

目 录

第一章 蒙特卡罗方法概述	3
§1.1 MC 方法的发展	3
§1.2 MC 方法的三个简单例子	3
§1.3 MC 的基本思想 (求平均)	6
§1.4 MC 的收敛性, 误差和费用	6
§1.5 MC 的特点	7
第二章 由已知分布的抽样	10
§2.1 由已知分布的随机抽样	10
§2.2 直接抽样方法	10
§2.3 舍选抽样法	12
§2.4 复合抽样方法	14
§2.5 变换抽样方法	16
§2.6 近似抽样方法	17
§2.7 若干重要分布的随机抽样	18
第三章 MC 的一般技巧	21
§3.1 降低实验方差的特性	21
§3.2 期望估计技巧	22
§3.3 重要抽样技巧	23
§3.4 分区间抽样技巧 (Stratification method)	26
§3.5 相关抽样技巧 (Correlation method)	27
§3.6 对偶抽样技巧 (Antithetic Variate)	28
第四章 非归一化 Metropolis 抽样方法	
(MC 在统计物理中的应用)	32
§4.1 预备知识	32
§4.2 正则系综平均量的计算	33
§4.3 Metropolis 抽样方法	34
§4.4 热浴法 (Heat bath)	36
§4.5 广义 Metropolis 抽样方法	38
第五章 蒙特卡罗方法模拟主方程 (跃迁过程)	40
§5.1 主方程差分解的 MC 模拟	40
§5.2 时间相关平均量 $\langle A_t(x) \rangle$ 的蒙特卡罗估计	43
§5.3 主方程的直接蒙特卡罗模拟	44

第六章	MC 求解随机微分方程 (Langevin Equations)	45
§6.1	噪声与随机力 (noise and random force)	45
§6.2	随机泰勒展开	46
§6.3	Fokker-Planck 方程 (FPE)	48
§6.4	色噪声驱动 LE	50
§6.5	随机龙格 - 库塔算法	52
§6.6	随机积分方法	54
§6.7	数值策略	54
第七章	Ising 模型的蒙特卡罗模拟	
	Metropolis 抽样的成功应用的例子	56
§7.1	Ising Model	56
§7.2	Ising 模型的 MC 模拟	59
§7.3	二元合金系统	60
§7.4	XY 模型	61
第八章	量子蒙特卡罗方法初步	63
§8.1	量子多体系统的最低能量	63
§8.2	McMillan-Metropolis 方法	64
§8.3	偏倚抽样法求极小能量	65

第一章 蒙特卡罗方法概述

Monte Carlo Methods(MC)

§1.1 MC 方法的发展

- 其它名称:

- (1) 统计实验方法 (Statistical Testing Method)

- (2) 随机抽样技术 (Random Sampling)

- (3) 计算机模拟 (Random Simulation)

- 20 世纪 40 年代计算机出现, 1946 年 Von Neumann 在计算机上用 R-S(Random Simulation) 方法模拟了中子链式反应 (应用于原子弹的研制)。

- 1777 年, 法国学者蒲丰 (Buffon) 用随机投针方法求 π :

$2a$ 平行线束, 投长为 $2l$ 的针 ($l < a$ 保证不同时与两线相交), 相交概率为

$$p = \frac{2l}{\pi a} \quad (1.1)$$

$$\pi \approx \frac{2l}{a} \left(\frac{1}{p} \right) = \frac{2l}{a} \left(\frac{N}{n} \right) \quad (1.2)$$

其中, n 为成功的交点数, N 为实验次数。

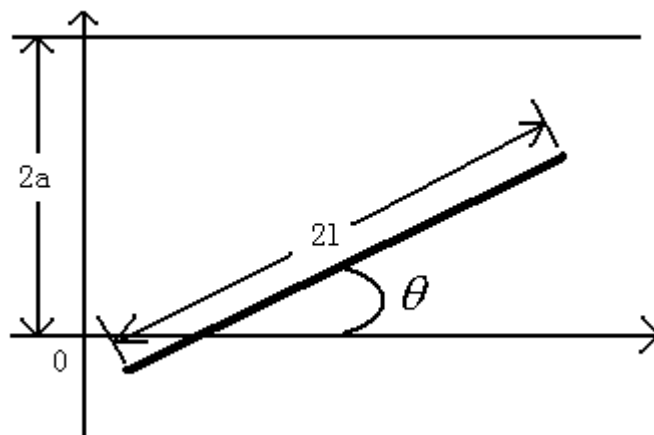
§1.2 MC 方法的三个简单例子

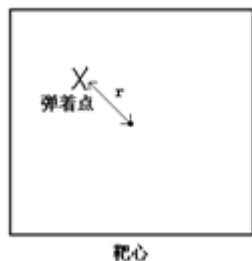
一. 射击问题 (MC 方法求定积分 \Rightarrow 求平均)

r : 弹着点到靶心的距离

$g(r)$: 弹着点的成绩

$f(r)$: 弹着点的分布。





则射击成绩:

$$\langle g \rangle = \int_0^\infty f(r)g(r)dr \quad (1.3)$$

概率语言: $\langle g \rangle$ 就是随机变量 $g(r)$ 的数学期望。 $f(r)$ 是一个分布密度函数, 满足非负归一:

$$f(r) > 0 \quad \int_0^\infty f(r)dr = 1 \quad (1.4)$$

假设射击运动员射击 N 次, 弹着点依次为 $\{r_1, r_2, \dots, r_N\}$, 简单子样 r_n 满足同一分布 $f(r)$, 得分平均值:

$$\hat{G}_N = \frac{1}{N} \sum_{n=1}^N g(r_n) \quad (1.5)$$

r_n 由 $f(r)$ 抽样得到, \hat{G}_N 是 $\langle g \rangle$ 的一个近似估计。

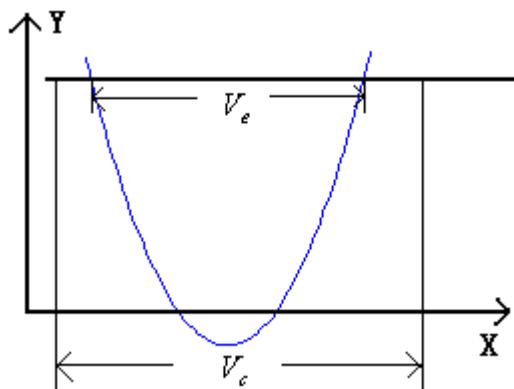
二. 线性代数方程组的求解 (根)

$$\sum_{j=1}^m a_{ij}x_j = b_i \quad (i = 1, 2, \dots, m) \quad (1.6)$$

转化为求如下二次式的最小值问题:

$$V(x_1, x_2, \dots, x_m) = \sum_{i=1}^m \alpha_i \left(\sum_{j=1}^m a_{ij}x_j - b_i \right)^2 \quad (1.7)$$

$\alpha_i > 0$ 为实数。 $y = V(x_1, x_2, \dots, x_m)$ 是一个 m 维抛物面, 被一超平面 $y = c$ 所截, $V(x_1, x_2, \dots, x_m) \leq c$ 形成一个椭圆体 V_e 。



- 如果: $(x_1^*, x_2^*, \dots, x_m^*)$ 是该椭圆体的对称中心点, 即对 V_e 内的任意点 (x_1, x_2, \dots, x_m) , 总有

$$V(x_1^* - x_1, x_2^* - x_2, \dots, x_m^* - x_m) = V(x_1^* + x_1, x_2^* + x_2, \dots, x_m^* + x_m),$$

结论：在此对称中心处二次式（1.7）取最小值。

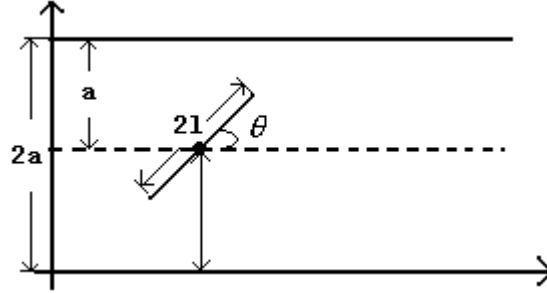
☆极值点的中值方法：

①确定一个 m 维长方体 V_c ， $A_i \leq x_i \leq B_i$ 使 $V_e \in V_c$ ， V_e 包含在 V_c 内。

②在 V_c 中产生 N 个均匀分布点 $P^{(1)} = (x_1^{(1)}, \dots, x_m^{(1)})$, $P^{(2)} = (x_1^{(2)}, x_2^{(2)}, \dots, x_m^{(2)})$, \dots , $P^{(N)} = (x_1^{(N)}, x_2^{(N)}, \dots, x_m^{(N)})$

③对于属于 V_e 的全部点 ($V(p^{(i)}) \leq c$)，按由小到大排队，取各自的中点 $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_m \Rightarrow$ 即为 (1.7) 的近似解。 \hat{x}_i 是 $\{x_i^{(N)}\}$ 中点。

三. 蒲丰氏问题 (Buffon)



- 一根针的位置用中心坐标 x ，与平行线夹角 θ 来决定 (x, θ) ， y 方向不影响相交。
- 任意投针，意味着 $\theta \in [0, \pi]$ 任意均匀， $x \in [0, a]$ 任意均匀。
- 针与平行线束相交的条件：

$$x \leq l \sin \theta, \quad 0 \leq x \leq a \quad (1.8)$$

●模拟投针

x 在 $[0, a]$ 中取任何值的概率都一样， x 的密度函数为

$$f_1(x) = \begin{cases} \frac{1}{a} & 0 \leq x \leq a \\ 0 & \text{else} \end{cases} \quad (1.9)$$

类似地， θ 的密度函数为

$$f_2(\theta) = \begin{cases} \frac{1}{\pi} & 0 \leq \theta \leq \pi \\ 0 & \text{else} \end{cases} \quad (1.10)$$

产生任意的 (x, θ) 的过程：由 $f_1(x)$ 抽样 x ，由 $f_2(\theta)$ 抽样 θ ，

$$\begin{cases} x = a\xi_1 \\ \theta = \pi\xi_2 \end{cases} \quad (1.11)$$

$\xi_1, \xi_2 \in (0, 1)$ 是两个独立的均匀随机数。

●最后，相交数 $S(x, \theta)$ 做为相交几率的估计

$$S(x_i, \theta_i) = \begin{cases} 1 & x_i \leq l \sin \theta_i \\ 0 & \text{else} \end{cases} \quad (1.12)$$

投针 N 次,

$$S_N = \frac{1}{N} \sum_{i=1}^N S_i(x_i, \theta_i) \quad (1.13)$$

是相交几率 P 的估计值, 则 $\pi \doteq \frac{2l}{a}(\frac{1}{p})$

圆周率 π 的实验值

实验值	年份	投针次数	π 的估计值
Wolf	1850	5000	3.1596
Smith	1855	3204	3.1553
Fox	1894	1120	3.1419
Lazzarini	1901	3408	3.1415929

§1.3 MC 的基本思想 (求平均)

(1) 确定一个统计量, 其数学期望正好等于所需的值。

$$G = E(g) = \int g(x)f(x)dx \quad (1.14)$$

$$\hat{G}_N = \frac{1}{N} \sum_{n=1}^N g(x_n), \quad x_n \text{ 由 } f(x) \text{ 抽样} \quad (1.15)$$

$$G = \hat{G}_N \quad (1.16)$$

(2) 构造一个人为的概率过程, 它的某些参量正好是所求问题的解。

§1.4 MC 的收敛性, 误差和费用

1. MC 的收敛性

$$P(\lim_{N \rightarrow \infty} \hat{G}_N = G) = 1 \quad (1.17)$$

的充要条件是无偏统计量 $g(x)$ 满足

$$E(|g|) = \int |g(x)|f(x)dx < +\infty \quad (1.18)$$

也就是说, 是否收敛, 取决于它的绝对值的期望值是否存在。

2. MC 的收敛速度

如果无偏统计量 $g(x)$ 满足

$$E(|g|^r) = \int |g(x)|^r dF(x) < +\infty, \quad 1 \leq r \leq 2 \quad (1.19)$$

则

$$P(\lim_{N \rightarrow \infty} N^{\frac{r-1}{r}} (\hat{G}_N - G) \rightarrow 0) = 1 \quad (1.20)$$

亦即 \hat{G}_N 依概率 1 收敛于 G 的速度为 $N^{-(\frac{r-1}{r})}$, $v \leq N^{-\frac{1}{2}}$ 收敛。

3. MC 的误差

根据中心极限定理，只要所确定的无偏估计量 $E(|g|)$ 是具有有限的方差 $\sigma^2 (\sigma^2 \neq 0)$ ，对于任意非负 X 均有

$$\lim_{N \rightarrow \infty} P\left(\frac{\sqrt{N}}{\sigma} |\hat{G}_N - G| < X\right) = \frac{1}{\sqrt{2\pi}} \int_{-X}^X e^{-\frac{t^2}{2}} dt < 1 \quad (1.21)$$

因此，当 N 足够大时，有如下近似等式

$$P(|\hat{G}_N - G| < \frac{X\sigma}{\sqrt{N}}) \approx \frac{1}{\sqrt{2\pi}} \int_{-X}^X e^{-\frac{t^2}{2}} dt = 1 - \alpha \quad (1.22)$$

α 为置信度， $1 - \alpha$ 为置信水平。

4. MC 的费用

设误差为 ε ，置信水平为 $1 - \alpha$ ， X 按正态积分表由 $1 - \alpha$ 确定，则

$$\frac{X\sigma}{\sqrt{N}} \leq \varepsilon \quad (1.23)$$

子样容量： $N \geq (\frac{X\sigma}{\varepsilon})^2$

进一步假设每计算一次无偏统计量所需要费用是 C ，则 MC 总费用为：

$$NC \approx (\frac{X\sigma}{\varepsilon})^2 C = (\frac{X}{\varepsilon})^2 \sigma^2 C \quad (1.24)$$

故 MC 总费用与方法所确定的无偏统计量的方差 σ^2 和其费用 C 的乘积成正比。

小结：

① MC 要点：

- 欲求数学期望： $G = E(g) = \int g(x)f(x)dx = \int g(x)dF(x)$
- 产生随机变量 x 的抽样 x_1, \dots, x_N ，其相应的统计量 $g(x_1), \dots, g(x_N)$ 的算术平均值：

$$\hat{G}_N = \frac{1}{N} \sum_{n=1}^N g(x_n) \quad (1.25)$$

作为 G 的近似估计。

② MC 最低要求，能计算一个与计算步数 N 有关的统计量 \hat{G}_N ，对任意的 $\varepsilon > 0$ ，应有

$$P(\lim_{N \rightarrow \infty} |\hat{G}_N - G| < \varepsilon) = 1 \quad (1.26)$$

- ③ 随机变量的期望 $E(g) = \bar{g} = \int_{\Omega} g(x)f(x)dx$
 方差 $0 \neq \sigma^2 = \int_{\Omega} (g(x) - \bar{g})^2 f(x)dx < +\infty$
 无偏统计量 $E(|g|) = \int |g(x)|f(x)dx < +\infty$

§1.5 MC 的特点

1. 收敛速度或误差与问题的维数无关

例如， $G = \int_0^1 \cdots \int_0^1 g(x_1, \dots, x_s) dx_1 \cdots dx_s$ S 重积分，它的 MC 估计是：

$$\hat{G}_N = \frac{1}{N} \sum_{n=1}^N g(\vec{X}_n) \quad (1.27)$$

其中 $\vec{X}_n = (\xi_1^{(n)}, \dots, \xi_s^{(n)})$ 为 S 维立方体内均匀分布抽样确定的点。如果 $\sigma^2(|g|)$ 不变, S 的变化 \rightarrow 引起每次观察 C 较小变化。所以 MC 的 ε 与 S 无关, MC 的收敛速度也与 S 无关。

MC 的误差公式: $|\hat{G}_N - G| < \frac{X\sigma}{\sqrt{N}}$, σ 是无偏统计的方差。置信水平 $1 - \alpha$ 确定后, X 也确定。

X	$1-\alpha$
0.6745	50%
1.96	95%
3	99.7%

2. 受问题的条件影响不大

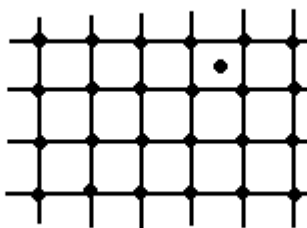
$$G^V = \int \dots \int_V g(x_1, \dots, x_s) dx_1 \dots dx_s \quad (1.28)$$

$$\hat{G}_N^V = \frac{1}{N} \sum_{X_n \in V} g(\vec{X}_n) \quad (1.29)$$

不管 S 维内区域 V 如何特殊, 投点成功几率 $= \frac{n}{N}$ 。

3. 不必进行离散化处理:

数值网络



插值 \Rightarrow 占大量存储单元。MC 是直接抽样, 精度提高了。(random walk)

MC 适用: ①高维 ②节省内存 ③程序方便。

4. 具有直接解决问题的能力

其它方法: 非确定统计性问题 \Rightarrow 确定性问题(方程) \Rightarrow 解。

MC 方法通过模拟原问题的实际过程而得到问题的解决(结果更合理)。

5. MC 便于并行计算

6. 误差容易估计确定 $\varepsilon = |\hat{G}_N - G| = \frac{X\sigma}{\sqrt{N}}$

① X 由置信水平 $1 - \alpha = 95\%$ 确定(查表) $\frac{1}{\sqrt{2\pi}} \int_{-X}^X e^{-\frac{1}{2}t^2} dt = 1 - \alpha$

② N 是实际的抽样数。

③无偏统计量 $g(x)$ 的异零方差 σ^2 , 可由计算给出

$$\hat{\sigma}_N = \left(\frac{1}{N} \sum_{n=1}^N g^2(x_n) - \hat{G}_N^2 \right)^{\frac{1}{2}} \quad (1.30)$$

本章引出三个问题:

① $[0, 1]$ 区间均匀随机数 ξ 的产生方法(独立、均匀、不重复)

乘同余方法, 对任意初值 x_1 , 有

$$x_{n+1} = ax_n \pmod{M} \quad (1.31)$$

$$\xi_{n+1} = \frac{x_{n+1}}{M} \quad (1.32)$$

$$a = \begin{cases} 1 & \pmod{2} & M = 2 \\ 3 & \pmod{4} & M = 4 \\ 5 & \pmod{8} & M = 8 \end{cases} \quad (1.33)$$

推荐 x_1 与 M 互素。

②如何由已知分布进行抽样。

③降低实验方差 σ^2 技巧, 使问题更加确定化。

第二章 由已知分布的抽样

$$G = E(g) = \int g(x)f(x)dx \quad \hat{G}_N = \frac{1}{N} \sum_{n=1}^N g(x_n) \quad (2.1)$$

§2.1 由已知分布的随机抽样

$$f(x) = \frac{dF}{dx} \quad dF(x) = f(x)dx \quad F(x) = \int_a^x f(x')dx' \leq 1 \quad (2.2)$$

一. 含义 是由已知分布的总体中产生简单子样 (分离散分布和连续分布两种)。

x_1, \dots, x_N 表示由总体 $F(x)$ 中产生容量为 N 的简单子样。

二. 特点

1. x_1, \dots, x_N 相互独立;

2. 具有相同分布, 随机抽样在计算机上进行, 最终由 $[0, 1]$ 内的均匀随机数来表示 (函数)。

三. 设法降低抽样费用 (或提高抽样效率)

§2.2 直接抽样方法

1. 离散型分布的直接抽样方法

一般形式:

$$F(x) = \sum_{x_i \leq x} P_i \quad (2.3)$$

其中 x_1, x_2, \dots 为离散型随机变量的跳跃点, P_1, P_2, \dots 为相应的概率。若

$$\sum_{i=1}^{n-1} P_i < \xi \leq \sum_{i=1}^n P_i \quad (2.4)$$

则直接抽样结果为:

$$X_F = x_n \quad (2.5)$$

2. 连续分布的直接抽样方法

$$F(x) = \int_a^x f(x')dx' \quad (2.6)$$

如果分布函数的反函数存在 F^{-1} , 则

$$F(X_F) = \int_a^{X_F} f(x')dx' = \xi \quad 0 < \xi \leq 1 \quad (2.7)$$

所以直接抽样结果为

$$X_F = F^{-1}(\xi) \quad (2.8)$$

3. 举例

例 1. 二项式分布 (离散型), $P(x=n) = P_n = C_N^n p^n (1-p)^{N-n}$, 其中 $0 < p < 1$, 随机变量是整数 n 。若

$$\sum_{i=1}^{n-1} P_i < \xi \leq \sum_{i=1}^n P_i \quad (2.9)$$

则直接抽样结果为

$$X_F = n$$

例 2. 泊松分布, $P(x=n) = P_n = e^{-\lambda} \frac{\lambda^n}{n!}$, 其中 $\lambda > 0$, 随机变量是整数 n , $\sum_{n=0}^{\infty} P_n = 1$ 。若

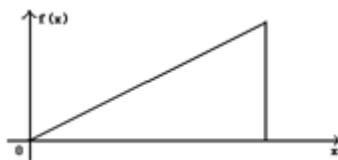
$$\sum_{i=0}^{n-1} \frac{\lambda^i}{i!} < e^\lambda \xi \leq \sum_{i=0}^n \frac{\lambda^i}{i!} \quad (2.10)$$

则直接抽样结果为

$$X_F = n$$

例 3. β -分布 (连续型)

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{else} \end{cases} \quad (2.11)$$



$$\int_0^{X_F} f(x') dx' = \xi \quad (2.12)$$

所以直接抽样结果为

$$X_F = \sqrt{\xi}$$

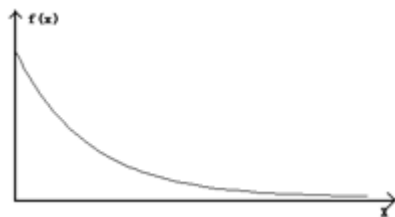
例 4. 倒数分布

$$f(x) = \begin{cases} \frac{1}{\ln \alpha} \cdot \frac{1}{x} & 1 \leq x \leq \alpha \\ 0 & \text{else} \end{cases} \quad (2.13)$$

$$\int_1^{X_F} \frac{1}{\ln \alpha} \frac{1}{x} dx = \frac{\ln X_F}{\ln \alpha} = \xi \rightarrow X_F = e^{\xi \ln \alpha} \quad (2.14)$$

例 5. 指数分布, $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$

$$\begin{aligned} \int_0^{X_F} \lambda e^{-\lambda x} dx &= 1 - e^{-\lambda X_F} = \xi \\ \rightarrow X_F &= -\frac{1}{\lambda} \ln(1 - \xi) = -\frac{1}{\lambda} \ln \xi \end{aligned} \quad (2.15)$$



例 6. 散射方位角余弦分布, $f(x) = \frac{1}{\pi} \cdot \frac{1}{\sqrt{1-x^2}}$, $-1 \leq x \leq 1$

$$\int_{-1}^{X_F} \frac{1}{\pi} \frac{dx}{\sqrt{1-x^2}} = \frac{1}{2\pi} \arcsin X_F = \xi \quad (2.16)$$

所以直接抽样结果为 $X_F = \sin 2\pi\xi$, 或 $X_F = \cos 2\pi\xi$

§2.3 舍选抽样法

对于连续分布的直接抽样方法的困难:

- ① 分布函数无法解析表达出来, 不存在 $F(x)$;
- ② 存在 $F(x)$, 但不存在 $F^{-1}(x)$, 给不出 $F^{-1}(x)$;
- ③ 反函数的计算量过大。

1. 简单分布的舍选法

考虑定义在 $[0, 1]$ 上的分布 $f(x)$, 并假设 $f(x)$ 是有上界的:

$$f(x) \leq M \quad M \geq 1 \quad (2.17)$$

● 对此分布的舍选抽样方法如下:

