



计算物理学

作者：张仁友

组织：中国科学技术大学

前言

与同学们相伴学习已然一学期，课堂教学笔记皆为黑板书。为方便同学复习参考，在往年电子版讲义基础上二次修订，成此。诸多课堂内容无法一一详尽于此讲义之中，希望后续能不断完善。讲义内容如有瑕疵或错误，欢迎指正。师生共同进步。

Ren-You Zhang

January 1, 2021

目录

1	随机变量及其分布	1
1.1	概率论简介	1
1.2	一维随机变量	5
2	蒙特卡洛方法	11
2.1	大数定律和中心极限定理	11
2.2	减小蒙特卡洛方差的方法	13
2.3	一维随机变量抽样方法	17
2.4	高维随机向量抽样方法	24
3	蒙特卡洛方法应用	26
3.1	多体末态相空间产生	26
4	微分方程数值求解方法	32
4.1	常微分方程	32
4.2	偏微分方程有限差分法	36
4.3	偏微分方程有限元素法	41

第一章 随机变量及其分布

1.1 概率论简介

所谓试验，应当在相同条件下可以重复进行。若试验的所有可能结果都已明确知晓，但每次试验结果无法确定，此类试验称为**随机试验**。随机试验的可能结果称为**随机事件**，简称**事件**。

随机试验的结果（即所有可能的事件）由**样本空间**描述。样本空间的任一子集即为随机试验的一个可能事件。记： S 为样本空间，

$$A = \{e\} \quad e \in S \quad (\text{基本事件})$$

$$A = S \quad (\text{必然事件})$$

$$A = \emptyset \quad (\text{不可能事件})$$

基于集合论，我们可以引入关于事件的如下概念：

定义 1.1

- (和事件) $A \cup B$: 事件 A 和事件 B 至少有一个发生。
- (积事件) $A \cap B$: 事件 A 和事件 B 同时发生。
- (差事件) $A - B$: 事件 A 发生而事件 B 不发生。
- (不相容) $A \cap B = \emptyset$
- (对立事件) $\bar{A} = S - A$

$A_i \neq \emptyset (i = 1, 2, \dots, n)$ 为随机试验的 n 个事件，满足

$$\begin{cases} A_i \cap A_j = \emptyset, & \forall i \neq j \\ A_1 \cup A_2 \cup \dots \cup A_n = S \end{cases}$$

称 $\{A_i | i = 1, 2, \dots, n\}$ 为样本空间 S 的一个划分。



1.1.1 概率及其性质

为定量描述随机试验中各随机事件发生的可能性的的大小，引入**概率**。

定义 1.2. (概率)

$P: A \subset S \mapsto P(A) \in \mathbb{R}$ ，满足

$$(1) \quad 0 \leq P(A) \leq 1, \quad \forall A \subset S$$

$$(2) \quad P(S) = 1$$

$$(3) \quad P(A_1 \cup A_2 \cup \dots \cup A_n \cup \dots) = \sum_k P(A_k), \quad \{A_k | k = 1, 2, \dots\} \text{两两不相容}$$



通常，先验概率基于某种假定而获得。如

等概率假定：样本空间为包含 n 个元素的有限集时，每个基本事件的概率为 $1/n$ 。

性质

- (1) $\sum_{i=1}^n P(A_i) = 1,$ $\{A_i | i = 1, 2, \dots, n\}$ 为分割
- (2) $P(A - B) = P(A) - P(B) \geq 0,$ $A \supset B$
- (3) $P(A \cup B) = P(A) + P(B) - P(AB)$
- $$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 A_2) - P(A_2 A_3) - P(A_3 A_1) + P(A_1 A_2 A_3)$$

1.1.2 条件概率

A, B 为一随机试验的两个事件 ($A \neq \emptyset$), 概率分别为 $P(A)$ 和 $P(B)$ 。问: 事件 A 发生的条件下事件 B 发生的概率是多少?

回答这个问题需要注意两件事情:

1. 事件 A 条件下的事件 B , 即为积事件 AB 。该积事件的概率为 $P(AB)$;
2. 以事件 A 为条件时, 样本空间为 A 。概率的归一性诱导出归一化因子 $1/P(A)$ 。

因此,

定理 1.1. (条件概率)

事件 A 条件下, 事件 B 的条件概率为

$$P(B|A) = \frac{P(AB)}{P(A)}$$



条件概率定理: 条件概率即为重新归一化后的积事件概率。反之, AB 积事件的概率亦可由 A 事件和 B 事件的概率相乘而得到。

定理 1.2. (积事件概率)

$$P(AB) = P(A|B)P(B) = P(B|A)P(A)$$

推论:

$$\begin{aligned} P(A_n A_{n-1} A_{n-2} \dots A_2 A_1) &= P(A_n | A_{n-1} A_{n-2} \dots A_2 A_1) \\ &\quad P(A_{n-1} | A_{n-2} \dots A_2 A_1) \\ &\quad \dots \\ &\quad P(A_2 | A_1) \\ &\quad P(A_1) \end{aligned}$$



什么是相互独立的事件? n 个事件 A_i ($i = 1, 2, \dots, n$) 之间彼此独立, 则

$$P(A_{j_1} A_{j_2} \dots A_{j_k}) = P(A_{j_1}) P(A_{j_2}) \dots P(A_{j_k})$$

$$\forall 0 \leq j_1 < j_2 < \dots < j_k \leq n$$

特别地, 对于两个相互独立的事件 A 和 B ,

$$P(AB) = P(A)P(B) \longrightarrow P(A|B) = P(A), \quad P(B|A) = P(B)$$

例 1.1 某种粒子探测器，信号粒子触发概率为 $\eta \sim 1$ ，偶然虚假触发概率为 $\varepsilon \sim 0$ 。问：如何提高粒子探测效率？

解 探测器随机试验为伯努利试验，基本事件为

$$\begin{cases} A = \{\text{触发}\} & P(A) = p \\ \bar{A} = \{\text{未触发}\} & P(\bar{A}) = q \end{cases} \quad (p + q = 1)$$

进行了多次彼此独立的相同伯努利试验。

(1) $\underline{P(A_1 \cup A_2 \cup \dots \cup A_n)}$:

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n) &= 1 - P(\overline{A_1 \cup A_2 \cup \dots \cup A_n}) \\ &= 1 - P(\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_n) \\ &= 1 - q^n \end{aligned}$$

(2) $\underline{P(A_1 \cap A_2 \cap \dots \cap A_n)}$:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = p^n$$

可以看到

$$P(A_1 \cup A_2 \cup \dots \cup A_n) > P(A) > P(A_1 \cap A_2 \cap \dots \cap A_n)$$

这一结论与触发概率 p 无关。显然，上面任何一种方案均无法实现增加 η 的同时减小 ε 。

考察如下事件：

$$D = \bigcap_{a=1}^m \left(\bigcup_{i=1}^n A_i^{(a)} \right) \longrightarrow P(D) = (1 - q^n)^m$$

• 情形一 ($p \sim 1, q \sim 0$):

$$P(D) = (1 - q^n)^m \simeq 1 - mq^n \quad P(D) > P(A)$$

• 情形二 ($p \sim 0, q \sim 1$):

$$P(D) = (1 - q^n)^m = [1 - (1 - p)^n]^m \simeq (np)^m \quad P(D) < P(A)$$


因此，定义事件 D 为试验成功，可以提高粒子探测效率。

1.1.3 贝叶斯公式

利用样本空间分割的性质，我们可以得到如下的全概率公式和贝叶斯公式。

定理 1.3. (全概率公式)

$$(1) \quad P(B) = \sum_{k=1}^n P(B|A_k)P(A_k) \quad (2) \quad P(B|C) = \sum_{k=1}^n P(B|A_k C)P(A_k|C)$$

$\{A_k | k = 1, 2, \dots, n\}$ 为样本空间的划分 

全概率公式 (2) 可以等价地写为

$$P(B|C) = \sum_{k=1}^n P(B|A_k C)P(A_k C|C)$$

与全概率公式 (1) 相比较,

样本空间: $S \rightarrow C$ 分割: $\{A_k | k = 1, 2, \dots, n\} \rightarrow \{A_k C | k = 1, 2, \dots, n\}$

需要说明的是, $A_k C$ 不能为 \emptyset , 否则从分割中去除之。

定理 1.4. (贝叶斯公式)

$$\begin{aligned} P(A_i|B) &= \frac{P(B|A_i)P(A_i)}{P(B)} \\ &= P(B|A_i)P(A_i) / \sum_{k=1}^n P(B|A_k)P(A_k) \end{aligned}$$

$\{A_k | k = 1, 2, \dots, n\}$ 为样本空间的划分

贝叶斯公式中, 有别于先验概率 $P(A_i)$, $P(A_i|B)$ 称为后验概率。

例 1.2 盒中装有 a 个黑球和 b 个白球。随机从盒中取出一个小球, 然后在盒中添加 c 个同种颜色的小球, 并将取出的小球放回。问: 这之后从盒中取出白球的概率是多少?

解

$A_1 = \{\text{第一次取出黑球}\}$ $A_2 = \{\text{第一次取出白球}\}$

$B = \{\text{第二次取出白球}\}$

等概率假设

$$\begin{cases} P(A_1) = \frac{a}{a+b} \\ P(A_2) = \frac{b}{a+b} \end{cases} \quad \begin{cases} P(B|A_1) = \frac{b}{a+b+c} \\ P(B|A_2) = \frac{b+c}{a+b+c} \end{cases}$$

全概率公式

$$\begin{aligned} P(B) &= \sum_{k=1}^2 P(B|A_k)P(A_k) \\ &= \frac{b}{a+b+c} \times \frac{a}{a+b} + \frac{b+c}{a+b+c} \times \frac{b}{a+b} = \frac{b}{a+b} \end{aligned}$$

例 1.3 有 5 个完全一样的盒子, 每个盒子分别装有 4 个小球, 黑球和白球数目分别为 $(4, 0), (3, 1), (2, 2), (1, 3), (0, 4)$ 。随机地从 5 个盒中取出 1 个小球, 之后继续从原来盒中取出一球。问: 在第一次取出黑球条件下, 第二次取出黑球的概率。

解

$A_i = \{\text{摸取第 } i \text{ 个盒子}\} \quad (i = 1, 2, 3, 4, 5)$

$B = \{\text{第一次取出黑球}\}$

$C = \{\text{第二次取出黑球}\}$

等概率假设

$$\begin{aligned} P(A_i) &= \frac{1}{5} \\ P(B|A_1) &= 1 \quad P(B|A_2) = \frac{3}{4} \quad P(B|A_3) = \frac{2}{4} \quad P(B|A_4) = \frac{1}{4} \quad P(B|A_5) = 0 \\ P(C|BA_1) &= 1 \quad P(C|BA_2) = \frac{2}{3} \quad P(C|BA_3) = \frac{1}{3} \quad P(C|BA_4) = 0 \end{aligned}$$

(1) $\underline{P(B)}$:

$$P(B) = \sum_{i=1}^5 P(B|A_i)P(A_i) = \frac{1}{2}$$

(2) $\underline{P(A_i|B)}$:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \begin{cases} 4/10 & i=1 \\ 3/10 & i=2 \\ 2/10 & i=3 \\ 1/10 & i=4 \\ 0 & i=5 \end{cases} \quad (BA_5 = \emptyset)$$

(3) $\underline{P(C|B)}$:

$$P(C|B) = \sum_{i=1}^4 P(C|BA_i)P(A_i|B) = \frac{2}{3}$$

1.2 一维随机变量

一维随机变量是定义在样本空间 S 上的单值实函数，分为离散型随机变量和连续型随机变量。

$$X : e \in S \mapsto X(e) = x \in \mathbb{R}$$

1.2.1 分布特征

定义 1.3

一维随机变量 X 的分布函数:

$$F(x) = P(X \leq x)$$

(如果存在) 分布密度函数:

$$f(x) = \frac{dF(x)}{dx} \quad F(x) = \int_{-\infty}^x f(\tau) d\tau$$



性质

$$(1) \quad 0 \leq F(x) \leq 1$$

$$(1) \quad f(x) \geq 0$$

$$(2) \quad F(+\infty) = 1$$

$$(2) \quad \int_{-\infty}^{+\infty} f(x) dx = 1$$

$$(3) \quad F(x_2) - F(x_1) \geq 0$$

$$(3) \quad \int_{x_1}^{x_2} f(x) dx = F(x_2) - F(x_1) \quad \forall x_2 > x_1$$

对于离散型随机变量，我们一般不采用分布函数描述，只需给出随机变量取各种不同值的概率:

$$P(X = x_i) = p_i \quad (i = 1, 2, \dots)$$

显然, 离散型随机变量的分布函数为右连续阶跃函数:

$$F(x) = \sum_{x_i \leq x} p_i$$

随机变量 Y 为随机变量 X 的函数, 则 Y 的分布由 X 的分布唯一决定。

- 离散型 $X = x \in \{x_i | i = 1, 2, \dots\}$:

$$y \in \{y(x_i) | i = 1, 2, \dots\} \quad P(Y = y) = \sum_{x=x(y)} P(X = x)$$

- 连续型 $X = x \in \mathbb{R}$:

$$g(y) = \sum_{x=x(y)} f(x) \left| \frac{dx}{dy} \right|$$

特别地, $y = y(x)$ 为一一映射时, 有

$$g(y) = f(x) \left| \frac{dx}{dy} \right| \quad \text{即:} \quad \begin{cases} g(y)dy = +f(x)dx & (\text{单调递增映射}) \\ g(y)dy = -f(x)dx & (\text{单调递减映射}) \end{cases}$$

例 1.4 X 为取值在 $[-1, +1]$ 区间上的均匀分布随机变量, Y 为 X 的函数:

$$Y = \tan \pi X$$

求 Y 的概率分布。

解 记 X 和 Y 的分布密度函数分别为 $f(x)$ 和 $g(y)$ 。由题意知: $f(x) = 1$ 。 $X \mapsto Y$ 为单调递增映射, 因此, 有

$$\begin{aligned} g(y)dy &= f(x)dx \\ &= d \left[\frac{1}{\pi} \arctan y \right] = \frac{1}{\pi} \frac{1}{1+y^2} dy \\ \therefore g(y) &= \frac{1}{\pi} \frac{1}{1+y^2} \end{aligned}$$

