

# 19: Metropolis抽样法

许传奇 PB16021546

## 1 题目

设体系的能量为  $H = x^2/2\sigma_x^2 + y^2/2\sigma_y^2$  (以  $kT$  为单位), 采用 Metropolis 抽样法计算  $\langle x^2 \rangle, \langle y^2 \rangle, \langle x^2 + y^2 \rangle$  并与解析结果进行比较。抽样时在 2 维平面上依次标出 Markov 链点分布, 从而形象地理解 Markov 链。

## 2 原理与算法

### 2.1 原理

#### 2.1.1 Markov 链

当满足条件:

$$p_n(x_1, \dots, x_n; t_1, \dots, t_n) = p_n(x_1, \dots, x_n; t_1 + t, \dots, t_n + t) \quad (1)$$

时, 称随机过程是平稳的, 这时的起点是无关紧要的。

我们称一个平稳的随机序列 Markov 的, 如果某一时刻  $x$  取值的条件几率是独立于上一时刻之前的所有  $x$  值的话, 即:

$$p(x_n | x_{n-1}, \dots, x_1) = p(x_n | x_{n-1}) \quad (2)$$

该式表示, 某一步的结果仅仅依赖于上一步, 与更前面的历史无关, 对应的态序列  $(x_1, \dots, x_n)$  称为 Markov 链。

Markov 链的极限分布是与初始分布的选择无关的, 仅取决于转移概率。设达到平衡后各个状态的概率组成的向量为  $\mathbf{p}$ , 跃迁矩阵为  $\mathbf{W}$ , 则达到平衡后, 应有:

$$\mathbf{p} = \mathbf{p}\mathbf{W} \quad (3)$$

$$p_i = \sum_j p_j W_{ji} \quad (4)$$

#### 2.1.2 Metropolis 抽样法

Metropolis 方法的实际做法有点像舍选法, 设  $p(x)$  为所考虑的几率密度分布, 并且已经产生了  $x_1, x_2, \dots, x_n$  个抽样点, 现在的问题是如何产生下一个抽样点  $x_{n+1}$ ?

可以在上一个点附近构造一个试探解,  $x_t = x_n + \delta$ ,  $\delta$  是试探步长 (可正可负, 例如可取  $\delta = (\xi - 0.5)\Delta x$ ,  $\Delta x$  是固定步长,  $\xi \in (0, 1)$  是均匀分布的随机数), 该点是否被选择取决于比值  $r = p(x_t)/p(x_n)$ :

1. 如果  $r > 1$  则选取，即  $x_{n+1} = x_t$  ( $r > 1$ )。
2. 否则，产生  $[0, 1]$  区间内均匀分布的随机数  $\xi$ ，如果  $\xi < r$  则选取；也即，使得  $x_{n+1} = x_t$  ( $r < 1$ ) 的选取概率为  $r$ 。否则，放弃；即  $x_{n+1} = x_n$ 。

通过 Metropolis 算法可以产生大量离散的  $x$  值，构成 Markov 链。其中序列中热化过程的前  $M$  个构型被舍去。注意在计算体系系综统计平均时，需要将所有有效步数统计在内（不包括热化阶段）而不能只保留选取成功的步数和构型，如第  $i$  个构型计算出物理量的值为  $A_i$ ，则系综平均为：

$$\langle A \rangle = \frac{1}{N-M} \sum_{i=M+1}^N A_i \quad (5)$$

### 2.1.3 题目的理论推导

由题中所设，坐标为  $(x, y)$  时的概率为：

$$\rho(x, y) = \frac{1}{Z} \exp\left(-\frac{H}{kT}\right) = \frac{1}{Z} \exp(-(x^2/2\sigma_x^2 + y^2/2\sigma_y^2)) \quad (6)$$

其中的  $Z$  为配分函数，其解析解为：

$$Z = \iint_{-\infty}^{+\infty} \exp(-(x^2/2\sigma_x^2 + y^2/2\sigma_y^2)) dx dy = 2\pi\sigma_x\sigma_y \quad (7)$$

于是概率为：

$$\rho(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp(-(x^2/2\sigma_x^2 + y^2/2\sigma_y^2)) \quad (8)$$

