Numerical Analysis

Zhiping Li

LMAM and School of Mathematical Sciences
Peking University

随机问题

人类对自然现象和社会经济活动的数学描述, 即有确定性的方法, 也有随机性的方法.

许多问题因为其内在的随机性, 应用随机方法描述和求解是一种自然的选择. 例如, 统计物理、量子化学、材料科学、生命科学等自然科学领域中都有大量内蕴随机性的问题.

另外, 在经济预测、金融产品定价等社会科学领域也大量随机性的问题.

还有许多问题, 自身并不一定有内蕴的随机性, 但用随机方法描述和处理更加得力. 例如, 在维数很高的空间中做数值积分.

随机问题与 Monte Carlo 方法

Monte Carlo 方法是一种处理随机问题的随机方法. 许多随机系统可以通过“概率密度函数”刻画. 从某种意义上说, Monte Carlo 方法是对这类系统的一种直接的离散数值模拟方法.

例如, 在统计物理中, 已知系统中大量粒子的（以概率分布函数的形式给出的）运动分布规律, 希望从中提取由这些大量粒子共同表现出来的宏观统计物理量, 比如温度、比热等.

从大量粒子的某种性质的分布规律提取相应的宏观统计物理量, 本质上是在粒子所在的相空间求某随机变量的数学期望. 由于粒子数目庞大, 因此相空间的维数自然也水涨船高.

这时, Monte Carlo 方法即可以认为是对大量粒子系统的直接离散数值模拟, 也可以视为求高维积分的一种数值积分方法.

随机方法的例 — 数值积分的 Monte Carlo 方法

以一维积分为例来说明用 Monte Carlo 方法计算积分的思想. 考虑

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

我们有许多传统的数值积分方法, 例如复合梯形公式:

$$I(f) = \left[\frac{1}{2}f(x_0) + \sum_{i=1}^{N-1} f(x_i) + \frac{1}{2}f(x_N)\right]h + O(h^2),$$

其中 $h = 1/N$, $x_i = ih$, $i = 0, 1, \cdots, N$.

而用 Monte Carlo 方法求积分 $I(f)$, 则是将其视为某个在区间 $[0, 1]$ 上服从均匀分布的随机变量 X 的函数 $f(X)$ 的数学期望 $I(f) = Ef(X)$.

随机方法的例 — 数值积分的 Monte Carlo 方法

由弱大数定律, 对独立同分布的随机变量 X_i, $i = 1, 2, \cdots$, 如果 $E|X_i| < +\infty$, 记 $S_N = \sum_{i=1}^{N} X_i$, 则当 $N \to \infty$ 时, S_N/N 依概率收敛于 EX_1.

因此, 为了计算 $I(f) = Ef(X)$, 取相互独立且同服从区间 $[0, 1]$ 上均匀分布的随机变量 X_i, $i = 1, 2, \cdots, N$, 构造数值积分公式

$$I(f) \approx I_N(f) \triangleq \frac{1}{N} \sum_{1}^{N} f(X_i).$$

这就是数值积分的 Monte Carlo 方法.

数值积分的 Monte Carlo 方法的误差

由弱大数定律, $I_N(f)$ 依概率收敛于 $I(f)$. 而且, 作为随机变量 $I_N(f)$ 满足

$$EI_N(f) = E\left[\frac{1}{N}\sum_{i=1}^{N} f(X_i)\right] = \frac{1}{N}\sum_{i=1}^{N} I(f) = I(f).$$

数值积分的 Monte Carlo 方法的误差 $e_N = |I_N(f) - I(f)|$ 仍然是一个随机变量, 其平方的期望称为均方误差:

$$E|e_N|^2 = E\left[\frac{1}{N}\sum_{i=1}^{N}(f(X_i) - I(f))\right]^2$$

$$= \frac{1}{N^2}\sum_{i,j=1}^{N} E[(f(X_i) - I(f))(f(X_j) - I(f))]$$

$$= \frac{1}{N}E[f(X_1) - I(f)]^2 = \frac{1}{N}E[f(X_1) - Ef(X_1)]^2 = \frac{1}{N}Var(f(X)).$$

数值积分的 Monte Carlo 方法的误差

由 Schwartz 不等式有

$$E|e_N| \leq \sqrt{E|e_N|^2} = \sqrt{\frac{Var(f(X))}{N}}.$$

由此知, 如果 $f(X)$ 有有限的方差, 则数值积分的 Monte Carlo 方法在如上意义下有半阶 $O(N^{-1/2})$ 的收敛性.

半阶收敛性对低维空间的积分当然算不上什么好方法. 但是注意这种收敛性与空间维数无关, N 是计算函数值并将其求和的次数.

