

CNY520501: 模式识别

习题 2 答案

2018 年 5 月 25 日

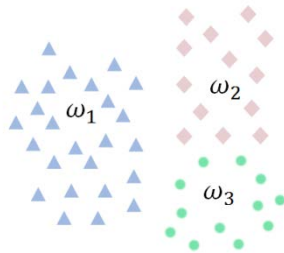
1. 什么是最小距离分类器? 分别计算下列两个模式类的平均样本:

$$\omega_1: X_1 = (-2, 2)^T, X_2 = (2, -2)^T; \omega_2: X_3 = (1, 1)^T, X_4 = (1, 2)^T$$

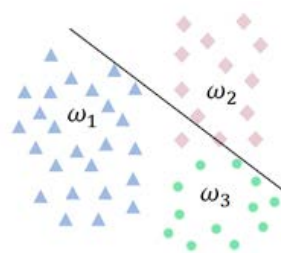
最小距离分类器: 利用两个模式在特征空间中的距离作为两者之间的相似性度量, 由此设计的度量函数就是最小距离分类器

$$\text{平均样本: } \omega_1: (0, 0)^T; \omega_2: (1, 1.5)^T$$

2. 下图所示的样本集合

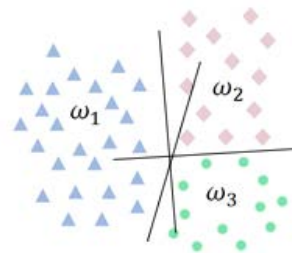


是否总体线性可分?



不是总体线性可分

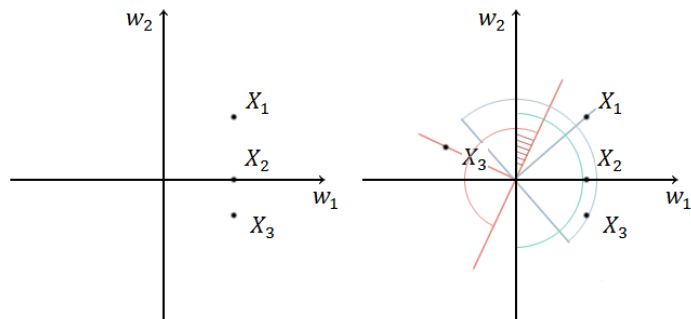
是否成对线性可分?



是成对线性可分

3. 画出下图中两个模式类的线性分类器 $G(X) = W^T X$ 的解区

$$\omega_1: X_1 = (1, 1)^T, X_2 = (1, 0)^T; \omega_2: X_3 = \left(1, -\frac{1}{\sqrt{3}}\right)^T$$



4. 对以下样本集合，使用感知器算法求解线性分类器（取初始 W 为全 0 向量，取 $\rho = 1$ ），写出迭代过程。

$$\omega_1: X_1 = (1, 0)^T; \omega_2: X_2 = (-1, 0)^T, X_3 = (0, 1)^T$$

- 模式类的增广形式: $\omega_1: X_1 = (1, 0, 1)^T; \omega_2: X_2 = (-1, 0, 1)^T, X_3 = (0, 1, 1)^T$
- $k = 0, g(X_1) = 0; \quad W(1) = W(0) + \rho X_1 = (1, 0, 1)^T; \quad N_c = 0$
- $k = 1, g(X_2) = 0; \quad W(2) = W(1) - \rho X_2 = (2, 0, 0)^T; \quad N_c = 0$
- $k = 2, g(X_3) = 0; \quad W(3) = W(2) - \rho X_3 = (2, -1, -1)^T; \quad N_c = 0$
- $k = 3, g(X_1) = 1; \quad W(4) = W(3) = (2, -1, -1)^T; \quad N_c = 1$
- $k = 4, g(X_2) = -3; \quad W(5) = W(4); \quad = (2, -1, -1)^T \quad N_c = 2$
- $k = 5, g(X_3) = -2; \quad W(6) = W(5); \quad = (2, -1, -1)^T \quad N_c = 3$

注:

1. 增广形式
2. 迭代终止的条件为 $N_c =$ 样本个数
3. 每次迭代只输入一个样本

5. 对以下样本集合，使用感知器算法求解线性分类器并保证不存在不确定区域（取初始 W 为全 0 向量，取 $\rho = 1$ ），写出迭代过程。

$$\omega_1: X_1 = (1, 1)^T; \omega_2: X_2 = (-2, 1)^T; \omega_3: X_3 = (2, -2)^T$$

- 初始化

模式类的增广形式: $\omega_1: X_1 = (1, 1, 1)^T; \omega_2: X_2 = (-2, 1, 1)^T; \omega_3: X_3 = (2, -2, 1)^T$

