马氏决策规划及其在管理中的应用

沙聚桢 编著

国防工业出版社

马氏决策规划及其在管理中的应用

沙聚桢 编著

国防工业出版社

（京）新登字 106 号

图书在版编目（CIP）数据

马氏决策规划及其在管理中的应用/沙聚桢编著．一北京：国防工业出版社，1994

ISBN 7-118-01231-9

Ⅰ．马…

Ⅱ．沙…

Ⅲ．①马尔柯夫决策-应用-管理学 ②管理学-应用-马尔柯夫决策

Ⅳ．C931．1

马氏决策规划及其在管理中的应用

沙聚桢 编著

国防工业出版社出版发行

（北京市海淀区紫竹院南路 23 号）

（邮政编码 100044）

新华书店经售

北京昌平长城印刷厂印刷

*

开本 787×1092 1/32 印张 4 3/4 100 千字

1994 年 8 月第 1 版 1994 年 8 月北京第 1 次印刷 印数 1—2000 册

ISBN7－118－01231－9/F·74 **定价**：4.80 元

前　言

在现代管理中,运筹学占据着极为重要的位置。在一定的人力、物力等条件下,如何给出最优决策,以获得最大的经济效益或最佳效用值,这是运筹学应用的最终目的。在运筹学中确定性的优化方法如线性规划、整数规划、图论与网络理论、动态规划等在管理中得到广泛的应用。

最近30多年来,随着科学技术的飞速发展与管理向更深层次发展的需要,人们开始注重于随机性系统的优化研究。马尔可夫决策规划(Markovian Decision Programming)或简称为马氏决策就是解决这方面问题的一个重要分支。当所考察的系统的状态转移具有无后效性(即马氏性),或近似地看作具有马氏性时,而且通过观察与统计可以计算出状态转移概率,在给定报酬结构之后,通过马氏决策规划可以给出在某种意义下系统目标使之最优的策略。

马氏决策规划在现代管理中有广阔的应用前景。它在汽车更换、机械维修、生产与存贮等方面都得到了许多成功的应用。可以深信,在未来的宏观经济调控中,马氏决策将得到更广泛的应用。

马氏决策是确定型动态规划与马尔可夫过程相结合的产物,其基本思想已于本世纪50年代形成,60年代已形成基本概念。尽管如此,其研究方法、概念、符号目前尚不统一。本人虽然作了些工作,但肯定还有不足之处。

目前，我国尚未出版一本关于马氏决策及其在管理中应用方面的教材。如何处理马尔可夫过程的非专业人员难以掌握的深邃的内容与突出它在管理中的应用，笔者作了点滴的试尝。赵达纲、刘克等同志对本书提出了许多宝贵的修改意见，特表谢意。

由于时间仓促，缺点与不足之处一定不少，敬请读者与有关专家不吝指正，以便改进。

目　录

第一章　马尔可夫过程

在讨论马尔可夫决策过程之前，首先介绍这部分内容涉及的基础知识 —— 随机过程中的马尔可夫过程的基本概念。随机过程是研究大量随机现象的重要工具之一。在现代管理课题的深入研究中必然地要涉及诸如设备损坏的规律，顾客到达服务系统的规律等。对这些现象的研究必须具备随机过程的一些基本概念。

§1　随机过程的基本概念

随机过程是概率论重要分支之一。由于随机现象的迁移特点不同，还可以细分为若干个具体的过程，我们的目的不在于讨论这方面深奥的数学内容，而是给出有关马尔可夫过程（简称马氏过程）的基本知识。

首先通过实例加以说明。

例 1　某百货商店上午 9 点开始营业，下午 6 点闭店。考虑一个变量 $N(t)$，$t \in [9,18]$。它表示该商店在时间区间 $[9,t]$ 内到达商店购货的顾客数。当 t 固定时，$N(t)$ 不是依 t 而唯一确定的值，是一个随机遇不同而变化的量。也就是说它是一个随机变量。由于 t 在 $[9,18]$ 上连续地变化，每一个时刻都对应一个随机变量，因而该过程涉及到无穷多个不可数的随机变量问题。

例 2 $x(t)$ 表示某工厂在第 t 个月的销售水平，$t \in \{1,2,\cdots,12\}$。根据对该工厂的过去销售量的统计可知 $X(t)$ 的概率分布规律。对该问题而言，要讨论随机变量 $x(1),x(2),\cdots,x(12)$，也就是说该过程涉及到有限个随机变量问题。

