

Αναγνώριση προτύπων
Εργασία 2
Σάββας Λιάπης 57403

Άσκηση 1

Για 2 ενδεχόμενα A και B, βρέθηκε ότι

- Μόνο το A συνέβη N_1 φορές
- Μόνο το B συνέβη N_2 φορές
- Το A και B συνέβη N_3 φορές
- Ούτε το A ούτε το B συνέβη N_4 φορές.

Ποιά είναι η δεσμευμένη πιθανότητα $P(A|B)$;

Απάντηση 1:

Η δεσμευμένη πιθανότητα $P[A|B]$ είναι η πιθανότητα να συμβεί το ενδεχόμενο A δεδομένου ότι έχει συμβεί το ενδεχόμενο B. μαθηματικά αυτό εκφράζεται από τον τύπο

$$P[A|B] = \frac{P[A \cap B]}{P[B]} \quad (1), \quad P[B] > 0$$

Δηλαδή ισούται με την πιθανότητα να συνέβη και το A και το B προς την πιθανότητα να συνέβη το B (δηλαδή και όταν συνέβη και το B μόνο και όταν συνέβη και A και B μαζί).

$$(1) \Leftrightarrow P[A|B] = \frac{\frac{N_3}{N_1 + N_2 + N_3 + N_4}}{\frac{N_2 + N_3}{N_1 + N_2 + N_3 + N_4}} \Leftrightarrow P[A|B] = \frac{N_3}{N_2 + N_3}$$

Άσκηση 2

Ο ακόλουθος πίνακας μας δίνει τις υπό συνθήκη πιθανότητες μιας τυχαίας μεταβλητής X για τρεις κατηγορίες ω_1 , ω_2 , και ω_3 (τρία ζάρια). Έστω ότι γνωρίζουμε τις a priori πιθανότητες $p(\omega_1) = 0.3$ και $p(\omega_2) = 0.3$. Δηλαδή, διαλέγουμε ένα ζάρι το ρίχνουμε και προσπαθούμε να μαντέψουμε ποιο ζάρι είχαμε διαλέξει. Υπολογίστε το ολικό σφάλμα της ταξινόμησης χρησιμοποιώντας τον κανόνα απόφασης Bayes.

$p(X=x \omega)$						
	$X=1$	$X=2$	$X=3$	$X=4$	$X=5$	$X=6$
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

Απάντηση 2:

Από τον κανόνα απόφασης Bayes γνωρίζουμε ότι :

$$P(\omega_j|x) = \frac{P(x|\omega_j)P(\omega_j)}{P(x)}$$

Οπότε χρησιμοποιώντας τα δεδομένα της άσκησης μπορούμε να καταλήξουμε στους παρακάτω πίνακες. Ο παρονομαστής δεν μας είναι γνωστός αλλά ουσιαστικά απλοποιείται κατά τις συγκρίσεις (κανονικά θα έπρεπε να έχουμε $P(\omega_j|x) P(x)$ άλλα αυτό το $P(x)$ το παραλείπουμε γιατί απλοποιείται).

	$P(\omega_1 x)$
1	0.09
2	0.06
3	0.03
4	0.03
5	0.06
6	0.03

	$P(\omega_2 x)$
1	0.06
2	0.06
3	0.12
4	0.015
5	0.03
6	0.015

	$P(\omega_3 x)$
1	0.04
2	0.12
3	0.06
4	0.02
5	0.12
6	0.04

Το συνολικό σφάλμα ταξινόμησης θα είναι :

$$\begin{aligned}
 P_{e,total} &= 1 - \sum_{i=1}^6 \left\{ \max_k [P(\omega_k|x_i)] \right\} = 1 - [P(\omega_3|1) + P(\omega_3|2) + P(\omega_2|3) + P(\omega_1|4) + \\
 &+ P(\omega_3|5) + P(\omega_3|6)] = 1 - (0.09 + 0.12 + 0.12 + 0.03 + 0.12 + 0.04) = 1 - 0.52 \Leftrightarrow \\
 &\Leftrightarrow P_{e,total} = 0.48
 \end{aligned}$$

