清华大学电子工程系

媒体与认知

2023-2024 学年春季学期

作业 2

毕嘉仪 2022010608 2024 年 5 月 17 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:

$$\omega_1 : \{(3,1),(2,2),(4,3),(3,2)\}$$
$$\omega_2 : \{(1,3),(1,2),(-1,1),(-1,2)\}$$

试利用 LDA，将样本特征维数压缩为一维。

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

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

$\mu=(1.5,2)$

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

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

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

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

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

类内散度矩阵 $S_w = \frac{1}{8}\left(\begin{bmatrix}2&1\\1&2\end{bmatrix}+\begin{bmatrix}4&2\\2&4\end{bmatrix}\right)$

$= \frac{1}{8}\begin{bmatrix}6&3\\3&6\end{bmatrix} = \frac{3}{8}\begin{bmatrix}2&1\\1&2\end{bmatrix}$

类间散度矩阵 $S_b = \frac{4}{8}\left(\begin{bmatrix}3/2\\0\end{bmatrix}\begin{bmatrix}3/2&0\end{bmatrix} + \begin{bmatrix}-3/2\\0\end{bmatrix}\begin{bmatrix}-3/2&0\end{bmatrix}\right)$

$= \frac{1}{2}\begin{bmatrix}9/2&0\\0&0\end{bmatrix}$

$= \frac{9}{4}\begin{bmatrix}1&0\\0&0\end{bmatrix} = \frac{3}{8}\begin{bmatrix}6&0\\0&0\end{bmatrix}$

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

求特征值分解：

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

$\left[\begin{array}{cc|cc}2&1&1&0\\1&2&0&1\end{array}\right] \to \left[\begin{array}{cc|cc}4&2&2&0\\1&2&0&1\end{array}\right] \to \left[\begin{array}{cc|cc}3&0&2&-1\\1&2&0&1\end{array}\right]$

$\to \left[\begin{array}{cc|cc}1&0&2/3&-1/3\\1&2&0&1\end{array}\right] \to \left[\begin{array}{cc|cc}1&0&2/3&-1/3\\0&2&-2/3&4/3\end{array}\right] \to \left[\begin{array}{cc}2/3&-1/3\\-1/3&2/3\end{array}\right]$

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

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

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

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

$\lambda = 0$ 或 4.

(舍)

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

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

$$v_1 = -2v_2$$

又 $v_1^2 + v_2^2 = 1 \Rightarrow v_1 = 2/\sqrt{5}$

$v_2 = -1/\sqrt{5}$.

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

投影后 $w_1 = \{\sqrt{5}, 2\sqrt{5}, \sqrt{5}, 4\sqrt{5}\}$

$w_2 = \{-1/\sqrt{5}, 0, -3/\sqrt{5}, -4/\sqrt{5}\}$.

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 $P(Y)=\sum_{i=1}^{N}P(Y|X_i)P(X_i)$)

$P(\omega_1)=0.8,\ P(\omega_2)=0.2.$

$P(x_2|\omega_2)=0.9,\quad P(x_2|\omega_1)=0.05,\quad P(\omega_2|x_2)?$

$$P(x_2)=P(x_2|\omega_1)P(\omega_1)+P(x_2|\omega_2)P(\omega_2)$$
$$=0.05\times0.8+0.9\times0.2=0.04+0.18=0.22.$$

$$P(\omega_2|x_2)=\frac{P(\omega_2\cap x_2)}{P(x_2)}=\frac{P(x_2|\omega_2)P(\omega_2)}{P(x_2)}$$
$$=\frac{0.9\times0.2}{0.22}=\frac{0.9}{1.1}=81.82\%.$$

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：$\Sigma_1=\Sigma_2=\Sigma=\begin{bmatrix}4&2\\2&\frac{4}{3}\end{bmatrix}$，$p(\omega_1)=p(\omega_2)$。试计算分类界面，并对特征向量 $x=[6,2]^T$ 分类。

