贝叶斯计量经济模型

朱慧明　林 静　/著

科学出版社
www.sciencep.com

贝叶斯计量经济模型

朱慧明　林　静　著

科　学　出　版　社
北　京

内 容 简 介

本书系统地研究了计量经济模型的贝叶斯理论及其应用，主要内容包括贝叶斯统计计算、统计分布理论、贝叶斯决策理论、贝叶斯线性回归模型、贝叶斯自回归移动平均模型、贝叶斯向量自回归模型、贝叶斯自回归条件异方差模型和贝叶斯随机波动模型.

本书可以作为统计学、计量经济学和管理科学与工程专业高年级本科生、硕士研究生和博士研究生的教材，也可以作为高校教师、研究人员和科技人员的参考书.

图书在版编目(CIP)数据

贝叶斯计量经济模型/朱慧明，林静著. —北京：科学出版社，2009
ISBN 978-7-03-025140-4
Ⅰ.贝… Ⅱ.①朱… ②林… Ⅲ.计量经济学-经济模型-研究 Ⅳ.F224.0
中国版本图书馆 CIP 数据核字(2009) 第 133515 号

责任编辑：陈玉琢／责任校对：陈玉凤
责任印制：钱玉芬／封面设计：王 浩

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

源海印刷有限责任公司 印刷

科学出版社发行 各地新华书店经销

*

2009 年 9 月第 一 版 开本：B5(720×1000)
2009 年 9 月第一次印刷 印张：13
印数：1—2 500 字数：243 000

定价：48.00 元

（如有印装质量问题，我社负责调换〈环伟〉）

前言

贝叶斯理论源于英国学者贝叶斯 (Thomas Bayes) 于 1763 年在英国皇家学会哲学学报上发表的论文 “论有关机遇问题的求解”, 并在宏观经济预测、金融工程与风险管理、市场营销、人工智能与模式识别、统计质量控制与可靠性工程等很多领域获得了广泛应用. 本书目的是向读者介绍计量经济模型的贝叶斯理论及其应用, 包括贝叶斯统计计算、统计分布理论、贝叶斯决策、贝叶斯线性回归模型和贝叶斯时间序列计量经济模型. 全书共分为 8 章, 具体内容安排如下:

第 1 章, 贝叶斯统计计算. 介绍贝叶斯理论的基本思想, 随机数生成方法, Monte Carlo 和 Markov chain Monte Carlo 计算技术.

第 2 章, 统计分布理论. 介绍贝叶斯计量经济模型理论体系中常用的统计分布及其性质, 包括伽玛分布族、正态分布族、Wishart 分布族, 以及多元 t 分布和矩阵 t 分布.

第 3 章, 贝叶斯决策理论. 探讨位置–尺度参数的扩散先验分布和参数共轭先验分布的构造, 以及贝叶斯风险决策解.

第 4 章, 贝叶斯线性回归模型. 研究线性回归模型的贝叶斯理论, 包括贝叶斯多元线性回归模型参数、参数线性假设的贝叶斯检验、随机误差序列自相关的贝叶斯诊断方法和多重线性回归模型的贝叶斯理论.

第 5 章, 贝叶斯自回归移动平均模型. 研究自回归移动平均模型的贝叶斯理论, 包括贝叶斯 AR 模型、贝叶斯 MA 模型、贝叶斯稳健 ARMA 模型和贝叶斯 ARFIMA 模型.

第 6 章, 贝叶斯向量自回归模型. 研究向量自回归模型的贝叶斯理论, 包括贝叶斯非限制性 VAR 模型、贝叶斯限制性 VAR 模型和贝叶斯 VARFIMA 模型.

第 7 章, 贝叶斯自回归条件异方差模型. 研究自回归条件异方差模型的贝叶斯理论, 包括贝叶斯 GARCH 模型、贝叶斯 AR-GJR-GARCH 模型和贝叶斯 SW-GARCH 模型.

第 8 章, 贝叶斯随机波动模型. 研究随机波动模型的贝叶斯理论, 包括贝叶斯 SV-N 模型, 贝叶斯 SV-T 模型, 贝叶斯 SV-GED 模型, 贝叶斯 SV-MN 模型和贝叶斯 SV-MT 模型.