Άσκηση 3

Υπολογίστε την εξίσωση της επιφάνειας απόφασης και το ολικό σφάλμα της κατηγοριοποίησης κατά Bayes για δυο κατηγορίες ω_1 και ω_2 , με $p(\omega_1)=0.25$, όπου θεωρούμε δείγματα σε 2- διαστάσεις από γκαουσιανές κατανομές που περιγράφονται από τις ακόλουθες πυκνότητες πιθανότητας:

$$p(\bar{x}|\omega_1) = N(\bar{\mu}_1, \bar{\Sigma}_1)$$

$$p(\bar{x}|\omega_2) = N(\bar{\mu}_2, \bar{\Sigma}_2)$$

Όπου :

$$\mu_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \Sigma_1 = \Sigma_2 = \Sigma = \begin{bmatrix} 4 & 1 \\ 1 & 9 \end{bmatrix}, \quad \mu_2 = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$$

Απάντηση 3:

εξίσωση της επιφάνειας απόφασης:

Για να υπολογίσουμε την εξίσωση της επιφάνειας απόφασης, αφού το Σ είναι ίδιο και για τις 2 κλάσεις θα χρησιμοποιήσουμε το:

$$\mathbf{w}^T(\mathbf{x} - \mathbf{x}_0) = 0 \quad (2)$$

Όπου :

$$\mathbf{w} = \Sigma^{-1}(\mu_1 - \mu_2) \quad (3)$$

$$\mathbf{x}_0 = \frac{1}{2}(\mu_1 + \mu_2) - \frac{\ln\left(\frac{P(\omega_1)}{P(\omega_2)}\right)}{(\mu_1 - \mu_2)^T(\mu_1 - \mu_2)}(\mu_1 - \mu_2) \quad (4)$$

Βρίσκω το \mathbf{w} :

$$(3) \Leftrightarrow \mathbf{w} = \begin{bmatrix} 4 & 1 \\ 1 & 9 \end{bmatrix}^{-1} [1 - (-2) \quad 2 - (-1)]$$

Βρίσκω τον Σ^{-1} :

$$\Sigma^{-1} = \frac{1}{4 * 9 - 1 * 1} \begin{bmatrix} 4 & -1 \\ -1 & 9 \end{bmatrix} = \frac{1}{35} \begin{bmatrix} 4 & -1 \\ -1 & 9 \end{bmatrix}$$

Άρα

$$\mathbf{w} = \frac{1}{35} \begin{bmatrix} 4 & -1 \\ -1 & 9 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \frac{1}{35} \begin{bmatrix} 24 \\ 9 \end{bmatrix} \Rightarrow \mathbf{w}^T = \frac{1}{35} [24 \quad 9] \quad (5)$$

$$(4) \Leftrightarrow \mathbf{x}_0 = \frac{1}{2} \begin{bmatrix} 1 - 2 \\ 2 - 1 \end{bmatrix} - \frac{\ln\left(\frac{0.25}{0.75}\right)}{\begin{bmatrix} 3 & 3 \end{bmatrix} \frac{1}{35} \begin{bmatrix} 4 & -1 \\ -1 & 9 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \end{bmatrix}} \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix} -$$

$$-35 \frac{1.0986}{\begin{bmatrix} 24 & 9 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \end{bmatrix}} \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{38.4514}{99} \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} + \begin{bmatrix} 1.1652 \\ 1.1652 \end{bmatrix} \Leftrightarrow \mathbf{x}_0 = \begin{bmatrix} 0.6652 \\ 1.6652 \end{bmatrix} \quad (6)$$

