Θεόδωρος Αλεξόπουλος, Αναπληρωτής Καθηγητής
ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ - ΤΟΜΕΑΣ ΦΥΣΙΚΗΣ
ΗΡΩΩΝ ΠΟΛΥΤΕΧΝΕΙΟΥ 9
ΑΘΗΝΑ 157 80
Τηλ: 210 772-3019, Fax: 210 772-3025
e-mail: theoalex@central.ntua.gr

Theodoros Alexopoulos, Associate Professor
NATIONAL TECHNICAL UNIVERSITY
DEPARTMENT OF PHYSICS
ZOGRAFOU CAMPUS
157 80 ATHENS - GREECE
Phone : +30 210 772-3019, Fax: +30 210 772-3025
e-mail: Theodoros.Alexopoulos@cern.ch
html://www.physics.ntua.gr/Faculty/theoalex

Αναγνώριση Προτύπων & Νευρωνικά Δίκτυα

Προβλήματα 1

(Κεφάλαιο: Στατιστική Θεωρία του Bayes)

(Επιστροφή: 24 Νοεμβρίου 2003)

0.1

Ο ακόλουθος πίνακας του σχήματος (1) μας δίνει τις υπό συνθήκη πιθανότητες μιας τυχαίας μεταβλητής X για τρεις κλάσεις ω_1, ω_2, και ω_3. Έστω ότι γνωρίζουμε τις a priori πιθανότητες $P(\omega_1) = 0,3$ και $P(\omega_2) = 0,3$.
Υπολογίστε την πιθανότητα λάθους της ταξινόμησης χρησιμοποιώντας τον κανόνα απόφασης Bayes.
Παρατηρείστε ότι $\sum_x p(x/\omega_i) = 1,0$ και η τυχαία μεταβλητή X παίρνει τιμές στο διάστημα $[1,6]$.

Λύση:

Για τις a priori πιθανότητες έχουμε:

$$P(\omega_1) = P(\omega_2) = 0,3$$

$$\Rightarrow P(\omega_3) = 1 - P(\omega_1) - P(\omega_2) = 0,4.$$

Οι τιμές $P(\omega_i)p(x/\omega_i)$, $\forall i = 1,2,3$ που θα χρησιμοποιήσουμε για την ταξινόμηση των κλάσεων φαίνονται στον πίνακα 1.
Παρατηρούμε ότι $\sum_x P(\omega_i)p(x/\omega_i) = P(\omega_i)$. Ο ταξινομητής Bayes θα αποφασίσει για την κλάση ω_i όταν:

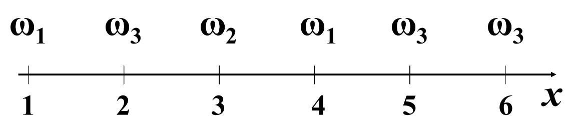
$$P(\omega_i)p(x/\omega_i) > P(\omega_j)p(x/\omega_j) \quad \forall\, j \neq i.$$

$p(X=x/\omega_i)$						
$\omega_i \backslash x$	X=1	X=2	X=3	X=4	X=5	X=6
ω_1	0,3	0,2	0,1	0,1	0,2	0,1
ω_2	0,2	0,2	0,4	0,05	0,1	0,05
ω_3	0,1	0,3	0,15	0,05	0,3	0,1

Σχήμα 1: Παρατηρείστε ότι $\sum_x p(x/\omega_i) = 1,0$ και η τυχαία μεταβλητή X παίρνει τιμές στο διάστημα $[1, 6]$.

$P(\omega_i)p(x/\omega_i)$	$x=1$	$x=2$	$x=3$	$x=4$	$x=5$	$x=6$
ω_1	$0,09$ *	0,06	0,03	$0,03$ *	0,06	0,03
ω_2	0,06	0,06	$0,12$ *	0,015	0,03	0,015
ω_3	0,04	$0,12$ *	0,06	0,02	$0,12$ *	$0,040$ *

Πίνακας 1:



Σχήμα 2:

Επομένως από τον πίνακα, η απόφαση στο χώρο των χαρακτηριστικών σημείων είναι αυτή που φαίνεται στο σχήμα (2). Η πιθανότητα λάθους ταξινόμησης είναι:

$$P_e = 1 - P_c = 1 - [0,09 + 0,12 + 0,12 + 0,03 + 0,12 + 0,04] = 0,48,$$

όπου P_c είναι η πιθανότητα της ορθής ταξινόμησης.

0.2

Να υπολογίσετε την πιθανότητα του λάθους ταξινόμησης κατά Bayes για δυο κλάσεις ω_1, ω_2, όπου θεωρούμε δείγματα σε 2-διαστάσεις από γκαουσιανές κατανομές που περιγράφονται από τις ακόλουθες πυκνότητες πιθανότητας:

$$p(\vec{x}/\omega_1) \sim N(\vec{\mu}_1, \overline{\overline{\Sigma}}_1),$$

με

$$\vec{\mu}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \overline{\overline{\Sigma}}_1 = \begin{pmatrix} 4 & 1 \\ 1 & 9 \end{pmatrix},$$

και

$$p(\vec{x}/\omega_2) \sim N(\vec{\mu}_2, \overline{\overline{\Sigma}}_2),$$

με

$$\vec{\mu}_1 = \begin{pmatrix} -2 \\ -1 \end{pmatrix}, \quad \overline{\overline{\Sigma}}_2 = \begin{pmatrix} 4 & 1 \\ 1 & 9 \end{pmatrix}.$$

Θεωρήστε ότι η a priori πιθανότητα $P(\omega_1) = 0,25$.

