

---

# 随机退化过程与统计建 模

---

徐安察 汤银才 庄亮亮

科学出版社  
2025 年 4 月



# 目录

<b>前言</b>	<b>1</b>
<b>第一章 引言</b>	<b>3</b>
1.1 退化建模方法回顾 . . . . .	4
1.2 一般路径模型 . . . . .	6
1.3 随机过程模型 . . . . .	6
1.3.1 维纳过程模型 . . . . .	7
1.3.2 伽马过程模型 . . . . .	8
1.3.3 逆高斯过程模型 . . . . .	9
1.3.4 指数分散过程模型 . . . . .	9
1.4 数据集介绍 . . . . .	10
1.4.1 激光退化数据 . . . . .	10
1.4.2 碳膜电阻器退化数据 . . . . .	10
1.4.3 有机发光二极管退化数据 . . . . .	11
1.4.4 列车车轮退化数据 . . . . .	12
1.4.5 设备-B 数据/集成电路数据 . . . . .	13
1.4.6 铝合金裂纹退化数据 . . . . .	13
1.4.7 锂离子电池容量退化数据 . . . . .	13
1.4.8 应力关系退化数据 . . . . .	14
1.4.9 两阶段模式的锂离子电池退化数据 . . . . .	15
<b>第二章 基于维纳过程的统计建模</b>	<b>17</b>
2.1 维纳过程 . . . . .	17
2.2 广义维纳过程 . . . . .	19

2.2.1	模型构建 . . . . .	20
2.2.2	统计推断 . . . . .	24
2.2.3	模拟实验 . . . . .	29
2.2.4	实例分析 . . . . .	30
2.2.5	附录 . . . . .	35
2.3	二阶段维纳退化过程 . . . . .	42
2.3.1	模型构建 . . . . .	42
2.3.2	多层贝叶斯推断 . . . . .	46
2.3.3	模拟实验 . . . . .	48
2.3.4	实例分析 . . . . .	49
2.3.5	附录 . . . . .	54
<b>第三章</b>	<b>基于伽马过程的统计建模</b>	<b>61</b>
3.1	伽马过程 . . . . .	61
3.2	伽马过程的贝叶斯分析 . . . . .	62
3.2.1	共轭先验 . . . . .	63
3.2.2	后验抽样 . . . . .	67
3.2.3	模拟实验 . . . . .	70
3.2.4	异质性 . . . . .	73
3.2.5	RUL 预测 . . . . .	77
3.2.6	实例分析 . . . . .	78
3.2.7	附录 . . . . .	79
3.3	重参化伽马过程的变分贝叶斯推断 . . . . .	82
3.3.1	重参化伽马过程 . . . . .	84
3.3.2	变分贝叶斯推断 . . . . .	86
3.3.3	模拟实验 . . . . .	94
3.3.4	实例分析 . . . . .	97
3.3.5	附录 . . . . .	100
<b>第四章</b>	<b>基于逆高斯过程的统计建模</b>	<b>116</b>
4.1	逆高斯过程 . . . . .	116
4.2	子总体异质性下的逆高斯过程 . . . . .	117

4.2.1	模型构建 . . . . .	118
4.2.2	统计推断 . . . . .	120
4.2.3	模拟实验 . . . . .	127
4.2.4	实例分析 . . . . .	127
4.2.5	附录 . . . . .	134
4.3	二阶段逆高斯过程 . . . . .	138
4.3.1	研究背景 . . . . .	138
4.3.2	模型构建 . . . . .	140
4.3.3	统计推断 . . . . .	145
4.3.4	模拟实验 . . . . .	150
4.3.5	实例分析 . . . . .	153
4.3.6	附录 . . . . .	157
4.4	在线估计与 RUL 预测 . . . . .	167
4.4.1	模型设定 . . . . .	167
4.4.2	在线估计 . . . . .	168
4.4.3	考虑随机效应的在线估计 . . . . .	174
4.4.4	模拟实验 . . . . .	180
4.4.5	实例分析 . . . . .	181
4.4.6	附录 . . . . .	183
<b>第五章</b>	<b>基于指数分散过程的统计建模</b>	<b>208</b>
5.1	指数分散过程 . . . . .	208
5.1.1	定义及相关性质 . . . . .	209
5.1.2	统计推断 . . . . .	213
5.1.3	加速模型 . . . . .	216
5.1.4	模拟实验 . . . . .	217
5.1.5	实例分析 . . . . .	218
5.1.6	带随机效应的指数扩散过程 . . . . .	222
5.2	动态使用环境下系统可靠性评估 . . . . .	224
5.2.1	模型构建 . . . . .	226
5.2.2	部件相关性 . . . . .	226
5.2.3	系统可靠度 . . . . .	231

5.2.4	参数冗余性	233
5.2.5	模拟实验	237
5.2.6	实例分析	239
5.3	部件寿命排序约束下系统可靠性评估	240
5.3.1	模型	241
5.3.2	统计推断	245
5.3.3	模拟实验	250
5.3.4	实例分析	250
<b>第六章 总结与展望</b>		<b>261</b>
<b>参考文献</b>		<b>262</b>
<b>索引</b>		<b>283</b>

# 前言

在当今技术迅猛发展的背景下,系统结构和运行环境日益复杂,可靠性评估作为衡量产品质量的核心指标,已成为保障系统长期稳定运行的重要环节.随着我国制造业的转型升级和国际竞争力的提升,可靠性评估对推动质量强国建设具有重要作用. 中共中央、国务院于 2023 年发布的《质量强国建设纲要》明确提出要提高大型装备、基础零部件及元器件的可靠性水平,以增强企业核心竞争力,推动制造业的高质量发展. 在这一背景下,如何在产品全生命周期中科学评估其可靠性,尤其是在产品未发生失效前的退化阶段进行有效监测和预测,成为可靠性工程中的关键任务.

本书旨在系统讨论如何利用统计模型来刻画和预测系统的退化行为. 与传统方法相比,随机退化过程模型因其能刻画性能随时间变化的不确定性,在可靠性建模中具有天然优势. 通过合理选择和构建这些模型,可有效识别系统在不同工作环境与使用条件下的性能退化模式,实现更精确的寿命预测与可靠性评估. 全书围绕四类具有代表性的随机过程展开建模与分析: 维纳过程、伽马过程、逆高斯过程与指数扩散过程. 这些模型具有良好的数学可解释性与实际适应性,能够有效捕捉退化行为的连续性、累积性和随机性,尤其适用于寿命预测与剩余寿命评估等关键任务. 书中的核心章节具体内容如下:

第一章主要介绍了研究背景和当前可靠性评估的难点. 首先回顾了基于失效数据的评估方法,并引入了基于性能退化数据的可靠性评估理念;然后,详细讨论了物理模型、人工智能方法与数据驱动统计方法三大类退化建模方法,并对相关文献进行了综述;最后,介绍了本书中所用的数据集以及相关的研究问题.

第二章聚焦于基于维纳过程的退化建模. 首先从标准维纳过程出发,构建用于刻画线性与非线性退化路径的基础模型;进而通过引入退化速率与波

动性的相关性,提出一类广义维纳过程模型及其统计推断方法;同时,为应对退化过程中的阶段性突变,构建了两阶段维纳过程模型,并设计了多层贝叶斯推断方法进行参数估计。

第三章主要介绍基于伽马过程的统计建模方法.首先介绍了伽马过程的定义及相关性质;然后在共轭先验下讨论伽马过程的贝叶斯推断以及 RUL 预测方法;最后,通过重参数变换讨论一类带随机效应的重参数化伽马过程及其变分贝叶斯推断方法.

第四章主要介绍基于逆高斯过程的统计建模方法.首先介绍了逆高斯过程的定义及基本性质;然后讨论一类具有子群体异质性的逆高斯过程;接着,针对退化变点现象,讨论一类阶段重参数化逆高斯过程,并结合 RUL 分布提出一种自适应替换策略;最后,针对逆高斯过程讨论了一种具有递归更新的在线算法,并用于系统 RUL 预测.

第五章主要介绍基于指数扩散过程的统计建模方法.首先介绍了非线性指数分散过程的定义及基本性质,并讨论了其统计推断方法和在退化数据建模中的应用;在此基础上,针对复杂使用环境和系统内部寿命排序约束条件,进一步构建了基于指数分散过程的系统可靠性评估模型及统计推断方法.

本书力图将理论推导与实际工程应用相结合,选取了多个具有代表性的数据集,包括激光器电流退化数据、碳膜电阻器温度应力退化数据、有机发光二极管亮度下降数据、列车车轮磨损数据、锂离子电池容量衰减数据等,分别用于模型的拟合、参数估计和寿命预测。通过具体的案例分析,详细展示了模型构建、推断与预测的完整流程,旨在帮助读者全面理解如何将理论模型有效应用于实际工程问题中。

本书主要面向从事可靠性工程、统计建模与智能制造等领域的科研人员、工程技术人员及研究生,力求在理论深度与工程实用性之间取得平衡,既提供系统的建模理论,也给出可操作的应用方法。由于作者水平有限,书中内容定有诸多不足之处,恳请各位专家学者与工程实践者批评指正,提出宝贵意见,以助后续修订与完善。

# 第一章 引言

可靠性作为评估产品质量的核心指标,能够全面反映产品的性能、稳定性和耐用程度,是衡量一个国家制造业发展水平的重要标准.为了加快推进质量强国建设,提升我国制造业的国际竞争力,中共中央、国务院于 2023 年 2 月联合发布了《质量强国建设纲要》.该纲要明确提出要实施质量可靠性提升计划,重点提高大型装备、基础零部件以及元器件的可靠性水平.这一计划的实施,将有助于推动我国制造业的转型升级,提升产品质量和可靠性,增强企业的核心竞争力.可靠性评估作为产品质量提升的关键环节,对于确保产品在全寿命周期内的稳定性和可靠性具有至关重要的作用.通过科学合理的可靠性评估方法,可以及时发现产品在设计、生产、使用过程中存在的问题,为改进和优化提供依据.同时,可靠性评估还可为企业制定产品标准和质量控制提供重要参考,促进产业整体水平的提升.因此,实施质量可靠性提升计划,加强可靠性评估工作,对于推动我国制造业高质量发展、提升国家整体竞争力具有重要意义.

传统的产品可靠性评估方法主要依赖于失效数据,即在产品出现故障或失效时收集的数据.这种方法通常基于一定的寿命分布假设,如指数分布、威布尔分布或对数正态分布等,来推断产品的可靠度(张志华, 2002; 范诗松等, 2008).然而,随着生产技术的进步和产品质量标准的提高,产品的可靠度不断提升,导致在一定的试验时间内很难收集到失效数据来支持这种传统的可靠度评估.虽然很多国内外学者提出了处理零失效数据的统计方法,如 Miller 等(1992)建立“黑盒子模型”获取失效概率的先验分布,减少了主观因素对可靠性评估结果的影响; 韩明(2003)提出一种可靠性指标估计的综合贝叶斯法,通过引入经验信息来估计产品失效概率,避免了可靠性评估中出现的异常现象; 李海洋等(2018)在置信限分析中结合失效信息来修正单侧置信限,得到产

品可靠性单侧置信限的稳健估计,但是这些方法都依赖大量的先验信息,并不能从根本上解决失效信息缺乏所带来的可靠性评估问题.为此,国内外学者开始关注基于产品性能退化数据的可靠度评估方法.与失效数据不同,性能退化数据是在产品正常使用过程中监测到的性能参数变化值.这些数据反映了产品性能随时间的退化趋势,从而提供有关产品可靠性的有用信息.基于性能退化数据的可靠度评估方法主要通过对产品性能参数的连续监测和数据分析来实现.通过收集产品在不同时间点的性能退化数据,建立性能退化模型来描述产品性能随时间的变化规律,推导出产品的可靠度函数,从而对产品的可靠性进行评估(Meeker 等, 1998; Peng 等, 2009; Ye, Xie, 2015).与基于失效数据的传统方法相比,基于性能退化数据的可靠度评估方法具有以下三个优势:首先,通过连续监测产品的性能参数变化,能够在产品尚未发生实际失效时提供更准确、更及时的可靠性评估.这不仅缩短了评估周期,还提高了评估效率和精度.其次,它可以在产品运行的早期阶段发现潜在问题,及时采取必要的维护措施,防止突发失效的发生,延长产品的使用寿命,降低维护成本.同时,这也为企业提供了更科学的维护策略制定依据,有助于优化维修计划和库存管理.最后,通过对性能退化数据的分析,可以深入了解产品的性能变化趋势和失效机理.这为产品的设计和改进提供了宝贵的反馈信息,有助于发现设计中的薄弱环节和潜在问题,指导产品的优化设计和持续改进.因此,研究基于性能退化数据的可靠度评估方法对于提高产品可靠性、指导维护和预防失效、以及促进产品设计和改进都具有重要意义.

## 1.1 退化建模方法回顾

基于系统性能特征退化数据的建模方法主要可分为三类:物理模型、人工智能方法和数据驱动的统计方法.基于物理模型的方法以理解产品的物理结构、工作原理和环境条件为基础,从而揭示潜在的物理故障过程.虽然物理模型能够准确描述系统失效规律和性能指标之间的相互关系,但由于现代工业设备的复杂性,获得失效物理模型的解析形式通常具有挑战性,因而在实际应用中受到一定限制.近年来,人工智能方法在性能退化建模方面取得了显著进展(Chao 等, 2022; Nguyen 等, 2022; Xia 等, 2022),它通过对数据进行模型训练,建立与可靠性相关的映射关系,并考虑各性能指标之间的相依性.然而,在实际退化数据分析中,该方法面临以下挑战:1. 要求大量高质量的训

练数据和准确的事件标记, 同时需要大量时间训练模型. 2. 主观假设故障发生时间已知, 但错误的时间标记可能引入噪声, 影响可靠性估计的准确性. 3. 模型泛化能力的有限, 特别是在面对新领域或未见场景时的预测准确性 (Ren 等, 2024). 数据驱动的统计方法通过分析系统的歷史故障信息或性能退化数据建立统计模型, 用于评估系统的可靠性, 并为确保系统稳定性、制定质保策略和预防性维修策略提供理论依据 (Wen 等, 2022). 在基于故障事件数据的建模中, 通过收集系统故障的时间, 运用统计方法来推断故障时间的概率分布, 并将其用于后续的决策分析 (Lewis-Beck 等, 2022; Peng Y 等, 2017). 尽管这种方法能够描述整体故障时间的分布趋势, 但它无法为特定系统提供准确的故障预测和可靠性评估. 在基于性能退化数据的建模中, 通过监控系统性能退化状态, 建立退化模型并学习系统性能退化的演化规律, 从而评估系统的可靠性 (Zhang 等, 2021). 该方法具有两个显著的优点: 1. 性能退化数据的易收集性: 随着传感器和监测技术的不断进步, 我们能够实时监控和收集性能退化数据. 这种随时获取退化特征的能力极大地简化了数据收集的过程, 并提高了数据的准确性. 2. 性能退化与系统寿命的直接关联: 性能退化所反映的是系统在使用过程中状态的变化. 这种变化与系统的寿命息息相关, 因此, 通过分析性能退化信息, 我们能够实时更新系统寿命的估计, 实现对系统可靠性的实时预测. 这为及时采取维修策略和管理系统可靠性提供了有力支持. 下面章节将对性能退化建模方法的研究现状进行分析, 对应的综述框架见图 1.1.

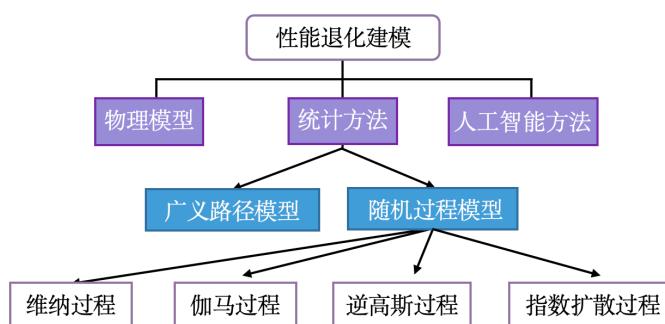


图 1.1: 性能退化建模方法框架.

## 1.2 一般路径模型

早期的退化数据建模主要借鉴了随机效应模型的思想: 通过物理或经验分析, 确定产品性能退化与时间的函数形式, 并在此基础上引入随机效应和观测误差, 从而建立产品寿命与性能退化之间的定量关系. Ye, Xie (2015) 将这种方法称为一般路径模型(General path model). Lu 等 (1993) 首次对一般路径模型进行了详细的阐述, 并说明了该模型可以灵活地刻画测量误差和产品间的异质性. 之后, 该模型得到了广泛关注, 并被应用到更复杂的退化数据建模中 (Robinson 等, 2000). 例如: Hong 等 (2015) 引入动态协变量来刻画多种应力 (如: 温度, 湿度和紫外线强度) 对某种环氧树脂涂层退化的影响. Liu 等 (2018) 将该模型拓展到时空退化数据中, 构建了一个由退化过程和传播过程相结合的时空模型. Si 等 (2018) 提出了一种多变量一般路径模型, 用于描述材料拉伸过程中的动态局部变形, 并通过两阶段方法验证其有效性和适用性. 在此基础上, Lu 等 (2020) 结合非线性路径协变量与特征间随机效应 对模型做了拓展, 并利用 EM 算法与 MCMC 模拟实现参数估计和系统可靠性分析. 一般路径模型能够灵活地融合协变量和单元间异质效应, 使其适用于多样化的实际场景, 但是它在计算上存在一定的复杂性, 尤其是在处理大规模多维退化数据时, 可能需要较高的计算资源. 此外, 在实际应用中, 产品的性能退化会受到环境变化、外部因素或未知因素的影响. 当这些因素不能被准确记录或量化时, 一般路径模型就难以刻画这种随机性引起的产品性能退化特征 (Ye, Xie, 2015; Zhai, Xu, 等, 2023).

## 1.3 随机过程模型

产品性能特征的退化通常由多次微小损伤逐步累积所导致, 这一过程既具有累积性又带有随机性. 退化过程不仅受到外部环境的影响, 还与材料本身的性质及其工作状态密切相关. 因此, 随机过程作为一种能够有效描述这种随机性和累积性的数学工具, 在物理意义上与产品退化特征高度契合, 并逐渐成为研究退化行为的核心方法之一.

近年来, 随着统计学和随机过程理论的不断发展, 维纳过程、伽马过程、逆高斯过程以及指数分散过程 (Exponential dispersion process) 等模型因其独特的数学性质和对实际问题的适用性而受到广泛关注. 这些模型能够有效地

捕捉和描述产品在长期使用过程中的退化行为,尤其是在描述不同类型的随机损伤积累机制上展现出了重要优势. 国内外学者通过大量研究, 已经在这些随机过程模型的拓展与应用、参数估计以及剩余使用寿命预测等方面取得了丰富的成果. 本节将对这些常用的随机过程模型进行文献综述, 具体的模型理论和相关应用将在后续章节中进一步展开。

### 1.3.1 维纳过程模型

维纳过程及其扩展形式最早用于描述产品性能的退化机制. 由于维纳过程在物理解释和数学性质上的优越性, 过去 20 年来, 它在退化数据分析方面得到了广泛而深入的研究. 基本的维纳过程假设所有样本的退化均值和波动性是同质的, 且退化路径为线性. Doksum 等 (1992) 使用该模型拟合绝缘电缆寿命数据, 并证明其寿命分布为逆高斯分布. 更多该模型的应用见 Guan 等 (2016). 然而, 许多产品的退化路径表现出非线性的特征. 为此, 部分学者提出了基于时间尺度变换的非线性维纳退化模型. 例如, Whitmore 等 (1997) 对维纳过程在时间尺度上进行幂变换, 并利用该模型分析了自控温加热电缆的寿命. Tseng 等 (2007) 也进行了类似研究, 但他们采用了指数变换对维纳过程进行时间尺度上的调整.

产品在实际使用过程中, 其性能退化往往受到多种潜在因素的共同影响, 导致退化路径呈现出显著的异质性. 为了在退化建模中刻画这种异质性, 通常会在模型中引入随机效应, 即假设某些模型参数服从特定的概率分布. Peng 等 (2009) 研究了带随机效应的维纳过程模型, 并系统分析了模型假设错误对参数估计的影响; Wang (2010) 在非线性维纳过程模型中引入随机的漂移和波动参数, 提出的模型能够较好地拟合桥梁裂纹数据. Ye, Chen, 等 (2015) 在激光设备性能退化数据中发现模型的漂移参数与波动参数之间存在较强的相依关系, 并提出了一种参数相依的维纳过程退化模型. Zhai, Chen, 等 (2018) 从加速失效机理的角度进一步拓展了 Ye, Chen, 等 (2015) 的模型. 随后, Zhou 等 (2021) 提出了更为一般的非线性参数相依维纳退化模型, 并给出了其寿命分布的近似表达式. 此外, Paroissin (2015) 观察到某些产品的初始退化值具有随机性, 因此提出了具有随机初始值的维纳退化模型. Shen 等 (2018) 在更一般的随机退化框架下进一步研究了该现象. 更多关于随机效应维纳退化模型的研究, 可参考 (Pan 等, 2017; Tang, Guo, 等, 2014; Wang 等, 2020; Zhang 等, 2018).

由于设备测量精度的限制以及试验设备的老化,退化测量过程中不可避免地会产生测量误差.针对这一现象,Whitmore (1995)建议在基本维纳过程模型中引入时间独立的测量误差,以更准确地描述退化数据. Peng 等 (2012) 在此基础上进一步研究了该模型,并推导出了其方差协方差矩阵的解析表达形式. Ye, Wang, 等 (2013) 对 Whitmore (1995) 的模型进行了扩展,提出了一种同时考虑测量误差与随机效应的非线性维纳退化模型. 随后, Peng 等 (2020) 在 Ye, Wang, 等 (2013) 模型的基础上进一步引入了多层随机效应,并提出了一种基于学生化  $t$  随机过程模型,可有效拟合具有厚尾特征的退化数据. 更多关于带测量误差的维纳退化模型研究及其应用,可参考 (Li J 等, 2017; Li 等, 2018; Tang, Yu, 等, 2014; Yan 等, 2020).

### 1.3.2 伽马过程模型

伽马过程在退化建模中常用于拟合具有单调退化路径的产品,其因能描述逐步累积的退化特性而备受关注. Abdel Hameed (1975) 最早提出使用伽马过程进行退化建模. 随后, Singpurwalla (1997) 从方法论角度对伽马过程进行了深入阐述,并指出其能够很好地拟合某种机翼合金材料的退化行为. Park 等 (2014) 基于齐次伽马过程推导了产品寿命分布的解析表达,并指出该分布可进一步用 Birnbaum-Sauders(BS) 分布近似. 为考虑产品间的异质性, Lawless 等 (2004) 提出了带随机效应的线性伽马退化模型. Tsai 等 (2012) 基于随机效应非线性伽马退化模型,研究了退化试验的优化设计问题. Guida 等 (2019) 从贝叶斯角度研究了非线性伽马退化过程的参数估计问题,并提出了一种估计剩余寿命的近似方法. 进一步地, Fan 等 (2017) 利用贝叶斯方法解决了步进应力加速退化试验下的伽马退化模型统计推断问题. 更多相关研究可参考 (Hao 等, 2015; Wang 等, 2021).

在测量误差方面, Pulcini (2016) 假设测量误差服从一阶随机游动过程,并提出了一种具有时间相关测量误差的伽马退化模型. Giorgio 等 (2019) 在非高斯测量误差的假设下进一步扩展了伽马退化模型. 然而,在伽马过程中添加时间独立的测量误差通常导致似然函数难以解析表示,从而增加了模型参数估计的难度. 为此, Hazra 等 (2020) 利用近似贝叶斯计算方法,解决了带时间独立测量误差的伽马退化模型的统计推断问题. 除了在退化建模中的广泛应用外,伽马过程还在其他领域发挥了重要作用,如统计过程控制 (Chen 等, 2018a) 和

预防性维修 (Van Noortwijk, 2009; Yuan 等, 2021) 等.

### 1.3.3 逆高斯过程模型

逆高斯过程是另一种常用来拟合单调退化路径的随机过程. Wang X 等 (2010) 最早将逆高斯过程引入退化建模领域, 并发现对于某些具有单调退化路径的高可靠性产品, 基于逆高斯过程的退化模型在拟合效果上优于伽马过程模型. Ye, Chen N (2014) 进一步证明了逆高斯过程是复合泊松过程的极限形式, 从而为逆高斯退化模型提供了清晰的物理解释. 他们还基于维纳过程构建了两种具有解析结构的随机效应逆高斯退化模型. Peng 等 (2014) 从贝叶斯的角度对 @ye2014inverse 提出的逆高斯退化模型进行了深入分析. 此外, Peng (2015) 研究了同时包含协变量和随机效应的逆高斯退化模型, 并推导了对应的寿命分布及平均失效时间的解析表达形式. 更多关于该模型的研究可参考 (Guo 等, 2018; Ma Z 等, 2019; Peng W 等, 2017; Ye, Chen L-P, 等, 2014).

在测量误差的研究方面, Sun 等 (2021) 在时间相关测量误差的假设下, 研究了带测量误差与随机效应的逆高斯退化模型. Hao 等 (2019) 则在不同的随机效应设置下进一步探讨了带测量误差的逆高斯退化建模方法. 此外, Peng 等 (2019) 提出了一种改进的逆高斯过程退化模型, 该模型假设观测量在特定的非线性变换下服从逆高斯过程, 从而增强了对复杂退化行为的刻画能力. 当系统具有多个性能退化指标时, Fang 等 (2022) 观察产品多个性能指标存在固有异质性和相依性, 并提出了一种具有多元随机效应的逆高斯退化模型, 以更全面地描述系统性能的退化行为.

### 1.3.4 指数分散过程模型

指数分散过程是一类广义的随机过程, 前述的维纳过程、伽马过程和逆高斯过程均为其特例 (Tseng 等, 2016). 当指数分散过程的单位方差函数呈幂函数形式时, 被称为 Tweedie 指数分散过程, 其中参数  $p$  的不同取值对应特定的随机过程:  $p = 0, 2, 3$  时, 分别对应维纳过程, 伽马过程和逆高斯过程. 目前, 指数分散过程在退化建模中的研究尚不够深入. Tseng 等 (2016) 基于指数分散过程研究了加速退化试验中的样本分配问题, 并针对三种常用的优化准则 (V-最优, D-最优, 和 A-最优), 给出了最优样本分配方案. Lee 等 (2020) 进一步探讨了基于指数分散过程的加速退化试验设计, 在两个和三个应力水平下,

通过最小化产品寿命分布分位点估计的渐近方差, 给出了各应力水平下的全局最优样本分配方案. Hong 等 (2017) 将指数分散过程推广应用于加速退化试验设计, 统一了基于维纳过程、伽马过程和逆高斯过程的加速退化试验最优设计方法. 随后, Zhou 等 (2019) 系统阐述了非齐次指数分散过程的统计性质, 并探讨了具有随机效应的指数分散过程的统计推断方法. 在考虑产品异质性方面, Duan 等 (2018) 研究了具有随机效应的线性指数分散过程, 以更好地刻画产品间的变异性. Chen 等 (2021) 则进一步探讨了具有随机效应的非线性指数分散过程, 同时研究了同时包含加速退化和随机效应的指数分散过程的统计推断问题. 此外, Xu 等 (2021) 利用指数分散过程的性质, 提出了适用于随机环境下的系统可靠性评估框架.

## 1.4 数据集介绍

### 1.4.1 激光退化数据

该数据集主要用于本书的第 2.2 节、3.2 节和 5.1 节, 来源于 Meeker 等 (1998). 数据集记录了在 80°C 高温条件下测试的 15 个 GaAs (砷化镓) 激光器的运行电流随时间增加的百分比变化情况. 在时间  $t = 0$  时, 所有样本的电流增加百分比均设定为 0, 此后每隔 250 小时测量一次运行电流的增加量, 直至试验终止时间 4000 小时. 所有激光器的性能退化过程如图 1.2 所示. 激光器的性能退化表现为运行电流随时间的逐渐增加, 以维持恒定的光输出. 当电流增加量超过初始值的 10% 时, 激光器被认为失效. 该数据集在分析激光器的退化模式、寿命预测以及识别潜在失效机制等方面具有重要研究价值.

### 1.4.2 碳膜电阻器退化数据

该数据集主要用于本书的第 2.2 节, 数据来源于 Meeker 等 (1998), 列于其书中表 C.3. 数据集记录了碳膜电阻器在不同温度条件下电阻随时间的变化情况, 包括 30 个电阻器单元在 83°C、133°C 和 173°C 三个温度水平下的电阻增加百分比. 试验首先记录了每个电阻器的初始电阻值, 随后分别在 452 小时、1030 小时、4341 小时和 8084 小时时, 以初始电阻值的百分比形式记录电阻增加情况. 该数据集可用于评估不同温度条件下碳膜电阻器的性能退化趋势, 帮助研究人员分析温度对电阻器退化速率的影响. 图 1.3 展示了 9 个碳膜

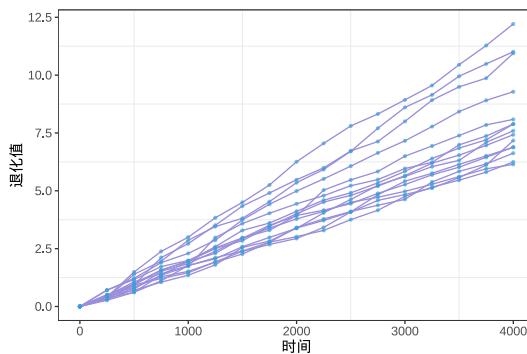


图 1.2: 激光工作电流的退化路径.

电阻器在 83°C 条件下的电阻增加百分比退化路径.

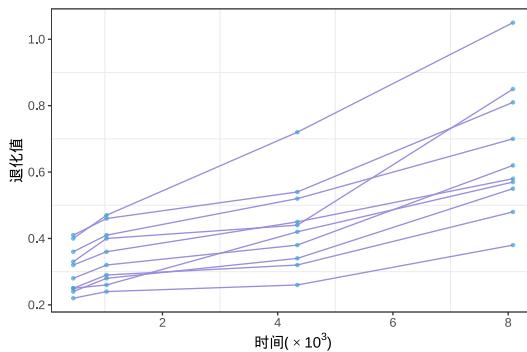


图 1.3: 9 个碳膜电阻的电阻百分比增加值

### 1.4.3 有机发光二极管退化数据

该数据集主要用于本书的第 2.3 节, 来源于 Bae 等 (2008). 该数据集主要用于评估有机发光二极管 (Organic light-emitting diode, OLED) 在恒定应力水平 (直流电流 =25mA) 下的可靠性. 该数据集包含了 6 个用于移动电话显示元件的 OLED 样本单元, 在 20 个时间点的光度测量值. 测试样本被分配在在四种加速应力水平: 25 mA, 32 mA, 40 mA 和 50 mA. 各直流电流下的退化路径如图 1.4 所示. 行业标准规定, 在相对光度低于 0.5 或低于 50% 时失效.

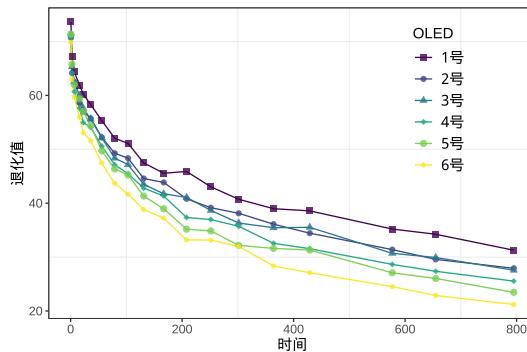


图 1.4: 6 个 OLED 的退化数据

#### 1.4.4 列车车轮退化数据

该数据集主要用于本书的第 3.2 节, 来源于 (Peng 等, 2020), 记录了 14 个列车车轮在行驶过程中直径磨损情况. 测试以 50 千公里为间隔, 直至累积里程达到 1000 千公里. 在每个里程节点, 分别测量每个车轮的直径. 当直径磨损达到 60 毫米时, 即视为车轮失效. 由于其中三个样品的车轮直径磨损过快, 很早便超过了 60 毫米, 故在数据集中移除了这三组数据. 剩余 11 个样品的磨损数据如图 1.5 所示. 该数据集可用于分析列车车轮的退化趋势和剩余使用寿命, 有助于科学规划车轮的维护与更换策略, 从而提升铁路运输的安全性与运营效率.

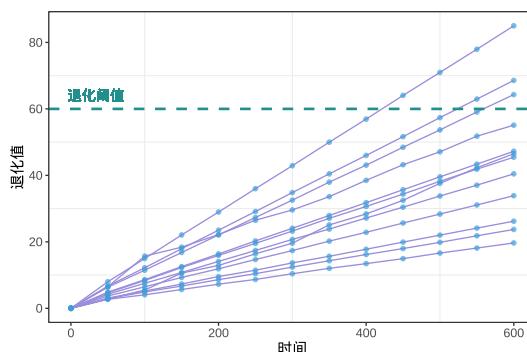


图 1.5: 列车车轮退化路径.

### 1.4.5 设备-B 数据/集成电路数据

数据集主要用于本书的第 3.3 节、第 4.1 节、4.4 节, 来源于 Meeker 等 (1998). 该数据集专注于通信卫星中使用的射频功率放大器在不同温度条件下的性能和可靠性. 工程师们将功率输出下降 0.5 分贝 (dB) 以下的初始输出时间点定义为故障时间, 并随机选取了 67,000 个设备-B 样本位置, 在 150°C、195°C 和 237°C 这三个极端温度条件下, 以每 125 小时为间隔进行测量, 直至累计达到 4,000 小时. 图 1.6 展示了在所有温度条件下产品的退化路径.

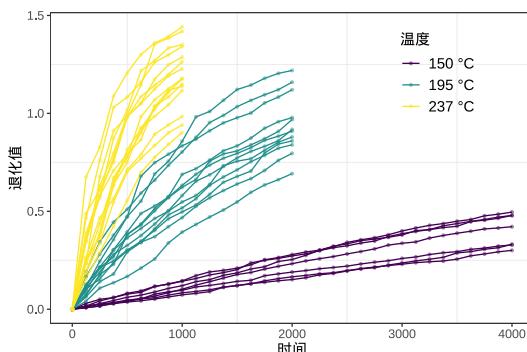


图 1.6: 所有温度下设备-B 数据退化路径.

### 1.4.6 铝合金裂纹退化数据

该数据集主要运用于本书的第 4.1 节、5.1 节. 数据可以在 (Wu 等, 2003) 的表 1 中找到. 实验中, 选用了 30 个抛光处理的 2024-T351 铝合金样本, 在室温环境下利用统一测量设备进行测试. 实验设计以 10,000 次载荷循环为时间单位, 测试范围覆盖 1 至 4 单位, 步长设定为 0.5 单位. 当样本裂纹长度超出 15 毫米时, 即判定为失效并终止测试, 此时的循环次数记录为样本寿命, 其退化路径如图 1.7 所示.

### 1.4.7 锂离子电池容量退化数据

该数据集主要运用于本书的第 4.4 节, 数据来源于 (Severson 等, 2019). 该数据集包含 124 个商业锂铁磷 (LFP) /石墨电池 (A123 Systems, APR18650M1A, 1.1 Ah 名义容量) 在快速充电条件下的循环测试数据. 这些电池样本在一个温

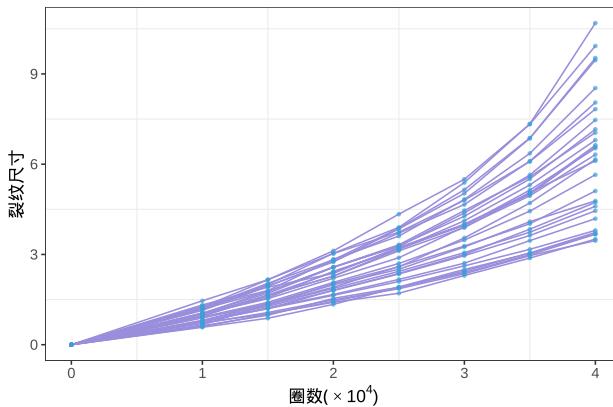


图 1.7: 铝合金裂纹退化数据.

度控制的环境室 (30°C) 中进行测试, 在放电条件保持一致的情况下采用 72 种不同的快速充电策略. 研究中定义电池寿命终止为电池容量降至其名义容量的 80%. 退化路径如图 1.8 所示. 分析该数据集有助于评估和预测电池的循环寿命.

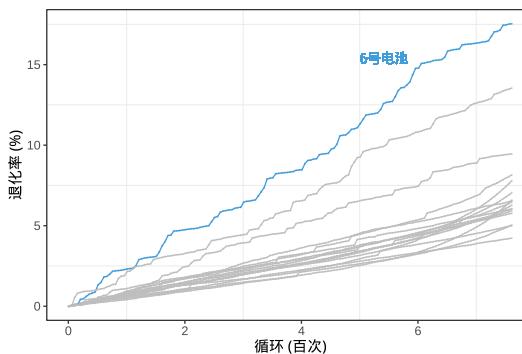


图 1.8: 17 个锂离子电池的容量退化率.

#### 1.4.8 应力关系退化数据

该数据集主要运用于本书的第 5.1 节, 原始数据来源于 Yang (2008) 并在 Ye, Chen L-P, 等 (2014) 中表 IV 中列出. 应力松弛是指在恒定应变下随时间发生的组件应力损失. 该数据集展示了电气连接器在不同温度条件下的应力松

弛行为,记录了在 65°C、85°C 和 100°C 三个不同温度下,18 个电气连接器样本的应力松弛测试结果. 这些样本被随机分为三组,每组在指定的温度下进行测试,以模拟连接器在实际使用过程中可能遇到的不同环境条件. 当应力松弛超过预定义的水平(例如 30%)时,认为连接器已经失效. 退化路径如图 1.9 所示. 这些信息对于连接器的设计改进和可靠性评估至关重要,有助于制造商和工程师更好地理解产品在实际应用中的性能和寿命.

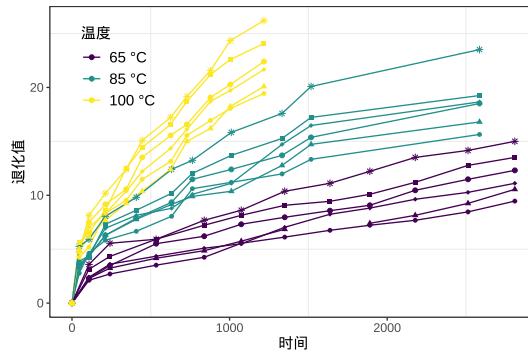


图 1.9: 应力关系退化数据.

#### 1.4.9 两阶段模式的锂离子电池退化数据

该数据集主要运用于本书的第 4.3 节, 数据来源于 (Zhuang 等, 2024). 图 1.10 展示了六个电池容量随周期变化的退化数据. 从图中可以看出, 每个电池的容量退化呈现出两个阶段的特征: 初始阶段退化速率较低, 而后进入退化速率较高的阶段. 分析这类两阶段退化类型的数据有助于评估和预测电池的循环寿命.

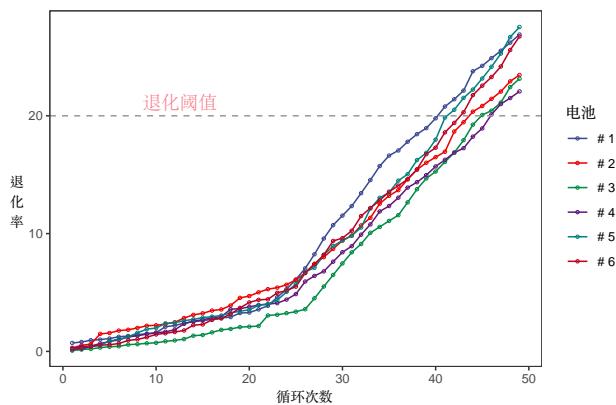


图 1.10: 6 个锂电池的容量退化数据.

## 第二章 基于维纳过程的统计建模

### 2.1 维纳过程

若连续时间随机过程  $\{B(t), t \geq 0\}$  满足以下条件:

- (i)  $B(0) = 0$ ;
- (ii)  $B(t)$  有平稳独立增量;
- (iii) 对任意  $t, s \geq 0$ ,  $B(t+s) - B(s)$  服从均值为 0, 方差为  $t$  的正态分布;
- (iv)  $B(t)$  关于  $t$  是连续函数,

则称  $\{B(t), t \geq 0\}$  服从标准布朗运动过程或标准维纳过程.

根据上述定义, 带时间漂移的维纳过程可表示为

$$X(t) = vt + \sigma B(t), \quad (2.1)$$

其中参数  $v$  和  $\sigma$  分别为漂移参数和扩散参数,  $\{B(t), t \geq 0\}$  为标准布朗运动过程. 根据定义, 随机过程  $\{X(t), t \geq 0\}$  具有以下性质: (I) 在任意时间点  $t, s > 0$ , 增量  $\Delta X = X(t+s) - X(s)$  服从正态分布  $N(vt, \sigma^2 t)$ ; (II) 对任意两个不相交时间区间  $[t_1, t_2]$  和  $[t_3, t_4]$  ( $0 \leq t_1 < t_2 \leq t_3 < t_4$ ), 增量  $X(t_2) - X(t_1)$  与  $X(t_4) - X(t_3)$  相互独立.

上述模型主要适用于退化路径呈线性变化的情形. 然而, 对于某些产品, 性能退化可能随时间呈现非线性趋势. 为描述这一特性, 可采用以下形式的非线性维纳过程 (Whitmore, 1995)

$$Y(t) = v\Lambda(t) + \sigma B(\Lambda(t)), \quad (2.2)$$

其中  $\Lambda(t)$  为时间尺度函数. 通过选择合适的函数形式, 如  $\Lambda(t) = t^\alpha$ , 模型能够灵活地描述线性、凹形或凸形的退化路径, 从而更准确地拟合不同类型的退化数据. 这一改进的模型在保留经典维纳过程优良数学性质的同时, 显著增强了其在非线性退化问题中的适用性, 为复杂退化数据的建模与分析提供了更强有力的工具.

退化建模的主要目的是为了评估产品的可靠性与寿命. 与传统产品失效不同的是, 基于性能退化的失效被定义为退化值首次超过预设失效阈值  $\omega$  的时间, 该时间即为产品的寿命. 若产品性能退化过程由随机过程  $\{X(t), t \geq 0\}$  来描述, 则产品寿命  $T$  定义为

$$T = \inf\{t \mid X(t) \geq \omega\},$$

即, 随机过程  $\{X(t), t \geq 0\}$  的首达时. 假设  $\{X(t), t \geq 0\}$  是维纳过程 (2.1), 根据维纳过程的性质可知产品寿命  $T$  服从逆高斯分布  $IG(\omega/v, \omega^2/\sigma^2)$ , 其对应的累积分布函数 (Cumulative distribution function, CDF) 和概率密度函数 (Probability density function, PDF) 分别为

$$F_T(t) = \Phi\left(\frac{vt - \omega}{\sigma\sqrt{t}}\right) + \exp\left(\frac{2v\omega}{\sigma^2}\right)\Phi\left(\frac{-\omega - vt}{\sigma\sqrt{t}}\right), \quad (2.3)$$

$$f_T(t) = \frac{\omega}{\sqrt{2\pi\sigma^2 t^3}} \exp\left[-\frac{(\omega - vt)^2}{2\sigma^2 t}\right]. \quad (2.4)$$

由逆高斯分布的性质, 产品寿命  $T$  的期望和方差分别为

$$\mathbb{E}(T) = \frac{\omega}{v}, \quad \text{Var}(T) = \frac{\omega\sigma^2}{v^3}. \quad (2.5)$$

同理, 若产品性能退化由非线性维纳过程(2.2)刻画, 则产品寿命  $T$  服从时间尺度变换的逆高斯分布, 即,  $\Lambda(T) \sim IG(\omega/v, \omega^2/\sigma^2)$ .

上述两类维纳过程模型(2.1)-(2.2)在描述较为简单的退化模式方面表现出色, 已广泛应用于工程实践并取得了良好的效果. 然而, 在面对更加复杂的退化行为时, 这些模型的适用性可能受到一定限制. 为此, 本章进一步拓展模型, 以增强对复杂退化过程的解释能力: 1. 广义维纳过程模型:

针对退化速率和波动率随时间动态变化的情况, 该模型通过引入灵活的时间尺度函数  $\Lambda(t)$ , 使得均值和方差能够随时间进行动态调整, 从而更精确地刻

画复杂的退化路径及产品性能变化趋势. 2. 两阶段维纳过程模型: 针对某些产品在退化过程中存在明显的阶段性变化 (如退化速率在特定时间点发生突变), 该模型将退化过程划分为两个阶段, 并分别采用独立的维纳过程进行建模, 以更准确地反映退化速率变化对寿命分布的影响. 本章将在第 2.2 节和第 2.3 节中详细阐述这两类扩展模型的理论推导及其在寿命预测中的实际应用.

## 2.2 广义维纳过程

维纳模型(2.1)-(2.2)假设所有产品性能退化的波动率  $\sigma$  是相同的. 然而, 实际中退化波动率往往并非固定不变, 且可能与退化速率  $v$  存在显著相关性. 例如, 在对镓砷激光器数据的研究中, Ye, Chen, 等 (2015) 发现退化速率  $v$  与波动率  $\sigma$  的皮尔逊相关系数达到 0.487. 这表明, 退化过程中的波动程度可能随退化速度的加快或减缓而相应增强或减弱. 为更准确地描述这种依赖关系, Ye, Chen, 等 (2015) 提出了一种改进模型, 用  $\sigma v$  替代原模型中的扩散参数  $\sigma$ , 并成功应用于疲劳裂纹数据的拟合分析. 此后, 相关研究进一步拓展了这一思路: Wang 等 (2019) 在时间尺度  $\Lambda(t) = t$  的假设下, 将该改进模型应用于在线剩余使用寿命 (Remaining useful life, RUL) 预测; Zhai, Chen, 等 (2018) 结合加速失效时间原理对模型进行了改进, 假设退化速率  $v$  服从逆高斯分布, 以简化数学推导过程; Yan 等 (2020) 在线性退化路径的假设下, 基于 Zhai, Chen, 等 (2018) 的模型引入测量误差, 并提出 EM 算法来估计模型参数.

尽管上述改进模型在一定程度上考虑了退化速率  $v$  与波动性  $\sigma$  的相关性, 但仍假设退化过程的均值与方差比保持恒定. 然而, 这一假设在实际中可能过于严格, 因为退化路径的动态变化往往会导致均值与方差比随时间而改变. 为更准确地刻画复杂的退化行为, 本节提出了一种广义维纳过程模型, 并在现有研究的基础上作出以下改进: (1) 引入退化速率  $v$  与波动性  $\sigma$  的动态依赖关系, 并允许均值-方差比随时间变化, 从而更准确地刻画复杂退化行为. (2) 通过模型的改进, 推导出产品寿命分布的近似解析表达式, 从而为模型的灵活应用和可靠性评估提供理论基础. 本节内容安排如下: 第 2.2.1 节提出广义维纳退化模型, 并给出产品寿命分布的近似解析表达式; 第 2.2.2 节介绍该模型参数的统计推断方法, 包括 EM 算法的构造、初值选择方法以及时间尺度函数的确定; 第 2.2.3 节通过模拟研究验证算法性能, 第 2.2.4 节展示两个实际退化数据的分析案例, 说明该模型在复杂退化场景中的实际应用效果.

### 2.2.1 模型构建

令  $Y(t)$  表示产品在时间  $t$  的退化值. 记如下带漂移参数  $v$  和扩散函数  $\sigma(v)$  的维纳过程为  $M_0$ :

$$M_0 : Y(t) = v\Lambda(t) + \sigma(v)B(\tau(t)),$$

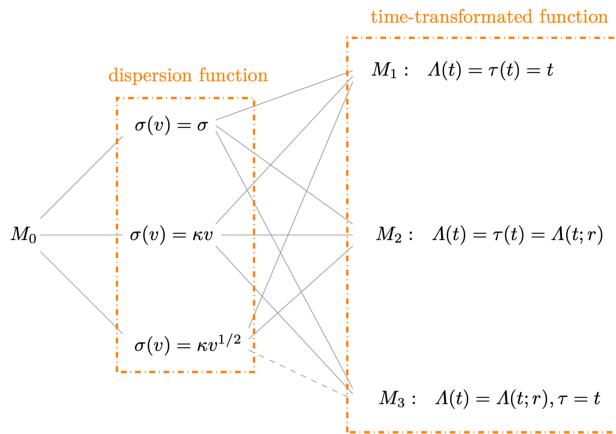
其中  $\Lambda(t) = \Lambda(t; r)$  和  $\tau(t) = \tau(t; b)$  为两个单调递增的时间变换函数,  $\Lambda(0) = \tau(0) = 0$ ,  $B(\cdot)$  为标准布朗运动. 从物理意义上讲, 漂移参数  $v$  代表退化率或退化斜率, 而扩散参数  $\sigma(v)$  表示退化波动性. 由于  $\tau(t)$  在区间  $[0, \infty]$  上是单调递增的, 可证明  $B(\tau(t))$  保持了标准布朗运动的基本性质. 因此,  $Y(t)$  具有以下性质: (i)  $Y(0) = 0$ ; (ii)  $Y(t)$  有独立增量; (iii) 对任意  $t > s > 0$ , 增量  $\Delta Y(t) = Y(t) - Y(s)$  服从正态分布, 均值为  $v\Delta\Lambda(t)$ , 方差为  $\sigma(v)\Delta\tau(t)$ , 其中  $\Delta\Lambda(t) = \Lambda(t) - \Lambda(s)$ ,  $\Delta\tau(t) = \tau(t) - \tau(s)$ .

与现有的维纳退化模型相比, 所提模型  $M_0$  具有两个主要特点: 1. 扩散参数  $\sigma(v)$  被设定为漂移参数  $v$  的函数, 从而能够刻画退化波动性与退化速率之间的相关性; 2. 根据模型  $M_0$ , 可得  $\mathbb{E}(Y(t)) = v\Lambda(t)$ ,  $\text{Var}(Y(t)) = \sigma^2(v)\tau(t)$ , 则其均值-方差比为  $v\Lambda(t)/(\sigma^2(v)\tau(t))$ . 该比值随时间变化, 表明模型能够有效捕捉复杂退化过程的时序特征. 通过对  $\sigma(v)$  和  $\tau(t)$  选择合适的函数形式, 模型  $M_0$  不仅能够覆盖现有的维纳退化模型, 还可为尚未深入研究的潜在模型提供理论支持与构建框架. 图 2.1 展示了所提模型及其特例, 其中实线灰色线条表示已包含在模型  $M_0$  中的现有维纳退化模型; 虚线灰色线条则代表尚未深入研究但可视为模型  $M_0$  潜在扩展的模型.

在模型  $M_0$  中,  $\sigma(v)$ ,  $\Lambda(t)$  和  $\tau(t)$  不同选择方式可对应文献中不同的维纳过程模型.

(1) **传统维纳退化模型 (TWM)** (当  $\sigma(v) = \sigma$  时):

- 当  $\Lambda(t) = \tau(t) = t$  时, 该模型适用于线性漂移的退化过程 (模型 (2.1)), 可参考 Ye, Shen, 等 (2012) 和 Tsai 等 (2011).
- 当  $\Lambda(t) = \tau(t) = \Lambda(t; r)$  时, 该模型适用于非线性漂移的退化过程 (模型 (2.2)), 相关研究可见 Whitmore 等 (1997) 和 Tsai 等 (2011).
- 当  $\Lambda(t) = \Lambda(t; r)$ ,  $\tau(t) = t$  时, 该模型是非线性 TWM 的简化形式, 旨在降低数学推导的复杂性, 已被 Si 等 (2011) 研究.

图 2.1: 所提模型  $M_0$  以及拓展模型的层次结构图.

(2) **广义维纳退化模型-类型 I (NWM-I)** (当  $\sigma(v) = v\kappa$  时):

- 当  $\Lambda(t) = \tau(t) = \Lambda(t; r)$  时, 记该模型为非线性 NWM-I, 由 Ye, Chen, 等 (2015) 提出.
- 当  $\Lambda(t) = \Lambda(t; r), \tau(t) = t$  时, 记该模型为 NWM-I. Wang 等 (2019) 将其应用于轴承振动信号的在线 RUL 预测. 研究表明, 该模型在预测稳定性和精度方面优于相应的 TWM.

(3) **广义维纳退化模型-类型 II (NWM-II)** (当  $\sigma(v) = v^{1/2}\kappa$  时), 该模型保留了 NWM-I 的所有特性, 同时具备更优越的数学性质:

- 当  $\Lambda(t) = \tau(t) = \Lambda(t; r)$  时, 记该模型为非线性 NWM-II, 最早由 Zhai, Chen, 等 (2018) 结合加速失效时间原理提出并研究.
- 当  $\Lambda(t) = \Lambda(t; r), \tau(t) = t$  时, 记该模型为 NWM-II. 此外, 对于  $\Lambda(t)$  和  $\tau(t)$  的一般函数形式, 这些模型尚未被深入研究.

通过对模型  $M_0$  的讨论可看出, NWM-II 模型不仅同时具备时间变化的均值-方差比, 以及退化率与波动性之间的依赖关系, 还具有更优越的数学性质. 因此, 本节将重点研究 NWM-II 模型. 在实际应用中, 除了共同的退化趋势外, 不同样本之间往往存在个体异质性. 从统计学的角度来看, 这种不确定性通常可以通过引入随机效应来刻画. 为进一步扩展 NWM-II 模型, 本文假设漂移参数  $v$  服从逆高斯分布  $\mathcal{IG}(\alpha, \beta)$ . 为简化记号, 我们仍将新模型记为模型

$M_0$ , 则有

$$M_0 : \begin{cases} Y(t) = v\Lambda(t) + \kappa v^{1/2} B(\tau(t)), \\ v \sim \mathcal{IG}(\alpha, \beta). \end{cases} \quad (2.6)$$

基于模型  $M_0$ , 可得出定理 2.1, 其具体的数学推导过程详见本节附录 2.2.5.

**定理 2.1.** 在模型  $M_0$  下,  $Y(t)$  的边际 PDF 为:

$$f_{Y(t)}(y_t) = \sqrt{\frac{\beta a_1(t)}{\pi^2 \kappa^2 \tau(t) a_2(t)}} e^{a_3(t)} K_{-1} \left( \sqrt{a_1(t) a_2(t)} \right), \quad (2.7)$$

其中

$$a_1(t) = \frac{\Lambda(t)^2}{\tau(t)\kappa^2} + \frac{\beta}{\alpha^2}, \quad a_2(t) = \frac{y_t^2}{\tau(t)\kappa^2} + \beta, \quad a_3(t) = \frac{y_t\Lambda(t)}{\tau(t)\kappa^2} + \frac{\beta}{\alpha},$$

并且

$$K_p(x) = \frac{1}{2} \int_0^\infty \mu^{p-1} \exp \left( -\frac{x}{2} (\mu + \mu^{-1}) \right) d\mu$$

是第二类  $p$  阶修正贝塞尔函数.  $Y(t)$  的无条件均值和方差分别为

$$\begin{aligned} \mathbb{E}(Y(t)) &= \mathbb{E}_v[\mathbb{E}(Y(t) | v)] = \alpha\Lambda(t), \\ \mathbb{Var}(Y(t)) &= \mathbb{E}_v[\mathbb{Var}(Y(t) | v)] + \mathbb{Var}_v[\mathbb{E}(Y(t) | v)] = \alpha k^2 \tau(t) + \frac{\alpha^3}{\beta} \Lambda(t). \end{aligned}$$

可得带随机效应的  $M_0$  模型的均值-方差比为  $\Lambda(t) / [\kappa^2 \tau(t) + \frac{\alpha^2}{\beta} \Lambda(t)]$ , 该比值随时间动态变化.

基于模型  $M_0$ , 产品寿命可定义为  $T = \inf\{t \mid Y(t) \geq \omega\}$ . 寿命  $T$  的 CDF 可表示为

$$\begin{aligned} F_T(t) &= 1 - P(T > t) = 1 - P(Y(t) \leq \omega) \\ &= \sqrt{\frac{\beta a_1(t)}{\pi^2 \kappa^2 \tau(t)}} \int_\omega^\infty \frac{e^{a_3(t)}}{\sqrt{a_2(t)}} K_{-1} \left( \sqrt{a_1(t) a_2(t)} \right) dy_t. \end{aligned} \quad (2.8)$$

由于式 (2.8) 中被积函数过于复杂,  $F_T(t)$  难以获得解析表达式. 为此, 本节提出了一种近似表达式, 具体步骤如下: 首先, 在给定  $v$  的条件下, 基于模型  $M_1$

和  $M_3$  (见图 2.1) 中得到  $T$  的条件 PDF; 然后, 进一步得到在模型  $M_2$  和  $M_0$  下  $T$  的条件 PDF. 相关结果见定理 2.2. 详细的推导过程见本节附录 2.2.5.

**定理 2.2.** 在模型  $M_0$  下, 产品寿命  $T$  的条件条件 PDF 可近似为

$$f_{M_0}(t | v, \omega) \cong g_{M_0}(t | v, \omega) / A, \quad (2.9)$$

其中  $A = \int_0^\infty g_{M_0}(t | v, \omega) dt$  是正则化常数,

$$g_{M_0}(t | v, \omega) = \frac{1}{\sqrt{2\pi\tau(t)}} \left( \frac{S_2(\tau(t))}{\tau(t)} + \frac{vH'(\tau(t))}{\sigma(v)} \right) \exp \left( -\frac{(S_2(\tau(t)))^2}{2\tau(t)} \right) \tau'(t),$$

$$S_2(\tau(t)) = \frac{1}{\sigma(v)} (\omega - vH(\tau(t))), \quad H'(\tau(t)) = \frac{\partial H(\tau(t))}{\partial \tau(t)}, \quad \tau'(t) = \frac{\partial \tau(t)}{\partial t}.$$

在得到  $T$  的条件 PDF 后, 可通过随机效应  $v$  来处理寿命分布中的不确定性. 具体方法是对  $f_{M_0}(t | v, \omega)$  关于  $v$  求期望, 从而推导出  $T$  的边际 PDF. 相关结果见定理 2.3, 详细推导过程参见本节附录 2.2.5.

**定理 2.3.** 在模型  $M_0$  下, 失效时间的近似无条件 PDF 为

$$f_{M_0}(t | \omega) \cong \int_0^\infty f_{M_0}(t | v, \omega) f(v) \Leftrightarrow v = w(t | \omega) / A_1, \quad (2.10)$$

其中  $A_1 = \int_0^\infty w(t | \omega) dt$  是与  $\omega$  相关的正则化常数, 且

$$w(t | \omega) = \frac{\tau'(t)\sqrt{\beta}}{\pi\kappa\sqrt{\tau(t)}} e^{a_3(t)} \left[ \left( H'(\tau(t)) - \frac{H(\tau(t))}{\tau(t)} \right) K_0(\sqrt{a_1(t)a_2(t)}) \right. \\ \left. + \frac{\omega}{\tau(t)} K_{-1}(\sqrt{a_1(t)a_2(t)}) \right],$$

$$a_1(t) = \frac{(H(\tau(t)))^2}{\kappa^2} + \frac{\beta}{\alpha^2}, \quad a_2(t) = \frac{\omega^2}{\kappa^2} + \beta, \quad a_3(t) = \frac{\omega H(\tau(t))}{\kappa^2} + \frac{\beta}{\alpha}.$$

在上述公式中, 复合函数  $H(\tau(t))$  及其导数  $H'(\tau(t))$  的形式依赖于时间变换函数  $\Lambda(t)$  和  $\tau(t)$  的具体组合. 为便于理解, 表 2.1 对不同  $\Lambda(t)$  和  $\tau(t)$  的组合下,  $H(s)$  及  $H'(\tau(t))$  的具体表达式进行了汇总.

表 2.1: 不同  $\Lambda(t)$  和  $\tau(t)$  设置下, 复合函数  $H(s)$  及  $H'(\tau(t))$  的表达式.

场景	$(\Lambda(t), \tau(t))$	$H(s)$	$H'(s)$	$H'(\tau(t))$
I	$(t^r, t^b)$	$s^{r/b}$	$\frac{r}{b} s^{r/b-1}$	$\frac{r}{b} t^{r-b}$
II	$(t^r, \exp(tb))$	$\left(\frac{\log(s)}{b}\right)^r$	$\frac{r}{bs} \left(\frac{\log(s)}{b}\right)^{r-1}$	$\frac{r}{b} \frac{t^{r-1}}{\exp(tb)}$
III	$(\exp(tr), t^b)$	$\exp(rs^{1/b})$	$\frac{rs^{1/b-1}}{r_2} \exp(rs^{1/b})$	$\frac{r}{b} \frac{\exp(rt)}{t^{b-1}}$
IV	$(\exp(tr), \exp(tb))$	$s^{r/b}$	$\frac{r}{b} s^{r/b-1}$	$\frac{r}{b} \exp(tb)^{r/b-1}$

## 2.2.2 统计推断

假设退化试验中共有  $n$  个随机样品, 令  $y_{ij} = Y(t_{ij})$  表示第  $i$  个样品在时间点  $t_{ij}$  处测得的退化量, 且第  $i$  个样品的测量次数为  $m_i$ . 定义退化增量  $\Delta Y = (\Delta Y_1, \dots, \Delta Y_n)', \Delta Y_i = (\Delta y_{i1}, \Delta y_{i2}, \dots, \Delta y_{im_i})'$ , 且  $\Delta y_{ij} = y_{ij} - y_{ij-1}, j = 1, 2, \dots, m_i$ . 所有样品的初始时间点为  $t_{i0} = 0$ , 初始退化值为  $y_{i0} = 0$ . 假设性能退化规律服从模型  $M_0$ , 则有漂移参数  $v_i$  服从逆高斯分布  $IG(\alpha, \beta)$ , 且  $v_i, i = 1, \dots, n$  之间相互独立. 在给定  $v_i$  的条件下, 退化增量  $\Delta y_{ij}$  服从均值  $v_i \Delta \Lambda_{ij}$  和方差  $\sigma(v_i)^2 \Delta \tau_{ij}$  的正态分布, 其中  $\Delta \Lambda_{ij} = \Lambda(t_{ij}) - \Lambda(t_{ij-1})$ ,  $\Delta \tau_{ij} = \tau(t_{ij}) - \tau(t_{ij-1})$ . 基于此,  $\Delta Y_i$  和  $v_i$  的联合 PDF 为

$$\begin{aligned}
 & f_{\Delta Y_i, v_i | \Theta}(\Delta y_i, v_i) \\
 &= \sqrt{\frac{\beta}{2\pi v_i^3}} \exp\left[-\frac{\beta(v_i - \alpha)^2}{2\alpha^2 v_i}\right] \prod_{j=1}^{m_i} \left\{ \frac{1}{\sqrt{2\pi \sigma(v_i)^2 \Delta \tau_{ij}}} \exp\left[-\frac{(\Delta y_{ij} - v_i \Delta \Lambda_{ij})^2}{2\sigma(v_i)^2 \Delta \tau_{ij}}\right] \right\} \\
 &= \sqrt{\frac{\beta}{(2\pi)^{m_i+1} v_i^{m_i+3} \kappa^{2m_i}}} e^{a_{3i}} \prod_{j=1}^{m_i} (\Delta h_{2ij}^{-1/2}) \exp\left[-\frac{1}{2} \left(\frac{a_{2i}}{v_i} + a_{1i} v_i\right)\right],
 \end{aligned}$$

其中  $\Theta = (\alpha, \beta, \kappa^2, r, b)$  为模型的待估参数,  $r$  和  $b$  分别是  $\Lambda(t) = \Lambda(t; r)$  和  $\tau(t) = \tau(t; b)$  中的未知参数;  $a_{1i}, a_{2i}$  和  $a_{3i}$  分别为

$$a_{1i} = \sum_{j=1}^{m_i} \frac{\Delta \Lambda_{ij}^2}{\Delta \tau_{ij} \kappa^2} + \frac{\beta}{\alpha^2}, \quad a_{2i} = \sum_{j=1}^{m_i} \frac{\Delta y_{ij}^2}{\Delta \tau_{ij} \kappa^2} + \beta, \quad a_{3i} = \sum_{j=1}^{m_i} \frac{\Delta y_{ij} \Delta \Lambda_{ij}}{\Delta \tau_{ij} \kappa^2} + \frac{\beta}{\alpha}. \quad (2.11)$$

对联合分布  $f_{\Delta Y_i, v_i | \Theta}(\Delta y_i, v_i)$  求积分消去  $v_i$  后, 即可得到  $\Delta Y_i$  的边际 PDF

$$f_{\Delta Y_i | \Theta}(\Delta y_i) = \sqrt{\frac{\beta a_{2i}^{p_i}}{2^{m_i-1}(\pi)^{m_i+1}\kappa^{2m_i}a_{1i}^{p_i}}} e^{a_{3i}} \prod_{j=1}^{m_i} (\Delta \tau_{ij}^{-1/2}) K_{p_i}(\sqrt{a_{1i}a_{2i}}), \quad (2.12)$$

其中  $p_i = -\frac{m_i+1}{2}$  是第二类修正贝塞尔函数的阶数. 因此, 基于观测数据  $\Delta Y$ , 未知参数  $\Theta$  的极大似然估计 (Maximum Likelihood Estimator, MLE) 可通过最大化对数似然函数

$$\ell(\Theta | \Delta Y) = \sum_{i=1}^n \log(f_{\Delta Y_i | \Theta}(\Delta y_i)) \quad (2.13)$$

获得. 然而, 由于涉及到第二类修正贝塞尔函数  $K_{p_i}(\cdot)$ , 且其对数似然函数的导数包含所有未知参数  $\Theta$ , 直接优化  $\ell(\Theta | \Delta Y)$  并不容易. 为了克服这个难题, 考虑采用期望最大化 (Expectation Maximization, EM) 算法来求解参数估计.

### 2.2.2.1 EM 算法

EM 算法通过迭代执行 E 步和 M 步来估计参数  $\Theta$ . 将漂移参数  $v = (v_1, v_2, \dots, v_n)$  视为隐变量, 则基于完全数据  $Z = (\Delta Y, v)$  的对数似然函数为

$$\begin{aligned} \ell_c(\Theta | \Delta Y, v) &= \sum_{i=1}^n \log(f_{\Delta Y_i, v_i | \Theta}(\Delta y_i, v_i)) \\ &= \frac{1}{2} \sum_{i=1}^n \left\{ \log(\beta) - (m_i + 3) \log(v_i) - \sum_{j=1}^{m_i} \log(\Delta \tau_{ij}) \right. \\ &\quad \left. - m_i \log(\kappa^2) + 2a_{3i} - a_{1i}v_i - \frac{a_{2i}}{v_i} \right\} + C, \end{aligned} \quad (2.14)$$

其中  $C = \log(2\pi)/2(n + \sum_{i=1}^n m_i)$  为常数. 假设  $\Theta^{(s)}$  是在第  $s$  次迭代中的 M 步的最优解. 在第  $(s+1)$  次迭代中, 需要计算以下 Q 函数

$$\begin{aligned} Q(\Theta | \Theta^{(s)}) &= \sum_{i=1}^n \mathbb{E}_{v_i | \Delta Y_i, \Theta^{(s)}} [\log (f_{\Delta Y_i, v_i}(\Delta y_i, v_i) | \Theta)] \\ &= \frac{1}{2} \sum_{i=1}^n \left\{ \log(\beta) - (m_i + 3) \mathbb{E}_{v_i | \Delta Y_i, \Theta^{(s)}} [\log(v_i)] \right. \\ &\quad - \sum_{j=1}^{m_i} \log(\Delta \tau_{ij}) - m_i \log(\kappa^2) + 2a_{3i} - a_{1i} \mathbb{E}_{v_i | \Delta Y_i, \Theta^{(s)}} [v_i] \\ &\quad \left. - a_{2i} \mathbb{E}_{v_i | \Delta Y_i, \Theta^{(s)}} [v_i^{-1}] \right\} + C, \end{aligned}$$

其中数学期望是关于条件密度函数  $p(\mathbf{v} | \Delta Y, \Theta^{(s)})$  来计算. 条件 PDF  $p(\mathbf{v} | \Delta Y, \Theta^{(s)})$  及条件期望值  $\mathbb{E}_{v_i | \Delta Y_i, \Theta^{(s)}} [v_i]$  和  $\mathbb{E}_{v_i | \Delta Y_i, \Theta^{(s)}} [v_i^{-1}]$  的详细推导见本节附录 2.2.5. 在得到式(2.15)中的 Q 函数后, 在 M 步中通过以下更新公式计算最优解:

$$\Theta^{(s+1)} = \arg \max_{\Theta} Q(\Theta | \Theta^{(s)}). \quad (2.15)$$

M 步的详细推导可见本节附录 2.2.5. 当 E 步和 M 步的迭代满足指定的收敛准则后, 即可得到  $\Theta$  的 MLE. EM 算法可通过以下步骤来实现:

### EM 算法

- (1) 设定参数  $\Theta$  的初始值  $\Theta^{(0)}$  (见第 2.2.2.2 节), 并设置容差误差  $\epsilon$ .
- (2) 根据第  $s$  次迭代的参数估计  $\Theta^{(s)}$ , 通过式 (3.51) - (3.52) 计算第  $(s+1)$  次参数估计  $\Theta^{(s+1)}$ .
- (3) 重复步骤 2, 直到满足  $|\Theta^{(s+1)} - \Theta^{(s)}| < \epsilon$ , 其中  $|\cdot|$  表示欧式距离.
- (4) 最终获得  $\Theta$  的 MLE 为  $\hat{\Theta} = \Theta^{(s+1)}$ .

### 2.2.2.2 确定初始值

在 EM 算法的实现中, 初始值的选择至关重要. 它不仅影响算法的收敛速度, 还可能决定算法是否能够找到全局最优解. 为此, 本小节基于 Ye, Wang, 等 (2013) 的方法, 提出了一种基于三步法的参数初始值选取策略, 以提高算法的稳定性与求解精度. 具体步骤如下:

#### 初始值确定

(1) 估计  $\mathbf{v}$  和  $r$ : 考虑到  $v_i \Lambda(t)$  表示单元  $i$  的平均退化路径, 可以将其视为一个非线性回归问题. 通过最小化  $\mathbf{v} \Lambda(t)$  和观测退化值  $\mathbf{Y}(t)$  之间的均方误差 (Mean squared error, MSE), 估计漂移参数  $\mathbf{v} = (v_1, v_2, \dots, v_n)$  和时间尺度参数  $r$ .

$$MSE = \sum_{i=1}^n (\mathbf{Y}_i(t) - \mathbf{v} \Lambda(t))' (\mathbf{Y}_i(t) - \mathbf{v} \Lambda(t)).$$

(2) 估计  $\alpha$  和  $\beta$ : 根据第一步得到的  $\mathbf{v}$  的估计值, 利用其一阶和二阶矩估计分布参数  $\alpha$  和  $\beta$ .

(3) 估计  $\kappa^2$  和  $b$ : 结合前两步得到的估计值, 进一步通过最大化以下轮廓对数似然函数, 初步确定扩散参数  $\kappa^2$  和时间变换参数  $b$  的初值:

$$\ell(\kappa^2, b) = \frac{1}{2} \sum_{i=1}^n \left\{ \sum_{j=1}^{m_i} \log(\Delta \tau_{ij}) - m_i \log(\kappa^2) + 2a_{3i} - a_{1i}v_i - \frac{a_{2i}}{v_i} \right\}.$$

### 2.2.2.3 确定 $\Lambda(t), \tau(t)$ 形式

在模型构建中, 时间变换函数  $\Lambda(t)$  和  $\tau(t)$  的选择至关重要. 它们直接决定了模型对退化路径及其波动特性的刻画能力, 也对模型的拟合效果和预测精度产生重要影响. 针对具有递增退化特性的产品 (递减特性可通过适当变换转化为递增形式), 退化路径通常呈现三种典型形态: 凹形、线性和凸形 (见图 2.2). 为适应不同形态的退化路径, 本小节基于退化路径的平均趋势和波动

特性, 提出一种系统化的框架, 指导时间变换函数的选择.

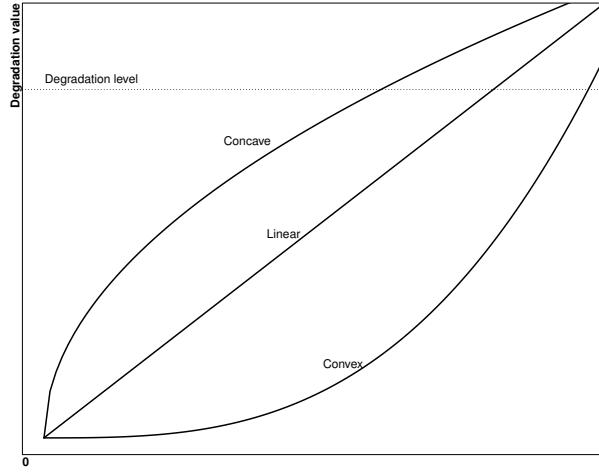


图 2.2: 退化曲线的三种典型形态.

对于模型中的平均退化路径  $v\Lambda(t)$ , 由于其递增且退化速率  $v > 0$ , 因此可以将其视为比例参数, 此时路径的形状完全由  $\Lambda(t)$  中的参数  $r$  决定. 常见的  $\Lambda(t)$  形式包括幂律函数 ( $t^r$ , 当  $r = 1$  时为线性) 和指数律函数 ( $\exp(rt) - 1$ ), 如表 2.2 所示. 从表中可以看出, 幂律函数相比指数律函数更灵活, 能够涵盖各种可能的退化趋势. 然而, 对于许多表现为凸形退化路径的产品, 这两种函数形式在实际应用中难以区分. 为解决上述问题, 本小节提出了一种基于皮尔逊相关系数的方法来选择最优的时间变换函数形式. 给定变量  $\{y_i\}_{i=1}^n$  及其预测值  $\{\hat{y}_i\}_{i=1}^n$ , 相关系数定义为

$$\rho_{xy} = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{y})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}}, \quad (2.16)$$

其中  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ . 当  $\rho_{xy}$  的绝对值越接近 1 时 (Kvalseth, 1985), 表明  $\{\hat{y}_i\}_{i=1}^n$  与  $\{y_i\}_{i=1}^n$  之间的相关性越强. 基于上述讨论, 确定  $\Lambda(t)$  和  $\tau(t)$  形式的具体步骤如下:

表 2.2: 不同时间尺度函数下,  $\Lambda(t)$  的形式.

$\Lambda(t)$	$\Lambda''(t)$	形式
$t^r$	$r(r-1)t^{r-2}$	$r=1$ 时为线性 $r \in (0, 1)$ 时为凹形 $r > 1$ 时为凸形
$\exp(rt)$	$r^2 \exp(rt)$	$r > 0$ 时恒为凸形

### $\Lambda(t)$ 和 $\tau(t)$ 形式确定.

(1) 基于平均退化量  $\mathbb{E}_{Y(t)|v} = v\Lambda(t)$  确定  $\Lambda(t)$ :

- 计算每个检测时间  $t_j$  的经验平均退化量  $\{\bar{y}_j\}_{j=1}^m$ , 其中  $\bar{y}_j = \frac{1}{n} \sum_{i=1}^n y_{ij}$ , 且  $t_j = t_{ij}, i = 1, \dots, n$ , 假设所有样品在相同的时间序列  $\{t_j\}_{j=1}^m$  进行检测.
- 对于每种形式的  $\Lambda(t)$ , 将  $\{\bar{y}_j\}_{j=1}^m$  拟合为  $\{v\Lambda(t_j)\}_{j=1}^m$ , 得到参数估计值  $\hat{v}$  和  $\hat{r}$ , 然后根据式 (2.16) 计算相应的相关系数  $\rho_{xy}$ .

(2) 基于退化方差  $\text{Var}(Y(t) - v\Lambda(t) \mid v) = \sigma(v)^2 \tau(t)$  和  $\hat{v}$  确定  $\tau(t)$ :

- 计算时间序列  $\{t_j\}_{j=1}^m$  的经验方差  $\{\sigma_{y_j}\}_{j=1}^m$ , 其中  $\sigma_{y_j} = \frac{1}{n} \sum_{i=1}^n (y_{ij} - \bar{y}_j)^2, j = 1, 2, \dots, m$ .
- 对于每种形式的  $\tau(t)$ , 将  $\{\sigma_{y_j}\}_{j=1}^m$  拟合为  $\{\sigma(\hat{v}\tau(t_j))\}_{j=1}^m$ , 然后根据式 (2.16) 计算相应的  $\rho_{xy}$ .

(3) 选择具有最佳相关系数的  $\Lambda(t)$  和  $\tau(t)$ .

### 2.2.3 模拟实验

本节通过模拟研究验证模型  $M_0$  及其所提算法的性能. 由于模型  $M_i$  ( $i = 1, 2, 3$ ) 是  $M_0$  的特殊形式, 因此在实验中未单独考虑. 本实验基于表 2.1 中的四种场景, 分别设定样本量  $n = 25, 50$  和  $100$ , 以及测量次数  $m = 15$  和  $30$ . 检测从第 1 小时开始, 每小时记录一次数据. 依据模型  $M_0$ , 结合时间尺度函数  $(\Lambda(t), \tau(t))$  和表 2.3-2.6 中的参数, 在不同  $(n, m)$  设置下生成退化数据, 并对

表 2.3: 场景 I 中不同  $(n, m)$  下估计量的均值与 MSE.

$(n, m)$	$\alpha = 1$		$\beta = 20$		$\kappa^2 = 0.1$		$r = 2$		$b = 0.1$	
	均值	MSE	均值	MSE	均值	MSE	均值	MSE	均值	MSE
(25,15)	0.991	0.484	22.250	3.913	0.098	0.327	1.990	0.040	0.097	0.277
(50, 15)	1.001	0.310	21.956	2.300	0.098	0.023	1.994	0.018	0.099	0.238
(100,15)	0.999	0.200	20.582	1.031	0.099	0.016	2.000	0.007	0.100	0.107
(25,30)	1.001	0.449	21.891	2.713	0.098	0.022	1.995	0.015	0.099	0.202
(50,30)	1.001	0.311	21.835	1.976	0.099	0.015	1.999	0.008	0.100	0.150
(100,30)	1.000	0.200	20.129	0.875	0.100	0.001	2.000	0.002	0.100	0.089

表 2.4: 场景 II 中不同  $(n, m)$  下估计量的均值与 MSE.

$(n, m)$	$\alpha = 1$		$\beta = 20$		$\kappa^2 = 0.1$		$r = 0.2$		$b = 0.1$	
	均值	MSE	均值	MSE	均值	MSE	均值	MSE	均值	MSE
(25,15)	0.980	2.115	22.109	4.651	0.096	1.093	2.074	2.562	0.097	0.314
(50, 15)	1.006	0.647	20.952	1.535	0.097	0.979	2.053	1.165	0.098	0.301
(100,15)	0.998	0.544	20.645	0.386	0.097	0.334	2.050	0.425	0.099	0.255
(25,30)	1.012	1.626	21.330	4.184	0.098	0.787	2.042	0.705	0.098	0.271
(50,30)	0.994	0.549	21.293	1.488	0.099	0.366	2.033	0.214	0.100	0.248
(100,30)	0.999	0.502	20.505	0.188	0.099	0.344	2.017	0.130	0.100	0.221

每种模拟设置重复 1000 次. 模拟数据生成后, 利用所提出的 EM 算法对模型参数  $\Theta$  进行估计, 并计算每种  $(\Lambda(t), \tau(t))$  组合及  $(n, m)$  设置下参数估计值的均值和均方误差(Mean squared error, MSE). 模拟结果列于表 2.3-2.6 中. 结果显示, 在所有实验场景中, 参数估计值与真实值的差异较小. 此外, 随着样本量  $n$  或测量次数  $m$  的增加, 参数估计的 MSE 逐渐减小, 表明所提算法具有较好的收敛性和估计精度.

## 2.2.4 实例分析

### 2.2.4.1 激光设备退化数据

为了验证模型在实际数据中的应用效果, 本节对激光器的工作电流退化数据进行建模. 首先, 为确定时间尺度函数  $\Lambda(t)$  和  $\tau(t)$  的形式, 计算了每个

表 2.5: 场景 III 中不同  $(n, m)$  下估计量的均值与 MSE.

$(n, m)$	$\alpha = 1$		$\beta = 20$		$\kappa^2 = 0.1$		$r = 0.2$		$b = 0.05$	
	均值	MSE	均值	MSE	均值	MSE	均值	MSE	均值	MSE
(25,15)	1.279	0.513	23.579	5.020	0.132	0.149	0.190	0.018	0.052	0.329
(50, 15)	1.251	0.506	21.853	4.380	0.130	0.137	0.191	0.011	0.051	0.302
(100,15)	1.085	0.269	21.499	2.104	0.113	0.079	0.196	0.010	0.050	0.278
(25,30)	1.005	0.455	21.212	4.258	0.097	0.117	0.199	0.013	0.050	0.029
(50,30)	0.999	0.281	21.339	3.874	0.098	0.108	0.200	0.010	0.051	0.292
(100,30)	1.000	0.247	20.686	1.784	0.100	0.039	0.200	0.001	0.050	0.214

表 2.6: 场景 IV 中不同  $(n, m)$  下估计量的均值与 MSE.

$(n, m)$	$\alpha = 1$		$\beta = 20$		$\kappa^2 = 0.1$		$r = 0.1$		$b = 0.001$	
	均值	MSE	均值	MSE	均值	MSE	均值	MSE	均值 $\times 10^{-3}$	MSE
(25,15)	0.994	0.464	22.627	5.603	0.111	0.332	0.099	0.233	0.928	0.423
(50, 15)	0.996	0.294	21.155	3.841	0.109	0.218	0.099	0.205	0.957	0.388
(100,15)	0.997	0.231	20.827	2.245	0.101	0.155	0.099	0.135	0.971	0.245
(25,30)	1.011	0.233	21.477	3.313	0.106	0.222	0.100	0.206	0.963	0.349
(50,30)	1.003	0.207	20.401	2.123	0.106	0.164	0.100	0.121	0.974	0.208
(100,30)	1.001	0.186	20.133	1.582	0.100	0.158	1.000	0.107	0.997	0.202

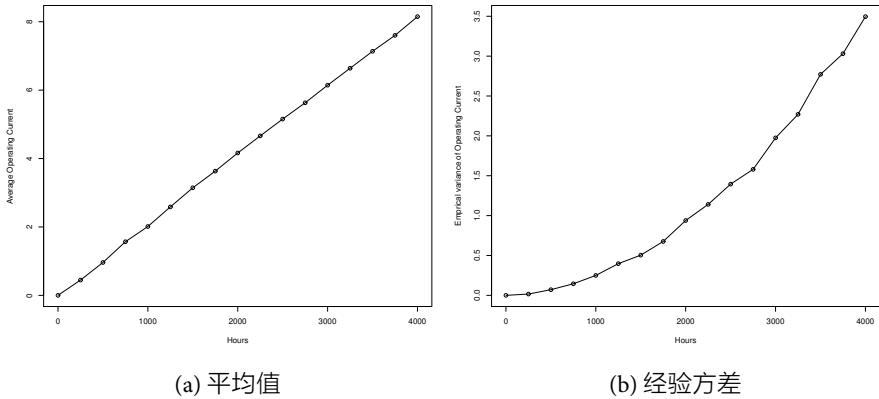


图 2.3: 激光退化数据中工作电流的平均值和经验方差.

表 2.7: 不同拟合函数下的相关系数.

函数形式	相关系数	
	$\Lambda(t)$	$b(t)$
线性	0.9996	0.9416
幂律	0.9986	0.9991
指数	0.8473	0.8676

检测时间点激光器工作电流的平均值和经验方差, 并在图 2.3 中直观展示了其变化趋势. 随后, 分别评估了线性函数、幂函数和指数函数作为时间尺度函数的相关系数, 结果汇总于表 2.7 中. 综合图 2.3 和表 2.7 的分析结果可得: 线性函数最适合作为  $\Lambda(t)$  的形式, 而幂函数更适合作为  $\tau(t)$  的形式. 因此, 针对激光数据的最优模型  $M_0$  可表示为  $vt + \kappa v^{-1/2} B(t^b)$ .

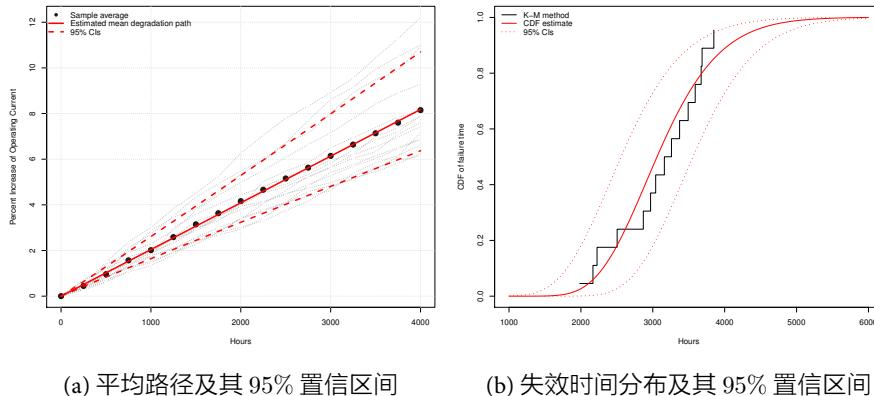
在确定了  $\Lambda(t)$  和  $\tau(t)$  的形式后, 使用模型  $M_d$  ( $d = 0, 1, 2, 3$ ) 对激光退化数据进行拟合. 此外, 还考虑了另一种时间尺度组合  $(\Lambda(t), \tau(t)) = (t^r, t^b)$ , 并将其定义为模型  $M_0^{(1)}$ . 通过赤池信息准则 (Akaike information criterion, AIC) 对各模型进行比较, AIC 值最小的模型被认为是最优模型. 相关结果汇总于表 2.8. 从该表中可见, 模型  $M_0$  的 AIC 值最小, 为 -144.178. 相比之下, 即使忽略退化速率与波动性之间的关系 (如  $\sigma(v) = \sigma$ ), 具有两个随机效应的非线性维纳退化模型 (Wang, 2010) 的 AIC 值为 -130.3, 明显高于  $M_0$ , 进一步凸显了模型

表 2.8: 基于激光退化数据的不同模型 MLE 和 AIC 值.

模型	$\alpha$	$\beta$	$\kappa$	$r$	$b$	AIC 值
$M_0$	2.037	5.091	0.236	1.000	1.001	-144.178
$M_0^{(1)}$	2.031	5.138	0.209	0.998	1.029	-142.635
$M_1$	2.037	5.094	0.236	1.000	1.000	-142.183
$M_2$	1.991	4.982	0.227	1.003	1.003	-142.189
$M_3$	2.036	5.088	0.238	1.001	1.000	-143.191

$M_0$  的优越性.

图 2.4(a) 展示了样本均值、估计的平均退化路径 ( $\hat{\alpha}t^{\hat{r}_1}$ ) 及其 95% 置信区间. 可以看出, 样本均值路径完全落在 95% 置信区间内, 且与估计的平均退化路径高度吻合. 针对模型  $M_0$ , 图 2.4(b) 给出了退化阈值  $\omega = 6$  时的 Kaplan-Meier 经验分布函数、失效时间分布及其 95% 置信区间. 从图中可以看出, 估计的失效时间分布能够准确反映数据特性, 进一步验证了模型  $M_0$  的优越性.

图 2.4: 模型  $M_0$  下激光退化数据估计结果.

#### 2.2.4.2 碳膜电阻器退化数据

这里以碳膜电阻退化数据为例, 使用所提模型进行建模分析. 首先, 计算电阻百分比增加值的平均值和经验方差, 并利用这些结果来确定  $\Lambda(t)$  和  $\tau(t)$  形式, 如图 2.5 所示. 接着, 对不同类型  $\Lambda(t)$  和  $\tau(t)$  下的电阻百分比增加值平

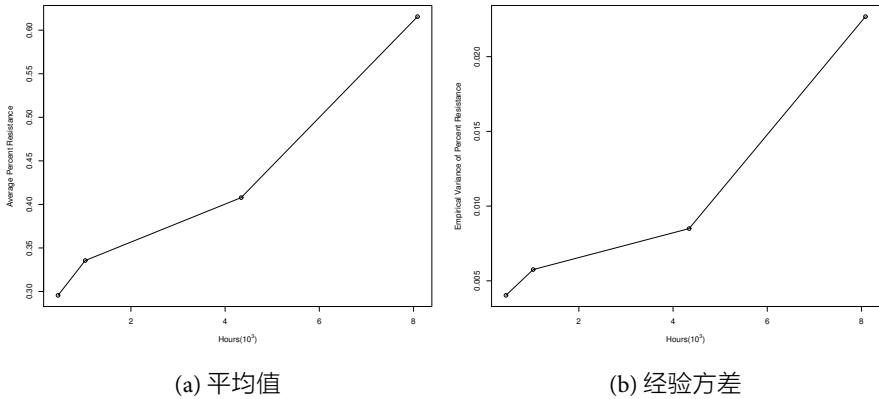


图 2.5: 电阻百分比增加值的平均值和经验方差.

表 2.9: 不同拟合函数下的相关系数.

函数形式	相关系数	
	$\Lambda(t)$	$\tau(t)$
线性函数	0.7598	0.8671
幂函数	0.8097	0.7989
指数函数	0.9688	0.9506

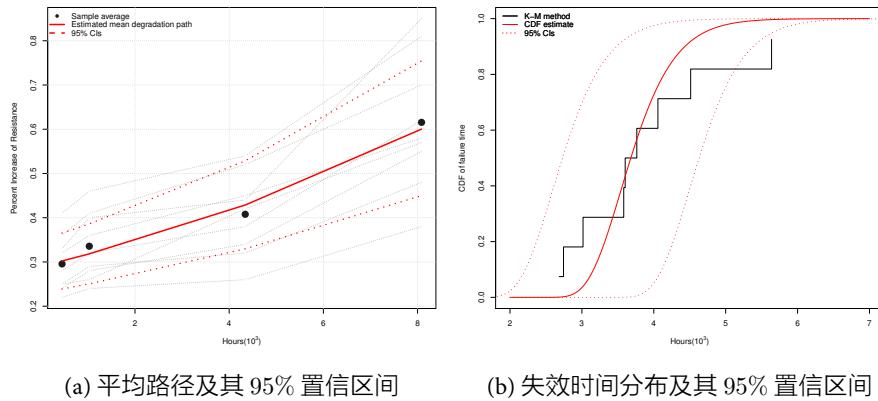
均曲线及其经验方差曲线进行拟合, 拟合结果汇总在表 2.9 中. 通过分析确定  $\Lambda(t)$  和  $\tau(t)$  的最优形式均为指数函数. 因此, 最优模型  $M_0$  为  $\exp(rt) + \kappa v^{-1/2} B(\exp(bt))$ .

接下来, 将碳膜电阻的退化数据分别使用  $M_d$  和  $M_0^{(1)}$  在  $(\Lambda(t), \tau(t)) = (t, \exp(bt))$  的条件下进行拟合. 表 2.10 给出了各模型的参数估计值及对应的 AIC 值. 结果显示, 模型  $M_0$  的 AIC 值最小, 为 -104.506. 相比之下, 对于  $\sigma(v) = \sigma$  的情况, 广义随机效应维纳退化模型 (Li J 等, 2017) 的 AIC 值为 -95.352, 即使考虑测量误差, 该模型的 AIC 值也仅为 -99.979, 仍然高于模型  $M_0$ . 与前例类似, 图 2.4(a) 展示了碳膜电阻退化数据的样本均值、估计的平均退化路径及其 95% 置信区间. 可以看出, 样本均值路径完全落在 95% 置信区间内, 与估计的退化路径高度吻合. 图 2.4(b) 则给出了退化阈值  $\omega = 0.5$  下 Kaplan-Meier 经验分布函数、失效时间分布及其 95% 置信区间. 结果表明, 模型  $M_0$  能够准确反

表 2.10: 基于电阻退化数据的不同模型 MLE 和 AIC 值.

模型	$\alpha$	$\beta$	$\kappa$	$r$	$b$	AIC 值
$M_0$	0.291	13.943	0.117	0.091	0.175	-104.560
$M_0^{(1)}$	0.042	0.595	5.659	1.000	0.001	-51.000
$M_1$	0.085	20.950	0.718	1.000	1.000	4.875
$M_2$	0.299	10.235	0.174	0.089	0.089	-97.456
$M_3$	0.288	9.007	0.066	1.089	1.000	-100.778

映碳膜电阻的退化数据特征, 再次验证了其优越性.

图 2.6: 模型  $M_0$  下碳膜电阻退化数据估计结果.

## 2.2.5 附录

### 定理 2.1 证明

在给定漂移参数  $v$  的条件下,  $Y(t)$  的条件 PDF 为

$$f_{Y(t)|v}(y_t) = \frac{1}{\sqrt{2\pi\sigma(v)^2\tau(t)}} \exp\left[-\frac{(y_t - v\Lambda(t))^2}{2\sigma(v)^2\tau(t)}\right].$$

因此,  $Y(t)$  和  $v$  的联合 PDF 可表示为:

$$f_{Y(t),v}(y_t, v) = \frac{1}{\sqrt{2\pi\sigma(v)^2\tau(t)}} \exp\left[-\frac{(y_t - v\Lambda(t))^2}{2\sigma(v)^2\tau(t)}\right] \cdot \sqrt{\frac{\beta}{2\pi v^3}} \exp\left[-\frac{\beta(v - \alpha)^2}{2\alpha^2 v}\right].$$

通过对  $v$  积分, 可以得到  $Y(t)$  的边际 PDF

$$\begin{aligned} f_{Y(t)}(y_t) &= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma(v)^2\tau(t)}} \exp\left[-\frac{(y_t - v\Lambda(t))^2}{2\sigma(v)^2\tau(t)}\right] \cdot \sqrt{\frac{\beta}{2\pi v^3}} \exp\left[-\frac{\beta(v - \alpha)^2}{2\alpha^2 v}\right] dv \\ &= \frac{\sqrt{\beta}}{2\pi\kappa\sqrt{\tau(t)}} \cdot \int_0^\infty \frac{1}{v^2} \exp\left[-\frac{1}{2v} \left( \frac{y_t^2}{\kappa^2\tau(t)} + \beta \right) - \frac{v}{2} \left( \frac{\Lambda^2(t)}{\kappa^2\tau(t)} + \frac{\beta}{\alpha^2} \right) + \left( \frac{y_t\Lambda(t)}{\kappa^2\tau(t)} + \frac{\beta}{\alpha} \right) \right] dv. \end{aligned}$$

其中

$$a_1(t) = \frac{\Lambda(t)^2}{\tau(t)\kappa^2} + \frac{\beta}{\alpha^2}, \quad a_2(t) = \frac{y_t^2}{\tau(t)\kappa^2} + \beta, \quad a_3(t) = \frac{y_t\Lambda(t)}{\tau(t)\kappa^2} + \frac{\beta}{\alpha},$$

化简得到

$$f_{Y(t)}(y_t) = \frac{\sqrt{\beta}}{2\pi\kappa\sqrt{\tau(t)}} \exp(a_3(t)) \cdot \int_0^\infty \frac{1}{v^2} \exp\left[-\frac{1}{2} \left( \frac{a_2(t)}{v} + va_1(t) \right)\right] dv. \quad (2.17)$$

之后令  $v = \sqrt{\frac{a_2(t)}{a_1(t)}}\mu$ , 式 (2.17) 进一步化简为:

$$\begin{aligned} f_{Y(t)}(y_t) &= \frac{\sqrt{\beta}}{2\pi\kappa\sqrt{\tau(t)}} \exp(a_3(t)) \sqrt{\frac{a_1(t)}{a_2(t)}} \\ &\quad \times \int_0^\infty \frac{1}{\mu^2} \exp\left[-\frac{\sqrt{a_1(t)a_2(t)}}{2} \left( \frac{1}{\mu} + \mu \right)\right] d\mu. \end{aligned}$$

基于阶数为  $p$  的第二类修正贝塞尔函数

$$K_p(x) = \frac{1}{2} \int_0^\infty \mu^{p-1} \exp\left(-\frac{x}{2}(\mu + \mu^{-1})\right) d\mu,$$

可得到

$$f_{Y(t)}(y_t) = \sqrt{\frac{\beta a_1(t)}{\pi^2 \kappa^2 h_2(t) a_2(t)}} e^{a_3(t)} K_{-1} \left( \sqrt{a_1(t) a_2(t)} \right).$$

### 定理 2.2 证明

(1) 对于模型  $M_1$ , 假设已知  $\omega$  和  $v$ , 则失效时间  $T$  服从均值为  $\omega/v$  和形状参数为  $(\omega/\sigma(v))^2$  的 IG 分布, 其条件 PDF 为

$$f_{M_1}(t | v, \omega) = \frac{\omega}{\sqrt{2\sigma(v)^2 t^3}} \exp \left( -\frac{(\omega - vt)^2}{2\sigma(v)^2 t} \right). \quad (2.18)$$

(2) 对于模型  $M_2$ , 结合  $\Lambda(t)$  和式 (2.18), 其条件 PDF 可表示为

$$\begin{aligned} f_{M_2}(t | v, \omega) &= f_{M_1}(\Lambda(t; r) | v) \Lambda'(t) \\ &= \frac{\omega}{\sqrt{2\sigma(v)^2 \Lambda(t)^3}} \exp \left( -\frac{(\omega - v\Lambda(t))^2}{2\sigma(v)^2 \Lambda(t)} \right) \Lambda'(t), \end{aligned}$$

其中  $\Lambda'(t)$  是  $\Lambda(t)$  对  $t$  的一阶导数.

(3) 对于模型  $M_3$ , 假设已知  $\omega$  和  $v$ , 则  $T$  的条件 PDF 为 (Si 等, 2012):

$$f_{M_3}(t | v, \omega) = \frac{1}{\sqrt{2\pi t}} \left( \frac{S(t)}{t} + \frac{v\Lambda'(t)}{\sigma} \right) \exp \left( -\frac{(S(t))^2}{2t} \right), \quad (2.19)$$

其中  $S(t) = [\omega - v\Lambda(t)]/\sigma$  被称为标准布朗运动的时变边界. 接着, 通过在式 (2.19) 中设置  $\sigma = \sigma(v)$ , 可以得到  $M_3$  中  $T$  的条件 PDF 为

$$f_{M_3}(t | v, \omega) = \frac{1}{\sqrt{2\pi t}} \left( \frac{S_1(t)}{t} + \frac{v\Lambda'(t)}{\sigma(v)} \right) \exp \left( -\frac{(S_1(t))^2}{2t} \right), \quad (2.20)$$

其中  $S_1(t) = [\omega - v\Lambda(t)]/\sigma(v)$ . 实际上, 式 (2.18) 是式 (2.20) 在特定情况下的简化形式, 当  $\Lambda(t) = t$  时成立.

(4) 对于模型  $M_0$ , 为推导其失效时间的条件 PDF, 可通过变量替换将  $M_0$  转换为模型  $M_3$  的形式. 设  $s = \tau(t)$ , 则  $t = \tau^{-1}(s)$ , 相应地, 退化过程  $M_0$

可表示为

$$X(s) = vH(s) + \sigma(v)B(s),$$

其中  $H(s) = \Lambda(\tau^{-1}(s))$  是复合时间尺度函数. 根据 Wang, Balakrishnan, 等 (2014) 的假设, 模型  $M_0$  的失效时间条件 PDF 近似为

$$f_{M_0}(t | v, \omega) \cong g_{M_0}(t | v, \omega) / A,$$

其中  $A = \int_0^\infty g_{M_0}(t | v, \omega) dt$  是正则化常数, 且

$$g_{M_0}(t | v, \omega) = \frac{1}{\sqrt{2\pi\tau(t)}} \left( \frac{S_2(\tau(t))}{\tau(t)} + \frac{vH'(\tau(t))}{\sigma(v)} \right) \exp \left( -\frac{(S_2(\tau(t)))^2}{2\tau(t)} \right) \tau'(t),$$

$$S_2(\tau(t)) = \frac{1}{\sigma(v)} (\omega - vH(\tau(t))), \quad H'(\tau(t)) = \frac{\partial H(\tau(t))}{\partial \tau(t)}, \quad \tau'(t) = \frac{\partial \tau(t)}{\partial t}.$$

显然, 当  $\tau(t) = t$  时, 上式退化为模型  $M_3$  的条件 PDF.

