隐藏马尔科夫模型及其在金融中的应用

胡淑兰 ◎著

華中師範大学出版社

隐藏马尔科夫模型
及其在金融中的应用

胡淑兰　著

华中师范大学出版社

新出图证(鄂)字 10 号

图书在版编目(CIP)数据

隐藏马尔科夫模型及其在金融中的应用/胡淑兰著. —武汉：华中师范大学出版社，2014.8

ISBN 978-7-5622-6735-5

Ⅰ. ①隐… Ⅱ. ①胡… Ⅲ. ①统计模型—应用—金融学—研究 Ⅳ. ①F830

中国版本图书馆 CIP 数据核字(2014)第 175985 号

隐藏马尔科夫模型及其在金融中的应用

©胡淑兰 著

责任编辑：	李晓璐 袁正科	电 话：	027—67867362
编 辑 室：	第二编辑室	责任校对：	刘 峥 封面设计：张 蕾
出版发行：	华中师范大学出版社		
社 址：	湖北省武汉市珞喻路 152 号	邮 编：	430079
销售电话：	027—67863426/67863280		
邮购电话：	027—67861321	传 真：	027—67863291
网 址：	http://www.ccnupress.com	电子信箱：	hscbs@public.wh.hb.cn
印 刷：	虎彩印艺股份有限公司	督 印：	章光琼
开 本：	710mm×1000mm 1/16	印 张：	13 字 数：220 千字
版 次：	2014 年 8 月第 1 版	印 次：	2014 年 8 月第 1 次印刷
定 价：	32.00 元		

前　言

隐藏马尔科夫模型最初由 Baum 和 Petrie(1966)作为马尔科夫模型的某种概率函数形式提出之后，得到了广泛的应用。自此，专家学者们从各个角度丰富了隐藏马尔科夫模型的知识体系。近年来，隐藏马尔科夫模型在语音识别、经济学、生物学、医学、计量经济学、工程等领域有着广泛的应用。随着大数据时代的到来，互联网与金融相互融合，隐藏马尔科夫模型在海量的金融数据中的应用也受到了人们的关注。

目前国内介绍隐藏马尔科夫模型的书籍较少，部分随机过程的书籍中虽有提及隐藏马尔科夫模型，但所占篇幅都不多。在这样的背景下，我们特别希望有一本关于隐藏马尔科夫模型基本理论以及给出了模型在金融领域中应用案例的书籍。相信随着研究人员的逐步探索和科技的进步，隐藏马尔科夫模型在金融、经济计量等领域中会有更多的应用研究。

本书主要介绍隐藏马尔科夫模型基础理论以及在金融、经济生活中的应用，书中提供了大量的案例分析。为了便于读者理解，其中的定理和推论，并没有全部给出严格的证明，阅读本书并不需要有较多的数学背景。

本书适合作为统计学、金融学、经济计量学等相关领域研究人员和工作者的参考书，同时也可以作为研究生学习隐藏马尔科夫模型理论的入门教材。本书主要分为六章：第 1 章介绍了隐藏马尔科夫模型的研究意义、背景以及发展历史；第 2 章介绍了马尔科夫过程的定义、性质和理论，是学习隐藏马尔科夫模型的基础；第 3 章介绍了隐藏马尔科夫链的基础知识、性质以及其广泛的应用领域，并且结合 Matlab 软件介绍了隐藏马尔科夫模型中的三个基本问题以及解决这三种问题的最常用的算法；第 4 章介绍了隐藏马尔科夫模型的极限理论，包括大偏差原理、传输信息不等式等；第 5 章主要结合 EViews 软件和 Matlab 软件，对比介绍了传统条件异方差模型与马尔科夫转移模型；第 6 章基于模型，对世界原油期货价格的波动性进行了实证分析。

由于隐藏马尔科夫过程的研究领域和应用范围很广，本书仅对隐藏马尔科夫过程作了基本的介绍，许多内容都未涉及，还有许多未给出详细证明的定理和性质等需要读者自己查阅更多的文献。书中难免有不尽完美之处，我们欢迎各界朋友能对本书提出宝贵的意见，希望能在大家的帮助下，不断完善本书的内容。

