

分类号 F83
UDC 591

密级 公开
编号 10741



硕士学位论文
(专业学位)

论文题目 基于 EGARCH-GA-KMV 模型的我国
上市公司信用风险研究

研究生姓名: 黄令根

指导教师姓名、职称: 姬新龙、教授

学科、专业名称: 应用经济学、金融硕士

研究方向: 金融投资与金融风险

提交日期: 2024 年 6 月 3 日

独创性声明

本人声明所提交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。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。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名： 黄气根 签字日期： 2024.6.3

导师签名： Hebing 签字日期： 2024.6.3

导师(校外)签名： _____ 签字日期： _____

关于论文使用授权的说明

本人完全了解学校关于保留、使用学位论文的各项规定， 同意 (选择“同意” / “不同意”) 以下事项：

1. 学校有权保留本论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文；

2. 学校有权将本人的学位论文提交至清华大学“中国学术期刊(光盘版)电子杂志社”用于出版和编入 CNKI《中国知识资源总库》或其他同类数据库，传播本学位论文的全部或部分內容。

学位论文作者签名： 黄气根 签字日期： 2024.6.3

导师签名： Hebing 签字日期： 2024.6.3

导师(校外)签名： _____ 签字日期： _____

Research on Credit Risk of Listed Companies in China Based on EGARCH-GA-KMV Model

Candidate : Huang Linggen

Supervisor: Ji Xinlong

摘 要

信用风险作为金融市场上最重要且最古老的金融风险,对其进行预测和控制显得尤为重要。近些年受国内外因素的影响,我国上市公司信用风险问题暴露愈加严重,也备受瞩目。通过研究国际上流行的信用风险评价方法,试图找到适合我国上市公司信用风险管控的模型和评价体系,这不仅有助于我国信用风险管理水平的提升,还对我国金融市场的健康稳定发展具有重要的理论和实践意义。

本文首先对信用风险、信用风险度量模型及其发展演进、相关理论基础等进行详细阐述,并对信用风险度量模型进行综合比较分析,考虑 KMV 模型在我国的适用性。其次介绍了 KMV 模型的建立基础及理论,再针对传统 KMV 模型中估算股权价值波动率以及确定违约点系数方法在我国适用性不强的问题,分别应用 EGARCH 模型和遗传学算法(GA)对股权价值波动率和违约点系数重新确定,从而构建了精度更高的 EGARCH-GA-KMV 模型。实证方面,选择 40 家高风险组(ST 类公司)和低风险组(非 ST 类公司)作为研究样本,并进行研究比较。结果证明,重构之后的模型能比较有效的辨别高低风险组,其中高风险组(ST 类公司)的违约距离显著小于低风险组(非 ST 类公司),且经过验证,发现重构后的模型较传统 KMV 模型能够更好地体现信用风险高低的差异。因此,重构后改进的模型在我国上市公司的违约风险测度方面具有良好效果,适用于我国特定的国情。最后,对未来我国上市公司信用风险预测提出一些展望。

关键词: 信用风险 KMV 模型 遗传学算法 EGARCH 模型 违约距离

Abstract

Credit risk, as the most important and oldest financial risk in the financial market, is particularly important to predict and control. In recent years, influenced by internal and external factors, credit risk issues of listed companies in China have become increasingly serious and have attracted significant attention. By studying internationally popular credit risk assessment methods, we attempt to find models and evaluation systems suitable for credit risk management of listed companies in China. This not only helps improve the level of credit risk management in China but also holds important theoretical and practical significance for the healthy and stable development of China's financial market.

This paper first elaborates on credit risk, credit risk measurement models, their development and evolution, relevant theoretical foundations, and conducts a comprehensive comparative analysis of credit risk measurement models, considering the applicability of the KMV model in China. Secondly, it introduces the establishment foundation and theory of the KMV model, and addresses the weak applicability of estimating equity value volatility and determining default point coefficient methods in the traditional KMV model in China. It applies the EGARCH model and Genetic Algorithm (GA) to re-determine equity value volatility and default point coefficients, thus constructing a more accurate

EGARCH-GA-KMV model. In terms of empirical research, 40 high-risk group (ST companies) and low-risk group (non-ST companies) are selected as research samples for comparison. The results demonstrate that the reconstructed model can effectively distinguish between high and low-risk groups, with the default distance of the high-risk group (ST companies) significantly smaller than the low-risk group (non-ST companies). It is verified that the reconstructed model can better reflect the differences in credit risk levels compared to the traditional KMV model. Therefore, the improved model after reconstruction has a good effect on measuring default risk of listed companies in China and is suitable for the specific national conditions of China. Finally, some prospects for future credit risk prediction of listed companies in China are proposed.

Keywords: Credit risk; KMV model; Genetic algorithm; EGARCH model; Default distance

目 录

1 绪论	1
1.1 研究背景与意义	1
1.1.1 研究背景	1
1.1.2 研究意义	2
1.2 国内外文献综述	2
1.2.1 信用风险度量研究	2
1.2.2 KMV 模型信用风险度量研究	6
1.2.3 KMV 模型修正和应用研究	7
1.2.4 文献评述	8
1.3 研究方法与内容	9
1.3.1 研究方法	9
1.3.2 研究内容	9
1.3.3 研究框架	11
1.4 创新与不足	12
1.4.1 创新之处	12
1.4.2 不足之处	12
2 相关理论基础与信用风险度量模型	13
2.2 相关理论基础	13
2.2.1 信用风险理论	13
2.2.2 期权定价理论	15
2.2.3 遗传学算法	18
2.3 信用风险度量模型	21
2.3.1 信用风险度量模型发展	21
2.3.2 信用风险度量模型比较分析	22
2.4 本章小结	27
3 EGARCH-GA-KMV 模型的构建与参数选择	28
3.1 传统 KMV 模型的构建	28
3.1.1 KMV 模型理论基础	28

3.1.2	KMV 模型参数选取	30
3.1.3	传统 KMV 模型建立	33
3.1.4	传统 KMV 模型评价	34
3.2	EGARCH-GA-KMV 模型的构建	35
3.2.1	基于 EGARCH 模型的 KMV 模型参数修正	35
3.2.2	基于遗传学算法的 KMV 模型参数修正	38
3.2.3	EGARCH-GA-KMV 模型计算过程	41
3.3	本章小结	42
4	基于 EGARCH-GA-KMV 模型的上市公司信用风险实证分析	43
4.1	样本选取和参数设定	43
4.1.1	样本选取	43
4.1.2	参数设定	44
4.2	实证分析	45
4.2.1	基于 EGARCH 的股权价值波动率检验和确定	45
4.2.2	资产市场价值及波动率确定	51
4.2.3	最优违约点确定	51
4.2.4	违约距离确定	54
4.3	实证结果检验和分析	55
4.3.1	EGARCH-GA-KMV 模型实证结果检验	55
4.3.2	样本组间实证结果对比分析	57
4.3.3	时间序列分析	58
4.4	本章小结	60
5	研究结论及展望	61
5.1	研究结论	61
5.2	未来展望	62
	参考文献	63
	致 谢	67

1 绪论

1.1 研究背景与意义

1.1.1 研究背景

近三十年来,我国经济始终保持着高速增长,到 2022 年 GDP 总量已达 121 万亿,相比 1978 年经济总量提高了近 327 倍,一跃成为世界第二大经济体。中国金融市场也持续发展,但受国家去杠杆化政策等多方面的影响,金融市场近几年来违约风险不断提高,各类风险事件也相继频繁发生,尤其我国上市公司信用风险暴雷问题渐突,如 2014 年“11 日超日债”违约、2021 年“恒大地产暴雷”事件、2023 年千亿企业河南“建业地产”违约等等。对我国上市公司而言,主要存在两个信用风险,一个是银行等金融机构对公司授信贷款公司违约风险,另一个是公司发债违约风险,无论哪种风险爆发都将对经济建设带来不可挽回的损失。

商业银行贷款方面,据国家金融监督管理总局历年发布的报告显示,2020-2022 年我国商业银行不良贷款余额分别是 2.7 万亿元、2.8 万亿元、2.98 万亿元,不良贷款率分别是 1.92%、1.81%、1.71%,数据显示我国商业银行不良贷款率在缓慢下降,而不良贷款余额却在上升,说明总贷款量在持续上升。公司发债违约风险来看,根据 WIND 数据库数据显示,我国 2022 年信用债共有 27 家企业 84 只信用债发生违约,违约金额达到 500.61 亿元,2020 年和 2021 年违约金额分别是 1076 亿元、1603.46 亿元,虽然 2022 年的违约金额相比 2021 年下降近七成,违约幅度大幅减少,但是债券展期却显历史新高。由此可见,从信用风险主要的两个方面出发,我国现存的信用风险违约金额是相对较大的,且潜在的未来信用违约风险具有上升的可能性。

其次,从国家层面上来看,2014 年国务院印发《社会信用体系建设规划纲要(2014-2020 年)》(后简称“纲要”),该纲要统筹兼顾、细致入微,从宏微观角度出发,明确了每个发展阶段的主体任务及建立合理有效的信用风险监管机制;2022 年党的二十大报告明确要求完善市场主体的信用监管机制,在新时代背景下,要更好地治理市场主体的违法、违约等失信现象;2023 年中国人民银行发布《中国金融稳定报告 2022》提出要有力推进防范化解金融风险体制机制建设,

加强风险预警、防控机制和能力建设,防止系统性金融风险发生。以上情况说明,提高信用风险的管理水平,防止引发系统性金融风险是国家金融风险监管的重要任务。

综上所述,监管层面不到位以及衡量信用风险能力缺乏导致上市公司在面临信用风险时风险管理能力不强,且国家层面对信用风险的管理也十分重视。在此背景下,有效预测和掌控公司信用风险,提前做好风险预警的重要性不言而喻,同时也是学者、监管者、金融机构与投资者所共同关注的紧要问题。所以,本文在充分参考国内外优秀文献的研究基础上,结合我国金融市场实际特殊状况,对 KMV 模型进行富有针对性的改进,并对改进后的 KMV 模型准确适用性进行实证分析,力求改进后的模型在对我国上市公司信用风险预测和度量方面准确率更高、适应度更强。

1.1.2 研究意义

本文研究具有理论意义和现实意义两部分内容。

理论意义:第一,经济新常态下,经济产业结构面临转型升级,产业转向高质量发展,对 KMV 模型进行改进升级,加强对上市公司信用风险预测,有利于减少行业成本、增强稳定性与产业优化升级、实现信贷资源有效配置。

第二,基于我国国情对 KMV 模型参数针对性改进,通过 M-WU 等检验证明改进型 KMV 模型识别准确度高,有助于 KMV 模型的运用以及拓展,完善我国信用风险评价体系。

现实意义:首先,投资者视角下,提升公司信用风险的估算准确率,可帮助投资者评估和把握投资机会,减少信用风险事件发生的频率,减少无效投资和不必要损失,进而提高整体投资者收益率。

其次,就金融风险防范而言,准确及时评估上市公司的信用风险,有助于增强监管机构对企业债务风险的预警能力,推动建立实用的信用风险缓解途径、风险管控工具以及预防体系。

1.2 国内外文献综述

1.2.1 信用风险度量研究

信用风险是金融风险的分支,是信用的延伸。信用风险,即违约风险,是指

债务人未能按时履约导致债权人生产经营活动带来损失的风险。从时间线来看，关于信用风险的度量从银行的诞生开始就已经产生；从国内外角度来看，国外信用风险度量发展优于国内，目前主流信用风险度量方法均来自国外研究；国内信用风险度量研究迟缓于国外，国内学者主要通过借鉴国外信用风险度量方法并结合我国实际情况进行改进、创新、应用。因此，本文结合时间线，把信用风险度量模型分为三种：古典信用分析法、基于统计学的信用风险度量模型和信用风险量化管理模型。

古典信用分析法，又称专家分析法、定性分析法，是训练有素、经验丰富的专家以贷款公司的定性信息为基准，通过主观判断对信贷风险进行大小判断和决策，主要包括 7 种方法：5C、5P、5W、4F、LAPP、CAMPARI 分析法和 CAMEL 评估法。5C 分析法即从借款人的道德、还款能力、资本实力、担保及经营环境状况五个方面来定性评判借款人的信用风险和还款能力，C 则为各个方面所代表意思英文字母的首字母；5P 分析法则从个人、资金用途、还款来源、债权保障及公司发展前景等五个因素（Factor）对债务人进行判断，因其每个因素均为英文 P 开头，故名为 5P 分析法；5W 分析法利用借款受信方（Who）、借款使用用途（Why）、还款时间期限（When）、担保抵押物（What）和如何还款（How）等债务人主观意图来对信用风险进行定性；4F 分析法具体含义为 Organization、Economic、Financial 和 Management 等四个要素（Factor）分别代表组织要素、经济要素、财务要素、管理要素。剩下的其他三个方法，如 LAPP 分析法、CAMPARI 分析法、CAMEL 评估法也都基于其字母所代表含义的要素出发，对其中一个要素指标进行评分，定性分析，因此古典信用分析法具有操作性强、推广使用容易的优点，但是评估结果容易受到专家知识经验、素质高低等因素影响，使得古典信用分析法取得信用风险评估结果缺乏可比性。

基于统计学的信用风险度量模型优化了古典信用分析法只能从定性角度衡量信用风险的缺点，通过引入财务定量数据和统计学方法，对一些趋向恶化的财务特征指标进行预测，进而确定其信用等级。基于统计学的信用风险度量模型主要包含单变量判别分析法、多元变量判别的 Z-score 评分模型、多元判别的 Zeta 模型、Logistic 回归模型（LR 模型）、Probit 回归模型等。单变量判别分析法即利用单一的财务比率来预测公司的财务困境，如 Beaver（1966）选取 1954 年到

1964 年连续十年间两对 79 家公司（79 家成功公司和 79 家失败公司）分别对应低风险和高风险组，从中挑选 30 个重要财务指标来对此研究，研究结果证明，在不考虑行业和资产规模的前提下，使用财务指标对公司分析财务失败是有效的，一定程度上验证了单变量判别分析法的实用性。随后，单变量判别分析法研究到多变量判定的阶段，美国学者 Altman(1968)首次利用多元判别分析，以 1940-1960 年代间的破产公司和对应数量的正常公司作为样本，利用 22 个财务指标生成 5 个比率变量加权汇总后产生总判断分指，建立起 Z-score 评分模型；Altman, Haldeman & Narayanan (1977) 在 Z-score 基础上把 5 个变量延伸拓展成 7 个变量，建立了 Zeta 模型；相较于国外学者，国内学者主要对多元判别分析法进行应用或者改良，邹昆仑等(2018)使用 SPSS 中的判别分析法对 Altman 的 Z-score 评分模型加以改进来研究中国企业债信用评级，但由于中国债券市场还不够成熟，使得改进后的模型准确率不高，为此提出对策建议；唐静和郝洪常(2020)通过运用“Z-Score”模型对西藏医药上市公司的财务危机进行预警研究，计算得出“Z-Score”模型比较适合西藏医药上市公司的财务预警且准确率达 83.33%；陈茜和田治威(2017)利用多元变量判别分析法之一的因子分析法对主要财务指标进行分析构建林业上市公司的财务风险评价模型，并使用聚类分析法来划分林业上市公司的财务风险程度，使用因子分析法计算出的结果跟实际结果一致，具有非常好的参考性；余静文等(2021)运用计算的 Z 值代替被解释变量商业银行风险承担研究与数字金融之间的关系，最后发现数字金融有助于抑制商业银行系统性金融风险。

多元变量分析法，比如 Z-score 评分模型、Zeta 模型等在验证公司信用风险方面有较高的准确率，但是其组内分布近似标准正态分布的假设、高风险组和低风险组配对标准在现实中难以满足。因此，Ohlson(1980)首次在研究中使用 Logistic 回归模型分析 105 家和 2058 家公司组成的非配对样本连续 7 年间的的历史数据，发现使用公司规模、资本结构等分析方法预测信用风险准确率达到 96%，从而突破了多元变量分析法的局限；Zavgren(1985)使用因子分析获得输入变量，继续使用 Logistic 回归模型分析同样获得相同的预测准确率；李森和赵轩维等(2018)基于融资项目和融资企业两个维度构建股权众筹项目融资成功率判别指标体系，继而比较 Logistic 回归模型和神经网络模型的判别能力，最后得出 Logistic 回归

模型具有较高准确率的判别能力,从侧面 Logistic 回归模型分析法优越性。Probit 回归模型与 Logistic 回归模型类似,唯一不同之处在于 Logistic 回归模型使用逻辑分布,而 Probit 回归模型则使用标准正态分布,Zmijewski (1984) 最早提出将 Probit 回归模型应用信用风险分析。对整个基于统计学的信用风险度量模型进行梳理可以发现,其摆脱了古典信用分析法古板定性分析,从定量角度并且在定量信息基础上加上各种统计回归模型,将违约或者不违约的黑白问题转化成概率问题,在很大程度上提升了信用风险测量的正确率。

信用风险量化管理模型起源于金融全球化趋势加强以及全球风险波动加剧的背景下,传统的因素分析法和统计学信用分析法已经不再适用于信用风险管理的要求,信用风险量化管理模型主要包含四种形式:KMV 模型、Credit Metrics 模型、Credit Risk+模型、Credit Portfolio View 模型(CPV 模型)。

KMV 模型于 1997 年在美国建立,主要用途是来测算借款受贷公司的信用风险违约概率。Altman 等(1997)在比对 KMV 模型和 Probit 回归模型实证结果基础上,发现相较于其他风险识别方法,KMV 模型识别效果更好;M.Roger Stein (2000)在论文中对 KMV 模型、ROA、Z-score 信用风险评估模型进行对比论证,最后得出 KMV 模型是其中准确率最高的模型;Credit Metrics 模型(1997)则结合信用评级体系和信用等级的贴现曲线计算出的 VAR 值来判定信用风险级别;Credit Risk+模型 Credit Suisse First Boston 于 1997 年研发的一种建立在保险精算法基础上的信用风险模型,但 Crouhy 等(2000)觉得 Credit Metrics 和 Credit Risk+两种模型缺乏预测的实时性,因为两种模型框架类似且均以历史数据为基础进行预测;Credit Portfolio View 模型是以实证统计研究为基础,进行宏观模拟的信用风险模型,但经 Oskari 等(2008)使用压力测试进行研究表明,该模型并不能理想的真实反映银行资本约束情况。从国内学者来看,王珂等(2016)认为 Credit Metrics 模型具有相同的信用等级转移概率假设在现实中是难以估计的,在其提出模糊 Credit Metrics 模型后通过实例分析验证了其有效性;吕志华等(2011)年从 Credit Risk+模型的前提采用泊松分布出发验证,理论上证明 Credit Risk+模型采用泊松分布计算会高估实际信用风险,加大估计误差;曹裕等(2017)认为在考虑信用风险因素时,CPV 模型只从宏观经济因素对主体信用风险有所侧重导致计算结果存在误差。

综上所述,从古典信用分析法到现代信用风险量化管理模型,信用风险度量经历直观定性分析、概率化定量分析到引入期权定价模型,基于现代金融理论基础信用风险量化管理模型更能度量集中信用风险。从信用风险度量模型选择上来看,现代信用风险量化管理模型优于其他模型, KMV 模型以资本市场为数据,是一种实时动态模型,更适合本文的研究。

1.2.2 KMV 模型信用风险度量研究

KMV 模型采用欧式看涨期权模式,把权益价值 Q 和公司负债 F 分别看作是公司资产的欧式看涨期权和看涨期权协议价格,若债务到期时 $Q \geq F$,则股东执行看涨期权,因此不会发生违约,反之, $Q < F$ 违约发生。该模型的有效性经过国内外学者的大量验证,较为可靠。Morris 等(2005)经过实证分析,认为 KMV 模型的信用风险判别及预测能力具有突出性。Dwyer and Woo(2007)使用 KMV 模型对投资信托和抵押贷款总共 210 家机构进行研究论证,结果表明其中某些机构在一年内的违约风险概率超过了 10%,尤其是新世纪金融公司表现明显,它在违约风险事件发生之前的一年里面信用质量下降急速,甚至超过了 90%的同组分位数,也最终证明 KMV 模型能够有效辨别信用风险具有问题的次级贷款机构。Jackson 等(2013)分别使用 KMV 模型、单一模型和 Z 评分模型预测部分公司信用风险违约情况,结果表明,相较于其他两种模型, KMV 准确率更高。同样, Zhao 等(2016)也证明了 KMV 模型可以更好的准确识别信用风险,在其对电子银行业务是否会存在风险进行的研究中发现, KMV 模型是很好的衡量指标,且同时证明违约距离跟不良负债率的反向关系。曾诗鸿和李萌(2014)在辨别节能行业上市公司信用风险状况的研究中发现 KMV 模型确实具有强区分度。杨志安等(2022)基于 KMV 模型测算地方政府债务违约风险,研究结果显示我国地方政府债务风险整体来看尚在安全范围内,但是从时间演变来看波动较大且 KMV 模型具有较好的度量效果。凌江怀和刘燕媚(2013)选择 10 家上市银行作为研究样本,将样本数据代入 KMV 模型,再根据 KMV 模型得出的违约距离和违约概率与评级权威机构标普得出的结果对比,发现两者并无区别,几乎一致。而关晓宇等(2023)同样验证了此结果,他们在研究环境违约跟公司债务违约两者之间关系的时候,首先使用 KMV 模型和 PSM 方法两种论证方法检验环境违规行为对信守债务履约的影响,最后得出环境违规跟债务履约两者之间影响不大,

