

现代应用数学方法丛书 8

拟生灭过程与矩阵几何解

田乃硕 岳德权 著

科学出版社

2002

内 容 简 介

本书系统地论述了位相型分布、拟生灭过程和矩阵几何解的基本理论及其在复杂随机模型分析中的应用. 通过含有位相分布的排队、多服务台休假排队、随机环境、供应链等随机模型的分析, 着重介绍使用拟生灭过程建模和使用矩阵几何解进行指标分析的方法和技巧. 书中包含电子商务营运分析、高峰期分析、供应链管理等现代随机模型应用的数值例子.

本书可供高等院校数学、运筹与管理、应用概率论等专业本科生、研究生及教师阅读, 也可供企业管理和经济分析等领域的工程技术人员阅读参考.

图书在版编目 (CIP) 数据

拟生灭过程与矩阵几何解 / 田乃硕, 岳德权著. —北京: 科学出版社, 2002

(现代应用数学方法丛书 / 方开泰主编)

ISBN 7-03-010176-6

I. 拟… II. ①田…②岳… III. 运筹学-计算方法 IV. O22

中国版本图书馆 CIP 数据核字 (2002) 第 009613 号

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>

印刷

科学出版社发行 各地新华书店经销

*

2002 年 5 月第 一 版 开本: 850×1168 1/32

2002 年 5 月第一次印刷 印张: 3 7/8

印数: 1—3 000 字数: 96 000

定价: 8.00 元

(如有印装质量问题, 我社负责调换〈新欣〉)

《现代应用数学方法丛书》编委会

名誉主编 胡国定

主 编 方开泰

副 主 编 程 侃

编 委 (以姓氏笔画为序)

井竹君 方开泰 冯士雍

毕 颖 沈世镒 应隆安

杨德庄 周子康 赵凤治

顾基发 程 侃

总 序

应用数学的发展与自然科学和社会科学广泛的交叉和渗透密切相关。一方面，它为形形色色的物理、化学、生物、社会等现象提供描述和分析的数学工具。另一方面，这些实际问题的解决又为数学学科的发展提供了动力和永不枯竭的源泉。许多成功的应用数学方法，如解非线性方程的牛顿-高斯法、曲线拟合的最小二乘法、线性规划的单纯形法等，成了当今应用数学工作者手中不可缺少的工具。它们之所以有如此强大的生命力，原因在于方法本身有坚实的理论基础，同时又有鲜明的应用背景，能用于不同的领域。因此，成功的应用数学方法是理论联系实际的桥梁和纽带。

我国的数学要达到世界先进水平，要对人类有所贡献，重要的一点是要有一批独创的应用数学方法。《现代应用数学方法丛书》的出版，希望能为鼓励和促进我国的数学工作者创造或介绍更多的现代应用数学方法增加一个舞台。

这套丛书的宗旨是介绍现代应用数学方法。这些方法应该是目前世界上最先进的，或者是我国独创的，或者是国外已经普遍使用但国内知之甚少的方法。丛书着重阐明所介绍方法的应用背景和思想，避开深奥的数学论证，力求深入浅出、图文并茂，有数值及应用性的例子，使读者易于理解和使用。丛书要求短小精练，突出新的方法，不求齐全。书末所附的文献将指出方法的理论背景以及最近的进展，以便读者进一步深入研究。

本丛书的出版得到国家“天元”项目的资助，得到科学出版社的大力支持，是全体编委努力的成果。我们要特别感谢许多作者在百忙中为丛书撰写文稿，付出了辛勤的劳动。我们希望这套丛书的出版对我国应用数学的发展起到促进作用，衷心地希望丛

书成为广大读者的良师益友.

胡国定 (南开大学)

方开泰 (中国科学院应用数学研究所)

前 言

以指数分布为基础的经典生灭过程，是分析处理随机现象的基本工具之一。不仅在排队论、可靠性、库存论等学科的发展中起着重要作用，还广泛应用于工程技术、管理科学、生命科学、生态系统、人口理论、行为科学及社会科学等领域。

20 世纪中期以来，随着计算机通讯网络、柔性制造系统 (FMS)、异步转换模式 (ATM)、电子商务 (EC) 等高新技术领域的发展，提出了大量复杂随机系统的设计与控制问题。与传统随机模型不同，这些系统表现出多层次、多位相、变动参数、随机环境等复杂特性。经典生灭过程在处理这些问题时存在着极大的局限性。70 年代以来，Neuts 等系统地发展了结构矩阵分析方法，使随机模型分析由以指数分布为核心发展到广泛使用位相型 (phase type, 简记 PH) 分布的新阶段。经典生灭过程发展为拟生灭过程，导致生灭过程稳态分布的古老方法发展为矩阵几何解方法。这一变化为复杂随机模型的分析提供了强有力的工具。在随机模型的研究和应用中广泛使用这些新思想和新方法，已经成为当代流行的趋势。

本书的目的是系统地介绍拟生灭过程与矩阵几何解方法。我们将力求说明这些新方法产生的背景和来龙去脉。从简单的指数分布族出发类比地引出 PH 分布；从经典生灭过程的剖析入手直观地过渡到拟生灭过程；从 $M/M/c$ 排队系统的处理方法自然地发展到矩阵几何解。重点是阐明方法的理论实质和使用这些方法处理问题的技巧。我们回避了一些比较艰深的数学论证，但在多数情况下通过直观和类比的说明加以补救，从而保证了论述的连贯性和系统性。通过处理含 PH 分布的排队、各种休假策略的 $M/M/c$ 排队、随机环境下的排队及供应链模型，反复示范用拟

生灭过程为复杂系统建模的具体步骤和用矩阵几何解计算系统指标的方法. 同时, 这些复杂系统的分析结果也充分显示了拟生灭过程和矩阵几何解方法的强大威力. 书中还包含一些在电子商务营运分析、高峰期分析和供应链管理等领域应用的例子, 并对它们给出了详细的数值结果.

第一章是 PH 分布的简要介绍. 第二章引入拟生灭过程和矩阵几何解, 叙述了基本理论和方法. 第三、四、五、六章使用这些方法处理含有 PH 分布的排队、 $M/M/c$ 休假排队、随机环境下的排队及供应链中的随机模型. 在论述上努力避免涉及更多的数学理论和技巧. 阅读本书只需要 Markov 过程和矩阵分析的基础知识.

第一、二章中少量结果是作者首先给出的. 第四章对各种休假策略的 $M/M/c$ 排队系统的分析和条件随机分解理论都是作者近几年的研究成果, 其中只允许部分服务台休假的模型及其在电子商务营运中的应用, 是与美国 Western Washington 大学、加拿大 Simon Fraser 大学 (兼职) George Zhang (张哲) 教授合作完成的. 第六章使用排队论方法研究供应链中的随机模型, 主要是美国 Columbia 大学 David D. Yao (姚大卫) 教授及 Jing-sheng Song (宋京生) 教授开创的, 供应链分析已经成为当代随机模型研究的热点.

作者的研究连续地得到国家自然科学基金支持, 本书的出版又获“数学天元基金”资助, 特此表示衷心地感谢.

中国科学院数学与系统科学研究院徐光辉、程侃、曹晋华研究员, 长期以来对作者的研究工作给予诸多关怀和帮助, 谨向他们致以真诚的感谢.

