

基于随机过程理论的多状态系统建模与可靠性评估

王丽英 崔利荣 著



科学出版社

基于随机过程理论的多状态 系统建模与可靠性评估

王丽英 崔利荣 著



科学出版社

北京

内 容 简 介

本书是一本基于随机过程理论的多状态系统建模与可靠性评估方面的专著. 全书共 11 章. 内容包括: 多状态系统的基本概念、建模及可靠性评估的基础知识; 多运行水平系统建模与可靠性分析; 历史相依、环境相依系统建模与可靠性评估; 冗余相依、故障相依、空间相依系统建模与可靠性分析. 本书内容新颖、全面、系统, 可为大型复杂系统的可靠性分析和优化设计提供理论依据.

本书适用于从事可靠性维修性相关研究的学者及可靠性数学、应用数学、管理科学与工程、系统工程等专业的研究生学习使用, 也可供从事相关工作的工程师在具体的工程实践中参考.

图书在版编目(CIP)数据

基于随机过程理论的多状态系统建模与可靠性评估/王丽英, 崔利荣著.
—北京: 科学出版社, 2017.12

ISBN 978-7-03-055381-2

I. ①基… II. ①王… ②崔… III. ①随机过程-系统建模②随机过程-系统可靠性-评估 IV. ①O211.6

中国版本图书馆 CIP 数据核字(2017) 第 281381 号

责任编辑: 胡庆家 / 责任校对: 邹慧卿
责任印制: 张 伟 / 封面设计: 迷底书装

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

北京建宏印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2017 年 12 月第 一 版 开本: 720 × 1000 1/16

2017 年 12 月第一次印刷 印张: 12 1/4

字数: 240 000

定价: 78.00 元

(如有印装质量问题, 我社负责调换)

前 言

复杂系统及其组成元件在寿命周期内往往有明显的退化特性,故障过程呈现出多状态的特征.以“二态”假设为基础的传统可靠性理论难以描述大多数复杂系统的故障规律.多状态系统既能表征复杂系统多状态的特点,也能反映系统与元件之间的关系,因而成为学术界和工业界所共同关注的热点,并取得了一定的研究成果.

合理压缩随元件个数增加而急剧增长的状态空间、精确描述系统状态之间及元件之间普遍存在的相依性是多状态系统可靠性研究中亟待解决的重点问题.近十年来,作者及合作者一直关注上述问题,并开展了基于随机过程理论的多状态系统建模与可靠性评估.本书是作者及合作者已有研究成果的总结.

本书分为 11 章.第 1 章是应用随机过程理论进行多状态系统建模与可靠性评估所需的基本知识和基础理论.第 2 章构建多运行水平马尔可夫可修系统模型,运用聚合随机过程理论压缩状态空间,给出常用及多状态系统特有的可靠性度量指标计算方法.第 3, 4 章构建状态历史相依多状态系统模型,运用马尔可夫、半马尔可夫及离子通道理论对其进行可靠性评估.第 5, 6 章考虑多状态系统的环境相依性,构建多运行机制马尔可夫可修系统及二机制半马尔可夫可修系统模型并进行可靠性分析.第 7, 8 章考虑多状态系统元件间的冗余相依性,构建冗余相依多状态马尔可夫可修系统及故障相依多状态半马尔可夫可修系统模型,并进行可靠性评估.第 9—11 章考虑拓扑结构对多状态系统部件间相依性的影响,构建空间相依圆形、星形及网格马尔可夫可修系统模型并对其进行可靠性分析.

读者通过本书的学习,可以掌握运用随机过程理论对多状态系统进行可靠性评估的基本理论和方法,了解国内外相关研究前沿.因此,本书对可靠性数学、应用数学、管理科学与工程、系统工程等相关专业的大学高年级学生和研究生深入学习和掌握多状态系统可靠性的相关理论知识很有帮助,同时对相关工作的工程师有一定指导意义.

本书第 2—6 章是王丽英和崔利荣教授合作的研究成果.第 7—11 章是王丽英和杨清、田玉然等合作的研究成果.在本书的编著过程中,裴朝娜、杨艳妹、李康乐做了大量的编辑、排版、校订等工作,在此一并感谢.

