

Monte Carlo Simulation

田付阳

应用物理研究所
数理学院物理系
北京科技大学

May 14, 2018



- 1 Basic knowledge for MC
- 2 随机数和随机抽样
- 3 蒙特卡罗方法的应用

1 Basic knowledge for MC

2 随机数和随机抽样

3 蒙特卡罗方法的应用

计算机数值计算

在对物理体系的计算机模拟过程中，可计算得到体系中某时刻各个分子(个体)在坐标和动量组成的相空间中的轨迹，因而可得出体系在某一时刻的相空间位形；进而由各个时刻的相空间位形可得出体系在相空间随时间的演变。根据体系在相空间中获得位形变化的方法，可将计算机模拟方法分成确定性模拟方法和随机模拟方法。

一个设定概率模型下，不断产生随机数序列来模拟物理过程或求解确定性问题，物理体系在相空间中的位移转变，是通过马尔科夫(Markov)过程由一个概率性的随机演化引起的。马尔科夫过程是内禀动力学在概率论方面的对应物。另外二个应用，直接蒙特卡罗模拟(采用随机数序列来模拟复杂随机过程的效应)，蒙特卡罗积分(利用随机数序列计算积分)；

而确定性模拟方法是通过直接求解物理系统中每一个粒子所遵守的运动方程，获得相空间中的位形变化，由此来模拟整个系统的行为。

什么是蒙特卡罗方法(Monte Carlo, MC)

也称为随机抽样法(random sampling)、随机模拟(random simulation)或统计试验法(statistic testing)

由von Neumann等把计算机随机模拟方法定名为Monte Carlo方法，明显带有随机性。一种利用随机统计规律进行计算和模拟的方法，可用于数值计算和数字仿真。

- 数值计算方面，可用于多重积分、代数方程求解、矩阵求逆以及用于微分方程求解，包括常微分方程、偏微分方程、本征方程和非奇次线性积分方程和非线性方程；
- 数字仿真方面，常用于核系统临界条件模拟、反应堆模拟以及实验核物理、高能物理、统计物理、真空、地震、生物物理和信息物理等领域。

Many-body problems often involve the calculation of integrals of very high dimension which cannot be treated by standard methods. For the calculation of thermodynamical averages, Monte Carlo methods are very useful which sample the integration volume at randomly chosen points.

Three examples

- 蒲丰(Buffon)投针问题:

$$p = \frac{1}{0.5a\pi} \int_0^\pi \frac{l}{2} \sin \phi d\phi = \frac{2l}{\pi a}$$

- 射击游戏:

$$\bar{g} = \int_0^\infty g(r)f(r)dr \quad \tilde{g} = \frac{1}{N} \sum_{i=1}^N g_i$$

- 计算积分:

$$I = \int_0^1 f(x)dx, f(x) \leq 1$$

求解事件出现的概率, 或者某个随机变量的期望值, 可通过某种“试验”方法, 得到这个事件出现的频率, 或者这个随机变量的平均值, 并以此平均值作为问题的解。

随机变量及其分布函数

事件发生的可能性大小用概率 p 表示, 则必然事件的发生概率 $p = 1$, 不可能事件的概率为 $p = 0$; 随机事件发生的概率为 $0 \leq p \leq 1$ 。描述随机事件 A 发生的概率用 $p(A)$ 表示, 显然, $0 \leq p(A) \leq 1$ 。经常碰到的随机变量有两类: 一类是离散型随机变量, 这种随机变量只能取有限个数值, 可以一一列举出来; 另一类是连续型随机变量, 这种随机变量的可能值是连续的分布在某个区间。

- 连续型分布: 对连续型随机变量 ξ , 其取值在区间 $[x, x + \Delta x]$ 内的概率表示为 $p(x \leq \xi < x + \Delta x)$, 如果极限

$$\lim_{\Delta x \rightarrow 0} \frac{p(x < \xi < x + \Delta x)}{\Delta x} = f(x)$$

存在, 则函数 $f(x)$ 表示随机变量 ξ 取值 x 的概率密度, 把 $f(x)$ 称为随机变量 ξ 的概率分布密度, 简称分布密度或密度函数。于是, 随机变量 ξ 落在 $[a, b]$ 内的概率 $p(a \leq \xi < b)$ 可写为

$$p(a \leq \xi < b) = \int_a^b f(x)dx$$

此式只有当 $f(x)$ 可积时才有意义。由此可以定义随机变量的分布函数来描写随机变量的概率分布规律。分布函数 $F(x)$ 定义为

$$F(x) = p(\xi \leq x) = \int_{-\infty}^x f(x)dx$$

i.e. 分布函数 $F(x)$ 在 x 处的值等于随机变量 ξ 取值小于或等于 x 这样一个随机事件的概率。显然，

有 $F(x = -\infty) = 0$, $F(x \leq +\infty) = 1$ 。所以分布函数 $f(x)$ 和分布函数 $F(x)$ 满足如下性质：

