Pattern Recognition

Homework 1

王淮慕
學號 : 606415050
系所 : 電機系

November 9, 2017

0.1 Consider the minimax criterion for a two-category classification problem.

1. Assume $p(x|\omega_1) =\sim N(0,1)$ and $p(x|\omega_2) =\sim N(1/2, 1/4)$ under a zero-one loss. Find the decision threshold x^* and the resulting overall risk .

1-1 sol :

$$\int_{R_2} p(x|\omega_1)dx = \int_{R_1} p(x|\omega_2)dx$$
$$\Rightarrow \int_{\theta}^{\infty} \frac{1}{\sqrt{2\pi}} \mathsf{exp}[\frac{-1}{2}x^2]dx = \int_{-\infty}^{\theta} \frac{1}{\sqrt{2\pi}\times\frac{1}{2}} \mathsf{exp}[\frac{-1}{2\times\frac{1}{2}}(x-\frac{1}{2})^2]dx$$
$$\Rightarrow \Phi(-\theta) = \Phi(\frac{\theta-\frac{1}{2}}{\frac{1}{2}})$$
$$\Rrightarrow \theta = \frac{\theta-\frac{1}{2}}{\frac{1}{2}}$$
$$\Rrightarrow \theta = 1$$

1-2 sol :

$$R = \int_{R_1} p(x|\omega_2)dx$$
$$= \int_{-\infty}^{1} p(x|\omega_2)dx$$
$$= \Phi(\frac{1-\frac{1}{2}}{\frac{1}{2}})$$
$$= \Phi(-1)$$
$$= 0.1587$$

0.2 Consider the Neyman-Person criterion for two univariate normal distributions : $p(x|\omega_1) \sim N(-1,2), p(x|\omega_2) \sim N(2,4)$ $and\ P(\omega_1) = 0.6$.

1. Suppose the maximum acceptable error rate for classifying a pattern that is actually in ω_1 as if it were in ω_2 is 0.05. What is the resulting single-point decision boundary?

$$E_1 = 1 - \int_{-\infty}^{\theta} p(x|\omega_1)dx$$

$$\Rightarrow E_1 = 1 - \Phi(\frac{\theta - \mu_1}{\sigma_1})$$

$$E_1 = 0.05 = 1 - \Phi(\frac{\theta + 1}{\sqrt{2}})$$

$$\theta = \Phi^{-1}(0.95) \times \sqrt{2} - 1 = 1.326(by\ calculas)$$

2. For this boundary, what is the error rate for classifying ω_1 as ω_2

$$E_2 = \int_{\theta}^{-\infty} p(x|\omega_2)dx$$

$$= \Phi(\frac{\theta - \mu_2}{\sigma_2})$$

$$= \Phi(\frac{1.326 - 2}{2})$$

$$= \Phi(-0.337)$$

$$= 0.3681$$

$$error\ rate = E_2 \times P(\omega_2) = 0.3681 \times 0.4 = 0.14724$$

3. What is the overall error rate? Also, compare your result to the Bayes error rate.

$$Overall\ error\ rate\ of\ NP = E_1 \times P(w_1) + E_2 \times P(\omega_2) = 0.17724$$

$$\frac{p(x|\omega_1)P(\omega_1)}{p(x|\omega_2)P(\omega_2)} = \frac{\sigma_2 P(\omega_1)e^{-\frac{1}{2}(\frac{x+1}{\sigma_1})^2}}{\sigma_1 P(\omega_2)e^{-\frac{1}{2}(\frac{x-2}{\sigma_2})^2}}$$

$$\Rightarrow \ln(\frac{\sigma_2 P(\omega_1)}{\sigma_1 P(\omega_2)}) - \frac{1}{2}[(\frac{x+1}{\sigma_1})^2 - (\frac{x-2}{\sigma_2})^2]$$

$$0.752 - \frac{x^2 + 8x - 2}{8} = 0 \qquad then, x = -8.474, 0.474$$

$$Overall\ error\ rate\ of\ Bayes = E_1 \times P(\omega_1) + E_2 \times P(\omega_2)$$

$$\Rightarrow (1 - \int_{-8.474}^{0.474} p(x|\omega_1)dx) \times P(\omega_1) + \int_{-8.474}^{0.474} p(x|\omega_2)dx \times P(\omega_2) = 0.17824$$

0.3 Consider the problem of classifying 10 samples from Table 1. Assume that the underlying distributions are normal.For each category, the mean and the covariance matrix are given by

$$\mu = \frac{1}{10}\sum_{k=1}^{10} x_k$$

$$\sum = \frac{1}{10}\sum_{k=1}^{10}(x_k - \mu)(x_k - \mu)^t$$

where x_k denotes the k-th samples in that category

1. Assume that the prior probabilities for the first two categories are equal $p(\omega_1) = p(\omega_2) = 1/2$ and $p(\omega_3) = 0$. Determine mean vectors and covariance matrices for these two categories using x_1 and x_2 feature values.

3-1 sol :

We substituted $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ to the $\mu = \frac{1}{10}\sum_{k=1}^{10} x_k$, then find $\mu_1,\ \mu_2,\ \mu_3$.

$$\mu_1 = \begin{bmatrix} -0.4400 \\ -1.7490 \end{bmatrix} \quad \mu_2 = \begin{bmatrix} -0.5430 \\ -0.7620 \end{bmatrix} \quad \mu_3 = \begin{bmatrix} 3.8330 \\ 1.3760 \end{bmatrix}$$

According to the formula : $\sum = \frac{1}{10}\sum_{k=1}^{10}(x_k - \mu)(x_k - \mu)^t$
to find $\sum_1,\ \sum_2,\ \sum_3,$

$$\sum_1 = \begin{bmatrix} 12.942 & 6.9258 \\ 6.9258 & 13.161 \end{bmatrix} \quad \sum_2 = \begin{bmatrix} 33.146 & 8.9828 \\ 8.9828 & 11.852 \end{bmatrix} \quad \sum_3 = \begin{bmatrix} 7.4743 & 6.7005 \\ 6.7005 & 7.7044 \end{bmatrix}$$

2. Calculate the percentage of misclassified samples.

3-2 sol : According to the below formula

$$g_i(x) = \mathbf{x^t W_i x} + \mathbf{w_i^t x} + \mathbf{w_{i0}}$$

$$\mathbf{W}_i = \frac{-1}{2}\sum_i^{-1}$$

$$\mathbf{W}_i = \sum_i^{-1} \mu_i$$

$$w_{i0} = \frac{-1}{2}\mu_i^t \sum_i^{-1} \mu_i - \frac{1}{2}\ln|\sum_i| + \ln P(\omega_i)$$

$$error\ rate = 1 - \frac{correct\ sample}{total\ sample}$$

$error\ rate\ \omega_1 = 0.3$
$error\ rate\ \omega_2 = 0.7$

$error\ rate\ \omega_3$ isn't exist ,because $p(\omega_3) = 0$

so the $total\ error\ rate = \frac{0.3+0.7}{3} = 0.5$

3. Repeat all of the above, but now use three feature values ($i.e., x_1, x_2, and\, x_3$)

3-3 sol : We re-calculate $\mu_1, \mu_2, \mu_3, \sum_1, \sum_2, \sum_3$

$$\mu_1 = \begin{bmatrix} -0.4400 \\ -1.7490 \\ -0.7660 \end{bmatrix} \quad \mu_2 = \begin{bmatrix} -0.5430 \\ -0.7620 \\ -0.5420 \end{bmatrix} \quad \mu_3 = \begin{bmatrix} 3.8330 \\ 1.3760 \\ 1.5800 \end{bmatrix}$$

$$\sum_1 = \begin{bmatrix} 12.942 & 6.9258 & 3.7101 \\ 6.9258 & 13.161 & 3.5162 \\ 3.7101 & 3.5162 & 17.7521 \end{bmatrix} \quad \sum_2 = \begin{bmatrix} 33.146 & 8.9828 & -14.7301 \\ 8.9828 & 11.852 & 0.3681 \\ -14.7301 & 0.3681 & 16.5791 \end{bmatrix}$$

$$\sum_3 = \begin{bmatrix} 7.4743 & 6.7005 & 11.8346 \\ 6.7005 & 7.7044 & 10.4478 \\ 11.8346 & 10.4478 & 42.5586 \end{bmatrix}$$

$error\ rate\ \omega_1 = 0.4$
$error\ rate\ \omega_2 = 0.6$

so the $total\ error\ rate = \frac{0.4+0.6}{3} = 0.5$

Table 1: Computer exercise 2 relies on the following data.

	ω_1			ω_2			ω_3		
sample	x_1	x_2	x_3	x_1	x_2	x_3	x_1	x_2	x_3
1	-5.01	-8.12	-3.68	-0.91	-0.18	-0.05	$+5.35$	$+2.26$	$+8.13$
2	-5.43	-3.48	-3.54	$+1.30$	-2.06	-3.53	$+5.12$	$+3.22$	-2.66
3	$+1.08$	-5.52	$+1.66$	-7.75	-4.54	-0.95	-1.34	-5.31	-9.87
4	$+0.86$	-3.78	-4.11	-5.47	$+0.50$	$+3.92$	$+4.48$	$+3.42$	$+5.19$
5	-2.67	$+0.63$	$+7.39$	$+6.14$	$+5.72$	-4.85	$+7.11$	$+2.39$	$+9.21$
6	$+4.94$	$+3.29$	$+2.08$	$+3.60$	$+1.26$	$+4.36$	$+7.17$	$+4.33$	-0.98
7	-2.51	$+2.09$	-2.59	$+5.37$	-4.63	-3.65	$+5.75$	$+3.97$	$+6.65$
8	-2.25	-2.13	-6.94	$+7.18$	$+1.46$	-6.66	$+0.77$	$+0.27$	$+2.41$
9	$+5.56$	$+2.86$	-2.26	-7.39	$+1.17$	$+6.30$	$+0.90$	-0.43	-8.71
10	$+1.03$	-3.33	$+4.33$	-7.50	-6.32	-0.31	$+3.52$	-0.36	$+6.43$

0.4 Consider the data points generated from two categories as shown in Table 2.

1. Write your own code to plot the receiver operating characteristic (ROC) curve.
2. Write your own code to determine the area under the ROC curve (often referred to as AUC).

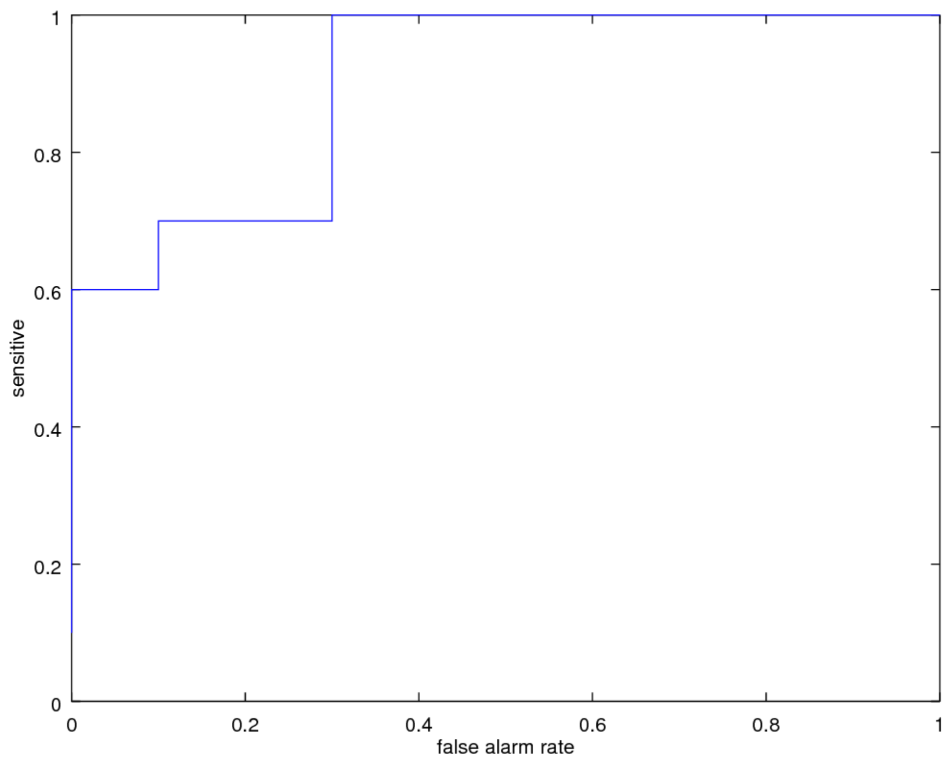
Table 2: Data points generated from two categories.

sample	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
ω_1	-8.1	-3.48	-5.52	-3.78	$+0.63$	$+3.29$	$+2.09$	-2.13	$+2.86$	-3.33
ω_2	$+5.35$	$+5.12$	-1.34	$+4.48$	$+7.11$	$+7.17$	$+5.75$	$+0.77$	$+0.90$	$+3.52$

Mathlab code :

```
clear all;
x = -8:0.1:8;
len = length(x);
omega 1=[-8.1,-3.48,-5.52,-3.78,0.63,3.29,2.09,-2.13,2.86,-3.33];
omega 2=[5.35,5.12,-1.34,4.48,7.11,7.17,5.75,0.77,0.90,3.52];
x axis = zeros(1,len);
y axis = zeros(1,len);
for i=1:len
x start=double([(x(1,i)*ones(1,length(omega 2)))>=omega 2]);
sum x=sum(x start);
x axis(1,i)=sum x/length(omega 2);
y start=double([(x(1,i)*ones(1,length(omega 1)))>=omega 1]);
```

```
sum y=sum(y start);
y axis(1,i)=sum y/length(omega 1);
end
plot(x axis,y axis,'-');
xlabel('false alarm rate');
ylabel('sensitive');
auc area = 0.0;
for i=2:length(y axis)
auc area = ((y axis(1,i)+y axis(1,i-1)) * (x axis(1,i)-x axis(1,i-1))/2)+auc
area;
end
auc area=abs(auc area);
disp('AUC=');
disp(auc area);
```



AUC(area under curve)=0.900