甚至环境违规被处罚的公司，其信守承诺履约竟优于未违规公司。

综上，KMV 模型在信用风险度量方面得到国内外学者研究认同，且在信用风险度量方面具有较突出的准确率。

1.2.3 KMV 模型修正和应用研究

KMV 模型的修正集中于结合我国的实际情况进行研究，将 KMV 模型与其他模型进行融合研究对我国公司信用风险度量适用情况，且随着计算机和网络技术的飞速发展，部分计算机技术也融合进 KMV 模型中来。从对 KMV 模型修正以及融合来看，研究有：Zhang & Shi (2016) 粒子群优化和极大似然估计法引入 KMV 模型建立了 PSO-KMV 模型并进行实证分析，结果显示：PSO-KMV 模型的性能显著优于 KMV 模型。B.J Christensen (2012) 在研究核密度与 GARCH 模型两者结合的基础上，加入局部多项式，最后得出结论修正后的模型效果更好。J.C Duan (2017) 提出新的假设来修正 KMV 模型违约点参数，在假设违约点都服从相似规律基础上，使用密度递增扩展数据序列蒙特卡罗方法来修正违约点参数可以更好的获得拟合效果。谢远涛等(2014)选择我国上市公司作为数据样本，对波动率采用 GARCH(1,1) 模型来进行重新估计得出新的股权价值波动率，最后该方法效果较好，相较于静态估计法，对我国上市公司具有更高的适用性。在这一点上，孙亮和吕丹妮(2021)具有不同的看法，他们认为由于市场存在非对称效应，所以 GARCH 方法本身在运用过程中是存在缺陷的，市场的非对称效应导致对数收益率会存在“肥尾”现象。谢远涛和杨娟等(2018)也试图对模型进行修正，但在使用同等的房地产样本基础上，分组度量并比较，发现使用违约距离 DD 和违约概率 EDF 来度量方法不可行。王佳等(2022)则采用新的方法进行修正，引入跳跃的风险资产价格构建跳跃-扩散 KMV 模型，最后证明该方法可行且拟合我国上市公司信用风险较好，结果与我国实际情况相差无几。张慧和王宁等(2021)构建分形 KMV 模型以测度 31 个省份地方政府债券的安全发行规模，然后利用 GM(1,1)与回归模型得到财政收入预测值，结果表明我国政府债务发行信用风险较小。王莉(2017)发现传统模型中违约概率跟违约距离之间并不能进行函数关联，于是利用理论违约点的设置解决了该问题，最后证明修正违约距离的 KMV 模型通过不同违约利率对应的违约成本影响财务危机。谢赤等(2018)同样也通过修正的 KMV 模型来测度银行 ABS 产品不同期的信用风险，

并使用 MST 算法考察信用风险在其之间的传染机制，最后得出政策性银行和大型商业银行的 ABS 产品不同时期均处于信用低风险水平。

KMV 模型在多个行业也都有应用。孙亮等（2021）分别采用 GARCH、TGARCH 等波动率预测方法对传统 KMV 模型股权价值波动率进行重新估算，并计算出最终结果，通过对比各波动率预测方法得出的结果进行比较分析，修正后的结果相较于传统结果公司风险更高。潜力和冯雯静（2020）也认为 KMV 模型结果可信，并使用该模型对地方政府在 2020 年到 2023 年发行的专项债券经营测算并预测其违约风险概率。同样，张建同等（2018）和陈致远等（2019）都分别应用方法修正传统 KMV 模型，并使用修正后的 KMV 模型对样本公司进行预测，结果均都显示修正后的模型可以有效识别信用风险。王慧等（2018）则考虑修正后的 KMV 模型在上市房地产公司间的识别率，选取 2009 年到 2016 年间 147 家公司作为样本，在传统和修正的 KMV 模型计算下，表明修正后的 KMV 模型效果更好。在制造业方面，赵平和孙志峰（2022）在论证中国银行业贷款利率是否完全反映企业违约风险水平基础上，利用 KMV 模型进行实证分析，最后结果得出，如若在名义贷款利率完全开放的背景下，制造业上市公司基本上遵循信贷风险定价原则。

1.2.4 文献评述

KMV 模型在现代信用风险量化管理模型中具有较高的准确率，其度量信用风险的能力得到大量国内外学者的论证，国内学者也将 KMV 模型运用于各行业的信用风险的度量研究。从国内外文献综述的研究来看，KMV 模型进入了跟其他模型结合进行参数修正的发展历程，为此在结合我国金融市场发展情况基础上对 KMV 模型的参数进行针对性改进是可行的。正如孙亮和吕丹妮（2021）研究表明，对股权波动率参数修正部分学者使用 GARCH(1,1)模型进行修正可以完美解决股票收益率“尖峰厚尾”中的“尖峰”问题，但无法解决“厚尾”问题，基于此，本文将考虑使用 EGARCH 模型来解决股票收益率的“尖峰厚尾”问题；其次，对违约点的参数修正问题上，国内部分学者使用对“长期负债”的系数采用穷举法或密度递增扩展数据序列蒙特卡罗方法进行修正取得比较好的效果，随着机器学习方法在金融领域的运用，本文考虑使用遗传学算法（GA）对违约点进行修正。因此，基于 EGARCH 模型以及遗传学算法（EGARCH-GA-KMV 模

型，同改进型 KMV 模型）分别对股权价值波动率和违约点进行改进，并在综合比较的基础上，论证改进后的 KMV 模型是否具有信用风险识别能力可研究性。

1.3 研究方法 with 内容

1.3.1 研究方法

文献参考法。充分整理及阅读信用风险管理的国内外相关文献，全面了解现代信用风险度量模型，探寻 KMV 模型最新研究成果。系统分类梳理国内外学者 KMV 模型的研究成果。因此，本文将清晰展现国内外学者的相关研究成果，奠定本文进一步探讨研究及创新的基础。

实证分析法。通过应用遗传学算法和 EGARCH 模型对 KMV 模型的参数进行针对性优化，并在我国 80 家上市公司连续 5 年间的样本数据上进行实证分析和检验。比较改进后的 KMV 模型与传统 KMV 模型在衡量上市公司信用风险方面的效果，并做进一步详细分析。

1.3.2 研究内容

本文基于已有研究，探索运用参考资料、学习关于信用风险管理工具和技术等方面的信息，明确适用于我国各种模型的情况。根据我国独特的情况，对 KMV 模型进行升级与改进，旨在提升上市公司信用风险管理的效能。全文分为五个章节，内容如下：

第一章为绪论。先叙述本文的研究背景与意义，再在充分梳理文件基础上阐述国内外文献研究成果，厘清目前的研究基础，为本文搭建研究框架、方法及探寻创新点等构建合理前提。

第二章为相关理论基础和信用风险度量模型。首先为下文建立理论框架，包括信用风险、期权定价、遗传学算法等理论基础；其次通过对信用风险度量模型发展历史和模型对比进行分析，为下文选定 KMV 模型奠定基础。

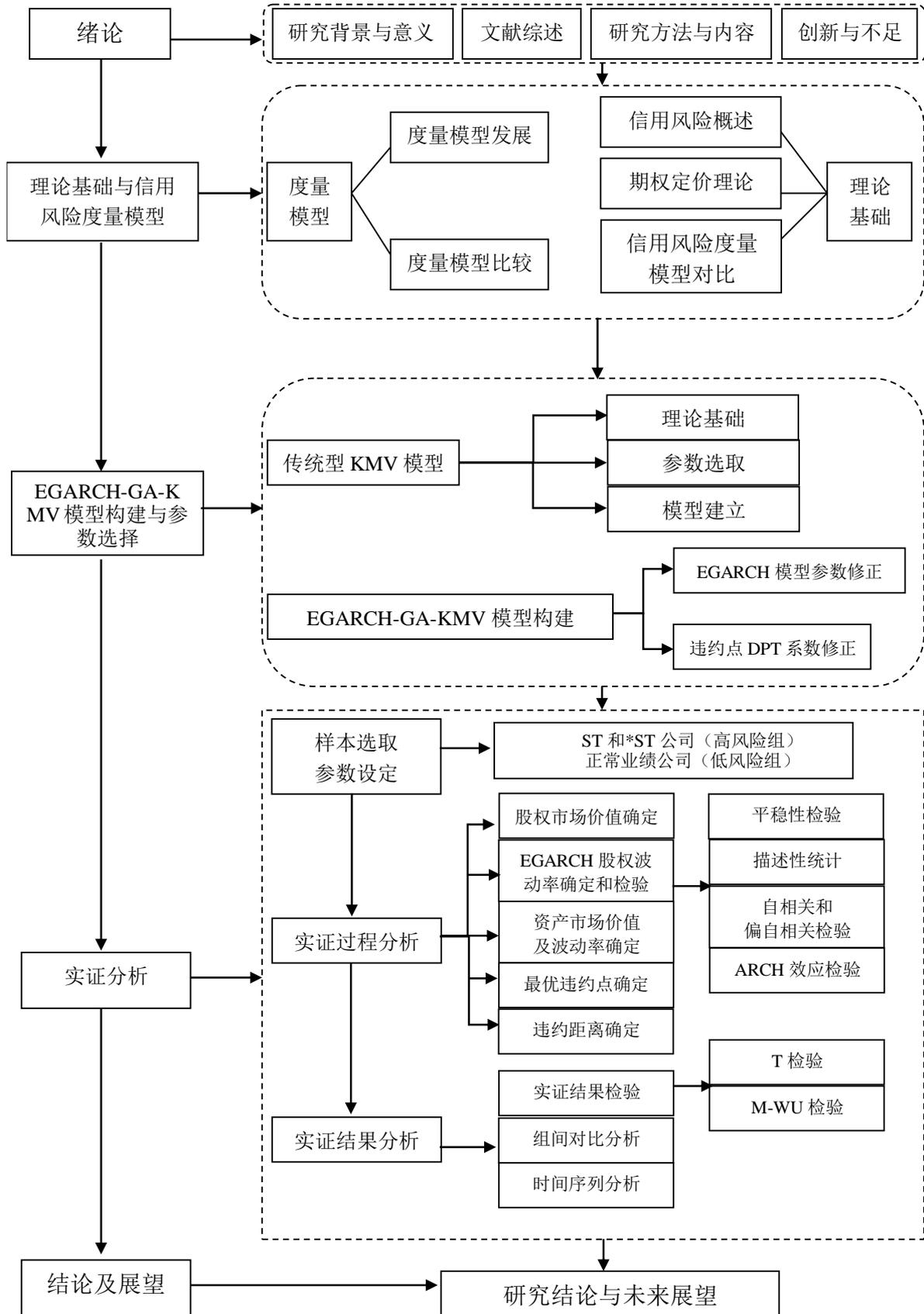
第三章为 EGARCH-GA-KMV 模型的构建与参数选择。先对 KMV 模型的理论基础进行论述以及定义参数选取，最后论述本文所研究主题模型的建立以及计算过程。

第四章是实证分析部分。该章节旨在优化 KMV 模型以适应我国金融市场的发展。首先详细描述了选择样本数据、处理流程和参数设置的方法。其次，使用

遗传学算法确定模型基于样本数据的违约点最佳系数；其次，使用 EGARCH 模型重新估算 KMV 模型股权价值波动率并对结果进行检验；再次，使用优化后的 KMV 模型推导各上市公司违约距离；最后对实证结果进行综合比较分析。

第五章为全文总结。基于全文的基础上进行整理和归纳，同时说明设计的不足以及对未来进行展望。

1.3.3 研究框架



1.4 创新与不足

1.4.1 创新之处

通过对国内外学者研究分析，本文存在的研究创新点如下：

(1) 使用遗传学算法优化违约点 (DPT) 系数。利用遗传学算法对该模型的违约点系数进一步调整，以解决传统违约点确定方法的局部缺陷。为了提高结果的准确率，将使用样本数据代入遗传学算法，在确定为最佳适应度函数基础上输出最优结果。

(2) 对模型中股权价值波动率参数进行重新估算。KMV 模型的传统参数选取中，股权价值波动率来源于股票收益率静态估算的历史波动率，该方法应用不了我国特殊的金融市场状况，我国金融市场并不满足传统假设中股票收益率呈现正态分布且为白噪声的前提。部分学者探索使用 GARCH 模型来修正误差，但该方法未考虑我国金融市场股票收益的“杠杆效应”，杠杆效应即收益率跟未来波动之间强显著的负相关性。因此，为了克服“杠杆效应”造成的“肥尾”现象，这里考虑使用 EGARCH 模型对股权价值波动率进行重新估算。

1.4.2 不足之处

基于理论基础、实证分析模型和内容框架，本文存在以下可能不足：

(1) 在使用遗传学算法确定最优违约点 (DPT) 过程中，数据使用的是选取的开发样本，得出的结论并不能使用某个上市公司直接套用计算，同时，由于我国没有公司违约数据库来确定最优违约点的取值，所以基于遗传学算法计算出来的取值是基于样本计算的最优取值。

(2) KMV 模型自身的缺陷。虽然本文是对传统 KMV 模型进行改进，但是 KMV 模型的基本原理主要是从股票市场信息和财务指标信息角度进行分析，可能会受到数据不全面、财务造假、缺乏长期考虑等事件因素影响。

2 相关理论基础与信用风险度量模型

2.2 相关理论基础

2.2.1 信用风险理论

信用风险是金融领域一个重要的概念。从债权主体来看，商业银行是贷款的大头，可能成为信用风险损失的最大主体，信用风险也贯穿银行业务的始终，从债务主体来看，如若债务主体选择信用违约，将牵一发动全身，风险将随着直接链条或者间接链条风险传播至经济体系，因此，管理信用风险以及发现潜在的信用风险就显得尤为重要。

1. 信用

从信用的基本含义来看，信用具有三大基本含义，第一，信用是基本道德规范，指日常交往中信守承诺行为守则；第二，信用是经济活动的基本要求，指在授信人对受信人还款承诺的信心基础上，无须付现就能获得交易成功的能力；第三，从法律角度来看，信用作为一种法律制度，遵守信用即合法依法实现利益期待，违反则应当承担相关法律责任。

除基本含义之外，信用还具有广义和狭义之分。从广义角度来看，广义信用是指一个人或实体在经济活动中所拥有的声誉和信誉，包括其在社会中的信任度和声望。在经济领域，广义的信用还可以指代一个人或实体的还款能力和履约能力，以及其在市场中的信誉和形象。广义的信用不仅仅是一种经济概念，更是社会交往中的一种重要价值观。从狭义的角度上讲，狭义信用则是指在金融领域中，借款人按时履行借款合同义务的能力和意愿。狭义的信用主要关注借款人的信用记录、信用评级和信用分数等指标，用于评估借款人的信用风险和确定借款条件。在金融市场中，信用是一种重要的资产，能够直接影响借款人的融资成本和融资条件，银行和金融机构通过评估借款人的信用状况来决定是否向其提供贷款，以及贷款的利率和额度等具体条件。因此，维护良好的信用记录对于个人和企业来说至关重要，可以为其未来的金融活动提供更多的机会和便利。

2. 信用风险成因

信用风险也称违约风险，指受信方未能按时履行合同约定的责任，从而给授信方带来价值损失的风险。信用风险的成因从主客观方向讲主要有以下几种。

(1) 市场风险。市场风险是系统风险，不可分散的风险，对于公司主体来说，该风险来自于公司外部，不能利用资产组合分散的风险。包括宏观经济走势、国家经济政策变动、财政税收政策改革等从外部因素影响市场风险，外部因素影响整个市场经济运行，导致资金价格的市场波动，从而影响资金持有者的市场收益，进而导致其还款能力及还款意愿降低，如此发生信用风险。

(2) 经济周期的影响。客观上面，经济周期影响发展大环境，当经济处于扩张周期时，社会总体经济情况好，债务主体经营较好，资本充足，有较强的盈利能力及还款意愿，信用风险降低；当经济处于下行周期时，社会总体盈利情况恶化，债务主体不能按时还款的因素增加。

(3) 特殊事件的发生。特殊事件和经济周期在很大程度上是无关的、相互独立的，但是与受信方发生信用风险具有很大的关联。特殊事件发生的因素诸如债务主体人（受信方）内部的财务状况恶化、道德风险等，外部的战争、政策环境变化、市场环境变化等都会导致债务主体经营状况恶化进而产生信用风险。

信用风险的产生因素是多种多样的，其产生于经济交易活动的各个环节，基于信用风险产生的主观及客观原因导致信用风险产生不可避免，因此，如何防范以及管控信用风险是个重要议题，但如若能对作为市场经济主体的公司进行信用风险预测，再加以管控则能有效的减少信用风险的发生。

其次，按照风险的发生形式，信用风险分为结算前风险、结算风险两种。结算前风险是指在交易结算前未能完成交易而产生的信用风险，如贷款、证券、衍生品交易等的违约；结算风险则指在结算过程中，结算工作失误或者不符合结算规定造成损失的一种风险，这种风险具有短期性。根据目前的定义，信用风险广义指代结算前风险。

3. 信用风险特征

信用风险成因受主客观方面原因影响。客观上面由于经济周期性特征影响经济发展大环境，当经济处于下行周期时，债务主体经营不佳，信用风险增加。主观上面债务主体（借款人）违约主要是造成借款人经营有影响的特殊事件发生，这些因素包含财务状况恶化、政策环境变化、市场环境变化等都会导致借款人违约增加。同时，信用风险具有如下特征。

(1) 损失收益的非对称性。债权主体在授信贷款的过程中将会面临预期

收益和预期风险的非对称性，非对称性在信用合约生效之前就已经存在，债权主体需要承担债务主体的道德风险，如若债务主体正常还款付息，那么债权主体将会得到预期收益，但是当债务主体选择违约时，债权主体将会收到比预期收益更的损失。也就是说，债权主体的收益是有上限的，而风险是无限的，风险跟收益的非对称性造成债务主体的信用风险概率呈现“左侧肥尾”现象。

(2) 信用风险具有累积性。信用风险违约事件不是突然发生的，而是存在前期风险的。信用风险将不断累计、恶性循环、发生连环反应，当信用风险超过债务主体临界点的时候会突然爆发造成金融危机。因为信用风险累积性特性，在信用风险达到临界点之前早预防早发现，会降低信用风险发生的预期损失。

(3) 信用风险是非系统性的。非系统性风险相对于系统性风险而言，系统性风险是由政治、经济、社会等因素整体把控的，系统性风险不化解。相对于经济周期波动性风险、利率风险、汇率风险等系统性风险，信用风险受债务主体主观因素影响，债务主体为不损失利益，可能会隐瞒真实影响，因此信用风险观察数据少且不易被察觉，具有非系统性的特征。

信用风险除了以上主要特征之外，还具有潜在性、长期性、破坏性、可控制性以及控制艰巨性等特点。

2.2.2 期权定价理论

KMV 模型以现代期权定价理论为基础，引入欧式看涨期权和看跌期权，分别将企业股权看作是看涨期权、将债权看作是看跌期权，进而得出股权市场价值和波动率，同时把投资回报率定义为无风险利率，据此得出 KMV 模型的预测违约率。1973 年，布莱克（Fischer Black）和斯科尔斯（Myron Scholes）共同推出著名的布莱克-斯科尔斯（B-S）期权定价模型。几乎同时，默顿（Merton）也发现了相同的定价理论，并在布莱克和斯科尔斯的基础上优化期权定价模型，于 1974 年发表出来建立了 BSM（Black- Scholes- Merton）期权定价模型。因此，期权定价理论的演进史包含两个重要部分，那就是 Black- Scholes 模型和 BSM 模型。

1. Black- Scholes 期权定价模型

由于期权交易方式、方向和标的物等方面的不同，期权分为看涨期权和看跌期权。买方在期权的合约期内有权按照执行价格买入特定数量标的资产，这被称

为看涨期权；卖方则有权在期权合约期内以特定执行价格售出资产，这便是看跌期权。所以期权是一种选择权，期权的买方或卖方的权利和义务是不等价的，而期权价格则是期权合约中的核心问题。因此，布莱克和斯科尔斯以看涨期权作为切入点探索期权定价方式，也就是 Black-Scholes 期权定价模型。该模型表明，只有股价才与未来预测有关，其他皆无关；期权定价的因素极其复杂，合约期限、无风险利率水平、交割价格等都是影响期权定价的因素。

1973 年，布莱克和斯科尔斯依据看涨期权推导出期权定价模型，其推导需要符合较为严苛的假设，假设及前提条件如下：

(1) 市场是完全有效的。该模型假设市场不存在困难或限制，所有交易都是无困难的，并且所有信息都是公开和免费获取的。

(2) 资产价格的对数正态分布。该模型假设资产价格的对数服从正态分布，这意味着资产价格的波动是随机的且连续的。

(3) 无风险利率是固定的。该模型假设无风险利率在期权的整个期限内是固定不变的。

(4) 无交易费用和税收。该模型假设在交易时不存在交易费用或税收。

(5) 标的资产的价格遵从几何布朗运动。

(6) 无债务支付。该模型假设标的资产在期权达成之前不支付任何支出。

(7) 期权为欧式期权。该模型默认期权为欧式期权，即期权到期前无计可施，必须到期才能行权。

综上前提和假设，推导出看涨期权的 B-S 期权定价模型公式为：

$$C(S, T) = S \cdot N(d_1) - K \cdot e^{-rT} \cdot N(d_2) \quad (2-1)$$

$$d_2 = d_1 - \sigma\sqrt{T} \quad (2-2)$$

$$d_1 = \frac{\ln \frac{S}{K} + rT - \frac{\sigma^2 T}{2}}{\sigma\sqrt{T}} \quad (2-3)$$

