作业

1．一个三类问题，其判别函数如下：

$$d_1(x) = -x_1, d_2(x) = x_1 + x_2 - 1, d_3(x) = x_1 - x_2 - 1$$

a）设这些函数是在多类情况1条件下确定的，绘出其判别界面和每个模式类别的区域。

b）设为多类情况2，并使$d_{12}(x) = d_1(x), d_{13}(x) = d_2(x), d_{23}(x) = d_3(x)$。绘出其判别界面和多类情况2的区域。

c）设$d_1(x), d_2(x)$和$d_3(x)$是在多类情况3的条件下确定的，绘出其判别界面和每类的区域。

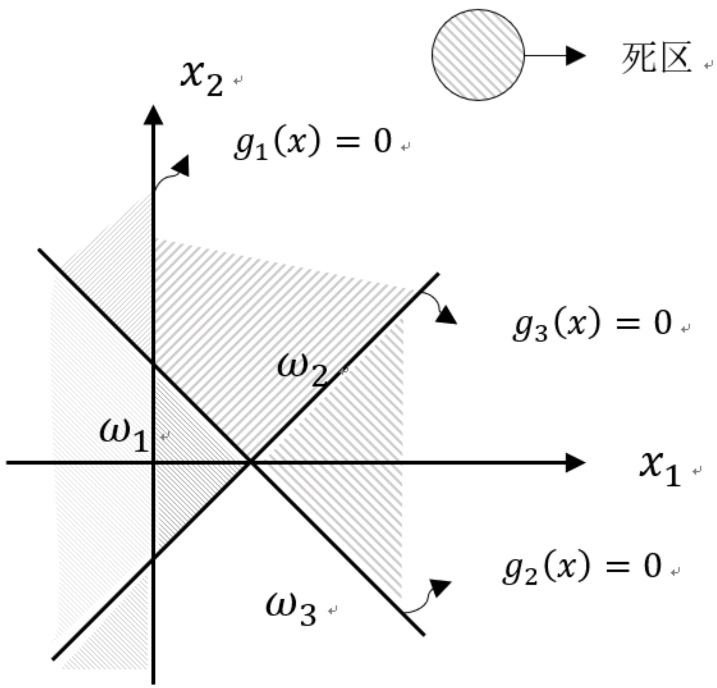
问题：没有具体画出或者画错每个模式类别的区域。

作业

解：a) $\begin{cases} g_1(x) = -x_1 = 0 \\ g_2(x) = x_1 + x_2 - 1 = 0 \\ g_3(x) = x_1 - x_2 - 1 = 0 \end{cases}$，由$x_1 + x_2 - 1 > 0$等，得：

$$\omega_1:\begin{cases} g_1(x) > 0 \\ g_2(x) < 0 \\ g_3(x) < 0 \end{cases},\quad \omega_2:\begin{cases} g_1(x) < 0 \\ g_2(x) > 0 \\ g_3(x) < 0 \end{cases},\quad \omega_3:\begin{cases} g_1(x) < 0 \\ g_2(x) < 0 \\ g_3(x) > 0 \end{cases}$$

