Αναγνώριση προτύπων
Εργασία 3
Σάββας Λιάπης 57403

Άσκηση 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
$$g_i(\boldsymbol{x}) = -\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_i) - \frac{d}{2}\ln(2\pi) - \frac{1}{2}ln(|\Sigma_i|) + \ln\big(P(\omega_i)\big)$$
για μια δεδομένη κανονική κατανομή d διαστάσεων και εκ των προτέρων πιθανότητα P(ωi).

β. Γράψτε ένα πρόγραμμα για τον υπολογισμό της Ευκλείδειας απόστασης μεταξύ δύο αυθαίρετων σημείων d-διαστάσεων **x1** και **x2**

γ. Τι είναι η Mahalanobis distance? Γράψτε ένα πρόγραμμα για τον υπολογισμό της απόστασης Mahalanobis μεταξύ της μέσης **μ** και ενός αυθαίρετου σημείου **x**, δεδομένου του πίνακα συνδιασποράς **Σ.**

Απάντηση 3.1(α):

```
function [ discr_func ] = discriminant(x,mean,cov, prior )
%returns the discriminant function
%it takes as inputs:
%1.the vector of random variables x
%2.the vector of mean values mean
%3.the covariance matrices Σ
%and computes the discriminant function g(x)

%the dimension of the distribution
dim=length(x);

%we compute the vector and matrices product and the constant part
%separately and then we sum them to
matrixprod= ((x-mean)'*inv(cov))*(x-mean);
constpart= -(dim/2)*log(2*pi)-(1/2)*log(det(cov))+log(prior);

discr_func=-(1/2)*matrixprod +constpart;
end
```

η συνάρτηση αυτή λαμβάνει ως εισόδους τον διάνυσμα των τυχαίων μεταβλητών το διάνυσμα του μέσου όρου και τον πίνακα συνδιασποράς και βγάζει σαν αποτέλεσμα την κατάλληλη συνάρτηση διάκρισης αφού υπολογίσει πρώτα το γινόμενο των πινάκων και το σταθερό μέρος ξεχωριστά και τα προσθέσει στο τέλος.

Απάντηση 3.1(β):

```
function [ eucl_dist ] = euclidean(point1,point2)
%euclidean distance
%it gets as inputs the vector x of the point to wich we want
%compute the distance and returns the euclidian distance
eucl_dist=sqrt(((point1-point2)')*(point1-point2));
end
```

η συνάρτηση αυτή λαμβάνει ως εισόδους τις θέσεις 2 σημείων και υπολογίζει την ευκλείδεια απόσταση τους

Απάντηση 3.1(γ):

```
function [ mah_dist ] = mahalanobis(x,mean,cov )
%computes the mahalanobis distance of a point to a distribution
%it takes as inputs the vectors of the point we want to examine
%the mean and the covariance matrix and computes the mahalanobis
distance

mah_dist=sqrt(((x-mean)'*inv(cov))*(x-mean));
end
```

Η συνάρτηση αυτή υπολογίζει την απόσταση μεταξύ της μέσης τιμής μ και ενός αυθαίρετου σημείου x, για έναν δεδομένο πίνακα συνδιασποράς Σ.

Άσκηση 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(ω1) = P(ω2) = 1/2 και P(ω3) = 0) και βρείτε το σημείο διαχωρισμού για τις δύο αυτές κατηγορίες χρησιμοποιώντας μόνο το χαρακτηριστικό x1.
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, x1, x2.
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(ω1) = 0.8 και P(ω2) = P(ω3) = 0.1, βρείτε τις συναρτήσεις διαχωρισμού για τις τρεις αυτές κατηγορίες χρησιμοποιώντας όλα τα χαρακτηριστικά x1. x2, και x3.