如果是看跌期权，B-S 期权定价模型公式为：

$$P(S, T) = K \cdot e^{-rT} \cdot N(-d_2) - S \cdot N(-d_1) \quad (2-4)$$

其中， C 为看涨期权当前价格， P 为看跌期权当前价格； S 是股票当前价格， T 是期权的到期时间， K 是期权的行权价格， σ 股票的波动率， $N(x)$ 表示标准正态分布的累计概率分布函数，指代行权概率。

根据 Black- Scholes 期权定价理论，在无风险的对冲交易中，当企业债务到期时，若资产市场价值小于负债，则企业会选择违约，相反，则企业选择偿还负债。这一条理论也为 KMV 模型的建立奠定了基础。

2.BSM 期权定价模型

BSM 期权定价模型是默顿 1974 年在 Black- Scholes 期权定价模型的基础上建立起来的，是对布莱克-斯科尔斯期权定价模型的完善。从公式的推导过程来看，两者并没有巨大差别，但值得一提的是，在模型的假设上，BSM 期权定价模型比 Black- Scholes 期权定价模型多了一条假设，即标的资产可以支付股息。这个假设在实际市场中更为贴近现实，因为很多标的资产（如股票）在期权到期前会支付股息。BSM 模型通过考虑股息支付对期权价格的影响，提供了更准确的定价结果。

从企业的资产负债表来看，企业的资产由股东和债权人分别拥有的股权价值与债务价值两部分组成，公司的责任是有限的，股东承担的最大损失为股权价值，这就意味着股东承担的损失风险是有限的，而只要对债权人的债务按时偿还，股东就享有剩余资产价值的支配权。因此，股东的价值就可以用看涨期权的价值来代替。

假设 V_T 为公司债务到期时的资产价值， D 为公司的债务面值。则当公司的债务到期时，债权人的最大损益额为 $[V_T, D]$ ，股东的最大损益额为 $[0, V_T - D]$ 。其构成如下图 2.1 和图 2.2 所示。

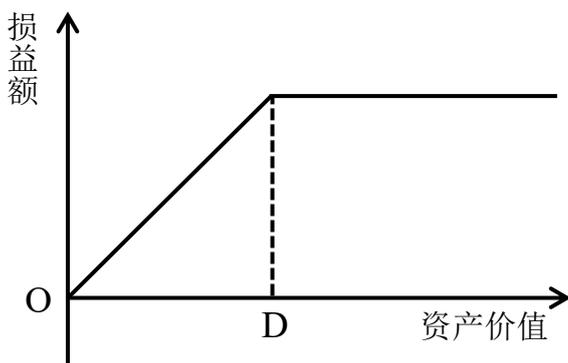


图 2.1 债权人到期损益图

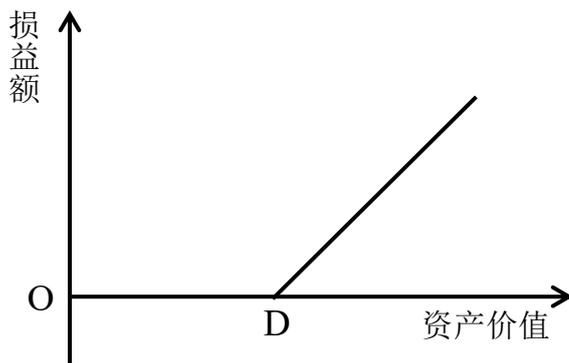


图 2.2 股东到期损益图

默顿发现个人持有公司股权所获得的收入跟个人买入一份以公司资产作为标的欧式看涨期权是一个恒等态射。于是，默顿把公司股权看作是一份以公司资产为标的的欧式看涨期权，公司债务账面价值作为该份看涨期权的执行价格，期权的执行期限则为公司债务的期限，再将此原理代入 Black-Scholes 期权定价模型，就推导出股票市场价值跟公司资产价值两者之间的关系，继而推出BSM期权定价模型。

$$E = V_A \cdot N(d_1) - D \cdot e^{-rT} \cdot N(d_2)$$

$$d_2 = d_1 - \sigma_A \sqrt{T} \quad (2-5)$$

$$d_1 = \frac{\ln \frac{V_A}{D} + (r + \frac{\sigma_A^2}{2}) \cdot T}{\sigma_A \sqrt{T}}$$

其中， E 为股权价值； D 为债务价值； V_A 是公司的资产价值； T 是公司债务的到期期限； σ_A 公司资产价值的波动率； r 为无风险利率， $N(x)$ 表示标准正态分布的累计概率分布函数。

2.2.3 遗传学算法

遗传学算法 (Genetic Algorithm, GA) 是由 John Holland 于上世纪七十年代提出，根据达尔文“物竞天择，适者生存，优胜劣汰”的自然择优选择及生物进化遗传理论的精髓，在数学计算中插入该精髓来模拟生物进化的计算过程。该算法使用数学方式进行模拟仿真计算，将数学中计算的求解问题换个角度转变成生物

的染色体基因交叉、感染、变异等的遗传问题，最终根据择优遗传取得最优解。将一个多项式的复杂组合取值优化问题，变成为相对常规的数学计算问题，该方法目前已经广泛应用于机器学习、组合优化求解、工程设计等各项领域。

遗传学算法通用框架如图 2.3 所示：

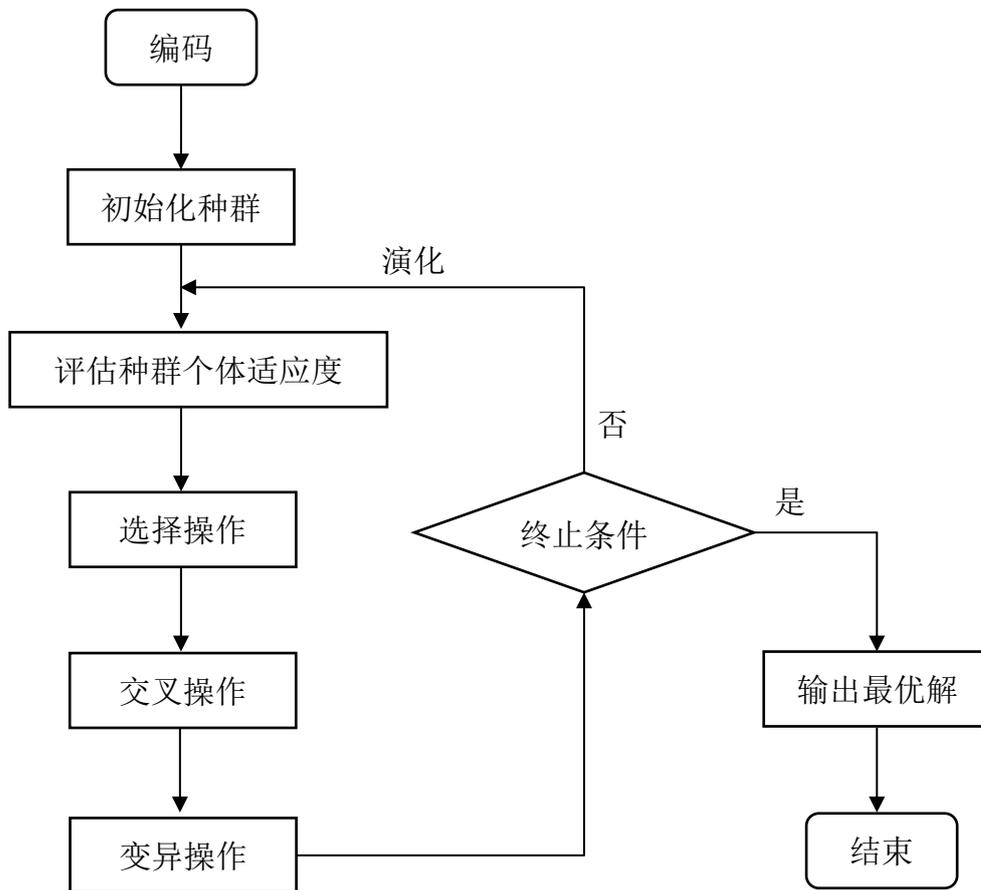


图 2.3 遗传学算法流程图

1. 遗传学算法基本原理

遗传学算法是模拟生物进化过程进而求得全局优化解的一个计算过程。首先对数据样本进行编码生成初始种群，初始种群由遗传基因中的个体染色体组成，对可行解进行编码可以得到该个体。随着时间的演进，通过设置适应度函数来评估每个种群的优劣度，将适应度差的个体进行“优胜劣汰”，借助“物竞天择”原则，选择基因较好的，也即适应度函数取值最优的个体遗传至下一代。这些个体经过优化操作（选择、交叉、变异）生成新下一代初始种群，如此反复操作，每一次迭代都是对上一代最优结果的延续，如同“适者生存”进化一般，随着迭代进化的进行，最优解得出。因此，遗传学算法简单来说就是一个对可行解进行

重新编码，最终群体迭代优化的过程。

遗传学算法通用执行程序如下：

(1) 字符编码

字符编码即把字符编码为指定集合中的某一对象，把问题可行解从解空间处理成计算机算法所能理解的搜索空间的一种处理方法。在计算机中，我们在显示器上面看到的图片、文字、数字等信息存储并不是我们所看到的样子，而是通过进制编码成字符形式保存在计算机存储器中。同样，在遗传学算法中，进行多维或多极值函数、高精度的连续函数优化求解，使用二进制编码无法得到最优解，二进制编码具有局限性，而浮点编码则克服了该缺点，也就是说本文中所指代的编码使用的是浮点编码方式。

(2) 初始化种群

字符编码完成之后，通过设置条件，初始化一定数目的个体作为初始种群，每个个体都代表一个可行解，也都具有“基因”携带染色体。初始化个体数目可以自由设置，一般 20-100 个个体即可，在此次计算中，迭代的种群个体遍历所有，使用矩阵方式来进行最优组合。

(3) 适应度评价

适应度评价即是对个体优劣设置适应度函数来评价每个个体的适应度值，当前某一代种群中能够遗传到下一代的个体是跟适应度函数值成正比的概率来决定的。适应度评价在遗传学算法过程中代表的是“优胜劣汰”的标准，达到适应度最优的个体则携带其“基因染色体”通过一定的感染率传播至下一代，如此反复直至达到终止条件最优解。

(4) 遗传算子

遗传代表的是选择操作、交叉操作和变异操作，遗传算子具有三种形式，分别是选择、交叉、变异。选择算子也即有概率的选取部分个体进入到下一代生成的新种群中的操作；交叉算子即有概率的两两交叉交换基因的操作；变异算子即表达个体基因遗传至下代时的变异概率。

(5) 终止条件判断

检验是否得到停止计算标准，如果满足则停止计算，输出最优解并返回结果，整个过程结束。设置终止条件目的在于利用遗传学算法计算过程中得出来的可行

解是靠近适应度函数值最优的解,该可行解在遇到多维计算可能有很多种组合方式,并不是唯一解,所以一直计算下去都会收敛于某个值范围内,因此需要设置终止条件目标。计算中,一般将终止条件判断设定为遗传迭代 200 代。

2. 遗传学算法特点

在计算复杂计算问题中,使用遗传学算法跟传统优化算法具有很大的区别,遗传学算法并不是从解开单个解出发,而是从串集开始搜索,覆盖面比较大,是全局的最优解,不容易误入单个局部最优解的局限。因此,遗传学算法具有以下特点。

(1) 自学习性。在遗传学算法执行计算过程中,需要给定适应度函数、最大遗传代数、种群规模、交叉和重组概率等要素,随着过程执行,适应度大的个体被保留下来,这使得个体具有一定交叉感染概率,这种自学习性的特点最大程度保留了自然选择的特点和规律。

(2) 高效率性。遗传学算法可以计算多维度、高精度的多元函数极值,并且可以通过高性能实用型进化算法工具箱来实现进化算法中各项重要操作库函数,通过模块化的进化算法框架来求解单目标优化、多目标优化、组合优化和复杂约束优化问题,尤其是在超大种群规模的进化优化中。

(3) 并行性。遗传学算法并不是从某单个解开始的,而是从问题可行解的串集开始。这样可以使得到的最优解是全局最优解而不是单个局部最优解。同时,遗传学算法在计算过程中可以被分为多个子任务并行进行,在提高计算效率的同时,还实现了并行化。

2.3 信用风险度量模型

2.3.1 信用风险度量模型发展

信用风险度量是对企业信用履约能力及信用信守程度进行预测偿还债务意愿及能力的一种评估方法。从信用风险度量发展历程来看,信用风险度量可以分为三个阶段:20 世纪 70 年代以前的古典信用分析法、20 世纪 60-90 年代的基于统计学的信用风险度量模型和面向 21 世纪的信用风险量化管理模型。

古典信用分析法又称专家分析法,主要通过使用经验丰富的专家主观判断作为评判信用主体的信用风险,通常有 5C 分析法、5P 分析法、5W 分析法、4F 分析法和 CAMEL 评估法等信用风险度量方法。古典信用分析法是一种定性分析法,

即逐一对要素估算并打分，量化受信人的信用，将此作为进行贷款及后续行为的依据。

基于统计学的信用风险度量模型克服古典信用分析法只能定性角度衡量信用风险的缺点，多元化了信用风险度量标准。通过引入财务数据和统计学方法，确定信用主体财务恶化的财务指标，据此判定信用主体的财务状况，确定其信用等级，进而评判其信用风险。该信用风险度量模型的代表是美国学者 Altman

(1968)首次利用多元判别分析，以 1940-1960 年代间的破产公司和对应数量的正常公司作为样本，利用 22 个财务指标生成 5 个比率变量加权汇总后产生总判断分指，建立起 Z-score 评分模型，后续学者并在此基础上建立了 Zeta 评分模型。随着统计学的进步，Logistic 回归模型、Probit 回归模型也相继建立出现。

21 世纪以来，全球贸易化进程加快，金融风险波动也随着全球化加剧蔓延，计算机新兴学科的发展，传统的古典信用分析法和统计学信用分析法已经不再适用于信用风险管理的要求，因此，面向 21 世纪的现代信用风险量化管理模型诞生。现代信用风险量化管理模型主要利用数理统计方法，使用建模来评估信用风险的概率密度函数。目前主流的信用风险度量管理模型主要有 KMV 模型、Credit Metrics 模型、Credit Risk+模型、Credit Portfolio View 模型等。

2.3.2 信用风险度量模型比较分析

从时间线来看，信用风险度量总共经历了三个阶段，随着统计学和计算机科学的发展，每一个阶段的发展都是在原有基础上的进步。基于统计学的信用风险度量克服了专家分析法在定性分析上的漏洞，并开始引入信用主体的财务状况，使得信用风险度量更加科学，而不是单纯依靠有经验的专家的主观评判。全球金融风险具有链接性、持久性、波动性、多元性等特点，现代信用风险量化管理模型更符合现代信用风险管理的要求，为此，这里只介绍对比现代信用风险量化管理模型。

1.KMV 模型

上世纪八十年代，许多学者在 Black- Scholes 期权定价模型和默顿的 BSM 期权定价模型基础上进一步完善且优化了期权定价模型，KMV 模型于 1993 年首次被作为一个正式模型概念来开始用于企业的信用风险评估，经过众多学者数次调整实验后，该模型于 1997 年正式用于评估企业信用风险的违约概率。换句话说，

KMV 模型是基于BSM期权定价模型理论框架被建立起来的，同时兼顾企业自身的财务实况和股票资本市场数据，对上市公司当前潜在的信用风险隐患进行评估。

在给定负债的前提下，根据 KMV 模型所表明的思想，信贷信用风险由受信方即债务人的资产市场价值定夺的。不过，市场并不能直接观察到资产的市场价值，因此从银行角度来看考虑信贷的偿还问题是不合理的，但若考虑企业偿还信贷的信用风险问题，该问题就变得简单易懂。理论层面上，KMV 模型权益价值当作是以公司资产为标的的欧式看涨期权，协议价格则为公司负债的账面价值，如果在债务到期时权益价值大于或等于负债的账面价值，那么股东将执行期权权益选择拥有公司，不会发生违约，反之，违约发生。由于使用的数据是股票资本市场交易数据和企业的实际财务数据，KMV 模型精度较高，可以通过更改使用数据来实时动态计算企业的信用风险，克服了以前过度依赖历史数据问题。此外，KMV 模型也具有一定缺陷，使用的财务数据跟股票市场交易数据需要企业披露，所以不能概括所有公司，只局限于上市公司。其次，KMV 模型前提假设较为严格，世界各国经济制度或发展模式有差异，KMV 模型并不适合所有世界所有国家或者地区。

2.Credit Metrics 模型

Credit Metrics 模型也称信用计量模型，于 1997 年由 J.P.摩根推出的用于量化信用风险的模型，该模型运用 VaR 框架，对信用贷款和未交易的资产进行信用风险评估。根据 VaR 理论，如果某已信用评级公司债务在置信度 95%下为 493 元，则表示在 95%可能性下该笔债务损失低于 493 元，反过来则在 5%可能性下损失超过 493 元，总之，置信度跟模型准确性成正比，这是该模型的基本原理。对于Credit Metrics模型而言，影响信贷资产价值的因素既有违约事件，也有信贷资产质量的变化，其基本思想是信用风险取决于债务主体的信用状况，而信用等级直接与信用状况挂钩；债务主体的信用工具（包含贷款和债券）的市场价值也直接与债务主体的信用等级挂钩。

从模型的技术框架来看，该模型主要包含三个关键环节。其基本框架如图 2.4 所示：

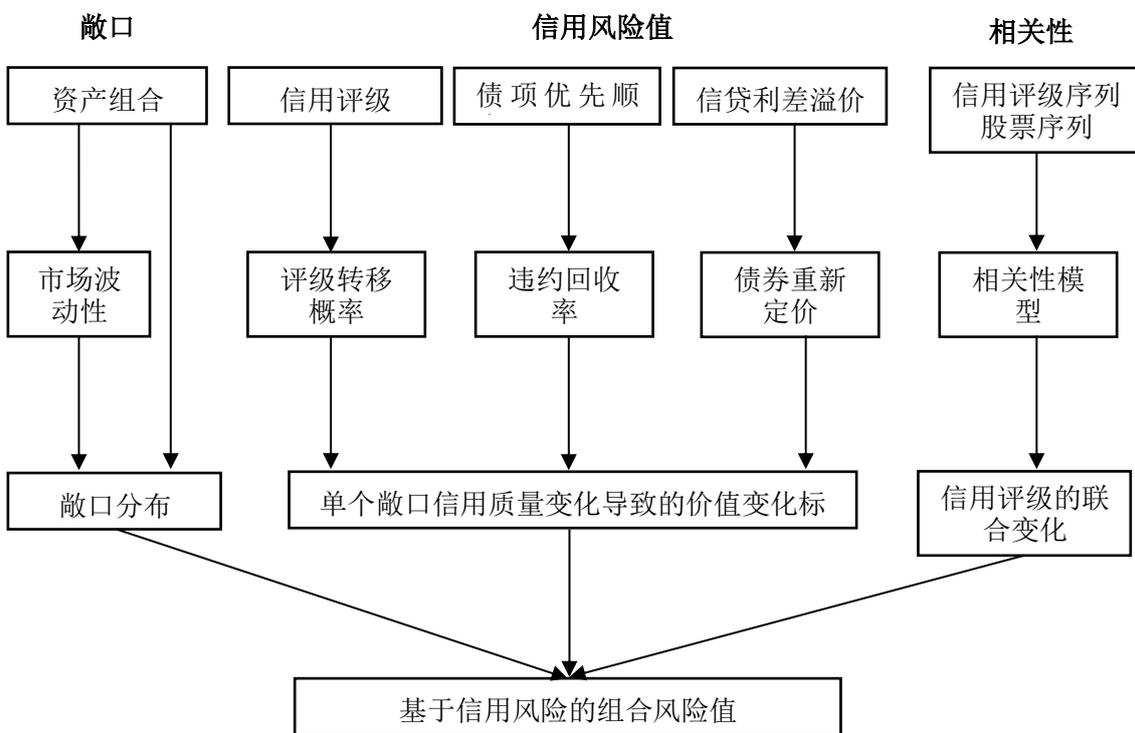


图 2.4 Credit Metrics 模型框架

虽然 Credit Metrics 模型即考虑了宏观因素，也考虑了微观因素，但是其核心在于根据企业信用等级评估来评价信用风险，也为该模型套上一定的局限。

3. Credit Risk+模型

在保险学的思想和方法的启发下，瑞士信贷银行金融产品部于 1993 年开发出基于财险精算的 Credit Risk+ 模型，以违约率和波动性计算投资组合。该模型单考虑信用风险两种状态：违约或者不违约，并以假设违约率是随机为前提来预测度量信用风险损失变化。根据该模型的思想，信用风险是市场风险的一种，所以信用风险变化也可以看作是市场风险的部分，且信用组合的风险价值跟信用利差的风险价值之间不具有关联性。在计算 Credit Risk+ 模型过程中，使用泊松分布概率密度函数来推理信用事件的违约概率，只需要将参数输入就可以获得所需的违约风险概率，使得该模型使用起来相对容易。

