

计算物理 B

期末考试试卷及参考解答

中国科学技术大学

汇编自 USTC 评课社区 (icourse.club) 公开资料

共收录 8 套期末试卷 (2021 秋 – 2025 秋)

2021 年秋	吴雨生班
2022 年秋	吴雨生班
2022 年秋	张志勇班
2023 年秋	吴雨生班
2024 年秋	魏逸丰班
2025 年春	张志勇班
2025 年秋	高阳班
2025 年秋	魏逸丰班

2026 年 9 月 12 日

目录

1	2021 年秋季学期（吴雨生班）	2
2	2022 年秋季学期（吴雨生班）	6
3	2022 年秋季学期（张志勇班）	11
4	2023 年秋季学期（吴雨生班）	14
5	2024 年秋季学期（魏逸丰班）	20
6	2025 年春季学期（张志勇班）	24
7	2025 年秋季学期（高阳班）	28
8	2025 年秋季学期（魏逸丰班）	32

1 2021 年秋季学期（吴雨生班）

试卷

题 1. 蒙特卡洛方法

(1) 给定 Gamma 分布

$$f(x) = \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x}, \quad x \geq 0, \quad n \in \mathbb{N}^+, \quad \lambda > 0,$$

用变换法对其进行抽样并证明。

(2) 给定二维随机变量的联合概率密度 $f(x, y)$ ，写出利用边缘分布与条件分布进行抽样的步骤。

题 2. 蒙特卡洛方法

(1) 给定积分 $I = \int_D g(x) dx$ 以及两个不同的分布密度函数 $p_1(x), p_2(x)$ 。设对 $k = 1, 2$, $p_k(x) > 0$ 覆盖 $g(x)$ 的非零区域, $\int_D p_k(x) dx = 1$ 。用重要抽样法写出两种情况下的蒙特卡洛估计量、方差并比较, 写出步骤。

(2) 解释方差不一样的原因。

题 3. 有限差分法

用五点差分格式解 Laplace 方程 $\nabla^2 u = 0$, 计算三个内点的函数值。取网格

$$x = 0, 2, 4, 6, 8, \quad y = 0, 2, 4,$$

边界条件为

$$u(0, y) = u(x, 0) = u(x, 4) = 0, \quad u(8, y) = \sin \frac{\pi y}{4}.$$

求 $(2, 2), (4, 2), (6, 2)$ 三个内点处的函数值。

题 4. 蒙特卡洛方法

写出减除法进行蒙特卡洛抽样的步骤。设目标密度可写为

$$p(x) = \frac{g(x) - h(x)}{C}, \quad 0 \leq h(x) \leq g(x), \quad C = \int [g(x) - h(x)] dx,$$

并且容易从 $q(x) = g(x)/G$, $G = \int g(x) dx$ 抽样。

题 5. 分子动力学

(1) 用最小像力和周期性边界条件计算给定粒子受到多少个粒子的作用。设二维正方形周期区域边长 $L = 10$, 均匀划分为 10×10 个小格, 每个小格中心放置一个粒子, 目标粒子为右下角粒子, 截断距离 $r_c = 0.35L$ 。

(2) 写出 Lennard-Jones 势的表达式, 说明 ϵ 和 σ 的物理意义, 并说明势能极小值的位置。

(3) 给定 r^n 和 v^n , 写出用 Verlet 算法计算 r^{n+1} 和 v^{n+1} 的步骤。

题 6. 变分法与 Mathematica

用 Mathematica 解球坐标系下的量子力学变分问题。设

$$\hat{H} = -\frac{\nabla^2}{2\mu} + ar^3, \quad \Psi_0(r, \lambda) = \frac{\lambda^{3/2}}{\sqrt{\pi}} e^{-\lambda r}, \quad \Psi_1(r, \lambda) = \frac{\lambda^{3/2}}{\sqrt{3\pi}} (3 - 2\lambda r) e^{-\lambda r}.$$

- (1) 采用自然单位制, 分析题目中三个参数的量纲 (表示为能量量纲 $\dim[E]$ 的幂次)。
- (2) 根据要求用 Mathematica 代码定义 Ψ_0, Ψ_1 及哈密顿量。
- (3) 证明两个 s 态波函数的正交性与归一性。
- (4) 用 Mathematica 表示哈密顿量矩阵, 然后用变分法求基态能量。

题 7. 机器学习 (原卷此题信息缺失, 从略。)

参考解答

题 1.

- (1) 当形状参数 n 为正整数时, Gamma 分布可看作 n 个独立同分布指数随机变量之和。设 $\xi_1, \dots, \xi_n \sim U(0, 1)$ 相互独立, 并令

$$X_i = -\frac{1}{\lambda} \ln \xi_i, \quad X = \sum_{i=1}^n X_i = -\frac{1}{\lambda} \ln \left(\prod_{i=1}^n \xi_i \right).$$

由反函数抽样可知每个 X_i 服从指数分布 $f_1(x) = \lambda e^{-\lambda x}$ ($x \geq 0$)。 X 的矩母函数为

$$M_X(t) = \prod_{i=1}^n \mathbb{E}(e^{tX_i}) = \left(\frac{\lambda}{\lambda - t} \right)^n, \quad t < \lambda,$$

这正是 Gamma 分布 $f(x) = \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x}$ 的矩母函数, 因此上述变换得到目标分布。

- (2) 对二维密度 $f(x, y)$, 先求边缘密度 $f_X(x) = \int f(x, y) dy$ 及边缘分布函数 $F_X(x) = \int_{-\infty}^x f_X(t) dt$ 。产生 $\xi_1 \sim U(0, 1)$, 令 $x = F_X^{-1}(\xi_1)$ 。再由条件密度

$$f_{Y|X}(y | x) = \frac{f(x, y)}{f_X(x)}$$

求条件分布函数 $F_{Y|X}(y | x) = \int_{-\infty}^y f_{Y|X}(s | x) ds$ 。产生 $\xi_2 \sim U(0, 1)$, 令 $y = F_{Y|X}^{-1}(\xi_2 | x)$, 则 (x, y) 服从给定二维联合密度。若反函数不易显式写出, 可在边缘抽样或条件抽样步骤中使用舍选法或数值反演。

题 2.

- (1) 对 $k = 1, 2$, 把积分写成

$$I = \int_D \frac{g(x)}{p_k(x)} p_k(x) dx = \mathbb{E}_{p_k} \left[\frac{g(X)}{p_k(X)} \right].$$

从 p_k 抽取独立样本 $X_1^{(k)}, \dots, X_N^{(k)}$, 重要抽样估计量为

$$\hat{I}_k = \frac{1}{N} \sum_{i=1}^N \frac{g(X_i^{(k)})}{p_k(X_i^{(k)})}.$$

它是无偏估计, 因为 $\mathbb{E}(\hat{I}_k) = I$ 。方差为

$$\text{Var}(\hat{I}_k) = \frac{1}{N} \left[\int_D \frac{g(x)^2}{p_k(x)} dx - I^2 \right].$$

哪个积分 $\int g^2/p_k$ 更小, 哪个重要抽样方案的方差就更小。

- (2) 方差不同的原因是权重 $w_k(x) = g(x)/p_k(x)$ 的波动不同。若 p_k 在 $|g|$ 大的区域取值太小, 则少数样本会产生很大的权重, 导致方差增大; 若 p_k 与 $|g|$ 的形状接近, 权重更平稳, 方差较小。若 $g(x) \geq 0$, 理论最优选择为

$$p^*(x) = \frac{g(x)}{\int_D g(x) dx},$$

此时 $g(x)/p^*(x)$ 为常数, 方差为 0。

题 3. 三个内点记为 $u_1 = u(2, 2)$, $u_2 = u(4, 2)$, $u_3 = u(6, 2)$ 。因为 $\Delta x = \Delta y = 2$, 五点差分格式为

$$u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j} = 0.$$

边界上有 $u(0, y) = u(x, 0) = u(x, 4) = 0$, $u(8, 2) = \sin \frac{\pi}{2} = 1$ 。于是

$$-4u_1 + u_2 = 0, \quad u_1 - 4u_2 + u_3 = 0, \quad u_2 - 4u_3 + 1 = 0,$$

即

$$4u_1 - u_2 = 0, \quad -u_1 + 4u_2 - u_3 = 0, \quad -u_2 + 4u_3 = 1.$$

解得

$$u(2, 2) = \frac{1}{56}, \quad u(4, 2) = \frac{1}{14}, \quad u(6, 2) = \frac{15}{56}.$$

题 4. 减除法可看成用 $g(x)$ 对 $g(x) - h(x)$ 作舍选抽样。步骤如下:

- (1) 从容易抽样的密度 $q(x) = \frac{g(x)}{G}$, $G = \int g(x) dx$ 产生候选样本 X 。
- (2) 产生 $u \sim U(0, 1)$ 。
- (3) 以概率

$$A(X) = \frac{g(X) - h(X)}{g(X)} = 1 - \frac{h(X)}{g(X)}$$

接受 X , 否则拒绝并重新抽样。

- (4) 重复上述步骤, 直到得到足够多的样本。

接受样本的密度正比于

$$q(x)A(x) = \frac{g(x)}{G} \cdot \frac{g(x) - h(x)}{g(x)} = \frac{g(x) - h(x)}{G},$$

归一化后就是 $p(x) = \frac{g(x) - h(x)}{C}$ 。因此减除法得到目标密度。

题 5.

- (1) 正方形边长 $L = 10$, 划分为 10×10 个小格, 格距为 1。截断半径 $r_c = 0.35L = 3.5$ 。在周期性边界和最小像力约定下, 其他粒子相对右下角目标粒子的最小像位移可写成整数对 (p, q) , 并满足

$$p^2 + q^2 < 3.5^2 = 12.25, \quad (p, q) \neq (0, 0).$$

逐项计数:

$$\begin{aligned} p = 0 : q = \pm 1, \pm 2, \pm 3, & \quad 6 \text{ 个}; \\ p = \pm 1 : q = 0, \pm 1, \pm 2, \pm 3, & \quad 14 \text{ 个}; \\ p = \pm 2 : q = 0, \pm 1, \pm 2, & \quad 10 \text{ 个}; \\ p = \pm 3 : q = 0, \pm 1, & \quad 6 \text{ 个}. \end{aligned}$$

因此目标粒子受到 $6 + 14 + 10 + 6 = 36$ 个其他粒子的作用。

(2) Lennard-Jones 势为

$$U(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right].$$

其中 ϵ 表示势阱深度, σ 是势能过零点对应的长度尺度。 $r < \sigma$ 时排斥项占主导, 势能迅速升高; $r > \sigma$ 后吸引项起主要作用并逐渐趋于 0^- 。由 $\frac{dU}{dr} = 0$ 得

$$r_{\min} = 2^{1/6}\sigma, \quad U_{\min} = -\epsilon.$$

粒子 j 对粒子 i 的力可写为

$$\mathbf{F}_{ij} = 24\epsilon \left[\frac{2\sigma^{12}}{r_{ij}^{14}} - \frac{\sigma^6}{r_{ij}^8} \right] \mathbf{r}_{ij}, \quad \mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j.$$

(3) 速度 Verlet 格式为: 先由当前坐标计算力和加速度 $\mathbf{a}_n = \mathbf{F}(\mathbf{r}_n)/m$ 。位置更新为

$$\mathbf{r}_{n+1} = \mathbf{r}_n + \mathbf{v}_n \Delta t + \frac{1}{2} \mathbf{a}_n \Delta t^2.$$

再由新位置计算 $\mathbf{a}_{n+1} = \mathbf{F}(\mathbf{r}_{n+1})/m$ 。最后更新速度

$$\mathbf{v}_{n+1} = \mathbf{v}_n + \frac{1}{2}(\mathbf{a}_n + \mathbf{a}_{n+1})\Delta t.$$

若使用位置 Verlet, 则位置递推也可写为 $\mathbf{r}_{n+1} = 2\mathbf{r}_n - \mathbf{r}_{n-1} + \mathbf{a}_n \Delta t^2$ 。

题 6.

