Στατιστική Μοντελοποίηση και Αναγνώριση Προτύπων
1η Σειρά Ασκήσεων

Μαρία Μαραγκάκη Α.Μ.: 2015030153

11/04/20

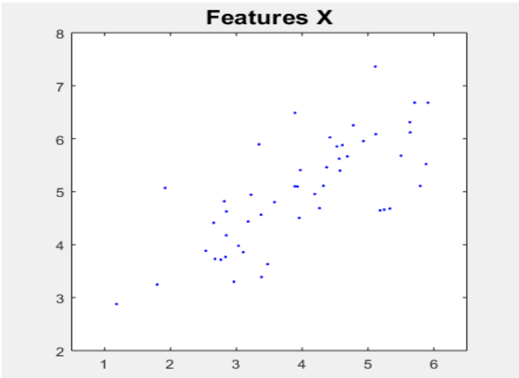
Θέμα 1: Principal Component Analysis (PCA)

Σκοπός της συγκεκριμένης άσκησης ήταν η εξοικείωση με την μέθοδο Principal Component Analysis, η οποία χρησιμοποιείται για την μείωση των διαστάσεων των δεδομένων.

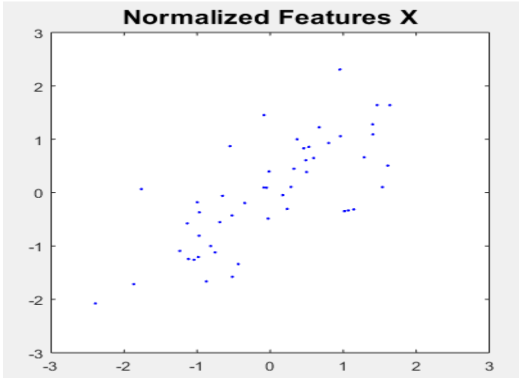
Μέρος 1
Στο πρώτο μέρος της άσκησης ζητήθηκε η εφαρμογή της μεθόδου PCA σε ένα μικρό σύνολο δεδομένων δύο διαστάσεων.

1. Σε πρώτη φάση, μέσω της εντολής *load('ex1_1_data.mat')* έγινε φόρτωση των δεδομένων στην μεταβλητή X και απεικόνηση αυτών στον χώρο. Το αποτέλεσμα φαίνεται στο παρακάτω figure.

2. Έπειτα πραγματοποιήθηκε κανονικοποίηση των δεδομένων με μέση τιμή μηδέν και διασπορά 1. Η κανονικοποίηση των δεδομένων έγινε με σκοπό την μείωση του εύρους τιμών των δειγμάτων, ορισμένα δείγματα μπορεί να έχουν πολύ μεγαλύτερες τιμές απο άλλα με αποτέλεσμα να επηρεάζουν το γενικό αποτέλεσμα, οπότε μέσω της κανονικοποίησης περιορίζονται οι τιμές των δειγμάτων εντός κάποιων ορίων. Το αποτέλσμα της κανονικοποίησης φαίνονται στο παρακάτω figure.



(a) Regular Features



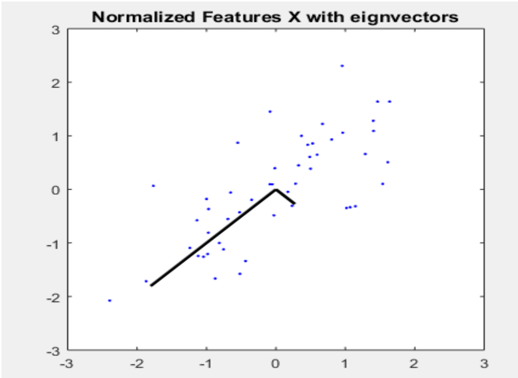
(b) Normalized Features

Όπως ήταν αναμενόμενο παρατηρείται η μείωση του εύρους τιμών σε σχέση με τα προτότυπα δεδομένα.

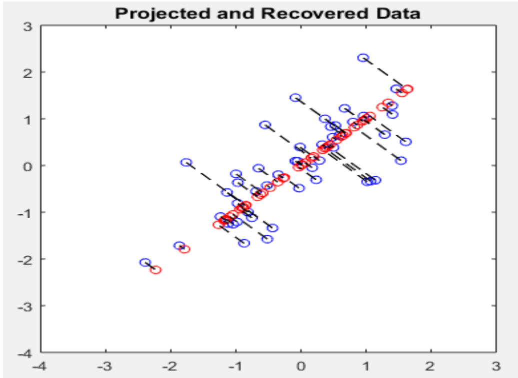
3. Εν συνεχεία, ζητήθηκε η απεικόνηση των κύριων συνιστωσών του αλγορίθμου PCA και ο σχεδιασμός μαζί με τα αρχικά δείγματα. Για τον υπολογισμό του πίνακα συνδιασποράς χρησιμοποιήθηκε η εξίσωσση $\Sigma = 1/m * X^T * X$. Για τον υπολογισμό των συνιστωσών ακολουθήθηκε η εξής διαδιακασία υπολογίστηκε ο πίνακας συνδιασποράς σύμφωνα με τον παραπάνω τύπο, στη συνέχεια χρησιμοποιήθηκε η συνάρτηση *eig* του Matlab η οποία επιστρέφει έναν διαγώνιο πίνακα με τα eigenvalues καθώς και έναν πίνακα του οποίου οι στήλες είναι τα eigenvectors. Έπειτα έγινε ταξινόμηση των των eigenvalues κατά φθήνουσα σειρά και επιλέχθηκαν τα κατάλληλα eigenvectors. Το αποτέλεσμα φαίνεται στο παρακάτω figure.

4. Ακόμη, υπολογίστηκε η συνεισφορά που έχει κάθε κύρια συνιστώσα στην συνολική διακύμανση έτσι ώστε να πραγματοποιηθεί μείωση των διαστάσεων απο 2D σε 1D. Συμπληρώθηκε κατάλληλα ο κώδικας στην συνάρτηση projectData έτσι ώστε να επιστρέφονται τα προβαλλόμενα δεδομένα. Η προβολή έγινε σύμφωνα με τον τύπο $z_i = U^T * X$ όπου U είναι οι K κύριες συνιστώσες που έχουν επιλεγεί για την προβολή.

5. Τέλος, πραγματοποιήθηκε ανάκτηση των προβαλλόμενων δεδομένων μέσω του τύπου $x_{rec} = U * z$ με σκοπό την μείωση των διαστάσεων όπως αναφέρθηκε αρχικά. Τα αποτελέσματα φαίνονται στο παρακάτω figure.



(a) Normalized Features with the eigenvectors



(b) Projected and recovered data

Στην αριστερή εικόνα παρατηρούνται τα eigenvectors με τη μεγαλύτερη διακύμανση. Στην δεξιά εικόνα παρατείται η προβολή των δεδομένων, η μείωση των διαστάσεων απο 2D σε 1D και η ανάκτηση των δεδομένων (κόκκινες κουκίδες). Όπως παρατηρείται έγινε προβολή των δεδομένων κατά την διεύθυνση του eigenvector με τη μεγαλύτερη διακύμανση για την βέλτηστη προβολή των δεδομένων.