Credit Risk+模型由三个部分组成，包含信用风险管理、经济资本配置和积极的组合管理。信用风险管理模块需要设定相关的输入变量、违约率及波动性、风险暴露和回收率，违约率及波动性是根据信用评级统计得出的经验数据；经济资本配置模块是在一定置信水平下非预期违约损失配置的经济资本；积极的组合管理模块是根据风险偏好设计的限额系统和积极的组合管理。其中信用风险管理模块如下图 2.5 所示：

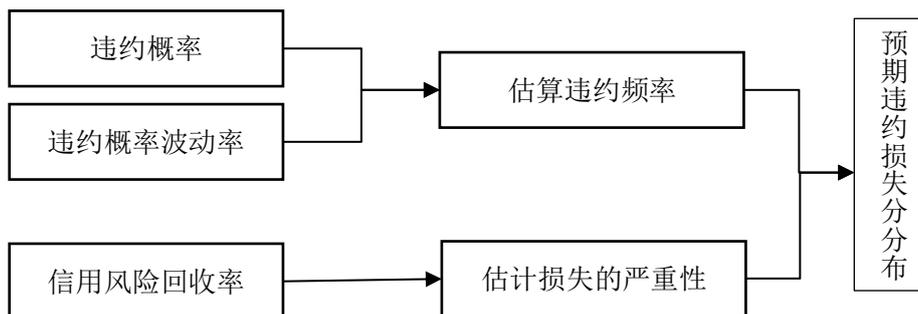


图 2.5 Credit Risk+模型信用风险管理模块

4. Credit Portfolio View 模型

Credit Portfolio View 模型也被称为 CPV 模型，是麦肯锡咨询公司 1997 年开发的用于衡量、评价违约风险的宏观经济模拟的离散多时期模型。该模型假定信用等级迁移概率在经济的同一阶段是稳定的，然而对经济周期比较敏感。为了避免误差影响，Credit Portfolio View 模型试图将如经济增长率、行业因素、利率和汇率等因素与转移概率间的关系进行模型量化，使用蒙特卡罗模拟技术模拟周期性因素的“冲击”来测定评级转移概率的变化。在本质上，该方法先是分析宏观经济周期波动的态势，再对其信用周期进行评价判断，最后得出信用风险。该操作的具体方法是，模拟并构建多种宏观经济环境下，一个国家内不同产业与不同信用等级金融工具违约的联合条件概率分布及位移概率。通过对这些分布和概率的分析，深入理解宏观经济形势的变动影响信用违约的概率及其位移情况。进一步地，评估不同行业或部门中不同信用级别的借款人所面临的信用风险程度，并据此绘制出信用损失的分布情况。

Credit Portfolio View 模型在实质上是对 Credit Metrics 模型的补充，因为它弥补了 Credit Metrics 模型里面不同时段评级转移矩阵保持不变的缺陷。尽管如此，

该模型还是有一定的不足，它从宏观角度出发却忽略了微观角度的影响，企业本身的资本市场及偿债能力等微观因素也会影响企业的信用风险。

5. 信用风险度量模型比较

现代信用风险量化管理模型更合适现代经济发展对信用风险度量的需要，但是每个模型都具有局限性，这在某种程度上限制了具体信用风险的度量。KMV 模型从股票资本市场和企业自身财务出发考虑企业权益价值和债务价值的违约阈值。Credit Portfolio View 模型虽是在 Credit Metrics 模型的基础上延伸发展而来，但都是基于信用等级评价不同程度的度量企业的信用风险，具有片面性。Credit Risk+ 模型受保险学思想和方法的启发，在计算违约概率和预期损失时，先要满足违约事件服从泊松分布的假设前提事实，即使加入违约概率波动率来尽量使模型符合现实情况，但仍然具有忽略信用等级变化、市场风险等局限性。因此，从现实市场情况出发，基于股票资本市场和企业自身财务状况出发的 KMV 模型更符合本文的需要，更适合度量上市公司信用风险。信用风险量化管理模型比较如表 2.1 所示。

表 2.1 信用风险量化管理模型比较

模型名称	应用环境	基本特征
KMV 模型	赋予企业破产和违约的理论基础，基于股票市场动态价格数据；使用基数衡量法评估和预测企业的违约事件	基于资产市场数据，动态计算破产概率和违约概率，数据受管理层影响小；不依赖样本，具有一般性。
Credit Metrics 模型	相比 KMV 模型差异较大；数据分析基础使企业信用评级变化及其概率的历史数据。	适用于小样本数据、所需参数少，仅需违约和风险暴露分布就能推导出违约概率。
Credit Risk+ 模型	受财产保险启迪开发；利用该模型可得贷款组合损失的概率分布	模型应用简单，仅需债务工具和风险暴露数据；仅考虑了违约风险，不考虑信用等级降级风险
Credit Portfolio View 模型	完善版的 Credit Metric 模型；非违约的转移概率则需基于调整的历史数据补充	因不使用历史数据，所以不能完整计量出等级转移矩阵

2.4 本章小结

本章主要对后续研究所需要用到的理论进行一个梳理，为下文构建 EGARCH-GA-KMV 模型搭建好理论基础，包括解释信用风险、期权定价模型、遗传学算法。从上文可知，KMV 模型是基于 B-S 期权定价模型跟 BSM 期权定价模型发展而来的，因此跟两个模型具有很大的关联性，也为后续结果处理提供依据；其次，遗传学算法可提供全局最优解取值，并不像传统数学统计，其提供最优解具有自学性等特点。最后，概括信用风险度量模型衍生发展，通过对比信用风险量化管理模型，将本文信用风险度量的切入点模型确定为 KMV 模型。

3 EGARCH-GA-KMV 模型的构建与参数选择

3.1 传统 KMV 模型的构建

3.1.1 KMV 模型理论基础

KMV 模型基本思想是 Black- Scholes 期权定价模型和 BSM 期权定价模型的信用风险度量模型，这一点在前面第二章的相关理论基础中有介绍。KMV 模型将公司的股票价值和债务价值看作是欧式看涨期权和看跌期权的衍生金融工具，在假设以无风险利率作为投资回报的风险中性程度中去预测企业的信用风险违约概率。也就是说企业的债务价值看作是一份看涨期权，该份看涨期权在债权方银行手里面则看跌期权，那么企业所拥有的资产市场价值就是期权标的，企业的债务市场价值就是执行价格。因此，企业是否行权，则跟企业的债务价值跟市场价值有关。如图 3.1 所示，我们把 OB 当作公司的债务市场价值，OA 是资产的市场价值，当企业债务到期时，此时假如企业的资产市场价值达到 OC 阶段，则 $OC < OB$ ，也就是企业的债务市场价值大于资产市场价值，企业如若行权，那么收益为亏损状态，因此企业将选择不执行权力，而选择信用违约不偿还债务；如若企业的资产市场价值达到 OD 阶段时，这时 $OD > OB$ ，也就是企业债务市场价值处于资产市场价值下方，企业将行使期权，偿还贷款后依然能够获得正的收益，企业会选择偿还贷款。

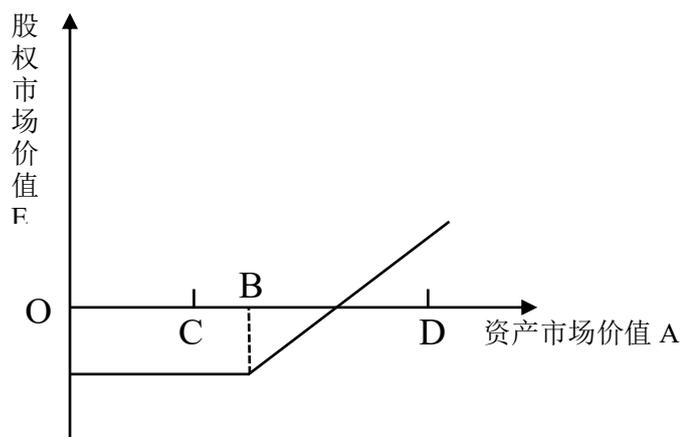


图 3.1 KMV 模型期权定价原理

BSM 期权定价模型可以看作是对 Black- Scholes 期权定价模型的发展，因此，KMV 模型是将 BSM 期权定价模型的理论应用于债务估值的一个信用风险监控

系统。KMV 模型的估值主要分为以下几步：第一，获取企业资本市场数据，这些数据包含企业股权市场价值及波动率、负债市场价值；第二，根据 KMV 相关关系式计算出企业的资产市场价值 V_A 及波动率；第三，根据 KMV 模型设定的为违约点 DPT（即企业选择违约的触发点）计算出违约距离 DD；第四，根据标准正态分布概率公式计算出违约概率 EDF。KMV 模型的框架图如图 3.2 所示。

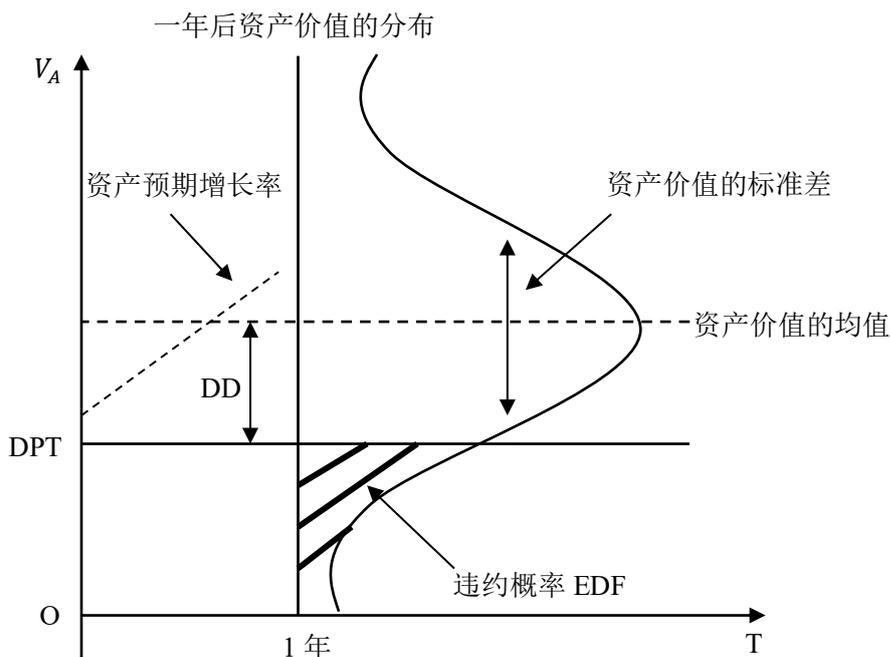


图 3.2 KMV 模型框架图

在 KMV 模型的框架图中，给定的时间为 T （一般为一年），假设企业资产市场价值服从一个标准的正态分布这个大前提。此刻债务的市场价值处于违约点 DPT 处，资产价值的均值到达违约点处的距离就是违约距离 DD，根据 KMV 模型的原定假设资产市场的均值线下降到表示债务的市场价值 DPT 以下时就会发生违约，因此，阴影处面积表示的就是违约概率 EDF。从上图可以看出，违约概率的大小的关键点在于违约点，也即表示企业债务市场价值，同时，违约距离 DD 越大，即资产市场价值均值跟违约点之间的距离越远，阴影处面积表示的违约概率 EDF 越小，也就是说违约距离 DD 跟违约概率 EDF 成反比。

除此之外，KMV 模型虽然在 BSM 期权定价模型基础上发展而来，但与 BSM 期权定价模型的假设有些许差别，KMV 模型假设如下。

(1) 符合 BSM 模型的基本假设, 即完美市场, 没有交易成本、没有税收; 无风险利率固定不变; 资本市场交易连续、可分且不存在套利机会等。

(2) 企业资产市场价值大于债务市场价值时不会违约, 反之则反。

(3) 企业的资产市场价值满足指数布朗运动。

(4) 企业的资产市场收益符合标准正态分布。

(5) 企业资本价值由所有者权益、长短期负债和可转换优先股组成。

3.1.2 KMV 模型参数选取

根据 KMV 模型的原理及假设, 该模型的参数包含股权市场价值 E、股权价值波动率 σ_E 、资产市场价值 V_A 、资产价值波动率 σ_E 、债务期限 T、无风险利率 r、负债 D、违约点 DPT、违约距离 DD 等。其中股权市场价值、股权价值波动率、负债都是根据企业的股票资本市场情况跟财务状况确定的, 无风险利率跟债务期限根据原定假设提取。在原定假设中假定企业的资产市场收益符合标准正态分布, 所以资产市场价值跟资产价值波动率由已知数根据 KMV 模型公式计算得出。下面介绍 KMV 模型中主要参数的选取及计算。

1. 股权市场价值 E

KMV 模型跟 BSM 期权定价模型相差无几, 都假设市场是完美的, 企业的所有股票都必须在市场交易, 没有限制销售情况存在, 为此:

$$\text{股权市场价值 } E = \text{股票价格} \times \text{在售股票数量}$$

2. 股权价值波动率 σ_E

KMV 模型在计算股权价值波动率 σ_E 时使用的是历史数据法。历史数据法假定未来与过去相一致, 以长期历史数据为基础推测未来数据, 据此在此处 KMV 模型也使用历史时期的波动率情况来预测下一期波动。同时, 根据 KMV 模型的企业资产市场收益符合标准正态分布假设, 因此, 在计算股权价值波动率 σ_E 采用对数收益法对股票日收益率取对数计算股票的日收益率, 进一步计算股票日波动率, 最后得到股权价值波动率 σ_E 年度数据。具体方法如公式 (3-1) 至 (3-4) 所示:

$$\sigma_E = \sigma_n \times \sqrt{n} \quad (3-1)$$

其中:

$$\sigma_n = \sqrt{\frac{1}{n-1} \times \sum_{i=1}^n (u_i - \bar{u})^2} \quad (3-2)$$

$$\bar{u} = \frac{1}{n} \sum_{i=1}^n u_i \quad (3-3)$$

$$u_i = \ln\left(\frac{S_i}{S_{i-1}}\right) \quad (3-4)$$

公式中， n 为股票交易天数， S_i 为股票第 i 个交易日的收盘价格， S_{i-1} 为股票第 $i-1$ 个交易日的收盘价格， \bar{u} 为一定时间内的股票的平均收益率， u_i 为第 i 日股票日收益率， σ 为股票收益日波动率。

3. 债务期限 T

KMV 模型使用历史数据法预测，因此根据一个会计区间一年为债务期限，使用前一年的数据去预测下一年度企业的违约概率，判断其是否会违约。

4. 负债 D

负债 D 即企业的负债价值，根据 KMV 原理负债 D 为长期负债跟短期负债之和。

5. 资产市场价值 V_A 和资产价值波动率 σ_E

确定了股权市场价值 E 、股权价值波动率 σ_E 、债务 D 、无风险利率 r 和债务期限 T 之后可根据 KMV 模型推定的公式联立方程组推导出来。公式如 (3-5) 至 (3-8) 所示：

$$E = V_A N(d_1) - D e^{-rT} N(d_2) \quad (3-5)$$

$$\sigma_E = \frac{V_A}{E} \times N(d_1) \sigma_A \quad (3-6)$$

$$d_1 = \frac{\ln\left(\frac{V_A}{D}\right) + \left(r + \frac{1}{2}\sigma_A^2\right)T}{\sigma_A \sqrt{T}} \quad (3-7)$$

$$d_2 = d_1 - \sigma_A \sqrt{T} \quad (3-8)$$

公式中， $N(x)$ 表示标准正态分布下的累计分布函数。

6. 违约点 DPT

根据 KMV 模型定义，违约点 DPT 为企业选择违约的触发点，即阈值，是企业选择违约还是偿还债务之间的一个临界点，当企业债务价值达到这个临界点时会优先选择违约。KMV 模型定义的传统违约点 DPT 等于短期负债和二分之一

的长期负债之和。公式如（3-9）所示：

$$DPT = STD + \frac{1}{2}LTD \quad (3-9)$$

公式中，STD代表短期负债，LTD代表长期负债。

7. 违约距离 DD 和违约概率 EDF

违约距离 DD，即根据图 3.2 中企业资产价值和违约临界值之间的标准偏差。和违约概率 EDF 概念相同，一样是衡量信用违约风险的指标。它代表的是资产预期价值和违约点 DPT 之间的距离相异于未来资产标准差的大小，即指代公司从当前正常状态到要发生信用违约所要求的最低资产收益额，其中违约点设在当前负债（见公式（3-9））的价值处。违约间距绝对值（即违约距离 DD）的公式如下：

$$DD = \frac{V_A - DPT}{V_A \sigma_A} \quad (3-10)$$

违约概率 EDF 是指企业在特定时期内信用违约风险发生概率的大小，在图 3.2 中的阴影部分面积即表示是违约概率 EDF，根据 KMV 模型定义及假设，资产收益符合标准正态分布，违约概率 EDF 跟违约距离 DD 具有如下关系：

$$EDF = N(-DD) = 1 - N(DD) \quad (3-11)$$

同时，根据 KMV 模型的计算，在选取一定时期企业违约的数据基础上计算出来的违约概率 EDF 跟违约距离 DD 之间成一个映射关系，发现违约距离 DD 跟违约概率 EDF 之间是反相关关系，该关系如图 3.3 所示：

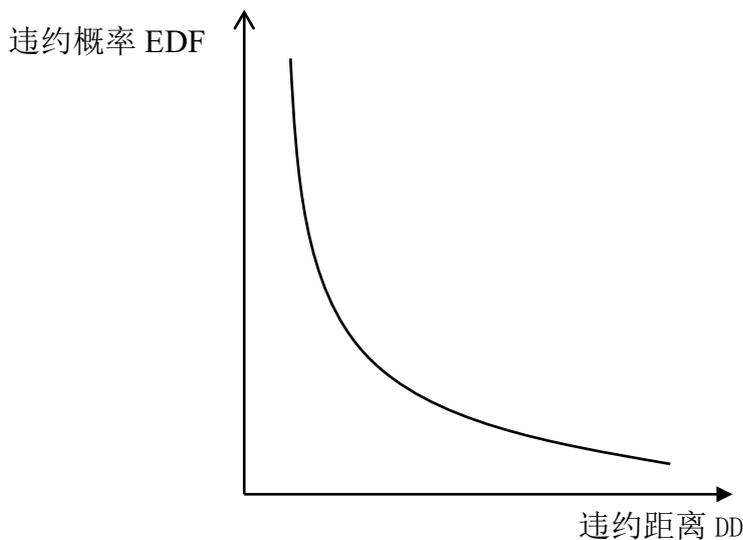


图 3.3 违约距离跟违约概率关系图

3.1.3 传统 KMV 模型建立

已知根据资本市场股票的市场价值、股价的波动率和负债估算出企业资产市场价值 V_A 及其波动率 σ_A 。

则设企业的资产市场价值为 V_A ，企业资产市场价值的变动为 dV_A 作为看涨期权的股价 E ，因为在 KMV 模型的基本原理中，股价被当作是以企业资产市场价值为标的的看涨期权，执行价格为企业的债务价值，则公式的变动过程为：

$$\frac{dV_A}{V_A} = udt + \sigma_A dW \quad (3-12)$$

公式中， u 、 σ_A 代表的是 V_A 的漂移率跟波动率， dW 则代表的是标准维纳过程，维纳过程是一个重要的独立增量过程，也即布朗运动过程，这符合 KMV 模型的假设定义，则定义股权价值为 E ，负债为 D ，无风险利率为 r ，根据 KMV 模型理论模型 BSM 期权定价模型推导可知：

$$E = V_A \cdot N(d_1) - D \cdot e^{-rT} \cdot N(d_2)$$

$$d_2 = d_1 - \sigma_A \sqrt{T} \quad (3-13)$$

$$d_1 = \frac{\ln \frac{V_A}{D} + (r + \frac{\sigma_A^2}{2}) \cdot T}{\sigma_A \sqrt{T}}$$

再根据伊藤引理：

$$dE = \left(\frac{\partial E}{\partial t} + \mu \frac{\partial E}{\partial V_A} + \frac{1}{2} \sigma_A^2 V_A^2 \frac{\partial^2 E}{\partial V_A^2} \right) dt + \sigma_A V_A \frac{\partial E}{\partial V_A} dD \quad (3-14)$$

接着：

$$\begin{aligned} \text{var}(dE) &= \left(\frac{\partial E}{\partial V_A} \times \sigma_A V_A \right)^2 dt = \sigma_E^2 E^2 dt \Rightarrow \\ \frac{\partial E}{\partial V_A} \times \sigma_A V_A &= \sigma_E E \Rightarrow \\ \sigma_E &= \frac{V_A}{E} \times \frac{\partial E}{\partial V_A} \times \sigma_A V_A \end{aligned} \quad (3-15)$$

由避险参数解开可知：

$$\frac{\partial E}{\partial V_A} = N(d_1) \Rightarrow \sigma_E = \frac{V_A}{E} \times N(d_1) \sigma_A \quad (3-16)$$

