783764

35
600

数学地质丛书

马尔柯夫过程在地质学中的应用

景毅　王世称　苑清扬　著

地质出版社

783764

数 学 地 质 丛 书

马尔柯夫过程在地质学中的应用

景 毅 王世称 苑清扬 著

地 质 出 版 社

Series on Mathematical Geology

THE APPLICATIONS OF MARKOVPROCESSES IN GEOLOGY

Jing Yi Wang Shichen

Yuan Qingyang

GEOLOGICAL PUBLISHING HOUSE

BEIJING 1985

1Ap8 /07

数学地质丛书

马尔柯夫过程在地质学中的应用

景 毅 王世称 苑清扬 著

*

责任编辑：高书平、刘承祚

地质出版社出版

（北京西四）

地质出版社印刷厂印刷

（北京海淀区学院路29号）

新华书店北京发行所发行·各地新华书店经售

*

开本：850×1168 $^{1}/_{32}$ 印张：8 $^{1}/_{2}$ 插页：3个 字数：222,000

1986年2月北京第一版·1986年2月北京第一次印刷

印数：1—2,090册 定价：2.55元

统一书号：13038·新158

前　　言

本书为数学地质丛书的一个组成部分，编写本书的目的主要在于介绍马尔柯夫过程的基本原理及其在地质学中的应用。尽管马尔柯夫过程在很多学科中已经得到广泛地应用，但是把它应用于地质学中还是一个新的探索。

预计到本世纪末，“纯数学”和“应用数学”将会有一个很大的发展。数学正以坚定的步伐渗透到生物学、化学、经济学、语言学、地质学、社会学等当中去。例如研究较为通用的拓扑空间的马尔柯夫过程及其有关的从一般连续函数的巴拿赫空间变换成正半群的全部算符。这一课题对于扩散理论、热传导理论和岩浆热力学等问题的研究都具有很大意义。马尔柯夫链为近代概率论的一个新领域，即马尔柯夫随机过程理论。这一领域从理论研究到实际应用方面都在飞速发展。

长春地质学院数学地质研究室，在景毅教授和王世称教授的指导下，几年来围绕这方面的课题，结合师资进修和培养研究生开展了此项研究工作。我们应用马尔柯夫过程分析的原理和方法研究了变质铁矿和沉积磷矿的层位对比、岩相分析、层状矿体的资源统计预测和航空照片的定量解释等。我们将这初步研究成果介绍给广大地质工作者，希望起到抛砖引玉的作用。

本书共分五部分：

一、马尔柯夫过程概论；

二、马尔柯夫过程在地层分析上的应用；

三、半马尔柯夫链在地层分析上的应用；

四、时间连续状态离散的马尔柯夫过程方法的应用；

五、国内外的一些研究实例。

所选的例子都是有关的理论或重要方法的应用。关于时间连

此为试读，需要完整PDF请访问：www.ertongbook.com

续状态也连续的沉积扩散问题，我们尚在研究中。

本书前四章由景毅执笔，有关地质部分由王世称和苑清扬执笔。夏立显对本书原稿提出了一些有益的意见。参加本项研究工作的还有刘安洲、范继章、曹绪言及研究生张光前。

本书初稿承蒙张文佑教授、赵鹏大教授、裴荣富所长和刘承祚副研究员审阅并提出宝贵意见，在此表示感谢。

由于我们水平有限，书中难免有错误及不足之处，希望读者提出批评指正。

编　者

一九八四年九月

目　　录

第一章　马尔柯夫过程概论

§1. 引　言

科学体系的基础是在整个宇宙中存在着的某种秩序或格局。科学研究的目的就在于揭示这种秩序或格局的实质，并作有条件的利用。地质科学的研究也是如此。预计在1980—2000年期间，应用数学将有一个很大的发展。由于电子计算机技术的发展，数学正在被应用到许多学科中去。许多新的设想和许多崭新的技术如同雨后春笋一般涌现出来，一幅光辉的远景已展现在我们面前。但是，地质学和其他学科相比，在采用电子计算机技术方面显得不那么快。这是因为地质学研究过去主要限于定性的和描述性的。然而，在今天，地质学的领域中，应用各种数学方法研究和模拟问题日益增多，这是一个良好的新局面。应该理解到，并不是为了改进技术而是为了从根本上过渡到一个新的发展阶段，将数学的原理和方法引用到地质学中来。地质学需要其自己的数学科学，而不是仅仅从物理、化学或力学中借用方法。尽管这种借用也很有益处，但对地质学本身的发展不能起到带有根本性质的作用。

