部分作业

解答

三、判别函数分类法

3.5 已知两类训练样本为

ω_1： $[0,0,0]^{\mathrm{T}}$， $[1,0,0]^{\mathrm{T}}$， $[1,0,1]^{\mathrm{T}}$， $[1,1,0]^{\mathrm{T}}$

ω_2： $[0,0,1]^{\mathrm{T}}$， $[0,1,1]^{\mathrm{T}}$， $[0,1,0]^{\mathrm{T}}$， $[1,1,1]^{\mathrm{T}}$

设 $\boldsymbol{W}(1)=[-1,-2,-2,0]^{\mathrm{T}}$，用感知器算法求解判别函数，并绘出判别界面。

解答：感知器算法为：

$$\boldsymbol{W}(k+1)=\begin{cases}\boldsymbol{W}(k), & 若\boldsymbol{W}^{\mathrm{T}}(k)\boldsymbol{X}_i>0\\ \boldsymbol{W}(k)+c\boldsymbol{X}_i, & 若\boldsymbol{W}^{\mathrm{T}}(k)\boldsymbol{X}_i\le 0\end{cases}，\quad c为正的校正增量。$$

首先，将所有样本写成增广向量的形式并编号，属于 ω_2 的样本乘以(-1)：

$$\boldsymbol{X}_1=[0,0,0,1]^{\mathrm{T}}，\quad \boldsymbol{X}_2=[1,0,0,1]^{\mathrm{T}}，\quad \boldsymbol{X}_3=[1,0,1,1]^{\mathrm{T}}，\quad \boldsymbol{X}_4=[1,1,0,1]^{\mathrm{T}}$$

$$\boldsymbol{X}_5=[0,0,-1,-1]^{\mathrm{T}}，\quad \boldsymbol{X}_6=[0,-1,-1,-1]^{\mathrm{T}}，\quad \boldsymbol{X}_7=[0,-1,0,-1]^{\mathrm{T}}，\quad \boldsymbol{X}_8=[-1,-1,-1,-1]^{\mathrm{T}}$$

取c =1开始迭代：

第一轮：

$$\boldsymbol{W}^{\mathrm{T}}(1)\boldsymbol{X}_1=[-1\quad -2\quad -2\quad 0]\begin{bmatrix}0\\0\\0\\1\end{bmatrix}=0\le 0，故\boldsymbol{W}(2)=\boldsymbol{W}(1)+\boldsymbol{X}_1=[-1,-2,-2,1]^{\mathrm{T}}$$

$$\boldsymbol{W}^{\mathrm{T}}(2)\boldsymbol{X}_2=[-1\quad -2\quad -2\quad 1]\begin{bmatrix}1\\0\\0\\1\end{bmatrix}=0\le 0，故\boldsymbol{W}(3)=\boldsymbol{W}(2)+\boldsymbol{X}_2=[0,-2,-2,2]^{\mathrm{T}}$$

$$\boldsymbol{W}^{\mathrm{T}}(3)\boldsymbol{X}_3=[0\quad -2\quad -2\quad 2]\begin{bmatrix}1\\0\\1\\1\end{bmatrix}=0\le 0，故\boldsymbol{W}(4)=\boldsymbol{W}(3)+\boldsymbol{X}_3=[1,-2,-1,3]^{\mathrm{T}}$$

$$\boldsymbol{W}^{\mathrm{T}}(4)\boldsymbol{X}_4=[1\quad -2\quad -1\quad 3]\begin{bmatrix}1\\1\\0\\1\end{bmatrix}=2>0，故\boldsymbol{W}(5)=\boldsymbol{W}(4)$$

$$\boldsymbol{W}^{\mathrm{T}}(5)\boldsymbol{X}_5=[1\quad -2\quad -1\quad 3]\begin{bmatrix}0\\0\\-1\\-1\end{bmatrix}=-2\le 0\text{，故}\boldsymbol{W}(6)=\boldsymbol{W}(5)+\boldsymbol{X}_5=[1,-2,-2,2]^{\mathrm{T}}$$

$$\boldsymbol{W}^{\mathrm{T}}(6)\boldsymbol{X}_6=[1\quad -2\quad -2\quad 2]\begin{bmatrix}0\\-1\\-1\\-1\end{bmatrix}=2>0\text{，故}\boldsymbol{W}(7)=\boldsymbol{W}(6)$$

$$\boldsymbol{W}^{\mathrm{T}}(7)\boldsymbol{X}_7=[1\quad -2\quad -2\quad 2]\begin{bmatrix}0\\-1\\0\\-1\end{bmatrix}=0\le 0\text{，故}\boldsymbol{W}(8)=\boldsymbol{W}(7)+\boldsymbol{X}_7=[1,-3,-2,1]^{\mathrm{T}}$$

$$\boldsymbol{W}^{\mathrm{T}}(8)\boldsymbol{X}_8=[1\quad -3\quad -2\quad 1]\begin{bmatrix}-1\\-1\\-1\\-1\end{bmatrix}=3>0\text{，故}\boldsymbol{W}(9)=\boldsymbol{W}(8)$$

