Statistical ML
Assignment - I

VENKTESH V
Roll No : PhD 19016
PHD, CSE

Question 1 Consider following decision rule for 2 category 1D problem :-

Decide ω_1 if $x > \theta$; or decide ω_2

a) S.T. probability of error for this rule is

$$P(\text{error}) = P(\omega_1)\int_{-\infty}^{\theta} P(x|\omega_1)\,dx + P(\omega_2)\int_{\theta}^{\infty} P(x|\omega_2)\,dx$$

b) By differentiating show a necessary condition to minimize P(error) is that θ satisfy $P(\theta|\omega_1)P(\omega_1) = P(\theta|\omega_2)P(\omega_2)$

Answer : a) $$P(\omega_i|x) = \frac{p(x|\omega_i)\,P(\omega_i)}{p(x)}$$

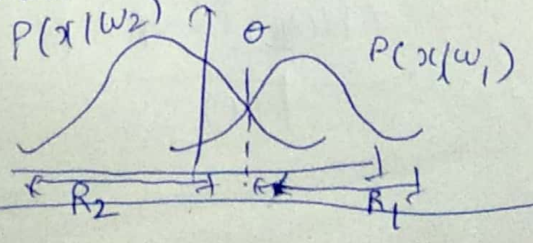
ω_1 if $x > \theta$ or ω_2

$$P(\text{error}) = \int_{-\infty}^{\infty} P(\text{error}, x)\,dx$$

$$= \int_{-\infty}^{\infty} P(\text{error}|x)\,p(x)\,dx$$

$$P(\text{error}|x) = \begin{cases} P(\omega_1|x) & \text{if } \omega_2 \text{ is decision} \\ P(\omega_2|x) & \text{if } \omega_1 \text{ is decided} \end{cases}$$

θ is the threshold based. Assuming likelihood to be normal dist.



Case I: In region R_1, though some points belongs to category ω_2 it might get misclassified as ω_1 as $x > \theta$ is the decision rule. So here error is $\int_{\theta}^{\infty} P(\omega_2|x)\, p(x)\, dx$

Case II: In region R_2, though some points belongs to category ω_1 in some cases, it will be classified as ω_2 in these cases as $x < \theta$ classifies sample to category ω_2.

$\therefore$ error here is $\int_{-\infty}^{\theta} P(\omega_1|x)\, p(x)\, dx$

Adding the above 2

$$P(\text{error}) = \int_{-\infty}^{\theta} P(\omega_1|x)\, p(x)\, dx + \int_{\theta}^{\infty} P(\omega_2|x)\, p(x)\, dx$$

$P(\omega_i|x)$ can be written as $\dfrac{p(x|\omega_i)\, P(\omega_i)}{p(x)}$

$$\Rightarrow P(\text{error}) = \int_{-\infty}^{\theta} \frac{P(x|\omega_1)\, P(\omega_1)}{p(x)}\, p(x)\, dx + \int_{\theta}^{\infty} \frac{P(x|\omega_2)\, P(\omega_2)}{p(x)}\, p(x)\, dx$$

Scanned by CamScanner

$$\Rightarrow P(\text{error}) = \int_{-\infty}^{\theta} p(x|w_1)\, P(w_1)\, dx + \int_{\theta}^{\infty} p(x|w_2)\, P(w_2)\, dx$$

$$P(\text{error}) = P(w_1) \int_{-\infty}^{\theta} p(x|w_1)\, dx + P(w_2) \int_{\theta}^{\infty} p(x|w_2)\, dx$$

b) Differentiating the above equation w.r.t θ to minimize error such that $P(\text{error}) = 0$

$$\Rightarrow \frac{d}{d\theta}\left(P(w_1) \int_{-\infty}^{\theta} p(x|w_1)\, dx + P(w_2) \int_{\theta}^{\infty} p(x|w_2)\, dx \right) = 0 \quad \text{—(1)}$$

By rule of differentiation under integral sign

$$\frac{d}{dx} \int_{0}^{x} f(y)\, dy = f(x) \quad \Rightarrow \frac{d}{dx} \int_{x}^{0} f(y)\, dy = -f(x)$$