产生一对 $[0, 1]$ 内的均匀随机数 ξ, ξ' ,

$$M\xi' \leq f(\xi) \begin{cases} \text{yes} \rightarrow X_f = \xi \rightarrow \text{continue} \\ \text{no} \rightarrow \text{continue} \end{cases} \quad (2.18)$$

实质上是在 X 域 ($0 \leq x \leq 1, 0 \leq y \leq M$) 内产生均匀的相互独立的随机点列:

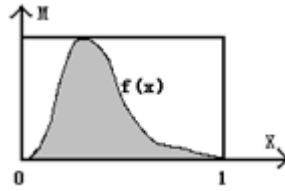
$(\xi_1, M\xi_2), (\xi_3, M\xi_4), \dots, (\xi_{2N-1}, M\xi_{2N})$, 抛弃在 $f(x)$ 之上的所有点, 保留 $f(x)$ 之下的所有点。

● 抽样效率: 是指挑选抽样方法中进行挑选的被选中的概率。

$$E = P(M\xi' \leq f(\xi)) \quad (2.19)$$

因为 $\int_0^1 f(x)dx = 1$, 所以

$$E = \frac{\int_0^1 f(x)dx}{M \cdot 1} = \frac{1}{M} \quad (2.20)$$



2. 乘分布的舍选法

$$f(x) = H(x)f_1(x) \quad (2.21)$$

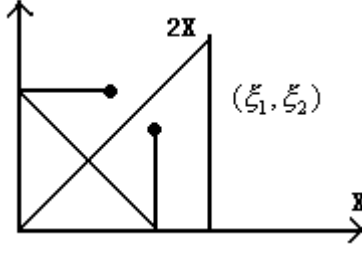
f_1 任意, $H(x) \leq M$, 于是

$$M\xi \leq H(x_{f1}) \begin{cases} \text{yes} \rightarrow X_f = X_{f1} \rightarrow \text{continue} \\ \text{no} \rightarrow \text{continue} \end{cases} \quad (2.22)$$

其中 X_{f1} 由 $f_1(x)$ 抽样。

3. 例子

例 1. β 分布, $f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & else \end{cases}$



$$M = \max[f(x)] = 2, \quad M\xi_2 \leq f(\xi_1) \rightarrow 2\xi_2 \leq 2\xi_1 \begin{cases} yes \rightarrow X_F = \xi_1 \rightarrow continue \\ no \rightarrow continue \end{cases} \quad (2.23)$$

抽样效率 $E = \frac{1}{2} = \frac{1}{M}$, 很显然, $\xi_1 > \xi_2$ 情况没有必要舍弃。 $X_F = \max(\xi_1, \xi_2)$

例 2. 乘分布舍选法例子

Maxwell 分布 $f(x) = \frac{2}{\sqrt{\pi}}\beta^{\frac{3}{2}}\sqrt{x}e^{-\beta x} \quad 0 \leq x$

取

$$f_1(x) = \frac{2\beta}{3}e^{-\frac{2\beta x}{3}} \quad (2.24)$$

$$H(x) = \frac{f(x)}{f_1(x)} = \frac{\frac{2}{\sqrt{\pi}}\beta^{\frac{3}{2}}\sqrt{x}e^{-\beta x}}{\frac{2\beta}{3}e^{-\frac{2\beta x}{3}}} = \frac{3\sqrt{x}}{\sqrt{\pi}}\sqrt{\beta}e^{-\frac{1}{3}\beta x} \quad (2.25)$$

① 先由 $f_1(x)$ 抽样 x_{f1}

$$\begin{aligned} \int_0^{X_{f1}} f_1(x)dx &= \int_0^{X_{f1}} \frac{2\beta}{3}e^{-\frac{2\beta x}{3}}dx = -e^{-\frac{2}{3}\beta X_{f1}} + 1 = \xi_1 \\ \rightarrow e^{-\frac{2}{3}\beta X_{f1}} &= 1 - \xi_1 = \xi_1 \\ \rightarrow X_{f1} &= -\frac{3}{2\beta}\ln\xi_1 \end{aligned} \quad (2.26)$$

② 确定 H 的极大值 M

$$\begin{aligned} H'(x) &= \frac{3}{\sqrt{\pi}}\left\{\frac{\sqrt{\beta}}{2\sqrt{x}}e^{-\frac{1}{3}\beta x} - \frac{1}{3}\beta\sqrt{\beta x}e^{-\frac{1}{3}\beta x}\right\} = 0 \\ \rightarrow M &= H\left(\frac{3}{2\beta}\right) = \frac{3}{\sqrt{\pi}}\sqrt{\frac{3}{2}}e^{-\frac{1}{2}} \end{aligned} \quad (2.27)$$

③ 抽样 X_f

$$\begin{aligned} M\xi_2 &\leq H(x_{f1}) \\ \rightarrow \sqrt{\frac{3}{2}}e^{-\frac{1}{2}}\xi_2 &\leq \sqrt{\beta}\sqrt{X_{f1}}e^{-\frac{1}{3}\beta X_{f1}} \\ \rightarrow \xi_2^2 &\leq -e\xi_1\ln\xi_1 \\ \rightarrow X_f &= X_{f1} = -\frac{3}{2\beta}\ln\xi_1 \end{aligned} \quad (2.28)$$

作业：①

正态分布 $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, 根据乘分布的舍选抽样方法, 选 $f_1(x) = \frac{1}{2} e^{-|x|}$, 则抽样过程为

$$X_{f1} = \ln \xi_1 \rightarrow \frac{1}{2}(X_{f1} + 1)^2 \leq -\ln \xi_2 \rightarrow X_f = \text{sign}(\xi_3 - \frac{1}{2})X_{f1}$$

抽样效率 $E = \sqrt{\frac{\pi}{2}} e^{-\frac{1}{2}} \approx 0.76$

§2.4 复合抽样方法

1. 复合抽样方法的一般形式

$$f(x) = \int_{-\infty}^{+\infty} f_2(x|y) dF_1(y) = \int_{-\infty}^{+\infty} f_2(x|y) f_1(y) dy \quad (2.29)$$

$$\Rightarrow X_f = X_{f2}(x|Y_{f1}) \quad (2.30)$$

其中 Y_{f1} 是由分布 $f_1(y)$ 抽样得到, $X_{f2}(x|Y_{f1})$ 是由分布 $f_2(x|Y_{f1})$ 中抽样确定。

2. 加分布的复合抽样方法

$$f(x) = \sum_{n=1}^{\infty} P_n f^{(n)}(x) \quad (2.31)$$

$$\int_{-\infty}^{+\infty} f(x) dx = \sum_{n=1}^{\infty} P_n \int_{-\infty}^{+\infty} f^{(n)}(x) dx = \sum_{n=1}^{\infty} P_n = 1 \quad (2.32)$$

$$P_n \geq 0, \quad \sum_{n=1}^{\infty} P_n = 1 \quad (2.33)$$

$f^{(n)}(x)$ 为与参数 n 有关的分布密度函数。

若

$$\sum_{i=1}^{n-1} P_i < \xi \leq \sum_{i=1}^n P_i \quad (2.34)$$

则取

$$F_1(x) = \sum_{n < x} P_n, \quad X_f = X_f^{(n)} \quad (2.35)$$

其中 $X_f^{(n)}$ 由 $f^{(n)}(x)$ 中抽样。

3. 举例

例 1. 环带均匀分布, $f(R) = \frac{2R}{R_1^2 - R_0^2}$, $R_0 \leq R \leq R_1$,

引入变换 $x = \frac{R - R_0}{(R_1 - R_0)^2}$, 则

$$f(x) = \frac{R_1 - R_0}{R_1 + R_0} \cdot 2x + \frac{2R_0}{R_1 + R_0} \cdot 1 \quad (2.36)$$

其中 $\frac{R_1 - R_0}{R_1 + R_0} = P_1$, $2x = f^{(1)}(x)$, $\frac{2R_0}{R_1 + R_0} = P_2$, $1 = f^{(2)}(x)$, $P_1 + P_2 = 1$

于是, 根据加分布的复合抽样方法有:

$$X_f = X_f^{(2)}, \quad \sum_{i=1}^1 P_i < \xi \leq \sum_{i=1}^2 P_i = 1 \quad (2.37)$$

$$\frac{R_1 - R_0}{R_1 + R_0} < \xi_1 \Rightarrow R_f = R_0 + \xi_2(R_1 - R_0) \quad (2.38)$$

例 2. 指数函数分布 $E_n(x) = \int_1^\infty \frac{ne^{-xy}}{y^n} dy, \quad 0 \leq x,$

引入如下两个分布:

$$f_1(y) = ny^{-n-1}, \quad 1 \leq y \quad (2.39)$$

$$f_2(x|y) = ye^{-xy}, \quad 0 \leq x \quad (2.40)$$

则显然有如下等式:

$$E_n(x) = \int_1^\infty f_2(x|y)f_1(y)dy \quad (2.41)$$

$$X_f = X_{f2}(x|Y_{f1}) \quad (2.42)$$

$$\begin{aligned} \int_1^{Y_{f1}} f_1(y)dy &= \int_1^{Y_{f1}} ny^{-n-1}dy = -Y_{f1}^{-n} + 1 = \xi_1 \\ \rightarrow Y_{f1}^{-n} &= 1 - \xi_1 = \xi_1 \\ \rightarrow Y_{f1} &= \xi_1^{-\frac{1}{n}} \end{aligned} \quad (2.43)$$

$$\begin{aligned} \int_0^{X_{f2}} f_2(x|Y_{f1})dx &= \int_0^{X_{f2}} ye^{-xy}dx \Big|_{y=Y_{f1}} = -e^{-Y_{f1}X_{f2}} + 1 = \xi_2 \\ \rightarrow e^{-Y_{f1}X_{f2}} &= 1 - \xi_2 = \xi_2 \\ \rightarrow X_{f2} &= -\frac{1}{Y_{f1}} \ln \xi_2 = -\xi_1^{\frac{1}{n}} \ln \xi_2 \end{aligned} \quad (2.44)$$

所以

$$X_f = X_{f2}(x|Y_{f1}) = -\xi_1^{\frac{1}{n}} \ln \xi_2 \quad (2.45)$$

作业: ②

正态分布 $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, 引入

$$f_1(y) = ye^{-\frac{y^2}{2}}, \quad 0 \leq y \quad (2.46)$$

$$f_2(x) = \frac{1}{\pi} \frac{1}{\sqrt{1-x^2}}, \quad -1 \leq x \leq 1 \quad (2.47)$$

则

$$f(x) = \int_{-\infty}^{\infty} \frac{1}{|y|} f_2\left(\frac{x}{y}\right) f_1(y) dy \quad (2.48)$$

$$X_f = \sqrt{-2 \ln \xi_1} \cos 2\pi \xi_2 \quad \text{or} \quad \sqrt{-2 \ln \xi_1} \sin 2\pi \xi_2 \quad (2.49)$$

§ 2.2 直接法:

① 离散法;

② 反函数法, $x_f = F^{-1}(\xi), F(X_f) = \int_a^{X_f} f(x)dx = \xi$

§ 2.3 舍选法:

① 简单分布 $f(x), M = \max f(x)$, 产生一对 $(\xi, \xi'), x \in [0, 1]$

② 乘分布 $f(x) = H(x)f_1(x), \quad M\xi' \leq f(\xi) \Rightarrow X_f = \xi$

§ 2.4 复合法:

①一般 $f(x) = \int_{-\infty}^{\infty} f_2(x|y)f_1(y)dy$

②加分布复合法 $X_f = X_{f_2}(x|Y_{f_1})$

③复合舍选法

乘分布舍选法证明: 设 $f_0(x, y)$ 为任意二维分布密度函数, $H(x)$ 为任意函数, 则对于如下积分形式的分布

$$f(x) = \frac{\int_0^{H(x)} f_0(x, y)dy}{\int_0^{\infty} dx \int_0^{H(x)} f_0(x, y)dy} \quad (2.50)$$

则满足 $\int_0^{\infty} f(x)dx = 1$ 有舍选抽样方法:

$$Y_{f_0} \leq H(X_{f_0}) \Rightarrow X_f = X_{f_0} \quad (2.51)$$

其中 y 的变化不能大于 $H(x)$ 。

证明, 对于任意的 x 有

$$\begin{aligned} P(x \leq X_f < x + dx) &= P(x \leq X_{f_0} < x + dx | Y_{f_0} \leq H(X_{f_0})) \\ &= \frac{P(x \leq X_{f_0} \leq x + dx, Y_{f_0} \leq H(X_{f_0}))}{P(Y_{f_0} \leq H(X_{f_0}))} \\ &= \frac{\int_0^{H(x)} f_0(x, y)dy}{\int_0^{\infty} dx \int_0^{H(x)} dy f_0(x, y)} \end{aligned} \quad (2.52)$$

对乘分布的舍选法抽样: $f(x) = H(x)f_1(x)$, 令

$$f_0(x, y) = \begin{cases} \frac{1}{M}f_1(x) & 0 \leq y \leq M \\ 0 & else \end{cases} \quad (2.53)$$

其中 M 为 $H(x)$ 的上界,

$$M\xi \leq H(X_{f_1}) \Rightarrow X_f = X_{f_1} \quad (2.54)$$

§2.5 变换抽样方法

●随机变量 x 的密度函数 $f_1(x)$, 用以下方式将 x 变换为随机变量 y , 设 $y = \varphi(x)$, 则 y 的密度函数为

$$f(y)|dy| = f_1(x)|dx| \quad (2.55)$$

$$f(y) = f_1(x) \left| \frac{dx}{dy} \right| = f_1(\psi(y)) \left| \frac{dx}{dy} \right| \quad (2.56)$$

式中 $\psi(y) = \varphi^{-1}(y) = x$, 即 $\psi(y)$ 为 $\varphi(x)$ 的反函数。

●变换抽样方法的一般过程是:

为了由分布 $f(y)$ 中抽样产生 y , 可先由分布 $f_1(x)$ 中抽样产生 X_{f_1} , 然后通过变换 $X_f = \varphi(X_{f_1})$ 得到。

例: 用变换法产生标准正态随机数 $f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ 。

解: 设两个独立的、 $(0, 1)$ 区间内均匀的随机数 ξ_1, ξ_2 , 将它们做如下变换:

$$x_1 = \sqrt{-2 \ln \xi_1} \cos 2\pi \xi_2 \quad x_2 = \sqrt{-2 \ln \xi_1} \sin 2\pi \xi_2 \quad (2.57)$$

其反变换为

$$\xi_1 = e^{-\frac{1}{2}(x_1^2+x_2^2)} \quad \xi_2 = \frac{1}{2\pi}[\arctan(\frac{x_2}{x_1}) + c] \quad (c = \text{constant}) \quad (2.58)$$

可导出 (x_1, x_2) 的密度

$$f(x_1, x_2) = f[\psi_1(x_1, x_2), \psi_2(x_1, x_2)]|J| \quad (2.59)$$

雅可比行列式

$$J = \begin{vmatrix} \frac{\partial \xi_1}{\partial x_1} & \frac{\partial \xi_1}{\partial x_2} \\ \frac{\partial \xi_2}{\partial x_1} & \frac{\partial \xi_2}{\partial x_2} \end{vmatrix} = \begin{bmatrix} \frac{\partial \xi_1}{\partial x_1} & \frac{\partial \xi_2}{\partial x_2} \\ \frac{\partial \xi_2}{\partial x_1} & \frac{\partial \xi_1}{\partial x_2} \end{bmatrix} \quad (2.60)$$

$$\frac{\partial \xi_1}{\partial x_1} = -x_1 e^{-\frac{1}{2}(x_1^2+x_2^2)} \quad (2.61)$$

$$\frac{\partial \xi_1}{\partial x_2} = -x_2 e^{-\frac{1}{2}(x_1^2+x_2^2)} \quad (2.62)$$

$$\frac{\partial \xi_2}{\partial x_2} = \frac{1}{2\pi} \cdot \frac{\frac{1}{x_1}}{1 + (\frac{x_2}{x_1})^2} = \frac{1}{2\pi} \cdot \frac{x_1}{x_1^2 + x_2^2} \quad (2.63)$$

$$\frac{\partial \xi_2}{\partial x_1} = \frac{1}{2\pi} \cdot \frac{-\frac{x_2}{x_1^2}}{1 + (\frac{x_2}{x_1})^2} = -\frac{1}{2\pi} \cdot \frac{x_2}{x_1^2 + x_2^2} \quad (2.64)$$

$$|J| = \left| -\frac{e^{-\frac{1}{2}(x_1^2+x_2^2)}}{2\pi} \cdot \frac{x_1^2}{x_1^2 + x_2^2} - \frac{e^{-\frac{1}{2}(x_1^2+x_2^2)}}{2\pi} \cdot \frac{x_2^2}{x_1^2 + x_2^2} \right| = \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2+x_2^2)} \quad (2.65)$$

所以

$$f(x_1, x_2) = \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2+x_2^2)} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_1^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x_2^2}{2}} = f(x_1)f(x_2) \quad (2.66)$$

§2.6 近似抽样方法

1. 动机：对舍选、复合、变换抽样方法（直接抽样方法）

① 抽样可以实现，但效率很低；

② 能实现， E 也高，但运算量过大。

2. 一般原理，取一级近似 $f_a(x) \approx f(x)$ ，然后对 $f_a(x)$ 抽样来代替原来分布的抽样。

(1) 线性近似（插值）

$$f_a(x) = f_{i-1} + \frac{x - x_{i-1}}{x_i - x_{i-1}} \cdot (f_i - f_{i-1}) \quad x_{i-1} < x \leq x_i \quad (2.67)$$

其中 x_0, \dots, x_n 为任意点， f_0, \dots, f_n 为相应的密度函数值。

如果 $F_a(x_{i-1}) < \xi \leq F_a(x_i)$ 线性近似分布抽样为：

$$X_{fa} = x_{i-1} + \frac{\xi - F_a(x_{i-1})}{F_a(x_i) - F_a(x_{i-1})}(x_i - x_{i-1}) \quad (2.68)$$

(2) 由连续分布直接抽样，对反函数 $F^{-1}(\xi)$ 做某种近似， $F^{-1}(\xi)$ 代替 $X_f = F^{-1}(\xi)$

例子：正态分布 $F(X_f) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{X_f} e^{-\frac{1}{2}x^2} dx = \xi$

现对分布函数的反函数进行有理逼近, 令 $t = \xi - 0.5, y = \sqrt{-2 \ln t}$, 则

$$X_f = \text{sign}(t) \left\{ y - \frac{a_0 + a_1 y + a_2 y^2}{1 + b_1 y + b_2 y^2 + b_3 y^3} \right\} \quad (2.69)$$

其中 $a_0 = 2.515517, a_1 = 0.802853, a_2 = 0.010328$

$b_1 = 1.432788, b_2 = 0.189269, b_3 = 0.001308$

(3) 渐近分布 $f_n(x)$ 的极限分布为已知分布 $f(x)$, 当 n 足够大, 用 $f_n(x)$ 的简单子样作为原分布的简单子样。

例子: 正态分布,

根据概率统计理论, 随机变量

$$X_{fn} = \frac{\frac{1}{n} \sum_{i=1}^n \xi_i - \frac{1}{2}}{\sqrt{\frac{1}{12n}}} \quad (2.70)$$

是渐近正态分布 $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ 的随机变量。在实际应用中, 取 $n = 4 \sim 12$ 可得到满意的结果。特别当 $n=12$ 时,

$$X_{f_{12}} = \sum_{i=1}^{12} \xi_i - 6 = \sum_{i=1}^{12} \left(\xi_i - \frac{1}{2} \right) = \sum_{i=1}^6 (\xi_{2i} - \xi_{2i-1}) \quad (2.71)$$

§2.7 若干重要分布的随机抽样

习题 1: 复合抽样的三个引理: 已知 f_1 和 f_2 是两个分布密度函数, 求证:

(1) 对于分布 $f(x) = \int f_2(x \mp y) f_1(y) dy$ 的抽样方法为: $X_f = X_{f_2} \pm Y_{f_1}$

(2) 对于分布 $f(x) = \int \frac{1}{|y|} f_2\left(\frac{x}{y}\right) f_1(y) dy$, 则: $X_f = X_{f_2} X_{f_1}$

(3) 对于分布 $f(x) = \int |y| f_2(xy) f_1(y) dy$, 则: $X_f = \frac{X_{f_2}}{Y_{f_1}}$

习题 2: 非弹性散射能分布如下:

$$f(x) = \begin{cases} a^2 x e^{-ax} & 0 \leq x \\ 0 & \text{else} \end{cases} \quad (2.72)$$

引入如下分布

$$f_1 = \begin{cases} a e^{-ax} & 0 \leq x \\ 0 & \text{else} \end{cases} \quad (2.73)$$

对于 $f_2(x)$ 选成同样分布, 则

$$f(x) = \int_{-\infty}^{\infty} f_2(x-y) f_1(y) dy \quad (2.74)$$

因此有如下复合分布抽样方法

$$X_f = -\frac{1}{a} \ln(\xi_1 \xi_2) \quad (2.75)$$

更一般的情况是

$$X_f = -\frac{1}{a} \ln(\xi_1 \xi_2 \dots \xi_n) \quad (2.76)$$

它所遵守的分布是

$$f(x) = \begin{cases} \frac{a^n}{(n-1)!} x^{(n-1)} e^{-ax} & 0 \leq x \\ 0 & else \end{cases} \quad (2.77)$$

习题 3：对数分布的一般形式如下：

$$f(x) = \begin{cases} -\ln x & 0 \leq x < 1 \\ 0 & else \end{cases} \quad (2.78)$$

可以表示成复合分布形式

$$f(x) = \int_{-\infty}^{+\infty} \frac{1}{|y|} f_2\left(\frac{x}{y}\right) f_1(y) dy \quad (2.79)$$

①问 f_1 和 f_2 选为何分布时，对数分布的复合抽样如下：

$$X_f = \xi_1 \xi_2 \quad (2.80)$$

②它的更一般形式如下：

$$f(x) = \begin{cases} \frac{(-\ln x)^n}{n!} & 0 \leq x < 1 \\ 0 & else \end{cases} \quad (2.81)$$

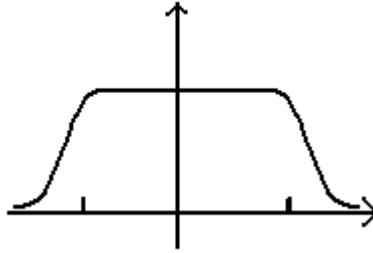
有如下复合抽样方法：

$$X_f = \xi_1 \xi_2 \dots \xi_{n+1} \quad (2.82)$$

试证明之。

习题 4：考虑正态分布 $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ ，利用乘抽样方法，选

$$f_1(x) = \begin{cases} \frac{1}{4\lambda} & -\lambda \leq x < \lambda \\ \frac{\lambda}{4x^2} & else \end{cases} \quad (2.83)$$



抽样效率为 $E = \frac{1}{4} \sqrt{e\pi} \approx 0.730$, $\lambda = \sqrt{\frac{2}{e}}$

习题 5：广义复合舍选抽样法

$$f(x) = \int_{-\infty}^{\infty} H(x, y) f_2(x|y) f_1(y) dy \quad (2.84)$$

这里 $f_2(x|y)$ 为条件分布密度函数， $f_1(y)$ 是一分布密度函数， $H(x, y)$ 的上界为 M ，其抽样过程为：

$$M\xi \leq H(X_{f_2}(x|Y_{f_1}), Y_{f_1}) \begin{cases} yes \rightarrow X_f = X_{f_2}(x|Y_{f_1}) \rightarrow continue \\ no \rightarrow continue \end{cases} \quad (2.85)$$

试用概率法证明其正确性。

典型例子 1. 用坐标变换法对正态分布直接随机抽样 $f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$, x, y 独立,

$$f(x, y) = f(x)f(y) = \frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)} \quad (2.86)$$