公式中， σ_E 指股价波动率，因为 E 、 D 、 r 、 σ_E 已知，则解得如公式（3-17）方程组：

$$\begin{aligned}
 E &= V_A N(d_1) - D e^{-rT} N(d_2) \\
 \sigma_E &= \frac{V_A}{E} \times N(d_1) \sigma_A \\
 d_1 &= \frac{\ln\left(\frac{V_A}{D}\right) + \left(r + \frac{1}{2}\sigma_A^2\right)T}{\sigma_A \sqrt{T}} \\
 d_2 &= d_1 - \sigma_A \sqrt{T}
 \end{aligned} \tag{3-17}$$

即推导出 KMV 模型公式，已知部分数据，解公式（3-17）联合方程组即得出资产市场价值 V_A 与其波动率 σ_A 。

3.1.4 传统 KMV 模型评价

综上为经典 KMV 模型的基本思想跟原理假设，在确定了各参数之后，便可以用 KMV 模型推导出的公式进行计算预测相关企业的信用风险概率。据此，推导出本文所疑虑的两个关键问题。

第一，在 KMV 模型的原定假设中，假定企业的资产价值收益符合标准正态分布，但是通过学者们大量研究，我国资本市场存在特殊性，资本市场交易上的股票收益率并不完全服从标准正态分布，反而股票收益率存在波动聚集性（ARCH 效应），同时股票收益率波动会呈现“尖峰肥尾”特征，EGARCH 模型可以较好的修正此种特征，所以股权价值波动率考虑使用该模型进行修正。

第二，在 KMV 模型计算中，违约点 DPT 起到了重要作用，根据假设 DPT 代表企业会选择违约的临界点，一直以来，违约点 DPT 的计算使用 KMV 模型固定公式（即短期负债跟二分之一长期负债之和），该违约点是 KMV 公司在 KMV 模型创立之初依据美国本土的违约数据库进行计算出来的，虽然 KMV 模型具有众多优点，但是我国信用风险度量起步比较晚，尚未有建立起包含大量数据的历史违约数据库，因此，直接使用 KMV 模型原本定义的违约点 DPT 公式来计算我国上市公司信用风险会存在误差，具有不可考量性。所以本文考虑使用遗传学算法（GA）来修正违约点 DPT 公式，建立 ST 类公司和 *ST 类公司分别代表低风险组和高风险组当作数据库，使用遗传学算法来确定违约点 DPT 的最优解。其次，违约距离 DD 跟违约概率 EDF 之间成反相关关系，因为有部分学者认为违约概率 EDF 公式并不适合我国，因此，本文只使用违约距离 DD 来表示上市公司的信用风险大小，违约距离 DD 越大，则表示信用风险越小，反之则反。

3.2 EGARCH-GA-KMV 模型的构建

3.2.1 基于 EGARCH 模型的 KMV 模型参数修正

在 KMV 模型中参数股权价值波动率 σ_E 对计算结果较大, 股权价值波动率 σ_E 取自于股票收益率的波动年度数据, 波动率是股票收益率不确定性的波动程度, 是衡量金融风险的重要指标, 已广泛用于金融投资和金融风险的操作实务中。对波动率进行研究则能够更好的识别风险准确度、管控风险、了解变化趋势、提高收益率等。但以国内外的研究为依据, 波动率具有明显的集聚性和连续性, 具有很强的相关性, 经验验证, 这种情况广泛出现于我国资本市场中, 结合 KMV 模型原定假设, 这就造成对我国上市信用风险度量中直接使用原始的股权价值波动进行计算无法准确预估信用风险。

现实需求与理论需求强结合, 已催生了大量的波动率预测模型的研究, 自回归条件异方差类模型 (ARCH 类模型) 就是其中的一种, 并且随着研究的发展, ARCH 类模型已逐步变形和发展出诸如 GARCH 模型、EGARCH 模型、TGARCH 模型等一系列模型。从金融时间序列角度以及综合我国资本市场实际情况, 本文采取 EGARCH 模型对 KMV 模型中的股权价值波动率 σ_E 进行修正。

1. 时间序列跟波动率

时间序列是指一定时间内按时间顺序测量的某个变量的取值序列, 如果变量是股票收益率, 则股票收益率随着时间变化也是一种时间序列 (或金融时间序列), 波动率是股票收益率的条件标准差 (或者条件方差), 因此, 波动率也是一种金融时间序列。在资本市场中, 波动率 (volatility) 指资产市场价格的波动程度, 其具有强弱性且不能直接观察到, 但可以从资产收益率的上下摇摆中观察出波动率的特征, 其特征如下:

- (1) 波动率存在波动聚集性, 也可以表示平稳性。
- (2) 波动率一般不出现跳跃式变化, 是持续且连续变动的。
- (3) 波动率是平稳的, 一般在固定范围内变动。
- (4) 波动率具有杠杆效应 (leverage effect), 在资产价格大幅度上涨或者下跌的情况下, 波动率也会更大。

时间序列分析是指祈求通过过去来预测未来的变化, 波动率模型的提出便是如此。ARCH 模型的提出助力波动率模型的发展, 但实际情况是多变的, 为了满

足上述波动率的特征，许多波动率模型在原有诊断的基础上被不断的被提出，GARCH 模型在 ARCH 模型基础上解决了方差恒定所引起的拟合精度不准的问题，EGARCH 模型则在 GARCH 模型理论上更能表现出波动率在价格上扬和下跌时的不对称性。

2.波动率模型建立和变形

波动率是股票收益率的标准差，也是一种金融时间序列，为了使用历史信息研究未来，对资产收益率序列(时间序列)建立了波动率模型。1982 年，R. F. Engle 提出了 ARCH 模型（自回归条件异方差模型）。

ARCH 模型能准确地模拟时间序列变量的波动性的变化，在实证研究中应用甚广，能使人们更加精准的掌控风险（波动性），根据定义，ARCH (m) 模型如下所示：

$$y_t = \beta x_t + \varepsilon_t \quad (3-18)$$

$$\sigma_t^2 = x_0 + x_1 \varepsilon_{t-1}^2 + \dots + x_m \varepsilon_{t-m}^2 \quad (3-19)$$

公式 (3-18) 为条件均值模型， y_t 和 x_t 分别表示因变量和自变量， ε_t 为无序列相关性的随机扰动项。公式 (3-19) 为方差回归模型， σ_t^2 为 t 时刻随机扰动项的方差。其中 $x_0 > 0$ ， $x_j \geq 0, j = 1, 2, \dots, m$ ， m 为 ARCH 模型的滞后阶数。

但是在实际情况中，ARCH (m) 模型通常需要滞后很多阶数才能得到精确的拟合效果，因此，Bollerslev 在 1986 年公布 ARCH 模型的推广模型，称为 GARCH 模型（广义自回归条件异方差模型）。GARCH (m, s) 模型如下所示：

$$y_t = \beta x_t + \varepsilon_t \quad (3-20)$$

$$\sigma_t^2 = w + \sum_{i=1}^m \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^s \theta_j \sigma_{t-j}^2 \quad (3-21)$$

GARCH (m, s) 模型中， w 为常数项， m 和 s 分别代表 ARCH 项的滞后阶数和 GARCH 项滞后阶数，其中 $w > 0$ ， $\alpha_i \geq 0$ ， $\theta_j \geq 0$ ， $0 < \sum_{i=1}^m \alpha_i + \sum_{j=1}^s \theta_j < 1$ 。

为了填补 GARCH 模型处理金融时间序列过程中的缺陷，即 GARCH 模型不能解释股票收益和收益变化波动之间出现的负相关现象（非对称效应）以及假定条件方差模型 (3-21) 所有系数均大于 0。1991 年，Nelson 提出的 EGARCH 模型（指数 GARCH 模型）允许正负资产对波动率有非对称性影响，他考虑如下：

$$g(\varepsilon_t) = \alpha\varepsilon_t + \gamma[|\varepsilon_t| - E|\varepsilon_t|] \quad (3-22)$$

其中 α 和 γ 是实常数, ε_t 和 $|\varepsilon_t| - E|\varepsilon_t|$ 都分别是零均值独立同分布白噪声, 且连续分布, 所以 $E[g(\varepsilon_t)]=0$, $g(\varepsilon_t)$ 的非对称性可以从如下看出:

$$g(\varepsilon_t) = \begin{cases} (\alpha + \gamma)\varepsilon_t - \gamma E|\varepsilon_t|, & \varepsilon_t \geq 0 \\ (\alpha - \gamma)\varepsilon_t - \gamma E|\varepsilon_t|, & \varepsilon_t < 0 \end{cases} \quad (3-23)$$

进而得出 EGARCH (m, s) 模型写成:

$$a_t = \sigma_t \varepsilon_t \quad (3-24)$$

$$\ln \sigma_t^2 = \alpha_0 + \sum_{j=1}^m [\alpha_j \varepsilon_{t-j} + \gamma_j (|\varepsilon_{t-j}| - E|\varepsilon_{t-j}|)] + \sum_{i=1}^s \beta_i \ln \sigma_{t-i}^2 \quad (3-25)$$

式(3-25)左边表示条件方差的对数, 表明杠杆效应的影响是指数级的, α_j 代表的是对数收益率的正负扰动对波动率的不同影响, 如果 $\alpha_j = 0$ 则正负扰动对波动率影响就相同了。

从上面可以看出, GARCH 模型是对 ARCH 模型的推广, 能够更准确的拟合波动率的精度, 但是 GARCH 模型假定系数必须非负, 同时认定不存在杠杆效应, 也就是说在资产收益率的“尖峰肥尾”现象中, GARCH 模型能很好的解释“尖峰”现象, 但是“肥尾”现象还不能够很好的解释, 因此为了更为准确的描述资本市场股票收益率波动的特征, 本文选用 EGARCH 模型来对 KMV 模型中股权价值波动率 σ_E 进行修正。

3. EGARCH 模型参数修正

使用 EGARCH 模型对 KMV 模型股权价值波动率 σ_E 的修正主要分为以下几个步骤。

- (1) 数据预处理。对选择样本数据的资产收益率取对数等。
- (2) 平稳性检验。对金融时间序列变量进行平稳性检验, 可采用单位根检验进行检验, 同时金融时间序列变量还得满足白噪声序列要求, 即具有自相关性。
- (3) 建立条件均值模型。
- (4) ARCH 效应检验。EGARCH 模型是 GARCH 模型的改进, GARCH 模型的建立需要金融时间序列变量存在 ARCH 效应, 为此, 建立 EGARCH 模型的前提是存在 ARCH 效应。
- (5) 建立 EGARCH 模型的条件波动率模型。包含扰动项分布的选择、ARCH

项和 EGARCH 项最优阶数确定, 是否存在杠杆效应等, 最后对最优模型进行建模, 并检验最优模型是否还存在 ARCH 效应, 残差的系列检验等。

(6) 预测分析。通过模型预测分析未来股权价值波动 σ_E 。

3.2.2 基于遗传学算法的 KMV 模型参数修正

在 KMV 模型中, 违约点 DPT 是计算违约距离的重要数据, 对于预测违约风险概率来说至关重要, 不过 KMV 模型原定的违约点 DPT 公式是根据美国历史违约数据库经验决定的, 因此 KMV 模型原定的违约点不一定适用于我国上市公司。虽然众多学者在知悉存在这个问题之后, 相继采用穷举法来确定违约点短期负债或长期负债的系数, 但是此方法得出来的系数存在一定的主观性或者不是整体全局最优的。综上, 结合遗传学算法理论介绍, 本文考虑使用遗传学算法来确定违约点 DPT 的短期负债和长期负债双变量的系数。

1. 问题设定

根据第二章相关理论基础中遗传学算法的介绍, 可知遗传学算法包含字符编码、初始化种群、适应度评价、遗传算子和终止条件判断等步骤。遗传学算法对 KMV 模型违约点修正的基本思想是以我国上市公司作为样本数据, 通过求得判断上市公司类型失误率的极小值得到违约点的最优系数。

在我国资本市场上, 根据证监会规定, 当上市公司连续两年亏损或者净资产低于股票面值的时候其股票会被打上“ST”(特殊处理), 更进一步的还会打上“*ST”, 意为具有退市风险, 因其违约风险程度较高, 所以我们定义“ST 和*ST”上市公司为 ST 类公司, 当作高风险样本; 正常经营的上市公司定义为非 ST 类公司, 因其违约风险程度低, 当作低风险样本。本文仅采用违约距离 DD 来代表违约风险, 从上文可知, 违约距离 DD 跟违约概率成反向变动, 违约距离 DD 愈大, 则违约风险概率愈小。理论上, 当上市公司的违约距离大于或者等于 0 时, 上市公司不会违约, 而当违约距离小于 0 时, 上市公司可能已经违约未被披露或者即将选择违约, 违约潜在可能性大。基于此, 将违约距离 DD 大于或者等于 0 的上市公司判断为 L 类 (Low-Risk) 上市公司, 违约距离 DD 小于 0 的上市公司判断为 H 类 (High-Risk) 上市公司。那么, 对于选取的样本数据将会出现以下四种判断。

判断 1: ST 类公司被判断为 L 类上市公司;

判断 2：ST 类公司被判断为 H 类上市公司；

判断 3：非 ST 类公司被判断为 L 类上市公司；

判断 4：非 ST 类公司被判断为 H 类上市公司。

在判断过程中，判断 2 跟 3 的出现则表明判断准确，而判断 1 跟 4 则表明判断出现失误。那么判断 1 跟 4 出现的最少，判断的失误率最小时的违约点 DPT 的系数值就是全局最优值。

对违约点 DPT 公式进行重新设定，设定为：

$$DPT = \alpha STD + \beta LTD \quad (3-26)$$

其中 α 和 β 分别为系数，STD 为短期负债，LTD 为长期负债。对于每一组 (α , β) 代入违约距离 DD 公式重新计算违约距离，再重新判定公司类型。

2. 框架搭建

根据问题设定中的逻辑，结合遗传学算法通用框架搭建属于 GA-KMV 模型的计算框架。假设样本数据总数为 N ，样本数据为平衡面板数据，则 ST 类和非 ST 类公司数量分别为 $N/2$ ，在四种判断情况中，假设判断 2 数量为 m ，判断 3 数量为 n ，判断 1 跟 4 出现的概率就分别是 $(\frac{1}{2} - \frac{m}{N})$ ， $(\frac{1}{2} - \frac{n}{N})$ ，因此，判定违约点的最优系数问题就转化为了求解 $(1 - \frac{m+n}{N})$ 最小。遗传学算法对违约点 DPT 修正求解框架如图 3.4 所示。图 3.4 中，判定规则代表对样本数据 ST 类和非 ST 类的判定，根据前小节的问题设定，如若违约距离 DD 大于或者等于 0 则应判定为 L 类上市公司，否则是 H 类上市公司；适应度函数即每代种群中子代达到最优的标准，筛选出来的子代将作为下一代父代经过选择、交叉、变异等演化产生新的优化种群，继续进入循环，最终达到设定的终止条件遗传迭代 200 代为止停止迭代。

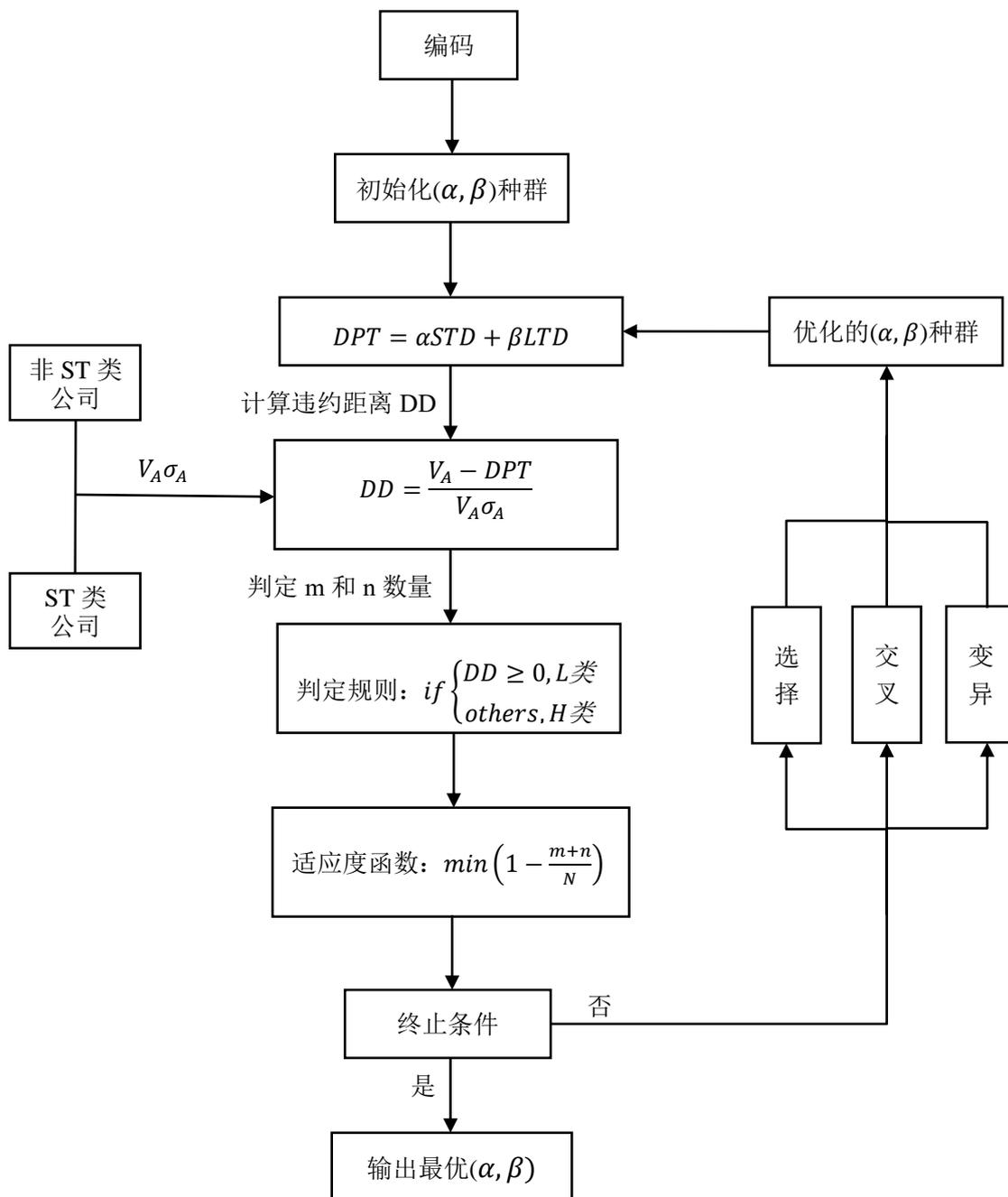


图 3.4 违约点修正框架图

3.2.3 EGARCH-GA-KMV 模型计算过程

改进型 KMV 模型跟传统 KMV 模型在计算过程一致，都是确定好股权市场价值 E、股权价值波动率 σ_E 、负债 D、无风险利率 r，接着利用已知数据计算出资产市场价值 V_A 和资产价值波动率 σ_A ，然后根据违约点 DPT 计算出违约距离 DD。但是为了适应度量我国上市公司信用风险，结合 KMV 模型原定假设以及我国资本市场实际情况，以我国上市公司公开数据为样本，分别使用 EGARCH 模型和遗传学算法（GA）对 KMV 模型中参数股权价值波动率 σ_E 及违约点 DPT。具体的改进型 KMV 模型计算框架图如图 3.5 所示。