随机变量的分布特征可由分布函数、分布密度函数或随机变量的各阶矩描述。本质上, 矩是分布密度函数傅立叶像的泰勒展开系数。常用的随机变量数字特征包括: 期望值、方差、偏度、峰度, 分别为随机变量的一、二、三、四阶矩。

定义 1.4

期望值:	$\mu = \langle X \rangle$	偏度:	$\gamma_1 = \frac{\langle (X - \mu)^3 \rangle}{\sigma^3}$
方差:	$\sigma^2 = \langle (X - \mu)^2 \rangle$	峰度:	$\gamma_2 = \frac{\langle (X - \mu)^4 \rangle}{\sigma^4} - 3$



其中, $\langle \dots \rangle$ 定义为

$$\langle X \rangle = \begin{cases} \int x f(x) dx & (\text{连续型}) \\ \sum_i x_i p_i & (\text{离散型}) \end{cases}$$

σ^2 定量地反映出了随机变量相对其期望值的离散程度, γ_1 描述了分布密度函数相对于期望值的不对称程度, 而 γ_2 则给出了分布密度函数相对于正态分布 $N(\mu, \sigma^2)$ 的尖锐程度。

1.2.2 均匀分布

均匀分布是最简单的连续型随机变量分布。顾名思义, $[a, b]$ 区间上均匀分布随机变量在任意等长度子区间内取值的概率相同, 因此, 在 $[a, b]$ 区间内分布密度函数为常数。

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{else} \end{cases}$$

几个主要数字特征分别为

$$\mu = \frac{a+b}{2} \quad \sigma^2 = \frac{(b-a)^2}{12} \quad \gamma_1 = 0 \quad \gamma_2 = -1.2$$

均匀分布随机变量在蒙特卡洛方法中是非常基本而重要的。原则上, 任意分布的随机数序列均可通过对 $[0, 1]$ 区间上均匀分布的随机数序列的改造来得到。

1.2.3 指数衰减分布

以不稳定粒子的寿命 T 为例, 指数衰减分布密度函数为

$$f(t; \lambda) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases} \quad (\lambda > 0)$$

几个主要数字特征分别为

$$\mu = \frac{1}{\lambda} \quad \sigma^2 = \frac{1}{\lambda^2} \quad \gamma_1 = 2 \quad \gamma_2 = 6$$

衰变宽度 $\lambda = 1/t_0$, t_0 为该不稳定粒子的平均寿命。不稳定粒子寿命大于 t 的概率为

$$P(T > t) = \int_t^\infty f(\tau, \lambda) d\tau = 1 - F(t; \lambda) = e^{-\lambda t}$$

记 t 时刻不稳定粒子数目为 $N(t)$ 。若初始时刻粒子数目足够大, 则终了与初始时刻不稳定粒子数目之比为

$$\frac{N(t_2)}{N(t_1)} \simeq P(T > t_2 - t_1) = e^{-\lambda(t_2 - t_1)}$$

由此可得衰变宽度

$$\lambda = \frac{1}{t_2 - t_1} \text{Ln} \frac{N(t_1)}{N(t_2)}$$

命题 1.1

不稳定粒子数目足够大时

$$N(t) = N_0 e^{-\lambda t} \quad \frac{1}{N} \frac{dN}{dt} = -\lambda$$

$$\therefore \frac{1}{N} \frac{dN}{dt} = -\lambda = \text{const.} \quad \therefore \frac{N(t_1)}{N(t_2)} = \frac{dN(t_1)/dt}{dN(t_2)/dt}$$

衰变宽度又可写为

$$\lambda = \frac{1}{t_2 - t_1} \text{Ln} \left[\frac{dN(t_1)/dt}{dN(t_2)/dt} \right]$$

显然, dN/dt 比 N 更易于测量。

1.2.4 二项分布

只有两个基本事件的随机试验称为**伯努利试验**。方便起见，我们将两个基本事件 A 和 \bar{A} 分别记为“成功”和“失败”。 $P(A) = p$, $P(\bar{A}) = q \equiv 1 - p$ 。

定义 1.5

伯努利变量 R : 单次伯努利试验中试验成功的次数。

$$R = r \in \{0, 1\}$$

伯努利分布 $B(p)$: 成功概率为 p 的伯努利变量服从的分布。

$$B(r; p) = p^r q^{1-r} = \begin{cases} p & r = 1 \\ q & r = 0 \end{cases}$$



伯努利变量的期望值和方差分别为

$$\mu = p \quad \sigma^2 = pq$$

A 为随机试验 E 的一个事件，对立事件为 \bar{A} 。 $P(A) = p$, $P(\bar{A}) = q$ 。事件 A 诱导出伯努利试验 E_A ，其基本事件为 A 和 \bar{A} 。记事件 A 为“成功”，则 E_A 试验成功概率为 p 。

定理 1.5

R_i ($i = 1, 2, \dots, n$) 为彼此独立的同分布（试验成功概率为 p ）伯努利随机变量。

$$R = R_1 + R_2 + \dots + R_n$$

则 R 服从二项分布 $B(n, p)$:

$$B(r; n, p) = C_n^r p^r q^{n-r} \quad (r = 0, 1, 2, \dots, n)$$



显然， R 是独立进行 n 次相同伯努利试验中试验成功的次数。二项分布随机变量的期望值、方差、偏度、峰度分别为

$$\mu = np \quad \sigma^2 = npq \quad \gamma_1 = -\frac{p-q}{\sigma} \quad \gamma_2 = \frac{1-6pq}{\sigma^2}$$

推论 1.1

X 和 Y 相互独立， $X \sim B(n, p)$, $Y \sim B(m, p)$ 。则

$$X + Y \sim B(n + m, p)$$



注 (1) 伯努利分布为 $n = 1$ 的二项分布；(2) 二项分布为伯努利分布的和分布。

例 1.5 LEP (Large Electron Positron Collider) 上 $e^+e^- \rightarrow f\bar{f}$ 过程前后不对称性 (Forward-Backward Asymmetry, A_{FB}) 和左右不对称性 (Left-Right Asymmetry, A_{LR}) 的精确测量是研究弱相互作用宇称不守恒、确定弱混合角的重要实验方法。

$$A_{FB} = \frac{N_F - N_B}{N} \quad A_{LR} = \frac{N_L - N_R}{N}$$

$e^+e^- \rightarrow f\bar{f}$ 事例总数为 N ，前后向事例数和左右手事例数分别为 $N_{F,B}$ 和 $N_{L,R}$ ，方向及手征性由末态费米子 f 决定。试讨论 A_{FB} 和 A_{LR} 的期望值及方差。

解 以 A_{FB} 为例, 记

$$E: e^+e^- \rightarrow f\bar{f} \quad \begin{cases} F = \{f \text{ 前向}: \cos\theta > 0\} & P(F) = p \\ B = \{f \text{ 后向}: \cos\theta < 0\} & P(B) = q \end{cases}$$

事件 F 诱导伯努利试验 E_F 。显然, N_F 为 N 次独立伯努利试验 E_F 中试验成功的次数, $N_F \sim B(N, p)$ 。 $\langle N_F \rangle = Np$, $V(N_F) = Npq$ 。

命题 1.2

$X_i (i = 1, 2, \dots, n)$ 为彼此独立的随机变量, 期望值和方差分别为 μ_i 和 σ_i^2 。

$$Y = \sum_{i=1}^n \alpha_i X_i$$

则 Y 的期望值和方差分别为

$$\langle Y \rangle = \sum_{i=1}^n \alpha_i \mu_i \quad V(Y) = \sum_{i=1}^n \alpha_i^2 \sigma_i^2$$

由上述命题,

$$\therefore A_{FB} = \frac{2N_F}{N} - 1 \quad \therefore \begin{cases} \langle A_{FB} \rangle = \frac{2}{N} \langle N_F \rangle - 1 = p - q \\ V(A_{FB}) = \left(\frac{2}{N}\right)^2 V(N_F) = \frac{4pq}{N} \end{cases}$$

1.2.5 泊松分布

泊松分布是一种统计与概率学里常见到的离散机率分布, 适合于描述单位时间 (或空间) 内随机事件发生的次数。

定理 1.6

$R(t)$: $[0, t]$ 时间段内随机事件发生的次数。 $P(r; t)$: $R(t) = r (r \in \mathbb{N})$ 的概率。

假设 (Poisson)

1. $R(dt)$ 为伯努利变量。

$$P(r; dt) = \begin{cases} 1 - \lambda dt & r = 0 \\ \lambda dt & r = 1 \\ 0 & r \geq 2 \end{cases} \quad \lambda = \text{const.}$$

2. $\Delta t > 0$ 时, $R(t)$ 与 $R(t + \Delta t) - R(t)$ 相互独立。

则

$$P(r; t) = \frac{1}{r!} (\lambda t)^r e^{-\lambda t}$$

证明

$$R(t + dt) = R(t) + R(dt)$$

由独立性可知

$$\begin{aligned} P(r; t + dt) &= P(r; t) \otimes P(r; dt) \\ &= P(r; t)P(0; dt) + P(r-1; t)P(1; dt) \end{aligned}$$

(1) $r = 0$:

$$P(0; t + dt) = P(0; t)P(0; dt) \quad \text{说明: } P(-1; t) = 0$$

得到如下一阶微分方程

$$\begin{cases} \frac{dP(0; t)}{dt} = -\lambda P(0; t) \\ P(0; 0) = 1 \end{cases}$$

方程解为

$$P(0; t) = e^{-\lambda t}$$

注 $P(0; t)$ 为 $[0, t]$ 时间段内事件不发生的概率。由此可知, 事件首次发生的时刻 t 服从指数衰减分布: $f(t) = \lambda e^{-\lambda t}$

(2) $r \geq 1$:

$$P(r; t + dt) = P(r; t)P(0; dt) + P(r-1; t)P(1; dt)$$

由此得到如下递推微分方程

$$\frac{dP(r; t)}{dt} = -\lambda [P(r; t) - P(r-1; t)]$$

方程解为

$$P(r; t) = \frac{1}{r!} (\lambda t)^r e^{-\lambda t}$$

定理 1.7

满足泊松假设的随机变量 R 称为泊松变量, 服从泊松分布 $P(\mu)$:

$$P(r; \mu) = \frac{1}{r!} \mu^r e^{-\mu} \quad r \in \mathbb{N}$$



泊松变量的期望值、方差、偏度、峰度分别为

$$\mu = \mu \quad \sigma^2 = \mu \quad \gamma_1 = \frac{1}{\sqrt{\mu}} \quad \gamma_2 = \frac{1}{\mu}$$

泊松分布可以看作二项分布在保持期望值恒定条件下 $n \rightarrow +\infty$ 的极限。

推论 1.2

$$\lim_{n \rightarrow +\infty} B(n, \mu/n) = P(\mu)$$



证明

$$\begin{aligned} P(r; n, \mu/n) &= \frac{n!}{r!(n-r)!} \left(\frac{\mu}{n}\right)^r \left(1 - \frac{\mu}{n}\right)^{n-r} \\ &= \frac{1}{r!} \mu^r \frac{n!}{n^r(n-r)!} \left(1 - \frac{\mu}{n}\right)^{n-r} \end{aligned}$$

藉由如下极限

$$\lim_{n \rightarrow +\infty} \frac{n!}{n^r(n-r)!} = 1 \quad \lim_{n \rightarrow +\infty} \left(1 - \frac{\mu}{n}\right)^{n-r} = e^{-\mu}$$

得证结论。

第二章 蒙特卡洛方法

蒙特卡洛 (Monte Carlo) 方法是一种统计试验方法, 其基本思想是: 选择或建立合适的概率模型或概率过程, 将待求解问题关联至相关随机变量的期望值, 通过对概率模型进行随机模拟或随机抽样, 得到期望值的估计值, 从而得到所求问题的近似值。

2.1 大数定律和中心极限定理

蒙特卡洛方法的数学理论基础为大数定律和中心极限定理。大数定律反映的是相互独立随机变量序列的代数平均所具有的稳定性; 而中心极限定理则告诉我们大样本统计蒙特卡洛估计值的具体分布。

2.1.1 大数定律

首先, 我们介绍什么是随机变量序列依概率收敛和依分布收敛。

定义 2.1

$X_n (n = 1, 2, \dots)$ 为随机变量序列, X 为随机变量。

若

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} P(|X_n - X| < \varepsilon) = 1$$

称随机变量序列 $X_n (n = 1, 2, \dots)$ 依概率收敛于随机变量 X 。



定义 2.2

$X_n (n = 1, 2, \dots)$ 为随机变量序列, X 为随机变量, 分布函数分别为 $F_n(x)$ 和 $F(x)$ 。

若

$$\lim_{n \rightarrow +\infty} F_n(x) = F(x)$$

称随机变量序列 $X_n (n = 1, 2, \dots)$ 依分布收敛于随机变量 X 。



需要说明的是: (1) 依分布收敛是依概率收敛的必要条件;

(2) 当 X 为常数时, 依分布收敛与依概率收敛等价。

定理 2.1. 切比雪夫 (Chebyshev) 大数定律

$X_i (i = 1, 2, \dots)$ 为相互独立随机变量序列, 期望值为 μ_i , 方差存在且一致有上界。

则

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n \mu_i\right| < \varepsilon\right) = 1$$



注 切比雪夫大数定律可以由切比雪夫不等式方便证明。

$$\langle X \rangle = \mu, \quad V(X) = \sigma^2 \quad \longrightarrow \quad P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

定理 2.2. 辛钦 (Khinchin) 大数定律 (同分布大数定律)

$X_i (i = 1, 2, \dots)$ 为相互独立同分布随机变量序列, 期望值为 μ 。

则

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| < \varepsilon\right) = 1$$

即: $Y_n \equiv \frac{X_1 + X_2 + \dots + X_n}{n}$ 依概率收敛于 μ 。

(有别于切比雪夫大数定律, 辛钦大数定律中并不要求存在有限的方差。)