作为对比, 当空间维数是 d, 每个坐标轴方向的网格尺度均为 $h = 1/n$, 且采用复合梯形公式时, 需要计算的函数值及其求和次数是 $N = n^d$, 而收敛阶是 $O(n^{-2}) = O(N^{-2/d})$. 在分子动力学或统计物理中, $d = 6M$, M 为系统的粒子数. 对有成千上万甚至更多粒子的系统, Monte Carlo 方法恐怕是唯一合理的选择了.

数值积分的 Monte Carlo 方法的例

用 Monte Carlo 方法计算积分 $\int_0^{\frac{\pi}{2}} \sin x \, dx$, 有

$$\int_0^{\frac{\pi}{2}} \sin x \, dx \approx \frac{1}{N} \sum_{i=1}^{N} \frac{\pi}{2} \sin\left(\frac{\pi}{2} X_i\right),$$

其中取 X_i, $i = 1, \cdots, N$, 为 $i.i.d.\ \mathcal{U}[0,1]$ 随机变量.

为了减小方差对计算结果的影响, 更好地在图形上显示算法的收敛性, 通常会指定一正整数 m, 对每个 N, 令 $e_N = \frac{1}{m}\sum_{i=1}^{m} e_N^i$, 即将 m 个样本数为 N 的 Monte Carlo 方法的数值积分结果的误差 e_N^i, $i = 1, \cdots, m$, 的平均值作为样本数为 N 的数值积分结果的误差.

图 7.1 给出的是在 $\ln N \sim \ln e_N$ 坐标下, $m = 100$ 时的误差图, 它清楚地显示了算法的半阶收敛速度.

一般概率型积分数值积分的 Monte Carlo 方法

对一般概率型积分 $\int_\Omega f(x)p(x)\,dx$, 其中 $p(x)$ 为 Ω 上的概率密度函数, 满足 $p(x) \geq 0$ 和 $\int_\Omega p(x)\,dx = 1$. 则 Monte Carlo 方法的数值积分公式为

$$\int_\Omega f(x)p(x)\,dx \approx \frac{1}{N}\sum_{i=1}^{N} f(X_i),$$

其中 $X_i, i = 1, \cdots, N$, 是以 $p(x)$ 为其分布密度函数的 $i.i.d.$ 的随机变量.

随机数与伪随机数

要在计算机上实现 Monte Carlo 方法的首要任务是以较高的效率产生服从指定分布的随机数. 在计算机上产生理论意义上的随机数是难以实现的, 况且如果真的实现了, 那么数值实验的“可重复性”就无法实现了.

在实际计算中, 通常是在计算机上产生所谓的“伪随机数序列”(pseudo random number sequence), 即采用某种确定性算法产生貌似“随机”的数列, 这种数列的性态可以通过相应的统计检验, 从而可以作为真正随机数的一种替代, 得到与理论预期一致的数值计算结果.

$\mathcal{U}[0,1]$ 伪随机数的产生 — 线性同余法

目前已经有一些可靠的 $\mathcal{U}[0,1]$ 伪随机数发生器. 衡量伪随机数发生器好坏的一个重要指标所谓最大循环长度.

比较简单的 $\mathcal{U}[0,1]$ 伪随机数发生器是由线性同余法

$$X_{n+1} = aX_n + b \pmod{m},$$

其中 a, b, m 是事先取定的自然数, 给出的.

定理: 对线性同余法, 如果 a, b, m 满足

1. b 与 m 互素;
2. $(a-1)$ 是m 的任一奇数因子的倍数;
3. 若 4 是 m 的因子, 则 4 也必定是 $(a-1)$ 的因子,

则该伪随机数发生器的最大循环长度为 m, 即满长度.

例如, 可取 $m = 2^k$, $a = 4c+1$, b 为奇数.

$\mathcal{U}[0,1]$ 伪随机数的产生 — 神奇的16807

线性同余法的一个特殊例子是由 Lewis, Goodman 和 Miller 于 1969 年提出的以下伪随机数发生器：

$$X_{n+1} = aX_n \pmod{m},$$

其中 $a = 7^5 = 16807$, $m = 2^{31} - 1 = 2147483647$. 该伪随机数发生器最大循环长度可达$2.1 \times 10^9$. Shrage 给出了在计算机上高效实现上述乘法同余的算法.

更神奇的 Bays-Durham **洗牌算法** — L'Ecuyer 给出了基于该算法的最大循环长度达到约 2.3×10^{18} 的伪随机数发生器.

可靠的(伪)随机数发生器是可靠的随机模拟的基本保障, 所以, 在做实际随机模拟时应选用成熟的随机数发生器程序包.

一般分布的随机变量的产生

有了 $\mathcal{U}[0,1]$ 随机数发生器, 理论上说, 我们就可以给出一般分布的随机数发生器.

命题： 设随机变量 Y 的分布函数为 $F(y)$, 即

$$P\{Y \leq y\} = F(y).$$

如果随机变量 $X \sim \mathcal{U}[0,1]$, 则随机变量 $Y = F^{-1}(X)$ 的分布函数为 $F(y)$.

