UNIVERSITY OF PATRAS
SCHOOL OF ENGINEERING
Department of Computer Engineering & Informatics

Division of Applications and Foundations of Computer Science
Pattern Recognition Laboratory
Name:Παναγιώτης Κωνσταντίνος Αθανασόπουλος
A.M:1058112

Εργαστηριακή Άσκηση για το μάθημα Θεωρία Αποφάσεων

2019- 2020

Μέρος Α’

Ερώτημα 1.

Θεωρείστε τα 2 διαστάσεων δεδομένα από δύο κλάσεις ω_1 και ω_2 και κάθε μία από αυτές ακολουθεί την Gaussian κατανομή $p(x/\omega_k) \sim N(\mu_k, \Sigma_k)$.

Πίνακας 1: Δεδομένα από τις κλάσεις ω_1 και ω_2

ω_1	ω_2
(0, 0)	(6, 9)
(0, 1)	(8, 9)
(2, 2)	(9, 8)
(3, 1)	(9, 9)
(3, 2)	(9,10)
(3, 3)	(8, 11)

1. Ποια είναι η εκ των προτέρων πιθανότητα για κάθε κλάση; ($P(\omega_1)$ και $P(\omega_2)$).
   Η εκ των προτέρων πιθανότητα για κάθε κλάση είναι αντίστοιχα:

$P(\omega_1)=\frac{6}{12}=\frac{1}{2}=0{,}5$

$P(\omega_2)=\frac{6}{12}=\frac{1}{2}=0{,}5$

2. Να υπολογίσετε τη μ.τ. και τον πίνακα συνδιασποράς, για κάθε κλάση.

η μέση τιμή της κλάσης ω_1:

$$\mu_{\omega 1}=\begin{pmatrix}0 & 0\\ 0 & 1\\ 2 & 2\\ 3 & 1\\ 3 & 2\\ 3 & 3\end{pmatrix}$$

Για το διάνυσμα της 1ης στήλης ισχύει: $\frac{0+0+2+3+3+4}{6}=$

$\mu_{\omega 1}=(1{,}83 \quad 1{,}5)$

η μέση τιμή της κλάσης ω_2:

$$\mu_{\omega 2}=\begin{pmatrix}6 & 9\\ 8 & 9\\ 9 & 8\\ 9 & 9\\ 9 & 10\\ 8 & 11\end{pmatrix}$$

$\mu_{\omega 2}=(8{,}1 \quad 9{,}3)$

Ο πίνακας συνδιασποράς για την κλάση ω_1:

$\Sigma_1=E[(X-\mu\omega 1)^T(X-\mu\omega 1)]$

=

$$\begin{pmatrix}0-1{,}83 & 0-1{,}83 & 2-1{,}83 & 3-1{,}83 & 3-1{,}83 & 3-1{,}83\\ 0-1{,}5 & 1-1{,}5 & 2-1{,}5 & 1-1{,}5 & 2-1{,}5 & 3-1{,}5\end{pmatrix}\begin{pmatrix}0-1{,}83 & 0-1{,}5\\ 0-1{,}83 & 1-1{,}5\\ 2-1{,}83 & 2-1{,}5\\ 3-1{,}83 & 1-1{,}5\\ 3-1{,}83 & 2-1{,}5\\ 3-1{,}83 & 3-1{,}5\end{pmatrix}$$

$=\begin{pmatrix} 2.1667 & 1.1000 \\ 1.1000 & 1.1000 \end{pmatrix}$

Ο πίνακας συνδιασποράς για την κλάση ω_2:

