

UNIVERSITY OF PATRAS

SCHOOL OF ENGINEERING

Department of Computer Engineering & Informatics

Division of Applications and Foundations

of Computer Science

Pattern Recognition Laboratory

Director: S. Likothanassis, Professor,

e-mail: likothan@ceid.upatras.gr

URL: http://prlab.ceid.upatras.gr/~likothan/

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

Διδάσκοντες: Σ. Λυκοθανάσης, Α. Ανδρικόπουλος

Ιωάννης Βελγάκης

1047071

Ακ. Έτος 2019-20

ANTIKEIMENO:

Σε αυτή την εργαστηριακή άσκηση καλείστε να χρησιμοποιήσετε βασικά στοιχεία πιθανοθεωρίας, αλλά και να εξασκηθείτε στη θεωρία αποφάσεων του Bayes, λύνοντας απλά προβλήματα (κανόνες απόφασης, όρια απόφασης και λάθος απόφασης). Για κάποιους υπολογισμούς και το σχεδιασμό διαγραμμάτων, μπορείτε να χρησιμοποιήσετε το Matlab.

Ερώτημα 1.

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

Πίνακας 1: Δεδομένα από τις κλάσεις ωι και ω2

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)

1. Ποια είναι η εκ των προτέρων πιθανότητα για κάθε κλάση; (P(ω1) και P(ω2)).

Υπάρχουν N=12 2-d διανύσματα εκπαίδευσης, 6 σε κάθε κλάση, άρα οι εκ των προτέρων πιθανότητες είναι ισοπίθανες: $P(\omega_1)=P(\omega_2)=\frac{N_i}{N}=0.5$.

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

Οι δύο κλάσεις ακολουθούν Gaussian κατανομή, άρα τα μετρικά μέση τιμή (μ_i) και πίνακας συνδιασπορά (Σ_i) βρίσκονται από τους παρακάτω τύπους.

$$\boldsymbol{\mu}_i = \frac{1}{N} \sum_{k=1}^{N} \boldsymbol{x}_k \tag{1}$$

$$\mu_i - \frac{1}{N} \sum_{k=1}^{N} x_k
\Sigma_i = \frac{1}{N} \sum_{k=1}^{N} x_k x_k^t - \mu_i \mu_i^t$$
(2)

Χρησιμοποιώντας τον MATLAB κώδικα (Appendix I) παίρνουμε τις παρακάτω τιμές:

Πίνακας 1: Πίνακες μέσης τιμής και Συνδιασποράς για τις δύο κλάσεις

	μ_i	Σ_i
ω_1	$\binom{1.83}{1.50}$	$\begin{pmatrix} 2.17 & 1.10 \\ 1.10 & 1.10 \end{pmatrix}$
ω_2	$\binom{8.17}{9.33}$	$\begin{pmatrix} 1.37 & -0.07 \\ -0.07 & 1.07 \end{pmatrix}$

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 $p(\omega_i/x)$).

Για την εύρεση του ορίου απόφασης, θεωρώντας συνάρτηση κόστους μηδέν ένα, πρέπει να χρησιμοποιηθεί η διακρίνουσα συνάρτηση διχοτόμος:

$$g(x) = g_1(x) - g_2(x) = 0 : \omega_1 \text{ if } g(x) > 0 \text{ ; else } \omega_2$$

$$g(x) = p(\omega_1|x) - p(\omega_2|x) \xrightarrow{f(\cdot) \equiv ln} g(x) = ln \frac{p(x|\omega_1)}{p(x|\omega_2)} + ln \frac{P(\omega_1)}{P(\omega_2)} \tag{1}$$

Εφόσον τα δεδομένα εκπαίδευσης ακολουθούν 2-d Gaussian κατανομή, η εξίσωση πιθανοφάνειας που προκύπτει για κάθε κλάση είναι:

$$p(x|\omega_i) = \frac{1}{(2\pi)^{2/2} |\Sigma_i|^{1/2}} exp\left[-\frac{1}{2} (x - \mu_i)^t \Sigma_i^{-1} (x - \mu_i)\right]$$
(2)

