

第三章 蒙特卡洛方法的若干应用

蒙特卡洛方法是利用随机变量的一个数值序列来得到特定问题的近似解的数值计算方法。

蒙特卡洛方法的应用可以大致分为两类：第一类是所求问题具有严格确定的数学形式，另一类是本身就是具有统计性质的问题。

3.1 蒙特卡洛方法在积分计算中的应用

一、一维定积分计算的平均值法（期望值估计法）

一维积分计算

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

在 x 的定义域 $[0, 1]$ 上均匀地随机取点，该均匀分布的随机变量记为 ξ 。我们定义一个随机变量 η_1 为

$$\eta_1 = f(\xi).$$

则显然有

$$E\{\eta_1\} = E\{f(\xi)\} = I.$$

η_1 的期望值等于积分值 I 。只要抽取足够多的随机点，即取随机点数 n 足够大时， $f(\xi)$ 的平均值

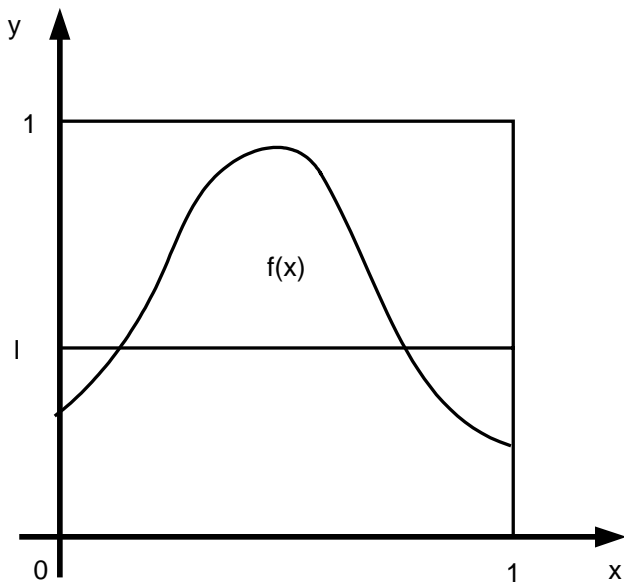
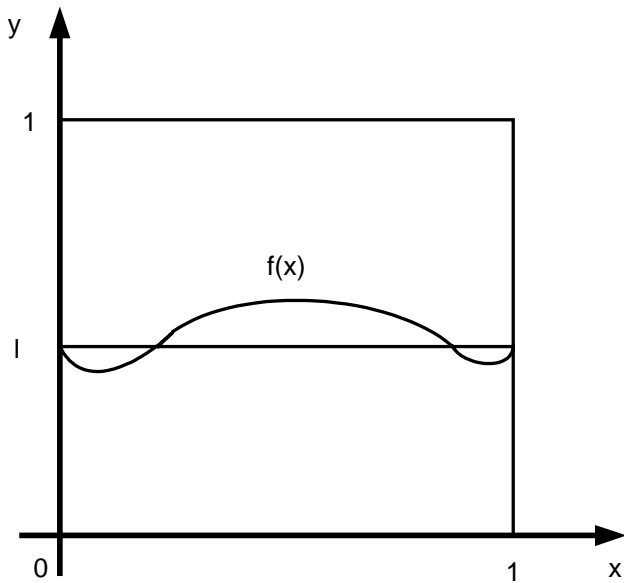
$$I_n = \frac{1}{n} \sum_{i=1}^n f(\xi_i)$$

就是积分 I 的一个无偏估计值。

η_1 的方差。

$$V\{\eta_1\} = \int_0^1 [f(x) - I]^2 dx .$$

显然 $V\{\eta_1\}$ 依赖于被积函数 $f(x)$ 在积分域上的方差。当 $f(x)$ 在 x 的定义域内变化平坦，即和 I 的差处处都较小时，方差也小；反之，则方差较大。



从这里可以看出：尽量减小被积函数在积分域上的方差，可以减小积分估计值的方差，加速收敛。推而广之来说，就是要减少模拟量在模拟范围内的方差。

根据这样的原则，当被积函数 $f(x)$ 在积分域内的方差较大时，可以采用各种抽样技巧。如采用重要抽样法，将 $f(x)$ 的方差吸收到 $g(x)$ 中去，这样模拟量——记录函数 $f^*(x) = f(x)/g(x)$ 在定义域内相当平坦，则我们将(3.1.1)式的计算变为

$$I = \int_0^1 f(x)dx = \int_0^1 \frac{f(x)}{g(x)} g(x)dx = \int_0^1 f^*(x)g(x)dx$$

若选取 η' 为服从分布密度函数 $g(x)$ 的函数 $f^*(x)$ 的抽样值。这里 $g(x)$ 称为偏倚分布密度函数。我们得到

$$I = E\{\eta'\}.$$

因此它的平均值

$$I_n = \frac{1}{n} \sum_{i=1}^n \eta'_i = \frac{1}{n} \sum_{i=1}^n f^*(x_i).$$

给出了 I 的一个无偏估计值。这时的方差为：

$$V\{\eta'\} = \int_0^1 [f^*(x) - I]^2 g(x)dx = \int_0^1 \left[\frac{f(x)}{g(x)} - I \right]^2 g(x)dx = \int_0^1 \frac{f^2(x)}{g(x)} dx - I^2.$$

