Εργαστηριακή Άσκηση
για το μάθημα Θεωρία Αποφάσεων
2019 - 2020
Μέρος Α’

Παναγιώτης Χριστόπουλος 1054409

Ερώτημα 1:

ω1	ω2
(0, 0)	(6, 9)
(0, 1)	(8, 9)
(2, 2)	(9, 8)
(3, 1)	(9, 9)
(3, 2)	(9,10)
(3, 3)	(8, 11)

Η εκ των προτέρων πιθανότητα για κάθε κλάση προκύπτει από τα στοιχεία της κάθε κλάσης σε σχέση με τα συνολικά στοιχεία. Επομένως έχουμε για την κλάση Ω1 P(Ω1)=6/12 και αντίστοιχα για την Ω_2 P(Ω2)=6/12

Ερώτημα 2:

Για την κλάση Ω1: Αφού ακολουθείται κανονική κατανομή ξέρουμε ότι E(X)=μ=Σ xP(X=x). Δηλαδή το βρίσκουμε πολλαπλασιάζοντας το κάθε δεδομένο της κλάσης με την αντίστοιχη πιθανότητά του:

$\frac{1}{6}[(0,0)+(0,1)+(2,2)+(3,1)+(3,2)+(3,3)]=\frac{1}{6}(11,9)=\left(\frac{11}{6},\frac{9}{6}\right)$.

Ο πίνακας συνδιασποράς για το Ω1 προκύπτει απο τον τύπο $Cov(X)=\Sigma=E\{(x-\mu)(x-\mu)^T\}$. Αν αναλύσουμε τα δεδομένα της κλάσης Ω1 σε Ω1x και Ω2y για τον πίνακα συνδιασποράς θα έχουμε : $\Sigma_{11}=cov(\Omega1x1,\Omega1x1)$, $\Sigma_{12}=cov(\Omega1x1,\Omega1x2)\Sigma_{21}=cov(\Omega1x2,\Omega1x1),\Sigma_{22}=cov(\Omega1x2,\Omega1x2)$ και θα έχουμε τον εξής πίνακα:

2.1667	1.1000
1.1000	1.1000

Για την κλάση Ω_2: Αντίστοιχα για την Ω_2, ξέρουμε ότι $E(X)=\mu=\sum x\,P(X=x)$, και έχουμε: $\frac{1}{6}[(6,9)+(8,9)+(9,8)+(9,9)+(9,10)+(8,11)]=\left(\frac{49}{6},\frac{56}{6}\right)$. Για τον πίνακα συνδιασποράς , αφού αναλύσουμε τα δεδομένα της Ω_2 σε Ω_2x και Ω_2y χρησιμοποιούμε τα στοιχεία: Σ11=cov(Ω2x1, Ω2x1), Σ12= cov(Ω2x1, Ω2x2) Σ21= cov(Ω2x2, Ω2x1), Σ22= cov(Ω2x2, Ω2x2) και προκύπτει:

1.3667	-0.0667
-0.0667	1.0667

Ερώτημα 3:

Για να βρούμε το όριο απόφασης αρχικά πρέπει να βρούμε τις συναρτήσεις διαχωρισμού της κάθε κλάσης και μετά να τις εξισώσουμε. Διαπιστώνουμε ότι οι πίνακες συνδιασποράς είναι διαφορετικοί οπότε θα χρησιμοποιήσουμε τον τύπο: $g_i(x)=\frac{-1}{2}(x-\mu_i)^T\Sigma_i^{-1}(x-\mu_i)-\frac{1}{2}ln|\Sigma_i|+lnP(\omega_i)$.