$\Sigma_2 = E[(X - \mu\omega 2\ )^T (X - \ \mu\omega 2\ )]$

=

$$\begin{pmatrix} 6-8{,}1 & 8-8{,}1 & 9-8{,}1 & 9-8{,}1 & 9-8{,}1 & 8-8{,}1 \\ 9-9{,}3 & 9-9{,}3 & 8-9{,}3 & 9-9{,}3 & 10-9{,}3 & 11-9{,}3 \end{pmatrix} \begin{pmatrix} 6-8{,}1 & 9-9{,}3 \\ 8-8{,}1 & 9-9{,}3 \\ 9-8{,}1 & 8-9{,}3 \\ 9-8{,}1 & 9-9{,}3 \\ 9-8{,}1 & 10-9{,}3 \\ 8-8{,}1 & 11-9{,}3 \end{pmatrix} =$$

$$\begin{pmatrix} 1.3667 & -0.0667 \\ -0.0667 & 1.0667 \end{pmatrix}$$

3. Να παράγετε την εξίσωση για το όριο απόφασης που διαχωρίζει τις δύο κλάσεις και να σχεδιάσετε το όριο απόφασης (σημείωση: Αν θέλετε μπορείτε να χρησιμοποιήσετε την εκ των υστέρων πιθανότητα $p(\omega_i/x)$).

Η διακρίνουσα συνάρτηση για την κλάση ω1 είναι:

$g_1(x,y) = -1/2(x - \mu\omega 1)^T \Sigma 1^{-1}(x - \mu\omega 1) - 1/2\ln|\Sigma 1| + \log P(\omega 1)$

$= -\frac{1}{2}(x - 1{,}83 \quad y - 1{,}5) \begin{pmatrix} 1/2.1667 & 1/1.1000 \\ 1/1.1000 & 1/1.1000 \end{pmatrix} \begin{pmatrix} x - 1{,}83 \\ y - 1{,}5 \end{pmatrix} - \frac{1}{2}\ln 1{,}17 + \log\frac{1}{2}$

$= -\frac{1}{2}\,[(x-1{,}83)*(\frac{x}{2{,}1667} + \frac{y}{1{,}1} - \frac{5{,}26}{2{,}38}) + (y-1{,}5)*(\frac{x}{1{,}1} + \frac{y}{1{,}1} * \frac{3{,}33}{1{,}1})] - \frac{1}{2}\ln 1{,}17 + \log\frac{1}{2}$

Η διακρίνουσα συνάρτηση για την κλάση ω2 είναι:

$g_2(x,y) = -\frac{1}{2}(x - \mu\omega 2)^T \Sigma 1^{-1}(x - \mu\omega 2) - 1/2\ln|\Sigma 2| + \log P(\omega 2)$

$= -\frac{1}{2}(x - 8{,}1 \quad y - 9{,}3) \begin{pmatrix} 1/1.3667 & 1/-0.0667 \\ 1/-0.0667 & 1/1.0667 \end{pmatrix} \begin{pmatrix} x - 8{,}1 \\ y - 9{,}3 \end{pmatrix} - \frac{1}{2}\ln 1.45 + \log\frac{1}{2}$

$= -\frac{1}{2}[(y-9{,}3)*(\frac{y}{1{,}0667} - \frac{x}{0{,}0667} + \frac{8{,}01}{7{,}11}) + (x-8{,}1)(\frac{x}{1{,}3667} - \frac{y}{0{,}0667} + \frac{1{,}21}{9{,}11})] - \frac{1}{2}\ln 1{,}45 + \log\frac{1}{2}$

Η εξίσωση του οριού απόφασης που διαχωρίζει τις δύο κλάσεις είναι:

$g_1(x,y) = g_2(x,y) \Leftrightarrow$

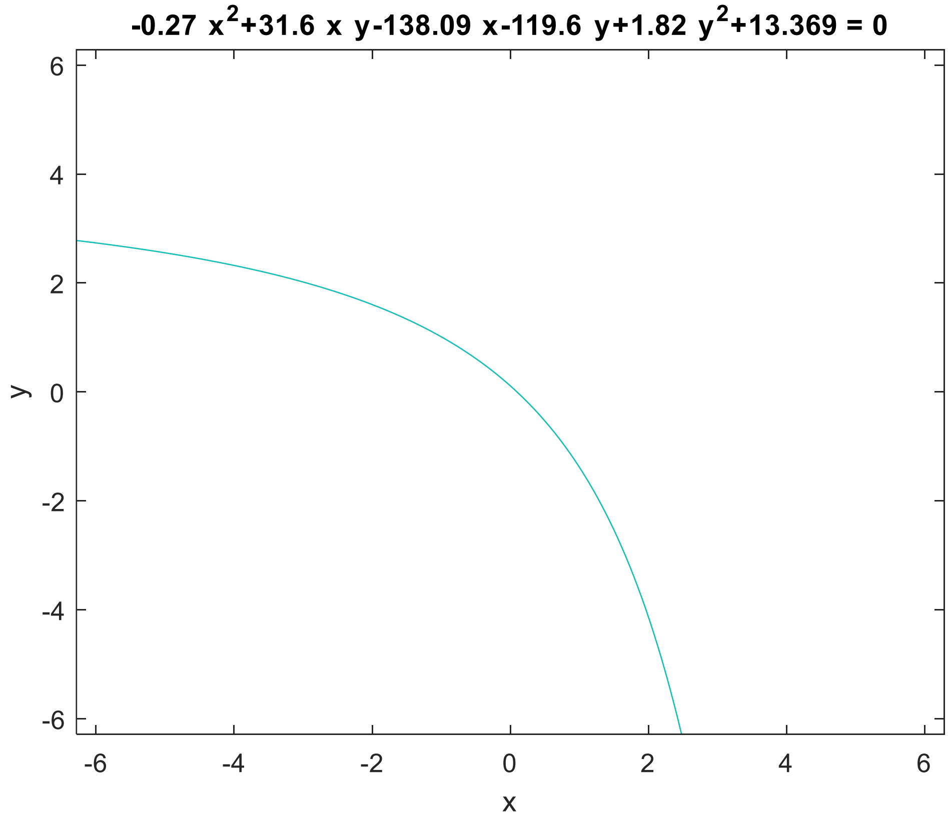
$-\frac{1}{2}$ [(x-1,83)*($\frac{x}{2,1667}+\frac{y}{1,1}-\frac{5,26}{2,38}$)+(y-1,5)*($\frac{x}{1,1}+\frac{y}{1,1}*\frac{3,33}{1,1}$)]-$\frac{1}{2}$ln1,17 +log$\frac{1}{2}$=-$\frac{1}{2}$[(y-9,3)*($\frac{y}{1,0667}$-$\frac{x}{0,0667}+\frac{8,01}{7,11}$)+(x-8,1)($\frac{x}{1,3667}-\frac{y}{0,0667}+\frac{1,21}{9,11}$)]-$\frac{1}{2}$ln1,45+log$\frac{1}{2}$⇔

Καθως P(ω1)=P(ω2) απαλοίφω τα log$\frac{1}{2}$

[(x-1,83)*($\frac{x}{2,1667}+\frac{y}{1,1}-\frac{5,26}{2,38}$)+(y-1,5)*($\frac{x}{1,1}+\frac{y}{1,1}*\frac{3,33}{1,1}$)]+ ln1,17= [(y-9,3)*($\frac{y}{1,0667}-\frac{x}{0,0667}+\frac{8,01}{7,11}$)+(x-8,1)($\frac{x}{1,3667}-\frac{y}{0,0667}+\frac{1,21}{9,11}$)]+ ln1,45⇔

(x – 1,83)*(x/2,1667 + y/1,1 -2,21) - (x – 8,1)*(x/1,3667 - y/0,0667 +0,13) + (x/1,1 + 2,75y)*(y - 3/2) - (y – 9,3)*(y/1,0667 - x/0,0667 +1,12) -2,14=0⇔

