


MCMC BOOK
(Markov Chain Monte Carlo)

Renhe W.

Preface

蒙特卡罗的方法在工程、金融以及统计等领域应用十分广泛，蒙特卡罗技术的研究需要广泛领域的详细知识；例如，描述随机实验和过程的**概率**，分析数据的**统计**，有效实现算法的**计算科学**，以及制定和解决优化问题的**数学规划**。一般运用起来，若没有充足的理论基础，很难去实施，本书的目的就是为了便于查找和运用相关方法。

这本手册的目的是提供蒙特卡洛技术及相关主题的易于理解和全面的汇编。它包含了**理论（概述）**、**算法（伪代码和实际代码）**以及**应用**。该书旨在成为蒙特卡洛方法的基本指南，供高年级本科生和研究生/研究人员使用，以快速查阅想法、流程、公式、图表等内容，而不仅仅是一本研究专著或教材。

由于蒙特卡洛方法可以用于多种方式和多种不同目的，因此该手册的组织方式是作为独立章节的集合，每个章节专注于一个单独的主题，而不是按照数学推导的方式组织。理论部分与书中讨论相关主题的其他部分进行了交叉引用，边栏的符号  指向相应的页码。理论部分以示例和 MATLAB 代码进行了说明。

Contents

第一章 均匀随机数的生成	1
1.1 随机数	1
1.1.1 好的随机数生成器的性质	2
第二章 Markov Chain Monte Carlo	3
2.1 METROPOLIS-HASTINGS 算法	3
第三章 Variance Reduction	5
3.1 Importance Sampling	5

List of Figures

List of Tables

CHAPTER 1

均匀随机数的生成

这一章概述了生成均匀随机数的主要技术和算法，包括基于线性递归、模 2 算术以及这些方法的组合。提供了一系列理论和实证测试来评估均匀随机数生成器的质量。将在第三章中讨论从任意分布生成随机变量的方法，这些方法无一例外地基于均匀随机数生成器。

1.1 随机数

任何蒙特卡罗方法的核心都是一个随机数生成器：一个产生无限流的过程

$$U_1, U_2, U_3, \dots, \sim Dist$$

是根据某些概率分布 $Dist$ 的独立同分布 (iid) 的随机变量。当该分布是在区间 $(0,1)$ (即 $Dist = (0,1)$) 上的均匀分布时，该生成器被称为均匀随机数生成器。大多数计算机语言已经包含了一个内置的统一随机数生成器。用户通常只被要求输入一个初始数字，称为 **Seed**，在调用时，随机数生成器在区间 $(0,1)$ 上产生一系列独立的均匀随机变量序列。例如，在 Matlab 中，由 `rand` 函数提供。

无限独立同分布随机变量序列的概念是一个数学抽象，可能无法在计算机上实现。在实践中，人们最多只能希望产生一系列具有与真实独立同分布随机变量序列无法区分的统计特性的“随机”数字。虽然基于普遍背景辐射或

量子力学的物理生成方法似乎提供了这种真正随机性的稳定来源，但当前绝大多数随机数生成器都基于可以在计算机上轻松实现的简单算法。根据[?]的观点，这种算法可以表示为一个元组 $(S, f, \mu, \mathcal{U}, g)$ ，其中

- S 是有限状态集，
- f 是从 S 到 S 的函数，
- μ 是 S 上的概率分布，
- \mathcal{U} 是输出空间；对于均匀随机数生成器， \mathcal{U} 是区间 $(0, 1)$ ，除非另有说明，
- g 是从 S 到 \mathcal{U} 的函数。

1.1.1 好的随机数生成器的性质

CHAPTER 2

Markov Chain Monte Carlo

Markov chain Monte Carlo (MCMC) 从任意分布近似抽样的通用方法。其主要思想是生成一个极限分布等于期望分布的马尔可夫链。在本章中，我们将描述最杰出的 MCMC 算法：

1. Metropolis-Hastings 算法，特别是独立采样器和随机游走采样器；
2. 吉布斯采样器，在贝叶斯分析中特别有用；
3. Hit-and-run 采样器——通常用于具有高度约束参数空间的贝叶斯设置和一般的罕见事件模拟问题；
4. Shake-and-bake 算法是一种在多面体表面均匀分布点的实用方法；
5. Metropolis-Gibbs 混合算法和多重尝试的 Metropolis-Hastings 方法，其中结合了来自不同的 MCMC 算法的思想；
6. 辅助可变采样器，如切片采样器和 Swendsen-Wang 算法；
7. Reversible-jump 采样器，在贝叶斯模型选择中有应用。

2.1 METROPOLIS-HASTINGS 算法

3.1 Importance Sampling

重要性抽样 (Importance Sampling) 是一种用于估计概率分布性质的统计方法，特别是在计算期望值或概率密度函数的归一化常数时非常有用。它通过从一个不同的提议分布中抽样，来估计原始分布的性质。重要性抽样是最重要的方差减小技术之一。这种技术对于估计罕见事件的概率（见第 10 章）特别有用。标准设置是估计一个量：

$$\ell = \mathbb{E}_f H(\mathbf{X}) = \int H(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}, \quad (3.1)$$

其中 H 是一个实值函数， f 是随机向量 \mathbf{X} 的概率密度，称为名义概率密度。下标 f 被添加到期望算子以表示它是相对于密度 f 进行的。

令 g 是另一个概率密度，使得 Hf 被 g 所主导。也就是说， $g(\mathbf{x}) = 0 \Rightarrow H(\mathbf{x})f(\mathbf{x}) = 0$ 。使用密度 g ，我们可以表示 ℓ 如下：

$$\ell = \int H(\mathbf{x}) \frac{f(\mathbf{x})}{g(\mathbf{x})} g(\mathbf{x}) d\mathbf{x} = \mathbb{E}_g H(\mathbf{X}) \frac{f(\mathbf{X})}{g(\mathbf{X})}. \quad (3.2)$$

因此，如果 $\mathbf{X}_1, \dots, \mathbf{X}_N \sim_{\text{idd}} g$ ，那么

$$\hat{\ell} = \frac{1}{N} \sum_{k=1}^N H(\mathbf{X}_k) \frac{f(\mathbf{X}_k)}{g(\mathbf{X}_k)} \quad (3.3)$$

是 ℓ 的无偏估计。这个估计被称为重要性抽样估计, g 被称为重要性抽样密度. 密度比值为:

$$W(\mathbf{x}) = \frac{f(\mathbf{x})}{g(\mathbf{x})}, \quad (3.4)$$

被称为似然比 (本书定义), 因为似然通常在统计学中视为参数的函数 (见第 B.2 节) .

算法 3.1 (重要性抽样估计). 重要性抽样估计的具体算法步骤为:

1. 选择一个主导 Hf 的重要性抽样密度 g .
2. 生成独立同分布的样本 $\mathbf{X}_1, \dots, \mathbf{X}_N \stackrel{iid}{\sim} g$, 并令 $Y_i = H(\mathbf{X}_i) f(\mathbf{X}_i) / g(\mathbf{X}_i), i = 1, \dots, N$.
3. 通过 $\hat{\ell} = \bar{Y}$ 估计 ℓ , 并确定一个近似的 $1 - \alpha$ 置信区间, 如下所示:

$$\left(\hat{\ell} - z_{1-\alpha/2} \frac{S}{\sqrt{N}}, \hat{\ell} + z_{1-\alpha/2} \frac{S}{\sqrt{N}} \right),$$

其中 z_γ 表示 $N(0, 1)$ 分布的 γ 分位数, S 是 Y_1, \dots, Y_N 的样本标准差。