Λύση:

Παρατηρούμε ότι για τους πίνακες διασποράς $\overline{\overline{\Sigma}}_1 = \overline{\overline{\Sigma}}_2$ έχουμε: $\overline{\overline{\Sigma}}_1 = \overline{\overline{\Sigma}}_2 = \overline{\overline{\Sigma}}$. Ο αντίστροφος πίνακας διασποράς είναι:

$$\overline{\overline{\Sigma}}^{-1} = \frac{1}{35} \begin{pmatrix} 9 & -1 \\ -1 & 4 \end{pmatrix}.$$

Οι συναρτήσεις διάκρισης, $g_1(\overrightarrow{x}),\ g_2(\overrightarrow{x})$, είναι:

$$g_i(\overrightarrow{x}) = \left(\overline{\overline{\Sigma}}_i^{-1} \overrightarrow{\mu}_i\right)^t \overrightarrow{x} - \frac{1}{2} \overrightarrow{\mu}_i^t \overline{\overline{\Sigma}}_i^{-1} \overrightarrow{\mu}_i + \ln P(\omega_i),$$

με

$$P(\omega_2) = 0,25, \qquad P(\omega_2) = 0,75,$$

και η συνάρτηση απόφασης είναι:

$$d(\overrightarrow{x}) = g_1(\overrightarrow{x}) - g_2(\overrightarrow{x})$$

$$\Rightarrow d(\overrightarrow{x}) = \left(\overline{\overline{\Sigma}}^{-1}\overrightarrow{\mu}_1\right)^t \overrightarrow{x} - \frac{1}{2}\overrightarrow{\mu}_1^t\overline{\overline{\Sigma}}^{-1}\overrightarrow{\mu}_1 -$$

$$-\left(\overline{\overline{\Sigma}}^{-1}\overrightarrow{\mu}_2^t\right)\overrightarrow{x} + \frac{1}{2}\overrightarrow{\mu}_2^t\overrightarrow{\Sigma}\,\overrightarrow{\mu}_2 + \ln\left(\frac{P(\omega_1)}{P(\omega_2)}\right)$$

$$\Rightarrow d(\overrightarrow{x}) = 24x_1 + 9x_2 - 31,$$

όπου $\overrightarrow{x} = (\ x_1 \ \ x_2\ )^t$.
Η επιφάνεια απόφασης είναι:

$$d(\overrightarrow{x}) = 0$$

$$\Rightarrow 24x_1 + 9x_2 - 31 = 0$$

$$\Rightarrow \overrightarrow{w}^t\overrightarrow{x} + w_0 = 0, \qquad \text{όπου}$$

$$\overrightarrow{w} = \begin{pmatrix} 24/35 \\ 9/35 \end{pmatrix}, \qquad w_0 = -31/35 = -0,885.$$

Ορίζουμε μια νέα μεταβλητή $y = \overrightarrow{w}^t\overrightarrow{x}$. Τότε οι υπό συνθήκη πιθανότητες είναι:

$$p(\overrightarrow{x}/\omega_i) \sim N\left(\overrightarrow{\mu}_i, \overline{\overline{\Sigma}}_i\right)$$

$$\Rightarrow p(\overrightarrow{w}^t\overrightarrow{x}/\omega_i) \sim N\left(\overrightarrow{w}^t\overrightarrow{\mu}_i, \overrightarrow{w}^t\overline{\overline{\Sigma}}_i\overrightarrow{w}^t\right),$$

ή

$$p(\overrightarrow{w}^t\overrightarrow{x}/\omega_1) \sim N\left(1,20, \ \ 2,829\right),$$

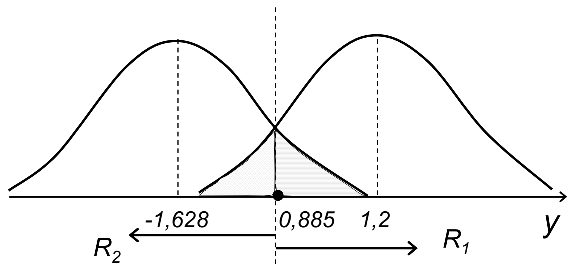
$$p(\overrightarrow{w}^t\overrightarrow{x}/\omega_2) \sim N\left(-1,63, \ \ 2,829\right).$$

Δηλαδή το πρόβλημά μας έγινε πρόβλημα μίας διάστασης, όπως φαίνεται στο σχήμα (3). Η πιθανότητα ορθής ταξινόμησης είναι:

$$P_c = P(\omega_1)\int_{R_1} p(\overrightarrow{w}^t\overrightarrow{x}/\omega_1)d(\overrightarrow{w}^t\overrightarrow{x}) + P(\omega_2)\int_{R_2} p(\overrightarrow{w}^t\overrightarrow{x}/\omega_2)d(\overrightarrow{\omega}^t\overrightarrow{x})$$

$$\Rightarrow P_c = \frac{1}{4}\int_{0,885}^{\infty} N\left(1,2, \ \sqrt{2,829}\right)dy + \frac{3}{4}\int_{-\infty}^{0,885} N\left(-1,628, \ \sqrt{2,829}\right)dy$$

$$\Rightarrow P_c = \frac{1}{4} - \frac{1}{4}\Phi\left(\frac{0,885-1,2}{\sqrt{2,83}}\right) + \frac{3}{4}\Phi\left(\frac{0,885+1,628}{\sqrt{2,83}}\right)$$



Σχήμα 3:

$$\Rightarrow P_c = \frac{1}{4} - \frac{1}{4}\Phi(-0,187) + \frac{3}{4}\Phi(1,494)$$

$$\Rightarrow P_c = \frac{1}{4} - \frac{1}{4} \times (0,425) + \frac{3}{4} \times (0,933)$$

$$\Rightarrow P_c = 0,84.$$

0.3

Θεωρείστε μια μέτρηση σε τρεις διαστάσεις $\vec{x} = (x_1\ x_2\ x_3)^t$. Έστω ότι έχουμε 4 δείγματα από την κλάση ω_1 και 4 δείγματα από την κλάση ω_2:

$$\omega_1 : \left\{(1\ 0\ 1)^t,\ (0\ 0\ 0)^t,\ (1\ 0\ 0)^t,\ (1\ 1\ 0)^t\right\},$$

$$\omega_2 : \left\{(0\ 0\ 1)^t,\ (0\ 1\ 1)^t,\ (1\ 1\ 1)^t,\ (0\ 1\ 0)^t\right\}.$$

Να υποθέσετε ότι η μεταβλητή $\vec{x}$ είναι γκαουσιανή, όπου μπορείτε να χρησιμοποιήσετε τις ακόλουθες σχέσεις για τον προσδιορισμό της μέσης τιμής και του πίνακα διασποράς:

$$\vec{\mu} = \frac{1}{N}\sum_{k=1}^{N} \vec{x}_k,$$

$$\overline{\overline{\Sigma}} = \frac{1}{N}\sum_{k=1}^{N} \vec{x}_k\vec{x}_k^t - \vec{\mu}\vec{\mu}^t,$$

όπου N είναι το πλήθος των δειγμάτων.