Από τα προηγούμενα έχουμε :

$$(2) \xLeftrightarrow{(5),(6)} \frac{1}{35} [24 \quad 9] \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0.6652 \\ 1.6652 \end{bmatrix} \right\} = 0 \Leftrightarrow [24 \quad 9] \begin{bmatrix} x_1 - 0.6652 \\ x_2 - 1.6652 \end{bmatrix} = 0 \Leftrightarrow$$

$$\Leftrightarrow 24(x_1 - 0.6652) + 9(x_2 - 1.6652) = 0 \Leftrightarrow x_2 = -2.6667x_1 + 3.4391$$

ολικό σφάλμα της κατηγοριοποίησης κατά Bayes:

η επίλυση του προβλήματος με 2 μεταβλητές είναι πολύ δύσκολη για αυτό είναι αναγκαίο να το γυρίσουμε σε 1 μεταβλητή : $y = \mathbf{w}^T \mathbf{x}$

Έτσι θα έχουμε για την πρώτη κλάση :

$$P(\mathbf{x}|\omega_1) = N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}) \Rightarrow P(\mathbf{w}^T \mathbf{x}|\omega_1) = N(\mathbf{w}^T \boldsymbol{\mu}_1, \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}) \Leftrightarrow$$

$$\Leftrightarrow P(y|\omega_1) = N(\mathbf{w}^T \boldsymbol{\mu}_1, \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w})$$

Όπου :

$$\mathbf{w}^T \boldsymbol{\mu}_1 = \frac{1}{35} [24 \quad 9] \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{42}{35} = 1.2$$

$$\mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} = \frac{1}{35} [24 \quad 9] \begin{bmatrix} 4 & 1 \\ 1 & 9 \end{bmatrix} \frac{1}{35} \begin{bmatrix} 24 \\ 9 \end{bmatrix} = \frac{3465}{1225} = 2.8286$$

$$\text{Άρα } P(y|\omega_1) = N(1.2, 2.8286)$$

Για την δεύτερη κλάση :

$$P(\mathbf{x}|\omega_2) = N(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}) \Rightarrow P(\mathbf{w}^T \mathbf{x}|\omega_2) = N(\mathbf{w}^T \boldsymbol{\mu}_2, \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}) \Leftrightarrow$$

$$\Leftrightarrow P(y|\omega_2) = N(\mathbf{w}^T \boldsymbol{\mu}_2, \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w})$$

Όπου :

$$\mathbf{w}^T \boldsymbol{\mu}_2 = \frac{1}{35} [24 \quad 9] \begin{bmatrix} -2 \\ -1 \end{bmatrix} = \frac{-57}{35} = -1.6286$$

$$\mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} = 2.8286 \text{ όπως παραπάνω}$$

$$\text{Άρα } P(y|\omega_2) = N(-1.6286, 2.8286)$$

Βρίσκουμε το σημείο απόφασης των 2 μονοδιάστατων πλέον κατανομών χρησιμοποιώντας την παρακάτω εξίσωση :

$$P(y|\omega_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\frac{1}{2}(y-\mu_i)^2}{2\sigma^2}} \quad (7)$$

$$\text{Οπότε για την } P(y|\omega_1) = N(1.2, 2.8286)$$

$$(7) \Leftrightarrow P(y|\omega_1) = \frac{1}{\sqrt{2.8286}\sqrt{2\pi}} e^{-\frac{\frac{1}{2}(y-1.2)^2}{2*2.8286}} = 0.2372 e^{-\frac{1}{11.34}(y-1.2)^2}$$

Οπότε για την $P(y|\omega_2) = N(1.6286, 2.8286)$

$$(7) \Leftrightarrow P(y|\omega_2) = \frac{1}{\sqrt{2.8286}\sqrt{2\pi}} e^{-\frac{\frac{1}{2}(y-1.6286)^2}{2*2.8286}} = 0.2372 e^{-\frac{1}{11.34}(y-1.6286)^2}$$