Applying the above to equation (1) we get

$$= P(w_1)\,(p(\theta|w_1)) + P(w_2)(-p(\theta|w_2)) = 0$$

$$\boxed{P(w_1)\, p(\theta|w_1) = P(w_2)\, p(\theta|w_2)}$$

Question : 2 Let the conditional densities for a 2 category 1 dim problem be given by Cauchy dist.

$$p(x|w_i) = \frac{1}{\pi b} \frac{1}{1+\left(\frac{x-a_i}{b}\right)^2}, i=1,2.$$

Assuming $P(w_1)=P(w_2)$, S.T. $P(w_1|x) = P(w_2|x)$ if $x=(a_1+a_2)/2$ i.e. minimum decision boundary is a point midway between peaks of 2 dist. regardless of b.

Answer : Given & assumptions : 2 category so 2 classes w_1 & w_2 priors are equal (so both are equally likely) i.e. $P(w_1) = P(w_2)$.

given $x = \frac{a_1+a_2}{2}$, substituting it in $p(x|w_1)$

$$p(x|w_1) = \frac{1}{\pi b} \frac{1}{1+\left(\frac{\frac{a_1+a_2}{2}-a_1}{b}\right)^2} \quad \boxed{\text{i.e.}(i=1)}$$

$$\Rightarrow p(x|w_1) = \frac{1}{\pi b} \frac{1}{1+\left(\frac{a_1+a_2-2a_1}{b}\right)^2}$$

$$= \frac{1}{\pi b} \frac{1}{1+\left(\frac{a_2-a_1}{b}\right)^2} = \frac{1}{\pi b} \frac{1}{1+(-1)^2\left(\frac{a_1-a_2}{b}\right)^2}$$

P.T.O

$$\Rightarrow p(x|w_1) = \frac{1}{\pi b} \frac{1}{1+\left(\frac{a_1-a_2}{2b}\right)^2}$$

Now again substituting $x = \frac{a_1+a_2}{2}$ in $p(x|w_2)$

$$p(x|w_2) = \frac{1}{\pi b} \frac{1}{1+\left(\frac{\frac{a_1+a_2}{2}-a_2}{b}\right)^2}$$ Here $i = 2$

$$= \frac{1}{\pi b} \frac{1}{1+\left(\frac{a_1-a_2}{2b}\right)^2}$$

Now posterior = $\frac{\text{likelihood} * \text{prior}}{\text{evidence}}$ $\Rightarrow P(w_i/x) = \frac{p(x|w_i)P(w_i)}{p(x)}$

but $p(x)$ is just present so left side adds to 1 ∴

$$P(w_i|x) \propto p(x|w_i)\, P(w_i)$$

$$P(w_1|x) \propto p(x|w_1)\, P(w_1)$$

$$\Rightarrow P(w_1|x) \propto \frac{1}{\pi b} \frac{1}{1+\left(\frac{a_1-a_2}{2b}\right)^2} P(w_1) \quad - (1)$$

$$P(w_2|x) \propto \frac{1}{\pi b} \frac{1}{1+\left(\frac{a_1-a_2}{2b}\right)^2} P(w_2) \quad - (2)$$

Q2 -- continued

From ① & ② & since $P(\omega_1) = P(\omega_2)$ (equally likely) $= \frac{1}{2}$

~~So we derive P(~~ we get $\boxed{P(\omega_1) P(x|\omega_1) = P(x|\omega_2) P(\omega_2)}$

$\Rightarrow \boxed{P(\omega_1|x) = P(\omega_2|x)}$

Another proof:

Also the above problem can be proved another way.

~~Assu~~ We can prove that for $P(\omega_1|x)$ to be equal to $P(\omega_2|x)$

x should be $\frac{a_1 + a_2}{2}$

So

$$P(\omega_1|x) = P(\omega_2|x)$$

given $P(\omega_i|x) \propto p(x|\omega_i) P(\omega_i)$

$$p(x|\omega_1) P(\omega_1) = p(x|\omega_2) P(\omega_2)$$

$$\frac{1}{\pi b} \frac{1}{1+\left(\frac{x-a_1}{b}\right)^2} \cdot P(\omega_1) = \frac{1}{\pi b} \frac{1}{1+\left(\frac{x-a_2}{b}\right)^2} \cdot P(\omega_2)$$