(1) 采用自然单位 $\hbar = 1$ 。指数 λr 必须无量纲, 因此

$$[r] = E^{-1}, \quad [\lambda] = E.$$

动能项 $-\nabla^2/(2\mu)$ 具有能量量纲, 而 $\nabla^2 \sim E^2$, 所以 $[\mu] = E$ 。势能项 ar^3 也具有能量量纲, 因此 $[a]E^{-3} = E$, 即 $[a] = E^4$ 。

(2) Mathematica 中可按如下方式定义波函数、内积和哈密顿量:

```
Clear[r, lambda, mu, a]
$Assumptions = lambda > 0 && mu > 0 && a > 0 && r >= 0;
psi0[r_] := lambda^(3/2)/Sqrt[Pi] Exp[-lambda r];
psi1[r_] := lambda^(3/2)/Sqrt[3 Pi] (3 - 2 lambda r) Exp[-lambda r];
ip[f_, g_] := FullSimplify[Integrate[4 Pi r^2 f[r] g[r], {r, 0, Infinity}]]
lap[f_] := 1/r^2 D[r^2 D[f, r], r];
ham[f_] := -lap[f]/(2 mu) + a r^3 f;
```

(3) 正交归一性由球对称内积

$$\langle f|g\rangle = 4\pi \int_0^\infty r^2 f(r)g(r) dr$$

给出。计算得

$$\langle \psi_0|\psi_0\rangle = 1, \quad \langle \psi_1|\psi_1\rangle = 1, \quad \langle \psi_0|\psi_1\rangle = 0.$$

Mathematica 代码为 `ip[psi0,psi0]`, `ip[psi1,psi1]`, `ip[psi0,psi1]`。

(4) 哈密顿量矩阵元为

$$H_{ij} = 4\pi \int_0^\infty r^2 \psi_i(r) \hat{H} \psi_j(r) dr.$$

对 $\{\psi_0, \psi_1\}$ 得

$$H = \begin{pmatrix} \frac{\lambda^2}{2\mu} + \frac{15a}{2\lambda^3} & \sqrt{3}\left(\frac{\lambda^2}{3\mu} - \frac{15a}{2\lambda^3}\right) \\ \sqrt{3}\left(\frac{\lambda^2}{3\mu} - \frac{15a}{2\lambda^3}\right) & \frac{7\lambda^2}{6\mu} + \frac{75a}{2\lambda^3} \end{pmatrix}.$$

基态能量的二维变分上界为

$$E_0^{\text{var}} = \min_{\lambda>0} \lambda_{\min}(H(\lambda)),$$

其中 $\lambda_{\min}(H(\lambda))$ 表示矩阵的较低本征值。Mathematica 中可用 `Eigenvalues[H]` 与 `NMinimize` 实现。

题 7. 原卷此题信息缺失, 从略。

2 2022 年秋季学期 (吴雨生班)

试卷

题 1. 蒙特卡洛方法

考虑积分

$$I = \int_{-\infty}^{+\infty} \frac{1}{2} x^2 e^{-|x|} dx.$$

(1) 取概率密度 $p(x) = \frac{1}{2} e^{-|x|}$, 用直接抽样法估计 I , 写出抽样步骤和估计量。

(2) 取 $q(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ 为正态重要抽样密度, 写出抽样步骤、估计量, 并说明与直接抽样的差别。

题 2. 蒙特卡洛方法

给定二维分布

$$f(x, y) = 6(x - y), \quad 0 \leq y \leq x \leq 1,$$

用条件抽样法抽取 (x, y) , 画流程图或写步骤。

题 3. 蒙特卡洛方法

对目标分布 $\pi(x)$ 用 Metropolis 方法抽样。写出使用对称试探分布 $T(x' | x) = T(x | x')$ 时的伪代码流程。

题 4. 有限差分法

设 $D = [0, 1] \times [0, 1]$, 将区域分成 16 个格子, 并给定

$$\nabla^2 \phi = 0, \quad \phi|_{\partial D} = g(x, y) = xy.$$

设 $x_i = i/4$, $y_j = j/4$, $\phi_{ij} \approx \phi(x_i, y_j)$, 求内点 ϕ_{11} 、 ϕ_{33} 的函数值。

题 5. 分子动力学

- (1) 在周期性边界条件和最小像力约定下, 求目标分子受到多少个其他分子的作用。设二维正方形区域边长为 $L = 10$, 均匀划分为 10×10 个小格, 每个小格中心有一个粒子, 截断距离 $r_c = 0.35L$, 目标粒子为右下角粒子。
- (2) 对 Lennard-Jones 势作解释, 讨论相互作用随距离变化的情形。
- (3) 对 Lennard-Jones 势求粒子所受力 F_i 。

题 6. 变分法与 Mathematica

考虑三维球对称量子体系, 采用自然单位 $\hbar = 1$ 。哈密顿量为

$$\hat{H} = -\frac{\nabla^2}{2\mu} + ar^3.$$

取两个 s 态试探基函数

$$\psi_0(r, \lambda) = \frac{\lambda^{3/2}}{\sqrt{\pi}} e^{-\lambda r}, \quad \psi_1(r, \lambda) = \frac{\lambda^{3/2}}{\sqrt{3\pi}} (3 - 2\lambda r) e^{-\lambda r}.$$

球对称内积定义为 $\langle f|g \rangle = 4\pi \int_0^\infty r^2 f(r)g(r) dr$ 。用 $\{\psi_0, \psi_1\}$ 张成二维变分子空间, 构造哈密顿量矩阵并对其较低本征值关于 λ 取极小, 从而给出基态能量的变分上界。

- (1) 对 λ, μ, a 作量纲分析。
- (2) 用 Mathematica 代码证明 $\langle \psi_0|\psi_1 \rangle = 0$, 并验证正交归一。
- (3) 求哈密顿量矩阵 $H_{ij} = \langle \psi_i|\hat{H}|\psi_j \rangle$, 并给出用 Mathematica 求基态能量上界的代码。

题 7. 机器学习

- (1) (原卷信息缺失, 从略。)
- (2) 解释 ROC 曲线和 AUC 的含义。
- (3) 对如下训练数据画出决策树, 划分数据并解释:

x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

参考解答

题 1.

(1) 令 $p(x) = \frac{1}{2}e^{-|x|}$, $-\infty < x < +\infty$, 则原积分就是

$$I = \int_{-\infty}^{+\infty} x^2 p(x) dx = \mathbb{E}_p[X^2].$$

直接抽样可先产生 $\xi_1, \xi_2 \sim U(0, 1)$, 令

$$R = -\ln \xi_1, \quad X = \begin{cases} R, & \xi_2 < 1/2, \\ -R, & \xi_2 \geq 1/2. \end{cases}$$

则 X 服从 $p(x)$ 。蒙特卡洛估计量为 $\hat{I}_N = \frac{1}{N} \sum_{k=1}^N X_k^2$ 。解析值为

$$I = 2 \int_0^{\infty} \frac{1}{2} x^2 e^{-x} dx = \int_0^{\infty} x^2 e^{-x} dx = 2.$$

(2) 若取标准正态密度 $q(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, 可用 Box-Muller 方法抽样:

$$X = \sqrt{-2 \ln \xi_1} \cos(2\pi \xi_2) \sim N(0, 1).$$

对每个 X_k 计算权重

$$w(X_k) = \frac{\frac{1}{2} X_k^2 e^{-|X_k|}}{q(X_k)} = \frac{\sqrt{2\pi}}{2} X_k^2 e^{X_k^2/2 - |X_k|},$$

重要抽样估计量为

$$\hat{I}_N = \frac{1}{N} \sum_{k=1}^N \frac{\sqrt{2\pi}}{2} X_k^2 e^{X_k^2/2 - |X_k|}, \quad X_k \sim N(0, 1).$$

这种写法形式上无偏, 但正态分布尾部比 $e^{-|x|}$ 衰减更快, 权重在大 $|x|$ 处很大, 方差很差, 甚至会出现二阶矩发散的问题。因此本题直接按 $p(x)$ 抽样更稳定。

题 2. 对 $f(x, y) = 6(x - y)$, $0 \leq y \leq x \leq 1$, 先求边缘密度

$$f_X(x) = \int_0^x 6(x - y) dy = 3x^2, \quad 0 \leq x \leq 1,$$

因而 $F_X(x) = \int_0^x 3t^2 dt = x^3$ 。取 $\xi_1 \sim U(0, 1)$, 令 $x = \xi_1^{1/3}$ 。条件密度为

$$f_{Y|X}(y | x) = \frac{f(x, y)}{f_X(x)} = \frac{2(x - y)}{x^2}, \quad 0 \leq y \leq x,$$

其条件分布函数为

$$F_{Y|X}(y | x) = \int_0^y \frac{2(x - t)}{x^2} dt = 1 - \left(1 - \frac{y}{x}\right)^2.$$

再取 $\xi_2 \sim U(0, 1)$, 令 $\xi_2 = 1 - (1 - y/x)^2$, 得 $y = x(1 - \sqrt{1 - \xi_2})$ 。因此

$$(x, y) = \left(\xi_1^{1/3}, \xi_1^{1/3}(1 - \sqrt{1 - \xi_2})\right)$$

即为所需样本。

题 3. 使用对称试探分布时, Metropolis 抽样伪代码可写为:

给定初始状态 x_0 , 步长 δ , 总步数 N

for $i = 0, 1, \dots, N-1$:

从对称试探分布 $T(x'|x_i)$ 产生候选点 x'

计算接受率 $A = \min(1, \pi(x')/\pi(x_i))$

产生 $u \sim U(0,1)$

if $u < A$:

$x_{i+1} = x'$

else:

$x_{i+1} = x_i$

丢弃热化段后, 将保留的 x_i 作为目标分布 $\pi(x)$ 的样本

关键是接受率只含 $\pi(x')/\pi(x_i)$, 所以归一化常数会相消。

题 4. 网格步长为 $h = \frac{1}{4}$, $x_i = \frac{i}{4}$, $y_j = \frac{j}{4}$ 。五点差分格式为

$$\phi_{i+1,j} + \phi_{i-1,j} + \phi_{i,j+1} + \phi_{i,j-1} - 4\phi_{ij} = 0.$$

因为边界函数为 $g(x,y) = xy$, 而 xy 满足连续 Laplace 方程 $\nabla^2(xy) = 0$, 并且五点差分对双线性函数也给出零离散 Laplace 值, 所以离散解由唯一性确定为

$$\phi_{ij} = x_i y_j = \frac{ij}{16}.$$

因此 $\phi_{11} = \frac{1}{16}$, $\phi_{33} = \frac{9}{16}$ 。

题 5.

- (1) 正方形边长 $L = 10$, 划分为 10×10 个小格, 格距 $a_0 = L/10 = 1$ 。截断距离 $r_c = 0.35L = 3.5$ 。在周期性边界和最小像力约定下, 以右下角粒子为目标, 其他粒子相对于它的最小像位移可写为整数对 (p, q) , 并满足 $p^2 + q^2 < 3.5^2 = 12.25$, $(p, q) \neq (0, 0)$ 。逐项计数:

$$p = 0 : 6; \quad p = \pm 1 : 14; \quad p = \pm 2 : 10; \quad p = \pm 3 : 6.$$

因此相互作用粒子数为 $6 + 14 + 10 + 6 = 36$ 。

- (2) Lennard-Jones 势通常写为 $U(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6]$ 。其中 ϵ 表示势阱深度, σ 是势能为零时的距离尺度。 $r < \sigma$ 时排斥项占优, 势能迅速升高; $r = \sigma$ 时 $U = 0$; 势能在 $r = 2^{1/6}\sigma$ 处取最小值 $U_{\min} = -\epsilon$ 。当 $r > 2^{1/6}\sigma$ 时主要表现为吸引, 远距离时 $U(r) \rightarrow 0^-$ 。

- (3) 设 $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, $r_{ij} = |\mathbf{r}_{ij}|$ 。粒子 j 对粒子 i 的力为

$$\mathbf{F}_{ij} = -\nabla_{\mathbf{r}_i} U(r_{ij}) = 24\epsilon \left[\frac{2\sigma^{12}}{r_{ij}^{14}} - \frac{\sigma^6}{r_{ij}^8} \right] \mathbf{r}_{ij}.$$

因而粒子 i 所受总力为 $\mathbf{F}_i = \sum_{j \neq i} \mathbf{F}_{ij}$, 实际分子动力学计算中只对截断半径内的最小像粒子求和。

题 6.

- (1) 采用自然单位 $\hbar = 1$ 。指数 λr 必须无量纲, 所以 $[\lambda] = L^{-1}$ 。由动能项 $-\nabla^2/(2\mu)$ 具有能量量纲, 且 $\nabla^2 \sim L^{-2}$, 在自然单位下有

$$[\mu] = E, \quad [r] = E^{-1}, \quad [\lambda] = E.$$

势能项 ar^3 也具有能量量纲, 因此 $[a]E^{-3} = E$, 即 $[a] = E^4$ 。

(2) 球对称内积为 $\langle \psi_i | \psi_j \rangle = 4\pi \int_0^\infty r^2 \psi_i(r) \psi_j(r) dr$ 。Mathematica 代码:

```
Clear[r, lambda]
$Assumptions = lambda > 0 && r >= 0;
psi0[r_] := lambda^(3/2)/Sqrt[Pi] Exp[-lambda r];
psi1[r_] := lambda^(3/2)/Sqrt[3 Pi] (3 - 2 lambda r) Exp[-lambda r];
ip[f_, g_] := FullSimplify[Integrate[4 Pi r^2 f[r] g[r], {r, 0, Infinity}]]
ip[psi0, psi0]; ip[psi1, psi1]; ip[psi0, psi1]
```

输出应为 1, 1, 0, 即两个基函数正交归一。

(3) 对球对称函数, $\nabla^2 \psi = \frac{1}{r^2} \frac{d}{dr} (r^2 \frac{d\psi}{dr})$ 。令

$$H_{ij} = 4\pi \int_0^\infty r^2 \psi_i(r) \hat{H} \psi_j(r) dr,$$

可得

$$H = \begin{pmatrix} \frac{\lambda^2}{2\mu} + \frac{15a}{2\lambda^3} & \sqrt{3} \left(\frac{\lambda^2}{3\mu} - \frac{15a}{2\lambda^3} \right) \\ \sqrt{3} \left(\frac{\lambda^2}{3\mu} - \frac{15a}{2\lambda^3} \right) & \frac{7\lambda^2}{6\mu} + \frac{75a}{2\lambda^3} \end{pmatrix}.$$

用 `Eigenvalues[H]` 与 `NMinimize[{Min[evals], lambda > 0}, lambda]` 即可得到基态能量的变分上界

$$E_0^{\text{var}} = \min_{\lambda > 0} \lambda_{\min}(H(\lambda)).$$

题 7.