可以通过系综的分布函数求出系综物理量的期望值：

$$\langle x^2 \rangle = \iint_{-\infty}^{+\infty} x^2 \rho(x, y) = \sigma_x^2 \quad (9)$$

$$\langle y^2 \rangle = \iint_{-\infty}^{+\infty} y^2 \rho(x, y) = \sigma_y^2 \quad (10)$$

$$\langle x^2 + y^2 \rangle = \iint_{-\infty}^{+\infty} (x^2 + y^2) \rho(x, y) = \sigma_x^2 + \sigma_y^2 \quad (11)$$

(12)

其中，第三项的期望也可以直接由前两项的期望相加得到，因为期望满足线性性。

## 2.2 算法

根据原理部分所述的 Metropolis 算法即可。

对于该系综，需要选择总次数  $N$  和热化阶段的次数  $M$ 。首先设置初始的坐标为  $(x_0, y_0)$ 。

然后进行 Metropolis 抽样法。即：

1. 设当前坐标为  $(x_n, y_n)$  抽取两个  $[-1, 1]$  的随机数  $\xi_x, \xi_y$ ，认为此步下的坐标变化到  $(x_t, y_t) = (x_n + \xi_x \Delta x, y_n + \xi_y \Delta y)$ 。计算变化后与变化前的分布概率之比，即  $r = \exp(-(x_t^2/2\sigma_x^2 + y_t^2/2\sigma_y^2) - (x_n^2/2\sigma_x^2 + y_n^2/2\sigma_y^2))$ 。
2. 判断  $r$  与 1 的大小关系：

- (a) 如果 $r > 1$ , 则选取, 即 $x_{n+1} = x_t, y_{n+1} = y_t (r > 1)$ 。
- (b) 如果 $r < 1$ , 则产生 $[0, 1]$ 区间内均匀分布的随机数 $\xi$ 。判断 $\xi$ 与 $r$ 的关系:
- 如果 $\xi < r$ , 则选取; 也即, 使得 $x_{n+1} = x_t, y_{n+1} = y_t (r < 1)$ 的选取概率为 $r$ 。
  - 如果 $\xi > r$ , 则放弃; 即 $x_{n+1} = x_n, y_{n+1} = y_n$ 。

3. 若达到规定的迭代次数, 则退出; 否则返回第一步。

按照上述方法可以得到该系综的抽样。

对于物理量期望的求法, 直接运用离散分布的定义即可。即:

$$\langle A \rangle = \frac{1}{N - M} \sum_{i=M+1}^N A_i \quad (13)$$

对于本题, 有:

$$\langle x^2 \rangle = \frac{1}{N - M} \sum_{i=M+1}^N x_i^2 \quad (14)$$

$$\langle y^2 \rangle = \frac{1}{N - M} \sum_{i=M+1}^N y_i^2 \quad (15)$$

$$\langle x^2 + y^2 \rangle = \frac{1}{N - M} \sum_{i=M+1}^N (x_i^2 + y_i^2) \quad (16)$$

### 3 源文件使用说明

编译并运行“19\_Metropolis.cpp”, 将弹出命令行, 按要求输入`sigma_x`、`sigma_y`、基准值步长`step`、固定步长`step`、起始横坐标`x0`和起始纵坐标`y0`, 程序将运行。

