

作业 2

毕嘉仪 2022010608

2024 年 5 月 18 日

理论部分

1 单选题（15 分）

1.1 D

1.2 C

1.3 D

1.4 D

1.5 B

2 计算题 (15 分)

2.1 给定两个类别的样本分别为:

84

$$\omega_1: \{(3, 1), (2, 2), (4, 3), (3, 2)\}$$

$$\omega_2: \{(1, 3), (1, 2), (-1, 1), (-1, 2)\}$$

$$N_1 = 4, \mu_1 = (3, 2)$$

$$N_2 = 4, \mu_2 = (0, 2)$$

$$\mu = (1.5, 2)$$

试利用 LDA, 将样本特征维数压缩为一维。

$$S_1 = \frac{1}{4} \left(\begin{bmatrix} 0 \\ -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \end{bmatrix} \right)$$

$$= \frac{1}{4} \left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + 0 \right) = \frac{1}{4} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$S_2 = \frac{1}{4} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \end{bmatrix} \begin{bmatrix} -1 & -1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \end{bmatrix} \right)$$

$$= \frac{1}{4} \left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right)$$

$$= \frac{1}{4} \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}$$

散射矩阵: $S_w = \frac{1}{8} \left(\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} \right)$
 $= \frac{1}{8} \begin{bmatrix} 6 & 3 \\ 3 & 6 \end{bmatrix} = \frac{3}{8} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

散射矩阵 $S_b = \frac{4}{8} \left(\begin{bmatrix} 3/2 \\ 0 \end{bmatrix} \begin{bmatrix} 3/2 & 0 \end{bmatrix} + \begin{bmatrix} -3/2 \\ 0 \end{bmatrix} \begin{bmatrix} -3/2 & 0 \end{bmatrix} \right)$
 $= \frac{1}{2} \begin{bmatrix} 9/2 & 0 \\ 0 & 0 \end{bmatrix}$
 $= \frac{9}{4} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \frac{3}{8} \begin{bmatrix} 6 & 0 \\ 0 & 0 \end{bmatrix}$

$$S_T = S_w + S_b = \frac{3}{8} \begin{bmatrix} 8 & 1 \\ 1 & 2 \end{bmatrix}$$

求特征值分解.

$$S_w^{-1} S_b v = \lambda v \Rightarrow S_w^{-1} S_b - \lambda I = 0.$$

$$\begin{array}{c|c} 2 & 1 \\ \hline 1 & 2 \end{array} \begin{array}{c} 1 \\ 0 \\ \hline 0 \\ 1 \end{array} \rightarrow \begin{array}{c|c} 4 & 2 \\ \hline 1 & 2 \end{array} \begin{array}{c} 2 \\ 0 \\ \hline 0 \\ 1 \end{array} \rightarrow \begin{array}{c|c} 3 & 0 \\ \hline 1 & 2 \end{array} \begin{array}{c} 2 \\ -1 \\ \hline 0 \\ 1 \end{array}$$

$$\rightarrow \begin{array}{c|c} 1 & 0 \\ \hline 1 & 2 \end{array} \begin{array}{c} 2/3 \\ -1/3 \\ \hline 0 \\ 1 \end{array} \rightarrow \begin{array}{c|c} 1 & 0 \\ \hline 0 & 2 \end{array} \begin{array}{c} 2/3 \\ -1/3 \\ \hline -2/3 \\ 4/3 \end{array} \rightarrow \begin{array}{c} 2/3 \\ -1/3 \\ \hline -1/3 \\ 2/3 \end{array}$$

$$S_w^{-1} = \frac{8}{3} \cdot \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \frac{8}{9} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$S_w^{-1} \cdot S_b = \frac{8}{9} \times \frac{3}{8} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 6 & 0 \\ 0 & 2 \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} 12 & 0 \\ -6 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ -2 & 0 \end{bmatrix}$$

$$\begin{vmatrix} 4-\lambda & 0 \\ -2 & -\lambda \end{vmatrix} = 0 \Rightarrow (4-\lambda)\lambda = 0$$

$\lambda = 0$ 或 4 .