(1) 原卷信息缺失, 从略。

(2) ROC 曲线描述分类阈值改变时真正率和假正率之间的关系:

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad \text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}.$$

阈值从严格到宽松变化时, 可得到一条 (FPR, TPR) 曲线。AUC 是 ROC 曲线下的面积, 可理解为随机抽取一个正样本和一个负样本时, 分类器把正样本排在负样本前面的概率。AUC 越接近 1, 排序能力越强; AUC 约为 0.5 时接近随机猜测。

(3) 对给定数据, 标签 $y = 1$ 当且仅当 $x_1 = 1, x_2 = 1$, 因此可得到如下决策树:

$$\begin{aligned} x_1 = 0? & \quad \text{是} \rightarrow y = 0 \\ & \quad \text{否} \rightarrow \text{再看 } x_2: \quad x_2 = 0 \rightarrow y = 0, \quad x_2 = 1 \rightarrow y = 1 \end{aligned}$$

划分解释为: 先按 x_1 划分, $x_1 = 0$ 的两个样本均为 $y = 0$, 已经是纯节点; $x_1 = 1$ 的两个样本再按 x_2 划分, $x_2 = 0$ 得到 $y = 0$, $x_2 = 1$ 得到 $y = 1$ 。等价地, 也可以先按 x_2 划分, 结果对称。

3 2022 年秋季学期（张志勇班）

试卷

题 1. 蒙特卡洛方法

用线性同余法产生伪随机数。给定参数 $a = 5$, $c = 1$, $m = 16$, 初始值 $x_0 = 1$ 。写出递推公式, 并计算前两个归一化随机数。

题 2. 蒙特卡洛方法

考虑分布 $f(x) = 3x^2$, $0 \leq x \leq 1$ 。

- (1) 用直接抽样法抽取服从该分布的随机变量, 写出抽样步骤或流程图。
- (2) 用第一类舍选法抽取服从该分布的随机变量, 写出抽样步骤或流程图。
- (3) 比较直接抽样法和第一类舍选法的优势与不足。

题 3. 蒙特卡洛方法

说明为什么 Metropolis 方法可以应用于无法显式归一化的分布密度函数。用 Metropolis 方法抽样自由度为 4 的 χ^2 分布

$$f(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}, \quad x \geq 0,$$

写出抽样步骤或流程图。

题 4. 有限差分法

用有限差分法求解二维 Laplace 方程。给定网格节点 $x = [0, 2, 4, 6, 8]$, $y = [0, 2, 4]$, 形成 5×3 个格点。边界条件为

$$u(0, y) = u(x, 0) = u(x, 4) = 0, \quad u(8, y) = \sin \frac{\pi y}{4}, \quad \Delta u = 0.$$

求三个内部格点 $(2, 2)$, $(4, 2)$, $(6, 2)$ 处的函数值。

题 5. 数值微分方程

比较有限元法和有限差分法的优势与不足。

题 6. 分子动力学

- (1) 周期性边界条件的目的是什么? 最小像力约定的定义是什么?
- (2) 二维 10×10 正方形区域划分为 100 个小格, 每个小格中心放置一个粒子。采用周期性边界条件和最小像力约定, 截断条件为 $r < r_c = 3$ 。求与右下角粒子发生相互作用的粒子数。
- (3) 写出 Lennard-Jones 势的公式, 说明参数的物理意义, 并指出该势能在何处取到负的极小值。
- (4) 写出微正则系综的定义以及速度 Verlet 算法的基本步骤。

题 7. Mathematica

用 Mathematica 计算带线性阻尼的自由落体运动。设竖直向上为正方向, 受力为 $F = -mg - bv$ 。写出对应微分方程, 并给出 Mathematica 求解代码。

题 8. 机器学习

(1) 判断下列说法是否正确, 并简要说明理由。

- (a) 聚类是监督学习。
- (b) 批量梯度下降适用于特征数 n 较大的情况。
- (c) 袋装算法需要较大数据量。
- (d) 减小正则化系数可以防止过拟合。
- (e) 决策树的关键是如何分割。

(2) 写出 BP 神经网络训练的基本流程。

(3) 说明 PCA 的主要思想; 若给定二维散点图, 应如何画出两个 PCA 方向?

题 9. 并行计算

说明单程序多数据 (SPMD) 和多程序多数据 (MPMD) 分别是什么。

参考解答

题 1. 线性同余法递推公式为 $x_{n+1} = (ax_n + c) \bmod m$ 。代入 $a = 5, c = 1, m = 16, x_0 = 1$, 得

$$x_1 = (5 \cdot 1 + 1) \bmod 16 = 6, \quad x_2 = (5 \cdot 6 + 1) \bmod 16 = 15.$$

因而前两个归一化随机数为

$$\xi_1 = \frac{x_1}{16} = \frac{3}{8}, \quad \xi_2 = \frac{x_2}{16} = \frac{15}{16}.$$

题 2.

(1) 分布函数为 $F(x) = \int_0^x 3t^2 dt = x^3$ 。令 $F(x) = \xi$, 得到直接抽样公式

$$x = \xi^{1/3}, \quad \xi \sim U(0, 1).$$

(2) 第一类舍选法可在矩形 $0 \leq x \leq 1, 0 \leq y \leq 3$ 内均匀取点。令 $x = \xi_1, y = 3\xi_2$, 若 $y \leq 3x^2$ 则接受 x , 否则拒绝并重新抽样。接受率为

$$\eta = \frac{\int_0^1 3x^2 dx}{3} = \frac{1}{3}.$$

(3) 直接抽样法效率高, 没有拒绝样本, 但要求分布函数可积分且反函数容易求出。舍选法更通用, 不要求显式反函数, 但存在拒绝样本, 效率取决于包络区域与目标分布的贴合程度。

题 3. Metropolis 方法只用到目标分布在新旧状态处的比值 $\pi(x')/\pi(x_i)$ 。若 $\pi(x) = C\tilde{\pi}(x)$, 归一化常数 C 在比值中相消, 因此即使 C 不知道, 也可以按未归一化密度 $\tilde{\pi}(x)$ 抽样。

自由度为 4 的 χ^2 分布为 $\pi(x) \propto xe^{-x/2}, x \geq 0$ 。可采用对称随机游走试探: 给定当前状态 x_i , 产生

$$x' = x_i + \delta(2\xi_1 - 1).$$

若 $x' < 0$, 直接拒绝; 若 $x' \geq 0$, 计算接受率

$$A = \min\left\{1, \frac{x' e^{-x'/2}}{x_i e^{-x_i/2}}\right\} = \min\left\{1, \frac{x'}{x_i} e^{-(x'-x_i)/2}\right\}.$$

再产生 $\xi_2 \sim U(0, 1)$ 。若 $\xi_2 < A$, 令 $x_{i+1} = x'$; 否则令 $x_{i+1} = x_i$ 。重复迭代, 丢弃热化段后得到样本。

题 4. 设 $u_1 = u(2, 2), u_2 = u(4, 2), u_3 = u(6, 2)$ 。网格步长为 $\Delta x = \Delta y = 2$ 。五点差分格式为

$$u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j} = 0.$$

由于上下边界均为 0, 左边界为 0, 右边界在 $y = 2$ 处为 $u(8, 2) = \sin \frac{\pi}{2} = 1$, 得到线性方程组

$$4u_1 - u_2 = 0, \quad -u_1 + 4u_2 - u_3 = 0, \quad -u_2 + 4u_3 = 1.$$

解得

$$u(2, 2) = \frac{1}{56}, \quad u(4, 2) = \frac{1}{14}, \quad u(6, 2) = \frac{15}{56}.$$

题 5. 有限差分法直接在规则网格上用差商近似微分算子, 形式简单、实现方便, 适合矩形区域和规则网格; 不足是处理复杂边界、复杂区域和非均匀材料时不够灵活。有限元法先写出弱形式或变分形式, 再选取基函数并组装刚度矩阵, 适合复杂几何、非均匀网格和复杂边界条件; 不足是推导、网格剖分和矩阵组装更复杂, 计算实现成本较高。

题 6.

- (1) 周期性边界条件的目的是用有限大小的模拟盒近似无限体系, 减小表面效应和边界效应。最小像力约定是指计算两个粒子间相互作用时, 在所有周期性镜像中选取距离最近的一对作为实际相互作用距离, 通常要求截断半径不超过盒长的一半。
- (2) 10×10 区域分为 10×10 个小格, 格距为 1。对右下角粒子, 在周期性边界和最小像力约定下, 其他粒子对应整数位移 (p, q) 。截断条件 $r < 3$ 等价于 $p^2 + q^2 < 9$, $(p, q) \neq (0, 0)$ 。逐项计数: $p = 0$ 时 $q = \pm 1, \pm 2$, 共 4 个; $p = \pm 1$ 时 $q = 0, \pm 1, \pm 2$, 共 10 个; $p = \pm 2$ 时 $q = 0, \pm 1, \pm 2$, 共 10 个。因此相互作用粒子数为 $4 + 10 + 10 = 24$ 。若采用 $r \leq r_c$ 的截断约定, 则还包括 $(\pm 3, 0), (0, \pm 3)$, 总数为 28。
- (3) Lennard-Jones 势为 $U(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6]$ 。其中 ϵ 表示势阱深度, σ 是势能过零点对应的长度尺度。势能在 $r = 2^{1/6}\sigma$ 处取最小值 $U_{\min} = -\epsilon$ 。
- (4) 微正则系综为 N, V, E 不变的系综, 即粒子数、体积和总能量保持不变。速度 Verlet 算法步骤为: 由当前位置计算力 $F_i(t)$, 先更新位置

$$r_i(t + \Delta t) = r_i(t) + v_i(t)\Delta t + \frac{F_i(t)}{2m_i}\Delta t^2,$$

再由新位置计算 $F_i(t + \Delta t)$, 最后更新速度

$$v_i(t + \Delta t) = v_i(t) + \frac{F_i(t) + F_i(t + \Delta t)}{2m_i}\Delta t.$$

题 7. 设 $y(t)$ 为竖直坐标, 竖直向上为正, $v(t) = y'(t)$ 。由 $F = -mg - bv$ 得到

$$my''(t) = -mg - by'(t).$$

Mathematica 代码可写为:

```

Clear[y, t, m, g, b, y0, v0]
eqn = m y'[t] == -m g - b y'[t];
ics = {y[0] == y0, y'[0] == v0};
sol = DSolve[{eqn, ics}, y[t], t]
pars = {m -> 1, g -> 9.8, b -> 0.2, y0 -> 10, v0 -> 0};
nsol = NDSolve[Evaluate[{eqn, ics} /. pars], y, {t, 0, 5}]
Plot[Evaluate[y[t] /. nsol], {t, 0, 5}]

```

题 8.

- (1) (a) ×: 聚类通常属于无监督学习。
 - (b) ✓: 批量梯度下降可用于特征数较大的优化问题。
 - (c) ✓: 袋装法通常需要足够数据以产生有差异的 bootstrap 训练集。
 - (d) ×: 防止过拟合通常应增大正则化系数而不是减小。
 - (e) ✓: 决策树训练的核心是选择划分特征和划分点, 使节点不纯度下降。
- (2) BP 训练流程为: 初始化各层权重和偏置; 输入训练样本并进行前向传播, 逐层计算神经元输出; 用损失函数计算预测值与标签的误差; 从输出层向前逐层反向传播误差, 计算损失函数对各参数的梯度; 用梯度下降或其变体更新参数; 重复训练直到损失收敛或达到停止条件。
 - (3) PCA 的主要思想是把数据投影到方差最大的正交方向上, 以较少的新变量保留尽可能多的信息。具体做法是先对数据中心化, 计算协方差矩阵, 再求其特征值和特征向量。最大特征值对应的特征向量为第一主成分方向, 第二大特征值对应且与第一方向正交的特征向量为第二主成分方向。在二维散点图中, 应从数据中心画出沿最大延展方向的第一 PCA 方向, 再画出与其正交的第二 PCA 方向。

题 9. SPMD 是 single program multiple data, 即多个处理单元运行同一个程序, 但处理不同的数据分块; 常见于 MPI 程序中按进程编号区分数据和任务。MPMD 是 multiple program multiple data, 即不同处理单元可以运行不同程序, 并分别处理不同数据; 适合耦合多个功能模块或多物理过程程序的并行计算。

4 2023 年秋季学期 (吴雨生班)

试卷

题 1. 蒙特卡洛方法

考虑在高能物理中描述不稳定粒子的不变质量谱的 Breit-Wigner 分布

$$f(x) = \frac{k}{(x^2 - m^2)^2 + m^2 \Gamma^2}, \quad x \in [m, +\infty),$$

其中 k 为归一化常数。

- (1) 使用第一类舍选法对 $f(x)$ 抽样, 给出抽样效率并探讨如何优化抽样范围。
- (2) 使用第二类舍选法对 $f(x)$ 抽样, $h(x) = \frac{m}{x^2}$, 给出抽样效率并探讨与 (1) 的区别。

题 2. 蒙特卡洛方法

(1) 对 $f(x) = n(1 - e^{-x})^{n-1}e^{-x}$, $n \in \mathbb{N}^+$, $x \in [0, \infty)$, 证明 x 的抽样方法为

$$x = \max\{-\ln \xi_1, -\ln \xi_2, \dots, -\ln \xi_n\}, \quad \xi_i \in [0, 1].$$

(2) 利用 (1) 中结论, 结合叠加原理, 对满足如下分布密度函数的随机变量进行抽样, 画出流程图或写出抽样步骤:

$$f(x) = \frac{3}{2}(e^{-x} - 2e^{-2x} + e^{-3x}) + \frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

题 3. 蒙特卡洛方法

在 2D-Ising Model 中, $H(\vec{s}, \vec{J}) = -\frac{1}{2} \sum_{i \neq j} J_{ij} s_i s_j$ 。系统在温度 T 下为平衡态, 系统态满足 Boltzmann 分布 $\sim \frac{1}{Z} e^{-H(\vec{s}, \vec{J})/(k_B T)}$ 。写出用 Metropolis 方法完成系统态抽样的步骤, 给出系统总磁矩的期望, 探讨其与温度的依赖关系。

题 4. 数值微分方程

(1) 推导二阶 Runge-Kutta 公式。

(2) 用有限元法求解一类 Poisson 方程

$$\nabla^2 \phi = -\frac{\rho}{\epsilon}, \quad \phi|_L = \phi_0,$$

其 $\delta I(\phi) = 0$ 的泛函为

$$I(\phi) = \int_D \left[\frac{\epsilon}{2} (\nabla \phi)^2 - \rho \phi \right] dS.$$

题 5. 分子动力学

(1) 将边长为 L 的正方形区域均匀划分为 100 格子, 每格中央有一粒子, $r_c = 0.35L$ 。使用周期性边界条件和最小像力约定, 请问多少粒子与最右侧粒子有相互作用。

(2) 什么是对关联函数? 写出 Lennard-Jones 势下对关联函数随距离变化的特征。

(3) 写出求解 Verlet 速度形式的基本流程。

题 6. 变分法与 Mathematica

有球对称势能函数

$$\hat{H} = -\frac{\nabla^2}{2\mu} + ar^3,$$

试探基函数为

$$\Psi_0(r, \lambda) = \frac{\lambda^{3/2}}{\sqrt{\pi}} e^{-\lambda r}, \quad \Psi_1(r, \lambda) = \frac{\lambda^{3/2}}{\sqrt{3\pi}} (3 - 2\lambda r) e^{-\lambda r}.$$

(1) 写出 λ 与 a 的量纲。

(2) 写出验证基函数归一化与正交性的 Mathematica 代码。

(3) 写出用变分法计算系统基态能量上限的 Mathematica 代码。

题 7. 机器学习

(1) 以下哪些选项是正确的:

- (a) 无监督学习可以预言数据类别。
 (b) 随机梯度下降法需要更大样本完成训练。
 (c) 数据个数远多于特征数时为稀疏情形。
 (d) 逻辑回归可以通过将 Softmax 函数作用线性回归实现。
 (e) 决策树的末端叶子处 Gini 杂度为 0。
 (f) 神经网络反向传播用来计算代价函数梯度。
- (2) 写出决定系数公式，并说明如何用其判断训练优度。
- (3) 有如下数据集，用其训练一神经网络对数据进行分类。写出具体流程、网络结构、代价函数与传递系数。

数据 X	标签
$\{1, 2, 3\}$	A
$\{2, 3, 5\}$	A
$\{3, 4, 5\}$	A
$\{7, 8, 9\}$	A
$\{-1, 2, 3\}$	B
$\{2, -4, 5\}$	B
$\{3, 7, -9\}$	B
$\{-2, 8, -7\}$	B

参考解答

题 1.