极坐标: $x = r \sin \theta$ $y = r \cos \theta$

$$f(x, y)dxdy = |J|f(r, \theta)drd\theta = f(r, \theta)rdrd\theta \quad (2.87)$$

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \sin \theta & r \cos \theta \\ \cos \theta & -r \sin \theta \end{vmatrix} = r \quad (2.88)$$

$$\begin{aligned} \int_{-\infty}^{X_f} \int_{-\infty}^{Y_f} f(x, y)dxdy &= \int_0^{R_f} \int_0^{\theta_f} f(r, \theta)rdrd\theta \\ &= \int_0^{R_f} \int_0^{\theta_f} \frac{1}{2\pi}e^{-\frac{1}{2}r^2}rdrd\theta \\ &= \int_0^{R_f} e^{-\frac{1}{2}r^2}d\left(\frac{r^2}{2}\right) \int_0^{\theta_f} \frac{d\theta}{2\pi} \\ &= \xi \\ &= \xi_1\xi_2 \end{aligned} \quad (2.89)$$

典型例子 2. 各向同性散射角余弦分布 $x = \cos \alpha$ 遵从如下分布

$$f(x) = \begin{cases} \frac{1}{2} & -1 \leq x \leq 1 \\ 0 & else \end{cases} \quad (2.90)$$

方法一. 直接抽样方法如下:

$$\int_{-1}^{X_f} \frac{1}{2}dx = \xi \rightarrow \cos \alpha = 2\xi - 1 \rightarrow \sin \alpha = (1 - \cos^2 \alpha)^{\frac{1}{2}} \quad (2.91)$$

方法二. 乘分布抽样方法

$f(x) = H(x)f_1(x)$, 其中

$$f_1(x) = \begin{cases} \frac{1}{2} \cdot \frac{1}{(1+|x|)\sqrt{1-x^2}} & -1 \leq x \leq 1 \\ 0 & else \end{cases} \quad (2.92)$$

$$H(x) = (1 + |x|)\sqrt{1 - x^2}, \quad -1 \leq x \leq 1 \quad (2.93)$$

在所考虑的区域上, $H(x)$ 上界 $M = \frac{3}{4}\sqrt{3}$

$$M\xi \leq H(X_{f1}) \begin{cases} yes \rightarrow X_f = X_{f1} \rightarrow continue \\ no \rightarrow continue \end{cases} \quad (2.94)$$

第三章 MC 的一般技巧

§3.1 降低实验方差的特性

MC 的误差公式:

$$\varepsilon = X_\alpha \sigma / \sqrt{N} \quad (3.1)$$

常数 X_α 由置信水平 $1 - \alpha$ 确定。现讨论 MC 方法改进的一些方向。

利用非独立随机变数序列: 简单子样 - 独立同分布 \Rightarrow 非独立但同分布

1. 使 MC 估计 $\bar{\theta}_N$ 依概率收敛于真值 Θ , 不需要随机变量间相互独立, 依照马尔科夫 (Markov) 定理, 只要随机变数 $\theta(\omega_1), \theta(\omega_2), \dots, \theta(\omega_N)$ 满足

$$\sigma^2\left(\frac{1}{N} \sum_{n=1}^N \theta(\omega_n)\right) \rightarrow 0 \quad (3.2)$$

对于任意正数 $\varepsilon > 0$ 就有

$$P\left(\left|\frac{1}{N} \sum_{n=1}^N \theta(\omega_n) - \Theta\right| < \varepsilon\right) \rightarrow 1 \quad (3.3)$$

因此, 序列 $\{\theta(\omega_n), n = 1, 2, \dots\}$ 满足 (3.2) 式, 那么 $\bar{\theta}_N$ 总能依概率收敛到真值。

2. 其次, 根据切贝雪夫不等式, 有

$$P(|\bar{\theta}_N - \Theta| < \varepsilon' \sigma(\bar{\theta}_N)) \geq 1 - \frac{1}{\varepsilon'} \quad (3.4)$$

如果令 $\varepsilon' = 1/\alpha$ 则有 $P(|\bar{\theta}_N - \Theta| < \sigma(\bar{\theta}_N)/\alpha) \geq 1 - \alpha$

因此在一定置信水平下, 误差直接取决于 $\bar{\theta}_N$ 的均方差 $\sigma(\bar{\theta}_N)$ 的大小。

3. 设 $\sigma^2(\theta(\omega_n)) = \sigma^2, n = 1, 2, \dots,$

$$\text{cov}(\theta(\omega_m), \theta(\omega_n)) = \rho_{mn} \sigma(\theta(\omega_m)) \sigma(\theta(\omega_n)) = \rho_{mn} \sigma^2 \quad (3.5)$$

于是

$$\begin{aligned} \sigma^2(\bar{\theta}_N) &= \frac{1}{N^2} \left(\sum_{m=1}^N \sum_{n=1}^N \text{cov}(\theta(\omega_m), \theta(\omega_n)) \right) \\ &= \frac{\sigma^2}{N} \left(1 + \sum_{m \neq n}^N \rho_{mn} \right) \end{aligned} \quad (3.6)$$

相互独立序列, $\rho_{mn} = 0$, 则

$$\sigma^2(\bar{\theta}_N) = \frac{1}{N} \sigma^2 \quad (3.7)$$

若不独立且使 $\sum_{m \neq n} \rho_{mn} < 0$ 则

$$\sigma^2(\bar{\theta}_N) < \frac{\sigma^2}{N} \quad (3.8)$$

§3.2 期望估计技巧

欲求积分

$$\theta = \int \int_{V_2} g(x, y) f(x, y) dx dy \quad (3.9)$$

其中 $f(x, y)$ 是 V_2 上联合密度函数, 即 $f(x, y) \geq 0$, $\int \int_{V_2} f(x, y) dx dy = 1$, 令

$$f_1(x) = \int f(x, y) dy \quad \text{边缘分布} \quad (3.10)$$

$$f_2(y|x) = \frac{f(x, y)}{f_1(x)} \quad \text{条件分布} \quad (3.11)$$

又令

$$\theta_x = \int g(x, y) f_2(y|x) dy \Rightarrow \theta = \int \theta_x f_1(x) dx \quad (3.12)$$

$$\sigma_x^2 = \int (g(x, y) - \theta_x)^2 f_2(y|x) dy \quad (3.13)$$

则由 $f(x, y)$ 抽样, 以 $g(x, y)$ 为估计的方差

$$\begin{aligned} \sigma_g^2 &= \int \int_{V_2} [g(x, y) - \theta]^2 f(x, y) dx dy \\ &= \int \int_{V_2} [(g - \theta_x) + (\theta_x - \theta)]^2 f(x, y) dx dy \\ &= \int \int_{V_2} (g - \theta_x)^2 f(x, y) dx dy + \int \int_{V_2} (\theta_x - \theta)^2 f(x, y) dx dy \\ &\quad + 2 \int \int_{V_2} (g - \theta_x)(\theta_x - \theta) f(x, y) dx dy \end{aligned} \quad (3.14)$$

等号右边第一项等于

$$\int \int_{V_2} (g - \theta_x)^2 f_2(y|x) f_1(x) dx dy = \int f_1(x) dx \int (g - \theta_x)^2 f_2(y|x) dy = \int \sigma_x^2 f_1(x) dx \quad (3.15)$$

第二项等于

$$\int (\theta_x - \theta)^2 dx \int f(x, y) dy = \int (\theta_x - \theta)^2 f_1(x) dx \quad (3.16)$$

第三项等于

$$\int \int g(x, y) \theta_x f(x, y) dx dy - \int \int \theta_x^2 f(x, y) dx dy + \theta \int \int \theta_x f(x, y) dx dy - \theta^2 \quad (3.17)$$

$$\begin{aligned} \int \int g(x, y) \theta_x f(x, y) dx dy &= \int \int g(x, y) \theta_x f_2(y|x) f_1(x) dx dy \\ &= \int \theta_x f_1(x) dx \int g(x, y) f_2(y|x) dy \end{aligned}$$

$$= \int \theta_x^2 f_1(x) dx \quad (3.18)$$

$$\int \theta_x^2 dx \int f(x, y) dy = \int \theta_x^2 f_1(x) dx \quad (3.19)$$

$$\theta \int \int \theta_x f(x, y) dx dy = \theta \int \theta_x f_1(x, y) dx dy = \theta^2 \quad (3.20)$$

所以

$$\sigma_g^2 = \int \sigma_x^2 f_1(x) dx + \int (\theta_x - \theta)^2 f_1(x) dx \quad (3.21)$$

另一方面, 用 MC 方法, 以 θ_x 为估计量, 其方差为

$$\sigma^2 = \int (\theta_x - \theta)^2 f_1(x) dx \quad (3.22)$$

显然, $\sigma_g^2 \geq \sigma^2$, 用解析 (或数值) 方法代替一次随机抽样, 就能导致方差的减小, 也称期望估计技巧。

例 1. 考虑积分 $\frac{\pi}{4} = \int_0^1 \int_0^1 g(x, y) dx dy$, 这里

$$g(x, y) = \begin{cases} 1 & x^2 + y^2 \leq 1 \\ 0 & x^2 + y^2 > 1 \end{cases} \quad (3.23)$$

$$\sigma_g^2 = \frac{\pi}{4} - \left(\frac{\pi}{4}\right)^2 = 0.168 \quad (3.24)$$

先求出一个积分 $y = \sqrt{1-x^2}$ $G = \int_0^1 \sqrt{1-x^2} dx$,

$$\sigma_{expect}^2 = \int_0^1 (1-x^2) dx - \left(\frac{\pi}{4}\right)^2 = 0.050 \quad (3.25)$$

§3.3 重要抽样技巧

$$G = \int g(\vec{X}) f(\vec{X}) d\vec{X} \quad (3.26)$$

引入新的密度函数 $f^*(\vec{X})$, 当 $f(\vec{X}) \neq 0$ 时, $f^*(\vec{X}) \neq 0$ 则上述 n 重积分可以表示形式如下:

$$\begin{aligned} G &= \int g(\vec{X}) \frac{f(\vec{X})}{f^*(\vec{X})} f^*(\vec{X}) d\vec{X} \\ &= \int \int g_{IS}(\vec{X}) f^*(\vec{X}) d\vec{X} \end{aligned} \quad (3.27)$$

1. 重要抽样技巧是: 依 $g(\vec{X})f(\vec{X})$ 的重要性确定合适的密度函数 $f^*(\vec{X})$, 使 $\sigma^2(g_{IS}(\vec{X})|f^*(\vec{X}))$ 尽量小。

2. 重要抽样的方差

$$\sigma^2(g_{IS}(\vec{X})|f^*(\vec{X})) = \int \int g^2(\vec{X}) \frac{f^2(\vec{X})}{f^*(\vec{X})} d\vec{X} - \left(\int \int g(x, y) f(x, y) dx dy\right)^2 \quad (3.28)$$

用拉格朗日乘子法确定最佳的 $f^*(\vec{X})$ 由下式的极小值来确定

$$L\{f^*\} = \left[\int \frac{g^2(\vec{X}) f^2(\vec{X})}{f^*(\vec{X})} dx + \lambda \int f^*(\vec{X}) d\vec{X} \right] = \min \quad (3.29)$$

$$\frac{\delta}{\delta f^*}[\dots] = 0 \quad (3.30)$$

完成泛函微分得到

$$-\frac{g^2(\vec{X})f^2(\vec{X})}{f^{*2}(\vec{X})} + \lambda = 0 \quad (3.31)$$

则 $f^*(\vec{X}) = \lambda' |g(\vec{X})f(\vec{X})|$ λ' 由 $\int f^*(\vec{X})d\vec{X} = 1$ 来确定, 假设 $g(\vec{X}) \geq 0$, 则

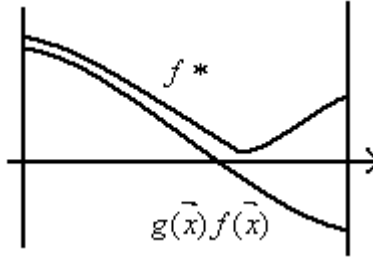
$$1 = \int f^*(\vec{X})d\vec{X} = \lambda' \int g(\vec{X})f(\vec{X})d\vec{X} \quad (3.32)$$

所以

$$\lambda' = \frac{1}{\int g(\vec{X})f(\vec{X})d\vec{X}} \quad (3.33)$$

故

$$f^*(\vec{X}) = \frac{g(\vec{X})f(\vec{X})}{\int g(\vec{X})f(\vec{X})d\vec{X}} = \frac{g(\vec{X})f(\vec{X})}{G} \quad (3.34)$$



3. MC 法求 $G = \int g(\vec{X})f(\vec{X})d\vec{X} = \int [\frac{g(\vec{x})f(\vec{x})}{f^*(\vec{x})}]f^*(\vec{x})d\vec{x}$ 或为

$$\hat{G}_N = \frac{1}{N} \sum_{n=1}^N \frac{g(X_{f^*}^{(n)})f(X_{f^*}^{(n)})}{f^*(X_{f^*}^{(n)})} = \frac{1}{N} \sum_{n=1}^N \frac{g(X_{f^*}^{(n)})f(X_{f^*}^{(n)})}{g(X_{f^*}^{(n)})f(X_{f^*}^{(n)})/G} = \frac{1}{N} \sum_{n=1}^N G = G \quad (3.35)$$

4. 应用举例

例 1. $G = \int_0^1 \cos(\frac{\pi x}{2})dx$ $\hat{G}_N = \frac{1}{N} \sum_{n=1}^N \cos(\frac{\pi \xi_n}{2})$ $f(x) = 1$ 在 $(0,1)$ 均匀抽样, 求和 $g_1(x) = \cos(\frac{\pi x}{2})$

$$\sigma_{g_1}^2 = \int_0^1 \cos^2(\frac{\pi x}{2})dx - G^2 = 0.0947$$

用重要抽样技巧 $\cos(\frac{\pi x}{2}) = 1 - \frac{\pi^2}{8}x^2 + \frac{\pi^4 x^4}{2^4 4!} - \dots$

令 $f^*(x) = \frac{3}{2}(1-x^2)$, $f^* \geq 0$, $\int_0^1 f^*(x)dx = 1$

$$G = \int \frac{gf}{f^*} f^* dx = \int g^* f^* dx$$

$$g^* = \frac{g_1}{f^*} = \frac{2\cos(\frac{\pi x}{2})}{3(1-x^2)} \quad \sigma^2(g^*) = 0.000990$$

例 2. $\int_0^1 \sqrt{1-x^2}dx = \frac{\pi}{4}$

一般方法: $g_1 = \sqrt{1-x^2}$, $f_1 = 1$, $\sigma^2(g_1) = 0.050$

重要抽样方法: $g(x) = 1 - \frac{1}{2}x^2 + \dots$

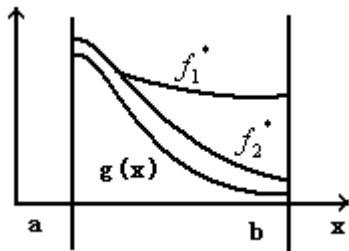
$$f^* = c(1 - \beta x^2) \quad \int_0^1 c(1 - \beta x^2)dx = c(1 - \frac{1}{3}\beta) = 1$$

$$f^* = \frac{1 - \beta x^2}{1 - \frac{1}{3}\beta}$$

β 待定, 由 $\sigma^2(g^*)$ 极小来确定, $g^* = \frac{g_1}{f^*} = (1 - \frac{1}{3}\beta) \frac{\sqrt{1-x^2}}{1-\beta x^2}$

$$\sigma^2(g^*) = (1 - \frac{1}{3}\beta) [\frac{1}{\beta} - \frac{1-\beta}{\beta\sqrt{\beta}} \arctan(\sqrt{\beta})] - (\frac{\pi}{4})^2$$

$\frac{\partial \sigma^2(g^*)}{\partial \beta} = 0$ 数值求解, $\beta = 0.74$, $\sigma^2(g^*) = 0.0029$, 不妨选 $\beta = \frac{1}{2}$, $\sigma^2(g_{\beta=\frac{1}{2}}^*) = 0.011$
 $G = \int_a^b g(x)dx$, f_2^* 好于 f_1^* (泰勒展开)



例 3. 如果 $\int g(\bar{x})f(\bar{x})d\bar{x}$ 是单值的, 但 $\sigma^2(g(x)|f(x))$ 不存在; 但可以证明 $g(x)f(x)/f^*(x)$ 的方差存在。

$$G = \int_0^1 \frac{dx}{x^{\frac{1}{2}}} \quad g(x) = \frac{1}{\sqrt{x}}$$

则

$$\langle g^2 \rangle = \int_0^1 g^2 dx = \int_0^1 \frac{1}{x} dx = \infty$$

我们试用函数 $f^*(x) = (1-r)x^{-r}$, $r < 1$, 先求 G

则

$$g^* = \frac{gf}{f^*} = \frac{x^{r-\frac{1}{2}}}{1-r}, \quad f = 1$$

g^* 的 n 次矩

$$\langle g^{*n} \rangle = (1-r)^{-n+1} \int_0^1 x^{nr-\frac{n}{2}} x^{-r} dx$$

此积分若存在, 需 $(n-1)r - \frac{n}{2} > -1$, 实际上所有矩存在, 为 $\frac{1}{2} \leq r < 1$, 最佳的 $r = \frac{1}{2}$

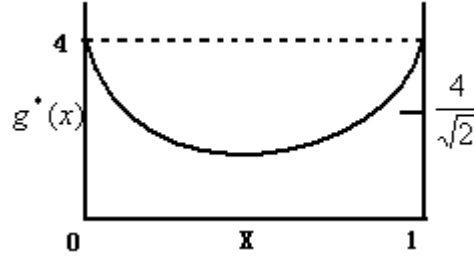
例 4. $G = \int_0^1 \frac{dx}{[x(1-x)]^{\frac{1}{2}}}$, $\sigma^2(g) \rightarrow \infty$,

引入

$$f^*(x) = \frac{1}{4\sqrt{x}} + \frac{1}{4\sqrt{1-x}}, \quad 0 \leq x \leq 1$$

$$g^*(x) = \frac{g}{f^*} = \frac{\frac{1}{\sqrt{x(1-x)}}}{\frac{1}{4\sqrt{x}} + \frac{1}{4\sqrt{1-x}}} = \frac{4}{\sqrt{x} + \sqrt{1-x}}$$

$$\sigma^2(g^*) = \int (g^* - G)^2 f^*(x) < (4 - \frac{4}{\sqrt{2}})^2 = 1.37$$



§3.4 分区间抽样技巧 (Stratification method)

● 要点:

- (1) 将抽样区间分成若干子区间, 对积分值贡献大的抽样更多的出现; \Leftarrow (与重要抽样一样)
- (2) 不改变原来的概率分布; \Leftarrow (与重要抽样不同)

● 作法: $I = \int_0^1 f(x)dx$

将区间 $[0, 1]$ 用分点 a_i ($i = 0, 1, 2, \dots, m$) 分成 m 个互不相交的子区间 (不一定等分), 每个子区间的长度 $l_i = a_i - a_{i-1}$ ($i = 1, 2, \dots, m; a_0 = 0, a_m = 1$), l_1, l_2, \dots, l_m 不一定等区间分。于是

$$I = \int_0^1 f(x)dx = \sum_{i=1}^m \int_{a_{i-1}}^{a_i} f(x)dx = \sum_{i=1}^m I_i \quad (3.36)$$

- 在每个小区间 $[a_{i-1}, a_i]$ 用 MC 法直接求 I_i ($i = 1, 2, \dots, m$)
- 在 $[a_{i-1}, a_i]$ 长度为 l_i 上产生 n_i 个均匀随机数 ξ_j ($j = 1, 2, \dots, n_i$). 则

$$x_j^{(i)} = a_{i-1} + l_i \xi_j \quad (j = 1, 2, \dots, n_i; \quad \xi_j \in (0, 1)) \quad (3.37)$$

$$\begin{aligned} I_i &= \int_{a_{i-1}}^{a_i} f(x)dx = l_i \int_{a_{i-1}}^{a_i} \frac{f(x)}{l_i} dx \\ &= \frac{l_i}{n_i} \sum_{j=1}^{n_i} f(x_j^{(i)}) = \theta^{(i)} \end{aligned} \quad (3.38)$$

$$I = \sum_{i=1}^m \theta^{(i)} = \sum_{i=1}^m \frac{l_i}{n_i} \sum_{j=1}^{n_i} f(x_j^{(i)}) \quad (3.39)$$

方差为

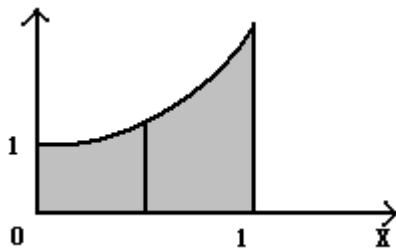
$$\begin{aligned} \sigma_I^2 &= \sigma^2 \left(\sum_{i=1}^m \frac{l_i}{n_i} \sum_{j=1}^{n_i} f(x_j^{(i)}) \right) \\ &= \sum_{i=1}^m \frac{l_i^2}{n_i^2} \sum_{j=1}^{n_i} \sigma^2(f(x_j^{(i)})) \\ &= \sum_{i=1}^m \frac{l_i^2}{n_i} \sigma^2(f(x^{(i)})) \end{aligned} \quad (3.40)$$

其中

$$\begin{aligned} \sigma^2(f(x^{(i)})) &= E[f^2(x^{(i)})] - [Ef(x^{(i)})]^2 \\ &= \int_{a_{i-1}}^{a_i} \frac{1}{l_i} f^2(x)dx - \left(\frac{I_i}{l_i} \right)^2 \end{aligned} \quad (3.41)$$

例：用分层抽样法求 $I = \int_0^1 e^x dx$

将积分 $[0, 1]$ 分成 $[0, 0.5]$ 和 $[0.5, 1]$ 两个区间



在 $[0, 0.5]$ 上抽样 4 次： $x_j^{(1)} = 0.5\xi_j$, $(j = 1, 2, 3, 4)$, 在 $[0.5, 1]$ 上抽样 6 次： $x_j^{(2)} = 0.5 + 0.5\xi_j$ $(j = 1, 2, \dots, 6)$, 由分层抽样公式,

$$\begin{aligned}\theta &= \sum_{i=1}^2 \frac{l_i}{n_i} \sum_{j=1}^{n_i} f(x_j^{(i)}) \\ &= \frac{0.5}{4} \sum_{j=1}^4 f(x_j^{(1)}) + \frac{0.5}{6} \sum_{j=1}^6 f(x_j^{(2)}) \\ &= \frac{1}{8} \sum_{j=1}^4 e^{x_j^{(1)}} + \frac{1}{12} \sum_{j=1}^6 e^{x_j^{(2)}}\end{aligned}\quad (3.42)$$

估计量 θ 的方差为

$$\begin{aligned}\sigma^2(\theta) &= \sum_{i=1}^2 \frac{l_i^2}{n_i} \sigma^2(f(x^{(i)})) \\ &= \frac{1}{16} \sigma^2(e^{x^{(1)}}) + \frac{1}{24} \sigma^2(e^{x^{(2)}})\end{aligned}\quad (3.43)$$

$$\begin{aligned}\sigma^2(e^{x^{(1)}}) &= \int_0^{0.5} \frac{1}{l_1} f^2(x) dx - \left(\frac{1}{l_1} \int_0^{0.5} f(x) dx \right)^2 \\ &= \int_0^{0.5} \frac{1}{0.5} e^{2x} dx - \left(\int_0^{0.5} \frac{1}{0.5} e^x dx \right)^2 \\ &= (e - 1) - 4(\sqrt{e} - 1)^2 \\ &= 0.03492\end{aligned}\quad (3.44)$$

$$\sigma^2(e^{x^{(2)}}) = 0.09493 \quad (3.45)$$

则

$$\sigma^2(\theta) = \frac{1}{16} \times 0.03492 + \frac{1}{24} \times 0.09493 = 0.006138 \quad (3.46)$$

