ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΤΥΠΩΝ (PATTERN RECOGNITION)

Σέργιος Θεοδωρίδης

Κωνσταντίνος Κουτρούμπας

Version 2

Εκφράζοντας τον ταξινομητή Bayes

(a) Με χρήση **συναρτήσεων διάκρισης** (**discriminant functions**)

- Έστω $g_q(\mathbf{x})=f(P(\omega_q)p(\mathbf{x}|\omega_q))$, $q=1,\ldots,M$, όπου f ***γνησίως αύξουσα*** συνάρτηση.

- Ταξινόμησε δεδομένο $\mathbf{x}$ στην κλάση ω_j για την οποία $g_j(\mathbf{x}) = max_{q=1,\ldots,M}\, g_q(\mathbf{x})$.

(b) Με χρήση **επιφανειών απόφασης** (**decision surfaces**)

- Ταξινόμησε δεδομένο $\mathbf{x}$ στην κλάση ω_j για την οποία $x \in R_j$, όπου

$$R_j=\{\mathbf{x} \in R^l \colon g_j(\mathbf{x}) = max_{q=1,\ldots,M}\, g_q(\mathbf{x}) \} = \{\mathbf{x} \in R^l \colon P(\omega_j|\mathbf{x}) = max_{q=1,\ldots,M}\, P(\omega_q|\mathbf{x}) \} =$$

$$\{\mathbf{x} \in R^l \colon P(\omega_j)p(\mathbf{x}|\omega_j) = max_{q=1,\ldots,M}\, P(\omega_q)p(\mathbf{x}|\omega_q) \}$$

ΣΗΜΕΙΩΣΗ: Οι R_js μπορούν να ταυτοποιηθούν μέσω των συνόρων τους με τις γειτονικές τους περιοχές.

Σύνορο κλάσεων ω_j, ω_q που αντιστοιχούν σε γειτ. περιοχές.: $g_{jq}(\mathbf{x}) \equiv g_j(\mathbf{x}) - g_q(\mathbf{x})=0$

Περιοχές απόφασης: $R_j=\{\mathbf{x} \in R^l \colon g_{jq}(\mathbf{x})>0\}$, $R_q=\{\mathbf{x} \in R^l \colon g_{jq}(\mathbf{x}) <0\}$

Σημαντική παρατήρηση: Στην πράξη η εύρεση της μίας έκφρασης από την άλλη είναι μια καθόλου προφανής διαδικασία. Αυτός είναι ένας βασικός λόγος για τον οποίο έχουν αναπτυχθεί ταξινομητές χρησιμοποιώντας ανεξάρτητα τις παραπάνω εκφράσεις.

Ο ταξινομητής Bayes: Η περίπτωση των κανονικών κατανομών

❖Επίσης $g_{jk}(\boldsymbol{x})=0 \Leftrightarrow g_j(\boldsymbol{x})-g_k(\boldsymbol{x})=0$ είναι **ισοδύναμη** με

