

2025 春计算物理 B 期末考试

一、蒙特卡洛抽样 (10 分, 每题 5 分)

- 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 ;
- 考虑 Metropolis 方法生成正态分布序列:

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

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.

二、蒙特卡洛抽样 (15 分, 每题 5 分)

考虑分布函数:

$$f(x) = 3x^2, \quad x \in [0, 1].$$

- 直接抽样抽取 $f(x)$ 的样本, 写出抽样步骤或画出流程图;
- 利用第一类舍选法进行抽样, 写出抽样步骤或画出流程图;
- 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, $i = 1, 2, \dots, n$.

三、有限差分法 (10 分, 每题 5 分)

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 $h = \frac{1}{2}$. 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 $x = \frac{1}{2}$ 处数值解与解析解之间的误差. 若步长改为 $h = \frac{1}{3}$, 误差是否有变化? 说明原因.

四、有限元法 (15 分, 每题 5 分)

考虑微分方程:

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

利用有限元素法进行求解, 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 $\Delta x = \frac{1}{4}$.

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 $u(0.25), u(0.5), u(0.75)$.

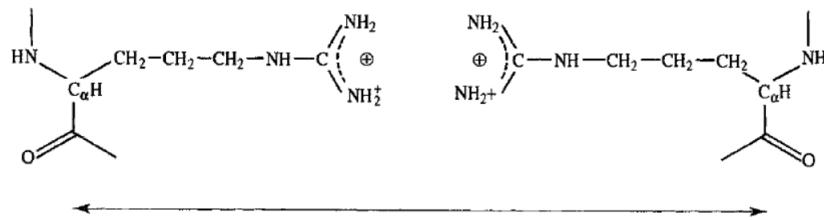
五、分子动力学 (25 分)

- 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；
- 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 $\varepsilon = \sigma = 1$ 。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 方向上的分量可以写为：

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

- 在考虑下面基团的相互作用时，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？



- 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，有各自的优缺点。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。
- 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 1fs？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？

六、机器学习 (25 分)

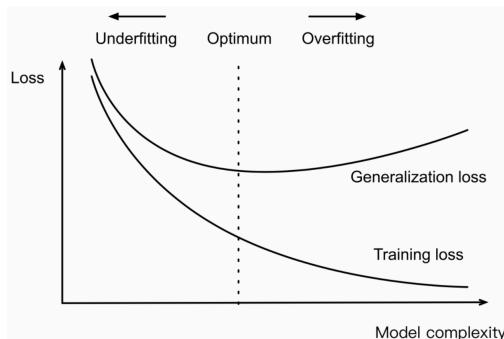
1. 判断题 (5 分)

- (1) 决策树输入数据前要对数据进行归一化/标准化.
- (2) 集成多个线性分类器也无法进行非线性分类.
- (3) Sigmoid 函数容易使梯度消失.
- (4) 回归算法给出连续输出, 分类算法给出离散输出.
- (5) L1 正则化用于特征选择.

2. 填空题 (10 分)

- (1) 写出三个降维的方法: _____、_____、_____.
- (2) BP 预测中误差传递的顺序为 _____ 层、_____ 层、_____ 层.
- (3) (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 _____.
- (4) 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.



注: 原试卷上没有标明 Test Set 和 Training Set, 只标注了 Bias 和 Variance.

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