⇒ 判决函数：$g_i(x)=-\frac{1}{2}(x-\mu_i)^T\Sigma_i^{-1}(x-\mu_i)+\ln P(\omega_i)$

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

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

$$\begin{bmatrix} 6 & 3 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2/3 & 0 \\ 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2/3 & -1 \\ 0 & 1 \end{bmatrix}$$

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

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

$$\Rightarrow g_1(x) = \frac{1}{2}\left(\begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right)^T\vec{x}$$

$$-\frac{1}{2}\cdot\frac{1}{2}\begin{bmatrix} 2 & -1 \end{bmatrix}\begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} 7 & -12 \end{bmatrix}\vec{x} - \frac{1}{4}\begin{bmatrix} 2 & -1 \end{bmatrix}\begin{bmatrix} 7 \\ -12 \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} 7 & -12 \end{bmatrix}\vec{x} - \frac{1}{4}(14+12)$$

$$= \frac{1}{2}\begin{bmatrix} 7 & -12 \end{bmatrix}\vec{x} - \frac{13}{2}$$

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

$$-\frac{1}{2}\times\frac{1}{2}\begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} -1 & 3 \end{bmatrix}\vec{x} - \frac{1}{4}\times 2$$

$$= \frac{1}{2}\begin{bmatrix} -1 & 3 \end{bmatrix}\vec{x} - \frac{1}{2}.$$

分类界面: $g(x)=g_1(x)-g_2(x)$

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

当 $\vec{x}=\begin{bmatrix}6\\2\end{bmatrix}$ $=2\begin{bmatrix}3\\1\end{bmatrix}$

$g_1\left(\begin{bmatrix}6\\2\end{bmatrix}\right)=[7\ -12]\begin{bmatrix}3\\1\end{bmatrix}-\frac{13}{2}$

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

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

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

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

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

2.4 给定异或的样本集

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

该样本集是线性不可分的，可采用如下所示的多项式函数 $\phi(\mathbf{x})$ **将样本** $D=\{(\mathbf{x}_n,y_n)\}$ **映射为** $D_\phi=\{(\phi(\mathbf{x}_n),y_n)\}$**，其中** $\phi(\mathbf{x})$ **满足**

$$\phi_1(\mathbf{x})=2(x_1-0.5) \quad =2x_1-1$$

$$\phi_2(\mathbf{x})=4(x_1-0.5)(x_2-0.5) \quad =4x_1x_2-2(x_1+x_2)+1.$$

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

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

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

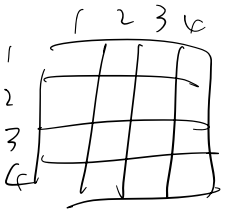
$\vec{x}_3 = \begin{bmatrix}1\\0\end{bmatrix}$, $\phi(\vec{x}_3) = \begin{bmatrix}1\\-1\end{bmatrix}$.

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

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

2) 有目标 fit

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



求解 $\max_{\alpha}(L(\vec{w}, b, \vec{\alpha}))$.

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

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

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

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

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

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

$$\frac{\partial L}{\partial \alpha_2} = 1 - \left(\sum_{j=1,3,4} \alpha_j y_2 y_j \vec{x}_2^T \vec{x}_j\right) - 2\alpha_2$$

$$= 1 - \left(\alpha_1(-1)\begin{bmatrix}-1 & -1\end{bmatrix}\begin{bmatrix}-1\\1\end{bmatrix} + \alpha_3(+1)\begin{bmatrix}-1 & -1\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix} + \alpha_4(-1)\begin{bmatrix}-1 & -1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\right) - 2\alpha_2$$