❖Εξίσωση συνόρου συνεχόμενων κλάσεων

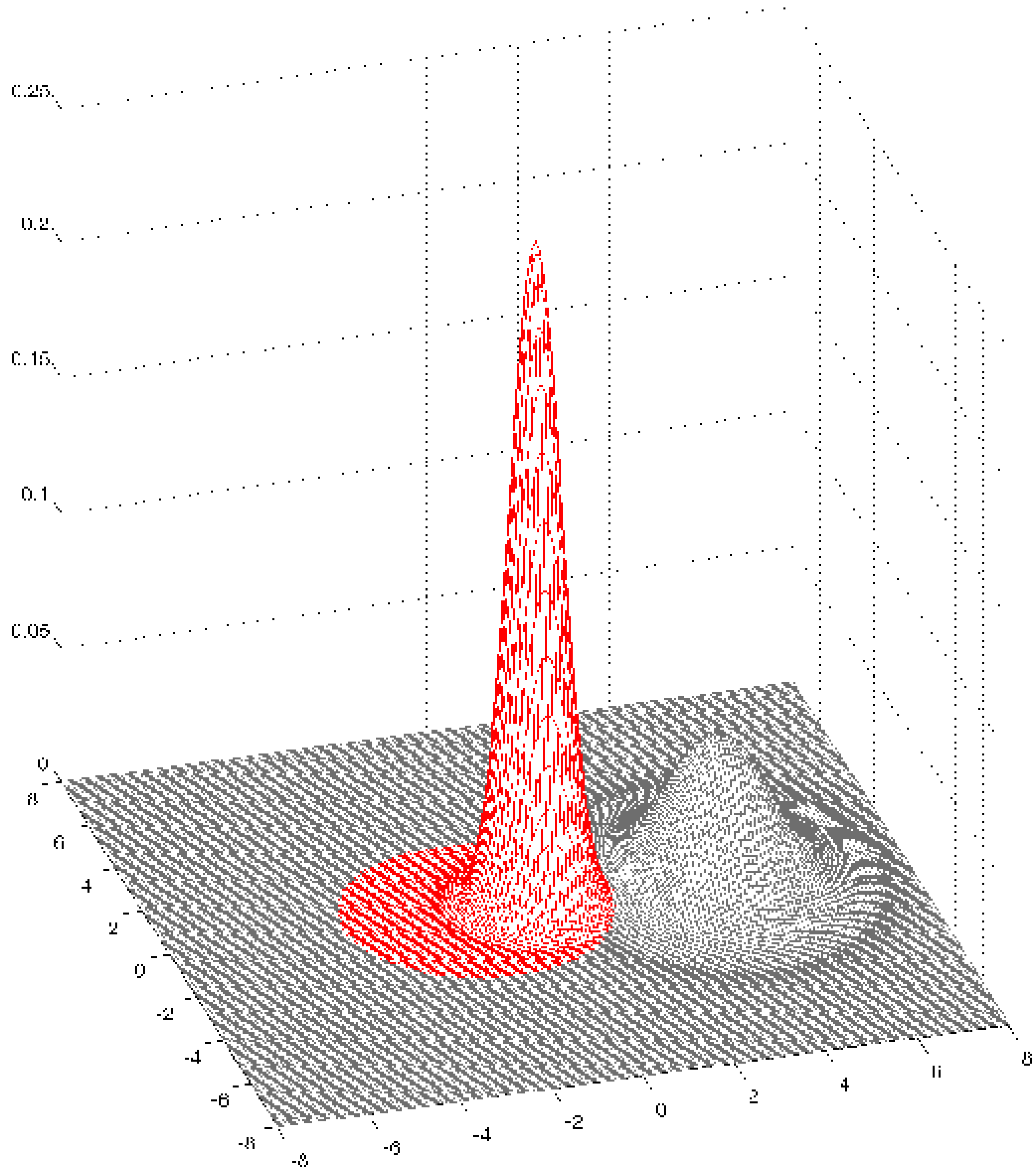
$$\begin{cases} -\frac{1}{2}x^T\Sigma_j^{-1}x+\mu_j^T\Sigma_j^{-1}x-\frac{1}{2}\mu_j^T\Sigma_j^{-1}\mu_j+c_j+\ln P(\omega_j)= \\ -\frac{1}{2}x^T\Sigma_k^{-1}x+\mu_k^T\Sigma_k^{-1}x-\frac{1}{2}\mu_k^T\Sigma_k^{-1}\mu_k+c_k+\ln P(\omega_k) \end{cases}$$

(i) $\Sigma_i \neq \Sigma_k$:

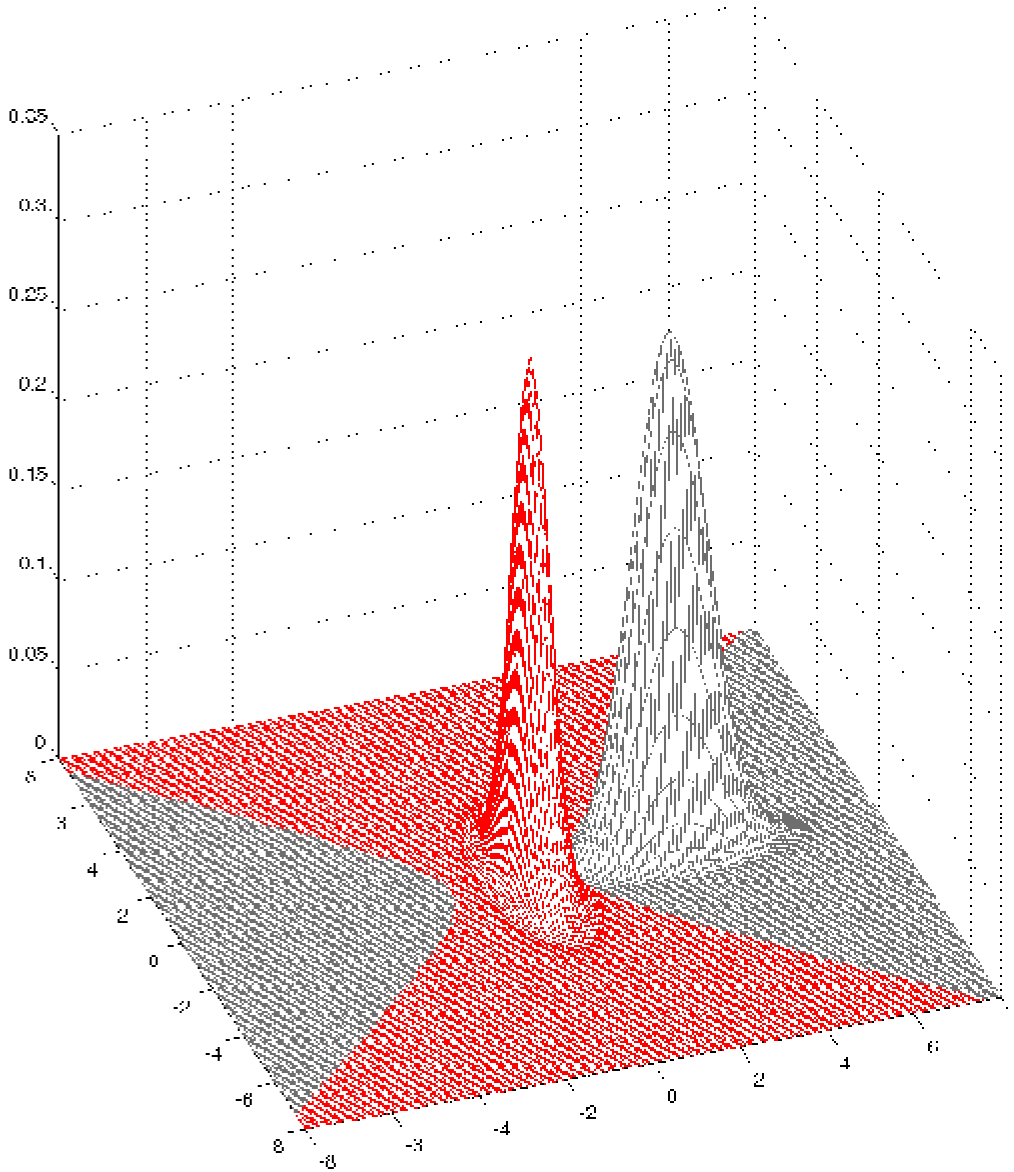
Σ' αυτή την περίπτωση η $g_{jk}(\boldsymbol{x})=0$ αναπαριστά μια καμπύλη δευτέρου βαθμού (π.χ. υπερ-έλλειψη, υπερ-παραβολή) (τετραγωνικός ταξινομητής - quadratic discriminant analysis)