本书受国家自然科学基金(No. 32212112002)、中央高校基本科研业务费专项资金（No. 31541311210）及中南财经政法大学统计与数学学院学科建设资金资助，同时得到了许多同仁朋友的鼓励和支持，谨此致谢！特别对研究生张璇同学、赵预立同学、魏伊同学的参与编写排版，表示感谢！

胡淑兰

2014 年 6 月 20 日

目　　录

第1章 概　　论

1.1 隐藏马尔科夫模型的研究背景

Baum 和 Petrie (1966)① 笼统地介绍了隐藏马尔科夫模型(Hidden Markov Models,简称 HMMs),在文中,它们被称为马尔科夫链的概率函数。实际上是指,观测值的状态序列的概率分布具有马尔科夫性,由于状态序列一般为有限个,故而是马尔可夫链。Baum和Petrie(1966—1969)进一步研究了平稳遍历的有限状态下的隐藏马尔科夫模型的统计性质。他们证明了遍历性,给出了一个隐藏马尔科夫模型关于另一个隐藏马尔科夫模型的相对熵密度的几乎处处收敛性下的遍历定理;同时还证明了极大似然估计的相合性和渐近正态性。Petrie (1969)② 提出了隐藏马尔科夫模型可识别性的一个充分条件,同时弱化了 1966 年的著作中的一些假定。对于给定的隐藏马尔科夫模型中的观测序列,Baum, Petrie, Soules 和 Weiss (1970) 为了估计观测值所在状态的概率,建立了向前- 向后迭代算法。并且在建立更一般的隐藏马尔科夫模型参数的极大似然估计方法时,他们也利用向前- 向后迭代算法,建立了一个有效的数值迭代方法:著名的最大期望算法(Expectation-Maximum,简称 EM 算法). Dempster, Laird 和 Rubin (1977)③ 将这个方法应用于隐藏马尔科夫模型。Baum, Petrie, Soules 和 Weiss (1970) 又在此

---

① L. E. Baum and T. Petrie. Statistical inference for probabilistic functions of finite state Markov chains, *Annals of Mathematical Statistics*, 1966(37), pp. 1554-1563.

② T. Petrie. Probabilistic functions of finite state Markov chains, *Annals of Mathematical Statistics*, 1969(40), pp. 97-115.

③ A. P. Dempster et al. Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion), *Journal of the Royal Statistical Society Series*, 1977, pp. 1-38.

基础上,给出了算法的局部收敛性质,故此算法通常名为 Baum 算法,或者 Baum-Petrie 算法。而为了纪念 Lloyd Welch,也称这个算法为 Baum-Welch 算法。类似地,向前-向后迭代算法在更早的时候亦由 Chang 和 Hancock (1966)[①] 在处理码间干扰信道的最优解码问题中提出。

在引入马尔科夫链的概率函数之前,马尔科夫链的确定性函数得到了广泛研究。在相关统计文献中,它们通常被称为聚合马尔科夫过程。因为函数可能使马尔科夫链的几个状态转移压缩到一个状态上。而隐藏马尔科夫模型的状态空间有限,那么具有马尔科夫性的有限状态空间的确定性函数和概率函数是相关的。所以,马尔科夫链的任何确定性函数都可以被描述为一个有限简单状态的隐藏马尔科夫模型,而任何有限状态的隐藏马尔科夫模型都可以被描述为一个带有扩张的状态空间的马尔科夫链的确定性函数。

Shannon (1948)[②] 使用马尔科夫链的确定性函数作为信号源的模型。Ash (1965)[③] 把这类函数称为马氏源,但是这个术语常常与 Gallagher (1968)[④] 引入的单线源关联到一起。Shannon, McMillan 和 Breiman 三位学者将隐藏马尔科夫模型分别扩展在 L^1 情形、几乎处处收敛意义下的基本遍历有限状态过程上,被称为 Shannon-McMillan-Breiman 定理,或者渐近均分性质[⑤]。这个定理适用于任何平稳遍历并具有有限状态的隐藏马尔科夫模型。马尔科夫链的确定性函数在 Blackwell (1957), Blackwell 和 Koopmans (1957), Burke 和 Rosenblatt (1958), Gilbert (1959), Fox (1962), Dharmadhikari (1963—1965), Heller (1965) 和 Carlyle (1967) 等学者的研究中得到了更深入地探讨。他们研究了马尔科夫链的确定性函数成为马尔科夫链的可识别性条件。

隐藏马尔科夫模型构成了一个参数随机过程的庞大家族,也常常与工程学、统

① R. Chang and J. C. Hancock. On receiver structures for channels having memory, *IEEE Transactions on Information Theory*, 1966, 12(4), pp. 463-468.