### 定理 2.3 证明

根据条件分布的推导,  $f_{M_0}(t | v, \omega)$  可看作  $T$  的条件 PDF, 其中  $f_{M_i}(t | v, \omega), i = 1, 2, 3$  是其特殊形式. 因此, 只需推导出模型  $M_0$  的近似无条件 PDF 即可, 其结果为

$$\begin{aligned} f_{M_0}(t | \omega) &= \frac{1}{A} \int_0^\infty \frac{1}{\sqrt{2\pi\tau(t)}} \left( \frac{S_2(\tau(t))}{h_2(t)} + \frac{vH'(\tau(t))}{\sigma(v)} \right) \\ &\quad \times \exp \left( -\frac{(S_2(\tau(t)))^2}{2\tau(t)} \right) \tau'(t) \sqrt{\frac{\beta}{2\pi v^3}} \exp \left[ -\frac{\beta(v-\alpha)^2}{2\alpha^2 v} \right] dv \\ &= \frac{1}{A} \frac{\tau'(t)\sqrt{\beta}}{\pi\sqrt{\tau(t)}} \int_0^\infty v^{-2/3} \exp \left( -\frac{\beta(v-\alpha)^2}{2\alpha^2 v} \right) \\ &\quad \times \left( \frac{\omega - vH(\tau(t))}{\sqrt{v}\kappa\tau(t)} + \frac{\sqrt{v}H'(\tau(t))}{\kappa} \right) \exp \left( -\frac{(\omega - vH(\tau(t)))^2}{2v\kappa^2} \right) dv \\ &= \frac{1}{A} \frac{\tau'(t)\sqrt{\beta}}{\pi\kappa\sqrt{\tau(t)}} \int_0^\infty \left( v^{-2} \frac{\omega}{\tau(t)} + v^{-1} \left( H'(\tau(t)) - \frac{H(\tau(t))}{\tau(t)} \right) \right) \\ &\quad \times \exp \left( -\frac{v}{2} \left( \frac{(H(\tau(t)))^2}{\kappa^2} + \frac{\beta}{\alpha^2} \right) - \frac{v^{-1}}{2} \left( \frac{\omega^2}{\kappa^2} + \beta \right) + \left( \frac{\omega H(\tau(t))}{\kappa^2} + \frac{\beta}{\alpha} \right) \right) dv \end{aligned}$$

令

$$a_1(t) = \frac{(H(\tau(t)))^2}{\kappa^2} + \frac{\beta}{\alpha}, \quad a_2(t) = \frac{\omega^2}{\kappa^2} + \beta, \quad a_3(t) = \frac{\omega H(\tau(t))}{\kappa^2} + \frac{\beta}{\alpha},$$

则

$$f_{M_0}(t \mid \omega) \cong \int_0^\infty f_{M_0}(t \mid v, \omega) f(v) dv = w(t \mid \omega) / A_1,$$

其中  $A_1 = \int_0^\infty w(t \mid \omega) dt$  是正则化常数, 且

$$\begin{aligned} w(t \mid \omega) = & \frac{\tau'(t)\sqrt{\beta}}{\pi\kappa\sqrt{\tau(t)}} e^{a_3(t)} \left[ \left( H'(\tau(t)) - \frac{H(\tau(t))}{\tau(t)} \right) K_0(\sqrt{a_1(t)a_2(t)}) \right. \\ & \left. + \frac{\omega}{\tau(t)} K_{-1}(\sqrt{a_1(t)a_2(t)}) \right]. \end{aligned}$$

之后, 通过指定  $\Lambda(t)$  和  $\tau(t)$  的特殊情况, 从  $f_{M_0}(t \mid \omega)$  推导出  $f_{M_i}(t \mid \omega)$ , \$i=1,2,3\$. 例如, 当  $\Lambda(t) = t^r$  且  $\tau(t) = t^b$  时,  $f_{M_0}(t \mid \omega)$  的近似表达式为

$$f_{M_0}(t \mid \omega) \cong w^*(t \mid \omega) / A_1^*,$$

其中  $A_1^* = \int_0^\infty w_1^*(t \mid \omega) dt$ ,

$$w^*(t \mid \omega) = \frac{bt^{b-1}\sqrt{\beta}}{\pi\kappa\sqrt{t^b}} e^{a_3^*(t)} \left( \frac{t^{r-b}(r-b)}{r_2} K_0(\sqrt{a_1^*(t)a_2(t)}) + \frac{\omega K_{-1}(\sqrt{a_1^*(t)a_2(t)})}{t^b} \right),$$

$a_1^*(t) = \frac{t^{2r}}{\kappa^2} + \frac{\beta}{\alpha^2}$ ,  $a_3^*(t) = \frac{\omega r t^{r-b}}{\kappa^2 b} + \frac{\beta}{\alpha}$ . 之后, 通过将参数向量  $(r, b)$  分别设为  $(1, 1)$ 、 $(r, r)$  和  $(r, 1)$ , 可以得到  $M_1$ 、 $M_2$  和  $M_3$  的无条件 PDF.

## EM 算法技术细节

### 1) E 步中的条件期望

对于不同样品, 其观测到的退化值和漂移速率相互独立. 因此, 联合分布可以分解为

$$f(\mathbf{v} \mid \Delta \mathbf{Y}, \Theta^{(s)}) = \prod_{i=1}^n f(v_i \mid \Delta \mathbf{Y}_i, \Theta^{(s)}),$$

其中  $f(v_i | \Delta Y_i, \Theta^{(s)})$  可以通过贝叶斯公式得到

$$\begin{aligned} f(v_i | \Delta Y_i, \Theta^{(s)}) &= \frac{f_{v_i, \Delta Y_i | \Theta^{(s)}}(\Delta y_i, v_i)}{f_{\Delta Y_i | \Theta^{(s)}}(\Delta y_i)} \\ &= \frac{\left(a_{1i}^{(s)} / a_{2i}^{(s)}\right)^{p_i/2}}{2K_{p_i} \left(\sqrt{a_{1i}^{(s)} a_{2i}^{(s)}}\right)} v_i^{p_i-1} \exp \left[-\frac{1}{2} \left(a_{1i}^{(s)} v_i + a_{2i}^{(s)} v_i^{-1}\right)\right]. \end{aligned}$$

该分布为广义逆高斯分布 (Generalized Inverse Gaussian, GIG) 分布, 其聚集 (concentration) 参数为  $\sqrt{a_{1i}^{(s)} a_{2i}^{(s)}}$ , 尺度参数为  $\sqrt{a_{1i}^{(s)} / a_{2i}^{(s)}}$ , 阶数参数为  $p_i$ , 记为  $\text{GIG}(a_{1i}^{(s)}, a_{2i}^{(s)}, p_i)$ .  $a_{1i}^{(s)}$  和  $a_{2i}^{(s)}$  可通过将  $\Theta^{(s)}$  代入式 (5.7) 中得到. 根据 GIG 分布的性质,  $\text{GIG}(a_{1i}^{(s)}, a_{2i}^{(s)}, p_i)$  的一阶矩 (期望值) 和一阶逆矩 (倒数的期望值) 为

$$\begin{aligned} E_{v_i | \Delta Y_i, \Theta^{(s)}}[v_i] &= \frac{\sqrt{a_{2i}^{(s)}} K_{p_i+1}(\sqrt{a_{1i}^{(s)} a_{2i}^{(s)}})}{\sqrt{a_{1i}^{(s)}} K_{p_i}(\sqrt{a_{1i}^{(s)} a_{2i}^{(s)}})}, \\ E_{v_i | \Delta Y_i, \Theta^{(s)}}[v_i^{-1}] &= \frac{\sqrt{a_{1i}^{(s)}} K_{p_i-1}(\sqrt{a_{1i}^{(s)} a_{2i}^{(s)}})}{\sqrt{a_{2i}^{(s)}} K_{p_i}(\sqrt{a_{1i}^{(s)} a_{2i}^{(s)}})}. \end{aligned}$$

## 2) M 步中的参数求解

在 M 步中, 通过直接最大化 E 步中计算得到的  $Q$  函数来更新  $\Theta$ . 对  $\{\alpha, \beta, \kappa^2, r_1, r_2\}$  求一阶偏导数, 结果如下:

$$\begin{aligned} \frac{\partial Q(\Theta | \Theta^{(s)})}{\partial \alpha} &= \frac{\beta}{\alpha^3} \sum_{i=1}^n \left\{ E_{v_i | \Delta Y, \Theta^{(s)}}[v_i] - \alpha \right\}, \\ \frac{\partial Q(\Theta | \Theta^{(s)})}{\partial \beta} &= \frac{n}{2\beta} - \frac{1}{2} \sum_{i=1}^n \left\{ E_{v_i | \Delta Y, \Theta^{(s)}}[v_i] \frac{1}{\alpha^2} + E_{v_i | \Delta Y, \Theta^{(s)}}[v_i^{-1}] - \frac{2}{\alpha} \right\}, \\ \frac{\partial Q(\Theta | \Theta^{(s)})}{\partial \kappa^2} &= -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{m_i} \left\{ \frac{1}{\kappa^2} - E_{v_i | \Delta Y, \Theta^{(s)}}[v_i] \frac{\Delta \Lambda_{ij}^2}{\Delta \tau_{ij}} \frac{1}{\kappa^4} \right. \\ &\quad \left. - E_{v_i | \Delta Y, \Theta^{(s)}}[v_i^{-1}] \frac{\Delta y_{ij}^2}{\Delta \tau_{ij}} \frac{1}{\kappa^4} + 2 \frac{\Delta y_{ij} \Delta \Lambda_{ij}}{\Delta \tau_{ij}} \frac{1}{\kappa^4} \right\}, \end{aligned}$$

$$\begin{aligned}\frac{\partial Q(\Theta | \Theta^{(s)})}{\partial r} &= -\sum_{i=1}^n \sum_{j=1}^{m_i} \left\{ \frac{\Delta \Lambda'_{ij}}{\Delta \tau_{ij} \kappa^2} (\Delta \Lambda_{ij} E_{v_i | \Delta Y, \Theta^{(s)}}[v_i] - \Delta y_{ij}) \right\}, \\ \frac{\partial Q(\Theta | \Theta^{(s)})}{\partial b} &= -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{m_i} \left\{ \frac{\Delta \tau'_{ij}}{\Delta \tau_{ij}^2 \kappa^2} (\Delta \tau_{ij} \kappa^2 - \Lambda_{ij}^2 E_{v_i | \Delta Y, \Theta^{(s)}}[v_i] \right. \\ &\quad \left. - \Delta y_{ij}^2 E_{v_i | \Delta Y, \Theta^{(s)}}[v_i^{-1}] + 2 \Delta y_{ij} \Delta \Lambda_{ij}) \right\},\end{aligned}$$

其中  $\Delta \Lambda'_{ij} = \frac{\partial \Delta \Lambda_{ij}}{\partial r}$ ,  $\Delta \tau'_{ij} = \frac{\partial \Delta \tau_{ij}}{\partial b}$ . 然后, 令这些导数为零, 可以得到以下方程:

$$\begin{aligned}\alpha^{(s+1)} &= \frac{1}{n} \sum_{i=1}^n E_{v_i | \Delta Y_i, \Theta^{(s)}}[v_i], \\ \beta^{(s+1)} &= \left( \frac{1}{n} \sum_{i=1}^n E_{v_i | \Delta Y_i, \Theta^{(s)}}[v_i^{-1}] - 1/\alpha^{(s+1)} \right)^{-1}, \\ k^{2(s+1)} &= \frac{1}{\sum_{i=1}^n m_i} \sum_{i=1}^n \sum_{j=1}^{m_i} \left\{ \frac{\Delta y_{ij}^2}{\Delta \tau_{ij}^{(s)}} E_{v_i | \Delta Y_i, \Theta^{(s)}}[v_i^{-1}] \right. \\ &\quad \left. + \frac{\Delta \Lambda_{ij}^{(s)2}}{\Delta \tau_{ij}^{(s)}} E_{v_i | \Delta Y_i, \Theta^{(s)}}[v_i] - 2 \frac{\Delta y_{ij} \Delta \Lambda_{ij}^{(s)}}{\Delta \tau_{ij}^{(s)}} \right\}.\end{aligned}$$

$r^{(s+1)}$  可通过求解以下方程得到:

$$\sum_{i=1}^n \sum_{j=1}^{m_i} \left\{ \frac{\Delta \Lambda'_{ij}^{(s)}}{\Delta \tau_{ij}^{(s)} k^{2(s+1)}} (\Delta h_{ij}^{(s)} E_{v_i | \Delta Y, \Theta^{(s)}}[v_i] - \Delta y_{ij}) \right\} = 0,$$

$b^{(s+1)}$  可通过求解以下方程得到:

$$\begin{aligned}\sum_{i=1}^n \sum_{j=1}^{m_i} \left\{ \frac{\Delta \tau'_{ij}^{(s)}}{\Delta \tau_{ij}^{(s)2} k^{2(s+1)}} (\Delta \tau_{ij}^{(s)} k^{2(s+1)} - \Lambda_{ij}^{(s)2} E_{v_i | \Delta Y, \Theta^{(s)}}[v_i] \right. \\ \left. - \Delta y_{ij}^2 E_{v_i | \Delta Y, \Theta^{(s)}}[v_i^{-1}] + 2 \Delta y_{ij} \Delta \Lambda_{ij}^{(s)}) \right\} = 0,\end{aligned}$$

其中  $\Delta \Lambda_{ij}^{(s)} = \Delta \Lambda_{ij}|_{r=r^{(s)}}$ ,  $\Delta \tau_{ij}^{(s)} = \Delta \tau_{ij}|_{b=b^{(s)}}$ ,  $\Delta \Lambda'_{ij}^{(s)} = \Delta \Lambda'_{ij}|_{r=r^{(s)}}$ ,  $\Delta \tau'_{ij}^{(s)} = \Delta \tau'_{ij}|_{b=b^{(s)}}$ .

## 2.3 二阶段维纳退化过程

某些产品在测试或使用过程中表现出典型的两阶段退化特征, 例如: 退化初期的“磨合阶段”通常退化速率较快, 可能与设备内部杂质或缺陷的逐步消除有关; 随后进入稳定的退化阶段 (Bae 等, 2006), 如等离子显示面板 (Plasma display panels, PDP) 和 OLED 等产品. 为了更准确地描述这类退化特性并估计产品的寿命分布, 选择适合的模型尤为关键. 近年来, 研究者提出了多种模型来刻画两阶段退化情形. 例如, Bae 等 (2006) 提出变点回归模型描述了此类变点问题, 同时考虑了 PDP 等产品间的异质性退化机制; Bae 等 (2008) 所提出的随机系数双指数模型在拟合 PDP 数据时要比 Bae 等 (2006) 的变点回归模型更优. 基于贝叶斯框架的模型进一步扩展了对两阶段退化路径的刻画能力, 例如: Bae 等 (2015) 的贝叶斯变点回归模型和 Yuan 等 (2016) 的多层次贝叶斯双指数模型. 这些模型显著提升了产品寿命分布的估计精度. 然而, 能够同时捕捉退化路径内在随机性与两阶段特性的随机过程模型仍较为少见.

为了解决上述问题, 本节研究一种带测量误差的变点维纳过程模型 (Change-point Wiener process with measurement error, CPWPME), 用于建模两阶段退化数据. 该模型通过设定维纳过程的漂移项为时间的两阶段线性函数来捕捉退化特性, 并引入单元特定的系数和变点, 以反映不同退化路径间的差异. 与传统随机效应模型不同, 所提模型采用多层次贝叶斯方法, 通过单元间信息共享提升估计精度, 为可靠性分析提供了灵活高效的工具. 本节的结构如下: 第 2.3.1 节介绍所提模型并推导失效时间和 RUL 分布; 第 2.3.2 节说明多层次贝叶斯方法及参数推断的具体实现; 第 2.3.3 和 2.3.4 节分别通过模拟研究和案例分析验证模型性能.

### 2.3.1 模型构建

设  $W(t)$  为产品性能的退化特征,  $Y(t) = W(0) - W(t)$  为时间  $t$  时的退化值.  $Y(t)$  用维纳过程建模为

$$Y(t) = m(t) + \sigma B(t), \quad (2.21)$$

其中  $m(t)$  为漂移函数,  $\sigma$  为扩散参数,  $B(t)$  为标准维纳过程. 为了刻画退化速率的变化, 假设漂移函数  $m(t)$  在未知时间点  $\tau$  处发生变化. 具体而言, 第  $i$  个

单元 ( $i = 1, \dots, N$ ) 的漂移函数  $m_i(t; \beta_i^H, \beta_i^L, \tau_i)$  定义为

$$m_i(t; \beta_i^H, \beta_i^L, \tau_i) = \begin{cases} \beta_i^H t, & t \leq \tau_i \\ \beta_i^L(t - \tau_i) + \beta_i^H \tau_i, & t > \tau_i, \end{cases} \quad (2.22)$$

其中  $\beta_i^H$  和  $\beta_i^L$  分别为第一阶段和第二阶段的退化速率,  $\tau_i$  为第  $i$  个单元退化速率的变化时间点 (变点). 假定扩散参数  $\sigma$  在所有单元中相同. 设  $\mathbf{t}_i = (t_{i,1}, \dots, t_{i,n_i})$  为第  $i$  个单元的按顺序排列的测量时间点, 其中  $t_{i,1} = 0$ ,  $\mathbf{y}_i = (y_{i,1}, \dots, y_{i,n_i})$  为在  $\mathbf{t}_i$  处观测到的退化值,  $\mathbf{Y}_i = (Y_{i,1}, \dots, Y_{i,n_i})$  为真实退化值, 总测量次数为  $n_i$ . 由于观测到的退化值可能因测试环境或仪器缺陷而存在测量误差, 在模型中引入误差项

$$Y_{i,j} = Y(t_{i,j}) + \epsilon_{i,j},$$

其中  $\epsilon_{i,j}$  为测量误差, 假设服从独立同分布的  $\mathcal{N}(0, \gamma^2)$ . 定义在时间间隔  $(t_{i,j}, t_{i,j+1})$  上退化增量  $\Delta Y_{i,j} = Y_{i,j+1} - Y_{i,j}$  的观测值为  $\Delta y_{i,j} = y_{i,j+1} - y_{i,j}$ . 根据漂移函数  $m_i(t)$  的定义,  $\Delta Y_{i,j}$  的期望可表示为

$$\Delta m_{i,j} = \begin{cases} \beta_i^H \Delta t_{i,j}, & \tau_i \geq t_{i,j+1}, \\ (\beta_i^H - \beta_i^L) \tau_i + \beta_i^L t_{i,j+1} - \beta_i^H t_{i,j}, & t_{i,j} \leq \tau_i < t_{i,j+1}, \\ \beta_i^L \cdot \Delta t_{i,j}, & \tau_i < t_{i,j}, \end{cases}$$

其中  $\Delta t_{i,j} = t_{i,j+1} - t_{i,j}$ . 此外,  $\Delta Y_{i,g}$  和  $\Delta Y_{i,k}$  之间的协方差为

$$\text{Cov}(\Delta Y_{i,g}, \Delta Y_{i,k}) = \begin{cases} \sigma^2 \Delta t_{i,1} + \gamma^2, & k = g = 1, \\ \sigma^2 \Delta t_{i,k} + 2\gamma^2, & k = g > 1, \\ -\gamma^2, & k = g + 1 \text{ 或 } g = k + 1, \\ 0, & \text{其他情况,} \end{cases}$$

其中  $k, g = 1, \dots, n_i - 1$ . 定义  $\Delta \mathbf{Y}_i = (\Delta Y_{i,1}, \dots, \Delta Y_{i,n_i-1})$  的均值向量为  $\Delta \mathbf{m}_i = (\Delta m_{i,1}, \dots, \Delta m_{i,n_i-1})$ , 协方差矩阵为  $\Sigma_i$ , 其第  $(k, g)$  项为  $\text{cov}(\Delta Y_{i,g}, \Delta Y_{i,k})$ . 则可知  $\Delta \mathbf{Y}_i$  服从多元正态分布, 其均值向量为  $\Delta \mathbf{m}_i$ , 协方

差矩阵为  $\Sigma_i$ , 即  $\mathcal{MVN}(\Delta \mathbf{m}_i, \Sigma_i)$ .  $\Delta \mathbf{Y}_i$  的联合 PDF 为

$$f_{\Delta \mathbf{Y}_i}(\Delta \mathbf{y}_i) = (2\pi)^{-\frac{n_i-1}{2}} |\Sigma_i|^{-\frac{1}{2}} \exp \left[ -\frac{(\Delta \mathbf{y}_i - \Delta \mathbf{m}_i)^\top \Sigma_i^{-1} (\Delta \mathbf{y}_i - \Delta \mathbf{m}_i)}{2} \right], \quad (2.23)$$

其中  $\Delta \mathbf{y}_i = (\Delta y_{i,1}, \dots, \Delta y_{i,n_i-1})$  为第  $i$  个单元观测到的退化增量. 令  $\beta^H = (\beta_1^H, \dots, \beta_N^H)$  和  $\beta^L = (\beta_1^L, \dots, \beta_N^L)$  分别表示第一阶段和第二阶段的退化速率参数向量,  $\tau = (\tau_1, \dots, \tau_N)$  为变点向量, 则 CPWPME 模型的参数集合为  $(\beta^H, \beta^L, \tau, \sigma^2, \gamma^2)$ , 其似然函数为

$$\begin{aligned} \mathcal{L}(\beta^H, \beta^L, \tau, \sigma^2, \gamma^2) &= \prod_{i=1}^N (2\pi)^{-\frac{n_i-1}{2}} |\Sigma_i|^{-\frac{1}{2}} \\ &\quad \exp \left[ -\frac{(\Delta \mathbf{y}_i - \Delta \mathbf{m}_i)^\top \Sigma_i^{-1} (\Delta \mathbf{y}_i - \Delta \mathbf{m}_i)}{2} \right]. \end{aligned}$$

### 2.3.1.1 失效时间和 RUL 分布

**定理 2.4.** 如果性能退化服从模型(2.21), 漂移参数如式 (2.22), 则第  $i$  个测试单元的失效时间  $T_i = \inf\{t \mid Y_i(t) \leq \mathcal{F}_i\}$  的 CDF 为

$$F_{T_i}(t) = \begin{cases} F_{IG} \left( t; \frac{\mathcal{F}_i}{\beta_i^H}, \frac{\mathcal{F}_i^2}{\sigma^2} \right), & t \leq \tau_i, \\ F_{IG} \left( t; \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i}{\beta_i^L}, \frac{(\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i)^2}{\sigma^2} \right), & t > \tau_i, \end{cases} \quad (2.24)$$

其中  $F_{IG}(x; \mu, \lambda)$  表示逆高斯分布  $IG(\mu, \lambda)$  的 CDF.

**定理 2.5.** 基于定理 2.4 的失效时间分布, 第  $i$  个单元的平均失效时间 (Mean time to failure, MTTF) 为

$$\begin{aligned} \mathbb{E}[T_i] &= \frac{\mathcal{F}_i}{\beta_i^H} \left[ 1 - F_{IG} \left( \frac{\mathcal{F}_i^2}{\tau_i \beta_i^H}, \frac{\mathcal{F}_i}{\beta_i^H}, \frac{\mathcal{F}_i^2}{\sigma^2} \right) \right] + \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i}{\beta_i^L} \\ &\quad \times F_{IG} \left( \frac{[\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i]^2}{\tau_i \beta_i^L}, \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i}{\beta_i^L}, \frac{[\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i]^2}{\sigma^2} \right). \end{aligned}$$

定理 2.4 和 2.5 的证明见本节附录 2.3.5. 通过将参数的后验估计代入式 (2.24) 和 (2.5), 可以得到失效时间分布和 MTTF 的估计. 此外, 还可推导每个测试单元的 RUL 分布及平均剩余寿命 (Mean residual life, MRL).

**定理 2.6.** 定义第  $i$  个单元在时间  $t$  的剩余寿命  $R_{it}$  为  $R_{it} = \inf\{X_t > 0; Y_i(t + X_t) \geq \mathcal{F}_i \mid Y_i(t) \leq \mathcal{F}_i\}$ . 则其 PDF 如下: 当  $t \leq \tau_i$  时,

$$f_{R_{it}}(x) = \begin{cases} f_{IG}\left(x; \frac{\mathcal{F}_i - \beta_i^H t}{\beta_i^H}, \frac{(\mathcal{F}_i - \beta_i^H t)^2}{\sigma^2}\right), & x + t \leq \tau_i, \\ f_{IG}\left(x; \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i - \beta_i^L t}{\beta_i^L}, \frac{[\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i - \beta_i^L t]^2}{\sigma^2}\right), & x + t > \tau_i \end{cases} \quad (2.25)$$

当  $t > \tau_i$  时,

$$f_{R_{it}}(x) = f_{IG}\left(x; \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i - \beta_i^L t}{\beta_i^L}, \frac{[\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i - \beta_i^L t]^2}{\sigma^2}\right). \quad (2.26)$$

**定理 2.7.** 第  $i$  个测试单元在时间  $t$  的 MRL 为  $\mu_i(t) = \mathbb{E}_t[R_{it}]$ , 则当  $t \leq \tau_i$  时,

$$\begin{aligned} \mu_i(t) = & \frac{\mathcal{F}_i - \beta_i^H t}{\beta_i^H} \left[ 1 - F_{IG}\left(\frac{(\mathcal{F}_i - \beta_i^H t)^2}{(\tau_i - t)\beta_i^{H2}}, \frac{\mathcal{F}_i - \beta_i^H t}{\beta_i^H}, \frac{(\mathcal{F}_i - \beta_i^H t)^2}{\sigma^2}\right) \right] \\ & + \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i - \beta_i^L t}{\beta_i^L} F_{IG}\left(\frac{[\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i - \beta_i^L t]^2}{(\tau_i - t)\beta_i^{L2}}, \right. \\ & \left. \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i - \beta_i^L t}{\beta_i^L}, \frac{[\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i - \beta_i^L t]^2}{\sigma^2}\right). \end{aligned}$$

当  $t > \tau_i$  时,

$$\mu_i(t) = \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i - \beta_i^L t}{\beta_i^L}. \quad (2.27)$$

定理 2.7 的证明见本节附录 2.3.5. 有时, 生产者可能更关注产品整体失效时间的分布. 基于所提模型, 可以推出如下结论:

**定理 2.8.** 假设  $\mathcal{F}$  是产品的失效阈值, 则产品失效时间  $T = \inf\{t \mid Y(t) \leq \mathcal{F}\}$  的预测分布为

$$p(T \mid \Delta \mathbf{y}) = \int p(T \mid \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \Delta \mathbf{y}) d\boldsymbol{\theta}. \quad (2.28)$$

式(2.28)中涉及到多维积分, 一般不具有解析表达式, 可采用蒙特卡洛方法获得数值结果, 同时可通过数值近似方法计算  $T$  的后验均值.

### 2.3.2 多层贝叶斯推断

本节将采用多层次贝叶斯方法对模型参数进行统计推断. 通过共享单元间的数据信息, 这种方法不仅能提升参数估计的精度, 还可以有效应对数据中的异质性问题, 从而为可靠性分析提供稳健的工具.

设第  $i$  个单元的漂移参数集合为  $\boldsymbol{\eta}_i = (\beta_i^H, \beta_i^L, \tau_i)$ . 通过为  $\boldsymbol{\eta}_i$  指定一个共同的分布, 构建多层次贝叶斯结构, 使漂移参数向量之间能够共享信息. 这种方法已在 Gelman 等 (2014) 中得到详细阐述. 为了尽可能减少先验对后验分布的影响, 本章节将优先选择共轭先验分布, 并为超参数设置较大的方差, 从而体现较弱的先验信息, 并让数据信息主导模型参数的推断过程. 具体的先验分布设置如下:

(1) 针对  $\boldsymbol{\eta}_i$ , 假设  $\boldsymbol{\eta}_i$  服从截断的多元正态分布

$$\boldsymbol{\eta}_i \sim \mathcal{MVN}(\boldsymbol{\omega}, \boldsymbol{\Omega}) \mathcal{I}_{\{\beta_i^H > 0, \beta_i^L > 0, \tau_i > 0\}},$$

其中  $\boldsymbol{\omega}$  是均值向量,  $\boldsymbol{\Omega}$  是协方差矩阵, 其先验PDF为

$$\begin{aligned} \pi(\boldsymbol{\eta}_i \mid \boldsymbol{\omega}, \boldsymbol{\Omega}) &= (2\pi)^{-\frac{3}{2}} |\boldsymbol{\Omega}|^{-\frac{1}{2}} \exp \left[ -\frac{(\boldsymbol{\eta}_i - \boldsymbol{\omega})^\top \boldsymbol{\Omega}^{-1} (\boldsymbol{\eta}_i - \boldsymbol{\omega})}{2} \right] \\ &\quad \times \mathcal{I}_{\{\beta_i^H > 0, \beta_i^L > 0, \tau_i > 0\}}. \end{aligned}$$

若已知退化速率在阶段间存在变化, 可在  $\beta_i^L$  和  $\beta_i^H$  上添加顺序约束条件, 对  $\pi(\boldsymbol{\eta}_i \mid \boldsymbol{\omega}, \boldsymbol{\Omega})$  做相应的修改.

(2) 针对超参数  $\boldsymbol{\omega}$ , 假设  $\boldsymbol{\omega}$  服从多元正态分布  $\mathcal{MVN}(\boldsymbol{\kappa}, \boldsymbol{\Psi})$

$$\pi(\boldsymbol{\omega} \mid \boldsymbol{\kappa}, \boldsymbol{\Psi}) = (2\pi)^{-\frac{3}{2}} |\boldsymbol{\Psi}|^{-\frac{1}{2}} \exp \left[ -\frac{(\boldsymbol{\omega} - \boldsymbol{\kappa})^\top \boldsymbol{\Psi}^{-1} (\boldsymbol{\omega} - \boldsymbol{\kappa})}{2} \right].$$

考虑到对  $\boldsymbol{\omega}$  缺乏有效信息, 设置均值向量为  $\boldsymbol{\kappa} = \mathbf{0}_3$ , 协方差矩阵为  $\boldsymbol{\Psi} = 10^6 \mathbf{I}_3$ , 其中  $\mathbf{0}_3$  是三维零向量,  $\mathbf{I}_3$  是  $3 \times 3$  的单位矩阵.

(3) 针对超参数  $\boldsymbol{\Omega}$ , 由于  $\boldsymbol{\Omega}$  通常为正定矩阵, 则其可分解为  $\boldsymbol{\Omega} = \boldsymbol{\Theta} \boldsymbol{Q} \boldsymbol{\Theta}$ , 其中  $\boldsymbol{\Theta} = \text{diag}(\theta_1, \theta_2, \theta_3)$  且  $\theta_k > 0$ ,  $\boldsymbol{Q}$  为对称正定矩阵. 进一步假设  $\boldsymbol{Q}$  服

从逆 Wishart 分布  $\mathcal{IW}(\rho, \mathbf{S})$

$$\pi(\mathbf{Q} \mid \rho, \mathbf{S}) = \left[ 2^{\frac{3\rho}{2}} \pi^{\frac{3}{2}} \prod_{g=1}^3 \Gamma\left(\frac{\rho+1-g}{2}\right) \right]^{-1} |\mathbf{S}|^{\frac{\rho}{2}} |\mathbf{Q}|^{-\frac{\rho+4}{2}} \exp\left[-\frac{\text{tr}(\mathbf{S}\mathbf{Q}^{-1})}{2}\right],$$

其中  $\rho$  是自由度,  $\mathbf{S}$  是对称正定矩阵. 此外, 为  $\theta_k$  指定伽马分布作为先验

$$\pi(\theta_k \mid a_k, b_k) = \frac{b_k^{a_k}}{\Gamma(a_k)} \theta_k^{a_k-1} e^{-b_k \theta_k}, k = 1, 2, 3.$$

设置  $\rho = 4$ ,  $\mathbf{S} = \mathbf{I}_3$ ,  $a_k = 0.0001$ ,  $b_k = 0.0001$ , 以表示较弱的先验信息 (Ntzoufras, 2011).

(4) 针对  $\sigma^2$  和  $\gamma^2$ , 分别假设两者服从逆伽马分布  $\mathcal{IG}(a_\sigma, b_\sigma)$  和  $\mathcal{IG}(a_\gamma, b_\gamma)$ , 其形式为

$$\pi(\sigma^2 \mid a_\sigma, b_\sigma) = \frac{b_\sigma^{a_\sigma}}{\Gamma(a_\sigma)} \sigma^{2-(a_\sigma+1)} e^{-\frac{b_\sigma}{\sigma^2}},$$

$$\pi(\gamma^2 \mid a_\gamma, b_\gamma) = \frac{b_\gamma^{a_\gamma}}{\Gamma(a_\gamma)} \gamma^{2-(a_\gamma+1)} e^{-\frac{b_\gamma}{\gamma^2}}.$$

若缺乏先验信息, 可以设置  $a_\sigma = b_\sigma = a_\gamma = b_\gamma = 0.0001$ .

通过上述设置, 模型参数的先验分布能够体现单元间异质性, 同时减少主观假设的影响. 定义模型参数集合为  $\boldsymbol{\theta} = (\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_N, \sigma^2, \gamma^2, \boldsymbol{\omega}, \boldsymbol{\Omega})$ . 根据贝叶斯定理,  $\boldsymbol{\theta}$  的联合后验分布可表示为

$$\begin{aligned} \pi(\boldsymbol{\theta} \mid \Delta \mathbf{y}) &\propto \mathcal{L}(\boldsymbol{\beta}^H, \boldsymbol{\beta}^L, \boldsymbol{\tau}, \sigma^2, \gamma^2) \left[ \prod_{i=1}^N \pi(\boldsymbol{\eta}_i \mid \boldsymbol{\omega}, \boldsymbol{\Omega}) \right] \pi(\boldsymbol{\omega} \mid \boldsymbol{\kappa}, \boldsymbol{\Psi}) \pi(\mathbf{Q} \mid \rho, \mathbf{S}) \\ &\times \left[ \prod_{k=1}^3 \pi(\theta_k \mid a_k, b_k) \right] \pi(\sigma^2 \mid a_\sigma, b_\sigma) \pi(\gamma^2 \mid a_\gamma, b_\gamma) \\ &\propto \left\{ \prod_{i=1}^N (2\pi)^{-\frac{n_i-1}{2}} |\boldsymbol{\Sigma}_i|^{-\frac{1}{2}} \exp\left[-\frac{(\Delta \mathbf{y}_i - \Delta \mathbf{m}_i)^\top \boldsymbol{\Sigma}_i^{-1} (\Delta \mathbf{y}_i - \Delta \mathbf{m}_i)}{2}\right] \right. \\ &\quad \times (2\pi)^{-\frac{3}{2}} |\boldsymbol{\Omega}|^{-\frac{1}{2}} \exp\left[-\frac{(\boldsymbol{\eta}_i - \boldsymbol{\omega})^\top \boldsymbol{\Omega}^{-1} (\boldsymbol{\eta}_i - \boldsymbol{\omega})}{2}\right] \mathcal{I}_{\{\beta_i^H > 0, \beta_i^L > 0, \tau_i > 0\}} \Big\} \\ &\quad \times (2\pi)^{-\frac{3}{2}} |\boldsymbol{\Psi}|^{-\frac{1}{2}} \exp\left[-\frac{(\boldsymbol{\omega} - \boldsymbol{\kappa})^\top \boldsymbol{\Psi}^{-1} (\boldsymbol{\omega} - \boldsymbol{\kappa})}{2}\right] \end{aligned}$$

$$\begin{aligned}
& \times \left[ 2^{\frac{3\rho}{2}} \pi^{\frac{3}{2}} \prod_{g=1}^3 \Gamma\left(\frac{\rho+1-g}{2}\right) \right]^{-1} |\mathbf{S}|^{\frac{\rho}{2}} |\mathbf{Q}|^{-\frac{\rho+4}{2}} \exp\left[-\frac{\text{tr}(\mathbf{S}\mathbf{Q}^{-1})}{2}\right] \\
& \times \left[ \prod_{k=1}^3 \frac{b_k^{a_k}}{\Gamma(a_k)} \theta_k^{a_k-1} e^{-b_k\theta_k} \right] \frac{b_\sigma^{a_\sigma}}{\Gamma(a_\sigma)} \sigma^{2-(a_\sigma+1)} e^{-\frac{b_\sigma}{\sigma^2}} \\
& \times \frac{b_\gamma^{a_\gamma}}{\Gamma(a_\gamma)} \gamma^{2-(a_\gamma+1)} e^{-\frac{b_\gamma}{\gamma^2}}. \tag{2.29}
\end{aligned}$$

后验推断通过从上述联合后验分布中抽取后验样本实现. 常用方法是基于完全条件后验分布的 Gibbs 采样算法. 在本研究中, 使用 R 语言中的 *R2OpenBUGS* 包调用 **OpenBUGS** 实现后验样本的生成与推断.

### 2.3.3 模拟实验

为评估多层贝叶斯 CPWPME 模型的估计性能, 本小节通过模拟仿真生成不同场景的退化数据并进行分析. 研究包括以下三个场景: 场景 I: ( $N = 5$ ,  $n_i = 16$ ), 场景 II ( $N = 5$ ,  $n_i = 21$ ), 场景 III ( $N = 10$ ,  $n_i = 21$ ). 在每个场景中, 测量时间从 0 到 18, 以等间隔取值. 退化增量根据多元正态分布 (2.23) 生成. 表 2.11 汇总了模拟中参数的真实值与估计结果, 而场景 III 的部分退化路径展示于图 2.7.

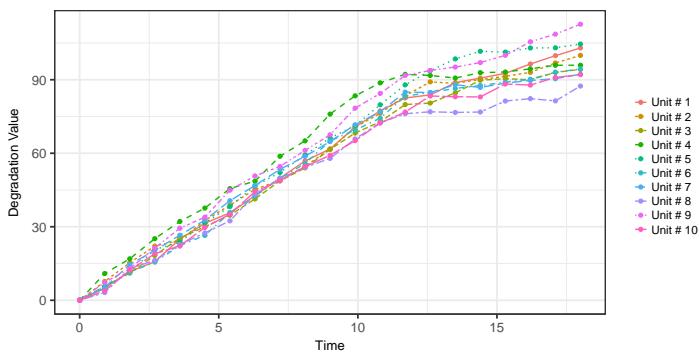


图 2.7: 场景 III 下某一组模拟数据的退化路径.

### 2.3.3.1 后验估计性能评估

多层次贝叶斯方法用于估计上述场景中的模型参数. 后验样本的获取分为两个阶段: 首先进行 15,000 次迭代作为预烧 (burn-in) 样本 (即, 抽样初始阶段需丢弃的样本), 随后进行 10,000 次迭代以生成用于后验推断的样本. 为确保马尔可夫链的收敛性, 通过追踪图、遍历均值图、自相关图以及 Gelman-Rubin 比率对每条链进行监控 (Gelman-Rubin 比率使用 *R* 的 *coda* 包计算). 在每种场景的设置下, 重复上述模拟过程 500 次. 所有参数的后验估计值由后验样本的均值给出.

表 2.11 列出了多层次贝叶斯方法在不同场景下的估计性能指标, 包括偏差、均方根误差 (Root mean squared error, RMSE) 和 95% 频率覆盖概率 (Coverage probability, CP). 结果表明, 即使模型参数较多, 该方法在参数估计上的表现依然出色, 尤其在样本量较小时亦具有良好的估计性能. 从场景 I 和场景 II 的比较来看, 随着测量次数的增加, RMSE 显著减小, 但偏差变化不大; 而从场景 II 和场景 III 的结果比较可发现, 尽管样本数量增加, 偏差未见显著改善, 这可能是由于样本增多的同时未知参数的数量也随之增加. 因此, 模拟仿真结果表明, 多层次贝叶斯方法能够在有限样本条件下保持较高的估计精度.

### 2.3.3.2 失效时间和 RUL 估计性能评估

基于定理 2.4 和模拟数据, 可以推导出退化路径的失效时间分布, 结果如图 2.7 所示, 其中失效阈值设定为  $\mathcal{F} = 200$ . 基于定理 2.5 计算的模拟数据中 10 个样品的 MTTF 分别为 (54.961, 66.311, 65.415, 105.771, 71.484, 73.549, 86.744, 78.291, 53.001, 71.163). 此外, 基于定理 2.6, 计算了第一个样品在不同时间点 RUL 的 PDF, 并与真实 RUL 的 PDF 进行了比较, 结果如图 2.8(b) 所示. 从图中可以看出, 随着第一个样品退化接近生命周期末端, 估计的 PDF 与真实情况非常接近. 十个样品在不同时间点的对应 MRL 结果列于表 2.12. 同时, 根据定理 2.8, 通过蒙特卡洛方法计算得到设备整体失效时间的预测均值为 85.57.

### 2.3.4 实例分析

本节通过 OLED 退化测试数据, 展示多层次贝叶斯方法在 CPWPME 模型中的应用与优势. 以下将从参数估计、模型比较以及失效时间与 RUL 预测三个方面展开分析.

表 2.11: 基于多层次贝叶斯方法, 三种场景下估计性能比较.

场景	$\beta_1^H$	$\beta_2^H$	$\beta_3^H$	$\beta_4^H$	$\beta_5^H$	$\beta_1^L$	$\beta_2^L$	$\beta_3^L$	$\beta_4^L$	$\beta_5^L$	
	真值	6.720	7.082	6.626	7.713	7.147	1.741	2.154	2.233	2.182	1.903
	偏差	0.142	0.017	0.177	-0.259	-0.030	0.223	-0.061	-0.096	-0.122	0.128
	RMSE	0.348	0.312	0.384	0.474	0.326	0.516	0.434	0.427	0.439	0.496
	CP	0.928	0.990	0.910	0.876	0.960	0.960	0.984	0.968	0.960	0.966
I	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$	$\omega[1]$	$\omega[2]$	$\omega[3]$	$\sigma^2$	$\gamma^2$	
	真值	12.828	12.214	11.660	10.787	12.616	7.000	2.000	12.000	2.000	1.000
	偏差	-0.284	-0.085	0.074	0.291	-0.209	0.067	0.057	-0.021	0.048	0.185
	RMSE	0.566	0.422	0.527	0.498	0.508	0.213	0.324	0.257	0.705	0.520
	CP	0.924	0.966	0.972	0.940	0.940	0.994	0.996	1.000	0.974	0.970
	$\beta_1^H$	$\beta_2^H$	$\beta_3^H$	$\beta_4^H$	$\beta_5^H$	$\beta_1^L$	$\beta_2^L$	$\beta_3^L$	$\beta_4^L$	$\beta_5^L$	
	真值	6.720	7.082	6.626	7.713	7.147	1.741	2.154	2.233	2.182	1.903
	偏差	0.149	-0.007	0.193	-0.297	-0.024	0.239	-0.034	-0.120	-0.136	0.119
	RMSE	0.353	0.294	0.378	0.477	0.276	0.521	0.432	0.427	0.438	0.429
	CP	0.934	0.972	0.924	0.884	0.986	0.946	0.974	0.972	0.974	0.986
	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$	$\omega[1]$	$\omega[2]$	$\omega[3]$	$\sigma^2$	$\gamma^2$	
II	真值	12.828	12.214	11.660	10.787	12.616	7.000	2.000	12.000	2.000	1.000
	偏差	-0.292	-0.052	0.069	0.331	-0.155	0.060	0.057	0.002	0.159	0.085
	RMSE	0.521	0.406	0.453	0.559	0.379	0.200	0.308	0.221	0.733	0.402
	CP	0.930	0.974	0.970	0.918	0.970	1.000	1.000	0.998	0.944	0.960
	Stat.	$\beta_1^H$	$\beta_2^H$	$\beta_3^H$	$\beta_4^H$	$\beta_5^H$	$\beta_6^H$	$\beta_7^H$	$\beta_8^H$	$\beta_9^H$	$\beta_{10}^H$
	真值	6.720	7.082	6.626	7.713	7.147	6.633	7.218	7.330	7.257	6.863
	偏差	0.179	-0.022	0.259	-0.362	-0.064	0.249	-0.094	-0.152	-0.113	0.109
	RMSE	0.311	0.228	0.358	0.475	0.228	0.349	0.251	0.294	0.270	0.250
	CP	0.908	0.978	0.894	0.814	0.980	0.890	0.978	0.946	0.966	0.976
	$\beta_1^L$	$\beta_2^L$	$\beta_3^L$	$\beta_4^L$	$\beta_5^L$	$\beta_6^L$	$\beta_7^L$	$\beta_8^L$	$\beta_9^L$	$\beta_{10}^L$	
III	真值	2.478	2.123	1.804	1.300	2.356	1.986	1.995	2.298	2.260	2.188
	偏差	-0.174	0.031	0.195	0.347	-0.108	0.077	0.052	-0.162	-0.125	-0.038
	RMSE	0.416	0.361	0.384	0.555	0.348	0.331	0.307	0.379	0.371	0.328
	CP	0.952	0.984	0.960	0.878	0.990	0.986	0.990	0.972	0.976	0.990
	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$	$\tau_6$	$\tau_7$	$\tau_8$	$\tau_9$	$\tau_{10}$	
	真值	12.503	12.428	12.041	10.910	12.339	11.969	11.915	11.194	11.738	12.229
	偏差	-0.220	-0.199	-0.129	0.272	-0.128	-0.078	-0.012	0.329	0.093	-0.141
	RMSE	0.430	0.369	0.355	0.442	0.355	0.336	0.292	0.486	0.301	0.372
	CP	0.930	0.946	0.970	0.894	0.962	0.968	0.992	0.858	0.980	0.960
	$\omega[1]$	$\omega[2]$	$\omega[3]$	$\sigma^2$	$\gamma^2$						
	真值	7.000	2.000	12.000	2.000	1.000					
	偏差	0.058	0.088	-0.094	0.063	0.041					
	RMSE	0.147	0.224	0.185	0.520	0.293					
	CP	0.992	0.986	0.996	0.952	0.956					

表 2.12: 不同时间点下产品的 MRL (符号'-'表示设备已失效).

单元编号	时间点							
	0	10	20	30	40	50	60	70
#1	54.961	44.961	34.961	24.961	14.961	4.961	-	-
#2	66.311	56.311	46.311	36.311	26.311	16.311	6.311	-
#3	65.415	55.415	45.415	35.415	25.415	15.415	5.415	-
#4	105.771	95.771	85.771	75.771	65.771	55.771	45.771	35.771
#5	71.484	61.484	51.484	41.484	31.484	21.484	11.484	1.484
#6	73.549	63.549	53.549	43.549	33.549	23.549	13.549	3.549
#7	86.744	76.744	66.744	56.744	46.744	36.744	26.744	16.744
#8	78.291	68.291	58.291	48.291	38.291	28.291	18.291	8.291
#9	53.001	43.001	33.001	23.001	13.001	3.001	-	-
#10	71.163	61.163	51.163	41.163	31.163	21.163	11.163	1.163

### 2.3.4.1 参数估计

首先, 将所提模型应用于六个 OLED 样本的退化数据, 并基于第 2.3.2 节定义的先验分布, 采用多层贝叶斯方法进行推断. 具体而言, 首先进行 20,000 次迭代作为马尔可夫链的 burn-in 样本, 以减少初始值对结果的影响; 随后再进行 20,000 次迭代, 生成后验样本用于参数推断. 图 2.9 展示了部分参数的后验样本诊断结果, 表明马尔可夫链已达到收敛. 表 2.13 汇总了所提模型的估计结果, 其中  $\gamma^2$  和  $\sigma^2$  的后验估计值分别为 0.4069 和 0.4005. 漂移参数的总体均值向量与协方差矩阵的后验估计值为

$$\hat{\omega} = (3.76, 9.74, 4.36), \quad \hat{\Omega} = \begin{pmatrix} 0.16350 & 0.00353 & -0.00368 \\ 0.00353 & 0.22370 & -0.00108 \\ -0.00368 & -0.00108 & 0.09092 \end{pmatrix}.$$

### 2.3.4.2 模型比较

为了比较 OLED 退化数据的拟合性能, 考虑三种基准模型作比较分析: (1) 变点维纳过程 (Change-point Wiener process, CPWP) 模型 (Wang, Tang, Bae, He, 2018), 类似于 CPWPME, 但未考虑测量误差; (2) 两阶段线性变点 (Two-phase

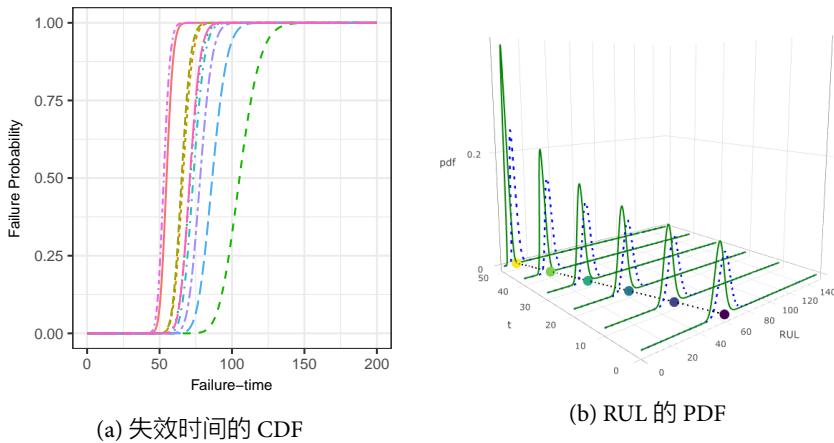


图 2.8: 模拟数据的失效时间和 RUL 估计结果.

表 2.13: 基于 OLED 数据下, 模型参数估计结果.

参数	单元编号					
	#1	#2	#3	#4	#5	#6
$\hat{\beta}^H$	3.67 (3.20, 4.10)	3.65 (3.18, 4.10)	3.70 (3.24, 4.14)	3.81 (3.40, 4.26)	3.78 (3.32, 4.25)	3.93 (3.50, 4.49)
$\hat{\beta}^L$	9.80 (9.22, 10.44)	9.56 (8.82, 10.12)	9.82 (9.27, 10.46)	9.64 (8.99, 10.17)	9.81 (9.25, 10.42)	9.80 (9.23, 10.40)
$\hat{\tau}$	4.48 (4.21, 4.75)	4.48 (4.22, 4.80)	4.36 (4.10, 4.59)	4.43 (4.19, 4.67)	4.17 (3.87, 4.45)	4.27 (4.01, 4.51)

linear regression with change-point, TPLCP) 模型 (Bae 等, 2015), 以分段线性回归描述退化路径; (3) 双指数 (Bi-exponential, BE) 模型 (Yuan 等, 2016), 以指数函数描述退化路径. 具体而言, 第  $i$  个测试样品的 TPLCP 模型定义为

$$y_{i,j} = \begin{cases} \zeta_i t_{i,j} - \kappa_i t_{i,j} + \epsilon_{i,j}, & j = 1, \dots, \gamma_i, \\ \zeta_i t_{i,j} - \kappa_i \varsigma_i + \epsilon_{i,j}, & j = \gamma_i + 1, \dots, n_i, \end{cases} \quad (2.30)$$

其中  $y_{i,j}$  是第  $i$  个样品在时间  $t_{i,j}$  测量的退化值,  $\varsigma_i \in [t_{\gamma_i}, t_{\gamma_i+1})$  是第  $i$  个样品退化的变点位置. 误差项  $\epsilon_{i,j}$  假定为独立同分布的正态分布. 进一步定义参数向量  $\zeta = (\zeta_1, \dots, \zeta_N)^\top$ ,  $\kappa = (\kappa_1, \dots, \kappa_N)^\top$  和  $\varsigma = (\varsigma_1, \dots, \varsigma_N)^\top$ . 接下来, BE

表 2.14: 不同模型下, OLED 退化数据的参数估计值 (括号内表示标准误差).

单元编号	CPWPME			CPWP		
	$\beta^H$	$\beta^L$	$\tau$	$\beta^H$	$\beta^L$	$\tau$
#1	3.67 (0.22)	9.80 (0.30)	4.48 (0.13)	3.80 (0.52)	9.27 (0.64)	4.37 (0.28)
#2	3.65 (0.23)	9.56 (0.33)	4.48 (0.14)	3.80 (0.52)	9.30 (0.63)	4.37 (0.26)
#3	3.70 (0.22)	9.82 (0.29)	4.36 (0.12)	3.82 (0.52)	9.31 (0.63)	4.31 (0.27)
#4	3.81 (0.22)	9.64 (0.30)	4.43 (0.12)	3.85 (0.52)	9.33 (0.64)	4.36 (0.24)
#5	3.78 (0.23)	9.81 (0.29)	4.17 (0.15)	3.85 (0.52)	9.43 (0.64)	4.21 (0.27)
#6	3.93 (0.26)	9.80 (0.29)	4.27 (0.13)	3.90 (0.53)	9.41 (0.64)	4.29 (0.26)

	TPLCP			BE		
	$\zeta$	$\kappa$	$\varsigma$	$\phi$	$\gamma$	$\Delta\gamma$
#1	-8.95 (0.20)	-5.44 (0.21)	4.51 (0.11)	-0.65 (0.02)	-4.74 (0.20)	-4.68 (0.20)
#2	-9.14 (0.17)	-5.64 (0.18)	4.49 (0.10)	-0.65 (0.02)	-4.74 (0.20)	-4.68 (0.20)
#3	-9.09 (0.17)	-5.56 (0.18)	4.26 (0.11)	-0.63 (0.01)	-4.74 (0.20)	-4.68 (0.20)
#4	-9.45 (0.17)	-5.56 (0.18)	4.44 (0.10)	-0.62 (0.01)	-4.74 (0.20)	-4.68 (0.20)
#5	-9.67 (0.19)	-5.94 (0.23)	4.03 (0.11)	-0.61 (0.01)	-4.75 (0.20)	-4.68 (0.20)
#6	-9.82 (0.16)	-5.65 (0.18)	4.22 (0.10)	-0.60 (0.02)	-4.75 (0.20)	-4.68 (0.20)

模型表示为

$$y_{i,j} = \phi_i \exp(-(\gamma_i + \Delta\gamma_i)t_{ij}) + (1 - \phi_i) \exp(-\gamma_i t_{ij}) + \epsilon_{i,j},$$

其中  $i = 1, \dots, N, j = 1, \dots, n_i$ , 误差项  $\epsilon_{i,j}$  服从独立同分布的正态分布. 定义参数向量  $\phi = (\phi_1, \dots, \phi_N)^\top$ ,  $\gamma = (\gamma_1, \dots, \gamma_N)^\top$  和  $\Delta\gamma = (\Delta\gamma_1, \dots, \Delta\gamma_N)^\top$ .

需要指出的是, CPWPME、CPWP 和 TPLCP 模型是在对数时间尺度上定义的, 而 BE 模型则在常规时间尺度上定义. 模型参数的估计结果总结于表 2.14, 不同模型下拟合的退化路径见图 2.11 和图 2.10. 从表 2.14 可以看出, CPWPME 和 CPWP 模型提供了较为接近的参数估计值, 但 CPWPME 模型的参数标准误差显著低于 CPWP 模型, 这反映了 CPWPME 在精度上的优势. 尽管 TPLCP 模型对个体退化路径的拟合效果较好 (图 2.10), 但未考虑路径之间的相关结构, 可能导致过拟合问题. 而 BE 模型在退化终点处的拟合表现欠佳 (图 2.11).

表 2.15: 四个模型下第 7 个 OLED 退化路径的 MSPE.

模型	CPWPME	CPWP	TPLCP	BE
MSPE	360.04	406.04	436.29	678.53

此外, 为评估模型的预测能力, 计算了均方预测误差 (Mean squared prediction error, MSPE), 定义为

$$\text{MSPE} = \sum_{j=1}^{n_i} (y_{i,j} - \hat{y}_{i,j})^2,$$

其中  $\mathbf{y}_i = \{y_{i,1}, \dots, y_{i,n_i}\}$  是第 7 个样品的退化数据, 即  $i = 7$ ,  $\hat{\mathbf{y}}_i = \{\hat{y}_{i,1}, \dots, \hat{y}_{i,n_i}\}$  是预测值, 这些预测值根据 Gelman 等 (2014) 中的多层次预测算法获得. 四个模型计算的均方预测误差列在表 2.15 中. 从表可知, CPWPME 模型的 MSPE 最小, 表明其预测能力最优.

### 2.3.4.3 失效时间和 RUL 预测

根据行业标准, 当光强退化到初始光强的 50% 时, OLED 设备被视为失效 (Bae 等, 2016). 基于所提模型的后验参数估计, 通过式 (2.24) 推导出 OLED 样本的失效时间分布, 结果见图 2.12(a). 进一步地, 基于定理 2.6, 计算出不同时间点下的 RUL 分布, 并将估计结果与真实值比较, 结果如图 2.12(b) 所示. 同时, 每个 OLED 设备在不同时间点的 MRL 由定理 2.7 计算, 结果列于表 2.16. 若以六个 OLED 样本的失效阈值均值作为设备总体的失效阈值, 总体的平均失效时间通过定理 2.8 的方法计算为 6.33.

## 2.3.5 附录

### 定理 2.4 证明

当  $t \leq \tau_i$  时, 退化路径可表示为  $Y(t) = \beta_i^H t + \sigma B(t)$ . 根据 Schrödinger (1915) 的研究, 首达时  $T_i$  服从逆高斯分布, 其参数为

$$T_i \sim \mathcal{IG} \left( \frac{\mathcal{F}_i}{\beta_i^H}, \frac{\mathcal{F}_i}{\sigma^2} \right).$$

表 2.16: 六个 OLED 在不同时间点的平均剩余寿命.

单元编号	时间点						
	0	1	2	3	4	5	6
#1	6.558	5.558	4.558	3.558	2.558	1.558	0.558
#2	6.467	5.467	4.467	3.467	2.467	1.467	0.467
#3	6.357	5.357	4.357	3.357	2.357	1.357	0.357
#4	6.344	5.344	4.344	3.344	2.344	1.344	0.344
#5	6.201	5.201	4.201	3.201	2.201	1.201	0.201
#6	6.117	5.117	4.117	3.117	2.117	1.117	0.117

类似地, 当  $t > \tau_i$  时, 退化路径表示为  $Y(t) = (\beta_i^H - \beta_i^L)\tau_i + \beta_i^L t + \sigma B(t)$ , 此时,  $T_i$  的分布为

$$T_i \sim \mathcal{IG} \left( \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i}{\beta_i^L}, \frac{[\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i]^2}{\sigma^2} \right).$$

### 定理 2.5 证明

失效时间的 MTTF 可以分为两个部分: 当  $t \leq \tau_i$  时和  $t > \tau_i$  时, 表达式为

$$\begin{aligned} \mathbb{E}[T_i] &= \int_0^{\tau_i} t dF_{IG} \left( t; \frac{\mathcal{F}_i}{\beta_i^H}, \frac{\mathcal{F}_i^2}{\sigma^2} \right) \\ &\quad + \int_{\tau_i}^{\infty} t dF_{IG} \left( t; \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i}{\beta_i^L}, \frac{(\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i)^2}{\sigma^2} \right). \end{aligned}$$

记两部分的积分分别为  $I_1$  和  $I_2$ :

$$I_1 = \int_0^{\tau_i} t dF_{IG} \left( t; \frac{\mathcal{F}_i}{\beta_i^H}, \frac{\mathcal{F}_i^2}{\sigma^2} \right),$$

和

$$I_2 = \int_{\tau_i}^{\infty} t dF_{IG} \left( t; \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i}{\beta_i^L}, \frac{(\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i)^2}{\sigma^2} \right).$$

利用 Shuster (1968) 的结果可得:  $F_{IG}(t; \mu, \lambda) = F_{IG}(t/\mu, 1, \lambda/\mu)$ , 将  $I_1$  转换为:

$$I_1 = \int_0^{\tau_i} t dF_{IG} \left( t; \frac{\mathcal{F}_i}{\beta_i^H}, \frac{\mathcal{F}_i^2}{\sigma^2} \right) = \int_0^{\tau_i} t dF_{IG} \left( \frac{t\beta_i^H}{\mathcal{F}_i}; 1, \frac{\mathcal{F}_i\beta_i^H}{\sigma^2} \right).$$

令  $z = t\beta_i^H/\mathcal{F}_i$ , 则  $I_1$  转化为

$$I_1 = \frac{\mathcal{F}_i}{\beta_i^H} \int_0^{\frac{\tau_i\beta_i^H}{\mathcal{F}_i}} \left( \frac{\mathcal{F}_i\beta_i^H}{2\pi\sigma^2 z} \right)^{1/2} \exp \left( -\frac{\mathcal{F}_i\beta_i^H(z-1)^2}{2\sigma^2 z} \right) dz.$$

进一步令  $y = \sqrt{z}$ , 则  $I_1$  表示为

$$I_1 = \frac{\mathcal{F}_i}{\beta_i^H} \int_0^{\left( \frac{\tau_i\beta_i^H}{\mathcal{F}_i} \right)^{1/2}} \left( \frac{2\mathcal{F}_i\beta_i^H}{\pi\sigma^2} \right)^{1/2} \exp \left( -\frac{\mathcal{F}_i\beta_i^H(y^2-1)^2}{2\sigma^2 y^2} \right) dy.$$

根据 Shuster (1968) 的公式

$$F_{IG}(t; 1, \lambda) = \int_{t^{-1/2}}^{\infty} (2\lambda/\pi)^{1/2} \exp \left[ -\frac{\lambda(y^2-1)^2}{y^2} \right] dy.$$

得到

$$I_1 = \frac{\mathcal{F}_i}{\beta_i^H} \left[ 1 - F_{IG} \left( \frac{\mathcal{F}_i^2}{\tau_i\beta_i^H}, \frac{\mathcal{F}_i}{\beta_i^H}, \frac{\mathcal{F}_i^2}{\sigma^2} \right) \right].$$

类似地, 当  $t > \tau_i$  时, 有

$$\begin{aligned} I_2 &= \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i}{\beta_i^L} \\ &\times F_{IG} \left( \frac{[\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i]^2}{\tau_i\beta_i^L}, \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i}{\beta_i^L}, \frac{[\mathcal{F}_i - (\beta_i^H - \beta_i^L)\tau_i]^2}{\sigma^2} \right). \end{aligned}$$

将  $I_1$  和  $I_2$  相加, 即得 MTTF 的最终表达式.

### 定理 2.7 证明

利用与 MTTF 类似的分段积分方法, 剩余寿命的 MRL 表达如下. 当  $t \leq \tau_i$  时,

$$\begin{aligned}
 \mu_i(t) &= \int_0^\infty x dF_{R_{it}(x)} = \int_0^{\tau_i-t} x dF_{IG} \left( x; \frac{\mathcal{F}_i - \beta_i^H t}{\beta_i^H}, \frac{(\mathcal{F}_i - \beta_i^H t)^2}{\sigma^2} \right) \\
 &\quad + \int_{\tau_i-t}^\infty x dF_{IG} \left( x; \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L) \tau_i - \beta_i^L t}{\beta_i^L}, \frac{(\mathcal{F}_i - (\beta_i^H - \beta_i^L) \tau_i - \beta_i^L t)^2}{\sigma^2} \right) \\
 &= \frac{\mathcal{F}_i - \beta_i^H t}{\beta_i^H} \left[ 1 - F_{IG} \left( \frac{(\mathcal{F}_i - \beta_i^H t)^2}{(\tau_i - t) \beta_i^H}, \frac{\mathcal{F}_i - \beta_i^H t}{\beta_i^H}, \frac{(\mathcal{F}_i - \beta_i^H t)^2}{\sigma^2} \right) \right] \\
 &\quad + \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L) \tau_i - \beta_i^L t}{\beta_i^L} F_{IG} \left( \frac{(\mathcal{F}_i - (\beta_i^H - \beta_i^L) \tau_i - \beta_i^L t)^2}{(\tau_i - t) \beta_i^L}, \right. \\
 &\quad \left. \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L) \tau_i - \beta_i^L t}{\beta_i^L}, \frac{(\mathcal{F}_i - (\beta_i^H - \beta_i^L) \tau_i - \beta_i^L t)^2}{\sigma^2} \right).
 \end{aligned}$$

当  $t > \tau_i$  时,

$$\mu_i(t) = \frac{\mathcal{F}_i - (\beta_i^H - \beta_i^L) \tau_i - \beta_i^L t}{\beta_i^L}.$$

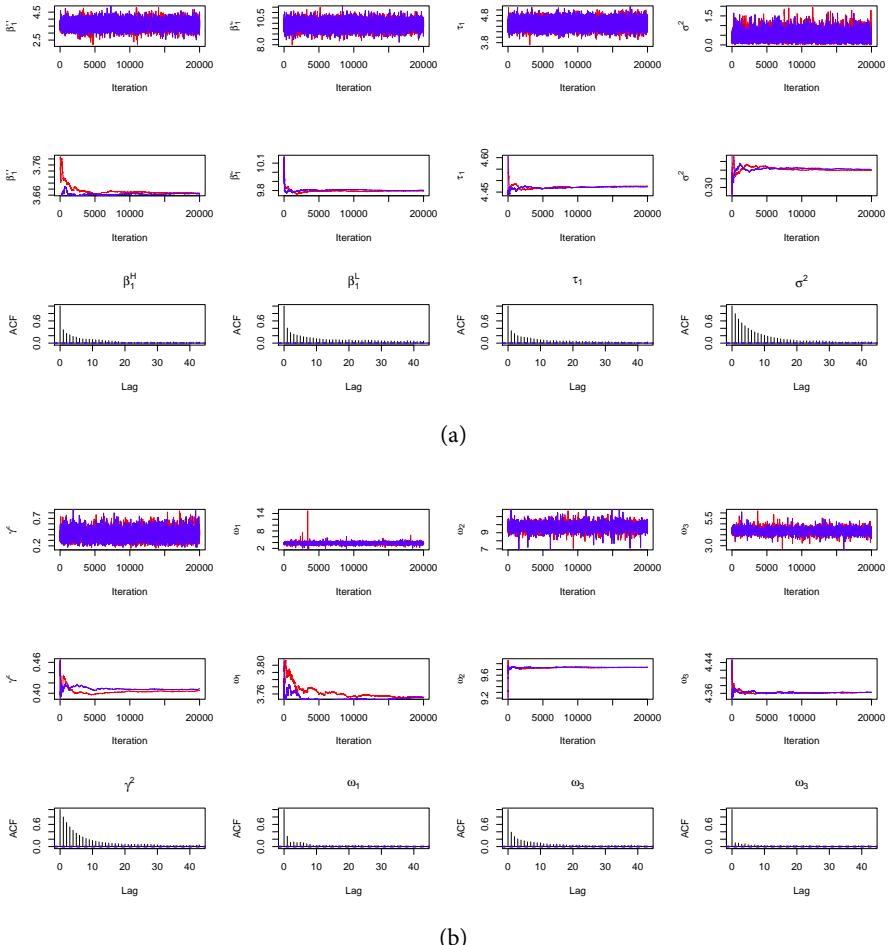


图 2.9: 追踪图、遍历均值图和自相关图.

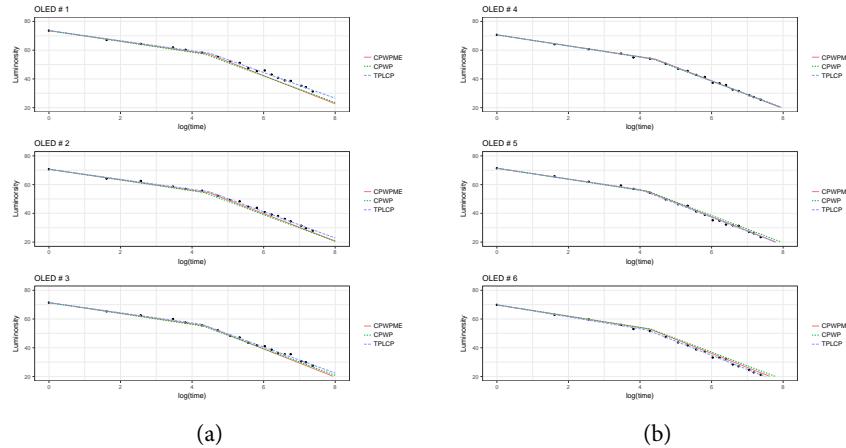


图 2.10: 每个 OLED 退化路径的拟合结果.

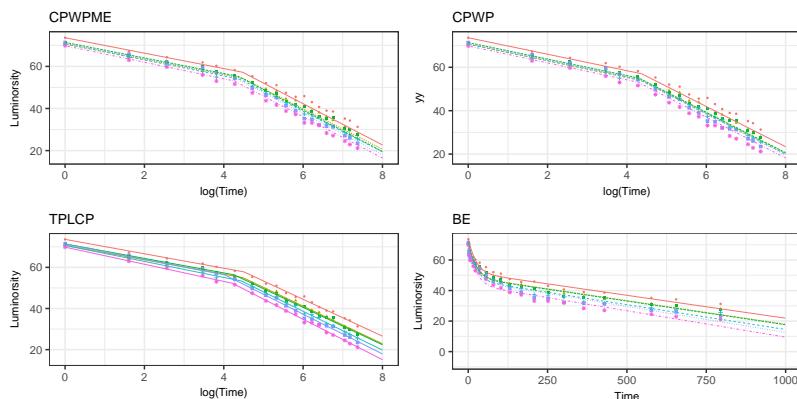


图 2.11: OLED 退化路径的整体拟合结果.

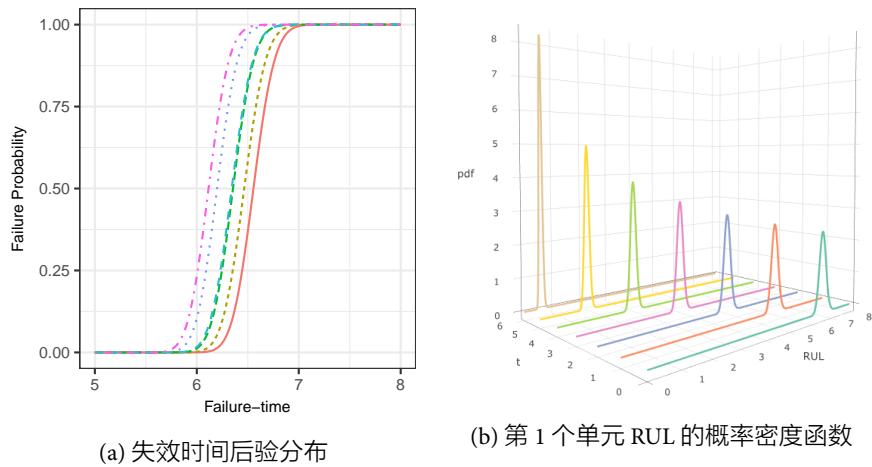


图 2.12: 基于 OLED 退化数据, 失效时间和 RUL 估计结果.

# 第三章 基于伽马过程的统计建模

## 3.1 伽马过程

如果随机过程  $\{Y(t), t \geq 0\}$  满足以下性质:

- (i)  $Y(0) = 0$ ;
- (ii)  $\{Y(t), t \geq 0\}$  具有平稳和独立增量;
- (iii) 对任意的两个时间  $t > s > 0$ , 增量  $Y_{ts} = Y(t) - Y(s)$  服从伽马分布伽马分布  $Ga(\alpha(t-s), \beta)$ , 其 PDF 为

$$f(y | \alpha, \beta) = \frac{\beta^{\alpha(t-s)} y^{\alpha(t-s)-1}}{\Gamma(\alpha(t-s))} \exp\{-\beta y\},$$

其中  $\Gamma(\cdot)$  表示伽马函数,  $\alpha$  和  $\beta$  是未知参数, 且取值大于 0. 称随机过程  $\{Y(t), t \geq 0\}$  为齐次伽马过程, 记作  $Y(t) \sim \mathcal{GP}(\alpha t, \beta)$ .

当产品性能退化服从伽马过程  $\mathcal{GP}(\alpha t, \beta)$  且失效阈值为  $\omega$  时, 定义产品寿命为  $T = \inf\{t | Y(t) \geq \omega\}$ . 此时, 寿命  $T$  的 CDF 可表示为

$$F_T(t | \alpha, \beta) = P(T < t) = P(Y(t) > \omega) = \frac{\Psi(\beta\omega, \alpha t)}{\Gamma(\alpha t)}, \quad (3.1)$$

其中  $\Psi(k, \alpha)$  是上不完全伽马函数, 定义为  $\Psi(k, \alpha) = \int_k^\infty x^{\alpha-1} \exp(-x) dx$ . 本章采用贝叶斯框架进行统计推断, 故将参数  $\alpha$  和  $\beta$  视为随机变量, 并统一使用条件分布的形式. 尽管  $F_T(t | \alpha, \beta)$  可通过常见软件 (如 R 或 MATLAB) 直接计算, 但产品的 MTTF 与寿命分位数缺乏显式解析解, 导致后验推断效率低下, 尤其在实时在线预测场景中计算负担显著. 为此, 可采用 Park 等 (2005) 推荐的

方法, 使用两参数 Birnbaum-Saunders (BS) 分布  $BS(\alpha^*, \beta^*)$  来近似  $F_T(t | \alpha, \beta)$ . BS 分布的 CDF 为

$$\Phi\left(\frac{1}{\alpha^*}\left[\sqrt{\frac{x}{\beta^*}} - \sqrt{\frac{\beta^*}{x}}\right]\right),$$

其中  $\alpha^* = \sqrt{\frac{1}{\beta\omega}}$ ,  $\beta^* = \frac{\beta\omega}{\alpha}$ . 此外, 产品的 MTTF 可近似为

$$\beta^* \left(1 + \frac{(\alpha^*)^2}{2}\right) = \frac{1 + 2\beta\omega}{2\alpha}. \quad (3.2)$$

本章围绕伽马过程展开建模与统计推断. 第 3.2 节构建伽马过程的贝叶斯分析框架, 并提出多种后验采样算法以适应不同场景下的参数推断需求. 针对产品异质性带来的模型复杂度提升问题, 第 3.3 节提出一种重参数化伽马过程及其变分贝叶斯 (Variational Bayesian, VB) 算法. 该算法通过赋予参数明确物理意义, 显著提升推断效率, 为复杂退化模型的求解提供高效方案.

## 3.2 伽马过程的贝叶斯分析

在线 RUL 预测通过退化建模与统计推断实现参数估计. 在伽马过程退化模型领域, Paroissin (2017) 和 Xu, Shen (2018) 提出了递推式线性估计方法, 但其在 RUL 预测及置信区间估计中的应用仍存在局限性. 传统离线方法 (如基于贝叶斯框架和极大似然法的策略 (Ling 等, 2019; Wang, 2008; Wang 等, 2021)) 需依赖过去全部的退化数据, 当新增观测时需重新计算, 对存储空间与计算资源需求较高. 为此, 本节介绍了一种低计算量、高效的伽马过程在线 RUL 预测框架. 内容涵盖了以下几个方面:

1. **贝叶斯推断方法:** 推导伽马过程中模型参数的共轭先验, 并以此为基础提出贝叶斯推断方法. 相较于无信息先验与混合先验形式 (Bousquet 等, 2015; Ling 等, 2019), 该方法通过超参数动态更新简化后验计算, 并针对异质性退化场景提供超参数设定准则.
2. **算法开发:** 开发三种后验分布随机样本生成算法, 克服共轭先验结构复杂性与后验采样效率问题. 通过数值模拟实验证各算法的计算性能与精度优势.

3. 在线 RUL 预测算法: 基于共轭先验的解析递推特性, 构建在线 RUL 预测算法. 其流程 (详见第 3.2.5 节) 仅需两步:

- **后验递推更新:** 利用当前观测值与历史推断结果更新后验分布;
- **后验样本生成:** 调用预设算法输出参数后验样本. 相比传统方法, 该机制避免全量数据存储与重复计算, 显著提升在线推断效率

本节结构安排如下: 第 3.2.1 节介绍伽马过程的共轭先验及其性质; 第 3.2.2 节介绍三种基于共轭先验的后验样本生成算法; 第 3.2.3 节通过模拟评估算法的估计精度和计算效率; 第 3.2.4 节讨论共轭先验在具有异质效应伽马过程中的扩展; 第 3.2.5 节提出一种在线 RUL 预测算法; 第 3.2.6 节展示新方法在实际案例中的应用.

### 3.2.1 共轭先验

假设产品性能的退化路径服从伽马过程  $\mathcal{GP}(\alpha t, \beta)$ . 从总体中随机选择  $n$  个产品进行测试, 测量时间为  $T_1 < T_2 < \dots < T_m$ . 记第  $i$  个产品在时间点  $T_j$  的退化值为  $Y_i(T_j)$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ . 设  $Y_{ij} = Y_i(T_j) - Y_i(T_{j-1})$  且  $t_j = T_j - T_{j-1}$ , 其中  $Y_i(T_0) = 0$  且  $T_0 = 0$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ . 根据伽马过程的性质 (iii), 有  $Y_{ij} \sim \text{Ga}(\alpha t_j, \beta)$ . 令  $y_{ij}$  表示  $Y_{ij}$  的观测值, 并将观测数据表示为  $\mathbf{y} = \{y_{ij}, i = 1, \dots, n, j = 1, \dots, m\}$ . 此时, 基于数据  $\mathbf{y}$ , 似然函数为

$$\begin{aligned} L(\mathbf{y} \mid \alpha, \beta) &= \prod_{i=1}^n \prod_{j=1}^m \frac{\beta^{\alpha t_j}}{\Gamma(\alpha t_j)} y_{ij}^{\alpha t_j - 1} \exp\{-\beta y_{ij}\} \\ &\propto \frac{\beta^{nT_m \alpha} \left[ \prod_{i=1}^n \prod_{j=1}^m y_{ij}^{t_j} \right]^\alpha}{\left[ \prod_{j=1}^m \Gamma(\alpha t_j) \right]^n} \exp\left\{-\beta \sum_{i=1}^n \sum_{j=1}^m y_{ij}\right\} \\ &\propto \frac{\beta^{nm\bar{T}_m \alpha} \exp\{-mn\bar{y}_a \beta\}}{\left[ \prod_{j=1}^m (\Gamma(\alpha t_j))^{1/m} \right]^{mn}} \left[ \prod_{i=1}^n \prod_{j=1}^m y_{ij}^{\frac{t_j}{nm\bar{T}_m}} \right]^{nm\bar{T}_m \alpha}, \end{aligned} \quad (3.3)$$

其中  $\bar{T}_m = \frac{T_m}{m}$ ,  $\bar{y}_a = \frac{1}{mn} \sum_{i=1}^n \sum_{j=1}^m y_{ij}$ .

**定理 3.1.** 基于似然函数 (3.3),  $\alpha$  和  $\beta$  的共轭先验为

$$\pi(\alpha, \beta) = C \frac{(\beta \omega)^{\delta \bar{T}_m \alpha}}{\left[ \prod_{j=1}^m (\Gamma(\alpha t_j))^{1/m} \right]^\delta} \exp\{-\delta \lambda \beta\}, \quad (3.4)$$

其中  $C$  是正则化常数,  $\delta$ 、 $\omega$  和  $\lambda$  是非负超参数, 分别为先验分布的扩散、形状以及尺度参数.

定理 3.1 的证明见本节附录 3.2.7. 共轭先验  $\pi(\alpha, \beta)$  与测量时间相关. 尽管其形式较为复杂, 但可以通过将  $\pi(\alpha, \beta)$  分解如下形式, 使其更为简洁和易于处理.

$$\begin{aligned} \pi(\alpha, \beta) &\propto \frac{(\delta \lambda)^{\delta \bar{T}_m \alpha + 1}}{\Gamma(1 + \delta \bar{T}_m \alpha)} \beta^{\delta \bar{T}_m \alpha} \exp\{-\delta \lambda \beta\} \\ &\times \frac{\left( \frac{\omega}{\delta \lambda} \right)^{\delta \bar{T}_m \alpha} \Gamma(1 + \delta \bar{T}_m \alpha)}{\left[ \prod_{j=1}^m (\Gamma(\alpha t_j))^{1/m} \right]^\delta}. \end{aligned} \quad (3.5)$$

可以发现, 在给定  $\alpha$  时,  $\beta$  的条件先验  $\pi(\beta | \alpha)$  服从伽马分布  $\text{Ga}(\delta \bar{T}_m \alpha + 1, \delta \lambda)$ . 因此,  $\pi(\beta | \alpha)$  的众数、方差和变异系数分别为  $\bar{T}_m \alpha / \lambda$ 、 $(\delta \bar{T}_m \alpha + 1) / (\delta \lambda)^2$  和  $(\delta \bar{T}_m \alpha + 1)^{-1/2}$ , 其中超参数  $\lambda$  是标准的尺度参数, 而超参数  $\delta$  则在给定  $\alpha$  时决定了方差和变异系数. 当  $\delta$  较大时,  $\pi(\beta | \alpha)$  的分布曲将集中于众数附近. 因此, 将  $\delta$  定义为扩散参数.  $\alpha$  的边际先验形式正比于:

$$h(\alpha) = \frac{\left( \frac{\omega}{\delta \lambda} \right)^{\delta \bar{T}_m \alpha} \Gamma(1 + \delta \bar{T}_m \alpha)}{\left[ \prod_{j=1}^m (\Gamma(\alpha t_j))^{1/m} \right]^\delta}.$$

根据 Stirling 公式, 可知:

$$\lim_{u \rightarrow +\infty} \frac{\sqrt{2\pi} u^{u+1/2} \exp\{-u\}}{\Gamma(u+1)} = 1.$$

当  $\alpha \rightarrow +\infty$  时, 用 Stirling 公式对伽马函数做近似处理, 得到以下结果:

$$\begin{aligned}
h(\alpha) &\equiv \frac{\left(\frac{\omega}{\delta\lambda}\right)^{\delta\bar{T}_m\alpha} \sqrt{2\pi}(\delta\bar{T}_m\alpha)^{\delta\bar{T}_m\alpha+1/2} \exp\{-\delta\bar{T}_m\alpha\}}{\left[\prod_{j=1}^m \left(\sqrt{2\pi}(\alpha t_j)^{\alpha t_j-1/2} \exp\{-\alpha t_j\}\right)^{1/m}\right]^\delta} \\
&= O\left(\alpha^{\frac{\delta+1}{2}} \exp\left\{-\alpha\delta\bar{T}_m \log\left(\frac{\lambda}{\omega} \frac{\prod_{j=1}^m t_j^{t_j/T_m}}{\bar{T}_m}\right)\right\}\right), \quad (3.6)
\end{aligned}$$

其中  $h(\alpha) = O(g(\alpha))$  表示  $h(\alpha)$  和  $g(\alpha)$  同阶. 可以证明  $\log\left(\frac{\prod_{j=1}^m t_j^{t_j/T_m}}{\bar{T}_m}\right) \geq 0$  (详见本节附录 3.2.7). 为了保证  $\pi(\alpha)$  是正常的 PDF, 一个充分条件是  $\omega < \lambda$ . 同时根据 (3.6) 可知  $\pi(\alpha)$  的右尾特性类似于伽马分布

$$\text{Ga}\left(\frac{\delta+3}{2}, \delta\bar{T}_m \left[\log\left(\frac{\lambda}{\omega}\right) + \log\left(\frac{\prod_{j=1}^m t_j^{t_j/T_m}}{\bar{T}_m}\right)\right]\right).$$

在  $\pi(\alpha)$  中,  $\omega$  起到尺度参数的作用, 进一步影响  $\pi(\beta | \alpha)$  的形状. 因此,  $\omega$  被称为形状参数. 由于  $\pi(\beta | \alpha)$  的伽马条件先验和  $\pi(\alpha)$  的右尾特性, 这种共轭先验  $\pi(\alpha, \beta)$  被称为近似伽马-伽马 (Approximated-gamma-gamma, AGG) 分布, 记为  $\text{AGG}(\delta, \omega, \lambda)$ .

图 3.1 展示了当  $t_j = 1$  且  $j = 1, \dots, m$  时, 不同  $(\delta, \omega, \lambda)$  值下  $\pi(\alpha, \beta)$  的函数图和等高线图. 以  $\delta = 2, \omega = 0.5, \lambda = 1.5$  为基准, 从图中可以看出, 当  $\delta$  的值增加到 5 且其他两个超参数保持不变时, 众数的位置几乎相同, 但等高线更集中于众数附近. 增大  $\lambda$  的值同样有类似现象, 但众数的位置有所变化.  $\omega$  则会改变  $\pi(\alpha, \beta)$  的形状, 并使其众数的位置发生偏移. 综上所述, 该图展示了不同参数对 AGG 分布的形状和众数位置有显著影响, 为选择合理的超参数提供了依据.

**注 1:** 当采用等间隔测量 (即  $t_j = l$ ) 时,  $\pi(\alpha, \beta)$  的形式简化为

$$\pi(\alpha, \beta) = C \frac{(\beta\omega)^{\delta l \alpha}}{[\Gamma(l\alpha)]^\delta} \exp(-\delta\lambda\beta). \quad (3.7)$$

当  $l = 1$  时,  $\pi(\alpha, \beta)$  简化为伽马分布  $\text{Ga}(\alpha, \beta)$  的共轭先验 (Damsleth, 1975).

**注 2:** 超参数的取值可以基于先验信息量的强弱来选择. 如图 3.1 所示,

不同的超参数组合会显著影响  $(\alpha, \beta)$  的分布特性: 较大的  $\delta$ 、较小的  $\omega$  或较大的  $\lambda$  会导致  $(\alpha, \beta)$  的方差较小, 表示强的先验信息. 在先验知识较少的情况下, 可以选择较小的  $\delta$ 、较大的  $\omega$  或较小的  $\lambda$ . 在实际应用中, 可以通过调整  $\delta$  控制先验信息的强度. 例如, 根据式 (3.20),  $\alpha$  和  $\beta$  的后验分布为  $\text{AGG}(\delta_p, \omega_p, \lambda_p)$ . 特别地, 对于  $\omega$  和  $\lambda$ , 可以使用以下基于观测数据的选择方案:

$$\omega = \prod_{i=1}^n \prod_{j=1}^m y_{ij}^{\frac{t_j}{n^m T_m}}, \quad \lambda = \bar{y}_a. \quad (3.8)$$

这种超参数选择方法属于数据驱动先验的范畴, 其合理性已在贝叶斯统计中得到广泛验证. 例如: 回归系数的先验 (Zellner, 1986), 阈值参数的有效先验 (Hall 等, 2005), 线性退化路径模型的参考先验 (Xu 等, 2012) 等. 式 (3.8) 所建议的超参数设置具有以下优点: 1. 共轭性保证: 当观测增量中至少有两个值不相等时, 自动满足共轭先验的条件  $\omega < \lambda$ . 2. 众数合理性:  $\omega$  和  $\lambda$  决定了先验分布  $\pi(\alpha, \beta)$  的众数位置, 通过数据驱动方式可确保众数值的合理性. 3. 信息量可解释性: 超参数  $\delta$  类似于测量次数, 其大小可解释为等效的先验信息量:  $\delta = 1$ : 等效于单个产品的单次测量信息;  $\delta = 0$ : 非信息性先验, 表示完全不依赖先验信息;  $\delta = mn$ : 与所有观测数据的信息量等效, 表示非常强的先验信息. 基于上述优点, 本章在模拟研究与数据分析中采用自动化策略 (3.8) 指定超参数值, 从而简化共轭先验的设定过程.

**注 3:** 在实际工程应用中, 可能已经掌握了一些关于参数  $\alpha$  和  $\beta$  的先验信息 (如众数、均值、分位数和方差等). 然而, 由于联合先验分布  $\pi(\alpha, \beta)$  的形式复杂, 缺乏解析表达式, 直接确定其超参数  $(\delta, \lambda, \omega)$  可能较为困难. 以下是一种将这些先验信息合理融入到超参数设定中的方法, 分为以下三个步骤:

- (1) 设定  $\delta$  的值: 如注释 2 所述, 可以通过等效测量信息量, 主观地设定  $\delta$ ;
- (2) 确定  $\lambda$  的值: 利用  $\alpha$  的信息, 结合  $\beta$  的信息以及条件先验分布  $\pi(\beta | \alpha)$  来确定  $\lambda$ ;
- (3) 确定  $\omega$  的值: 根据  $\alpha$  的近似分布

$$Ga\left(\frac{\delta + 3}{2}, \delta \bar{T}_m \left[ \log\left(\frac{\lambda}{\omega}\right) + \log\left(\frac{\prod j = 1^m t_j^{t_j/T_m}}{\bar{T}_m}\right) \right]\right),$$

结合  $\alpha$  的信息设定  $\omega$ . 例如, 对于  $\alpha$  和  $\beta$  的众数  $\alpha_M$  和  $\beta_M$ , 若工程师对其置信度较低, 则可以选择一个相对较小的  $\delta$  值. 给定  $\alpha = \alpha_M$  时, 可知  $\pi(\beta | \alpha)$

的众数为  $\bar{T}_m \alpha_M / \lambda$ . 之后可通过求解  $\beta_M = \bar{T}_m \alpha_M / \lambda$  得到  $\lambda = \bar{T}_m \alpha_M / \beta_M$ . 由于  $\alpha$  的边际先验的近似具有众数为如下形式:

$$\frac{\delta + 1}{2\delta \bar{T}_m \left[ \log\left(\frac{\lambda}{\omega}\right) + \log\left(\frac{\prod_{j=1}^m t_j^{t_j/T_m}}{\bar{T}_m}\right) \right]}.$$

给定众数  $\alpha_M$ , 可以得到

$$\omega = \frac{\prod_{j=1}^m t_j^{t_j/T_m}}{\bar{T}_m} \exp\left\{\log \lambda - \frac{\delta + 1}{2\delta \bar{T}_m \alpha_M}\right\}.$$

### 3.2.2 后验抽样

由于后验分布  $\pi(\alpha, \beta \mid \mathbf{y})$  无法得到解析形式, 直接计算  $\alpha$  和  $\beta$  的贝叶斯估计变得不可行. 蒙特卡罗方法为处理这种难以解析的后验分布提供了一种可行的替代推断手段. 该方法的核心思想是从参数的联合后验分布中抽取随机样本, 并利用这些样本对参数或其函数进行点估计和区间估计. 本节将介绍三种用于产生 AGG 分布随机数的算法.

#### 3.2.2.1 Gibbs 采样

Gibbs 采样是一种特殊的马尔可夫链蒙特卡罗 (Monte Carlo Markov chain, MCMC) 方法, 其通过从完全条件后验密度  $\pi(\beta \mid \alpha, \mathbf{y})$  和  $\pi(\alpha \mid \beta, \mathbf{y})$  迭代抽样来实现. 根据式 (3.5), 完全条件后验密度  $\pi(\beta \mid \alpha, \mathbf{y})$  是一个伽马分布  $Ga(\delta_p \bar{T}_m \alpha, \delta_p \lambda_p)$ ; 而  $\pi(\alpha \mid \beta, \mathbf{y})$  的完全条件后验密度与下式成比例

$$\frac{(\beta \omega_p)^{\delta_p \bar{T}_m \alpha}}{\left[ \prod_{j=1}^m (\Gamma(\alpha t_j))^{1/m} \right]^{\delta_p}}.$$

可以证明  $\pi(\alpha \mid \beta, \mathbf{y})$  具有对数凹的性质. 因此, 可以利用自适应拒绝采样 (Adaptive synthetic sampling, ARS) 算法生成  $\pi(\alpha \mid \beta, \mathbf{y})$  的随机样本 (Gilks 等, 1992). 在获得  $\alpha$  和  $\beta$  的后验样本后, 可构造任何参数的函数  $\eta = p(\alpha, \beta)$  (例如: 产品的可靠度、MTTF) 的贝叶斯估计. Gibbs 采样的后验推断过程详见算法 3.2.

### 3.2.2.2 离散网格采样

后验分布  $\pi(\alpha, \beta \mid \mathbf{y})$  可表示为条件分布  $\pi(\beta \mid \alpha, \mathbf{y})$  与边缘分布  $\pi(\alpha \mid \mathbf{y})$  的乘积, 即  $\pi(\alpha, \beta \mid \mathbf{y}) = \pi(\beta \mid \alpha, \mathbf{y})\pi(\alpha \mid \mathbf{y})$ , 其中  $\pi(\beta \mid \alpha, \mathbf{y})$  为  $Ga(\delta_p \bar{T}_m \alpha, \delta_p \lambda_p)$ ,  $\pi(\alpha \mid \mathbf{y})$  与以下函数成比例:

$$h_p(\alpha) = \frac{\left(\frac{\omega_p}{\delta_p \lambda_p}\right)^{\delta_p \bar{T}_m \alpha} \Gamma(1 + \delta_p \bar{T}_m \alpha)}{\left[\prod_{j=1}^m (\Gamma(\alpha t_j))^{1/m}\right]^{\delta_p}}. \quad (3.9)$$

根据上述分解, 当给定从  $\pi(\alpha \mid \mathbf{y})$  生成的随机样本  $\alpha_0$  时, 可直接从  $Ga(\delta_p \bar{T}_m \alpha_0, \delta_p \lambda_p)$  中生成  $\beta$  的随机样本. 因此, 目前主要挑战在于生成  $\alpha$  的随机样本. 由于  $\pi(\alpha \mid \mathbf{y})$  具有复杂性, 本节提出一种简便的离散网格采样法 (Discrete grid sampling, DGS): 通过在一系列网格点上构建离散分布来近似  $\alpha$  的边缘后验分布, 从而实现  $\alpha$  的统计推断. 具体方法如下: 首先, 选取一个区间  $[A_1, A_2]$ , 使  $\alpha$  落入该区间的概率接近 1, 即, 计算定积分  $\int_{A_1}^{A_2} \pi(\alpha \mid \mathbf{y}) d\alpha$  的值是否接近 1. 区间  $[A_1, A_2]$  可利用六西格玛法来确定, 具体步骤如下:

(a) 确定  $\tilde{\alpha} = \arg \max_{\alpha} \log h_p(\alpha)$ , 以及

$$I(\tilde{\alpha}) = -\frac{\partial^2 \log h_p(\alpha)}{\partial \alpha^2} \bigg|_{\alpha=\tilde{\alpha}}.$$

(b) 根据 (Berger, 2013) 的研究,  $\pi(\alpha \mid \mathbf{y})$  可以用正态分布  $N(\tilde{\alpha}, \tilde{\sigma}^2)$  近似, 其中  $\tilde{\sigma} = \sqrt{1/I(\tilde{\alpha})}$ .

(c) 构建区间  $A_1 = \max(0, \tilde{\alpha} - 6\tilde{\sigma})$  和  $A_2 = \tilde{\alpha} + 6\tilde{\sigma}$ . 根据正态分布的性质, 可知  $\alpha$  落在区间  $[A_1, A_2]$  内的概率几乎为 1.

在区间  $[A_1, A_2]$  内选择  $M$  个等间距网格点  $\{A_1 = \alpha^{(1)}, \alpha^{(2)}, \dots, A_2 = \alpha^{(M)}\}$ , 并使用未标准化的后验密度  $h_p(\alpha)$  计算每个网格点的概率

$$P(\alpha = \alpha^{(s)}) = \frac{h_p(\alpha^{(s)})}{\sum_{i=1}^M h_p(\alpha^{(i)})}, \quad s = 1, \dots, M. \quad (3.10)$$

通过选择足够大的  $M$ , 可以确保离散分布的近似精度. 通过离散分布采样显著简化了从  $\pi(\alpha \mid \mathbf{y})$  中抽样的难度. 例如, 在 R 语言中, 可以直接使用 `sample()` 函数来模拟随机样本. 上述方法的核心是将复杂的连续分布转换为易操作的离散分布, 显著降低了抽样的计算复杂度. 基于该方法的后验推断过程可见算法 3.3.

### 3.2.2.3 重要性重采样

第三种算法与第二种算法的主要区别在于生成  $\pi(\alpha \mid \mathbf{y})$  后验样本的方法, 该算法采用了重要性重采样 (Sampling Importance Resampling, SIR) 方法. 在 SIR 中, 采样步骤首先从辅助分布  $g(\alpha)$  中生成一组随机数, 然后通过重新计算权重, 将其近似为来自  $\pi(\alpha \mid \mathbf{y})$  的样本. 辅助分布  $g(\alpha)$  的选择较为灵活, 可以从一组易于抽样的分布中选取. 然而, SIR 方法的效率取决于  $g(\alpha)$  和  $\pi(\alpha \mid \mathbf{y})$  的相似程度, 尤其是分布尾部的相似性会影响近似的精度. 从式 (3.6) 可知,  $\pi(\alpha \mid \mathbf{y})$  的尾部形式 (当  $\alpha \rightarrow \infty$  时) 与形状参数  $(\delta_p + 3)/2$  和尺度参数

$$\nu = \delta_p \bar{T}_m \left[ \log \left( \frac{\lambda_p}{\omega_p} \right) + \log \left( \frac{\prod j = 1^m t_j^{t_j/T_m}}{\bar{T}_m} \right) \right]$$

的伽马分布尾部同阶. 因此, 伽马分布  $Ga(a, b)$  是一个合适的辅助分布, 其参数  $a$  和  $b$  可以通过以下步骤来确定:

(a) 令  $\tilde{\alpha} = \arg \max_{\alpha} \log h_p(\alpha)$ , 并计算

$$I(\tilde{\alpha}) = -\frac{\partial^2 \log h_p(\alpha)}{\partial \alpha^2} \bigg|_{\alpha=\tilde{\alpha}}.$$

(b) 初始化  $b$  为  $b_0 = \nu$ , 并初始化  $a$  为  $a_0 = \tilde{\alpha}b_0$ . 确保辅助分布  $Ga(a_0, b_0)$  的均值为  $\tilde{\alpha}$ .

(c) 计算精度比  $R = \frac{b_0^2/a_0}{I(\tilde{\alpha})}$ , 然后更新参数  $a = a_0/R$  和  $b = b_0/R$ . 该步骤不会改变辅助分布的均值, 但可以使  $Ga(a, b)$  的方差与  $\pi(\alpha \mid \mathbf{y})$  的渐近方差一致.

通过以上步骤确定了辅助分布  $Ga(a, b)$ , 并从中产生  $M$  个随机样本

$\{\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(M)}\}$ . 计算权重  $w_i = h_p(\alpha^{(i)}) / f_{Ga}(\alpha^{(i)} | a, b)$ ,  $i = 1, \dots, M$ , 其中  $f_{Ga}(\alpha^{(i)} | a, b)$  表示  $Ga(a, b)$  在  $\alpha^{(i)}$  处的 PDF 值. 对权重进行归一化, 得到  $\tilde{w}_i = w_i / \sum_{j=1}^M w_j$ . 最后, 从离散分布中生成  $\alpha$  的随机样本, 其分布律为

$$P(\alpha = \alpha^{(i)}) = \tilde{w}_i, \quad i = 1, 2, \dots, M. \quad (3.11)$$

SIR 算法的后验推断过程可见算法 3.4.

### 3.2.3 模拟实验

在开展模拟研究之前, 首先利用所提方法分析激光退化数据, 其退化路径如图 1.2 所示, 其中  $n = 15, m = 16, \omega = 10$ . 在应用伽马过程拟合数据之前, 首先, 通过 Kolmogorov-Smirnov 检验来验证数据是否符合伽马过程的假设 (Marsaglia 等, 2003). 由于该数据集中所有  $t_j$  的值均为 250 小时, 因此  $y_{ij} \sim Ga(250\alpha, \beta)$ . 使用 R 软件中的 `ks.test()` 函数进行检验, 得到 p 值为 0.13, 大于显著性水平 0.05. 这表明伽马过程能够用来拟合该数据. 因此, 假设激光器件的退化路径服从伽马过程  $\mathcal{GP}(\alpha t, \beta)$ . 接着, 基于共轭先验 (3.4) 进行贝叶斯推断, 其中设定  $\delta = 1$  且  $\lambda = \bar{y}_a$ . 由于数据是等间隔测量的, 因此  $\omega = \bar{y}_g = \prod_{i=1}^n \prod_{j=1}^m y_{ij}^{1/(mn)}$ , 即  $\{y_{ij}, i = 1, \dots, 15, j = 1, \dots, 16\}$  的几何平均值. 在第 3.2.1 节中曾讨论过,  $\delta = 1$  表示先验信息仅相当于一次测量的信息量, 与总计  $mn$  次测量数据相比, 先验信息较弱. 因此,  $\alpha$  和  $\beta$  的后验分布为  $AGG(mn + 1, \prod_{i=1}^n \prod_{j=1}^m y_{ij}^{1/(mn)}, \bar{y}_a)$ . 利用上述三种算法分别得到  $\alpha$  和  $\beta$  的点估计及其 95% 区间估计, 同时估计 4500 (小时) 时设备的可靠度  $R(4500)$ . 估计结果列在表 3.1 中 (其中“GS”表示基于 Gibbs 采样的算法). 在 Gibbs 采样中, 设定迭代次数  $K_1$  为 3000, 其中前 1000 次迭代作为预烧样本舍弃, 并设定稀释 (thinning) 间隔为 2, 最终保留 1000 个有效样本用于后验推断. 在 DGS 中, 离散化区间为  $[0, 10]$ , 网格点数量设为 10000, 后验推断的样本量同样为 1000. 在 SIR 中, 设定  $M = 10,000$ , 并从中重采样  $K_3 = 1,000$  个样本. 正如表 3.1 所示, 三种算法下  $\alpha$ 、 $\beta$  和  $R(4500)$  的贝叶斯点估计及 95% 区间估计上均表现出高度一致性.

为了更全面地比较这三种算法, 本节在具有不同信息量的共轭先验下进行了模拟研究. 数据通过伽马过程  $\mathcal{GP}(\alpha t, \beta)$  生成, 参数设定为  $\alpha = 0.031$  和  $\beta = 15.35$  (接近表 3.1 中的估计值). 实验中, 共对  $n = 15$  个样本进行

表 3.1:  $\alpha$ 、 $\beta$  和  $R(4500)$  的点估计及其 95% 置信区间.

估计值	GS 算法			DGS 算法			SIR 算法		
	$\alpha$	$\beta$	$R(4500)$	$\alpha$	$\beta$	$R(4500)$	$\alpha$	$\beta$	$R(4500)$
Point	0.0309	15.342	0.879	0.0308	15.325	0.878	0.0310	15.438	0.882
2.5%	0.0258	12.693	0.740	0.0260	12.698	0.737	0.0256	12.677	0.743
97.5%	0.0366	18.332	0.963	0.0370	18.328	0.962	0.0366	18.368	0.964

了测试, 每个样本以 250 小时的间隔进行测量, 总计进行了  $m = 16$  次测量, 并将失效阈值设定为 10. 采用共轭先验分布  $\text{AGG}(\delta, \bar{y}_g, \bar{y}_a)$ , 通过设定  $\delta = 0, 1, m/4, m/2$  来评估先验信息量对结果的影响.

随机生成了  $N = 10000$  个数据集, 并应用所提出的算法进行分析, 得到  $\alpha$ 、 $\beta$ 、 $R(4500)$  和 MTTF 的贝叶斯点估计及其 95% 区间估计. 基于这些估计结果, 计算了相对偏差 (Relative bias, RB) 和 RMSE, 结果分别列于表 3.2 和表 3.3. 在所有情况下, 参数估计均表现出令人满意的效果:  $\alpha$ 、 $\beta$  和  $R(4500)$  的贝叶斯估计的 RB 约为 2%, 而 MTTF 的贝叶斯估计的 RB 约为 0.1%. 三种算法在 RB 和 RMSE 上的表现几乎一致, 且先验信息量  $\delta$  的变化对估计结果的影响不显著.