作者 2001 年 4 月
于秦皇岛, 燕山大学

目 录

第一章 从指数分布到位相型分布	(1)
§ 1.1 指数分布族	(1)
§ 1.2 连续 PH 分布	(3)
§ 1.3 离散 PH 分布	(6)
§ 1.4 PH 分布的性质.....	(8)
§ 1.5 应用例子	(12)
第二章 从生灭过程到拟生灭过程	(16)
§ 2.1 从生灭过程谈起	(16)
§ 2.2 通向矩阵几何解—— $M/M/c$ 排队	(18)
§ 2.3 拟生灭过程	(22)
§ 2.4 矩阵几何解的性质	(26)
第三章 PH 分布的排队系统	(31)
§ 3.1 $M/PH/1$ 排队	(31)
§ 3.2 $PH/M/c$ 排队	(35)
第四章 服务员休假的 $M/M/c$ 排队	(39)
§ 4.1 休假排队系统	(39)
§ 4.2 同步休假 $M/M/c$ 排队	(41)
§ 4.3 异步休假 $M/M/c$ 排队	(48)
§ 4.4 部分服务台同步休假 $M/M/c$ 排队	(56)
§ 4.5 部分服务台异步休假 $M/M/c$ 排队	(63)
§ 4.6 应用例子——电子商务营运分析	(70)
第五章 随机环境	(77)
§ 5.1 Markov 环境	(77)
§ 5.2 随机环境下的 $M/M/1$ 排队	(79)
§ 5.3 高峰期分析	(82)
第六章 供应链中的随机模型	(87)
§ 6.1 供应链和订单装配系统	(87)

§ 6.2	ATO 的拟生灭过程模型	(91)
§ 6.3	ATO 系统的指标分析	(95)
§ 6.4	ATO 的 $M/G/\infty$ 排队群模型	(99)
§ 6.5	上、下界和优化问题	(104)
参考文献	(110)

第一章 从指数分布到位相型分布

§ 1.1 指数分布族

众所周知,对各种随机模型进行解析处理的主要障碍,是条件概率引起的高度复杂性.例如,要计算正在使用中的部件的剩余寿命,或正在接受服务顾客的剩余服务时间,我们需要对部件已用过多长时间,或服务从何时开始的信息有全面的了解.一方面,这些关于历史过程的详细信息在许多应用中很难确切知道;另一方面,既使有了这些信息,条件分布的表达式通常也极为复杂.经典随机模型理论中广泛使用指数分布的随机变量,通过指数分布的无后效性绕过这一困难.

称随机变量 X 服从参数 λ 的指数分布,是指它有分布函数和密度函数

$$F(x) = 1 - e^{-\lambda x}, \quad f(x) = \lambda e^{-\lambda x}, \quad x \geq 0.$$

X 的数学期望和 Laplace-Stieltjes 变换(LST)是

$$E(X) = \frac{1}{\lambda}, \quad \tilde{f}(s) = \int_0^{\infty} e^{-sx} f(x) dx = \frac{\lambda}{s + \lambda}, \quad \operatorname{Re} s \geq 0.$$

指数分布最独特的性质是无后效(无记忆)性.设随机变量 X 表示一个部件的使用寿命,对 $t > 0$ 和 $s > 0$,我们有

$$P\{X > s + t \mid X > t\} = \frac{1 - F(s + t)}{1 - F(t)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s}.$$

(1.1.1)

(1.1.1)式表明,无论部件已经使用了多长时间,其剩余寿命仍如 X 一样服从参数 λ 的指数分布.无后效性的本质是“完全随机性”(见 Feller, 1972 第一章),即部件失效完全随机地发生,不受使用时间和折旧行为的影响.当然,只要假定所研究的随机变量服从指

数分布,由条件概率产生的种种麻烦就荡然无存了.

让我们稍微扩大一下范围,考虑下列指数分布族.设 $0 \leq \alpha \leq 1$, 形如

$$F(x) = 1 - \alpha e^{-\lambda x}, \quad x \geq 0 \quad (1.1.2)$$

的全体分布函数称为(修正的)指数分布族.特别地,若 $\alpha = 0$, (1.1.2)退化为集中于原点的分布,其分布函数记为

$$U_0(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases}$$

$U_0(x)$ 属于指数分布族.若 $\alpha = 1$, (1.1.2)对应于标准的指数分布.当 $0 < \alpha < 1$ 时,注意到 $F(0) = 1 - \alpha$, $F(x)$ 在 $x = 0$ 集中着概率质量 $1 - \alpha$,并在 $(0, +\infty)$ 上有不完全概率密度函数

$$f(x) = \alpha \lambda e^{-\lambda x}, \quad x > 0. \quad (1.1.3)$$

这时, $F(x)$ 是一个混合型分布.它在 $x = 0$ 处具有离散分布的特性,在 $(0, +\infty)$ 上又具有连续随机变量分布的形态.指数分布族的密度函数和分布函数如图 1.1 所示.形如(1.1.2)的分布有数学期望和 LST:

$$E(X) = \frac{\alpha}{\lambda}, \quad \tilde{f}(s) = 1 - \alpha + \frac{\alpha \lambda}{s + \lambda}. \quad (1.1.4)$$

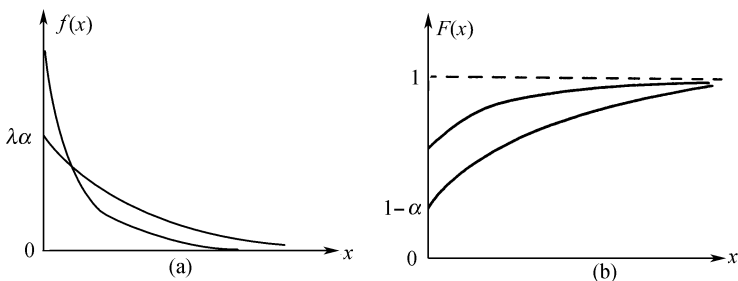


图 1.1 指数分布族的密度函数与分布函数

注意到(1.1.2)可改写成

$$F(x) = (1 - \alpha) U_0(x) + \alpha(1 - e^{-\lambda x}), \quad x \geq 0, \quad (1.1.5)$$

这表明指数分布族中的任何分布,都是集中于原点的分布与一个标准指数分布的混合(mixture). $F(x)$ 以概率 $(1-\alpha)$ 是集中于原点上的退化分布,以概率 α 是一个标准指数分布.如果 X 的分布函数如(1.1.2)所示,则 X 以概率 $1-\alpha$ 集中于原点,以概率 α 服从参数 λ 的标准指数分布,并具有指数分布的无后效性.

§ 1.2 连续 PH 分布

指数分布具有易于进行解析处理的优点,但也有明显的局限性.假定顾客服务时间、部件使用寿命服从指数分布,其不合理性是显而易见的.在广泛使用指数分布的经典随机模型分析中,我们在一定程度上牺牲了模型的真实性和实际应用价值,才换得了解析处理的方便.长期以来,随机模型分析中一直存在一个突出矛盾:一方面,离开指数分布,随机模型的定量分析就面临着实质性困难;另一方面,假定所研究的随机变量服从指数分布,在多数情况下将与客观事实大相径庭.

寻求保持指数分布易于进行解析处理的优点,又能包罗更广的新分布类的努力可以追溯到 Erlang (1918)的早期工作.他从指数分布的独立和出发构造了 Erlang 分布.在此后漫长的岁月里,许多作者尝试过各种混合型分布,例如广义 Erlang 分布,超指数分布等.这些分布都具有有理函数的 LST,这就推动了 Cox (1955)关于有理变换分布类的研究.但是, Cox 分布类又过于宽泛,失去了解析处理方便的优点.位相型(PH)分布正是有理变换分布类中能较好地保持指数分布方便性的一个子类.