在实际计算中，方差通过下式得到计算结果：

$$\sigma^2 = \left\langle \left(\frac{1}{n} \sum_{i=1}^n f^*(x_i) \right)^2 \right\rangle - \left\langle \left(\frac{1}{n} \sum_{i=1}^n f^*(x_i) \right) \right\rangle^2.$$

式中角型括号 $\langle \rangle$ 表示对括号内所有可能的 $[0, 1]$ 区间，按 $g(x)$ 分布的随机坐标数序列 $\{x_i\}$ 对应的数值求平均。方程右边第一项对 $\{f^{*2}(x_i)\}$ 求平均 $\left(\overline{f^{*2}}\right)$ ，第二项表示求 $\{f^*(x_i)\}$ 平均值的平方 $\left(\overline{f^*}^2\right)$ 。上

式可以经推导得到：

$$\sigma^2 = \frac{1}{n}(\overline{f^{*2}} - \bar{f}^{*2}) = \frac{V\{f^*\}}{n}.$$

由此我们看出其误差平方与 f^* 在 $[0, 1]$ 区间的方差成正比，并且 $\sigma \propto 1/\sqrt{n}$ 。这与中心极限定理所得到的结果一致。

二、一维定积分计算的掷点法

计算积分也可以这样做：

在单位正方形内均匀投点，每个点的坐标为 (x_i, y_i) ，共做 N 个投点。如果投点满足不等式 $y_i \leq f(x_i)$ ，即点落在 $f(x)$ 曲线下，则记录下投点次数（认为试验成功）；反之，则认为试验失败。

用蒙特卡洛的语言来讲，就是产生随机数 ξ_1, ξ_2 。如果 $\xi_1 \leq f(\xi_2)$ ，则认为试验成功；如果 $\xi_1 > f(\xi_2)$ ，则试验失败。若在 N 次试验中有 m 次成功，则比值 m/N 就给出 I 的一个无偏估计值：

$$I \approx \frac{m}{N}.$$

引入随机变量

$$\eta(\xi_1, \xi_2) = \begin{cases} 1, & \xi_1 \leq f(\xi_2) \\ 0, & \xi_1 > f(\xi_2) \end{cases}$$

则 $I = E\{\eta(\xi_1, \xi_2)\}$ 。

它在 N 次试验下的一个 I 的无偏估计值为

$$I_N = \frac{1}{N} \sum_{i=1}^N \eta(\xi_{2i-1}, \xi_{i2}) = \frac{m}{N}.$$

这是 I 的一个近似值，它的方差为

$$V\{\eta\} = E\{\eta^2\} - [E\{\eta\}]^2 = I - I^2.$$

容易证明掷点法的方差比平均值法的方差大

$$\begin{aligned} V\{\eta\} - V\{\eta_1\} &= I - I^2 - \int_0^1 [f(x) - I]^2 dx = I - I^2 - \int_0^1 f^2(x) dx + 2I \int_0^1 f(x) dx - I^2 \\ &= \int_0^1 f(x)[1 - f(x)] dx \geq 0. \end{aligned}$$

证明：

$$\int_0^1 \eta(x, \xi_2) dx = \int_0^{f(\xi_2)} \eta(x, \xi_2) dx + \int_{f(\xi_2)}^1 \eta(x, \xi_2) dx = f(\xi_2)$$

而在平均值法中 $I = E\{\eta_1\} = E\{f(\xi)\}$ ，恰恰用了 $\eta(\xi_1, \xi_2)$ 对 ξ_1 的期望值代替了 $\eta(\xi_1, \xi_2)$ 。

这里可以反应出减小方差，加快收敛的又一个原则。这就是要尽量使用理论分析得到的期望值来代替模拟估计值。这个原则也同样适用于所有的蒙特卡洛模拟过程。

实际上使用这个原则可以减小方差、加快收敛的原因是显然的。因为一切随机模拟量总会有误差的，如果以精确的理论值来代替 $\eta(\xi_1, \xi_2)$ ，就必然会减小方差。所以在一切模拟过程中，能用理论计算值的地方应当尽量使用。

以上我们介绍的这两个减小方差，加速收敛的原则，也正是重要抽样法、分层抽样法、对偶变量法、相关抽样法等的基本出发点。

三、多重定积分的计算

物理上的许多问题都会涉及多重定积分。例如，一个粒子衰变到 n 体末态的相空间积分，由于每个末态粒子都有动量和能量四个分量，考虑到每个粒子满足质能公式和所有粒子的总能、动

