

## 第二章 蒙特卡洛方法

计算机模拟采用的方法来看，它大致可以分为两种类型：

- (1) 随机模拟方法或统计试验方法，又称蒙特卡洛(Monte Carlo)方法。它是通过不断产生随机数序列来模拟过程。自然界中有的过程本身就是随机的过程，物理现象中如粒子的衰变过程、粒子在介质中的输运过程...等。当然蒙特卡洛方法也可以借助概率模型来解决不直接具有随机性的确定性问题。
- (2) 确定性模拟方法。它是通过数值求解一个个的粒子运动方程来模拟整个系统的行为。在统计物理中称为分子动力学(Molecular Dynamics)方法。关于分子动力学方法我们将在第六章中介绍。此外，近年来还发展了神经网络方法和原胞自动机方法。

从蒙特卡洛模拟的应用来看，该类型的应用可以分为三种形式：

- (1) 直接蒙特卡洛模拟。它采用随机数序列来模拟复杂随机过程的效应。
- (2) 蒙特卡洛积分。这是利用随机数序列计算积分的方法。积分维数越高，该方法的积分效率就越高。
- (3) Metropolis 蒙特卡洛模拟。这种模拟是以所谓“马尔科夫”(Markov)链的形式产生系统的分布序列。该方法可以使我们能够研究经典和量子多粒子系统的问题。

## 2.1 蒙特卡洛方法的基础知识

### 一、基本思想

对求解问题本身就具有概率和统计性的情况，例如中子在介质中的传播，核衰变过程等，我们可以使用**直接蒙特卡洛模拟方法**。该方法是按照实际问题所遵循的概率统计规律，用电子计算机进行直接的抽样试验，然后计算其统计参数。直接蒙特卡洛模拟法最充分体现蒙特卡洛方法无可比拟的特殊性和优越性，因而在物理学的各种各样问题中得到广泛的应用。该方法也就是通常所说的“计算机实验”。

蒙特卡洛方法也可以人为地构造出一个合适的概率模型，依照该模型进行大量的统计实验，使它的某些统计参量正好是待求问题的解。这也就是所谓的**间接蒙特卡洛方法**。下面我们举两个最简单的例子来说明间接蒙特卡洛方法应用的内涵。

### 巴夫昂(Buffon)投针实验。

该试验方案是：在平滑桌面上划一组相距为  $s$  的平行线，向此桌面随意地投掷长度  $l = s$  的细针，那末从针与平行线相交的概率就可以得到  $\pi$  的数值。

数学统计理论的简单地计算：

设针与平行线的垂直方向的夹角为  $\alpha$ ，那么针在与平行线垂直的方向上投影的长度为  $l \cdot |\cos \alpha|$ 。对于确定的  $\alpha$  夹角，细针与平行线相交的概率为投影长度与平行线间距之比，即  $\frac{l \cdot |\cos \alpha|}{s} = |\cos \alpha|$ 。由于  $\alpha$  是在  $[0, \pi]$  区间均匀分布的，所以  $|\cos \alpha|$  的平均值为

$$\frac{1}{\pi} \int_0^{\pi} |\cos \alpha| d\alpha = \frac{2}{\pi} .$$

假如在  $N$  次投针中，有  $M$  次和平行线相交。当  $N$  充分大时，相交的频数  $M/N$  就近似为细针与平行线相交的概率。因此，我们得到

$$\pi \approx \frac{2N}{M} .$$

然而，上述投针法得到试验结果的效率和精度都很差。

经过  $n$  次投针后得到  $\pi$  值的精度。

设  $p = 2/\pi$ ，则针与平行线相交的次数应满足二项式分布，其期望值为  $np$ ，方差应为  $np(1-p)$ 。因而  $2/\pi$  值的方差为  $p(1-p)/n$ ，标准误差为  $\sqrt{\frac{p(1-p)}{n}}$ 。将  $p = 2/\pi$  的标准误差改写为  $\pi$  的标准误差  $2.37/\sqrt{n}$ 。这意味着试验所得的  $\pi$  值的不确定性的范围如下：