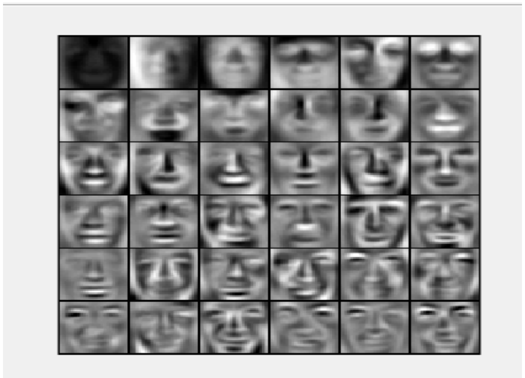
Μέρος 2

Στο δεύτερο μέρος της άσκησης ζητήθηκε η επανάληψη της παραπάνω διαδικασίας σε δεδομένα 5000 προσώπων που δόθηκαν.

1. Αρχικά, πραγματοποιήθηκε φόρτωση των δεδομένων στην μεταβλητή X μέσω της εντολής *load('ex1_1_faces.mat')* και απεικόνηση των πρώτων 100 προσώπων με τη χρήση της συνάρτησης *displayData()*. Το αποτέλεσμα φαίνεται στο παρακάτω figure.

2. Έπειτα, ζητήθηκε η εφαρμογή standardization καθώς και η εφαρμογή του αλγορίθμου PCA μέσω των συναρτήσεων που δημιουργήθηκαν παραπάνω. Κατόπιν, έγινε σχεδιασμός των πρώτων 36 κύριων συνιστωσών του αλγορίθμου. Το αποτέλεσμα φαίνεται στο παρακάτω figure. Παρατηρείται οτι χρησιμοποιόντας τις πρώτες 36 κύριες συνιστώσες η προβολή των εικόνων είναι αρκετά θολή με αποτέλεσμα να μην μπορεί να γίνει σωστή ανάκτηση.

(a) Displaying of the first 100 faces



(b) Displaying the first 36 eignvectors

3. Στη συνέχεια, πραγματοποιήθηκε μείωση των διαστάσεων μέσω της συνάρτησης που δημιουργήθηκε παραπάνω χρησιμοποιόντα τις 100 πρώτες κύριες συνιστώσες.

4. Τέλος, σχεσιάστηκαν τα δείγματα μειωμένης διάστασης αφού πρώτα πραγματοποιήθηκε ανάκτηση. Η διαδικασία επαναλήφθηκε για αριθμό συνιστωσών 10, 20 ,200 και τα αποτελέσματα φαίνονται παρακάτω.
Παρατηρείται ότι αυξάνονται τον αριθμό των κύριων συνιστωσών του αλγορίθμου αυξάνεται και η ποιότητα ανάκτησης των προσώπων όπως ήταν αναμενόμενο λόγω της σημαντικότητας των κύριων συνιστωσών στην ανάκτηση του δείγματος των εικόνων.

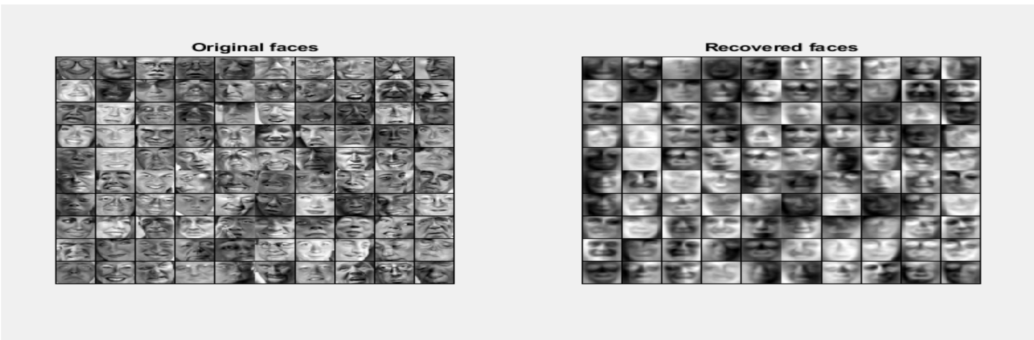


Figure 4: Displaying the first 10 eignvectors

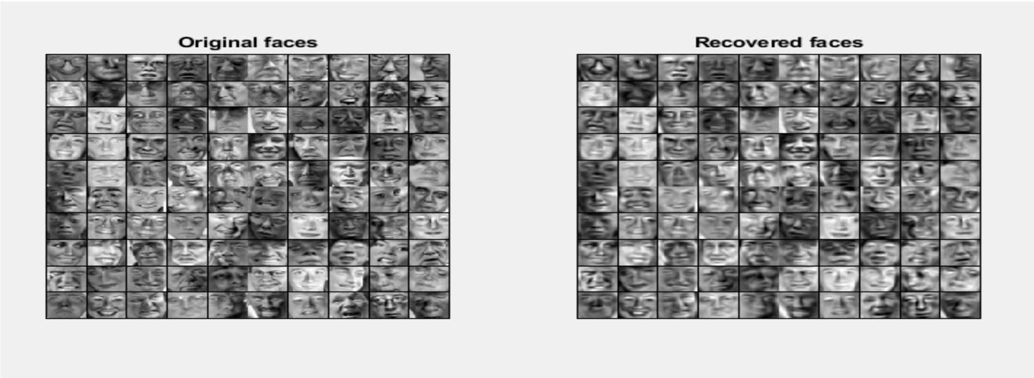


Figure 5: Displaying the first 100 eignvectors

Figure 6: Displaying the first 200 eignvectors

Θέμα 2: Σχεδιάστε ένα ταξινομητή LDA (Linear Discriminant Analysis)

$$S_w^{-1} = 1/2 * (S1 + S2) = 1/2 * \begin{bmatrix} 11 & 9 \\ 9 & 11 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 1/2 * \begin{bmatrix} 13 & 9 \\ 9 & 13 \end{bmatrix} = \begin{bmatrix} 6.5 & 4.5 \\ 4.5 & 6.5 \end{bmatrix}$$

$$S_w^{-1} = 1/(det(Sw)) * \begin{bmatrix} 6.5 & -4.5 \\ -4.5 & 6.5 \end{bmatrix}$$

$$det(Sw) = \begin{vmatrix} 6.5 & 4.5 \\ 4.5 & 6.5 \end{vmatrix} = 6.5^2 - 4.5^2 == 42.25 - 20.25 = 22$$

$$S_w^{-1} = 1/22 * \begin{bmatrix} 6.5 & -4.5 \\ -4.5 & 6.5 \end{bmatrix} = \begin{bmatrix} 0.29 & -0.2 \\ -0.2 & 0.29 \end{bmatrix}$$

$$w = S_w^{-1} * (m_1 - m_2) = \begin{bmatrix} 0.29 & -0.2 \\ -0.2 & 0.29 \end{bmatrix} * (\begin{bmatrix} -5 \\ 5 \end{bmatrix} - \begin{bmatrix} 10 \\ 15 \end{bmatrix}) = \begin{bmatrix} 0.29 & -0.2 \\ -0.2 & 0.29 \end{bmatrix} * \begin{bmatrix} -15 \\ -10 \end{bmatrix} = \begin{bmatrix} ((-15) * 0.29) + ((-0.2) * (-10)) \\ ((-15) * (-0.2)) + (0.29 * (-10)) \end{bmatrix} =$$

$$= \begin{bmatrix} -2.35 \\ 0.1 \end{bmatrix}$$

Θέμα 3: Linear Discriminant Analysis(LDA) vs PCA

Σκοπός της άσκηση ήταν η εφαρμογή του αλγορίθμου LDA και η σύγκριση με τον αλγόριθμο PCA.

