Pattern Recognition Homework 2

Ali Mehmandoost (963624020)

November 10, 2017

Contents

1 Cost of substitution

Risk of choosing class ω_{i} is cost of choosing the wrong class (Substitution Error).

$$R(\omega_i|\mathbf{x}) = \lambda_s(1 - P(\omega_i|\mathbf{x}))$$

Risk of associating x with the class ω_{i} should be less than cost of rejection

$$\lambda_s(1 - P(\omega_i|\mathbf{x})) \leq \lambda_r \implies P(\omega_i|\mathbf{x}) \geq 1 - \frac{\lambda_r}{\lambda_s}$$

2 Gaussian assumption

2.1 A

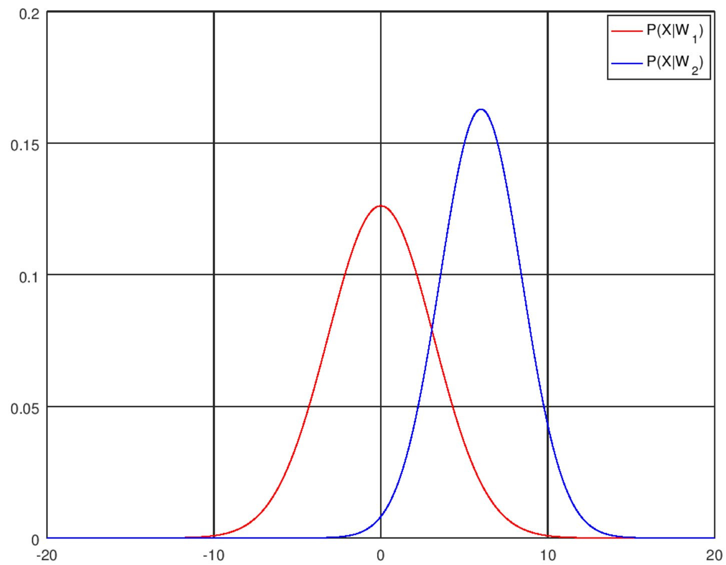
I assume both P(x|ω_1) and P(x|ω_2) are in the form of a Guassian distribution.

$$P(x|\omega_1) = k_1 e^{-\frac{x^2}{20}} = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \implies \mu = 0,\ \sigma^2 = 10,\ k_1 = \frac{1}{\sqrt{20\pi}}$$

$$P(x|\omega_2) = k_2 e^{-\frac{(x-6)^2}{12}} = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \implies \mu = 6,\ \sigma^2 = 6,\ k_2 = \frac{1}{\sqrt{12\pi}}$$

The following octave code will provide a graph of these two densities.

```
X = [-20:0.1:20];
plot (X, normpdf(X,0,sqrt(10)), "r;P(X|W_{1});");
hold on;
plot (X, normpdf(X,6,sqrt(6)), "b;P(X|W_{2});");
grid on;
```



2.2 B

I assume that cost of choosing the right class is zero therefore λ_{11}=λ_{22}=0.

$$R(\omega_1|\mathbf{x}) = \lambda_{11}P(\omega_1|\mathbf{x})+\lambda_{12}P(\omega_2|\mathbf{x}) = 0\times P(\omega_1|\mathbf{x})+\sqrt{3}P(\omega_2|\mathbf{x}) = \sqrt{3}P(\omega_2|\mathbf{x})$$

$$R(\omega_2|\mathbf{x}) = \lambda_{22}P(\omega_2|\mathbf{x})+\lambda_{21}P(\omega_1|\mathbf{x}) = 0\times P(\omega_2|\mathbf{x})+\sqrt{5}P(\omega_1|\mathbf{x}) = \sqrt{5}P(\omega_1|\mathbf{x})$$

