad \lambda \in [0, 1] \quad (3.27)$$

从图 3.6 中可以清楚地看到这种函数关系：

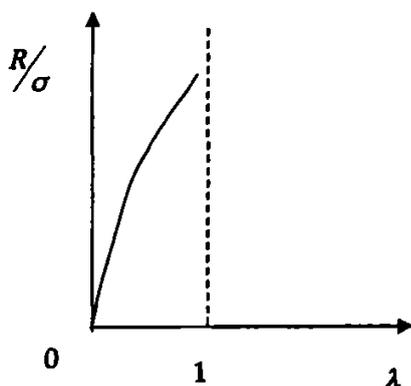


图 3.6 风险偏好系数和风险收益比率之间的关系

Fig.3.6 The relationship between risk taste data coefficient and rate of risk and income

从公式 3.29 可以得到，风险偏好系数和单位风险超额收益率成反比，联立公式 3.21 与公式 3.27 得：

$$\lambda = \frac{\delta}{\delta + 1} = 1 - \frac{1}{\delta + 1} \quad (3.28)$$

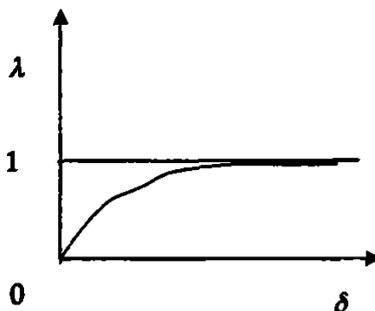


图 3.7 风险偏好系数和单位收益的超额风险系数的关系

Fig.3.7 The relationship between risk taste coefficient and excess risk coefficient

由公式 3.28 可以分析得到这样的结论，风险偏好系数 λ 和超额风险系数 δ 之间呈现正向关系，随着 δ 的增加， λ 也随着增加。通过前面对基金资产分类可以了解到，基金资产的分类是按照资产的超额风险系数与加权平均值作比较，得到的按照超额风险进行分类的基金资产。通过图 3.7 可以看出，随着 δ 的增加，即投资组合中的高风险资产所占比例增加，从图中可以看出，风险偏好系数也在增加，逐渐接近于 1，风险偏好类

型逐渐转为风险喜好，同样，随着 δ 的减少，风险偏好系数也在降低，这样风险偏好类型就逐渐转为风险厌恶。在此，通过超额风险系数可以判断风险偏好系数的大小，一个投资者可以根据自身选取的投资范围，通过计算投资范围内的资产的超额风险系数，通过公式 3.28 来计算自身的风险偏好系数。

本论文的模型的建立主要依据 CAPM 模型以及夏普指数和遗传算法中的适应度函数来建立，其中， $E(r)$ 为投资组合的期望收益，

$$E(r) = \eta_1 R_F + \eta_2 (E_H - R_F) + \eta_3 (E_H - R_F) + \varepsilon_i \quad (3.29)$$

R_F 无风险资产收益 E_H 中风险资产预期收益 E_H 高风险资产预期收益

η_1, η_2, η_3 为投资于三种资产的比例，且 $\eta_1 + \eta_2 + \eta_3 = 1$

$$f(r) = \frac{E(r) - R_F}{\beta} = \frac{(\lambda \eta_1 - 1) R_F + (1 - \lambda)(\eta_2 E_H + \eta_3 E_H)}{\beta} \quad (3.30)$$

方程 $f(r)$ 的经济含义为风险调整后的超额收益，即每承担单位风险而获得的超额收益，以此来衡量投资组合的绩效。 $f(r)$ 为具有风险偏好系数的投资组合模型。简称风险偏好投资组合模型。将风险调整后的超额收益，即 $f(r)$ 作为目标函数作为遗传操作的评估准则。取 $f(r)$ 的最大值为适应值的选取条件。

β 为投资组合的风险大小。并且 $\beta = \frac{\alpha_1 \beta_H + \alpha_2 \beta_H}{\alpha_1 + \alpha_2}$ ，其中 α_1, α_2 为各中风险资产和高风险资产占总资产的权重， β_H, β_H 各自为中风险资产和高风险资产产生的风险值。 λ 为风险偏好系数。

$$\max f(r) = \max \frac{E(r) - R_F}{\beta} = \max \frac{(\lambda \eta_1 - 1) R_F + (1 - \lambda)(\eta_2 E_H + \eta_3 E_H)}{\beta} \quad (3.31)$$

$$s.t. \quad \eta_1 + \eta_2 + \eta_3 = 1$$

$$\eta_{i(i=1,2,3)} \geq 0$$

公式 (3.31) 即为基于遗传算法的基金投资组合模型。由于投资者对待风险的态度不同，投资者在选择投资组合资产和投资组合类型时都会受风险喜好（厌恶）程度这一因素影响，投资者会根据此因素选择投资组合类型，在一定程度上影响了投资组合的收益，所以，在模型中加入了风险厌恶系数这一参数，投资者根据自身风险喜好程度来决定此参数的大小。这其中 $f(r)$ 作为模型的目标函数求其最大值，其经济意义为每承担单位风险而获得的超额回报率，其数值越高，表明承担单位风险而获得的超额收益越高，首先，设定不同的风险厌恶系数，之后利用遗传算法进行选择、交叉和变异，将 $f(r)$ 作

为适应度函数, 最终当求解过程中以 $f(r)$ 最大值作为整个寻优过程的终止条件, 若没有, 将选择在最大代数 1000 时终止程序, 并选择其中次优解作为最优解。本论文采取达到最大代数终止程序的方法, 即是达到最大代数后, 选取 $f(r)$ 最大值作为最优解。