PH分布是一个有限状态 Markov 过程吸收时间的分布. Jensen (1954)首先把这一分布用于一类经济模型. Neuts (1975)成功地发展了处理这类分布的矩阵分析技巧,其分布函数被表示成矩阵指数形式,清晰地显示出 PH 分布类与指数分布类的深刻类比.另一方面, PH 分布的不同位相适合于描述具有多层次和变动参数的复杂随机现象.因此, PH 分布类迅速成为当代随机模型分

析的有力工具.

考虑状态集 $\{1, \dots, m, m+1\}$ 上的 Markov 过程, 状态 $1, \dots, m$ 都是非常返的, 状态 $m+1$ 是吸收的. 过程的无穷小生成元可写成下列分块形式:

$$Q = \begin{pmatrix} T & T^0 \\ \mathbf{0} & 0 \end{pmatrix},$$

m 阶方阵 $T = (T_{ij})$ 满足 $T_{ii} < 0, T_{ij} \geq 0, i \neq j, 1 \leq i, j \leq m$. $T^0 = (T_1^0, \dots, T_m^0)^T$ 是非负列向量, 满足 $T\mathbf{e} + T^0 = \mathbf{0}$. 在本书中 \mathbf{e} 表示适当维数的元素全为 1 的列向量. 过程的初始概率是 (α, α_{m+1}) , 其中 $\alpha = (\alpha_1, \dots, \alpha_m)$, $\alpha\mathbf{e} + \alpha_{m+1} = 1$. 可以证明, 状态 $1, \dots, m$ 都是非常返的当且仅当矩阵 T 是非奇异的. 今后总设 T 是可逆方阵.

引理 1.1 Markov 过程 Q 以初始分布 (α, α_{m+1}) 出发直到吸收于状态 $m+1$ 为止的时间有分布函数

$$F(x) = 1 - \alpha \exp(Tx) \mathbf{e}, \quad x \geq 0. \quad (1.2.1)$$

证明 $v_j(x)$ 表示时刻 x 过程处在状态 j 上的绝对概率, $j = 1, \dots, m$. 常规的转移率分析可得

$$\begin{aligned} v_j(x + \Delta x) &= \sum_{k=1, k \neq j}^m v_k(x) T_{kj} \Delta x + (1 + T_{jj} \Delta x) v_j(x) \\ &\quad + o(\Delta x), \quad j = 1, \dots, m. \end{aligned} \quad (1.2.2)$$

由此给出

$$v_j'(x) = \sum_{k=1}^m v_k(x) T_{kj}, \quad j = 1, \dots, m.$$

引入行向量 $v(x) = (v_1(x), \dots, v_m(x))$, 上式可写成矩阵形式

$$v'(x) = v(x) T,$$

初始条件是 $v(0) = (\alpha_1, \dots, \alpha_m)$. 它的解是 $v(x) = \alpha \exp(Tx)$. 注意到 $v(x)\mathbf{e}$ 是时刻 x 尚未吸收的概率, 于是得到 (1.2.1). \square

分布函数 (1.2.1) 也称为矩阵指数分布, 它是指数分布族 (1.1.2) 从数值参数到矩阵参数的一种推广. 矩阵指数函数由形式幂级数定义:

$$\exp(\mathbf{T}x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \mathbf{T}^k.$$

可以证明,对任何实数 x 和方阵 \mathbf{T} 级数都是收敛的.矩阵指数函数 $\exp(\mathbf{T}x)$ 具有与指数函数 $e^{-\lambda x}$ 完全类似的性质,详见任何一本矩阵分析方面的著作,例如 Bellman(1960).

引理 1.2 a_j 表示吸收前过程在位相 j 上的平均逗留时间, $j=1, \dots, m$, 则

$$\mathbf{a} = (a_1, \dots, a_m) = -\alpha \mathbf{T}^{-1}. \quad (1.2.3)$$

证明 由引理 1 的证明,过程时刻 x 处在位相 j 的概率 $v_j(x)$ 是 $\alpha \exp(\mathbf{T}x)$ 的第 j 个分量, $j=1, \dots, m$. 引入印记

$$I_j(x) = \begin{cases} 1, & \text{时刻 } x \text{ 过程在位相 } j \text{ 上,} \\ 0, & \text{否则.} \end{cases}$$

于是 $E(I_j(x)) = v_j(x)$, 并且

$$a_j = \int_0^{\infty} E(I_j(x)) dx, \quad j = 1, \dots, m.$$

这就给出

$$\mathbf{a} = \alpha \int_0^{\infty} \exp(\mathbf{T}x) dx = -\alpha (\mathbf{T}^{-1}). \quad \square$$

定义 1.1 $[0, +\infty)$ 上的概率分布 $F(\cdot)$ 称为(连续)PH 分布,当且仅当它是一个有限状态 Markov 过程吸收时间的分布,并且形如(1.2.1). (α, \mathbf{T}) 称为它的 m 阶 PH 表示.

容易验证 $\mathbf{T}^* = \mathbf{T} + \mathbf{T}^0 \alpha$ 是 m 阶生成元.若 \mathbf{T}^* 是不可约的,称 (α, \mathbf{T}) 是不可约 PH 表示.今后总假设所用的 PH 表示是不可约的.由定义,下列事实是显然的.

1° $F(0) = 1 - \alpha \mathbf{e} = \alpha_{m+1}$. 若 α 是零向量, (1.2.1) 退化为 $U_0(x)$, 集中于原点的分布是一个 PH 分布; 若 $0 < \alpha \mathbf{e} < 1$, $F(x)$ 在 $x=0$ 处集中一个概率质量 α_{m+1} , 并在 $(0, +\infty)$ 上有不完全概率密度

$$f(x) = \alpha \exp(\mathbf{T}x) \mathbf{T}^0, \quad x \geq 0.$$

2° $F(x)$ 的 LST 是

$$\tilde{f}(x) = \int_0^{+\infty} e^{-sx} \alpha \exp(\mathbf{T}x) dx \mathbf{T}^0 = \alpha_{m+1} + \alpha(s\mathbf{I} - \mathbf{T})^{-1} \mathbf{T}^0. \quad (1.2.4)$$

3° PH 分布的 i 阶原点矩是

$$\mu_i = \int_0^{\infty} x^i \alpha \exp(\mathbf{T}x) dx \mathbf{T}^0 = (-1)^i i! \alpha \mathbf{T}^{-i} \mathbf{e}, \quad i = 1, 2, \dots.$$

一阶 PH 分布就是指数分布族(1.1.2), 广义 Erlang 分布和超指数分布都是 PH 分布. 我们只介绍了 PH 分布理论的最基本内容, 详见田乃硕等(1995)或 Neuts(1981)第二章, 田乃硕(2001)第一章. 表 1.1 列出指数分布族与 PH 分布类的类比关系.

