实验三　Fisher 线性判别实验

姓名：徐维坚 学号：2220103484 日期：2012/7/7

一、实验目的：

1）加深对 Fisher 线性判别的基本思想的认识和理解。

2）编写实现 Fisher 线性判别准则函数的程序。

二、实验原理：

1.基本原理：

一般情况下，我们总可以找到某个方向，使得这个方向的直线上，样本的投影能分开的最好，而 Fisher 法所要解决的基本问题就是找到这条最好的、最易于分类的投影线。

先从 d 维空间到一维空间的一维数学变换方法。假设有一集合 X 包含 N 个 d 维样本 $x_1, x_2, ..., x_N$，其中 N_1 个属于 ω_1 类的样本记为子集 X_1，N_2 个属于 ω_2 类的样本记为 X_2。若对 x_N 的分量做线性组合可得标量

$$y_n = w^T x_n，\quad n = 1,2,...,N_i$$

这样便得到 N 个一维样本 y_n 组成的集合，并可分为两个子集 Y_1 和 Y_2。w 的绝对值是无关紧要的，它仅使 y_n 乘上一个比例因子，重要的是选择 w 的方向，从而转化为寻找最好的投影方向 w^*，是样本分开。

2.基本方法：

先定义几个基本参量：

（1）各类样本均值向量 m_i

$$m_i = \frac{1}{N} \sum_{x \in X_i} x, i = 1,2$$

（2）样本类内离散度矩阵 S_i 和总类内离散度矩阵 S_ω

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

$$S_\omega = S_1 + S_2$$

（3）样本类间离散度矩阵 S_b

$$S_b = (m_1 - m_2)(m_1 - m_2)^T$$

我们希望投影后，在低维空间里个样本尽可能的分开些，即希望两类均值 $(m_1 - m_2)$ 越大越

好，同时希望各类样本内部尽量密集，即 S_i 越小越好。因此，我们定义 Fisher 准则函数为

$$J_F(w)=\frac{(m_1-m_2)^2}{S_1+S_2}$$

但 $J_F(w)$ 不显含 w，因此必须设法将 $J_F(w)$ 变成 w 的显函数。

由式子

$$m_i=\frac{1}{N_i}\sum_{y\in Y_i}y=\frac{1}{N_i}\sum_{x\in X_i}w^T x=w^T(\frac{1}{N_i}\sum_{x\in X_i}x)=w^T m_i$$

$$(m_1-m_2)^2=(w^T m_1-w^T m_2)^2=w^T(m_1-m_2)(m_1-m_2)^T w=w^T S_b w$$

$$S_i=\sum_{y\in Y_i}(y-m_i)^2=\sum_{y\in Y_i}(w^T x-w^T m_i)^2=w^T(x-m_i)(x-m_i)^T w=w^T S_i w$$

从而得到

$$J_F(w)=\frac{w^T S_b w}{w^T S_\omega w},$$

采用 Lagrange 乘子法求解它的极大值 w^*

$$L(w,\lambda)=w^T S_b w-\lambda(w^T S_\omega w-c)$$

对其求偏导，得 $S_b w^*-\lambda S_\omega w^*=0$，即

$$S_b w^*=\lambda S_\omega w^*$$

从而我们很容易得到

$$\lambda w^*=S_\omega^{-1}(S_b w^*)=S_\omega^{-1}(m_1-m_2)R,\text{其中}R=(m_1-m_2)^T w^*$$

$$w^*=\frac{R}{\lambda}S_\omega^{-1}(m_1-m_2)$$

忽略比例因子 R/λ，得

$$w^*=S_\omega^{-1}(m_1-m_2)$$

这就是我们 Fisher 准则函数 $J_F(w)$ 取极大值时的解。

三、实验内容：

依据实验基本原理和基本方法，对下面表 3-1 样本数据中的类别 ω_1 和 ω_2 计算最优方向 w，画出最优方向 w 的直线，并标记出投影后的点在直线上的位置。选择决策边界，实现新样本 xx1=(-0.7,0.58,0.089)，xx2=(0.047,-0.4,1.04)的分类。