在评估参数的区间估计时, 计算了 95% 可信区间的平均长度和 CP. 结果分别列在表 3.4 和表 3.5. 分析发现, 随着先验信息量的增加, 95% 置信区间的长度逐渐减小, 且三种算法在这一趋势上表现相似, 差异不显著. 然而, 在覆盖概率方面, 不同算法表现出一定差异. 对于模型参数  $\alpha$  和  $\beta$ , 无论先验信息量  $\delta$  的变化如何, 基于 DGS 和 SIR 算法的覆盖概率更接近理想的 95% 水平, 而基于 GS 算法的覆盖概率则相对较低. 对于  $R(4500)$  和 MTTF 的估计, 三种算法的覆盖概率均接近名义水平 (即, 95%), 表明所提出的后验采样算法在这些指标估计中具有较高的准确性.

在计算效率方面, 三种算法 (Gibbs 采样、DGS 以及 SIR) 在每个数据集上的平均运行时间分别为 0.602 秒、0.00341 秒以及 0.00499 秒. 测试环境为运行 Windows 11 操作系统的台式机, 配备 Intel(R) Core(TM) i7-10700 CPU(2.9 GHz) 和 16 GB RAM. 该结果显示, DGS 和 SIR 算法的计算效率相近, 并且都显著快于 Gibbs 算法, 速度提升了超过一百倍. 在实时推断的应用场景中, 计算效率是一个关键因素. 随着新观测数据的不断收集, 需要实时更新后验分布, 同时在保持估计精度的情况下, 尽可能快速完成推断过程. 正如表 3.2 至表 3.5 所

表 3.2: 参数点估计的相对偏差.

算法	$\delta = 0$				$\delta = 1$			
	$\alpha$	$\beta$	$R(4500)$	MTTF	$\alpha$	$\beta$	$R(4500)$	MTTF
GS	0.0245	0.0256	0.0161	0.00109	0.0243	0.0254	0.0161	0.00108
DGS	0.0245	0.0256	0.0161	0.0011	0.0245	0.0256	0.0161	0.00109
SIR	0.0245	0.0256	0.0161	0.00109	0.0245	0.0256	0.0161	0.00109

算法	$\delta = \frac{m}{4}$				$\delta = \frac{m}{2}$			
	$\alpha$	$\beta$	$R(4500)$	MTTF	$\alpha$	$\beta$	$R(4500)$	MTTF
GS	0.0233	0.0247	0.0153	0.00136	0.0234	0.0248	0.0151	0.00137
DGS	0.0234	0.0249	0.0152	0.00137	0.0233	0.0247	0.0151	0.00136
SIR	0.0234	0.0249	0.0152	0.00138	0.0232	0.0246	0.0151	0.00136

表 3.3: 参数点估计的 RMSE.

算法	$\delta = 0$				$\delta = 1$			
	$\alpha$	$\beta$	$R(4500)$	MTTF	$\alpha$	$\beta$	$R(4500)$	MTTF
GS	0.00302	1.547	0.0601	115.218	0.00302	1.547	0.0601	115.205
DGS	0.00301	1.539	0.0601	115.267	0.00301	1.538	0.0601	115.258
SIR	0.00301	1.539	0.0601	115.267	0.00301	1.538	0.0601	115.258

算法	$\delta = \frac{m}{4}$				$\delta = \frac{m}{2}$			
	$\alpha$	$\beta$	$R(4500)$	MTTF	$\alpha$	$\beta$	$R(4500)$	MTTF
GS	0.00302	1.546	0.0599	115.204	0.00296	1.512	0.0607	116.192
DGS	0.00301	1.537	0.0600	115.250	0.00294	1.501	0.0607	116.205
SIR	0.00301	1.537	0.0600	115.250	0.00294	1.501	0.0607	116.205

表 3.4: 参数 95% 可信区间的长度.

算法	$\delta = 0$				$\delta = 1$			
	$\alpha$	$\beta$	$R(4500)$	MTTF	$\alpha$	$\beta$	$R(4500)$	MTTF
GS	0.0109	5.588	0.224	447.444	0.0109	5.585	0.224	446.609
DGS	0.0109	5.629	0.224	446.274	0.0109	5.620	0.223	445.596
SIR	0.0110	5.630	0.224	446.414	0.0109	5.624	0.223	445.318

算法	$\delta = \frac{m}{4}$				$\delta = \frac{m}{2}$			
	$\alpha$	$\beta$	$R(4500)$	MTTF	$\alpha$	$\beta$	$R(4500)$	MTTF
GS	0.0108	5.541	0.222	444.082	0.0107	5.502	0.220	440.493
DGS	0.0108	5.582	0.221	443.013	0.0107	5.535	0.219	439.562
SIR	0.0109	5.581	0.221	443.131	0.0108	5.536	0.219	439.373

表 3.5: 参数 95% 可信区间的覆盖概率.

算法	$\delta = 0$				$\delta = 1$			
	$\alpha$	$\beta$	$R(4500)$	MTTF	$\alpha$	$\beta$	$R(4500)$	MTTF
GS	0.9372	0.9369	0.9474	0.9492	0.9344	0.9336	0.9458	0.9475
DGS	0.9424	0.9424	0.9463	0.9481	0.9416	0.9396	0.9460	0.9468
SIR	0.9418	0.9406	0.9467	0.9473	0.9411	0.9413	0.9471	0.9484
算法	$\delta = \frac{m}{4}$				$\delta = \frac{m}{2}$			
	$\alpha$	$\beta$	$R(4500)$	MTTF	$\alpha$	$\beta$	$R(4500)$	MTTF
GS	0.9340	0.9337	0.9455	0.9478	0.9385	0.9369	0.9399	0.9413
DGS	0.9394	0.9384	0.9454	0.9466	0.9434	0.9428	0.9411	0.9423
SIR	0.9406	0.9384	0.9451	0.9456	0.9451	0.9424	0.9415	0.9427

示, DGS 和 SIR 算法不仅在计算效率上具有明显优势, 而且在估计精度方面表现优异, 充分满足在线推断的需求. 基于上述结果, 在后续章节中, 将主要采用 DGS 和 SIR 算法用于在线 RUL 的预测研究.

### 3.2.4 异质性

产品间的异质性通常源于内在因素和外在因素的差异. 内在因素包括原材料和生产工艺的变化, 而外在因素则涉及操作环境和使用习惯的不同. 这种异质性导致每个产品的性能退化轨迹各不相同. 尽管如此, 由于这些产品源自同一总体, 它们的失效机制通常是一致的. 基于此, 假设第  $i$  个产品的性能退化服从伽马过程  $\mathcal{GP}(\alpha t, \beta_i)$ . 这里形状参数  $\alpha$  相同是由于产品失效机制一致, 而不同的尺度参数  $\beta_i$  则刻画产品间的异质性. 为简化符号, 假设采用等间隔测量, 即相邻测量时间点的间隔为  $l$ . 试验中测试了  $n$  个产品, 在时间点  $T_m = ml$  时, 每个产品进行了  $m$  次测量. 令  $Y_{ij}$  表示第  $i$  个产品在时间点  $T_j = jl$  的退化值, 其中  $i = 1, \dots, n, j = 1, \dots, m$ . 定义退化增量为  $y_{ij} = Y_{ij} - Y_{ij-1}$ , 并设  $Y_{i0} = 0$ . 在时间点  $T_m$  处, 记观测数据为  $\mathbf{y}_{(m)} = \{y_{ij}, i = 1, \dots, n, j = 1, \dots, m\}$ . 由于  $Y_i(t) \sim \mathcal{GP}(\alpha t, \beta_i)$ , 可得  $y_{ij} \sim Ga(\alpha l, \beta_i)$ . 基于观测数据  $\mathbf{y}_{(m)}$ , 对应的似然函数为

$$L(\mathbf{y}_{(m)} | \alpha, \beta_1, \dots, \beta_n) = \prod_{i=1}^n \prod_{j=1}^m \frac{\beta_i^{\alpha l}}{\Gamma(\alpha l)} y_{ij}^{\alpha l-1} \exp\{-\beta_i y_{ij}\}$$

$$\propto \frac{\bar{\beta}_g^{mn l \alpha}}{[\Gamma(\alpha)]^{mn}} \bar{y}_{g(m)}^{mn l \alpha} \exp \left\{ - \sum_{i=1}^n m \bar{y}_{i(m)} \beta_i \right\}, \quad (3.12)$$

其中

$$\begin{aligned} \bar{\beta}_g &= \left[ \prod_{i=1}^n \beta_i \right]^{1/n}, \bar{y}_{g(m)} = \left[ \prod_{i=1}^n \prod_{j=1}^m y_{ij} \right]^{\frac{1}{mn}}, \\ \bar{y}_{i(m)} &= \frac{1}{m} \sum_{j=1}^m y_{ij}, i = 1, \dots, n. \end{aligned}$$

**定理 3.2.** 基于似然函数(3.12),  $(\alpha, \beta_1, \dots, \beta_n)'$  的共轭先验为

$$\pi(\alpha, \beta_1, \dots, \beta_n) = C \frac{(\bar{\beta}_g \omega)^{\delta_1 l \alpha}}{[\Gamma(l \alpha)]^{\delta_1}} \exp \left\{ - \sum_{i=1}^n \delta_2 \lambda_i \beta_i \right\}, \quad (3.13)$$

其中  $C$  是正则化常数,  $\delta_1$ 、 $\delta_2$ 、 $\omega$  和  $\lambda_i$  是非负超参数.

定理 3.2 的证明见本节附录 3.2.7. 当  $\beta_1 = \dots = \beta_n = \beta$ ,  $\delta_1 = n \delta_2 = \delta$  且  $\lambda_1 = \dots = \lambda_n = \lambda$  时, 共轭先验 (3.13) 简化为 (3.7). 为了更好地理解共轭先验 (3.13), 可将  $\pi(\alpha, \beta_1, \dots, \beta_n)$  分解为

$$\begin{aligned} \pi(\alpha, \beta_1, \dots, \beta_n) &= \prod_{i=1}^n \pi(\beta_i | \alpha) \pi(\alpha) \\ &\propto \prod_{i=1}^n \frac{(\delta_2 \lambda_i)^{1+\delta_1 l \alpha / n} \beta_i^{\delta_1 l \alpha / n}}{\Gamma(1 + \delta_1 l \alpha / n)} \exp\{-\delta_2 \lambda_i \beta_i\} \\ &\quad \times \frac{[\Gamma(1 + \delta_1 l \alpha / n)]^n}{[\Gamma(l \alpha)]^{\delta_1}} \exp \left\{ -\alpha \delta_1 l \left[ \log \left( \frac{\delta_2}{\omega} \right) + \frac{1}{n} \sum_{i=1}^n \log \lambda_i \right] \right\}. \end{aligned}$$

因此, 给定  $\alpha$  时,  $\beta_i$  的条件分布为  $Ga(1 + \delta_1 l \alpha / n, \delta_2 \lambda_i)$ , 而  $\alpha$  的边缘 PDF 与下式成比例

$$\begin{aligned} g(\alpha) &= \frac{[\Gamma(1 + \delta_1 l \alpha / n)]^n}{[\Gamma(l \alpha)]^{\delta_1}} \\ &\quad \times \exp \left\{ -\alpha \delta_1 l \left[ \log \left( \frac{\delta_2}{\omega} \right) + \frac{1}{n} \sum_{i=1}^n \log \lambda_i \right] \right\}. \quad (3.14) \end{aligned}$$

利用 Stirling 公式, 当  $\alpha \rightarrow \infty$  时, 有

$$g(\alpha) \equiv O\left(\alpha^{\frac{\delta_1+n}{2}} \exp\{-A\alpha\}\right),$$

其中

$$A = \delta_1 l \left[ \log\left(\frac{n\delta_2}{\delta_1}\right) + \frac{1}{n} \sum_{i=1}^n \log\left(\frac{\lambda_i}{\omega}\right) \right].$$

因此, 当  $\alpha \rightarrow \infty$  时,  $\pi(\alpha)$  趋于 0 的速度与伽马分布  $\text{Ga}\left(\frac{\delta_1+n+2}{2}, K\right)$  同阶. 因此, 称  $\pi(\alpha, \beta_1, \dots, \beta_n)$  为近似伽马-多元伽马 (Approximated-gamma-multivariate gamma, AGMG) 分布, 维度为  $n$ , 记作  $\text{AGMG}_n(\boldsymbol{\gamma}, \omega, \boldsymbol{\xi})$ , 其中  $\boldsymbol{\gamma} = (\delta_1, \delta_2)', \boldsymbol{\xi} = (\lambda_1, \dots, \lambda_n)'$ .

由式 (3.21) 可知, 参数  $(\alpha, \beta_1, \dots, \beta_n)'$  的后验分布服从  $\text{AGMG}_n(\boldsymbol{\gamma}_{p(m)}, \omega_{p(m)}, \boldsymbol{\xi}_{p(m)})$ , 其中  $\boldsymbol{\gamma}_{p(m)} = (\delta_{1p(m)}, \delta_{2p(m)})'$  和  $\boldsymbol{\xi}_{p(m)} = (\lambda_{1p(m)}, \dots, \lambda_{np(m)})'$ . 超参数  $\omega$  和  $\boldsymbol{\xi}$  的具体选择分别为  $\bar{y}g(m)$  和  $\bar{\mathbf{y}}(m) = (\bar{y}1(m), \dots, \bar{y}n(m))'$ . 由于后验分布的形式为  $\text{AGMG}_n(\boldsymbol{\gamma}_{p(m)}, \bar{y}g(m), \bar{\mathbf{y}}(m))$ . 在该设置下, 超参数  $\delta_1$  和  $\delta_2$  与测量次数有类似的解释. 由等式 (3.7) 可知,  $\delta_1$  和  $\delta_2$  主要决定 AGMG 分布的峰度, 从而控制先验信息的权重. 生成 AGMG 分布的随机样本可以通过适当修改算法 3.3 和 3.4 来实现. 主要区别在于目标分布变更为  $\text{AGMG}_n(\boldsymbol{\gamma}_{p(m)}, \omega_{p(m)}, \boldsymbol{\xi}_{p(m)})$  中  $\alpha$  的后验边缘分布, 而在给定  $\alpha$  时,  $\beta_i$  则从伽马分布  $\text{Ga}\left(1 + \delta_{1p(m)}l\alpha/n, \delta_{2p(m)}\lambda_{ip(m)}\right)$  中产生随机数, 其中  $i = 1, \dots, n$ . 这两种算法 (DGS 和 SIR) 的计算时间与维度  $n$  成正比. 为了说明这一点, 本小节在相同的参数设置下, 对  $n$  从 2 到 50 的 AGMG 分布应用了这两种算法, 并记录了计算时间, 结果如图 3.2 所示. 从图中可以看出, 计算时间随着  $n$  的增加而线性增长. 例如, 当  $n$  从 2 增加到 50 时, DGS 的计算时间从 0.00328 秒增加到 0.00895 秒, 而 SIR 的计算时间从 0.00448 秒增加到 0.0101 秒. 这表明即使在较大的  $n$  值下, 这两种算法仍然保持着较高的计算效率.

RUL 预测是产品故障诊断与健康管理的核心基础. 在实际应用中, 当产品处于正常工作状态时 (即退化量尚未达到失效阈值), 需要基于实时监测数据动态评估其剩余使用寿命. 具体而言, 假设在时间  $t_m$  监测到第  $i$  个产品的退化轨迹  $\{Y_{i1}, \dots, Y_{im}\}$  均满足  $Y_{ij} < \omega$  (其中  $\omega$  为失效阈值), 则其在时间  $t_m$

的 RUL 定义为

$$Z_{it_m} = \inf\{z : Y_i(z + t_m) > \omega | Y_{ij} < \omega, j = 1, \dots, m\},$$

其分布函数为

$$\begin{aligned} F_{Z_{it_m}}(z | \alpha, \beta_i) &= P(Y_i(z + t_m) > \omega) = P(Y_i(z) > \omega - Y_{im}) \\ &= \frac{\Psi(\beta(\omega - Y_{im}), \alpha z)}{\Gamma(\alpha z)}. \end{aligned} \quad (3.15)$$

由伽马过程的齐次性可知, 最后两个等式成立. 由于  $F_{Z_{it_m}}(z | \alpha, \beta_i)$  较为复杂, 可用两参数 BS 分布  $BS(\alpha_{im}^*, \beta_{im}^*)$  近似 (Park 等, 2005), 其中

$$\alpha_{im}^* = \sqrt{\frac{1}{\beta_i(\omega - Y_{im})}}, \quad \beta_{im}^* = \frac{\beta_i(\omega - Y_{im})}{\alpha}.$$

因此, 在时间点  $t_m$ , 第  $i$  个产品 RUL 的均值可近似为

$$\mu_{im}(\alpha, \beta_i) = \beta_{im}^* \left(1 + (\alpha_{im}^*)^2 / 2\right) = \frac{1 + 2\beta_i(\omega - Y_{im})}{2\alpha}.$$

此外,  $Z_{it_m}$  的分布  $\rho$  分位数可近似为

$$\mu_{im}^\rho(\alpha, \beta_i) = \frac{\beta_{im}^*}{4} \left[ u_\rho \alpha_{im}^* + \sqrt{(u_\rho \alpha_{im}^*)^2 + 4} \right]^2,$$

其中  $u_\rho$  是标准正态分布的  $\rho$  分位数. 在时间  $t_m$ , 第  $i$  个产品 RUL 的贝叶斯点预测为

$$\tilde{\mu}_{im} = \int_0^\infty \int_0^\infty \mu_{im}(\alpha, \beta_i) \pi(\alpha, \beta_i | \mathbf{y}_{(m)}) d\alpha d\beta_i. \quad (3.16)$$

在时间  $t_m$  时, 第  $i$  个产品 RUL 的  $1 - \rho$  可信水平的贝叶斯区间预测为

$$(\tilde{\mu}_{im}^{\rho/2}, \tilde{\mu}_{im}^{1-\rho/2}), \quad (3.17)$$

其中

$$\tilde{\mu}_{im}^\rho = \int_0^\infty \int_0^\infty \mu_{im}^\rho(\alpha, \beta_i) \pi(\alpha, \beta_i | \mathbf{y}_{(m)}) d\alpha d\beta_i.$$

给定后验样本  $(\alpha^{(k)}, \beta_i^{(k)}), k = 1, \dots, K$ , 式 (3.16) 和 (3.17) 可通过蒙特卡罗方法近似:

$$\tilde{\mu}_{im} \approx \frac{1}{K} \sum_{k=1}^K \mu_{im}(\alpha^{(k)}, \beta_i^{(k)}), \quad \tilde{\mu}_{im}^\rho \approx \frac{1}{K} \sum_{k=1}^K \mu_{im}^\rho(\alpha^{(k)}, \beta_i^{(k)}). \quad (3.18)$$

**注 4:** DGS 和 SIR 算法主要用于产生模型参数  $(\alpha, \beta_1, \dots, \beta_n)'$  的后验样本. 基于这些后验样本, 可通过式 (3.18) 对  $n$  个产品进行 RUL 预测. 算法具备高度灵活性, 既适用于单个产品的预测, 也适用于多个产品的协同预测. 当  $n = 1$  时, 算法仅利用单个产品的信息, 退化为第 3.2.2 节中的内容. 当  $n \geq 2$  时, 算法通过共享多个产品的信息来共同估计参数  $\alpha$ , 并利用其他产品的信息提高参数  $\beta_i$  的估计精度.

### 3.2.5 RUL 预测

假设在时间点  $t_{m+1} = (m+1)l$  收集到了  $n$  个产品的新退化增量  $\mathbf{y}_{(m+1)} = (y_{1m+1}, \dots, y_{nm+1})'$ . 在获取新的观测数据后,  $(\alpha, \beta_1, \dots, \beta_n)'$  的后验分布需要更新. 对于具有共轭先验的贝叶斯推断, 可使用递推公式来实现更新. 由式 (3.21) 可知, 在时间  $t_{m+1} = (m+1)l$  时,  $(\alpha, \beta_1, \dots, \beta_n)'$  的后验分布为  $\text{AGMG}_n(\gamma_{p(m+1)}, \omega_{p(m+1)}, \xi_{p(m+1)})$ , 其中参数  $\gamma_{p(m+1)}$ 、 $\omega_{p(m+1)}$  和  $\xi_{p(m+1)}$  可实现递推更新. 即

$$\begin{aligned} \gamma_{p(m+1)} &= \gamma_{p(m)} + (n, 1)', \\ \omega_{(m+1)} &= \omega_{(m)}^{\frac{mn+\delta_1}{(m+1)n+\delta_1}} \left[ \prod_{i=1}^m y_{im+1} \right]^{\frac{1}{(m+1)n+\delta_1}}, \\ \lambda_{(m+1)} &= \frac{m+\delta_2}{m+1+\delta_2} \lambda_{(m)} + \frac{1}{m+1+\delta_2} \mathbf{y}_{(m+1)}. \end{aligned} \quad (3.19)$$

由式 (3.19) 可知, 后验分布的更新仅依赖  $t_m$  时刻后验分布中的参数值和新增观测数据, 可显著提升计算和数据存储效率. 本章提出的高效 DGS 和 SIR 算法, 可快速生成 AGMG 分布样本, 并用于多个产品的在线 RUL 预测. 具体过程如下:

---

**算法 3.1: 在线 RUL 预测**


---

**Input:** 时间  $t_m = ml$  时  $(\alpha, \beta_1, \dots, \beta_n)'$  的后验分布参数值, 以及时间  $t_{m+1} = (m+1)l$  时的新观测数据  $(y_{1m+1}, \dots, y_{nm+1})$ .

**Output:**  $n$  个产品的 RUL 点估计和  $100(1 - \rho)\%$  置信区间.

- 1 根据式 (3.19) 更新  $(\alpha, \beta_1, \dots, \beta_n)'$  的后验分布. 通过 DGS 或 SIR 从更新后的后验分布中生成随机样本.
- 2 根据式 (3.18) 计算  $n$  个产品的 RUL 点估计和  $100(1 - \rho)\%$  区间估计.

---

### 3.2.6 实例分析

#### 3.2.6.1 激光退化数据

基于第 3.2.3 节的激光退化数据集 (见图 1.2), 我们重点分析第 1、6 和 10 个激光设备. 观测数据显示, 这三个设备在运行 4000 小时后均因性能退化量超过阈值  $\omega = 10$  而发生失效. 由于监测间隔为 250 小时, 其失效时间区间可精确锁定为相邻监测时间点之间. 例如, 第一个设备的退化值在 3750 至 4000 小时之间超过了阈值. 因此, 可采用线性插值法来估计其失效时间. 第一个设备在 3750 小时时记录的退化值为 9.87, 而在 4000 小时时为 10.94. 通过线性插值法, 其失效时间估计为

$$3750 + \frac{10 - 9.87}{10.94 - 9.87} \times (4000 - 3750) = 3785.75 \text{ 小时.}$$

同理, 第 6 和第 10 个设备的失效时间分别估计为 3506.75 小时和 3351.25 小时.

通过 Kolmogorov-Smirnov 检验验证各设备退化数据与伽马过程  $\mathcal{GP}(\alpha t, \beta_i)$  的拟合优度. 如图 3.3 所示, 所有设备的检验  $p$  值均大于 0.05 的显著性水平, 统计上无法拒绝原假设, 表明伽马过程模型能够有效描述退化轨迹的随机特性.

基于上述模型验证结果, 应用第 3.2.5 节提出的在线预测算法进行实时 RUL 预测. 具体实施过程如下: 从  $t = 500$  小时开始, 仅基于前两次监测数据初始化预测模型; 随着监测数据累积, 通过算法 3.1 递归更新参数后验分布; 计算出每个时间点三台设备的 RUL 点预测值及其 95% 预测区间. 图 3.4 展示了随时间变化的  $\alpha$ 、 $\beta_1$ 、 $\beta_6$  和  $\beta_{10}$  的估计值. 从图中可发现,  $\beta_1$ 、 $\beta_6$  和  $\beta_{10}$  的估计值呈上升趋势, 且它们之间差异显著, 表明设备间具有较强的异质性. 图 3.5 展示了在各个时间点上三台设备 RUL 的点预测值及其 95% 预测区间, 并与每

个时间点的实际 RUL 进行了对比. 如图所示, 几乎所有实际 RUL 均被 95% 预测区间覆盖, 且点预测值与实际 RUL 非常接近, 显示了预测的高准确性.

### 3.2.6.2 列车车轮数据

本节将使用所提出的模型和方法对列车车轮数据 (见图 1.5) 进行分析, 并预测每个测量点上车轮的 RUL. 从图 1.5 可以看出, 退化路径呈线性并且单调递增. 有三个车轮在 600 千公里前已经失效. 通过线性插值法计算第 5、第 9 和第 11 个车轮的失效时间, 结果分别为 523.537、558.861 和 421.508 千公里.

基于 Kolmogorov-Smirnov 检验结果 (图 3.6 显示所有车轮退化数据的检验  $p$  值均大于 0.05), 验证了伽马过程模型对列车车轮退化数据的适用性. 采用考虑异质性效应的伽马过程模型进行建模, 并从第二次监测时点开始在线 RUL 预测. 如图 3.7 所示, 三个车轮的 RUL 点预测值与实际退化轨迹高度吻合, 且所有时间节点的真实 RUL 均被 95% 预测区间完整覆盖, 表明所提算法在不确定性量化与工程实用性方面具有显著优势, 可为轨道交通关键部件的预防性维护决策提供可靠依据.

## 3.2.7 附录

### 定理 3.1 的证明

基于似然函数 (3.3) 和先验 (3.4),  $\alpha$  和  $\beta$  的联合后验密度为

$$\begin{aligned} \pi(\alpha, \beta | \mathbf{y}) &\propto L(\mathbf{y} | \alpha, \beta) \pi(\alpha, \beta) \\ &\propto \frac{(\beta \omega_p)^{\delta_p \bar{T}_m \alpha}}{\left[ \prod_{j=1}^m (\Gamma(\alpha t_j))^{1/m} \right]^{\delta_p}} \exp\{-\delta_p \lambda_p \beta\}, \end{aligned} \quad (3.20)$$

其中

$$\begin{aligned} \delta_p &= mn + \delta, \\ \omega_p &= \omega^{\frac{\delta}{mn+\delta}} \left[ \prod_{i=1}^n \prod_{j=1}^m y_{ij}^{\frac{t_j}{mn\bar{T}_m}} \right]^{\frac{mn}{mn+\delta}}, \\ \lambda_p &= \frac{mn}{mn+\delta} \bar{y}_a + \frac{\delta}{mn+\delta} \lambda. \end{aligned}$$

因此,  $\pi(\alpha, \beta)$  和  $\pi(\alpha, \beta | \mathbf{y})$  均属于同一分布族.

### 不等式证明

由  $T_m = \sum_{j=1}^m t_j$ , 得到

$$\begin{aligned} \log \left( \frac{\prod_{j=1}^m t_j^{t_j/T_m}}{\bar{T}_m} \right) &= \sum_{j=1}^m \frac{t_j}{T_m} \log(t_j) - \log(\bar{T}_m) \\ &= \frac{\sum_{j=1}^m t_j \log(t_j)}{\sum_{j=1}^m t_j} - \log \left( \frac{1}{m} \sum_{j=1}^m t_j \right) \end{aligned}$$

令  $q(x) = x \log(x)$ , 则  $q(x)$  是凸函数. 利用 Jensen 不等式得

$$\begin{aligned} \sum_{j=1}^m t_j \log(t_j) &\geq m \cdot \frac{1}{m} \sum_{j=1}^m t_j \cdot \log \left( \frac{1}{m} \sum_{j=1}^m t_j \right) \\ &= \left( \sum_{j=1}^m t_j \right) \cdot \log \left( \frac{1}{m} \sum_{j=1}^m t_j \right). \end{aligned}$$

因此, 可以得出:  $\log \left( \frac{\prod_{j=1}^m t_j^{t_j/T_m}}{\bar{T}_m} \right) \geq 0$ .

### 三种后验抽样方法

---

#### 算法 3.2: Gibbs 采样

---

**Input:** 观测数据  $\mathbf{y}$ .

**Output:**  $\alpha, \beta$  和  $\eta$  的点估计和  $100(1 - \rho)\%$  置信区间.

- 1 随机初始化  $\alpha$  和  $\beta$ (例如,  $\alpha_0$  和  $\beta_0$ ), 并计算  $\delta_p$ 、 $\omega_p$  和  $\lambda_p$ .
- 2 **for**  $k$  in  $\{1, 2, \dots, K_1\}$  **do**
- 3   从  $\text{Ga}(\delta_p \bar{T}_m \alpha_{k-1}, \delta_p \lambda_p)$  中生成  $\beta_k$ ;
- 4   通过 ARS 算法从  $\pi(\alpha | \beta_k, \mathbf{y})$  中生成  $\alpha_k$ ;
- 5   计算  $\eta_k = p(\alpha_k, \beta_k)$ .
- 6 **end**
- 7 舍弃前  $B$  个预烧样本, 并选择间隔  $L$  来消除后验样本中的自相关性.
- 8 通过后验样本  $\{(\alpha_k, \beta_k, \eta_k), k = B+1, B+L+1, B+2L+1, \dots, K_1\}$  计算这些参数的点估计和  $100(1 - \rho)\%$  区间估计.

---

---

**算法 3.3: 离散网格采样**

---

**Input:** 观测数据  $y$ .**Output:**  $\alpha, \beta$  和  $\eta$  的点估计和  $100(1 - \rho)\%$  置信区间.

- 1 计算  $\delta_p, \omega_p$  和  $\lambda_p$ .
- 2 通过六西格玛法则确定区间  $[A_1, A_2]$ .
- 3 选择  $M$  个网格点  $\{A_1 = \alpha^{(1)}, \alpha^{(2)}, \dots, A_2 = \alpha^{(M)}\}$ , 计算每个网格的概率 (3.10).
- 4 **for**  $k$  in  $\{1, 2, \dots, K_2\}$  **do**
- 5   从离散分布 (3.10) 中生成  $\alpha_k$ ;
- 6   从  $\text{Ga}(\delta_p \bar{T}_m \alpha_k, \delta_p \lambda_p)$  中生成  $\beta_k$ ;
- 7   计算  $\eta_k = p(\alpha_k, \beta_k)$ .
- 8 **end**
- 9 通过后验样本  $\{(\alpha_k, \beta_k, \eta_k), k = 1, \dots, K_2\}$  计算这些参数的点估计和  $100(1 - \rho)\%$  区间估计.

---



---

**算法 3.4: 采样重要性重采样**

---

**Input:** 观测数据  $y$ .**Output:**  $\alpha, \beta$  和  $\eta$  的点估计和  $100(1 - \rho)\%$  置信区间.

- 1 计算  $\delta_p, \omega_p$  和  $\lambda_p$ .
- 2 按照上述三步确定  $a$  和  $b$ .
- 3 从  $\text{Ga}(a, b)$  中生成  $M$  个随机样本, 并构建离散分布 (3.11).
- 4 **for**  $k$  in  $\{1, 2, \dots, K_3\}$  **do**
- 5   从离散分布 (3.11) 中生成  $\alpha_k$ ;
- 6   从  $\text{Ga}(\delta_p \bar{T}_m \alpha_k, \delta_p \lambda_p)$  中生成  $\beta_k$ ;
- 7   计算  $\eta_k = p(\alpha_k, \beta_k)$ .
- 8 **end**
- 9 通过后验样本  $\{(\alpha_k, \beta_k, \eta_k), k = 1, \dots, K_3\}$  计算这些参数的点估计和  $100(1 - \rho)\%$  区间估计.

---

### 定理 3.2 的证明

基于似然函数 (3.12) 和先验 (3.13),  $(\alpha, \beta_1, \dots, \beta_n)'$  的联合后验密度为

$$\begin{aligned}\pi(\alpha, \beta_1, \dots, \beta_n \mid \mathbf{y}) &\propto L(\mathbf{y} \mid \alpha, \beta_1, \dots, \beta_n) \pi(\alpha, \beta_1, \dots, \beta_n) \\ &\propto \frac{\bar{\beta}_g^{(mn+\delta_1)l\alpha} \bar{y}_{g(m)}^{mn l \alpha} \omega^{\delta_1 l \alpha}}{[\Gamma(l\alpha)]^{mn+\delta_1}} \exp \left\{ -\sum_{i=1}^n (m\bar{y}_{i(m)} + \delta_2 \lambda_i) \beta_i \right\} \\ &\propto \frac{(\bar{\beta}_g \omega_{p(m)})^{\delta_{1p(m)} l \alpha}}{[\Gamma(l\alpha)]^{\delta_{1p(m)}}} \exp \left\{ -\sum_{i=1}^n \delta_{2p(m)} \lambda_{ip(m)} \beta_i \right\}, \quad (3.21)\end{aligned}$$

其中

$$\begin{aligned}\delta_{1p(m)} &= mn + \delta_1, \delta_{2p(m)} = m + \delta_2, \omega_{p(m)} = \omega^{\frac{\delta_1}{\delta_{1p(m)}}} \bar{y}_{g(m)}^{\frac{mn}{\delta_{1p(m)}}}, \\ \lambda_{ip(m)} &= \frac{m}{\delta_{2p(m)}} \bar{y}_{i(m)} + \frac{\delta_2}{\delta_{2p(m)}} \lambda_i, i = 1, \dots, n.\end{aligned}$$

因此,  $\pi(\alpha, \beta_1, \dots, \beta_n)$  和  $\pi(\alpha, \beta_1, \dots, \beta_n \mid \mathbf{y})$  属于同一分布族.

## 3.3 重参数化伽马过程的变分贝叶斯推断

随着传感器技术的快速发展, 退化数据采集能力的提升与模型复杂度的增加对统计推断方法提出了双重挑战:

一方面需处理高维异构数据, 另一方面需满足实时性计算需求. 传统 MCMC 方法虽为复杂退化模型提供了严格的贝叶斯推断框架, 但其存在三方面本质性局限 (Zaidan 等, 2016):

- 计算效率的维度诅咒: 即使采用高性能计算集群, 在处理中等规模数据集 (如  $n > 50$  的异质退化轨迹) 时, 马尔可夫链的混合速度随参数维度呈指数级下降. 以双随机效应伽马模型为例, 单次迭代耗时可达  $O(nm^2)$  量级, 导致实际工程场景中难以获得实时推断结果.
- 收敛诊断的主观依赖性: 预烧样本长度与抽样间隔的设定高度依赖经验性准则 (如 Gelman-Rubin 统计量或自相关分析). 在退化建模中, 由于参数空间存在强相关性 (如形状参数  $\alpha$  与速率参数  $\beta$ ), 往往需要超过  $10^4$  次迭代才能达到平稳分布, 显著增加计算资源消耗.

- 在线更新的结构性障碍: MCMC的批处理特性导致其无法利用增量数据实现递归更新. 当新增监测数据时, 需重新初始化并执行全链采样, 这种“推倒重来”的计算模式严重制约其在 PHM 系统中的应用.

变分贝叶斯 (Variational Bayesian, VB) 方法通过将贝叶斯推断转化为优化问题, 利用易处理的代理分布近似真实的后验分布, 并以 Kullback-Leibler (KL) 散度最小化为目标函数 (Blei 等, 2017; Jaakkola 等, 2000). 相较于 MCMC, VB 方法在保持精度的同时显著提升计算效率 (例如在双随机效应维纳模型中效率提升达 87% (Zhou 等, 2020)). 尽管 VB 方法在复杂模型中展现优势, 但在经典随机效应伽马退化模型中仍面临双重挑战:

- 参数可解释性缺失: 传统模型通过尺度参数  $\beta$  引入异质性效应, 但  $\beta$  的统计意义 (影响退化增量均值  $E[Y(t)] = \alpha t / \beta$ ) 与工程认知 (如退化速率、环境应力加速效应) 缺乏直接对应, 导致先验分布设定与后验解释存在障碍.
- 数值稳定性缺陷: 由于伽马分布的共轭结构限制, 传统变分后验需通过数值积分处理随机效应, 这容易出现浮点溢出或精度丢失问题.

针对上述问题, 本节提出重参数化伽马过程 (Reparametrized gamma process, RGa) 及相应的 VB 算法:

- 提出了一种 RGa 模型, 它为模型参数赋予了明确的物理意义, 使得随机效应能够直接且清晰地融入模型, 同时影响退化过程的均值和波动性.
- 结合高斯-赫尔米特 (Gauss-Hermite, GH) 求积法 (Liu 等, 1994) 和拉普拉斯近似 (Geisser 等, 1990), 介绍了一种具有解析形式的 VB 推断方法. 同时还基于 GH 求积法计算的观测信息矩阵, 提出了一种构建区间估计的高效策略, 补充了 VB 在区间估计方面的局限性.

本节结构如下: 第 3.3.1 节将介绍重参数化伽马过程及产品寿命分布的推导. 第 3.3.2 节将提出基于 GH 求积法和拉普拉斯近似的 VB 方法, 并详细介绍区间估计方法. 第 3.3.3 节和第 3.3.4 节将分别进行数值实验和案例研究, 以验证所提方法的有效性.

### 3.3.1 重参数化伽马过程

令  $\{Y(t), t \geq 0\}$  表示系统性能的退化过程. 称  $\{Y(t), t \geq 0\}$  服从 RGa 过程, 若其满足以下性质: (i)  $Y(0) = 0$ ; (ii) 对于任意的  $0 \leq t_i, i = 1, \dots, 4$ ,  $Y(t_1) - Y(t_2)$  与  $Y(t_3) - Y(t_4)$  相互独立; (iii) 每个增量  $\Delta Y(t) = Y(t) - Y(t + \Delta t)$  服从具有形状参数  $\alpha \Delta t$  和尺度参数  $\eta/\alpha$  的伽马分布, 记为  $\text{Ga}(\alpha \Delta t, \eta/\alpha)$ , 其中  $\alpha > 0, \eta > 0$ . 根据伽马分布的可加性, 可以推导出  $Y(t) \sim \text{Ga}(\alpha t, \eta/\alpha)$ , 其 PDF 为

$$p_{Y(t)}(y | \eta, \alpha) = \frac{1}{\Gamma(\alpha t)} \left(\frac{\eta}{\alpha}\right)^{-\alpha t} y^{\alpha t-1} e^{-\frac{\eta y}{\alpha}}. \quad (3.22)$$

可算出  $Y(t)$  的均值、方差以及变异系数分别为  $\eta t, \eta^2 t/\alpha$ , 和  $1/\sqrt{\alpha t}$ . 因此, 参数  $\eta$  表示平均退化速率 (或斜率), 而参数  $\alpha$  则量化总体样本退化过程的波动程度. 经典伽马过程的一个重要特性是其对退化均值和波动性的依赖性 (Tsai 等, 2012). 正如 Ye, Chen, 等 (2015) 所讨论的, 这种依赖关系在建模时有助于提高拟合精度. 从均值和方差可知, 所提出的 RGa 过程保持了这一特性, 但参数的解释性更加清晰.

在退化试验中, 由于材料内在差异性和操作环境的多样性, 同批次产品的性能退化路径可能存在差异. 为体现这种差异性, RGa 模型引入随机效应, 假设退化速率参数  $\eta$  在不同单元间服从对数正态分布, 其 PDF 为

$$p(\eta | \mu, \sigma^2) = \frac{1}{\eta \sigma \sqrt{2\pi}} e^{-\frac{(\ln(\eta) - \mu)^2}{2\sigma^2}}, \quad (3.23)$$

其中位置参数为  $\mu \in (-\infty, \infty)$ , 尺度参数为  $\sigma > 0$ . 参数  $\eta$  的均值和方差分别为  $e^{\mu + \sigma^2/2}$  和  $(e^{\sigma^2} - 1) e^{2\mu + \sigma^2}$ . 当有关于参数  $\eta$  的先验信息时, 超参数  $\mu$  和  $\sigma^2$  可以通过以下方式确定:  $\mu = \mathbb{E}_\eta[\ln(\eta)]$ ,  $\sigma^2 = \mathbb{E}_\eta[(\ln(\eta) - \mu)^2]$  (Krishnamoorthy, 2016). 随机效应采用对数正态分布有如下理由:

- 对数正态分布的定义域为正实数, 适用于单调递增且为正的退化量. 单调递减的退化过程可通过简单的数学变换处理为单调递增的形式.
- 对数正态分布属于一类偏态分布, 具有较高的灵活性. 在实例分析中, 该分布对两组真实数据中随机效应变量  $\eta$  的经验估计具有良好的拟合效果.
- 在变分推断框架下, 假设  $\eta$  服从对数正态分布时, 参数  $(\mu, \sigma^2)$  的联合变分后验分布很容易构造. 这一特性放宽了 VB 推断中参数完全独立的假

设, 允许考虑  $(\mu, \sigma^2)$  之间的相关性. 若  $\eta$  服从其他正实数域分布, 则参数  $(\mu, \sigma^2)$  的联合后验通常难以解析, 增加了推断复杂性.

- 相比经典伽马过程中通过尺度参数引入个体异质性的方式, RGa 模型中的随机效应设置针对退化速率, 更具物理意义.
- 所提模型的退化均值仅依赖于  $\eta$ , 而不受形状参数  $\alpha$  的影响. 在 VB 推断中, 假设  $\alpha$  VB 与  $(\mu, \sigma^2)$  独立有一定的合理性, 间接减少了由近似带来的信息损失.
- 在固定检测频率  $\Delta t$  下, 包含对数正态随机效应的 RGa 模型结构与伽马对数正态模型 (Turlapaty, 2020) 相似, 其在无线通信系统的建模中有广泛应用.

基于带随机效应的 RGa 模型, 可得到  $Y(t)$  的边际 PDF (见定理 3.3), 详细数学推导请参见本节附录 3.3.5.

**定理 3.3.**  $Y(t)$  的边际 PDF 为

$$p_{Y(t)}(y_t | \Theta) = \frac{\alpha^{\alpha t} y_t^{\alpha t - 1}}{\Gamma(\alpha t) \sqrt{\pi}} \int_{-\infty}^{+\infty} q(z, y_t | \Theta) \exp(-z^2) dz, \quad (3.24)$$

其中  $q(z, y_t | \Theta) = \exp \left\{ -\alpha t (\mu + z \sqrt{2\sigma^2}) - \alpha y_t \exp [-(\mu + z \sqrt{2\sigma^2})] \right\}$ .

在式 (3.24) 中, 积分项难以直接解析计算, 因此需要采用数值方法进行近似. 不同与蒙特卡洛积分通过生成大量随机点来估计积分值, GH 求积方法提供了一种高效的近似手段, 其通过选择最优的节点和权重, 能够精确计算特定阶数的多项式积分. 在计算效率和精度方面, GH 求积法都远高于蒙特卡洛积分 (Seo 等, 2017). 利用 GH 求积法, 式 (3.24) 可数值近似为

$$\int_{-\infty}^{+\infty} q(z, y_t | \Theta) \exp(-z^2) dz \approx \sum_{l=1}^L w(z_l) q(z_l, y_t | \Theta), \quad (3.25)$$

其中  $z_l, l = 1, \dots, L$  是  $L$  阶 Hermite 多项式  $H_L(z_l)$  的根 (Davis 等, 2007), 其表达式为

$$H_L(z_l) = L! \sum_{k=1}^{\lfloor \frac{L}{2} \rfloor} \frac{(-1)^k (2z_l)^{L-2k}}{k! (L-2k)!},$$

而权重函数  $w(z_l)$  (称为 GH 因子) 定义为  $w(z_l) = (2^{n-1} n! \sqrt{\pi}) (n^2 (H_{L-1}(z_l)))$ . 图 3.8 展示了基于 GH 近似的 PDF (3.25) 与四种不同情况下的经验 PDF. 结果

表明, GH 方法近似的 PDF 在所有情况下均表现良好.

接下来对产品寿命分布进行分析. 设退化阈值为  $\omega$ , 产品寿命  $T$  的 CDF 表示为

$$F_T(t) = 1 - Pr(Y(t) \leq \omega) = \int_{\omega}^{\infty} p_{Y(t)}(y_t | \Theta) dy_t. \quad (3.26)$$

利用  $p_{Y(t)}(y_t | \Theta)$  的精确近似, 式 (3.26) 可通过梯形积分法计算. 在每个检测点  $\tilde{t}$  处, 定义  $G(y | \tilde{t}, \Theta) = \tilde{p}_{Y(\tilde{t})}(y_{\tilde{t}} | \Theta)$ , 其中  $\tilde{p}_{Y(\tilde{t})}(y_{\tilde{t}} | \Theta)$  是  $p_{Y(t)}(y_t | \Theta)$  的 GH 积分近似. 随后, 使用 R 软件 *pracma* (Borchers, 2019) 包中的 *trapz()* 函数计算积分  $\int_{\omega}^{\infty} G(y | \tilde{t}, \Theta) dy$ , 该方法与梯形规则积分一致. 为验证积分效果, 图 3.9 比较了基于 GH 的积分框架与不同参数  $\Theta$  和退化阈值  $\omega$  下寿命  $T$  的 BS 分布近似. 结果显示, 两种方法几乎无差异.

### 3.3.2 变分贝叶斯推断

在退化试验中, 假设产品的性能特征具有递增的退化路径. 设  $\mathbf{y}_n$  为第  $n$  个样本的观测退化数据, 该样本在检测时间点  $t_{n,m}$  处的观测值  $y_{n,m}$ , 其中  $n = 1, \dots, N, m = 1, \dots, M_n, N$  表示样本量,  $M_n$  为第  $n$  个样本的测量次数. 为简化表示, 定义  $\mathbf{y} = \{\mathbf{y}_n\}_{n=1}^N$ , 其中  $\mathbf{y}_n = \{y_{n,m}\}_{m=1}^{M_n}$ . 假设  $\mathbf{y}$  是来自 RGA 过程  $\{Y(t), t \geq 0\}$  的一组观测值. 定义增量  $x_{n,m} = \Delta y_{n,m} = y_{n,m} - y_{n,m-1}$ , 时间间隔  $h_{nm} = \Delta t_{n,m} = t_{n,m} - t_{n,m-1}$ , 其中  $y_{n,0} = t_{n,0} = 0$ . 记  $\mathbf{x}_n = \{x_{n,m}\}_{m=1}^{M_n}, \mathbf{x} = \{\mathbf{x}_n\}_{n=1}^N$ . 为符号方便, 假设  $M_n = M, h_{n,m} = h_c$ , 即, 试验采用等间隔测量, 且所有样本的测量次数相同. 根据 RGA 过程的性质 (iii), 增量  $\mathbf{x}$  相互独立, 且服从伽马分布. 在给定随机效应  $\boldsymbol{\eta} = \{\eta_n\}_{n=1}^N$  时, 增量  $x_{n,m}$  的条件 PDF 为

$$p_{x|\eta_n}(x_{n,m} | \alpha h_c, \eta_n) = \frac{x_{n,m}^{\alpha h_c - 1}}{\Gamma(\alpha h_c)} \left(\frac{\eta_n}{\alpha}\right)^{-\alpha h_c} e^{-\frac{\alpha x_{n,m}}{\eta_n}}. \quad (3.27)$$

随机效应  $\boldsymbol{\eta}$  为独立同分布的对数正态分布, 其 PDF 为

$$p(\boldsymbol{\eta} | \mu, \sigma^2) = \prod_{n=1}^N p(\eta_n | \mu, \sigma^2) = \prod_{n=1}^N \frac{1}{\eta_n \sigma \sqrt{2\pi}} e^{-\frac{(\ln(\eta_n) - \mu)^2}{2\sigma^2}}. \quad (3.28)$$

令  $\Theta = (\alpha, \mu, \sigma^2)$ , 则  $(\eta, \Theta)$  的联合后验分布为

$$p(\eta, \Theta | \mathbf{x}) \propto \prod_{n=1}^N \left[ p(\eta_n | \mu, \sigma^2) \prod_{m=1}^M p_{x|\eta}(x_{n,m} | \alpha h_c, \eta_n) \right] p(\alpha) p(\mu, \sigma^2), \quad (3.29)$$

其中  $p(\alpha)$  和  $p(\mu, \sigma^2)$  分别是参数  $\alpha$  和  $(\mu, \sigma^2)$  的先验分布. 式 (3.29) 中的先验设置隐含了以下假设: (i) 允许参数  $\mu$  和  $\sigma^2$  存在一定的相关性, 可选用二元分布作为其先验; (ii) 假设  $\alpha$  与  $(\mu, \sigma^2)$  相互独立. 这一假设基于上一节中关于随机效应设置的讨论, 认为  $\alpha$  主要描述形状参数的变化, 与刻画单元间异质性的参数  $(\mu, \sigma^2)$  无关.

### 3.3.2.1 变分贝叶斯估计

由于式 (3.29) 中参数的后验分布非常复杂, 直接进行统计推断较为困难, VB 方法通过引入一组可处理的变分后验密度 (Variational posterior density, VPD)  $q(\Theta, \eta)$  来近似真实的后验密度  $p(\eta, \Theta | \mathbf{x})$ , 即  $p(\eta, \Theta | \mathbf{x}) \approx q(\eta, \Theta)$ , 该近似可通过求解以下约束优化问题确定.

$$q^*(\Theta, \eta) = \arg \max_{q(\Theta, \eta) \in \mathcal{Q}} -F(q(\Theta, \eta)), \text{ s.t. } \int q(\Theta, \eta) d\Theta d\eta = 1, \quad (3.30)$$

其中  $F(q(\eta, \Theta)) = E_{\Theta, \eta} \log \left( \frac{q(\eta, \Theta)}{p(\eta, \Theta | \mathbf{x})} \right)$ , 即,  $q(\eta, \Theta)$  和  $p(\eta, \Theta | \mathbf{x})$  的 Kullback-Liebler(KL) 散度.  $\mathcal{Q}$  表示所有潜在 VPD 的集合. 通常而言, 集合  $\mathcal{Q}$  的选择直接决定了 VPD 的复杂性, 从而影响优化问题 (3.30) 的求解难度, 并进一步影响 VB 推断的计算效率. 因此, 合理选择  $\mathcal{Q}$  至关重要, 它应足够简单, 以确保优化问题 (3.30) 可行, 同时又要具备足够的灵活性, 以准确逼近真实的后验分布. 一种常用的选择是均值场变分族 (Mean-field variational family, MFVF), 该方法假设模型参数之间相互独立, 或者在一定条件下表现出独立性. 在 MFVF 下, 可以进一步假设:

$$q(\Theta, \eta) = \prod_{n=1}^N q(\eta_n) q(\alpha) q(\mu, \sigma^2). \quad (3.31)$$

其中  $q(\eta_n)$ 、 $q(\alpha)$  和  $q(\mu, \sigma^2)$  分别是  $\eta_n$ 、 $\alpha$  和  $(\mu, \sigma^2)$  的 VPD. 该分解形式在降低计算复杂度的同时, 仍能有效逼近真实的后验分布. 将 MFVF (3.31) 代入优

化目标 (3.30) 中, 然后使用拉格朗日乘数法求解式(3.30), 得到

$$F_1(q(\boldsymbol{\Theta}, \boldsymbol{\eta})) = F(q(\boldsymbol{\Theta}, \boldsymbol{\eta})) + \boldsymbol{\lambda}^T \left( \int (q(\boldsymbol{\Theta}, \boldsymbol{\eta}) - 1) d\boldsymbol{\Theta} d\boldsymbol{\eta} \right), \quad (3.32)$$

其中向量  $\boldsymbol{\lambda} = (\{\lambda_{\eta_n}\}_{n=1}^N, \lambda_\alpha, \lambda_\mu, \lambda_{\sigma^2})$  是拉格朗日乘数, 表示  $q(\boldsymbol{\Theta}, \boldsymbol{\eta})$  对目标函数的边际效应.

VB 算法可视为传统 EM 方法和贝叶斯推断在计算时间和计算效率上的一种折中, 通过使用易处理的代理分布来近似复杂的贝叶斯后验. 它利用 KL 散度将贝叶斯推断转化为数值优化问题, 从随机模拟转向确定性优化. 如 Blei 等 (2017) 所述, VB 的优化分为两个迭代步骤: 针对潜在变量的 VB 后验更新 (变分贝叶斯期望 (VB-E) 步骤) 和针对目标参数的 VB 后验更新 (变分贝叶斯最大化 (VB-M) 步骤), 与 EM 算法的迭代过程相似.

### 3.3.2.1.1 VB-E: 最优潜变量

对式 (3.32) 中关于  $q(\boldsymbol{\eta})$  求导, 可以得到  $\boldsymbol{\eta}$  的 VPD 的对数形式

$$\ln(q(\boldsymbol{\eta})) = \sum_{n=1}^N \mathbb{E}_{\mu, \sigma^2} [\ln(p_\eta(\eta_n | \mu, \sigma^2))] + \sum_{n=1}^N \sum_{m=1}^M \mathbb{E}_\alpha [p_{x|\eta}(x_{n,m} | \alpha h_c, \eta_n)] + C,$$

其中  $C$  表示常数. 为简便符号, 本节后续内容涉及到常数都用  $C$  来表示, 这不影响各参数的 VPD 形式. 由独立性假设,  $\eta_n$  的 VPD 的对数形式为

$$\ln(q(\eta_n)) = \mathbb{E}_{\mu, \sigma^2} [\ln(p_\eta(\eta_n | \mu, \sigma^2))] + \sum_{m=1}^M \mathbb{E}_{q(\alpha)} [p_{x|\eta}(x_{n,m} | \alpha h_c, \eta_n)] + C.$$

利用式 (3.27) 和 (3.28), 可以得到

$$\ln(q(\eta_n)) = (w_1 w_2 - w_0 M h_c - 1) \ln(\eta_n) - \frac{y_{n,M} w_0}{\eta_n} - w_2 (\ln(\eta_n))^2 / 2 + C, \quad (3.33)$$

其中  $y_{n,M} = \sum_{m=1}^M x_{n,m}$ ,  $w_0 = \mathbb{E}_\alpha [\alpha]$ ,  $w_1 = \mathbb{E}_{\mu, \sigma^2} [\mu]$ , 以及  $w_2 = \mathbb{E}_{\mu, \sigma^2} [\sigma^{-2}]$ . 则  $\eta_n$  的 VPD 可表示为  $q(\eta_n) = g_0(\eta_n) / A_{\eta_n}$ , 其中

$$g_0(\eta_n) = \eta_n^{w_1 w_2 - w_0 M h_c - 1} e^{-\frac{y_{n,M} w_0}{\eta_n} - \frac{w_2 (\ln(\eta_n))^2}{2}},$$

$A_{\eta_n} = \int_0^{+\infty} g_0(\eta_n) d\eta_n$  为正则化常数. 令  $s = \ln(\eta_n) \sqrt{w_2/2}$ . 将  $s$  代入  $g_0(\eta_n)$  得到

$$g_0(\eta_n) = q(s, y_{n,M} \mid w_0, w_1, w_2) e^{-s^2}, \quad (3.34)$$

其中

$$q(s, y_{n,M} \mid w_0, w_1, w_2) = \sqrt{2/w_2} e^{s\sqrt{2/w_2}(w_1 w_2 - w_0 M h_c) - e^{-y_{n,M} w_0 s \sqrt{2/w_2}}}.$$

注意到  $g_0(\eta_n)$  与式 (3.25) 中的被积函数具有相同的结构. 因此,  $A_{\eta_n}$  也可以通过第 3.3.1 节中介绍的 GH 求积方法来计算. 基于式 (3.33) 和 (3.34), 任何连续函数  $f(\eta_n)$  的期望可以表示为

$$\begin{aligned} \mathbb{E}_{\eta_n}(f(\eta_n)) &= \int_0^{+\infty} f(\eta_n) q(\eta_n) d\eta_n \\ &= \frac{1}{A_{\eta_n}} \int_{-\infty}^{+\infty} f(\exp(s\sqrt{2/w_2})) e^{-s^2} q(s, y_{n,M} \mid w_0, w_1, w_2) ds, \end{aligned} \quad (3.35)$$

这一积分也可通过 GH 求积法进行数值计算.

### 3.3.2.1.2 VB-M: 最优模型参数

对式 (3.32) 中关于  $q(\Theta)$  求导, 可以得到  $\Theta$  的 VPD 的对数形式为

$$\begin{aligned} \ln(q(\alpha, \mu, \sigma^2)) &= \sum_{n=1}^N \left[ E_{\eta_n} [\ln(p(\eta_n \mid \mu, \sigma^2))] \right] \\ &\quad + \sum_{n=1}^N \sum_{m=1}^M \left[ E_{q(\eta_n)} [\ln(p(x_{n,m} \mid \eta_n, \alpha))] \right] \\ &\quad + \ln(p(\alpha)) + \ln(p(\mu, \sigma^2)) + C. \end{aligned}$$

可发现  $\alpha$  和  $(\mu, \sigma^2)$  是相互独立的. 因此, 给定  $q(\eta_n), n = 1, 2, \dots, N$  后, 可以依次推导出  $q(\alpha)$  和  $q(\mu, \sigma^2)$  的 VPD 表达式.

i)  $\alpha$  的 VPD:  $\alpha$  的 VPD 的对数形式为

$$\ln(q(\alpha)) = \sum_{n=1}^N \sum_{m=1}^M \left[ E_{q(\eta_n)} [\ln(p(x_{n,m} | \eta_n, \alpha))] \right] + \ln(p(\alpha)) + C. \quad (3.36)$$

再将 (3.27) 和 (3.28) 代入 (3.36), 得到

$$\begin{aligned} \ln(q(\alpha)) = & - (h_c(Mw_3 - w_5) + w_4)\alpha + \alpha h_c NM \ln(\alpha) \\ & - NM \ln(\Gamma(\alpha h_c)) + \ln(p(\alpha)) + C, \end{aligned}$$

其中

$$w_3 = \sum_{n=1}^N E_{\eta_n} [\ln(\eta_n)], w_4 = \sum_{n=1}^N y_{nM} E_{\eta_n} [\eta_n^{-1}], w_5 = \sum_{n=1}^N \sum_{m=1}^M \log(x_{nm}),$$

并且  $E_{\eta_n} [\ln(\eta_n)]$  和  $E_{\eta_n} [\eta_n^{-1}]$  可以通过式 (3.35) 更新. 对 (3.37) 两边取指数变换, 可得到

$$q(\alpha) \propto p(\alpha) \frac{\alpha^{a\alpha}}{\Gamma(\alpha h_c)^b} \exp(-c\alpha), \quad (3.37)$$

其中  $a = h_c NM$ ,  $b = NM$ ,  $c = h_c(Mw_3 - w_5) + w_4$ . 为方便数学推导, 使用共轭先验可简化计算 (Gelman 等, 2014). 可验证: 当先验  $p(\alpha)$  取以下形式时, 其与  $q(\alpha)$  属于相同的分布族:

$$p(\alpha) \propto \frac{\alpha^{a_0}}{\Gamma(\alpha h_c)^{b_0}} \exp(-c_0\alpha), \quad (3.38)$$

其中  $a_0$ 、 $b_0$  和  $c_0$  是先验  $p(\alpha)$  的超参数. 将 (3.38) 代入 (3.37) 并计算正则化常数, 可得

$$q(\alpha) = g_\alpha(\alpha)/A_\alpha \quad (3.39)$$

其中

$$g_\alpha(\alpha) = \frac{\alpha^{a_1\alpha}}{\Gamma(\alpha h_c)^{b_1}} \exp(-c_1\alpha), \quad A_\alpha = \int_{\Omega_\alpha} g_\alpha(\alpha) d\alpha,$$

$a_1 = a_0 + a$ ,  $b_1 = b_0 + b$ ,  $c_1 = c_0 + c$ . 然而, 由于  $g_\alpha(\alpha)$  的积分可能在数值上不稳定, 通过 (3.39) 计算  $\alpha$  的 VPD 的均值和方差比较困难. 作为替

代方案, 可对 (3.39) 进行拉普拉斯近似. 此时,  $\alpha$  的 VPD 的均值和方差可近似为

$$m(\alpha) \approx \arg \min_{\alpha} \{a_1 \log(\alpha) - b_1 h_c \psi(\alpha h_c) + a_1 - c_1\},$$

$$v(\alpha) \approx \frac{1}{b h_c^2 \psi'(m(\alpha) h_c) - \frac{a_1}{m(\alpha)}}.$$

ii) 参数  $(\mu, \sigma^2)$  的联合 VPD: 基于 (3.36),  $(\mu, \sigma^2)$  的联合 VPD 的对数形式为

$$\ln(q(\mu, \sigma^2)) = \sum_{n=1}^N \left[ E_{q(\eta_n)} [\ln(p(\eta_n | \mu, \sigma^2))] \right] + \ln(p(\mu, \sigma^2)) + C. \quad (3.40)$$

将 (3.27) 和 (3.28) 代入 (3.40) 并对结果取指数, 可得

$$q(\mu, \sigma^2) \propto p(\mu, \sigma^2) (\sigma^2)^{-N/2} e^{-\frac{w_6 - w_3^2/N}{2\sigma^2} - \frac{N(\mu - w_3/N)^2}{2\sigma^2}}, \quad (3.41)$$

其中

$$w_6 = \sum_{n=1}^N E_{q(\eta_n)} [(\ln(\eta_n))^2], E_{\eta_n} [(\ln(\eta_n))^2]$$

可以通过式 (3.35) 计算. 若  $p(\mu, \sigma^2)$  服从正态逆伽马 (Normal inverse gamma, NIGa) 分布, 式 (3.41) 恰好是 NIGa 密度函数的核 (Bolstad 等, 2016). 通常, 随机变量  $(X, Y)$  服从  $\text{NIGa}((x, y); \rho, \varsigma, \nu, \xi)$ , 其 PDF 为:

$$p(x, y) = \frac{\sqrt{\varsigma}}{\sqrt{y} \sqrt{2\pi}} \frac{\xi^\nu}{\Gamma(\nu)} y^{-\nu-1} e^{-\frac{2\xi + \varsigma(x-\rho)^2}{2y}}, \quad (3.42)$$

其中  $x \in (-\infty, +\infty)$ ,  $y \in (0, +\infty)$ ,  $\nu > 0$ ,  $\xi > 0$ ,  $\varsigma > 0$ , 以及  $\rho \in (-\infty, +\infty)$  是参数. 根据 NIGa 分布的定义 (Denison 等, 2002), (3.42) 可分解为在给定  $\sigma^2$  条件下,  $\mu$  的均值为  $\rho$ 、方差为  $\sigma^2/\varsigma$  的条件正态分布和形状参数为  $\nu$ 、尺度参数为  $\xi$  的  $\sigma^2$  的逆伽马分布. 因此, 可假设参数  $(\mu, \sigma^2)$  的联合先验为

$$p(\mu, \sigma^2) = \text{NIGa}((\mu, \sigma^2); \rho_0, \varsigma_0, \nu_0, \xi_0). \quad (3.43)$$

将式 (3.43) 代入 (3.41) 后, 可得到  $(\mu, \sigma^2)$  的联合 VPD 为

$$q(\mu, \sigma^2) = \text{NIGa}((\mu, \sigma^2); \rho_1, \varsigma_1, \nu_1, \xi_1), \quad (3.44)$$

其中

$$\begin{aligned} \rho_1 &= \frac{1}{\varsigma_1} [\varsigma_0 \rho_0 + w_3], \varsigma_1 = \varsigma_0 + N, \nu_1 = \nu_0 + \frac{N}{2}, \\ \xi_1 &= \xi_0 + \frac{1}{2} [\varsigma_0 \rho_0^2 + w_6 - \varsigma_1 \rho_1^2]. \end{aligned}$$

根据 NIGa 分布的性质 (Bolstad 等, 2016),  $\mu$  和  $\sigma^2$  的 VPD 均值和方差为

$$\begin{aligned} m(\mu) &= \rho_1, \quad v(\mu) = \frac{\xi_1}{\varsigma_1(\nu_1 - 1)}, \\ m(\sigma^2) &= \frac{\xi_1}{\nu_1 - 1}, \quad v(\sigma^2) = \frac{\xi_1^2}{(\nu_1 - 1)^2(\nu_1 - 2)}, \end{aligned}$$

并且  $\sigma^2$  的一阶逆矩等于  $\frac{\nu_1}{\xi_1}$ .

通过当前模型的 VB-E 和 VB-M 步的更新方程, 可对这两步进行迭代, 直到估计方差不再显著改善, 即

$$\varepsilon(\Theta^{(s)}) = \|v(\Theta^{(s)}) - v(\Theta^{(s-1)})\|_1 \leq \varepsilon_0, \quad (3.45)$$

其中  $\|\cdot\|_1$  为绝对值范数,  $\varepsilon_0$  是确定所需精度的预设值.

**注 5:** VB 算法是否成功与高效, 关键在于均值场变分族的假设和模型参数的共轭特性. 均值场变分族假设使得能够基于 KL 散度和拉格朗日优化算法求得近似贝叶斯后验分布. 共轭参数特性则提供了解析形式的近似贝叶斯后验. 对于非共轭模型参数, 也可以基于自然共轭先验通过拉普拉斯近似得到相应的变分后验. 这不仅减少了数值优化需求, 提高了计算效率, 而且在 VB 和贝叶斯后验中保持了结构的相似性, 从而提升了精度. 然而, 当模型参数如  $\eta$  服从非共轭偏态分布时, 非共轭参数的增加则会降低模型的计算效率和精度.

### 3.3.2.2 初始值确定

在 VB 算法中, 参数初始值的选择会影响算法的收敛时间及其局部最优解. 在本节中, 参数  $\Theta$  的初始值将通过以下步骤确定:

1. 对于第  $n$  个样品, 分别计算增量  $\mathbf{X}_n$  的均值和方差为  $\bar{x}_n = \frac{1}{M} \sum_{m=1}^M x_{n,m}$  和  $V_n = \frac{1}{M-1} \sum_{m=1}^M (x_{n,m} - \bar{x}_n)^2$ .
2. 由于增量  $\mathbf{X}_n$  服从  $Ga(\alpha h_c, \eta_n/\alpha)$ , 其均值为  $\eta_n h_c$ , 方差为  $\eta_n^2 h_c/\alpha$ , 因此可估计  $\eta_n$  和  $\alpha$  为  $\bar{x}_n/h_c$  和  $\bar{x}_n^2/(V_n h_c)$ .
3. 由  $\{\eta_n\}_{n=1}^N$ , 参数  $\mu$  和  $\sigma^2$  的初始值可以通过以下公式确定:  $\bar{\eta}_n = \frac{1}{N} \sum_{n=1}^N \ln(\eta_n)$  和  $V_{\ln(\eta_n)} = \frac{1}{N-1} \sum_{n=1}^N (\ln(\eta_n) - \bar{\eta}_n)^2$ .

### 3.3.2.3 置信区间

通过上述 VB 推断方法, 可得到每个  $\theta_i \in \Theta$  的 VPD 均值  $m(\theta_i)$  和方差  $v(\theta_i)$ . 然后构造  $\theta_i$  的  $100(1 - \gamma)$  近似区间估计为

$$m(\theta_i) \pm Z_{1-\gamma/2} \sqrt{v(\theta_i)}, \quad (3.46)$$

其中  $Z_\gamma$  为标准正态分布的  $\gamma$  分位数. 然而, 式 (3.46) 存在不足, 因为 VB 方法通常会低估后验分布的散度, 从而导致方差估计值比真实值小 (Blei 等, 2017). 为了解决这个问题, Wang 等 (2005) 对 VB 推断方法下的区间估计进行改进, 其中 VPD 的方差可通过原始似然函数下的 Fisher 信息矩阵计算. 当 Fisher 信息矩阵计算困难时, 则可进一步用观测 Fisher 信息矩阵替代 (Meeker 等, 1998). 基于此思想, 当前模型的观测 Fisher 信息矩阵为

$$\hat{I}_{\Theta} = -\frac{\partial^2 \log(\ell(\Theta))}{\partial \Theta^T \partial \Theta}, \quad (3.47)$$

其中

$$\begin{aligned} \ell(\Theta) = & -NM \log(\Gamma(\alpha h_c)) - NM \alpha h_c \log(\alpha) \\ & + \sum_{n=1}^N \sum_{m=1}^M (\alpha h_c - 1) \log(x_{n,m}) - \frac{N}{2} \log(\pi) \\ & + \sum_{n=1}^N \log \left[ \int_{-\infty}^{\infty} q(z, y_{n,M}, \Theta) \exp(-z^2) dz \right], \end{aligned}$$

其中  $z = \frac{\log(\eta_n - \mu)}{\sqrt{2\sigma^2}}$ ,

$$q(z, y_{n,M}, \Theta) = e^{-\alpha h_c M(\sqrt{2\sigma^2}z + \mu) - \alpha y_{n,M} e^{-\sqrt{2\sigma^2}z - \mu}}, \quad (3.48)$$

详细导数可见本节附录 3.3.5. 本节附录 3.3.5 还给出其他两种基准估计方法, 分别为 EM 算法和贝叶斯方法. 后续会在数值模拟和案例研究中与所提 VB 推断方法进行比较.

### 3.3.3 模拟实验

模拟实验将验证 VB 算法的计算精度, 同时比较 VB、EM 和 MCMC 三种算法的计算效率. 三种方法的基本设置如下:

1. VB 算法: 先验分布  $p(\alpha)$  和  $p(\mu, \sigma^2)$  的超参数使用弱信息先验设定, 具体为  $a_0 = b_0 = c_0 = 1, \xi_0 = \varsigma_0 = \rho_0 = 1, \nu_0 = 0$ , 以便与 EM 算法进行公平比较.
2. EM 算法: 采用最近两次迭代估计值的绝对值范数作为收敛停止准则.
3. MCMC 算法: 使用与 VB 相同的超参数设定, 后验采样中包含 3 条马尔科夫链, 每条链迭代 20,000 次, 前半部分作为燃烧期.
4. 通用设置: VB 和 EM 算法的目标精度为  $10^{-6}$ . 所有算法均通过 R 软件实现, 运行于 Windows 10 桌面电脑 (3.5 GHz Intel Xeon E3-1241 v3 处理器, 16 GB 内存).

样本量  $N$  选为 15、30 和 60, 对于每个  $N$ , 测量次数  $M = 10$  和 20. 对于每种  $(N, M)$  组合, 从参数  $\Theta = (\alpha, \mu, \sigma^2) = (10, -1, 0.01)$  的 RGa 过程中随机生成 2,000 个退化样本  $y$ . 对每个样本  $y$ , 通过 VB、EM 和 MCMC 算法获得  $\Theta$  的估计值. 因此, 每种算法均得到 2,000 个估计值. 基于这些估计值, 计算平均值和 MSE. 结果列在表 3.6 中. 同时还计算了每种算法的平均 CPU 运行时间 (秒,  $\bar{t}$ ), 结果如图 3.10 所示. 此外, 还计算了 95% 可信区间的覆盖率和区间长度 (Interval length, IL), 见表 3.7. 从这两个表格和三种算法的平均 CPU 运行时间图中, 可以得出以下结果: (1) 所有情况下, 估计值的平均值接近真实值, 且样本量增加时, 平均值与真实值的差异逐渐缩小. (2) 随着样本量增加, 各模拟设置下的 MSE 总体呈下降趋势. VB 和 EM 的 MSE 小于贝叶斯方法, 且样本量增加时, 三者差距逐渐减小. (3) 所有算法的 IL 随样本量增加而缩小, 同时基于

表 3.6: 不同算法和样本量下, 估计值的平均值和均方误差.

$(N, M)$	方法	$\alpha = 10$		$\mu = -1$		$\sigma^2 = 0.01$	
		MSE	均值	MSE ( $\times 10^{-3}$ )	均值	MSE ( $\times 10^{-5}$ )	均值
$(15, 10)$	VB	3.5689	10.2939	2.1490	-0.9921	12.7358	0.0118
	EM	3.7443	10.4057	2.4983	-0.9990	13.8214	0.0118
	Bayes	1.2705	10.2116	1.2181	-1.0007	9.3512	0.0179
$(15, 20)$	VB	3.0234	10.3048	1.3635	-0.9973	8.9276	0.0118
	EM	3.4744	10.3870	1.3836	-1.0011	9.3441	0.0114
	Bayes	0.7495	10.1547	1.1617	-1.0011	5.1783	0.0116
$(30, 10)$	VB	2.4774	10.2291	1.1519	-0.9990	4.1917	0.0097
	EM	3.4369	10.3056	1.1763	-0.9986	3.0103	0.0108
	Bayes	0.6956	10.2142	0.8215	-1.0083	1.8619	0.0103
$(30, 20)$	VB	2.2640	10.1209	0.8165	-0.9974	2.4172	0.0097
	EM	2.6755	10.1871	0.8234	-0.9988	2.6170	0.0104
	Bayes	0.3813	10.0223	0.5394	-1.0001	1.1418	0.0102
$(30, 30)$	VB	2.1271	10.0525	0.6582	-1.0154	1.9043	0.0104
	EM	2.042	10.0147	0.4325	-1.0137	1.2131	0.0107
	Bayes	0.2032	10.0125	0.3142	-1.0015	0.9324	0.0102
$(60, 20)$	VB	2.0017	10.093	0.3627	-1.0097	1.0121	0.0098
	EM	2.0014	10.0104	0.3674	-1.0117	1.1374	0.0103
	Bayes	0.1747	10.0041	0.2873	-1.0009	0.7213	0.0101

VB 与贝叶斯方法的 IL 差异也逐渐减小. (4) 在覆盖率方面, VB 略优于 EM, 表现出更接近 0.95 的名义水平, 而贝叶斯方法在小样本 (如  $N = 15, M = 10, 20$ ) 时表现最佳. (5) 在计算效率上, VB 显著优于 EM 和贝叶斯方法, 部分情况下运行时间仅为贝叶斯方法的百分之一或千分之一. 总体而言, 贝叶斯方法提供最高的估计精度, 但计算效率低; 三种算法的估计精度随样本量增加而提高; EM 提高了效率, 但牺牲了部分精度; VB 在精度与效率之间实现了良好平衡, 是贝叶斯与 EM 方法的折中方案.

表 3.7: 在不同算法和样本量下, 区间估计的 95% 覆盖概率和区间长度.

(N,M)	方法	$\alpha = 10$		$\mu = -1$		$\sigma^2 = 0.01$	
		CP	IL	CP	IL	CP	IL
(15,10)	VB	0.9123	3.9051	0.9234	0.7675	0.9014	0.7234
	EM	0.9136	4.5865	0.9216	0.7921	0.8945	0.8653
	Bayes	0.9208	3.8839	0.9307	0.1688	0.9134	0.5053
(15,20)	VB	0.9293	2.7575	0.9356	0.6060	0.9345	0.4209
	EM	0.9294	3.2482	0.9341	0.7459	0.9301	0.5678
	Bayes	0.9406	2.7325	0.9406	0.1471	0.9406	0.3847
(30,10)	VB	0.9323	3.8864	0.9245	0.5655	0.9471	0.4515
	EM	0.9289	3.9691	0.9223	0.7319	0.9412	0.5317
	Bayes	0.9501	3.3626	0.9307	0.1092	0.9505	0.2628
(30,20)	VB	0.9457	2.7569	0.9451	0.4971	0.9487	0.3124
	EM	0.9470	2.7751	0.9412	0.5777	0.9465	0.4018
	Bayes	0.9503	2.2675	0.9504	0.0947	0.9507	0.1987
(60,10)	VB	0.9473	2.6321	0.9462	0.4737	0.9503	0.2132
	EM	0.9469	2.7747	0.9451	0.5153	0.9495	0.3314
	Bayes	0.9505	1.8278	0.9507	0.0879	0.9510	0.1637
(60,20)	VB	0.9489	2.4143	0.9475	0.4687	0.9504	0.2047
	EM	0.9475	2.6032	0.9457	0.5096	0.9502	0.3176
	Bayes	0.9509	1.8021	0.9511	0.0813	0.9513	0.1521

### 3.3.4 实例分析

#### 3.3.4.1 激光退化数据

以激光退化数据为例 (见图 1.2) 进行分析. 拟合前需验证两个目标: 一是检查退化率的单元间变异性是否符合对数正态分布, 二是研究激光数据中 RGa 过程的形状参数与尺度参数的相关性. 为验证第一个目标, 计算经验估计量  $\hat{\eta} = \{\hat{\eta}_n = y_{nM}/t_{nM}\}_{n=1}^{15}$ , 并利用式 (3.23) 估计对数正态分布的参数, 得到  $\hat{\eta} \sim LN(-0.697, 0.045)$ . 图 3.11(a) 展示了带 95% CI 的经验分布函数和估计的对数正态分布, 表明退化率  $\hat{\eta}$  符合对数正态分布的特征. 此外, 通过 Anderson-Darling 检验得到的 P 值为 0.633, 进一步说明使用对数正态分布的合理性. 在研究 RGa 过程形状参数与尺度参数的相关性时, 定义经验估计量为

$$\hat{\alpha}, \hat{\beta} = \left\{ \hat{\alpha}_n = \frac{\hat{\eta}_n^2}{v(\mathbf{x}_n / \sqrt{h_c})}, \hat{\beta}_n = \frac{\hat{\eta}_n}{\hat{\alpha}_n} \right\}_{n=1}^{15}.$$

图 3.11(b) 展示了  $\hat{\alpha}^{-1}$  和  $\hat{\beta}$  的散点分布情况以及回归曲线, 相关系数  $\rho = 0.958$  表明两者之间存在显著正相关性.

接下来, 采用 VB 算法对 RGa 退化模型进行拟合, 并将其应用于激光数据集, 同时并与 EM 和贝叶斯方法进行比较, 以衡量计算效率和拟合精度. 表 3.8 展示了各算法的参数估计值、95% 区间估计以及运行时间. 结果显示, 不同算法的参数估计值相近, 但 VB 算法在 CPU 运行时间上显著优于 EM 算法, 且远低于贝叶斯算法. 图 3.12(a) 显示了样本平均值、估计的平均退化路径 ( $\hat{\eta}t$ ) 以及各算法的 95% 区间估计, 表明所有算法均能有效地拟合样本平均值.

本文提出的模型还与五个基准模型进行了比较, 包括伽马过程、逆高斯过程、维纳过程、具有随机效应的同质伽马过程和 Tsai 等 (2012) 所提的 RGa 过程. 为简化表述, 将本文提出的随机效应伽马过程模型记为 RGaLn, Tsai 等 (2012) 的模型记为 GaGa. 鉴于 EM 和 VB 算法在估计结果上无显著差异, 模型选择基于 EM 算法的结果. 同时 VB 算法作为一种确定性近似方法, 不适用于计算涉及有效样本数量的 BIC, 因此使用 AIC 准则进行评估. 表 3.9 列出了各模型的 AIC 值, 显示随机效应模型 (GaGa 和 RGaLn) 的 AIC 值低于其他模型, 表明在激光退化数据中考虑随机效应是合理的. 然而, RGaLn 和 GaGa 的 AIC 值差异不显著. 图 3.12(b) 展示了经验分布函数, 以及阈值  $\omega = 6$  下基于 GaGa

表 3.8: 激光数据在不同方法下的参数估计值、95% 置信区间及计算时间.

方法	$\alpha$		$\mu$		$\sigma^2$		耗时
	估计值	95% CIs	估计值	95% CIs	估计值	95% CIs	
VB	9.682	(8.161,11.792)	-0.696	(-0.854, -0.549)	0.059	(0.019, 0.141)	0.1842
EM	9.713	(8.159,11.802)	-0.696	(-0.856, -0.550)	0.058	(0.019, 0.143)	0.7451
Bayes	9.699	(8.478,11.702)	-0.694	(-0.820, -0.561)	0.057	(0.021, 0.138)	12.381

表 3.9: 激光退化数据基于不同模型的对数似然值和 AIC 值.

模型	对数似然值	AIC 值
RGaLn	93.291	-180.581
GaGa	93.738	-181.476
RGa	69.609	-135.219
伽玛	69.609	-135.219
逆高斯	73.031	-148.062
维纳	45.563	-87.126

和 RGaLn 模型的产品寿命分布. 经验分布函数通过中位数秩估计对伪失效时间进行非参数估计 (Meeker 等, 1998). 可以看出, GaGa 和 RGaLn 模型均能很好地拟合经验分布函数.

### 3.3.4.2 设备-B 数据

这里以 7 个设备-B 的退化数据为例 (见图 1.6 中 150°C 条件下的 7 个设备-B 样本) 进行分析. 图 3.13(a) 显示了退化率  $\hat{\eta}$  的经验分布函数及其 95% 置信区间, 以及拟合的对数正态分布. 结果表明, 对数正态分布能够很好地拟合退化率分布. 同时 Anderson-Darling 检验的  $P$  值为 0.883, 进一步验证用对数正态分布刻画样品间变异性的合理性. 图 3.13(b) 显示了形状参数与尺度参数倒数的回归线及其相关系数  $\rho$ , 结果表明两者之间存在显著的正相关关系.

随后采用 VB、EM 和贝叶斯方法对设备-B 数据进行拟合, 并使用所提 RGa 退化模型进行分析. 拟合结果汇总于表 3.10. 从表中可以看出, VB 方法在计算效率上显著优于其他方法, 同时在精度方面表现良好. 图 3.14(a) 展示了样本平均值及每种算法的估计平均退化路径 ( $\hat{\eta}t$ ). 可以看出, 各算法的估计

表 3.10: 设备-B 数据下, 不同方法的模型参数估计值、95% 置信区间和计算时间.

方法	$\alpha$		$\mu$		$\sigma^2$		耗时
	估计值	95% CI	估计值	95%CI	估计值	95%CI	
VB	7.681	(6.381, 9.171)	-4.384	(-4.653, -4.119)	0.112	(0.021, 0.195)	0.643
EM	7.679	(6.380, 9.162)	-4.384	(-4.653, -4.117)	0.110	(0.023, 0.97)	1.423
Bayes	7.689	(6.376, 9.207)	-4.385	(-4.655, -4.120)	0.116	(0.019, 0.201)	20.506

表 3.11: 设备-B 数据基于不同模型的对数似然值和 AIC 值.

模型	对数似然值	AIC 值
RGaLn	893.009	-1780.018
GaGa	888.156	-1770.312
RGa	876.595	-1749.190
伽马	876.595	-1749.190
逆高斯	851.529	-1699.058
维纳	881.747	-1759.494

平均退化路径均能很好地拟合样本平均值.

接下来, 将设备-B 数据分别用五种基准模型拟合, 其对数似然值和 AIC 值列于表 3.11. 结果显示, 提出的 RGaLn 模型在所有候选模型中具有最小的 AIC 值, 可认为是最优模型. 图 3.14(b) 中展示了经验分布函数以及退化阈值  $\omega = 0.3$  下基于 GaGa 和 RGaLn 模型估计的寿命分布. 结果表明, RGaLn 模型的拟合效果显著优于 GaGa 模型.

### 3.3.5 附录

#### 定理 3.3 证明

根据式(3.22)和(3.23),  $Y(t)$  和  $\eta$  的联合 PDF 可表示为

$$p_{Y(t),\eta}(y_t, \eta \mid \Theta) = p_{Y(t)}(y_t \mid \eta, \alpha)p(\eta \mid \mu, \sigma^2), \quad (3.49)$$

其中  $\Theta = (\alpha, \mu, \sigma^2)$ . 接下来, 对式 (3.49) 关于  $\eta$  求积分, 可得到  $Y(t)$  的边际 PDF:

$$p_{Y(t)}(y_t \mid \Theta) = \frac{\alpha^{\alpha t} y_t^{\alpha t - 1}}{\Gamma(\alpha t) \sqrt{2\pi\sigma}} \int_0^{+\infty} \eta^{-\alpha t - 1} e^{-\frac{\alpha y_t}{\eta} - \frac{(\ln(\eta) - \mu)^2}{2\sigma^2}} d\eta.$$

由于积分函数的复杂性和积分区域的约束, 这一积分的求解可能会导致数值不稳定的结果. 可通过重参数化或变量变换  $z = (\log(\eta) - \mu)/\sqrt{2\sigma^2}$ , 从而得到

$$p_{Y(t)}(y_t \mid \Theta) = \frac{\alpha^{\alpha t} y_t^{\alpha t - 1}}{\Gamma(\alpha t) \sqrt{\pi}} \int_{-\infty}^{+\infty} q(z, y_t \mid \Theta) e^{-z^2} dz, \quad (3.50)$$

其中

$$q(z, y_t \mid \Theta) = \exp \left[ -\alpha t(\mu + z\sqrt{2\sigma^2}) - \alpha y_t e^{-(\mu + z\sqrt{2\sigma^2})} \right].$$

#### 观测 Fisher 信息矩阵

基于对数似然函数, 对  $\ell(\Theta)$  关于  $\Theta$  中每个参数求一阶导数可得

$$\begin{aligned} \frac{\partial \ell(\Theta)}{\partial \alpha} &= -NMh_c(\psi(\alpha h_c) + \log(\alpha) + 1) + \sum_{n=1}^N \sum_{m=1}^M h_c \log(x_{n,m}) + \sum_{n=1}^N A_n(\alpha), \\ \frac{\partial \ell(\Theta)}{\partial \theta_d} &= \sum_{n=1}^N A_n(\theta_d), \end{aligned}$$

其中

$$A_n(s) = \frac{\int_{-\infty}^{\infty} \frac{\partial q(z, y_{n,M}, \Theta)}{\partial s} \exp(-z^2) dz}{\int_{-\infty}^{\infty} q(z, y_{n,M}, \Theta) \exp(-z^2) dz}, \theta_d \in \{\mu, \sigma^2\}.$$

$q(z, y_{n,M}, \Theta)$  中每个元素的一阶导数为

$$\begin{aligned}\frac{\partial q(z, y_{n,M}, \Theta)}{\partial \alpha} &= q(z, y_{n,M}, \Theta) \cdot \log(q(z, y_{n,M}, \Theta)) \cdot \alpha^{-1}, \\ \frac{\partial q(z, y_{n,M}, \Theta)}{\partial \mu} &= q(z, y_{n,M}, \Theta) \cdot \left( -\log(q(z, y_{n,M}, \Theta)) - \alpha M h_c(\sqrt{2\sigma^2}z + \mu + 1) \right), \\ \frac{\partial q(z, y_{n,M}, \Theta)}{\partial \sigma^2} &= \frac{\partial q(z, y_{n,M}, \Theta)}{\partial \mu} \cdot \frac{z}{\sqrt{2\sigma^2}}.\end{aligned}$$

接下来, 对  $\ell(\Theta)$  关于  $\Theta$  求二阶偏导数可得

$$\begin{aligned}\frac{\partial \ell(\Theta)}{\partial \alpha^2} &= -NM h_c(h_c \psi'(\alpha h_c) + 1/\alpha) + \sum_{n=1}^N [B_n(\alpha, \alpha) + A_n(\alpha)^2], \\ \frac{\partial \ell(\Theta)}{\partial \theta_d \partial \theta_s} &= \sum_{n=1}^N [B_n(\theta_d, \theta_s) - A_n(\theta_d)A_n(\theta_s)],\end{aligned}$$

其中

$$B_n(s_1, s_2) = \frac{\int_{-\infty}^{\infty} \frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial s_1 \partial s_2} \exp(-z^2) dz}{\int_{-\infty}^{\infty} q(z, y_{n,M}, \Theta) \exp(-z^2) dz}, \theta_d, \theta_s \in \Theta, (\theta_d, \theta_s) \neq (\alpha, \alpha).$$

$q(z, y_{n,M}, \Theta)$  对  $\Theta$  中每个元素的二阶导数为

$$\begin{aligned}\frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial \alpha^2} &= \frac{\partial q(z, y_{n,M}, \Theta)}{\partial \alpha} \cdot \log(q(z, y_{n,M}, \Theta)) \cdot \alpha^{-1}, \\ \frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial \alpha \partial \mu} &= \frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial \mu \partial \alpha} = (\log(q(z, y_{n,M}, \Theta)) + 1) \times \frac{\partial q(z, y_{n,M}, \Theta)}{\partial \mu} \cdot \alpha^{-1}, \\ \frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial \alpha \partial \sigma^2} &= \frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial \sigma^2 \partial \alpha} = \frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial \alpha \partial \mu} \cdot \frac{z}{\sqrt{2\sigma^2}}, \\ \frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial \mu^2} &= \frac{\partial q(z, y_{n,M}, \Theta)}{\partial \mu} \left( -\log(q(z, y_{n,M}, \Theta)) - \alpha M h_c(\sqrt{2\sigma^2}z + \mu + 1) - 1 \right) \\ &\quad - q(z, y_{n,M}, \Theta) \cdot \alpha M h_c, \\ \frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial \mu \partial \sigma^2} &= \frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial \sigma^2 \partial \mu} = \frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial \mu^2} \cdot \frac{z}{\sqrt{2\sigma^2}},\end{aligned}$$

$$\frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial(\sigma^2)^2} = \left( \frac{\partial^2 q(z, y_{n,M}, \Theta)}{\partial \mu \partial \sigma^2} - \frac{\partial q(z, y_{n,M}, \Theta)}{\partial \mu} \cdot \frac{1}{2\sigma^2} \right) \times \frac{z}{\sqrt{2\sigma^2}}.$$

上述推导过程表明, 对式 (3.47) 进行数值计算并不困难. 首先, 式 (3.47) 中涉及的所有积分项都具有 GH 求积的结构. 其次, (3.47) 中的二阶偏导数项可以通过 (3.48) 转化为简单的算术运算, 这也可以通过 GH 求积计算. 计算式 (3.47) 仅涉及一个低维的  $3 \times 3$  矩阵的求逆运算, 这可以通过多种成熟的集成算法完成, 例如 R 软件中的 `solve()` 函数. 在实际操作中, 由于计算精度溢出可能导致矩阵求逆的误差, 这种情况可以通过将参数间的协方差设定为较小值来处理.

## 基准估计方法

EM 算法通常用于对不可解析的似然函数求 MLE, 特别是对于具有缺失数据的模型 (McLachlan 等, 2007). 对于所提伽马退化模型, 设隐变量  $\eta$  为缺失数据, 然后可以使用 EM 算法来获得  $\Theta$  的 MLE, 其流程如下:

1. **E 步:** 给定第  $s$  次迭代后的估计值  $\Theta^{(s)}$ , 计算后续 M 步所需的  $\eta_n$  的矩为

$$\begin{aligned} E_{1n}^{(s)} &= E_{q_1(\eta_n)}[\ln(\eta_n)]|_{\Theta=\Theta^{(s)}}, \\ E_{2n}^{(s)} &= E_{q_1(\eta_n)}[1/(\eta_n)]|_{\Theta=\Theta^{(s)}}, \\ E_{3n}^{(s)} &= E_{q_1(\eta_n)}[(\ln(\eta_n))^2]|_{\Theta=\Theta^{(s)}}, \end{aligned}$$

其中

$$q_1(\eta_n) \propto \eta_n^{\mu^{(s)} / (\sigma^2)^{(s)} - \alpha^{(s)} M h_c - 1} e^{-\frac{\alpha^{(s)} y_{n,M}}{\eta_n} - \frac{(\ln(\eta_n))^2}{2(\sigma^2)^{(s)}}}. \quad (3.51)$$

注意, 式 (3.51) 也可以转换为 GH 求积, 并用于在 E 步中计算  $E_{dn}^{(s)}, d = 1, 2, 3$ .

2. **M 步:** 使用  $E_{dn}^{(s)}$ ,  $d = 1, 2, 3$  更新参数  $\Theta$  为

$$\begin{aligned}\alpha^{(s+1)} &= \arg \min_{\alpha \in \Omega_\alpha} \left\{ |NM [h_c(\ln(\alpha) + 1) - \psi(\alpha h_c) \right. \\ &\quad \left. + h_c \left( w_5 - M \sum_{n=1}^N E_{1n} \right) - w_6 \sum_{n=1}^N E_{2n} ]| \right\}, \\ \mu^{(s+1)} &= \frac{1}{N} \sum_{n=1}^N E_{1n}^{(s)}, \\ (\sigma^2)^{(s+1)} &= \frac{1}{N} \sum_{n=1}^N [E_{3n} - (\mu^{(s+1)})^2].\end{aligned}$$

EM 算法的实现从参数的初始值  $\Theta^{(0)}$  开始, 然后在 (3.51) 和 (3.52) 之间迭代, 直到满足给定的收敛条件.

贝叶斯方法通过结合数据与先验信息, 推断目标参数的后验分布. 然而, 在贝叶斯分析中, 直接获得每个边际后验的解析形式通常较为困难. 为此, 可采用 MCMC 方法, 特别是 Gibbs 抽样算法, 从参数的满条件后验分布中生成样本, 逐步迭代直到马尔可夫链达到平稳时终止. 对于当前的 RGa 退化模型, 基于 MCMC 的框架如下:

1. 在当前后验样本  $\Theta^{(s)}$  下, 从  $\eta_n$  的满条件后验中生成  $\eta_n^{(s+1)}$ , 即

$$p_2(\eta_n \mid \Theta^{(s)}) = \prod_{m=1}^M p_{x|\eta}(x_{n,m} \mid \alpha^{(s)}). \quad (3.52)$$

2. 在当前后验样本  $\eta^{(s+1)}$  下, 从  $\alpha$  的满条件后验中生成  $\Theta^{(s+1)}$ , 即

$$p_2(\alpha \mid \eta^{(s+1)}) = \prod_{n=1}^N \prod_{m=1}^M p_{x|\eta}(x_{n,m} \mid \eta_n^{(s+1)}, \alpha) p(\alpha), \quad (3.53)$$

以及从  $(\mu, \sigma^2)$  的满条件后验中采样:

$$p_2(\mu, \sigma^2 \mid \eta^{(s+1)}) = \prod_{n=1}^N p_\eta(\eta_n^{(s+1)} \mid \mu, \sigma^2) p(\mu, \sigma^2), \quad (3.54)$$

其中  $p(\alpha)$  和  $p(\mu, \sigma^2)$  是  $\alpha$  和  $(\mu, \sigma^2)$  的先验, 与 VB 算法中的设置相同. 贝

叶斯方法的实现从设置初始值  $\Theta^{(0)}$  和  $\eta^{(0)}$  开始, 并基于式 (3.52)-(3.54) 迭代生成后验样本, 直到所有链达到稳定且充分混合.

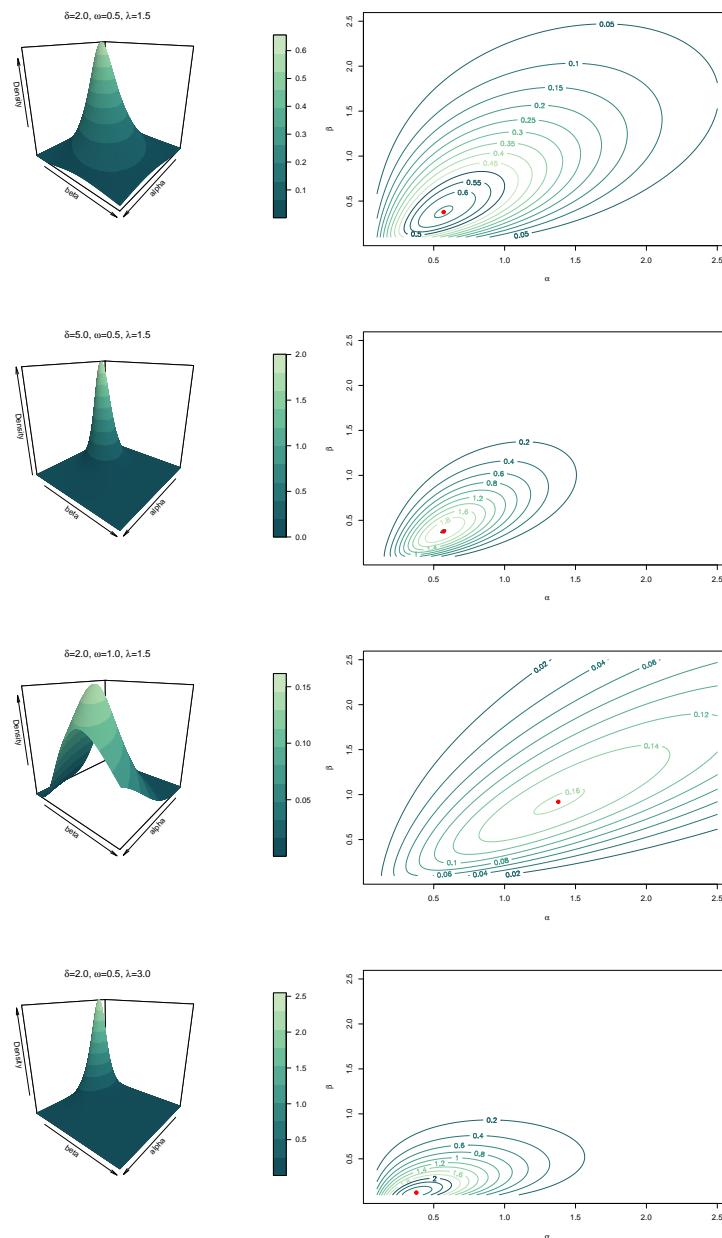


图 3.1: 共轭先验的 PDF 特性概览. 左图为 PDF 的完整形态, 右图为其对应的等高线表示.

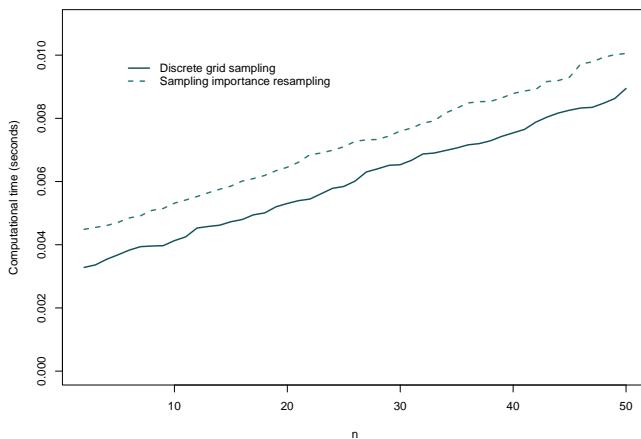


图 3.2: 不同维度下对 AGMG 分布进行采样的两种算法的计算时间.

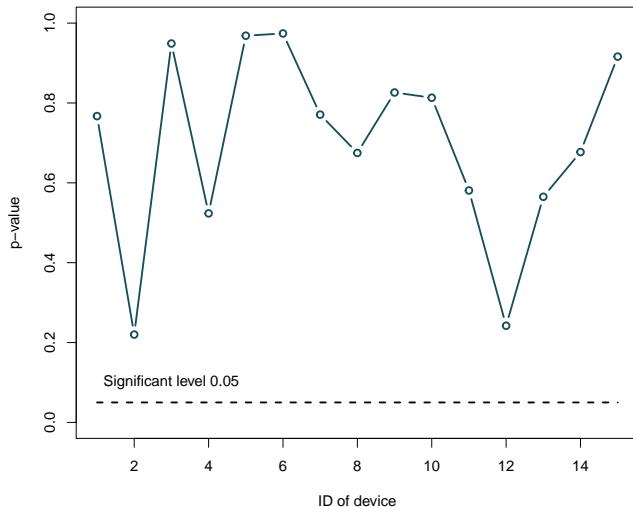
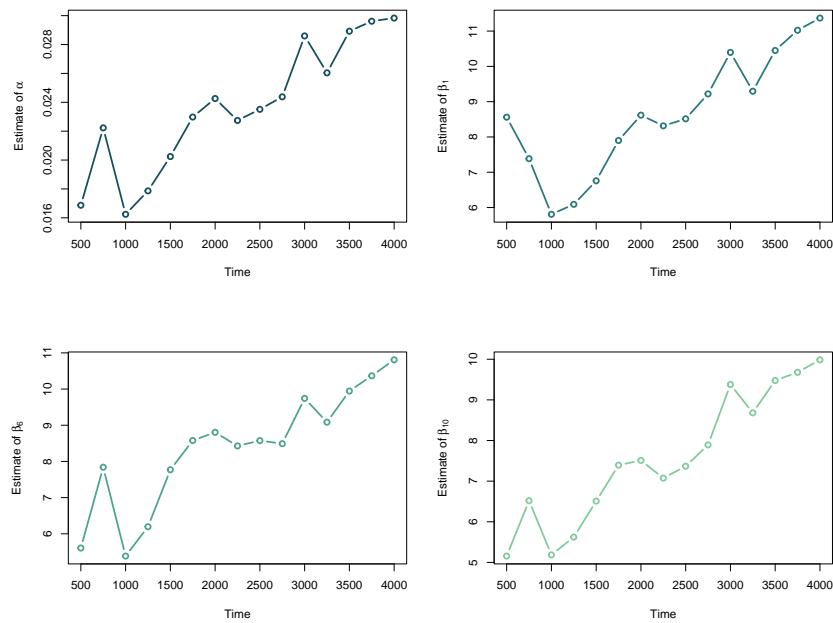


图 3.3: 每个激光设备的 Kolmogorov-Smirnov 检验  $p$  值.

图 3.4: 激光退化数据中  $\alpha$ 、 $\beta_1$ 、 $\beta_6$  和  $\beta_{10}$  的在线估计值.

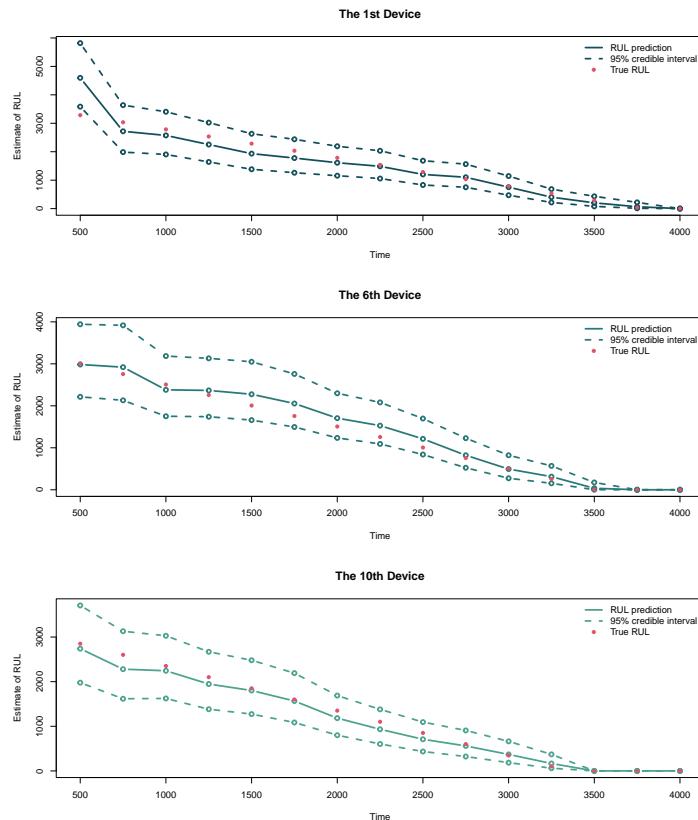


图 3.5: 第 1、6 和 10 个设备的 RUL 点预测值及其 95% 预测区间.