第二轮：

$\boldsymbol{W}^{\mathrm{T}}(9)\boldsymbol{X}_1=1>0$，故$\boldsymbol{W}(10)=\boldsymbol{W}(9)$

$\boldsymbol{W}^{\mathrm{T}}(10)\boldsymbol{X}_2=2>0$，故$\boldsymbol{W}(11)=\boldsymbol{W}(10)$

$\boldsymbol{W}^{\mathrm{T}}(11)\boldsymbol{X}_3=0\le 0$，故$\boldsymbol{W}(12)=\boldsymbol{W}(11)+\boldsymbol{X}_3=[2,-3,-1,2]^{\mathrm{T}}$

$\boldsymbol{W}^{\mathrm{T}}(12)\boldsymbol{X}_4=1>0$，故$\boldsymbol{W}(13)=\boldsymbol{W}(12)$

$\boldsymbol{W}^{\mathrm{T}}(13)\boldsymbol{X}_5=-1\le 0$，故$\boldsymbol{W}(14)=\boldsymbol{W}(13)+\boldsymbol{X}_5=[2,-3,-2,1]^{\mathrm{T}}$

$\boldsymbol{W}^{\mathrm{T}}(14)\boldsymbol{X}_6=4>0$，故$\boldsymbol{W}(15)=\boldsymbol{W}(14)$

$\boldsymbol{W}^{\mathrm{T}}(15)\boldsymbol{X}_7=2>0$，故$\boldsymbol{W}(16)=\boldsymbol{W}(15)$

$\boldsymbol{W}^{\mathrm{T}}(16)\boldsymbol{X}_8=2>0$，故$\boldsymbol{W}(17)=\boldsymbol{W}(16)$

第三轮：

$\boldsymbol{W}^{\mathrm{T}}(17)\boldsymbol{X}_1=1>0$，故$\boldsymbol{W}(18)=\boldsymbol{W}(17)$

$\boldsymbol{W}^{\mathrm{T}}(18)\boldsymbol{X}_2=3>0$，故$\boldsymbol{W}(19)=\boldsymbol{W}(18)$

$\boldsymbol{W}^{\mathrm{T}}(19)\boldsymbol{X}_3=1>0$，故$\boldsymbol{W}(20)=\boldsymbol{W}(19)$

$\boldsymbol{W}^{\mathrm{T}}(20)\boldsymbol{X}_4=0\le 0$，故$\boldsymbol{W}(21)=W(20)+\boldsymbol{X}_4=[3,-2,-2,2]^{\mathrm{T}}$

$\boldsymbol{W}^{\mathrm{T}}(21)\boldsymbol{X}_5=0\le 0$，故$\boldsymbol{W}(22)=\boldsymbol{W}(21)+\boldsymbol{X}_5=[3,-2,-3,1]^{\mathrm{T}}$

$\boldsymbol{W}^{\mathrm{T}}(22)\boldsymbol{X}_6=4>0$，故$\boldsymbol{W}(23)=\boldsymbol{W}(22)$

$\boldsymbol{W}^{\mathrm{T}}(23)\boldsymbol{X}_7=1>0$，故$\boldsymbol{W}(24)=\boldsymbol{W}(23)$

$\boldsymbol{W}^{\mathrm{T}}(24)\boldsymbol{X}_8=1>0$，故$\boldsymbol{W}(25)=\boldsymbol{W}(24)$

第四轮：

$\boldsymbol{W}^{\mathrm{T}}(25)\boldsymbol{X}_1=1>0$，故$\boldsymbol{W}(26)=\boldsymbol{W}(25)$

$\boldsymbol{W}^{\mathrm{T}}(26)\boldsymbol{X}_2=4>0$，故$\boldsymbol{W}(27)=\boldsymbol{W}(26)$

$\boldsymbol{W}^{\mathrm{T}}(27)\boldsymbol{X}_3=1>0$，故$\boldsymbol{W}(28)=\boldsymbol{W}(27)$

$\boldsymbol{W}^{\mathrm{T}}(28)\boldsymbol{X}_4=2>0$，故$\boldsymbol{W}(29)=\boldsymbol{W}(28)$

$\boldsymbol{W}^{\mathrm{T}}(29)\boldsymbol{X}_5=2>0$，故$\boldsymbol{W}(30)=\boldsymbol{W}(29)$