在 $N = 10$ 次抽样中, MC 一般法, $\sigma_{MC}^2 = 0.242$, 重要抽样 $\sigma_{IS}^2 = 0.00269$

§3.5 相关抽样技巧 (Correlation method)

思想：改变统计量 g , 确定一个相关量 g^* , 使 $\int g^* f(x) dx$ 可以积分, 而 $\sigma^2((g - g^*)|f)$ 尽量小

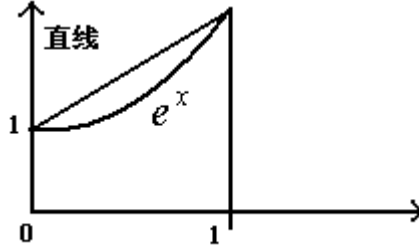
$$G = \int g(x) f(x) dx$$

$$\begin{aligned}
&= \int [g(x) - h(x) + h(x)]f(x)dx \\
&= \int [g(x) - h(x)]f(x) + \int h(x)f(x)dx
\end{aligned} \tag{3.47}$$

其中 $\int h(x)f(x)dx$ 可以解析积分出, 则 G 的 MC 估计:

$$G \approx \int h(x)f(x)dx + \frac{1}{N} \sum_{i=1}^N [g(x_i^{(f)}) - h(x_i^{(f)})] \tag{3.48}$$

x_i 由 $f(x)$ 抽样, 而 $\sigma^2[(g(x) - h(x))|f(x)] \ll \sigma^2(g(x)|f(x))$
 例 1. $I = \int_0^1 e^x dx$, MC 直解 $\sigma_{MC}^2 = 0.242$



$$\sigma_I^2 = \int_0^1 e^{2x} dx - [\int_0^1 e^x dx]^2 = 0.242 \tag{3.49}$$

$$I_{CS} = \int_0^1 [e^x - (1+x) + (1+x)]dx = \int_0^1 (e^x - (1+x)) + \frac{3}{2} \tag{3.50}$$

方差

$$\begin{aligned}
\sigma_{CS}^2 &= \sigma^2(e^x - (1+x)) + (\frac{3}{2})^2 - (\frac{3}{2})^2 \\
&= \int_0^1 [e^x - (1+x)]^2 dx - [\int_0^1 (e^x - (1+x))dx]^2 \\
&= 0.0437
\end{aligned} \tag{3.51}$$

进一步, 最佳化估计 $\int_0^1 (e^x - (1+\beta x) + (1+\beta x))dx$

$$\begin{aligned}
&\int_0^1 (e^x - 1 - \beta x)^2 dx - (\int_0^1 (e^x - 1 - \beta x)dx)^2 = \min \\
\Rightarrow \quad \beta &= 1.09, \quad \sigma_{min}^2 = 0.0039
\end{aligned} \tag{3.52}$$

§3.6 对偶抽样技巧 (Antithetic Variate)

$$E(g) = \iint g(x, y)f(x, y)dx dy \tag{3.53}$$

$$\rho_{xy} = \frac{cov(x, y)}{\sqrt{\sigma^2(x)} \cdot \sqrt{\sigma^2(y)}} \tag{3.54}$$

令 x, y 由 $f(x, y)$ 抽样, x', y' 是由同分布变换确定的, 即

$$\begin{cases} x' = \varphi(x, y) \\ y' = \psi(x, y) \end{cases} \quad (3.55)$$

$$cov(x, y) = \langle (x - \bar{x})(y - \bar{y}) \rangle \quad (3.56)$$

x', y' 和 x, y 服从同分布,

$$g^*(x, y) = g(\varphi(x, y), \psi(x, y)) \quad (3.57)$$

则

$$E(g) = \iint \frac{1}{2} [g(x, y) + g^*(x, y)] f(x, y) dx dy \quad (3.58)$$

对偶技巧:

$$E(g) = \iint g_{AS}(x, y) f(x, y) dx dy \quad (3.59)$$

● 确定合适的变换, 使 $g^*(x, y)$ 与 $g(x, y)$ 是负相关。 $g^*(x, y)$ 和 $g(x, y)$ 在 $f(x, y)$ 下有相同的方差, 而 $g_{AS}(x, y)$ 的方差

$$\sigma_{AS}^2 = \frac{1 + \rho}{2} \sigma^2(g(x, y) | f(x, y)) \quad (3.60)$$

当相关系数 $\rho = -1$ 时, $\sigma_{AS}^2 = 0$ 。

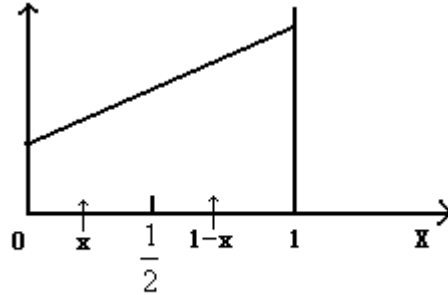
例 1. $G = \int_0^1 g(x) dx$, 这里 $g(x)$ 是线性函数,

$$G = \int_0^1 \frac{1}{2} [g(x) + g(1-x)] dx \quad (3.61)$$

$$G_N = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} [g(x_i) + g(1-x_i)] \quad (3.62)$$

$$\sigma_G^2 = 0$$

近似线性函数可以使用。



例 2. $G = \int_0^1 e^x dx$

$$G_N = \frac{1}{2} \sum_{i=1}^N [g(x_i) + g(1-x_i)] \quad (3.63)$$

$$\sigma_{AS}^2 = 0.0039, \text{ 而 } \sigma_{conv}^2 = 0.0437, \quad \sigma_{MC}^2 = 0.242$$

本章小结： MC 的一般技巧： 减小方差的方法

$$E(g) = \int g(x)f(x)dx \quad (3.64)$$

$$\hat{G}_N = \frac{1}{N} \sum_{i=1}^N g(x_f^{(i)}) \quad (3.65)$$

$$|\hat{G}_N - E(g)| = \varepsilon = \frac{X_\alpha \sigma(g)}{\sqrt{N}} \quad (3.66)$$

$$\sigma^2(g) = \int g^2 f(x)dx - (\int g f(x)dx)^2 \quad (3.67)$$

(1) 期望估计技巧 (Expected-Value method) 多维积分，先将有些维度的积分解析或数值积分 $E(g) = \int g_1(x)f_1(x)dx$ 。

(2) 重要抽样技巧 (Importance Sampling)

$$E(g) = \int g(x)f(x)dx = \int \left[\frac{gf}{f^*}\right] f^* dx \quad (3.68)$$

$\frac{gf}{f^*}$ 为统计量， f^* 为分布密度。

(3) 分层抽样技巧 (Stratification method)

$$E(g) = \int_a^b g(x)f(x)dx = \sum_{i=1}^{n_i} \int_{a_{i-1}}^{a_i} g(x)f(x)dx \quad (3.69)$$

对积分贡献大的区间要多抽样。

(4) 相关抽样技巧 (Correlation method)

$$\begin{aligned} E(g) &= \int g(x)f(x)dx \\ &= \int [g(x) - h(x) + h(x)]f(x)dx \\ &= \int [g(x) - h(x)]f(x)dx + \int h(x)f(x)dx \end{aligned} \quad (3.70)$$

前者尽可能小，后者可解析积分 $\sigma^2 = 0$ 。

(5) 对偶抽样技巧 (Antithetic Variate)

$$\begin{aligned} cov(g, g^*) &= \int_0^1 (e^x - \int_0^1 e^{x'} dx') (e^{1-x} - \int_0^1 e^{x'} dx') dx \\ &= \int_0^1 [e^x - (e-1)][e^{1-x} - (e-1)] dx \\ &= \int_0^1 [e + (e-1)^2 - (e-1)(e^x + e^{1-x})] dx \\ &= e + (e-1)^2 - 2(e-1)^2 \\ &= e - (e-1)^2 \\ &= -0.2342 \end{aligned} \quad (3.71)$$

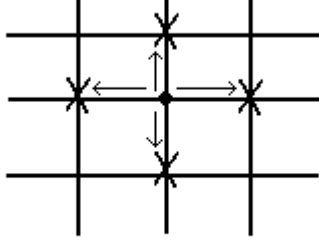
$$\begin{aligned}
\rho &= \frac{cov(g, g^*)}{\sqrt{\sigma^2(g)}\sqrt{\sigma^2(g^*)}} \\
&= \frac{-0.2342}{\sqrt{0.242}\sqrt{0.242}} \\
&= -0.9678
\end{aligned} \tag{3.72}$$

第四章 非归一化 Metropolis 抽样方法 (MC 在统计物理中的应用)

§4.1 预备知识

马尔科夫链 跃迁概率 随机矩阵

一. 马尔科夫过程: 最近的将来的统计性质由现时完全决定, 与过去无关 (缺乏记忆)



定义: 序列 $x_0, x_1, \dots, x_n, \dots$, 如果对于任何 n 都有

$$P(x_n | x_{n-1}, \dots, x_1, x_0) = P(x_n | x_{n-1}) \quad (\text{条件概率}) \quad (4.1)$$

则称此序列是一个马尔科夫链。序列 x_0, \dots, x_n 发生的概率可分解因式为

$$P(x_0, \dots, x_n) = P(x_n | x_{n-1}) P(x_{n-1} | x_{n-2}) \cdots P(x_1 | x_0) P(x_0) \quad (4.2)$$

二. 一步跃迁概率简记为

$$P_{ij} = P(x_j | x_i) = P(x_i \rightarrow x_j) \quad (4.3)$$

三. 随机矩阵

将随机概率排成一个矩阵, 其本征值为 1 (各态历经)。考虑一个具有四个状态的系统 $P = (P_{ij})$, 其跃迁概率排成随机矩阵

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix} \quad (4.4)$$

四. 计算机模拟出发点:

从一个初始态 x_0 出发 ($P(x_0) = 1$), 最终诸状态遵从某一特定分布, 如

$$P(x_i) \propto e^{-H(x_i)/k_B T} \quad (4.5)$$

§4.2 正则系综平均量的计算

一. 平均观察量

在 Statistical Physics , 物理量 A , 平均观察量 $\langle A \rangle$,

$$\langle A \rangle = \int_{\Omega} A(\vec{x}) P(\vec{x}) d\vec{x} \quad (4.6)$$

其中 \vec{x} 是相空间 Ω 中的一个点 (状态, 组态, 位形), $P(\vec{x})$ 是系综分布, 且满足 $P(\vec{x}) \geq 0$, $\int_{\Omega} P(\vec{x}) d\vec{x} = 1$

假设这个系统有 N 个全同粒子, 一个粒子状态用动力学参数 $\{\vec{r}, \vec{v}, s\}$ (s 为自旋) 等描述, 用 $\alpha = (\vec{r}, \vec{v}, s, \dots)$, 于是, Ω 中的点 \vec{x} 可表示为 $\vec{x} = \{\alpha_1, \alpha_2, \dots, \alpha_N\}$, 即系统 (N 个粒子) 的一个状态。

正则系综: 粒子数 N , 温度 T 和体积 V 固定, 其分布密度函数是 Boltzmann 分布

$$P(\vec{x}) = \frac{e^{-H_N(\vec{x})/k_B T}}{\int_{\Omega} e^{-H_N(\vec{x})/k_B T} d\vec{x}} \quad (4.7)$$

$H_N(\vec{x})$ 为系统的哈密顿量, 则系综平均量改写为

$$\langle A \rangle = \frac{\int_{\Omega} A(\vec{x}) e^{-H_N(\vec{x})/k_B T} d\vec{x}}{\int_{\Omega} e^{-H_N(\vec{x})/k_B T} d\vec{x}} \quad (4.8)$$

如果系统状态 \vec{x} 取离散点上的值, 则

$$\langle A \rangle = \frac{\sum_i A(\vec{x}_i) e^{-H_N(\vec{x}_i)/k_B T}}{\sum_i e^{-H_N(\vec{x}_i)/k_B T}} \quad (4.9)$$

除此, 还有微正则系综 (孤立的, 参量固定, 统计平衡态)、巨正则系综 (化学势 μ , V , T 固定, 粒子数 N 变化)

二. 内能, 比热, 自由能和熵

内能 U_N 定义为

$$U_N = \langle H_N(\vec{x}) \rangle = \frac{\int_{\Omega} H_N(\vec{x}) e^{-H_N(\vec{x})/k_B T} d\vec{x}}{\int_{\Omega} e^{-H_N(\vec{x})/k_B T} d\vec{x}} \quad (4.10)$$

定容热容量

$$C_V = \frac{\partial U_N}{\partial T} = k_B \frac{\partial U_N}{\partial (k_B T)} = [\langle H_N^2(\vec{x}) \rangle - \langle H_N(\vec{x}) \rangle^2] / k_B T^2 \quad (4.11)$$

自由能

$$F_N = -k_B T \ln Z_N \quad (4.12)$$

熵

$$S_N = \frac{(U_N - F_N)}{T} \quad (4.13)$$

其中配分函数

$$Z_N = \int_{\Omega} e^{-H_N(\vec{x})/k_B T} d\vec{x} \quad (4.14)$$

三. 蒙特卡罗计算

● 2 维 Ising 模型, $N = 10 \times 10$ 点阵, 系统的一个状态点 $\vec{x} = (\sigma_1, \sigma_2, \dots, \sigma_{100})$, 其中 σ_j 是第 j 个地址上的自旋, 每个自旋取 $\sigma_j = 1$, $\sigma_j = -1$ 两个值. $\langle A \rangle$ 要在 2^{100} 项上求和, $2^{100} = 1.3 \times 10^{30}$, 对于 1 台每秒亿次 CP, 至少需计算 4×10^{14} 年, 因此需用 MC.

● MC 求系统平均 $\langle A \rangle$ 的思想是:

从某个已知分布密度 $\pi(\vec{x})$, ($\pi(\vec{x}) \geq 0$, $\int_{\Omega} \pi(\vec{x}) d\vec{x} = 1$) 抽样得到

$$A = \frac{\int_{\Omega} A(\vec{x}) e^{-H_N(\vec{x})/k_B T} \frac{1}{\pi(\vec{x})} \pi(\vec{x}) d\vec{x}}{\int_{\Omega} e^{-H_N(\vec{x})/k_B T} \frac{1}{\pi(\vec{x})} \pi(\vec{x}) d\vec{x}} \quad (4.15)$$

抽样 $\vec{x}_m, m = 1, 2, \dots, M$, 然后

$$\langle A \rangle = \frac{\sum_m A(\vec{x}_m) e^{-H_N(\vec{x}_m)/k_B T} [\pi(\vec{x}_m)]^{-1}}{\sum_m e^{-H_N(\vec{x}_m)/k_B T} [\pi(\vec{x}_m)]^{-1}} \quad (4.16)$$

$\pi(\vec{x})$ 的选法有多种, 一种好的选择是取系综分布本身, 即

$$\pi(\vec{x}) = \frac{e^{-H_N(\vec{x})/k_B T}}{\int_{\Omega} e^{-H_N(\vec{x})/k_B T} d\vec{x}} \quad (4.17)$$

则

$$\langle A \rangle \approx A_M = \frac{1}{M} \sum_{m=1}^M A(\vec{x}_m) \quad (4.18)$$

因为在 $\pi(\vec{x})$ 中有一个未知的归一化常数 (配分函数) $Z = \int_{\Omega} e^{-H_N(\vec{x})/k_B T} d\vec{x}$ 因此, 直接抽样难以应用:

- ① $H_N(\vec{x})$ 复杂, 甚至取负值, 难以求得 $\pi(\vec{x})$ 的最大值;
 - ② $M = \max[\pi(\vec{x})]$, 但 $\pi(\vec{x})$ 随 \vec{x} 变化很大, 舍选法无效。
- 1953 年, Metropolis 抽样方法在 S.P 中广泛应用。

§4.3 Metropolis 抽样方法

一. 离散分布的 Metropolis 抽样方法

研究由离散分布

$$\pi_i = \frac{e^{-H_N(\vec{x}_i)/k_B T}}{\sum_i e^{-H_N(\vec{x}_i)/k_B T}}, \quad i = 1, 2, \dots, I \quad (4.19)$$

的抽样问题, 设 $\pi_i > 0$ ($i = 1, 2, \dots, I$),

☆代表点, 或子样 $\vec{x}_{i0}, \vec{x}_{i1}, \dots, \vec{x}_{im}, \vec{x}_{im+1}, \dots$ 可由以下手续产生:

- ① 初始点 \vec{x}_{i0} 由任意初始分布 $S(\vec{x}_{i0})$ 抽样, 这里 $S_{i0} = S(\vec{x}_{i0}) \geq 0$, $\sum_{i0=1}^I S(\vec{x}_{i0}) = 1$
- ② 给定一个对称的转移概率矩阵, $P^* = (P_{ij}^*)$ 满足

$$P_{ij}^* = P^*(\vec{x}_j | \vec{x}_i) = P^*(\vec{x}_i | \vec{x}_j) = P_{ji}^* > 0 \quad (4.20)$$

$$\sum_{j=1}^I P_{ij}^* = 1, \quad i = 1, 2, \dots, I \quad (4.21)$$

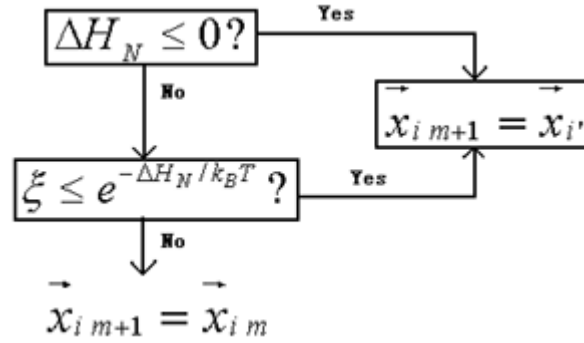
③在 \vec{x}_{im} 已知的条件下, \vec{x}_{im+1} 可如下产生:

从对称转移矩阵 P^* 的第 im 所在行 (同行抽)

$$P_{im,i'} = P^*(\vec{x}_{i'}|\vec{x}_{im}), \quad i' = 1, 2, \dots, I \quad (4.22)$$

抽样产生 $\vec{x}_{i'}$, 计算能量差

$$\Delta H_N(\vec{x}_{im}, \vec{x}_{i'}) = H_N(\vec{x}_{i'}) - H_N(\vec{x}_{im}) \quad (4.23)$$



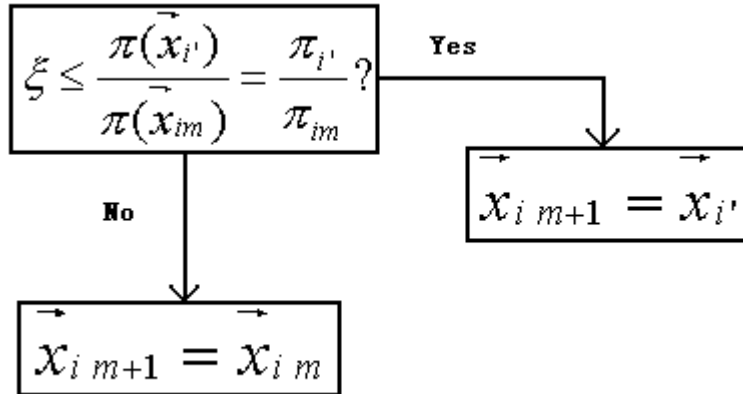
注意到

$$e^{-\Delta H_N/k_B T} = \frac{e^{-H_N(\vec{x}_{i'})/k_B T}}{e^{-H_N(\vec{x}_{im})/k_B T}} = \frac{\pi_{i'}}{\pi_{im}} \quad (4.24)$$

$$\Delta H_N \leq 0 \Rightarrow e^{-\Delta H_N/k_B T} \geq 1 \Rightarrow \frac{\pi_{i'}}{\pi_{im}} \geq 1 \quad (\pi_{i'} \geq \pi_{im}) \quad (4.25)$$

④上述抽样过程重写为:

- a) 从对称转移矩阵 P^* 的第 im 行, $P_{im,i'} = P^*(\vec{x}_{i'}|\vec{x}_{im})$, $i' = 1, 2, \dots, I$ 中抽样产生 $\vec{x}_{i'}$
- b)



- c) 不断重复 b).

二. Metropolis 抽样的物理解释。

1) 若 $H_N(\vec{x}_{i'}) < H_N(\vec{x}_{im}) \Rightarrow \pi(\vec{x}_{i'}) > \pi(\vec{x}_{im})$ 则 $\vec{x}_{im+1} = \vec{x}_{i'}$

若 $H_N(\vec{x}_{i'}) > H_N(\vec{x}_{im}) \Rightarrow \pi(\vec{x}_{i'}) < \pi(\vec{x}_{im})$ 则 $\vec{x}_{im+1} = \vec{x}_{i'}$ (概率为 $\pi(\vec{x}_{i'})/\pi(\vec{x}_{im})$) 这是符合概率密度 $\pi(\vec{x}_{i'})$ 大时, 抽 $\vec{x}_{i'}$ 的机会多这一特点。

2) 已知 \vec{x}_{im} 后, \vec{x}_{im+1} 的确定只与 \vec{x}_{im} 有关, 即与 $\vec{x}_{i0}, \vec{x}_{i1}, \dots, \vec{x}_{im-1}$ 无关, 因此, \vec{x}_{im} 组成一个有限状态的均匀马尔可夫链。

3) 这样抽样过程与归一化常数无关。

§4.4 热浴法 (Heat bath)

一. 对于正则系综, 观察量的平均值

$$\langle A(\vec{x}) \rangle_T = \frac{1}{Z} \int d\vec{x} e^{-H(\vec{x})/k_B T} A(\vec{x}) \quad (4.26)$$

$$Z = \int d\vec{x} e^{-H(\vec{x})/k_B T} \quad (4.27)$$

● 简单抽样: 积分是在一切态 $\{x_l\}$ 按每个态的固有权重 $e^{-H(\vec{x})/k_B T}$ 来求积。

$$M \rightarrow \infty \quad \overline{A(\vec{x})} = \frac{\sum_{l=1}^M \exp[-H(x_l)/k_B T] A(x_l)}{\sum_{l=1}^M \exp[-H(x_l)/k_B T]} \quad (4.28)$$

$H(\vec{x})$ 只包括势能项。动能项是二次项, 可解析积分。

● 重要抽样: 以某一概率 $P(x_l)$ 选取相应的点 x_l ($l = 1, 2, \dots, M$)

$$\overline{A(\vec{x})} = \frac{\sum_{l=1}^M \exp[-H(x_l)/k_B T] A(x_l) [P(x_l)]^{-1}}{\sum_{l=1}^M \exp[-H(x_l)/k_B T] [P(x_l)]^{-1}} \quad (4.29)$$

不是完全随机地对各个位形 x_l 进行抽样。而是优先在相互间在温度 T 下为重要的那一区域中对 x_l 进行抽样。

二. Metropolis 抽样

若选

$$P(x_l) \propto e^{-H(x_l)/k_B T} \Rightarrow \overline{A(\vec{x})} = \frac{1}{M} \sum_{l=1}^M A(x_l) \quad (4.30)$$

基本思想: 不独立选取相继诸状态 x_l , 而是构造一个 Markov 过程, 每一状态 x_{l+1} 由前一个状态 x_l , 通过一个适当的跃迁概率 $W(x_l \rightarrow x_{l+1})$ 得到, ($l = 1, 2, \dots, m$), 在 $M \rightarrow \infty$, Markov 过程产生的状态的分布函数 $P(x_l)$ 趋于平衡分布

$$P_{eq} = \frac{1}{Z} e^{-H(x_l)/k_B T} \quad (4.31)$$

达到这一点的充分条件是加上细致平衡条件

$$P_{eq}(x_l) W(x_l \rightarrow x_{l'}) = P_{eq}(x_{l'}) W(x_{l'} \rightarrow x_l) \quad (4.32)$$