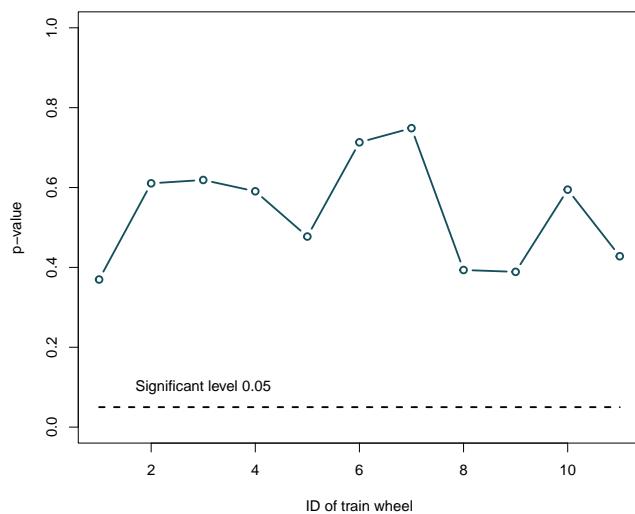


图 3.6: 每个列车车轮的 Kolmogorov-Smirnov 检验  $p$  值.

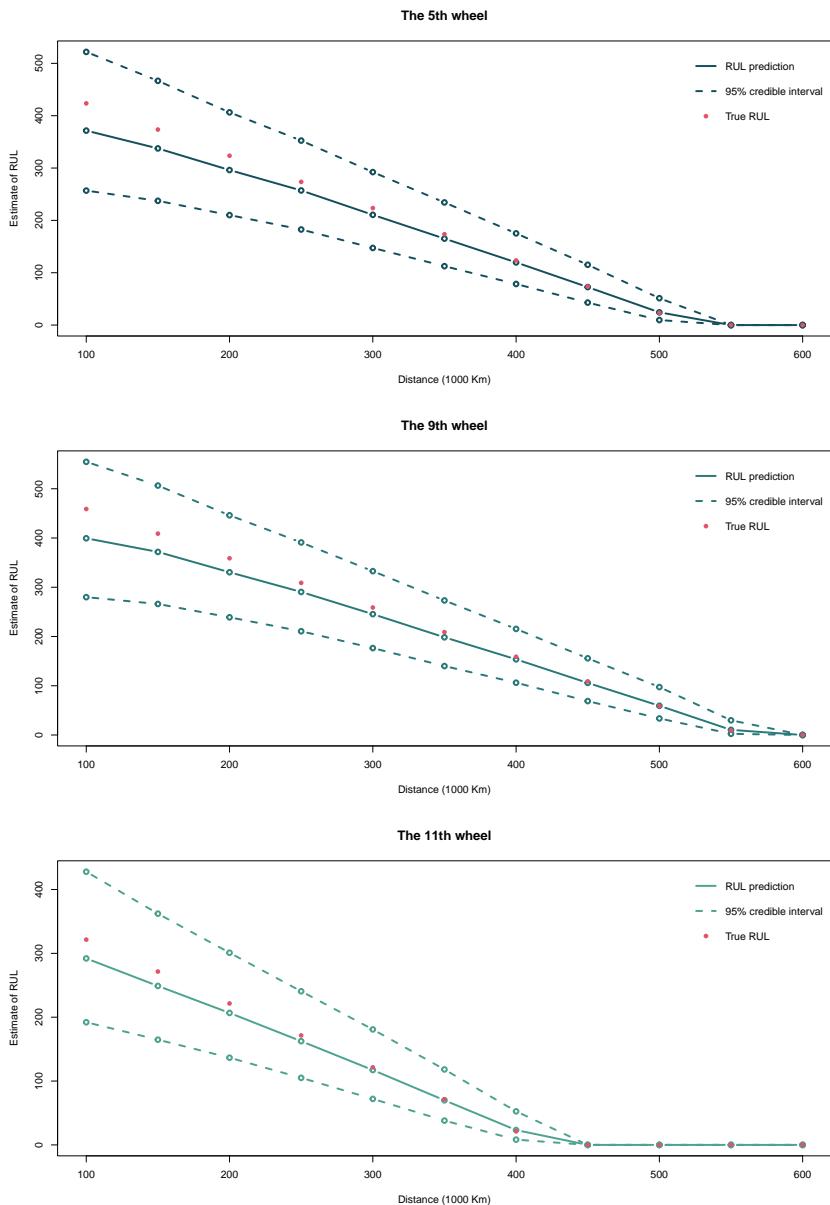


图 3.7: 第 5、9 和 11 个车轮的 RUL 点预测值及其 95% 预测区间.

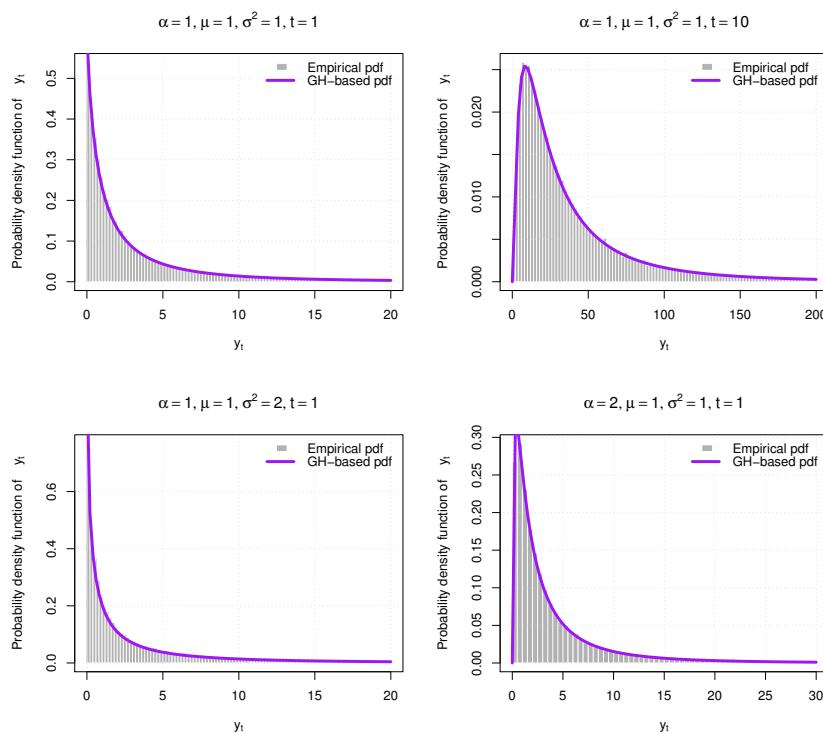


图 3.8: 不同设置下的  $y_t$  的 PDF.

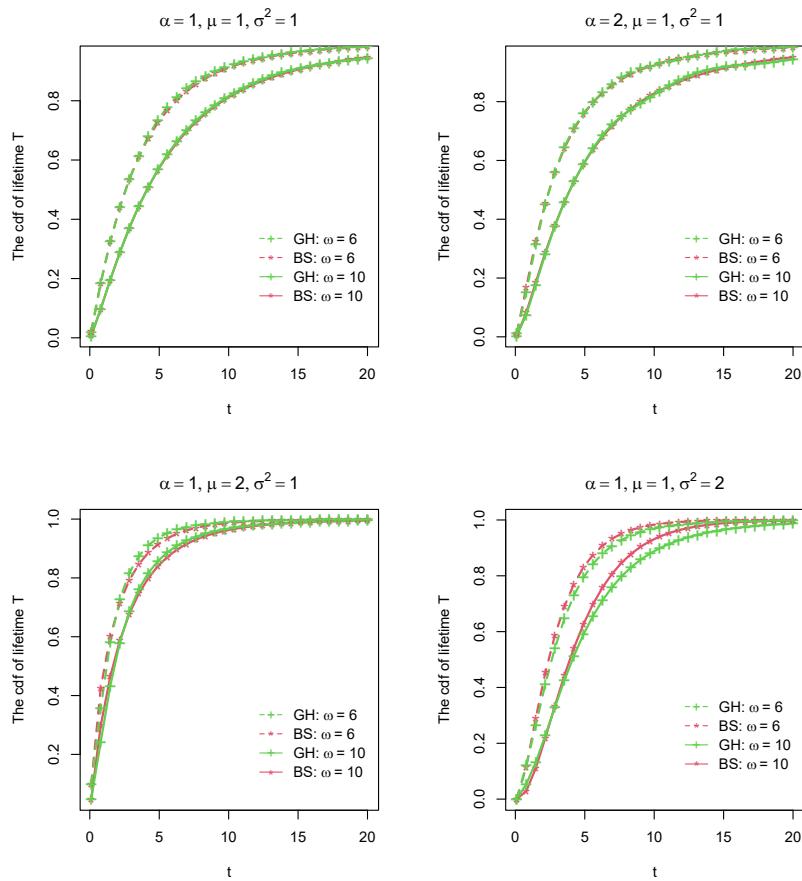


图 3.9: 不同设置下寿命  $T$  的 CDF.

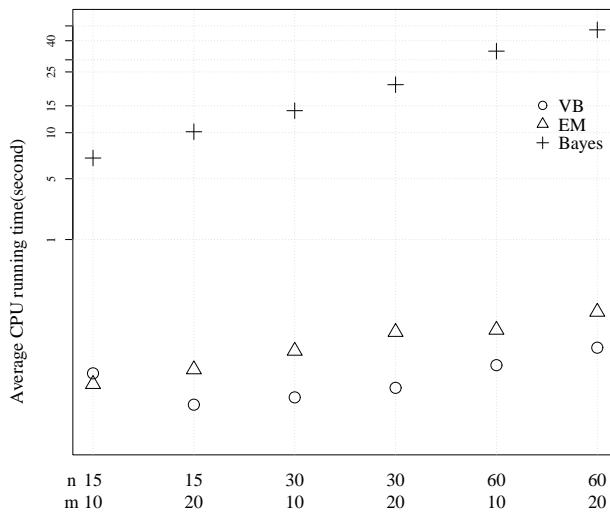


图 3.10: 在不同算法和样本量下的平均 CPU 运行时间.

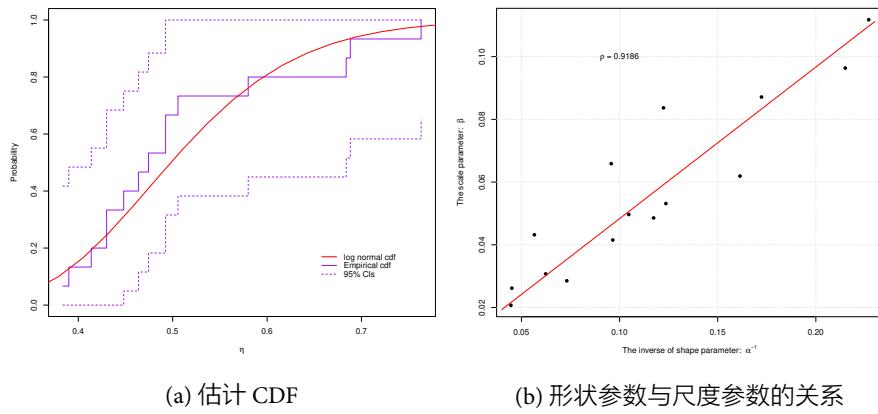
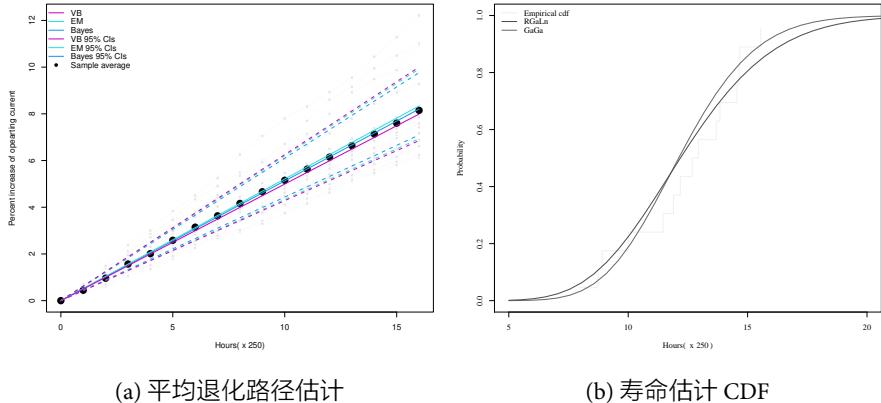


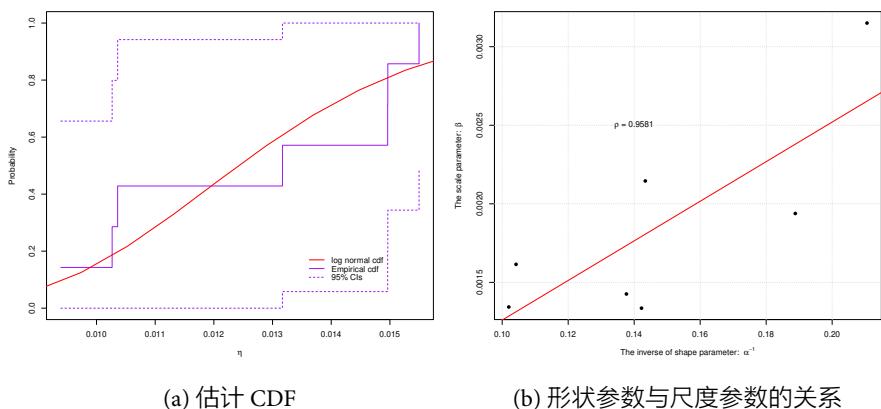
图 3.11: 模型验证相关结果.



(a) 平均退化路径估计

(b) 寿命估计 CDF

图 3.12: 激光数据下寿命分布相关结果.



(a) 估计 CDF

(b) 形状参数与尺度参数的关系

图 3.13: 模型验证相关结果.

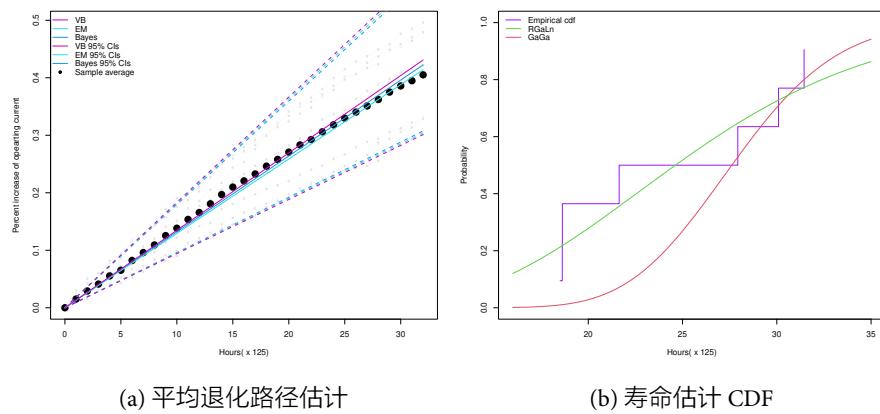


图 3.14: 设备-B 数据下寿命分布相关结果.

# 第四章 基于逆高斯过程的统计建模

## 4.1 逆高斯过程

尽管维纳过程和伽马过程在退化建模领域得到了广泛应用, 但在处理复杂退化数据时, 其拟合能力可能受到一定限制. 尤其是当退化路径存在显著异质性或动态变化特征时, 传统方法可能难以精准刻画系统的退化规律. 相比之下, 逆高斯 (Inverse Gaussian, IG) 过程凭借其灵活的分布形式和参数结构, 能够更有效地捕捉退化数据中的异质性和动态演化特征. 其在单调退化建模场景中的适用性已在多个研究中得到验证, 并展现出优越的拟合能力和预测性能 (Peng 等, 2014; Wang X 等, 2010; Ye, Chen N, 2014).

若随机过程  $\{Y(t), t \geq 0\}$  满足以下性质:

- (i)  $Y(0) = 0$ , 概率为 1;
- (ii) 对于  $t > s > u$ , 增量  $Y(t) - Y(s)$  与  $Y(s) - Y(u)$  相互独立;
- (iii) 对于  $t > s \geq 0$ , 增量  $Y(t) - Y(s)$  服从 IG 分布  $IG(\alpha(t-s), \lambda(t-s)^2)$ .

其 PDF 为

$$f_{IG}(y; \alpha, \lambda) = \sqrt{\frac{\lambda t^2}{2\pi y^3}} \exp\left\{-\frac{\lambda}{2y} \left(\frac{y}{\alpha} - t\right)^2\right\}.$$

则称该过程为 {IG 过程, 记作  $\{Y(t), t \geq 0\} \sim IG(\alpha t, \lambda t^2)$ .

IG 过程的均值和方差分别为  $\mathbb{E}[Y(t)] = \alpha t$  和  $\text{Var}[Y(t)] = \alpha^3 t / \lambda$ . 从其数学表达形式可以看出, IG 过程在刻画产品性能退化规律时具有清晰的物理解

释. 其中, 参数  $\alpha$  表示退化速率, 直接反映系统性能随时间推移的衰减速度, 而参数  $\lambda$  作为扩散系数, 衡量退化过程的波动程度, 表征退化路径的随机性和不确定性. 假设产品失效阈值为  $\omega$ , 则产品寿命定义为  $T = \inf\{t \mid Y(t) \geq \omega\}$ . 此时, 产品寿命  $T$  的 CDF 为

$$\begin{aligned} F(t) &= P(T \leq t) = P(Y(t) > \omega) \\ &= \Phi \left[ \sqrt{\frac{\lambda}{\omega}} \left( t - \frac{\omega}{\alpha} \right) \right] - \exp \left( \frac{2\lambda t}{\alpha} \right) \Phi \left[ -\sqrt{\frac{\lambda}{\omega}} \left( \frac{\omega}{\alpha} + t \right) \right], \end{aligned} \quad (4.1)$$

其中  $\Phi(\cdot)$  为标准正态分布的 CDF.

在退化建模中, IG 过程  $\mathcal{IG}(\alpha t, \lambda t^2)$  假设了线性的平均退化路径. 然而, 实际退化数据通常表现为非线性特征, 并且存在显著的异质性, 即不同产品或系统的退化行为存在差异. 为了克服这些局限性, 本章对经典 IG 模型进行了扩展: 首先, 引入了子群体异质性(详见第 4.2 节), 以便更精准地刻画不同群体之间的异质性, 从而提高模型在复杂数据中的适应性. 其次, 针对某些产品在退化过程中存在明显的阶段性变化, 提出一类新的两阶段重参数化 IG 过程(详见第 4.3 节), 并基于 RUL 分布给出了一种自适应替换策略; 最后, 针对非线性退化和带随机效应的 IG 过程, 介绍一种高效的在线算法, 实现动态更新参数估计并 RUL 分布的预测(详见 4.4 节).

## 4.2 子总体异质性下的逆高斯过程

退化数据通常因不可观测的内生因素(如原材料的初始缺陷)和外部因素(如使用模式)表现出显著的群体内异质性(Ye, Hong, 等, 2013). 随机效应模型通过引入服从特定分布(如正态分布(Si 等, 2012)、截断正态分布(Ye, Chen N, 2014)、伽马分布(Peng 等, 2014; Peng W 等, 2017))的随机参数来表征单元间的差异. 然而, 这些模型通常仅适用于描述同一群体内的异质性, 难以处理多个子群体的情况. 在实际应用中, 多子群体共存的现象非常普遍. 例如, 制造过程中的高度变异可能导致微机电系统设备的失效时间呈现多峰分布(Hartzell 等, 2011; Yuan 等, 2012), 或者在激光单元数据中, 假设多个子群体的模型优于同质群体模型(Yuan 等, 2015). 类似情况也出现在口腔冲洗器(Erişoğlu 等, 2011)和汽车铅酸电池(Kontar 等, 2017)的研究中. 为处理这种子群体异

质性, 研究者广泛采用混合分布随机效应模型 (Al-Hussaini 等, 1989; Erişoğlu 等, 2011; Li M 等, 2017). 然而, 由于此类数据通常较为稀缺, 这些方法在高可靠性产品的失效数据建模中效果有限. 随着传感器技术的进步, 退化数据的可用性显著提高, 但如何建模来自异质群体且具有子群体结构的退化数据, 现有研究仍显不足. 例如, 基于混合高斯模型的广义路径模型 (Yuan 等, 2015) 虽然能够表征单元间的变异性, 但仅适用于环境随机性较小的情况; 而基于混合正态分布参数的维纳退化过程 (Zhang 等, 2017) 则无法有效建模单调退化过程. 因此, 针对子群体异质性退化数据的建模方法仍存在较大局限性, 需进一步研究和改进.

基于上述讨论, 本节提出了一种基于 IG 过程的新模型, 用于建模存在子群体异质性的退化数据. 该模型引入混合高斯分布作为随机效应分布, 能够近似任意形式的分布 (Komárek 等, 2008; Lesaffre 等, 1991), 同时利用 IG 过程来有效建模单调退化过程. 针对模型参数增多时对初始值敏感的问题 (Lindstrom 等, 1988), 提出一种高效的 EM 算法, 通过边际化方法实现稳健的参数估计, 并结合偏差修正的百分位数自助法进行区间估计. 本节的结构如下: 第 4.2.1 节介绍带有混合高斯分布随机效应的 IG 过程; 第 4.2.2 节详细介绍了统计推断方法; 第 4.2.3 节通过模拟研究验证模型与算法的性能; 第 4.2.4 节通过两个案例研究展示模型与算法的实际效果.

### 4.2.1 模型构建

本小节首先将第 4.1 节中定义的 IG 过程扩展到非线性退化路径的情形. 令  $\Lambda(t, \beta)$  为时间尺度变换函数, 用于描述退化路径的非线性特征, 其中  $\Lambda(0, \beta) = 0$ . 与上一节中定义的 IG 过程不同, 本节中的增量  $Y(t + \Delta t) - Y(t)$  服从  $IG(\alpha \Delta \Lambda(t, \beta), \lambda \Delta \Lambda(t, \beta)^2)$ , 其中  $\Delta \Lambda(t, \beta) = \Lambda(t + \Delta t, \beta) - \Lambda(t, \beta)$ . 为了同时刻画单元与子总体的异质性, 定义  $\nu = 1/\alpha$  为随机效应, 并假设其服从混合高斯分布. 在这种设置下,  $\nu$  同时影响退化过程的均值和方差, 从而刻画了单元间的变异性, 并灵活地将总体划分为多个子总体. 设第  $k$  个高斯分布的

均值和方差分别为  $\mu_k$  和  $\sigma_k^2/\lambda$ , 则具有子总体异质性的 IG 退化模型为

$$\begin{aligned} Y(t) \mid \nu &\sim \text{IG}(\Lambda(t, \beta)/\nu, \lambda\Lambda(t, \beta)^2), \\ \nu &\sim \sum_{k=1}^K p_k N(\mu_k, \sigma_k^2/\lambda), \end{aligned} \quad (4.2)$$

其中  $p_k$  为第  $k$  个子总体的比例,  $N(\cdot)$  表示高斯分布,  $K$  为子群体总数.

**注 2:** 模型(4.2)中子总体的方差设定为  $\sigma_k^2/\lambda$ , 目的是为了简化数学推导. 实际上, 从模型的角度来看, 这一设定等同于假设  $\nu \sim \sum_{k=1}^K p_k N(\mu_k, \delta_k^2)$ . 通过参数变换  $\delta_k^2 = \sigma_k^2/\lambda$  可知两者是等价的. 然而, 模型 (4.2) 的设定在数学表达和后续统计推断中会更加简洁和方便.

基于模型(4.2), 定义产品寿命  $T = \inf\{t \mid Y(t) \geq \omega\}$ . 则  $T$  的分布可由以下定理给出.

**定理 4.1.** 在模型(4.2)下, 产品寿命  $T$  的 CDF 为

$$\begin{aligned} F_T(t) &= \int_{-\infty}^{\infty} [1 - P(Y(t) \leq \omega \mid \nu)] f(\nu) d\nu \\ &= \sum_{k=1}^K p_k \Phi \left( \frac{\sqrt{\lambda}(\Lambda(t, \beta) - \omega\mu_k)}{\sqrt{\omega(1 + \omega\sigma_k^2)}} \right) \\ &\quad - \sum_{k=1}^K p_k \exp [2\lambda\Lambda(t, \beta)(\sigma_k^2\Lambda(t, \beta) + \mu_k)] \\ &\quad \times \Phi \left( -\frac{\sqrt{\lambda}[(1 + 2\omega\sigma_k^2)\Lambda(t, \beta) + \omega\mu_k]}{\sqrt{\omega(1 + \omega\sigma_k^2)}} \right). \end{aligned} \quad (4.3)$$

根据定理 4.1 可知, 当  $K = 1$  时, 式 (4.3) 就简化为 Peng (2015) 的结果, 表示所有单元来自同一总体; 当  $K > 1$  时, 则刻画了包含  $K$  个子总体的一般情况.

### 4.2.2 统计推断

假设在退化试验中有  $n$  个样品. 设  $y_{ij}$  为第  $i$  个样品在测量时间  $t_{ij}$  ( $j = 1, \dots, m_i, i = 1, \dots, n$ ) 时所观测到的退化值. 令  $\Delta \mathbf{y}_i = (\Delta y_{i1}, \dots, \Delta y_{im_i})'$ , 其中  $\Delta y_{i1} = y_{i1}$ , 且  $\Delta y_{ij} = y_{ij} - y_{i(j-1)}$ ,  $i = 1, \dots, n$ ,  $\mathbf{y} = (\Delta \mathbf{y}_1, \dots, \Delta \mathbf{y}_n)'$ . 设  $h_{ij}(\beta) = \Lambda(t_{ij}, \beta) - \Lambda(t_{i(j-1)}, \beta)$ , 并且  $t_{i0} = 0$ ,  $j = 1, \dots, m_i, i = 1, \dots, n$ . 假设样品性能的退化过程服从模型 (4.2), 则基于第  $i$  个样品的退化数据, 给定  $\nu_i$ , 对任意的  $j$ , 有  $\Delta y_{ij} \sim IG(h_{ij}(\beta)/\nu_i, \lambda h_{ij}(\beta)^2)$ , 且  $\nu_i \sim \sum_{k=1}^K p_k N(\mu_k, \sigma_k^2/\lambda)$ . 令  $\Theta = (\lambda, \beta, p_k, \mu_k, \sigma_k^2, k = 1, \dots, K)$ . 对于第  $i$  个退化路径的完整数据  $(\Delta \mathbf{y}_i, \nu_i)$ , 其似然函数为

$$L_i(\Delta \mathbf{y}_i, \nu_i | \Theta) = \prod_{j=1}^{m_i} \sqrt{\frac{\lambda h_{ij}(\beta)^2}{2\pi \Delta y_{ij}^3}} \exp \left\{ -\frac{\lambda}{2\Delta y_{ij}} (\nu_i \Delta y_{ij} - h_{ij}(\beta))^2 \right\} \\ \times \sum_{k=1}^K p_k \sqrt{\frac{\lambda}{2\pi \sigma_k^2}} \exp \left\{ -\frac{\lambda(\nu_i - \mu_k)^2}{2\sigma_k^2} \right\}. \quad (4.4)$$

基于  $\prod_{i=1}^n L_i(\Delta \mathbf{y}_i, \nu_i | \Theta)$  求解参数  $\Theta$  的估计通常有两种方法:

- (1) EM 算法: Wang 等 (2007) 提出了一类 EM 算法, 用于求非线性随机效应混合模型的参数估计. 然而, 该方法在数学推导上较为复杂, 难以求解, 并且需要处理两类潜在变量, 这显著降低了计算效率.
- (2) 贝叶斯分析: 引入潜在变量简化 (4.4) 中的求和结构, 并为参数  $\Theta$  指定先验分布, 结合 MCMC 算法生成后验样本, 从而估计  $\Theta$ . 然而, 在实际实施过程中, 后验抽样可能面临一些挑战. 特别是参数  $\sigma_k^2$  的先验分布难以确定, 一些弱信息先验 (如均匀分布或逆伽马分布) 在特定情况下可能导致后验分布不存在, 相关问题已在文献中得到详细讨论 (Gelman, 2006; Polson 等, 2012).

基于以上分析, 本小节提出一种新的 EM 算法: 首先对  $\nu_i$  做边际化处理, 然后构造 EM 算法来实现参数估计.

### 4.2.2.1 EM 算法

由于  $\nu_i$  不可观测, 对  $L_i(\Delta\mathbf{y}_i, \nu_i \mid \Theta)$  中的  $\nu_i$  进行积分后, 基于第  $i$  个样品退化数据的似然函数为

$$\begin{aligned}
 L_i(\Delta\mathbf{y}_i \mid \Theta) &= \int_{-\infty}^{\infty} L_i(\Delta\mathbf{y}_i, \nu_i \mid \Theta) d\nu_i \\
 &= (2\pi\lambda)^{m_i/2-1} y_{im_i}^{-1/2} \prod_{j=1}^{m_i} \frac{h_{ij}(\beta)}{\Delta y_{ij}^{3/2}} \\
 &\quad \times \exp \left\{ -\frac{\lambda}{2} \left[ \sum_{j=1}^{m_i} h_{ij}(\beta)^2 / \Delta y_{ij} - \Lambda(t_{im_i}, \beta)^2 / y_{im_i} \right] \right\} \\
 &\quad \times \sum_{k=1}^K \frac{p_k \sqrt{\lambda}}{\sqrt{2\pi(1/y_{im_i} + \sigma_k^2)}} \exp \left\{ -\frac{\lambda(\Lambda(t_{im_i}, \beta)/y_{im_i} - \mu_k)^2}{2(1/y_{im_i} + \sigma_k^2)} \right\} \\
 &= G_i(\Delta\mathbf{y}_i, \beta, \lambda) \sum_{k=1}^K \frac{p_k \sqrt{\lambda}}{\sqrt{2\pi(1/y_{im_i} + \sigma_k^2)}} \\
 &\quad \times \exp \left\{ -\frac{\lambda(\Lambda(t_{im_i}, \beta)/y_{im_i} - \mu_k)^2}{2(1/y_{im_i} + \sigma_k^2)} \right\}, \tag{4.5}
 \end{aligned}$$

其中

$$\begin{aligned}
 G_i(\Delta\mathbf{y}_i, \beta, \lambda) &= (2\pi\lambda)^{m_i/2-1} y_{im_i}^{-1/2} \\
 &\quad \times \prod_{j=1}^{m_i} \frac{h_{ij}(\beta)}{\Delta y_{ij}^{3/2}} \exp \left\{ -\frac{\lambda}{2} \left[ \sum_{j=1}^{m_i} h_{ij}(\beta)^2 / \Delta y_{ij} - \Lambda(t_{im_i}, \beta)^2 / y_{im_i} \right] \right\}.
 \end{aligned}$$

为简化式 (4.5), 引入潜在变量  $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{iK})'$  表示第  $i$  个样品对应的子总体标签.  $\mathbf{Z}_i$  为多项分布变量, 当  $Z_{ik} = 1$  且  $Z_{ij} = 0, \forall j \neq k$  时, 第  $i$  个样品属于第  $k$  个子总体. 给定  $Z_{ik} = 1$  时, 似然函数为

$$\begin{aligned}
 L_i(\Delta\mathbf{y}_i \mid Z_{ik} = 1, \Theta) &= G_i(\Delta\mathbf{y}_i, \beta, \lambda) \\
 &\quad \times \frac{\sqrt{\lambda}}{\sqrt{2\pi(1/y_{im_i} + \sigma_k^2)}} \exp \left\{ -\frac{\lambda(\Lambda(t_{im_i}, \beta)/y_{im_i} - \mu_k)^2}{2(1/y_{im_i} + \sigma_k^2)} \right\}.
 \end{aligned}$$

因此, 给定第  $i$  个样品的完整数据  $(\Delta \mathbf{y}_i, \mathbf{Z}_i)$ , 似然函数可表示为

$$L_i(\Delta \mathbf{y}_i, \mathbf{Z}_i \mid \Theta) = G_i(\Delta \mathbf{y}_i, \beta, \lambda) \times \prod_{k=1}^K \left[ \frac{p_k \sqrt{\lambda}}{\sqrt{2\pi(1/y_{im_i} + \sigma_k^2)}} \exp \left\{ -\frac{\lambda(\Lambda(t_{im_i}, \beta)/y_{im_i} - \mu_k)^2}{2(1/y_{im_i} + \sigma_k^2)} \right\} \right]^{Z_{ik}}.$$

对于所有样品的完整数据  $(\Delta \mathbf{y}, \mathbf{Z})$ , 参数  $\Theta$  的对数似然函数为

$$\begin{aligned} l(\Delta \mathbf{y}, \mathbf{Z} \mid \Theta) &= \sum_{i=1}^n \log(L_i(\Delta \mathbf{y}_i, \mathbf{Z}_i \mid \Theta)) \\ &= C + \frac{1}{2} \sum_{i=1}^n m_i \log(\lambda) + \sum_{i=1}^n \sum_{j=1}^{m_i} \log h_{ij}(\beta) \\ &\quad - \frac{\lambda}{2} \left[ \sum_{i=1}^n \sum_{j=1}^{m_i} h_{ij}(\beta)^2 / \Delta y_{ij} - \sum_{i=1}^n \Lambda(t_{im_i}, \beta)^2 / y_{im_i} \right] \\ &\quad + \sum_{i=1}^n \sum_{k=1}^K Z_{ik} \left[ \log p_k - \frac{1}{2} \log(1/y_{im_i} + \sigma_k^2) \right. \\ &\quad \left. - \frac{\lambda(\Lambda(t_{im_i}, \beta)/y_{im_i} - \mu_k)^2}{2(1/y_{im_i} + \sigma_k^2)} \right], \end{aligned} \quad (4.6)$$

其中

$$C = - \sum_{i=1}^n \left[ m_i \log(2\pi) + \log(y_{im_i}) + 3 \sum_{j=1}^{m_i} \log(\Delta y_{ij}) \right] / 2.$$

EM 算法是一个迭代算法, 包含 E 步和 M 步. E 步的计算是在观测数据  $\Delta \mathbf{y}$  和上次迭代参数值  $\Theta^{(s)}$  条件下, 对  $\mathbf{Z}$  的对数似然函数的期望. 可知  $\mathbf{Z}_i$  的条件分布为多项分布:  $\mathbf{Z}_i \mid \Delta \mathbf{y}_i, \Theta \sim \mathcal{MN}(1; w_{i1}(\Theta), \dots, w_{iK}(\Theta))$ , 其中

$$w_{ik}(\Theta) = \frac{p_k(1/y_{im_i} + \sigma_k^2)^{-1} \exp \left\{ -\frac{\lambda(\Lambda(t_{im_i}, \beta)/y_{im_i} - \mu_k)^2}{2(1/y_{im_i} + \sigma_k^2)} \right\}}{\sum_{k=1}^K p_k(1/y_{im_i} + \sigma_k^2)^{-1} \exp \left\{ -\frac{\lambda(\Lambda(t_{im_i}, \beta)/y_{im_i} - \mu_k)^2}{2(1/y_{im_i} + \sigma_k^2)} \right\}},$$

表示第  $i$  个样品属于第  $k$  个子总体的概率. 在 E 步中, 结合  $Z_i$  的条件分布, 对  $l(\Delta y, Z | \Theta)$  关于  $Z$  求期望可得

$$\begin{aligned} Q(\Theta, \Theta^{(s)}) &= C + \frac{1}{2} \sum_{i=1}^n m_i \log(\lambda) + \sum_{i=1}^n \sum_{j=1}^{m_i} \log h_{ij}(\beta) \\ &\quad - \frac{\lambda}{2} \left[ \sum_{i=1}^n \sum_{j=1}^{m_i} h_{ij}(\beta)^2 / \Delta y_{ij} - \sum_{i=1}^n \Lambda(t_{im_i}, \beta)^2 / y_{im_i} \right] \\ &\quad + \sum_{i=1}^n \sum_{k=1}^K w_{ik}^{(s)} \left[ \log p_k - \frac{1}{2} \log(1/y_{im_i} + \sigma_k^2) \right. \\ &\quad \left. - \frac{\lambda(\Lambda(t_{im_i}, \beta)/y_{im_i} - \mu_k)^2}{2(1/y_{im_i} + \sigma_k^2)} \right], \end{aligned}$$

其中  $w_{ik}^{(s)} = w_{ik}(\Theta^{(s)})$ . 在 M 步中通过最大化  $Q(\Theta, \Theta^{(s)})$  得到  $\Theta^{(s+1)}$ . 具体更新公式见式 (2.15), 详细推导见本节附录 4.2.5. EM 算法的实施过程与 2.2.2.1 节类似, 这里不再重复介绍.

**注 3:** 混合模型存在识别问题, 即交换子总体标签不会影响似然函数  $\prod_{i=1}^n L_i(\Delta y_i | \Theta)$  的值 (Celeux 等, 2000). 为确保参数可识别性, 在计算过程中可对子总体均值  $\mu_k$  按递增顺序排列.

**注 4:** 所提出的 EM 算法基于给定的子总体数  $K$  进行计算.  $K$  的选择属于模型选择问题, 本节通过 AIC 来确定最优的  $K$ . AIC 定义为

$$AIC = -2 \sum_{i=1}^n \log L_i(\Delta y_i | \hat{\Theta}) + 2 \times (3K + 1).$$

最小的 AIC 值对应最优的  $K$ . 在实际操作中, 可预设一个最大值  $K$ , 该值取决于样本大小  $n$ , 确保  $n$  大于模型参数个数, 以保证参数的可识别性.

**注 5:** 本节中样品的标签信息不可观测. 基于参数估计值  $\hat{\Theta}$ , 可通过后验概率  $w_{ik}(\hat{\Theta})$  估计样品的标签. 若  $w_{ij}(\hat{\Theta}) = \max\{w_{ik}(\hat{\Theta}), k = 1, \dots, K\}$ , 则样品  $i$  被归为第  $j$  个子总体. 因此, EM 算法的实施过程同时也是样品聚类的学习过程.

### 4.2.2.2 收敛速率

EM 算法通过迭代不断优化参数估计. 根据 Wu (1983) 的定理 1, 该算法能够使对数似然函数收敛到一个稳定点或局部极大值. 为实现该算法, 需要随机初始化参数值, 通过多次迭代寻找最优解. 最终, 选取使似然函数值达到最大的点作为参数估计.

对于混合高斯分布, Xu 等 (1996) 的定理 1 表明, EM 算法本质上是一种梯度上升算法, 且收敛速度与对数似然函数的 Hessian 矩阵的条件数有关. 在本模型中, Hessian 矩阵  $H(\Theta)$  可由对数似然函数  $l(\Theta) = \sum_{i=1}^n \log L_i(\Delta y_i | \Theta)$  求得, 具体为  $H(\Theta) = \frac{\partial^2 l(\Theta)}{\partial \Theta \partial \Theta}$ . 在参数更新中, 使用牛顿法求解  $\sigma_k^2$  和  $\beta$  (见式 (4.10) 和 (4.11)). 因此, 所提的 EM 算法也可视为一种梯度上升算法.

**定理 4.2.** 在 EM 算法的每次迭代中有

$$\begin{aligned} P^{(s+1)} - P^{(s)} &= G_P^{(s)} \frac{\partial l(\Theta)}{\partial P} \Big|_{\Theta=\Theta^{(s)}}, \\ \mu_k^{(s+1)} - \mu_k^{(s)} &= G_{\mu_k}^{(s)} \frac{\partial l(\Theta)}{\partial \mu_k} \Big|_{\Theta=\Theta^{(s)}}, \\ \lambda^{(s+1)} - \lambda^{(s)} &= G_{\lambda}^{(s)} \frac{\partial l(\Theta)}{\partial \lambda} \Big|_{\Theta=\Theta^{(s)}}, \\ (\sigma_k^2)^{(s+1)} - (\sigma_k^2)^{(s)} &= G_{\sigma_k}^{(s)} \frac{\partial l(\Theta)}{\partial \sigma_k^2} \Big|_{\Theta=\Theta^{(s)}}, \\ \beta^{(s+1)} - \beta^{(s)} &= G_{\beta}^{(s)} \frac{\partial l(\Theta)}{\partial \beta} \Big|_{\Theta=\Theta^{(s)}}, \end{aligned}$$

其中  $P = (p_1, \dots, p_K)',$

$$\begin{aligned} h_{ij} &= h_{ij}(\beta^{(s)}), h_{ij}' = h_{ij}'(\beta^{(s)}), h_{ij}'' = h_{ij}''(\beta^{(s)}), \\ \Lambda_i &= \Lambda(t_{im_i}, \beta^{(s)}), \Lambda_i' = \Lambda'(t_{im_i}, \beta^{(s)}), \Lambda_i'' = \Lambda''(t_{im_i}, \beta^{(s)}), \\ G_P^{(s)} &= \frac{1}{n} \left[ \text{diag}\{p_1^{(s)}, \dots, p_K^{(s)}\} - P^{(s)} (P^{(s)})' \right], \\ G_{\mu_k}^{(s)} &= \left[ \sum_{i=1}^n \lambda^{(s)} w_{ik}^{(s)} / (1/y_{im_i} + (\sigma_k^2)^{(s)}) \right]^{-1}, \\ G_{\lambda}^{(s)} &= 2\lambda^{(s)} \left\{ \sum_{i=1}^n \left[ \sum_{j=1}^{m_i} \frac{h_{ij}^2}{\Delta y_{ij}} - \frac{\Lambda_i^2}{y_{im_i}} + \sum_{k=1}^K w_{ik}^{(s)} \frac{(\Lambda_i/y_{im_i} - \mu_k^{(s)})^2}{2(1/y_{im_i} + (\sigma_k^2)^{(s)})} \right] \right\}^{-1}, \end{aligned}$$

$$\begin{aligned}
G_{\sigma_k}^{(s)} &= - \left\{ \sum_{i=1}^n w_{ik}^{(s)} \left[ \frac{\lambda^{(s)}(\Lambda_i/y_{im_i} - \mu_k^{(s)})^2}{\left(1/y_{im_i} + (\sigma_k^2)^{(s)}\right)^3} - \frac{1}{2\left(1/y_{im_i} + (\sigma_k^2)^{(s)}\right)^2} \right] \right\}^{-1}, \\
G_{\beta}^{(s)} &= - \left\{ \sum_{i=1}^n \left[ \sum_{j=1}^{m_i} \frac{h_{ij}''h_{ij} - (h_{ij}')^2}{h_{ij}^2} \right. \right. \\
&\quad \left. \left. - \lambda^{(s)} \left( \sum_{j=1}^{m_i} \frac{h_{ij}''h_{ij} + (h_{ij}')^2}{\Delta y_{ij}} - \frac{\Lambda_i''\Lambda_i - (\Lambda_i')^2}{y_{im_i}} \right) \right] \right. \\
&\quad \left. - \sum_{i=1}^n \sum_{j=1}^{m_i} \frac{\lambda^{(s)} w_{ik}^{(s)}}{1 + y_{im_i}(\sigma_k^2)^{(s)}} \left( \frac{\Lambda_i''\Lambda_i^2 + 2(\Lambda_i')^2\Lambda_i}{y_{im_i}} - \mu_k^{(s)}(\Lambda_i''\Lambda_i + (\Lambda_i')^2) \right) \right\}^{-1}.
\end{aligned}$$

定理 4.2 的证明见本节附录 4.2.5. Xu 等 (1996) 表明, 在约束条件  $\sum_{k=1}^K p_k^{(s)} = 1$  且  $p_k^{(s)} \geq 0$  (对所有  $k$ ) 的情况下,  $G_P^{(s)}$  是正定矩阵,  $G_{\mu_k}^{(s)}$  和  $G_{\lambda}^{(s)}$  是正值, 这一点易于验证. 在定理 4.2 的证明中,  $G_{\sigma_k}^{(s)}$  和  $G_{\beta}^{(s)}$  分别为

$$G_{\sigma_k}^{(s)} = -\frac{\partial^2 l(\Theta)}{\partial \sigma_k^2 \partial \sigma_k^2} \Big|_{\Theta = \Theta^{(s)}} \quad \text{和} \quad G_{\beta}^{(s)} = -\frac{\partial^2 l(\Theta)}{\partial \beta \partial \beta} \Big|_{\Theta = \Theta^{(s)}},$$

对于足够大的  $n$ , 这些值均为正. 定义

$$\mathcal{G}(\Theta) = \text{diag}\{G_{\lambda}, G_{\beta}, G_P, G_{\mu_1}, \dots, G_{\mu_K}, G_{\sigma_1}, \dots, G_{\sigma_K}\},$$

这是一个正定矩阵. 根据定理 4.2 可写为

$$\Theta^{(s+1)} = \Theta^{(s)} + \mathcal{G}(\Theta^{(s)}) \frac{\partial l(\Theta)}{\partial \Theta} \Big|_{\Theta = \Theta^{(s)}}. \quad (4.7)$$

因此, 在每次 EM 算法迭代中, 搜索方向  $\Theta^{(s+1)} - \Theta^{(s)}$  在对数似然函数梯度上具有正投影. 令  $\mathcal{E} = \{\mathcal{E}_1, \dots, \mathcal{E}_d\}$  为一组独立的单位基向量, 覆盖  $\Theta$  的参数空间. 令  $\gamma_{s+1} = \frac{|\Theta^{(s+1)} - \Theta^*|}{|\Theta^{(s)} - \Theta^*|}$ , 其中  $\Theta^*$  是  $\Theta$  的真实值. 根据 Xu 等 (1996) 的式 (16), EM 算法的收敛速度由以下不等式界定:

$$\gamma_{s+1} \leq \sqrt{1 + \kappa_M^2 [\mathcal{E}' \mathcal{G}(\Theta^*) H(\Theta^*) \mathcal{E}] - 2\kappa_m [\mathcal{E}' \mathcal{G}(\Theta^*) H(\Theta^*) \mathcal{E}]}, \quad (4.8)$$

其中  $\kappa_m[A]$  和  $\kappa_M[A]$  分别表示矩阵  $A$  的最小和最大特征值. EM 算法的收敛速度依赖于矩阵  $\mathcal{E}' \mathcal{G}(\Theta^*) H(\Theta^*) \mathcal{E}$  的条件数, 定义为

$$\mathcal{C} = \kappa_M [\mathcal{E}' \mathcal{G}(\Theta^*) H(\Theta^*) \mathcal{E}] / \kappa_m [\mathcal{E}' \mathcal{G}(\Theta^*) H(\Theta^*) \mathcal{E}]. \quad (4.9)$$

较大的  $\mathcal{C}$  会导致较慢的收敛速度. 当  $\mathcal{C} = 1$  且  $\kappa_M [\mathcal{E}' \mathcal{G}(\Theta^*) H(\Theta^*) \mathcal{E}] = 1$  时, 式(4.8)中的上界收敛至 0, 实现超线性收敛速度. 在算法实施中, 用每次迭代的估计值代替 (4.9) 中的  $\Theta^*$ , 以监控条件数  $\mathcal{C}$  的变化, 从而评估不同算法的收敛速度.

#### 4.2.2.3 自助法计算区间

Louis (1982) 提出了推导参数 Fisher 信息矩阵的方法. 但该模型获得其解析形式较为困难. 因此, 可采用偏差校正百分位自助法计算参数的区间估计 (Efron 等, 1993). 对于参数的任意函数  $\eta = g(\Theta)$ , 具体实施步骤如下:

#### 基于偏差校正的自助法

- (1) 给定观测退化数据  $\Delta y$ , 使用 EM 算法基于模型 (4.2) 计算 MLE  $\hat{\Theta}$ , 并获得  $\hat{\eta} = g(\hat{\Theta})$ .
- (2) 设置自助法样本数  $B$ , 通过以下步骤生成每个自助法样本的估计值  $\hat{\eta}_b, b = 1, \dots, B$ :
  - (a) 根据  $\nu_i \sim \sum k = 1^K \hat{p}_k N(\hat{\mu}_k, \hat{\sigma}_k^2 / \hat{\lambda})$ ,  $i = 1, \dots, n$  生成  $\nu_i$ , 确保每个子群体至少包含两个  $\nu_i$ , 以避免信息不足影响模型估计.
  - (b) 根据  $\Delta y_{ij} \sim \text{IG}(h_{ij}(\hat{\beta}) / \nu_i, \hat{\lambda} h_{ij}(\hat{\beta})^2)$ ,  $j = 1, \dots, m_i, i = 1, \dots, n$  生成退化数据.
- (3) 对估计值  $\hat{\eta}(1) \leq \hat{\eta}(2) \leq \dots \leq \hat{\eta}(B)$  进行排序, 计算  $\eta$  的  $100(1 - \gamma)\%$  置信区间为  $(\hat{\eta}_{(L)}, \hat{\eta}_{(U)})$ , 其中  $L = B \times \Phi(2\Phi^{-1}(p_0) + \Phi^{-1}(\gamma/2))$ ,  $U = B \times \Phi(2\Phi^{-1}(p_0) + \Phi^{-1}(1 - \gamma/2))$ ,  $p_0$  是小于  $\hat{\eta}$  的自助法估计值的比例.

### 4.2.3 模拟实验

本节将通过蒙特卡洛仿真来验证 EM 算法的有效性. 设定总体包含两个子群体 ( $K = 2$ ), 异质性参数为  $p_1 = 0.4$ 、 $p_2 = 0.6$ 、 $\mu_1 = 2$ 、 $\mu_2 = 4$ 、 $\sigma_1^2 = 0.2$ 、 $\sigma_2^2 = 0.5$ . 模型的其他参数包括  $\Lambda(t, \beta) = t^{1.5}$  和  $\lambda = 1$ , 测量时间点设置为  $(t_1, t_2, \dots, t_m) = (2, 4, \dots, 2m)$ . 样本大小和测量次数分别选取  $n = 50, 100$  和  $m = 10, 20$ . 为评估不同  $(n, m)$  组合对估计结果的影响, 每个组合重复模拟 2000 次, EM 算法的容忍误差设为  $\epsilon = 10^{-5}$ .

表 4.1 列出了基于 2000 次模拟的模型参数估计结果. 可发现:

- (1) 在固定  $m$  的情况下, 随着样本量  $n$  增加, 参数估计的偏差和 RMSE 显著降低.
- (2) 在固定  $n$  的情况下, 随着测量次数  $m$  增加, 参数  $\lambda$  和  $\beta$  的 RMSE 明显减少, 而其他模型参数的 RMSE 几乎保持不变. 这是因为异质性的信息主要由样本量决定, 而参数  $\lambda$  和  $\beta$  的估计精度依赖于总测量次数  $n \times m$ . 例如, 当  $(n, m) = (50, 20)$  和  $(100, 10)$  时, 两者总测量次数均为 1000, 对应的  $\lambda$  和  $\beta$  的估计 RMSE 非常接近.

假设阈值为  $\omega = 10$ , 根据式 (4.3) 绘制产品失效时间的分布函数. 图 4.1 展示了基于估计的分布函数, 其中实线为真实分布函数, 虚线为表 4.1 中平均估计得到的分布函数, 灰色虚线为每个数据集的估计结果 (仅绘制前 100 条以便说明). 从图中可以看出, 随着样本量增加, 分布函数的估计精度逐步提高. 对于固定样本量, 不同  $m$  的分布函数估计结果相似, 这与表 4.1 的分析一致.

### 4.2.4 实例分析

本节应用所提模型分析两个实际数据集. 第 4.2.4.1 节展示模型可准确识别子总体数量, 并对每个样品实现分类. 第 4.2.4.2 节通过与其他两种考虑子总体的现有模型进行 AIC 比较, 验证模型有效性, 同时评估 EM 算法的计算效率.

#### 4.2.4.1 集成电路器件退化数据

为验证模型的有效性, 选取了一个集成电路器件退化数据的实际案例进行分析 (详见图 1.6). 根据测试条件, 设备被分为三个子总体, 且每个设备的子

表 4.1: 参数估计性能评估.

$(n, m)$	估计值	$p_1$	$\mu_1$	$\mu_2$	$\sigma_1^2$	$\sigma_2^2$	$\lambda$	$\beta$
(50,10)	均值	0.346	2.066	3.665	0.178	0.563	1.012	1.499
	Bias	0.054	0.066	0.335	0.022	0.063	0.012	0.001
	RMSE	0.172	0.589	0.621	0.256	0.397	0.0951	0.0119
(50,20)	均值	0.349	2.005	3.657	0.182	0.554	1.004	1.500
	Bias	0.051	0.005	0.343	0.018	0.054	0.004	0.000
	RMSE	0.174	0.533	0.535	0.211	0.390	0.0675	0.0074
(100,10)	均值	0.377	2.029	3.821	0.186	0.511	1.004	1.500
	Bias	0.023	0.029	0.179	0.014	0.011	0.004	0.000
	RMSE	0.128	0.412	0.456	0.161	0.313	0.0664	0.0076
(100,20)	均值	0.380	2.022	3.792	0.189	0.515	1.001	1.500
	Bias	0.020	0.022	0.208	0.011	0.015	0.001	0.000
	RMSE	0.118	0.408	0.451	0.153	0.315	0.0468	0.0050

总体归属已知. 图 4.2-(a) 显示了设备分类结果, 不同颜色代表不同子总体. 采用第 4.2.2.1 节的方法, 通过 AIC 确定子总体数量, 并在算法运行后估计设备标签. 由于真实数据已知, 可直接评估模型和算法的分类准确性.

数据采用线性退化路径  $\Lambda(t, \beta) = t$  进行拟合. 表 4.2 列出了不同子总体数量  $K = 1, 2, 3, 4$  时的参数估计和 AIC 值. 结果显示,  $K = 3$  对应最小的 AIC 值, 表明数据中存在三个子群体, 与实际情况相符. 图 4.2-(b) 展示了基于  $K = 3$  的分类结果. 此外, 本方法与广义路径模型 (Yuan 等, 2015) (图 4.2-(c)) 和维纳过程模型 (Zhang 等, 2017) (图 4.2-(d)) 进行了比较. 在本方法中, 仅有一个  $237^{\circ}C$  组的样品被错误分类为  $195^{\circ}C$  (用粗长虚线标出), 分类准确率达到 97.06. 相比之下, 其他两种模型各有三个样品分类错误, 表明本方法分类效果更佳. 值得注意的是, Wang 等 (2007) 提出的 EM 算法不适用于此例, 其复杂的数学表达式使得推导困难. 同样, 基于 MCMC 的贝叶斯分析也不适用, 因为在没有先验信息的情况下指定九个参数的先验分布可能导致不适当的后验分布 (Gelman, 2006; Polson 等, 2012), 且 MCMC 算法计算负担重, 收敛时间长. 这些结果进一步证实了模型和算法在确定子总体数量和准确分类方面的有

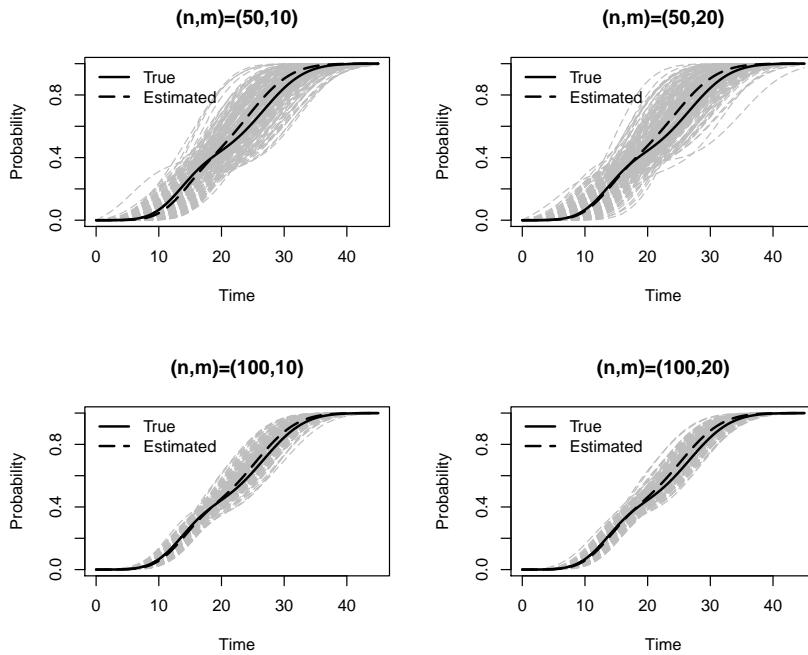


图 4.1: 不同  $(n, m)$  条件下失效时间的 CDF 估计值.

效性.

#### 4.2.4.2 铝合金裂纹退化数据

本小节以铝合金裂纹退化数据为例 (详见图 1.7), 考虑非线性退化路径函数  $\Lambda(t, \beta) = t^\beta$  和  $\Lambda(t, \beta) = \exp(\beta t) - 1$ , 并分别设置  $K = 1, 2, 3, 4$ . 各模型的点估计结果和 AIC 值列于表 ???. 结果显示,  $\Lambda(t, \beta) = \exp(\beta t) - 1$  的模型在 AIC 值上显著优于其他模型. 为方便表述, 本小节后续将直接以“模型”代指“使用  $\Lambda(t, \beta) = \exp(\beta t) - 1$  的模型”. 在子总体数量的选择上,  $K = 2$  的模型 AIC 值最小, 其次是  $K = 3$  的模型, 两者 AIC 值非常接近.

图 4.3 展示了不同  $K$  值下估计的失效时间分布, 其中黑点为拟合随机效应非线性模型的结果, 可视为“伪”失效时间点. 可以看出,  $K = 3$  的模型拟合效果更好. 然而, AIC 综合考虑了拟合优度与模型复杂度, 表明  $K = 2$  的模型更合适. 这可能是因为  $K = 3$  的模型存在过拟合现象. 表 4.4 列出了  $K = 2$  和  $K = 3$  模型参数的 95% 自助法置信区间, 图 4.3 同时展示了  $K = 2$  模型

表 4.2: 集成电路器件退化数据的参数估计结果 ( $\Lambda(t, \beta) = t$ ).

$K$	$k$	$p_k$	$\mu_k$	$\sigma_k^2$	$\lambda(\times 10^{-5})$	AIC 值
1	1	1	2910.2	173.98	1.234	-612.61
	2	0.793	1138.8	1.199		
2	2	0.207	9599.7	59.56	1.285	-668.39
	1	0.342	686.8	0.069		
	3	0.449	1444.9	2.248	1.349	<b>-683.28</b>
3	3	0.209	9747.3	62.16		
	1	0.189	618.6	$1.36 \times 10^{-2}$		
	2	0.136	725.6	$1.62 \times 10^{-3}$		
	4	0.466	1425.6	2.242	1.354	-679.96
	4	0.209	9747.3	62.16		

的 95% 点估计自助法置信区间. 似然比检验结果显示: 原假设为  $K = 1$  对备择假设  $K = 2$  的检验 P 值为 0.021, 而原假设为  $K = 2$  对备择假设  $K = 3$  的检验 P 值为 0.12. 这表明在显著性水平 0.05 下, 30 个样本中存在两个子总体. 对于  $K = 2$  的情况, 广义路径模型 (Yuan 等, 2015) 和维纳过程模型 (Zhang 等, 2017) 的 AIC 值分别为 -105.69 和 -117.83, 而该模型的 AIC 值更低, 显示了更优的性能. 这可能是因为广义路径模型未能捕捉时间变化的退化波动, 而维纳过程更适合非单调退化路径. 需要注意的是, 混合高斯分布可能导致负值, 与单调递增退化路径的假设冲突, 但根据估计参数,  $\nu$  为负的概率仅为  $2.031 \times 10^{-6}$ , 可忽略不计.

通过监测对数似然函数值评估所提 EM 算法的收敛性, 容差设定为  $\epsilon = 1 \times 10^{-5}$ . 图 4.4(a) 显示了每次迭代中对数似然函数的变化, 当  $|\sum_{i=1}^n [\log L_i(\Delta y_i | \Theta^{(s+1)}) - \log L_i(\Delta y_i | \Theta^{(s)})]| < \epsilon$  时, 算法停止. 在本例中, 收敛在第 45 次迭代后实现.

图 4.4(b) 对比了所提 EM 算法与简单梯度上升算法每次迭代的条件数, 其中简单梯度上升算法基于矩阵  $H(\Theta)$  计算条件数. 从图中可以看出, 所提 EM 算法的条件数更小, 表明其收敛速度更快. 图 4.5 显示了每次迭代的模型参数估计值变化. 可以看到, 所提 EM 算法的参数估计值收敛非常迅速, 计算时间仅为 0.5 秒 (测试环境: Windows 7 系统, Intel Core 2 Duo 处理器, 2.4 GHz, 4 GB 内

表 4.3: 对疲劳裂纹数据使用不同的  $\Lambda(t, \beta)$  进行参数估计.

K	k	$\Lambda(t, \beta) = t^\beta$					
		$p_k$	$\mu_k$	$\sigma_k^2$	$\lambda$	$\beta$	AIC
1	1	1	1.207	2.935	19.304	1.377	-51.34
	2	0.612	0.924	0.453	19.343	1.377	-54.94
2	1	0.388	1.655	0.697			
	2	0.180	1.455	0.205	19.866	1.377	-53.74
	3	0.204	1.842	0.264			
3	1	0.185	0.716	0.085			
	2	0.415	1.002	0.134	19.866	1.377	-51.18
	3	0.192	1.447	0.251			
	4	0.208	1.842	0.264			
K	k	$\Lambda(t, \beta) = \exp(\beta t) - 1$					
		$p_k$	$\mu_k$	$\sigma_k^2$	$\lambda$	$\beta$	AIC
1	1	1	0.344	4.950	391.00	0.268	-133.52
	2	0.608	0.263	0.804	390.85	0.268	-137.26
2	1	0.392	0.471	1.314			
	2	0.181	0.416	0.016	396.23	0.268	-137.12
	3	0.204	0.207	0.692			
3	1	0.400	0.288	0.135			
	2	0.196	0.413	0.047	395.73	0.268	-133.69
	3	0.204	0.526	0.078			
	4	0.200	0.207	0.692			

表 4.4: 当  $\Lambda(t, \beta) = \exp(\beta t) - 1$  时, 参数估计的 95% 置信区间.

	$p_1$	$\mu_1$	$\sigma_1^2$	$\lambda$
	(0.432, 0.768)	(0.230, 0.298)	(0.608, 1.172)	(297.29, 543.82)
$K = 2$	$p_2$	$\mu_2$	$\sigma_2^2$	$\beta$
	(0.231, 0.567)	(0.407, 0.530)	(0.746, 2.012)	(0.248, 0.285)
	$p_1$	$\mu_1$	$\sigma_1^2$	$\lambda$
	(0.366, 0.734)	(0.225, 0.301)	(0.620, 1.183)	(305.18, 554.02)
$K = 3$	$p_2$	$\mu_2$	$\sigma_2^2$	$\beta$
	(0.068, 0.466)	(0.208, 0.469)	$(3.06 \times 10^{-4}, 0.117)$	(0.249, 0.289)
	$p_3$	$\mu_3$	$\sigma_3^2$	
	(0.097, 0.435)	(0.423, 0.583)	$(2.36 \times 10^{-4}, 0.247)$	

表 4.5: 不同方法下参数估计及计算时间 ( $\Lambda(t, \beta) = \exp(\beta t) - 1, K = 2$ ).

方法	$k$	$p_k$	$\mu_k$	$\sigma_k^2$	$\lambda$	$\beta$	耗时
所提方法	1	0.608	0.263	0.804	390.85	0.268	<b>0.5 秒</b>
	2	0.392	0.471	1.314			
Wang et al.	1	0.596	0.260	0.792	390.87	0.268	1.63 秒
	2	0.404	0.466	1.326			
Marin et al.	1	0.626	0.270	0.827	390.98	0.270	23.15 秒
	2	0.374	0.483	1.297			

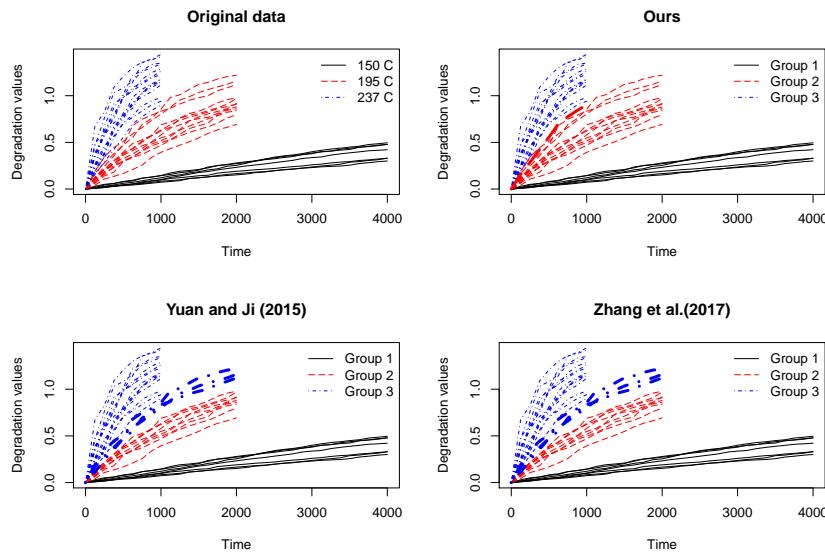


图 4.2: (a) 图为原始数据, (b)-(d) 分别展示了所提模型与其他模型的分类结果.

存). 相比之下, (Wang 等, 2007) 的 EM 算法也需 45 次迭代, 但计算时间为 1.63 秒; 而基于 MCMC 方法 (Marin 等 (2005) 的第 4.3 节) 的算法需要 20,000 次迭代, 计算时间达 23.15 秒. 三种方法的参数估计结果列于表 4.5. 图 4.6 展示了在假设存在两个子群体时, 30 个单元的聚类情况. 结果显示, 第 1 组样品的性能退化速度明显快于第 2 组.

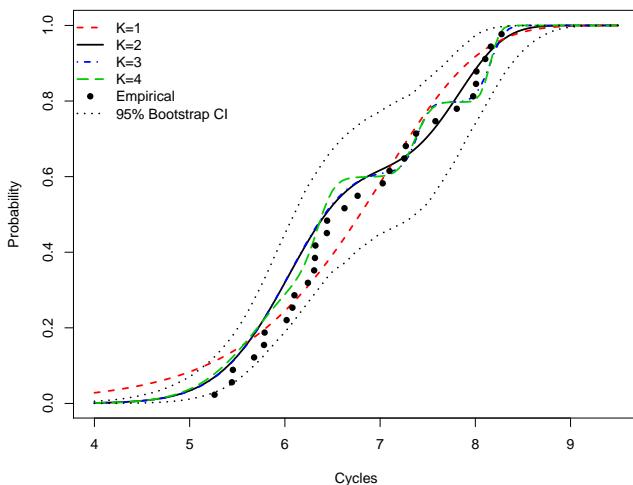


图 4.3: 基于疲劳裂纹数据的不同  $K$  的 CDF 估计.

#### 4.2.5 附录

##### EM 算法技术细节

M 步可以通过求解以下方程实现.

$$\begin{aligned}
 \frac{\partial Q(\boldsymbol{\Theta}, \boldsymbol{\Theta}^{(s)})}{\partial p_k} &= \frac{\sum_{i=1}^n w_{ik}^{(s)}}{p_k} = 0, \\
 \frac{\partial Q(\boldsymbol{\Theta}, \boldsymbol{\Theta}^{(s)})}{\partial \mu_k} &= \sum_{i=1}^n w_{ik}^{(s)} \lambda(\Lambda(t_{im_i}, \beta)/y_{im_i} - \mu_k) / (1/y_{im_i} + \sigma_k^2) = 0, \\
 \frac{\partial Q(\boldsymbol{\Theta}, \boldsymbol{\Theta}^{(s)})}{\partial \sigma_k^2} &= -\frac{1}{2} \sum_{i=1}^n \left[ \frac{w_{ik}^{(s)}}{1/y_{im_i} + \sigma_k^2} - \frac{w_{ik}^{(s)} \lambda(\Lambda(t_{im_i}, \beta)/y_{im_i} - \mu_k)^2}{(1/y_{im_i} + \sigma_k^2)^2} \right] = 0, \\
 k &= 1, \dots, K, \\
 \frac{\partial Q(\boldsymbol{\Theta}, \boldsymbol{\Theta}^{(s)})}{\partial \lambda} &= \frac{1}{2} \sum_{i=1}^n \left[ m_i / \lambda - \sum_{j=1}^{m_i} h_{ij}(\beta)^2 / \Delta y_{ij} + \Lambda(t_{im_i}, \beta)^2 / y_{im_i} \right. \\
 &\quad \left. - \sum_{k=1}^K w_{ik}^{(s)} (\Lambda(t_{im_i}, \beta)/y_{im_i} - \mu_k)^2 / 2(1/y_{im_i} + \sigma_k^2) \right] = 0,
 \end{aligned}$$

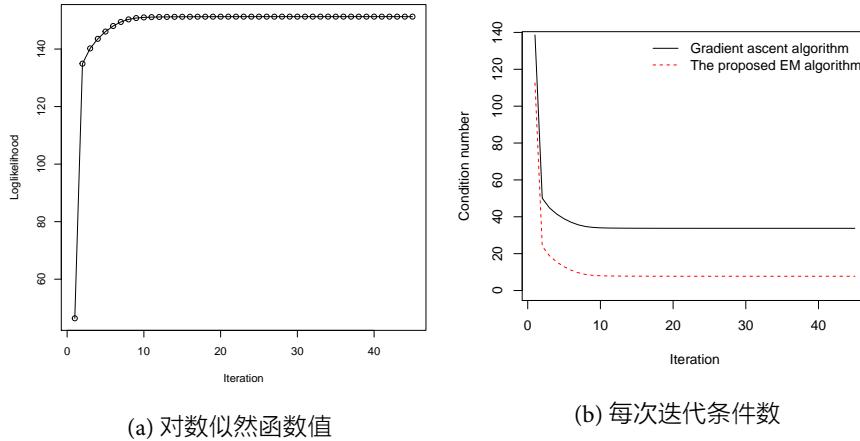


图 4.4: 所提算法迭代效果.

$$\frac{\partial Q(\Theta, \Theta^{(s)})}{\partial \beta} = \sum_{i=1}^n \left[ \sum_{j=1}^{m_i} \frac{h'_{ij}(\beta)}{h_{ij}(\beta)} - \lambda \left( \sum_{i=1}^{m_i} h'_{ij}(\beta) h_{ij}(\beta) / \Delta y_{ij} - \Lambda'(t_{im_i}, \beta) \Lambda(t_{im_i}, \beta) / y_{im_i} \right) \right] - \sum_{i=1}^n \sum_{k=1}^K w_{ik}^{(s)} \lambda \Lambda'(t_{im_i}, \beta) \Lambda(t_{im_i}, \beta) (\Lambda(t_{im_i}, \beta) / y_{im_i} - \mu_k) / (1 + y_{im_i} \sigma_k^2) = 0.$$

通过求解上述方程, 可求得第  $(s+1)$  次迭代的参数估计值  $\Theta^{(s+1)}$ , 如下所示

$$p_k^{(s+1)} = \frac{\sum_{i=1}^n w_{ik}^{(s)}}{n},$$

$$\mu_k^{(s+1)} = \frac{\sum_{i=1}^n w_{ik}^{(s)} \Lambda(t_{im_i}, \beta^{(s)}) / (y_{im_i} / (\sigma_k^2)^{(s)} + 1)}{\sum_{i=1}^n w_{ik}^{(s)} / (1 / y_{im_i} + (\sigma_k^2)^{(s)})},$$

$$\lambda^{(s+1)} = \frac{\sum_{i=1}^n m_i}{\sum_{i=1}^n \left[ \sum_{j=1}^{m_i} h_{ij}(\beta^{(s)})^2 / \Delta y_{ij} - \Lambda(t_{im_i}, \beta^{(s)})^2 / y_{im_i} + \sum_{k=1}^K w_{ik}^{(s)} \frac{(\Lambda(t_{im_i}, \beta^{(s)}) / y_{im_i} - \mu_k^{(s)})^2}{2(1 / y_{im_i} + (\sigma_k^2)^{(s)})} \right]}$$

$(\sigma_k^2)^{(s+1)}$  可以通过求解以下方程得到.

$$\sum_{i=1}^n \left[ \frac{w_{ik}^{(s)}}{1/y_{im_i} + \sigma_k^2} - \frac{w_{ik}^{(s)} \lambda^{(s)}(\Lambda(t_{im_i}, \beta^{(s)})/y_{im_i} - \mu_k^{(s)})^2}{(1/y_{im_i} + \sigma_k^2)^2} \right] = 0, \quad (4.10)$$

$\beta^{(s+1)}$  可以通过求解以下方程得到.

$$\begin{aligned} & \sum_{i=1}^n \left[ \sum_{j=1}^{m_i} \frac{h'_{ij}(\beta)}{h_{ij}(\beta)} - \lambda^{(s)} \left( \sum_{i=1}^{m_i} h'_{ij}(\beta) h_{ij}(\beta) / \Delta y_{ij} - \Lambda'(t_{im_i}, \beta) \Lambda(t_{im_i}, \beta) / y_{im_i} \right) \right] \\ & - \sum_{i=1}^n \sum_{k=1}^K w_{ik}^{(s)} \lambda^{(s)} \Lambda'(t_{im_i}, \beta) \Lambda(t_{im_i}, \beta) \frac{\Lambda(t_{im_i}, \beta) / y_{im_i} - \mu_k^{(s)}}{1 + y_{im_i} (\sigma_k^2)^{(s)}} = 0. \end{aligned} \quad (4.11)$$

通过使用 R 软件中的 `uniroot()` 函数, 求解式 (4.10) 和 (4.11) 的零点.

### 定理 4.2 的证明

对数似然函数定义为  $l(\Theta) = \sum_{i=1}^n \log L_i(\Delta y_i \mid \Theta)$ , 其中  $L_i(\Delta y_i \mid \Theta)$  的表达式如式 (4.5) 所示. 定义

$$H_{ik}(\Theta) = \frac{\sqrt{\lambda}}{\sqrt{2\pi(1/y_{im_i} + \sigma_k^2)}} \exp \left\{ -\frac{\lambda(\Lambda(t_{im_i}, \beta) / y_{im_i} - \mu_k)^2}{2(1/y_{im_i} + \sigma_k^2)} \right\}.$$

则对  $P$  的一阶导数为

$$\frac{\partial l(\Theta)}{\partial P} \Big|_{\Theta=\Theta^{(s)}} = \sum_{i=1}^n \frac{(H_{i1}(\Theta^{(s)}), \dots, H_{iK}(\Theta^{(s)}))'}{\sum_{k=1}^K p_k^{(s)} H_{ik}(\Theta^{(s)})},$$

进一步计算得到:

$$\begin{aligned} G_P^{(s)} \frac{\partial l(\Theta)}{\partial P} \Big|_{\Theta=\Theta^{(s)}} &= \frac{1}{n} \sum_{i=1}^n \frac{(p_1^{(s)} H_{i1}(\Theta^{(s)}), \dots, p_K^{(s)} H_{iK}(\Theta^{(s)}))'}{\sum_{k=1}^K p_k^{(s)} H_{ik}(\Theta^{(s)})} - P^{(s)} \\ &= \frac{1}{n} \sum_{i=1}^n (w_{i1}^{(s)}, \dots, w_{iK}^{(s)})' - P^{(s)} \\ &= P^{(s+1)} - P^{(s)}. \end{aligned}$$

类似地, 对  $\mu_k$  的一阶导数为

$$\frac{\partial l(\Theta)}{\partial \mu_k} \Big|_{\Theta=\Theta^{(s)}} = \sum_{i=1}^n \lambda^{(s)} w_{ik}^{(s)} \frac{\Lambda_i/y_{im_i} - \mu_k^{(s)}}{1/y_{im_i} + (\sigma_k^2)^{(s)}}.$$

因此

$$G_{\mu_k}^{(s)} \frac{\partial l(\Theta)}{\partial \mu_k} \Big|_{\Theta=\Theta^{(s)}} = \mu_k^{(s+1)} - \mu_k^{(s)}.$$

对  $\lambda$  进行一阶导数, 并将  $\lambda = \lambda^{(s)}$  代入, 得到

$$\frac{\partial l(\Theta)}{\partial \lambda} \Big|_{\Theta=\Theta^{(s)}} = \frac{\sum_{i=1}^n m_i}{2\lambda^{(s)}} - \frac{1}{2} \sum_{i=1}^n \left[ \sum_{j=1}^{m_i} \frac{h_{ij}^2}{\Delta y_{ij}} - \frac{\Lambda_i^2}{y_{im_i}} + \sum_{k=1}^K w_{ik}^{(s)} \frac{(\Lambda_i/y_{im_i} - \mu_k^{(s)})^2}{2(1/y_{im_i} + (\sigma_k^2)^{(s)})} \right].$$

因此

$$G_{\lambda}^{(s)} \frac{\partial l(\Theta)}{\partial \lambda} \Big|_{\Theta=\Theta^{(s)}} = \lambda^{(s+1)} - \lambda^{(s)}.$$

令

$$g_k(\sigma_k) = -\frac{1}{2} \sum_{i=1}^n \left[ \frac{w_{ik}^{(s)}}{1/y_{im_i} + \sigma_k^2} - \frac{w_{ik}^{(s)} \lambda^{(s)} (\Lambda(t_{im_i}, \beta^{(s)})/y_{im_i} - \mu_k^{(s)})^2}{(1/y_{im_i} + \sigma_k^2)^2} \right],$$

则

$$\frac{\partial l(\Theta)}{\partial \sigma_k^2} \Big|_{\Theta=\Theta^{(s)}} = g_k(\sigma_k^{(s)}).$$

$\sigma_k^2$  的 EM 更新通过牛顿法计算:

$$(\sigma_k^2)^{(s+1)} - (\sigma_k^2)^{(s)} = -\frac{g_k(\sigma_k^{(s)})}{g'_k(\sigma_k^{(s)})},$$

其中

$$g'_k(\sigma_k^{(s)}) = \sum_{i=1}^n w_{ik}^{(s)} \left[ \frac{\lambda^{(s)} (\Lambda_i/y_{im_i} - \mu_k^{(s)})^2}{(1/y_{im_i} + (\sigma_k^2)^{(s)})^3} - \frac{1}{2(1/y_{im_i} + (\sigma_k^2)^{(s)})^2} \right].$$

设  $G_{\sigma_k}^{(s)} = -1/g'_k(\sigma_k^{(s)})$ , 则结果成立. 对  $\beta$  的推导与  $\sigma_k^2$  类似, 这里省略.

## 4.3 二阶段逆高斯过程

### 4.3.1 研究背景

在 2.3 节中, 介绍了两阶段维纳过程模型, 其用来描述产品性能呈现两个不同退化的规律. 由维纳过程的性质可知, 这类模型对非单调退化数据的拟合有显著效果. 然而, 对于单调退化数据, 例如钢铝合金金属的疲劳裂纹扩展 (Lawless 等, 2004)、碳膜电阻器的电阻变化 (Park 等, 2006) 以及 LED 的退化 (Zhai, Ye, 2018), 伽马过程和 IG 过程可能会更合适. 这两类随机过程能够有效刻画单调退化模式, 并为系统剩余寿命的准确估计提供有力支持. 对于伽马过程, Ling 等 (2019) 提出了一种两阶段伽马退化过程模型, 并使用 ML 方法和贝叶斯方法估计未知参数. 同样, Lin 等 (2021) 提出了一个固定变点的两阶段伽马退化过程模型, 用于分析恒定电流下电池老化的电压-放电曲线. 而对于 IG 过程, Duan 等 (2017) 研究了一类两阶段 IG 退化过程模型, 但其工作存在以下局限:

- 1) 变点位置的约束: 在两阶段退化模型中 (如图 4.7(a) 所示),  $Y(t)$  在时间  $\tau$  之前服从 IG 过程  $\mathcal{IG}(\alpha_1 t, \lambda t^2)$ , 在  $\tau$  之后则过渡到另一个 IG 过程  $\mathcal{IG}(\alpha_2 t, \lambda t^2)$ . 假设  $t_j < \tau < t_{j+1}$ , 表示  $Y_j$ 、 $Y_\tau$  和  $Y_{j+1}$  分别为时间  $t_j$ 、 $\tau$  和  $t_{j+1}$  的退化值. 退化增量  $Y_{j+1} - Y_j$  可分解为  $Y_{j+1} - Y_\tau$  和  $Y_\tau - Y_j$  两个子增量, 且已知  $Y_{j+1} - Y_\tau \sim IG(a, b)$  和  $Y_\tau - Y_j \sim IG(c, d)$ , 其中:  $a = \alpha_2(t_{j+1} - \tau)$ ,  $b = \lambda(t_{j+1} - \tau)^2$ ,  $c = \alpha_1(\tau - t_j)$ ,  $d = \lambda(\tau - t_j)^2$ . 然而, 由于 IG 分布的可加性在此问题中不适用, 即便  $\alpha_1 = \alpha_2$ , 推导  $Y_{j+1} - Y_j$  的分布仍具有挑战性. 为简化问题, Duan 等 (2017) 假设变点  $\tau$  与测量时间点重合, 即  $\tau = t_j$  或  $\tau = t_{j+1}$ . 尽管这种假设在理论上可行, 但实际情况下变点通常是随机的, 这一约束可能导致参数估计偏差、RUL 预测不准确, 进而影响维修决策.
- 2) 寿命分布推导的局限: Duan 等 (2017) 假设变点对应的退化状态为固定值, 并在此基础上推导失效时间的分布. 然而, 变点位置的估计本质上依赖于样本信息, 其不确定性将直接影响退化状态的估计结果. 因此, 有必要进一步考虑变点的随机性, 并推导失效时间的边际分布. 此外, Duan 等 (2017) 仅聚焦于统计推断问题, 尚未充分探讨 RUL 估计及其对后续维修策略制定的潜在影响.

3) **忽略估计中的不确定性**: Duan 等 (2017) 采用 ML 法对模型参数进行点估计, 未考虑参数估计的不确定性. 在中小样本情形下, 参数估计方差可能较大, 而 MTTF、可靠性及分位寿命等关键指标一般为参数的函数, 其估计方差也会较大, 从而影响对产品可靠性的准确评估. 因此, 引入区间估计方法以量化参数不确定性尤为关键. 这不仅有助于全面反映估计量的波动, 也为后续的维修决策提供更具参考价值的信息 (Wu 等, 2023).

为更好地解决这些问题, 本节首先介绍一类新的两阶段重参数化 IG (reparameterized IG, rIG) 过程模型, 并在此基础上考虑变点的异质性以及基于 RUL 的维修决策问题. 具体包括以下三个方面:

- 1) **提出新的两阶段 rIG 退化模型**. 为克服传统两阶段 IG 退化模型中涉及的非可加问题, 引入一种新的两阶段 rIG 退化模型. 如图 4.7(b) 所示, 当变点位于时间间隔  $(t_j, t_{j+1})$  内时, 通过两个子增量  $Y_{j+1} - Y_\tau$  和  $Y_\tau - Y_j$  的分布, 推导出退化增量  $Y_{j+1} - Y_j$  的分布. 具体而言, 若  $Y_{j+1} - Y_\tau \sim rIG(a, b)$  且  $Y_\tau - Y_j \sim rIG(c, b)$ , 则  $Y_{j+1} - Y_j \sim rIG(a + c, b)$ . 该模型允许变点发生在任意时刻, 而不局限于特定的测量时间, 从而显著提升模型的灵活性和准确性, 使其能够更全面地刻画复杂退化模式.
- 2) **引入系统特定的变点和参数以刻画系统间的异质性**. 通过为每个系统设定独立的变点和模型参数, 反映设备之间的差异性, 并充分考虑变点处退化状态的不确定性. 在此基础上, 推导每个系统的故障时间分布及 RUL 分布. 为实现精准推断, 提供两种统计推断方法: i) 基于 ML 方法的参数估计, 通过自助法获得区间估计. ii) 采用自适应拒绝抽样 (Adaptive Rejection Metropolis Sampling, ARMS) 结合 Gibbs 抽样生成后验样本, 构建可信区间. 这些区间量化了参数估计和 RUL 预测中的不确定性, 为工程应用提供全面的信息支持.
- 3) **提出基于 RUL 分布的自适应替换策略**. 针对两阶段或多阶段退化的维修策略, 尽管已有研究 (Yang 等, 2017; Zhang 等, 2016; Zhang 等, 2024) 基于固定变点考虑了成本最优化决策, 但对未知参数和变点的动态检测关注较少. 文献 (Fouladirad 等, 2008, 2011; Grall 等, 2008) 考虑了已知模型参数情况下的变点检测问题. 相较之下, 本节假设参数和变点位置未知, 通过连续获取的观测数据动态更新参数估计, 并以此制定自适应替

换策略. 这种自适应策略更贴近实际应用, 可以显著提升系统动态演化背景下的维修决策效果.

本节内容安排如下: 第 4.3.2 节介绍两阶段 rIG 退化模型, 推导出相应的失效时间和剩余寿命分布, 并提出基于 RUL 的自适应替换策略. 第 4.3.3 节描述用于估计模型未知参数的两种方法: EM 方法和贝叶斯方法. 第 4.3.4 节通过模拟研究比较不同方法的统计推断性能. 第 4.3.5 节进行案例研究, 用于验证所提方法的适用性.

## 4.3.2 模型构建

### 4.3.2.1 rIG 过程

传统的 IG 过程广泛用于刻画产品性能的退化规律 (Fan 等, 2024; Hao 等, 2019; Pan 等, 2016). 然而, 当性能退化呈现多个阶段时, 传统 IG 过程对不同阶段退化的可加性质将不满足, 这限制了其应用. 为解决这一问题, 本节首先引入 rIG 分布  $rIG(\delta, \gamma)$ , 为相关的 rIG 过程提供数学基础.  $rIG(\delta, \gamma)$  与传统的 IG 分布  $IG(a, b)$  之间的参数关系为  $a = \delta/\gamma$  和  $b = \delta^2$  (Barndorff-Nielsen 等, 1998).  $rIG(\delta, \gamma)$  的 PDF 表达式为:

$$f_{rIG}(y|\delta, \gamma) = \frac{\delta}{\sqrt{2\pi}} e^{\delta\gamma} y^{-3/2} e^{-(\delta^2 y^{-1} + \gamma^2 y)/2}, \quad y > 0, \delta > 0, \gamma > 0. \quad (4.12)$$

对应的 CDF 为:

$$F_{rIG}(y|\delta, \gamma) = \Phi\left[\sqrt{y}\gamma - \frac{\delta}{\sqrt{y}}\right] + e^{2\delta\gamma} \Phi\left[-\sqrt{y}\gamma - \frac{\delta}{\sqrt{y}}\right], \quad (4.13)$$

其中  $\Phi(\cdot)$  是标准正态分布的 CDF. 若随机变量  $Y$  服从 rIG 分布  $rIG(\delta, \gamma)$ , 则其矩母函数 (moment-generating function, MGF) 为

$$M_Y(t) = \mathbb{E}(e^{ty}) = e^{\delta\gamma(1 - \sqrt{1 - \frac{2t}{\gamma^2}})}. \quad (4.14)$$

根据 (4.14) 的表达式, 可知 rIG 分布具有可加性质, 即, 若两个独立随机变量  $Y_1 \sim rIG(\delta_1, \gamma)$  和  $Y_2 \sim rIG(\delta_2, \gamma)$ , 则有  $Y_1 + Y_2 \sim rIG(\delta_1 + \delta_2, \gamma)$ .

由此, 可定义一类 rIG 过程. 若随机过程  $Z(t), t \geq 0$  满足以下条件, 则称其

为 rIG 过程: i)  $Z(0) = 0$ , 概率为 1; ii) 对于  $t_2 > t_1 \geq s_2 > s_1 \geq 0$ ,  $Z(t_2) - Z(t_1)$  与  $Z(s_2) - Z(s_1)$  相互独立; iii) 对于  $t > s \geq 0$ , 增量  $Z(t) - Z(s)$  服从 rIG 分布  $rIG(\delta(\Lambda(t) - \Lambda(s)), \gamma)$ , 其中  $\Lambda(t)$  是单调增函数, 满足  $\Lambda(0) = 0$ . 称该 rIG 过程为  $rIG(\delta\Lambda(t), \gamma)$ , 其中  $\delta$  是漂移参数,  $\gamma$  是扩散参数. 基于上述定义, 可以推导出  $Z(t)$  的均值和方差分别为  $\delta\Lambda(t)/\gamma$  和  $\delta\Lambda(t)/\gamma^3$ .

#### 4.3.2.2 两阶段 rIG 退化模型

假设系统的退化过程由两个不同阶段组成, 并通过一个变点进行区分. 在这两个阶段中, 假定退化模式符合 rIG 过程, 每个阶段使用不同的漂移参数来描述变点前后的退化行为. 由于测试系统来自相同的总体, 可以推断这些系统的故障机制是一致的. 由于扩散参数  $\gamma$  反映系统的故障机制, 进一步假定它在不同系统中保持一致. 令  $\Lambda(t) = t$  来描述退化速率随时间线性增加的情形 (Kong 等, 2017; Wang, Tang, Bae, Xu, 2018). 此外, 不同系统的变点可能不同. 为了刻画这种变异性, 将变点  $\tau$  设为随机变量, 并假定其服从正态分布, 记其 PDF 为  $g_\tau(\cdot | \mu_\tau, \sigma_\tau^2)$ . 这种处理方式既能刻画系统间的异质性, 又便于数学推导(可以得到估计  $\mu_\tau$  和  $\sigma_\tau^2$  的解析形式), 因此在退化建模领域被广泛采用 (Lu 等, 2020; Shen 等, 2018). 由此, 可考虑以下的两阶段 rIG 退化模型:

$$Y(t) | \tau \sim rIG(m(t; \delta_1, \delta_2, \tau), \gamma), \quad \tau \sim N(\mu_\tau, \sigma_\tau^2),$$

$$m(t; \delta_1, \delta_2, \tau) = \begin{cases} \delta_1 t, & t \leq \tau, \\ \delta_2(t - \tau) + \delta_1 \tau, & t > \tau, \end{cases} \quad (4.15)$$

其中  $\delta_1$  和  $\delta_2$  分别是  $t \leq \tau$  和  $t > \tau$  时的漂移参数.

#### 4.3.2.3 失效时间和 RUL 分布

基于两阶段 rIG 退化模型 4.15, 接下来对系统的失效时间和 RUL 分布进行推导, 这对于维修决策和优化资源分配具有重要意义 (Zhang 等, 2018). 失效时间  $T$  定义为系统的退化值  $Y(t)$  首次超过失效阈值  $\mathcal{D}$  的时间, 即  $T = \inf\{t | Y(t) \geq \mathcal{D}\}$ .  $T$  的可靠度函数和 MTTF 计算公式见定理 4.3.

**定理 4.3.** 系统失效时间  $T$  的可靠度函数为

$$\begin{aligned} R(t) &= P(Y(t) < \mathcal{D}, \tau \geq t) + P(Y(t) < \mathcal{D}, 0 < \tau < t) \\ &= \bar{F}_1(t | \tau) \bar{G}_\tau(t) + \int_0^t g_\tau(\tau | \mu_\tau, \sigma_\tau^2) \bar{F}_2(t | \tau) d\tau, \end{aligned} \quad (4.16)$$

其中  $\bar{G}_\tau(t)$  是随机变量  $\tau$  的生存函数,  $\bar{F}_1(t | \tau) = P(T > t | \tau \geq t)$ ,  $\bar{F}_2(t | \tau) = P(T > t | \tau < t)$ . 给定可靠性函数, MTTF 为

$$MTTF = \mathbb{E}(T) = \int_0^\infty R(t) dt. \quad (4.17)$$

对于工程师来讲, 可能更关注系统在时刻  $t$  未失效时, 还能继续正常工作多长时间. 令  $y_t$  为时刻  $t$  系统性能退化的观测值. 系统在  $t$  时刻的 RUL 定义为:  $S_t = \inf \{x; Y(t+x) \geq \mathcal{D} | Y_t < \mathcal{D}\}$ .

**定理 4.4.** RUL 的可靠度函数为

$$\begin{aligned} R_{S_t}(x) &= \bar{F}_{S_t,1}(x | \tau) \bar{G}_\tau(x+t) + \int_t^{x+t} g_\tau(\tau | \mu_\tau, \sigma_\tau^2) \bar{F}_{S_t,2}(\tau | \tau) d\tau \\ &\quad + \int_0^t g_\tau(\tau) \bar{F}_{S_t,3}(x | \tau) d\tau, \end{aligned} \quad (4.18)$$

其中,  $\bar{F}_{S_t,i}$  ( $i = 1, 2, 3$ ) 为  $S_t$  的条件可靠度函数, 由时间  $t, t+x$  和变点  $\tau$  的关系定义. RUL 的 PDF 为:

$$f_{S_t}(x) = -\frac{\partial R_{S_t}(x)}{\partial x}. \quad (4.19)$$

在时刻  $t$  的平均剩余寿命 (Mean Residual Life, MRL) 为:

$$MRL = \mathbb{E}(S_t) = \int_0^\infty R_{S_t}(x) dx. \quad (4.20)$$

定理 4.3 和 4.4 的证明见本节附录 A.

#### 4.3.2.4 自适应替换策略

本节将详细阐述自适应替换策略, 并建立了一个基于单周期准则的维修成本模型 (Lu 等, 2022; Sheu 等, 2019). 假设工程师在确定的检测时间点上对第  $i$  个系统的性能退化进行观测. 令  $0 = t_{i,0} < t_{i,1} < \dots < t_{i,j} < \dots < t_{i,n_i}$  表示检测时间点, 且  $y_{i,j}$  表示在时间点  $t_{i,j}$  所观测到的退化值. 记  $y_{i,1:j} = \{y_{i,1}, y_{i,2}, \dots, y_{i,j}\}$ . 为了充分利用每次新增的观测数据, 可采用动态更新的方式, 通过更新模型参数的估计值 (将在第 4.3.3 节中详细介绍) 给出 RUL 分布预测, 并记在  $t_{i,j}$  时刻系统的 RUL 分布为  $f_{S_t}(x|y_{i,1:j})$ . 在动态维修决策中, 需要在每个检测时间点重新评估候选维修操作, 并结合新收集的数据, 确定最优的预定备件方案和维修策略. 通过这种序贯更新过程, 企业可以实时获取系统状态信息, 动态优化维修决策, 从而实现主动预防系统失效.

假设系统的失效只能通过定期检测来发现, 每次检测的成本为  $c_i$ . 当系统发生失效并进行维修时, 工程师会将其更换为全新的同型号备件. 这种维修方式称为完美维修, 即系统在维修后被完全恢复至初始状态. 为了保障系统的可靠运行, 通常需提前准备备件. 在实际维修开始前, 还需完成一系列维修前的准备工作, 以避免操作失误或不必要的延误. 这些准备工作包括但不限于: 工具与设备的配置、技术人员的调配以及系统的关闭等. 将准备所需的时间记为  $\varpi$ .

在替换策略中, 决策者可选择纠正性替换或预防性替换两种策略. 设系统在  $t_{i,j}$  时刻正常运行, 当预测其即将发生故障时, 决策者可选择执行预防性替换, 以避免突发性故障带来的损失. 预防性替换的成本记为  $c_p$ , 且替换操作需在指定的准备时间  $\varpi$  后完成. 相反, 若在检测时发现系统已发生故障, 则必须执行纠正性替换, 对应的成本为  $c_c$ . 此外, 在系统失效后, 维修准备期间的停机将会产生额外成本, 记停机成本为  $c_b$ . 因此, 对于每个系统, 其在特定时刻  $t_{i,j}$  的最优替换时间  $\mathcal{T}_{i,j}$  可通过最小化期望成本来确定:

$$\begin{aligned} \mathcal{T}_{i,j} = \inf_{T_{i,j}} \left\{ \int_0^{T_{i,j}-t_{i,j}} \frac{c_c + c_i \lfloor x + t_{i,j} \rfloor + c_b}{x + t_{i,j} + \varpi} f_{S_t}(x|y_{i,1:j}) dx \right. \\ \left. + \int_{T_{i,j}-t_{i,j}}^{+\infty} f_{S_t}(x|y_{i,1:j}) \frac{c_p + c_i \lfloor T_{i,j} - \varpi \rfloor}{T_{i,j}} dx \right\}, \end{aligned} \quad (4.21)$$

其中,  $\lfloor \psi \rfloor = \max\{h \in \mathbb{Z} \mid t_{i,h} \leq \psi\}$  表示在时间  $\psi$  之前已完成的检测次数. 这

种决策过程在预防性替换的即时成本与等待下一次检测所可能带来的风险及相关成本之间实现了权衡. 需要注意的是,  $\mathcal{T}_{i,j}$  被视为候选的替换时间点. 这是因为在较长的时间跨度内, RUL 的预测精度可能较低. 然而, 随着检测次数的增加和观测数据的累积, 基于 RUL 的自适应替换策略将随着时间推移不断提升预测准确性. 由此, 能够实施更为可靠的维修策略, 从而优化系统的整体运行效率和安全性.

具体而言, 随着  $\mathcal{T}_{i,j}$  的动态更新, 应在首次满足条件  $\mathcal{T}_{i,j} - t_{i,j} \leq \varpi$  时, 确定最佳的准备时间点. 一旦该条件成立, 即可启动相应的准备工作, 并在准备工作完成后立即执行替换操作. 换言之, 最佳的准备时间和替换时间可通过以下公式来确定:

$$\mathcal{T}'_i = \inf_{t_{i,j}} \{\mathcal{T}_{i,j} - t_{i,j} \leq \varpi\}, \quad \text{和} \quad \mathcal{T}_i^* = \mathcal{T}'_i + \varpi. \quad (4.22)$$

替换完成后, 新安装的部件将投入运行, 并开启新的维修决策周期.

上述内容描述了所提自适应替换策略的基本原理, 接下来将重点评估该策略的实际性能. 设有  $I$  个系统, 每个系统只运行一个维修周期. 定义  $\mathbb{X}_i = \min\{\mathcal{T}_i^*, \mathcal{T}_i^f\}$ , 其中  $\mathcal{T}_i^*$  表示预测的最优替换时间,  $\mathcal{T}_i^f$  表示系统的实际失效时间. 此时, 第  $i$  个系统的实际成本率可通过以下公式计算:

$$CR_i = \begin{cases} \frac{c_p + c_i \lfloor \mathbb{X}_i - \varpi \rfloor}{\mathcal{T}_i^*}, & \mathbb{X}_i = \mathcal{T}_i^*, \\ \frac{c_c + c_i \lfloor \mathbb{X}_i \rfloor + c_b}{\mathcal{T}_i^f + \varpi}, & \mathbb{X}_i = \mathcal{T}_i^f. \end{cases} \quad (4.23)$$

因此, 可定义所有系统的平均成本率为:

$$\overline{CR} = \frac{\sum_{i=1}^I CR_i}{I}. \quad (4.24)$$

算法 4.1 给出了所提出的动态自适应替换决策过程. 在实际应用中可采用贝叶斯方法进行统计推断分析. 第 4.3.4 节的仿真结果表明, 相较于 ML 方法, 贝叶斯方法在预测准确性与不确定性量化方面表现更优. 为进一步验证基于 RUL 的自适应替换策略的有效性, 本节将其与以下两种基准维修策略进行了对比: i). 经典替换策略 (Classical Replacement Policy, CRP): 基于历史可靠性数据确定预防性维修时间, 通常以系统的 MTTF 为准. 在该策略下, 第  $i$  个系统

的成本率形式与式 (4.23) 类似, 但将  $\mathcal{T}_i^*$  替换为  $\bar{\mathcal{T}}^F$ , 且不计入检测成本. ii). 理想替换策略 (Ideal Replacement Policy, IRP): 假设可准确预测失效时间  $\mathcal{T}_i^P$ . 在这种策略下, 第  $i$  个系统的成本率为  $c_p/\mathcal{T}_i^P$ . 最终, 基于式 (4.24), 计算三种策略下所有系统的平均成本率, 以评估所提方法的相对优势.

---

#### 算法 4.1: 基于 RUL 的自适应替换策略

---

```

Input:  $y, c_c, c_p, c_b, \varpi, \mathcal{D}, j$ .
Output:  $\mathcal{T}_i^*, CR_i, i = 1, \dots, I$ , 和  $\bar{CR}$ .
1 for  $i = 1$  to  $I$  do
2   while 未执行维修 do
3     if 系统正在运行 then
4       收集新的检测数据  $Y_{i,j}$ ;
5       使用第 4.3.3 节中的贝叶斯方法更新模型参数估计;
6       根据式 (4.19) 计算 RUL 分布  $f_{S_t}(x|y_{i,1:j})_{x=0}^{+\infty}$ ;
7       利用式 (4.21) 确定  $\mathcal{T}_{i,j}$ , 并通过式 (4.22) 找到  $\mathcal{T}'_i$ ;
8       if  $t_{i,j} = \mathcal{T}'_i$  then
9         | 检测完成, 预防性维修在  $\mathcal{T}_i^*$  时执行.
10      end
11    end
12    else
13      | 纠正性维修;
14      | 设置  $\mathcal{T}_i^f = t_{i,j}$ .
15    end
16     $j = j + 1$ .
17  end
18  通过式 (4.23) 计算  $CR_i$ .
19 end
20  通过式 (4.24) 计算  $\bar{CR}$ .

