数学实验：第八次作业

计算机系　计 73　2017011620　李家昊

2020 年 5 月 14 日

1　实验目的

- 掌握概率统计的基本概念及用 MATLAB 实现的方法。
- 用这些方法解决实际问题。

2　问题求解

2.1　Chap11-Ex5 炮弹问题（计算题）

2.1.1　算法设计

模型建立　设目标中心为 $x=0, y=0$，圆形区域半径为 $a=100$，则圆形区域可表示为 $\Omega: x^2+y^2=a^2$，炮弹的落点服从二维正态分布 $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，其中，

$$\boldsymbol{\mu}=\begin{pmatrix}0\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}=\begin{pmatrix}\sigma_x^2 & r\sigma_x\sigma_y\\ r\sigma_x\sigma_y & \sigma_y^2\end{pmatrix} \tag{1}$$

记 $\boldsymbol{x}=(x,y)^T$，其维度 $m=2$，则概率密度函数如下，图像如图 1。

$$f(\boldsymbol{x})=\left(\frac{1}{2\pi}\right)^{\frac{m}{2}}\frac{1}{\sqrt{|\boldsymbol{\Sigma}|}}\exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right] \tag{2}$$

炮弹命中圆形区域内部的概率 p 为，

$$p=\iint_{\Omega} f(x,y)dxdy \tag{3}$$

方程 (3) 即为本题的模型。

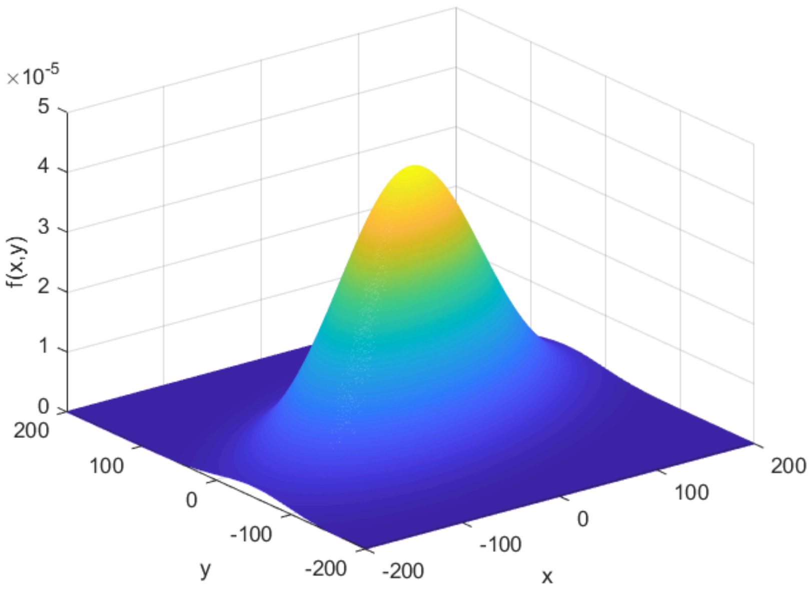

图 1: 概率密度函数 $f(x,y)$ 的图像

算法实现 考虑到方程 (3) 的二重积分无解析解，可以采用两种方法求解。

一种方法是采用蒙特卡洛方法求解，取 n 个独立均匀分布在区域 $[-a,a]\times[-a,a]$ 的随机点 $\{(x_i,y_i)\}_{i=1}^{n}$，记 I 为示性函数，则可近似计算得出，

$$p=\frac{4a^2}{n}\sum_{i=1}^{n}f(x_i,y_i)I_{(x_i,y_i)\in\Omega} \tag{4}$$

另一种方法是采用数值积分方法求解，即，

$$p=\int_{-a}^{a}\int_{-a}^{a}f(x,y)I_{(x,y)\in\Omega}dxdy \tag{5}$$

在 MATLAB 的实现中，可采用 `mvnpdf` 命令计算二维正态分布的概率密度函数 f，采用 `unifrnd` 命令生成独立均匀分布的随机数，采用 `integral2` 命令计算二维数值积分。

2.1.2 程序

请参见附录4.1。

2.1.3 计算结果

蒙特卡洛方法 分别选取 $n=10^4,10^5,10^6,10^7$，在每个 n 值下计算 5 次，得到结果如表 1。取 $n=10^7$ 时 5 次计算的均值作为最终结果，得到 $p=0.6980$。