2.3 C

$$\sqrt{3}P(\omega_2|\mathbf{x}) = \sqrt{5}P(\omega_1|\mathbf{x}) \implies \frac{\sqrt{3}P(\mathbf{x}|\omega_2)P(\omega_2)}{P(\mathbf{x})} = \frac{\sqrt{5}P(\mathbf{x}|\omega_1)P(\omega_1)}{P(\mathbf{x})}$$

note: P(ω_1) = P(ω_1)

$$\sqrt{3}P(\mathbf{x}|\omega_2) = \sqrt{5}P(\mathbf{x}|\omega_1)$$

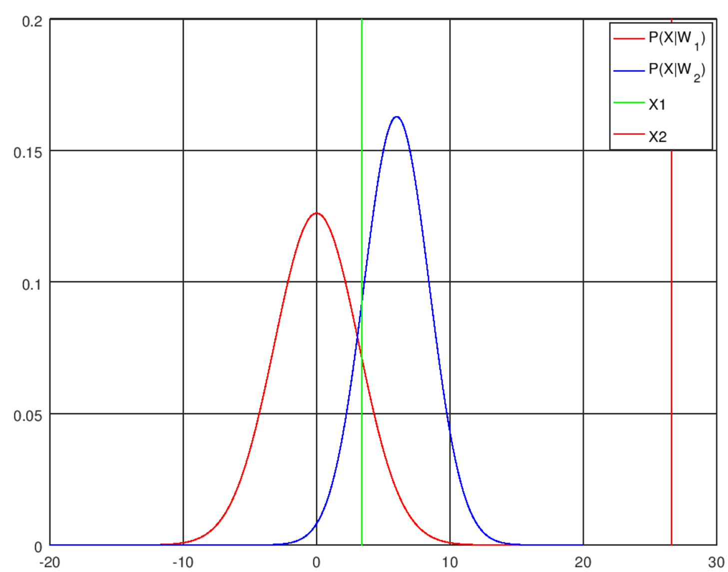
now lets replace likelihoods with their expressions :

$$\sqrt{3}\frac{1}{\sqrt{12\pi}}e^{-\frac{(x-6)^2}{12}} = \sqrt{5}\frac{1}{\sqrt{20\pi}}e^{-\frac{x^2}{20}} \implies e^{-\frac{(x-6)^2}{12}} = e^{-\frac{x^2}{20}} \implies$$

$$8x^2 - 240x + 720 = 0 \implies x = [3.38104, 26.61895]$$

The following octave code will add these two decision boundaries to the previous graph.

```
X = [-20:0.1:20];
plot (X, normpdf(X,0,sqrt(10)), "r;P(X|W_{1});");
hold on;
plot (X, normpdf(X,6,sqrt(6)), "b;P(X|W_{2});");
x=3.38104
plot([x,x],[0,0.2], "g;X1;")
x=26.61895
plot([x,x],[0,0.2], "r;X2;")
grid on;
```



2.4 D

Let aria_1 be the aria before first decision boundary x_1=3.38104, aria_2 aria in between of two decision boundaries, and aria_3 aria after the second decision boundary x_2=26.61895.
in aria_1 decision rule will choose $\omega_1 \implies R_1$.
in aria_2 decision rule will choose $\omega_2 \implies R_2$.
in aria_3 decision rule will choose $\omega_1 \implies R_1$.

$$\int_{-\infty}^{x_1} 0P(\omega_1|x)+\lambda_{12}P(\omega_2|x)+\int_{x_1}^{x_2} \lambda_{21}P(\omega_1|x)+0P(\omega_2|x)+\int_{x_2}^{\infty} 0P(\omega_1|x)+\lambda_{12}P(\omega_2|x) =$$

$$\int_{-\infty}^{x_1} \lambda_{12}P(x|\omega_2)P(\omega_2) + \int_{x_1}^{x_2} \lambda_{21}P(x|\omega_1)P(\omega_1) + \int_{x_2}^{\infty} \lambda_{12}P(x|\omega_2)P(\omega_2) =$$

$$\int_{-\infty}^{x_1} \sqrt{3}N(6,6)\frac{1}{2} + \int_{x_1}^{x_2} \sqrt{5}N(0,10)\frac{1}{2} + \int_{x_2}^{\infty} \sqrt{3}N(6,6)\frac{1}{2} \approx 0.28$$

3 Communication noise

3.1 A

We decide m is one when

$$P(r|m=1)P(m=1) > P(r|m=0)P(m=0) \implies P(r|m=1)\frac{1}{4} > P(r|m=0)\frac{3}{4}$$