本书是在王丽英主持的国家自然科学基金项目“基于聚合随机过程的元件相依多状态系统可靠性研究”,河北省自然科学基金项目“分区载荷共享多状态系统可靠性研究”及崔利荣教授主持的国家自然科学基金项目“系统可靠性建模与分析

的理论与方法研究”资助下完成的。限于作者水平，遗漏或错误之处在所难免，恳请读者批评指正。

王丽英 崔利荣

2017年8月

目 录

前言

第 1 章 多状态系统的概念、模型与可靠性分析	1
1.1 多状态系统概念	1
1.2 多状态系统建模	1
1.2.1 多状态系统建模方法	1
1.2.2 多状态系统可靠性研究的“状态空间爆炸”问题	2
1.2.3 多状态系统可靠性研究的相依问题	2
1.3 多状态系统可靠性度量	3
1.3.1 常用可靠性度量	3
1.3.2 多状态系统特有可靠性度量	4
1.4 基于随机过程法的多状态系统建模	4
1.4.1 马尔可夫和半马尔可夫可修多状态系统	4
1.4.2 基于状态聚合的多状态系统建模	9
1.5 拉普拉斯变换和拉普拉斯-斯蒂尔切斯变换	13
1.5.1 拉普拉斯变换	13
1.5.2 非负随机变量的拉普拉斯-斯蒂尔切斯变换	14
参考文献	15
第 2 章 多运行水平马尔可夫可修系统建模与可靠性分析	20
2.1 引言	20
2.2 模型描述	23
2.3 可用度、维修频度和故障频度	26
2.3.1 可用度	26
2.3.2 系统的故障频度和维修频度	27
2.4 多变量半马尔可夫过程	28
2.4.1 多变量半马尔可夫过程的定义	28
2.4.2 多变量半马尔可夫过程的核	29
2.5 随机时间分布	32
2.5.1 开工时间	33
2.5.2 周期长度	33
2.5.3 不同运行水平集的逗留时间	33

2.5.4 一个开工时间区间的总输出	34
2.6 访问次数分布	35
2.7 数值算例	35
2.7.1 系统的可靠性评估	35
2.7.2 维修策略对系统的影响分析	41
2.8 结论	43
参考文献	44
第 3 章 状态历史相依马尔可夫可修系统建模与可靠性分析	45
3.1 引言及系统描述	45
3.2 系统可靠性度量	47
3.2.1 一些基本结论	47
3.2.2 访问可变状态集的概率	50
3.2.3 故障频度	50
3.3 随机时间分布	51
3.3.1 可变状态集总逗留时间分布	51
3.3.2 单个状态集逗留时间的分布	53
3.4 数值算例	55
3.5 结论	57
参考文献	57
第 4 章 状态历史相依半马尔可夫可修系统建模与可靠性分析	59
4.1 引言及系统描述	59
4.2 一些基本结论	60
4.3 首次故障前时间及故障频度	63
4.3.1 首次故障前时间	64
4.3.2 故障频度	65
4.4 聚合半马尔可夫过程及其核	66
4.4.1 聚合半马尔可夫过程	66
4.4.2 半马尔可夫核	67
4.5 随机时间分布	69
4.5.1 总逗留时间分布	69
4.5.2 单个状态集逗留时间分布	71
4.6 数值算例	73
4.6.1 系统的可靠性度量	73
4.6.2 状态历史相依半马尔可夫和马尔可夫系统可靠性比较分析	76
4.7 结论	78