```

---

### 4.3.3 统计推断

假设在退化试验中共有  $I$  个系统. 每个系统的性能退化过程在某个特定时刻  $\tau_i$  发生明显变化, 即,  $\tau_i$  为第  $i$  个系统的性能退化变点. 假设系统的性能退化规律服从两阶段 rIG 退化模型 4.15. 记  $Y_{i,j}$  表示在测量时间  $t_{i,j}$  处观测到的退化值, 其中  $i = 1, \dots, I$  且  $j = 1, \dots, n_i$ , 满足  $0 < t_{i,1} < \dots < t_{i,n_i}$ . 定义退化增量  $\Delta y_{i,j} = Y_{i,j} - Y_{i,j-1}$ , 其中  $Y_{i,0} = 0$ . 进一步定义  $\Delta \mathbf{Y}_i = (\Delta y_{i,1}, \dots, \Delta y_{i,n_i})^\top$ ,  $\Delta \mathbf{Y} = (\Delta \mathbf{Y}_1^\top, \dots, \Delta \mathbf{Y}_I^\top)^\top$ . 变点  $\tau_i$  决定了各时刻下退

化增量  $\Delta y_{i,j}$  的分布形式. 如图 4.8 所示, 变点  $\tau_i$  与测量时间点存在三种潜在关系:  $k = 1$ , 对应于  $\tau_i \geq t_{i,j}$ ;  $k = 2$ , 对应于  $t_{i,j-1} \leq \tau_i < t_{i,j}$ ;  $k = 3$ , 对应于  $\tau_i < t_{i,j-1}$ . 因此, 对于每个观测点, 退化增量  $\Delta y_{i,j}$  的分布可统一表示为  $rIG(\Delta m_{i,j}^{(k)}(\delta_{1,i}, \delta_{2,i}, \tau_i), \gamma)$ , 其中

$$\Delta m_{i,j}^{(k)}(\delta_{1,i}, \delta_{2,i}, \tau_i) = \begin{cases} \delta_{1,i} \Delta t_{i,j} & k = 1, \\ (\delta_{1,i} - \delta_{2,i}) \tau_i + \delta_{2,i} t_{i,j} - \delta_{1,i} t_{i,j-1}, & k = 2, \\ \delta_{2,i} \Delta t_{i,j}, & k = 3, \end{cases}$$

$\Delta t_{i,j} = t_{i,j} - t_{i,j-1}$  且  $t_{i,0} = 0$ ,  $i = 1 \dots, I$ ,  $j = 1 \dots, n_i$ . 为简化表达, 令  $\lambda_{i,j}^{(1)} = \mathcal{I}(\tau_i \geq t_{i,j})$ ,  $\lambda_{i,j}^{(2)} = \mathcal{I}(t_{i,j-1} \leq \tau_i < t_{i,j})$ , 和  $\lambda_{i,j}^{(3)} = \mathcal{I}(\tau_i < t_{i,j-1})$ , 其中  $\mathcal{I}(\cdot)$  为示性函数, 进一步可得:

$$\Delta m_{i,j}(\delta_{1,i}, \delta_{2,i}, \tau_i) = \Delta m_{i,j}^{(1)}(\delta_{1,i}, \delta_{2,i}, \tau_i)^{\lambda_{i,j}^{(1)}} \times \Delta m_{i,j}^{(2)}(\delta_{1,i}, \delta_{2,i}, \tau_i)^{\lambda_{i,j}^{(2)}} \times \Delta m_{i,j}^{(3)}(\delta_{1,i}, \delta_{2,i}, \tau_i)^{\lambda_{i,j}^{(3)}}.$$

因此, 给定变点  $\tau_i$ ,  $\Delta y_{i,j}$  的条件 PDF 为:

$$f_{i,j}(\Delta y_{i,j} | \delta_{1,i}, \delta_{2,i}, \tau_i, \gamma) = \frac{\Delta m_{i,j}(\delta_{1,i}, \delta_{2,i}, \tau_i)}{\sqrt{2\pi}} \exp\{\gamma \Delta m_{i,j}(\delta_{1,i}, \delta_{2,i}, \tau_i)\} \Delta y_{i,j}^{-3/2} \times \exp\left\{-\frac{[\Delta m_{i,j}(\delta_{1,i}, \delta_{2,i}, \tau_i)]^2 \Delta y_{i,j}^{-1} + \gamma^2 \Delta y_{i,j}}{2}\right\}. \quad (4.25)$$

记  $\boldsymbol{\delta}_1 = (\delta_{1,1}, \dots, \delta_{1,I})^\top$ ,  $\boldsymbol{\delta}_2 = (\delta_{2,1}, \dots, \delta_{2,I})^\top$  和  $\boldsymbol{\tau} = (\tau_1, \dots, \tau_I)^\top$ . 定义  $\boldsymbol{\eta} = (\boldsymbol{\delta}_1^\top, \boldsymbol{\delta}_2^\top, \gamma)^\top$ ,  $\boldsymbol{\theta}_\tau = (\mu_\tau, \sigma_\tau^2)^\top$  和  $\boldsymbol{\vartheta} = (\boldsymbol{\theta}_\tau^\top, \boldsymbol{\eta}^\top)^\top$ . 给定观测数据  $\Delta \mathbf{Y}$ , 模型参数  $\boldsymbol{\vartheta}$  的似然函数可表示为:

$$L_{obs}(\Delta \mathbf{Y} | \boldsymbol{\vartheta}) = \prod_{i=1}^I \int_{-\infty}^{\infty} \prod_{j=1}^{n_i} f_{i,j}(\Delta y_{i,j} | \delta_{1,i}, \delta_{2,i}, \tau_i, \gamma) g_\tau(\tau_i | \boldsymbol{\theta}_\tau) d\tau_i. \quad (4.26)$$

显然, 该似然函数涉及对变点  $\tau_i$  的积分, 难以获得参数  $\boldsymbol{\vartheta}$  的 MLE 的解析形式. 为此, 将引入一种基于 EM 算法的迭代优化方法来估计参数. EM 算法是一种处理含有潜在变量问题的经典方法, 广泛应用于可靠性建模与统计推断

中 (Xiao 等, 2023). 该算法通过在每次迭代中交替执行期望步 (E 步) 和最大化步 (M 步), 最终获得模型参数的估计值.

#### 4.3.3.1 EM 算法与自助法

在 E 步中, 构建一个 Q 函数, 用于表示在当前参数估计值下, 完全数据  $(\Delta Y, \tau)$  的对数似然函数的条件期望. 该期望值基于当前参数向量  $\boldsymbol{\vartheta}$  和  $\tau$  的条件分布进行计算. 随后, 在 M 步中, 通过最大化该 Q 函数以更新模型参数  $\boldsymbol{\vartheta}$  的估计值. 上述 E 步与 M 步的迭代过程将持续进行, 直至参数更新的幅度小于设定的收敛阈值, 从而获得稳定的参数估计. 在该框架下, 完全数据的对数似然函数可表示为:

$$l_c(\Delta Y, \tau | \boldsymbol{\vartheta}) = \sum_{i=1}^I l_i(\boldsymbol{\theta}_\tau) + \sum_{i=1}^I \sum_{j=1}^{n_i} l_{i,j}(\boldsymbol{\eta}, \tau), \quad (4.27)$$

其中,

$$\begin{aligned} l_i(\boldsymbol{\theta}_\tau) &= \log g_\tau(\tau_i | \boldsymbol{\theta}_\tau) = -\log \sqrt{2\pi} \sigma_\tau - \frac{(\tau_i - \mu_\tau)^2}{2\sigma_\tau^2}, \\ l_{i,j}(\boldsymbol{\eta}, \tau) &= \log f_{i,j}(\Delta y_{i,j} | \boldsymbol{\eta}, \tau) \\ &= -\log \sqrt{2\pi} + \log \Delta m_{i,j} + \gamma \Delta m_{i,j} - \frac{3}{2} \log \Delta y_{i,j} - \frac{\Delta m_{i,j}^2}{2\Delta y_{i,j}} - \frac{\gamma^2 \Delta y_{i,j}}{2}, \end{aligned}$$

$\Delta m_{i,j} = \Delta m_{i,j}(\delta_{1,i}, \delta_{2,i}, \tau_i)$ . 假设在第  $s$  次迭代中, M 步得到了参数估计的最优解  $\boldsymbol{\vartheta}_{(s)}$ , 则在第  $s+1$  次迭代的 E 步中, 需计算如下形式的 Q 函数:

$$\begin{aligned} \mathbf{Q}_{(s)}(\boldsymbol{\vartheta}) &= \mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} [l_c(\Delta Y, \tau | \boldsymbol{\vartheta})] \\ &= \sum_{i=1}^I \mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} [l_i(\boldsymbol{\theta}_\tau) | \Delta Y] + \sum_{i=1}^I \sum_{j=1}^{n_i} \mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} [l_{i,j}(\boldsymbol{\eta}, \tau) | \Delta Y], \end{aligned} \quad (4.28)$$

其中, 公式右侧的两项分别为  $\mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} [l_i(\boldsymbol{\theta}_\tau) | \Delta Y]$  和  $\mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} [l_{i,j}(\boldsymbol{\eta}, \tau) | \Delta Y]$ , 其详细推导见本节附录 B. 一旦得到 Q 函数, 通过以下 M 步更新参数估计:

$$\boldsymbol{\vartheta}_{(s+1)} = \arg \max \mathbf{Q}_{(s)}(\boldsymbol{\vartheta}). \quad (4.29)$$

可利用数值优化算法(如牛顿或拟牛顿算法)获得  $\boldsymbol{\vartheta}_{(s+1)}$  (Jamshidian 等, 1997). 在此基础上, 模型参数  $\boldsymbol{\vartheta}$  的 MLE 可通过迭代计算直至收敛, 变点  $\tau_i$  可通过计算  $E_{\hat{\boldsymbol{\vartheta}}} \{\tau_i \mid \Delta \mathbf{y}_i\}, i = 1, \dots, I$  获得. EM 算法的详细技术细节可见本节附录 B.

除了点估计  $\boldsymbol{\vartheta}$ , 还需要为参数函数  $h(\boldsymbol{\vartheta})$  构造置信区间. 通常的区间估计基于渐近理论. 然而, 考虑到所提出模型 Fisher 信息矩阵的复杂性, 采用参数化自助法 (Efron, 2012; Zhai, Ye, 2023) 作为替代方法以量化参数的不确定性. 自助法的实施步骤可见算法 4.2. 在获得自助法估计  $\{\hat{\boldsymbol{\vartheta}}_1^*, \dots, \hat{\boldsymbol{\vartheta}}_{\mathcal{B}}^*\}$  后, 可为参数函数  $h(\boldsymbol{\vartheta})$  构造近似  $100(1 - \zeta)\%$  的自助置信区间, 形式如下:

$$\left[ h\left(\hat{\boldsymbol{\vartheta}}^*\right)_{(\zeta \mathcal{B}/2)}, h\left(\hat{\boldsymbol{\vartheta}}^*\right)_{((1-\zeta/2)\mathcal{B})} \right],$$

其中  $h\left(\hat{\boldsymbol{\vartheta}}^*\right)_{(b)}$  表示  $\{h\left(\hat{\boldsymbol{\vartheta}}^*\right)_1, \dots, h\left(\hat{\boldsymbol{\vartheta}}^*\right)_{\mathcal{B}}\}$  中的第  $b$  个统计量.

---

#### 算法 4.2: 参数化自助法

---

**Input:** 点估计  $\hat{\boldsymbol{\vartheta}}$ .

**Output:**  $\mathcal{B}$  个自助估计值  $\{\hat{\boldsymbol{\vartheta}}_1^*, \dots, \hat{\boldsymbol{\vartheta}}_{\mathcal{B}}^*\}$ .

```

1 for  $b = 1$  to  $\mathcal{B}$  do
2   | 从  $\mathcal{N}(\hat{\mu}_\tau, \hat{\sigma}_\tau^2)$  中产生  $\tau$ ;
3   | for  $i = 1$  to  $I$  do
4     |   | for  $j = 1$  to  $n_i$  do
5       |     |   从  $rIG\left(\Delta m_{i,j}^{(k)}\left(\hat{\delta}_{1,i}, \hat{\delta}_{2,i}, \hat{\tau}_i\right), \hat{\gamma}\right), k = 1, 2, 3$  中产生退化
6       |     |   样本  $\Delta \tilde{Y}_{i,j}$ .
7       |   | end
8     |   | end
9   | 利用所提 EM 算法, 以  $\Delta \tilde{Y}$  为输入, 计算得到  $\hat{\boldsymbol{\vartheta}}_b^*$ .
end

```

---

#### 4.3.3.2 贝叶斯分析

贝叶斯分析在 PHM 领域中具有重要地位, 因其能够有效融合先验知识并量化模型参数的不确定性, 受到广泛关注 (Taylor 等, 2024; Zhou 等, 2023; Zhu 等, 2022). 其显著优势之一在于通过先验分布整合已有信息, 这些先验分布体现了对参数的先验认知. 在结合观测数据后, 贝叶斯分析可提供更为稳健的参数估计与不确定性评估. 本节将采用贝叶斯方法对两阶段退化模型中

的参数进行估计, 所考虑的模型框架如下:

$$Y_i(t|\tau_i) \sim r\mathcal{IG} (m(t; \delta_{1,i}, \delta_{2,i}, \tau_i), \gamma), \tau_i \sim N(\mu_\tau, \sigma_\tau^2), i = 1, \dots, I, \quad (4.30)$$

$$m(t; \delta_{1,i}, \delta_{2,i}, \tau_i) = \begin{cases} \delta_{1,i}t, & t \leq \tau_i, \\ \delta_{2,i}(t - \tau_i) + \delta_{1,i}\tau_i, & t > \tau_i, \end{cases} \quad (4.31)$$

$$(\mu_\tau, \sigma_\tau^2) \sim NIGa(\beta_\tau, \eta_\tau, v_\tau, \xi_\tau), \gamma \sim N(\omega, \kappa^2), \quad (4.32)$$

$$\delta_{1,i} \sim N(\mu_1, \sigma_1^2), \delta_{2,i} \sim N(\mu_2, \sigma_2^2), \quad (4.33)$$

$$(\mu_1, \sigma_1^2) \sim NIGa(\beta_1, \eta_1, v_1, \xi_1), (\mu_2, \sigma_2^2) \sim NIGa(\beta_2, \eta_2, v_2, \xi_2), \quad (4.34)$$

其中,  $NIGa(\cdot)$  表示正态逆伽马分布. 式 (4.30) 和式 (4.31) 是与第 4.3.3.1 节的模型设定一致. 在式 (4.32) 中, 模型的共享参数 (如  $\tau_i$ ) 设定先验分布, 以整合样本间信息并提高  $\tau_i$  的估计精度. 接下来, 为漂移参数设定先验分布. 由于不同系统的退化路径可能存在差异, 其漂移参数可能不同. 然而, 这些系统来自同一总体, 存在一定的共同特性. 因此, 引入分层先验方法, 在漂移参数中同时考虑总体层面的共同影响和个体层面的异质性. 这种分层建模方法首先通过式 (4.33) 描述漂移参数的个体特性, 然后在式 (4.34) 中进一步引入整体先验, 以捕捉系统间的整体相关性. 这种方法能够更全面地描述系统退化模式的复杂性.

**注 1:** 在贝叶斯框架中, 本文为参数  $\gamma$ 、 $\delta_{1,i}$  和  $\delta_{2,i}$  指定正态先验分布. 虽然这些参数可能出现负值, 但需要注意的是, 当先验分布的标准差与均值之比足够小时, 这种情况的发生概率会变得极低 (Chen 等, 2013; Wang, Tang, Joo Bae, 等, 2018). 这一假设符合贝叶斯建模的标准做法, 即在精心构造的先验分布下, 罕见的极端值通常被赋予较低权重, 从而对后验推断的影响较小. 此外, 选择正态先验不仅具有数学上的便利性, 还适合为超参数构建分层结构. 正如 (Bernardo 等, 2009) 所述, 正态逆伽马分布是正态分布的均值和方差参数的共轭先验, 这一特性极大地简化了推断过程, 使分析更高效且易于处理.

定义  $\boldsymbol{\theta} = (\boldsymbol{\vartheta}, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2)^\top$  为两阶段贝叶斯模型的参数向量. 根据贝叶斯定理, 参数  $\boldsymbol{\theta}$  的联合后验分布可表示为:

$$\begin{aligned} \pi(\boldsymbol{\theta} | \Delta \mathbf{Y}) &\propto \pi(\mu_\tau, \sigma_\tau^2) \pi(\mu_1, \sigma_1^2) \pi(\mu_2, \sigma_2^2) \pi(\gamma | \omega, \kappa) \pi(\tau | \mu_\tau, \sigma_\tau^2) \\ &\quad \times \pi(\boldsymbol{\delta}_1 | \mu_1, \sigma_1^2) \pi(\boldsymbol{\delta}_2 | \mu_2, \sigma_2^2) f_{\Delta Y}(\Delta \mathbf{Y} | \boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \tau, \gamma). \end{aligned} \quad (4.35)$$

由于  $\pi(\boldsymbol{\theta} \mid \Delta \mathbf{Y})$  形式复杂, 难以直接解析求解其贝叶斯估计. 为此, 采用 MCMC 算法 (见算法 4.3) 生成后验样本. 这里,  $\boldsymbol{\theta}_{\setminus \boldsymbol{\eta}}$  表示从  $\boldsymbol{\theta}$  中去除  $\boldsymbol{\eta}$  后的剩余参数, 满条件后验分布的推导细节可以在本节附录 C 中找到. 需要注意的是,  $\boldsymbol{\theta}$  中除  $\tau_i, \delta_{1,i}$  和  $\delta_{2,i}$  外的参数, 其完整条件后验分布是已知的, 因此这些参数的后验样本可直接通过统计软件生成. 而对于  $\tau_i, \delta_{1,i}$  和  $\delta_{2,i}$  ( $i = 1, \dots, I$ ), 需要采用 ARMS 算法 (Gilks 等, 2022).

---

### 算法 4.3: ARMS-Gibbs 采样算法.

---

**Input:** 观测数据:  $(\Delta \mathbf{Y}, \Delta \mathbf{t})$ .

**Output:**  $\boldsymbol{\theta}$  的后验样本.

- 1 设置初始值  $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\vartheta}^{(0)}, \sigma_{\tau}^{2(0)}, \mu_1^{(0)}, \sigma_1^{2(0)}, \mu_2^{(0)}, \sigma_2^{2(0)})^{\top}$ .
- 2 **for**  $s = 1$  **to**  $\mathcal{S}$  **do**
- 3     从  $NIGa(\beta_{\tau}^{\prime(s)}, \eta_{\tau}^{\prime(s)}, v_{\tau}^{\prime(s)}, \xi_{\tau}^{\prime(s)})$ ,  $NIGa(\beta_1^{\prime(s)}, \eta_1^{\prime(s)}, v_1^{\prime(s)}, \xi_1^{\prime(s)})$  和  $NIGa(\beta_2^{\prime(s)}, \eta_2^{\prime(s)}, v_2^{\prime(s)}, \xi_2^{\prime(s)})$  分别产生后验样本  $(\mu_{\tau}^{(s)}, \sigma_{\tau}^{2(s)})$ ,  $(\mu_1^{(s)}, \sigma_1^{2(s)})$  和  $(\mu_2^{(s)}, \sigma_2^{2(s)})$  来自  $(\mu_{\tau}, \sigma_{\tau}^2)$ ,  $(\mu_1, \sigma_1^2)$ ,  $(\mu_2, \sigma_2^2)$ ;
- 4     从  $N(\omega', \kappa')$  中产生  $\gamma^{(s)}$  的后验样本  $\gamma$ ;
- 5     使用 ARMS 算法从  $\pi(\delta_{1,i} \mid \boldsymbol{\theta}_{\setminus \delta_{1,i}}^{(s)}, \Delta \mathbf{Y})$ ,  $\pi(\delta_{2,i} \mid \boldsymbol{\theta}_{\setminus \delta_{2,i}}^{(s)}, \Delta \mathbf{Y})$  和  $\pi(\tau_i \mid \boldsymbol{\theta}_{\setminus \tau_i}^{(s)}, \Delta \mathbf{Y})$ ,  $i = 1, \dots, I$  中生成参数  $\delta_{1,i}^{(s)}$ ,  $\delta_{2,i}^{(s)}$  和  $\tau_i^{(s)}$  的后验样本.
- 6 **end**
- 7 丢弃前  $\mathcal{L}$  个预烧样本后, 保留每个参数的  $\mathcal{S} - \mathcal{L}$  个后验样本. 基于这些后验样本, 可以计算参数的点估计 (例如后验均值或后验中位数) 以及构建区间估计 (例如最高密度区间或等尾置信区间), 以量化参数的不确定性.

---

#### 4.3.4 模拟实验

本节通过模拟研究对所提出的模型与参数推断方法的性能进行评估. 为此, 考虑三种不同的系统数量  $I$  和观测点数量  $n_i$ : 情形 (I):  $I = 5, n_i = 20$ ; 情形 (II):  $I = 5, n_i = 40$ ; 情形 (III):  $I = 8, n_i = 20$ . 扩散参数  $\gamma$  设定为 2. 为体现系统间的异质性, 漂移参数和变点按照如下方式随机生成: 从  $N(4, 1)$  分布中生成  $\delta_{1,1}, \dots, \delta_{1,I}$ ; 从  $N(15, 1)$  分布中生成  $\delta_{2,1}, \dots, \delta_{2,I}$ ; 从  $N(10, 1)$  分布中生成变点  $\tau_1, \dots, \tau_I$ . 在给定变点  $\tau_i$  的条件下, 每个系统的退化增量根据式 (4.25) 所定义的 rIG 分布模拟生成. 为降低随机抽样对结果的影响, 每种情形均重复模拟 500 组样本, 用于后续的统计分析与性能评估

#### 4.3.4.1 参数估计的性能评估

首先, 使用所提出的模型和方法对模拟数据进行拟合. 对于贝叶斯方法, 先验分布设定为无信息先验, 具体如下:  $(\mu_\tau, \sigma_\tau^2) \sim NIGa(8, 100, 0.01, 0.01)$ ,  $(\mu_1, \sigma_1^2) \sim NIGa(1, 100, 0.01, 0.01)$ ,  $(\mu_2, \sigma_2^2) \sim NIGa(2, 100, 0.01, 0.01)$ ,  $\gamma \sim N(5, 100)$ . 后验样本的生成采用第 4.3.3.2 节中介绍的 ARMS-Gibbs 算法. 设定烧蚀期样本长度为  $\mathcal{L} = 5000$ , 进行  $\mathcal{S} - \mathcal{L} = 5000$  次迭代以获得后验样本. 所有参数的贝叶斯点估计均取对应后验样本的均值. 对于 ML 方法, 使用上述贝叶斯估计结果作为 EM 算法的初始值. 参数的点估计通过第 4.3.3.1 节中描述的 EM 算法获得, 参数的区间估计则通过参数化自助法实现, 自助样本数量设为  $\mathcal{B} = 500$ . EM 算法的收敛标准为:  $|\boldsymbol{\vartheta}_{(s+1)} - \boldsymbol{\vartheta}_{(s)}| < 10^{-3}$ , 其中  $|\cdot|$  表示  $L_1$  距离.

表 ?? 展示了两种推断方法的评估结果, 包括 RB、RMSE 和 95% 区间估计的 CP. 从点估计的角度来看, 贝叶斯方法与 ML 方法均表现出较小的相对偏差和合理的 RMSE, 表明两种方法在参数点估计方面均具备较好的性能. 需要指出的是, 在情形 I 与情形 II 中, 随着单个系统观测次数  $n_i$  的增加, 两种方法的 RMSE 均显著降低, 说明增加观测频率有助于提升参数估计的准确性. 而在情形 III 中, 由于试验中包含更多的系统数量, RMSE 有一定程度的下降, 进一步表明系统间信息的融合对估计精度具有积极影响. 然而, 在区间估计方面, 贝叶斯方法在所有情形下均表现出更优性能, 其覆盖概率更接近 0.95 的名义水平, 显示出较强的区间置信性. 相比之下, ML 方法在所有情形中的 CP 均明显低于 0.95, 表明其在不确定性量化方面存在一定不足. 综上所述, 建议在两阶段退化模型的参数估计中优先采用贝叶斯方法, 以获得更为准确的点估计和更具可靠性的区间估计.

#### 4.3.4.2 可靠性估计的性能评估

本小节进一步开展模拟研究, 以评估所提模型在可靠性估计中的优势. 选择情形 I 和情形 III, 假设系统故障均发生于第二阶段, 且失效阈值设定为 75. 为进行对比分析, 引入三种不考虑变点结构的基准模型, 包括: 1) 线性 rIG 模型, 退化路径设为  $\Lambda(t) = t$ ; 2) 幂律模型:  $\Lambda(t; \varsigma) = t^\varsigma$ ; 3) 指数模型:  $\Lambda(t; \varsigma) = \exp(\varsigma t) - 1$ . 对于所有基准模型, 均采用贝叶斯方法进行推断, 并设参数  $\varsigma$  的先验为正态分布  $N(5, 100)$ , 其中较大的方差表示参数  $\varsigma$  的先验信息较弱. 其

表 4.6: 三种情景下不同参数估计方法结果比较

场景	方法	$\delta_{1,1}$	$\delta_{1,2}$	$\delta_{1,3}$	$\delta_{1,4}$	$\delta_{1,5}$	$\delta_{2,1}$	$\delta_{2,2}$	$\delta_{2,3}$	$\delta_{2,4}$	$\delta_{2,5}$	$\gamma$		
I	ML	RB	0.024	0.029	-0.007	0.015	0.012	-0.026	0.019	0.023	0.056	0.003	0.011	
		HB	RMSE	1.326	1.363	1.357	1.332	1.330	0.422	0.424	0.476	0.422	0.431	0.168
		CP	0.956	0.953	0.946	0.953	0.957	0.941	0.925	0.900	0.928	0.926	0.964	
		RB	0.057	0.039	0.040	0.057	0.050	0.065	0.071	0.057	0.078	0.060	0.057	
		RMSE	1.315	1.381	1.302	1.401	1.508	0.641	0.645	0.576	0.667	0.739	0.308	
		CP	0.889	0.922	0.878	0.900	0.833	0.922	0.922	0.900	0.889	0.867	0.811	
	HB	Stat.	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$							
		RB	0.002	0.001	0.002	0.001	-0.009							
		RMSE	0.248	0.224	0.240	0.191	0.243							
		CP	0.915	0.937	0.937	0.961	0.961							
场景	方法	$\delta_{1,1}$	$\delta_{1,2}$	$\delta_{1,3}$	$\delta_{1,4}$	$\delta_{1,5}$	$\delta_{2,1}$	$\delta_{2,2}$	$\delta_{2,3}$	$\delta_{2,4}$	$\delta_{2,5}$	$\gamma$		
II	ML	RB	-0.005	0.007	0.023	0.011	-0.005	-0.019	0.000	0.016	0.000	0.012	0.001	
		HB	RMSE	1.068	1.011	1.065	1.015	1.044	0.349	0.283	0.275	0.355	0.332	0.124
		CP	0.930	0.945	0.950	0.944	0.927	0.902	0.925	0.947	0.885	0.902	0.914	
		RB	0.036	0.035	0.017	0.032	0.039	0.029	0.041	0.036	0.025	0.042	0.039	
		RMSE	0.944	1.010	0.880	0.900	0.985	0.331	0.358	0.323	0.328	0.346	0.150	
		CP	0.905	0.890	0.905	0.920	0.900	0.895	0.890	0.930	0.930	0.920	0.865	
	HB	Stat.	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$							
		RB	0.002	0.000	-0.001	0.003	-0.004							
		RMSE	0.225	0.214	0.218	0.185	0.189							
		CP	0.951	0.941	0.929	0.966	0.942							
场景	方法	$\delta_{1,1}$	$\delta_{1,2}$	$\delta_{1,3}$	$\delta_{1,4}$	$\delta_{1,5}$	$\delta_{1,6}$	$\delta_{1,7}$	$\delta_{1,8}$					
III	ML	RB	-0.024	-0.010	-0.004	-0.010	0.010	-0.002	0.015	0.029				
		HB	RMSE	1.121	1.096	1.087	1.083	1.083	1.221	1.124	1.155			
		CP	0.946	0.953	0.942	0.951	0.947	0.911	0.943	0.940				
		RB	0.089	0.073	0.086	0.079	0.066	0.076	0.074	0.073				
		RMSE	1.098	1.095	1.087	1.179	1.015	1.028	0.993	1.018				
		CP	0.887	0.900	0.913	0.880	0.887	0.887	0.867	0.893				
	HB		$\delta_{2,1}$	$\delta_{2,2}$	$\delta_{2,3}$	$\delta_{2,4}$	$\delta_{2,5}$	$\delta_{2,6}$	$\delta_{2,7}$	$\delta_{2,8}$	$\gamma$			
		RB	0.011	-0.060	0.012	-0.073	0.030	-0.022	0.021	0.107	-0.001			
		RMSE	0.463	0.432	0.314	0.485	0.327	0.356	0.379	0.494	0.138			
		CP	0.915	0.909	0.977	0.916	0.960	0.931	0.947	0.918	0.946			
		RB	0.087	0.095	0.087	0.085	0.102	0.070	0.087	0.097	0.091			
	ML	RMSE	0.642	0.629	0.606	0.604	0.623	0.604	0.545	0.569	0.230			
		CP	0.880	0.887	0.893	0.887	0.873	0.900	0.920	0.900	0.893			
			$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$	$\tau_6$	$\tau_7$	$\tau_8$				
		RB	0.002	0.003	-0.005	0.004	0.006	0.003	-0.012	-0.001				
	HB	RMSE	0.193	0.230	0.226	0.208	0.188	0.183	0.333	0.174				
		CP	0.957	0.951	0.917	0.959	0.960	0.962	0.912	0.979				

余模型参数的先验分布与第 4.3.4.1 节的设定一致. 图 4.9 给出了各模型对系统 MTTF 估计上的 RMSE 对比结果. 从图中可以看出, 所提模型在贝叶斯方法下所得的 RMSE 显著低于其他基准模型. 这一结果表明, 所提模型在预测系统 MTTF 方面具有更高的精度, 从而验证了其在可靠性估计中的显著优势.

#### 4.3.4.3 变点估计的性能评估

为进一步突出所提模型在实时环境下对变点检测的优势, 本研究基于情形 II ( $n_i = 40$ ) 进行模拟实验. 设退化数据为动态获取, 在每个阶段使用当前观测数据  $y_{i,1:j}$  对模型参数进行更新估计, 并据此判断变点位置. 图 4.10 展示了在  $j = 20, 30, 40$  三个时间点下, 变点估计的平均 RMSE. 结果表明, 随着可用观测数据量的增加, 变点估计的 RMSE 持续下降, 且两种方法在各阶段均保持较小的 RMSE, 说明所提模型能够有效识别变点位置. 值得注意的是, 贝叶斯方法在变点检测方面表现出更高的精度, 其在所有阶段下的平均 RMSE 均显著低于 ML 方法. 该结果进一步验证了贝叶斯方法在动态监测退化过程中的鲁棒性与优越性

### 4.3.5 实例分析

本小节采用锂离子电池退化数据集对所提方法的实际应用性能进行验证 (见图 1.10). 锂离子电池广泛应用于手机、电动汽车等各类商业产品中, 其性能退化或故障可能导致设备效能下降, 甚至完全失效. 因此, 准确预测电池的 RUL 分布, 并据此制定科学的维修与更换策略, 对于提升设备运行的可靠性和降低维护成本具有重要意义 (Peng Weiwen 等, 2018; Zhang 等, 2023).

#### 4.3.5.1 模型拟合与可靠性分析

首先, 采用所提的两阶段 rIG 模型对退化数据进行拟合, 并分别通过贝叶斯方法和 ML 方法进行参数估计. 两种方法的设置与模拟实验中的保持一致. 为验证 EM 算法和 ARMS-Gibbs 采样算法的收敛性, 本节附录 D 展示了参数估计的迭代过程、后验样本的轨迹图以及遍历均值图. 结果表明, 两种算法均具有良好的收敛性, 表现出快速且稳定的收敛趋势. 表 ?? 给出了模型参数和变点的估计结果. 以扩散参数  $\gamma$  为例, 贝叶斯方法和 ML 方法的点估计分别为 2.930 和 3.001, 对应的 95% 可信区间 (贝叶斯方法) 与置信区间 (ML 方法)

表 4.7: 基于贝叶斯和 ML 方法的参数和变点估计

电池	贝叶斯方法			极大似然方法		
	$\beta_1$	$\beta_2$	$\tau$	$\beta_1$	$\beta_2$	$\tau$
# 1	2.5%	0.422	2.198	22.257	0.497	2.511
	均值	0.536	2.437	23.187	0.510	2.632
	97.5%	0.645	2.851	24.664	0.518	2.713
# 2	2.5%	0.523	2.013	24.365	0.638	2.113
	均值	0.608	2.356	25.336	0.658	2.215
	97.5%	0.785	2.615	26.557	0.670	2.282
# 3	2.5%	0.336	2.161	26.316	0.405	2.412
	均值	0.468	2.424	26.761	0.414	2.531
	97.5%	0.518	2.831	27.381	0.420	2.61
# 4	2.5%	0.467	1.993	24.151	0.561	2.12
	均值	0.569	2.345	25.008	0.576	2.221
	97.5%	0.703	2.595	26.060	0.587	2.288
# 5	2.5%	0.495	2.162	23.184	0.624	2.382
	均值	0.588	2.418	23.893	0.642	2.496
	97.5%	0.752	2.809	25.370	0.654	2.572
# 6	2.5%	0.464	2.130	24.722	0.559	2.324
	均值	0.566	2.408	25.576	0.574	2.440
	97.5%	0.697	2.769	26.306	0.585	2.517

分别为 (2.615, 3.556) 和 (2.804, 3.165). 这些结果表明, 两种方法均能够提供稳定且可靠的参数估计

此外, 为评估所提模型的预测能力, 选取前 30 个数据点用于模型拟合, 并预测后续 19 个充放电循环后的电池容量变化. 作为对比, 本文引入 Duan 等 (2017) 提出的两阶段 IG 模型 (以下简称 “Duan” 模型), 该模型假设变点仅发生在观测时间点上. 模型中的未知参数通过 ML 方法估计, 并借助 Schwarz 信息准则在拟合精度与模型复杂度之间进行权衡以确定变点位置. 需注意的是, “Duan” 模型的变点位置仅限定于离散的观测点上, 而所提模型则可在连续时间域中直接估计变点位置, 并允许其具备一定的不确定性.

表 ?? 汇总了五种模型在训练阶段、预测阶段及整体性能方面的 RB 与

表 4.8: 不同模型的 RMSE 和 RB.

模型	训练集		预测集		全部数据集	
	RMSE	RB	RMSE	RB	RMSE	RB
所提模型	0.448	0.248	1.538	0.060	1.020	0.175
线性模型	3.476	1.442	3.685	0.156	3.558	0.943
逆幂模型	2.057	0.568	2.475	0.113	2.229	0.391
指数模型	0.908	0.313	1.611	0.065	1.230	0.217
Duan 模型	0.434	0.239	1.976	0.075	1.276	0.175

RMSE, 其中 “Proposed” 表示采用贝叶斯方法进行参数估计的所提模型. 从表中可以看出, 三种不考虑变点的基准模型在预测精度方面表现较差, RMSE 与 RB 值均显著偏高. 例如, 图 4.11 展示了各模型对电池 #2 的拟合与预测路径. 可以观察到, 不包含变点机制的模型拟合能力不足, 导致其在未来退化路径的预测中存在较大误差. 相较之下, 两阶段模型能够有效识别退化过程中的变点, 从而生成与实际观测更为一致的预测结果. 与 “Duan” 模型相比, 所提模型在预测性能上表现更优, 体现为更低的 RMSE 与 RB 值. 这一优势主要得益于其对变点位置的更精准估计. 以电池 #2 为例, 图 4.11 进一步对比了两种模型在变点检测方面的差异. 可以明显看出, 变点估计结果的差异直接影响了第二阶段退化速率的判断, 进而对 RUL 预测的准确性产生显著影响.

基于所提模型的估计结果, 可推导每个电池的失效时间和 RUL 分布, 方法如第 4.3.2.3 节所述. 以贝叶斯估计为例, 使用前 30 个循环数据, 图 4.12 展示了每个电池失效时间的可靠度函数和 PDF, 阈值为  $\mathcal{D} = 20\%$ . 根据式 (4.17), 各电池的 MTTF 分别为 41.984、43.208、43.658、44.588、43.320 和 42.257. 图 4.13 展示了第 30 个循环时 RUL 的可靠度函数和 PDF, 各电池的 MRL 分别为 9.352、13.375、14.354、15.111、13.180 和 11.925.

#### 4.3.5.2 自适应替换策略

基于 RUL 预测结果, 可利用所提自适应替换策略确定每个电池的最佳替换时间. 为展示模型在实时场景中的适用性, 选取前 30 的数据点作为历史数据, 并随着时间推移逐步获取新数据. 在此过程中, 每当获取新的观测数据时, 模型参数和 RUL 分布的估计均会动态更新. 依据期望成本率式 (4.21), 可以

表 4.9: 在连续数据采集周期中, 电池 #2 和 #3 的候选替换时间及最佳准备时间

循环次数 ( $\times 300$ )	电池 #2			电池 #3		
	真实 RUL	MRL	$\mathcal{T}_{2,j}$	真实 RUL	MRL	$\mathcal{T}_{3,j}$
31	12	13.865	43	13	13.228	46
33	10	11.219	41	11	10.278	43
35	8	7.624	41	9	8.389	42
37	6	5.986	41	7	6.884	42
39	4	4.040	42	5	4.206	43
41	2	2.764	43	3	2.318	44
<b>42</b>	1	1.235	43	2	1.556	44
<b>43</b>	-	-	-	1	0.380	44

确定持续数据收集期间每个电池的候选替换时间. 为说明效果, 表 ?? 展示了电池 #2 和 #3 的候选替换时间, 其中各项成本设定为  $c_i = 2, c_c = 600, c_p = 200, c_b = 100$ , 维修准备时间  $\varpi = 1$ . 此外, 表中还包括真实的 RUL 和预测的 MRL. 从表中可以看出, 候选替换时间根据 RUL 预测结果进行动态调整. 电池 #2 和 #3 的最佳准备时间分别为 42 和 43. 一旦准备完成, 最佳替换时间分别为 43 和 44, 均与电池实际失效时间高度一致, 表明应在失效发生前进行预防性维修. 这一结果进一步验证了所提自适应替换策略在动态环境下的有效性与实用性.

为突出模型准确性对自适应维修策略的影响, 将所提两阶段模型的结果与其他模型进行比较. 需要注意, “Duan” 模型假设变点已知, 且未推导出 RUL 分布, 因此仅将所提模型与其他三个不考虑变点的模型 (线性、幂律和指数) 进行比较. 表 ?? 展示了不同模型下 6 个电池的最佳替换时间, 其中 “FC” 表示电池的真实失效时间, “P” 和 “C” 分别对应预防性维修和纠正性维修. 从表中可见, 在自适应替换策略下, 除幂律模型对五个电池执行纠正性维修外, 其他模型中所有电池的最终确定的最佳替换时间均小于 FC, 即成功执行预防性维修.

图 4.14 展示了每种策略的平均成本率, 其中 ARP 代表所提策略. 在 ARP 下, 使用两阶段 rIG 模型的策略称为 ARP-TP. 从图中可以看出, 除 ARP-Power 策略外, 其他基于 RUL 的 ARP 策略均明显优于 CRP. 此外, 值得注意的是, 与

表 4.10: 自适应更换政策下 6 块电池的维修成本率

电池	FC	所提模型		线性模型			逆幂模型			指数模型			
		$\mathcal{T}_i^*$	动作	CR	$\mathcal{T}_i^*$	动作	CR	$\mathcal{T}_i^*$	动作	CR	$\mathcal{T}_i^*$	动作	CR
1	40	37	P	7.351	37	P	7.351	40	P	6.950	35	P	7.657
2	43	43	P	6.605	42	P	6.714	-	C	17.909	40	P	6.950
3	44	44	P	6.500	44	P	6.500	-	C	17.556	42	P	6.714
4	45	44	P	6.500	43	P	6.605	-	C	17.217	41	P	6.829
5	41	40	P	6.950	39	P	7.077	-	C	18.667	38	P	7.211
6	42	42	P	6.714	41	P	6.829	-	C	18.326	40	P	6.950

其他策略相比, ARP-TP 的表现最接近 IRP. 结合表 ??, ARP-TP 提供了接近系统实际故障时间的准确  $\mathcal{T}_i^*$  值, 而不会超过其实际寿命. ARP-TP 的平均成本率相对较低, 归因于所提两阶段模型的有效性, 该模型能够准确捕捉变更点的位置并拟合退化路径, 显著提升了预测和决策的精度.

## 4.3.6 附录

### 4.3.6.1 附录 A: 定理 4.3 和 4.4 证明

首先证明定理 4.3 的结果. 设  $Y_1(t)$  和  $Y_2(t)$  分别是变点  $\tau$  之前和之后的退化过程. 那么有

$$Y(t) = \begin{cases} Y_1(t), & t \leq \tau, \\ Y_1(\tau) + Y_2(t - \tau), & t > \tau. \end{cases}$$

当  $0 \leq t \leq \tau$  时, 在给定  $\tau$  的条件下,  $T$  的可靠度函数, 记作  $\bar{F}_1(t|\tau)$ , 可以写为

$$\bar{F}_1(t | \tau) = P(T > t | \tau \geq t) = P(Y_1(t) < \mathcal{D} | \tau \geq t) = F_{rIG}(\mathcal{D} | \delta_1 t, \gamma). \quad (4.36)$$

当  $t > \tau$  时, 给定  $\tau$  的  $T$  的可靠度函数可以表示为

$$\begin{aligned}\bar{F}_2(t \mid \tau) &= P(Y(t) < \mathcal{D} \mid \tau < t) = P(Y_1(\tau) + Y_2(t - \tau) < \mathcal{D} \mid \tau < t) \\ &= \int_0^{\mathcal{D}} P(Y_2(t - \tau) < \mathcal{D} - y_\tau \mid \tau < t) f_1(y_\tau \mid \tau) dy_\tau \\ &= \int_0^{\mathcal{D}} F_{rIG}(\mathcal{D} - y_\tau \mid \delta_2(t - \tau), \gamma) f_1(y_\tau \mid \tau) dy_\tau,\end{aligned}\quad (4.37)$$

其中  $y_\tau$  表示变点时间  $\tau$  处的退化值,  $f_1(y_\tau \mid \tau)$  是  $y_\tau$  的 PDF. 根据 rIG 过程的性质可知,  $f_1(y_\tau \mid \tau) = f_{rIG}(y_\tau \mid \delta_1 \tau, \gamma)$ . 由式 (4.36) 和式 (4.37) 得到  $T$  的无条件可靠度函数为

$$\begin{aligned}R(t) &= P(Y(t) < \mathcal{D}, \tau \geq t) + P(Y(t) < \mathcal{D}, 0 < \tau < t) \\ &= \bar{F}_1(t \mid \tau) \bar{G}_\tau(t) + \int_0^t g_\tau(\tau \mid \mu_\tau, \sigma_\tau^2) \bar{F}_2(t \mid \tau) d\tau,\end{aligned}\quad (4.38)$$

其中  $\bar{G}_\tau(t)$  是随机变量  $\tau$  的生存函数.

接下来, 证明定理 4.4 的结果. 设  $y_t$  是时间  $t$  处观测到的退化值. 在时间  $t$  处系统  $S_t$  的 RUL 定义为  $S_t = \inf\{x; Y(t+x) \geq \mathcal{D} \mid y_t < \mathcal{D}\}$ , 表示在条件  $y_t < \mathcal{D}$  下系统继续正常运行的最短时间. 为了计算系统在时间  $t+x$  处正常运行的概率, 需计算系统剩余使用寿命的可靠性. 考虑到时间  $t$ 、 $t+x$  和  $\tau$  之间的不同关系, 首先根据  $\tau$  的条件得到了三种不同的可靠度.

(i) 当  $x+t \leq \tau$  时,  $S_t$  的条件可靠度函数为

$$\begin{aligned}\bar{F}_{S_t,1}(x \mid \tau) &= P(Y(t+x) < \mathcal{D} \mid y_t < \mathcal{D}, x+t \leq \tau) \\ &= P(Y(t+x) - y_t < \mathcal{D} - y_t \mid y_t < \mathcal{D}, x+t \leq \tau) \\ &= F_{rIG}(\mathcal{D} - y_t \mid \delta_1 x, \gamma).\end{aligned}\quad (4.39)$$

(ii) 当  $t < \tau < x+t$  时,  $S_t$  的条件可靠度函数为

$$\begin{aligned}\bar{F}_{S_t,2}(x \mid \tau) &= P(Y(t+x) < \mathcal{D} \mid y_t < \mathcal{D}, t < \tau < x+t) \\ &= P(Y_2(t+x-\tau) + Y_1(\tau) < \mathcal{D} \mid y_t < \mathcal{D}, t < \tau < x+t) \\ &= \int_0^{\mathcal{D}} F_{rIG}(\mathcal{D} - y_\tau \mid \delta_2(t+x-\tau), \gamma) f_1(y_\tau \mid \tau) dy_\tau.\end{aligned}\quad (4.40)$$

(iii) 当  $\tau \leq t$  时,  $S_t$  的条件可靠度函数为

$$\bar{F}_{S_t,3}(x \mid \tau) = F_{r\mathcal{IG}}(\mathcal{D} - y_t \mid \delta_2 x, \gamma). \quad (4.41)$$

基于式 (4.39) - 式 (4.41), RUL 的无条件可靠度函数是

$$\begin{aligned} R_{S_t}(x) &= P(Y(t+x) < \mathcal{D}, t < x+t \leq \tau) \\ &\quad + P(Y(t+x) < \mathcal{D}, t \leq \tau < x+t) + P(Y(t+x) < \mathcal{D}, t > \tau) \\ &= \bar{F}_{S_t,1}(x \mid \tau) \bar{G}_\tau(x+t) + \int_t^{x+t} g_\tau(\tau \mid \mu_\tau, \sigma_\tau^2) \bar{F}_{S_t,2}(x \mid \tau) d\tau \\ &\quad + \int_0^t g_\tau(\tau) \bar{F}_{S_t,3}(x \mid \tau) d\tau. \end{aligned} \quad (4.42)$$

#### 4.3.6.2 附录 B: EM 算法技术细节

为了进一步解释 EM 算法的技术细节, 首先定义一组符号. 对数似然函数 (4.27) 可以根据  $\tau_i$  分为两部分, 即

$$l_i(\boldsymbol{\theta}_\tau) = \mathbf{v}_i^\top(\tau_i) \mathbf{w}_i(\boldsymbol{\theta}_\tau) \quad \text{和} \quad l_{i,j}(\boldsymbol{\eta}, \tau_i) = \sum_{k=1}^3 \lambda_{i,j}^{(k)} \mathbf{v}_{i,j}^{(k)\top}(\tau_i) \mathbf{w}_{i,j}^{(k)}(\boldsymbol{\eta}),$$

其中

$$\begin{aligned} \mathbf{v}_i(\tau_i) &= (1, \tau_i, \tau_i^2)^\top, v_{i,j}^{(1)}(\tau_i) = 1, \\ \mathbf{v}_{i,j}^{(2)}(\tau_i) &= (1, \log(\Delta A_{i,j} + \Delta B_i \tau_i), \tau_i, \tau_i^2)^\top, \\ v_{i,j}^{(3)}(\tau_i) &= 1, \\ \mathbf{w}_i(\boldsymbol{\theta}_\tau) &= \left( -\log \sqrt{2\pi} \sigma_\tau - \frac{\mu_\tau^2}{2\sigma_\tau^2}, \frac{\mu_\tau}{\sigma_\tau^2}, -\frac{1}{2\sigma_\tau^2} \right)^\top, \end{aligned}$$

$$\begin{aligned}
w_{i,j}^{(1)}(\boldsymbol{\eta}) &= -\log \sqrt{2\pi} + \log \delta_{1,i} \Delta t_{i,j} + \gamma \delta_{1,i} \Delta t_{i,j} - \frac{3}{2} \log \Delta y_{i,j} \\
&\quad - \frac{(\delta_{1,i} \Delta t_{i,j})^2}{2 \Delta y_{i,j}} - \frac{\gamma^2 \Delta y_{i,j}}{2}, \\
\mathbf{w}_{i,j}^{(2)}(\boldsymbol{\eta}) &= \left( -\log \sqrt{2\pi} - \frac{3}{2} \log \Delta y_{i,j} - \frac{\gamma^2 \Delta y_{i,j}}{2} \right. \\
&\quad \left. + \gamma \Delta A_{i,j} - \frac{\Delta A_{i,j}^2}{2 \Delta y_{i,j}}, 1, \Delta B_i \gamma - \frac{\Delta A_{i,j} \Delta B_i}{\Delta y_{i,j}}, -\frac{\Delta B_i^2}{2 \Delta y_{i,j}} \right)^\top, \\
w_{i,j}^{(3)}(\boldsymbol{\eta}) &= -\log \sqrt{2\pi} + \log \delta_{2,i} \Delta t_{i,j} + \gamma \delta_{2,i} \Delta t_{i,j} - \frac{3}{2} \log \Delta y_{i,j} \\
&\quad - \frac{[\delta_{2,i} \Delta t_{i,j}]^2}{2 \Delta y_{i,j}} - \frac{\gamma^2 \Delta y_{i,j}}{2},
\end{aligned}$$

$$\Delta A_{i,j} = \delta_{2,i} t_{i,j} - \delta_{1,i} t_{i,j-1} \text{ 和 } \Delta B_i = \delta_{1,i} - \delta_{2,i}.$$

#### 4.3.6.2.1 附录 B-1: E 步中条件期望的推导

在 E 步中, 需要计算关于  $p(\tau_i \mid \Delta \mathbf{y}_i)$  的期望. 为简化说明, 省略对参数  $\boldsymbol{\vartheta}$  的依赖. 根据退化增量的独立性,  $\Delta \mathbf{Y}_i$  和  $\tau_i$  的联合 PDF 为

$$f_{\Delta \mathbf{Y}_i, \tau_i}(\Delta \mathbf{y}_i, \tau_i) = \prod_{j=1}^{n_i} f_{i,j}(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i) g_\tau(\tau_i \mid \boldsymbol{\theta}_\tau). \quad (4.43)$$

从式 (4.43) 中积分掉  $\tau_i$  后,  $\Delta \mathbf{Y}_i$  的边际 PDF 为

$$f_{\Delta \mathbf{Y}_i}(\Delta \mathbf{y}_i) = \int_{-\infty}^{+\infty} \prod_{j=1}^{n_i} f_{i,j}(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i) g_\tau(\tau_i \mid \boldsymbol{\theta}_\tau) d\tau_i. \quad (4.44)$$

在计算条件 PDF 时, 需考虑变点  $\tau_i$ 、时间点  $t_{i,j}$  以及  $t_{i,j-1}$  之间的三种关系, 因为会得到不同形式的  $\Delta m_{i,j}(\delta_{1,i}, \delta_{2,i}, \tau_i)$ . 为简化表示, 根据三种情形分别

引入了不同的条件 PDF. 具体来说, 当  $\tau_i < t_{i,0}$  时,  $\Delta y_{i,j}$  的条件 PDF 为:

$$f_{i,j}(\Delta y_{i,j} | \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i) = \frac{\Delta m_{i,j}^{(3)}(\delta_{1,i}, \delta_{2,i}, \tau_i)}{\sqrt{2\pi}} \exp \left\{ \gamma \Delta m_{i,j}^{(3)}(\delta_{1,i}, \delta_{2,i}, \tau_i) \right\} \Delta y_{i,j}^{-3/2} \\ \times \exp \left\{ -\frac{[\Delta m_{i,j}^{(3)}(\delta_{1,i}, \delta_{2,i}, \tau_i)]^2}{2} \Delta y_{i,j}^{-1} + \gamma^2 \Delta y_{i,j} \right\}, \quad (4.45)$$

将  $\Delta y_{i,j}$  的条件 PDF 记为  $f_{i,j|(3)}(\Delta y_{i,j} | \delta_{2,i}, \gamma, \tau_i)$ , 其中  $\Delta m_{i,j}(\delta_{1,i}, \delta_{2,i}, \tau_i) = \Delta m_{i,j}^{(3)}(\delta_{1,i}, \delta_{2,i}, \tau_i)$ . 类似地, 对于  $t_{i,j-1} \leq \tau_i < t_{i,j}$  和  $\tau_i \geq t_{i,n_i}$  的情况,  $\Delta y_{i,j}$  的条件 PDF 分别记为  $f_{i,j|(2)}(\Delta y_{i,j} | \delta_{2,i}, \gamma, \tau_i)$  和  $f_{i,j|(1)}(\Delta y_{i,j} | \delta_{2,i}, \gamma, \tau_i)$ , 其中

$$\Delta m_{i,j}(\delta_{1,i}, \delta_{2,i}, \tau_i) = \Delta m_{i,j}^{(2)}(\delta_{1,i}, \delta_{2,i}, \tau_i), \text{ 以及 } \Delta m_{i,j}(\delta_{1,i}, \delta_{2,i}, \tau_i) = \Delta m_{i,j}^{(1)}(\delta_{1,i}, \delta_{2,i}, \tau_i).$$

接下来, 将式 (4.44) 中的  $\prod_{j=1}^{n_i} f_{i,j}(\Delta y_{i,j} | \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i)$  分解为以下三种情况:

(1)  $\Delta \tau_i < t_{i,0}$ ,

$$\prod_{j=1}^{n_i} f_{i,j}(\Delta y_{i,j} | \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i) = \prod_{j=1}^{n_i} f_{i,j|(3)}(\Delta y_{i,j} | \delta_{2,i}, \gamma, \tau_i), \quad (4.46)$$

$$\triangleq L_i(\Delta \mathbf{y}_i | \boldsymbol{\eta}, \tau_i).$$

(2)  $\Delta t_{i,j-1} \leq \tau_i < t_{i,j}$ ,  $j = 1, \dots, n_i$ ,

$$\prod_{j=1}^{n_i} f_{i,j}(\Delta y_{i,j} | \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i) = \left\{ \prod_{j'=1}^{j-1} f_{i,j'|1}(\Delta y_{i,j'} | \delta_{1,i}, \gamma, \tau_i) \right\} \\ \times f_{i,j|(2)}(\Delta y_{i,j} | \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i), \\ \times \left\{ \prod_{j'=j+1}^{n_i} f_{i,j'|3}(\Delta y_{i,j'} | \delta_{2,i}, \gamma, \tau_i) \right\} \\ \triangleq M_{ij}(\Delta \mathbf{y}_i | \boldsymbol{\eta}, \tau_i). \quad (4.47)$$

(3) 当  $\tau_i \geq t_{i,n_i}$ ,

$$\begin{aligned} \prod_{j=1}^{n_i} f_{i,j}(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i) &= \prod_{j=1}^{n_i} f_{i,j|(1)}(\Delta y_{i,j} \mid \delta_{2,i}, \gamma, \tau_i), \\ &\triangleq R_i(\Delta \mathbf{y}_i \mid \boldsymbol{\eta}, \tau_i). \end{aligned} \quad (4.48)$$

因此, 式 (4.44) 中的  $\Delta \mathbf{Y}_i$  的边际 PDF 可以重写为:

$$\begin{aligned} f_{\Delta \mathbf{Y}_i}(\Delta \mathbf{y}_i) &= \int_{-\infty}^{+\infty} \prod_{j=1}^{n_i} f_{i,j}(\Delta y_{i,j} \mid \delta_{1,i}, \delta_{2,i}, \gamma, \tau_i) g_{\tau}(\tau_i \mid \boldsymbol{\theta}_{\tau}) d\tau_i \\ &= \int_{-\infty}^{t_{i,0}} L_i(\Delta \mathbf{y}_i \mid \boldsymbol{\eta}, \tau_i) g_{\tau}(\tau_i \mid \boldsymbol{\theta}_{\tau}) d\tau_i + \sum_{j=1}^{n_i} \int_{t_{i,j-1}}^{t_{i,j}} M_{i,j}(\Delta \mathbf{y}_i \mid \boldsymbol{\eta}, \tau_i) g_{\tau}(\tau_i \mid \boldsymbol{\theta}_{\tau}) d\tau_i \\ &\quad + \int_{t_{i,n_i}}^{+\infty} R_i(\Delta \mathbf{y}_i \mid \boldsymbol{\eta}, \tau_i) g_{\tau}(\tau_i \mid \boldsymbol{\theta}_{\tau}) d\tau_i, \end{aligned} \quad (4.49)$$

根据贝叶斯定理, 可以计算条件 PDF  $p(\tau_i \mid \Delta \mathbf{y}_i)$  为:

$$p(\tau_i \mid \Delta \mathbf{y}_i) = \frac{f_{\Delta \mathbf{Y}_i, \tau_i}(\Delta \mathbf{y}_i, \tau_i)}{f_{\Delta \mathbf{Y}_i}(\Delta \mathbf{y}_i)}. \quad (4.50)$$

然后, 可以推导出条件期望  $\mathbb{E}_{\boldsymbol{\theta}_{(s)}}[\mathbf{v}_i \mid \Delta \mathbf{y}_i]$  和  $\mathbb{E}_{\boldsymbol{\theta}_{(s)}}[\lambda_{i,j}^{(k)} v_{i,j}^{(k)} \mid \Delta \mathbf{y}_i]$ . 对于  $\mathbf{v}_i$  的条件期望,  $i = 1, \dots, I$ , 可表示为:

$$\begin{aligned} \mathbb{E}_{\boldsymbol{\theta}_{(s)}}\{\mathbf{v}_i \mid \Delta \mathbf{y}_i\} &= \frac{1}{f_{\Delta \mathbf{Y}_i}(\Delta \mathbf{y}_i)} \left( \int_{-\infty}^{t_{i,0}} \mathbf{v}_i(\tau_i) L_i(\Delta \mathbf{y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i) g_{\tau}(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}) d\tau_i \right. \\ &\quad + \sum_{j=1}^{n_i} \int_{t_{i,j-1}}^{t_{i,j}} \mathbf{v}_i(\tau_i) M_{i,j}(\Delta \mathbf{y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i) g_{\tau}(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}) d\tau_i \\ &\quad \left. + \int_{t_{i,n_i}}^{+\infty} \mathbf{v}_i(\tau_i) R_i(\Delta \mathbf{y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i) g_{\tau}(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}) d\tau_i \right). \end{aligned} \quad (4.51)$$

$\lambda_{i,j}^{(k)} v_{i,j}^{(k)}$  的条件期望, 其中  $i = 1, \dots, I, j = 1, \dots, n_i, k = 1, 2, 3$ , 如下所示: 当

$k = 1$  时:

$$\begin{aligned} & \mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} \left\{ \lambda_{i,j}^{(1)} v_{i,j}^{(1)} \mid \Delta \mathbf{y}_i \right\} \\ &= \frac{1}{f_{\Delta \mathbf{Y}_i}(\Delta \mathbf{y}_i)} \left( \sum_{j'=j+1}^{n_i} \int_{t_{i,j'-1}}^{t_{i,j'}} v_{i,j'}^{(1)}(\tau_i) M_{i,j'}(\Delta \mathbf{y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i) g_{\tau}(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}) d\tau_i \right. \\ & \quad \left. + \int_{t_{i,n_i}}^{\infty} v_{i,j}^{(1)}(\tau_i) R_i(\Delta \mathbf{y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i) g_{\tau}(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}) d\tau_i \right), \end{aligned} \quad (4.52)$$

当  $k = 2$  时:

$$\begin{aligned} & \mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} \left\{ \lambda_{i,j}^{(2)} \mathbf{v}_{i,j}^{(2)} \mid \Delta \mathbf{y}_i \right\} \\ &= \frac{1}{f_{\Delta \mathbf{Y}_i}(\Delta \mathbf{y}_i)} \left( \int_{t_{i,j-1}}^{t_{i,j}} \mathbf{v}_{i,j}^{(2)}(\tau_i) M_{i,j}(\Delta \mathbf{y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i) g_{\tau}(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}) d\tau_i \right), \end{aligned} \quad (4.53)$$

当  $k = 3$  时:

$$\begin{aligned} & \mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} \left\{ \lambda_{i,j}^{(3)} v_{i,j}^{(3)} \mid \Delta \mathbf{y}_i \right\} \\ &= \frac{1}{f_{\Delta \mathbf{Y}_i}(\Delta \mathbf{y}_i)} \left( \int_{-\infty}^{t_{i,0}} v_{i,j}^{(3)}(\tau_i) L_i(\Delta \mathbf{y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i) g_{\tau}(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}) d\tau_i \right. \\ & \quad \left. + \sum_{j'=1}^{j-1} \int_{t_{i,j'-1}}^{t_{i,j'}} v_{i,j'}^{(3)}(\tau_i) M_{i,j'}(\Delta \mathbf{y}_i \mid \boldsymbol{\eta}_{(s)}, \tau_i) g_{\tau}(\tau_i \mid \boldsymbol{\theta}_{\tau(s)}) d\tau_i \right). \end{aligned} \quad (4.54)$$

因此, E 步中的 Q 函数为:

$$\begin{aligned} \mathbf{Q}_{(s)}(\boldsymbol{\vartheta}) &= \sum_{i=1}^I \mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} \{ \mathbf{v}_i \mid \Delta \mathbf{y} \}^{\top} \mathbf{w}_i(\boldsymbol{\theta}_{\tau}) \\ & \quad + \sum_{i=1}^I \sum_{j=1}^{n_i} \sum_{k=1}^3 \mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} \left\{ \lambda_{i,j}^{(k)} v_{i,j}^{(k)} \mid \Delta \mathbf{y} \right\}^{\top} \mathbf{w}_{i,j}^{(k)}(\boldsymbol{\eta}). \end{aligned} \quad (4.55)$$

#### 4.3.6.2.2 附录 B-2: M 步中关于 $\boldsymbol{\vartheta}$ 的一阶偏导数

对式 (4.55) 中的 Q 函数关于  $\boldsymbol{\vartheta}$  求一阶偏导数, 并将导数设为零, 从而得到

$\boldsymbol{\vartheta}_{(s+1)}$  的估计.

$$\begin{aligned}
 \frac{\partial \mathbf{Q}_{(s)}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\theta}_\tau} &= \sum_{i=1}^I \left[ \frac{\partial \mathbf{w}_i(\boldsymbol{\theta}_\tau)}{\partial \boldsymbol{\theta}_\tau} \right]^\top \mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} \{ \mathbf{v}_i \mid \Delta \mathbf{y} \} = \mathbf{0}, \\
 \frac{\partial \mathbf{Q}_{(s)}(\boldsymbol{\vartheta})}{\partial (\delta_{1,i}, \delta_{2,i})^\top} &= \sum_{j=1}^{n_i} \sum_{k=1}^3 \left[ \frac{\partial \mathbf{w}_{i,j}^{(k)}(\boldsymbol{\eta})}{\partial (\delta_{1,i}, \delta_{2,i})^\top} \right]^\top \mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} \{ \lambda_{i,j}^{(k)} v_{i,j}^{(k)} \mid \Delta \mathbf{y} \} = \mathbf{0}, \quad i = 1, \dots, I, \\
 \frac{\partial \mathbf{Q}_{(s)}(\boldsymbol{\vartheta})}{\partial \gamma} &= \sum_{i=1}^I \sum_{j=1}^{n_i} \sum_{k=1}^3 \left[ \frac{\partial \mathbf{w}_{i,j}^{(k)}(\boldsymbol{\eta})}{\partial \gamma} \right]^\top \mathbb{E}_{\boldsymbol{\vartheta}_{(s)}} \{ \lambda_{i,j}^{(k)} v_{i,j}^{(k)} \mid \Delta \mathbf{y} \} = 0,
 \end{aligned} \tag{4.56}$$

其中

$$\begin{aligned}
 \frac{\partial \mathbf{w}_i(\boldsymbol{\theta}_\tau)}{\partial \boldsymbol{\theta}_\tau} &= \begin{pmatrix} -\frac{\mu_\tau}{\sigma_\tau^2}, & -\frac{1}{2\sigma_\tau^2} + \frac{\mu_\tau^2}{2\sigma_\tau^4} \\ \frac{1}{\sigma_\tau^2}, & \frac{-\mu_\tau}{\sigma_\tau^4} \\ 0, & \frac{1}{2\sigma_\tau^4} \end{pmatrix}, \\
 \frac{\partial w_{i,j}^{(1)}(\boldsymbol{\eta})}{\partial (\delta_{1,i}, \delta_{2,i})^\top} &= \left( \left( \frac{1}{\delta_{1,i}} + \gamma \right) \Delta t_{i,j} - \frac{\delta_{1,i} \Delta t_{i,j}^2}{\Delta y_{i,j}}, 0 \right), \\
 \frac{\partial \mathbf{w}_i(\boldsymbol{\theta}_\tau)}{\partial \boldsymbol{\theta}_\tau} &= \begin{pmatrix} -\frac{\mu_\tau}{\sigma_\tau^2}, & -\frac{1}{2\sigma_\tau^2} + \frac{\mu_\tau^2}{2\sigma_\tau^4} \\ \frac{1}{\sigma_\tau^2}, & \frac{-\mu_\tau}{\sigma_\tau^4} \\ 0, & \frac{1}{2\sigma_\tau^4} \end{pmatrix}, \\
 \frac{\partial w_{i,j}^{(1)}(\boldsymbol{\eta})}{\partial (\delta_{1,i}, \delta_{2,i})^\top} &= \left( \left( \frac{1}{\delta_{1,i}} + \gamma \right) \Delta t_{i,j} - \frac{\delta_{1,i} \Delta t_{i,j}^2}{\Delta y_{i,j}}, 0 \right), \\
 \frac{\partial \mathbf{w}_{i,j}^{(2)}(\boldsymbol{\eta})}{\partial (\delta_{1,i}, \delta_{2,i})^\top} &= \begin{pmatrix} -\gamma t_{i,j-1} + \frac{\Delta A_{i,j} t_{i,j-1}}{\Delta y_{i,j}}, & \gamma t_{i,j} - \frac{\Delta A_{i,j} t_{i,j}}{\Delta y_{i,j}} \\ 0, & 0 \\ \gamma + \frac{\Delta B_i t_{i,j-1} - \Delta A_{i,j}}{\Delta y_{i,j}}, & -\gamma + \frac{\Delta A_{i,j} - \Delta B_i t_{i,j}}{\Delta y_{i,j}} \\ \frac{\delta_{2,i} - \delta_{1,i}}{\Delta y_{i,j}}, & \frac{\delta_{1,i} - \delta_{2,i}}{\Delta y_{i,j}} \end{pmatrix}, \\
 \frac{\partial w_{i,j}^{(3)}(\boldsymbol{\eta})}{\partial (\delta_{1,i}, \delta_{2,i})^\top} &= \left( 0, \left( \frac{1}{\delta_{2,i}} + \gamma \right) \Delta t_{i,j} - \frac{\delta_{2,i} \Delta t_{i,j}^2}{\Delta y_{i,j}} \right), \\
 \frac{\partial w_{i,j}^{(1)}(\boldsymbol{\eta})}{\partial \gamma} &= \delta_{1,i} \Delta t_{i,j} - \gamma \Delta y_{i,j},
 \end{aligned}$$

$$\begin{aligned}\frac{\partial \mathbf{w}_{i,j}^{(2)}(\boldsymbol{\eta})}{\partial \gamma} &= \left( -\gamma \Delta y_{i,j} + \Delta A_{i,j}, 0, \Delta B_i, 0 \right)^\top, \\ \frac{\partial w_{i,j}^{(3)}(\boldsymbol{\eta})}{\partial \gamma} &= \delta_{2,i} \Delta t_{i,j} - \gamma \Delta y_{i,j}.\end{aligned}$$

#### 4.3.6.2.3 附录 B-3: EM 算法的步骤

EM 算法可以通过以下步骤实现:

- **步骤 1.** 设定参数  $\boldsymbol{\vartheta}$  的初始值  $\boldsymbol{\vartheta}_{(0)}$ , 并设置容差误差  $\epsilon$ .
- **步骤 2.** 基于第  $s$  次迭代的解  $\boldsymbol{\vartheta}_{(s)}$ , 计算  $\mathbb{E}_{\boldsymbol{\vartheta}_{(s)}}[l_i(\boldsymbol{\theta}_\tau) | \Delta \mathbf{y}]$  和  $\mathbb{E}_{\boldsymbol{\vartheta}_{(s)}}[l_{i,j}(\boldsymbol{\eta}, \boldsymbol{\tau}) | \Delta \mathbf{y}]$ .
- **步骤 3.** 根据式 (4.29) 计算第  $(s+1)$  次迭代的解  $\boldsymbol{\vartheta}_{(s+1)}$ .
- **步骤 4.** 重复步骤 2 和步骤 3, 直到  $|\boldsymbol{\vartheta}_{(s+1)} - \boldsymbol{\vartheta}_{(s)}| < \epsilon$ , 其中  $|\cdot|$  是  $L_1$  距离.
- **步骤 5.** 参数  $\boldsymbol{\vartheta}$  的 MLE 可以通过  $\hat{\boldsymbol{\vartheta}} = \boldsymbol{\vartheta}_{(s+1)}$  获得.

#### 4.3.6.3 附录 C: 贝叶斯分析技术细节

每个参数的满条件后验分布计算如下

- 给定  $\boldsymbol{\theta}_{\setminus(\mu_\tau, \sigma_\tau^2)}$  和  $\Delta \mathbf{Y}, (\mu_\tau, \sigma_\tau^2)$  的满条件后验分布是

$$(\mu_\tau, \sigma_\tau^2) | \boldsymbol{\theta}_{\setminus(\mu_\tau, \sigma_\tau^2)}, \Delta \mathbf{Y} \sim NIGa(\beta'_\tau, \eta'_\tau, v'_\tau, \xi'_\tau),$$

其中  $\beta'_\tau = \beta_\tau + I$ ,  $\eta'_\tau = \left( \beta_\tau \eta_\tau + \sum_{i=1}^I \tau_i \right) / (\beta_\tau + I)$ ,  $v'_\tau = I/2 + v_\tau$ , 和  $\xi'_\tau = \xi_\tau + \beta_\tau \eta_\tau^2/2 + \sum_{i=1}^I \tau_i^2/2 - \left( \beta_\tau \eta_\tau + \sum_{i=1}^I \tau_i \right)^2 / (2(\beta_\tau + I))$ .

- 给定  $\boldsymbol{\theta}_{\setminus(\mu_1, \sigma_1^2)}$  和  $\Delta \mathbf{Y}, (\mu_1, \sigma_1^2)$  的满条件后验分布是

$$(\mu_1, \sigma_1^2) | \boldsymbol{\theta}_{\setminus(\mu_1, \sigma_1^2)}, \Delta \mathbf{Y} \sim NIGa(\beta'_1, \eta'_1, v'_1, \xi'_1),$$

其中  $\beta'_1 = \beta_1 + I$ ,  $\eta'_1 = \left( \beta_1 \eta_1 + \sum_{i=1}^I \delta_i^1 \right) / (\beta_1 + I)$ ,  $v'_1 = I/2 + v_1$  和  $\xi'_1 = \xi_1 + \beta_1 \eta_1^2/2 + \sum_{i=1}^I \delta_{1,i}^2/2 - \left( \beta_1 \eta_1 + \sum_{i=1}^I \delta_i^1 \right)^2 / (2(\beta_1 + I))$ .

- 给定  $\boldsymbol{\theta}_{\setminus(\mu_2, \sigma_2^2)}$  和  $\Delta \mathbf{Y}, (\mu_2, \sigma_2^2)$  的满条件后验分布是

$$(\mu_2, \sigma_2^2) | \boldsymbol{\theta}_{\setminus(\mu_2, \sigma_2^2)}, \Delta \mathbf{Y} \sim NIGa(\beta'_2, \eta'_2, v'_2, \xi'_2),$$

其中  $\beta'_2 = \beta_2 + I$ ,  $\eta'_2 = \left( \beta_2 \eta_2 + \sum_{i=1}^I \delta_{2,i} \right) / (\beta_2 + I)$ ,  $v'_2 = I/2 + v_2$  和  $\xi'_2 = \xi_2 + \beta_2 \eta_2^2 / 2 + \sum_{i=1}^I \delta_{2,i}^2 / 2 - \left( \beta_2 \eta_2 + \sum_{i=1}^I \delta_{2,i} \right)^2 / (2(\beta_2 + I))$ .

- 给定  $\theta_{\setminus \gamma}$  和  $\Delta Y$ ,  $\gamma$  的满条件后验分布是

$$\gamma \mid \theta_{\setminus \gamma}, \Delta Y \sim N(\omega', \kappa') .$$

其中  $\omega' = (\omega + \kappa N) / (1 + \kappa N)$ ,  $\kappa' = \kappa^2 / (1 + \kappa^2 N)$  和  $N = \sum_{i=1}^I \sum_{j=1}^{n_i} \Delta y_{i,j}$ .

- $\delta_{1,i}, i = 1, \dots, I$  的满条件后验分布为:

$$\pi(\delta_{1,i} \mid \theta_{\setminus \delta_{1,i}}, \Delta Y) \propto \exp \left\{ \frac{2\mu_1 \delta_{1,i} - \delta_{1,i}^2}{2\sigma_1^2} + \gamma \sum_{j=1}^{n_i} \Delta \mathcal{M}_{1,i,j} - \sum_{i=1}^I \sum_{j=1}^{n_i} \frac{\Delta \mathcal{M}_{1,i,j}^2}{2\Delta y_{i,j}} \right\} \prod_{j=1}^{n_i} \Delta \mathcal{M}_{1,i,j},$$

其中  $\Delta \mathcal{M}_{1,i,j} = \delta_{1,i} t_{i,j} \lambda_{i,j}^{(1)} + [(\delta_{1,i} - \delta_{2,i}) \tau_i + \delta_{2,i} t_{i,j+1} - \delta_{1,i} t_{i,j}] \lambda_{i,j}^{(2)}$ .

- $\delta_{2,i}, i = 1, \dots, I$  的满条件后验分布为:

$$\pi(\delta_{2,i} \mid \theta_{\setminus \delta_{2,i}}, \Delta Y) \propto \exp \left\{ \frac{2\mu_2 \delta_{2,i} - \delta_{2,i}^2}{2\sigma_2^2} + \gamma \sum_{j=1}^{n_i} \Delta \mathcal{M}_{2,i,j} - \sum_{i=1}^I \sum_{j=1}^{n_i} \frac{\Delta \mathcal{M}_{2,i,j}^2}{2\Delta y_{i,j}} \right\} \prod_{j=1}^{n_i} \Delta \mathcal{M}_{2,i,j},$$

其中  $\Delta \mathcal{M}_{2,i,j} = [(\delta_{1,i} - \delta_{2,i}) \tau_i + \delta_{2,i} t_{i,j+1} - \delta_{1,i} t_{i,j}] \lambda_{i,j}^{(2)} + \delta_{2,i} t_{i,j} \lambda_{i,j}^{(3)}$ .

- 对于  $\tau_i, i = 1, \dots, I$ , 满条件后验分布如下:

$$\pi(\tau_i \mid \theta_{\setminus \tau_i}, \Delta Y) \propto \exp \left\{ \frac{2\mu_\tau \tau_i - \tau_i^2}{2\sigma_\tau^2} + \gamma \sum_{j=1}^{n_i} \Delta \mathcal{M}_{3,i,j} - \sum_{i=1}^I \sum_{j=1}^{n_i} \frac{\Delta \mathcal{M}_{3,i,j}^2}{2\Delta y_{i,j}} \right\} \prod_{j=1}^{n_i} \Delta \mathcal{M}_{3,i,j},$$

其中  $\Delta \mathcal{M}_{3,i,j} = [(\delta_{1,i} - \delta_{2,i}) \tau_i + \delta_{2,i} t_{i,j+1} - \delta_{1,i} t_{i,j}] \lambda_{i,j}^{(2)}$ .

#### 4.3.6.4 附录 D: 案例研究附加结果

对于贝叶斯方法, 通过轨迹图和遍历均值图来监控 ARMS-Gibbs 算法的收敛性, 见图 4.15 和图 4.16. 通过这些图可以确认马尔科夫链已经收敛. 图 4.17 则展示了基于 EM 算法的模型参数估计迭代过程. 从图中可以看出, 经过 100 次迭代后, 参数估计值已经收敛到一个相对稳定的状态.

## 4.4 在线估计与 RUL 预测

现有大多关于 IG 过程的研究主要集中于离线退化数据分析, 侧重于评估分位寿命、平均寿命和可靠度等总体特征 (Chen 等, 2018b; Peng, 2015; Wang X 等, 2010; Ye, Chen N, 2014). 然而, 随着实际应用中实时获取的退化数据日益增多, 对 RUL 预测的动态更新需求日益迫切. 在此背景下, Pan 等 (2016) 基于 IG 过程探讨了 RUL 预测问题. 然而, 其模型假设和估计方法存在以下局限性:

- 平稳性假设: 在大多数应用中, 性能退化通常呈现非线性趋势, 平稳性假设可能导致 RUL 预测结果产生较大偏差.
- 单一系统数据: 当前方法仅基于单个系统的退化数据进行 RUL 预测, 而实际中往往需同时监测多个系统, 其他系统的退化数据同样蕴含对目标系统 RUL 有价值的信息. 合理利用这些信息有助于提升预测精度.
- 计算负担: 在估计 RUL 分布函数中未知参数时, 当前方法要求将所有历史退化数据作为 EM 算法的输入. 这种非递归估计方法在处理海量数据时会显著增加计算负担, 从而难以满足决策过程中对快速 RUL 预测的需求.

为克服上述局限性, 实现对实时退化数据的动态处理和 RUL 的快速预测, 本节介绍一种适用于非平稳 IG 过程的在线 RUL 预测方法. 本节结构如下: 第 4.4.1 节介绍模型设定及相关问题; 第 4.4.2 节提出参数估计与 RUL 预测的在线算法; 第 4.4.3 节将方法扩展至随机效应模型; 第 4.4.4 节通过模拟验证方法性能; 第 4.4.5 节展示实际应用.

### 4.4.1 模型设定

假设系统的性能退化指标  $Y(t) \sim \text{IG}(\Lambda_\beta(t)/\nu, \lambda\Lambda_\beta^2(t))$ , 其中  $\Lambda_\beta(t)$  是时间  $t$  的单调递增函数, 且满足  $\Lambda_\beta(0) = 0$ ,  $\beta$  为待估未知参数. 则退化增

量  $\Delta Y_{ts} = Y(t) - Y(s)$  服从逆高斯分布  $IG(\Delta\Lambda_{ts}/\nu, \lambda\Delta\Lambda_{ts}^2)$ , 其中  $\Delta\Lambda_{ts} = \Lambda_\beta(t) - \Lambda_\beta(s)$ . 假设当前有  $n$  个系统, 在时间点  $0 = t_0 < t_1 < t_2 < \dots < t_m < \dots$ , 对所有系统的性能退化值进行测量. 令  $y_{i,j}$  表示第  $i$  个系统在时间  $t_j$  的退化值,  $\mathbf{Y}_{0:m}^{(i)} = (y_{i,0}, \dots, y_{i,m})$  表示第  $i$  个系统在  $t_m$  之前收集的退化观测数据. 所有系统的观测数据集合表示为  $\mathbf{Y}_{0:m} = (\mathbf{Y}_{0:m}^{(1)}, \dots, \mathbf{Y}_{0:m}^{(n)})$ . 不失一般性, 假设所有系统在相同的时间点进行检测, 本节方法同样适用于系统检测时间点不同的情形. 假设当前时间点为  $t_m$ , 在线 RUL 预测需要解决以下两个问题:

- **离线训练:** 基于历史数据  $\mathbf{Y}_{0:m}$ , 估计未知参数  $\boldsymbol{\theta} = (\nu, \lambda, \beta)$ . 假设某个特定系统的历史退化观测值为  $(y_1, \dots, y_m)$ , 且失效阈值为  $\omega$ . 在  $t_m$  时刻该系统的 RUL 定义为  $\mathcal{X}_m = \inf\{x : Y(x + t_m) \geq \omega \mid y_m < \omega\}$ , 其分布可以利用当前参数估计值  $\hat{\boldsymbol{\theta}}^{(m)} = (\hat{\nu}^{(m)}, \hat{\lambda}^{(m)}, \hat{\beta}^{(m)})$  进行估计.
- **在线更新:** 记  $(y_{1,m+1}, \dots, y_{n,m+1})$  为下一个时间点  $t_{m+1}$  所收集到的系统退化值. 基于新数据以及之前的估计值  $\hat{\boldsymbol{\theta}}^{(m)}$ , 如何高效更新参数估计  $\hat{\boldsymbol{\theta}}^{(m+1)}$ . 该在线算法应通过递归方式更新参数, 避免直接使用完整历史数据  $\mathbf{Y}_{0:m}$  进行重新计算, 以降低计算复杂度.

由于 RUL 分布是参数  $\boldsymbol{\theta}$  的函数, 在每次参数更新后, 可根据新的参数估计值实时预测 RUL. 这种方法不仅适用于单一系统, 还能高效整合多系统间的信息, 提升预测的精度与时效性.

## 4.4.2 在线估计

### 4.4.2.1 当 $\beta$ 已知时, 递归更新 $\hat{\nu}$ 和 $\hat{\lambda}$

高效的在线 RUL 估计需要快速的递归算法来更新参数  $\boldsymbol{\theta} = (\nu, \lambda, \beta)$  的估计值. 为此, 可先基于历史数据  $\mathbf{Y}_{0:m} = (\mathbf{Y}_{0:m}^{(1)}, \dots, \mathbf{Y}_{0:m}^{(n)})$  考察参数估计量. 设  $\Delta\Lambda_j = \Lambda_\beta(t_j) - \Lambda_\beta(t_{j-1})$  和  $\Delta y_{i,j} = y_{i,j} - y_{i,j-1}$ , 其中  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ . 根据 IG 过程的性质, 给定  $\mathbf{Y}_{0:m}$  时,  $\boldsymbol{\theta}$  的似然函数为

$$\begin{aligned} L(\boldsymbol{\theta} \mid \mathbf{Y}_{0:m}) &= \prod_{i=1}^n \prod_{j=1}^m \sqrt{\frac{\lambda \Delta\Lambda_j^2}{2\pi \Delta y_{i,j}^3}} \exp \left\{ -\frac{\lambda}{2\Delta y_{i,j}} (\nu \Delta y_{i,j} - \Delta\Lambda_j)^2 \right\} \\ &= C_0 \lambda^{nm/2} \prod_{j=1}^m \Delta\Lambda_j^n \exp \left\{ -\lambda \sum_{i=1}^n \sum_{j=1}^m \frac{1}{2\Delta y_{i,j}} (\nu \Delta y_{i,j} - \Delta\Lambda_j)^2 \right\}, \end{aligned} \tag{4.57}$$

其中常数  $C_0 = \prod_{i=1}^n \prod_{j=1}^m \sqrt{\frac{1}{2\pi\Delta y_{i,j}^3}}$ . 假设  $\beta$  已知, 则  $\nu$  和  $\lambda$  的 MLE 为 (推导见本节附录 4.4.6)

$$\hat{\nu}^{(m)} = \frac{n\Lambda_\beta(t_m)}{\sum_{i=1}^n y_{i,m}}, \quad \hat{\lambda}^{(m)} = \frac{nm}{\sum_{i=1}^n \sum_{j=1}^m \frac{\Delta\Lambda_j^2}{\Delta y_{i,j}} - \frac{n^2\Lambda_\beta^2(t_m)}{\sum_{i=1}^n y_{i,m}}}. \quad (4.58)$$

上述公式表明  $\nu$  和  $\lambda$  的估计可通过递归公式更新. 当收集到新的退化测量值  $y_{m+1} = (y_1, m+1, \dots, y_{n,m+1})$  时:

$$\hat{\nu}^{(m+1)} = \frac{n\Lambda_\beta(t_{m+1})}{\sum_{i=1}^n y_{i,m+1}}. \quad (4.59)$$

这不需要使用  $\mathbf{Y}_{0:m}$  中的任何信息. 对于  $\hat{\lambda}^{(m+1)}$ , 其分母包含两部分:  $\sum_{i=1}^n \sum_{j=1}^{m+1} \frac{\Delta\Lambda_j^2}{\Delta y_{i,j}}$  和  $\frac{n^2\Lambda_\beta^2(t_{m+1})}{\sum_{i=1}^n y_{i,m+1}}$ . 前者可分解为  $\sum_{i=1}^n \sum_{j=1}^m \frac{\Delta\Lambda_j^2}{\Delta y_{i,j}} + \sum_{i=1}^n \frac{\Delta\Lambda_{m+1}^2}{\Delta y_{i,m+1}}$ , 后者等于  $[\hat{\nu}^{(m+1)}]^2 \sum_{i=1}^n y_{i,m+1}$ . 基于此,  $\hat{\lambda}^{(m+1)}$  的递归公式为

$$\begin{aligned} \hat{\lambda}^{(m+1)} = n(m+1) & \left[ \frac{nm}{\hat{\lambda}^{(m)}} + [\hat{\nu}^{(m)}]^2 \sum_{i=1}^n y_{i,m} \right. \\ & \left. - [\hat{\nu}^{(m+1)}]^2 \sum_{i=1}^n y_{i,m+1} + \sum_{i=1}^n \frac{\Delta\Lambda_{m+1}^2}{\Delta y_{i,m+1}} \right]^{-1}. \end{aligned} \quad (4.60)$$

因此, 只需存储历史数据  $\mathbf{Y}_{0:m}$  中的最后一次退化观测量  $y_m = (y_{1,m}, \dots, y_{n,m})$ , 以及  $\hat{\nu}^{(m)}$  和  $\hat{\lambda}^{(m)}$ , 即可实现参数估计的更新.

**注 6:** Pan 等 (2016) 考虑的单个系统情形, 即  $n = 1$  (单个系统). 根据(4.59), 可知

$$\hat{\nu}^{(m+1)} = \frac{\Lambda_\beta(t_{m+1})}{y_{1,m+1}}.$$

令

$$\psi_m = \sum_{j=1}^m \frac{\Delta\Lambda_j^2}{\Delta y_{1,j}} - \frac{\Lambda_\beta^2(t_m)}{y_{1,m}},$$

则有

$$\psi_{m+1} = \psi_m + \frac{\Delta y_{1,m+1} y_{1,m+1}}{y_{1,m}} \left[ \hat{\nu}^{(m+1)} - \frac{\Delta\Lambda_{m+1}}{\Delta y_{1,m+1}} \right]^2.$$

根据和(4.60), 则可得  $\hat{\lambda}^{(m+1)}$  的更新公式

$$\begin{aligned}\hat{\lambda}^{(m+1)} &= \frac{m+1}{\psi_{m+1}} = \hat{\lambda}^{(m)} \frac{m+1}{m} \frac{\psi_m}{\psi_{m+1}} \\ &= \hat{\lambda}^{(m)} \frac{m+1}{m} \left[ 1 + \frac{\Delta y_{1,m+1} y_{1,m+1}}{\psi_m y_{1,m}} \left[ \hat{\nu}^{(m+1)} - \frac{\Delta \Lambda_{m+1}}{\Delta y_{1,m+1}} \right]^2 \right]^{-1}.\end{aligned}$$

#### 4.4.2.2 基于一步近似的 $\hat{\beta}$ 递归更新

在上一小节中, 给定参数  $\beta$  时, 得到参数估计  $\hat{\nu}$  和  $\hat{\lambda}$  的递归公式. 当  $\beta$  未知时, 可通过历史数据  $\mathbf{Y}_{0:m}$  推导  $\beta$  的轮廓似然函数(Profile likelihood function), 并利用  $\hat{\beta}^{(m)}$  的 MLE 更新  $\hat{\nu}^{(m+1)}$  和  $\hat{\lambda}^{(m+1)}$ . 另一方面,  $\hat{\beta}^{(m+1)}$  也可通过更新后的  $\hat{\nu}^{(m+1)}$  和  $\hat{\lambda}^{(m+1)}$  进一步优化. 然而, 由于  $\hat{\beta}$  的 MLE 缺乏解析表达式, 直接更新算法较为复杂. 因此, 本小节介绍了一种基于一步近似 (one-step approximation) 的高效递归算法来更新  $\hat{\beta}$ . 其基本思想是借用一步估计 (one-step estimator) 的方法: 给定一个初始估计  $\tilde{\theta}$ , 一步估计量  $\hat{\theta}$  表示为

$$\hat{\theta} = \tilde{\theta} + [I(\tilde{\theta})]^{-1} \dot{L}(\tilde{\theta}), \quad (4.61)$$

其中  $I(\cdot)$  是 Fisher 信息矩阵,  $\dot{L}(\cdot)$  是得分函数, 定义为  $\dot{L}(\theta) = \frac{\partial \log L(\theta)}{\partial \theta}$ ,  $L(\cdot)$  为似然函数. 如果  $\tilde{\theta}$  是  $\sqrt{n}$  一致的, 并且  $\theta \mapsto \dot{L}(\theta)$  满足一定的可微性条件, 则一步估计量  $\hat{\theta}$  也是  $\sqrt{n}$  一致的, 且其渐近方差可达到 Cramér-Rao 下界 (Van der Vaart, 1998).

这里假设  $\nu$  和  $\lambda$  已知来说明更新  $\hat{\beta}$  的方法, 具体实施时可通过循环迭代依次更新  $(\hat{\nu}, \hat{\lambda})$  和  $\hat{\beta}$ . 设  $\hat{\beta}^{(m)}$  为基于观测数据  $\mathbf{Y}_{0:m}$  得到的参数  $\beta$  的一致估计. 若初始估计为 MLE 估计, 则该估计的一致性可以得到保障. 将  $\hat{\beta}^{(m)}$  作为式 (4.61) 中的初始估计, 可以得到一步估计量  $\hat{\beta}^{(m+1)}$  的表达式为

$$\hat{\beta}^{(m+1)} = \hat{\beta}^{(m)} + V_{m+1}(\hat{\beta}^{(m)}) \frac{1}{n} \dot{L}(\nu, \lambda, \hat{\beta}^{(m)} \mid \mathbf{Y}_{0:m+1}), \quad (4.62)$$

其中  $V_{m+1}(\beta)$  是由  $\mathbf{Y}_{0:m+1}$  贡献的 Fisher 信息矩阵的逆,  $L(\cdot)$  为式 (4.57) 中的似然函数. 给定时刻  $t_j$  测量的增量  $\Delta \mathbf{y}_j = (\Delta y_{1,j}, \dots, \Delta y_{n,j})$ ,  $1 \leq j \leq m$ ,  $\beta$  的

对数似然函数为

$$l_j(\beta | \Delta \mathbf{y}_j, \nu, \lambda) = \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) + n \log(\Delta \Lambda_j) - \frac{3}{2} \sum_{i=1}^n \log(\Delta y_{i,j}) \\ - \sum_{i=1}^n \frac{\lambda}{2\Delta y_{i,j}} (\nu \Delta y_{i,j} - \Delta \Lambda_j)^2.$$

由第  $j$  个增量贡献的  $\beta$  的 Fisher 信息为

$$I_j(\beta | \nu, \lambda) = -\mathbb{E} \left[ \frac{\partial^2 l_j(\beta | \Delta \mathbf{y}_j, \nu, \lambda)}{\partial \beta^2} \right] \\ = n \left[ \frac{\nu \lambda (\Delta \Lambda'_j)^2}{\Delta \Lambda_j} + 2 \left( \frac{\Delta \Lambda'_j}{\Delta \Lambda_j} \right)^2 \right], \quad (4.63)$$

其中  $\Delta \Lambda'_j$  是  $\Delta \Lambda_j$  关于  $\beta$  的导数 (详见本节附录 4.2.5). 根据 IG 过程的性质, 增量  $\Delta y_{i,j}, i = 1, \dots, n, j = 1, \dots, m$ , 是独立的但非同分布. 历史数据  $\mathbf{Y}_{0:m}$  对  $\beta$  的 Fisher 信息贡献为  $\sum_{j=1}^m I_j(\beta | \nu, \lambda)$ , 因此式 (4.62) 中的  $V_{m+1}(\hat{\beta}^{(m)})$  可以表示为  $\left[ \sum_{j=1}^{m+1} I_j(\hat{\beta}^{(m)} | \nu, \lambda) \right]^{-1}$ . 由于  $I_j$  需要针对不同时刻下的  $\beta$  估计值进行重新计算,  $V_{m+1}(\hat{\beta}^{(m)})$  无法通过  $V_m(\hat{\beta}^{(m-1)})$  迭代更新. 为解决这一问题, 需考虑一个高效递归公式来近似  $V_{m+1}(\hat{\beta}^{(m)})$ : 结合每个时间点上  $\nu$  和  $\lambda$  的估计, 可用  $\tilde{V}_{m+1} = \left[ \sum_{j=1}^{m+1} I_j(\beta^{(j-1)} | \nu^{(j-1)}, \lambda^{(j-1)}) \right]^{-1}$  近似  $V_{m+1}(\hat{\beta}^{(m)})$ .

可发现这种处理方式会有以下的迭代关系

$$\tilde{V}_{m+1}^{-1} = \tilde{V}_m^{-1} + I_{m+1}(\hat{\beta}^{(m)} | \hat{\nu}^{(m)}, \hat{\lambda}^{(m)}). \quad (4.64)$$

基于式 (4.64),  $\beta$  估计的递归公式 (4.62) 可近似为

$$\hat{\beta}^{(m+1)} = \hat{\beta}^{(m)} + \frac{1}{n} \tilde{V}_{m+1} \dot{L}(\nu, \lambda, \hat{\beta}_n^{(m)} | \mathbf{Y}_{0:m+1}), \quad (4.65)$$

另一方面, 基于  $\mathbf{Y}_{0:m}$  所得的  $\hat{\beta}^{(m)}$  具有  $\sqrt{n}$  一致性, 则有

$$\dot{L}(\nu, \lambda, \hat{\beta}_n^{(m)} | \mathbf{Y}_{0:m}) = o_P(n^{-1/2}),$$

这表明可以忽略与历史数据相关的部分, 仅需考虑新增观测  $\Delta \mathbf{y}_{m+1}$  对似然

函数的贡献. 因此, 式 (4.65) 可进一步简化为

$$\hat{\beta}^{(m+1)} = \hat{\beta}^{(m)} + \frac{1}{n} \tilde{V}_{m+1} S_{m+1}(\Delta \mathbf{y}_{m+1}, \hat{\beta}^{(m)}), \quad (4.66)$$

其中  $S_j(\mathbf{y}, \beta) = \frac{\partial l_j(\beta | \mathbf{y}, \nu, \lambda)}{\partial \beta}$ . 显然,  $S_{m+1}(\Delta \mathbf{y}_{m+1}, \hat{\beta}^{(m)})$  仅依赖于新增观测  $\Delta \mathbf{y}_{m+1}$ .

**注 7:** 对于非线性和非高斯模型, 粒子滤波器常用来处理在线估计问题 (Ma Y 等, 2019). 然而, 随着粒子数量的增加, 其计算复杂度显著提高, 这对实时应用构成挑战. 此外, 粒子退化可能导致估计偏差, 而重采样过程中的噪声可能进一步降低性能 (Peng 等, 2024). 相比之下, 上述提出的一步更新方法具有更高的计算效率, 通过递归计算, 仅利用新增观测和上一步的估计即可实现估计的更新, 显著减少了存储需求和计算负担.

#### 4.4.2.3 在线算法和 RUL 估计

本小节介绍估计参数  $\theta = (\nu, \lambda, \beta)$  的整体在线算法. 首先需要初始化参数值. 正如本节附录 4.4.6 所示, 当系统的性能退化仅被测量一次 (即  $m = 1$ ) 时, 参数  $\beta$  是不可识别的. 同时数值实验也表明, 基于前三个检测时间点的观测数据进行离线训练比  $m = 2$  的情况更稳健. 因此将基于前三次观测数据得到的 MLE  $\hat{\theta}^{(3)} = (\hat{\nu}^{(3)}, \hat{\lambda}^{(3)}, \hat{\beta}^{(3)})$  作为算法的初始值, 其具体计算步骤详见本节附录 4.4.6. 之后, 当新的退化测量数据到来时, 可通过式 (4.59), (4.60) 和 (4.66) 依次更新  $\hat{\nu}$ ,  $\hat{\lambda}$  和  $\hat{\beta}$ . 以下给出了整体在线估计算法的流程:

#### IG 过程的在线算法

1. 在收集至少三次测量的退化数据后, 执行离线估计以获得参数的初始值  $\hat{\nu}^{(3)}$ ,  $\hat{\lambda}^{(3)}$  和  $\hat{\beta}^{(3)}$ . 同时计算  $\tilde{V}_3 = \left[ \sum_{j=1}^3 I_j(\beta^{(3)} | \hat{\nu}^{(3)}, \hat{\lambda}^{(3)}) \right]^{-1}$ .
2. 在第  $m$  次迭代 ( $m \geq 3$ ) 后, 当收到新的观测值  $\mathbf{y}_{m+1}$  时, 使用式 (4.59) 和 (4.60) 更新  $\nu$  和  $\lambda$  的估计值, 其中  $\beta$  替换为  $\hat{\beta}^{(m)}$ , 记更新后的估计值为  $\hat{\nu}^{(m+1)}$  和  $\hat{\lambda}^{(m+1)}$ .
3. 使用式 (4.64) 更新  $\tilde{V}_{m+1}$ . 然后将  $\hat{\nu}^{(m+1)}$  和  $\hat{\lambda}^{(m+1)}$  代入式 (4.66), 计算  $\hat{\beta}^{(m+1)}$ .

4. 重复步骤 2 和 3, 直到没有新的观测值为止.

定理 4.5 给出了估计量  $\hat{\theta}^{(m)}$  的一致性和渐近正态性结论, 为提出的迭代过程提供了理论支撑. 相比经典估计理论 (如 M 估计和一步估计), 该算法由于增加近似和迭代特性使得证明渐近性质时有一定区别, 基本思路是通过将 MLE  $\hat{\theta}^{(3)}$  作为算法初始值, 逐步建立  $(\hat{\nu}^{(m)}, \hat{\lambda}^{(m)})$  和  $\hat{\beta}^{(m)}$  的收敛性; 进一步, 基于 Donsker 类性质 (见本节附录 4.4.6), 证明了估计量的渐近正态性.

**定理 4.5.** 对于每个  $m \geq 3$ , 当  $n \rightarrow \infty$  时,  $(\hat{\nu}^{(m)}, \hat{\lambda}^{(m)}, \hat{\beta}^{(m)})$  依概率收敛于  $(\nu_0, \lambda_0, \beta_0)$ . 此外, 估计量序列  $\sqrt{n}\{(\hat{\nu}^{(m)}, \hat{\lambda}^{(m)}, \hat{\beta}^{(m)}) - (\nu_0, \lambda_0, \beta_0)\}$  依分布收敛于均值为零且协方差矩阵为  $\Sigma_m$  的三维正态随机向量, 其中矩阵  $\Sigma_m$  可用递归公式更新, 其具体形式见本节附录 4.4.6.

在确定不同检测时间点的估计量后, 可对 RUL 进行实时更新. 定义系统在时间  $t_m$  的 RUL 为  $\mathcal{X}_m = \inf\{x : Y(x + t_m) \geq \omega \mid y_m < \omega\}$ , 其中  $\omega$  为失效阈值. 若  $t_m$  时刻的退化值  $y_m$  已达到  $\omega$ , 则 RUL 为零. 由于 IG 过程是单调递增的, 事件  $\{\mathcal{X}_m < x\}$  等价于  $\{Y(x + t_m) \geq \omega\}$ , 因此 RUL 的 CDF 可表示为

$$\begin{aligned} F_{\mathcal{X}_m}(x \mid y_m) &= P\{Y(x + t_m) \geq \omega\} = P\{Y(x + t_m) - y_m \geq \omega - y_m\} \\ &= \Phi\left(\frac{\sqrt{\lambda}[\Delta\Lambda_x - \nu(\omega - y_m)]}{\sqrt{\omega - y_m}}\right) \\ &\quad - \exp(2\nu\lambda\Delta\Lambda_x)\Phi\left(-\frac{\sqrt{\lambda}[\Delta\Lambda_x + \nu(\omega - y_m)]}{\sqrt{\omega - y_m}}\right), \end{aligned} \quad (4.67)$$

其中  $\Delta\Lambda_x = \Lambda_\beta(x + t_m) - \Lambda_\beta(t_m)$ ,  $\Phi(\cdot)$  是标准正态分布的 CDF. 使用估计量  $(\hat{\nu}^{(m)}, \hat{\lambda}^{(m)}, \hat{\beta}^{(m)})$  可对  $F_{\mathcal{X}_m}(\cdot)$  进行实时更新. 此外, 还可推导系统的其他可靠性特征. 例如, RUL 的均值可表示为

$$\mathbb{E}(\mathcal{X}_m \mid y_m) = \int_0^\infty [1 - F_{\mathcal{X}_m}(x \mid y_m)] dx.$$

然而, 由于  $\mathbb{E}(\mathcal{X}_m \mid y_m)$  涉及到非线性时间尺度变换  $\Lambda_\beta(\cdot)$ , 其解析表达式难以获得, 需要借助数值方法计算 (Huynh, 2021).

### 4.4.3 考虑随机效应的在线估计

本小节研究带随机效应IG 过程的在线 RUL 估计. 由于原材料差异、生产过程波动和运行环境变化等因素, 不同系统之间可能存在异质性. 随机效应模型常被用来刻画这类异质特征 (Fang 等, 2022; Wang 等, 2021; Wu 等, 2020; Zhai, Chen, 等, 2018). 在 IG 过程中, 可将漂移参数  $\nu$  设为正态分布随机变量  $\nu \sim N(\mu, \sigma^2)$ , 并假定  $\mu \gg \sigma$ , 以忽略  $\nu$  取负值的概率 (Pan 等, 2016; Ye, Chen N, 2014).

#### 4.4.3.1 估计量

在带随机效应IG 过程的模型中, 参数为  $\theta = (\lambda, \beta, \mu, \sigma)$ . 基于退化数据  $Y_{0:m}$ , 似然函数的形式较为复杂, 无法得到参数  $\theta$  的估计量的解析表达式. 本小节将介绍一个两阶段的离线估计方法, 并为模型参数提供结合偏差校正的解析形式估计量.