即

$$\frac{W(x_l \rightarrow x_{l'})}{W(x_{l'} \rightarrow x_l)} = \frac{P_{eq}(x_{l'})}{P_{eq}(x_l)} = \frac{e^{-H(x_{l'})/k_B T}}{e^{-H(x_l)/k_B T}} = e^{-(H(x_{l'}) - H(x_l))/k_B T} \quad (4.33)$$

上式不能唯一地确定 $W(x_l \rightarrow x_{l'})$, 两种经常用的选择:

①

$$W(x_l \rightarrow x_{l'}) = \begin{cases} \frac{1}{\tau_s} e^{-\Delta H/k_B T}, & \Delta H > 0 \\ \frac{1}{\tau_s}, & \text{else} \end{cases} \quad (4.34)$$

②

$$W(x_l \rightarrow x_{l'}) = \frac{1}{\tau_s} \cdot \frac{e^{-\Delta H/k_B T}}{1 + e^{-\Delta H/k_B T}} \quad (4.35)$$

$$\Delta H = H(x_{l'}) - H(x_l) \quad (4.36)$$

式中 τ_s : MC 时间, 则 $W(x_l \rightarrow x_{l'})$ 是单位时间的跃迁概率, $\tau_s = 1$ 。

三. 渐近分布

$$\begin{array}{ccccccc} x_1 & \rightarrow & x_2 & \rightarrow & \dots & \rightarrow & x_l & \rightarrow & x_{l'} & \rightarrow & \dots \\ P(x_1) & & P(x_2) & & & & P(x_l) & & P(x_{l'}) & & \end{array}$$

概率分布收敛于 $P_{eq}(x)$ 。

证明:

$N_r, N_s < N$, 假设 $H(x_r) < H(x_s)$, 用随机数可以产生变迁 $x_r \rightarrow x_s$, 若不考虑能量变化, 令 $\Delta H = 0$ (自由空间), 则 $x_r \rightarrow x_s$ 与 $x_s \rightarrow x_r$ 是相同的, 即跃迁概率是对称的: $W_{\Delta H=0}(x_r \rightarrow x_s) = W_{\Delta H=0}(x_s \rightarrow x_r)$ 。根据细致平衡原理,

$$W(x_r \rightarrow x_s) = W_{\Delta H=0}(x_r \rightarrow x_s) e^{-[H(x_s) - H(x_r)]/k_B T} \quad (4.37)$$

$$W(x_s \rightarrow x_r) = W_{\Delta H=0}(x_s \rightarrow x_r) = W_{\Delta H=0}(x_r \rightarrow x_s) \quad (4.38)$$

在 Markov 链的这一点, 从 x_r 到 x_s 的跃迁总数 $N_{r \rightarrow s}$ 为:

$$N_{r \rightarrow s} = N_r W(x_r \rightarrow x_s) = N_r W_{\Delta H=0}(x_r \rightarrow x_s) e^{-[H(x_s) - H(x_r)]/k_B T} \quad (4.39)$$

而反向跃迁的总数为:

$$N_{s \rightarrow r} = N_s W(x_s \rightarrow x_r) = N_s W_{\Delta H=0}(x_s \rightarrow x_r) = N_s W_{\Delta H=0}(x_r \rightarrow x_s) \quad (4.40)$$

于是净跃迁数

$$\begin{aligned} \Delta N_{r \rightarrow s} &= N_{r \rightarrow s} - N_{s \rightarrow r} \\ &= W_{\Delta H}(x_r \rightarrow x_s) \{ N_r e^{-[H(x_s) - H(x_r)]/k_B T} - N_s \} \\ &= N_r W_{\Delta H}(x_r \rightarrow x_s) \left[\frac{e^{-H(x_s)/k_B T}}{e^{-H(x_r)/k_B T}} - \frac{N_s}{N_r} \right] \end{aligned} \quad (4.41)$$

只要 $\frac{N_s}{N_r} < \frac{e^{-H(x_s)/k_B T}}{e^{-H(x_r)/k_B T}}$ (平衡概率之比), 即可得到 $\Delta N_{r \rightarrow s} > 0$

四. 正则系综的蒙特卡罗步骤

(1) 规定一个初始位形

(2) 产生一个新位形 \vec{x}' $\begin{cases} \text{随机数} \\ \text{规则数} \end{cases}$

(3) 计算能量变化 $\Delta H = H(\vec{x}') - H(\vec{x}_m)$

(4) 若 $\Delta H < 0$, 接受新位形 $\vec{x}_{m+1} = \vec{x}'$, 并回到第 2 步, 否则继续进行下一步

(5) 计算 $e^{-\Delta H/k_B T}$

(6) 产生一个随机数 $R \in (0, 1)$

(7) 若 $R < e^{-\Delta H/k_B T}$, 接受新位形 $\vec{x}_{m+1} = \vec{x}'$, 并回到第 2 步

(8) 否则, 保持原位形作为新位形并回到第 2 步 $\vec{x}_{m+1} = \vec{x}_m$

第 4 步表明我们对一个能量低于早先之值的新位形无条件地接受, 而对升高能量的位形则仅按 Boltzmann 概率接受。

五. 意义

☆人们最先并不知道重要贡献来自何处, 但人们可以设计一个算法, 它会肯定把人们引导到哪里。

☆必须产生大量的位形 $\{x_l\}$ 之后, 对初始状态的“记忆”才会消失。

六. 各态历经定理

①

$$\overline{A(x)} = \frac{1}{M} \sum_{l=1}^M A(\vec{x}_l) \quad (4.42)$$

可理解为沿着相空间的随机轨迹的时间平均

$$\overline{A} = \frac{1}{t_M - t_{M0}} \int_{t_{M0}}^{t_M} A(t) dt = \frac{1}{M - M_0} \sum_{\nu=M_0+1}^M A(\vec{x}_\nu) \quad (4.43)$$

$t > M_0$ 系统驰豫平衡

②

$$\lim_{t \rightarrow \infty} \overline{A} = \langle A \rangle_T = \langle A(\infty) \rangle_T \quad (4.44)$$

$$\langle A \rangle_T = \int_{\Omega} A(\vec{x}) P_{eq}(\vec{x}) d\vec{x} \quad (4.45)$$

③也就是 Metropolis 重要抽样将引发遍历性结论。

§4.5 广义 Metropolis 抽样方法

1. 主方程

收支平衡情况:

$$\frac{\partial \rho(x, t)}{\partial t} = - \sum_{x'} K(x \rightarrow x') \rho(x, t) + \sum_{x'} K(x' \rightarrow x) \rho(x', t) \quad (4.46)$$

其中 $\rho(x', t)$ 可理解为粒子数密度。

Stationary State : $\frac{\partial \rho(x, t)}{\partial t} \equiv 0$

$$\sum_{x'} K(x \rightarrow x') \rho(x, t) = \sum_{x'} K(x' \rightarrow x) \rho(x', t) \quad (4.47)$$

主方程的正解

$$K(x \rightarrow x') \rho(x) = K(x' \rightarrow x) \rho(x') \quad (4.48)$$

设转移概率

$$K(x \rightarrow x') = A(x'|x) T(x'|x) \quad (4.49)$$

其中 $A(x'|x)$ 是接受移动 $x \rightarrow x'$ 的几率, $T(x'|x)$ 是在 x 条件下, 发现 x' 的概率。

$$A(x'|x)T(x'|x)\rho(x) = A(x|x')T(x|x')\rho(x') \quad (4.50)$$

对所有的 x, x' , 都有

$$\sum_{x'} T(x'|x) = 1, \quad T(x'|x) = T(x|x') > 0 \quad (4.51)$$

接受几率比:

$$\frac{A(x'|x)}{A(x|x')} = \frac{\rho(x')}{\rho(x)} \quad (4.52)$$

如果 $\rho(x') < \rho(x)$, 则 $A(x'|x) = \frac{\rho(x')}{\rho(x)}$; 如果 $\rho(x') \geq \rho(x)$, 则 $A(x'|x) = 1$ 。

接受几率:

$$q = \min(1, \frac{\rho(x')}{\rho(x)}) \quad (4.53)$$

第五章 蒙特卡罗方法模拟主方程 (跃迁过程)

§5.1 主方程差分解的 MC 模拟

一. 时间相关平均量

$\vec{x}(t)$ 是系统的某个状态, $P(\vec{x}, t)$ 是系统 t 时在状态 \vec{x} 的几率分布,

$\langle A_t(\vec{x}) \rangle = \int_{\Omega} A(\vec{x}) P(\vec{x}, t) d\vec{x}$ 是在某固定时间下的平均。

问题: ① 如何从 $P(\vec{x}, t)$ 中抽样 \vec{x} ?

② $P(\vec{x}, t)$ 满足什么方程?

二. 主方程

设 $P(\vec{x}, t) \geq 0$, $\int_{\Omega} P(\vec{x}, t) d\vec{x} = 1$, 对任意 t 都满足 Master equation

$$\frac{\partial P(\vec{x}, t)}{\partial t} = \int_{\Omega} W(\vec{x}' \rightarrow \vec{x}, t) P(\vec{x}', t) d\vec{x}' - \int_{\Omega} W(\vec{x} \rightarrow \vec{x}', t) P(\vec{x}, t) d\vec{x}' \quad (5.1)$$

$$P(\vec{x}, t) \Big|_{t=0} = S(\vec{x}) \quad (5.2)$$

$W(\vec{x}' \rightarrow \vec{x}, t)$ 为跃迁函数, $S(\vec{x})$ 为系统的初始分布, $S(\vec{x}) \geq 0$, $\int_{\Omega} S(\vec{x}) d\vec{x} = 1$ 。令

$$W_t(\vec{x}, t) = \int_{\Omega} W(\vec{x} \rightarrow \vec{x}', t) d\vec{x}' \quad (5.3)$$

表示 t 时刻系统在 \vec{x} 状态, 单位时间内发生跃的总次数 (跃迁率)。则主方程变为

$$\frac{\partial}{\partial t} P(\vec{x}, t) + W_t(\vec{x}, t) P(\vec{x}, t) = \int_{\Omega} W(\vec{x}' \rightarrow \vec{x}, t) P(\vec{x}', t) d\vec{x}' \quad (5.4)$$

$$P(\vec{x}, t) \Big|_{t=0} = S(\vec{x}) \quad (5.5)$$

① 如果跃迁函数 $W(\vec{x}' \rightarrow \vec{x}, t)$ 为各向同性, 即 $W(\vec{x}' \rightarrow \vec{x}, t) = c$, 则

$$W_t(\vec{x}, t) = c\Omega \quad (5.6)$$

$$P(\vec{x}, t) = S(\vec{x}) e^{-c\Omega t} + \frac{1}{\Omega} (1 - e^{-c\Omega t}) \quad (5.7)$$

那么

$$\lim_{t \rightarrow \infty} P(\vec{x}, t) = \frac{1}{\Omega} \quad (5.8)$$

② 如果 $t \geq t_d$, (从某时间 t_d 起), $\frac{\partial}{\partial t} P(\vec{x}, t) = 0$ 则

$$W_t(\vec{x}, t) P(\vec{x}, t) = \int_{\Omega} W(\vec{x}' \rightarrow \vec{x}, t) P(\vec{x}', t) d\vec{x}' \quad (5.9)$$

反之也成立。特别地, 如果从某个时刻起, 系统细致平衡成立 (Detail Balance), 即

$$W(\vec{x} \rightarrow \vec{x}', t) P(\vec{x}, t) = W(\vec{x}' \rightarrow \vec{x}, t) P(\vec{x}', t) \quad (5.10)$$

得到 $\frac{\partial P(\vec{x}, t)}{\partial t} = 0 \Rightarrow P(\vec{x}, t) = P(\vec{x})$ 稳定态。

三. 主方程的差分解

$$\frac{P(\vec{x}, t + \Delta t) - P(\vec{x}, t)}{\Delta t} + W_t(\vec{x}, t)P(\vec{x}, t) = \int_{\Omega} W(\vec{x}' \rightarrow \vec{x}, t)P(\vec{x}', t)d\vec{x}' \quad (5.11)$$

稳定后有

$$P(\vec{x}, t + \Delta t) = [1 - \Delta t W_t(\vec{x}, t)]P(\vec{x}, t) + \Delta t \int_{\Omega} W(\vec{x}' \rightarrow \vec{x}, t)P(\vec{x}', t)d\vec{x}' \quad (5.12)$$

分三种情况:

(1) $W_t(\vec{x}, t) = W_0 \equiv c$ (常数), 取 $\Delta t = \frac{1}{W_0}$, 系统平均发生一次跃迁所要的时间

$$P(\vec{x}, t + \Delta t) = \int_{\Omega} W^*(\vec{x}' \rightarrow \vec{x}, t)P(\vec{x}', t)d\vec{x}' \quad (5.13)$$

其中

$$W^*(\vec{x}' \rightarrow \vec{x}, t) = \frac{1}{W_0} W(\vec{x}' \rightarrow \vec{x}, t) \geq 0 \quad (5.14)$$

$$\int_{\Omega} W^*(\vec{x}' \rightarrow \vec{x}, t)d\vec{x} = 1 \quad (5.15)$$

取 $t = m \cdot \Delta t$, 令 $P(\vec{x}, t) = P(\vec{x}, m\Delta t) = P_m(\vec{x})$, $m = 0, 1, 2, \dots$, 于是

$$\begin{cases} P_{m+1}(\vec{x}) = \int_{\Omega} W^*(\vec{x}' \rightarrow \vec{x}, t_m)P_m(\vec{x}')d\vec{x}', \\ P_0(\vec{x}) = S(\vec{x}) \end{cases} \quad m = 0, 1, 2, \dots \quad (5.16)$$

这样, $P_m(\vec{x})$ 的样本 \vec{x}_m 可用以下方法产生:

(a) \vec{x}_0 由初始分布 $S(\vec{x}_0)$ 抽样产生。

(b) 当 \vec{x}_m 给定后, \vec{x}_{m+1} 由 $W^*(\vec{x} \rightarrow \vec{x}_{m+1}, t_m)$ 抽样产生。

显然, 上述由 $\{P_m(\vec{x}_m)\}_m$ 顺序抽样产生 $\{\vec{x}_m\}_m$ 的过程与 Metropolis 方法的过程相似。换句话说, Metropolis 方法中的每一次抽样, 是某个满足主方程的系统, 在 Δt 时间内, 由一个状态到另一个状态的一次跃迁。跃迁函数 = 马尔科夫转移概率函数。

(2) $W_t(\vec{x}, t) = W_0(t)$, 这时, 取 $\Delta t = \frac{1}{W_0(t)}$, 与时间有关。

$$t_m = t_{m-1} + \Delta t_m = \sum_{l=1}^m \Delta t_l \quad \Delta t_l = \frac{1}{W_0(t_{l-1})} \quad (5.17)$$

(3) $W_t(\vec{x}, t)$ 与 \vec{x}, t 均有关时, 设 $W_t(\vec{x}, t)$ 有界, 令 $W_{max} = \max W_t(\vec{x}, t)$, 这时, 主方程写为:

$$\begin{cases} \frac{\partial}{\partial t} P(\vec{x}, t) + W_{max}P(\vec{x}, t) = \int_{\Omega} \widetilde{W}(\vec{x}' \rightarrow \vec{x}, t)P(\vec{x}', t)d\vec{x}' \\ P(\vec{x}, t)|_{t=0} = S(\vec{x}) \end{cases} \quad (5.18)$$

这里,

$$\widetilde{W}(\vec{x}' \rightarrow \vec{x}, t) = W(\vec{x}' \rightarrow \vec{x}, t) + (W_{max} - W_t(\vec{x}', t))\delta(\vec{x} - \vec{x}') \quad (5.19)$$

显然 $\int_{\Omega} \widetilde{W}(\vec{x}' - \vec{x}, t)d\vec{x} = W_{max} = \text{常数}$, ($\vec{x}' \in \Omega$)

利用假想的跃迁函数 $\widetilde{W}(\vec{x}' - \vec{x}, t)$, 与第一种情况完全类似。取 $\Delta t = \frac{1}{W_{max}}$ 这时有

$$\begin{cases} P_{m+1}(\vec{x}) = \int_{\Omega} \widetilde{W}^*(\vec{x}' - \vec{x}, t_m)P_m(\vec{x}')d\vec{x}' \\ P_0(\vec{x}) = S(\vec{x}) \end{cases} \quad m = 0, 1, 2, \dots \quad (5.20)$$

其中,

$$\widetilde{W}^*(\vec{x}' \rightarrow \vec{x}, t_m) = \frac{W(\vec{x}' \rightarrow \vec{x}, t)}{W_{max}} + \frac{W_{max} - W_t(\vec{x}', t)}{W_{max}} \delta(\vec{x} - \vec{x}') \quad (5.21)$$

则

$$\begin{aligned} & \int_{\Omega} \widetilde{W}^*(\vec{x}' \rightarrow \vec{x}, t_m) d\vec{x} \\ &= \int_{\Omega} \left\{ \frac{W(\vec{x}' \rightarrow \vec{x}, t_m)}{W_{max}} + \frac{W_{max} - W_t(\vec{x}', t_m)}{W_{max}} \delta(\vec{x} - \vec{x}') \right\} d\vec{x} \\ &= 1 + \int_{\Omega} \frac{W(\vec{x}' \rightarrow \vec{x}, t_m)}{W_{max}} d\vec{x} - \int \frac{W_t(\vec{x}', t_m)}{W_{max}} \delta(\vec{x} - \vec{x}') d\vec{x} \\ &= 1 + \frac{W_t(\vec{x}', t_m)}{W_{max}} - \frac{W_t(\vec{x}', t_m)}{W_{max}} \\ &= 1 \end{aligned} \quad (5.22)$$

因此可用第一种情况的抽样方法求解。所不同的是:

- ①以概率 $\frac{W_t(\vec{x}', t)}{W_{max}}$, 由 $W^*(\vec{x}' \rightarrow \vec{x}, t_m)$ 抽样 \vec{x} ;
- ②以概率 $1 - \frac{W_t(\vec{x}', t_m)}{W_{max}}$ 取 $\vec{x} = \vec{x}'$

以上见: 张孝译, 蒙特卡罗模拟主方程的一般技巧, 计算物理, 1(1984)133。

四. Nanbu 方法 (K.Nanbu, J.Phys.Soc.Japan, 52(1983) 2654)

从出发 $\rightarrow P(\vec{x}, t + \Delta t) = (1 - \Delta t W_t(\vec{x}, t)) P(\vec{x}, t) + \Delta t \int_{\Omega} W(\vec{x}' \rightarrow \vec{x}, t) P(\vec{x}', t) d\vec{x}'$

解的条件: $0 \leq 1 - \Delta t W_t(\vec{x}, t) < 1$

$$\begin{aligned} P(\vec{x}, t + \Delta t) &= \int_{\Omega} [1 - \Delta t W_t(\vec{x}', t)] \delta(\vec{x} - \vec{x}') P(\vec{x}', t) d\vec{x}' \\ &\quad + \int_{\Omega} \Delta t W_t(\vec{x}', t) \cdot \frac{W(\vec{x}' \rightarrow \vec{x}, t)}{W_t(\vec{x}', t)} P(\vec{x}', t) d\vec{x}' \end{aligned} \quad (5.23)$$

这里 $W_t(\vec{x}', t) = \int_{\Omega} W(\vec{x}' \rightarrow \vec{x}'', t) d\vec{x}''$ 是跃迁率, 即单位时间内发生跃迁的总数。令

$$\begin{aligned} W_N^*(\vec{x}' \rightarrow \vec{x}, t) &= [1 - \Delta t W_t(\vec{x}', t)] \delta(\vec{x} - \vec{x}') \\ &\quad + \Delta t W_t(\vec{x}', t) \cdot W^*(\vec{x}' \rightarrow \vec{x}, t) \end{aligned} \quad (5.24)$$

式中

$$W^*(\vec{x}' \rightarrow \vec{x}, t) = \frac{W(\vec{x}' \rightarrow \vec{x}, t)}{W_t(\vec{x}', t)} \quad (5.25)$$

则

$$P(\vec{x}, t + \Delta t) = \int_{\Omega} W_N^*(\vec{x}' \rightarrow \vec{x}, t) P(\vec{x}', t) d\vec{x}' \quad (5.26)$$

容易证明

$$W_N^*(\vec{x}' \rightarrow \vec{x}, t) \geq 0 \quad (5.27)$$

$$\int_{\Omega} W_N^*(\vec{x}' \rightarrow \vec{x}, t) d\vec{x} = 1, \quad \vec{x}' \in \Omega \quad (5.28)$$

步骤:

- ①在给定 \vec{x}' 后, 以概率 $p = \Delta t W_t(\vec{x}', t)$, 由 $W_N^*(\vec{x}' \rightarrow \vec{x}, t)$ 抽样 \vec{x}
- ②以概率 $1 - p$ 取 $\vec{x} = \vec{x}'$

§5.2 时间相关平均量 $\langle A_t(x) \rangle$ 的蒙特卡罗估计

由主方程的解确定状态的时间序列为

$$\begin{pmatrix} \vec{x}_0^{(i)}, & \vec{x}_1^{(i)}, & \dots, & \vec{x}_m^{(i)}, & \vec{x}_{m+1}^{(i)}, & \dots \\ t_0^{(i)}, & t_1^{(i)}, & \dots, & t_m^{(i)}, & t_{m+1}^{(i)}, & \dots \end{pmatrix} \quad (5.29)$$

其中

$$\begin{aligned} t_0^{(i)} &= 0, & i &= 1, 2, \dots, I \\ t_{m+1}^{(i)} &= t_m^{(i)} + \Delta t_{m+1}^{(i)} \end{aligned} \quad (5.30)$$

i 表示利用不同的初值 $(t_0^{(i)}, \vec{x}_0^{(i)})$ 做的抽样轮数。为得到固定 t 下 $\langle A_t(\vec{x}) \rangle$ 的 MC 估计值，必须得到固定 t 下 $P(\vec{x}, t)$ 的子样。

$$\langle A_t(\vec{x}) \rangle = \frac{1}{2\varepsilon I} \sum_{i=1}^I \sum_{m=0}^{\infty} A(\vec{x}_m^{(i)}) \Delta\varepsilon(t - t_m^{(i)}) \quad (5.31)$$

这里,

$$\Delta\varepsilon(t - t_m^{(i)}) = \begin{cases} 1, & |t - t_m^{(i)}| \leq \varepsilon \\ 0, & \text{else} \end{cases} \quad (5.32)$$

其中 ε 为预先给定的小正数。当 Δt_m 为常数情况下, $t = m_\Delta t$,

$$\langle A_t(\vec{x}) \rangle = \frac{1}{I} \sum_{i=1}^I A(\vec{x}_m^{(i)}) \quad (t = m_\Delta t) \quad (5.33)$$

相关系数和结构函数的计算 (与实验结果进行比较)

由主方程确定了一个 Markov 过程, 因此, 当 $t \geq t_0$ 时,

$$P(\vec{x}, t) = \int_{\Omega} K(\vec{x}' \rightarrow \vec{x}; t, t_0) P(\vec{x}', t_0) d\vec{x}' \quad (5.34)$$

在时刻 t

$$\langle A_t(\vec{x}) \rangle = \int P(\vec{x}, t) A(\vec{x}) d\vec{x} = \int_{\Omega} \int_{\Omega} A(\vec{x}) K(\vec{x}' \rightarrow \vec{x}; t, t_0) P(\vec{x}', t_0) d\vec{x}' d\vec{x} \quad (5.35)$$

相关量:

$$\langle A_t(\vec{x}) B_{t_0}(\vec{x}') \rangle = \int_{\Omega} \int_{\Omega} A(\vec{x}) K(\vec{x}' \rightarrow \vec{x}; t, t_0) P(\vec{x}', t_0) B(\vec{x}') d\vec{x} d\vec{x}' \quad (5.36)$$

