ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΤΥΠΩΝ

ΠΡΟΑΙΡΕΤΙΚΗ ΕΡΓΑΣΙΑ 1

Σταματόπουλος Βασίλειος – 1115201400188

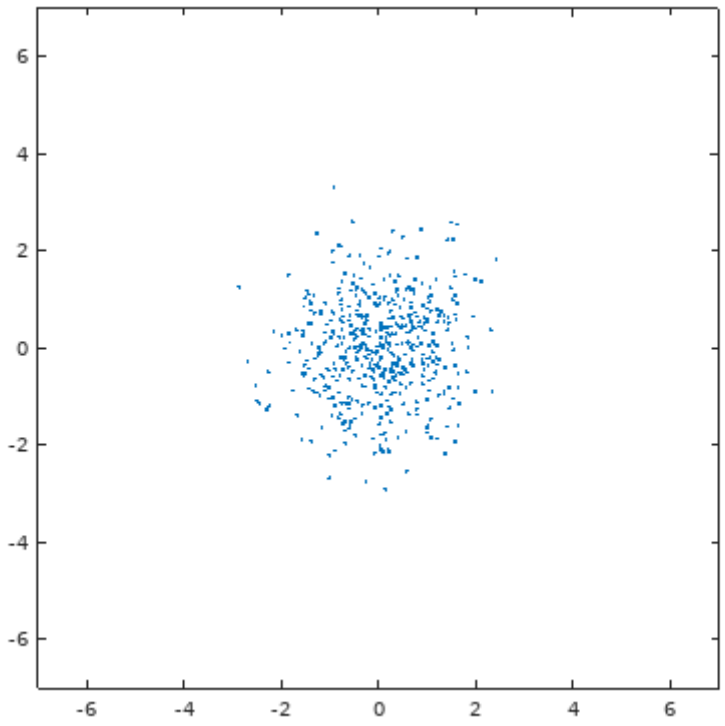
1.

Στην άσκηση αυτή χρησιμοποιήθηκε ο κώδικας που δίνεται, για κάθε ένα από τα παραδείγματα.

Στη συνέχεια αναλύονται τα σχήματα που προκύπτουν για κάθε ομάδα. Αφού η μέση τιμή είναι μηδενική, τα δεδομένα θα κατανομούνται κανονικά γύρω από το (0,0). Συνεπώς, οι διαφορές κάθε σχήματος είναι ανάλογες τις συνδιασποράς που χρησιμοποιείται. Οι τιμές των σ, δίνουν στο γράφημα την πυκνότητα και την κατεύθυνση των στοιχείων. Το ιδιοδιάνυσμα που αντιστοιχεί στην μεγαλύτερη ιδιοτιμή, ορίζει την κατεύθυνση κατά την οποία τα στοιχεία παρουσιάζουν τη μέγιστη διασπορά. Το αμέσως επόμενο, είναι κάθετο στο πρώτο και αντιστοιχεί στη κατεύθυνση στην οποία τα δεδομένα παρουσιάζουν την αμέσως επόμενη μέγιστη διασπορά. **Έστω,** $\boldsymbol{S} = \begin{bmatrix} \boldsymbol{\alpha} & \boldsymbol{\beta} \\ \boldsymbol{\gamma} & \boldsymbol{\delta} \end{bmatrix}$

α)