Για τον πίνακα συνδιασποράς έγουμε

$$\Sigma_i = \begin{bmatrix} \sigma_{11i} & \sigma_{12i} \\ \sigma_{21i} & \sigma_{22i} \end{bmatrix} = \begin{bmatrix} \sigma_{1i}^2 & \rho \sigma_{1i} \sigma_{2i} \\ \rho \sigma_{1i} \sigma_{2i} & \sigma_{2i}^2 \end{bmatrix}, \text{\'o}\pi o v \ \rho_i = \frac{\sigma_{12i}}{\sigma_{1i} \sigma_{2i}}$$

και ρ: συντελεστής συσχέτισης

Για την ορίζουσα του πίνακα συνδιασποράς:

$$|\Sigma_i| = \sigma_{1i}^2 \sigma_{2i}^2 (1 - \rho_i^2)$$

Ο αντίστροφος πίνακας συνδιασποράς:

$$\Sigma_{i}^{-1} = \frac{1}{\sigma_{1i}^{2}\sigma_{2i}^{2}(1-\rho_{i}^{2})} \begin{bmatrix} \sigma_{2i}^{2} & -\rho_{i}\sigma_{1i}\sigma_{2i} \\ -\rho_{i}\sigma_{1i}\sigma_{2i} & \sigma_{1i}^{2} \end{bmatrix} = \frac{1}{1-\rho_{i}^{2}} \begin{bmatrix} 1/\sigma_{1i}^{2} & -\rho_{i}/\sigma_{1i}\sigma_{2i} \\ -\rho_{i}/\sigma_{1i}\sigma_{2i} & 1/\sigma_{2i}^{2} \end{bmatrix}$$

Και η τετραγωνική απόσταση Mahalanobis, θεωρώντας το διάνυσμα: **x** =

$$(x - \boldsymbol{\mu})^t \Sigma^{-1} (x - \boldsymbol{\mu}) = \begin{bmatrix} x - \mu_{ix} & y - \mu_{iy} \end{bmatrix} \frac{1}{1 - \rho_i^2} \begin{bmatrix} 1/\sigma_{1i}^2 & -\rho_i/\sigma_{1i}\sigma_{2i} \\ -\rho_i/\sigma_{1i}\sigma_{2i} & 1/\sigma_{2i}^2 \end{bmatrix} \begin{bmatrix} x - \mu_{ix} \\ y - \mu_{iy} \end{bmatrix}$$

Χρησιμοποιώντας την παραπάνω ανάλυση η εξίσωση πιθανοφάνειας (2) γίνεται:

$$p(x|\omega_{i}) = \frac{1}{2\pi\sigma_{1i}\sigma_{2i}\sqrt{1-\rho_{i}^{2}}} exp\left[-\frac{1}{2(1-\rho_{i}^{2})}\left[\left(\frac{x-\mu_{ix}}{\sigma_{1i}}\right)^{2} - 2\rho_{i}\left(\frac{x-\mu_{ix}}{\sigma_{1i}}\right)\left(\frac{y-\mu_{iy}}{\sigma_{2i}}\right) + \left(\frac{y-\mu_{iy}}{\sigma_{2i}}\right)^{2}\right]\right]$$

Η g(x) μέσω την (2) γράφεται:

$$g(x) μέσω την (2) γράφεται:$$

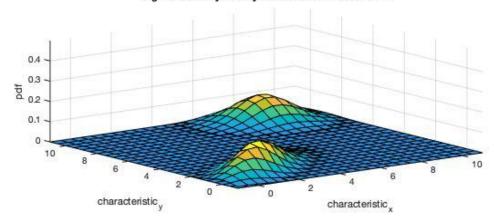
$$g(x) = -\frac{1}{2(1-\rho_1^2)} (*για i = 1) + \frac{1}{2(1-\rho_2^2)} (*για i = 2) + \ln \frac{\sigma_{1\omega_2}\sigma_{2\omega_2}\sqrt{1-\rho_2^2}}{\sigma_{1\omega_1}\sigma_{2\omega_1}\sqrt{1-\rho_1^2}}$$

