ΤΟΜΕΑΣ ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΕΡΓΑΣΤΗΡΙΟ ΑΝΑΛΥΣΗΣ ΗΛΕΚΤΡΙΚΩΝ ΚΥΚΛΩΜΑΤΩΝ
ΜΑΘΗΜΑ: ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΤΥΠΩΝ – 9ο ΕΞΑΜΗΝΟ

Εργασία 3

(Προθεσμία: Κυριακή 15 Νοεμβρίου 2020)

ΑΣΚΗΣΗ 3.1

Για πολλές ασκήσεις θα χρειαστείτε τις ακόλουθες διαδικασίες. Χρησιμοποιώντας MATLAB (καλό είναι να έχετε και το statistics toolbox), γράψτε τα παρακάτω προγράμματα.

(α) Γράψτε ένα πρόγραμμα για τον υπολογισμό της συνάρτησης διάκρισης της μορφής

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln(|\boldsymbol{\Sigma}_i|) + \ln(P(\omega_i))$$

για μια δεδομένη κανονική κατανομή d διαστάσεων και εκ των προτέρων πιθανότητα $P(\omega_i)$.

(β) Γράψτε ένα πρόγραμμα για τον υπολογισμό της Ευκλείδειας απόστασης μεταξύ δύο αυθαίρετων σημείων d-διαστάσεων $\mathbf{x}_1$ και $\mathbf{x}_2$.

(γ) Τι είναι η **Mahalanobis distance?** Γράψτε ένα πρόγραμμα για τον υπολογισμό της απόστασης **Mahalanobis** μεταξύ της μέσης $\boldsymbol{\mu}$ και ενός αυθαίρετου σημείου $\mathbf{x}$, δεδομένου του πίνακα συνδιασποράς $\boldsymbol{\Sigma}$.

ΑΣΚΗΣΗ 3.2

Δίνονται τα παρακάτω δεδομένα

	ω_1			ω_2			ω_3		
Δείγμα	x_1	x_2	x_3	x_1	x_2	x_3	x_1	x_2	x_3
1	−5.01	−8.12	−3.68	−0.91	−0.18	−0.05	5.35	2.26	8.13
2	−5.43	−3.48	−3.54	1.30	−2.06	−3.53	5.12	3.22	−2.66
3	1.08	−5.52	1.66	−7.75	−4.54	−0.95	−1.34	−5.31	−9.87
4	0.86	−3.78	−4.11	−5.47	0.50	3.92	4.48	3.42	5.19
5	−2.67	0.63	7.39	6.14	5.72	−4.85	7.11	2.39	9.21
6	4.94	3.29	2.08	3.60	1.26	4.36	7.17	4.33	−0.98
7	−2.51	2.09	−2.59	5.37	−4.63	−3.65	5.75	3.97	6.65
8	−2.25	−2.13	−6.94	7.18	1.46	−6.66	0.77	0.27	2.41
9	5.56	2.86	−2.26	−7.39	1.17	6.30	0.90	−0.43	−8.71
10	1.03	−3.33	4.33	−7.50	−6.32	−0.31	3.52	−0.36	6.43

Χρησιμοποιήστε τον ταξινομητή σας από το πρόβλημα 3.1.α για να κατατάξετε τα ακόλουθα 10 δείγματα από τον παραπάνω πίνακα με τον ακόλουθο τρόπο. Υποθέτουμε ότι οι υποκείμενες κατανομές είναι κανονικές (Gaussian).

1. Υποθέστε ότι οι εκ των προτέρων πιθανότητες για τις δύο πρώτες κατηγορίες είναι ίσες ($P(\omega_1) = P(\omega_2) = 1/2$ και $P(\omega_3) = 0$) και βρείτε το σημείο διαχωρισμού για τις δύο αυτές κατηγορίες χρησιμοποιώντας μόνο το χαρακτηριστικό x_1.
2. Προσδιορίζεται το εμπειρικό σφάλμα κατάρτισης σχετικά με τα δείγματα, δηλαδή σας, το ποσοστό των σημεία ταξινομείται εσφαλμένα
3. Επαναλάβετε τα παραπάνω, αλλά χρησιμοποιώντας τώρα δύο χαρακτηριστικά, x_1, x_2 και
4. Επαναλάβετε, αλλά χρησιμοποιώντας τώρα και τις τρεις τιμές των χαρακτηριστικών
5. Συζητήστε τα αποτελέσματά σας. Ειδικότερα, είναι ποτέ δυνατόν για ένα πεπερασμένο σύνολο των στοιχείων το εμπειρικό σφάλμα να είναι μεγαλύτερο όταν αυξάνεται η διάσταση των χαρακτηριστικών;
6. Υποθέστε ότι οι εκ των προτέρων πιθανότητες για τις κατηγορίες είναι $P(\omega_1) = 0.8$ και $P(\omega_2) = P(\omega_3) = 0.1$, βρείτε τις συναρτήσεις διαχωρισμού για τις τρεις αυτές κατηγορίες χρησιμοποιώντας όλα τα χαρακτηριστικά x_1. x_2, και x_3.