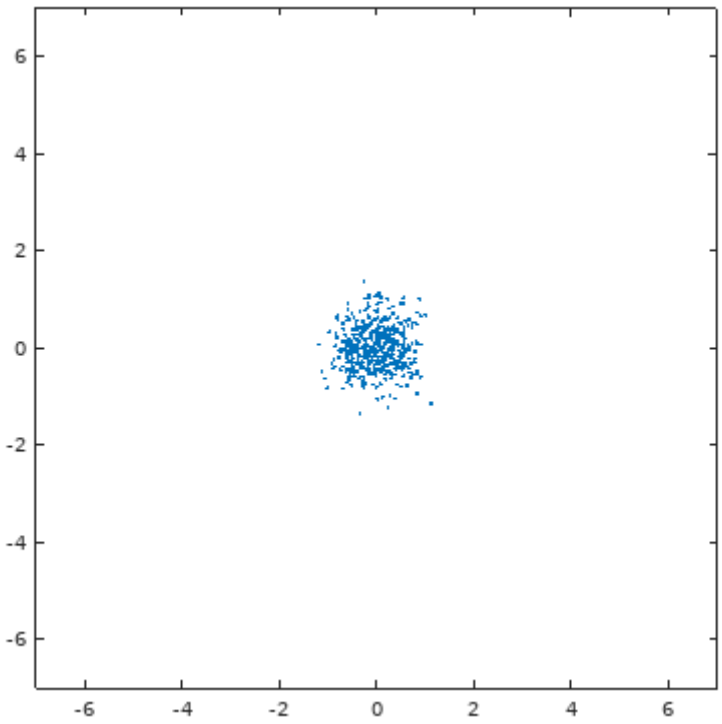
$$\boldsymbol{m} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \; \boldsymbol{S} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$$



Στην περίπωση αυτή, καθώς α = δ =1 και > β = γ = 0 δηλαδή ο πίνακας είναι διαγώνιος με ίσα διαγώνια στοιχεία, τα δεδομένα έχουν ομοιόμορφη κατανομή.

β)

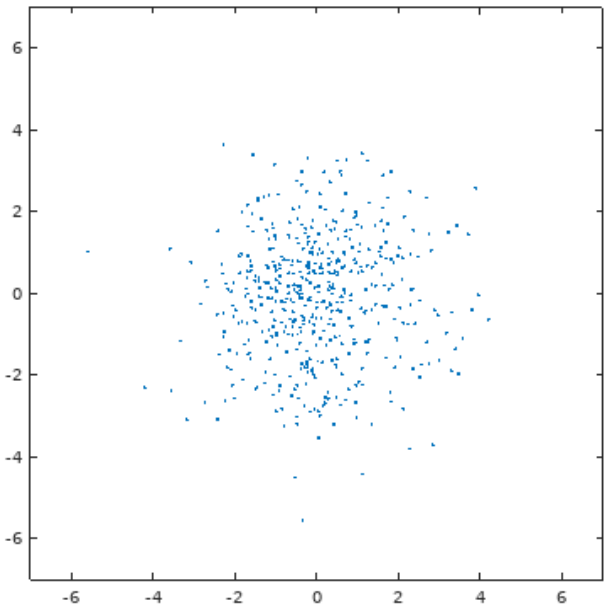
$$\boldsymbol{m} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \boldsymbol{S} = \begin{bmatrix} \mathbf{0.2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0.2} \end{bmatrix}$$



Η περίπτωση αυτή είναι αντίστοιχη με τη προηγούμενη, καθώς όμως το α = δ = 0.2 < 1, τα δεδομένα βρίσκονται πιο κοντά στο (0,0).

γ)

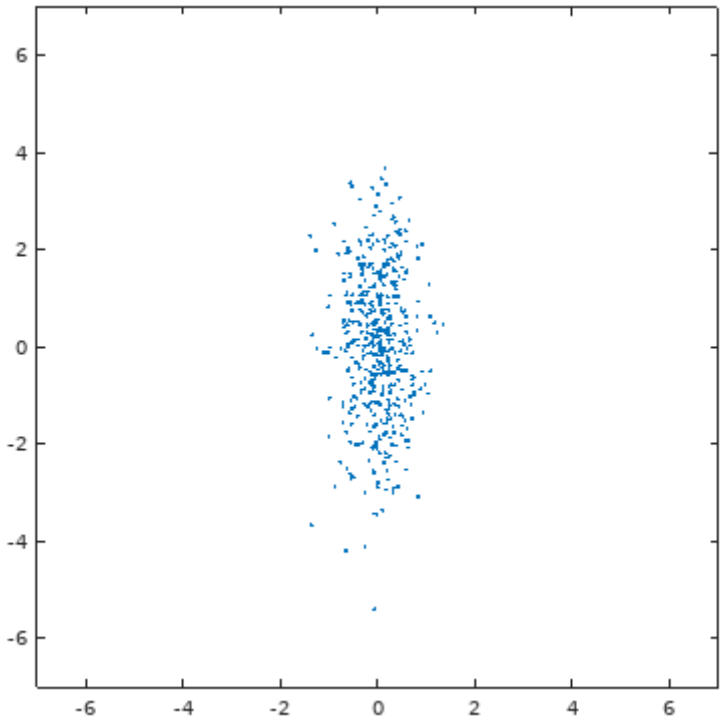
$$\boldsymbol{m} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \boldsymbol{S} = \begin{bmatrix} \mathbf{2} & \mathbf{0} \\ \mathbf{0} & \mathbf{2} \end{bmatrix}$$