初始化三个决策函数 $g_1(X) = W_1^T X, g_2(X) = W_2^T X, g_3(X) = W_3^T X$

- $k = 0, g_1(X_1) = 0, g_2(X_1) = 0, g_3(X_1) = 0, N_c = 0$

$$\begin{cases} W_1(1) = W_1(0) + \rho X_1 = (1, 1, 1)^T \\ W_2(1) = W_2(0) - \rho X_1 = (-1, -1, -1)^T \\ W_3(1) = W_3(0) - \rho X_1 = (-1, -1, -1)^T \end{cases}$$
- $k = 1, g_1(X_2) = 0, g_2(X_2) = 0, g_3(X_2) = 0, N_c = 0$

$$\begin{cases} W_1(2) = W_1(1) - \rho X_2 = (3, 0, 0)^T \\ W_2(2) = W_2(1) + \rho X_2 = (-3, 0, 0)^T \\ W_3(2) = W_3(1) - \rho X_2 = (1, -2, -2)^T \end{cases}$$
- $k = 2, g_1(X_3) = 6, g_2(X_3) = -6, g_3(X_3) = 4, N_c = 0$

$$\begin{cases} W_1(3) = W_1(2) - \rho X_3 = (1, 2, -1)^T \\ W_2(3) = W_2(2) = (-3, 0, 0)^T \\ W_3(3) = W_3(2) = (1, -2, -2)^T \end{cases}$$
- $k = 3, g_1(X_1) = 2, g_2(X_1) = -3, g_3(X_1) = -3, N_c = 1$

$$\begin{cases} W_1(4) = W_1(3) = (1, 2, -1)^T \\ W_2(4) = W_2(3) = (-3, 0, 0)^T \\ W_3(4) = W_3(3) = (1, -2, -2)^T \end{cases}$$
- $k = 4, g_1(X_1) = -1, g_2(X_1) = 6, g_3(X_1) = -6, N_c = 2$

$$\begin{cases} W_1(5) = W_1(4) = (1, 2, -1)^T \\ W_2(5) = W_2(4) = (-3, 0, 0)^T \\ W_3(5) = W_3(4) = (1, -2, -2)^T \end{cases}$$
- $k = 5, g_1(X_1) = -3, g_2(X_1) = -6, g_3(X_1) = 4, N_c = 3$

$$\begin{cases} W_1(5) = W_1(4) = (1, 2, -1)^T \\ W_2(5) = W_2(4) = (-3, 0, 0)^T \\ W_3(5) = W_3(4) = (1, -2, -2)^T \end{cases}$$

- 无不确定区域的解:

$$\begin{cases} G_{12}(X) = g_1(X) - g_2(X) = (4, 2, -1)^T \\ G_{13}(X) = g_1(X) - g_3(X) = (0, 4, 1)^T \\ G_{23}(X) = g_2(X) - g_3(X) = (-4, 2, 2)^T \end{cases}$$

注:

1. 增广形式
2. 注意不确定区域的消除方法

6. 对于二维线性判别函数 $g(X) = 4x_1 - 3x_2 + 5$

- a) 将判别函数写成矩阵形式 $g(x) = W^T X + w_{n+1}$

$$W = [4, -3]^T, \quad X = [x_1, x_2]^T, \quad w_{n+1} = 5, \quad g(X) = [4, -3] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + 5$$

- b) 映射成广义线性函数 $f(Y) = W^T Y, Y = (y_1, y_2, y_3)^T = (2x_1, x_2, 1)^T$

$$W = [2, -3, 5]^T, \quad f(Y) = [2, -3, 5] \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

7. 用广义线性判别函数法解决以下两个模式类的分类:

$$\omega_1: X_1 = (1, 0)^T, X_2 = (-1, 0)^T; \quad \omega_2: X_3 = (0, 1)^T, X_4 = (0, -1)^T$$

- a) 设计变换函数 $y_i = f_i(x_1, x_2), i = 1, 2, \dots$ 使变换后的样本在 Y 空间线性可分。

$$Y = [y_1, y_2]^T = [x_1, x_2^2]^T$$

- b) 给出一个决策面函数 $g(Y)$

$$g(Y) = [0, 1] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - 0.5$$

注:

1. 不少同学写 $g(Y) = \frac{1}{2}$, 注意 $g(Y)$ 是关于 Y 的函数, 正确应为 $g(Y) = y_1 - 1/2$
2. 本题解法不唯一

8. 若准则函数的形式为 $J(W, X) = \frac{1}{2}(1 - W^T X)^2$. $W = (2, 1)^T, X_1 = (2, 3)^T, X_2 = (1, 4)^T$, 在用梯度下降法时, 求 $\nabla J(W, X_1)$ 和 $\nabla J(W, X_2)$

$$\nabla J(W, X) = \frac{\partial J}{\partial W} = (1 - W^T X)(-X)$$

$$\nabla J(W, X_1) = (1 - W^T X_1)(-X_1) = (-6)[-2, -3]^T = [12, 18]^T$$

$$\nabla J(W, X_2) = (1 - W^T X_2)(-X_2) = (-5)[-1, -4]^T = [5, 20]^T$$

注: 很多同学误写为 $\nabla J(W, X) = (1 - W^T X)(-W)$

9. 利用位势法对以下模式类进行分类, 位势函数选 $K(X, X_n) = \exp(-||x - x_n||^2)$, $K_{A,0}(x) = 0$

$$\omega_1: X_1 = (0,0)^T, X_2 = (1,2)^T; \omega_2: X_3 = (1,-1)^T, X_4 = (3,0)^T$$