本书第 1 章和第 2 章由朱慧明和林静完成, 第 3~8 章由朱慧明完成. 在本书的撰写过程中, 武汉大学数学与统计学院邹新堤教授, 英国皇家统计学会院士、英国 Brunel 大学数学系教授 Yu Keming 博士提出了许多宝贵意见; 同时, 郑进城、邓

迎春、李锋、曾惠芳、郝立亚和李素芳等研究生做了大量相关研究工作. 本书部分研究成果得到了国家自然科学基金项目 (70771038)、教育部新世纪优秀人才支持计划项目 (NCET050704) 和教育部人文社科规划项目 (06JA910001) 的资助. 在本书的出版过程中, 得到了科学出版社陈玉琢老师的大力支持及帮助. 作者在此表示衷心感谢!

由于作者水平有限, 书中必定存在疏漏与不足, 恳请专家和读者提出宝贵的意见.

作　者

2009 年 5 月

符号缩写及说明

tr	矩阵的迹
$\mathbf{0}$	零向量或零矩阵
$\boldsymbol{A}^-$	矩阵 $\boldsymbol{A}$ 的广义逆
$\propto$	正比符号
$\theta_{\sim i}$	向量 θ 中不含分量 θ_i
sup	上确界
inf	下确界
vec	矩阵的向量化算子
$\theta_{\sim i}=(\theta_1,\theta_2,\cdots,\theta_{i-1},\theta_{i+1},\cdots,\theta_m)$	不包含第 i 个分量的参数向量
$\mathbf{1}_n$	分量均为 1 的 n 维列向量
$\boldsymbol{J}_n$	各元素均为 1 的 $n\times n$ 矩阵
$\lvert\boldsymbol{A}\rvert_+$	矩阵 $\boldsymbol{A}$ 行列式的绝对值
$\pi(\theta)$	参数 θ 的先验分布
$\pi(\theta\lvert\cdot)$	参数 θ 的后验分布
i.i.d.	相互独立同分布
$\mathrm{rank}(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 的秩
$\boldsymbol{A}\otimes\boldsymbol{B}$	矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 的 Kronecker 积
$X\overset{d.f}{=}Y$	随机变量 X 与 Y 服从相同的统计分布
$J(\boldsymbol{A}\to\boldsymbol{B})$	矩阵 $\boldsymbol{A}$ 到矩阵 $\boldsymbol{B}$ 变换的 Jacobian 行列式
$o(n^{-1/2})$	$\lim\limits_{n\to\infty} o(n^{-1/2})/n^{-1/2}=0$
$\boldsymbol{X}\sim Mt_k(v,\boldsymbol{\theta},\boldsymbol{P})$	k 维随机向量 $\boldsymbol{X}$ 服从多元 t 分布
$\boldsymbol{X}\sim Mt_{k_1\times k_2}(v,\boldsymbol{\theta},\boldsymbol{\Sigma}\otimes\boldsymbol{Q})$	$k_1\times k_2$ 随机矩阵 $\boldsymbol{X}$ 服从矩阵 t 分布
$\boldsymbol{X}\sim MN_{k_1\times k_2}(\boldsymbol{\theta},\boldsymbol{\Sigma}\otimes\boldsymbol{\Psi})$	$k_1\times k_2$ 随机矩阵 $\boldsymbol{X}$ 服从矩阵正态分布
$\boldsymbol{X}\sim W_m(v,\boldsymbol{\Sigma})$	$m\times m$ 随机矩阵 $\boldsymbol{X}$ 服从 Wishart 分布
$\boldsymbol{X}\sim IW_m(v,\boldsymbol{\Sigma})$	$m\times m$ 随机矩阵 $\boldsymbol{X}$ 服从逆 Wishart 分布

目　　录

第1章　贝叶斯统计计算

本章介绍贝叶斯理论的基本思想, 随机数的生成方法, Monte Carlo 计算, MH 抽样、Gibbs 抽样和收敛性诊断等贝叶斯计算理论与方法.

1.1　贝叶斯理论

贝叶斯 (Thomas Bayes), 英国数学家, 1702 年出生于伦敦, 1742 年成为英国皇家学会会员, 1763 年 4 月 7 日逝世. 贝叶斯在数学方面主要研究概率论. 他首先将归纳推理法用于概率论基础理论, 并创立了贝叶斯统计理论, 对于统计决策函数、统计推断、统计估算等作出了重要贡献. 贝叶斯于 1763 年在皇家学会学报上发表了著名论文 "机会的学说概论"(*An essay towards solving a problem in the doctrine of chances*), 该文提出了从二项分布的观察值出发对其参数进行概率推断的方法; 一般认为, 贝叶斯假定参数服从单位区间上的均匀分布, 他所提出的有关二项分布参数推断的方法后来被称为贝叶斯定理, 并且被推广到其他统计分布推断问题.