΄Ιδια περίπτωση με τις δύο προηγούμενες αλλά τα δεδομένα διασκορπίζονται στον χώρο περισσότερο καθώς α = δ = 2.

δ)

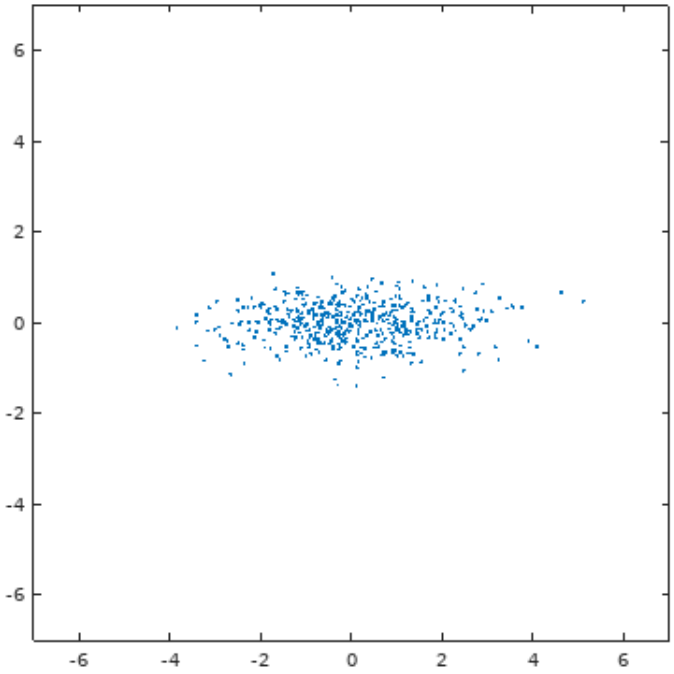
$$\boldsymbol{m} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \boldsymbol{S} = \begin{bmatrix} \mathbf{0.2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0.2} \end{bmatrix}$$



Η περίπτωση αυτή είναι διαγώνιος πίνακας με α<<δ που σημαίνει πως η διασπορά των δεδομένων στον οριζόντιο άξονα είναι αρκετά μικρότερη απο εκείνη στον κάθετο.

ε)

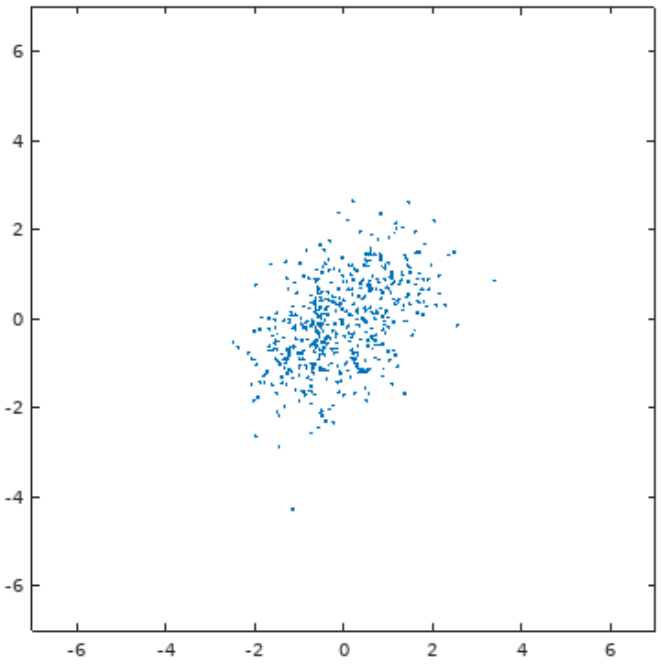
$$\boldsymbol{m} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \boldsymbol{S} = \begin{bmatrix} \mathbf{2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0.2} \end{bmatrix}$$