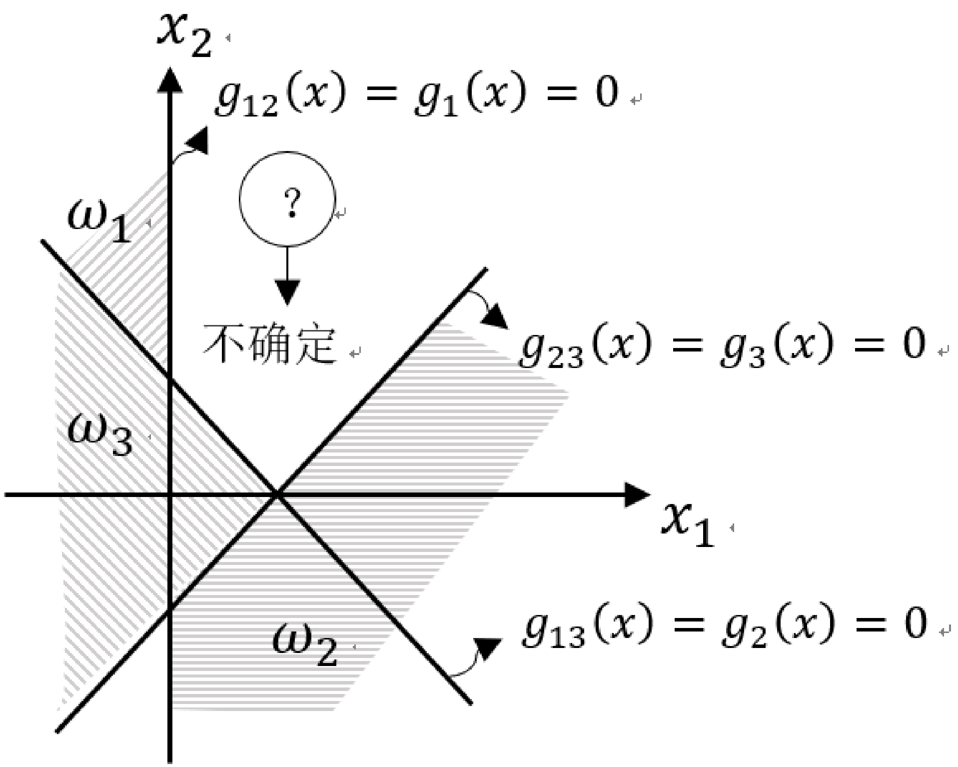
如下图所示：



作业

b) $\omega_1:\begin{cases} g_{12}(x)>0 \\ g_{13}(x)>0 \end{cases}$, $\omega_2:\begin{cases} g_{21}(x)>0 \\ g_{23}(x)>0 \end{cases}$, $\omega_3:\begin{cases} g_{31}(x)>0 \\ g_{32}(x)>0 \end{cases}$ $(g_{ij}(x)=-g_{ji}(x))$

如下图所示：

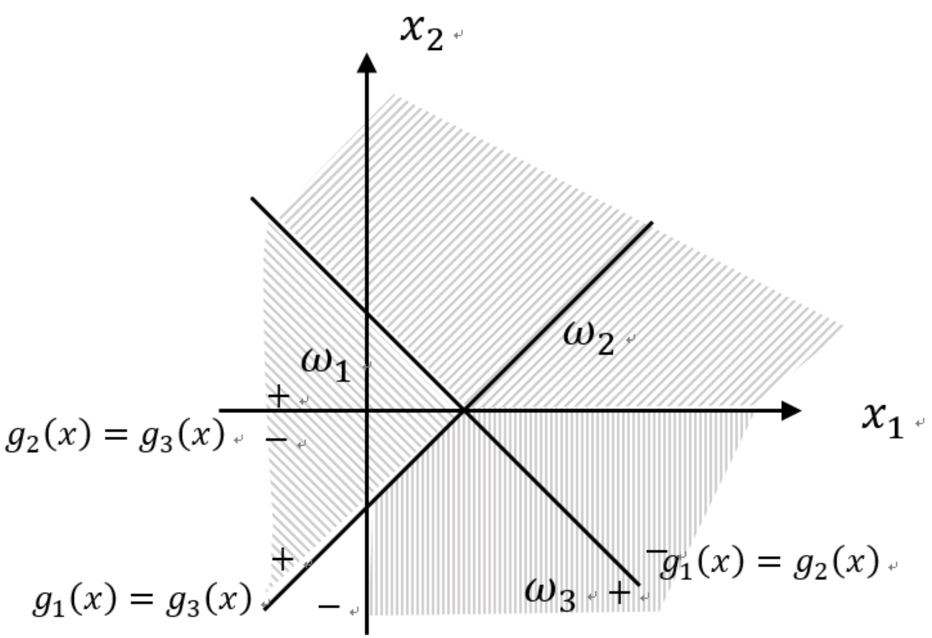


作业

c） $\omega_1:\begin{cases} g_1(x) > g_2(x) \\ g_1(x) > g_3(x) \end{cases}$， $\omega_2:\begin{cases} g_2(x) > g_1(x) \\ g_2(x) > g_3(x) \end{cases}$， $\omega_3:\begin{cases} g_3(x) > g_1(x) \\ g_3(x) > g_2(x) \end{cases}$