切比雪夫大数定律中, 若 $X_i (i = 1, 2, \dots)$ 为伯努利变量, 则

定理 2.3. 泊松 (Poisson) 大数定律

$X_i \sim B(p_i) (i = 1, 2, \dots)$ 为相互独立伯努利变量序列。

则

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \frac{p_1 + p_2 + \dots + p_n}{n}\right| < \varepsilon\right) = 1$$



泊松大数定律中, 若 $X_i (i = 1, 2, \dots)$ 服从相同的伯努利分布, 则

定理 2.4. 伯努利 (Bernoulli) 大数定律

$R_n \sim B(n, p)$ 。则

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow +\infty} P\left(|R_n/n - p| < \varepsilon\right) = 1$$

即: R_n/n 依概率收敛于 p 。

(大样本统计 $n \rightarrow +\infty$ 下, 试验成功的频率 R_n/n 趋于试验成功的概率。)



简而言之,

$$Y_n \equiv \frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow{P} \mu \equiv \langle X \rangle \quad (\text{同分布大数定律})$$

同分布大数定律揭示了大量相互独立同分布随机变量的算数平均结果, 但没有涉及该算数平均的分布问题; 这一问题的答案由中心极限定理给出。

2.1.2 中心极限定理**定理 2.5. 列维-林德伯格 (Levy-Lindeberg) 中心极限定理 (同分布中心极限定理)**

$X_i (i = 1, 2, \dots)$ 为相互独立同分布随机变量序列, 期望值和方差分别为 μ 和 σ^2 。

则

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow +\infty} P\left(\frac{(X_1 + X_2 + \dots + X_n) - n\mu}{\sqrt{n}\sigma} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$



从中心极限定理看到: 无论 $X_i (i = 1, 2, \dots)$ 服从什么样的分布, 大样本算数平均近似地服从正态分布。

$$n \rightarrow +\infty: \quad Y_n \equiv \frac{X_1 + X_2 + \dots + X_n}{n} \sim N(\mu, \sigma^2/n) \quad (\text{同分布中心极限定理})$$

若 $X_i (i = 1, 2, \dots)$ 为同分布伯努利变量, 则

定理 2.6. 棣莫佛-拉普拉斯 (de Moivre-Laplace) 中心极限定理

$R_n \sim B(n, p)$ 。则

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow +\infty} P\left(\frac{R_n - np}{\sqrt{np(1-p)}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int^x e^{-t^2/2} dt$$



实际问题中, 常遇到 $X_i (i = 1, 2, \dots)$ 相互独立但分布并不相同的情形, 例如: 实验测量误差的合成。此种情形下, 也存在类似的中心极限定理, 这里不再赘述。

基于大数定律和中心极限定理, 有如下 (关于蒙特卡洛方法的) 命题:

命题 2.1. 蒙特卡洛方法

$Y_n = \frac{X_1 + X_2 + \dots + X_n}{n}$, $X_i (i = 1, 2, \dots)$ 相互独立并与 X 服从相同分布。 $I = \langle X \rangle$ 。
则

1. 蒙特卡洛估计: $I_n = y_n \equiv \frac{1}{n} (x_1 + x_2 + \dots + x_n)$

$\{x_i | i = 1, 2, \dots, n\}$: X 抽样序列

2. 蒙特卡洛方差: $\sigma^2 = \frac{V(X)}{n}$



2.2 减小蒙特卡洛方差的方法

待求解问题为 I 。若 $I = \langle X \rangle$, 由此诱导出的蒙特卡洛估计值的方差为 $\sigma^2 = \frac{V(X)}{n}$ 。

$$\begin{cases} \frac{\partial \sigma^2}{\partial n} < 0 \\ \lim_{n \rightarrow \infty} \sigma^2 = 0 \end{cases} \quad \lim_{n \rightarrow \infty} \frac{\partial \sigma^2}{\partial n} = 0$$

可以看到: 增加样本数 n 是减小蒙特卡洛方差的普适方法之一, 但蒙特卡洛方差减小的速度随 n 的增大趋近于 0。

2.2.1 重要抽样法

选择方差较小的随机变量 X , 是减小蒙特卡洛方差的另一普适方法。连续型随机变量的期望值问题为定积分问题, 因此, 下面的讨论中均以定积分为例。

$$I = \int_a^b f(x) dx$$

选择函数 $h(x)$ 和 $g(x)$, 要求: (1) $g(x)$ 为分布密度函数; (2) $h(x)g(x) = f(x)$ 。则

$$I = \int_a^b h(x)g(x)dx = \langle h(X) \rangle$$

$X = x \in [a, b]$, 分布密度函数为 $g(x)$

$$\therefore I_n = \frac{1}{n} \sum_{i=1}^n h(x_i), \quad \{x_i | i = 1, 2, \dots, n\}: \text{服从 } g(x) \text{ 分布}; \quad \sigma^2 = \frac{V(h)}{n}$$

其中

$$V(h) = \int_a^b \left(\frac{f(x)}{g(x)} - I \right)^2 g(x) dx$$

显然, 选择合适的分布密度函数 $g(x)$, 要求 $\frac{f(x)}{g(x)}$ 在 $[a, b]$ 区间内的变化不要太大, 尽量接近于常数, 则可以有效地减小蒙特卡洛方差。

例 2.1 如图 (2.1) 所示, 曲线 $y = f(x)$ 定域于边长为 1 的正方形区域内, 利用蒙特卡洛方法计算该曲线下面积。

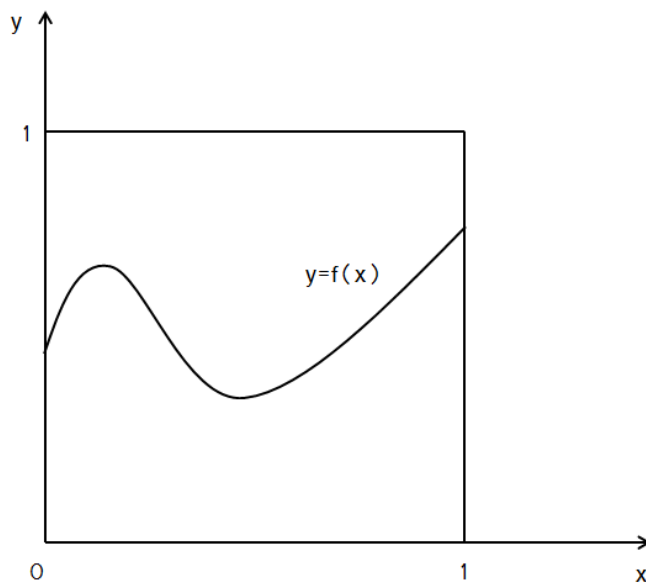


图 2.1: 定积分计算

解 所求解问题为如下一维定积分

$$I = \int_0^1 f(x) dx \quad 0 \leq f(x) \leq 1$$

(1) 蒙特卡洛估计一:

伯努利试验 E : 在 $[0, 1] \times [0, 1]$ 区域内随机投点, 落在曲线 $y = f(x)$ 下方为成功事件; 落在曲线 $y = f(x)$ 上方为失败事件。

注 二维随机向量 $(X, Y) = (x, y) \in [0, 1] \times [0, 1]$ 服从均匀分布。上述的单次伯努利试验相当于对该随机向量 (X, Y) 的一个抽样。 $y \leq f(x)$ 为成功事件; $y > f(x)$ 为失败事件。

伯努利试验 E 诱导出伯努利随机变量 R :

$$R = r \in \{0, 1\} \quad r = h(x, y) = \begin{cases} 1 & y \leq f(x) \\ 0 & y > f(x) \end{cases}$$

$$\begin{aligned} p &\equiv P(r = 1) = P(\{(x, y) \mid 0 \leq y \leq f(x), 0 \leq x \leq 1\}) \\ &= \int_0^1 dx \int_0^{f(x)} dy g(x, y) \\ &= \int_0^1 f(x) dx \end{aligned}$$

其中, $g(x, y) = 1$ 为二维随机向量 (X, Y) 的分布密度函数。 $\therefore I = p = \langle R \rangle$

- 蒙特卡洛估计值:

$$I_n = \frac{1}{n} (r_1 + r_2 + \dots + r_n) = \frac{m}{n}$$

$\{r_i | i = 1, 2, \dots, n\}$ 服从伯努利分布。伯努利变量 R 的抽样可由如下变换法实现。

$$(X, Y) \mapsto R = h(X, Y)$$

- 蒙特卡洛方差:

$$\sigma^2 = \frac{V(R)}{n}$$

(2) 蒙特卡洛估计二:

引入随机变量 Z , 它是 $[0, 1]$ 区间上均匀分布随机变量 X 的函数: $Z = f(X)$ 。则

$$I = \int_0^1 f(x) dx = \langle Z \rangle$$

- 蒙特卡洛估计值:

$$I'_n = \frac{1}{n} (z_1 + z_2 + \dots + z_n)$$

$z_i = f(x_i)$ ($i = 1, 2, \dots, n$), $\{x_i | i = 1, 2, \dots, n\}$ 为 $[0, 1]$ 区间上均匀分布随机数序列。

- 蒙特卡洛方差:

$$\sigma'^2 = \frac{V(Z)}{n}$$

(3) 蒙特卡洛方差比较:

$$\begin{aligned} V(R) - V(Z) &= (I - I^2) - (\langle Z^2 \rangle - I^2) \\ &= I - \langle Z^2 \rangle \\ &= \int_0^1 f(x)[1 - f(x)] dx \geq 0 \quad \therefore \sigma^2 \geq \sigma'^2 \end{aligned}$$

注

I_n 计算:

$$I = \langle R \rangle = \langle h(X, Y) \rangle$$

(X, Y) : \square 上均匀分布

$$= \int h(x, y) dx dy$$

I'_n 计算:

$$I = \langle Z \rangle = \langle f(X) \rangle$$

X : $[0, 1]$ 上均匀分布

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

$$\int_{\square} h(x, y) dx dy = \int_0^1 dx \int_0^{f(x)} dy h(x, y) + \int_0^1 dx \int_{f(x)}^1 dy h(x, y) = \int_{[0,1]} f(x) dx$$

- I_n 是 $h(x, y)$ 二重积分的蒙特卡洛估计;

- $h(x, y)$ 对变量 y 积分解析计算后得到 $f(x)$, I'_n 是 $f(x)$ 一重积分的蒙特卡洛估计。

结论 “可以精确计算的不作蒙特卡洛计算”是多重积分计算中的基本指导原则, 能够帮助有效地减小方差。

2.2.2 分层抽样法

以一维定积分为例,

$$I = \langle Y \rangle = \int_a^b h(x)g(x)dx$$

$X = x \in [a, b]$, 分布密度函数为 $g(x)$, $Y = f(X)$

此期望值诱导的蒙特卡洛估计值和方差分别为

- 估计值:

$$I_n = \frac{1}{n} \sum_{i=1}^n h(x_i) \quad \{x_i | i = 1, 2, \dots, n\}: \text{服从 } g(x) \text{ 分布}$$

- 方差:

$$\sigma^2 = \frac{V(Y)}{n} \quad V(Y) = \int_a^b [h(x) - I]^2 g(x)dx$$

如果对积分区间作分割: $a = x_0 < x_1 < \dots < x_{J-1} < x_J = b$, 则有

$$I = \sum_{j=1}^J p_j I_j \quad I_j = \langle Y_j \rangle = \int_{x_{j-1}}^{x_j} h(x)g_j(x)dx$$

$X_j = x \in [x_{j-1}, x_j]$, 分布密度函数为 $g_j(x)$, $Y_j = h(X_j)$

其中

$$\begin{cases} g_j(x) = \frac{1}{p_j} g(x), & x \in [x_{j-1}, x_j] \\ p_j = \int_{x_{j-1}}^{x_j} g(x)dx \end{cases}$$

给出的蒙特卡洛估计值和方差分别为 (要求: $n = n_1 + n_2 + \dots + n_J$)

- 估计值:

$$I'_n = \sum_{j=1}^J p_j I_{j,n_j}$$

$$I_{j,n_j} = \frac{1}{n_j} \sum_{i=1}^{n_j} h(x_{ij}) \quad \{x_{ij} | i = 1, 2, \dots, n_j\}: \text{服从 } g_j(x) \text{ 分布}$$

- 方差:

$$\sigma'^2 = \sum_{j=1}^J p_j^2 \sigma_j^2$$

$$\sigma_j^2 = \frac{V(Y_j)}{n_j} \quad V(Y_j) = \int_{x_{j-1}}^{x_j} [h(x) - I_j]^2 g_j(x)dx$$

推导:

$$\begin{aligned} V(Y) &= \int_a^b [h(x) - I]^2 g(x)dx \\ &= \sum_{j=1}^J p_j \int_{x_{j-1}}^{x_j} [(h(x) - I_j)^2 + (I_j - I)^2 + 2(I_j - I)(h(x) - I_j)] g_j(x)dx \\ &= \sum_{j=1}^J p_j V(Y_j) + \sum_{j=1}^J p_j (I_j - I)^2 \end{aligned}$$

因此，有

$$\begin{aligned}\sigma^2 - \sigma'^2 &= \frac{1}{n} \left[\sum_{j=1}^J p_j V(Y_j) + \sum_{j=1}^J p_j (I_j - I)^2 \right] - \sum_{j=1}^J \frac{p_j^2}{n_j} V(Y_j) \\ &= \sum_{j=1}^J p_j \left(\frac{1}{n} - \frac{p_j}{n_j} \right) V(Y_j) + \frac{1}{n} \sum_{j=1}^J p_j (I_j - I)^2\end{aligned}$$

可以看到：若取

$$\frac{1}{n} = \frac{p_j}{n_j} \longrightarrow n_j = np_j$$

分层抽样可以减小蒙特卡洛方差。



笔记

减除法：被积函数 $f(x)$ 在积分区间内存在 (pseudo-)singularity 是常见的。例如：当我们利用量子场论来研究基本粒子散射时，需要计算圈积分及末态相空间积分。这些积分中会遇到（赝）极点问题，直接采用蒙特卡洛方法计算是不可行的。如果我们清楚被积函数在（赝）极点附近的解析行为，则可以采用合适的减除方法来完成蒙特卡洛计算，减小蒙特卡洛方差。