$$\therefore \begin{bmatrix} 4 & 0 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 4 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$\begin{bmatrix} 4v_1 \\ -2v_1 \end{bmatrix} = 4 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \quad \begin{matrix} -2v_1 = 4v_2 \\ v_1 = -2v_2 \end{matrix}$$

$$\text{又 } v_1^2 + v_2^2 = 1 \Rightarrow \begin{matrix} v_1 = 2/\sqrt{5} \\ v_2 = -1/\sqrt{5} \end{matrix}$$

$$\therefore W_{LDA} = \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}$$

投影后

$$\begin{cases} w_1 = \{ \sqrt{5}, 2/\sqrt{5}, \sqrt{5}, 4/\sqrt{5} \} \\ w_2 = \{ -1/\sqrt{5}, 0, -3/\sqrt{5}, -4/\sqrt{5} \} \end{cases}$$

2.2 模型训练通常需要大量的数据，假设某采集的数据集包含 80% 的有效数据和 20% 的无效数据。采用一种算法判断数据是否有效，其中无效数据被成功判别为无效数据的概率为 90%，而有效数据被误判为无效数据的概率为 5%。如果某条数据经过该算法被判别为无效数据，则根据贝叶斯定理，这条数据是无效数据的概率是多少？

(提示：全概率公式 $P(Y) = \sum_{i=1}^N P(Y|X_i)P(X_i)$)

$$P(w_1) = 0.8, P(w_2) = 0.2$$

$$P(x_2|w_2) = 0.9, P(x_2|w_1) = 0.05, P(w_2|x_2)?$$

$$P(x_2) = P(x_2|w_1)P(w_1) + P(x_2|w_2)P(w_2) \\ = 0.05 \times 0.8 + 0.9 \times 0.2 = 0.04 + 0.18 = 0.22$$

$$P(w_2|x_2) = \frac{P(w_2 \cap x_2)}{P(x_2)} = \frac{P(x_2|w_2)P(w_2)}{P(x_2)} \\ = \frac{0.9 \times 0.2}{0.22} = \frac{0.9}{1.1} = 81.82\%$$

2.3 设有两类正态分布的样本集，第一类均值为 $\mu_1 = [2, -1]^T$ ，第二类均值为 $\mu_2 = [1, 1]^T$ 。两类样本集的协方差矩阵和出现的先验概率都相等： $\Sigma_1 = \Sigma_2 = \Sigma = \begin{bmatrix} 4 & 2 \\ 2 & \frac{4}{3} \end{bmatrix}$ ， $p(w_1) = p(w_2)$ 。试计算分类界面，并对特征向量 $x = [6, 2]^T$ 分类。

$$\Rightarrow \text{判决函数: } g_i(x) = -\frac{1}{2}(x - \mu_i)^T \Sigma_i^{-1}(x - \mu_i) + \ln p(w_i)$$

$$\Rightarrow g_{iLX}(x) = (\Sigma^{-1} \mu_i)^T \vec{x} - \frac{1}{2} \mu_i^T \Sigma^{-1} \mu_i$$

$$\Sigma^{-1}: \quad \Sigma = \frac{2}{3} \begin{bmatrix} 6 & 3 \\ 3 & 2 \end{bmatrix}$$

$$\begin{array}{c|c} 6 & 3 \\ \hline 3 & 2 \end{array} \begin{array}{c} 1 \\ 0 \end{array} \begin{array}{c} 0 \\ 1 \end{array} \rightarrow \begin{array}{c|c} 4 & 2 \\ \hline 3 & 2 \end{array} \begin{array}{c} 2/3 \\ 0 \end{array} \begin{array}{c} 0 \\ 1 \end{array} \rightarrow \begin{array}{c|c} 1 & 0 \\ \hline 3 & 2 \end{array} \begin{array}{c} 2/3 \\ 0 \end{array} \begin{array}{c} -1 \\ 1 \end{array}$$

$$\rightarrow \begin{array}{c|c} 0 & 2/3 \\ \hline 0 & 2 \end{array} \begin{array}{c} 2/3 \\ -1 \end{array} \begin{array}{c} -1 \\ 4 \end{array} \rightarrow \Sigma^{-1} = \frac{3}{2} \begin{bmatrix} 2/3 & -1 \\ -1 & 2 \end{bmatrix} \\ = \frac{1}{2} \begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}$$

$$\begin{aligned} \Rightarrow g_1(x) &= \frac{1}{2} \left(\begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right)^T \vec{x} \\ &\quad - \frac{1}{2} \frac{1}{2} \begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \\ &= \frac{1}{2} [7 \ -12] \vec{x} - \frac{1}{4} [2 \ -12] \begin{bmatrix} 7 \\ -12 \end{bmatrix} \\ &= \frac{1}{2} [7 \ -12] \vec{x} - \frac{1}{4} (14 + 12) \\ &= \frac{1}{2} [7 \ -12] \vec{x} - \frac{13}{2} \end{aligned}$$

$$\begin{aligned} g_2(x) &= \frac{1}{2} \left(\begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)^T \vec{x} \\ &\quad - \frac{1}{2} \times \frac{1}{2} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \frac{1}{2} [-1 \ 3] \vec{x} - \frac{1}{4} \times 2 \\ &= \frac{1}{2} [-1 \ 3] \vec{x} - \frac{1}{2} \end{aligned}$$

分类界面: $g(x) = g_1(x) - g_2(x)$

$$= \frac{1}{2} [8 \ -15] \vec{x} - 6$$

当 $\vec{x} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$ 时, $g_1(\begin{bmatrix} 6 \\ 2 \end{bmatrix}) = [7 \ -12] \begin{bmatrix} 3 \\ 1 \end{bmatrix} - \frac{13}{2}$

$$= 21 - 12 - \frac{13}{2} = \frac{5}{2}$$

$g_2(\begin{bmatrix} 6 \\ 2 \end{bmatrix}) = [-1 \ 3] \begin{bmatrix} 3 \\ 1 \end{bmatrix} - \frac{1}{2}$

$$= -3 + 3 - \frac{1}{2} = -\frac{1}{2}$$

$g_1(\begin{bmatrix} 6 \\ 2 \end{bmatrix}) > g_2(\begin{bmatrix} 6 \\ 2 \end{bmatrix})$

$\therefore \vec{x} = \begin{bmatrix} 6 \\ 2 \end{bmatrix} \in \text{第1类}$

2.4 给定异或的样本集

$$D = \{((0,0)^T, -1), ((0,1)^T, 1), ((1,0)^T, 1), ((1,1)^T, -1)\}$$

该样本集是线性不可分的, 可采用如下所示的多项式函数

$\phi(x)$ 将样本 $D = \{(x_n, y_n)\}$ 映射为 $D_\phi = \{(\phi(x_n), y_n)\}$, 其

中 $\phi(x)$ 满足

$$\begin{aligned} \phi_1(x) &= 2(x_1 - 0.5) \approx 2x_1 - 1 \\ \phi_2(x) &= 4(x_1 - 0.5)(x_2 - 0.5) \approx 4x_1x_2 - 2(x_1 + x_2) + 1 \end{aligned}$$

(1) 给出映射后的样本集;

(2) 在映射后的样本集中, 设计一个线性 SVM 分类器, 给出支持向量及分类界面。

$$1) \vec{x}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \phi_1(\vec{x}_1) = -1, \phi_2(\vec{x}_2) = 1.$$

$$\phi(\vec{x}_1) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

$$\vec{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \phi(\vec{x}_2) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

$$\vec{x}_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \phi(\vec{x}_3) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

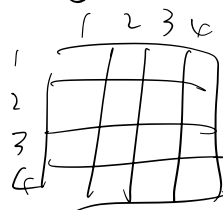
$$\vec{x}_4 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \phi(\vec{x}_4) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$$D\phi = \{(\begin{bmatrix} -1 \\ 1 \end{bmatrix}, -1), (\begin{bmatrix} -1 \\ -1 \end{bmatrix}, 1), (\begin{bmatrix} 1 \\ 1 \end{bmatrix}, 1), (\begin{bmatrix} 1 \\ -1 \end{bmatrix}, -1)\}$$

2) 有目标 fit

$$L(w, b, \vec{\alpha}) = \sum_{i=1}^4 \alpha_i - \frac{1}{2} \sum_{i=1}^4 \sum_{j=1}^4 \alpha_i \alpha_j y_i y_j \vec{x}_i^T \vec{x}_j$$

求解 maximize $L(w, b, \vec{\alpha})$



$$\frac{\partial L}{\partial \alpha_1} = 1 - \frac{1}{2} \left(\sum_{j=1}^4 \alpha_j y_j y_1 \vec{x}_1^T \vec{x}_j - \sum_{j=1}^4 \alpha_j y_j y_1 \vec{x}_j^T \vec{x}_1 \right) \text{ 一样的}$$

$$-2 \cdot \frac{1}{2} \alpha_1 y_1^2 \vec{x}_1^T \vec{x}_1 = 2 \quad (\vec{x}_1^T \vec{x}_1 = 2, y_1^2 = 1)$$

$$= 1 - (\alpha_2 (-1) \cdot 1 \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix}^T \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$+ \alpha_3 (-1) \cdot 1 \cdot \begin{bmatrix} -1 \\ -1 \end{bmatrix}^T \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$+ \alpha_4 (-1) (-1) \begin{bmatrix} 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ -1 \end{bmatrix}) - \alpha_1 \cdot 1 \cdot 2$$

$$= 1 - (2\alpha_3) - 4\alpha_1 = 0 \Rightarrow 2\alpha_1 + \alpha_3 = 1/2$$

$$\begin{aligned}
 \frac{\partial L}{\partial \alpha_2} &= 1 - \left(\sum_{j=1,3,4} \alpha_j y_j \vec{x}_2^T \vec{x}_j \right) - 2\alpha_2 \\
 &= 1 - \left(\alpha_1(-1) \cancel{[-1 -1]^T} + \alpha_3(+1) \cancel{[-1 -1]^T} \right. \\
 &\quad \left. + \alpha_4(+1) [-1 -1]^T \right) - 2\alpha_2
 \end{aligned}$$

$$= 1 - 2\alpha_4 - 2\alpha_2 = 0 \Rightarrow 2\alpha_2 + \alpha_4 = 1/2.$$

$$\begin{aligned}
 \frac{\partial L}{\partial \alpha_3} &= 1 - y_3 \left(\sum_{j=1,2,4} \alpha_j y_j \vec{x}_3^T \vec{x}_j \right) - 2\alpha_3 \\
 &= 1 - \left(\alpha_1(-1) [1 -1]^T \cancel{[-1]} + \alpha_2(+1) \cancel{[1 -1]^T} \right. \\
 &\quad \left. + \alpha_4(+1) [1 -1]^T \right) - 2\alpha_3 \\
 &= 1 - 2\alpha_1 - 2\alpha_3 = 0 \Rightarrow 2\alpha_3 + \alpha_1 = 1/2.
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial L}{\partial \alpha_4} &= 1 - y_4 \left(\sum_{j=1,2,3} \alpha_j y_j \vec{x}_4^T \vec{x}_j \right) - 2\alpha_4 \\
 &= 1 + \left(\alpha_1(-1) \cancel{[1 1]^T} + \alpha_2(+1) [1 1]^T \right. \\
 &\quad \left. + \alpha_3(+1) \cancel{[1 1]^T} \right) - 2\alpha_4 \\
 &= 1 - 2\alpha_2 - 2\alpha_4 = 0 \Rightarrow \alpha_2 + \alpha_4 = 1/2.
 \end{aligned}$$

$$\begin{aligned}
 &\Rightarrow 2\alpha_1 + \alpha_3 = 2\alpha_2 + \alpha_4 = 2\alpha_3 + \alpha_1 = 2\alpha_4 + \alpha_2 = 1/2 \\
 &\Rightarrow \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1/6. \quad \boxed{\text{均为支持向量}}
 \end{aligned}$$

$$\begin{aligned}
 \therefore \vec{w} &= \sum_{i=1}^4 \alpha_i y_i \vec{x}_i \\
 &= \frac{1}{6} (-[-1] + [-1] + [-1] - [-1]) \\
 &= \frac{1}{6} \begin{bmatrix} 0 \\ -4 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \end{bmatrix}
 \end{aligned}$$

$$\text{取 } \vec{x}_1, y_1 = ([-1], -1)$$

$$\begin{aligned}
 -1 &= \begin{bmatrix} 0 & -2/3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} + b \\
 &= -\frac{2}{3} + b \quad b = -\frac{1}{3}
 \end{aligned}$$

$$\begin{aligned}
 \therefore g(\vec{x}_\phi) &= \vec{w}^T \vec{x}_\phi + b \\
 &= \begin{bmatrix} 0 & -2/3 \end{bmatrix} \vec{x}_\phi - \frac{1}{3}
 \end{aligned}$$

$$\text{分类界面 } g(\vec{x}_\phi) = 0$$

$$\Rightarrow -\frac{2}{3}x_2 - \frac{1}{3} = 0$$

$$2x_2 = -1$$

$$x_2 = -\frac{1}{2}$$

$$\text{在原空间中, 即 } 4x_1x_2 - 2(x_1 + x_2) + 1 = \frac{1}{2}$$

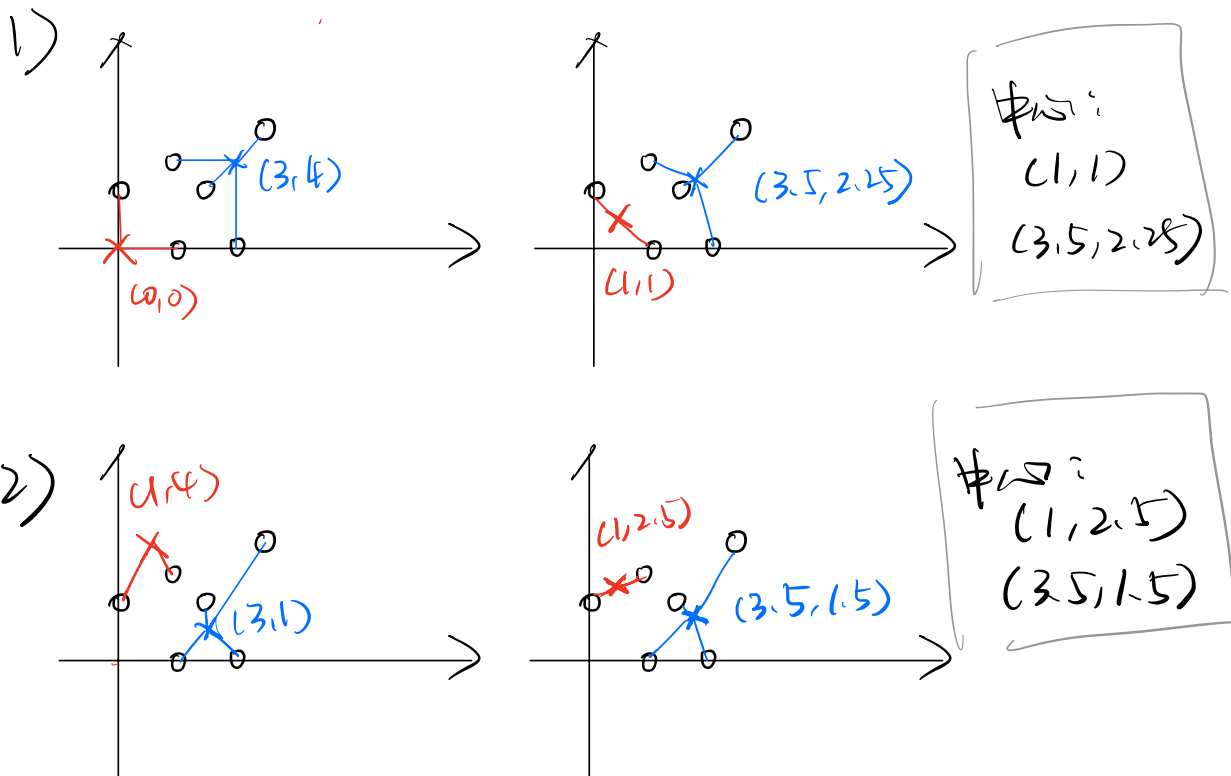
2.5 使用 KMeans 算法对 2 维空间中的 6 个点

$(0, 2), (2, 0), (2, 3), (3, 2), (4, 0), (5, 4)$ 进行聚类, 距离函数选择

欧氏距离 $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ 。

(1) 起始聚类中心选择 $(0, 0)$ 和 $(4, 3)$, 计算聚类中心;

(2) 起始聚类中心选择 $(1, 4)$ 和 $(3, 1)$, 计算聚类中心。



编程部分

2 编程作业报告

2.1 程序验证

输入：

1 python check.py

结果图像：

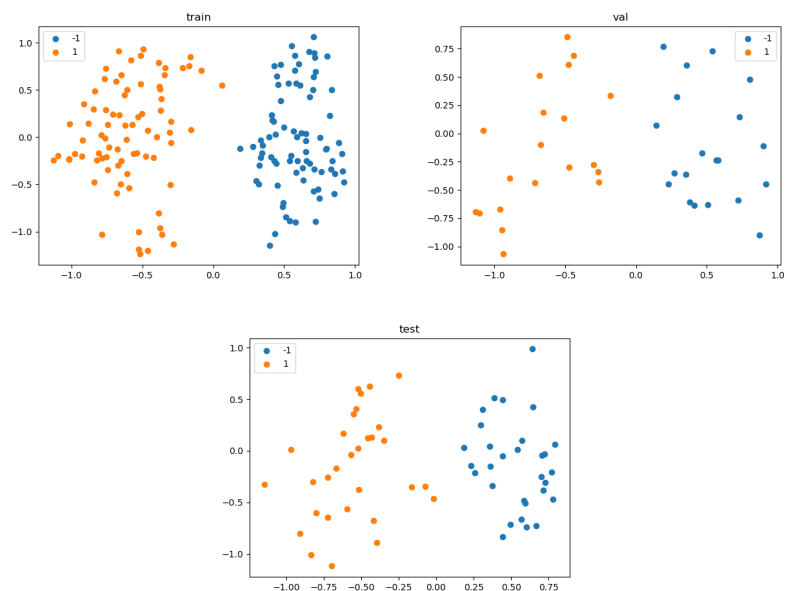
```
(media_cognition) C:\0000ijlay\011Media_Cognition_HW\hw3\hw3_release\code>python check.py
C:\0000ijlay\011Media_Cognition_HW\hw3\hw3_release\code\svm_hw.py:133: UserWarning: To copy construct from a tensor, it is recommended to use sourceTensor.clone().detach() or sourceTensor.clone().detach().requires_grad_(True), rather than torch.tensor(sourceTensor).
  self.W = torch.tensor(torch.randn(1, in_channels), requires_grad=True)
C:\0000ijlay\011Media_Cognition_HW\hw3\hw3_release\code\svm_hw.py:134: UserWarning: To copy construct from a tensor, it is recommended to use sourceTensor.clone().detach() or sourceTensor.clone().detach().requires_grad_(True), rather than torch.tensor(sourceTensor).
  self.b = torch.tensor(torch.randn(1), requires_grad=True)
C:\0000ijlay\011Media_Cognition_HW\hw3\hw3_release\code\svm_hw.py:37: UserWarning: To copy construct from a tensor, it is recommended to use sourceTensor.clone().detach() or sourceTensor.clone().detach().requires_grad_(True), rather than torch.tensor(sourceTensor).
  y = torch.tensor(x, dtype=torch.float64) @ torch.tensor(W, dtype=torch.float64).t() + b
C:\0000ijlay\011Media_Cognition_HW\hw3\hw3_release\code\svm_hw.py:37: UserWarning: To copy construct from a tensor, it is recommended to use sourceTensor.clone().detach() or sourceTensor.clone().detach().requires_grad_(True), rather than torch.tensor(sourceTensor).
  y = torch.tensor(x, dtype=torch.float64) @ torch.tensor(W, dtype=torch.float64).t() + b
Linear successfully tested!
Hinge successfully tested!
SVM_HINGE successfully tested!
```