输入数据后, 程序运行并将热化阶段后的每步的横、纵坐标输出到文件“`sx=输入的sigma_x,sy=输入的sigma_y,step=输入的步长.txt`”中。

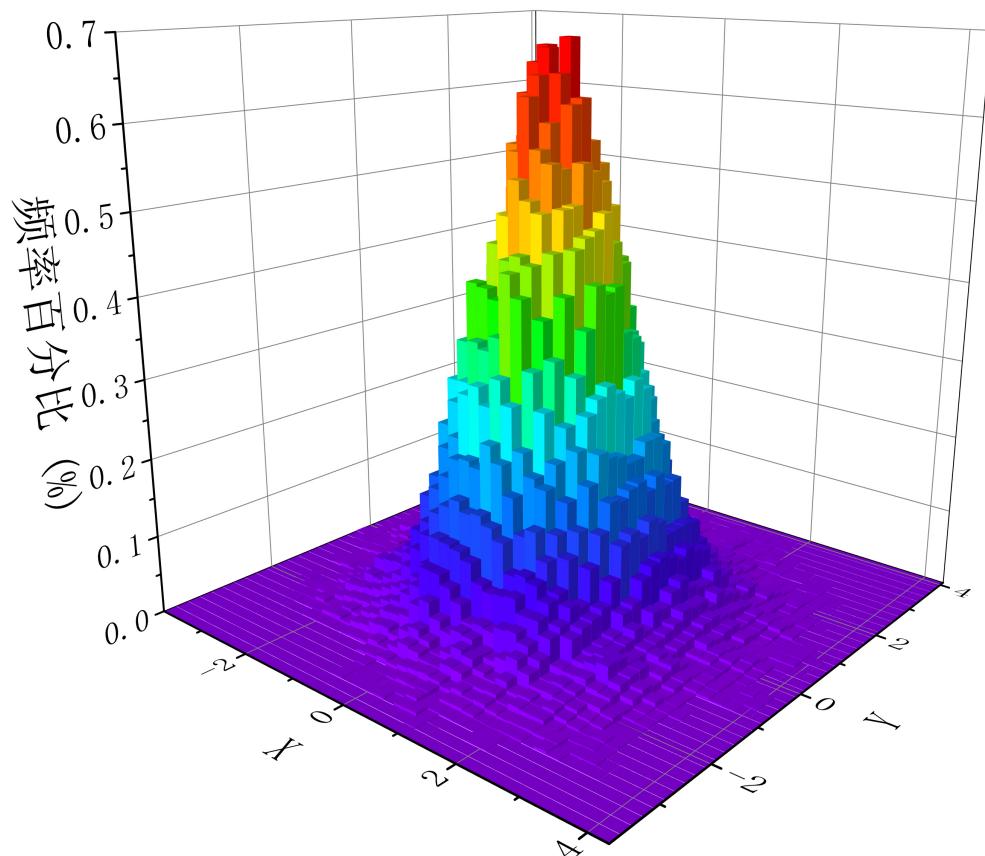
程序运行结束时, 在命令行打印计算的 $\langle x^2 \rangle, \langle y^2 \rangle, \langle x^2 + y^2 \rangle$ 。

## 4 计算结果及具体分析

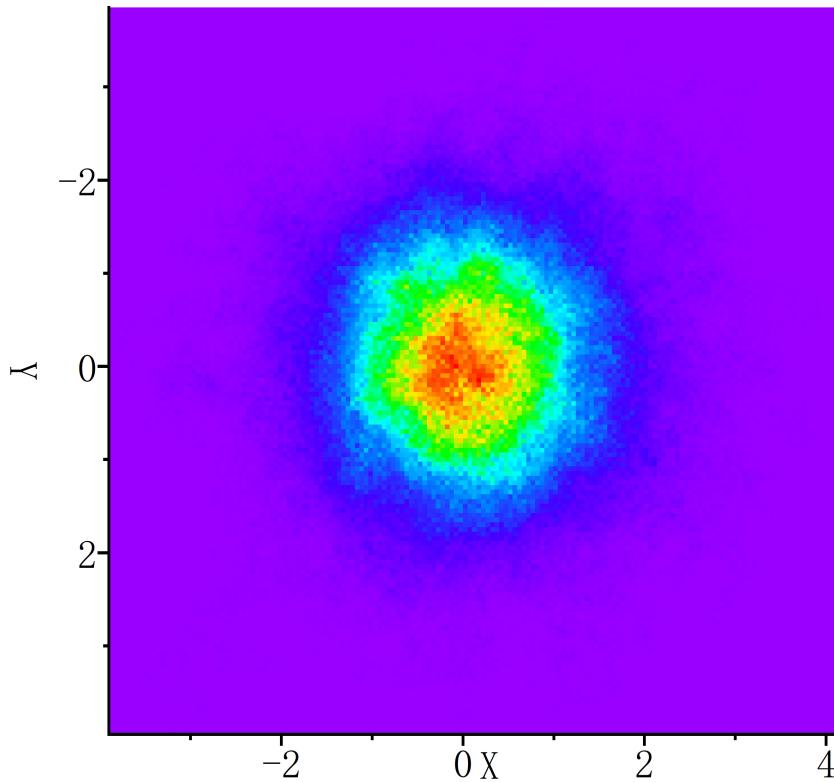
### 4.1 正则系综分布图

选择 $N = 10^6, M = 10^5, \sigma_x = 1, \sigma_y = 1, \Delta x = 0.1$ , 得到的正则系综分布示意图在下页展示。