参考文献	79
第 5 章 多运行机制马尔可夫可修系统建模及可靠性分析	80
5.1 引言	80
5.2 系统描述	81
5.3 首次故障前时间	83
5.3.1 工作状态集可变情形的首次故障前时间	83
5.3.2 没有可变状态集情形的首次故障前时间	86
5.4 系统可用度	87
5.5 数值算例	88
5.5.1 系统的可靠性评估	88
5.5.2 机制逗留时间对系统的影响分析	91
5.6 结论	93
参考文献	93
第 6 章 交替环境中的半马尔可夫可修系统建模及可靠性分析	94
6.1 引言及系统描述	94
6.2 系统可用度	95
6.2.1 转移概率函数	95
6.2.2 系统的瞬时和稳态可用度	97
6.3 首次故障前时间	99
6.3.1 在工作状态集内的逗留概率和逗留时间分布	99
6.3.2 首次故障前时间分布	102
6.4 机制逗留时间为常数情形下的半马尔可夫可修系统	104
6.4.1 常数情形下的可用度	104
6.4.2 常数情形下的首次故障前时间分布	105
6.5 机制逗留时间为指数分布情形下的半马尔可夫可修系统	106
6.6 数值算例	106
6.7 结论	112
参考文献	112
第 7 章 冗余相依多状态马尔可夫可修系统可靠性分析	113
7.1 引言	113
7.2 系统假设	114
7.2.1 冗余相依多状态马尔可夫可修系统	114
7.2.2 系统状态转移分析	114
7.3 访问各个状态集的概率及首次故障前时间	116
7.3.1 访问各个状态集概率	116

7.3.2 首次故障前时间	117
7.4 可接受状态集逗留时间分布	119
7.5 数值算例	119
7.5.1 系统描述	119
7.5.2 可用度及访问各个状态集的概率	121
7.5.3 首次故障前时间及一个周期中系统在可接受状态集的逗留时间	123
7.6 结论	124
参考文献	124
第 8 章 故障相依载荷共享多状态系统可靠性分析	126
8.1 引言	126
8.2 系统假设	127
8.2.1 故障相依载荷共享系统	127
8.2.2 系统对应的半马尔可夫过程	128
8.2.3 过程的半马尔可夫核	128
8.3 首次故障前时间	131
8.4 系统可用度	132
8.5 数值算例	134
8.5.1 系统描述及半马氏核	134
8.5.2 首次故障前时间分布及可用度	137
8.5.3 寿命和修理时间都是指数情形下的系统可靠性分析	140
8.6 结论	142
参考文献	142
第 9 章 空间相依圆形马尔可夫可修系统可靠性分析	145
9.1 引言	145
9.2 模型假设	146
9.2.1 基本模型	146
9.2.2 四部件和五部件空间相依圆形马尔可夫可修系统状态	146
9.3 四部件和五部件空间相依圆形马尔可夫可修系统状态转移分析	149
9.4 四部件和五部件空间相依圆形马尔可夫可修系统可用度	151
9.5 数值算例	152
9.6 结论	154
参考文献	154
第 10 章 空间相依星形马尔可夫可修系统可靠性分析	156
10.1 引言	156
10.2 模型假设	156

10.3 六部件空间相依星形系统状态及状态间分析	157
10.3.1 系统状态	157
10.3.2 系统状态转移分析	159
10.4 系统可用度	164
10.5 数值算例	165
10.5.1 可用度分析	165
10.5.2 系统访问四种状态集的概率	166
10.5.3 系统可靠性对部件故障率的敏感性分析	167
10.6 结论	170
参考文献	170
第 11 章 空间相依网格马尔可夫可修系统可靠性分析	172
11.1 系统假设	172
11.1.1 六部件空间相依网格马尔可夫可修系统假设	172
11.1.2 系统的状态	173
11.2 系统的状态转移分析	174
11.3 系统可用度分析	178
11.4 数值算例	179
11.4.1 瞬时可用度	179
11.4.2 系统访问四类状态集概率	179
11.4.3 系统可靠性的敏感性分析	180
11.5 结论	183
参考文献	183