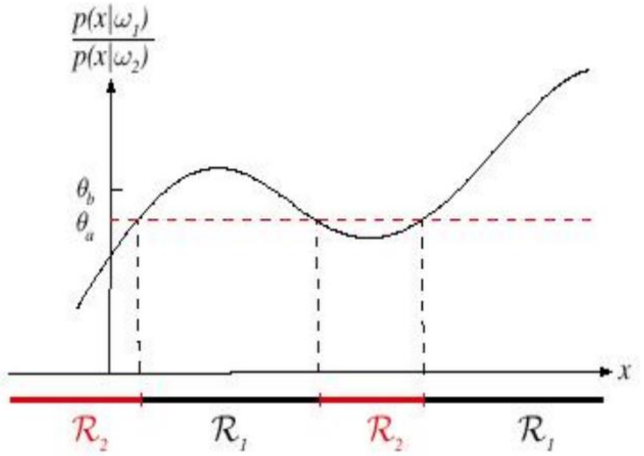
-0,27x^2+31,6xy-138,09x-119,6y+1,82y^2+13,369=0



4. Να θεωρήσετε την περίπτωση όπου τα κόστη λάθους ταξινόμησης είναι διαφορετικά για τις 2 κλάσεις (δηλαδή δεν είναι 0-1). Θα επηρεάσει αυτό το όριο απόφασης και πως;

Όταν χρησιμοποιείται η μηδέν-ένα συνάρτηση κόστους,τα όρια απόφασης καθορίζονται από τιμή κατωφλίου ίση με θ.
Εάν το λάθος της ταξινόμησης στην κλάση ω1, δειγμάτων που ανήκουν στην κλάση ω2, τιμωρείται περισσότερο από το αντίστροφο, δηλαδή εάν λ12 > λ21, τότε η εξίσωση $\frac{p(x/\omega1)}{p(x/\omega2)} > \frac{\lambda12-\lambda22}{\lambda21-\lambda11}\frac{P(\omega2)}{P(\omega1)}$οδηγεί στο κατώφλι θb. Σημειώνεται ότι το εύρος των τιμών του x για τις οποίες ένα πρότυπο ταξινομείται στην κλάση ω1 μικραίνει.Εάν η συνάρτηση κόστους τιμωρεί τη λάθος κατηγοριοποίηση στην ω1 δειγμάτων που ανήκουν

στην ω2 περισσότερο από το αντίστροφο, προκύπτει το κατώφλι θb, και ως αποτέλεσμα η περιοχή απόφασης R1 γίνεται μικρότερη.



5. Θεωρήστε δύο κλάσεις ω_1 και ω_2 στο χώρο προτύπων Ω με συνεχή κατανομή πιθανότητας $p_1(x)$ και $p_2(x)$ αντίστοιχα. Το πρόβλημα ταξινόμησης δύο κατηγοριών μπορεί να διατυπωθεί σαν διαίρεση του

   χώρου Ω σε δύο εξαντλητικά και μη επικαλυπτόμενα σύνολα Ω_1 και Ω_2, έτσι ώστε $\Omega_1 \cup \Omega_2 = \Omega$ και

   $\Omega1 \cap \Omega2 = 0$. Αν το xi $\in \Omega_k$ τότε αντιστοίχισε το x_i στην κλάση ω_k.

   5.1 θεωρείστε ότι δίνεται μια διακρίνουσα συνάρτηση f(.). Να αναφέρετε τα δύο λάθη που μπορεί να κάνει αυτή η συνάρτηση.

Οι περιπτώσεις για τις οποίες μπορεί να συμβεί κάποιο λάθος ταξινόμησης είναι δύο.Είτε ένα δείγμα x βρίσκεται στην περιοχή R2 ενώ η πραγματική κατάσταση της φύσης είναι η ω1,είτε βρίσκεται στην περιοχή R1 ενώ η πραγματική κατάσταση της φύσης είναι η ω2.

5.2 Να γράψετε την πιθανότητα λάθους που αντιστοιχεί σε αυτά τα δύο λάθη.