Απάντηση 3.2(1):

Σε αυτή την περίπτωση έχουμε να ταξινομήσουμε τα δείγματά μας βάσει ενός χαρακτηριστικού. Σημειωτέων ότι το prior της τρίτης κλάσης είναι μηδενικό (ουσιαστικά δεν υπάρχει), οπότε έχουμε 20 δείγματα προς μελέτη. Για να βρούμε την εξίσωση διαχωρισμού των 2 κλάσεων εξισώνουμε τις συναρτήσεις discriminant των 2 κλάσεων και λύνουμε ως προς x1. Η ισότητα ικανοποιείται στα :

```
boundary =

 -5.09
  4.33
```

Απάντηση 3.2(2):

```
class =

     1     0
     0     1
     1     0
     1     0
     1     0
     0     1
     1     0
     1     0
     0     1
     1     0
     1     0
     1     0
     0     1
     0     1
     0     1
     1     0
     0     1
     0     1
     0     1
     0     1

wrong =

     6

error =

    0.3000
```

Χρησιμοποιώντας την συνάρτηση discriminant που φτιάξαμε για το ερώτημα 3.1.α εξάγουμε την συνάρτηση διάκρισης για κάθε κλάση και έπειτα τις συγκρίνουμε για να δούμε που κατατάσσεται το δείγμα που μελετάμε. Κάθε δείγμα λέμε ότι ανήκει στην κλάση 1 εάν η συνάρτηση διάκρισης της κλάσης 1 είναι μεγαλύτερη από την συνάρτηση διάκρισης της κλάσης 2. Για υπολογίσουμε αν έχει ταξινομηθεί κάτι λάθος επειδή γνωρίζουμε σε ποια κλάση ανήκουν και έτσι μπορούμε να εξετάσουμε εύκολα αν έχουν μπει στην σωστή.
Τρέχοντας το πρόγραμμα βλέπουμε πως έχουν ταξινομηθεί τα δεδομένα μας και πληροφορούμαστε ότι το σφάλμα είναι 30% δηλαδή στα 20 δείγματα 6 κατανεμήθηκαν λάθος

Απάντηση 3.2(3):

Ακολουθούμε πάλι την παρόμοια με παραπάνω διαδικασία όμως τώρα για 2 διαστάσεις. Εδώ τα boundaries βλέπουμε ότι προκύπτουν πιο πολύπλοκα αν λύσουμε ως προς χ1 θα πάρουμε :

```
boundary1 =

 0.401*x2 + 0.72 - 1.59e-7*(6.51e12*x2^2 - 8.43e13*x2 + 4.92e14)^(1/2)
 0.401*x2 + 1.59e-7*(6.51e12*x2^2 - 8.43e13*x2 + 4.92e14)^(1/2) + 0.72
```

Και ως προς χ2 :

```
boundary2 =

 0.00173*(2.6e9*x1^2 - 1.73e10*x1 + 2.82e10)^(1/2) - 87.1*x1 + 295.0
 295.0 - 87.1*x1 - 0.00173*(2.6e9*x1^2 - 1.73e10*x1 + 2.82e10)^(1/2)
```



```
wrong =

     9

error =

    0.4500
```

Επίσης βλέπουμε ότι 9 δείγματα ταξινομήθηκαν λανθασμένα και το σφάλμα αυξάνεται στα 45 %

Απάντηση 3.2(4):

