数值分析上机习题报告 (16)

张宏毅 1500017736

May 27, 2017

1 Problem A

1.1 Description

产生以函数

$$p(x)=\begin{cases}2x, & 0\le x<1/2,\\ 2-2x, & 1/2\le x<1,\\ 2x-2, & 1\le x<3/2,\\ 4-2x, & 3/2\le x\le 2,\end{cases}$$

为概率密度的随机变量，并用 Monte-Carlo 方法计算积分 $\int_0^2 \mathrm{e}^{-x}\mathrm{d}x$ 的数值，并和理论计算得到的精确解比较。计算算法的数值收敛阶。

1.2 Solution

经计算可知随机变量对应的分布函数为

$$F(x)=\begin{cases}x^2, & 0\le x<1/2,\\ 1/2-(1-x)^2, & 1/2\le x<1,\\ 1/2+(x-1)^2, & 1\le x<3/2,\\ 1-(2-x)^2, & 3/2\le x\le 2.\end{cases}$$

进而可求出分布函数的反函数

$$F^{-1}(x)=\begin{cases}\sqrt{x}, & 0\le x<1/4,\\ 1-\sqrt{1/2-y}, & 1/4\le x<1/2,\\ 1+\sqrt{y-1/2}, & 1/2\le x<3/4,\\ 2-\sqrt{1-y}, & 3/4\le x\le 1.\end{cases}$$

于是 $X=F^{-1}(U)$, $U\sim U(0,1)$ 即为以 $p(x)$ 为概率密度的随机变量。

下面利用 Monte-Carlo 方法求积分 $I=\int_0^2 \mathrm{e}^{-x}\mathrm{d}x$ 的值。在区间 $[0,1]$ 上产生 N 个独立的服从 $[0,1]$ 上均匀分布的随机变量 $X_1,X_2,\cdots,X_N$，并构造格式

$$I_N=\frac{1}{N}\sum_{i=1}^{N}2\mathrm{e}^{-2X_i},$$

则 I_N 有数学期望

$$\mathrm{E}I_N=\mathrm{E}(2\mathrm{e}^{-2X})=\int_0^1 2\mathrm{e}^{-2x}\mathrm{d}x=\int_0^2 \mathrm{e}^{-t}\mathrm{d}t.$$

因此可通过不断增大样本的选取量 N 来求积分的近似值。再来考虑算法的数值收敛阶。已知积分的精确值为 $1-\mathrm{e}^{-2}\approx 0.864665$，设选取 N 个样本点进行计算时的误差为 $\epsilon_N\approx CN^{-\alpha}$，则收敛的阶可近似为

$$\alpha\approx\frac{\ln\epsilon_N-\ln\epsilon_{Nk}}{\ln k},$$

其中 k 为正整数。

表 1 给出了利用上述方法求解数值积分的结果，每个样本容量下的积分近似及误差均为重复采样 100 次得到的平均值，收敛阶的计算公式中 $k=10$。从表中可以看出 Monte-Carlo 方法的半阶收敛速度。

表 1: 利用 Monte-Carlo 方法计算积分近似值

N	积分近似值	平均误差	数值收敛阶
10	0.848746	0.125981	
100	0.864352	0.041031	0.4872
1000	0.863911	0.012481	0.5169
10000	0.864679	0.004014	0.4926
100000	0.864786	0.001258	0.5039

2 Problem B

2.1 Description

用 Monte-Carlo 方法计算积分

$$I=\int_0^1\cos\left(\frac{x}{5}\right)\mathrm{e}^{-5x}\mathrm{d}x,$$

并分别对不使用和使用重要性抽样减小方差技术所获得的结果进行比较。

2.2 Solution

记 $f(x)=\cos(x/5)\mathrm{e}^{-5x}$，则求积分 I 的最直接的方法为，在区间 $[0,1]$ 内产生 N 个独立的服从 $U(0,1)$ 分布的随机变量 $X_1,X_2,\cdots,X_N$，并构造格式

$$I_N=\frac{1}{N}\sum_{i=1}^{N}\cos\left(\frac{X_i}{5}\right)\mathrm{e}^{-5X_i},$$

则有数学期望 $\mathrm{E}I_N = I$。

若考虑使用重要性抽样法，取 $[0,1]$ 上的概率密度 $p(x) = C\mathrm{e}^{-5x}$，$C = 5/(1-\mathrm{e}^{-5})$，则可构造格式

$$I'_N = \frac{1}{N}\sum_{i=1}^{N}\frac{f(Y_i)}{p(Y_i)} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{C}\cos\left(\frac{Y_i}{5}\right),$$

其中 $Y_1, Y_2, \cdots, Y_N$ 为相互独立的以 $p(x)$ 为密度的随机变量，对应的分布函数为 $F(x) = \frac{C}{5}(1-\mathrm{e}^{-5x})$，因而可采用变换法，利用

$$Y = -\frac{1}{5}\ln\left(1 - \frac{5U}{C}\right),\ U \sim U(0,1),$$

来生成符合要求的随机变量 Y。

已知积分的精确值为 $5(25\mathrm{e}^5 - 25\cos(1/5) + \sin(1/5))/(626\mathrm{e}^5) \approx 0.198373$。表 2 给出了以上两种方法得到的积分近似及误差（重复 100 次后取平均），从中可以看出，合理的重要性抽样策略可以明显提高 Monte-Carlo 方法的精度。