从上面两例可知，描述一个随机遇而变化的过程，仅用一个随机变量进行描述是不够的。需要研究带有某个参数(一般表示时间)的一族随机变量的变化规律。对于一个经济系统或管理系统而言，其状态确实是随时间变化而变化的。为便于研究一个经济系统随时间的演变规律，可借助于随机过程这一工具，以便于对经济系统进行最优控制，达到提高经济效益的目的。

定义 设 $T \subset R_1$ 是已给的实数集，有穷或无穷、可列或不可列。设 T 中每一元 t 对应于一随机变量 $x_t(\omega)$，称随机变量族 $x_t(\omega)(t \in T)$ 为一随机过程。记此过程为

$$X(\omega) = \{x_t(\omega) \mid t \in T\}$$

$x_t(\omega)$ 可表示为 $x(t)$ 或 x_t。

T 一般表示时间参数集，x_t 的取值集合用 X 表示，称为状态空间。如按 T、X 集合的特点可把随机过程分为四类：①X、T 皆连续；②X 连续，T 离散；③X 离散，T 连续；④X 离散、T 离散。本书主要涉及 ④ 的内容。

例 1 之随机过程属于类型 ③，例 2 之随机过程属于类型 ④。

随机过程的理论在近代物理、无线电技术、自动控制和管理科学中均有重要应用。但对管理工作者而言，应用上占重要地位的是时间参数取值集合为离散的，状态集合也是离散的情形。著名数学家波尔·哈尔马斯(Paul Halmos)曾说："在不久的将来(而且就在现在)，离散数学对于我们研究工作，对于

了解世界，将成为越来越重要的工具。相反地，分析数学（即连续数学）则起到次要作用了。”这正是我们强调类型 ④ 的原因。

马尔可夫过程是一类特殊的随机过程。该过程的概念首先由俄国数学家马尔可夫（A. A. Марков）所给出。在应用上占有重要地位。

我们通过考察一个实例，给出马氏过程的概念。

在某一个养殖场中，假设有一箱兔子，已知在某一个时刻 t 有 i 只兔子，而在时刻 $\tau(\tau > t)$ 有 j 只兔子的概率为

$$P_{ij}(t,\tau) = P\{x(\tau) = j \mid x(t) = i\}$$

在该问题中，随机变量 $x(t)$ 有这样的一个特点：概率 $P_{ij}(t,\tau)$ 与在时刻 t 以前系统所处的状态无关，也就是说和系统如何演变成时刻 t 时的状态 i 的过程无关。具有该性质的过程称为该过程满足无后效性。详细地可以叙述为：

定义 如果一个随机过程 $\{x_t(\omega) \mid t \in T\}$，$(T \subset R_1)$，具有这样的性质，即若当 $\tau > t$，且对任意的 $t_{i_1}, t_{i_2}, \cdots, t_{i_k}$ 满足

$$t > t_{i_1} > t_{i_2} \cdots > t_{i_k}$$

时，如图 1-1 所示，有

$$\begin{aligned} P\{x(\tau) = j \mid x(t_{i_k}) = j_k, x(t_{i_{k-1}}) = j_{k-1}, \cdots, \\ x(t_{i_1}) = j_1, x(t) = i\} = P\{x(\tau) = j \mid x(t) = i\} \end{aligned}$$

成立，其中 $j_k, j_{k-1}, \cdots, j_1 \in X$。则称该随机过程满足无后效性或马氏性，满足无后效性的随机过程称为马尔可夫过程，称 $P_{ij}(t,\tau)$ 为马氏过程的转移概率。

图 1-1

马氏过程描述了客观世界中具有某类特征的事物的演变过程，简单地讲，该特征即事物未来的演变过程仅仅依赖于目前的状态，而与过去该事物的发展历史(指状态及演变过程)无关。换言之，事物的发展历史仅通过事物的当前状态影响事物的将来发展。我们所研究的马氏过程仅限于前面所讲的④情形。因为在现代管理应用中尤为重要的是状态集合和时间集合皆为离散的有限集。为此，时间集合可以表示为

$$T = \{0,1,2,\cdots,m\}$$

状态集合可以表示为

$$X = \{0,1,2,\cdots,n\}$$

或

$$X = \{1,2,\cdots,n\}$$

例 3 (具有吸引壁的随机徘徊问题) 考虑在一个数轴上的一列整数点 $0,1,2,\cdots,2a$(a 为某一个正整数)组成的集合 X。今有一个质点 Q 自某一状态出发自由地在数轴上随机游动，经过单位时间可以由一个整数点跳跃到另一个整数点位置。假定质点由当前的位置经过单位时间跳跃到另一点的概率仅与当前位置有关，而与该质点过去在何处无关。从这一假设可知随机徘徊问题具有马氏性。