Όπως και παραπάνω :

```
boundary1 =

 2.63*x3 - 0.219*x2 + 8.77e-8*(1.11e14*x2^2 - 5.44e14*x2*x3 - 1.15e15*x2 + 1.16e15*x3^2 + 2.27e15*x3 + 1.3e15)^(1/2) + 2.86
 2.63*x3 - 0.219*x2 - 8.77e-8*(1.11e14*x2^2 - 5.44e14*x2*x3 - 1.15e15*x2 + 1.16e15*x3^2 + 2.27e15*x3 + 1.3e15)^(1/2) + 2.86

boundary2 =

 0.272*x1 + 1.88*x3 + 4.36e-7*(6.93e12*x1^2 - 2.9e13*x1*x3 - 2.39e13*x1 + 5.44e12*x3^2 + 7.72e13*x3 + 1.04e14)^(1/2) + 4.7
 0.272*x1 + 1.88*x3 - 4.36e-7*(6.93e12*x1^2 - 2.9e13*x1*x3 - 2.39e13*x1 + 5.44e12*x3^2 + 7.72e13*x3 + 1.04e14)^(1/2) + 4.7

boundary3 =

 0.75*x2 - 1.3*x1 + 1.74e-7*(7.27e13*x1^2 - 5.76e13*x1*x2 - 4.22e13*x1 + 5.46e12*x2^2 + 9.43e13*x2 - 1.8e13)^(1/2) - 0.597
 0.75*x2 - 1.3*x1 - 1.74e-7*(7.27e13*x1^2 - 5.76e13*x1*x2 - 4.22e13*x1 + 5.46e12*x2^2 + 9.43e13*x2 - 1.8e13)^(1/2) - 0.597

wrong =

     3
```

Όσον αφορά το σφάλμα που προκύπτει εδώ , πετυχαίνουμε το καλύτερο σφάλμα μέχρι στιγμής με 15% δηλαδή μόνο 3 λάθη.

```
error =

    0.1500
```

Απάντηση 3.2(5):

Όπως είδα και παραπάνω είναι δυνατόν δυνατόν για ένα πεπερασμένο σύνολο των στοιχείων το εμπειρικό σφάλμα να είναι μεγαλύτερο όταν αυξάνεται η διάσταση των χαρακτηριστικών. Εδώ παρατηρήσαμε ότι όταν κατηγοριοποιούσαμε τα δεδομένα με βάση 2 χαρακτηριστικά πετύχαμε μεγαλύτερο σφάλμα . Αυτό μπορεί να οφείλεται στο ότι κάποιο χαρακτηριστικό να λαμβάνει κοντινές τιμές για τις 2 κλάσεις όπως εισάγοντας αυτό το χαρακτηριστικό ως επιπλέον κριτήριο ταξινόμησης να «μπερδέψουμε τον ταξινομιτή μας».

Απάντηση 3.2(6):

Εδώ θα έχουμε :

```
boundary1 =

 2.63*x3 - 0.219*x2 + 8.77e-8*(1.11e14*x2^2 - 5.44e14*x2*x3 - 1.15e15*x2 + 1.16e15*x3^2 + 2.27e15*x3 + 2.03e16)^(1/2) + 2.86
 2.63*x3 - 0.219*x2 - 8.77e-8*(1.11e14*x2^2 - 5.44e14*x2*x3 - 1.15e15*x2 + 1.16e15*x3^2 + 2.27e15*x3 + 2.03e16)^(1/2) + 2.86


boundary2 =

 0.272*x1 + 1.88*x3 - 4.36e-7*(6.93e12*x1^2 - 2.9e13*x1*x3 - 2.39e13*x1 + 5.44e12*x3^2 + 7.72e13*x3 - 8.5e14)^(1/2) + 4.7
 0.272*x1 + 1.88*x3 + 4.36e-7*(6.93e12*x1^2 - 2.9e13*x1*x3 - 2.39e13*x1 + 5.44e12*x3^2 + 7.72e13*x3 - 8.5e14)^(1/2) + 4.7


boundary3 =

 0.75*x2 - 1.3*x1 + 1.74e-7*(7.27e13*x1^2 - 5.76e13*x1*x2 - 4.22e13*x1 + 5.46e12*x2^2 + 9.43e13*x2 - 2.41e15)^(1/2) - 0.597
 0.75*x2 - 1.3*x1 - 1.74e-7*(7.27e13*x1^2 - 5.76e13*x1*x2 - 4.22e13*x1 + 5.46e12*x2^2 + 9.43e13*x2 - 2.41e15)^(1/2) - 0.597
```