Η πιθανότητα λάθους για την περίπτωση που το δείγμα x βρίσκεται στην περιοχή R2 ενώ η πραγματική κατάσταση της φύσης είναι η ω1 είναι:

$P(x \in R2, \omega 1) = \int_{R2} p(x/\omega 1)P(\omega 1)\, dx$

Η πιθανότητα λάθους για την περίπτωση που το δείγμα x βρίσκεται στην περιοχή R1 ενώ η πραγματική κατάσταση της φύσης είναι η ω2 είναι:

$P(x \in R1, \omega 2) = \int_{R1} p(x/\omega 2)P(\omega 2)\, dx$

5.3 Υποθέστε ότι τα κόστη για τους δύο τύπους λαθών είναι c1 και c2. Να γράψετε το συνολικό αναμενόμενο κόστος.

Το κόστος για το 1 είναι:

R1 = c1*P(ω1|x)

Το κόστος για το 2 είναι:

R2 = c2*P(ω2|x)

Το συνολικό αναμενόμενο ρίσκο είναι:

R = c1*P(ω1|x)+c2*P(ω2|x)

6. Υποθέστε ότι έχουμε ένα πρόβλημα ταξινόμησης δύο κατηγοριών, σολομός (ω_1) και πέρκα (ω_2).

   6.1Πρώτα, υποθέτουμε ότι έχουμε ένα χαρακτηριστικό, και οι σ.π.π. είναι Gaussians $N(0, \sigma^2)$ και $N(1, \sigma^2)$ για τις δύο κλάσεις αντίστοιχα. Να δείξετε ότι το κατώφλι που ελαχιστοποιεί το ελάχιστο ρίσκο είναι:

   $$r = \frac{1}{2} - \sigma^2 ln \frac{\lambda_{12} P(\omega_2)}{\lambda_{21} P(\omega_1)}$$

   Όπου έχουμε υποθέσει ότι $\lambda_{11}=\lambda_{22}=0$.

Από:

$R(\omega_1|r)=\lambda_{11}P(\omega_1|r)+\lambda_{12}P(\omega_2|r)=\lambda_{12}P(\omega_2|r)$

$R(\omega_2|r)=\lambda_{22}P(\omega_2|r)+\lambda_{21}P(\omega_1|r)=\lambda_{21}P(\omega_1|r)$

Πρέπει:

$\lambda_{12}P(\omega_2|r)= \lambda_{21}P(\omega_1|r)\Leftrightarrow$

$\lambda_{12}P(\omega_2|r)\ P(\omega_2)= \lambda_{21}P(\omega_1|r)\ P(\omega_1)\Leftrightarrow$

$$\frac{\lambda_{12}P(\omega_2)}{\lambda_{21}P(\omega_1)}=\frac{P(r|\omega_1)}{P(r|\omega_2)}\Leftrightarrow$$

$$\frac{\lambda_{12}P(\omega_2)}{\lambda_{21}P(\omega_1)}=\frac{\frac{1}{\sqrt{2\pi\sigma}}\exp[-\frac{1}{2\sigma^2}(r-\mu_1)\verb|^|2]}{\frac{1}{\sqrt{2\pi\sigma}}\exp[-\frac{1}{2\sigma^2}(r-\mu_2)\verb|^|2]}\Leftrightarrow$$

$$\ln\frac{\lambda_{12}P(\omega_2)}{\lambda_{21}P(\omega_1)}=\ln(\frac{1}{\sqrt{2\pi\sigma}})-\frac{1}{2\sigma^2}(r-0)\verb|^|2-[\ln(\frac{1}{\sqrt{2\pi\sigma}})-\frac{1}{2\sigma^2}(r-1)\verb|^|2]\Leftrightarrow$$

$$\ln\frac{\lambda_{12}P(\omega_2)}{\lambda_{21}P(\omega_1)}= -\frac{r^2}{2\sigma^2}+\frac{r^2}{2\sigma^2}-\frac{2r}{2\sigma^2}+\frac{1}{2\sigma^2}\Leftrightarrow$$

$$\frac{r}{\sigma^2}=\frac{1}{2\sigma^2} -\ln\frac{\lambda_{12}P(\omega_2)}{\lambda_{21}P(\omega_1)}\Leftrightarrow$$

$$r=\frac{1}{2} -\sigma^2\ln\frac{\lambda_{12}P(\omega_2)}{\lambda_{21}P(\omega_1)}$$