地质过程很多属于随机过程。而马尔柯夫过程属于近代概率论的一个重要的新领域—随机过程理论，在今天已得到迅速发展。马尔柯夫过程分析方法逐渐地、更广泛地被应用到地质领域中来。

本章将介绍马尔柯夫过程的概念，主要理论和基本方法。

此为试读，需要完整PDF请访问：www.ertongbook.com

§2. 随机过程概念

在自然界和人类社会中存在两种不同的现象。第一种现象如在没有外力作用的条件下，物体永远保持匀速直线运动状态，可表示为$S=vt$，其中S表距离，v表速度，t表时间。S与v，t之间存在着确定的函数关系。我们可以根据自变量的数值准确地计算因变量的值。我们称这类数字模型为确定性模型，而与其相应的过程称为“因果过程”。第二种现象与第一种不同，在因变量和自变量之间不存在确定性关系。但在它们之间存在着某种相关关系，这类过程的进行受在体系内部所存在的特定概率支配。这种过程称为“随机过程”，这里先举几个例子。

例1. 在［0，1］这段时间内来观测液面上作布朗运动的微粒。若$X(t)$，$Y(t)$分别表示质点在时刻t的X坐标和Y坐标，则对一固定的t，$X(t)$，$Y(t)$显然是个二元的随机变量。若t变动时，就得到二维空间内一族无限多个随机变量。

显然，在空间内的微粒运动，是由$X(t)$，$Y(t)$，$Z(t)$三个坐标组成的随机变量。若t变动时，就得到三维空间内一族无限多个随机变量。

例2. 在波尔（Bohr）氢原子模型中，电子可以在允许的轨道之一上面运动，我们以$X(t)=i$表示“电子在时刻t，在第i条轨道上运动”。假设电子轨道的变化只在时刻t_1，t_2，t_3，…发生，显然，当t分别取t_1，t_2，t_3，……这些值时，就得到一维空间内一族无限多个随机变量。

例3. 在地层分析问题中，若对某一地层剖面的n个层面，从下而上在每层面上都取一点P_k，并在每个点上测定m种元素的含量

$P_1\ (x_{11},\ x_{12},\ \cdots,\ x_{1m})$，

$P_2\ (x_{21},\ x_{22},\ \cdots,\ x_{2m})$，

……………………

$$P_n\ (x_{n1},\ x_{n2},\ \cdots,\ x_{nm})$$

组成一族n个m维随机变量。

例4. 一个水系由不同方向不同级别的河道组成。可以将沿水系由上游向下游不同方向河道的变化，看成为一种状态向另一种状态的随机变化过程，它们的变化受一定概率支配。

由此可见，从实际需要来看，不仅要求我们考虑一个随机变量或同时考虑有限个随机变量，而且还要求我们考虑一族无限多个随机变量。今后我们将用"随机变量"一词来表示依赖于一个变动参数的一族随机变量。上述四个例子在概率论上都是随机过程，它们都是我们即将讨论的问题和研究的对象。

定义 依赖于参数t的随机变量集合$\{\xi(t)\}$为随机过程，其中t属于一个实数集T。我们将它记为$\{\xi(t),\ t\in T\}$，称参数t为时间。一个随机变量序列$\{\xi_k\}$ $(k=1,\ 2,\ 3,\ \cdots\cdots)$也是一个随机过程。这里$T$只是所有正整数组成的集。这种过程称为离散时刻的过程；若T是不可数集，而仅限T为有限或无穷区间情形，我们称它为具有连续时刻的过程。

在随机过程的许多问题中，参数t的确代表实际观测的时刻，但也可能表示其它的物理量。例如过程所处的空间—地层堆积的

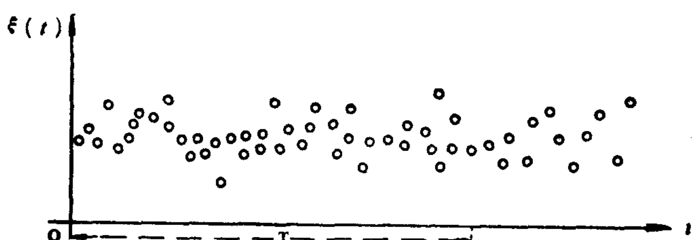

图 1　随机过程变化图

t—表示时间；$\xi(t)$ —随机过程；
T—规定时间；0—点的位置

过程。按照概率论理论，多元随机变量不同于几个随机变量的简单的组合。在多元随机变量问题中，除了要研究其中每个随机变量的性质外，还必须研究这些随机变量之间的关系。同样，在随

机过程的研究中，也是如此。由于对随机过程讲，随机变量的数目变为无限，情况就要复杂得多。虽然如此，我们仍可借助于以分布函数完整地描述随机变量的统计规律的方法来完整地描述随机过程的统计规律性。