given $P(\omega_1) = P(\omega_2) = \frac{1}{2}$ (equally likely)

$$\frac{1}{\pi b} \frac{1}{1+\left(\frac{x-a_1}{b}\right)^2} = \frac{1}{\pi b} \frac{1}{1+\left(\frac{x-a_2}{b}\right)^2}$$

we don't consider b so we can cancel it out

$$\frac{1}{1+\left(\frac{x-a_1}{b}\right)^2} = \frac{1}{1+\left(\frac{x-a_2}{b}\right)^2}$$

$$\Rightarrow \left(\frac{x-a_2}{b}\right)^2 = \left(\frac{x-a_1}{b}\right)^2$$

$$\frac{x^2 - 2a_2x + a_2^2}{\cancel{b}} = \frac{x^2 - 2xa_1 + a_1^2}{\cancel{b}}$$

$$\cancel{x} \Rightarrow -2a_2x + a_2^2 = -2xa_1 + a_1^2$$

$$2x(a_1 - a_2) = a_1^2 - a_2^2$$

$$2x = \frac{(a_1 + a_2)(a_1 - a_2)}{(a_1 - a_2)}$$

$$2x = (a_1 + a_2) \Rightarrow x = \frac{a_1 + a_2}{2}$$

Hence the converse that for $P(\omega_1|x) = P(\omega_2|x)$ x must be $\frac{a_1 + a_2}{2}$

Question 3 Suppose we have 3 equi probable categories in 2-D with following underlying distributions:

$$p(x|w_1) \sim N(0, I)$$
$$p(x|w_2) \sim N\left(\begin{bmatrix}1\\1\end{bmatrix}, I\right)$$
$$p(x|w_3) \sim 0.5\, N\left(\begin{bmatrix}0.5\\0.5\end{bmatrix}, I\right) + 0.5\, N\left(\begin{bmatrix}-0.5\\0.5\end{bmatrix}, I\right)$$

By explicit calculation of posterior probabilities, classify point $\begin{bmatrix}0.3\\0.3\end{bmatrix}$ for min prob. of error.

Answer:

$$\text{Posterior} = \frac{\text{likelihood} * \text{prior}}{\text{evidence}}$$

$$P(w_i|x) = \frac{p(x|w_i)\,P(w_i)}{p(x)} \Rightarrow P(w_i|x) \propto p(x|w_i)\,P(w_i)$$

Here given w_1, w_2, w_3 are equally probable the posterior probabilities are directly dependent on likelihoods. So the sample category w_i for which $P(x|w_i)$ is largest will be the most likely true category.

$p(x|w_1) = N(0, I)$. Now the form of multivariate normal density is

$$N(\mu, \Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}}\, e^{\left[-\frac{1}{2}(x-\mu)^t \Sigma^{-1}(x-\mu)\right]}$$

where x - d column vect, μ - d dim mean vector.

Σ - $d \times d$ - covariance matrix.

Scanned by CamScanner

So now solving for $p(x|w_1)$ where $d = 2$ (given)

$$p(x|w_1) = \frac{1}{(2\pi)^{2/2}|\Sigma|^{1/2}} e^{\left\{-\frac{1}{2}\left(\begin{bmatrix}0.3\\0.3\end{bmatrix}^t \Sigma^{-1}\left(\begin{bmatrix}0.3\\0.3\end{bmatrix}\right)\right)\right\}} \quad (1)$$

In above μ is omitted as μ is 0 vector.

Now given $\Sigma = I$ so $\Sigma = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$

$$\Sigma^{-1} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \quad |\Sigma|^{d} = \begin{vmatrix}1 & 0\\0 & 1\end{vmatrix} = 1$$