Άρα θα χρειαστούμε τους αντίστροφους πίνακες συνδιασποράς για τις κλάσεις ω_1 και ω_2. Για την κλάση Ω_1 έχουμε:Σ_1^{-1}=

0,9375	-0,9375
-0,9375	1,8466

Για την κλάση Ω2 έχουμε :Σ_2^{-1}=

0,7339	0,0459
0,0459	0,9403

$$g_1(x) = \frac{-1}{2}\begin{pmatrix} x-\frac{11}{6} \\ y-\frac{9}{6} \end{pmatrix}^T \begin{pmatrix} 0.9375 & -0.9375 \\ -0.9375 & 1.8466 \end{pmatrix}\begin{pmatrix} x-\frac{11}{6} \\ y-\frac{9}{6} \end{pmatrix} - \frac{1}{2}ln(2.6101) + ln\frac{1}{2}$$

$$g_2(x) = \frac{-1}{2}\begin{pmatrix} x-\frac{49}{6} \\ y-\frac{56}{6} \end{pmatrix}^T \begin{pmatrix} 0.7339 & 0.0459 \\ 0.0459 & 0.9403 \end{pmatrix}\begin{pmatrix} x-\frac{49}{6} \\ y-\frac{56}{6} \end{pmatrix} - \frac{1}{2}ln(0.6880) + ln\frac{1}{2}$$

Στη συνέχεια θα εξισώσουμε τις 2 παραπάνω συναρτήσεις:

$$g_1(x) = g_2(x) \Rightarrow$$

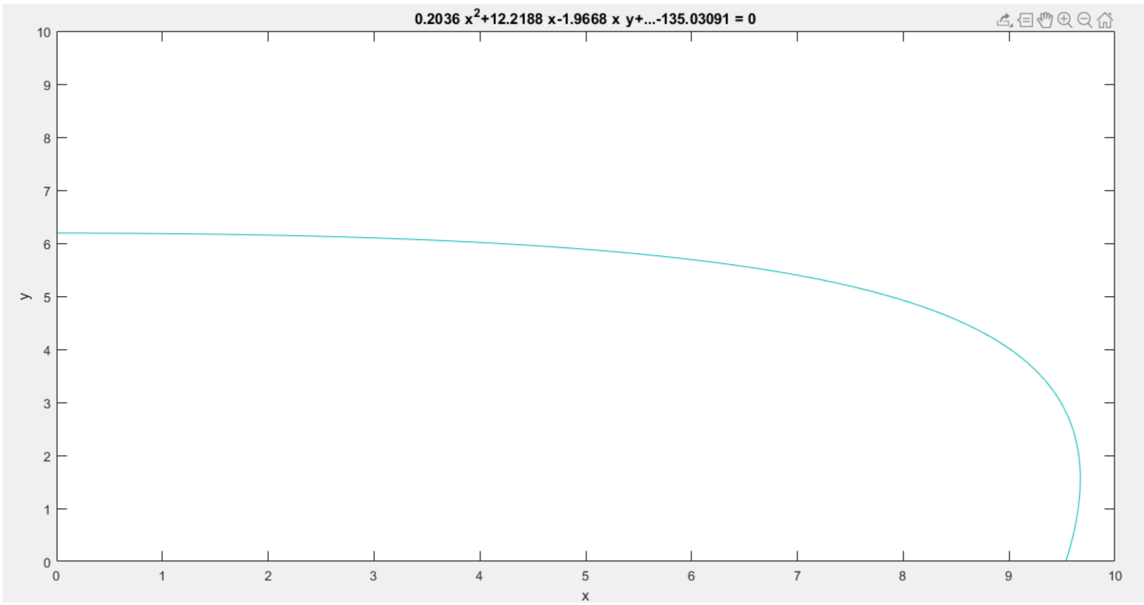
$$\frac{-1}{2}\begin{pmatrix} x-\frac{11}{6} \\ y-\frac{9}{6} \end{pmatrix}^T \begin{pmatrix} 0.9375 & -0.9375 \\ -0.9375 & 1.8466 \end{pmatrix}\begin{pmatrix} x-\frac{11}{6} \\ y-\frac{9}{6} \end{pmatrix} - \frac{1}{2}ln(2.6101) + ln\frac{1}{2}$$
$$= \frac{-1}{2}\begin{pmatrix} x-\frac{49}{6} \\ y-\frac{56}{6} \end{pmatrix}^T \begin{pmatrix} 0.7339 & 0.0459 \\ 0.0459 & 0.9403 \end{pmatrix}\begin{pmatrix} x-\frac{49}{6} \\ y-\frac{56}{6} \end{pmatrix} - \frac{1}{2}ln(0.6880) + ln\frac{1}{2} \Rightarrow$$