② C. E. Shannon. A mathematical theory of communication, *The Bell System Technical Journal*, 1948, 27, pp. 379-423, 623-656.

③ R. B. Ash. *Information Theory* (New York: Dover, 1965), p. 185.

④ R. G. Gallager. *Information Theory and Reliable Communication* (New York: Wiley, 1968).

⑤ R. M. Gray. *Entropy and Information Theory* (New York: Springer-Verlag, 1990).

计学、计量经济学的许多随机过程产生关联。隐藏马尔科夫模型与混合过程有着显著关联,它的每个观察值有一个混合分布。但与混合过程不同的是,隐藏马尔科夫模型的观察值不需要满足独立性条件。另外,马尔科夫机制转移自回归过程是隐藏马尔科夫模型在时间序列中的应用。马尔科夫机制转移自回归过程的每一时刻的观测值都依赖于那一时刻马尔科夫链的状态,当自回归阶数为零时,转移自回归过程即退化为一个隐藏马尔科夫模型。从控制论的角度来看,隐藏马尔科夫模型就是动力系统。当状态空间是有限的或者是可数无限的,每个状态可以用欧几里得空间的一个单位矢量来表示。有关研究有马尔科夫-泊松过程,是指泊松过程的抵达率受不可观测的连续时间情形下马尔科夫链的控制。马尔科夫-泊松过程可以被视为一个马尔科夫更新过程或一个隐藏马尔科夫模型。在这两种情况中,离散时间下的马尔科夫链由在泊松事件原点的连续时间链抽样决定,而观察序列由间隔时长决定。

隐藏马尔科夫模型最早的一个主要应用是文字语言识别。Raviv (1967)① 在 IBM T. J. 沃森研究中心研究了相关问题。语言的字符是由马尔科夫链的状态表示,测量值由观察过程表示。Raviv 检测了识别过程中的最少字符错误率。为了检测,Raviv 对给定观察值下对应状态的条件概率建立了新的算法。

20 世纪 70 年代中期,IBM T. J. 沃森研究中心研究了隐藏马尔科夫模型的另一个主要应用。此应用由 Jelinek (1976), Baker (1975), Jelinek, Bahl 和 Mercer (1975), Bahl 和 Jelinek (1975) 以及他们的合作者通过隐藏马尔科夫模型模拟语音信号建立了一个语音识别系统。其中模型词汇中的每个单词由个人语音模型组成,而个人语音模型采用了 Baum 算法。另外, Viterbi 算法或 Jematic 的堆栈图搜索算法能够解决语音解码问题。还有,美国国防分析研究所的 Ferguson 小组,美国电话电报公司(AT&T) 贝尔实验室的 Rabiner 小组都对语音识别进行了更深入一步的研究。这些研究均使用了隐藏马尔科夫模型弗格森理论,马尔科夫链的概率函数可能就是在此时第一次被称为隐藏马尔科夫模型。

隐藏马尔科夫模型以良好的数据建模作为支撑,在语音识别等自然语言处理领域得到了很好的应用,被公认为语音识别领域中最成功的统计模型之一。在其他领域中隐藏马尔科夫模型同样有着普遍的适用性。通过假设一个"隐藏" 结构的存在,隐藏马尔科夫模型能够把简单的词序信息和更高层的语言信息联系起

① J. Raviv. Decision making in Markov chains applied to the problem of pattern recognition, *IEEE Transactions on Informaction Theory*, 1967, IT-3, pp. 536-551.