Να χρησιμοποιήσετε τον κανόνα απόφασης Bayes για τον προσδιορισμό της επιφάνειας απόφασης. Προσδιορίστε την ολική πιθανότητα για την ορθή ταξινόμηση του κανόνα απόφασης Bayes.

Λύση:

Οι μέσες τιμές είναι:

$$\overrightarrow{\mu}_1 = \frac{1}{4}\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \qquad \overrightarrow{\mu}_2 = \begin{pmatrix} 1 \\ 3 \\ 3 \end{pmatrix},$$

και οι πίνακες διασποράς:

$$\overline{\overline{\Sigma}}_1 = \overline{\overline{\Sigma}}_2 = \frac{1}{16}\begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & -1 & 3 \end{pmatrix}.$$

Εφόσον $\overline{\overline{\Sigma}}_1 = \overline{\overline{\Sigma}}_2$, ο κανόνας του Bayes θα μας δώσει ένα ταξινομητή ελάχιστης απόστασης Mahalanobis με συναρτήσεις διάκρισης:

$$g_i(\overrightarrow{x}) = \left(\overline{\overline{\Sigma}}^{-1}\overrightarrow{\mu}_i\right)^t \overrightarrow{x} - \frac{1}{2}\overrightarrow{\mu}_i^t\overline{\overline{\Sigma}}^{-1}\overrightarrow{\mu}_i,$$

όπου

$$\overline{\overline{\Sigma}}^{-1} = \begin{pmatrix} 8 & -4 & -4 \\ -4 & 8 & 4 \\ -4 & 4 & 8 \end{pmatrix} \qquad \text{και} \qquad \overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Επομένως:

$$g_1(\overrightarrow{x}) = 4x_1 - \frac{3}{2}, \quad g_2(\overrightarrow{x}) = -4x_1 + 8x_2 + 8x_3 - \frac{1}{2}.$$

Η επιφάνεια απόφασης είναι:

$$d(\overrightarrow{x}) = g_1(\overrightarrow{x}) - g_2(\overrightarrow{x}) = 0$$

$$\Rightarrow 8x_1 - 8x_2 - 8x_3 + 4 = 0$$

$$\Rightarrow 2\begin{pmatrix} 1 & -1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + 1 = 0$$

$$\Rightarrow \overrightarrow{w}^t\overrightarrow{x} + w_0 = 0,$$

όπου

$$\overrightarrow{w} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \qquad \text{και} \qquad w_0 = 1/2.$$

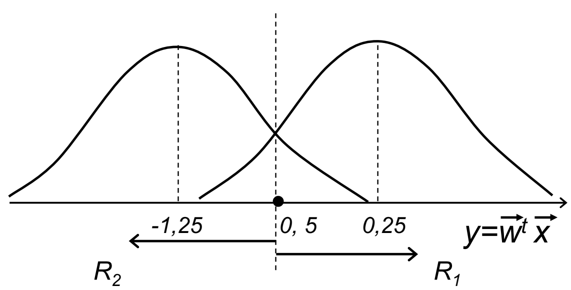
Εφόσον $\overline{\overline{\Sigma}}_1 = \overline{\overline{\Sigma}}_2$ και $p(\overrightarrow{x}/\omega_i) \sim N(\overrightarrow{\mu}_i, \overline{\overline{\Sigma}}_i)$, τότε:

$$p(\overrightarrow{w}^t\overrightarrow{x}/\omega_i) \sim N(\overrightarrow{w}^t\overrightarrow{\mu}_i, \overrightarrow{w}^t\overline{\overline{\Sigma}}_i\overrightarrow{w}).$$

Επομένως οι νέες μονοδιάστατες πυκνότητες πιθανότητας είναι:

$$p(\overrightarrow{w}^t\overrightarrow{x}/\omega_1) \sim N(0,25,\ 0,1875),$$

$$p(\overrightarrow{w}^t\overrightarrow{x}/\omega_2) \sim N(-1,25,\ 0,1875).$$



Σχήμα 4:

Η πιθανότητα του λάθους ταξινόμησης είναι:

$$P_e = \frac{1}{2}\left[\Phi\left(\frac{-0,5+1,25}{\sqrt{0,1875}}\right) + \Phi\left(\frac{-0,5-1,2}{\sqrt{0,1875}}\right)\right] = 0,041.$$

0.4

Έστω τα δείγματα:

$$(1\ 2)^t,\ (2\ 2)^t,\ (3\ 1)^t,\ (3\ 2)^t,\ (2\ 3)^t$$

ανήκουν σε μια κλάση ω_1. Επιπλέον, έστω τα δείγματα:

$$(7\ 9)^t,\ (8\ 9)^t,\ (9\ 8)^t,\ (9\ 9)^t,\ (8\ 10)^t$$

ανήκουν σε μια κλάση ω_2. Έστω ότι τα δείγματα των δυο κλάσεων προέρχονται από γκαουσιανές πυκνότητες πιθανότητας. Να βρείτε την επιφάνεια απόφασης εφαρμόζοντας τον κανόνα Bayes. Θεωρείστε ισοπίθανες κλάσεις.