对 100 次投针为，0.2374

对 10,000 次投针为，0.0237

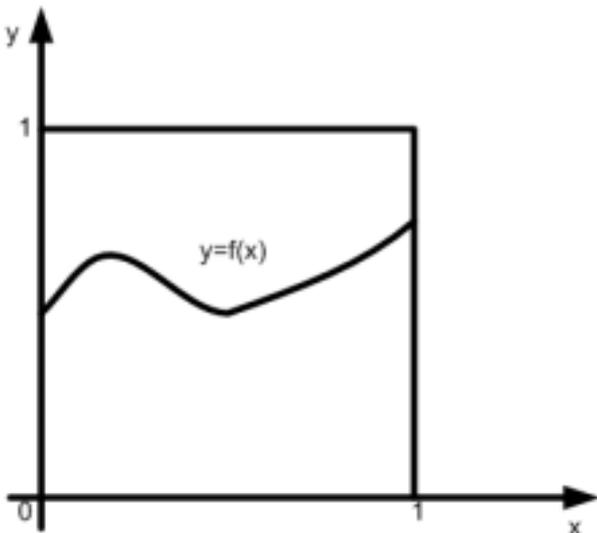
对 1,000,000 次投针为，0.0024。

显然用这种方法比用其它方法计算  $\pi$  值所引起的不确定范围要大得多。

例：定积分计算

$$I = \int_0^1 f(x) dx \quad . \quad 0 \leq f(x) \leq 1$$

这时我们可以随机地向正方形内投点，最后统计落在曲线下的点数  $M$ ，当总的掷点数  $N$  充分大时， $M/N$  就近似等于积分值  $I$ 。



### 蒙特卡洛方法的基本思想：

当问题可以抽象为某个确定的数学问题时，应当首先建立一个恰当的概率模型，即确定某个随机事件 A 或随机变量 X，使得待求的解等于随机事件出现的概率或随机变量的数学期望值。然后进行模拟实验，即重复多次地模拟随机事件 A 或随机变量 X。最后对随机实验结果进行统计平均，求出 A 出现的频数或 X 的平均值作为问题的近似解。这种方法也叫做**间接蒙特卡洛模拟**。

## 二、随机变量和随机变量的分布

$$g(u)du = P[u < u' < u + du].$$

$g(u)$  称为  $u$  的概率分布密度函数，它表示随机变量  $u'$  取  $u$  到  $u + du$  之间值的概率。

$g(u)$  分布函数定义为：

$$G(u) = \int_{-\infty}^u g(x)dx .$$

则  $g(u) = dG(u) / du$  .

注意： $G(u)$  是一个在  $[0,1]$  区间取值的单调递增函数。通常  $g(u)$  是归一化的分布密度函数，因而该函数对所有的  $u$  值范围的积分值应当为 1。

## 三、随机变量的独立性

假如我们考虑两个随机变量  $u'$  和  $v'$  的分布，则必须引进这两个变量的联合分布密度函数  $h(u, v)$ ，此时带来的数学问题就更为复杂。

在  $h(u, v) = p(u) \cdot q(v)$  这种特殊情况下， $u'$  和  $v'$  是彼此独立的随机变量。

对于两个以上的变量来说，随机变量独立性的概念就更复杂了。

如： $r$  和  $s$  是两个均匀分布在  $[0,1]$  区间的相互独立的随机变量，由此我们可以构造三个新的变量

$$\begin{aligned}x &= r, \\y &= s, \\z &= (r + s) \bmod 1.\end{aligned}$$

此时  $x, y, z$  也都是均匀分布在区间  $[0,1]$  的随机变量，并且所有的  $(x, y), (y, z)$  和  $(x, z)$  组合都是独立的(括号内任一个变量值的选取并不对括号中另一个变量的取值有影响)。但是  $(x, y, z)$  中，任意两个变量的值可以确定出第三个变量的值。此时它们之间存在明显的相关性。

#### 四、期望值、方差和协方差

一个函数  $f(u')$  的**数学期望值**是定义为该函数的平均值

$$E\{f\} = \int f(u) dG(u) = \int f(u)g(u)du .$$

$$E\{f\} = \frac{1}{b-a} \int_a^b f(u)du .$$

类似地，可以定义变量  $u'$  的期望值为  $u$  的平均值

$$E\{u'\} = \int udG(u) = \int ug(u)du .$$

一个函数或变量的**方差**:

$$V\{f\} = E\{(f - E\{f\})^2\} = \int [f - E\{f\}]^2 dG .$$

方差的平方根叫做**标准误差**。由于标准误差与其真值有相同的量纲，因而它比方差更具有物理意义。

如果将求期望值和求方差的运算作为算符，我们可以证明出这些算符作用在随机变量的线性组合式上的一些简单规则。假如  $x$  和  $y$  是随机变量， $c$  是一个常数，则