2.2 数据预处理

输入：

1 python data_preprocess.py

可视化结果：



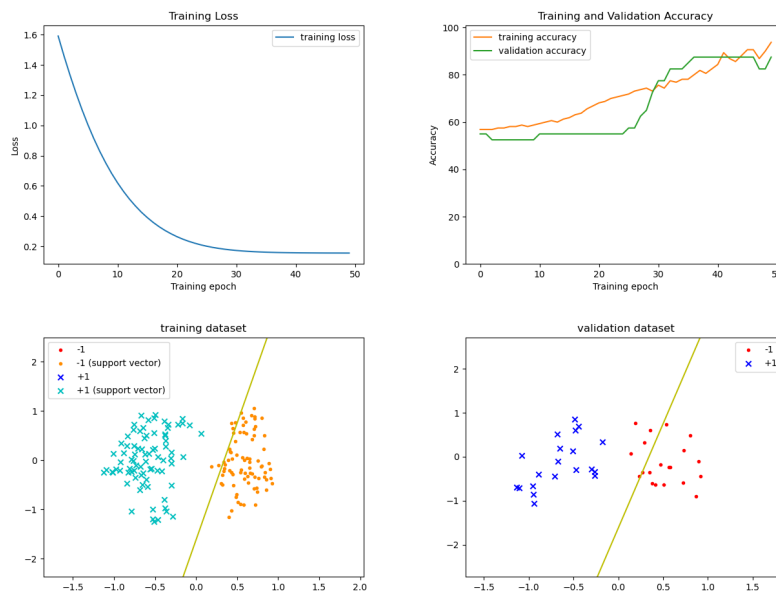
2.3 训练、验证及测试

(1) 训练模型

输入:

```
1 python train_svm.py
```

输出:



(2) 测试模型

输入:

```
1 python test_svm.py
```

输出:

```
1 Test accuracy = 88.3%
```

2.4 调整正则化系数 C ，分析不同的 C 值对分类效果的影响。

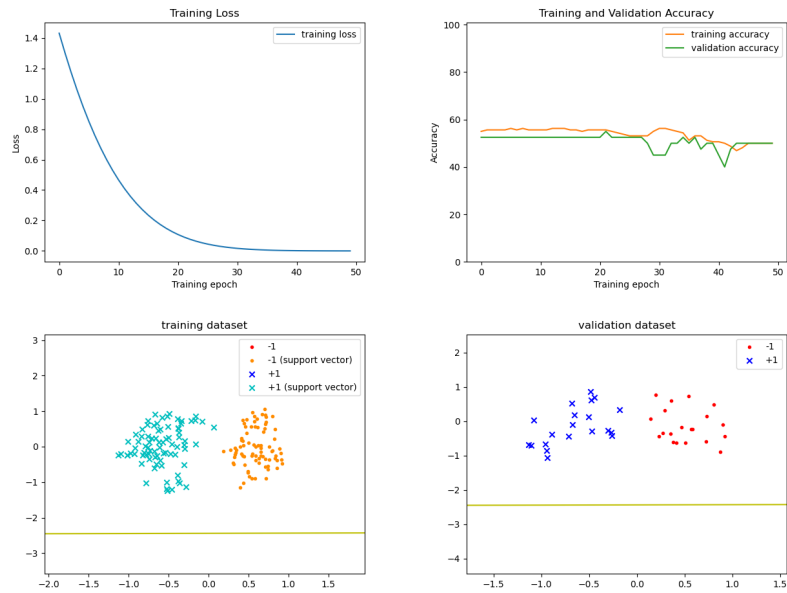
(1) 分别输入:

```
1 python train_svm.py --C 1e-6
2 python test_svm.py
3
4 python train_svm.py --C 1e-3
5 python test_svm.py
6
```

