

## 第三次 课后作业

---

提交截止时间：2022 年 11 月 27 日 20: 00

**问题一：**（20 分）线性回归最小二乘法获得最优参数值的必要条件证明：预测误差为**零均值**且**与输入样本无关**。试说明“零均值”和“无关”的根据是什么？

**问题二：**训练样本集合  $\{(x_i, y_i)\}_{i=1}^{10}$  由 10 个样本组成，其中  $y_i = -\sin(x_i/5) + \cos(x_i) + \epsilon$ ,  $x_i \sim \mathcal{U}[-5, 5]$ , 其中  $\epsilon \sim \mathcal{N}(0, 0.04)$ 。

编程解决下述问题：

- 1) （10 分）绘出训练样本散布图和  $y_i = f(x_i)$  曲线；
- 2) （30 分）实现解决线性回归问题的最小二乘法和梯度下降法；
- 3) （10 分）使用 2) 中实现的两种方法求解训练样本集的线性回归方程，并给出训练样本集合的均方误差；
- 4) （10 分）解释 3) 中你所得到的结果，并对两种求解方法进行比较；
- 5) （20 分）使用 2) 中实现的最小二乘法解决训练数据集多项式回归问题的最优解。