表 2: 利用重要性抽样法求数值积分

N	不抽样的近似	抽样的近似	不抽样的误差	抽样的误差
10	0.195275	0.198384	0.056010	0.000117
100	0.196846	0.198367	0.020155	0.000041
1000	0.197895	0.198373	0.005892	0.000014
10000	0.198741	0.198373	0.001886	0.000004

3 Problem C

3.1 Description

设计一种随机数抽样方法，产生满足概率密度为

$$p(x_1, x_2, x_3) = \frac{1}{8\pi}\mathrm{e}^{-r} \quad \left(r = (x_1^2 + x_2^2 + x_3^2)^{1/2}\right)$$

的随机向量 $\boldsymbol{X} = (X_1, X_2, X_3)$。

3.2 Solution

考虑球坐标变换

$$(x_1, x_2, x_3) = (r\sin\varphi\cos\theta, r\sin\varphi\sin\theta, r\cos\varphi),$$

其中 $r \in [0, +\infty), \varphi \in [0, \pi), \theta \in [0, 2\pi)$，则概率密度微元可化为

$$\begin{aligned} p(x_1, x_2, x_3)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 &= \frac{1}{8\pi}\mathrm{e}^{-r} \cdot r^2\sin\varphi\mathrm{d}r\mathrm{d}\varphi\mathrm{d}\theta \\ &= \left(\frac{r^2\mathrm{e}^{-r}}{2}\mathrm{d}r\right)\left(\frac{\sin\varphi}{2}\mathrm{d}\varphi\right)\left(\frac{1}{2\pi}\mathrm{d}\theta\right). \end{aligned}$$

因此我们只需考虑生成三个独立的随机变量 r,φ,θ，使得它们分别以以上三式为概率密度即可。对随机变量 θ，易见其服从 $[0,2\pi]$ 上的均匀分布，故只需令 $\theta=2\pi U_\theta$（$U_\theta\sim U(0,1)$）。对随机变量 φ，有分布

$$F_\varphi(\varphi)=\int_0^\varphi\frac{\sin s}{2}\mathrm{d}s=\frac{1-\cos\varphi}{2},$$

于是可利用变换法，令 $\varphi=\arccos(1-2U_\varphi)$（$U_\varphi\sim U(0,1)$）即可。对随机变量 r，由于其分布函数

$$F_r(r)=\int_0^r\frac{s^2\mathrm{e}^{-s}}{2}\mathrm{d}s=-\frac{1}{2}(r^2+2r+2)\mathrm{e}^{-r}$$

的反函数不容易计算，所以我们这里不采用变换法生成 r。注意到 r 服从 $\Gamma(3,1)$ 分布，故可考虑先生成三个独立的服从 $\Gamma(1,1)$ 分布（即参数为 1 的指数分布）的变量 e_1,e_2,e_3，再令 $r=e_1+e_2+e_3$ 即可。指数分布 $\mathcal{E}(1)$ 容易通过变换法生成，即令 $e_i=-\ln U_{e_i}$（$U_{e_i}\sim U(0,1)$）。由此可总结出以下步骤:

$$\begin{cases}U_1,U_2,U_3,U_4,U_5\sim \text{i.i.d. } U(0,1),\\ r=-\ln U_1-\ln U_2-\ln U_3,\\ \varphi=\arccos(1-2U_4),\\ \theta=2\pi U_5,\\ X_1=r\sin\varphi\cos\theta,\\ X_2=r\sin\varphi\sin\theta,\\ X_3=r\cos\varphi.\end{cases}$$

这样产生的随机向量 $\boldsymbol{X}=(X_1,X_2,X_3)$ 便符合我们的要求。图 1 是利用上述方法产生了 $N=200$ 个随机点的空间分布图，各点分布的概率和它到原点距离的指数成反比。

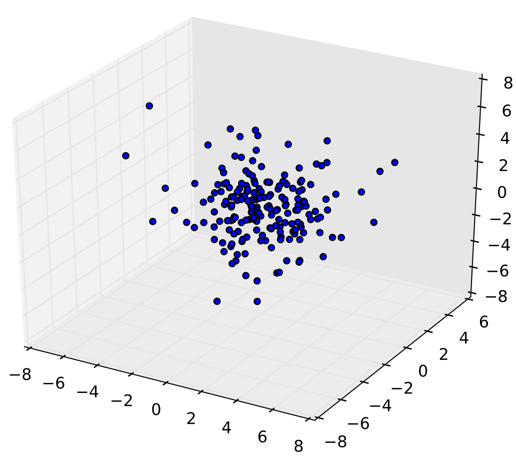

图 1: 对随机向量 $\boldsymbol{X}$ 的采样 $(N=200)$