Μέρος 1

1. Σε πρώτη φάση μέσω της εντολής *load('ex1_3_data1.mat')* έγινε φόρτωση των δεδομένων στην μεταβλητή X.

2. Στη συνέχεια πραγματοποιήθηκε standardization των δεδομένων και το αποτέλεσμα φαίνεται στο παρακάτω figure.

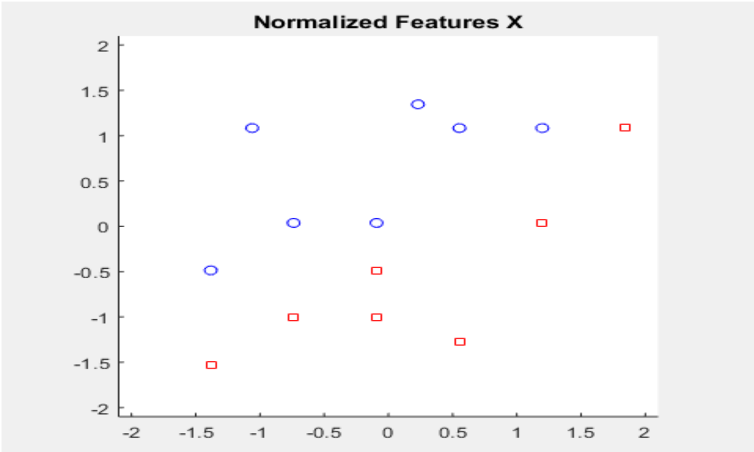


Figure 7: Normalized Features

3. Έπειτα, προστέθηκε ο κατάλληλος κώδικας στη συνάρτηση *fisherLinearDiscriminant* ώστε να εφαρμοστέι ο αλγόριθμος. Για την υλποποίση του αλγορίθμου χρησιμοποιήθηκε ο τύπος $w = S_w^{-1} * (m_1 - m_2)$. Το αποτέλεσμα που επιστρέφεται είναι το διάνυσμα πάνω στο οποίο θα πραγματοποιηθεί η προβολή των δεδομένων. Το αποτέλεσμα φαίνεται στο πρακάτω figure.

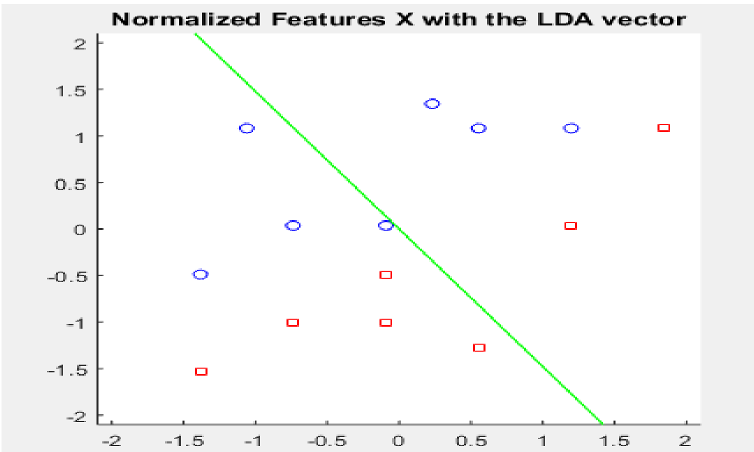


Figure 8: Normalized Features with the LDA vector

4. Ακόμη, πραγματοποιήθηκε το projection ώστε να δημιουργηθούν τα δείγματα μειωμένης διάστασης.

5. Κατόπιν, πραγματοποιήθηκε ανάκτηση των δειγμάτων στον αρχικό χώρο διαστάσεων προβάλλοντας τα πάνω στην κατεύθυνση του διανύσματος προβολής LDA. Παρατηρείται η προβολή των αρχικών δεδομένων πάνω στην κατεύθυνση του διανύσματος (μπλε και κόκκινες έντονες κουκίδες). Η προβολή πραγματοποιήθηκε σύμφωνα με τον τρόπο διαχωρισμού των κλάσεων. Το αποτέλεσμα φαίνεται στο παρακάτω figure.

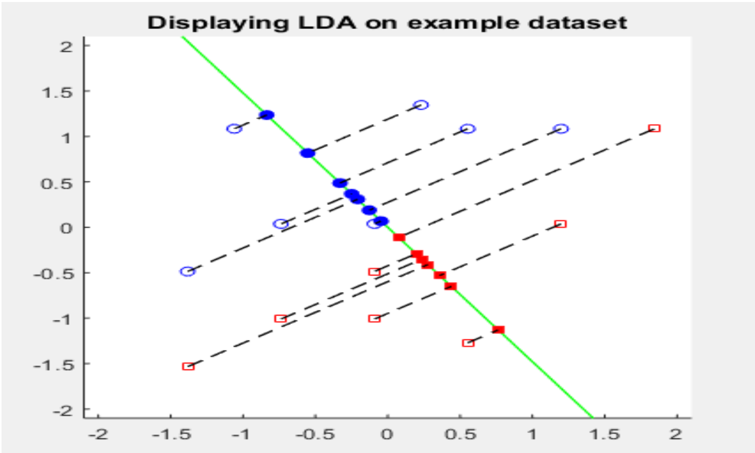


Figure 9: Displaying LDA on features

6. Τέλος, εφαρμόστηκε η ίδια διαδικασία με την μέθοδο PCA. Το αποτέλεσμα φαίνεται στο παρακάτω figure.