由于由状态 i 转移到状态 j 经过单位时间达到，可以把这个概率记为 $P^{(1)}\{x(l+1)=j|x(l)=i\}$，称这个概率为状态转移概率，简称转移概率。对该概率加强假设条件为

$$P^{(1)}\{x(l+1)=i+1|x(l)=i\}=p_i(l)$$

$$P^{(1)}\{x(l+1)=i-1|x(l)=i\}=q_i(l)$$

$$p_i(l)+q_i(l)=1,\quad 0<i<2a$$

由上式可知当 $|i-j|>1$ 时，

$$P^{(1)}\{x(l+1)=j|x(l)=i\}=0$$

称 $p_i(l)$ 为前进一步概率，$q_i(l)$ 为后退一步概率。

如果再加假设条件，质点 Q 一旦处于状态 0 或 $2a$，则永远处于该状态，即有

$$P^{(1)}\{x(l+1)=0 | x(l)=0\}=1$$

$$P^{(1)}\{x(l+1)=2a | x(l)=2a\}=1$$

从直观上易观察出经过转移的步数无限增加时，质点处于 0 与 $2a$ 的概率和为 1，称 0 与 $2a$ 为吸引（收）壁。上面所讨论的问题称为具有吸引壁 0 与 $2a$ 的随机徘徊问题。

为了更详细地描述该问题的数学特征可以用矩阵工具。如果用 $p_{ij}(l)$ 表示质点 Q 在时刻 l 处于状态 i，而在时刻 $l+1$ 处于状态 j 的概率，则有

$$p_{ij}(l)\geqslant 0,\sum_{j=0}^{2a}p_{ij}(l)=1,\quad i,j\in X \tag{1-1}$$

以 $p_{ij}(l)$ 为元素的矩阵称为随机矩阵，记为 $P(l)$。应用(1-1)式的符号有

$$P(l)=\begin{bmatrix} p_{00}(l) & p_{01}(l) & \cdots & p_{0,2a}(l) \\ p_{10}(l) & p_{11}(l) & \cdots & p_{1,2a}(l) \\ p_{2a-1,0}(l) & p_{2a-1,1}(l) & \cdots & p_{2a-1,2a}(l) \\ p_{2a,0}(l) & p_{2a,1}(l) & \cdots & p_{2a,2a}(l) \end{bmatrix}$$

$$=\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ q_1(l) & 0 & p_1(l) & \cdots & 0 & 0 & 0 \\ 0 & q_2(l) & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & q_{2a-1}(l) & 0 & p_{2a-1}(l) \end{bmatrix}$$

特别重要的情况是 $P(l)$ 不依赖于 l，此时 $P(m)=P(n)$，$m,n\in T$。设某养殖场在某一时刻 t 有 i 个兔子，经过时间 τ 之后，在时刻 $t+\tau$ 有 j 个兔子的概率为 $P_{ij}(t,t+\tau)$，在营养、防疫等条件不变的情况下，兔子在时刻 $t+\tau$ 的数量仅与初始时

刻的状态和所经历时间长短有关，而与初始时刻无关。具有这样性质的事物运动过程称为关于时间是齐次的，其涵义用数学式子可表示为

$$P\{x(t+\Delta t)=j|x(t)=i\}=P\{x(\Delta t)=j|x(0)=i\}\quad i,j\in X$$

上式可以简记为 $p_{ij}(\Delta t)$。

在应用中，有一类特殊的马尔可夫过程 — 马尔可夫链占有重要的位置。在管理中经常遇到的是随机变量 x_t 的时间取值集合 T 为离散的有限集。在管理中往往把周、月、季、年作为时段的单位。考察一个企业的经济效益时，可以把整个考察期分为若干个时段进行，如年经济效益计划由 12 个月的组成。另一方面，在管理中遇到的大量状态集合 X 也为有限集。如机械维修问题，机器的性能状态假如可以分为状态良好，有小毛病，损坏严重，完全损坏，分别用数字 1、2、3、4 表示。马尔可夫链的概念是具体事例高度抽象的结果，一般定义如下：

定义 （马尔可夫链） 定义在概率空间$(\Omega,\mathscr{F},\mathscr{P})$上的取非负整数值的随机变量序列 $x_n(x_n=x_n(\omega),\omega\in\Omega),n=0,1,2,\cdots$，如果$\{x_n\}$满足等式

$$\begin{aligned}P(x_{m+k}&=i_{m+k}|x_m=i_m,x_{j_l}=i_{j_l},\cdots,x_{j_2}=i_{j_2},x_{j_1}=i_{j_1})\\&=P(x_{m+k}=i_{m+k}|x_m=i_m)\end{aligned}$$