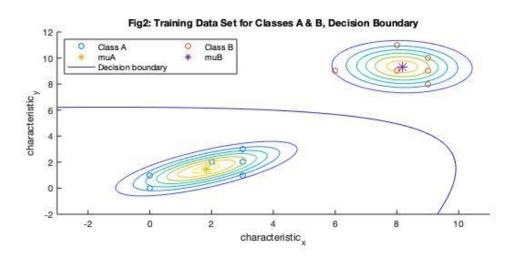
Αντικαθιστώντας από τον Πίνακα 1:

$$g(x) = 0.17x^2 - 1.90xy + 0.95y^2 + 12.26x + 15.98y - 136.05$$

Λύνοντας την g(x) = 0 βρίσκω την εξίσωση του ορίου απόφασης. Παρατηρώ ότι η g(x) είναι εξίσωση κωνικής τομής $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$ εφόσον $B^2 - 4AC = 2.96 > 0$ υπερβολής συγκεκριμένα Συμπερασματικά, οτιδήποτε κάτω από το όριο απόφασης ταξινομείται στην κλάση 1 και αντίστοιχα οτιδήποτε πάνω από το όριο στην κλάση 2, όπως φαίνεται στο παρακάτω σχήμα (σχήμα 1).







Σχήμα 1: Fig1 3d γράφημα των πυκνοτήτων πιθανότητας για τις 2 κλάσεις. Fig2 διάγραμμα των δεδομένων εκπαίδευσης του ταξινομητή για τις δύο κλάσεις (input data), τις μέσες τιμές για την κάθε μία, το όριο απόφασης και το 2d γράφημα των πυκνοτήτων πιθανότητας για τι 2 κλάσεις.

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;

Εφόσον, υπάρχουν κόστη λάθους ταξινόμησης, το όριο απόφασης βρίσκεται από τον έλεγχο λόγου πιθανοφάνειας (likelihood ratio test):

$$l_{12} \equiv \frac{p(x|\omega_1)}{p(x|\omega_2)} = \frac{P(\omega_2)(\lambda_{21} - \lambda_{22})}{P(\omega_1)(\lambda_{12} - \lambda_{11})}$$

όπου η απόφαση καθορίζεται: $ω_1$ if $l_{12} > 0$; else $ω_2$ και προφανώς επηρεάζεται λόγω της ύπαρξης των βαρών (κόστη) $λ_{ij}$. Ο λόγος πιθανοφάνειας l_{12} θα καθορίσει ένα (ή περισσότερα ∞ διάστασης) κατώφλι $θ_a$ το όποιο θα διαμερίσει το επίπεδο στους χώρους απόφασης R_i κάθε κλάσης. Υποθέτοντας ότι $λ_{ij}$, $i=j << λ_{ij}$, $i\neq j$ μπορούμε να διακρίνουμε τις δύο ακόλουθες περιπτώσεις:

Πρώτον, ότι η συνάρτηση κόστους τιμωρεί περισσότερο τη λάθος κατηγοριοποίηση στην ωι δειγμάτων της ω2 συγκριτικά με το αντίστροφο (λάθος κατηγοριοποίηση δείγματος στην ω2 ενώ ανήκει στην ω1), το κατώφλι θα μεγαλώσει αυξάνοντας την περιοχή απόφασης της ω1 και ταυτόχρονα μειώνοντας της ω2. Ομοίως το αντίστροφο θα μικρύνει το κατώφλι μειώνοντας την περιοχή απόφασης της ω1 και αυξάνοντας της ω2.