Λύση:

Όι μέσες τιμές των δύο κατανομών είναι:

$$\vec{\mu}_1 = \begin{pmatrix} 11/5 \\ 2 \end{pmatrix}, \qquad \vec{\mu}_2 = \begin{pmatrix} 41/5 \\ 9 \end{pmatrix},$$

όπου χρησιμοποιήσαμε τη σχέση:

$$\vec{\mu}_i = \frac{1}{5}\sum_{k=1}^{5} \vec{x}_k \qquad \forall\ i = 1,2 \text{ (για τις δύο κλάσεις)},$$

Για τον προσδιορισμό των μέσων τιμών και για τους πίνακες διασποράς:

$$\overline{\overline{\Sigma}}_i = \frac{1}{5}\sum_{k=1}^{5} \vec{x}_k \vec{x}_k^t - \vec{\mu}_i - \vec{\mu}_i^t,$$

Επομένως θα έχουμε:

$$\overline{\overline{\Sigma}}_1 = \frac{1}{5}\left\{\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} + \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} + \begin{pmatrix} 9 & 3 \\ 3 & 1 \end{pmatrix} + \begin{pmatrix} 9 & 6 \\ 6 & 4 \end{pmatrix} + \right.$$

$$\left. + \begin{pmatrix} 4 & 6 \\ 6 & 9 \end{pmatrix}\right\} - \begin{pmatrix} (11/5)^2 & (11/5)\times 2 \\ (11/5)\times 2 & 2^2 \end{pmatrix}$$

$$\Rightarrow \overline{\overline{\Sigma}}_1 = \begin{pmatrix} 11/25 & -1/5 \\ -1/5 & 2/5 \end{pmatrix}$$

$$\Rightarrow \overline{\overline{\Sigma}}_1^{-1} = \begin{pmatrix} 2,17 & 1,09 \\ 1,09 & 3,04 \end{pmatrix}.$$

Δείξτε ότι $\overline{\overline{\Sigma}}_2 = \overline{\overline{\Sigma}}_1$. Οι συναρτήσεις διάκρισης $g_1(\vec{x}),\ g_2(\vec{x})$ και η συνάρτηση απόφασης είναι:

$$d(\vec{x}) = g_1(\vec{x}) - g_2(\vec{x})$$

$$\Rightarrow d(\vec{x}) = \left\{ (\ \overline{\overline{\Sigma}}_1^{-1}\ \ \vec{\mu}_1\ )^t \vec{x} - \frac{1}{2}\vec{\mu}_1^t \overline{\overline{\Sigma}}_1 \vec{\mu}_1 \right\} - \left\{ \left(\overline{\overline{\Sigma}}_2^{-1}\ \ \vec{\mu}_2 \right)^t \vec{x} - \frac{1}{2}\vec{\mu}_2 \overline{\overline{\Sigma}}_2 \vec{\mu}_2 \right\}$$

$$\Rightarrow d(\overrightarrow{x}) = (\overrightarrow{\mu}_1 - \overrightarrow{\mu}_2)^t \overline{\overline{\Sigma}}_1 \overrightarrow{x} - \frac{1}{2}\left(\overrightarrow{\mu}_1^t \overline{\overline{\Sigma}}_1^{-1} - \overrightarrow{\mu}_2 \overline{\overline{\Sigma}}_2^{-1} \overrightarrow{\mu}_2\right),$$

όπου $\overrightarrow{x} = (\ x_1 \ \ x_2\ )^t$. Επομένως η επιφάνεια απόφασης είναι:

$$d(\overrightarrow{x}) = x_1 + 1,35x_2 - 12,61 = 0$$

$$\Rightarrow x_1 + 1,35x_2 - 12,61 = 0.$$

0.5

(α) Η επιφάνεια απόφασης για δυο κλάσεις ω_1 και ω_2 που είναι κατανεμημένες κατά Gauss με πίνακες διασποράς:

$$\overline{\overline{\Sigma}}_1 = \overline{\overline{\Sigma}}_2 \neq \sigma^2 \overline{\overline{I}},$$

και μέσες τιμές μ_1, μ_2, και a priori πιθανότητες:

$$P(\omega_1) \neq P(\omega_2),$$

είναι μια υπερεπιφάνεια που περιγράφεται από την ακόλουθη γραμμική εξίσωση:

$$\vec{A}^t(\vec{x} - A_0) = 0.$$

Να εκφράσετε τις παραμέτρους $\vec{A}$ και A_0 ως συνάρτηση των $\vec{\mu}_1$, $\vec{\mu}_2$, $\overline{\overline{\Sigma}}$, $P(\omega_1)$, και $P(\omega_2)$.
(β) Ο ίδιος κανόνας απόφασης Bayes για δυο κατηγορίες με πίνακες διασποράς:

$$\overline{\overline{\Sigma}}_1 \neq \overline{\overline{\Sigma}}_2, \quad P(\omega_1) \neq P(\omega_2)$$

θα δώσει μια επιφάνεια απόφασης που περιγράφεται από την εξίσωση:

$$\vec{x}^t \overline{\overline{B}} \vec{x} + \vec{A}^t \vec{x} + C = 0.$$

Να καθορίσετε τις παραμέτρους $\overline{\overline{B}}$, $\vec{A}$, και C.

Λύση:

(α)

Η συνάρτηση διάκρισης $g_i(\overrightarrow{x})$ είναι:

$$g_i(\overrightarrow{x}) = -\frac{1}{2}\overrightarrow{x}^t \overline{\overline{\Sigma}}^{-1} \overrightarrow{x} + \left(\overrightarrow{\mu}_i^t \overline{\overline{\Sigma}}^{-1}\right) \overrightarrow{x} -$$
$$-\frac{1}{2}\overrightarrow{\mu}_i^t \overline{\overline{\Sigma}}^{-1} \overrightarrow{\mu}_i + \ln P(\omega_i),$$

με επιφάνεια απόφασης

$$d(\overrightarrow{x}) = g_1(\overrightarrow{x}) - g_2(\overrightarrow{x}) = 0$$

$$\Rightarrow g_1(\overrightarrow{x}) = g(\overrightarrow{x})$$

$$\Rightarrow \left(\overrightarrow{\mu}_1^t - \overrightarrow{\mu}_2^t\right) \overline{\overline{\Sigma}}^{-1} \overrightarrow{x} -$$

$$-\frac{1}{2}\left(\overrightarrow{\mu}_1^t \overline{\overline{\Sigma}}^{-1} \overrightarrow{\mu}_1 - \overrightarrow{\mu}_2^t \overline{\overline{\Sigma}}^{1} \overrightarrow{\mu}_2\right) + \ln\frac{P(\omega_1)}{P(\omega_2)} = 0. \tag{1}$$