来,以完成诸如词性标注这样的任务。

近年来,隐藏马尔科夫模型已经引起统计学家和信息理论学家的广泛关注。Baum 和 Petrie 完成了大量的隐藏马尔科夫模型的基础理论研究工作,包括有限状态的隐藏马尔科夫模型,扩张到连续状态空间的隐藏马尔科夫模型,进而到一般有限状态空间的隐藏马尔科夫模型的情形等。特别地,Leroux (1992); Finesso(1990);Le Gland,Mevel (2000) 以及 Douc,Matias (2001) 等学者建立了隐藏马尔科夫模型相对熵密度的新遍历定理。Leroux (1992);Bickel, Ritov, Ryden (1998);Le Gland,Mevel (2000), Jensen,Petersen (1999) 以及 Douc, Matias (2001) 等学者证明了隐藏马尔科夫模型极大似然参数估计的相合性和渐近正态性。Francq,Roussignol (1998);Krishnamurthy,Ryden (1998) 以及 Douc, Moulines,Ryden (2004) 等学者也证明了马尔科夫机制转移自回归过程中极大似然参数估计的遍历定理与渐近最优性。Ryden (1994—1995) 在马尔科夫-泊松过程中也发现了类似结果。Le Gland,Mevel (2000) 以及 Douc,Matias (2001) 等学者研究了隐藏马尔科夫模型的指数遗忘性以及几何遍历性。作者在 Hu (2011)① 中,证明了隐藏马尔科夫模型的信息传输费用不等式,同时推广了 Kontorovich (2006) 的结果。同时也在 Hu,Wu (2011)②中,证明了动力系统的大偏差原理,并进一步将此原理应用到隐藏马尔科夫模型中,给出了大偏差原理速率函数的显示性结果。Ito, Amari 和 Kobayashi (1992) 给出了非平稳马尔科夫链的确定性函数识别的完全解法。Ryden (1996) 给出了马尔科夫调制泊松过程的识别条件。Lindgren (1978); Askar,Derin (1981) 建立了新的稳定的递归方法,这个递归方法为马尔科夫链提供了条件均值滤波器以及平滑器,且这个通过信道观测的马尔科夫链并不要求服从高斯分布。

除了 Baum 和 Petrie 的一系列相关研究,近些年还有许多关于隐藏马尔科夫模型的研究成果。如 Elliot, Aggoun 和 Moore (1994)③ 对一般的隐藏马尔科夫模

① Shulan Hu. Transportation inequalities for hidden Markov chains and applications, *Science China Mathematics*, 2011,54(5), pp. 1027-1042.

② Shulan Hu and Liming Wu. Large deviations for random dynamical systems and applications to hidden Markov models, *Stochastic Processes and Their Applications*, 2011,121(1), pp. 61-90.

③ R. J. Elliott et al. *Hidden Markov Models: Estimation and Control* (New York: Springer-Verlag, 1994).

型建立了全面的动态系统方法。特别是他们对于一般隐藏马尔科夫模型的条件均值估计量建立了有限维递归方法,其中主要的参数估计方法采用了最大似然估计(MLE)方法。同时,上述学者也研究了离散时间状态、连续时间状态、离散时间观察过程、连续时间观察过程、有限状态和连续状态的隐藏马尔科夫模型。Finesso (1990);Kieffer (1993) 以及 Liu,Narayan (1994) 研究了针对有限离散状态的隐藏马尔科夫模型的强一致顺序估计量的信息理论方法。Ryden (1991) 对隐藏马尔科夫模型建立了顺序估计量。Ryden (1997) 研究了隐藏马尔科夫模型的参数的一致渐近正态递归估计量。Robert, Celeux 和 Diebold (1993) 等学者研究了一般隐藏马尔科夫模型的 Gibbs 抽样的贝叶斯参数估计方法。