- (1) 第一类舍选法若直接在 $[m, \infty)$ 上取均匀试探，矩形面积无穷大，接受率为 0，因此实际必须先截断到 $[m, b]$ 。在该区间内有

$$f_{\max} = f(m) = \frac{k}{m^2 \Gamma^2}.$$

可令 $x = m + (b - m)\xi_1$, $y = f_{\max}\xi_2$, 若 $y \leq f(x)$ 则接受 x , 否则重抽。若目标仍按原来的归一化 f 计算，则接受率为

$$\eta_1(b) = \frac{\int_m^b f(x) dx}{(b - m)f_{\max}} = \frac{m^2 \Gamma^2}{k(b - m)} \int_m^b f(x) dx.$$

取 b 越大，尾部遗漏越小，但矩形面积增大、效率下降；通常令尾概率 $\int_b^\infty f(x) dx < \varepsilon$ 后选择尽可能小的 b ，或把峰附近和尾部分段抽样。

- (2) 第二类舍选法取归一化比较函数

$$h(x) = \frac{m}{x^2}, \quad x \in [m, \infty), \quad \int_m^\infty h(x) dx = 1.$$

其分布函数为 $H(x) = 1 - m/x$, 故可由 $x = \frac{m}{1 - \xi_1}$ 抽取 $h(x)$ 的样本。令

$$M = \max_{x \geq m} \frac{f(x)}{h(x)} = \frac{k}{m} \max_{x \geq m} \frac{x^2}{(x^2 - m^2)^2 + m^2 \Gamma^2}.$$

极值点满足 $x_*^2 = m\sqrt{m^2 + \Gamma^2}$, 因而

$$M = \frac{k(\sqrt{m^2 + \Gamma^2} + m)}{2m^2\Gamma^2}.$$

用 ξ_2 判断 $\xi_2 \leq \frac{f(x)}{Mh(x)}$ 时接受。由于 h 已归一化, 抽样效率为 $\eta_2 = 1/M$ 。与第一类舍选法相比, 第二类舍选法不需要人为截断无穷区间, 比较函数还能覆盖长尾; 缺点是需要构造合适的 $h(x)$ 并求出上界 M 。

题 2.

- (1) 令 $Y_i = -\ln \xi_i$, $i = 1, \dots, n$ 。则 Y_i 服从参数为 1 的指数分布, $P(Y_i \leq x) = 1 - e^{-x}$, $x \geq 0$ 。若 $X = \max(Y_1, \dots, Y_n)$, 则

$$P(X \leq x) = \prod_{i=1}^n P(Y_i \leq x) = (1 - e^{-x})^n.$$

对 x 求导得到 $p_X(x) = n(1 - e^{-x})^{n-1}e^{-x}$, 即题设分布。

- (2) 在 $x \geq 0$ 上,

$$\frac{3}{2}(e^{-x} - 2e^{-2x} + e^{-3x}) = \frac{1}{2} \cdot 3(1 - e^{-x})^2 e^{-x},$$

而 $\int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{2}$ 。因此该分布可看成两个权重均为 1/2 的分布叠加: 一半来自 $n = 3$ 的最大指数样本, 一半来自半正态分布。抽样步骤为:

$$\xi_0 \sim U(0, 1).$$

若 $\xi_0 < 1/2$, 产生 ξ_1, ξ_2, ξ_3 , 取 $x = \max\{-\ln \xi_1, -\ln \xi_2, -\ln \xi_3\}$ 。若 $\xi_0 \geq 1/2$, 产生标准正态 $Z \sim N(0, 1)$, 取 $x = |Z|$ 。

题 3. Metropolis 抽样步骤如下: 先给定初始自旋构型 $\vec{s}$, 其中 $s_i = \pm 1$, 并设定温度 T 。每一步随机选取一个格点 i , 提出翻转 $s_i \rightarrow -s_i$ 。只需计算翻转前后的能量差

$$\Delta H = H(\vec{s}') - H(\vec{s}) = 2s_i \sum_j J_{ij} s_j.$$

若 $\Delta H \leq 0$, 接受翻转; 若 $\Delta H > 0$, 以概率 $\exp(-\frac{\Delta H}{k_B T})$ 接受, 否则保持原状态。经过充分热化后, 对构型进行采样统计。系统总磁矩可写为 $M(\vec{s}) = \sum_i s_i$, 其平衡期望为

$$\langle M \rangle = \frac{1}{Z} \sum_{\vec{s}} M(\vec{s}) \exp\left[-\frac{H(\vec{s})}{k_B T}\right].$$

在零外场有限系统中, 由于正负磁化态对称, 常用 $\langle |M| \rangle / N$ 作为序参量。铁磁耦合 $J_{ij} > 0$ 时, 低温下自旋趋于同向排列, $\langle |M| \rangle / N$ 接近 1; 高温下热涨落占主导, 自旋近似随机, $\langle |M| \rangle / N$ 接近 0。在临界温度附近磁矩快速变化, 磁化率涨落通常出现峰值。

题 4.

- (1) 对常微分方程 $y' = f(t, y)$, 设二阶 Runge-Kutta 格式为 $y_{n+1} = y_n + h(\alpha k_1 + \beta k_2)$, 其中 $k_1 = f(t_n, y_n)$, $k_2 = f(t_n + ph, y_n + qhk_1)$ 。与 Taylor 展开匹配到二阶, 需满足

$$\alpha + \beta = 1, \quad \beta p = \frac{1}{2}, \quad \beta q = \frac{1}{2}.$$

取 $\alpha = 0, \beta = 1, p = q = 1/2$, 得到中点格式

$$k_1 = f(t_n, y_n), \quad k_2 = f\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_1\right), \quad y_{n+1} = y_n + hk_2.$$

也可取 $\alpha = \beta = 1/2, p = q = 1$, 得到 Heun 格式。

- (2) 令近似解写为 $\phi_h(r) = \sum_j \phi_j N_j(r)$, 其中 N_j 为有限元基函数, 边界节点满足给定 Dirichlet 条件 $\phi|_L = \phi_0$ 。由

$$I(\phi) = \int_D \left[\frac{\epsilon}{2} |\nabla \phi|^2 - \rho \phi \right] dS$$

作变分, 得到弱形式

$$\int_D \epsilon \nabla \phi \cdot \nabla v dS = \int_D \rho v dS.$$

代入基函数后形成线性方程组 $K\phi = b$, 其中

$$K_{ij} = \int_D \epsilon \nabla N_i \cdot \nabla N_j dS, \quad b_i = \int_D \rho N_i dS.$$

实际计算时在每个单元上计算 K^e, b^e , 组装到全局矩阵, 施加边界条件后求解内部节点的 ϕ_j 。

题 5.

- (1) 正方形被分成 10×10 个格子, 格距为 $a = L/10$ 。截断距离为 $r_c = 0.35L = 3.5a$ 。在周期性边界和最小像约定下, 与任意一个粒子相互作用的其他粒子对应整数位移 (p, q) , 满足 $p^2 + q^2 \leq (3.5)^2$, $(p, q) \neq (0, 0)$ 。因为 p, q 为整数, 所以等价于 $p^2 + q^2 \leq 12$ 。逐项计数: $p = 0 : 6$; $p = \pm 1 : 14$; $p = \pm 2 : 10$; $p = \pm 3 : 6$ 。总数为 $6 + 14 + 10 + 6 = 36$ 。因此不计自身时, 有 36 个粒子与最右侧粒子有相互作用。
- (2) 对关联函数 $g(r)$ 描述在距离 r 处找到另一粒子的概率相对于理想均匀气体的增强或减弱。在二维体系中可写为

$$g(r) = \frac{1}{\rho N} \left\langle \sum_{i \neq j} \frac{\delta(r - r_{ij})}{2\pi r} \right\rangle, \quad \rho = N/L^2.$$

Lennard-Jones 势下, 短距离因强排斥而 $g(r) \approx 0$; 在势能极小值附近 $r = 2^{1/6}\sigma$ 附近出现第一峰; 液体中随后出现逐渐衰减的振荡峰, 远距离 $g(r) \rightarrow 1$ 。低温或固体相中峰更尖锐且延伸更远。

- (3) 速度 Verlet 基本流程为: 先给定 $r_i(t), v_i(t)$, 由势能计算 $F_i(t)$; 再更新位置

$$r_i(t + \Delta t) = r_i(t) + v_i(t)\Delta t + \frac{F_i(t)}{2m_i}\Delta t^2.$$

对新位置施加周期性边界条件, 重新计算 $F_i(t + \Delta t)$, 然后更新速度

$$v_i(t + \Delta t) = v_i(t) + \frac{F_i(t) + F_i(t + \Delta t)}{2m_i}\Delta t.$$

重复上述步骤, 并在热化后统计所需物理量。

题 6.

- (1) 指数项要求 λr 无量纲, 因此 $[\lambda] = L^{-1}$ 。势能项 ar^3 具有能量量纲, 故 $[a] = E L^{-3}$ 。
- (2) 验证归一化与正交性的 Mathematica 代码可写为:

```

$Assumptions = lambda > 0 && r >= 0;
psi0[r_] := lambda^(3/2)/Sqrt[Pi] Exp[-lambda r];
psi1[r_] := lambda^(3/2)/Sqrt[3 Pi] (3 - 2 lambda r) Exp[-lambda r];
norm0 = FullSimplify[Integrate[4 Pi r^2 psi0[r]^2, {r, 0, Infinity}]]
norm1 = FullSimplify[Integrate[4 Pi r^2 psi1[r]^2, {r, 0, Infinity}]]
orth01 = FullSimplify[Integrate[4 Pi r^2 psi0[r] psi1[r], {r, 0, Infinity}]]

```

运行结果应分别为 1,1,0。

- (3) 对球对称函数, 径向 Laplace 算子为 $\nabla^2 \psi = \frac{1}{r^2} \frac{d}{dr} (r^2 \frac{d\psi}{dr})$ 。用 Ψ_0 作试探函数时, 能量期望为

$$E(\lambda) = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle = \frac{\lambda^2}{2\mu} + \frac{15a}{2\lambda^3}.$$

令 $dE/d\lambda = 0$, 可得 $\lambda^5 = \frac{45a\mu}{2}$ 。对应 Mathematica 代码可写为:

```

$Assumptions = lambda > 0 && a > 0 && mu > 0 && r > 0;
psi0[r_] := lambda^(3/2)/Sqrt[Pi] Exp[-lambda r];
lap[u_] := 1/r^2 D[r^2 D[u, r], r];
Hpsi0 = -lap[psi0[r]]/(2 mu) + a r^3 psi0[r];
E0 = FullSimplify[Integrate[4 Pi r^2 psi0[r] Hpsi0, {r, 0, Infinity}]]
lambdaOpt = Solve[D[E0, lambda] == 0, lambda, Reals]
Emin = FullSimplify[E0 /. lambdaOpt[[1]]]

```

题 7.