从某种意义上讲, 贝叶斯方法中的许多思想, 可以追溯到 1713 年 Bernoulli 的研究成果; 在其概率专著 *Ars conjectandi* 中, Bernoulli 不仅发展了二项式定理, 提出了交换律和结合律, 而且还提出了贝叶斯逆概率 (inverse probability) 问题; 尽管如此, Bernoulli 并没有提出逆概率的数学结构. 根据 Stigler 的观点, Laplace 于 1774 年论述了一般形式下的逆概率定理. 1939 年, Jeffreys 重新发现了 Laplace 的研究工作. 但是, 1774~1939 年的科学文献中似乎没有令人信服的东西, 因此, 可以说此段时间内的贝叶斯方法是悄无声息、无人问津的. 如果把 1900 年算作是近代数理统计开始的第一年, 则到 Cramer 的专著 *Mathematical Methods of Statistic* 出版为止的随后 50 多年中, 可以说基本上是经典统计 (即频率统计) 方法一统天下的局面.

20 世纪 50 年代以后, 随着统计分析技术应用范围的扩大, 决策问题在统计应用中占有越来越重要的地位, 贝叶斯方法在实践中也逐步得到应用, 并得到了迅速发展; 而对决策分析而言, 先验知识的利用是不可或缺的, 在这类问题中主观概率的提法也比较自然, 它反映了决策者掌握信息的程度, 因此贝叶斯观点易于接受. 20 世纪初, 数学的公理化倾向影响到统计界, 概率的公理化最终由 Kolmogorov 完成, 并得到了普遍认同, 即概率是正则化的测度; 然而, 对主观概率应该用什么样的公理来描述, 经过很长时间的讨论, 还没有比较一致的看法. 意大利学者 de Finetti 认

为 Kolmogorov 公理中互斥事件之和的概率等于各事件概率之和的结论, 对于信念而言是不成立, 只能是有限可加; de Finetti 关于可换随机变量的研究成果为先验分布的客观性提供了理论基础; Fisher 的似然原理促进了贝叶斯理论的发展, 似然函数是贝叶斯方法应用的基点和支柱, 从最大似然估计到最大后验估计, 从似然比到后验比, 贝叶斯推断理论和方法获得了系统的论述. 英国学者 Jeffreys 在无信息先验分布理论领域的重要突破, 形成了后来称为的 Jeffreys 准则, 其专著 *Probability Theory* 的出版则标志着贝叶斯理论体系的形成; 后来 Good 对主观概率的研究, Savage 对效应函数、主观概率以及它们之间相互关系的研究, 把贝叶斯理论推进到一个新阶段.

在统计推断的基本理论和方法方面, 贝叶斯方法与频率统计方法之间存在本质性差异, 这主要表现在三个方面.

首先, 频率统计方法在进行推断分析过程时, 主要依据两类信息: 一是模型信息, 即总体服从何种概率分布, 这是制定统计方法的基础; 另一个是样本信息, 即观测值或试验结果. 贝叶斯方法则除了以上两类信息外, 还利用另外一类信息, 即总体分布中未知参数的分布信息.

其次, 频率统计方法坚持概率的频率解释, 并在这个基础上去理解一切统计推断的结论, 如在 Neyman 的区间估计理论中, "某区间估计 $(\underline{\theta},\overline{\theta})$ 的置信水平为 $1-\alpha$" 这一推断, 此处 θ 应理解为一无随机性的未知参数, 当区间估计 $(\underline{\theta},\overline{\theta})$ 反复大量使用时, 100 次中大约平均有 $(1-\alpha)\times 100$ 次包含了参数 θ. 与此相反, 贝叶斯方法赞成主观概率; 也就是说, 概率是认识主体对事件出现可能性大小的相信程度, 它并不依赖事件能否重复.

第三, 二者的推断理念之间存在着根本性差异. 统计学奠基人 Fisher 在其 1921 年发表的论文 "理论统计的数学基础" 中将统计学的任务概括为三个问题: 选定模型、确定统计量和决定统计量的分布. 根据 Fisher 的观点, 信息包含在样本中, 但样本为数众多, 因此须用少数几个统计量把信息集中起来, 而抽样分布则决定了统计量的全部性质. 目前, 频率统计方法基本上都是按照这种思路来处理统计推断问题的.