设某新类别 ω_3 数据如表 3-2 所示，用自己的函数求新类别 ω_3 分别和 ω_1、ω_2 分类的投

影方向和分类阀值。

表 3-1 Fisher 线性判别实验数据

类别		1	2	3	4	5	6	7	8	9	10
ω_1	x1	-0.4	-0.31	-0.38	-0.15	-0.35	0.17	-0.011	-0.27	-0.065	-0.12
	x2	0.58	0.27	0.055	0.53	0.47	0.69	0.55	0.61	0.49	0.054
	x3	0.089	-0.04	-0.035	0.011	0.034	0.1	-0.18	0.12	0.0012	-0.063
ω_2	x1	0.83	1.1	-0.44	0.047	0.28	-0.39	0.34	-0.3	1.1	0.18
	x2	1.6	1.6	-0.41	-0.45	0.35	-0.48	-0.079	-0.22	1.2	-0.11
	x3	-0.014	0.48	0.32	1.4	3.1	0.11	0.14	2.2	-0.46	-0.49

表 3-2 新类别实验数据

类别		1	2	3	4	5	6	7	8	9	10
ω_3	x1	1.58	0.67	1.04	-1.49	-0.41	1.39	1.2	-0.92	0.45	-0.76
	x2	2.32	1.58	1.01	2.18	1.21	3.61	1.4	1.44	1.33	0.84
	x3	-5.8	-4.78	-3.63	-3.39	-4.73	2.87	-1.89	-3.22	-4.38	-1.96

四、实验程序及其说明：

1）Fisher 准则函数算法：

其中 w 为我们要找到的投影方向 w^*，w1、w2 是我们的样本 ω_1、ω_2，s1、s2 是相应样本的类内离散度矩阵 S_i，sw 是总类内离散度矩阵 S_ω，m1、m2 是相应样本均值。在进行 ω_3 分别和 ω_1、ω_2 分类的投影方向和分类阀值时，将对应的 s1、s2；sw；m1、m2 以及输出坐标换成相应的样本符号。由于代码除此之外均相同，没有必要再重复列出，只需在运行时修改上述值即可。

注意：在画出 w 时（即 w^*），由于样本的不同，输出系数应作相应的调整。如：

ω_1、ω_2 时，plot3(30*x,30*x*w(2,:)/w(1,:),30*x*w(3,:)/w(1,:),'k');

ω_1、ω_3 时，plot3(x,x*w(2,:)/w(1,:),x*w(3,:)/w(1,:),'k');

ω_2、ω_3 时，plot3(5*x,5*x*w(2,:)/w(1,:),5*x*w(3,:)/w(1,:),'k');

