Homework-6 Solutions

Question 1

Consider the following training data:

x_1	x_2	y
1	1	+
2	1	+
1	2	+
0	0	−
1	0	−
2	0	−
3	0	−
0	3	−
3	3	−

−	.	.	−
.	+	.	.
.	+	+	.
−	−	−	−

1. Assume Gaussian distribution where both covariance matrices are a multiple of the identity matrix (Case 1.). What is the discriminat function?

Answer:

$$\mu_1 = \begin{pmatrix} 4/3 \\ 4/3 \end{pmatrix}, \quad \mu_2 = \begin{pmatrix} 1.5 \\ 1 \end{pmatrix},$$
$$w = \mu_1 - \mu_2 = \begin{pmatrix} -1/6 \\ 1/3 \end{pmatrix}, \quad b = 0.$$

The value of b was computed by looking at the sorted products of $w^T x$.

label:	−	−	−	−	+	+	+	−	−
$w^T x_i$	-1/2	-1/3	-1/6	0	0	1/6	1/2	1/2	1

Compute the threshold that gives smallest number of errors. We can't have less than 3 errors, for example with $t = 0$. With these values the discriminant function is:

$$d(x) = -x_1/6 + x_2/3, \quad \text{or} \quad d(x) = 2x_2 - x_1$$

2. Assume equal priors and Gaussian distribution where the covariance matrix is the same for both classes (Case 2.). What is the discriminat function?

Answer:

$$\mu_1 = \begin{pmatrix} 4/3 \\ 4/3 \end{pmatrix}, \quad \mu_2 = \begin{pmatrix} 1.5 \\ 1 \end{pmatrix}, \quad \mu = \begin{pmatrix} 13/9 \\ 10/9 \end{pmatrix},$$
$$C = \sum_{i=1}^{9} (x_i - \mu)(x_i - \mu)^T = \begin{pmatrix} 1.136 & -0.0494 \\ -0.0494 & 1.432098 \end{pmatrix}$$

$$\text{Solve: } Cw = (\mu_1 - \mu_2) \quad \Rightarrow \quad w = \begin{pmatrix} -0.137 \\ 0.228 \end{pmatrix}$$

Calculate b:

i	1	2	3	4	5	6	7	8	9
label:	+	+	+	−	−	−	−	−	−
$w^T x_i$	0.091	-0.046	0.319	0	-0.137	-0.274	-0.4105	0.684	0.274

Sorted:

i	7	6	5	2	4	1	9	3	8
label:	−	−	−	+	−	+	−	+	−
$w^T x_i$	-0.4105	-0.274	-0.137	-0.046	0	0.091	0.274	0.319	0.684

Select a threshold that gives smallest number of errors. The smallest numbe of errors is 3, for example with $t = -0.0915$. The corresponding b is 0.0915. This gives:

$$d(x) = -0.137x_1 + 0.228x_2 + 0.0915$$

Case 3:

$$C_1 = \frac{1}{9}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad C_1^{-1} = 3\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 19/12 & 0 \\ 0 & 2 \end{pmatrix} \quad C_2^{-1} = \begin{pmatrix} 12/19 & 0 \\ 0 & 1/2 \end{pmatrix}$$

$$\begin{aligned} q_1 &= (x-\mu_1)^T C_1^{-1}(x-\mu_1) = 6(x_1^2 + x_2^2 + x_1x_2 - 4x_1 - 4x_2) + 32 \\ q_2 &= (x-\mu_2)^T C_2^{-1}(x-\mu_2) = \frac{3}{19}(2x_1-3)^2 + \frac{1}{2}(x_2-1)^2 \\ q_2 - q_1 &= -\frac{102}{19}x_1^2 - \frac{11}{2}x_2^2 - 6x_1x_2 + \frac{420}{19}x_1 + 23x_2 - 30.07.. = Q(x) - 30.07.. \end{aligned}$$

Therefore, the discriminant function is

$$= Q(x) + b$$

Calculate b:

i	1	2	3	4	5	6	7	8	9
label:	+	+	+	−	−	−	−	−	−
$Q(x)$	28.24	28.24	28.74	0	16.74	22.74	18	19.5	-16.5

Sorted:

i	9	4	5	7	8	6	1	2	3
label:	−	−	−	−	−	−	+	+	+
$Q(x)$	-16.5	0	16.75	18	19.5	22.74	28.24	28.24	28.74

$$t = (28.24 - 22.74)/2 = 3, \quad b = -3$$

Therefore the discriminant function is:

$$-\frac{102}{19}x_1^2 - \frac{11}{2}x_2^2 - 6x_1x_2 + \frac{420}{19}x_1 + 23x_2 - 3$$