Άρα το σημείο απόφασης θα είναι :

$$P(y|\omega_1) = P(y|\omega_1) \Leftrightarrow 0.2372 e^{\frac{1}{11.34}(y-1.2)^2} = 0.2372 e^{\frac{1}{11.34}(y-1.6286)^2} \Leftrightarrow$$

$$\stackrel{ln}{\Leftrightarrow} (y-1.2)^2 = (y-1.6286)^2 \Leftrightarrow y^2 - 2.4y + 1.44 = y^2 - 3.2572y + 2.6523 \Leftrightarrow$$

$$\Leftrightarrow y = 1.4142$$

Οπότε το ολικό σφάλμα θα είναι :

$$\begin{aligned} P_{total\ error} &= \int_{-\infty}^{1.4142} P(y|\omega_2)P(\omega_2)dy + \int_{1.4142}^{\infty} P(y|\omega_1)P(\omega_1)dy = \\ &= \int_{-\infty}^{1.4142} 0.1779 e^{-\frac{1}{11.34}(y-1.6286)^2} dy + \int_{1.4142}^{\infty} 0.0593 e^{-\frac{1}{11.34}(y-1.2)^2} dy = \\ &= 0.4298 + 0.1643 \Leftrightarrow P_{total\ error} = 0.4928 + 0.1643 = 0.6571 \end{aligned}$$

Ο υπολογισμός των ολοκληρωμάτων έγινε με την βοήθεια του matlab όπου χρησιμοποίησα τις παραπάνω εντολές στο command window :

```
syms x
funct1=0.1779*(exp(-(1/11.34)*((x-1.6286)^2)))
funct2=0.0593*(exp(-(1/11.34)*((x-1.2)^2)))
int(funct2, -inf, 1.4142)
int(funct1, 1.4142, inf)
```

Άσκηση 4

(α) Να ευρεθεί η βέλτιστη λύση (απόφαση), όταν

- $\Omega = \{\omega_1, \omega_2\}$
- $P(x|\omega_1) = N(2, 0.5)$
- $P(x|\omega_2) = N(1.5, 0.2)$
- $P(\omega_1) = 1/3$
- $P(\omega_2) = 2/3$
- με $\lambda = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$

και να υπολογισθεί το ελάχιστο κόστος.

(β) Να προσομοιωθεί η διαδικασία υπολογιστικά, δημιουργώντας τυχαίους αριθμούς που ακολουθούν την κανονική κατανομή με την εντολή randn, και να εκτιμηθεί αριθμητικά το κόστος από την λύση (α).

Απάντηση 4:

Υπολογισμός βέλτιστης λύσης :

Έχουμε δύο κατηγορίες οπότε καταλήγουμε στο ακόλουθο ρίσκο :

$$R(a_1|x) = \lambda_{11}P(\omega_1|x) + \lambda_{12}P(\omega_2|x)$$

$$R(a_2|x) = \lambda_{21}P(\omega_1|x) + \lambda_{22}P(\omega_2|x)$$

Το κριτήριο θα δίνει απόφαση ω_1 αν $R(a_1|x) < R(a_2|x) \Leftrightarrow$

$$\Leftrightarrow (\lambda_{21} - \lambda_{11})P(\omega_1|x) < (\lambda_{12} - \lambda_{22})P(\omega_2|x) \Leftrightarrow$$

$$\Leftrightarrow (\lambda_{21} - \lambda_{11}) \frac{P(x|\omega_1)P(\omega_1)}{P(x)} < (\lambda_{12} - \lambda_{22}) \frac{P(x|\omega_2)P(\omega_2)}{P(x)} \Leftrightarrow$$

$$\Leftrightarrow \frac{P(x|\omega_1)}{P(x|\omega_2)} > \frac{(\lambda_{12} - \lambda_{22})P(\omega_2)}{(\lambda_{21} - \lambda_{11})P(\omega_1)} \Leftrightarrow \frac{P(x|\omega_1)}{P(x|\omega_2)} > 2 \frac{2-1}{3-1} \Leftrightarrow \frac{P(x|\omega_1)}{P(x|\omega_2)} > 1 \Leftrightarrow$$

$$\Leftrightarrow P(x|\omega_1) > P(x|\omega_2) \quad (8)$$

Τα $P(x|\omega_1)$ και $P(x|\omega_2)$ ακολουθούν κανονική κατανομή οπότε η συνάρτησή τους ακολουθεί την :

$$P(x|\omega_i) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_i|^{\frac{1}{2}}} e^{-\frac{1}{2} \{ [x - \bar{\mu}_i]^T \Sigma_i^{-1} [x - \bar{\mu}_i] \}} \quad (9)$$

