模式识别及应用

实验指导书

电子工程学院

目录

实验一　Bayes 分类器的设计（1）

一、　实验目的：

1. 对模式识别有一个初步的理解，能够根据自己的设计对贝叶斯决策理论算法有一个深刻地认识；

2. 理解基于最小错误率贝叶斯分类器的设计原理。

二、　实验条件：

PC 微机一台和 MATLAB 软件。

三、　实验原理：

最小错误率贝叶斯决策可按下列步骤进行：

1. 在已知 $P(\omega_i)$，$P(X\mid\omega_i)$，$i=1,\cdots,c$ 及给出待识别的 X 的情况下，根据贝叶斯公式计算出后验概率：

$$P(X\mid\omega_i)=\frac{P(X\mid\omega_i)P(\omega_i)}{\sum_{j=1}^{c}P(X\mid\omega_j)P(\omega_j)}\qquad j=1,\cdots,c$$

2. 对 1 中得到的 c 个后验概率进行比较，找出使后验概率最大的决策，即：

$$P(X\mid\omega_k)=\max_{i=1\ldots c}P(X\mid\omega_i),X\in\omega_k\ ,$$

则 ω_k 就是最小错误率贝叶斯决策。

四、　实验内容：

假定某个局部区域细胞识别中正常（ω_1）和非正常（ω_2）两类先验概率分别为：

正常状态：$P(\omega_1)$=0.9；

异常状态：$P(\omega_2)$=0.1。

现有一系列待观察的细胞，其观察值为 x：

-3.9847	-3.5549	-1.2401	-0.9780	-0.7932	-2.8531
-2.7605	-3.7287	-3.5414	-2.2692	-3.4549	-3.0752
-3.9934	2.8792	-0.9780	0.7932	1.1882	3.0682
-1.5799	-1.4885	-0.7431	-0.4221	-1.1186	4.2532

$P(x|\omega_1)$ $P(x|\omega_2)$ 类条件概率分布正态分布分别为（-2，0.5）（2,2）。

试利用基于最小错误率的贝叶斯准则对以上 24 个细胞进行分类，给出并观察分类的结果。

参考实验程序：

实验主程序如下：

函数：

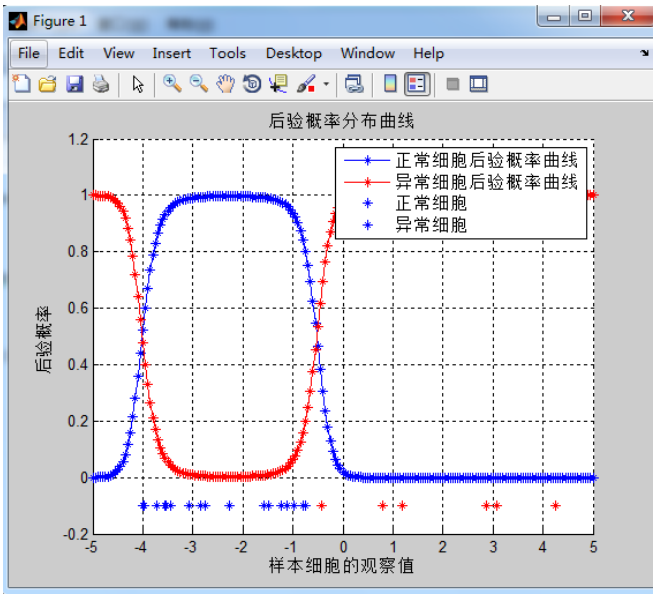
```
function results=bayes(x,pw1,pw2)


m=numel(x);%得到待测细胞数

pw1_x=zeros(1,m);%存放对w1的后验概率

pw2_x=zeros(1,m);%存放对w2的后验概率

results=zeros(1,m);%存放比较结果矩阵

e1=-2; a1=0.5;

e2=2;a2=2;


for i=1:m

pw1_x(i)=(pw1*normpdf(x(i),e1,a1))/(pw1*normpdf(x(i),e1,a1)+pw2*n

ormpdf(x(i),e2,a2));%计算w1下的后验概率

pw2_x(i)=(pw2*normpdf(x(i),e2,a2))/(pw1*normpdf(x(i),e1,a1)+pw2*n

ormpdf(x(i),e2,a2));%计算w2下的后验概率

end


for i=1:m

    if pw1_x(i)> pw2_x(i)%比较两类后验概率

        results(i)=0;%正常细胞

       xi=num2str(x(i));

        str=sprintf('x= %s 此细胞为正常细胞',xi);

        disp(str)

       else
```

```
        results(i)=1;%异常细胞
         xi=num2str(x(i));
         str=sprintf('x= %s 此细胞为异常细胞',xi);
        disp(str)
    end
end

a=[-5:0.05:5];
n=numel(a);
pw1_plot=zeros(1,n);
pw2_plot=zeros(1,n);
for j=1:n
pw1_plot(j)=(pw1*normpdf(a(j),e1,a1))/(pw1*normpdf(a(j),e1,a1)+pw
2*normpdf(a(j),e2,a2));
pw2_plot(j)=(pw2*normpdf(a(j),e2,a2))/(pw1*normpdf(a(j),e1,a1)+pw
2*normpdf(a(j),e2,a2));
end

figure(1)
hold on
plot(a,pw1_plot,'b*-',a,pw2_plot,'r*-')
for k=1:m
    if results(k)==0
        plot(x(k),-0.1,'b*')%正常细胞用*表示
    else
        plot(x(k),-0.1,'r*')%异常细胞用五角星表示
    end
end

%legend('正常细胞','异常细胞','Location','Best')
```

```
legend('正常细胞后验概率曲线','异常细胞后验概率曲线','正常细胞','异常细胞')
xlabel('样本细胞的观察值')
ylabel('后验概率')
title('后验概率分布曲线')
grid on
return
```

