

# 课后练习1

-

2023 年 9 月 22 日

## 1 问题一

### 1.1

$$\text{令 } x = \begin{bmatrix} 0 & 1 & 1 \\ 2 & 2 & 1 \\ 3 & 4 & 1 \end{bmatrix}, y = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}, \hat{w} = \begin{bmatrix} w_1 \\ w_2 \\ b \end{bmatrix}, \text{ 可以计算 } x^T x = \begin{bmatrix} 13 & 16 & 5 \\ 16 & 21 & 7 \\ 5 & 7 & 3 \end{bmatrix}$$

$$\text{那么 } \hat{w} = (x^T x)^{-1} x^T y = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \text{ 故 } w_1 = 1, w_2 = 0, b = 1$$

### 1.2

$$\text{此时 } x = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 2 & 2 & 4 & 1 \\ 3 & 4 & 8 & 1 \end{bmatrix}, \hat{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ b \end{bmatrix}, \text{ 计算 } x^T x = \begin{bmatrix} 13 & 16 & 32 & 5 \\ 16 & 21 & 42 & 7 \\ 32 & 42 & 84 & 14 \\ 5 & 7 & 14 & 3 \end{bmatrix}, \text{ 可以发现这是一}$$

个奇异矩阵(有重复的行), 没有逆矩阵, 因此此时线性回归没有唯一解

### 1.3

$$\text{已知 } \lambda = 1, \text{ 故 } x^T x + \lambda I = \begin{bmatrix} 14 & 16 & 32 & 5 \\ 16 & 22 & 42 & 7 \\ 32 & 42 & 85 & 14 \\ 5 & 7 & 14 & 4 \end{bmatrix}$$

$$\text{那么 } \hat{w} = (x^T x + \lambda I)^{-1} x^T y = \begin{bmatrix} 0.378 \\ 0.140 \\ 0.280 \\ 0.304 \end{bmatrix}, \text{ 因此可得 } w_1 = 0.378, w_2 = 0.140, w_3 =$$

0.280,  $b = 0.304$

## 2 问题二

### 2.1

令  $X = \begin{bmatrix} x_1^T, 1 \\ \vdots \\ x_n^T, 1 \end{bmatrix}$ ,  $R = \begin{bmatrix} r_1 & & \\ & \ddots & \\ & & r_n \end{bmatrix}$ , 则我们可以将损失函数写成矩阵的形式:

$$L(\hat{w}) = (y - X\hat{w})^T R (y - X\hat{w}) + \lambda \hat{w}^T \hat{w}$$

计算偏导数:

$$\begin{aligned} \frac{\partial L(\hat{w})}{\partial \hat{w}} &= \frac{y^T R y - y^T R X \hat{w} - \hat{w}^T X^T R y + \hat{w}^T X^T R X \hat{w} + \lambda \hat{w}^T \hat{w}}{\partial \hat{w}} \\ &= \frac{y^T R y - 2y^T R X \hat{w} + \hat{w}^T X^T R X \hat{w} + \lambda \hat{w}^T \hat{w}}{\partial \hat{w}} \\ &= -2X^T R y + 2X^T R X \hat{w} + 2\lambda I \hat{w} \\ &= 0 \end{aligned}$$

解得:

$$\hat{w} = (X^T R X + \lambda I)^{-1} X^T R y$$

### 2.2

通过人为加入权重可以对不同的数据的重要性进行评估和加权, 能够更准确地进行评估和预测, 这样更重要的数据权重更大, 能更好地反映实际情况

加权计算和直接将  $(x_i, y_i)$  复制  $r_i$  次的结果不一样, 因为复制数据后  $X$  矩阵会有重复行, 虽然有L2正则化, 但是矩阵也和加权计算的不一样, 最后结果也会不一样

## 3 问题三

### 3.1

## 4 问题四

### 4.1