$$E\{cx + y\} = cE\{x\} + E\{y\} .$$

$$V\{cx + y\} = c^2V\{x\} + V\{y\} + 2cE\{(y - E\{y\})(x - E\{x\})\} .$$

期望值算符是一个线性算符，而方差算符是非线性算符。

#### 五、大数法则和中心极限定理

概率论中的大数法则和中心极限定理是蒙特卡洛方法的基础。

**大数法则反映了大量随机数之和的性质。**

如果函数  $h$  在  $[a, b]$  区间，以均匀的概率分布密度随机地取  $n$  个数  $u_i$ ，对每个  $u_i$  计算出函数值  $h(u_i)$ 。大数法则告诉我们这些函数值之和除以  $n$  所得的值将收敛于函数  $h$  的期望值，即

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n h(u_i) \equiv \lim_{n \rightarrow \infty} I_n = \frac{1}{b-a} \int_a^b h(u) du \equiv I.$$

大数法则保证了在抽取足够多的随机样本后，计算得到的积分的蒙特卡洛估计值将收敛于该积分的正确结果。若要对收敛的程度进行研究，并做出各种误差估计，则要用到中心极限定理。

**中心极限定理**告诉我们：在有足够大，但又有限的抽样数  $n$  的情况下，蒙特卡洛估计值是如何分布的。

该定理指出：无论随机变量的分布如何，它的若干个独立随机变量抽样值之和总是满足正则分布(即高斯分布)。例如我们有一个随机变量  $\eta$ ，它满足分布密度函数  $f(x)$ 。如果我们将  $n$  个满足分布密度函数  $f(x)$  的独立随机数相加： $R_n = \eta_1 + \eta_2 + \dots + \eta_n$ ，则  $R_n$  满足高斯分布。高斯分布可以由给定的期望值  $\mu$  和方差  $\sigma^2$  完全确定下来，通常用  $N(\mu, \sigma^2)$  来表示

$$N(\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-(x-\mu)^2 / 2\sigma^2\right].$$

中心极限定理可以给出蒙特卡洛估计值的偏差。如果上面公式右边积分的期望值为  $I$ ，公式左边用  $n$  次抽样的蒙特卡洛估计值为  $I_n$ ，标准误差为  $\sigma$ ，则当  $n$  充分大时，对任意的  $\lambda(\lambda > 0)$ ，有

$$\lim_{n \rightarrow \infty} \text{Pr ob} \left\{ -\lambda \frac{\sigma(f)}{\sqrt{n}} \leq I_n - I < \lambda \frac{\sigma(f)}{\sqrt{n}} \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} e^{-t^2/2} dt = 1 - \alpha.$$

这说明：该积分的期望值与蒙特卡洛估计值之差在范围

$$|I_n - I| < \lambda \frac{\sigma(f)}{\sqrt{n}}$$

内的概率为  $1 - \alpha$ ， $\alpha$  称为显著水平， $1 - \alpha$  称为置信水平。 $\sigma$  为蒙特卡洛估计值的标准误差， $\sigma^2 = V\{f\}/n$ 。 $\alpha$  与  $\lambda$  的关系可以有上面积分公式求得。也有专门的数学用表可查。例如取置信水平  $1 - \alpha = 99\%$ ，可以查得  $\lambda = 3$ 。这可以解释为：不等式  $|I_n - I| < 3 \frac{\sigma}{\sqrt{n}}$

成立的概率为99%。同样 $1-\alpha=95\%$ 时， $\lambda=2$ 。

从上面的分析看到，蒙特卡洛方法的误差与 $\sigma^2$ 和 $n$ 有关。为了减小误差，就应当选取最优的随机变量，使其方差最小。对同一个问题，往往会有多个可供选择的随机变量，这时就应当择优而用之。在方差固定时，增加模拟次数可以有效地减小误差。如试验次数增加100倍，精度提高10倍。当然这样做就增加了计算的机时，提高了费用。所以在考虑蒙特卡洛方法的精确度时，不能只是简单地减少方差和增加模拟次数，还要同时兼顾计算费用，即机时耗费。通常以方差和费用的乘积作为衡量方法优劣的标准。