Επίσης όπως ήταν αναμενόμενο, μεγαλώνοντας το prior της πρώτης κλάσης και μικραίνοντας το prior των υπόλοιπων 2 στην ουσία γιγαντώνουμε την κατανομή 1 και ο ταξινομιτής κατατάσσει πιο πολλά αντικείμενα στην κλάση 1 ακόμα και αν δεν ανήκουν σε αυτή, με αποτέλεσμα το σφάλμα να γίνεται πολύ μεγάλο έχουμε 11 λάθη στα 30 δείγμα ποσοστό που πλησιάζει το

```
wrong =

    11


error =

    0.4783
```

50 %

Άσκηση 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

$$p(\theta|D^0) = \begin{cases} Asin(\pi\theta), & 0 \leq \theta \leq 1 \\ 0, & \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$$

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

Απάντηση 3.3(1):

Εκμεταλλευόμαστε το στοιχείο :

$$\int_0^1 \mathrm{Asin}(\pi\theta) = 1 \Rightarrow A[-\cos(\pi\theta)]_0^1 = 1 \Rightarrow A = \frac{\pi}{2}$$

Απάντηση 3.3(2):

για να σχεδιάσουμε τα $p(\theta|D^1), p(\theta\big|D^5), p(\theta|D^{10})$ θα βασιστούμε στην αναδρομική σχέση:

$$p(\theta|D^n) = \frac{p(x_n|\theta)p(\theta|D^{n-1})}{\int p(x_n|\theta)p(\theta|D^{n-1})d\theta}$$

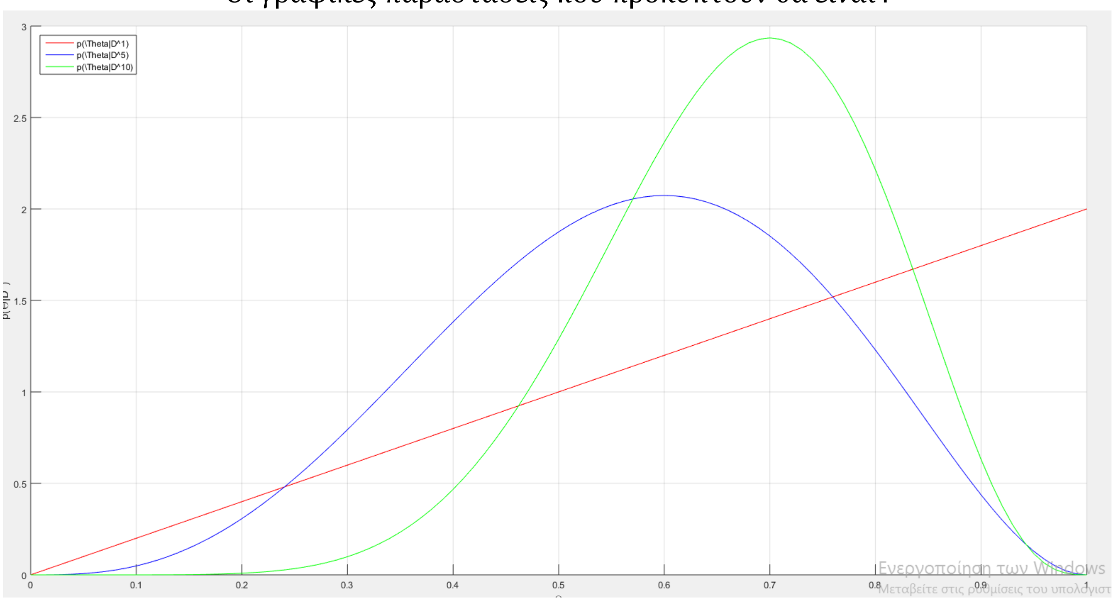
Επειδή το πρόβλημα έχει να κάνει με κορώνα γράμματα ξέρουμε ότι :

$$p(x_n|\theta) = \begin{Bmatrix} \theta\ , \kappa\varepsilon\varphi\acute{\alpha}\lambda\iota \\ (1-\theta)\gamma\rho\acute{\alpha}\mu\mu\alpha\tau\alpha \end{Bmatrix}$$