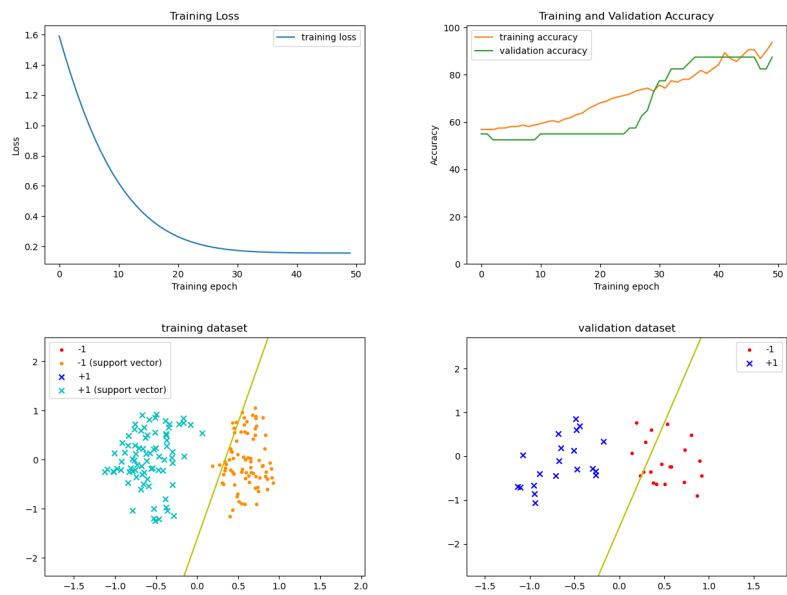
```
7 python train_svm.py --C 1
8 python test_svm.py
```

图像输出见下：

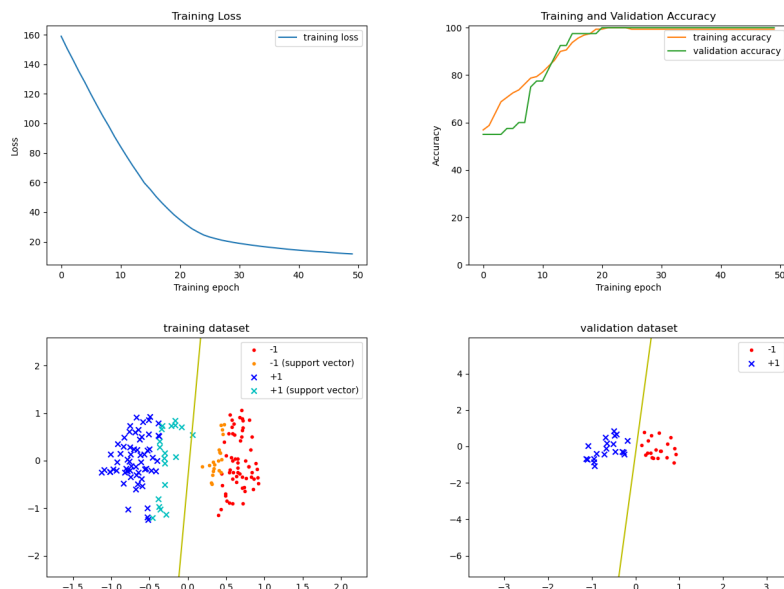
1) $C = 1e - 6$



2) $C = 1e - 3$



3) $C = 1$



测试准确率输出分别如下：

```
1 Test accuracy = 50.0%
2 Test accuracy = 88.3%
3 Test accuracy = 100.0%
```

(2) 分析

C 越大，即要求模型的误差越小，进入间隔区间的点越少，容易过拟合； C 越小，即模型的误差越大，进入间隔区间的点越多，训练集上容易出现欠拟合。

2.5 问题与解决方案

补齐代码过程中遇到的最大问题在于如何载入数据集，认真研读前两次作业的代码以后摸索出了方法。另外还在如何取出支持向量处卡了一阵子，重新研读了课上 ppt 后得以解决。

本次作业最后实现的 PCA 所得到的特征数值都很小（绝对值基本都在 1 左右），初步分析是由于作业二中 CNN 提取出的特征量级都比较小。此外，hinge loss 的 C 在默认值 $1e-3$ 处分类效果不够好，经过试验发现只要把 C 增大一点点、改成 $C = 1e-2$ 效果就很好。（从报告中所附结果图像可以看出，当 $C = 1e-3$ ，SVM 会把所有数据点都算在间隔区间内，而 $C = 1$ 时则不会）通过初步和试验分析，认为这和 PCA 结果量级太小有关。