- (1) a. ×; b. ×; c. ×; d. ✓; e. ×; f. ✓。

理由: 无监督学习可发现簇结构, 但没有标签时不能直接预言给定类别; 随机梯度下降每步只用一个或小批样本, 适合大样本问题, 但不是“需要更大样本”; 稀疏通常指特征向量中大量分量为零或特征数远多于样本数; 多分类逻辑回归可由线性得分接 Softmax 实现; 反向传播的核心用途就是高效计算代价函数对参数的梯度。若题目默认决策树完全生长到纯叶, 则第 (e) 项也可作为理想情形下的正确表述; 严格说, 实际剪枝或停止条件下叶节点 Gini 杂度不一定为 0。

- (2) 决定系数为

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}.$$

其中分子为残差平方和, 分母为总平方和。 R^2 越接近 1, 说明模型解释的方差比例越高, 训练拟合越好; $R^2 = 0$ 表示与总是预测均值 $\bar{y}$ 相当; $R^2 < 0$ 表示模型比均值预测还差。

- (3) 可用一个三输入、单输出的逻辑神经元分类。把标签编码为 $A \rightarrow 0, B \rightarrow 1$ 。对样本 $x = (x_1, x_2, x_3)^T$, 网络结构为

$$z = w^T x + b, \quad \hat{y} = \sigma(z) = \frac{1}{1 + e^{-z}}.$$

代价函数可取二元交叉熵

$$J(w, b) = -\frac{1}{N} \sum_{i=1}^N [y_i \ln \hat{y}_i + (1 - y_i) \ln(1 - \hat{y}_i)].$$

传递函数取 Sigmoid；传递系数为权重 w_1, w_2, w_3 与偏置 b 。训练流程为：整理数据并编码标签，初始化参数；前向传播计算 $\hat{y}_i$ ；计算 J ；反向传播求

$$\frac{\partial J}{\partial w} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i) x_i, \quad \frac{\partial J}{\partial b} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i);$$

用学习率 α 更新 $w \leftarrow w - \alpha \partial J / \partial w$, $b \leftarrow b - \alpha \partial J / \partial b$ 。迭代到收敛后，用阈值 $\hat{y} \geq 0.5$ 判为 B，否则判为 A。

5 2024 年秋季学期（魏逸丰班）

试卷

题 1. 蒙特卡洛方法

取 $a = 137$, $c = 187$, $m = 256$, $x_0 = 1$ ，用线性同余法产生伪随机数。令

$$x_{n+1} = (ax_n + c) \bmod m, \quad \xi_n = \frac{x_n}{m}.$$

写出二维点列 $\{(\xi_n, \xi_{n+1})\}_{n \geq 0}$ 的生成方法，并给出前若干个点的。

题 2. 蒙特卡洛方法

描述抽样与估计步骤。

(1) 抽样指数分布 $f(x) = \lambda e^{-\lambda x}$, $x \in [0, \infty)$, $\lambda > 0$ 。

(2) 重要抽样法：用第一问的指数随机数产生器求积分 $\int_0^{+\infty} x^{3/2} e^{-x} dx$ 。

题 3. 蒙特卡洛方法

二维 Ising 模型中，设 $H(s) = -J \sum_{\langle i, j \rangle} s_i s_j$, $P(s) \propto e^{-H(s)/(k_B T)}$ 。用 Metropolis 方法求总磁矩 $M = \sum_i s_i$ 的期望。讨论配分函数 Z 未知是否会影响抽样。

题 4. 数值微分方程

(1) 对 $\nabla^2 u = f(x, y)$ ，设二维均匀直角网格步长为 h_x, h_y ，写出二阶中心差分形式和五点差分格式。

(2) 比较有限差分法和有限元法的异同。

题 5. 分子动力学

设二维正方形周期元胞边长为 $L = 10$ ，均匀划分为 10×10 个小格，每个小格中心放置一个粒子，目标粒子为右下角粒子。考虑最小像力约定和周期性边界条件，截断距离为 $r_c = 3$ 。判断目标粒子和哪些粒子有相互作用，并写出 Verlet 方法。

题 6. 机器学习

判断下列操作可以改善欠拟合还是过拟合：增大正则化系数、减小正则化系数、增加训练样本、添加新特征、增大模型复杂度。

题 7. 机器学习

给定 $x, w_1, w_2, w_3, b_1, b_2, b_3, y$ ，网络结构取为一维前馈链

$$x \longrightarrow a^{[1]} \longrightarrow a^{[2]} \longrightarrow \hat{y},$$

其中隐藏层 $g(z) = \frac{1}{1+e^{-z}}$ ，输出层 $f(z) = z$ ，误差函数 $J = \frac{1}{2}(\hat{y} - y)^2$ 。

- (1) 求隐藏层和输出层的值 $a^{[1]}, a^{[2]}, \hat{y}$ 。
- (2) 写出反向传播的表达式，并说明其中超参数 α 是什么。
- (3) 计算 J 对 $\hat{y}, a^{[2]}, a^{[1]}$ 的偏导： $\frac{\partial J}{\partial \hat{y}}, \frac{\partial J}{\partial a^{[2]}}, \frac{\partial J}{\partial a^{[1]}}$ 。
- (4) 计算 J 对 $w_1, w_2, w_3, b_1, b_2, b_3$ 的偏导。

参考解答

题 1. 线性同余递推为

$$x_{n+1} = (137x_n + 187) \bmod 256, \quad \xi_n = \frac{x_n}{256}.$$

由 $x_0 = 1$ 开始，前若干项为

n	0	1	2	3	4	5	6	7
x_n	1	68	31	82	157	192	123	142

因此二维点列的前若干个数为

$$\left(\frac{1}{256}, \frac{68}{256}\right), \left(\frac{68}{256}, \frac{31}{256}\right), \left(\frac{31}{256}, \frac{82}{256}\right), \left(\frac{82}{256}, \frac{157}{256}\right), \dots$$

因为 $c = 187$ 与 256 互素，且 $a - 1 = 136$ 可被 4 整除，该线性同余序列对模 2^8 满足满周期条件，周期为 256 。生成二维数组时只需按递推依次生成 x_n ，再把相邻归一化随机数组成点 (ξ_n, ξ_{n+1}) 。

题 2.

- (1) 指数分布的分布函数为 $F(x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}$ 。令 $F(x) = \xi$ ，其中 $\xi \sim U(0, 1)$ ，得到反函数抽样公式

$$x = -\frac{1}{\lambda} \ln(1 - \xi).$$

由于 $1 - \xi$ 与 ξ 同分布，实际计算中常写为 $x = -\frac{1}{\lambda} \ln \xi$ 。

- (2) 令 $I = \int_0^\infty x^{3/2} e^{-x} dx$ 。用指数分布 $p_\lambda(x) = \lambda e^{-\lambda x}$ 做重要抽样，则

$$I = \int_0^\infty \frac{x^{3/2} e^{-x}}{\lambda e^{-\lambda x}} p_\lambda(x) dx = \mathbb{E}_{p_\lambda} \left[\frac{1}{\lambda} X^{3/2} e^{-(1-\lambda)X} \right].$$

生成 N 个指数样本 X_k 后，估计量为

$$\hat{I}_N = \frac{1}{N} \sum_{k=1}^N \frac{1}{\lambda} X_k^{3/2} e^{-(1-\lambda)X_k}.$$

若取 $\lambda = 1$, 则 $X_k \sim e^{-x}$, 估计量化为 $\hat{I}_N = \frac{1}{N} \sum_{k=1}^N X_k^{3/2}$ 。解析值可作为校验:

$$I = \Gamma\left(\frac{5}{2}\right) = \frac{3\sqrt{\pi}}{4}.$$

题 3. 二维 Ising 模型的目标分布可写为

$$\pi(s) = \frac{1}{Z} e^{-H(s)/(k_B T)}, \quad H(s) = -J \sum_{\langle i,j \rangle} s_i s_j.$$

Metropolis 抽样步骤为:

- (1) 给定初始自旋构型 s , 每个 $s_i = \pm 1$ 。
- (2) 随机选取一个格点 i , 提出翻转 $s_i \rightarrow -s_i$, 得到候选构型 s' 。
- (3) 只计算局部能量差 $\Delta H = H(s') - H(s) = 2J s_i \sum_{j \in \text{nn}(i)} s_j$ 。
- (4) 以概率 $A(s \rightarrow s') = \min\{1, e^{-\Delta H/(k_B T)}\}$ 接受翻转; 否则保持原构型。
- (5) 热化后记录样本构型的总磁矩 $M(s) = \sum_i s_i$, 用样本平均估计 $\langle M \rangle \approx \frac{1}{N} \sum_k M(s^{(k)})$ 。

配分函数 Z 未知不影响 Metropolis 抽样, 因为接受率只用到概率比

$$\frac{\pi(s')}{\pi(s)} = \frac{e^{-H(s')/(k_B T)}/Z}{e^{-H(s)/(k_B T)}/Z} = e^{-\Delta H/(k_B T)},$$

归一化常数完全抵消。

题 4.

- (1) 对二维网格函数 $u_{i,j}$, 二阶中心差分为

$$u_{xx}(x_i, y_j) \approx \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h_x^2}, \quad u_{yy}(x_i, y_j) \approx \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h_y^2}.$$

因而

$$\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h_x^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h_y^2} = f_{i,j}.$$

若 $h_x = h_y = h$, 得到五点差分格式

$$\frac{u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j}}{h^2} = f_{i,j},$$

或 $u_{i,j} = \frac{1}{4}(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - h^2 f_{i,j})$ 。

- (2) 有限差分法直接用差商近似微分算子, 推导和程序实现简单, 适合规则区域、规则网格和简单边界; 不足是复杂几何、非均匀材料和复杂边界条件处理不灵活。有限元法从弱形式或变分形式出发, 在单元上选取基函数并组装整体矩阵, 适合复杂区域、非均匀网格和多种边界条件; 不足是网格剖分、基函数选择和矩阵组装更复杂。

题 5. 设目标粒子在右下角格心。由于盒长 $L = 10$ 、格距为 1, 在最小像约定下, 其他粒子相对目标粒子的最小像位移可写为整数对 (p, q) 。截断条件为 $p^2 + q^2 < 3^2$, $(p, q) \neq (0, 0)$ 。满足条件的相对位移为

$$(\pm 1, 0), (0, \pm 1), (\pm 1, \pm 1), (\pm 2, 0), (0, \pm 2), (\pm 2, \pm 1), (\pm 1, \pm 2), (\pm 2, \pm 2).$$

共 $4 + 4 + 4 + 8 + 4 = 24$ 个粒子与目标粒子相互作用。若采用 $r \leq r_c$ 的截断约定, 则还会包括 $(\pm 3, 0), (0, \pm 3)$, 总数变为 28。

速度 Verlet 方法可写为

$$r_i(t + \Delta t) = r_i(t) + v_i(t)\Delta t + \frac{1}{2}a_i(t)\Delta t^2,$$

先由新位置计算 $F_i(t + \Delta t)$ 和 $a_i(t + \Delta t) = F_i(t + \Delta t)/m_i$, 再更新速度

$$v_i(t + \Delta t) = v_i(t) + \frac{1}{2}[a_i(t) + a_i(t + \Delta t)]\Delta t.$$

若使用位置 Verlet, 也可写成 $r_i(t + \Delta t) = 2r_i(t) - r_i(t - \Delta t) + a_i(t)\Delta t^2$ 。

题 6. 各操作对应的主要作用为

操作	主要改善的问题
增大正则化系数	过拟合
减小正则化系数	欠拟合
增加训练样本	过拟合
添加新特征	欠拟合
增大模型复杂度	欠拟合

理由是: 过拟合通常表现为模型方差过大、训练误差低但测试误差高, 因此可通过更强正则化或更多训练样本降低方差; 欠拟合通常表现为模型表达能力不足、训练误差和测试误差都较高, 因此可通过减弱正则化、添加有效特征或增大模型复杂度降低偏差。

题 7.

(1) 记

$$z^{[1]} = w_1x + b_1, \quad a^{[1]} = g(z^{[1]}), \quad z^{[2]} = w_2a^{[1]} + b_2, \quad a^{[2]} = g(z^{[2]}), \quad z^{[3]} = w_3a^{[2]} + b_3, \quad \hat{y} = f(z^{[3]}) = z^{[3]}.$$

因此

$$a^{[1]} = \frac{1}{1 + e^{-(w_1x + b_1)}}, \quad a^{[2]} = \frac{1}{1 + e^{-(w_2a^{[1]} + b_2)}}, \quad \hat{y} = w_3a^{[2]} + b_3.$$

(2) 反向传播就是从损失函数开始, 按链式法则由输出层向输入层逐层计算梯度。令 $\delta^{[3]} = \partial J / \partial z^{[3]}$, $\delta^{[2]} = \partial J / \partial z^{[2]}$, $\delta^{[1]} = \partial J / \partial z^{[1]}$, 先求 $\delta^{[3]}$, 再求 $\delta^{[2]}$, 最后求 $\delta^{[1]}$, 并由这些量得到各权重和偏置的梯度。超参数 α 是学习率, 用于控制梯度下降更新步长: $\theta \leftarrow \theta - \alpha \partial J / \partial \theta$ 。

(3) 由于 $J = \frac{1}{2}(\hat{y} - y)^2$, 有

$$\frac{\partial J}{\partial \hat{y}} = \hat{y} - y, \quad \frac{\partial J}{\partial a^{[2]}} = (\hat{y} - y)w_3,$$

$$\frac{\partial J}{\partial a^{[1]}} = (\hat{y} - y)w_3 g'(z^{[2]}) w_2 = (\hat{y} - y)w_3 w_2 a^{[2]}(1 - a^{[2]}).$$

(4) 因 $f(z) = z$, 有 $\delta^{[3]} = \hat{y} - y$ 。输出层参数梯度为

$$\frac{\partial J}{\partial w_3} = \delta^{[3]}a^{[2]}, \quad \frac{\partial J}{\partial b_3} = \delta^{[3]}.$$

第二层隐藏层误差项为 $\delta^{[2]} = \delta^{[3]}w_3 g'(z^{[2]}) = (\hat{y} - y)w_3 a^{[2]}(1 - a^{[2]})$, 因而

$$\frac{\partial J}{\partial w_2} = \delta^{[2]}a^{[1]}, \quad \frac{\partial J}{\partial b_2} = \delta^{[2]}.$$

第一层隐藏层误差项为 $\delta^{[1]} = \delta^{[2]}w_2 g'(z^{[1]}) = \delta^{[2]}w_2 a^{[1]}(1 - a^{[1]})$, 所以

$$\frac{\partial J}{\partial w_1} = \delta^{[1]}x, \quad \frac{\partial J}{\partial b_1} = \delta^{[1]}.$$

6 2025 年春季学期（张志勇班）

试卷

题 1. 蒙特卡洛方法

- (1) 利用线性同余法生成随机数序列，参数 $a = 5$, $c = 3$, $m = 32$, 初始值 $x_0 = 6$, 写出前三个数 ξ_1, ξ_2, ξ_3 。
- (2) 考虑 Metropolis 方法生成正态分布序列 $f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ 。初始值取 x_0 , 试探区间 $[x_i - \delta, x_i + \delta]$, $\delta = 2$ 。利用 (1) 生成的随机数 ξ_1 进行试探, ξ_2 决定是否接受。