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

$$\frac{\partial L}{\partial \alpha_3} = 1 - y_3\left(\sum_{j=1,2,4} \alpha_j y_j \vec{x}_3^T \vec{x}_j\right) - 2\alpha_3$$

$$= 1 - \left(\alpha_1(-1)\begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}-1\\1\end{bmatrix} + \alpha_2(+1)\begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}-1\\-1\end{bmatrix} + \alpha_4(+1)\begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\right) - 2\alpha_3$$

$$= 1 - 2\alpha_1 - 2\alpha_3 = 0 \Rightarrow 2\alpha_3 + \alpha_1 = 1/2.$$

$$\frac{\partial L}{\partial \alpha_4} = 1 - y_4\left(\sum_{j=1,2,3} \alpha_j y_j \vec{x}_4^T \vec{x}_j\right) - 2\alpha_4$$

$$= 1 + \left(\alpha_1(-1)\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}-1\\1\end{bmatrix} + \alpha_2 \cdot 1\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}-1\\-1\end{bmatrix} + \alpha_3 \cdot 1\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix}\right) - 2\alpha_4$$

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

$$\Rightarrow 2\alpha_1 + \alpha_3 = 2\alpha_2 + \alpha_4 = 2\alpha_3 + \alpha_1 = 2\alpha_4 + \alpha_2 = 1/2$$

$$\Rightarrow \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1/6$$ 均为支持向量.

$$\vec{w} = \sum_{i=1}^{4} \alpha_i y_i \vec{x}_i$$

$$= \frac{1}{6}\left(-\begin{bmatrix}-1\\1\end{bmatrix}+\begin{bmatrix}-1\\-1\end{bmatrix}+\begin{bmatrix}1\\-1\end{bmatrix}-\begin{bmatrix}1\\1\end{bmatrix}\right)$$

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

取 $\vec{x}_1, \vec{y}_1 = \left(\begin{bmatrix}-1\\1\end{bmatrix}, -1\right)$

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

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

$$\therefore g(\vec{x}_\phi) = \vec{w}^T \vec{x}_\phi + b$$

$$= \begin{bmatrix}0 & -2/3\end{bmatrix}\vec{x}_\phi - \frac{1}{3}$$

分类界面 $g(\vec{x}_\phi) = 0$

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

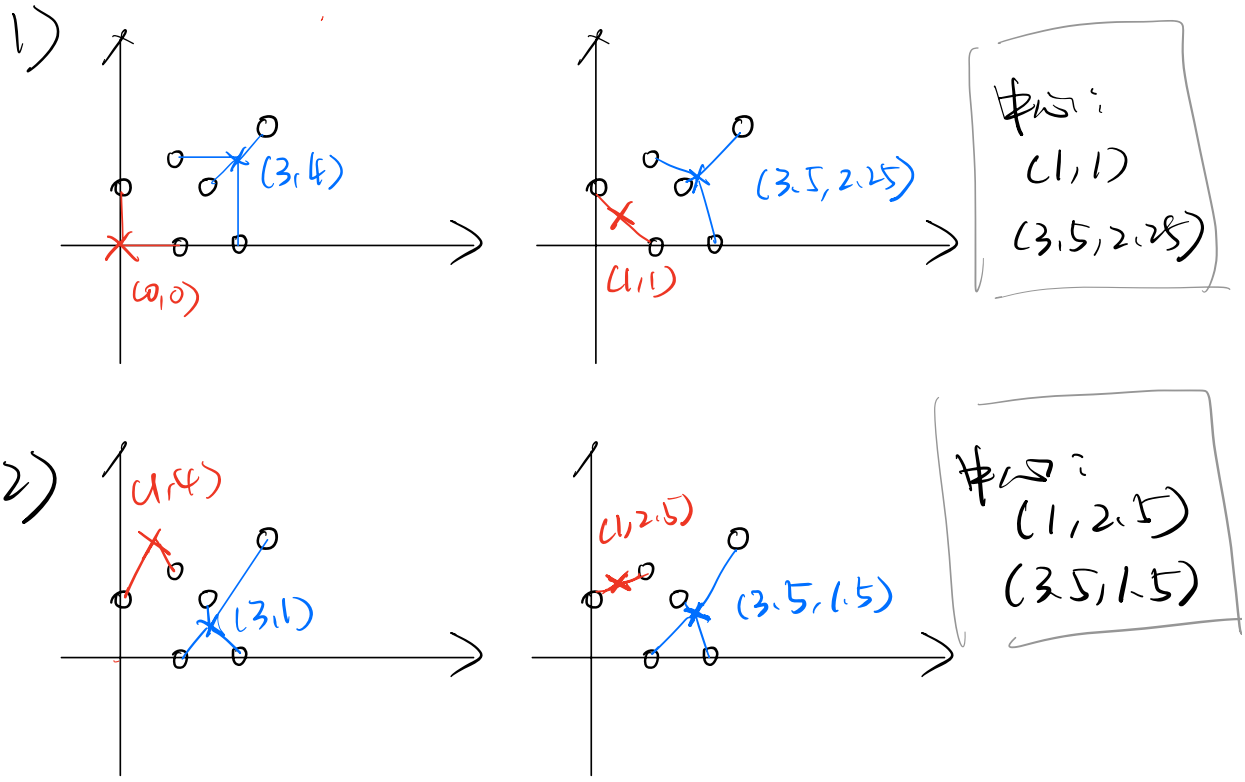
$$2x_{\phi 2} = 1 \quad x_{\phi 2} = \frac{1}{2}$$

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：

```
python check.py
```

结果图像：

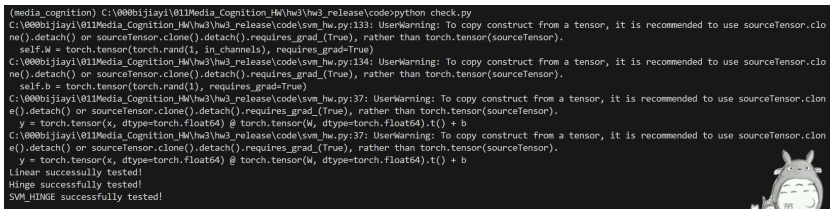


图 1

2.2 数据预处理

输入：

```
python data_preprocess.py
```

可视化结果：

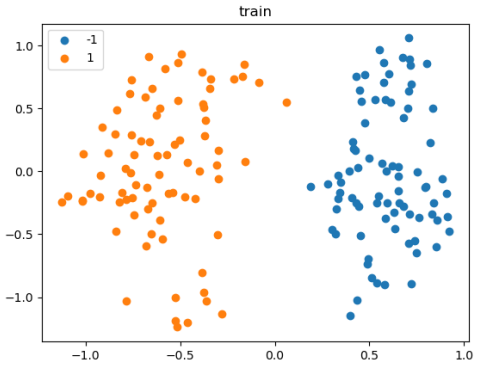


图 2

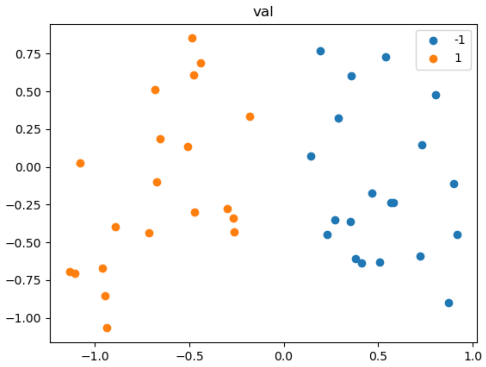


图 3

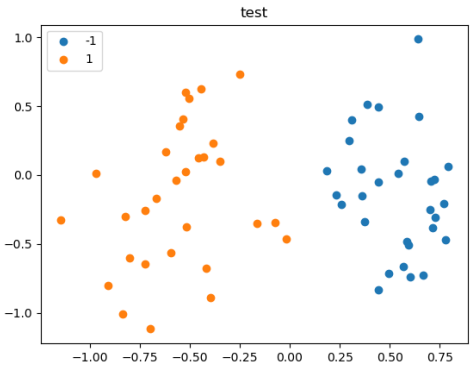


图 4

2.3 训练、验证及测试

(1) 训练模型

(2) 测试模型

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) $C = 1e - 6$

(2) $C = 1e - 3$

(3) $C = 1$

(4) 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。