表 1: 不同 n 取值下计算 5 次得到的概率 p 值

n	1	2	3	4	5	Mean	Variance
10^4	0.7032	0.7030	0.6991	0.6903	0.7019	0.6995	2.9125×10^{-5}
10^5	0.6977	0.6962	0.6970	0.6978	0.7014	0.6980	3.9820×10^{-6}
10^6	0.6973	0.6985	0.6978	0.6979	0.6986	0.6980	2.8700×10^{-7}
10^7	0.6980	0.6981	0.6977	0.6979	0.6984	0.6980	6.7000×10^{-8}

数值积分方法 利用数值积分方法求出的值为 $p = 0.6980$。

2.1.4 结果分析

从表 1可以看出，采用蒙特卡洛方法时，随着试验次数 n 的增大，计算结果的方差逐渐减小，逐步稳定在真值附近，当 n 达到 10^5 时，蒙特卡洛结果与数值积分结果已经非常接近，可作为真值的一个近似解。

2.1.5 结论

炮弹命中圆形区域的概率为 69.80%。

2.2 Chap11-Ex7 报童问题（计算题）

2.2.1 算法设计

模型建立 记报童每天购入报纸数量为 n，需求量为 r，服从正态分布 $N(\mu, \sigma^2)$，其概率密度函数为 f，批发价为 $a = A(1 - n/K)$，每份报纸的零售价为 b，退回价为 c，则报童每天的利润 V 为，

$$V(n) = \sum_{r=0}^{n-1} [(b-a)r - (a-c)(n-r)] f(r) + \sum_{r=n}^{\infty} [(b-a)n] f(r) \tag{6}$$

其图像如图 2，可以看到，随着购入报纸数量 n 的增大，利润 V 首先在 $n = 2000$ 附近出现了一个极大值，然后降低到负值，然而，随着 n 的继续增大，V 的值出现了急速的攀升，甚至超过了之前的极大值。这是因为随着 n 的增大，批发价格 a 逐渐降低，当 $n > 15000$ 时，批发价格竟然低于回收价格，当 $n > 50000$ 时，批发价格甚至变成了负数，在这些情况下，报童只要多购进报纸，再统一进行回收，就能赚到差价，显然是不符合实际情况的。为了保证批发价与回收价之间有一定的差价，这里对 n 的范围做如下规定，

$$0 \leq n \leq 5000 \tag{7}$$

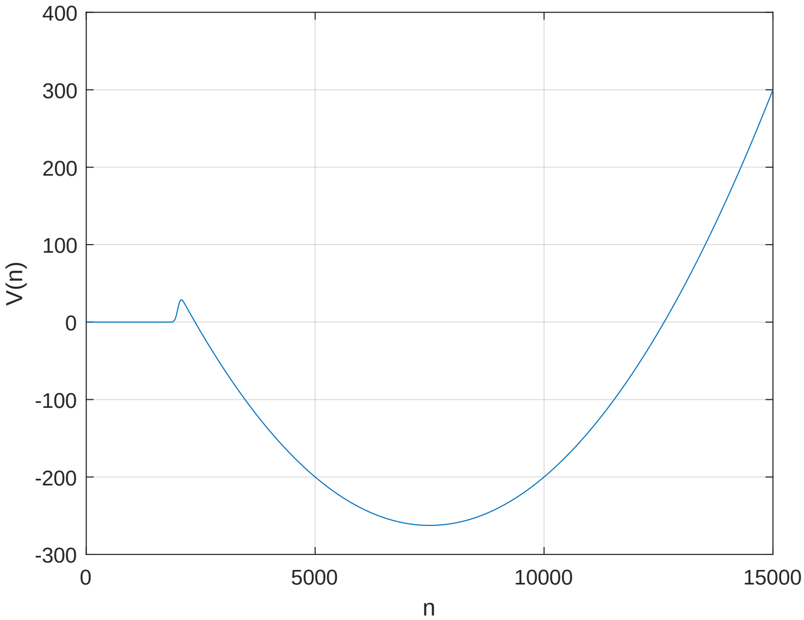

图 2: 利润 V 随购入报纸数量 n 变化的图像

将 r 和 n 看作连续变量，则方程 (6) 可写作，

$$V(n) = \int_0^n [(b-a)r - (a-c)(n-r)]f(r)dr + \int_n^{+\infty} (b-a)nf(r)dr \tag{8}$$