<u>程序运行</u>

```
x=[-3.9847,-3.5549,-1.2401,-0.9780,
    -0.7932,-2.8531,-2.7605,-3.7287,
-3.5414,-2.2692,-3.4549,-3.0752,
-3.9934,2.8792,-0.9780,0.7932,
1.1882,3.0682,-1.5799,-1.4885,
-0.7431,-0.4221,-1.1186,4.2532];
disp(x);
pw1=0.9;
pw2=0.1;
result=bayes(x,pw1,pw2);
disp('result中正常细胞为“0”，不正常细胞为“1”');
result
```



实验二　Bayes 分类器的设计（2）

一、　实验目的：

1. 对模式识别有一个初步的理解，能够根据自己的设计对贝叶斯决策理论算法有一个深刻地认识；

2. 理解基于最小风险贝叶斯分类器的设计原理。

二、　实验条件：

PC 微机一台和 MATLAB 软件。

三、　实验原理：

最小风险贝叶斯决策可按下列步骤进行：

1. 在已知 $P(\omega_i)$，$P(X\mid\omega_i)$，$i=1,\cdots,c$ 及给出待识别的 X 的情况下，根据贝叶斯公式计算出后验概率：

$$P(X\mid\omega_i)=\frac{P(X\mid\omega_i)P(\omega_i)}{\sum_{j=1}^{c}P(X\mid\omega_j)P(\omega_j)}\qquad j=1,\cdots,c$$

2. 利用计算出的后验概率及决策表，按下式计算出采取 α_i 决策的条件风险：

$$R(\alpha_i\mid X)=\sum_{j=1}^{c}\lambda(\alpha_i,\omega_j)P(\omega_j\mid X)\qquad i=1,\cdots,a$$

3. 对 2 中得到的 a 个条件风险值 $R(\alpha_i\mid X)$（$i=1,\cdots,a$）进行比较，找出使条件风险最小的决策 α_k，即：

$$R(\alpha_k\mid X)=\min_{i=1,\cdots,c}R(\alpha_k\mid X),$$

则 α_k 就是最小风险贝叶斯决策。

四、　实验内容：

假定某个局部区域细胞识别中正常（ω_1）和非正常（ω_2）两类先验概率分别为：

正常状态：$P(\omega_1)$=0.9；

异常状态：$P(\omega_2)$=0.1。

现有一系列待观察的细胞，其观察值为x：

-3.9847	-3.5549	-1.2401	-0.9780	-0.7932	-2.8531
-2.7605	-3.7287	-3.5414	-2.2692	-3.4549	-3.0752
-3.9934	2.8792	-0.9780	0.7932	1.1882	3.0682
-1.5799	-1.4885	-0.7431	-0.4221	-1.1186	4.2532

$P(x|\omega_1)$ $P(x|\omega_2)$类条件概率分布正态分布分别为（-2，0.5）（2,2）。决策表为$\lambda_{11}=0$(λ_{11}表示$\lambda(\alpha_i,\omega_j)$的简写)，$\lambda_{12}$=2，$\lambda_{21}$=4，$\lambda_{22}$=0。

试利用基于最小风险的贝叶斯决策准则对以上 24 个细胞进行分类，写出并观察分类结果。

与实验一中利用基于最小错误率的贝叶斯准则所得分类结果相比较，分析得到两种不同分类结果的原因。

参考实验程序：

实验主程序如下：

函数：

```
function [R1_x,R2_x,result]=risk(x,pw1,pw2)


m=numel(x);%得到待测细胞数

R1_x=zeros(1,m);%存放把样本 X 判为正常细胞所造成的整体损失

R2_x=zeros(1,m);%存放把样本 X 判为异常细胞所造成的整体损失

result=zeros(1,m);%存放比较结果

e1=-2; a1=0.5;

e2=2;a2=2;%类条件概率正态分布的参数

r11=0;

r12=2;

r21=4;

r22=0; %风险决策表
```