判别边界： $\begin{cases} g_1(x) - g_2(x) = 0 \\ g_1(x) - g_3(x) = 0 \\ g_2(x) - g_3(x) = 0 \end{cases}$

如下图所示：



作业

2. 设两类样本的类内离散度矩阵分别 $S_1 = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}, S_2 = \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix}$，两类的中心点分别为$\mathrm{m}_1 = (2,0)^T$，$\mathrm{m}_2 = (2,2)^T$, 试用$Fisher$准则求其决策面阈值。

解：

$$S_w = S_1 + S_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, S_w^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}, \mu = S_w^{-1}(m_1 - m_2) = (0, -1)^T$$

$$\tilde{m}_1 = \mu^T m_1 = 0, \tilde{m}_2 = \mu^T m_2 = -2, y_t = \frac{\tilde{m}_1 + \tilde{m}_2}{2} = -1$$

（存在一些计算失误的问题）

作业

3. 有两类样本，$\begin{matrix} \omega_1:\{(0 \quad 0 \quad 1)^T,(0 \quad 1 \quad 0)^T,(0 \quad 1 \quad 1)^T\} \\ \omega_2:\{(1 \quad 1 \quad 0)^T,(1 \quad 0 \quad 1)^T,(1 \quad 1 \quad 1)^T\} \end{matrix}$，用$Fisher$准则进行降维分类（分类阈值为$W_0=\frac{\bar{Y}_1+\bar{Y}_2}{2}$，其中$\bar{Y}_1$是$\omega_1$类降维的类样本均值、 $\bar{Y}_2$是ω_2类降维的类样本均值），并分析其分类性能。

解：（1）计算类内散度矩阵：（出现计算失误）

$$S_\omega=\sum_{x_i\in\omega_1}(\overrightarrow{x_i}-\overrightarrow{\mu_1})(\overrightarrow{x_i}-\overrightarrow{\mu_1})^T+\sum_{x_i\in\omega_2}(\overrightarrow{x_i}-\overrightarrow{\mu_2})(\overrightarrow{x_i}-\overrightarrow{\mu_2})^T$$

$$=\begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{6}{9} & -\frac{3}{9} \\ 0 & -\frac{3}{9} & \frac{6}{9} \end{bmatrix}\times 2=\begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{4}{3} & -\frac{2}{3} \\ 0 & -\frac{2}{3} & \frac{4}{3} \end{bmatrix}$$

作业

（2）计算类间散度矩阵：

$$S_B=(\overrightarrow{\mu_2}-\overrightarrow{\mu_1})(\overrightarrow{\mu_2}-\overrightarrow{\mu_1})^T=(1,0,0)(1,0,0)^T=\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}$$

（3）求解特征值：$S_\omega{}^{-1}S_B\vec{\omega}=\lambda\vec{\omega}$

由于$|S_\omega|=0$，即S_ω为奇异矩阵，S_ω的逆不存在。因此，该问题无法用Fisher线性判别方法求解。

（忽略了用Fisher线性判别的前提要是S_ω 可逆）

作业

4.　设两类数据的线性判别函数为$y = 3x_1 + 2x_2 + 6$，则该判别函数对应判别界面到原点的距离为？原点在判别界面的哪侧？

解：根据公式

$$d(x_0, g(x)) = \frac{g(x_0)}{\|w\|} = \frac{6}{\sqrt{9+4}} = \frac{6}{\sqrt{13}}$$