令 $\{\xi(t), t\in T\}$ 是依赖于参数 t 的一族随机变量，其中 T 表示随机过程 $\xi(t)$ 的参数 t 的变化范围。如何选定 T，这将随所研究的具体对象而定。上图1则表示在范围 $[0, T]$ 内一族44个随机变量的变化过程。例如，在例1中 T 表示有限区间 $[0, 1]$，在例2中的 T 为可数集合。当 T 中元素为可数时，我们也将随机过程 $\{\xi(t), t\in T\}$ 称为随机序列。按前边所述对于每一个固定的 t $(\in T)$，$\xi(t)$ 是一个随机变量，因而对任何一固定的 t，随机变量 $\xi(t)$ 的分布函数

$$F_t(x) = P\{\xi(t) < x\} \tag{1.1}$$

应该是确定的。此外，为了要描述 $\xi(t)$，$t\in T$ 相互之间的关系，自然要求对于任意指定的 n 及任意的 n 个 T 中的元素 $t_1, t_2\cdots$, t_n 对应于 $t_1, t_2, \cdots, t_n$ 的随机变量 $\xi(t_1), \xi(t_2), \cdots, \xi(t_n)$，它们的联合分布函数

$$F_{t_1, t_2, \cdots t_n}(x_1, x_2, \cdots, x_n) =$$

$$P\{\xi(t_1) < x_1, \xi(t_2) < x_2, \cdots, \xi(t_n) < x_n\} \tag{1.2}$$

应该是确定的。由于上式中的 n 及 $t_1, t_2, \cdots, t_n$ 都是任意的，因此，等式（1.2）给出一族无限多分布函数记为 $\{F_{t_1, t_2, \cdots, t_n}(x_1, x_2, \cdots, x_n)\}$，称为随机过程 $\xi(t)$ 的有穷维分布函数族。可以看出，这一族分布函数不仅刻划出了对应于每一个 t 的随机变量 $\xi(t)$ 的统计规律性，而且也刻划出各个 $\xi(t)$ 之间的关系。因此，随机过程的统计规律性可由它的有穷维分布函数族完整地描述出来。

不难看出，这种由有穷维分布函数族来刻划出随机过程的统计性的方法，也适用于当 T 是一个多维欧几里德空间中的点集的情形，如例3。此时只需将（1.2）中的 $t_1, t_2, \cdots, t_n$ 了解为多维欧几里德空间中的点即可。

对于二维情形，如例1中所考虑的在液面上作布朗运动的微粒的坐标（$x(t)$，$y(t)$），则可将它的有穷维分布函数族定义为：

$$
\begin{aligned}
&F_{t_1, t_2, \cdots, t_n}(x_1, y_1; x_2, y_2; \cdots, x_n, y_n) \\
&\quad = P\{x(t_1) < x_1, y(t_1) < y_1; x(t_2) < x_2, y(t_2) \\
&\qquad < y_2, \cdots, x(t_n) < x_n, y(t_n) < y_n\} \qquad (1.3)
\end{aligned}
$$

不难看出，二维过程（$x(t)$，$y(t)$）的统计特性可由分布函数族$\{F_{t_1, t_2, \cdots t_n}(x_1, y_1; x_2, y_2; \cdots, x_n, y_n\}$来表示。

对于k维随机过程，可相仿地建立它的k维分布函数族为

$$
\begin{aligned}
&F_{t_1, t_2, \cdots, t_n}(x_{11}, x_{12}, \cdots, x_{1k}; x_{21}, x_{22}, \cdots, \\
&\quad x_{2k}; \cdots, x_{n1}, x_{n2}, \cdots, x_{nk}) \\
&\quad = P\{x_{11}(t_1) < x_{11}, x_{12}(t) < x_{12}, \cdots, x_{1n}(t) \\
&\qquad < x_{1n}; \cdots\cdots x_{n1}(t) < x_{n1}, x_{n2}(t) < x_{n2}, \cdots, x_{nk} \\
&\qquad (t) < x_{nk}\}
\end{aligned}
$$

随机过程的理论在近代物理、地质科学、经济学、公用事业等方面均有广泛的应用。本书的重点是介绍应用最广而理论亦比较完善的马尔柯夫过程及其在地质科学上的应用。

地球上所发生的种种地质现象都曾经历过一个历史的演变过程，并且还在不断向前发展着。任何一种地质过程的演变和发展既遵循着某种确定的物理学的和化学的规律，又受种种随机因素的支配和控制。由于地质条件异常复杂，随机因素往往起着重要作用，因而多数地质过程都属于随机过程。