- $k = 0$, $X' = X_1 = (0,0)^T$, $K_{A,0}(X') = 0$, $N_c = 0$

因为 $X' \in \omega_1$, 且 $K_{A,0}(X') \leq 0$

$$K_{A,1}(X) = K_{A,0}(X) + K(X, X') = \exp\{-||X||^2\}$$

- $k = 1$, $X' = X_2 = (1,2)^T$, $K_{A,1}(X') = \exp\{-1\} > 0$, $N_c = 1$

因为 $X' \in \omega_1$, 且 $K_{A,0}(X') > 0$

$$K_{A,2}(X) = K_{A,1}(X) = \exp\{-||X||^2\}$$

- $k = 2$, $X' = X_3 = (1,-1)^T$, $K_{A,2}(X') = \exp\{-2\} > 0$, $N_c = 0$

因为 $X' \in \omega_2$, 且 $K_{A,0}(X') > 0$

$$K_{A,3}(X) = K_{A,2}(X) - K(X, X') = \exp\{-||X||^2\} - \exp\{-||X - (1,-1)^T||^2\}$$

- $k = 3$, $X' = X_4 = (3,0)^T$, $K_{A,3}(X') = \exp\{-9\} - \exp\{-5\} < 0$, $N_c = 1$

因为 $X' \in \omega_2$, 且 $K_{A,0}(X') < 0$

$$K_{A,4}(X) = K_{A,3}(X) = \exp\{-||X||^2\} - \exp\{-||X - (1,-1)^T||^2\}$$

- $k = 4$, $X' = X_1 = (0,0)^T$, $K_{A,4}(X') = 1 - \exp\{-2\} > 0$, $N_c = 2$

因为 $X' \in \omega_1$, 且 $K_{A,0}(X') > 0$

$$K_{A,5}(X) = K_{A,4}(X) = \exp\{-||X||^2\} - \exp\{-||X - (1,-1)^T||^2\}$$

- $k = 5$, $X' = X_2 = (1,2)^T$, $K_{A,5}(X') = \exp\{-5\} - \exp\{-9\} > 0$, $N_c = 3$

因为 $X' \in \omega_1$, 且 $K_{A,0}(X') > 0$

$$K_{A,6}(X) = K_{A,5}(X) = \exp\{-||X||^2\} - \exp\{-||X - (1,-1)^T||^2\}$$

- $k = 6$, $X' = X_3 = (1,-1)^T$, $K_{A,6}(X') = \exp\{-2\} - 1 < 0$, $N_c = 4$

$\therefore N_c = 4$ \therefore 迭代结束

- 最终位势函数为:

$$\exp\{-||X||^2\} - \exp\{-||X - (1,-1)^T||^2\}$$

10. 用 Fisher 线性判别法对以下模式类构造分类器, 确定最佳投影方向

$$\omega_1: X_1 = (0,0)^T, X_2 = (1,2)^T; \omega_2: X_3 = (1,-1)^T, X_4 = (3,0)^T$$

- 计算两个样本的均值向量:

$$m_x^1 = \frac{1}{n_1} \sum_{X_k \in \omega_1} X_k = [0.5, 1]^T$$

$$m_x^2 = \frac{1}{n_2} \sum_{X_k \in \omega_2} X_k = [2, -0.5]^T$$

- 计算类内总离散度矩阵:

$$\begin{aligned} S_w &= \sum_{X_k \in \omega_1} (X_k - m_x^1)(X_k - m_x^1)^T + \sum_{X_k \in \omega_2} (X_k - m_x^2)(X_k - m_x^2)^T \\ &= \begin{bmatrix} 0.5 & 1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ 1 & 0.5 \end{bmatrix} = \begin{bmatrix} 2.5 & 2 \\ 2 & 2.5 \end{bmatrix} \end{aligned}$$

- 计算 S_w 的逆矩阵 S_w^{-1} :

$$S_w^{-1} = \begin{bmatrix} \frac{10}{9} & -\frac{8}{9} \\ -\frac{9}{9} & \frac{10}{9} \end{bmatrix}$$

- 最佳投影:

$$\begin{aligned} W^* &= S_w^{-1}(m_x^1 - m_x^2) \\ &= [-3, 3]^T \end{aligned}$$

注: 有疑问或习题有错误请联系助教
陈纪伟: cjwbdw6@mail.ustc.edu.cn