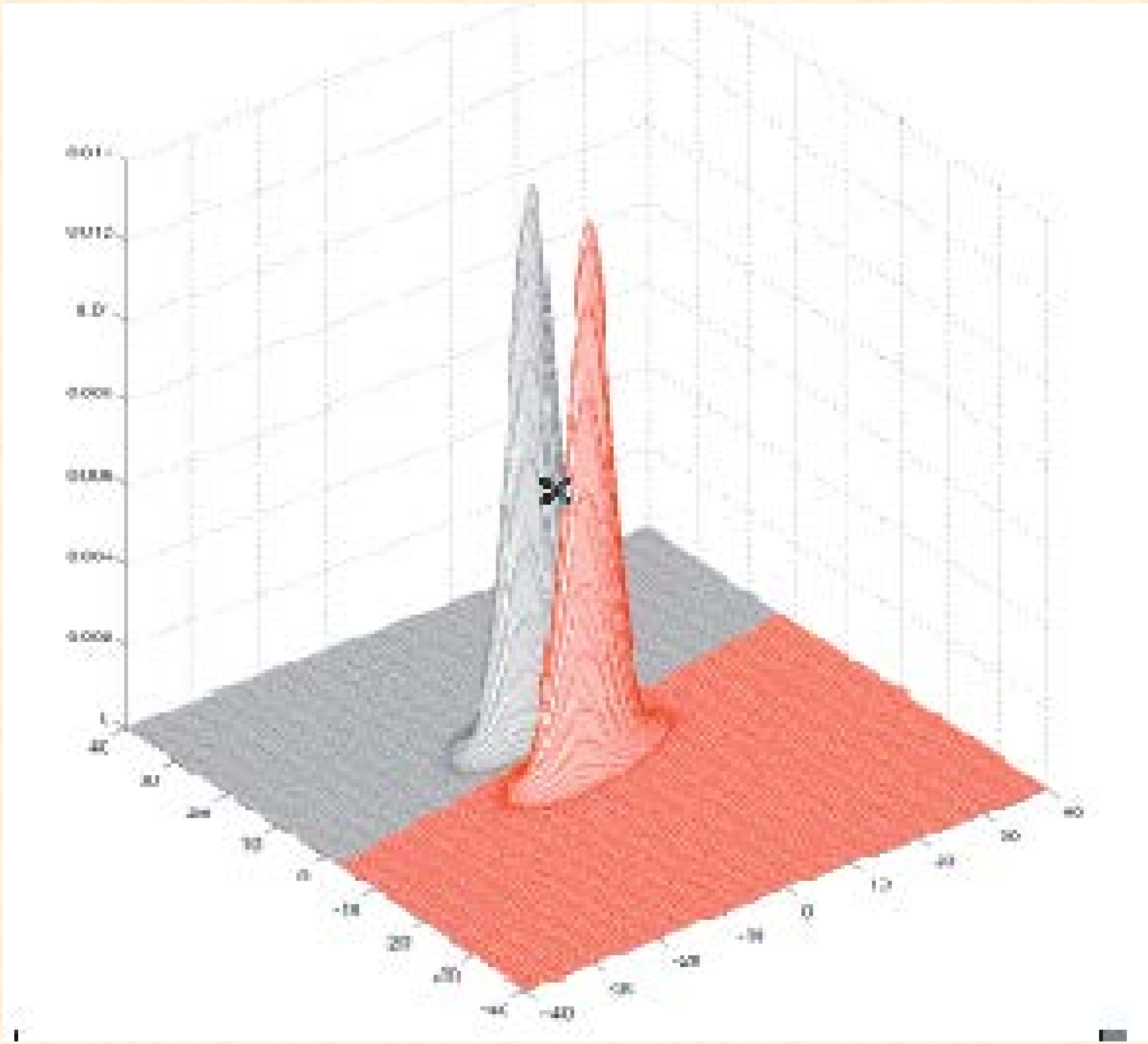
(ii) $\Sigma_i=\Sigma_k$:

Σ' αυτή την περίπτωση η $g_{jk}(\boldsymbol{x})=0$ αναπαριστά μια καμπύλη πρώτου βαθμού (υπερεπίπεδο) (γραμμικός ταξινομητής - linear discriminant analysis)



0.25
0.2
0.15
0.1
0.05
0
8
6
4
2
0
-2
-4
-6
-8
-8
-6
-4
-2
0
2
4
6
8


0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
8
6
4
2
0
-2
-4
-6
8
-8
-6
-4
-2
0
2
4
6
8



➢ Παράδειγμα: $\Sigma_i = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$

$$g_i(\underline{x}) = -\frac{1}{2\sigma^2}(x_1^2 + x_2^2) + \frac{1}{\sigma^2}(\mu_{i1}x_1 + \mu_{i2}x_2)$$