注意这里 a 是关于 n 的函数，两边对 n 求导，化简得，

$$V'(n) = \int_0^n \left(\frac{2An}{K} - A + c\right) f(r)dr + \int_n^{+\infty} \left(\frac{2An}{K} - A + b\right) f(r)dr \tag{9}$$

令 $V'(n) = 0$，得到，

$$\frac{\int_0^n f(r)dr}{\int_n^{+\infty} f(r)dr} = \frac{-2An + AK - bK}{2An - AK + cK} \tag{10}$$

考虑到均值 μ 比标准差 σ 大得多时，有 $\int_0^n f(r)dr \approx \int_{-\infty}^n f(r)dr$，由对称性可知 $\int_n^{+\infty} f(r)dr = 1 - \int_{-\infty}^n f(r)dr$，因此得到，

$$\int_{-\infty}^n f(r)dr = \frac{2An - AK + bK}{K(b-c)} \tag{11}$$

方程 (11) 即为本题的模型。

算法实现 方程 (11) 是一个超越方程，无解析解，可采用 `fzero` 命令求解方程，采用 `normcdf` 命令计算正态分布的累积分布函数。

2.2.2 程序

请参见附录4.2。

2.2.3 计算结果

取初值为 $n_0 = 2000$，计算得到最优购入报纸份数为 $n = 1968$。

2.2.4 结果分析

当限制 $0 \le n \le 5000$ 时，批发价格 a 的取值范围为 $0.45 \le a \le 0.5$，此时始终满足 $c \le a \le b$，且批发价格与回收价格之间至少有 0.1 的差价，这是符合实际情况的。在这种限制下，求得的结果为全局最大值，可以作为实际应用的参考。

2.2.5 结论

为了获得最大利润，报童每天购进的报纸数应为 1968 份。

2.3 Chap11-Ex9 轧钢问题（计算题）

2.3.1 算法设计

模型建立 由题意，钢材的规定长度为 l，粗轧得到的钢材长度服从正态分布 $N(m, \sigma^2)$，记其概率密度函数为 $f(x)$，累积分布函数为 $F(x)$。

对于第 1 个目标函数，每粗轧一根钢材的浪费长度为，

$$u(m) = \int_{-\infty}^{l} x f(x) dx + \int_{l}^{+\infty} (x - l) f(x) dx = m - l(1 - F(l)) \tag{12}$$

对于第 2 个目标函数，每得到一根规定长度钢材的浪费长度为，

$$v(m) = \frac{u(m)}{1 - F(l)} = \frac{m}{1 - F(l)} - l \tag{13}$$

分别求出最佳均值 m，使得 $u(m)$ 和 $v(m)$ 取最小值即可。

算法实现 对于累积分布函数，可采用 `normcdf` 命令，对于最小值求解，可采用 `fminunc` 命令。

2.3.2 程序

请参见附录4.3。

2.3.3 计算结果

为了生产出合格的钢材并减少浪费，最佳均值 m 应当在 2 附近，首先画出区间 [2,3] 中的 $u(m)$ 和 $v(m)$ 图像，如图 3。可以看出，最小值在 $m = 2.3$ 附近，因此将其作为初值进行求解。