在通信和信息理论中,隐藏马尔科夫模型的许多方面也得到了研究。Bahl, Cocke, Jelined 和 Raviv (1974)① 提出运用 Chang 和 Hang 的向前-向后递归方法来对卷积和线性码进行最小误码率解码,这个算法又被称为 BCJR 算法,是 Turbo 码译码的稳定版本。Turbo 码使用了多个级联卷积码和一个反馈机制,减少了误码率的迭代。他们达到了 Shannon 容量,超过了高斯信道。Fontana, Gray 和 Kieffer (1980)② 研究了复合源的特性,即隐藏马尔科夫模型。1978 年他们将 Lempel-Ziv 通用数据压缩算法应用到了有限状态隐藏马尔科夫模型的通用解码中。这个算法渐近优于来自任何来源(不局限于隐藏马尔科夫模型)的压缩序列的任何有限状态编码方案。Merhav (1991)③ 建立了关于隐藏马尔科夫模型的 Lempel-Ziv 算法的大偏差性质。Ziv (1988);Gutman(1989) 以及 Zeitounin, Ziv, Merhav (1992) 等学者运用实际观察序列,在任意阶的马尔科夫链的通用分类上取得了显著进步。Merhav (1991);Ephraim,Merhav (1991) 以及 Kieffer (1993) 运用实际观察训练序列建立了隐藏马尔科夫模型通用分类方法。Ziv (1985); Lapidoth,Ziv (1998) 对有限状态信道建立了通用编译算法。Kaleh 和 Vallet (1994) 运用 Baum 算法建立了解码未知的码间干扰信道的算法。

总体来说,隐藏马尔科夫模型具有以下几点特征:

---

① L. R. Bahl et al. Optimal decoding of linear codes for minimizing symbol error rate, *IEEE Transactions on Information Theory*, 1974,1T-20, pp. 284-287.

② R. J. Fontana et al. Limit theorems for slowly varying composite sources, *IEEE Transactions on Information Theory*, 1980,1T-26, pp. 702-709.

③ N. Merhav. Universal classification for hidden Markov models, *IEEE Transactions on Information Theory*, 1991,37, pp. 1586-1594.

第一，基于马尔科夫性的双重随机过程，隐藏马尔科夫模型有较广的应用范围与较强的适用性；

第二，隐藏马尔科夫模型是一种不完全数据的统计模型（包含了非线性滤波模型），使得其既能反映现象的随机性，又能反映现象潜在的结构，便于运用结构与局部联系性质等方面的知识；

第三，隐藏马尔科夫模型是很少的几个既能从物理模型出发，又能与数据拟合相直接联系的算法之一；

第四，隐藏马尔科夫模型虽然具有黑箱的某些特点，又与纯黑箱操作的人工神经网络等算法相比有较明显的优点，其参数往往有较为实质的含义。

随着隐藏马尔科夫模型相关理论研究的进步，近年来在许多领域出现了大量关于隐藏马尔科夫模型的应用。比如生物学、经济学、控制论、谱估计、雷达、声呐和图像信号处理、故障检测、计算机视觉、人工智能和计量学等。

1.2 相关问题

隐藏马尔科夫模型是一种重要的统计分析模型，也是一种能更好地描述现实中的事件的方法。近十几年来，该模型的相关理论不仅在工程控制、人工智能和生物医疗等多个领域得到广泛应用，在金融、经济等领域中的应用也越来越多。

本书介绍了隐藏马尔科夫模型的基础知识，包括马尔科夫过程、隐藏马尔科夫模型的基本定义和性质，重点介绍了隐藏马尔科夫模型的三个基本算法、极限理论以及条件异方差性等内容。

书中通过实际案例提出了三个基本问题，分别是评估问题、解码问题和学习问题，进而导出了对应的三个经典算法，分别为向前-向后算法、Viterbi 算法和 Baum-Welch 算法。通过解决以上三个基本问题，可以实现隐藏马尔科夫模型在金融领域中的应用，弥补计量回归模型的某些缺陷，完成许多较为复杂的工作和任务。目前，Matlab软件和R软件等都能较好地解决这些问题，给出具体问题的分析结果。书中给出了一些金融中的实证分析案例，结果显示，隐藏马尔科夫模型在金融领域的数据序列波动性分析中有更好的拟合效果。