Οπότε η παραπάνω σχέση γίνεται :

$$p(\theta|D^n) = \frac{\theta^\kappa(1-\theta)^{N-\kappa}p(\theta|D^{n-1})}{\int \theta^\kappa(1-\theta)^{N-\kappa}p(\theta|D^{n-1})d\theta}$$

Οι γραφικές παραστάσεις που προκύπτουν θα είναι :



Απάντηση 3.3(3):

Εκτελώντας τον αλγόριθμο παίρνουμε ότι το $p(x = k|D^{10})$ θα είναι το $p(\theta^{*}|D^{10})$ όπου για $\theta = \theta^{*}$ η συνάρτηση μεγιστοποιείται (βέλτιστος εκτιμητής) . παίρνουμε λοιπόν ότι η συνάρτηση μεγιστοποιείται για θ=0.7 και παίρνει την τιμή 2.9351

Άσκηση 3.4

Έστω πρόβλημα με 3 κλάσεις (ω1, ω2, ω3) με δεδομένα από κανονικές κατανομές και με πιθανότητες :

- (μ_1, Σ_1) και $\mathrm{P}(\omega_1) = \mathrm{p1}$
- (μ_2, Σ_2) και $\mathrm{P}(\omega_2) = \mathrm{p2}$
- (μ_3, Σ_3) και $\mathrm{P}(\omega_3) = \mathrm{p3}$

Α. Δημιουργήστε 2 τυχαία δείγματα {X} και {X1}, 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 {X1} ως «σύνολο δοκιμής».

- $\mathrm{P}(\omega_1) = \mathrm{P}(\omega_2) = \mathrm{P}(\omega_3) = \frac{1}{3}$
- $\mu_1 = [0 \quad 0 \quad 0]^T \quad \mu_2 = [1 \quad 2 \quad 2]^T \quad \mu_3 = [3 \quad 3 \quad 4]^T$
- $\Sigma = \Sigma_1 = \Sigma_2 = \Sigma_3 = \begin{bmatrix} 0.8 & 0.2 & 0.1 \\ 0.2 & 0.8 & 0.2 \\ 0.1 & 0.2 & 0.8 \end{bmatrix}$

Β. Χρησιμοποιώντας τις παραπάνω παραμέτρους, επιλέξτε τον κατάλληλο ταξινομητή και ταξινομήστε το σύνολο {X1} και υπολογίστε την πιθανότητα λάθους.

Γ. Χρησιμοποιώντας το σύνολο {X} και την μέθοδο της μεγίστης πιθανοφάνειας, εκτιμήστε τα μ1, μ2, μ3, και το Σ, και μετά επιλέξτε τον κατάλληλο ταξινομητή και ταξινομήστε το σύνολο {X1} και υπολογίστε την πιθανότητα λάθους.

Δ. Χρησιμοποιώντας :

- $\mathrm{P}(\omega_1) = \frac{1}{6} \quad \mathrm{P}(\omega_2) = \frac{1}{6} \quad \mathrm{P}(\omega_3) = \frac{2}{3}$
- $\mu_1 = [0 \quad 0 \quad 0]^T \quad \mu_2 = [1 \quad 2 \quad 2]^T \quad \mu_3 = [3 \quad 3 \quad 4]^T$
- $\Sigma_1 = \begin{bmatrix} 0.8 & 0.2 & 0.1 \\ 0.2 & 0.8 & 0.2 \\ 0.1 & 0.2 & 0.8 \end{bmatrix} \quad \Sigma_2 = \begin{bmatrix} 0.6 & 0.2 & 0.01 \\ 0.2 & 0.8 & 0.01 \\ 0.01 & 0.01 & 0.6 \end{bmatrix} \quad \Sigma_3 = \begin{bmatrix} 0.6 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.1 \\ 0.1 & 0.1 & 0.6 \end{bmatrix}$

