2024 秋季学期人工智能中的数学编程作业

截止时间 2025 年 1 月 9 号，提交实验报告与源代码（打包）
实验报告包括结果，主要伪码与适当分析过程，建议总体 4 页以内

1. (概率密度变换实践)

   (1) 假设 $u \sim \mathcal{U}[-1,1]$ 为 $[-1,1]$ 上的均匀分布，$\theta \sim \arccos(u), \phi \sim \mathcal{U}[0,2\pi]$ 为 $[0,2\pi]$ 上的均匀分布。请通过以上方法生成 5000 个独立同分布的 u_i，ϕ_i，绘制出对应的

   $$\boldsymbol{\varphi}_i = (\sin\theta_i\cos\phi_i, \sin\theta_i\sin\phi_i, \cos\theta_i)$$

   的图像。

   (2) 从 3 维标准正态分布中产生 5000 个独立样本 $\mathbf{x}_i$，画图 $\mathbf{x}_i/\|\mathbf{x}_i\|$ 的图像与（1）对比。

2. (大数律与中心极限定理验证)

   对于均匀分布 $\mathcal{U}[-1,1]$、指数分布 $\lambda = 1$ 和标准正态分布，分别在样本量 n=1, 5,100 时，重复以下实验 10000 次：

   - 生成样本 $x_1, \ldots, x_n$,
   - 计算样本均值 $\bar{x}_n$,
   - 标准化样本均值 $z = \frac{\bar{x}_n - \mathbb{E}(x)}{\sqrt{\mathrm{Var}(x)/n}}$,

   绘制标准化样本均值 z 的直方图，并与标准正态分布对比，验证中心极限定理。

3. (线性回归实践)

   请考虑如下的线性回归模型：

   $$y = \mathbf{x}^\top \mathbf{w} + \epsilon,$$

   其中 $\mathbf{w} = [2,1], \mathbf{x}_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_1), \mathbf{x}_2 = 1, \epsilon \sim \mathcal{N}(0,1)$。

   (1) 生成 100 对协变量 $\mathbf{x}$ 和对应的响应变量 y。

   (2) 使用最小二乘法来求参数 $\mathbf{w}$ 的估计量 $\hat{\mathbf{w}}$，并绘制拟合数据的图像。

   (3) 重复步骤（1）和（2）1000 次，记录每次估计得到的 $\hat{\mathbf{w}}$。比较这些估计值 $\hat{\mathbf{w}}_1$ 的直方图与其渐近分布 $\mathcal{N}\left(\mathbf{0}, \frac{1}{100}\mathbf{I}_1\right)$ 进行比较。

   (4) 对 $n = 10000$ 的情况重复相同的步骤。$\hat{\mathbf{w}}_1$ 的渐近分布是否提供了更好的拟合?

4. (机器学习实际问题实践)

   使用线性回归在 MNIST 数据集上进行手写数字分类：MNIST 数据集是机器学习领域中最著名的基准数据集之一，包含了大量的手写数字图像（0-9）。每个图像是 28x28 像素的灰度图像，常用

于图像分类任务。虽然线性回归通常用于回归问题（预测连续值），但在本练习中，我们将探索如何将其应用于分类问题（预测离散类别）。

- 加载 MNIST 数据集，仅选择两个数字进行二分类，例如“0”和“1”。
- 将图像数据从二维（28x28）展平成一维（784 维），将像素值标准化到 0 到 1 之间，以提高模型训练的效率，将标签转换为数值格式（0 和 1）。
- 随机划分训练集与测试集（百分 80 用于训练）。
- 使用训练集训练线性回归模型。
- 使用训练好的模型对测试集进行预测，得到连续的预测值。由于线性回归输出的是连续值，需要设定一个阈值（如 0.5）将其转换为类别标签（0 或 1）。
- 比较预测标签与真实标签，计算分类准确率。

(hint: 可以用 tensorflow.keras.datasets.mnist 或者 torchvision.datasets.MNIST 调用 MNIST 数据集，或寻找官网下载)

5. (谱图理论实践)

   对于五十阶完全图 K_{50} 与环形图 C_{50}，进行以下计算：

   (1) 计算对应的邻接矩阵和拉普拉斯矩阵，而后计算拉普拉斯矩阵的特征值。

   (2) 对于五十阶图的第 i 个顶点赋值为 $v_i^0 = x_i$，$\mathbf{x}$ 服从 50 维标准正态分布。而后执行迭代平均过程，其中每个顶点在每一步迭代中将其当前值更新为自身与所有相邻顶点值的平均值。请编程比较五十阶完全图 K_{50} 与环形图 C_{50} 的迭代平均过程满足 $\sum_{i=1}^{50} \left|v_i^t - v_i^{t-1}\right| \leq 0.000001$ 所需要的迭代步数 t。

   (3) 重复 30 次实验，在平均意义下比较两种图的收敛速度。

6. (规律探索) 假如一个矩阵是随机的，那么随着矩阵的扩大，其特征值个数增多，特征值的分布会不会出现一些规律?

   Wigner 半圆定律（Wigner’s Semicircle Law）是随机矩阵理论中的一个基本定理，描述了大规模对称随机矩阵的特征值分布在特定条件下的统计规律。该定律由匈牙利物理学家 Eugene Wigner 在 1950 年代提出，最初用于研究原子核中的能级分布，现在已广泛应用于数学、物理、统计学、金融学等多个领域。我们称一个对称矩阵 $\mathbf{M} \in \mathbb{R}^{d\times d}$ 为 Wigner 矩阵，如果其满足以下性质：

   - $\mathbf{M} = \mathbf{M}^\top$，且对于 $1 \leq i \leq j \leq d$，$\mathbf{M}_{ij}$ 都是独立的，
   - 对于 $1 \leq i < j \leq d$，$\mathbf{M}_{ij} \overset{\text{i.i.d}}{\sim} \mathbf{y}$，并且满足 $\mathbb{E}(\mathbf{y}) = 0$，$\mathrm{Var}(\mathbf{y}) = 1$，
   - 对于 $1 \leq i \leq d$，$\mathbf{M}_{ii} \overset{\text{i.i.d}}{\sim} \mathbf{z}$，并且满足 $\mathbb{E}(\mathbf{z}) = 0$，$\mathrm{Var}(\mathbf{z}) < +\infty$。

   Wigner 半圆定律表明 $\frac{1}{\sqrt{d}}\mathbf{M}$ 的特征值的联合分布在 $d \to +\infty$ 时几乎必然收敛于 Wigner 半圆分布：

   $$p(x) = \frac{1}{2\pi}\sqrt{4-x^2}, |x| \leq 2.$$

   (1) 请生成 1 个 $\mathbb{R}^{1000\times 1000}$ 的 Wigner 矩阵，其每个元素服从标准正态分布（上半矩阵元素独立生成，下半部分对称），而后画出其特征值的直方图并与 Wigner 半圆分布做对比。

   (2) 若矩阵元素改为中心在 0 点的均匀分布，画出特征值的直方图并与 Wigner 半圆分布做对比。