•

$$f(x) \geq 0$$

•

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

- 对于任意实数 x_1, x_2 ($x_1 \leq x_2$)

$$p(x_1 < x \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)dx$$

- 若 $f(x)$ 在 x 点连续, 则有 $F'(x) = f(x)$.
为了描述随机变量的统计特征需引入数学期望和方差的概念。数学期望定义为

$$E\xi = \int_{-\infty}^{+\infty} xf(x)dx$$

方差定义为

$$D\xi = \int_{-\infty}^{+\infty} (x - E\xi)^2 f(x)dx$$

2. 离散型分布

对于离散型随机变量 ξ , 可用分布列

$$\begin{cases} x_1, x_2, \cdots, x_n \\ p_1, p_2, \cdots, p_n \end{cases}$$

来描写, 它表示 ξ 取值 x_i 的概率为 $p_i (i = 1, 2, \cdots, n)$, i.e.

$$p_i = p(\xi = x_i)$$

分布列描写了离散型随机变量的概率分布。对于离散型随机变量 ξ 的分布函数 $F(x)$ 定义为阶梯函数

$$F(x) = p(\xi \leq x) = \sum_{x_i \leq x} p_i$$

其数学期望和方差被定义为

$$E\xi = \sum_{i=1}^n x_i p_i, \quad D\xi = \sum_{i=1}^n p_i (x_i - E\xi)^2$$

3. 概率密度分布

对于一组 N 个实验观测数据，相应于某一个随机变量的一个样本，可以用直方图形象地表示样本中数据分布的规律性，i.e. 概率分布函数。

先将随机变量的取值范围划分为若干个区间，将落入每一区间的数据个数 m (称为频数)与随机变量的取值区间之间的关系画成阶跃曲线，即构成直方图。数 m/N 是观测数据落入这个小区间的频率。直方图的横坐标为随机变量的取值范围，比例坐标为频数。

- 均匀密度分布函数 在区间 $[a, b]$ 均匀密度分布定义为

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & x < a, x > b \end{cases}$$

其中重要的特殊情况是 $[0,1]$ 均匀密度分布

$$f(x) = \begin{cases} 0 & x < 0, x > 1 \\ 1 & 0 \leq x \leq 1 \end{cases}$$

- 正态分布：具有平均值 μ 和标准偏差 σ 的正态分布(normal density)为

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

特别重要的是对应 $\mu = 0, \sigma = 1$ 的分布

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

- 指数分布：具有参数 λ 的指数分布

$$f(x) = \lambda e^{-\lambda x}, 0 \leq x < \infty, \lambda > 0$$

具有平均值 λ 和方差 λ^2 。

大数定理和中心极限定理

- 大数定理：设 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 为一随机变量序列，独立同分布，数学期望值 $E\xi = a$ 存在，则对任意 $\varepsilon > 0$ ，有关系

$$\lim_{n \rightarrow \infty} p\{|\bar{\xi} - a| < \varepsilon\} = 1$$

$$\bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i$$

为算术平均。此定理指出，当 $n \rightarrow \infty$ 时，算术平均收敛到数学期望(或统计平均) a 。也就是说当 n 很大时，可以用算术平均代替数学期望。其产生的误差可用中心极限定理确定。

- 中心极限定理：设 $\xi_1, \xi_2, \dots, \xi_n, \dots$ 为一随机变量序列，独立同分布，数学期望值 $E\xi = a$ ，方差 $D\xi = \sigma^2$ ，则当 $n \rightarrow \infty$ 时，

$$p\left(\frac{\bar{\xi} - a}{\sigma/\sqrt{n}} < X_a\right) \rightarrow \frac{2}{\sqrt{2\pi}} \int_0^{X_a} e^{-x^2/2} dx$$