对任意正整数 l,m,k 及任意非负整数 $j_l>\cdots>j_2>j_1(m>j_l),i_{m+k},i_m,i_{j_l},\cdots,i_{j_2},i_{j_1}$ 成立，且所有构成条件的事件概率皆大于 0，称随机变量序列 x_n 为一马尔可夫链。

由上述定义可知如果一个随机过程的时间集合与状态集合为离散的，且过程满足无后效性，这样的过程就是马尔可夫链。由前述理由我们仅讨论 T、X 是离散的有限集情形。

为了应用上的方便，把条件概率

$$P(x_{l+k}=j|x_l=i)$$

叙述的第 l 步时系统处于状态 i，经过 k 步后系统处于状态 j 的概率。此概率简记为 $_lp_{ij}^{(k)}$。由 $_lp_{ij}^{(k)}$ 所构成的矩阵

$$_lp^{(k)}=[_lp_{ij}^{(k)}]_{m\times m}$$

叫经 k 步转移的随机矩阵。显然有

$$_lp_{ij}^{(k)}\geqslant 0,\sum_{j=1}^{m}{}_lp_{ij}^{(k)}=1$$

如果条件概率 $_lp_{ij}^{(k)}$ 与 l 无关，称这样的马氏链为齐次马氏链。经过 k 步的齐次马氏链的转移概率记为 $p_{ij}^{(k)}$，其相应的随机矩阵，即经过 k 步的转移概率矩阵记为 $p^{(k)}$。在管理中主要应用齐次马氏链。以后无特别声明简称齐次马氏链为马氏链。

例 4 假设某一个人的工作调转仅有 3 种工作可供选择：① 技术工作；② 销售工作；③ 机械制造工作。工作调转过程满足无后效性，且假设转移概率与初始时刻无关。经过单位时间状态转移矩阵为

$$P^{(1)}=\begin{bmatrix}\frac{1}{3}&\frac{1}{3}&\frac{1}{3}\\ \frac{1}{2}&\frac{1}{4}&\frac{1}{4}\\ \frac{2}{5}&\frac{1}{5}&\frac{2}{5}\end{bmatrix}$$

对这个状态转移矩阵加以说明，第一行中 $p_{11}^{(1)}=\frac{1}{3}$ 说明一开始从事技术工作，经过单位时间后仍作技术工作的可能性为 $\frac{1}{3}$；$p_{12}^{(1)}=\frac{1}{3}$ 说明转到销售工作的可能性也为 $\frac{1}{3}$；对于 $p_{13}^{(1)}=\frac{1}{3}$ 也可以作类似的解释。状态转移矩阵可以形象地用状态转

移图表示出来，如图 1-2 所示。顶点表示状态，每一条弧的权表示转移概率。始端与终端重合的弧（即环）表示由起始状态又转到本状态。图 1-2 更一般地表示为图 1-3。

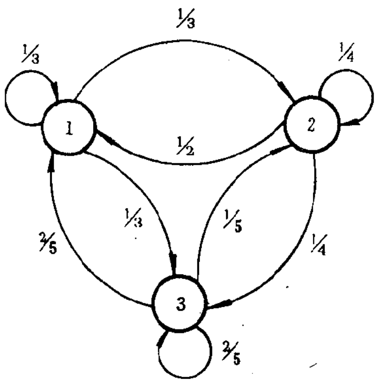

图 1-2

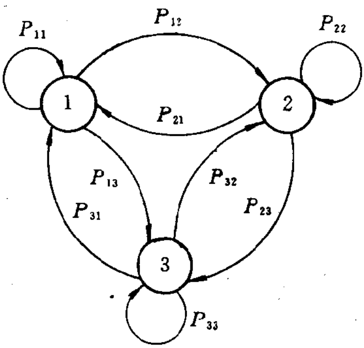

图 1-3

§2 状态转移概率的计算

对于给定的一个马尔可夫链而言，已知其一步状态转移矩阵，如何计算经过 $n(\geqslant 2)$ 步转移后的状态转移矩阵，这是本节所要讨论的问题。首先通过实例给出具体算法。