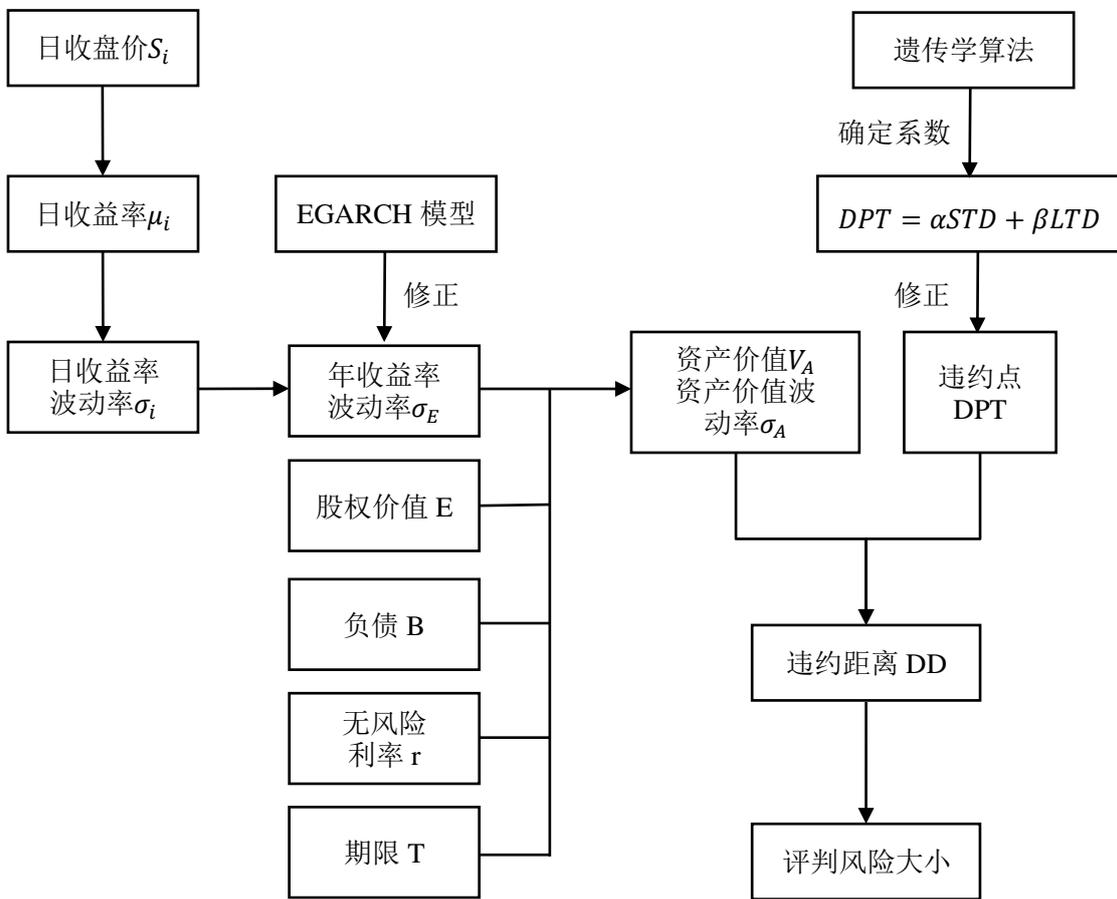


图 3.5 EGARCH-GA-KMV 模型计算过程

3.3 本章小结

本章主要介绍了四个方面。首先介绍了 KMV 模型，厘清理论基础、前提假设、确定参数选取、公式推导等，并给出 KMV 模型中股权价值波动率跟违约点需要重新修正的原因。其次，从时间序列和波动率介绍到波动率模型的发展，并选取 EGARCH 模型修正股权价值波动率，并说明理由。再次，按照第二章 2.2.3 章节的遗传学算法通用框架设定 KMV 模型中遗传学算法对违约点 DPT 修正的问题设计和逻辑框架。最后，给出传统 KMV 和基于 EGARCH 模型和遗传学算法改进型 KMV 模型计算过程。

4 基于 EGARCH-GA-KMV 模型的上市公司信用风险实证分析

4.1 样本选取和参数设定

4.1.1 样本选取

本文数据来源于 RESSET 及 CSMAR 数据库, 样本选取的区间为 2018 年 1 月 1 日至 2022 年 12 月 31 日连续 5 年数据, 选取的是各上市公司每年年度数据。本文将样本数据划分为 ST 类上市公司和非 ST 类上市公司当作高风险组和低风险组, 高风险组为特殊处理或者具有退市风险的被标记为“ST”或“*ST”的上市公司, 低风险组为正常经营的上市公司。本文将选取区间内连续三年及以上被标记为“ST”或“*ST”上市划入高风险组。

ST 类上市公司包含被证监会特殊处理或者面临退市风险的上市公司, 相较于其他正常公司, 具有较高的信用风险, 因此被当作高风险违约组, 正常类上市公司被当作低风险组, 然后分别计算违约距离 DD, 并将传统 KMV 模型计算跟改进型 KMV 模型计算的相比较。对于 ST 类上市公司和非 ST 类上市公司具有如下原则。

- (1) 均为 A 股上市公司, 剔除同时在 B 股或者 H 股一同上市的公司;
- (2) ST 类上市公司配对的非 ST 类上市公司处于同一或者相近行业, 行业标准按照《国民经济行业分类和代码表》执行;
- (3) ST 类上市公司配对的非 ST 类上市公司资产规模大致相近;
- (4) 所有样本在时间区间内均没有长时间停牌或者重大事件发生, 股票交易应连续。

基于以上原则, 本文选取总样本 80 家上市公司, 其中高风险组 (ST 类) 低风险 (非 ST 类) 组各 40 家。选取的样本如下表 4.1 所示, 带“ST”或“*ST”标签为高风险组。

表 4.1 选取的样本

证券代码	证券名称	证券代码	证券名称	证券代码	证券名称	证券代码	证券名称
000410	*ST 沈机	600084	ST 中葡	000021	深科技	002927	泰永长征
000422	*ST 宜化	600112	ST 天成	000030	富奥股份	600071	凤凰光学
000707	*ST 双环	600150	*ST 船舶	000545	金浦钛业	600076	康欣新材
000752	*ST 西发	600228	ST 昌九	000590	启迪古汉	600099	林海股份
000806	*ST 银河	600234	*ST 山水	000635	英力特	600105	永鼎股份

续表 4.1 选取的样本

证券代码	证券名称	证券代码	证券名称	证券代码	证券名称	证券代码	证券名称
000820	*ST 节能	600238	ST 椰岛	000731	四川美丰	600110	诺德股份
000920	*ST 汇通	600290	ST 华仪	000822	山东海化	600130	波导股份
000953	*ST 河化	600301	*ST 南化	000856	冀东装备	600143	金发科技
000972	ST 中基	600319	*ST 亚星	000919	金陵药业	600230	沧州大化
000981	*ST 银亿	600408	ST 安泰	000929	兰州黄河	600249	两面针
002005	ST 德豪	600421	ST 仰帆	000960	锡业股份	600367	红星发展
002021	ST 中捷	600518	ST 康美	002008	大族激光	600513	联环药业
002089	*ST 新海	600539	ST 狮头	002104	恒宝股份	600764	中国海防
002102	ST 冠福	600595	ST 中孚	002112	三变科技	600847	万里股份
002427	*ST 尤夫	600608	ST 沪科	002154	报喜鸟	600962	国投中鲁
002501	*ST 利源	600666	*ST 瑞德	002199	东晶电子	601579	会稽山
002513	*ST 蓝丰	600725	*ST 云维	002274	华昌化工	603218	日月股份
002569	ST 步森	600781	*ST 辅仁	002472	双环传动	603811	诚意药业
002692	ST 远程	600877	ST 电能	002576	通达动力	603878	武进不锈
002766	*ST 索菱	601113	ST 华鼎	002702	海欣食品	603936	博敏电子

资料来源：锐思数据库、国泰安数据库

4.1.2 参数设定

按照 3.1.2 章节设定，EGARCH-GA-KMV 模型（后简称改进型 KMV 模型）跟传统 KMV 模型一样，模型的参数包含股权市场价值 E 、股权价值波动率 σ_E 、资产市场价值 V_A 、资产价值波动率 σ_E 、债务期限 T 、无风险利率 r 、负债 D 、违约点 DPT 、违约距离 DD 。因此，本节确定股权市场价值 E 、债务期限 T 、无风险利率 r 、负债 D 的参数值，其他参数在 4.2 章节进行实证分析。

1. 股权市场价值 E

按照传统 KMV 模型公式，股权市场价值 E 等于股票价格和在售股票数量的乘积。我国自从股权分置改革之后资本市场存在流通股和非流通股，因此使用原假定的公式计算股权市场价值 E 会对结果有影响，所以本文所确定的股权市场价值 E 采用以下计算方法。如未加以说明，则本文所算结果均基于此数据计算。

$$\text{股权市场价值 } E = \text{期末收盘价} \times \text{流通股股数} + \text{每股净资产} \times \text{非流通股股数}$$

2. 负债 D

负债 D 依然按照章节 3.1.2 所定义的参数进行确定，即负债等于长期负债跟

短期负债之和。如未加以说明，则本文所算结果均基于此数据计算。

3. 期限 T

期限 T 按照传统假定，本文均设定为 1 年，如计算违约距离为季度数据，则 T 选取 0.25。

4. 无风险利率 r

本文使用的无风险利率 r 为一年期中债国债收益率，如下表 4.2 所示。

表 4.2 无风险利率

年份	无风险利率（百分比）
2018	2.6
2019	2.5
2020	2.6
2021	2.3
2022	1.8

资料来源：中国国债信息网

4.2 实证分析

4.2.1 基于 EGARCH 的股权价值波动率检验和确定

根据前文分析，本节按照章节 3.2.1 中 EGARCH 模型的构建使用 EGARCH(m, s)模型修正股权价值波动率 σ_E ，其中 m 和 s 分别为 ARCH 项和 GARCH 项阶数。因为 EGARCH 模型是基于历史数据对未来进行预测，所以选定区间为 2016 年 1 月 1 日至 2022 年 12 月 30 日样本公司日收益率，依次按照年度使用 EGARCH(m, s)模型预测 2018 年 1 月 1 日至 2022 年 12 月 30 日样本公司收益率波动率，最后转换成年波动率，也即股权价值波动率。

1. 平稳性检验

首先，要判断 EGARCH 模型是否可用，因为指数条件异方差模型可用的前提是平稳的残差序列，如若数据不平稳，可先建立对收益率序列的对数收益率序列。这里以*ST 沈机（股票代码 000410）为例，选取样本区间为 2016 年 1 月 1 日至 2018 年 12 月 31 日，预测区间为 2018 年度，具体介绍其波动率修正过程。数据来源于国泰安数据库，使用 EViews 软件进行操作。

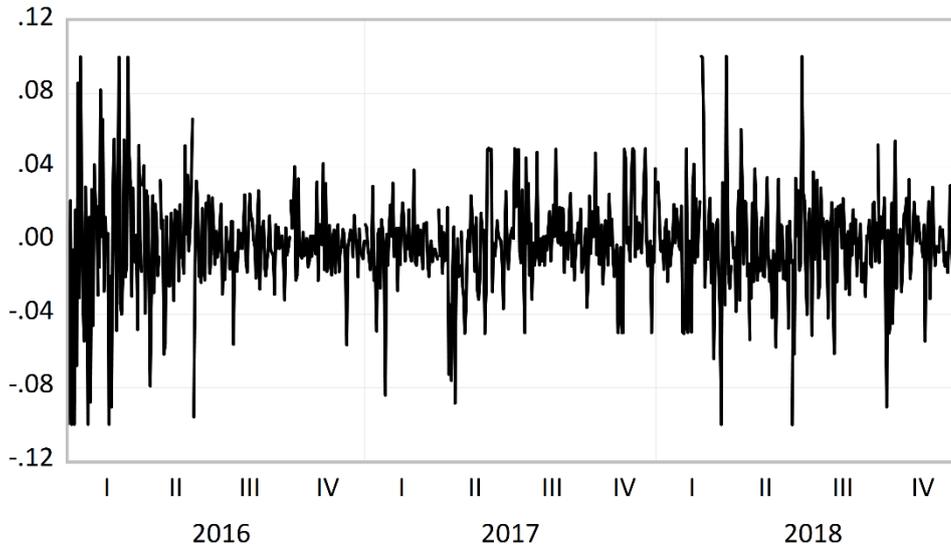


图 4.1 *ST 沈机对数收益率时序图

图 4.1 为进行平稳性检验的第一种方法，可视图形化，从图形上可以看出，收益率围绕数值 0 上下波动，波动幅度上下一致，为白噪声序列，即平稳序列。其次，也说明*ST 沈机（000410）在 2016 年上半年及 2018 年上半年波动较大。

除此之外，还可以进行单位根检验，单位根检验原假设 H_0 ：存在单位根，序列不平稳，备择假设 H_1 ：不存在单位根，序列平稳。结果如表 4.3。可知 $p < 0.05$ ，拒绝原假设，序列平稳，可以进行下一步操作。如果数据不平稳，则需要降阶进行操作。

表 4.3 *ST 沈机单位根平稳性检验表

	T 统计量	P 值
Augmented Dickey-Fuller test statistic	-25.95247	0.0000
Test critical values: 1% level	-3.970957	
5% level	-3.416123	
10% level	-3.130350	

2.描述性统计

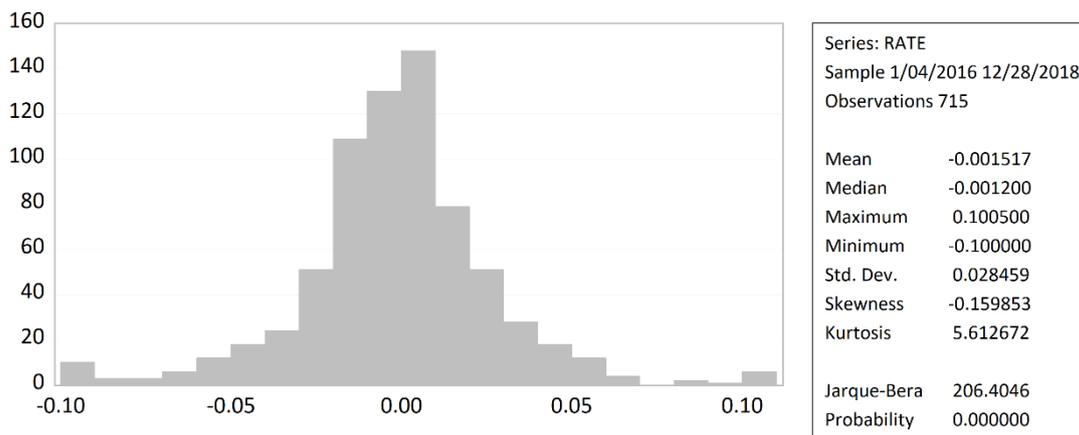


图 4.2 *ST 沈机描述性统计

根据 1986 年 Bollerslev 提出的定义，GARCH 模型应具备尖峰分布及条件变异异质性，因此 EGARCH 模型也应具备上述定义。从图 4.2 可以看出，偏度 Skewness 为 $-0.159853 < 0$ ，则说明收益率序列左偏，峰度 Kurtosis 为 $5.612672 > 3$ ，说明收益率序列尖峰，因此此收益率序列存在“尖峰肥尾”现象，不服从正态分布，满足尖峰条件。

3.自相关和偏自相关检验

根据定义，序列平稳之后，不存在自相关性是 EGARCH 模型建立的前提，如果存在自相关性，可以更换或者梳理数据解决，或者对平稳序列嵌套 ARMA 模型来消除序列相关性。*ST 沈机对数收益率序列自相关和偏自相关性如表 4.4 所示。

表 4.4 *ST 沈机自相关和偏自相关检验表

阶数	AC	PAC	Q-Stat	Prob
1	0.033	0.033	0.7614	0.383
2	0.062	0.061	3.4765	0.176
3	0.009	0.005	3.5288	0.317
4	0.038	0.034	4.5434	0.337
5	-0.02	-0.024	4.8381	0.436
6	-0.041	-0.045	6.0457	0.418

续表 4.4 *ST 沈机自相关和偏自相关检验表

阶数	AC	PAC	Q-Stat	Prob
7	-0.036	-0.032	6.9796	0.431
8	-0.063	-0.057	9.7979	0.279
9	0.024	0.034	10.199	0.335
10	-0.091	-0.084	16.131	0.096
11	0.062	0.067	18.856	0.064
12	0.042	0.049	20.094	0.065
13	0.026	0.01	20.58	0.082

如表 4.4 所示，其中 AC 和 PAC 分别代表自相关性和偏自相关性，Q-Stat 和 Prob 代表 Q 统计量和 P 值，对于 AC 和 PAC 而言，越接近于 1 自相关性程度越高，同时，Q 统计量所对应的 P 值均大于 0.05，所以序列在 5%水平上显著不存在自相关性。

4.ARCH 效应检验

在检验完对数收益率序列符合 EGARCH 模型建立条件之后，还需要建立均值方程，建立的均值方程应该具有 ARCH 效应。这里使用 Eviews 软件的 ARCH-LM 检验对均值方程进行 ARCH 效应检验，分别滞后阶数 1, 2, 4, 6 阶数检查。如表 4.4 所示，可以看到 P 值均小于 0.05，存在严重的 ARCH 效应，可以进行建立 EGARCH 模型。

表 4.5 *ST 沈机 ARCH 效应检验

阶数	F-statistic	Prob. Chi-Square(1)	Prob. F(1,697)
1	11.13003	0.0009	0.0009
2	25.51956	0.0000	0.0000
4	13.18347	0.0000	0.0000
6	8.601900	0.0000	0.0000

5. 建立 EGARCH 模型

在 EGARCH 模型中，完整为 EGARCH(m, s)模型，其中 m 为 ARCH 项阶数， s 为 GARCH 项阶数，因此在建立 EGARCH 模型之前需要判断模型的最优阶数，同时，基于 EGARCH 模型承认存在杠杆效应的预测，还需要加入非对称阶数 (r)，阶数要求 $r < m$ 。并选择 T 分布来建立模型。在 Eviews 软件中做出操作，阶数定阶需要一步一步判断，ARCH 项滞后阶数跟 GARCH 项滞后阶数选取一般不宜过高，其中 r 为 0 则代表不存在非对称效应，为 1 则代表加入非对称效应。部分阶数选取及结果如表 4.6 所示。

表 4.6 最优阶数选取

阶数取值 (m, r, s)	AIC	SC	HQ	GARCH 项 显著性	ARCH 项 显著性	非对称项 显著性
(1,1,1)	-4.532372	-4.493363	-4.517293	显著	显著	不显著
(2,1,2)	-4.535610	-4.483598	-4.515504	显著	显著	不显著
(6,1,1)	-4.576341	-4.504824	-4.548696	显著	显著	显著

其中 AIC、SC、HQ 为衡量模型拟合优良性的标准，分别是赤池信息准则、施瓦兹准则、HQ 信息准则，其值越小越好。通过表 4.6 对比，当 ARCH 项阶数跟 GARCH 项阶数分别取 6 跟 1 时，且存在非对称效应时，最显著，所以这里选择第三种建立 EGARCH(6, 1)模型的条件均值方程跟条件方差方程，同时该模型还存在非对称效应，模型拟合效果更好。

最后可进一步对建立的 EGARCH 模型检验，EGARCH 模型建立之后应不存在 ARCH 效应，ARCH 效应检验如下，分别滞后阶数 1，2，4，6 阶数检查。

表 4.7 建模后 ARCH 效应检验

阶数	F-statistic	Prob. Chi-Square(1)	Prob. F(1,697)
1	0.522133	0.4702	0.4695
2	0.350188	0.7047	0.7037
4	0.504813	0.7322	0.7306
6	0.650397	0.6898	0.6875

如表 4.7 所示, 建模之后, P 值大于 0.05, 说明均不存在 ARCH 效应。同时建模完成之后, 残差项不应处在自相关性, 使用 Eviews 进行自相关 Q 检验发现滞后 36 阶均不存在自相关性, 建模成功。

6. 股权价值波动率预测

通过上面步骤建模之后得到股票收益率的日波动率, 取日波动率长期均值, 最后乘以该年实际交易天数, 按照章节 3.1.2 股权价值波动率方法计算得到*ST 沈机 2018 年年度股权价值波动率 $\sigma_E = 0.409608739$ 。

重复上文所说操作, 验证符合建模前提且建模完成之后相继检验, 统计出 80 家上市公司连续 5 个年度数值, 最后得到基于 EGARCH 模型的股权价值波动率。但是股权价值波动率的修正结果是基于历史数据进行修正预测的, 因此每次在修正预测过程中, 需要对样本的时间进行不断的取值。修正后的股权价值波动率如表 4.8 所示, 因数据太多, 此处只展示 2018 年样本修正后的股权价值波动率数据。

表 4.8 样本数据 2018 年基于 EGARCH 模型的股权价值波动率

证券名称	修正 σ_E	证券名称	修正 σ_E	证券名称	修正 σ_E	证券名称	修正 σ_E
*ST 沈机	0.409609	ST 中葡	0.319801	深科技	0.434321	泰永长征	0.568047
*ST 宜化	0.436629	ST 天成	0.470589	富奥股份	0.314112	凤凰光学	0.448463
*ST 双环	0.442626	*ST 船舶	0.353818	金浦钛业	0.247079	康欣新材	0.333752
*ST 西发	0.481158	ST 昌九	0.369539	启迪古汉	0.487819	林海股份	0.439935
*ST 银河	0.523909	*ST 山水	0.38442	英力特	0.419922	永鼎股份	0.425461
*ST 节能	0.572887	ST 椰岛	0.450302	四川美丰	0.345471	诺德股份	0.379131
*ST 汇通	0.3846	ST 华仪	0.450696	山东海化	0.408277	波导股份	0.481604
*ST 河化	0.440285	*ST 南化	0.321704	冀东装备	0.496430	金发科技	0.345033
ST 中基	0.3403	*ST 亚星	0.430655	金陵药业	0.284193	沧州大化	0.483944
*ST 银亿	0.210559	ST 安泰	0.408902	兰州黄河	0.368558	两面针	0.425206
ST 德豪	0.221522	ST 仰帆	0.347438	锡业股份	0.424910	红星发展	0.365222
ST 中捷	0.40297	ST 康美	0.250228	大族激光	0.451644	联环药业	0.322650
*ST 新海	0.411622	ST 狮头	0.324524	恒宝股份	0.427093	中国海防	0.283282
ST 冠福	0.423964	ST 中孚	0.393127	三变科技	0.540332	万里股份	0.422145
*ST 尤夫	0.538215	ST 沪科	0.349986	报喜鸟	0.299345	国投中鲁	0.488957
*ST 利源	0.359434	*ST 瑞德	0.481282	东晶电子	0.491270	会稽山	0.271146