相对于隐藏马尔科夫模型的广泛应用，模型的理论研究成果仍然较少，且均属于经典型结果，还有较大的理论研究空间，同时这些研究都将具有相当的难度和挑战性。

第 2 章　马尔科夫过程

2.1　离散时间的马尔科夫链

概率论最重要的一个分支就是随机过程，它被广泛地应用于物理学、生物学、通信科学以及金融学等领域，几乎无所不在。

最初安德烈·马尔科夫(A. A. Markov)研究了有特定相依性的随机变量序列的性质，后人称具有这样性质的随机过程为马尔科夫链。马尔科夫链按照状态分类可以分为离散时间的马尔科夫链和连续时间的马尔科夫链。其中离散时间的马尔科夫链按照性质又可以分为时齐、常返性与暂态等类别。

安德烈·马尔科夫所建立的概率模型，不但具有深刻的哲学意义，而且具有客观的物质背景。与此同时，更复杂的随机过程也逐步出现在相关研究中，只不过当时人们还很难认识到这类模型的普遍意义或无法用精确的数学语言表述出来罢了。例如苏格兰植物学家布朗(R. Brown, 1773—1858)于 1827 年发现的悬浮微粒的无规则运动、英格兰遗传学家高尔顿(F. Galton, 1822—1911)于1889年提出的家族遗传规律、荷兰物理学家埃伦费斯特(P. Ehrenfest, 1880—1933)于1907年研究的关于容器中分子扩散的实验以及传染病感染的人数、谣言的传播、原子核中自由电子的跃迁、人口增长的过程等，都可以用马尔科夫过程来描述。这些应用也正是在统计物理、量子力学、遗传学以及社会科学的若干新课题、新事物中所需要的。

马尔科夫过程是一个典型的随机过程。设 $\xi(t)$ 是一个随机过程，当该过程在时刻 t_0 所处的状态为已知，在时刻 $t(t>t_0)$ 所处的状态与在 t_0 时刻之前的状态无关时，这个特性称为无后效性。无后效的随机过程称为马尔科夫过程。马尔科夫过程中的时间和状态既可以是连续的，又可以是离散的。人们又称状态离散的马尔科夫过程为马尔科夫链。马尔科夫链中各个时刻的状态的转变由一个状态转移的概率矩阵控制。马尔科夫随机过程的发展史说明了理论与实际之间的密切关系。

2.1.1 离散马尔科夫性质和状态分类

1. 离散定义与马尔科夫性质

定义 2.1 随机过程可能取到的值(状态)组成的集合记为 S,称为**随机过程的状态空间**。随机序列$\{\xi_n:n\geqslant 0\}$ 称为**马尔科夫链**,如果这些随机变量都是离散的(也就是说状态空间 S 至多是一个可数集,即或者 S 是有限集,或者 S 可以与自然数一一对应),而且对于 $\forall n\geqslant 0$ 及任意状态 $i,j,i_0,\cdots,i_{n-1}$,都有

$$P(\xi_{n+1}=j\mid\xi_n=i,\xi_{n-1}=i_{n-1},\cdots,\xi_0=i_0)=P(\xi_{n+1}=j\mid\xi_n=i)=P_{ij}\text{。}\tag{2-1}$$

对式(2-1)可以理解为,马尔科夫链中,在已知过去状态 $\xi_{n-1},\cdots,\xi_0$ 和现在状态 ξ_n 的条件下,将来的状态 ξ_{n+1} 与过去的状态相互独立,ξ_{n+1} 只依赖于现在状态 ξ_n。这个性质称为马尔科夫性质。其中 P_{ij} 表示处于状态 i 时下一步转移到状态 j 的概率。