题 2. 蒙特卡洛方法

考虑分布函数 $f(x) = 3x^2$, $x \in [0, 1]$ 。

- (1) 直接抽样抽取 $f(x)$ 的样本, 写出抽样步骤或画出流程图。
- (2) 利用第一类舍选法进行抽样, 写出抽样步骤或画出流程图。
- (3) 将第一类舍选法的对象推广到 $f(x) = nx^{n-1}$, $x \in [0, 1]$ 。证明抽取的对象可以写为 $\eta = \max(\xi_1, \dots, \xi_n)$, 其中 $\xi_i \in [0, 1]$ 是随机数, $i = 1, 2, \dots, n$ 。

题 3. 有限差分法

考虑微分方程

$$\begin{cases} u''(x) = 2, & x \in [0, 1], \\ u(0) = 0, & u(1) = 1. \end{cases}$$

- (1) 将微分方程离散化, 取步长 $h = 1/2$ 。写出离散后的差分方程并求出数值解。
- (2) 已知方程的解析解为 $u(x) = x^2$, 计算 $x = 1/2$ 处数值解与解析解之间的误差。若步长改为 $h = 1/3$, 误差是否有变化? 说明原因。

题 4. 有限元法

考虑微分方程

$$\frac{d^2u}{dx^2} = 1, \quad x \in [0, 1], \quad u(0) = u(1) = 0.$$

利用有限元法进行求解。已知微分方程对应的泛函写为

$$I = \int_0^1 \left[\frac{1}{2} \left(\frac{du}{dx} \right)^2 + u \right] dx.$$

将区间平均分为四格, 即 $\Delta x = 1/4$ 。

- (1) 进行元素内线性插值, 写出 $u(x)$ 与 $u'(x)$ 。
- (2) 写出每个元素内的子泛函。
- (3) 求总泛函的泛函极值, 以此写出节点的函数值 $u(0.25), u(0.5), u(0.75)$ 。

题 5. 分子动力学

- (1) 对于微正则系综进行分子动力学模拟, 写出速度 Verlet 算法中位置与速度的递推公式。

- (2) 考虑二维平面内的分子动力学模拟，平面尺寸为 10×10 ，分为 100 个小格。两个粒子分别位于 $A(2,3)$ 和 $B(3,4)$ ，且初始速度均为 0。假设粒子质量为 1，粒子间相互作用为 Lennard-Jones 势，参数设为 $\epsilon = \sigma = 1$ 。取时间步长为 $\Delta t = 0.1$ ，求一步之后 B 的位置与速度。LJ 势下两个粒子 i, j 之间 i 的受力在 x 方向上的分量为

$$F_{i,x} = 48(x_i - x_j) \left(\frac{\epsilon}{\sigma^2} \right) \left[\left(\frac{\sigma}{r_{ij}} \right)^{14} - \frac{1}{2} \left(\frac{\sigma}{r_{ij}} \right)^8 \right].$$

- (3) 在考虑基团相互作用时，记 C_α 原子为 marker 原子。比较 marker 原子之间的距离与截断距离，判断是否需要考虑基团之间的相互作用。这种操作可能带来什么问题？如何解决？
- (4) 分子动力学模拟 (MD) 与蒙特卡洛模拟 (MC) 是两种不同的模拟方式。考虑一种结合 MD 与 MC 的模拟方式：在进行一个时间步长的 MD 后，利用 Metropolis 方法判断是否接受这一步，再进行下一步 MD。试与 MD 和 MC 的特点比较，写出这种模拟可能存在的优点与缺点。
- (5) 为什么在现实的分子动力学模拟中常常将时间步长设置为 1 fs？实验中，往往通过“增加”氢原子质量来增加时间步长，这么做的依据是什么？

题 6. 机器学习

- (1) 判断题：
- (a) 决策树输入数据前要对数据进行归一化/标准化。
 - (b) 集成多个线性分类器也无法进行非线性分类。
 - (c) Sigmoid 函数容易使梯度消失。
 - (d) 回归算法给出连续输出，分类算法给出离散输出。
 - (e) L1 正则化用于特征选择。
- (2) 填空题：
- (a) 写出三个降维的方法：____、____、____。
 - (b) BP 预测中误差传递的顺序为____层、____层、____层。
 - (c) 在 (a) 增大正则化系数；(b) 添加新特征；(c) 增加训练集；(d) 增大模型复杂度四个操作中，用于处理欠拟合的是____；用于处理过拟合的是____。
 - (d) 机器学习中模型、____、____ 和模型评估方法是机器学习的核心要素。
- (3) 下图体现了机器学习中的核心原则，试对此做出解释，并阐述集成学习的优势。
- (4) 利用聚合聚类对下面数据进行聚类。采用欧式距离衡量样本间的相似性，最小距离衡量待合并的两类间的相似性。

	P1	P2	P3	P4	P5	P6
X	3	1	7	5	7	6
Y	2	3	2	6	5	6

参考解答

题 1.

(1) 线性同余递推为 $x_{n+1} = (5x_n + 3) \bmod 32$ 。由 $x_0 = 6$ 得

$$x_1 = 1, \quad x_2 = 8, \quad x_3 = 11,$$

因而

$$\xi_1 = \frac{1}{32}, \quad \xi_2 = \frac{1}{4}, \quad \xi_3 = \frac{11}{32}.$$

(2) 用 ξ_1 在 $[x_0 - 2, x_0 + 2]$ 内作均匀试探, 可写为

$$x^* = x_0 - 2 + 4\xi_1 = x_0 - \frac{15}{8}.$$

对正态分布, 接受率为

$$A = \min\left\{1, \frac{f(x^*)}{f(x_0)}\right\} = \min\left\{1, \exp\left[-\frac{(x^*)^2 - x_0^2}{2}\right]\right\}.$$

若 $\xi_2 = 1/4 < A$, 则接受 x^* ; 否则保持在 x_0 。若题面取 $x_0 = 0$, 则

$$A = \exp\left(-\frac{225}{128}\right) \approx 0.172 < \frac{1}{4},$$

本步应拒绝。

题 2.

(1) 分布函数为 $F(x) = \int_0^x 3t^2 dt = x^3$ 。令 $F(x) = \xi$, 直接抽样公式为 $x = \xi^{1/3}$, $\xi \sim U(0, 1)$ 。

(2) 第一类舍选法可在矩形 $0 \leq x \leq 1$, $0 \leq y \leq 3$ 内均匀取点, 若 $y \leq 3x^2$ 则接受 x , 否则重新抽样。接受率为

$$\eta = \frac{\int_0^1 3x^2 dx}{3} = \frac{1}{3}.$$

(3) 若 $\eta = \max(\xi_1, \dots, \xi_n)$, 则

$$P(\eta \leq x) = P(\xi_1 \leq x, \dots, \xi_n \leq x) = x^n, \quad 0 \leq x \leq 1.$$

因而 $p_\eta(x) = \frac{d}{dx}x^n = nx^{n-1}$, 正好得到目标分布。

题 3.

(1) 中心差分格式为 $\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = 2$ 。取 $h = 1/2$ 时只有一个内部节点 $x_1 = 1/2$, 且 $u_0 = 0, u_2 = 1$, 故

$$\frac{1 - 2u_1}{(1/2)^2} = 2,$$

解得 $u(1/2) \approx u_1 = \frac{1}{4}$ 。

(2) 解析解在 $x = 1/2$ 处为 $u(1/2) = 1/4$, 误差为 0。若 $h = 1/3$, 中心差分对二次函数的二阶导数是精确的, 数值解仍等于节点上的解析解: $u(1/3) = 1/9$, $u(2/3) = 4/9$ 。因此节点误差仍为 0。

题 4.

(1) 在元素 $[x_i, x_{i+1}]$ 上令 $h = x_{i+1} - x_i = 1/4$, 线性插值为

$$u(x) = u_i \frac{x_{i+1} - x}{h} + u_{i+1} \frac{x - x_i}{h}, \quad u'(x) = \frac{u_{i+1} - u_i}{h}.$$

(2) 元素子泛函为

$$I_i = \int_{x_i}^{x_{i+1}} \left[\frac{1}{2}(u')^2 + u \right] dx = \frac{(u_{i+1} - u_i)^2}{2h} + \frac{h}{2}(u_i + u_{i+1}).$$

(3) 记内部节点值为 $u = (u_1, u_2, u_3)^T$, 由 $\partial I / \partial u_i = 0$ 得

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = -\frac{1}{16} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

解得

$$u(0.25) = -\frac{3}{32}, \quad u(0.5) = -\frac{1}{8}, \quad u(0.75) = -\frac{3}{32}.$$

题 5.

(1) 速度 Verlet 格式为

$$r_i(t + \Delta t) = r_i(t) + v_i(t)\Delta t + \frac{1}{2}a_i(t)\Delta t^2, \quad v_i(t + \Delta t) = v_i(t) + \frac{1}{2}[a_i(t) + a_i(t + \Delta t)]\Delta t,$$

其中 $a_i = F_i/m_i$ 。

(2) 初始时 $r_B - r_A = (1, 1)$, $r = \sqrt{2}$ 。B 粒子受到 A 粒子的力为

$$F_B = 48(1, 1) \left[\left(\frac{1}{\sqrt{2}} \right)^{14} - \frac{1}{2} \left(\frac{1}{\sqrt{2}} \right)^8 \right] = \left(-\frac{9}{8}, -\frac{9}{8} \right).$$

因 $m = 1$, 一步后

$$r_B(0.1) = (3, 4) + \frac{1}{2} \left(-\frac{9}{8}, -\frac{9}{8} \right) (0.1)^2 = (2.994375, 3.994375).$$

同时 $r_A(0.1) = (2.005625, 3.005625)$, 重新计算可得 $F_B(0.1) \approx (-1.189208, -1.189208)$ 。因而

$$v_B(0.1) = \frac{1}{2}[F_B(0) + F_B(0.1)]\Delta t \approx (-0.115710, -0.115710).$$

(3) 用 C_α 作为 marker 时, 可以先比较 $d(C_\alpha, C_\alpha)$ 与截断半径 r_c : 若距离明显大于截断半径, 则跳过该基团对; 若不大于截断半径, 则进一步计算基团内原子对相互作用。问题在于 marker 距离不能完全代表基团真实几何范围, 可能出现 marker 较远但侧链原子接近的漏判, 也可能出现 marker 较近但所有原子对均超出截断的误判。常用修正是为每个基团建立包围球或包围盒, 用 $d(C_\alpha, C_\alpha) \leq r_c + R_1 + R_2$ 作为预筛选条件, 再对通过筛选的基团做精确原子对距离判断。

(4) 这种方法相当于用 MD 产生相关但较大的候选构型, 再用 Metropolis 接受率校正统计分布。优点是候选移动比普通随机 MC 更有物理方向性, 可能降低随机游走, 提高构型空间探索效率; 若接受率按目标系综构造, 还能保持正确平衡分布。缺点是每次候选移动都要做力计算, 代价高; MD 积分误差过大时接受率会下降; 若积分器不满足时间可逆和相空间体积守恒, Metropolis 校正会失去严格依据; 接受或拒绝也会破坏真实动力学轨迹的连续物理意义。

(5) 常用 1 fs 是因为分子中最快的振动通常来自含氢键, 振动周期约为若干到十几飞秒, 时间步长必须显著小于最快周期才能稳定积分并保持能量守恒。增大氢原子质量会增大相关振动的约化质量, 使 $\omega = \sqrt{k/\mu}$ 下降, 最快振动变慢, 因此允许更大的时间步长。代价是动力学时间尺度和动力学性质会被改变, 通常只应在关注平衡采样或经过验证的近似中使用。

题 6.