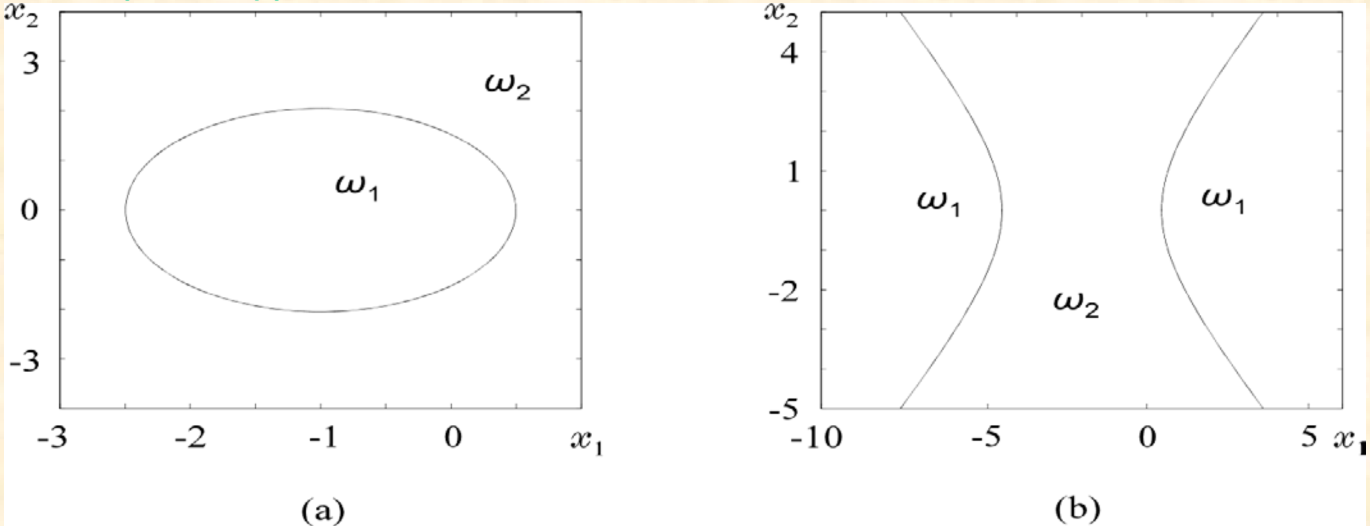
$$-\frac{1}{2\sigma^2}(\mu_{i1}^2 + \mu_{i2}^2) + \ln P(\omega_i) + C_i$$

Δηλαδή, η $g_i(x)$ είναι τετραγωνική. Σημειώστε ότι οι επιφάνειες

$$g_i(\underline{x}) - g_j(\underline{x}) = 0$$

ΔΕΝ είναι, απαραίτητα, τετραγωνικής μορφής (ελλείψεις, παραβολές,υπερβολές, ζεύγη ευθειών).

Για παράδειγμα:



Ο ταξινομητής Bayes: Η περίπτωση των κανονικών κατανομών

(ii)-a $\Sigma_i=\Sigma_k$ ***και*** $P(\omega_j)=P(\omega_k)$:

Αγνοώντας τις κοινές ποσότητες που εμφανίζονται σε όλες τις $g_j(\boldsymbol{x})$s, η $g_j(\boldsymbol{x})$ γίνεται

$$g_j(x) = -\frac{1}{2}(x-\mu_j)^T \Sigma^{-1}(x-\mu_j)$$

Θυμίζουμε ότι

$$d_M(x,\mu_j) = \left((x-\mu_j)^T \Sigma^{-1}(x-\mu_j)\right)^{1/2}$$

είναι η γνωστή **Mahalanobis απόσταση.**

Σ' αυτή την περίπτωση, ο ταξινομητής Bayes μπορεί να γραφτεί **ισοδύναμα** ως

- ❖Καταχώρησε το $\boldsymbol{x}$ στην ω_i με:
- ❖ $d_M(\boldsymbol{x},\boldsymbol{\mu}_j) = \boldsymbol{min}_{q=1,\ldots M}\, d_M(\boldsymbol{x},\boldsymbol{\mu}_q)$

Ταξινομητής ελάχιστης Mahalanobis απόστασης

Ο ταξινομητής Bayes: Η περίπτωση των κανονικών κατανομών

(ii)-b $\Sigma_i = \Sigma_k = \sigma^2 I$ ***και*** $P(\omega_j) = P(\omega_k)$:

Αγνοώντας τις κοινές ποσότητες που εμφανίζονται σε όλες τις $g_j(\boldsymbol{x})$s, η $g_j(\boldsymbol{x})$ γίνεται

$$g_j(x) = -\frac{1}{2}(x-\mu_j)^T(x-\mu_j) = -\frac{1}{2}\| x-\mu_j \|^2$$

όπου $||\boldsymbol{x}-\boldsymbol{\mu}_j||$ είναι η γνωστή **Ευκλείδεια απόσταση.**

Σ' αυτή την περίπτωση, ο ταξινομητής Bayes μπορεί να διατυπωθεί **ισοδύναμα** ως εξής

❖ Κατχώρησε το ***x*** στην ω_i με:
❖ $||\boldsymbol{x} - \boldsymbol{\mu}_j|| = \boldsymbol{min}_{q=1,\ldots M} ||\boldsymbol{x} - \boldsymbol{\mu}_q||$