- $d=1$ (γιατί έχουμε 1 διάσταση) $\Rightarrow (2\pi)^{\frac{d}{2}} = 2\pi^{\frac{1}{2}} = 2.5066$

Τώρα για $P(x|\omega_1)$:

- $\mu_1 = 2$
- $\Sigma_1 = 0.5$ άρα $|\Sigma_1|^{\frac{1}{2}} = 0.7071$ και $\Sigma_1^{-1} = 2$

$$(9) \Leftrightarrow P(x|\omega_1) = \frac{1}{2.5066 * 0.7071} e^{-\frac{1}{2} \{ (x-2)^2 (x-2) \}} \Leftrightarrow P(x|\omega_1) = 0.5642 e^{-(x-2)^2} \quad (10) \Leftrightarrow$$

Για $P(x|\omega_2)$:

- $\mu_2 = 1.5$
- $\Sigma_2 = 0.5$ άρα $|\Sigma_2|^{\frac{1}{2}} = 0.4472$ και $\Sigma_2^{-1} = 5$

$$(9) \Leftrightarrow P(x|\omega_2) = \frac{1}{2.5066 \cdot 0.4472} e^{-\frac{5}{2}(x-1.5)(x-1.5)} \Leftrightarrow P(x|\omega_2) = 0.8921 e^{-\frac{5}{2}(x-1.5)^2} \quad (11)$$

Οπότε:

$$(8) \stackrel{(10),(11)}{\Leftrightarrow} 0.5642 e^{-(x-2)^2} > 0.8921 e^{-\frac{5}{2}(x-1.5)^2} \Leftrightarrow$$

$$\stackrel{\ln}{\Leftrightarrow} \ln(0.5642) - (x^2 - 4x + 4) > \ln(0.8921) - \frac{5}{2}(x^2 - 3x + 2.25) \Leftrightarrow$$

$$\Leftrightarrow 3x^2 - 7x + 2.3338 > 0$$

οι ρίζες του παραπάνω πολυωνύμου είναι : 0.403 και 1.9303 και το πολυώνυμο είναι >0 για $x \in (-\infty, 0.403) \cup (1.9303, +\infty)$

Οπότε το κριτήριο θα δίνει βέλτιστη απόφαση

- ω_1 για $x \in (-\infty, 0.403) \cup (1.9303, +\infty)$
- ω_2 για $x \in [0.403, 1.9303]$

Το παραπάνω αποτέλεσμα είναι θεωρητικό. Ωστόσο ξέρουμε ότι στην κανονική κατανομή η πιθανότητα ένα σημείο να ανήκει σε αυτή είναι σχεδόν 1 στο διάστημα $[\mu-3\sigma, \mu+3\sigma]$.

Για την κατανομή του likelihood 1 : $\mu=2$, $\sigma=0.7071$ οπότε η πιθανότητα να υπάρξει σημείο έξω από το διάστημα $[2-3 \cdot 0.7071, 2+3 \cdot 0.7071]$ ή $[-0.1213, 4.1213]$ είναι πρακτικά 0.

Για την κατανομή του likelihood 2 : $\mu=1.5$, $\sigma=0.4472$ οπότε η πιθανότητα να υπάρξει σημείο έξω από το διάστημα $[1.5-3 \cdot 0.4472, 1.5+3 \cdot 0.4472]$ ή $[0.1584, 2.8416]$ είναι πρακτικά 0.