επαναλάβετε τα Α, Β και Γ. Σχολιάστε τα αποτελέσματα σας

Απάντηση 3.4(Α):

Δημιουργούμε τα δύο σύνολα χρησιμοποιώντας την εντολή mvnrnd με την οποία παράγουμε σύνολα πολλών μεταβλητών που ακολουθούν γκαουσιανή κατανομή με ορισμένα μ και Σ (το σύνολο 3 έχει 1 δείγμα παραπάνω από τα άλλα για να βγαίνουν 1000 και 10000 τα νούμερα) . Στις μεταβλητές mean1, mean2,mean3, cov, prior έχουμε δώσει τις κατάλληλες τιμές όπως ορίζει η εκφώνηση για να βγουν οι επιθυμητές κατανομές.

```
train=[mvnrnd(mean1,cov,floor(traindata*prior));
      mvnrnd(mean2,cov,floor(traindata*prior));
      mvnrnd(mean3,cov,ceil(traindata*prior))];
%train dataset
x=train(randperm(size(train,1)),:);

%we use mvrnd to generate multivariabe data that follow
%gaussian distribution
test=[mvnrnd(mean1,cov,floor(testdata*prior));
      mvnrnd(mean2,cov,floor(testdata*prior));
      mvnrnd(mean3,cov,ceil(testdata*prior))];
%test dataset
x1=test(randperm(size(test,1)),:);
```

Απάντηση 3.4(Β):

Παρατηρούμε ότι ο πίνακας συν διασποράς είναι ένας και για τις 3 κλάσεις. Αυτό μας επιτρέπει να ταξινομήσουμε τα δεδομένα χρησιμοποιώντας τον mahalanobis ταξινομιτή για να γλυτώσουμε υπολογιστική ισχύ. Θα ήταν ακόμα πιο βολικό να χρησιμοποιούσαμε ευκλείδεια απόσταση αλλά εδώ δεν έχουμε αυτή την δυνατότητα γιατί ο Σ δεν είναι της μορφής σ*Ι. Για να βρούμε το σφάλμα στην ουσία μετράμε πόσα αντικείμενα τοποθετήθηκαν σε κάθε κλάση και τα συγκρίναμε με τον αριθμό αντικειμένων που έπρεπε να τοποθετηθούν (είχα την αμφιβολία ότι αυτή η τεχνική μπορεί να μην έβγαζε σωστά αποτελέσματα ωστόσο όταν το δοκίμασα και χωρίς να ανακατέψω το dataset που τεστάρω χρησιμοποιώντας rng ‘default’ έβγαλε το ίδιο αποτέλεσμα που μου έδινε και το μπερδεμένο dataset οπότε δουλεύει). Τελικά το σφάλμα που προκύπτει είναι για μία τυχαία φορά που έτρεξα το πρόγραμμα 1,2% (me rng ‘default’ δίνει 1%)

```
error1 =

     5

error2 =

     6

error3 =

     1

totalerror =

    0.0120
```

Απάντηση 3.4(Γ):

Σε αυτό το ερώτημα βρισκω βασιζόμενος στα data εκπαίδευσης, υποθέτοντας ότι δεν ξέρω τι κατανομές ακολουθούν τα δεδομένα, εκτιμητές μέσης τιμής μ και συνδιασποράς Σ. Εφόσον έχω κανονική κατανομή για να βρώ αυτά τα μεγέθη βασίστηκα στους τύπους :

$$\hat{\mu} = \frac{1}{n}\sum_{k=1}^{n} x_k \ \ \kappa\alpha\iota \ C = \frac{1}{n-1}\sum_{k=1}^{n} (\boldsymbol{x}_k - \boldsymbol{\mu})(\boldsymbol{x}_k - \widehat{\boldsymbol{\mu}})^{\boldsymbol{T}}$$