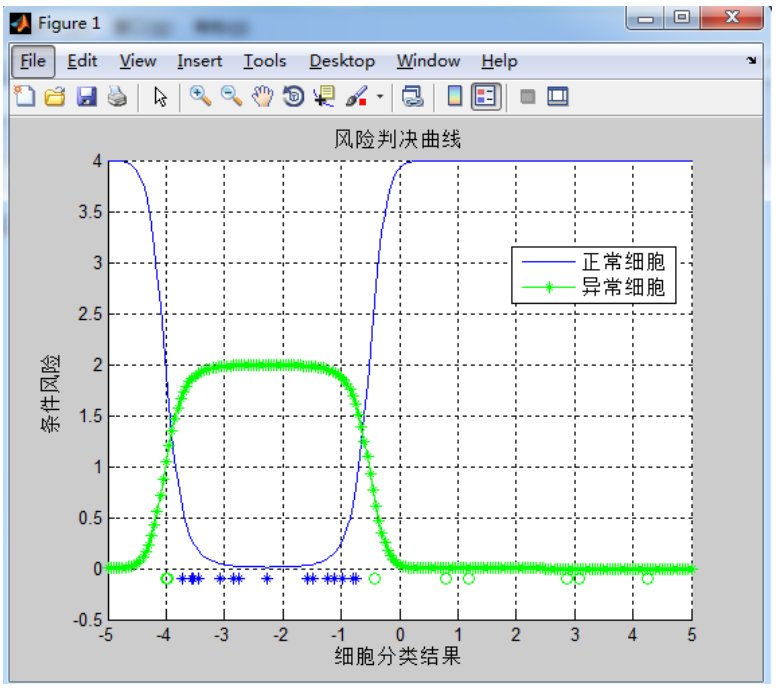
```
for i=1: m %计算两类风险值
R1_x(i)=(r11*pw1*normpdf(x(i),e1,a1))/(pw1*normpdf(x(i),e1,a1)+pw2*normpdf(x(i),e2,a2))+r2
1*pw2*normpdf(x(i),e2,a2)/(pw1*normpdf(x(i),e1,a1)+pw2*normpdf(x(i),e2,a2));

R2_x(i)=(r12*pw1*normpdf(x(i),e1,a1))/(pw1*normpdf(x(i),e1,a1)+pw2*normpdf(x(i),e2,a2))+r2
2*pw2*normpdf(x(i),e2,a2)/(pw1*normpdf(x(i),e1,a1)+pw2*normpdf(x(i),e2,a2));
end

for i=1:m
    if R2_x(i)> R1_x(i)%第二类比第一类风险大
        result(i)=0;%判为正常细胞（损失较小），用 0 表示
          xi=num2str(x(i));
        str=sprintf('x= %s  此细胞为正常细胞',xi);
        disp(str)
    else
        result(i)=1;%判为异常细胞，用 1 表示
          xi=num2str(x(i));
          str=sprintf('x= %s  此细胞为异常细胞',xi);
        disp(str)
    end
end

a=[-5:0.05:5];%取样本点以画图
n=numel(a);
R1_plot=zeros(1,n);
R2_plot=zeros(1,n);

for j=1:n

R1_plot(j)=(r11*pw1*normpdf(a(j),e1,a1))/(pw1*normpdf(a(j),e1,a1)+pw2*normpdf(a(j),e2,a2))
```

```
+r21*pw2*normpdf(a(j),e2,a2)/(pw1*normpdf(a(j),e1,a1)+pw2*normpdf(a(j),e2,a2));

R2_plot(j)=(r22*pw2*normpdf(a(j),e2,a2))/(pw1*normpdf(a(j),e1,a1)+pw2*normpdf(a(j),e2,a2))
+r12*pw1*normpdf(a(j),e1,a1)/(pw1*normpdf(a(j),e1,a1)+pw2*normpdf(a(j),e2,a2));
end%计算各样本点的风险以画图

figure(1)
hold on
plot(a,R1_plot,'b-',a,R2_plot,'g*-')
for k=1:m
    if result(k)==0
        plot(x(k),-0.1,'b*')%正常细胞用*表示
    else
        plot(x(k),-0.1,'go')%异常细胞用圆表示
    end
end

legend('正常细胞','异常细胞','Location','Best')
xlabel('细胞分类结果')
ylabel('条件风险')
title('风险判决曲线')
grid on
return
```

程序运行

```
x=[-3.9847,-3.5549,-1.2401,-0.9780,
    -0.7932,-2.8531,-2.7605,-3.7287,
-3.5414,-2.2692,-3.4549,-3.0752,
-3.9934,2.8792,-0.9780,0.7932,
1.1882,3.0682,-1.5799,-1.4885,
```

```
-0.7431,-0.4221,-1.1186,4.2532];
disp(x);
pw1=0.9;
pw2=0.1;
[R1_x,R2_x,result]=risk(x,pw1,pw2);
disp('result中正常细胞为“0”，不正常细胞为“1”');
result
```



实验三 基于 Fisher 准则的线性分类器设计

一、 实验目的：

1. 进一步了解分类器的设计概念，能够根据自己的设计对线性分类器有更深刻地认识；

2. 理解 Fisher 准则方法确定最佳线性分界面方法的原理，以及拉格朗日乘子求解的原理。

二、 实验条件：

PC 微机一台和 MATLAB 软件。

三、 实验原理：

设有一个集合包含 N 个 d 维样本 $x_1, x_2, \cdots, x_N$，其中 N_1 个属于 ω_1 类，N_2 个属于 ω_2 类。线性判别函数的一般形式可表示成 $g(x) = W^T x + w_0$，其中

$$W = (w_1, \cdots, w_d)^T \text{。}$$

根据 Fisher 选择投影方向 W 的原则，即使原样本向量在该方向上的投影能兼顾类间分布尽可能分开，类内样本投影尽可能密集的要求，用以评价投影方向 W 的函数为：

$$J_F(W) = \frac{W^T S_b W}{W^T S_w W}$$

$$W^* = S_W^{-1}(m_1 - m_2)^T$$

其中：

$m_i = \frac{1}{N_i}\sum_{j=1}^{N_i} x_j \quad i = 1,2$ 　　x_j 为 N_i 类中的第 j 个样本

S_w 为类内离散度，定义为：$S_w = \sum_{i=1}^{2}\sum_{j=1}^{N_i}(x_j - m_1)(x_j - m_i)^T$

S_b 为类间离散度，定义为：$S_b = (m_1 - m_2)(m_1 - m_2)^T$

上面的公式是使用 Fisher 准则求最佳法线向量的解，我们称这种形式的运算为线性变换，其中 $(m_1 - m_2)$ 是一个向量，S_W^{-1} 是 S_W 的逆矩阵，如 $(m_1 - m_2)$ 是 d 维，

S_W^{-1} 和 S_W 都是 $d \times d$ 维，得到的 W^* 也是一个 d 维的向量。

向量 W^* 就是使 Fisher 准则函数 $J_F(W)$ 达极大值的解，也就是按 Fisher 准则将 d 维 X 空间投影到一维 Y 空间的最佳投影方向，该向量 W^* 的各分量值是对原 d 维特征向量求加权和的权值。