在减除法中，我们需要引入减除项 $g(x)$ ，不同的减除项对应不同的减除方案。

$$\int f(x) dx = \int [f(x) - g(x)] dx + \int g(x) dx$$

减除项 $g(x)$ 需要满足如下条件：

1. $g(x)$ 与 $f(x)$ 有相同的极点行为；
2. $g(x)$ 在整个区间上的积分可以精确计算。

2.3 一维随机变量抽样方法

随机变量的抽样方法很多，不同分布随机变量采用的抽样方法不尽相同。其中， $[0, 1]$ 区间上均匀分布随机变量的抽样是最基本的，其它各种分布随机变量的抽样可在此基础上实现。

真随机数序列除了统计规律外完全无法预计，因而真随机数序列是不可重复产生的。真随机数序列需要借助物理随机过程来产生，例如不稳定粒子的衰变。实际蒙特卡洛模拟中使用的是按照某种算法产生的伪随机数。伪随机数的产生方法可以表述为如下形式的递推公式：

$$\xi_{n+k} = G(\xi_n, \xi_{n+1}, \dots, \xi_{n+k-1}) \quad (n = 1, 2, \dots)$$

其中， $\xi_{1,2,\dots,k}$ 为给定的初始值。

例 2.2 线性同余伪随机数产生器。

$$\begin{cases} x_n = (\lambda x_{n-1} + c) \bmod M \\ \xi_n = x_n / M \end{cases} \quad (n = 1, 2, \dots)$$

由此得到 $[0, 1]$ 区间上均匀分布伪随机数序列 $\{\xi_n \mid n = 1, 2, \dots\}$ 。

注 伪随机数序列的两个特点：

1. 整个伪随机数序列由递推公式和初始值唯一确定，因此，伪随机数序列不满足随机数相互独立的要求。
2. 计算机上所能表示的 $[0, 1]$ 区间上的数是有限多的，而伪随机数序列由递推公式给出，因此，伪随机数序列存在周期性。

(有些递推公式本身就会给出周期性。例如：线性同余法给出的伪随机数序列，周期不会大于 M 。)

选择合适的递推公式，可以让伪随机数序列的周期足够大（大于蒙特卡洛计算/模拟需要的随机数的个数），且相互独立性也近似满足。

2.3.1 变换抽样法

(1) 离散型随机变量

定理 2.7

$X = x \in [0, 1]$, $R = r \in \{r_i | i = 1, 2, \dots, n\}$ 。

$$\text{映射 } \phi: X \mapsto R \quad \phi(x) = \begin{cases} r_1 & q_0 \leq x \leq q_1 \\ r_2 & q_1 < x \leq q_2 \\ \dots & \\ r_n & q_{n-1} < x \leq q_n \end{cases}$$

$$(0 = q_0 < q_1 < \dots < q_{n-1} < q_n = 1)$$

若: $f(x) = 1$, 则

$$P(R = r_i) = q_i - q_{i-1}$$



(2) 连续型随机变量

定理 2.8

$$X \mapsto Y: \text{一一映射} \quad \begin{cases} X: \text{服从 } f(x) \text{ 分布} \\ Y: \text{服从 } g(y) \text{ 分布} \end{cases}$$

则

$$f(x)dx = \pm g(y)dy$$



特别地，若 X 为 $[0, 1]$ 区间上均匀分布随机变量，即：

$$X = x \in [0, 1], \quad f(x) = 1$$

1. 单调递增映射：

$$dx = +g(y)dy \xrightarrow{y(x=0)=y_{\min}} \int_0^x dx' = + \int_{y_{\min}}^y g(y')dy'$$

2. 单调递减映射：

$$dx = -g(y)dy \xrightarrow{y(x=1)=y_{\min}} \int_1^x dx' = - \int_{y_{\min}}^y g(y')dy'$$

因此, 映射为

$$\begin{cases} Y = G^{-1}(X) & (\text{单调递增}) \\ Y = G^{-1}(1 - X) & (\text{单调递减}) \end{cases}$$

注 X 为 $[0, 1]$ 区间上均匀分布随机变量时, $(1 - X)$ 也是 $[0, 1]$ 区间上均匀分布随机变量。

例 2.3 指数衰减分布随机变量抽样。

解 衰减常数为 λ 的指数分布:

$$F(x) = 1 - e^{-\lambda x} \longrightarrow F^{-1}(x) = -\frac{1}{\lambda} \ln(1 - x)$$

定义映射:

$$\phi: X \longmapsto Y$$

$$y = \phi(x) \equiv F^{-1}(1 - x) = -\frac{1}{\lambda} \ln x$$

X 为 $[0, 1]$ 区间上均匀分布随机变量时, Y 为指数衰减分布随机变量, 衰减常数为 λ 。



笔记

变换法: 更一般的变换法是随机向量变换, 且不必是一一映射。例如下面两个例子。

命题 2.2

X_i ($i = 1, 2, \dots, n$) 为相互独立的 $[0, 1]$ 区间上均匀分布随机变量。

$$X = \max\{X_1, X_2, \dots, X_n\}$$

则

$$X = x \in [0, 1], \text{ 分布密度函数为: } f(x) = nx^{n-1}$$



证明

$$\begin{aligned} F(x) &\equiv P(X = \max\{X_1, X_2, \dots, X_n\} \leq x) \\ &= P(X_1 \leq x, X_2 \leq x, \dots, X_n \leq x) \\ &= x^n \end{aligned}$$

命题 2.3

$(U, V) = (u, v) \in [0, 1]^2$, $(X, Y) = (x, y) \in \mathbb{R}^2$ 。

$$\text{映射 } \phi: (U, V) \longmapsto (X, Y) \quad \begin{cases} x = \sqrt{-2 \ln u} \cos 2\pi v \\ y = \sqrt{-2 \ln u} \sin 2\pi v \end{cases}$$

若: $g(u, v) = 1$, 则

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



证明

$$\begin{cases} x^2 + y^2 = -2 \ln u \\ \frac{y}{x} = \tan 2\pi v \end{cases} \longrightarrow \begin{cases} du = -u(xdx + ydy) \\ dv = \frac{1}{2\pi(x^2 + y^2)}(-ydx + xdy) \end{cases}$$

$$du \wedge dv = -\frac{u}{2\pi(x^2 + y^2)} (xdx + ydy) \wedge (-ydx + xdy) = -\frac{u}{2\pi} dx \wedge dy$$

$$\therefore \mathcal{J} = \left| \det \frac{\partial(u, v)}{\partial(x, y)} \right| = \frac{u}{2\pi} \quad f(x, y) = g(u, v) \times \mathcal{J}$$

2.3.2 舍选抽样法

(1) 第一类舍选法

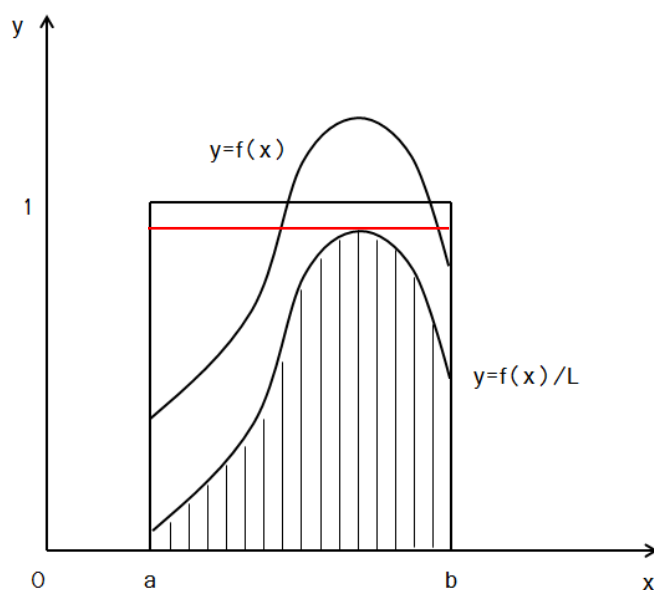


图 2.2: 第一类舍选法

$$X = x \in [a, b], \text{ 分布密度为 } f(x). \quad L = \max \{f(x) | a \leq x \leq b\}$$

抽样步骤

1. 在 $[a, b] \times [0, 1]$ 长方形区域内均匀投点，即：对 $[a, b] \times [0, 1]$ 区域内均匀分布二维随机向量 (U, V) 抽样。

$$\{(u_i, v_i) | i = 1, 2, \dots, n\}$$

2. 保留曲线 $v = f(u)/L$ 下方的点，取其横坐标。

$$\{u_i | v_i \leq f(u_i)/L, i = 1, 2, \dots, n\} = \{x_1, x_2, \dots, x_m\}$$

★ 舍选效率期望值：

$$E = \frac{1}{(b-a)L}$$

证明

$$\begin{aligned} P(X \leq x) &= P(U \leq x | V \leq f(U)/L) = \frac{P(U \leq x, V \leq f(U)/L)}{P(V \leq f(U)/L)} \\ &= \left[\int_a^x du \int_0^{f(u)/L} dv g(u, v) \right] \bigg/ \left[\int_a^b du \int_0^{f(u)/L} dv g(u, v) \right] \\ &= \int_a^x du f(u) \quad \text{其中: } g(u, v) = \frac{1}{(b-a)} \end{aligned}$$

$$E = P(V \leq f(U)/L) = \int_a^b du \int_0^{f(u)/L} dv g(u, v) = \frac{1}{(b-a)L}$$

例 2.4 $X = x \in [0, 1]$, 分布密度函数为 $f(x) = 2x$ 。利用舍选法对该随机变量抽样。

解 分布密度函数定出判选函数:

$$L = \max\{f(x) | 0 \leq x \leq 1\} = 2 \quad f(x)/L = x$$

由第一类舍选法, 抽样结果为

$$\{u_i | v_i \leq u_i, i = 1, 2, \dots, n\}$$

其中, $\{u_i | i = 1, 2, \dots, n\}$ 和 $\{v_i | i = 1, 2, \dots, n\}$ 均为 $[0, 1]$ 区间上均匀分布随机数序列。

注 上述抽样中, $\{u_i | i = 1, 2, \dots, n\}$ 为备选序列, $\{v_i | i = 1, 2, \dots, n\}$ 为判选序列。考虑到这两个序列是同分布的, 我们也可将 $\{v_i | i = 1, 2, \dots, n\}$ 当作备选序列, $\{u_i | i = 1, 2, \dots, n\}$ 当作判选序列。这样得到的抽样结果为

$$\{v_i | u_i \leq v_i, i = 1, 2, \dots, n\}$$

显然, 优化后的抽样结果为二者之结合:

$$\{\max\{u_i, v_i\} | i = 1, 2, \dots, n\}$$

(2) 第二类舍选法

$$X = x \in \mathbb{R}, \text{ 分布密度为 } f(x) = Lh(x)g(x). \quad \max\{h(x) | x \in \mathbb{R}\} \leq 1$$

$g(x)$ 为分布密度函数

抽样步骤

1. $U = u \in \mathbb{R}$, 分布密度为 $g(u)$ 。对随机变量 U 抽样, 得到**备选序列**。

$$\{u_i | i = 1, 2, \dots, n\}$$

2. $V = v \in [0, 1]$, 分布密度为 1。对随机变量 V 抽样, 得到**判选序列**。

$$\{v_i | i = 1, 2, \dots, n\}$$

3. 舍选: 保留备选序列中满足判选条件 $v_i \leq h(u_i)$ 的随机数。

$$\{u_i | v_i \leq h(u_i), i = 1, 2, \dots, n\} = \{x_1, x_2, \dots, x_m\}$$

★ 舍选效率期望值:

$$E = \frac{1}{L}$$

证明

$$\begin{aligned} P(X \leq x) &= P(U \leq x | V \leq h(U)) = \frac{P(U \leq x, V \leq h(U))}{P(V \leq h(U))} \\ &= \left[\int_{-\infty}^x du g(u) \int_0^{h(u)} dv \right] / \left[\int_{-\infty}^{+\infty} du g(u) \int_0^{h(u)} dv \right] \\ &= \int_{-\infty}^x du f(u) \\ E &= P(V \leq h(U)) = \int_{-\infty}^{+\infty} du g(u) \int_0^{h(u)} dv = \frac{1}{L} \end{aligned}$$

例 2.5 采用舍选法对随机变量 X 抽样。 $X = x \in [0, +\infty)$, 分布密度函数为

$$f(x) = 2 \times \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

解 X 的取值范围 $[0, +\infty)$ 为无穷区间, 因此, 第一类舍选法无法实现, 需要采用第二类舍选法对其抽样。考虑到 $f(x)$ 在 $[0, +\infty)$ 内单调递减, 我们要求备选序列服从衰减常数为 1 的指数衰减分布。这样选择, 不仅备选序列容易抽样得到, 而且舍选效率较高。

$$f(x) = Lh(x)g(x), \quad g(x) = e^{-x}$$

判选函数 $h(x)$:

$$Lh(x) = \frac{f(x)}{g(x)} = \sqrt{\frac{2e}{\pi}} e^{-\frac{1}{2}(x-1)^2} \longrightarrow \begin{cases} h(x) = e^{-\frac{1}{2}(x-1)^2} \\ L = \sqrt{\frac{2e}{\pi}} \end{cases}$$

抽样结果:

$$\{u_i | v_i \leq h(u_i), i = 1, 2, \dots, n\} \quad \text{其中, } \{u_i | i = 1, 2, \dots, n\}: [0, +\infty] \text{ 上指数衰减分布}$$

$$\{v_i | i = 1, 2, \dots, n\}: [0, 1] \text{ 上均匀分布}$$

结论

- 第一类舍选法是第二类舍选法的特例:

$$g(x) = \frac{1}{b-a}$$

优点是易于抽样, 但当分布密度函数 $f(x)$ 在 $[a, b]$ 区间变化剧烈时, 舍选效率较低。

- 第二类舍选法是第一类舍选法的改进。选择

$$\frac{g(x)}{f(x)} \sim \text{const.}$$

舍选效率会大大提高, 但备选序列的抽样较为复杂。

2.3.3 叠加原理

定理 2.9

$$\mathcal{G}(x) = \begin{cases} \{g_i(x) | i = 1, 2, \dots, n\} & \text{概率: } p_i \quad (\text{离散型}) \\ \{g(x|y) | y \in \mathbb{R}\} & \text{分布密度: } h(y) \quad (\text{连续型}) \end{cases}$$