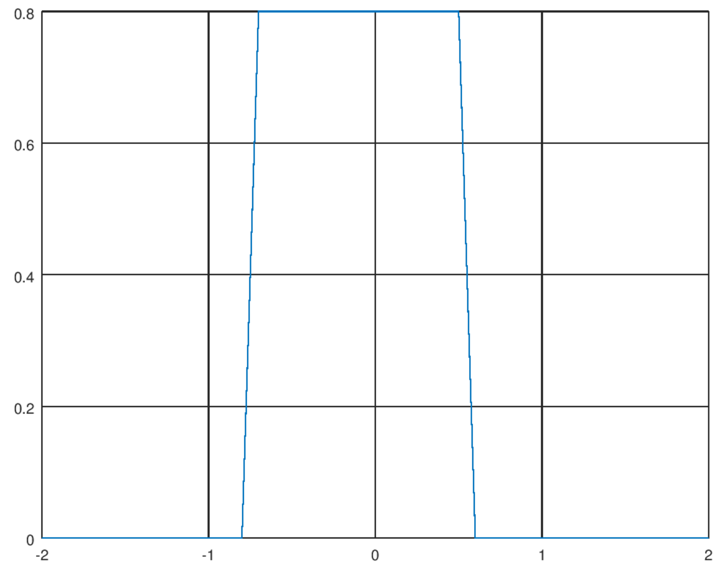
$$\implies P(r|m=1) > 3P(r|m=0)$$

otherwise m is zero.

3.2 B

The uniform distribution can be produce by the following octave code:

```
X = [-2:0.1:2];
plot (X, unifpdf(X,(-3/4),(2/4)));
grid on;
```



Our decision rule in last part still stands.

We will define the new likelihoods for the uniform pdf.

$$P(r|m=1)\begin{cases}\frac{4}{5}, & \text{if } \frac{1}{4}<r<\frac{6}{4}\\ 0, & \text{otherwise}\end{cases}$$

$$P(r|m=0)\begin{cases}\frac{4}{5}, & \text{if } \frac{-3}{4}<r<\frac{2}{4}\\ 0, & \text{otherwise}\end{cases}$$

from our rule:

$$m=\begin{cases}0, & \frac{-3}{4}<r<\frac{2}{4}\\ 1, & \frac{2}{4}<r<\frac{6}{4}\end{cases}$$

And for the Probability of the error:

$$P(Error)=P(m=1|m=0).P(m=0)+P(m=0|m=1)P(m=1)=$$
$$p(\frac{1}{4}<r<\frac{2}{4}|m=1).P(m=1)=(\frac{2}{4}-\frac{1}{4})\times\frac{4}{5}\times\frac{1}{4}=\frac{1}{20}$$

4 Linear Transformation and Mahalanobis distance

```
S = [1,0,0;0,5,4;0,4,5;];
[Phi, Lambda] = eig(S);
```

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.77011 & 0.77011 \\ 0 & 0.77011 & 0.77011 \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

4.1 A

```
S = [1,0,0;0,5,4;0,4,5;];
mu = [3;1;2];
X0 = [5;6;3;];

(1/( (2*pi)^(3/2) * det(S)^0.5  )) * exp( ( -(X0-mu)' * inv(S) * (X0-mu) )/2 )
```

ans = 1.9300e-05

4.2 B

I Wasn't sure if I should suggest an orthonormal transformation or use the Φ calculated in the first part.

$$\Phi = \begin{bmatrix} \frac{2}{7} & \frac{6}{7} & \frac{3}{7} \\ \frac{3}{7} & \frac{2}{7} & \frac{-6}{7} \\ \frac{6}{7} & \frac{-3}{7} & \frac{2}{7} \end{bmatrix}$$

anyhow I suggest the previous matrix and continue with the Φ in the first part:

$$y = \Phi^T x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.77011 & 0.77011 \\ 0 & 0.77011 & 0.77011 \end{bmatrix}^T x$$

$$||y||^2 = y^T y = (\Phi^T x)^T \Phi^T x = x^T \Phi \Phi^T x = x^T x = ||x||^2$$