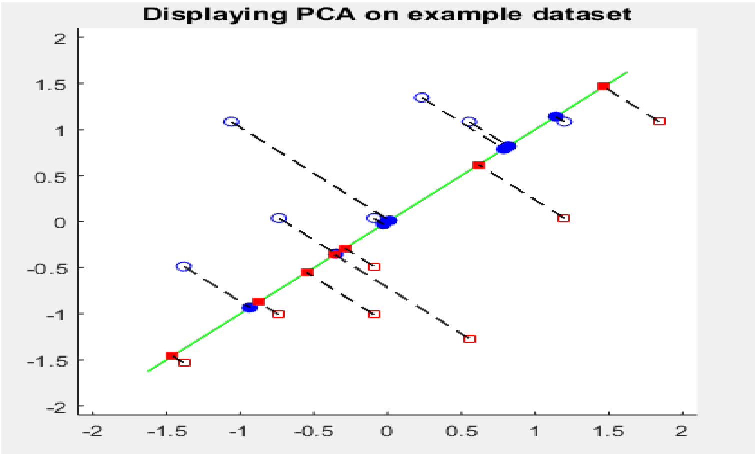


Figure 10: Displaying PCA on features

Είναι αξιοσημείωτο το γεγονός ότι τα διανύσματα προβολής των δύο αλγορίθμων είναι σχεδόν κάθετα μεταξύ τους και αυτό οφείλεται στον διαφορετικό τρόπο διαχωρισμού των κλάσεων μεταξύ του. Ο αλγόριθμος LDA βρίσκει την προβολή εκείνη η οποία θα μεγιστοποιεί την between-class απόσταση και θα ελαχιστονποιεί την within-class απόσταση. Εκεί οφείλεται και η μεγάλη απόσταση των δειγμάτων απο τη γραμμή διαχωρισμού. Σε αντίθεση με τον αλγόριθμο PCA ο οποίος μεγιστοποιεί τη διακύμανση των κλάσεων οι αποστάσεις δηλαδή των δειγμάτων απο τη γραμμή διαχωριμού είναι μικρότερες. Ο αλγόριθμος LDA λειτουργεί καλύτερα όταν υπάρχει μεγάλο αριθμός κλάσεων.

Μέρος 2

1. Σε πρώτη φάση μέσω της εντολής *load('fisheriris.mat')* έγινε φόρτωση των δεδομένων στην μεταβλητή X.

2. Στη συνέχεια πραγματοποιήθηκε standardization των δεδομένων και το αποτέλεσμα φαίνεται στο παρακάτω figure.

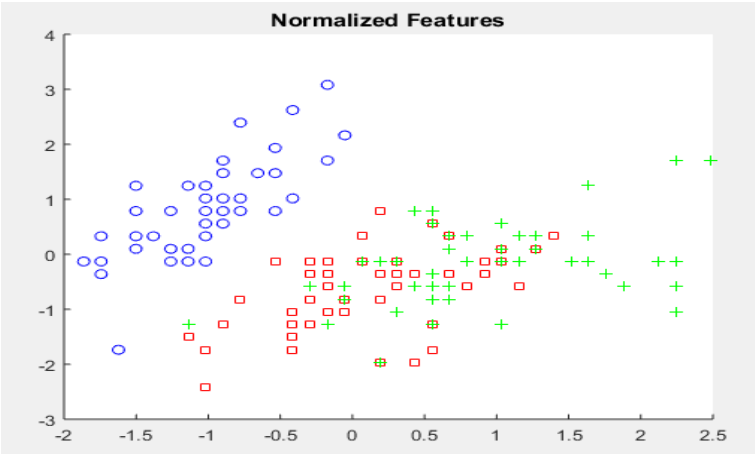


Figure 11: Normalized features

3. Έπειτα προστέθηκε ο κατάλληλος κώδικας στην συνάρτηση *myLDA* με σκοπό την εφαρμογή του αλγορίθμου για την μείωση των διαστάσεων των δειγμάτων σε 2. Αυτό που επιστρέφεται απο τη συνάρτηση είναι ο πίνακας με τα κατάλληλα διανύσματα προβολής του LDA. Για τον υπολογισμό των διανυσμάτων ήταν αναγκαίος ο υπολογισμός των παρακάτω παραμέτρων:

   (a) Τις prior πιθανότητες των κλάσεων

   (b) Τις μέσες τιμές των κλάσεων

   (c) Τον ολικό μέσο

   (d) Τον πίνακα σκέδασης S_w (Whithin-Class Scatter Matrix)

   (e) Τον πίνακα σκέδασης S_b (Between-Class Scatter Matrix)

   (f) Τον πίνακα $S_w^{-1} * S_b$

   Κατόπιν, στον πίνακα $S_w^{-1} * S_b$ εφαρμόστηκε eigendecomposition υπολογίστηκαν δηλαδή τα ιδιοδιανύσματα που αντιστοιχούν στι μεγαλύτερες 2 ιδιοτιμές με σκοπό την μείωση των διαστάσεων. Το eigendecomposition έγινε όπως αναφέρθηκε και σε αντίστοιχη άσκηση παραπάνω.

4. Τέλος, εφαρμόστηκαν τα διανύσματα προβολής που υπολογίστηκαν απο την *myLDA* στα αρχικά δείγματα μέσω της συνάρτηση *projectDataLDA* με σκοπό την μείωση των διαστάσεων

σε 2 και το αποτέλεσμα φαίνεται στο παρακάτω figure. Όπως παρατηρείται έχει γίνει διαχωρισμός των 3ων κλάσεων και ομαδοποίηση των δεδομένων όπως ήταν αναμενόμενο, καθώς όπως αναφέρθηκε ο LDA λειτουργεί καλύτερα για μεγάλο αριθμό κλάσεων.

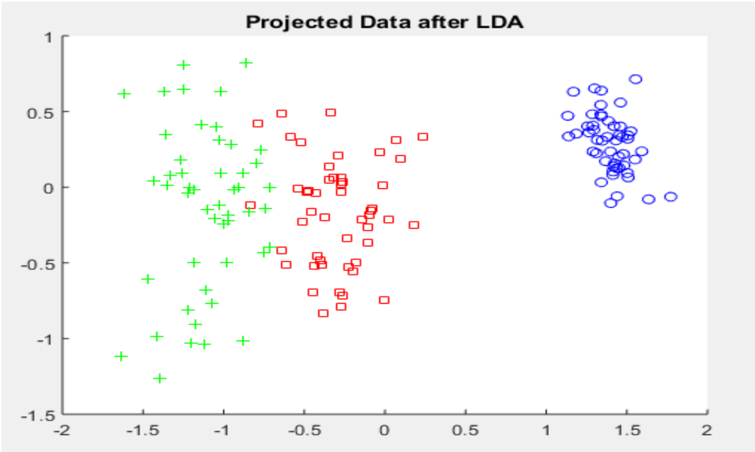


Figure 12: Features after projection on LDA vector

Θέμα 4: Εξαγωγήχαρακτηριστικλων υπολογισμός των εκ των υστέρων πιθανοτήτων με τον κανόνα του Bayes.

Στη συγκεκριμένη άσκηση ζητήθηκε η εύρεση του λόγου όψεως (aspect ratio), ο οποίος ορίζεται ως ο λόγος width/hight, όπου witdh και hight είναι το μήκος και το ύψος του ελαχίστου ορθογωνίου που περικλείει ένα ψηφίο. Τα ψηφία των οποίων υπολογίστηκε το aspect ratio ήταν τα ψηφία 1,2.

