

练习题

1. 考虑线性决策函数 $g(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b$, 令 $\Pi: g(\mathbf{x}) = 0$ 则获得决策超平面。请计算任意一个样本点 \mathbf{x} 到决策超平面 $\Pi: g(\mathbf{x}) = 0$ 的距离, 并结合所得到的计算结果揭示 $g(\mathbf{x})$ 和 b 的几何含义。

2. 通过计算验证两种方式所推导的逻辑斯蒂回归最大似然估计的等价性。(提示: 计算对数似然度)

3. 请在给定数据集 Toydata.mat 上完成, 训练感知器:

- (1) 给出不同初始值或不同提交样本顺序时, 所得到的感知器权值(算法收敛时), 并尝试画出决策面。
- (2) 重复(1)中的结果 10 次, 把所得到的 10 个感知器的权值取平均, 分别画出由平均权值向量所对应的决策面和 10 个感知器的决策面。

4. 从 2 个二维高斯分布中分别随机选择 100 个样本, 假设高斯分布的参数如下:

$$\boldsymbol{\mu}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \Sigma_1 = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix} \text{ 和 } \boldsymbol{\mu}_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \Sigma_2 = \begin{pmatrix} 1 & -0.8 \\ -0.8 & 2 \end{pmatrix}$$

分别对应于类别 C1 和类别 C2。请完成下列任务:

- (1) 生成两组数据, 显示数据点的散布图;
- (2) 编程实现感知器算法, 在数据散布图中把感知器算法所得到的决策超平面画出;
- (3) 编程实现逻辑斯蒂回归, 在数据散布图中把逻辑斯蒂回归得到的决策结果展示出来;
- (4) 在训练样本中随机添加几个离群点(outliers), 请观察离群值对感知器、线性回归和逻辑斯蒂回归的影响。

5. 在 TwoMoon 数据集上, 完成下列任务:

- (1) 显示数据点的散布图;
- (2) 编程实现感知器算法, 在数据散布图中把感知器算法所得到的决策超平面画出;
- (3) 编程实现逻辑斯蒂回归, 在数据散布图中把逻辑斯蒂回归得到的决策结果展示出来。