Έστω ότι:

$$C = \frac{1}{2}\left(\overrightarrow{\mu}_1^t \overline{\overline{\Sigma}}^{-1} \overrightarrow{\mu}_1 - \overrightarrow{\mu}_2^t \overline{\overline{\Sigma}}^{-1} \overrightarrow{\mu}_2\right) - \ln\frac{P(\omega_1)}{P(\omega_2)}.$$

Τότε, η σχέση (1) θα γίνει:

$$(\overrightarrow{\mu}_1 - \overrightarrow{\mu}_2)^t \overline{\overline{\Sigma}}^{-1} \left(\overrightarrow{x} - \frac{C(\overrightarrow{\mu}_1 + \overrightarrow{\mu}_2)}{(\overrightarrow{\mu}_1 - \overrightarrow{\mu}_2)^t \overline{\overline{\Sigma}}^{-1} (\overrightarrow{\mu}_1 + \overrightarrow{\mu}_2)}\right) = 0$$

$$\Rightarrow \overrightarrow{A}^t (\overrightarrow{x} - A_0) = 0,$$

όπου

$$\overrightarrow{A} = \overline{\overline{\Sigma}}^{-1} \left(\overrightarrow{\mu}_1 - \overrightarrow{\mu}_2\right),$$

$$A_0 = \frac{1/2\ \left(\overrightarrow{\mu}_1^t \overline{\overline{\Sigma}}^{-1} \overrightarrow{\mu}_1 - \overrightarrow{\mu}_2^t \overline{\overline{\Sigma}}^{-1} \overrightarrow{\mu}_2\right) - \ln\frac{P(\omega_1)}{P(\omega_2)}}{(\overrightarrow{\mu}_1 - \overrightarrow{\mu}_2)^t \overline{\overline{\Sigma}}^{-1} (\overrightarrow{\mu}_1 + \overrightarrow{\mu}_2)} (\overrightarrow{\mu}_1 + \overrightarrow{\mu}_2).$$

Παρατηρούμε ότι:

$$(\overrightarrow{\mu}_1 - \overrightarrow{\mu}_2)^t \overline{\overline{\Sigma}}^{-1} (\overrightarrow{\mu}_1 + \overrightarrow{\mu}_2) = \overrightarrow{\mu}_1^t \overline{\overline{\Sigma}}^{-1} \overrightarrow{\mu}_1 - \overrightarrow{\mu}_2^t \overline{\overline{\Sigma}}^{-1} \overrightarrow{\mu}_2.$$

Επομένως,

$$A_0 = \frac{1}{2}(\overrightarrow{\mu}_1 + \overrightarrow{\mu}_2) - \frac{\ln\left(P(\omega_1)/P(\omega_2)\right)}{(\overrightarrow{\mu}_1 - \overrightarrow{\mu}_2)^t \overline{\overline{\Sigma}}^{-1} (\overrightarrow{\mu}_1 + \overrightarrow{\mu}_2)} (\overrightarrow{\mu}_1 + \overrightarrow{\mu}_2).$$

(β)

Για $\overline{\overline{\Sigma}}_1 \neq \overline{\overline{\Sigma}}_2$ και $P(\omega_1) \neq P(\omega_2)$, η συνάρτηση διάκρισης είναι:

$$g_i(\overrightarrow{x}) = -\frac{1}{2}\overrightarrow{x}^t\overline{\overline{\Sigma}}_i^{-1}\overrightarrow{x} + \left(\overrightarrow{\mu}_i^t\overline{\overline{\Sigma}}_i^{-1}\right)\overrightarrow{x} -$$

$$-\frac{1}{2}\overrightarrow{\mu}_i^t\overline{\overline{\Sigma}}_i^{-1} - \frac{1}{2}\ln|\overline{\overline{\Sigma}}_i| + \ln P(\omega_i).$$

Η επιφάνεια απόφασης $d_{12}(\overrightarrow{x})$ είναι:

$$g_1(\overrightarrow{x}) = g_2(\overrightarrow{x})$$

$$\Rightarrow -\frac{1}{2}\overrightarrow{x}^t(\overline{\overline{\Sigma}}_1^{-1} - \overline{\overline{\Sigma}}_2^{-1})\overrightarrow{x} + (\overrightarrow{\mu}_1^t\overline{\overline{\Sigma}}_1^{-1} - \overrightarrow{\mu}_2^t\overline{\overline{\Sigma}}_2^{-1})\overrightarrow{x} -$$

$$-\frac{1}{2}\left(\overrightarrow{\mu}_1^t\overline{\overline{\Sigma}}_1^{-1}\overrightarrow{\mu}_1 - \overrightarrow{\mu}_2^t\overline{\overline{\Sigma}}_2^{-1}\overrightarrow{\mu}_2\right) +$$

$$+\frac{1}{2}\ln\frac{|\overline{\overline{\Sigma}}_1|}{|\overline{\overline{\Sigma}}_2|} + \ln\frac{P(\omega_1)}{P(\omega_2)} = 0$$

$$\Rightarrow \overrightarrow{x}^t\overline{\overline{B}}\overrightarrow{x} + \overrightarrow{A}^t\overrightarrow{x} + C = 0,$$

όπου προφανώς έχουμε ορίσει τις ακόλουθες παραμέτρους:

$$\overline{\overline{B}} = -\frac{1}{2}(\overline{\overline{\Sigma}}_1^{-1} - \overline{\overline{\Sigma}}_2^{-1}),$$

$$\overrightarrow{A} = \overline{\overline{\Sigma}}_1^{-1}\overrightarrow{\mu}_1 - \overline{\overline{\Sigma}}_2^{-1}\overrightarrow{\mu}_2,$$

και

$$C = -\frac{1}{2}\left(\overrightarrow{\mu}_1^t\overline{\overline{\Sigma}}_1^{-1}\overrightarrow{\mu}_1 - \overrightarrow{\mu}_2^t\overline{\overline{\Sigma}}_2^{-1}\overrightarrow{\mu}_2\right) +$$

$$+\frac{1}{2}\ln\frac{|\overline{\overline{\Sigma}}_1|}{|\overline{\overline{\Sigma}}_2|} + \ln\frac{P(\omega_1)}{P(\omega_2)}.$$