由于  $\beta$  可通过上一节中的一步估计方法获得, 这里先假设其已知. 两阶段估计方法的基本策略是先估计缺失参数  $\nu_1, \dots, \nu_n$ , 然后利用基于这些估计量获取  $\mu$  和  $\sigma$  的估计. 第一步中, 在给定  $\beta$  和观测退化增量  $\Delta y_1, \dots, \Delta y_m$  的条件下,  $\lambda$  和  $\nu_i$  的 MLE 分别为

$$\hat{\lambda}^{(m)} = \frac{nm}{\sum_{i=1}^n \phi_{i,m}}, \quad \hat{\nu}_i^{(m)} = \frac{\Lambda_\beta(t_m)}{y_{i,m}}, \quad i = 1, \dots, n, \quad (4.68)$$

其中  $\phi_{i,m} = \sum_{j=1}^m \frac{\Delta \Lambda_j^2}{\Delta y_{i,j}} - \frac{\Lambda_\beta^2(t_m)}{y_{i,m}}$ . 鉴于参数数量较多而单个系统的样本量有限, 当  $m$  较小时, 这些估计量可能存在较大偏差. 因此, 引入以下方法来减少估计的偏差. 首先, 对  $\hat{\lambda}^{(m)}$  进行校正. 注意到

$$\Delta y_{i,j} \sim \text{IG}(\Delta \Lambda_j / \nu_i, \lambda \Delta \Lambda_j^2), \quad y_{i,m} \sim \text{IG}(\Lambda_\beta(t_m) / \nu_i, \lambda \Lambda_\beta^2(t_m)).$$

同时有

$$\mathbb{E} \left[ \frac{1}{\Delta y_{i,j}} \right] = \frac{\nu_i}{\Delta \Lambda_j} + \frac{1}{\lambda \Delta \Lambda_j^2}, \quad \mathbb{E} [y_{i,m}] = \frac{\nu_i}{\Lambda_\beta(t_m)} + \frac{1}{\lambda \Lambda_\beta^2(t_m)}.$$

由此可得

$$\begin{aligned}\mathbb{E} [\phi_{i,m}] &= \sum_{j=1}^m \mathbb{E} \left[ \frac{\Delta \Lambda_j^2}{\Delta y_{i,j}} \right] - \mathbb{E} \left[ \frac{\Lambda_\beta^2(t_m)}{y_{i,m}} \right] \\ &= \sum_{j=1}^m \left( \nu_i \Delta \Lambda_j + \frac{1}{\lambda} \right) - \left( \nu_i \Lambda_\beta(t_m) + \frac{1}{\lambda} \right) = \frac{m-1}{\lambda}.\end{aligned}$$

$\sum_{i=1}^n \phi_{i,m}$  的期望为

$$\mathbb{E} \left[ \sum_{i=1}^n \phi_{i,m} \right] = \frac{n(m-1)}{\lambda},$$

从而可得  $1/\lambda$  的无偏估计量:  $T_m = \sum_{i=1}^n \phi_{i,m} / [n(m-1)]$ . 对于任何可微函数  $h(\cdot)$ ,  $h(T_m)$  都可以作为  $h(1/\lambda)$  的估计量, 但通常会有偏差. 通过泰勒展开近似可得

$$\begin{aligned}\mathbb{E} [h(T_m)] &\approx \mathbb{E} \left[ h(1/\lambda) + h'(1/\lambda)(T_m - 1/\lambda) + h''(1/\lambda)(T_m - 1/\lambda)^2 / 2 \right] \\ &= h(1/\lambda) + \text{Var}(T_m)h''(1/\lambda)/2.\end{aligned}$$

当  $h(x) = 1/x$  时,  $1/T_m$  的期望可近似为

$$\mathbb{E} [1/T_m] \approx \lambda + \text{Var}(T_m)\lambda^3. \quad (4.69)$$

为了找到  $\lambda$  的渐近无偏估计量, 需要找到  $\text{Var}(T_m)$  的近似. 由于  $\phi_{im}, i = 1, \dots, n$  独立同分布, 因此  $\text{Var}(T_m) = \frac{\text{Var}(\phi_{im})}{n(m-1)^2}$ . 给定  $\beta$ ,  $\text{Var}(\phi_{im})$  可以通过  $\frac{1}{n-1} \sum_{i=1}^n (\phi_{im} - \bar{\phi}_m)^2$  进行估计, 其中  $\bar{\phi}_m = \frac{1}{n} \sum_{i=1}^n \phi_{im}$ . 因此, 方差的渐近估计量为

$$\widehat{\text{Var}}(T_m) = \frac{\frac{1}{n-1} \sum_{i=1}^n (\phi_{im} - \bar{\phi}_m)^2}{n(m-1)^2}. \quad (4.70)$$

结合式 (4.68) 和 (4.69), 可得带偏差校正的  $\lambda$  的解析形式估计量为

$$\tilde{\lambda}^{(m)} = \frac{n(m-1)}{\sum_{i=1}^n \phi_{im}} - \frac{\sum_{i=1}^n (\phi_{im} - \bar{\phi}_m)^2}{n(n-1)(m-1)^2} \left( \frac{nm}{\sum_{i=1}^n \phi_{i,m}} \right)^3. \quad (4.71)$$

对于  $\hat{\nu}_i^{(m)}$  的偏差校正, 有

$$\mathbb{E} \left[ \hat{\nu}_i^{(m)} \right] = \mathbb{E} \left[ \mathbb{E} \left[ \hat{\nu}_i^{(m)} \mid \nu_i \right] \right] = \mathbb{E} \left[ \nu_i + \frac{1}{\lambda \Lambda_\beta(t_m)} \right] = \mu + \frac{1}{\lambda \Lambda_\beta(t_m)}, \quad (4.72)$$

这意味着偏差为  $1/(\lambda \Lambda_\beta(t_m))$ . 因此, 带有偏差校正的  $\nu$  的解析形式估计量为

$$\tilde{\nu}_i^{(m)} = \hat{\nu}_i^{(m)} - \frac{1}{\tilde{\lambda}^{(m)} \Lambda_\beta(t_m)} = \frac{\Lambda_\beta(t_m)}{y_{i,m}} - \frac{1}{\tilde{\lambda}^{(m)} \Lambda_\beta(t_m)}, \quad i = 1, \dots, n. \quad (4.73)$$

在第二阶段, 通过估计量  $\tilde{\nu}^{(m)} = (\tilde{\nu}_1^{(m)}, \dots, \tilde{\nu}_n^{(m)})$  来估计随机效应参数  $\mu$  和  $\sigma$ , 这些参数可被视为来自  $N(\mu, \sigma^2)$  的伪样本.  $\mu$  的估计量可以直接通过  $\tilde{\nu}^{(m)}$  的均值计算, 具体为

$$\tilde{\mu}^{(m)} = \frac{1}{n} \sum_{i=1}^n \tilde{\nu}_i^{(m)}. \quad (4.74)$$

同时,  $\tilde{\nu}_i^{(m)}$  的方差近似为

$$\begin{aligned} \mathbb{Var} \left[ \tilde{\nu}_i^{(m)} \right] &\approx \mathbb{Var} \left[ \hat{\nu}_i^{(m)} \right] = \mathbb{Var} \left[ \mathbb{E} \left[ \hat{\nu}_i^{(m)} \mid \nu_i \right] \right] + \mathbb{E} \left[ \mathbb{Var} \left[ \hat{\nu}_i^{(m)} \mid \nu_i \right] \right] \\ &= \mathbb{Var} \left[ \nu_i + \frac{1}{\lambda \Lambda_\beta(t_m)} \right] + \mathbb{E} \left[ \frac{2}{\lambda^2 \Lambda_\beta^4(t_m)} + \frac{\nu_i}{\lambda \Lambda_\beta^3(t_m)} \right] \\ &= \sigma^2 + \frac{2}{\lambda^2 [\Lambda_\beta(t_m)]^4} + \frac{\mu}{\lambda [\Lambda_\beta(t_m)]^3}. \end{aligned} \quad (4.75)$$

基于此,  $\sigma$  可用以下估计量来估计

$$\begin{aligned} \tilde{\sigma}^{(m)} &= \left\{ \frac{1}{n-1} \sum_{i=1}^n \left( \tilde{\nu}_i^{(m)} - \tilde{\mu}^{(m)} \right)^2 \right. \\ &\quad \left. - \frac{2}{[\tilde{\lambda}^{(m)}]^2 [\Lambda_\beta(t_m)]^4} - \frac{\tilde{\mu}^{(m)}}{[\tilde{\lambda}^{(m)}]^3 [\Lambda_\beta(t_m)]^3} \right\}^{-1/2}. \end{aligned} \quad (4.76)$$

#### 4.4.3.2 递归更新

上节推导的解析形式估计量的一大优势在于: 它们能够利用历史退化数据的汇总 (summary) 统计量进行高效更新. 为说明这一点, 本小节首先给出了

$\phi_{i,m}$  的递归表达式. 对于  $i = 1, \dots, n$ , 有

$$\begin{aligned}\phi_{i,m+1} &= \sum_{j=1}^{m+1} \frac{\Delta\Lambda_j^2}{\Delta y_{i,j}} - \frac{\Lambda_\beta^2(t_{m+1})}{y_{i,m+1}} \\ &= \phi_{i,m} + \frac{y_{i,m+1}\Delta y_{i,m+1}}{y_{i,m}} \left[ \frac{\Lambda_\beta(t_{m+1})}{y_{i,m+1}} - \frac{\Delta\Lambda_{m+1}}{\Delta y_{i,m+1}} \right]^2,\end{aligned}\quad (4.77)$$

其中  $y_{i,m+1}$  和  $\Delta y_{i,m+1}$  分别为第  $i$  个系统的新退化测量值和退化增量. 基于递归公式,  $\tilde{\lambda}^{(m+1)}$  可以直接通过式 (4.71) 更新, 而只需存储  $n$  个系统的汇总统计量  $\phi_{i,m}$  和最后一次退化测量值  $y_{i,m}$ , 这些统计量可由历史退化数据  $Y_{0:m}$  计算.

更新后的  $\tilde{\lambda}^{(m+1)}$  可用于递归更新  $\tilde{\nu}^{(m+1)}$  (参见式 (4.73)). 随后, 利用式 (4.74) 和 (4.76) 可分别得到  $\tilde{\mu}^{(m+1)}$  和  $\tilde{\sigma}^{(m+1)}$ . 上述的递归过程需要输入  $\beta$ . 为了进一步优化, 采用一步估计法得到  $\beta$  的递推公式, 从而实现  $(\tilde{\lambda}, \tilde{\mu}, \tilde{\sigma}, \tilde{\beta})$  的高效迭代更新. 类似于在不考虑随机效应时的递推公式 (4.66), 在存在随机效应时可用以下递推公式.

$$\tilde{\beta}^{(m+1)} = \tilde{\beta}^{(m)} + \frac{1}{n} \widetilde{RV}_{m+1} R S_{m+1}(\Delta y_{m+1}, \tilde{\beta}^{(m)}), \quad (4.78)$$

其中  $\widetilde{RV}_{m+1} = \left[ \sum_{j=1}^{m+1} R I_j(\tilde{\beta}^{(j-1)} \mid \tilde{\nu}^{(j-1)}, \tilde{\lambda}^{(j-1)}) \right]^{-1}$  是 Fisher 信息逆矩阵的近似, 且

$$R I_j(\tilde{\beta}^{(j)} \mid \tilde{\nu}^{(j)}, \tilde{\lambda}^{(j)}) = \sum_{i=1}^n \left[ 2 \left( \frac{\Delta \tilde{\Lambda}_j'}{\Delta \tilde{\Lambda}_j} \right)^2 + \frac{\tilde{\nu}_i^{(j)} \tilde{\lambda}^{(j)} (\Delta \tilde{\Lambda}_j')^2}{\Delta \tilde{\Lambda}_j} \right],$$

其中  $\Delta \tilde{\Lambda}_j$  和  $\Delta \tilde{\Lambda}_j'$  分别是  $\Delta \Lambda_j$  和  $\Delta \Lambda_j'$ , 只需将  $\beta$  替换为  $\tilde{\beta}^{(j)}$ . 另一方面,  $R S_j(\Delta y_j, \beta)$  表示第  $j$  个退化增量的对数似然函数关于  $\beta$  的导数, 其表达式为

$$R S_j(\Delta y_j, \tilde{\beta}^{(j-1)}) = \frac{n \Delta \tilde{\Lambda}_j'}{\Delta \tilde{\Lambda}_j} + \sum_{i=1}^n \left[ \tilde{\nu}_i^{(j)} \tilde{\lambda}^{(j)} \Delta \tilde{\Lambda}_j' - \frac{\tilde{\lambda}^{(j)} \Delta \tilde{\Lambda}_j' \Delta \tilde{\Lambda}_j}{y_{ij}} \right].$$

容易验证, Fisher 信息逆矩阵的近似  $\widetilde{RV}_m$  可通过以下公式递归更新.

$$\widetilde{RV}_{m+1}^{-1} = \widetilde{RV}_m^{-1} + RI_{m+1}(\tilde{\beta}^{(m)} \mid \tilde{\nu}^{(m)}, \tilde{\lambda}^{(m)}) \quad (4.79)$$

并且  $RS_{m+1}(\Delta \mathbf{y}_{m+1}, \tilde{\beta}^{(m)})$  仅依赖于新退化增量  $\Delta \mathbf{y}_{m+1}$ . 因此, 式 (4.78) 中的  $\tilde{\beta}$  具有递归结构, 可实现高效更新.

#### 4.4.3.3 在线算法和 RUL 估计

基于上述递推公式的推导, 可构建一个全面的在线算法. 类似于不考虑随机效应的情形, 建议采用至少  $m = 3$  次退化测量来获取参数的初始估计. 算法的具体步骤如下:

##### 含随机效应的 IG 过程在线算法

1. 使用前三次退化测量进行离线估计. 根据式 (4.71)、(4.73) 和 (4.85), 计算初始参数  $(\tilde{\lambda}^{(3)}, \tilde{\nu}_1^{(3)}, \dots, \tilde{\nu}_n^{(3)}, \tilde{\beta}^{(3)})$ , 并基于伪样本  $\tilde{\nu}^{(3)}$  通过式 (4.74) 和 (4.76) 获得  $\tilde{\mu}^{(3)}$  和  $\tilde{\sigma}^{(3)}$ . 同时计算 Fisher 信息逆矩阵的初始值:  $\widetilde{RV}_3 = [\sum_{j=1}^3 RI_j(\tilde{\beta}^{(3)} \mid \tilde{\nu}^{(3)}, \tilde{\lambda}^{(3)})]^{-1}$ .
2. 对于第  $m$  次迭代后 ( $m \geq 3$ ), 在获得新的观测值  $\mathbf{y}_{m+1}$  后, 按照以下步骤更新: a) 利用式 (4.77) 更新  $\phi_{i,m+1}$ . b) 根据式 (4.71)、(4.73)、(4.74) 和 (4.76), 依次更新  $\tilde{\lambda}^{(m+1)}$ 、 $\tilde{\nu}^{(m+1)}$ 、 $\tilde{\mu}^{(m+1)}$  和  $\tilde{\sigma}^{(m+1)}$ , 其中  $\beta$  替换为  $\tilde{\beta}^{(m)}$ .
3. 使用式 (4.79) 更新  $\widetilde{RV}_{m+1}$ , 并将  $\tilde{\lambda}^{(m+1)}$  和  $\tilde{\nu}^{(m+1)}$  代入式 (4.78), 更新  $\beta^{(m+1)}$ .
4. 重复步骤 2 和 3, 直到不再有新的观测值被收集.

在推导  $t_m$  时刻的 RUL 分布  $\mathcal{X}_m = \inf\{x : Y(x + t_m) \geq \omega \mid y_m < \omega\}$  时, 需要在式 (4.67) 中对未知的  $\nu$  进行积分, 这增加了计算的复杂性. 具体来

说, 给定  $\nu$  和  $y_m$  的情况下,  $\mathcal{X}_m$  的 CDF 为

$$F_{\mathcal{X}_m}(x | \nu, y_m) = \Phi \left( \frac{\sqrt{\lambda} [\Delta \Lambda_x - \nu(\omega - y_m)]}{\sqrt{\omega - y_m}} \right) - \exp(2\nu\lambda\Delta\Lambda_x) \Phi \left( -\frac{\sqrt{\lambda} [\Delta \Lambda_x + \nu(\omega - y_m)]}{\sqrt{\omega - y_m}} \right),$$

其中  $\Delta \Lambda_x = \Lambda_\beta(x + t_m) - \Lambda_\beta(t_m)$ . 通过对  $F_{\mathcal{X}_m}(x | \nu, y_m)$  关于  $\nu$  的条件分布取期望, 可以得到  $\mathcal{X}_m$  的边际 CDF:

$$F_{\mathcal{X}_m}(x | y_m) = \int_{\nu} F_{\mathcal{X}_m}(x | \nu, y_m) \pi(\nu | y_m) d\nu, \quad (4.80)$$

其中  $\pi(\nu | y_m)$  表示  $\nu$  的后验 PDF. 利用本节附录 4.4.6 中  $n = 1$  的似然函数和贝叶斯公式,  $\pi(\nu | y_m)$  为

$$\begin{aligned} \pi(\nu | y_m) &\propto \exp \left\{ -\frac{\lambda y_m (\nu - \Lambda_\beta(t_m)/y_m)^2}{2} \right\} \exp \left\{ -\frac{(\nu - \mu)^2}{2\sigma^2} \right\} \\ &\propto \exp \left\{ -\frac{\left( \nu - \frac{\lambda \Lambda_\beta(t_m) + \mu \sigma^{-2}}{\lambda y_m + \sigma^{-2}} \right)^2}{2(\lambda y_m + \sigma^{-2})^{-1}} \right\}, \end{aligned}$$

化简后可得  $\pi(\nu | y_m)$  为正态分布  $\mathcal{N}(\mu_m, \tau_m)$ , 其中

$$\mu_m = \frac{\lambda \Lambda_\beta(t_m) + \mu \sigma^{-2}}{\lambda y_m + \sigma^{-2}}, \quad \tau_m = (\lambda y_m + \sigma^{-2})^{-1}.$$

因此, 边际 CDF 有显式表达式, 如定理 4.6 所示, 其证明见本节附录 4.4.6.

**定理 4.6.** 假设当前  $t_m$  时刻的观测退化测量值为  $0 < y_1 < \dots < y_m < \omega$ , 且这些值服从带随机效应  $\nu \sim N(\mu, \sigma^2)$  的 IG 过程. 则  $\mathcal{X}_m$  的 CDF 为

$$\begin{aligned} F_{\mathcal{X}_m}(x | y_m) &= \Phi \left( \frac{-K_1 \mu_m + K_2}{\sqrt{1 + K_1^2 \tau_m}} \right) - \exp \left( K_3 \mu_m + \frac{K_3^2 \tau_m}{2} \right) \\ &\quad \times \Phi \left( \frac{-K_1 \mu_m - K_2 - K_1 K_3 \tau_m}{\sqrt{1 + K_1^2 \tau_m}} \right), \end{aligned} \quad (4.81)$$

其中

$$\mu_m = \frac{\lambda \Lambda_\beta(t_m) + \mu \sigma^{-2}}{\lambda y_m + \sigma^{-2}}, \quad \tau_m = (\lambda y_m + \sigma^{-2})^{-1},$$

$$K_1 = \sqrt{\lambda(\omega - y_m)}, \quad K_2 = \frac{\sqrt{\lambda} \Delta \Lambda_x}{\sqrt{(\omega - y_m)}}, \quad K_3 = 2\lambda \Delta \Lambda_x.$$

此外,  $\mathcal{X}_m$  的 PDF 为

$$\begin{aligned} f_{\mathcal{X}_m}(x | y_m) &= \varphi \left( \frac{-K_1 \mu_m + K_2}{\sqrt{1 + K_1^2 \tau_m}} \right) \frac{K_2'}{\sqrt{1 + K_1^2 \tau_m}} \\ &\quad - (K_3' \mu_m + K_3' K_3 \tau_m) \exp \left( K_3 \mu_m + \frac{K_3^2 \tau_m}{2} \right) \\ &\quad \times \Phi \left( \frac{-K_1 \mu_m - K_2 - K_1 K_3 \tau_m}{\sqrt{1 + K_1^2 \tau_m}} \right) \\ &\quad + \frac{K_2' + K_1 K_3' \tau_m}{\sqrt{1 + K_1^2 \tau_m}} \exp \left( K_3 \mu_m + \frac{K_3^2 \tau_m}{2} \right) \\ &\quad \times \varphi \left( \frac{-K_1 \mu_m - K_2 - K_1 K_3 \tau_m}{\sqrt{1 + K_1^2 \tau_m}} \right), \end{aligned} \quad (4.82)$$

其中  $\varphi(\cdot)$  为标准正态分布的 PDF,  $K_2' = \frac{\sqrt{\lambda} \Delta \Lambda_x'}{\sqrt{(\omega - y_m)}}$ ,  $K_3' = 2\lambda \Delta \Lambda_x'$ ,  $\Delta \Lambda_x'$  为  $\Delta \Lambda_x$  关于  $x$  的一阶导数.

#### 4.4.4 模拟实验

本节通过模拟研究评估所提出在线算法的性能. 参数设置:

$$\nu = 3, \lambda = 10, \Lambda_\beta(t) = t^2 (\beta = 2).$$

试验中随机选取  $n = 15$  个系统, 采用等间隔测量, 测量时间在区间  $[0.1, 10]$  内, 测量间隔为 0.1. 因此, 每个系统的测量次数  $m = 100$ .

采用第 4.4.2 节提出的在线算法分析数据, 并记录每个时间点的估计值. 该过程重复 10000 次, 总计算耗时 3.27 分钟 (运行环境: 2.7 GHz Intel Core i5, 8 GB RAM). 通过 10000 次重复实验, 计算每个测量时间点  $t_j$  处不同参数的 RB 和 RMSE. 图 4.18 展示参数  $\nu$ 、 $\lambda$  和  $\beta$  的估计结果. 结果表明, 估计值在几次迭代

后迅速稳定, 并最终收敛于真实值.

接下来, 进一步考虑带随机效应的 IG 过程, 假设  $\nu \sim N(3, 0.8^2)$ , 即  $\mu = 3$  且  $\sigma = 0.8$ . 为说明偏差校正估计量的效果, 本小节对比了带偏差校正和不带偏差校正的在线算法. 本次模拟重复进行 10000 次, 总计算时间为 5.45 分钟. 图 4.19 展示了不同参数在不同测量时间点的 RMSE, 其中黑色曲线代表带偏差校正的在线算法, 红色曲线代表不带偏差校正的算法. 结果表明, 采用偏差校正后, 估计精度显著提升. 此外, 即使在收集到少量退化测量后, 基于偏差校正的在线算法的估计值也能快速收敛至真实值.

本节还评估了基于含随机效应的 IG 过程的 RUL 预测性能. 不失一般性, 选取第一个系统的 RUL 作为研究对象, 并将失效阈值设定为  $\omega = 60$ . 在每次模拟中, 持续监测系统的退化水平直至达到失效阈值, 从而在每个检测时间点获得真实的 RUL. 基于 10000 次重复实验, 图 4.20 展示了第一个系统 RUL 预测的 RB 和 RMSE. 结果表明, 所提出的 RUL 预测方法在仅收集到少量测量值后就能达到较高的预测精度. 后续章节将结合真实数据集进一步分析 RUL 预测性能.

#### 4.4.5 实例分析

本节通过两个真实的退化数据集, 验证了所提出方法的有效性. 这些数据表明, IG 过程是一个合适的模型, 且在线算法可以高效地估计模型参数并预测 RUL.

##### 4.4.5.1 集成电路器件退化数据

本研究选取了在  $195^{\circ}\text{C}$  下收集的集成电路器件退化数据作为分析对象 (详见图 1.6). 在该温度下, 共测试了  $n = 12$  个器件, 测量次数  $m = 16$ , 测量时间点均匀分布在 0 至 2000 小时的区间内. 选取非线性时间函数为  $\Lambda_{\beta}(t) = t^{\beta}$ . 图 4.21 展示了退化数据以及拟合的带随机效应和无随机效应的 IG 过程模型. 从图中可以看出, 这两种模型均能够较好地刻画退化路径特征.

随后, 运用所提出的在线算法计算参数估计值. 图 4.22 和图 4.23 展示了在两种 IG 模型下参数的递归更新估计值. 可看出, 参数估计普遍在 0.75 千小时左右后达到稳定. 为了对比两种 IG 模型, 计算了每个时间点下模型的 AIC 值, 其中 AIC 值较低的模型更优. 此外, 将无随机效应的 IG 模型作为原假设, 含随

机效应的 IG 模型作为备择假设, 运用似然比检验, 同时计算每个时间下的检验  $p$  值. 图 4.24 展示每个时间点的 AIC 值和  $p$  值.

结果显示, 在所有测量时间点, 含随机效应的 IG 模型具有更低的 AIC 值, 表明其拟合效果更佳. 同时,  $p$  值均小于设定的显著性水平 0.05, 说明拒绝原假设. 这些发现表明, 带随机效应的 IG 模型更适用于分析该退化数据集.

#### 4.4.5.2 锂离子电池容量退化数据

以图 1.8 中的 17 个电池退化数据为例, 首先使用所提出的两种 IG 模型对数据进行拟合. 设定  $\Lambda_\beta(t) = t^\beta$ , 使用所提出的算法获得参数的在线估计, 结果如图 4.25 和图 4.26 所示. 可发现参数估计在 100 次循环后趋于稳定. 图 4.27 显示了 AIC 值和似然比检验的  $p$  值. 结果表明, 在显著性水平 0.05 下, 带随机效应的 IG 过程更适合描述该数据. 随后, 通过 Shapiro-Wilk 检验评估了随机效应服从正态分布的假设. 如图 4.28 所示, 除前几个时间点外, 几乎所有正态性检验的  $p$  值均大于 0.05, 这表明随机效应  $\nu$  的正态性假设是合理的.

为了进一步展示随机效应模型在 RUL 预测中的优势, 本小节以 #6 锂离子电池为例, 设定失效阈值为  $\omega = 17$ . 基于此阈值, #6 电池的寿命为 731 个充放电循环, 因此在每个循环中都可以获得其真实 RUL. 采用第 4.4.3 节提出的算法, 在 251、326、401、476 和 551 个循环时对 #6 电池的 RUL 进行预测. 图 4.29(a) 展示了基于两种 IG 模型的 RUL 预测均值和 PDF. 结果表明, 带随机效应的 IG 模型提供了更准确的预测, 且随着退化数据量的增加, 预测精度不断提高. 图 4.29(b) 展示了所有 125 个测量时间点的 RUL 预测均值及近似 95% 置信区间. 这些置信区间利用  $F_{\mathcal{X}_m}(x | y_m)$  的 2.5% 和 97.5% 分位数和参数的估计值来构造, 见公式 (4.81). 结果显示, 基于随机效应模型的置信区间中, 125 个真实 RUL 点中有 7 个超出置信区间, 对应的频率覆盖概率为  $(125 - 7) / 125 = 94.4\%$ , 接近名义水平 95%. 相比之下, 基于无随机效应的 IG 模型所构造的置信区间明显更宽, 仅有 2 个真实 RUL 点超出区间, 对应覆盖概率为 98.4%. 尽管覆盖概率较高, 但较宽的置信区间反映了更大的预测不确定性. 综合来看, 随机效应模型能更有效地利用退化数据, 为 RUL 预测提供更高的精度和更具解释性的信息.

#### 4.4.6 附录

当  $\beta$  已知时,  $\nu$  和  $\lambda$  的 MLE

在式 (4.57) 中, 似然函数可写为

$$L(\boldsymbol{\theta} \mid \mathbf{Y}_{0:m}) = C_0 \lambda^{nm/2} \prod_{j=1}^m \Delta \Lambda_j^n \exp \left\{ -\frac{\lambda}{2} \left[ \sum_{i=1}^n \sum_{j=1}^m \frac{\Delta \Lambda_j^2}{\Delta y_{ij}} - \frac{n^2 \Lambda_\beta(t_m)^2}{\sum_{i=1}^n y_{im}} \right] \right\} \\ \times \exp \left\{ -\frac{\lambda \sum_{i=1}^n y_{im}}{2} \left( \nu - \frac{n \Lambda_\beta(t_m)}{\sum_{i=1}^n y_{im}} \right)^2 \right\}. \quad (4.83)$$

在 (4.83) 中, 第一行是  $\lambda$  的函数 (记为  $G(\lambda)$ ), 而第二行是一个指数函数, 当  $\nu = n \Lambda_\beta(t_m) / \sum_{i=1}^n y_{im}$  时达到最大值. 另一方面, 当

$$\lambda = \frac{nm}{\sum_{i=1}^n \sum_{j=1}^m \frac{\Delta \Lambda_j^2}{\Delta y_{ij}} - \frac{n^2 \Lambda_\beta^2(t_m)}{\sum_{i=1}^n y_{im}}}$$

时,  $G(\lambda)$  达到最大值. 此时,  $\nu$  和  $\lambda$  的 MLE 如 (4.58) 所示.

#### $\beta$ 的 Fisher 信息的推导

给定第  $j$  个增量  $\Delta y_j = (\Delta y_{i,j}, \dots, \Delta y_{n,j})$  时,  $\beta$  的对数似然函数为  $l_j(\beta)$ . 对  $l_j(\beta)$  关于  $\beta$  求一阶导数, 可得

$$\frac{\partial l_j(\beta \mid \Delta \mathbf{y}_j)}{\partial \beta} = \frac{n \Delta \Lambda_j'}{\Delta \Lambda_j} + n \nu \lambda \Delta \Lambda_j' - \sum_{i=1}^n \frac{\lambda \Delta \Lambda_j' \Delta \Lambda_j}{\Delta y_{i,j}},$$

对  $l_j(\beta)$  关于  $\beta$  求二阶导数, 可得

$$\frac{\partial^2 l_j(\beta \mid \Delta \mathbf{y}_j)}{\partial \beta^2} = n \left[ \frac{\Delta \Lambda_j''}{\Delta \Lambda_j} - \left( \frac{\Delta \Lambda_j'}{\Delta \Lambda_j} \right)^2 + \nu \lambda \Delta \Lambda_j'' \right] \\ - \sum_{i=1}^n \left[ \frac{\lambda \Delta \Lambda_j'' \Delta \Lambda_j}{\Delta y_{i,j}} + \frac{\lambda (\Delta \Lambda_j')^2}{\Delta y_{i,j}} \right],$$

其中  $\Delta\Lambda_j'$  和  $\Delta\Lambda_j''$  分别是  $\Delta\Lambda_j$  关于  $\beta$  的一阶和二阶导数. 由于  $\Delta y_{i,j} \sim \text{IG}(\Delta\Lambda_j/\nu, \lambda\Delta\Lambda_j^2)$ , 可以计算出  $1/\Delta y_{i,j}$  的期望为

$$E\left[\frac{1}{\Delta y_{i,j}}\right] = \frac{\nu}{\Delta\Lambda_j} + \frac{1}{\lambda\Delta\Lambda_j^2}.$$

因此, 给定观测值  $\Delta y_{i,j}, \beta$  的 Fisher 信息为

$$I_j(\beta | \nu, \lambda) = -E\left[\frac{\partial^2 l_j(\beta | \Delta y_j)}{\partial \beta^2}\right] = n \left[ \frac{\nu\lambda(\Delta\Lambda_j')^2}{\Delta\Lambda_j} + 2 \left(\frac{\Delta\Lambda_j'}{\Delta\Lambda_j}\right)^2 \right].$$

### 参数的可识别性和离线估计

需要证明当  $n$  个系统仅被测量一次时, 参数是不可识别的. 给定  $y_1 = (y_{1,1}, \dots, y_{n,1})$ ,  $\theta$  的对数似然函数为

$$l(\theta | y_1) = C + \frac{n}{2} \log \lambda + n \log \Lambda_\beta(t_1) - \lambda \sum_{i=1}^n \frac{1}{2y_{i,1}} (\nu y_{i,1} - \Lambda_\beta(t_1))^2.$$

根据式 (4.58), 在给定  $\beta$  的情况下,  $\nu$  和  $\lambda$  的 MLE 分别为

$$\frac{n\Lambda_\beta(t_1)}{\sum_{i=1}^n y_{i,1}} \quad \text{和} \quad \frac{n}{\sum_{i=1}^n \frac{1}{y_{i,1}} (\nu y_{i,1} - \Lambda_\beta(t_1))^2}.$$

将这两个估计量代入  $l(\theta | y_1)$  中, 得到  $\beta$  的轮廓似然函数为

$$pl_1(\beta) = C_1 - n \log \Lambda_\beta(t_1) + n \log \Lambda_\beta(t_1) - n/2 = C_1 - n/2,$$

其中  $C_1$  是一个与  $\beta$  无关的常数. 因此,  $\beta$  是不可识别的, 因为  $pl_1(\beta)$  对于任意的  $\beta$  值均为常数. 这个结果是比较直观的, 因为一条曲线的形状仅通过两个点无法确定. 因此, 为了获得参数的 MLE 并运用在线算法, 至少需要两次测量. 在实际应用中, 通过前三个测量时间点  $t_1, t_2$  和  $t_3$  收集的观测数据,  $\nu$  和  $\lambda$  的 MLE 分别为

$$\hat{\nu}^{(3)} = \frac{n\Lambda_{\hat{\beta}^{(3)}}(t_3)}{\sum_{i=1}^n y_{i,3}}, \quad \hat{\lambda}^{(3)} = \frac{3n}{\sum_{i=1}^n \sum_{j=1}^3 \frac{\Delta\tilde{\Lambda}_j^2}{\Delta y_{i,j}} - \frac{n^2 \Lambda_{\hat{\beta}^{(3)}}^2(t_3)}{\sum_{i=1}^n y_{i,3}}},$$

其中  $\tilde{\Delta\Lambda}_j$  是  $\Delta\Lambda_j$  的估计,  $\beta$  被替换为  $\hat{\beta}^{(3)}$ , 而  $\hat{\beta}^{(3)}$  可通过最大化以下轮廓似然函数得到.

$$pl_3(\beta) = C_3 - \frac{3n}{2} \log \left[ \sum_{i=1}^n \sum_{j=1}^3 \frac{\Delta\Lambda_j^2}{\Delta y_{i,j}} - \frac{n^2 \Lambda_\beta^2(t_3)}{\sum_{i=1}^n y_{i,3}} \right] + n \sum_{j=1}^3 \log \Delta\Lambda_j,$$

其中  $C_3$  是与  $\beta$  无关的常数.

对于含随机效应的 IG 模型,  $\Theta = (\lambda, \nu_1, \dots, \nu_n, \beta)$  的似然函数为

$$\begin{aligned} L(\Theta \mid \mathbf{Y}_{0:m}) &= C_0 \lambda^{nm/2} \prod_{j=1}^m \Delta\Lambda_j^n \exp \left\{ -\frac{\lambda}{2} \sum_{i=1}^n \left[ \sum_{j=1}^m \frac{\Delta\Lambda_j^2}{\Delta y_{ij}} - \frac{n^2 \Lambda_\beta(t_m)^2}{\sum_{i=1}^n y_{im}} \right] \right\} \\ &\quad \times \exp \left\{ -\frac{\lambda}{2} \sum_{i=1}^n y_{im} \left( \nu_i - \frac{\Lambda_\beta(t_m)}{y_{im}} \right)^2 \right\}. \end{aligned} \quad (4.84)$$

此时可以得到(4.68) 中的  $\lambda$  和  $\nu_i$  的 MLE. 再将  $\lambda$  和  $\nu_i$  替换为 MLE, 可得  $\beta$  的轮廓似然函数, 即,

$$pl_m^*(\beta) = C_4 - \frac{mn}{2} \log \sum_{i=1}^n \left[ \sum_{j=1}^m \frac{\Delta\Lambda_j^2}{\Delta y_{i,j}} - \frac{\Lambda_\beta^2(t_m)}{y_{i,m}} \right] + n \sum_{j=1}^m \log \Delta\Lambda_j, \quad (4.85)$$

其中  $C_4$  是与  $\beta$  无关的常数.  $\beta$  (即  $m = 3$ ) 的初始估计量可通过最大化轮廓似然函数  $pl_m^*(\beta)$  进行数值求解.

## 理论 4.5 证明

为了方便表示, 做以下定义.

$$\ell^{(j)}(\boldsymbol{\theta}; \Delta y_j) \equiv \frac{1}{2} \log \left( \frac{\lambda}{2\pi} \right) + \log(\Delta\Lambda_j) - \frac{3}{2} \log(\Delta y_j) - \frac{\lambda}{2\Delta y_j} (\nu \Delta y_j - \Delta\Lambda_j)^2,$$

并且  $\ell^m(\boldsymbol{\theta}; \Delta Y) \equiv \sum_{j=1}^m \ell^{(j)}(\boldsymbol{\theta}; \Delta y_j)$  表示单系统的对数似然函数.

令  $\mathcal{J}^{(j)} = P\ddot{\ell}^{(j)}(\nu_0, \lambda_0, \beta_0)$  表示基于第  $j$  次检测时间点的 Fisher 信息矩阵, 它可以表示为矩阵块的形式

$$\mathcal{J}^{(j)} \equiv \begin{pmatrix} \mathcal{J}_{(\nu, \lambda)}^{(j)} & \mathcal{J}_{(\nu, \lambda), \beta}^{(j)} \\ \mathcal{J}_{\beta, (\nu, \lambda)}^{(j)} & \mathcal{J}_\beta^{(j)} \end{pmatrix}, \quad \mathcal{J}_{(\nu, \lambda)}^{(j)} \in \mathbb{R}^{2 \times 2}, \quad \mathcal{J}_{(\nu, \lambda), \beta}^{(j)} = \mathcal{J}_{\beta, (\nu, \lambda)}^{(j)} \in \mathbb{R}^{2 \times 1},$$

其中下标  $(\nu, \lambda)$  对应于  $(\nu, \lambda)$  的维度, 而下标  $\beta$  对应于  $\beta$  的维度. 令  $\mathcal{I}^m = \sum_{j=1}^m \mathcal{I}^{(j)}$ , 并以类似的方式定义  $\mathcal{I}(\nu, \lambda)^m$ 、 $\mathcal{I}(\nu, \lambda)$ 、 $\beta^m$  和  $\mathcal{I}_\beta^m$ .

首先, 通过经典的大样本理论, 可以确立初始 MLE  $\nu^{(3)}$ 、 $\lambda^{(3)}$  和  $\beta^{(3)}$  的一致性. 目标是在  $\nu^{(m)}$ 、 $\lambda^{(m)}$  和  $\beta^{(m)}$  的一致性基础上, 证明对于每个  $m \geq 3$ ,  $\nu^{(m+1)}$ 、 $\lambda^{(m+1)}$  和  $\beta^{(m+1)}$  也具有一致性. 利用大数定律, 可以得到以下结论.

$$\begin{aligned} n^{-1} \sum_{i=1}^n y_{i,m} &\rightarrow_P \frac{\Lambda_{\beta_0}(t_m)}{\nu}, \\ n^{-1} \sum_{i=1}^n \frac{1}{\Delta y_{i,m+1}} &\rightarrow_P \frac{\nu_0}{\Delta \Lambda_{\beta_0}(t_{m+1})} + \frac{1}{\lambda_0 \Delta \Lambda_{\beta_0}(t_{m+1})^2}. \end{aligned}$$

如果  $\beta^{(m)} \rightarrow_P \beta_0$ , 由于  $\Lambda_\beta(t)$  关于  $\beta$  的可微性, 利用连续映射定理可以得到

$$\begin{aligned} \Lambda_{\beta^{(m)}}(t_m) &\rightarrow_P \Lambda_{\beta_0}(t_m), \\ \Delta \Lambda_{\beta^{(m)}}(t_{m+1}) &\rightarrow_P \Delta \Lambda_{\beta_0}(t_{m+1}). \end{aligned}$$

再次使用连续映射定理, 得到

$$\nu^{(m+1)} = \Lambda_{\beta^{(m)}}(t_m) / (n^{-1} \sum_{i=1}^n y_{i,m}) \rightarrow_P \nu_0$$

$$\begin{aligned} \lambda^{(m+1)} &= (m+1) \left\{ m / \hat{\lambda}^{(m)} + [\hat{\nu}^{(m)}]^2 n^{-1} \sum_{i=1}^n y_{i,m} \right. \\ &\quad \left. - [\hat{\nu}^{(m+1)}]^2 n^{-1} \sum_{i=1}^n y_{i,m+1} + \Delta \Lambda_{\beta^{(m)}}^2(t_{m+1}) n^{-1} \sum_{i=1}^n / \Delta y_{i,m+1} \right\}^{-1} \\ &\rightarrow_P (m+1) \left\{ m / \lambda_0 - \nu n^{-1} \Delta \Lambda_{\beta_0}(t_{m+1}) \right. \\ &\quad \left. + \Delta \Lambda_{\beta_0}^2(t_{m+1}) [\nu_0 / \Delta \Lambda_{\beta_0}^2(t_{m+1}) + (\lambda_0 \Delta \Lambda_{\beta_0}^2(t_{m+1}))^{-1}] \right\}^{-1} = \lambda_0, \end{aligned}$$

其中, 后一项的弱收敛依赖于  $\nu^{(m+1)}$  的一致性. 要利用一致的  $\nu^{(m+1)}$  和  $\lambda^{(m+1)}$  去证明  $\beta^{(m+1)}$  的一致性, 需要以下引理.

**引理 4.1.** 若  $\Delta \Lambda_j$  关于  $\beta$  在  $\beta_0$  的邻域内是二阶连续可微的, 存在  $\delta > 0$  使得

函数类  $\mathcal{F}_\delta$  是 *Donsker* 类, 其中

$$\mathcal{F}_\delta \equiv \left\{ \dot{\ell}^m(\boldsymbol{\theta}; \Delta Y) - \dot{\ell}^m(\boldsymbol{\theta}_0; \Delta Y) : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq \delta \right\},$$

且  $\dot{\ell}^m(\boldsymbol{\theta}; \Delta Y)$  是  $\ell^m(\boldsymbol{\theta}; \Delta Y)$  关于  $(\nu, \lambda, \beta)$  的导数. 对于每一个满足  $\tilde{\boldsymbol{\theta}}_n \rightarrow_P \boldsymbol{\theta}_0$  的序列  $\tilde{\boldsymbol{\theta}}_{nn} = 1^\infty$  有

$$\mathbb{G}_n \left[ \dot{\ell}^m(\tilde{\boldsymbol{\theta}}_n) - \dot{\ell}^m(\boldsymbol{\theta}_0) \right] \equiv \sqrt{n}(\mathbb{P}_n - P) \left[ \dot{\ell}^m(\tilde{\boldsymbol{\theta}}_n) - \dot{\ell}^m(\boldsymbol{\theta}_0) \right] = o_P(1).$$

证明. 对于任意  $M > 0$ , 存在  $\delta > 0$ , 使得当  $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq \delta$  时,  $\Delta\Lambda_j$ 、 $\Delta\Lambda'_j$  和  $\Delta\Lambda''_j$  都是上有界且远离零的. 因此可以得到

$$\left\| \dot{\ell}^m(\boldsymbol{\theta}; \Delta Y) \right\| \lesssim \left( \sum_{j=1}^m \Delta y_j \right) \bigvee \left( \sum_{j=1}^m \Delta 1/y_j \right),$$

其中右侧是可积函数. 由于二阶导数在  $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq \delta$  上一致有界, 存在一个函数  $h(y)$  使得以下的 Lipschitz 条件成立.

$$\|\dot{\ell}^m(\boldsymbol{\theta}_1; \Delta Y) - \dot{\ell}^m(\boldsymbol{\theta}_2; \Delta Y)\| \leq h(y) \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|, \text{ 任意 } \boldsymbol{\theta}_1, \boldsymbol{\theta}_2.$$

因此, 类似于 (Van der Vaart, 1998) 中的例子 19.7,  $\mathcal{F}_\delta$  的包围数满足

$$N_{[]}(\epsilon \|h\|_{L_2(P_0)}, \mathcal{F}_\delta, L_2(P_0)) \lesssim (\delta/\epsilon)^3, \forall 0 < \epsilon < \delta.$$

因此, 包围积分是有限的, 这意味着  $\mathcal{F}_\delta$  是 Donsker 类.

接下来证明  $\mathbb{G}_n \left[ \dot{\ell}^m(\tilde{\boldsymbol{\theta}}_n) - \dot{\ell}^m(\boldsymbol{\theta}_0) \right] = o_P(1)$ . 类似于  $\dot{\ell}^m(\boldsymbol{\theta}; \Delta Y)$  的构造, 可以证明函数类  $\mathcal{F}_\delta$  被平方可积函数所界定. 因此, 通过对  $\tilde{\boldsymbol{\theta}}_n$  的几乎处处收敛子序列进行讨论, 使用支配收敛定理, 得到

$$P_0 \left[ \dot{\ell}^m(\tilde{\boldsymbol{\theta}}_n) - \dot{\ell}^m(\boldsymbol{\theta}_0) \right]^2 \rightarrow_P 0.$$

因此, 4.1 引理的第二部分可以通过 [Van der Vaart (1998)] 中的定理 19.24 推导得出.  $\square$

此时可利用  $(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)})$  依概率收敛于  $(\nu_0, \lambda_0)$  来证明  $\hat{\beta}^{(m+1)}$  的一

致性. 根据在线算法, 估计量  $\hat{\beta}^{(m+1)}$  表达式如下.

$$\hat{\beta}^{(m+1)} = \hat{\beta}_n^{(m)} + \widetilde{V}_{m+1}^{-1} n \mathbb{P}_n \dot{\ell}^{(m+1)}(\hat{\nu}_n^{(m+1)}, \hat{\lambda}_n^{(m+1)}, \hat{\beta}_n^{(m)}).$$

根据假设, 第一项  $\hat{\beta}^{(m)}$  以概率收敛到  $\beta_0$ . 类似于引理 4.1 的证明, 序列

$$\ell^{(m+1)}(\boldsymbol{\theta}) - \ell^{(m+1)}(\boldsymbol{\theta}_0) : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq \delta$$

形成一个 Donsker 类, 这意味着:

$$\mathbb{P}_n \dot{\ell}^{(m+1)}(\hat{\nu}_n^{(m+1)}, \hat{\lambda}_n^{(m+1)}, \hat{\beta}_n^{(m)}) = o_P(n^{-1/2}).$$

由于每个  $I_j$  都是关于  $\boldsymbol{\theta}$  的连续函数, 利用连续映射定理得到

$$n \widetilde{V}_{m+1}(\hat{\nu}_n^{(m+1)}, \hat{\lambda}_n^{(m+1)}, \hat{\beta}_n^{(m)}) = O_P(1).$$

因此, 第二项是  $O_P(n^{-1/2}) = o_P(1)$ . 综上所述,  $\hat{\beta}_n^{(m+1)} \rightarrow_P \beta_0$ .

接下来, 通过利用  $(\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}, \hat{\beta}^{(j)})$  在  $j = 3, \dots, m$  时的收敛速度, 建立  $(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}, \hat{\beta}^{(m+1)})$  的渐近正态性. 为此, 提供一个关于  $(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}, \hat{\beta}^{(m+1)})$  渐近分布的递推关系, 并建立

$$\sqrt{n} \left[ (\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}, \hat{\beta}^{(m+1)}) - (\nu_0, \lambda_0, \beta_0) \right]$$

的弱收敛结果.  $(\hat{\nu}^{(3)}, \hat{\lambda}^{(3)}, \hat{\beta}^{(3)})$  的  $\sqrt{n}$ -一致性可通过 MLE 估计的  $\sqrt{n}$ -一致性来保证, 后者可通过经典统计方法得到验证. 在此基础上, 通过数学归纳法假设, 对于每个  $j = 3, \dots, m$ , 估计量  $(\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}, \hat{\beta}^{(j)})$  均满足  $\sqrt{n}$ -一致性. 对于  $(\nu, \lambda)$  的估计量, 因为  $(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)})$  最大化了  $\mathbb{P}_n \dot{\ell}^{m+1}(\nu, \lambda, \hat{\beta}^{(m)})$ , 对于固定的  $\hat{\beta}^{(m)}$  的值有

$$\begin{aligned} & \mathbb{P}_n \dot{\ell}_{(\nu, \lambda)}^{m+1}(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}, \hat{\beta}^{(m)}) \\ & \equiv \mathbb{P}_n \left. \frac{\partial}{\partial(\nu, \lambda)} \right|_{(\nu, \lambda) = (\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)})} \ell^{m+1}(\nu, \lambda, \hat{\beta}^{(m)}) = 0. \end{aligned}$$

通过向量  $(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}, \hat{\beta}^{(m)})$  的一致性, 并使用引理 4.1 得到

$$\mathbb{G}_n \dot{\ell}_{(\nu, \lambda)}^{m+1}(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}, \hat{\beta}^{(m)}) = \mathbb{G}_n \dot{\ell}_{(\nu, \lambda)}^{m+1}(\nu_0, \lambda_0, \beta_0) + o_P(1).$$

由  $P\dot{\ell}^{m+1}(\nu_0, \lambda_0, \beta_0) = 0$  得

$$\begin{aligned} \sqrt{n} \left[ P\dot{\ell}_{(\nu, \lambda)}^{m+1}(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}, \hat{\beta}^{(m)}) - P\dot{\ell}_{(\nu, \lambda)}^{m+1}(\nu_0, \lambda_0, \beta_0) \right] \\ = -\sqrt{n} \mathbb{P}_n \dot{\ell}_{(\nu, \lambda)}^{m+1}(\nu_0, \lambda_0, \beta_0) + o_P(1) \end{aligned}$$

对于每一个  $j$ , 得分函数的期望可以写为

$$P\dot{\ell}_{(\nu, \lambda)}^{(j)}(\boldsymbol{\theta}) = \begin{pmatrix} -\lambda(\Delta\Lambda_{j;0} - \Delta\Lambda_j) \\ \frac{\Delta\Lambda_{j;0}^2 - \Delta\Lambda_j^2 - \nu\lambda(\Delta\Lambda_{j;0} - \Delta\Lambda_j)^2}{2\lambda\Delta\Lambda_{j;0}^2} \end{pmatrix},$$

函数  $\boldsymbol{\theta} \mapsto P\dot{\ell}_{(\nu, \lambda)}^{m+1}(\boldsymbol{\theta})$  是连续可微的. 因此有

$$\begin{aligned} \sqrt{n} \left[ P\dot{\ell}_{(\nu, \lambda)}^{m+1}(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}, \hat{\beta}^{(m)}) - P\dot{\ell}_{(\nu, \lambda)}^{m+1}(\nu_0, \lambda_0, \beta_0) \right] \\ = -\mathcal{I}_{(\nu, \lambda)}^{m+1} \sqrt{n} \left[ (\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}) - (\nu_0, \lambda_0) \right] - \mathcal{I}_{(\nu, \lambda), \beta}^{m+1} \sqrt{n} (\hat{\beta}^{(m)} - \beta_0) \\ + O_P \left( \sqrt{n} \|(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}) - (\nu_0, \lambda_0)\|^2 + \sqrt{n} \|\hat{\beta}^{(m)} - \beta_0\|^2 \right). \end{aligned}$$

使用  $\hat{\beta}^{(m)}$  的  $\sqrt{n}$  一致性和  $\sqrt{n} \mathbb{P}_n \dot{\ell}_{(\nu, \lambda)}^{m+1}(\nu_0, \lambda_0, \beta_0)$ , 得到

$$(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}) - (\nu_0, \lambda_0) = O_p(n^{-1/2}),$$

这进一步意味着如下近似关系.

$$\begin{aligned} \mathcal{I}_{(\nu, \lambda)}^{m+1} \sqrt{n} \left[ (\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}) - (\nu_0, \lambda_0) \right] \\ = \sqrt{n} \mathbb{P}_n \dot{\ell}_{(\nu, \lambda)}^{m+1}(\nu_0, \lambda_0, \beta_0) - \mathcal{I}_{(\nu, \lambda), \beta}^{m+1} \sqrt{n} (\hat{\beta}^{(m)} - \beta_0) + o_P(1). \end{aligned} \quad (4.86)$$

对于  $\beta$  的估计, 迭代算法由以下公式给出

$$\hat{\beta}^{(m+1)} = \hat{\beta}^{(m)} + \tilde{V}_{m+1} n \mathbb{P}_n \dot{\ell}_{\beta}^{(m+1)}(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}, \hat{\beta}^{(m)}).$$

假设以下方程对  $m \geq 4$  都成立

$$\mathbb{P}_n \dot{\ell}_\beta^3(\hat{\nu}^{(3)}, \hat{\lambda}^{(3)}, \hat{\beta}^{(m)}) + \sum_{j=4}^m \mathbb{P}_n \dot{\ell}_\beta^{(j)}(\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}, \hat{\beta}^{(m)}) = o_P(n^{-1/2}). \quad (4.87)$$

实际上, 根据 MLE 估计的定义, 有  $\mathbb{P}_n \dot{\ell}_\beta^3(\hat{\nu}^{(3)}, \hat{\lambda}^{(3)}, \hat{\beta}^{(3)}) = o_P(n^{-1/2})$ . 接着, 将通过递归关系证明该等式在  $m = 4$  时成立, 并假设当方程 (4.87) 对  $m$  成立时, 它对  $m + 1$  也成立. 再将利用递推关系证明该等式在  $m = 4$  的情况下成立, 并在此基础上, 假设方程 (4.87) 对  $m$  成立, 则其对  $m + 1$  也成立. 若方程 (4.87) 成立, 可以推导出以下结论.

$$\begin{aligned} & n \widetilde{V}_{m+1}^{-1}(\hat{\beta}^{(m+1)} - \beta_0) - n \widetilde{V}_{m+1}^{-1}(\hat{\beta}^{(m)} - \beta_0) \\ &= \mathbb{P}_n \dot{\ell}_\beta^3(\hat{\nu}^{(3)}, \hat{\lambda}^{(3)}, \hat{\beta}^{(m)}) + \sum_{j=4}^{m+1} \mathbb{P}_n \dot{\ell}_\beta^{(j)}(\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}, \hat{\beta}^{(m)}) + o_P(n^{-1/2}). \end{aligned} \quad (4.88)$$

根据  $(\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}, \hat{\beta}^{(m)})$  的一致性以及引理 4.1, 可得以下结论.

$$\mathbb{G}_n \dot{\ell}_\beta^{(j)}(\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}, \hat{\beta}^{(m)}) - \mathbb{G}_n \dot{\ell}_\beta^{(j)}(\boldsymbol{\theta}_0) = o_P(1), \text{ 任意 } 1 \leq j \leq m+1,$$

这意味着  $\mathbb{P}_n \dot{\ell}_\beta^{(j)}(\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}, \hat{\beta}^{(m)})$  可以表示为

$$\begin{aligned} \mathbb{P}_n \dot{\ell}_\beta^{(j)}(\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}, \hat{\beta}^{(m)}) &= \mathbb{P}_n \dot{\ell}_\beta^{(j)}(\boldsymbol{\theta}_0) - \mathcal{J}_\beta^{(j)}(\hat{\beta}^{(m)} - \beta_0) \\ &\quad - \mathcal{J}_{(\nu, \lambda), \beta}^{(j)} \left[ (\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}) - (\nu_0, \lambda_0) \right] + O_P(n^{-1/2}). \end{aligned}$$

因此, (4.88) 的右侧可以表示为

$$\begin{aligned} & \mathbb{P}_n \dot{\ell}_\beta^{m+1}(\boldsymbol{\theta}_0) - \mathcal{J}_\beta^{m+1}(\hat{\beta}^{(m)} - \beta_0) + \sum_{j=4}^{m+1} \mathcal{J}_{\beta, (\nu, \lambda)}^{(j)} \left[ (\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}) - (\nu_0, \lambda_0) \right] \\ &+ \mathcal{J}_{\beta, (\nu, \lambda)}^3 \left[ (\hat{\nu}^{(3)}, \hat{\lambda}^{(3)}) - (\nu_0, \lambda_0) \right] + o_P(n^{-1/2}). \end{aligned}$$

借助  $n \widetilde{V} m + 1^{-1} \rightarrow_P \mathcal{J} \beta^{m+1}$  这一事实, 可以得到以下近似关系.

$$\mathcal{J}_\beta^{m+1} \sqrt{n} (\hat{\beta}^{(m+1)} - \beta_0) = \sqrt{n} \mathbb{P}_n \dot{\ell}_\beta^{m+1}(\boldsymbol{\theta}_0) - \mathcal{J}_{\beta, (\nu, \lambda)}^3 \sqrt{n} \left[ (\hat{\nu}^{(3)}, \hat{\lambda}^{(3)}) - (\nu_0, \lambda_0) \right]$$

$$-\sum_{j=4}^{m+1} \mathcal{J}_{\beta,(\nu,\lambda)}^{(j)} \sqrt{n} \left[ (\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}) - (\nu_0, \lambda_0) \right] + o_P(1). \quad (4.89)$$

因此证得  $\hat{\beta}^{(m+1)} - \beta_0 = O_P(n^{-1/2})$ .

接下来证明渐近性质可以从  $m$  传递到  $m+1$ , 即

$$\mathbb{P}_n \dot{\ell}_\beta^3(\hat{\nu}^{(3)}, \hat{\lambda}^{(3)}, \hat{\beta}^{(m+1)}) + \sum_{j=4}^{m+1} \mathbb{P}_n \dot{\ell}_\beta^{(j)}(\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}, \hat{\beta}^{(m+1)}) = o_P(n^{-1/2}).$$

再次使用引理 4.1, 对于每一个  $j$  有

$$\begin{aligned} & \mathbb{P}_n \dot{\ell}_\beta^{(j)}(\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}, \hat{\beta}^{(m+1)}) \\ &= \mathbb{P}_n \dot{\ell}_\beta^{(j)}(\boldsymbol{\theta}_0) - \mathcal{J}_\beta^{(j)} \left[ \hat{\beta}^{(m+1)} - \beta_0 \right] - \mathcal{J}_{\beta,(\nu,\lambda)}^{(j)} \left[ (\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}) - (\nu_0, \lambda_0) \right] + o_P(n^{-1/2}). \\ & \mathbb{P}_n \dot{\ell}_\beta^3(\hat{\nu}^{(3)}, \hat{\lambda}^{(3)}, \hat{\beta}^{(m+1)}) + \sum_{j=4}^{m+1} \mathbb{P}_n \dot{\ell}_\beta^{(j)}(\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}, \hat{\beta}^{(m+1)}) \\ &= \mathbb{P}_n \dot{\ell}_\beta^{m+1}(\boldsymbol{\theta}_0) - \mathcal{J}_\beta^{m+1} \left( \hat{\beta}^{(m+1)} - \beta_0 \right) - \mathcal{J}_{\beta,(\nu,\lambda)}^3 \left[ (\hat{\nu}^{(3)}, \hat{\lambda}^{(3)}) - (\nu_0, \lambda_0) \right] \\ & - \sum_{j=4}^{m+1} \mathcal{J}_{\beta,(\nu,\lambda)}^{(j)} \left[ (\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}) - (\nu_0, \lambda_0) \right] + o_P(n^{-1/2}). \end{aligned}$$

再利用式 (4.89) 将  $\hat{\beta}^{(m+1)} - \beta_0$  进行替换, 可以得到右侧等于  $o_P(n^{-1/2})$ , 从而证明了式 (4.87) 对于  $m+1$  成立. 因此, 公式 (4.87) 对于每个  $m \geq 4$  都成立.

最后, 需推导  $\Sigma_{m+1}$  的渐近协方差. 对于  $m=3$ , 利用 MLE 的渐近性质, 可以得到

$$\begin{aligned} \sqrt{n} \left[ (\hat{\nu}^{(3)}, \hat{\lambda}^{(3)}) - (\nu_0, \lambda_0) \right] &= (\tilde{\mathcal{J}}_{(\nu,\lambda)}^3)^{-1} \left( I_2, -\mathcal{J}_{(\nu,\lambda),\beta}^3 \left( \mathcal{J}_\beta^3 \right)^{-1} \right) \sqrt{n} \mathbb{P} \dot{\ell}^3(\boldsymbol{\theta}_0) + o_P(1), \\ \sqrt{n} \left( \hat{\beta}^{(3)} - \beta_0 \right) &= \left( -(\tilde{\mathcal{J}}_\beta^3)^{-1} \mathcal{J}_{\beta,(\nu,\lambda)}^3 \left( \mathcal{J}_{(\nu,\lambda)}^3 \right)^{-1}, (\tilde{\mathcal{J}}_\beta^3)^{-1} \right) \sqrt{n} \mathbb{P} \dot{\ell}^3(\boldsymbol{\theta}_0) + o_P(1), \end{aligned}$$

其中  $I_2$  是  $2 \times 2$  的单位矩阵,  $\tilde{\mathcal{J}}_{(\nu,\lambda)}^3$  和  $\tilde{\mathcal{J}}_\beta^3$  分别表示有效信息矩阵, 其定义如下.

$$\begin{aligned} \tilde{\mathcal{J}}_{(\nu,\lambda)}^3 &= \mathcal{J}_{(\nu,\lambda)}^3 - \mathcal{J}_{(\nu,\lambda),\beta}^3 \left( \mathcal{J}_\beta^3 \right)^{-1} \mathcal{J}_{\beta,(\nu,\lambda)}^3, \\ \tilde{\mathcal{J}}_\beta^3 &= \mathcal{J}_\beta^3 - \mathcal{J}_{\beta,(\nu,\lambda)}^3 \left( \mathcal{J}_{(\nu,\lambda)}^3 \right)^{-1} \mathcal{J}_{(\nu,\lambda),\beta}^3. \end{aligned}$$

对于  $m = 4$ , 根据递推关系有

$$\begin{aligned}
& \sqrt{n} \left[ (\hat{\nu}^{(4)}, \hat{\lambda}^{(4)}) - (\nu_0, \lambda_0) \right] \\
&= \left( \mathcal{J}_{(\nu, \lambda)}^4 \right)^{-1} \left[ \sqrt{n} \mathbb{P}_n \dot{\ell}_{(\nu, \lambda)}^4(\boldsymbol{\theta}_0) - \mathcal{J}_{(\nu, \lambda), \beta}^4 \sqrt{n} (\hat{\beta}^{(3)} - \beta_0) \right] + o_P(1) \\
&= \left( \mathcal{J}_{(\nu, \lambda)}^4 \right)^{-1} \left( I_2 + \mathcal{J}_{(\nu, \lambda), \beta}^4 \left( \tilde{\mathcal{J}}_{\beta}^3 \right)^{-1} \mathcal{J}_{\beta, (\nu, \lambda)}^3 \left( \mathcal{J}_{(\nu, \lambda)}^3 \right)^{-1}, \mathcal{J}_{(\nu, \lambda), \beta}^4 \left( \tilde{\mathcal{J}}_{\beta}^3 \right)^{-1} \right) \sqrt{n} \mathbb{P}_n \dot{\ell}^3(\boldsymbol{\theta}_0) \\
&\quad + \left( \mathcal{J}_{(\nu, \lambda)}^4 \right)^{-1} (I_2, 0) \sqrt{n} \mathbb{P}_n \dot{\ell}^{(4)}(\boldsymbol{\theta}_0) + o_P(1), \\
& \sqrt{n} (\hat{\beta}^{(4)} - \beta_0) \\
&= \left( \mathcal{J}_{\beta}^4 \right)^{-1} \left\{ \sqrt{n} \mathbb{P}_n \dot{\ell}_{\beta}^4(\boldsymbol{\theta}_0) - \mathcal{J}_{\beta, (\nu, \lambda)}^3 \sqrt{n} \left[ (\hat{\nu}^{(3)}, \hat{\lambda}^{(3)}) - (\nu_0, \lambda_0) \right] \right. \\
&\quad \left. - \mathcal{J}_{\beta, (\nu, \lambda)}^4 \sqrt{n} \left[ (\hat{\nu}^{(4)}, \hat{\lambda}^{(4)}) - (\nu_0, \lambda_0) \right] \right\} + o_P(1) \\
&= \left( \begin{array}{c} -\left( \mathcal{J}_{\beta}^4 \right)^{-1} \mathcal{J}_{\beta, (\nu, \lambda)}^3 (\tilde{\mathcal{J}}_{\beta}^3)^{-1} - \mathcal{J}_{\beta, (\nu, \lambda)}^4 - \mathcal{J}_{\beta, (\nu, \lambda), \beta}^4 \mathcal{J}_{(\nu, \lambda)}^4 (\tilde{\mathcal{J}}_{\beta}^3)^{-1} \mathcal{J}_{\beta, (\nu, \lambda)}^3 (\tilde{\mathcal{J}}_{\beta}^3)^{-1} \\ \left( \mathcal{J}_{\beta}^4 \right)^{-1} + \left( \mathcal{J}_{\beta}^4 \right)^{-1} \mathcal{J}_{\beta, (\nu, \lambda)}^3 (\tilde{\mathcal{J}}_{\beta}^3)^{-1} \mathcal{J}_{(\nu, \lambda), \beta}^3 (\tilde{\mathcal{J}}_{\beta}^3)^{-1} - \mathcal{J}_{\beta, (\nu, \lambda)}^4 \mathcal{J}_{(\nu, \lambda), \beta}^4 (\tilde{\mathcal{J}}_{\beta}^3)^{-1} \end{array} \right)^{\top} \\
&\quad \times \sqrt{n} \mathbb{P}_n \dot{\ell}^3(\boldsymbol{\theta}_0) + \left( -\left( \mathcal{J}_{\beta}^4 \right)^{-1} \mathcal{J}_{\beta, (\nu, \lambda)}^4 (\mathcal{J}_{(\nu, \lambda)}^4)^{-1}, \left( \mathcal{J}_{\beta}^4 \right)^{-1} \right) \sqrt{n} \mathbb{P}_n \dot{\ell}^{(4)}(\boldsymbol{\theta}_0) + o_P(1),
\end{aligned}$$

其中  $\sqrt{n} \mathbb{P}_n \dot{\ell}^3(\boldsymbol{\theta}_0) \rightsquigarrow N(0, \mathcal{J}^3)$  和  $\sqrt{n} \mathbb{P}_n \dot{\ell}^{(4)}(\boldsymbol{\theta}_0) \rightsquigarrow N(0, \mathcal{J}^4)$  是独立的. 结合这两个近似结果, 可以得到  $\Sigma_4$  的表达式.

要建立对于每个  $m \geq 4$  的渐近协方差  $\Sigma_{m+1}$ , 注意  $\sqrt{n} (\hat{\beta}^{(m)} - \beta_0)$  可以表示为

$$\begin{aligned}
\sqrt{n} (\hat{\beta}^{(m)} - \beta_0) &= A_{m,3} \sqrt{n} \mathbb{P}_n \dot{\ell}^3(\boldsymbol{\theta}_0) + A_{m,4} \sqrt{n} \mathbb{P}_n \dot{\ell}^{(4)}(\boldsymbol{\theta}_0) \\
&\quad + \cdots + A_{m,m} \sqrt{n} \mathbb{P}_n \dot{\ell}^{(m)}(\boldsymbol{\theta}_0),
\end{aligned}$$

其中  $A_{m,3}, \dots, A_{m,m}$  是  $3 \times 1$  的矩阵. 接下来为这些系数矩阵建立一个递推关系. 根据  $\sqrt{n} (\hat{\beta}^{(m)} - \beta_0)$  的近似式, 得到

$$\begin{aligned}
& \sqrt{n} \left[ (\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}) - (\nu_0, \lambda_0) \right] \\
&= \left( \mathcal{J}_{(\nu, \lambda)}^{m+1} \right)^{-1} (I_2, 0) \sqrt{n} \mathbb{P}_n \dot{\ell}^{m+1}(\boldsymbol{\theta}_0) - \left( \mathcal{J}_{(\nu, \lambda)}^{m+1} \right)^{-1} \mathcal{J}_{(\nu, \lambda), \beta}^{m+1} \sqrt{n} (\hat{\beta}^{(m)} - \beta_0) \\
&= \left( \mathcal{J}_{(\nu, \lambda)}^{m+1} \right)^{-1} \left\{ [(I_2, 0) - A_{m,3}] \sqrt{n} \mathbb{P}_n \dot{\ell}^3(\boldsymbol{\theta}_0) + \sum_{j=4}^m [(I_2, 0) - A_{m,j}] \sqrt{n} \mathbb{P}_n \dot{\ell}^{(j)}(\boldsymbol{\theta}_0) \right\}
\end{aligned}$$

$$+ \sum_{j=4}^m \left( \mathcal{J}_{(\nu, \lambda)}^{m+1} \right)^{-1} (I_2, 0) \sqrt{n} \mathbb{P}_n \dot{\ell}^{(m+1)}(\boldsymbol{\theta}_0).$$

再利用式 (4.87) 中关于  $\sqrt{n} [(\hat{\nu}^{(j)}, \hat{\lambda}^{(j)}) - (\nu_0, \lambda_0)]$  的近似, 可以得到

$$\begin{aligned} A_{m+1,3} &= \left( \mathcal{J}_{\beta}^{m+1} \right)^{-1} \left\{ (0, 0, 1) - \mathcal{J}_{\beta, (\nu, \lambda)}^3 A_{3,3} \right. \\ &\quad \left. - \sum_{j=4}^m \mathcal{J}_{\beta, (\nu, \lambda)}^{(j)} \left( \mathcal{J}_{(\nu, \lambda)}^j \right)^{-1} [(I_2, 0) - A_{m,3}] \right\}, \\ A_{m+1,j} &= \left( \mathcal{J}_{\beta}^{m+1} \right)^{-1} \left\{ (0, 0, 1) - \mathcal{J}_{\beta, (\nu, \lambda)}^{(j)} \left( \mathcal{J}_{(\nu, \lambda)}^j \right)^{-1} (I_2, 0) \right. \\ &\quad \left. - \sum_{l=j+1}^m \mathcal{J}_{\beta, (\nu, \lambda)}^{(l)} \left( \mathcal{J}_{(\nu, \lambda)}^l \right)^{-1} [(I_2, 0) - A_{m,l}] \right\}, \\ A_{m+1,m+1} &= \left( \mathcal{J}_{\beta}^{m+1} \right)^{-1} (0, 0, 1), \end{aligned}$$

其中  $j = 4, \dots, m$ . 结合  $\sqrt{n} [(\hat{\nu}^{(m+1)}, \hat{\lambda}^{(m+1)}) - (\nu_0, \lambda_0)]$  和  $\sqrt{n}(\hat{\beta}^{(m)} - \beta_0)$  的近似, 可以得到协方差矩阵  $\Sigma_{m+1}$ .

### 定理 4.6 证明

基于 (4.80),  $\mathcal{X}_m$  的 CDF 可以表示为  $\mathbb{E}_{\pi(\nu|y_m)}[F_{\mathcal{X}_m}(x | \nu, y_m)]$ , 并且已知  $\pi(\nu | y_m)$  是一个正态分布, 其均值和方差分别为

$$\mu_m = \frac{\lambda \Lambda_{\beta}(t_m) + \mu \sigma^{-2}}{\lambda y_m + \sigma^{-2}}, \quad \tau_m = (\lambda y_m + \sigma^{-2})^{-1}.$$

另一方面,  $F_{\mathcal{X}_m}(x | \nu, y_m)$  在 (4.67) 中给出, 由两个部分组成, 其形式为  $\Phi(aX + b)$  和  $\exp(cX)\Phi(aX + b)$ , 其中  $a, b$  和  $c$  为实数. 因此,  $F_{\mathcal{X}_m}(x | \nu)$  可以根据以下结果推导 (Si 等, 2013):

- (i) 如果  $X \sim \mathcal{N}(w, \delta^2)$ , 且  $a, b \in \mathbb{R}$ , 则有  $\mathbb{E}[\Phi(aX + b)] = \Phi\left(\frac{aw+b}{\sqrt{1+a^2\delta^2}}\right)$ .
- (ii) 如果  $X \sim \mathcal{N}(w, \delta^2)$ , 且  $a, b, c \in \mathbb{R}$ , 则

$$\mathbb{E}[\exp(cX)\Phi(aX + b)] = \exp\left(cw + c^2\delta^2/2\right) \Phi\left(\frac{aw+b+ac\delta^2}{\sqrt{1+a^2\delta^2}}\right).$$

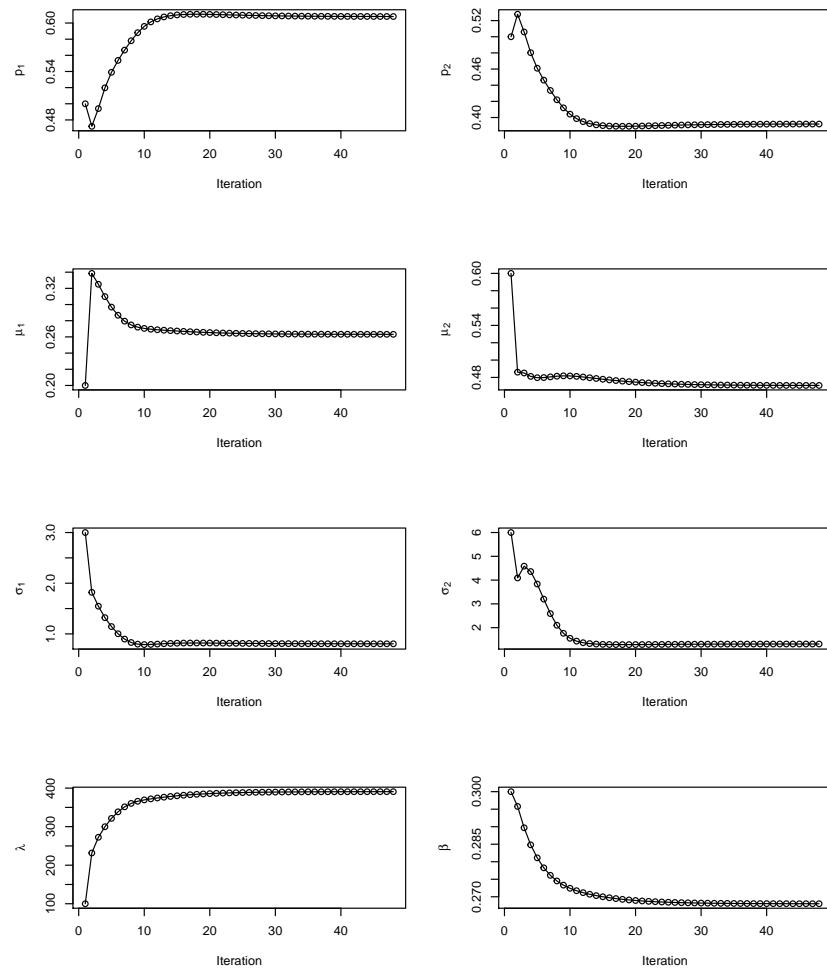


图 4.5: 模型参数估迭代图.

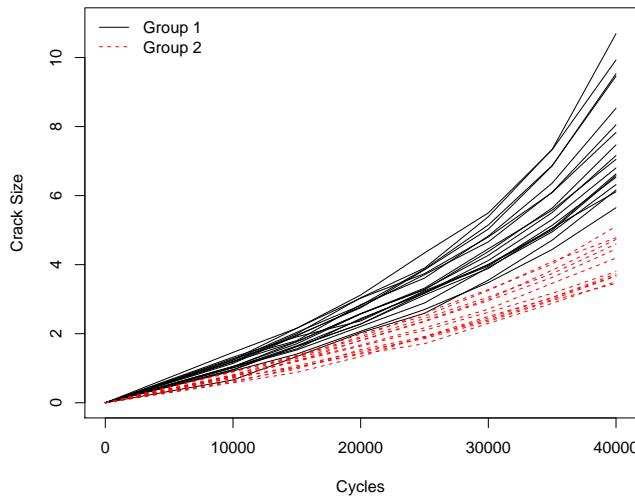
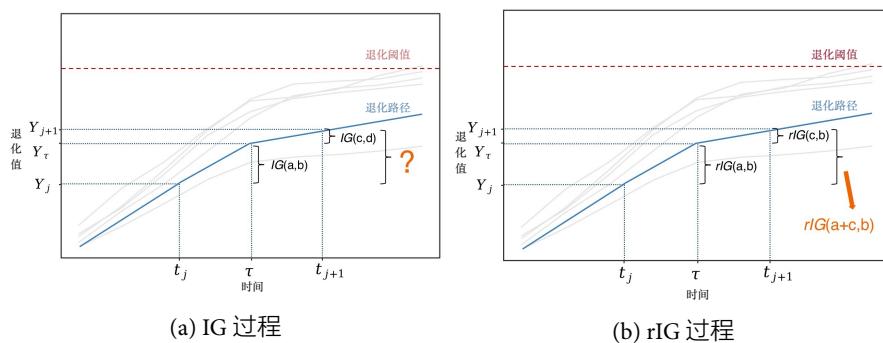
图 4.6: 当  $K = 2$  时, 疲劳裂纹数据的单元聚类.

图 4.7: 两种类型的两阶段随机过程.

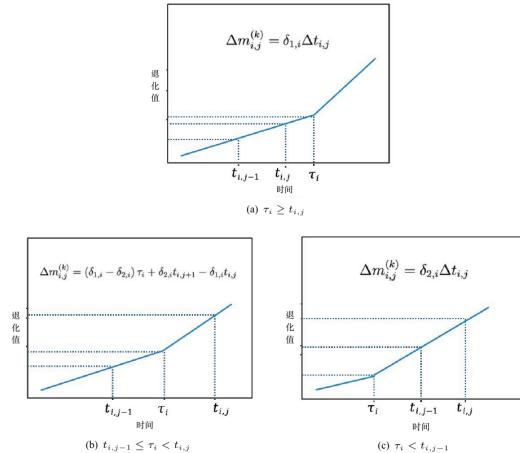


图 4.8: 变点与检测时间之间的三种关系.

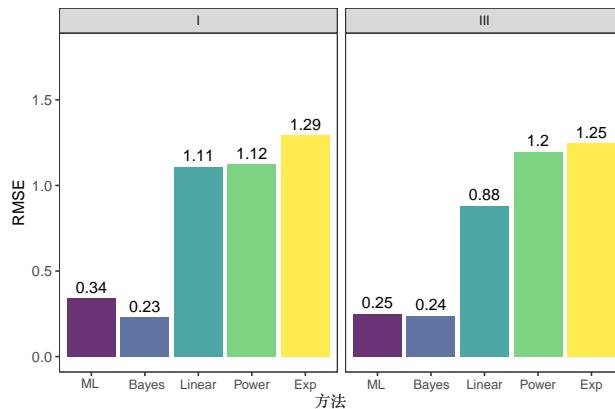


图 4.9: 基于不同模型的 MTTF 估计结果比较.

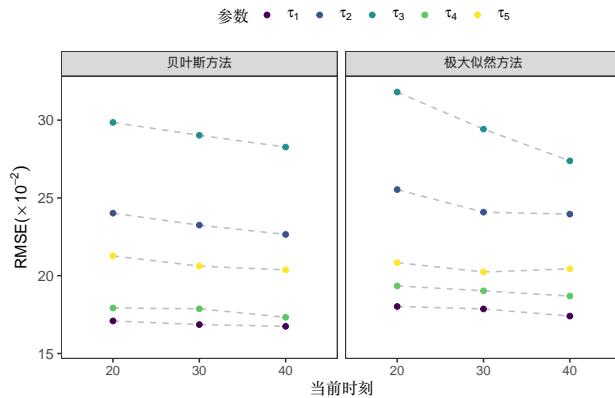


图 4.10: 不同时间点下变点估计的平均 RMSE.

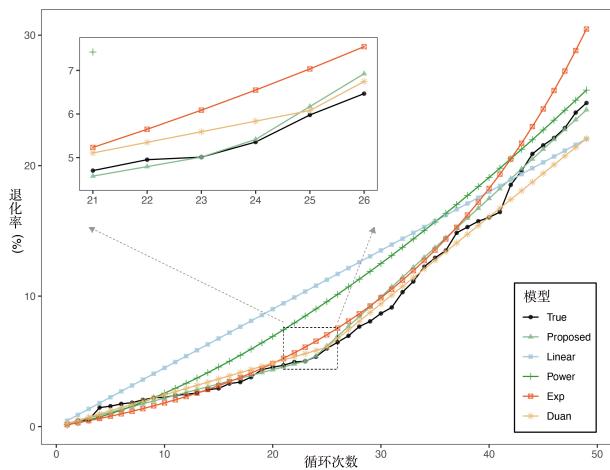


图 4.11: 使用不同方法对电池 #2 进行退化路径训练和预测结果.

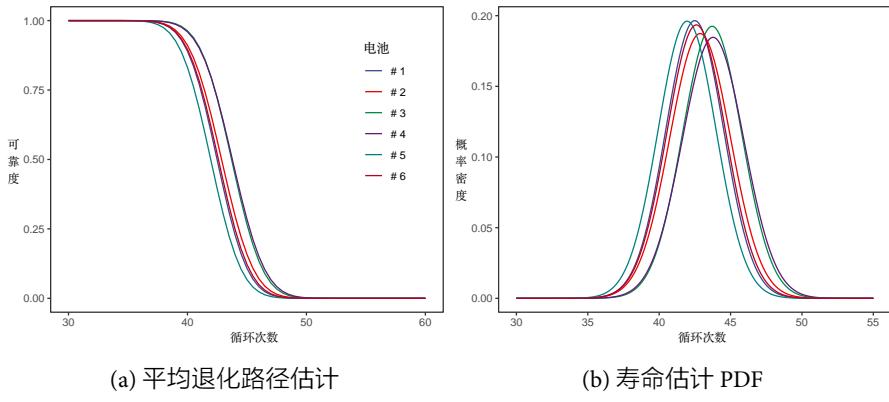


图 4.12: 基于贝叶斯方法的可靠度和失效时间密度函数.

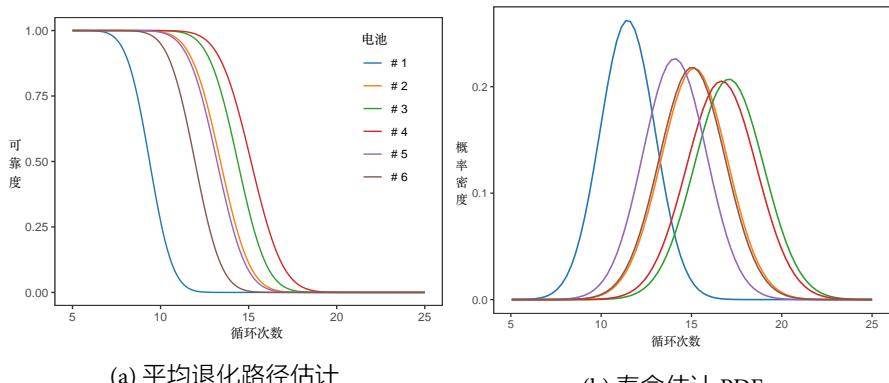


图 4.13: 基于贝叶斯方法的 RUL 在第 30 个周期的可靠度和密度函数.

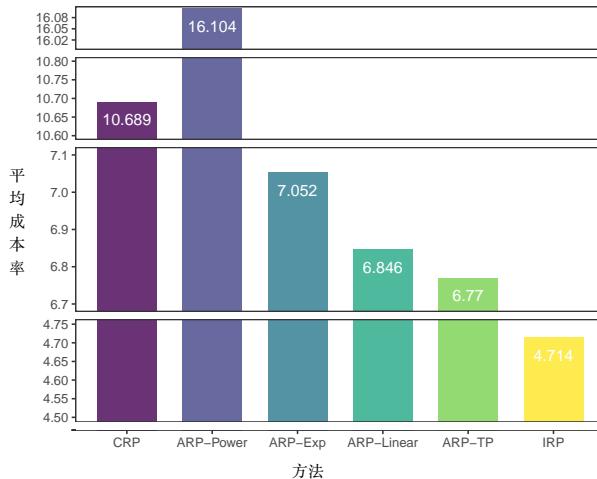


图 4.14: 每个策略的平均成本率.

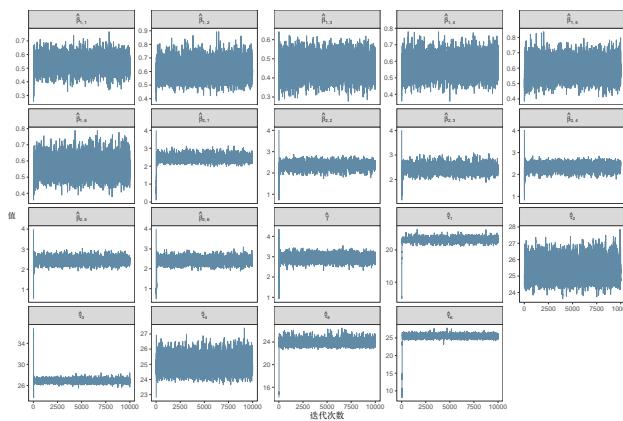


图 4.15: 模型参数后验样本轨迹图.

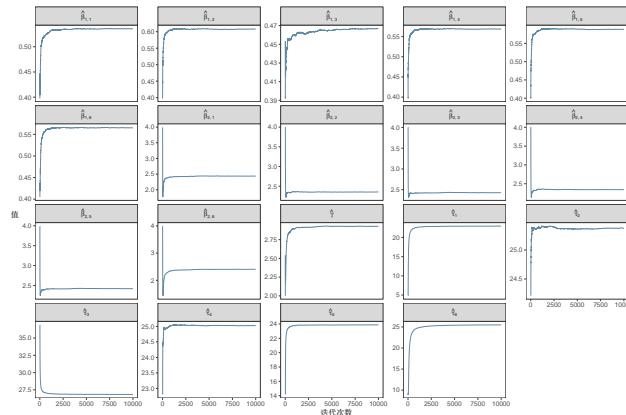


图 4.16: 模型参数后验样本遍历平均图.

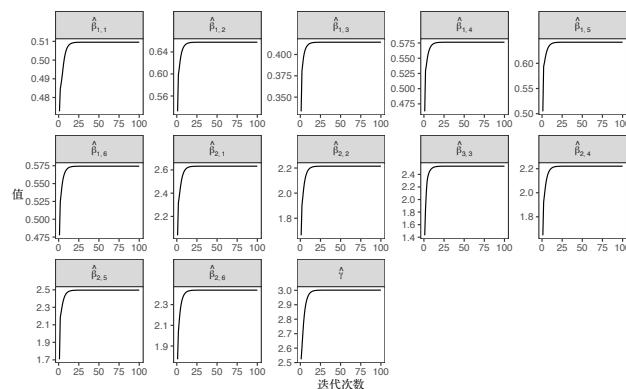
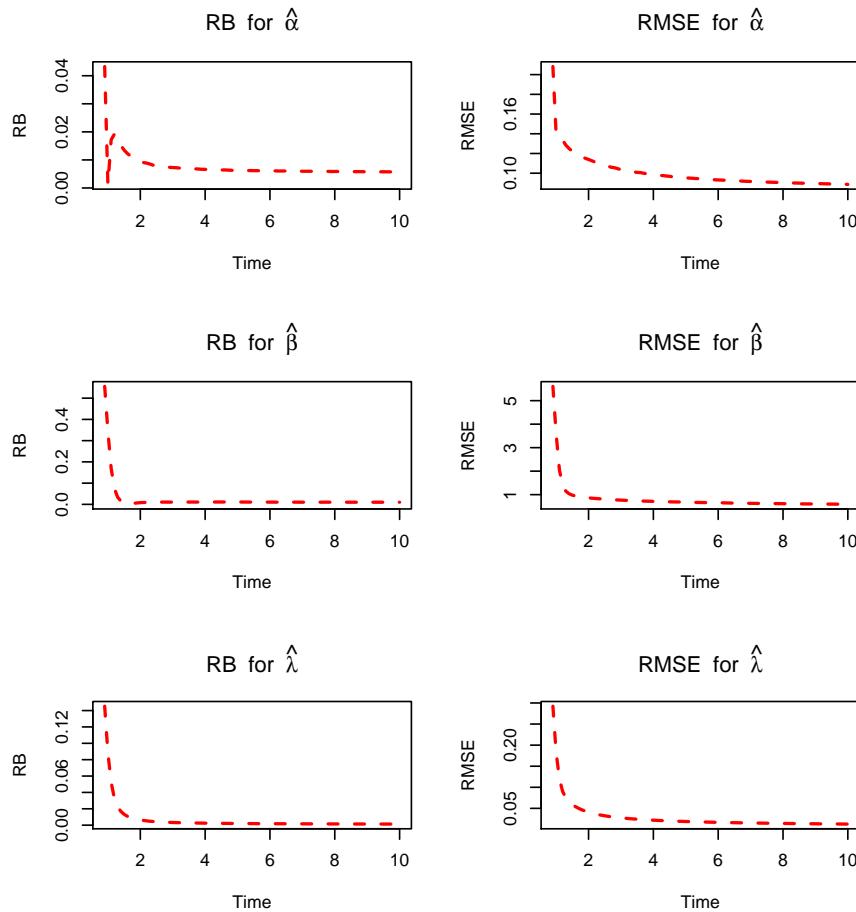


图 4.17: EM 算法中各参数迭代过程.

图 4.18:  $\nu$ 、 $\lambda$  和  $\beta$  参数的估计性能.

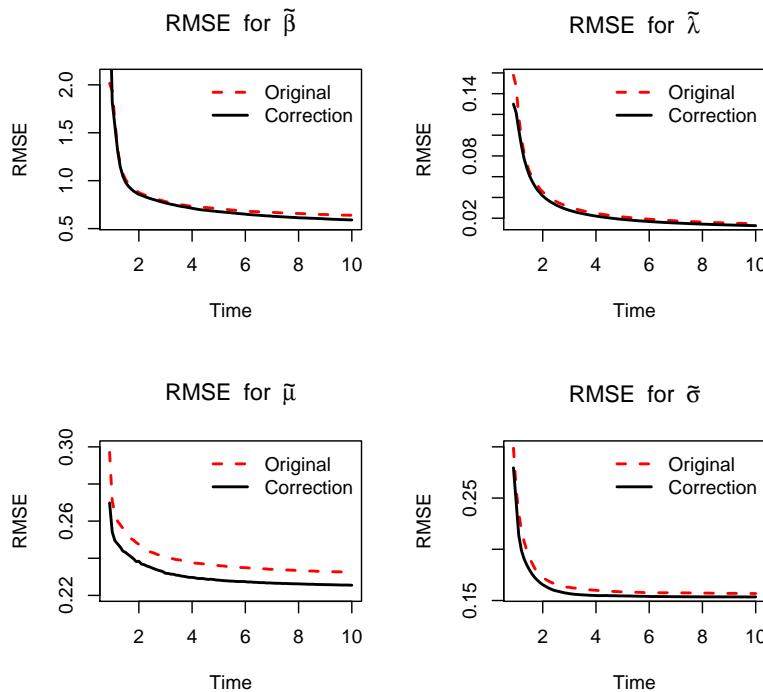


图 4.19: 基于含随机效应的 IG 模型, 不同方法下参数估计的 RMSE.

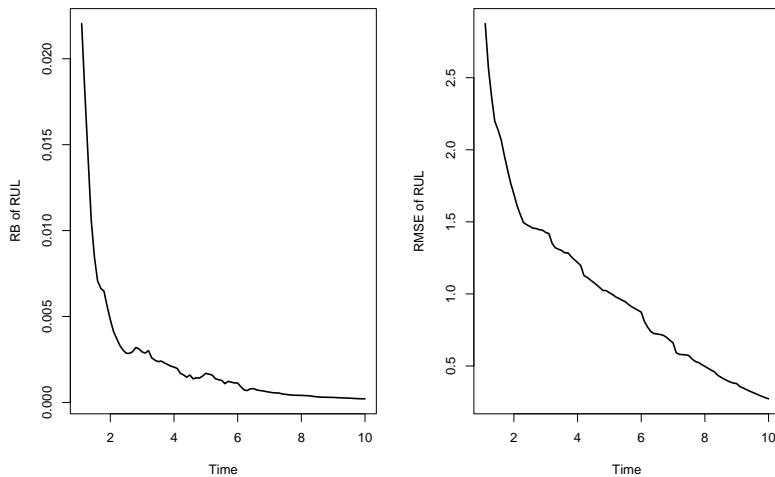


图 4.20: 第一个系统 RUL 预测的 RB 和 RMSE.

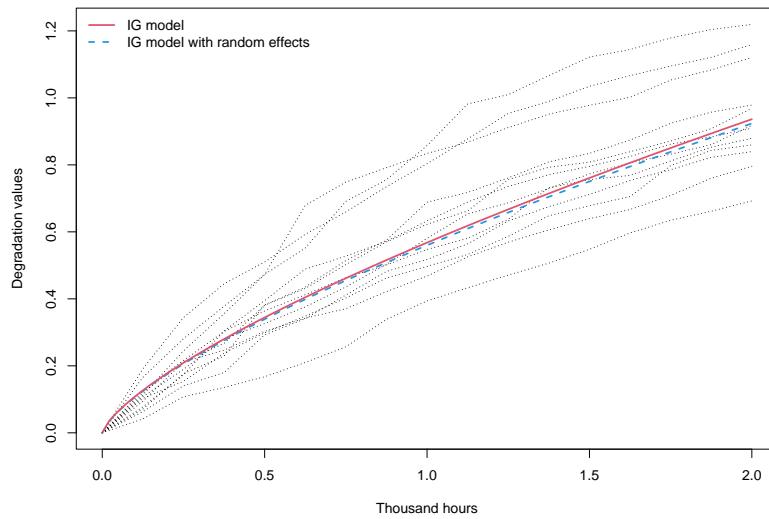


图 4.21: 基于集成电路器件的退化数据的不同模型拟合情况.

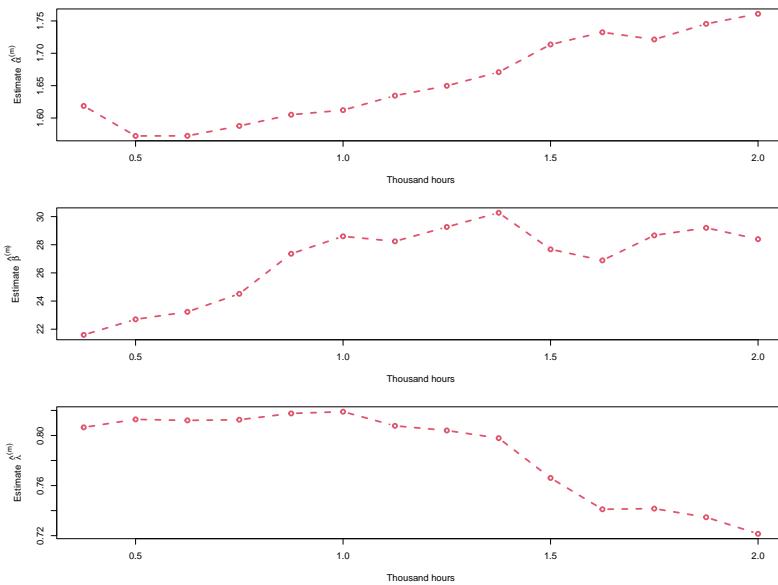


图 4.22: 基于集成电路器件退化数据, 无随机效应的 IG 过程的参数估计演变轨迹.

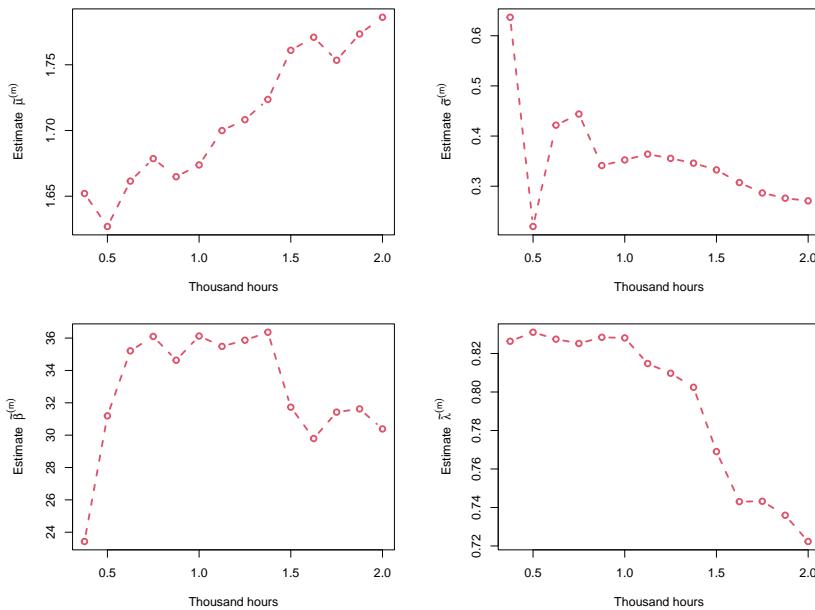


图 4.23: 基于集成电路器件退化数据, 含随机效应的 IG 过程的参数估计值演变轨迹.

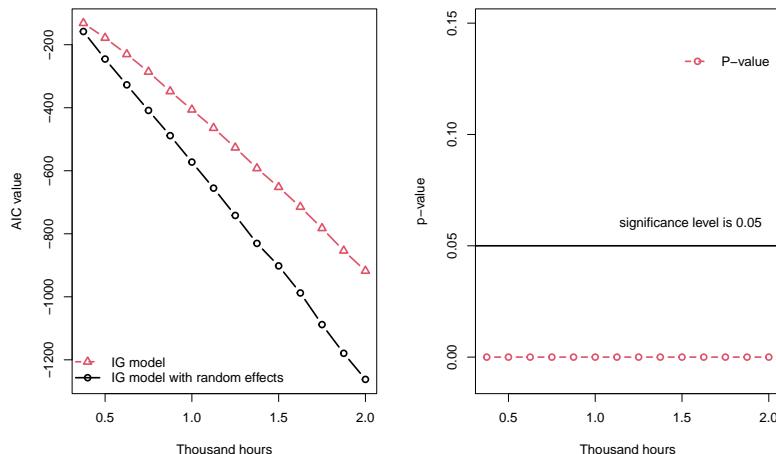


图 4.24: 基于集成电路器件退化数据, 两个 IG 模型的 AIC 值和  $p$  值.

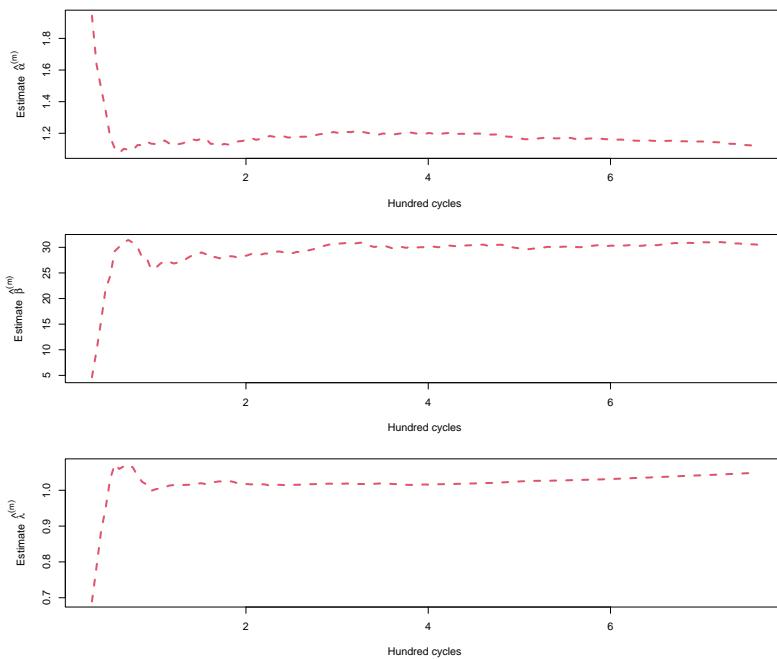


图 4.25: 基于锂离子电池退化数据, 无随机效应的 IG 过程的参数估计演变轨迹.

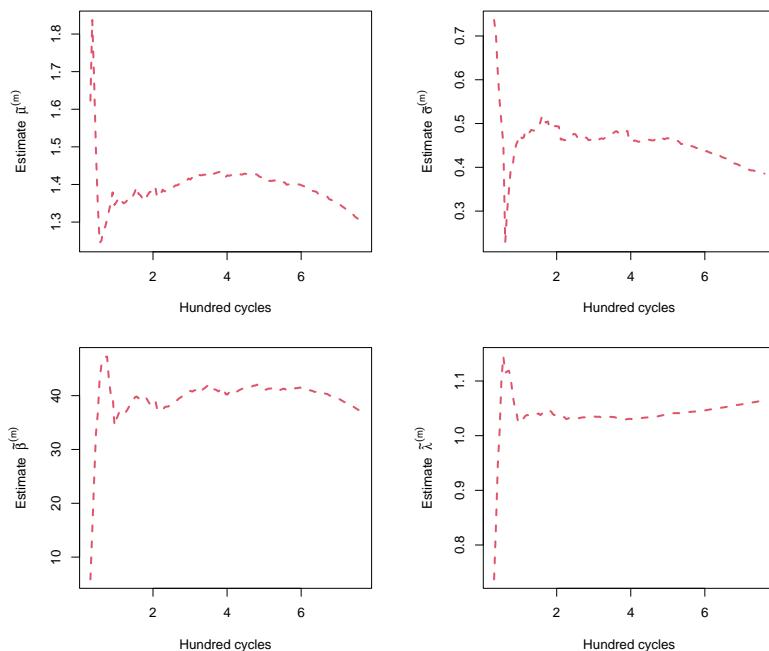


图 4.26: 基于锂离子电池退化数据, 带有随机效应的 IG 过程的参数估计值演变轨迹.