续表 4.8 样本数据 2018 年基于 EGARCH 模型的股权价值波动率

证券名称	修正 σ_E	证券名称	修正 σ_E	证券名称	修正 σ_E	证券名称	修正 σ_E
*ST 蓝丰	0.414207	*ST 云维	0.366815	华昌化工	0.323803	日月股份	0.438316
ST 步森	0.508109	*ST 辅仁	0.388456	双环传动	0.340686	诚意药业	0.413920
ST 远程	0.467489	ST 电能	0.32636	通达动力	0.520421	武进不锈	0.322790
*ST 索菱	0.445998	ST 华鼎	0.326829	海欣食品	0.414949	博敏电子	0.473863

4.2.2 资产市场价值及波动率确定

基于 EGARCH 模型确定股权价值波动率之后，得到修正的股权价值波动率 σ_E ，按照第三章 3.1.2 章节公式 (3-5) 至公式 (3-8) 建立联立方程组，且已知 4.2 章节中股权市场价值 E、负债 D、无风险利率 r，代入联立方程组，使用 Matlab 中的 fsolve 函数计算得到改进型 KMV 模型所需的资产市场价值 V_A 和资产价值波动率 σ_A 。

4.2.3 最优违约点确定

对违约点 DPT 的修正采用遗传学算法进行确定，遗传学算法能提供一个全局的最优解。按照第二章 2.2.3 章节对遗传学算法进行框架搭建，违约点确定流程按照第三章 3.2.2 章节进行修正。这里将 Geatpy 库运用到 Python 中进行计算，Geatpy 为进化算法工具箱，具有性能高、实用性强、演算速度快等特点，其中包含众多已实务操作化的库函数，并能提供高度模块化、耦合度低的面向对象的进化算法框架。

遗传学算法确定违约点过程中需要的用到的数据有股权市场价值 E、基于 EGARCH 模型修正的股权价值波动率 σ_E 、修正股权价值波动率后计算得出的资产市场价值 V_A 和资产价值波动率 σ_A 、短期负债 STD、长期负债 LTD、总负债 D 以及无风险利率 r。同时重新设定违约点公式为：

$$DPT = \alpha STD + \beta LTD \quad (4-1)$$

根据第三章 3.2.2 章节确立好的违约点修正框架，将 ST 类和非 ST 类各 40 家上市公司连续 5 年总共 400 个数据当作高风险组（40 家 200 个）和低风险组（40 家 200 个）代入数据，输出结果为适应度函数 $\left(1 - \frac{m+n}{N}\right)$ 最小，则此时的 (α, β) 即为最优。进行计算之前需要设定好遗传学算法参数，这里将终止条件

设定为遗传迭代 200 代，其他参数设定如表 4.9 所示。

表 4.9 遗传学算法执行参数设定

设定名称	数值	设定名称	数值
适应度函数	$\min\left(1 - \frac{m+n}{N}\right)$	决策变量类型	实数
初始值	(0, 0)	交叉概率	0.8
决策变量	(α , β)	重组概率	0.05
决策变量范围	[0,10]	种群规模	50

其中，决策变量即 α 和 β 取值；决策变量范围即 α 和 β 取值范围；决策变量类型一般分为实数和整数，为了取值更方便选择实数；交叉概率即父辈基因在下一代个体间染色体交叉组合概率；重组概率即基因突变概率，基因不宜过大突变，一般选择 0.01-0.1；种群规模即生成初代数组个体数量。

经过反复多次尝试迭代，分别得出如表 4.10 遗传学算法计算出来的最优违约点系数。其中 m 表示 ST 类上市公司被判定为 H 类，n 表示非 ST 类上市公司被判定为 L 类，N 为总数 400。

表 4.10 遗传学算法最优违约点判定表

α	β	N	m	n	适应度函数值
1.1916	4.999	400	99	148	0.3824
1.1648	6.2082	400	103	159	0.345
1.1655	5.1794	400	99	162	0.3475

表 4.10 为其中三次使用遗传学算法进行计算的情况汇总，三次运算结果适应度函数值几乎围绕在数值 0.35 附近，每一次都代表该次运算迭代内最优值，根据定义适应度函数最小即为最优，因此这里取 $\alpha = 1.1648$ ， $\beta = 6.2082$ 作为本文最优违约点系数，改进型 KMV 模型违约点公式如下：

$$DPT = 1.1648 * STD + 6.2082 * LTD \quad (4-2)$$

据第三章 3.1.2 章节可知，传统 KMV 模型违约点公式如下：

$$DPT = STD + \frac{1}{2} * LTD \tag{4-3}$$

最优违约点 $\alpha = 1.1648$ ， $\beta = 6.2082$ 的运行迭代收敛过程如图 4.3。

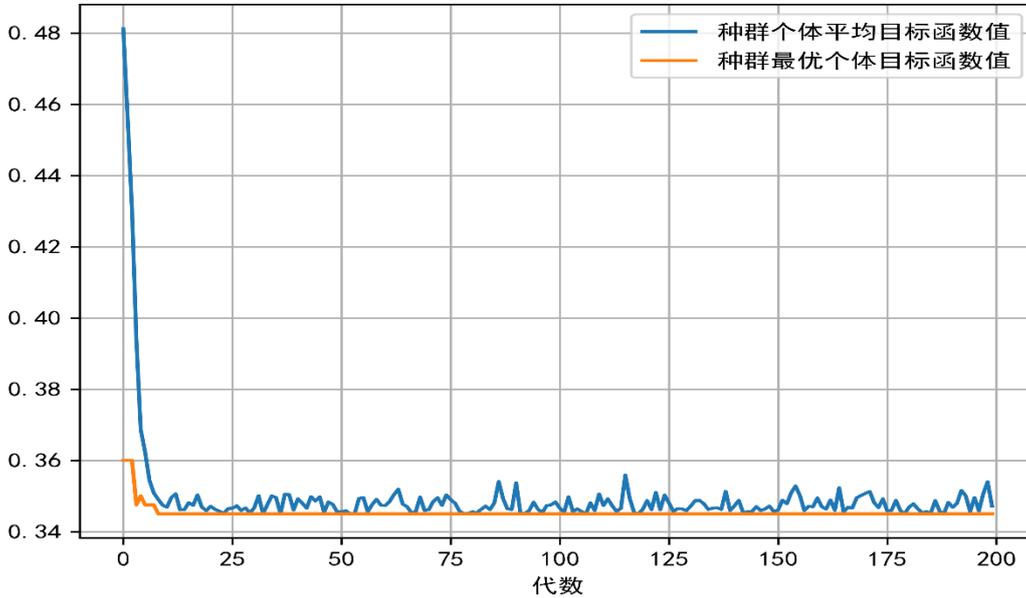


图 4.3 遗传学算法优化违约点运行迭代过程

从图 4.3 中可以看出，种群个体平均适应度函数值在前 10 代内迅速收敛下降，最优个体适应度函数值在第 9 代开始收敛于 0.345，第 25 代以后种群最优跟个体最优均在 0.345 附近上下浮动，最终经过 200 次迭代，均都收敛于 0.345，得到最优违约点 $\alpha = 1.1648$ ， $\beta = 6.2082$ 。最终得到改进型 KMV 模型违约点如表 4.11 所示，因数据过多，此处只展示 2018 年样本公司数据。

表 4.11 改进型 KMV 模型 2018 年样本公司最优违约点

证券名称	违约点 (万元)	证券名称	违约点 (万元)	证券名称	违约点 (万元)	证券名称	违约点 (万元)
*ST 沈机	3487933	ST 中葡	76516	深科技	1138250	泰永长征	18632
*ST 宜化	4234413	ST 天成	191912	富奥股份	931355	凤凰光学	61775
*ST 双环	928912	*ST 船舶	6612710	金浦钛业	165469	康欣新材	467142
*ST 西发	86430	ST 昌九	35119	启迪古汉	75937	林海股份	13703
*ST 银河	154851	*ST 山水	100247	英力特	67036	永鼎股份	381009
*ST 节能	306718	ST 椰岛	115742	四川美丰	358858	诺德股份	1113575

续表 4.11 改进型 KMV 模型 2018 年样本公司最优违约点

证券名称	违约点 (万元)	证券名称	违约点 (万元)	证券名称	违约点 (万元)	证券名称	违约点 (万元)
*ST 汇通	386191	ST 华仪	512279	山东海化	278858	波导股份	23377
*ST 河化	75496	*ST 南化	27891	冀东装备	204290	金发科技	3026786
ST 中基	98422	*ST 亚星	164138	金陵药业	89945	沧州大化	116306
*ST 银亿	3880164	ST 安泰	707797	兰州黄河	35792	两面针	165615
ST 德豪	624338	ST 仰帆	5731	锡业股份	5401031	红星发展	71345
ST 中捷	886537	ST 康美	14842084	大族激光	2411696	联环药业	55160
*ST 新海	1428962	ST 狮头	9388	恒宝股份	40124	中国海防	76441
ST 冠福	977328	ST 中孚	4702724	三变科技	98139	万里股份	11739
*ST 尤夫	1863590	ST 沪科	15709	报喜鸟	215490	国投中鲁	183692
*ST 利源	362944	*ST 瑞德	862201	东晶电子	14106	会稽山	226952
*ST 蓝丰	100828	*ST 云维	25376	华昌化工	595214	日月股份	155448
ST 步森	116824	*ST 辅仁	1004639	双环传动	1506837	诚意药业	19052
ST 远程	389828	ST 电能	342663	通达动力	31611	武进不锈	110095
*ST 索菱	627287	ST 华鼎	882412	海欣食品	42714	博敏电子	237550

4.2.4 违约距离确定

根据第三章 3.2.3 章节构建的 EGARCH-GA-KMV 模型计算过程, 使用数据股权市场价值 E 、基于 EGARCH 模型修正的股权价值波动率 σ_E 、修正股权价值波动率后计算得出的资产市场价值 V_A 和资产价值波动率 σ_A 、总负债 D 、基于遗传学算法确定的最优违约点 DPT 以及无风险利率 r 使用第三章公式 (3-10) 得到改进型 KMV 模型违约距离, 如表 4.12 所示, 此处为 2018 年样本公司数据。

表 4.12 改进型 KMV 模型 2018 年样本公司违约距离

证券名称	违约距离	证券名称	违约距离	证券名称	违约距离	证券名称	违约距离
*ST 沈机	-15.00067	ST 中葡	2.4469858	深科技	-0.614224	泰永长征	1.5211983
*ST 宜化	-33.64038	ST 天成	0.1511373	富奥股份	-1.084709	凤凰光学	1.4998592
*ST 双环	-12.8665	*ST 船舶	-7.216476	金浦钛业	1.9245876	康欣新材	-0.196086
*ST 西发	1.1552912	ST 昌九	2.2545181	启迪古汉	1.1536806	林海股份	2.0383483
*ST 银河	0.8751627	*ST 山水	1.0210252	英力特	1.7954532	永鼎股份	0.5613767

续表 4.12 改进型 KMV 模型 2018 年样本公司违约距离

证券名称	违约距离	证券名称	违约距离	证券名称	违约距离	证券名称	违约距离
*ST 节能	-1.447153	ST 椰岛	1.0517136	四川美丰	-0.513078	诺德股份	-3.449009
*ST 汇通	-1.903367	ST 华仪	-2.211643	山东海化	0.9278302	波导股份	1.907174
*ST 河化	0.8661751	*ST 南化	2.5168998	冀东装备	0.8302873	金发科技	-3.818757
ST 中基	1.6694929	*ST 亚星	-0.155938	金陵药业	2.5681163	沧州大化	1.7199059
*ST 银亿	-17.24058	ST 安泰	-3.907879	兰州黄河	1.9217002	两面针	0.4703588
ST 德豪	-3.113148	ST 仰帆	2.7730331	锡业股份	-6.860264	红星发展	1.8547107
ST 中捷	-2.826662	ST 康美	-10.28563	大族激光	0.4998925	联环药业	2.1569489
*ST 新海	-6.546957	ST 狮头	2.8936825	恒宝股份	2.0887665	中国海防	3.1994288
ST 冠福	-1.854091	ST 中孚	-31.71785	三变科技	0.4304692	万里股份	2.2189261
*ST 尤夫	-10.44169	ST 沪科	2.5358887	报喜鸟	1.2952355	国投中鲁	0.1070495
*ST 利源	-3.090363	*ST 瑞德	-5.047387	东晶电子	1.9134449	会稽山	1.425153
*ST 蓝丰	1.2977208	*ST 云维	2.4911618	华昌化工	-2.429622	日月股份	0.2333447
ST 步森	0.2845928	*ST 辅仁	-8.490842	双环传动	-5.84885	诚意药业	2.0822094
ST 远程	-0.679246	ST 电能	-0.064639	通达动力	1.5755981	武进不锈	0.2622907
*ST 索菱	-7.725507	ST 华鼎	-0.807588	海欣食品	1.6499479	博敏电子	0.1060796

4.3 实证结果检验和分析

4.3.1 EGARCH-GA-KMV 模型实证结果检验

上文 4.3 章节建立了改进型 KMV 模型（同 EGARCH-GA-KMV 模型），在章节 3.1.3 构建了传统 KMV 模型，为了验证改进型 KMV 模型跟传统 KMV 模型计算出的违约距离 DD 具有显著差异，接下来进行配对样本 T 检验跟 M-WU 检验。

1. 配对样本 T 检验

配对样本 T 检验需要服从配对性假设及正态分布假设，如果为大样本则可忽略正态分布假设。配对样本 T 检验原假设 H_0 ：两样本没有显著差异，备择假设 H_1 ：两样本具有显著差异。

为此，我们将改进型 KMV 模型计算出的违约距离 DD 跟传统 KMV 模型计算出的违约距离 DD 作为配对，KMV 模型作为变量进行配对样本 T 检验，检验结果如表 4.13 所示。

表 4.13 违约距离 T 检验结果

模型	ST 均值±标准差	均值差	T 值	P 值
改进前	0.499827±2.90548	4.619977	7.7052	0.0000
改进后	-4.120149±11.03088			
模型	非 ST 均值±标准差	均值差	T 值	P 值
改进前	1.655683±0.6269914	1.111779	8.3443	0.0000
改进后	0.5439039±2.247347			

从表 4.13 可以看出，配对样本 T 检验的 P 值改进前后均小于 0，具有统计学意义，拒绝原假设，ST 类和非 ST 类均存在较大差异，但是 ST 类差值更大，非 ST 类差值更小，说明 ST 类作为高风险违约样本传统 KMV 模型不能够较好的预测违约风险，而非 ST 类上市公司作为低风险正常经营公司其违约风险较小，所以有差异但是差异不大。这也从侧面说明改进型 KMV 模型能够更好的满足 ST 类上市公司的违约信用风险预测。

2. 曼-惠特尼检验 (M-WU 检验)

曼-惠特尼检验也称 wilcoxon 秩和检验，是检验两个独立样本之间中位数是否相同。M-WU 检验是一种非参数检验，当检验样本数据正态性和方差齐性不满足的时候就可以使用 M-WU 检验。其主要效用是反应两个样本中位数之间的差异程度，为此我们分别对传统 KMV 模型计算的违约距离 DD 和改进型 KMV 模型计算的违约距离 DD 分别进行 M-WU 检验。结果如表 4.14 所示。

表 4.14 违约距离 M-WU 检验结果

模型	ST 类中位数	非 ST 类中位数	Z 值	P 值
传统 KMV	1.397337	1.631321	3.556	0.0004
改进型 KMV	-0.0873428	1.265771	5.055	0.0000

从表 4.14 可以看出，两个模型 M-WU 检验的 P 值均小于 0，具有统计学意义，均存在显著性差异。但传统 KMV 模型计算出的 ST 类违约距离 DD 中位数为正，不符合违约组理论小于 0 假设，同时改进型 KMV 模型的 Z 值比传统 KMV

模型大,说明改进型 KMV 模型更能体现高风险组 ST 类跟低风险组 ST 类之间的差异。综上检验,我们可以发现,改进型 KMV 模型相较于传统 KMV 模型可以较好的显示违约样本跟不违约样本之间的差异,体现了改进型 KMV 模型较好的拟合性。

4.3.2 样本组间实证结果对比分析

根据 3.1.2 章节的图 3.3 可知,违约距离 DD 跟预测违约概率 EDF 成反相关关系,即违约距离越大,则预测违约概率越小,反之则反。因此我们在样本间分别进行对比分析。

首先我们通过分析传统 KMV 模型跟基于 EGARCH-GA 改进的 KMV 模型的违约距离在样本组间进行对比分析,对比违约风险拟合度方面的比较。ST 类跟非 ST 类违约距离传统 KMV 模型跟改进型 KMV 模型对比图分别如图 4.4 跟图 4.5 所示。

从图 4.4 中可以看出,在 ST 类上市公司违约距离绝大部分位于数值 0 上方,ST 类上市公司作为违约样本,违约距离应接近于 0 或者小于 0,在这一点上,传统 KMV 模型存在小于 0 的数值,但不多,而改进 KMV 模型在这一点优于传统 KMV 模型。从图 4.5 可以看出,非 ST 类上市公司违约距离传统 KMV 模型跟改进型 KMV 模型之间差别不大,绝大多数都位于数值 0 上方,但改进型 KMV 模型存在部分异突值,位于数值 0 下方,仔细查看发现异突值原传统 KMV 模型违约距离本身就接近于 0,这说明传统 KMV 模型在测定正常样本的时候并不能区分差异性,而改进型 KMV 模型则区分出这种差异性,这部分异突值可能代表传统 KMV 预测准确率过低或者已经违约存在重大情况,但市场并没有正确反应。

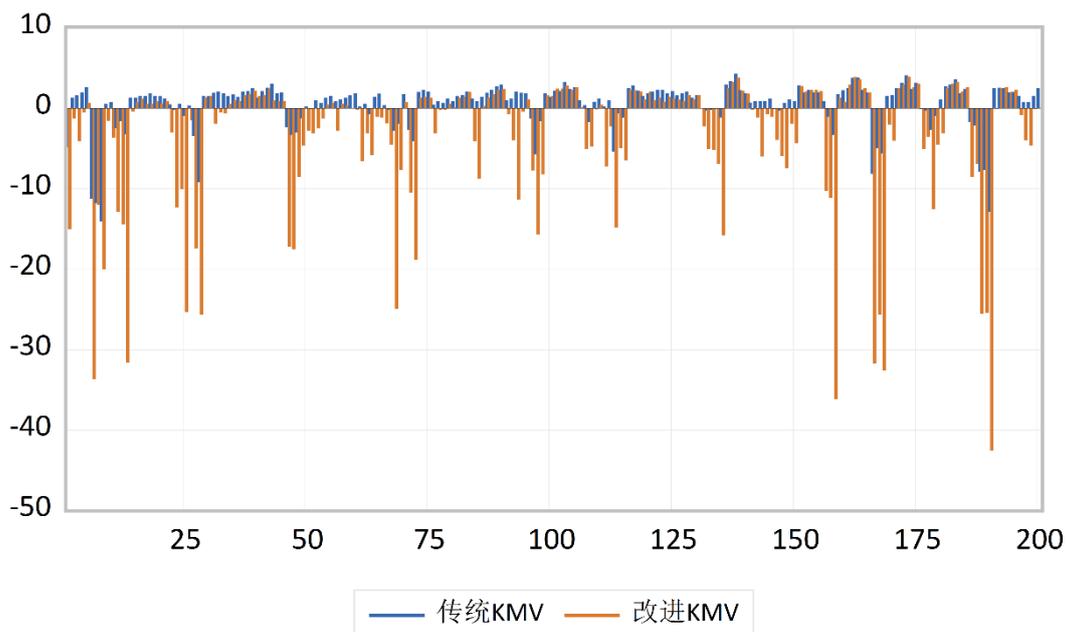


图 4.4 ST 类违约距离传统 KMV 模型跟改进 KMV 模型对比

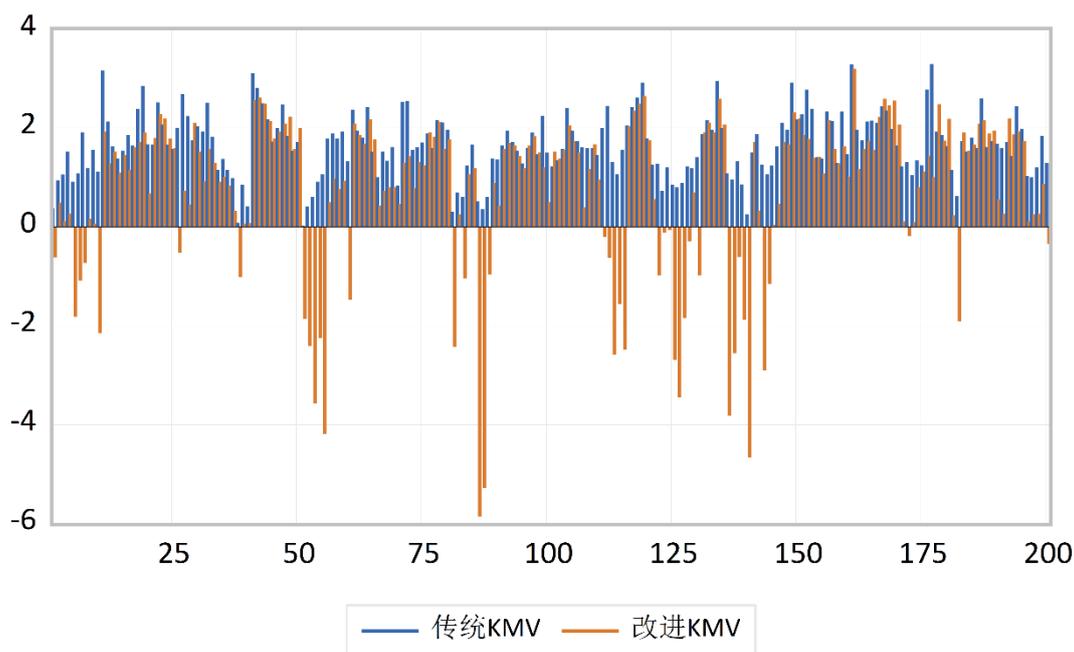


图 4.5 非 ST 类违约距离传统 KMV 模型跟改进 KMV 模型对比

4.3.3 时间序列分析

为清晰追踪上市公司信用状况动态变化状况，本文基于时间序列变化，更进一步计算违约距离，以便直观的观察单个上市公司的信用风险动态变化状况。同样，以*ST 沈机（000410）作为样本，根据该公司 2018 年至 2022 年连续 5 年内