0.6

(α) Θεωρείστε ένα πρόβλημα αναγνώρισης προτύπων με M-κατηγορίες μιας διάστασης όπου κάθε κατηγορία χαρακτηρίζεται από μια κατανομή πυκνότητας πιθανότητας Rayleigh:

$$p(x/\omega_i) = \begin{cases} (x/\sigma_i^2)\, e^{-x^2/2\sigma_i^2} & \textbf{για} \quad x \geq 0 \\ 0 & \textbf{για} \quad x < 0. \end{cases}$$

Βρείτε τη συνάρτηση απόφασης Bayes του προβλήματος, υποθέτοντας ότι οι a priori πιθανότητες είναι $P(\omega_i) = 1/M$.
(β) Επαναλάβατε το ίδιο πρόβλημα στην περίπτωση για την κατανομή πυκνότητας πιθανότητας:

$$p(x/\omega_i) = \begin{cases} (x/T_i)\, e^{-x/T_i} & \textbf{για} \quad x \geq 0 \\ 0 & \textbf{για} \quad x < 0. \end{cases},$$

<u>Λύση:</u>

(α)

Η συνάρτηση διάκρισης είναι:

$$g_i(x) = p(x/\omega_i)P(\omega_i),$$

ή μπορούμε να χρησιμοποιήσουμε το φυσικό λογάριθμο,

$$g_i(x) = \ln p(x/\omega_i) + \ln P(\omega_i)$$

$$\Rightarrow g_i(x) = \ln x - 2\ln\sigma_i - \frac{x^2}{2\sigma_i^2} + \ln\frac{1}{M}$$

$$\Rightarrow g_i(x) = \ln x - 2\ln\sigma_i - \frac{x^2}{2\sigma_i^2} - \ln M.$$

Για δύο κλάσεις i, j ο ταξινομητής θα μας δώσει μια επιφάνεια απόφασης:

$$d_{ij}(x) = g_i(x) - g_j(x) = 0$$

$$\Rightarrow -2\ln\left(\frac{\sigma_i}{\sigma_j}\right) - \frac{x^2}{2}\left(\frac{1}{\sigma_i^2} - \frac{1}{\sigma_j^2}\right) = 0$$

$$\Rightarrow x = \sqrt{4\ln\left(\frac{\sigma_j}{\sigma_i}\right) \Bigg/ \left(\frac{1}{\sigma_i^2} - \frac{1}{\sigma_j^2}\right)}.$$

(β)

Η συνάρτηση διάκρισης είναι:

$$g_i(x) = \ln p(x/\omega_i) - \ln P(\omega_i),$$

όπου

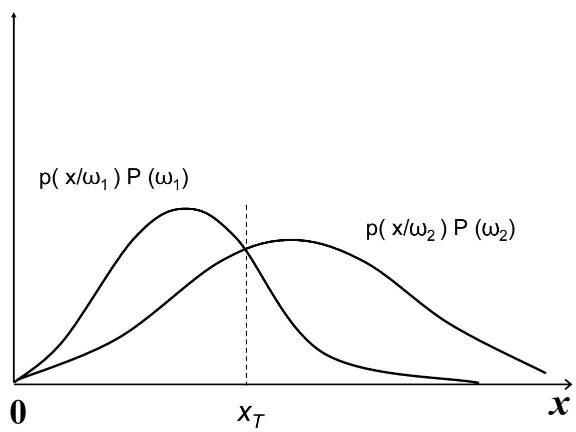
$$p(x/\omega_i) = \frac{x}{T_i^2} e^{-x/T_i}, \qquad P(\omega_i) = \frac{1}{2},$$

και η επιφάνεια απόφασης είναι:

$$d_{ij}(x) = g_i(x) - g_j(x)$$

$$\Rightarrow \left(-2\ln T_i - \frac{x}{T_i}\right) - \left(-2\ln T_j - \frac{x}{T_j}\right) = 0$$

$$\Rightarrow x_T = \frac{2\ln(T_j/T_i)}{\dfrac{1}{T_i} - \dfrac{1}{T_j}}.$$



Σχήμα 5:

Η πιθανότητα του λάθους ταξινόμησης είναι:

$$P_e = \frac{1}{2}\left[\int_0^{x_T} \frac{x}{T_j^2} e^{-x/T_j} dx + \int_{x_T}^{\infty} \frac{x}{T_j^2} e^{-x/T_i} dx\right]$$

$$\Rightarrow P_e = \frac{1}{2}\left[x_T\left(-\frac{e^{-x_T/T_j}}{T_j} + \frac{e^{-x_T/T_i}}{T_i}\right) + \left(-e^{-x_T/T_j} + e^{-x_T/T_i}\right)\right]$$

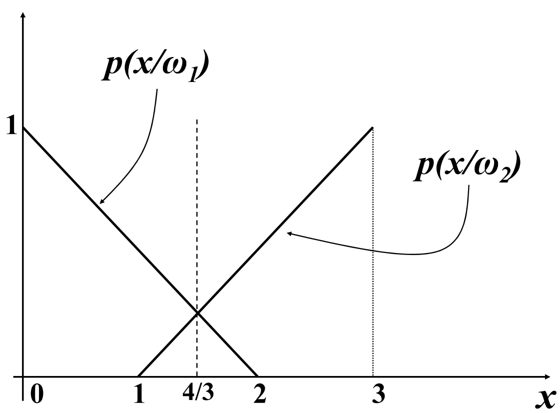
$$\Rightarrow P_e = \frac{1}{2}\left[e^{-x_T/T_i}\left(\frac{x_T}{T_i} + 1\right) - e^{-x_T/T_j}\left(\frac{x_T}{T_j} + 1\right)\right],$$

όπου x_T είναι το σημείο απόφασης:

$$x_T = \frac{2\ln(T_j/T_i)}{\dfrac{1}{T_i} - \dfrac{1}{T_j}}.$$

0.7

Έστω ότι δυο κλάσεις περιγράφονται από τις συναρτήσεις πυκνότητας πιθανότητας του σχήματος (6). Βρείτε το σύνορο απόφασης Bayes. Να υποθέσετε ότι έχουμε ισοπίθανες κλάσεις ω_1 και ω_2.