4.3 C

$$\Lambda^{-1/2}\Phi^T x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 9 \end{bmatrix}^{-1/2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.77011 & 0.77011 \\ 0 & 0.77011 & 0.77011 \end{bmatrix}^T x$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.33333 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.77011 & 0.77011 \\ 0 & 0.77011 & 0.77011 \end{bmatrix}^T x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -0.77011 & 0.77011 \\ 0 & 0.23570 & 0.23570 \end{bmatrix} x$$

Because ($\Lambda^{-1/2}$ Φ^T x) is a linear transformation of x and p(x|ω)~N(μ, Σ) we can conclude

$$P((\Lambda^{-1/2}\Phi^T x)|\omega) \sim N((\Lambda^{-1/2}\Phi^T \mu), (\Lambda^{-1/2}\Phi^T x)\Sigma(\Lambda^{-1/2}\Phi^T x)^T)$$

This transformation will cause Σ to became I and to move the distribution to the center we need to move the points according to ($\Lambda^{-1/2}$ Φ^T)(x-μ).

4.4 D

```
S = [1,0,0;0,5,4;0,4,5;];
mu = [3;1;2];
X0 = [5;6;3;];
[Phi, Lambda] = eig(S);

A = Phi*(Lambda^(-0.5));

X_w = A'*(X0-mu)
```

$$X(w) = (\Lambda^{-1/2}\Phi^T)(x - \mu) = \begin{bmatrix} 2 \\ -2.8284 \\ 1.4142 \end{bmatrix} \tag{19}$$

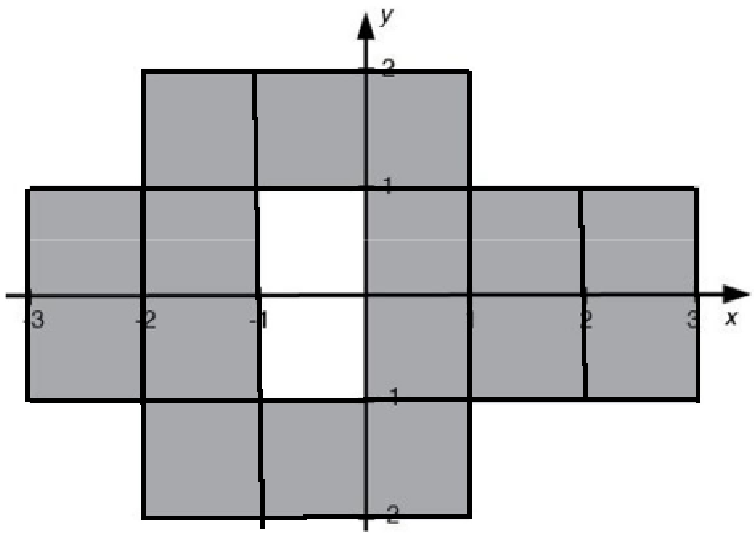
4.5 E

```
S = [1,0,0;0,5,4;0,4,5;];
mu = [3;1;2];
X0 = [5;6;3;];
[Phi, Lambda] = eig(S);

A = Phi*(Lambda^(-0.5));

X_w = A'*(X0-mu);

function m = mahalanobis_dist(X, Sigma, Mu)
m = (X-Mu)' * inv(Sigma) * (X-Mu);
endfunction

mahalanobis_dist(X0, S, mu)
mahalanobis_dist(X_w, eye(3), 0)
```

Ans = 14
Ans = 14
Yes, Linear Transformation doesn't change Mahalanobis distance.

5 Gray Squares

5.1 A



they are equal gray aria on both side of x axis so P(ω_1) = P(ω_2) = 0.5.

$$P(1 < Y < 2|\omega_1) = \frac{P(Y|\omega_1)}{P(\omega_1)} = \frac{\frac{2}{16}}{\frac{1}{2}} = \frac{1}{4}$$

$$P(0 < Y < 1|\omega_1) = \frac{P(Y|\omega_1)}{P(\omega_1)} = \frac{\frac{2}{16}}{\frac{1}{2}} = \frac{1}{4}$$

$$P(-1 < Y < 0|\omega_1) = \frac{P(Y|\omega_1)}{P(\omega_1)} = \frac{\frac{2}{16}}{\frac{1}{2}} = \frac{1}{4}$$

$$P(-2 < Y < 1|\omega_1) = \frac{P(Y|\omega_1)}{P(\omega_1)} = \frac{\frac{2}{16}}{\frac{1}{2}} = \frac{1}{4}$$

$$P(1 < Y < 2|\omega_2) = \frac{P(Y|\omega_2)}{P(\omega_2)} = \frac{\frac{1}{16}}{\frac{1}{2}} = \frac{1}{8}$$