性质 2.1 (随机事件的马尔科夫性)若 A 为

$$\{\xi_0(\omega)=i_{0,k},\xi_1=i_{1,k},\cdots,\xi_{n-1}=i_{n-1,k}\}$$

的事件的并,即 A 为$n+1$时刻之前所有会转移到状态 k 的随机变量的集合,其中 k 也是一个随机变量,A 可以表示为

$$A=\bigcup_k\{\xi_0=i_{0,k},\xi_1=i_{1,k},\cdots,\xi_{n-1}=i_{n-1,k}\},$$

则有

$$P(\xi_{n+1}=j\mid A,\xi_n=i)=P(\xi_{n+1}=j\mid\xi_n=i)\text{。}\tag{2-2}$$

式(2-2)说明过去的状态无论怎样转移,未来的状态只与现在的状态有关。

性质 2.2 (等价性质)对于过程在时刻 $n+1$ 及其后的时刻所确定的事件 $B=\{\xi_{n+m}=i_{n+m}\}$ 及性质 2.1 中的 A 有

$$P(B\mid A,\xi_n=i)=P(B\mid\xi_n=i)\text{。}\tag{2-3}$$

性质 2.3 (等价性质)在已知“现在”的条件下,“将来”与“过去”是条件独立的,即

$$P(AB\mid\xi_n=i)=P(A\mid\xi_n=i)P(B\mid\xi_n=i)\text{。}\tag{2-4}$$

式(2-4)的含义为:过程在时刻 n 处于状态 i 的概率,不管该过程以前处于什么状态,过程以后处于什么状态,是一样的。这说明,马尔科夫链在已知“现在”的条件下,“将来”与“过去”是相互独立的。等价性质 2.3 其实是等价性质 2.2 的一个变

化形式，故式(2-4)与式(2-3)表示的含义实际上是一样的。

性质 2.4　(等价性质) 对马尔科夫链$\{\xi_n\}$及$\forall m \geqslant 1, n \geqslant 0$及任意状态$i, j$,

$i_0, \cdots, i_{n-1}$，有

$$P(\xi_{n+m} = j \mid \xi_n = i, \xi_{n-1} = i_{n-1}, \cdots, \xi_0 = i_0) = P(\xi_{n+m} = j \mid \xi_n = i)。\tag{2-5}$$

性质 2.5　(函数的马尔科夫性) 对状态空间S上的任意有界实值函数f有

$$E(f(\xi_{n+1}) \mid \xi_0 = i_0, \cdots, \xi_n = i_n) = E(f(\xi_{n+1}) \mid \xi_n = i_n)。\tag{2-6}$$

式(2-1)是式(2-6)的特例，即$f(x) = I_{\{j\}}(x)$的情形。

性质 2.6　(一般形式) 对于常见的实数集合Λ，及由随机序列$\{\xi_m\}$在时刻n及其后的信息所决定的随机变量η_n，恒有

$$P(\eta_n \in \Lambda \mid \xi_0 = i_0, \cdots, \xi_n = i_n) = P(\eta_n \in \Lambda \mid \xi_n = i_n) \tag{2-7}$$

或

$$E(\eta_n \mid \xi_0 = i_0, \cdots, \xi_n = i_n) = E(\eta_n \mid \xi_n = i_n)。$$

注　这个等价条件的内涵十分丰富，其等价性的证明在测度论中非常典型，本书不再详述，有兴趣的读者可以阅读测度论的相关书籍。

由马尔科夫性可知，当给定马尔科夫链的初始分布，其特性完全由条件概率$P(\xi_{n+1} = j \mid \xi_n = i)$决定。因此，在马尔科夫链的理论与应用研究中，人们最关心的是，当经过一定时间之后，随机变量到达某些状态的概率有多大，以及需要多长时间才能到达这些状态。这类问题的描述首先涉及马尔科夫链的状态转移特性——马尔科夫链的n步转移概率矩阵族。

转移概率矩阵$\boldsymbol{P}$是一个随机矩阵，矩阵内的元素是每一次的状态转移概率，因此元素需要满足概率的性质：

(1) $\boldsymbol{P}$中的元素均为非负，即$p_{ij} \geqslant 0$；

(2) $\boldsymbol{P}$中的每一行的元素之和均为1，即$\sum_j p_{ij} = 1$。

马尔科夫链的转移概率$\{p_{ij}, i, j \in S\}$(或者说转移概率矩阵$\boldsymbol{P}$)是刻画马尔科夫链的统计特征的基本量。

定义 2.2　马尔科夫链称为时齐的，如果其转移概率$P(n, n+1)$与n无关，我们把此矩阵简记为$\boldsymbol{P} = (p_{ij})$。

由式(2-7)得到

$$P(n, n+m) = P(n, n+1)P(n+1, n+2)\cdots P(n+m-1, n+m) = \boldsymbol{P}^{(m)}。$$