Σχήμα 6:

<u>Λύση:</u>

Από το σχήμα (6) έχουμε:

$$p(x/\omega_1) = -\frac{x}{2} + 1, \qquad p(x/\omega_2) = x - 1.$$

$$P(\omega_1) = P(\omega_2) = \frac{1}{2}.$$

Η συνάρτηση διάκρισης είναι:

$$g_i(x) = p(x/\omega_i)P(\omega_i),$$

και η επιφάνεια απόφασης είναι:

$$d_{12}(x) = g_1(x) - g_2(x) = 0$$

$$\Rightarrow \frac{1}{2}\left(-\frac{x}{2}+1\right) - (x-1) = 0$$

$$\Rightarrow x = \frac{4}{3}.$$

Η πιθανότητα του λάθους της ταξινόμησης είναι:

$$P_e = P(\omega_1)\int_1^{4/3}(x-1)dx + P(\omega_2)\int_{4/3}^{2} -\left(\frac{x}{2}+1\right)dx$$

$$\Rightarrow P_e = \frac{1}{2}\frac{1}{2}\left(1 \times \frac{1}{3}\right) = \frac{1}{12}.$$

0.8

Υποθέτουμε ότι οι ακόλουθες δυο κατηγορίες (ω_1, ω_2) περιγράφονται από γκαουσιανές συναρτήσεις πυκνότητας πιθανότητας:

$$\omega_1 : \left\{(0\ 0)^t,\ (2\ 0)^t,\ (2\ 2)^t,\ (0\ 2)^t\right\},$$

και

$$\omega_2 : \left\{(4\ 4)^t,\ (6\ 4)^t,\ (6\ 6)^t,\ (4\ 6)^t\right\}.$$

Έστω ότι οι a priori πιθανότητες είναι ίσες, $P(\omega_1) = P(\omega_2)$. Να βρείτε την εξίσωση που περιγράφει το σύνορο απόφασης Bayes μεταξύ των δυο αυτών κατηγοριών.

Λύση:

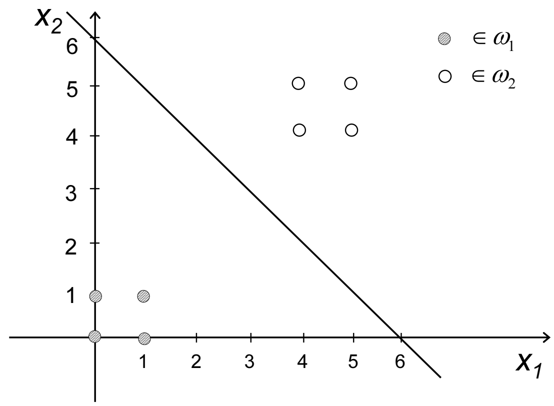
Έχουμε

$$P(\omega_1) = P(\omega_2) = \frac{1}{2},$$

$$\vec{\mu}_1 = \frac{1}{4}\sum_{j=1}^{4}\vec{x}_{1j} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

όπου $\vec{x}_{1j}$ είναι τα πρότυπα της κλάσης ω_1,

$$\vec{\mu}_2 = \frac{1}{4}\sum_{j=1}^{4}\vec{x}_{2j} = \begin{pmatrix} 5 \\ 5 \end{pmatrix},$$



Σχήμα 7:

όπου $\overrightarrow{x}_{2j}$ είναι τα πρότυπα της κλάσης ω_2. Οι πίνακες διασποράς είναι:

$$\overline{\overline{\Sigma}}_1 = \frac{1}{4}\sum_{j=1}^{4} \overrightarrow{x}_{1j}\overrightarrow{x}_{1j}^t - \overrightarrow{\mu}_1\overrightarrow{\mu}_1^t$$

$$\Rightarrow \overline{\overline{\Sigma}}_1 = \frac{1}{4}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \end{pmatrix}\begin{pmatrix} 2 & 0 \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \end{pmatrix}\begin{pmatrix} 2 & 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \end{pmatrix}\right] -$$

$$-\begin{pmatrix} 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I,$$

και ομοίως βρείτε οτι ο πίνακας διασποράς $\overline{\overline{\Sigma}}_2$ είναι:

$$\overline{\overline{\Sigma}}_2 = \frac{1}{4}\sum_{j=1}^{4} \overrightarrow{x}_{2j}\overrightarrow{x}_{2j}^t = I.$$

Εφόσον $\overline{\overline{\Sigma}}_1 = \overline{\overline{\Sigma}}_2 = \overline{\overline{\Sigma}} = I$, οι συναρτήσεις διάκρισης είναι:

$$g_1(\overrightarrow{x}) = \overrightarrow{x}^t\overline{\overline{\Sigma}}^{-1}\overrightarrow{\mu}_1 - \frac{1}{2}\overrightarrow{\mu}_1^t\overline{\overline{\Sigma}}^{-1}\overrightarrow{\mu}_1$$

$$\Rightarrow g_1(\overrightarrow{x}) = \overrightarrow{x}^t\overrightarrow{\mu}_1 - \frac{1}{2}\overrightarrow{\mu}_1^t\overrightarrow{\mu}_1 = \begin{pmatrix} x_1 & x_2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\Rightarrow g_1(\overrightarrow{x}) = x_1 + x_2 - 1.$$

Όμοια για τη $g_2(\overrightarrow{x})$ έχουμε:

$$g_2(\overrightarrow{x}) = \overrightarrow{x}^t \overline{\overline{\Sigma}} \overrightarrow{\mu}_2 - \frac{1}{2} \overrightarrow{\mu}_2^t \overline{\overline{\Sigma}} \overrightarrow{\mu}_2$$

$$\Rightarrow g_2(\overrightarrow{x}) = 5x_1 + 5x_2 - 25.$$

Επομένως, η επιφάνεια απόφασης είναι:

$$d_{12}(\overrightarrow{x}) = g_1(\overrightarrow{x}) - g_2(\overrightarrow{x}) = 0$$

$$\Rightarrow x_1 + x_2 - 6 = 0,$$

όπως φαίνεται στο σχήμα (7).