Παρόμοια περίπτωση με την προηγούμενη αλλά τώρα ισχύει α>>δ. Συνεπώς οι διασπορές των δεδομένων στον χώρο είναι αντεστραμμένες.

στ)

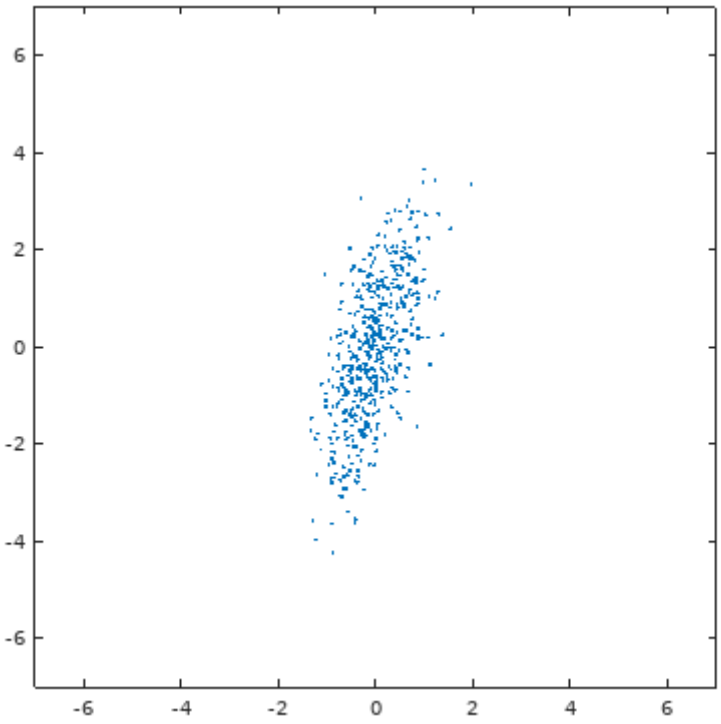
$$\boldsymbol{m} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \boldsymbol{S} = \begin{bmatrix} \mathbf{1} & \mathbf{0.5} \\ \mathbf{0.5} & \mathbf{1} \end{bmatrix}$$



Καθώς τώρα το β και το γ έχουν τιμή, τα δεδομένα έχουν κλίση στον χώρο.

ζ)

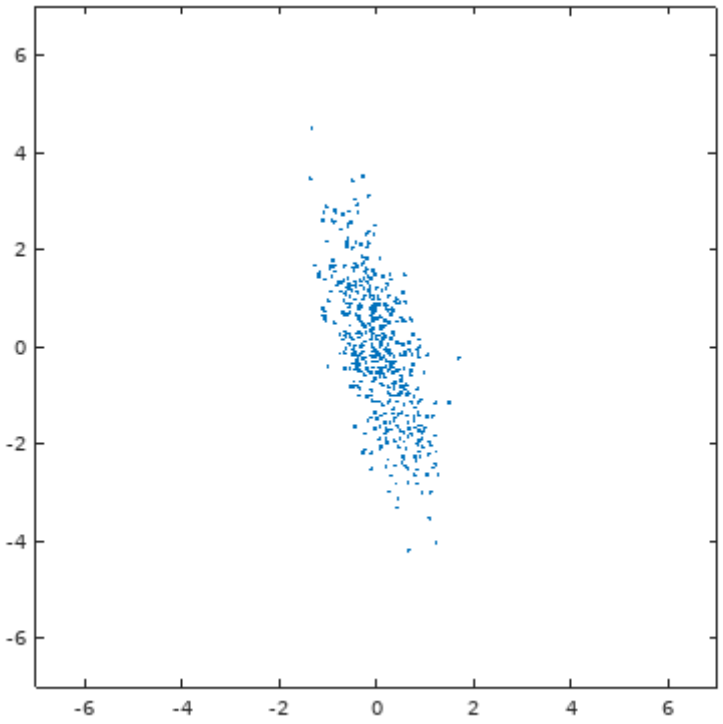
$$\boldsymbol{m} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \; \boldsymbol{S} = \begin{bmatrix} \mathbf{0.3} & \mathbf{0.5} \\ \mathbf{0.5} & \mathbf{0.3} \end{bmatrix}$$