与此不同, 贝叶斯推断的一般模式是: 先验信息 $\oplus$ 样本信息 $\Longrightarrow$ 后验信息, 此处符号 "$\oplus$" 表示为贝叶斯定理的作用. 在贝叶斯理论体系中, 先验分布反映了试验前对总体参数分布的认识, 在获得样本信息后, 人们对这个认识有了改变, 其结果就反映在后验分布中, 即后验分布综合了参数先验分布和样本信息. 由此可以看出, 频率统计推断是 "从无到有" 的过程: 在试验前, 关于未知参数的情况是一无所知, 而试验后则有些了解, 但对了解多少并无普遍的表述方法, 在实践中有赖于所使用的统计量的针对性. 贝叶斯推断则不然, 它是一个 "从有到有" 的过程, 结果清楚自然, 符合人们的思维习惯: 根据所获得的信息修正以前的看法, 不一定从零开

始. 因此, 从本质上说, 贝叶斯推理方法概括了成人的学习过程.

第四, 贝叶斯方法只能基于参数的后验分布来分析问题. 也就是说, 在获得后验分布后, 如果把样本、原来的统计模型 (包括总体分布和先验分布) 都丢掉, 一点也不会影响将来的统计推断问题, 凡是符合这个准则的推断就是贝叶斯推断. 因此, 频率统计中的矩估计、显著性统计检验和置信区间估计都不属于贝叶斯推断的范畴; 但是, MLE 估计则可视为均匀先验分布之下的贝叶斯估计, 也就是说, 作为频率学派中一个很重要的极大似然估计, 不过是在一种很特殊先验分布下的贝叶斯估计而已.

贝叶斯理论的哲理有相当大的吸引力, 而且方法简单, 它在统计推断模式上与频率统计的不同之处在于: 在频率统计中, 似然函数概括了模型参数的全部信息, 因此关于参数 θ 的统计推断只要利用似然函数就够了; 而贝叶斯方法既利用了似然函数, 又利用了参数先验信息. 如果先验信息很少或没有先验信息, 这时基于贝叶斯方法与基于频率统计方法所得到的结论基本相同.

与频率统计方法比较, 贝叶斯方法具有如下几个方面的优点:

(1) 充分利用了样本信息和模型参数先验信息, 贝叶斯估计量具有更小的方差或平方误差, 能得到更精确的预测结果;

(2) 贝叶斯 HPD 置信区间比不考虑参数先验信息的频率置信区间短;

(3) 能对假设检验或估计问题所作出的判断结果进行量化评价, 而不是频率统计中的 "接受原假设" 或者 "拒绝原假设" 的简单判断.

目前, 贝叶斯方法在计量经济模型, 保险精算与风险管理, 质量控制与可靠性分析等领域获得了广泛应用.

1.2 随机数的生成

在贝叶斯计量经济模型理论研究领域中, 常常需要利用随机数的生成进行仿真分析, 而大多数概率分布随机数的产生均基于均匀分布 $U(0,1)$ 的随机数. 本节介绍一些常用随机数生成方法, 包括逆变换法、合成方法和筛选抽样等, 以及这些方法在连续随机抽样和多变量抽样中的应用.

1.2.1 逆变换法

逆变换法 (inverse transform method) 是一种应用广泛的随机数生成方法, 它利用了分布函数的一个基本性质: 如果随机变量 X 的分布函数 $F_X(x)$ 是严格单调上升的连续函数, 定义

$$F_X^{-1}(y) = \inf\{x : F_X(x) \geqslant y\}, \quad 0 \leqslant y \leqslant 1 \tag{1.2.1}$$

并设随机变量 U 服从区间 $(0,1)$ 上的均匀分布, 则 $X = F_X^{-1}(U)$ 的分布函数为 $F_X(x)$, 因为

$$P(X \leqslant x) = P(F_X^{-1}(U) \leqslant x) = P(U \leqslant F_X(x)) = F_X(x), \tag{1.2.2}$$

所以, 如果要产生来自分布函数 $F(x)$ 的随机数, 只要先生成均匀分布 $U(0,1)$ 的随机数 u, 然后计算 $F^{-1}(u)$ 即可, 具体过程如下:

(1) 由 $U(0,1)$ 抽取随机数 u;

(2) 计算 $x = F^{-1}(u)$.

例如, 如果 $X \sim U(a,b)$, 则其分布函数为

$$F(x) = \begin{cases} 0, & x < a, \\ \dfrac{x-a}{b-a}, & a \leqslant x < b, \\ 1, & x \geqslant b, \end{cases} \tag{1.2.3}$$

从而 $F^{-1}(y) = a + (b-a)y$, $0 \leqslant y \leqslant 1$, 若由 $U(0,1)$ 抽取 u, 则 $a + (b-a)u$ 就是来自 $U(a,b)$ 的一个随机数.