第1章 多状态系统的概念、模型与可靠性分析

1.1 多状态系统概念

传统的可靠性理论认为系统及其组成元件仅存在正常工作和故障两个状态。然而,现代工业生产中的许多系统是由有不同运行水平 (Operating Level) 和不同故障模式的多状态元件组成。元件的不同水平或不同故障模式对整个系统有不同的影响。这样的系统被称为多状态系统 (Multi-State System, MSS)(Lisnianski, Levitin, 2003)。多状态系统既能真实地表征复杂系统多状态的特点,又能反映出系统性能与元件性能、系统可靠性与系统性能的关系 (李春洋, 2011),因而成为学术界和工业界所共同关注的热点问题,并在机械工程、计算机和网络系统、网格、通讯系统、能源系统、供给系统、城市基础设施、战略和防御等众多领域得到了迅速发展 (刘宇, 2011)。各国学者的不断努力使得多状态系统的建模、表示及定量分析等方面有了很大的发展。Lisnianski 和 Levitin(2003), Zio (2009) 分别对已有研究做了总结和归纳。

1.2 多状态系统建模

1.2.1 多状态系统建模方法

目前,多状态系统可靠性理论的研究主要集中在三个方面:多状态系统的可靠性建模及评估、可靠性优化设计以及维修和保修管理策略。通过可靠性建模与评估可以从可靠性角度出发,对不同的设计方案进行分析比较,为设计决策提供依据;可以定量地预计或评价系统的可靠性,发现薄弱环节,为系统设计改进或生产过程控制提供依据;同时可靠性建模与评估又是进行故障模式、影响及危害性分析的基础 (蒋仁言, 左明健, 1999),因而多状态系统的可靠性建模及评估是多状态可靠性理论研究的基础。然而已有的多状态系统可靠性模型存在不能满足实际要求、覆盖面远远不够、模型与实际有一定差距等方面的问题。本书拟从工程实际出发,围绕多状态系统可靠性建模及评估来展开。

多状态系统可靠性分析的方法主要有以下四类:二态布尔代数扩展法、随机过程法 (马尔可夫和半马尔可夫过程)、通用生成函数法 (Universal Generating Function) 和随机模拟法 (Lisnianski, 2007; Soro et al., 2010; Schoenig et al., 2006)。四种

方法的比较如表 1.1 所示.

表 1.1 多状态系统主要研究方法比较

方法名称	优点	存在问题
二态布尔代数扩展法	可参照二态布尔代数法	适用于状态数目较少的系统, 无法描述系统的动态变化过程
随机过程法	理论成熟, 可以得到概率型、时间分布型可靠性度量指标	适用于状态数目较少的系统
通用生成函数法	适用于元件数目多但结构相对简单的系统, 计算速度快	不适用结构、相依模式复杂的系统, 不能得到时间分布型可靠性度量指标
随机模拟法	可以很好地描述复杂系统的动态演化过程	模型的执行和构造费时、造价高, 对系统的状态数目很敏感, 不能得到精确解

1.2.2 多状态系统可靠性研究的“状态空间爆炸”问题

对于一个由 n 个元件组成的多状态系统而言, 若元件 j ($j = 1, 2, \dots, n$) 有 k_j 个不同的状态, 则系统有 $K = \prod_{j=1}^n k_j$ 个基本状态. 例如, 由 8 个三态元件组成的多状态系统有 $3^8 = 6561$ 个基本状态. 而大多数复杂系统的元件个数远远大于 8 个, 并且每个元件一般有工作、劣化、故障等状态. 由此可见, 即使元件数目相对较小的多状态系统, 基本状态的数目也很大. 这将给可靠性模型的求解、评估指标的选取及计算带来很大的困难. 从表 1.1 也可以看出, 二态布尔代数扩展法、随机过程法、随机模拟法的应用都受到系统状态数目的限制. 因此, 如何合理地进行状态空间压缩, 缓解“状态空间爆炸”带来的诸多问题, 是多状态系统可靠性建模及评估所面临的首要问题.

已有文献给出了一些“降维”的方法. 如宋月 (2006), Jiang (2003) 等运用截断法, 把出现概率较小的状态忽略掉, 以减少状态空间的数目. 该方法在高可靠性系统分析中有很好的效果, 不足之处在于容易把出现概率小, 但破坏性强的状态忽略掉, 从而高估系统的可靠性. Lisniansk 和 Ding (2009) 把随机过程法和通用生成函数法结合起来进行可靠性评估. 这种方法在一定程度上克服了随机过程法受系统状态数目限制的缺点, 但是很难描述元件间的相依关系, 并且只能得到概率型可靠性度量指标. 最近一些学者运用状态聚合法, 开展基于聚合随机过程的系统可靠性研究.