代码：

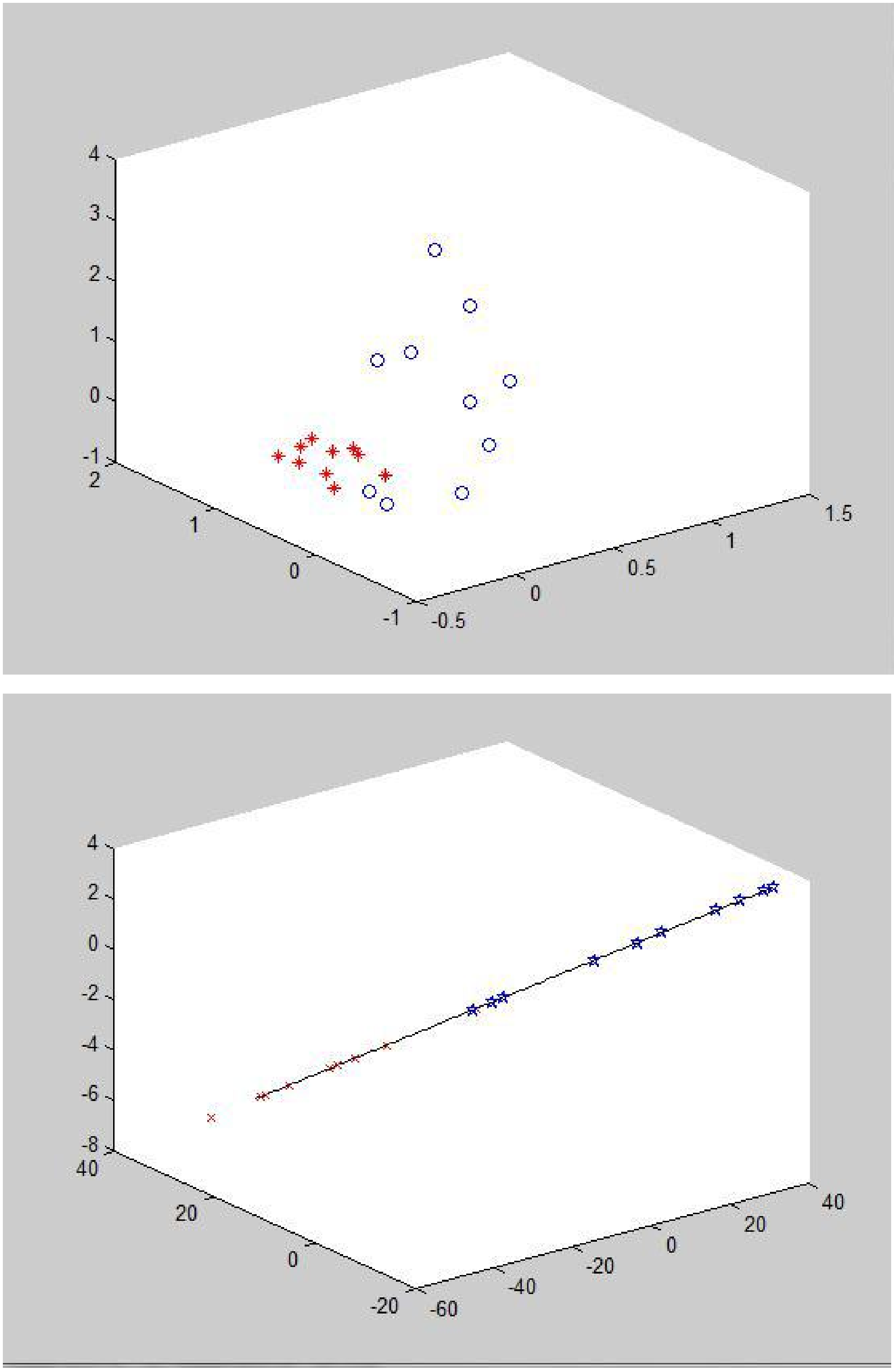
```
clear;
w1=[-0.4 0.58 0.089;-0.31 0.27 -0.04;-0.38 0.055 -0.035;-0.15 0.53 0.011;-0.35 0.47 0.034;
    -0.17 0.69 0.1;-0.011 0.55 -0.18;-0.27 0.61 0.12;-0.065 0.49 0.0012;-0.12 0.054 -0.063];
w2=[0.8 1.6 -0.014;1.1 1.6 0.48;-0.44 -0.41 0.32;0.047 -0.45 1.4;0.28 0.35 3.1;
    -0.39 -0.48 0.11;0.34 -0.079 0.14;-0.3 -0.22 2.2;1.1 1.2 -0.46;0.18 -0.11 -0.49];
w3=[1.58 2.32 -5.8;0.67 1.58 -4.78;1.04 1.01 -3.63;-1.49 2.18 -3.39;-0.41 1.21 -4.73;
    1.39 3.61 2.87;1.2 1.4 -1.89;-0.92 1.44  -3.22;0.45 1.33 -4.38;-0.76 0.84 -1.96];
```

```
xx1=[-0.7 0.58 0.089];
xx2=[0.047 -0.4 1.04];
s1=cov(w2,1);%样本类间离散度 S1
m1=mean(w2);%样本均值 m1
s2=cov(w3,1);%样本类间离散度 S2
m2=mean(w3);%样本均值 m2
sw=s1+s2;%总类内离散度 Sw
w=inv(sw)*(m1-m2)';%Fisher 准则函数 w*
h1=figure(1);
for i=1:1:10%打印样本
    plot3(w2(i,1),w2(i,2),w2(i,3),'r*');
    hold on;
    plot3(w3(i,1),w3(i,2),w3(i,3),'bo');
end;
figure(2)
%画出 fisher 判别函数
xmin=min(min(w2(:,1)),min(w3(:,1)));
xmax=max(max(w2(:,1)),max(w3(:,1)));
x=xmin-1:(xmax-xmin)/100:xmax;
plot3(5*x,5*x*w(2,:)/w(1,:),5*x*w(3,:)/w(1,:),'k');
hold on;
%将样本投影到 fisher 判别函数上
y1=w'*w2';%yn=w^*T * xn,n=1,2
y2=w'*w3';
figure(2)
for i=1:1:10
    plot3(y1(i)*w(1),y1(i)*w(2),y1(i)*w(3),'rx');
    hold on;
    plot3(y2(i)*w(1),y2(i)*w(2),y2(i)*w(3),'bp');
end;
```