其中: $g_i(x)$, $g(x|y)$ 为分布密度函数。

$$f(x) \equiv \langle \mathcal{G}(x) \rangle = \begin{cases} \sum_{i=1}^n p_i g_i(x) & (\text{离散型}) \\ \int_{\mathbb{R}} dy h(y) g(x|y) & (\text{连续型}) \end{cases}$$

叠加原理:

1. 依概率 p_i 或分布密度 $h(y)$ 抽样, 得 i 或 y ;
2. 依分布密度 $g_i(x)$ 或 $g(x|y)$ 抽样, 得 x 。

则: x 服从 $f(x)$ 分布。



例 2.6 $X = x \in [a, b]$, 分布密度函数为 $f(x) \sim x^3$ 。对该随机变量抽样。

解 $[0, 1]$ 区间上幂函数分布随机变量的抽样方法是已知的。因此, 引入如下的线性映射:

$$X = a + (b - a)Y$$

$Y = y \in [0, 1]$, 分布密度函数

$$\begin{aligned} g(y) &\sim f(x) \sim [a + (b - a)y]^3 \\ &= a^3 + 3a^2(b - a)y + 3a(b - a)^2y^2 + (b - a)^3y^3 \\ &= \sum_{i=1}^4 \lambda_i g_i(y) \end{aligned}$$

其中

$$\begin{cases} \lambda_1 = a^3 \\ \lambda_2 = \frac{3}{2}a^2(b - a) \\ \lambda_3 = a(b - a)^2 \\ \lambda_4 = \frac{1}{4}(b - a)^3 \end{cases} \quad g_i(y) = iy^{i-1} \quad (i = 1, 2, 3, 4)$$

$$\therefore g(y) = \sum_{i=1}^4 p_i g_i(y) \quad \text{权重因子: } p_i = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4}$$

抽样方法简述如下:

1. 依权重因子 (概率) p_i ($i = 1, 2, 3, 4$) 抽样, 确定分布密度函数 $g_i(y)$;
2. 依分布密度函数 $g_i(y)$ 抽样, 得 y ;
3. 作映射: $x = a + (b - a)y$ 。



笔记

近似抽样法: 利用叠加原理, 可以方便地对复杂分布随机变量抽样。

命题 2.4

$$f(x) = pf_1(x) + qf_2(x)$$

1. 分布密度函数 $f_1(x)$:

$$\frac{f_1(x) \text{ (简单, 易于抽样)}}{f(x) \text{ (复杂, 不易抽样)}} \sim 1$$

2. 权重因子 p :

$$pf_1(x) \leq f(x) \quad \forall x \in \mathbb{R} \quad \xrightarrow{\text{尽可能大}} \quad p = \min \left\{ \frac{f(x)}{f_1(x)} \mid f_1(x) \neq 0, x \in \mathbb{R} \right\}$$

3. 修正项 $qf_2(x)$:

$$q = 1 - p \quad f_2(x) = \frac{f(x) - pf_1(x)}{q}$$

结论:

大部分成员依 $f_1(x)$ 抽样 (approximation); 少部分成员依 $f_2(x)$ 抽样 (correction)。



2.4 高维随机向量抽样方法

高维随机向量的每个坐标均为一维随机变量，因此，高维随机向量抽样问题可以转化为一维随机变量抽样问题。此外，第一类舍选抽样法也适用于高维随机向量。

2.4.1 条件分布密度法

$(X_1, X_2, \dots, X_n) = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ ，分布密度为 $f(x_1, x_2, \dots, x_n)$ 。

定义 2.3

(1) 边缘分布密度:

$$f(x_{i \in \mathcal{A}}) = \int f(x_1, x_2, \dots, x_n) \prod_{j \notin \mathcal{A}} dx_j \quad \mathcal{A} \subset \{1, 2, \dots, n\}$$

(2) 条件分布密度:

$$f(x_{i \in \mathcal{A}} | x_{j \in \mathcal{B}}) = \frac{f(x_{i \in \mathcal{A}}, x_{j \in \mathcal{B}})}{f(x_{j \in \mathcal{B}})} \quad \mathcal{A} \cap \mathcal{B} = \emptyset, \quad \mathcal{A}, \mathcal{B} \subset \{1, 2, \dots, n\}$$



据上述条件分布密度定义，得到如下的乘法定理：

定理 2.10

$$f(x_1, x_2, \dots, x_n) = f(x_1) f(x_2 | x_1) f(x_3 | x_1, x_2) \cdots f(x_n | x_1, x_2, \dots, x_{n-1})$$



抽样步骤

1. 依分布密度 $f(x_1)$ 抽样，得 x_1 ；

2. For $i = 2$ to n ,

依条件分布密度 $f(x_i | x_1, x_2, \dots, x_{i-1})$ 抽样，得 x_i 。

注 条件分布密度法下，高维随机向量的抽样即为其各个坐标（一维随机变量）的抽样。乘法定理具体形式决定了各坐标的抽样顺序。

例 2.7 中子入射角 (Φ, Θ) 抽样。

$$f(\phi, \theta) \sim (1 + \sqrt{3} \sin \phi \sin \theta) \sin \phi \sin^2 \theta \quad \phi, \theta \in [0, \pi/2]$$

解 作变换：

$$(\Phi, \Theta) \mapsto (X, Y) \quad \begin{cases} x = \cos \phi \\ y = \cos \theta \end{cases} \quad x, y \in [0, 1]$$

(X, Y) 的分布密度：

$$g(x, y) = f(\phi, \theta) \times \left| \det \frac{\partial(x, y)}{\partial(\phi, \theta)} \right|^{-1} \sim \left[1 + \sqrt{3} \sqrt{(1-x^2)(1-y^2)} \right] \sqrt{1-y^2}$$

X 的边缘分布密度：

$$g(x) = \int_0^1 g(x, y) dy \sim \frac{\pi}{4} + \frac{2}{\sqrt{3}} \sqrt{1-x^2}$$

Y 的条件分布密度：

$$g(y|x) = \frac{g(x, y)}{g(x)} \sim g(x, y)$$

叠加原理:

$$g(x) = p_1 g_1(x) + p_2 g_2(x)$$

$$\begin{cases} g_1(x) = 1 & p_1 \sim 1 \\ g_2(x) = \frac{4}{\pi} \sqrt{1-x^2} & p_2 \sim \frac{2}{\sqrt{3}} \end{cases}$$

$$g(y|x) = q_1 h_1(y) + q_2 h_2(y)$$

$$\begin{cases} h_1(y) = \frac{4}{\pi} \sqrt{1-y^2} & q_1 \sim \frac{\pi}{4} \\ h_2(y) = \frac{3}{2}(1-y^2) & q_2 \sim \frac{2}{\sqrt{3}} \sqrt{1-x^2} \end{cases}$$

抽样流程简述:

- (1) X 抽样: 依分布密度 $g_1(x)$ 或 $g_2(x)$ 抽样 (概率分别为 p_1 和 p_2), 得 x ;
- (2) Y 抽样: 依分布密度 $h_1(x)$ 或 $h_2(x)$ 抽样 (概率分别为 q_1 和 q_2), 得 y ;
- (3) (Φ, Θ) 抽样值: $(\phi, \theta) = (\arccos x, \arccos y)$ 。

两点说明:

- 条件分布密度函数 $g(y|x)$ 对条件的依赖反映于权重因子的 x -dependence。

$$q_1 = \frac{\frac{\pi}{4}}{\frac{\pi}{4} + \frac{2}{\sqrt{3}} \sqrt{1-x^2}} \quad q_2 = \frac{\frac{2}{\sqrt{3}} \sqrt{1-x^2}}{\frac{\pi}{4} + \frac{2}{\sqrt{3}} \sqrt{1-x^2}}$$

- 一维随机变量抽样。

$g_1(x):$ (均匀分布)	$g_2(x), h_1(x):$ (第一类舍选法) $x^2 + y^2 \leq 1$	$h_2(x):$ (第一类舍选法) $x^2 + y \leq 1$
---------------------	--	--

2.4.2 舍选抽样法

第一类舍选法可以很方便地推广至高维随机向量抽样。

$$\vec{X} = \vec{x} \in \mathcal{D}, \text{ 分布密度为 } f(\vec{x}). \quad L = \max\{f(\vec{x}) \mid \vec{x} \in \mathcal{D}\}$$

$$\mathcal{D} = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$$

抽样步骤

1. 在 $\mathcal{D} \times [0, 1] \subset \mathbb{R}^{n+1}$ 区域内均匀投点, 即: 对 $\mathcal{D} \times [0, 1]$ 区域内均匀分布 $(n+1)$ 维随机向量 (\vec{U}, V) 抽样。

$$\{(\vec{u}_i, v_i) \mid i = 1, 2, \dots, N\}$$

2. 保留超曲面 $v = f(\vec{u})/L$ 下方的点, 取其前 n 坐标。

$$\{\vec{u}_i \mid v_i \leq f(\vec{u}_i)/L, i = 1, 2, \dots, N\} = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_M\}$$

★ 舍选效率期望值:

$$E = \left[L \prod_{i=1}^n (b_i - a_i) \right]^{-1}$$

第三章 蒙特卡洛方法应用

3.1 多体末态相空间产生

与经典物理不同，量子理论中描述一个自由粒子运动的状态参量为

$$\left(\vec{p} \text{ (外部运动: 动量)}, \quad \lambda \text{ (内禀运动: 极化)} \right)$$

不考虑粒子的内禀自由度（例如：自旋为 0 的标量和赝标量粒子）， n 个自由粒子的运动状态 $(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n)$ 即为 $\text{dim} = 3n$ 相空间中一点。粒子相互作用过程分为如下两类：

$$\text{衰变: } A \longrightarrow 1 + 2 + \dots + n$$

$$\text{散射: } a + b \longrightarrow 1 + 2 + \dots + n$$

反应末态 $(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n)$ 为随机向量，分布密度由相互作用动力学给出。

定义 3.1

$\text{dim} = 3n$ 末态相空间体积元：

$$d\Phi_n(P; p_1, \dots, p_n) = (2\pi)^4 \delta^4(P - \sum_{i=1}^n p_i) \prod_{i=1}^n \frac{d^3 \vec{p}_i}{(2\pi)^3 2E_i}$$

其中

$$\text{初态总能动量: } P = (E, \vec{P})$$

$$\text{末态各粒子能动量: } p_i = (E_i, \vec{p}_i), \quad E_i = \sqrt{\vec{p}_i^2 + m_i^2}$$



衰变和散射过程的反应率分别由微分宽度和微分截面描述，且依赖于参考系。质心系下，

- 微分宽度：

$$d\Gamma = \frac{1}{2M} |\mathcal{M}|^2 d\Phi_n(P; p_1, \dots, p_n) \quad P = p_A = (M, \vec{0})$$

- 微分截面：

$$d\sigma = \frac{1}{4|\vec{p}_a| M} |\mathcal{M}|^2 d\Phi_n(P; p_1, \dots, p_n) \quad P = p_a + p_b = (M, \vec{0})$$

因此，

$$f(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n) \prod_{i=1}^n d^3 \vec{p}_i \sim \begin{cases} d\Gamma & (\text{衰变}) \\ d\sigma & (\text{散射}) \end{cases}$$

力学量 \mathcal{O} 的期望值为

$$\langle \mathcal{O} \rangle = \int \mathcal{O}(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n) f(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n) \prod_{i=1}^n d^3 \vec{p}_i$$

3.1.1 事例

相空间点也称为事例，记为 $e \equiv (\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n)$ 。依分布密度 $f(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n)$ 抽样得到的事例称为非加权事例，否则称为加权事例。

定理 3.1

$$\langle \mathcal{O} \rangle = \int O(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n) f(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n) \prod_{i=1}^n d^3 \vec{p}_i$$

1. 非加权事例：

$$\langle \mathcal{O} \rangle \simeq \frac{1}{N} \sum_{i=1}^N O(e_i)$$

2. 加权事例：

$$\langle \mathcal{O} \rangle \simeq \frac{1}{N} \sum_{i=1}^N \omega(e_i) O(e_i) \quad \omega(e) = \frac{f(e)}{g(e)} \quad (\text{权重因子})$$

加权事例服从 $g(e)$ 分布



3.1.2 两体末态相空间

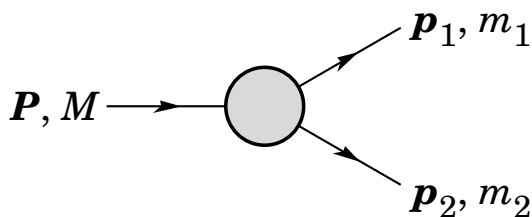


图 3.1: 两体末态相空间

两体末态相空间体积元：

$$\begin{aligned} d\Phi_2(P; p_1, p_2) &= (2\pi)^4 \delta^4(p_1 + p_2 - P) \frac{d^3 \vec{p}_1}{(2\pi)^3 2E_1} \frac{d^3 \vec{p}_2}{(2\pi)^3 2E_2} \\ &= \frac{1}{(2\pi)^2} \frac{1}{4E_1 E_2} \delta(E_1 + E_2 - E) |\vec{p}_1|^2 d|\vec{p}_1| d\Omega_1 \\ &= \frac{1}{(2\pi)^2} \frac{|\vec{p}_1|^2}{4E_1 E_2} \left[1 / \frac{\partial(E_1 + E_2)}{\partial|\vec{p}_1|} \right] d\Omega_1 \end{aligned}$$

其中

$$\vec{p}_1 + \vec{p}_2 = \vec{P} \quad E_1 + E_2 = E$$

定义 3.2

以 \hat{n} 为参考方向，

$$\text{投影算符: } \begin{cases} \mathcal{P}_L \equiv \hat{n}\hat{n} & (\text{纵向}) \\ \mathcal{P}_T \equiv 1 - \hat{n}\hat{n} & (\text{横向}) \end{cases}$$