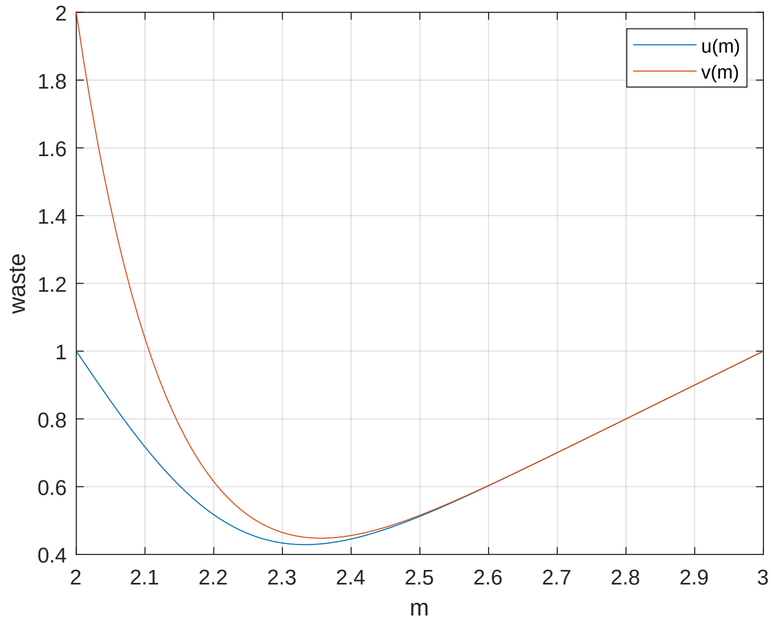

图 3: 两种目标函数的图像

对于第 1 个目标函数，求解得到 $m = 2.3327$，此时浪费长度最小，为 $u(m) = 0.4289$。

对于第 2 个目标函数，求解得到 $m = 2.3562$，此时浪费长度最小，为 $v(m) = 0.4479$。

2.3.4 结果分析

由图 3可以看出，当 m 低于最优值时，粗轧得到钢管长度较短，导致整根报废的钢管数量较多，均摊到每根钢管的浪费长度较大；当 m 高于最优值时，粗轧得到的每根钢管长度较长，精轧时的浪费长度也较大；当 m 处于最优值时，达到了两者的平衡，此时平均浪费长度最小。

实际生产以利润为最终目标，生产利润往往与合格钢材的数量成正比，当生产每根合格钢材的浪费最小时，生产总利润最大，因此，在实际应用中，应当采用第 2 个目标函数。

2.3.5 结论

要求每粗轧一根钢材的浪费最小时，最优的粗轧钢材长度均值为 2.3327 米，此时平均浪费长度最小，为 0.4289 米。

要求每得到一根规定长度钢材的浪费最小时，最优的粗轧钢材长度均值为 2.3562 米，此时平均浪费长度最小，为 0.4479 米。

3 收获与建议

在本次实验中，我掌握了 MATLAB 求解概率统计问题的方法，用概率统计方法建立了实际问题的模型，并进行求解，在解决实际问题的过程中，我对数学方法的原理和应用有了更深刻的理解。

希望助教能对每次的实验进行详细的解答，希望老师在未来的课堂上介绍更多数学应用的前沿知识。

4 附录：程序代码

4.1 Chap11-Ex5

```
global a mu sigma;
a = 100;
sx = 80;
sy = 50;
r = 0.4;
mu = [0 0];
sigma = [sx^2  r * sx * sy; r * sx * sy  sy^2];

n = 100000;
x = unifrnd(-a, a, n, 2);
idx = x(:,1).^2 + x(:,2).^2 <= a^2;
x = x(idx, :);

p = 4 * a^2 * sum(mvnpdf(x, mu, sigma)) / n;
p

p = integral2(@pdf, -a, a, -a, a);
p

function f = pdf(x, y)
    global a mu sigma;
    idx = x.^2 + y.^2 <= a^2;
    f = zeros(size(x));
    f(idx) = mvnpdf([x(idx) y(idx)], mu, sigma);
end
```

4.2 Chap11-Ex7

```
mu = 2000;
sigma = 50;
```

```
A = 0.5;
K = 50000;
b = 0.5;
c = 0.35;

fun = @(n) normcdf(n,mu,sigma) - (2*A*n/K - A + b) / (b-c);

fzero(fun, 2000)
```

4.3 Chap11-Ex9

```
sigma = 0.2;
l = 2;

m = 2:0.01:3;
func_u = @(m) m - l.*(1-normcdf(l,m,sigma));
func_v = @(m) m ./ (1-normcdf(l,m,sigma)) - l;

u = func_u(m);
v = func_v(m);
plot(m, u, m, v);
xlabel('m'); ylabel('waste'); grid on;
legend('u(m)', 'v(m)');

[m, fval] = fminunc(func_u, 2.3);
m, fval
[m, fval] = fminunc(func_v, 2.3);
m, fval
```