τα αποτελέσματα που βρήκαμε είναι :

```
estmean1 =

    0.0172    0.0027   -0.0090


estmean2 =

   -0.0088    1.9977    2.0020


estmean3 =

    2.9976    2.9842    4.0042


estcov =

    0.7951    0.2102    0.1062
    0.2102    0.7867    0.2021
    0.1062    0.2021    0.7965
```

όπου :

- estmean1: $\hat{\boldsymbol{\mu}}_1$
- estmean1: $\hat{\boldsymbol{\mu}}_2$
- estmean1: $\hat{\boldsymbol{\mu}}_3$
- estcov: C

Χρησιμοποιώντας έπειτα αυτόν τον ταξινομητή για να ταξινομήσουμε τα δεδομένα του συνόλου test δηλαδή του X1 βρίσκουμε 3% σφάλμα, γεγονός που δείχνει ότι και αυτός ο τρόπος ταξινόμισης λειτουργεί πολύ καλά.

```
totalerror =

    0.0300
```


Απάντηση 3.4(Δ):

Αρχικά φτιάχνουμε τα νέα σύνολα με τα καινούργια δεδομένα όπως στο πρώτο ερώτημα

```
prior1=1/6;
prior2=1/6;
prior3=2/3;

mean1=[0 0 0]';
mean2=[0 2 2]';
mean3=[3 3 4]';

cov1=[0.8 0.2 0.1;0.2 0.8 0.2;0.1 0.2 0.8];
cov2=[0.6 0.2 0.01;0.2 0.8 0.01;0.01 0.01 0.6];
cov3=[0.6 0.1 0.1;0.1 0.6 0.1;0.1 0.1 0.6];

train=[mvnrnd(mean1,cov1,floor(traindata*prior1));
        mvnrnd(mean2,cov2,floor(traindata*prior2));
        mvnrnd(mean3,cov3,ceil(traindata*prior3))];
%train dataset
x=train(randperm(size(train,1)),:);


test=[mvnrnd(mean1,cov1,floor(testdata*prior1));
      mvnrnd(mean2,cov2,floor(testdata*prior2));
      mvnrnd(mean3,cov3,ceil(testdata*prior3))];
%test dataset
x1=test(randperm(size(test,1)),:);
```

τώρα για τον υπολογισμό του λάθους δεν μπορούμε να χρησιμοποιήσουμε ούτε την απόσταση mahalanobis αφού πλέον οι πίνακες συνδιασποράς και οι prior πιθανότητες είναι διαφορετικάς μεταξύ τους (αντίστοιχα). Έτσι χρησιμοποιούμε την συνάρτηση διάκρισης που δημιουργήσαμε στο ερώτημα 3.1.α και κατατάσουμε το κάθε δείγμα στην τάξη όπου η συνάρτηση διάκρισης του έχει μεγαλύτερη τιμή. Τρέχοντας το πρόγραμμα προκύπτει ένα σφάλμα 6.7%.

```
totalerror =

    0.0678
```

Τώρα εφαρμόζουμε την μέθοδο μέγιστης πιθανοφάνειας . Βρίσκουμε τις νέες τιμές μ και Σ όπως περιγράφεται παραπάνω και έχουμε :

```
estmean1 =

   -0.0298   -0.0073   -0.0149


estmean2 =

    0.0246    1.9752    1.9859


estmean3 =

    3.0035    2.9859    4.0064
```

```
estcov1 =

    0.8496    0.1985    0.0762
    0.1985    0.8006    0.1849
    0.0762    0.1849    0.8292


estcov2 =

    0.6487    0.2060    0.0251
    0.2060    0.8314    0.0222
    0.0251    0.0222    0.5998


estcov3 =

    0.6001    0.1110    0.1143
    0.1110    0.6042    0.1081
    0.1143    0.1081    0.5984
```

- Όπου estmean1: $\hat{\boldsymbol{\mu}}_1$
- estmean1: $\hat{\boldsymbol{\mu}}_2$
- estmean1: $\hat{\boldsymbol{\mu}}_3$
- estcov1: C1
- estcov2: C2
- estcov3: C3