任意向量 \vec{r} ，有

$$\vec{r} = \vec{r}_L + \vec{r}_T \quad \begin{cases} \vec{r}_L \equiv \mathcal{P}_L \cdot \vec{r} = r_L \hat{n} & r_L \equiv \vec{r} \cdot \hat{n} \\ \vec{r}_T \equiv \mathcal{P}_T \cdot \vec{r} = \vec{r} - \vec{r}_L \end{cases}$$



命题 3.1

雅可比因子:

$$\frac{\partial(E_1 + E_2)}{\partial |\vec{p}_1|} = \left(\frac{\vec{p}_1}{E_1} - \frac{\vec{p}_2}{E_2} \right) \cdot \hat{p}_1$$

**证明**

$$E^2 = \vec{p}^2 + m^2 \quad \longrightarrow \quad d(E_1 + E_2) = \frac{1}{2E_1} d|\vec{p}_1|^2 + \frac{1}{2E_2} d|\vec{p}_2|^2$$

以 \hat{p}_1 为参考方向,

$$\vec{p}_2 = \vec{p}_{2,L} + \vec{p}_{2,T}$$

由动量守恒

$$\vec{p}_2 = \vec{P} - \vec{p}_1 \quad \longrightarrow \quad \begin{cases} \vec{p}_{2,L} = \vec{P}_L - \vec{p}_1 & p_{2,L} = P_L - |\vec{p}_1| \\ \vec{p}_{2,T} = \vec{P}_T \end{cases}$$

$$\mathcal{P}_L = \hat{p}_1 \hat{p}_1$$

$$\mathcal{P}_T = 1 - \hat{p}_1 \hat{p}_1$$

 $\longrightarrow \vec{P}_{L,T}$ 依赖 $\cos \theta_1$ 和 ϕ_1 , 但与 $|\vec{p}_1|$ 无关。

因此,

$$\begin{cases} dP_L \wedge d\Omega_1 = 0 \\ d|\vec{P}_T| \wedge d\Omega_1 = 0 \end{cases} \quad \longrightarrow \quad \begin{cases} dp_{2,L} \wedge d\Omega_1 = -d|\vec{p}_1| \wedge d\Omega_1 \\ d|\vec{p}_{2,T}| \wedge d\Omega_1 = 0 \end{cases}$$

代入 $d(E_1 + E_2)$, 有

$$\begin{aligned} d(E_1 + E_2) \wedge d\Omega_1 &= \left(\frac{1}{2E_1} d|\vec{p}_1|^2 + \frac{1}{2E_2} d|\vec{p}_2|^2 \right) \wedge d\Omega_1 \\ &= \left[\frac{1}{E_1} |\vec{p}_1| d|\vec{p}_1| + \frac{1}{E_2} \left(p_{2,L} dp_{2,L} + |\vec{p}_{2,T}| d|\vec{p}_{2,T}| \right) \right] \wedge d\Omega_1 \\ &= \left(\frac{|\vec{p}_1|}{E_1} - \frac{p_{2,L}}{E_2} \right) d|\vec{p}_1| \wedge d\Omega_1 \end{aligned}$$

即

$$d(E_1 + E_2) \wedge d\Omega_1 = \left(\frac{\vec{p}_1}{E_1} - \frac{\vec{p}_2}{E_2} \right) \cdot \hat{p}_1 d|\vec{p}_1| \wedge d\Omega_1$$

定理 3.2

质心系下,

$$d\Phi_2(P; p_1, p_2) = \frac{1}{(2\pi)^2} \frac{|\vec{p}_1|}{4M} d\Omega_1$$

**注** 能动量守恒:

$$\text{映射: } (\cos \theta_1, \phi_1) \longmapsto (\vec{p}_1, \vec{p}_2)$$

- 在变换法意义下, (\vec{p}_1, \vec{p}_2) 的抽样即为 $(\cos \theta_1, \phi_1)$ 的抽样。
- 上述映射与参考系有关。

因此, 求解出上述映射是变换法抽样 (\vec{p}_1, \vec{p}_2) (即: 两体末态事例) 的前提。

定理 3.3

质心系下, $\vec{p}_1 = -\vec{p}_2$ 。

$$|\vec{p}_1| = \frac{1}{2M} \sqrt{\lambda(M^2, m_1^2, m_2^2)} \quad \begin{cases} E_1 = \frac{1}{2M}(M^2 + m_1^2 - m_2^2) \\ E_2 = \frac{1}{2M}(M^2 + m_2^2 - m_1^2) \end{cases}$$

$$\lambda(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2yz - 2zx$$

映射: $(\cos \theta_1, \phi_1) \mapsto (\vec{p}_1, \vec{p}_2)$

$$\begin{cases} \vec{p}_1 = |\vec{p}_1| (\sin \theta_1 \cos \phi_1, \sin \theta_1 \sin \phi_1, \cos \theta_1) \\ \vec{p}_2 = -\vec{p}_1 \end{cases}$$

**加权事例产生**

- 依均匀分布抽样 $(\cos \theta_1, \phi_1)$, 再作映射: $(\cos \theta_1, \phi_1) \mapsto (\vec{p}_1, \vec{p}_2)$ 。
- 权重因子:

$$\omega = \frac{f(\vec{p}_1, \vec{p}_2) d^3 \vec{p}_1 d^3 \vec{p}_2}{\frac{1}{4\pi} d\Omega_1} = \frac{d\Phi_2(P; p_1, p_2)}{\frac{1}{4\pi} d\Omega_1} \times \begin{cases} \frac{1}{2M} |\mathcal{M}|^2 / \Gamma & (\text{衰变}) \\ \frac{1}{4|\vec{p}_a| M} |\mathcal{M}|^2 / \sigma & (\text{散射}) \end{cases}$$

(质心系)

注 散射过程:

$$M = \sqrt{P^2} \quad (\text{初态不变质量}) \quad |\vec{p}_{a,\text{cm}}| = \frac{1}{2M} \sqrt{\lambda(M^2, m_a^2, m_b^2)}$$

3.1.3 多体末态相空间

因子化是多体末态相空间产生的普适方法。

定理 3.4

$$d\Phi_n(P; p_1, \dots, p_n) = d\Phi_{n-j+1}(P; q, p_{j+1}, \dots, p_n) d\Phi_j(q; p_1, \dots, p_j) \frac{1}{2\pi} dq^2$$



$j = 2$, 有

$$d\Phi_n = d\Phi_{n-1} \otimes d\Phi_2$$

反复利用上述因子化定理, 得

推论 3.1

$$d\Phi_n(P; p_1, \dots, p_n) = \prod_{j=2}^n d\Phi_2(q_j; q_{j-1}, p_j) \times \prod_{j=2}^{n-1} \frac{1}{2\pi} dM_j^2$$

其中

$$q_n = P \quad q_1 = p_1$$

$$M_j = \sqrt{q_j^2} \quad (j = 2, 3, \dots, n-1)$$



因此,

$$d\Phi_n = \underbrace{d\Phi_2 \otimes d\Phi_2 \otimes \cdots \otimes d\Phi_2}_{(n-1)}$$

n 体相空间因子化为 $(n-1)$ 个两体相空间。

命题 3.2

$$d\Phi_n(P; p_1, \dots, p_n) = (2\pi)^4 \delta^4(P - \sum_{i=1}^n p_i) \prod_{i=1}^n \theta(p_i^0) \delta(p_i^2 - m_i^2) \frac{d^4 p_i}{(2\pi)^3}$$

(顺时正洛伦兹不变)



证明

$$\delta(p^2 - m^2) = \delta(p^{02} - E^2) = \frac{1}{2E} [\delta(p^0 - E) + \delta(p^0 + E)] \quad E = \sqrt{\vec{p}^2 + m^2}$$

因此, 有

$$\theta(p^0) \delta(p^2 - m^2) \frac{d^4 p}{(2\pi)^3} = \frac{d^3 \vec{p}}{(2\pi)^3 2E}$$

注 利用相空间因子化及顺时正洛伦兹不变性, 可以实现多体末态相空间产生。

(1) 两体相空间

$$(\text{洛伦兹变换}) \quad \Lambda_j : \begin{cases} q_j & \mapsto q'_j = (M_j, 0, 0, 0) \\ q_{j-1} & \mapsto q'_{j-1} \\ p_j & \mapsto p'_j \end{cases}$$

$$d\Phi_2(q_j; q_{j-1}, p_j) = d\Phi_2(q'_j; q'_{j-1}, p'_j) = \frac{1}{(2\pi)^2} \frac{|\vec{p}'_j|}{4M_j} d\Omega'_j \quad (\text{顺时正洛伦兹不变性})$$

$$\downarrow$$

$$d\Omega'_j = d(\cos \theta'_j) d\phi'_j$$

(2) 不变质量

$$dM_j^2 = 2M_j dM_j$$

结论

• $(3n-4)$ 维随机向量:

$$\left(\underbrace{\cos \theta'_j, \phi'_j}_{j=n, \dots, 2}, \underbrace{M_j}_{j=n-1, \dots, 2} \right)$$

• 抽样分布密度函数:

1. 不变质量

$$g(M_j) = \frac{1}{M_{j+1} - \hat{m}_{j+1}} \quad M_j \in [\hat{m}_j, M_{j+1} - m_{j+1}] \quad \hat{m}_j \equiv \sum_{i=1}^j m_i$$

2. 方位角

$$g(\cos \theta'_j, \phi_j) = \frac{1}{4\pi} \quad (\cos \theta'_j, \phi'_j) \in [-1, 1] \times [0, 2\pi]$$

加权事例产生

For $j = n$ to 2,

- 确定 Λ_j 。

$$(\text{洛伦兹变换}) \quad \Lambda_j : \quad q_j \quad \longmapsto \quad q'_j = (M_j, 0, 0, 0)$$

- 抽样 M_{j-1} 。

$$(\text{均匀分布}) \quad M_{j-1} \in [\widehat{m}_{j-1}, M_j - m_j] \quad \quad M_1 \equiv m_1 \quad M_n \equiv M$$

- 抽样 $(\cos \theta'_j, \phi'_j)$ 。

$$(\text{均匀分布}) \quad (\cos \theta'_j, \phi'_j) \in [-1, 1] \times [0, 2\pi]$$

- 确定 p_j 和 q_{j-1} 。

$$p'_j = (E'_j, \vec{p}'_j) \quad \quad q'_{j-1} = q'_j - p'_j$$

其中

$$\begin{cases} \vec{p}'_j = |\vec{p}'_j| (\sin \theta'_j \cos \phi'_j, \sin \theta'_j \sin \phi'_j, \cos \theta'_j) & |\vec{p}'_j| = \frac{1}{2M_j} \sqrt{\lambda(M_j^2, M_{j-1}^2, m_j^2)} \\ E'_j = \frac{1}{2M_j} (M_j^2 + m_j^2 - M_{j-1}^2) \end{cases}$$

$$(\text{逆洛伦兹变换}) \quad \Lambda_j^{-1} : \quad \begin{cases} q'_{j-1} & \longmapsto & q_{j-1} \\ p'_j & \longmapsto & p_j \end{cases}$$

第四章 微分方程数值求解方法

4.1 常微分方程

高阶常微分方程组通过降阶可以转化为一阶常微分方程组。因此，我们首先介绍一阶常微分方程（组）的各种数值求解方法，并对比讨论各种方法的截断误差。

4.1.1 一阶常微分方程求解

考虑一阶常微分方程初值问题：

$$\begin{cases} \dot{x} = f(t, x) \\ x(0) = x_0 \end{cases}$$

将时间离散化，选择步长为 τ ，记

$$t_{i+p} = (i+p)\tau, \quad (i = 0, 1, 2, \dots \quad 0 \leq p \leq 1)$$

命题 4.1

时间格点上函数值 $x(t_i)$ 的递推公式：

$$x(t_{i+1}) = x(t_i) + K\tau$$

其中

$$K \equiv \frac{x(t_{i+1}) - x(t_i)}{\tau} \quad (\text{割线斜率}) = \dot{x}(\xi) \quad (\text{切线斜率}) \quad \xi \in [t_i, t_{i+1}]$$



$$K_{i+p} = \dot{x}(t_{i+p}) = f(t_{i+p}, x_{i+p})$$

基于数值积分

(1) 矩形近似

$$x_{i+p} = \bar{x}_{i+p} \equiv x_i + p\tau K_i, \quad K_i = f(t_i, x_i) \quad K \simeq K_i$$

截断误差： $\mathcal{O}(\tau^2)$

(2) 梯形近似

$$x_{i+1} = x_i + \frac{\tau}{2} [f(t_i, x_i) + f(t_{i+1}, x_{i+1})] \quad K \simeq \frac{1}{2} (K_i + K_{i+1})$$

↓
(迭代法)

$$x_{i+1}^{(k)} = x_i + \frac{\tau}{2} [f(t_i, x_i) + f(t_{i+1}, x_{i+1}^{(k-1)})], \quad x_{i+1}^{(0)} = \bar{x}_{i+1}$$

$k = 1$:

$$x_{i+1}^{(1)} = x_i + \frac{\tau}{2} (K_i + K_{i+1}), \quad \begin{cases} K_i = f(t_i, x_i) \\ K_{i+1} = f(t_{i+1}, \bar{x}_{i+1}) \end{cases}$$

截断误差： $\mathcal{O}(\tau^3)$

(3) 中点近似

$$x_{i+1} = x_i + \tau f(t_{i+1/2}, x_{i+1/2}) \quad K \simeq K_{i+1/2}$$

↓ (迭代法)

$$x_{i+1}^{(k)} = x_i + \tau f(t_{i+1/2}, x_{i+1/2}^{(k-1)}), \quad x_{i+1/2}^{(0)} = \bar{x}_{i+1/2}$$

 $k = 1:$

$$x_{i+1}^{(1)} = \boxed{x_i + \tau K_{i+1/2}, \quad K_{i+1/2} = f(t_{i+1/2}, \bar{x}_{i+1/2})}$$