中心极限定理表明, 当 n 很大时, 用算术平均近似数学期望的误差为

$$|\bar{\xi} - a| < \frac{\sigma X_a}{\sqrt{n}} = \varepsilon$$

的概率为 $1 - \alpha$, 所以误差 ε 有时也称为概率误差。 α 定义为

$$\alpha = 1 - \frac{2}{\sqrt{2\pi}} \int_0^{X_a} e^{-x^2/2} dx$$

α 称为置信度, $1 - \alpha$ 为置信水平。 α 和 X_α 的关系可以计算得到

α	0.5	0.05	0.02	0.01
X_α	0.6745	1.96	2.3863	2.5758

当要求置信度 $\alpha = 0.05$ 时, 误差估计为

$$\left| \frac{1}{n} \sum_{i=1}^n \xi_i - a \right| < \varepsilon = 1.96 \frac{\sigma}{\sqrt{n}}$$

的概率是95%.

1 Basic knowledge for MC

2 随机数和随机抽样

3 蒙特卡罗方法的应用

均匀分布随机数的产生

对一个随机过程，需要产生满足这个随机过程概率分布的随机变量。最简单和最基本的随机变量是 $[0,1]$ 区间上均匀分布的随机变量，这些随机变量的抽样值称为随机数。真正的随机数除统计规律外无任何其他规律可循。

在科学计算中通常按照某种算法给出随机数，称为伪随机数或称为赝随机数。其两个主要特性为：

- 具有一定的周期 n ，通常 n 足够大，以使其在整个使用过程中不表现其周期性；
- 具有一定的统计性，对均匀分布的随机数，既要求随机数产生的随机性，又要求产生的随机数分布的均匀性。

乘同余法公式产生随机数

$$x_n = cx_{n-1}(\text{mod}N) \quad \text{or} \quad x_n = cx_{n-1} - N \text{int}\left(\frac{cx_{n-1}}{N}\right)$$

其中 c , N 为给定的常数。给定 x_0 后，就可以用上式依次给出 x_1, x_2, \dots 一系列随机数。其一般原则为

- 关于 N 的取值一般取 $N = 2^{m-1}$, 其中 m 为计算机中二进制数的字长, $N - 1$ 则为计算所能表示的最大整数。
- 关于 c 的取值, 一般取 $c = 8M \pm 3$, 其中 M 为任一正整数, $c \sim N^{1/2}$ 。
- 关于 x_0 的取值, 一般取 x_0 为奇数。

在matlab中, 均匀分布 $U(a, b)$:

产生 $m \cdot n$ 阶 $[a, b]$ 均匀分布 $U(a, b)$ 的随机数矩阵: `unifrnd(a, b, m, n)`

产生一个 $[a, b]$ 均匀分布的随机数: `unifrnd(a, b)`

0-1分布 $U(0, 1)$:

产生 $m \cdot n$ 阶 $[0, 1]$ 均匀分布的随机数矩阵: `rand(m, n)`

产生一个 $[0, 1]$ 均匀分布的随机数: `rand`

正态分布的随机数: `normrnd(μ , σ , m, n)`, μ 为均值, 方差为 σ

随机性统计检验

好的随机数发生器或一个好的随机数生成程序必须满足两个条件：

- ① 所生成的随机数序列应当具有足够长的周期；
- ② 生成的随机数序列应具有真正随机序列所具有的统计性质。

其周期的长短比较容易测试和判断。通常对统计性质的检验方法是采用频数分布检验：对于一个均匀分布的随机数发生器，设所产生随机数序列的值域为 $[0,1]$ ，则所产生的随机数字应与 $0 \sim 1$ 均匀的频数分布相一致。

为了检验频数分布情况，可按画统计直方图的方法，将整个值域分成 M 个宽度相等的子区间，设 x_i 是第 i 区间内出现的随机数的个数，即第 i 个子区间的频数，则所有子区间中随机数个数的平均频数

$$\bar{x} = \frac{1}{M} \sum_{i=1}^M x_i$$

第 i 子区间频数的偏差

$$\varepsilon_i = x_i - \bar{x}$$

和频数的方均根偏差

$$\sigma = \sqrt{\frac{1}{M} \sum_{i=1}^M \varepsilon_i^2}$$

如果所产生的 N 个随机数均匀分布于整个值域, 则 $x_i = \bar{x} = N/M$. 且在任一子区间内出现 x 个数字的概率服从高斯分布规律, *i.e.* $e^{-(x-\bar{x})^2/(2\sigma^2)}$, 其中 $\sigma = \sqrt{\bar{x}}$ 为标准偏差。可见, 一个均匀分布的随机数序列, 其最可几频数应为 \bar{x} , 而其频数 x 在 $x \pm \sigma$ 范围内的概率应为0.68, 其频数的方均根偏差 σ_{rms} 应接近于标准偏差 σ 。所以检查均匀分布随机数分布的均匀程度就是判断各频数是否接近于平均频数和频数的方均偏差是否接近于标准偏差。