$|\Sigma|^{1/2} = 1$. Substituting this in (1)

$$p(x|w_1) = \frac{1}{2\pi} e^{\left(-\frac{1}{2}\left(\begin{bmatrix}0.3 & 0.3\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}0.3\\0.3\end{bmatrix}\right)\right)}$$

$$= \frac{1}{2\pi} e^{\left(-\frac{1}{2}\left(\begin{bmatrix}0.3 & 0.3\end{bmatrix}\begin{bmatrix}0.3\\0.3\end{bmatrix}\right)\right)}$$

$$= \frac{1}{2\pi} e^{-\frac{1}{2}\left(\left[0.09+0.09\right]\right)}$$

$$= \frac{1}{2\pi} e^{-\frac{1}{2}\left[0.18\right])}$$

$$= \frac{1}{2\pi} e^{-\frac{1}{2}(0.18)}$$

$$= \frac{1}{2\pi} e^{-(0.09)}$$

(numpy np.exp(-0.09) gives 0.9139)

$$\therefore p(x|w_1) = \frac{0.9139}{2 \times 3.14} = 0.145525$$

Now again $p(x|w_2) = \frac{1}{(2\pi)^{2/2}|\Sigma|^{1/2}} e^{\left(-\frac{1}{2}\left(\begin{bmatrix}0.3-1\\0.3-1\end{bmatrix}^t \Sigma^{-1} \begin{bmatrix}0.3-1\\0.3-1\end{bmatrix}\right)\right)}$

$= \frac{1}{2\pi(1)} e^{\left(-\frac{1}{2}\left(\begin{bmatrix}-0.7\\-0.7\end{bmatrix}^t \Sigma^{-1} \begin{bmatrix}-0.7\\-0.7\end{bmatrix}\right)\right)}$

$= \frac{1}{2\pi} e^{\left(-\frac{1}{2}\left\{\begin{bmatrix}-0.7\\-0.7\end{bmatrix}^t \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \begin{bmatrix}-0.7\\-0.7\end{bmatrix}\right)\right)}$

$= \frac{1}{2\pi} e^{\left(-\frac{1}{2}\left(\begin{bmatrix}-0.7 & -0.7\end{bmatrix} \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \begin{bmatrix}-0.7\\-0.7\end{bmatrix}\right)\right)}$

$= \frac{1}{2\pi} e^{\left(-\frac{1}{2}\left(\begin{bmatrix}-0.7 & -0.7\end{bmatrix} \begin{bmatrix}-0.7\\-0.7\end{bmatrix}\right)\right)}$

$= \frac{1}{2\pi} e^{\left(-\frac{1}{2}\left(\left[0.49 + 0.49\right]\right)\right)}$

$= \frac{1}{2\pi} e^{\left(-\frac{1}{2}\left[0.98\right]\right)}$

$= \frac{1}{2\pi} e^{-(0.49)}$

$= \frac{1}{2\pi} (0.61262)$

$= \frac{0.61262}{6.28} = 0.09755$

$$\boxed{p(x|w_2) = 0.09755}$$

Now calculating $P(x|W_3) = \frac{1}{2\sqrt{\pi}} e^{\left(\frac{-1}{2}\right.}$

$$\approx 0.5\left[\frac{1}{2\pi(1)} e^{\left(-\frac{1}{2}\left(\begin{bmatrix}0.3-0.5\\0.3-0.5\end{bmatrix}^T \Sigma^{-1}\begin{bmatrix}0.3-0.5\\0.3-0.5\end{bmatrix}\right)\right)}\right]$$

$$+ 0.5\left[\frac{1}{2\pi(1)} e^{\left(-\frac{1}{2}\left(\begin{bmatrix}0.3-(-0.5)\\0.3-0.5\end{bmatrix}^T \Sigma^{-1}\begin{bmatrix}0.3-(-0.5)\\0.3-0.5\end{bmatrix}\right)\right)}\right]$$

$$= \frac{1}{4\pi}\left[e^{-\frac{1}{2}\left(\begin{bmatrix}-0.2 & -0.2\end{bmatrix}\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}-0.2\\-0.2\end{bmatrix}\right)}\right] + \frac{1}{4\pi}\left[e^{-\frac{1}{2}\left(\begin{bmatrix}-0.8 & -0.2\end{bmatrix}\begin{bmatrix}-0.8\\-0.2\end{bmatrix}\right)}\right]$$