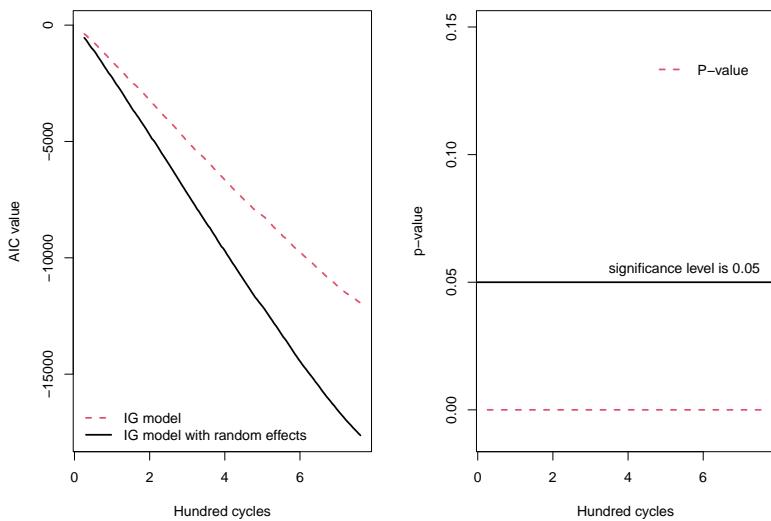


图 4.27: 基于锂离子电池退化数据, 两个 IG 模型的 AIC 值和  $p$  值.

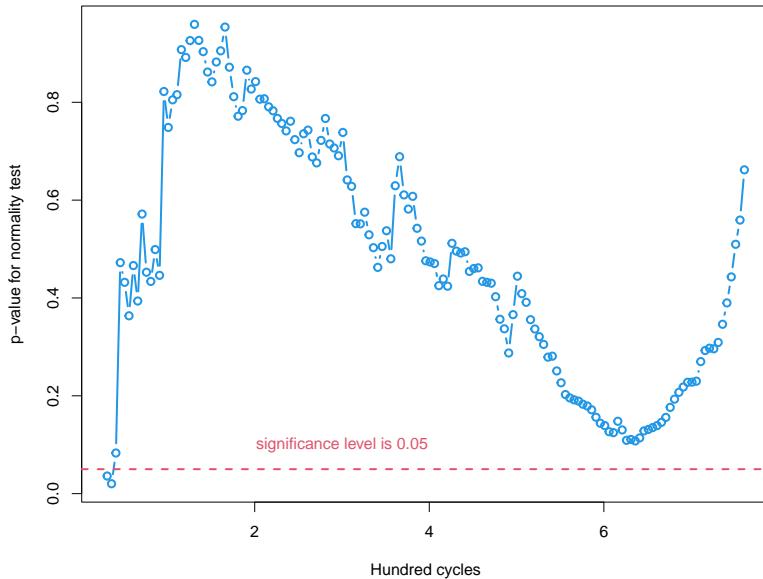


图 4.28: 基于锂离子电池退化数据的随机效应正态性检验  $p$  值.

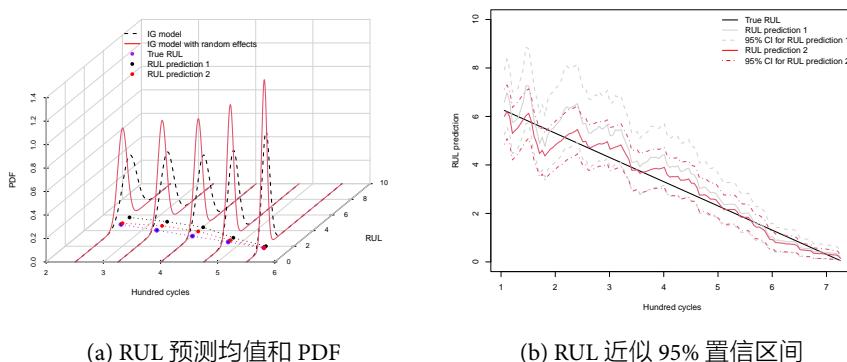


图 4.29: 基于锂离子电池退化数据, 两种 IG 模型的锂离子电池 #6 的 RUL 预测结果.

# 第五章 基于指数分散过程的统计建模

## 5.1 指数分散过程

尽管随机过程模型能够较好地拟合大多数退化数据, 但数据分析通常依赖于特定的模型假设, 或者通过 AIC 准则进行模型选择. 然而, AIC 准则存在以下局限性: 当所有候选模型均不适用时, AIC 无法提供有效的选择; 且只有真实模型包含在候选模型中时, AIC 能选出合适的模型. 因此, 为了更有效地分析退化数据, 开发一种通用的退化模型显得尤为重要. Tseng 等 (2016) 提出的指数分散 (ED) 过程是一种通用的退化建模框架, 涵盖了维纳过程、伽马过程和逆高斯过程. 但其假设的平均退化路径是线性的, 且主要研究加速退化试验的最优设计问题. 线性假设虽然适用于某些场景, 但无法拟合疲劳裂纹扩展 (Wu 等, 2003) 和应力松弛数据 (Yang, 2008) 等具有非线性特性的情况.

为克服这些局限性, 本文将构建了一类非线性 ED 过程, 其主要包括: (1) 推广 Tseng 等 (2016) 模型至非线性退化路径, 以适应更广泛的数据特性; (2) 采用单位偏差 (unit deviance) 与鞍点近似简化似然函数, 便于进行最大似然估计, 并讨论模型的统计性质及参数的物理意义; (3) 在实际应用中, 引入加速应力与随机效应, 构建更具灵活性的扩展模型. 本节结构安排如下: 第 5.1.1 节介绍所提的 ED 过程及其性质; 第 5.1.2 节讨论参数估计方法; 第 5.1.3 节扩展至带解释变量的 ED 模型; 第 5.1.4 节进行仿真研究; 第 5.1.5 节分析三个真实数据集; 第 5.1.6 节探讨带随机效应的 ED 模型.

### 5.1.1 定义及相关性质

若随机过程  $\{Y(t) \mid t\}$  满足以下条件:

(i)  $Y(0) = 0$ ;

(ii)  $\{Y(t) \mid t\}$  具有独立增量;

(iii) 每个增量  $\Delta Y_j = Y(t_j) - Y(t_{j-1})$  服从 ED 分布, 其 PDF 为

$$f(\Delta y_j; \theta, \lambda, \alpha) = c(\Delta y_j \mid \Delta \Lambda(t_j; \alpha), \lambda) \exp \lambda [\Delta y_j \theta - \Delta \Lambda(t_j; \alpha) \kappa(\theta)], \quad (5.1)$$

其中  $\Delta \Lambda(t_j; \alpha) = \Lambda(t_j; \alpha) - \Lambda(t_{j-1}; \alpha)$ ,  $\Lambda(\cdot)$  是时间  $t$  的单调递增函数,  $\alpha$  为未知参数,  $c(\cdot)$  是正则化常数, 确保 PDF 在其支撑上的积分为 1;  $\kappa(\cdot)$  是累积量函数 (cumulant function), 其导数决定了 ED 分布的连续累积量.

称满足上述条件的随机过程为 ED 过程, 记为  $Y(t) \sim \mathcal{ED}(\mu \Lambda(t; \alpha), \lambda)$ , 其中  $\mu$  为漂移参数,  $\lambda$  为扩散参数. 需要注意的是, 参数  $\mu$  是  $\theta$  的函数, 两者存在一一对应关系. 在符号上采用  $\mu$  而非  $\theta$ , 是因为  $\mu$  有更清晰的物理意义, 这后面的均值和方差函数中可以体现. 当  $\Lambda(t; \alpha) = t$  时, ED 过程可简化为 Tseng 等 (2016) 提出的平稳 ED 退化模型.

通过矩母函数可得,  $\mathbb{E}(Y(t)) = \mu \Lambda(t; \alpha)$ ,  $\text{Var}(Y(t)) = V(\mu) \Lambda(t; \alpha) / \lambda$ , 其中  $\mu = \kappa'(\theta)$  和  $V(\mu) = \kappa''(\theta)$ .  $\kappa'(\theta)$  和  $\kappa''(\theta)$  关于  $\kappa(\theta)$  对  $\theta$  的一阶和二阶导数. 因此, 称  $V(\mu)$  为方差函数. 假设方差函数为

$$V(\mu) = \mu^d, d \in (-\infty, 0] \cup [1, \infty), \quad (5.2)$$

则该族 ED 模型也称为 Tweedie 模型 (Tweedie, 1984). 其特例包括: 维纳过程 ( $d = 0$ )、泊松过程 ( $d = 1$ )、复合泊松过程 ( $1 < d < 2$ )、伽马过程 ( $d = 2$ ) 和 IG 过程 ( $d = 3$ ). 根据上述假设,  $\kappa(\theta)$  的解析形式为 (Dunn 等, 2005)

$$\theta = \begin{cases} \frac{\mu^{1-d}}{1-d}, & d \neq 1, \\ \log \mu, & d = 1, \end{cases} \quad \kappa(\theta) = \begin{cases} \frac{\mu^{2-d}}{2-d}, & d \neq 2, \\ \log \mu, & d = 2. \end{cases} \quad (5.3)$$

表 5.1: Tweedie 模型的单位偏差.

随机过程	单位偏差
泊松过程 ( $d = 1$ )	$2 \left\{ \frac{y}{\Lambda(t; \alpha)} \log \left( \frac{y}{\Lambda(t; \alpha)\mu} \right) - \left( \frac{y}{\Lambda(t; \alpha)} - \mu \right) \right\}$
伽马过程 ( $d = 2$ )	$2 \left\{ \frac{y}{\Lambda(t; \alpha)} \log \left( \frac{y}{\Lambda(t; \alpha)\mu} \right) + \frac{y}{\Lambda(t; \alpha)\mu} - 1 \right\}$
其他过程 ( $d \neq 1, 2$ )	$2 \left\{ \frac{(y/\Lambda(t; \alpha))^{2-d}}{(1-d)(2-d)} - \frac{y\mu^{1-d}}{\Lambda(t; \alpha)(1-d)} + \frac{\mu^{2-d}}{2-d} \right\}$

因此,  $\kappa(\theta)$  也可以表示为

$$\kappa(\theta) = \begin{cases} \exp(\theta), & d = 1, \\ -\log(-\theta), & d = 2, \\ \frac{[(1-d)\theta]^{(2-d)/(1-d)}}{2-d}, & d \neq 1 \text{ 和 } 2. \end{cases}$$

参数  $d$  不仅决定方差函数的形式, 还反映了数据的离散程度. 例如, 当  $d = 0$  时,  $\text{Var}(Y(t)) = \Lambda(t; \alpha)/\lambda$ , 方差与  $\mu$  无关; 而当  $d \geq 1$  时, 方差与  $\mu$  相关, 且  $\mu$  越大, 方差越大. 对于  $\mu > 1$  的情况, 即使其理论平均退化路径相同, 由伽马过程或 IG 过程生成的退化数据相比于维纳过程会更分散. 在后续讨论中, 假设方差函数的形式为  $V(\mu) = \mu^d$ .

虽然 Tweedie 模型可作为拟合退化数据的通用工具, 但参数估计仍具有一定的难度. 这是因为在式 (5.1) 中, 函数  $c(\cdot)$  和  $\kappa(\cdot)$  的解析形式仅在某些特殊的  $d$  值下存在. 关于  $c(\cdot)$  和  $\kappa(\cdot)$  的详细讨论可参考 (Jørgensen, 1998). 通过 Jørgensen (1986)、Jørgensen (1987)、Jørgensen (1992) 等提出的单位偏差方法, 式 (5.1) 可写为

$$f(\Delta y_j; \Theta) = a(\Delta y_j \mid \Delta \Lambda(t_j; \alpha); \lambda) \exp \left\{ -\frac{\lambda \Delta \Lambda(t_j; \alpha)}{2} D(\Delta y_j \mid \Delta \Lambda(t_j; \alpha); \mu) \right\}, \quad (5.4)$$

其中  $\Theta = (\lambda, d, \mu, \alpha)$ , 单位偏差  $D(\Delta y_j \mid \Delta \Lambda(t_j; \alpha); \mu)$  定义为

$$D(\Delta y_j \mid \Delta \Lambda(t_j; \alpha); \mu) = 2 \left\{ \sup_{\theta \in \Theta} \{ \Delta y_j / \Delta \Lambda(t_j; \alpha) \theta - \kappa(\theta) \} \right\}$$

$$- \Delta y_j / \Delta \Lambda(t_j; \alpha) \theta + \kappa(\theta) \Big\}, \quad (5.5)$$

而正则化常数  $a(\Delta y_j | \Delta \Lambda(t_j; \alpha); \lambda)$  的表达形式为

$$\begin{aligned} a(\Delta y_j | \Delta \Lambda(t_j; \alpha); \lambda) &= c(\Delta y_j | \Delta \Lambda(t_j; \alpha), \lambda) \\ &\times \exp \left[ \lambda \Delta \Lambda(t_j; \alpha) \sup_{\theta \in \Theta} \{ \Delta y_j / \Delta \Lambda(t_j; \alpha) \theta - \kappa(\theta) \} \right]. \end{aligned} \quad (5.6)$$

其中  $D(\Delta y_j | \Delta \Lambda(t_j; \alpha); \mu)$  可通过对  $\theta$  求二阶导数得到. 表 5.1 列出了 Tweedie 模型中单位偏差的具体形式. 例如, 维纳过程对应  $d = 0$ , 其单位偏差为  $D(y | \Lambda(t; \alpha); \mu) = \left( \frac{y}{\Lambda(t; \alpha)} - \mu \right)^2$ . 需要注意, 式 (5.6) 中的  $a(\Delta y_j | \Delta \Lambda(t_j; \alpha); \lambda)$  与函数  $c(\cdot)$  相关. 为了简化  $f(\Delta y_j; \mu, \sigma, \alpha)$  的形式, 可采用鞍点近似. 根据 Jørgensen (1998),  $a(\Delta y_j | \Delta \Lambda(t_j; \alpha); \lambda)$  可近似表达为

$$a(\Delta y_j | \Delta \Lambda(t_j; \alpha); \lambda) \approx \{2\pi\lambda^{-1}\Delta\Lambda(t_j; \alpha)V(\Delta y_j / \Delta\Lambda(t_j; \alpha))\}^{-1/2}. \quad (5.7)$$

图 5.1 展示了当  $\Lambda(t; \alpha) = t$  时, 在四种不同参数设置下基于鞍点近似的 PDF. 可以看出, 鞍点近似的 PDF 能捕捉到真实 PDF 的峰值以及尾部特征, 效果较好.

记  $\omega$  为退化路径的失效阈值. 则产品寿命  $T$  定义为退化路径  $Y(t)$  首次穿过阈值  $\omega$  的时间:  $T = \inf\{t | Y(t) \geq \omega\}$ . 记可靠度函数为  $R(t; \Theta) = P(T \geq t; \Theta)$ . 通常,  $R(t; \Theta)$  的解析表达式难以获得, 但在某些特殊的  $d$  值下可进行解析评估. 例如, 当  $d = 0$  时, ED 模型简化为维纳过程, 此时产品寿命分布具有解析形式 (Xu, Shen, Wang, 等, 2018). 当  $d \geq 1$  时, ED 分布定义在非负值范围内, 可靠性函数可表示为:

$$R(t; \Theta) = P(T \geq t; \Theta) = P(Y(t) \geq \omega; \Theta) = R_{ED}(\omega; t, \Theta),$$

其中  $R_{ED}(\cdot)$  表示均值为  $\mu\Lambda(t; \alpha)$ 、方差为  $\mu^d\Lambda(t; \alpha)/\lambda$  的 ED 分布的可靠性函数. 当  $d < 0$  时, 直接计算  $R(t; \Theta)$  较为困难, 此时可采用蒙特卡洛方法. 通过从假设模型生成大量退化路径, 记录每条路径首次穿越阈值  $\omega$  的时间, 并以失效时间大于  $t$  的比例作为可靠性的估计值. 具体步骤如下:

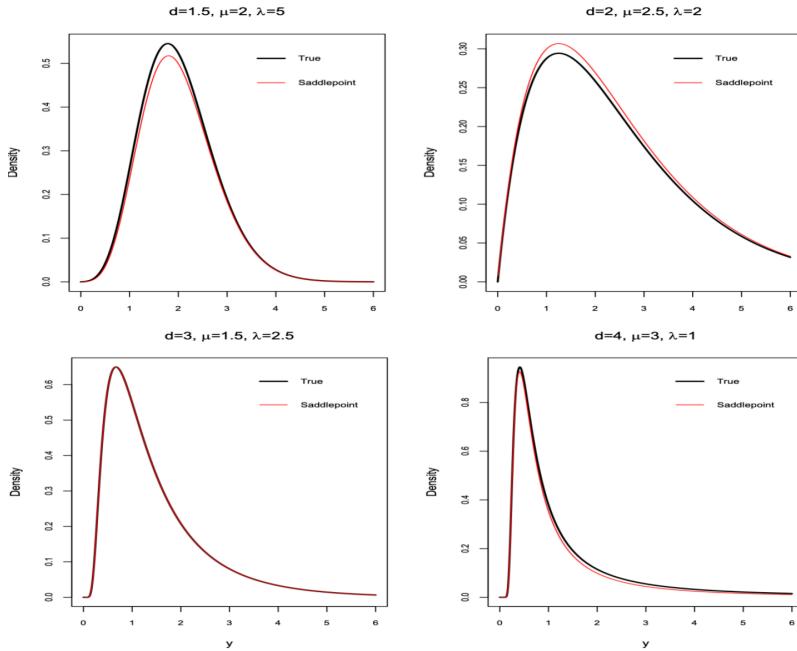


图 5.1: 不同设置下的鞍点近似.

### 基于蒙特卡洛模拟的可靠性计算

1. 设置检查时间  $t_{(1)}, \dots, t_{(M)}$ , 依次从  $ED(\hat{\mu} \Delta \Lambda(t_{(i)}; \hat{\alpha}), \hat{\lambda})$  中生成  $\Delta y_i, i = 1, \dots, M$ . 在时间间隔  $[t_{(L-1)}, t_{(L)}]$  内随机选取一个值, 作为累积退化值  $\sum_{i=1}^L \Delta y_i$  首次穿越阈值  $\omega$  时的失效时间.
2. 重复上述过程  $K$  次, 生成失效时间  $T_1, \dots, T_K$ .
3. 对于任意  $t$  值, 利用  $\frac{\sum_{i=1}^K I(T_i > t)}{K}$  估计  $R(t; \Theta)$ , 其中  $I(T_i > t) = 1$  表示  $T_i > t$ ; 否则  $I(T_i > t) = 0$ .

**注 1:** 事实上, 第 1 步计算出的失效时间存在误差: 实际失效时间应在测量区间  $[t_{(L-1)}, t_{(L)}]$  内. 这是由于退化轨迹首次超越预设阈值的精确时刻无法通过离散观测完全还原. 然而在实际应用中, 当满足  $\Delta t = t_{(L)} - t_{(L-1)}$  远远小于平均寿命时, 由此引入的系统误差具有量级可控特性. 例如, 假设产品的平

均寿命为 5000 小时, 当测量间隔选为 5 小时, 则理论偏差率小于 0.1%. 这种较基准值降低两个数量级的偏差幅度, 完全满足工程分析中对模型精度的容忍阈值.

### 5.1.2 统计推断

假设在退化试验中, 随机选择  $n$  个单元进行测试.  $y_{ij}$  表示第  $i$  个单元在测量时间  $t_{ij}$  的退化值,  $j = 1, \dots, m_i$ ,  $i = 1, \dots, n$ , 即每条退化路径的测量时间点不同. 设  $\Delta y_i = (\Delta y_{i1}, \dots, \Delta y_{im_i})'$ , 其中  $\Delta y_{i1} = y_{i1}$ ,  $\Delta y_{ij} = y_{ij} - y_{i(j-1)}$ ,  $i = 1, \dots, n$ , 并记  $y = (\Delta y_1, \dots, \Delta y_n)'$ . 假设  $y_{ij} \sim ED(\mu\Lambda(t_{ij}; \alpha), \lambda)$ , 则有  $\Delta y_{ij} \sim ED(\mu\Delta\Lambda(t_{ij}; \alpha), \lambda)$ . 根据式 (5.4) 和 (5.7),  $\Delta y_{ij}$  的 PDF 可近似为

$$f(\Delta y_{ij}; \Theta) \approx \left\{ 2\pi\lambda^{-1}\Delta\Lambda(t_{ij}; \alpha)V(\Delta y_{ij}/\Delta\Lambda(t_{ij}; \alpha)) \right\}^{-1/2} \times \exp \left\{ -\frac{\lambda\Delta\Lambda(t_{ij}; \alpha)}{2}D(\Delta y_{ij} | \Delta\Lambda(t_{ij}; \alpha); \mu) \right\}. \quad (5.8)$$

尽管式 (5.8) 是一个近似形式, 但如图 5.1 所示, 其与式 (5.4) 的偏差较小. 此外, 式 (5.8) 具有解析形式, 便于后续的统计推断. 如表 5.1 所示, 单位偏差分为三种情况: 对于  $d = 1$  和  $d = 2$  的情形, 分别对应泊松过程和伽马过程, 相关统计推断可参考 (Jørgensen, 1998); 这里仅使用  $d \neq 1$  和  $d \neq 2$  时单位偏差的表达式. 在实际数据分析中, 可先做假设检验, 以确定是否  $d = 1$  或  $d = 2$ . 因此, 不失一般性, 假设单位偏差的形式为

$$D(\Delta y_{ij} | \Delta\Lambda(t_{ij}; \alpha); \mu) = 2 \left( \frac{(\Delta y_{ij}/\Delta\Lambda(t_{ij}; \alpha))^{2-d}}{(1-d)(2-d)} - \frac{\Delta y_{ij}\mu^{1-d}}{\Delta\Lambda(t_{ij}; \alpha)(1-d)} + \frac{\mu^{2-d}}{2-d} \right).$$

参数  $\Theta$  的对数似然函数可写为

$$l(\Theta | y) = \sum_{i=1}^n \sum_{j=1}^{m_i} \left\{ \frac{1}{2} \log(\lambda) - \frac{1}{2} \log(2\pi) + \frac{d-1}{2} \log(\Delta\Lambda(t_{ij}; \alpha)) - \frac{d}{2} \log(\Delta y_{ij}) \right. \\ \left. - \lambda \left[ \frac{(\Delta y_{ij})^{2-d}(\Delta\Lambda(t_{ij}; \alpha))^{d-1}}{(1-d)(2-d)} - \frac{\Delta y_{ij}\mu^{1-d}}{1-d} + \frac{\Delta\Lambda(t_{ij}; \alpha)\mu^{2-d}}{2-d} \right] \right\}. \quad (5.9)$$

$l(\Theta | y)$  关于参数  $\mu$  和  $\lambda$  求一阶导数, 并建立似然方程求解可得

$$\begin{aligned}\mu &= \sum_{i=1}^n y_{im_i} \left( \sum_{i=1}^n \Lambda(t_{im_i}; \alpha) \right)^{-1}, \\ \lambda &= (1-d)(2-d) \sum_{i=1}^n \frac{m_i}{2} A^{-1},\end{aligned}\quad (5.10)$$

其中  $A$  表达式为

$$A = \sum_{i=1}^n \sum_{j=1}^{m_i} (\Delta y_{ij}^{d-2} \Delta \Lambda(t_{ij}; \alpha)^{d-1}) - \left( \sum_{i=1}^n y_{im_i} \right)^{2-d} \left( \sum_{i=1}^n \Lambda(t_{im_i}; \alpha) \right)^{d-1}.$$

将式 (5.10) 的  $\mu$  和  $\lambda$  代入到似然函数 (式 (5.9)) 中, 可以得到关于  $\alpha$  和  $d$  的轮廓似然函数

$$\begin{aligned}l(\alpha, d) &= \sum_{i=1}^n \left\{ m_i \left[ \log \left( \sum_{i=1}^n \frac{m_i}{2} \right) + \log((2-d)(1-d)) + \log(A) \right] \right. \\ &\quad \left. + \sum_{j=1}^{m_i} [(d-1) \log(\Delta \Lambda(t_{ij}; \alpha)) - d \log(\Delta(y_{ij}))] \right\} + \sum_{i=1}^n \frac{m_i}{2}. \quad (5.11)\end{aligned}$$

最终,  $\alpha$  和  $d$  的 MLE 可以通过优化  $l(\alpha, d)$  得到:

$$\{\hat{\alpha}, \hat{d}\} = \arg \max_{\alpha \in \Omega_\alpha, d \in \Omega_d} l(\alpha, d).$$

将  $(\hat{\alpha}, \hat{d})$  代回式 (5.10), 即可进一步获得  $\mu$  和  $\lambda$  的 MLE.

### 5.1.2.1 确定初始值

在使用 Newton-Raphson 方法求解参数  $\alpha$  和  $d$  的 MLE 时, 结果对初始值高度敏感, 选择不当可能导致收敛到局部最优解或根本无法收敛. 因此, 合理地提供初始参数值尤为重要. 设  $t_1, \dots, t_m$  为  $m$  个有序且互不相同的检验时间点. 假设在时间  $t_j$  时, 记录  $n_{t_j}$  个单元的测量数据, 其序号集合记为  $A_{t_j}$ . 以下步骤可用于确定初始值:

### 参数初始值确定

- 对于每个  $t_j$ , 计算平均退化值  $\bar{y}_{t_j} = \frac{1}{n_{t_j}} \sum_{i \in A_{t_j}} y_{ij}, j = 1, \dots, m.$
- 基于  $\mathbb{E}(Y(t)) = \mu \Lambda(t; \alpha)$ , 利用  $(t_j, \bar{y}_{t_j}), j = 1, \dots, m$ , 通过非线性拟合得到  $\alpha$  和  $\mu$  的初始估计值, 记为  $\tilde{\alpha}_{ini}$  和  $\tilde{\mu}_{ini}$ .
- 若  $n_{t_j} \geq 2$ , 计算退化值的方差  $S_{y_{t_j}}^2 = \frac{1}{n_{t_j}-1} \sum_{i \in A_{t_j}} (y_{ij} - \bar{y}_{t_j})^2.$
- 在  $\text{Var}(Y(t)) = \frac{\mu^d \Lambda(t; \alpha)}{\lambda}$  中, 使用  $\tilde{\alpha}_{ini}$  和  $\tilde{\mu}_{ini}$  替代  $\alpha$  和  $\mu$ . 基于  $(t_j, S_{y_{t_j}}^2), j = 1, \dots, m$ , 通过非线性拟合得到  $\lambda$  和  $d$  的初始估计值, 记为  $\tilde{\lambda}_{ini}$  和  $\tilde{d}_{ini}$ .

#### 5.1.2.2 区间估计

在获取参数的 MLE 后, 理论上可基于估计量的渐近正态性构建置信区间. 然而, 对于复杂模型, 直接推导 Fisher 信息矩阵可能非常困难. 此外, 基于观测 Fisher 信息矩阵进行计算时, 还可能出现方差估计值为负的情况, 导致结果不合理. 为解决上述问题, 这里采用自助法构造置信区间(Efron 等, 1993). 该方法的鲁棒性体现在: 通过重采样生成估计量的经验分布, 绕开对 Fisher 信息矩阵的理论依赖. 具体实施时, 注意到退化数据的时序特征, 采用基于独立同分布假设的单元层面重采样策略. 这种设计有效保持退化路径的时序完整性, 同时满足自助法的原始数据可交换性要求.

### 基于自助法构建置信区间

- 从原始数据集中有放回地随机抽取  $n$  个单元, 构建自助样本  $\{(t_{ij}^*, y_{ij}^*), i = 1, \dots, n; j = 1, \dots, m_i\}.$
- 基于自助样本计算参数的 MLE:  $\hat{\Theta}^{\text{boot}} = (\hat{\lambda}^{\text{boot}}, \hat{\mu}^{\text{boot}}, \hat{\alpha}^{\text{boot}}, \hat{d}^{\text{boot}}).$
- 重复上述两个步骤  $B$  次, 获得  $B$  个自助估计量

$$\{\hat{\Theta}_i^{\text{boot}} = (\hat{\lambda}_i^{\text{boot}}, \hat{\mu}_i^{\text{boot}}, \hat{\alpha}_i^{\text{boot}}, \hat{d}_i^{\text{boot}}), i = 1, \dots, B\}.$$

4. 参数函数  $G(\Theta)$  的近似  $(1 - \gamma)$  自助置信区间可以表示为  $(G_{\gamma/2}, G_{1-\gamma/2})$ , 其中  $G_{\gamma/2}$  和  $G_{1-\gamma/2}$  分别对应样本集合  $\{G(\hat{\Theta}_i^{\text{boot}}), i = 1, \dots, B\}$  的第  $(\gamma/2)$  和  $(1 - \gamma/2)$  分位数.

### 5.1.3 加速模型

在某些情况下, 常规退化试验可能无法收集到产品性能退化信息, 此时可通过加速退化试验 (Accelerated degradation testing, ADT) 加速产品性能的退化进程. 为此, 在 ED 退化模型中引入加速应力 (如温度、电压), 通过将式 (5.8) 中的  $\mu$  替换为  $\mu \exp(\beta' \phi(S))$ , 将这些加速应力加入到模型, 其中  $S$  和  $\beta$  分别表示应力水平和回归系数. 假设试验总共选取  $k$  个应力水平, 且在应力水平  $S_l$  下有  $n_l$  个单元进行测试, 其中  $l = 1, \dots, k$ . 因此, 总样本数为  $N = \sum_{l=1}^k n_l$ . 设  $y_{ijl}$  表示应力水平  $S_l$  下第  $i$  个单元在测量时间  $t_{ijl}$  的退化值, 其中  $i = 1, \dots, n_l$ ,  $j = 1, \dots, m_{il}$ . 为简化符号, 假设每组的测量时间相同, 即  $m_{1l} = \dots = m_{n_l l} = m_l$ .

假设  $y_{ijl} \sim ED(\mu(S_l)\Lambda(t_{ijl}; \alpha), \lambda)$ , 其中  $\mu(S_l) = \mu \exp(\beta' \phi(S_l))$ . 定义增量  $\Delta y_{ijl} = y_{ijl} - y_{i(j-1)l}$ , 其中  $y_{i0l} = 0$ . 则  $\Delta y_{ijl} = y_{ijl} - y_{i(j-1)l}$  服从扩散参数为  $\lambda$ , 漂移参数为  $\mu(S_l)$  的 ED 模型, 记作  $\Delta Y_i(t_{ijl} \mid x_l) \sim ED(\mu(x_l)\Delta\Lambda(t_{ijl}; \alpha), \lambda)$ , 其中  $\Delta\Lambda(t_{ijl}; \alpha) = \Delta\Lambda(t_{ijl}; \alpha) - \Delta\Lambda(t_{i(j-1)l}; \alpha)$ . 此时,  $\Delta y_{ijl}$  的 PDF 可写为

$$f(\Delta y_{ijl}; \Phi) = \{2\pi\lambda^{-1}\Delta\Lambda(t_{ijl}; \alpha)V(\Delta y_{ijl}/\Delta\Lambda(t_{ijl}; \alpha))\}^{-1/2} \times \exp\left\{-\frac{\lambda\Delta\Lambda(t_{ijl}; \alpha)}{2}D(\Delta y_{ijl} \mid \Delta\Lambda(t_{ijl}; \alpha); \mu(S_l))\right\}. \quad (5.12)$$

其中  $\Phi = (\lambda, d, \alpha, \mu, \beta)$ , 单位偏差

$$D(\Delta y_{ijl} \mid \Delta\Lambda(t_{ijl}; \alpha); \mu(S_l)) = \frac{(\Delta y_{ijl}/\Delta\Lambda(t_{ijl}; \alpha))^{2-d}}{(1-d)(2-d)} - \frac{\Delta y_{ijl}\mu(S_l)^{1-d}}{\Delta\Lambda(t_{ijl}; \alpha)(1-d)} + \frac{\mu(S_l)^{2-d}}{2-d}.$$

基于观测数据  $\{\Delta y_{ijl}, i = 1, \dots, n_l, j = 1, \dots, m_l, l = 1, \dots, k\}$ , 参数  $\Phi$  的对数

似然函数为

$$l(\Phi) = \sum_{l=1}^k \sum_{i=1}^{n_l} \sum_{j=1}^{m_l} \left[ \frac{1}{2} \log(\lambda) - \frac{1}{2} \log(2\pi) + \frac{d-1}{2} \log(\Delta\Lambda(t_{ijl}; \alpha)) \right. \\ \left. - \frac{d}{2} \log(\Delta y_{ijl}) - \lambda \left( \frac{(\Delta y_{ijl})^{2-d} (\Delta\Lambda(t_{ijl}; \alpha))^{d-1}}{(1-d)(2-d)} \right. \right. \\ \left. \left. - \frac{\Delta y_{ijl} \mu(S_l)^{1-d}}{1-d} + \frac{\Delta\Lambda(t_{ijl}; \alpha) \mu(S_l)^{2-d}}{2-d} \right) \right]. \quad (5.13)$$

$l(\Phi)$  关于  $\lambda$  求一阶导, 并令其等于 0, 可解出

$$\lambda = (1-d)(2-d) \sum_{l=1}^k \frac{m_l n_l}{2} B^{-1}, \quad (5.14)$$

其中

$$B = \sum_{l=1}^k \sum_{i=1}^{n_l} \left\{ \sum_{j=1}^{m_l} \Delta(y_{ijl})^{2-d} (\Delta\Lambda(t_{ijl}; \alpha))^{d-1} \right. \\ \left. - (2-d) \mu(S_l)^{1-d} [y_{im_l l} - (1-d) \mu(S_l) t_{im_l l}^\alpha] \right\}.$$

将(5.14)代入到  $l(\Phi)$  可得到参数的  $(d, \alpha, \mu, \beta)$  轮廓对数似然函数

$$l(d, \alpha, \mu, \beta) = \sum_{l=1}^k \left\{ m_l n_l \left[ \log \left( \sum_{l=1}^k \frac{m_l n_l}{2} \right) + \log((2-d)(1-d)) + \log(B) \right. \right. \\ \left. \left. + \sum_{i=1}^{n_l} \sum_{j=1}^{m_l} [(d-1) \log(\Delta\Lambda(t_{ijl}; \alpha)) - d \log(\Delta y_{ijl})] \right] \right\} + \sum_{l=1}^k \frac{m_l n_l}{2}. \quad (5.15)$$

通过对上式求最优化计算可得到参数  $(d, \alpha, \mu, \beta)$  的 MLE  $(\hat{d}, \hat{\alpha}, \hat{\mu}, \hat{\beta})$ , 进而代入式 (5.14) 得到  $\lambda$  的 MLE. 参数初始值和置信区间的选择方法与上一节类似.

#### 5.1.4 模拟实验

本节通过模拟研究评估所提方法的性能. 模型参数  $(d, \lambda, \mu)$  设定为  $(1.5, 0.05, 5)$ 、 $(2, 1, 2)$  和  $(3, 2, 2)$ , 分别对应复合泊松过程、伽马过程和 IG 过程. 假定

$\Lambda(t; \alpha) = t$ , 样本量  $n = 15, 30$  和  $60$ , 测量时间点为  $2, 4, \dots, 2m$ , 测量次数  $m = 5, 10$  和  $15$ . 每种组合重复模拟  $1000$  次, 以研究不同样本量和测量次数对参数估计性能的影响. 表 5.2 给出了估计量的均值和 MSE. 结果显示, 随着样本量和测量次数的增加, 估计量的 MSE 显著降低; 即使在样本量较小且测量次数较少的情况下 (如  $n = 15, m = 5$ ), 估计量的均值依然接近真实值. 为了进一步验证 ED 过程的优越性, 针对每组数据分别用 ED 过程、维纳过程、伽马过程和 IG 过程拟合, 比较可靠度函数估计的平均值, 并绘制拟合结果 (见图 5.2). 分析结果总结如下:

- (1) 可靠性接近真实值: 在所有情况下, 基于 ED 过程的可靠度函数估计均与产品的真实可靠度函数高度吻合.
- (2) **常用模型拟合不足:** 当  $d = 1.5$  (复合泊松过程) 时, 维纳过程、伽马过程和 IG 过程的拟合效果显著低于 ED 过程, 表明这些模型并不能很好地处理复合泊松退化数据.
- (3) **ED 过程的优越性能:** 即使在模型假设正确的情况下 (如伽马过程拟合  $d = 2$  的数据), ED 过程的拟合表现依然优于伽马过程. 尤其是在小样本量与测量次数不多 (如  $n = 15, m = 5$ ) 的场景下, 基于 ED 过程拟合的可靠度函数与真实函数的吻合度更高.

## 5.1.5 实例分析

### 5.1.5.1 激光退化数据

本节以一组  $80^{\circ}\text{C}$  下测试的 GaAs 激光器运行电流随时间的变化数据为例, 研究其退化特征. 当运行电流超过  $10$  时, 设备被判定为失效. 文献中分别使用维纳过程、伽马过程和 IG 过程对这组数据进行分析. 此外, Peng (2015) 比较了七种模型, 并基于 AIC 准则认为 IG 过程是最佳模型. 按照 Meeker 等 (1998) 的方法, 假设  $\Lambda(t; \alpha) = t$ , 利用 ED 退化模型对数据进行拟合. 参数的 MLE 和 95% CI 列于表 5.3 中. 图 5.3(a) 显示了估计的平均退化路径 ( $\hat{\mu}t$ ) 及其 95% CI. 可以发现, 虽然部分观测值落在置信带之外, 但样本平均路径完全位于置信带内, 说明模型较好地捕捉了退化趋势. 为进一步比较不同模型的拟合效果, 本文基于 AIC 指标对拟合质量进行了评价 (结果见表 5.4). 分析表明, ED 模型的拟合效果最佳 ( $AIC = -148.84$ ), 其次为 IG 模型 ( $AIC = -142.21$ )

表 5.2: 在  $(n, m)$  和  $(d, \lambda, \mu)$  的不同设置下, 模型参数估计的均值和 MSE ( $\times 1/100$ ).

$(n, m)$	$(d, \lambda, \mu)$	$d$		$\lambda$		$\mu$	
		均值	MSE	均值	MSE	均值	MSE
(15, 5)	(1.5, 0.05, 5)	1.717	52.392	0.055	0.172	5.004	7.567
	(2, 1, 2)	2.012	1.324	1.002	5.561	2.093	15.687
	(3, 2, 2)	2.975	22.521	2.002	11.205	2.067	8.942
(15, 10)	(1.5, 0.05, 5)	1.567	27.308	0.059	0.134	4.998	1.474
	(2, 1, 2)	2.006	0.058	1.001	2.952	2.034	6.366
	(3, 2, 2)	2.982	9.190	2.003	5.487	2.054	4.111
(15, 15)	(1.5, 0.05, 5)	1.553	19.961	0.057	0.114	5.001	0.526
	(2, 1, 2)	2.004	0.038	1.005	1.969	2.024	4.509
	(3, 2, 2)	2.983	9.053	1.997	3.393	2.029	3.025
(30, 5)	(1.5, 0.05, 5)	1.638	28.701	0.054	0.133	4.995	0.717
	(2, 1, 2)	2.007	0.058	0.996	2.529	2.047	6.894
	(3, 2, 2)	2.986	9.392	1.997	5.126	2.043	4.048
(30, 10)	(1.5, 0.05, 5)	1.531	15.278	0.057	0.099	5.001	0.360
	(2, 1, 2)	2.003	0.026	0.998	1.355	2.028	3.189
	(3, 2, 2)	2.986	4.094	1.996	2.442	2.029	2.258
(30, 15)	(1.5, 0.05, 5)	1.521	15.277	0.056	0.075	4.997	0.259
	(2, 1, 2)	2.003	0.016	0.998	0.929	2.022	2.008
	(3, 2, 2)	2.998	2.834	2.009	1.739	2.019	1.437
(60, 5)	(1.5, 0.05, 5)	1.531	14.813	0.056	0.091	4.996	0.348
	(2, 1, 2)	2.003	0.027	1.001	1.385	2.018	3.077
	(3, 2, 2)	2.996	4.615	1.998	2.846	2.037	2.224
(60, 10)	(1.5, 0.05, 5)	1.516	9.257	0.055	0.064	4.997	0.172
	(2, 1, 2)	2.001	0.012	1.003	0.717	2.009	1.677
	(3, 2, 2)	2.989	2.065	2.005	1.323	2.015	1.061
(60, 15)	(1.5, 0.05, 5)	1.509	6.177	0.053	0.043	4.998	0.119
	(2, 1, 2)	2.000	0.008	1.002	0.491	2.006	1.044
	(3, 2, 2)	2.996	1.507	2.001	0.927	2.004	0.677

表 5.3: 基于激光退化数据, 参数的 MLE 和 95% 置信区间.

$\hat{d}$	$\hat{\mu}$	$\hat{\lambda}$
3.61 (2.011, 4.305)	2.037 (1.927, 2.323)	0.013 (0.008, 0.107)

表 5.4: 基于激光退化数据, 不同退化模型的 AIC 值.

	维纳过程	伽马过程	IG 过程	ED 过程
AIC 值	-85.14	-130.22	-142.21	-148.84

. 此外, 从表 5.3 可以看到, 参数  $d$  的 95% 置信区间为  $(2.011, 4.305)$ , 其中包含 IG 模型的特定值  $d = 3$ , 这表明 IG 模型同样适用于该激光退化数据. 此外, 图 5.3(b) 展示了激光器设备的可靠度函数及其 95% 置信区间.

### 5.1.5.2 铝合金裂纹退化数据

这里以铝合金疲劳裂纹退化为例来说明当退化路径呈现非线性时 ED 模型的性能表现. Peng (2015) 使用维纳过程、伽马过程和 IG 过程对这组数据进行拟合, 并发现 IG 过程的拟合效果优于其他模型. 在此基础上, 我们采用 ED 退化模型, 并假设  $\Lambda(t; \alpha) = t^\alpha$ , 这与 Peng (2015) 的设定一致. 参数的 MLE 和 95% 自助置信区间列于表 5.5 中. 结果显示,  $d$  的 MLE 为 4.388, 表明拟合结果与现有模型存在显著差异. 表 5.6 给出了各模型的 AIC 值, 结果显示, ED 模型的 AIC 最低, 说明其在拟合疲劳裂纹数据方面具有明显优势. 图 5.4(a) 显示了估计的平均退化路径 ( $\hat{\mu}t^{\hat{\alpha}}$ ) 及其 95% 置信带. 与激光数据结果类似, 置信带能够很好地覆盖样本平均路径. 图 5.4(b) 显示了金属的可靠度函数及其 95% 置信区间.

表 5.5: 基于铝合金裂纹退化数据, 参数的 MLE 和 95% 置信区间

$\hat{d}$	$\hat{\alpha}$	$\hat{\mu}$	$\hat{\lambda}$
4.388 (2.036, 4.752)	1.333 (1.263, 1.473)	0.978 (0.736, 1.093)	0.187 (0.082, 0.267)

表 5.6: 基于铝合金裂纹退化数据, 不同退化模型的 AIC 值.

	维纳过程	伽马过程	IG 过程	ED 过程
AIC 值	632.49	123.39	105.69	94.77

表 5.7: 基于应力松弛数据, 参数的 MLE 和 95% 置信区间

$\hat{d}$	$\hat{\alpha}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\lambda}$
1.471	0.458	-2.034	1.84	0.434
(1.054, 1.487)	(0.438, 0.487)	(-2.179, -1.882)	(1.549, 1.953)	(0.234, 0.767)

### 5.1.5.3 应力关系退化数据

为了进一步验证模型的适用性, 我们分析了一组电气连接器在不同温度下测试的退化数据. 本试验随机选取了 18 个样本, 分别在  $65^{\circ}\text{C}$ 、 $85^{\circ}\text{C}$  和  $100^{\circ}\text{C}$  的应力水平下进行测试, 以评估在实际使用条件 ( $S_0 = 40^{\circ}\text{C}$ ) 下的可靠性.

当温度作为加速应力时, 根据 Arrhenius 模型可知

$$\mu(S_l) = \mu \exp(\beta\phi(S_l)) \quad \text{或} \quad \log(\mu(S_l)) = \log(\mu) + \beta\phi(S_l),$$

其中  $\phi(S_l) = 1/(273.15 + S_l)$ . 为便于分析, 将应力水平标准化为

$$x_1 = \frac{\phi(65) - \phi(40)}{\phi(100) - \phi(40)} = 0.46, x_2 = \frac{\phi(85) - \phi(40)}{\phi(100) - \phi(40)} = 0.78, x_3 = \frac{\phi(100) - \phi(40)}{\phi(100) - \phi(40)} = 1.$$

从而得到

$$\log(\mu(S_l)) = \beta_0 + \beta_1\phi(S_l),$$

其中  $\beta_0 = \log(\mu) + \beta\phi(40)$ ,  $\beta_1 = \beta(\phi(100) - \phi(40))$ .

表 5.8: 基于应力松弛数据, 不同退化模型的 AIC 值

	维纳模型	伽马模型	IG 模型	ED 模型
AIC 值	2320.309	470.311	452.770	401.703

参数的 MLE 及其 95% 置信区间列于表 5.7 中. 结果显示,  $d$  的 MLE 为 1.471, 95% 置信区间为 (1.054, 1.487), 表明复合泊松过程更适合拟合这些数据. 表 5.8 的 AIC 值同样支持这一结论. 图 5.5(a) 显示了各应力水平下估计的平均退化路径 ( $\hat{\mu}t^{\hat{\alpha}}$ ) 及其 95% 置信带, 置信带能够很好地覆盖样本平均路径. 图 5.5(b) 显示了应力松弛数据的可靠性及其 95% 自助置信区间, 进一步验证了复合泊松过程的适用性.

### 5.1.6 带随机效应的指数扩散过程

在实际应用中, 由于不可观测的因素 (如原材料初始缺陷的变异和使用模式的差异 (Ye, Hong, 等, 2013)), 同一群体内产品的退化路径可能存在显著差异. 为描述这种现象, 常采用随机效应模型. 在 ED 过程中, 可以通过漂移参数  $\mu$  或扩散参数  $\lambda$  的变异性引入随机效应, 以下是三种带随机效应的扩展模型.

(1) **随机漂移模型**: 当产品之间退化速率存在异质性时, 可通过漂移参数  $\mu$  的随机性进行建模. 假设  $\mu$  服从截断正态分布  $TN(\eta, \tau^{-2})$ , 其 PDF 为

$$f_{TN}(\mu) = \frac{\tau\phi(\tau(\mu - \eta))}{1 - \Phi(-\tau\eta)}, \quad \mu > 0, \tau > 0,$$

其中  $\phi(\cdot)$  和  $\Phi(\cdot)$  分别是标准正态分布的 PDF 和 CDF. 基于退化数据  $y = (\Delta y_1, \dots, \Delta y_n)'$ , 参数  $\Theta_1 = (\lambda, d, \alpha, \eta, \tau)'$  的似然函数为

$$\begin{aligned} L(\Theta_1 \mid y) &= \prod_{i=1}^n \int_0^\infty f_{ED}(\Delta y_i \mid \mu_i) f_{TN}(\mu_i) d\mu_i \\ &= \prod_{i=1}^n \int_0^\infty \prod_{j=1}^{m_i} f_{ED}(\Delta y_{ij} \mid \mu_i) f_{TN}(\mu_i) d\mu_i \\ &= \left( \frac{\tau}{1 - \Phi(-\tau\eta)} \right)^n \prod_{i=1}^n \prod_{j=1}^{m_i} \{2\pi\lambda^{-1}\Delta\Lambda(t_{ij}; \alpha)V(\Delta y_{ij}/\Delta\Lambda(t_{ij}; \alpha))\}^{-1/2} \\ &\quad \times \exp \left\{ -\frac{\lambda \sum_{i=1}^n \sum_{j=1}^{m_i} (\Delta y_{ij})^{2-d} (\Delta\Lambda(t_{ij}; \alpha))^{d-1}}{(1-d)(2-d)} \right\} \\ &\quad \times \prod_{i=1}^n \int_0^\infty \phi(\tau(\mu_i - \eta)) \end{aligned}$$

$$\times \exp \left\{ -\lambda \sum_{j=1}^{m_i} \left( \frac{\Delta \Lambda(t_{ij}; \alpha) \mu^{2-d}}{2-d} - \frac{\Delta y_{ij} \mu^{1-d}}{(1-d)} \right) \right\} d\mu_i, \quad (5.16)$$

(2) **随机扩散模型:** 当所有单元具有相同的平均退化路径, 但退化过程的方差存在差异时, 可假设扩散参数  $\lambda$  服从伽马分布  $Ga(\alpha_\lambda, \beta_\lambda)$ , 其 PDF 为

$$f_{Ga}(\lambda) = \frac{\beta_\lambda^{\alpha_\lambda}}{\Gamma(\alpha_\lambda)} \lambda^{\alpha_\lambda-1} \exp\{-\beta_\lambda \lambda\}.$$

给定退化数据  $y$ , 参数  $\Theta_2 = (\lambda, d, \alpha, \alpha_\lambda, \beta_\lambda)'$  的似然函数为

$$\begin{aligned} L(\Theta_2 \mid y) &= \prod_{i=1}^n \int_0^\infty f_{ED}(\Delta y_i \mid \lambda_i) f_{Ga}(\lambda_i) d\lambda_i \\ &= \prod_{i=1}^n \int_0^\infty \prod_{j=1}^{m_i} f_{ED}(\Delta y_{ij} \mid \lambda_i) f_{Ga}(\lambda_i) d\lambda_i \\ &= \frac{\Gamma(\tilde{\alpha}) \beta_\lambda^{\alpha_\lambda}}{\Gamma(\alpha_\lambda) \tilde{\beta}^{\tilde{\alpha}}} \prod_{i=1}^n \prod_{j=1}^{m_i} (2\pi \Delta \Lambda(t_{ij}; \alpha) V(\Delta y_{ij} / \Delta \Lambda(t_{ij}; \alpha)))^{-1/2}, \end{aligned} \quad (5.17)$$

其中

$$\begin{aligned} \tilde{\alpha} &= \alpha_\lambda + \sum_{i=1}^n m_i / 2, \\ \tilde{\beta} &= \beta_\lambda + \sum_{i=1}^n \sum_{j=1}^{m_i} \left( \frac{\Delta y_{ij}^{2-d} (\Delta \Lambda(t_{ij}; \alpha))^{d-1}}{(1-d)(2-d)} - \frac{\Delta y_{ij} \mu^{1-d}}{1-d} + \frac{\Delta \Lambda(t_{ij}; \alpha) \mu^{2-d}}{2-d} \right). \end{aligned}$$

(3) **随机漂移-扩散模型:** 当  $\mu$  和  $\lambda$  同时具有随机性时, 可采用截断正态-伽马 (Truncated normal-gamma, TNG) 混合分布来描述两者的相关性, 即

$$\mu \mid \lambda \sim TN(\eta, \tau^{-2}/\lambda), \quad \lambda \sim Ga(\alpha_\lambda, \beta_\lambda).$$

给定退化数据  $y$ , 参数  $\Theta_3 = (\lambda, d, \alpha, \eta, \tau, \alpha_\lambda, \beta_\lambda)'$  的似然函数为

$$\begin{aligned}
 L(\Theta_3 | y) &= \prod_{i=1}^n \int_0^\infty \int_0^\infty \prod_{j=1}^{m_i} f_{ED}(\Delta y_{ij} | \mu_i, \lambda_i) f_{TN}(\mu_i | \lambda_i) f_{Ga}(\lambda_i) d\mu_i d\lambda_i \\
 &= \prod_{i=1}^n \int_0^\infty \int_0^\infty \prod_{j=1}^{m_i} \left\{ 2\pi \lambda_i^{-1} \Delta \Lambda(t_{ij}; \alpha) V(\Delta y_{ij} / \Delta \Lambda(t_{ij}; \alpha)) \right\}^{-1/2} \\
 &\quad \times \exp \left\{ -\lambda_i \Delta \Lambda(t_{ij}; \alpha) \left( \frac{(\Delta y_{ij} / \Delta \Lambda(t_{ij}; \alpha))^{2-d}}{(1-d)(2-d)} \right. \right. \\
 &\quad \left. \left. - \frac{\Delta y_{ij} \mu_i^{1-d}}{\Delta \Lambda(t_{ij}; \alpha)(1-d)} + \frac{\mu_i^{2-d}}{2-d} \right) \right\} \\
 &\quad \times \frac{\tau \sqrt{\lambda_i} \phi(\tau \sqrt{\lambda_i} (\mu_i - \eta))}{1 - \Phi(-\tau \sqrt{\lambda_i} \eta)} \times \frac{\beta_{\lambda_i}^{\alpha_{\lambda_i}}}{\Gamma(\alpha_{\lambda_i})} \lambda_i^{\alpha_{\lambda_i}-1} \exp\{-\beta_{\lambda_i} \lambda_i\} d\mu_i d\lambda_i.
 \end{aligned} \tag{5.18}$$

通过优化上述模型的对数似然函数, 可以获得参数的最大似然估计. 然而, 由于似然函数的复杂性, 直接最大化可能导致估计偏差. EM 算法作为一种迭代优化方法, 能有效处理潜在变量并提高参数估计精度, 具体讨论可参考 (Chen 等, 2022).

## 5.2 动态使用环境下系统可靠性评估

系统可靠性的评估涉及多方面因素, 包括系统结构、工作条件和部件间的相关性, 较为复杂. 传统研究通常假设部件寿命相互独立, 且工作条件是静态或确定的 (Peng W 等, 2018; Xu 等, 2014; Yalaoui 等, 2005). 然而, 这种假设可能导致系统可靠性评估出现偏差 (Kotz 等, 2003; Song 等, 2014). 为刻画部件间相关性, 文献中常用以下三种方法: (1) 多元分布: 如多元对数正态分布 (Fan 等, 2015) 和 Marshall-Olkin Weibull 分布 (Kundu 等, 2013), 但验证部件寿命是否服从特定分布较为复杂. (2) Copula 函数: 通过引入相关性结构来构建部件相关性 (Peng 等, 2016; Wang, Guo, 等, 2014), 然而, Copula 方法的物理解释较为有限, 且选择合适的函数形式具有挑战性. (3) 脆弱因子 (Frailty) 模型: 通过引入脆弱因子来描述系统异质性及部件间的相关性 (Hougaard, 1995; Liu, 2012). 然而, 脆弱因子只是一个观测不到的随机比例因子, 它能调整故障率函数, 却不

能反映环境的时变效应.

实际上, 动态环境是影响系统寿命不可忽视的重要因素. 例如, 锂电池的寿命取决于充电和放电次数、使用习惯以及使用温度等; 车辆刹车片和轮胎的老化或磨损过程与道路条件、驾驶习惯和使用率相关. 这些动态因素会导致部件的等效工作时间  $A(t)$  呈现随机性, 进而反映出动态环境对老化过程的累积影响. 为描述这种累积效应, Nelson (1980) 提出了累积损伤模型, 随后 Hong 等 (2019) 基于此假设  $A(t)$  是一个随机时间尺度, 并提出以下模型

$$F_i(t) = 1 - \exp\{-\eta_i A(t)\}, \quad (5.19)$$

其中  $\eta_i > 0$  为基故障率,  $A(t)$  则描述运行环境应力对时间尺度的影响. 从物理角度来看,  $A(t)$  可以解释为部件老化过程的度量.  $F_i(t)$  包含一些常见的寿命分布作为特例, 例如, 威布尔分布 ( $A(t) = t^\alpha$ )、Gompertz 分布 ( $A(t) = \exp(\alpha t) - 1$ )、Lomax 分布 ( $A(t) = \log(1 + t/\alpha)$ ). 在恒定环境下,  $A(t)$  通常被设定为确定性的时间尺度. 但实际情况中, 系统部件往往处于动态环境中, 其工作条件可能随着时间变化而显现出随机性. 为更准确地描述这种复杂性, 将时间尺度扩展为随机过程  $Y(t), t \geq 0$  是更合理的选择. 接下来, 将基于随机过程  $Y(t)$  建立动态环境中的可靠性模型.

在给定  $Y(t)$  的条件下, 部件寿命  $T_i$  的 CDF 为

$$F_i(t | Y(t)) = 1 - \exp\{-\eta_i Y(t)\}. \quad (5.20)$$

Hong 等 (2019) 通过复合泊松过程、伽马过程和 IG 过程对  $Y(t)$  进行建模, 以捕捉动态环境的影响. 由于复合泊松过程的复杂性, Hong 等 (2019) 在实际分析中采用了非齐次泊松过程以简化计算. 本文在更一般的框架下, 假设  $Y(t)$  服从 ED 过程, 其优点包括: 1. 可将 Hong 等 (2019) 的模型作为特例, 但仍具有其相关性和系统可靠性的结果. 2. 提供统一的数据分析框架, 有助于选择最优模型. 本节安排如下: 第 5.2.1 节介绍所提出的模型. 第 5.2.2 节研究部件寿命相关性, 推导联合 PDF. 第 5.2.3 节分析系统可靠性, 并讨论动态环境和相关性对系统可靠性的影响. 第 5.2.4 节探讨参数冗余问题. 第 5.2.5 节进行模拟研究, 验证模型性能.

表 5.9: 当  $\mu = 1$  时  $\theta$  的值以及不同  $d$  下的  $\kappa(\theta)$ .

$d$	随机过程	$\theta$	$\kappa(\theta)$
0	维纳过程	1	$\theta^2/2$
1	泊松过程	0	$\exp(\theta)$
(1, 2)	复合泊松过程	$\frac{1}{1-d}$	$\frac{[(1-d)\theta]^{(2-d)/(1-d)}}{2-d}$
2	伽马过程	-1	$-\log(-\theta)$
3	逆高斯过程	$-\frac{1}{2}$	$-\sqrt{-2\theta}$

### 5.2.1 模型构建

基于第 5.1 节中对 ED 过程的定义, 本节进一步构建一个能够体现动态环境影响的统一模型, 具体形式如下

$$\begin{cases} T_i \mid Y(t) \sim F_i(t \mid Y(t)), \\ Y(t) \sim \mathcal{ED}(\mu\Lambda(t; \alpha), \lambda), \end{cases} \quad (5.21)$$

其中  $T_i$  表示部件寿命,  $F_i(t \mid Y(t))$  为其条件分布;  $Y(t)$  为反映时间尺度的随机过程, 服从 ED 过程  $\mathcal{ED}(\mu\Lambda(t; \alpha), \lambda)$ . 由于时间尺度的性质, 随机过程  $Y(t), t \geq 0$  应满足非递减性, 并且初始值  $Z(0) = 0$ . 然而, 当参数  $d \leq 0$  时, ED 过程可能失去单调性, 因此本研究选择方差函数  $V(\mu) = \mu^d$ , 并限定  $d \geq 1$ . 同时, 为避免模型 (5.21) 中的参数出现不可识别性, 进一步假设漂移参数  $\mu = 1$ . 在此条件下, 不同  $d$  值下的参数  $\theta$  以及对应的累积量函数  $\kappa(\theta)$  的解析表达式列于表 5.9 中. 这些参数和函数形式对于理解 ED 过程的统计性质有重要参考作用.

### 5.2.2 部件相关性

首先考虑一个双部件系统, 并令  $T_1$  和  $T_2$  分别表示部件 1 和 2 的寿命. 设  $t_1$  和  $t_2$  分别为  $T_1$  和  $T_2$  的观测值. 按升序排列  $t_i$  时, 用  $r_i$  表示第  $i$  个部件寿命值  $t_i$  的秩 (排序). 这一秩的定义同样适用于后续  $K$  部件系统的情况. 基于寿命的秩重新排列基故障率  $\eta_i$ , 令  $\eta_{(r_i)} = \eta_i$ . 在此基础上, 可以得出定理 5.1.

**定理 5.1.** 若  $T_i$  的条件 CDF 定义为式 (5.20) (其中  $i = 1, 2$ ), 且  $Y(t) \sim$

$\mathcal{ED}(\Lambda(t; \alpha), \lambda)$ ), 则  $T_1$  和  $T_2$  的联合生存函数为

$$S(t_1, t_2) = \exp \left\{ \lambda \Lambda(t_{r_1}; \alpha) \left[ \kappa \left( \theta - \frac{\eta_1 + \eta_2}{\lambda} \right) - \kappa \left( \theta - \frac{\eta_{(2)}}{\lambda} \right) \right] + \lambda \Lambda(t_{r_2}; \alpha) \left[ \kappa \left( \theta - \frac{\eta_{(2)}}{\lambda} \right) - \kappa(\theta) \right] \right\}. \quad (5.22)$$

**证明:** 在 ED 过程假设下,  $Y(t_{r_1})$  和  $\Delta Y(t_{r_2}) = Y(t_{r_2}) - Y(t_{r_1})$  相互独立, 其中  $Y(t_{r_1}) \sim ED(\Lambda(t_{r_1}; \alpha), \lambda)$ ,  $\Delta Y(t_{r_2}) \sim ED(\Delta \Lambda(t_{r_2}; \alpha), \lambda)$ . 在给定  $Y(t_{r_1})$  和  $\Delta Y(t_{r_2})$  的条件下,

$$P(T_1 > t_1, T_2 > t_2 | Y(t_{r_1}), \Delta Y(t_{r_2})) = \exp \left\{ -(\eta_1 + \eta_2)Y(t_{r_1}) \right\} \exp \left\{ -\eta_{(2)}\Delta Y(t_{r_2}) \right\}.$$

接着, 以  $Z(t_{r_1})$  和  $\Delta Z(t_{r_2})$  为变量, 对条件概率  $P(T_1 > t_1, T_2 > t_2 | Z(t_{r_1}), \Delta Z(t_{r_2}))$  求期望, 得出

$$\begin{aligned} S(t_1, t_2) &= P(T_1 > t_1, T_2 > t_2) \\ &= \mathbb{E}_{Y(t_{r_1})} \left[ \exp \left\{ -(\eta_1 + \eta_2)Y(t_{r_1}) \right\} \right] \mathbb{E}_{\Delta Y(t_{r_2})} \left[ \exp \left\{ -\eta_{(2)}\Delta Y(t_{r_2}) \right\} \right] \\ &= \exp \left\{ \lambda \Lambda(t_{r_1}; \alpha) [\kappa(\theta - (\eta_1 + \eta_2)/\lambda) - \kappa(\theta)] \right\} \\ &\quad \times \exp \left\{ \lambda \Delta \Lambda(t_{r_2}; \alpha) [\kappa(\theta - \eta_{(2)}/\lambda) - \kappa(\theta)] \right\} \\ &= \exp \left\{ \lambda \Lambda(t_{r_1}; \alpha) \left[ \kappa(\theta - (\eta_1 + \eta_2)/\lambda) - \kappa(\theta - \eta_{(2)}/\lambda) \right] \right. \\ &\quad \left. + \lambda \Lambda(t_{r_2}; \alpha) \left[ \kappa(\theta - \eta_{(2)}/\lambda) - \kappa(\theta) \right] \right\}. \end{aligned}$$

□

**注 1:** 定理 5.1 中的  $S(t_1, t_2)$  是一般化形式, Hong 等 (2019) 的结果可被视为其特例. 例如, 当  $d = 1$  时, 根据表 5.9, 可以得到联合生存函数为

$$S(t_1, t_2) = \exp \left\{ \lambda \Lambda(t_{r_1}; \alpha) \left[ \exp \left( -\frac{\eta_1 + \eta_2}{\lambda} \right) - \exp \left( -\frac{\eta_{(2)}}{\lambda} \right) \right] + \lambda \Lambda(t_{r_2}; \alpha) \left[ \exp \left( -\frac{\eta_{(2)}}{\lambda} \right) - 1 \right] \right\}. \quad (5.23)$$

当  $\lambda = 1$  时, 即为 Hong 等 (2019) 中的复合泊松过程. 若  $1 < d < 2$ , 根据定理

5.1 和表 5.9, 联合生存函数为

$$S(t_1, t_2) = \exp \left\{ \frac{\lambda \Lambda(t_{r_1}; \alpha)}{2-d} \left[ \left( 1 + \frac{(\eta_1 + \eta_2)(d-1)}{\lambda} \right)^{\frac{2-d}{1-d}} - \left( 1 + \frac{\eta_{(2)}(d-1)}{\lambda} \right)^{\frac{2-d}{1-d}} \right] \right\} \\ \times \exp \left\{ \frac{\lambda \Lambda(t_{r_2}; \alpha)}{2-d} \left[ \left( 1 + \frac{\eta_{(2)}(d-1)}{\lambda} \right)^{\frac{2-d}{1-d}} - 1 \right] \right\}. \quad (5.24)$$

类似地,  $d = 2$  和  $d = 3$  分别对应伽马过程和 IG 过程的结果.

ED 过程有效地捕捉了动态环境对两部件累积影响的共同作用, 从而自然地刻画了两部件寿命的相关性. 为验证模型的适用性, 借鉴 Hong 等 (2019) 的研究, 可分析  $T_1$  和  $T_2$  之间的相关性. 当对于所有  $t_1$  和  $t_2$ , 满足

$$P(T_1 > t_1, T_2 > t_2) \geq P(T_1 > t_1)P(T_2 > t_2),$$

则称两个随机变量  $T_1$  和  $T_2$  具有正象限相关 (positive quadrant dependent).

**定理 5.2.** 在定理 5.1 的条件下, 部件寿命  $T_1$  和  $T_2$  具有正象限相关.

**证明:** 以下仅针对  $d \neq 1, 2$  的情况进行证明, 因为 Hong 等 (2019) 已证明在  $d = 1$  和  $d = 2$  时, 结论同样成立. 设  $T_i$  的生存函数为

$$P(T_i > t_i) = \mathbb{E}_{Y(t_i)} [\exp \{-\eta_i Y(t_i)\}] \\ = \exp \left\{ \lambda \Lambda(t_i; \alpha) \left[ \kappa \left( \theta - \frac{\eta_i}{\lambda} \right) - \kappa(\theta) \right] \right\}, \quad i = 1, 2. \quad (5.25)$$

由定理 5.1 可得

$$Q(t_1, t_2) = \frac{P(T_1 > t_1, T_2 > t_2)}{P(T_1 > t_1)P(T_2 > t_2)} \\ = \exp \left\{ \lambda \Lambda(t_{r_1}; \alpha) \left[ \kappa \left( \theta - \frac{\eta_1 + \eta_2}{\lambda} \right) - \kappa \left( \theta - \frac{\eta_1}{\lambda} \right) - \kappa \left( \theta - \frac{\eta_2}{\lambda} \right) + \kappa(\theta) \right] \right\} \\ = \exp \left\{ \frac{\lambda \Lambda(t_{r_1}; \alpha) g(d)}{2-d} \right\},$$

其中

$$g(d) = \left( 1 + \frac{(\eta_1 + \eta_2)(d-1)}{\lambda} \right)^{\frac{2-d}{1-d}} - \left( 1 + \frac{\eta_1(d-1)}{\lambda} \right)^{\frac{2-d}{1-d}} - \left( 1 + \frac{\eta_2(d-1)}{\lambda} \right)^{\frac{2-d}{1-d}} + 1.$$

定义  $q(x, y) = y^x, y \geq 1$ . 对于任何  $x < 0$ ,  $q(x, y)$  是关于  $y$  的凸函数. 因此, 对于任意  $1 \leq y_1 < y_2 \leq y_3 < y_4$ , 根据凸函数性质有:

$$\frac{y_2^x - y_1^x}{y_2 - y_1} \leq \frac{y_4^x - y_3^x}{y_4 - y_3}. \quad (5.26)$$

当  $1 < d < 2$  时,  $(2-d)/(1-d) < 0$ . 假设  $\eta_1 \leq \eta_2$ . 根据上述性质得到

$$\left(1 + \frac{\eta_1(d-1)}{\lambda}\right)^{\frac{2-d}{1-d}} - 1 \leq \left(1 + \frac{(\eta_1 + \eta_2)(d-1)}{\lambda}\right)^{\frac{2-d}{1-d}} - \left(1 + \frac{\eta_2(d-1)}{\lambda}\right)^{\frac{2-d}{1-d}},$$

即  $g(d) \geq 0$ . 因此  $Q(t_1, t_2) \geq 1$ .

类似地, 对于  $d > 2$  的情况, 由于  $0 < (2-d)/(1-d) < 1$ , 且对于任意  $0 < x < 1$ ,  $q(x, y)$  是关于  $y$  的凹函数, 因此可得  $g(d) \leq 0$ . 由此推得  $Q(t_1, t_2) \geq 1$ .  $\square$

**注 2:** 定理 5.2 表明, 随机时间尺度  $\{Y(t), t \geq 0\}$  引入了部件寿命之间的相关性. 当  $\{Y(t), t \geq 0\}$  退化为确定性函数  $\Lambda(t; \alpha)$  时 (即  $\lambda \rightarrow \infty$ ) 时, 通过洛必达法则, 可以证明式 (??) 中  $\lambda \cdot g(d) \rightarrow 0$ . 那么  $Q(t_1, t_2) \rightarrow 1$ , 这意味着  $T_1$  和  $T_2$  独立.

**注 3:** 定理 5.2 的结果也适用于某些非单调随机过程. 对于维纳过程 ( $d = 0$ ),  $\kappa(\theta) = \theta^2/2$ , 可推导出  $Q(t_1, t_2) = \exp\{\Lambda(t_{r_1}; \alpha)\eta_1\eta_2/\lambda\} \geq 1$ .

对于  $K$  部件系统, 其联合生存函数  $(T_1, \dots, T_K)$  的形式如定理 5.3 所示.

**定理 5.3.** 若  $T_i$  的条件 CDF 定义为式 (5.20) ( $i = 1, \dots, K$ ), 且  $Y(t) \sim \mathcal{ED}(\Lambda(t; \alpha), \lambda)$ , 那么  $(T_1, \dots, T_K)$  的联合生存函数为

$$S(t_1, \dots, t_K) = \prod_{i=1}^K \exp\left\{-\lambda\Lambda(t_{(r_i)}; \alpha)\bigtriangledown\kappa_i\right\},$$

其中  $\bigtriangledown\kappa_i = \kappa\left(\theta - \sum_{k=i+1}^K \eta_{(r_k)}/\lambda\right) - \kappa\left(\theta - \sum_{k=i}^K \eta_{(r_k)}/\lambda\right)$ , 且  $\sum_{k=K+1}^K \eta_{(r_k)} = 0$ .

**证明:** 设  $\Delta Y(t_{r_i}) = Y(t_{r_i}) - Y(t_{r_{i-1}})$ , 其中  $i = 1, \dots, K$  且  $t_0 = 0$ . 增量  $\Delta Y(t_{r_1}), \dots, \Delta Y(t_{r_K})$  相互独立, 且  $\Delta Y(t_{r_i}) \sim ED(\Delta\Lambda(t_{(r_i)}; \alpha), \lambda)$ , 其中  $i = 1, \dots, K$ , 且  $\Delta\Lambda(t_{(r_i)}; \alpha) = \Lambda(t_{r_i}; \alpha) - \Lambda(t_{r_{i-1}}; \alpha)$ . 在给定  $\Delta Y(t_{r_i}), i = 1, \dots, K$

的条件下,  $(T_1, \dots, T_K)$  的条件联合生存函数为

$$S(t_1, \dots, t_K \mid \Delta Y(t_{r_i}), i = 1, \dots, K) = \prod_{i=1}^K \exp \left\{ -\Delta Y(t_{r_i}) \sum_{k=i}^K \eta_{(r_k)} \right\}.$$

对  $S(t_1, \dots, t_K \mid \Delta Y(t_{r_i}), i = 1, \dots, K)$  关于  $\Delta Y(t_{r_i})$  取期望, 可以得到

$$\begin{aligned} S(t_1, \dots, t_K) &= \prod_{i=1}^K \mathbb{E}_{\Delta Y(t_{r_i})} \left[ \exp \left\{ -\Delta Y(t_{r_i}) \sum_{k=i}^K \eta_{(r_k)} \right\} \right] \\ &= \prod_{i=1}^K \exp \left\{ \lambda \Delta \Lambda(t_{(r_i)}; \alpha) \left[ \kappa \left( \theta - \sum_{k=i}^K \eta_{(r_k)} / \lambda \right) - \kappa(\theta) \right] \right\} \\ &= \prod_{i=1}^K \exp \left\{ -\lambda \Lambda(t_{(r_i)}; \alpha) \bigtriangledown \kappa_i \right\}, \end{aligned} \quad (5.27)$$

其中  $\bigtriangledown \kappa_i = \kappa \left( \theta - \sum_{k=i+1}^K \eta_{(r_k)} / \lambda \right) - \kappa \left( \theta - \sum_{k=i}^K \eta_{(r_k)} / \lambda \right)$ , 且  $\sum_{k=K+1}^K \eta_{(r_k)} = 0$ .  $\square$

**注 4:** 当  $d = 1, 2$  和  $3$  时, 定理 5.3 包括了 Hong 等 (2019) 的结果, 分别对应非齐次泊松过程、伽马过程和逆高斯过程. 此外, 该定理还涵盖了一些非递减随机过程的结果, 例如复合泊松过程 ( $1 < d < 2$ ) 和具有正增量的随机过程 ( $2 < d < 3$  及  $d > 3$ ).

**注 5:** 根据定理 5.2, 任意两个部件的寿命  $T_i$  和  $T_j$  之间具有正象限相关. 当  $\lambda \rightarrow \infty$  时, 即随机时间尺度  $\{Y(t), t \geq 0\}$  退化为确定性函数  $\Lambda(t; \alpha)$ , 由洛必达法则可证明

$$\lambda \left[ \kappa \left( \theta - \sum_{k=i}^K \frac{\eta_{(r_k)}}{\lambda} \right) - \kappa \left( \theta - \sum_{k=i+1}^K \frac{\eta_{(r_k)}}{\lambda} \right) \right] \rightarrow \eta_{r_k}.$$

因此, 联合生存函数退化为:  $S(t_1, \dots, t_K) = \prod_{i=1}^K \exp \{-\eta_i \Lambda(t_i; \alpha)\}$ , 这表明部件寿命  $T_1, \dots, T_K$  之间相互独立.

### 5.2.3 系统可靠度

对于  $K$  部件串联系统, 其寿命为  $T = \min\{T_1, \dots, T_K\}$ . 根据定理 5.3, 系统可靠度函数可表示为

$$R_S(t) = P(T_1 > t, \dots, T_K > t) = \exp \left\{ -\lambda \Lambda(t; \alpha) \left[ \kappa(\theta) - \kappa \left( \theta - \sum_{i=1}^K \eta_i / \lambda \right) \right] \right\}.$$

若假设  $T_1, \dots, T_K$  之间相互独立, 则系统可靠度函数为

$$R_{IS}(t) = \exp \left\{ -\lambda \Lambda(t) \sum_{i=1}^K [\kappa(\theta) - \kappa(\theta - \eta_i / \lambda)] \right\}.$$

当随机时间尺度  $Y(t)$  退化为确定性函数  $\Lambda(t; \alpha)$  (即  $\lambda \rightarrow \infty$ ) 时, 系统可靠度函数可以简化为

$$R_{DS}(t) = \exp \left\{ -\Lambda(t; \alpha) \sum_{i=1}^K \eta_i \right\}.$$

进一步, 定理 5.4 给出了  $R_S(t)$ ,  $R_{IS}(t)$  与  $R_{DS}(t)$  之间的关系.

**定理 5.4.**  $R_{DS}(t) < R_{IS}(t) < R_S(t)$ .

**证明:** 定义  $h_i(\lambda) = \lambda [\kappa(\theta) - \kappa(\theta - \eta_i / \lambda)]$ . 根据拉格朗日中值定理, 有  $h_i(\lambda) = \eta_i \kappa'(\theta_i^*)$ , 其中  $\theta - \eta_i / \lambda < \theta_i^* < \theta$ . 从表 5.9 中可知: 当  $d = 1$  时, 有  $\kappa'(\theta_i^*) = \exp(\theta_i^*)$ , 且  $-\eta_i / \lambda < \theta_i^* < 0$ . 当  $d = 2$  时, 有  $\kappa'(\theta_i^*) = -1/\theta_i^*$ , 且  $-1 - \eta_i / \lambda < \theta_i^* < -1$ . 因此, 无论  $d = 1$  或  $d = 2$ , 均可得  $\kappa'(\theta_i^*) < 1$ . 对于  $d \neq 1, 2$  的情况, 有  $\kappa'(\theta_i^*) = [(1-d)\theta_i^*]^{\frac{1}{1-d}}$ , 且  $\frac{1}{1-d} - \frac{\eta_i}{\lambda} < \theta_i^* < \frac{1}{1-d}$ , 由此, 同样可得  $\kappa'(\theta_i^*) < 1$ . 于是  $\sum_{i=1}^K h_i(\lambda) < \sum_{i=1}^K \eta_i$ , 从而证明了第一个不等式.

接下来证明第二个不等式. 首先, 需证明  $\kappa(\theta)$  是  $\theta$  的严格凸函数, 这一结论显然成立. 由此有

$$\kappa(\theta) - \kappa \left( \theta - \frac{\sum_{i=1}^K \eta_i}{\lambda} \right) = \sum_{i=1}^K \left[ \kappa \left( \theta - \frac{\sum_{k=1}^{i-1} \eta_k}{\lambda} \right) - \kappa \left( \theta - \frac{\sum_{k=1}^i \eta_k}{\lambda} \right) \right],$$

其中  $\sum_{k=1}^0 \eta_k / \lambda = 0$ . 根据凸函数的性质, 可得

$$\kappa \left( \theta - \frac{\sum_{k=1}^{i-1} \eta_k}{\lambda} \right) - \kappa \left( \theta - \frac{\sum_{k=1}^i \eta_k}{\lambda} \right) < \kappa(\theta) - \kappa \left( \theta - \frac{\eta_i}{\lambda} \right), \quad i = 1, \dots, K.$$

因此, 第二个不等式成立.  $\square$

定理 5.4 揭示了两个重要结论: (1) 如果未考虑动态环境的影响, 则会低估系统可靠性 ( $R_{DS}(t) < R_{IS}(t)$ ). (2) 忽略部件寿命之间的正相关性, 也会导致系统可靠性的低估 ( $R_{IS}(t) < R_S(t)$ ).

对于图 5.6 所示的串并联系统, 假设系统由  $K$  个子系统串联组成, 第  $i$  个子系统包含  $m_i$  个并联的部件. 令  $\omega_{ij}$  表示第  $i$  个子系统中第  $j$  个部件在时间  $t$  的状态, 即  $\omega_{ij} = 1$  表示该部件正常工作,  $\omega_{ij} = 0$  表示失效, 其中  $i = 1, \dots, K, j = 1, \dots, m_i$ . 定义  $\omega_i = (\omega_{i1}, \dots, \omega_{im_i})'$  为第  $i$  个子系统的状态.  $\omega = (\omega_1, \dots, \omega_K)'$  和  $\eta = (\eta_1, \dots, \eta_K)'$  为系统的整体状态,  $\Omega$  为所有正常工作状态的集合, 则系统可靠度函数为

$$\begin{aligned} R_{SP}(t) &= \mathbb{E}_{Y(t)} \left[ \sum_{\omega \in \Omega} \left\{ \exp\{-\omega' \eta Y(t)\} \prod_{i=1}^K (-1)^{|\omega_i|+1} \right\} \right] \\ &= \sum_{\omega \in \Omega} \exp \left\{ -\lambda \Lambda(t; \alpha) [\kappa(\theta) - \kappa(\theta - \omega' \eta / \lambda)] \right\} \prod_{i=1}^K (-1)^{|\omega_i|+1}. \end{aligned} \quad (5.28)$$

当  $\{Y(t), t \geq 0\}$  退化为确定性函数  $\Lambda(t; \alpha)$  时, 即  $\lambda \rightarrow \infty$ , 由洛必达法则可得

$$\lambda [\kappa(\theta) - \kappa(\theta - \omega' \eta / \lambda)] \rightarrow \omega' \eta.$$

此时, 系统可靠度函数变为

$$R_{ISP} = \sum_{\omega \in \Omega} \exp \left\{ -\omega' \eta \Lambda(t; \alpha) \right\} \prod_{i=1}^K (-1)^{|\omega_i|+1},$$

这表明, 当随机性消失时, 系统的各部件和子系统的性能将表现为相互独立.

对于  $K$  部件并联系统, 系统可靠度函数可以表示为

$$R_P(t) = \sum_{\omega \in \Omega} (-1)^{|\omega|+1} \exp \left\{ -\lambda \Lambda(t; \alpha) [\kappa(\theta) - \kappa(\theta - \omega' \eta / \lambda)] \right\}.$$

如 Kotz 等 (2003) 所示, 忽略并联系统中部件寿命的正相关性会高估系统可靠性. 然而, 对于串并联系统, 情况更加复杂, 值得进一步研究.

### 5.2.4 参数冗余性

当有观测数据时, 所提模型可用于评估系统可靠性. 参数估计的常用方法是极大似然法, 即通过最大化对数似然函数获得估计值. 为此, 我们需要推导  $(T_1, \dots, T_K)$  的联合 PDF. 根据定理 5.3,  $(T_1, \dots, T_K)$  的联合 PDF 可通过对联合生存函数  $S(t_1, \dots, t_K)$  关于  $t_i$  求偏导数得到

$$\begin{aligned} f(t_1, \dots, t_K) &= \frac{(-1)^K \partial^K S(t_1, \dots, t_K)}{\partial t_1 \dots \partial t_K} \\ &= \prod_{i=1}^K \left\{ \lambda \nabla \kappa_i \exp \left[ -\lambda \Lambda(t_{(r_i)}; \alpha) \nabla \kappa_i \right] \frac{\partial \Lambda(t_{(r_i)}; \alpha)}{\partial t_{(r_i)}} \right\}. \end{aligned} \quad (5.29)$$

在实际应用中, 部分部件寿命可能被观测到, 而其余部分则会因试验时间的限制而无法完全观测. 设  $T_{r_1}, \dots, T_{r_m}$  的观测值分别为  $t_{r_1}, \dots, t_{r_m}$ , 其余部件寿命  $T_{r_{m+1}}, \dots, T_{r_R}$  在截尾时间  $\tau$  处被截尾 ( $\tau \geq t_{r_m}$ ). 此时,  $P(T_{r_1} = t_{r_1}, \dots, T_{r_m} = t_{r_m}, T_{r_{m+1}} > \tau, \dots, T_{r_R} > \tau)$  可通过将式 (5.29) 对  $T_{r_{m+1}}, \dots, T_{r_R}$  在  $(\tau, \infty)$  上积分得到

$$\begin{aligned} P(T_{r_1} = t_{r_1}, \dots, T_{r_m} = t_{r_m}, T_{r_{m+1}} > \tau, \dots, T_{r_R} > \tau) \\ &= \prod_{i=1}^m \left\{ \lambda \nabla \kappa_i \exp \left[ -\lambda \Lambda(t_{(r_i)}; \alpha) \nabla \kappa_i \right] \frac{\partial \Lambda(t_{(r_i)}; \alpha)}{\partial t_{(r_i)}} \right\} \\ &\quad \times \exp \left\{ \lambda \Lambda(\tau; \alpha) \left[ \kappa \left( \theta - \sum_{k=m+1}^K \eta_{(r_k)} / \lambda \right) - \kappa(\theta) \right] \right\}. \end{aligned} \quad (5.30)$$

上述条件要求  $t_i$  彼此不同. 当某些  $t_i$  相等时, 联合 PDF 会更加复杂. 以  $K = 2$  为例, 此时  $T_1$  和  $T_2$  的联合生存函数为

$$S(t_1, t_2) = \exp \left\{ -\lambda \Lambda(t_{r_1}; \alpha) \left[ \kappa(\theta - \eta_{(2)}/\lambda) - \kappa(\theta - (\eta_1 + \eta_2)/\lambda) \right] - \lambda \Lambda(t_{r_2}; \alpha) \left[ \kappa(\theta) - \kappa(\theta - \eta_{(2)}/\lambda) \right] \right\}.$$

与 Marshall-Olkin 二元 Weibull 分布类似 (Kundu 等, 2013),  $S(t_1, t_2)$  包含绝对连续部分和奇异部分, 奇异部分对应于正概率事件  $T_1 = T_2$ . 这一现象的物理解释是, 极端操作条件可能导致部件同时失效. 为了进一步分解  $S(t_1, t_2)$ , 定义如下符号

$$\bigtriangledown \kappa_{3i} = \kappa(\theta - \eta_i/\lambda) - \kappa(\theta - (\eta_1 + \eta_2)/\lambda), \quad \bigtriangledown \kappa_{i0} = \kappa(\theta) - \kappa(\theta - \eta_i/\lambda), \quad i = 1, 2,$$

以及

$$\bigtriangledown \kappa_{30} = \kappa(\theta) - \kappa(\theta - (\eta_1 + \eta_2)/\lambda).$$

基于这些符号, 则有以下结论.

**定理 5.5.** 令  $\alpha = \frac{\bigtriangledown \kappa_{10} - \bigtriangledown \kappa_{32}}{\bigtriangledown \kappa_{30}}$ , 则有

$$S(t_1, t_2) = \alpha S_s(t_1, t_2) + (1 - \alpha) S_c(t_1, t_2),$$

其中  $S_s(t_1, t_2) = \exp\{-\lambda \Lambda(\max(t_1, t_2); \alpha) \bigtriangledown \kappa_{30}\}$  是奇异生存函数,  $S_c(t_1, t_2) = \frac{1}{1 - \alpha} S(t_1, t_2) - \frac{\alpha}{1 - \alpha} S_s(t_1, t_2)$  是绝对连续的生存函数.

**证明:** 为求得  $S(t_1, t_2)$  的绝对连续部分, 需对  $S(t_1, t_2)$  关于  $t_1$  和  $t_2$  求二阶偏导数, 有

$$(1 - \alpha) f_c(t_1, t_2) = \frac{\partial^2 S(t_1, t_2)}{\partial t_1 \partial t_2}.$$

接下来进行具体计算

$$(1-\alpha)f_c(t_1, t_2) = \begin{cases} \exp\{-\lambda\Lambda(t_1; \alpha)\bigtriangledown\kappa_{32} - \lambda\Lambda(t_2; \alpha)\bigtriangledown\kappa_{20}\} \\ \quad \times \lambda^2 \bigtriangledown\kappa_{32} \bigtriangledown\kappa_{20} \prod_{i=1}^2 \frac{\partial\Lambda(t_i; \alpha)}{\partial t_i}, & t_1 < t_2, \\ \exp\{-\lambda\Lambda(t_2; \alpha)\bigtriangledown\kappa_{31} - \lambda\Lambda(t_1; \alpha)\bigtriangledown\kappa_{10}\} \\ \quad \times \lambda^2 \bigtriangledown\kappa_{31} \bigtriangledown\kappa_{10} \prod_{i=1}^2 \frac{\partial\Lambda(t_i; \alpha)}{\partial t_i}, & t_1 > t_2. \end{cases}$$

对于  $t_1 < t_2$  的情况, 绝对连续部分  $S_c(t_1, t_2)$  可写为

$$(1-\alpha)S_c(t_1, t_2) = \exp\{-\lambda[\Lambda(t_2; \alpha)\bigtriangledown\kappa_{32} + \Lambda(t_1; \alpha)\bigtriangledown\kappa_{20}]\} \\ - \frac{\bigtriangledown\kappa_{20} - \bigtriangledown\kappa_{31}}{\bigtriangledown\kappa_{30}} \exp\{-\lambda\Lambda(t_2; \alpha)\bigtriangledown\kappa_{30}\}.$$

对于  $t_1 > t_2$  的情况,

$$(1-\alpha)S_c(t_1, t_2) = \int_{t_2}^{\infty} \int_{t_1}^{\infty} (1-\alpha)f_c(x, y) dx dy \\ = \int_{t_1}^{\infty} \int_{t_1}^y (1-\alpha)f_c(x, y) dx dy + \int_{t_1}^{\infty} \int_{t_2}^x (1-\alpha)f_c(x, y) dy dx \\ = \exp\{-\lambda[\Lambda(t_1; \alpha)\bigtriangledown\kappa_{31} + \Lambda(t_2; \alpha)\bigtriangledown\kappa_{10}]\} \\ - \frac{\bigtriangledown\kappa_{10} - \bigtriangledown\kappa_{32}}{\bigtriangledown\kappa_{30}} \exp\{-\lambda\Lambda(t_1; \alpha)\bigtriangledown\kappa_{30}\},$$

结合两种情况, 可统一表示为

$$(1-\alpha)S_c(t_1, t_2) = S(t_1, t_2) - \frac{\bigtriangledown\kappa_{10} - \bigtriangledown\kappa_{32}}{\bigtriangledown\kappa_{30}} \exp\{-\lambda\Lambda(\max(t_1, t_2); \alpha)\bigtriangledown\kappa_{30}\}.$$

根据  $S_c(t_1, t_2)$  和  $\alpha$  的已知条件  $S_c(0, 0) = S(0, 0) = 1$ , 可得

$$\alpha = \frac{\bigtriangledown\kappa_{10} - \bigtriangledown\kappa_{32}}{\bigtriangledown\kappa_{30}}.$$

奇异部分可以通过相减得到, 即

$$\begin{aligned} S_s(t_1, t_2) &= \frac{S(t_1, t_2) - (1 - \alpha)S_c(t_1, t_2)}{\alpha} \\ &= \exp \{-\lambda\Lambda(\max(t_1, t_2); \alpha)\triangleleft\kappa_{30}\}. \end{aligned}$$

□

该定理可进一步得到  $T_1$  和  $T_2$  的联合 PDF 表达式

$$f(t_1, t_2) = \begin{cases} \exp \{-\lambda\Lambda(t_1; \alpha)\triangleleft\kappa_{32} - \lambda\Lambda(t_2; \alpha)\triangleleft\kappa_{20}\} \\ \quad \times \lambda^2 \triangleleft\kappa_{32} \triangleleft\kappa_{20} \prod_{i=1}^2 \frac{\partial\Lambda(t_i; \alpha)}{\partial t_i}, & t_1 < t_2, \\ \exp \{-\lambda\Lambda(t_2; \alpha)\triangleleft\kappa_{31} - \lambda\Lambda(t_1; \alpha)\triangleleft\kappa_{10}\} \\ \quad \times \lambda^2 \triangleleft\kappa_{31} \triangleleft\kappa_{10} \prod_{i=1}^2 \frac{\partial\Lambda(t_i; \alpha)}{\partial t_i}, & t_1 > t_2, \\ \exp \{-\lambda\Lambda(t; \alpha)\triangleleft\kappa_{30}\} \\ \quad \times \lambda(\triangleleft\kappa_{10} - \triangleleft\kappa_{32}) \frac{\partial\Lambda(t; \alpha)}{\partial t}, & t_1 = t_2 = t. \end{cases} \quad (5.31)$$

在实际应用中, 给定观测数据后, 可以基于式 (5.30) 和 (5.31) 构建似然函数. 然而, 在某些特定情况下, 模型参数可能会产生冗余. 例如, 当观测数据满足  $t_{i1} < t_{i2}$  且  $\forall i = 1, \dots, n$  时, 模型参数  $\Theta = (\lambda, d, \eta_1, \eta_2)$  的似然函数为

$$L(\Theta) = \prod_{i=1}^n \exp \{-\lambda\Lambda(t_{i1}; \alpha)\triangleleft\kappa_{32} - \lambda\Lambda(t_{i2}; \alpha)\triangleleft\kappa_{20}\} \lambda^2 \triangleleft\kappa_{32} \triangleleft\kappa_{20} \prod_{j=1}^2 \frac{\partial\Lambda(t_{ij}; \alpha)}{\partial t_{ij}}.$$

为简化分析, 定义  $\beta_1 = \lambda\triangleleft\kappa_{32}$  和  $\beta_2 = \lambda\triangleleft\kappa_{20}$ , 此时似然函数变为

$$L(\beta_1, \beta_2) = \prod_{i=1}^n \exp \{-\beta_1\Lambda(t_{i1}; \alpha) - \beta_2\Lambda(t_{i2}; \alpha)\} \beta_1 \beta_2 \prod_{j=1}^2 \frac{\partial\Lambda(t_{ij}; \alpha)}{\partial t_{ij}}, \quad (5.32)$$

这表明  $\lambda, d, \eta_1$  和  $\eta_2$  无法同时被识别, 因为对任意的  $d$ , 即使设定  $\lambda = 1$ ,  $L(\beta_1, \beta_2)$  的最大值都相同. 这说明该类数据没有提供任何关于部件寿命相关性的信息. 进一步地, 定义  $\nu_1 = \lambda\triangleleft\kappa_{32}$ ,  $\nu_2 = \lambda\triangleleft\kappa_{31}$ ,  $\nu_0 = \lambda(\triangleleft\kappa_{10} - \triangleleft\kappa_{32})$ . 当  $\Lambda(t) = t$  或  $\Lambda(t) = t^\alpha$  时, 模型 (5.31) 分别退化为 Marshall-Olkin 双变量指数分布和 Weibull 分布 (Marshall 等, 1967; Xu 等, 2017).

观测数据的分布形式会影响参数估计的存在性. 令  $n_1, n_2$  和  $n_0$  分别表示观测数据中  $t_{i1} < t_{i2}$ ,  $t_{i1} > t_{i2}$  和  $t_{i1} = t_{i2}$  的数量. 根据 (Bemis 等, 1972), 当  $n_j = 0$  ( $j = 0, 1, 2$ ) 时,  $(\nu_0, \nu_1, \nu_2)$  的 MLE 不存在; 当  $n_j > 0$  时, 虽然 MLE 存在, 但模型 (5.31) 可能仍存在参数冗余问题. 为避免此问题, 建议在实际数据分析中设定  $\lambda = 1$ , 并基于观测数据优化参数  $d$  的取值.

### 5.2.5 模拟实验

本节通过模拟研究评估所提出模型的性能. 假设一个系统由两个部件构成, 且该系统在动态环境下运行. 同时假定有  $n$  个系统进行测试, 其中  $n$  分别取 15、20、30 和 50. 假设部件的基失效率别为  $\eta_1 = 1$  和  $\eta_2 = 1.5$ . 部件的随机时间尺度由 ED 过程描述, 其中  $\lambda(t) = t^\alpha$ , 且  $\alpha = 2$ . 此时, 部件寿命的联合 PDF 满足式 (5.31).

为研究模型在不同退化模式下的表现, 分别模拟了  $d = 1, 1.5, 2$  和  $3$  的情形, 分别对应于非齐次泊松过程、复合泊松过程、伽马过程和 IG 过程. 对于每组  $(n, d)$  组合, 生成了 10000 个独立数据集. 采用统一的模型框架 (5.31) 来编写似然函数, 并使用极大似然法估计参数  $(\alpha, d, \eta_1, \eta_2)$ . 同时, 基于渐近正态性定理构建参数的区间估计. 为评价参数估计性能, 计算了偏差、RMSE 和 95% 覆盖概率, 结果列于表 5.10 和表 5.11 中. 从结果可以发现, 所有情况下 RB 均接近 0, RMSE 随样本量  $n$  增加而显著减小. 此外, 经验 CP 接近名义水平 0.95, 表明拟合数据的模型能够准确估计参数. 这一结果验证了所提模型的可行性, 同时凸显了模型在不同数据条件下的鲁棒性与灵活性.

为了进一步说明模型的应用价值, 我们绘制了复合泊松过程 ( $d = 1.5$ ) 下的估计系统可靠度函数 (如图 5.7 所示). 对于其他  $d$  的取值, 估计结果具有类似趋势, 这里不再赘述. 从图 5.7 可以看出, 当样本量  $n = 15$  和  $20$  时, 系统可靠度函数的估计相较于真实可靠度函数略有偏差; 然而, 当样本量增加至  $n = 30$  和  $50$  时, RB 显著减小, 估计结果与真实可靠度函数高度一致. 这表明, 即使在中等样本量下, 模型也能够提供合理且准确的参数估计和可靠度函数估计.

表 5.10: 参数估计的性能评估.

模型		n = 15				n = 20			
		$\alpha$	$d$	$\eta_1$	$\eta_2$	$\alpha$	$d$	$\eta_1$	$\eta_2$
$d = 1$	偏差	0.140	0.136	0.144	0.209	0.107	0.102	0.063	0.156
	RMSE	0.365	0.540	0.609	0.849	0.316	0.457	0.432	0.801
$d = 1.5$	偏差	0.152	0.121	0.085	0.204	0.108	0.089	0.049	0.123
	RMSE	0.387	0.795	0.594	0.902	0.293	0.676	0.414	0.701
$d = 2$	偏差	0.143	0.312	0.094	0.235	0.114	0.256	0.062	0.104
	RMSE	0.372	1.367	0.478	0.908	0.308	1.136	0.414	0.633
$d = 3$	偏差	0.142	0.674	0.075	0.191	0.111	0.655	0.055	0.140
	RMSE	0.358	2.396	0.461	0.857	0.300	2.168	0.373	0.641

模型		n = 30				n = 50			
		$\alpha$	$d$	$\eta_1$	$\eta_2$	$\alpha$	$d$	$\eta_1$	$\eta_2$
$d = 1$	偏差	0.084	0.083	0.049	0.104	0.036	0.0057	0.030	0.067
	RMSE	0.254	0.318	0.355	0.601	0.173	0.180	0.241	0.415
$d = 1.5$	偏差	0.083	0.050	0.040	0.099	0.044	0.028	0.025	0.059
	RMSE	0.240	0.404	0.324	0.548	0.184	0.299	0.248	0.408
$d = 2$	偏差	0.075	0.146	0.037	0.086	0.041	0.071	0.023	0.047
	RMSE	0.247	0.781	0.304	0.521	0.179	0.495	0.233	0.375
$d = 3$	偏差	0.087	0.523	0.021	0.079	0.044	0.274	0.020	0.044
	RMSE	0.231	1.891	0.294	0.472	0.174	1.316	0.219	0.347

表 5.11: 参数估计的 95% 区间覆盖概率.

模型	$n = 15$				$n = 20$			
	$\alpha$	$d$	$\eta_1$	$\eta_2$	$\alpha$	$d$	$\eta_1$	$\eta_2$
$d = 1$	0.951	0.976	0.928	0.920	0.962	0.990	0.930	0.912
$d = 1.5$	0.942	0.986	0.914	0.888	0.957	0.986	0.919	0.927
$d = 2$	0.937	0.924	0.917	0.921	0.944	0.914	0.923	0.917
$d = 3$	0.948	0.889	0.919	0.918	0.931	0.890	0.922	0.919

模型	$n = 30$				$n = 50$			
	$\alpha$	$d$	$\eta_1$	$\eta_2$	$\alpha$	$d$	$\eta_1$	$\eta_2$
$d = 1$	0.944	0.984	0.926	0.932	0.950	0.964	0.948	0.948
$d = 1.5$	0.952	0.967	0.929	0.929	0.949	0.967	0.937	0.940
$d = 2$	0.937	0.924	0.917	0.921	0.948	0.941	0.940	0.939
$d = 3$	0.956	0.914	0.927	0.932	0.950	0.921	0.937	0.941

### 5.2.6 实例分析

本节采用所提出的 ED 模型对糖尿病视网膜病变研究中双眼失明时间数据进行分析 (Csorgo 等, 1989). 该研究由国家眼科研究所发起, 旨在评估激光治疗对延缓糖尿病视网膜病变患者失明进程的效果. 研究共纳入 71 名患者, 每位患者随机选择一只眼睛接受激光治疗, 另一只眼睛作为未治疗的对照组. 设  $T_1$  表示接受激光治疗眼睛的失明时间,  $T_2$  表示未治疗眼睛的失明时间. 为评估两只眼睛失明时间的相关性以及激光治疗的有效性, 本文采用  $\Lambda(t) = t^\alpha$  的 ED 模型对数据进行拟合, 结果如表 5.12 所示. 从表中可以看到, 估计参数  $d = 1.315$ , 其 95% 置信区间为  $(1.140, 1.490)$ . 这一结果表明  $T_1$  和  $T_2$  之间存在显著的相关性, 且其依赖性符合复合泊松过程的特性. 此外, 计算得到  $P(T_1 < T_2)$  (即激光治疗眼睛的失明时间早于未治疗眼睛的概率) 估计值为 0.394, 95% 置信区间为  $(0.281, 0.507)$ . 该结果表明, 激光治疗能够显著延缓患者失明的进程. 具体而言, 概率  $P(T_1 < T_2)$  小于 0.5, 说明激光治疗眼睛的失明时间通常晚于未治疗眼睛. 这不仅验证了激光治疗的有效性, 也凸显了所提模型在处理双变量失效时间数据中的优势.

表 5.12: 参数的估计值、标准差及 95% 置信区间.

	$\alpha$	$d$	$\eta_1$	$\eta_2$
估计值	1.595	1.315	0.312	0.370
标准差	0.103	0.0892	0.0593	0.0696
95% CI	(1.393, 1.798)	(1.140, 1.490)	(0.196, 0.428)	(0.234, 0.507)

### 5.3 部件寿命排序约束下系统可靠性评估

在上一节讨论了动态使用环境下部件寿命之间的相关关系后, 本节进一步考虑部件寿命带排序约束的情形. 这一问题的提出来源于对刹车片寿命数据的分析. 图 5.8 展示了 35 辆出租车的前后刹车片寿命数据, 这些车辆由同一公司运营, 并于 2013 年在同一城市行驶. 刹车片作为制动系统的关键部件, 以串联方式协同工作. 其主要失效模式是因长期使用导致磨损. 车辆的制动频率受到多种因素的影响, 例如道路条件、速度限制、障碍物以及驾驶员的操作习惯, 这些因素体现了系统运行环境随时间变化的动态特性. 此外, 图中还发现以下规律: (1) 正相关性: 前后刹车片的寿命具有正相关性, 表现为寿命较短或较长的刹车片更倾向成对出现; (2) 寿命排序约束: 前刹车片的寿命始终短于后刹车片. 这是因为在制动过程中, 更多的车辆重量施加在前轮上, 导致前刹车片承受更高的负载并更快磨损. 这一现象在文献中被称为排序约束, 在许多系统中都普遍存在, 尤其是某些部件因负载分配不均而承受更高应力的情况.

针对上述分析, 本章将系统研究动态操作环境、寿命依赖以及排序约束共同作用下的系统可靠性评估问题. 在 Hong 等 (2019) 和 Xu 等 (2021) 的研究基础上, 假设所有部件在同一动态操作环境下运行, 通过随机过程来刻画累积风险函数随环境变化的特征, 将动态环境的影响和部件寿命间的相依关系统一模型中. 同时, 为解决排序约束问题, 对联合寿命分布的支撑进行合理化截断, 从而准确描述部件间的排序关系. 此外, 在统计推断方面, 提出了广义推断方法构建参数和系统可靠性的区间估计. 本章结构安排如下: 第 5.3.1 节介绍模型框架; 第 5.3.2 节介绍模型参数和系统可靠性的统计推断方法; 第 5.3.3 节通过模拟仿真评估区间估计方法的性能; 第 5.3.4 节通过实际数据来说明模型和方法的可行性.