截断误差: $\mathcal{O}(\tau^3)$ **定义 4.1**

(Runge-Kutta 方法)

$$x_{i+1} = x_i + \tau \sum_{m=1}^N \lambda_m K_m$$

$$K_m = f(t_i + p_m \tau, x_i + q_m \tau), \quad p_1 = q_1 = 0$$

$$q_m = \sum_{n=1}^{m-1} \beta_{mn} K_n \quad (m = 2, 3, \dots, N)$$

**例 4.1** 二阶 Runge-Kutta 方法。

解

递推公式: $x_{i+1} = x_i + \tau (\lambda_1 K_1 + \lambda_2 K_2)$ 选择: $K_2 = f(t_{i+p}, \bar{x}_{i+p})$ λ_1, λ_2, p 由截断误差优化确定。

$$\begin{aligned} K_2 &= f(t_i + p\tau, x_i + p\tau K_1) \\ &= \left[1 + p\tau \left(\frac{\partial}{\partial t} + K_1 \frac{\partial}{\partial x} \right) \right] f(t_i, x_i) + \mathcal{O}(\tau^2) \\ K_1 &= f(t_i, x_i) = \dot{x}(t_i) \\ &= \left[1 + p\tau \frac{1}{dt} \right] \dot{x}(t_i) + \mathcal{O}(\tau^2) \\ &= \dot{x}(t_i) + p\tau \ddot{x}(t_i) + \mathcal{O}(\tau^2) \end{aligned}$$

递推结果为

$$\begin{aligned} x_{i+1} &= x_i + \tau (\lambda_1 K_1 + \lambda_2 K_2) \\ &= x(t_i) + (\lambda_1 + \lambda_2) \tau \dot{x}(t_i) + p\lambda_2 \tau^2 \ddot{x}(t_i) + \mathcal{O}(\tau^3) \end{aligned}$$

截断误差 $\mathcal{O}(\tau^3)$, 应取

$$\begin{cases} \lambda_1 + \lambda_2 = 1 \\ \lambda_2 p = \frac{1}{2} \end{cases}$$

两个常用的特解为

$$\begin{aligned} p = 1, \quad \lambda_1 = \lambda_2 = \frac{1}{2} & \quad (\text{梯形近似}) \\ p = \frac{1}{2}, \quad \lambda_1 = 0, \quad \lambda_2 = 1 & \quad (\text{中点近似}) \end{aligned}$$

例 4.2 三阶 Runge-Kutta 方法。

解

递推公式: $x_{i+1} = x_i + \tau(\lambda_1 K_1 + \lambda_2 K_2 + \lambda_3 K_3)$

$$K_2 = f(t_{i+p}, \bar{x}_{i+p}), \quad \bar{x}_{i+p} = x_i + \tau p K_1 \quad (\text{Euler 法})$$

选择:

$$K_3 = f(t_{i+q}, \tilde{x}_{i+q}), \quad \tilde{x}_{i+q} = x_i + \tau q(\rho_1 K_1 + \rho_2 K_2)$$

(二阶 Runge-Kutta 法)

可以证明, 截断误差 $\mathcal{O}(\tau^4)$, 应取

$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 1 \\ \rho_1 + \rho_2 = 1 \\ \lambda_2 p + \lambda_3 q = \frac{1}{2} \\ \lambda_2 p^2 + \lambda_3 q^2 = \frac{1}{3} \\ \lambda_3 \rho_2 p q = \frac{1}{6} \end{cases}$$

常用的特解为

$$p = \frac{1}{2}, \quad q = 1, \quad \lambda_1 = \frac{1}{6}, \quad \lambda_2 = \frac{2}{3}, \quad \lambda_3 = \frac{1}{6}, \quad \rho_1 = -1, \quad \rho_2 = 2$$



笔记 常用龙格-库塔算法:

命题 4.2. (二阶 Runge-Kutta 算法)

(梯形近似)

$$\begin{cases} x_{i+1} = x_i + \frac{\tau}{2}(K_1 + K_2) \\ K_1 = f(t_i, x_i) \\ K_2 = f(t_i + \tau, x_i + \tau K_1) \end{cases}$$

(中点近似)

$$\begin{cases} x_{i+1} = x_i + \tau K_2 \\ K_1 = f(t_i, x_i) \\ K_2 = f(t_i + \frac{\tau}{2}, x_i + \frac{\tau}{2} K_1) \end{cases}$$

截断误差: $\mathcal{O}(\tau^3)$



命题 4.3. (三阶 Runge-Kutta 算法)

(三点近似)

$$\begin{cases} x_{i+1} = x_i + \frac{\tau}{6}(K_1 + 4K_2 + K_3) \\ K_1 = f(t_i, x_i) \\ K_2 = f(t_i + \frac{\tau}{2}, x_i + \frac{\tau}{2} K_1) \\ K_3 = f(t_i + \tau, x_i + \tau(-K_1 + 2K_2)) \end{cases}$$

截断误差: $\mathcal{O}(\tau^4)$



命题 4.4. (四阶 Runge-Kutta 算法)

(改进三点近似)

$$\begin{aligned}
 x_{i+1} &= x_i + \frac{\tau}{6} (K_1 + 2K_2 + 2K_3 + K_4) \\
 \begin{cases} K_1 &= f(t_i, x_i) \\ K_2 &= f(t_i + \frac{\tau}{2}, x_i + \frac{\tau}{2}K_1) \\ K_3 &= f(t_i + \frac{\tau}{2}, x_i + \frac{\tau}{2}K_2) \\ K_4 &= f(t_i + \tau, x_i + \tau K_3) \end{cases}
 \end{aligned}$$

截断误差: $\mathcal{O}(\tau^5)$ 

上述一阶常微分方程数值求解方法同样适用于一阶常微分方程组, 这里不再赘述。

4.1.2 一类特殊二阶常微分方程求解

二阶线性常微分方程初值问题:

$$\begin{cases} \ddot{x} = \boxed{H(t)} \dot{x} + F(t)x + G(t) \\ x(t_0) = x_0 \\ \dot{x}(t_0) = v_0 \end{cases} \quad (\text{不含一阶微商})$$

数值求解方法: Numerov 算法。

注 对于任意高阶常微分方程, 可以采用“降阶 + Runge-Kutta 算法”数值求解。

解

由 $x_{i\pm 1}$ 精确至 $\mathcal{O}(\tau^5)$ 的泰勒展开式, 知

$$x_{i+1} + x_{i-1} - 2x_i = \tau^2 \left(x_i^{(2)} + \frac{1}{12} \tau^2 x_i^{(4)} \right) + \mathcal{O}(\tau^6)$$

对于任意函数 $x(t)$, 定义

$$\Delta^2 x_i = x_{i+1} + x_{i-1} - 2x_i$$

则

$$\left. \begin{aligned} \Delta^2 x_i &= \tau^2 \left(x_i^{(2)} + \frac{1}{12} \tau^2 x_i^{(4)} \right) + \mathcal{O}(\tau^6) \\ \downarrow \\ \Delta^2 x_i^{(2)} &= \tau^2 \left(x_i^{(4)} + \frac{1}{12} \tau^2 x_i^{(6)} \right) + \mathcal{O}(\tau^6) \end{aligned} \right\} \longrightarrow \Delta^2 \left(x_i - \frac{1}{12} \tau^2 x_i^{(2)} \right) = \tau^2 x_i^{(2)} + \mathcal{O}(\tau^6)$$

$$\text{其中: } x_i^{(2)} = F_i x_i + G_i$$

因此, 有

$$\Delta^2 \left(x_i - \frac{1}{12} \tau^2 F_i x_i \right) = \tau^2 \left(F_i x_i + G_i + \frac{1}{12} \Delta^2 G_i \right) + \mathcal{O}(\tau^6)$$

至此, 我们得到不含一阶微商的二阶线性常微分方程的 Numerov 算法。

命题 4.5. (Numerov 算法)

定义: $y = \left(1 - \frac{1}{12}\tau^2 F\right)x$

$$y_{i+1} - 2y_i + y_{i-1} = \tau^2 \left[\frac{F_i}{1 - \frac{1}{12}\tau^2 F_i} y_i + G_i + \frac{1}{12}(G_{i+1} - 2G_i + G_{i-1}) \right]$$

(y_0, y_1 由初始条件给出)

截断误差: $\mathcal{O}(\tau^6)$



4.2 偏微分方程有限差分法

差分法的基本思想是将微分用差分、微商用差商来代替,从而将微分方程转化为代数方程。

4.2.1 差分和差商

以一元函数 $f(x)$ 为例,步长 h 的一阶、二阶差分定义如下:

定义 4.2

$$\begin{aligned} \text{一阶差分: } \Delta f(x) &= \begin{cases} f(x+h/2) - f(x-h/2) & (\text{中心}) \\ f(x+h) - f(x) & (\text{向前}) \\ f(x) - f(x-h) & (\text{向后}) \end{cases} \\ \text{二阶差分: } \Delta^2 f(x) &= f(x+h) - 2f(x) + f(x-h) \quad (\text{中心}) \end{aligned}$$



由泰勒展开式,知

命题 4.6

$$\begin{aligned} \text{一阶差商: } \frac{\Delta f(x)}{h} - f'(x) &= \begin{cases} \mathcal{O}(h^2) & (\text{中心}) \\ \mathcal{O}(h) & (\text{向前/向后}) \end{cases} \\ \text{二阶差商: } \frac{\Delta^2 f(x)}{h^2} - f''(x) &= \mathcal{O}(h^2) \quad (\text{中心}) \end{aligned}$$



注 对于中心差分、差商格式,有

$$\Delta^2 = \Delta\Delta, \quad \frac{\Delta^2}{h^2} = \frac{\Delta}{h} \frac{\Delta}{h}$$

证明

$$\begin{aligned} \Delta\Delta f(x) &= \Delta f(x+h/2) - \Delta f(x-h/2) \\ &= [f(x+h) - f(x)] - [f(x) - f(x-h)] \\ &= \Delta^2 f(x) \end{aligned}$$

物理上，基本的动力学方程是二阶偏微分方程；因此，我们仅引入一阶和二阶差分格式。本节将以例分别讨论如下三类简单的二阶偏微分方程：

- 抛物型：

$$\left[\frac{\partial}{\partial t} - k \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \right] \varphi(x, y, z, t) = q(x, y, z, t)$$

- 双曲型：

$$\left[\frac{\partial^2}{\partial t^2} - c^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \right] \varphi(x, y, z, t) = f(x, y, z, t)$$

- 椭圆型：

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \varphi(x, y, z) = -\rho(x, y, z)$$

上述微分方程求解需要附加初始条件和边界条件。抛物型方程需要一个初始条件：初始时刻场；双曲型方程需要两个初始条件：初始时刻场及其时间变化率；椭圆型方程为定态方程，无需初始条件。边界条件有以下三种类型：

- 第一类 (Dirichlet)：

$$\varphi(x, y, z, t) \Big|_{\partial V} = g(x, y, z, t)$$

- 第二类 (Neumann)：

$$\frac{\partial}{\partial \hat{n}} \varphi(x, y, z, t) \Big|_{\partial V} = g(x, y, z, t)$$

- 第三类：

$$\left(\alpha \frac{\partial}{\partial \hat{n}} + \beta \right) \varphi(x, y, z, t) \Big|_{\partial V} = g(x, y, z, t)$$

4.2.2 一维无源热传导方程

$q = 0$ (无热源)：

$$\left(\frac{\partial}{\partial t} - k \frac{\partial^2}{\partial x^2} \right) \varphi(x, t) = 0 \quad (0 \leq x \leq L, t \geq 0)$$

离散化：

$$\varphi_{i,j} = \varphi(x_i, t_j), \quad \begin{cases} x_i = ih & (i = 0, 1, 2, \dots, N) \\ t_j = j\tau & (j = 0, 1, 2, \dots) \end{cases} \quad h = \frac{L}{N}$$

差分方程：

$$\left(\frac{\Delta t}{\tau} - k \frac{\Delta_x^2}{h^2} \right) \varphi_{i,j} = 0 \quad \longrightarrow \quad \Delta t \varphi_{i,j} = \boxed{\frac{\tau}{h^2} k} \Delta_x^2 \varphi_{i,j} \quad \equiv \lambda$$

其中，取定

$$\Delta_x^2 \varphi_{i,j} = \varphi_{i-1,j} - 2\varphi_{i,j} + \varphi_{i+1,j}$$

(1) 显式格式

$$\varphi_{i,j+1} - \varphi_{i,j} = \lambda \Delta_x^2 \varphi_{i,j} \quad \longrightarrow \quad \varphi_{i,j+1} = (1 + \lambda \Delta_x^2) \varphi_{i,j}$$

(2) 隐式格式

$$\varphi_{i,j} - \varphi_{i,j-1} = \lambda \Delta_x^2 \varphi_{i,j} \quad \longrightarrow \quad (1 - \lambda \Delta_x^2) \varphi_{i,j+1} = \varphi_{i,j}$$

(3) Crank-Nicolson 格式

$$(1 - \lambda \Delta_x^2/2) \varphi_{i,j+1} = (1 + \lambda \Delta_x^2/2) \varphi_{i,j}$$

命题 4.7. (一维无源热传导方程差分格式)

(显式格式)

$$\varphi_{i,j+1} = (1 + \lambda \Delta_x^2) \varphi_{i,j}$$

(隐式格式)

$$\varphi_{i,j+1} = \frac{1}{1 - \lambda \Delta_x^2} \varphi_{i,j}$$

(Crank-Nicolson 格式)

$$\varphi_{i,j+1} = \frac{1 + \frac{1}{2} \lambda \Delta_x^2}{1 - \frac{1}{2} \lambda \Delta_x^2} \varphi_{i,j}$$

其中

$$\lambda = \frac{\tau}{h^2} k,$$

$$\Delta_x^2 \varphi_{i,j} = \varphi_{i-1,j} - 2\varphi_{i,j} + \varphi_{i+1,j}$$

注 显式和隐式格式分别对应于 Δ_t 的向前及向后差分格式；Crank-Nicolson 六点格式则为显式格式和隐式格式的算术平均。上述求解格式为时间扫描递推公式， $t = 0$ 时刻由初始条件给出。显式格式中，稳定性条件要求 $\lambda \in [0, 1/2]$ ；隐式格式及 Crank-Nicolson 格式中，每一扫描时刻均面临代数线性方程组求解。