一、连续型分布的抽样方法

- 直接抽样方法:

已知随机变量 η 具有分布密度 $f(x)$.如何随机抽样出满足分布密度 $f(x)$ 的随机变量 η , 根据分布函数 $F(x)$ 的定义, $F(\eta)$ 表示随机变量取值小于和等于 η 的概率

$$F(\eta) = \int_{-\infty}^{\eta} f(x)dx$$

$F(\eta)$ 为单调增函数, 而且存在反函数。令均匀分布的随机数 ξ 等于分布函数, *i.e.* $\xi = F(\eta)$, 则得到 $\eta = F^{-1}(\xi)$ 就是具有分布密度 $f(x)$ 的随机变量的抽样。由分布函数定义

$$p(\eta \leq x) = \int_{-\infty}^x f(x)dx = F(x)$$

如果 $\xi = F(\eta)$ 是 $[0, 1]$ 均匀分布的随机变量

$$\begin{aligned}
 p\{\eta \leq x\} &= p\{F^{-1}(\xi) \leq x\} = p\{\xi \leq F(x)\} = \int_{-\infty}^{F(x)} f_u(t) dt \\
 &= \int_{-\infty}^0 0 \cdot dx + \int_0^{F(x)} 1 \cdot dx = F(x)
 \end{aligned}$$

其中用到了 $[0, 1]$ 均匀分布的随机变量的分布密度函数

$$f_u(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{other} \end{cases}$$

可以利用 $[0, 1]$ 均匀分布的随机数产生具有给定分布密度 $f(x)$ 的随机变量 η 序列。由此得抽样步骤为：

1. 给定分布密度 $f(x)$; 2. 计算其分布函数 $F(x)$; 3. 产生随机数 ξ ; 4. 计算 $F^{-1}(\xi)$, 令 $\eta = F^{-1}(\xi)$; 5. 重复3, 4.

Two examples

- ① 通常用指数密度分布描述电子元件的稳定时间，系统的可靠性和粒子游动的自由程等。求按指数密度分布的随机抽样。

$$f(x) = \begin{cases} \lambda^{-1}e^{-x/\lambda} & x > 0, \lambda > 0 \\ 0, & \text{other} \end{cases}$$

- ② 求均匀密度函数

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{other} \end{cases}$$

在区间 $[a, b]$ 上均匀分布的随机数。

● 变换抽样方法

将复杂的分布抽样变换为已知的或比较简单的抽样。for instance, 如果分布密度为 $f(x)$ 的随机变量 x 的抽样比较复杂, 而分布密度为 $g(y)$ 的随机变量 y 的抽样已知, 由概率守恒, 通过变换

$$f(x) = \left| \frac{dy}{dx} \right| g(y)$$

实际上, 这里如果取 $g(y)$ 是 $[0,1]$ 区间的均匀分布

$$g(y) = \begin{cases} 0, & y < 0, y > 1 \\ 1, & y \in [0, 1] \end{cases}$$

$$y(x) = \int_{-\infty}^x f(t) dt, \left| \frac{dy}{dx} \right| = f(x)$$

$x(y)$ 正是满足密度分布 $f(x)$.

同样对两个随机变量问题：设随机变量 x 和 y 的联合分布密度函数是 $f(x, y)$, 随机变量 u 和 v 的联合分布密度函数是 $g(u, v)$, 则变换规律为

$$f(x, y) = |J|g(u, v), J = \frac{\partial(u, v)}{\partial(x, y)} = \left(\frac{\partial u}{\partial x} \frac{\partial v}{\partial y} - \frac{\partial u}{\partial y} \frac{\partial v}{\partial x} \right)$$

例：给出下列正态分布的描述

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, -\infty < x < +\infty$$

● 离散型随机变量抽样法

设 η 是离散型随机变量，分别以概率 p_i 取值 $\eta_i (i = 1, 2, \dots; p_1 + p_2 + \dots = 1)$, 抽样步骤为

