

# 数据科学与工程数学基础

## 作业提交规范及第 16 次作业

教师：黄定江

助教：陈诺、刘文辉

2022 年 5 月 8 日

### 作业提交规范

1. 作业提交形式：**练习本或笔记本**（建议统一使用一般的**练习本**即可，不接收以纸张的方式书写的作业）。
2. 作业书写说明：
  - (a) 可以讨论，**禁止抄袭！**
  - (b) 练习本封面至少包含两方面信息：**姓名和学号**
  - (c) 每一次的作业**请另起一页**，并在**第一行标明第几次作业**。例如“第 16 次作业”；
  - (d) 每一题请**标注题号**，无需抄题，直接解答；
  - (e) 题与题之间**请空一行**；
  - (f) 不要求字好，但要求书写整体清晰易读。
3. 作业提交途径：纸质作业交给**学习委员**，由学习委员**按学号顺序**收齐后统一在截止日期前交到**助教实验室**。**单数周**布置的作业交到助教刘文辉处**数学馆西 109**；**双数周**布置的作业交到助教陈诺处**地理馆 353**。
4. 作业评分说明：正常提交作业的按照实际评分记录；逾期补交作业的根据逾期情况在实际评分基础上酌情扣分；**未交作业的当次作业记为 0 分**。

### 第 16 次作业



提交截至时间：**暂定 2022/04/29 周五 20:00（晚上）**

## 理论部分

**习题 1.** 求激活函数  $\sigma(x) = \frac{1}{1+e^{-x}}$  的导数。

**解.**  $\sigma'(x) = \frac{e^{-x}}{(1+e^{-x})^2} = \sigma(x)(1 - \sigma(x))$

**习题 2.** 构建模型使得预测值与真实值的误差最小常用向量 2-范数度量, 求解模型过程中需要计算梯度, 求梯度:

- $f(A) = \frac{1}{2}\|Ax + b - y\|_2^2$ , 求  $\frac{\partial f}{\partial A}$
- $f(x) = \frac{1}{2}\|Ax + b - y\|_2^2$ , 求  $\frac{\partial f}{\partial x}$

其中  $A \in R^{m \times n}$ ,  $x \in R^n$ ,  $b, y \in R^m$

**解.**

$$\begin{aligned}\frac{\partial}{\partial A} f &= \frac{\partial}{\partial A} \frac{1}{2} (x^T A^T A x + 2(b - y)^T A x + (b - y)^T (b - y)) \\ &= \frac{\partial}{\partial A} \frac{1}{2} (x^T A^T A x + 2(b - y)^T A x) \\ &= A x x^T + (b - y) x^T\end{aligned}$$

$$\frac{\partial}{\partial x} f = A^T A x + A^T (b - y)$$

**习题 3.** 二次型是数据分析中常用函数, 求  $\frac{\partial x^T A x}{\partial x}$ ,  $\frac{\partial x^T A x}{\partial A}$ , 其中  $A \in R^{m \times m}$ ,  $x \in R^m$

**解.**  $\frac{\partial x^T A x}{\partial x} = (A + A^T)x$

$$\frac{\partial x^T A x}{\partial A}_{ij} = x_i x_j, \frac{\partial x^T A x}{\partial A} = x x^T$$

**习题 4.** 利用迹微分法求解  $\frac{\partial \text{Tr}(W^{-1})}{\partial W}$ , 其中  $W \in R^{m \times m}$

**解.** 因为

$$0 = dI = d(WW^{-1}) = dWW^{-1} + WdW^{-1}$$

$$WdW^{-1} = -dWW^{-1}$$

$$dW^{-1} = -W^{-1}dWW^{-1}$$

所以

$$\begin{aligned}d\text{Tr}(W^{-1}) &= \text{Tr}(dW^{-1}) \\ &= \text{Tr}(-W^{-1}dWW^{-1}) \\ &= \text{Tr}(-(W^{-1})^2 dW)\end{aligned}$$

即

$$\frac{\partial \text{Tr}(W^{-1})}{\partial W} = -(W^{-T})^2$$