1.2.2 合成抽样

合成抽样法 (composition sampling method) 是由 Butler(1956) 提出的, 其基本思想是: 如果从随机变量 X 的概率密度函数 $f(x)$ 抽样比较困难, 而 X 关于 Y 的条件密度函数 $f(x|y)$ 及 Y 的密度函数 $g(y)$ 均易于抽样, 则可按如下步骤生成 X 的随机数: 首先由 Y 的分布 $g(y)$ 抽取 y, 然后由条件分布 $f(x|y)$ 抽取 x. Rubinstein(2007) 证明了据此所获得的随机数 X 的概率分布为 $f(x)$.

合成抽样法解决了一些复杂分布的随机数生成问题. 例如, 随机变量 X 的密度函数如下:

$$f_X(x) = n\int_1^{\infty} y^{-n}\mathrm{e}^{-xy}\mathrm{d}y. \tag{1.2.4}$$

显然, 对此分布来说, 很难直接利用逆变换方法生成随机数, 但是利用合成法很容易生成随机数. 为此, 记 $g(y) = ny^{-(n+1)}$, $1 < y < \infty$, 则 $f(x|y) = y\mathrm{e}^{-xy}$; 根据合成抽样, X 的随机数可按如下步骤抽取: 首先由 $U(0,1)$ 抽取随机数 u_1 和 u_2, 然后取 $y = u_1^{-1/n}$, 最后计算 $x = -y^{-1}\ln u_2$ 即可.

1.2.3 筛选抽样

假设要从 $f(x)$ 抽样, 如果可以将 $f(x)$ 分解成三部分: $f(x) = cg(x)h(x)$, $c \geqslant 1$, $0 < g(x) \leqslant 1$, $h(x)$ 是一个概率分布密度函数且易于抽样, 则 X 的随机数抽取程序如下:

(1) 由 $U(0,1)$ 和 $h(y)$ 分别抽取随机数 u 和 y;

(2) 如果 $u \leqslant g(y)$, 则 $x=y$, 并停止抽样;

(3) 如果 $u > g(y)$, 则回到步骤 (1).

这就是 von Neuman(1951) 提出的筛选抽样 (acceptance-rejection method), 其理论依据是如下定理:

定理 1.2.1 设 X 的密度函数为 $f(x)$, 且 $f(x)=cg(x)h(x), c\geqslant 1, 0<g(x)\leqslant 1$, $h(x)$ 是一个密度函数且易于抽样; 令 U 和 Y 分别服从 $U(0,1)$ 和 $h(y)$, 则在 $U \leqslant g(y)$ 的条件下, Y 的条件分布为 $f(x)$, 即 $f_Y(x|U \leqslant g(Y)) = f(x)$.

定理的详细证明请见 Rubinstein(2007). 筛选抽样是一个相当重要的抽样方法, 它解决了许多难以直接抽样分布的抽样问题.

1.2.4 正态分布抽样

假设随机变量 X 服从均值 μ, 方差 σ^2 的正态分布, 即 $X \sim N(\mu,\sigma^2)$, μ,σ^2 已知, 则 $(X-\mu)/\sigma \sim N(0,1)$. 因此, 只需要讨论标准正态分布 $N(0,1)$ 的抽样问题.

标准正态随机数的产生方法有多种, 例如综合利用合成抽样法和筛选法, 近似法等. Box 和 Muller 在 1958 年给出了产生 $N(0,1)$ 随机数的一种比较简单的方法, 其理论依据是以下引理:

引理 1.2.1 设 U_1, U_2 是独立同分布的 $U(0,1)$ 变量, 记

$$\begin{cases} X_1 = (-2\ln U_1)^{1/2}\cos(2\pi U_2), \\ X_2 = (-2\ln U_2)^{1/2}\cos(2\pi U_1), \end{cases} \tag{1.2.5}$$

则 X_1 和 X_2 相互独立, 并且均服从标准正态分布 $N(0,1)$.

根据引理 1.2.1, 生成 $N(0,1)$ 的随机数包括两个步骤: 首先由 $U(0,1)$ 独立抽取 u_1, u_2, 然后应用 (1.2.5) 计算 x_1 和 x_2, 据此可以产生两个标准正态随机数.