$A_t(\vec{x} + \vec{x}_0)$ 与 $B_{t_0}(\vec{x})$ 的相关函数

$$\begin{aligned} G_{AB}(\vec{x}_0, t_0) &= \langle [A_t(\vec{x}' + \vec{x}_0) - \langle A_t(\vec{x}' + \vec{x}_0) \rangle] [B_{t_0}(\vec{x}) - \langle B_{t_0}(\vec{x}) \rangle] \rangle \\ &= \langle A_t(\vec{x}' + \vec{x}_0) B_{t_0}(\vec{x}') \rangle - \langle A_t(\vec{x}' + \vec{x}_0) \rangle \langle B_{t_0}(\vec{x}') \rangle \end{aligned} \quad (5.37)$$

结构函数

$$S(\vec{k}, t) = \int_{\Omega} G(\vec{x}, t) e^{-i\vec{k} \cdot \vec{x}} d\vec{x} \quad (5.38)$$

§5.3 主方程的直接蒙特卡罗模拟

一. 主方程:

$$\begin{cases} \frac{\partial P(\vec{x}, t)}{\partial t} + W_t(\vec{x}, t)P(\vec{x}, t) = \int_{\Omega} W(\vec{x}' \rightarrow \vec{x}, t)P(\vec{x}', t)d\vec{x}' = Q(\vec{x}, t) \\ P(\vec{x}, t)|_{t=0} = S(\vec{x}) \end{cases} \quad (5.39)$$

方程的形式解:

$$P(\vec{x}, t) = \int_0^t Q(\vec{x}, t')e^{-\int_{t'}^t W_t(\vec{x}, t_s)dt_s}dt' + S(\vec{x})e^{-\int_0^t W_t(\vec{x}, t_s)dt_s} \quad (5.40)$$

引入函数

$$\eta(s) = \begin{cases} 1, & s \geq 0 \\ 0, & s < 0 \end{cases} \quad (5.41)$$

则

$$P(\vec{x}, t) = \int_0^\infty [Q(\vec{x}, t') + \delta(t')S(\vec{x})]e^{-\int_{t'}^t W_t(\vec{x}, t_s)dt_s} \cdot \eta(t - t')dt' \quad (5.42)$$

代入

$$\begin{aligned} Q(\vec{x}, t) &= \int_{\Omega} W(\vec{x}' \rightarrow \vec{x}, t)P(\vec{x}', t)d\vec{x}' \\ &= \int_{\Omega} \left\{ \int_0^\infty [Q(\vec{x}', t') + \delta(t')S(\vec{x}')]e^{-\int_{t'}^t W_t(\vec{x}', t_s)dt_s} \right. \\ &\quad \left. \cdot \eta(t - t')W(\vec{x}' \rightarrow \vec{x}, t)dt' \right\} d\vec{x}' \end{aligned} \quad (5.43)$$

令

$$\overline{Q}(\vec{x}, t) = Q(\vec{x}, t) + S(\vec{x})\delta(t) \quad (5.44)$$

得到主方程的积分形式

$$\begin{cases} \overline{Q}(\vec{x}, t) = \int_{\Omega} \left\{ \int_0^\infty \overline{Q}(\vec{x}', t')e^{-\int_{t'}^t W_t(\vec{x}', t_s)dt_s} \cdot \eta(t - t')W(\vec{x}' \rightarrow \vec{x}, t)dt' \right\} d\vec{x}' \\ P(\vec{x}, t) = \int_0^\infty \overline{Q}(\vec{x}, t')e^{-\int_{t'}^t W_t(\vec{x}, t_s)dt_s} \cdot \eta(t - t')dt' \end{cases} \quad (5.45)$$

二. 模拟过程

(1) $t_0 = 0$, \vec{x}_0 由初始分布 $S(\vec{x})$ 抽样。

(2) 当 (\vec{x}_m, t_m) 给定后, (\vec{x}_{m+1}, t_{m+1}) 由分布

$$W_t(\vec{x}_m, t_{m+1})e^{-\int_{t_m}^{t_{m+1}} W_t(\vec{x}_m, t_s)dt_s} \eta(t_{m+1} - t_m) \cdot \frac{W(\vec{x}_m \rightarrow \vec{x}_{m+1}, t_{m+1})}{W_t(\vec{x}_m, t_{m+1})} \quad (5.46)$$

抽样得到, 也就是:

由 $\int_{t_m}^{t_{m+1}} W_t(\vec{x}_m, t_s)dt_s = -\ln \xi$ 决定 t_{m+1} 。 ($e^{-\int_{t_m}^{t_{m+1}} W_t(\vec{x}_m, s)ds} = \xi$)

由 $\frac{W(\vec{x}_m \rightarrow \vec{x}_{m+1}, t_{m+1})}{W_t(\vec{x}_m, t_{m+1})}$ 决定 \vec{x}_{m+1}

则

$$\langle A_t(\vec{x}) \rangle = \frac{1}{I} \sum_{i=1}^I \sum_{m=0}^\infty A(\vec{x}_m^{(i)})e^{-\int_{t_m^{(i)}}^t W_t(\vec{x}_m^{(i)}, t_s)dt_s} \cdot \eta(t - t_m^{(i)}) \quad (5.47)$$

第六章 MC 求解随机微分方程 (Langevin Equations)

§6.1 噪声与随机力 (noise and random force)

1. 什么叫随机动力学 Stochastic dynamics

$$\text{力: } m\ddot{x}, \quad \beta\dot{x}, \quad -U'(x), \quad \xi(t)$$

布朗运动

2. 内部噪声 (internal noise) : 涨落与耗散来自于液体分子对布朗粒子的无规碰撞。

$$D = \beta k_B T, \quad \langle E_k \rangle = \langle \frac{1}{2} m \dot{x}^2 \rangle + \frac{1}{2} m \omega_0^2 \langle x^2 \rangle = k_B T \quad (6.1)$$

外部噪声 (external noise) : 由外部输入或由外部参数控制的 Fluctuation.

3. 白噪声与色噪声 (Gaussian distribution)

白噪声, 谱

$$S(\omega) = C \quad (6.2)$$

$$\langle \xi(t) \rangle = 0 \quad (6.3)$$

$$\langle \xi(t) \xi(s) \rangle = 2\delta(t-s) \quad (6.4)$$

$$S(\omega) = \int_0^\infty d\omega \cos \omega t \langle \xi(t) \xi(s) \rangle = C \quad (6.5)$$

色噪声, 谱

$$\langle \xi(t) \xi(s) \rangle = \frac{1}{\tau_c} e^{-\frac{|t-s|}{\tau_c}} \quad (6.6)$$

平衡过程

$$S = \int_0^\infty \cos \omega t \frac{1}{\tau_c} e^{-\frac{t}{\tau_c}} d\omega = \frac{1}{1 + \omega^2 \tau_c^2} \quad (6.7)$$

τ_c : 色噪声关联时间,

$$\tau_c \rightarrow 0, \quad \langle \xi(t) \xi(s) \rangle \rightarrow 2\delta(t-s) \quad (6.8)$$

4. 无规力: 出现在动力学方程中, 驱动粒子无规运动 (扩散, 迁移) 的动力

$$m\ddot{x} = -\beta\dot{x} - \frac{\partial U(x)}{\partial x} + \xi(t) \quad (6.9)$$

$$\langle \xi(t) \xi(s) \rangle = 2\beta k_B T \delta(t-s) \quad (6.10)$$

涨落耗散定理。

§6.2 随机泰勒展开

一. Euler 方法 (1/2 阶)

$$\dot{x}(t) = f(x) + \xi(t) \quad \langle \xi(t) \rangle = 0 \quad \langle \xi(t)\xi(s) \rangle = 2D\delta(t-s) \quad (6.11)$$

$$\Rightarrow \int_t^{t+\Delta t} \dot{x}(s)ds = \int_t^{t+\Delta t} f(x(s))ds + \int_t^{t+\Delta t} \xi(s)ds \quad (6.12)$$

$$\Rightarrow x(t+\Delta t) - x(t) = f(x(t))\Delta t + \int_t^{t+\Delta t} \xi(s)ds \quad (6.13)$$

令

$$\Gamma(t) = \int_t^{t+\Delta t} \xi(s)ds \quad (6.14)$$

$$\langle \Gamma(t) \rangle = 0 \quad (6.15)$$

$$\begin{aligned} \langle \Gamma^2(t) \rangle &= \int_t^{t+\Delta t} \int_t^{t+\Delta t} \langle \xi(s_1)\xi(s_2) \rangle ds_1 ds_2 \\ &= 2D \int_t^{t+\Delta t} \int_t^{t+\Delta t} \delta(s_1 - s_2) ds_1 ds_2 \\ &= 2D \int_t^{t+\Delta t} ds_2 \\ &= 2D\Delta t \end{aligned} \quad (6.16)$$

$$\Gamma(t) = \sqrt{2D\Delta t} \omega_t \quad (6.17)$$

其中, ω_t 是标准高斯随机数, $\langle \omega_t \rangle = 0$, $\langle \omega_t^2 \rangle = 1$ 。

二. 3/2 阶

$$\begin{aligned} x(t+\Delta t) - x(t) &= \int_t^{t+\Delta t} [f(x(t)) + f'(x(t))(x(s) - x(t))]ds + \int_t^{t+\Delta t} \xi(s)ds \\ &= f(x(t))\Delta t + f'(x(t)) \int_t^{t+\Delta t} [x(s) - x(t)]ds + \Gamma_1(t) \end{aligned} \quad (6.18)$$

$$x(s) - x(t) = f(x(t))(s-t) + \int_t^s \xi(s')ds' \quad (6.19)$$

$$x(t+\Delta t) - x(t) = f(x(t))\Delta t + \frac{1}{2}f'(x(t))f(x(t))\Delta t^2 \quad (6.20)$$

$$+ \Gamma_1(t) + f'(x(t))\Gamma_2(t) \quad (6.21)$$

$$\Gamma_1(t) = \sqrt{2D\Delta t} \omega_1 \quad (6.22)$$

$$\Gamma_2(t) = \int_t^{t+\Delta t} ds \int_t^s \xi(s')ds' \quad (6.23)$$

$$\langle \Gamma_2(t) \rangle = 0 \quad (6.24)$$

$$\begin{aligned} \langle \Gamma_2^2(t) \rangle &= \int_t^{t+\Delta t} ds_1 \int_t^{t+\Delta t} ds_2 \int_t^{s_1} \int_t^{s_2} \langle \xi(s'_1)\xi(s'_2) \rangle ds'_1 ds'_2 \\ &= 2D \int_t^{t+\Delta t} ds_1 \int_t^{t+\Delta t} ds_2 \int_t^{s_1} \int_t^{s_2} \delta(s'_1 - s'_2) ds'_1 ds'_2 \\ &= \frac{2}{3}D\Delta t^3 \end{aligned} \quad (6.25)$$

三. 乘性噪声

$$\dot{x}(t) = f(x) + g(x)\xi(t) \quad (6.26)$$

$$\langle \xi(t) \rangle = 0 \quad \langle \xi(t)\xi(s) \rangle = 2D\delta(t-s) \quad (6.27)$$

$$\begin{aligned} \int_t^{t+\Delta t} \dot{x}(s)ds &= \int_t^{t+\Delta t} f(x(s))ds + \int_t^{t+\Delta t} g(x(s))\xi(s)ds \\ &= \int_t^{t+\Delta t} [f(x(s)) + \frac{\partial f}{\partial x}(x(s) - x(t))]ds \\ &\quad + \int_t^{t+\Delta t} [g(x(t)) + \frac{\partial g}{\partial x}(x(s) - x(t))]\xi(s)ds \end{aligned} \quad (6.28)$$

利用

$$x(s) - x(t) = f(x(t))(s-t) + g(x(t)) \int_t^s \xi(s_1)ds_1 \quad (6.29)$$

得

$$\begin{aligned} &x(t+\Delta t) - x(t) \\ &= f(x(t))\Delta t + \frac{1}{2}\frac{\partial f}{\partial x}f(x(t))\Delta t^2 \\ &\quad + g(x(t))\frac{\partial f}{\partial x} \int_t^{t+\Delta t} ds \int_t^s \xi(s_1)ds_1 + g(x(t)) \int_t^{t+\Delta t} \xi(s)ds \\ &\quad + \frac{\partial g}{\partial x} \int_t^{t+\Delta t} [f(x(t))(s-t) + g(x(t)) \int_t^s \xi(s_1)ds_1]\xi(s)ds \end{aligned} \quad (6.30)$$

$$\begin{aligned} &= f(x(t))\Delta t + \frac{1}{2}\frac{\partial f}{\partial x}f(x(t))\Delta t^2 \\ &\quad + g(x(t))\frac{\partial f}{\partial x}\Gamma_2(t) + g(x(t))\Gamma_1(t) \\ &\quad + f\frac{\partial g}{\partial x} \int_t^{t+\Delta t} (s-t)\xi(s)ds + g\frac{\partial g}{\partial x} \int_t^{t+\Delta t} \xi(s)ds \int_t^s \xi(s_1)ds_1 \end{aligned} \quad (6.31)$$

最后两项:

$$\begin{aligned} a) &= f\frac{\partial g}{\partial x} \int_t^{t+\Delta t} (s-t)d\Gamma_1(s) = f\frac{\partial g}{\partial x} [\Gamma_1(s)(s-t)|_t^{t+\Delta t} - \int_t^{t+\Delta t} \Gamma_1(s)ds] \\ &= f\frac{\partial g}{\partial x} [\Delta t\Gamma_1(t+\Delta t) - \Gamma_2(t)] = f\frac{\partial g}{\partial x} [\Delta t\Gamma_1(t) - \Gamma_2(t)] \end{aligned} \quad (6.32)$$

$$\begin{aligned} b) &= g\frac{\partial g}{\partial x} \int_t^{t+\Delta t} \xi(s)ds\Gamma_1(s) = g\frac{\partial g}{\partial x} \int_t^{t+\Delta t} \Gamma_1(s)d\Gamma_1(s) \\ &= \frac{g}{2}\frac{\partial g}{\partial x} [\Gamma_1^2(t+\Delta t) - \Gamma_1^2(t)] = 0 \end{aligned} \quad (6.33)$$

四. 一般阻尼情况

$$m\ddot{x} + \beta\dot{x} + U'(x) = \Gamma(t) \quad (6.34)$$

$$\langle \Gamma(t) \rangle = 0 \quad \langle \Gamma(t)\Gamma(s) \rangle = \sqrt{2\beta k_B T} \delta(t-s) \quad (6.35)$$

令

$$f(x) = -\frac{\partial U}{\partial x}, \quad v = \dot{x} \quad (6.36)$$

$$\begin{cases} \dot{x} = v \\ \dot{v} = -\frac{\beta}{m}v + \frac{1}{m}f(x) + \frac{1}{m}\Gamma(t) \end{cases} \quad (6.37)$$

欧拉方法（最低阶）

$$\begin{cases} x(t') - x(t) = v(t') - v(t) \\ v(t') - v(t) = -\frac{\beta}{m}v(t)(t' - t) + \frac{1}{m}f(x(t))(t' - t) + \frac{1}{m} \int_t^{t'} \Gamma(s')ds' \end{cases} \quad (6.38)$$

$$\begin{aligned} & x(t + \Delta t) - x(t) \\ &= \int_t^{t+\Delta t} v(s)ds \\ &= \int_t^{t+\Delta t} [v(t) + v(s) - v(t)]ds \\ &= v(t)\Delta t - \frac{\beta}{2m}v(t)\Delta t^2 + \frac{1}{2m}f(x(t))\Delta t^2 + \frac{1}{m} \int_t^{t+\Delta t} ds \int_t^s \Gamma(s')ds' \end{aligned} \quad (6.39)$$

$$\begin{aligned} & v(t + \Delta t) - v(t) \\ &= -\frac{\beta}{m} \int_t^{t+\Delta t} [v(t) + v(s) - v(t)]ds \\ &\quad + \frac{1}{m} \int_t^{t+\Delta t} [f(x(t)) + f'(x(t))(x(s) - x(t))]ds + \frac{1}{m} \int_t^{t+\Delta t} \Gamma(s)ds \\ &= -\frac{\beta}{m}v(t)\Delta t - \frac{\beta}{m} \int_t^{t+\Delta t} (v(s) - v(t))ds \\ &\quad + \frac{1}{m}f(x(t))\Delta t + \frac{1}{2m}f'(x(t))v(x(t))(\Delta t)^2 + \frac{1}{m} \int_t^{t+\Delta t} \Gamma(s)ds \end{aligned} \quad (6.40)$$

$$\begin{aligned} & \int_t^{t+\Delta t} (v(s) - v(t))ds \\ &= -\frac{\beta}{2m}v(x(t))\Delta t^2 + \frac{1}{2m}f(x(t))\Delta t^2 + \frac{1}{m} \int_t^{t+\Delta t} ds \int_t^s \Gamma(s')ds' \end{aligned} \quad (6.41)$$

$$\text{令 } \Gamma_1(t) = \int_t^{t+\Delta t} \Gamma(s)ds \quad \Gamma_2(t) = \int_t^{t+\Delta t} ds \int_t^s \xi(s')ds'$$

则

$$\begin{cases} x(t + \Delta t) = x(t) + v(t)\Delta t - \frac{\beta}{2m}v(x(t))\Delta t^2 + \frac{1}{2m}f(x(t))\Delta t^2 + \frac{1}{m}\Gamma_2(t) \\ v(t + \Delta t) = v(t) - \frac{\beta}{m}v(t)\Delta t + \frac{1}{m}f(x(t))\Delta t + \frac{1}{2}(\frac{\beta}{m})^2v(t)\Delta t^2 - \frac{\beta}{2m}f(x(t))\Delta t^2 \\ \quad + \frac{1}{2m}f'(x(t))v(t)\Delta t^2 + \frac{1}{m}\Gamma_1(t) - \frac{\beta}{m^2}\Gamma_2(t) \end{cases} \quad (6.42)$$

$$\begin{cases} \Gamma_1(t) = \sqrt{2\beta k_B T \Delta t} \omega_1 \\ \Gamma_2(t) = \sqrt{2\beta k_B T \Delta t} \Delta t (\frac{1}{2}\omega_1 + \frac{1}{2\sqrt{3}}\omega_2) \end{cases} \quad (6.43)$$

§6.3 Fokker-Planck 方程 (FPE)

·Langevin equation(LE)

$$\dot{x}(t) = a(x, t) + b(x, t)\Gamma(t) \quad \langle \Gamma(t) \rangle = 0 \quad \langle \Gamma(t)\Gamma(t') \rangle = \delta(t - t') \quad (6.44)$$

· 等价的 FP 方程

$$\frac{\partial P(x, t)}{\partial t} = -\frac{\partial}{\partial x}[D^{(1)}(x, t)P(x, t)] + \frac{\partial^2}{\partial x^2}(D^{(2)}(x, t)P(x, t)) \quad (6.45)$$

其中 $D^{(1)}$ 是漂移系数, $D^{(2)}$ 是扩散系数

· Kramers-Moyal 展开系数 $D^n(x, t)$ 定义

$$D^{(n)}(x, t) = \frac{1}{n!} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle (x(t + \Delta t) - x(t))^n \rangle \quad n = 1, 2, \dots \quad (6.46)$$

由 LE 出发,

$$x(t + \Delta t) - x(t) = \int_t^{t+\Delta t} a(x(t'), t') dt' + \int_t^{t+\Delta t} b(x(t'), t') \Gamma(t') dt' \quad (6.47)$$

将 $a(x(t'), t')$ 和 $b(x(t'), t')$ 展开有

$$a(x(t'), t') = a(x(t), t') + a'_x(x(t), t')(x(t') - x(t)) + \dots \quad (6.48)$$

$$b(x(t'), t') = b(x(t), t') + b'_x(x(t), t')(x(t') - x(t)) + \dots \quad (6.49)$$

代入 LE

$$\begin{aligned} & x(t + \Delta t) - x(t) \\ = & \int_t^{t+\Delta t} a(x(t), t') dt' + \int_t^{t+\Delta t} a'_x(x(t), t')(x(t') - x(t)) dt' + \dots \\ & + \int_t^{t+\Delta t} b(x(t), t') \Gamma(t') dt' + \int_t^{t+\Delta t} b'_x(x(t), t')(x(t') - x(t)) \Gamma(t') dt' + \dots \end{aligned} \quad (6.50)$$

再对 $x(t') - x(t)$ 进行一次叠代后得到

$$x(t') - x(t) = \int_t^{t'} a(x(t), t'') dt'' + \int_t^{t'} b(x(t), t'') \Gamma(t'') dt'' \quad (6.51)$$

$$\begin{aligned} & x(t + \Delta t) - x(t) \\ = & \int_t^{t+\Delta t} a(x(t), t') dt' + \int_t^{t+\Delta t} a'_x(x(t), t') dt' \int_t^{t'} a(x(t), t'') dt'' \\ & + \int_t^{t+\Delta t} a'_x(x(t), t') dt' \int_t^{t'} b(x(t), t'') \Gamma(t'') dt'' \\ & + \int_t^{t+\Delta t} b(x(t), t') \Gamma(t') dt' + \int_t^{t+\Delta t} b'_x(x(t), t') \Gamma(t') dt' \int_t^{t'} a(x(t), t'') dt'' \\ & + \int_t^{t+\Delta t} b'_x(x(t), t') \Gamma(t') dt' \int_t^{t'} b(x(t), t'') \Gamma(t'') dt'' \end{aligned} \quad (6.52)$$

注意: 上式右端含有奇数个 $\Gamma(t)$ 的项, 其均值为零, 含有偶数个 $\Gamma(t)$ 的项除 $\int_t^{t+\Delta t} b'_x(x(t), t') \Gamma(t') dt' \int_t^{t'} b(x(t), t'') \Gamma(t'') dt''$ 外, 他们的均值至少是 $o(\Delta t)^2$, 因此

$$\begin{aligned} & \langle x(t + \Delta t) - x(t) \rangle \\ = & \int_t^{t+\Delta t} a(x(t), t') dt' \\ & + \langle \int_t^{t+\Delta t} b'_x(x(t), t') \Gamma(t') dt' \int_t^{t'} b(x(t), t'') \Gamma(t'') dt'' \rangle + o(\Delta t^2) \end{aligned} \quad (6.53)$$

$$\text{由于第二项} = \begin{cases} 0 & \text{在随机积分意义下} \\ \frac{1}{2}b'_x(x(t),t)b(x(t),t)\Delta t & \text{在 Stratonovich 意义下} \end{cases} \quad (6.54)$$

则

$$D^{(1)}(x, t) = \begin{cases} a(x, t) & \hat{I}to\text{意义下} \\ a(x, t) + \frac{1}{2}b'_x(x, t)b(x, t) & \text{Stratonovich意义下} \end{cases} \quad (6.55)$$

根据同样分析

$$\begin{aligned} \langle (x(t+\Delta t) - x(t))^2 \rangle &= \langle \left(\int_t^{t+\Delta t} b(x(t), t')\Gamma(t')dt' \right)^2 \rangle + o(\Delta t)^2 \\ &= \Delta t b^2(x(t), t) + o(\Delta t)^2 \end{aligned} \quad (6.56)$$

因而有

$$D^{(2)}(x, t) = \frac{1}{2!} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle (x(t+\Delta t) - x(t))^2 \rangle = \frac{1}{2}b^2(x, t) \quad (6.57)$$

同理可得

$$D^{(n)}(x, t) = \frac{1}{n!} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle (x(t+\Delta t) - x(t))^n \rangle = 0, \quad n \geq 3 \quad (6.58)$$

§6.4 色噪声驱动 LE

$$\begin{cases} \dot{x}(t) = f(x) + y(t) & (1) \\ \dot{y} = -\frac{1}{\tau}y + \frac{\sqrt{2D}}{\tau}\xi(t) & (2) \end{cases} \quad (6.59)$$

$$\langle \xi(t) \rangle = 0, \quad \langle \xi(t)\xi(s) \rangle = \delta(t-s) \quad (6.60)$$