- ① 选取随机数 ξ ;
- ② 确定满足不等式 $\sum_{i=0}^{j-1} p_i \leq \xi < \sum_{i=0}^j p_i$ 的 j 值(这里约定 $p_0 = 0$);
- ③ 对应 j 的 η_j 就是所求的抽样值, 即 $x_p = \eta_j$;

$$p(x_p = \eta_j) = p\left(\sum_{i=0}^{j-1} p_i \leq \xi < \sum_{i=0}^j p_i\right) = \sum_{i=0}^j p_i - \sum_{i=0}^{j-1} p_i = p_j$$

当 $l_j = \sum_{i=0}^j p_i$, 则 $l_{j-1} \leq \xi < l_j \Rightarrow x_p = \eta_j$ 。

MC方法的基本步骤

一. 随机性问题：采用直接模拟方法：首先必须根据物理问题的规律，建立一个概率模型(随机向量或随机过程)，然后用计算机进行抽样试验，得出对应于物理问题的随机变量分布。

假定随机变量 $y = g(x_1, x_2, \dots, x_m)$ 是研究对象，它是 m 个互相独立的随机变量 (x_1, x_2, \dots, x_m) 的函数， (x_1, x_2, \dots, x_m) 概率分布密度分别为 $(f(x_1), f(x_2), \dots, f(x_m))$ ，因此蒙特卡罗模拟的基本方法是：根据概率分布密度

$$(f(x_{j1}), f(x_{j2}), \dots, f(x_{jm})), j = 1, 2, \dots, N$$

在计算机上用随机抽样的方法抽样产生 N 组随机变量 $(x_{j1}, x_{j2}, \dots, x_{jm})$ ，计算

$$y_j = g(x_{j1}, x_{j2}, \dots, x_{jm}), j = 1, 2, \dots, N$$

的值，用这样的样本分布来近似 y 的函数，由此可以计算出这些量的统计值。

对于随机性问题，蒙特卡罗方法的计算过程是用统计方法模拟实际的物理过程，它主要是在计算机上产生已知分布的随机变量样本，以代替昂贵的甚至难以实现的实验，又被视为用计算机来完成物理实验的一种方法。

二. 确定性问题：

- ① 首先建立一个有关这个确定性问题的概率统计模型，使所求的解是这个模型的概率分布或数学期望；
- ② 然后对这个模型作随机抽样，
- ③ 最后用其算术平均值作为所求解的近似值。

1 Basic knowledge for MC

2 随机数和随机抽样

3 蒙特卡罗方法的应用

- ① 方程求根
- ② 计算定积分
- ③ 求解拉普拉斯方程
- ④ 核链式反应的模拟
- ⑤ 中子贯穿概率问题

随机行走问题

- 设 X_k 是一个独立同分布的随机变量序列, 对于每一个正整数 n , 设 S_n 表示和 $X_1 + X_2 + \cdots + X_n$, 序列 S_n 称为随机行走。

典型例子: 醉汉行走问题, 醉汉开始从一根电线杆的位置出发, 其坐标向右为正, 向左为负, 假定醉汉的步长为 l , 他走的每一步的取向是随机的, 与前一步的方向无关。如果醉汉在每个时间间隔内向右行走一步的概率为 p , 而向左走一步的概率为 q . 记录向右走的步数, 向左走的步数, 总共走了 N 步, 那么醉汉在行走了 N 步以后, 离电线杆的距离为 x 的概率 $P_N(x)$ 。醉汉在走了 N 步后的位移和方差的平均值的计算公式为

$$\langle x_N \rangle = \sum_x x P_N(x)$$

$$\langle x_N^2 \rangle = \sum_x x^2 P_N(x)$$

$$\langle \Delta x_N^2 \rangle = \langle x_N^2 \rangle - \langle x_N \rangle^2$$

- ① 求连续型概率密度 $f(x) = x, 0 \leq x \leq 1$ 的随机抽样公式。
- ② 求按下列指数密度分布的抽样公式

$$\begin{aligned} f(x) &= \lambda e^{-\lambda x}, x > 0, \lambda > 0 \\ &= 0, x \leq 0 \end{aligned}$$

- ③ 设能量为 E 的 γ 光子与物质相互作用产生光电效应的相对截面 $\sigma_e = 0.20$, 产生康普顿散射的相对截面 $\sigma_s = 0.45$, 产生电子对效应的相对截面 $\sigma_p = 0.35$, 如果某一次产生的随机数 $\varepsilon = 0.3$ 确定就这一次相互作用的抽样类型? 说明判断原理。

- MC基础知识;

总结

- MC基础知识;
- 随机抽样;

总结

- MC基础知识;
- 随机抽样;
- 应用.

感谢关注!