4.2.3 一维无源波动方程

 $f = 0$ (无振源):

$$\left(\frac{\partial^2}{\partial t^2} - c^2 \frac{\partial^2}{\partial x^2} \right) \varphi(x, t) = 0 \quad (0 \leq x \leq L, t \geq 0)$$

(离散化)

$$\left(\frac{\Delta_t^2}{\tau^2} - c^2 \frac{\Delta_x^2}{h^2} \right) \varphi_{i,j} = 0$$

取二阶微商的中心差商格式，得

命题 4.8. (一维无源波动方程差分格式)

(显式格式)

$$\varphi_{i,j+1} = \lambda (\varphi_{i-1,j} + \varphi_{i+1,j}) + 2(1 - \lambda) \varphi_{i,j} - \varphi_{i,j-1}$$

$$\lambda = \frac{\tau^2}{h^2} c^2$$

稳定性条件: $\lambda < 1$

注 与热传导方程不同，波动方程需要两个初始条件。初始条件确定 $\varphi_{i,0}$ 和 $\varphi_{i,1}$ ，进而由上面递推公式得到 $\varphi_{i,j}$ ($j \geq 2$)。

$$\text{初始条件: } \begin{cases} \varphi(x, 0) = u(x) \\ \frac{\partial \varphi}{\partial t}(x, 0) = v(x) \end{cases} \xrightarrow{\text{差分近似}} \begin{cases} \varphi_{i,0} = u_i \\ \frac{\Delta_t}{\tau} \varphi_{i,0} = v_i \end{cases}$$

 Δ_t 的不同差分格式给出不同的初值 $\varphi_{i,1}$ 。

- 向前差分格式

$$\Delta_t \varphi_{i,0} = \varphi_{i,1} - \varphi_{i,0} \quad \therefore \quad \varphi_{i,1} = u_i + v_i \tau$$

- 中心差分格式

$$\Delta_t \varphi_{i,0} = \frac{\varphi_{i,1} - \varphi_{i,-1}}{2}$$

联立

$$\begin{cases} \frac{\varphi_{i,1} - \varphi_{i,-1}}{2} = v_i \tau & (\text{差分格式}) \\ \varphi_{i,1} = \lambda(\varphi_{i-1,0} + \varphi_{i+1,0}) + 2(1-\lambda)\varphi_{i,0} - \varphi_{i,-1} & (\text{递推公式}) \end{cases}$$

$$\downarrow \quad \text{消元: } \varphi_{i,-1}$$

$$\therefore \varphi_{i,1} = \frac{\lambda}{2}(u_{i-1} + u_{i+1}) + (1-\lambda)u_i + v_i \tau$$

因此, 有

$$\varphi_{i,0} = u_i, \quad \varphi_{i,1} = \left[u_i \text{ or } \frac{\lambda}{2}(u_{i-1} + u_{i+1}) + (1-\lambda)u_i \right] + v_i \tau$$

4.2.4 泊松方程

	时间维数	空间维数	自变量数
热传导方程	1	n	$1+n$
波动方程	1	n	$1+n$
泊松方程	0	n	n

表 4.1: 常见偏微分方程比较。

注 热传导和波动问题为含时问题, 而泊松方程描述的是静态问题。表 4.1 表明: 热传导方程和波动方程必为偏微分方程; 而泊松偏微分方程则要求 $n \geq 2$ 。

二维泊松方程:

$$\nabla^2 \varphi(x, y) = -\rho(x, y) \quad (x, y) \in \mathcal{D}$$

$n \geq 2$ 时, 边界条件是非平庸的。假定

$$\varphi(x, y) \Big|_{\partial \mathcal{D}} = g(x, y) \quad (\text{边界条件}) \quad \mathcal{D} = [0, L]^2_{(\text{区域})}$$

离散化:

$$\varphi_{i,j} = \varphi(x_i, y_j) = \varphi(ih_x, jh_y), \quad h_x = h_y = h = \frac{L}{N}, \quad (i, j = 0, 1, \dots, N)$$

取

$$\nabla^2 \varphi_{i,j} = \frac{1}{h^2} (\varphi_{i+1,j} + \varphi_{i,j+1} + \varphi_{i-1,j} + \varphi_{i,j-1} - 4\varphi_{i,j}) \quad (\text{Laplace 算子})$$

则有

$$\begin{cases} \nabla^2 \varphi_{i,j} = -\rho_{i,j} & (i, j = 1, 2, \dots, N-1) \\ \varphi_{i,j} = g_{i,j} & (i, j = 0, N) \end{cases}$$

这是 $(N-1)^2$ 个未知数 $\varphi_{i,j}$ ($i, j = 1, 2, \dots, N-1$) 的代数线性方程组。

(1) Monte Carlo 方法

0 为区域 \mathcal{D} 内正则节点 (相邻节点都在区域 \mathcal{D} 内或边界上), 相邻 (右、上、左、下) 四个节点记为 1, 2, 3, 4。

$$\varphi_0 = \frac{1}{4}(\varphi_1 + \varphi_2 + \varphi_3 + \varphi_4) + \frac{h^2}{4}\rho_0$$

↓ Monte Carlo 估计值

$$= \begin{cases} \varphi_1 & \xi \leq \frac{1}{4} \\ \varphi_2 & \frac{1}{4} < \xi \leq \frac{1}{2} \\ \varphi_3 & \frac{1}{2} < \xi \leq \frac{3}{4} \\ \varphi_4 & \xi > \frac{3}{4} \end{cases} \quad \xi \in [0, 1] \text{ (均匀分布)}$$

因此, 有

$$\varphi_0 = \varphi_{0+\Delta} + \frac{h^2}{4}\rho_0 \quad \text{游走方向: } \Delta \in \{\rightarrow, \uparrow, \leftarrow, \downarrow\} \text{ (等概率离散型随机变量)}$$

↓

$$0 + \Delta = 1, 2, 3, 4$$

继续游走, 直至边界节点, 得到 φ_0 的一个 Monte Carlo 估计值。重复上述随机游走过程, 即可得到 φ_0 的估计值序列:

$$\{\bar{\varphi}_0^{(m)} \mid m = 1, 2, \dots, M\}$$

其中

$$\bar{\varphi}_0^{(m)} = g_{f^{(m)}} + \frac{h^2}{4} \sum_{v \in m\text{-path}} \rho_v, \quad \begin{cases} m\text{-path}: \text{第 } m \text{ 条随机游走路径} \\ v \in m\text{-path}: m\text{-path 的正则节点} \\ f^{(m)}: m\text{-path 的终点 (边界节点)} \end{cases}$$

为减小 Monte Carlo 方差, 取

$$\bar{\varphi}_0 = \frac{1}{M} \sum_{m=1}^M \bar{\varphi}_0^{(m)}$$

(2) 迭代法

迭代法是代数线性方程组数值求解的普适方法。

• 同时位移迭代法

$$\varphi_{i,j}^{(k+1)} = \frac{1}{4} \left(\varphi_{i+1,j}^{(k)} + \varphi_{i,j+1}^{(k)} + \varphi_{i-1,j}^{(k)} + \varphi_{i,j-1}^{(k)} + h^2 \rho_{i,j} \right)$$

(缺点: 收敛速度极慢; 占用内存单元多)

• 高斯-塞德尔 (Gauss-Seidal) 迭代法

$$\varphi_{i,j}^{(k+1)} = \frac{1}{4} \left(\varphi_{i+1,j}^{(k)} + \varphi_{i,j+1}^{(k)} + \varphi_{i-1,j}^{(k+1)} + \varphi_{i,j-1}^{(k+1)} + h^2 \rho_{i,j} \right)$$

(扫描顺序: 沿 x 和 y 正方向)

• 超松驰 (Over-Relaxation) 迭代法

$$\varphi_{i,j}^{(k+1)} = (1 - \omega) \varphi_{i,j}^{(k)} + \omega \varphi_{i,j}^{(k+1)}$$

($1 < \omega < 2$: 超松驰迭代)

迭代法的初次启动值 $\varphi_{i,j}^{(k=0)}$ 可以取为边界节点处数值的平均。



笔记

二维拉普拉斯算子五点差商格式:

$$0 = (x, y), \quad 1 = (x + h_1, y), \quad 2 = (x, y + h_2), \quad 3 = (x - h_3, y), \quad 4 = (x, y - h_4)$$

有

$$\nabla^2 \varphi_0 = 2 \left[\frac{h_3 (\varphi_1 - \varphi_0) + h_1 (\varphi_3 - \varphi_0)}{h_1 h_3 (h_1 + h_3)} + \frac{h_4 (\varphi_2 - \varphi_0) + h_2 (\varphi_4 - \varphi_0)}{h_2 h_4 (h_2 + h_4)} \right]$$

↓

$$h_1 = h_3 = h_x, \quad h_2 = h_4 = h_y$$

$$\nabla^2 \varphi_0 = \frac{\varphi_1 + \varphi_3 - 2\varphi_0}{h_x^2} + \frac{\varphi_2 + \varphi_4 - 2\varphi_0}{h_y^2}$$

↓

$$h_x = h_y = h$$

$$\nabla^2 \varphi_0 = \frac{1}{h^2} (\varphi_1 + \varphi_2 + \varphi_3 + \varphi_4 - 4\varphi_0)$$

一阶微商差商格式:

$$f'(x) = \frac{1}{h_1 + h_3} \left[\frac{h_3}{h_1} f(x + h_1) - \frac{h_1}{h_3} f(x - h_3) \right] + \frac{h_1 - h_3}{h_1 h_3} f(x)$$

↓

$$h_1 = h_3 = h$$

$$f'(x) = \frac{1}{2h} [f(x + h) - f(x - h)]$$

当区域 \mathcal{D} 和边界条件较为复杂时, 会用到上述非等步长差分格式。这里将不再展开讨论。

4.3 偏微分方程有限元素法

有限元素法是把变分法与有限差分法结合起来的一种方法。理论上, 它以变分原理为基础, 将偏微分方程边值问题转化为泛函极值问题; 计算方法上, 它利用了差分法的离散化思想。

二维区域 \mathcal{D} 上、电荷密度为 $\rho(x, y)$ 的齐次第一类边界泊松方程可以归结为如下作用量泛函的极值问题:

$$A[\varphi] = \int_{\mathcal{D}} dx dy \left[\frac{1}{2} (\nabla \varphi)^2 - \rho \varphi \right]$$

(1) 单元划分

将求解区域 \mathcal{D} 划分为有限个互不重迭的三角形基本单元, 三角形的顶点称为节点。相邻两个三角形元素有两个公共顶点; 边界三角形元素有两个顶点落于边界。

(2) 元素内线性插值函数的确定

三角形元素 e 的顶点为 (x_i, y_i) , (x_j, y_j) , (x_m, y_m) , 顶点处的函数值为 $\varphi_i, \varphi_j, \varphi_m$ 。在三角形内作线性插值:

$$\varphi(x, y) = g_1 + g_2x + g_3y, \quad (x, y) \in e$$

其中

$$\begin{cases} g_1 = \frac{1}{2\Delta} (a_i\varphi_i + a_j\varphi_j + a_m\varphi_m) \\ g_2 = \frac{1}{2\Delta} (b_i\varphi_i + b_j\varphi_j + b_m\varphi_m) \\ g_3 = \frac{1}{2\Delta} (c_i\varphi_i + c_j\varphi_j + c_m\varphi_m) \end{cases}$$

$$\Delta = \frac{1}{2} (b_ic_j - b_jc_i) = \frac{1}{2} \begin{vmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_m & y_m \end{vmatrix}, \quad \begin{cases} a_i = x_jy_m - x_my_j \\ b_i = y_j - y_m \\ c_i = x_m - x_j \end{cases} \quad (i, j, m \text{ 轮换})$$

引入三角形型函数

$$N_l(x, y) = \frac{1}{2\Delta} (a_l + b_lx + c_ly), \quad (l = i, j, m)$$

则

$$\varphi(x, y) = \sum_{l=i,j,m} \varphi_l N_l(x, y)$$

(3) 作用量函数

$$A[\varphi] = \sum_e A_e[\varphi] = \sum_e (A_{1e}[\varphi] - A_{2e}[\varphi])$$

其中

$$A_{1e}[\varphi] = \int_e dx dy \frac{1}{2} (\nabla \varphi)^2, \quad A_{2e}[\varphi] = \int_e dx dy \rho \varphi$$

定义

$$\Phi^{(e)} = \begin{pmatrix} \varphi_i \\ \varphi_j \\ \varphi_m \end{pmatrix} \quad N^{(e)} = \begin{pmatrix} N_i \\ N_j \\ N_m \end{pmatrix} \quad B^{(e)} = \frac{1}{2\Delta} \begin{pmatrix} b_i & b_j & b_m \\ c_i & c_j & c_m \end{pmatrix}$$

得

$A_{1e}[\varphi] = \frac{1}{2} \Phi^{(e)T} \left[\int_e dx dy B^{(e)T} B^{(e)} \right] \Phi^{(e)}$ <div style="text-align: center; margin-top: 10px;"> \downarrow defined as $K^{(e)}$ </div>	$A_{2e}[\varphi] = \Phi^{(e)T} \left[\int_e dx dy \rho N^{(e)} \right]$ <div style="text-align: center; margin-top: 10px;"> \downarrow defined as $P^{(e)}$ </div>
---	--

$$A_e[\varphi] = \frac{1}{2} \Phi^{(e)T} K^{(e)} \Phi^{(e)} - \Phi^{(e)T} P^{(e)}$$

定义网格节点列向量 Φ ，则

$$\Phi^{(e)} = R^{(e)} \Phi$$

作用量泛函为

$$A[\varphi] = \frac{1}{2} \Phi^T K \Phi - \Phi^T P$$

其中

$$K = \sum_e R^{(e)T} K^{(e)} R^{(e)}, \quad P = \sum_e R^{(e)T} P^{(e)}$$

(4) 代数线性方程组

$$\frac{\partial A}{\partial \Phi^T} = 0 \quad \longrightarrow \quad K \Phi = P \quad (\text{极值条件})$$

利用迭代法，即可求得此线性方程组的解。