$$= \frac{1}{4\pi}\left[\left(e^{-\frac{1}{2}\left(\begin{bmatrix}-0.2 & -0.2\end{bmatrix}\begin{bmatrix}-0.2\\-0.2\end{bmatrix}\right)}\right)\right] + \frac{1}{4\pi}\left[e^{-\frac{1}{2}\left(\begin{bmatrix}-0.8 & -0.2\end{bmatrix}\begin{bmatrix}-0.8\\-0.2\end{bmatrix}\right)}\right]$$

$$= \frac{1}{4\pi}\left[e^{-\frac{1}{2}\left(\left[0.04+0.04\right]\right)} + e^{-\frac{1}{2}\left(\left[0.64+0.04\right]\right)}\right]$$

$$= \frac{1}{4\pi}\left[e^{-1/2\left[0.08\right]} + e^{-1/2\left[0.68\right]}\right]$$

$$= \frac{1}{4\pi}\left[e^{-\left[0.04\right]} + e^{-\left[0.34\right]}\right]$$

$$= \frac{1}{4\pi}\left(0.9607 + 0.7117\right)$$

$$= \frac{1.6724}{4 \times 3.14} = 0.13315$$

$$\boxed{P(x|W_3) = 0.13315}$$

Problem 3 continued

$\therefore$ $p(x|w_1) = 0.145525$, $p(x|w_2) = 0.09755$

& $p(x|w_3) = 0.13315$

As $p(x|w_1)$ is highest here $p(x|w_1) > p(x|w_3) > p(x|w_2)$

So the point (x $\begin{bmatrix} 0.3 \\ 0.3 \end{bmatrix}$ ~~so~~ is classified as class

/category 1. (w_1))

Scanned by CamScanner

Question 4 a

question_4_a.py

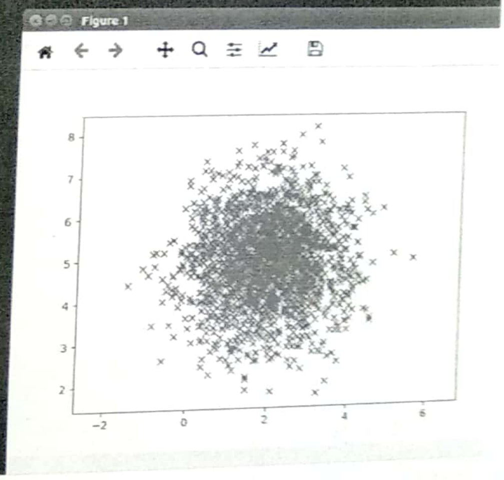
```python
import numpy as np
import matplotlib.pyplot as plt

def generate_random_variables(mean, covariance_matrix, size = 2000):
    # checks could be made to enforce that both are matrices ,
    # that is d >1 but since it is not specified and d could also be 1
    # this method is open ended
    return np.random.multivariate_normal(mean, covariance_matrix, size)


if __name__ == '__main__':
    shape = int(input("Enter the number of entries of mean"))
    print("please enter mean entries  in one line (separated by space):  ")
    entries = list(map(int, input().split()))
    mean = np.array(entries).reshape(shape)
    print(mean)
    rows_cov = int(input("Enter the number of rows of covariance_matrix"))
    columns_cov = int(input("Enter the number of columns of covariance_matrix "))
    print("please enter covariance matrix entries in one line (separated by space):  ")
    cov = list(map(int, input().split()))
    covariance_matrix = np.array(cov).reshape(rows_cov, columns_cov)
    print(covariance_matrix)
    x = generate_random_variables(mean, covariance_matrix)
    print("generated output is : ", x)
    print("shape of output : ", x.shape)
    x_vector, y_vector = x.T
    plt.plot(x_vector, y_vector, 'x')
    plt.axis('equal')
    plt.show()
```

Ln 7, Col 8   Spaces: 4   UTF-8   LF   Python

```
(pysyft) venktesh@venktesh-G7-7590:~/iiit-journey-books-papers/Statistical Machine learning/assignment1$ python question_4_a.py
Enter the number of entries of mean2
please enter mean entries  in one line (separated by space):
2 5
[2 5]
Enter the number of rows of covariance_matrix2
Enter the number of columns of covariance_matrix 2
please enter covariance matrix entries in one line (separated by space):
1 0 0 1
[[1 0]
 [0 1]]
generated output is :  [[1.01743088 4.9031965 ]
 [1.5524852  4.34629623]
 [1.39604602 5.78029152]
 ...
 [0.96562378 6.17077628]
 [0.01639439 3.37787314]
 [1.66793477 4.59813154]]
shape of output :  (2000, 2)
```