Τώρα η κλίση διατηρείται, αλλά αλλάζει η πυκνότητα στον κάθετο και οριζόντιο άξονα.

η)

$$\boldsymbol{m} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \; \boldsymbol{S} = \begin{bmatrix} \mathbf{0.3} & \mathbf{-0.5} \\ \mathbf{-0.5} & \mathbf{0.3} \end{bmatrix}$$



Η αντίστροφη περίπτωση της προηγούμενης. (Η κλίση είναι από την άλλη μεριά).

2.

Έχουμε m1 = 1, m2 = 4, s =1.

α)

Ισχύει ότι η a priori πιθανότητες των κλάσεων είναι p1 = p2 = 0.5. Θα υπολογίσουμε τις a-posteriori πιθανότητες P1, P2.

$$\boldsymbol{P1} = \frac{\mathbf{1}}{\sqrt{\mathbf{2\pi}}} * \boldsymbol{exp}\left(-\frac{(\mathbf{1.7-1})^{\mathbf{2}}}{\mathbf{2}}\right), \qquad \boldsymbol{P2} = \frac{\mathbf{1}}{\sqrt{\mathbf{2\pi}}} * \boldsymbol{exp}(-\frac{(\mathbf{1.7-4})^{\mathbf{2}}}{\mathbf{2}})$$

Καθώς p1 = p2, η κλάση που θα επιλεχθεί έχει να κάνει μονάχα μόνο με τις a-posteriori πιθανότητες.

Έστω, $\boldsymbol{P1} > \boldsymbol{P2} \rightarrow \boldsymbol{ln}(\boldsymbol{P1}) > \boldsymbol{ln}(\boldsymbol{P2}) \rightarrow \left(-\frac{\mathbf{0.49}}{\mathbf{2}}\right) > \left(-\frac{\mathbf{5.29}}{\mathbf{2}}\right) \rightarrow \mathbf{5.29} > \mathbf{0.49}$

Η υπόθεση ισχύει, συνεπώς το x = 1.7 θα κατηγοριοποιηθεί στην πρώτη κλάση.

β)

Στο ερώτημα αυτό, χρησιμοποιήθηκε ο δωθέντας κώδικας με την προσθήκη του παρακάτω κομματιού, για τον υπολογισμό του λάθους.

```
...

error = nnz(bayes_res);  %Ypologismos ton mi-midenikon arithmon ston pinaka
fprintf("Error is %d percent\n", (error*100)/N);


...
```

Ο κώδικας αυτός υπολογίζει τον αριθμό των μη μηδενικών αριθμών στον πίνακα σύγκρισης και βάση αυτού δίνει το ποσοστό του σφάλματος, που για N = 100 ήταν αρκετά καλό (3%).

γ)

Για N=1000 και N=10000 το σφάλμα κυμαίνεται στις ίδιες τιμές (2.9 και 3.19 αντίστοιχα)

δ)

Η πιθανότητα λάθους για τη κάθε μία πιθανότητα είναι το ποσοστό των 1 στα δύο μισά του πίνακα σύγκρισης. Πιο συγκεκριμένα, για την πρώτη κλάση και για N=100 είναι 6% και 8% αντίστοιχα. Ενώ για N=1000 και N = 10000, τα ποσοστά αυτά δεν αλλάζουν σημαντικά.(N = 1000: e1 = 5.8%, e2 = 8.2%. N=10000: e1 = 6.38%, e2 = 7.06%.)