$$\begin{pmatrix} x-\frac{11}{6} \\ y-\frac{9}{6} \end{pmatrix}^T \begin{pmatrix} 0.9375 & -0.9375 \\ -0.9375 & 1.8466 \end{pmatrix}\begin{pmatrix} x-\frac{11}{6} \\ y-\frac{9}{6} \end{pmatrix} + ln(2.6101)$$
$$= \begin{pmatrix} x-\frac{49}{6} \\ y-\frac{56}{6} \end{pmatrix}^T \begin{pmatrix} 0.7339 & 0.0459 \\ 0.0459 & 0.9403 \end{pmatrix}\begin{pmatrix} x-\frac{49}{6} \\ y-\frac{56}{6} \end{pmatrix} + ln(0.6880) \Rightarrow$$

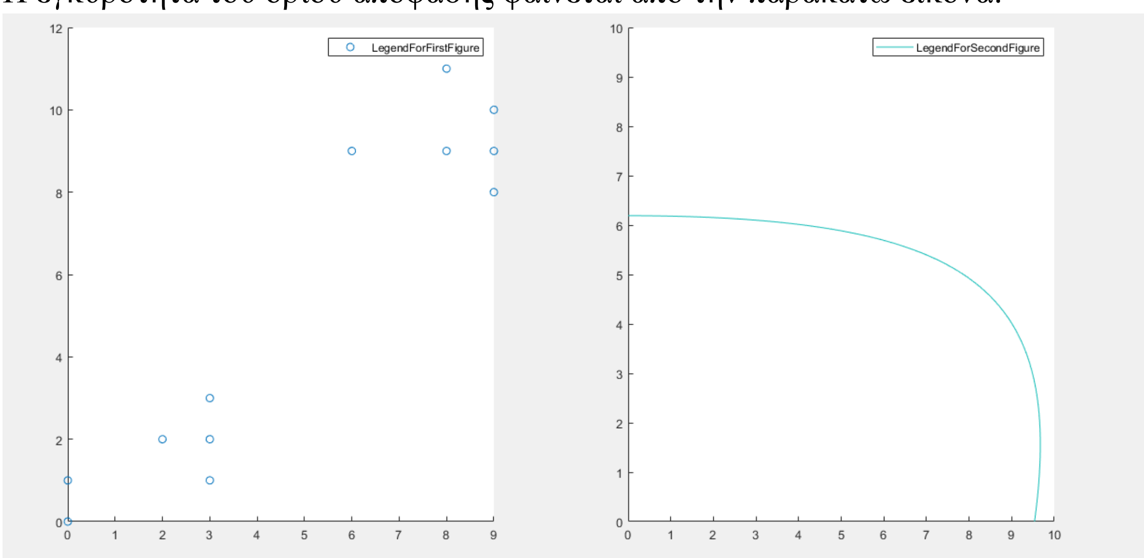
$$0.9375x^2 - 0.625x - 1.875xy + 1.8466y^2 - 2.1023y + 2.14969 + 0.9594$$
$$= 0.7339x^2 - 12.8438x + 0.0918xy - 18.3018y + 0.9403y^2 + 137.761$$
$$+ 0.3739 \Rightarrow$$

$$0.2036x^2 + 12.2188x - 1.9668xy + 16.1995y + 0.9063y^2 = 135.03091$$

Σύμφωνα με τη Matlab το παραπάνω όριο απόφασης απεικονίζεται ως εξής:



Η εγκυρότητα του ορίου απόφασης φαίνεται από την παρακάτω εικόνα:



Ερώτημα 4

Στις παραπάνω περιπτώσεις (με τα λάθη ταξινόμησης να είναι 0 και 1 αντίστοιχα), το μοναδικό κριτήριο για το όριο απόφασης ήταν οι εκ των υστέρων πιθανότητες. Αν όμως είναι διαφορετικά τα κόστη λάθος ταξινόμησης, θα ορίσουμε ένα κατώφλι το οποίο προκύπτει από τη σχέση $\frac{p(x/\omega_1)}{p(x/\omega_2)} > \frac{\lambda_{12}-\lambda_{22}}{\lambda_{21}-\lambda_{11}} \frac{P(\omega_2)}{P(\omega_1)}$ και ένα άλλο βασισμένο στην υπόθεση λάθους 0-1. Ανάλογα με την απόσταση των παραπάνω κατωφλιών στον πίνακα πιθανοφάνειας, θα μεταβληθεί το όριο απόφασης είτε προς την Ω_1 είτε προς την Ω_2 και η περιοχή απόφασης μιας από τις 2 κλάσεις θα γίνει μικρότερη.

Ερώτημα 5.1

Επειδή οι κλάσεις ω_1 και ω_2 έχουν συνεχή κατανομή πιθανότητας $p_1(x)$ και $p_2(x)$, υπάρχουν δύο περιπτώσεις λάθους. Τα δύο λάθη που μπορεί να κάνει η διακρίνουσα συνάρτηση f(.) είναι η ταξινόμηση του χ στην κλάση ω_1 ενώ ανήκει στην ω_2 και η ταξινόμηση του χ στην κλάση ω_2 ενώ ανήκει στην ω_1

Ερώτημα 5.2

Για την πρώτη περίπτωση, η πιθανότητα λάθους βάσει του διαγράμματος της εκ των υστέρων πιθανότητας είναι η ακόλουθη:

$$P(Erorr_1) = \int_{R_1} p\,(x \vee \omega_2) P(\omega 2) dx$$

Αντίστοιχα για τη δεύτερη, η πιθανότητα λάθους είναι:

$$P(Erorr_2) = \int_{R_2} p\,(x \vee \omega_1) P(\omega 1) dx$$

Άρα, η συνολική πιθανότητα λάθους αντιστοιχεί στο

$$P(Error_{συνολικό}) = \int_{R_1} p\,(x \vee \omega_2) P(\omega 2) dx + \int_{R_2} p\,(x \vee \omega_1) P(\omega 1) dx$$

Ερώτημα 5.3

Στην πρώτη περίπτωση που εξετάσαμε (ταξινόμηση χ στην ω_1 ενώ ανήκει στην ω_2) έχουμε:

$$c_1 = \lambda_{12} P(\omega_1 \vee x) = \lambda(\alpha_1 \vee \omega_2) P(\omega_1 \vee x)$$

Στη δεύτερη περίπτωση, αντίστοιχα έχουμε

$$c2 = \lambda 21 P(\omega 2 \vee x) = \lambda(\alpha 2 \vee \omega 1) P(\omega 2 \vee x)$$

Το συνολικό αναμενόμενο κόστος όμως περιέχει όχι μόνο τα c_1 και c_2, αλλά και τα κόστη σωστής επιλογής:

$$c_{(σωστήςεπιλογής)} = \lambda_{11} P(\omega_1 \vee x) + \lambda_{22} P(\omega_2 \vee x)$$

Άρα $c_{ολικό} = c_{(σωστήςεπιλογής)} + c_1 + c_2 \Rightarrow$

$$c_{ολικό} = \lambda_{11} P(\omega_1 \vee x) + \lambda_{12} P(\omega_1 \vee x) + \lambda_{21} P(\omega_2 \vee x) + \lambda_{22} P(\omega_2 \vee x)$$

