

文章主要介绍了如下方面：

- **MCMC的动机**：MCMC方法可以用来解决高维空间中的积分和优化问题，这些问题在贝叶斯推断、统计物理、模型选择等领域中经常出现。
- **蒙特卡罗原理**：MCMC方法的基本思想是从目标分布 $p(x)$ 中抽取一组独立同分布的样本 $x(i) \ N i = 1$ ，然后用这些样本来近似 $p(x)$ 或者关于 $p(x)$ 的积分。
- **拒绝采样和重要性采样**：当 $p(x)$ 的标准形式不可得时，可以使用拒绝采样和重要性采样来生成样本。拒绝采样需要找到一个易于采样的建议分布 $q(x)$ ，使得 $p(x) \leq Mq(x)$ ，然后根据一定的概率接受或拒绝从 $q(x)$ 中抽取的候选样本。重要性采样需要找到一个支撑集包含 $p(x)$ 的建议分布 $q(x)$ ，然后根据权重函数 $w(x) \propto p(x)/q(x)$ 来修正从 $q(x)$ 中抽取的样本的贡献。这两种方法都有各自的局限性，例如拒绝采样的接受率可能很低，重要性采样的方差可能很大。
- **MCMC算法**：MCMC方法是一种更复杂的采样策略，它利用马尔可夫链机制来探索状态空间 X ，并使得样本 $x(i)$ 模拟从目标分布 $p(x)$ 中抽取。这种机制需要构造一个转移核 K ，使得 $p(x)$ 是它的不变分布。一种常用的方法是满足细致平衡条件 $p(x)K(y|x) = p(y)K(x|y)$ 。此外，还需要保证马尔可夫链是不可约的和非周期的，以确保收敛性。

在此基础上，本篇文章进一步讨论了：

- **MCMC-EM算法**：一种用于最大似然和最大后验估计的标准算法，利用MCMC采样来处理E步中的期望计算，当期望是一个指数级别的求和或者一个难以求解的积分时。
- **辅助变量采样器**：一种通过引入辅助变量 u 来增加采样效率的方法，例如混合蒙特卡罗（HMC）和切片采样（slice sampling）。
- **可逆跳跃MCMC**：一种用于模型选择的方法，可以在不同维度的模型空间之间跳跃，通过定义合适的扩展空间、确定性变换和接受率来保证可逆性和细致平衡条件。这篇文档的第二部分主要讨论了以下内容：
- **MCMC-EM算法**：一种用于最大似然和最大后验估计的标准算法，利用MCMC采样来处理E步中的期望计算，当期望是一个指数级别的求和或者一个难以求解的积分时。
- **辅助变量采样器**：一种通过引入辅助变量 u 来增加采样效率的方法，例如混合蒙特卡罗（HMC）和切片采样（slice sampling）。
- **可逆跳跃MCMC**：一种用于模型选择的方法，可以在不同维度的模型空间之间跳跃，通过定义合适的扩展空间、确定性变换和接受率来保证可逆性和细致平衡条件。

第三部分

- **蒙特卡罗方法**：介绍了蒙特卡罗方法的历史、原理、应用和变种，包括**Metropolis算法**、**Gibbs采样**、**重要性采样**、**模拟退火**、**切片采样**等。
- **贝叶斯推断**：介绍了如何利用蒙特卡罗方法进行贝叶斯推断，包括**EM算法**、**MCMC算法**、**Rao-Blackwell化**、**变分推断**、**粒子滤波**等。
- **应用领域**：介绍了蒙特卡罗方法在各个领域的应用，包括**图像处理**、**语音信号**、**神经网络**、**动态概率网络**、**数据挖掘**等。

其中Metropolis-Hastings 算法抽样为提供了机器学习及相关内容非常好的基础。不需要归一化极大地减少了算法的计算量，同时，不同抽样之间形成的抽样链是完全独立的，所以极大利好了抽样的并行计算，为大规模的机器学习研究提供了基础。

"A very powerful property of MCMC is that it is possible to combine several samplers into mixtures and cycles of the individual samplers (Tierney, 1994)."MCMC算法同时还强大在于，它可以将几个不同的独立采样器结合起来，混合参与循环。

例如如果采样器1、2存在独立分布 $p(\cdot)$ ，那么其混合也是独立分布 $p(\cdot)$ ，得到的新分布 $K = \mu K_1 + (1 - \mu) K_2$

这样，对于一个高维度的多向量采样优化问题，我们可以将其拆解成几个低维度向量的MCMC采样问题，这样对于一个难以收敛的高维问题可以分解成几个易于收敛的低维度问题。同时可以在全局最优的基础上进一步探索局部最优，更好实现机器学习的相关优化。