   6.2 Μετά, υποθέτουμε ότι έχουμε δύο χαρακτηριστικά $\mathbf{x} = (x_1, x_2)$ και οι υπό συνθήκη πυκνότητες δύο κλάσεων p(χ/ω=1) και p(χ/ω=2), είναι 2-Δ gaussians κατανομές με κέντρα στα σημεία (4, 11) και (10, 3) αντίστοιχα με τον ίδιο πίνακα συνδιασποράς Σ=3I (όπου I είναι ο μοναδιαίος πίνακας). Υποθέστε ότι οι a priori πιθανότητες είναι P(ω=1) = 0.6 και P(ω=2) = 0.4.

   (α) Υποθέστε ότι χρησιμοποιούμε τον Κανόνα Απόφασης του Bayes. Να γράψετε τις διακρίνουσες συναρτήσεις $g_1(x)$ και $g_2(x)$.

Η διακρίνουσα συνάρτηση του της κλάσης ω1 είναι:

$g1(x)=-\frac{1}{2\sigma^2}(x-\mu i)^T(x-\mu i)+log(\omega i)$

$= -\frac{1}{2}\begin{pmatrix} x-\mu x \\ y-\mu y \end{pmatrix}^T \begin{pmatrix} 1/3 & 0 \\ 0 & 1/3 \end{pmatrix} \begin{pmatrix} x-\mu x \\ y-\mu x \end{pmatrix} + logP(\omega 1) = -\frac{1}{2}\begin{pmatrix} x-4 \\ y-11 \end{pmatrix}^T \begin{pmatrix} 1/3 & 0 \\ 0 & 1/3 \end{pmatrix} \begin{pmatrix} x-4 \\ y-11 \end{pmatrix} + logP(\omega 1)$

$= -\frac{1}{2}\begin{pmatrix} x-4 \\ y-11 \end{pmatrix}^T \begin{pmatrix} 1/3 & 0 \\ 0 & 1/3 \end{pmatrix} \begin{pmatrix} x-4 \\ y-11 \end{pmatrix} + log(0,6) = -\frac{1}{2}(x/3-4/3)*(x-4)-\frac{1}{2}(y/3-11/3)*(y-11)+log(0,6)$

$= -x^2+8x-y^2+22y-137+6log(0,6)$

Η διακρίνουσα συνάρτηση του της κλάσης ω2 είναι:

$g2(x,y)=-\frac{1}{2\sigma^2}(x-\mu i)^T(x-\mu i)+log(\omega i)= = -\frac{1}{2}\begin{pmatrix} x-\mu x \\ y-\mu y \end{pmatrix}^T \begin{pmatrix} 1/3 & 0 \\ 0 & 1/3 \end{pmatrix} \begin{pmatrix} x-\mu x \\ y-\mu x \end{pmatrix} + logP(\omega 2)$

$= -\frac{1}{2}\begin{pmatrix} x-10 \\ y-3 \end{pmatrix}^T \begin{pmatrix} 1/3 & 0 \\ 0 & 1/3 \end{pmatrix} \begin{pmatrix} x-10 \\ y-3 \end{pmatrix} + logP(0,4)$

$=-\frac{1}{2}(x/3-10/3)*(x-10)-\frac{1}{2}(y/3-1)*(y-3)+log(0,4)=-x^2+20x-y^2+6y-109+6log(0,4)$

(β) Να βρείτε την εξίσωση για το όριο απόφασης.