3.4.4 测定投资组合的最优解

由公式 (1) 和 $\eta_1 + \eta_2 + \eta_3 = 1$ 可以得知:

$$\eta_1 = \frac{E(r) - E_{rh} + \eta_2[E_{rh} - E_{rl}]}{R_F - E_{rh}} \quad (3.32)$$

$$\eta_2 = \frac{E(r) - E_{rh} + \eta_1(R_F - E_{rh})}{E_{rh} - E_{rl}} \quad (3.33)$$

$$\eta_3 = 1 - \eta_1 - \eta_2 \quad (3.34)$$

将求解问题的目标函数 $\max f(r) = \frac{E(r) - R_F}{\beta} = \frac{(\lambda\eta_1 - 1)R_F + (1 - \lambda)(\eta_2 E_{rl} + \eta_3 E_{rh})}{\beta}$ 作为遗传操作的评估准则, 求 η_1, η_2, η_3 的具体数值。

(1) 随机确定初始种群规模, 在随机生成初始染色体种群时, 为保证满足约束条件, 设计一个染色体可行性的检验函数。 $\sum_{j=1}^n w_j = 1$ 返回 0; 否则, 返回 1。其中, 检验函数值为 0 表示不可行, 为 1 表示可行。然后, 在初始解群的可行集中, 随机抽取 300 个可行解, 作为初始解群。

(2) 计算当前解的适应度函数值。这里采用方程 3.31 作为目标函数。方程 3.32 作为统计函数, 并通过判断适应度的大小决定某个体进入下一代的概率。由于本文目标函数为求极大值, 因此, 适应度值越大, 进入下一代的概率越大。在此计算 $\lambda = 0.2$ 和 $\lambda = 0.8$ 两个不同风险厌恶系数下的最大适应值, 来分析两种不同风险喜好类型下的最优投资比例。

(3) 选取遗传操作的对象, 并进行遗传操作, 形成下一代群体。简单遗传算法中的遗传操作有选择、交叉和变异三种。在选择时, 本文采用适应值比例选择方法, 用 Holland(1975)的赌轮盘方式实现。对于给定规模为 l 的群体 $P = \{a_1, a_2, \dots, a_l\}$, 首先计算个体 $a_j \in P$ 的适应度值 $f(a_j)$, 然后计算此适应度值在群体适应度值总和中的比重, 表示该个体被选择的概率, 则个体被选择的期望数目为 $P(a_j) = l \times P_r(a_j)$, 其中 $j = 1, 2, \dots, l$ 。本文的选择概率为 0.39。对于交叉操作来说, 首先从交配池中随机抽取要交配的一对个体, 根据串长度 L , 对要交配的一对个体, 随机选取 $[1, L - 1]$ 中的某个位置作为交叉

位, 根据交叉概率 $P_c (0 < P_c \leq 1)$ 实施交叉操作, 配对个体在交叉位置处, 相互交换各自的部分内容(权重值), 从而形成新的一对个体。本文设交叉概率为 0.7。在变异操作中, 按照变异概率 P_m 随机选取要进行变异操作的个体, 然后随机选取变异的位, 并将两个基本点位的权重值进行互换。为防止接近最优的解个体结构被变异操作所破坏, 在问题的求解中, 设定变异概率 $P_m = 0.01$ 。

(4) 重复步骤 2-步骤 4, 直到满足运行的最大代数 1000 ($\max g = 1000$)

(5) 计算 1000 代中适应值最高的代数的相应值。

最终将染色体各向量进行归一化运算, 所得既为各部分资产的投资比例 η_1, η_2, η_3 , 在得到无风险偏好系数下的最优的投资比例之后, 将 $R_f, E_{r1}, E_{r2}, \beta$ 值代入方程 $f(r)$ 中, 得到一个关于风险偏好系数的方程:

$$f(r) = \frac{(\lambda\eta_1 - 1)R_f + (1 - \lambda)(\eta_2 E_{r1} + \eta_3 E_{r2})}{\beta} \quad (3.35)$$

设 $A = \lambda\eta_1 - 1, B = (1 - \lambda)\eta_2, C = (1 - \lambda)\eta_3$, 则方程 3.30 变为:

$$f(r) = A \frac{R_f}{\beta} + B \frac{E_{r1}}{\beta} + C \frac{E_{r2}}{\beta} \quad (3.36)$$

则在不同的风险偏好系数下, 投资组合的投资于三种资产的最优投资比例为 (A, B, C)。

4 基于遗传算法的投资组合模型的应用

通过选择合适的证券或证券组合,在承担相同的投资风险时,获得最高投资收益;或者在获得相同证券投资收益时,承担最低投资风险;或者在一定的风险水平下,追求特定的收益,从而达到最优投资的目的。本实证研究从两种情况,即风险偏好型、风险厌恶型下讨论投资组合的最优解。本实证研究首先从投资者三种不同的风险喜好程度出发,根据风险厌恶系数来选择投资类别,通过不同的风险厌恶系数来确定最优的投资组合比例,最后与现有的基金建仓时的投资组合比例作比较,得到根据遗传算法建立的投资组合模型优于现有的投资组合模型的结论。

4.1 数据的选取与处理

数据选取上证 180 指数,数据选取的原因为,第一,数据规模,容量适中,指数可以很好的反应证券市场的变化,第二,数据时间比较长,选取的数据点足够计量应用,满足样本空间的最小要求。