3.

Για τους υπολογισμούς χρησιμοποιήθηκε ο δοθέντας κώδικας, σε συνδυασμό με το παρακάτω κομμάτι κώδικα.

```
t=[ones(1,N/2) 2*ones(1,N/2)];
out_eucl = euclidean(m,X);
eucl_res = (t~=out_eucl);
out_mahal = mahalanobis(m,S,X);
mahal_res = (t~=out_mahal);
error1 = nnz(eucl_res);
error2 = nnz(mahal_res);
fprintf("Error euclidean is %d percent\n", (error1*100)/N);
fprintf("Error mahalanobis is %d percent\n", (error2*100)/N);
```

β)

Και πήραμε τα αποτελέσματα που περιγράφονται στον πίνακα:

	N =100	**N = 1000**	**N = 10000**
Euclidean	23%	25.1%	25.51%
Mahalanobis	23%	22.2%	21.04%

γ)

Όπως φαίνεται, για χαμηλό αριθμό δεδομένων, οι δύο αλγόριθμοι δίνουν παρόμοια αποτελέσματα. Όσο αυξάνει όμως το N, φαίνεται πιο αισθητά η υπεροχή του αλγορίθμου Mahalanobis.

δ)

Όσον αφορά την απόδοση του ταξινομητή Bayes, χρησιμοποιήθηκε το παρακάτω κομμάτι κώδικα για τον υπολογισμό της. Τα αποτελέσματα όπως ήταν αναμενόμενο, ήταν ίδια με του αλγόριθμου ελάχιστης απόστασης Mahalanobis. Αυτό γιατί, οι δύο αλγόριθμοι ταυτίζονται σε αυτό το παράδειγμα.

```
[l,c]=size(m);
for i=1:N
        variable = X(:,i);
         for j=1:c
                mean = m(:,j);
                exp_component = (variable - mean)'*(pinv(S))*(variable - mean);
                val(j) = (1/sqrt(2*pi))*exp(-1/2*exp_component);
        end
         [num,z(i)]=max(val);
end

out_bayes = z;
bayes_res = (t~=out_bayes);
error3 = nnz(bayes_res);
fprintf("Error bayes is %d percent\n", (error3*100)/N);
```

4.

γ)

Αφού ακολουθήσουμε τα βήματα του α και του β, με την υλοποίηση του γ, παίρνουμε m1 = 0.83590 και m2 = 2.6196.

δ)

Οι a priori πιθανότητες των κλάσεων είναι, p1 = p2 =0.5. Με αυτά τα στοιχεία ο ταξινομητής Bayes δίνει πιθανότητα λάθους 0.163, που είναι ίση με την πιθανότητα λάθους του αλγορίθμου μέγιστης πιθανοφάνειας.

ε)

Για N1 = N2 = 500, ο ταξινομητής Bayes δίνει ίδια πιθανότητα λάθους, ενώ η ML, δίνει p = 0.164.

ζ)

Δεν θα είχε νόημα ο υπολογισμός των μέσων τιμών με τη βοήθεια της μεθόδου MAP, καθώς ισχύει ότι $\mu_{MAP} \cong \mu_{ML}$

5.

Για μικρό αριθμό δεδομένων (N=40), ο ταξινομητής Bayes δίνει καλύτερα αποτελέσματα από τα παράθυρα Parzen ($p_{e1} = 0.163$ $p_{e2} = 0.227$ αντίστοιχα). Με την αύξηση των δεδομένων κάθε κλάσης στα 200, η πιθανότητα λάθους του Bayes δεν αλλάζει, ενώ των Parzen windows μειώνεται σημαντικά ($p_{e2} = 0.166$). Γενικότερα, ο δεύτερος από τους αλγορίθμους αυτούς δίνει καλύτερα αποτελέσματα εάν έχουμε μεγαλύτερο σύνολο δεδομένων, καθώς όσο περισσότερα τα δεδομένα, τόσα περισσότερα θα είναι και τα τεμνόμενα παράθυρα, αρά θα μεγαλώσει και η ακρίβεια της ταξινόμησης.