1.2.3 多状态系统可靠性研究的相依问题

2009 年, 国际可靠性工程领域著名学者 Zio 教授, 在可靠性领域著名期刊 *Reliability Engineering and System Safety* 中的《可靠性工程: 老问题和新挑战》

一文中指出“系统状态之间以及各个元件的状态之间存在的相依性,是多状态系统建模困难的原因所在”(Zio, 2009).

目前描述状态之间相依性的系统有马氏相依系统(Yun et al., 2007)、环境相依系统(Soszynska, 2010)、历史相依系统等. 描述元件相依性的研究主要集中在共因失效系统(王正, 谢里阳, 2008)、经济相依系统、冗余相依系统、故障相依系统、空间相依系统等. 东北大学以谢里阳教授为首的科研团队,在共因失效的机械系统可靠性方面进行了广泛而深入的研究(王正, 谢里阳, 2008). 李春洋(2011), 武小悦(2006), 肖刚和李志忠(2008), 薛云和曹晋华(2006)等也做了大量有意义的工作.

1.3 多状态系统可靠性度量

1.3.1 常用可靠性度量

可靠性是指系统在规定的时间内、规定的条件下,完成规定功能的概率.当系统可以进行维修时,通过修理或更换,使系统恢复其规定功能,这种过程称为维修.在维修工作完成以后,系统又能继续完成其规定的功能.可修系统中的部件除了工作状态外,还可以处于检修状态,故障的部件经过维修后能再度投入工作,因此可修系统的可靠性问题比不可修系统要复杂得多.人们从不同侧面、不同角度出发,对系统的可靠性进行定量分析.常用的可靠性指标有:首次故障前时间、可用度、故障频度、开工时间、停工时间.

可修系统的运行随时间的进程是正常与故障交替出现的.用 X_i 和 $Y_i (i = 1, 2, \dots)$ 分别表示第 i 个周期的开工时间和停工时间.一般情况下, $X_1, X_2, \dots$ 或 $Y_1, Y_2, \dots$ 的分布不同.系统的首次故障前时间分布指 X_1 的分布.首次故障前平均时间为

$$\text{MTTFF} = E(X_1) = \int_0^{+\infty} t dF_1(t).$$

当可修系统的故障产生灾难性后果时,首次故障前时间分布及其均值是该系统的最重要的可靠性指标.

系统的瞬时可用度 $A(t)$ 是系统在时刻 t 工作的概率.系统的瞬时可用度 $A(t)$ 只涉及时刻 t 系统是否正常,并不关心以前是否发生过故障.在工程实际中,有时更需要了解系统进入稳定状态后的可用度 $A = \lim_{t \rightarrow \infty} A(t)$, 以及 $(0, t]$ 时间内的平均

可用度 $\bar{A}(t) = \frac{1}{t} \int_0^t A(u) du$. 若极限 $\bar{A} = \lim_{t \rightarrow \infty} \bar{A}(t)$ 存在,则称 $\bar{A}$ 为极限平均可用度.它的物理解释是描述系统在一个在很长的时间过程内,能维持运行状态所占的时间份额,它是衡量系统可用度的重要参数.

令 $N(t)$ 为 $(0, t]$ 时间内系统的故障次数,系统在时间段 $(0, t]$ 内的平均故障次

数 $M(t) = E\{N(t)\}$. 当 $M(t)$ 的导数存在时, 称 $m(t) = \frac{d}{dt}M(t)$ 为系统的瞬时故障频度. 在工程应用中, 更感兴趣的是系统的稳态故障频度 $M = \lim_{t \rightarrow \infty} \frac{M(t)}{t}$. $M(t)$ 和 M 也是重要的可靠性指标. 在更换问题的研究中, 它告诉我们大约需要多少备件 (曹晋华, 程侃, 2006).