五、实验结果及其说明：

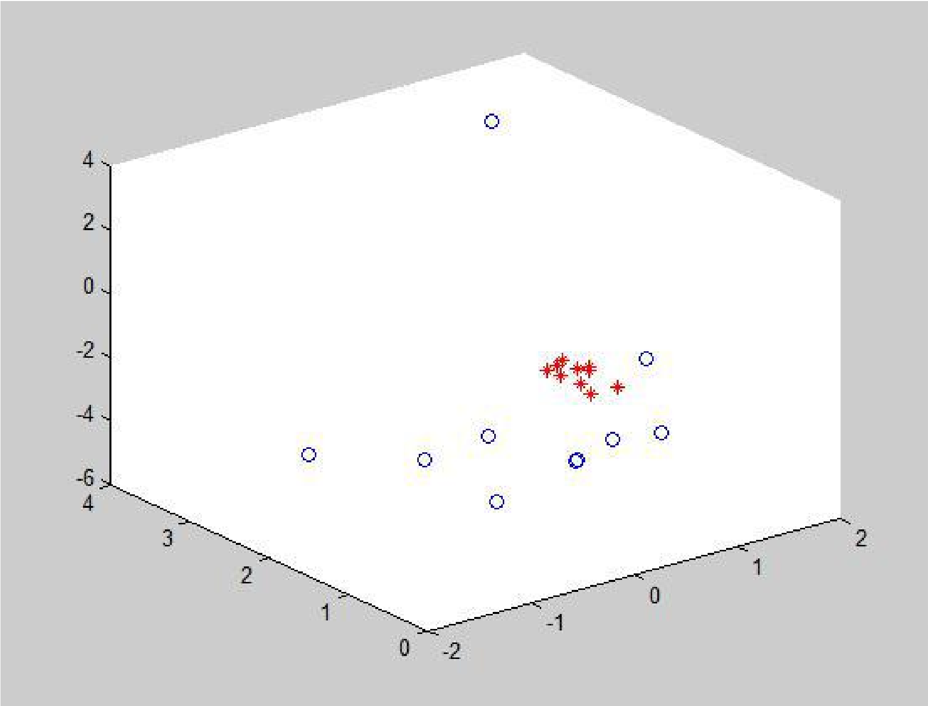
1）Fisher 线性判别算法输出结果：

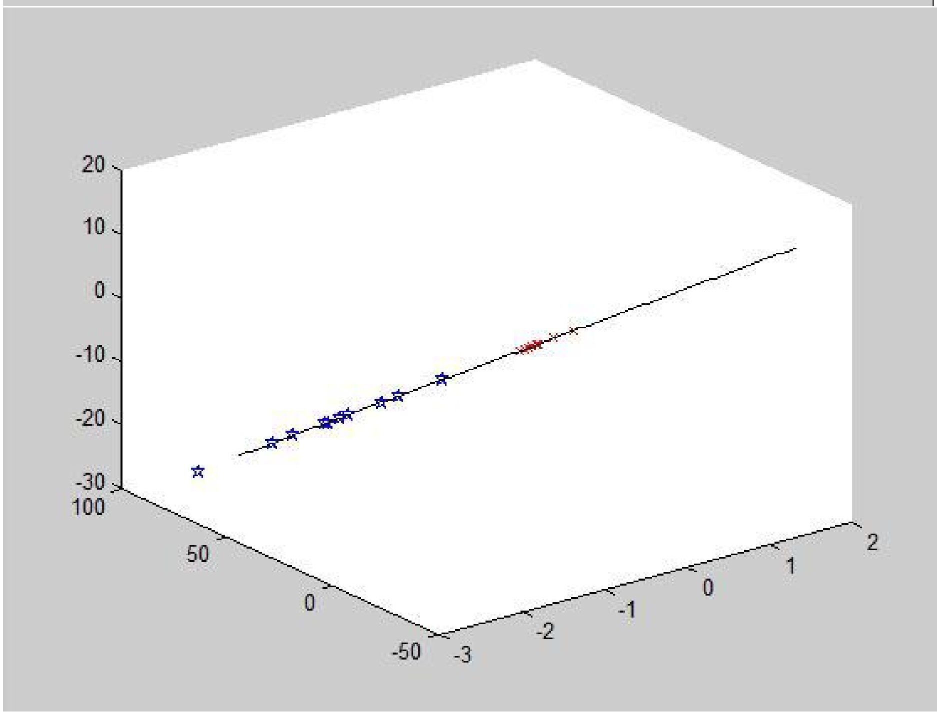
1.样本ω_1、ω_2



实验结果图 3-1

2.样本 ω_1 、 ω_3

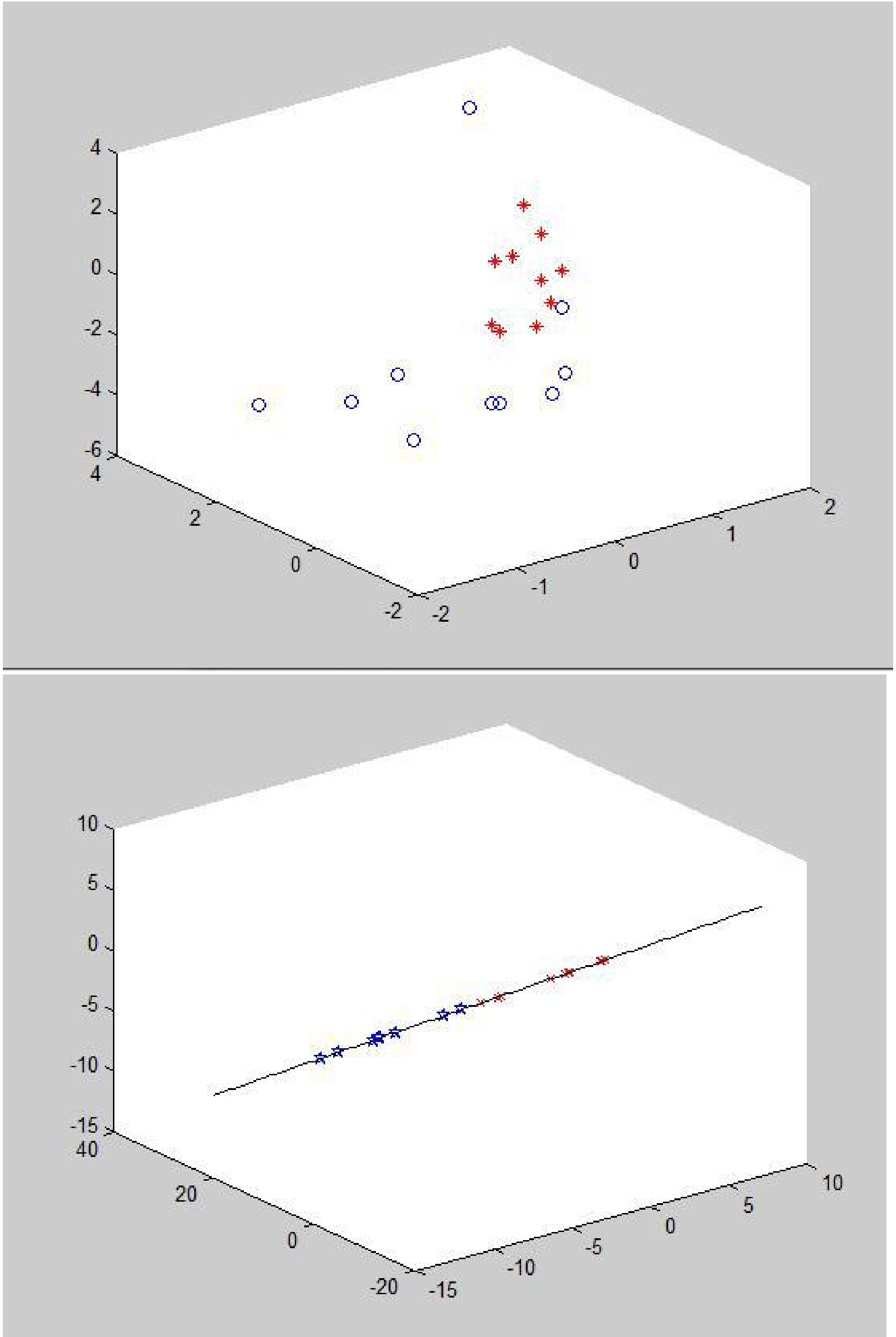




实验结果图 3-2

Authord: Vivid Xu;

3.样本 ω_2、ω_2



实验结果图 2-3

六、实验小结：

试验时阅读课本发现公式很多，中间推导式较复杂，以为会很难写代码。但参考例代码，并结合课本发现所需要的公式就那几个基本参量，以及最终推导出来的 w^*。为了画出最优方

向 w^*，我想起了高数里讲过三维空间直线 $\frac{x}{x_0}=\frac{y}{y_0}=\frac{z}{z_0}$，从而引导我利用此式，画出了 w^*，如上所示（ω_2、ω_3 时，plot3(5*x,5*x*w(2,:)/w(1,:),5*x*w(3,:)/w(1,:),'k');）。

特别值得说的是试验中 w=inv(sw)*(m1-m2)';后面是(m1-m2)'，而不是(m1-m2)，从推导式我们可以看出这无关紧要，只不过要将后续矩阵——y1=w'*w2';相应的求其转置，以使得矩阵可以相乘。最终我们可以得到 w^*，以使得样本 X 从 d 维空间投影到一维 Y 空间。