则

$$\langle y(t) \rangle = 0 \quad \{ \langle y(t)y(s) \rangle \} = \frac{D}{\tau} e^{-\frac{|t-s|}{\tau}} \quad (6.61)$$

$$y(t) \text{ is Gaussian } \langle (y(0))^2 \rangle = \frac{D}{\tau}$$

(2) 式的解

$$y(t) = e^{-t/\tau} y(0) + \frac{\sqrt{2D}}{\tau} \int_0^t e^{\frac{s-t}{\tau}} \xi(s) ds \quad (6.62)$$

则递推叠代关系

$$y(t+\Delta t) = e^{-\frac{\Delta t}{\tau}} y(t) + \frac{\sqrt{2D}}{\tau} \omega_0 \quad (6.63)$$

$$\omega_0 = \int_t^{t+\Delta t} e^{\frac{s-t-\Delta t}{\tau}} \xi(s) ds \quad (6.64)$$

定义

$$\omega_1 \equiv \int_t^{t+\Delta t} dt' \int_t^{t'} e^{\frac{s-t'}{\tau}} \xi(s) ds \quad (6.65)$$

$$\omega_2 \equiv \int_t^{t+\Delta t} dt' \int_t^{t'} ds \int_t^s e^{\frac{s'-s}{\tau}} \xi(s') ds' \quad (6.66)$$

$$\begin{aligned}
x(t + \Delta t) - x(t) &= \int_t^{t+\Delta t} f(x(t'))dt' + \int_t^{t+\Delta t} y(t')dt' \\
&= f(x(t))\Delta t + \frac{\partial f}{\partial x} \int_t^{t+\Delta t} (x(t') - x(t))dt' + \int_t^{t+\Delta t} y(t')dt'
\end{aligned} \tag{6.67}$$

其中，第二项：

$$\begin{aligned}
\int_t^{t+\Delta t} (x(t') - x(t))dt' &= \frac{1}{2}f(x(t))(\Delta t)^2 + \int_t^{t+\Delta t} dt' \int_t^{t'} y(s)ds \\
&= \frac{1}{2}f(\Delta t)^2 + \tau^2(\Delta t/\tau + e^{-\frac{\Delta t}{\tau}} - 1)y(t) + \frac{\sqrt{2D}}{\tau}\omega_2
\end{aligned} \tag{6.68}$$

第三项：

$$\begin{aligned}
\int_t^{t+\Delta t} y(t')dt' &= \int_t^{t+\Delta t} [e^{-\frac{t'-t}{\tau}}y(t) + \frac{\sqrt{2D}}{\tau} \int_t^{t'} e^{\frac{s-t'}{\tau}}\xi(s)ds]dt' \\
&= \tau[1 - e^{-\frac{\Delta t}{\tau}}]y(t) + \frac{\sqrt{2D}}{\tau}\omega_1
\end{aligned} \tag{6.69}$$

整理得：

$$\begin{cases} y(t + \Delta t) = e^{-\frac{\Delta t}{\tau}}y(t) + \frac{\sqrt{2D}}{\tau}\omega_0 \\ x(t + \Delta t) = x(t) + \tau[1 - e^{-\frac{\Delta t}{\tau}}]y(t) + \frac{\sqrt{2D}}{\tau}\omega_1 + f(x(t))\Delta t + \frac{1}{2}(\Delta t)^2 f(x(t))f'(x(t)) \\ \quad + f'(x(t))[\tau^2(\frac{\Delta t}{\tau} + e^{-\frac{\Delta t}{\tau}} - 1)y(t) + \frac{\sqrt{2D}}{\tau}\omega_2] \end{cases} \tag{6.70}$$

随机变量 $\omega_0, \omega_1, \omega_2$ 的自关联和交叉关联为：

$$\langle \omega_0^2 \rangle = \frac{\tau}{2}(1 - e^{-2\alpha}) \quad \alpha = \Delta t/\tau \tag{6.71}$$

$$\langle \omega_1^2 \rangle = \tau^3(2\alpha - 3 - e^{-2\alpha} + 4e^{-\alpha}) \tag{6.72}$$

$$\langle \omega_2^2 \rangle = \frac{\tau^5}{2}(1 + 2\alpha - 2\alpha^2 + \frac{2}{3}\alpha^3 - 4\alpha e^{-\alpha} - e^{-2\alpha}) \tag{6.73}$$

$$\langle \omega_0\omega_1 \rangle = \frac{\tau^2}{2}(1 - 2e^{-\alpha} + e^{-2\alpha}) \tag{6.74}$$

$$\langle \omega_0\omega_2 \rangle = \frac{\tau^3}{2}(1 - 2\alpha e^{-\alpha} - e^{-2\alpha}) \tag{6.75}$$

$$\langle \omega_1\omega_2 \rangle = \frac{\tau^4}{2}(1 + \alpha^2 - 2\alpha + 2\alpha e^{-\alpha} - 2e^{-\alpha} + e^{-2\alpha}) \tag{6.76}$$

$$\begin{cases} \omega_0 = \sqrt{\langle \omega_0^2 \rangle} Y_0 \\ \omega_1 = \frac{\langle \omega_0\omega_1 \rangle}{\sqrt{\langle \omega_0^2 \rangle}} Y_0 + [\langle \omega_1^2 \rangle - \frac{(\langle \omega_0\omega_1 \rangle)^2}{\langle \omega_0^2 \rangle}]^{\frac{1}{2}} Y_1 \\ \omega_2 = AY_0 + BY_1 + CY_2 \end{cases} \tag{6.77}$$

$$\langle Y_i \rangle = 0, \quad \langle Y_i Y_j \rangle = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad i, j = 0, 1, 2 \tag{6.78}$$

作业：求 $\langle \omega_1^2 \rangle, \langle \omega_2^2 \rangle$ 和噪声系数 A,B,C

§6.5 随机龙格 - 库塔算法

一. 常微分方程的二阶 R - K 算法

$$\dot{x} = F(x) \quad (6.79)$$

$$x(t + \Delta t) = x(t) + \frac{1}{2}\Delta t[F_1 + F_2] \quad (6.80)$$

式中

$$F_1 = F(x(t)) \quad (6.81)$$

$$F_2 = F(x(t) + \Delta t F_1) = F(x(t)) + F'(x(t))\Delta t F_1 \quad (6.82)$$

$$\begin{aligned} x(t + \Delta t) &= x(t) + \frac{1}{2}[2F + F'(x(t))\Delta t F]\Delta t \\ &= x(t) + F\Delta t + \frac{1}{2}F'F\Delta t^2 \quad 2 \text{ 阶} \end{aligned} \quad (6.83)$$

二. 含白噪声的随机微分方程

$$\dot{x} = f(x) + \xi(t) \quad \langle \xi(t) \rangle = 0 \quad \langle \xi(t)\xi(s) \rangle = 2D\delta(t-s) \quad (6.84)$$

$$\begin{aligned} x(t + \Delta t) - x(t) &= \int_t^{t+\Delta t} f(x(s))ds + \int_t^{t+\Delta t} \xi(s)ds \\ &= \frac{1}{2}[F_1 + F_2]\Delta t + \sqrt{2D\Delta t} \phi_0 \end{aligned} \quad (6.85)$$

其中

$$F_1 = f(x(t) + \sqrt{2D\Delta t} \phi_1) \quad (6.86)$$

$$F_2 = f(x(t) + \Delta t F_1 + \sqrt{2D\Delta t} \phi_2) \quad (6.87)$$

ϕ_0, ϕ_1, ϕ_2 是三个互相独立的标准高斯变量

展开

$$F_1 = f(x(t)) + f'(x(t))\sqrt{2D\Delta t} \phi_1 \quad (6.88)$$

$$\begin{aligned} F_2 &= f(x(t)) + f'(x(t))[\Delta t F_1 + \sqrt{2D\Delta t} \phi_2] \\ &= f(x(t)) + f'(x(t))[\Delta t f(x(t)) + \Delta t f'(x(t))\sqrt{2D\Delta t} \phi_1] \\ &\quad + f'(x(t))\sqrt{2D\Delta t} \phi_2 \end{aligned} \quad (6.89)$$

$$x(t + \Delta t) = x(t) + \Delta t f(x(t)) + \frac{1}{2}(\Delta t)^2 f'f' + R(t) \quad (6.90)$$

$$\begin{aligned} R(t) &= \sqrt{2D\Delta t} \phi_0 + \frac{1}{2}\Delta t \sqrt{2D\Delta t} f'(x(t))(\phi_1 + \phi_2) \\ &\quad + \frac{1}{2}(\Delta t)^2 f'^2 \sqrt{2D\Delta t} \phi_1 \end{aligned} \quad (6.91)$$

$$\langle R(t) \rangle = 0 \quad (6.92)$$

$$\langle R^2(t) \rangle = 2D\Delta t \langle \phi_0^2 \rangle + 2D(\Delta t)^2 f'(x(t)) \langle \phi_0(\phi_1 + \phi_2) \rangle + o(\Delta t^2) \quad (6.93)$$

$$\langle \phi_0^2 \rangle = 1 \quad \langle \phi_0(\phi_1 + \phi_2) \rangle = 1/2 \quad (6.94)$$

$$R(t) = \sqrt{2D_{\Delta}t} \omega_1 + f'(x(t))\sqrt{2D_{\Delta}t} \Delta t \left(\frac{1}{2}\omega_1 + \frac{1}{2\sqrt{3}}\omega_2 \right) \quad (6.95)$$

$$\langle R^2(t) \rangle = 2D_{\Delta}t + D(\Delta t)^2 f'(x(t)) + o(\Delta t^2) \quad (6.96)$$

小结:

$$\begin{cases} x(t + \Delta t) = x(t) + \frac{1}{2}\Delta t[F_1 + F_2] + \sqrt{2D_{\Delta}t} \psi \\ F_1 = f(x(t)) \\ F_2 = f(x(t) + \Delta t F_1 + \sqrt{2D_{\Delta}t} \psi) \end{cases} \quad (6.97)$$

$$\begin{cases} x(t + \Delta t) = x(t) + \frac{1}{2}\Delta t(f(x(t)) + f(x^*(t))) + \sqrt{2D_{\Delta}t} \omega \\ x^*(t) = x(t) + \Delta t f(x(t)) + \sqrt{2D_{\Delta}t} \omega \end{cases} \quad (6.98)$$

验证:

$$\begin{aligned} x(t + \Delta t) &= x(t) \\ &\quad + \frac{1}{2}\Delta t\{f(x(t)) + f(x(t)) + f'(x(t))[\Delta t f(x(t)) + \sqrt{2D_{\Delta}t} \psi]\} \\ &\quad + \sqrt{2D_{\Delta}t} \psi \\ &= x(t) + \Delta t f + \frac{1}{2}(\Delta t)^2 f f' + \frac{1}{2}\Delta t f' \sqrt{2D_{\Delta}t} \psi + \sqrt{2D_{\Delta}t} \psi \end{aligned} \quad (6.99)$$

$$R(t) = (\sqrt{2D_{\Delta}t} + \frac{1}{2}\Delta t f' \sqrt{2D_{\Delta}t}) \psi \quad (6.100)$$

$$\langle R^2 \rangle = 2D_{\Delta}t + (\Delta t)^2 D f'(x(t)) + o(\Delta t^2) \quad (6.101)$$

三. 含噪声 (多维) LE 的 R-K 算法

$$\begin{cases} \dot{x} = f(x, \varepsilon) & f(x, \varepsilon) = f(x) + \varepsilon \\ \dot{\varepsilon} = h(\varepsilon) + \lambda g_{\omega}(t) & h(\varepsilon) = -\lambda \varepsilon, \quad \lambda = \frac{1}{\tau} \end{cases} \quad (6.102)$$

$$\langle g_{\omega}(t) \rangle = 0 \quad \langle g_{\omega}(t) g_{\omega}(s) \rangle = 2D\delta(t - s) \quad (6.103)$$

Stochastic RKII algorithm:

$$\begin{cases} x(t + \Delta t) = x(t) + \frac{1}{2}\Delta t(F_1 + F_2) \\ \varepsilon(t + \Delta t) = \varepsilon(t) + \frac{1}{2}\Delta t(H_1 + H_2) + (2D\lambda_{\Delta}^2 t)^{\frac{1}{2}} \psi \end{cases} \quad (6.104)$$

式中

$$H_1 = h(\varepsilon(t)) \quad (6.105)$$

$$H_2 = h[\varepsilon(t) + \Delta t H_1 + \sqrt{2D\lambda_{\Delta}^2 t} \psi] \quad (6.106)$$

$$F_1 = f(x(t), \varepsilon(t)) \quad (6.107)$$

$$F_2 = f[x(t) + \Delta t F_1, \varepsilon(t) + \Delta t H_1 + \sqrt{2D\lambda_{\Delta}^2 t} \psi] \quad (6.108)$$

$\langle \psi \rangle = 0, \langle \psi^2 \rangle = 1$ 标准高斯随机数

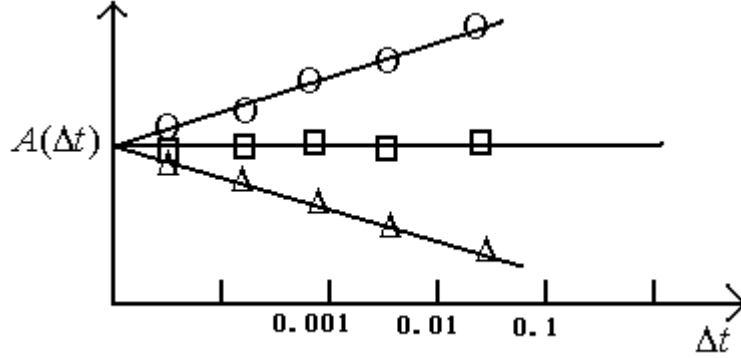
§6.6 随机积分方法

§6.7 数值策略

1. 数值结果对时间步长的依赖关系或寻找时间步长平台

$$A(\Delta t) = a + b_{\Delta t} \quad (6.109)$$

$$\Delta t \rightarrow 0 \quad \lim_{\Delta t \rightarrow 0} A(\Delta t) = a \quad (6.110)$$



2. 二阶龙格-库塔方法或预估-修正方法对非线性势场

3. 刚度问题 (Stiff problem) 用来隐式算法

$$\begin{cases} \dot{x} = v \\ m\dot{v} = -\nu v(t) + f(x(t)) + \sqrt{2\nu k_B T} \eta(t) \end{cases} \quad (6.111)$$

$$\langle \eta(t) \rangle = 0, \quad \langle \eta(t)\eta(s) \rangle = \delta(t-s) \quad (6.112)$$

即

$$\begin{cases} \dot{x} = v \\ \dot{v} = -\frac{\nu}{m} v + \frac{1}{m} \{f + \sqrt{2\nu T} \eta(t)\} \end{cases} \quad (6.113)$$

$m \rightarrow 0$ 或 $\nu \rightarrow \infty$ 系数 $\frac{\nu}{m} \rightarrow \infty$ 称为 Stiff prob.

$$\begin{cases} x(t+\Delta t) = x(t) + \frac{1}{2}\Delta t[v(t) + v^*(t)] \\ v(t+\Delta t) = (1 - \frac{\nu}{m}\Delta t)v(t) + \frac{\Delta t}{2m}[f(x) + f(x^*)] + \frac{\sqrt{2\nu T\Delta t}}{m}\omega_0 \end{cases} \quad (6.114)$$

Δt 必须很小 $0 < 1 - \frac{\nu}{m}\Delta t < 1$

$$\begin{aligned} v(t+\Delta t) - v(t) &= -\frac{\nu}{m} \int_t^{t+\Delta t} v(s)ds + \frac{1}{m} \int_t^{t+\Delta t} [f(x(s)) + \sqrt{2\nu T} \eta(s)]ds \\ &= -\frac{\nu}{m} v(t+\Delta t)\Delta t + \frac{1}{m} \int_t^{t+\Delta t} [f(x(s)) + \sqrt{2\nu T} \eta(s)]ds \end{aligned} \quad (6.115)$$

$$(1 + \frac{\nu}{m}\Delta t)v(t+\Delta t) = v(t) + \frac{1}{m} \int_t^{t+\Delta t} [f(x(s)) + \sqrt{2\nu T} \eta(s)]ds \quad (6.116)$$

所以

$$\begin{cases} v(t +_{\Delta} t) = \frac{m}{m + \nu_{\Delta} t} v(t) + \frac{1}{m + \nu_{\Delta} t} \int_t^{t +_{\Delta} t} [f(x(s)) + \sqrt{2\nu T} \eta(s)] ds \\ \quad = \frac{m}{m + \nu_{\Delta} t} v(t) + \frac{1}{m + \nu_{\Delta} t} \left\{ \frac{1}{2} [f + f^*]_{\Delta} t + \sqrt{2\nu T_{\Delta} t} \omega_0 \right\} \\ x(t +_{\Delta} t) = x(t) + v(t)_{\Delta} t \end{cases} \quad (6.117)$$

$m \rightarrow 0$ 方程变成 $v\dot{x} = f(x) + \sqrt{2\nu T}\eta(t)$

$$v(t +_{\Delta} t) = \frac{1}{v_{\Delta} t} \left\{ \frac{1}{2} [f + f^*]_{\Delta} t + \sqrt{2\nu T_{\Delta} t} \omega_0 \right\} \quad (6.118)$$

代人 x

$$x(t +_{\Delta} t) = x(t) + \frac{1}{2\nu} [f + f^*]_{\Delta} t + \frac{1}{\nu} \sqrt{2\nu T_{\Delta} t} \omega_0 \quad (6.119)$$

3. 混合算法：线性部分用积分，非线性部分用 R-KII

To see *J.D.Bao, J.Stat.phys.(USA)* 99, 595(2000) *phys.Rev.E* 60, 7572(1999)

J.D.Bao, S.J.Lin

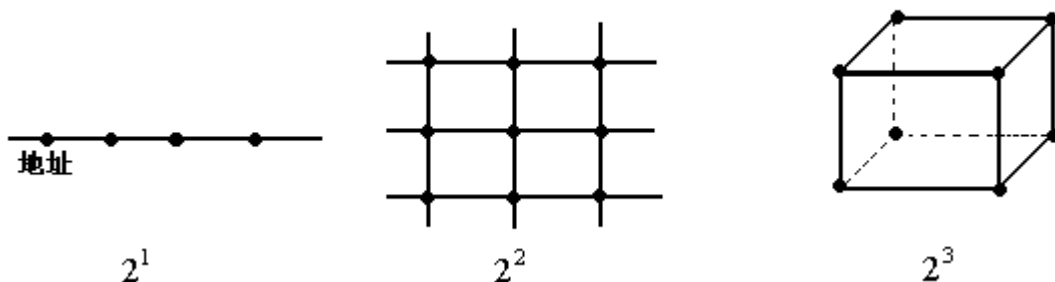
第七章 Ising 模型的蒙特卡罗模拟

Metropolis 抽样的成功应用的例子

§7.1 Ising Model

一. Ising Model

- 一个由 N 个固定点（地址）构成的 S 维（ $S=1,2,3$ ）规则磁性点阵



- 与每个地址 n 联系的是自旋变数 $\sigma_n, \sigma_n = \begin{cases} + & \uparrow \\ - & \downarrow \end{cases}$
- 因此系统的一个状态，由一个序列 $\vec{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_n, \dots, \sigma_N)$ 来描述
- 系统的哈密顿：

$$H_N(\vec{\sigma}) = - \sum_{(n,n')} J_{nn'} \sigma_{nn'} - \sum_{n=1}^N H_n \sigma_n \quad (7.1)$$

其中， $J_{nn'}$ 为耦合常数，表示 σ_n 和 $\sigma_{n'}$ 的相互作用， H_n 为外加磁场，其方向为自旋向上的方向，强度为 H_n 通常假定

$$\textcircled{1} \quad J_{nn'} = \begin{cases} J, & (n,n') \text{ 为最近邻对} \\ 0, & \text{else} \end{cases} \quad (7.2)$$

$$\textcircled{2} \quad H_n = H \text{ (均匀磁场)} \quad (7.3)$$

评论：

- (1) Ising 用来研究铁磁体性质；
 - (2) 温度由高降低时，二维和三维 Ising 模型均显现出相变
 - (3) 一维、二维 Ising 模型可以解析，三维只能用 MC 确定
- 这时，正则系综的平均量 $\langle A(\vec{\sigma}) \rangle$ 为：

$$\langle A(\vec{\sigma}) \rangle = \frac{\sum_{\{\sigma\}} e^{-H_N(\vec{\sigma})/k_B T} A(\vec{\sigma})}{\sum_{\{\sigma\}} e^{-H_N(\vec{\sigma})/k_B T}} \quad (7.4)$$

二. 主要物理量和方法

- 在 Ising 模型中, 除了系统能量 $\langle H_N(\vec{\sigma}) \rangle$ 外, 还有平均磁化强度:

$$\langle M \rangle = \langle \frac{1}{N} \sum_{n=1}^N \sigma_n \rangle \quad (7.5)$$

平均磁化率 χ :

$$\chi = \frac{\partial \langle M \rangle}{\partial H} = \frac{N}{k_B T} [\langle M^2 \rangle - \langle M \rangle^2] \quad (7.6)$$

- 解决方法:

① 用主方程描写 Ising 模型系统随时间演化的动力学过程

② 对 Ising 系统, 用 Metropolis 方法抽样产生的每一状态, 看作真实动力学过程演变出的状态

三. Metropolis 方法

① 用 Metropolis 抽样方法得到状态序列:

$$\vec{\sigma}_0, \vec{\sigma}_1, \dots, \vec{\sigma}_m, \dots \quad \begin{cases} \vec{\sigma}_0 = (\sigma_{10}, \sigma_{20}, \dots, \sigma_{n0}, \dots, \sigma_{N0}) & t = 0 \\ \dots & \dots \\ \vec{\sigma}_m = (\sigma_{1m}, \sigma_{2m}, \dots, \sigma_{nm}, \dots, \sigma_{Nm}) & t = m\tau \end{cases} \quad (7.7)$$

每一个状态 $\vec{\sigma}_m$, 又看作是时间 $t = m\tau$ 时系统的状态。则除了计算 $\langle M \rangle$ 和 χ 这一类正则系综的平均量以外, 还可计算在动力学过程中, 任意时刻 $t = m\tau, m = 1, 2, \dots$ 的磁化强度

$$\langle \mu(\tau) \rangle = \langle \frac{1}{N} \sum_{n=1}^N \sigma_n(m\tau) \rangle \quad (7.8)$$

这里, τ 表示系统换转一次状态所要的平均时间 (寿命)

② 在时间 $t = m\tau$ 对系统状态平均, 必须产生固定 $t = m\tau$ 下的各种状态, 可用多轮 Metropolis 抽样方法

$$\vec{\sigma}_0(k), \vec{\sigma}_1(k), \dots, \vec{\sigma}_m(k), \vec{\sigma}_{m+1}(k), \dots \quad k = 1, 2, \dots, K$$

m 为时间离散, K 为系统轨道, $\vec{\sigma}_m(k) = (\sigma_{m1}(k), \sigma_{m2}(k), \dots, \sigma_{mn}(k), \dots, \sigma_{mN}(k))$

☆ $\sigma_{mn}(k)$ 表示第 k 轮 Metropolis 抽样时, 第 n 个地址上的自旋, 在时刻 $t = m\tau$ 的值 $\vec{\sigma}_0(k), k = 1, 2, \dots, K$ 遵守同一初始分布

- 磁化强度

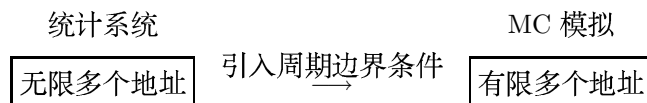
$$\langle \mu(t) \rangle = \mu_{mK} = \frac{1}{K} \sum_{k=1}^K (\frac{1}{N} \sum_{n=1}^N \sigma_{mn}(k)) \quad (7.9)$$