- 5. Θεωρήστε δύο κλάσεις ωι και ω2 στο χώρο προτύπων Ω με συνεχή κατανομή πιθανότητας p1(x) και p2(x) αντίστοιχα. Το πρόβλημα ταξινόμησης δύο κατηγοριών μπορεί να διατυπωθεί σαν διαίρεση του χώρου Ω σε δύο εξαντλητικά και μη επικαλυπτόμενα σύνολα Ω1 και Ω2, έτσι ώστε Ω1∪Ω2 = Ω και Ω1∩ Ω2 = 0. Αν το xi ∈ Ωk τότε αντιστοίχισε το xi στην κλάση ωk.
 - 5.1 θεωρείστε ότι δίνεται μια διακρίνουσα συνάρτηση f(.). Να αναφέρετε τα δύο λάθη που μπορεί να κάνει αυτή η συνάρτηση.

Να ταξινομήσει δείγμα της κλάσης 2 στην κλάση 1 και το αντίστροφο.

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

Σε κάθε περίπτωση η πιθανότητα λάθους είναι:

$$P(\lambda \dot{\alpha} \theta o \varsigma) = \begin{cases} P(\omega_1 | x), \varepsilon \alpha v \ \alpha \pi o \varphi \alpha \sigma i \zeta o v \mu \varepsilon \ \omega_2 \\ P(\omega_2 | x), \varepsilon \alpha v \ \alpha \pi o \varphi \alpha \sigma i \zeta o v \mu \varepsilon \ \omega_1 \end{cases}$$

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

$$R(a_i|x) = \sum_{j=1}^{2} \lambda(\alpha_i|\omega_j) P(\omega_j|x) = (c1 + c2) P(\omega_j|x)$$

6. Υποθέστε ότι έχουμε ένα πρόβλημα ταξινόμησης δύο κατηγοριών, σολομός (ω1) και πέρκα (ω2).

6.1Πρώτα, υποθέτουμε ότι έχουμε ένα χαρακτηριστικό, και οι σ.π.π. είναι Gaussians N(0, σ2) και N(1, σ2) 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 λ11=λ22=0.

Για την κλάση ωι έχουμε:
$$N(0,\sigma_2)$$
 άρα $pdf_1 = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right)$ (3)

Για την κλάση ω2 έχουμε:
$$N(1,\sigma_2)$$
 άρα $pdf_1 = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{1}{2}\left(\frac{x-1}{\sigma}\right)^2\right)$ (4)

Χρησιμοποιώντας το κριτήριο το Bayes, την υπόθεση $\lambda_{11}=\lambda_{22}=0$ και τις εξισώσεις (3) & (4) προκύπτει:

$$\frac{p(x|\omega_1)}{p(x|\omega_2)} > \frac{\lambda_{12}P(\omega_2)}{\lambda_{21}P(\omega_1)} \Rightarrow \frac{\frac{1}{\sqrt{2\pi}\sigma^2}\exp\left(-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right)}{\frac{1}{\sqrt{2\pi}\sigma^2}\exp\left(-\frac{1}{2}\left(\frac{x-1}{\sigma}\right)^2\right)} > \frac{\lambda_{12}P(\omega_2)}{\lambda_{21}P(\omega_1)}$$

$$\Leftrightarrow \exp\left(-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2 + \frac{1}{2}\left(\frac{x-1}{\sigma}\right)^2\right) > \frac{\lambda_{12}P(\omega_2)}{\lambda_{21}P(\omega_1)} \Leftrightarrow \frac{-2x+1}{\sigma^2} > \ln\frac{\lambda_{12}P(\omega_2)}{\lambda_{21}P(\omega_1)}$$

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

το κατώφλι που ελαχιστοποιεί το ρίσκο!