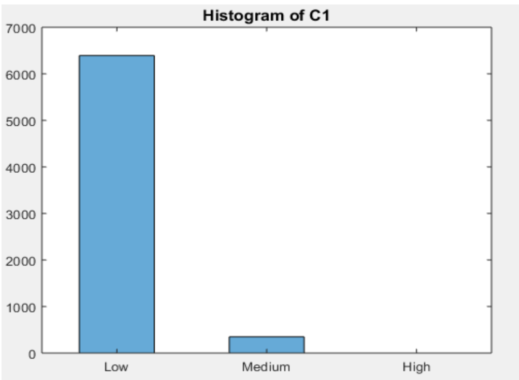
1. Αρχικά, υπολογίστηκε το aspect ratio των ψηφίων 1,2. Η διαδικασία για τον υπολογισμό του aspect ratio είναι η εξής:

   - Αρχικά υπολογίζεται το άθροισμα των γραμμών κάθε στήλης και το άθροισμα των στηλών κάθε γραμμής όπου αναφέροντα στα pixels που βρίσκονται κατα ύψος και κατά πλάτος αντίστοιχα.

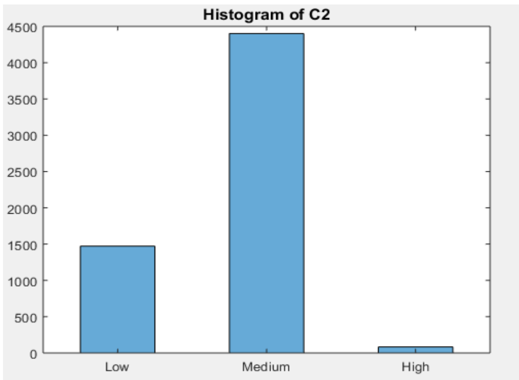
- Κατόπιν μέσω της συνάρτησης *find* του Matlab βρίσκεται η θέση του πρώτου μη μηδενικού στοιχείου απο τον πίνακα με τα αθροίσματα, για τις γραμμές και τις στήλες, και η θέση του τελέυταίου μη μηδενικού στοιχείου, για τις γραμμές και τις στήλες. Όπου για τις γραμμές υποδηλώνει τη θέση του πρώτου και του τελευταίου pixel κατά ύψος και για τις στήλες τη θέση του πρώτου και του τελευταίου pixel κατά πλάτος.
- Έπειτα, υπολογίζεται το width και το hight αφαιρόντας το max από το min και υπολογίζεται το aspect ratio.

Για τα aspect ratio των ψηφίων 1,2 βρέθηκε ότι: $minimum\_aspect\_ratio = 0.1$ και $maximum\_aspect\_ratio = 2.22$

2. Έπειτα, χωρίστηκε το διάστημα [$minimum\_aspect\_ratio = 0.1$, $maximum\_aspect\_ratio = 2.22$] σε τρία ίσα υποδιαστήματα, **Low**, **Medium**, **High**, και σχεδιάστηκαν τα ιστογράμματα των εικόνων των κλάσεων C_1 και C_2 τα οποία φαίνονται στο παρακάτω figure.



(a) Histogram of C1 class



(b) Histogram of C2 class

Όπως ηταν αναμενόμενο στην πρώτη κλάση, οι περισσότεερες εικόνες έχουν aspect ratio στο μικρότερο διάστημα αφού ο αριθμός 1 έχει μεγαλύτερο ύψος από ότι πλάτος συνεπώς το κλάσμα θα βγαίνει μικρό. Ενώ στην δεύτερη κλάση τα περίσσοτερα ψηφία βγαίνουν στο δεύτερο διάστημα, αφου ο αριθμός 2 έχει σχεδόν το ίδιο πλάτος και ύψος που σημαίνει ότι το ελάχιστον ορθογώνιο που περικλέιει το ψηφίο θα είναι τετράγωνο.

3. Στη συνέχεια υπολογίστηκαν οι πιθανότητες: $P(C_1), P(C_2), P(LgivenC_1), P(MgivenC_1), P(HgivenC_1)$

$P(LgivenC_2), P(MgivenC_2), P(HgivenC_2), P(L), P(M), P(H)$. Τα αποτελέσματα είναι τα εξής:

- $P(C_1) = C1\_images/(C1\_images + C2\_images) = 0.5309$
- $P(C_2) = C2\_images/(C1\_images + C2\_images) = 0.4691$
- $P(LgivenC_1) = (C1\_bins(1,1)/size(index\_C1,1)) * PC1 = 0.5034$
- $P(MgivenC_1) = (C1\_bins(2,1)/size(index\_C1,1)) * PC1 = 0.0275$
- $P(HgivenC_1) = (C1\_bins(2,1)/size(index\_C1,1)) * PC1 = 0$
- $P(LgivenC_2) = (C2\_bins(1,1)/size(index\_C2,1)) * PC2 = 0.1160$
- $P(MgivenC_2) = (C2\_bins(2,1)/size(index\_C2,1)) * PC2 = 0.3466$
- $P(HgivenC_2) = (C2\_bins(3,1)/size(index\_C2,1)) * PC2 = 0.0065$
- $P(L) = P(LgivenC_1) + P(LgivenC_2) = 0.6194$
- $P(M) = P(MgivenC_1) + P(MgivenC_2) = 0.3741$
- $P(H) = P(HgivenC_2) + P(HgivenC_2) = 0.0065$

4. Τέλος υπολογίστηκαν οι εκ των υστέρων πιθανότητες $P(C_1givenL)$ και $P(C_2givenL)$. Από το κανόνα του Bayes γνωρίζουμε ότι $P(YgivenX) = P(XgivenY) * P(Y)/P(X)$. Συνεπώς, $P(C_1givenL) = P(LgivenC_1) * P(C_1)/P(L) = 0.4315$ και $P(C_2givenL) = P(LgivenC_2) * P(C_2)/P(L) = 0.0879$