### 5.3.1 模型

设  $T$  表示部件寿命,  $S(t)$  表示生存函数. 则有  $S(t) = \exp\{-C(t)\}$ , 其中  $C(t)$  是累积风险函数. 假设  $C(t)$  服从 ED 过程  $\{Y(t), t \geq 0\}$ , 记作  $\mathcal{ED}(\mu\Lambda(t; \alpha), \lambda)$ . 在双部件情形下, 模型表达为

$$T_i | C(t) \sim S(t | C(t)), \quad C(t) \sim \mathcal{ED}(\mu\Lambda(t; \alpha), \lambda) \quad i = 1, 2, \quad (5.33)$$

其中  $T_i$  表示第  $i$  个部件的寿命.

基于模型 (5.33),  $T_i$  的边际生存函数为:

$$S_{T_i}(t) = \mathbb{E}[\exp(-C(t))] = \exp\{\lambda\Lambda(t; \alpha)[\kappa(\theta - 1/\lambda) - \kappa(\theta)]\}.$$

通过选取不同形式的  $\Lambda(t; \alpha)$ , 可得不同寿命分布, 如  $\Lambda(t; \alpha) = t$  的指数分布,  $\Lambda(t; \alpha) = t^\alpha$  的威布尔分布,  $\Lambda(t; \alpha) = \exp(\alpha t) - 1$  的 Gompertz 分布, 以及  $\Lambda(t; \alpha) = \log(1 + t/\alpha)$  的 Lomax 分布等. 基于模型 (5.33), 部件寿命  $T_1$  和  $T_2$  的联合生存函数可由下面的定理 5.6 给出.

**定理 5.6.** 寿命  $T_1$  和  $T_2$  的联合生存函数为

$$S(t_1, t_2) = \exp \left\{ -\lambda\Lambda(t_{(1)}; \alpha) \left[ \kappa\left(\theta - \frac{1}{\lambda}\right) - \kappa\left(\theta - \frac{2}{\lambda}\right) \right] - \lambda\Lambda(t_{(2)}; \alpha) \left[ \kappa(\theta) - \kappa\left(\theta - \frac{1}{\lambda}\right) \right] \right\},$$

其中  $t_{(1)} < t_{(2)}$  是将  $t_1$  和  $t_2$  按升序排序得到的值. 此外, 部件寿命  $T_1$  和  $T_2$  是正象限相关的.

**证明:** 根据 ED 过程的性质,  $C(t_{(1)})$  和  $\Delta C(t_{(2)}) = C(t_{(2)}) - C(t_{(1)})$  相互独立, 且满足:  $C(t_{(1)}) \sim ED(\mu\Lambda(t_{(1)}; \alpha), \lambda)$  和  $\Delta C(t_{(2)}) \sim ED(\mu\Delta\Lambda(t_{(2)}; \alpha), \lambda)$ . 在  $C(t_{(1)})$  和  $\Delta C(t_{(2)})$  的条件下, 有

$$P(T_1 > t_1, T_2 > t_2 | C(t_{(1)}), \Delta C(t_{(2)})) = \exp\{-2C(t_{(1)})\} \exp\{-\Delta C(t_{(2)})\}.$$

对上述条件概率关于  $C(t_{(1)})$  和  $\Delta C(t_{(2)})$  取期望, 可得

$$\begin{aligned}
 S(t_1, t_2) &= P(T_1 > t_1, T_2 > t_2) \\
 &= \mathbb{E}_{C(t_{(1)})} [\exp\{-2C(t_{(1)})\}] \cdot \mathbb{E}_{\Delta C(t_{(2)})} [\exp\{-\Delta C(t_{(2)})\}] \\
 &= \exp \left\{ \lambda \Lambda(t_{(1)}; \alpha) \left[ \kappa \left( \theta - \frac{2}{\lambda} \right) - \kappa(\theta) \right] \right\} \\
 &\quad \times \exp \left\{ \lambda \Delta \Lambda(t_{(2)}; \alpha) \left[ \kappa \left( \theta - \frac{1}{\lambda} \right) - \kappa(\theta) \right] \right\} \\
 &= \exp \left\{ -\lambda \left[ \Lambda(t_{(1)}; \alpha) \left( \kappa \left( \theta - \frac{1}{\lambda} \right) - \kappa \left( \theta - \frac{2}{\lambda} \right) \right) \right. \right. \\
 &\quad \left. \left. + \Lambda(t_{(2)}; \alpha) \left( \kappa(\theta) - \kappa \left( \theta - \frac{1}{\lambda} \right) \right) \right] \right\}.
 \end{aligned}$$

对  $T_i$  的生存函数进行推导, 可得

$$\begin{aligned}
 P(T_i > t_i) &= \mathbb{E}_{C(t_i)} [\exp\{-C(t_i)\}] \\
 &= \exp \left\{ -\lambda \Lambda(t_i; \alpha) \left[ \kappa(\theta) - \kappa \left( \theta - \frac{1}{\lambda} \right) \right] \right\}, \quad i = 1, 2.
 \end{aligned}$$

进一步计算得

$$\begin{aligned}
 Q(t_1, t_2) &= \frac{P(T_1 > t_1, T_2 > t_2)}{P(T_1 > t_1)P(T_2 > t_2)} \\
 &= \exp \left\{ \lambda \Lambda(t_{(1)}; \alpha) \left[ \kappa(\theta) - 2\kappa \left( \theta - \frac{1}{\lambda} \right) + \kappa \left( \theta - \frac{2}{\lambda} \right) \right] \right\}.
 \end{aligned}$$

由于  $\kappa(\cdot)$  是凸函数, 满足

$$\kappa(\theta) - 2\kappa \left( \theta - \frac{1}{\lambda} \right) + \kappa \left( \theta - \frac{2}{\lambda} \right) \geq 0,$$

可知  $Q(t_1, t_2) \geq 1$ , 即  $T_1$  和  $T_2$  是正象限相关的.  $\square$

类似于上一节的内容, 可以证明  $S(t_1, t_2)$  在  $t_1 = t_2$  时是不连续的, 并表

现出奇异性. 其对应的联合 PDF 为

$$f(t_1, t_2) = \begin{cases} \lambda^2 \nabla \kappa_1 \nabla \kappa_2 \prod_{i=1}^2 \frac{\partial \Lambda(t_i; \alpha)}{\partial t_i} \exp\{-\lambda \Lambda(t_1; \alpha) \nabla \kappa_1 - \lambda \Lambda(t_2; \alpha) \nabla \kappa_2\}, & t_1 < t_2, \\ \lambda^2 \nabla \kappa_1 \nabla \kappa_2 \prod_{i=1}^2 \frac{\partial \Lambda(t_i; \alpha)}{\partial t_i} \exp\{-\lambda \Lambda(t_2; \alpha) \nabla \kappa_1 - \lambda \Lambda(t_1; \alpha) \nabla \kappa_2\}, & t_1 > t_2, \\ \lambda (\nabla \kappa_2 - \nabla \kappa_1) \frac{\partial \Lambda(t; \alpha)}{\partial t} \exp\{-\lambda \Lambda(t; \alpha) \nabla \kappa_0\}, & t_1 = t_2 = t, \end{cases}$$

其中

$$\begin{aligned} \nabla \kappa_0 &= \kappa(\theta) - \kappa(\theta - 2/\lambda), \\ \nabla \kappa_1 &= \kappa(\theta - 1/\lambda) - \kappa(\theta - 2/\lambda), \\ \nabla \kappa_2 &= \kappa(\theta) - \kappa(\theta - 1/\lambda). \end{aligned}$$

在此模型下,  $T_1$  和  $T_2$  之间可能存在以下三种关系: 第一部分为  $T_1 < T_2$ , 即第一个部件先于第二个部件失效; 第二部分为  $T_1 > T_2$  即第二个部件先于第一个部件失效; 第三部分为  $T_1 = T_2$ , 表示外部环境的极端影响导致两个部件同时失效. 然而, 模型 (5.33) 无法描述  $T_1$  严格小于  $T_2$  的现象, 例如在刹车片寿命数据中所观察到的情况. 为了解决这一问题, 可以对联合 PDF 的支撑进行截断, 将其限制在  $t_1 < t_2$  的区域内.

**定理 5.7.** 基于模型 (5.33), 若  $T_1$  严格小于  $T_2$ , 那么  $T_1$  和  $T_2$  的联合 PDF 为

$$\begin{aligned} f_{TR}(t_1, t_2) &= \lambda^2 \nabla \kappa_0 \nabla \kappa_2 \\ &\times \prod_{i=1}^2 \frac{\partial \Lambda(t_i; \alpha)}{\partial t_i} \exp\{-\lambda \Lambda(t_1; \alpha) \nabla \kappa_1 - \lambda \Lambda(t_2; \alpha) \nabla \kappa_2\}, & t_1 < t_2. \end{aligned} \tag{5.34}$$

**证明:** 首先, 计算  $T_1 < T_2$  的概率

$$\begin{aligned} P(T_1 < T_2) &= \int_0^\infty \int_0^{t_2} f(t_1, t_2) dt_1 dt_2 \\ &= \int_0^\infty \int_0^{t_2} \lambda^2 \nabla \kappa_1 \nabla \kappa_2 \end{aligned}$$

$$\begin{aligned}
& \times \prod_{i=1}^2 \frac{\partial \Lambda(t_i; \alpha)}{\partial t_i} \exp \{-\lambda \Lambda(t_1; \alpha) \nabla \kappa_1 - \lambda \Lambda(t_2; \alpha) \nabla \kappa_2\} dt_1 dt_2 \\
& = \int_0^\infty \lambda \nabla \kappa_2 \frac{\partial \Lambda(t_2; \alpha)}{\partial t_2} \exp \{-\lambda \Lambda(t_2; \alpha) \nabla \kappa_2\} \\
& \quad \times [1 - \exp \{-\lambda \Lambda(t_2; \alpha) \nabla \kappa_1\}] dt_2 \\
& = 1 - \frac{\nabla \kappa_2}{\nabla \kappa_1 + \nabla \kappa_2} = \frac{\nabla \kappa_1}{\nabla \kappa_0}.
\end{aligned}$$

因此, 区域  $T_1 < T_2$  下,  $T_1$  和  $T_2$  的联合 PDF 为

$$\begin{aligned}
f_{TR}(t_1, t_2) &= \frac{f(t_1, t_2)}{P(T_1 < T_2)} \\
&= \lambda^2 \nabla \kappa_0 \nabla \kappa_2 \prod_{i=1}^2 \frac{\partial \Lambda(t_i; \alpha)}{\partial t_i} \\
&\quad \times \exp \{-\lambda \Lambda(t_1; \alpha) \nabla \kappa_1 - \lambda \Lambda(t_2; \alpha) \nabla \kappa_2\}, \quad t_1 < t_2.
\end{aligned}$$

当  $T_1$  严格大于  $T_2$  时, 可以采用与式 (5.31) 的推导类似的方法, 得到  $T_1$  和  $T_2$  的联合 PDF. 虽然新的联合 PDF 能够反映寿命排序约束, 但仍需验证其是否保留了  $T_1$  和  $T_2$  之间的相关性. 这种相关性来源于部件所处的共同随机环境, 忽略这一依赖性可能导致系统可靠性评估出现偏差, 如过高或过低的估计 (Hong 等, 2019; Kotz 等, 2003). 一种常用的验证方法是分析  $T_1$  和  $T_2$  的二阶全正序 (total positive of order 2) 关系 (Shaked, 1977), 它是一类强相关关系, 可推出正象限相关与皮尔逊线性相关 (Lai 等, 2006).

**定义 5.1.** 设随机变量  $X$  和  $Y$  的联合 PDF 为  $p(x, y)$ . 如果对于所有  $x_1 < x_2$  和  $y_1 < y_2$ , 有  $p(x_1, y_1)p(x_2, y_2) \geq p(x_1, y_2)p(x_2, y_1)$ , 则称  $p(x, y)$  是二阶全正序的.

根据定义 5.1, 需要验证对于任意  $t_{11} < t_{12}$  和  $t_{21} < t_{22}$ , 不等式

$$f_{TR}(t_{11}, t_{21}) f_{TR}(t_{12}, t_{22}) \geq f_{TR}(t_{11}, t_{22}) f_{TR}(t_{12}, t_{21}) \quad (5.35)$$

成立, 其中  $t_{i1}$  和  $t_{i2}$  分别为  $T_i$  的任意观测值. 考虑如下两种情形:

- (1) 当  $t_{12} \geq t_{21}$  时: 由于  $t_{12}$  和  $t_{21}$  不满足  $t_1 < t_2$  的条件,  $f_{TR}(t_{12}, t_{21}) = 0$ . 显然, 不等式(5.35)成立.

(2) 当  $t_{12} < t_{21}$  时: 根据  $f_{TR}$  的表达式, 计算得到

$$\begin{aligned} f_{TR}(t_{11}, t_{21}) f_{TR}(t_{12}, t_{22}) &= f_{TR}(t_{11}, t_{22}) f_{TR}(t_{12}, t_{21}) \\ &= \lambda^4 [\nabla \kappa_0]^2 [\nabla \kappa_2]^2 \prod_{i=1}^2 \prod_{j=1}^2 \frac{\partial H_\alpha(t_{ij})}{\partial t_{ij}} \\ &\quad \times \exp \{ -\lambda [H_\alpha(t_{11}) + H_\alpha(t_{12})] \nabla \kappa_1 \\ &\quad - \lambda [H_\alpha(t_{21}) + H_\alpha(t_{22})] \nabla \kappa_2 \}. \end{aligned}$$

结果两边相等, 因此不等式成立.

由此可知,  $T_1$  和  $T_2$  是二阶全正序的. 截断后的联合 PDF  $f_{TR}(t_1, t_2)$  能同时刻画随机变量间的相关性与排序关系. 然而, 该模型的相关性特征无法通过单一参数完全描述. 由于  $T_1$  和  $T_2$  受共同 ED 过程影响, 其依赖结构呈现非线性复杂性, 传统的皮尔逊相关系数难以有效反映这一特点. 事实上, 该模型揭示了一种非线性相关结构, 而这种相关性可能无法通过模型参数直接量化. 因此, 在实际应用中 (见第 5.3.4 节), 我们采用斯皮尔曼秩相关系数来验证  $T_1$  和  $T_2$  之间是否存在依赖性, 以更准确地刻画两者的相关特征.

### 5.3.2 统计推断

本节将从以下两个方面展开详细讨论: 一是点估计, 重点在于通过优化似然函数获取模型参数的估计值; 二是区间估计, 利用参数的不确定性信息构建合理的置信区间, 从而为推断结果提供更全面的可信度描述.

#### 5.3.2.1 点估计

假设观测数据为  $\{(t_{1j}, t_{2j}), j = 1, \dots, n\}$ , 并且对于每个  $j$ , 满足  $t_{1j} < t_{2j}$ . 设模型参数为  $\boldsymbol{\theta} = (\lambda, \theta, d, \alpha)$ . 基于式 (5.34) 给出的  $f_{TR}(t_1, t_2)$ , 其似然函数可表示为

$$\begin{aligned} L_1(\boldsymbol{\theta}) &= \lambda^{2n} [\nabla \kappa_0]^n [\nabla \kappa_2]^n \\ &\times \prod_{i=1}^2 \prod_{j=1}^n \Lambda'(t_{ij}; \alpha) \exp \left\{ -\lambda \nabla \kappa_1 \sum_{j=1}^n \Lambda(t_{1j}; \alpha) - \lambda \nabla \kappa_2 \sum_{j=1}^n \Lambda(t_{2j}; \alpha) \right\}, \end{aligned} \quad (5.36)$$

其中  $\Lambda'(t_{ij}; \alpha) = \frac{\partial \Lambda(t; \alpha)}{\partial t} \Big|_{t=t_{ij}}$ . 为简化表示, 定义  $\varsigma_1 = \lambda \nabla \kappa_0$ ,  $\varsigma_2 = \lambda \nabla \kappa_2$ , 以及  $\omega = (\varsigma_1, \varsigma_2, \alpha)$ . 则似然函数变为

$$L_2(\omega) = \varsigma_2^n \varsigma_1^n \prod_{i=1}^2 \prod_{j=1}^n \Lambda'(t_{ij}; \alpha) \\ \times \exp \left\{ -\varsigma_1 \sum_{j=1}^n \Lambda(t_{1j}; \alpha) - \varsigma_2 \sum_{j=1}^n [\Lambda(t_{2j}; \alpha) - \Lambda(t_{1j}; \alpha)] \right\}. \quad (5.37)$$

通过最大化对数似然函数  $\log L_2(\omega)$ , 可得到  $\varsigma_1$  和  $\varsigma_2$  的 MLE

$$\hat{\varsigma}_1 = \frac{n}{\sum_{j=1}^n \Lambda(t_{1j}; \hat{\alpha})}, \quad \hat{\varsigma}_2 = \frac{n}{\sum_{j=1}^n [\Lambda(t_{2j}; \hat{\alpha}) - \Lambda(t_{1j}; \hat{\alpha})]}.$$

此外,  $\alpha$  的 MLE 可通过最大化以下目标函数获得

$$-n \log \left( \sum_{j=1}^n [\Lambda(t_{2j}; \alpha) - \Lambda(t_{1j}; \alpha)] \right) - n \log \left( \sum_{j=1}^n \Lambda(t_{1j}; \alpha) \right) + \sum_{i=1}^2 \sum_{j=1}^n \log \Lambda'(t_{ij}; \alpha).$$

**注 6:** 由公式可知,  $(\hat{\varsigma}_1, \hat{\varsigma}_2)$  的值仅依赖于观测数据, 而不依赖于  $\kappa(\cdot)$  的具体形式. 这意味着参数  $\lambda$ ,  $\theta$  和  $d$  无法同时被完全识别, 只能基于观测数据估计其中两个. 为解决参数识别问题, 建议采用以下两种统计推断方案: (1) 固定  $d$ : 即选择一个指定的随机过程  $\{Y(t), t \geq 0\}$  将动态环境的影响纳入模型. (2) 固定  $\lambda = 1$ : 即通过对扩散参数进行缩放, 使得数据能够学习最优的  $d$  值, 从而选择合适的随机过程来描述动态环境的随机影响.

### 5.3.2.2 区间估计

在样本量较小时, 基于正态近似或自助法的传统区间构造方法可能会表现不佳, 特别是在复杂模型的场景下, 可能导致置信区间覆盖率不足或长度过大. 为提高参数估计的精确性, 采用广义推断方法构建模型参数的置信区间. 这一方法具有更高的精度, 尤其适用于处理复杂模型和小样本数据的场景.

构建  $\alpha$  的精确置信区间需要依赖以下两个引理 (Wang B 等, 2010).

**引理 5.1.** 如果  $V_{1:n} < \dots < V_{n:n}$  是从标准指数分布中抽取的样本容量为  $n$  的

顺序统计量, 设  $W_1 = nV_{1:n}$ ,  $W_i = (n-i+1)(V_{i:n} - V_{i-1:n})$ ,  $i = 2, \dots, n$ , 则  $W_1, \dots, W_n$  是相互独立的标准指数随机变量.

**引理 5.2.** 如果  $Q_i = \sum_{j=1}^i W_j$ ,  $i = 1, \dots, n$ , 且  $U_i = Q_i/Q_n$ ,  $i = 1, \dots, n-1$ , 则  $U_1 < \dots < U_{n-1}$  是从均匀  $(0, 1)$  分布中抽取的样本容量为  $n-1$  的顺序统计量.

基于引理 5.1 和 5.2, 可得到以下结果.

**定理 5.8.** 设  $Z_1 = \Lambda(T_1; \alpha)$  和  $Z_2 = \Lambda(T_2; \alpha) - \Lambda(T_1; \alpha)$ . 则  $Z_1 \sim \text{Exp}(\varsigma_1)$  和  $Z_2 \sim \text{Exp}(\varsigma_2)$ , 其中  $\text{Exp}(\varsigma)$  表示具有速率  $\varsigma$  的指数分布. 此外,  $Z_1$  和  $Z_2$  是独立的,

**证明:** 设  $T_1 = H_\alpha^{-1}(Z_1)$  和  $T_2 = H_\alpha^{-1}(Z_1 + Z_2)$ , 可知  $(Z_1, Z_2)'$  与  $(T_1, T_2)'$  为一一变换关系. 令  $\mathcal{J}$  表示该变换的雅可比矩阵, 其表达式为

$$\mathcal{J} = \begin{vmatrix} \frac{\partial(T_1, T_2)'}{\partial(Z_1, Z_2)'} \end{vmatrix} = \frac{1}{H'_\alpha(H_\alpha^{-1}(Z_1)) \cdot H'_\alpha(H_\alpha^{-1}(Z_1 + Z_2))}.$$

因此, 随机向量  $Z = (Z_1, Z_2)'$  的联合 PDF 为

$$\begin{aligned} f_Z(z_1, z_2) &= f_{TR}(t_1, t_2) \cdot \mathcal{J} \\ &= \lambda^2 \nabla \kappa_0 \nabla \kappa_2 \exp\{-\lambda \nabla \kappa_1 z_1 - \lambda \nabla \kappa_2 (z_1 + z_2)\} \\ &= \lambda \nabla \kappa_0 \exp\{-\lambda \nabla \kappa_0 z_1\} \cdot \lambda \nabla \kappa_2 \exp\{-\lambda \nabla \kappa_2 z_2\} \\ &= (\varsigma_1 + \varsigma_2) \exp\{-(\varsigma_1 + \varsigma_2) z_1\} \cdot \varsigma_2 \exp\{-\varsigma_2 z_2\}. \end{aligned}$$

由此可见, 结论成立.  $\square$

根据定理 5.8, 可知  $\varsigma_1 Z_1$  和  $\varsigma_2 Z_2$  服从标准指数分布, 且相互独立. 设

$$z_{1j} = \Lambda(t_{1j}; \alpha), z_{2j} = \Lambda(t_{2j}; \alpha) - \Lambda(t_{1j}; \alpha) (j = 1, \dots, n).$$

记  $z_{i:1:n} < \dots < z_{i:n:n}$  为  $z_{ij}$  的升序排列 ( $j = 1, \dots, n, i = 1, 2$ ). 定义

$$\begin{aligned} W_{i1} &= nz_{i:1:n}, \quad W_{ij} = (n - j + 1)(z_{i:j:n} - z_{i:j-1:n}), \quad j = 2, \dots, n, \quad i = 1, 2. \\ Q_{ij} &= \sum_{k=1}^j W_{ik}, \quad j = 1, \dots, n, \quad U_{i:j} = Q_{ij}/Q_{in}, \quad j = 1, \dots, n - 1. \end{aligned} \quad (5.38)$$

根据引理 5.1、5.2 和定理 5.8 可知,  $U_{1:1} < \dots < U_{1:n-1}$  和  $U_{2:1} < \dots < U_{2:n-1}$  是两组从均匀分布  $(0, 1)$  抽取的样本容量为  $n - 1$  的顺序统计量, 且相互独立. 构造如下枢轴量 (Casella 等, 2021).

$$M(\alpha) = \sum_{j=1}^{n-1} (-2 \log U_{1:j}) + \sum_{j=1}^{n-1} (-2 \log U_{2:j}) = M_1(\alpha) + M_2(\alpha). \quad (5.39)$$

可证明  $M_1(\alpha)$  和  $M_2(\alpha)$  是  $\alpha$  的严格增函数, 并且  $M_1(\alpha)$  和  $M_2(\alpha)$  独立且服从自由度为  $2(n - 1)$  的  $\chi^2$  分布. 因此,  $M(\alpha)$  具有自由度为  $4(n - 1)$  的  $\chi^2$  分布. 则参数  $\alpha$  的精确  $1 - \nu$  置信区间为

$$\left( M^{-1} \left[ \chi_{1-\nu/2}^2(4(n-1)) \right], M^{-1} \left[ \chi_{\nu/2}^2(4(n-1)) \right] \right), \quad (5.40)$$

其中  $\chi_{\nu}^2(m)$  是自由度为  $m$  的  $\chi^2$  分布的上  $\nu$  百分位数,  $M^{-1}(x)$  是关于  $\alpha$  方程  $M(\alpha) = x$  的唯一解 ( $x > 0$ ). 对于参数  $\varsigma_1$ , 已知  $E_1 = 2\varsigma_1 \sum_{j=1}^n z_{1j} = 2\varsigma_1 \sum_{j=1}^n \Lambda(t_{1j}; \alpha)$  服从自由度为  $2n$  的  $\chi^2$  分布, 因此有

$$\varsigma_1 = \frac{E_1}{2 \sum_{j=1}^n \Lambda(t_{1j}; \alpha)}. \quad (5.41)$$

令  $g(M, T)$  为  $M(\alpha) = x$  的唯一解, 结合替代法 (Weerahandi, 1993), 将  $\alpha$  替换为  $g(M, T)$  后, 得到参数  $\varsigma_1$  的广义枢轴量

$$P_1 = \frac{E_1}{2 \sum_{j=1}^n \Lambda(t_{1j}; g(M, t))}, \quad (5.42)$$

其中  $t = \{(t_{1j}, t_{2j}), j = 1, \dots, n\}$  为  $T = \{(T_{1j}, T_{2j}), j = 1, \dots, n\}$  的观测值. 显然,  $P_1$  的分布不依赖于未知参数. 类似地, 对于参数  $\varsigma_2$ , 已知  $E_2 =$

$2\varsigma_2 \sum_{j=1}^n z_{2j} = 2\varsigma_2 \sum_{j=1}^n [\Lambda(t_{2j}; \alpha) - \Lambda(t_{1j}; \alpha)]$  服从自由度为  $2n$  的  $\chi^2$  分布, 因此有

$$\varsigma_2 = \frac{E_2}{2 \sum_{j=1}^n [\Lambda(t_{2j}; \alpha) - \Lambda(t_{1j}; \alpha)]}. \quad (5.43)$$

使用替代方法, 得到  $\varsigma_2$  的广义枢轴量

$$P_2 = \frac{E_2}{2 \sum_{j=1}^n [\Lambda(t_{2j}; g(M, t)) - \Lambda(t_{1j}; g(M, t))]} \quad (5.44)$$

同样可知  $P_2$  的分布不依赖于任何未知参数.

对于  $\alpha, \varsigma_1$  和  $\varsigma_2$  的任意连续函数  $J(\alpha, \varsigma_1, \varsigma_2)$ , 可构造其广义枢轴量

$$P_J = J(g(M, t), P_1, P_2), \quad (5.45)$$

其分布亦不依赖于未知参数. 因此, 参数  $\varsigma_1, \varsigma_2$  和  $J(\alpha, \varsigma_1, \varsigma_2)$  的近似  $(1 - \nu)$  CI 可以通过以下步骤计算获得.

### $\varsigma_1, \varsigma_2$ 和 $J(\alpha, \varsigma_1, \varsigma_2)$ 的广义置信区间计算

- 生成随机样本:** 独立生成随机变量  $M \sim \chi^2(4(n - 1))$ ,  $E_1 \sim \chi^2(2n)$  和  $E_2 \sim \chi^2(2n)$ .
- 计算广义枢轴量:** 根据式 (5.42)、(5.44) 和 (5.45) 分别计算  $P_1, P_2$  和  $P_J$ .
- 重复采样:** 重复步骤 1 和 2 共  $N$  次, 得到  $P_1, P_2, P_J$  的  $N$  个样本值, 分别记为  $P_{1,1}, \dots, P_{1,N}, P_{2,1}, \dots, P_{2,N}$  和  $P_{J,1}, \dots, P_{J,N}$ .
- 计算置信区间:** 对每个枢轴量  $P_i$  ( $i = 1, 2, J$ ), 计算样本的上  $\nu/2$  和  $1 - \nu/2$  百分位数, 分别记为  $P_{i,\nu/2}$  和  $P_{i,1-\nu/2}$ . 相应的近似  $1 - \nu$  置信区间为  $[P_{i,\nu/2}, P_{i,1-\nu/2}]$ .

### 5.3.3 模拟实验

假设系统包含两个部件, 随机环境的影响由 ED 过程  $\mathcal{ED}(\mu\Lambda(t; \alpha), 1)$  描述, 其中  $\Lambda(t; \alpha) = t^\alpha$ , 模型参数  $d$  取值 1.5、2 和 3, 分别对应复合泊松过程、伽马过程和逆高斯过程. 不失一般性, 设  $\mu = 1$ . 参数  $(\varsigma_1, \varsigma_2)$  分别为  $(1.000, 0.667)$ 、 $(1.098, 0.693)$  和  $(1.236, 0.732)$ . 参数  $\alpha$  设置为 0.8 和 3, 分别对应部件失效率递减和递增的情形. 样本量  $n$  取为 20、30 和 50.

对于每个  $(n, d, \alpha)$  的组合, 生成 2000 个样本, 然后对每个样本用所提出的方法估计模型参数以及两个部件的 0.05 和 0.95 分位寿命 (记作  $t_{1,0.05}, t_{1,0.95}, t_{2,0.05}, t_{2,0.95}$ ), 以此计算估计的 RB 和 RMSE. 模拟结果如表 5.13-5.14 所示. 结果表明: 当  $\alpha = 3$  时, 在不同样本量下参数估计的 RB 和 RMSE 都非常小. 然而, 当  $\alpha = 0.8$  且  $n \leq 30$  时, 0.05 分位寿命的 RB 相对较大 (10%-20%). 这是因为 0.05 分位寿命的真实值较小, 参数估计的偏差会对寿命分位数估计产生较大影响. 尽管如此, 从 RMSE 的角度来看, 估计性能仍可以接受, 当  $n = 50$  时, 0.05 分位寿命的 RB 均小于 10%.

为评估广义枢轴量方法在构建模型参数和分位寿命置信区间方面的表现, 设置  $N = 10000$ , 通过经验分位数获得 95% 的置信区间. 基于 2000 次重复模拟, 覆盖概率列于表 5.15. 从表 5.15 中可发现, 在不同参数和样本量设置下, 覆盖概率始终接近名义水平 95%. 这表明所提出的区间估计方法具有良好性能, 可为部件可靠性提供准确的不确定性量化.

### 5.3.4 实例分析

本节分析两个实例: 一是刹车片寿命数据, 二是欧洲冠军联赛 (UEFA Champion's League) 数据集. 在这两个案例中, 对  $\Lambda(t; \alpha)$  采用三种函数形式, 即  $t^\alpha$ 、 $\exp(\alpha t) - 1$  和  $\log(1 + t/\alpha)$ . 依据 AIC 和贝叶斯信息准则 (Bayesian information criterion, BIC) 选择合适的模型, 即,  $\Lambda(t; \alpha)$  的函数形式.

#### 5.3.4.1 刹车片寿命数据

首先分析前后刹车片寿命的相关性, 与传统皮尔逊相关系数不同, 采用斯皮尔曼秩相关系数来表征其非线性相关性. 该数据集的斯皮尔曼秩相关系数为  $\rho = 0.336$ , 且相关性检验  $p$  值为 0.048, 表明在显著性水平 0.05 下前后刹车片寿命之间存在相关性. 接下来, 在模型 (5.34) 中采用  $\Lambda(t; \alpha) = t^\alpha$ .

表 5.13: 不同样本量和模型下估计参数的 RB.

$\alpha = 0.8$									
模型	n	参数							
		$\alpha$	$\varsigma_1$	$\varsigma_2$	$\mu$	d	$t_{1,0.05}$	$t_{1,0.95}$	$t_{2,0.05}$
$d = 1.5$	20	4.836	1.154	5.533	2.613	0.308	16.874	7.102	23.153
	30	3.562	0.126	4.486	2.243	0.935	12.854	4.475	17.946
	50	1.431	0.234	2.054	1.389	0.702	5.648	2.197	9.575
$d = 2$	20	5.154	1.126	5.644	2.737	0.928	15.520	7.214	23.650
	30	3.582	0.452	4.311	1.931	1.306	12.865	4.431	17.905
	50	1.634	0.246	2.253	1.312	1.865	4.642	2.867	9.617
$d = 3$	20	5.207	1.379	5.016	5.684	5.046	15.688	7.351	23.363
	30	3.583	0.754	3.968	5.137	4.556	12.772	4.396	17.502
	50	1.694	0.430	1.953	2.102	1.318	5.673	2.378	9.847

$\alpha = 3$									
模型	n	参数							
		$\alpha$	$\varsigma_1$	$\varsigma_2$	$\mu$	d	$t_{1,0.05}$	$t_{1,0.95}$	$t_{2,0.05}$
$d = 1.5$	20	4.924	1.179	5.754	2.263	0.313	4.491	1.951	6.112
	30	3.528	0.304	4.291	2.986	0.905	3.497	1.138	3.898
	50	1.711	0.438	2.151	1.551	0.763	1.534	0.637	2.539
$d = 2$	20	4.975	1.462	5.557	3.083	0.796	4.477	1.934	6.054
	30	3.571	0.748	4.105	2.442	1.224	3.565	1.147	3.981
	50	1.710	0.497	1.983	1.818	1.842	1.517	0.641	2.564
$d = 3$	20	5.026	1.718	4.917	4.107	5.046	4.388	1.931	5.897
	30	3.584	0.991	3.798	3.638	4.296	3.605	1.135	4.057
	50	1.704	0.543	1.844	1.012	1.211	1.566	0.656	2.576

表 5.14: 不同样本量和模型下估计参数的 RMSE.

$\alpha = 0.8$									
模型	n	参数							
		$\alpha$	$\varsigma_1$	$\varsigma_2$	$\mu$	d	$t_{1,0.05}$	$t_{1,0.95}$	$t_{2,0.05}$
$d = 1.5$	20	0.133	0.276	0.224	0.247	1.856	0.023	1.099	0.053
	30	0.107	0.205	0.181	0.187	1.431	0.018	0.865	0.039
	50	0.074	0.149	0.128	0.138	0.893	0.012	0.775	0.025
$d = 2$	20	0.135	0.296	0.216	0.268	2.152	0.022	1.116	0.049
	30	0.106	0.224	0.185	0.232	1.823	0.016	0.857	0.037
	50	0.078	0.167	0.130	0.176	1.396	0.011	0.761	0.024
$d = 3$	20	0.136	0.336	0.229	0.293	2.330	0.018	1.105	0.048
	30	0.106	0.240	0.181	0.261	1.930	0.013	0.740	0.036
	50	0.079	0.183	0.132	0.186	1.813	0.009	0.769	0.023

$\alpha = 3$									
模型	n	参数							
		$\alpha$	$\varsigma_1$	$\varsigma_2$	$\mu$	d	$t_{1,0.05}$	$t_{1,0.95}$	$t_{2,0.05}$
$d = 1.5$	20	0.512	0.280	0.220	0.232	1.925	0.070	0.119	0.097
	30	0.408	0.204	0.183	0.193	1.425	0.057	0.095	0.079
	50	0.297	0.156	0.141	0.135	0.940	0.042	0.077	0.060
$d = 2$	20	0.511	0.309	0.224	0.245	2.375	0.068	0.115	0.095
	30	0.407	0.223	0.187	0.207	1.985	0.055	0.092	0.078
	50	0.297	0.171	0.144	0.183	1.524	0.041	0.075	0.059
$d = 3$	20	0.510	0.351	0.231	0.261	2.822	0.065	0.111	0.093
	30	0.407	0.251	0.193	0.234	2.515	0.053	0.089	0.076
	50	0.296	0.192	0.148	0.208	2.413	0.039	0.072	0.057

表 5.15: 不同样本量和模型下估计的覆盖概率 (%).

		$\alpha = 0.8$								
模型	n	参数								
		$\alpha$	$\varsigma_1$	$\varsigma_2$	$\mu$	d	$t_{1,0.05}$	$t_{1,0.95}$	$t_{2,0.05}$	$t_{2,0.95}$
$d = 1.5$	20	95.45	94.35	95.15	96.05	95.55	95.65	94.45	95.35	94.65
	30	94.70	95.30	94.95	95.90	94.65	95.05	94.90	94.85	95.05
	50	95.15	94.90	95.05	95.50	94.80	95.05	95.05	94.95	95.00
$d = 2$	20	95.35	94.50	95.25	96.15	95.95	95.60	94.70	95.35	94.55
	30	94.70	95.30	94.95	95.95	94.65	95.05	94.90	95.10	95.15
	50	94.85	94.80	94.95	95.05	95.20	94.90	95.20	95.10	95.05
$d = 3$	20	95.30	94.55	95.35	95.40	95.40	95.50	94.75	95.35	94.50
	30	94.70	95.35	95.00	95.20	94.65	95.10	95.00	94.75	95.05
	50	94.85	94.80	95.05	94.95	94.85	95.15	94.95	94.85	95.05
$\alpha = 3$										
模型	n	参数								
		$\alpha$	$\varsigma_1$	$\varsigma_2$	$\mu$	d	$t_{1,0.05}$	$t_{1,0.95}$	$t_{2,0.05}$	$t_{2,0.95}$
$d = 1.5$	20	95.15	94.35	95.25	96.45	96.65	95.85	94.80	95.35	94.60
	30	94.45	94.60	94.50	95.60	95.70	94.30	94.65	94.40	93.80
	50	94.45	95.05	95.00	94.55	93.90	94.60	94.20	94.20	95.50
$d = 2$	20	95.15	94.50	95.20	96.15	96.90	95.90	94.80	95.35	94.45
	30	94.45	94.85	94.45	95.65	95.85	94.30	94.70	94.35	93.90
	50	94.45	94.70	94.95	95.05	93.90	94.65	94.20	94.25	95.45
$d = 3$	20	95.10	94.45	95.45	95.65	96.15	95.90	94.75	95.30	94.55
	30	94.45	94.90	94.50	95.20	94.80	94.30	94.75	94.40	93.85
	50	94.45	94.60	95.15	94.85	93.85	94.60	94.25	94.30	95.65

表 5.16: 基于刹车片寿命数据, 参数的点估计和 95% 置信区间.

$\Lambda(t; \alpha)$	估计值	$\alpha$	$\mu$	$d$	$\varsigma_1$	$\varsigma_2 \times 100$	AIC	BIC
$t^\alpha$	均值	4.105	0.106	11.506	0.266	0.272		
	2.5%	3.372	0.063	6.381	0.158	0.069	<b>67.431</b>	<b>72.097</b>
	97.5%	4.928	0.163	19.362	0.407	0.870		
$\exp(\alpha t) - 1$	均值	1.067	0.132	15.184	0.325	0.928		
	2.5%	0.384	0.064	6.393	0.159	0.070	83.778	88.444
	97.5%	1.654	0.166	19.586	0.410	1.170		
$\log(1 + t/\alpha)$	均值	0.234	0.121	12.557	0.251	0.303		
	2.5%	0.084	0.067	5.273	0.152	0.074	121.120	125.786
	97.5%	0.369	0.172	18.569	0.398	0.883		

$\exp(\alpha t) - 1$  和  $\log(1 + t/\alpha)$  三种形式分别进行拟合. 针对每种  $\Lambda(t; \alpha)$  的形式, 计算了模型参数的点估计及 95% 置信区间, 结果列于表 5.16. 基于 AIC 和 BIC 准则, 可知当  $\Lambda(t; \alpha) = t^\alpha$  时, 模型具有最好的拟合效果.

在选定  $\Lambda(t; \alpha) = t^\alpha$  后, 采用极大似然法评估前后刹车片的可靠性. 例如, 图 5.9 展示了边际 CDF 估计与经验 CDF 的对比图, 可发现两者高度一致. 这验证了模型和估计方法的有效性. 估计的可靠性信息在许多后续决策活动中具有重要价值, 包括新车的可靠性预测、保修成本预测、维护活动规划以及设计变更影响的评估等 (Chen 等, 2017; Hu 等, 2020).

### 5.3.4.2 欧洲冠军联赛足球数据

第二个实例来自 Meintanis (2007) 的欧洲冠军联赛数据集, 其中包含两个变量: 第一个变量  $T_1$  表示主队打进首个进球 (任何类型) 的时间 (分钟), 第二个变量  $T_2$  表示任一球队通过任意球打进首球的时间. 原始数据包含 37 个样本. 在该数据集中, 存在  $T_1 < T_2$ 、 $T_1 > T_2$  和  $T_1 = T_2$  三种情况, 且每种情况的发生均具有非零概率. 因此, 可以使用联合 PDF (5.31) 进行建模分析. 为统一单位 (小时) 及缩小模型参数的量级, 对所有数据点进行了量纲处理 (除以 60). 基于完整数据, 分别计算了  $\Lambda(t; \alpha) = t^\alpha$ 、 $\exp(\alpha t) - 1$  和  $\log(1 + t/\alpha)$  对应模型的 AIC 和 BIC 值. 结果显示,  $\Lambda(t; \alpha) = \exp(\alpha t) - 1$  的模型具有最小的 AIC 和 BIC 值, 分别为 91.337 和 97.781.

表 5.17: 基于联赛足球数据, 参数的点估计及 95% 置信区间.

$\Lambda(t; \alpha)$	估计类型	$\alpha$	$\mu$	$d$	$\varsigma_1$	$\varsigma_2$	AIC	BIC
$t^\alpha$	均值	1.567	8.351	11.697	3.450	2.085		
	2.5%	1.040	2.051	3.634	1.893	1.213	11.00	13.50
	97.5%	2.164	12.223	19.818	5.542	3.223		
$\exp(\alpha t) - 1$	均值	1.694	0.540	2.297	0.783	0.449		
	2.5%	1.045	0.055	1.329	0.468	0.203	<b>8.20</b>	<b>10.70</b>
	97.5%	2.147	1.154	2.948	1.475	0.867		
$\log(1 + t/\alpha)$	均值	1.673	3.156	2.214	3.289	1.886		
	2.5%	1.142	2.042	1.351	1.788	1.222	22.2	24.70
	97.5%	2.679	4.861	3.460	5.312	3.151		

为进一步验证所提模型有效性, 筛选出满足  $T_1 < T_2$  的数据子集进行分析, 子集包含 17 个观测值. 首先检查两个变量之间的相关性. 基于该子集计算的斯皮尔曼秩相关系数为  $\rho = 0.556$ , 其相关性检验  $p$  值为 0.021, 表明在显著性水平 0.05 下,  $T_1$  和  $T_2$  之间具有显著相关性. 随后, 使用所提模型对该子集进行分析. 通过广义推断方法, 获得了参数的点估计与区间估计, 估计结果以及对应的 AIC 和 BIC 值列于表 5.17. 结果表明, 当  $\Lambda(t; \alpha) = \exp(\alpha t) - 1$  时, 模型具有最优的拟合效果, 与基于完整数据的结果一致. 动态环境的影响通过 ED 过程  $\mathcal{ED}(0.54[\exp(1.695t) - 1], 1)$  描述, 其中  $d = 2.297$ . 由于  $d = 2$  (对应于伽马过程) 落在  $d$  的 95% 置信区间 (1.329, 2.948) 内, 可认为伽马过程能够合理反映动态环境的影响. 图 5.10 展示了  $T_1$  和  $T_2$  的边际 CDF 估计及其 95% 置信区间. 结果显示, 估计的 CDF 与经验 CDF 高度吻合, 进一步验证了模型的有效性.

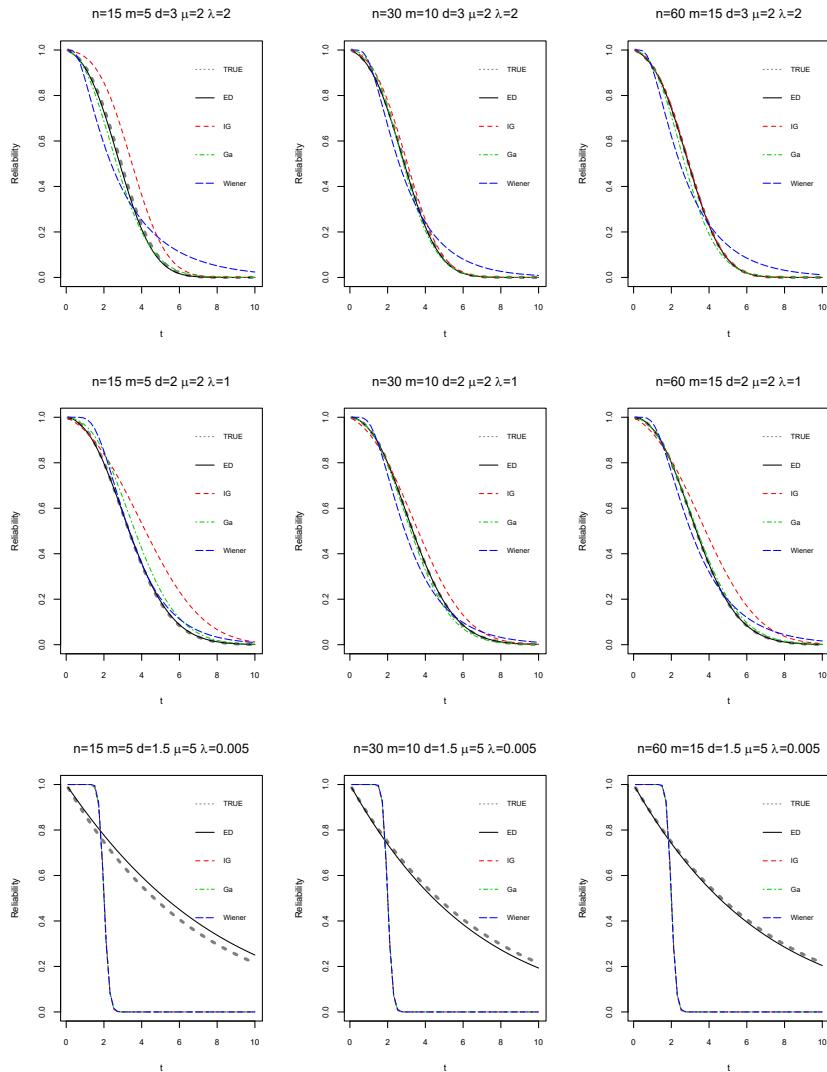


图 5.2: 不同模型和模拟设置下的拟合可靠性函数.

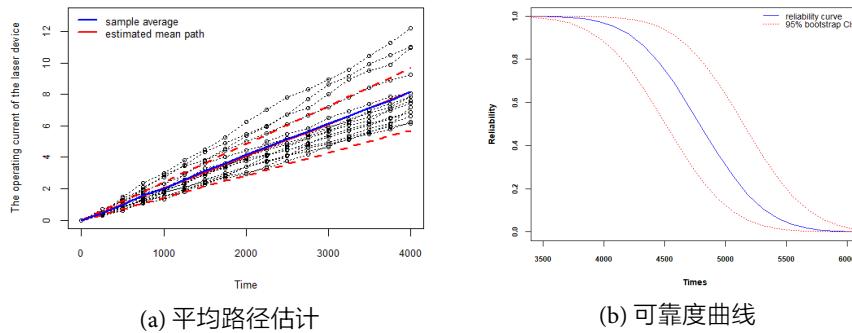


图 5.3: 基于激光退化数据, 平均路径和可靠度曲线估计.

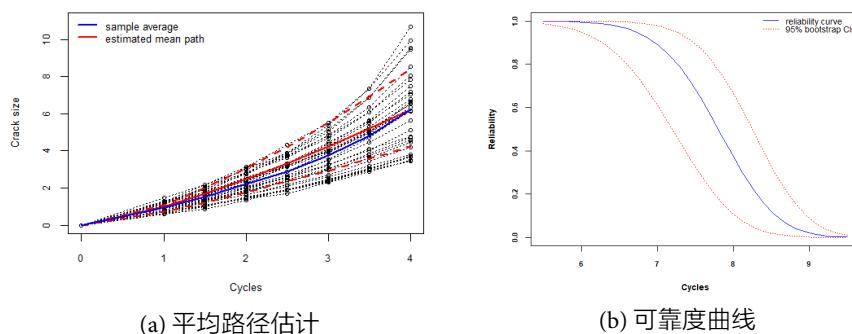


图 5.4: 基于铝合金裂纹退化数据, 平均路径和可靠度曲线估计.

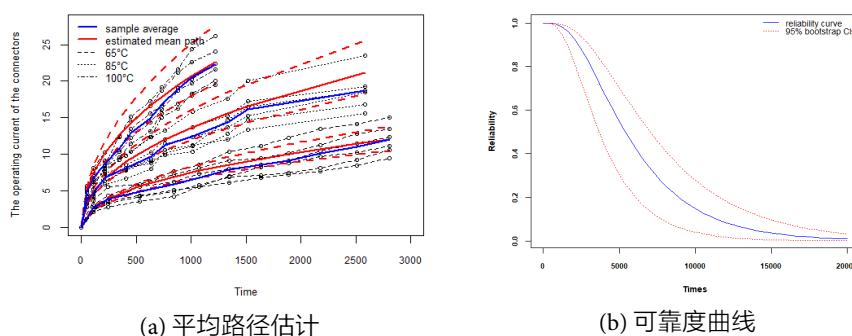


图 5.5: 基于应力松弛数据, 平均路径和可靠度曲线估计.

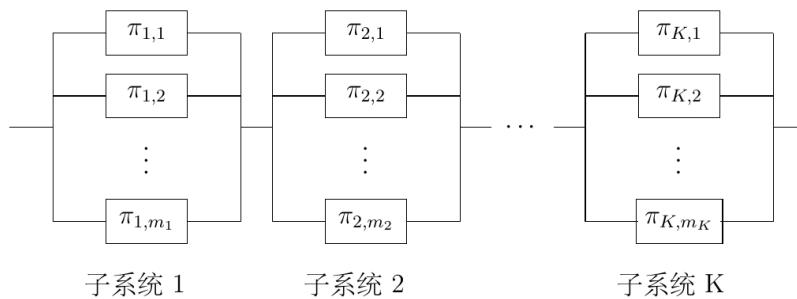


图 5.6: 一般串并联系统.

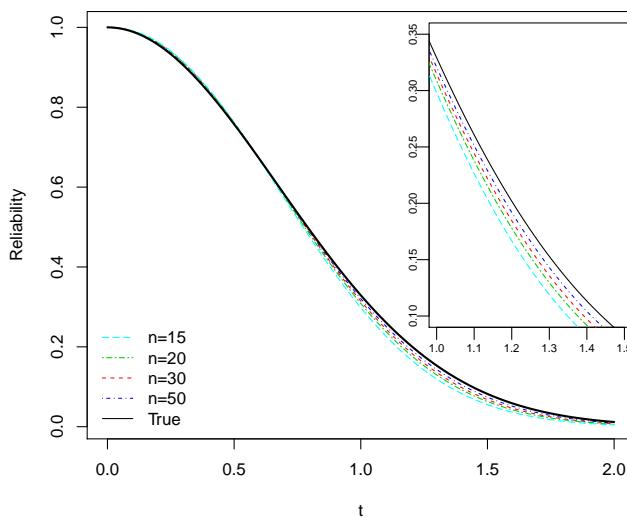


图 5.7: 基于复合泊松过程的系统可靠性函数估计.

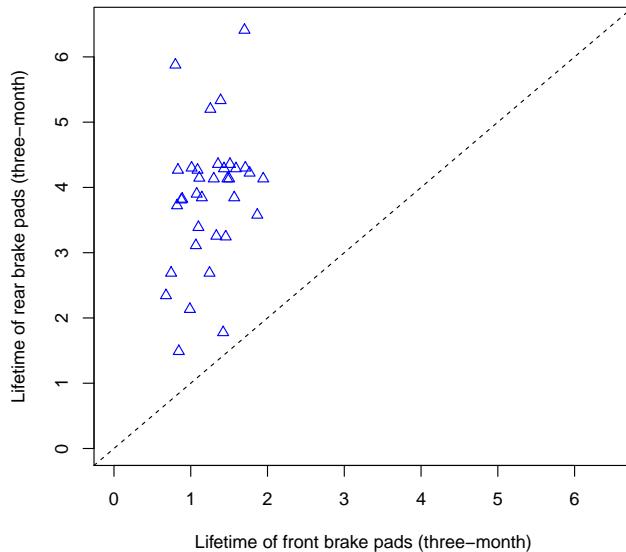


图 5.8: 后刹车片寿命随前刹车片寿命变化的函数关系.

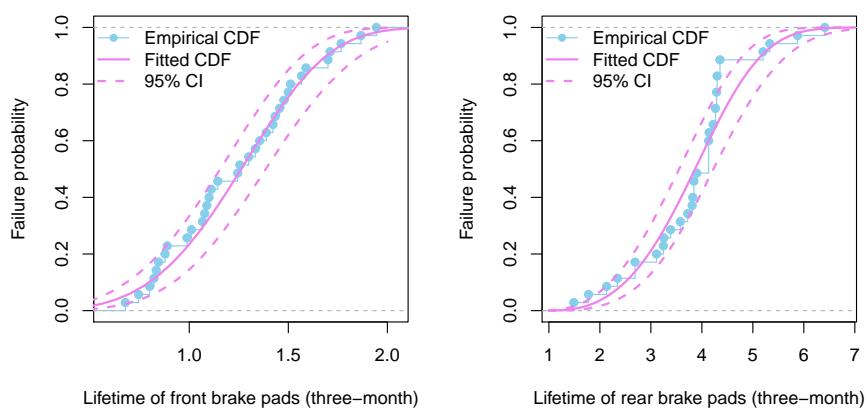


图 5.9: 基于刹车片寿命数据, 边际 CDF 的拟合结果.

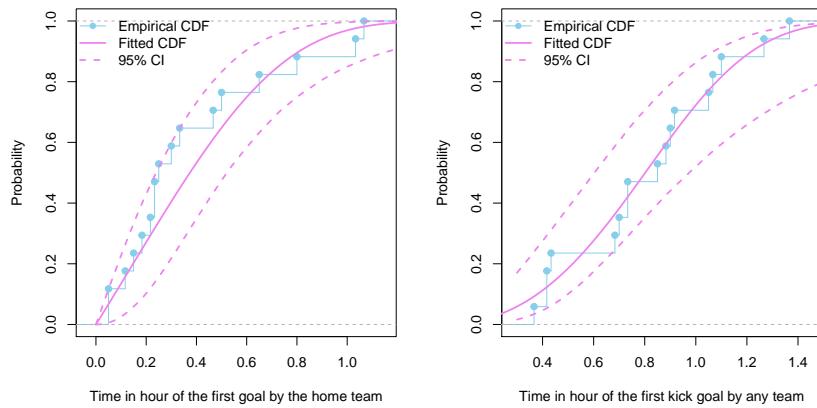


图 5.10: 基于联赛足球数据, 边际 CDF 的拟合结果.

# 第六章 总结与展望

在本书中, 我们深入探讨了不同随机退化过程的统计建模方法, 重点介绍了维纳过程、伽玛过程、逆高斯过程和指数分散过程等在可靠性分析中的应用. 这些方法为产品和系统的退化建模提供了科学的理论基础与实践指导, 能够有效预测和评估产品的寿命和可靠性, 帮助工程师识别潜在的风险并采取有效的预防措施.

通过对各类随机过程的详细分析, 本书阐述了不同建模方法的优势与适用场景, 并展示了如何在实际工作中结合具体产品的特性选择合适的建模工具. 我们特别强调了这些方法和模型在刻画不同性能退化特征、RUL 预测以及系统可靠性分析中的应用, 旨在为可靠性工程师和研究人员提供更加全面、系统的建模框架.

尽管本书提供了多种退化过程建模的方法论, 但在实际应用中仍面临诸多挑战, 比如, 如何应对复杂系统中的多维退化数据建模问题、如何提高建模过程的计算效率等. 未来的研究可以进一步深化这些建模方法的理论基础, 优化现有模型, 提高计算效率和预测精度.

此外, 随着技术的不断进步, 尤其是人工智能、大数据和机器学习技术的发展, 退化过程建模的研究也可以借助这些新兴技术来拓展. 结合数据驱动的模型与传统的物理建模方法, 将有助于提升产品质量管理和系统可靠性评估的准确性. 未来, 我们可以期待更多先进的建模方法和工具在可靠性工程中的应用, 推动产品生命周期管理和故障预防机制的进一步优化.

## 参考文献

Abdel Hameed M, 1975. A gamma wear process[J]. IEEE Transactions on Reliability, 24(2): 152–153.

Al-Hussaini E K, Abd-El-Hakim N S, 1989. Failure rate of the inverse Gaussian-Weibull mixture model[J]. Annals of the Institute of Statistical Mathematics, 41(3): 617–622.

Bae S J, Kim S-J, Kim M S, 等, 2008. Degradation analysis of nano-contamination in plasma display panels[J]. IEEE Transactions on Reliability, 57(2): 222–229.

Bae S J, Kvam P H, 2006. A change-point analysis for modeling incomplete burn-in for light displays[J]. IIE Transactions, 38(6): 489–498. DOI:[10.1080/074081791009068](https://doi.org/10.1080/074081791009068).

Bae S J, Yuan T, Kim S, 2016. Bayesian degradation modeling for reliability prediction of organic light-emitting diodes[J]. Journal of Computational Science, 17: 117–125. DOI:[10.1016/j.jocs.2016.08.006](https://doi.org/10.1016/j.jocs.2016.08.006).

Bae S J, Yuan T, Ning S, 等, 2015. A Bayesian approach to modeling two-phase degradation using change-point regression[J]. Reliability Engineering & System Safety, 134: 66–74.

Barndorff-Nielsen O E, Koudou A E, 1998. Trees with random conductivities and the (Reciprocal) inverse Gaussian distribution[J]. Advances in Applied Probability, 30(2): 409–424.

Bemis B M, Bain L J, Higgins J J, 1972. Estimation and hypothesis testing for the parameters of a bivariate exponential distribution[J]. Journal of the American Statistical Association, 67(340): 927–929.

Berger J O, 2013. Statistical decision theory and Bayesian analysis[M]. Springer Science & Business Media.

Bernardo J M, Smith A F, 2009. Bayesian Theory[M]. John Wiley & Sons.

Blei D M, Kucukelbir A, McAuliffe J D, 2017. Variational inference: A review for statisticians[J]. *Journal of the American Statistical Association*, 112(518): 859–877.

Bolstad W M, Curran J M, 2016. Introduction to Bayesian statistics[M]. John Wiley & Sons.

Borchers H W, 2019. Pracma: practical numerical math functions[J]. R package version, 2(9): 519.

Bousquet N, Fouladirad M, Grall A, 等, 2015. Bayesian gamma processes for optimizing condition-based maintenance under uncertainty[J]. *Applied Stochastic Models in Business and Industry*, 31(3): 360–379.

Casella G, Berger R L, 2021. Statistical Inference[M]. Cengage Learning.

Celeux G, Hurn M, Robert C P, 2000. Computational and inferential difficulties with mixture posterior distributions[J]. *Journal of the American Statistical Association*, 95(451): 957–970.

Chao M A, Kulkarni C, Goebel K, 等, 2022. Fusing physics-based and deep learning models for prognostics[J]. *Reliability Engineering & System Safety*, 217: 107961.

Chen N, Tsui K L, 2013. Condition monitoring and remaining useful life prediction using degradation signals: revisited[J]. *IIE Transactions*, 45(9): 939–952. DOI:[10.1080/0740817X.2012.706376](https://doi.org/10.1080/0740817X.2012.706376).

Chen Z, Xia T, Li Y, 等, 2021. Random-effect models for degradation analysis based on nonlinear Tweedie exponential-dispersion processes[J]. *IEEE Transactions on Reliability*, 71(1): 47–62.

Chen Z, Xia T, Li Y, 等, 2022. Random-Effect Models for Degradation Analysis Based on Nonlinear Tweedie Exponential-Dispersion Processes[J]. *IEEE Transactions on Reliability*, 71(1): 47–62.

Chen P, Ye Z-S, 2017. Estimation of field reliability based on aggregate lifetime data[J]. *Technometrics*, 59(1): 115–125.

Chen P, Ye Z-S, 2018a. A systematic look at the gamma process capability indices[J]. *European Journal of Operational Research*, 265(2): 589–597.

Chen P, Ye Z-S, 2018b. Uncertainty quantification for monotone stochastic

degradation models[J]. *Journal of Quality Technology*, 50(2): 207–219.

Csorgo S, Welsh A. H., 1989. Testing for exponential and Marshall-Olkin distributions[J]. *Journal of Statistical Planning and Inference*, 23: 287–300.

Damsleth E, 1975. Conjugate classes for gamma distributions[J]. *Scandinavian Journal of Statistics*, 2(2): 80–84.

Davis P J, Rabinowitz P, 2007. *Methods of numerical integration*[M]. Courier Corporation.

Denison D G, Holmes C C, Mallick B K, 等, 2002. *Bayesian methods for nonlinear classification and regression*[M]. John Wiley & Sons.

Doksum K A, Hbyland A, 1992. Models for variable-stress accelerated life testing experiments based on Wiener processes and the inverse Gaussian distribution[J]. *Technometrics*, 34(1): 74–82.

Duan F, Wang G, 2017. Reliability modeling of two-phase inverse Gaussian degradation process[C]// 2017 Second International Conference on Reliability Systems Engineering (ICRSE). IEEE: 1–6.

Duan F, Wang G, 2018. Exponential-dispersion degradation process models with random effects and covariates[J]. *IEEE Transactions on Reliability*, 67(3): 1128–1142.

Dunn P K, Smyth G K, 2005. Series evaluation of Tweedie exponential dispersion model densities[J]. *Statistics and Computing*, 15(4): 267–280.

Efron B, 2012. Bayesian inference and the parametric bootstrap[J]. *The Annals of Applied Statistics*, 6(4): 1971–1997. DOI:[10.1214/12-AOAS571](https://doi.org/10.1214/12-AOAS571).

Efron B, Tibshirani R, 1993. *An introduction to the bootstrap*[M]. Chapman & Hall: 49–54.

Erişoğlu Ü, Erişoğlu M, Erol H, 2011. A mixture model of two different distributions approach to the analysis of heterogeneous survival data[J]. *International Journal of Computational and Mathematical Sciences*, 5(6): 544–548.

Fan T-H, Chen C-H, 2017. A Bayesian predictive Analysis of step-stress accelerated tests in Gamma degradation-based processes[J]. *Quality and Reliability Engineering International*.

Fan T-H, Dong Y-S, Peng C-Y, 2024. A complete Bayesian degradation analysis based

on inverse Gaussian processes[J]. *IEEE Transactions on Reliability*, 73(1): 536–548. DOI:[10.1109/TR.2023.3304673](https://doi.org/10.1109/TR.2023.3304673).

Fang G, Pan R, Wang Y, 2022. Inverse Gaussian processes with correlated random effects for multivariate degradation modeling[J]. *European Journal of Operational Research*, 300(3): 1177–1193.

Fan T, Hsu T, 2015. Statistical inference of a two-component series system with correlated log-normal lifetime distribution under multiple type-I censoring[J]. *IEEE Transactions on Reliability*, 64(1): 376–385.

Fouladirad M, Grall A, 2011. Condition-based maintenance for a system subject to a non-homogeneous wear process with a wear rate transition[J]. *Reliability Engineering & System Safety*, 96(6): 611–618. DOI:[10.1016/j.ress.2010.12.008](https://doi.org/10.1016/j.ress.2010.12.008).

Fouladirad M, Grall A, Dieulle L, 2008. On the use of on-line detection for maintenance of gradually deteriorating systems[J]. *Reliability Engineering & System Safety*, 93(12): 1814–1820. DOI:[10.1016/j.ress.2008.03.020](https://doi.org/10.1016/j.ress.2008.03.020).

Geisser S, Hodges J, Press S, 等, 1990. The validity of posterior expansions based on Laplace's method[J]. *Bayesian and likelihood methods in statistics and econometrics*, 7: 473.

Gelman A, 2006. Prior distributions for variance parameters in hierarchical models[J]. *Bayesian Analysis*, 1(3): 515–533.

Gelman A, Carlin J B, Stern H S, 等, 2014. *Bayesian Data Analysis*[M]. Taylor & Francis.

Gilks W R, Best N G, Tan K K C, 2022. Adaptive rejection Metropolis sampling within Gibbs sampling[J]. *Applied Statistics*, 44: 455–472. DOI:[10.2307/2986138](https://doi.org/10.2307/2986138).

Gilks W R, Wild P, 1992. Adaptive rejection sampling for Gibbs sampling[J]. *Journal of the Royal Statistical Society: Series C*, 41(2): 337–348.

Giorgio M, Mele A, Pulcini G, 2019. A perturbed gamma degradation process with degradation dependent non-Gaussian measurement errors[J]. *Applied Stochastic Models in Business and Industry*, 35(2): 198–210.

Grall A, Fouladirad M, 2008. Maintenance decision rule with embedded online Bayesian change detection for gradually non-stationary deteriorating systems[J]. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk*

and Reliability, 222(3): 359–369. DOI:[10.1243/1748006XJRR141](https://doi.org/10.1243/1748006XJRR141).

Guan Q, Tang Y, Xu A, 2016. Objective Bayesian analysis accelerated degradation test based on Wiener process models[J]. Applied Mathematical Modelling, 40(4): 2743–2755.

Guida M, Postiglione F, Pulcini G, 2019. A Bayesian approach for non-homogeneous gamma degradation processes[J]. Communications in Statistics-Theory and Methods, 48(7): 1748–1765.

Guo J, Wang C, Cabrera J, 等, 2018. Improved inverse Gaussian process and bootstrap: Degradation and reliability metrics[J]. Reliability Engineering & System Safety, 178: 269–277.

Guo C, Wang W, Guo B, 等, 2013. A maintenance optimization model for mission-oriented systems based on Wiener degradation[J]. Reliability Engineering & System Safety, 111: 183–194.

Hall P, Wang J Z, 2005. Bayesian likelihood methods for estimating the end point of a distribution[J]. Journal of the Royal Statistical Society: Series B, 67(5): 717–729.

Hao H, Su C, Li C, 2015. LED lighting system reliability modeling and inference via random effects gamma process and copula function[J]. International Journal of Photoenergy, 2015.

Hao S, Yang J, Berenguer C, 2019. Degradation analysis based on an extended inverse Gaussian process model with skew-normal random effects and measurement errors[J]. Reliability Engineering & System Safety, 189: 261–270. DOI:[10.1016/j.ress.2019.04.031](https://doi.org/10.1016/j.ress.2019.04.031).

Hartzell A L, Da Silva M G, Shea H R, 2011. Lifetime Prediction[M]// MEMS Reliability. Springer, New York: 9–42.

Hazra I, Pandey M D, Manzana N, 2020. Approximate Bayesian computation method for estimating parameters of the gamma process using noisy data[J]. Reliability Engineering & System Safety, 198: 106780.

Hong Y, Duan Y, Meeker W Q, 等, 2015. Statistical methods for degradation data with dynamic covariates information and an application to outdoor weathering data[J]. Technometrics, 57(2): 180–193.

Hong L, Ye Z-S, 2017. When is acceleration unnecessary in a degradation test?[J].

Statistica Sinica, 27(3): 1461–1483.

Hong L, Zhai Q, Wang X, 等, 2019. System reliability evaluation under dynamic operating conditions[J]. IEEE Transactions on Reliability, 68(3): 800–809.

Hougaard P, 1995. Frailty models for survival data[J]. Lifetime Data Analysis, 1(3): 255–273.

Hu J, Chen P, 2020. Predictive maintenance of systems subject to hard failure based on proportional hazards model[J]. Reliability Engineering & System Safety, 196: 106707.

Huynh K T, 2021. An adaptive predictive maintenance model for repairable deteriorating systems using inverse Gaussian degradation process[J]. Reliability Engineering & System Safety, 213: 107695.

Jaakkola T S, Jordan M I, 2000. Bayesian parameter estimation via variational methods[J]. Statistics and Computing, 10(1): 25–37.

Jamshidian M, Jennrich R I, 1997. Acceleration of the EM algorithm by using quasi-Newton methods[J]. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 59(3): 569–587. DOI:[10.1111/1467-9868.00083](https://doi.org/10.1111/1467-9868.00083).

Jørgensen B, 1986. Some properties of exponential dispersion models[J]. Scandinavian Journal of Statistics, 13(3): 187–197.

Jørgensen B, 1987. Exponential dispersion models[J]. Journal of the Royal Statistical Society, 49(2): 127–162.

Jørgensen B, 1992. Exponential dispersion models and extensions: a review[J]. International Statistical Review, 60(1): 5–20.

Jørgensen B, 1998. The theory of exponential dispersion models[M]. Chapman & Hall.

Komárek A, Lesaffre E, 2008. Generalized linear mixed model with a penalized Gaussian mixture as a random effects distribution[J]. Computational Statistics & Data Analysis, 52(7): 3441–3458.

Kong D J, Balakrishnan N, Cui L R, 2017. Two-phase degradation process model with abrupt jump at change point governed by Wiener process[J]. IEEE Transactions on Reliability, 66(4): 1345–1360. DOI:[10.1109/Tr.2017.2711621](https://doi.org/10.1109/Tr.2017.2711621).

Kontar R, Son J, Zhou S, 等, 2017. Remaining useful life prediction based on the mixed effects model with mixture prior distribution[J]. IIE Transactions, 49(7): 682–

697.

Kotz S, Lai C D, Xie M, 2003. On the effect of redundancy for systems with dependent components[J]. *IIE Transactions*, 35(12): 1103–1110.

Krishnamoorthy K, 2016. *Handbook of statistical distributions with applications*[M]. CRC Press.

Kundu D, Gupta A K, 2013. Bayes estimation for the Marshall–Olkin bivariate Weibull distribution[J]. *Computational Statistics & Data Analysis*, 57(1): 271–281.

Kvalseth T O, 1985. Cautionary Note about  $R^2$ [J]. *The American Statistician*, 39(4): 279–285.

Lai C D, Xie M, 2006. *Stochastic Ageing and Dependence for Reliability*[M]. Springer.

Lawless J, Crowder M, 2004. Covariates and random effects in a gamma process model with application to degradation and failure[J]. *Lifetime Data Analysis*, 10: 213–227. DOI:[10.1023/B:LIDA.0000036389.14073.dd](https://doi.org/10.1023/B:LIDA.0000036389.14073.dd).

Lee I-C, Tseng S-T, Hong Y, 2020. Global planning of accelerated degradation tests based on exponential dispersion degradation models[J]. *Naval Research Logistics*, 67(6): 469–483.

Lesaffre E, Molenberghs G, 1991. Multivariate probit analysis: a neglected procedure in medical statistics[J]. *Statistics in Medicine*, 10(9): 1391–1403.

Lewis-Beck C, Tian Q, Meeker W Q, 2022. Prediction of future failures for heterogeneous reliability field data[J]. *Technometrics*, 64(1): 125–138.

Li M, Meng H, Zhang Q, 2017. A nonparametric Bayesian modeling approach for heterogeneous lifetime data with covariates[J]. *Reliability Engineering & System Safety*, 167: 95–104.

Lindstrom M J, Bates D M, 1988. Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data[J]. *Journal of the American Statistical Association*, 83: 1014–1022.

Ling M H, Ng H, Tsui K L, 2019. Bayesian and likelihood inferences on remaining useful life in two-phase degradation models under gamma process[J]. *Reliability Engineering & System Safety*, 184: 77–85.

Lin C P, Ling M H, Cabrera J, 等, 2021. Prognostics for lithium-ion batteries using a two-phase gamma degradation process model[J]. *Reliability Engineering &*

System Safety, 214: 107797. DOI:[10.1016/j.ress.2021.107797](https://doi.org/10.1016/j.ress.2021.107797).

Liu X, 2012. Planning of accelerated life tests with dependent failure modes based on a gamma frailty model[J]. *Technometrics*, 54(4): 398–409.

Liu Q, Pierce D A, 1994. A note on Gauss—Hermite quadrature[J]. *Biometrika*, 81(3): 624–629.

Liu X, Yeo K, Kalagnanam J, 2018. A statistical modeling approach for spatio-temporal degradation data[J]. *Journal of Quality Technology*, 50(2): 166–182.

Li J, Wang Z, Zhang Y, 等, 2017. Degradation data analysis based on a generalized Wiener process subject to measurement error[J]. *Mechanical Systems and Signal Processing*, 94: 57–72.

Li J, Wang Z, Zhang Y, 等, 2018. A nonlinear Wiener process degradation model with autoregressive errors[J]. *Reliability Engineering & System Safety*, 173: 48–57.

Louis T A, 1982. Finding the observed information matrix when using the EM Algorithm[J]. *Journal of the Royal Statistical Society, Series B*, 44: 226–233.

Lu B, Chen Z, Zhao X, 2022. Data-driven dynamic adaptive replacement policy for units subject to heterogeneous degradation[J]. *Computers & Industrial Engineering*, 171: 108478. DOI:[10.1016/j.cie.2022.108478](https://doi.org/10.1016/j.cie.2022.108478).

Lu C J, Meeker W O, 1993. Using degradation measures to estimate a time-to-failure distribution[J]. *Technometrics*, 35(2): 161–174.

Lu L, Wang B, Hong Y, 等, 2020. General path models for degradation data with multiple characteristics and covariates[J]. *Technometrics*, 63(3): 354–369.

Ma Y, Chen Y, Zhou X, 等, 2019. Remaining useful life prediction of lithium-ion battery based on Gauss—Hermite particle filter[J]. *IEEE Transactions on Control Systems Technology*, 27(4): 1788–1795.

Marin J-M, Mengersen K, Robert C P, 2005. Bayesian modelling and inference on mixtures of distributions[J]. *Handbook of Statistics*, 25: 459–507.

Marsaglia G, Tsang W W, Wang J, 2003. Evaluating Kolmogorov's distribution[J]. *Journal of Statistical Software*, 8: 1–4.

Marshall A W, Olkin I, 1967. A multivariate exponential distribution[J]. *Journal of the American Statistical Association*, 62(317): 30–44.

Ma Z, Wang S, Liao H, 等, 2019. Engineering-driven performance degradation

analysis of hydraulic piston pump based on the inverse Gaussian process[J]. Quality and Reliability Engineering International, 35(7): 2278–2296.

McLachlan G J, Krishnan T, 2007. The EM algorithm and extensions[M]. John Wiley & Sons.

Meeker W Q, Escobar L A, Pascual F G, 1998. Statistical Methods for Reliability Data[M]. John Wiley & Sons.

Meintanis S G, 2007. Test of fit for Marshall–Olkin distributions with applications[J]. Journal of Statistical Planning and Inference, 137: 3954–3963.

Miller K W, Morell L J, Noonan R E, 等, 1992. Estimating the probability of failure when testing reveals no failures[J]. IEEE Transactions on Software Engineering, 18(1): 33.

Nelson W, 1980. Accelerated life testing-step-stress models and data analyses[J]. IEEE Transactions on Reliability, 29(2): 103–108.

Nguyen K T, Medjaher K, Gogu C, 2022. Probabilistic deep learning methodology for uncertainty quantification of remaining useful lifetime of multi-component systems[J]. Reliability Engineering & System Safety, 222: 108383.

Ntzoufras I, 2011. Bayesian modeling using WinBUGS[M]. John Wiley & Sons.

Pan D, Liu J-B, Cao J, 2016. Remaining useful life estimation using an inverse Gaussian degradation model[J]. Neurocomputing, 185: 64–72.

Pan D, Liu J-B, Huang F, 等, 2017. A Wiener process model with truncated normal distribution for reliability analysis[J]. Applied Mathematical Modelling, 50: 333–346.

Park C, Padgett W, 2005. Accelerated degradation models for failure based on geometric Brownian motion and gamma processes[J]. Lifetime Data Analysis, 11(4): 511–527.

Park C, Padgett W J, 2006. Stochastic degradation models with several accelerating variables[J]. IEEE Transactions on Reliability, 55(2): 379–390. DOI:[10.1109/TR.2006.874937](https://doi.org/10.1109/TR.2006.874937).

Park C, Padgett W J, 2014. Cumulative damage models based on gamma processes[J]. Wiley StatsRef: Statistics Reference Online.

Paroissin C, 2015. Inference for the Wiener process with random initiation time[J].

IEEE Transactions on Reliability, 65(1): 147–157.

Paroissin C, 2017. Online estimation methods for the gamma degradation process[J]. IEEE Transactions on Reliability, 66(4): 1361–1367.

Peng C-Y, 2015. Inverse Gaussian processes with random effects and explanatory variables for degradation data[J]. Technometrics, 57(1): 100–111.

Peng C-Y, Cheng Y-S, 2020. Student-t processes for degradation analysis[J]. Technometrics, 62(2): 223–235.

Peng W, Chen Y, Xu A, 等, 2024. Collaborative online RUL prediction of multiple assets with analytically recursive Bayesian inference[J]. IEEE Transactions on Reliability, 73(1): 506–520.

Peng C-Y, Hsu S-C, 2012. A note on a Wiener process with measurement error[J]. Applied Mathematics Letters, 25(4): 729–732.

Peng W, Li Y-F, Yang Y-J, 等, 2014. Inverse Gaussian process models for degradation analysis: A Bayesian perspective[J]. Reliability Engineering & System Safety, 130: 175–189.

Peng W W, Li Y F, Yang Y J, 等, 2016. Bivariate analysis of incomplete degradation observations based on inverse Gaussian processes and copulas[J]. IEEE Transactions on Reliability, 65(2): 624–639.

Peng W, Li Y-F, Yang Y-J, 等, 2017. Bayesian degradation analysis with inverse Gaussian process models under time-varying degradation rates[J]. IEEE Transactions on Reliability, 66(1): 84–96. DOI:[10.1109/TR.2016.2635149](https://doi.org/10.1109/TR.2016.2635149).

Peng W, Shen L, Shen Y, 等, 2018. Reliability analysis of repairable systems with recurrent misuse-induced failures and normal-operation failures[J]. Reliability Engineering & System Safety, 171: 87–98.

Peng C-Y, Tseng S-T, 2009. Mis-specification analysis of linear degradation models[J]. IEEE Transactions on Reliability, 58(3): 444–455.

Peng Y, Wang Y, Zi Y, 等, 2017. Dynamic reliability assessment and prediction for repairable systems with interval-censored data[J]. Reliability Engineering & System Safety, 159: 301–309.

Peng Weiwen, Ye Z-S, Chen N, 2018. Joint online RUL prediction for multivariate deteriorating systems[J]. IEEE Transactions on Industrial Informatics, 15(5):

2870–2878. DOI:[10.1109/TII.2018.2869429](https://doi.org/10.1109/TII.2018.2869429).

Peng W, Zhu S-P, Shen L, 2019. The transformed inverse Gaussian process as an age-and state-dependent degradation model[J]. *Applied Mathematical Modelling*, 75: 837–852.

Polson N G, Scott J G, 2012. On the half-cauchy prior for a global scale parameter[J]. *Bayesian Analysis*, 7(4): 887–901.

Pulcini G, 2016. A perturbed gamma process with statistically dependent measurement errors[J]. *Reliability Engineering & System Safety*, 152: 296–306.

Ren L, Jia Z, Laili Y, 等, 2024. Deep learning for time-series prediction in IIoT: progress, challenges, and prospects[J]. *IEEE transactions on neural networks and learning systems*, 35(11): 15072–15091.

Robinson M E, Crowder M J, 2000. Bayesian methods for a growth-curve degradation model with repeated measures[J]. *Lifetime Data Analysis*, 6: 357–374.

Schrödinger E, 1915. Notiz über den Kapillardruck in Gasblasen[J]. *Annalen der Physik*, 351(3): 413–418.

Seo K, Pan R, 2017. Data analysis of step-stress accelerated life tests with heterogeneous group effects[J]. *IIE Transactions*, 49(9): 885–898.

Severson K A, Attia P M, Jin N, 等, 2019. Data-driven prediction of battery cycle life before capacity degradation[J]. *Nature Energy*, 4(5): 383–391.

Shaked M A, 1977. Concept of positive dependence for exchangeable random variables[J]. *Annals of Statistics*, 5(3): 505–515.

Shen L, Wang Y, Zhai Q, 等, 2018. Degradation modeling using stochastic processes with random initial degradation[J]. *IEEE Transactions on Reliability*, 68(4): 1320–1329. DOI:[10.1109/TR.2018.2885133](https://doi.org/10.1109/TR.2018.2885133).

Sheu S-H, Tsai H-N, Sheu U-Y, 等, 2019. Optimal replacement policies for a system based on a one-cycle criterion[J]. *Reliability Engineering & System Safety*, 191: 106527. DOI:[10.1016/j.ress.2019.106527](https://doi.org/10.1016/j.ress.2019.106527).

Shuster J, 1968. On the inverse Gaussian distribution function[J]. *Journal of the American Statistical Association*, 63(324): 1514–1516.

Si X-S, Hu C, Wang W, 等, 2011. An adaptive and nonlinear drift-based Wiener

process for remaining useful life estimation[C]. 2011 Prognostics; System Health Managment Confernece: 1–5.

Singpurwalla N, 1997. Gamma processes and their generalizations: an overview[J]. Engineering Probabilistic Design and Maintenance for Flood Protection, : 67–75.

Si X-S, Wang W, Hu C-H, 等, 2012. Remaining useful life estimation based on a nonlinear diffusion degradation process[J]. IEEE Transactions on Reliability, 61(1): 50–67.

Si W, Yang Q, Wu X, 等, 2018. Reliability analysis considering dynamic material local deformation[J]. Journal of Quality Technology, 50(2): 183–197.

Si X-S, Zhou D, 2013. A generalized result for degradation model-based reliability estimation[J]. IEEE Transactions on Automation Science and Engineering, 11(2): 632–637.

Song S, Coit D W, Feng Q, 等, 2014. Reliability analysis for multi-component systems subject to multiple dependent competing failure processes[J]. IEEE Transactions on Reliability, 63(1): 331–345.

Sun B, Li Y, Wang Z, 等, 2021. An improved inverse Gaussian process with random effects and measurement errors for RUL prediction of hydraulic piston pump[J]. Measurement, 173: 108604.

Tang S, Guo X, Yu C, 等, 2014. Accelerated degradation tests modeling based on the nonlinear Wiener process with random effects[J]. Mathematical Problems in Engineering, 2014(1): 560726.

Tang S, Yu C, Wang X, 等, 2014. Remaining useful life prediction of lithium-ion batteries based on the Wiener process with measurement error[J]. Energies, 7(2): 520–547.

Taylor D, Rigdon S E, Pan R, 等, 2024. Bayesian D-optimal design for life testing with censoring[J]. Quality and Reliability Engineering International, 40(1): 71–90. DOI:[10.1002/qre.322](https://doi.org/10.1002/qre.322).

Tsai C, Tseng S, Balakrishnan N, 2011. Mis-specification analyses of gamma and Wiener degradation processes[J]. Journal of Statistical Planning and Inference, 141(12): 3725–3735.

Tsai C-C, Tseng S-T, Balakrishnan N, 2012. Optimal design for degradation tests

based on gamma processes with random effects[J]. *IEEE Transactions on Reliability*, 61(2): 604–613.

Tseng S-T, Lee I-C, 2016. Optimum allocation rule for accelerated degradation tests with a class of exponential-dispersion degradation models[J]. *Technometrics*, 58(2): 244–254.

Tseng S-T, Peng C-Y, 2007. Stochastic diffusion modeling of degradation data[J]. *Journal of Data Science*, 5(3): 315–333.

Turlapaty A C, 2020. Variational Bayesian estimation of statistical properties of composite gamma log-normal distribution[J]. *IEEE Transactions on Signal Processing*, 68: 6481–6492.

Tweedie M C, 1984. An index which distinguishes between some important exponential families[J]. 579: 579–604.

Van der Vaart A W, 1998. *Asymptotic Statistics*[M]. Cambridge University Press.

Van Noortwijk J M, 2009. A survey of the application of gamma processes in maintenance[J]. *Reliability Engineering & System Safety*, 94(1): 2–21.

Wang X, 2008. A pseudo-likelihood estimation method for nonhomogeneous gamma process model with random effects[J]. *Statistica Sinica*, 18(3): 1153–1163.

Wang X, 2010. Wiener processes with random effects for degradation data[J]. *Journal of Multivariate Analysis*, 101(2): 340–351.

Wang X, Balakrishnan N, Guo B, 2014. Residual life estimation based on a generalized Wiener degradation process[J]. *Reliability Engineering & System Safety*, 124: 13–23.

Wang X, Guo B, Cheng Z, 2014. Residual life estimation based on bivariate Wiener degradation process with time-scale transformations[J]. *Journal of Statistical Computation and Simulation*, 84(3): 545–563.

Wang H, Ma X, Zhao Y, 2019. An improved Wiener process model with adaptive drift and diffusion for online remaining useful life prediction[J]. *Mechanical Systems and Signal Processing*, 127: 370–387.

Wang X, Schumitzky A, D'Argenio D Z, 2007. Nonlinear random effects mixture models: Maximum likelihood estimation via the EM algorithm[J]. *Computational Statistics & Data Analysis*, 51: 6614–6623.

Wang P, Tang Y, Bae S J, He Y, 2018. Bayesian analysis of two-phase degradation data based on change-point Wiener process[J]. *Reliability Engineering & System Safety*, 170: 244–256.

Wang P, Tang Y, Bae S J, Xu A, 2018. Bayesian approach for two-phase degradation data based on change-point Wiener process with measurement errors[J]. *IEEE Transactions on Reliability*, 67(2): 688–700. DOI:[10.1109/tr.2017.2785978](https://doi.org/10.1109/tr.2017.2785978).

Wang P, Tang Y, Joo Bae S, 等, 2018. Bayesian analysis of two-phase degradation data based on change-point Wiener process[J]. *Reliability Engineering & System Safety*, 170: 244–256. DOI:[10.1016/j.ress.2017.09.027](https://doi.org/10.1016/j.ress.2017.09.027).

Wang B, Titterington D M, 2005. Inadequacy of interval estimates corresponding to variational Bayesian approximations[C]//International Workshop on Artificial Intelligence and Statistics. PMLR: 373–380.

Wang X, Wang B X, Jiang P H, 等, 2020. Accurate reliability inference based on Wiener process with random effects for degradation data[J]. *Reliability Engineering & System Safety*, 193: 106631.

Wang X, Xu D, 2010. An inverse Gaussian process model for degradation data[J]. *Technometrics*, 52(2): 188–197.

Wang B, Yu K, Jones M C, 2010. Inference under progressively type II right-censored sampling for certain lifetime distributions[J]. *Technometrics*, 52(4): 453–460.

Wang Z, Zhai Q, Chen P, 2021. Degradation modeling considering unit-to-unit heterogeneity-A general model and comparative study[J]. *Reliability Engineering & System Safety*, 216: 107897.

Weerahandi S, 1993. Generalized confidence intervals[J]. *Journal of the American Statistical Association*, 88: 899–905.

Wen Y, Rahman M F, Xu H, 等, 2022. Recent advances and trends of predictive maintenance from data-driven machine prognostics perspective[J]. *Measurement*, 187: 110276.

Whitmore G, 1995. Estimating degradation by a Wiener diffusion process subject to measurement error[J]. *Lifetime data analysis*, 1(3): 307–319.

Whitmore G A, Schenkelberg F, 1997. Modelling accelerated degradation data using Wiener diffusion with a time scale transformation[J]. *Lifetime data analysis*, 3(1):

27–45.

Wu C, 1983. On the convergence properties of the EM algorithm[J]. *The Annals of Statistics*, 11(1): 95–103.

Wu S, Castro I T, 2020. Maintenance policy for a system with a weighted linear combination of degradation processes[J]. *European Journal of Operational Research*, 280(1): 124–133.

Wu W F, Ni C C, 2003. A study of stochastic fatigue crack growth modeling through experimental data[J]. *Probabilistic Engineering Mechanics*, 18(2): 107–118.

Wu W, Wang B X, Chen J, 等, 2023. Interval estimation of the two-parameter exponential constant stress accelerated life test model under Type-II censoring[J]. *Quality Technology & Quantitative Management*, 20(6): 751–762. DOI:[10.1080/16843703.2022](https://doi.org/10.1080/16843703.2022.2022)

Xia J, Feng Y, Teng D, 等, 2022. Distance self-attention network method for remaining useful life estimation of aeroengine with parallel computing[J]. *Reliability Engineering & System Safety*, 225: 108636.

Xiao T, Park C, Lin C, 等, 2023. Hybrid reliability analysis with incomplete interval data based on adaptive Kriging[J]. *Reliability Engineering & System Safety*, 237: 109362. DOI:[10.1016/j.ress.2023.10936](https://doi.org/10.1016/j.ress.2023.10936).

Xu L, Jordan M I, 1996. On convergence properties of the EM algorithm for gaussian mixtures[J]. *Neural Computation*, 8: 129–151.

Xu A, Sanjib B, Tang Y, 2014. A full Bayesian approach for masked data in step-stress accelerated life testing[J]. *IEEE Transactions on Reliability*, 63(3): 798–806.

Xu A, Shen L, 2018. Improved on-line estimation for gamma process[J]. *Statistics & Probability Letters*, 143: 67–73.

Xu A, Shen L, Wang B, 等, 2018. On modeling bivariate Wiener degradation process[J]. *IEEE Transactions on Reliability*, 67(3): 897–906.

Xu A, Tang Y, 2012. Objective Bayesian analysis for linear degradation models[J]. *Communications in Statistics-Theory and Methods*, 41(21): 4034–4046.

Xu A, Zhou S, 2017. Bayesian analysis of series system with dependent causes of failure[J]. *Statistical Theory and Related Fields*, 1: 128–140.

Xu A, Zhou S, Tang Y, 2021. A unified model for system reliability evaluation under dynamic operating conditions[J]. *IEEE Transactions on Reliability*, 70(1): 65–72.

Yalaoui A, Chu C, Chatelet E, 2005. Reliability allocation problem in a series-parallel system[J]. *Reliability Engineering & System Safety*, 90(1): 55–61.

Yang G, 2008. *Life Cycle Reliability Engineering*[M]. John Wiley & Sons, Inc.

Yang L, Ma X, Zhao Y, 2017. A condition-based maintenance model for a three-state system subject to degradation and environmental shocks[J]. *Computers & Industrial Engineering*, 105: 210–222. DOI:[10.1016/j.cie.2017.01.012](https://doi.org/10.1016/j.cie.2017.01.012).

Yan B, Ma X, Yang L, 等, 2020. A novel degradation-rate-volatility related effect Wiener process model with its extension to accelerated ageing data analysis[J]. *Reliability Engineering & System Safety*, 204: 107138.

Ye Z-S, Chen N, 2014. The inverse Gaussian process as a degradation model[J]. *Technometrics*, 56(3): 302–311.

Ye Z-S, Chen N, Shen Y, 2015. A new class of Wiener process models for degradation analysis[J]. *Reliability Engineering & System Safety*, 139: 58–67.

Ye Z-S, Chen L-P, Tang L C, 等, 2014. Accelerated degradation test planning using the inverse Gaussian process[J]. *IEEE Transactions on Reliability*, 63(3): 750–763.

Ye Z-S, Hong Y, Xie Y, 2013. How do heterogeneities in operating environments affect field failure predictions and test planning?[J]. *The Annals of Applied Statistics*, 7(4): 2249–2271.

Ye Z-S, Shen Y, Xie M, 2012. Degradation-based burn-in with preventive maintenance[J]. *European Journal of Operational Research*, 221(2): 360–367.

Ye Z-S, Wang Y, Tsui K-L, 等, 2013. Degradation data analysis using Wiener processes with measurement errors[J]. *IEEE Transactions on Reliability*, 62(4): 772–780.

Ye Z-S, Xie M, 2015. Stochastic modelling and analysis of degradation for highly reliable products[J]. *Applied Stochastic Models in Business and Industry*, 31(1): 16–32.

Ye Z-S, Xie M, Tang L-C, 等, 2012. Degradation-based burn-in planning under competing risks[J]. *Technometrics*, 54(2): 159–168.

Yuan T, Bae S J, Zhu X, 2016. A Bayesian approach to degradation-based burn-in optimization for display products exhibiting two-phase degradation patterns[J]. *Reliability Engineering & System Safety*, 155: 55–63.

Yuan X-X, Higo E, Pandey M D, 2021. Estimation of the value of an inspection

and maintenance program: A Bayesian gamma process model[J]. *Reliability Engineering & System Safety*, 216: 107912.

Yuan T, Ji Y, 2015. A hierarchical Bayesian degradation model for heterogeneous data[J]. *IEEE Transactions on Reliability*, 64(1): 63–70.

Yuan T, Zhu X, 2012. Reliability study of ultra-thin dielectric films with variable thickness levels[J]. *IIE Transactions*, 44(9): 744–753.

Zaidan M A, Mills A R, Harrison R F, 等, 2016. Gas turbine engine prognostics using Bayesian hierarchical models: A variational approach[J]. *Mechanical Systems and Signal Processing*, 70: 120–140.

Zellner A, 1986. On assessing prior distributions and Bayesian regression analysis with g-prior distributions[J]. *Bayesian Inference and Decision Techniques*.

Zhai Q, Chen P, Hong L, 等, 2018. A random-effects Wiener degradation model based on accelerated failure time[J]. *Reliability Engineering & System Safety*, 180: 94–103.

Zhai Q, Xu A, Yang J, 等, 2023. Statistical modeling and reliability analysis for degradation processes indexed by two scales[J]. *IEEE Transactions on Industrial Informatics*, 20(3): 3675–3684.

Zhai Q, Ye Z-S, 2018. Degradation in common dynamic environments[J]. *Technometrics*, 60(4): 461–471. DOI:[10.1080/00401706.2017.1375994](https://doi.org/10.1080/00401706.2017.1375994).

Zhai Q, Ye Z-S, 2023. A multivariate stochastic degradation model for dependent performance characteristics[J]. *Technometrics*, 65(3): 315–327. DOI:[10.1080/00401706.2022.222222](https://doi.org/10.1080/00401706.2022.222222).

Zhang H, Chen M, Shang J, 等, 2021. Stochastic process-based degradation modeling and RUL prediction: from Brownian motion to fractional Brownian motion[J]. *Science China Information Sciences*, 64(7): 171201.

Zhang Y, Feng F, Wang S, 等, 2023. Joint nonlinear-drift-driven Wiener process-Markov chain degradation switching model for adaptive online predicting lithium-ion battery remaining useful life[J]. *Applied Energy*, 341: 121043. DOI:[10.1016/j.apenergy.2023.121043](https://doi.org/10.1016/j.apenergy.2023.121043).

Zhang J, Huang X, Fang Y, 等, 2016. Optimal inspection-based preventive maintenance policy for three-state mechanical components under competing failure modes[J]. *Reliability Engineering & System Safety*, 152: 95–103.

DOI:[10.1016/j.ress.2016.02.007](https://doi.org/10.1016/j.ress.2016.02.007).

Zhang Z, Hu C, Si X, 等, 2017. Stochastic degradation process modeling and remaining useful life estimation with flexible random-effects[J]. *Journal of the Franklin Institute*, 354(6): 2477–2499.

Zhang Y, Ouyang L, Meng X, 等, 2024. Condition-based maintenance considering imperfect inspection for a multi-state system subject to competing and hidden failures[J]. *Computers & Industrial Engineering*, 188: 109856. DOI:[10.1016/j.cie.2023.109856](https://doi.org/10.1016/j.cie.2023.109856).

Zhang Z, Si X, Hu C, 等, 2018. Degradation data analysis and remaining useful life estimation: A review on Wiener-process-based methods[J]. *European Journal of Operational Research*, 271(3): 775–796.

Zhou S, Tang Y, Xu A, 2021. A generalized Wiener process with dependent degradation rate and volatility and time-varying mean-to-variance ratio[J]. *Reliability Engineering & System Safety*, 216: 107895.

Zhou S, Xu A, 2019. Exponential dispersion process for degradation analysis[J]. *IEEE Transactions on Reliability*, 68(2): 398–409.

Zhou S, Xu A, Lian Y, 等, 2020. Variational Bayesian analysis for Wiener degradation model with random effects[J]. *Communications in Statistics-Theory and Methods*, : 1–21.

Zhou S, Xu A, Tang Y, 等, 2023. Fast Bayesian inference of reparameterized gamma process with random effects[J]. *IEEE Transactions on Reliability*, 73(1): 399–412. DOI:[10.1109/TR.2023.3263940](https://doi.org/10.1109/TR.2023.3263940).

Zhuang L, Xu A, Wang Y, 等, 2024. Remaining useful life prediction for two-phase degradation model based on reparameterized inverse Gaussian process[J]. *European Journal of Operational Research*, 319(3): 877–890. DOI:<https://doi.org/10.1016/j.ejor.2024.06.032>.

Zhu R, Chen Y, Peng W, 等, 2022. Bayesian deep-learning for RUL prediction: An active learning perspective[J]. *Reliability Engineering & System Safety*, 228: 108758. DOI:[10.1016/j.ress.2022.108758](https://doi.org/10.1016/j.ress.2022.108758).

张志华, 2002. 加速寿命试验及其统计分析 [M]. 加速寿命试验及其统计分析.

李海洋, 谢里阳, 李铭, 等, 2018. 一种新的无失效数据可靠性评估方法研究 [J].

兵工学报, 39(8): 1622.

茆诗松, 汤银才, 王玲玲, 2008. 可靠性统计 [M]. 可靠性统计.

韩明, 2003. 产品无失效数据的综合处理 [J]. 机械工程学报, 39(2): 129–132.

# 索引

AIC, 32, 34, 97, 99, 123, 127–130, 181, 182, 208, 218, 220, 222, 250, 254, 255

EM 算法, 6, 19, 26, 88, 94, 102, 103, 118, 120, 124, 125, 127, 128, 146, 159, 167

Gibbs 抽样, 103, 139

MCMC, 6, 67, 82, 83, 94, 103, 120, 128, 133, 150

RUL 预测, 21, 49, 62, 63, 75, 138, 139, 155, 156, 167, 168, 181

一般路径模型, 6

似然函数, 8, 25, 27, 44, 63, 64, 73, 74, 79, 82, 93, 100, 102, 120–125, 130, 136, 146, 147, 159, 168, 170–172, 174, 177, 179, 183–185, 208, 213, 214, 217, 222–224, 233, 236, 237, 245, 246

伽马分布, 47, 61, 64, 65, 67, 69, 75, 83, 84, 86, 117, 223

伽马过程, 6, 8–10, 61–63, 70, 73, 76, 78, 79, 84, 85, 97, 116, 138, 208–210, 213, 217, 218, 220, 225, 228, 230, 237, 250, 255

先验, 3, 4, 46, 47, 51, 62–67, 70, 71, 74, 75, 77, 79, 82–84, 87, 90–92, 94, 103, 120, 128, 148, 149, 151, 153

剩余寿命, 8, 44, 45, 57, 138, 140, 142

协变量, 6, 9

变分贝叶斯, 62, 83, 88

变点, 42–44, 51, 52, 138, 139, 141, 142, 145, 146, 148, 150, 151, 153–158, 160

可信区间, 71, 94, 139, 153

可靠度, 3, 4, 67, 70, 158, 167, 198

可靠度函数, 4, 141, 142, 155, 157–159, 211, 218, 220, 231, 232, 237

可靠性, 3–6, 9, 11, 13, 18, 42, 46, 118, 281

- 139, 142, 144, 146, 151, 153, 158, 173, 211, 218, 221, 222, 224, 225, 232, 233, 240, 250, 254, 261
- 可靠性评估, 3–5, 10, 15, 19, 224, 240, 244, 261
- 均方误差, 30
- 多维退化数据, 6
- 失效数据, 3, 4, 118
- 失效时间, 9, 19, 21, 23, 33, 34, 37, 38, 42, 44, 45, 49, 54, 55, 78, 79, 98, 117, 127, 129, 138, 140–142, 144, 145, 155, 156, 211, 212, 239
- 寿命, 3–10, 13–15, 18, 19, 22, 23, 42, 61, 83, 86, 98, 99, 117, 119, 138, 139, 157, 167, 182, 211–213, 224–226, 228–233, 236, 237, 240, 241, 243, 244, 250, 261
- 异质性, 6–8, 10, 21, 42, 46, 47, 62, 73, 78, 79, 83, 85, 87, 116–119, 127, 139, 141, 149, 150, 174, 222, 224
- 性能退化, 4–7, 9, 10, 17–19, 24, 44, 61, 73, 78, 84, 116, 133, 140, 142, 143, 145, 153, 167, 168, 172, 216
- 指数分散过程, 6, 9, 10
- 故障, 3–5, 13, 75, 139, 141, 143, 151, 153, 157, 224–226, 261
- 时间尺度函数, 18, 19, 29, 30, 32, 38
- 正态分布, 3, 17, 20, 24, 43, 46, 48, 52, 53, 68, 76, 84, 86, 91, 93, 97, 98, 117, 118, 140, 141, 149, 151, 173, 174, 179, 180, 182, 193, 222, 224
- 测量误差, 6, 8, 9, 19, 34, 42, 43, 51
- 相依, 4, 7, 9, 240
- 相关系数, 19, 28, 29, 32, 97, 98, 245, 250, 255
- 维纳过程, 6, 7, 9, 10, 17–20, 42, 51, 97, 116, 128, 130, 138, 208–211, 218, 220, 229, 261
- 置信区间, 33, 34, 62, 71, 98, 129, 130, 148, 150, 153, 182, 215, 217, 220, 222, 239, 245, 246, 248, 250, 254, 255
- 自适应替换, 117, 139, 140, 143, 144, 155, 156
- 贝叶斯, 3, 8, 9, 40, 42, 47, 61, 62, 66, 67, 70, 71, 76, 77, 82, 83, 88, 92, 94, 95, 97, 98, 103, 104, 120, 128, 138, 140, 144, 145, 148–151, 153, 155, 162, 167, 179, 250
- 退化过程, 6, 8, 10, 18–20, 37, 83, 84, 117, 118, 120, 138, 141, 145, 153, 155, 157, 223, 261
- 退化阈值, 33, 34, 86, 99

逆高斯分布, 7, 18, 19, 21, 24, 40, 44, 82–87, 97, 117, 118, 120,  
54, 168 129, 167, 174, 176–179, 181,  
逆高斯过程, 6, 9, 10, 97, 208, 230, 182, 185, 208, 222  
250, 261  
预烧样本, 70, 82  
随机效应, 6–10, 21–23, 32, 34, 42,