ΑΣΚΗΣΗ 3.3 Εκτιμητές Bayes

Πετάμε ένα νόμισμα N=10 φορές και φέρνουμε {κ,κ,γ,κ,γ,κ,κ,κ,γ,κ} (κ=κεφάλι,γ=γράμματα). Ποια είναι η πιθανότητα **θ** να φέρουμε κεφάλι; Η αρχική κατανομή του θ είναι

$$p(\theta \mid D^0) = \begin{cases} A\sin(\pi\theta) & \text{για } 0 \le \theta \le 1 \\ 0 & \text{αλλού} \end{cases} .$$

1. Υπολογισθεί το A
2. Να σχεδιασθεί το $p(\theta \mid D^1),\ p(\theta \mid D^5), p(\theta \mid D^{10})$
3. Να βρεθεί (αριθμητικά) το $p(x = k \mid D^{10})$

ΑΣΚΗΣΗ 3.4

Έστω πρόβλημα με 3 κλάσεις (ω_1, ω_2, ω_3) με δεδομένα από κανονικές κατανομές με ((μ_1, Σ_1), (μ_2, Σ_2), (μ_3, Σ_3)), και με πιθανότητες ($P(\omega_1)=p_1$, $P(\omega_2)=p_2$, $P(\omega_3)=p_3$)

Α. Δημιουργήστε 2 τυχαία δείγματα {X} και {X_1}, με 10000 και 1000 δείγματα (πρότυπα), αντίστοιχα που να ακολουθούν τις παραπάνω κατανομές, σύμφωνα με τις δεδομένες εκ των προτέρων πιθανότητες. Χρησιμοποιήστε το δείγμα {X} ως «σύνολο εκμάθησης» και το {X_1} ως «σύνολο δοκιμής».

$p_1= p_2= p_3= 1/3$, $\mu_1=[0,0,0]^T$, $\mu_2=[1,2,2]^T$, $\mu_3=[3,3,4]^T$

και $\Sigma=\Sigma_1= \Sigma_2= \Sigma_3=$ $\begin{bmatrix} 0{,}8 & 0{,}2 & 0{,}1 \\ 0{,}2 & 0{,}8 & 0{,}2 \\ 0{,}1 & 0{,}2 & 0{,}8 \end{bmatrix}$

Β. Χρησιμοποιώντας τις παραπάνω παραμέτρους, επιλέξτε τον κατάλληλο ταξινομητή και ταξινομήστε το σύνολο {X_1} και υπολογίστε την πιθανότητα λάθους.

Γ. Χρησιμοποιώντας το σύνολο {X} και την μέθοδο της μεγίστης πιθανοφάνειας, εκτιμήστε τα μ_1, μ_2, μ_3, και το Σ, και μετά επιλέξτε τον κατάλληλο ταξινομητή και ταξινομήστε το σύνολο {X_1} και υπολογίστε την πιθανότητα λάθους.

Δ. Χρησιμοποιώντας,

$p_1= 1/6$ $p_2= 1/6$ $p_3= 2/3$, $\mu_1=[0,0,0]^T$, $\mu_2=[1,2,2]^T$, $\mu_3=[3,3,4]^T$

και $\Sigma_1= \begin{bmatrix} 0,8 & 0,2 & 0,1 \\ 0,2 & 0,8 & 0,2 \\ 0,1 & 0,2 & 0,8 \end{bmatrix}$ $\Sigma_2= \begin{bmatrix} 0,6 & 0,2 & 0,01 \\ 0,2 & 0,8 & 0,01 \\ 0,01 & 0,01 & 0,6 \end{bmatrix}$ $\Sigma_3= \begin{bmatrix} 0,6 & 0,1 & ,0,1 \\ 0,1 & 0,6 & 0,1 \\ 0,1 & 0,1 & 0,6 \end{bmatrix}$

επαναλάβετε τα Α, Β και Γ.

Σχολιάστε τα αποτελέσματα σας.

Ξάνθη, 4/11/2020