$\boldsymbol{W}^{\mathrm{T}}(30)\boldsymbol{X}_6=4>0$，故$\boldsymbol{W}(31)=\boldsymbol{W}(30)$

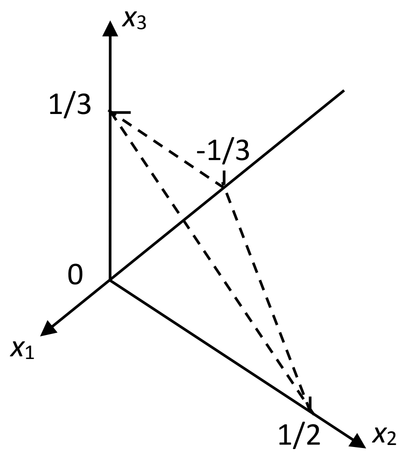
$\boldsymbol{W}^{\mathrm{T}}(31)\boldsymbol{X}_7=1>0$，故$\boldsymbol{W}(32)=\boldsymbol{W}(31)$

$\boldsymbol{W}^{\mathrm{T}}(32)\boldsymbol{X}_8=1>0$，故$\boldsymbol{W}(33)=\boldsymbol{W}(32)$

该轮迭代分类结果全部正确，故解向量$\boldsymbol{W}=[3,-2,-3,1]^{\mathrm{T}}$，对应的判别函数为：

$$d(\boldsymbol{X})=3x_1-2x_2-3x_3+1$$

判别界面$d(\boldsymbol{X})=0$如解图3.4所示，图中虚线为判别界面与坐标面x_1ox_2，x_1ox_3，x_2ox_3的交线。



解图 3.4　判别界面

4.2　假设在某个地区的疾病普查中，异常细胞（ω_1）和正常细胞（ω_2）的先验概率分别为$P(\omega_1)=0.1$，$P(\omega_2)=0.9$。现有一待识别细胞，其观察值为 X，从类概率密度分布曲线上查得$p(X\mid\omega_1)=0.4$，$p(X\mid\omega_2)=0.2$试对该细胞利用最小错误率贝叶斯决策规则进行分类。

解 1： $P(\omega_2 \mid X) = \dfrac{p(X \mid \omega_2)P(\omega_2)}{\sum_{i=1}^{2} p(X \mid \omega_i)P(\omega_i)} = \dfrac{0.2\times0.9}{0.2\times0.9+0.4\times0.1} \approx 0.818$

$P(\omega_1 \mid X) = \dfrac{0.4\times0.1}{0.2\times0.9+0.4\times0.1} \approx 0.182$

Q $P(\omega_2 \mid X) > P(\omega_1 \mid X)$ $\therefore X \in \omega_2$（正常）

解 2： $p(X \mid \omega_2)P(\omega_2) = 0.2\times0.9 = 0.18$， $p(X \mid \omega_1)P(\omega_1) = 0.4\times0.1 = 0.04$

Q $p(X \mid \omega_2)P(\omega_2) > p(X \mid \omega_1)P(\omega_1)$ $\therefore X \in \omega_2$（正常）

4.4 对 4.2 题中两类细胞的分类问题（异常细胞 ω_1 和正常细胞 ω_2 ），除已知的数据外，若损失函数的值分别为 $L_{11}=0$， $L_{21}=6$， $L_{12}=1$， $L_{22}=0$，试用最小风险贝叶斯决策规则对细胞进行分类。

解 1：当 X 被判为 ω_1 类时：

$$d_1(X) = L_{11}p(X \mid \omega_1)P(\omega_1) + L_{12}p(X \mid \omega_2)P(\omega_2) = 0\times0.4\times0.1 + 1\times0.2\times0.9 = 0.18$$

当 X 被判为 ω_2 类时：

$$d_2(X) = L_{21}p(X \mid \omega_1)P(\omega_1) + L_{22}p(X \mid \omega_2)P(\omega_2) = 6\times0.4\times0.9 + 0\times0.2\times0.1 = 2.16$$

$\because d_1(X) < d_2(X)$， $\therefore X \in \omega_1$（异常）

解 2： $l_{12}(X) = \dfrac{p(X \mid \omega_1)}{p(X \mid \omega_2)} = \dfrac{0.4}{0.2} = 2$

$\theta_{12} = \dfrac{(L_{12}-L_{22})P(\omega_2)}{(L_{21}-L_{11})P(\omega_1)} = \dfrac{(1-0)\times0.9}{(6-0)\times0.1} = 1.5$

Q $l_{12}(X) > \theta_{12}$， $\therefore X \in \omega_1$（异常）