距离为$\frac{6}{\sqrt{13}}$在判别函数的正侧

作业

5. 请写出支持向量机的拉格朗日对偶优化问题的代价函数？如何理解在迭代优化过程中，先固定拉格朗日乘子α，最小化代价优化判别函数w和b，然后固定w和b，最大化代价优化乘子α？

解：支持向量机的拉格朗日对偶优化问题的代价函数为：

$$\min_{w,b} \max_{\alpha} L(w,b,\alpha)$$

其中

$$L(w,b,\alpha)=\frac{1}{2}w^T w+\sum_{n=1}^{N}\alpha_n[1-y_n(w^T x_n+b)],\alpha_n\geq 0$$

要点：1）存在多个变量时，可选择优化某一部分变量，固定其他变量的方式解决问题。

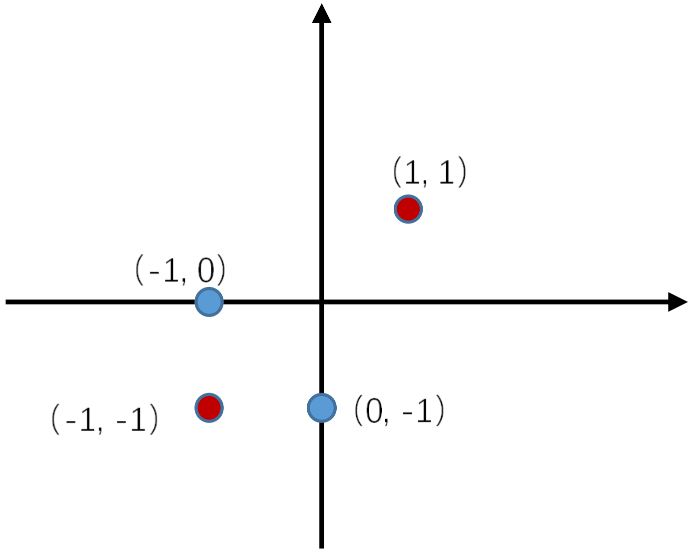
2）当固定α时，w和b的最优解可通过求导等于0获得，进而简化原问题的$minmax$问题为α的最大化问题。

作业

6. 考虑一个2维空间中的有监督学习问题，假设有2个正样本点，坐标是 $(1,1)^T$ 和 $(-1,-1)^T$，以及2个负样本点，坐标是 $(0,-1)^T$ 和 $(-1,0)^T$。请在空间中画出这4个样本点，请问这两类样本线性可分吗？考虑映射函数 $\phi(x) = (1, x_1, x_2, x_1 x_2)$ 将两类样本投影到新的4维特征空间，请写出投影空间中4个样本点的坐标，请问它们线性可分吗？采用SVM方法进行分类，如果判别函数形式是 $y(x) = \boldsymbol{w}^T \phi(x)$，对应 $\boldsymbol{w}$ 的解是什么？

解：原空间中的两类样本不是线性不可分的。

选择投影函数进行投影后，对应正样本两个点的坐标是：$(1,1,1,1)^T$ 和 $(1,-1,-1,1)^T$，负样本两个点的坐标是：$(1,0,-1,0)^T$ 和 $(1,-1,0,0)^T$。投影后的两类样本线性可分，对应SVM分类器的权向量 $\boldsymbol{w} = [0,0,0,1]^T$。



作业

7. 给定两类样本$X = \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} 1.2 \\ 1.3 \end{bmatrix} \begin{bmatrix} 1.8 \\ 3 \end{bmatrix} \begin{bmatrix} 2.3 \\ 1.5 \end{bmatrix} \right\}$,

$Y = \left\{ \begin{bmatrix} 1.5 \\ 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1.3 \end{bmatrix} \begin{bmatrix} 2.5 \\ 0.5 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}$，如下图所示，请按照$AdaBoost$的思想给出三个弱分类器，并对样本$[2.5 \quad 1]^T$和$[1 \quad 4]^T$进行分类。