(1) a. $\times$; b. $\times$; c. $\checkmark$; d. $\checkmark$; e. $\checkmark$ 。

理由：决策树按特征阈值划分，通常不要求归一化；多个线性分类器通过投票、分段或提升可以形成非线性决策边界；Sigmoid 在饱和区导数很小，容易造成梯度消失；回归和分类输出类型分别对应连续值和离散类别；L1 正则化能使部分参数变为零，因此可用于特征选择。

(2) 填空可写为：

- PCA、LDA、t-SNE;
- 输出层、隐藏层、输入层;
- 欠拟合: (b)、(d); 过拟合: (a)、(c);
- 模型、策略、算法和模型评估方法。

(3) 图中体现的是偏差和方差的权衡。模型过简单时偏差大，训练集和测试集表现都差，属于欠拟合；模型过复杂时训练误差低但方差大，测试误差上升，属于过拟合。集成学习通过组合多个模型改善泛化：Bagging 和随机森林主要通过降低模型相关性来减小方差，Boosting 通过逐步修正前一轮错误来降低偏差，最终常能得到比单个模型更稳定的预测。

(4) 使用欧氏距离和最小距离准则，首先有 $d(P_4, P_6) = 1$ ，合并 $\{P_4, P_6\}$ 。随后

$$d(P_5, \{P_4, P_6\}) = \min(\sqrt{5}, \sqrt{2}) = \sqrt{2},$$

合并为 $\{P_4, P_5, P_6\}$ 。接着 $d(P_1, P_2) = \sqrt{5}$ ，合并为 $\{P_1, P_2\}$ 。再有 $d(P_3, \{P_4, P_5, P_6\}) = 3$ ，合并为 $\{P_3, P_4, P_5, P_6\}$ 。最后两类 $\{P_1, P_2\}$ 与 $\{P_3, P_4, P_5, P_6\}$ 的最小距离为 $d(P_1, P_3) = 4$ ，在距离 4 处合并为一类。

7 2025 年秋季学期（高阳班）

试卷

题 1. 蒙特卡洛方法

宽度为 1 的无限深方势阱，有 0.8 概率位于基态，0.2 概率位于第一激发态。如何根据加分布和舍选法抽样位置？如何获得 $F(x) = x^2$ 的观测量？

题 2. 蒙特卡洛方法

M 为 100 阶三对角方阵，主对角元为 $1 - a$ ，次对角元为 $-b$ ，且 $a + 2b = 1$ 。问如何用随机游走解方程 $MX = Y$ 。

题 3. 有限差分法

有限差分法解极坐标系 Laplace 方程，区域为圆环 $0.2 < r < 1$ 。内部节点为 $r = 0.6$, $\theta = 0, \frac{2\pi}{3}, \frac{4\pi}{3}$ 。边界条件为

$$\phi(0.2, \theta) = 0.2 \sin \theta, \quad \phi(1, \theta) = \sin \theta.$$

题 4. 数值线性代数

Gauss-Seidel 迭代解线性方程组（三阶方阵）。给定初值，迭代两次：

$$A = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 15 \\ 10 \\ 10 \end{pmatrix}, \quad x^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

题 5. 有限元法

给定单位圆上的 8 个点 $(\cos \frac{n\pi}{4}, \sin \frac{n\pi}{4})$, $n = 0, \dots, 7$, 和圆内 2 个点 $(\pm 0.1, 0)$ 。请给出合适的三角剖分，并阐述如何在计算机中存储此剖分。

题 6. 有限元法

一维有限元法：在 $[0, 1]$ 上 $y'' + \pi^2 y = 0$ ，边界条件为 $y(0) = \frac{1}{2}$, $y(1) + 0.2y'(1) = -0.67$ 。给出对应泛函并证明其取极值的条件确为此方程成立；在取 $0, 0.2, 0.5, 1$ 四个节点时求解问题。

题 7. 分子动力学

类似于 Ising 模型，只有相邻位置对哈密顿量有贡献。给定

$$\frac{dS_i}{dt} = \alpha(S_i \times B_i) + \beta\left(S_i \times \frac{dS_i}{dt}\right), \quad B_i = \frac{\partial H}{\partial S_i}.$$

请根据分子动力学设计算法，实现求解一维问题中的平衡态，其中 H 为已给定的相邻自旋哈密顿量。

题 8. 分子动力学

最小像力约定：五个点为 $1(0.1, 0.1)$, $2(0.2, 0.3)$, $3(0.3, 0.9)$, $4(0.4, 0.4)$, $5(0.8, 0.8)$ 。正方形区域为 $[0, 1] \times [0, 1]$ ，截断距离为 0.4。问每个粒子受到的力都来自哪些粒子（或其像），并给出坐标。

参考解答

题 1. 无限深方势阱取区间 $0 < x < 1$ ，本征态可写为 $\psi_n(x) = \sqrt{2} \sin(n\pi x)$ 。题设位置分布为混合分布

$$p(x) = 0.8|\psi_1(x)|^2 + 0.2|\psi_2(x)|^2 = 1.6 \sin^2(\pi x) + 0.4 \sin^2(2\pi x).$$

按加分布抽样时，先产生 $\xi \sim U(0, 1)$ ：若 $\xi < 0.8$ ，选择基态分布 $p_1(x) = 2 \sin^2(\pi x)$ ；否则选择第一激发态分布 $p_2(x) = 2 \sin^2(2\pi x)$ 。对给定的 $p_n(x)$ ，可用舍选法抽样：取 $x \sim U(0, 1)$, $u \sim U(0, 2)$ ，若 $u \leq 2 \sin^2(n\pi x)$ 则接受，否则重抽。得到样本 x_k 后，观测量

$$\langle F \rangle = \langle x^2 \rangle \approx \frac{1}{N} \sum_{k=1}^N x_k^2.$$

解析校验为

$$\langle x^2 \rangle_n = \int_0^1 2x^2 \sin^2(n\pi x) dx = \frac{1}{3} - \frac{1}{2n^2\pi^2},$$

因此混合态下

$$\langle x^2 \rangle = 0.8\left(\frac{1}{3} - \frac{1}{2\pi^2}\right) + 0.2\left(\frac{1}{3} - \frac{1}{8\pi^2}\right) = \frac{1}{3} - \frac{17}{40\pi^2}.$$

题 2. 将方程写成随机游走适用的形式。第 i 行为 $(1-a)X_i - bX_{i-1} - bX_{i+1} = Y_i$ 。因为 $a+2b=1$, 有 $1-a=2b$, 于是内点满足

$$X_i = \frac{1}{2}X_{i-1} + \frac{1}{2}X_{i+1} + \frac{Y_i}{1-a}.$$

记 $g_i = \frac{Y_i}{1-a}$, 则 X_i 等于从 i 出发的对称随机游走沿途源项 g 的期望累积。具体算法为: 从待求分量 i 出发, 令粒子每步以概率 $1/2$ 向 $i-1$ 走、以概率 $1/2$ 向 $i+1$ 走; 走到 0 或 101 时吸收, 对应边界贡献取 0。每条路径记录

$$W = \sum_{k=0}^{\tau-1} g_{S_k},$$

其中 S_k 为第 k 步所在位置, τ 为吸收时刻。重复 N 条路径后

$$X_i \approx \frac{1}{N} \sum_{\ell=1}^N W_{\ell}.$$

对所有 $i=1, \dots, 100$ 分别估计即可得到向量 X 。

题 3. 极坐标 Laplace 方程为

$$\frac{\partial^2 \phi}{\partial r^2} + \frac{1}{r} \frac{\partial \phi}{\partial r} + \frac{1}{r^2} \frac{\partial^2 \phi}{\partial \theta^2} = 0.$$

在 $r=0.6$ 上取 $\Delta r=0.4$, $\Delta \theta=2\pi/3$, 中心差分给出

$$\frac{\phi_j^+ - 2\phi_j + \phi_j^-}{\Delta r^2} + \frac{1}{r} \frac{\phi_j^+ - \phi_j^-}{2\Delta r} + \frac{1}{r^2} \frac{\phi_{j+1} - 2\phi_j + \phi_{j-1}}{\Delta \theta^2} = 0,$$

其中角向按周期边界处理, $\phi_j^- = 0.2 \sin \theta_j$, $\phi_j^+ = \sin \theta_j$ 。由于边界条件均正比于 $\sin \theta$, 设内点 $\phi(0.6, \theta_j) = K \sin \theta_j$ 。对三点角网格有 $\phi_{j+1} - 2\phi_j + \phi_{j-1} = -3K \sin \theta_j$ 。代入差分方程得

$$K = \frac{1.2/\Delta r^2 + 0.8/(2r\Delta r)}{2/\Delta r^2 + 3/(r^2\Delta \theta^2)} \approx 0.636584.$$

因此三个内点近似为

$$\phi(0.6, 0) = 0, \quad \phi\left(0.6, \frac{2\pi}{3}\right) \approx 0.551298, \quad \phi\left(0.6, \frac{4\pi}{3}\right) \approx -0.551298.$$

题 4. 对线性方程组 $Ax = b$, Gauss-Seidel 迭代格式为

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j<i} a_{ij} x_j^{(k+1)} - \sum_{j>i} a_{ij} x_j^{(k)} \right).$$

第一次迭代由 $x^{(0)} = (0, 0, 0)^T$ 得

$$x_1^{(1)} = \frac{15}{4} = 3.75, \quad x_2^{(1)} = \frac{10 + x_1^{(1)}}{4} = 3.4375, \quad x_3^{(1)} = \frac{10 + x_2^{(1)}}{3} \approx 4.479167.$$

第二次迭代为

$$x_1^{(2)} = \frac{15 + x_2^{(1)}}{4} = 4.609375, \quad x_2^{(2)} = \frac{10 + x_1^{(2)} + x_3^{(1)}}{4} \approx 4.772135, \quad x_3^{(2)} = \frac{10 + x_2^{(2)}}{3} \approx 4.924045.$$

所以两次迭代后 $x^{(2)} \approx (4.609375, 4.772135, 4.924045)^T$ 。

题 5. 记边界点 $P_n = \left(\cos \frac{n\pi}{4}, \sin \frac{n\pi}{4}\right)$, $n = 0, \dots, 7$, 内点 $Q_- = (-0.1, 0)$ 、 $Q_+ = (0.1, 0)$ 。一种合适剖分为

$$(Q_+, P_0, P_1), (Q_+, P_1, P_2), (Q_+, P_2, Q_-), (Q_-, P_2, P_3), (Q_-, P_3, P_4), \\ (Q_-, P_4, P_5), (Q_-, P_5, P_6), (Q_-, P_6, Q_+), (Q_+, P_6, P_7), (Q_+, P_7, P_0).$$

这 10 个三角形覆盖圆内多边形近似区域, 且不交叉。计算机中可用三类表存储: 节点坐标表 $\text{coord}[i] = (x_i, y_i)$, 单元连接表 $\text{elem}[e] = (i_e, j_e, k_e)$, 以及边界边表 $\text{bdryEdge} = (P_0P_1, P_1P_2, \dots, P_7P_0)$ 。若需快速查邻接关系, 可再由单元连接表生成“边到单元”的哈希表或每个单元的邻居表。

题 6. 取泛函

$$I[y] = \int_0^1 \left[\frac{1}{2} (y')^2 - \frac{\pi^2}{2} y^2 \right] dx + \frac{5}{2} y(1)^2 + \frac{67}{20} y(1),$$

并将 $y(0) = 1/2$ 作为本质边界条件。对任意满足 $\delta y(0) = 0$ 的变分, 有

$$\delta I = - \int_0^1 (y'' + \pi^2 y) \delta y dx + \left[y'(1) + 5y(1) + \frac{67}{20} \right] \delta y(1).$$

因而极值条件给出

$$y'' + \pi^2 y = 0, \quad y'(1) + 5y(1) + \frac{67}{20} = 0,$$

后者等价于 $y(1) + 0.2y'(1) = -0.67$ 。

对节点 0, 0.2, 0.5, 1 采用分片线性基函数。元素 $[x_i, x_{i+1}]$ 上

$$K^{(e)} = \frac{1}{h} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} - \frac{\pi^2 h}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

组装后把 $y_0 = 0.5$ 移到右端, 并加入右端边界项, 未知量 $(y_1, y_2, y_3)^T$ 满足

$$\begin{pmatrix} 6.688399 & -3.826814 & 0 \\ -3.826814 & 2.701439 & -2.822467 \\ 0 & -2.822467 & 5.355066 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2.664493 \\ 0 \\ -3.35 \end{pmatrix}.$$

解得节点近似值

$$y(0) = 0.5, \quad y(0.2) \approx 0.539789, \quad y(0.5) \approx 0.247160, \quad y(1) \approx -0.495307.$$

题 7. 不指定 H 的具体形式, 只要求它是相邻自旋项之和, 例如程序中提供函数 $H = H(S_1, \dots, S_N)$, $B_i = \partial H / \partial S_i$ 。初始化一维链长 N 、边界条件、参数 α, β 、时间步长 Δt , 并给出满足 $|S_i| = 1$ 的初始自旋构型。每一步先由当前自旋计算局域场 B_i 。为避免方程右端含有 $\dot{S}_i$, 可将

$$\dot{S}_i = \alpha(S_i \times B_i) + \beta(S_i \times \dot{S}_i)$$

改写为

$$\dot{S}_i = \frac{\alpha}{1 + \beta^2} [S_i \times B_i + \beta S_i \times (S_i \times B_i)],$$

其中默认 $|S_i| = 1$ 。用显式四阶 Runge-Kutta、中点法或预测-校正法推进所有自旋, 每步后归一化 $S_i^{n+1} = S_i^{n+1,*} / |S_i^{n+1,*}|$ 以抑制数值误差。重复迭代直到满足平衡判据, 例如 $\max_i |\dot{S}_i| < \varepsilon$ 或 $|H^{n+1} - H^n| < \varepsilon$ 。若要避免陷入局部极小, 可从多组随机初态出发, 或在早期加入退火噪声, 最后取能量最低且满足收敛条件的构型。

题 8. 对边长 $L = 1$ 的周期正方形, 粒子 i 看粒子 j 时的最小像位移为

$$\Delta x = x_j - x_i - \text{round}(x_j - x_i), \quad \Delta y = y_j - y_i - \text{round}(y_j - y_i).$$

若 $\sqrt{\Delta x^2 + \Delta y^2} \leq 0.4$, 则二者相互作用。逐对计算得到有效相互作用对为 $(1, 2), (1, 3), (2, 4)$ 。因此各粒子看到的相互作用对象及其像坐标为:

粒子	截断半径内的其他粒子或像
1	2(0.2, 0.3), 3(0.3, -0.1)
2	1(0.1, 0.1), 4(0.4, 0.4)
3	1(0.1, 1.1)
4	2(0.2, 0.3)
5	无

例如 1 到 3 的最小像位移为 $(0.2, -0.2)$, 距离 $0.2828 < 0.4$; 而 1 到 5 的最小像距离为 $0.4243 > 0.4$, 故不相互作用。

8 2025 年秋季学期 (魏逸丰班)

试卷

题 1. 蒙特卡洛方法

- (1) 取 $a = 5, c = 3, m = 32, x_0 = 6$ 。用线性同余法产生 $[0, 1]$ 区间内均匀分布的伪随机数, 写出产生的前三个数 (ξ_1, ξ_2, ξ_3) 。
- (2) 用 Metropolis 方法对标准正态分布抽样: $f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ 。取起点 $x_0 = 0$, 试探点范围为 $[x_i - \delta, x_i + \delta]$, $\delta = 2$ 。用 (1) 中得到的随机数 ξ_1 产生试探点, 用 ξ_2 判断是否接受。
- (3) 上述哪一步体现 Metropolis 方法可以不考虑分布密度函数的归一化?

