

习题 2-2.

i) $r^{(n)}$ 的作用: 把每个样本对参数影响的程度进行加权, 让更重要的样本对模型产生更大的影响.

ii) 求 \vec{w}^*

令 $R = \text{diag}(r^{(1)}, r^{(2)}, \dots, r^{(n)})$

则 $R(\vec{w}) = \frac{1}{2} R \|\vec{y} - X^T \vec{w}\|^2$

\vec{y} 表示标签组成的列向量

X 是特征矩阵

$$\frac{\partial R(\vec{w})}{\partial \vec{w}} = -X R (\vec{y} - X^T \vec{w})$$

$$\text{令 } \frac{\partial R(\vec{w})}{\partial \vec{w}} = 0$$

$$\text{解得 } \vec{w}^* = (X R X^T)^{-1} X R \vec{y}$$

习题 2-4

$$R(\vec{w}) = \frac{1}{2} \|\vec{y} - X^T \vec{w}\|^2 + \frac{\lambda}{2} \|\vec{w}\|^2$$

$$\frac{\partial R(\vec{w})}{\partial \vec{w}} = -X (\vec{y} - X^T \vec{w}) + \lambda I \vec{w}$$

$$\text{令 } \frac{\partial R(\vec{w})}{\partial \vec{w}} = 0$$

$$\text{则 } \lambda I \vec{w} = X \vec{y} - X X^T \vec{w}$$

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