综上所述，我们可以将随机过程简单地概括为：任意一个受概率定律支配的真实过程就是一个物理的随机过程。数学上的随机过程就是物理随机过程的数学模型。

§3. 马尔柯夫过程的定义与类型

马尔柯夫过程是下述的这样一种过程，在已知时刻t_0系统所处状态的条件下，在时刻t_0以后系统将到达的情况与时刻t_0以前

系统所处的状态无关。这个特性称为无后效性；就是说，过程的将来发展完全取决于其现在状态，而不依赖于现在状态如何由过去发展而来，即与现在以前的状态无关，这种性质也称为“马尔柯夫性质”。

在地质科学中，山势倾斜，水系游动，地层堆积，沉积扩散或岩浆活动等过程都是与马尔柯夫性质有关的地质过程。

为了进一步讨论具有这种特性的随机过程，我们给出一个严格的数学定义

定义3.1 在马尔柯夫过程条件下，用$P_{ij}(t,\ \tau)$表示“已知在时刻t，系统处于状态E_i的条件下，在时刻τ（$\tau>t$）系统处于状态Ej”的概率，称为转移概率。

如果我们已经知道马尔柯夫过程$x(t)$的转移概率$P_{ij}(t,\ \tau)$，则它的❶有穷维分布

$$F\{x(t_1)=r_{i1},\ x(t_2)=r_{i2},\ \cdots,\ x(t_n)=r_{in}\}$$
$$(0\leqslant t_1<t_2<\cdots<t_n)$$

也就知道了。事实上，如果假定在初始时刻系统处于某状态x_{i0}：$x(0)=r_{i0}$，则因为马尔柯夫过程的无后效性，有：

$$\begin{aligned}P\{x(t_1)=r_{i1},\ x(t_2)=r_{i2},\ \cdots,\ x(t_n)=r_{in}\}\\ =P_{i0}i_1(0,\ t_1)\,p_{i1}i_2(t_1,\ t_2)\,p_{i2}i_3(t_2,\ t_3)\cdots\\ \cdots P_{in-1}i_n(t_{n-1},\ t_n)\end{aligned}\tag{1.5}$$

因此，对马尔柯夫过程讲，转移概率就完全描述了它的统计性质。

如何决定转移概率的问题，正是马尔柯夫过程理论中研究的重要问题之一。研究的问题不同，确定问题转移概率的方法也有差异。

特别，当上述过程的转移概率只与$i,j,\tau-t$有关时，就称这个过程是时齐的马尔柯夫过程。这时可以用$P_{ij}(t)$来表示“已

---

❶显然，若随机过程只取可列个值，则有穷维分布函数即可用此处的有穷维分布来代替。

知在时刻τ系统处于状态Ei的条件下，经过一段时间t后系统处于状态E_j''的概率，用原来的记号，它可表成：$P_{ij}(t)=P_{ij}(\tau, t+\tau)$成立。

有时为了简化，往往只考虑在时刻t_1，t_2，t_3，……才能发生转移，那么时间也就离散化了。

定义3.2 状态的转移只能在$t=t_n$（$n=1, 2, 3, \cdots$）发生的马尔柯夫过程称为马尔柯夫链。即马尔柯夫链是时间与状态都离散的马尔柯夫过程。

定义3.3 设有随机过程$X(t)$，如果已知在时刻t系统处于状态X的条件下，在时刻τ（$\tau>t$）系统所处状态和时刻t以前系统所处的状态无关，则称这个随机过程为马尔柯夫过程。

定义3.4 一个只能取可列个值r_1，r_2，r_3，………的随机过程$X(t)$，我们把$X(t)=r_n$称为在时刻t系统处于状态E_n（$n=1, 2, 3, \cdots\cdots$），若在已知时刻t系统处于状态Ei的条件下，在时刻τ（$\tau>t$）系统所处的状态与时刻t以前系统所处的状态无关，那么称这个随机过程为时间连续、状态离散的马尔柯夫过程。

一般马尔科夫过程的转移概率随时间而变，可称为非时齐的或非平稳马尔科夫过程；转移概率不随时间而变的马尔科夫过程称为平稳马尔科夫过程，它们是马尔科夫过程的一种特殊类型。

马尔柯夫过程分类：按照时间和状态的离散、连续情况分为三类

（1）时间与状态（空间）都离散的过程；

（2）时间连续与状态（空间）离散的过程；

（3）时间与状态（空间）都连续的过程。

在地质学的领域内，目前主要侧重于前两类的应用，而我们对第三类问题还是很感兴趣的，目前还在继续研究。

此为试读，需要完整PDF请访问：www.ertongbook.com