Θέμα 5: Bayes

1. Αρχικά, η έκφραση για το σύνορο απόφασης(διαχωρισμού) είναι:

$$P(\omega_1 givenx) = P(\omega_2 givenx) \Leftrightarrow P(xgiven\omega_1) * P(\omega_1) = P(xgiven\omega_2) * P(\omega_2)(1)$$

$$P(xgiven\omega_1) = 1/(2 * \pi * |\Sigma_1|^{1/2}) * exp(-1/2 * (\overline{x} - \overline{\mu}_1)^T * \Sigma_1^{-1} * (\overline{x} - \overline{\mu}_1))$$

$$det(\Sigma_1) = \begin{vmatrix} 1.2 & -0.4 \\ -0.4 & 1.2 \end{vmatrix} = 1.2^2 - 0.4^2 = 1.28$$

$$\Sigma_1^{-1} = 1/1.28 * \begin{bmatrix} 1.2 & 0.4 \\ 0.4 & 1.2 \end{bmatrix} = \begin{bmatrix} 0.94 & 0.31 \\ 0.31 & 0.94 \end{bmatrix}$$

$$P(xgiven\omega_1) = 1/(2*\pi*1.28^{1/2})*exp(-1/2*\begin{bmatrix} x_1-3 & x_2-3 \end{bmatrix} * \begin{bmatrix} 0.94 & 0.31 \\ 0.31 & 0.94 \end{bmatrix} * \begin{bmatrix} x_1-3 \\ x_2-3 \end{bmatrix})$$

$$= 0.14*exp(-1/2*\begin{bmatrix} 0.94*x_1-2.82+0.31*x_2-0.93 & 0.31*x_1-0.93+0.94*x_2-2.82 \end{bmatrix} * \begin{bmatrix} x_1-3 \\ x_2-3 \end{bmatrix})$$

$$= 0.14*exp(-1/2*\begin{bmatrix} 0.94*x_1+0.31*x_2-3.75 & 0.31*x_1+0.94*x_2-3.75 \end{bmatrix} * \begin{bmatrix} x_1-3 \\ x_2-3 \end{bmatrix})$$

$$= 0.14*exp((-1/2)*[(x_1-3)*(0.94*x_1+0.31*x_2-3.75)+(x_2-3)*(0.31*x_1+0.94*x_2-3.75)]$$

$$= 0.14*exp((-1/2)[0.94*x_1^2+0.31*x_1*x_2-3.75*x_1-2.82*x_1-0.93*x_2+11.25+0.31*x_1*x_2+0.94*x_2^2$$

$$-3.75*x_2-0.93*x_1-2.82*x_2+11.25)]$$

$$= 0.14*exp((-1/2)*[0.94*x_1^2+0.94*x_2^2+0.62*x_1*x_2-7.5*x_1-7.5*x_2+22.5])$$

$$P(xgiven\omega_2) = 1/(2*\pi*|\Sigma_2|^{1/2})*exp((-1/2)*(\overline{x}-\overline{\mu}_2)^T*\Sigma_2^{-1}*(\overline{x}-\overline{\mu}_1))$$

$$det(\Sigma_2) = \begin{vmatrix} 1.2 & 0.4 \\ 0.4 & 1.2 \end{vmatrix} = 1.2^2-0.4^2 = 1.28$$

$$\Sigma_2^{-1} = 1/1.28*\begin{bmatrix} 1.2 & -0.4 \\ -0.4 & 1.2 \end{bmatrix} = \begin{bmatrix} 0.94 & -0.31 \\ -0.31 & 0.94 \end{bmatrix}$$

$$P(xgiven\omega_2) = 0.14*exp((-1/2)*\begin{bmatrix} x_1-6 & x_2-6 \end{bmatrix} * \begin{bmatrix} 0.94 & -0.31 \\ -0.31 & 0.94 \end{bmatrix} * \begin{bmatrix} x_1-6 \\ x_2-6 \end{bmatrix})$$

$$= 0.14*exp((-1/2)*\begin{bmatrix} 0.94*x_1-5.64-0.31*x_2+1.86 & -0.31*x_1+1.86+0.94*x_2-5.64 \end{bmatrix} * \begin{bmatrix} x_1-6 \\ x_2-6 \end{bmatrix})$$

$$= 0.14*exp((-1/2)*\begin{bmatrix} 0.94*x_1-0.31*x_2-3.78 & -0.31*x_1+0.94*x_2-3.78 \end{bmatrix} * \begin{bmatrix} x_1-6 \\ x_2-6 \end{bmatrix})$$

$$= 0.14*exp((-1/2)[(x_1-6)*(0.94*x_1-0.31*x_2-3.78)+(x_2-6)*(-0.31*x_1+0.94*x_2-3.78)])$$

$$= 0.14*exp((-1/2)*[0.94*x_1^2-0.31*x_1*x_2-3.78*x_1-5.64*x_1+1.86x_2+22.68-0.31*x_1*x_2+0.94*x_2^2-3.78*x_2$$

$$+1.86*x_1-5.64*x_2+22.68])$$

$$= 0.14*exp((-1/2)*[0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-7.5*x_1-7.5*x_2+45.36])$$

$$(1) \Leftrightarrow 0.14*exp((-1/2)*(0.94*x_1^2+0.94*x_2^2+0.62*x_1*x_2-7.5*x_1-7.5*x_2+22.5))*P(\omega_1)$$

$$= 0.14*exp((-1/2)*(0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-7.5*x_1-7.5*x_2+45.36))*P(\omega_2) \Leftrightarrow$$

$$ln(exp((-1/2)*(0.94*x_1^2+0.94*x_2^2+0.62*x_1*x_2-7.5*x_1-7.5*x_2+22.5))*P(\omega_1)) =$$

$$ln(exp((-1/2)*(0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-7.5*x_1-7.5*x_2+45.36))*P(\omega_2))$$

$$\Leftrightarrow ln(P(\omega_1))+((-1/2)*(0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-7.5*x_1-7.5*x_2+45.36)) =$$

$$ln(P(\omega_2))+((-1/2)*(0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-7.5*x_1-7.5*x_2+45.36))$$

$$ln(P(\omega_1))-ln(P(\omega_2))-0.62*x_1*x_2-0.03*x_1-0.03*x_2+11.43=0$$

$$\Leftrightarrow x_1 = (ln(P(\omega_1))-ln(P(\omega_2))-0.03x_1-0.03x_2+11.43)/(0.62*x_2+0.03)$$

2. Κατόπιν, σχεδιάστηκαν μερικές ισοϋψει καμπύλες των δεσμευμένων πιθανοτήτων. Το αποτέλεσμα φαίνεται στο παρακάτω figure.

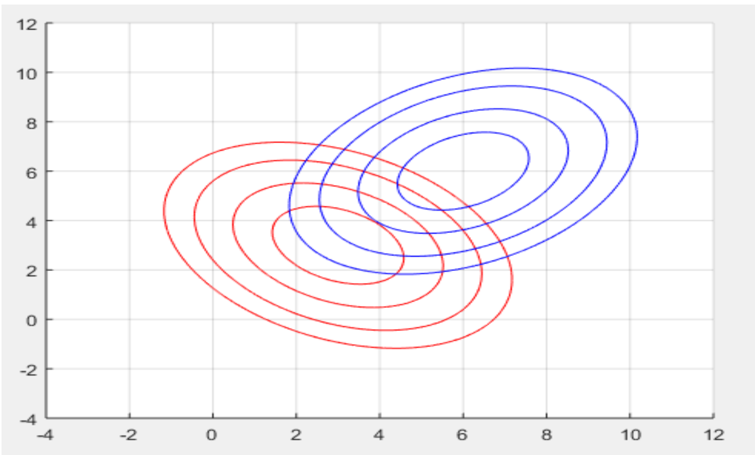


Figure 14: Equal curves of committed probabilities

3. Έπειτα υποθέτοντας ότι $P(\omega_1) = 1 - P(\omega_2)$ σχεδιάστηκαν τα σύνορα απόφασης και το αποτέλεσμα που προκύπτει είναι το εξής:

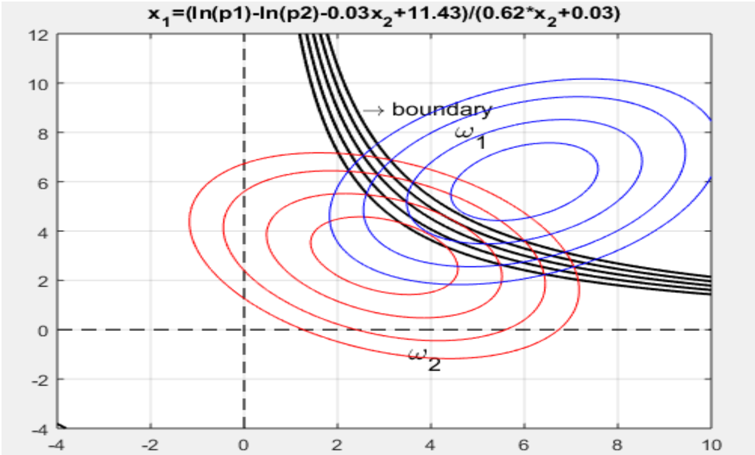


Figure 15: Dicision boundaries 1

4. Στην παραπάνω εικόνα παρατηρείται ότι η μορφή των συνόρων απόφαση είναι παραβολική καθώς η εξίσωση η οποία βρέθηκε περιέχει την μεταβλητή x_2 στον παρονομαστή. Οι διαφορετικές μορφές των εκ των προτέρων πιθανοτήτων έχουν σαν αποτέλεσμα όσο αυξάνεται η πιαθανότητα $P(\omega_1)$ τόση το όριο απόφασης να μεταφέρεται προς τα πάνω.

5. Ακολουθήθηκε η ίδια διαδικασία με $\Sigma_1 = \Sigma_2 = \begin{bmatrix} 1.2 & 0.4 \\ 0.4 & 1.2 \end{bmatrix}$ οπότε προέκυψε το παρακάτω αποτέλεσμα.

$$P(\omega_1 given x) = P(\omega_2 given x) \Leftrightarrow P(x given \omega_1) * P(\omega_1) = P(x given \omega_2) * P(\omega_2)(1)$$

$$P(x given \omega_1) = 1/(2 * \pi * |\Sigma_1|^{1/2}) * exp(-1/2 * (\overline{x} - \overline{\mu}_1)^T * \Sigma_1^{-1} * (\overline{x} - \overline{\mu}_1))$$

$$det(\Sigma_1) = \begin{vmatrix} 1.2 & 0.4 \\ 0.4 & 1.2 \end{vmatrix} = 1.2^2 - 0.4^2 = 1.28$$

$$\Sigma_1^{-1} = 1/1.28 * \begin{bmatrix} 1.2 & -0.4 \\ -0.4 & 1.2 \end{bmatrix} = \begin{bmatrix} 0.94 & -0.31 \\ -0.31 & 0.94 \end{bmatrix}$$

$$P(x given \omega_1) = 1/(2 * \pi * 1.28^{1/2}) * exp(-1/2 * \begin{bmatrix} x_1 - 3 & x_2 - 3 \end{bmatrix} * \begin{bmatrix} 0.94 & -0.31 \\ -0.31 & 0.94 \end{bmatrix} * \begin{bmatrix} x_1 - 3 \\ x_2 - 3 \end{bmatrix})$$

$$= 0.14*exp(-1/2*\begin{bmatrix}0.94*x_1-2.82-0.31*x_2+0.93 & -0.31*x_1+0.93+0.94*x_2-2.82\end{bmatrix}*\begin{bmatrix}x_1-3\\x_2-3\end{bmatrix})$$

$$= 0.14*exp(-1/2*\begin{bmatrix}0.94*x_1-0.31*x_2-1.89 & -0.31*x_1+0.94*x_2-1.89\end{bmatrix}*\begin{bmatrix}x_1-3\\x_2-3\end{bmatrix})$$

$$= 0.14*exp((-1/2)*[(x_1-3)*(0.94*x_1-0.31*x_2-1.89)+(x_2-3)*(-0.31*x_1+0.94*x_2-1.89)]$$

$$= 0.14*exp((-1/2)[0.94*x_1^2-0.31*x_1*x_2-1.89*x_1-2.82*x_1+0.93*x_2+5.67-0.31*x_1*x_2+0.94*x_2^2$$

$$-1.89*x_2+0.93*x_1-2.82*x_2+5.67)]$$

$$= 0.14*exp((-1/2)*[0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-3.78*x_1-3.78*x_2+11.34])$$

$$P(xgiven\omega_2) = 1/(2*\pi*|\Sigma_2|^{1/2})*exp((-1/2)*(\overline{x}-\overline{\mu}_2)^T*\Sigma_2^{-1}*(\overline{x}-\overline{\mu}_1))$$

$$det(\Sigma_2) = \begin{vmatrix}1.2 & 0.4\\0.4 & 1.2\end{vmatrix} = 1.2^2-0.4^2 = 1.28$$

$$\Sigma_2^{-1} = 1/1.28*\begin{bmatrix}1.2 & -0.4\\-0.4 & 1.2\end{bmatrix} = \begin{bmatrix}0.94 & -0.31\\-0.31 & 0.94\end{bmatrix}$$

$$P(xgiven\omega_2) = 0.14*exp((-1/2)*\begin{bmatrix}x_1-6 & x_2-6\end{bmatrix}*\begin{bmatrix}0.94 & -0.31\\-0.31 & 0.94\end{bmatrix}*\begin{bmatrix}x_1-6\\x_2-6\end{bmatrix})$$

$$= 0.14*exp((-1/2)*\begin{bmatrix}0.94*x_1-5.64-0.31*x_2+1.86 & -0.31*x_1+1.86+0.94*x_2-5.64\end{bmatrix}*\begin{bmatrix}x_1-6\\x_2-6\end{bmatrix})$$

$$= 0.14*exp((-1/2)*\begin{bmatrix}0.94*x_1-0.31*x_2-3.78 & -0.31*x_1+0.94*x_2-3.78\end{bmatrix}*\begin{bmatrix}x_1-6\\x_2-6\end{bmatrix})$$

$$= 0.14*exp((-1/2)[(x_1-6)*(0.94*x_1-0.31*x_2-3.78)+(x_2-6)*(-0.31*x_1+0.94*x_2-3.78)])$$

$$= 0.14*exp((-1/2)*[0.94*x_1^2-0.31*x_1*x_2-3.78*x_1-5.64*x_1+1.86x_2+22.68-0.31*x_1*x_2+0.94*x_2^2-3.78*x_2$$