例 5 （人才预测问题） 某研究院 1985 年初有科技人才（中专文化程度以上者）1200 人，其中技术人员为 500 人，助理工程师为 400 人，工程师为 200 人，高级工程师为 100 人。该研究院各种科技人才的构成层次可以用状态变量 $x_i(t)(i=1,2,3,4)$ 分别表示在时刻 t 技术员、助理工程师、工程师、高级工程师的人数。1985 年初该研究院的人才构成可以表示为

$$x_1(1985)=500$$
$$x_2(1985)=400$$
$$x_3(1985)=200$$
$$x_4(1985)=100$$

根据过去历史统计资料可以给出各级职称人员经过一年晋升的概率，以及调离、退休、死亡等原因造成的脱离事件的概率，用第五状态表示脱离研究院这一状态。综上所述可以得到状态转移矩阵

$$P^{(1)}=\begin{bmatrix}0.70 & 0.20 & 0 & 0 & 0.10\\ 0 & 0.70 & 0.25 & 0 & 0.05\\ 0 & 0 & 0.80 & 0.15 & 0.05\\ 0 & 0 & 0 & 0.90 & 0.10\\ 0 & 0 & 0 & 0 & 1\end{bmatrix}$$

如果人才不补充，试求 1988 年初该研究院各层次科技人才的数量。

例 6 某车间有同一型号的车床 4 台。今欲考察该车间 3 年来的机床运行状态。机床的可能运行状态为 x_1— 所有的机床完好；x_2— 仅有一台机床损坏；x_3— 仅有两台机床损坏；x_4— 仅有三台机床损坏；x_5— 全部损坏。如果根据过去统计资料得到的状态转移图如图 1-4。求第三年末该车间机床处于状态 x_i 的期望值为多少，$i = 1,2,3,4,5$。

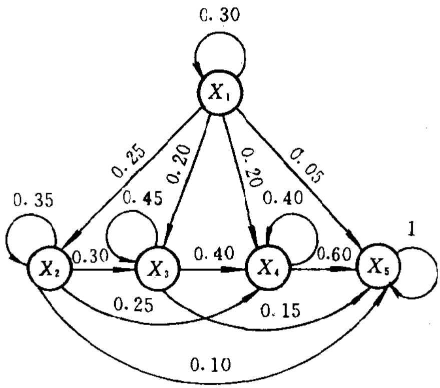

图 1-4

例 5 与例 6 提出了管理工作者所关心的一个问题：一个系统经过一定的时段之后，在某一个时段 t 系统处于各个状态如何。也就是说

$$X(t) = [x_1(t), x_2(t), \cdots, x_m(t)]$$

等于什么。这个状态向量称为系统在时段 t 的状态。例 5 要求求出状态向量

$$X(1988) = [x_1(1988), x_2(1988), x_3(1988), x_4(1988), x_5(1988)]$$

例 6 要求计算状态向量

$$X(3)=[x_1(3),x_2(3),x_3(3),x_4(3),x_5(3)]$$

我们可以首先计算

$$X(1986)=[x_1(1986),x_2(1986),x_3(1986),x_4(1986),x_5(1986)]$$

$$X_1(1986)=[500,400,200,100,0]\cdot[0.70,0,0,0,0]^T=350$$

$$X_2(1986)=[500,400,200,100,0]\cdot[0.20,0.70,0,0,0]^T=380$$

类似地可以算得

$$x_3(1986)=260,x_4(1986)=120,x_5(1986)=90$$

如果用 $P_j^{(1)}$ 表示状态转移矩阵的第 j 列向量，则有一般公式

$$X_j(1986)=X(1985)\cdot P_j \qquad (1\leqslant j\leqslant 5)$$

再加以推广为一般情形为

$$X_j(1)=X(0)P_j^{(1)} \quad (1\leqslant j\leqslant m) \tag{1-2}$$

用矩阵运算表示 $X(1986)$ 有

$$X(1986)=X(1985)\cdot P^{(1)}=[500,400,200,100,0]\times\begin{bmatrix}0.70 & 0.20 & 0 & 0 & 0.1\\ 0 & 0.70 & 0.25 & 0 & 0.05\\ 0 & 0 & 0.80 & 0.15 & 0.05\\ 0 & 0 & 0 & 0.90 & 0.10\\ 0 & 0 & 0 & 0 & 1\end{bmatrix}$$

该问题的一般提法为：假设所给的随机过程为一马尔可夫链，状态向量 $X(0)$ 及一步状态转移矩阵为已知，如何求出经过 k 步状态转移后的状态向量 $X(k)$。值得指出的是这里所指状态向量为多维动态规划中的状态向量。

以例 5 加以说明，1985 年初记为时刻 0，研究院科技人才