Question 4 b

a1_sml_phd19016.png

A1_Phd19016_question4_b.py - A1_Phd19016 - Visual Studio Code

EXPLORER

OPEN EDITORS

A1_Phd19016_question4_b.py

A1_PHD19016_question_4_a.py

A1_PHD19016

A1_PHD19016_question_4_a.py

A1_Phd19016_question4_b.py

A1_Phd19016_question4_b.py × A1_PHD19016_question_4_a.py

```
import numpy as np
def discriminant_fun(feature_vector, mean,variance, prior):
    #ignoring unimportant additive constants
    # params feature_vector : x a d * 1 dim array
    # mean d * 1 dim array of mean
    #prior(unless all priors are equal this must be included)
    g = (0.5 * (1/(variance **2)) * mean.T).dot(feature_vector)
    - (1/(2*(variance **2))* ((mean).T).dot(mean))
    + np.log(prior)
    return float(g)


if __name__ == '__main__':
    rows_mean = int(input("Enter the number of rows of mean"))
    columns_mean = int(input("Enter the number of columns of mean "))
    print("please enter mean entries  in one line (separated by space):  ")
    entries = list(map(int, input().split()))
    mean = np.array(entries).reshape(rows_mean,columns_mean)
    rows_x = int(input("Enter the number of rows of feature_matrix"))
    columns_x = int(input("Enter the number of columns of feature matrix ")) space: str
    print("please enter covariance matrix entries in one line (separated by space):  ")
    x_vec = list(map(int, input().split()))
    x_vector = np.array(x_vec).reshape(rows_x, columns_x)
    prior = float(input("enter prior probability"))
    variance = float(input("enter variance values"))
    discriminant_value = discriminant_fun(x_vector, mean,variance, prior )
    print(discriminant_value)
```

OUTLINE

MAVEN PROJECTS

10/2020

Question 4 b

sml-a1-phd19016-q4_b-solved.png

```
venktesh@venktesh-G7-7590: ~/iiit-journey-books-papers/Statistical Machine learning/A1_Phd19016
1 1
Enter the number of rows of feature_matrix2
Enter the number of columns of feature_matrix 1
please enter covariance matrix entries in one line (separated by space):
2 4
enter prior probability1
enter variance values4
0.125
(base) venktesh@venktesh-G7-7590:~/iiit-journey-books-papers/Statistical Machine learning/A1_Phd19016$ python A1_Phd19016_question4_b.py
Enter the number of rows of mean2
Enter the number of columns of mean 1
please enter mean entries  in one line (separated by space):
1 1
Enter the number of rows of feature_matrix2
Enter the number of columns of feature_matrix 1
please enter covariance matrix entries in one line (separated by space):
1 10
enter prior probability1
enter variance values2
1.375
(base) venktesh@venktesh-G7-7590:~/iiit-journey-books-papers/Statistical Machine learning/A1_Phd19016$ clear
(base) venktesh@venktesh-G7-7590:~/iiit-journey-books-papers/Statistical Machine learning/A1_Phd19016$ python A1_Phd19016_question4_b.py
Enter the number of rows of mean2
Enter the number of columns of mean 1
please enter mean entries  in one line (separated by space):
1 2
Enter the number of rows of feature_matrix2
Enter the number of columns of feature_matrix 1
please enter feture matrix entries in one line (separated by space):
3 5
enter prior probability1
enter variance values4
0.40625
(base) venktesh@venktesh-G7-7590:~/iiit-journey-books-papers/Statistical Machine learning/A1_Phd19016$ █
```

Question 4 c: compare the discriminants function's values for two different distributions. $N(\mu_1 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \Sigma_1 = \sigma^2 I)$

& $N(\mu_2 = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}; \Sigma_2 = \sigma^2 I)$ in $d = 2$ dimensions

$X = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $P(\omega_1) = 1/3$ & $P(\omega_2) = 2/3$.