题 2. 蒙特卡洛方法

产生服从如下概率分布的随机变量: $p(x) = 6x(1-x), x \in [0, 1]$ 。

- (1) 使用第一类舍选法抽样, 写出步骤, 指出 L 值和舍选效率。
- (2) 用第二类舍选法, 比较函数 $g(x) = 2x, x \in [0, 1]$, 写出 L 值、舍选效率, 并比较两种方法及原因。

题 3. 有限差分法

给定

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = f(x, y), \quad \phi|_{\Gamma} = F(s).$$

取步长 h , 进行二维离散化。

- (1) 推导二阶微商的差分格式, 并给出 0 点位置的五点差分格式。
- (2) 比较有限元法和有限差分法的异同。

题 4. 有限元法

考虑 $\frac{d^2u}{dx^2} = 1$, $0 \leq x \leq 1$, $u(0) = u(1) = 0$, 泛函形式为

$$I[u] = \int_0^1 \left[\frac{1}{2} \left(\frac{du}{dx} \right)^2 + u \right] dx.$$

将区间 $[0, 1]$ 平均划分为 4 个单元, $\Delta x = 1/4$ 。

- (1) 进行元素内线性插值, 写出 $u(x)$ 和 $u'(x)$ 。
- (2) 做元素分析, 写出每个元素上的子泛函表达式。
- (3) 对总泛函求极值, 求解线性方程组, 求节点处 $u(0.25), u(0.5), u(0.75)$ 的值。

题 5. 分子动力学

- (1) 设二维正方形区域尺寸为 10×10 , 分成 100 个小格子, 将五个粒子分别放在 $A(2, 3)$, $B(7, 1)$, $C(5, 6)$, D 考虑周期性边界条件和最小像力约定, 若 $r_c = 3.0$, 计算或画图说明需要考虑哪些粒子间相互作用。
- (2) 利用 Verlet 速度形式求解分子动力学问题的基本流程。
- (3) 沿用第 (1) 问中的区域、粒子坐标、周期性边界条件、最小像力约定和截断距离 $r_c = 3.0$ 。在 Lennard-Jones 势作用下, 粒子 i 受到粒子 j 的相互作用力在 x 方向上的分量为

$$F_{ij,x} = 48 \left(\frac{\epsilon}{\sigma^2} \right) (x_i - x_j) \left[\left(\frac{\sigma}{r_{ij}} \right)^{14} - \frac{1}{2} \left(\frac{\sigma}{r_{ij}} \right)^8 \right],$$

y 方向同理。令 $\epsilon = \sigma = 1$, $m = 1$, 所有粒子初速度均为 0, $\Delta t = 0.1$ 。只保留第 (1) 问判定出的截断半径内相互作用, 用速度 Verlet 形式将体系推进一步, 求一步后 B 粒子的位置和速度。

题 6. 机器学习

- (1) 判断题:
 - (a) 神经网络是非监督学习的一个典型例子。
 - (b) 增加模型复杂性可降低偏差。
 - (c) 梯度下降算法的最适学习率可由代价函数梯度计算。
 - (d) 正则化方法旨在人为干预避免过拟合问题的出现。
 - (e) 决策树分类方案中, 每个节点样本的分类杂度越大越好。
- (2) 画出单层 3 个变量、一个隐藏层 4 个神经元、输出层 2 个 Softmax 变量的神经网络, 写出训练的具体步骤。

题 7. 机器学习

给定数据如下, 要求训练一个以温度 T 预测长度变化 ΔL 的模型。

$T/^{\circ}\text{C}$	20	40	60	80	100	120	140
$\Delta L/\text{mm}$	0.12	0.25	0.37	0.52	0.61	0.78	0.95

- (1) 计算回归树根节点样本方差。

- (2) 以下 3 个分裂点谁最优? A: $T \leq 30^\circ\text{C}$, B: $T \leq 70^\circ\text{C}$, C: $T \leq 110^\circ\text{C}$ 。
- (3) 用最佳分裂点构建第一层, 计算左右两个子节点的预测值。
- (4) 装袋构建样本训练集 (有放回地抽 7 次), 问其作用和优势。
- (5) 为什么集成学习 (随机森林) 通常比单个回归树或决策树表现更好?

参考解答

题 1.

- (1) 线性同余法采用 $x_{n+1} = (5x_n + 3) \bmod 32$, $\xi_n = x_n/32$ 。由 $x_0 = 6$ 得

$$x_1 = 1, \quad x_2 = 8, \quad x_3 = 11,$$

因而

$$(\xi_1, \xi_2, \xi_3) = \left(\frac{1}{32}, \frac{1}{4}, \frac{11}{32}\right).$$

- (2) Metropolis 抽样中, 若用 $y = x_i - \delta + 2\delta\xi$ 产生试探点, 则从 $x_0 = 0$ 出发有

$$y = -2 + 4 \cdot \frac{1}{32} = -\frac{15}{8}.$$

接受概率为

$$A = \min\left\{1, \frac{f(y)}{f(0)}\right\} = \exp\left(-\frac{y^2}{2}\right) = \exp\left(-\frac{225}{128}\right) \approx 0.173.$$

因为 $\xi_2 = 1/4 > A$, 故拒绝试探点, 本步后仍取 $x_1 = 0$ 。

- (3) 归一化常数在比值 $f(y)/f(x_i)$ 中相消, 因此接受率只依赖未归一化密度的比值。这一步体现了 Metropolis 方法不需要知道分布密度的归一化因子。

题 2.

- (1) 第一类舍选法可取 $x \sim U(0, 1)$, $u \sim U(0, L)$, 当 $u \leq p(x)$ 时接受。由于

$$p(x) = 6x(1-x), \quad p_{\max} = p(1/2) = \frac{3}{2},$$

取 $L = 3/2$, 舍选效率为

$$\eta_1 = \frac{\int_0^1 p(x) dx}{L} = \frac{2}{3}.$$

- (2) 第二类舍选法取比较分布 $g(x) = 2x$, 先按 g 抽样, 即 $x = \sqrt{\xi}$, 再按

$$\frac{p(x)}{Lg(x)} = \frac{6x(1-x)}{3 \cdot 2x} = 1-x$$

接受。这里 $L = \max_{0 < x \leq 1} \frac{p(x)}{g(x)} = 3$, $\eta_2 = \frac{1}{L} = \frac{1}{3}$ 。本题中第一类舍选法效率更高; 原因是 $g(x) = 2x$ 在 x 接近 1 处权重较大, 而目标分布在 $x = 1$ 处为零, 比较分布与目标分布形状不够贴合。

题 3.

(1) 对内点 (x_i, y_j) , 二阶中心差分为

$$\frac{\partial^2 \phi}{\partial x^2} \Big|_{i,j} \approx \frac{\phi_{i+1,j} - 2\phi_{i,j} + \phi_{i-1,j}}{h^2}, \quad \frac{\partial^2 \phi}{\partial y^2} \Big|_{i,j} \approx \frac{\phi_{i,j+1} - 2\phi_{i,j} + \phi_{i,j-1}}{h^2}.$$

因而五点差分格式为

$$\frac{\phi_{i+1,j} + \phi_{i-1,j} + \phi_{i,j+1} + \phi_{i,j-1} - 4\phi_{i,j}}{h^2} = f_{i,j}.$$

若把中心点记为 0, 四个相邻点记为 E, W, N, S, 则 $\phi_0 = \frac{\phi_E + \phi_W + \phi_N + \phi_S - h^2 f_0}{4}$, 边界上的点直接由 $\phi|_\Gamma = F(s)$ 给出。

(2) 有限差分法直接在规则网格上离散微分算子, 形式简单、实现方便; 有限元法先构造弱形式和基函数, 可自然处理复杂区域、非均匀网格和复杂边界, 但组装矩阵与积分过程更复杂。

题 4.

(1) 在单元 $[x_i, x_{i+1}]$ 上令 $h = 1/4$, 线性插值为

$$u(x) = N_i(x)u_i + N_{i+1}(x)u_{i+1}, \quad N_i = \frac{x_{i+1} - x}{h}, \quad N_{i+1} = \frac{x - x_i}{h}, \quad u'(x) = \frac{u_{i+1} - u_i}{h}.$$

(2) 元素子泛函为

$$I_i = \int_{x_i}^{x_{i+1}} \left[\frac{1}{2}(u')^2 + u \right] dx = \frac{(u_{i+1} - u_i)^2}{2h} + \frac{h}{2}(u_i + u_{i+1}).$$

(3) 边界条件 $u_0 = u_4 = 0$ 。令未知量为 u_1, u_2, u_3 , 对总泛函求偏导为零, 得到

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} -1/16 \\ -1/16 \\ -1/16 \end{pmatrix}.$$

解得

$$u(0.25) = -\frac{3}{32}, \quad u(0.5) = -\frac{1}{8}, \quad u(0.75) = -\frac{3}{32}.$$

题 5.

(1) 对周期边界, 最小像位移可写为

$$\Delta r_{ij} = r_j - r_i - L \text{round}\left(\frac{r_j - r_i}{L}\right), \quad L = 10.$$

逐对计算距离并与 $r_c = 3$ 比较, 只有

$$B(7, 1) \leftrightarrow E(6, 2), \quad D(8, 4) \leftrightarrow E(6, 2)$$

两对需要考虑相互作用; 其余粒子对的最小像距离均大于 3。因此 A 与 C 无相互作用对象, B 与 E 相互作用, D 与 E 相互作用, E 与 B, D 相互作用。

(2) 速度 Verlet 的基本递推为

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{F(t)}{2m}\Delta t^2, \quad v(t + \Delta t) = v(t) + \frac{F(t) + F(t + \Delta t)}{2m}\Delta t.$$

流程为: 初始化位置和速度; 按周期边界与最小像约定计算力; 更新位置; 重新计算新位置处的力; 更新速度; 进入下一步。

- (3) 沿用第 (1) 问条件, 初始时 B 只与 E 相互作用, D 也与 E 相互作用。对 B 粒子, $r_B - r_E = (1, -1)$, $F_B^{(0)} = (-1.125, 1.125)$ 。因为 $v_B^{(0)} = 0$, $m = 1$, $\Delta t = 0.1$, 位置更新为

$$r_B^{(1)} = r_B^{(0)} + \frac{F_B^{(0)}}{2} \Delta t^2 = (6.994375, 1.005625).$$

为了按全体速度 Verlet 重算新时刻的力, 还需同步更新与 B 力有关的 E 粒子。初始时 $F_D^{(0)} \approx (-0.011673, -0.011673)$, $F_E^{(0)} \approx (1.136673, -1.113327)$, 因而 $r_E^{(1)} \approx (6.005683, 1.994433)$ 。用新位置和截断条件重新计算 B 受到 E 的力, 得 $F_B^{(1)} \approx (-1.189138, 1.189278)$ 。因此速度更新为

$$v_B^{(1)} = v_B^{(0)} + \frac{F_B^{(0)} + F_B^{(1)}}{2} \Delta t \approx (-0.115707, 0.115714).$$

题 6.

- (1) a. $\times$; b. $\checkmark$; c. $\times$; d. $\checkmark$; e. $\times$ 。

理由: 神经网络既可用于监督学习, 也可用于无监督学习, 但普通前馈分类网络不是无监督学习的典型例子; 增加模型复杂度通常会降低偏差但可能增大方差; 梯度下降的梯度给出更新方向, 学习率通常是超参数, 不能简单由梯度自动给出; 正则化用于约束模型复杂度、缓解过拟合; 决策树分类希望划分后节点不纯度降低, 而不是越大越好。

- (2) 单隐藏层网络结构为: 输入层含 3 个变量 x_1, x_2, x_3 , 隐藏层含 4 个神经元 $h_1, \dots, h_4$, 输出层含 2 个 Softmax 变量 $\hat{y}_1, \hat{y}_2$:

$$x \in \mathbb{R}^3, \quad h = \sigma(W_1 x + b_1), \quad W_1 \in \mathbb{R}^{4 \times 3}, \quad z = W_2 h + b_2, \quad W_2 \in \mathbb{R}^{2 \times 4}, \quad \hat{y} = \text{softmax}(z).$$

训练步骤为: 准备带标签样本并初始化 W_1, b_1, W_2, b_2 ; 前向传播计算 $\hat{y}$; 用交叉熵等损失函数 J 衡量误差; 反向传播求各参数梯度; 用梯度下降或其变体更新参数; 在训练集/验证集上重复迭代并评估分类准确率。

题 7.

- (1) 数据均值为

$$\bar{y} = \frac{0.12 + 0.25 + 0.37 + 0.52 + 0.61 + 0.78 + 0.95}{7} \approx 0.514286.$$

回归树中常用的根节点均方差为

$$\frac{1}{7} \sum_{i=1}^7 (y_i - \bar{y})^2 \approx 0.073682.$$

若按无偏样本方差定义, 则结果为 0.085962。

- (2) 三个分裂点的左右节点预测值与平方误差和为

分裂点	$\bar{y}_L$	$\bar{y}_R$	SSE
$T \leq 30^\circ\text{C}$	0.120000	0.580000	0.334400
$T \leq 70^\circ\text{C}$	0.246667	0.715000	0.139767
$T \leq 110^\circ\text{C}$	0.374000	0.865000	0.171370

因此最优分裂点为 $T \leq 70^\circ\text{C}$ 。

- (3) 第一层回归树为：当 $T \leq 70^{\circ}\text{C}$ 时预测 $\Delta L = 0.246667$ ；当 $T > 70^{\circ}\text{C}$ 时预测 $\Delta L = 0.715000$ 。
- (4) 装袋法用有放回抽样构造多个大小为 n 的 bootstrap 训练集，并分别训练基学习器，最后对回归结果取平均。这样做能增加基学习器之间的差异，降低模型方差，并可利用未被抽中的袋外样本估计泛化误差。
- (5) 随机森林进一步在节点分裂时随机选择特征子集，使树之间相关性更低，因此通常比单棵树有更好的泛化能力和更稳定的预测结果。