

## 第四讲 课后作业

---

1. (20)在讲稿第 9 页最下面的“注意”中指出：令两个偏导数为零获得了最优参数值的必要条件，它们意味着预测误差是零均值且与输入不相关。

试说明“零均值”和“不相关”的根据是什么？

2. 线性回归问题：假设存在直线方程  $y = 1.4x + 0.9$ ，噪声服从  $\mathcal{N}(0, 5)$ 。
  - 1) (10)生成满足条件的 100 个独立同分布样本，选择其中 80%作为训练样本，剩余为测试样本。
  - 2) (10)绘出训练样本和测试样本的散布图。
  - 3) (30)编程实现求解线性回归的最小二乘法和采用 Huber 损失函数的鲁棒线性回归方法。
  - 4) (10)使用 3)中实现的线性回归方法求解训练样本的线性回归方程，并给出测试样本的均方差。
  - 5) (20)解释 4)中你所得到的结果，并对两种求解方法进行比较。