无风险利率的确定在国外研究中,常以一年期的短期国债利率或银行同业拆借利率来代替无风险利率。但是我国目前利率还没有市场化,且国债以长期品种为多,因此无法用国债利率来代表无风险利率。在本论文中,决定采用三个月的定期储蓄存款利率为无风险利率。股票的收益率选取 2006 年 6 月 1 日到 2006 年 9 月 1 日三个月的数据,通过这三个月的收益率可以计算出这三个月中,每只股票的标准差,用于应用遗传算法计算投资组合最优投资比例时候作为染色体的基因值。

4.2 计算基金投资组合模型的最优解

根据遗传算法的步骤,首先,选取初始种群,之后选择,交叉和变异。最后得到最大的适应值,以最大适应值作为最优解。

(1) 随机确定初始种群规模,在随机生成初始染色体种群时,为保证满足约束条件,设计一个染色体可行性的检验函数。 $\sum_{j=1}^n w_j = 1$ 返回 0; 否则,返回 1。其中,检验函数值为 0 表示不可行,为 1 表示可行。取 $x_1 = \mu(0, 1)$, $x_2 = \mu(0, 1)$, ..., $x_N = \mu(0, 1)$, 其中, $\mu(0, 1)$ 表示在 $[0, 1]$ 上产生的均匀分布的随机数。然后作归一化处理,并记 $x = (x_1, x_2, \dots, x_N)$ 为一个染色体。经过有限次抽样后,可得到 M 个可行的染色体。记 $V(0) = (v_1, v_2, \dots, v_M)$ 为初始解群; $V(g) = (v_1, v_2, \dots, v_M)$ 为第 g 代解群。

(2) 在此计算 $\lambda = 0.2$ 和 $\lambda = 0.8$ 两个不同风险厌恶系数下的最大适应值, 这里采用方程 2 作为目标函数。方程 3 作为统计函数, 本文采用适应值比例选择方法, 用 Holland(1975)的赌轮盘方式实现。对于给定规模为 l 的群体 $P = \{a_1, a_2, \dots, a_l\}$, 首先计算个体 $a_j \in P$ 的适应度值 $f(a_j)$, 然后计算此适应度值在群体适应度值总和中的比重, 表示该个体被选择的概率, 则个体被选择的期望数目为 $P(a_j) = l \times P_j(a_j)$, 其中 $j = 1, 2, \dots, l$ 。本文的选择概率为 0.39。

(3) 在下一代前 $P_{CM} = 300$ 个染色体中随机选择两个染色体 v_p 、 v_q , 在 $[1, N]$ 范围内产生一个随机整数 r 。在 r 处交叉产生两个新染色体, 作归一化处理, 加入下一代解群。重复(3)中的操作, 直至产生下一代解群 $V(g+1)$ 的后 $P_{CM} = 700$ 个染色体。本文设交叉概率为 0.7。

(4)a、随机选择一个染色体 v_r , b、换 v_r 中随机选择的两个基因位, 重复 a-b、直到已经完成 $P_{MM} = 10$ 次变异操作。设定变异概率 $P_m = 0.01$ 。

(4) 重复步骤 2-步骤 4, 直到满足运行的最大代数 1000 ($\max g = 1000$)

(5) 计算 1000 代中适应值最高的代数的相应值。

根据第三章建立的模型, 根据遗传算法求得在不同风险厌恶系数下的最优投资组合比例, 具体结果为:

表 4.1 不同风险厌恶系数下的求解结果
Tab4.1 The result with different risk aversion coefficient

λ	0.2	0.8
收益	3.5E-04	2.974E-03
风险	3.98566E-04	5.5236E-04
适应值	0.01217	0.081597
无风险资产比例 η_1	65%	48%
中风险资产比例 η_2	20%	21%
高风险资产比例 η_3	15%	31%

4.2.1 风险偏好下的最优解

目前, 风险偏好的投资基金主要泛指投资于股票的投资基金, 这类基金也叫做股票型投资基金, 在基金持仓中, 股票占据了很大的份额, 基金净值的增长和降低主要来自

于股票的增长和降低，现在很多基金通过其名称来表明其投资类别和其持仓股票的性质，比如说景顺资源垄断（LOF）基金，其持仓中更倾向于投资垄断资源类的股票，通过投资垄断性资源行业股票，来获得超额收益；基金景顺成长，其持仓中更倾向于投资成长企业类的股票，主要投资成长类的企业的股票，通过企业股票长期的增长，来带动基金净值的增长；基金上证 180ETF，完全按照上证 180 指数中，各类别企业在指数组成中的比例，来建立自身基金的持仓比例，完成自身的投资组合，通过上证指数的增长，来带动基金净值的增长。

在风险厌恶系数为 0.8 时，根据遗传算法，分别独立运行，建立不同的初始种群，将每次运算的收益、风险和适应值求算术平均数，并计算相应的方差。得到这样的结果：投资组合的收益为 0.002974，方差为 5.5236E-04，适应值为 0.081597，三种资产的投资比例为 48%、21%、31%。

4.2.2 风险厌恶下的最优解

风险厌恶型的基金主要表现在债券型基金上，顾名思义，债券型基金的主要投资对象是不同类型的债券，但是从基金净值可以随时间增长的条件来看，债券型基金还必须投资除债券以外的其他金融衍生品，这其中最主要的是股票，但是由于其招募目的为保值增值，其基金资产的安全性尤为重要，所以，在其基金净值中，债券占据绝大部分的份额，股票只占基金净值中的很小的一部分。比如说嘉实债券，其股票资产只占其净值总额的 13.59%，净值的 82.41%都是投资于债券等金融工具，一方面，股票资产带来了基金净值的增长，另一方面，巨大比例的债券资产给基金的安全性带来了很大的保障。富国天利增长债券，其股票资产占基金净值的比例仅仅为 3.94%，比例十分的小，可以说，在基金的操作中，安全性要比净值增长能力要高。