平均开工时间 (MUT 或 MTBF) 为 $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i$, 平均停工时间 $MDT = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n Y_i$, 平均周期 $MCT = MUT + MDT$. 除了反映可修系统自身的可靠性指标外, 有时还需要反映修理设备忙闲程度的指标, 如修理设备忙的瞬时概率和修理设备忙的稳态概率等.

还有一些文献从工程实际出发定义了其他可靠性指标 (Rausand, Hoyland, 2003; Gnedenenko, Ushakov, 1995; Csenki, 1995), 如无故障运行概率 (Probability of a Failure-Free Operation)、访问频度 (Visiting Frequency)、平均访问逗留时间 (Mean Duration of a Visit)、区间可用度 (Interval Availability)、区间可靠度 (Interval Reliability)、联合区间可靠度 (Joint Interval Reliability)、运行额外时间资源 (Extra Time Resource for Performance)、可接受空闲时间 (Acceptable Idle Interval) 等.

1.3.2 多状态系统特有可靠性度量

对多状态系统而言, 除了一般的可靠性度量外, 人们更感兴趣的是状态集有关的度量指标. 在工程实践中, 通常把系统的整个状态空间划分成不同的状态集, 并把每个状态集看做一个状态处理, 这种方法称为状态聚合法 (Chan, Asgarpoor, 2006; 蒋仁言, 左明健, 1999). 如按照系统输出能否满足客户需求, 把基本状态分为可接受集和不可接受集两类; 按照系统运行水平的高低, 把基本状态分为完美工作状态集、劣化状态集、警戒状态集及故障等几个集类. 这类系统称为状态聚合系统.

1.4 基于随机过程法的多状态系统建模

1.4.1 马尔可夫和半马尔可夫可修多状态系统

1. 马尔可夫随机过程

设 $\{X(t), t \geq 0\}$ 是取值在 $E = \{0, 1, \dots\}$ 或 $E = \{0, 1, \dots, N\}$ 上的一个随机过程. 假设过程在时刻 s 的状态为 $X(s) = i$. 过程在时刻 $s+t$ 将在状态 j 的条件概率为

$$P\{X(t+s) = j | X(s) = i, X(u) = x(u), 0 \leq u < s\},$$

其中 $\{x(u), 0 \leq u < s\}$ 表示系统到时刻 s 之前, 不包括时刻 s 的历史. 若

$$P\{X(t+s) = j | X(s) = i, X(u) = u, 0 \leq u < s\} = P\{X(t+s) = j | X(s) = i\} \quad (1.1)$$

对所有可能的 $x(u) (0 \leq u < s)$ 都成立, 则称该过程具有马尔可夫性 (Rausand, Hoyland, 2003; 曹晋华, 程侃, 2006). 式 (1.1) 可做如下的直观解释: 已知现在状态的条件, 过程将来的发展和过去独立. 若过程 $\{X(t), t \geq 0\}$ 具有马尔可夫性, 则称之为离散状态空间 E 上的连续时间马尔可夫过程.

如果在任意的时刻 $t, u \geq 0$ 均有

$$P\{X(t+u) = j | X(u) = i\} = P_{ij}(t)$$

与 u 无关, 则称马尔可夫过程 $\{X(t), t \geq 0\}$ 是时齐的 (齐次的). 如果没有特殊声明, 下面的马尔可夫随机过程均是时齐的有限状态空间 E 上的随机过程.

对固定的 $i, j \in E$, 函数 $P_{ij}(t)$ 称为转移概率函数. $|E| \times |E|$ 维矩阵 $P(t) = [P_{ij}(t)] (i, j \in E)$ 称为转移概率矩阵. 假定

$$\lim_{t \rightarrow 0} P_{ij}(t) = \delta_{ij} = \begin{cases} 1, & \text{如果 } i = j, \\ 0, & \text{如果 } i \neq j. \end{cases}$$

转移概率函数有如下的性质:

$$P_{ij}(t) \geq 0, \quad \sum_{j \in E} P_{ij}(t) = 1, \quad \sum_{k \in E} P_{ik}(u) P_{kj}(v) = P_{ij}(u+v).$$

若令 $p_j(t) = P\{X(t) = j\} (j \in E)$, 它表示时刻 t 系统处于状态 j 的概率, 并且

$$p_j(t) = \sum_{k \in E} p_k(0) P_{kj}(t).$$

对于有限状态空间 E 上的齐次马尔可夫过程有下列重要性质.