- **6.2** Μετά, υποθέτουμε ότι έχουμε δύο χαρακτηριστικά $\mathbf{x} = (x_1, x_2)$ και οι υπό συνθήκη πυκνότητες δύο κλάσεων $p(\chi/\omega=1)$ και $p(\chi/\omega=2)$, είναι 2- Δ gaussians κατανομές με κέντρα στα σημεία (4, 11) και (10, 3) αντίστοιχα με τον ίδιο πίνακα συνδιασποράς $\Sigma=3I$ (όπου I είναι ο μοναδιαίος πίνακας). Υποθέστε ότι οι a priori πιθανότητες είναι $P(\omega=1)=0.6$ και $P(\omega=2)=0.4$.
- (α) Υποθέστε ότι χρησιμοποιούμε τον Κανόνα Απόφασης του Bayes. Να γράψετε τις διακρίνουσες συναρτήσεις $g_1(x)$ και $g_2(x)$.

$$\Sigma = 3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \Sigma^{-1}, \mu_1 = \begin{bmatrix} 4 \\ 11 \end{bmatrix}, \mu_2 = \begin{bmatrix} 10 \\ 3 \end{bmatrix}, |\Sigma| = 3$$

Χρησιμοποιώντας την εξίσωση (2) (βλέπε ερώτημα 3) και έστω $x = \begin{bmatrix} x & y \end{bmatrix}$

$$p(x|\omega_i) = \frac{1}{(2\pi)^{2/2} |\Sigma_i|^{1/2}} exp\left[-\frac{1}{2}(x - \mu_i)^t \Sigma_i^{-1}(x - \mu_i)\right]$$

$$\Rightarrow p(x|\omega_i) = \frac{1}{2\pi\sqrt{3}}exp\left(-\frac{1}{2}\begin{bmatrix}x-\boldsymbol{\mu}_i\\y-\boldsymbol{\mu}_i\end{bmatrix}^t3\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}x-\boldsymbol{\mu}_i\\y-\boldsymbol{\mu}_i\end{bmatrix}\right)$$

Για τις διακρίνουσες συναρτήσεις:

$$g_{i}(x) = lnp(x|\omega_{i}) + lnP(\omega_{i})$$

$$g_{1}(x) = lnp(x|\omega_{1}) + lnP(\omega_{1})$$

$$= ln\frac{1}{2\pi\sqrt{3}} - \frac{3}{2}[x - 4 \quad y - 11]\begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix}\begin{bmatrix} x - 4\\ y - 11 \end{bmatrix} + ln0.6 \Leftrightarrow$$

$$g_{1}(x) = ln\frac{1}{2\pi\sqrt{3}} - \frac{3}{2}((x - 4)^{2} + (y - 11)^{2}) + ln0.6$$

Αντίστοιχα η $g_2(x)$:

$$g_2(x) = ln \frac{1}{2\pi\sqrt{3}} - \frac{3}{2}((x-10)^2 + (y-3)^2) + ln0.4$$

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

Χρησιμοποιώντας τη διακρίνουσα συνάρτηση διχοτόμου:

$$g(x) = g_1(x) - g_2(x) : \omega_1 \text{ if } g(x) > 0 \text{ ; else } \omega_2$$

$$g(x) = -\frac{3}{2}((x-4)^2 - (x-10)^2 - (y-3)^2 + (y-11)^2) - 1.43$$

$$g(x) = 0 \Rightarrow -12x + 16y - 27.73 = 0$$

$$\Leftrightarrow g(x) = 3x - 4y + 6.93 = 0$$

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

Η διακρίνουσα συνάρτηση διχοτόμος για όλα τα προβλήματα ταξινόμησης που αφορούν 2 κλάσεις με δισδιάστατο διάνυσμα χαρακτηριστικών και που ακολουθούν Gaussian κατανομή είναι η ακόλουθη(βλέπε ερωτ. 3):