- 进一步求得磁化强度的时间平均值

$$\langle \mu \rangle = \frac{1}{t_{m'} - t_{n'}} \int_{t_{n'}}^{t_{m'}} \langle \mu(s) \rangle ds \quad (7.10)$$

这里, $(t_{n'}, t_{m'})$ 是系统处于亚稳态的时间范围。

四. 周期边界条件



三维点阵的地址排列方式有三种：①简立方（SC）；②面心立方（FCC）；③体心立方（BCC）
讨论简立方点阵的周期边界条件，设点阵地址数为 N ， $N = L \times L \times L$ 于是，简立方点阵中任一地址 n ，可有三个坐标 i, j, k

$$n = (i, j, k) \quad 1 \leq i, j, k \leq L \quad (7.11)$$

周期边界条件可用下式表示：

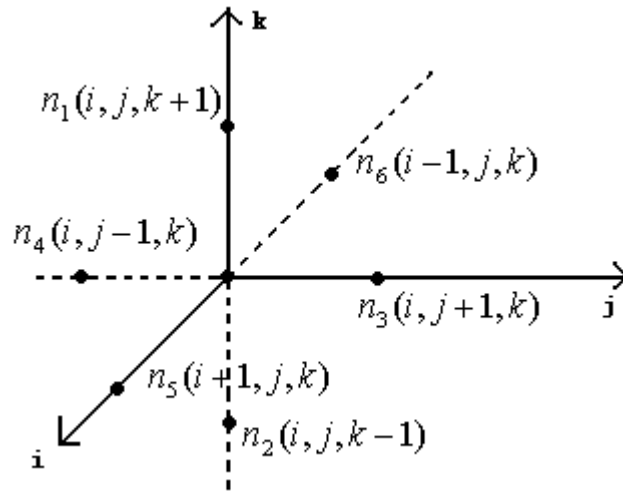
$$A(i, j, k) = A(i + m_i L, j + m_j L, k + m_k L), \quad m, m_j, m_k = 0, \pm 1, \pm 2, \dots \quad (7.12)$$

这里， $A(i, j, k)$ 为地址 $n = (i, j, k)$ 上的任意一个观察物理量。特别地（如自旋），

$$\sigma_n = \sigma(i, j, k) = \sigma(i + m_i L, j + m_j L, k + m_k L) \quad (7.13)$$

- L 越大，越接近真实（无限）情况。

五. 最近邻相互作用



- 最近邻定义：两个地址间的距离是一个间距单位。

与地址 $n = (i, j, k)$ 最近邻的地址为六个地址：

$$\begin{aligned} n_1 : & (i, j, k+1), & n_2 : & (i, j, k-1) \\ n_3 : & (i, j+1, k), & n_4 : & (i, j-1, k) \\ n_5 : & (i+1, j, k), & n_6 : & (i-1, j, k) \end{aligned} \quad (7.14)$$

在计算哈密顿 $H_N(\vec{\sigma})$ 时，要对所有最近邻相互作用求和，又要求不重复，因此

$$\begin{aligned} H_N(\vec{\sigma}) = & - \sum_{i=1}^L \sum_{j=1}^L \sum_{k=1}^L (J_{nn1} \sigma_n \sigma_{n1} + J_{nn3} \sigma_n \sigma_{n3} + J_{nn5} \sigma_n \sigma_{n5}) \\ & - \sum_{i=1}^L \sum_{j=1}^L \sum_{k=1}^L H_n \sigma_n \end{aligned} \quad (7.15)$$

§7.2 Ising 模型的 MC 模拟

从系综分布

$$\pi(\vec{\sigma}) = \frac{e^{-H_N(\vec{\sigma})/k_B T}}{\sum_{\{\vec{\sigma}\}} e^{-H_N(\vec{\sigma})/k_B T}} \quad (7.16)$$

产生子样或代表状态: $\vec{\sigma}_0, \vec{\sigma}_1, \dots, \vec{\sigma}_m, \vec{\sigma}_{m+1}, \dots$ 其中 $\sigma = (\sigma_1, \dots, \sigma_n, \dots, \sigma_N)$ 经过下列三步:

(1) 初始状态 $\vec{\sigma}_0$: 对每一地址 n 的自旋, 以某概率取 $\sigma_n = 1$ 或 $\sigma_n = -1$ ($n = 1, 2, \dots, N$), 例如简单地令 $\vec{\sigma}_0 = (1, 1, \dots, 1, \dots, 1)$ 。一般:

$$\vec{\sigma}_0 = (\sigma_{01}, \sigma_{02}, \dots, \sigma_{0n}, \dots, \sigma_{0N}) \quad (7.17)$$

(2) 对任意 $m \geq 0$, 知道 $\vec{\sigma}_m$ 后, $\vec{\sigma}_{m+1}$ 的产生如下 (Pass 方法): 令

$$\vec{\sigma}_{m+1}^{(0)} = \vec{\sigma}_m = (\sigma_{m+1\ 1}^{(0)}, \dots, \sigma_{m+1\ n}^{(0)}, \dots, \sigma_{m+1\ N}^{(0)}) = (\sigma_{m1}, \dots, \sigma_{mn}, \dots, \sigma_{mN}) \quad (7.18)$$

先翻转第一个地址上的自旋, 并保持其它地址上的自旋不变, 组成一个预选状态 $\vec{\sigma}_{m+1\ 1}^*$:

$$\vec{\sigma}_{m+1\ 1}^* = (-\sigma_{m+1\ 1}^{(0)}, \sigma_{m+1\ 2}^{(0)}, \dots, \sigma_{m+1\ n}^{(0)}, \dots, \sigma_{m+1\ N}^{(0)}) \quad (7.19)$$

利用 Metropolis 规则判断新状态 $\vec{\sigma}_{m+1}^{(1)}$, 即

① 计算

$$\Delta H_N(\vec{\sigma}_{m+1\ 1}^*, \vec{\sigma}_{m+1}^{(0)}) = H_N(\vec{\sigma}_{m+1\ 1}^*) - H_N(\vec{\sigma}_{m+1}^{(0)}) \quad (7.20)$$

如果 $\Delta H_N \leq 0$, 那么 $\vec{\sigma}_{m+1}^{(1)} = \vec{\sigma}_{m+1\ 1}^*$

② 否则, 利用随机数 ξ 进一步判断, 如果

$$\xi \leq e^{-\Delta H_N/k_B T} \quad (7.21)$$

成立, 同样取 $\vec{\sigma}_{m+1}^{(1)} = \vec{\sigma}_{m+1\ 1}^*$, 否则 $\vec{\sigma}_{m+1}^{(1)} = \vec{\sigma}_{m+1}^{(0)}$

③ 一般, 当 $0 \leq n < N$, 知道 $\vec{\sigma}_{m+1}^{(n)}$ 后, 可进一步得到 $\vec{\sigma}_{m+1}^{(n+1)}$, 设

$$\vec{\sigma}_{m+1}^{(n)} = (\sigma_{m+1\ 1}^{(n)}, \dots, \sigma_{m+1\ n}^{(n)}, \sigma_{m+1\ n+1}^{(n)}, \dots, \sigma_{m+1\ N}^{(n)}) \quad (7.22)$$

翻转第 $n+1$ 个地址上的自旋 $S_{n+1} = \pm 1 \rightarrow \mp 1$, 同时保持其它各地址上的自旋不变, 组成一个预选状态 $\vec{\sigma}_{m+1\ n+1}^*$

$$\vec{\sigma}_{m+1\ n+1}^* = (\sigma_{m+1\ 1}^{(n)}, \dots, \sigma_{m+1\ n}^{(n)}, -\sigma_{m+1\ n+1}^{(n)}, \dots, \sigma_{m+1\ N}^{(n)}) \quad (7.23)$$

重复①和②

$$\vec{\sigma}_{m+1}^{(n+1)} = \begin{cases} \vec{\sigma}_{m+1\ n+1}^* \\ \vec{\sigma}_{m+1}^{(n)} \end{cases} \quad (7.24)$$

④ 持续翻转到第 N 个地址上的自旋, 得到状态 $\vec{\sigma}_{m+1}^{(N)}$ 为止: $\vec{\sigma}_{m+1}^{(1)}, \vec{\sigma}_{m+1}^{(2)}, \dots, \vec{\sigma}_{m+1}^{(n)}, \dots, \vec{\sigma}_{m+1}^{(N)}$ 问题: 如何计算 $H_N(\vec{\sigma})$?

§7.3 二元合金系统

- 目的：用 Ising 模型描述二元合金中的有序无序现象
- 什么是二元合金系统

设在合金点阵的各个地址上，可由两种原子 A 和 B 中的一个去占据，而形式 A 和 B 原子保持固定的组成比例。

- 系统的哈密顿量 H 为

$$H = \sum_{(i,j)} [V_{AA}(r_{ij})C_i^A C_j^A + V_{BB}(r_{ij})C_i^B C_j^B + 2V_{AB}(r_{ij})C_i^A C_j^B] - \sum_j g_A(\vec{r}_j)C_j^A - \sum_j g_B(\vec{r}_j)C_j^B + H_0 \quad (7.25)$$

这里，

$$C_i^A = \begin{cases} 1, & \text{地址 } i \text{ 由 A 原子占领} \\ 0, & \text{else} \end{cases} \quad (7.26)$$

$$C_j^B = \begin{cases} 1, & \text{地址 } j \text{ 由 B 原子占领} \\ 0, & \text{else} \end{cases} \quad (7.27)$$

$V_{AA}(r_{ij})$ 是两个 A 原子相距 $r_{ij} = |\vec{r}_i - \vec{r}_j|$ 时相互作用势

$V_{BB}(r_{ij})$ 是两个 B 原子相距 $r_{ij} = |\vec{r}_i - \vec{r}_j|$ 时相互作用势

$V_{AB}(r_{ij})$ 是 A 原子和 B 原子相距 r_{ij} 时的相互作用势。

$g_A(\vec{r}_i)$ 和 $g_B(\vec{r}_j)$ 分别是 A 原子和 B 原子在位置 \vec{r}_j 时的化学势， H_0 是其他自由度的贡献。

- 将 H 写成 Ising model 的 H 的形式，令

$$C_j^A = \frac{1 + \sigma_j}{2}, \quad C_j^B = \frac{1 - \sigma_j}{2} \quad \sigma_j = \pm 1 \quad (7.28)$$

$$\sigma_j = 1 \Rightarrow \begin{cases} C_j^A = 1 \\ C_j^B = 0 \end{cases} \text{ 表示地址 } j \text{ 由 A 占领} \quad (7.29)$$

$$\sigma_j = -1 \Rightarrow \begin{cases} C_j^A = 0 \\ C_j^B = 1 \end{cases} \text{ 表示地址 } j \text{ 由 B 占领} \quad (7.30)$$

那么，

$$H = - \sum_{(i,j)} J_{ij} \sigma_i \sigma_j - \sum_j H_j \sigma_j + H'_0 \quad (7.31)$$

其中，

$$J_{ij} = \frac{1}{4} [2V_{AB}(r_{ij}) - V_{AA}(r_{ij}) - V_{BB}(r_{ij})] \quad (7.32)$$

$$H_j = \frac{1}{2} \left\{ \sum_{i \neq j} [V_{BB}(r_{ij}) - V_{AA}(r_{ij})] + \frac{1}{2} [g_A(\vec{r}_j) - g_B(\vec{r}_j)] \right\} \quad (7.33)$$

$$H'_0 = H_0 + \frac{1}{4} \sum_{(i,j)} [V_{AA}(r_{ij}) + V_{BB}(r_{ij}) + 2V_{AB}(r_{ij})] - \frac{1}{2} \sum_j [g_A(\vec{r}_j) + g_B(\vec{r}_j)] \quad (7.34)$$

§7.4 XY 模型

一. XY model

Ising 模型：每个地址上只取某一方向上的两个值，一维可以推广到每个地址上的自旋在高维空间连续取值。

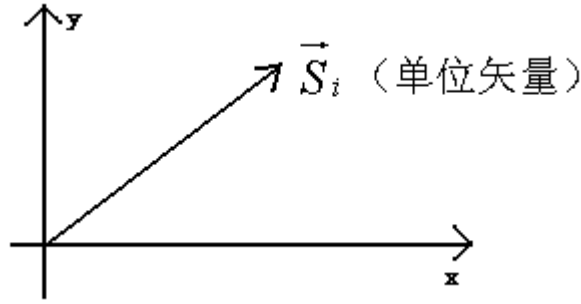
Heisenberg 模型：地址上的自旋在三维空间连续取值，这时两个自旋 \vec{S}_i 和 \vec{S}_j 的相互作用为：

$$-J_x^{ij} S_x^i S_x^j - J_y^{ij} S_y^i S_y^j - J_z^{ij} S_z^i S_z^j, \quad \vec{S} = (S_x, S_y, S_z) \text{ 单位向量} \quad (7.35)$$

如果 $J_x^{ij} = J_y^{ij} = J_z^{ij} = J^{ij}$ ，称之为各向同性的海森堡模型

$$H = - \sum_{(i,j)} J_{ij} \vec{S}_i \cdot \vec{S}_j - \sum_j H_j S_z^j \quad (7.36)$$

如果 $J_x = J_y, J_z = 0 \Rightarrow XY \text{ model}$



$$H_{xy} = - \sum_{(ij)} J_{ij} (S_x^i S_x^j + S_y^i S_y^j) \quad (7.37)$$

$$\begin{cases} S_x = \cos \varphi \\ S_y = \sin \varphi \end{cases} \quad 0 \leq \varphi \leq 2\pi \quad (7.38)$$

则

$$H_{xy} = - \sum_{(ij)} J_{ij} (\cos \varphi_i \cos \varphi_j + \sin \varphi_i \sin \varphi_j) = - \sum_{(ij)} J_{ij} \cos(\varphi_i - \varphi_j) \quad (7.39)$$

二. MC 计算 XY 模型

计算系统的内能 E 和比热 C_V 随温度 T 的变化

$$E = \frac{\int H_{XY}(\vec{\varphi}) e^{-H_{XY}(\vec{\varphi})/k_B T} d\vec{\varphi}}{\int e^{-H_{XY}(\vec{\varphi})/k_B T} d\vec{\varphi}} \quad (7.40)$$

$$C_V = \frac{\partial E}{\partial T} = (\langle H_{XY}^2 \rangle - \langle H_{XY} \rangle^2) / k_B T^2 \quad (7.41)$$

$$\vec{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_N), \quad 0 \leq \varphi_n \leq 2\pi, \quad n = 1, 2, \dots, N \quad (7.42)$$

N 为点阵中地址的总数。

Ising model \rightarrow 顺序改变每个地址上的自旋, XY model \rightarrow 改变每个地址上的 φ_i 角, 通常预选

$$\varphi_j^* = \varphi_i + (2\xi - 1)\Delta\varphi \quad \Delta\varphi = \text{常数} \quad (7.43)$$

在计算时，可先将 φ 离散化来做，即自旋取单位圆上若干个方向，

$$\vec{\varphi}_k = \frac{2\pi}{K} k, \quad k = 1, 2, \dots, K \quad (7.44)$$

$$\vec{\varphi} = (\vec{\varphi}_{k1}, \vec{\varphi}_{k2}, \dots, \vec{\varphi}_{kN}) \quad (7.45)$$

$$E = \frac{\sum_{k_1=1}^K \dots \sum_{k_N=1}^K H_{XY}(\vec{\varphi}) e^{-H_{XY}(\vec{\varphi})/k_B T}}{\sum_{k_1=1}^K \dots \sum_{k_N=1}^K e^{-H_{XY}(\vec{\varphi})/k_B T}} \quad (7.46)$$

$$H_{XY}(\vec{\varphi}) = -J \sum_{(ij)} \cos(\vec{\varphi}_{ki} - \vec{\varphi}_{kj}) \quad (7.47)$$

$\cos(\vec{\varphi}_{ki} - \vec{\varphi}_{kj})$ 可能取值只有 K 个，相应的 $e^{-J \cos(\vec{\varphi}_{ki} - \vec{\varphi}_{kj})/k_B T}$ 也只有 K 个。在计算时，可以事先将它们算好存放起来。

第八章 量子蒙特卡罗方法初步

§8.1 量子多体系统的最低能量

一. 量子多体系统

- N 粒子量子系统满足薛定谔方程

$$H\Psi(R) = E\Psi(R) \quad (8.1)$$

$$H = -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla_i^2 + V(R) \quad (8.2)$$

$$R = (\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \in D, \quad \vec{r}_i \in \bar{V}_0, \quad i = 1, 2, \dots, N \quad (8.3)$$

其中 E 为系统能量特征值, V(R) 为 N 个粒子间的相互作用势, D 为 3N 维区域, \bar{V}_0 为粒子 \vec{r}_i 的取值区域, $|\Psi(R)|^2 dR$ 为 N 个粒子占有状态在 R 附近 dR 的几率, $\Psi(R)$ 满足一定的边界条件。

二. 变分原理

考虑泛函

$$E_t = E[\Psi_t] = \frac{\int_D \Psi_t^* H \Psi_t dR}{\int_D \Psi_t^* \Psi_t dR} \quad (*) \quad (8.4)$$

$$\int_D |\Psi_t(R)|^2 dR < +\infty \quad (8.5)$$

- 思想: 当波函数 $\Psi_t(R)$ 满足薛定谔方程时, E_t 有极值; 反过来, 当 E_t 在 $\Psi_t(R)$ 取到极值时, E 和 $\Psi_t(R)$ 满足薛定谔方程。

推导变分原理:

设 $\Psi = \Psi_t(R)$ 时, E_t 有极值

$$E = \frac{\int_D \Psi^* H \Psi dR}{\int_D \Psi^* \Psi dR} \quad (8.6)$$

那么, 对于任何与 Ψ 相差 $\delta\Psi$ 的函数 $\Psi_t = \Psi + \delta\Psi$ 总有

$$\delta E_t = E[\Psi + \delta\Psi] - E[\Psi] = 0 \quad (8.7)$$

略去二阶小量, 并注意到算符 H 的厄米性, E 为实数, 有

$$\begin{aligned} \delta E &= \frac{\int_D (\Psi^* + \delta\Psi^*) H (\Psi + \delta\Psi) dR}{\int_D (\Psi^* + \delta\Psi^*) (\Psi + \delta\Psi) dR} - \frac{\int_D \Psi^* H \Psi dR}{\int_D \Psi^* \Psi dR} \\ &= \frac{\int_D \Psi^* (H - E) \delta\Psi dR + \int_D \delta\Psi^* (H - E) \Psi dR}{\int_D \Psi^* \Psi dR} \end{aligned} \quad (8.8)$$

由于 $\int_D \Psi^* (H - E) \delta\Psi dR = \int_D \delta\Psi^* (H - E) \Psi dR$ 及 $\delta\Psi$ 的任意性, 立即由 $\delta E_t = 0$ 得到 $(H - E)\Psi = 0$, 即 $H\Psi = E\Psi$ ☆求 S.E 的基态能量转化为求泛函 (*) 的最小值问题。

§8.2 McMillan-Metropolis 方法

一. M-M 方法

为求泛函 (*) 的极小值, 先取某一试探函数 $\Psi_t(R, a)$, a 为参数向量, 从其中找出泛函的极小值, 计算若干 $a_j, j = 1, 2, \dots, J$

设 a^* 满足

$$E[\Psi_t(R, a^*)] = \min_{1 \leq j \leq J} E[\Psi_t(R, a_j)] \quad (8.9)$$

那么 $E[\Psi_t(R, a^*)]$ 和 $\Psi_t(R, a^*)$ 就是 S.E 基态能量问题的近似解。

$$E_t = \frac{\int_D \Psi_t^* H \Psi_t dR}{\int_D \Psi_t^* \Psi_t dR} = \frac{\int_D |\Psi_t|^2 \frac{1}{\Psi_t} H \Psi_t dR}{\int_D |\Psi_t|^2 dR} = \langle \frac{1}{\Psi_t} H \Psi_t \rangle \quad (8.10)$$

分布密度函数

$$\pi(R) = \frac{|\Psi_t(R)|^2}{\int_D |\Psi_t(R)|^2 dR} \quad (8.11)$$

用 Metropolis 抽样方法。

二. 试探函数族 $\Psi_t(R, a)$

多取乘积型,

① 比如对玻色液体取为

$$\Psi_t(R) = \prod_{i < j=1}^N f(r_{ij}) = \prod_{i < j=1}^N e^{-\frac{1}{2}u(r_{ij})} = e^{-\frac{1}{2}\sum_{i < j=1}^N u(r_{ij})} \quad (8.12)$$

$$r_{ij} = |\vec{r}_i - \vec{r}_j| \quad (8.13)$$

② N 个粒子的经典系统, R 的分布为 Boltzmann 分布

$$f(R) = \frac{e^{-E/k_B T}}{\int_D e^{-E/k_B T} dR}, \quad E = \sum_{i < j=1}^N \Phi(r_{ij}) \quad (\text{势能}) \quad (8.14)$$

③ $|\Psi_t(R)|^2$ 与 $f(R)$ 相类似,

$$|\Psi_t(R)|^2 = e^{-\sum_{i < j=1}^N u(r_{ij})} \quad (8.15)$$

$u(r_{ij})$ 与 $\Phi(r_{ij})$ 相对应, 称为伪势。

④ 例子: McMillan 计算 ^4He 液体基态能量时, 相互作用势

$$V(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad \text{Lennard-Jones 势} \quad (8.16)$$

$\varepsilon = 10.22, \sigma = 2.556$, 而伪势 $u(r)$ 取为

$$u(r) = \left(\frac{b\sigma}{r} \right)^m \quad (8.17)$$

m, b 是两个参数, 对多组 $(m_j, b_j) \rightarrow$ 计算 $E_t \rightarrow$ 得到使 E_t 取最小值的 $(m_s, b_s) \rightarrow E_t$ 和 $\Psi_t(R)$ 。
 $m = 5, \quad b = 1.17$ 。

$3N, N=32, N=108$, 周期边界条件。

§8.3 偏倚抽样法求极小能量

为求 E_{tmin} , 要计算若干组参数 a_j 下的能量

$$E_{tj} = E_t[\Psi_t(R, a_j)], j = 1, 2, \dots, J \quad (8.18)$$

偏倚抽样法基本思想: 对某一个固定的参数 (比如 a_1), 从相应的分布函数抽出代表点, 在得到 E_{t1} 的估计值的同时, 得到其他 $E_{tj} (j \neq 1)$ 的估计值。

$$\begin{aligned} E_t[\Psi_t(R, a_j)] &= \frac{\int_D \Psi_t^*(R, a_j) H \Psi_t(R, a_j) dR}{\int_D \Psi_t^*(R, a_j) \Psi_t(R, a_j) dR} \\ &= \frac{\int_D \left| \frac{\Psi_t(R, a_j)}{\Psi_t(R, a_1)} \right|^2 \frac{H \Psi_t(R, a_j)}{\Psi_t(R, a_j)} |\Psi_t(R, a_1)|^2 dR}{\int_D \left| \frac{\Psi_t(R, a_j)}{\Psi_t(R, a_1)} \right|^2 |\Psi_t(R, a_1)|^2 dR} \end{aligned} \quad (8.19)$$

因此从分布 $\frac{|\Psi_t(R, a_1)|^2}{\int_D |\Psi_t(R, a_1)|^2 dR}$ 抽出代表点 R_1, R_2, \dots, R_M , 那么在得到 E_{t1} 同时, 也可得到诸 E_{tj} 的估计如下:

$$E_{tj} = \frac{\sum_{m=1}^M W_{mj} H \Psi_t(R_m, a_j) / \Psi_t(R_m, a_j)}{\sum_{m=1}^M W_{mj}}, \quad j = 1, 2, \dots, J \quad (8.20)$$

其中

$$W_{mj} = \left| \frac{\Psi_t(R_m, a_j)}{\Psi_t(R_m, a_1)} \right|^2 \quad (8.21)$$