从图中可以看出, 抽样得到的分布基本上满足正则系综的分布, 为二维的正态分布曲线(未进行归一化)。



(a) 正则系综分布示意图



(b) 正则系综分布俯视图

图 1: 正则系综分布示意图

## 4.2 计算 $\langle x^2 \rangle, \langle y^2 \rangle, \langle x^2 + y^2 \rangle$

选择 $N = 10^6, M = 10^5, \sigma_x = 1, \sigma_y = 1, \Delta x = 0.1$ , 命令行中得到的结果如下

```
 $\langle \hat{x}^2 \rangle = 0.976666$ 
 $\langle \hat{y}^2 \rangle = 1.030696$ 
 $\langle \hat{x}^2 + \hat{y}^2 \rangle = 2.007363$ 
```

图 2: 计算的 $\langle x^2 \rangle, \langle y^2 \rangle, \langle x^2 + y^2 \rangle$

$$\langle x^2 \rangle = 0.976666, \langle y^2 \rangle = 1.030696, \langle x^2 + y^2 \rangle = 2.007363 \quad (17)$$

对应的标准值和误差为:

$$\langle x^2 \rangle = 1, \langle y^2 \rangle = 1, \langle x^2 + y^2 \rangle = 2 \quad (18)$$

$$\eta_x = 2.3334\%, \eta_y = 3.0696\%, \eta_{xy} = 0.3682\% \quad (19)$$

具体的误差和步数的关系如下图所示:

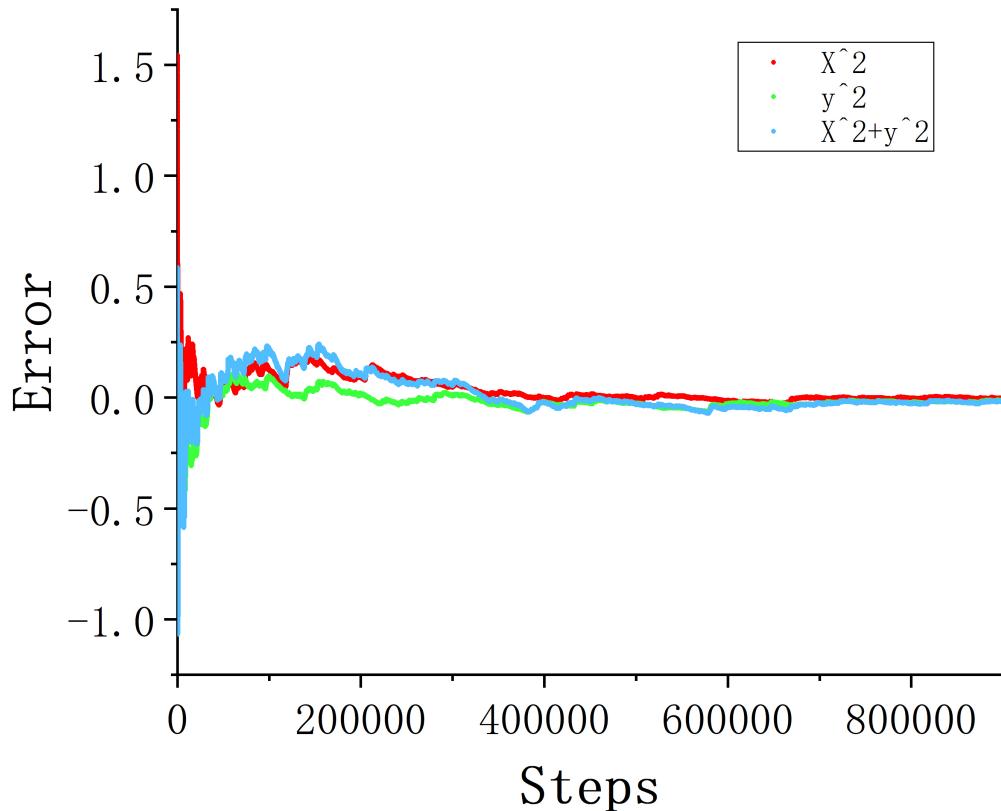


图 3: 误差和步数的关系

与理论值的偏差也较小。

当把热化步数和总步数选择更多的时候，得到的结果误差应该更小，因为系综更趋近于正则系综。

### 4.3 Markov链点分布

由于要展示Markov链点，因此热化步数选择0。

起始坐标选择为(100,100)，得到下图所示的变化：

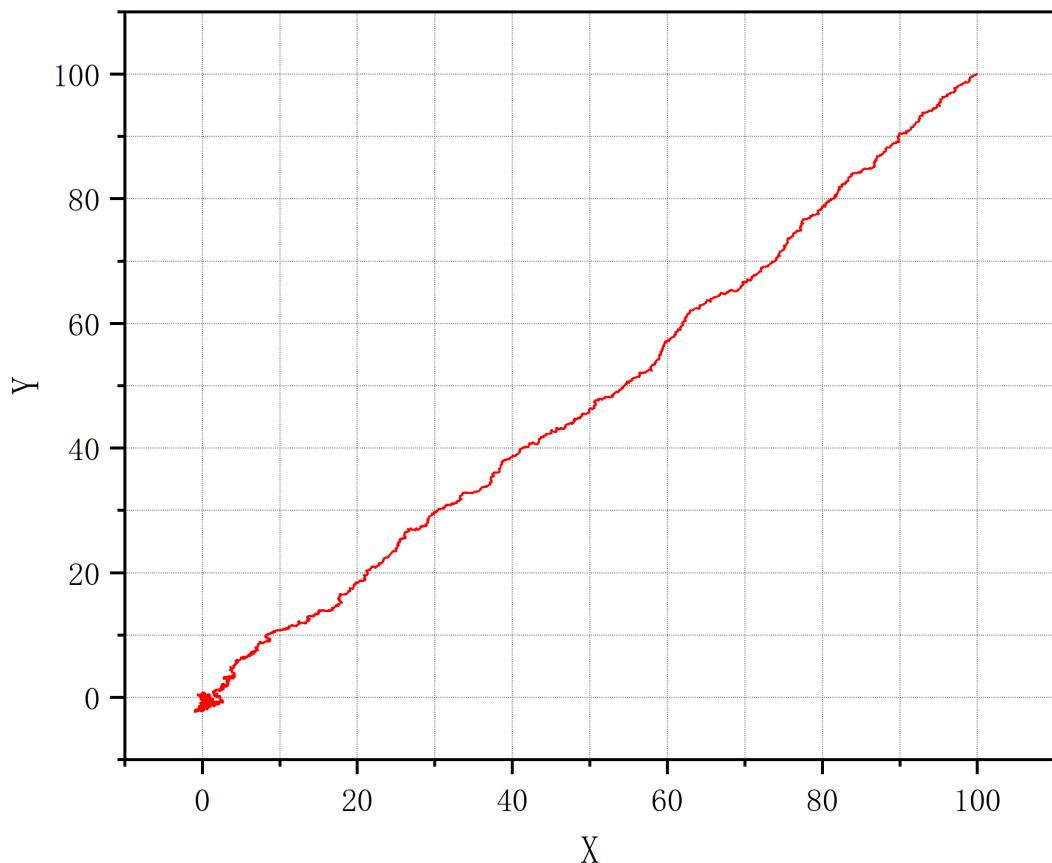


图 4: 起始坐标为(100,100)时的链点变化

起始坐标选择为(10,10)，得到下图所示的变化：

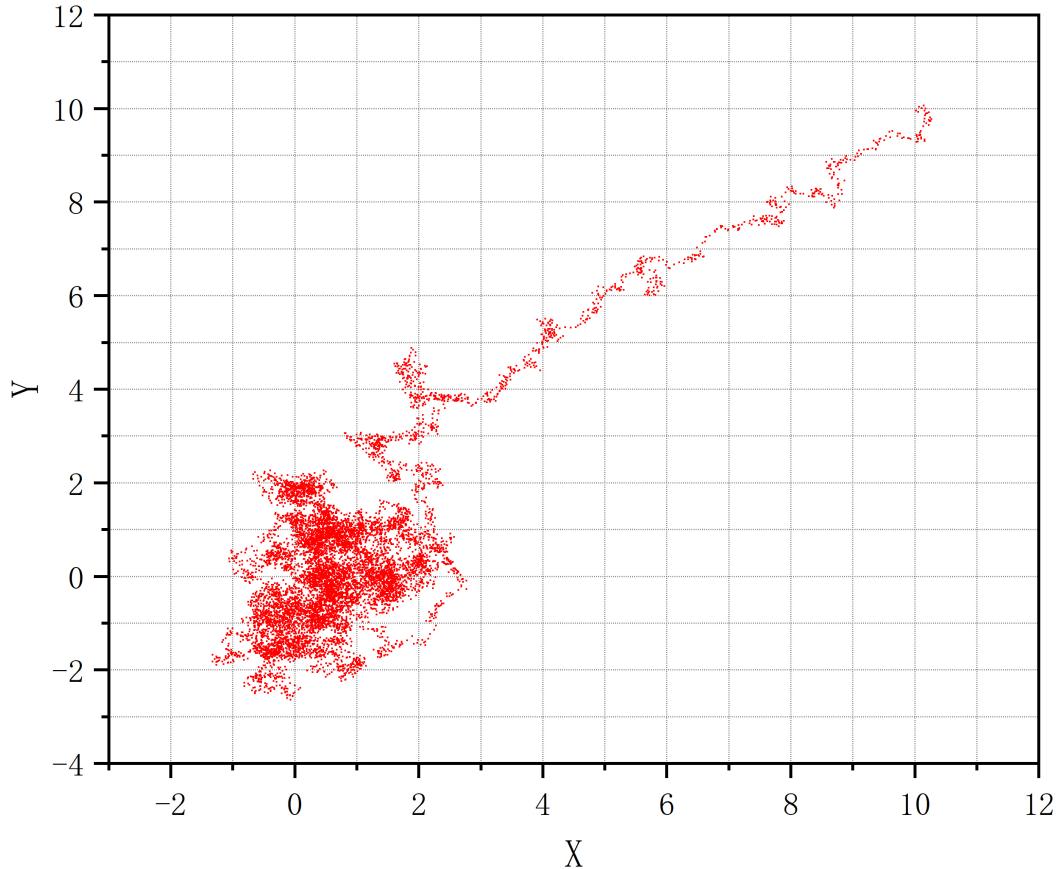


图 5: 起始坐标为(10,10)时的链点变化

从上图中可以看出，初始坐标选择较大的(100,100)，得到的变化几乎是一条直线到原点。进一步调整初始坐标，设置为(10,10)，也是很快地到原点后在原点附近变化。

主要是因为该题中的正则系综的分布可以认为是二维正态分布，原点位于中心，因此链点几乎都在原点附近。

## 5 总结

本次作业用Metropolis方法抽取了正则系综。

由实验的结果可知，最后抽样得到的粒子位置分布接近于二维正态分布，而这个分布的具体形式跟初始时给定的 $\sigma_x, \sigma_y$ 有关。

计算的结果也跟初始时给定的 $\sigma_x, \sigma_y$ 有关，本次选择的 $\sigma_x, \sigma_y$ 并无特殊性，而与理论值的误差很小，因此程序编写正确。

初始时的起始点位置对系统最后状态的影响不大，因为有热化过程，而且算法的设置让Markov链的点迅速从初始点移动到原点附近，主要是因为这个路径上接受改变的概率高，而离开原点的接受概率

低。

迭代的步长也会影响程序运行的情况：在理论上可知，如果迭代步长太小的话，系统变化得十分缓慢，会延长程序运行到近似正则系统的时间；而迭代步长太大的话，则系统运动过于剧烈，离散性更大，与真实情况相差较大。同时，系统接收的概率也会减小，这也增加了程序的运行时间和效率。

因此需要选择合适的步长，这是从选择上提高程序运行效率的方法。