

模式识别实验一

实验人：叶平

实验内容：

p.127 Prob. 1

- (a)、编写程序，对实验数据中的 $W1$ 的三个特征进行计算，求解最大似然估计 μ 和 δ ；
- (b)、处理二维数据，处理 $W1$ 中的任意两个特征的组合；
- (c)、处理三维数据，处理 $W1$ 中的三个特征的组合；
- (d)、在这三维高斯模型可分离的条件下，编写程序估计类别 $w2$ 中的均值和协方差矩阵中的 3 个参数；
- (e)、比较前 4 种方式计算出来的均值的异同，并加以解释；
- (f)、比较前 4 种方式计算出来的方差的异同，并加以解释。

p.129 Prob. 9

- (a)、编写用 FISHER 线性判别方法，对三维数据求最优方向 w 的通用程序；
- (b)、对表格中的类别 $W2$ 和 $W3$ ，计算最优方向 w ；
- (c)、画出表示最优方向 w 的直线，并且标记出投影后的点在直线上的位置；
- (d)、在这个子空间中，对每种分布用一维高斯函数拟合，并且求分类决策面；
- (e)、(b)中得到的分类器的训练误差是什么？
- (f)、为了比较，使用非最优方向 $w=(1.0,2.0,-1.5)'$ 重复(d)(e)两个步骤。在这个非最优子空间中，训练误差是什么。

实验结果:

Prob. 1

(a)、运行实验一下 Res.m 文件, 可以得到运行结果为:

```
x1的均值, 方差分别为:
-0.070900    0.906177

x2的均值, 方差分别为:
-0.604700    4.200715

x3的均值, 方差分别为:
-0.911000    4.541949
```

(b)、运行实验一下 Res.m 文件, 可以得到运行结果为:

```
x1, x2 特征组合的均值为:
```

```
E12 =
```

```
-0.0709    -0.6047
```

```
x1, x2 特征组合的协方差为:
```

```
S12 =
```

```
0.9062    0.5678
0.5678    4.2007
```

```
x1, x3 特征组合的均值为:
```

```
E13 =
```

```
-0.0709    -0.9110
```

```
x1, x3 特征组合的协方差为:
```

```
S13 =
```

```
0.9062    0.3941
0.3941    4.5419
```

```
x2, x3 特征组合的均值为:
```

```
E23 =
```

```
-0.6047    -0.9110
```

```
x2, x3 特征组合的协方差为:
```

```
S23 =
```

```
4.2007    0.7337
0.7337    4.5419
```

(c)、运行实验一下 Res.m 文件，可以得到运行结果为：

```
三维情况下的均值为：

E3d =

    -0.0709    -0.6047    -0.9110

三维情况下的协方差为：

S3d =

    0.9062    0.5678    0.3941
    0.5678    4.2007    0.7337
    0.3941    0.7337    4.5419
```

(d)、运行实验一下 Res.m 文件，可以得到运行结果为：

```
W2在三维情况下的均值为：

Ew2 =

    -0.1126    0.4299    0.0037

W2在三维情况下的协方差为：

Sw2 =

    0.0539         0         0
         0    0.0460         0
         0         0    0.0073
```

整理上述结果可得：
对单个特征求解均值和方差时

W1 的特征	均值	方差
X1	-0.070900	0.906177
X2	-0.604700	4.200715
X3	-0.911000	4.541949

对俩个组合特征求解均值和协方差

W1 的特征组合	均值	协方差	
X1 X2	(-0.0709 , -0.6047)	0.9062	0.5678
		0.5678	4.2007
X1 X3	(-0.0709 , -0.9110)	0.9062	0.3941
		0.3941	4.5419
X2 X3	(-0.6047 , -0.9110)	4.2007	0.7337
		0.7337	4.5419

三个特征时的均值和协方差

	均值			协方差		
W1	-0.0709	-0.6047	-0.9110	0.9062	0.5678	0.3941
				0.5678	4.2007	0.7337
				0.3941	0.7337	4.5419
W2	-0.1126	0.4299	0.0037	0.0539	0	0
				0	0.0460	0
				0	0	0.0073

(e) (f)

由表可以看出，一维和多维求均值都相等，说明均值独立。协方差对角元素和一维的方差相等，若相互独立，则协方差为对角阵。

Prob. 2

(a)、Fisher 线性判别的方向 \mathbf{W} ，计算公式如下：

$$\mathbf{W} = S_w^{-1}(\vec{m}_1 - \vec{m}_2),$$

其中：

$$S_w = S_1 + S_2,$$

$$S_i = \sum_{D_i} (X - m_i) (X - m_i)^T$$

$$m_i = \frac{1}{n_i} \sum_{D_i} X$$

其中详细代码见实验一下 FLDA.m 文件。

(b)、运行实验一下 FisherRes.m 文件，求得最优方向 \mathbf{w} 为：

由fisher线性判别计算出来的最优方向为：

$\mathbf{w} =$

-0.3832

0.2137

-0.0767

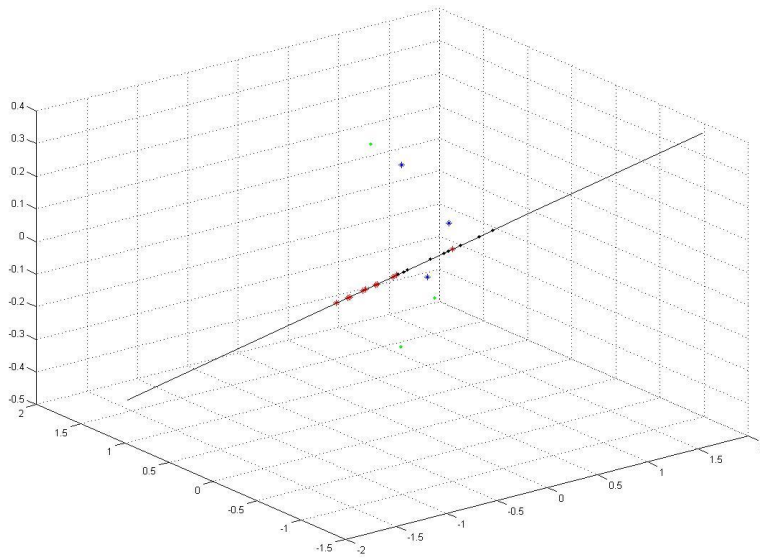
fisher条件下最优 \mathbf{w} 分类面为 $x=0.020478$