The form of a multivariate normal density function is

$$\cancel{g_i(x)}\ N(\mu, \Sigma) = p(x) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} e^{\left(-\frac{1}{2}(x-\mu)^t \Sigma^{-1} (x-\mu)\right)}$$

The discriminant function is of form.

$$g_i(x) = \ln\left(\frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}}\right) - \frac{1}{2}(x-\mu_i)^t \Sigma_i^{-1} (x-\mu_i) + \ln P(\omega_i)$$

(prior should be included if all priors are not equal).

ignoring additive constants as $(\Sigma = \sigma^2 I)$

$$g_i(x) = -\frac{1}{2}\Big(\underbrace{x^T \Sigma_i^{-1} x}_{\text{can be ignored}} - x^T \Sigma_i^{-1} \mu - \mu_i^T \Sigma_i^{-1} x + \mu_i^T \Sigma^{-1} \mu_i\Big) + \ln(P(\omega_i))$$

$$= \frac{1}{2}\left(2\mu_i^T \Sigma^{-1} x - \mu_i^T \Sigma^{-1} \mu_i\right) + \ln(P(\omega_i))$$

$$= \frac{\mu_i^t x}{\sigma^2} - \frac{\mu_i^T \mu_i}{2\sigma^2} + \ln(P(\omega_i)) \quad -①$$

Substituting given values in ① & assuming non-uniform priors

$$g_1(x) = \frac{1}{\sigma^2}\begin{bmatrix}0.5\\0.5\end{bmatrix}^t \begin{bmatrix}1\\1\end{bmatrix} - \frac{1}{2\sigma^2}\begin{bmatrix}0.5\\0.5\end{bmatrix}^t \begin{bmatrix}0.5\\0.5\end{bmatrix} + \ln(P(\omega_1))$$

$$= \frac{1}{\sigma^2}\left[[0.5 \ 0.5]\begin{bmatrix}1\\1\end{bmatrix} - \frac{1}{2}[0.5 \ 0.5]\begin{bmatrix}0.5\\0.5\end{bmatrix}\right] + \ln(P(\omega_1))$$

$$= \frac{1}{\sigma^2}\left[1 - \frac{1}{2}(0.5)\right] = \frac{0.75}{\sigma^2} + \ln(P(\omega_1))$$

$$= \frac{0.75}{\sigma^2} + \ln(1/3)$$

$$g_2(x) = \frac{1}{\sigma^2}\left[\begin{bmatrix}-0.5\\0.5\end{bmatrix}^t \begin{bmatrix}1\\1\end{bmatrix} - \frac{1}{2}\begin{bmatrix}-0.5\\0.5\end{bmatrix}^t \begin{bmatrix}-0.5\\0.5\end{bmatrix}\right] + \ln(P(\omega_2))$$

$$= \frac{1}{\sigma^2}\left[[-0.5 + 0.5] - \frac{1}{2}(0.25 + 0.25)\right] + \ln\left(\frac{2}{3}\right)$$

$$= \frac{-0.25}{\sigma^2} + \log\left(\frac{2}{3}\right)$$

Scanned by CamScanner

$\therefore g_1(x) = \frac{0.75}{\sigma^2} + \ln\left(\frac{1}{3}\right)$

$g_2(x) = \frac{-0.25}{\sigma^2} + \ln\left(\frac{2}{3}\right)$

~~prior for w_2 is higher decision boundary might favour w_1 but an interesting comparison of disc. functions is that when $\sigma = 1$ $g_1(x) > g_2(x)$ but when $\sigma = 2, 3$ & so on $g_2(x) > g_1(x)$.~~

$\ln(2/3) = -0.4054$

$\ln(1/3) = -1.098$

when $\sigma = 1$ $g_1(x) = 0.75 - 1.098 = -0.348$

$g_2(x) = -0.25 - 0.4054 = -0.6554$

when $\sigma = 2$ $g_1(x) = \frac{0.75}{4} - 1.098 = -0.9105$

$g_2(x) = \frac{-0.25}{4} - 0.4054 = -0.4679$

As σ increases $g_2(x) > g_1(x)$

As $P(w_2) > P(w_1)$ decision boundary would favour class 2 more.