Ταξινομητής ελάχιστης Ευκλείδειας απόστασης

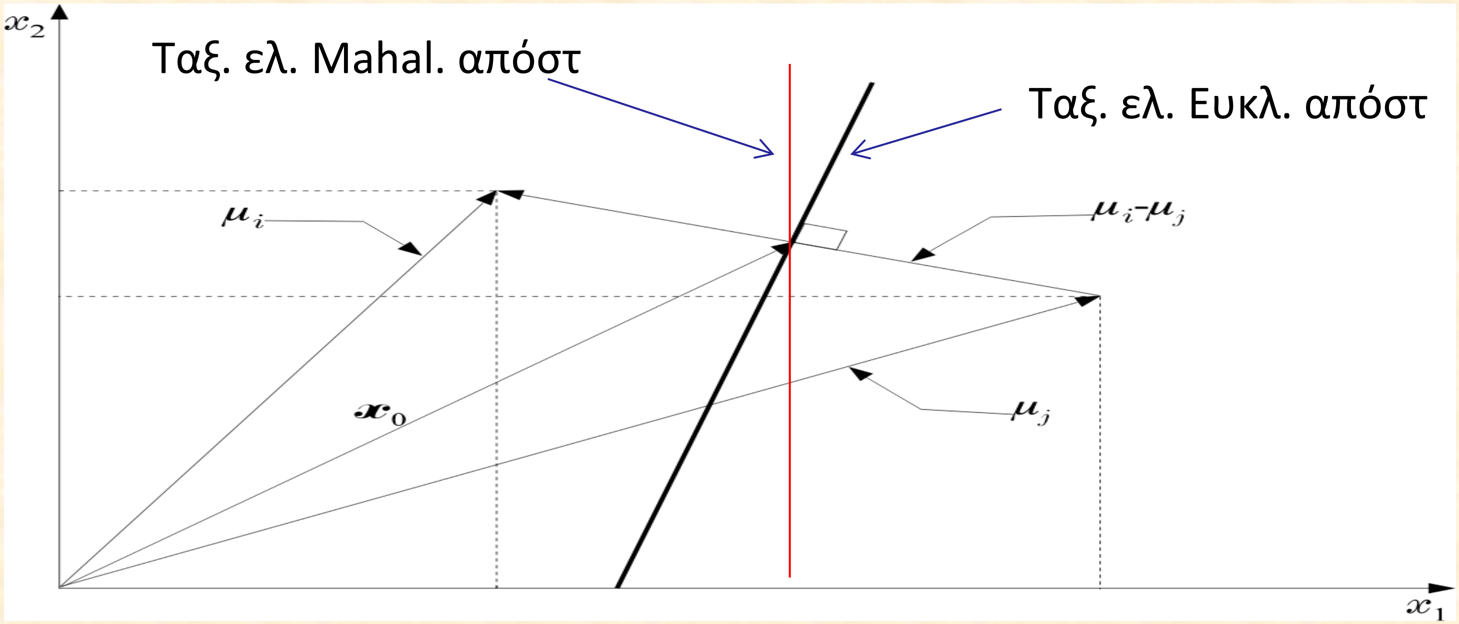
Ο ταξινομητής Bayes: Η περίπτωση των κανονικών κατανομών

(ii)-a-b

<u>Παρατήρηση:</u> Για προβλήματα ταξινόμησης <u>δύο κλάσεων</u>:

- ο ταξινομητής ελαχ. Ευκλείδειας απόστασης οριοθετεί τις περιοχές απόφασης των δύο κλάσεων με τη μεσοκάθετο του τμήματος που ενώνει τα μέσα διανύσματα των κλάσεων.

- ο ταξινομητής ελαχ. Mahalanobis απόστασης οριοθετεί τις περιοχές απόφασης των δύο κλάσεων με ευθεία που διέρχεται από το μέσο του τμήματος που ενώνει τα μέσα διανύσματα των κλάσεων, αλλά δεν είναι κάθετη σ' αυτό.