Ερώτημα 6.1

Σε μία κανονική κατανομή ξέρουμε ότι αν τα χαρακτηριστικά είναι ανεξάρτητα και έχουν την ίδια διασπορά(Ω_1 N(0, σ^2) και Ω_2 N(1, σ^2)), για την ταξινόμηση ελάχιστου ρυθμού λάθους ισχύει η σχέση $x_0 = \frac{1}{2}(\mu_2 + \mu_1) - \frac{\sigma^2}{\|\mu_2 - \mu_1\|^2} ln \frac{P(\omega_2)}{P(\omega_1)} (\mu_2 - \mu_1)$ για $P(\Omega_2) > P(\Omega_1) \Rightarrow$

$$x_0 = \frac{1}{2}(1+0) - \frac{\sigma^2}{\|1-0\|^2} ln \frac{P(\omega_2)}{P(\omega_1)}(1-0) \Rightarrow$$

$$x_0 = \frac{1}{2} - \sigma^2 ln \frac{P(\omega_2)}{P(\omega_1)}$$

Επίσης ισχύει η σχέση :

$$\frac{p(x/\omega_1)}{p(x/\omega_2)} > \frac{\lambda_{12}-\lambda_{22}}{\lambda_{21}-\lambda_{11}} \frac{P(\omega_2)}{P(\omega_1)}$$

Γνωρίζοντας ότι λ_{22} και λ_{11} είναι 0, έχουμε ότι

$$\frac{p(x/\omega_1)}{p(x/\omega_2)} > \frac{\lambda_{12}}{\lambda_{21}} \frac{P\omega_2}{P\omega_1}$$

Από τις παραπάνω σχέσεις βγαίνει το ελάχιστο κατώφλι

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

Ερώτημα 6.2α

Ξέρουμε ότι για τις δύο κατανομές ισχύει $\sigma_1 = \sigma_2$. Οπότε για να βρούμε τις διακρίνουσες συναρτήσεις θα χρησιμοποιήσουμε τον τύπο

$$g_i(x) = -\frac{1}{2} \begin{bmatrix} x-\mu_x \\ y-\mu_y \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} x-\mu_x \\ y-\mu_y \end{bmatrix} + ln\big(P(\omega_i)\big)$$

Ξέρουμε ότι τα κέντρα των υπο συνθήκη πιθανοτήτων των δύο κλάσεων είναι (4,11) και (10,3) οπότε συμπεραίνουμε ότι μ_1=(4,11) και μ_2=(10,3)

Άρα έχουμε:

$$g_1(x) = -\tfrac{1}{2}\begin{bmatrix} x-4 \\ y-11 \end{bmatrix}^T \tfrac{1}{3}\Sigma \begin{bmatrix} x-4 \\ y-11 \end{bmatrix} + ln(0{,}6) =$$

$$-\tfrac{1}{2}\left[\tfrac{1}{3}(x-4)^2 + \tfrac{1}{3}(y-11)^2\right] + ln(0{,}6) =$$

$$-\frac{1}{6}[(x-4)^2 + (y-11)^2] + ln(0{,}6)$$

και

$$g_2(x) = -\tfrac{1}{2}\begin{bmatrix} x-10 \\ y-3 \end{bmatrix}^T \tfrac{1}{3}\Sigma \begin{bmatrix} x-10 \\ y-3 \end{bmatrix} + ln(0{,}4) =$$

$$-\tfrac{1}{2}\left[\tfrac{1}{3}(x-10)^2 + \tfrac{1}{3}(y-3)^2\right] + ln(0{,}4) =$$

$$-\frac{1}{6}[(x-10)^2 + (y-3)^2] + ln(0{,}4)$$

Ερώτημα 6.2β

Για να βρούμε την εξίσωση για το όριο απόφασης αρκεί να εξισώσουμε τις διακρίνουσες συναρτήσεις $g_1(x)$ και $g_2(x)$:

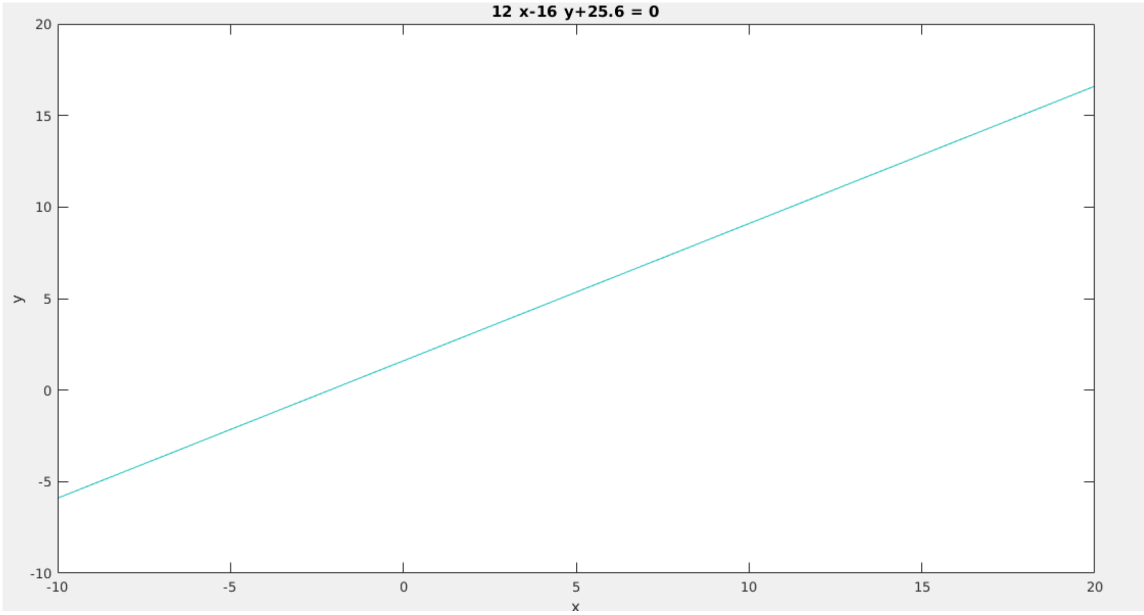
$$g_1(x) = g_2(x) \Rightarrow$$

$$-\frac{1}{6}[(x-4)^2 + (y-11)^2] + ln(0{,}6) = -\frac{1}{6}[(x-10)^2 + (y-3)^2] + ln(0{,}4) \Rightarrow$$

$$-\frac{1}{6}[(x-4)^2 + (y-11)^2] - \frac{1}{6}\left(-6ln(0{,}6)\right) = -\frac{1}{6}[(x-10)^2 + (y-3)^2] - \frac{1}{6}\left(-6ln(0{,}4)\right) \Rightarrow$$

$$[(x-4)^2 + (y-11)^2] - 6ln(0{,}6) = [(x-10)^2 + (y-3)^2] - 6ln(0{,}4) \Rightarrow$$

$$12x - 18y + 25.6 = 0$$



Ερώτημα 6.γ

Σύμφωνα με τον τύπο των διακρινουσών συναρτήσεων $g_i(x) = -\frac{1}{2}\begin{bmatrix} x-\mu_x \\ y-\mu_y \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} x-\mu_x \\ y-\mu_y \end{bmatrix} + ln\big(P(\omega_i)\big)$, η αύξηση της a priori πιθανότητας μιας κλάσης μεταβάλλει το όριο απόφασης υπέρ της, ενώ η αύξηση της συνδιασποράς το μεταβάλλει εις βάρος της.

Ερώτημα 6.2δ

Σύμφωνα με τη Matlab, 100 τυχαία σημεία από κάθε κατανομή και το όριο απόφασης(έγκυρο) για το ερώτημα είναι τα παρακάτω