表 1.1 从指数分布到连续 PH 分布

分布类	指数分布类	连续 PH 分布类
分布参数	$(\alpha, -\lambda); 0 \leq \alpha \leq 1, \lambda > 0$	$(\alpha, \mathbf{T}); 0 \leq \alpha \mathbf{e} \leq 1, \mathbf{T} \mathbf{e} < 0$
分布函数	$F(x) = 1 - \alpha e^{-\lambda x}$	$F(x) = 1 - \alpha \exp(\mathbf{T}x) \mathbf{e}$
原点跳跃	$F(0) = 1 - \alpha$	$F(0) = 1 - \alpha \mathbf{e} = \alpha_{m+1}$
密度函数	$f(x) = \alpha [e^{-\lambda x}] \lambda$	$f(x) = \alpha \exp(\mathbf{T}x) \mathbf{T}^0$
数学期望	$\alpha \lambda^{-1}$	$\alpha (\mathbf{T}^{-1}) \mathbf{e}$
LST	$\tilde{f}(s) = (1 - \alpha) + \alpha (s + \lambda)^{-1} \lambda$	$\tilde{f}(s) = \alpha_{m+1} + \alpha (s\mathbf{I} - \mathbf{T})^{-1} \mathbf{T}^0$

§ 1.3 离散 PH 分布

在离散随机变量中, 惟一具有无后效性的是服从标准几何分布的随机变量:

$$P\{X = j\} = (1 - p)p^j, \quad 0 < p < 1, \quad j = 0, 1, \dots. \quad (1.3.1)$$

数学期望和概率母函数(PGF)是

$$E(X) = \frac{1}{1 - p}, \quad X(z) = \frac{1 - p}{1 - zp}, \quad |z| \leq 1.$$

我们用 Bernoulli 试验中首次出现“成功”前的等待时间(试验次数)解释几何分布的无后效性. 进行独立重复试验, 每次试验出

现“成功”的概率是 $1-p$ ，“失败”的概率是 p . X 表示首次出现“成功”前的试验次数,也称为等待时间. X 的分布如(1.3.1).对正整数 m 和 n ,我们有

$$P\{X = m+n | x \geq n\} = \frac{(1-p)p^{m+n}}{p^n} = (1-p)p^m. \quad (1.3.2)$$

(1.3.2)表明,无论我们已经等待了多长时间,剩余等待时间都与 X 一样服从参数 p 的几何分布.这一性质称为几何分布的无后效性.

现在讨论几何分布族.设 $0 \leq \alpha \leq 1$,非负整数集上的离散分布

$$P\{X = k\} = \begin{cases} 1-\alpha, & k=0, \\ \alpha p^{k-1}(1-p), & k \geq 1 \end{cases} \quad (1.3.3)$$

的全体,称为(修正的)几何分布族. $\alpha=0$ 时,(1.3.3)退化为集中于原点的分布; $\alpha=p$ 时,(1.3.3)给出标准的几何分布; $\alpha=1$ 时,(1.3.3)为对应于自然数集上的几何分布.分布(1.3.3)有均值和 PGF:

$$E(X) = \frac{\alpha}{1-p}, \quad X(z) = 1 - \alpha + \alpha z \frac{1-p}{1-zp}. \quad (1.3.4)$$

离散 PH 分布是几何分布族由数值参数到矩阵参数的推广.考虑有限状态集 $\{1, \dots, m, m+1\}$ 上的 Markov 链, $m+1$ 是吸收状态.转移概率矩阵表示成分块形式

$$\tilde{P} = \begin{pmatrix} T & T^0 \\ \mathbf{0} & 1 \end{pmatrix},$$

其中, m 阶方阵 T 是子随机阵,满足 $T_{ij} \geq 0$, $Te \leq e$.非负列向量 $T^0 = (I - T)e$, $I - T$ 是非奇异的. Markov 链的初始分布是 (α, α_{m+1}) ,满足 $\alpha e + \alpha_{m+1} = 1$.

定义 1.2 非负整数集上的离散概率分布 $\{p_k, k \geq 0\}$ 称为离散 PH 分布,当且仅当它是上述 Markov 链吸收前转移步数 X 的分布,即

$$p_k = P\{X = k\} = \begin{cases} \alpha_{m+1}, & k=0, \\ \alpha T^{k-1} T^0, & k \geq 1. \end{cases} \quad (1.3.5)$$

称 (α, T) 是 X 或分布 $\{p_k, k \geq 0\}$ 的 m 阶 PH 表示.

形如 (1.3.5) 的分布也称为矩阵几何分布, 它是几何分布族 (1.3.3) 从数值参数到矩阵参数的一种推广. 分布 (1.3.5) 的数学期望、PGF 与 k 阶阶乘矩是

$$\begin{aligned} E(X) &= \alpha(I - T)^{-1}e, \quad X(z) = \alpha_{m+1} + z\alpha(I - zT)^{-1}T^0, \\ \mu_k &= X^{(k)}(1) = k! \alpha T^{k-1}(I - T)^{-k}e. \end{aligned} \quad (1.3.6)$$

表 1.2 给出离散 PH 分布与几何分布族的类比.

表 1.2 从几何分布到离散 PH 分布

分布类	几何分布族	离散 PH 分布类
分布参数	$(\alpha, p); 0 \leq \alpha \leq 1, 0 < p < 1$	$(\alpha, T); 0 \leq \alpha e \leq 1, Te < e$
分布形式	$p_0 = 1 - \alpha$ $p_k = \alpha p^{k-1}(1-p), k \geq 1$	$p_0 = 1 - \alpha e = \alpha_{m+1}$ $p_k = \alpha T^{k-1}(I - T)e, k \geq 1$
数学期望	$\alpha(1-p)^{-1}$	$\alpha(I - T)^{-1}e$
PGF	$(1-\alpha) + z\alpha(1-zp)^{-1}(1-p)$	$\alpha_{m+1} + z\alpha(I - zT)^{-1}(I - T)e$

§ 1.4 PH 分布的性质

PH 分布除了保持指数(几何)分布易于进行解析处理的优点, 还有许多其他良好的性质. 首先是封闭性: 若干个 PH 分布的运算结果通常是一个新的 PH 分布; 含有 PH 分布的随机模型, 其指标通常也服从 PH 分布. 封闭性在模型的理论分析和实际应用中都带来了极大的方便.

定理 1.1 F 和 G 是相互独立的连续(离散)PH 分布, 分别有 m, n 阶 PH 表示 (α, T) 和 (β, S) , 则卷积 $F * G$ 是 $m+n$ 阶 PH 分布, 有 PH 表示 (γ, L) , 这里

$$\begin{aligned} \gamma &= (\alpha, \alpha_{m+1}\beta), \quad \gamma_{m+n+1} = \alpha_{m+1}\beta_{n+1}, \\ L &= \begin{pmatrix} T & T^0\beta \\ \mathbf{0} & S \end{pmatrix}, \quad L^0 = \begin{pmatrix} \beta_{n+1}T^0 \\ S^0 \end{pmatrix}. \end{aligned} \quad (1.4.1)$$

证明 使用矩阵分块求逆法则,有

$$(sI-L)^{-1} = \begin{pmatrix} (sI-T)^{-1} & (sI-T)^{-1} T^0 \beta(sI-S)^{-1} \\ \mathbf{0} & (sI-S)^{-1} \end{pmatrix},$$

代入 PH 表示 (γ, L) 的 LST 表达式,我们有

$$\begin{aligned} \gamma_{m+n+1} + \gamma(sI-L)^{-1} L^0 \\ &= \alpha_{m+1} \beta_{n+1} + \beta_{n+1} \alpha(sI-T)^{-1} T^0 + \alpha_{m+1} \beta(sI-S)^{-1} S^0 \\ &\quad + \alpha_{m+1} \alpha(sI-T)^{-1} T^0 \beta(sI-S)^{-1} S^0 \\ &= \{ \alpha_{m+1} + \alpha(sI-T)^{-1} T^0 \} \{ \beta_{n+1} + \beta(sI-S)^{-1} S^0 \}. \quad \square \end{aligned}$$

定理 1.2 F 是连续 PH 分布,有 m 阶 PH 表示 (α, T) 和均值 $\mu = \alpha(-T^{-1})e$, $\llbracket WTBX \rrbracket$ 则 F 的平稳剩余寿命分布

$$\tilde{F}(x) = \frac{1}{\mu} \int_0^x (1 - F(u)) du, \quad x \geq 0,$$

有 PH 表示 (π, T) , 其中 $\pi = \mu^{-1} \alpha(-T^{-1})$ 是生成元 $T^* = T + T^0 \alpha$ 的平稳概率向量.

证明 对 $\tilde{F}(x)$, 我们有

$$\begin{aligned} \tilde{F}(x) &= \frac{1}{\mu} \int_0^x \alpha \exp(Tu) du e \\ &= \frac{\alpha}{\mu} T^{-1} [\exp(Tx) - I] e \\ &= 1 - \pi \exp(Tx) e, \quad x \geq 0. \end{aligned}$$

容易验证 $\pi T^* = 0$ 和 $\pi e = 1$, 因此 π 是生成元 T^* 的平稳概率向量. \square

定理 1.3 PH 分布的有限混合仍是 PH 分布. (p_1, \dots, p_k) 是一个离散概率分布, F_j 有 PH 表示 $(\alpha(j), T(j))$, 则有限混合 $F = p_1 F_1 + \dots + p_k F_k$ 有 PH 表示 (α, T) , 其中

$$\begin{aligned} \alpha &= (p_1 \alpha(1), \dots, p_k \alpha(k)), \\ T &= \text{diag}(T(1), \dots, T(k)). \end{aligned}$$

证明 显然.

以后经常用到集中于原点的分布与一个 PH 分布的混合. 设 $0 < p < 1$, F 有 m 阶 PH 表示 (α, T) , 则混合

$$F_0 = pU_0 + (1-p)F$$

仍服从 m 阶 PH 分布,有 PH 表示 (γ, T) . 这里

$$\gamma = (1-p)\alpha, \quad \gamma_{m+1} = p + (1-p)\alpha_{m+1}. \quad (1.4.2)$$

PH 分布的无限混合一般不再服从 PH 分布,一个重要的例外是连续 PH 分布卷积的离散 PH 混合. 设 $\{p_k, k \geq 0\}$ 是 n 阶离散 PH 分布,有 PH 表示 (β, S) ; F 是 m 阶连续 PH 分布,有不可约表示 (α, T) , 则无限混合

$$F_s(x) = \sum_{k=0}^{\infty} p_k F^{(k)}(x)$$

服从 mn 阶 PH 分布,其中 $F^{(k)}(x)$ 是 F 的 k 重卷积. 给出 F_s 的 PH 表示需使用矩阵的 Kronecker 乘积, 详见 Neuts (1981) 定理 2.2.5. 在可靠性理论中经常使用下列极值分布: 设 X 与 Y 是相互独立的随机变量, 有分布函数 F 和 G , 则它们的极值 $\max(X, Y)$ 和 $\min(X, Y)$ 有分布函数

$$F_1(x) = F(x)G(x)$$

和

$$F_2(x) = 1 - (1 - F(x))(1 - G(x)).$$

如果 F 和 G 分别是 m 和 n 阶 PH 分布, 则 F_1 和 F_2 分别是 $mn + m + n$ 和 mn 阶 PH 分布. 在这些性质的证明中涉及复杂的 Kronecker 乘积运算, 详见田乃硕 (2001) 第 1.4 节.

在许多应用中需要讨论一个随机长度 $[0, X]$ 上强度 λ 的 Poisson 计数过程, X 是独立于 Poisson 过程的随机变量, 有分布函数 $F(x)$. N 表示随机区间 $[0, X]$ 上 Poisson 过程发生事件的次数, 我们有

$$a_k = P\{N = k\} = \int_0^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} dF(t), \quad k = 0, 1, \dots, \quad (1.4.3)$$

对分布 $\{a_k, k \geq 0\}$ 成立着下列封闭性定理.

定理 1.4 若 F 有 m 阶 PH 表示 (α, T) , 则 $\{a_k, k \geq 0\}$ 是 m 阶离散 PH 分布, 有 PH 表示 (γ, L) . 这里

$$\begin{aligned}\gamma &= \lambda \alpha (\lambda \mathbf{I} - \mathbf{T})^{-1}, \quad \gamma_{m+1} = \alpha_{m+1} + \alpha (\lambda \mathbf{I} - \mathbf{T})^{-1} \mathbf{T}^0, \\ \mathbf{L} &= \lambda (\lambda \mathbf{L} - \mathbf{T})^{-1}, \quad \mathbf{L}^0 = (\lambda \mathbf{I} - \mathbf{T})^{-1} \mathbf{T}^0.\end{aligned}\quad (1.4.4)$$

证明 容易验证 L 是子随机阵. 首先

$$\begin{aligned}a_0 &= \int_0^\infty e^{-\lambda t} dF(t) \\ &= \alpha_{m+1} + \alpha \int_0^\infty \exp[-(\lambda \mathbf{I} + \mathbf{T})t] dt \mathbf{T}^0 \\ &= \alpha_{m+1} + \alpha (\lambda \mathbf{I} - \mathbf{T})^{-1} \mathbf{T}^0 = \gamma_{m+1}.\end{aligned}$$

现在计算 $\{a_k, k \geq 0\}$ 的 PGF, 我们有

$$\begin{aligned}A(z) &= \gamma_{m+1} + \sum_{k=1}^\infty \int_0^\infty \frac{(\lambda t)^k}{k!} z^k e^{-\lambda t} \alpha \exp(\mathbf{T}t) dt \mathbf{T}^0 \\ &= \gamma_{m+1} + \alpha \int_0^\infty \exp[-(\lambda(1-z)\mathbf{I} - \mathbf{T})t] dt \mathbf{T}^0 \\ &\quad - \alpha (\lambda \mathbf{I} - \mathbf{T})^{-1} \mathbf{T}^0 \\ &= \gamma_{m+1} - \alpha (\lambda \mathbf{I} - \mathbf{T})^{-1} \mathbf{T}^0 + \alpha [\lambda(1-z)\mathbf{I} - \mathbf{T}]^{-1} \mathbf{T}^0 \\ &= \gamma_{m+1} + \alpha (\lambda \mathbf{I} - \mathbf{T})^{-1} \{-\mathbf{I} + \lambda z (\lambda \mathbf{I} - \mathbf{T})^{-1} \\ &\quad + \mathbf{I}\} [\mathbf{I} - \lambda z (\lambda \mathbf{I} - \mathbf{T})^{-1}]^{-1} \mathbf{T}^0 \\ &= \gamma_{m+1} + z \gamma (\mathbf{I} - z \mathbf{L})^{-1} \mathbf{L}^0.\end{aligned}\quad \square$$

定理 1.4 的离散时间类比是考虑 $\{0, 1, \dots, N\}$ 上的二项分布, 其中 N 是一个随机变量. 设 X 的分布是

$$b_k = P\{X = k\} = \binom{N}{k} p^k (1-p)^{N-k}, \quad 0 \leq k \leq N, N = 0, 1, \dots, \quad (1.4.5)$$

其中 $0 < p < 1$, 并记 $\bar{p} = 1 - p$.

定理 1.5 若 N 有 m 阶离散 PH 表示 (α, \mathbf{T}) , 则 $\{b_k, k \geq 0\}$ 服从 m 阶离散 PH 分布, 有不可约表示 (γ, \mathbf{L}) . 这里

$$\begin{aligned}\gamma &= p \alpha (\mathbf{I} - \bar{p} \mathbf{T})^{-1}, \quad \gamma_{m+1} = \alpha_{m+1} + \alpha \bar{p} (\mathbf{I} - \bar{p} \mathbf{T})^{-1} \mathbf{T}^0, \\ \mathbf{L} &= p \mathbf{T} (\mathbf{I} - \bar{p} \mathbf{T})^{-1}, \quad \mathbf{L}^0 = (\mathbf{I} - \bar{p} \mathbf{T})^{-1} \mathbf{T}^0.\end{aligned}\quad (1.4.6)$$

证明 容易验证 L 是子随机阵. X 的无条件分布是

$$b_k = P\{X = k\} = \sum_{n=k}^{\infty} P\{N = n\} \binom{n}{k} p^k p^{n-k}, \quad k = 0, 1, \dots$$

取 PGF, 我们有

$$X(z) = \sum_{k=0}^{\infty} z^k \sum_{n=k}^{\infty} P\{N = n\} \binom{n}{k} p^k p^{n-k} = N(1 - p(1 - z)),$$

其中 $N(z)$ 是 N 的 PGF:

$$N(z) = \alpha_{m+1} + z \alpha (I - zT)^{-1} T^0. \quad (1.4.7)$$

将 (1.4.7) 中的 z 换成 $1 - p(1 - z)$, 经一系列矩阵运算给出

$$X(x) = \gamma_{m+1} + z \gamma (I - zL)^{-1} L^0. \quad \square$$

除封闭性外, PH 分布类还有另外一些优点. 可以证明, PH 分布类在 $[0, +\infty)$ 上一切概率分布的类中稠密. 这意味着, 无论我们研究什么样的随机分布, 总可选择一个适当的 PH 分布来代替它, 并且使误差达到我们所要求的精度. 指数分布易于处理, 但在大多数情况下与客观实际相差甚远; PH 分布在很大程度上保持了指数的分布容易进行解析处理的优点, 又可与客观实际任意接近. 这正是 PH 分布类的诱人之处, 也是 PH 分布迅速普及的重要原因之一. 然而, PH 分布的这种一般性迄今还只是理论上的, 在具体应用中怎样构造一个 PH 分布来逼近已知分布? 如何用 PH 分布拟合统计数据? 这将涉及到 PH 分布的统计理论, 这方面的研究还不多.

PH 分布都是渐近指数的. 对任何一个 PH 分布 F , 都存在常数 $K > 0$ 和 $\eta > 0$, 使

$$1 - F(x) = Ke^{-\eta x} + o(e^{-\eta x}), \quad x \rightarrow +\infty.$$

关于 PH 分布及其在随机模型研究中的应用, 见田乃硕与李泉林 (1995). PH 分布理论的一般性论述见 Neuts (1981) 第二章或田乃硕 (2001) 第一章. 对 PH 分布类理论研究的一些新进展见 O'Cinneide (1990, 1993) 等.

§ 1.5 应用例子

本节通过简单例子说明如何用 PH 分布刻画复杂随机现象.

例 1 (有折旧行为的部件寿命) 部件寿命按折旧程度分三个阶段.在第 1、2、3 阶段上的使用时间分别服从均值 $\frac{1}{10}$, $\frac{1}{20}$ 和 $\frac{1}{30}$ (年)的指数分布.处在寿命第一阶段的部件以 0.2 的概率失效,并以 0.8 的概率因老化进入第二阶段.处在寿命第二阶段的部件,以 0.4 的概率失效,以 0.6 的概率因老化进入寿命的第三阶段.处在第三阶段的部件老化后将全部失效或更新.另一方面,由于内在的质量原因,一部件在投入使用时处在第 1, 2, 3 阶段的概率分别是 0.62, 0.28 和 0.09, 并以 0.01 的概率根本不能使用.部件寿命折旧行为如图 1.2 所示.以 X 表任取一部件的使用寿命, X 服从一

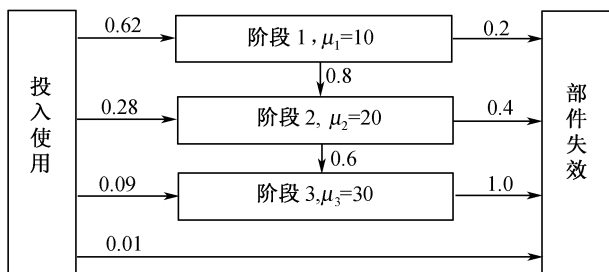


图 1.2 部件寿命的折旧行为

个 3 阶 PH 分布,有 PH 表示 (α, T) , 其中

$$\alpha = (\alpha_1, \alpha_2, \alpha_3) = (0.62, 0.28, 0.09), \quad \alpha_4 = 0.01,$$

$$T_{11} = -10, \quad T_{12} = 10 \times 0.8, \quad T_{13} = 0, \quad T_1^0 = 10 \times 0.2,$$

$$T_{21} = 0, \quad T_{22} = -20, \quad T_{23} = 20 \times 0.6, \quad T_2^0 = 20 \times 0.4,$$

$$T_{31} = 0, \quad T_{32} = 0, \quad T_{33} = -30, \quad T_3^0 = 30$$

于是

$$T = \begin{pmatrix} -10 & 8 & 0 \\ & -20 & 12 \\ & & -30 \end{pmatrix}, \quad T^0 = \begin{pmatrix} 2 \\ 8 \\ 30 \end{pmatrix}.$$

平均使用寿命是

$$E(X) = \alpha(-T^{-1})e = 0.11932(\text{年}).$$

例 2 (柔性制造系统中的加工时间) 柔性制造系统中某产品加工流程如图 1.3 所示. 加工从第一道工序开始, 并在每一工序加工过程中自动测试各种指标, 以决定进一步加工的需要. 在完成第一道工序后, 有 10% 的半成品可送到第三工序稍作补充加工即可出厂, 另有 90% 半成品需送第二道工序进一步加工. 第二道工序

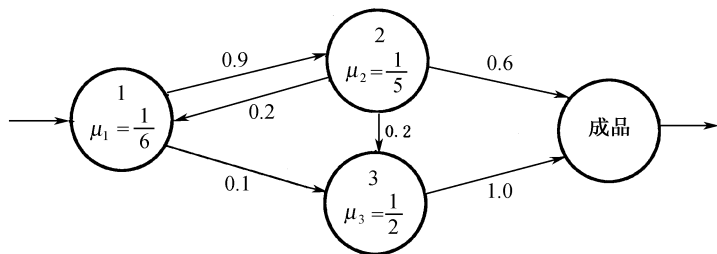


图 1.3 柔性制造系统中某产品加工流程

上的半成品加工后有 60% 可达到成品要求, 20% 需退回第一道工序返工, 另有 20% 需送到第三工序进行补充加工后达到成品要求. 产品在工序 1、2、3 的加工时间依次服从数学期望为 6, 5, 2(分钟)的指数分布. 为了简单, 假定产品到达每一工序都不需要等待. 以 X 表示一个产品的加工时间, X 服从 3 阶 PH 分布, 有 PH 表示 (α, T) :

$$\alpha = (\alpha_1, \alpha_2, \alpha_3) = (1, 0, 0), \quad \alpha_1 = 0,$$

$$T_{11} = -\frac{1}{6}, \quad T_{12} = 0.9 \times \frac{1}{6}, \quad T_{13} = 0.1 \times \frac{1}{6}, \quad T_1^0 = 0,$$

$$T_{21} = 0.2 \times \frac{1}{5}, \quad T_{22} = -\frac{1}{5}, \quad T_{23} = 0.2 \times \frac{1}{5}, \quad T_2^0 = 0.6 \times \frac{1}{5},$$

$$T_{31} = 0, \quad T_{32} = 0, \quad T_{33} = -2, \quad T_3^0 = 2.$$

于是

$$\mathbf{T} = \begin{pmatrix} -\frac{1}{6} & \frac{3}{20} & \frac{1}{60} \\ \frac{1}{25} & -\frac{1}{5} & \frac{1}{25} \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}, \quad \mathbf{T}^0 = \begin{pmatrix} 0 \\ \frac{3}{25} \\ \frac{1}{2} \end{pmatrix}.$$

产品的平均加工时间是

$$E(X) = \alpha(-\mathbf{T}^{-1})\mathbf{e} = 13.4878(\text{分}).$$

第二章 从生灭过程到拟生灭过程

§ 2.1 从生灭过程谈起

经典生灭过程的研究和应用已经有近百年的历史,不仅形成了相当完整的理论体系,而且广泛应用于自然科学、工程技术、生命科学、人口理论、行为科学及社会科学各种领域.我国学者王梓坤等在生灭过程理论研究中曾做出重要贡献,见王梓坤(1980)、侯振廷(2000)等.

为了引入拟生灭过程的概念和术语,并说明其来龙去脉,我们简要地叙述经典生灭过程的一些最基本内容.

非负整数集 $\{0, 1, \dots\}$ 上的 Markov 过程 $\{X(t), t \geq 0\}$,如果其无穷小生成元有下列三对角形式:

$$\tilde{Q} = \begin{pmatrix} -\lambda_0 & \lambda_0 & & & & \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & & & \\ & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \mu_n & -(\lambda_n + \mu_n) & \lambda_n \\ & & & & \dots & \dots \end{pmatrix}, \quad (2.1.1)$$

则称 $\{X(t), t \geq 0\}$ 是一个生灭过程. $\{\lambda_n, n \geq 0\}$ 称为生率, $\{\mu_n, n \geq 1\}$ 称为灭率.引入瞬时状态概率

$$p_j(t) = P\{X(t) = j\}, \quad j = 0, 1, \dots.$$

生灭过程的最基本特征是:在充分短的时间内,过程几乎只能转移到相邻的状态上,即当 $\Delta t \rightarrow 0$ 时,有

$$P\{X(t+\Delta t)=k \mid X(t)=j\} = \begin{cases} \lambda_j \Delta t + o(\Delta t), & k=j+1, \\ \mu_j \Delta t + o(\Delta t), & k=j-1, j \geq 1, \\ 1 - (\lambda_j + \mu_j) \Delta t + o(\Delta t), & k=j, \\ o(\Delta t), & |k-j| \geq 2. \end{cases} \quad (2.1.2)$$

由(2.1.2)出发,容易建立瞬态概率 $p_j(t)$ 的下列微分方程:

$$\begin{cases} p_0'(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t), \\ p_j'(t) = \lambda_{j-1} p_{j-1}(t) - (\lambda_j + \mu_j) p_j(t) + \mu_{j+1} p_{j+1}(t), & j \geq 1. \end{cases} \quad (2.1.3)$$

(2.1.3)称为生灭过程的 Kolmogorov 方程.在给定的初始条件下求解微分方程(2.1.3)是十分困难的.在大多数应用中,我们更关心过程运行相当长时间以后的形态,如果极限

$$p_j = \lim_{t \rightarrow \infty} p_j(t), \quad j=0,1,\dots \quad (2.1.4)$$

存在,并且 $p_0 + p_1 + \dots = 1$,则称过程是正常返的,并称 $\{p_j, j \geq 0\}$ 为过程平稳概率分布或稳态分布.经典生灭过程的重要问题之一是如何从生成元(2.1.1)出发,判断过程的正常返性并给出它的稳态分布.如果稳态分布存在,由(2.1.4),必有

$$\lim_{t \rightarrow \infty} p_j'(t) = 0, \quad j=0,1,\dots$$

在(2.1.3)的两端取极限,给出

$$\begin{cases} \lambda_0 p_0 - \mu_1 p_1 = 0, \\ \lambda_{j-1} p_{j-1} - (\lambda_j + \mu_j) p_j + \mu_{j+1} p_{j+1} = 0, & j \geq 1. \end{cases} \quad (2.1.5)$$

如果把稳态分布写成无穷维向量 $\pi = (p_0, p_1, \dots)$, (2.1.5)可改写成更紧凑的矩阵形式

$$\pi \tilde{Q} = \mathbf{0}, \quad \pi e = 1. \quad (2.1.6)$$

(2.1.5)和(2.1.6)统称为过程的平衡方程(及正规化条件).求解差分方程(2.1.5)可给出过程的平稳概率分布.另一个更直接的方法由 Karlin 与 McGregor 给出:

定理2.1 (Karlin-McGregor) 生灭过程(2.1.1)正常返当

且仅当

$$\sum_{k=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k} < \infty,$$

过程的稳态分布可表为

$$p_0 = K, \quad p_j = \frac{\lambda_0 \cdots \lambda_{j-1}}{\mu_1 \cdots \mu_j} K, \quad j \geq 1.$$

其中

$$K = \left[1 + \sum_{k=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k} \right]^{-1}.$$

我们只罗列一些最基本的事实,关于经典生灭过程的理论和应用,国内外已出版了许多著作.理论研究的专著见王梓坤(1980)、侯振廷(2000);以应用为目的的一般论述见赖特(1979)、费勒(1994);在排队论、可靠性、库存论及相关随机模型中的应用,见徐光辉(1988)、曹晋华与程侃(1986)、邓永录(1994)等.在生命科学中的应用见蒋庆琅(1987),在生态系统中的应用见皮洛(1988),在行为科学及社会科学中的应用见赖特(1979)和 Bhat (1984).

§ 2.2 通向矩阵几何解——M/M/c 排队

考虑一个二维 Markov 过程 $\{X(t), J(t)\}$, 有状态空间

$$\Omega = \{(k, j): k \geq 0, j = 1, 2, \dots, m\}.$$

状态集 $\{(k, 1), \dots, (k, m)\}$ 称为水平 $k, k \geq 0$. 如果将状态按字典序排列后, 其生成元可写成下列分块三对角形式:

$$\tilde{Q} = \begin{pmatrix} A_0 & C_0 & & & \\ B_1 & A_1 & C_1 & & \\ & B_2 & A_2 & C_2 & \\ & & \ddots & \ddots & \ddots \\ & & & B_n & A_n & C_n \\ & & & & \ddots & \ddots & \ddots \end{pmatrix}, \quad (2.2.1)$$

则称 $\{X(t), J(t)\}$ 是一个拟生灭过程 (quasi birth and death process, 简记 QBD), 其中所有子块是 m 阶方阵, $A_k, k \geq 0$, 有负的对角线元素和非负的非对角线元素, 并且行和非正. B_k 和 C_k 都是非负阵, 满足

$$(A_0 + C_0)e = (A_k + B_k + C_k)e = 0, \quad k \geq 1.$$

由 (2.2.1) 与 (2.1.1) 的比较可知, QBD 是经典生灭过程从一维状态空间到二维状态空间的推广. 经典生灭过程的状态 $X(t) = k$, 现在被分解成 m 个子状态 $\{(k, 1), \dots, (k, m)\}$. 这一结构顺应了刻画过程在多层次、多位相及变动参数情况下演化的要求. Evans (1967) 首先在排队论分析中使用这类过程, Wallace (1969) 在计算机系统研究中也遇到形如 (2.2.1) 的过程, 并由他引入了“拟生灭过程”这一术语. 20 世纪 70 年代以来, Neuts (1981) 等系统地发展了处理 QBD 的矩阵分析方法, 有力地推动了 QBD 过程在随机模型分析中的广泛应用.

对于形如 (2.2.1) 的一般 QBD, 尚未建立起平行于经典生灭过程的理论和处理方法. 关于经典生灭过程平稳分布的 Karlin-McGregor 定理, 目前还不能推广到 QBD 过程. 迄今广泛使用的只是一类特殊的 QBD, 即生成元 \tilde{Q} 中的子块从某一个水平开始不再发生变化的情形. 这时, 过程生成元形如:

$$\tilde{Q} = \begin{pmatrix} A_0 & C_0 & & & \\ B_1 & A_1 & C_1 & & \\ & \ddots & \ddots & \ddots & \\ & & B_{c-1} & A_{c-1} & C_{c-1} \\ & & & B & A & C \\ & & & & B & A & C \\ & & & & & \ddots & \ddots & \ddots \end{pmatrix}. \quad (2.2.2)$$

处理 QBD 的方法, 完全依赖于生成元 \tilde{Q} 的特殊结构. 为了更自然地引入这种方法, 我们注意到在经典生灭过程模型中与 (2.2.2) 对

应的典型例子是 $M/M/c$ 排队系统. 下面使用一种特殊的技巧处理 $M/M/c$ 排队系统, 这种处理自然地导致了对 (2.2.2) 型 QBD 解析分析的矩阵几何解方法.

在到达率 λ 和服务率 μ 的 $M/M/c$ 排队系统 (见徐光辉, 1988) 中, $L(t)$ 表示时刻 t 系统内顾客数, 也称为系统的队长. $\{L(t), t \geq 0\}$ 是一个经典生灭过程, 其无穷小生成元是

$$\tilde{Q} = \begin{pmatrix} -\lambda & \lambda & & & & \\ \mu & -(\lambda + \mu) & \lambda & & & \\ & 2\mu & -(\lambda + 2\mu) & \lambda & & \\ & & \ddots & \ddots & & \\ & & & (c-1)\mu & -[\lambda + (c-1)\mu] & \lambda \\ & & & & c\mu & -(\lambda + c\mu) & \lambda \\ & & & & & c\mu & -(\lambda + c\mu) & \lambda \\ & & & & & & \ddots & \ddots & \ddots \end{pmatrix}. \quad (2.2.3)$$

在 (2.2.3) 中, 自状态 $k = c$ 以后生率和灭率都不再随状态变化. 过程的状态转移如图 2.1 所示. 设 $\rho = \lambda(c\mu)^{-1} < 1$, 过程是正常返的. 记

$$\pi_k = \lim_{t \rightarrow \infty} P\{L(t) = k\}, \quad k \geq 0,$$

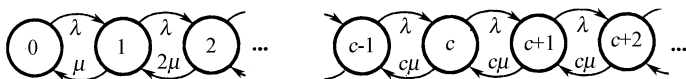


图 2.1 $M/M/c$ 排队的状态转移

当 $k > c$ 时, 平衡方程 (2.1.6) 和生成元 (2.2.3) 给出

$$\lambda \pi_{k-1} - (\lambda + c\mu) \pi_k + c\mu \pi_{k+1} = 0, \quad k > 0. \quad (2.2.4)$$

现在考虑二次方程

$$c\mu z^2 - (\lambda + c\mu)z + \lambda = 0, \quad (2.2.5)$$

它的两个根是 $z = \rho = \lambda(c\mu)^{-1} < 1$ 和 $z = 1$. $z = \rho$ 是方程 (2.2.5) 的最小非负解. 代入 (2.2.4) 验证, 可知 $\pi_k = K\rho^{k-c}$, $k > c$, 满足

(2.2.4), 其中 K 是待定常数. 在(2.2.4)中, 取 $k=c+1$, 给出

$$\lambda\pi_c - \lambda K + (\lambda K - (\lambda + c\mu)K\rho + c\mu K\rho^2) = \lambda(\pi_c - K) = 0,$$

于是 $K = \pi_c$. 对 $k \geq c$ 时, 我们有

$$\pi_k = \pi_c \rho^{k-c}, \quad k \geq c. \quad (2.2.6)$$

现在, 只要再确定 $\pi_0, \pi_1, \dots, \pi_c$. 由平衡方程(2.1.6)及(2.2.3), 并注意到 $\pi_{c+1} = \pi_c \rho$, 得到

$$\begin{cases} -\lambda\pi_0 + \mu\pi_1 = 0, \\ \lambda\pi_{k-1} - (\lambda + k\mu)\pi_k + (k+1)\mu\pi_{k+1} = 0, \quad 1 \leq k \leq c-1, \\ \lambda\pi_{c-1} + [-(\lambda + c\mu) + c\mu\rho]\pi_c = 0. \end{cases}$$

如果引入 $c+1$ 阶方阵

$$B[\rho] = \begin{pmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda + \mu) & \lambda & & \\ & \ddots & \ddots & \ddots & \\ & & (c-1)\mu - [\lambda + (c-1)\mu] & \lambda & \\ & & c\mu & -(\lambda + c\mu) + c\mu\rho & \end{pmatrix}. \quad (2.2.7)$$

我们并不算出 $c\mu\rho = \lambda$, 是为了在后面的讨论中加以比较. π_0, \dots, π_c 满足

$$(\pi_0, \dots, \pi_c) B[\rho] = \mathbf{0}. \quad (2.2.8)$$

求解方程组(2.2.8)并应用表达式(2.2.6), 给出经典 $M/M/c$ 排队系统的稳态分布

$$\pi_k = \begin{cases} K \frac{1}{k!} \left(\frac{\lambda}{\mu} \right)^k, & 0 \leq k \leq c, \\ K \frac{1}{c!} \left(\frac{\lambda}{\mu} \right)^c \rho^{k-c}, & k > c, \end{cases} \quad (2.2.9)$$

其中常数因子 K 可由正规化条件确定

$$K = \left\{ \sum_{k=0}^{c-1} \frac{1}{k!} \left(\frac{\lambda}{\mu} \right)^k + \frac{1}{c!} \left(\frac{\lambda}{\mu} \right)^c (1 - \rho)^{-1} \right\}^{-1}.$$

上述结果可总结为下列定理:

定理 2.2 经典生灭过程(2.2.3)正常返, 当且仅当方程