fisher条件下最优 \mathbf{w} 下误差率为0.200000

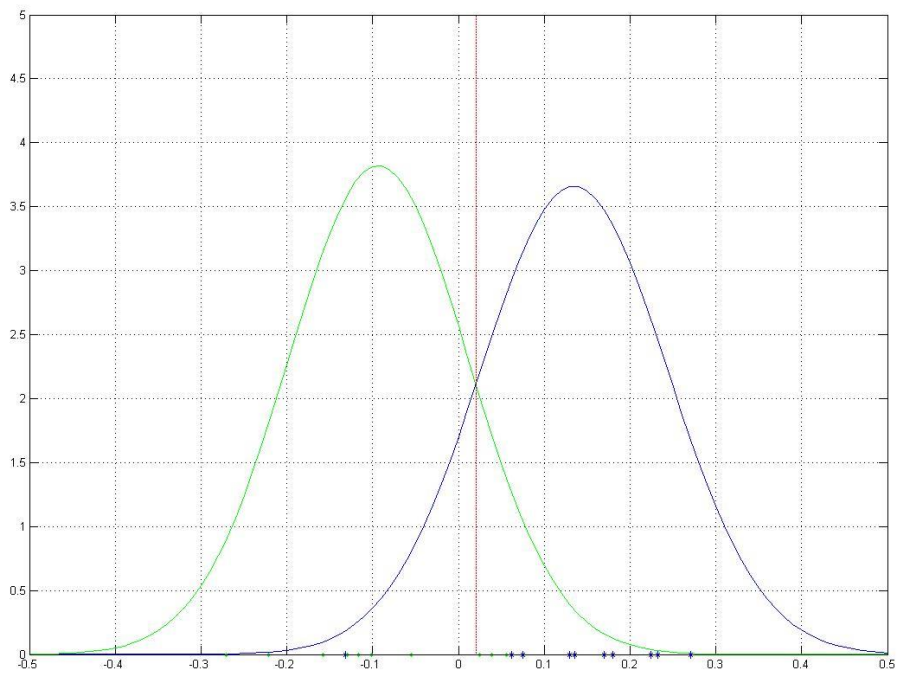
fisher条件下非最优 \mathbf{w} 的分类面为 $x=-0.062467$

fisher条件下非最优 \mathbf{w} 下误差率为0.250000

(c)、最优方向 \mathbf{w} 的直线如下，其中绿色，蓝色的点表示原来的点，红色的点表示为绿色的点在直线上的投影，黑色的点为蓝色的点在直线上的投影。

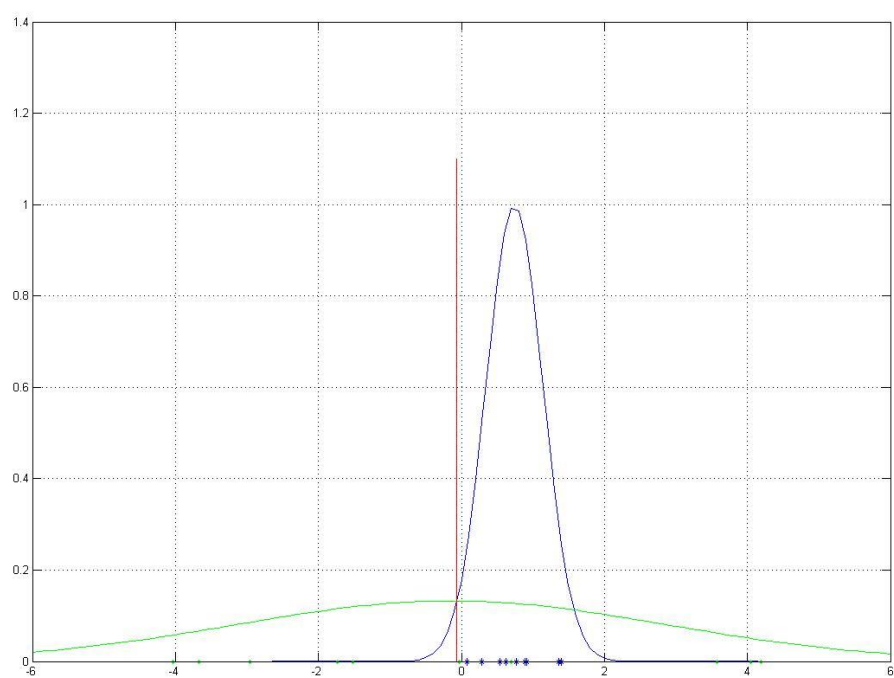


(d) (e)、在这个子空间中，对每种分布用一维高斯函数拟合，并且求分类决策面；



最优 w 的分类面为 $x=0.020478$ ；误差率为 0.200000；

(f)、为了比较，使用非最优方向 $w=(1.0,2.0,-1.5)'$ 重复(d)(e)两个步骤。在这个非最优子空间中，训练误差是什么。



分类面为 $x=-0.062467$ ；误差率为 0.250000；