Pattern Recognition Homework 1

İpek Erdoğan

November 29, 2020

Question 1

Part 1

Let α_i be the action of deciding W_i and α_j be the action of deciding W_j. And $\lambda_{i_j} = \lambda(\alpha_i|W_j)$ is the loss for taking α_i as the action and deciding on W_i, when the true choice is W_j.

Then, our conditional risks would be:

$$R(\alpha_i|x) = \lambda_{i_i} P(W_i|x) + \lambda_{i_j} P(W_j|x)$$

$$R(\alpha_j|x) = \lambda_{j_j} P(Wj_i|x) + \lambda_{j_i} P(W_i|x)$$

Let's choose zero-one function as our loss function to simplify the conditional risks. In that case, λ_{i_i} and λ_{j_j} values will be zero and λ_{i_j} and λ_{j_i} values will be one.

Now we can see that:

$$R(\alpha_i|x) = P(W_j|x) = P(W_j)p(x|W_j)$$

and

$$R(\alpha_j|x) = P(W_i|x) = P(W_i)p(x|W_i)$$

Considering our problem, the error would be:

$$P(W_j|x) if x > k \rightarrow \int_{k}^{\infty} P(W_2)p(x|W_2)\,dx$$
$$P(W_j|x) if x \leq k \rightarrow \int_{-\infty}^{k} P(W_1)p(x|W_1)\,dx$$

And our average probability of error would be:

$$\int_{-\infty}^{k} P(W_1)p(x|W_1)\,dx + \int_{k}^{\infty} P(W_2)p(x|W_2)\,dx$$

Summarized explanation: In the area that we needed to choose W_1 while x>k (let's call it R1), our error probability would be $P(W_2|x)$ and in the area that we needed to choose W_2 while x≤k (let's call it R2), our error probability would be $P(W_1|x)$.

Part 2

The derivative of the average error probability with respect to k needs to be equal to zero, to find the minimum error.

$$\frac{\mathrm{d}P(error)}{\mathrm{d}k}$$

$$P(error) = P(W_1)\int_{-\infty}^{k} p(x|W_1)\,dx + P(W_2)\int_{k}^{\infty} p(x|W_2)\,dx$$

I will use the fundamental method to take the derivative of the second integral above.

$$P'(error) = P(W_1)p(k|W_1) - P(W_2)p(k|W_2)$$

The equation above needs to be equal to zero, to make P(error) minimum.

$$P(W_1)p(k|W_1) = P(W_2)p(k|W_2)$$

$$\frac{P(W_1)}{p(k|W_1)} = \frac{P(W_2)}{p(k|W_2)}$$

Question 2

For the critical region

$$C_\alpha = \left[x | \frac{p(x|W_1)}{p(x|W_0)} \geq k_\alpha \right]$$

Since our distributions are univariate Gaussian, we can take the ln of both sides to get rid of the exponentials and simplify our solution.

$$C_\alpha = \left[x | \ln(\frac{p(x|W_1)}{p(x|W_0)}) \geq \ln(k_\alpha) \right]$$

$$p(x|W_0) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu_0)^2 \Big/ 2\sigma^2}$$

$$p(x|W_1) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu_1)^2 \Big/ 2\sigma^2}$$

$$\frac{p(x|W_1)}{p(x|W_0)} = \frac{e^{-(x-\mu_1)^2 \Big/ 2\sigma^2}}{e^{-(x-\mu_0)^2 \Big/ 2\sigma^2}}$$

$$\ln(\frac{e^{-(x-\mu_1)^2 \Big/ 2\sigma^2}}{e^{-(x-\mu_0)^2 \Big/ 2\sigma^2}})$$

$$ln(e^{-(x-\mu_1)^2 \Big/ 2\sigma^2}) - ln(e^{-(x-\mu_0)^2 \Big/ 2\sigma^2})$$

$$= -(x-\mu_1)^2 \Big/ 2\sigma^2 + (x-\mu_0)^2 \Big/ 2\sigma^2$$

$$C_\alpha = \left[\frac{(x-\mu_0)^2 - (x-\mu_1)^2}{2\sigma^2} \geq \ln(k_\alpha) \right]$$

$$2(\mu_1 - \mu_0)x \geq 2\sigma^2 \ln(k_\alpha) + (\mu_1)^2 - (\mu_0)^2$$

$$x \geq \frac{2\sigma^2 \ln(k_\alpha) + (\mu_1)^2 - (\mu_0)^2}{2(\mu_1 - \mu_0)}$$

$if \mu_1 > \mu_0$ we reject H_0 if

$$x \geq \frac{2\sigma^2 \ln(k_\alpha) + (\mu_1)^2 - (\mu_0)^2}{2(\mu_1 - \mu_0)}$$

$if \mu_1 < \mu_0$ we reject H_0 if

$$x \leq \frac{2\sigma^2 \ln(k_\alpha) + (\mu_1)^2 - (\mu_0)^2}{2(\mu_1 - \mu_0)}$$

And for the power of test:

$$1 - \beta = \int_{X_0}^{\infty} p(x|W_1)\, dx$$
$$= \frac{1}{\sigma\sqrt{2\pi}} e^{-(X_0-\mu_1)^2 \big/ 2\sigma^2}$$

Question 3

Part 1

We need to derive the minimum risk decision rule.

$$R(\alpha_i|x) = \sum_{j=1}^{c} \lambda(\alpha_i|W_j)P(W_j|x)$$

For α_r:

$$R(\alpha_r|x) = \sum_{j=1}^{c} \lambda_r P(W_j|x)$$

For α_i:

$$R(\alpha_i|x) = \sum_{j=1, i\neq j}^{c} \lambda_s P(W_j|x)$$

Normally we use posterior probabilities to decide between the categories. So it's normal to use $p(x|W_i)p(W_i)$ as the discriminator function between the normal categories. But here, we have also a rejection option.