股价及相应的财务数据通过传统 KMV 模型和改进型 KMV 模型，计算每个季度违约距离，并绘制季度违约距离变化趋势图。季度违约距离 KMV 模型参数设定为评估期限为 $\frac{1}{4}$ 、其他均为季度参数。趋势图如图 4.6 所示。

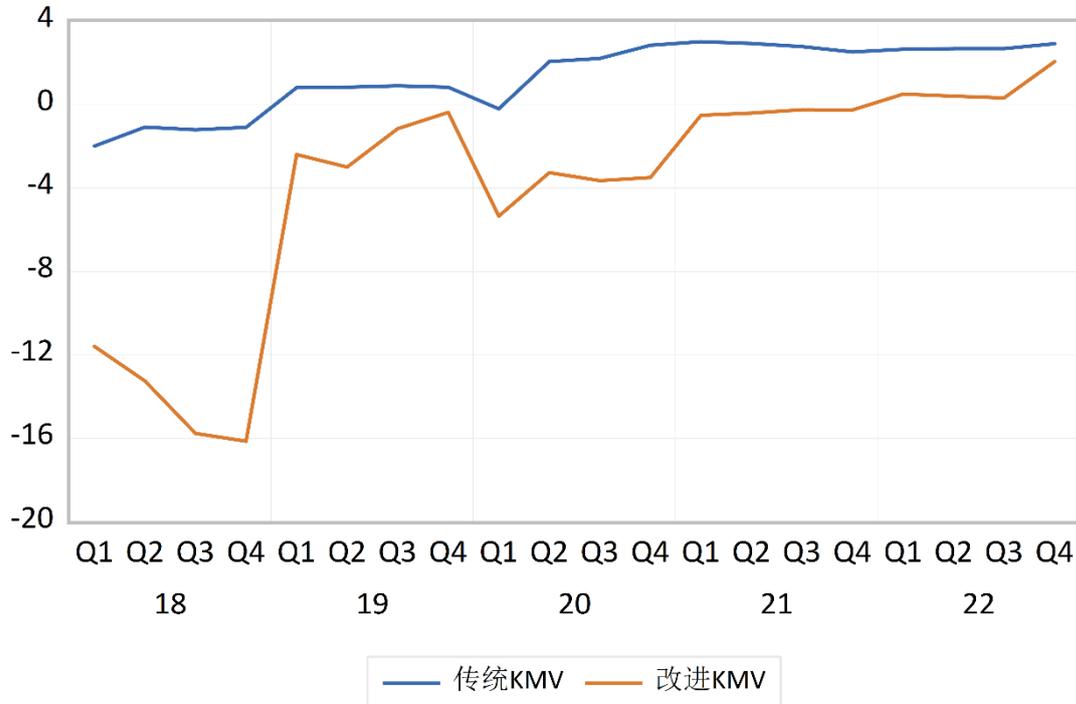


图 4.6 *ST 沈机传统 KMV 模型跟改进 KMV 模型季度违约距离

*ST 沈机未被标注风险之前为沈阳机床，主营业务机床制造，2018 年受宏观经济形势影响，国内机床业务市场规模持续萎缩，沈阳机床由盈转亏，亏损 7.88 亿元，于 2018 年 3 月 1 日“脱星摘帽”一年半再度在 2019 年 8 月 20 重新“披星戴帽”，被实施退市风险警示处理。2019 年 7 月 13 日沈阳机床（集团）收到法院《通知书》，因未能按时清偿债务，缺乏偿债能力，公司债权人向法院申请公司重整，最终，沈阳机床进行了破产重组，于 2019 年加入中国通用技术集团控股有限公司，2015 年至 2019 年已经是沈阳机床第十个亏损的年头。重组之后的沈阳机床在 2022 年再度“戴帽”，最终于 2023 年 5 月份“脱帽”。

配合图 4.6 可以看到，在 2019 年重组前年，改进型 KMV 模型违约距离有个大幅度波折，而传统 KMV 模型却反响平平，从违约距离来看，意味着沈阳机床会有一个潜在的违约风险，这一点改进型 KMV 模型比传统 KMV 模型表现更加明显。在 2019 年第四季度传统 KMV 模型违约距离依然在数值 0 以上，但是改

改进型 KMV 小幅度下降且持续在数值 0 以下，沈阳机床重组之后总体上才有一个缓慢上升的趋势，直至 2023 年*ST 沈机“摘星脱帽”，改进型 KMV 模型在 2022 年第四季度才开始缓慢上升，且浮动到数值 0 以上，传统 KMV 模型从 2018 年开始一直维持在数值 0 以上，波折较少，既不符合已经违约重组的现实因素，也不灵敏反应市场波动。以上从而有效证实，改进型 KMV 模型在信用风险识别方面具有良好的前瞻性且在灵敏度方面比传统 KMV 模型更具有反应能力。

4.4 本章小结

本章第一节首先对样本选取进行了详细的介绍，根据筛选原则选取出 80 家上市公司，并将 80 家上市公司分为 ST 类上市公司和非 ST 类上市公司，同时对实证部分所需的数据进行设定。第二节详细论证了基于 EGARCH 模型和遗传学算法的改进型 KMV 模型的计算过程，包括股权价值波动率的预测和最优违约点的确定。最后对实证结果进行检验和分析，经过检验和分析发现改进型 KMV 模型具有较好的拟合优良性。

5 研究结论及展望

5.1 研究结论

本文以对上市公司信用风险的理论分析作为基础,梳理信用风险预测和度量模型的衍化过程,并选取 80 家上市公司连续五年间数据(高风险组和低风险组各 40 家),使用 EGARCH 模型修正各公司股权价值波动率,同时构建遗传学算法框架对 KMV 模型最优违约点进行确定,最终搭建 EGARCH-GA-KMV 模型,然后利用 EGARCH-GA-KMV 模型测算的违约距离跟传统 KMV 模型测算的违约距离进行对比,以此来评价模型之间适合我国上市公司信用风险预测的拟合优良性。

通过对上文的写作和理解,在综合分析的基础上,得出如下结论:

第一,EGARCH 模型能更好的修正股权价值波动率。在计算 KMV 模型过程中,部分学者使用 GARCH 模型建立 GARCH(1,1)来修正股权价值波动率存在的“尖峰肥尾”问题,但在实际情况中,我国资本市场是明显的“高峰肥尾”问题,同时资本市场可能存在“杠杆效应”,这是 GARCH 模型捕捉不到的。因此,EGARCH 模型更符合波动率的三个特征,更适合修正股权价值波动率。

第二,遗传学算法能很好的模拟出全局最优解。作为一种机器学习算法,遗传学算法不仅可以应用于统计学领域,还可以运用于人工智能领域,该算法合理及未来发展领域宽广。遗传学算法基于其“物竞天择、适者生存、优胜劣汰”原则在构建的遗传学算法最优违约点确定模型中,基于样本数据,遗传学算法很好的得出最优值,在短时间收敛迭代出最优值,也为构建的 EGARCH-GA-KMV 模型更快的得出结果。

第三,修正股权价值波动率跟违约点之后的 KMV 模型能更好的预测我国上市公司信用风险。我国资本市场情况较为特殊,不仅存在限售股情况,资本收益率序列并不满足传统 KMV 模型在资本市场情况的假设,所以使用传统 KMV 模型在预测上市公司信用风险时会出现误差。通过实证研究部分可以发现,修正过后的改进型 KMV 模型相较于传统 KMV 模型在 ST 类中具有更好的拟合性,在 ST 类中相差无几,这说明传统 KMV 模型未能正确反应我国上市公司情况。除此之外,也说明改进型 KMV 模型在样本间具有更好的拟合效果,因此,在使用 KMV 模型对我国上市公司信用风险进行预测时,应注重对股权价值波动率跟违

约点的确定，并不能直接套用。

5.2 未来展望

本文利用遗传学算法和 EGARCH 模型对 KMV 模型予调整，研究结果证明，经过修正之后的模型，在风险评估具有一定的效果。但通过对上文的梳理及分析，本文还具有一定的不足。

首先，样本数量限制导致参数优化不足，修正后的 KMV 模型中关键点的最优违约点是基于样本量得出的全局最优解，具有一定的局限性，并不适合所有上市公司，也不适合非上市公司，样本量跟结果准确度之间存在一定的正向关系。从传统 KMV 模型的构建来看，传统 KMV 模型不仅跟美国实际资本市场相契合，而且基于美国丰富的企业违约数据库确定了违约点，这从侧面提出对我国未来信用风险评估的展望。其次，数据处理不足，改进型 KMV 模型不断优化需要全市场大量的、动态按时更新的违约数据信息，便能做到信用风险动态实时的更新。因此，以上这些都会影响模型优化的效果。同时，基于本文出发有如下展望。

第一，在完善上市公司信息披露机制的同时，更应着手搭建企业违约数据库，不管是从模型的预测，还是事实需求，历史数据在未来发展中具有很大作用。鉴于我国信用风险体系建立时间并不长，我国更应要健全和完善信用违约数据库的建设，减少经济的“暴雷”，增加对经济的“把握”。违约数据库的建设，不仅利于上市公司，更是我国商业银行的重要信息共享平台，我国上市公司最主要的债权主体依然是商业银行，经济牵一发动全身，把握上市公司，某种程度上就是把握银行。

第二，加强 KMV 模型及相似理论的探索及摸索。国情不同，实践跟需要也不同，KMV 模型来源于西方国家的风险管理摸索实践，于中国资本市场而言，并不一定具有全实用性及适用性。所以，该模型修正后才更可能满足中国上市公司信用风险预测德尔需要。如若追求该模型在中国的适用性，则还需进行以下两点理论研究：（1）关系模型应用中国实际效果的合理违约点选取；（2）公司资产价值的预处理方式。同时，对于中国上市公司信用风险的预测还需进行摸索。

总之，本文提供了一个改进 KMV 模型测算的研究思路，其研究是过渡性的，另外，任何模型都存在前提假设，因此，对中国上市公司信用风险预测，满足假设条件及检验，应是我们未来的研究方向。

参考文献

- [1] Altman E I, SAUNDERS A. Credit Risk Measurement: Developments over the Last 20 Years[J]. *Journal of Banking & Finance*, 1997, 21(11/12): 1721-1742.
- [2] Altman, E. I. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy[J]. *The Journal of Finance*, 1968, 23(4): 589-609.
- [3] Altman, E. I., Haldeman R. G. & Narayanan, P. ZETATM analysis A new model to identify bankruptcy risk of corporations[J]. *Journal of banking & finance*, 1977, 1(1): 29-54.
- [4] ANUWAR M H, JAFFAR M M. Grading the probabilities of credit default risk for Malaysian listed companies by using the KMV-Merton model[C]. *AIP Conference Proceedings*. AIP Publishing, 2017, 1870(1): 040025
- [5] C.V. Zavgren. Assessing the vulnerability to failure of American industrial firm: a logistic analysis[J]. *Journal of Business Finance & Accounting*, 1985, 12(1): 19-45.
- [6] CHRISTENSEN B.J, DAHL C M, Iglesias E M. Semiparametric inference in a GARCH-in-mean model[J]. *Journal of Econometrics*, 2012, 167(2): 458-472.
- [7] Crouhy, M., Galai, D. & Mark, R. A comparative analysis of current credit risk models[J]. *Journal of Banking & Finance*, 2000, 24(1): 59-117.
- [8] Douglas Dwyer, Sarah Woo. Analyzing The Subprime Market Fallout Using EDF Credit Measures [Z]. Moody's KMV Corporation, 2007. 9.
- [9] Esa, J., Kimmo, E. & Oskari, V. Macro-model-based stress testing of Basel II capital requirements[J/OL]. *Bank of Finland Research Discussion Paper*, Available at SSRN: <http://dx.doi.org/10.2139/ssrn.1267194>, 2008.
- [10] Jackson, R. H. & Wood, A. The performance of insolvency prediction and credit risk models in the UK: A comparative study[J]. *The British Accounting Review*, 2013, 45(3): 183-202.
- [11] Morgan, J. P. *Creditmetrics-technical document*[M]. New York: JP Morgan, 1997, 4: 2
- [12] Morris, M. & Carter, C. R. Relationship marketing and supplier logistics performance: An extension of the key mediating variables model[J]. *Journal of*

- Supply Chain Management, 2005, 41(4): 32-43.
- [13] Ohlson, J. A. Financial ratios and the probabilistic prediction of bankruptcy[J]. Journal of accounting research, 1980: 109-131.
- [14] SOBCHART J.R.,KEENAN S.,STEN R.M.”Benchmarking Quantitative Default Risk Model:A Validation Methodology”[J].Moody Investors Service document,2000,24(8):23-56.
- [15] W. Beaver. Financial ratios as predictors of failure[J].Journal of Accounting Research,1966,4(Supplement):71-111.
- [16] Zhang, Y., Shi, B. Non-tradable shares pricing and optimal default point based on hybrid KMV models: Evidence from China[J]. Knowledge-Based Systems, 2016, 110: 202-209.
- [17] Zhao, Z., Lan. Y. & Wu. X. The Impact of Electronic Banking on the Credit Risk of Commercial Banks —An Empirical Study Based on KMV Model[J]. Journal of Mathematical Finance, 2016, 6: 778-791.
- [18] Zhiheng Xu,Wen Fan,Fan Zhu. Research on Regional Debt Risk in Hubei Province Based on Modified KMV Model[J]. IOP Conference Series: Materials Science and Engineering,2020,768(5).
- [19] Zmijewski, M. E. Methodological issues related to the estimation of financial distress prediction models[J]. Journal of Accounting research, 1984: 59-82.
- [20] 曹裕,陈霞,刘小静.违约距离视角下的开发性金融信用风险评估[J].财经理论与实践,2017,38(05):14-19.
- [21] 曾诗鸿,李萌.基于KMV法研究节能环保与新能源产业上市公司信用风险[J].中国人口·资源与环境,2014(24):54-57.
- [22] 陈茜,田治威.林业上市企业财务风险评价研究—基于因子分析法和聚类分析法[J].财经理论与实践,2017,38(01):103-108.
- [23] 陈致远,刘成昆,陈蔚.我国保理业务信用风险的识别——基于修正的KMV模型[J].学术研究,2019(2):96-102.
- [24] 段翀.基于零价格概率模型的违约概率测算[J].统计与决策,2020,36(13):61-66.
- [25] 关晓宇,韩淑亚,周昊明.环境违规披露对公司债务违约的影响研究[J].税务与

- 经济,2023,No.248(03):99-105.
- [26] 郭立仑,周升起.商业银行信用风险主要影响因素来自内部还是外部?—基于 KMV 及随机森林模型的实证研究[J].会计与经济研究,2022,36(01):105-124.
- [27] 李森,赵轩维,夏恩君.股权众筹项目融资成功率判别—Logistic 回归与神经网络模型比较分析[J].技术经济,2018,37(09):80-91.
- [28] 凌江怀,刘燕媚.基于 KMV 模型的中国商业银行信用风险实证分析—以 10 家上市商业银行为例[J].华南师范大学学报(社会科学版),2013(05):142-148+209.
- [29] 刘庆富,陈志伟,何畅.中国绿色信贷风险的评估与监测——基于新能源汽车产业的视角[J].复旦学报(社会科学版),2020,62(02):192-200.
- [30] 吕志华,彭建刚.Credit Risk+模型采用 Poisson 分布所产生的经济资本计量误差分析[J].管理评论,2011,23(01):33-40.
- [31] 潜力,冯雯静.地方政府专项债券违约风险——基于 KMV 模型的分析[J].统计与信息论坛,2020,35(07):35-44.
- [32] 苏振兴,扈文秀,夏元婷.基于机器学习的地方政府隐性债务风险先导预警模型[J].统计与决策,2022,38(07):20-25.
- [33] 孙亮,吕丹妮.我国共享经济公司信用风险度量的案例分析——基于 KMV 修正模型的实证研究[J].技术经济,2021,40(06):132-139.
- [34] 唐静,郝洪常.新冠疫情背景下西藏医药上市公司财务危机预警研究——基于“Z-Score”模型[J].西藏民族大学学报(哲学社会科学版),2020,41(04):35-39+76+155.
- [35] 王慧,张国君.KMV 模型在我国上市房地产公司信用风险度量中的应用[J].经济问题,2018(3):36-40.
- [36] 王佳,曹琮予.基于跳跃-扩散 KMV 模型的上市公司信用风险评估[J].技术经济,2022,41(01):160-168.
- [37] 王珂,孟海丽,杨全.模糊 Credit Metrics 模型及其在信用风险评估中的应用[J].金融理论与实践,2016,No.439(02):59-64.
- [38] 王莉.引入违约距离的修正 KMV 模型在财务危机预警中的应用[J].统计与决策,2017,No.485(17):88-91.
- [39] 谢赤,胡扬斌,龙剑友.证券与投资证券投资基金市场风险与信用风险度量及

- 其关系研究[J].财经理论与实践,2019,40(03):52-58.
- [40] 谢赤,凌毓秀.银行信贷资产证券化信用风险度量及传染研究——基于修正 KMV 模型和 MST 算法的实证[J].财经理论与实践,2018,39(03):2-8.
- [41] 谢远涛,蒋涛,杨娟.基于尾部依赖的保险业系统性风险度量[J].系统工程理论与实践,2014,34(08):1921-1931.
- [42] 谢远涛,罗润方,杨娟.基于修正的 KMV 模型的信用风险度量[J].统计与决策,2018,34(15):169-173.
- [43] 杨志安,杨枫.我国地方政府债务风险测算及可持续性分析[J].地方财政研究,2022,No.216(10):55-63.
- [44] 余静文,吴滨阳.数字金融与商业银行风险承担——基于中国商业银行的实证研究[J].产经评论,2021,12(04):108-128.
- [45] 张慧,王宁,孟纹羽.基于分形 KMV 模型的地方政府债券安全发行规模测度[J].统计与决策,2021,37(15):133-136.
- [46] 张建同,张敏,郭卓琦.基于修正 KMV 模型的汽车供应链金融风险分析[J].工业工程与管理,2019,24(01):128-135+143.
- [47] 赵平,孙志峰.中国银行业贷款利率是否反映企业违约风险水平——来自制造业上市公司数据的经验证据[J].金融经济研究,2022,37(02):40-55.
- [48] 邹昆仑,陆萍.基于 Z 模型对中国企业债信用评级的实证研究[J].宏观经济研究,2018,No.234(05):72-83+175.

致 谢

思往事，愁如织；念良久，白驹过隙，遁去三四载。自此往前四年，风华正茂、踌躇满志，然事与愿违，终成南柯一梦；后疫疾肆虐，民不聊生，百念皆灰，而为者常成，行者常至，遂心如铁石，再战沙场，终见分晓。行文至此，聊表谢意。

鹤瘦松青，精神与秋月争明。首先，致三年前的自己，是在绝境下愈战愈勇的自己；是在夜深时挑灯苦读的自己；是在冬日里早起背书的自己；是在夕阳后坚持锻炼的自己；是在祖国各地日夜奔波的自己...。再致兰州财经大学，背负着“仰望星空”的期望，在彷徨迷茫时，是兰州财经大学抛出橄榄枝，给予我希望。

仰之弥高，钻之弥坚。是导师带我进入研究生“仰”“钻”的世界，是导师用他那深厚的人文情怀引导我学习、教导我为人处世之道。同时，也谢各位答辩专家、课程教学及辅导老师，是各位专家老师给予我中肯的意见，开拓思路，是各位课程教学老师让我畅游研究生学习的世界，是辅导老师给予学校生活上的帮助。片纸有尽，而感激之情无尽！

愿岁并谢，与友长兮。研究生三年，得到很多人的帮助，是同学，是好友，亦是挚友！各位或来自山川湖海、或来自天南海北，但皆万众一心、志同道合，汇聚于共同的梦想，在人生的大道上奔赴。独学无友，则孤陋寡闻；此经别离，愿相逢不语，芙蓉著秋雨。

流年似水，岁月蹉跎；愿各位以梦为马，不负韶华；愿祖国祥风时雨，麦穗两岐；愿未来宜诗宜歌，无疾无疫。愿来日方长，万事胜意！