1.2.5 随机向量抽样

在利用 Monte Carlo 计算中, 有时需要考虑随机向量的随机数生成问题. 若随机向量的各个分量相互独立, 则它等价于多个一元随机向量的抽样. 下面考虑分量非独立的随机向量抽样.

设 $X_1, X_2, \cdots, X_m$ 的联合概率密度函数为

$$f(x_1,x_2,\cdots,x_m) = f_1(x_1)f_2(x_2|x_1)\cdots f_m(x_m|x_1,x_2,\cdots,x_{m-1}), \tag{1.2.6}$$

其中 $f_1(x_1)$ 为 X_1 的边缘分布密度函数, $f_i(x_i|x_1,x_2,\cdots,x_{i-1})$ 为给定 $X_1, X_2, \cdots, X_{i-1}$ 的条件下 X_i 的条件分布密度函数, $i=2,3,\cdots,m$.

定理 1.2.2 设 $U_1, U_2, \cdots, U_m$ 是独立同分布的 $U(0,1)$ 变量, F_i 是与 f_i 相对

应的分布函数, $X_1, X_2, \cdots, X_m$ 是方程

$$\begin{cases} F_1(X_1) = U_1, \\ F_i(x_i|x_1, x_2, \cdots, x_{i-1}) = U_i, \quad i = 2, \cdots, m \end{cases} \tag{1.2.7}$$

的解, 则 $X_1, X_2, \cdots, X_m$ 的联合分布密度函数为 $f(x_1, x_2, \cdots, x_m)$.

证明 请读者自己证明.

定理 1.2.2 给出了随机向量的逆变换抽样方法：首先由 $U(0,1)$ 独立地抽取 $u_1, u_2, \cdots, u_m$, 然后解方程 (1.2.7) 即可获得 $x_1, x_2, \cdots, x_m$. 需要注意的是, 由于联合分布密度函数的分解 (1.2.6) 可能并不唯一, 因此方程 (1.2.7) 的解可能不一样, 从而抽样的难易程度也会有所不同.

作为特定分布随机向量的抽样方法, 我们在此介绍多元正态变量的抽样.

设 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_m$ 服从多元正态分布 $N(\boldsymbol{\mu}, \Sigma)$, $\boldsymbol{\mu} \in \mathbf{R}$, $\boldsymbol{\Sigma} = (\sigma_{ij}) > 0$ 已知, 其分布密度函数为

$$f(x_1, x_2, \cdots, x_m) = \frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}(\boldsymbol{x}-\boldsymbol{\mu}) \right). \tag{1.2.8}$$

由于 $\boldsymbol{\Sigma}$ 是一个正定对称矩阵, 因此存在下三角阵

$$\boldsymbol{C} = \begin{pmatrix} c_{11} & 0 & \cdots & 0 \\ c_{21} & c_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mm} \end{pmatrix} \tag{1.2.9}$$

使得 $\boldsymbol{\Sigma} = \boldsymbol{C}\boldsymbol{C}'$, 于是, 若 $\boldsymbol{Z} = (Z_1, Z_2, \cdots, Z_m)' \sim N(\boldsymbol{0}, \boldsymbol{I}_m)$, 则 $\boldsymbol{X} = \boldsymbol{C}\boldsymbol{Z} + \boldsymbol{\mu} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. 为了获得多元正态分布的随机数, 首先, 由 $X_1 = c_{11}Z_1 + \mu_1$, 有 $\sigma_{11} = c_{11}^2$, 因此 $c_{11} = \sigma_{11}^{1/2}$; 其次, $X_2 = c_{21}Z_1 + c_{22}Z_2 + \mu_2$, 于是

$$\sigma_{22} = c_{21}^2 + c_{22}^2, \quad \sigma_{12} = \mathrm{Cov}(X_1, X_2) = c_{11}c_{21}. \tag{1.2.10}$$

因此

$$c_{21} = \frac{\sigma_{12}}{\sigma_{11}^{1/2}}, \quad c_{22} = \left(\sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}\right)^{1/2}. \tag{1.2.11}$$

如此反复进行, 可以得到一般的迭代公式:

$$c_{ij} = \frac{\sigma_{ij} - \sum\limits_{l=1}^{j-1} c_{il}c_{jl}}{\left(\sigma_{jj} - \sum\limits_{t=1}^{j-1} c_{jt}^2\right)^{1/2}}, \quad i = 1, 2, \cdots, m, \quad j = 1, 2, \cdots, i. \tag{1.2.12}$$