0.9

Να επαναλάβετε το πρόβλημα (0.8) για τις ακόλουθες κλάσεις:

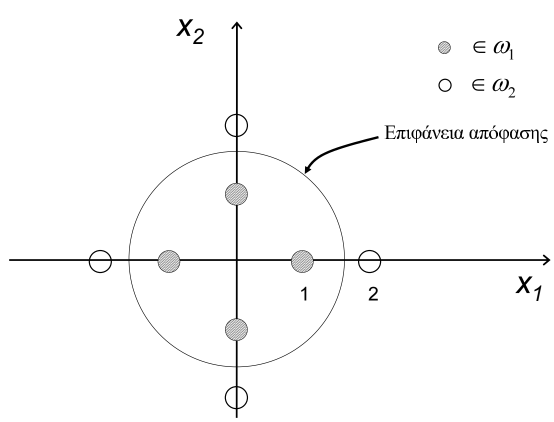
$$\omega_1 : \left\{(-1\ 0)^t,\ (0\ -1)^t,\ (1\ 0)^t,\ (0\ 1)^t\right\},$$

και

$$\omega_2 : \left\{(-2\ 0)^t,\ (0\ -2)^t,\ (2\ 0)^t,\ (0\ 2)^t\right\}.$$

Παρατηρείστε ότι οι δυο αυτές κλάσεις δεν είναι γραμμικά διαχωρίσιμες.

Λύση:



Σχήμα 8:

Οι μέσες τιμές των δύο κλάσεων είναι:

$$\overrightarrow{\mu}_1 = \frac{1}{4}\sum_{j=1}^{4} \overrightarrow{x}_{1j} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\overrightarrow{\mu}_2 = \frac{1}{4}\sum_{j=1}^{4} \overrightarrow{x}_{2j} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

όπου τα δείγματα $\overrightarrow{x}_{1j}$ και $\overrightarrow{x}_{2j}$ είναι:

$$\overrightarrow{x}_{1j} \in \left\{ \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\},$$

και

$$\overrightarrow{x}_{2j} \in \left\{ \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}.$$

Οι πίνακες διασποράς είναι:

$$\overline{\overline{\overline{\Sigma}}}_1 = \frac{1}{4}\sum_{j=1}^{4} \overrightarrow{x}_{1j}\overrightarrow{x}_{1j}^{t} - \overrightarrow{\mu}_1\overrightarrow{\mu}_1^{t}$$

$$\Rightarrow \overline{\overline{\overline{\Sigma}}}_1 = \frac{1}{4}\left[\begin{pmatrix} -1 \\ 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \end{pmatrix} \begin{pmatrix} 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} + \right.$$

$$\left. + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} \right] - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \end{pmatrix}$$

$$\Rightarrow \overline{\overline{\overline{\Sigma}}}_1 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}I,$$

και

$$\overline{\overline{\overline{\Sigma}}}_2 = \frac{1}{4}\left[\begin{pmatrix} -2 \\ 0 \end{pmatrix} \begin{pmatrix} -2 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -2 \end{pmatrix} \begin{pmatrix} 0 & -2 \end{pmatrix} + \right.$$

$$\left. + \begin{pmatrix} 2 \\ 0 \end{pmatrix} \begin{pmatrix} 2 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \end{pmatrix} \begin{pmatrix} 0 & 2 \end{pmatrix} \right] - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \end{pmatrix}$$

$$\overline{\overline{\overline{\Sigma}}}_2 = 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 2I.$$

Παρατηρούμε ότι $\overline{\overline{\overline{\Sigma}}}_1 \neq \overline{\overline{\overline{\Sigma}}}_2$ επομένως οι δύο αυτές κλάσεις ω_1, ω_2 δεν είναι γραμμικά διαχωρίσιμες. Οι συναρτήσεις διάκρισης είναι:

$$g_1(\overrightarrow{x}) = -\frac{1}{2}\ln|\overline{\overline{\overline{\Sigma}}}_1| - \frac{1}{2}(\overrightarrow{x} - \overrightarrow{\mu}_1)\overline{\overline{\overline{\Sigma}}}_1^{-1}(\overrightarrow{x} - \overrightarrow{\mu}_1)$$

$$\Rightarrow g_1(\overrightarrow{x}) = -\frac{1}{2}\ln(1/4) - \frac{1}{2}(\overrightarrow{x} - \overrightarrow{\mu}_1)^t(\overrightarrow{x} - \overrightarrow{\mu}_2)2$$

$$\Rightarrow g_1(\overrightarrow{x}) = \frac{\ln 4}{2} - \left(x_1^2 + x_2^2\right) = \ln 2 - \left(x_1^2 + x_2^2\right),$$

όπου

$$\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Όμοια:

$$g_2(\overrightarrow{x}) = -\frac{1}{2}\ln|\overline{\overline{\Sigma}}_2| - \frac{1}{2}(\overrightarrow{x} - \overrightarrow{\mu}_2)^t\overline{\overline{\Sigma}}_2^{1}(\overrightarrow{x} - \overrightarrow{\mu}_2)$$

$$\Rightarrow g_2(\overrightarrow{x}) = -\frac{1}{2}\ln 4 - \frac{1}{4}\overrightarrow{x}^t\overrightarrow{x}$$

$$\Rightarrow g_2(\overrightarrow{x}) = -\frac{\ln 4}{2} - \frac{1}{4}\left(x_1^2 + x_2^2\right) = \ln 2 - \frac{1}{4}\left(x_1^2 + x_2^2\right).$$

Επομένως, η επιφάνεια απόφασης είναι:

$$d_{12}(\overrightarrow{x}) = g_1(\overrightarrow{x}) - g_2(\overrightarrow{x}) = 0$$

$$\Rightarrow 2\ln 2 - \frac{3}{4}(x_1^2 + x_2^2) = 0$$

$$\Rightarrow x_1^2 + x_2^2 = \frac{8}{3}\ln 2.$$