在风险厌恶系数为 0.2 时，根据遗传算法，分别独立运行，建立不同的初始种群，将每次运算的收益、风险和适应值求算术平均数，并计算相应的方差。得到这样的结果：投资组合的收益为 0.00035，风险大小为 3.98566E-04，适应值为 0.01217，三种资产的投资比例为 65%、20%、15%。

4.3 与现有的基金投资组合模型投资比例的对比

应用遗传算法对基金投资组合模型进行了求解，得到了最大适应值，求得了三种资产在具有不同风险系数下的最优投资比例。应用风险调整后的收益来作为原有投资组合绩效与应用新建投资组合绩效的标准。

现有的开放式基金共分为股票型，货币型，混合型，债券型四类，其中货币型主要投资于货币市场，如短期国债、回购、央行票据、银行存款等等，由于其投资对象单一，

本文不对这类基金进行分析。其他三类基金的投资对象为股票和债券，其中股票型基金的风险最大，收益也相对较高，债券型基金主要投资对象为各个类型的国债，其中有少量的股票，风险相对较小，收益也相对较小，混合型基金是介于股票型基金和债券型基金的一种基金类型，投资对象也是股票和债券，只是股票投资比例相对股票型基金的投资比例小，但是比债券型基金中股票的投资比例要大，风险介于股票型基金和债券型基金之间，但是收益相对债券型基金大的多。

本论文选取基金的标准是，将每类基金中的基金按照基金的业绩，日收益率，基金公司整体业绩，三项指标来选取最好的五只基金，其中股票型基金选取富国天益，华夏成长，上投，易放达价值精选，广发小盘；混合型基金选取了广发聚富，易方达策略成长，上投摩根中国优势，富国天瑞，华夏大盘精选；由于本身债券型基金规模和基金数量上就比较少，所以不能按照基金业绩和日收益率等指标来选取，只能按照债券型基金得发行公司的规模，以及发行公司的其他基金得历史业绩来作为参考，所以主要选取了，嘉实债券，长盛中信全债，银河收益，富国天利增长债券，国投瑞银融华债券。由于每类基金的基金规模不同，单单按照每类基金的股票投资比例来计算同一类基金的股票投资比例的话，那么计算的结果势必要与实际值有很大的偏差，故在此在求解同一类别基金的股票投资比例的时候，采用加权平均的方法来进行计算，降低误差产生的概率。

股票型基金中选取的 5 支基金，其总共的基金资产为 3955765.4 万元，其中富国天益的资产规模是 801927.7 万元，占总资产规模的 20.27%；华夏成长的资产规模为 1047554 万元，占总资产规模的 26.48%；上投摩根阿尔法的资产规模是 655937 万元，占总资产的 16.58%；易放达价值精选的资产规模是 944216.5 万元，占总资产的 23.87%；广发小盘的资产规模为 506130.5 万元，占样本总资产规模的 12.8%。其中资产规模最大的为华夏成长，为 1047553.73 万元，其中股票资产占自身资产比例在所选样本最高的是富国天益，为 56.41%。按照前面的股票分类，中风险股票资产占总资产规模的 7.231%，高风险股票资产占总资产规模的 30.125%。

表 4.2 股票型基金样本数据
Tab4.2 Stylebook data of stock fund

基金名称	基金代码	资产规模	股票占 基金资 产比例	自身资产占 总资产规模 的比例	加权股票 资产占资 产比例
富国天益	100020	801927.71	56.41%	0.202724	0.114356
华夏成长	000001	1047553.73	34.38%	0.264817	0.091044
上投 α	373010	655937.01	35.71%	0.165818	0.059214
易放达价值精选	110009	944216.524	31.44%	0.238694	0.075045
广发小盘	162703	506130.47	26.50%	0.127948	0.033906
总资产规模		3955765.44			0.373566

表 4.3 股票型基金收益方差数据
Tab4.3 Stock fund data of income and risk

基金名称	自身资产占 总资产规模 的比例	收益率	方差	加权平均 收益率	加权方差
富国天益	0.202724	16.82e-04	3.58e-04	3.41E-04	7.26E-05
华夏成长	0.264817	9.25e-04	1.66e-04	2.45E-04	4.40E-05
上投	0.165818	6.649e-03	3.38e-04	1.10E-03	5.60E-05
易放达价值精选	0.238694	16.26e-04	5.14e-04	3.88E-04	1.23E-04
广发小盘	0.127948	-5.2e-04	5.41e-04	-6.65E-04	6.92E-05
总计	1	1.04E-02		2.01E-03	3.64E-04

混合型基金选取的样本总资产为 2548920 万元，其中广发聚富资产规模达到了 574911.2 万元，占总资产规模的 22.35%，易方达策略成长的资产规模为 653383.6 万元，占总资产的 25.4%，其中，其股票资产占其自身资产比例为 37.52%，为选取的样本数据中股票资产占自身资产最高的一支基金。上投摩根中国优势的资产规模为选取的 5 支基金中资产规模最大的，为 763693.6 万元，占总资产规模的 26.69%，兴业趋势的资产规模较小，为 247155.3 万元，占总资产规模的 9.6%，但是其股票资产占其自身的资产比例较高，为 35.49%，华夏大盘精选由于其目前净值较高，日收益率一直以来排在总基

金排名的靠前位置,故将其选中作为一个样本,其资产规模为 333068.9 万元,占总资产规模比例为 12.95%。在选取的五个样本中,上投摩根中国优势的资产规模最大,为 763693.6 万元,而股票资产占自身资产比例最高的是易方达策略成长,其股票资产占自身资产比例达到了 37.52%。按照前面的股票分类,中风险股票占总资产规模的 6.86%,高风险股票占总资产规模的 26.34%。

表 4.4 混合型基金样本数据

Tab 4.4 The mix fund data

基金名称	基金代码	资产规模	自身资产规模 占总资产规模 的比例	股票占基 金资产比 例	股票加权 占资产比 例
广发聚富	270001	574911.2	0.223508447	30.39%	0.0679242
易方达策略成长	110002	653383.6	0.254016171	37.52%	0.0953069
上投摩根中国优势	375010	763693.6	0.29690143	33.25%	0.0987197
兴业趋势	163402	247155.3	0.096086647	35.49%	0.0341012
华夏大盘精选	000011	333068.9	0.129487306	27.73%	0.0359068
总资产规模		2572212			0.3319588

表 4.5 混合型基金收益方差数据

Tab4.5 Mix fund data of income and risk

基金名称	自身资产规 模占总资产 规模的比例	收益率	方差	加权平均 收益率	加权方差
广发聚富	0.223508447	0.0018	0.000186	0.000402	4.15726E-05
易方达策略成长	0.254016171	0.0011	0.000495	0.000279	0.000125738
上投摩根中国优势	0.29690143	0.0006	0.000655	0.000178	0.00019447
兴业趋势	0.096086647	0.0021	0.00024	0.000202	2.30608E-05
华夏大盘精选	0.129487306	-0.001	0.000655	-0.00013	8.48142E-05
总计	1	0.0046	0.00223	0.000932	0.000469656

债券型基金选取的样本总资产规模为 321782.4954 万元,整体资产规模较股票型基金和混合型基金小很多,一般资产规模在 5 亿元到 10 亿元左右,其中嘉实债券的资产规模为 91973.7 万元,占总资产比例为 28.58%;长盛中信全债的资产规模为 45245.8 万

元，占总资产规模的比例 14.06%；银河收益的资产规模为 50890.3 万元，占总资产规模的比例为 15.82%；富国天利增长债券的资产规模为 81424.9 万元，占总资产规模的比例为 25.3%；国投瑞银融华债券的资产规模为 52248.3 万元，占总资产规模的比例为 16.23%。五个样本中，资产规模最大的为嘉实债券，占总资产比例为 28.58%，股票占基金资产比例最高的为国投瑞银融华债券，其股票资产占自身资产比例为 15.47%。按照前面的股票分类，中风险股票占总资产规模的 4.367%，高风险股票占总资产规模的 6.176%。

表 4.6 债券型基金样本数据
Tab4.6 The bond fund stylebook data

基金名称	基金代码	资产规模	占总资产规模的比例	股票占基金资产比例	股票加权占资产比例
嘉实债券	070005	91973.72973	0.285826	13.59%	0.038844
长盛中信全债	510080	45245.28302	0.140608	8.89%	0.0125
银河收益	151002	50890.32258	0.158151	11.96%	0.018915
富国天利增长债券	100018	81424.90909	0.253043	3.94%	0.00997
国投瑞银融华债券	121001	52248.25095	0.162371	15.47%	0.025119
总资产规模		321782.4954			0.105347

表 4.7 债券型基金收益方差数据
Tab4.7 The bond fund data of income and risk

基金名称	自身资产 规模占总 资产规模 的比例	收益率	方差	加权平均收 益率	加权方差
嘉实债券	0.285826	7.58245E-06	1.57E-05	2.17E-06	4.49E-06
长盛中信 全债	0.140608	0.000227211	2.13216E-05	3.19477E-05	3.00E-06
银河收益	0.158151	0.000147659	7.39998E-05	2.33524E-05	1.17E-05
富国天利 增长债券	0.253043	-0.000128245	1.22915E-05	-3.24515E-05	3.11E-06
国投瑞银 融华债券	0.162371	0.000411436	5.87396E-05	6.68053E-05	9.54E-06
总计	1			9.18E-05	3.18E-05

比较分析可得，在样本的选取上，股票型基金的总资产规模最大，其次是混合型基金，最少的是债券型基金，这在一定程度上也反映了三类基金在市场上的销售情况，在三类基金中，股票型基金的股票资产占总资产的比例最高，为 37.36%，其次为混合型基金，其股票资产占总资产比例为 33.2%，债券型基金中的股票资产占总资产比例最小，仅仅为 10.5%。

表 4.8 三类基金的风险收益及风险资产投资比例
Tab4.8 Risk assets investment scale and the risk and income of three kind of fund

	股票型基金	混合型基金	债券型基金
收益	2.01E-03	9.32E-04	9.18E-05
风险	3.64E-04	4.69656E-04	3.18E-05
无风险资产比例 η_1	62.644%	66.8%	89.457%
中风险资产比例 η_2	7.231%	6.86%	4.367%
高风险资产比例 η_3	30.125%	26.34%	6.176%

通过对比表 4.1 和表 4.8，在风险厌恶系数为 0.8 时，投资组合偏好倾向于获得风险调整后的收益，这类投资组合中的资产高风险资产占有的比例较大，通过遗传算法求得

的投资组合在风险调整后的收益上比现在的股票型基金和混合型基金要高, 风险又低于股票型基金, 在风险厌恶系数为 0.2 时, 投资组合偏好倾向于获得稳定的收益, 风险和收益的比较中, 投资者更看重投资的低风险, 这类投资组合中的资产中风险资产和无风险资产占有的比例较大, 通过遗传算法求得的投资组合在收益上比现在的债券型基金要高很多, 风险又低于债券型基金, 可以说明通过遗传算法求得投资组合模型得到的投资比例, 优于现有的基金投资比例。也证明了, 通过遗传算法求解的投资组合模型在计算结果的精确度方面上优于现有的通过其他方法求得的投资组合模型。

4.4 与应用遗传算法求解投资组合模型比较

在应用遗传算法求解投资组合模型的过程中, 首先是建立目标函数模型, 将目标函数模型作为遗传操作的评估准则, 现有的目标函数模型的建立, 完全采用了统计学中的二次规划模型, 只对风险最小, 或者收益最大来作为评价的目标函数, 这样的模型有利于统计上的应用, 但是经济意义不明确, 风险最小或者收益最大只能简单描述投资者的最终目标。

以下面模型为例:

$$\min T = (1 - \mu) \left[\sum x_i \beta_i \right] \sigma_m^2 - \mu \sum x_i A_i \quad (4.1)$$

$$s.t.: \sum x_i = 1 \quad x_i \text{ 代表第 } i \text{ 种资产的权重}$$

$$x_i \geq 0 \quad \text{表示我国不允许卖空证券}$$

此模型引入权重参数 μ 和 $(1 - \mu)$ 分别表示投资净收益和总体投资风险的权重值, 权重数值和 $(1 - \mu)$ 分别表示决策者对投资净收益和总体投资风险的重视程度, μ 值越大, 表示决策者更注重投资净收益, 敢于冒风险; 反之, 表示投资者比较保守, 不敢冒太大的投资风险。 $\left[\sum x_i \beta_i \right] \sigma_m^2$ 表示投资的风险, $\sum x_i A_i$ 表示投资的收益, 整个模型的含义就是投资的风险与收益的差最小, 但是从量纲分析来看, 风险和收益的量纲是不同的, 二者的比较不是在一个相同的位子上来衡量的, 所以这种现在应用的比较多的线性模型其统计意义很明显, 但是经济意义不是很明确, 如果说, 只是衡量投资组合的风险最小, 并不能说明投资组合的绩效就是最好的, 所以现在的模型在衡量投资组合绩效方面存在着评价目标不明确的问题。

本文建立的模型, 无论从统计学上构建, 还是从经济学意义方面, 都比现在的投资组合模型优越。

$$f(r) = \frac{E(r) - R_F}{\beta} = \frac{(\lambda\eta_1 - 1)R_F + (1 - \lambda)(\eta_2 E_{rl} + \eta_3 E_{rh})}{\beta} \quad (4.2)$$

$$\max f(r) = \frac{E(r) - R_F}{\beta} = \frac{(\lambda\eta_1 - 1)R_F + (1 - \lambda)(\eta_2 E_{rl} + \eta_3 E_{rh})}{\beta} \quad (4.3)$$

$$s.t. \quad \eta_1 + \eta_2 + \eta_3 = 1$$

$$\eta_{i(i=1,2,3)} \geq 0$$

从经济意义来讲，方程 $f(r)$ 的经济含义为风险调整后的超额收益，即每承担单位风险而获得的超额收益，以此来衡量投资组合的绩效。用此标准衡量的原因为，每承担单位风险而获得的超额收益要比单从风险大小或者收益大小来衡量投资绩效要全面，单从风险角度或者收益方面来衡量投资组合绩效并不全面，如果投资选择了一个风险最小的投资组合，那么可能这个投资组合的收益要比市场上风险次小的投资组合收益低很多；若投资组合的收益很高的话，那么要获得高收益，就要承担高风险，这样在承担单位风险上获得的收益并不比承担此高风险而获得的收益高。这样，单从风险和收益两方面来衡量投资组合的绩效的话，都不是完全有效的。本论文中的风险偏好系数是基于超额风险系数来选取的，而不是原有的投资者去主观臆测的，超额风险系数为所选的投资对象的风险收益比，用单位收益承担的风险值来衡量，投资者可以根据选取得投资对象的历史数据计算投资对象的风险值和收益，用风险值比收益，得到超额风险系数。利用公式 3.28，计算风险偏好系数，这样就避免了原来的主观臆测风险偏好带来的误差。有利于选取更好的投资组合。公式 3.28 建立风险偏好系数的数学表达式。为测量风险偏好建立了数学模型。

5 结论

5.1 本文的研究结论

首先,从风险角度对基金资产进行了分类。分析影响股票资产风险的因素,从宏观和微观两个角度选取了16个指标,这16个指标囊括了影响股票资产风险的因素,通过聚类分析,最终选取了9个显著性指标,建立了由这9个指标组成的指标体系。得到了股票资产风险受股票盈利能力,企业偿债能力,市场操纵行为,市场价格趋向四个因素影响,股票盈利能力对股票资产风险呈正相关关系;企业偿债能力与股票资产风险呈负相关关系,市场操纵行为与股票资产风险呈负相关关系,市场价格趋向与股票资产风险呈正相关关系。通过建立的股票风险模糊数学模型,按照隶属度矩阵和评语集,将股票资产明确的分为中风险资产和高风险资产。中风险资产是:为获得单位资产所要承受的风险值低于平均承受风险值的那部分资产,高风险资产是:为获得单位资产所要承受的风险值高于平均承受风险值的那部分资产。

其次,引入了风险偏好系数的投资组合的绩效要好于传统的投资组合。将风险偏好系数引入到基金投资模型中,原有的模型的适用范围更大了,原来的模型只能首先假定投资者是风险中性的,基于此来进行投资组合。当在投资组合模型中加入了风险偏好系数后,模型的适用范围就扩大了。并且定义了超额风险系数,其经济意义是为了获得单位收益而要承担的风险值。将超额风险系数和风险偏好系数建立了联系,风险偏好系数和超额风险系数为正相关关系,也就是说,为了获得更高的收益,必须要承担更大的风险,这与经济原理是一致的。

再次,以遗传算法中的目标函数的方式建立了基金投资组合模型。将基金投资组合模型作为遗传算法中的适应度函数,应用遗传算法,解决了原有投资组合模型求解结果精确度不高的问题,求得了在不同风险偏好系数下的最优值。从而使不同风险偏好的投资者可以根据自身的风险偏好,来选择三种资产的投资比例,使投资组合模型的应用范围更广了。通过对现有基金的绩效对比分析,应用基于遗传算法建立的模型进行投资得到的投资绩效比现有基金的投资绩效要好。

5.2 进一步研究的建议

本文的不足之处在于,在对基金资产进行分类的时候,指标的选取范围不够全面,只是针对微观层面来选取得指标,如果能够从宏观角度来选取一些影响指标的话,那么指标的解释意义更强。在进行与现有基金进行比较时,由于存在交易费用,所以得到

的结果要比现有基金强许多的结论，若加上交易费用这一限制，得到的结果会和现实有所接近。

随着我国证券市场的不断发展，将各种金融资产进行投资组合，能够有效地降低证券市场得波动风险，为证券市场的健康发展提供一定的保障。在不久的将来，一些限制投资的法规将进一步放宽，股指期货也将成为基金的资产，股指期货与现有资产的负相关性的关系，可以在投资组合中加入股指期货，这样可以达到降低投资组合风险，提高收益的作用。未来，我国的金融全球一体化的趋势会越来越强，投资国外的资本市场也是一个很好的投资方向，随着新的QDII产品的出现以及国家对QDII的管制进一步放开，QDII将会投资海外的资本市场，如何将海外资本市场和国内资本市场相结合，也将成为一个研究方向。

参 考 文 献

- [1]Markowitz H. Portfolio selection. *Journal of Finance*, 1952 (3) : 77- 91.
- [2]韩战钢. 遗传算法及在经济中的应用. *北京师范大学学报*. 2000. 23-26
- [3]丰雪. 投资组合理论的脉络及发展. *辽宁大学学报*. 2005, 32 (2)
- [4]严明义, 路维春. CAPM的历史渊源及附加假设条件研究. *经济问题*. 2006 (2)
- [5]景乃权. 资本资产定价模型及其评述. *经济学家*. 2004. 116-118
- [6] 埃德温·J·艾尔顿, 马丁·J·格鲁伯, 斯蒂芬·J·布朗, 威廉·J·格茨曼. 现代投资组合理论和投资分析. 中国人民大学出版社. 2000. 221-241
- [7]牛庆莲, 张霞, 杨月巧. 套利定价理论及应用. *山西财经大学学报*. 2002 (24)
- [8]东朝辉. 对套利定价理论及应用的再认识. *数量经济技术经济研究*. 2003, 05: 144
- [9]王硕, 唐小我, 曾勇. 基于加速遗传算法的组合证券投资决策. *中国工程科学*, 2002, 4(9) :59-62.
- [10]金汉均, 洪峰. 遗传算法的证券组合投资优化问题的模拟分析. *华中师范大学学报*. 2004, 38(12)
- [11]刘超. 现代证券投资组合理论在我国应用的局限和思考. *经济经纬*. 2006, 02: 139
- [12]朱衍强. 中国股票市场统计指标的构建. *理论新探*. 2004, 08
- [13]韦廷权. 风险度量和投资组合构造的进一步实证. *南开经济研究*, 2001, 2:3-6
- [14]蔡乙萍, 万力, 范旭东. 各种指数基金模型的实证比较分析. *数量经济技术经济研究*. 2006 (10) .
- [15]李惠勇. 对股价变动的几点思考. *学术交流*. 2001 (1)
- [16]徐忠兰, 许永龙, 赵亮. 股票价格影响因素研究. *天津师范大学学报(自然科学版)*. 2004, 24(2)
- [17]吴世农, 陈斌. 风险度量方法与金融资产配置模型理论和实证研究. *经济研究*, 1999, 9:30-38
- [18]徐涤龙. 资本资产定价模型与上海股票市场的实证分析. *南昌大学学报(理科版)*, 2005, (02).
- [19]威廉·夏普, 戈登·J·亚历山大. 投资学. 中国人民大学出版社. 1998. 225-234
- [21]王欣荣, 樊治平. 上市公司财务状况的动态多指标综合评价方法. *系统工程理论与实践*, 2002, 22(4) :54-57.
- [20]马立肖, 王江晴. 遗传算法在组合优化问题中的应用. *计算机工程与科学*. 2005, 27(7)
- [22] Nawrocki David. A brief History of Downside Risk Measure. *Journal of Investing*, 1999, 8(3) 9-25
- [24]张鸿雁, 任叶庆. 有风险控制的最优资产组合的数学建模与计算分析. *系统工程*, 2002, 20(5) : 40- 42.
- [25] 攀登, 吴冲锋. 用遗传算法直接搜索证券组合投资的有效边界. *系统工程学报*, 2002, 17(4) : 364-367.
- [26]王庆石, 肖俊喜. 风险调整的投资组合绩效测度指标综合评价. *世界经济*. 2001 (10)
- [27]卢有杰, 卢家仪. 项目风险管理. 清华大学出版社. 1998. 40-47
- [28] 任叶庆, 张鸿雁. 有风险控制的最优组合的数学建模与计算分析. *系统工程*, 2002(9), 20(5) :40-42
- [29]许谨度, 周江雄. 风险管理. 中国出版社. 1998. 56-74.
- [30]陈彦斌, 邹恒甫. 基于风险基金的资本资产定价模型. *经济学季刊*. 2003 3(1)