$$\begin{split} g(x) &= -\frac{1}{2(1-\rho_1^2)} \left[\left(\frac{x - \mu_{x\omega_1}}{\sigma_{1\omega_1}} \right)^2 - 2\rho_i \left(\frac{x - \mu_{x\omega_1}}{\sigma_{1\omega_1}} \right) \left(\frac{y - \mu_{y\omega_1}}{\sigma_{2\omega_1}} \right) + \left(\frac{y - \mu_{y\omega_1}}{\sigma_{2\omega_1}} \right)^2 \right] \\ &+ \frac{1}{2(1-\rho_2^2)} \left[\left(\frac{x - \mu_{x\omega_2}}{\sigma_{1\omega_2}} \right)^2 - 2\rho_i \left(\frac{x - \mu_{x\omega_2}}{\sigma_{1\omega_2}} \right) \left(\frac{y - \mu_{y\omega_2}}{\sigma_{2\omega_2}} \right) + \left(\frac{y - \mu_{y\omega_2}}{\sigma_{2\omega_2}} \right)^2 \right] \\ &+ \ln \frac{\sigma_{1\omega_2} \sigma_{2\omega_2} \sqrt{1 - \rho_2^2}}{\sigma_{1\omega_1} \sigma_{2\omega_1} \sqrt{1 - \rho_1^2}} + \ln \frac{P(\omega_1)}{P(\omega_2)} \end{split}$$

Το πολυώνυμο αυτό θα είναι της μορφής: $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$, και είναι κωνική τομή. Ο γεωμετρικός τόπος καθορίζεται από την έκφραση $B^2 - 4AC$ και ανήκει σε μία από τις παρακάτω περιπτώσεις.

Περίπτωση	$B^2 - 4AC$	γ.τ.
A	<0	Έλλειψη
В	<0	Κύκλος
Γ	<0	Σημείο
Δ	=0	Παραβολή
Е	=0	2 Παράλληλες γραμμές
Z	=0	1 γραμμή
Н	>0	Υπερβολή
Θ	>0	2 τεμνόμενες ευθείες

Για τις καμπύλες (όχι ευθείες γραμμές) ο πίνακας συνδιασποράς μιας και συμβάλει άμεσα στους όρους Α-Ε θα επηρεάσει τα σχήμα τους, δηλ. την εκκεντρότητα, κατεύθυνση, πόσο γρήγορα συγκλίνει και μέγεθος αξόνων, οι εκ των προτέρων πιθανότητες και έμμεσα και ο πίνακας συνδιασποράς (συντελεστής συσχέτισης) θα επηρεάσουν την κατακόρυφη μετατόπιση της καμπύλης στον άξονα γγ'. Για τις γραμμές αντίστοιχα ο πίνακας συνδιασποράς θα επηρεάσει την κλίση των γραμμών και πάλι οι εκ των προτέρων πιθανότητες και έμμεσα και ο πίνακας συνδιασποράς (συντελεστής συσχέτισης) θα επηρεάσουν την κατακόρυφη μετατόπιση.

(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.

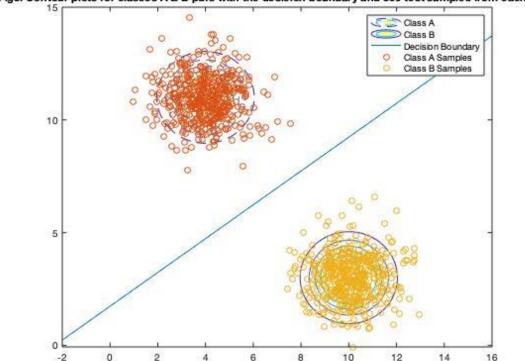


Fig3: Contour plots for classes A & B pdfs with the desicion boundary and 500 test samples from each cla

Βιβλιογραφία

- 1. "Documentation." *MATLAB Documentation*, https://www.mathworks.com/help/.
- 2. Duda, Richard O., et al. Pattern Recognition. Second ed.
- 3. Αυκοθανάσης, Σπυρίδων. Θεωρία Αποφάσεων. Πανεπιστήμιο Πατρών, 2015.

Appendix I

https://github.com/johnvelgakis/Decision Theory ex1