So we can basically say: Decide α_r if

$$R(\alpha_r|x) < R(\alpha_i|x)$$

$$\sum_{j=1}^{c} \lambda_r P(W_j|x) < \sum_{j=1, i\neq j}^{c} \lambda_s P(W_j|x)$$

Let's add and substract $\lambda_s P(W_i|x)$ to the right side.

$$\sum_{j=1}^{c} \lambda_r P(W_j|x) < \sum_{j=1, i\neq j}^{c} \lambda_s P(W_j|x) + \lambda_s P(W_i|x) - \lambda_s P(W_i|x)$$

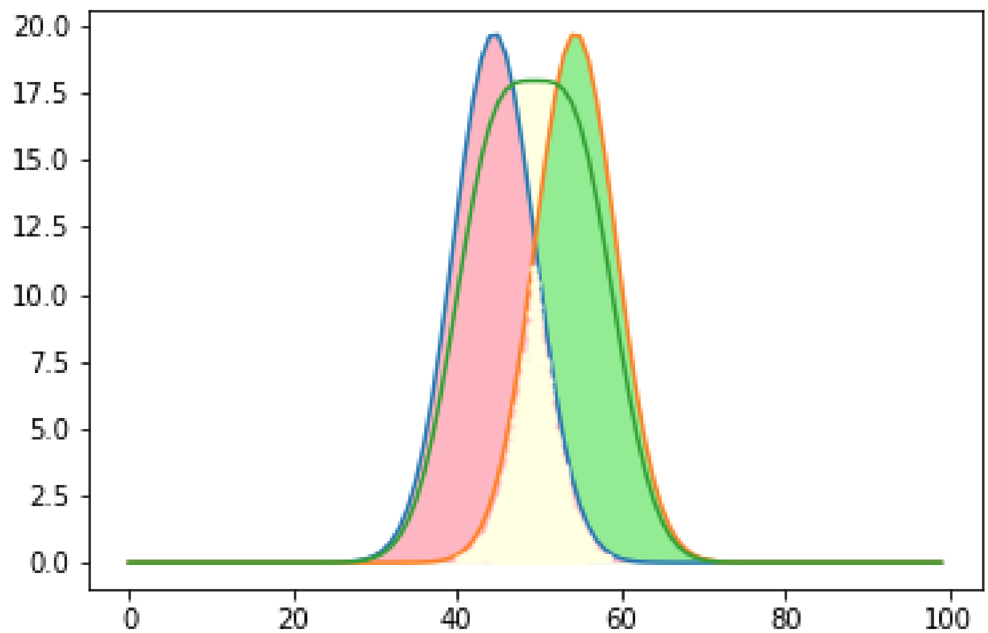
$$\sum_{j=1}^{c} \lambda_r P(W_j|x) - \sum_{j=1}^{c} \lambda_s P(W_j|x) < -\lambda_s P(W_i|x)$$

And devide both sides with $-\lambda_s$

$$\frac{\lambda_s - \lambda_r}{\lambda_s} \sum_{j=1}^{c} P(W_j|x) > P(W_i|x)$$

Since my decision rule includes the same discriminant functions, I can say that tey are optimal.

Part 2



There is a graph of the discriminant functions and decision regions above. In this graph, blue line is the discriminant function for W_2, blue line is the discriminant function for W_1 and the green line is the discriminant function for rejection option.

Red area refers to the decision region of W_2, green area refers to the decision region of W_1 and yellow area refers to the decision region for rejection option.

My code is down below:

```
x_values = np.linspace(-10, 10, 100)
def class_density(x,mu,sig):
    return np.exp(-np.power(x - mu, 2.) / (2 * np.power(sig, 2.))) / np.power((2

mu1=1
sig1=1
mu2=-1
sig2=1

mus = [mu1,mu2]
sigs= [sig1,sig2]
alpha_r=1
alpha_s=4
c = 2 #W1 and W2

def g1x(x,i):
    return (1/2)*class_density(x,mus[i],sigs[i]) #1/2 = P(W1) = P(W2)
def g2x(x):
    return ((alpha_s-alpha_r)/alpha_s)*sum([g1x(x,i) for i in range(0,c)])

def discriminant_func(x,i):
    if (1<=i<=c):
        return g1x(x,i-1)
    elif(i==c+1):
        return g2x(x)

#1,2 and 3 down below are the class numbers.
g_w1=discriminant_func(x_values,1)
g_w2=discriminant_func(x_values,2)
g_reject=discriminant_func(x_values,3)

mp.plot(g_w2)
mp.plot(g_w1)
mp.plot(g_reject)
mp.fill(g_reject,'lightyellow')
mp.fill(g_w1,'lightgreen')
mp.fill(g_w2,'lightpink')

mp.savefig("decision.png")
mp.show()
```