模式识别导论上机题3-线性判别函数

薛犇 1500012752

1. 程序实现说明

本次实验采用Matlab作为编程语言，使用的版本为2016b。

在实验的一开始，利用importdata函数读取hw3_data.txt中的数据，保存在向量 raw_x中。

```
raw_x = importdata('hw3_data.txt');
```

（1）用Fisher准则求线性分类器

Fisher准则对两类问题给出了分类器的标准，假设分类后两类投影的均值为μ_1和μ_2，离散值为σ_1和σ_2，其中：

$$\mu_i = \frac{1}{n_i}\sum_{x\in w_i} w^T x, i = 1,2$$

$$\sigma_i^2 = \sum_{x\in w_i}(w^T x - \mu_i)^2, i = 1,2$$

Fisher准则要求：

$$J_F(w) = \frac{(\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}$$

这个准则函数的值最小

若假设两类样本的均值向量为m_1, m_2， 类内离散度矩阵为S_1, S_2

$$m_i = \frac{1}{n_i}\sum_{x\in w_i} x, i = 1,2$$

$$S_i = \sum_{x\in w_i}(x-m_i)(x-m_i)^T, i = 1,2$$

利用拉格朗日插值算法，可以得到最后的线性分类器的方向为：

$$\hat{w} = (S_1+S_2)^{-1}(m_1-m_2)$$

偏移值b的值也可以相应的算出

$$b = -\hat{w}^T m$$

代码的实现如下：

```
[n, d] = size(raw_x);
y = raw_x(:, d);
x = raw_x(:, 1:d-1);
x_negative = x(y ==0 ,:);
x_positive = x(y ==1 ,:);
m_negative = mean(x_negative).';
m_positive = mean(x_positive).';
S1 = cov(x_negative);
S2 = cov(x_positive);
w_fisher = inv(S1 + S2) * (m_negative - m_positive)
[n_negative, dd] = size(x_negative);
[n_positive, dd] = size(x_positive);
b_fisher = - w_fisher.' * mean(x).'
```

（2）利用最小平方误差准则求线性分类器

最小平方误差准则首先定义了样本和权重的增广向量：

$$a = (w^1, w^2, ..., w^d, 1)^T$$

$$z = (x^1, x^2, ..., x^d, 1)^T$$

求

$$\min_a J_{MSE}(a) = \min_a ||za - y||^2$$

在矩阵$z^T z$可逆的情况下，a的解析解为：

$$\hat{a} = (z^T z)^{-1} z^T y$$

代码实现如下：

```
z = [x, repmat([1], n, 1)];
a = inv(z.' * z)*z.'*y;
w_mse = a(1:3)
b_mse = a(4)
```

2.实验结果分析

（1）Fisher

得到的结果如下：

```
w_fisher =
```

```
    -2.4266
    -0.7226
     0.0045


b_fisher =

   10.1253
```

（2）MSE

得到的结果如下：

```
w_mse =

    0.1635
    0.0487
   -0.0003


b_mse =

   -0.1822
```

可以发现，两种方案解出的权重方向是几乎一致的，只是相差了几倍的尺度，若进行一下归一化，那么可以发现两者几乎一致。

但是归一化之后，得到的偏移量的大小，b，却相差了将近7倍，说明两种发式可能在不同的平面上收敛，或者说这个问题本身就有多个解。