模式识别与机器学习作业

BraveY

2020 年 1 月 17 日

1 第四章特征选择和提取

1.1 题目1

设有如下三类模式样本集$_1$，$_2$和$_3$，其先验概率相等，求S_w和S_b

$$
\begin{aligned}
&w_1 : \{(1,0)^T, (2,0)^T, (1,1)^T\} \\
&w_2 : \{(-1,0)^T, (0,1)^T, (-1,1)^T\} \\
&w_3 : \{(-1,-1)^T, (0,-1)^T, (0,-2)^T\}
\end{aligned}
$$

1.1.1 解

首先求出三类对应的均值向量分别为：

$$
\begin{aligned}
m_1 &= (\frac{4}{3}, \frac{1}{3})^T \\
m_2 &= (\frac{-2}{3}, \frac{2}{3})^T \\
m_3 &= (\frac{-1}{3}, \frac{-4}{3})^T
\end{aligned}
$$

分别计算出每一类对应的有偏协方差矩阵为：

$$C_1 = \begin{bmatrix} \frac{2}{9} & -\frac{1}{9} \\ -\frac{1}{9} & \frac{2}{9} \end{bmatrix}$$

$$C_2 = \begin{bmatrix} \frac{2}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{2}{9} \end{bmatrix}$$

$$C_3 = \begin{bmatrix} \frac{2}{9} & -\frac{1}{9} \\ -\frac{1}{9} & \frac{2}{9} \end{bmatrix}$$

所以带入公式：

$$\begin{aligned} S_w &= \sum_{i=1}^{c} P(\omega_i) E\{(x-m_i)(x-m_i)^T|\omega_i\} = \sum_{i=1}^{c} P(\omega_i) C_i \\ &= \frac{1}{3}C_1 + \frac{1}{3}C_2 + \frac{1}{3}C_3 \\ &= \begin{bmatrix} \frac{2}{9} & -\frac{1}{27} \\ -\frac{1}{27} & \frac{2}{9} \end{bmatrix} \end{aligned}$$

所有样本的总体均值为：

$$m_0 = (\frac{1}{9}, \frac{1}{9})^T$$

带入公式有：

$$\begin{aligned} S_b &= \sum_{i=1}^{c} P(\omega_i)(m_i - m_0)(m_i - m_0)^T \\ &= \frac{1}{3} \begin{bmatrix} \frac{11}{9} \\ -\frac{4}{9} \end{bmatrix} \begin{bmatrix} \frac{11}{9} & -\frac{4}{9} \end{bmatrix} + \frac{1}{3} \begin{bmatrix} -\frac{7}{9} \\ \frac{7}{9} \end{bmatrix} \begin{bmatrix} -\frac{7}{9} & \frac{7}{9} \end{bmatrix} + \frac{1}{3} \begin{bmatrix} -\frac{4}{9} \\ -\frac{11}{9} \end{bmatrix} \begin{bmatrix} -\frac{4}{9} & -\frac{11}{9} \end{bmatrix} \\ &= \begin{bmatrix} \frac{62}{81} & \frac{13}{81} \\ \frac{13}{81} & \frac{62}{81} \end{bmatrix} \end{aligned}$$

1.2 题目2

设有如下两类样本集，其出现的概率相等：

$$w_1 : \{(1,0,0)^T, (1,0,0)^T, (1,0,1)^T, (1,1,0)^T\}$$

$$w_2 : \{(0,0,1)^T, (0,1,0)^T, (0,1,1)^T, (1,1,1)^T\}$$

用K-L变换，分别把特征空间维数降到二维和一维，并画出样本在该空间中的位置。

1.2.1 解

计算所有样本的总体均值$m_0 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})^T$,因为不是零均值，所以直接K-L变换不是最佳的变换。先对所有样本进行均值平移有：

$$
\begin{aligned}
&w_1 : \{(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})^T, (\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})^T, (\frac{1}{2}, -\frac{1}{2}, \frac{1}{2})^T, (\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})^T\} \\
&w_2 : \{(-\frac{1}{2}, -\frac{1}{2}, \frac{1}{2})^T, (-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})^T, (-\frac{1}{2}, \frac{1}{2}, \frac{1}{2})^T, (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})^T\}
\end{aligned}
$$

计算所有样本的协方差矩阵为：

$$
\begin{aligned}
R &= \sum_{i=1}^{2} P(\omega_i) E(xx^T) \\
&= \frac{1}{2} \times \frac{1}{4} \sum_{j=1}^{3} x_{1j} x_{1j}^T + \frac{1}{2} \times \frac{1}{4} \sum_{j=1}^{3} x_{2j} x_{2j}^T \\
&= \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}
\end{aligned}
$$

求得R的3个特征值为$\lambda_1 = \frac{1}{4}, \lambda_2 = \frac{1}{4}, \lambda_2 = \frac{1}{4}$,把特征值带回去后求得特征向量可以取任意的向量。为了简化计算选择三个向量构成的矩阵为：

$$
\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

因此二维的时候选择前两个特征向量构成变换矩阵：

$$
\Phi_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}
$$

根据公式$y = \Phi^T x$得到变换后两维的新样本：

$$w_1 : \{(-\frac{1}{2}, -\frac{1}{2})^T, (\frac{1}{2}, -\frac{1}{2})^T, (\frac{1}{2}, -\frac{1}{2})^T, (\frac{1}{2}, \frac{1}{2})^T\}$$
$$w_2 : \{(-\frac{1}{2}, -\frac{1}{2})^T, (-\frac{1}{2}, \frac{1}{2})^T, (-\frac{1}{2}, \frac{1}{2})^T, (\frac{1}{2}, \frac{1}{2})^T\}$$

画出的图为：Figure 1所示。 同理降到1维的时候变换矩阵取：

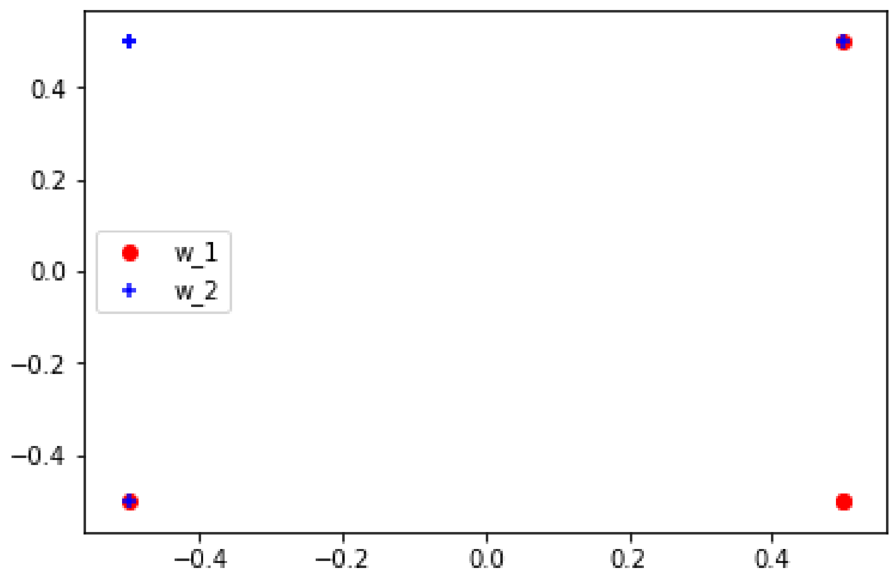

图 1: 降维到2维

$$\Phi_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

根据公式$y = \Phi^T x$得到变换后一维的新样本：

$$w_1 : \{(-\frac{1}{2})^T, (\frac{1}{2})^T, (\frac{1}{2})^T, (\frac{1}{2})^T\}$$
$$w_2 : \{(-\frac{1}{2})^T, (-\frac{1}{2})^T, (-\frac{1}{2})^T, (\frac{1}{2})^T\}$$

画出的图为：Figure 2所示。

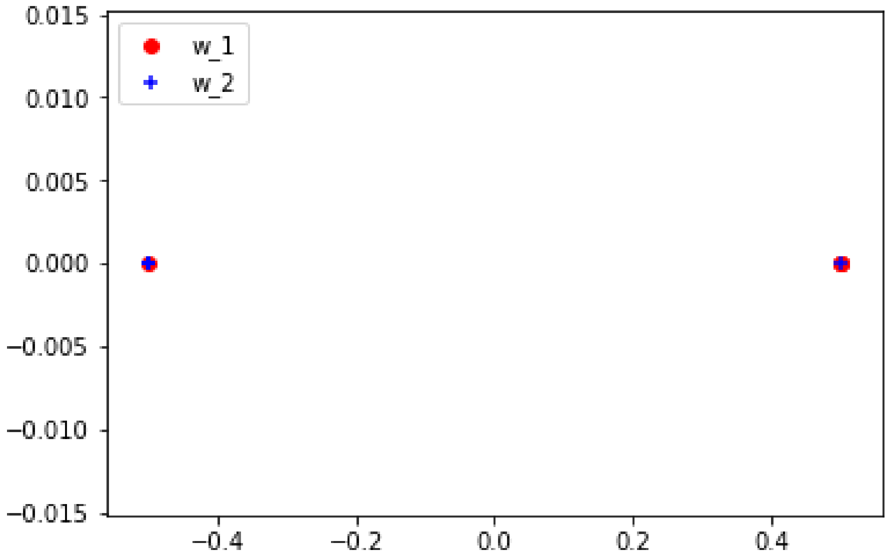

图 2: 降维到1维