Παράδειγμα:

Δοθέντων των ω_1, ω_2 : $P(\omega_1) = P(\omega_2)$ και $p(\underline{x}|\omega_1) = N(\underline{\mu}_1, \quad),$

$$p(\underline{x}|\omega_2) = N(\underline{\mu}_2, \quad),\ \underline{\mu}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ \underline{\mu}_2 = \begin{bmatrix} 3 \\ 3 \end{bmatrix},\ \Sigma = \begin{bmatrix} 1.1 & 0.3 \\ 0.3 & 1.9 \end{bmatrix}$$

κατηγοριοποίησε το διάνυσμα $\underline{x} = \begin{bmatrix} 1.0 \\ 2.2 \end{bmatrix}$ με χρήση Bayesian ταξινομητή:

-> Είναι: Ταξινομητής Bayes $\Leftrightarrow$ Ταξινομητής ελαχ. Mahalanobis απόστασης

- $\Sigma^{-1} = \begin{bmatrix} 0.95 & -0.15 \\ -0.15 & 0.55 \end{bmatrix}$

- Υπόλ. Mahalanobis απόστ. d_m από μ_1, μ_2 : $d^2_{m,1} = \begin{bmatrix} 1.0, & 2.2 \end{bmatrix}$

$$\Sigma^{-1} \begin{bmatrix} 1.0 \\ 2.2 \end{bmatrix} = 2.952,\ d^2_{m,2} = \begin{bmatrix} -2.0, & -0.8 \end{bmatrix} \Sigma^{-1} \begin{bmatrix} -2.0 \\ -0.8 \end{bmatrix} = 3.672$$

- Καταχώρησε $\underline{x}$ -> ω_1.

Παράδειγμα:

Δοθέντων των ω_1, ω_2 : $P(\omega_1) = \Sigma P(\omega_2)$ και $p(\underline{x}|\omega_1) = N(\underline{\mu}_1, \quad),$

$$p(\underline{x}|\Sigma\omega_2) = N(\underline{\mu}_2, \quad), \ \underline{\mu}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \ \underline{\mu}_2 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \ \Sigma = \begin{bmatrix} 2.0 & 0.0 \\ 0.0 & 2.0 \end{bmatrix}$$

κατηγοριοποίησε το διάνυσμα $\underline{x} = \begin{bmatrix} 1.0 \\ 2.2 \end{bmatrix}$ με χρήση Bayesian ταξινομητή:

-> Είναι: Ταξινομητής Bayes $\Leftrightarrow$ Ταξινομητής ελαχ. Ευκλείδειας απόστασης

- Υπολ. Ευκλείδειας απόστ. του $\underline{x}$ από τα $\underline{\mu}_1$, $\underline{\mu}_2$:

$$d_{e,1}^2 = \begin{bmatrix} 1.0 & 2.2 \end{bmatrix} \begin{bmatrix} 1.0 \\ 2.2 \end{bmatrix} = 5.84, \qquad d_{e,2}^2 = \begin{bmatrix} -2.0 & -0.8 \end{bmatrix} \begin{bmatrix} -2.0 \\ -0.8 \end{bmatrix} = 4.64$$

- Καταχώρησε $\underline{x}$ -> ω_2.

<u>Παρατήρηση:</u> Οι δύο ταξινομητές καταχωρούν το ΙΔΙΟ διάνυσμα σε διαφορετικές κλάσεις. Αυτό οφείλεται στον διαφορετικό τρόπο μοντελοποίησης των κλάσεων σε κάθε περίπτωση.

Ο ταξινομητής Bayes: Η περίπτωση των κανονικών κατανομών

Ερώτηση: Κάτω από ποιες συνθήκες ο ταξινομητής ελάχιστης Ευκλείδειας απόστασης είναι βέλτιστος ως προς το κριτήριο της πιθανότητας λάθους;

Απάντηση: Όταν οι κατηγορίες

(a) είναι ισοπίθανες,

(b) Μοντελοποιούνται από κανονικές κατανομές με πίνακες συνδιασποράς της μορφής $\sigma^2 I$.

Ερώτηση: Κάτω από ποιες συνθήκες ένας γραμμικός ταξινομητής **μπορεί** να είναι βέλτιστος ως προς το κριτήριο της πιθανότητας λάθους;

Απάντηση: Όταν οι κατηγορίες μοντελοποιούνται από κανονικές κατανομές με κοινό μητρώο συνδιασποράς.

Ερώτηση: Κάτω από ποιες συνθήκες ένας τετραγωνικός ταξινομητής **μπορεί** να είναι βέλτιστος ως προς το κριτήριο της πιθανότητας λάθους;

Απάντηση: Όταν οι κατηγορίες μοντελοποιούνται από κανονικές κατανομές με διαφορετικές κατανομές συνδιασποράς.