Η εξίσωση για το όριο απόφασης είναι:

$g1(x)=g2(x) \Leftrightarrow$

$-x^2+8x-y^2+22y-137+6log(0,6)= -x^2+20x-y^2+6y-109+6log(0,4) \Leftrightarrow$

$-x^2+8x-y^2+22y-137+6log(0,6)-(-x^2+20x-y^2+6y-109+6log(0,4))=0 \Leftrightarrow$

$-x^2+8x-y^2+22y-137+6log(0,6)+ x^2-20x+y^2-6y+109-6log(0,4)=0 \Leftrightarrow$

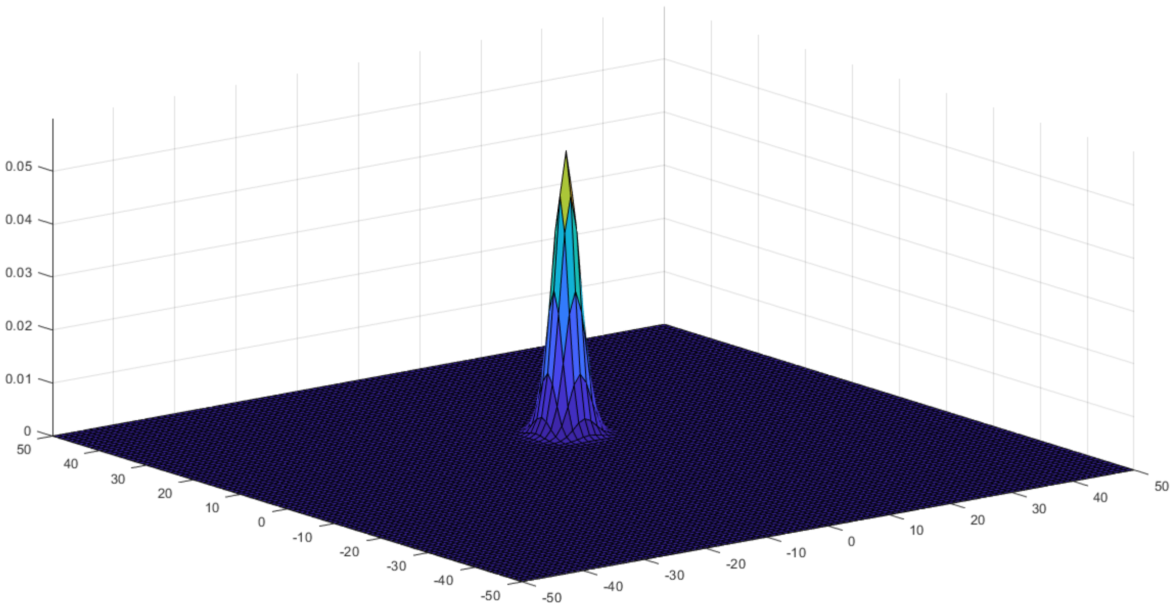
$-12x+16y-28-6log(3/2)=0 \Leftrightarrow y=(12*x+28+6*log(3/2))/16$

(γ) Πως θα αλλάξει το όριο απόφασης αν αλλάξουν οι a priori πιθανότητες και η συνδιασπορά;