- [31] 王小平, 曹立明. 遗传算法—理论、应用与软件实现. 西安交通大学出版社, 2002.
- [32] 马永开, 唐小我. 允许持有无风险资产的 β 值证券组合投资决策模型研究. 系统工程理论与实践. 2002, 12.
- [33] 李善明, 徐沛. 马克维茨投资组合理论模型应用研究. 经济科学. 2000, (1):42-51
- [34] 岳意定, 肖赛君. 影响股票价格的宏观因素研究. 湖南财经高等专科学校学报. 2006, 22(101)

附录 A 上证 180 指数中选取的 40 家上市公司数据

名称	代号	隶属度值	实际值	每股收益	流通股本
贵州茅台	600519	0.57	110.31	4.08	6.08
沪东重机	600150	0.38	104.06	1.02	1.58
天马股份	002122	0.36	79.99	0.27	3
驰宏锌锗	600497	0.77	71.16	1.78	9.63
张裕 A	000869	0.37	62.84	0.83	0.93
吉林敖东	000623	0.67	69.8	4.36	25.86
中国平安	601318	0.52	64.48	5.75	10.09
马应龙	600993	0.75	73.79	0.26	0.63
大商股份	600694	0.28	51.99	2.11	0.82
天地科技	600582	0.12	56.94	0.76	0.26
中信证券	600030	0.42	64.17	17.1	25.34
中兴通讯	000063	0.04	46.99	4.87	9
XD 中财国	600970	0.38	61.21	0.73	0.71
苏宁电器	002024	0.09	49.2	5.88	5.39
宏达股份	600331	0.38	40.96	2.5	4.38
宝泰股份	600456	0.21	41.82	1.39	2.97
招商地产	000024	0.09	46.8	3.17	1.09
烟台万华	600309	0.27	48.03	5.88	3.07
中金黄金	600489	0.2	45.79	1.28	3.49
航天信息	600271	0.26	46.29	1.95	2.1
辽宁成大	600739	0.46	47.95	7.16	22.28
王府井	600859	0.18	40.56	1.92	0.36
首旅股份	600258	0.15	30.8	0.91	0.63
盐湖钾肥	000792	0.24	40.78	3.36	2.08
新安股份	600596	0.28	40.53	1.97	1.08
用友软件	600588	0.22	49.35	1.36	0.64
同仁堂	600085	0.17	36.6	1.94	1.21
中国卫星	600118	0.06	47.94	1.11	4.28
保利地产	600048	0.07	39.13	3	3.34
三一重工	600031	0.38	37.34	3.46	7.11
上海机场	600009	0.2	34.11	9.97	8.38

海油工程	600583	0.07	45.01	4.76	4.58
中联重科	000157	0.28	32.62	3.35	2.54
美的电器	000527	0.15	48.23	9.13	3.32
云南白药	000538	0.15	36.04	2.21	1.33
浦发银行	600000	0.22	29.65	32.6	9.01
五粮液	000858	0.2	27.55	17.14	18.24
招商银行	600036	0.17	21.31	47.1	19.51
万科 A	000002	0.09	18.45	48.8	17.77
上港集团	600018	0.04	11.5	24.2	17.12

根据前面股票风险隶属度矩阵的求解形式,选取 13 个板块中的 40 只股票进行分类。通过计算得到这样的结论,在 13 个板块之中,采掘业、初级产品工业、医药卫生、交通运输和仓储业、公共事业和综合类、社会服务业为中风险板块;耐用消费品制造、金融业、房地产业、信息技术业、贸易和零售业、投资品工业和经常性消费品工业为高风险板块。

攻读硕士学位期间发表学术论文情况

[1]梁艳, 吴阳。基于遗传算法的基金投资组合模型研究。中国科技论文在线, 200712-62

致 谢

本人于 2005 年开始，在大连理工大学经济系攻读金融学专业硕士学位。在学习期间，我有幸得到了经济系各位老师的精心栽培，得以深入研究专业知识，为论文写作打下了深厚的理论基础。在毕业论文的选题、开题、收集资料及写作的过程中，我得到了导师梁艳副教授和经济系各位老师的悉心指导。导师对我的指导和鼓励使我不但顺利的完成了课题的研究和论文的撰写，掌握了科学的学习和研究方法，也使我鼓起了面对生活挑战的勇气。导师渊博的学识，严谨的科学态度，以及对问题清晰敏锐的判断力使我受益匪浅。在此谨向梁艳老师致以最诚挚的感谢！

我还要感谢经济系的全体教师，他们是原毅军老师、金镛老师、侯铁珊老师、刘凤朝老师、兆文军老师、逯羽铎老师、任曙光老师和陈艳莹老师，他们在我的论文选题和预答辩中以高度负责的态度，对我的论文工作给予了指导和热情的帮助，感谢在学术上的悉心指导！

同时感谢师姐于秋景，以及我的同窗好友们对我的帮助、支持和鼓励！

最后，我要感谢我的父母和各位亲人在攻读学位期间对我的大力支持，使我能够全身心地投入研究工作！

感谢两年以来所有关心、支持和帮助过我的人！