证明： 由 $X \sim \mathcal{U}[0,1]$, $Y = F^{-1}(X)$ 有

$$P\{Y \leq y\} = P\{F^{-1}(X) \leq y\} = P\{X \leq F(y)\} = F(y).$$

■

基于变换法的一般分布的随机变量发生器

如果 F^{-1} 易求, 用 $\mathcal{U}[0,1]$ 随机数发生器产生序列 X_i, $i=1,2\cdots$, 则序列 $Y_i=F^{-1}(X_i)$, $i=1,2\cdots$, 就是分布函数为 $F(y)$ 的随机数序列. 这种利用 $\mathcal{U}[0,1]$ 随机数发生器和 F^{-1} 直接给出分布函数 $F(y)$ 的随机数发生器的方法称为变换法.

例： 指数分布的概率密度为

$$p(y)=\begin{cases}0, & y<0,\\ \lambda e^{-\lambda y}, & y\geq 0,\end{cases}$$

其分布函数 $F(y)=0$, $\forall y\leq 0$; $F(y)=\int_0^y p(y)dy=1-e^{-\lambda y}$, $\forall y>0$, 从而 $F^{-1}(x)=-\lambda^{-1}\ln(1-x)$, $\forall x\in(0,1)$. 由变换法,

$$Y_i=-\lambda^{-1}\ln(1-X_i),\quad i=1,2,\cdots,\quad X_i\sim\mathcal{U}(0,1),$$

是服从指数分布的随机数序列.

基于变换法的标准正态分布随机变量发生器

正态分布 (概率密度为 $p(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$) 的分布函数的逆没有解析表达式, 也没有高效算法, 直接应用变换法不是理想的选择.

Box-Muller 方法：利用 $(\int_{-\infty}^{+\infty}e^{-x^2}dx)^2=\int_0^{+\infty}\int_0^{2\pi}e^{-r^2}rdrd\theta=\pi$, 及积分变量替换 $(x_1,x_2)=(r\cos\theta,r\sin\theta)$, 有

$$\frac{1}{2\pi}e^{-\frac{x_1^2+x_2^2}{2}}dx_1dx_2=\frac{1}{2\pi}e^{-\frac{r^2}{2}}rdrd\theta=\left(\frac{1}{2\pi}d\theta\right)\cdot\left(e^{-\frac{r^2}{2}}rdr\right).$$

这将一个二维正态分布密度函数转化为两个一维分布密度函数, 其中一个给出的是均匀分布 $\mathcal{U}(0,2\pi)$ 的分布密度函数 $\frac{1}{2\pi}$, 另一个给出的是分布函数 $F(r)=\int_0^r e^{-\frac{s^2}{2}}sds=1-e^{-\frac{r^2}{2}}$ 的分布密度函数 $e^{-\frac{r^2}{2}}r$. 特别有 $F^{-1}(x)=\sqrt{-2\ln(1-x)}$.

基于变换法的标准正态分布随机变量发生器

于是二维正态分布随机数 (Y_1, Y_2) 的产生, 可通过先选取相互独立的随机数 X_1, $X_2 \sim \mathcal{U}(0,1)$, 然后 利用变换

$$\begin{cases} Y_1 = \sqrt{-2\ln X_1}\cos(2\pi X_2), \\ Y_2 = \sqrt{-2\ln X_1}\sin(2\pi X_2) \end{cases}$$

来实现. 同时 Y_1 (或 Y_2) 给出了一维正态分布随机数.

基于这种算法的标准正态分布伪随机数发生器可在许多程序包中找到.

更一般分布的随机变量的产生— 舍选法

对于定义在 $[a,b]$ 上概率密度有界的更一般的分布, 下面介绍的舍选法 (acceptance-rejection method) 是一个不错的选择.

舍选法利用了以下基本事实: 如果随机变量 (X_1, X_2) 的概率密度为 $p(x_1, x_2)$, 则 X_1 的概率密度为 $\int_{-\infty}^{+\infty} p(x_1, x_2)\, dx_2$.

要生成定义在 $[a,b]$ 上概率密度为 $p(x)$ $(\leq d < \infty)$ 的随机变量 X, 首先定义集合

$$A \triangleq \{(x,y)|x \in [a,b],\ \ y \in [0, p(x)]\},$$

然后, 取二维随机变量 $(X,Y) \sim \mathcal{U}(A)$, 则 X 分量服从概率密度为

$$\int_0^d \chi_A(x,y)dy = \int_0^{p(x)} 1dy = p(x)$$

的分布. 这就将问题转化为二维随机变量 $(X,Y) \sim \mathcal{U}(A)$ 的生成.