$$+1.86*x_1-5.64*x_2+22.68])$$

$$= 0.14*exp((-1/2)*[0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-7.5*x_1-7.5*x_2+45.36])$$

$$(1) \Leftrightarrow 0.14*exp((-1/2)*[0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-3.78*x_1-3.78*x_2+11.34])*P(\omega_1)$$

$$= 0.14*exp((-1/2)*(0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-7.5*x_1-7.5*x_2+45.36))*P(\omega_2) \Leftrightarrow$$

$$ln(exp((-1/2)*(0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-3.78*x_1-3.78*x_2+11.34))*P(\omega_1)) =$$

$$ln(exp((-1/2)*(0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-7.5*x_1-7.5*x_2+45.36))*P(\omega_2))$$

$$\Leftrightarrow ln(P(\omega_1))+((-1/2)*(0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-3.78*x_1-3.78*x_2+11.34) =$$

$$ln(P(\omega_2))+((-1/2)*(0.94*x_1^2+0.94*x_2^2-0.62*x_1*x_2-7.5*x_1-7.5*x_2+45.36))$$

$$ln(P(\omega_1))-ln(P(\omega_2))-1.89*x_1-1.89*x_2+17.01=0$$

$$x_1 = (ln(P(\omega_1))-ln(P(\omega_2))-1.89+17.01)/1.89$$

6. Σύμφωνα με την παρπάνω εξίσωση και υποθέτοντας ότι $P(\omega_1) = 1 - P(\omega_2)$ σχεδιάστηκαν τα σύνορα απόφασης και το αποτέλεσμα που προκύπτει είναι το εξής:

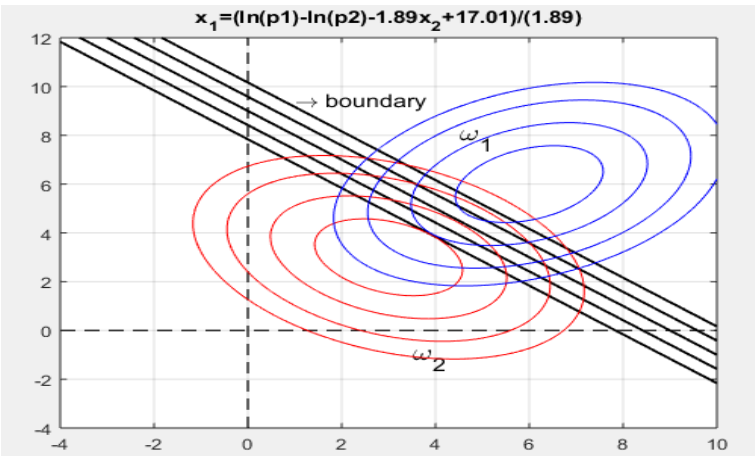


Figure 16: Dicision boundaries 2

7. Στην παραπάνω εικόνα παρατηρείται ότι η μορφή των συνόρων απόφαση είναι γραμμική καθώς η εξίσωση η οποία βρέθηκε είναι γραμμική εξίσωση πρώτου βαθμού. Οι διαφορετικές μορφές των εκ των προτέρων πιθανοτήτων έχουν σαν αποτέλεσμα όσο αυξάνεται η πιαθανότητα $P(\omega_1)$ τόσο το όριο απόφασης να μεταφέρεται προς τα πάνω.

Θέμα 6: Minimum Risk

Το κρητήριο μέσου ρίσκου είναι:

$$l_1 = \lambda_{11} * P(\omega_1) * P(xgiven\omega_1) + \lambda_{21} * P(\omega_2) * P(xgiven\omega_2)$$

$$l_2 = \lambda_{12} * P(\omega_1) * P(xgiven\omega_1) + \lambda_{22} * P(\omega_2) * P(xgiven\omega_2)$$

Για να ανήκει το x στην ω_1 πρέπει:

$$l_1 < l_2 \Leftrightarrow \lambda_{11}*P(\omega_1)*P(xgiven\omega_1)+\lambda_{21}*P(\omega_2)*P(xgiven\omega_2) < \lambda_{12}*P(\omega_1)*P(xgiven\omega_1)+\lambda_{22}*P(\omega_2)*P(xgiven\omega_2)$$

$$\Leftrightarrow 0 + P(xgiven\omega_1) < 0.5 * P(xgiven\omega_2) + 0 \Leftrightarrow P(xgiven\omega_2)/P(xgiven\omega_1) < 0.5$$

Για να ανήκει το x στην ω_2 πρέπει:

$$l_1 > l_2 \Leftrightarrow \lambda_{11}*P(\omega_1)*P(xgiven\omega_1)+\lambda_{21}*P(\omega_2)*P(xgiven\omega_2) > \lambda_{12}*P(\omega_1)*P(xgiven\omega_1)+\lambda_{22}*P(\omega_2)*P(xgiven\omega_2)$$

$$\Leftrightarrow 0 + P(xgiven\omega_1) > 0.5 * P(xgiven\omega_2) + 0 \Leftrightarrow P(xgiven\omega_2)/P(xgiven\omega_1) > 0.5$$

Το υπερεπίπεδο απόφασης είναι:

$$l_1 = l_2 \Leftrightarrow \lambda_{11}*P(\omega_1)*P(xgiven\omega_1)+\lambda_{21}*P(\omega_2)*P(xgiven\omega_2) = \lambda_{12}*P(\omega_1)*P(xgiven\omega_1)+\lambda_{22}*P(\omega_2)*P(xgiven\omega_2)$$

$$\Leftrightarrow 0 + P(xgiven\omega_1) = 0.5 * P(xgiven\omega_2) + 0 \Leftrightarrow P(xgiven\omega_2) = P(xgiven\omega_1) * 0.5$$

$$\Leftrightarrow x/_1^2 * e^{-x^2/2*_1^2} = 1/2*x/_2^2*e^{-x^2/2*_2^2} \Leftrightarrow x/4*e^{-x^2/8} = x/2*e^{-x^2/2} \Leftrightarrow ln(x/4*e^{-x^2/8}) = ln(x/2*e^{-x^2/2})$$

$$\Leftrightarrow ln(x/4 * e^{-x^2/8}) = ln(x/2 * e^{-x^2/2}) \Leftrightarrow ln(x/4) + ln(e^{-x^2/8}) = ln(x/2) + ln(e^{-x^2/2})$$

$$\Leftrightarrow -x^2/8 + x^2/2 = ln(x/2/x/4) \Leftrightarrow (-x^2 + 4 * x^2)/8 = ln(2) \Leftrightarrow 3 * x^2/8 = ln(2) \Leftrightarrow x = + - 1.35$$

Αφού όμως γωνρίζουμε απο την εκφώνηση $x >= 0$ αποδεκτή $x = 1.35$

Παραπομπές

- Διαφάνειες μαθήματος και Φροντιστηρίου
- https://www.mathworks.com/matlabcentral/answers/513482-how-to-measure-the-aspect-ratio-of-a-certain-part-of-an-image