(1) 下列极限

$$\begin{cases} \lim_{\Delta t \rightarrow 0} \frac{P_{ij}(\Delta t)}{\Delta t} = q_{ij}, & i \neq j, i, j \in E, \\ \lim_{\Delta t \rightarrow 0} \frac{1 - P_{ii}(\Delta t)}{\Delta t} = -q_{ii}, & i \in E \end{cases}$$

存在且满足 $\sum_{j \in E} q_{ij} = 0, i \in E$.

(2) 马尔可夫过程在任何状态 i 的逗留时间服从参数为 $-q_{ii}$ 的指数分布, 不依赖于下一个将要转入的状态, 从状态 i 转移到状态 j 的概率为 $q_{ij}/(-q_{ii})$.

定义 $|E| \times |E|$ 维矩阵 $Q = (q_{ij}) (i, j \in E)$, 称 Q 为马尔可夫过程的转移率矩阵或无穷小生成矩阵. 令 $p(t) = (p_0(t), p_1(t), \dots, p_N(t))$, $P(t) = (P_{ij}(t))$, $i, j \in E$, 则有下列两个微分方程组成立:

$$\frac{d\mathbf{p}(t)}{dt} = \mathbf{p}(t)Q, \quad (1.2)$$

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{P}(t)Q. \quad (1.3)$$

式 (1.2) 称为马尔可夫过程的状态方程组, 用它可以得到系统时刻 t 处于各个状态的概率. 式 (1.3) 称为柯尔莫哥洛夫前进方程组 (Kolmogorov Forward Equations), 通过它可以得到系统的转移概率函数. 通常用 L 变换的方法得到上述方程组的数值解. 有关马尔可夫随机过程的其他性质可参见文献 (Ross, 1996).

2. 半马尔可夫过程

设 $\{(Z_n, R_n), n \geq 0\}$ 是一个二维离散时间随机过程. 假设 $R_0 = 0$, 序列 $\{Z_n, n \geq 0\}$ 表示系统相继访问的状态, 其状态空间为 E . $\{R_n, n \geq 0\}$ 是系统在各个状态的逗留时间, 更精确地说, R_n 表示系统在状态 Z_{n-1} 的逗留时间. 若状态转移时刻用序列 $\{T_n, n \geq 0\}$ 表示, 其中 $T_0 = 0$, 则 $T_1 = R_1, \dots, T_n = \sum_{r=1}^n R_r$, $R_n = T_n - T_{n-1}$.

若对所有的 $n = 0, 1, \dots, j \in E, t \geq 0$, 都有

$$\begin{aligned} &P\{Z_{n+1} = j, R_{n+1} \leq t | Z_0, Z_1, \dots, Z_n, R_0, R_1, \dots, R_n\} \\ &= P\{Z_{n+1} = j, R_{n+1} \leq t | Z_n\}, \end{aligned}$$

则称 $\{(Z_n, R_n), n \geq 0\}$ 为状态空间 E 上的马尔可夫更新过程 (Janssen, Manca, 2006). 如果对所有 $i, j \in E, t \geq 0$,

$$P\{Z_{n+1} = j, R_{n+1} \leq t | Z_n\} = Q_{ij}(t)$$

与 n 无关, 则称 $\{(Z_n, R_n), n \geq 0\}$ 是时齐的 (齐次的), 称 $\{Q_{ij}(t), i, j \in E\}$ 为半马尔可夫核. 时齐的马尔可夫更新过程的性质由它的半马尔可夫核完全确定. 如果没有特殊声明, 下面的马尔可夫更新过程都是时齐的.