解：设分类器结果为1时标记为圆圈，结果为−1时标记为三角。给定样本分别描述为$X_0 \sim X_3$和$Y_0 \sim Y_3$。样本(α, β)对应的类型为γ。

对于给定数据，初始化样本权重：$D_1(i) = \frac{1}{m} = 0.125(m = 8, i = 1,2,3, \ldots, m)$，假设三个弱分类器为：

$$h_1(\alpha, \beta) = \begin{cases} 1, \alpha < 1.3 \\ -1, \alpha > 1.3 \end{cases}, \quad h_2(\alpha, \beta) = \begin{cases} 1, \alpha < 2.4 \\ -1, \alpha > 2.4 \end{cases},$$

$$h_3(\alpha, \beta) = \begin{cases} -1, \beta < 2.5 \\ 1, \beta > 2.5 \end{cases}。$$

作业

①弱分类器$t = 1$：

误差最小的阈值取在横轴方向1.2和1.5之间，此时采用第1个弱分类器，即

$$h_1(\alpha,\beta) = \begin{cases} 1, \alpha < 1.3 \\ -1, \alpha > 1.3 \end{cases},$$

则样本X_2，X_3为错分样本，错误率为$\varepsilon_1 = P(h_1(\alpha,\beta) \neq \gamma) = 0.25$，由错误率计算得弱分类器1的权重为$w_1 = \frac{1}{2} ln \frac{1-\varepsilon_1}{\varepsilon_1} = 0.5493$。

由式

$$D_2(i) = \frac{D_1(i)\, exp[-w_1 \gamma_i h_1(\alpha_i,\beta_i)]}{Z_1}$$

更新训练样本权重为

[0.0833，0.0833，0.25，0.25，0.0833，0.0833，0.0833，0.0833]。

作业

②弱分类器$t = 2$：

误差最小的阈值取在横轴方向2.3和2.5之间，此时采用第2个弱分类器，即

$$h_2(\alpha,\beta) = \begin{cases} 1, \alpha < 2.4 \\ -1, \alpha > 2.4 \end{cases},$$

问题:误差计算错误，导致阈值选错

则样本Y_0，Y_1，为错分样本，错误率为$\varepsilon_2 = P(h_2(\alpha,\beta) \neq \gamma) = 0.1666$，由错误率计算得弱分类器2的权重为$w_2 = \frac{1}{2} ln \frac{1-\varepsilon_2}{\varepsilon_2} = 0.805$。

由式

$$D_3(i) = \frac{D_2(i)\, exp[-w_2 \gamma_i h_2(\alpha_i,\beta_i)]}{Z_2}$$

更新训练样本权重为

[0.05， 0.05， 0.15， 0.15， 0.25， 0.25， 0.05， 0.05]。

作业

③弱分类器$t = 3$：

误差最小的阈值取在纵轴方向2和3之间，此时采用第3个弱分类器，即

$$h_3(\alpha, \beta) = \begin{cases} -1, \beta < 2.5 \\ 1, \beta > 2.5 \end{cases},$$

则样本X_1，X_3为错分样本，错误率为$\varepsilon_3 = P(h_3(\alpha, \beta) \neq \gamma) = 0.2$，由错误率计算得到弱分类器3的权重为$w_3 = \frac{1}{2} ln \frac{1-\varepsilon_3}{\varepsilon_3} = 0.6931$。

作业

三个弱分类器组成的强分类器为$sign(H(\alpha,\beta)=\sum_{t=1}^{3} w_t h_t(\alpha,\beta))=sign(0.5493\times h_1(\alpha,\beta)+0.805\times h_2(\alpha,\beta)+0.6931\times h_3(\alpha,\beta))$

样本$[2.5 \quad 1]^T$和$[1 \quad 4]^T$带入计算分别得到结果−1和1，分别判定为三角和圆圈。

对于训练数据而言，给定的X类和Y类样本带入计算分别得到结果1、1、1、−1和−1、−1、−1、−1，故仅用3个弱分类器组成的强分类器对训练样本X_3的判别结果是错误的，其余正确。