Άρα πρακτικά το κριτήριο θα δίνει βέλτιστη απόφαση :

- ω_1 για $x \in [-0.1213, 0.403) \cup (1.9303, 4.1213]$
- ω_2 για $x \in [0.403, 1.9303]$

Υπολογισμός ελάχιστου κόστους :

Έστω $\lambda_{12} - \lambda_{22}$ το κόστος αν επιλέξω την υπόθεση H_1 δεδομένου ότι η H_2 είναι σωστή
Έστω επίσης $\lambda_{21} - \lambda_{11}$ το κόστος αν επιλέξω την υπόθεση H_2 δεδομένου ότι η H_1 είναι σωστή.

Επιλέγουμε την H_2 με βάση το ελάχιστο κόστος αν ισχύει :

$$\begin{aligned}(\lambda_{21} - \lambda_{11}) \frac{P(x | H_2)P(H_2)}{P(x)} &> (\lambda_{12} - \lambda_{22}) \frac{P(x | H_1)P(H_1)}{P(x)} \xLeftrightarrow{H_1=\omega_1, H_2=\omega_2} \\ \Leftrightarrow \frac{P(x | \omega_2)}{P(x | \omega_1)} &> \frac{\lambda_{12} - \lambda_{22}}{\lambda_{21} - \lambda_{11}} \frac{P(\omega_1)}{P(\omega_2)} \xLeftrightarrow{(10),(11)} \frac{0.8921e^{-\frac{5}{2}(x-1.5)^2}}{0.5642e^{-(x-2)^2}} > \frac{1}{2} \frac{2-1}{3-1} \Leftrightarrow \\ \Leftrightarrow \frac{0.8921e^{-\frac{5}{2}(x-1.5)^2}}{0.5642e^{-(x-2)^2}} &> \frac{1}{4} \Leftrightarrow 0.8921e^{-\frac{5}{2}(x-1.5)^2} > \frac{0.5642}{4} e^{-(x-2)^2} \xLeftrightarrow{\ln} \\ \Leftrightarrow \ln(0.8921) - \frac{5}{2}(x^2 - 3x + 2.25) &> \ln(0.1410) - (x^2 - 4x + 4) \Leftrightarrow \\ \Leftrightarrow 3x^2 - 7x - 0.4396 &< 0\end{aligned}$$

οι ρίζες του παραπάνω πολυωνύμου είναι : -0.0612 και 2.3945 και το πολυώνυμο είναι <0 για $x \in (-0.0612, 2.3945)$

Οπότε το κριτήριο θα δίνει, με βάση το ελάχιστο κόστος, την υπόθεση :

- H_1 για $x \in (-\infty, -0.0612) \cup (2.3945, +\infty)$
- H_2 για $x \in (-0.0612, 2.3945)$

Πρακτικά όμως το κριτήριο θα δίνει (από τα διαστήματα που υπολογίσαμε στο προηγούμενο ερώτημα) , με βάση το ελάχιστο κόστος, την υπόθεση :

- H_1 για $x \in [-0.1213, -0.0612) \cup (2.3945, 4.1213]$
- H_2 για $x \in [-0.0612, 2.3945]$

Υλοποίηση υπολογιστικά του υπολογισμού ελαχίστου κόστους (σύντομη περιγραφή)

Στον αλγόριθμο question4b δημιουργούμε τα 2 datasets βάσει των ρυθμίσεων που προκαθορίζονται από την άσκηση, τα συγχωνεύουμε και τρέχουμε τον αλγόριθμο επιλογής κατά bayes με βάση το ελάχιστο κόστος σε κάθε επανάληψη του αλγορίθμου συγκεντρώνουμε τα δεδομένα που όντως κατατάχτηκαν σε λάθος ενδεχόμενο . Μετά το τέλος της επανάληψης υπολογίζουμε την αριθμητική εκτίμηση του κόστους που προκύπτει από τα μετρήσιμα ενδεχόμενα που μπήκαν σε λάθος κλάση από αυτή που έπρεπε να μπουν .