令 $\lim_{t \rightarrow \infty} Q_{ij}(t) = P_{ij} (i, j \in E)$, 设 $\mathbf{P} = (P_{ij}) (i, j \in E)$. 由马尔可夫更新过程的定义可得 $\{Z_n, n \geq 0\}$ 是状态空间 E 上具有转移概率矩阵 $\mathbf{P}$ 的离散时间马尔可夫链 (Ross, 1996).

若令 $X(t) = Z_n$, 当 $T_n \leq t \leq T_{n+1}$ 时, 称 $\{X(t), t \geq 0\}$ 是与马尔可夫更新过程 $\{(Z_n, R_n), n \geq 0\}$ 相联系的半马尔可夫随机过程.

$X(t)$ 可以看成是过程在时刻 t 所处的状态. 过程在时刻 $T_1, T_2, \dots$ 发生状态转移. 在时刻 T_n 转入状态 Z_n , 在 Z_n 的逗留时间为 $T_{n+1} - T_n$, 它的分布依赖于正在访问的状态 Z_n 和下一步要访问的状态 Z_{n+1} . 相继访问的状态 $\{Z_n\}$ 组成一个马尔可夫链. 在已知 $Z_n (n = 0, 1, 2, \dots)$ 的条件下, 相继的状态逗留时间是条件独立的.

在马尔可夫随机过程中, 系统在每个状态的逗留时间服从指数分布. 由于指数分布的无记忆性, 系统在任意时刻 t 都具有马尔可夫性. 在半马尔可夫过程中, 系统在各个状态的逗留时间服从一般分布, 因此不是所有时刻 t 都是再生点, 只在状态转移时刻 $T_n (n \geq 0)$ 具有马尔可夫性. 当 $Q_{ij}(t)/P_{ij}(i, j \in E)$ 是仅和 i 有关的指数分布时, 半马尔可夫过程 $\{X(t), t \geq 0\}$ 成为马尔可夫过程. 当过程只有一个状态时, $\{R_n, n \geq 0\}$ 是独立同分布随机变量序列, 在这个特殊情形下, 马尔可夫更新过程成为更新过程. 因此半马尔可夫过程是马尔可夫过程和更新过程完美结合的产物, 它广泛应用于可靠性、风险评估、排队论、社会保险等领域.

在应用半马尔可夫过程解决实际问题时, 通常会遇到马尔可夫更新方程组

$$h_i(t) = g_i(t) + \sum_{j \in E} \int_0^t h_j(t) dQ_{ij}(u), \quad i, j \in E.$$

当已知 $g_i(t)$ 和 $Q_{ij}(t)$ 的条件下, 通常用 L 变换或 L-S 变换求解上述方程组.

3. 基于马尔可夫和半马尔可夫过程的多状态系统建模

马尔可夫和半马尔可夫可修系统作为两类最主要的可修系统, 一直是可靠性研究的热点. 许多学者把它们引入不同的领域, 进行可靠性评估、生产过程可靠性控制等. 同时结合实际问题和工程背景, 以两类系统为基础, 建立了大量的切合实际的模型.

曹晋华和程侃 (2006) 详细讨论了马尔可夫可修系统的一般模型, 归纳了用马尔可夫过程理论求解马尔可夫可修系统的一般方法, 给出了系统的可用度、可靠度、故障频度、开工时间、停工时间、周期等可靠性指标的计算方法. 同时讨论了单部件系统、串联系统、并联系统、表决系统、冷储备系统、温储备系统等几大类典型系统的可靠性问题, 列出了各种可靠性指标的计算公式. 对马尔可夫可修系统进行可靠性评估的关键步骤有两个: 一是如何定义系统的状态以使用马尔可夫随机过程描述其运行过程; 二是如何得到系统的转移率矩阵.

可用马尔可夫更新过程描述的可修系统称为半马尔可夫可修系统. 对半马尔可夫可修系统进行可靠性评估的关键步骤有两个: 一是如何定义系统的状态; 二是如何利用半马尔可夫核, 通过马尔可夫更新方程表示可用度、首次故障前时间和故障频度等可靠性指标. 曹晋华和程侃 (2006) 分析了两个同型部件的冷储备系统、两