更一般分布的随机变量的产生— 舍选法

算法:（用舍选法生成 $[a, b]$ 上概率密度 $\leq d$ 的分布的随机变量）

1. 按分布 $\mathcal{U}[a, b]$ 生成随机数 X_i;
2. 按分布 $\mathcal{U}[0, d]$ 生成随机数 Y_i, 称为决策变量;
3. 决策：如果 $0 \leq Y_i < p(X_i)$, 则接受 X_i, 否则拒绝;
4. 转步 1.

以上算法产生的序列 X_j 是服从概率密度为 $p(x)$ 的随机数序列.

对定义域无界以及密度无界的分布, 可做适当变换将其转化为有界情形再应用舍选法.

注：当然, 也有一些更一般的随机变量的生成方法.

减小方差是提高精度的需要

由前所述, Monte-Carlo 方法求积的误差由 $\frac{\sigma}{\sqrt{N}}$ 表征, 这里 $\sigma = \sqrt{Var(f(X))}$ 为标准差, N 为样本数量.

因此, 为了提高精度, 有两个选择, 其一是增大样本数量 N, 其二是减小方差 $Var(f(X))$.

下面就来介绍如何利用理论技巧和对问题的先验知识来减小方差的一些方法.

重要性抽样法

要计算积分 $\int_0^1 f(x)\,dx$, 基本的 Monte Carlo 算法是产生随机数列 $X_i \sim$ i.i.d $\mathcal{U}[0,1]$, 然后计算

$$I(f) \approx I_N(f) \triangleq \frac{1}{N}\sum_{i=1}^{N} f(X_i).$$

我们也可以换一个视角来看 $I(f)$:

$$I(f) = \int_0^1 f(x)\,dx = \int_0^1 \frac{f(x)}{p(x)} p(x)\,dx,$$

其中 $p(x)$ 是 $[0,1]$ 上的一个正概率密度. 这样我们就得到了另一种形式的 Monte Carlo 算法:

$$I(f) \approx \frac{1}{N}\sum_{i=1}^{N} \frac{f(Y_i)}{p(Y_i)},$$

其中 $Y_i,\ i = 1, 2, \cdots, N$, 是概率密度为 $p(x)$ 的 i.i.d. 随机变量.

重要性抽样法的直观解释

直观地说, 重要性抽样法就是希望在 $f(x)$ 较大的区域多取些点, 而在 $f(x)$ 较小的区域少取些点. 如果能做到完全按比例取得话, 可以想象, 得到的数值结果将是很精确的.

举一个极端的例子，设 $f(x)=0,\ \forall x\in[0,1/2)$, $f(x)=1$, $\forall x\in[1/2,1]$. 如果按基本的 Monte Carlo 算法是产生随机数列 $X_i\sim$ i.i.d $\mathcal{U}[0,1]$, 则数值结果为:

$$I_N(f)\triangleq\frac{1}{N}\sum_{i=1}^{N}f(X_i)=\frac{N_>}{N},$$

其中 $N_>$ 是 $\geq 1/2$ 的 X_i 的个数. 而按 “重要性抽样法” 则有

$$I_N(f)\triangleq\frac{1}{2N}\sum_{i=1}^{N}f(X_i)=\frac{1}{2}=I(f),$$

其中随机数列 $X_i\sim$ i.i.d $\mathcal{U}[1/2,1]$. 取一个点就能得到精确值 !

重要性抽样法的方差分析

比较两个随机变量$f(X)$ 和 $f(Y)/p(Y)$ 的方差

$$Var_X(f) = \int_0^1 |f(x) - I(f)|^2 dx = \int_0^1 |f(x)|^2 dx - I^2(f),$$

$$Var_Y\left(\frac{f}{p}\right) = \int_0^1 \left|\frac{f(y)}{p(y)}\right|^2 p(y)dy - I^2(f) = \int_0^1 \frac{f^2(y)}{p(y)} dy - I^2(f),$$

可见，若取 $p(x)$ 使 $\int_0^1 \left(\frac{f^2(x)}{p(x)} - f^2(x)\right) dx < 0$，则方差减少。特别地，当 $f > 0$ 时，若取 $p(x) = f(x)/I(f)$，则 $Var_Y(f/p) = 0$.

重要性抽样法的例

取 $f(x) = x^2$ 和 $p(x) = 2x$ (或 $p(x) = 4x^3$), 则有

$$Var_X(f) = \int_0^1 |f(x)|^2 dx - I^2(f) = \frac{1}{5} - \frac{1}{9},$$

$$Var_Y\left(\frac{f}{p}\right) = \int_0^1 \frac{f^2(y)}{p(y)} dy - I^2(f) = \frac{1}{8} - \frac{1}{9},$$

此时有, $Var_X(f) > 6 Var_Y(f/p)$.

习题七：1, 2, 3;　上机习题七：1, 2, 3.

Thank You!