

第八章作业 王晨曦

Created @April 8, 2022 11:45 PM

姓名：王晨曦

班级：计算机96

学号：2196123413

得分：

1、给定一个损失函数如下所示

$$L(\mathbf{x}) = (\mathbf{x} - \mathbf{m}_1)^T \Lambda_1 (\mathbf{x} - \mathbf{m}_1) + (\mathbf{x} - \mathbf{m}_2)^T \Lambda_2 (\mathbf{x} - \mathbf{m}_2)$$

上式中 \mathbf{x} , \mathbf{m}_1 和 \mathbf{m}_2 均为 $n \times 1$ 的向量， Λ_1 和 Λ_2 均为 $n \times n$ 的实对称矩阵，且为正定的矩阵，试求：

(1)、 $\frac{\partial L}{\partial \mathbf{x}}$ 的形式

(2)、使得 $L(\mathbf{x})$ 最小的 \mathbf{x} 的形式

$$\begin{aligned} 1. (1) \quad \frac{\partial L}{\partial \mathbf{x}} &= (\mathbf{x} - \mathbf{m}_1)^T \Lambda_1 + (\Lambda_1 (\mathbf{x} - \mathbf{m}_1))^T + (\mathbf{x} - \mathbf{m}_2)^T \Lambda_2 + (\Lambda_2 (\mathbf{x} - \mathbf{m}_2))^T \\ &= (\mathbf{x} - \mathbf{m}_1) (\Lambda_1 + \Lambda_1^T) + (\mathbf{x} - \mathbf{m}_2) (\Lambda_2 + \Lambda_2^T) \\ &= 2\mathbf{x} (\Lambda_1 + \Lambda_2) - 2\mathbf{m}_1 \Lambda_1 - 2\mathbf{m}_2 \Lambda_2 \end{aligned}$$

$$\begin{aligned} (2) \quad \because \frac{\partial L}{\partial \mathbf{x}} &= 0 \\ \therefore 2\mathbf{x} (\Lambda_1 + \Lambda_2) - 2\mathbf{m}_1 \Lambda_1 - 2\mathbf{m}_2 \Lambda_2 &= 0 \\ \therefore \mathbf{x} &= (\mathbf{m}_1 \Lambda_1 + \mathbf{m}_2 \Lambda_2) (\Lambda_1 + \Lambda_2)^{-1} \end{aligned}$$

2、如果给定一组样本 $\{(\mathbf{x}_i, y_i) | i = 1, \dots, N\}$ ，其中 $y_i \in \{0, 1\}$ ，现在我们采用逻辑函数 $\sigma(t) = \frac{1}{1 + e^{-t}}$ 来对输出 y 进行逼近，其中 $t = \mathbf{x}_i \cdot \mathbf{w} + b$ ，损失函数如下

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N (y_i - \sigma(\mathbf{x}_i \cdot \mathbf{w} + b))^2$$

1) 尝试推导损失函数相对于权重 \mathbf{w} 的偏导数

2) 尝试推导损失函数相对于偏置 b 的偏导数

Handwritten derivation of the partial derivatives of the loss function \mathcal{L} with respect to weight w and bias b .

(1)
$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial w} &= \frac{1}{N} \sum_{i=1}^N \frac{\partial}{\partial w} (y_i - \sigma(x_i w + b))^2 \\ &= -\frac{1}{N} \sum_{i=1}^N 2(y_i - \sigma(x_i w + b)) \frac{\partial \sigma}{\partial t} \frac{\partial t}{\partial w} \\ &= \frac{1}{N} \sum_{i=1}^N 2x_i (y_i - \sigma(x_i w + b)) \frac{e^{-x_i w - b}}{(e^{-x_i w - b} - 1)^2} \end{aligned}$$

(2)
$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial b} &= \frac{1}{N} \sum_{i=1}^N \frac{\partial}{\partial b} (y_i - \sigma(x_i w + b))^2 \\ &= -\frac{1}{N} \sum_{i=1}^N 2(y_i - \sigma(x_i w + b)) \frac{\partial \sigma}{\partial t} \frac{\partial t}{\partial b} \\ &= \frac{1}{N} \sum_{i=1}^N 2(y_i - \sigma(x_i w + b)) \frac{e^{-x_i w - b}}{(e^{-x_i w - b} - 1)^2} \end{aligned}$$