以上讨论了线性判别函数加权向量 W 的确定方法，并讨论了使 Fisher 准则函数极大的 d 维向量 W^* 的计算方法，但是判别函数中的另一项 w_0 尚未确定，一般可采用以下几种方法确定 w_0 如

$$w_0 = -\frac{W^{*T}(m_1 + m_2)}{2}$$

或者 $w_0 = -\dfrac{W^{*T}(N_1 m_1 + N_2 m_2)}{N_1 + N_2}$

或当 $P(\omega_1)$ 与 $P(\omega_2)$ 已知时可用

$$w_0 = -[\frac{W^{*T}(m_1 + m_2)}{2} - \frac{\ln[P(\omega_1)/P(\omega_2)]}{N_1 + N_2 - 2}]$$

当 w_0 确定之后，则可按以下规则分类，

$$W^{*T} X > -w_0 \rightarrow x \in \omega_1$$

$$W^{*T} X < -w_0 \rightarrow x \in \omega_2$$

四、 实验内容：

已知有两类数据 ω_1 和 ω_2 二者的概率已知 $P(\omega_1)$=0.6， $P(\omega_2)$=0.4。

ω_1 中数据点的坐标对应一一如下：

$x1$ =	0.2331	1.5207	0.6499	0.7757	1.0524	1.1974
	0.2908	0.2518	0.6682	0.5622	0.9023	0.1333
	-0.5431	0.9407	-0.2126	0.0507	-0.0810	0.7315
	0.3345	1.0650	-0.0247	0.1043	0.3122	0.6655
	0.5838	1.1653	1.2653	0.8137	-0.3399	0.5152
	0.7226	-0.2015	0.4070	-0.1717	-1.0573	-0.2099
x2 =	2.3385	2.1946	1.6730	1.6365	1.7844	2.0155
	2.0681	2.1213	2.4797	1.5118	1.9692	1.8340

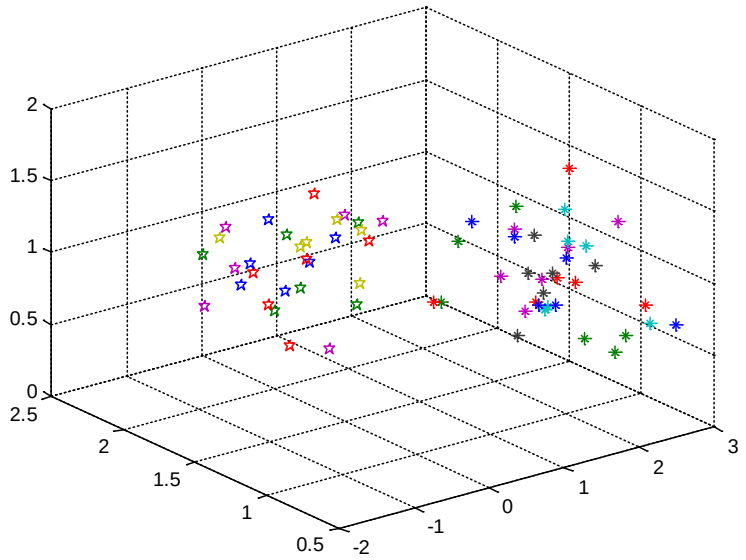
1.8704	2.2948	1.7714	2.3939	1.5648	1.9329
2.2027	2.4568	1.7523	1.6991	2.4883	1.7259
2.0466	2.0226	2.3757	1.7987	2.0828	2.0798
1.9449	2.3801	2.2373	2.1614	1.9235	2.2604

x3 =	0.5338	0.8514	1.0831	0.4164	1.1176	0.5536
	0.6071	0.4439	0.4928	0.5901	1.0927	1.0756
	1.0072	0.4272	0.4353	0.9869	0.4841	1.0992
	1.0299	0.7127	1.0124	0.4576	0.8544	1.1275
	0.7705	0.4129	1.0085	0.7676	0.8418	0.8784
	0.9751	0.7840	0.4158	1.0315	0.7533	0.9548

ω_2 数据点的对应的三维坐标为:

x4=	1.4010	1.2301	2.0814	1.1655	1.3740	1.1829
	1.7632	1.9739	2.4152	2.5890	2.8472	1.9539
	1.2500	1.2864	1.2614	2.0071	2.1831	1.7909
	1.3322	1.1466	1.7087	1.5920	2.9353	1.4664
	2.9313	1.8349	1.8340	2.5096	2.7198	2.3148
	2.0353	2.6030	1.2327	2.1465	1.5673	2.9414
x5=	1.0298	0.9611	0.9154	1.4901	0.8200	0.9399
	1.1405	1.0678	0.8050	1.2889	1.4601	1.4334
	0.7091	1.2942	1.3744	0.9387	1.2266	1.1833
	0.8798	0.5592	0.5150	0.9983	0.9120	0.7126
	1.2833	1.1029	1.2680	0.7140	1.2446	1.3392
	1.1808	0.5503	1.4708	1.1435	0.7679	1.1288
x6=	0.6210	1.3656	0.5498	0.6708	0.8932	1.4342
	0.9508	0.7324	0.5784	1.4943	1.0915	0.7644
	1.2159	1.3049	1.1408	0.9398	0.6197	0.6603
	1.3928	1.4084	0.6909	0.8400	0.5381	1.3729
	0.7731	0.7319	1.3439	0.8142	0.9586	0.7379
	0.7548	0.7393	0.6739	0.8651	1.3699	1.1458

数据的样本点分布如下图：



根据所得结果判断（1，1.5，0.6），(1.2，1.0，0.55)，(2.0，0.9，0.68)，(1.2，1.5，0.89)，（0.23，2.33，1.43），分别属于哪个类别，并画出数据分类相应的结果图，要求画出其在W上的投影，并写出使Fisher准则函数 $J_F(W)$ 达极大值的解W。

参考实验程序：

实验主程序如下：

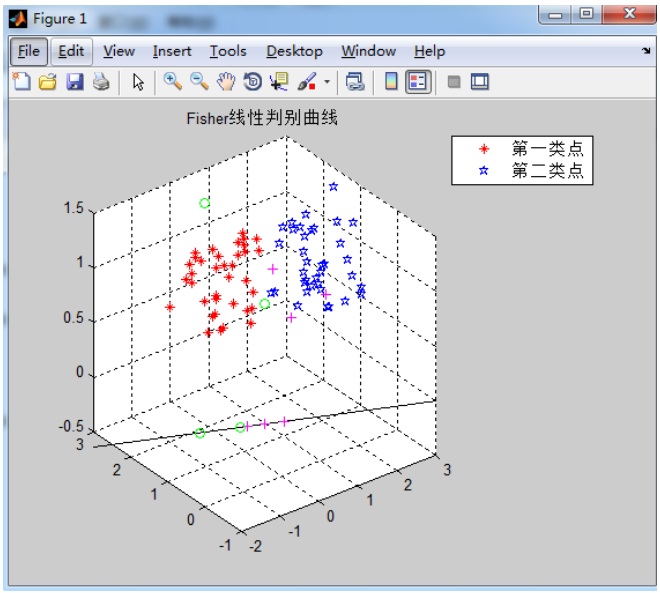
```
x1=[0.2331    1.5207    0.6499    0.7757    1.0524    1.1974
    0.2908    0.2518    0.6682    0.5622    0.9023    0.1333
   -0.5431    0.9407   -0.2126    0.0507   -0.0810    0.7315
    0.3345    1.0650   -0.0247    0.1043    0.3122    0.6655
    0.5838    1.1653    1.2653    0.8137   -0.3399    0.5152
    0.7226   -0.2015    0.4070   -0.1717   -1.0573   -0.2099];
x2=[2.3385    2.1946    1.6730    1.6365    1.7844    2.0155
    2.0681    2.1213    2.4797    1.5118    1.9692    1.8340
    1.8704    2.2948    1.7714    2.3939    1.5648    1.9329
    2.2027    2.4568    1.7523    1.6991    2.4883    1.7259
    2.0466    2.0226    2.3757    1.7987    2.0828    2.0798
    1.9449    2.3801    2.2373    2.1614    1.9235    2.2604];
x3=[0.5338    0.8514    1.0831    0.4164    1.1176    0.5536
    0.6071    0.4439    0.4928    0.5901    1.0927    1.0756
    1.0072    0.4272    0.4353    0.9869    0.4841    1.0992
    1.0299    0.7127    1.0124    0.4576    0.8544    1.1275
    0.7705    0.4129    1.0085    0.7676    0.8418    0.8784
    0.9751    0.7840    0.4158    1.0315    0.7533    0.9548];
```

```
%将 x1、x2、x3 变为行向量
x1=x1(:);
x2=x2(:);
x3=x3(:);
%计算第一类样本均值向量 m1
m1(1)=mean(x1);
m1(2)=mean(x2);
m1(3)=mean(x3);

%计算第一类样本类内离散度矩阵 S1
S1=zeros(3,3);
for i=1:36
    S1=S1+[-m1(1)+x1(i)    -m1(2)+x2(i)    -m1(3)+x3(i)]'*[-m1(1)+x1(i)    -m1(2)+x2(i)
-m1(3)+x3(i)];
end

%w2 的数据点坐标

  x4=[1.4010    1.2301    2.0814    1.1655    1.3740    1.1829
     1.7632    1.9739    2.4152    2.5890    2.8472    1.9539
     1.2500    1.2864    1.2614    2.0071    2.1831    1.7909
     1.3322    1.1466    1.7087    1.5920    2.9353    1.4664
     2.9313    1.8349    1.8340    2.5096    2.7198    2.3148
     2.0353    2.6030    1.2327    2.1465    1.5673    2.9414];
  x5=[1.0298    0.9611    0.9154    1.4901    0.8200    0.9399
     1.1405    1.0678    0.8050    1.2889    1.4601    1.4334
     0.7091    1.2942    1.3744    0.9387    1.2266    1.1833
     0.8798    0.5592    0.5150    0.9983    0.9120    0.7126
     1.2833    1.1029    1.2680    0.7140    1.2446    1.3392
     1.1808    0.5503    1.4708    1.1435    0.7679    1.1288];
  x6=[0.6210    1.3656    0.5498    0.6708    0.8932    1.4342
     0.9508    0.7324    0.5784    1.4943    1.0915    0.7644
     1.2159    1.3049    1.1408    0.9398    0.6197    0.6603
     1.3928    1.4084    0.6909    0.8400    0.5381    1.3729
     0.7731    0.7319    1.3439    0.8142    0.9586    0.7379
     0.7548    0.7393    0.6739    0.8651    1.3699    1.1458];
x4=x4(:);
x6=x6(:);
x5=x5(:);
%计算第二类样本均值向量 m2
m2(1)=mean(x4);
m2(2)=mean(x5);
m2(3)=mean(x6);
```

```
%计算第二类样本类内离散度矩阵 S2

S2=zeros(3,3);

for i=1:36
S2=S2+[-m2(1)+x4(i) -m2(2)+x5(i) -m2(3)+x6(i)]'*[-m2(1)+x4(i) -m2(2)+x5(i) -m2(3)+x6(i)];
end

%总类内离散矩阵 Sw
Sw=zeros(3,3);
Sw=S1+S2;

%样本类间离散度矩阵 Sb
Sb=zeros(3,3);
Sb=(m1-m2)'*(m1-m2);

%最优解 W
W=Sw^-1*(m1-m2);

%将 W 变为单位向量以方便计算投影
W=W/sqrt(sum(W.^2));

%计算一维 Y 空间中的各类样本均值 M1 及 M2

for i=1:36
    y(i)=W'*[x1(i) x2(i) x3(i)]';
end
M1=mean(y);
for i=1:36
    y(i)=W'*[x4(i) x5(i) x6(i)]';
end
M2=mean(y);

%利用当 P(W1)与 P(W2)已知时的公式计算 W0
p1=0.6;p2=0.4;

W0=-(M1+M2)/2+(log(p2/p1))/(36+36-2);

%计算将样本投影到最佳方向上以后的新坐标

X1=[x1*W(1)+x2*W(2)+x3*W(3)]';     %将样本投影到最佳方向的投影长度
X2=[x4*W(1)+x5*W(2)+x6*W(3)]';
XX1=[W(1)*X1;W(2)*X1;W(3)*X1];     %将样本投影到最佳方向的新坐标
```

```
XX2=[W(1)*X2;W(2)*X2;W(3)*X2];


%绘制样本点
figure(1);
plot3(x1,x2,x3,'r*'); %第一类
hold on
plot3(x4,x5,x6,'bp');%第二类
legend('第一类点','第二类点');
title('Fisher 线性判别曲线');
W1=5*W;

%画出最佳方向
line([-W1(1),W1(1)],[-W1(2),W1(2)],[-W1(3),W1(3)],'color','b');

%判别已给点的分类
a1=[1,1.5,0.6]';a2=[1.2,1.0,0.55]';a3=[2.0,0.9,0.68]';
a4=[1.2,1.5,0.89]';a5=[0.23,2.33,1.43]';

A=[a1 a2 a3 a4 a5];
n=size(A,2);


%下面代码在改变样本时都不必修改

%绘制待测数据投影到最佳方向上的点
for k=1:n
    A1=A(:,k)'*W;
    A11=W*A1; %得到待测数据投影
y=W'*A(:,k)+W0;      %计算后与 0 相比以判断类别，大于 0 为第一类，小于 0 为第二类
    if y>0
        plot3(A(1,k),A(2,k),A(3,k),'go');%点为‘rp’对应第一类
        plot3(A11(1),A11(2),A11(3),'go');%投影为‘r+’对应 go 类

        x=A(:,k)
        disp('x 属于第一类')
    else
        plot3(A(1,k),A(2,k),A(3,k),'m+'); %点为‘bh’对应 m+类
        plot3(A11(1),A11(2),A11(3),'m+'); %投影为‘b*’对应 m+类

        x=A(:,k)
        disp('x 属于第二类')
    end
end
```

```
%画出最佳方向
 line([-W1(1),W1(1)],[-W1(2),W1(2)],[-W1(3),W1(3)],'color','k');
view([-37.5,30]);
axis([-2,3,-1,3,-0.5,1.5]);
grid on
hold off
```



图中，红色的*是给出的第一类样本点，蓝色的五角星是给出的第二类样本点。下方的实直线是最佳投影方向，待测数据投影在其上，绿色的圆圈是被分为第一类的待测样本点，十字是被分为第二类的待测样本点。

实验四 基于感知函数准则线性分类器设计

一、 实验目的：

本实验旨在让同学理解感知准则函数的原理，通过软件编程模拟线性分类器，理解感知函数准则的确定过程，掌握梯度下降算法求增广权向量，进一步深刻认识线性分类器。

二、 实验条件：

PC 微机一台和 MATLAB 软件。

三、 实验原理：

1.假设已知一组容量为 N 的样本集 y_1， y_2，…， y_N，其中 y_N 为 d 维增广样本向量，分别来自 ω_1 和 ω_2 类。如果有一个线性机器能把每个样本正确分类，即存在一个权向量 a，使得对于任何 $y \in \omega_1$，都有 $a^T y$>0，而对一任何 $y \in \omega_2$，都有 $a^T y$<0，则称这组样本集线性可分；否则称线性不可分。若线性可分，则必存在一个权向量 a，能将每个样本正确分类。

2.基本方法：

由上面原理可知，样本集 y_1， y_2，…， y_N 是线性可分，则必存在某个权向量 a，使得 $\begin{cases} a^T y_i > 0, \text{对一切} \mathrm{y}_i \in \omega_1 \\ a^T y_j < 0, \text{对一切} \mathrm{y}_j \in \omega_2 \end{cases}$

如果我们在来自 ω_2 类的样本 y_j 前面加上一个负号，即令 y_j =— y_j，其中 $y_j \in \omega_2$，则也有 $a^T y$>0。因此，我们令

$$y'_n = \begin{cases} y_i, \text{对一切} \mathrm{y}_i \in \omega_1 \\ -y_j, \text{对一切} \mathrm{y}_j \in \omega_2 \end{cases}$$

那么，我们就可以不管样本原来的类型标志，只要找到一个对全部样本 y'_n 都满足 $a^T y$>0， $n = 1,2,3,\ldots\ldots,N$ 的权向量 a 就行了。此过程称为样本的规范化， y'_n 成为规范化增广样本向量，后面我们用 y 来表示它。

我们的目的是找到一个解向量 a^*，使得

$$a^T y_n > 0, n = 1,2,\ldots,N$$

为此我们首先考虑处理线性可分问题的算法。先构造这样一个准则函数

$$J_p(a)=\sum_{y\in\gamma^k}(\delta_y a^T y)$$

式中γ^k是被权向量a错分类的样本集合。δ_y的取值保证因此$J_p(a)$总是大于等于0。即错分类时有$a^T y\le 0\,(y\in\omega_1)$，$a^T y\ge 0\,(y\in\omega_2)$，此时的$\delta_y$分别为-1,1。