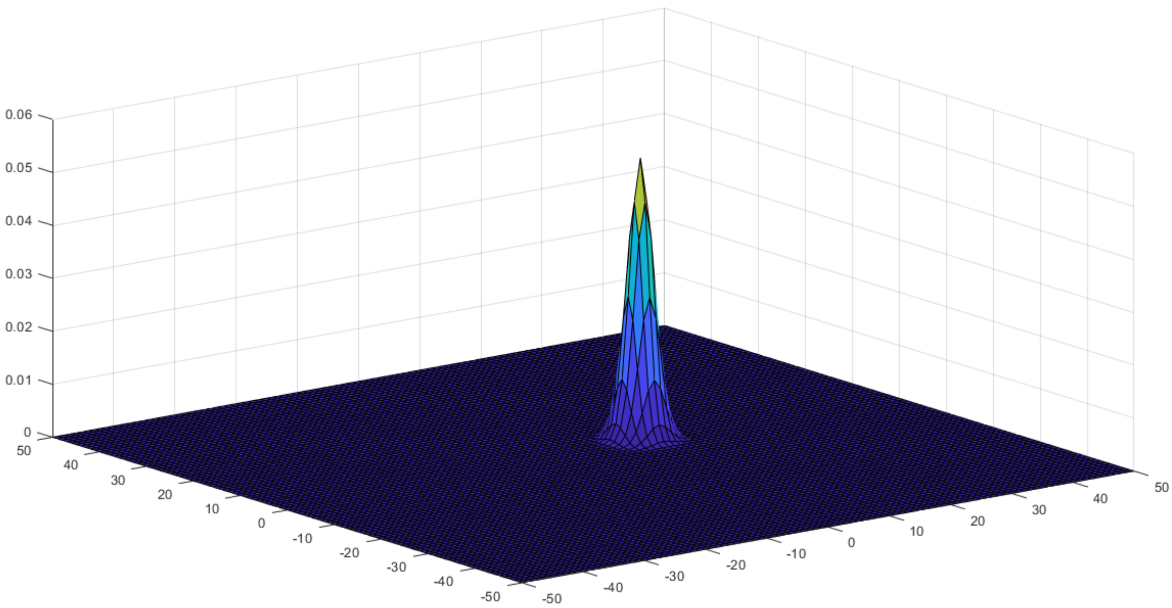
Επειδή w = μi - μj, το υπερεπίπεδο που χωρίζει τις περιοχές απόφασης Ri και Rj είναι ορθογώνιο στη γραμμή που συνδέει τα μέσα τους. Εάν P(ωi) = P(ωj), ο δεύτερος όρος στο δεξιό μέρος της εξίσωσης $x0=1/2(\mu i+\mu j)-\frac{\sigma^2}{||\mu i-\mu j||}\ln\frac{P(\omega i)}{P(\omega j)}(\mu i-\mu j)$ μηδενίζεται και επομένως το σημείο x0 βρίσκεται στο μέσο του τμήματος που ενώνει τα μέσα των κατηγοριών και το υπερεπίπεδο αποτελεί ένα κάθετο διχοτομητή αυτού του τμήματος. Εάν P(ωi) ≠ P(ωj), το σημείο x0 απομακρύνεται από το μέσο της πιο πιθανής κατηγορίας και πλησιάζει το μέσο της λιγότερο πιθανής.Εάν η διασπορά σ^2 είναι μικρή σε σχέση με την τετραγωνική απόσταση $||\mu i - \mu j||^2$, η θέση του ορίου απόφασης δεν επηρεάζεται σημαντικά από τις τιμές των εκ των προτέρων πιθανοτήτων. Εάν οι εκ των προτέρων πιθανότητες P(ωi) είναι ίδιες για όλες τις κατηγορίες , ο όρος lnP(ωi) μπορεί επίσης να απαλειφθεί. Σε αυτήν την περίπτωση, ο βέλτιστος κανόνας απόφασης παίρνει την εξής απλή μορφή: Για την ταξινόμηση ενός διανύσματος χαρακτηριστικών x, πρέπει να υπολογιστεί η Ευκλείδεια απόσταση του x από καθένα από τα c διανύσματα μέσων τιμών (που αντιστοιχούν στις c διαφορετικές κατηγορίες) και να τοποθετηθεί το x στην κατηγορία του κοντινότερου διανύσματος μέσης τιμής.

(δ) Χρησιμοποιώντας τη Matlab, να πάρετε 100 σημεία από κάθε μία κατανομή πυκνότητας. Να σχεδιάσετε τις δύο πυκνότητες (από τα δείγματα) και το όριο απόφασης στον 2-Δ χώρο.

Κατανομή πυκνότητας για την κλάση ω1:



Κατανομή πυκνότητας για την κλάση ω2:



Το όριο απόφασης: