模式识别作业2

张蔚桐 2015011493 自55

2017 年 3 月 16 日

1

1.1

根据题目中的提示，我们将误差函数分别对$\mathbf{w}$和w_0求导可以得到

$$\frac{\partial E}{\partial \mathbf{w}} = \sum_{i=1}^{n}((\mathbf{w}^T\mathbf{x}_i + w_0 - t_i)\mathbf{x}_i^T) \tag{1}$$

$$\frac{\partial E}{\partial w_0} = \sum_{i=1}^{n}(\mathbf{w}^T\mathbf{x}_i + w_0 - t_i) \tag{2}$$

整理可得误差函数极小值点的方程

$$\frac{\partial E}{\partial \mathbf{w}} = \mathbf{w}^T\sum_{i=1}^{n}(\mathbf{x}_i\mathbf{x}_i^T) + w_0\sum_{i=1}^{n}(\mathbf{x}_i^T) - \sum_{i=1}^{n}(t_i\mathbf{x}_i^T) \ = 0 \tag{3}$$

$$\frac{\partial E}{\partial w_0} = \mathbf{w}^T\sum_{i=1}^{n}(\mathbf{x}_i) + nw_0 - \sum_{i=1}^{n}(t_i) = 0 \tag{4}$$

根据题目中的设定$t_1 = \frac{n}{n_1}$以及$t_2 = -\frac{n}{n_2}$可以迅速得到$\sum_{i=1}^{n}(t_i) = 0$

因此化简4式得到

$$w_0 = -\frac{\mathbf{w}^T\sum_{i=1}^{n}(\mathbf{x}_i)}{n} = -\mathbf{w}^T\mathbf{m} \tag{5}$$

由5，此问得证

1.2

将5带入3可以得到如下表述

$$\begin{aligned}\mathbf{w}^T\sum_{i=1}^{n}(\mathbf{x}_i\mathbf{x}_i^T) - \mathbf{w}^T\mathbf{m}\sum_{i=1}^{n}(\mathbf{x}_i^T) - \sum_{i=1}^{n}(t_i\mathbf{x}_i^T) = 0 \\ \mathbf{w}^T\sum_{i=1}^{n}((\mathbf{x}_i - \mathbf{m})\mathbf{x}_i^T) = \sum_{i=1}^{n}(t_i\mathbf{x}_i^T)\end{aligned} \tag{6}$$

下面着重研究6式右侧的表述，按照类别可以划分为

$$
\begin{aligned}
\sum_{i=1}^{n}(t_i\mathbf{x}_i^T) = \frac{n}{n_1}\sum_{\mathcal{C}_1}(\mathbf{x}_i^T) - \frac{n}{n_2}\sum_{\mathcal{C}_2}(\mathbf{x}_i^T) \\
= n(\mathbf{m_1} - \mathbf{m_2})
\end{aligned} \tag{7}
$$

结合7，对6式取转置可得

$$
\sum_{i=1}^{n}(\mathbf{x}_i(\mathbf{x}_i - \mathbf{m})^T)\mathbf{w} = n(\mathbf{m_1} - \mathbf{m_2}) \tag{8}
$$

下面处理8式$\mathbf{w}$前系数问题，根据统计学知识，我们可以得到$\mathbf{m} = \frac{n_1\mathbf{m_1}+n_2\mathbf{m_2}}{n}$，由此，将8式左侧系数按照类别展开，得到

$$
\begin{aligned}
\sum_{i=1}^{n}(\mathbf{x}_i(\mathbf{x}_i - \mathbf{m})^T) = \sum_{\mathcal{C}_1}(\mathbf{x}_i(\mathbf{x}_i - \frac{n_1\mathbf{m_1} + n_2\mathbf{m_2}}{n})^T) \\
+\sum_{\mathcal{C}_2}(\mathbf{x}_i(\mathbf{x}_i - \frac{n_1\mathbf{m_1} + n_2\mathbf{m_2}}{n})^T) \\
= \frac{n_1}{n}\sum_{\mathcal{C}_1}(\mathbf{x}_i(\mathbf{x}_i - \mathbf{m_1})^T) + \frac{n_2}{n}\sum_{\mathcal{C}_1}(\mathbf{x}_i(\mathbf{x}_i - \mathbf{m_2})^T) \\
+\frac{n_1}{n}\sum_{\mathcal{C}_2}(\mathbf{x}_i(\mathbf{x}_i - \mathbf{m_1})^T) + \frac{n_2}{n}\sum_{\mathcal{C}_2}(\mathbf{x}_i(\mathbf{x}_i - \mathbf{m_2})^T) \\
= \frac{n_1}{n}\sum_{\mathcal{C}_1}((\mathbf{x}_i - \mathbf{m_1})(\mathbf{x}_i - \mathbf{m_1})^T) + \frac{n_1}{n}\sum_{\mathcal{C}_1}(\mathbf{m_1}(\mathbf{x}_i - \mathbf{m_1})^T) \\
+\frac{n_2}{n}\sum_{\mathcal{C}_1}((\mathbf{x}_i - \mathbf{m_1})(\mathbf{x}_i - \mathbf{m_2})^T) + \frac{n_2}{n}\sum_{\mathcal{C}_1}(\mathbf{m_1}(\mathbf{x}_i - \mathbf{m_2})^T) \\
+\frac{n_1}{n}\sum_{\mathcal{C}_2}((\mathbf{x}_i - \mathbf{m_2})(\mathbf{x}_i - \mathbf{m_1})^T) + \frac{n_1}{n}\sum_{\mathcal{C}_2}(\mathbf{m_2}(\mathbf{x}_i - \mathbf{m_1})^T) \\
+\frac{n_2}{n}\sum_{\mathcal{C}_2}((\mathbf{x}_i - \mathbf{m_2})(\mathbf{x}_i - \mathbf{m_2})^T) + \frac{n_2}{n}\sum_{\mathcal{C}_2}(\mathbf{m_2}(\mathbf{x}_i - \mathbf{m_2})^T) \\
= \frac{n_1}{n}\sum_{\mathcal{C}_1}((\mathbf{x}_i - \mathbf{m_1})(\mathbf{x}_i - \mathbf{m_1})^T) + \frac{n_2}{n}\sum_{\mathcal{C}_2}((\mathbf{x}_i - \mathbf{m_2})(\mathbf{x}_i - \mathbf{m_2})^T) \\
+\frac{n_2n_1}{n}\mathbf{m_1}(\mathbf{m_1} - \mathbf{m_2})^T - \frac{n_1n_2}{n}\mathbf{m_2}(\mathbf{m_1} - \mathbf{m_2})^T \\
+\frac{n_2}{n}\sum_{\mathcal{C}_1}((\mathbf{x}_i - \mathbf{m_1})(\mathbf{x}_i - \mathbf{m_1})^T) + \frac{n_2}{n}\sum_{\mathcal{C}_1}((\mathbf{x}_i - \mathbf{m_1})(\mathbf{m_1} - \mathbf{m_2})^T) \\
+\frac{n_1}{n}\sum_{\mathcal{C}_2}((\mathbf{x}_i - \mathbf{m_2})(\mathbf{x}_i - \mathbf{m_2})^T) + \frac{n_1}{n}\sum_{\mathcal{C}_2}((\mathbf{x}_i - \mathbf{m_2})(\mathbf{m_2})^T - \mathbf{m_1})^T) \\
= S_w + \frac{n_1n_2}{n}S_B
\end{aligned} \tag{9}
$$

于是可以得证本题

$$
(S_w + \frac{n_1n_2}{n}S_B)\mathbf{w} = n(\mathbf{m_1} - \mathbf{m_2}) \tag{10}
$$

1.3

设若$\mathbf{w} = kS_w^{-1}(\mathbf{m_2} - \mathbf{m_1})$满足上式10，将其带入10式可得

$$k\frac{n_1 n_2}{n} S_B S_w^{-1}(\mathbf{m_2} - \mathbf{m_1}) = (n + k)(\mathbf{m_1} - \mathbf{m_2}) \tag{11}$$

上式成立因此可得

$$S_B S_w^{-1}(\mathbf{m_2} - \mathbf{m_1}) = (n + k + k\frac{n_1 n_2}{n})(\mathbf{m_2} - \mathbf{m_1}) \tag{12}$$

经过上式分析可以发现，本题相当于证明$S_B S_w^{-1}$矩阵定有$(\mathbf{m_2} - \mathbf{m_1})$的特征向量且其特征值不为$n$

$$\begin{aligned} S_B S_w^{-1}(\mathbf{m_2} - \mathbf{m_1}) &= (\mathbf{m_2} - \mathbf{m_1})(\mathbf{m_2} - \mathbf{m_1})^T S_w^{-1}(\mathbf{m_2} - \mathbf{m_1}) \\ &= (\mathbf{m_2} - \mathbf{m_1})((\mathbf{m_2} - \mathbf{m_1})^T S_w^{-1}(\mathbf{m_2} - \mathbf{m_1})) \end{aligned} \tag{13}$$

可以发现右侧括号中为一二次型，结果为标量。因此可以证明$S_B S_w^{-1}$矩阵定有$(\mathbf{m_2} - \mathbf{m_1})$的特征向量。并且显然一般情况下其不为$n$ 因此此问得证。

2

其中MATLAB代码已经在附件中包含，平台采用的是MATLAB R2015b，执行时，请采用以下顺序

```
% load data, please do first
dataload;

%%%%%Fisher method%%%%%
% Fisher method
mainFisher;

% Check for Fisher method
% Please run after Fisher method
checkFisher;

%%%%%Logistic method%%%%%
% Logistic method
mainLogistic;

% Check for Logistic method
```

```
% Please run after Logistic method
checkLogistic;
%You can decide the order of Logistic method
% and Fisher method freely%
```

使用Logistics方法测试集准确性在0.97以上（三次测试），使用Fisher方法测试集准确性在0.96以上