下一步便是求解使代价函数$J_p(a)$达到极小值时的解向量a^*。这里我们采用梯度下降法。首先对a求梯度，这是一个纯量函数对向量的求导问题，不难看出

$$\nabla J_p(a)=\frac{\partial J_p(a)}{\partial a}=\sum_{y\in\gamma^k}(\delta_y y)$$

梯度下降法的迭代公式为$a(k+1)=a(k)-\rho_k\nabla J$，将上式代入得

$$a(k+1)=a(k)-\rho_k\sum_{y\in\gamma^k}\delta_y y$$

这样，经过有限次修改，一定能找到一个解向量a^*。其中算法在运行之前必须人为任意给定权向量$a(1)$，而常量ρ_k的选取也十分讲究。

四、 实验内容：

1、 用 matlab 完成感知准则函数确定程序的设计。

2、 对所提供的的数据，采用感知机算法设计分类器，画出决策面。

3、 分别取$\rho_k=0.001,\ 0.01,\ 0.5,\ 1,\ 5,\ 10$，以上 6 个不同值时，观察分类结果，讨论算法中参数ρ_k设置对分类结果的影响。

4、 根据实验结果请说明感知准则函数是否是唯一的，为什么？

参考实验程序及部分结果：

1数据读取子函数

```
function [x1,x2]=getts()%分别读取分类 x1 和 x2
x1=zeros(45,2);
x2=zeros(55,2);
x1(1,1)=5.1418; x1(1,2)=0.5950;
x1(2,1)=5.5519; x1(2,2)=3.5091;
x1(3,1)=5.3836; x1(3,2)=2.8033;
x1(4,1)=3.2419; x1(4,2)=3.7278;
x1(5,1)=4.4427; x1(5,2)=3.8981;
x1(6,1)=4.9111; x1(6,2)=2.8710;
x1(7,1)=2.9259; x1(7,2)=3.4879;
x1(8,1)=4.2018; x1(8,2)=2.4973;
x1(9,1)=4.7629; x1(9,2)=2.5163;
x1(10,1)=2.7118; x1(10,2)=2.4264;
x1(11,1)=3.0470; x1(11,2)=1.5699;
x1(12,1)=4.7782; x1(12,2)=3.3504;
x1(13,1)=3.9937; x1(13,2)=4.8529;
x1(14,1)=4.5245; x1(14,2)=2.1322;
```

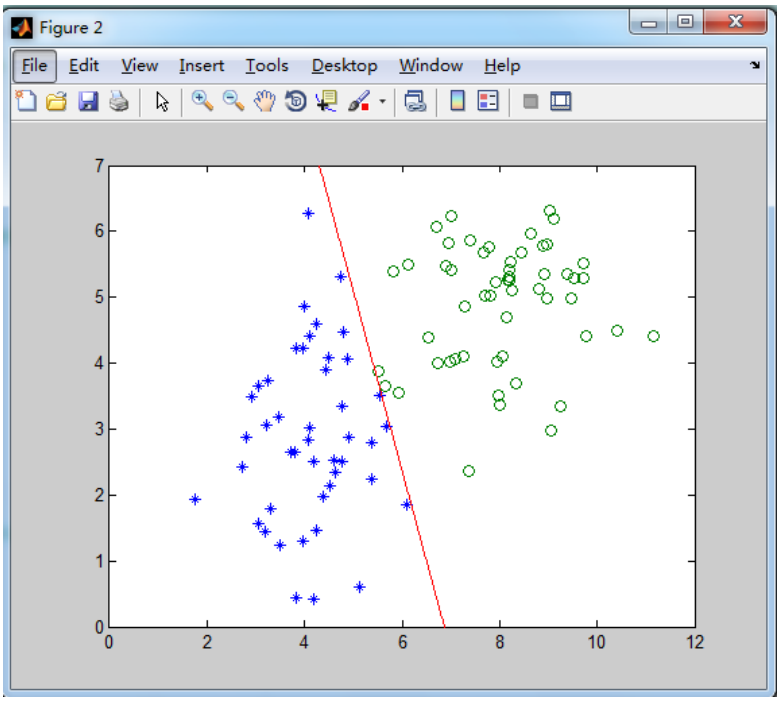
x1(15,1)=5.3643; x1(15,2)=2.2477;
x1(16,1)=4.4820; x1(16,2)=4.0843;
x1(17,1)=3.2129; x1(17,2)=3.0592;
x1(18,1)=4.7520; x1(18,2)=5.3119;
x1(19,1)=3.8331; x1(19,2)=0.4484;
x1(20,1)=3.1838; x1(20,2)=1.4494;
x1(21,1)=6.0941; x1(21,2)=1.8544;
x1(22,1)=4.0802; x1(22,2)=6.2646;
x1(23,1)=3.0627; x1(23,2)=3.6474;
x1(24,1)=4.6357; x1(24,2)=2.3344;
x1(25,1)=5.6820; x1(25,2)=3.0450;
x1(26,1)=4.5936; x1(26,2)=2.5265;
x1(27,1)=4.7902; x1(27,2)=4.4668;
x1(28,1)=4.1053; x1(28,2)=3.0274;
x1(29,1)=3.8414; x1(29,2)=4.2269;
x1(30,1)=4.8709; x1(30,2)=4.0535;
x1(31,1)=3.8052; x1(31,2)=2.6531;
x1(32,1)=4.0755; x1(32,2)=2.8295;
x1(33,1)=3.4734; x1(33,2)=3.1919;
x1(34,1)=3.3145; x1(34,2)=1.8009;
x1(35,1)=3.7316; x1(35,2)=2.6421;
x1(36,1)=2.8117; x1(36,2)=2.8658;
x1(37,1)=4.2486; x1(37,2)=1.4651;
x1(38,1)=4.1025; x1(38,2)=4.4063;
x1(39,1)=3.9590; x1(39,2)=1.3024;
x1(40,1)=1.7524; x1(40,2)=1.9339;
x1(41,1)=3.4892; x1(41,2)=1.2457;
x1(42,1)=4.2492; x1(42,2)=4.5982;
x1(43,1)=4.3692; x1(43,2)=1.9794;
x1(44,1)=4.1792; x1(44,2)=0.4113;
x1(45,1)=3.9627; x1(45,2)=4.2198;

x2(1,1)=9.7302; x2(1,2)=5.5080;
x2(2,1)=8.8067; x2(2,2)=5.1319;
x2(3,1)=8.1664; x2(3,2)=5.2801;
x2(4,1)=6.9686; x2(4,2)=4.0172;
x2(5,1)=7.0973; x2(5,2)=4.0559;
x2(6,1)=9.4755; x2(6,2)=4.9869;
x2(7,1)=9.3809; x2(7,2)=5.3543;
x2(8,1)=7.2704; x2(8,2)=4.1053;
x2(9,1)=8.9674; x2(9,2)=5.8121;
x2(10,1)=8.2606; x2(10,2)=5.1095;
x2(11,1)=7.5518; x2(11,2)=7.7316;
x2(12,1)=7.0016; x2(12,2)=5.4111;
x2(13,1)=8.3442; x2(13,2)=3.6931;
x2(14,1)=5.8173; x2(14,2)=5.3838;
x2(15,1)=6.1123; x2(15,2)=5.4995;
x2(16,1)=10.4188; x2(16,2)=4.4892;
x2(17,1)=7.9136; x2(17,2)=5.2349;
x2(18,1)=11.1547; x2(18,2)=4.4022;
x2(19,1)=7.7080; x2(19,2)=5.0208;
x2(20,1)=8.2079; x2(20,2)=5.4194;
x2(21,1)=9.1078; x2(21,2)=6.1911;
x2(22,1)=7.7857; x2(22,2)=5.7712;
x2(23,1)=7.3740; x2(23,2)=2.3558;
x2(24,1)=9.7184; x2(24,2)=5.2854;
x2(25,1)=6.9559; x2(25,2)=5.8261;
x2(26,1)=8.9691; x2(26,2)=4.9919;
x2(27,1)=7.3872; x2(27,2)=5.8584;
x2(28,1)=8.8922; x2(28,2)=5.7748;
x2(29,1)=9.0175; x2(29,2)=6.3059;
x2(30,1)=7.0041; x2(30,2)=6.2315;
x2(31,1)=8.6396; x2(31,2)=5.9586;
x2(32,1)=9.2394; x2(32,2)=3.3455;
x2(33,1)=6.7376; x2(33,2)=4.0096;
x2(34,1)=8.4345; x2(34,2)=5.6852;
x2(35,1)=7.9559; x2(35,2)=4.0251;
x2(36,1)=6.5268; x2(36,2)=4.3933;
x2(37,1)=7.6699; x2(37,2)=5.6868;
x2(38,1)=7.8075; x2(38,2)=5.0200;
x2(39,1)=6.6997; x2(39,2)=6.0638;
x2(40,1)=5.6549; x2(40,2)=3.6590;
x2(41,1)=6.9086; x2(41,2)=5.4795;
x2(42,1)=7.9933; x2(42,2)=3.3660;
x2(43,1)=5.9318; x2(43,2)=3.5573;
x2(44,1)=9.5157; x2(44,2)=5.2938;
x2(45,1)=7.2795; x2(45,2)=4.8596;
x2(46,1)=5.5233; x2(46,2)=3.8697;
x2(47,1)=8.1331; x2(47,2)=4.7075;
x2(48,1)=9.7851; x2(48,2)=4.4175;
x2(49,1)=8.0636; x2(49,2)=4.1037;
x2(50,1)=8.1944; x2(50,2)=5.2486;
x2(51,1)=7.9677; x2(51,2)=3.5103;
x2(52,1)=8.2083; x2(52,2)=5.3135;
x2(53,1)=9.0586; x2(53,2)=2.9749;
x2(54,1)=8.2188; x2(54,2)=5.5290;
x2(55,1)=8.9064; x2(55,2)=5.3435;

2感知器算法

```
clear all;
[x1,x2]=getts();%读取训练样本 x1,x2
w=[5.1 1.9 -34]';%设置权向量初值
Dcost=[0;0;0];%用来存储代价函数的梯度和
i=0;n=0;
cost=1;%初始化循环条件
while(cost~=0)
      i=i+1;cost=0;
    for j=1:45
            if (w'*[x1(j,:) 1])>0
                  Dcost=Dcost+[x1(j,1:2) 1]';%  用来存储代价函数的梯度和
                    cost=cost+w'*[x1(j,1:2) 1]';%代价函数值
            end
      end
    for k=1:55
            if (w'*[x2(k,:) 1])<0
                  Dcost=Dcost-[x2(k,1:2) 1]'; %  用来存储代价函数的梯度和
                  cost=cost+w'*[x2(k,:) 1]'; %代价函数值
            end
    end
          w=w-0.001*Dcost/i;%权向量更新
          n=n+1;
end

for i=1:45 r1(i)=x1(i,1);end;
for i=1:45 r2(i)=x1(i,2);end;
for i=1:55 r3(i)=x2(i,1);end;
for i=1:55 r4(i)=x2(i,2);end;

figure(1);
plot(r1,r2,'*',r3,r4,'o');
title('the training set');

figure(2)
x=0:0.01:12;
y=-x*w(1)/w(2)-w(3)/w(2);%画决策面
plot(r1,r2,'*',r3,r4,'o',x,y);
axis([0 12 0 7]);
```

部分结果



$\rho_k = 0.001$时决策面及分类结果