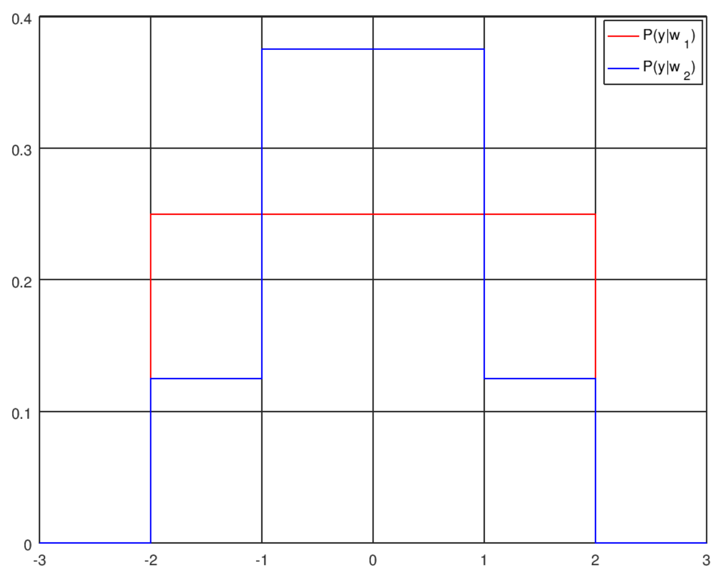
$$P(0 < Y < 1|\omega_2) = \frac{P(Y|\omega_2)}{P(\omega_2)} = \frac{\frac{3}{16}}{\frac{1}{2}} = \frac{3}{8}$$

$$P(-1 < Y < 0|\omega_2) = \frac{P(Y|\omega_2)}{P(\omega_2)} = \frac{\frac{3}{16}}{\frac{1}{2}} = \frac{3}{8}$$

$$P(-2 < Y < 1|\omega_2) = \frac{P(Y|\omega_2)}{P(\omega_2)} = \frac{\frac{1}{16}}{\frac{1}{2}} = \frac{1}{8}$$

5.2 B

```
Y = [-3, -2, -2+0.00001, -1, 0, 1, 2-0.00001, 2, 3];
P_Y_W1 = [0, 0, 1/4,1/4, 1/4, 1/4, 1/4, 0, 0];
plot(Y, P_Y_W1, "r;P(y|w_1);")
hold on
Y = [-3, -2, -2+0.00001, -1,-1+0.00001,1, 1+0.00001,2,2+0.00001,3]
P_Y_W2 = [0, 0, 1/8, 1/8, 3/8,3/8,1/8,1/8,0,0]
plot(Y, P_Y_W2, "b;P(y|w_2);")
```



5.3 C

We classify a point as ω_1 when

$$P(y|\omega_1)P(\omega_1) > P(y|\omega_1)P(\omega_1)$$

$$note: \; P(\omega_1) = P(\omega_2)$$

$$P(y|\omega_1) > P(y|\omega_2)$$

and We classify a point as ω_2 when

$$P(y|\omega_2) > P(y|\omega_1)$$

therefore

$$c = \begin{cases} \omega_1, & -2 < y < 1 or 1 < y < 2 \\ \omega_2, & -1 < y < 1 \end{cases} \tag{24}$$

and for the probability of error:

$$P(error) = P(c = \omega 1|\omega_2)P(\omega_1) + P(c = \omega 2|\omega_1)P(\omega_2) =$$
$$P(\omega_1)P(-1 < y < 1|\omega_2) + P(\omega_2)P(-2 < y < 1 or 1 < y < 2|\omega_1) =$$
$$\frac{1}{2} \times \frac{4}{8} + \frac{1}{2} \times \frac{2}{8} = \frac{3}{8}$$