量守恒，则总的相空间积分重数为 $3n-4$ 。这样的物理问题往往都需要做数值积分。

前面讲的一维定积分计算的平均值法和掷点法都可以推而广之，应用于多重定积分的计算。

对于 s 维多重积分，我们也可以用前面讲述的“归一化”方法，使得积分变量 $x_i \in [0,1]$, $(i=1,\dots,s)$ ，被积函数在积分范围内满足 $0 \leq f(x_1, x_2, \dots, x_s) \leq 1$ 。然后再做积分

$$I = \int_0^1 \int_0^1 \dots \int_0^1 f(x_1, x_2, \dots, x_s) dx_1 dx_2 \dots dx_s \quad .$$

在实际蒙特卡洛积分计算中，在积分的超立方体内不同区域的积分贡献可能强烈地变化。如果我们在积分的超立方体内均匀抽样，积分的贡献可能主要来自少数仅仅只有几个蒙特卡洛投点的小区域，这就会导致很大的统计误差。

当在积分域内 $f(x_1, x_2, \dots, x_s)$ 的方差很大时，就会产生这个效应。为了减少这些对误差的贡献，我们将随机投点更多地投在 $|f(x_1, x_2, \dots, x_s)|$ 取值大的区间。这就是说，采用重要抽样的蒙特卡洛积分方法。

具体操作步骤是：

(1) 选一个抽样比较简单的概率分布密度函数 $g(x_1, x_2, \dots, x_s)$ ，并定义

$$f^*(x_1, x_2, \dots, x_s) = \begin{cases} \frac{f(x_1, x_2, \dots, x_s)}{g(x_1, x_2, \dots, x_s)}, & g(x_1, x_2, \dots, x_s) \neq 0 \\ 0, & g(x_1, x_2, \dots, x_s) = 0 \end{cases}$$

使得 $f^*(x_1, x_2, \dots, x_s)$ 在积分域内的方差较小，则、

$$I = E\{f^*(x_1, x_2, \dots, x_s)\} = \int_0^1 \int_0^1 \dots \int_0^1 f^*(x_1, x_2, \dots, x_s) g(x_1, x_2, \dots, x_s) dx_1 dx_2 \dots dx_s \quad .$$

按照偏倚密度函数 $g(x_1, x_2, \dots, x_s)$ 在 $0 \leq x_i \leq 1, (i=1, \dots, s)$ 空间中抽取 N 个子样

$(x_{i1}, x_{i2}, \dots, x_{is}), i=1, 2, \dots, N$, 则记录函数 $f^*(x_1, x_2, \dots, x_s)$ 的平均值为

$$I_N = \frac{1}{N} \sum_{i=1}^N f^*(x_{i1}, x_{i2}, \dots, x_{is}) .$$

它给出了 I 的一个无偏估计值, 并可以作为 I 的近似值。

如果在 s 维体积 Ω 内做多重积分 $I = \int_{\Omega} \dots \int f(x_1, x_2, \dots, x_s) dx_1 dx_2 \dots dx_s$ 时,

如果在积分域 Ω 内 $f(x_1, x_2, \dots, x_s)$ 的方差并不大, 为了简化抽样, 就

取

$$g(x_1, x_2, \dots, x_s) = \begin{cases} 1/\Omega, (x_1, x_2, \dots, x_s) \in \Omega \\ 0, \text{其它} \end{cases}$$

这时记录函数为

$$f^*(x_1, x_2, \dots, x_s) = \frac{f(x_1, x_2, \dots, x_s)}{g(x_1, x_2, \dots, x_s)} = \Omega f(x_1, x_2, \dots, x_s) .$$

在 s 维体积 Ω 内抽取随机样本 $(x_{i1}, x_{i2}, \dots, x_{is})$ 是容易的, 若抽得 N 个

样本之后,

$$I_N = \frac{\Omega}{N} \sum_{i=1}^N f(x_{i1}, x_{i2}, \dots, x_{is}) .$$

就给出了 I 的近似值。

从前面介绍的减小方差的第二个原则可以看出: 在采用蒙特卡洛方法计算多重积分时, 如果能够将其中的某几重积分解析地求出时, 应当尽量地使用解析方法。这样便能减小方差, 加速收敛。

为了使在积分的高维体积内的投点更加均匀, 我们可以将积分空间分成许多相同体积的子空间, 在每个子空间中都投以相同数目的随机点, 从而减少蒙特卡洛积分误差。这就是采用前面第 2.4

节中介绍的“分层抽样方法”。这种积分方法也叫做分层蒙特卡洛积分法。

蒙特卡洛方法用于计算定积分时的特点：

- (1) 蒙特卡洛方法计算定积分的收敛速度与积分的重数无关。
- (2) 蒙特卡洛方法求定积分的误差仅仅与方差 $v\{f\}$ 和子样容量 n 有关，而与子样中的元素所在的集合空间 Ω 的组成无关。
- (3) 被求定积分的维数变化，除了引起抽样及计算时间有变化外，对计算结果的精度没有影响。

优点：

- (1) 利用该方法处理多重积分问题时，维数越高，其优越性越明显。
- (2) 利用蒙特卡洛计算定积分问题时受积分域的限制较小。只要积分空间 Ω 可以用数学形式描述出其范围，不论它的形状如何复杂，我们都可以用该方法给出该积分的估计值。因而蒙特卡洛方法是解决复杂几何空间定积分的有效方法。